A critical review of recent trends in sample classification using Laser-Induced Breakdown Spectroscopy (LIBS)

L. Brunnbauer[*], Z. Gajarska, H. Lohninger, A. Limbeck[**]

TU Wien, Institute of Chemical Technologies and Analytics, Getreidemarkt 9/164-I²AC, 1060, Vienna, Austria

ARTICLE INFO

ABSTRACT

LIBS-based classification has experienced an ever-increasing interest in the last few years. LIBS is a well-suited technique for classification tasks based on elemental fingerprinting, providing fast simultaneous multi-element analysis with stand-off, online, and portable capabilities. The topic of classification gained even more momentum due to the current hype on machine learning, big data, and chemometrics. Nevertheless, with many LIBS users not being data scientists by training, classification algorithms are often used and considered "black boxes," hindering the adequate application of these tools. This review provides a comprehensive introduction and overview of the steps necessary (e.g., normalization, background correction, feature selection) to go from recorded data to a well-performing classifier. Additionally, the basic principles, advantages, and limitations of the most used machine learning algorithms reported in LIBS-classification literature are discussed. Finally, the review offers an overview of the literature published in the field, highlighting the great diversity of applications.

1. Introduction

In various fields of modern science but also for routine applications in industry, food inspection, and health care, there is an increasing need to discriminate samples with differences in behavior or origin. The main goal of the efforts is to ensure the quality of produced goods (e.g., to guarantee specifications) or to protect them against fraud or imitation. For example, the determination of provenance, adulteration, and mislabeling is an important problem in many areas of the food industry, affecting the credibility of producers and traders and the rights of consumers [1]. Reliable authentication of samples is not limited to commercial food products such as wine, liquors, or ham and cheese, the determination of the authenticity is also imperative in archaeology, where counterfeit ancient artifacts must be differentiated from extremely similar authentic objects [2]. In the field of forensic analysis, the samples collected at a crime scene, including but not limited to broken glass, automotive paint chips, gunshot residues, drugs, or blood were often compared with reference samples to solve crimes, uncover mysteries, and convict criminals [3]. Such discrimination and, if possible, identification of samples is also required in many industrial procedures. For example, for effective recycling of valuable metals, a comprehensive screening of e-waste is mandatory [4]. Nowadays, the reuse of synthetic materials (e.g., plastic bottles) has become an important issue. However, the quality of recycled plastic is poor when no reliable sorting method is used for separating dumped plastics [5].

In recent years much progress has been made in sample classification based on elemental fingerprinting techniques [6–8]. For this purpose, elemental analysis using optical or mass spectrometric techniques are combined with multivariate statistical analysis of the obtained data to gain information about differences or variations within the investigated samples [9,10]. Determination of the prevailing elemental contents is usually accomplished using atomic absorption spectrometry (AAS), inductively coupled plasma optical emission spectrometry (ICP-OES), and inductively coupled plasma mass spectrometry (ICP-MS). The main benefit of the ICP-based approaches is that more than one element can be analyzed simultaneously, enabling fast and sensitive measurement of multi-elemental fingerprints. In the case of ICP-MS, additional information about the isotopic composition of the sample is accessible, providing further evidence about the geographical origin of the sample [11]. However, the conventional application of these techniques requires converting solid samples into liquid solutions. For this purpose, various kinds of digestion, combustion, or fusion

* Corresponding author.
** Corresponding author.
E-mail addresses: Lukas.brunnbauer@tuwien.ac.at (L. Brunnbauer), Andreas.Limbeck@tuwien.ac.at (A. Limbeck).

procedures were reported [12]. Although these approaches are well established, the need for a sample dissolution step prior to analysis is related to some drawbacks. In particular, the time demand for sample preparation and the risk of sample contamination or analyte losses must be mentioned [13,14]. These undesired sample modifications could bias the quality of sample classification, provenance determination, or food authenticity.

Direct analysis of solid samples allows to overcome the problems associated with sample dissolution. Besides, applying solid-sampling techniques often permits improvements in sensitivity since unnecessary sample dilution is avoided. Moreover, solid sampling techniques such as X-ray fluorescence analysis (XRF) [15], electron micro probe analysis (EMPA) [15], laser ablation-inductively coupled plasma-mass spectrometry (LA-ICP-MS) [16,17] or laser-induced breakdown spectroscopy (LIBS) [18,19] offer the possibility of spatially resolved analysis, providing information about the distribution of major, minor, and trace constituents within the sample. Compared to LIBS, LA-ICP-MS excels with its high sensitivity and capabilities of gathering isotopic information whereas EMPA can provide sub μm lateral resolution. Nevertheless, LA-ICP-MS usually allows only targeted analysis whereas EMPA only provides limited sensitivity. Even though, LA-ICP-MS [20–22], EMPA [23–25], and XRF [26–29] are also used for classification based on elemental fingerprinting, LIBS has gained increasing popularity in this field due to its many advantages which are outlined in the next paragraphs.

LIBS analysis is based on applying a high-power laser pulse on the sample surface leading to a certain amount of mass being ablated and a localized plasma forming on the ablation spot. In this plasma, the ablated material is partly atomized, and excited atoms and ions are generated. When these excited states decay back to their ground levels, the energy of the corresponding transitions is emitted in the form of electromagnetic radiation. Collection and detection of the emitted light provides information about the elemental composition of the investigated sample spot. In contrast to most other techniques used for elemental analysis LIBS offers access to the whole periodic table of elements, thus the main constituents of all biological and geological materials, the nonmetals H, C, N and O can be measured with this technique. Another major benefit of LIBS is that this analytic technique provides not only information about the elemental sample composition, to some extent also molecular information is accessible [30]. Incomplete atomization of organic sample constituents but also the recombination of atoms within the laser-induced plasma results in the formation of excited molecular species – in particular the C_2 swan band and the CN violet band. The use of these molecular emission signals is widely reported for the classification and discrimination of polymers [31,32].

LIBS allows a non-targeted analysis since no prior knowledge about sample composition and no preliminary definition of emission lines detected is necessary. However, the simultaneous measurement of all elements of the periodic table is only possible with the use of broadband spectrometers covering the wavelength range from approximately 200 to 900 nm. Besides recording broadband spectra, only selected spectral regions are monitored in some cases. With this approach, an improved resolution can be achieved, but the number of simultaneously accessible elements is usually limited. Thus, prior to sample analysis a careful optimization of the measurement conditions is crucial, which is a prerequisite for sensitive but also selective analysis. In this context instrumental parameters such as laser energy, gate delay, gate width and the applied atmosphere (e.g., ambient air, Ar or He) are important. Using optimized conditions detection limits ranging from the μg/g level for alkaline and earth alkaline metals to the per mill range for non-metals can be achieved. Further improvements in the signal to noise ratio as well as in the reproducibility of analysis are possible by accumulating the LIBS spectra from multiple laser shots. Moreover, frequently various data preprocessing steps are performed, such as background correction or data normalization, to further increase the quality of the obtained results.

An attribute that makes LIBS especially attractive for provenance and authentication studies and the detection of product adulterations is its minimal invasive character, which is in contrast to non-destructive techniques such as XRF or EMPA. Thus, separating surface layers or coatings is possible, providing access to the underlying bulk material. This characteristic feature of LIBS is vital for the analysis of samples with a layered structure such as technological materials containing special surface coatings necessary to achieve the aspired material properties (e.g., strength, hardness, stiffness, …) but also all kinds of samples which undergo surface reactions in the environment leading to the formation of passivation or corrosion layers.

One of the major challenges in the field of LIBS is quantitative analysis of the elemental composition. To achieve reliable quantification, typically an external calibration with matrix-matched standards is required [33], or other approaches such as calibration-free LIBS (CF-LIBS) [34] are used. Since quantification is typically not performed in classification studies, this topic is not further discussed within this review.

With the above-mentioned advantages, LIBS has become a great candidate for classification applications in a wide range of fields. Moreover, the current boom in the machine learning field and a growing number of advanced chemometric tools available for processing of LIBS data have boosted the application potential even more. Nevertheless, as outlined by Hahn and Omenetto [35], if applied incorrectly, these tools might provide statistically significant outcomes (e.g., high classification accuracy) even in cases with no chemical/elemental difference present in the samples. Classification of LIBS data is a complex task requiring expertise in analytical chemistry, plasma physics, and chemometrics/machine learning. Neglecting one of these aspects may easily result in overfitted models and improper classification results in general which is, unfortunately, often found in the literature. The following paragraph aims to summarize the most common malpractices.

Even though recent instrumental developments (lasers with a higher repetition rate, spectrometers with a higher resolution) facilitate collection of more and more data, the amount of data used for classification studies is still limited which may easily lead to over-fitting. Fig. 1 shows both the total number of raw LIBS spectra recorded and the number of sample classes reported in the revised
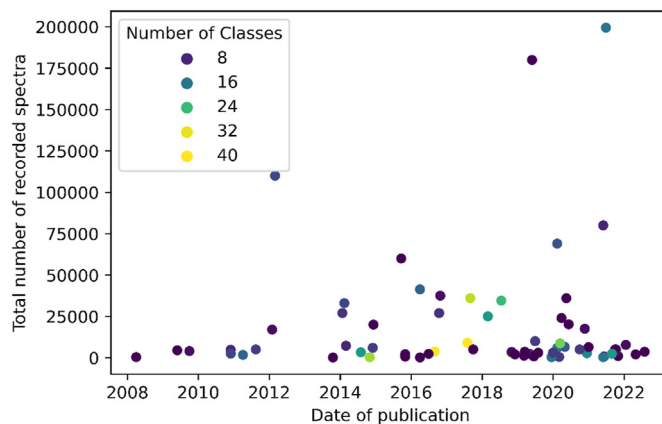

Fig. 1. Total number of recorded LIBS spectra and the number of investigated sample classes reported in the literature in the last decade.

literature over the last decade. Although, the amount of data used for classification studies has steadily increased over the last years, in most works, today's modern instrumentation is still not used at its full potential.

The other aspects that are applied suboptimal in the literature are mostly based on data handling and performance evaluation. Here, especially the use of the whole LIBS spectrum without proper data preprocessing and feature selection should be mentioned. This approach may lead to over-fitted models (e.g., fitted to the noise of the data) and the *"curse of dimensionality"* [36]. Additionally, hyperparameters are often tuned inadequately resulting in non-optimal performing models. Lastly, in many cases, the performance of the established classification model is evaluated by cross-validation only resulting in overestimated performance metrics. Whereas, using independent data (e.g., recorded on a different day) provides a more realistic performance and reflects the generalizability of the model.

In this review we provide an overview of good practices in the field of LIBS-based classification ranging from data pre-processing to the training of a reliable classifier applicable in the real-life scenario. Therefore, the review provides an intuitive guide to the supervised classification methods most frequently applied in the field. Finally, we present an extensive overview of publications within the last 10 years and application fields in which LIBS is expected to play a major role in the future.

2. Introduction to chemometrics/classification

From a chemometric perspective, LIBS spectra (Fig. 2 a)) can be seen as points in a multi-dimensional space. Whereas the analyzed wavelengths define the coordinate system, the corresponding spectral intensities specify the position within the space (Fig. 2 b)).

Assuming that substances of the same class deliver similar spectral fingerprints and that distances in the space are a reliable measure of the spectral similarity, spectra belonging to the same class are expected to cluster in a particular region of the LIBS space representative of the class (Fig. 2 b)). Allocating a spectrum of an unknown substance in such a space, one gains information about its identity.

Given a classification problem, it is common to first explore the relationships in the data, e.g., by employing exploratory tools such as PCA to find out whether there are chemical differences in the classes reflected in the LIBS spectra and, if so, whether they are sufficiently distinct to provide a reliable class separation. Once the discriminability of the classes by means of LIBS has been proven,

one can proceed with the development of a classification model (Fig. 3), enabling the mapping of the spectra to their classes.

As the development of a reliable classifier comprises multiple steps that profoundly influence its quality, each step shall be performed with care and ideally with the ultimate goal of the analysis − a reliable prediction of new samples − in mind. Even before the LIBS analysis, one shall ensure that the selected type of samples and
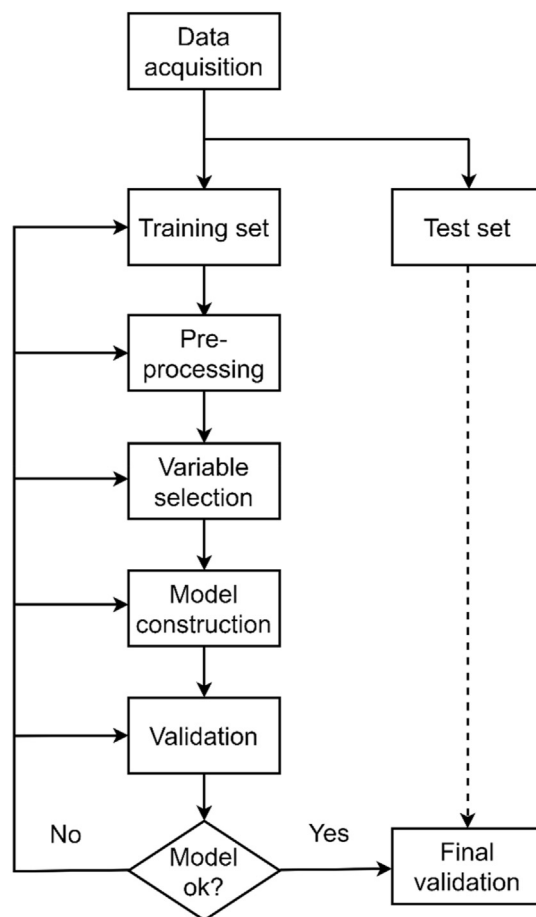

Fig. 3. Development of a classifier. The dashed line between the test set and final validation indicates that the data processing strategy "learned" during the training phase is to be applied to the test data before the final evaluation can begin.
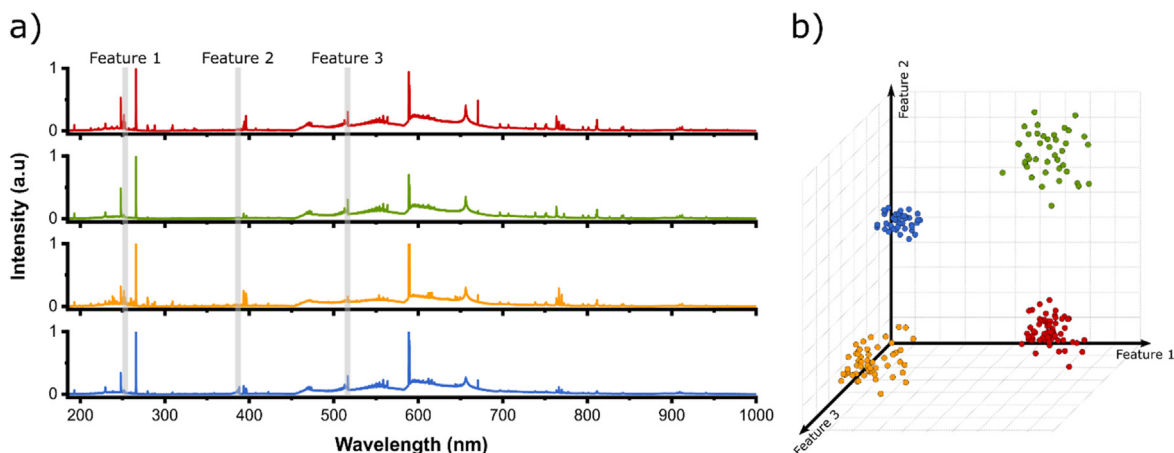

Fig. 2. a) Representative LIBS spectra and b) Intuitive representation of the chemometric (LIBS) space.

experimental conditions are representative of the future identification scenario. After the data acquisition, it is common to divide the available dataset into a training (approx. 80%) and a test set (approx. 20%). Whereas the training set is employed in the further process of classifier development and optimization, the test set shall be left "untouched" until the final evaluation of the classifier's performance.

In the subsequent step, it is good to check for anomalies or inconsistencies in the training data by a visual inspection of the raw spectra and/or use of various visualization and outlier detection techniques such as PCA [37]. Depending on their nature and severity, the detected outliers can be removed. Nevertheless, too harsh outlier filtering is not recommended as it can distort the natural data structure crucial for the training. The data selection is followed by a series of pre-processing steps aiming to reduce irrelevant spectral variations, which might complicate the training. Before letting the classifier "learn", it is common to reduce the dimensionality of the "learning" space by a down-selection or engineering of spectral variables representative of the involved classes. Throughout these steps, it is often helpful to track changes in the data structure by means of an unsupervised technique such as PCA. Once the data become representative of the true class distributions, one can employ a suitable classification algorithm to "learn" the function which maps the spectra to the corresponding classes.

In this process, the classifier's characteristics (hyperparameters) often need to be "tuned" for the particular scenario, which is typically done by either using a validation set (a subset of training data set aside for the purpose of model optimization) or a process of k-fold cross-validation [38,39]. The latter is based on splitting the training data into k subsets. The k-1 subsets are then employed for the training, and the left-out (k-th) subset is used for validation. This process is repeated k times until all subsets have been employed for the validation once. As the prediction results of the individual subsets are averaged to provide an estimate of the classifier's performance, k-fold cross-validation is typically preferred and perceived as a more reliable estimate of the model's performance. Additionally, by reporting the corresponding standard deviation, it is possible to gain insight into the classifier's stability (sensitivity to outliers/overfitting). As Fig. 3 demonstrates, the process of classifier development is iterative - if the classifier's performance does not meet the expectations, it is possible (and recommended) to return to any of the preceding steps and alter the employed strategy. The different settings can be compared by means of a k-fold cross-validation, which allows for the selection of the best-performing model.

Having the final classifier, its ability to correctly identify new samples shall be tested on an independent set of samples representative of the future application scenario. Nevertheless, such an additional set of data is often difficult to obtain (e.g., high cost). One therefore, accommodates the scenario of an independent validation set by splitting the data into training (approx. 80%) and test set (approx. 20%) as described above. Only at this point does the test set become revealed, subjected to the pre-processing strategy developed during the training, and used for the final evaluation of the model's prediction ability. As previously outlined, given the high sensitivity of the spectral fingerprint to the instrument characteristics and experimental conditions, conclusions about the generalization ability of the LIBS classifiers trained on spectra coming from the same measurement as the validation set could be disputable and might require additional training data obtained under different conditions.

To shed more light on the individual stages of analysis, the following sections provide a detailed discussion of each stage together with the methods employed for their handling, references to well-known reviews on the topic, as well as specific application examples from the LIBS literature.

2.1. Data preprocessing

After the data selection, one typically involves a series of pre-processing steps aiming at the reduction or complete removal of spectral variations/artifacts arising from phenomena other than the chemical nature of the classes (e.g., fluctuations of the laser energy or sample homogeneity). Whereas a suitable pre-processing strategy can greatly enhance the effectiveness of the training and stability of the corresponding classifier, improper treatment can distort the class distributions and negatively impact the classifier's performance. As outlined before, the suitability of the different strategies for the given scenario can be investigated by means of the k-fold cross-validation. In the field of LIBS, three steps discussed below are commonly employed for data preprocessing.

- Background correction (e.g., to compensate for continuum background in LIBS spectra)
- Data normalization (e.g., to address varying measurement conditions)
- Removal of noise

2.1.1. Background correction
Background correction is not only of interest in the field of LIBS but also in all other spectroscopic techniques. Schulze et al. provide an excellent review of approaches for background correction in various spectroscopic techniques [40]. LIBS spectra often exhibit noticeable background signals originating from various sources within the plasma. For reliable classification, only the contribution of the analytes to the observed signal intensities should be considered. The two major sources of continuous background signals are caused by Bremsstrahlung and radiation caused by recombination of species present in the plasma [41−43]. Although continuous background signals can be drastically reduced by careful optimization of the time and duration used for detection (gate delay and gate width of a gated detection system), background correction of LIBS spectra is still an important step in data preprocessing [44]. Gornushkin et al. proposed a method using polynomial fitting through a number of intensity minima in the spectrum for background correction [45]. This approach was improved and updated by Sun and Yun [46]. Yaroshchyk and Eberhardt introduced a modification of Friedrichs' method [47] for the correction of LIBS background [48]. Other approaches for background correction reported in LIBS literature include a spline interpolation [49] or Lorentz fitting [50]. In the work of Képeš et al. [51], different background correction approaches to LIBS datasets are investigated and comprehensively discussed.

2.1.2. Data normalization
One of the main goals of spectral normalization in the field of LIBS is the reduction of signal fluctuations observed during measurement. The main causes for unwanted signal fluctuations are often found in changing measurement conditions such as laser energy, efficiency of light collection, or defocusing of the laser. Additionally, insufficient sample homogeneity or changes in the morphological properties of the sample (e.g., roughness) may also lead to unwanted signal fluctuations [52−54]. In the field of LIBS, there are several established approaches for signal normalization summarized in a review article by Guezenoc et al. [55]. Pořízka et al. investigated the influence of different data normalization strategies on classification performance using LIBS [56]. Zorov et al. provide a review article on different normalization strategies for LIBS [57].

