

BRNO UNIVERSITY OF TECHNOLOGY

Faculty of Mechanical Engineering

HABILITATION THESIS

Brno, 2020

Ing. Pavel Pořízka, Ph.D.



**BRNO UNIVERSITY OF TECHNOLOGY**

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

**FACULTY OF MECHANICAL ENGINEERING**

FAKULTA STROJNÍHO INŽENÝRSTVÍ

**DEPARTMENT OF PHYSICAL ENGINEERING**

ÚSTAV FYZIKÁLNÍHO INŽENÝRSTVÍ

**PROCESSING THE OPTICAL EMISSION SIGNAL OF  
LASER-INDUCED PLASMA**

ZPRACOVÁNÍ OPTICKÉHO SIGNÁLU LASEREM BUZENÉHO PLAZMATU

**HABILITATION THESIS**

HABILITAČNÍ PRÁCE

**AUTHOR**

AUTOR PRÁCE

**Ing. PAVEL POŘÍZKA, Ph.D.**

BRNO 2020

## Abstract

This habilitation thesis brings thorough overview of the Laser-Induced Breakdown Spectroscopy (LIBS), theory and applications, with the aim on the approaches to the processing of spectroscopic signal. Thus, its outline goes through the instrumentation (both ablation and detection parts), laser-matter interaction, dynamics of ablation and laser-induced plasma formation. This is done in order to show the origin of the data and to build up the basis for consecutive processing part. Standard (univariate) as well as novel, advanced chemometrics (linear and non-linear multivariate algorithms) ways in signal processing are discussed. The perspective of data science and machine learning is connected to the basics of plasma physics and analytical chemistry.

Body of publications of the author is summarized, discussed and put in the context with the rest of the thesis. Finally, potential and further evolution of the field of LIBS is given with respect to other reference and complementary techniques and mainly to the long-term development of the LIBS technique by the spectroscopic community.

## Abstrakt

Tato habilitační práce přináší zevrubný přehled o teorii a aplikacích spektroskopie laserem buzeného plazmatu (z angl. Laser-Induced Breakdown Spectroscopy; LIBS), se zaměřením na způsoby zpracování spektroskopických dat. V práci je postupně pojednáno o základní instrumentaci (jak ablační, tak sběrné části), o interakci laserového záření s látkou, dynamice laserové ablace a vzniku laserem buzeného plazmatu. Širší popis základních poznatků o LIBS je podán záměrně proto, aby bylo zřejmé, co jest zdrojem spektroskopických dat a co ovlivňuje jejich kvalitu. Část práce o zpracování spektroskopických dat je rozdělena na dva základní směry: i) klasický přístup jednorozměrné statistiky a ii) pokročilé vícerozměrné algoritmy (lineární i nelineární). Přístupy datové vědy a strojového učení jsou nakonec úzce propojeny se základy fyziky plazmatu a analytické chemie.

Vědecké publikace autora jsou diskutovány v průběhu této práce a jejich přínos je zasezen do diskutované problematiky. Nakonec je představena určitá vize o budoucím směřování, ponciálu a rozvoji v oblasti techniky LIBS; toto je podáno s odkazem na vývoj v referenčních technikách a na dlouhodobý vývoj ve světové komunitě zabývající se technikou LIBS.

## Keywords

Laser-Induced Breakdown Spectroscopy, Analytical Chemistry, Plasma Physics, Data Processing, Multivariate Data Analysis, Machine Learning, Chemometrics.

## Klíčová slova

Spektroskopie laserem buzeného plazmatu, analytická chemie, fyzika plazmatu, zpracování dat, vícerozměrná statistická analýza, strojové učení, chemometrie.

## Reference

POŘÍZKA, Pavel. *Processing the optical emission signal of laser-induced plasma*. Brno, 2020. Habilitation thesis. Brno University of Technology, Faculty of Mechanical Engineering.

## **Declaration**

Hereby I declare that this habilitation thesis was prepared as an original author's work on my own. All the relevant information sources, which I have used during preparation of this thesis, are properly cited and included in the list of references.

.....  
Pavel Pořízka  
March 17, 2020

## **Acknowledgements**

Let me acknowledge the support of the members of research group RG1-6: Materials Characterization and Advanced Coatings at CEITEC BUT, namely prof. Ing. Jozef Kaiser, Ph.D., and all the researchers and developers from related research subgroup of Laser Spectroscopy.

I am grateful to my wife who granted me endless patience and support during the preparation of this thesis.

This work is devoted to  
my wife Martina and my son David.

*"An expert is a man who has made all the mistakes that can be made  
in a very narrow field."*

Niels Bohr

# Contents

<b>Preface</b>	<b>3</b>
<b>1 Introduction to spectroscopy</b>	<b>6</b>
<b>2 Spectroscopy in analytical chemistry</b>	<b>9</b>
2.1 Laser-Induced Breakdown Spectroscopy . . . . .	10
2.1.1 Background . . . . .	10
2.1.2 Instrumentation . . . . .	11
2.1.3 Applications . . . . .	12
2.2 Hyphenated Systems . . . . .	13
2.2.1 LIBS–LA-ICP-MS . . . . .	14
2.2.2 LIBS–Raman . . . . .	14
2.2.3 LIBS–XCT . . . . .	15
<b>3 Theory of laser-induced plasma</b>	<b>17</b>
3.1 Laser-induced plasma . . . . .	17
3.2 Spectrum and emission lines . . . . .	21
3.3 Thermodynamic properties of Laser-Induced Plasma . . . . .	24
<b>4 Processing the emission signal</b>	<b>27</b>
4.1 Data Preprocessing . . . . .	27
4.1.1 Background subtraction . . . . .	27
4.1.2 Signal standardization . . . . .	29
4.1.3 Outliers filtering . . . . .	30
4.1.4 Line selection and intensity estimation . . . . .	31
4.1.5 Organization of a data set . . . . .	33
4.2 Statistics . . . . .	34
4.2.1 Normal Distribution . . . . .	34
4.2.2 Extreme Value Distribution . . . . .	35
4.2.3 Central Limit Theorem . . . . .	36
4.2.4 Hypothesis testing . . . . .	37
4.3 Univariate Data Analysis . . . . .	38
4.3.1 Quantification . . . . .	38
4.3.2 Mapping . . . . .	40
4.3.3 Depth Profiling . . . . .	42
4.4 Chemometrics . . . . .	43
4.4.1 Visualization . . . . .	45
4.4.2 Dimensionality reduction . . . . .	46

4.4.3	Multivariate quantification . . . . .	47
4.4.4	Multivariate clustering and classification . . . . .	48
4.4.5	Multivariate mapping . . . . .	50
4.4.6	Data libraries . . . . .	51
4.5	Figures of merit . . . . .	52
4.5.1	Quantification . . . . .	53
4.5.2	Classification . . . . .	55
	<b>Conclusion</b>	<b>57</b>
	<b>Bibliography</b>	<b>61</b>
	<b>A List of authored publications</b>	<b>76</b>
	<b>B Printed versions of most significant authored publications</b>	<b>82</b>

# Preface

Our society is overloaded with data. Nowadays, anyone can effortlessly access desired information. However, the information may be obtained from various sources with alterations and no reference to the original source. It is getting hard to distinguish true information from the fake one. All this results in redundant data with non-informative features. Moreover, information is often distracting; showing a lot of noise.

The spectroscopic community also faces a similar issue. Contemporary progress in analytical instrumentation enabled an enormous throughput in data collection. Numbers of analyzed samples and obtained multivariate data reach limits of standard signal processing procedures and even computing devices. Standard data handling and storage prior the data mining itself demands unprecedented computing power. Thus, the big data revolution is fully in its gears and biases the classical approach of analytical chemistry. The current needs for signal processing challenge the potential of chemometrics.

In the motivation of the book entitled *Statistics in spectroscopy*, Mark and Workman [1] stressed that “*the field of spectroscopy had departed from the ‘classical’ approach*”. Data are processed more often using multivariate algorithms that provide supreme performance over the univariate ones. However, basic univariate algorithms can, in many cases, provide the same performance and lead to comparable results if they are judiciously performed.

The data analysis has shifted from *estimation* of classical chemistry to *prediction* of data science. This paradigm shift is polarizing the spectroscopic community and induces a lot of controversy. Standard approaches of data science are accepted by spectroscopists only very cautiously or they are being completely misused.

The main issue of today is a direct processing of raw untreated spectra by using chemometric algorithms. This issue was also reflected by Hahn and Omenetto [2]: “*the use of sophisticated chemometric algorithms may lead to apparently successful discrimination among samples in which the key identifiers are not related to the nature of the samples but rather to contaminant and/or background features*”. Multivariate algorithms are misused as a *black-box* without any proper prior knowledge of the fundamentals of plasma physics and analytical chemistry on one side and the limitations of mathematical apparatus on the other. What is even more striking is the use of non-linear, more sophisticated statistical algorithms (such as variations of neural networks) to discriminate statistically insignificant data sets (low number of objects and/or variables, noisy spectra with poor signal-to-noise ratio, *etc.*).

Discussed controversy in data handling and the whole paradigm shift has set a solid motivation for my thesis in which I would like to dissect the standard and novel approaches in signal processing. The goal of my thesis is therefore to bridge the gap between traditional univariate approaches with strong background in plasma physics and analytical chemistry on one side and advanced data science, heuristic approaches on the other. This topic has

been of my paramount research interest and following chapters are based on an extensive literature research and my own personal long-term experience.

The outline of my thesis has a logical structure that builds up from the fundamentals of plasma physics through the understanding of emission signal to the univariate and multivariate algorithms. I am convinced that it is absolutely necessary to understand signal (spectra) as a representation of its source (plasma and related sample). Thus, with great excitement I follow the scientific work of many authors in the LIBS community who are forging novel data processing algorithms together with the assumptions obtained from plasma physics and analytical chemistry.

First of all, I give a brief historical overview of the evolution of spectroscopy, chapter 1. The timeline shows an increasing frequency of technical inventions and scientific discoveries, coming to the modern age of laser spectroscopy and computer science when it is getting hard to keep up with the pace of contemporary state-of-the-art. Chapter 2 brings considerations about the analytical chemistry and its two basic approaches: i) bulk and ii) surface analysis. The second approach induced the invention of several analytical techniques that are tangentially mentioned.

I would like to stress that in my work I focus exclusively on the elemental analysis of sample surfaces using the Laser-Induced Breakdown Spectroscopy (LIBS) technique. In spite of that, the conclusions that I draw from the literature research and my scientific work are applicable to the processing of spectroscopic signal obtained using any other complementary or reference analytical techniques. Background and motivation for LIBS are introduced together with overview of basic instrumentation and related parameters, chapter 2.1. Surely, the most crucial instrument is laser itself, being responsible for material ablation and laser-induced plasma (LIP) formation. Properties of a laser pulse have direct, the most significant, influence on the quality of LIP and obtained signal.

List of applications where we can fully exploit the benefits of LIBS are given in chapter 2.1.3. LIBS gained its status among other analytical techniques due to the high-repetition rate analysis with multielemental capability. This is the main reason why it has been intensively developed for elemental imaging (mapping) of large-scale sample surfaces. There is, of course, no will to develop a one-for-all LIBS technique and substitute its analytical counterparts. On the contrary, LIBS is used as a complementary technique extending the information gain. Tandem approaches are described in chapter 2.2, where LIBS is accompanied with mass spectrometry, Raman spectroscopy or computed tomography.

Chapter 3 delivers a complete overview of knowledge, fundamentals of plasma physics, that is necessary to master prior the judicious data processing of spectroscopic data. The matrix effect is described as a product of complex processes beyond laser-matter interaction. Assumptions for consequent laser ablation of material to be stoichiometric and induced plasma to be optically thin and in the local thermodynamic equilibrium are formed. In the second part of this chapter, spectrum and its emission lines are described in detail. Line emission and broadening mechanisms are bound to mathematical equations enabling further computations of LIP properties, electron number density and temperature.

The core of my thesis is embedded in the processing of emission signal, chapter 4. To the best of my knowledge there is no unified approach, algorithm, in data processing and individual steps have no strict order. The data preprocessing can consist of many partial steps or data can be processed directly in their raw format. Background subtraction, signal estimation and standardization, and outliers filtering are influencing the data structure and latent relationships within data sets and, thus, their impact on the analytical performance is considered. In my authored and co-authored publications I have invested a great deal of

interest into the individual steps of data preprocessing. Their impact on resulting figures of merit have been shown.

Statistical analysis of obtained signal is also dissected. We have been among first research groups to point out that typical LIBS data are not normally distributed. This limits the implementation of basic statistics over obtained data sets and leads to biased results. It was proved that normal distribution gets distorted and data rather follow the extreme value distribution.

Univariate data analysis approaches are listed and critically commented. Basic quantitative analysis is mentioned as one of the pain points and limitations of LIBS technique. On the other hand, qualitative mapping and depth profiling are promising applications of LIBS while fully exploiting its benefits.

Chemometrics are extensively exploited in LIBS data processing. I describe their application for quantitative analysis of selected analytes as well as for classification of complex sample matrices. Novel approach is being developed, the so-called multivariate mapping. My co-authored publication have also targeted this issue when implementing linear and non-linear algorithms. Lastly, I also raise discussion over potential data library/model transfer which can trigger further development in the field of LIBS and related signal processing.

Robust models are of paramount interest to the data scientists and processing of LIBS data is not excluded. Recently, we have introduced a benchmark data set to test the capabilities of individual LIBS groups in classification of LIBS data. This benchmark data set was evaluated during the EMSLIBS 2019 conference, held and organized by our LIBS group in Brno. Finally, the figures of merit used to quantify the performance of the whole analysis are listed.

In each chapter I complement current literature with my own scientific work. References are given and the list of my publications is attached as Appendix A of this thesis. As a conclusion, I also intended to predict the future development of the field of laser-based spectroscopy and related data processing.

My scientific career is bound mainly to the Brno University of Technology; first being affiliated to the Faculty of Mechanical Engineering and then also to the Central European Institute of Technology. As a representative of Brno University of Technology I stayed (2011-2014) at Federal Institute for Materials Research and Testing (Berlin, Germany) in the group of prof. Ulrich Panne and Dr. Igor B. Gornushkin. Several years later (2017-2018), I had the opportunity to become a Fulbright fellow at the University of Florida, (Gainesville, Florida, US). On this journey, I have published 35 scientific articles with impact factor and reached H-index 11 according to the Web of Science.

I have also pursued an academic career along the scientific one. My primary affiliation is the Faculty of Mechanical Engineering at Brno University of Technology. At the Institute of Physical Engineering, I gained the status of assistant professor and where I am involved in teaching. The main lectured course is the Engineering Optics which focuses on laser sources and Gaussian beams. My secondary courses target elementary physics and optics in the undergraduate level.

Recently, in both career paths, my position shifted while being the leader of the laboratory of Laser Spectroscopy. I am supervising undergraduate and doctoral students in writing their theses and leading junior scientists in their research and development projects.

*P.P.*

# Chapter 1

## Introduction to spectroscopy

Many texts regarding the history of optics and spectroscopy have been written<sup>1</sup>. Thus, I will not bore the reader by an extensive description of consecutive historical events that led us to the present state-of-the-art instrumentation and its application. Albeit, with respect to our predecessors, the most important findings, especially bound to optical emission spectroscopy, will be briefly mentioned.<sup>2</sup>

It is generally accepted that Sir Isaac Newton was the first one to use the term spectrum, in his book *Opticks* from 1704. This term was used to describe his observation when a thin ray of sunlight was dispersed by a prism to a specific rainbow-like structure, i.e. the dispersion of white light into its constituent colors. Newton concluded that each colour of light travels with a different velocity in a transparent media and is thus refracted under different angles. This phenomenon is applied in spectroscopy when a characteristic radiation of the emitter, *i.e.* light source, is studied. The spectrum of light was then a subject to an intensive investigation and, in 1802, William Hyde Wollaston noticed dips in the intensity of solar spectrum, *i.e.* the position of absorption lines, using an improved spectrometer based on the Newton's model.

In 1814, Joseph von Fraunhofer observed, independently on Wollaston, and assigned absorption lines in the emission spectrum of the Sun, see Figure 1.1. A similar spectrum can also be observed in the emission of other stellar objects. Those spectra are collectively referred to as the absorption spectrum. Fraunhofer used spectral lines namely for measurements of optical properties of transparent materials. In 1859, Gustav Kirchhoff and Robert W.E. von Bunsen discovered that the position of lines in the spectrum (their wavelengths) differ for individual materials. Those wavelengths coincided with the wavelength of absorption lines found in the solar and stellar spectra. It was thus obvious that the spectrum is characteristic to each material and related to its atomic structure. In the meantime, characteristic spectra of various materials were observed using excitation sources, such as flame (Talbot in 1826) and spark (Volta in 1776). Those findings triggered the investigation of characteristic optical emission of elements and was at the beginning of modern spectroscopy.

The spectroscopy further evolved when, later on, the experimental findings were described by theory. In 1860s, James C. Maxwell bound the behaviour of electric and magnetic field forming the electromagnetic waves by differential equations. In 1900, Max Planck

---

<sup>1</sup>This chapter was obtained mainly from following sources [3–5].

<sup>2</sup>At this point it is worth mentioning Jan Marek Marci (1595 – 1667), a Bohemian scientist and a professor at the Charles university in Prague. His work included the theory that light changes colour only by refraction and he also explained the origin of rainbow, in 1648.

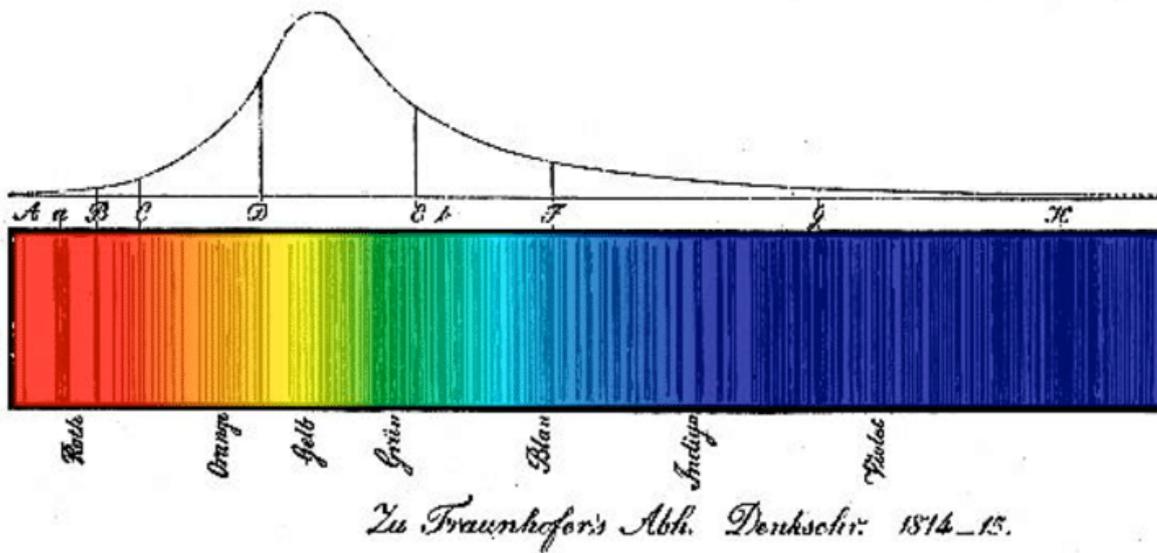


Figure 1.1: An example of the absorption lines in the spectrum of the Sun as observed by J. von Fraunhofer, 1814. Obtained from [6]

changed the cornerstones of physics and stood at the beginning of quantum physics. In his work, he brought a generalization of the energy in order to explain the phenomena occurring in the black body radiation, *i.e.* to describe the distribution of its radiation. From this point, the quantum physics developed dynamically and many groundbreaking theoretical findings were delivered in the beginning of the 20<sup>th</sup> century. Selected historical milestones, related namely to spectroscopy, are chronologically depicted in Figure 1.2.

The advancement in technology led to the invention of the first laser, in 1960, and the first inductively coupled plasma instrument, in 1963. This is where the historical prologue ends and phenomena (fundamentals of LIP, processing of obtained signal and applications of LIBS) discussed in this thesis begin.

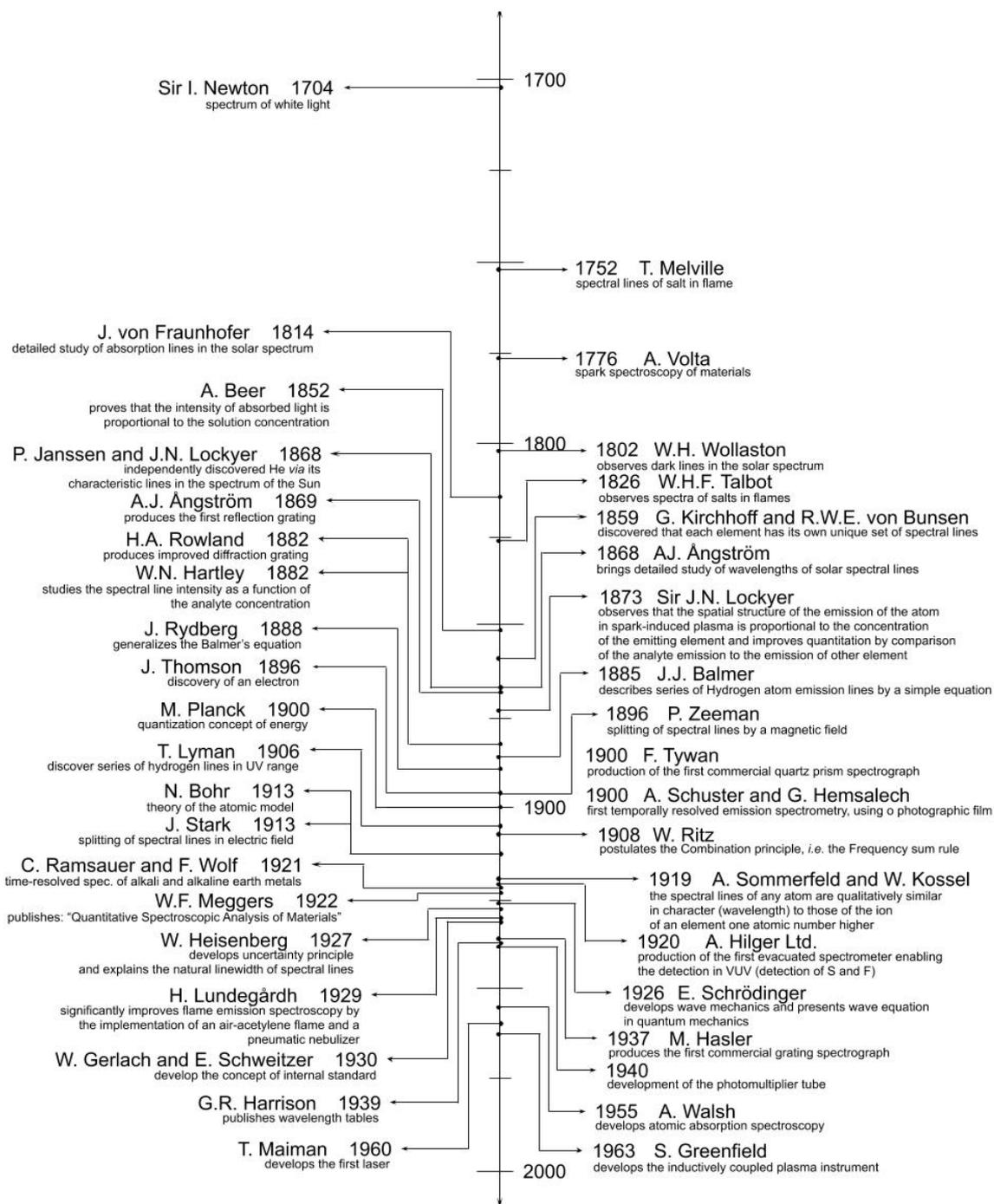


Figure 1.2: This figure chronologically depicts the most important experimental and theoretical findings that were done from Newton's work *Opticks* (1704) to the invention of the laser (1960) and the inductively coupled plasma instrument (1963). Note that presented historical milestones, taken from [5], were selected only in respect to optical emission spectroscopy and partly to optics and quantum mechanics. X-Ray spectroscopy was not considered at all.

## Chapter 2

# Spectroscopy in analytical chemistry

This chapter describes techniques used for elemental analysis typically applied in analytical chemistry and a comparison of their performance. The motivation for further development and utilization of LIBS technique is stated, including the introduction to the LIBS instrumentation and applications. Considerations on the capability of hyphenated systems are mentioned, forming future perspectives in materials' chemical and structural characterization.

**Analytical chemistry** The motivation to keep improving the analytical apparatus has always been firm from pioneering beginnings, reviewed in chapter 1, to the present day. The intense need drives researchers to understand natural processes in their full complexity. For those purposes, various techniques of analytical chemistry are being developed in order to identify and quantify the composition of matter with the highest possible accuracy.

The first issue of chemical analysis is the shift from sample to measurable signal response. There are many alternative ways to provide sampling of the material; as reviewed in the book of Skoog *et al.* [7]. Historically, most common sampling approach was devoted to dilution of obtained material in acids. Due to the dilution step, classical wet chemistry enabled only the analysis of sample bulk and the information about the distribution of elements within was lost.

The second issue is embedded in the excitation of the signal regardless of the state of matter. Basic excitation sources, flame and spark discharge, are being substituted by modern laser sources offering a wide range of parameters and performance. Finally, new capabilities of sampling, spot-by-spot analysis of sample surface, were gained after the introduction of laser-ablation based techniques. This opened new perspectives for the chemical analysis.

**Spectroscopic techniques** Spectroscopy is a vital tool of analytical chemistry and complements the information obtained using optical microscopy and especially mass spectrometry. I have to stress that this thesis deals solely with laser-based analytical methods, especially Laser-Induced Breakdown Spectroscopy (LIBS), enabling qualitative and quantitative elemental analysis of sample surfaces. Wet chemistry prior to Inductively Coupled Plasma Optical Emission Spectroscopy or Mass Spectrometry (ICP-OES/MS) are consid-

ered only as reference techniques to LIBS, while Raman spectroscopy and Laser Ablation Inductively Coupled Plasma Mass Spectrometry (LA-ICP-MS) as complementary ones.

There are several basic parameters to quantify the performance of individual techniques and related analytical systems:

- **multielemental capability** - simultaneous detection of various elements,
- **sensitivity** in the detection of trace elements,
- **throughput** related to the repetition rate,
- **lateral resolution** - number of measurements per unit area.

The output of aforementioned techniques is similar; series of signal responses originating from various sources (*e.g.* an emission spectrum) is obtained from a single analysis. Thus, it is possible to use mathematical and statistical methods and data processing algorithms developed for an application of selected analytical technique and to transfer it to another one. Such mathematical and statistical methods are collectively called chemometrics and are described in detail in chapter 4.

## 2.1 Laser-Induced Breakdown Spectroscopy

This chapter brings preliminary motivation for the development and utilization of LIBS as a method of analytical chemistry. Background of the LIBS technique and its history, variety in instrumentation and applications are also briefly discussed.

### 2.1.1 Background

Laser-Induced Breakdown Spectroscopy (LIBS) is an analytical technique of atomic emission spectroscopy capable of real-time qualitative and quantitative elemental analysis of a sample under investigation. A detailed description of LIBS technology, instrumentation, theory and applications may be found in related books [8–10] and extensive review articles [2, 11].

Constant development in instrumentation made LIBS a well-established technique with complementary capabilities to its analytical counterparts. Over time, LIBS has transformed from simply tinkered set-ups to a sophisticated instrumentation. LIBS systems are now produced as compact hand-held devices, state-of-the-art laboratory setups and robust stand-off systems. The sky was not a limit for LIBS and its most famous instrumental variation became the ChemCam device on the Mars Curiosity Rover. [12–14]

LIBS has many benefits (namely repetition rate and multi-elemental capability) that are fully exploited in large-scale elemental mapping and other applications, see chapter 2.1.3. The biggest drawback of LIBS, i.e. the sensitivity to trace elements, is being continuously mitigated with the advancement in technology. However, a necessity for no or minute sample preparation became an unfortunate cliché. Surely, there is no need to prepare a sample prior to the LIBS analysis, but this stands only in several applications. However, the best performance is achieved for the analysis of flat surfaces of solid samples. [15]

Analyzing a sample and obtaining a set of spectra is only the beginning. Data handling and signal processing is an inevitable part of each application and related elemental analysis. The amount of data obtained in a single experiment is constantly rising resulting in bulky

data sets. This phenomenon opens new challenges for further research while there is no unified, robust approach in data processing.

In brevity, this intense demand on data processing defines the core (chapter 4) of my thesis.

### 2.1.2 Instrumentation

During the decades of LIBS development, many diverse LIBS systems have been introduced.

A relative simplicity of LIBS instrumentation led to a construction of compact hand-held [16] and robust stand-off [17] analytical systems. The most famous LIBS system, the Chem-Cam, is a part of the Mars Science Laboratory (MSL) on the Curiosity rover. [14] Industrial LIBS systems were designed to achieve the most optimal cost-to-performance trade-off with respect to given application and its well-known needs and limitations. [18] Lab-based LIBS systems evolved without any restrictions and according to dreams of researchers enabling the multi-laser pulse ablation and simultaneous multi-spectrometer detection. Finally, there are also commercially available LIBS systems promoting LIBS between other standard analytical methods and not-a-technique developed by *connoisseurs*.

A typical LIBS system may be divided into two essential parts: ablation and detection units. Ablation unit consists of a laser and focusing optics which guide a laser pulse onto the sample surface. Detection unit is further decomposed to collection optics, incl. optical fibers, and a spectrometer equipped with a non-/intensified detector. More detailed information about LIBS systems and individual parts are discussed by Noll [9]. Castle *et al.* [19] studied variables influencing LIBS analysis and came to a conclusion that inter- and intra-measurement precision is optimized under different conditions.

**Lasers** Laser source is the most essential part of each LIBS system and therefore it deserves special attention. Basic parameters of lasers influencing the laser-matter interaction (see chapter 3.1) are:

- laser pulse duration (from ns to fs laser pulses),
- laser pulse wavelength (most often the harmonics of Nd:YAG),
- energy (from tenths to hundreds of mJ).<sup>1</sup>
- intensity profile of a laser pulse (typically Gaussian or flat-top),
- repetition rate (from units to hundreds of Hz).

**Detection unit** The detection unit is composed of spectrometers equipped with detectors and other opto-mechanical parts. Basic parameters relevant to detection:

- optical throughput of the whole detection unit (influencing the analytical sensitivity and limits of detection),
- resolving power of a spectrometer ( $\lambda/\Delta\lambda$ ) and its spectral range,

---

<sup>1</sup>Energy alone is not considered to be a relevant variable when it comes to the laser-matter interaction and laser ablation. Estimating the irradiance (the radiant flux received by unit area; in units of  $\text{W}\cdot\text{cm}^{-2}$ ) from laser energy and spot size is more appropriate. The irradiance is often confused with intensity, *i.e.* power transferred per unit area.

- readout time of the detector, its dark noise and quantum efficiency.

One of the biggest guiding principles of LIBS technology is its instrumentation flexibility. However, it may be turned against itself into the biggest nightmare. Differences in existing LIBS instrumentations (used laser sources, spectrometers and detectors, geometries in laser ablation and emission collection) and their individual parameters lead to difficulties when one wants to compare their analytical performance. Surely, the utilization of various figures of merit (as listed in Chapter 4.5) is natural. However, comprising two significantly different LIBS systems (for instance, i) a low-cost system for the on-line detection of impurities in plastics and ii) a high-end lab-based system) in terms of sensitivity to trace elements has to be accepted cautiously.

### 2.1.3 Applications

In the last decade, LIBS has undergone a dynamic evolution and, from my personal point of view, it finally became an established analytical method. However, it has been an arduous journey when finding suitable applications in order to fully make use of its potential. Galbács [20] brought a concise review of LIBS applications regardless of the field of interest.

The benefits and namely the limitations of the LIBS technique have to be critically evaluated when finding its suitable real-life applications. Especially, industrial applications are demanding *the performance-price trade-off* to be maximally optimized. LIBS systems find a place directly on production lines due to their instrumental simplicity and robustness. The goal is to find an adequate distance from the sample in order to ensure the protection of optical and opto-mechanical parts together with a sufficient solid angle for detection.

Industrial applications are ranging from the direct analysis of liquid steel slag, through depth profiling of thin Zn-layers, to sorting of scrap material.[18, 21] Recently, sorting of waste electronic and electrical equipment has been gaining particular attention [21, 22]. The increasing production and usage of plastics induced also new applications of LIBS for sorting of plastic material and quality control with respect to the trace analysis of toxic metals.[23] In civil engineering, LIBS outperforms other analytical methods in direct detection of chlorine on the structure surface which is correlated to the level of concrete corrosion.[24] Geology provides an endless number of sample matrices to be classified [25]. The robustness of LIBS instrumentation is proved by its *in-situ* implementation on conveyor belts or stand-off devices; such as ChemCam device on Mars Curiosity rover [14].

LIBS is being intensively developed in biological applications and for a detection of various chemical substances; this is also influenced by its unique capability to detect major organic elements (*i.e.* C, H, O, and N). This advantage is employed in the classification of artificial chemical products; such as plastics[23] or explosives[26]. Identification and quantification of biological samples was studied in the case of bacteria [27], algae [28], plants [29, 30], and food products [31].

Most recently, medical and clinical applications have been of a particular interest.[32–34] LIBS is capable to provide elemental mapping of large scale areas with a high lateral resolution, whole slide imaging. Elemental imaging provided by using LIBS is being exploited across applications, see a selection in Figure 2.1. This makes LIBS analysis a vital alternative to current analytical solution with complementary performance.

In my personal opinion, *mapping of large surface areas will become an essential part of elemental analysis and will be a prominent advantage of LIBS method.* Thus, a further improvement in the LIBS performance, namely repetition rate, are inevitable.

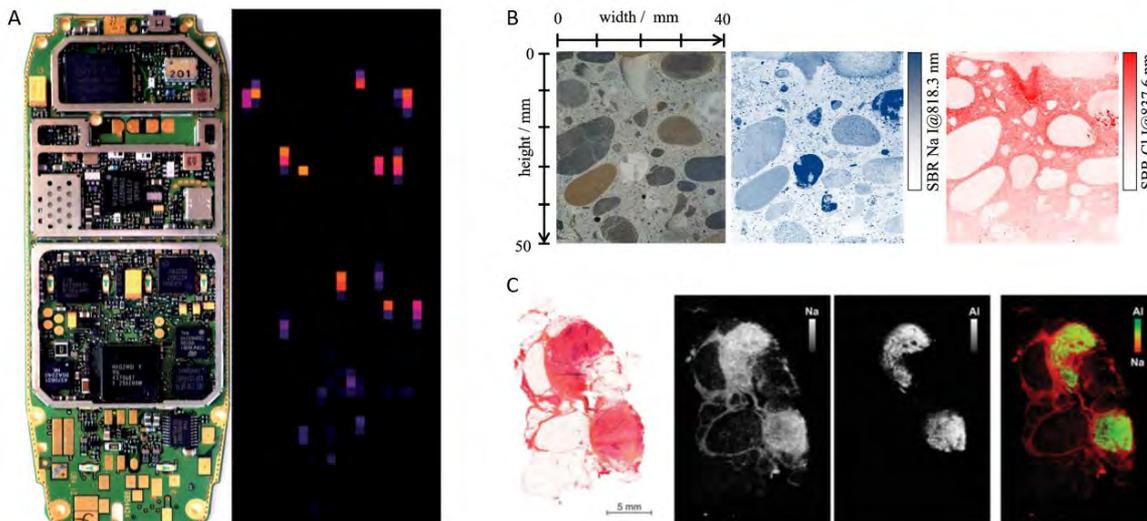


Figure 2.1: A) Photograph of the front side of a printed circuit board from a mobile phone (left) and raster scan in false colours of spectroscopic signatures of tantalum (right). Obtained from [21]. B) Cross section of a concrete sample (left) and element distributions (standardized intensities) of sodium and chlorine (middle and right). Obtained from [35]. C) Aluminum accumulation in skin reactions to Al-adsorbed immunotherapies. Histopathological morphology of a cutaneous granuloma (left) with corresponding elemental images. Obtained from [36].

## 2.2 Hyphenated Systems

Progress in instrumentation and data processing enables a less demanding utilization of individual techniques. Their benefits are combined together in the so-called hyphenated <sup>2</sup> systems in order to overcome their individual drawbacks.

Nowadays, coupling of individual techniques is done by using separate systems. However, there are plenty of efforts to deliver instrumental solutions enabling a spectroscopic analysis with more than one technique. Thus, hyphenated systems are getting an increasing popularity and are being intensively developed. Hyphenated analytical systems have been reviewed elsewhere [2, 20].

The utilization of more techniques for investigation of the same set of samples provides more complex information as well as data that have to be processed. In this chapter, I describe three main combinations that I deal with. Thus, I refer mainly to publications that I have co-/authored. Despite that, obstacles to be overcome in the implementation of selected combinations as well as in the consecutive data processing are rather general and give foundations for further work.

There are other combinations that are being developed but they are out of scope of this thesis and out of the capacities of RG1-6 at CEITEC. One of such combinations is the laser ablation LIBS (LA-LIBS); it might be rather considered as a tandem system. It is being developed by Prof. David W. Hahn from the University of Arizona, US-AZ

<sup>2</sup>Note that I have adopted the term *hyphenated systems* from [2]; it is used to describe systems where LIBS is coupled together with other analytical techniques to enhance the power of obtained information. The combinations of techniques that are using the same source<sup>3</sup> should be called *tandem systems*; such as LA-LIBS and LIBS-LA-ICP-MS.

(formerly University of Florida, US-FL) [37]. In this case the matter is pre-ablated and the aerosol is led to the interaction region where it is blasted with another laser pulse in order to induce plasma and consecutively to detect its optical emission. Such sampling enables to significantly mitigate matrix effect.

### 2.2.1 LIBS–LA-ICP-MS

In brevity, I describe the tandem LIBS–LA-ICP-MS technique. Their combination has been developed for past decade [29]. LIBS is gaining a stronger position; maturing to be an adequate source of information complementary to LA-ICP-MS.

Tandem utilization of both techniques is naturally centered around *the source* itself, the laser ablation of matter. Both techniques share the same plasma plume but each of them demands a different approach for spectroscopic feedback. In LIBS, plasma has to reach high temperature and excitation while hovering over the ablation spot; this enables an effective collection of characteristic emission. On the other hand in LA-ICP-MS, plasma plume is dragged by a carrying gas away from the interaction region into the ICP itself.

The quality of a laser-induced plasma and its parameters that is demanded by both techniques essentially differs. Contradictory experimental parameters force further restrictions on the ablation process which then becomes a trade-off between pure LIBS and LA-ICP-MS. This phenomenon was also of a paramount interest in our recent publication [38]; where the process of laser ablation was investigated with respect to the quality of aerosol formed. Moreover, the construction of the laser ablation cell enabling optical emission and mass collection is then more challenging. Thus, potential benefits of LIBS–LA-ICP-MS tandem is of intense discussions in the spectroscopic community.

From the analytical point of view, the combination of LIBS and LA-ICP-MS, both delivering elemental information, seems redundant. Yet still, both methods are developed into one tandem system. In general, LIBS is used for a detection of halogens, light elements (such as Li, Be, Mg) and macro-elements found in organisms (C, H, N, O, P, and S). LA-ICP-MS provides a sensitive detection of trace elements, such as metals in cross-section of soft tissues. The application of LIBS–LA-ICP-MS tandem was intensively studied by the group of Dr. Limbeck from the Technical University of Vienna, [39–41].

### 2.2.2 LIBS–Raman

Raman spectroscopy provides information about the molecular composition of the investigated sample. Briefly about Raman spectroscopy[42]; laser pulse illuminates the sample (or a region on its surface) and excites chemical bonds between atoms. Those chemical bonds then tend to vibrate and rotate. The frequency of laser photons is shifted by a certain amount with respect to the inelastic scattering induced by the interaction between photon and chemical bond. Molecular analysis is obtained from a detected optical spectrum, where the Raman shift is related to the type of chemical bond as a sort of fingerprint.

The combination of laser-based spectroscopic techniques is, thus, evident. LIBS, as a quasi-destructive technique, is applied after Raman scanning. The information obtained by using LIBS and Raman systems is complementary (elemental and chemical composition). Thus, the usage of LIBS–Raman hyphenated system provides complete chemical information about the investigated sample. Moreover, the combination of LIBS and Raman into one system with only one high-frequency pulsed laser is at hand since a Raman system utilizing gated detectors was introduced [43]. Lin *et al.* [44] tracked the development of LIBS and Raman hyphenated systems and foresaw its main future application in the field

analysis. In our review article [28], the potential utilization of combined LIBS–Raman systems, including LA-ICP-MS, was dissected with respect to the analysis of algae.

Applications of a LIBS–Raman system range from biology to mining. The biggest advantage that can be yielded from this system is when Raman spectroscopy detects changes in the molecular composition of a system as a response to a change in elemental composition; and vice versa. This was demonstrated in my feasibility study [45] where the ratio of saturated and unsaturated fatty acids in selected algal strains was indirectly related to Pb-nutrition stress. In another study [46], analytical performances of LIBS and Raman techniques were combined to achieve an improved detection of fluorine in a mineral sample. The detection of fluorine is challenging for other techniques, also in LIBS it is most conveniently detected *via* a molecular band (CaF) that is formed in later stages of laser-induced plasma formation. Raman spectroscopy was used to provide a reference analysis.

Lastly, Prochazka *et al.* [47] used LIBS–Raman hyphenated system for the classification of bacteria. A significant improvement in classification accuracy was achieved when LIBS and Raman spectra were merged together. Here, we arrive to the issue that has not been solved yet - processing of merged data set. A similar approach was also used by Hoehse *et al.* [48]; LIBS and Raman spectra were simply merged in their raw format prior the classification of inks. However, this injudicious data processing step opens further discussion. LIBS and Raman spectra significantly differ in their background and peak intensities, composition of spectral lines and bands, *etc.* This further extends requirements of data pre-/processing raised in chapter 4.

### 2.2.3 LIBS–XCT

In the previous two chapters, a combination of laser-based spectroscopic techniques was introduced, which provides complete elemental or even chemical information. A next step dwells in structural information which might be reached by using tomographic techniques; for instance X-Ray Computed Tomography (XCT).

Basically, a sample is exposed to X-Rays from consecutive directions resulting in images with varying degrees of absorption. Obtained images are then reconstructed using, for instance, the Radon transform. The volumetric information is then disassembled into individual parts *via* a tedious segmentation process. This serves as a basis for further investigation; XCT images are used to guide consecutive sample preprocessing in order to provide chemical mapping of regions of interest.

To the best of my knowledge, one of the first pioneering works in this particular field of development were delivered by Kaiser *et al.* [49, 50]. The synchrotron X-Ray source was used in their early work to achieve structural information distributed in 3D images of investigated objects. First, the uptake of toxic metals (Cd and Pb) in vegetal tissues of various species was checked using synchrotron-XCT (distribution of metals in plants) and consecutively using femtosecond LIBS (elemental analysis of hot-spots). [49] Second, a deformation of snake vertebrae was investigated. [50] It was confirmed that the deformation was a result of bone sickness; deficiency of calcium and abundance of phosphorus. In our recent study [51], lab-based XCT was utilized to provide 3D imaging of a Pb-bearing mineral. Sample was fixed in epoxy and its surface was polished for LIBS analysis. LIBS map and respective XCT slice were then compared to ascertain the presence and location of Pb in the mineral, see Figure 2.2. After the segmentation of XCT data and confirmation by using LIBS, total volume of Pb inclusions were determined.

Finally, the work of Galmed *et al.* [52] delivered a combination of LIBS and XCT. In this case, XCT served as a source of supplementary information in the investigation of LIBS ablation craters.

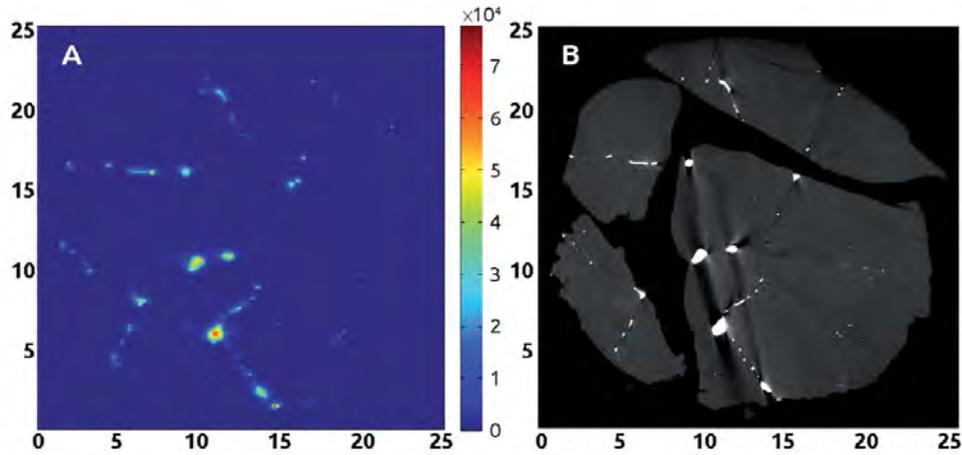


Figure 2.2: A sample of images obtained by using LIBS–XCT hyphenated system; a) chemical map of lead depicted as the intensity of spectral line Pb I 405.78 nm and b) tomogram top cross section. The orientation and size of both images were adjusted to fit together. The spots on the tomogram image with higher density (related to Pb) are depicted in white. The X and Y axes in both images are in millimeters and the intensities in the elemental map represented in the color bar are in arbitrary units. Obtained from [51].

## Chapter 3

# Theory of laser-induced plasma

Before I get to the discussion over the data processing, I would like to describe the source of our signal (spectroscopic data) - the Laser-Induced Plasma (LIP). I believe that it is absolutely crucial to understand its behavior and estimate its properties. Note that this is in contrast to the heuristic approach discussed later in the chemometrics chapter, chapter 4.4.

The motivation to this chapter may be postulated with the need for an accurate (quantitative) analysis. In order to achieve this, LIBS experiment has to meet following requirements [8]:

- a LIP has to be a product of a stoichiometric ablation,
- a LIP has to be found in the local thermodynamic equilibrium (LTE),
- spectral lines have to be optically thin.

Individual requirements will be step-by-step introduced and discussed in detail with references to classic literature sources.

Note that in this chapter I chiefly consulted LIBS books by Miziolek *et al.* [8] and Noll [9], LIBS-related reviews of Hahn and Omenetto [2, 11] and Russo *et al.* [53], and namely books on plasma spectroscopy by Griem [54, 55], Ingle and Crouch [56], and Fujimoto [57].

### 3.1 Laser-induced plasma

Plasma is generally considered to be the fourth state of matter. As Fujimoto expresses in the introduction to his book Plasma spectroscopy [57], plasma is an environment in which the particles co-exist. It influences their quantum behaviour, which leads to variation in their parameters and quality of detected spectral lines (intensities, spectral shift, *etc.*). In turn, detected signal provides information about the plasma (plasma diagnostics) and sample (qualitative and quantitative analysis).

Before I start with the introduction to the laser ablation, I would like to state several assumptions and, thus, reduce the number of parameters involved in the laser-matter interaction mechanisms. First, my thesis is focused solely on the analysis of solid samples with pretreated flat surfaces. Second, laser ablation induced only by using nanosecond laser pulses is considered. Laser sources generating nanosecond pulses are predominantly utilized in typical LIBS experiments. Recently, femtosecond laser sources have become popular for their supreme ablation performance and reducing costs. The femtosecond laser ablation

undergoes significantly different mechanisms compared to the nanosecond one. The power is transferred faster to the material resulting in the Coulomb explosion. Consult following articles [58, 59] for a detailed description of femtosecond laser ablation.

**Laser ablation** LIBS is a laser-ablation-based technique and the sampling is provided by using a pulsed laser source. Thus, in this section, I circle around the process of laser ablation and optical breakdown.

The simplification of the LIBS principle and related laser ablation is depicted in Figure 3.1.[9] High power laser pulse is delivered to the sample surface 3.1(1), energy is coupled to material during the duration of the pulse. Material starts to melt down and evaporate into its surrounding environment 3.1(2-3). Still persistent laser pulse heats up the vapor and luminous laser-induced plasma (LIP) is generated 3.1(4). Species within the plasma plume undergo mutual collisions, get excited and emit characteristic radiation 3.1(4-7). Plasma plume spreads freely into its ambient surroundings, cools down and decays 3.1(8). An ablation crater is left after analysis on the sample surface. Species formerly forming a LIP are removed from the interaction area naturally driven by their own motion or dragged away by an external flow of gas.

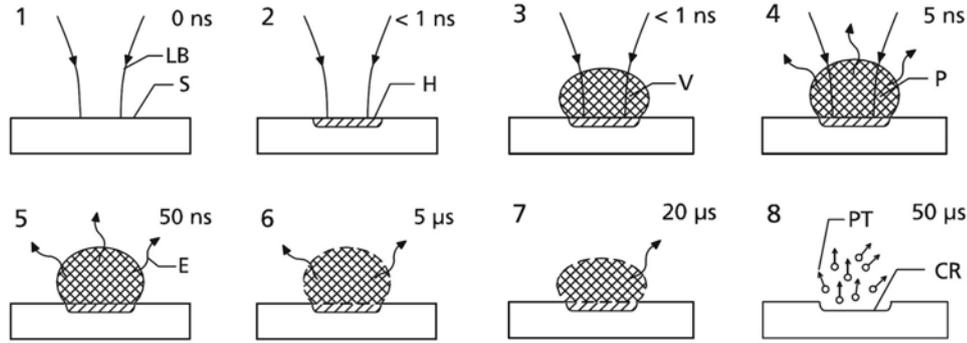


Figure 3.1: The schematic principle of laser ablation and plasma formation occurring during the LIBS analysis. LB = ablation laser pulse, S = sample, H = region of energy deposition, V = material vapor, P = plasma, E = element-specific emission, CR = crater, PT = particles. The times given depict the temporal evolution after the impact of laser pulse on the sample surface. Obtained from [9].

Russo *et al.* [53] delivered a concise description of the laser ablation process with regards to LIBS experimental parameters. Fast and dynamic mechanisms of laser ablation occur over a short time period; typically, the duration of a nanosecond laser pulse. They defined four thresholds with respect to the irradiance, the laser intensity coupled to material. Those thresholds characterize different changes in the ablation mechanisms.<sup>1</sup>

Laser-matter interaction and consecutive laser ablation is a product of experimental parameters (*e.g.* laser pulse wavelength, duration, and energy; ablation geometry and spot size; physical and chemical properties of a sample under investigation; ambient environment; *etc.*) Those parameters have a significant impact on the ablation and plasma formation, its

<sup>1</sup>The last two cases discussed in the review of Russo *et al.* [53] are not of interest to this thesis, since the laser pulse duration and intensities are out of formerly defined scope. However, they predicted femtosecond lasers to be significantly reducing the impact of the matrix effect and fractionation due to a fast LIP formation through Coulomb explosion. In turn, those sources are more suitable for stoichiometric ablation. As it turned out, the supreme performance of fs-lasers in LIBS was proved. [59]

parameters and quality of detected spectra. Several laser-related parameters have already been listed in chapter 2.1.2. The most important parameters seem to be the irradiance as Hahn and Omenetto [11] commented: „*The ablation process has a complex convoluted dependence on both the pulse energy and the pulse duration.*“ Thus, it is necessary to consider not only the total energy coupled to the interaction region but also the rate at which it was done.

First of all, the irradiance has to reach a certain level in order to start the process of ablation, *i.e.* *ablation threshold*.<sup>2</sup> The absorption of photons by free or bound electrons within the sample is considered to be the onset of laser ablation process. The energy is then transferred from electrons to the lattice through mutual collisions and the lattice heats up. The sample is irreversibly damaged<sup>3</sup> when the melting temperature is exceeded. Species from melted spot then get vaporized and the dense vapor forms a plume surrounded by ambient environment.

The optical breakdown of solid materials is initiated by the inverse Bremsstrahlung or multiphoton ionization, depending on the pressure/density of the vapor plume.[8, 53] Both processes are followed by the avalanche or cascade ionization of the ablated material. The ionization of the plasma plume is a result of the absorption of the laser energy by evaporated electrons and collisions with atoms in the plasma plume. This takes place when the electrons' kinetic energy exceeds their ionization potential.