The most commonly used approaches in LIBS literature are normalization to the total emission intensity, normalization to an internal standard, and normalization to the standard normal variate (SNV), which will be discussed in more detail below. Other normalization approaches include normalization to the Euclidean norm, normalization to the signal maximum or minimum, normalization to the background [58], normalization to the ablated mass [59], normalization to the acoustic signal measured after the laser pulse [57], and normalization to plasma images [60,61].

Normalization to the total emission intensity: The idea behind normalization to the total emission intensity (often also referred to as normalization to the total area) is the correlation between the total signal recorded in a LIBS spectrum and the laser energy [62]. Therefore, fluctuations in laser energy can be compensated by normalization to the total emission intensity.

Normalization to an internal standard: Normalization to an internal standard is an effective approach to compensate for instrumental drifts and measurement fluctuations. In this case, the signal from an element that is homogeneously distributed within the analyzed sample is used for normalization. Observed signals originating from this element should be constant during the measurement. If measurement fluctuations occur, the internal standard is equally affected as the analytes of interest and, therefore can be used for compensation. For proper application of an internal standard, some requirements have to be fulfilled, which is difficult in many cases [63].

- The internal standard must be homogeneously distributed within the sample
- The internal standard should have similar physical and chemical properties as the analyte of interest to be equally affected by measurement fluctuations

If no suitable internal standard is available within the investigated sample, some works have proposed the application of thin films (Au or spiked polymers) onto the surface of the compact sample, which can be used as an internal standard [53,64,65].

Normalization to the standard normal variate: This approach for data normalization is well-established in the field of IR- and Raman-Spectroscopy [66,67] and has recently also found its way to the field of LIBS [68,69]. When performing standard normal variate normalization, each spectrum is shifted to a mean value $\mu = 0$ and a standard deviation $\sigma = 1$. Using this approach, a significant reduction of shot-to-shot variations in LIBS signals was reported in the literature.

2.1.3. Removal of noise

A LIBS spectrum is always a superposition of useful spectral information and noise originating from the measurement. Tognoni et al. provide a comprehensive review of signal and noise in LIBS analysis [70]. Schlenke et al. describe three fundamental sources of noise present in LIBS spectra: photon noise, detector noise, and flicker noise [71], whereas Mermet et al. describe four different sources of noise in LIBS analysis: noise caused by plasma fluctuations, shot noise due to the random arrival of the photons on the detector, detector noise, and drifts [72]. To improve the performance of classification using LIBS data, one approach is to reduce noise in the data set as much as possible. There are several methods used for this task reported in the literature, most of which are based on a wavelet decomposition transformation (WDT) [73,74] of the LIBS spectra [71,75,76].

2.2. Extraction of spectral features

Since LIBS is a non-targeted analysis, observed spectra often have a great number of emission signals present, some of which can contribute to a classification task and others might not contain useful information. Broadband LIBS spectra covering wavelength ranges of >500 nm often consist of up to 10,000 data points per spectrum. Therefore, each measured LIBS spectrum can be represented as a point in a 10,000-dimensional space. Without proper reduction of spectral data, one could easily run into the so-called *"curse of dimensionality"* when establishing multivariate classification models. The *"curse of dimensionality"* is a term introduced by Bellman in 1961 [36] and describes various problems when analyzing data in high-dimensional spaces if the number of measured data points is much lower than the dimensionality of the space itself. As a rule of thumb, in the field of machine learning, at least five training samples are necessary for each dimension of the data space [77]. Therefore, without proper data reduction and extraction of significant spectral features, one would need to measure more than 50,000 LIBS spectra to circumvent the *"curse of dimensionality"* for proper classification results. When extracting, e.g., 20 significant variables from the LIBS spectra, the dimensionality of the data space is reduced, and the amount of necessary training samples is reduced to a reasonable number [78].

By now, a great variety of feature selection approaches has been reported in the field. The simplest is based on the extraction of all emission lines present in the spectrum. Although fulfilling the goal of dimensionality reduction, this approach might still retain spectral information irrelevant to the sample's identity, such as emission signals of impurities or measurement contributions (e.g., emission lines originating from the atmosphere − Ar, He, or air). Including such variables in the model might not only prolong the computational times but also result in "memorization of noise" (overfitting), leading to a poor classification of new samples (generalization ability). Thus, many works in the field use an alternative approach of manually selecting spectral intensities or spectral regions subjected to the analysis.

The concept of manual selection can be further extended to "feature engineering," which involves the concentration of the relevant spectral information in a set of derived features (so-called spectral descriptors) representing unique spectral traits of the investigated classes. In the field of LIBS, these are often spectral intensities or integrated spectral regions automatically accounting for the baseline. Nevertheless, definition of more complex nonlinear variables derived from multiple spectral regions is possible. As this approach requires chemical knowledge about the problem not always available to the analyst, it can be accompanied by information from another method such as PCA. For more information on spectral descriptors, we refer the readers to the work of Lohninger and Ofner [79].

Although the training of a classifier in a highly specific "chemical space" might greatly enhance the classifier's robustness and reduce the computational times, the manual definition of suitable descriptors requires expert knowledge and time. One can therefore generate a larger number of variables and employ a variable selection strategy to fully automatize the feature selection process. In general, three types of methods are recognized. The so-called filter methods select features prior to the modeling based on some kind of metrics evaluating their suitability, such as information gain, correlation, or Chi-square [80]. A recent example of a filter approach in the field of LIBS was presented in the work of Huffman et al. [81]. In contrast to filter methods, the so-called wrapper approach [82] bases the evaluation of different feature combinations on the performance of a classification algorithm. The generation of variable subsets in filter and wrapper methods can be governed by different algorithms such as genetic algorithm, forward or backward selection, or successive projection algorithm (SPA) [83]. This was demonstrated in the work of Pontes et al.

employing different modes of variable selection for the classification of Brazilian soils [84]. The last feature selection approach - the embedded method - selects the features automatically during the algorithm's execution, either as its normal (e.g., PLS-DA) or extended functionality (e.g., by including a penalty in the objective function of a classifier). Additionally, there is a number of hybrid modes [85] combining the different approaches. For more information on the available options, we refer the readers to general works on the topic [80,86−94].

Last but not least, many works report dimensionality reduction by means of PCA [78], which allows for a semi-automatized concentration of informative spectral content in a reduced set of variables. Nevertheless, these might not always represent an optimal space for the classifier training (see Section 2.3. for more details). Furthermore, as pointed out by Képeš et al. [95], with the increasing size of the modern LIBS data, the computational and memory requirements become significant. The authors, therefore, suggest Restricted Boltzmann Machine (RBM) - an unsupervised method based on artificial neural networks (ANN) - as an effective PCA alternative for the (pre-)processing of large LIBS datasets. With the increasing data volumes and recent advances in the deep learning field, the development of advanced ANN-based methods for automated feature extraction is currently on the rise (see the work of Zhao et al. presenting a convolutional neural network (CNN)-assisted strategy for classification of iron ores [96]) and is expected to further grow in the future [97−99].

Having the final variable set, it might be necessary to perform a column-wise (feature-wise) centering and scaling of the data matrix. This is especially important if the features have greatly different ranges (e.g., raw intensities and integrated peak areas), as otherwise, features with greater scales might dominate the analysis.

Highlighting the importance of proper data preprocessing and feature selection, PCA (introduced in more detail in the next chapter) is applied to an exemplary LIBS dataset of 23 different polymer composite materials (Fig. 4). The dataset was recorded using a commercial LIBS J200 system (Applied Spectra, Inc., Sacramento, CA) operating at a wavelength of 266 nm with a 6-channel Cherny Turner Spectrometer and CCD detection system under Ar

atmosphere. For each sample, 50 LIBS spectra on individual sample positions were recorded with each LIBS spectrum resulting from the accumulation of 50 laser shots (resulting in a total number of 2500 laser shots applied to each sample). Two different cases of data preprocessing are shown: In one case (proper data preprocessing), spectra are normalized (SNV) to compensate for instrumental drifts, atomic and molecular emission signals originating from the sample are identified and integrated with baseline subtraction, and the so-generated variables (total number of 41) are scaled prior to calculating the PCA. In the other case (no data preprocessing), all spectral intensities of the raw spectra are used as input variables for the PCA. Looking at the score-score plots, better separation of data from individual sample types is obtained after proper data preprocessing. The intra-class variability of each class is significantly reduced using data preprocessing (clusters are more concentrated at a location compared to more spread-out clusters obtained with no data preprocessing). Additionally, the number of outliers is reduced by data preprocessing. These two aspects are typically compensated by proper spectra normalization, noise removal, and background correction. Besides these considerations, score-score plots obtained after no data preprocessing might be biased showing information that is not chemically relevant since emission signals not originating from the sample are also included. Therefore, separation of two classes might be based on changing background or signals originating from the atmosphere (e.g., Ar, He, H, O, N). Using feature selection can not only circumvent this problem but can also simplify the interpretation of obtained results. Additionally, it should be noted that the computational time of the PCA increases significantly if no feature selection is carried out.

2.3. Exploratory data analysis (EDA)

Throughout the classifier development, it is often helpful to gain a better insight into the data. This is typically achieved by means of unsupervised techniques providing either a graphical representation of the multivariate space or other (non-visual) information about patterns and relationships within the data. Over the past few years, principal component analysis (PCA) has become one of the most widely employed techniques used for this purpose, which is
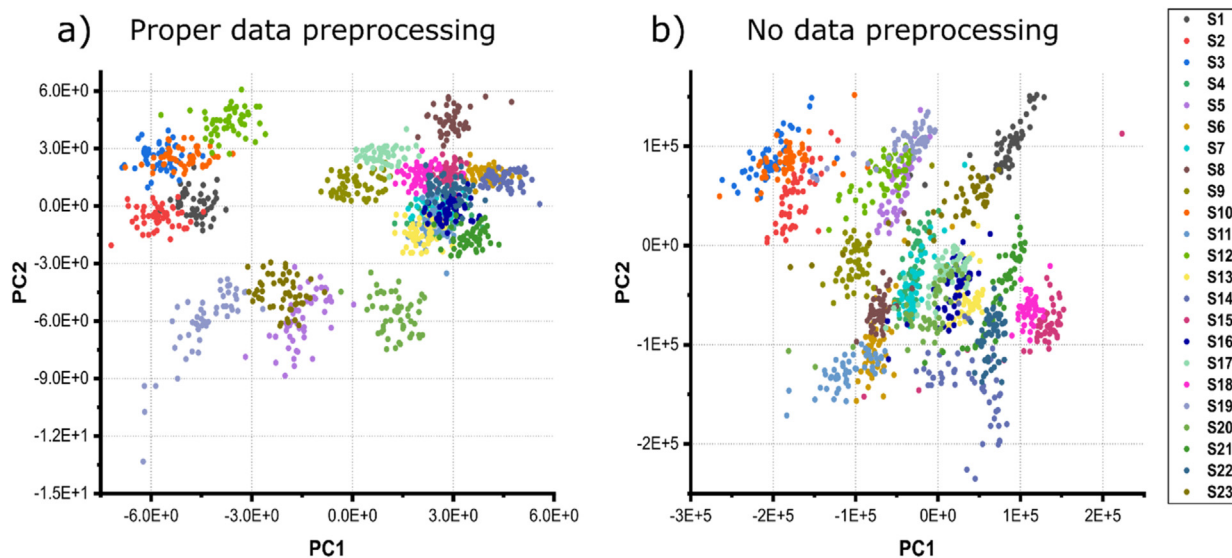
Fig. 4. Score-score plot of the same dataset of 23 polymer composite materials calculated with different data preprocessing. a) "Proper data preprocessing" including normalization of obtained spectra (SVM), feature selection and integration with baseline subtraction and scaling of the variables prior to calculating the PCA. b) "No data preprocessing" spectral intensities of raw LIBS spectra are used as input variables for the PCA.

well reflected in the LIBS classification literature where PCA occurs in a myriad of contexts, from data visualization to inference of variable importance and chemical relationships. As an extensive review of this topic was recently provided by Pořízka et al. [78], the following discussion is limited to the most important features and applications of the method.

Assuming that the maximum information is stored along with the directions of greatest data variation, PCA represents data in terms of a new coordinate system in which the individual dimensions (the so-called principal components) are mutually orthogonal and sorted in the descending order of the represented variance. The projections of data onto the planes spanned by different combinations of principal components can therefore result in greatly informative two-dimensional insights into the multivariate data (revealing data clusters and possible outliers). Furthermore, by displaying the loading/loading plots on top of the score plots (constructing the so-called bi-plot, Fig. 5), one can investigate the relationships between the clusters and spectral features contributing to their formation, which might greatly enhance the chemical understanding and interpretability of the data. Data shown in Fig. 5 is from the same experiment described in the previous section.

Nevertheless, when interpreting such visualizations, it is important to realize that PCA represents an insight into the data capturing the greatest variation in the spectral features, which might or might not represent the underlying variation among the classes. Thus, the clusters revealed in the plots generated by the first few PCs become representative of the true class distributions only if no other (interfering) sources of variation dominate. This is also important when employing PCA for dimensionality reduction (letting the classifier learn class distributions in the reduced space of PCA instead of the original one). Such an approach is based on the assumption that the first few retained principal components contain the information required for the class discrimination while the last components mainly contain noise.

The fact that the human brain is one of the best pattern recognizers in the world often leads to a tendency to use PCA as a classification tool. Nevertheless, as stressed in the work of Oliveri [100], this approach is erroneous as PCA is an unsupervised technique and therefore provides no trustworthy means of estimating the prediction error.
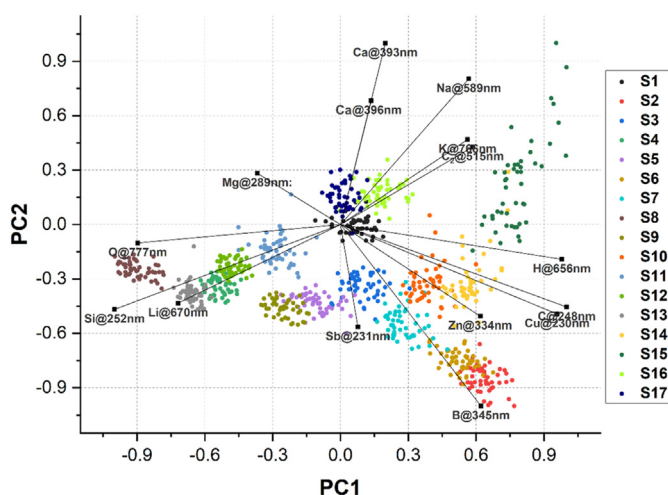

Fig. 5. Bi-plot showing the relationships between the clusters and emission lines responsible for their formation.

3. Supervised classification

Given the distribution of training data is representative of the true class distribution, one can employ a classification algorithm to estimate a mathematical function mapping the spectrum (input) to a class label (output). The "learned" function can subsequently be used for mapping new samples to their class labels. The process of training and classification can be perceived from different points of view.

From a probabilistic perspective, each class can be described in terms of a probability distribution. Thus, each point (which represents a spectrum) in the LIBS space can be assigned a probability of belonging to a particular class, which is greatest at the centroid of the distribution, evaluates to 0.5 at an interface of two classes (separation boundary) and further decreases to 0 as the distance of the point from the centroid increases. From this perspective, training can be seen as modeling the class distribution from the distribution of training data. Having established a model, one can identify new spectral points by evaluating their probability of belonging to each class and assigning them a label of the class exhibiting the highest probability.

From a geometric point of view, training can be perceived as a division of the LIBS space into subspaces representing the individual classes. Based on its spectral characteristics, a new point ends up in a particular region of the space (defined during training) and gets assigned a corresponding class label. Depending on the classifier employed, the class boundaries are modeled by a linear (e.g., LDA or PLS-DA) or a non-linear (e.g., kNN, ANN) function, which results either in a simple linear hyperplane or a more complex non-linear hypersurface separating the classes. The position and shape of the boundary are given by parameters learned during the training with the goal of achieving optimum discrimination of the training data. While linear functions provide robust discrimination in linearly separable cases, their lack of flexibility might result in poor separation of non-linear data. Although the non-linearities can be accounted for by non-linear classifiers, the increased flexibility of non-linear models combined with an effort to achieve perfect discrimination of the training data can result in boundaries perfectly separating the training data while delivering a poor prediction of new samples (known as overfitting resulting in poor generalization). The balance between the lack of flexibility (high bias) and poor generalization ability (high variance) of the classification model is known as the bias-variance trade-off (Fig. 6) and represents the ultimate challenge of designing a reliable classifier. In addition to the classifier type (linear or non-linear), it can be influenced by a set of hyperparameters controlling the training. Thus, each type of classifier is based on different assumptions about the data and comes with different hyperparameters to be tuned.

As implied by the "no free lunch" theorem [101], no single algorithm is able to deliver superior performance in all classification problems of the world. Therefore, the following section aims to provide an informative overview of the machine learning algorithms most commonly employed in the LIBS classification literature, together with specific application examples from the praxis. A note-worthy comparison of different data evaluation approaches and the performance of various classification algorithms was generated within the EMSLIBS 2019 contest. Here, the organizers provided an extensive LIBS dataset challenging the data handling skills of the participants. The results of the contest were published by Vrábel et al. [102] concluding "that even the simple classification (chemometric) techniques provided by a sufficient data preprocessing (spectroscopic data exploration, feature selection) can overcome the most recent machine learning approaches".
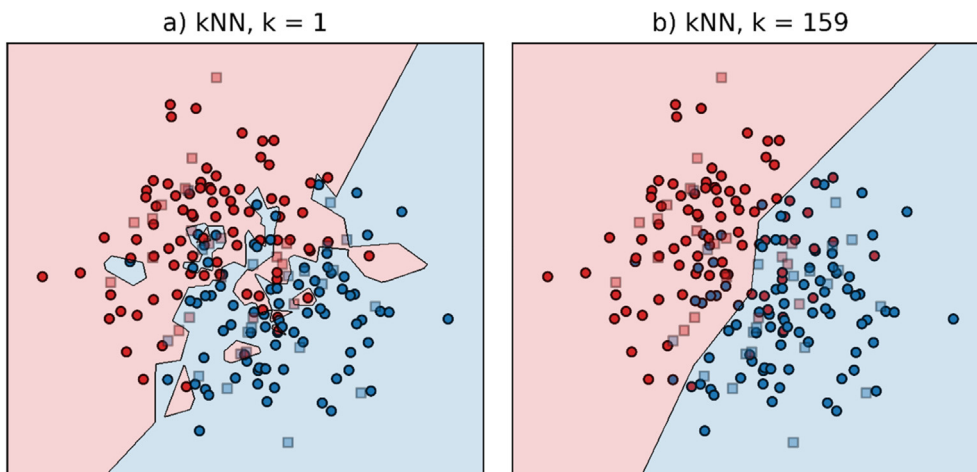
Fig. 6. Bias-variance trade-off. a) Overly flexible boundaries (overfitting, high variance). b) Lacking flexibility of the decision boundaries (high bias). Circles represent training data and squares represent test data.

3.1. Linear discriminant analysis (LDA)

In linear discriminant analysis (LDA) [103], the discrimination is based on a linear decision boundary optimally separating two-class distributions under the assumption that each class has a normal distribution and the same correlation structure (common variance-covariance matrix) (Fig. 7).

The probability density of each class is first estimated from the training data using the class centroid (position of the class in the space) and a pooled variance-covariance matrix (dispersion) common to all classes. The identification of new samples proceeds by computing the posterior probability of each class depending on the Mahalanobis distance (distance between the point and distribution mean corrected by the inverse covariance matrix) and the prior probability of the class (the fraction of all samples belonging to that class) and a subsequent assignment of the sample to the class with the greatest probability. The optimal decision boundary is formed where the density contours with equal probability of the class intersection (i.e., the Mahalanobis distance of the point to both class centroids adjusted by a "prior probability offset" is equal). In a multi-class scenario, one such linear delimiter is defined for each pair of the classes.

Despite the assumptions of class normality and common covariance (which are rather uncommon in real life), LDA can often result in sufficiently good discrimination of the classes. Nevertheless, in scenarios with a greater number of dimensions than samples (n > p), LDA has no mathematical solution. Thus, LDA is especially problematic in the field of spectroscopy, where the number of dimensions often exceeds the number of observations by far. This limitation can be addressed by combining LDA with a feature selection strategy (as done in the work of Pontes et al. on the classification of Brazilian soils [84]) or, more commonly, by a PCA dimensionality reduction. Examples of the PCA-LDA investigations from the field of LIBS include the classification of archaeological and paleontological materials [104], bricks from different localities [105], bone samples originating from different individuals [106], as well as differentiation of tissues as a real-time feedback mechanism during clinical laser surgery applications [107]. However, as discussed previously, PCA might or might not result in a training space representative of the true class distributions.

3.2. Partial least squares discriminant analysis (PLS-DA)

Partial least squares discriminant analysis (PLS-DA) is a classification technique combining dimensionality reduction (PLS) and linear discrimination (DA) in a single algorithm, which makes it especially suitable for handling the high-dimensional data common
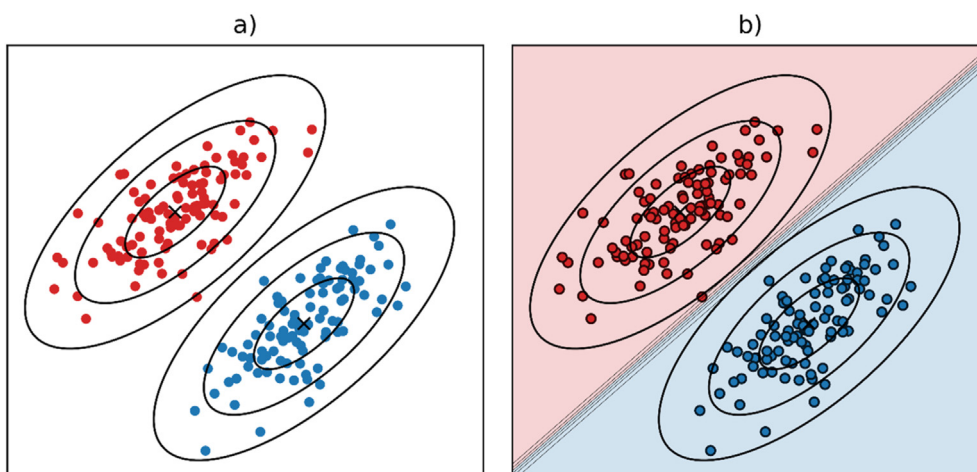
Fig. 7. Linear discriminant analysis. a) Distribution of the training data, classes have the same covariance matrix. b) Linear decision boundary found by the LDA.

in spectroscopy. In contrast to PCA, PLS searches for directions in space, maximizing the covariance between the spectral features and the class labels. The original set of features is reduced to a small number of so-called latent variables (LV) or factors capturing the greatest variation in the spectra relevant for the class discrimination. The DA part of the algorithm proceeds with finding a linear decision boundary allowing for the classification of new samples projected to the reduced space. The identification efficiency of a PLS-DA model is affected by the number of latent variables selected by the user. Whereas too few LVs might not retain enough information for constructing a reliable classifier, too many LVs might be counterproductive and result in overfitting. Obviously, the optimal number of LVs is characteristic of the given classification scenario and shall be determined by means of cross-validation. In the field of LIBS, PLS-DA was employed for the classification of explosives [108−110], consumer plastics [111], or minerals [109,111]. Furthermore, Merk et al. [112] developed an efficient approach for high-speed sorting of metal scrap based on PLS-DA. The authors provide an insightful description of developing a classifier for an industrial setting. The choice of PLS-DA was reasoned by the simplicity enabling for a fast evaluation of new samples − an aspect highly relevant for real-time analysis.

Despite the favorable ability of PLS-DA to handle entire LIBS spectra, the term PLS-DA encompasses an entire group of algorithm modifications, each suitable for a particular classification scenario. As discussed by Brereton and Lloyd [113], the lacking knowledge of the method might result in its inappropriate use and misinterpretation of the results. The interested readers are, therefore, greatly referred to a critical tutorial by Pomerantsev and Rodionova [114].

3.3. Support vector machines (SVM)

SVM is a very powerful method proposed by Vapnik and co-workers in 1992 [115], delivering a great performance in a wide range of classification scenarios, from linearly separable cases to highly complex non-linear problems. The binary (two-class) discrimination proceeds by means of a separating hyperplane constructed in a way to provide a good separation of the two classes while maximizing the distance (known as margin) between the hyperplane and the closest training observations on both sides of the plane (Fig. 8).

The exact position of the hyperplane is controlled by the so-called support vectors − usually a small number of training

observations located closest to the margin. The width of the margin is governed by the hyperparameter C reflecting the tolerance towards the misclassification of the training points. While high C values call for little misclassification tolerance resulting in a narrow margin highly fit to the data, lower C values allow for wider margins resulting in the poorer classification of the training data but a better generalization of the model (Fig. 9). Thus, the parameter C plays an important role in defining the variance-bias trade-off of the SVMs, which is important for achieving good performance. As the optimal value of C largely depends on the problem at hand, it shall be determined by means of cross-validation.

Once the separating hyperplane is defined, new samples can be assigned with a class label depending on which side of the boundary they fall.

Nevertheless, many real-life distributions are not linearly separable. To account for this fact, the SVM algorithm was extended to a non-linear approach by introducing the so-called kernel trick allowing the construction of separating hyperplanes in an expanded (non-linear) space without the need for its explicit construction (Fig. 10). This way, the inherently linear SVM achieves a non-linear separation of classes. One of the most widely applied kernels is the radial basis function (RBF), regulated by a hyperparameter gamma. In the case of applying RBF, one typically performs a simultaneous optimization of the hyperparameters C and gamma (e.g., by employing a grid search as demonstrated in the work of Sheng et al. [116]).

Although the original SVM concept applies to a two-class scenario, it can be extended to multiple classes by employing a one-versus-one or one-versus-all scenario followed by voting.

Due to their flexibility, SVMs can achieve superior performance in a wide range of scenarios, including high-dimensional spaces, which can be especially beneficial in the field of spectroscopy. However, the superior performance is conditioned by a suitable parameter tuning, which requires a good understanding of the algorithm and problem at hand. In the field of LIBS, SVMs were employed by Dingari et al. [117] for the analysis of pharmaceutical samples. The authors used RBF kernel and performed grid search over a wide range of hyperparameters to optimize the model. Additional works investigated SVMs for the classification of steel samples [118], iron ores [116], and sedimentary rocks [119]. Recently, Képeš et al. [120] demonstrated different approaches to the interpretation of SVM models on an example of 19 cyanobacterial strains.
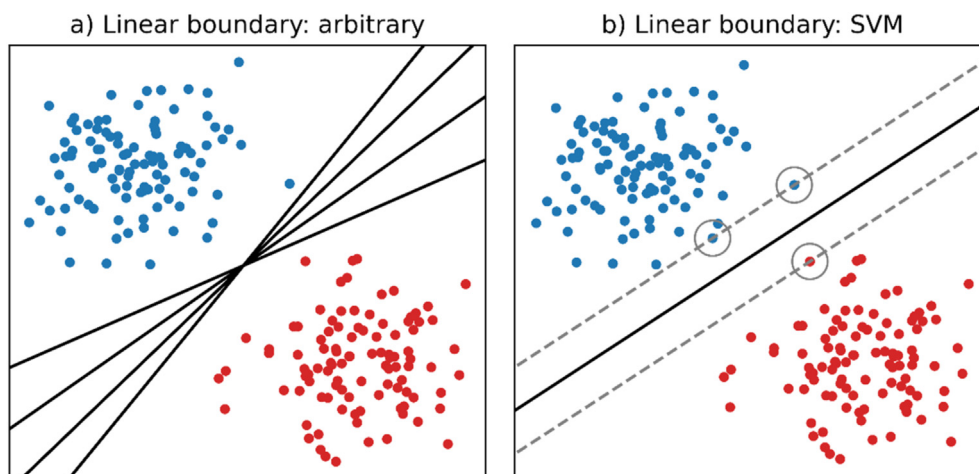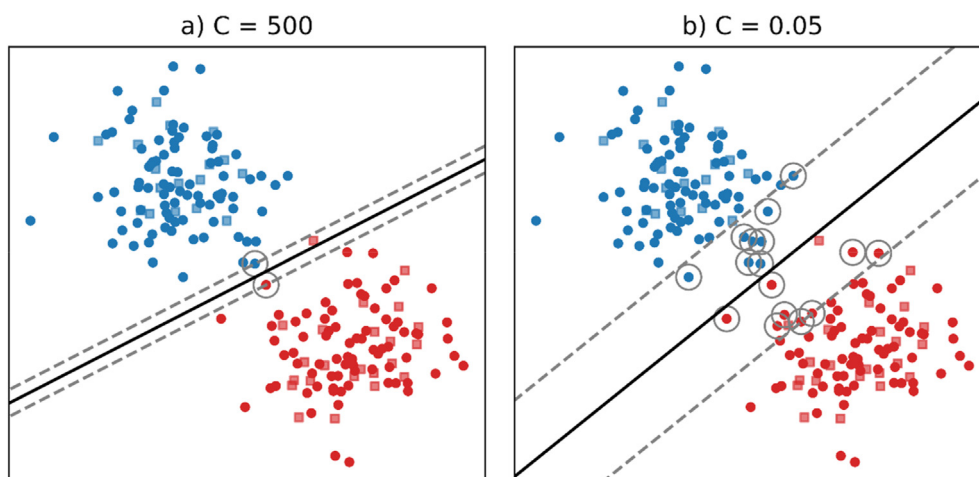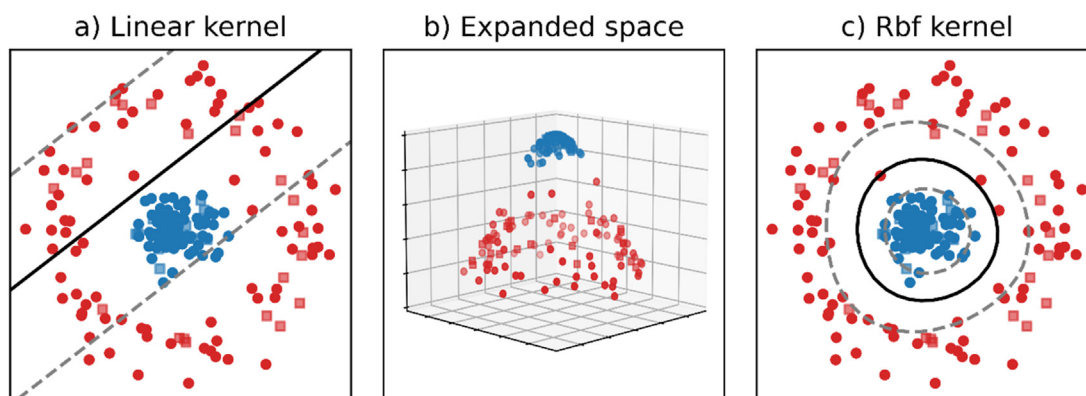
Fig. 8. Linear separation of the classes. a) There are infinitely many linear boundaries resulting in perfect separation of the training data. b) Decision boundary of the SVMs provides linear separation of the training data while maximizing the distance of the closest training samples to the boundary from both sides.

Fig. 9. Regularization of the linear SVMs. a) High C values result in a small number of support vectors defining the boundary and thus a higher tendency to overfit. b) Low C values allow for a greater number of support vectors and a better generalization of the decision boundary. Circles represent training data and squares represent test data.

Fig. 10. Kernel trick. a) The two classes cannot be separated by a linear decision boundary. b) Transformation of the space by a radial basis function allows for linear separation of the classes. c) Back-transformation of the data to the original space results in a non-linear decision boundary. Circles represent training data and squares represent test data.

3.4. K nearest neighbors (kNN)

kNN [121] is a very simple non-linear classification method based on the assumption that the similarity of the samples in the multivariate space is reliably represented by their mutual distances. The classification of a new sample is achieved by the identification of its k nearest neighbors in the training data (Fig. 11). The class of the unknown sample is derived from the class of the nearest neighbors with the greatest frequency.

As Fig. 6 shows, mapping the multivariate space according to this rule results in a non-linear decision boundary. While a k of 1 results in complex boundaries overfitting the training data, larger k results in their smoothening/linearization. Whereas the first scenario leads to misclassification of new samples due to poor generalization, the second scenario might result in wrong predictions due to insufficient flexibility of the decision boundary.

In addition to k, the efficiency of a kNN classifier greatly depends on the distance metrics: the better its ability to reflect the sample similarity, the better its performance. The distance-based nature of the kNN classifier additionally implies its sensitivity to measurement units and pre-processing steps. All of the above-mentioned (hyper)parameters shall be tuned by cross-validation to deliver an optimal classifier for the problem at hand. Nevertheless, the validity of these statements is conditioned by the basic assumption of kNN. As the dimensionality (the number of spectral features)

increases, the samples become spread across the outer range of the space, the mutual distances become similarly large, and the idea of the sample's "neighborhood" diminishes.

All in all, kNN is a simple and easy-to-implement method requiring no training. However, its reliability is conditioned by the preservation of local neighborhood (low-dimensional spaces or data "living" within a low-dimensional subspace or sub-manifold). Additionally, as the number of samples increases, kNN becomes computationally inefficient.

Specific application examples of kNN from the field of LIBS include classification of polymer e-waste [122], quality control of toys [123], and discrimination of soft tissues [124].

3.5. Random forests (RF)

Random forests [125] are robust non-linear classifiers based on the idea of "ensemble learning" — a method combining multiple weak learners (here decision trees) into a single classifier with superior performance. Training of a single decision tree involves the successive splitting of the multivariate space into subspaces according to a threshold on a variable resulting in the best separation of classes. The greedy nature of this process (i.e., its ultimate effort to achieve a perfect separation of the training data) results in a high tendency of the decision trees to overfit, which leads to a poor prediction power of the individual trees (Fig. 12).
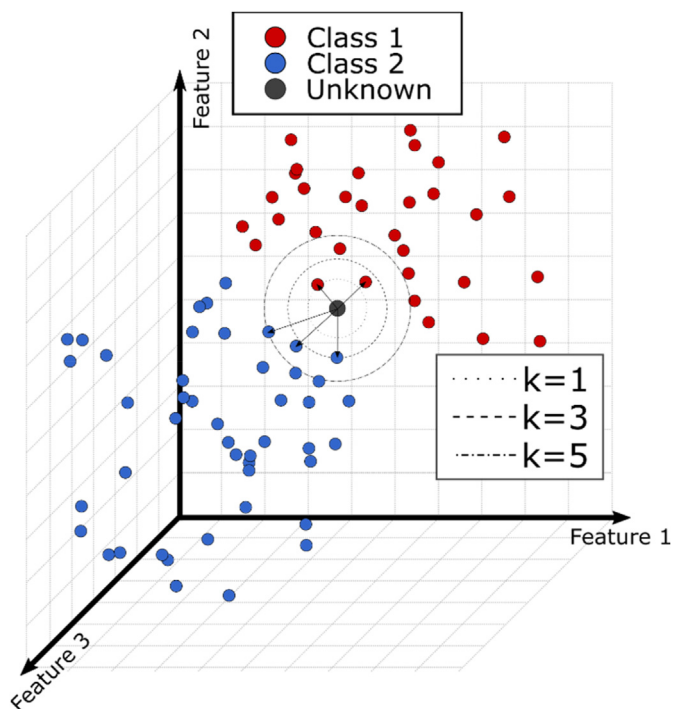
Fig. 11. Basic principle of kNN. Class assignment of new samples depends on the identity of k nearest training samples considered.

RF addresses this issue by combining multiple decision trees, each overfitting a different part of the data and averaging their results (Fig. 12). As the averaging effect improves with increasing tree diversity, each tree is trained on a slightly different subset (bootstrap sample) of the original data and considers only a random subset of all features at each splitting. The classification of new samples happens by majority voting of all trees. As each bootstrap sample includes only approx. 2/3 of the original observations, it is possible to estimate the predictive performance internally by means of the remaining 1/3 of observations left out from the training of each tree (known as out-of-bag sample (OOB)), which allows for even more efficient use of the training data than cross-validation. Additionally, the OOB sample is used to calculate the variable importance, which is another attractive trait of RF coming

"for free" with its training.

The main hyperparameters of RFs are the number of trees and the number of features considered at each splitting. The performance usually improves with the increasing number of trees, however, only up to a certain level (typically 70 to 100 trees). Adding more trees does not improve the performance but solely increases the computational complexity. The number of features considered has typically little effect and is often set to the empirically determined optimum of the square root of the total number of features [126].

All in all, the influence of hyperparameters on the predictive performance is usually smaller than in other advanced classification algorithms such as SVM. Thus, random forests often deliver great results even in the hands of non-experts, which gives them a unique position in the field. In the field of LIBS, RF was employed by Qi et al. [127] for the classification of archaeological ceramics. The authors used different pre-processing steps and optimized the model by trying different variable importance thresholds. Additional works employing RFs aimed at the classification of slag samples [128], iron ores [116], coal ash [129], and wines [130].

3.6. Artificial neural networks (ANN)

The term "artificial neural networks" (ANN) designates a set of highly complex and diverse classification algorithms which differ considerably in their architecture. The common feature of ANNs are rather small and simple processing units ("neurons") which are interconnected in various ways and which process one or several inputs to form an output signal which is then propagated to other connected neurons. While the transfer function (also called "activation function") of an individual neuron is rather simple (for example, a hyperbolic tangent applied to the weighted sum of the inputs), the complexity of the overall model function arises from the connection of many of these neurons.

Depending on the structure and the types of connections, different types of networks can be distinguished, for example, multi-layer perceptrons (MLP), convolutional neural networks (CNN), or self-organizing maps (SOM). As the MLP type predominates in the LIBS literature, the following discussion will be constrained to this architecture. Nevertheless, the fast development in the ANN field and the great range of application possibilities it has to offer result in ever greater interest of the LIBS community in this technique. In order to gain a greater insight into the topic, we
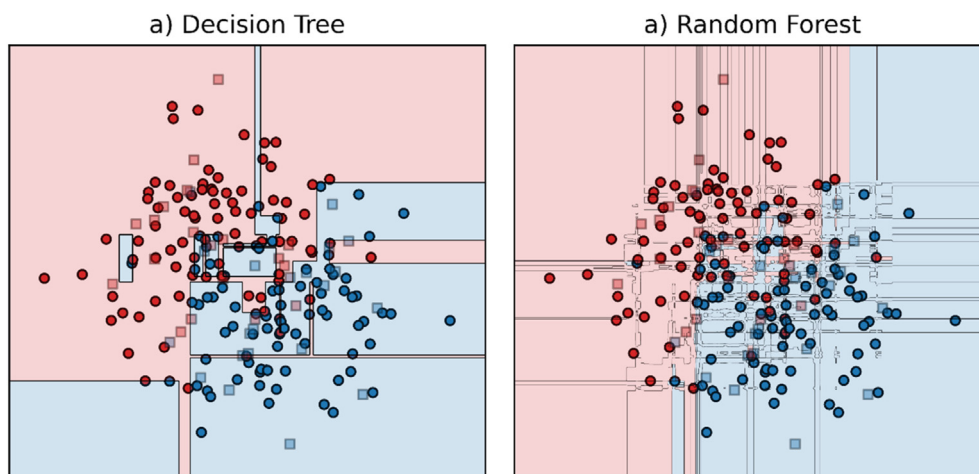
Fig. 12. Non-linear decision boundaries of tree-based methods. a) Decision boundaries of a single decision tree tend to overfit the training data. b) Averaging of 100 decision trees results in decision boundaries with greater generalization ability. Circles represent training data and squares represent test data.

refer the readers to a recent review of Li et al. [131], providing an exhaustive overview of the intersection between the LIBS and ANN fields.

The MLP architecture in its modern form was introduced by Werbos in 1974, who provided the backpropagation of errors as a then-new training method [181]. As depicted in Fig. 13, a typical MLP consists of an input layer (each neuron represents a particular spectral variable) and an output layer (each neuron represents a particular class) connected by one or several hidden layers. This structure enables a non-linear mapping of the input (spectral features) to the output (class label). When applying an ANN to classify a spectrum, each neuron of the input layer is connected to a particular spectral feature which is then fed forward to the next layers until the signal front arrives at the output layer. The individual layers are fully connected so that each neuron receives all output signals of the previous layers, weighting the received signals before processing the sum of these signals by the non-linear activation function and passing it on to the next layer. In the final step, the output is rescaled by the so-called softmax function to values representing the class probabilities.

The "strength" of the individual connections is governed by weights adjusted during the training by means of the backpropagation algorithm, which tries to minimize the output errors in an iterative way.

This process has several implications for the application: The basic ANN architecture and the wide range of options it offers (number of hidden layers, number of neurons in each layer, learning rate, etc.) results in great flexibility and ability of ANNs to learn any kind of discriminatory hypersurface (Fig. 14). Nevertheless, the more complicated the architecture becomes, the greater the number of parameters (weights) is which need to be adapted during the training phase and the greater the risk of overfitting becomes (especially if the number of training samples is small). Additionally, a higher number of neurons increases the computational demand resulting in long training times.

All in all, ANNs provide extreme flexibility and exciting application potential. However, their performance heavily depends on the knowledge and experience of the analyst as there is a considerable set of hyperparameters to be tuned. A detailed listing of specific ANN examples in the field of LIBS can be found in the above-mentioned review of Li et al. [131]. For more details we

Fig. 13. Typical structure of a multilayer perceptron (MLP). Each neuron in the input layer represents a particular spectral feature, and each neuron in the output layer is one of the target classes. Neurons in the hidden layer(s) are responsible for a non-linear mapping of the spectral features to the classes.

would like to refer the interested reader to the excellent comprehensive introduction and overview of ANNs applied for LIBS data analysis provided in the book chapter by Vrábel et al. [132].