If the irradiance is further increased, the ionized plasma plume causes a shielding effect and in turn reduces the *ablation efficiency*, see also [11]. Another increase of irradiance leads to the critical thermodynamic point of the material, see [53]. When the sample surface temperature approaches the critical temperature, vapor bubbles are generated within the melted material. The ablation of melted material into vapor is then accompanied by an ejection of melted droplets.

The expanding plasma plume absorbs laser intensity, grows and the atoms which are inside get ionized. Laser wavelength has an impact on the coupling of the energy. Longer wavelengths are easily absorbed by the plasma plume while the absorption is a function of the square of the wavelength. Moreover, as Miziolek *et al.* [8] suggests, the laser wavelength may have an impact on the ablation threshold. At this point I consider the plasma plume to be fully generated. Further discussion over the evolution of LIP follows in the section below.

Going back to the onset of laser ablation, the properties of ambient environment and mainly the sample itself has to be mentioned. The pressure of various environments surrounding the ablated material acts as a natural spatial confinement for the plasma plume.[60] Under vacuum conditions<sup>4</sup>, the expansion of the plasma plume is not restricted and the pear- or cigar-shape of the plume spreads freely into the surrounding space.

From another perspective, physical parameters of a sample and its chemical composition contribute to the complexity of laser ablation mechanisms. The contribution of physical and chemical properties is collectively described by **the matrix effect**. Surely, the *cliché* of no sample preparation is fortunately forgotten. Jantzi *et al.* [15] delivered a thorough review about strategies in preparation of various samples and their impact on consecutive LIBS

---

<sup>2</sup>Working with under-threshold irradiances enables to study laser-matter interaction and energy coupling, as commented by Russo *et al.* [53].

<sup>3</sup>Due to the low amount of material ablated with one pulse (typically less than a microgram), the laser ablation is considered to be quasi-non-destructive.

<sup>4</sup>Vacuum LIBS conditions are typically considered for reduced pressure to circa 1 mbar. Below this pressure, a LIP expands dramatically when quickly losing energy.

analysis. For the sake of brevity I consider the physical properties (surface roughness and hardness) as uniform during the whole sample set and I will not consider them at all in this thesis. However, chemical composition has a significant impact on the laser ablation.[8, 11] Over decades, an extensive investigation is in continuous process in order to deliver means to mitigate and overcome the matrix effect; one of them is also the signal standardization discussed in chapter 4.1.2.

Finally, there is a necessity to reach stoichiometric ablation, *i.e.* to avoid any fractionation. Fractionation biases the chemical analysis while the obtained LIP emission is not characteristic of the sample composition. Russo *et al.* [53] states that the fractionation occurring during laser ablation is a function of laser irradiance. The stoichiometric ablation occurs for laser intensities over  $\text{GW}/\text{cm}^2$ , conditions met in a typical LIBS experiment.[8] A more detailed discussion on the stoichiometry can be found in [61]

**Evolution of Laser-Induced Plasmas** The mechanisms of laser ablation as a function of laser intensity and of the formation of a plasma plume were discussed above. The complexity of laser-matter interaction significantly influences further expansion, evolution and decay of a LIP. The pulsed laser ablation makes the processes involved in plasma evolution dynamic and short in persistence. The description of evolutionary stages of a LIP is bound to its thermodynamic properties, temperature and number density. Detailed description can be found in [8, 9].

When the plasma plume is formed it immediately starts to expand from the interaction region with velocities reaching  $10^6$  cm/s. The persistence of a LIP is in units to tens of microseconds during which it emits analytically insignificant radiation (continuum and Bremsstrahlung), which occurs namely in the beginning of LIP persistence. The temporally resolved detection is optimized to avoid disturbing radiation and to reach the best possible signal-to-noise or -background ratios. The gating is, however, a trade-off between intense spectral lines sitting on high background or lower intensity spectral lines with mitigated background contribution.

Imaging plasma morphology and their spatial-temporal evolution is of interest to many researchers using various techniques, from high-speed photography through shadowgraphy to tomography. Plasma created with pulsed sources show distinct inhomogeneity and temperature and electron density gradients. The heterogeneity of elemental distribution within plasmas is also present. Bai *et al.* [62] proved the inhomogeneity in elemental distribution within a LIP using narrow spectral filters. The extension of this study shown the differences in plasma generated under various ambient gases (air and argon). Eschlböck-Fuchs *et al.* [63] provided a tomographic study using the Radon transform of double pulse LIPs. They have shown that the two-pulse technique improved the spatial homogeneity of the plasma. This is a promising result for further research of multi-pulse LIBS analyses.

**Local Thermodynamic Equilibrium** The species in the LIP are described by their kinetic, excitation, ionization and radiative energies; all of them following respective distribution (Maxwell, Boltzmann, Saha and Planck).[8] If all the distributions are determined with the same temperature  $T$ , then the plasma is found to be in a *complete thermodynamic equilibrium*. Such ideal conditions can be met only for stationary, stable and homogeneous plasma sources; none of which is the case of laser-induced plasmas.

Disequilibrium, a deviation from complete thermodynamic equilibrium, occurs and collisions between species forming up the plume are a dominant mechanism. This is typical for LIP sources, such as the case of LIBS plasmas. Those plasmas are found in the state

called *local thermodynamic equilibrium* (LTE). The most frequently used criterion to ensure the existence of LTE in a LIP is the McWhirter criterion [11]:

$$n_e \geq 1.6 \cdot 10^{12} \sqrt{T} (\Delta E)^3, \quad (3.1)$$

where  $n_e$  ( $\text{cm}^{-3}$ ) is electron density,  $T$  (K) is thermodynamic temperature and  $\Delta E$  (eV) is the highest possible energy transition for which the condition holds. This criterion demands the collision rate to be at least ten times higher than the radiative rate.

The McWhirter's criterion is used to ascertain whether the LIP is in LTE. However, Hahn and Omenetto [11] stressed that the criterion was derived for stationary, homogeneous, and optically thin plasmas. Therefore they urged to abandon this criterion as a proof of LTE.

**Optical thickness** The last requirement stated in the beginning of this chapter that demands further discussion is the optical thickness of a LIP.[9, 11]

A LIP is non-homogeneous with gradient distribution of temperature and electron density. This suggests that the core of a plasma plume has higher temperature compared to plasma envelope. The excitation of species is then more frequent within the plume core. Emitted radiation passes through the regions of cold plasma and can get absorbed by species of the same element. This phenomenon is called self-absorption and is reflected in the emission spectrum of the affected line by deviations from the theoretical line profile. A dip in the center of the line appears for extreme cases of self-absorption, *i.e.* self-reversed lines.

The rate of absorption is given by the self-absorption factor  $K_\nu$  (no units) [11]:

$$K_\nu \equiv \frac{\tau_\nu}{1 - \exp[-\tau_\nu]}, \quad (3.2)$$

which is a function of the optical depth  $\tau_\nu$  (no units):

$$\tau_\nu = k_\nu^* \ell, \quad (3.3)$$

where  $k_\nu^*$  ( $\text{cm}^{-1}$ ) is the absorption coefficient including the stimulated emission and  $\ell$  (cm) is the length of the plasma along the axis of observation/detection. If  $\tau_\nu \ll 1$  than the plasma is optically thin to the emission line of interest; otherwise the plasma is considered to be optically thick.

The self-absorption is more significant in the case of resonant lines originating as the transition from excited level to the ground state. Thus, resonant lines should be used only cautiously for any further analysis. The influence of the optical depth with respect to emission of spectral lines was studied by Aragon *et al.* [64] with perspective of estimating the degree of their self-absorption.

## 3.2 Spectrum and emission lines

A spectral line is an intense peak in the detected continuum<sup>5</sup> emission spectrum of a LIP.[8] In LIBS plasmas only neutral and singly ionized atoms are present together with molecules formed in later stages of LIP evolution. Spectroscopic analysis of detected spectrum provides qualitative and quantitative analysis of the sample composition.

---

<sup>5</sup>A result of Bremsstrahlung and electron-ion recombination.

**A spectrum of Laser-Induced Plasma** As I have discussed, a LIP is considered as a source of emission. This emission is detected over the spectral range of interest. Obtained spectrum is a combination of atomic and ionic spectral lines lying on the background continuum.

The emission mechanisms of electrons, ions and atoms are not strictly divided during the LIP persistence and their spectral features are combined. If the stoichiometric ablation is considered, the spectrum provides reliable qualitative and quantitative information about the elemental composition of the sample under investigation. The quality of a LIP spectrum depends on laser-matter interaction and experimental parameters as discussed above.

Typical LIP spectrum is:

- **sparse** - spectral features are overloaded with the background noise,
- **redundant** - each element can be detected on various wavelengths,
- **time dependent** - the composition of detected spectrum changes during the LIP evolution,
- **multi-elemental** - wide range of elements is concurrently detected,
- **chemical fingerprint** - spectrum represents the sample from which the LIP originated.

**Line emission** The emission of a LIP is dependent on interrelated parameters that may be factored out. The signal  $S$  (counts) of an analyte atom or ion is then expressed as follows (obtained from Miziolek *et al.* [8]; chapter 3):

$$S = A_{ji} f_{int} f_{exc} f_{det}, \quad (3.4)$$

where  $A_{ji}$  is the Einstein coefficient of the transition from the upper level  $j$  to the lower level  $i$ .  $f_{int}$  is a function of parameters collectively describing the laser-matter interaction and consecutive laser ablation of matter.  $f_{exc}$  describes mechanisms leading to the excitation/ionization of atomic or ionic species of an analyte and their consequent emission. The function  $f_{det}$  characterizes the radiation environment and the optical thickness of a LIP. This enigmatic equation shows that the detected spectral line carries not only the qualitative and quantitative information about the sample composition but also other information describing the properties of laser ablation and the properties of plasma itself.

More rigorous description of obtained signal is given by Noll [9]; chapter 9. The intensity of line emission is given as:

$$I_{ji} = N_a^z A_{ji} \frac{g_j}{u_a^z(T)} e^{-\frac{E_j}{kT}}, \quad (3.5)$$

where  $N_a^z$  is the density of species  $a$  with charge  $z$ ;  $g_j$  is the statistical weight of upper level  $j$ ;  $u_a^z(T)$  is the partition function<sup>6</sup> of species  $a$  with charge  $z$ ; and  $E_j$  is the energy of upper level  $j$ . This relation intrinsically contains the Boltzmann distribution relating the population density of excited level  $j$  of species  $a$ .

<sup>6</sup>The partition function  $u_a^z(T)$  of species  $a$  with charge  $z$  is expressed as follows [9]:

$$u_a^z(T) = \sum_{i=1}^{i=n^*} g_{a,i}^z \exp\left[-\frac{E_{a,i}^z}{kT}\right], \quad (3.6)$$

where  $i = 1$  denotes the ground state and  $i = n^*$  the highest level which is still bound.

The equation 3.5 is related to the line emission coefficient  $\epsilon_{ji}^{\nu, d\Omega}$  as follows:

$$\epsilon_{ji}^{\nu, d\Omega} = \frac{1}{4\pi} \Gamma(\nu) h\nu_{ji} I_{ji}, \quad (3.7)$$

where  $\Gamma(\nu)$  is the line profile as a function of the frequency,  $h\nu_{ji}$  is the energy of the transition from upper level  $j$  to a lower one,  $i$ . The unit of  $\epsilon_{ji}^{\nu, d\Omega}$  is  $\text{W}/(\text{m}^3 \text{sr s}^{-1})$ . The line profile function  $\Gamma(\nu)$  is normalized as follows:

$$\int_{-\infty}^{+\infty} \Gamma(\nu) d\nu = 1. \quad (3.8)$$

**Line broadening** From theory, the transition between energy levels is equal to exact energy difference and, in turn, the line profile should be described with a delta function. This is, however, not the case in real spectra obtained during spectroscopic analysis.

The line profile function  $\Gamma(\nu)$  depends on the broadening mechanisms involved. Several broadening mechanisms might be considered; they are in detail described elsewhere [9, 55–57]. However, not all of them have relevant impact on the broadening of LIP spectral lines.

*Natural broadening* is a result of the uncertainty principle relating the lifetime of the transition to the uncertainty in the energy, or intrinsically to the frequency/wavelength. Radiative lifetime of a transition is inversely proportional to the Einstein coefficient as  $\tau_r = A_{ji}^{-1}$ . The Einstein transition probabilities of atoms/ions in the LIP are in the range from  $10^6 \text{ s}^{-1}$  to  $10^8 \text{ s}^{-1}$ . The full width at half maximum (FWHM) of a spectral line is then from  $10^{-3} \text{ nm}$  to  $10^{-5} \text{ nm}$ , respectively. Thus, *natural broadening* is usually negligible in typical LIBS experiment.

*Doppler broadening* assumes that emitting species within the LIP are in constant movement along the observation path. The emitter velocities follow the Maxwellian distribution; for one-dimensional case the distribution becomes Gaussian. As adapted from Griem [55], the line half-width induced by the Doppler broadening is then:

$$\nu_D = \sqrt{\frac{2kT}{m}} \frac{\nu_m}{c}, \quad (3.9)$$

where  $k$  is the Boltzmann constant,  $c$  is the speed of light,  $\nu_m$  is the frequency at the line center,  $T$  is the temperature of the emitter and  $m$  its mass. The FWHM is obtained as  $\Delta\nu_D = 2\sqrt{\ln 2} \nu_D$ . The line profile function  $\Gamma(\nu_D)$  is then also Gaussian and reads:

$$\Gamma(\nu_D) = \frac{1}{\sqrt{\pi\nu_D}} \exp\left[-\left(\frac{\nu}{\nu_D}\right)^2\right], \quad (3.10)$$

where  $\nu$  is the frequency detuning.

*Pressure broadening* represents a set of mechanisms that are contributing to the broadening of a spectral line; literature sources provide a somewhat non-rigorous description of involved mechanisms. Griem [55] introduces the *pressure broadening* as a Stark effect caused by perturbers<sup>7</sup> in the proximity of the emitter. On the other hand, Fujimoto [57] describes the *pressure broadening* mechanism as a collision of plasma particles with emitters. During this collision, electric field is induced at the emitter resulting in a frequency shift.

<sup>7</sup>Electrons and ions nearby the emitting species are, in this case, collectively considered as perturbers. They induce the so-called linear Stark effect. Their collective fields represented with plasma waves are associated with the quadratic Stark effect.

*Pressure broadening* results in the Lorentz line profile  $\Gamma(\nu_L)$  as follows [55]:

$$\Gamma(\nu_L) = \frac{\nu_L / \pi}{\nu_L^2 + (\nu - d)^2}, \quad (3.11)$$

where  $\nu_L$  is the half width at half maximum (HWHM) and shift  $d$ . Finally, *pressure broadening* is a dominant mechanism and, thus, is typically used in order to estimate the electron density of a LIP. [56]

*Instrumental broadening* reflects the diffraction of collected plasma radiation on optomechanical parts of the instrument itself. It may be estimated by using a certified light source with known emission spectrum. However, in LIBS, usually it is not considered as a contributing broadening mechanisms. More concise information on the investigation of broadening mechanisms from the perspective of LIBS is given by Hahn and Omenetto [11]. The broadening mechanisms are divided according to a criterion whether or not the line profile is homogeneously broadened.[56] *Natural broadening* is always homogeneous; but *Doppler* and *pressure broadening* might also be inhomogeneous depending on the duration of disturbance of energy levels.

Naturally, the broadening mechanisms occur concurrently and their contribution to the overall broadening is combined. If the individual broadening mechanisms are statistically independent then their combination is given as a convolution. The convolution of Gauss and Lorentz line profiles is the Voigt profile that reads as [56]:

$$\Gamma(\nu_V) = \Gamma(\nu_D)_m \delta(a, \nu_r), \quad (3.12)$$

where  $\Gamma(\nu_D)_m$  is the maximum Doppler broadening given by relation  $\frac{2\sqrt{\ln 2}}{\nu_D \sqrt{\pi}}$ . The quantity  $\delta(a, \nu_r)$  is the Voigt integral expressed as:

$$\delta(a, \nu_r) = \frac{a}{\pi} \int_{-\infty}^{+\infty} \frac{\exp(-y^2)}{a^2 + (\nu_r - y)^2} dy, \quad (3.13)$$

where  $y$  is an integration variable.  $a$  is the damping parameter and gives the ratio of Lorentz and Doppler half-widths:

$$a = \sqrt{\ln 2} \frac{\nu_L}{\nu_D}, \quad (3.14)$$

and  $\nu_r$  is the relative frequency is respect to  $\nu_D$ :

$$\nu_r = \frac{2\sqrt{\ln 2}(\nu - \nu_m)}{\nu_D}. \quad (3.15)$$

Discussed line profiles are depicted in Figure 3.2.

### 3.3 Thermodynamic properties of Laser-Induced Plasma

Estimation of plasma properties, electron density and plasma temperature, is a standard application of spectroscopy.[55] In each case, many alternative approaches exist; their utilization often demands the knowledge of the other parameters. This leads to iterative estimation of electron density and temperature and consecutively improving their determination. However, it is more practical to compare obtained and synthetic spectrum. In this case, synthetic spectrum is built based on chosen thermodynamic parameters until a satisfactory fit is reached.

Overall, it is recommended to use multiple approaches and compare obtained values.

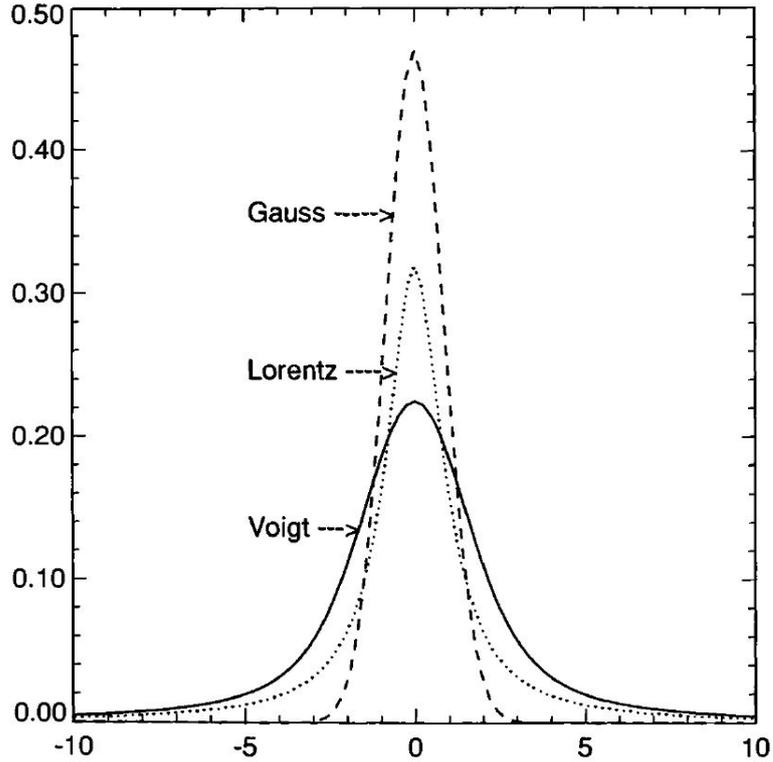


Figure 3.2: Normalized Gauss and Lorentz profiles. Depicted together with the Voigt profile as a product of the convolution of former two. The  $x$ -axis denotes frequency,  $y$ -axis denotes line profile as a function of frequency. Obtained from [55].

**Number density** It is necessary to obtain at least one spectral feature (spectral line width, absolute continuum intensity, absolute or relative line intensity) in order to determine the electron number density of a LIP.[55]

As the first options, the *Stark effect* is beneficially utilized to estimate the electron number density of a LIP. In this case, the *pressure broadening* is dominating mechanism in the line profile broadening or its contribution to the overall broadening is estimated; for instance, by means of deconvolution. As Hahn and Omenetto [11] highlighted, the *linear Stark effect* is reflected in the broadening of hydrogen lines and the *quadratic Stark effect* is considered for non-hydrogen spectral responses. They have also noted that all hydrogen lines have already been used to solve this issue; *e.g.*  $H_\alpha = 656.285$  nm,  $H_\beta = 486.133$  nm,  $H_\gamma = 434.047$  nm.

*Stark broadening* of a spectral line given as FWHM is following[9]:

$$\Delta\lambda_{Stark} = W_{FWHM}\left(\frac{n_e}{10^{16}}\right), \quad (3.16)$$

where  $W_{FWHM}$  is the electron impact half-width and  $n_e$  is the electron number density. This measure is temperature dependent and is tabulated in the book of Griem [54].

The second possible way of electron number density estimation is the *Saha-Boltzmann method*. This method demands the LIP to be in an LTE.[8] When this condition is met then the electron number density is obtained from the ratio of two spectral lines of the same element with different ionization stages.

**Temperature** As in the case of electron number density, there are alternative approaches to the estimation of plasma temperature. It is, however, necessary to distinguish between kinetic temperatures of electrons  $T_e$ , ions  $T_i$  and atoms  $T_a$ , while these temperatures may essentially differ.[55]

Assuming that the condition for LTE is satisfied, the plasma temperature can be estimated from the spectral lines' intensities of the same element and ionization stage. There is only the need to choose spectral lines close to each other and with the biggest possible energy difference of upper levels. The plasma temperature is then calculated as [8]:

$$T = \frac{E_i - E_m}{k \ln \frac{I_{nm} g_j A_{ji}}{I_{ji} g_n A_{nm}}}, \quad (3.17)$$

where  $\lambda_{ji}$  and  $\lambda_{nm}$  are two lines of the same species with significantly differing upper energy levels,  $E_j \neq E_n$ .

The *Boltzmann plot method* is the most popular approach in the determination of plasma temperature. This method builds on the Boltzmann distribution describing the population of individual energy levels of an element. The equation 3.5 may be further altered to the so-called Boltzmann plot equation that reads as follows [8]:

$$\ln \frac{I_{ji}}{g_j A_{ji}} = \ln \left( \frac{N_a^z}{u_a^z(T)} \right) - \frac{E_j}{kT}. \quad (3.18)$$

When plotting the left-hand side of this equation as a function of  $E_j$ , the slope has the value of  $-\frac{1}{kT}$ .

More advanced approach relies upon the so-called *Saha-Boltzmann distribution* that relates ionic and atomic populations of the same species. As in the former case, the distribution relation can be altered to the form of a linear regression as follows[8]:

$$\ln \left( \frac{I_{ji}^{\text{II}} A_{nm}^{\text{I}} g_n^{\text{I}}}{I_{nm}^{\text{I}} A_{ji}^{\text{II}} g_j^{\text{II}}} \right) = \ln \left( \frac{2(2\pi m_e kT)^{3/2}}{n_e h^3} \right) - \frac{(E_{ion} - \Delta E_{ion} + E_j^{\text{II}} - E_n^{\text{I}})}{kT}, \quad (3.19)$$

where superscripts I and II relate parameters to atomic and ionic species,  $E_{ion}$  is the first ionization potential,  $\Delta E_{ion}$  is the lowering correction parameter and  $m_e$  is the electron mass.

## Chapter 4

# Processing the emission signal

This chapter is the core of my thesis and therefore I would like to put it into a wider context including future perspectives and challenges in upcoming years.

In this chapter I discuss basic statistical apparatus describing obtained data (signal), means of data preprocessing and its impact on results (figures of merit). I make conclusions and recommendations that fit the LIBS experiment; however, this is not limiting factor for their implementation in processing of data obtained using other analytical techniques. This topic is also discussed in my articles, in the Appendix A of this thesis.

It has to be kept in mind that any data handling influences the structure of the data (its eigenstructure) and relationships between data points [65].

Obtained data, *i.e.* individual spectra, are detected in their raw form which is then a subject of further data handling and processing. Obviously, in raw spectra, the intensity is a function of the wavelength. The manipulation with detected spectra follows algorithms that are not strictly given and thus significantly differ between application or research groups, or even both combined. Thus, there is no golden standard in data processing that could be universally applied to any data obtained from various applications and LIBS systems. Despite that, judicious data preprocessing can lead to improved analytical figures of merit and is, therefore, important part of the complete analytical process.

### 4.1 Data Preprocessing

The list of preprocessing steps together with suggestions for their use is given namely in review articles [65, 66]. Motto-Ros *et al.* [67] discusses the influence of individual data processing steps and resulting nuances in final analytical performance that they induce. In this section I suggest a set of preprocessing steps that are usually used by the LIBS community and also Laser spectroscopy group at BUT. This algorithm is based on literature research and my personal experience. I will consider raw data as data provided directly from the commercial software of the spectrometer. Thus, I will not speak about a potential use of echellograms and customized way of readout of the echelle spectrometer that I have proposed in my recent study [68].

#### 4.1.1 Background subtraction

The first step in data preprocessing that I suggest is to provide subtraction of spectral background from each spectrum individually. Background intensity related to Bremsstrahlung radiation was introduced in chapter 3.1.

In the quantitative analysis, the background intensity of the close proximity to the analytical spectral line is subtracted from the intensity of this spectral line. Another challenge is rising from the fact that the background level is not constant in the whole spectral range. The background level is typically numerically estimated using mathematical algorithms. Moreover, the background is covered with analytically valuable spectral lines. Thus, the background subtraction has to be treated with care. In a systematic quantification study [67], the background subtraction was considered as a critical step with significant impact on final results.

There exist several available algorithms in the literature that can be used for the background subtraction purposes. Gornushkin *et al.* [69] used automatic background correction based on polynomial fitting. The algorithm introduced by Galloway is based on iterative wavelet transform [70]. Yaroshchuk and Eberhardt [71] applied moving minimum and Gaussian smoothing. The analytical figures of merit were improved when data were preprocessed using any of listed algorithms, this evidence may be found also in other studies [67, 72].

The comparison of listed algorithms was done in our recent work of Képeš *et al.* [73] by the means of LOD estimated in univariate calibration. Selected algorithms were based either on manual fitting, polynomial fitting of local minima (Gornushkin's alg.; [69]), iterative wavelet denoising (Galloway's alg.; [70]), or moving minimum with Gaussian smoothing (Yaroshchuk's alg.; [71]), see Figure 4.1. As a conclusion to this work, it was found that the difference between selected background subtraction algorithm is minimal and any background subtraction leads to significant improvement in observed figures of merit.

Note that the rise in the background intensity may be caused also by imperfectly resolved light. The detection of uranium might serve as perfect example. A typical spectrum of uranium is characteristic for dense spectral lines. Due to lower resolution of conventional spectrometers, the overload of non-resolved spectral lines leads increased background intensity [74]. This phenomenon was also described in one of my co-authored publication [75]. Background subtraction may be accompanied with the smoothing process (Savitzky-Golay algorithm, moving average, and other filters) in order to reduce the background noise.

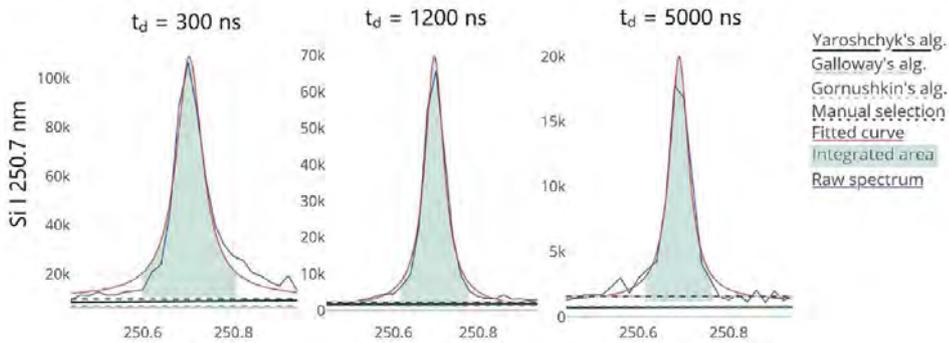


Figure 4.1: Representative emission of Si I 250.7 nm line at 3 various gate delays. Raw emission lines are shown including the estimated baselines by the four implemented algorithms. The grey area represents the integrated intensity of the emission lines. The red line shows the fitted curve, the blue line the measured spectrum. The X axes represent wavelengths in nm, the Y axes intensity in counts. Obtained from [73].

### 4.1.2 Signal standardization

As it was already stated in chapter 3.1, the signal obtained from pulse-to-pulse LIBS analysis is not stable and fluctuates around a certain value. Of course, not only from the statistical point of view, chapter 4.2, the fluctuation is naturally present in experimentally obtained data. Its origin is bound to the laser-matter interaction and in turn to the matrix effect. Therefore, any fluctuations in the stability of a laser source, sample local inhomogeneities or fluctuation of other parameters involved in the laser-ablation process significantly influence the consequently formed LIP and its thermodynamic parameters [2, 10]. Tognoni and Cristoforetti [76] delivered a detailed review on the origin of noise and fluctuations in LIBS analysis and suggested means of their mitigation.

Therefore, I suggest to perform a signal standardization (often inaccurately called normalization)<sup>1</sup> as a second step in the signal preprocessing. The signal standardization discussed in this chapter may be divided into two partial, yet connected issues:

- first, mitigating the fluctuation within a set of measurements obtained from one (homogeneous) sample,
- second, mitigating the impact of matrix effect on the analysis of a set of samples.

Both issues may be demonstrated on the calibration curve and related figures of merit, standard deviation and linearity respectively, see also chapter 4.3 and 4.4. In any case, the ablation process leads to the plasma with different properties (temperature and electron density) which in turn influences the irradiated intensity at individual wavelengths.<sup>2</sup>

Several possible ways to mitigate the influence of a matrix effect and natural fluctuations of the analytical system are frequently referenced in the literature. Hahn and Omenetto [2] suggested to use the ablated mass, plasma temperature and electron density. Their suggestion is based on the essential relation of laser-ablation and LIP formation process with the laser-matter interaction parameters and matrix effect. Consequent estimation of mass (volume) ablated by impact of each laser pulse is tedious and hardly applicable for measurements involving hundreds and thousands of repetitions. Calculation of plasma temperature and electron density is considered as inaccurate (up to 50% RSD as referenced by Miziolek *et al.* [8]). Therefore, other approaches are being implemented in LIBS application. I mention the example of the calibration curve because the signal standardization is of paramount interest in the quantitative analysis, but may be used also in other applications.

The most frequent approach is to use internal standard, spectral line intensity of a matrix element [2, 8–10]. Most dominant matrix element is used for the purposes of internal standardization. There are 9 criteria for the selection of line pairs as identified by Barnett *et al.* [78], including comparable ionization energies and atomic weights of elements and similar excitation energies and non-self-absorbed intensities of spectral lines. Body and Chadwick [79] proposed the ratio-ing of spectral line intensities by the total plasma emission for signal standardization. It is also worth mentioning the approach using the acoustic signal [80, 81].

Zorov *et al.* [82] compared the impact of acoustic signal, total emissivity, electric current and Mie scattering on the signal standardization. In another study, Castro and Pereira-Filho [83] evaluated the influence of twelve standardization strategies on uni- and multivariate quantification. In my former article [84], I have compared seven standardization

---

<sup>1</sup>Note that the terminology related to internal and external standardization differs and may be found contradictory, for instance referring to acoustic signal.

<sup>2</sup>Such phenomenon was beneficially used in my publication [77], in which the state of magnesium corrosion and related changes in the hardness of ablated matrix were correlated with LIP temperature and ratios of atomic and ionic magnesium lines.

procedures and their impact on classification of three sample sets with different matrices (steel, Al alloys, sedimentary ores). Obtained results indicated that there is no universal approach and the most appropriate standardization procedure has to be selected individually in each application. Despite indefinite standardization algorithm available in the LIBS community, the standardization of data prior to further data analysis (uni- and multivariate quantification or classification) is advised [66].

### 4.1.3 Outliers filtering

After the signal standardization I suggest to filter any potential outliers left in the data set.<sup>3</sup> Spectra are treated individually in background subtraction and signal standardization steps, but in this step spectra obtained on each sample are treated in respect to each other. In theory, all spectra obtained from the analysis of a heterogeneous sample should be the same; any spectral feature or characteristic of LIP and ablation process should correlate across the data set. That is, however, not possible in reality when the analysis is burdened with fluctuations coming from various sources as discussed in chapter 3.1.

The signal fluctuation (standard deviation of a spectral feature) can be significantly mitigated using data processing algorithms. However, certain data points can strongly deviate from rest of the data set; such data point is called as an outlier. At this step, it is necessary to find a characteristic describing the data set that is robust and able to precisely detect any outliers. Surely, the threshold level is still given by the operator/scientist.

Carranza and Hahn [85] rejected more than 60% of spectra based on the intensity values of Fe II spectral lines. Lazic *et al.* [86] estimated the maximum intensity of a selected spectral line and filtered up to 50% of outlying spectra. Sirven *et al.* [87] detected outliers in the space of principal components, the outliers were considered as the most distant points from the origin. Gornushkin *et al.* [88] proposed to use a simple linear correlation (Pearson's correlation coefficient) and to determine the level of similarity between two spectra (vectors of intensity values). Spectra, showing the lowest correlation to their counterparts, are discarded. Yaroshchuk *et al.* [89] determined the total emissivity of each spectrum (sum of intensities over the whole detected spectral range), then 10% of spectra were considered as outliers.

Mermet [90] defines the outliers according to their statistical distribution. In the case that the data points follow Gaussian distribution it is possible to mark all points outside the interval given by 3 times the standard deviation as outliers; discarding 0.1% of all data points from the population. There are several tests that can be used to ascertain outliers, for instance Dixon's, Grubb's and Hampel's tests [7].

In my recent article [91], I have compared three characteristics used for outliers filtering: i) total emissivity, ii) distance in the PC space, and iii) mutual linear correlation. Selected approaches were compared by the means of classification performance improvement of steel samples. First, it was found that each approach marks different data as outliers. Consequent filtering creates additional nuances in the final results. Second, the total emissivity approach led to the best improvement of classification accuracy. Thus, it may be stated that total emissivity is a powerful spectral characteristic applicable for both, signal standardization and outliers filtering.

---

<sup>3</sup>Of course, note that outliers filtering cannot be applied in the case of heterogeneous samples; for instance in the elemental mapping.

#### 4.1.4 Line selection and intensity estimation

As a motivation for this section, in an ideal case the elemental composition of a sample should be fully qualitatively and quantitatively estimated. Let's consider that the line selection and intensity estimation is the final step of the data preprocessing. Univariate and multivariate quantitative analyses are subject of discussion in chapters 4.3.1 and 4.4.3

Basically, the qualitative analysis (identification of each spectral line) should be the beginning of each LIBS analysis. It is, however, not necessary to identify all lines within a spectrum; each element emits characteristic radiation on several wavelengths. Some elements (Fe, Ti, U, *etc.*) emit characteristic radiation on hundreds of wavelengths. Identification of all spectral lines is thus tedious and irrelevant.

Let's consider a set of emitters (atoms or ions of the same analyte in a LIP plume) radiating characteristic signal on various wavelengths. Then, obtained spectral lines fully characterize the analyte of interest (qualitative analysis of the sample). Due to their origin from the same source (set of emitters), estimated intensities of those spectral lines and their pulse-to-pulse fluctuation has to be highly correlated. This phenomenon was described in detail, for instance, in my recent article [46]. This approach can be adapted in the decision making process and increase its robustness; wrongly assigned spectral lines will show lower correlation to other lines of the analyte.

A trained spectroscopist is able to swiftly localize and identify most prominent spectral lines while using elemental databases [92, 93] to tackle this task. Unfortunately, there is still no algorithm for spectral line localization and identification used in daily routine. Amato *et al.* [94] proposed an algorithm for unassisted identification of elements in a LIBS spectrum. Localization of the spectral line is, however, a challenging task complicated namely by spectral interferences, potentially low intensities of analytical lines and high background noise. Speaking about tools for localization of the most prominent spectral lines, neat approach is to investigate loadings of the principal component analysis, see chapter 4.4.2, or use variables importance in projection (VIP) as reviewed in my recent publication [65] and mentioned further in the chapter 4.4.2.

A spectrum has unnecessarily high dimension (number of variables) in its raw form. Taking advantage of the redundancy of spectral lines per element, it is necessary to identify at least one line of an analyte to decide whether it is present in the sample. The limit of detection dictates a threshold intensity that supports the presence of an analyte in the sample with statistical significance. The absence of spectral lines of an analyte suggests that

- selected spectral range is not adequate,
- the method is not sensitive enough (detected intensity is under the limit of detection),
- experimental apparatus is not suitable<sup>4</sup> or
- any combination of aforementioned options.

At this point of discussion it is appropriate to list a set of recommendations that are advised for accurate assignment of spectral lines to corresponding element:

---

<sup>4</sup>For instance, the detection of lines below 200 nm demands the absence of oxygen in the atmosphere in the interaction chamber.

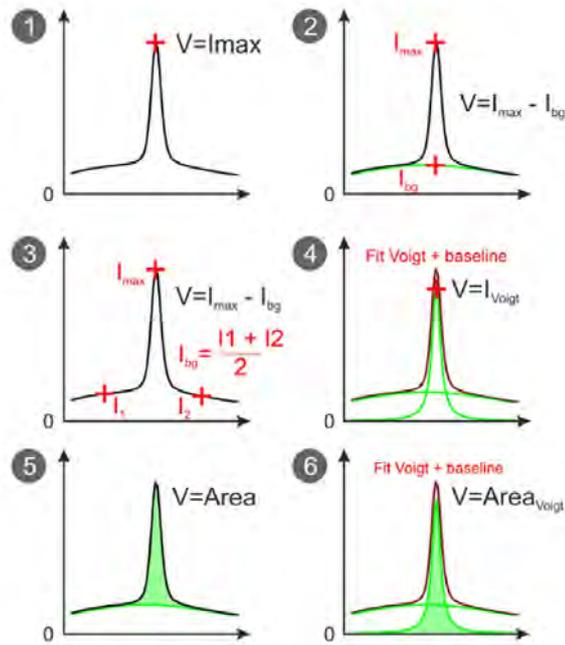


Figure 4.2: The graphical representation of sources of potential bias in the estimation of analytical signal. Obtained from [67].

- due to a relatively low temperature of LIP (see chapter 3.3), only the emission of neutral and singly ionized species may be observed,
- the Einstein coefficient (the transition strength) should be over  $10^6 \text{ s}^{-1}$ ,
- upper level energy should not reach more than several eV (which in turn might be related to the temperature of tens of thousands of K), and
- the presence of other spectral lines of the same element should be cross-checked.

Once the spectral line is assigned, another step is to estimate its intensity. The most simple and often used approach is to consider the maximum peak intensity for further analysis; finding a maximum intensity value in a wavelength range given by the position of analytical line. This approach is convenient but overlooks the shape of the spectral line. Another possibility is to sum up the intensity in the selected wavelength range or to fit the analytical spectral line with a mathematical function (Gaussian, Lorentzian or Voigt; see chapter 3.2). Fitting a line with a mathematical function enables to estimate other parameters (FWHM) and to judge on potential saturation of analytical line profile. Such data preprocessing step is crucial and significantly influences the resulting figures of merit, as concluded by Motto-Ros *et al.* [67] and schematically depicted in Figure 4.2.

It has to be kept in mind that only well selected spectral lines should be used in further spectroscopic analysis<sup>5</sup> (namely for the estimation of LIP temperature and electron density and quantitative analysis). The line should be free from any potential self-absorption (not to use a resonant line) or spectral interference.

<sup>5</sup>Speaking strictly about the univariate analysis. Unfortunately, recent trend in the LIBS analysis tends to deviate from a decent spectroscopic practice and lacks any solid basics discussed in this chapter. In chemometrics, the selection of spectral lines and handling of the data sets is so far accepted more freely.

Finally, finding the best possible wavelength range is of crucial importance from the sensitivity point of view. The sensitivity is typically lower for spectrometers (for instance in the echelle configuration) with broader wavelength range; using such spectrometers may be counter-effective and no signal from the element of interest is detected. Thus, it is advisable to use Czerny-Turner spectrometers with shorter wavelength ranges. In this case, prior knowledge of the sample and the analyte to be analyzed is absolutely necessary. I have tackled this issue in one of my recent articles [95].

#### 4.1.5 Organization of a data set

After the preprocessing, individual spectra are organized into a data matrix  $\mathbf{X}$  with  $n \times p$  dimensions; where  $n$  is the total number of obtained spectra and  $p$  is the number of variables/wavelengths organized as columns. When it comes to the clustering, classification, or quantification; spectra from each sample are simply put together to a data matrix  $\mathbf{X}$ . In the case of latter two, the rows of the data matrix are assigned with the class membership or the content of an analyte, respectively. The data set should be split before the classification or quantification analysis. This step divides the original data into subsets (model, validation, test) [66]. The data subsets have to be drawn cautiously in order to describe the variance within the original data with comparable power [96].

In the case of mapping the so called data cube<sup>6</sup> is created; rows are marked with a  $x, y$  position in the map and represent pixels in the image, see Figure 4.3. Data sets are reshaped to a form fitting the consecutive analysis namely by multivariate algorithms.

As another step, column-wise mean-centering and scaling are of interest. First, the mean-centering step in which the mean of each variable is subtracted and the center of mass is shifted to the beginning of the coordinate system. Second, scaling variables of different magnitudes may form a data set with their uniform impact on the model. When both steps are applied it turns the distribution of each variable to have mean equal to zero and its standard deviation equal to unity. Thus, those two steps are in the literature collectively called Standard Normal Variate (SNV) and its influence on LIBS data processing has already been investigated [98]. However, this step is not advised for data with significant contribution of variables carrying only noise, *e.g.* raw spectra, [84].

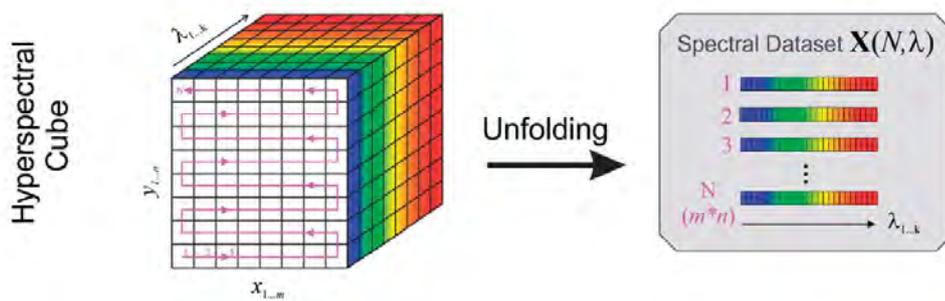


Figure 4.3: Organization of a data set into the hyperspectral cube. Obtained from [97].

<sup>6</sup>Also called a hyperspectral cube, see [97]

## 4.2 Statistics

First of all, it is necessary to adopt basic statistical terminology; for more detailed insight I refer the reader to, for instance, Mark and Workman [1] and Skoog *et al.* [7], which was used as a main source in writing following sections.

In the case of optical emission spectroscopy technique, a single measurement (such as that from a LIBS experiment) is represented by a detected spectrum, also called an object from statistical point of view. Each object is described by a set of parameters, characteristic wavelengths. The set of parameters is unique for each LIBS instrument, which is given namely by available wavelength range and spectral resolution to be detected.

Population is a complete collection of objects fully describing the studied system, in our case emission of LIP, via available parameters. Through the LIBS experiment one obtains a representative sample of objects that should satisfactorily describe the parent population. Henceforth, I will call a representative sample of objects as a data set with  $n$  spectra (objects) and  $p$  characteristic wavelengths (parameters<sup>7</sup>, variables). Moreover, the number of independent variables gives the number of degrees of freedom; independent variables are spectral responses of a LIBS measurement. To complete this list, dependent variables reflect properties of investigated sample and it is possible to construct a direct relationship between them and independent variables (for instance as a curve of growth). Any systematic change in the value of independent variable (intensity at certain spectral wavelength) enables to estimate a change in the dependent variable (content of an analyte).

Each variable fluctuates in a certain interval; any measurement is burdened with an error (systematic or random) giving the span (usually finite) to this interval. The probability of detecting a value from this interval range is non-uniform and is given by the distribution of the parent population. Thus, representative sample (repetitive detection of analytical line intensity) is drawn from its parent population and follows its distribution. Simply, the distribution tells us how probable it is to detect any intensity value in a repeated experiment. In the spectroscopic community, statistics are computed with non-tested assumption that obtained values follow normal distribution. However, this assumption is not instantly valid for all obtained variables and estimated statistical parameters tend to be biased. In following chapters, I will describe this phenomenon in further detail.

### 4.2.1 Normal Distribution

The importance of normal (or Gaussian) distribution, not only in spectroscopy, cannot be emphasized enough. Due to the general knowledge of the normal distribution, I do not give any other references than Mark and Workman [1].

As discussed above, a parent population is usually described by a distribution dictating the probability with which a value of a variable will be obtained. The probability density of the normal distribution is given by following continuous function:

$$f(x|\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad (4.1)$$

where  $\mu$  is the mean of the distribution (and also its median and mode),  $\sigma$  is the standard deviation and  $\sigma^2$  is the variance.

The mean and standard deviation are the fundamental characteristics used to describe the nature of the normal distribution. In the case of the whole population of  $N$  objects, the

---

<sup>7</sup>If the value of a parameter fluctuates then it is called a variable. Thus, I refer to measurement responses only as variables.

mean and variation (its root being the standard deviation) respectively, can be estimated as:

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i, \quad (4.2)$$

and

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2, \quad (4.3)$$

In the case of a representative sample drawn from its parent population (where  $n < N$ ), the mean and unbiased variation (once again, its root being the standard deviation) respectively, can be estimated as:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad (4.4)$$

and

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2. \quad (4.5)$$

Note, normal distribution characteristics are the most important and also the first statistics computed over the data set. Fluctuation of an analytical signal (intensity values detected at given wavelength) in the repeated (pulse-to-pulse) experiment is described by its mean and standard deviation; estimated as statistics for representative sample of a population. Moreover, the stability of the system, the measure of fluctuation, is estimated as the relative standard deviation (*RSD*, or the so called coefficient of variation when multiplied by 100%):

$$RSD = \frac{S}{\bar{X}}. \quad (4.6)$$

This coefficient enables to understand the data from the context of both, mean and standard deviation. There are also other statistical measures used to describe the performance of the system and quality of the data. They are collectively called figures of merit; I discuss them in detail in the chapter 4.5.

#### 4.2.2 Extreme Value Distribution

Typically, normal distribution is instantly considered to describe the fluctuation of spectral variables. This stands, especially, when repetitive (pulse-to-pulse) measurements are taken. However, certain departures from the normal distribution may be encountered; the mean might be shifted, leading to skewness of the probability density function. As Mark and Workman [1] suggested, frequently the distribution becomes log-normal. In the case of LIBS experiment, Michel and Chave [99] investigated the statistics of detected signal and concluded that obtained data significantly deviate from the normal distribution and follow rather the Extreme Value Distribution; Frechet's modification of the extreme value distribution to be precise.

Historically, there are three types of Extreme Value Distributions (Weibull, Gumbel, and Frechet) unified in the so-called Generalized Extreme Value Distribution (GEVD) [100]. GEVD is defined by the following probability density function:

$$G(x) = e^{-[1+\xi\frac{(x-\mu)}{\sigma}]^{\frac{-1}{\xi}}}, \quad (4.7)$$

where  $\mu$  is the location parameter in this case and  $\sigma$  is the scale parameter (analogous to the mean and standard deviation characteristics of the normal distribution), and  $\xi$  is the shape parameter of the GEVD. Values of GEVD parameters can be obtained through the maximum likelihood estimation.

Michel and Chave [99] urge the LIBS community to discontinue using mean and standard deviation (of normal distribution) as statistics characterizing the obtained spectroscopic data. They suggest the utilization of Extreme Value Distribution. However, their manuscript did not get significant impact. There are also other ways how to avoid the discrepancy in the estimation of normal statistics over spectroscopic data. One of them lies in the design of experiment and accumulating spectra; this benefits from the central limit theorem described in detail in the following paragraph. In spite of that, for analog accumulated spectra, the central limit theorem was not found to be valid and data still followed Extreme Value Distribution [99]. This might be due to high fluctuations occurring during the pulse-to-pulse LIBS analysis, which is in agreement with theory of the central limit theorem.

In the light of aforementioned research, we have run an experiment testing the righteousness of implementation of normal distribution over GEVD; see manuscript by Klus *et al.* [101] and Figure 4.4. Presented data did not fully support any of discussed distributions. In fact, the distribution, which each line follows, seems to be based on more complicated background involving quantum properties of spectral lines, dynamics of LIP and its properties, *etc.* The data were also preprocessed by various means of signal standardization; it was shown that standardization influences the data but no pattern was apparent. The relation between the estimated characteristics of both distributions were consistent throughout the experiment; Gaussian distribution in respect to GEVD tends to overestimate the  $\mu$  parameter. Finally, the central limit theorem was proved in our work, showing that accumulation had a positive effect on the distribution of data, which tend to be Gaussian.

Based on our unpublished experimental data and theoretical assumptions, this shift from normal to Extreme Value Distribution is given by non-linear relation between temperature of LIP and intensity of detected signal of analytical line. However, this phenomenon needs further investigation.

Concluding, the implementation of any of the two distributions, normal and GEVD, is justifiable under certain conditions. The experimental design and data processing should be done carefully with the feedback obtained from the statistical hypotheses tests.

### 4.2.3 Central Limit Theorem

The central limit theorem has already been tangentially mentioned. Its derivation and proof can be found in Mark and Workman [1].<sup>8</sup>

Basically, if many objects are added together (summed, accumulated) then the resulting distribution of individual variables tends to be Gaussian. Such assumption holds regardless of the original distributions of the variables. This summation can be done also from the pool of obtained data, as it was shown in our work [101]. The distribution of individual variable values, that are summed together, need not be the same so long as their errors are small. Then the distribution of this product of summation is getting closer to the normal distribution with increasing number of summarized objects. However, two restrictions must be met to the central limit theorem to be valid: i) the variance of any involved variable

---

<sup>8</sup>Historically, this proof was developed by P.S. de Laplace.

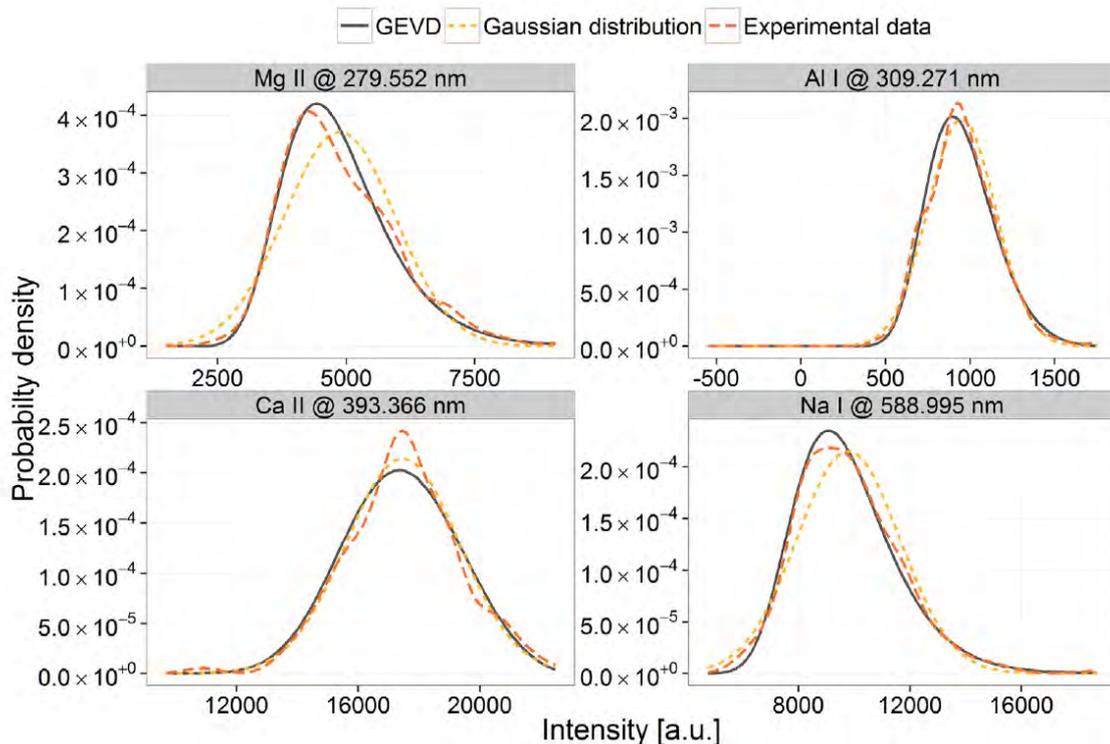


Figure 4.4: The Gaussian distribution fit and General Extreme Value Distribution fit for intensities of four spectral lines. The selection of lines is based on the Kolmogorov-Smirnov test performance. Obtained from [101]

must be finite and ii) the variances of individual variables must correspond to each other, so none of them is dominating over the others.