3.7. Soft independent modelling of class analogy (SIMCA)

Soft independent modeling of class analogy (SIMCA) is the only representative of the so-called class-modeling approaches presented in the current review. It is the oldest and most commonly used class-modeling approach [133,134] and gave a lot of impetus to classification methods in general. SIMCA results in a one-class classifier that positively identifies a class without looking at other (in contrast to, for example, discriminant techniques, such as PLS/DA). The class modeling approach is in some way similar to an "inverse" outlier detection, posing the question of whether a particular observation belongs to an assumed distribution or not.

The principle behind SIMCA is to create either a box or an ellipsoid around a particular class. Because different classes are normally recognized by different sets of variables, SIMCA tries to focus on the significant class-specific variables by processing the data for each class separately. The significant variables for a particular class are extracted by calculating the principal components (PCs) of the particular class, retaining only those components which contribute to the class model. The identified (latent) variables then form a p-dimensional coordinate system, where p can be different for each class (Fig. 15). The class box (or the ellipsoid) is then formed by establishing limits along each axis such that the majority of the class data is located within these limits. For ellipsoid models, the Mahalanobis distance from the class center is used instead of the hard limits along the axes. The critical distances of the classification boundaries are established by calculating the 95% (or 99%) confidence intervals for the class to be within the interval.

Unknown objects are classified by projecting them into the PC space of each class and checking whether this object falls into the particular class boundaries. Please note that an unknown observation may fall into several class boxes resulting in ambiguities (hence the expression "soft" in the name of the method because there may not be a hard clear-cut decision for a class). Thus, SIMCA can produce overlapping regions of classes, which may show up adversely for the classification process on the one hand but, on the other hand, can be used to indicate the quality of the model simply by counting the number of observations that fall into several class boxes.

Several modifications of the SIMCA approach have been proposed and evaluated [135,136]. A meanwhile common approach is to use distances obtained from a combination of the sum of squared residuals and the Mahalanobis distance of the sample to the class center [137]. This approach has the advantage that the classification rule for an unknown sample is very simple ($d^2 < 2$ if the T and Q statistics are normalized to the 0.95-quantile).

In the context of LIBS, SIMCA was applied for the classification of pharmaceutical tablets [138], characterization of historical building materials [139], polymer e-waste [122], and warfare agents [93].

3.8. Model validation

Having established a classifier, its prediction ability shall be validated using an independent set of samples that is representative of the future application scenario. As discussed previously, a new set of samples is often not readily available, and the evaluation is typically done on a small fraction of the original data (approx. 20%), which has not been used during the training. The comparison of the true class labels with the ones predicted by a binary classifier (class label of 1 if sample belongs to the class, 0 otherwise) results in 4 possible outcomes (true positive, true negative, false positive,
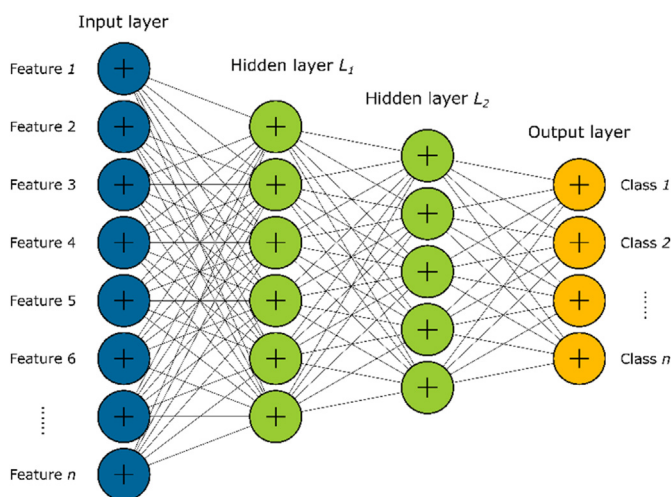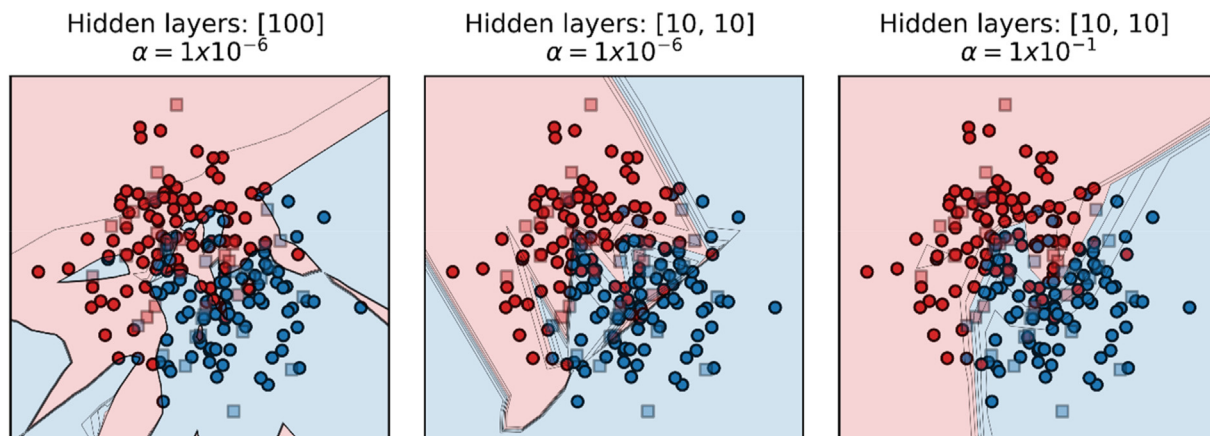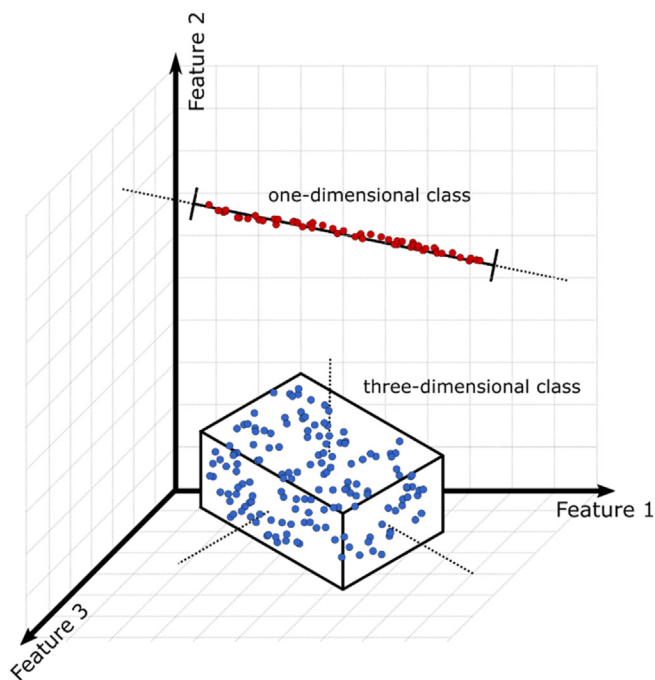
Fig. 14. Dependence of the MLP decision boundaries on the hyperparameter settings. The great number of options offers great flexibility of the classifier as well as great danger of overfitting in case of lacking expertise in the field. Circles represent training data and squares represent test data.

Fig. 15. Class models of SIMCA. Each class is modeled separately by calculating the principal components of the class and retaining only PCs which contribute to the class model.

**Fig. 16.** Confusion matrix for a binary classification.

misclassification rates and types of errors made for each class. Last but not least, in real-time applications (e.g., identification of metals in the recycling industry [112]), the time required for the identification might become equally or even more important than the prediction accuracy.

4. Applications

With LIBS providing fast and simultaneous multielement analysis, it is a very suitable technique for sample classification based on elemental fingerprinting and molecular information. Providing ease of use and the capability to analyze all kinds of solid samples without any restrictions (e.g., samples do not have to be stable in high vacuum or do not have to be electrically conductive), LIBS-based classification has been applied to a vast range of application fields ranging from materials science (e.g., polymers, steel, alloys, ceramics, …), life science applications (e.g., foodstuff, bacteria, tissue, plants, …), geology (rocks, minerals, ores, soil, …), to forensics (e.g., gunshot residuals, explosives, …). Fig. 17 provides an interactive figure enabling exploration of the revised LIBS classification literature of the last decade. The figure is based on a generated database containing relevant information (application field, number of classes, feature selection, applied algorithm, model validation, …) from all revised papers in this review. For exploration, the literature can be clustered by this information providing insights into the main chemometric approaches commonly employed in various fields or by different research groups.

and false negative), which are typically tracked in a confusion matrix (Fig. 16). This provides the basis for the calculation of different quality metrics summarized in Table 1.

Whereas in the field of LIBS, accuracy has become the most commonly employed metric to report the quality of a classifier, certain application fields such as medicine or detection of explosives might require reports of additional metrics such as specificity or sensitivity. If reporting a single metric to characterize the performance of a classifier, it is important to keep in mind that each metric has its limitations − e.g., accuracy is not suitable in imbalanced scenarios as the majority of a particular class will bias the overall result. Therefore, the use of alternatives (e.g., Matthews Correlation Coefficient (MCC)) and report of multiple metrics shall be considered. Additionally, in the case of multi-class scenarios, it is often convenient to report the entire confusion matrix tracking the

Table 1
Summary of the most commonly used metrics for evaluation of the classifier's performance.

Metric	Formula	Meaning
Accuracy (%)	$\dfrac{TP + TN}{TP + FP + TN + FN} \times 100$	Percentage of correct predictions out of all predictions made. Max: 100%
MCC ($-$)	$\dfrac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$	Correlation between the observed and predicted samples. Suitable for imbalanced scenarios. Max: 1
Sensitivity/Recall (%)	$\dfrac{TP}{TP + FN} \times 100$	Percentage of actual positives (e.g., explosives) identified correctly. Max: 100%
Specificity (%)	$\dfrac{TN}{TN + FP} \times 100$	Percentage of actual negatives (e.g., non-explosives) identified correctly. Max: 100%
Precision (%)	$\dfrac{TP}{TP + FP} \times 100$	Percentage of positive predictions which was correct. Max: 100%
F-score	$\dfrac{TP}{TP + \frac{1}{2}(FP + FN)}$	Harmonic mean of precision and sensitivity. Max: 1
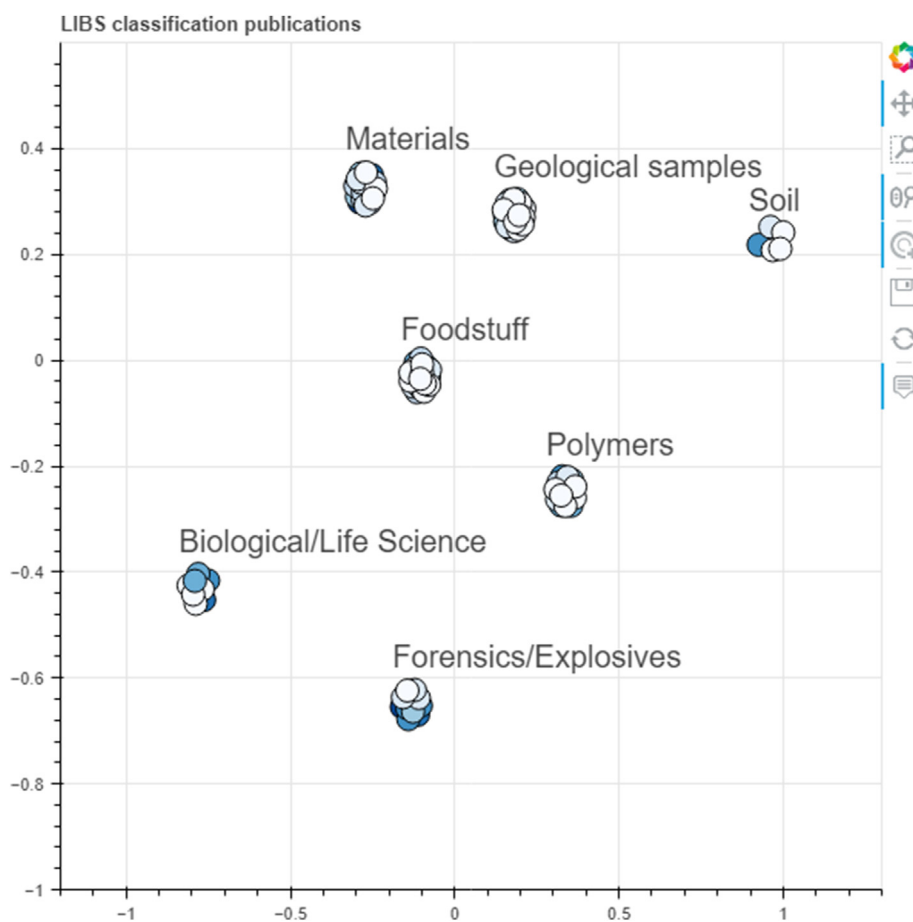
Fig. 17. Interactive figure providing insights into the revised LIBS classification literature of the last decade. All LIBS-based classification literature explicitly mentioned within this review as well as other works are included in the interactive figure [140–219]. Publications can be clustered by various variables (application field, applied algorithm, data pre-processing, …) and the node size and node color can be adjusted accordingly. The interactive figure can be accessed via this link. https://mybinder.org/v2/gh/gajzuz/binder-bokehh-test.git/HEAD?urlpath=%2Fproxy%2F5006%2Fbokeh-app.

Additionally, a table containing an extensive list of publications in LIBS-based classification, giving an overview of the various scientific fields where it is applied regularly, and which classification algorithms are commonly used is provided in the supporting file. The table also highlights works where data fusion with other techniques (e.g., Raman spectroscopy, LA-ICP-MS, …) was employed, adding information to the dataset used for training the classifier. Additionally, information is provided whether special instrumentation such as stand-off LIBS or double-pulse LIBS, or other approaches such as nanoparticle enhanced LIBS (NELIBS) was used.

The following chapter highlights the advantages of LIBS-based classification for specific research fields, and a selection of outstanding publications in the corresponding area are presented and discussed.

4.1. Geomaterials

The first reported works of using LIBS for the classification of geomaterials were already published at the beginning of the 21st century by McManus [220], McMilan [221,222], and Harmon [223,224]. An overview of the application of LIBS for the investigation of geological samples with a focus on elemental fingerprinting for the classification is provided by Hark et al. [225]. The unique capabilities making LIBS an excellent tool for the characterization of geomaterials include a sufficient sensitivity for light

elements (H, Li, Be, B, and C) and the possibility to detect F, O, S, and N, which is important in this field.

An extensive range of different geomaterials (carbonates, fluorites, silicate rocks, and soil) was investigated in the work of Gottfried et al. [226]. Additionally, the performance of a table-top instrument, a double-pulse system, and a stand-off LIBS device were compared. This work used PCA for exploratory data analysis, and PLS-DA was used for sample classification. The work demonstrates that PCA is a versatile tool to investigate large datasets of LIBS spectra of geological samples investigating composition, origin, and color. Moreover, PLS-DA can classify most of the investigated sample types correctly. The stand-off system's outstanding performance in terms of geomaterials classification enables investigations in the field in real-time.

The possibility of carrying out laterally resolved analysis using LIBS was exploited in the work of Moncayo et al. [227]. In this work, an area of 468 mm^2 of a thin section bloc of a turquoise sample consisting of three main mineral phases (turquoise, pyrite, and silica) was mapped with a lateral resolution of 15 μm. This mapping generated a large dataset of more than 2 million LIBS spectra. As a first step, the authors conducted a conventional data analysis investigating the intensity map of individual elements of geological interest (Fe, Si, Cu, P, and Al). In the next step, PCA was used for exploratory data analysis looking at intensity maps of individual scores. Looking at the loadings, it was concluded that PC1 showed the lateral distribution of pyrite and turquoise, whereas PC2 represents the silicate phase. Besides using PCA as a tool to identify the distribution of the different phases present, low concentration compounds were also identified. This is typically a difficult task as it includes finding emission signals which are only present in a very small number of spectra within a large dataset. Looking at higher-order PCs, the authors were able to identify Mg and Ti within their dataset, which might have easily been missed without exploratory data evaluation.

An outstanding contest in the field of LIBS-based classification was held at the EMSLIBS2019 conference in Brno, Czech Republic, based on data obtained from 138 different soil samples belonging to 12 distinct classes. A training dataset with labels and a test dataset without labels was provided to the participants to build a classification model which generalizes well [228]. Several different research groups and individuals participated in this contest, all using different strategies. The approaches of the five teams obtaining the best results are presented by Vrábel et al. [102]. This summary highlights that a wide range of machine learning algorithms can lead to satisfying results. Nevertheless, feature selection and dimensionality reduction were vital steps in all five approaches highlighting their importance when building reliable classification models. This demonstrates that for the classification of LIBS data, both expertise in spectroscopy as well as data science is of importance.

4.2. Materials

LIBS is widely applied to the analysis and classification of all kinds of different materials. Especially in metalworking, LIBS is used for the classification of steel, various alloys, slag, and metal scrap. Here the possibility of online measurements using LIBS with no required sample preparation and fast sample throughput, e.g., in recycling plants, is an interesting aspect for the industry.

In the work of Kong et al. [168], steel samples are analyzed and classified using LIBS. In this work, the authors investigated the influence of the variable selection from the LIBS spectra on the classification performance of PCA and ANN. The three spectral feature selection approaches include selecting all emission signals in the spectra, selecting only rather intensive emission signals, and using the whole spectrum. This study showed that using only the most intense spectral features resulted in the most robust classification model highlighting the importance of proper variable selection for a classification task.

Merk et al. [112] developed a LIBS-based system to identify metal scrap directly on a conveyor band. Therefore, two laser shots are applied to each sample: The first shot removes potential dirt, paint, or an oxide layer, and the second shot is used for the actual measurement. The samples analyzed in this study to train a classifier were metal scrape of 9 different alloys from a recycling facility representing a real-life classification scenario. Data pretreatment considering possible harsh environments at a recycling plant that might cause spectral shifts was carried out. The authors validated the performance of their developed PLS-DA model with a validation data set and investigated the robustness of the classifier by investigating data recorded with different laser and spectrometer systems. Finally, the model is tested for online classification of samples on a conveyer belt with satisfying results.

The possibility of aluminum alloy identification in post-consumer scrap on a moving conveyor belt (3 m s^{-1}) using 3D scanning LIBS was described by Werheit et al. [170]. To have access to a measurement volume of $600 \times 600 \times 100$ mm^3, the authors use galvo-scanner mirrors to deflect the laser beam within a few milliseconds onto the sample of interest enabling up to 6 measurements on each sample while passing by. For classification, a total of 60 aluminum cast and 168 aluminum wrought pieces of 8 different alloys were analyzed. Data preprocessing is used to discard spectra where insufficient plasma ignition due to the laser not hitting the desired measurement position occurred. Being able to classify the 8 different aluminum alloys with a correctness of >95%, an application in an industrial sorting facility is feasible.

4.3. Biological samples/Life science

The application of LIBS for the classification of biological material and samples from the field of life sciences ranges from plants, hard and soft tissue to bacteria. In these fields, LIBS's beneficial properties are especially the high sample throughput and the possibility to detect the most common elements in biological samples (C, H, N, O, S, …). Classification of bacteria using LIBS spectra is a field that has been steadily growing over the last 15 years. The interested reader is kindly referred to review articles by Steven Rehse [229] and Gaudiuso et al. [230], giving a comprehensive overview of the work performed in this field.

Lung tumor tissue and healthy boundary tissue were discriminated using LIBS in the work of Lin et al. [231]. Therefore, tissue sections of tumor tissue and boundary tissue of 45 different patients were prepared. In this work, recorded spectra are pretreated using a wavelet de-noising procedure and are normalized. The multivariate data analysis uses a selection of emission signals, including C, Mg, CN, Ca, Na, and K. Since measurements were carried out under air, signals from H, N, and O are neglected. The authors investigated the performance of different classification algorithms using 10-fold cross-validation, concluding the possibility of discriminating between tumor tissue and boundary tissue.

Another work dealing with the distinction between cancerous tissue and healthy tissue was carried out by Choi et al. [232] by mapping a tissue section of murine skin with embedded melanoma with 15 μm lateral resolution. Investigating LIBS spectra of melanoma and dermis, 12 signals were identified. Using a maximum likelihood estimation of intensity ratios of the 12 identified signals, each element's ability to discriminate between the two tissue types was estimated, and K was selected as the element showing the most significant discrimination power. Additionally, SVM with a 3rd order polynomial kernel function was used to build a classifier

for melanoma and dermis tissue classification. Applying the SVM to the LIBS mapping, the classification of each pixel reveals the lateral distribution of cancerous tissue, which is in good agreement with conventional H&E staining.

Archaeological bone samples were classified using LIBS in the work of Siozos et al. [233]. In this work, the authors investigated not only the classification capabilities but also examined the influence contamination of the burial soil has on the classification results. The work demonstrates that a careful feature selection is necessary to avoid overfitting of the developed classification model by excluding signals not originating from the sample under investigation itself but rather from contaminations.

4.4. Foodstuff

Characterization of all kinds of foodstuff is an important field regarding authentication, origin determination, and investigation of adulteration. In this field, elemental fingerprinting is a widely used approach since the elemental fingerprint is usually expected to be similar to the soil composition [234]. As LIBS enables simultaneous multielement analysis and provides access to all elements of the periodic table, it is a promising technique for this field. The preparation of a representative sample of foodstuff for LIBS analysis is often challenging since these samples often require drying or conversion of liquid samples into solids to improve the quality of the analysis.

Origin determination of 20 different kinds of rice was carried out in the work of Yang et al. [235]. In this work, emphasis was put on different sample preparation strategies and how these influence the classification performance. Sample preparation methods are based on pressed pellets, including different additives with ground and whole rice grains. The authors identified 90 spectral lines, including both atomic/ionic (C, H, O, N, K, Ca, Na, Mg, Al, Mn, Si) and molecular signals (C2, CN), which were used as input variables after normalization to the carbon line at 247.86 nm. Using PCA for variable reduction the computational time is reduced and overfitting is avoided. Therefore, the first 30 PCs were used as an input for SVM classification. Classification accuracies were determined by a 5-fold cross-validation ranging from 94.1% to 99.25%, depending on the sample preparation method.

Origin determination of thirty-eight different red wines from eleven protected designations of origin was carried out by Moncayo et al. [236]. Therefore, the wines under investigation were transformed into a gel using commercial collagen, which facilitated the subsequent LIBS analysis. The selection of spectral features was carried out by excluding signals originating from the collagen and using only signals from the wine samples (Mg, Ca, K, and Na). This decreases the discrimination power of the applied neural network but increases the ability to generalize. The authors tested their model's sensitivity, generalization ability, and robustness and concluded that LIBS can correctly classify the geographic origin of a wine.

Different meat species were identified by Sezer et al. [237]. Therefore, dried protein extracts of processed beef, chicken, and pork meat were prepared. Additional to samples containing only one type of meat, blends were also prepared for quantitative analysis. Even though a wide range of elemental emission signals was identified, the authors used the whole recorded LIBS spectrum as input for PCA and PLS analysis. PCA revealed a separation of the three different types of meat based on the recorded LIBS spectra. Additionally, a PLS model was calculated based on the different blends. Here the authors report satisfying RMSC, RMSEP, RSD, and REP for the validation data set, confirming the possibility of a quantitative assessment of beef, chicken, and pork mixtures.

4.5. Forensics/Explosives

The capabilities of LIBS to be used as a stand-off device enabling probing samples from a distance or using handheld instruments are beneficial for investigating potential hazardous and suspicious samples such as explosives. Using LIBS spectra has been proven to be able to distinguish between different explosive materials using their characteristic emission signals in the LIBS spectrum. Similar to the classification of polymers, C, H, N, and O, as well as the molecular fragments C_2 and CN can be used to distinguish between different explosive materials [238,239].

De Lucia et al. [240] used a stand-off LIBS instrument to distinguish between three different explosives (cyclo-trimethylenetrinitramine (RDX), trinitrotoluene (TNT), and Composition-B) deposited on various car panels. Additional to the application of explosives, non-explosive samples such as road dust, sand, diesel fuel, lubricant oil, and fingerprints were prepared. The main goal of this study was to investigate the influence of various substrates on the classification performance. Single-shot LIBS spectra were recorded with a stand-off instrument (25–30 m distance), and PLS-DA was calculated to classify samples either as explosive or nonexplosive. Investigating variable importance scores (VIP), the authors could identify signals only originating from the explosives. Using only these variables, computational time was significantly reduced, and classification performance was enhanced.

In forensics science, elemental fingerprinting is widely used to get additional information and insights for crime investigations. Twenty window glass samples from different crime scenes were analyzed using LIBS by El-Deftar et al. [241], and the discrimination capabilities were compared to LA-ICP-MS, μXRF, and SEM-EDX, techniques conventionally used for this task. First, the authors assessed LODs and precision of several elements (Ba, K, Sr, Ti) in LIBS analysis of glass samples using NIST610, NIST 612, and NIST1831. Additionally, the stability of LIBS analysis over seven days was investigated. Using ratios of emission signals with ANOVA and Turkey's HSD test, all three above-mentioned techniques showed a discrimination power greater than 96%. Finally, the authors state that LIBS analysis has significant advantages over LA-ICP-MS, μXRF, and SEM-EDX due to its fast sample throughput, simple instrumentation, and low cost.

Laterally resolved discrimination of latent fingerprints using elemental patterns obtained from LIBS analysis was investigated by Yang et al. [242]. Samples were prepared by applying individual and overlapping fingerprints on an Aluminum substrate. Different elements were identified in recorded LIBS spectra of fingerprints, including Fe, Al, Ca, Na, K, and O. Analyzing the data using PCA with the whole spectrum as an input, a clear separation of 4 different fingerprints in the score-score plot was observed. Using SIMCA, an average classification accuracy of 90.36% was achieved. PLS-DA was also investigated, showing similar performance. Finally, they mapped overlapping fingerprints and reconstructed the arrangement using their multivariate classification models.

4.6. Polymers

Conventionally, LIBS is considered a technique used for elemental analysis. Nevertheless, under certain circumstances, LIBS spectra do not only show elemental signals but also exhibit signals from molecules either originating from the sample itself due to incomplete atomization or from recombination within the plasma. Combining elemental signals from the main constituents of polymers H, C, N, and O with signals from molecular fragments proves to be capable of distinguishing different polymer types. Additionally, the possibility to perform online and stand-off analysis and

measurements not being constrained by sample color makes LIBS a promising tool for polymer sorting in recycling plants.

In the work of Junjuri et al. [209], five of the most common post-consumer plastics (PET, HDPE, LDPE, PP, PS) were analyzed using fs-LIBS. The investigated samples were obtained from a local recycling unit providing a close real-life scenario of polymer classification. This is especially important since pristine and well-defined samples are often used for method development. In contrast, in a real-life application, surface contaminations, uneven shapes, or other variations must be considered leading to a significant bias in the performance of a classifier. When training the classifier (ANN), the authors investigated the influence of feature selection by providing five different sets of variables generated from the recorded LIBS data. The authors reported identification rates up to 100%.

Gajarska et al. [243] successfully discriminated 20 different polymer types, the highest number reported in the literature. This work investigated the effect of different experimental conditions (laser energy, atmosphere, and gate delay), and careful optimization of feature selection based on a Random Forest was carried out. Additionally, in this work, the influence of various organic and inorganic additives commonly used in polymer manufacturing on the discrimination ability of polymer samples was investigated. It was concluded that additives present in the samples do not interfere with polymer type identification.

A recent work by Sommer et al. [244] applied LIBS to identify microplastics. Therefore, the authors analyzed environmental samples from the Lahn river (Marburg, Germany). The collected microplastics were first identified using FT-IR spectroscopy (PS, PE, PP, PA). In the next step, LIBS reference spectra of the corresponding polymer types were recorded. Comparing LIBS spectra of microplastics with pristine polymer samples show that the polymer-specific signals are still present in the microplastics. However, the spectrum usually has additional spectral features, which are often also present in naturally occurring particles resulting in challenging discrimination. Nevertheless, PCA analysis of the dataset shows that microplastics can be differentiated from naturally occurring particles, and PE and PS can be distinguished.

5. Conclusion

LIBS-based classification has experienced a rapidly growing interest in the last years. Providing unique characteristics such as simultaneous multi-element analysis with little to no sample preparation necessary, online and stand-off analysis capabilities, and imaging and depth profiling LIBS has many benefits over other techniques used for elemental fingerprinting. These properties make LIBS a suitable technique for classification tasks in a diverse range of application fields discussed within this review. Especially applications in the field of recycling and online material sorting, or online quality control in production facilities in general can pave the way of LIBS for routine industrial applications. Besides these promising industrial applications, rapid LIBS-based classification in life-science may have a significant impact on diagnostics in the medical field in the future.

Nevertheless, to successfully establish LIBS as a state-of-the-art technique capable of sample classification, a refinement of the community's chemometric awareness is necessary: designing the experiment, optimizing measurement conditions, and recording the data required for the successful development of a classifier is only one step. The other step, including data preprocessing, feature selection and training, and optimizing the classification model, is not less important. Each small step and decision made in this process of training a classification model can significantly influence the performance and generalizability of the model and should only be carried out after careful consideration.

Nowadays, a wide range of machine learning algorithms and tools are readily available even to non-experts enabling the training of a classifier with only a few lines of code. Nevertheless, one should keep in mind that building a classification model is not only about applying an algorithm to the dataset. Hyperparameters should be optimized, feature selection/data reduction/data normalization considered, different algorithms investigated, and the performance of the final model should be tested. Additionally, the classifier's performance should not only be evaluated with one test dataset recorded on the same day as the training dataset, but the stability of the model over a particular time should be considered.

With the development of new instrumentation, especially lasers providing higher repetition rates and detection systems providing more data points per spectra, the amount of data generated per time will increase even further. Therefore, data handling and proper data evaluation assisted by machine learning and other chemometric tools combined with exploratory data analysis will become even more critical in the future since manual data screening becomes impossible. Thus, it is of major importance that LIBS users can rely on and apply these tools properly.

With increasing complexity and size of the spectroscopic data, the typical LIBS users should possess basic understanding of analytical chemistry, plasma physics, and machine learning. It is the interdisciplinary approach to the LIBS data processing that gains the best possible performance. When combining this expert knowledge of experimental design, carefully optimized measurement parameters, and chemical knowledge of the operator for feature selection with a well-thought-out classification model construction, LIBS can provide reliable, robust, and reproducible results. Therefore, it is crucial to note that the successful development of a classification model requires both expert knowledge in the field of LIBS and in chemometrics and machine learning.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trac.2022.116859.

References

[1] C.A. Wallace, L. Manning, Food Provenance: Assuring Product Integrity and Identity, CABI Rev., 2020.

[2] I. Rouse, The classification of artifacts in archaeology, Am. Antiq. 25 (1960) 313−323.

[3] G. Zadora, Glass analysis for forensic purposes—a comparison of classification methods, J. Chemom. 21 (2007) 174−186. https://doi.org/10.1002/cem.1030.

[4] P. Kiddee, R. Naidu, M.H. Wong, Electronic waste management approaches: an overview, Waste Manag. 33 (2013) 1237−1250. https://doi.org/10.1016/j.wasman.2013.01.006.

[5] M. Chanda, Chemical aspects of polymer recycling, Adv. Ind. Eng. Polym. Res. 4 (2021) 133−150. https://doi.org/10.1016/j.aiepr.2021.06.002.

[6] S. Kelly, K. Heaton, J. Hoogewerff, Tracing the geographical origin of food: the application of multi-element and multi-isotope analysis, Trends Food Sci. Technol. 16 (2005) 555−567. https://doi.org/10.1016/j.tifs.2005.08.008.

[7] A. Gonzalvez, S. Armenta, M. de la Guardia, Trace-element composition and stable-isotope ratio for discrimination of foods with Protected Designation of Origin, TrAC Trends Anal. Chem. 28 (2009) 1295−1311. https://doi.org/10.1016/j.trac.2009.08.001.

[8] L. Li, C.E. Boyd, Z. Sun, Authentication of fishery and aquaculture products by multi-element and stable isotope analysis, Food Chem. 194 (2016) 1238−1244. https://doi.org/10.1016/j.foodchem.2015.08.123.

[9] Q. Xiong, Y. Lin, W. Wu, J. Hu, Y. Li, K. Xu, X. Wu, X. Hou, Chemometric intraregional discrimination of Chinese liquors based on multi-element determination by ICP-MS and ICP-OES, Appl. Spectrosc. Rev. 56 (2021) 115−127. https://doi.org/10.1080/05704928.2020.1742729.

[10] V.F. Taylor, H.P. Longerich, J.D. Greenough, Multielement analysis of Canadian wines by inductively coupled plasma mass spectrometry (ICP-MS) and multivariate statistics, J. Agric. Food Chem. 51 (2003) 856−860. https://doi.org/10.1021/jf025761v.

[11] S.A. Drivelos, C.A. Georgiou, Multi-element and multi-isotope-ratio analysis to determine the geographical origin of foods in the European Union, TrAC Trends Anal. Chem. 40 (2012) 38−51. https://doi.org/10.1016/j.trac.2012.08.003.

[12] A. González, S. Armenta, M. de la Guardia, Geographical traceability of Arròs de Valencia rice grain based on mineral element composition, Food Chem. 126 (2011) 1254−1260. https://doi.org/10.1016/j.foodchem.2010.11.032.

[13] Q. Jin, F. Liang, H. Zhang, L. Zhao, Y. Huan, Daqian Song, Application of microwave techniques in analytical chemistry, TrAC Trends Anal. Chem. 18 (1999) 479−484. https://doi.org/10.1016/S0165-9936(99)00110-7.

[14] C.A. Bizzi, M.F. Pedrotti, J.S. Silva, J.S. Barin, J.A. Nóbrega, E.M.M. Flores, Microwave-assisted digestion methods: towards greener approaches for plasma-based analytical techniques, J. Anal. At. Spectrom. 32 (2017) 1448−1466. https://doi.org/10.1039/C7JA00108H.

[15] G. Friedbacher, H. Bubert, Surface and Thin Film Analysis: A Compendium of Principles, Instrumentation, and Applications, John Wiley & Sons, 2011.

[16] R.E. Russo, X. Mao, J.J. Gonzalez, V. Zorba, J. Yoo, Laser ablation in analytical chemistry, Anal. Chem. 85 (2013) 6162−6177. https://doi.org/10.1021/ac4005327.

[17] J.S. Becker, A. Matusch, B. Wu, Bioimaging mass spectrometry of trace elements − recent advance and applications of LA-ICP-MS: a review, Anal. Chim. Acta 835 (2014) 1−18. https://doi.org/10.1016/j.aca.2014.04.048.

[18] L. Jolivet, M. Leprince, S. Moncayo, L. Sorbier, C.-P. Lienemann, V. Motto-Ros, Review of the recent advances and applications of LIBS-based imaging, Spectrochim. Acta Part B At. Spectrosc. 151 (2019) 41−53. https://doi.org/10.1016/j.sab.2018.11.008.

[19] A. Limbeck, L. Brunnbauer, H. Lohninger, P. Pořízka, P. Modlitbová, J. Kaiser, P. Janovszky, A. Kéri, G. Galbács, Methodology and applications of elemental mapping by laser induced breakdown spectroscopy, Anal. Chim. Acta 1147 (2021) 72−98. https://doi.org/10.1016/j.aca.2020.12.054.

[20] R. Scadding, V. Winton, V. Brown, An LA-ICP-MS trace element classification of ochres in the Weld Range environ, Mid West region, Western Australia, J. Archaeol. Sci. 54 (2015) 300−312. https://doi.org/10.1016/j.jas.2014.11.017.

[21] M.N.C. Grainger, M. Manley-Harris, S. Coulson, Classification and discrimination of automotive glass using LA-ICP-MS, J. Anal. At. Spectrom. 27 (2012) 1413−1422. https://doi.org/10.1039/C2JA30093A.

[22] A. van Es, J. de Koeijer, G. van der Peijl, Discrimination of document paper by XRF, LA−ICP−MS and IRMS using multivariate statistical techniques, Sci. Justice 49 (2009) 120−126. https://doi.org/10.1016/j.scijus.2009.03.006.

[23] G. Zadora, Z. Brożek-Mucha, SEM−EDX—a useful tool for forensic examinations, Mater. Chem. Phys. 81 (2003) 345−348. https://doi.org/10.1016/S0254-0584(03)00018-X.

[24] P. Fruhstorfer, R. Niessner, Identification and classification of airborne soot particles using an automated SEM/EDX, Microchim. Acta 113 (1994) 239−250. https://doi.org/10.1007/BF01243614.

[25] R.P. Alvarez, P.J.M. Van Espen, R. Rosa Plá, E. Montoya Rossi, R. Arrazcaeta Delgado, P.P. Godo Torres, M. Celaya González, Compositional classification of archaeological pottery based on INAA and SEM-EDX, J. Trace Microprobe Tech. 21 (2003) 677−695. https://doi.org/10.1081/TMA-120025818.

[26] V. Panchuk, I. Yaroshenko, A. Legin, V. Semenov, D. Kirsanov, Application of chemometric methods to XRF-data − a tutorial review, Anal. Chim. Acta 1040 (2018) 19−32. https://doi.org/10.1016/j.aca.2018.05.023.

[27] L. Bonizzoni, A. Galli, M. Gondola, M. Martini, Comparison between XRF, TXRF, and PXRF analyses for provenance classification of archaeological bricks, X Ray Spectrom. 42 (2013) 262−267. https://doi.org/10.1002/xrs.2465.

[28] I. Allegretta, B. Marangoni, P. Manzari, C. Porfido, R. Terzano, O. De Pascale, G.S. Senesi, Macro-classification of meteorites by portable energy dispersive X-ray fluorescence spectroscopy (pED-XRF), principal component analysis (PCA) and machine learning algorithms, Talanta 212 (2020), 120785. https://doi.org/10.1016/j.talanta.2020.120785.

[29] R. Padilla, P.V. Espen, P.P.G. Torres, The suitability of XRF analysis for compositional classification of archaeological ceramic fabric: a comparison with a previous NAA study, Anal. Chim. Acta 558 (2006) 283−289. https://doi.org/10.1016/j.aca.2005.10.077.

[30] M. Gaft, L. Nagli, I. Gornushkin, Y. Raichlin, Review on recent advances in analytical applications of molecular emission and modelling, Spectrochim. Acta Part B At. Spectrosc. 173 (2020), 105989. https://doi.org/10.1016/j.sab.2020.105989.

[31] K. Liu, D. Tian, C. Li, Y. Li, G. Yang, Y. Ding, A review of laser-induced breakdown spectroscopy for plastic analysis, TrAC Trends Anal. Chem. 110 (2019) 327−334. https://doi.org/10.1016/j.trac.2018.11.025.

[32] J.M. Anzano, C. Bello-Gálvez, R.J. Lasheras, Identification of polymers by means of LIBS, in: S. Musazzi, U. Perini (Editors), Laser-Induc. Breakdown Spectrosc. Theory Appl., Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, pp. 421−438. https://doi.org/10.1007/978-3-642-45085-3_15.

[33] G. Galbács, A critical review of recent progress in analytical laser-induced breakdown spectroscopy, Anal. Bioanal. Chem. 407 (2015) 7537−7562. https://doi.org/10.1007/s00216-015-8855-1.

[34] E. Tognoni, G. Cristoforetti, S. Legnaioli, V. Palleschi, Calibration-free laser-induced breakdown spectroscopy: state of the art, Spectrochim. Acta Part B At. Spectrosc. 65 (2010) 1−14. https://doi.org/10.1016/j.sab.2009.11.006.

[35] D.W. Hahn, N. Omenetto, Laser-induced breakdown spectroscopy (LIBS), Part II: review of instrumental and methodological approaches to material analysis and applications to different fields, Appl. Spectrosc. 66 (2012) 347−419. https://doi.org/10.1366/11-06574.

[36] R. Bellman, Dynamic Programming, Princet. Univ. Press 19S7, 1957.

[37] P. Pořízka, J. Klus, D. Prochazka, E. Képeš, A. Hrdlička, J. Novotný, K. Novotný, J. Kaiser, Laser-Induced Breakdown Spectroscopy coupled with chemometrics for the analysis of steel: the issue of spectral outliers filtering, Spectrochim. Acta Part B At. Spectrosc. 123 (2016) 114−120. https://doi.org/10.1016/j.sab.2016.08.008.

[38] C. Beleites, R. Salzer, Assessing and improving the stability of chemometric models in small sample size situations, Anal. Bioanal. Chem. 390 (2008) 1261−1271. https://doi.org/10.1007/s00216-007-1818-6.

[39] T. Hastie, R. Tibshirani, J.H. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, 2009. https://link.springer.com/book/10.1007/978-0-387-84858-7.

[40] G. Schulze, A. Jirasek, M.M.L. Yu, A. Lim, R.F.B. Turner, M.W. Blades, Investigation of selected baseline removal techniques as candidates for automated implementation, Appl. Spectrosc. 59 (2005) 545−574. https://doi.org/10.1366/0003702053945985.

[41] J.P. Singh, S.N. Thakur, Laser-Induced Breakdown Spectroscopy, Elsevier, 2007.

[42] A.W. Miziolek, V. Palleschi, I. Schechter, Laser Induced Breakdown Spectroscopy, Cambridge University Press, 2006.