Gauss had significant impact on the normal distribution itself, he proved that if the data are normally distributed then the fit of data provided by the least squares method gives the maximum likelihood. This fact, in turn, proves that this fit is correct.

#### 4.2.4 Hypothesis testing

From the statistics point of view, obtained data set is drawn from a certain population. Thus, each data set, and its features, should be subjected to hypothesis testing. Hypothesis testing is a standard data processing step in analytical chemistry where data points (*i.e.* observations) are tested in respect to experimental model, for instance whether the mean values of two populations are the same or not.<sup>9</sup> Further decision are then made based on results of hypothesis testing. For more detailed description of this topic refer the reader to the book of Skoog *et al.* [7] and Massart *et al.* [102]. They also advise to use the statistical tests cautiously, especially for data sets with low number of observations (below 50). Moreover, from their perspective, *„the application of good judgement based on broad experience with an analytical method is usually a sounder approach than the blind application of statistical tests“*.

<sup>9</sup>Analysis of variance (ANOVA) is used for the simultaneous tests of multiple populations.[102]

In general, applying statistics on obtained data sets usually demands that the observations follow normal distribution. This is, however, not valid as it is pointed out in the chapter 4.2.1. Kolmogorov-Smirnov test is used to estimate if the set of observations follows normal distribution. Such approach was used also in my co-authored publication [101]. In this case, the null hypothesis is that the set of observations follows normal distribution, the alternative hypothesis is that they are not normal.

In my recently co-authored publication [103], the difference between the control group and group of samples exposed to up-conversion nanoparticles was estimated using the Mann-Whitney U test. Potential outliers were estimated by using the Grubbs test.<sup>10</sup> However, in the LIBS community outliers filtering is provided using other algorithms, see chapter 4.1.3.

## 4.3 Univariate Data Analysis

### 4.3.1 Quantification

The most simple, yet neat and powerful, analysis that might be provided is a correlation of obtained analyte signal (a dependent variable, intensity of an analytical line) with the reference content of the analyte (independent variable) within the sample bulk. This leads us to the univariate calibration, based on a calibration curve, which is established as a relationship between independent and dependent variables using a regression analysis; intensity as a function of analyte content, see Figure 4.5. The concept of calibration is described in detail in following books [7, 102, 104, 105] and articles [90, 106–108] that are used as sources in this chapter and I recommend them for deeper study of this issue.

In analytical chemistry, various calibration approaches may be distinguished [7]. **External calibration** establishes the calibration curve from external standard samples that are independent from the sample of interest. Obviously, the matrix of an investigated sample and standard samples and the level of analyte content have to match. In the case of **internal calibration**, a known content of an analyte is added to all samples (including standards and blanks<sup>11</sup>).

Typically, the regression model describing the calibration curve, Figure 4.5, is a linear function:

$$y = mx + b, \quad (4.8)$$

where  $m$  is the slope of the calibration curve and  $b$  is the intercept of the curve on the y-axis. Data points do not sit on the line but are deviated due to an error of the measurement, see chapter 4.2 for deeper explanation of statistics beyond the typical LIBS analysis. The deviation of a point in respect to the y-axis is called a residual  $SS_i$  and can be estimated as:

$$SS_i = y_i - (mx_i + b). \quad (4.9)$$

In fact, the line interlaced through the data points is calculated using the least-squares method, in which the sum of squares of residuals  $SS_{resid}$  is minimized. The value of  $SS_{resid}$  for  $N$  data points is obtained from the equation:

$$SS_{resid} = \sum_{i=1}^N [y_i - (mx_i + b)]^2 \quad (4.10)$$

<sup>10</sup>There are also other tests, such as the Q test recommended by Skoog *et al.* [7].

<sup>11</sup>An analytical blank is a matrix-matched sample that does not contain the analyte of interest.

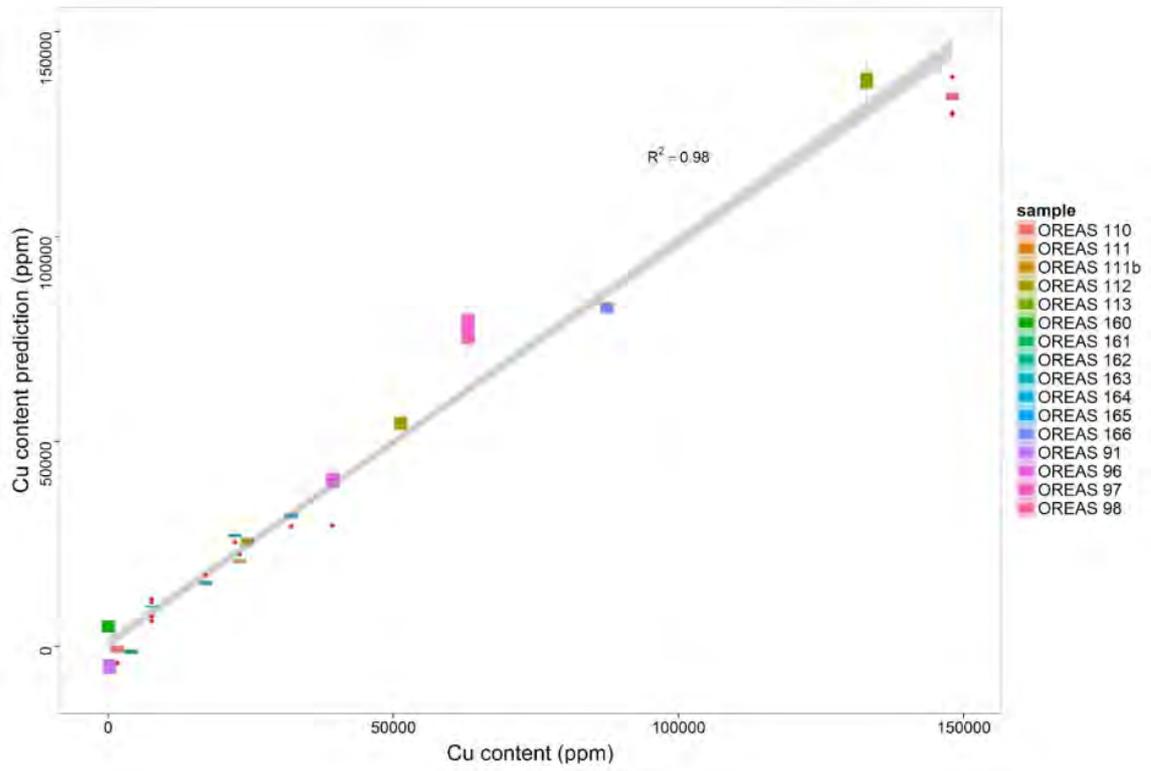


Figure 4.5: An example of a linear calibration curve with a confidence band. Constructed for the purposes of this thesis; unpublished data.

The closer the data points to the least-square line, the smaller the residuals. The total sum of squares  $SS_{tot}$  can be estimated from:

$$SS_{tot} = S_{yy} = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{N} \quad (4.11)$$

Those two measures give another important quantity, the coefficient of determination  $R^2$ ,

$$R^2 = 1 - SS_{resid}/SS_{tot}, \quad (4.12)$$

which measures the fraction of observed variation in  $y$  that is explained by the linear relationship. The closer the  $R^2$  to unity, the better the linear model explains the variation in  $y$ . [7]

In the case of linear relationship between the analytical response and the analyte content, the  $R$  is the Pearson correlation coefficient [107]:

$$R = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}. \quad (4.13)$$

Another figure of merit, the prediction ability  $Q^2$ , gives the ability of the calibration model to predict the content values close to the reference value [66].

The analytical response to variations in the analyte content (content range) follows a linear function; the linear dynamic range. However, this is valid only to a certain point where the linear trend levels off. This point is often called the limit of linearity and its

occurrence depends on several experimental parameters and selection of analytical line [2]. The non-linearity suggests that the response is saturated and the analytical line suffers from the self-absorption. This phenomenon affects namely resonant spectral lines and their use should be considered with care. The impact of selection of analytical line is discussed for instance by Motto-Ros *et al.* [67] and in my publication [109] in which the importance of matrix matched calibration is stressed.

The concept of the calibration curve is in some publications derived as a curve of growth (COG; [110]). However, it has to be stressed that calibration curve can be related to COG only if the atom number density is assumed to be linearly proportional to the analyte content in sample bulk. From the theory [8], the COG describes the analyte intensity as a function of the number density of ground state atoms, those are in turn related to the analyte content.

### 4.3.2 Mapping

LIBS is used for the elemental analysis of solid sample surfaces in the majority of applications. This then enables to investigate heterogeneity of elemental distribution and, in turn, bring additional information to researchers. This approach is collectively called the elemental mapping<sup>12</sup> or imaging<sup>13</sup> and in recent years gains particular attention.

Elemental mapping is summarized in review articles regarding for instance biological [29] and medical applications [32]. Also, the benefits of LIBS are intensively exploited in the mapping of the heterogeneous composition of mineral samples [111]. In my co-/authored articles, I have dealt with mapping of uranium [75] and fluorine in a mineral [46], and Cd-QDs on filtration paper [112] and in leaves of *Lemna minor* L. [113], and distribution of lanthanoid-based upconversion nanoparticles within an *R. sativus* plant [103], see Figure 4.6.

Most recently, high-end LIBS systems benefit from the advancement of instrumentation (lasers, spectrometers and detectors). The most crucial parameter<sup>14</sup> is, especially in the case of mapping, the repetition rate. The repetition rate dramatically influences the duration of an analysis and mapping of large-scale sample surfaces then becomes tedious. Currently, the repetition rate of routinely used LIBS system is 100 Hz [111]. A LIBS system has a potential to reach repetition rate in the order of kHz [68, 114]. Overall, the repetition rate is the main advantage of the LIBS technique in respect to its analytical counterparts. Consider to map the surface area of 1×1 mm with a 10 micron resolution giving 10 000 spots, this will take almost 3 hours with 1 Hz repetition and a bit more than 3 min with 50 Hz<sup>15</sup>.

The elemental imaging throughput is limited by the instrumentation repetition rate on one side and computing power able to process the multivariate data on the other [8]. Hence, dimensionality reduction is of great interest and reflects the need to process a spectrum line-by-line and estimate intensities of most prominent elemental lines. The issue of

---

<sup>12</sup>In the literature, a term “elemental imaging” can be found and in LIBS community are accepted as equivalents. This is, however, in contrast to the usage of terms in Raman spectroscopy where mapping refers to point-by-point analysis of the surface employing system with broad-range spectrometer and imaging to the analysis of the whole surface at once using directly a short-band filter and detector.

<sup>13</sup>Vadillo and Laserna (in the Chapter 6 of [8]) define the term imaging as information provided by an imaging method when the chemical identity is incorporated to spatial dimensions.

<sup>14</sup>Let sensitivity go aside for this time being; it is still the number one parameter and will be discussed as well. Other parameter that is worth mentioning is the spatial resolution.

<sup>15</sup>the current capability of Laser spectroscopy group at CEITEC BUT.

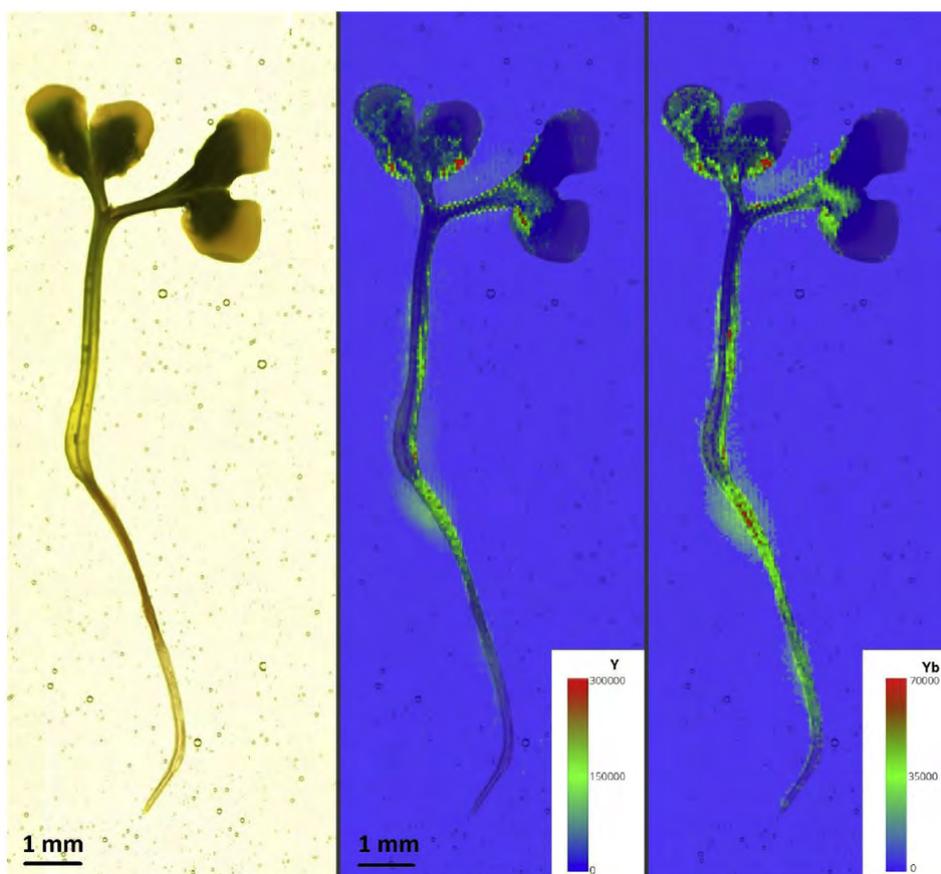


Figure 4.6: Left: a photograph of an *R. sativus* plant exposed to lanthanoid-based up-conversion nanoparticles at the nominal concentration 1000 mg/mL. LIBS maps depicting the distribution of Y I 437.49 nm (Middle) and Yb I 398.79 nm (Right). Obtained from [103].

dimensionality reduction is further discussed in the section related to multivariate mapping, chapter 4.4.5. In the case of multivariate mapping, the dimensionality reduction is provided by the variables downselection and estimation of intensities of several spectral lines. Then only most prominent elements are printed in the maps and their distribution is being investigated.

Albeit, the quantitative analysis of homogeneous samples is a standardized procedure, the analysis provided from a large spot-size has strictly limited usage. It was the improvement in the spatial (lateral) resolution that induced the paradigm shift, not only, in the LIBS analysis using laser-ablation methods. Typically, the spot size of 10 microns is reached as a reasonable trade-off between the ablated mass and analytical sensitivity. Sub-micron resolution can be reached as proved by the group of N. Jakubowski [115] in the field of LA-ICP-MS using a certain methodological trick<sup>16</sup>.

LIBS, as other laser-ablation based techniques, is a quasi-destructive technique. Thus, there might be only one attempt in providing the analysis when, for instance, valuable samples are under investigation. Choice of an adequate spectral region is of key importance and prior knowledge of the sample is necessary [8]. Moreover, the sensitivity of analysis is better when short-band high-throughput Czerny-Turner spectrometers are used [95].

To wrap this section up, the elemental mapping is one of the biggest opportunities of LIBS to gain its position in between other analytical techniques and to exploit its potential. Having a high repetition rate, LIBS may be used to provide overview analysis (2D elemental mapping) after other non-destructive techniques (XRF, optical microscopy, *etc.*).

### 4.3.3 Depth Profiling

The elemental mapping can be extend by consecutive analyses of the same area. Ablation of the surface layer-by-layer provides an insight into the sample bulk. The so-called depth profiling is one of interesting features of laser-ablation based methods. The depth profiling enables to selectively ablate thin layers of a sample surface when properly optimized. This methodological approach enables to study an abrupt interfaces on the sample surface, thin layers and coatings [8]. This is an indisputable advantage in respect to XRF which provides analysis from deeper layers of the sample surface but without any estimation of the depth.

Gaudioso [116] delivered a complex study of depth-profiling including considerations about ablation stoichiometry and fractionation. This assumption is valid since the conditions for laser-matter coupling are steadily changing with increasing depth of the crater. Laser then ablates matter not only from the bottom of the crater but also from its sides. Another influencing factor is the crater itself serving as a spatial container limiting the free spread of a plasma plume. Presented results did not show typical decrease in signal intensity with consecutive laser pulses. Thus, this study shows rather a way in which this analysis should be further taken than any valid conclusions.

Depth profiling can be provided in a single spot or connected with mapping, 3D imaging. Hou *et al.* [117] combined the ability of LIBS to provide depth profiling with the analysis of light elements; the distribution of Li in Li-ion electrolytes was demonstrated using a fs ablation. 50 layers were consecutively ablated and the ablation rate (depth resolution) was 700 nm. Gimenez *et al.* [118] run a feasibility study in 3D imaging of NPs injected in a soft tissue, murine kidney. Distribution of Gd-NPs was detected in a kidney fixed in epoxy and cut into a series of 200 micron slices; the depth resolution was then given not by the

---

<sup>16</sup>When the step size between two ablation spots is smaller than the diameter of an ablation spot.

ablation rate but by the slice thickness. In my co-authored articles, the depth profiling of historical painting models was done [119], see Figure 4.7.

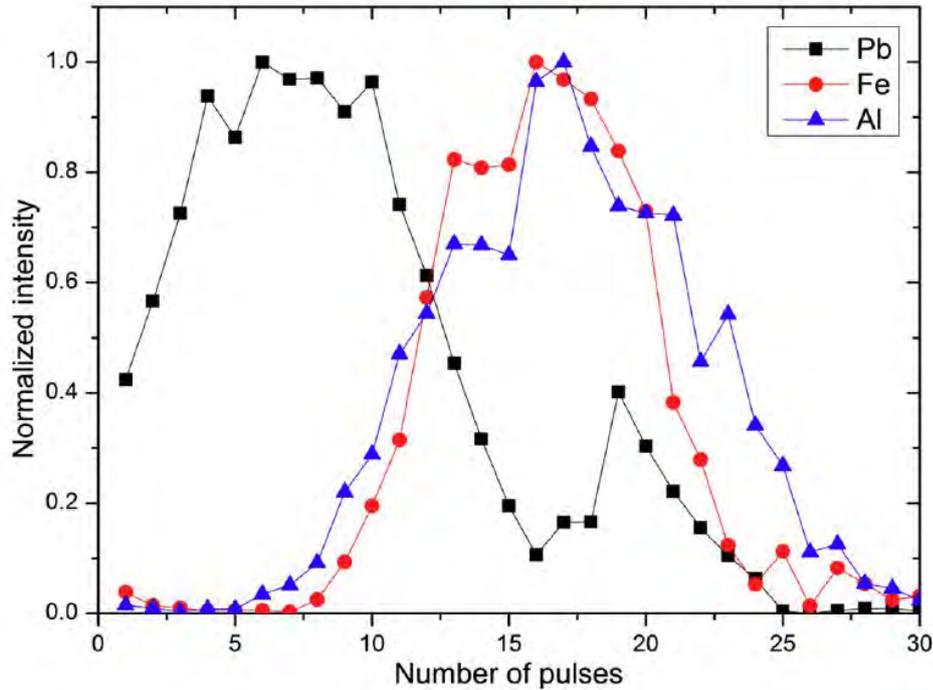


Figure 4.7: LIBS depth profile of the sample n. 242 B, based on the lines of Pb I 405.781 nm (layer 3a) and Fe I 438.354 nm and Al I 396.152 nm (layer 1). Obtained from [119].

## 4.4 Chemometrics

**Spectroscopic *versus* heuristic approach** To quote Hahn and Omenetto [2] once again: “*advanced chemometric algorithms must be used with knowledge of what emission features (e.g., atomic or molecular emission peaks) are providing the associated discrimination*”. This points back to fundamental aspects of LIP and their understanding prior the utilization of any multivariate algorithm.

Despite such clear statement, direct utilization of chemometric algorithms on raw LIBS data (high-dimensional, redundant, noisy spectroscopic data *etc.*) became frequent approach applied by many research groups across the LIBS community. Spectra are not inspected nor at least qualitatively analyzed, but simply introduced to chemometric algorithms in their complete wavelength range without any pretreatment. This is usually advocated with two basic reasons:

- using complete detected wavelength range (especially in the case of broad range spectrometers) might avoid losing any valuable information, and
- using raw spectra is considered as more convenient, non-demanding from the processing point of view (also enabled by more powerful computers and computing clusters).

This seemingly *ignorant* approach is understandable from the perspective of computer science. In computer science, the heuristic approach is usually applied. Data are processed

without the tendency to obtain further knowledge about physical/chemical phenomena involved or the way in which they were obtained.

Both fundamental approaches to data processing, that I have introduced, are polarizing the spectroscopic community and are in a strict contrast. From my personal point of view, those two approaches are extreme cases and further development will include combination of both. We are thus experiencing a significant shift from the classical approach (statistics, physics) which dictates to *estimate* mean value and variance of a variable to the computer science approach which enables to *predict* behaviour of a variable under certain conditions. [120]

Recently, in the LIBS community appeared a momentum driving further data-related research for a unified, standardized approach to data processing that will be generally accepted.<sup>17</sup>

**Assumptions for a sensible model** Building accurate and robust model (regardless of its purpose - quantification or classification) demands its validation. For the validation of a multivariate model, El Haddad *et al.* [66] advises to use leave-one-out or leave-many-out techniques or to divide the data set to model and validation subsets. This step in data processing is absolutely necessary while it mitigates the possibility of constructing an over-fitted model and yielding biased results. Moreover, any data scientist should strive for the highest possible degree of robustness and generalization during the creation of a multivariate model

**Basic literature** There are also many books directly dedicated to chemometric algorithms that are a valuable sources of information [102, 104, 121, 122].<sup>18</sup> More detailed introduction to the chemometrics<sup>19</sup> related to LIBS can be found in books [10, 124] and review articles [2, 65, 66].

There is a great variety of chemometric algorithms that can be implemented in the data processing methodology. Their usage significantly differs and there is no consensus on the universal approach. The methodology in data processing has to be optimized individually to each application and used LIBS system. From the literature research it seems that the only, generally valid, agreement is that non-linear more sophisticated chemometric algorithms outperform the linear ones. However, the use of simple linear algorithms is satisfactory in the majority of cases [65].

**Structure of chemometrics chapter** In this chapter I intend to convey basic approaches of chemometrics [102]; *unsupervised* and *supervised learning*.<sup>20</sup> Unsupervised learning is designed to find patterns within unlabeled data. I am concerned with unsupervised learning in sections 4.4.1, 4.4.2, 4.4.4 and 4.4.5. In contrast to that, supervised learning deals with data labeled by a particular information (*e.g.* content of an analyte in the case of quantification or class membership in the case of classification). Advances in supervised learning are summarized in sections 4.4.3, 4.4.4 and 4.4.6.

---

<sup>17</sup>Discussed problem was one of the reasons to deliver the benchmark data set designed to test various approaches in classification of spectroscopic data by using chemometrics. For more details see Chapter 4.4.6.

<sup>18</sup>To Czech speaking readers I also recommend books of prof. Meloun, for instance [123].

<sup>19</sup>Note that terms *multivariate data analysis* and *machine learning* are often used as equivalents of *chemometrics*. Chemometrics is then accepted to be a subsection of *computer science* in a broader sense.

<sup>20</sup>Note, I have no ambition to discuss *reinforcement learning*. [125]

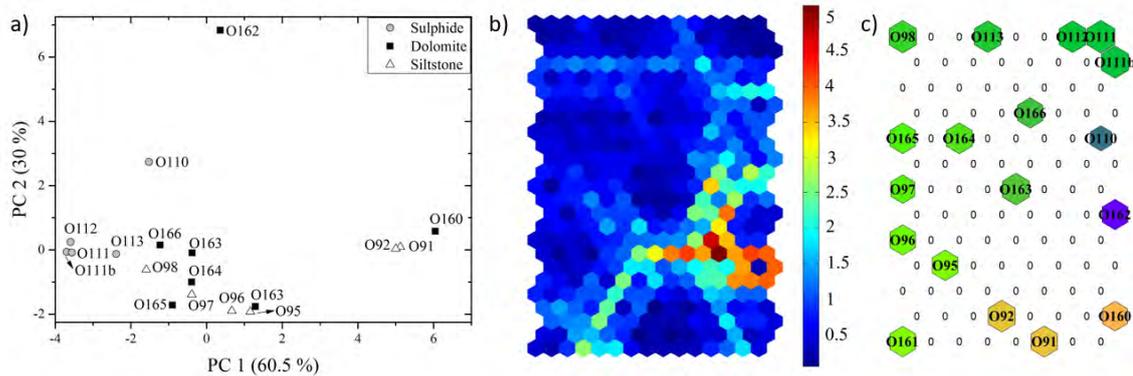


Figure 4.8: The comparison of PCA (a) and SOM (b and c) visualization of selected OREAS ore standards. a) Scores distributed in the space given by the first two principal components; b) distance matrix showing the discrimination of data points (samples) to three basic clusters and c) position of individual samples in the hexagonal matrix of SOM neurons (color is related to distances between SOM neurons and size is related to number of hits). Obtained and adapted from [128].

#### 4.4.1 Visualization

The first and foremost step in the chemometrics is to give an overview of the data and reveal mutual relationships (patterns, trends and outliers) between data points. Such step is called data visualization (pattern recognition or exploratory data analysis).

The reason for visualization is to get fast awareness of the data and the quality of the LIBS analysis prior any further data processing. Visualization may be done also in the case of univariate analysis when the intensities of selected lines are cross-plotted. This approach crashes when multidimensional data (high number of variables, wavelengths) are considered; the number of cross-plots for  $p$  variables reaches  $\frac{p(p-1)}{2}$ .

Due to the high dimension of the problem, it is beneficial to use statistical algorithms that are designed to find the most prominent variables within the data set and in turn to reduce the dimensionality. Clearly, it might be possible to find a new variable (let's call it a latent variable, or even better a principal component) that points in the direction of the biggest variance within the data set. This principal component can easily substitute multiple variables that described the original data set and were strongly correlated.

One of the most spread algorithms that is used for dimensionality reduction and visualization is the Principal Component Analysis (PCA) [65]. There are also other algorithms designed to provide visualization on a lower-dimensional scale, such as Independent Component Analysis (ICA) [126], Sammons's map [91], Kohonen's Self-Organizing Maps (SOM) [127], *etc.* The performance of introduced algorithms is often compared, see Figure 4.8, and simultaneous utilization of various algorithms is advised.

From Figure 4.8 it is obvious that the visualization performance of both algorithms (PCA and SOM) is similar; the presence of outliers (O162 and O110) is clear. I have used only a limited number of data points, thus, their performance was not fully tested. Moreover, Lasue *et al.* [126] compared the performance of PCA and ICA for the visualization of ChemCam data. They have stated a rationale for the utilization of ICA: „From the physics of LIBS plasma generation and emission, we can expect each elemental species to emit light independently from others.“ This statement is, however, too bold and not fully true; in-

dividual elemental species coexist in the same plasma plume and influence each other by series of random collisions, are influenced by the presence of local electric fields, *etc.* Thus, the emission of individual elements is not independent from others.

To briefly conclude, from the literature [65, 66], non-linear algorithms (SOM, Sammon's map) outperform the linear ones (PCA, ICA). Paradoxically, PCA is found to be the most spread algorithm despite having the lowest performance. I will thus concentrate my further discussion around PCA and PCA-related algorithms.

#### 4.4.2 Dimensionality reduction

This chapter is a loose continuation of discussion given in the chapter 4.1.4. However, in this case, manual and experienced selection of spectral lines is not considered in the first step. Described statistical algorithms are judiciously used to guide the hand of a spectroscopist and to ease his/her work.

Outputs derived from the PCA lead to fast, yet robust decision making. First, number of principal components may be judged from the Cattell's scree plot diagram. [65] From the essence of PCA, the first few PCs describe the most valuable variance within the data set and thus the latter PCs can be omitted from further data processing. This forms the first step in the dimensionality reduction. Lower number of PCs is vitally used for further unsupervised (clustering; section 4.4.4 and 4.4.5) and supervised analysis (quantification in section 4.4.3 and classification in section 4.4.4). Such step reduces dimension from thousands of original spectral variables to units of latent variables; leaving behind redundant and non-informative information.

Inspection of loadings plot reveals importance of individual variables forming the original spectra. Weights assigned to individual variables enables to select the most prominent spectral features responsible for the biggest variance within data set and in turn for the discrimination between individual samples. Loadings plot is used to obtain fast qualitative analysis when highlighting the most prominent spectral features. Then, only the spectral lines with the highest weights are qualitatively investigated.

Basically, this step in data processing is still of intense discussion. De Lucia and Gottfried [129] compared the performance of the discriminant analysis algorithm when fed with complete spectra and individual spectral features. It is noteworthy that they have used variable importance in projection (VIP) algorithm to further evaluate impact of individual variables. They have selected 38 spectral features (lines of C, H, N and O together with molecular bands of CN and C<sub>2</sub>) from the complete range of explosive residue spectra. Moreover, they have introduced ratios of spectral line intensities with respect to chemical relationships (stoichiometric ratios) to increase the number of selected spectral variables (177 in total). As a conclusion to their work, down-selecting the data or using the complete range are two extreme cases that might be considered. They have advised to introduce various ratios of spectral features in the down-selected model; however, this demands prior knowledge of the sample and data set.

In other study, Putnam *et al.* [130] proved that variable selection together with construction of well-selected line intensity ratios is of a paramount importance in the classification. They have used a data set comprising of 13 spectral lines and 67 line ratios which outperformed whole spectral range as well as 13 spectral lines alone in terms of classification of bacterial spectra. Surely, the construction of various spectral line ratios is tedious and demands firm understanding of the data set. This approach is potentially applicable only for routine analysis when it does not demand any case-to-case alterations. Kumar Myakalwar

*et al.* [131] explored simple utilization of narrow spectral ranges encompassing spectral lines of selected analytes. They have shown that individual ranges surpassed the performance of the classifier based on full range spectra.<sup>21</sup>

Finally I would like to comment my contribution to this issue. In my publication [95], the attention was given to a moving narrow spectral window scanning across the wide spectral range. Results proved the possibility to neatly assess the most prominent spectral window on which further analysis may be focused. This enables the shift from a low-sensitivity, broad range echelle spectrometer to high-sensitivity, narrow range Czerny-Turner spectrometer. Noteworthy is also my publication about the direct utilization of echellograms without the need for their conversion to spectra. In this way, dimensions are reduced by six orders of magnitude. This work is a continuation and extension of an approach suggested by Larsson *et al.* [132].

### 4.4.3 Multivariate quantification

In this chapter I put emphasis on the construction of a multivariate calibration model. Multivariate model then enables to make predictions of an analyte concentration (independent variable) based on a set of detected spectra (dependent multivariate variables). Thus, this discussion is an extension of chapter 4.3.1. The literature sources are books [102, 104] and articles [65, 66].

Many chemometric algorithms have already been introduced for the purposes of construction of a multivariate model [104], including simple cases of principal component regression (PCR) and partial least squares regression (PLSR). PLSR is the most spread algorithm for multivariate calibration across the LIBS community. [133]. Such algorithms combine the data matrix  $\mathbf{X}$  with a vector of analytes' contents or other sample parameters influencing the spectral response. Because of this construction, the multivariate models are generally accepted to outperform the univariate ones.

Collection of responses from various elements and their spectral lines (ionic and atomic) possesses more complex information about the sample itself and also the parameters of a LIP. In turn, multivariate models mitigate to certain extent the influence of the matrix effect. [66] This phenomenon is graphically depicted in Figure 4.9. The response of spectral lines (in this case was selected Cu I 521.82 nm) to the changes in the analyte content are strongly dependent on the content of major elements, resulting in different trends for individual matrices (*i.e.* minerals). Matrix effect is then significantly suppressed by using PCR and PLSR, both showing similar performances. Once again, judicious selection of spectral features is recommended over blind utilization of complete spectral range.

In the presented use case, the analytical response to the increasing copper content was burdened with the non-linearity resulting from the self-absorption for higher copper contents in respective LIPs. Safi *et al.* [133] recommend cross-checking the performance of a multivariate calibration model by using leave-one-out methods, *etc.* Surely, over-fitting of a multivariate model is frequently found in the literature leading to biased results. This issue is a result of the PLSR algorithm that forces the relation between independent variable and its analytical response to be strictly linear. Unfortunately as we can see, it is not the case of LIBS data.

---

<sup>21</sup>It is noteworthy that all three manuscripts used the partial least square discriminant analysis (PLS-DA) algorithm for classification of organic and inorganic material spectra. Such selection was not intentional, however, proves the capability of LIBS to classify challenging samples with judicious data pretreatment and simple chemometrics algorithm.

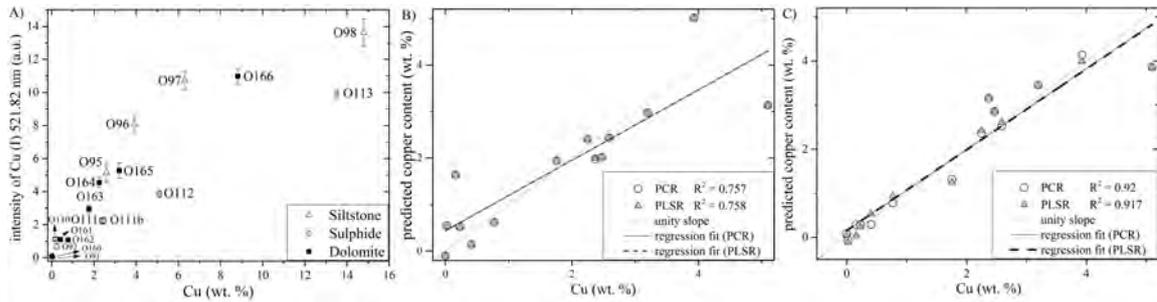


Figure 4.9: Univariate calibration of LIBS system to the response of a copper in OREAS samples (a). Multivariate regression based on Cu lines only (b) and Cu and matrix lines (c). Obtained and adapted from [128].

Recent shift to the utilization of non-linear algorithms, especially neural networks, is of no surprise. Their supreme performance in the prediction has already been demonstrated. Motto-Ros *et al.* [134] developed artificial neural networks (ANN) for quantification of multiple elements in rocks and soils. El Haddad *et al.* [135] used a series of ANN for the quantitative analysis of lead in soils. Neural networks are at the dawn and awaiting full exploitation of their performance capabilities in the analysis of LIBS data in near future.

An interesting use case of chemometric algorithms was suggested by Gottlieb *et al.* [136]. Certain elements suffer from their high ionization energy and low transition probability; *e.g.* such is the case of halogens (*i.e.* fluorine and chlorine), sulfur, *etc.*). Together with their low contents in the sample bulk their detection becomes a challenging task. However, in certain cases (corrosion of concretes as discussed in [136]) the matrix is significantly influenced by the presence of traces of investigated analyte. Moreover, the spectral information might be recovered from the background noise and making the quantitative prediction possible for lower contents.

Lastly, I would like to mention the publication by Yaroshchuk *et al.* [89]. They have compared the performance of chemometric algorithms with respect to combination of various sources of information (multiple spectrometers) in the prediction of iron in iron-bearing ores. Typically, spectra from different sources are glued together; this is unfortunate while the spectral lines intensities and background noise might significantly differ and thus influence the prediction. Serial partial least squares and multi-block partial least squares are alternative regression algorithms that process data set from various sources separately, but in parallel. In my personal opinion, this approach bears potential for its further exploitation in spectroscopic analysis based on data obtained from significantly various sources (such as LIBS and Raman spectroscopy).

#### 4.4.4 Multivariate clustering and classification

To start with a motivation for this chapter, a LIBS spectrum is understood as an elemental fingerprint representing the sample from which it was obtained. This neat consideration leads us to a possible discrimination of individual samples based on their characteristic spectra.

There are two possible cases [102]:

- the classes of modelled data are NOT KNOWN *a priori* then we speak about unsupervised pattern recognition (clustering), or

- the classes of modelled data are KNOWN *a priori* then we switch to supervised pattern recognition (classification or discriminant analysis).

In both cases, matrix effect is beneficially used and practically contributes to the discrimination of individual samples (represented by unique spectral fingerprints) from each other.

**Clustering** Clearly, the inspection of scores cross-plotted in a PC space effortlessly reveals potential similarities between individual samples. They tend to cluster together with respect to characteristic spectral fingerprints. Cluster analysis, thus, forms a simple way of discovering latent relationships between obtained data.

The measure of relationship is, most frequently, estimated as their mutual correlation by using the Pearson correlation coefficient. Gornushkin *et al.* [88] used linear correlation algorithm (Pearson) to classification of glass samples of forensic interest. They point out two problematic factors in the estimation of Pearson's correlation coefficient; insufficient number of observations and influence of non-correlated noise. While the first issue might be solved with improved experimental design, the second one is of more further spectroscopic interest and demands elaborate line selection *a priori*. Linear correlation between data points was also used for outliers filtering, see more in chapter 4.1.3.

Another measure of relationship reflects the distance of scores (data points) in the multidimensional space. Euclidean, Manhattan or Mahalanobis metrics are usually used. [104] The *a priori* dimensionality reduction of data matrix  $\mathbf{X}$  by using PCA is provided in order to speed up the process of inter-score distances estimations. Yueh *et al.* [137] introduced hierarchical cluster analysis based on the distance estimation between spectra. Afterwards, they depicted spectra of tissue samples *via* dendrogram, showing clear separation of individual organs. Andreason *et al.* [138] utilized *k*-means clustering, a non-hierarchical algorithm, for unsupervised classification of geo-samples.

Due to the unsupervised learning, the utilization of aforementioned algorithms is rather limited and the interest of spectroscopic community might be resurrected by the needs of multivariate mapping, see chapter 4.4.5.

**Classification** Many LIBS research groups invested a lot of efforts into the classification of various materials (from geological through steels and alloys to biological samples). LIBS analysis is advantageous due to the characteristic multielemental information provided in the so called fingerprint. Observing simultaneously light elements, majority of metals, halogens and biologically relevant elements (*i.e.* C, H, N, O, P, and S) in one spectrum opens unprecedented classification capabilities.

I have no ambition to cover all existing applications and alternative algorithms used for classification of LIBS spectra; the list is endless. The biggest area of interest seems to be the geology offering magnificent variety of minerals, ores and rocks to be classified. [139, 140] This application is also connected with the ChemCam instrument on the Mars Science Laboratory rover. It is noteworthy that the ChemCam team employs rather simple algorithm - partial least square discriminant analysis (PLS-DA) [141]. LIBS technique excels over its analytical counterparts also due to the possibility of C, O, H and N detection. This enables classification of biological samples (algae [28], bacterial strains [142]), explosives [129], plastics [23], *etc.*).

Most valuable publications bring a comparison of performance capabilities of various classification algorithms. Moncayo *et al.* [143] delivered a comparison of SIMCA, PLS-DA,

LDA, CART, ANN and SVM algorithms<sup>22</sup> by means of classification accuracy, sensitivity, generalization and robustness. Once again, supreme performance of ANN over other algorithms was proved.

**Benchmark data set** For the purposes of the classification contest organized under the auspices of EMSLIBS 2019 organizing committee, we have created a benchmark data set. [144] This data set consists in total of 12 classes, 138 samples, 690 000 spectra. Such data set is unique in the spectroscopic community and until today it was not beaten, the best accuracy reached slightly over 90 %.

The motivation for its construction was to provide a challenging data set to spectroscopic community and evaluate the performance of data processing algorithms used by individual research groups. The diversity in chemometric algorithms used for classification, together with variations in data preprocessing, is enormous and has no unified framework. Obtained results are still under evaluation and will be published in the conference proceedings.

#### 4.4.5 Multivariate mapping

Elemental mapping<sup>23</sup>, as discussed in chapter 4.3.2, is becoming the key benefit of LIBS in respect to its analytical counterparts. [32]

There are two assumptions defining this chapter:

- LIBS is applied to provide elemental mapping of large areas (relatively, in respect to the spot and step sizes; *i.e.* the whole slide images) of heterogeneous samples (most often geological samples),
- simple case of mapping the individual analytes (known *a priori*) is of no interest.

Contemporary state-of-the-art LIBS systems are capable of providing megapixel images (up to 3 hours with 100 Hz repetition rate laser operation). Each spot size (pixel of a map) is represented by a detected spectrum, *i.e.* multivariate information.<sup>24</sup> The number of variables in a raw spectrum is typically from 2048 (detectors mounted on Czerny-Turner spectrometers) to over 40 000 (in the case of echelle spectrometers). In turn, the size of data matrix can reach the magnitude of  $10^{10}$ . Moncayo *et al.* [97] organized data into the so-called hyperspectral cube, see Figure 4.3 which takes inspiration from the field of Hyperspectral imaging as described in detail by Grahn and Geladi [145].

The amount of obtained data are extending the capacity of a typical desktop computer which induces new challenges in their further processing. At this point, new approaches are being implemented in the analysis of large-scale maps. The biggest potential possesses the utilization of chemometric algorithms. Thus, the research focused is at revealing hidden relationships in bulky data sets obtained from mapping of heterogeneous samples. Simple dimensionality reduction and consecutive visualization in the space of newly constructed principal components reveals latent relationships between individual matrices forming the heterogeneous samples. Judicious handling of data then leads to achieve more elaborate description of a sample under investigation.

---

<sup>22</sup>SIMCA - soft independent modelling of class analogies; PLS-DA - partial least square discriminant analysis; LDA - linear discriminant analysis; CART - classification and regression tree; ANN - artificial neural networks; SVM - support vector machines

<sup>23</sup>In the literature, terms *elemental mapping* and *elemental imaging* are treated as equivalents. Note that this is in contrast to the terminology used in Raman spectroscopy.

<sup>24</sup>Therefore I call this chapter *Multivariate mapping*.

To the best of my knowledge, our publication [75] was one of the first pioneering studies in LIBS community concerned about multivariate mapping. In this work, the uranium-bearing mineral was investigated by using the PCA algorithms. Obtained results implied that the presence/absence of uranium was so significant that the first PC was dominated solely by the spectral lines of uranium. This point was proved by the correlation (estimating the Pearson correlation coefficient) between maps of uranium and the first PC. In the consecutive study [127], SOM was applied to the same data set and further relation between zirconium, titanium, uranium and niobium were established based on the inspection of the neuron map and composition of individual neurons.

In our case, the uniqueness of the sample was directly reflected in the composition of PCs. Following publications, however, dealt more with the distribution of data points in newly constructed PC space (with reduced dimensionality). The distribution of data points suggests their mutual relationships as in the case of clustering, chapter 4.4.4. But, individual clusters are not distinctly separated; they are interconnected forming a continuous cluster with several centers having higher density of points. Those centers, if correctly localized, lead directly to the pure spectra of individual materials (*i.e.* matrices) forming the heterogeneous sample.

Separation of such structure was of interest to Romppanen *et al.* [146], who investigated the possibility to use singular value decomposition algorithm to decompose rare earth element ore data set. Obtained data set was then decomposed according to the position in the 3D space given by the first three components; octants of the solid geometry. Gottlieb *et al.* [35] utilized expectation-maximization-clustering algorithm for clustering of concrete samples data showing great results in separating continuous clusters of heterogeneous data. Moncayo *et al.* [97] then *rediscovered* the necessity to cluster individual data points and relate them with corresponding matrices (minerals forming the sample), see Figure 4.10.

To conclude, the necessity to provide robust, yet simple and straight-forward algorithm enabling discrimination of individual clusters stands still unbeaten.

#### 4.4.6 Data libraries

This section is a loose continuation of the chapter 4.1.4. So far, extensive elemental line libraries have been introduced. Their construction yielded from other analytical methods, such as spark discharge, where qualitative elemental analysis of plasma radiation is under investigation. Several elemental line databases may be accessed [92, 93].

Intensive work has already been invested in creation of data libraries relevant to classification purposes based on LIBS spectra. Creation of LIBS data libraries with pure spectra dedicated to various materials is burdened with several drawbacks connected to parameters of laser-matter interaction, laser ablation and matrix effects; see chapter 3.1. All those aspects combined make the task of creation, implementation and data libraries transfer between LIBS systems almost impossible. Clearly speaking, data library is strictly bound the LIBS system on which it has been created.

The Mars Science Laboratory team of Roger Wiens started the creation of data libraries for the ChemCam LIBS system. They have collected numerous geological samples and ablated them with the ChemCam prior its Mars mission. In the case of a stand-off system, the variation in the sample-to-instrument distance is critical. Mezzacappa *et al.* [147] introduced distance corrections to ChemCam spectra. They are also aware of the need to stabilize the fluctuation of LIBS instrument as given in their publication [148].

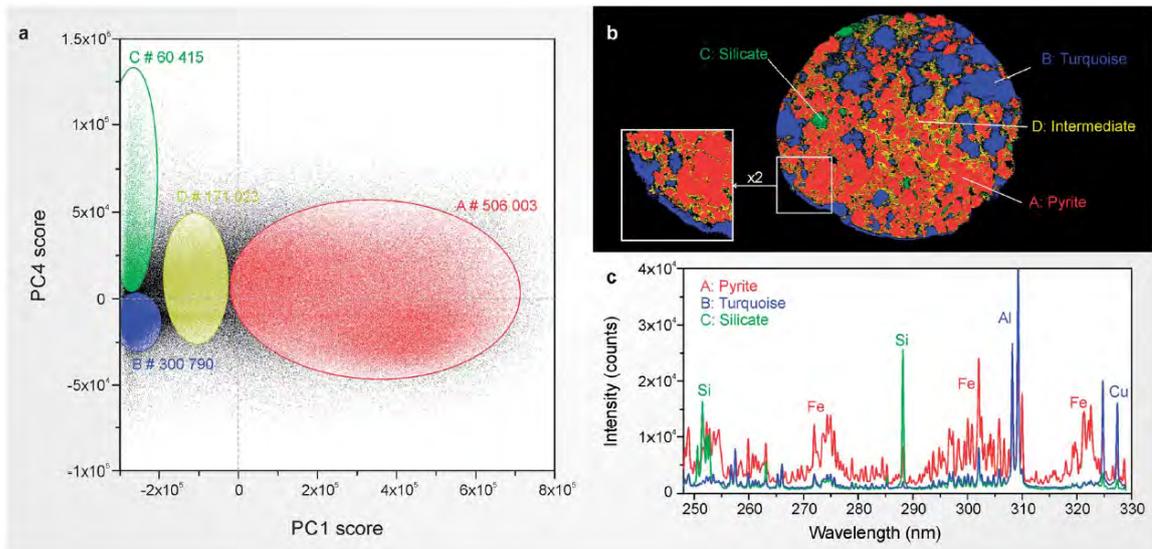


Figure 4.10: a) scatter plot showing four manually selected clusters in the PC space formed by crosplotting PC1 and PC4, b) image showing the distribution of selected clusters and corresponding mineral, and c) representative spectra of selected clusters. Obtained from [97].

Data library transfer is one of my recent research tasks that I have already presented on several conferences; including EMSLIBS 2019 in Brno [149].

## 4.5 Figures of merit

The quality of obtained results may be evaluated using several basic metrics, *i.e.* figures of merit. Each measurement is burdened with errors (random, systematic and gross)<sup>25</sup> leading

<sup>25</sup>Gross errors are outliers and has to be discarded from the data set. Systematic errors has to be corrected in the methodological approach. Random errors reflect fluctuation that is of analytical interest and needs to be evaluated.

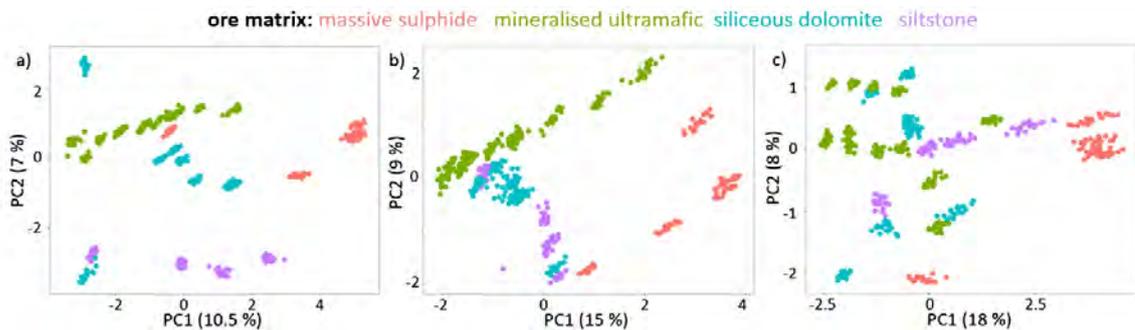


Figure 4.11: Projections of ore samples into PCA space based on the data set of echelle (a) and Czerny-Turner spectra centered at 305 nm (b) and at 405 nm (c). Obtained from [68] and adapted.

to deviations of obtained results from the true (reference)<sup>26</sup> value. The error reflects the variation and fluctuations occurring in an analytical system and caused by an operator. The figures of merit characterize the performance of an analytical instrument and compare it to its counterparts. Herein, I will describe basic figures of merit connected to the quantitative analysis and classification.

The terminology is anchored in the term accuracy of the analytical system. They are sometimes used freely, however, I will refer to this term in relation with the Gold book published by IUPAC<sup>27</sup> [150]. Moreover, I used the book *Fundamental of analytical chemistry* by Skoog *et al.* [7] as valuable source in this chapter.

As a motivation for following concept of figures of merit, an accurate analytical system or method demands high precision and trueness.

### 4.5.1 Quantification

The measure of sensitivity is used to characterize the calibration curve, chapter 4.5. The term sensitivity describes the change of analytical response per unit change of analyte content. In other words, the steeper the slope of the calibration curve, the more sensitive the calibration curve. The sensitivity is constant for the linear calibration line. To give more sense the value of sensitivity, the analytical sensitivity, it is given as a ratio of the calibration slope to the standard deviation. [7]

In the process of quantitative analysis it is necessary to estimate the least content that can be detected or quantified with a certain statistical significance. The limit of detection (LOD) is defined as [90, 106]:

$$LOD = \frac{k \cdot SD_B}{m}, \quad (4.14)$$

where  $SD_B$  is the standard deviation of the blank<sup>28</sup>,  $m$  is the calibration slope and  $k$  is the confidence factor (for  $k = 3$  it corresponds to 98.3% confidence level). The estimated value becomes the limit of quantification (LOQ) if the confidence factor is equal to 10 [90].

It is not a coincidence that the concept of  $LOD$  is closely related to the signal to noise ratio ( $SNR$ )<sup>29</sup>:

$$SNR = \frac{y_i}{SD_B}, \quad (4.15)$$

where  $y_i$  is the signal of the analytical line. Hahn and Omenetto [2] claim that the  $SNR$  and  $SBR$  can be used for the estimation of  $LOD$  and recommends to use following equation<sup>30</sup>:

$$LOD = k \cdot c_0 \frac{SD_B}{y_i}, \quad (4.16)$$

where  $c_0$  is the content of the analyte in the investigated sample. This approach is also called the one point calibration.

---

<sup>26</sup>Typically, ICP-MS/OES after acid digestion is considered as a technique providing reference analysis to LIBS.

<sup>27</sup>The abbreviation IUPAC refers to the International Union of Pure and Applied Chemistry (IUPAC), which is the World authority on chemical nomenclature and terminology.

<sup>28</sup>Background noise in the location of the analytical line or in its vicinity if the analytical blank is not available.

<sup>29</sup> $SNR$ , the standard deviation of the background noise in the vicinity of the analytical line, is often introduced together with the signal to background ratio (SBR), mean value of the background.

<sup>30</sup>Taken from [2] - equation 10b on page 362 and modified.

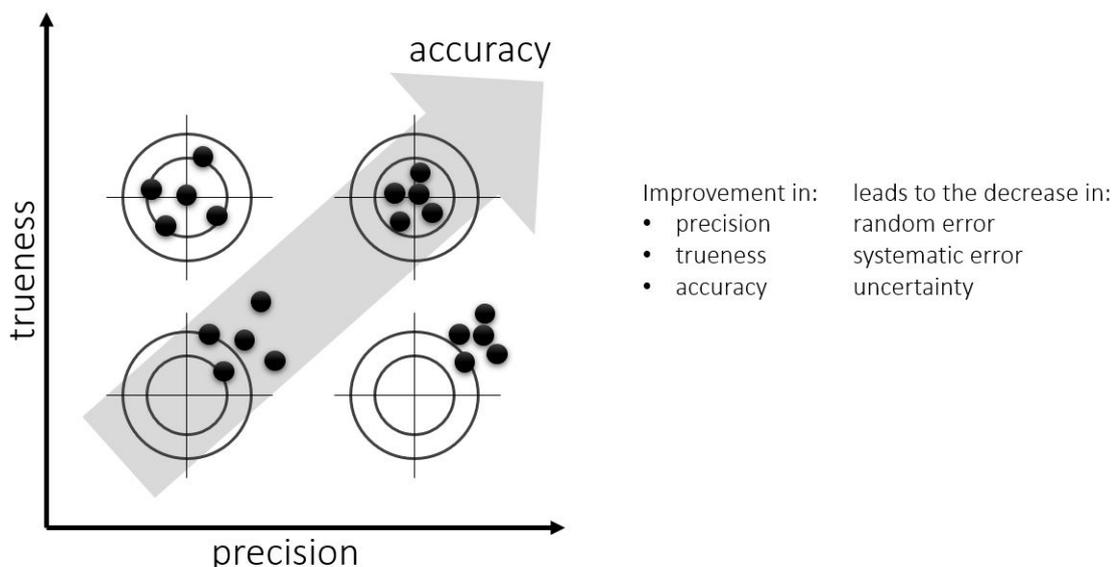


Figure 4.12: The graphical explanation of figures of merit.

I assume that the true, or reference, value is known when the analysis is performed, then the quality of quantitative analysis of an analyte may be characterized via following figures of merit:

- **Accuracy** is the closeness of the measured, observed value to the true value. From its definition, accuracy is defined by the trueness and precision of the measurement.
- **Precision** is related to the variability of results and is affected by random errors. Precision is defined as the closeness of fit between obtained results. It is expressed by the standard deviation of the results and is independent from the true value. High precision leads to low standard deviations, i.e. mitigated random error.
- **Trueness** is the closeness of fit between the mean estimated from observed values to the true value. The trueness might be related to **bias**<sup>31</sup>, which quantifies the systematic error.

The performance of analytical system in a repeated experiment is of interest when the optimized protocol is repeated between laboratories. The ability to repeat (for instance by the same technician using the same analytical system) and reproduce (by different technician using the same or similar analytical system and following the same methodology) the experiment gives repeatability, reproducibility, and robustness. Those measures are quantified by the values of accuracy, precision and trueness. Thus, the accuracy may be depicted as it is given in Figure 4.12.

Now, it is possible to predict the analyte content in the unknown sample once the calibration model is built. It is also possible to compare predicted values to the reference ones using a leave-one-out or leave-many-out (collectively referred to as cross-validation) approach. This in turn enables to estimate other figures of merit [66], such as: root mean

<sup>31</sup>It is noteworthy that I have used bias as a figure of merit in my publication aimed at the estimation of the copper content in the mineral samples of various matrices [109].

square error (*RMSE*):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\bar{x} - x_i)^2}{n}}, \quad (4.17)$$

relative error (*RE*):

$$RE(\%) = \frac{1}{n} \sum_{i=1}^n \frac{|\bar{x} - x_i|}{\bar{x}} \times 100, \quad (4.18)$$

maximum relative error (*MRE*):

$$MRE(\%) = \max\left(\frac{|\bar{x} - x_i|}{\bar{x}} \times 100\right). \quad (4.19)$$

#### 4.5.2 Classification

Classification is another analytical process in which the unknown samples are assigned to different classes based on their characteristics (ratios of wavelength intensities or element content). The quality of classification is assessed using the so called confusion matrix<sup>32</sup>, Table 4.1.

Table 4.1: The confusion matrix for simple case of two classes in binary classification. Obtained from [66] and modified.

confusion matrix		true condition	
		positive	negative
predicted condition	positive	true positive	false positive
	negative	false negative	true negative

To explain the cases described by the confusion matrix: true positive (TP) refers to the case when a sample is correctly classified and similarly for the true negative (TN), if the value is negative but it is predicted as positive then it is called false positive (FP) and vice versa for the false negative (FN) case. The confusion matrix can be constructed also for the case of more than two classes and is, thus, a powerful tool in estimation of the classification performance of the LIBS system and consequent data processing algorithms.

It is possible to determine following figures of merit<sup>33</sup> using the confusion matrix [66]:

- **Overall accuracy** = (TP + TN)/(total population),
- **Sensitivity** = TP/(TP + FN) for a given class,
- **Specificity** = TN/(TN + FP) for a given class.

Those figures of merit are frequently used in the literature related to the multivariate classification [65, 66]. It is worth mentioning that I have used those figures of merit for classification purposes also in my publications [84, 91, 95].

To assign an unknown value as positive or negative (or to assign a sample to any of model classes) it is necessary to define a threshold value. This threshold is called the decision limit and can be varied by the operator. The decision limit, thus, defines the

<sup>32</sup>The confusion matrix is also called 2x2 contingency table [102]

<sup>33</sup>Note that those figures of merit might be expressed in fractions or percentages.

sensitivity and specificity of the method. The Receiver Operating Characteristic (ROC) is a curve constructed by cross-plotting true positive rate (sensitivity) versus false positive rate (1 - specificity). ROC is used to assess the best possible performance (highest possible sensitivity and specificity) of the method for various thresholds of the decision limit. The diagonal line defines the method with a zero predictive value. Thus, the further the ROC curve is from the diagonal, the more predictive it is. [102]

# Conclusion

To conclude my thesis with a single sentence, *obtaining a spectrum is considered to be easy while providing robust and judicious data processing demands interdisciplinary knowledge.*

Continuous improvements in analytical instrumentation directly led to bulky data sets. Their storage, handling and processing is becoming a challenge of contemporary LIBS analysis considering the increasing demand on computing power. Classical approaches were designed to lower amounts of data and therefore they cannot withstand the data overload coming from state-of-the-art analysis. They can find place in well defined applications where, for instance, quantification of an analyte based on matrix-matched model is of interest. Those models break down and become inaccurate when the calibration sample set does not fully cover heterogeneity in sample matrices, jitter and fluctuation of the system is blurring the signal response, *etc.* Moreover, univariate algorithms provide insufficient model power because the matrix effects and non-linear behavior in the signal response are not considered. In spite of that I would like to stress that the use of univariate data processing algorithms still finds its place in contemporary applications of LIBS and deserves attention prior the use of more sophisticated approaches.

Classical, univariate algorithms are being abandoned and substituted with more and more advanced multivariate statistical algorithms, chemometrics. This paradigm shift polarized the LIBS community and induced long-lasting discussion on further development of the *status quo* of data processing. This discussion has been one heck of a seesaw recently.

To mitigate general anxiety related to the black-box-use of chemometrics, I recommend:

- to provide pre-flight optimization of the LIBS system to improve the signal to noise ratios of analytes of interest,
- to collect sufficient number of calibration/model samples,
- to optimize and generalize the multivariate models to avoid over-fitting,
- and namely, to invest efforts in the understanding of data from the perspective of the source (plasma physics and analytical chemistry).

The strictly data-driven approach has been, fortunately, recently enriched with thorough knowledge of laser spectroscopy and laser-induced plasma itself, thus, gaining essential touch with the source of data itself, *i.e.* laser-induced plasma. This, in turn, led to more elaborated and unbiased results that reflected the investigated phenomenon more plausibly. Multivariate data analysis thus became an inevitable part of data processing.

In my thesis I brought a road map navigating from classical chemical analysis to chemometrics and neural networks used in computer science. Let this thesis be a starting point when the goal of the data analysis is to establish robust, not-overfitted model with judiciously selected spectral features based on the firm knowledge of the original sample and

the structure of the data set together with a deeper understanding of plasma physics and analytical chemistry.

## Future perspectives

Having this opportunity I would like to give several remarks on future perspectives and potential development of the LIBS technique and related fields.

**Imaging** of sample surface is of a paramount interest to the LIBS community. Currently, many applications yield from the capability of LIBS to provide multi-elemental maps on a large-scale with micro-resolution, which is its most vital advantage of them all. The development of instrumentation is also enabling the transfer from 2D to 3D. Thus, handling and processing of megapixel hyper-spectral images is crucial task that needs to be solved.

**Hyphenated systems** are getting popular, they enable more complex understanding of a sample under investigation. I have shown several examples, when combination of LIBS with LA-ICP-MS and LIBS with Raman spectroscopy are well known. On the other hand, combination of LIBS with X-Ray Computed Tomography remains still unexplored and possesses great potential for future RD endeavor. The main drawback is in the sample preparation process and experimental approach, when the combination of *bulk* technique (XCT) and *surface* technique (LIBS) demands further considerations.

**Benchmark data set** is a standard in data science. However, it was missing in the LIBS community until 2019, when we have introduced it for the classification contest during EMSLIBS 2019 symposium. Based on the results from our contest, it is clear that accurate classification of challenging data set is possible only when data driven approach is judiciously combined with thorough knowledge of data and their relation to the source of origin. I believe that other groups will follow our example in construction of even more challenging benchmark data sets. Hopefully, this will also induce the discussion for unified algorithm in data processing generally accepted by the LIBS community.

**Data library transfer** has not yet been investigated. From my point of view, this issue is a time bomb waiting for LIBS devices to be fully exploited, especially in the classification regime. It was found that any minor fluctuations in the LIBS system performance or alteration of the obtained data set lead to drastic changes in the multivariate model. Then, utilization of such data sets biases multivariate models and obtained results. There are several possibilities how to avoid this issue, however, their implementation is burdened with the need for extensive measurement of the same sample set with various experimental parameters.

**The dawn of neural networks** is upon the LIBS community. The supreme performance of neural networks in terms of classification and quantification is indisputable. It is only a matter of time when robust, generalized neural network models will be tuned up and will substitute completely standard linear and non-linear chemometric approaches. Their implementation has already been demonstrated on numerous feasibility studies. Yet still, neural networks models demand more rigorous optimization in order to avoid over-fitting on non-informative signal and background noise.

## My contribution to the state-of-the-art

At this point I would like to briefly comment on my contribution to the state-of-the-art by the means of original scientific work published in impacted journals. To do so, I intended to bridge the gap between analytical chemistry and plasma physics on one side and data science on the other. In my authored and co-authored publications, I have targeted the issue of data processing from several perspectives. I have divided my publications into five categories, which I will comment in detail. Note that in the Appendix A I list my co-/authored publications and further comment my most important publications having the highest number of citations in Appendix B.

**Data analysis** is rather general class in which I have collected publications focused on individual steps of data processing. As I have discussed in chapter 4.1, there is no recommended algorithm for application of individual data processing steps. This depends on the experience of the operator and his heuristic ability. Let my review publication [65] serve as a collection of good practices found in LIBS literature and as a sort-of cookbook for well-balanced data processing.

The issue of *Gaussianity*, *i.e.* normal distribution of obtained signal, is often overlooked. As it was shown in my co-authored publication [101], distribution of an analytical signal oscillates from normal distribution to extreme value distribution. This depends on the experimental conditions and quantum properties of studied spectral line. Finally, bundle of publications dealt with impact of *outliers* [91], *standardization* approaches [84] and means of *background subtraction* [73] on the classification accuracy.

**Chemometrics** form a center of mass of this thesis and they are also of my paramount interest. Note that chemometric algorithms were also employed in aforementioned publications, namely PCA, SIMCA, and PLS-DA. In our pioneering study [151], a combination of principal component analysis and linear discriminant analysis was used to discriminate various brick samples with respect to their hardness (as a function of firing temperature). Those samples were measured using table-top and stand-off system. A stand-off system was used in my other study [95] where classification power of broad-range echelle spectra and short-band Czerny-Turner spectra were compared.

To the best of my knowledge, we stood at the beginning of multivariate mapping. In our research work we have fully exploited the potential of PCA [75] and SOM [127] to visualize big data, hyper-spectral images of uranium and zirconium bearing rocks. In both cases, the utilization of chemometrics enabled more efficient investigation of collected data and revealed hidden connections between data (individual sample matrices). In my other publication, I have shown that well-executed univariate analysis can provide similar performance (quantitative accuracy) as linear multivariate algorithms (PCR and PLSR). Proving my point that it is not absolutely necessary to always employ chemometrics but that it is crucial to know data and their relation to original samples.

**Laser ablation** is a complex process influenced by numerous physical, chemical and experimental parameters, as I have shown in chapter 3. We have contributed to the phenomenon of *NELIBS* in which nanoparticles are applied on the sample surface and affect the laser-matter interaction leading to significant enhancement of analytical signal. In our publication [152] the effect of reduced pressure of ambient atmosphere on the signal enhancement was studied. In another publication [38], nanoparticles were used for the first

time also in the LA-ICP-MS technique, the so-called *NE-LA-ICP-MS*. In this case, the presence of nanoparticles during laser-matter interaction led to improved laser ablation and fractionation.

**Hyphenated systems** possess great potential for further research in various applications. Surely, the combination of analytical techniques enables more rigorous investigation of samples. Together with my colleagues, I have contributed to recent state-of-the-art in hyphenated system as reviewed in chapter 2.2.

The combination of LIBS and Raman spectroscopy represents a vital solution providing complete chemical analysis of investigated samples. In my earlier works [28, 45], I have described the utilization of LIBS, LA-ICP-MS and Raman spectroscopy in the analysis of algae, potential classification of algal strains based on their spectral features and means of quantification of un-/saturated fatty acids. Next publication [47] brought direct combination of LIBS and Raman spectroscopy in the classification of various bacteria nested on agar. Higher classification accuracy was proved during joint LIBS–Raman experiment.

Most recently, the combination of LIBS and X-Ray Computed Tomography was developed; we have followed previous work where LIBS and synchrotron X-Ray radiation was used. Our publication [51] delivered a study of Pb bearing mineral, when the presence of Pb was validated by using LIBS and its location was marked in XCT data. Then, volumetric XCT data were segmented and total Pb content in the whole rock was estimated.

**Applications** of LIBS that I have pursued are discussed only in brevity. In the following publications we have intended to show the benefits of LIBS and fully exploited them in different applications, such as:

- *toxicology* where nanoparticles were mapped on a large-scale in whole plants or their parts [103, 112, 113],
- *depth profiling*, as a unique feature of laser-ablation techniques, was used to profile paint layers [119],
- *liquid* analysis and the detection of toxic heavy metal was tested [153],
- *braking tracks* can be visualized using LIBS indirectly by the detection of Zn [154, 155],
- *sulfur in concrete* was detected due to the presence of He atmosphere of various pressures [156],
- *corrosion* of Mg was quantified based on the ratios of ionic and atomic Mg spectral lines [77],
- *high-repetition rate* analysis when LIBS proved its fast turn-around time [68, 114].

# Bibliography

- [1] H. Mark and J. Workman. *Statistics in spectroscopy*. Academic press, 2<sup>nd</sup> edition, 2003. ISBN 9780124725317.
- [2] D. Hahn and N. Omenetto. Laser-induced breakdown spectroscopy (libs), part ii: Review of instrumental and methodological approaches to material analysis and applications to different fields. *Applied Spectroscopy*, 66:347–419, 2012. doi: 10.1366/11-06574.
- [3] A. Thorne, U Litzén, and S. Johansson. *Spectrophysics: principles and applications*. New York: Springer, 1<sup>st</sup> edition, 1999. ISBN 3-540-65117-9.
- [4] ChemTeam. A brief (incomplete) history of light and spectra [online]. URL <http://www.chemteam.info/Electrons/Spectrum-History.html>. [cit. 2019-12-01].
- [5] Editors. A timeline of atomic spectroscopy [online], 2006. URL <http://www.spectroscopyonline.com/timeline-atomic-spectroscopy?id=&pageID=1&sk=&date=>. [cit. 2019-12-01] *Spectroscopy online*. vol. 21. 2006.
- [6] R. Huth. History of spectroscopy - an essay [online], 2011. URL [https://www.researchgate.net/profile/Robert\\_Hunt9/publication/314285153/figure/fig1/AS:469562274652162@1488963778905/The-dark-lines-in-this-spectrum-are-absorption-lines-seen-and-drawn-by-Fraunhofer-HUJWeb.png](https://www.researchgate.net/profile/Robert_Hunt9/publication/314285153/figure/fig1/AS:469562274652162@1488963778905/The-dark-lines-in-this-spectrum-are-absorption-lines-seen-and-drawn-by-Fraunhofer-HUJWeb.png). [cit. 2019-12-01].
- [7] D. Skoog, D. West, F. Holler, and S. Crouch. *Fundamentals of analytical chemistry*. Belmont, CA: Cengage Brooks/Cole, ninth edition, 2014. ISBN 978-0-495-55828-6.
- [8] A. Miziolek, V. Palleschi, and I. Schechter. *Laser-Induced Breakdown Spectroscopy*. Cambridge: Cambridge University Press, 1<sup>st</sup> edition, 2006. ISBN 9780511541261.
- [9] R. Noll. *Laser-Induced Breakdown Spectroscopy: fundamentals and applications*. Berlin: Springer-Verlag, 1<sup>st</sup> edition, 2012. ISBN 978-3-642-20667-2.
- [10] D.A. Cremers and L.J. Radziemski. *Handbook of laser-induced breakdown spectroscopy*. Hoboken, NJ: John Wiley, 2<sup>nd</sup> edition, 2013. ISBN 978-1-11997-112-2.
- [11] D. Hahn and N. Omenetto. Laser-induced breakdown spectroscopy (libs), part i: Review of basic diagnostics and plasmaparticle interactions: Still-challenging issues within the analytical plasma community. *Applied Spectroscopy*, 64:335A–366A, 2010. doi: 10.1366/000370210793561691.

- [12] S. Maurice, R.C. Wiens, M. Saccoccio, B. Barraclough, O. Gasnault, O. Forni, N. Mangold, D. Baratoux, S. Bender, G. Berger, J. Bernardin, M. Berthé, N. Bridges, D. Blaney, M. Bouyé, P. Caïs, B. Clark, S. Clegg, A. Cousin, D. Cremers, A. Cros, L. Deflores, C. Derycke, B. Dingler, G. Dromart, B. Dubois, M. Dupieux, E. Durand, L. D’Uston, C. Fabre, B. Faure, A. Gaboriaud, T. Gharsa, K. Herkenhoff, E. Kan, L. Kirkland, D. Kouach, J.-L. Lacour, Y. Langevin, J. Lasue, S. Le Mouélic, M. Lescure, E. Lewin, D. Limonadi, G. Manhès, P. Mauchien, C. McKay, P.-Y. Meslin, Y. Michel, E. Miller, H.E. Newsom, G. Orttner, A. Paillet, L. Parès, Y. Parot, R. Pérez, P. Pinet, F. Poitrasson, B. Quartier, B. Sallé, C. Sotin, V. Sautter, H. Séran, J.J. Simmonds, J.-B. Sirven, R. Stiglich, N. Striebig, J.-J. Thocaven, M.J. Toplis, and D. Vaniman. The chemcam instrument suite on the mars science laboratory (msl) rover: Science objectives and mast unit description. *Space Science Reviews*, 170:95–166, 2012. doi: 10.1007/s11214-012-9912-2.
- [13] R.C. Wiens, S. Maurice, B. Barraclough, M. Saccoccio, W.C. Barkley, J.F. Bell III, S. Bender, J. Bernardin, D. Blaney, J. Blank, M. Bouyé, N. Bridges, N. Bultman, P. Caïs, R.C. Clanton, B. Clark, S. Clegg, A. Cousin, D. Cremers, A. Cros, L. Deflores, D. Delapp, R. Dingler, C. D’Uston, M. Darby Dyar, T. Elliott, D. Enemark, C. Fabre, M. Flores, O. Forni, O. Gasnault, T. Hale, C. Hays, K. Herkenhoff, E. Kan, L. Kirkland, D. Kouach, D. Landis, Y. Langevin, N. Lanza, F. Larocca, J. Lasue, J. Latino, D. Limonadi, C. Lindensmith, C. Little, N. Mangold, G. Manhes, P. Mauchien, C. McKay, E. Miller, J. Mooney, R.V. Morris, L. Morrison, T. Nelson, H. Newsom, A. Ollila, M. Ott, L. Pares, R. Perez, F. Poitrasson, C. Provost, J.W. Reiter, T. Roberts, F. Romero, V. Sautter, S. Salazar, J.J. Simmonds, R. Stiglich, S. Storms, N. Striebig, J.-J. Thocaven, T. Trujillo, M. Ulibarri, D. Vaniman, N. Warner, R. Waterbury, R. Whitaker, J. Witt, and B. Wong-Swanson. The chemcam instrument suite on the mars science laboratory (msl) rover: Body unit and combined system tests. *Space Science Reviews*, 170:167–227, 2012. doi: 10.1007/s11214-012-9902-4.
- [14] S. Maurice, S.M. Clegg, R.C. Wiens, O. Gasnault, W. Rapin, O. Forni, A. Cousin, V. Sautter, N. Mangold, L. Le Deit, M. Nachon, R.B. Anderson, N.L. Lanza, C. Fabre, V. Payré, J. Lasue, P.-Y. Meslin, R.J. Léveillé, B.L. Barraclough, P. Beck, S.C. Bender, G. Berger, J.C. Bridges, N.T. Bridges, G. Dromart, M.D. Dyar, R. Francis, J. Frydenvang, B. Gondet, B.L. Ehlmann, K.E. Herkenhoff, J.R. Johnson, Y. Langevin, M.B. Madsen, N. Melikechi, J.-L. Lacour, S. Le Mouélic, E. Lewin, H.E. Newsom, A.M. Ollila, P. Pinet, S. Schröder, J.-B. Sirven, R.L. Tokar, M.J. Toplis, C. D’Uston, D.T. Vaniman, and A.R. Vasavada. Chemcam activities and discoveries during the nominal mission of the mars science laboratory in gale crater, mars. *Journal of Analytical Atomic Spectrometry*, 31:863–889, 2016. doi: 10.1039/c5ja00417a.
- [15] S.C. Jantzi, V. Motto-Ros, F. Trichard, Y. Markushin, N. Melikechi, and A. De Giacomo. Sample treatment and preparation for laser-induced breakdown spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 115:52–63, 2016. doi: 10.1016/j.sab.2015.11.002.
- [16] J. Rakovský, P. Čermák, O. Musset, and P. Veis. A review of the development of portable laser induced breakdown spectroscopy and its applications. *Spectrochimica*

- Acta - Part B Atomic Spectroscopy*, 101:269–287, 2014. doi: 10.1016/j.sab.2014.09.015.
- [17] F.J. Fortes and J.J. Laserna. The development of fieldable laser-induced breakdown spectrometer: No limits on the horizon. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 65:975–990, 2010. doi: 10.1016/j.sab.2010.11.009.
- [18] R. Noll, V. Sturm, Ü. Aydin, D. Eilers, C. Gehlen, M. Höhne, A. Lamott, J. Makowe, and J. Vrenegor. Laser-induced breakdown spectroscopy-from research to industry, new frontiers for process control. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 63:1159–1166, 2008. doi: 10.1016/j.sab.2008.08.011.
- [19] B.C. Castle, K. Talabardon, B.W. Smith, and J.D. Winefordner. Variables influencing the precision of laser-induced breakdown spectroscopy measurements. *Applied Spectroscopy*, 52:649–657, 1998. doi: 10.1366/0003702981944300.
- [20] G. Galbács. A critical review of recent progress in analytical laser-induced breakdown spectroscopy. *Analytical and Bioanalytical Chemistry*, 407:7537–7562, 2017. doi: 10.1007/s00216-015-8855-3.
- [21] R. Noll, C. Fricke-Begemann, S. Connemann, C. Meinhardt, and V. Sturm. Libs analyses for industrial applications-an overview of developments from 2014 to 2018. *Journal of Analytical Atomic Spectrometry*, 33:945–956, 2018. doi: 10.1039/c8ja00076j.
- [22] V.C. Costa, J.P. Castro, D.F. Andrade, D. Victor Babos, J.A. Garcia, M.A. Sperança, T.A. Catelani, and E.R. Pereira-Filho. Laser-induced breakdown spectroscopy (libs) applications in the chemical analysis of waste electrical and electronic equipment (weee). *TrAC - Trends in Analytical Chemistry*, 108:65–73, 2018. doi: 10.1016/j.trac.2018.08.003.
- [23] K. Liu, D. Tian, C. Li, Y. Li, G. Yang, and Y. Ding. A review of laser-induced breakdown spectroscopy for plastic analysis. *TrAC - Trends in Analytical Chemistry*, 110:327–334, 2019. doi: 10.1016/j.trac.2018.11.025.
- [24] S. Millar, C. Gottlieb, T. Günther, N. Sankat, G. Wilsch, and S. Kruschwitz. Chlorine determination in cement-bound materials with laser-induced breakdown spectroscopy (libs) – a review and validation. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 147:1–8, 2018. doi: 10.1016/j.sab.2018.05.015.
- [25] R.S. Harmon, R.E. Russo, and R.R. Hark. Applications of laser-induced breakdown spectroscopy for geochemical and environmental analysis: A comprehensive review. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 87:11–26, 2013. doi: 10.1016/j.sab.2013.05.017.
- [26] E.S. Collins and J.L. Gottfried. Laser-induced deflagration for the characterization of energetic materials. *Propellants, Explosives, Pyrotechnics*, 42:592–602, 2017. doi: 10.1002/prop.201700040.
- [27] S.J. Rehse. A review of the use of laser-induced breakdown spectroscopy for bacterial classification, quantification, and identification. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 154:50–69, 2019. doi: 10.1016/j.sab.2019.02.005.

- [28] P. Pořízka, P. Prochazková, D. Prochazka, L. Sládková, J. Novotný, M. Petrilak, M. Brada, O. Samek, Z. Pilát, P. Zemánek, V. Adam, R. Kizek, K. Novotný, and J. Kaiser. Algal biomass analysis by laser-based analytical techniques—a review. *Sensors*, 14:17725–17752, 2014. doi: 10.3390/s140917725.
- [29] J. Kaiser, K. Novotný, M.Z. Martin, A. Hrdlička, R. Malina, M. Hartl, V. Adam, and R. Kizek. Trace elemental analysis by laser-induced breakdown spectroscopy - biological applications. *Surface Science Reports*, 67:233–243, 2012. doi: 10.1016/j.surfrep.2012.09.001.
- [30] A.G.G. De Carvalho, M.B. Bueno Guerra, A. Adame, C.S. Nomura, P.V. Oliveira, H.W. Pereira De Carvalho, D. Santos, L.C. Nunes, and F.J. Krug. Recent advances in libs and xrf for the analysis of plants. *Journal of Analytical Atomic Spectrometry*, 33:919–944, 2018. doi: 10.1039/c7ja00293a.
- [31] B. Sezer, G. Bilge, and I.H. Boyaci. Capabilities and limitations of libs in food analysis. *TrAC - Trends in Analytical Chemistry*, 97:345–353, 2017. doi: 10.1016/j.trac.2017.10.003.
- [32] B. Busser, S. Moncayo, J.-L. Coll, L. Sancey, and V. Motto-Ros. Elemental imaging using laser-induced breakdown spectroscopy: A new and promising approach for biological and medical applications. *Coordination Chemistry Reviews*, 358:70–79, 2018. doi: 10.1016/j.ccr.2017.12.006.
- [33] L. Jolivet, M. Leprince, S. Moncayo, L. Sorbier, C.-P. Lienemann, and V. Motto-Ros. Review of the recent advances and applications of libs-based imaging. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 151:41–53, 2019. doi: 10.1016/j.sab.2018.11.008.
- [34] R. Gaudio, N. Melikechi, Z.A. Abdel-Salam, M.A. Harith, V. Palleschi, V. Motto-Ros, and B. Busser. Laser-induced breakdown spectroscopy for human and animal health: A review. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 152:123–148, 2018. doi: 10.1016/j.sab.2018.11.006.
- [35] C. Gottlieb, S. Millar, S. Grothe, and G. Wilsch. 2d evaluation of spectral libs data derived from heterogeneous materials using cluster algorithm. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 134:58–68, 2017. doi: 10.1016/j.sab.2017.06.005.
- [36] S. Moncayo, F. Trichard, V. Bonnetterre, N. Pinel, F. Pelascini, P. Dugourd, J.-L. Coll, M. D’Incan, J. Charles, V. Motto-Ros, and L. Sancey. Characterization of foreign materials in paraffin-embedded pathological specimens using in situ multi-elemental imaging with laser spectroscopy. *Modern Pathology*, 31:378–384, 2018. doi: 10.1038/modpathol.2017.152.
- [37] B.C. Windom and D.W. Hahn. Laser ablation - laser induced breakdown spectroscopy (la-libs): A means for overcoming matrix effects leading to improved analyte response. *Journal of Analytical Atomic Spectrometry*, 24:1665–1675, 2009. doi: 10.1039/b913495f.
- [38] M. Holá, Z. Salajková, A. Hrdlička, P. Pořízka, K. Novotný, L. Čelko, P. Šperka, D. Prochazka, J. Novotný, P. Modlitbová, V. Kanický, and J. Kaiser. Feasibility of

- nanoparticle-enhanced laser ablation inductively coupled plasma mass spectrometry. *Analytical Chemistry*, 90:11820–11826, 2018. doi: 10.1021/acs.analchem.8b01197.
- [39] M. Bonta, J.J. Gonzalez, D. Quarles, R.E. Russo, B. Hegedus, and A. Limbeck. Elemental mapping of biological samples by the combined use of libs and la-icp-ms. *Journal of Analytical Atomic Spectrometry*, 31:252–258, 2016. doi: 10.1039/c5ja00287g.
- [40] M. Bonta, S. Török, B. Döme, and Limbeck A. Tandem la–libs coupled to icp-ms for comprehensive analysis of tumor samples [online], 2017. URL <http://www.spectroscopyonline.com/tandem-la-libs-coupled-icp-ms-comprehensive-analysis-tumor-samples-0>. [cit. 2019-12-01].
- [41] M. Bonta and A. Limbeck. Metal analysis in polymers using tandem la-icp-ms/libs: eliminating matrix effects using multivariate calibration. *Journal of Analytical Atomic Spectrometry*, 33:1631–1637, 2018. doi: 10.1039/c8ja00161h.
- [42] D.J. Gardiner, P.R. Graves, and H.J. Bowley. *Practical Raman spectroscopy*. New York: Springer-Verlag, 1<sup>st</sup> edition, 1989. ISBN 978-0-387-50254-0.
- [43] Q. Lin, S. Wang, G. Guo, Y. Tian, and Y. Duan. Novel laser induced breakdown spectroscopy–raman instrumentation using a single pulsed laser and an echelle spectrometer. *Instrumentation Science and Technology*, 46:163–174, 2017. doi: 10.1080/10739149.2017.1344702.
- [44] Q. Lin, G. Niu, Q. Wang, Q. Yu, and Y. Duan. Combined laser-induced breakdown with raman spectroscopy: Historical technology development and recent applications. *Applied Spectroscopy Reviews*, 48:487–508, 2013. doi: 10.1080/05704928.2012.751028.
- [45] P. Pořízka, D. Prochazka, Z. Pilát, L. Krajcarová, J. Kaiser, R. Malina, J. Novotný, P. Zemánek, J. Ježek, M. Šerý, S. Bernatová, V. Krzyžánek, K. Dobranská, K. Novotný, M. Trtílek, and O. Samek. Application of laser-induced breakdown spectroscopy to the analysis of algal biomass for industrial biotechnology. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 74-75:169–176, 2012. doi: 10.1016/j.sab.2012.06.014.
- [46] P. Pořízka, S. Kaski, A. Hrdlička, P. Modlitbová, L. Sládková, H. Häkkinen, D. Prochazka, J. Novotný, P. Gadas, L. Čelko, K. Novotný, and J. Kaiser. Detection of fluorine using laser-induced breakdown spectroscopy and raman spectroscopy. *Journal of Analytical Atomic Spectrometry*, 32:1966–1974, 2017. doi: 10.1039/c7ja00200a.
- [47] D. Prochazka, M. Mazura, O. Samek, K. Rebrošová, P. Pořízka, J. Klus, P. Prochazková, J. Novotný, K. Novotný, and J. Kaiser. Combination of laser-induced breakdown spectroscopy and raman spectroscopy for multivariate classification of bacteria. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 139: 6–12, 2018. doi: 10.1016/j.sab.2017.11.004.
- [48] M. Hoehse, A. Paul, I. Gornushkin, and U. Panne. Multivariate classification of pigments and inks using combined raman spectroscopy and libs. *Analytical and Bioanalytical Chemistry*, 402:1443–1450, 2012. doi: 10.1007/s00216-011-5287-6.

- [49] J. Kaiser, O. Samek, L. Reale, M. Liška, R. Malina, A. Ritucci, A. Poma, A. Tucci, F. Flora, A. Lai, L. Mancini, G. Tromba, F. Zanini, A. Faenov, T. Pikuz, and G. Cinque. Monitoring of the heavy-metal hyperaccumulation in vegetal tissues by x-ray radiography and by femto-second laser induced breakdown spectroscopy. *Microscopy Research and Technique*, 70:147–153, 2007. doi: 10.1002/jemt.20394.
- [50] M. Galiová, J. Kaiser, K. Novotný, M. Ivanov, M. Nývltová Fišáková, L. Mancini, G. Tromba, T. Vaculovič, M. Liška, and V. Kanický. Investigation of the osteitis deformans phases in snake vertebrae by double-pulse laser-induced breakdown spectroscopy. *Analytical and Bioanalytical Chemistry*, 398:1095–1107, 2010. doi: 10.1007/s00216-010-3976-1.
- [51] D. Prochazka, T. Zikmund, P. Pořízka, A. Břínek, J. Klus, J. Šalplachta, J. Kynický, J. Novotný, and J. Kaiser. Joint utilization of double-pulse laser-induced breakdown spectroscopy and x-ray computed tomography for volumetric information of geological samples. *Journal of Analytical Atomic Spectrometry*, 33: 1993–1999, 2018. doi: 10.1039/c8ja00232k.
- [52] A.H. Galmed, A. du Plessis, S.G. le Roux, E. Hartnick, H. Von Bergmann, and M. Maaza. Three dimensional characterization of laser ablation craters using high resolution x-ray computed tomography. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 139:75–82, 2018. doi: 10.1016/j.sab.2017.11.011.
- [53] R.E. Russo, X. Mao, and S.S. Mao. The physics of laser ablation in microchemical analysis. *Analytical Chemistry*, 74:70A–77A, 2002. doi: 10.1021/ac0219445.
- [54] H.R. Griem. *Spectral line broadening by plasmas*. Academic Press, 1<sup>st</sup> edition, 1974. ISBN 9780323150941.
- [55] H.R. Griem. *Principles of Plasma Spectroscopy*. New York: Cambridge University Press, 1<sup>st</sup> edition, 1997. ISBN 0-521-45504-9.
- [56] J.D. Ingle and S.R. Crouch. *Spectrochemical Analysis*. Englewood Cliffs, N.J.: Prentice Hall, 1<sup>st</sup> edition, 1988. ISBN 0-13-826876-2.
- [57] T. Fujimoto. *Plasma spectroscopy*. International series of monographs on physics. New York: Oxford University Press, 1<sup>st</sup> edition, 2004. ISBN 0-198-53028-5.
- [58] C.C. Garcia, H. Lindner, A. Von Bohlen, C. Vadla, and K. Niemax. Elemental fractionation and stoichiometric sampling in femtosecond laser ablation. *Journal of Analytical Atomic Spectrometry*, 23:470–478, 2008. doi: 10.1039/b718845e.
- [59] T.A. Labutin, V.N. Lednev, A.A. Ilyin, and A.M. Popov. Femtosecond laser-induced breakdown spectroscopy. *Journal of Analytical Atomic Spectrometry*, 31:90–118, 2016. doi: 10.1039/c5ja00301f.
- [60] J. Colgan, J.E. Barefield II, E.J. Judge, K. Campbell, H.M. Johns, D.P. Kilcrease, R. McInroy, and S.M. Clegg. Experimental and theoretical studies of laser-induced breakdown spectroscopy emission from iron oxide: studies of atmospheric effects. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 122:85–92, 2016. doi: 10.1016/j.sab.2016.05.016.

- [61] J. Hermann, L. Mercadier, E. Mothe, G. Socol, and P. Alloncle. On the stoichiometry of mass transfer from solid to plasma during pulsed laser ablation of brass. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 65:636–641, 2010. doi: 10.1016/j.sab.2010.03.015.
- [62] X. Bai, F. Cao, V. Motto-Ros, Q. Ma, Y. Chen, and J. Yu. Morphology and characteristics of laser-induced aluminum plasma in argon and air: A comparative study. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 113:158–166, 2015. doi: 10.1016/j.sab.2015.09.023.
- [63] S. Eschlböck-Fuchs, A. Demidov, I.B. Gornushkin, T. Schmid, R. Rössler, N. Huber, U. Panne, and J.D. Pedarnig. Tomography of homogenized laser-induced plasma by radon transform technique. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 123: 59–67, 2016. doi: 10.1016/j.sab.2016.07.007.
- [64] C. Aragon, J. Bengoechea, and J.A. Aguilera. Influence of the optical depth on spectral line emission from laser-induced plasmas. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 56:619–628, 2001. doi: 10.1016/S0584-8547(01)00172-0.
- [65] P. Pořízka, J. Klus, E. Képeš, D. Prochazka, D.W. Hahn, and J. Kaiser. On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 148: 65–82, 2018. doi: 10.1016/j.sab.2018.05.030.
- [66] J. El Haddad, L. Canioni, and B. Bousquet. Good practices in libs analysis: Review and advices. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 101:171–182, 2014. doi: 10.1016/j.sab.2014.08.039.
- [67] V. Motto-Ros, D. Syvilay, L. Bassel, E. Negre, F. Trichard, F. Pelascini, J. El Haddad, A. Harhira, S. Moncayo, J. Picard, D. Devismes, and B. Bousquet. Critical aspects of data analysis for quantification in laser-induced breakdown spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 140:54–64, 2018. doi: 10.1016/j.sab.2017.12.004.
- [68] P. Pořízka, J. Klus, J. Mašek, M. Rajnoha, D. Prochazka, P. Modlitbová, J. Novotný, R. Burget, K. Novotný, and J. Kaiser. Multivariate classification of echellograms: A new perspective in laser-induced breakdown spectroscopy analysis. *Scientific Reports*, 7:1–7, 2017. doi: 10.1038/s41598-017-03426-0.
- [69] I.B. Gornushkin, P.E. Eagan, A.B. Novikov, B.W. Smith, and J.D. Winefordner. Automatic correction of continuum background in laser-induced breakdown and raman spectrometry. *Applied Spectroscopy*, 57:187–207, 2003. doi: 10.1366/000370203321535123.
- [70] C.M. Galloway, E.C. Le Ru, and P.G. Etchegoin. An iterative algorithm for background removal in spectroscopy by wavelet transforms. *Applied Spectroscopy*, 63:1370–1376, 2009. doi: 10.1366/000370209790108905.
- [71] P. Yaroshchyyk and J.E. Eberhardt. Automatic correction of continuum background in laser-induced breakdown spectroscopy using a model-free algorithm. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 99:138–149, 2014. doi: 10.1016/j.sab.2014.06.020.

- [72] M.D. Dyar, S. Giguere, C.J. Carey, and T. Boucher. Comparison of baseline removal methods for laser-induced breakdown spectroscopy of geological samples. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 126:53–64, 2016. doi: 10.1016/j.sab.2016.10.018.
- [73] E. Képeš, P. Pořízka, J. Klus, P. Modlitbová, and J. Kaiser. Influence of baseline subtraction on laser-induced breakdown spectroscopic data. *Journal of Analytical Atomic Spectrometry*, 33:2107–2115, 2018. doi: 10.1039/c8ja00209f.
- [74] R.C. Chinni, D.A. Cremers, L.J. Radziemski, M. Bostian, and C. Navarro-Northrup. Detection of uranium using laser-induced breakdown spectroscopy. *Applied Spectroscopy*, 63:1238–1250, 2009. doi: 10.1366/000370209789806867.
- [75] J. Klus, P. Mikysek, D. Prochazka, P. Pořízka, P. Prochazková, J. Novotný, T. Trojek, K. Novotný, M. Slobodník, and J. Kaiser. Multivariate approach to the chemical mapping of uranium in sandstone-hosted uranium ores analyzed using double pulse laser-induced breakdown spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 123:143–149, 2016. doi: 10.1016/j.sab.2016.08.014.
- [76] E. Tognoni and G. Cristoforetti. Signal and noise in laser induced breakdown spectroscopy: An introductory review. *Optics and Laser Technology*, 79:164–172, 2016. doi: 10.1016/j.optlastec.2015.12.010.
- [77] P. Pořízka, I. Ročňáková, J. Klus, D. Prochazka, L. Sládková, P. Šperka, Z. Spötz, L. Čelko, K. Novotný, and J. Kaiser. Estimating the grade of mg corrosion using laser-induced breakdown spectroscopy. *Journal of Analytical Atomic Spectrometry*, 30:2099–2106, 2015. doi: 10.1039/c5ja00257e.
- [78] W.B. Barnett, V.A. Fassel, and R.N. Kniseley. Theoretical principles of internal standardization in analytical emission spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 23:643–664, 1968. doi: 10.1016/0584-8547(68)80045-X.
- [79] D. Body and B.L. Chadwick. Optimization of the spectral data processing in a libs simultaneous elemental analysis system. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 56:725–736, 2001. doi: 10.1016/S0584-8547(01)00186-0.
- [80] V. Kanicky, V. Otruba, and J.-M. Mermet. Use of internal standardization to compensate for a wide range of absorbance in the analysis of glasses by uv laser ablation inductively coupled plasma atomic emission spectrometry. *Applied Spectroscopy*, 52:638–642, 1998. doi: 10.1366/0003702981944201.
- [81] A. Hrdlička, L. Zaorálková, M. Galiová, T. Čtvrtníčková, V. Kanický, V. Otruba, K. Novotný, P. Krásenský, J. Kaiser, R. Malina, and K. Páleníková. Correlation of acoustic and optical emission signals produced at 1064 and 532 nm laser-induced breakdown spectroscopy (libs) of glazed wall tiles. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 64:74–78, 2009. doi: 10.1016/j.sab.2008.10.043.
- [82] N.B. Zorov, A.A. Gorbatenko, T.A. Labutin, and A.M. Popov. A review of normalization techniques in analytical atomic spectrometry with laser sampling: From single to multivariate correction. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 65:642–657, 2010. doi: 10.1016/j.sab.2010.04.009.

- [83] J.P. Castro and E.R. 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). *Journal of Analytical Atomic Spectrometry*, 31:2005–2014, 2016. doi: 10.1039/c6ja00224b.
- [84] P. Pořízka, J. Klus, A. Hrdlička, J. Vrábek, P. Škarková, D. Prochazka, J. Novotný, K. Novotný, and J. Kaiser. Impact of laser-induced breakdown spectroscopy data normalization on multivariate classification accuracy. *Journal of Analytical Atomic Spectrometry*, 32:277–288, 2017. doi: 10.1039/c6ja00322b.
- [85] J.E. Carranza and D.W. Hahn. Sampling statistics and considerations for single-shot analysis using laser-induced breakdown spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 57:779–790, 2002. doi: 10.1016/S0584-8547(02)00007-1.
- [86] V. Lazic, F. Colao, R. Fantoni, and V. Spizzicchino. Laser-induced breakdown spectroscopy in water: Improvement of the detection threshold by signal processing. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 60:1002–1013, 2005. doi: 10.1016/j.sab.2005.06.007.
- [87] J.-B. Sirven, B. Sallé, P. Mauchien, J.-L. Lacour, S. Maurice, and G. Manhès. Feasibility study of rock identification at the surface of mars by remote laser-induced breakdown spectroscopy and three chemometric methods. *Journal of Analytical Atomic Spectrometry*, 22:1471–1480, 2007. doi: 10.1039/b704868h.
- [88] I.B. Gornushkin, U. Panne, and J.D. Winefordner. Linear correlation for identification of materials by laser induced breakdown spectroscopy: Improvement via spectral filtering and masking. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 64:1040–1047, 2009. doi: 10.1016/j.sab.2009.07.038.
- [89] P. Yaroshchuk, D.L. Death, and S.J. Spencer. Comparison of principal components regression, partial least squares regression, multi-block partial least squares regression, and serial partial least squares regression algorithms for the analysis of Fe in iron ore using libs. *Journal of Analytical Atomic Spectrometry*, 27:92–98, 2012. doi: 10.1039/c1ja10164a.
- [90] J.-M. Mermet. Limit of quantitation in atomic spectrometry: An unambiguous concept? *Spectrochimica Acta - Part B Atomic Spectroscopy*, 63:166–182, 2008. doi: 10.1016/j.sab.2007.11.029.
- [91] P. Pořízka, J. Klus, D. Prochazka, E. Képeš, A. Hrdlička, J. Novotný, K. Novotný, and J. Kaiser. Laser-induced breakdown spectroscopy coupled with chemometrics for the analysis of steel: The issue of spectral outliers filtering. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 123:114–120, 2016. doi: 10.1016/j.sab.2016.08.008.
- [92] A. Kramida, Yu. Ralchenko, J. Reader, and NIST ASD Team. Nist atomic spectra database (version 5.7.1) [online], 2019. URL <https://physics.nist.gov/asd>. [cit. 2019-12-01].
- [93] P.L. Smith, C. Heise, J.R. Esmond, and R.L. Kurucz. Atomic spectral line database [online]. URL <https://www.cfa.harvard.edu/amp/ampdata/kurucz23/sekur.html>. [cit. 2019-12-01].

- [94] G. Amato, G. Cristoforetti, S. Legnaioli, G. Lorenzetti, V. Palleschi, F. Sorrentino, and E. Tognoni. Progress towards an unassisted element identification from laser induced breakdown spectra with automatic ranking techniques inspired by text retrieval. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 65:664–670, 2010. doi: 10.1016/j.sab.2010.04.019.
- [95] P. Pořízka, J. Klus, D. Prochazka, G. Vítková, M. Brada, J. Novotný, K. Novotný, and J. Kaiser. Assessment of the most effective part of echelle laser-induced plasma spectra for further classification using czerny-turner spectrometer. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 124:116–123, 2016. doi: 10.1016/j.sab.2016.09.004.
- [96] P. de Boves Harrington. Statistical validation of classification and calibration models using bootstrapped latin partitions. *Trends in Analytical Chemistry*, 25: 1112–1124, 2008. doi: 10.1016/j.trac.2006.10.010.
- [97] S. Moncayo, L. Duponchel, N. Mousavipak, G. Panczer, F. Trichard, B. Bousquet, F. Pelascini, and V. Motto-Ros. Exploration of megapixel hyperspectral libs images using principal component analysis. *Journal of Analytical Atomic Spectrometry*, 33: 210–220, 2018. doi: 10.1039/c7ja00398f.
- [98] D. Syvilay, N. Wilkie-Chancellier, B. Trichereau, A. Texier, L. Martinez, S. Serfaty, and V. Detalle. Evaluation of the standard normal variate method for laser-induced breakdown spectroscopy data treatment applied to the discrimination of painting layers. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 114:38–45, 2015. doi: 10.1016/j.sab.2015.09.022.
- [99] A.P.M. Michel and A.D. Chave. Analysis of laser-induced breakdown spectroscopy spectra: The case for extreme value statistics. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 62:1370–1378, 2007. doi: 10.1016/j.sab.2007.10.027.
- [100] S. Coles. *An Introduction to Statistical Modeling of Extreme Values*. New York: Springer, 1<sup>st</sup> edition, 2001. ISBN 978-1-85233-459-8.
- [101] J. Klus, P. Pořízka, D. Prochazka, J. Novotný, K. Novotný, and J. Kaiser. Effect of experimental parameters and resulting analytical signal statistics in laser-induced breakdown spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 126: 6–10, 2016. doi: 10.1016/j.sab.2016.10.002.
- [102] D.L. Massart. *Handbook of chemometrics and qualimetrics*. New York: Elsevier, 1<sup>st</sup> edition, 1998. ISBN 0-444-89724-0.
- [103] P. Modlitbová, A. Hlaváček, T. Švestková, P. Pořízka, L. Šimoníková, K. Novotný, and J. Kaiser. The effects of photon-upconversion nanoparticles on the growth of radish and duckweed: Bioaccumulation, imaging, and spectroscopic studies. *Chemosphere*, 225:723–734, 2019. doi: 10.1016/j.chemosphere.2019.03.074.
- [104] R.G. Brereton. *Chemometrics: data analysis for the laboratory and chemical plant*. Hoboken, NJ: John Wiley, 1<sup>st</sup> edition, 2003. ISBN 9780471489771.
- [105] E. Voigtman. *Limits of detection in chemical analysis*. Hoboken, NJ: John Wiley, 1<sup>st</sup> edition, 2017. ISBN 9781119188971.

- [106] E. Voigtman. Limits of detection and decision. part 4. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 63:154–165, 2008. doi: 10.1016/j.sab.2007.11.014.
- [107] J.-M. Mermet. Calibration in atomic spectrometry: A tutorial review dealing with quality criteria, weighting procedures and possible curvatures. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 65:509–523, 2010. doi: 10.1016/j.sab.2010.05.007.
- [108] J. Carlson, A. Wysoczanski, and E. Voigtman. Limits of quantitation - yet another suggestion. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 96:69–73, 2014. doi: 10.1016/j.sab.2014.03.012.
- [109] P. Pořízka, A. Demidov, J. Kaiser, J. Keivanian, I. Gornushkin, U. Panne, and J. Riedel. Laser-induced breakdown spectroscopy for in situ qualitative and quantitative analysis of mineral ores. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 101:155–163, 2014. doi: 10.1016/j.sab.2014.08.027.
- [110] I.B. Gornushkin, J.M. Anzano, L.A. King, B.W. Smith, N. Omenetto, and J.D. Winefordner. Curve of growth methodology applied to laser-induced plasma emission spectroscopy. *Spectrochimica Acta, Part B: Atomic spectroscopy*, 54: 491–503, 1999. doi: 10.1016/S0584-8547(99)00004-X.
- [111] J.O. Cáceres, F. Pelascini, V. Motto-Ros, S. Moncayo, F. Trichard, G. Panczer, A. Marín-Roldán, J.A. Cruz, I. Coronado, and J. Martín-Chivelet. Megapixel multi-elemental imaging by laser-induced breakdown spectroscopy, a technology with considerable potential for paleoclimate studies. *Scientific Reports*, 7:1–11, 2017. doi: 10.1038/s41598-017-05437-3.
- [112] P. Skarkova, K. Novotny, P. Lubal, A. Jebava, P. Porizka, J. Klus, Z. Farka, A. Hrdlicka, and J. Kaiser. 2d distribution mapping of quantum dots injected onto filtration paper by laser-induced breakdown spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 131:107–114, 2017. doi: 10.1016/j.sab.2017.03.016.
- [113] P. Modlitbová, K. Novotný, P. Pořízka, J. Klus, P. Lubal, H. Zlámalová-Gargošová, and J. Kaiser. Comparative investigation of toxicity and bioaccumulation of cd-based quantum dots and cd salt in freshwater plant lemna minor l. *Ecotoxicology and Environmental Safety*, 147:334–341, 2018. doi: 10.1016/j.ecoenv.2017.08.053.
- [114] P. Pořízka, B. Klessen, J. Kaiser, I. Gornushkin, U. Panne, and J. Riedel. High repetition rate laser-induced breakdown spectroscopy using acousto-optically gated detection. *Review of Scientific Instruments*, 85:1–9, 2014. doi: 10.1063/1.4890337.
- [115] H.A.O. Wang, D. Grolimund, C. Giesen, C.N. Borca, J.R.H. Shaw-Stewart, B. Bodenmiller, and D. Günther. Fast chemical imaging at high spatial resolution by laser ablation inductively coupled plasma mass spectrometry. *Analytical Chemistry*, 85:10107–10116, 2013. doi: 10.1021/ac400996x.
- [116] R. Gaudiuso. Calibration-free inverse method for depth-profile analysis with laser-induced breakdown spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 123:105–113, 2016. doi: 10.1016/j.sab.2016.08.002.
- [117] H. Hou, L. Cheng, T. Richardson, G. Chen, M. Doeff, R. Zheng, R. Russo, and V. Zorba. Three-dimensional elemental imaging of li-ion solid-state electrolytes

- using fs-laser induced breakdown spectroscopy (libs). *Journal of Analytical Atomic Spectrometry*, 30:2295–2302, 2015. doi: 10.1039/c5ja00250h.
- [118] Y. Gimenez, B. Busser, F. Trichard, A. Kulesza, J.M. Laurent, V. Zaun, F. Lux, J.M. Benoit, G. Panczer, P. Dugourd, O. Tillement, F. Pelascini, L. Sancey, and V. Motto-Ros. 3d imaging of nanoparticle distribution in biological tissue by laser-induced breakdown spectroscopy. *Scientific Reports*, 6:1–9, 2016. doi: 10.1038/srep29936.
- [119] E. Pospíšilová, K. Novotný, P. Pořízka, D. Hradil, J. Hradilová, J. Kaiser, and V. Kanický. Depth-resolved analysis of historical painting model samples by means of laser-induced breakdown spectroscopy and handheld x-ray fluorescence. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 147:100–108, 2018. doi: 10.1016/j.sab.2018.05.018.
- [120] P. Mehta, M. Bukov, C.-H. Wang, A.G.R. Day, C. Richardson, C.K. Fisher, and D.J. Schwab. A high-bias, low-variance introduction to machine learning for physicists. *Physics Reports*, 810:1–124, 2019. doi: 10.1016/j.physrep.2019.03.001.
- [121] R.G. Brereton. *Applied chemometrics for scientists*. Hoboken, NJ: John Wiley, 1<sup>st</sup> edition, 2007. ISBN 978-0-470-01686-2.
- [122] L. Eriksson. *Multi- and megavariable data analysis: basic principles and applications*. Malmö: Umetrics Academy, 3<sup>rd</sup> edition, 2013. ISBN 978-91-973730-5-0.
- [123] M. Meloun, J. Militký, and M. Hill. *Počítačová analýza vícerozměrných dat v příkladech*. Praha: Academia, 1<sup>st</sup> edition, 2005. ISBN 80-200-1335-0.
- [124] M. Baudelet. *Laser spectroscopy for sensing: fundamentals, techniques and applications*. Amsterdam: Woodhead Publishing, 1<sup>st</sup> edition, 2014. ISBN 978-0-85709-273-1.
- [125] R.S. Sutton and A.G. Barto. *Reinforcement learning: an introduction*. Cambridge, Mass.: MIT Press, 2<sup>nd</sup> edition, 1998. ISBN 0262193981.
- [126] J. Lasue, R.C. Wiens, T.F. Stepinski, O. Forni, S.M. Clegg, and S. Maurice. Nonlinear mapping technique for data visualization and clustering assessment of libs data: Application to chemcam data. *Analytical and Bioanalytical Chemistry*, 400: 3247–3260, 2011. doi: 10.1007/s00216-011-4747-3.
- [127] J. Klus, P. Pořízka, D. Prochazka, P. Mikysek, J. Novotný, K. Novotný, M. Slobodník, and J. Kaiser. Application of self-organizing maps to the study of u-zr-ti-nb distribution in sandstone-hosted uranium ores. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 131:66–73, 2017. doi: 10.1016/j.sab.2017.03.008.
- [128] P. Pořízka. *Using Laser-Induced Breakdown Spectroscopy (LIBS) for material analysis*. PhD thesis, Brno University of Technology, Faculty of Mechanical Engineering, Brno, 2014. prof. Ing. Jozef Kaiser, Ph.D.
- [129] F.C. De Lucia and J.L. Gottfried. Influence of variable selection on partial least squares discriminant analysis models for explosive residue classification. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 66:122–128, 2011. doi: 10.1016/j.sab.2010.12.007.