[43] A. De Giacomo, M. Dell'Aglio, R. Gaudiuso, S. Amoruso, O. De Pascale, Effects of the background environment on formation, evolution and emission spectra of laser-induced plasmas, Spectrochim. Acta Part B At. Spectrosc. 78 (2012) 1−19. https://doi.org/10.1016/j.sab.2012.10.003.

[44] J.-B. Sirven, P. Mauchien, B. Sallé, Analytical optimization of some parameters of a Laser-Induced Breakdown Spectroscopy experiment, Spectrochim. Acta Part B At. Spectrosc. 63 (2008) 1077−1084. https://doi.org/10.1016/j.sab.2008.08.013.

[45] I.B. Gornushkin, P.E. Eagan, A.B. Novikov, B.W. Smith, J.D. Winefordner, Automatic correction of continuum background in laser-induced breakdown and Raman spectrometry, Appl. Spectrosc. 57 (2003) 197−207.

[46] L. Sun, H. Yu, Automatic estimation of varying continuum background emission in laser-induced breakdown spectroscopy, Spectrochim. Acta Part B At. Spectrosc. 64 (2009) 278−287. https://doi.org/10.1016/j.sab.2009.02.010.

[47] M.S. Friedrichs, A model-free algorithm for the removal of baseline artifacts, J. Biomol. NMR 5 (1995) 147−153.

[48] P. Yaroshchyk, J.E. Eberhardt, Automatic correction of continuum background in Laser-induced Breakdown Spectroscopy using a model-free algorithm, Spectrochim. Acta Part B At. Spectrosc. 99 (2014) 138−149. https://doi.org/10.1016/j.sab.2014.06.020.

[49] B. Tan, M. Huang, Q. Zhu, Y. Guo, J. Qin, Detection and correction of laser induced breakdown spectroscopy spectral background based on spline interpolation method, Spectrochim. Acta Part B At. Spectrosc. 138 (2017) 64−71. https://doi.org/10.1016/j.sab.2017.10.012.

[50] L. Lt, L. Jg, Z. Nj, L. Cp, C.D., S.H., W. Cl, Z. Yj, L. Wq, Study on the automatic extraction method of spectral data features in laser induced breakdown spectroscopy, Guang Pu Xue Yu Guang Pu Fen Xi Guang Pu 31 (2011) 3285−3288.

[51] E. Képeš, P. Pořízka, J. Klus, P. Modlitbová, J. Kaiser, Influence of baseline subtraction on laser-induced breakdown spectroscopic data, J. Anal. At. Spectrom. 33 (2018) 2107−2115. https://doi.org/10.1039/C8JA00209F.

[52] D.W. Hahn, N. Omenetto, Laser-induced breakdown spectroscopy (LIBS), Part I: review of basic diagnostics and plasma−particle interactions: still-challenging issues within the analytical plasma community, Appl. Spectrosc. 64 (2010) 335A−366A.

[53] M. Bonta, H. Lohninger, M. Marchetti-Deschmann, A. Limbeck, Application of gold thin-films for internal standardization in LA-ICP-MS imaging experiments, Analyst 139 (2014) 1521–1531. https://doi.org/10.1039/c3an01511d.

[54] B.C. Castle, K. Talabardon, B.W. Smith, J.D. Winefordner, Variables influencing the precision of laser-induced breakdown spectroscopy measurements, Appl. Spectrosc. 52 (1998) 649–657. https://doi.org/10.1366/0003702981944300.

[55] J. Guezenoc, A. Gallet-Budynek, B. Bousquet, Critical review and advices on spectral-based normalization methods for LIBS quantitative analysis, Spectrochim. Acta Part B At. Spectrosc. 160 (2019), 105688. https://doi.org/10.1016/j.sab.2019.105688.

[56] P. Poŕízka, J. Klus, A. Hrdlička, J. Vrábel, P. Škarková, D. Prochazka, J. Novotný, K. Novotný, J. Kaiser, Impact of Laser-Induced Breakdown Spectroscopy data normalization on multivariate classification accuracy, J. Anal. At. Spectrom. 32 (2017) 277–288. https://doi.org/10.1039/C6JA00322B.

[57] N.B. Zorov, A.A. Gorbatenko, T.A. Labutin, A.M. Popov, A review of normalization techniques in analytical atomic spectrometry with laser sampling: from single to multivariate correction, Spectrochim. Acta Part B At. Spectrosc. 65 (2010) 642–657. https://doi.org/10.1016/j.sab.2010.04.009.

[58] J. Pricylla Castro, E. Rodrigues Pereira-Filho, Twelve different types of data normalization for the proposition of classification, univariate and multivariate regression models for the direct analyses of alloys by laser-induced breakdown spectroscopy (LIBS), J. Anal. At. Spectrom. 31 (2016). https://doi.org/10.1039/C6JA00224B, 2005–2014.

[59] S.I. Gornushkin, I.B. Gornushkin, J.M. Anzano, B.W. Smith, J.D. Winefordner, Effective normalization technique for correction of matrix effects in laser-induced breakdown spectroscopy detection of magnesium in powdered samples, Appl. Spectrosc. 56 (2002) 433–436. https://doi.org/10.1366/0003702021955088.

[60] Z.-B. Ni, X.-L. Chen, H.-B. Fu, J.-G. Wang, F.-Z. Dong, Study on quantitative analysis of slag based on spectral normalization of laser-induced plasma image, Front. Physiol. 9 (2014) 439–445.

[61] P. Zhang, L. Sun, H. Yu, P. Zeng, L. Qi, Y. Xin, An image auxiliary method for quantitative analysis of laser-induced breakdown spectroscopy, Anal. Chem. 90 (2018) 4686–4694. https://doi.org/10.1021/acs.analchem.7b05284.

[62] D. Body, B.L. Chadwick, Optimization of the spectral data processing in a LIBS simultaneous elemental analysis system, Spectrochim. Acta Part B At. Spectrosc. 56 (2001) 725–736. https://doi.org/10.1016/S0584-8547(01)00186-0.

[63] W.B. Barnett, V.A. Fassel, R.N. Kniseley, Theoretical principles of internal standardization in analytical emission spectroscopy, Spectrochim. Acta Part B At. Spectrosc. 23 (1968) 643–664. https://doi.org/10.1016/0584-8547(68)80045-X.

[64] I. Konz, B. Fernández, M.L. Fernández, R. Pereiro, H. González, L. Alvarez, M. Coca-Prados, A. Sanz-Medel, Gold internal standard correction for elemental imaging of soft tissue sections by LA-ICP-MS: element distribution in eye microstructures, Anal. Bioanal. Chem. 405 (2013) 3091–3096. https://doi.org/10.1007/s00216-013-6778-4.

[65] C. Austin, F. Fryer, J. Lear, D. Bishop, D. Hare, T. Rawling, L. Kirkup, A. McDonagh, P. Doble, Factors affecting internal standard selection for quantitative elemental bio-imaging of soft tissues by LA-ICP-MS, J. Anal. At. Spectrom. 26 (2011) 1494–1501. https://doi.org/10.1039/C0JA00267D.

[66] Y. Bi, K. Yuan, W. Xiao, J. Wu, C. Shi, J. Xia, G. Chu, G. Zhang, G. Zhou, A local pre-processing method for near-infrared spectra, combined with spectral segmentation and standard normal variate transformation, Anal. Chim. Acta 909 (2016) 30–40. https://doi.org/10.1016/j.aca.2016.01.010.

[67] P. Heraud, B.R. Wood, J. Beardall, D. McNaughton, Effects of pre-processing of Raman spectra on in vivo classification of nutrient status of microalgal cells, J. Chemom. 20 (2006) 193–197. https://doi.org/10.1002/cem.990.

[68] D. Syvilay, N. Wilkie-Chancellier, B. Trichereau, A. Texier, L. Martinez, S. Serfaty, V. Detalle, Evaluation of the standard normal variate method applied to the discrimination of painting layers, Spectrochim. Acta Part B At. Spectrosc. 114 (2015) 38–45. https://doi.org/10.1016/j.sab.2015.09.022.

[69] F. Liu, L. Ye, J. Peng, K. Song, T. Shen, C. Zhang, Y. He, Fast detection of copper content in rice by laser-induced breakdown spectroscopy with Uni- and multivariate analysis, Sensors 18 (2018) 705. https://doi.org/10.3390/s18030705.

[70] E. Tognoni, G. Cristoforetti, [INVITED] signal and noise in laser induced breakdown spectroscopy: an introductory review, Opt Laser. Technol. 79 (2016) 164–172. https://doi.org/10.1016/j.optlastec.2015.12.010.

[71] J. Schlenke, L. Hildebrand, J. Moros, J.J. Laserna, Adaptive approach for variable noise suppression on laser-induced breakdown spectroscopy responses using stationary wavelet transform, Anal. Chim. Acta 754 (2012) 8–19. https://doi.org/10.1016/j.aca.2012.10.012.

[72] J.M. Mermet, P. Mauchien, J.L. Lacour, Processing of shot-to-shot raw data to improve precision in laser-induced breakdown spectrometry microprobe, Spectrochim. Acta B At. Spectrosc. 63 (2008) 999–1005. https://doi.org/10.1016/j.sab.2008.06.003.

[73] C.M. Sundling, N. Sukumar, H. Zhang, M. Embrechts, C. Breneman, Wavelets in chemistry and cheminformatics, Rev. Comput. Chem. 22 (2006) 295–329.

[74] B. Zhang, L. Sun, H. Yu, Y. Xin, Z. Cong, Wavelet denoising method for laser-induced breakdown spectroscopy, J. Anal. At. Spectrom. 28 (2013) 1884–1893. https://doi.org/10.1039/C3JA50239B.

[75] P. Lu, Z. Zhuo, W. Zhang, J. Tang, H. Tang, J. Lu, Accuracy improvement of quantitative LIBS analysis of coal properties using a hybrid model based on a

[76] wavelet threshold de-noising and feature selection method, Appl. Opt. 59 (2020) 6443–6451. https://doi.org/10.1364/AO.394746.

[76] W. Wang, S. Li, H. Qi, B. Ayhan, C. Kwan, S. Vance, Revisiting the pre-processing procedures for elemental concentration estimation based on chemcam libs on mars rover 6th Workshop Hyperspectral Image Signal Process, Evol. Remote Sens. Whisp. (2014) 1–4. https://doi.org/10.1109/WHISPERS.2014.8077520, 2014.

[77] S. Theodoridis, K. Koutroumbas, Pattern recognition and neural networks, in: G. Paliouras, V. Karkaletsis, C.D. Spyropoulos (Editors), Mach. Learn. Its Appl. Adv. Lect., Springer, Berlin, Heidelberg, 2001, pp. 169–195. https://doi.org/10.1007/3-540-44673-7_8.

[78] P. Poŕízka, J. Klus, E. Képeš, D. Prochazka, D.W. Hahn, J. Kaiser, On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review, Spectrochim. Acta Part B At. Spectrosc. 148 (2018) 65–82. https://doi.org/10.1016/j.sab.2018.05.030.

[79] H. Lohninger, J. Ofner, Multisensor hyperspectral imaging as a versatile tool for image-based chemical structure determination, Spectrosc. Eur. 26 (2014) 6–10.

[80] A. Jovic, K. Brkic, N. Bogunovic, A review of feature selection methods with applications, in: 2015 38th Int. Conv. Inf. Commun. Technol. Electron. Microelectron. MIPRO, IEEE, Opatija, Croatia, 2015, pp. 1200–1205. https://doi.org/10.1109/MIPRO.2015.7160458.

[81] C. Huffman, H. Sobral, E. Terán-Hinojosa, Laser-induced breakdown spectroscopy spectral feature selection to enhance classification capabilities: a t-test filter approach, Spectrochim. Acta Part B At. Spectrosc. 162 (2019), 105721. https://doi.org/10.1016/j.sab.2019.105721.

[82] R. Kohavi, G.H. John, Wrappers for feature subset selection, Artif. Intell. 97 (1997) 273–324. https://doi.org/10.1016/S0004-3702(97)00043-X.

[83] S.F.C. Soares, A.A. Gomes, M.C.U. Araujo, A.R.G. Filho, R.K.H. Galvão, The successive projections algorithm, TrAC Trends Anal. Chem. 42 (2013) 84–98. https://doi.org/10.1016/j.trac.2012.09.006.

[84] M.J.C. Pontes, J. Cortez, R.K.H. Galvão, C. Pasquini, M.C.U. Araújo, R.M. Coelho, M.K. Chiba, M.F. de Abreu, B.E. Madari, Classification of Brazilian soils by using LIBS and variable selection in the wavelet domain, Anal. Chim. Acta 642 (2009) 12–18. https://doi.org/10.1016/j.aca.2009.03.001.

[85] F. Ruan, L. Hou, T. Zhang, H. Li, A novel hybrid filter/wrapper method for feature selection in archaeological ceramics classification by laser-induced breakdown spectroscopy, Analyst 146 (2021) 1023–1031. https://doi.org/10.1039/D0AN02045A.

[86] B.H. Menze, B.M. Kelm, R. Masuch, U. Himmelreich, P. Bachert, W. Petrich, F.A. Hamprecht, A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data, BMC Bioinf. 10 (2009) 213. https://doi.org/10.1186/1471-2105-10-213.

[87] G. Chandrashekar, F. Sahin, A survey on feature selection methods, Comput. Electr. Eng. 40 (2014) 16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024.

[88] Z. Zhao, F. Morstatter, S. Sharma, S. Alelyani, A. Anand, H. Liu, Advancing feature selection research, ASU Feature Sel. Repos (2010) 1–28.

[89] P. Jafari, F. Azuaje, An assessment of recently published gene expression data analyses: reporting experimental design and statistical factors, BMC Med. Inf. Decis. Making 6 (2006) 27. https://doi.org/10.1186/1472-6947-6-27.

[90] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, J. Mach. Learn. Res. 3 (2003) 1157–1182.

[91] J. Cai, J. Luo, S. Wang, S. Yang, Feature selection in machine learning: a new perspective, Neurocomputing 300 (2018) 70–79. https://doi.org/10.1016/j.neucom.2017.11.077.

[92] J. Li, K. Cheng, S. Wang, F. Morstatter, R.P. Trevino, J. Tang, H. Liu, Feature selection: a data perspective, ACM Comput. Surv. 50 (94) (2017) 1–94. https://doi.org/10.1145/3136625, 45.

[93] C.A. Munson, F.C. De Lucia, T. Piehler, K.L. McNesby, A.W. Miziolek, Investigation of statistics strategies for improving the discriminating power of laser-induced breakdown spectroscopy for chemical and biological warfare agent simulants, Spectrochim. Acta Part B At. Spectrosc. 60 (2005) 1217–1224. https://doi.org/10.1016/j.sab.2005.05.017.

[94] Y. Saeys, I. Inza, P. Larrañaga, A review of feature selection techniques in bioinformatics, Bioinformatics 23 (2007) 2507–2517. https://doi.org/10.1093/bioinformatics/btm344.

[95] E. Képeš, J. Vrábel, P. Poŕízka, J. Kaiser, Addressing the sparsity of laser-induced breakdown spectroscopy data with randomized sparse principal component analysis, J. Anal. At. Spectrom. 36 (2021) 1410–1421. https://doi.org/10.1039/D1JA00067E.

[96] W. Zhao, C. Li, C. Yan, H. Min, Y. An, S. Liu, Interpretable deep learning-assisted laser-induced breakdown spectroscopy for brand classification of iron ores, Anal. Chim. Acta 1166 (2021), 338574. https://doi.org/10.1016/j.aca.2021.338574.

[97] J. Acquarelli, T. van Laarhoven, J. Gerretzen, T.N. Tran, L.M.C. Buydens, E. Marchiori, Convolutional neural networks for vibrational spectroscopic data analysis, Anal. Chim. Acta 954 (2017) 22–31. https://doi.org/10.1016/j.aca.2016.12.010.

[98] X. Zhang, J. Xu, J. Yang, L. Chen, H. Zhou, X. Liu, H. Li, T. Lin, Y. Ying, Understanding the learning mechanism of convolutional neural networks in spectral analysis, Anal. Chim. Acta 1119 (2020) 41–51. https://doi.org/10.1016/j.aca.2020.03.055.

[99] J. Chen, J. Pisonero, S. Chen, X. Wang, Q. Fan, Y. Duan, Convolutional neural

network as a novel classification approach for laser-induced breakdown spectroscopy applications in lithological recognition, Spectrochim. Acta Part B At. Spectrosc. 166 (2020), 105801. https://doi.org/10.1016/j.sab.2020.105801.

[100] P. Oliveri, C. Malegori, E. Mustorgi, M. Casale, Qualitative pattern recognition in chemistry: theoretical background and practical guidelines, Microchem. J. 162 (2020), 105725. https://doi.org/10.1016/j.microc.2020.105725.

[101] D.H. Wolpert, W.G. Macready, No free lunch theorems for optimization, IEEE Trans. Evol. Comput. 1 (1997) 67–82. https://doi.org/10.1109/4235.585893.

[102] J. Vrábel, E. Képeš, L. Duponchel, V. Motto-Ros, C. Fabre, S. Connemann, F. Schreckenberg, P. Prasse, D. Riebe, R. Junjuri, M.K. Gundawar, X. Tan, P. Pořízka, J. Kaiser, Classification of challenging Laser-Induced Breakdown Spectroscopy soil sample data - EMSLIBS contest, Spectrochim. Acta Part B At. Spectrosc. 169 (2020), 105872. https://doi.org/10.1016/j.sab.2020.105872.

[103] R.A. Fisher, The use of multiple measurements in taxonomic problems, Ann. Eugen 7 (1936) 179–188. https://doi.org/10.1111/j.1469-1809.1936.tb02137.x.

[104] G. Vítková, K. Novotný, L. Prokeš, A. Hrdlička, J. Kaiser, J. Novotný, R. Malina, D. Prochazka, Fast identification of biominerals by means of stand-off laser-induced breakdown spectroscopy using linear discriminant analysis and artificial neural networks, Spectrochim. Acta Part B At. Spectrosc. 73 (2012) 1–6. https://doi.org/10.1016/j.sab.2012.05.010.

[105] G. Vítková, L. Prokeš, K. Novotný, P. Pořízka, J. Novotný, D. Všianský, L. Čelko, J. Kaiser, Comparative study on fast classification of brick samples by combination of principal component analysis and linear discriminant analysis using stand-off and table-top laser-induced breakdown spectroscopy, Spectrochim. Acta Part B At. Spectrosc. 101 (2014) 191–199. https://doi.org/10.1016/j.sab.2014.08.036.

[106] S. Moncayo, S. Manzoor, F. Navarro-Villoslada, J.O. Caceres, Evaluation of supervised chemometric methods for sample classification by Laser Induced Breakdown Spectroscopy, Chemometr. Intell. Lab. Syst. 146 (2015) 354–364. https://doi.org/10.1016/j.chemolab.2015.06.004.

[107] R. Kanawade, F. Mehari, C. Knipfer, M. Rohde, K. Tangermann-Gerk, M. Schmidt, F. Stelzle, Pilot study of laser induced breakdown spectroscopy for tissue differentiation by monitoring the plume created during laser surgery — an approach on a feedback Laser control mechanism, Spectrochim. Acta Part B At. Spectrosc. 87 (2013) 175–181. https://doi.org/10.1016/j.sab.2013.05.012.

[108] F.C. De Lucia, J.L. Gottfried, Influence of variable selection on partial least squares discriminant analysis models for explosive residue classification, Spectrochim. Acta Part B At. Spectrosc. 66 (2011) 122–128. https://doi.org/10.1016/j.sab.2010.12.007.

[109] F.C. De Lucia, J.L. Gottfried, C.A. Munson, A.W. Miziolek, Multivariate analysis of standoff laser-induced breakdown spectroscopy spectra for classification of explosive-containing residues, Appl. Opt. 47 (2008) G112–G121. https://doi.org/10.1364/AO.47.00G112.

[110] A. Kumar Myakalwar, N. Spegazzini, C. Zhang, S. Kumar Anubham, R.R. Dasari, I. Barman, M. Kumar Gundawar, Less is more: avoiding the LIBS dimensionality curse through judicious feature selection for explosive detection, Sci. Rep. 5 (2015), 13169. https://doi.org/10.1038/srep13169.

[111] J.-B. Sirven, B. Salle, P. Mauchien, J.-L. Lacour, S. Maurice, G. Manhes, 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 (2007) 1471–1480. https://doi.org/10.1039/b704868h.

[112] S. Merk, C. Scholz, S. Florek, D. Mory, Increased identification rate of scrap metal using Laser Induced Breakdown Spectroscopy Echelle spectra, Spectrochim. Acta Part B At. Spectrosc. 112 (2015) 10–15. https://doi.org/10.1016/j.sab.2015.07.009.

[113] R.G. Brereton, G.R. Lloyd, Partial least squares discriminant analysis: taking the magic away, J. Chemom. 28 (2014) 213–225. https://doi.org/10.1002/cem.2609.

[114] A.L. Pomerantsev, O.Ye Rodionova, Multiclass partial least squares discriminant analysis: taking the right way—a critical tutorial, J. Chemom. 32 (2018), e3030. https://doi.org/10.1002/cem.3030.

[115] B.E. Boser, I.M. Guyon, V.N. Vapnik, A training algorithm for optimal margin classifiers, New York, NY, USA, in: Proc. Fifth Annu. Workshop Comput. Learn. Theory, Association for Computing Machinery, 1992, pp. 144–152. https://doi.org/10.1145/130385.130401.

[116] L. Sheng, T. Zhang, G. Niu, K. Wang, H. Tang, Y. Duan, H. Li, Classification of iron ores by laser-induced breakdown spectroscopy (LIBS) combined with random forest (RF), J. Anal. At. Spectrom. 30 (2015) 453–458. https://doi.org/10.1039/C4JA00352G.

[117] N.C. Dingari, I. Barman, A.K. Myakalwar, S.P. Tewari, M.K. Gundawar, Incorporation of support vector machines in the LIBS toolbox for sensitive and robust classification amidst unexpected sample and system variability, Anal. Chem. 84 (2012) 2686–2694. https://doi.org/10.1021/ac202755e.

[118] L. Liang, T. Zhang, K. Wang, H. Tang, X. Yang, X. Zhu, Y. Duan, H. Li, Classification of steel materials by laser-induced breakdown spectroscopy coupled with support vector machines, Appl. Opt. 53 (2014) 544–552. https://doi.org/10.1364/AO.53.000544.

[119] X. Zhu, T. Xu, Q. Lin, L. Liang, G. Niu, H. Lai, M. Xu, X. Wang, H. Li, Y. Duan, Advanced statistical analysis of laser-induced breakdown spectroscopy data to discriminate sedimentary rocks based on Czerny–Turner and Echelle spectrometers, Spectrochim. Acta Part B At. Spectrosc. 93 (2014) 8–13.

[120] https://doi.org/10.1016/j.sab.2014.01.001.

[120] E. Képeš, J. Vrábel, O. Adamovsky, S. Střítežská, P. Modlitbová, P. Pořízka, J. Kaiser, Interpreting support vector machines applied in laser-induced breakdown spectroscopy, Anal. Chim. Acta 1192 (2022), 339352. https://doi.org/10.1016/j.aca.2021.339352.

[121] E. Fix, J.L. Hodges, Discriminatory analysis. Nonparametric discrimination: consistency properties, Int. Stat. Rev. Rev. Int. Stat. 57 (1989) 238–247. https://doi.org/10.2307/1403797.

[122] V.C. Costa, F.W. Batista Aquino, C.M. Paranhos, E.R. Pereira-Filho, Identification and classification of polymer e-waste using laser-induced breakdown spectroscopy (LISS) and chemometric tools, Polym. Test. 59 (2017) 390–395. https://doi.org/10.1016/j.polymertesting.2017.02.017.

[123] Q. Godoi, F.O. Leme, L.C. Trevizan, E.R. Pereira Filho, I.A. Rufini, D. Santos, F.J. Krug, Laser-induced breakdown spectroscopy and chemometrics for classification of toys relying on toxic elements, Spectrochim. Acta Part B-At. Spectrosc. 66 (2011) 138–143. https://doi.org/10.1016/j.sab.2011.01.001.

[124] X. Li, S. Yang, R. Fan, X. Yu, D. Chen, Discrimination of soft tissues using laser-induced breakdown spectroscopy in combination with k nearest neighbors (kNN) and support vector machine (SVM) classifiers, Opt Laser. Technol. 102 (2018) 233–239. https://doi.org/10.1016/j.optlastec.2018.01.028.

[125] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32. https://doi.org/10.1023/A:1010933404324.

[126] G. Biau, E. Scornet, A random forest guided tour, TEST 25 (2016) 197–227. https://doi.org/10.1007/s11749-016-0481-7.

[127] J. Qi, T. Zhang, H. Tang, H. Li, Rapid classification of archaeological ceramics via laser-induced breakdown spectroscopy coupled with random forest, Spectrochim. Acta Part B At. Spectrosc. 149 (2018) 288–293. https://doi.org/10.1016/j.sab.2018.09.006.

[128] H. Tang, T. Zhang, X. Yang, H. Li, Classification of different types of slag samples by laser-induced breakdown spectroscopy (LIBS) coupled with random forest based on variable importance (VIRF), Anal. Methods 7 (2015) 9171–9176. https://doi.org/10.1039/C5AY02208H.

[129] T. Zhang, C. Yan, J. Qi, H. Tang, H. Li, Classification and discrimination of coal ash by laser-induced breakdown spectroscopy (LIBS) coupled with advanced chemometric methods, J. Anal. At. Spectrom. 32 (2017) 1960–1965. https://doi.org/10.1039/C7JA00218A.

[130] Y. Tian, C. Yan, T. Zhang, H. Tang, H. Li, J. Yu, J. Bernard, L. Chen, S. Martin, N. Delepine-Gilon, J. Bocková, P. Veis, Y. Chen, J. Yu, Classification of wines according to their production regions with the contained trace elements using laser-induced breakdown spectroscopy, Spectrochim. Acta Part B At. Spectrosc. 135 (2017) 91–101. https://doi.org/10.1016/j.sab.2017.07.003.

[131] L.-N. Li, X.-F. Liu, F. Yang, W.-M. Xu, J.-Y. Wang, R. Shu, A review of artificial neural network based chemometrics applied in laser-induced breakdown spectroscopy analysis, Spectrochim. Acta Part B At. Spectrosc. 180 (2021), 106183. https://doi.org/10.1016/j.sab.2021.106183.

[132] J. Vrábel, E. Képeš, P. Pořízka, J. Kaiser, Artificial neural networks for classification, in: Chemom. Numer. Methods LIBS, John Wiley & Sons, Ltd, 2022, pp. 213–240. https://doi.org/10.1002/9781119759614.ch9.

[133] S. Wold, Pattern recognition by means of disjoint principal components models, Pattern Recogn. 8 (1976) 127–139. https://doi.org/10.1016/0031-3203(76)90014-5.

[134] S. Wold, M. Sjöström, SIMCA: a method for analyzing chemical data in terms of similarity and analogy, in: Chemom. Theory Appl., American Chemical Society, 1977, pp. 243–282. https://doi.org/10.1021/bk-1977-0052.ch012.

[135] R. Vitale, F. Marini, C. Ruckebusch, SIMCA modeling for overlapping classes: fixed or optimized decision threshold? Anal. Chem. 90 (2018) 10738–10747. https://doi.org/10.1021/acs.analchem.8b01270.

[136] A.L. Pomerantsev, O.Y. Rodionova, Popular decision rules in SIMCA: critical review, J. Chemom. 34 (2020), e3250. https://doi.org/10.1002/cem.3250.

[137] H.H. Yue, S.J. Qin, Reconstruction-based fault identification using a combined index, Ind. Eng. Chem. Res. 40 (2001) 4403–4414. https://doi.org/10.1021/ie000141+.

[138] A.K. Myakalwar, S. Sreedhar, I. Barman, N.C. Dingari, S. Venugopal Rao, P. Prem Kiran, S.P. Tewari, G. Manoj Kumar, Laser-induced breakdown spectroscopy-based investigation and classification of pharmaceutical tablets using multivariate chemometric analysis, Talanta 87 (2011) 53–59. https://doi.org/10.1016/j.talanta.2011.09.040.

[139] F. Colao, R. Fantoni, P. Ortiz, M.A. Vazquez, J.M. Martin, R. Ortiz, N. Idris, Quarry identification of historical building materials by means of laser induced breakdown spectroscopy, X-ray fluorescence and chemometric analysis, Spectrochim. Acta Part B-At. Spectrosc. 65 (2010) 688–694. https://doi.org/10.1016/j.sab.2010.05.005.

[140] C. Eum, D. Jang, S. Lee, K. Cha, H. Chung, Alternative selection of Raman or LIBS spectral information in hierarchical discrimination of raw sapphires according to geographical origin for accuracy improvement, Talanta 221 (2021), 121555. https://doi.org/10.1016/j.talanta.2020.121555.

[141] Y. Yang, X. Hao, L. Zhang, L. Ren, Application of Scikit and Keras libraries for the classification of iron ore data acquired by laser-induced breakdown spectroscopy (LIBS), Sensors 20 (2020) 1393. https://doi.org/10.3390/s20051393.

[142] Y. Bi, Y. Zhang, J. Yan, Z. Wu, Y. Li, Classification and discrimination of minerals using laser induced breakdown spectroscopy and Raman spectroscopy, Plasma Sci. Technol. 17 (2015) 923–927. https://doi.org/10.1088/1009-0630/17/11/06.

[143] R.S. Harmon, C.S. Throckmorton, R.R. Hark, J.L. Gottfried, G. Wörner,

K. Harpp, L. Collins, Discriminating volcanic centers with handheld laser-induced breakdown spectroscopy (LIBS), J. Archaeol. Sci. 98 (2018) 112–127. https://doi.org/10.1016/j.jas.2018.07.009.

[144] G. Saverio Senesi, P. Manzari, A. Consiglio, O.D. Pascale, Identification and classification of meteorites using a handheld LIBS instrument coupled with a fuzzy logic-based method, J. Anal. At. Spectrom. 33 (2018) 1664–1675. https://doi.org/10.1039/C8JA00224J.

[145] W. Zhang, Z. Zhuo, P. Lu, J. Tang, H. Tang, J. Lu, T. Xing, Y. Wang, LIBS analysis of the ash content, volatile matter, and calorific value in coal by partial least squares regression based on ash classification, J. Anal. At. Spectrom. 35 (2020) 1621–1631. https://doi.org/10.1039/D0JA00186D.

[146] J. Yu, Z. Hou, S. Sheta, J. Dong, W. Han, T. Lu, Z. Wang, Provenance classification of nephrite jades using multivariate LIBS: a comparative study, Anal. Methods 10 (2018) 281–289. https://doi.org/10.1039/C7AY02643A.

[147] P. Janovszky, K. Jancsek, D.J. Palásti, J. Kopniczky, B. Hopp, T.M. Tóth, G. Galbács, Classification of minerals and the assessment of lithium and beryllium content in granitoid rocks by laser-induced breakdown spectroscopy, J. Anal. At. Spectrom. (2021). https://doi.org/10.1039/D1JA00032B.

[148] C. Wang, J. Wang, J. Wang, H. Du, J. Wang, Classification of 13 original rock samples by laser induced breakdown spectroscopy, Laser Phys. 31 (2021), 035601. https://doi.org/10.1088/1555-6611/abdfc8.

[149] W.T. Li, Y.N. Zhu, X. Li, Z.Q. Hao, L.B. Guo, X.Y. Li, X.Y. Zeng, Y.F. Lu, In situ classification of rocks using stand-off laser-induced breakdown spectroscopy with a compact spectrometer, J. Anal. At. Spectrom. 33 (2018) 461–467. https://doi.org/10.1039/C8JA00001H.

[150] H. Peng, G. Chen, X. Chen, Z. Lu, S. Yao, Hybrid classification of coal and biomass by laser-induced breakdown spectroscopy combined with K-means and SVM, Plasma Sci. Technol. 21 (2018), 034008. https://doi.org/10.1088/2058-6272/aaebc4.

[151] R.H. El-Saeid, Z. Abdel-Salam, S. Pagnotta, V. Palleschi, M.A. Harith, Classification of sedimentary and igneous rocks by laser induced breakdown spectroscopy and nanoparticle-enhanced laser induced breakdown spectroscopy combined with principal component analysis and graph theory, Spectrochim. Acta Part B At. Spectrosc. 158 (2019) 105622. https://doi.org/10.1016/j.sab.2019.05.011.

[152] G. Yang, S. Qiao, P. Chen, Y. Ding, D. Tian, Rock and soil classification using PLS-DA and SVM combined with a laser-induced breakdown spectroscopy library, Plasma Sci. Technol. 17 (2015) 656–663. https://doi.org/10.1088/1009-0630/17/8/08.

[153] M. Yelameli, B. Thornton, T. Takahashi, T. Weerakoon, K. Ishii, Classification and statistical analysis of hydrothermal seafloor rocks measured underwater using laser-induced breakdown spectroscopy, J. Chemom. 33 (2019), e3092. https://doi.org/10.1002/cem.3092.

[154] S. Chen, H. Pei, J. Pisonero, S. Yang, Q. Fan, X. Wang, Y. Duan, Simultaneous determination of lithology and major elements in rocks using laser-induced breakdown spectroscopy (LIBS) coupled with a deep convolutional neural network, J. Anal. At. Spectrom. 37 (2022) 508–516. https://doi.org/10.1039/D1JA00406A.

[155] S. Müller, J.A. Meima, Mineral classification of lithium-bearing pegmatites based on laser-induced breakdown spectroscopy: application of semi-supervised learning to detect known minerals and unknown material, Spectrochim. Acta Part B At. Spectrosc. 189 (2022), 106370. https://doi.org/10.1016/j.sab.2022.106370.

[156] A. Ramil, A.J. López, A. Yáñez, Application of artificial neural networks for the rapid classification of archaeological ceramics by means of laser induced breakdown spectroscopy (LIBS), Appl. Phys. A 92 (2008) 197–202. https://doi.org/10.1007/s00339-008-4481-7.

[157] T. Zhang, D. Xia, H. Tang, X. Yang, H. Li, Classification of steel samples by laser-induced breakdown spectroscopy and random forest, Chemometr. Intell. Lab. Syst. 157 (2016) 196–201. https://doi.org/10.1016/j.chemolab.2016.07.001.

[158] A. dos Santos Augusto, É. Ferreira Batista, E. Rodrigues Pereira-Filho, Direct chemical inspection of eye shadow and lipstick solid samples using laser-induced breakdown spectroscopy (LIBS) and chemometrics: proposition of classification models, Anal. Methods 8 (2016) 5851–5860. https://doi.org/10.1039/C6AY01138A.

[159] X. Cui, Q. Wang, Y. Zhao, X. Qiao, G. Teng, Laser-induced breakdown spectroscopy (LIBS) for classification of wood species integrated with artificial neural network (ANN), Appl. Phys. B 125 (2019) 56. https://doi.org/10.1007/s00340-019-7166-3.

[160] A.M. Neiva, M.A.C. Jacinto, M.M. de Alencar, S.N. Esteves, E.R. Pereira-Filho, Proposition of classification models for the direct evaluation of the quality of cattle and sheep leathers using laser-induced breakdown spectroscopy (LIBS) analysis, RSC Adv. 6 (2016), 104827. https://doi.org/10.1039/C6RA22337K. –104838.

[161] T. Zhang, S. Wu, J. Dong, J. Wei, K. Wang, H. Tang, X. Yang, H. Li, Quantitative and classification analysis of slag samples by laser induced breakdown spectroscopy (LIBS) coupled with support vector machine (SVM) and partial least square (PLS) methods, J. Anal. At. Spectrom. 30 (2015) 368–374. https://doi.org/10.1039/C4JA00421C.

[162] H. Xia, M.C.M. Bakker, Reliable classification of moving waste materials with LIBS in concrete recycling, Talanta 120 (2014) 239–247. https://doi.org/10.1016/j.talanta.2013.11.082.

[163] E. Kim, Y. Kim, E. Srivastava, S. Shin, S. Jeong, E. Hwang, Soft classification scheme with pre-cluster-based regression for identification of same-base alloys using laser-induced breakdown spectroscopy, Chemometr. Intell. Lab. Syst. 203 (2020), 104072. https://doi.org/10.1016/j.chemolab.2020.104072.

[164] S. Awasthi, R. Kumar, G.K. Rai, A.K. Rai, Study of archaeological coins of different dynasties using libs coupled with multivariate analysis, Opt Laser. Eng. 79 (2016) 29–38. https://doi.org/10.1016/j.optlaseng.2015.11.005.

[165] H. Kim, J. Lee, E. Srivastava, S. Shin, S. Jeong, E. Hwang, Front-end signal processing for metal scrap classification using online measurements based on laser-induced breakdown spectroscopy, Spectrochim. Acta Part B At. Spectrosc. 184 (2021), 106282. https://doi.org/10.1016/j.sab.2021.106282.

[166] L. Zhan, X. Ma, W. Fang, R. Wang, Z. Liu, Y. Song, H. Zhao, A rapid classification method of aluminum alloy based on laser-induced breakdown spectroscopy and random forest algorithm, Plasma Sci. Technol. 21 (2019), 034018. https://doi.org/10.1088/2058-6272/aaf7bf.

[167] B. Campanella, E. Grifoni, S. Legnaioli, G. Lorenzetti, S. Pagnotta, F. Sorrentino, V. Palleschi, Classification of wrought aluminum alloys by Artificial Neural Networks evaluation of Laser Induced Breakdown Spectroscopy spectra from aluminum scrap samples, Spectrochim. Acta Part B At. Spectrosc. 134 (2017) 52–57. https://doi.org/10.1016/j.sab.2017.06.003.

[168] H. Kong, L. Sun, J. Hu, Y. Xin, Z. Cong, Selection of spectral data for classification of steels using laser-induced breakdown spectroscopy, Plasma Sci. Technol. 17 (2015) 964–970. https://doi.org/10.1088/1009-0630/17/11/14.

[169] A.J. López, G. Nicolás, M.P. Mateo, A. Ramil, V. Piñón, A. Yáñez, LIPS and linear correlation analysis applied to the classification of Roman pottery Terra Sigillata, Appl. Phys. A 83 (2006) 695–698. https://doi.org/10.1007/s00339-006-3556-6.

[170] P. Werheit, C. Fricke-Begemann, M. Gesing, R. Noll, Fast single piece identification with a 3D scanning LIBS for aluminium cast and wrought alloys recycling, J. Anal. At. Spectrom. 26 (2011) 2166–2174. https://doi.org/10.1039/C1JA10096C.

[171] B.G. Oztoprak, J. Gonzalez, J. Yoo, T. Gulecen, N. Mutlu, R.E. Russo, O. Gundogdu, A. Demir, Analysis and classification of heterogeneous kidney stones using laser-induced breakdown spectroscopy (LIBS), Appl. Spectrosc. 66 (2012) 1353–1361. https://doi.org/10.1366/12-06679.

[172] J. Liang, M. Li, Y. Du, C. Yan, Y. Zhang, T. Zhang, X. Zheng, H. Li, Data fusion of laser induced breakdown spectroscopy (LIBS) and infrared spectroscopy (IR) coupled with random forest (RF) for the classification and discrimination of compound salvia miltiorrhiza, Chemometr. Intell. Lab. Syst. 207 (2020), 104179. https://doi.org/10.1016/j.chemolab.2020.104179.

[173] J. Liang, C. Yan, Y. Zhang, T. Zhang, X. Zheng, H. Li, Rapid discrimination of Salvia miltiorrhiza according to their geographical regions by laser induced breakdown spectroscopy (LIBS) and particle swarm optimization-kernel extreme learning machine (PSO-KELM), Chemometr. Intell. Lab. Syst. 197 (2020), 103930. https://doi.org/10.1016/j.chemolab.2020.103930.

[174] X. Zhang, N. Li, C. Yan, J. Zeng, T. Zhang, H. Li, Four-metal-element quantitative analysis and pollution source discrimination in atmospheric sedimentation by laser-induced breakdown spectroscopy (LIBS) coupled with machine learning, J. Anal. At. Spectrom. 35 (2020) 403–413. https://doi.org/10.1039/C9JA00360F.

[175] S. Zhao, W. Song, Z. Hou, Z. Wang, Classification of ginseng according to plant species, geographical origin, and age using laser-induced breakdown spectroscopy and hyperspectral imaging, J. Anal. At. Spectrom. 36 (2021) 1704–1711. https://doi.org/10.1039/D1JA00136A.

[176] F.-Y. Yueh, H. Zheng, J.P. Singh, S. Burgess, Preliminary evaluation of laser-induced breakdown spectroscopy for tissue classification, Spectrochim. Acta Part B At. Spectrosc. 64 (2009) 1059–1067. https://doi.org/10.1016/j.sab.2009.07.025.

[177] R. Gaudiuso, E. Ewusi-Annan, W. Xia, N. Melikechi, Diagnosis of Alzheimer's disease using laser-induced breakdown spectroscopy and machine learning, Spectrochim. Acta Part B At. Spectrosc. 171 (2020), 105931. https://doi.org/10.1016/j.sab.2020.105931.

[178] Y.G. Mbesse Kongbonga, H. Ghalila, M.B. Onana, Z. Ben Lakhdar, Classification of vegetable oils based on their concentration of saturated fatty acids using laser induced breakdown spectroscopy (LIBS), Food Chem. 147 (2014) 327–331. https://doi.org/10.1016/j.foodchem.2013.09.145.

[179] B. Sezer, S. Durna, G. Bilge, A. Berkkan, A. Yetisemiyen, I.H. Boyaci, Identification of milk fraud using laser-induced breakdown spectroscopy (LIBS), Int. Dairy J. 81 (2018) 1–7. https://doi.org/10.1016/j.idairyj.2017.12.005.