- [130] R.A. Putnam, Q.I. Mohaidat, A. Daabous, and S.J. Rehse. A comparison of multivariate analysis techniques and variable selection strategies in a laser-induced breakdown spectroscopy bacterial classification. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 87:161–167, 2013. doi: 10.1016/j.sab.2013.05.014.
- [131] A. Kumar Myakalwar, N. Spegazzini, C. Zhang, S. Kumar Anubham, R.R. Dasari, I. Barman, and M. Kumar Gundawar. Less is more: Avoiding the libs dimensionality curse through judicious feature selection for explosive detection. *Scientific Reports*, 5(13169), 2015. doi: 10.1038/srep13169.
- [132] A. Larsson, H. Andersson, and L. Landström. Impact of data reduction on multivariate classification models built on spectral data from bio-samples. *Journal of Analytical Atomic Spectrometry*, 30:1117–1127, 2015. doi: 10.1039/c4ja00467a.
- [133] A. Safi, B. Campanella, E. Grifoni, S. Legnaioli, G. Lorenzetti, S. Pagnotta, F. Poggialini, L. Ripoll-Seguer, M. Hidalgo, and V. Palleschi. Multivariate calibration in laser-induced breakdown spectroscopy quantitative analysis: The dangers of a ‘black box’ approach and how to avoid them. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 144:46–54, 2018. doi: 10.1016/j.sab.2018.03.007.
- [134] V. Motto-Ros, A.S. Koulejev, G.R. Osinski, and A.E. Dudelzak. Quantitative multi-elemental laser-induced breakdown spectroscopy using artificial neural networks. *Journal of European Optical Society - rapid publications*, 3(08011), 2008. doi: 10.2971/jeos.2008.08011.
- [135] J. El Haddad, D. Bruyère, A. Ismaël, G. Gallou, V. Laperche, K. Michel, L. Canioni, and B. Bousquet. Application of a series of artificial neural networks to on-site quantitative analysis of lead into real soil samples by laser induced breakdown spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 97: 57–64, 2014. doi: 10.1016/j.sab.2014.04.014.
- [136] C. Gottlieb, S. Millar, T. Günther, and G. Wilsch. Revealing hidden spectral information of chlorine and sulfur in data of a mobile laser-induced breakdown spectroscopy system using chemometrics. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 132:43–49, 2017. doi: 10.1016/j.sab.2017.04.001.
- [137] F.-Y. Yueh, H. Zheng, J.P. Singh, and S. Burgess. Preliminary evaluation of laser-induced breakdown spectroscopy for tissue classification. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 64:1059–1067, 2009. doi: 10.1016/j.sab.2009.07.025.
- [138] R.B. Anderson, J.F. Bell III, R.C. Wiens, R.V. Morris, and S.M. Clegg. Clustering and training set selection methods for improving the accuracy of quantitative laser induced breakdown spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 70:24–32, 2012. doi: 10.1016/j.sab.2012.04.004.
- [139] J.L. Gottfried, R.S. Harmon, F.C. De Lucia, and A.W. Miziolek. Multivariate analysis of laser-induced breakdown spectroscopy chemical signatures for geomaterial classification. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 64: 1009–1019, 2009. doi: 10.1016/j.sab.2009.07.005.
- [140] S.M. Clegg, E. Sklute, M.D. Dyar, J.E. Barefield, and R.C. Wiens. Multivariate analysis of remote laser-induced breakdown spectroscopy spectra using partial least

- squares, principal component analysis, and related techniques. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 64:79–88, 2009. doi: 10.1016/j.sab.2008.10.045.
- [141] A.M. Ollila, J. Lasue, H.E. Newsom, R.A. Multari, R.C. Wiens, and S.M. Clegg. Comparison of two partial least squares-discriminant analysis algorithms for identifying geological samples with the chemcam laser-induced breakdown spectroscopy instrument. *Applied Optics*, 51:B130–B142, 2012. doi: 10.1364/AO.51.00B130.
- [142] R.A. Multari, D.A. Cremers, J.M. Dupre, and J.E. Gustafson. The use of laser-induced breakdown spectroscopy for distinguishing between bacterial pathogen species and strains. *Applied Spectroscopy*, 64:750–759, 2010. doi: 10.1366/000370210791666183.
- [143] S. Moncayo, S. Manzoor, F. Navarro-Villoslada, and J.O. Caceres. Evaluation of supervised chemometric methods for sample classification by laser induced breakdown spectroscopy. *Chemometrics and Intelligent Laboratory Systems*, 146: 354–364, 2015. doi: 10.1016/j.chemolab.2015.06.004.
- [144] J. Vrábel, E. Képeš, and P. Pořízka. Emslibs contest [online], 2019. URL <http://libs.ceitec.cz/libs-contest/>. [cit. 2019-12-01].
- [145] H. Grahn and P. Geladi. *Techniques and applications of hyperspectral image analysis*. Hoboken, NJ: John Wiley, 1<sup>st</sup> edition, 2007. ISBN 978-0-470-01086-0.
- [146] S. Romppanen, H. Häkkänen, and S. Kaski. Singular value decomposition approach to the yttrium occurrence in mineral maps of rare earth element ores using laser-induced breakdown spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 134:69–74, 2017. doi: 10.1016/j.sab.2017.06.002.
- [147] A. Mezzacappa, N. Melikechi, A. Cousin, R.C. Wiens, J. Lasue, S.M. Clegg, R. Tokar, S. Bender, N.L. Lanza, S. Maurice, G. Berger, O. Forni, O. Gasnault, M.D. Dyar, T. Boucher, E. Lewin, and C. Fabre. Application of distance correction to chemcam laser-induced breakdown spectroscopy measurements. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 120:19–29, 2016. doi: 10.1016/j.sab.2016.03.009.
- [148] S.M. Clegg, R.C. Wiens, R. Anderson, O. Forni, J. Frydenvang, J. Lasue, A. Cousin, V. Payré, T. Boucher, M.D. Dyar, S.M. McLennan, R.V. Morris, T.G. Graff, S.A. Mertzman, B.L. Ehlmann, I. Belgacem, H. Newsom, B.C. Clark, N. Melikechi, A. Mezzacappa, R.E. McInroy, R. Martinez, P. Gasda, O. Gasnault, and S. Maurice. Recalibration of the mars science laboratory chemcam instrument with an expanded geochemical database. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 129: 64–85, 2017. doi: 10.1016/j.sab.2016.12.003.
- [149] J. Novotný, P. Pořízka, J. Kaiser, V. Kanický, K. Novotný, and Z. Chládová. *EMSLIBS 2019 Book of abstracts*. Ioannes Marcus Marci Spectroscopic Society, 2019. ISBN 978-80-88195-13-9.
- [150] A.D. McNaught and A. Wilkinson. Iupac compendium of chemical terminology [online], 2009. URL <https://goldbook.iupac.org/>. [cit. 2019-12-01].

- [151] G. Vítková, L. Prokeš, K. Novotný, P. Pořízka, J. Novotný, D. Všianský, L. Čelko, and 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. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 101:191–199, 2014. doi: 10.1016/j.sab.2014.08.036.
- [152] L. Sládková, D. Prochazka, P. Pořízka, P. Škarková, M. Remešová, A. Hrdlička, K. Novotný, L. Čelko, and J. Kaiser. Improvement of the laser-induced breakdown spectroscopy method sensitivity by the usage of combination of ag-nanoparticles and vacuum conditions. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 127: 48–55, 2017. doi: 10.1016/j.sab.2016.11.005.
- [153] K. Skočovská, J. Novotný, D. Prochazka, P. Pořízka, K. Novotný, and J. Kaiser. Optimization of liquid jet system for laser-induced breakdown spectroscopy analysis. *Review of Scientific Instruments*, 87, 2016. doi: 10.1063/1.4947233.
- [154] D. Prochazka, M. Bilík, P. Prochazková, J. Klus, P. Pořízka, J. Novotný, K. Novotný, B. Ticová, A. Bradáč, M. Semela, and J. Kaiser. Detection of tire tread particles using laser-induced breakdown spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 108:1–7, 2015. doi: 10.1016/j.sab.2015.03.011.
- [155] D. Prochazka, M. Bilík, P. Prochazková, M. Brada, J. Klus, P. Pořízka, J. Novotný, K. Novotný, B. Ticová, A. Bradáč, M. Semela, and J. Kaiser. Detection of visually unrecognizable braking tracks using laser-induced breakdown spectroscopy, a feasibility study. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 118:90–97, 2016. doi: 10.1016/j.sab.2016.02.013.
- [156] A. Hrdlička, J. Hegrová, K. Novotný, V. Kanický, D. Prochazka, J. Novotný, P. Modlitbová, L. Sládková, P. Pořízka, and J. Kaiser. Sulfur determination in concrete samples using laser-induced breakdown spectroscopy and limestone standards. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 142:8–13, 2018. doi: 10.1016/j.sab.2018.01.015.

# Appendix A

## List of authored publications

In this section I present a complete list of my 36 publications having 340 citations (including self citations). I also add printed versions of the most cited ones together with a short description of their novelty and contribution to the state-of-the-art, Appendix B. Note that those ten publications form my H-index: 11, see the record in Figure A.1 obtained from the Web of Science on Mar 3<sup>rd</sup>, 2020.



Figure A.1: The citation record of my publications obtained from Web of Science on Mar 3<sup>rd</sup>, 2020.

**A list of authored publications** ordered chronologically. Note that the citations were obtained from the Web of Science on Mar 3<sup>rd</sup>, 2020.

1. P. Pořízka, D. Prochazka, Z. Pilát, L. Krajcarová, J. Kaiser, R. Malina, J. Novotný, P. Zemánek, J. Ježek, M. Šerý, S. Bernatová, V. Krzyžánek, K. Dobranská, K. Novotný, M. Trtílek and O. Samek. Application of laser-induced breakdown spectroscopy to the analysis of algal biomass for industrial biotechnology. *Spectrochimica Acta Part B*, 74–75:169–176, 2012.  
DOI: 10.1016/j.sab.2012.06.014; 20 citations
2. P. Pořízka, P. Prochazková, D. Prochazka, L. Sládková, J. Novotný, M. Petrilak, M. Brada, O. Samek, Z. Pilát, P. Zemánek, V. Adam, R. Kizek, K. Novotný, and J. Kaiser. Algal biomass analysis by laser-based analytical techniques—A review. *Sensors*, 14:17725–17752, 2014.  
DOI: 10.3390/s140917725; 36 citations
3. P. Pořízka, B. Klessen, J. Kaiser, I. Gornushkin, U. Panne, and J. Riedel. High repetition rate laser-induced breakdown spectroscopy using acousto-optically gated detection. *Review of Scientific Instruments*, 85:073104, 2014.  
DOI: 10.1063/1.4890337; 5 citations
4. G. Vítková, L. Prokeš, K. Novotný, P. Pořízka, J. Novotný, D. Všianský, L. Čelko, and 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. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 101:191–199, 2014.  
DOI: 10.1016/j.sab.2014.08.036, 32 citations
5. P. Pořízka, A. Demidov, J. Kaiser, J. Keivanian, I. Gornushkin, U. Panne, and J. Riedel. Laser-induced breakdown spectroscopy for in situ qualitative and quantitative analysis of mineral ores. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 101:155–163, 2014.  
DOI: 10.1016/j.sab.2014.08.027; 24 citations
6. P. Pořízka, I. Ročňáková, J. Klus, D. Prochazka, L. Sládková, P. Šperka, Z. Spatz, L. Čelko, K. Novotný, and J. Kaiser. Estimating the grade of Mg corrosion using laser-induced breakdown spectroscopy. *Journal of Analytical Atomic Spectrometry*, 30:2099–2106, 2015.  
DOI: 10.1039/c5ja00257e; 5 citations
7. D. Prochazka, M. Bilík, P. Prochazková, J. Klus, P. Pořízka, J. Novotný, K. Novotný, B. Ticová, A. Bradáč, M. Semela, and J. Kaiser. Detection of tire tread particles using laser-induced breakdown spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 108:1–7, 2015.  
DOI: 10.1016/j.sab.2015.03.011; 12 citations
8. D. Prochazka, M. Bilík, P. Prochazková, M. Brada, J. Klus, P. Pořízka, J. Novotný, K. Novotný, B. Ticová, A. Bradáč, M. Semela, and J. Kaiser. Detection of visually unrecognizable braking tracks using Laser-Induced Breakdown Spectroscopy, a feasibility study. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 118:90–97, 2016.

DOI: 10.1016/j.sab.2016.02.013; 1 citation

9. K. Skočovská, J. Novotný, D. Prochazka, P. Pořízka, K. Novotný, and J. Kaiser. Optimization of liquid jet system for laser-induced breakdown spectroscopy analysis. *Review of Scientific Instruments*, 87:043116, 2016.

DOI: 10.1063/1.4947233; 6 citations

10. J. Klus, P. Mikysek, D. Prochazka, P. Pořízka, P. Prochazková, J. Novotný, T. Trojek, K. Novotný, M. Slobodník, and J. Kaiser. Multivariate approach to the chemical mapping of uranium in sandstone-hosted uranium ores analyzed using double pulse Laser-Induced Breakdown Spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 123:143–149, 2016.

DOI: 10.1016/j.sab.2016.08.014; 17 citations

11. P. Pořízka, J. Klus, D. Prochazka, E. Képeš, A. Hrdlička, J. Novotný, K. Novotný, and J. Kaiser. Laser-Induced Breakdown Spectroscopy coupled with chemometrics for the analysis of steel: The issue of spectral outliers filtering. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 123:114–120, 2016.

DOI: 10.1016/j.sab.2016.08.008; 20 citations

12. P. Pořízka, J. Klus, D. Prochazka, G. Vítková, M. Brada, J. Novotný, K. Novotný, and J. Kaiser. Assessment of the most effective part of echelle laser-induced plasma spectra for further classification using Czerny-Turner spectrometer. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 124:116–123, 2016.

DOI: 10.1016/j.sab.2016.09.004; 6 citations

13. J. Klus, P. Pořízka, D. Prochazka, J. Novotný, K. Novotný, and J. Kaiser. Effect of experimental parameters and resulting analytical signal statistics in laser-induced breakdown spectroscopy. *Spectrochimica Acta Part B*, 123:143–149, 2016.

DOI: 10.1016/j.sab.2016.10.002; 10 citations

14. L. Sládková, D. Prochazka, P. Pořízka, P. Škarková, M. Remešová, A. Hrdlička, J. Novotný, L. Čelko, and J. Kaiser. Improvement of the Laser-Induced Breakdown Spectroscopy method sensitivity by the usage of combination of Ag-nanoparticles and vacuum conditions. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 127:48–55, 2017.

DOI: 10.1016/j.sab.2016.11.005; 8 citations

15. P. Pořízka, J. Klus, A. Hrdlička, J. Vrábel, P. Škarková, D. Prochazka, J. Novotný, K. Novotný, and J. Kaiser. Impact of Laser-Induced Breakdown Spectroscopy data normalization on multivariate classification accuracy. *Journal of Analytical Atomic Spectrometry*, 32:277–288, 2017.

DOI: 10.1039/c6ja00322b; 32 citations

16. P. Škarková, K. Novotný, P. Lubal, A. Jebavá, P. Pořízka, J. Klus, Z. Farka, A. Hrdlička, and J. Kaiser. 2d distribution mapping of quantum dots injected onto filtration paper by laser-induced breakdown spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 131:107–114, 2017.

DOI: 10.1016/j.sab.2017.03.016; 8 citations

17. J. Klus, P. Pořízka, D. Prochazka, P. Mikysek, J. Novotný, K. Novotný, M. Slobodník, and J. Kaiser. Application of self-organizing maps to the study of U-Zr-Ti-Nb distribution in sandstone-hosted uranium ores. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 131:66–73, 2017.  
DOI: 10.1016/j.sab.2017.03.008; 5 citations
18. P. Pořízka, S. Kaski, A. Hrdlička, P. Modlitbová, L. Sládková, H. Häkkinen, D. Prochazka, J. Novotný, P. Gadas, L. Čelko, K. Novotný, and J. Kaiser. Detection of fluorine using laser-induced breakdown spectroscopy and Raman spectroscopy. *Journal of Analytical Atomic Spectrometry*, 32:1966–1874, 2017.  
DOI: 10.1039/c7ja00200a; 10 citations
19. P. Pořízka, J. Klus, J. Mašek, M. Rajnoha, D. Prochazka, P. Modlitbová, J. Novotný, R. Burget, K. Novotný, and J. Kaiser. Multivariate classification of echellograms: A new perspective in Laser-Induced Breakdown Spectroscopy analysis. *Scientific Reports*, 7:3160, 2017.  
DOI: 10.1038/s41598-017-03426-0; 3 citations
20. D. Prochazka, M. Mazura, O. Samek, K. Rebrošová, P. Pořízka, J. Klus, P. Prochazková, J. Novotný, K. Novotný, and J. Kaiser. Combination of laser-induced breakdown spectroscopy and Raman spectroscopy for multivariate classification of bacteria. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 139:6–12, 2018.  
DOI: 10.1016/j.sab.2017.11.004; 10 citations
21. P. Modlitbová, K. Novotný, P. Pořízka, J. Klus, P. Lubal, H. Zlámalová-Gargošová, and J. Kaiser. Comparative investigation of toxicity and bioaccumulation of Cd-based quantum dots and Cd salt in freshwater plant *Lemna minor* L. *Ecotoxicology and Environmental Safety*, 147:334–341, 2018.  
DOI: 10.1016/j.ecoenv.2017.08.053; 17 citations
22. A. Hrdlička, J. Hegrová, K. Novotný, V. Kanický, D. Prochazka, J. Novotný, P. Modlitbová, L. Sládková, P. Pořízka, and J. Kaiser. Sulfur determination in concrete samples using laser-induced breakdown spectroscopy and limestone standards. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 142:8–13, 2018.  
DOI: 10.1016/j.sab.2018.01.015; 6 citations
23. P. Modlitbová, P. Pořízka, K. Novotný, J. Drbohlavová, I. Chamradová, Z. Farka, H. Zlámalová-Gargošová, T. Romih, and J. Kaiser. Short-term assessment of cadmium toxicity and uptake from different types of Cd-based Quantum Dots in the model plant *Allium cepa* L. *Ecotoxicology and Environmental Safety*, 153:23–31, 2018.  
DOI: 10.1016/j.ecoenv.2018.01.044; 15 citations
24. P. Modlitbová, K. Klepárník, Z. Farka, P. Pořízka, P. Skládal, K. Novotný, and J. Kaiser. Time-dependent growth of silica shells on CdTe quantum dots. *Nanomaterials*, 8:439, 2018.  
DOI: 10.3390/nano8060439; 1 citation

25. E. Pospíšilová, K. Novotný, P. Pořízka, D. Hradil, J. Hradilová, J. Kaiser, and V. Kanický. Depth-resolved analysis of historical painting model samples by means of laser-induced breakdown spectroscopy and handheld X-ray fluorescence. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 147:100–108, 2018.  
DOI: 10.1016/j.sab.2018.05.018; 5 citations
26. P. Pořízka, J. Klus, E. Képeš, D. Prochazka, D.W. Hahn, and J. Kaiser. On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 148:65–82, 2018.  
DOI: 10.1016/j.sab.2018.05.030; 16 citations
27. M. Holá, Z. Salajková, A. Hrdlička, P. Pořízka, K. Novotný, L. Čelko, P. Šperka, D. Prochazka, J. Novotný, P. Modlitbová, V. Kanický, and J. Kaiser. Feasibility of Nanoparticle-Enhanced Laser Ablation Inductively Coupled Plasma Mass Spectrometry. *Analytical Chemistry*, 90:11820–11826, 2018.  
DOI: 10.1021/acs.analchem.8b01197; 3 citations
28. D. Prochazka, T. Zikmund, P. Pořízka, A. Břínek, J. Klus, J. Šalplachta, J. Kynický, J. Novotný, and J. Kaiser. Joint utilization of double-pulse laser-induced breakdown spectroscopy and X-ray computed tomography for volumetric information of geological samples. *Journal of Analytical Atomic Spectrometry*, 33:1993–1999, 2018.  
DOI: 10.1039/c8ja00232k; 3 citations
29. E. Képeš, P. Pořízka, J. Klus, P. Modlitbová, and J. Kaiser. Influence of baseline subtraction on laser-induced breakdown spectroscopic data. *Journal of Analytical Atomic Spectrometry*, 33:2107–2115, 2018.  
DOI: 10.1039/c8ja00209f; 2 citation
30. P. Modlitbová, A. Hlaváček, T. Švestková, P. Pořízka, L. Šimoníková, K. Novotný, and J. Kaiser. The effects of photon-upconversion nanoparticles on the growth of radish and duckweed: Bioaccumulation, imaging, and spectroscopic studies. *Chemosphere*, 225:723–734, 2019.  
DOI: 10.1016/j.chemosphere.2019.03.074; 2 citation
31. P. Modlitbová, Z. Farka, M. Pastucha, P. Pořízka, K. Novotný, P. Skládal, and J. Kaiser. Laser-induced breakdown spectroscopy as a novel readout method for nanoparticle-based immunoassays. *Microchimica Acta*, 186:629, 2019.  
DOI: 10.1007/s00604-019-3742-9; 0 citations
32. J. Vrábel, P. Pořízka, J. Klus, D. Prochazka, J. Novotný, D. Koutný, D. Paloušek, and J. Kaiser. Classification of materials for selective laser melting by laser-induced breakdown spectroscopy. *Chemical Papers*, 73:2897–2905, 2019.  
DOI: 10.1007/s11696-018-0609-1; 0 citations
33. E. Pospíšilová, K. Novotný, P. Pořízka, J. Hradilová, J. Kaiser, and V. Kanický. Influence of laser wavelength and laser energy on depth profiling of easel painting samples. *Chemical Papers*, 73:2937–2943, 2019.  
DOI: 10.1007/s11696-019-00803-z; 0 citations

34. E. Képeš, P. Pořízka, and J. Kaiser. On the application of bootstrapping to laser-induced breakdown spectroscopy data. *Journal of Analytical Atomic Spectrometry*, 34:2411–2419, 2019.  
DOI: 10.1039/c9ja00304e; 0 citations
35. P. Modlitbová, P. Pořízka, and J. Kaiser. Laser-induced breakdown spectroscopy as a promising tool in the elemental bioimaging of plant tissues. *Trends in Analytical Chemistry*, 122:115729, 2020.  
DOI: 10.1016/j.trac.2019.115729; 0 citations

## Appendix B

# Printed versions of most significant authored publications

### Article 1

P. Pořízka, D. Prochazka, Z. Pilát, L. Krajcarová, J. Kaiser, R. Malina, J. Novotný, P. Zemánek, J. Ježek, M. Šerý, S. Bernatová, V. Krzyžánek, K. Dobranská, K. Novotný, M. Trtílek and O. Samek. Application of laser-induced breakdown spectroscopy to the analysis of algal biomass for industrial biotechnology. *Spectrochimica Acta Part B*, 74–75:169–176, 2012.

DOI: 10.1016/j.sab.2012.06.014; 20 citations

My very first publication was dedicated to the feasibility of algae analysis using LIBS. Despite any reasonable progress in this application, there is still considerable potential in direct, *in-situ* utilization of LIBS.

There exist several approaches in the analysis of algal samples. They are grown and harvested in the form of a dense suspension. Then, the analysis may be provided directly in the liquid phase or after certain preparation, such as deposition on the filter paper. In the publication, the liquid suspension was measured in two arrangements. First, surface of the bulk liquid was analyzed. This approach brings more disadvantages than advantages, namely focusing of the laser pulse on liquid surface is challenging. Moreover, orthogonal double pulse LIBS was used in order to improve sensitivity and plasma stability. Second, suspension was circulated in the specially designed vessel. Thin liquid jet was analyzed using collinear double pulse LIBS.

Un-/saturated fatty acids are of interest in the algal analysis. Calibrating LIBS response to fatty acid concentration will be a game-changer in this application.



Contents lists available at SciVerse ScienceDirect

## Spectrochimica Acta Part B

journal homepage: [www.elsevier.com/locate/sab](http://www.elsevier.com/locate/sab)

## Application of laser-induced breakdown spectroscopy to the analysis of algal biomass for industrial biotechnology

P. Pořízka<sup>a</sup>, D. Prochazka<sup>a</sup>, Z. Pilát<sup>b</sup>, L. Krajcarová<sup>c</sup>, J. Kaiser<sup>a,\*</sup>, R. Malina<sup>a</sup>, J. Novotný<sup>a</sup>, P. Zemánek<sup>b</sup>, J. Ježek<sup>b</sup>, M. Šerý<sup>b</sup>, S. Bernatová<sup>b</sup>, V. Krzyžánek<sup>b</sup>, K. Dobranská<sup>b</sup>, K. Novotný<sup>c</sup>, M. Trtílek<sup>d</sup>, O. Samek<sup>b</sup>

<sup>a</sup> X-ray micro CT and nano CT research group, CEITEC—Central European Institute of Technology, Brno University of Technology, Technická 3058/10, 616 00 Brno, Czech Republic

<sup>b</sup> Institute of Scientific Instruments of the ASCR v.v.i., Academy of Sciences of the Czech Republic, Královopolská 147, Brno 61669, Czech Republic

<sup>c</sup> Department of Chemistry, Faculty of Sciences, Masaryk University, Kotlářská 2, Brno 611 37, Czech Republic

<sup>d</sup> Photon Systems Instruments, Drásov 470, 664 24 Drásov, Czech Republic

### ARTICLE INFO

#### Article history:

Received 15 December 2011

Accepted 18 June 2012

Available online 26 June 2012

#### Keywords:

LIBS

Double-pulse

Water-jet

Algal biomass

Biotechnology

### ABSTRACT

We report on the application of laser-induced breakdown spectroscopy (LIBS) to the determination of elements distinctive in terms of their biological significance (such as potassium, magnesium, calcium, and sodium) and to the monitoring of accumulation of potentially toxic heavy metal ions in living microorganisms (algae), in order to trace e.g. the influence of environmental exposure and other cultivation and biological factors having an impact on them. Algae cells were suspended in liquid media or presented in a form of adherent cell mass on a surface (biofilm) and, consequently, characterized using their spectra. In our feasibility study we used three different experimental arrangements employing double-pulse LIBS technique in order to improve on analytical selectivity and sensitivity for potential industrial biotechnology applications, e.g. for monitoring of mass production of commercial biofuels, utilization in the food industry and control of the removal of heavy metal ions from industrial waste waters.

© 2012 Elsevier B.V. All rights reserved.

### 1. Introduction

One of the most promising alternatives to satisfy the increasing demands of the human population for energy sources is the production of carbon-based fuels from plants. By the process of photosynthesis plants convert the energy of solar radiation into the chemical energy stored in the molecular building blocks of life (proteins, lipids, carbohydrates, etc.), thus providing energy for most of the life forms on Earth. Moreover, photosynthesis has a crucial impact on Earth's atmosphere as it generates oxygen while simultaneously using up carbon dioxide, a prominent greenhouse-effect gas responsible for global climate changes.

In order to utilize photoautotrophic microorganisms (algae) for efficient biofuel, food industry, and bioremediation applications [1,2] the optimal cultivation parameters have to be determined for each given purpose, which results in a high production of oil in the selected cell line, increased production of carotenoids/omega-3 oils and potential absorption of heavy-metals in the selected cell line, respectively. This can be accomplished using small-scale photobioreactors that allow precise monitoring and control of the culture irradiance, temperature, pH, and gas composition in the medium. However, the ability to monitor the elemental composition of cells and consequently

the cellular response to external stimuli in real time (ratios/elemental compositions of algal cells might be significantly changed over short periods of time) is not provided. For this purpose qualitative/quantitative information about elemental composition of the algal cells that provides a deeper insight into the cellular physiology and enables more efficient optimization of the selected parameters for specified purposes outlined above is required. So far, only a few studies [3–6] have been performed – using dedicated chemical techniques for lipid determination – to analyze the effect of nutrient elements on oil production/bioremediation. Thus, a fast and remote technique which is capable to analyze elements *in-situ* within algal cells would be of an advantage.

The technique of laser-induced breakdown spectroscopy (LIBS) utilizes the high power densities employing focused radiation from a pulsed, fixed-frequency laser in order to generate luminous plasma from a sample (solid, liquid, and gaseous samples) [7]. In our experiments we assume the stoichiometric ablation so that the plasma composition reflects the elemental composition of the ablated target.

Here we report on the application of a double-pulse LIBS technique to the analysis of important elements including heavy metal ions in algal cells. The technique of double-pulse LIBS [8] was introduced by K. Niemax and his co-workers from ISAS Dortmund in 1991 [9], and nowadays it is a common technique used in many LIBS research laboratories. LIBS has been previously applied to the analysis of biological samples [7,10–12], namely to the determination of Sr in algal pellets [13].

\* Corresponding author.

E-mail address: [kaiser@fme.vutbr.cz](mailto:kaiser@fme.vutbr.cz) (J. Kaiser).

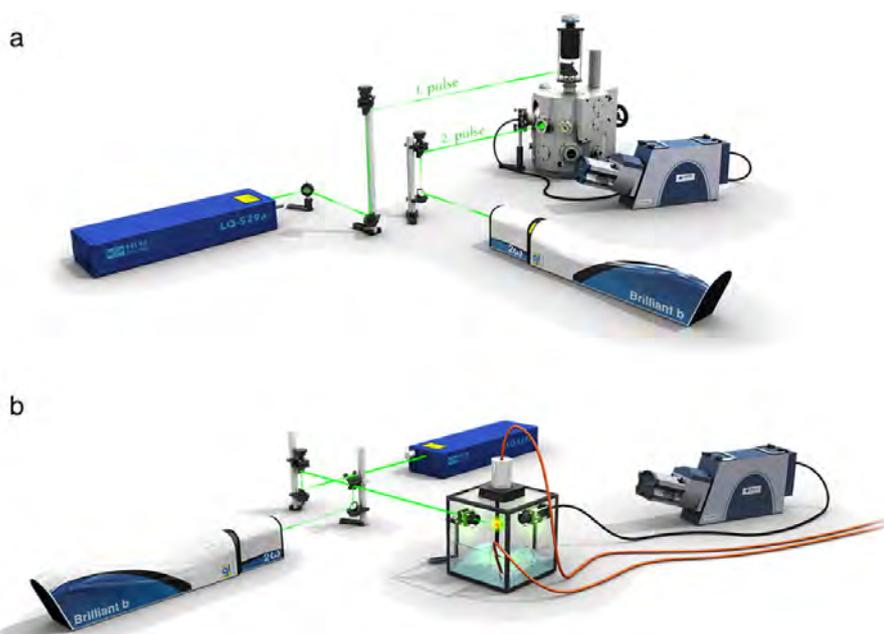


Fig. 1. a). Schematic diagram of LIBS system for ablation of biofilm. b) Schematic diagram of LIBS system employing laminar water jet setup.

The realization of LIBS apparatus employing a water jet for real-time, in situ and remote analysis of trace elements in liquid samples, which is potentially applicable to the analysis of pollutants in water in harsh or difficult-to-reach environments and used in our experiments, has been previously described [14–20].

Algal cells used for our investigations were present in liquid suspension (as was mentioned above for the use within the photobioreactor) or in a form of biofilm. Algal biomass in natural environment may act as a bio-indicator of pollution. If some algal species are selectively added into the polluted area, they might serve as a bioremediation tool.

We have utilized the LIBS set up in the three following modifications to follow different applications demands:

- (i) algal cells were suspended in a given volume of the liquid of interest and analyzed using double-pulse LIBS employing a water jet (wet procedure);
- (ii) algal cells were suspended in a given volume of the liquid of interest and analyzed using double-pulse LIBS employing bulk liquid measurements (wet procedure);
- (iii) algal biofilm was analyzed using double-pulse LIBS (dry procedure).

Valuable information about elemental composition of the cells brought by the LIBS technique can potentially help to elucidate important questions in algal biology (nutrition dynamics depending on the cultivation conditions) and identify the algal strains which have the potential for applications in metal-ion sorption (bioremediation), in food industry (source of omega-3 oils and proteins [21]), or in a biofuel industry. Here we would like to note, that bioremediation can be beneficially combined with the production of biofuels [22]. For our experiments we have selected the two algal species—*Trachydiscus minutus* (potential source of omega-3 fatty acids such as EPA and DHA) [23] and *Chlamydomonas* sp. (potential candidate for biofuel production) [24].

To the best of our knowledge these are the first LIBS measurements utilizing setup with the liquid jet so that algae samples could

be measured *in-vivo*, *on-line* and in *real-time*. Obviously, this is the area (real-world application) where LIBS technique can excel because the measurements can be performed directly e.g. at the bioreactor or remotely in *difficult-to-reach* environments at contaminated sites.

## 2. Experimental

As was mentioned above, in our investigations we used the three different experimental arrangements using double-pulse LIBS technique. Firstly, for fast, remote and *in situ* analysis we focused the laser beam onto the smooth vertical surface of a laminar jet stream of the liquid similar to our previously published arrangement [17],



Fig. 2. Schematic diagram of experimental setup for bulk liquid LIBS experiments involving the two lasers. Ablation laser I (from the top) at 266 nm and re-heating laser pulse II (from the side) at 1064 nm are shown.

secondly, onto the water surface (bulk liquid), very much as when routinely investigating solid targets. Thirdly, we set up a typical arrangement found in the majority of LIBS laboratories where laser is focused onto the surface of algal biofilm.

### 2.1. Preparation of algal samples for laser ablation experiments

The algal species were obtained from the Culture Collection of Autotrophic Organisms (CCALA), Institute of Botany, Academy of Sciences of the Czech Republic. This collection maintains about 650 strains of microalgae and cyanobacteria, a substantial portion of which are microorganisms from extreme environments. Recently, great effort has been put to the screening of microalgal strains suitable for biotechnological exploitation with respect to specific metabolites and production characteristics.

*T. minutus* (Bourrelly) was cultivated in 300 ml batch cultures at room temperature (approx. 22 °C), in a medium containing (in mg l<sup>-1</sup>) N 150, P<sub>2</sub>O<sub>5</sub> 50, K<sub>2</sub>O 300, MgO 30, SO<sub>3</sub> 75, B 0.3, Cu 0.1, Fe 0.7, Mn 0.4, Mo 0.04, and Zn 0.3, dissolved in 1:1 distilled and tap water mixture. The incident PAR flux density estimate was between 100 and 300 μmol photons m<sup>-2</sup> s<sup>-1</sup>. The cultivation apparatus consisted of 1000 ml Pyrex conical flasks fitted with a cotton plug, which were occasionally shaken manually. Lighting was supplied with two standard cool white 40 W fluorescent tubes. *Chlamydomonas* sp. was cultivated in 150 ml Erlenmeyer flasks in BBM medium at room temperature in daylight at a laboratory window with occasional manual

mixing. The cells were harvested in the stationary phase and subjected to the treatment with increased concentrations of some metal ions.

The culture aliquots were each supplied with solutions of CuSO<sub>4</sub> 10 mg ml<sup>-1</sup> in 1:1000 (v/v) ratio, and the cultures were incubated for 24 hours with occasional shaking. The sedimented cells were washed with distilled water and concentrated by a brief centrifugation. For biomass/biofilm formation the sediment was then transferred with spatula on the surface of a glass cover slip. The residual water was left to evaporate in the room conditions in a semi-closed dustproof container. Dried samples were subjected to the laser ablation. In order to produce algal suspension the sedimented cells were placed to a defined amount of liquid.

We would like to note that the question of biofilm definition has been debated in many discussions. In our study we favor the simple definition from a recently published review [25]—“a biofilm is a thin coating comprised of living material.”

### 2.2. LIBS experimental setup utilizing water jet (algal suspensions) and biofilm

Measurements of algal samples on two different setups for experiments with samples in solid state, as biofilms, and in a form of liquid, as the algal suspension, were performed at Brno University of Technology. We have employed setup with orthogonal pulses for biofilm experiments because this configuration enables better mapping of the sample [26,27] to setup with collinear pulses, which was used for water jet analysis.

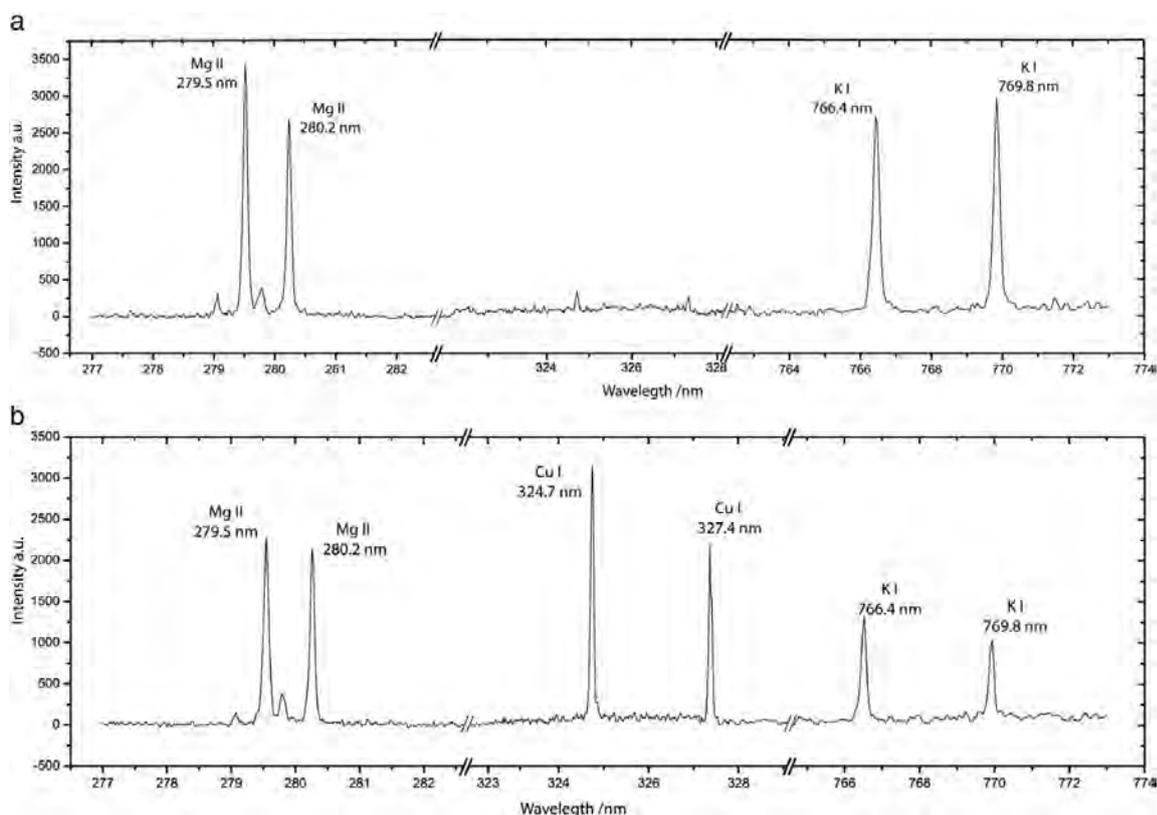


Fig. 3. a, b Details of LIBS spectra for the algae biofilm sample. (a) Batch 1 – Cu solution was not added to the batch—clearly only spectra of Mg and K are shown; (b) Batch 2 – algal biofilm was prepared with adding Cu to the batch—both Cu lines are visible indicating accumulation of Cu within algae cells.

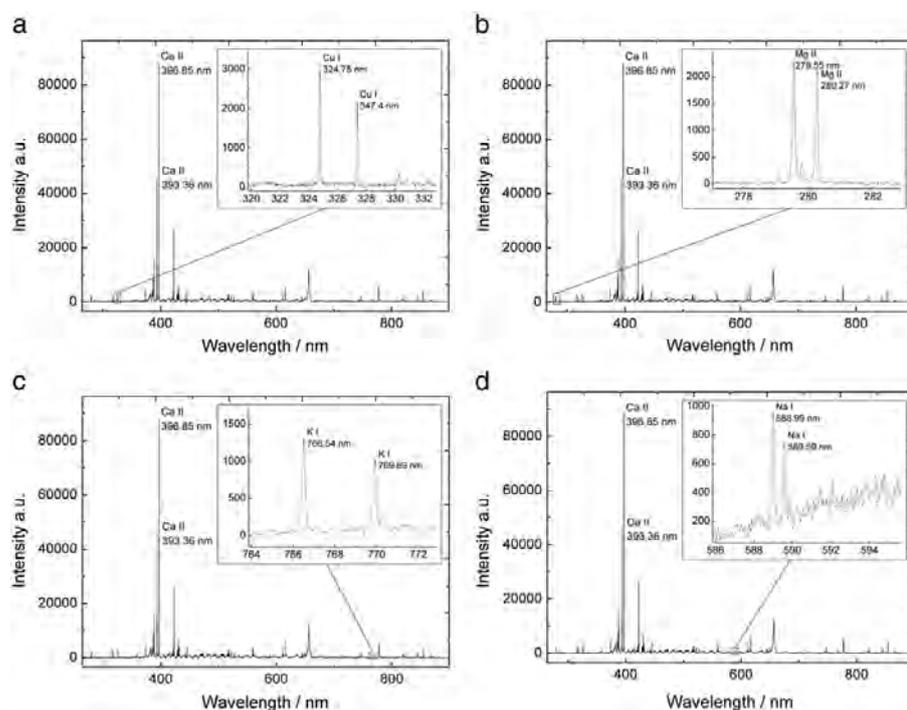


Fig. 4. a–d Segments of LIBS spectra, recorded from biofilm sample used in analysis highlighting individual elements—(a) Cu, (b) Mg, (c) K and (d) Na.

### 2.2.1. Biofilm experiments

The double-pulse LIBS (DP LIBS) measurements were carried out in a perpendicular geometry (Fig. 1a); this setup was previously described by our group [26,27]. The ablation laser was directed normal to the surface sample by mirrors and focused onto the sample surface by a 30 mm focal length glass triplet. The beam of the second laser was directed parallel to the sample surface in the plasma using mirrors and focused on the plasma by the lens of 40 mm focal length. The setup involved the two Q-switched Nd:YAG lasers both operating at 10 Hz. The first laser LQ 529a (SOLAR, BY) which operated at the second harmonic (532 nm) with the pulse width of about 10 ns was used as the ablation source. The second laser Brilliant B (Quintel, FR) operating at the first harmonic (1064 nm) with pulse width of about 6 ns was used for reheating the plasma. The energy of the ablation laser pulse was set on 12 mJ per pulse and the energy of the second laser was 110 mJ per pulse, the laser focal diameter was  $\sim 80 \mu\text{m}$ .

The sample was mounted on a stage with precision movements ( $2 \mu\text{m}$  resolution) inside the ablation chamber (Tesca, a.s. CZ). The ablation spot was targeted and controlled by a CCD camera placed outside of the ablation chamber.

### 2.2.2. Algal suspension

Setup for measuring liquid samples can be easily attached to DP LIBS setup described above. Liquid measurements were performed in home-made glass vessel instead of interaction chamber Fig. 1b.

Continuous and steady thin flow of liquid has been achieved for liquid LIBS measurements. For this we used peristaltic pump (PCD 81, Kouřil, CZ) working at 100 ml/min. Liquid sample was introduced to the nozzle (diameter of 0.6 mm) via silicone tubes and mounted to the XY positioning stage (ThorLabs, US).

Pulses of both lasers were led through the optical system (ThorLabs, US/Newport, UK) in the collinear geometry, employing harmonic separator (Eksma Optics, LT) reflecting 1064 nm and

transmitting 532 nm. The lens with 75 mm focal length focused the laser beams into the thin flow of liquid and, consequently, luminous micro-plasma was created.

In both above mentioned cases the LIBS plasma radiation was collected with UV-NIR achromatic collimating mirror system, the CC52 (ANDOR, UK) and transported via a fiber optic system ( $25 \mu\text{m}$  in diameter) connected to a spectrometer in the Echelle configuration (ME5000, Mechelle, ANDOR, UK). As a detector an ICCD camera (iStar 734i, ANDOR, UK) was employed.

The time-resolved studies were performed by controlling following parameters—the gate width  $t_w$  (time during which the spectra are integrated), the gate delay time  $t_d$  to reduce the effect of continuum signal in collected spectra (time delay after second laser pulse after which the spectra are acquired by the detector), and the delay between the two pulses  $\Delta t$ . The two lasers and ICCD camera were triggered via a delay generator (DG535, Stanford Research System, US). The best signal to noise ratio was obtained for the timing of optimized setup  $t_w = 16 \mu\text{s}$ ,  $t_d = 1.5 \mu\text{s}$ , and  $\Delta t = 1.5 \mu\text{s}$ .

The LIBS analysis was performed in the air at atmospheric pressure. For each spectral window – obtained from the biofilm or from algal suspension – 10 laser shots were averaged. The gain level on an ICCD camera was set on 80.

### 2.3. LIBS experimental setup utilizing bulk liquid arrangement (algal suspension)

The measurements of algal suspensions were performed at the Department of Chemistry, Masaryk University in Brno employing the orthogonal double-pulse LIBS setup (Fig. 2). The modified commercially available laser ablation system (UP 266 MACRO, New Wave, US), operating at the 266 nm with energy of 12 mJ per pulse, was used as the ablation laser. The energy of the ablation laser was set on 12 mJ using dedicated software. For the second, re-heating

laser pulse which propagated parallel to the sample surface a Nd:YAG laser (Brilliant, Quantel, FR) at fundamental wavelength 1064 nm (with energy of 100 mJ per pulse) was used. The laser beam was focused by the 120 mm focal length lens to intersect the path of the first laser beam and, consequently, to create microplasma placed 0.5 mm above the sample surface. The double pulse setup was controlled with two delay generators (DG 645, Stanford Research Systems, US) in order to provide the time synchronization of both lasers and the ICCD detector.

Each spectrum from the bulk liquid arrangement was obtained from single laser shot and with no gain set on the camera. As the optimal inter-pulse delay 600 ns was set according to the highest signal-to-noise ratio observed. The plasma emission was focused by the 80 mm focal length glass lens into the 3 m long optical fiber and transported into the entrance slit of Czerny-Turner spectrometer (TRIAx 320, Jobin Yvon, FR; 2400 g/mm; 50  $\mu$ m entrance slit) equipped with ICCD detector (Horiba, Jobin Yvon, FR). The gate delay time after second laser pulse and gate width of the detection were as well optimized to maximal signal to noise ratio and were set to 500 ns and 10  $\mu$ s, respectively.

### 3. Results and discussion

In this study, two different algal species—*T. minutus* and *Chlamydomonas* sp. were examined in terms of the elemental composition. We have detected magnesium, calcium, potassium, sodium, and copper. The corresponding atomic emission lines utilized were as follows: Mg lines at 279.5 nm and 280.2 nm, Ca lines at 393.3 nm and 396.8 nm, K lines at 766.5 nm and 769.8 nm, Na lines 588.9 nm and 589.6 nm and Cu lines at 324.7 nm and 327.4 nm.

#### 3.1. Measurements on algal biofilm—metal-ion binding experiments

LIBS experiments were performed on algal biofilm in order to reveal elemental composition of a biological sample. We targeted the dried biofilm (prepared according to the procedure described above), so that “real-world” LIBS applications can be followed. As an emerging issue, the mechanism of metal-ion binding to algae was studied by LIBS on the two algal samples—firstly “batch 1” was prepared following normal protocol and secondly “batch 2” was treated with specified amount of Cu (as described above). Examples of the

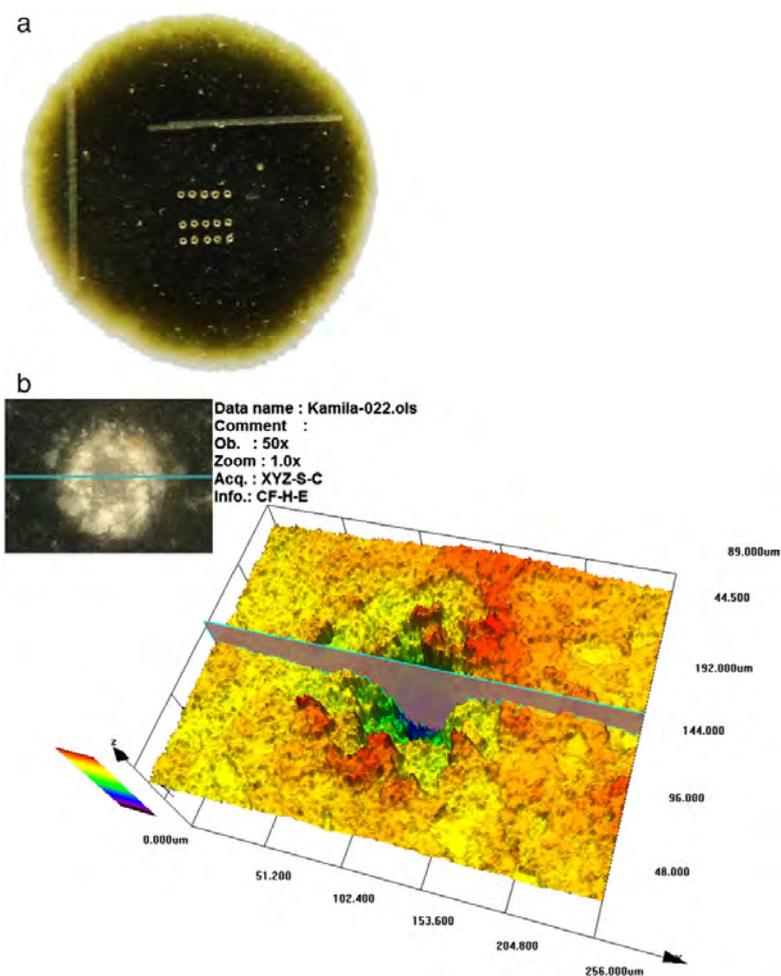


Fig. 5. a, b (a) Overall view on the biofilm sample clearly showing ablated craters during LIBS analysis (in the middle of the sample, where 15 craters are located on the area of about  $1 \times 1$  cm); (b) an image of the crater obtained from the confocal laser scanning microscope for the crater parameters estimation—depth of about 70  $\mu$ m and the diameter of about 100  $\mu$ m.

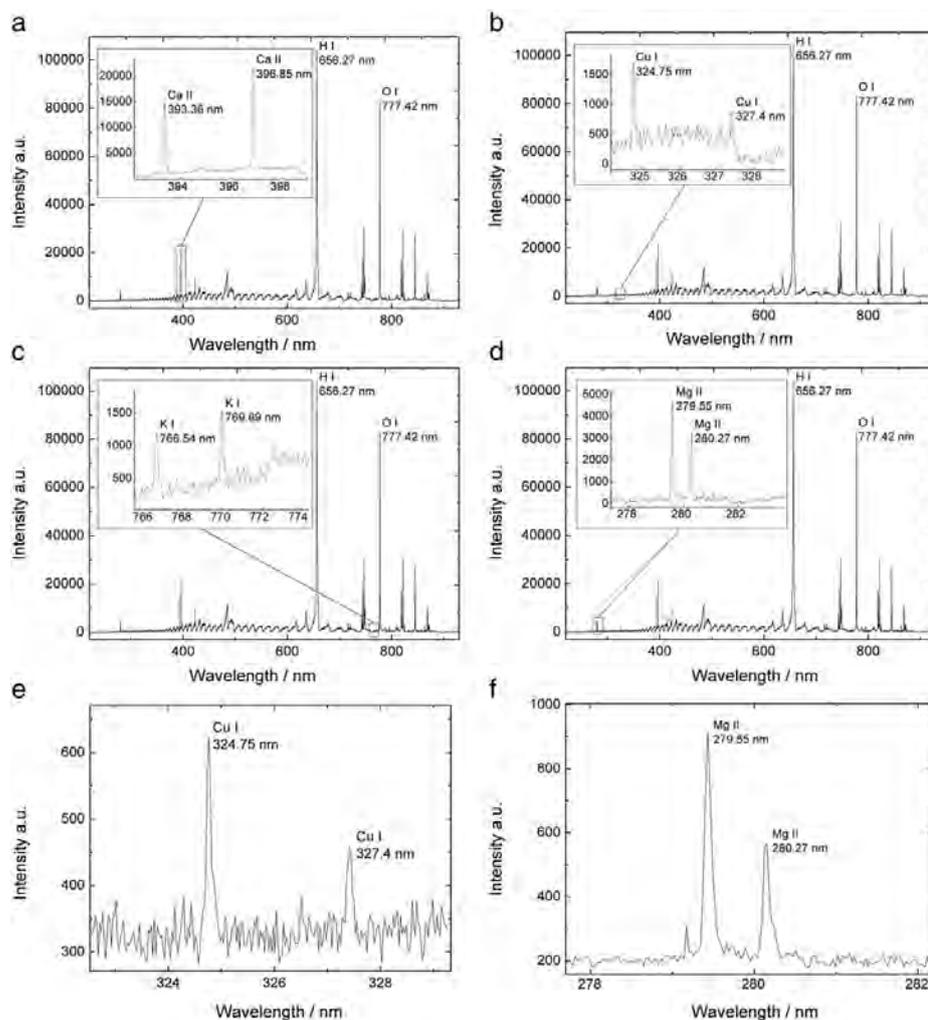


Fig. 6. a–f Segments of LIBS spectra, recorded from water jet and bulk liquid arrangements used in analysis highlighting individual elements—(a) Ca, (b) Cu, (c) K and (d) Mg for water jet and (e) Cu, and (f) Mg for bulk liquid setup (spectral range was limited by using the Czerny–Turner spectrometer).

spectra acquired from “batch 1” and “batch 2” are shown in Fig. 3a and b, respectively. The measured data clearly show the copper binding capacity of *Chlamydomonas* sp. The copper binding capacity was also measured for *T. minutus* (results are not shown) and was found slightly lower when monitoring Mg/Cu ratios (Mg belongs to the alkaline earth metals and its natural average concentration in algal biomass is about 9 mg/kg; Cu belongs to algal microelements with natural concentration of about 0.02 mg/kg [28]). Moreover, from the differences in the ratios of Mg/K (K belongs to the alkali metals and its natural average concentration is about 4 mg/kg [28]) estimated for both the batches we can speculate that during the incubation when Cu was added nutritional conditions within cells have been changed, that could mean Mg was preferentially absorbed in the cells. It could suggest that alkali metals and alkaline earth elements were exchanged with microelements ions from the added solution containing Cu. However, parameters such as concentration of metal in solution, pH, temperature, cations, anions and metabolic stage of the cells can affect above results. Also, a surface

effect image analysis would be very useful so that metal sorption can be thoroughly investigated. Such studies are currently under progression in our laboratories, exploiting LIBS and different imaging approaches.

These preliminary finding suggests that the algal biomass under our study might be potentially suitable for the removal of heavy metal ions from e.g. polluted waters. This was, indeed, as presented above, confirmed by our experiments which determined the binding of Cu metal-ions by the presence of increasing levels of LIBS signal for Cu lines. However, systematic studies are required to investigate mechanism of biofilm sorption further using statistical and chemometrics software, in order to improve on our analytical measurements. Such studies are currently under development in our laboratories, exploiting a selection of LIBS, Raman spectroscopy and LA-ICP-MS approaches. Specifically, we aim to identify and quantify elements within biofilms using LA-ICP-MS technique for the purpose of cross-validation and Raman spectroscopy for algal oil determination.

Fig. 4 shows segments of LIBS spectra, recorded from biofilm sample presenting the capability of LIBS technique to detect different elements within the algal cells. In order to estimate the ablation efficiency and the size of the ablation crater in the algal biofilm the confocal laser scanning microscope (Olympus, LEXT OLS3100, JP) images were obtained. Fig. 5a shows the ablated biofilm with the detail of the craters which were drilled during LIBS measurements. The ablation crater detail is shown on Fig. 5b presenting the cratered depth of about 70 μm and the diameter of about 100 μm (the total thickness of biofilm was estimated of about 150 μm). This translates to an average ablation rate to about 7 μm per pulse, considering that LIBS analysis was provided using 10 laser shots.

### 3.2. Measurements on algal suspension using water jet setup and bulk liquid setup

Here, LIBS experiments were performed on liquid suspensions of algal species. In the water jet example, shown in Fig. 6, biologically relevant elements (Ca, K, Mg) are clearly visible confirming the feasibility of the LIBS water jet setup for qualitative/semi-quantitative analysis of algal suspension in liquids (*T. minutus*). The spectra obtained from the experiments employing water jet setup suggest the possibility to perform detailed semi-quantitative analysis, which is now being investigated in our laboratories. Also, as described above for measurements on biofilms, experiments with algae treated with Cu were performed using the water jet. Fig. 6b displays two Cu lines corresponding to copper that was adsorbed by algae. Consequently, these results were confirmed utilizing the bulk liquid set up; spectra are shown in Fig. 6e–f, confirming that both arrangements are capable to measure traces of Cu in algae suspended in medium. We would like to point out the differences in the LIBS spectra obtained from the liquid jet and from the biofilm samples—intense H and O lines are visible for the liquid jet (Fig. 6a–d) in contrary to the strong lines of Ca for biofilm (Fig. 4).

The successful analysis of the elemental spectral signatures for the two algal samples confirms that LIBS spectroscopy could be used for fast *on-line* characterization of algal species within the industrial environment. Measurements of the selected ranges of interest (e.g. for estimation of Mg/Cu, Ca/Mg, Mg/K ratios for semi-quantitative analysis) can be acquired in a few seconds, depending on the selected range and instrumentation. Thus, LIBS might be capable to offer a real-time, in-situ recognition of algal species in the future. Consequently, these (preliminary) results warrant more extensive investigations with larger collections of algal strains to evaluate the power of LIBS spectroscopy compared to other analytical methods. This research is now being conducted in our laboratories.

## 4. Conclusion

In present study we have demonstrated that it is possible to perform LIBS analysis of macro- and micro-element concentrations in algal samples using the water jet setup, bulk liquid arrangement and by investigations utilizing biofilms. Additionally, the differences recognized using LIBS technique in the spectra of two algal batches, where one batch was exposed to medium containing elevated levels of copper have been shown.

From our feasibility experiments LIBS proves to be a useful tool for detecting absorption of pollutant elements by algae samples, which might be in turn considered as potential candidates for bioremediation. The present study was undertaken as a proof of principle only, and thus we only focused on selected results from the two algal species. The results of our experiments conducted for samples of algal suspensions and algal biofilms confirm LIBS as a fast, real-time technique for remote analysis.

As a natural development we plan to combine LIBS with a technique capable to analyze algal lipids (e.g. Raman spectroscopy) in

order to study the effects introduced by nutrient starvation – which can be detected/monitored using LIBS – on the oil production in selected algal samples. For this, we continue in our efforts to prove analytical potential of combining LIBS with Raman spectroscopy.

We believe that LIBS technique will be of significant assistance to research groups currently being involved in, or intending to join, the quest of sustainable biofuel generation, algal food applications, and bioremediation.

## Acknowledgments

The authors acknowledge the support for O. Samek by a Marie Curie Re-integration Grant (PERG 06-GA-2009-256526). This work also received support from the Ministry of Education, Youth and Sports of the Czech Republic together with the European Commission, the Czech Science Foundation, and the Ministry of Industry and Trade of the Czech Republic (projects ALISI No. CZ.1.05/2.1.00/01.0017, GA P205/11/1687, and FR-TI1/433). The work was also supported by the project “CEITEC—Central European Institute of Technology” (CZ.1.05/1.1.00/02.0068) from European Regional Development Fund.

## References

- [1] I. Khozin-Goldberg, U. Iskandarov, Z. Cohen, LC-PUFA from photosynthetic microalgae: occurrence, biosynthesis, and prospects in biotechnology, *Appl. Microbiol. Biotechnol.* 91 (2011) 905–915.
- [2] P. Jinfen, L. Ronggen, M. Li, A review of heavy metal adsorption by marine algae, *Chin. J. Oceanol. Limnol.* 18 (2000) 260–264.
- [3] X. Deng, X. Fei, Y. Li, The effects of nutritional restriction on neutral lipid accumulation in *Chlamydomonas* and *Chlorella*, *Afr. J. Microbiol. Res.* 5 (2011) 260–270.
- [4] B.N. Tripathi, S.K. Metha, A. Amar, J.P. Gaur, Oxidative stress in *Scenedesmus* sp. during short- and long-term exposure to Cu<sup>2+</sup> and Zn<sup>2+</sup>, *Chemosphere* 62 (2006) 538–544.
- [5] H.V. Perales-Vela, J.M. Pena-Castro, R.O. Canizares-Villanueva, Heavy metal detoxification in eukaryotic microalgae, *Chemosphere* 64 (2006) 1–10.
- [6] P.A. Terry, W. Stone, Biosorption of cadmium and copper contaminated water by *Scenedesmus abundans*, *Chemosphere* 47 (2002) 249–255.
- [7] A.W. Miziolek, V. Palleschi, I. Schechter, *Laser-Induced Breakdown Spectroscopy (LIBS) Fundamentals and Applications*, Cambridge University Press, Cambridge, 2006.
- [8] V.I. Babushok, F.C. DeLucia Jr., J.L. Gottfried, C.A. Munson, A.W. Miziolek, Double pulse laser ablation and plasmas: laser induced breakdown spectroscopy signal enhancement, *Spectrochim. Acta Part B* 61 (2006) 999–1014.
- [9] J. Uebbing, J. Brust, W. Sdorra, F. Leis, K. Niemax, Reheating of a laser produced plasma by a second pulse laser, *Appl. Spectrosc.* 45 (1991) 1419–1423.
- [10] S.J. Rehse, J. Diedrich, S. Palchaudhuri, Identification and discrimination of *Pseudomonas aeruginosa* bacteria grown in blood and bile by laser-induced breakdown spectroscopy, *Spectrochim. Acta Part B* 62 (2007) 1169–1176.
- [11] J. Kaiser, M. Galiova, K. Novotny, R. Cervenkova, L. Reale, J. Novotny, M. Liska, O. Samek, V. Kanicky, A. Hrdlicka, K. Stejskal, V. Adam, R. Kizek, Mapping of lead, magnesium and copper accumulation in plant tissues by laser-induced breakdown spectroscopy and laser-ablation inductively coupled plasma mass spectrometry, *Spectrochim. Acta Part B* 64 (2009) 67–73.
- [12] K. Novotny, J. Kaiser, M. Galiova, V. Konecna, J. Novotny, R. Malina, M. Liska, V. Kanicky, V. Otruba, Mapping of different structures on large area of granite sample using laser-ablation based analytical techniques, an exploratory study, *Spectrochim. Acta Part B* 63 (2008) 1139–1144.
- [13] L. Niu, H.H. Cho, K.S. Song, H.K. Cha, Y. Kim, Y.I. Lee, Direct determination of strontium in marine algal samples by laser-induced breakdown spectrometry, *Appl. Spectrosc.* 56 (2002) 1511–1514.
- [14] Y. Feng, J. Yang, J. Fan, G. Yao, X. Ji, X. Zhang, X. Zheng, Z. Cui, Investigation of laser-induced breakdown spectroscopy of a liquid jet, *Appl. Opt.* 49 (2010) 70–74.
- [15] M. Sadegh Cheri, S.H. Tavassoli, Quantitative analysis of toxic metals lead and cadmium in water jet by laser-induced breakdown spectroscopy, *Appl. Opt.* 50 (2011) 1227–1233.
- [16] P. Yaroshchuk, R.J.S. Morrison, D. Body, B.L. Cahdwick, Quantitative determination of wear metals in engine oils using laser-induced breakdown spectroscopy: a comparison between liquid jets and static liquids, *Spectrochim. Acta Part B* 60 (2005) 986–992.
- [17] O. Samek, D.C.S. Beddows, J. Kaiser, S.V. Kukhlevsky, M. Liška, H.H. Telle, J. Young, Application of laser-induced breakdown spectroscopy to in-situ analysis of liquid samples, *Opt. Eng.* 39 (2000) 2248–2262.
- [18] G. Arca, A. Ciucci, V. Palleschi, S. Rastelli, E. Tognoni, Trace element analysis in water by the laser-induced breakdown spectroscopy technique, *Appl. Spectrosc.* 50 (1997) 1102–1105.

- [19] S.D. Brown, M. Martin, S. Deshpande, S. Seal, K. Huang, E. Alm, Y. Yang, L. Wu, T. Yan, X. Liu, A. Arkin, K. Chourey, J. Zhou, D.K. Thompson, Cellular response of *Shewanella oneidensis* to strontium-stress, *Appl. Environ. Microbiol.* 72 (2006) 890–900.
- [20] N.K. Rai, A.K. Rai, LIBS—an efficient approach for the determination of Cr in industrial wastewater, *J. Hazard. Mater.* 150 (2008) 835–838.
- [21] B.A. Rasala, M. Muto, J. Sullivan, S.P. Mayfield, Improved heterologous protein expression in the chloroplast of *Chlamydomonas reinhardtii* through promoter and 5 untranslated region optimization, *Plant Biotechnol. J.* 9 (2011) 674–683.
- [22] J.B.K. Park, R.J. Craggs, A.N. Shilton, Wastewater treatment high rate algal ponds for biofuel production, *Bioresour. Technol.* 102 (2011) 35–42.
- [23] T. Rezanka, M. Petránková, V. Cepák, P. Přebyl, K. Singler, T. Cajthaml, *Trachydiscus minutus*, a new biotechnological source of eicosapentaenoic acid, *Folie Microbiol.* 55 (2010) 265–269.
- [24] O. Samek, A. Jonáš, Z. Pilát, P. Zemánek, L. Nedbal, J. Tříška, P. Kotas, M. Trtílek, Raman microspectroscopy of individual algal cells: sensing unsaturation of storage lipids *in vivo*, *Sensors* 10 (2010) 8635–8651.
- [25] E. Kataran, P. Watnick, Signals, regulatory networks, and materials that build and break bacterial biofilms, *Microbiol. Mol. Biol. Rev.* 73 (2009) 310–347.
- [26] M. Galiova, J. Kaiser, K. Novotny, M. Hartl, R. Kizek, P. Babula, Utilization of laser-assisted analytical methods for monitoring of lead and nutrition elements distribution in fresh and dried *Capsicum annuum* l. leaves, *Microsc. Res. Tech.* 74 (2011) 845–852.
- [27] M. Galiova, J. Kaiser, K. Novotny, M. Ivanov, M.N. Fisakova, L. Manicni, G. Tromba, T. Vaculovic, M. Liska, V. Kanicky, Investigation of the *osteitis deformans* phases in snake vertebrae by double-pulse laser-induced breakdown spectroscopy, *Anal. Bioanal. Chem.* 398 (2010) 1095–1107.
- [28] I. Michalak, K. Chojnacka, K. Marycz, Using ICP-OES and SEM-EDX in biosorption studies, *Microchim. Acta* 172 (2011) 65–74.

## Article 2

P. Pořízka, P. Prochazková, D. Prochazka, L. Sládková, J. Novotný, M. Petrilak, M. Brada, O. Samek, Z. Pilát, P. Zemánek, V. Adam, R. Kizek, K. Novotný, and J. Kaiser. Algal biomass analysis by laser-based analytical techniques—A review. *Sensors (Switzerland)*, 14:17725–17752, 2014.

DOI: 10.3390/s140917725; 36 citations

My first review article had the biggest impact on the LIBS community. It has summarized the utilization of laser-based spectroscopic techniques as perspective tools for the detection of chemical composition of algae, investigation of their nutrition and classification of individual algal strains. Spectroscopic techniques (LIBS, LA-ICP-MS and Raman spectroscopy) were selected for their complementarity in terms of provided information and possibility to implement them in a consequent experiment. A comprehensive review of algal biomass analysis was focused on the application of spectroscopy in the fields of biofuels and bioremediation.

As a feasibility study, principal component analysis (PCA) was used for visualization of four algal strains (*Chlarydomonas reinhardtii*, *Chlorococuum zurek*, *Desmodesmus quadratic*, *Haematococcus pluralis*.) Their full discrimination in LIBS spectra was observed based on variations in intensities of selected spectral lines of Ca, K, Mg, and Na.

Review

## Algal Biomass Analysis by Laser-Based Analytical Techniques—A Review

Pavel Pořízka <sup>1,2</sup>, Petra Prochazková <sup>3,5</sup>, David Prochazka <sup>1,3</sup>, Lucia Sládková <sup>3</sup>, Jan Novotný <sup>1,3</sup>, Michal Petrilak <sup>3</sup>, Michal Brada <sup>3</sup>, Ota Samek <sup>4</sup>, Zdeněk Pilát <sup>4</sup>, Pavel Zemánek <sup>4</sup>, Vojtěch Adam <sup>3</sup>, René Kizek <sup>3</sup>, Karel Novotný <sup>3,5</sup> and Jozef Kaiser <sup>1,3,\*</sup>

<sup>1</sup> Institute of Physical Engineering, Faculty of Mechanical Engineering, Brno University of Technology, Technická 2896/2, Brno 61669, Czech Republic; E-Mails: pavel.porizka@ceitec.vutbr.cz (P.P.); david.prochazka@ceitec.vutbr.cz (D.P.); jan.novotny@ceitec.vutbr.cz (J.N.)

<sup>2</sup> BAM—Federal Institute for Material Research and Testing, Richard Willstätter-Straße 11, Berlin D-12489, Germany

<sup>3</sup> CEITEC—Central European Institute of Technology, Brno University of Technology, Technická 3058/10, Brno 61600, Czech Republic; E-Mails: pekucero@seznam.cz (P.P.); lucia.sladkova@ceitec.vutbr.cz (L.S.); michal.petrilak@ceitec.vutbr.cz (M.P.); michal.brada@ceitec.vutbr.cz (M.B.); vojtech.adam@mendelu.cz (V.A.); kizek@sci.muni.cz (R.K.); codl@sci.muni.cz (K.N.)

<sup>4</sup> Institute of Scientific Instruments of the Academy of Science of the Czech republic, v.v.i., Královopolská 147, Brno 612 64, Czech Republic; E-Mails: osamek@isibrno.cz (O.S.); pilat@isibrno.cz (Z.P.); zemanek@isibrno.cz (P.Z.)

<sup>5</sup> Department of Chemistry, Faculty of Science, Masaryk University, Kotlářská 2, Brno 61137, Czech Republic

\* Author to whom correspondence should be addressed; E-Mail: kaiser@fme.vutbr.cz; Tel.: +420-54-114-2846.

Received: 28 April 2014; in revised form: 5 September 2014 / Accepted: 11 September 2014 /

Published: 23 September 2014

---

**Abstract:** Algal biomass that is represented mainly by commercially grown algal strains has recently found many potential applications in various fields of interest. Its utilization has been found advantageous in the fields of bioremediation, biofuel production and the food industry. This paper reviews recent developments in the analysis of algal biomass with the main focus on the Laser-Induced Breakdown Spectroscopy, Raman spectroscopy, and partly Laser-Ablation Inductively Coupled Plasma techniques. The advantages of the selected laser-based analytical techniques are revealed and their fields of use are discussed in detail.

**Keywords:** Laser-Induced Breakdown Spectroscopy; LIBS; Laser-Ablation Inductively Coupled Plasma coupled with Mass Spectroscopy and Optical Emission Spectroscopy; LA-ICP-MS; LA-ICP-OES; ICP-OES; Raman spectroscopy; algae; algal biomass; biofuel; bioremediation

---

## 1. Introduction

Expanding economies are still technologically dependent on crude oil, while the decreasing amount of oil deposits moves up prices of crude oil and related gasoline [1]. Trends are observed in searching for alternatives to fossil fuels. Another challenge lies in satisfying global energy needs in a way that would decrease the level of environmental pollution. This challenge can be faced with renewable energy sources [2]. Algae are a possible energy source which could solve both issues. Amongst prospective alternatives to fossil fuel, algae have become one of the most significant without competing for arable land [3–5]. Algae convert solar energy into lipids, carbohydrates and proteins via photosynthesis and then further processing of these primary metabolites can take place. Algae have a per-acre per-year yield that is 200 times higher than the best-performing plant/vegetable oils [6], while some algal strains are also capable of doubling their mass several times per day [7].

The complex study of algae as a future biodiesel and biomass feedstock was initiated by Williams and Laurens [3]. They summarized the collection of oil-producing microalgae. Great emphasis was placed on understanding the biochemistry of the algal strains and on the development of algal production systems. They pointed out the influence of the biochemical composition of the biomass (*i.e.*, the lipid content) on the economics of the biofuel production. They concluded that algal biofuels are potential valuable alternative to “traditional” biofuels. However, the non-profitable economics (considering the oil prices in the 1990s) of the biofuel production caused the termination of the long-term research program funded by the U.S. Department of Energy [4]. The possible replacement of fossil fuels by algal biofuels is a matter of future research and commercialization of the production process [8,9]. Algae may provide an effective solution, but several challenging aspects need to be overcome [3,10,11], e.g., light use efficiency, high amount of oil production in the algal cell, daily crop harvest, effective algal biomass to 3rd generation biofuel conversion, and the improvement of the entire system economics [3].

Algae can be grown in open pond systems [6,12–14] and bioreactors [15] with a possibility of a daily harvest. Algal ponds can as well serve for waste water treatment [13,14], which should result in the cost reduction. Every algal strain has to be grown under optimized conditions to obtain high amounts of crop harvest per day, e.g., sufficient sunlight, nutrients, and protection against natural pests. Individual algal strains have different properties and react differently to the conditions in which they grow. On the other hand, algal strains can easily adapt to their new environment [3]. Bioremediation is the ability of the algal strain to grow in polluted water and even prosper to increase its yielded lipid content [16–18]. Algae can also be used as bio-indicators of water pollution level, e.g., to determine the presence of heavy metals [19,20]. In order to reduce expenses for algae cultivation, it is possible to use various sources for nutrient control, including the agricultural waste waters [5]. Environmental and climate changes

can be traced in coralline algae—in mineralized algae species that are an excellent record of this information [21].

It is noteworthy that algae may be also used in other fields with potential economic impact [22,23], such as the food, cosmetic and pharmaceutical industries. A review covering majority of literature sources concerning marine algae products was done by Blunt *et al.* [24]. One of the studies is focused on the possible processing of an algal cell in order to yield more products at once [25]. Cardozo *et al.* [22] report on the importance of algae in the food industry in numerous countries, where the emphasis has moved from wild harvests to farming and controlled cultivation in order to produce valuable products on a large scale. The investigation of the algal chemical composition and related products are nowadays promising research areas in the pharmaceutical industry. Algal products may be used in cancer [26] and HIV [27] treatment. In civil engineering, flat panel airlift reactors for lipid production by the algal strain *Chlorella vulgaris* were already installed [28]. The reactor was designed as the renewable energy stock based on the algal biomass production. Another step in this research and development is the construction of a building with a bioreactor facade in Hamburg, Germany [29], where the energy is supplied by the growth of the algal biomass.

Despite the main aim of this article we briefly review the analysis of algae utilizing other techniques. Light Detection and Ranging (LIDAR) is an optical remote sensing technique that may be used for the analysis of larger areas [30]. Phytoplankton in the delta of the river Po was monitored utilizing fluorescence LIDAR systems placed on a van [31], oceanographic ship and airplane [32]. Laser-induced fluorescence has a long history of applications in the detection of marine algae; in the year 1972 a series of measurements was made employing a dye laser, ruggedized for airborne use [33].

Algae are mainly composed of carbohydrates, proteins, nucleic acids, and lipids, where carbohydrates and lipids are responsible for the energy storage [3]. It was reported that the lipid content in the lipid bodies depends on the growth and the nutrient status of algae. Various techniques are employed for obtaining elemental or molecular information of an algal strain. For instance, nuclear magnetic resonance (NMR) spectroscopy was applied to the analysis of plants, fungi and algae by Martin [34]. A non-invasive *in vivo* measurement employing NMR spectroscopy revealed details of the nitrogen and carbon metabolism in real time [35,36]. NMR spectroscopy was utilized to give characteristic fingerprints of the lipid extractions from algal samples, while marine algal strains and samples from the Lagoon of Venice were compared [37]. Danielewicz *et al.* [38] studied the intact triacylglycerol composition of four microalgae species using MALDI-TOF-MS (matrix-assisted laser desorption and ionization time-of-flight mass spectrometry) and <sup>1</sup>H-NMR spectroscopy. Moreover, MALDI-TOF-MS was employed in other studies for comparison of various algal strains [39–41].

The fast estimation of the algae lipid content is possible by employing Raman spectroscopy. Raman spectra—in the sense of a fingerprint—give information about the saturated and unsaturated fatty acids in the lipid body [42–45]. Samek *et al.* [42] showed that it is feasible to calculate the iodine value (IV) from Raman spectra. IV quantifies the degree of unsaturation and is mainly used in the biodiesel industry [46]. Moreover, the analysis of fatty acid composition in algae by gas chromatography—mass spectrometry (GC-MS) is also possible; however, it is a time-consuming technique [47,48].

Atomic spectroscopy techniques are in general the most commonly used for elemental analysis [49]. Inductively Coupled Plasma techniques, among others, have been the most commonly used technique in any field of interest. It was shown that Laser-Ablation Inductively Coupled Plasma Mass Spectrometry

(LA-ICP-MS) is an appropriate method for detection of environmental changes over the year (or even decades) in coralline algae [50–55]. This method is characterized by high spatial-resolution that is required to detect long-lived coralline algae, because seasonal growth increment widths range approximately from 230 to 330  $\mu\text{m}/\text{yr}$  [50].

Winefordner *et al.* [49] proposed another laser-based technique, Laser-Induced Breakdown Spectroscopy (LIBS), as a future superstar for elemental analysis of various samples in any state of matter. LIBS has the added advantage that the analysis can be performed remotely, provided optical access can be established between the instrument and the target. When analyzing a sample with other techniques the sample has to be presented to the instrument. This of course is not the case of LIBS, because all interactions between the system and the target can be purely optical. Therefore, LIBS has an extremely competitive position and excels if remote, *in-situ*, real time analysis is required. For instance, in analyzing a water jet containing algal suspension where only optical access using either an optical fiber or a telescope can be used. The most appropriate applications are thus found to be those which prefer remote quantitative or qualitative analysis, without any physical contact with the sample.

Moreover, the sample does not have to be prepared for analysis using solvents and any surface contaminants can be ablated off the sample before carrying out a measurement. This makes LIBS ideal for algal strain analysis where the target may be in the form of dried biomass as algal suspension. Specially engineered systems can be designed and assembled for each analytical problem allowing fast decisions to be made concerning the identification/analysis of target materials which can then be immediately analyzed, sorted and labelled. In this review, LIBS is critically evaluated and considered as a mature technique capable of competing with other techniques for elemental analysis.

Furthermore, the combination of Raman spectroscopy (chemical composition) and LIBS (elemental composition) can be complementary, increasing the information power [56]. This combination of techniques, called a hyphenated or tandem approach, has been already successfully used for the analysis of minerals [57] and cultural heritage objects [58]. Hoehse *et al.* [59] constructed a LIBS-Raman system with a two-arm Echelle spectrometer equipped with single CCD camera. Pořízka *et al.* [60] used laser-based techniques (LIBS and Raman spectroscopy) in tandem for obtaining both elemental and molecular information of the algal strain *Trachydiscus minutus*. Raman spectroscopy can be used to obtain the molecular composition of the sample under study, e.g., information on the lipid content inside algal cells. The elemental composition can be observed employing LIBS or LA-ICP-MS. As was mentioned above, LIBS is an emerging technique for elemental analysis, with its main advantage being the possibility of fast *in-situ* measurement. It has to be noted that LA-ICP-MS or LA-ICP coupled to Optical Emission Spectrometry (LA-ICP-OES) can be advantageously used also to validate the LIBS outcomes, mainly in the first stages of the research and development of new LIBS applications. The classification of the algal strains based on their spectra, *i.e.*, spectrochemical fingerprint in the sense of elemental or molecular composition, can be also provided by employing standard chemometric algorithms, such as principal component analysis (PCA) [61] and partial least squares (PLS) algorithms [61,62]. Chemometrics and their applications are further discussed in the text.

In this review the literature was surveyed for recent developments and results in research utilizing selected laser-based techniques (mainly LIBS and Raman spectroscopy, partly LA-ICP-MS) for analysis of the algal biomass or calcified coralline algae. ICP-OES was used as a supervising technique for analyzing LIBS results with chemometric.

Using LIBS and Raman spectroscopy techniques, one is able to monitor *in-situ*, on-line and in real time the spectral evolution of the major/minor elements (LIBS) and in addition chemical composition of the sample (Raman spectroscopy). Moreover, in these approaches, knowledge about the relation between elemental/chemical composition and spatial location is achieved. This enables to measure time-course data so that monitoring changes of sample over time (which could be related to spatial position) introduced by environmental/nutritional influences. On the contrary, when other approaches are used for analysis where the part of a sample is dissolved/pelleted and consequently analyzed using for instance methods based on atomic absorption spectroscopy (AAS) or MALDI information about spatial location and time evolution is hard to obtain or even completely lost.

## 2. Laser-Induced Breakdown Spectroscopy

The spectrochemical analytical technique LIBS, which is based on generating a laser-induced plasma (LIP) by high energy laser pulses and subsequent time-resolved spectral analysis of the LIP emission, can be used to analyze materials in any state of matter [63–67]. A LIP spectrum containing atomic and ionic emission lines may provide qualitative and quantitative information about the elemental composition of the sample in real-time and *in-situ*. In recent years, LIBS technique has gained its position among other spectroscopic techniques due to its advantages; such as simple and robust instrumentation, fast and precise analysis, no need for the sample preparation, the capability of on-site application and remote/stand-off detection [49]. Due to the relative simplicity of the whole measurement process, a movable remote (or stand-off) system can be constructed and employed for the analysis of environmental samples [68,69].

The utilization of LIBS in various fields is summarized in the review articles [56,70–72]. The technique has already proved its capability for the analysis of biological samples [73,74] and in biomedical applications [75]. Review on the femtosecond (fs) LIBS (physics of the laser-induced plasma, applications and perspectives) was introduced in [76]. The utilization of fs-laser source in a LIBS measurement leads to significant suppression of the matrix effect. Though, there exist multi ways how to overcome the matrix effect occurring in ns-LIP [56], e.g., Laser-Ablation LIBS (LA-LIBS).

To the best of our knowledge, only a few pioneering works have recently been performed in the analysis of algae employing LIBS. Surveying the literature for the LIBS measurements, multiple arrangements and several approaches can be found for direct measurements of algal strains. One possibility is to dry the algal biomass to produce a thin film [60] or to press dried algal biomass into the pellets [77,78]. Garcimuno *et al.* [77] measured natural watercourse algae with added standard solutions of Cu. The analytical figure of merit, limits of detection, were obtained and claimed to be in the units of ppm. Niu *et al.* [78] produced internal standards by adding known amount of Sr into dried biomass of two different algal strains, *Chlorella* and *Sargasso*. Both strains were standardized samples obtained from National Institute for Environmental Studies (NIES) in Japan. After drying, the algal powder was pressed into pellets, and measured by LIBS. The amount of Sr in the unknown sample was then successfully evaluated using the constructed calibration curves. The approach pressing the pellets prior the LIBS analysis, however, is time-demanding and not-applicable for *in-situ* and time-course analysis [77,78]. The most straightforward way is to measure the algae directly, in the water suspension. Pořízka *et al.* [60] observed the elemental composition of algal strain *Trachydiscus minutus* (Bourrelly)

measured with LIBS in three ways, (i) the algal sample was dried and deposited into the thin biofilm; (ii) the suspended algae was measured in the liquid jet; and (iii) in bulk (where laser-induced plasma was produced on the surface of the liquid). Elements of biological significance (Ca, K, Mg, and Na) were determined as well as the trace amounts of potentially toxic metal (Cu). The results of LIBS analysis of algal strains are listed in Table 1.

**Table 1.** The list of articles focused on the analysis of algae employing LIBS.

Algal Strain	Ref.	Pretreatment of Algae	Matrix Elements	Minor Elements	Trace Element	LOD (ppm)
<i>Trachydiscus minutus</i> (Bourrelly)	[60]	dried biofilm. liquid jet, surface of bulk liquid	Ca, K, Na, Mg	-	Cu	-
watercourse algae strain not specified	[77]	dried and pressed to pellets	Mg, K, Na, Fe, Si	Al, Mn, Ti, V	Cu, Cr, Pb, Zn	Cu: $9 \pm 2$
<i>Chlorella</i> , NIES No. 3 <i>Sargasso</i> , NIES No. 9	[78]	dried and pressed to pellets	Ca	-	Sr	-

One consideration that needs attention when using LIBS for quantitative analysis are “matrix-effects”, and these have been the subject of much discussion in all branches of spectrochemical analysis. These are the effects on the spectra associated with the combined physical and chemical properties of the target which result in different dynamics of laser/matter interaction and consequent ablated mass values, plasma formation, and its properties. They cause outliers in calibration plots if the samples are not chosen with similar composition. Extreme cases are seen when calibration plots for the same element are obtained from totally different materials. To get around these problems the composition of the calibration samples should closely match, in the sense of matrix elements content, that of the material to be analysed. Otherwise a different analytical approach has to be followed.

The composition of the sample, mainly in the sense of matrix or macro elements, is crucial in the laser/matter interaction and consecutive LIP formation and emission. The quantitative analysis of trace elements content in the algal samples is limited due to the significant influence of the matrix effect on the intensity of trace element lines. The same amount of an analyte in samples with various matrices may result in the significantly different intensity of corresponding spectral line. Then general calibration of the system for various matrices is therefore not possible. For this reason, the ways of compensating or even avoiding the matrix effect should be considered. The performance of LIBS in quantitative analysis may be improved by multivariate algorithms [64]. Though, multivariate algorithms, such as principal component regression (PCR) and partial least squares regression (PLSR), may compensate the matrix effect only to a certain extent. However, when the classification prior to the quantitative analysis is considered the main benefit may be to discriminate the samples based on their matrix elements content. Then the matrix effect may be suppressed while the calibration curve is constructed only for particular group, *i.e.*, samples with limited range of variation in the composition of matrix elements.

LIBS analysis of liquid samples is very challenging for the essential problem arising from the laser/liquid interaction. Improvement in the LIBS instrumentation for the measurements of samples in the form of liquid solutions and suspensions should lead to the improvement in the sensitivity and

repeatability of the technique. Moreover, the density of the liquid suspension has to be taken into account when different ratios (water to algae) affects also the matrix. The measurement can be performed in the matrix assisted mode, as was presented in [79], where algae were deposited on the surface which matrix is considered to be supreme, or simply measured in the form of dried biofilm [60]. With this approach the promising improvement in the sensitivity and repeatability of the LIBS setup is expected. This approach was adapted in our recent measurements [60] where alga *Trachydiscus minutus* (Bourrelly) was deposited on a microscope slide and dried. Thin film of algal biomass was then analyzed utilizing a double pulse LIBS (DP LIBS) technique. Elements of biological significance (Ca, K, Mg, and Na) were detected with the highest signal for Ca (II) doublet (393.4 and 396.9 nm). Furthermore, DP LIBS was utilized for higher sensitivity and lower amount of ablated mass. Moreover, table-top LIBS setup offers in general satisfying repeatability and reproducibility of the measurement with limits of detection under the ppm level.

The ratio of algae to water content may be controlled if the algal strains are prepared under the laboratory conditions. Further problems arise when the algal strains are collected from their natural environment, which affects the elemental composition of an algal strain. Though, influence of environmental parameters could be helpful when the classification of the algal strains is of an issue, *i.e.*, in the provenance study. Then, an algal strain may be classified when the composition of matrix elements (Ca, K, Mg, and Na) is affected by the surrounding environment.

Despite the increasing popularity of LIBS within the spectroscopic groups dealing with various application fields, the use of LIBS for the analysis of algal biomass remains still unexplored. Published papers on the analysis of algae samples employing LIBS, however, proved the capabilities of this analytical technique. Both reproducibility and good sensitivity were reached when algal pellets were measured with additional trace amounts of toxic heavy metals. The approach for measuring algae in liquid suspension further strengthened the position of LIBS for *in-situ* measurement of biological samples—from a liquid jet the instrument (see further in the text) can directly measure *in-situ*, in real-time and on-line changes in algal elemental composition within the bioreactor. LIBS is a promising technique for analyzing the algae elemental composition of samples *in-situ* and in real-time, in their natural habitat. It can be useful for monitoring the cultivation process of algae and for evaluating environmental pollution.

It should be noted that in different fields, e.g., in biofuel production, the elemental analysis of algal samples is not the primary objective with respect to the molecular analysis. As detailed in Section 2.2, LIBS, with certain limitations, is capable also of direct molecular analysis. Utilizing LIBS solely for elemental analysis, however, has applications in the analysis of drinking water, evaluation of the waste waters and its treatment control, the so-called bioremediation, trace elements detection in marine algae, *etc.*

### 2.1. Laser-Induced Breakdown Spectroscopy of Liquid Samples

An effective LIBS arrangement using a vertical thin liquid flow for measurement of trace amounts in water solutions has already been used [60,80]. Thus, LIBS has proved its applicability for detection of trace elements in aqueous samples, but for the rapid *in-situ* and real-time analysis of biological samples suspended in water, *i.e.*, algae, it is necessary to further develop LIBS systems, so that they are capable of measuring samples in the liquid state of matter.

Despite the general problems with plasma generation in liquid samples, the capability of LIBS for analysis of liquid samples has been tested and improved for more than thirty years. A detailed literature survey on the analysis of liquids is beyond the scope of this review, however, more comprehensive review on the analysis of liquid can be found in [56]. Practically, three different approaches of liquid sample measurements can be employed: (1) creation of the plasma in a bulk of the liquid; (2) on the steady surface of the liquid; or (3) on the surface of a thin laminar jet.

Both approaches, in bulk and on surface analysis, suffered from the sedimentation process when the suspended specimen in the liquid settled down to the bottom, making the sample inhomogeneous and the measurement irreproducible. This fact led to the introduction of the thin laminar vertical flow of a liquid [81–85]. Limits of detection found in selected articles are in the tenths or units of ppm for heavy metals (Pb, Cu, and Cr) and elements of biological significance, matrix elements of algae, (Ca, Na, Mg, and K). Preliminary LIBS measurements of algae suspended in liquid suspension utilizing the liquid jet and steady surface approaches were reported [60]. In this article, the peaks of H $\alpha$ , H $\beta$  and O lines were obtained with highest intensity.

### 2.2. Laser-Induced Breakdown Spectroscopy for Molecular Analysis

The analysis of a LIP emission provides information on the elemental composition of the sample. Molecular bands can be as well detected in a LIP radiation [56]. The molecular structural information of a sample is broken after the impact of a laser pulse. The newly formed LIP is composed of ions, atoms, and electrons. As the LIP cools down, ions recombine and molecules may be formed further in LIP temporal evolution. Consequently, the radiation corresponding to excitation states of molecular bands is detected. The molecular bands occur in later stages of the LIP formation. Debras-Guedón and Liodec [86] made series of measurements of molecular radiation, CN and AlO. However, CN bands were present in the plasma of carbon samples. Different timing of the LIBS experiment should be utilized for molecular analysis compared to conventional elemental analysis. Elements originating from the ambient gas surrounding the sample are as well ablated to create a consequent LIP. Those elements then react with the elements from the sample and form molecular bonds. Then, detected molecular radiation does not directly reflect the molecular structural information of the sample. Therefore, significant uncertainty may be introduced to the computation. Molecular analysis is not so frequent in LIBS applications, though this kind of analysis has already been tested. It was already proved that the CN band is a consequence of recombination among C<sub>2</sub> in the plasma, *i.e.*, ablated C from the sample, and N<sub>2</sub> originating from the ambient air [87]. Moreover, Baudelet *et al.* [88] in their work suggest the CN band as a reliable marker for the observation of biological samples.

Doucet *et al.* [89] coupled LIBS measurement with chemometrics (principal component regression and partial least squares regression) in order to obtain more reliable quantitative molecular prediction. In this work they analyzed 18 standardized pharmaceutical samples with emphasis on CN band emission. Kongbonga *et al.* [90] analyzed different types of oils and saccharose dissolved in water with an emphasis on direct detection CN (in the region 388 nm) and C<sub>2</sub> (516.6 nm) bands. The attempt to correlate the intensity of molecular band with the amount of fatty acid in the sample was done. However, no calibration curve was given due to a low range in the amounts of fatty acids. The detection of CN and C<sub>2</sub> bands together with the theory of the chemical processes involved in forming those molecular bands in LIP were presented by several authors [91–95]. In the wider context, the detection of molecular bands (e.g., C<sub>2</sub>) utilizing LIBS could be correlated under well specified circumstances to the real amount of fatty acids.

Utilization of laser-based methods, LIBS and Raman spectroscopy, for obtaining complete chemical information of algae has already been published [60]. In this work, the analysis of molecular information was delivered by Raman spectroscopy. Utilization of Raman spectroscopy with LIBS beneficially in tandem was already reviewed by Hahn and Omenetto [56] for various applications (such as archeology, cultural heritage, mineralogy and soil analysis). Nevertheless, based on the proposed theory, LIBS can—under well specified circumstances—provide both elemental and molecular information. There are several molecular bands (CN, C<sub>2</sub>, CO, CO<sub>2</sub>) which can be detected in the LIP radiation. The concentration of molecules within a LIP could be obtained when the LIBS measurement is supervised with Raman spectroscopy or GC-MS. Advanced statistical algorithms, such as PLS and PCR (principal components regression), could be used for that purpose.

Concluding the LIBS section, LIBS instrument could be used in bioremediation and environmental pollution monitoring. The sensitivity of the technique is satisfactory with the detection limits in the units of ppm. However, a robust LIBS setup with good reproducibility has to be constructed. As it was concluded by Hahn and Omenetto [56], the LIBS device is capable of quantitative analysis, however it is considered to be the only vulnerable feature of LIBS, therefore further research should be concentrated in this direction. Moreover, the matrix effect is of an issue when the quantitative analysis of the trace element is needed [56,64]. Nevertheless, it is possible to avoid or compensate the matrix effect to a certain extent in many ways, e.g., LA-LIBS, fs-LIBS, matrix assisted LIBS, and chemometric algorithms. Despite its limitations, LIBS is capable of direct and fast *in-situ* analysis without any need of sample preparation. Moreover, it is possible to classify various samples based on their chemical fingerprint provided by LIBS.

The performance of a LIBS device for *in-situ* analysis should be adapted to a case study rather than to a general use. Then, LIBS should be in the first stage of the research supervised with another technique (such as ICP-OES, GC-MS, *etc.*) to obtain the reference results and then to construct the supervised library of algal strains. Consequently, LIBS instrument can provide reliable real-time analysis.

### 3. Laser Ablation Inductively Coupled Plasma Based Techniques

To quantify the total content of elements in algal samples, numerous analytical techniques have recently been employed. Most of them, such as solution analysis by ICP-OES and ICP-MS, require sample decomposition and dissolution. The main drawback is the relatively demanding and laborious

sample pre-treatment. After rinsing in ultra-pure water to remove salts and oven-drying, the samples are homogenized, grinded and then weighed [96,97]. Some samples are extracted [98]. The next step is the decomposition procedure, usually the acid digestion [99] or microwave assisted acid digestion [98]. Other treatments such as the slurry sampling technique, acid leaching or the enzymatic hydrolysis can be used [100]. These procedures lead to a total dissolution of the biological materials. Due to the complete decomposition of biological materials, however, the spatially-resolved analysis of elements cannot be carried out.

The solution ICP-OES technique was used to determine major (such as Ca, K, Mg) and trace elements (such as Zn, Cr, Co) in edible algae [97–100]. Using ICP-OES Michalak *et al.* [100] observed the differences of concentrations of elements of marine edible algae during the annual period in different parts of Baltic Sea. Perez *et al.* [21] utilized ICP-OES of algae from three different stations from Patagonia (Argentina) to detect the degree of contamination caused by human activities and to study the seasonal differences between Cd and Pb. ICP-OES is distinguished by ability of multi-elemental analysis of biological (algal) samples with relatively high-sensitivity and rapidness over wide concentrations ranges [98]. Better sensitivity can be acquired by solution ICP-MS. It is a very sensitive and precise analytical technique that allows simultaneous determination of trace and ultra-trace elements in algae with detection limits in the order of  $\text{ng}\cdot\text{g}^{-1}$  [101]. Van Netten *et al.* [102] utilized ICP-MS to control heavy metals and radioactive isotopes in edible marine algae. The amounts of heavy metals in edible algae have to be controlled because of algal high affinity to heavy metals. Rodenas de la Rocha *et al.* [96] have also employed ICP-MS to analyze different elements in edible algae.

The analysis of biological samples without laborious decomposition can be performed using LA-ICP-MS/OES. These analytical techniques are widely used for a trace elemental analysis of solid samples with high spatial resolution (typically below  $20\ \mu\text{m}$  [50]). Moreover, LA-ICP-MS allows also a multi-element analysis of biological samples with no or little sample preparation and enables rapid analysis in real time with high spatial resolution. This advantage enables the observation of the evolution of appropriate biological samples (e.g., calcified algae) during their lifetime, which aids in the understanding their life cycle (series) or living conditions.

LA-ICP-MS was used on coralline red algae to detect climatic condition changes. Coralline red algae represent an ideal organism that occurs in mid- to high-latitude oceans. Their asset is their longevity and their incremental growth pattern. They are widely distributed in the coastal regions worldwide. Part of their skeleton is constituted of high content Mg-calcite. This skeleton grows over their lifespan. Coralline red algae do not suffer from the drawback of the ontogenetic growth trend [103]. Usually, to detect climatic condition changes, the most plentiful species of *Clathromorphum* from North Pacific Ocean are used; they can grow up to 850 years (based on radiometric dating). They can record climate information during an annual period [50–54].

The ratio of Mg/Ca is used to record temperature variation in different marine organisms [97,102]. For the detection of the sea-surface temperature Hetzinger [55] utilized coralline red alga, *Clathromorphum nereostratum*, which archives the environmental information of seawater with a high temporal resolution during its growth. Gamboa *et al.* [53] compared by LA-ICP-MS the Mg/Ca ratio in coralline red algae *Clathromorphum compactum* from two sites within the same region and showed that algae can be used as a recorders of past temperature variability. Halfar *et al.* [51] utilized *Clathromorphum compactum* to employ growth increment widths as a temperature proxy by

LA-ICP-MS. They investigated the relationship between the growth and the environmental parameters. Averaged results of multiple growth increments show strong correlations with annual sea surface temperature. They showed that the highest growth rates are observed during the summer months when the sea-surface temperatures and the light intensities are the highest.

Chan *et al.* [52] and Hetzinger *et al.* [54] used LA-ICP-MS to study the Ba/Ca ratio variations in *Clathromorphum nereostratum*, and investigated temporal salinity changes. Gamboa *et al.* [53] utilized LA-ICP-MS to determine Mg/Ca ratios of *Clathromorphum compactum* to understand the North Atlantic Oscillation. Hetzinger *et al.* [50] investigated algal species *Clathromorphum compactum* and *Clathromorphum nereostratum* by LA-ICP-MS in order to compare the ratios of Mg/Ca, Sr/Ca, U/Ca and Ba/Ca. The temperature dependence of Sr/Ca and Mg/Ca ratios was evidenced. The results show that the Sr content into algal calcite is dependent on the seawater temperature. The relationship between Mg/Ca, U/Ca and Ba/Ca ratios and the sea surface temperature was not proved.

The reported limits of LA-ICP-MS detection are under the ppm range, summarized in Table 2. For the analysis of algal samples, *i.e.*, measuring the concentrations of  $^{24}\text{Mg}$ ,  $^{43}\text{Ca}$ , and  $^{137}\text{Ba}$ , an Agilent 7500ce Quadrupole ICP-MS coupled with a New Wave Research UP 213 laser ablation system (213 nm, ND:YAG laser) was used [51–53]. The carrier gas was helium and the utilized laser energy density was  $6 \text{ J/cm}^2$ . The scan speed was  $10 \text{ }\mu\text{m/s}$ , the spot size was  $65 \text{ }\mu\text{m}$ , and the pulse rate was 10 Hz. NIST SRM 610 (U.S. National Institute of Standard and Technology Standard Reference Material) glass reference material was utilized as an external standard [51,52].

**Table 2.** Limits of detection (LOD) obtained by LA-ICP techniques taken from selected references.

Samples	Isotope	LOD [ppm]	Reference
<i>Clathromorphum compactum</i> , <i>Clathromorphum nereostratum</i>	$^{24}\text{Mg}$	0.16	[50]
	$^{43}\text{Ca}$	54.9	
	$^{88}\text{Sr}$	0.04	
	$^{238}\text{U}$	0.016	
	$^{137}\text{Ba}$	0.13	
<i>Clathromorphum compactum</i>	$^{24}\text{Mg}$	0.16	[51]
	$^{43}\text{Ca}$	54.9	
<i>Clathromorphum nereostratum</i>	$^{24}\text{Mg}$	0.02	[52]
	$^{43}\text{Ca}$	5.47	
	$^{137}\text{Ba}$	0.01	
<i>Clathromorphum compactum</i>	$^{24}\text{Mg}$	0.16	[53]
	$^{43}\text{Ca}$	54.9	

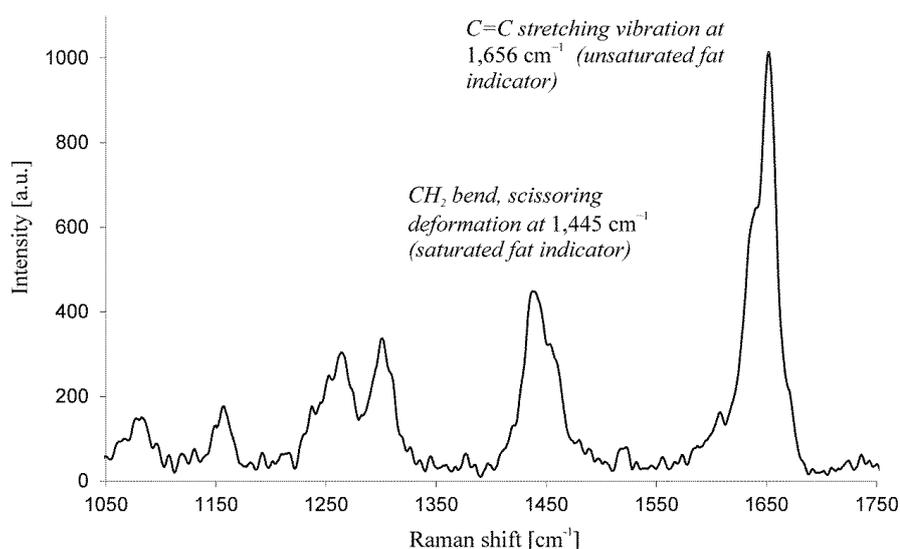
LA-ICP-MS is a prospective method to study solid algal samples with high resolution and relatively low limits of detection. This method seems to be more suitable to detect coralline algae than LIBS because LA-ICP-MS has in general higher spatial resolution and can detect isotopes ratio. This is important in the study of coralline algae, because they are optimal organisms to archive and detect climatic conditions of the environment. On the other hand, LA-ICP-MS cannot be used to study algal suspensions, and needs reference methods to measure the total concentration of elements in the samples, or standard reference materials to quantify elemental concentration in the samples.