[180] H.M. Velioglu, B. Sezer, G. Bilge, S.E. Baytur, I.H. Boyaci, Identification of offal adulteration in beef by laser induced breakdown spectroscopy (LIBS), Meat Sci. 138 (2018) 28–33. https://doi.org/10.1016/j.meatsci.2017.12.003.

[181] O. Gazeli, E. Bellou, D. Stefas, S. Couris, Laser-based classification of olive oils assisted by machine learning, Food Chem. 302 (2020), 125329. https://doi.org/10.1016/j.foodchem.2019.125329.

[182] Y. Lee, S.-H. Nam, K.-S. Ham, J. Gonzalez, D. Oropeza, D. Quarles, J. Yoo, R.E. Russo, Multivariate classification of edible salts: simultaneous laser-induced breakdown spectroscopy and laser-ablation inductively coupled plasma mass spectrometry analysis, Spectrochim. Acta Part B At. Spectrosc. 118 (2016) 102–111. https://doi.org/10.1016/j.sab.2016.02.019.

[183] N. Baskali-Bouregaa, M.-L. Milliand, S. Mauffrey, E. Chabert, M. Forrestier, N. Gilon, Tea geographical origin explained by LIBS elemental profile combined to isotopic information, Talanta 211 (2020), 120674. https://doi.org/10.1016/j.talanta.2019.120674.

[184] M. Yao, G. Fu, T. Chen, M. Liu, J. Xu, H. Zhou, X. He, L. Huang, A modified genetic algorithm optimized SVM for rapid classification of tea leaves using

laser-induced breakdown spectroscopy, J. Anal. At. Spectrom. 36 (2021) 361–367. https://doi.org/10.1039/D0JA00317D.

[185] S. Moncayo, S. Manzoor, J.D. Rosales, J. Anzano, J.O. Caceres, Qualitative and quantitative analysis of milk for the detection of adulteration by Laser Induced Breakdown Spectroscopy (LIBS), Food Chem. 232 (2017) 322–328. https://doi.org/10.1016/j.foodchem.2017.04.017.

[186] D. Stefas, N. Gyftokostas, S. Couris, Laser induced breakdown spectroscopy for elemental analysis and discrimination of honey samples, Spectrochim. Acta Part B At. Spectrosc. 172 (2020), 105969. https://doi.org/10.1016/j.sab.2020.105969.

[187] E. Bellou, N. Gyftokostas, D. Stefas, O. Gazeli, S. Couris, Laser-induced breakdown spectroscopy assisted by machine learning for olive oils classification: the effect of the experimental parameters, Spectrochim. Acta Part B At. Spectrosc. 163 (2020), 105746. https://doi.org/10.1016/j.sab.2019.105746.

[188] M.M. Tan, S. Cui, J. Yoo, S.-H. Han, K.-S. Ham, S.-H. Nam, Y. Lee, Feasibility of laser-induced breakdown spectroscopy (LIBS) for classification of sea salts, Appl. Spectrosc. 66 (2012) 262–271. https://doi.org/10.1366/11-06379.

[189] Z. Zhao, L. Chen, F. Liu, F. Zhou, J. Peng, M. Sun, Fast classification of geographical origins of honey based on laser-induced breakdown spectroscopy and multivariate analysis, Sensors 20 (2020) 1878. https://doi.org/10.3390/s20071878.

[190] N. Gyftokostas, D. Stefas, S. Couris, Olive oils classification via laser-induced breakdown spectroscopy, Appl. Sci. 10 (2020) 3462. https://doi.org/10.3390/app10103462.

[191] P. Yang, R. Zhou, W. Zhang, S. Tang, Z. Hao, X. Li, Y. Lu, X. Zeng, Laser-induced breakdown spectroscopy assisted chemometric methods for rice geographic origin classification, Appl. Opt. 57 (2018) 8297–8302. https://doi.org/10.1364/AO.57.008297.

[192] H. Sun, C. Song, X. Lin, X. Gao, Identification of meat species by combined laser-induced breakdown and Raman spectroscopies, Spectrochim. Acta Part B At. Spectrosc. 194 (2022), 106456. https://doi.org/10.1016/j.sab.2022.106456.

[193] B. Sezer, A. Unuvar, I.H. Boyaci, H. Köksel, Rapid discrimination of authenticity in wheat flour and pasta samples using LIBS, J. Cereal. Sci. 104 (2022) 103435. https://doi.org/10.1016/j.jcs.2022.103435.

[194] W. Huang, L. Guo, W. Kou, D. Zhang, Z. Hu, F. Chen, Y. Chu, W. Cheng, Identification of adulterated milk powder based on convolutional neural network and laser-induced breakdown spectroscopy, Microchem. J. 176 (2022), 107190. https://doi.org/10.1016/j.microc.2022.107190.

[195] F.C. De Lucia, J.L. Gottfried, C.A. Munson, A.W. Miziolek, Double pulse laser-induced breakdown spectroscopy of explosives: initial study towards improved discrimination, Spectrochim. Acta Part B At. Spectrosc. 62 (2007) 1399–1404. https://doi.org/10.1016/j.sab.2007.10.036.

[196] Q. Wang, G. Teng, C. Li, Y. Zhao, Z. Peng, Identification and classification of explosives using semi-supervised learning and laser-induced breakdown spectroscopy, J. Hazard Mater. 369 (2019) 423–429. https://doi.org/10.1016/j.jhazmat.2019.02.015.

[197] R.S. Harmon, F.C. DeLucia, A. LaPointe, R.J. Winkel, A.W. Miziolek, LIBS for landmine detection and discrimination, Anal. Bioanal. Chem. 385 (2006) 1140–1148. https://doi.org/10.1007/s00216-006-0513-3.

[198] K. Rzecki, T. Sośnicki, M. Baran, M. Niedźwiecki, M. Król, T. Łojewski, U.R. Acharya, Ö. Yildirim, P. Pławiak, Application of computational intelligence methods for the automated identification of paper-Ink samples based on LIBS, Sensors 18 (2018) 3670. https://doi.org/10.3390/s18113670.

[199] M. Hoehse, A. Paul, I. Gornushkin, U. Panne, Multivariate classification of pigments and inks using combined Raman spectroscopy and LIBS, Anal. Bioanal. Chem. 402 (2012) 1443–1450. https://doi.org/10.1007/s00216-011-5287-6.

[200] K. Menking-Hoggatt, L. Arroyo, J. Curran, T. Trejos, Novel LIBS method for micro-spatial chemical analysis of inorganic gunshot residues, J. Chemom. 35 (2021), e3208. https://doi.org/10.1002/cem.3208.

[201] V. Merk, D. Huber, L. Pfeifer, S. Damaske, S. Merk, W. Werncke, M. Schuster, Discrimination of automotive glass by conjoint Raman and laser-induced breakdown spectroscopy and multivariate data analysis, Spectrochim. Acta Part B At. Spectrosc. 180 (2021), 106198. https://doi.org/10.1016/j.sab.2021.106198.

[202] V. Lazic, A. Palucci, S. Jovicevic, M. Carpanese, Detection of explosives in traces by laser induced breakdown spectroscopy: differences from organic interferants and conditions for a correct classification, Spectrochim. Acta Part B At. Spectrosc. 66 (2011) 644–655. https://doi.org/10.1016/j.sab.2011.07.003.

[203] L. Pagnin, L. Brunnbauer, R. Wiesinger, A. Limbeck, M. Schreiner, Multivariate analysis and laser-induced breakdown spectroscopy (LIBS): a new approach for the spatially resolved classification of modern art materials, Anal. Bioanal. Chem. 412 (2020) 3187–3198. https://doi.org/10.1007/s00216-020-02574-z.

[204] L. Brunnbauer, S. Larisegger, H. Lohninger, M. Nelhiebel, A. Limbeck, Spatially resolved polymer classification using laser induced breakdown spectroscopy (LIBS) and multivariate statistics, Talanta 209 (2020), 120572. https://doi.org/10.1016/j.talanta.2019.120572.

[205] K.M. Shameem, K.S. Choudhari, A. Bankapur, S.D. Kulkarni, V.K. Unnikrishnan, S.D. George, V.B. Kartha, C. Santhosh, A hybrid LIBS–Raman system combined with chemometrics: an efficient tool for plastic identification and sorting, Anal. Bioanal. Chem. 409 (2017)

[206] Y. Yu, L.B. Guo, Z.Q. Hao, X.Y. Li, M. Shen, Q.D. Zeng, K.H. Li, X.Y. Zeng, Y.F. Lu, Z. Ren, Accuracy improvement on polymer identification using laser-induced breakdown spectroscopy with adjusting spectral weightings, Opt Express 22 (2014) 3895–3901. https://doi.org/10.1364/OE.22.003895.

[207] V.K. Unnikrishnan, K.S. Choudhari, S.D. Kulkarni, R. Nayak, V.B. Kartha, C. Santhosh, Analytical predictive capabilities of laser induced breakdown spectroscopy (LIBS) with principal component analysis (PCA) for plastic classification, RSC Adv. 3 (2013) 25872–25880. https://doi.org/10.1039/C3RA44946G.

[208] M. Banaee, S.H. Tavassoli, Discrimination of polymers by laser induced breakdown spectroscopy together with the DFA method, Polym. Test. 31 (2012) 759–764. https://doi.org/10.1016/j.polymertesting.2012.04.010.

[209] R. Junjuri, M.K. Gundawar, Femtosecond laser-induced breakdown spectroscopy studies for the identification of plastics, J. Anal. At. Spectrom. (2019). https://doi.org/10.1039/C9JA00102F.

[210] M. Vahid Dastjerdi, S.J. Mousavi, M. Soltanolkotabi, A. Nezarati Zadeh, Identification and sorting of PVC polymer in recycling process by laser-induced breakdown spectroscopy (LIBS) combined with support vector machine (SVM) model, Iran, J. Sci. Technol. Trans. Sci. 42 (2018) 959–965. https://doi.org/10.1007/s40995-016-0084-x.

[211] M. Boueri, V. Motto-Ros, W.-Q. Lei, Qain-LiMa, L.-J. Zheng, H.-P. Zeng, JinYu, Identification of polymer materials using laser-induced breakdown spectroscopy combined with artificial neural networks, Appl. Spectrosc. 65 (2011) 307–314.

[212] R.J. Lasheras, C. Bello-Gálvez, J. Anzano, Identification of polymers by libs using methods of correlation and normalized coordinates, Polym. Test. 29 (2010) 1057–1064. https://doi.org/10.1016/j.polymertesting.2010.07.011.

[213] Y. Tang, Y. Guo, Q. Sun, S. Tang, J. Li, L. Guo, J. Duan, Industrial polymers classification using laser-induced breakdown spectroscopy combined with self-organizing maps and K-means algorithm, Optik 165 (2018) 179–185. https://doi.org/10.1016/j.ijleo.2018.03.121.

[214] S. Grégoire, M. Boudinet, F. Pelascini, F. Surma, V. Detalle, Y. Holl, Laser-induced breakdown spectroscopy for polymer identification, Anal. Bioanal. Chem. 400 (2011) 3331–3340. https://doi.org/10.1007/s00216-011-4898-2.

[215] K. Liu, D. Tian, X. Deng, H. Wang, G. Yang, Rapid classification of plastic bottles by laser-induced breakdown spectroscopy (LIBS) coupled with partial least squares discrimination analysis based on spectral windows (SW-PLS-DA), J. Anal. At. Spectrom. (2019). https://doi.org/10.1039/C9JA00105K.

[216] J. Jasik, J. Heitz, J.D. Pedarnig, P. Veis, Vacuum ultraviolet laser-induced breakdown spectroscopy analysis of polymers, Spectrochim. Acta Part B At. Spectrosc. 64 (2009) 1128–1134. https://doi.org/10.1016/j.sab.2009.07.013.

[217] S. Jayaganthan, M.S. Babu, N.J. Vasa, R. Sarathi, T. Imai, Classification of coal deposited epoxy micro-nanocomposites by adopting machine learning techniques to LIBS analysis, J. Phys. Commun. 5 (2021), 105006. https://doi.org/10.1088/2399-6528/ac2b5d.

[218] X. Yan, X. Peng, Y. Qin, Z. Xu, B. Xu, C. Li, N. Zhao, J. Li, Q. Ma, Q. Zhang, Classification of plastics using laser-induced breakdown spectroscopy combined with principal component analysis and K nearest neighbor algorithm, Results Opt 4 (2021), 100093. https://doi.org/10.1016/j.rio.2021.100093.

[219] R. Junjuri, C. Zhang, I. Barman, M.K. Gundawar, Identification of post-consumer plastics using laser-induced breakdown spectroscopy, Polym. Test. 76 (2019) 101–108. https://doi.org/10.1016/j.polymertesting.2019.03.012.

[220] C.E. McManus, N.J. McMillan, R.S. Harmon, R.C. Whitmore, J. Frank C. De Lucia, A.W. Miziolek, Use of laser induced breakdown spectroscopy in the determination of gem provenance: beryls, Appl. Opt. 47 (2008) G72–G79. https://doi.org/10.1364/AO.47.000G72.

[221] N.J. McMillan, C.E. McManus, R.S. Harmon, F.C. De Lucia, A.W. Miziolek, Laser-induced breakdown spectroscopy analysis of complex silicate minerals—beryl, Anal. Bioanal. Chem. 385 (2006) 263–271. https://doi.org/10.1007/s00216-006-0374-9.

[222] N.J. McMillan, R.S. Harmon, F.C. De Lucia, A.M. Miziolek, Laser-induced breakdown spectroscopy analysis of minerals: carbonates and silicates, Spectrochim. Acta Part B At. Spectrosc. 62 (2007) 1528–1536. https://doi.org/10.1016/j.sab.2007.10.037.

[223] R.S. Harmon, F.C. DeLucia, C.E. McManus, N.J. McMillan, T.F. Jenkins, M.E. Walsh, A. Miziolek, Laser-induced breakdown spectroscopy – an emerging chemical sensor technology for real-time field-portable, geochemical, mineralogical, and environmental applications, Appl. Geochem. 21 (2006) 730–747. https://doi.org/10.1016/j.apgeochem.2006.02.003.

[224] R.S. Harmon, J. Remus, N.J. McMillan, C. McManus, L. Collins, J.L. Gottfried, F.C. DeLucia, A.W. Miziolek, LIBS analysis of geomaterials: geochemical fingerprinting for the rapid analysis and discrimination of minerals, Appl. Geochem. 24 (2009) 1125–1141. https://doi.org/10.1016/j.apgeochem.2009.02.009.

[225] R.R. Hark, R.S. Harmon, Geochemical fingerprinting using LIBS, in: S. Musazzi, U. Perini (Editors), Laser-Induc. Breakdown Spectrosc. Theory Appl., Springer, Berlin, Heidelberg, 2014, pp. 309–348. https://doi.org/10.1007/978-3-642-45085-3_12.

[226] J.L. Gottfried, R.S. Harmon, F.C. De Lucia, A.W. Miziolek, Multivariate analysis of laser-induced breakdown spectroscopy chemical signatures for geo-material classification, Spectrochim. Acta Part B At. Spectrosc. 64 (2009) 1009–1019. https://doi.org/10.1016/j.sab.2009.07.005.

[227] S. Moncayo, L. Duponchel, N. Mousavipak, G. Panczer, F. Trichard,

3299–3308.

B. Bousquet, F. Pelascini, V. Motto-Ros, Exploration of megapixel hyperspectral LIBS images using principal component analysis, J. Anal. At. Spectrom. 33 (2018) 210–220. https://doi.org/10.1039/C7JA00398F.

[228] E. Képeš, J. Vrábel, S. Střítežská, P. Pořízka, J. Kaiser, Benchmark classification dataset for laser-induced breakdown spectroscopy, Sci. Data 7 (2020) 53. https://doi.org/10.1038/s41597-020-0396-8.

[229] S.J. Rehse, A review of the use of laser-induced breakdown spectroscopy for bacterial classification, quantification, and identification, Spectrochim. Acta Part B At. Spectrosc. 154 (2019) 50–69. https://doi.org/10.1016/j.sab.2019.02.005.

[230] R. Gaudiuso, N. Melikechi, Z.A. Abdel-Salam, M.A. Harith, V. Palleschi, V. Motto-Ros, B. Busser, Laser-induced breakdown spectroscopy for human and animal health: a review, Spectrochim. Acta Part B At. Spectrosc. 152 (2019) 123–148. https://doi.org/10.1016/j.sab.2018.11.006.

[231] X. Lin, H. Sun, X. Gao, Y. Xu, Z. Wang, Y. Wang, Discrimination of lung tumor and boundary tissues based on laser-induced breakdown spectroscopy and machine learning, Spectrochim. Acta Part B At. Spectrosc. 180 (2021), 106200. https://doi.org/10.1016/j.sab.2021.106200.

[232] J.-H. Choi, S. Shin, Y. Moon, J.H. Han, E. Hwang, S. Jeong, High spatial resolution imaging of melanoma tissue by femtosecond laser-induced breakdown spectroscopy, Spectrochim. Acta Part B At. Spectrosc. 179 (2021), 106090. https://doi.org/10.1016/j.sab.2021.106090.

[233] P. Siozos, N. Hausmann, M. Holst, D. Anglos, Application of laser-induced breakdown spectroscopy and neural networks on archaeological human bones for the discrimination of distinct individuals, J. Archaeol. Sci. Rep. 35 (2021), 102769. https://doi.org/10.1016/j.jasrep.2020.102769.

[234] C.A. Georgiou, G.P. Danezis, Food Authentication: Management, Analysis and Regulation, John Wiley & Sons, 2017.

[235] P. Yang, Y. Zhu, X. Yang, J. Li, S. Tang, Z. Hao, L. Guo, X. Li, X. Zeng, Y. Lu, Evaluation of sample preparation methods for rice geographic origin classification using laser-induced breakdown spectroscopy, J. Cereal. Sci. 80 (2018) 111–118. https://doi.org/10.1016/j.jcs.2018.01.007.

[236] S. Moncayo, J.D. Rosales, R. Izquierdo-Hornillos, J. Anzano, J.O. Caceres, Classification of red wine based on its protected designation of origin (PDO) using Laser-induced Breakdown Spectroscopy (LIBS), Talanta 158 (2016) 185–191. https://doi.org/10.1016/j.talanta.2016.05.059.

[237] B. Sezer, A. Bjelak, H. Murat Velioglu, I. Hakkı Boyaci, Identification of meat species in processed meat products by using protein based laser induced breakdown spectroscopy assay, Food Chem. 372 (2022), 131245. https://doi.org/10.1016/j.foodchem.2021.131245.

[238] J.L. Gottfried, F.C. De Lucia, C.A. Munson, A.W. Miziolek, Laser-induced breakdown spectroscopy for detection of explosives residues: a review of recent advances, challenges, and future prospects, Anal. Bioanal. Chem. 395 (2009) 283–300. https://doi.org/10.1007/s00216-009-2802-0.

[239] F.C. DeLucia, A.C. Samuels, R.S. Harmon, R.A. Walters, K.L. McNesby, A. LaPointe, R.J. Winkel, A.W. Miziolek, Laser-induced breakdown spectroscopy (LIBS): a promising versatile chemical sensor technology for hazardous material detection, IEEE Sensor. J. 5 (2005) 681–689. https://doi.org/10.1109/JSEN.2005.848151.

[240] F.C.D. Lucia, J.L. Gottfried, Classification of explosive residues on organic substrates using laser induced breakdown spectroscopy, Appl. Opt. 51 (2012). https://doi.org/10.1364/AO.51.000B83. B83–B92.

[241] M.M. El-Deftar, N. Speers, S. Eggins, S. Foster, J. Robertson, C. Lennard, Assessment and forensic application of laser-induced breakdown spectroscopy (LIBS) for the discrimination of Australian window glass, Forensic Sci. Int. 241 (2014) 46–54. https://doi.org/10.1016/j.forsciint.2014.04.040.

[242] J.-H. Yang, J.J. Yoh, Forensic discrimination of latent fingerprints using laser-induced breakdown spectroscopy (LIBS) and chemometric approaches, Appl. Spectrosc. 72 (2018) 1047–1056. https://doi.org/10.1177/0003702818765183.

[243] Z. Gajarska, L. Brunnbauer, H. Lohninger, A. Limbeck, Identification of 20 polymer types by means of laser-induced breakdown spectroscopy (LIBS) and chemometrics, Anal. Bioanal. Chem. 413 (2021) 6581–6594. https://doi.org/10.1007/s00216-021-03622-y.

[244] C. Sommer, L.M. Schneider, J. Nguyen, J.A. Prume, K. Lautze, M. Koch, Identifying microplastic litter with laser induced breakdown spectroscopy: a first approach, Mar. Pollut. Bull. 171 (2021), 112789. https://doi.org/10.1016/j.marpolbul.2021.112789.