#### 4. Raman Spectroscopy

Raman spectroscopy (alternatively Raman tweezers—a combination of Raman microspectroscopy with optical trapping) is a powerful and robust technique for analyzing biochemical information and revealing the molecular composition of samples under study [42,104–113].

Raman spectroscopy is based on the phenomenon of Raman scattering of monochromatic light (laser), which is the inelastic scattering of a photon. When there is monochromatic light incident on a target there are several possibilities for the incoming photons, if they have sufficient energy the molecules of the target can be raised to an excited electronic state and the photons absorbed, they can pass through the target without interacting or they can undergo elastic or inelastic scattering. Elastic scattering is the type of scattering that occurs most often when light is incident upon a target. In Raman spectroscopy however, inelastic scattering is exploited. In this case the molecule is excited to a virtual energy state, however this time when it relaxes it returns to a different vibrational energy state than the one that it started from. Therefore the energy of the photon emitted during the relaxation is different to the energy of the photon that caused the excitation in the first place. The scattered photon consequently has a different frequency than the excitation source and this is what produces the Raman spectrum, a plot of the frequency shift between incident and scattered light (see Figure 1).

**Figure 1.** Raman spectrum of a lipid body within the algae (*Trachydiscus minutus*). From the ratio of intensities  $I_{1656}/I_{1445}$  the iodine value (IV) can be estimated [42]. Here algal lipid content is close to IV ~ 230.



Review articles on biological applications [114–118] and especially on algae [119] have been presented. The primary goal of this literature research is the utilization of Raman spectroscopy for obtaining the information about the amount of lipid content within the algal cell. Efficient production of algal strains with higher lipid yield could lead to lowering the prices of biofuels [3]. Therefore, techniques allowing for rapid characterization/identification of algae species are required, and specifically to determine the degree of unsaturation of constituent fatty acids in algal lipid bodies. Note that the third generation

biofuels technology is based on algae that contain high oil content. Also, concerning the modern fish industry, most fish consumed inland come from fish farms. In this aquaculture industry, as the number of fish farms grow (mainly control farmed salmon products), it becomes important to guarantee that the high content of precious omega-3 fatty acids find their way into fish oils. This therefore dictates some fish dietary requirements for dedicated aqua-cultural environment. Consequently, aquafarmers feed fish, soy, and chicken oil to fish, all of which could be eliminated using algal oil. This highlights algae as a potential source from which desired omega-3 fatty acids can be extracted [120].

In 1983, Brahma *et al.* [121] reported on the measurement of the marine algae phytoplankton, employing Raman spectroscopy. Algae were measured directly in the suspension and the emphasis was given to the observation of the carotenoid pigments and chlorophyll peaks. The application of Raman spectroscopy to the analysis of photo-synthesizing organisms, such as algae, is challenging due to the underlying strong fluorescence of omnipresent pigments that might obscure the characteristic Raman spectral features. Therefore, the use of Raman spectroscopy has been limited to relatively few algal species. Accordingly, the number of published papers is relatively small, but tends to increase in recent years. Because of being in an early stage of development, these publications on Raman spectroscopy of algae are scattered over a wide range of journals, for instance [42,122–127], with the majority of work published in the last five years.

Algal strains, which could be promising candidates for biofuel production, have been so far investigated by five groups worldwide (see Table 3). The most widely studied species is *Botryococcus braunii*. The species with the highest iodine value (IV) was found to be *Trachydiscus minutus*. Thus far only two groups have been involved in systematic research on estimating the unsaturation degree/IV within algal samples (see Table 3).

**Table 3.** The list of articles focused on the analysis of algae employing Raman spectroscopy.

Algae Species	Reference	Estimate of the Degree of Unsaturation/Iodine-Value of Algal Oil
<i>Dunaliella tertiolecta</i>	[122,126]	No
<i>Chlorella sorokiniana</i>	[123]	No
<i>Neochloris oleoabundans</i>	[124]	No
<i>Botryococcus braunii</i>	[125]	Yes
<i>Botryococcus braunii</i>	[125]	Yes
<i>Neochloris oleoabundans</i>	[125]	Yes
<i>Chlamydomonas reinhardtii</i>	[125]	Yes
<i>Trachydiscus minutus</i>	[42]	Yes
<i>Botryococcus sudeticus</i>	[42]	Yes
<i>Chlamydomonas sp.</i>	[42]	Yes

The pioneering work of Heraud *et al.* [122,126], performed in 2007 in Beardall's laboratories at Monash University (Australia), was focused namely on *in vivo* Raman spectroscopy to predict the nutrient status of individual algal cells. They found that the Raman spectra of cells revealed a range of Raman bands mainly attributed to chlorophyll and carotene when 780 nm laser beam was used for excitation.

Preliminary feasibility studies on using Raman spectroscopy of algae were reported by Huang *et al.* [123]. They performed the study on two algal species, namely *Chlorella sorokiniana* and *Neochloris oleoabundans*, which could be seen as potential candidates for the biofuel production. Nitrogen-starved *C. sorokiniana* and *N. oleoabundans* samples were measured and Raman signals due to storage lipid (specifically triglycerides) were detected. The fluorescence background interrupted by sudden high-intensity fluorescence events was observed in the Raman signals from the algae. The fluorescence was acquired as a consequence of photo-bleaching of cell pigments due to prolonged intense laser light exposure; but the occurrence of the sudden high-intensity fluorescence bursts eluded full understanding.

Weiss *et al.* [124] reported Raman spectroscopy on *Botryococcus braunii* algae. In this study, authors were focused mainly on mapping the presence and location of methylated *Botryococenes* within the colony. Specific Raman spectroscopic characteristics for *Botryococenes* of *Botryococcus braunii* have been identified. *In vivo* lipid profiling of oil-producing algae has been described, using single-cell laser-trapping Raman spectroscopy [125].

Finally, Samek *et al.* [42] have recently demonstrated spatially resolved Raman spectroscopy to determine the effective IV in lipid storage bodies of individual algal cells. The Raman spectra were collected from different algal species immobilized in agarose gel, thus preventing them from moving out from the tightly focused region of the probe laser beam in order to maintain high spatial resolution within lipid bodies. The principal parameter characterizing the algal lipid is the degree of unsaturation of the constituent fatty acids and can be quantified by the IV. Crucially, the IV is conveniently estimated from information contained within the Raman spectra, with no need to add any chemicals to the cells. They used the characteristic peaks in the Raman spectra at  $1656\text{ cm}^{-1}$  (*cis* C=C stretching mode) and  $1445\text{ cm}^{-1}$  ( $\text{CH}_2$  scissoring mode) as the markers defining the ratio of unsaturated-to-saturated carbon-carbon bonds of the fatty acids in the algal lipids (Figure 1). For the quantitative IV determination a calibration curve was generated based on pure fatty acids of known IV, when the IV differed significantly for the various algal species. These estimates based on Raman spectroscopy were validated using the established technique of gas chromatography mass spectroscopy (GC-MS); indeed, excellent agreement was found.

As was mentioned above, the technique of Raman spectroscopy could be an excellent candidate to follow the food chain in the aquaculture industry, enabling one to monitor the IV within the food chain. Similarly, the same procedure of IV determination can be applied to monitor algae samples for biofuel production, where IV must be kept below a given limit. Moreover, it has been demonstrated [42] that various algal oils exhibit significantly different IV, which may have important implications for the food/pharmaceutical industry in obtaining 3-omega fatty acids. The main advantage of oils obtained from algae is that they are not contaminated by industrial toxins/antibiotics as some oils obtained from fish possibly could be due to the contaminated environment (industrial farmed fish or wild-caught fish).

## 5. Chemometrics for the Recognition of Algal Strains

The discussed spectroscopic techniques (LIBS, LA-ICP-MS, and Raman spectroscopy) are able to analyze extensive samples set, where each sample is represented by complex spectral information. Chemometric algorithms are used in many fields including spectroscopy, for data mining and pattern

recognition within the bulky data sets [61]. In general, chemometrics are used for both qualitative and quantitative analysis.

Classification of the algal strain has already been performed. Zbikowski *et al.* [128] applied the factor analysis for the discrimination of algal strain collected in coastal and lagoon waters, based on their appearances in flame atomic absorption spectroscopy spectra (FAAS). The algal strains were collected in the Gulf of Gdansk in a short time period (2000–2003). Algal samples were then dried and digested with HNO<sub>3</sub> acid. The contents of four macroelements (Ca, Mg, Na and K) and six heavy metals (Cd, Cu, Ni, Pb, Zn and Mn) were determined and used for further statistical analysis. The correlation between the concentrations of Cu, Pb and Zn in green algae and the sampling sites was observed. Heraud *et al.* [126,129] utilized Raman spectroscopy and Fourier transform infrared spectroscopy (FTIR), respectively. Chemometrics were then applied on the measured data set for accurately predicting the nutrient status of an independent individual algal strain. PCA was successfully applied as well on the near-infrared (NIR) and FTIR spectra [130]. Laurens and Wolfrum [130] used the NIR and FTIR spectra of biomass from four species to predict accurately the levels of exogenously added lipids. Salomonsen *et al.* [131] presented an extensive comparative study of alginate, a salt of alginic acid distributed widely in the cell walls of brown algae, using IR, Raman spectroscopy, NIR and NMR techniques. Chemometric algorithms, partial least squares discriminant analysis (PLS-DA) and PCA were then used to accurately predict the nutrient status of the cells from the Raman spectral data.

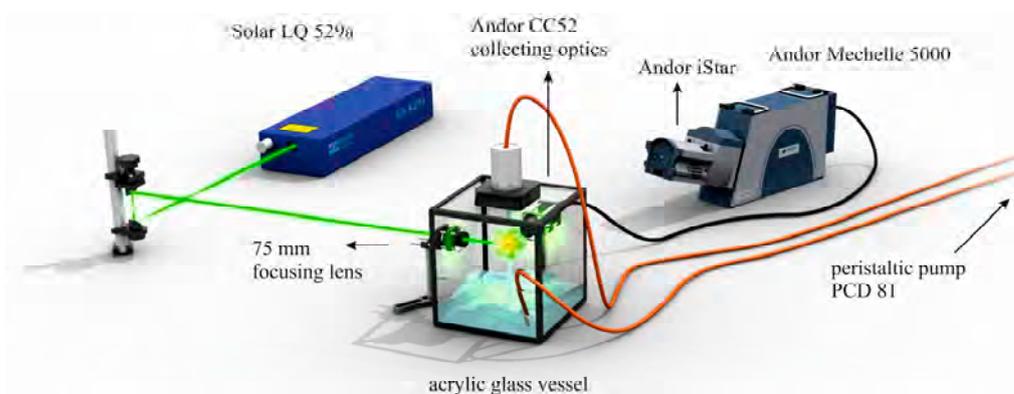
Concluding, chemometric algorithms may be of help in handling bulky data sets and revealing latent variables and relations among the biological samples. As stated above, LIBS is capable of providing information about the overall elemental composition of the sample, the so-called chemical fingerprint. The composition of matrix elements Ca, K, Mg, and Na could differ according to the particular measured algal strain [60]. The classification of algal strains based on their spectra is possible while employing the standard chemometric algorithm, such as principal component analysis. To the best of our knowledge, chemometric algorithms have not been used so far for the analysis and discrimination of algal strain based on their LIBS measurements. Nevertheless, successful utilization of LIBS for classification of biological samples has already been published [88,132–135].

### 5.1. Discrimination of Four Algal Strains by LIBS

Description of employed LIBS system, Figure 2, and related preparation of the four algal strains has already been published [60]. Four algal strains (*Chlarydomonas reinhardtii* (ChR), *Chlorococuum zurek* (ChZ), *Desmodesmus quadratic* (DQ), *Haematococcus pluralis* (HP)) were prepared under the same laboratory conditions. The samples were measured in the form of liquid suspensions with LIBS device where thin liquid jet was utilized. In this experiment, ns-laser pulse (Solar LQ 529a; operated at 532 nm, 10 ns, 50 mJ, ~65 GW/cm<sup>2</sup>) was focused with 75 mm planoconvex lens into a tight spot (100 μm). Radiation of a LIP was collected by using a large aperture collector-collimator (Andor CC52, F/2). Collected light was then spectrally resolved on the echelle grating (Andor Mechelle 5000; F/7,  $\lambda/\Delta\lambda = 6000$ ) and detected by an ICCD (Andor iStar 734). The temporal gating of the LIBS experiment was as follows: the gate delay of 3 μs and the gate width of 10 μs. Each measurement consists of 50 spectra in an accumulation while each measurement was 20 times repeated to obtain robust statistical dataset. A typical spectrum of an algal strain is depicted in Figure 3, where matrix elements (Ca, K, Mg,

and Na) are highlighted. During the data processing, the spectra were normalized to their integral intensities and averaged, and then four spectra per sample were obtained. Lines of matrix elements, listed in Table 4, were fitted with pseudo-Voigt profile and their intensities were calculated as the area under the peak with the background subtraction using custom MATLAB (version R2012a) software. Four spectra per each sample were organized as rows in the data matrix and its columns refer to individual variables. The range of each variable was normalized to unity and then mean-centered. The data matrix was analyzed with PCA to reveal possible latent variables among the data and to provide the discrimination of the samples (for detailed description of this procedure see further paragraph). This analysis was done employing MATLAB software customized with Self-Organizing MAP (SOM) toolbox [136] (Helsinki University of Technology, Finland) for multivariate analysis.

**Figure 2.** Schematic diagram of the so called liquid LIBS system [60].



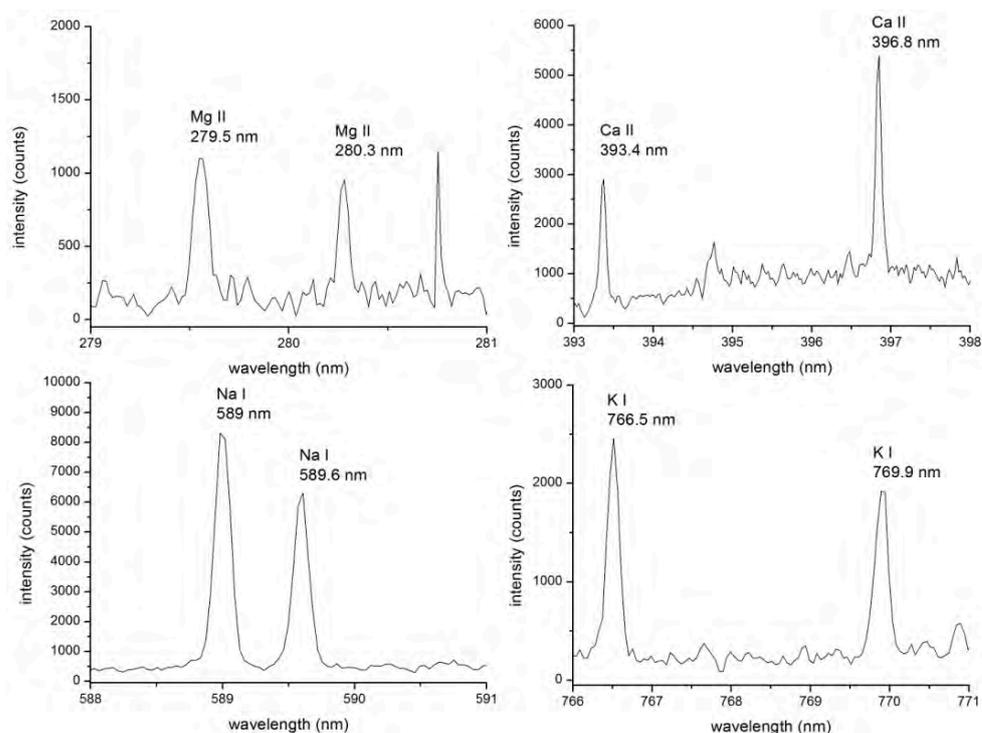
**Table 4.** Table of matrix elements utilized in multivariate analysis.

Element	Wavelength (nm)
Mg (II)	279.5
Mg (II)	280.3
Ca (II)	393.4
Ca (II)	396.8
Na (I)	589
Na (I)	589.6
K (I)	766.5
K (I)	769.9

Working with the whole data set can result in a PC space where various samples may be assigned to one group. In other words, the least squares property of PCA algorithm highlights the most significant variation among the data. For this reason, the less significant variation is overshadowed, *i.e.*, has lower impact on the classification in a newly constructed PC space. To overcome this problem one can utilize the approach suggested by Multari *et al.* [132] and used as well in related work by Ollila *et al.* [137]. There, any cluster is removed from the computation when it is successfully assigned to a distinct group in respect to the rest of the data set. Then the PCA algorithm is applied again on the reduced dataset. This leads to simplification of the variation in the dataset. In other words, the variation responsible for

the distinct separation of a group withdrawn from the computation is not present anymore. This results in the increase of the significance of a formerly less significant variation among the rest of the data. This process is repeated until all of the samples are successfully classified. By employing this algorithm for data classification we can proceed further in our investigation.

**Figure 3.** Typical spectrum of algae with the emphasis given to the matrix elements (Mg, Ca, Na, and K).

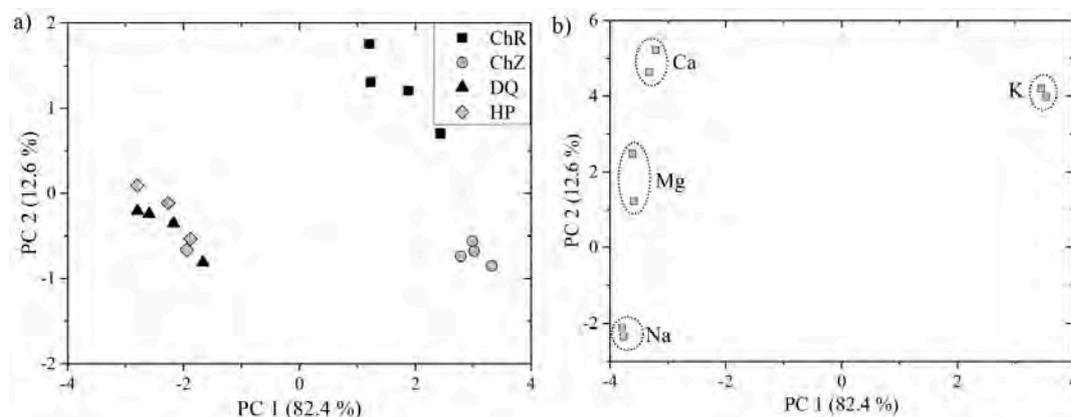


PCA was then applied on the data matrix constructed from the LIBS data. Two distinct groups are clearly visible when investigating resulted PCA scores in Figure 4a. The first two principal components describe 95% of overall variation among the data. This suggests that the discrimination of algal strains into three distinct groups is possible. However, data points representing algal strains *DQ* and *HP* are strongly overlapping. This may be a consequence of the similarity in the matrix elements and moderate repeatability of the LIBS measurement.

PCA analysis of LIBS spectra, *i.e.*, the clustering of the LIBS measurements, was then emulated by the PCA analysis of the ICP-OES measurement (not shown in this article). As in the case of LIBS measurement, the main emphasis was given to the signal intensities of the matrix element lines (namely Ca, K, Mg, and Na lines). The scores plot indicated that the composition of *DQ* and *HP* is more similar in the sense of matrix elements, *i.e.*, the data points are closer to each other, than the composition of *ChZ* and *ChR*, whose data points are distinctly separated in the newly created PC space. Those results coincided with the results of PCA applied to the LIBS data, where LIBS measurements of *ChR* and *ChZ* are distinctly separated compared to the overlapping LIBS data of *DQ* and *HP*. Therefore, the

repeatability of the LIBS measurement under given conditions still could be improved for reliable classification using PCA.

**Figure 4.** PCA of four algal strains based on their LIBS measurement: (a) scores; (b) loadings.



## 6. Conclusions and Future Prospects

Algae are considered to be promising alternative sources to corn and soybean for the next generation biofuel production. Third generation technology may be based on algal biomass, which is rich in polyunsaturated fatty acids. The algal biomass can be grown without competing for arable land (e.g., industrial waste waters can be used for the cultivation instead of the land suitable for growing food crops). Also, algae have the potential to decontaminate polluted water, because their cellular wall exhibits high affinity to metal cations. They are also widely used as a dietary supplement and in the drug industry. Moreover, calcified coralline algae can be utilized to detect climatic condition changes due to their longevity.

A comprehensive review on the analysis of algal biomass was given, preferably for the application in the fields of biofuels and bioremediation. The main aim of this review is focused on laser-based techniques for elemental and molecular analysis. It was shown that further development of methods for monitoring the elemental/chemical composition of the algal biomass is necessary. Each individual algal strain has different properties and reacts differently with its environment. Current laser-based spectroscopy techniques presented here such as LIBS, Raman spectroscopy and LA-ICP-OES/MS represent powerful tools for fast and complete analysis of biological samples and with certain limitations, can be adopted for effective analysis of algae biomass.

LIBS technique provides information primarily about the elemental composition. However limited information about the molecular structure can be also obtained. LIBS can serve as a robust, remote and rapid method for *in-situ*, on-line and real-time elemental analysis. Furthermore, portable LIBS equipment, employing a water jet, can be constructed for the fast elemental analysis or the algal identification in the field, mainly for the bioremediation application. Note that in the primary stage of the LIBS development LA-ICP-MS/OES can be also advantageously used to validate the LIBS outcomes. When the supervised spectral libraries are created, LIBS can stand alone as a robust, remote and rapid device for *in-situ* elemental analysis.

Raman spectroscopy (alternatively Raman tweezers – a combination of Raman microspectroscopy with optical trapping) is suitable for *in vivo* analysis of algae molecular composition in a non-destructive way. Recently, the primary goal of Raman spectroscopy was the determination of the lipid content within the algal cells. Spatially resolved Raman spectroscopy utilized for the determination of the iodine value, *i.e.*, lipid storage composition in the algal bodies, has a potential importance, especially in regard to third generation biofuels technologies. Raman spectroscopy as the only from the presented methods does not influence viability of living cells and can be combined with optical tweezers to sort individual cells according to their lipid content for subsequent breeding.

Chemometrics become more popular and irreplaceable in the spectral data mining. Chemometric algorithms may be an indispensable part of a robust analysis. Discrimination of different algal strains by LIBS or Raman spectroscopy using chemometric algorithms is also provided.

Pioneering works combining some of these approaches (e.g., LIBS and Raman spectroscopy) have already been published and the results, which are discussed above, show that those approaches can open new directions of bioanalytical remote measurement of algae.

### Acknowledgments

This work was supported by the project “CEITEC—Central European Institute of Technology” (CZ.1.05/1.1.00/02.0068) from European Regional Development Fund. J.N. acknowledges the project CZ.1.07/2.3.00/30.0005: “Support of Interdisciplinary Excellence Research Teams Establishment at BUT”. O.S., Z.P. and P.Z. acknowledge the support from Ministry of Education, Youth and Sports of the Czech Republic (LO1212) together with the European Commission (ALISI No. CZ.1.05/2.1.00/01.0017) and GACR P205/11/1687.

### Author Contributions

Pavel Pořízka, Ota Samek, Karel Novotný, Jan Novotný and Jozef Kaiser initiated, designed and structured the review. They were responsible for the six chapters of the manuscript and for writing and submission of the final manuscript.

David Procházka, Pavel Pořízka, Vojtěch Adam, Pavel Zemánek, and René Kízek participated in the study design, performed statistical analysis and analyzed extensive sample sets using in-house written routines. They significantly participated during the revision of the manuscript contributing to LIBS (chemometric), ICP and Raman chapters.

Petra Procházková, Lucia Sládková, and Zdeněk Pilát contributed to the study design, participated in data analysis and were responsible for cultivation and presentation of algae samples.

Michal Petrilak, Michal Brada contributed to manuscript writing, data presentation, literature survey and selection of relevant references for individual chapters.

### Conflicts of Interest

The authors declare no conflict of interest.

## References

1. I.E.A. IEA 2013 Key World Energy Statistics, 2013. Available online: [http://www.iea.org/publications/freepublications/publication/KeyWorld2013\\_FINAL\\_WEB.pdf](http://www.iea.org/publications/freepublications/publication/KeyWorld2013_FINAL_WEB.pdf) (accessed on 22 October 2013).
2. U.S. Department of Energy. Energy Efficiency and Renewable Energy, December 2007. Available online: <http://www.nrel.gov/docs/fy08osti/42168.pdf> (accessed on 22 October 2013).
3. Williams, P.J.I.B.; Laurens, L.M.L. Microalgae as biodiesel & biomass feedstocks: Review & analysis of the biochemistry, energetics & economics. *Energy Environ. Sci.* **2010**, *3*, 554–590.
4. Sheehan, J.; Dunahay, T.; Benemann, J.; Roessler, P. *A Look back at the U.S. Department of Energy's Aquatic Species Program-Biodiesel from Algae*; A National Laboratory of the U.S. Department of Energy: Golden, CO, USA, 1998.
5. Hannon, M.; Gimpel, J.; Tran, M.; Rasala, B.; Mayfield, S. Biofuels from algae: Challenges and potential. *Biofuels* **2010**, *1*, 763–784.
6. Demirbas, M.F. Biofuel from algae for sustainable development. *Appl. Energy* **2011**, *88*, 3437–3480.
7. U.S. Department of Energy, Energy Efficiency & Renewable Energy, 2013. Available online: <http://www1.eere.energy.gov/biomass/pdfs/algabiofuels.pdf> (accessed on 1 October 2013).
8. Chisti, Y. Constraints to commercialization of algal fuels. *J. Biotechnol.* **2013**, *167*, 201–214.
9. Elliott, L.G.; Feehan, C.; Laurens, L.M.L.; Pienkos, P.T.; Darzins, A.; Posewitz, M.C. Establishment of a bioenergy-focused microalgal culture collection. *Algal Res.* **2012**, *1*, 102–113.
10. Aguirre, M.; Bassi, A.; Saxena, P. Engineering challenges in biodiesel production from microalgae. *Crit. Rev. Biotechnol.* **2013**, *33*, 293–308.
11. Simionato, D.; Basso, S.; Giacometti, G.M.; Morosinotto, T. Optimization of light use efficiency for biofuel production in algae. *Biophys. Chem.* **2013**, *182*, 71–78.
12. Subhadra, B.G. Sustainability of algal biofuel production using integrated renewable energy park (IREP) and algal biorefinery approach. *Energy Policy* **2010**, *38*, 5892–5901.
13. Park, J.B.K.; Craggs, R.J.; Shilton, A.N. Wastewater treatment high rate algal ponds for biofuel production. *Biosource Technol.* **2011**, *102*, 35–42.
14. Park, J.B.K.; Craggs, R.J.; Shilton, A.N. Enhancing biomass energy yield from pilot-scale high rate algal ponds with recycling. *Water Res.* **2013**, *47*, 4422–4432.
15. Demirbas, A. Use of algae as biofuel sources. *Energy Convers. Manag.* **2010**, *51*, 2738–2749.
16. Christenson, L.; Sims, R. Production and harvesting of microalgae for wastewater treatment, biofuels, and biproducts. *Biotechnol. Adv.* **2011**, *29*, 686–702.
17. Rawat, R.R. Kumar Mutanda, T.; Bux, F. Dual role of microalgae: Phycoremediation of domestic wastewater and biomass production for sustainable biofuels production. *Appl. Energy* **2011**, *88*, 3411–3424.
18. Davis, T.A.; Volesky, B.; Mucci, A. A review of the biochemistry of heavy metal biosorption by brown algae. *Water Res.* **2003**, *37*, 4311–4330.
19. Durrieu, C.; Tran-Minh, C. Optical algal biosensor using alkaline phosphatase for determination of heavy metals. *Ecotoxicol. Environ. Saf.* **2002**, *51*, 206–209.

20. Campanella, L.; Cubadda, F.; Sammartino, M.P.; Saoncella, A. An algal biosensor for the monitoring of water toxicity in estuarine environments. *Water Res.* **2000**, *35*, 69–76.
21. Perez, A.A.; Farias, S.S.; Strobl, A.M.; Perez, L.B.; Lopez, C.M.; Pineiro, A.; Roses, O.; Fajardo, M.A. Levels of essential and toxic elements in *Porphyra columbina* and *Ulva* sp. from San Jorge Gulf, Patagonia Argentina. *Sci. Total Environ.* **2007**, *376*, 51–59.
22. Cardozo, K.H.M.; Guaratini, T.; Barros, M.; Falcao, V.R.; Tonon, A.P.; Lopes, N.P.; Campos, S.; Torres, M.A.; Souza, A.O.; Colepicolo, P.; et al. Metabolites from algae with economical impact. *Toxicol. Pharmacol.* **2007**, *146*, 60–78.
23. Singh, S.; Kate, B.N.; Banerjee, U.C. Bioactive compounds from cyanobacteria and microalgae: An overview. *Crit. Rev. Biotechnol.* **2005**, *25*, 73–95.
24. Blunt, J.W.; Copp, B.R.; Munro, M.H.G.; Northcote, P.T.; Prinsep, M.R. Marine natural products. *Nat. Prod. Rep.* **2005**, *22*, 15–61.
25. Skjånes, K.; Rebours, C.; Lindblad, P. Potential for green microalgae to produce hydrogen, pharmaceuticals and other high value products in a combined process. *Crit. Rev. Biotechnol.* **2013**, *33*, 172–215.
26. Kim, S.K.; Thomas, N.V.; Li, X. Anticancer compounds from marine macroalgae and their application as medical foods. *Adv. Food Nutr. Res.* **2011**, *64*, 213–224.
27. Kim, S.K.; Karadeniz, F. Anti-HIV activity of extracts and compounds from marine algae. *Adv. Food Nutr. Res.* **2011**, *64*, 255–265.
28. Munkel, R.; Schmid-Staiger, U.; Werner, A.; Hirth, T. Optimization of outdoor cultivation in flat panel airlift reactors for lipid production by *Chlorella vulgaris*. *Biotechnol. Bioeng.* **2013**, *110*, 2882–2893.
29. IBA, BIQ, Smart Material Houses, 2013. Available online: <http://www.iba-hamburg.de/en/themes-projects/the-building-exhibition-within-the-building-exhibition/smart-material-houses/biq/projekt/biq.html> (accessed on 22 October 2013).
30. Jaboyedoff, M.; Oppikofer, T.; Abellán, A.; Derron, M.H.; Loye, A.; Metzger, R.; Pedrazzini, A. Use of LIDAR in landslide investigations: A review. *Nat. Hazards* **2012**, *61*, 5–28.
31. Cracknell, A.P.; Hayes, L. *Introduction to Remote Sensing*, 2nd ed.; Taylor and Francis: London, UK, 2007.
32. Bazzani, M.; Cecchi, G. Algae and mucilage monitoring by fluorescence LIDAR experiments in field. *Adv. Remote Sens.* **1995**, *3*, 90–101.
33. Mumola, P.B.; Kim, H.H. Remote sensing of marine plankton by dye laser induced fluorescence. *Eng. Ocean Environ.* **1972**, *72*, 204–207.
34. Martin, F. Monitoring plant metabolism by <sup>13</sup>C, <sup>15</sup>N and <sup>14</sup>N nuclear magnetic resonance spectroscopy. A review of the applications to algae, fungi, and higher plants. *Plant Physiol.* **1985**, *23*, 463–490.
35. Schneider, B. *In-vivo* nuclear magnetic resonance spectroscopy of low-molecular-weight compounds in plant cells. *Planta* **1997**, *203*, 1–8.
36. Ratcliffe, R.G. *In-vivo* NMR studies of higher plants and algae. *Adv. Bot. Res.* **1994**, *20*, 43–123.
37. Pollesello, P.; Toffanin, R.; Murano, E.; Paoletti, S.; Rizzo, R.; Kvam, B.J. Lipid extracts from different algal species: <sup>1</sup>H- and <sup>13</sup>C-NMR spectroscopic studies as a new tool to screen differences in the composition of fatty acids, sterols and carotenoids. *J. Appl. Phycol.* **1992**, *4*, 315–322.

38. Danielewicz, M.A.; Anderson, L.A.; Franz, A.K. Triacylglycerol profiling of marine microalgae by mass spectrometry. *J. Lipid Res.* **2013**, *52*, 1–26.
39. Vieler, A.; Wilhelm, C.; Goss, R.; Süß, R.; Schiller, J. The lipid composition of the unicellular green alga *Chlamydomonas reinhardtii* and the diatom *Cyclotella meneghiniana*. *Chem. Phys. Lipids* **2007**, *150*, 143–155.
40. Wirth, H.; von Bergen, M.; Murugaiyan, J.; Rösler, U.; Stokowy, T.; Binder, H. MALDI-typing of infectious algae of the genus *Prototheca* using SOM portraits. *J. Microbiol. Methods* **2012**, *88*, 83–97.
41. Anastuyuk, S.D.; Shevchenko, N.M.; Dmitrenok, P.S.; Zvyagintseva, T.N. Structural similarities of fucoidans from brown algae *Silvetia babingtonii* and *Fucus evanescens*, determined by tandem MALDI-TOF mass spectrometry. *Carbohydr. Res.* **2012**, *358*, 78–81.
42. Samek, O.; Jonáš, A.; Pilát, Z.; Zemánek, P.; Nedbal, L.; Tríska, J.; Kotas, P.; Trtílek, M. Raman microspectroscopy of individual algal cells: Sensing unsaturation of storage lipids *in vivo*. *Sensors* **2010**, *10*, 8635–8651.
43. Bailey, G.F.; Horvat, R.J. Raman spectroscopic analysis of the cis/trans isomer composition of edible vegetable oils. *J. Am. Oil Chem. Soc.* **1972**, *4*, 494–498.
44. Sadeghi-Jorabchi, H.; Hendra, P.J.; Wilson, R.H.; Belton, P.S. Determination of the total unsaturation in oils and margarines by Fourier Transform Raman Spectroscopy. *J. Am. Oil Chem. Soc.* **1990**, *67*, 483–486.
45. Ozaki, Y.; Cho, R.; Ikegaya, K.; Muraishi, S.; Kawauchi, K. Potential of near-infrared Fourier Transform Raman Spectroscopy in food analysis. *Appl. Spectrosc.* **1992**, *46*, 1503–1507.
46. Schober, S.; Mittelbach, M. Iodine value and biodiesel: Is limitation still appropriate? *Lipid Technol.* **2007**, *17*, 281–284.
47. Gouveia, L.; Marques, A.E.; da Silva, T.L.; Reis, A. *Neochloris oleabundans*: A suitable renewable lipid source for biofuel production. *J. Ind. Microbiol. Biotechnol.* **2009**, *36*, 821–826.
48. Gouveia, L.; Oliveira, A.C. Microalgae as a raw material for biofuels production. *J. Ind. Microbiol. Biotechnol.* **2009**, *36*, 269–274.
49. Winefordner, J.D.; Gornushkin, I.B.; Correll, T.; Gibb, E.; Smith, B.W.; Omenetto, N. Comparing several atomic spectrometric methods to the super stars: Special emphasis on laser induced breakdown spectrometry, LIBS, a future superstar. *J. Anal. Atomic Spectrom.* **2004**, *19*, 1061–1083.
50. Hetzinger, S.; Halfar, J.; Zack, T.; Gamboa, G.; Jacob, D.E.; Kunz, B.E.; Kronz, A.; Adey, W.; Lebednik, P.A.; Steneck, R.S. High-Resolution analysis of trace elements in crustose coralline algae from the North Atlantic and North Pacific by laser ablation ICP-MS. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **2011**, *302*, 81–94.
51. Halfar, J.; Hetzinger, S.; Adey, W.; Zack, T.; Gamboa, G.; Kunz, B.; Williams, B.; Jacob, D.E. Coralline algal growth-increment widths archive North Atlantic climate variability. *Palaeogeography* **2011**, *302*, 71–80.
52. Chan, P.; Halfar, J.; Williams, B.; Hetzinger, S.; Steneck, R.; Zack, T.; Jacob, D.E. Freshening of the Alaska Coastal Current recorded by coralline algal Ba/Ca ratios. *J. Geophys. Res. Biogeosc.* **2011**, *116*, 2005–2012.

53. Gamboa, G.; Halfar, J.; Hetzinger, S.; Adey, W.; Zack, T.; Kunz, B.; Jacob, D.E. Mg/Ca ratios in coralline algae record NW Atlantic temperature variations and NAO relationships. *J. Geophys. Res. Oceans* **2010**, *115*, doi:10.1029/2010JC006262.
54. Hetzinger, S.; Halfar, J.; Zack, T.; Mecking, J.V.; Kunz, B.E.; Jacob, D.E.; Adey, W.H. Coralline algal Barium as indicator for 20th century northwestern North Atlantic surface ocean freshwater variability. *Sci. Rep.* **2013**, *3*, 1–8.
55. Hetzinger, S.; Halfar, J.; Kronz, A.; Steneck, S.; Walter, A.; Lebednik, P.A.; Schöne, B.R. High-resolution Mg/Ca ratios in a coralline red algae as a proxy for Bering Sea temperature variations from 1902 to 1967. *Soc. Sediment. Geol.* **2009**, *24*, 406–412.
56. Hahn, D.W.; Omenetto, N. Laser-Induced Breakdown Spectroscopy (LIBS), Part II: Review of instrumental and methodological approaches to material analysis and applications to different fields. *Appl. Spectrosc.* **2012**, *66*, 347–419.
57. Wiens, R.C.; Sharma, S.K.; Thomson, J.; Misra, A.; Lucey, P.G. Joint analysis by LIBS and Raman spectroscopy at stand-off distances. *Spectrochim. Acta Part A* **2005**, *61*, 2324–2334.
58. Giakoumaki, A.; Osticioli, I.; Anglos, D. Spectroscopic analysis using a hybrid LIBS-Raman system. *Appl. Phys. A* **2006**, *83*, 537–541.
59. Hoehse, M.; Mory, D.; Florek, S.; Weritz, F.; Gornushkin, I.; Panne, U. A combined laser-induced breakdown and Raman spectroscopy Echelle system for elemental and molecular microanalysis. *Spectrochim. Acta Part B* **2009**, *64*, 1219–1227.
60. Pořízka, P.; Prochazka, D.; Pilát, Z.; Krajcarová, L.; Kaiser, J.; Malina, R.; Novotný, J.; Zemánek, P.; Ježek, J.; Šerý, M.; *et al.* Application of laser-induced breakdown spectroscopy to the analysis of algal biomass for industrial biotechnology. *Spectrochim. Acta Part B* **2012**, *74–75*, 169–176.
61. Martens, H.; Naes, T. *Multivariate Calibration*; John Wiley & Sons Ltd.: Chichester, UK, 1989.
62. Barker, M.; Rayens, W. Partial least squares for discrimination. *J. Chemom.* **2003**, *17*, 166–173.
63. Cremers, D.A.; Radziemski, L.J. *Handbook of Laser-Induced Breakdown Spectroscopy*; John Wiley & Sons, Ltd.: New York, NY, USA, 2006.
64. Miziolek, A.W.; Palleschi, V.; Schechter, I. *Laser-Induced Breakdown Spectroscopy (LIBS): Fundamentals and Applications*; Cambridge University Press: Cambridge, UK, 2006.
65. Noll, R. *Laser-Induced Breakdown Spectroscopy: Fundamentals and Applications*; Springer-Verlag: Berlin/Heidelberg, Germany, 2012.
66. Hahn, D.W.; Omenetto, N. Laser-Induced breakdown spectroscopy (LIBS), Part I: Review of basic diagnostics and plasma-particle interactions: Still-Challenging issues within the analytical plasma community. *Appl. Spectrosc.* **2010**, *64*, 335–366.
67. Aragón, C.; Aguilera, J.A. Characterization of laser induced plasma by optical emission spectroscopy: A review of experiments and methods. *Spectrochim. Acta Part B* **2008**, *63*, 893–916.
68. López-Moreno, C.; Palanco, S.; Laserna, J.J. Remote laser-induced plasma spectrometry for elemental analysis of samples of environmental interest. *J. Anal. Atomic Spectrom.* **2004**, *19*, 1479–1484.
69. Fortes, F.J.; Laserna, J.J. The development of fieldable laser-induced breakdown spectrometer: No limits on the horizon. *Spectrochim. Acta Part B* **2010**, *65*, 975–990.

70. Gaudiuso, R.; Dell'Aglio, M.; de Pascale, O.; Senesi, G.S.; de Giacomo, A. Laser induced breakdown spectroscopy for elemental analysis in environmental, cultural heritage and space applications: A review of methods and results. *Sensors* **2010**, *10*, 7434–7468.
71. Harmon, R.S.; Russo, R.E.; Hark, R.R. Application of laser-induced breakdown spectroscopy for geochemical and environmental analysis: A comprehensive review. *Spectrochim. Acta Part B* **2013**, *87*, 11–26.
72. Michel, A.P.M. Review: Application of single-shot laser-induced breakdown spectroscopy. *Spectrochim. Acta Part B* **2010**, *65*, 185–191.
73. Santos, D., Jr.; Nunes, L.C.; de Carvalho, G.G.A.; da Silva Gomes, M.; de Souza, P.F.; de Oliveira Leme, F.; Cofani dos Santos, L.G.; Krug, F.J. Laser-Induced breakdown spectroscopy for analysis of plant materials: A review. *Spectrochim. Acta Part B* **2012**, *71*, 3–13.
74. Kaiser, J.; Novotný, K.; Martin, M.Z.; Hrdlička, A.; Malina, R.; Hartl, M.; Adam, V.; Kizek, R. Trace element analysis by laser-induced breakdown spectroscopy—Biological applications. *Surf. Sci. Rep.* **2012**, *67*, 233–243.
75. Rehse, S.J.; Salimnia, H.; Miziolek, A.W. Laser-Induced breakdown spectroscopy (LIBS): An overview of recent progress and future potential for biomedical applications. *J. Med. Eng. Technol.* **2012**, *36*, 77–89.
76. Gurevich, E.L.; Hergenröder, R. Femtosecond laser-induced breakdown spectroscopy: Physics, applications, and perspectives. *Appl. Spectrosc.* **2007**, *61*, 233–242.
77. Garcimuno, M.; Pace, D.M.D.; Bertucelli, G. Laser-Induced breakdown spectroscopy for quantitative analysis of copper in algae. *Opt. Laser Technol.* **2013**, *47*, 26–30.
78. Niu, L.; Cho, H.H.; Song, K.; Cha, H.; Kim, Y.; Lee, Y.I. Direct determination of strontium in marine algae samples by laser-induced breakdown spectrometry. *Appl. Spectrosc.* **2002**, *56*, 1511–1515.
79. Aguirre, M.A.; Legnaioli, S.; Almodóvar, F.; Hidalgo, M.; Palleschi, V.; Canals, A. Elemental analysis by surface-enhanced Laser-Induced Breakdown Spectroscopy combined with liquid-liquid microexcitation. *Spectrochim. Acta Part B* **2013**, *79*, 88–93.
80. Nakamura, S.; Ito, Y.; Sone, K. Determination of an iron suspension in water by laser-induced breakdown spectroscopy with two sequential laser pulses. *Anal. Chem.* **1996**, *68*, 2981–2986.
81. Ito, Y.; Ueki, O.; Nakamura, S. Determination of colloidal iron in water by laser-induced breakdown spectroscopy. *Anal. Chim. Acta* **1995**, *299*, 401–405.
82. Samek, O.; Beddows, D.C.; Kaiser, J.; Kukhlevsky, S.V.; Liska, M.; Telle, H.H.; Whitehouse, A.J. Application of laser-induced breakdown spectroscopy to *in situ* analysis of liquid samples. *Opt. Eng.* **2000**, *39*, 2248–2262.
83. Rai, V.N.; Yueh, F.Y.; Singh, J.P. Study of laser-induced breakdown emission from liquid under double-pulse excitation. *Appl. Opt.* **2003**, *42*, 2094–2101.
84. Feng, Y.; Yang, J.; Fan, J.; Yao, G.; Ji, X.; Zhang, X.; Zheng, X.; Cui, Z. Investigation of laser-induced breakdown spectroscopy of a liquid jet. *Appl. Opt.* **2010**, *49*, 70–74.
85. Cheri, M.S.; Tavassoli, S.H. Quantitative analysis of toxic metals lead and cadmium in water jet by laser-induced breakdown spectroscopy. *Appl. Opt.* **2011**, *50*, 1227–1233.

86. Debras-Guédon, J.; Liodec, N. De l'utilisation du faisceau issu d'un amplificateur à ondes lumineuses par émission induite par rayonnement (laser à rubis), comme source énergétique pour l'excitation des spectres d'émission des éléments. *Comptes Rendus Hebdomadaires des séances de l'Académie des Sci.* **1963**, *257*, 3336–3339.
87. Samuels, A.C.; DeLucia, F.C., Jr.; McNesby, K.L.; Miziolek, A.W. Laser-Induced breakdown spectroscopy of bacterial spores, molds, pollens, and protein: Initial studies of discrimination potential. *Appl. Opt.* **2003**, *42*, 6205–6209.
88. Baudalet, M.; Guyon, L.; Yu, J.; Wolf, J.P.; Amodeo, T.; Fréjafon, E.; Laloi, P. Femtosecond time-resolved laser-induced breakdown spectroscopy for detection and identification of bacteria: A comparison to the nanosecond regime. *J. Appl. Phys.* **2006**, *99*, 84701–84709.
89. Doucet, F.R.; Faustino, P.J.; Sabsabi, M.; Lyon, R.C. Quantitative molecular analysis with molecular bands emission using laser-induced breakdown spectroscopy and chemometrics. *J. Anal. At. Spectrom.* **2008**, *23*, 694–701.
90. Kongbonga, Y.G.M.; Ghalila, H.; Onana, M.B.; Lakhdar, Z.B. Classification of vegetable oils based on their concentration of saturated fatty acids using laser induced breakdown spectroscopy (LIBS). *Food Chem.* **2013**, *147*, 327–331.
91. Locke, R.J.; Morris, J.B.; Forch, B.E.; Miziolek, A.W. Ultraviolet laser microplasma-gas chromatography detector: detection of species-specific fragment emission. *Appl. Opt.* **1990**, *29*, 4987–4991.
92. St-Onge, L.; Sing, R.; Béchard, S.; Sabsabi, M. Carbon emission following 1.064  $\mu\text{m}$  laser ablation of graphite and organic samples in ambient air. *Appl. Phys. A* **1999**, *69*, 913–916.
93. Morel, S.; Leone, N.; Adam, P.; Amouroux, J. Detection of bacteria by time-resolved laser-induced breakdown spectroscopy. *Appl. Opt.* **2003**, *42*, 6184–6191.
94. Guyon, L.; Baudalet, M.; Yu, J.; Wolf, J.P.; Amodeo, T.; Fréjafon, E.; Laloi, P. Laser-Induced breakdown spectroscopy analysis of bacteria: What femtosecond lasers make possible. *Chem. Phys.* **2007**, *2007*, 193–195.
95. Berman, L.; Wolf, P.J. Laser-Induced breakdown spectroscopy of liquids: Aqueous solutions of nickel and chlorinated hydrocarbons. *Appl. Spectrosc.* **1998**, *52*, 438–443.
96. Rodenas de la Rocha, S.; Sanchez-Muniz, F.J.; Gomez-Juaristi, M.; Larrea Marin, M.T. Trace elements determination in edible seaweeds by an optimized and validated ICP-MS method. *J. Food Compos. Anal.* **2009**, *22*, 330–336.
97. Larrea Marin, M.T.; Pomares-Alfonso, M.S.; Gomez-Juaristi, M.; Sanchez-Muniz, F.J.; Rodenas de la Rocha, S. Validation of an ICP-OES method for macro and trace element determination in Laminaria and Porphyra seaweeds from four different countries. *J. Food Compos. Anal.* **2010**, *23*, 814–820.
98. Moreda-Pineiro, J.; Alonso-Rodriguez, E.; Lopez-Mahia, P.; Muniategui-Lorenzo, S.; Prada-Rodriguez, D.; Moreda-Pineiro, A.; Bermejo-Barrera, P. Development of a new sample pre-treatment procedure based on pressurized liquid extraction for the determination of metals in edible seaweed. *Anal. Chim. Acta* **2007**, *598*, 95–102.
99. Subba Rao, P.V.; Mantri, V.A.; Ganesan, K. Mineral composition of edible seaweed *Porphyra vietnamensis*. *Food Chem.* **2007**, *102*, 215–218.

100. Pena-Harfal, C.; Moreda-Pineiro, A.; Bermejo-Barrera, A.; Bermejo-Barrera, P.; Pinochet-Cancino, H.; Gregori-Henriquez, I. Speeding up enzymatic hydrolysis procedures for the multi-element determination in edible seaweed. *Anal. Chim. Acta* **2005**, *548*, 183–191.
101. Michalak, I.; Chojnacka, K. Multielemental analysis of macroalgae from Baltic Sea by ICP-OES to monitoring environmental pollution and assess their potential uses. *Int. J. Environ. Anal. Chem.* **2009**, *89*, 583–596.
102. Van Netten, C.; Hoption Cann, S.A.; Morley, D.R.; van Netten, J.P. Elemental and radioactive analysis of commercially available seaweed. *Sci. Total Environ.* **2000**, *255*, 169–175.
103. West, G.B.; Brown, J.H.; Enquist, B.J. A general model for ontogenetic growth. *Nature* **2001**, *413*, 628–631.
104. Čižmár, T.; Brzobohatý, O.; Dholakia, K.; Zemánek, P. The holographic optical micro-manipulation system based on counter-propagating beams. *Laser Phys. Lett.* **2011**, *8*, 50–56.
105. Vlasova, I.M.; Saletsky, A.M. Investigation of denaturation of human serum albumin under action of cethyltrimethylammonium bromide by raman spectroscopy. *Laser Phys.* **2011**, *21*, 239–244.
106. Gen, D.E.; Chernyshov, K.B.; Prokhorov, K.A.; Nikolaeva, G.Y.; Sagitova, E.A.; Pashinin, P.P.; Kovalchuk, A.A.; Klyamkina, A.N.; Nedorezova, P.M.; Optov, V.A.; *et al.* Polarized Raman study of random copolymers of propylene with olefins. *Laser Phys.* **2010**, *20*, 1354–1367.
107. Lademan, J.; Caspers, P.J.; van der Pol, A.; Richter, H.; Patzelt, A.; Zastrow, L.; Darvin, M.; Sterry, W.; Fluhr, J.W. *In vivo* Raman spectroscopy detects increased epidermal antioxidative potential with topically applied carotenoids. *Laser Phys. Lett.* **2009**, *6*, 76–79.
108. Sturm, M.; Schlösser, M.; Lewis, R.J.; Bornschein, B.; Drexlin, G.; Telle, H.H. Monitoring of all hydrogen isotopologues at tritium laboratory Karlsruhe using Raman spectroscopy. *Laser Phys.* **2010**, *20*, 493–507.
109. Samek, O.; Al-Marashi, J.F.M.; Telle, H.H. The potential of Raman spectroscopy for the identification of biofilm formation by *Staphylococcus epidermidis*. *Laser Phys. Lett.* **2010**, *7*, 378–383.
110. Samek, O.; Zemánek, P.; Jonáš, A.; Telle, H.H. Characterization of oil-producing microalgae using Raman spectroscopy. *Laser Phys. Lett.* **2011**, *8*, 701–709.
111. Jonáš, A.; Zemánek, P. Light at work: The use of optical forces for particle manipulation, sorting, and analysis. *Electrophoresis* **2008**, *29*, 4813–4851.
112. Petrov, D.V. Raman spectroscopy of optically trapped particles. *J. Opt.* **2007**, *9*, 139–156.
113. Pilát, Z.; Ježek, J.; Šerý, M.; Trtílek, M.; Nedbal, L.; Zemánek, P. Optical trapping of microalgae at 735–1064 nm: Photodamage assessment. *J. Photochem. Photobiol. B Biol.* **2013**, *121*, 27–31.
114. Kundu, P.P.; Narayana, C. Raman based imaging in biological application—A perspective. *J. Med. Allied Sci.* **2012**, *2*, 41–48.
115. Matousek, P.; Stone, N. Emerging concepts in deep Raman spectroscopy of biological tissue. *Analyst* **2009**, *134*, 1058–1066.
116. Downes, A.; Elfick, A. Raman spectroscopy and related techniques in biomedicine. *Sensors* **2010**, *10*, 1871–1889.
117. Movasaghi, Z.; Rehman, S.; Rehman, I.U. Raman spectroscopy of biological tissues. *Appl. Spectrosc.* **2007**, *42*, 493–541.
118. Notingher, I. Raman spectroscopy cell-based biosensors. *Sensors* **2007**, *7*, 1343–1358.

119. Parab, N.D.T.; Tomar, V. Raman spectroscopy of algae: A review. *J. Nanomed. Nanotechnol.* **2012**, *3*, 131–138.
120. Samek, O.; Zemánek, P.; Bernatová, S.; Pilát, Z.; Telle, H.H. Following lipids in the food chain: Determination of the iodine value using Raman micro-spectroscopy. *Spectrosc. Eur.* **2012**, *24*, 22–25.
121. Brahma, S.K.; Hargraves, P.E.; Howard, W.F., Jr.; Nelson, W.H. A Resonance raman method for the rapid detection and identification of algae in water. *Appl. Spectrosc.* **1983**, *37*, 55–58.
122. Heraud, P.; Wood, B.R.; Beardall, J.; McNaughton, D. Effects of pre-processing of Raman spectra on *in vivo* classification of nutrient status of microalgal cells. *J. Chemom.* **2006**, *20*, 193–197.
123. Huang, Y.Y.; Beal, C.M.; Cai, W.W.; Ruoff, R.S.; Terentjev, E.M. Micro-Raman spectroscopy of algae: Composition analysis and fluorescence background behavior. *Biotechnol. Bioeng.* **2010**, *105*, 889–898.
124. Weiss, T.L.; Chun, H.J.; Okada, S.; Vitha, S.; Holzenburg, A.; Laane, J.; Devarenne, T.P. Raman spectroscopy analysis of botryococcene hydrocarbons from the green microalga *Botryococcus braunii*. *J. Biol. Chem.* **2010**, *285*, 32458–32466.
125. Wu, H.; Volponi, J.V.; Oliver, A.E.; Parikh, A.N.; Simmons, B.A.; Singh, S. *In vivo* lipidomics using single-cell Raman spectroscopy. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 3809–3814.
126. Heraud, P.; Beardall, J.; McNaughton, D.; Wood, B.R. *In vivo* prediction of the nutrient status of individual microalgal cells using Raman microspectroscopy. *FEMS Microbiol. Lett.* **2007**, *275*, 24–30.
127. Pilát, Z.; Bernatová, S.; Ježek, J.; Šerý, M.; Samek, O.; Zemánek, P.; Nedbal, L.; Trtílek, M. Raman microspectroscopy of algal lipid bodies: Beta-Carotene quantification. *J. Appl. Phycol.* **2012**, *24*, 541–546.
128. Zbikowski, R.; Szefer, P.; Latala, A. Comparison of green algae *Cladophora* sp. and *Enteromorpha* sp. as potential biomonitors of chemical elements in the southern Baltic. *Sci. Total Environ.* **2007**, *387*, 320–337.
129. Heraud, P.; Stojkovic, S.; Beardall, J.; McNaughton, D.; Wood, B.R. Intercolonial variability in macromolecular composition in p-starved and p-replete scenedesmus populations revealed by infrared microspectroscopy. *J. Phycol.* **2008**, *44*, 1335–1339.
130. Laurens, L.M.L.; Wolfrum, E.J. Feasibility of spectroscopic characterization of algal lipids: Chemometric correlation of NIR and FTIR spectra with exogenous lipids in algal biomass. *Bioenergy Res.* **2011**, *4*, 22–35.
131. Salomonsen, T.; Jensen, H.M.; Stenbæk, D.; Engelsen, S.B. Chemometric prediction of alginate monomer composition: A comparative spectroscopic study using IR, Raman, NIR and NMR. *Carbohydr. Polym.* **2008**, *72*, 730–739.
132. Multari, R.A.; Cremers, D.A.; Dupre, J.M.; Gustafson, J.E. The use of laser-induced breakdown spectroscopy for distinguishing between bacterial pathogen species and strains. *Appl. Spectrosc.* **2010**, *64*, 750–759.
133. Kim, T.; Specht, Z.G.; Vary, P.S.; Lin, C.T. Spectral fingerprints of bacterial strains by laser-induced breakdown spectroscopy. *J. Phys. Chem. B* **2004**, *108*, 5477–5482.
134. Diedrich, J.; Rehse, S.J.; Palchadhuri, S. Pathogenic *Escherichia coli* strain discrimination using laser-induced breakdown spectroscopy. *J. Appl. Phys.* **2007**, *102*, 014702–014708.

135. Rehse, S.J.; Jeyasingham, N.; Diedrich, J.; Palchaudhuri, S. A membrane basis for bacterial identification and discrimination using laser-induced breakdown spectroscopy. *J. Appl. Phys.* **2009**, *105*, 1–13.
136. SOM Toolbox team, Helsinki University of Technology, 23 March 2005. Available online: <http://www.cis.hut.fi/somtoolbox/> (accessed on 1 November 2013).
137. Ollila, A.M.; Lasue, J.; Newson, H.E.; Multari, R.A.; Wiens, R.C.; Clegg, S.M. Comparison of two partial least squares-discriminant analysis algorithms for identifying geological samples with the ChemCam laser-induced breakdown spectroscopy instrument. *Appl. Opt.* **2012**, *51*, 130–142.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).

#### Article 4

G. Vítková, L. Prokeš, K. Novotný, P. Pořízka, J. Novotný, D. Všianský, L. Čelko, and 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. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 101:191–199, 2014.

DOI: 10.1016/j.sab.2014.08.036, 32 citations

In this co-authored work, the capability of stand-off LIBS system was tested. Several brick samples were measured at the distance of 6.2 m, having a spot-size of 1 mm. The emission was collected by a Newtonian telescope with a 10" mirror diameter. Obtained spectra were processed with simple linear chemometric algorithms, namely principal component analysis (PCA) for visualization and linear discriminant analysis (LDA) for classification. The stand-off performance was compared to the one of a table-top LIBS system.

The visualization in the PC space revealed clear separation of individual brick samples showing also a certain trend. This trend was attributed to the variation in the sample matrix. Individual bricks were fired under different conditions (temperatures) which led to variation in the sample hardness. Classification of the table-top LIBS system (100 % accuracy) slightly outperformed the stand-off one (86 % accuracy).



## 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<sup>☆</sup>



Gabriela Vítková<sup>a</sup>, Lubomír Prokeš<sup>a,b,c</sup>, Karel Novotný<sup>a,d,\*</sup>, Pavel Pořízka<sup>e</sup>, Jan Novotný<sup>f</sup>, Dalibor Všíanský<sup>g</sup>, Ladislav Čelko<sup>f</sup>, Jozef Kaiser<sup>e,f</sup>

<sup>a</sup> Department of Chemistry, Faculty of Science, Masaryk University, Kamenice 5/A14, 625 00 Brno, Czech Republic

<sup>b</sup> Department of Physical Electronics, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic

<sup>c</sup> CEPLANT, R&D Center for Low-Cost Plasma and Nanotechnology Surface Modifications, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic

<sup>d</sup> CEITEC – Central European Institute of Technology, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic

<sup>e</sup> Institute of Physical Engineering, Faculty of Mechanical Engineering, Brno University of Technology, Technická 2896/2, 616 69 Brno, Czech Republic

<sup>f</sup> CEITEC – Central European Institute of Technology, Brno University of Technology, Technická 3058/10, 616 00 Brno, Czech Republic

<sup>g</sup> Department of Geological Sciences, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic

### ARTICLE INFO

#### Article history:

Received 19 December 2013

Accepted 21 August 2014

Available online 30 August 2014

#### Keywords:

Stand-off LIBS

Classification

PCA

LDA

Cultural heritage

### ABSTRACT

Focusing on historical aspect, during archeological excavation or restoration works of buildings or different structures built from bricks it is important to determine, preferably in-situ and in *real-time*, the locality of bricks origin. Fast classification of bricks on the base of Laser-Induced Breakdown Spectroscopy (LIBS) spectra is possible using multivariate statistical methods. Combination of principal component analysis (PCA) and linear discriminant analysis (LDA) was applied in this case. LIBS was used to classify altogether the 29 brick samples from 7 different localities. Realizing comparative study using two different LIBS setups – *stand-off* and *table-top* it is shown that *stand-off* LIBS has a big potential for archeological *in-field* measurements.

© 2014 Elsevier B.V. All rights reserved.

### 1. Introduction

Laser-Induced Breakdown Spectroscopy (LIBS) [1] is a method capable of fast classification, qualitative and semi-quantitative analysis of samples under investigation. Briefly, high energy laser pulse is focused on the sample surface, small amount of the sample is ablated and luminous laser-induced plasma (LIP) is created. The persistence of LIP is in the order of microseconds. When the plasma plume starts to cool down the best conditions for realizing LIBS measurements occur because intensive ionic and atomic emission lines can be observed in the spectra. Plasma radiation is collected and transported to the spectrometer by an optical setup. The signal of elements can be processed in real time using various chemometric analyses.

LIBS has a great potential in many fields (including industrial, environmental, cultural heritage and extraterrestrial applications), because of its particular advantages over other analytical methods,

like quasi-nondestructivity, minimum or no need of sample preparation, relatively low instrumentation cost, experimental setup flexibility including the possibility of in-situ measurements [2]. Moreover, LIBS is a very promising technique for the elemental analysis of many types of cultural heritage objects [3], such as pottery, sculptures, pigments of paintings, glass, calcified tissues, geological samples or metals; as proved before [4–18].

For the analysis of historical walls or historic buildings the possibility of in-situ measurement has many advantages, including the preliminary determination of the origin of brick clay and estimation of chemical composition. The in-situ LIBS analysis of historical building material was presented by Laserna et al. [19], dealing with analysis of the Málaga cathedral walls with a *man-portable* LIBS and later in [20] with a *stand-off* LIBS. Remote sensing is a great advantage in this point, sometimes it might be even necessary. *Stand-off* LIBS could be the solution. In our work we focus on the analysis of historical walls in order to develop a simple and fast method for classification of bricks in-situ.

Focusing on building materials Xia et al. [21] used successfully multivariate analysis (partial least squares discriminant analysis PLS-DA and the hybrid combination PCA–Adaboost) to classify LIBS spectra of various materials for the purpose of concrete recycling.

<sup>☆</sup> Selected paper from the 7th Euro-Mediterranean Symposium on Laser Induced Breakdown Spectroscopy (EMSLIBS 2013), Bari, Italy, 16–20 September 2013.

\* Corresponding author. Department of Chemistry, Faculty of Science, Masaryk University, Kamenice 5/A14, 625 00 Brno, Czech Republic.

Generally, classification of bricks can be done by detailed investigation of their chemical composition. This composition can be revealed by LIBS, but it should be noted that due to the complicated brick matrices the LIBS spectra are very complex. Much information is ignored during the univariate analysis therefore multivariate methods were used with a great success. The chosen method for our purpose is LDA with a subset of PCA scores as inputs.

PCA, method generally applicable as a classification method, was applied here only to reduce the dimensionality in the data set and to discover latent variables among samples. This linear transformation of variables into lower number of so called principal components (PCs) filters out the irrelevant information about measurements and removes highly correlated variables [22]. Using scores of limited number of principal components as inputs for LDA prevents the overfitting [23].

LDA is a multivariate statistical method for discrimination of objects up to a finite number of categories, based on a certain subset of all objects (training set). The principle of this method is maximizing the ratio of the between-class variance to the within-class variance. The decision rules obtained by classification of training set are later applied to the testing set [24,25]. LDA was applied on LIBS measurements before, concerning alloys [26] and historical buildings [27]. The closest application to ours was the classification of IR spectra by means of LDA with principal components as inputs [28,29].

In our previous work [30] combination of PCA and LDA was introduced as an effective method for classification of LIBS spectra of archeological samples. The goal of work was to distinguish 7 types of materials (soil, brick, mortar, ceramics, shell, bone, bear tooth, and human tooth). Now, instead of classifying several different materials, we use the same method for only one of them – bricks, to see how specific this method can be among samples of similar composition, which differ in locality of origin.

The methodology was tested on both, data obtained by *stand-off* and *table-top* LIBS setup. The *table-top* arrangement used for experiment was chosen with respect to possible construction of portable device for in-field measurements. Lower pulse energy was used (as an alternative for diode-pumped laser) and CCD detector – such a device would be portable, low-cost and still effective enough for this kind of classification.

## 2. Experimental

### 2.1. Samples

Set of samples consisted of 29 samples containing bricks from 7 localities. 9 samples were laboratory manufactured for testing (localities Šlapanice, Hranice na Moravě and Tallinn) and the other 20 samples come from an archeological experiment, where bricks were prepared using different firing temperatures (localities Skalka u Velimi, Těšetice, Strážovice, Pohansko). Samples are listed in Table 1.

All samples were measured using two setups

- *stand-off* LIBS setup at Central European Institute of Technology (CEITEC), Brno University of Technology, Brno, Czech Republic,
- *table-top* LIBS setup at Department of Thermal Engineering, Tsinghua University, Beijing, China.

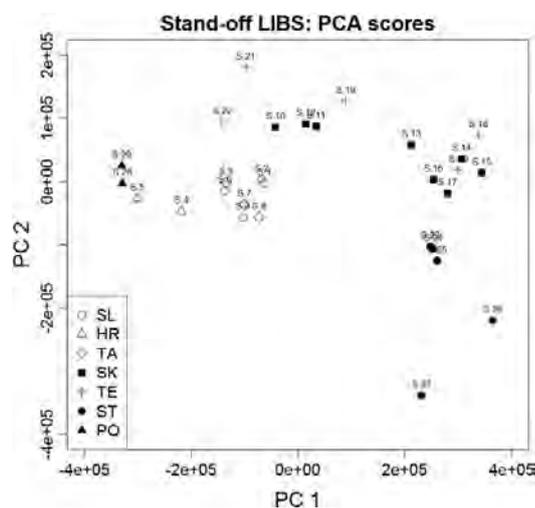
### 2.2. Stand-off LIBS instrumentation

The experimental setup was described in details elsewhere [31]. The *stand-off* LIBS setup contains a pulsed Nd:YAG laser (Solar Laser Systems, LQ 916) working on its second harmonic frequency (532 nm). Three-lens Galilean telescope optimized to have a focal spot with the smallest spherical aberration at the distance of ~10 m was used as a focusing system. The focusing system was projected to have the optimal spot size in the middle of the expected range of utilization (5–15 m) the *stand-off* LIBS apparatus is designed for. Using these parameters the diameter of ablation spot was ~1 mm at the analysis distance of

**Table 1**  
List of samples.

Sample number	Sample type name	Locality	Firing temperature (°C)	Sample set
S1	SL	Šlapanice	910	Training
S2	SL	Šlapanice	950	Training
S3	SL	Šlapanice	950	Test
S4	HR	Hranice na Moravě	1050	Training
S5	HR	Hranice na Moravě	1050	Test
S6	TA	Tallinn	950	Training
S7	TA	Tallinn	1050	Training
S8	TA	Tallinn	1100	Training
S9	TA	Tallinn	1060	Test
S10	SK	Skalka u Velimi	1000	Training
S11	SK	Skalka u Velimi	900	Training
S12	SK	Skalka u Velimi	800	Training
S13	SK	Skalka u Velimi	700	Training
S14	SK	Skalka u Velimi	600	Test
S15	SK	Skalka u Velimi	500	Training
S16	SK	Skalka u Velimi	400	Training
S17	SK	Skalka u Velimi	300	Training
S18	TE	Těšetice	900	Training
S19	TE	Těšetice	600	Training
S20	TE	Těšetice	500	Test
S21	TE	Těšetice	400	Training
S22	TE	Těšetice	300	Training
S23	ST	Strážovice	900	Training
S24	ST	Strážovice	800	Training
S25	ST	Strážovice	700	Test
S26	ST	Strážovice	600	Training
S27	ST	Strážovice	300	Training
S28	PO	Pohansko	600	Training
S29	PO	Pohansko	300	Test

6.2 m. The emission was collected by a Newtonian telescope (Sky-Watcher, Synta), with a primary mirror diameter of 10", and then transported through the optical cable (500 μm in diameter) to the spectrometer in Echelle configuration (Andor, Mechelle 5000) coupled to the ICCD detector (Andor, iStar 734i). The spectral resolution ( $\lambda/\Delta\lambda$ ) of the spectrometer (corresponding to 3 pixels FWHM) for a 50 μm in diameter entrance slit used in these measurements is about 4000. Measuring conditions were optimized in order to get the best signal to noise



**Fig. 1.** Stand-off LIBS results: clustering of localities on the first two principal components (projection pursuit method) counted from whole spectra of all samples.

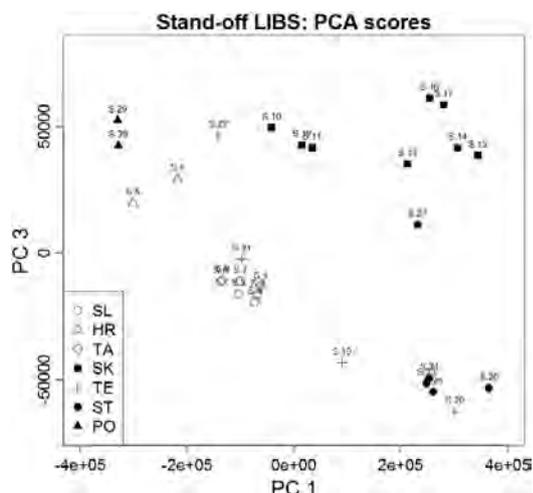


Fig. 2. Stand-off LIBS results: clustering of localities on scores of the first and the third principal component (projection pursuit method) counted from whole spectra of all samples.

ratio, the ICCD delay time and the gate width were set for 0.9  $\mu\text{s}$  and 15  $\mu\text{s}$ , respectively. Pulse energy was optimized to be 135 mJ per pulse at repetition rate of 1 Hz. Crater diameter was 1 mm that corresponds to power density of 3.4  $\text{GW}\cdot\text{cm}^{-2}$ .

### 2.3. Table-top LIBS instrumentation

The LIBS measurements were realized using a Spectrolaser 4000 (XRF, Australia). The instrument and configuration are described elsewhere [32]. The laser (Q-switched Nd:YAG laser, New Wave Research, US) has a wavelength of 532 nm and a pulse width of 5 ns. The laser pulse frequency is 1 Hz, and the gate time was fixed at 1 ms. The laser energy was optimized to be 40 mJ per pulse, with a crater diameter of 0.5 mm power density corresponding to 4.0  $\text{GW}\cdot\text{cm}^{-2}$ . Delay time

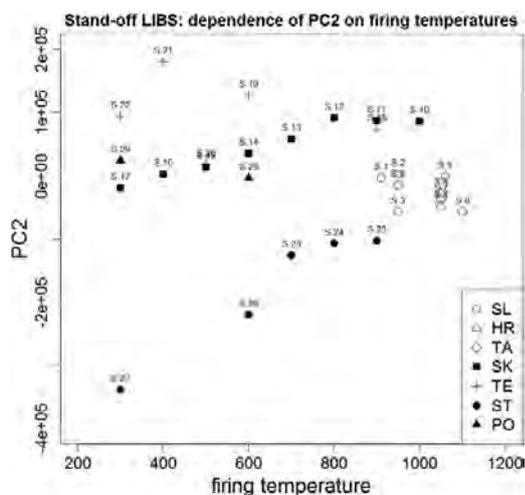


Fig. 4. Stand-off LIBS results: dependence of the second principal component (projection pursuit method) on firing temperatures of bricks.

was set for 1.5  $\mu\text{s}$ . The measuring conditions were optimized in order to get the best signal to noise ratio. The detection system is composed of four Czerny–Turner spectrometers and CCD detectors, which cover the spectral range from 190 to 940 nm with a nominal resolution of about 0.09 nm.

### 2.4. LIBS measurements

Each sample was measured at 10 spots. Surfaces were untreated because of maximal authenticity to in-field measurement; however, there were always 6 burn-off pulses before spectra recording. Spectra retrieved from the stand-off setup were then accumulated intensities coming from 10 pulses per one spot whereas spectra retrieved from table-top setup were the average intensities of the same number of pulses.

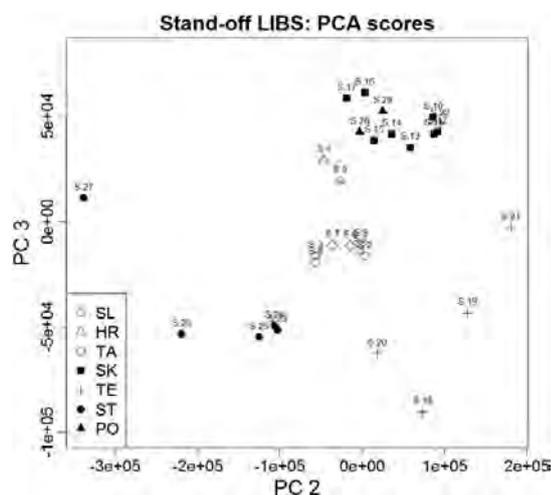


Fig. 3. Stand-off LIBS results: clustering of localities on scores of the second and the third principal component (projection pursuit method) counted from whole spectra of all samples.

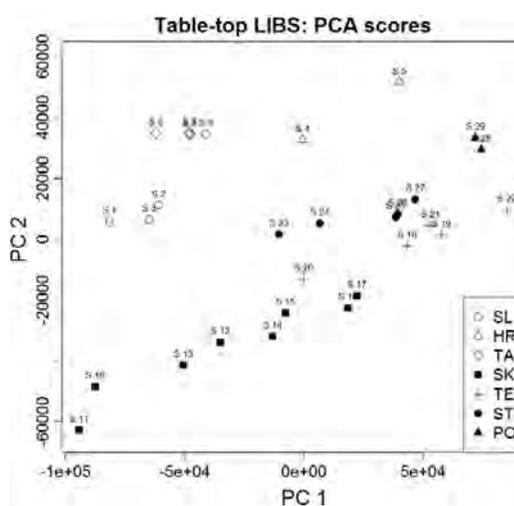


Fig. 5. Table-top LIBS results: clustering of localities on scores of first two principal components (projection pursuit method) counted from whole spectra of all samples.

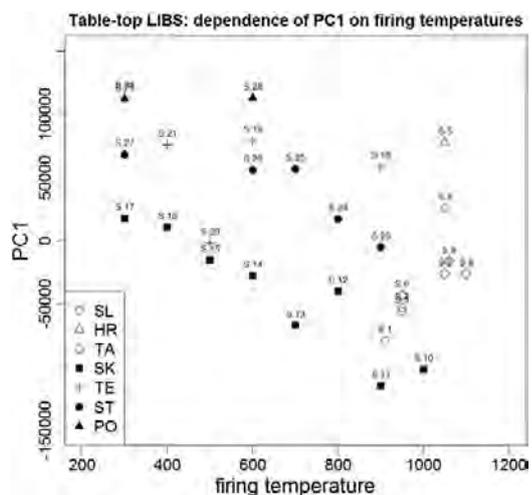


Fig. 6. Table-top LIBS results: dependence of first principal component (projection pursuit method) computed from whole spectra of all samples (no baseline pretreatment) on firing temperatures of bricks.

## 2.5. Statistical treatment

Experiment was as follows: 10 LIBS spectra per sample were averaged and centered. The whole spectra (200–1000 nm in case of *stand-off*, 190–950 nm in case of *table-top* LIBS) were used for PCA computing. Then first 5 PCs were inputted into LDA in order to classify the samples according to their locality of origin. 7 samples of 29 were taken as test samples (one from each locality) and the other 22 samples were used for training. Division into test and training set is obvious from Table 1.

During principal component analysis the influence of bricks firing temperature on LIBS spectra was observed. A strong temperature influence occurred in PCs computed from whole spectra. In PCs computed by using only segments of spectra (with lines of selected elements) most of this information was suppressed. This phenomenon was closely

investigated in order to find out how exactly the spectra are affected, what is the reason and how we could possibly use it.

All computations were done in the software R [33]. In case that baseline subtraction was done, “loess” method included in library “PROcess” was used. For calculation of PCA a robust method was used (projection pursuit method), which is included in library “rrcov”. For LDA it was maximum likelihood method that was included in library “MASS”.

The same statistical treatment was used to process data from both, *stand-off* and *table-top* LIBS setups and results were compared.

## 2.6. X-Ray Powder Diffraction instrumentation

X-ray Powder Diffraction (XRPD) was used in order to clarify the structural and chemical changes occurring during firing of bricks. Five samples from the same locality (ST) were examined. Analysis was performed using Bruker D8 Advance diffractometer with copper tube ( $\lambda K_{\alpha} = 0.15418$  nm), powered at 40 kV and 30 mA, 1-D position sensitive detector, and programmable divergence slits at the conventional Bragg–Brentano parafocussing  $\Theta$ – $\Theta$  reflection geometry; step size –  $0.02^{\circ}2\theta$ , time per step – 188 s, angular range  $5$ – $80^{\circ}2\theta$ . The measured data was processed using a DIFFRAC plus software and an ICDD PDF 2 database.

## 3. Results and discussion

### 3.1. Principal component analysis

Let's focus on *stand-off* LIBS results first. After baseline subtraction, PCA was applied on spectra. From scree plot, it was found out that the first score carries most of the information (87.25% of total variability) and about next 4 could be informative as well. First three of them are plotted against each other in Figs. 1–3. Individual localities seem to create well separated clusters. There are also visible certain continuities in sample orders among some groups. These happen mostly within the groups of samples SK and ST and in smaller extend in the group of TE samples. It indicates a dependence of firing temperature of bricks (see Table 1). For example S 27 in Figs. 1 and 3 lies far from other points related to ST. Its firing temperature was 300 °C so that a lot lower than of other samples. For a better view into the problem of this finding, firing temperature was plotted against first few principal components. The dependence was not only confirmed, but also localized: the most

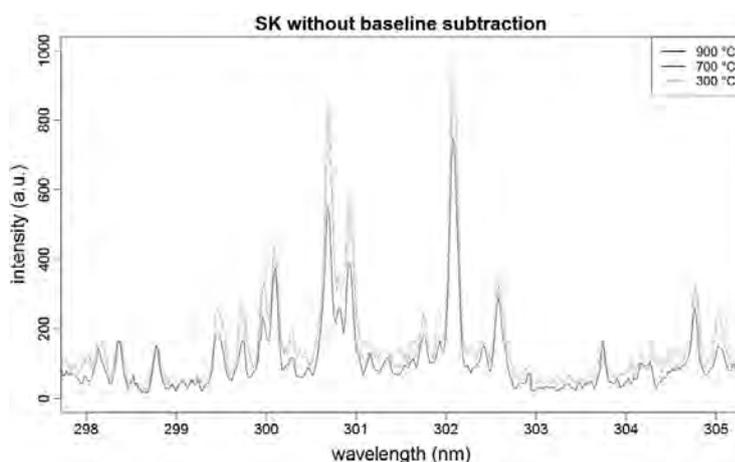


Fig. 7. Cut of *stand-off* LIBS spectra (with lines of Fe, Si, Ca): SK locality, without baseline subtraction.

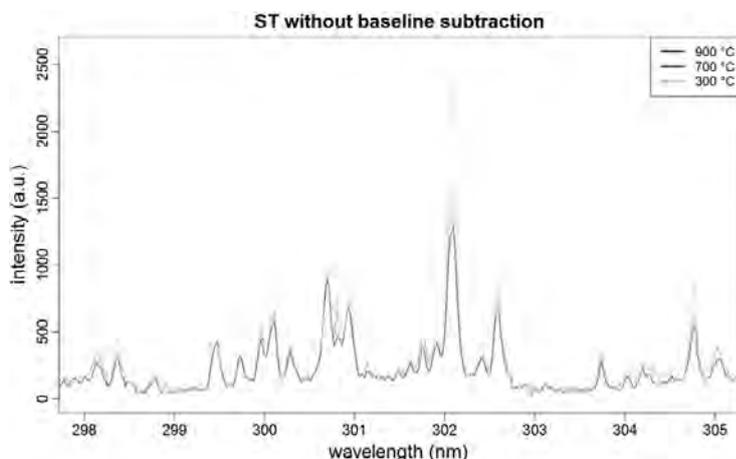


Fig. 8. Cut of *stand-off* LIBS spectra (with lines of Fe, Si, Ca): ST locality, without baseline subtraction.

significant dependence was observed on the second principal component (Fig. 4), though in the PC 1, PC 3 and PC 5 were observable too.

If we look at the *table-top* results, scree plot shows first 3 PCs containing most of the information. The first two principal components together account 75.51 + 20.23% of the total variability. When plotted against each other, first few principal components create well separated clusters of individual localities (see for example first two PCs shown in Fig. 5). The firing temperature dependence on PCs was observed only in spectra without baseline subtraction and was not so strong as in the case of *stand-off* LIBS. Fig. 6 shows the first principal component (no baseline subtraction) plotted against the firing temperature because there was a slight influence on ST data.

### 3.2. Temperature influence in LIBS spectra

None of the lines itself seems to be the bearer of the information on firing temperature. This dependence has more complex effect on spectra. Going through the raw spectra related to particular localities it was found out that this information is connected to the intensity of lines as well as to the baseline height. This is obvious from Fig. 7,

where there is the cut of three SK samples spectra measured using *stand-off* LIBS. At most of the wavelengths of SK samples spectra (*stand-off* LIBS) the same phenomenon occurs: increasing the firing temperature while decreasing baseline and maximum intensities.

Character of this dependence varies between classes of particular localities. ST *stand-off* LIBS spectra have slightly different patterns (Fig. 8). The reason can be in the mineralogical composition of clays, consequently in a degree of sintering or hardness of material. All this leads to a different course of ablation and can cause changes like this in spectra. Samples of SK were made of marl; samples of ST were made of clay.

Influence of temperature observable in *table-top* LIBS spectra is shown in Fig. 9 (the same range of wavelengths as in the case of *stand-off* LIBS before). Again, the consequence of maximum intensities according to firing temperature was almost the same in this range, as everywhere else in the spectra. This consequence is slightly different from ST samples (Fig. 10) and evidently very different to the spectra measured on *stand-off* LIBS (see Figs. 7 and 8). In fact it was surprisingly quite opposite of *stand-off* LIBS consequence.

The lines are not as well resolved as in the case of *stand-off* LIBS. This is caused by using Czerny–Turner spectrometer with CCD detector.

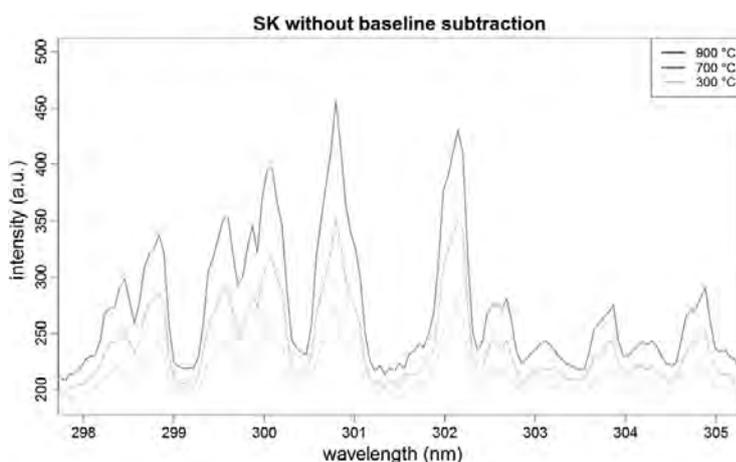


Fig. 9. Cut of *table-top* LIBS spectra (with lines of Fe, Si, Ca): SK locality, without baseline subtraction.

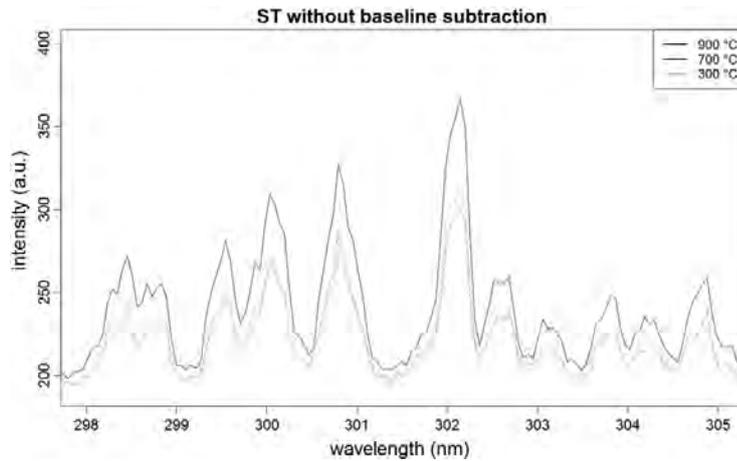


Fig. 10. Cut of *table-top* LIBS spectra (with lines of Fe, Si, Ca): ST locality, without baseline subtraction.

Another difference between spectra from these two setups is in the baselines. The Echelle spectrometer recorded spectra where baselines are less uniform. For that reason baseline subtraction was not so effective and the information about firing temperature of bricks appeared in principal components. After subtraction of baseline from the *table-top* spectra the firing temperature influence was suppressed.

Considering the assumption of hardness of material effecting the spectra (via different course of ablation) the ratios of the ionic magnesium lines at 280.27 nm to the neutral Mg lines at 285.21 nm were plotted against the firing temperature of bricks. This approach was used on the base of [34], because bricks fired on higher temperature should have higher hardness. Plasma created by ablation of materials which differ in hardness has different thresholds and thermodynamic properties. This will be reflected in line ratios. Focusing on *stand-off* LIBS data, samples of SK locality show the dependence (Fig. 11), which can be expressed by  $R^2 = 0.86$  and the samples of ST (Fig. 12) by  $R^2 = 0.8$ . The spectra after

baseline subtraction were used to make these relations, but we consider that the baseline shift is also caused by this effect. Even if MgII/MgI ratios confirm this “dependence of hardness” theory, from baselines we weren’t able to separate this influence.

Correlation of Mg II/MgI ratios and the firing temperature of bricks significantly deteriorated when using the *table-top* LIBS data. This can be caused by plasma conditions (lower energy, smaller crater) or by measuring conditions (longer delay time, CCD detector).

### 3.3. X-ray Powder Diffraction

In order to elucidate the structural and chemical changes occurring during brick manufactory XRPD analysis was performed. Samples of ST were measured and mineralogical changes with increasing temperature are shown in Figs. 13 and 14. Between 300 °C and 600 °C pyrite

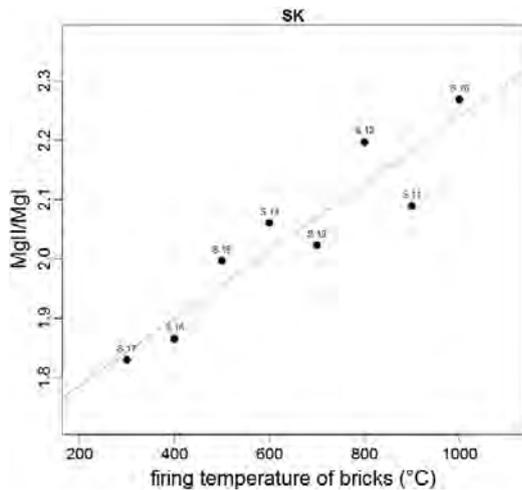


Fig. 11. *Stand-off* LIBS results: Dependence of MgII/MgI ratio on firing temperature of bricks. Locality SK.

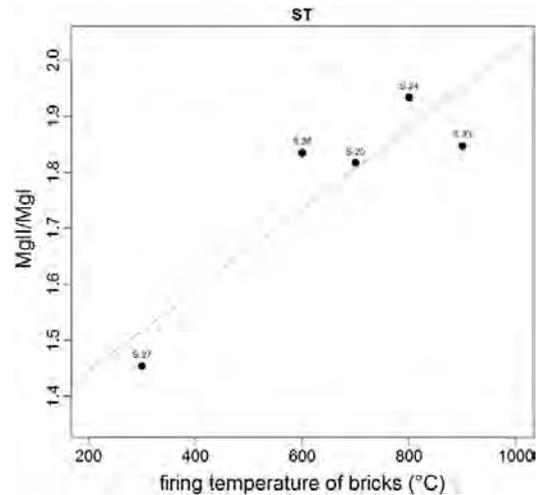


Fig. 12. *Stand-off* LIBS results: Dependence of MgII/MgI ratio on firing temperature of bricks. Locality ST.

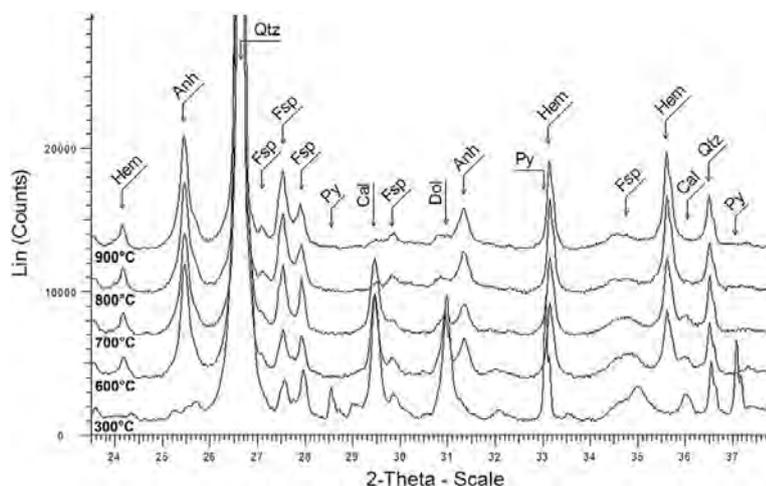


Fig. 13. X-ray Powder Diffraction results: Locality ST. On X-axis: diffraction angle  $2\theta$  (degrees), on Y-axis: intensity (counts). Hem = hematite, Anh = anhydrite, Qtz = quartz, Py = pyrite, Fsp = feldspar, Cal = calcite, Dol = dolomite.

decomposition and formation of hematite take place (Fig. 13) and also dehydroxylation of kaolinite followed by formation of amorphous metakaolinite, which usually starts below 500 °C and finishes up to 600 °C (Fig. 14). One of the major clay mineral constituents, illite, begins to decompose at 600 °C (dehydroxylation and collapse of the crystal structure takes place) what is observable due to shifted and split peaks of illite and mica from 600 to 800 °C (Fig. 14). Decarbonation starts at 600 °C: dolomite decomposes to calcite, MgO and CO<sub>2</sub> and at 800 °C dolomite is completely decomposed (Fig. 13). Calcite decomposition partly takes place simultaneously with the decomposition of dolomite, starts at a slightly higher temperature – see scans of 600–800 °C at Fig. 15. At 900 °C calcite is no longer present. Above 300 °C

and below 600 °C anhydrite II is formed (anhydrite II generally appears at temperatures above 400 °C). The content of anhydrite II does not change between 600 and 900 °C (Fig. 13). Anhydrite probably arises from gypsum and bassanite. In the 300 °C scan, however, anhydrite III, normally formed from bassanite at this temperature, is not present. That could be attributed to material inhomogeneity. In the temperature range of 600–800 °C also smectites and/or chlorite decompose (Fig. 14).

All these mineralogical changes affect the final properties of bricks, the degree of sintering and consequently hardness of the material. Using LIBS the combination of two factors affects the firing temperature symptoms in spectra: the composition of clay and measuring conditions.

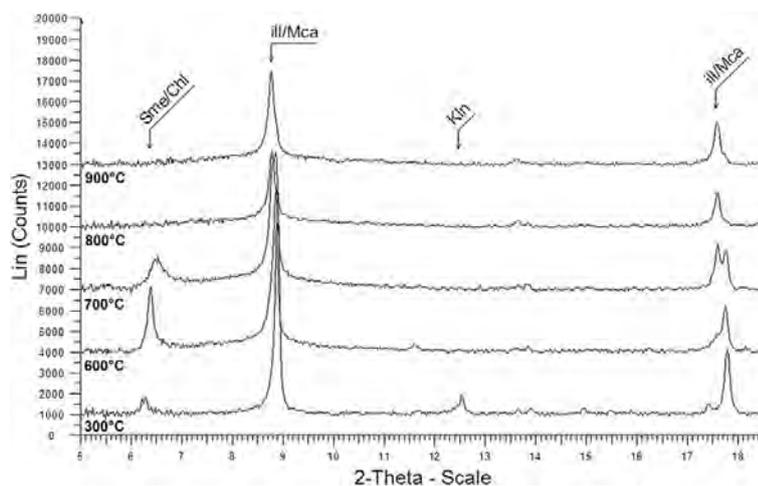


Fig. 14. X-ray Powder Diffraction results: Locality ST. On X-axis: diffraction angle  $2\theta$  (degrees), on Y-axis: intensity (counts). Sme = smectite, Chl = chlorite, Ill = illite, Mca = mica, Kln = kaolinite.

### 3.4. Linear discriminant analysis

After PCA computations linear discriminant analysis follows as a method chosen for estimation localities of bricks origin.

First comes the question of how many PCs should be used as inputs for LDA. As referenced before in [30], Defernez and Kemsley [23] stated that the onset of overfitting occurs when the dimensionality exceeds  $(n - g) / 3$ , where  $n$  is the number of specimens (training samples) and  $g$  is the number of groups (localities). This means that the number of adjustable parameters for our data set should be 5.

The predictions made for test samples are listed in Table 2. Using first 5 PCs computed from *stand-off* LIBS data, 6 of 7 test samples were classified correctly, whereas using *table-top* LIBS data provides 7 correct out of 7 predictions, which is a great success.

The wrong prediction appearing in the case of *stand-off* LIBS data processing was made by substitution sample of locality HR for the PO sample. Distribution of samples on PCA plots (Figs. 1 and 2) matches with this misclassification. HR test sample No. 5 clusters there with the PO samples on the first PC. The first PC was affected by firing temperature in smaller extend, but still it is difficult to decide whether the cause of the misclassification was this effect or simply the properties of samples.

The effect of firing temperature on principal components caused elongation of clusters of most localities (see Figs. 1–3 and 5). LDA is a method applicable for processing such elongated clusters [35]. Figs. 15 and 16 describe the geometric illustration of LDA. Discriminant coordinates display the primary differences between clusters and confirm that elongation of most clusters was eliminated and so was the effect of firing temperature on classification. Nevertheless HR and PO samples are still mixed together in case of *stand-off* LIBS and HR samples have great dispersion even in the case of *table-top* LIBS. Considering the fact that both groups consist of only two samples it is not easy to classify them correctly.

## 4. Conclusion

Before real application of *stand-off* or *portable* LIBS device, respectively, for archeological purposes there is a need to develop a fast and simple methodology of data processing. Focusing on historic walls we measured spectra of brick samples using two different LIBS setups

**Table 2**

The predictions for test samples created by LDA (ML). In the third column there are predictions given by LDA when *stand-off* LIBS data were put in, in the fourth when *table-top* LIBS data were used. The incorrect predictions are marked in bold.

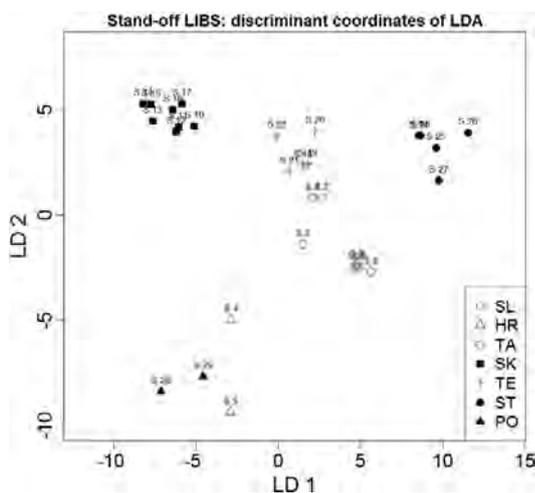
Test sample no.	Real locality	LDA prediction <i>stand-off</i> LIBS	LDA prediction <i>table-top</i> LIBS
3	SL	SL	SL
5	HR	<b>PO</b>	HR
9	TA	TA	TA
14	SK	SK	SK
20	TE	TE	TE
25	ST	ST	ST
29	PO	PO	PO

(*stand-off* and *table-top*) on which the proposed methodology was tested and compared.

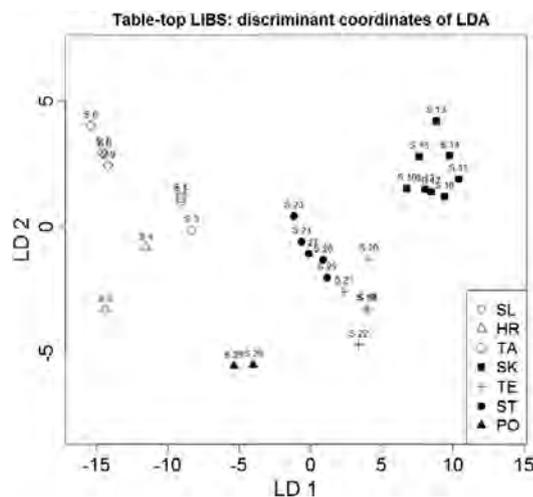
With the aim of bricks classification according to the place of origin via LDA, a robust PCA was applied on whole spectra (from about 200 to 1000 nm). PCA plots provided elongated clusters of particular localities, which showed signs of certain dependence in connection with bricks firing temperature. This effect was closely investigated with the prospect of estimation of firing temperature of unknown samples. This closer look showed that realizing this vision can be possible only using spectrometer with higher resolution and ICCD detector. On the other hand, it is achievable in *stand-off* LIBS configuration.

After dividing the sample set into training and test samples, the first five principal components were used as inputs and the prediction of localities were computed. Sample set contained 29 bricks from 7 different localities. Higher number of samples would be better for this multivariate statistic evaluation, but for methodology creation it was important to have samples with all possible information known. Information on firing temperature was confirmed to be valuable. To summarize the results of classification, LDA predicted one of 7 localities wrong in case of *stand-off* LIBS data, all predictions were right in case of inputting *table-top* LIBS spectra. LDA clusters had not showed the signs of elongation so the firing temperature influence of classification was mostly eliminated.

It is obvious that for classification of bricks the differences between setups used had not much crucial effect. On the contrary, lower resolution of CCD detector (contained in *table-top* LIBS used) helped to improve the LDA prediction compared to *stand-off* LIBS setup with Echelle spectrometer. This shows that proposed methodology will be



**Fig. 15.** *Stand-off* LIBS results: clustering of localities on the first two LDA discriminant coordinates.



**Fig. 16.** *Table-top* LIBS results: clustering of localities on the first two LDA discriminant coordinates.

convenient for our purpose of brick spectra classification using both setups and after creating a complex database of spectra it will definitely help with fast classification in situ. The databases have to be created separately for only one instrumentation and united measuring conditions.

### Acknowledgment

This work was supported by the project “CEITEC – Central European Institute of Technology” (CZ.1.05/1.1.00/02.0068) from the European Regional Development Fund. J.N. acknowledges the support from the European Regional Development Fund in the frame of project “Support of Interdisciplinary Excellence Research Teams Establishment at BUT” (CZ.1.07/2.3.00/30.0005).

### References

- [1] A.W. Miziolek, V. Palleschi, I. Schechter, *Laser-Induced Breakdown Spectroscopy, Fundamentals and Applications*, University Press, Cambridge, 2006.
- [2] J.D. Winefordner, I.B. Gornushkin, T. Correll, E. Gibb, B.W. Smith, N. Omenetto, Comparing several atomic spectrometric methods to the super stars: special emphasis on laser induced breakdown spectrometry, LIBS, a future superstar, *J. Anal. At. Spectrom.* 19 (2004) 1061–1083.
- [3] R. Gaudiuso, M. Dell’Aglia, O. De Pascale, G.S. Senesi, A. De Giacomo, Laser induced breakdown spectroscopy for elemental analysis in environmental, cultural heritage and space applications: a review of methods and results, *Sensors* 10 (2010) 7434–7468.
- [4] A. Giakoumaki, K. Melessanaki, D. Anglos, Laser-Induced Breakdown Spectroscopy (LIBS) in archaeological science – applications and prospects, *Anal. Bioanal. Chem.* 387 (2007) 749–760.
- [5] 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.
- [6] A.J. López, G. Nicolás, M.P. Mateo, A. Ramil, V. Piñón, A. Yáñez, LIBS and linear correlation analysis applied to the classification of Roman pottery terra sigillata, *Appl. Phys. A* 83 (2006) 695–698.
- [7] E.M. Rodríguez-Celis, I.B. Gornushkin, U.M. Heitman, J.R. Almirall, B.W. Smith, J.D. Winefordner, N. Omenetto, Laser induced breakdown spectroscopy as a tool for discrimination of glass for forensic applications, *Anal. Bioanal. Chem.* 391 (2008) 1961–1968.
- [8] I.B. Gornushkin, U. Panne, J.D. Winefordner, Linear correlation for identification of materials by laser induced breakdown spectroscopy: improvement via spectral filtering and masking, *Spectrochim. Acta Part B* 64 (2009) 1040–1047.
- [9] N.J. McMillan, R.S. Harmon, F.C. De Lucia, A.W. Miziolek, Laser-induced breakdown spectroscopy analysis of minerals: carbonates and silicates, *Spectrochim. Acta Part B* 62 (2007) 1528–1536.
- [10] S.M. Clegg, E. Sklute, M. Darby Dyar, J.E. Barefield, R.C. Wiens, Multivariate analysis of remote laser-induced breakdown spectroscopy spectra using partial least squares, principal component analysis, and related techniques, *Spectrochim. Acta Part B* 64 (2009) 79–88.
- [11] J.L. Luque-García, R. Soto-Ayala, M.D. Luque de Castro, Determination of the major elements in homogeneous and heterogeneous samples by tandem laser-induced breakdown spectroscopy – partial least squares regression, *Microchem. J.* 73 (2002) 355–362.
- [12] R.S. Harmon, J. Remus, N.J. McMillan, C. McManus, L. Collins, J.L. Gottfried Jr., F.C. De Lucia, A.W. Miziolek, LIBS analysis of geomaterials: geochemical fingerprinting for the rapid analysis and discrimination of minerals, *Appl. Geochem.* 24 (2009) 1125–1141.
- [13] J.L. Gottfried, R.S. Harmon, F.C. De Lucia Jr., A.W. Miziolek, Multivariate analysis of laser-induced breakdown spectroscopy chemical signatures for geomaterial classification, *Spectrochim. Acta Part B* 64 (2009) 1009–1019.
- [14] K. Novotný, J. Kaiser, M. Galiová, V. Konečná, J. Novotný, R. Malina, M. Liška, V. Kanický, V. Otruba, Mapping of different structures on large area of granite sample using laser-ablation based analytical techniques, an exploratory study, *Spectrochim. Acta Part B* 63 (2008) 1139–1144.
- [15] D.L. Death, A.P. Cunningham, L.J. Pollard, Multi-element analysis of iron ore pellets by laser-induced breakdown spectroscopy and principal components regression, *Spectrochim. Acta Part B* 63 (2008) 763–769.
- [16] J.M. Andrade, G. Cristoforetti, S. Legnaioli, G. Lorenzetti, V. Palleschi, A.A. Shaltout, Classical univariate calibration and partial least squares for quantitative analysis of brass samples by laser-induced breakdown spectroscopy, *Spectrochim. Acta Part B* 65 (2010) 658–663.
- [17] A. Jurado-López, M.D. Luque de Castro, Rank correlation of laser-induced breakdown spectroscopic data for the identification of alloys used in jewelry manufacture, *Spectrochim. Acta Part B* 58 (2003) 1291–1299.
- [18] P. Inakollu, T. Philip, A.K. Rai, F.-Y. Yueh, J.P. Singh, A comparative study of laser induced breakdown spectroscopy analysis for element concentrations in aluminum alloy using artificial neural networks and calibration methods, *Spectrochim. Acta Part B* 64 (2009) 99–104.
- [19] F.J. Fortes, J. Cuñat, L.M. Cabalín, J.J. Laserna, In situ analytical assessment and chemical imaging of historical buildings using a man-portable laser system, *Appl. Spectrosc.* 61 (2007) 558–564.
- [20] I. Gaona, P. Lucena, J. Moros, F.J. Fortes, S. Guirado, J. Serrano, J.J. Laserna, Evaluating the use of standoff LIBS in architectural heritage: surveying the Cathedral of Málaga, *J. Anal. At. Spectrom.* 28 (2013) 810–820.
- [21] H. Xia, M.C.M. Bakker, Reliable classification of moving waste materials with LIBS in concrete recycling, *Talanta*. <http://dx.doi.org/10.1016/j.talanta.2013.11.082>.
- [22] R.G. Brereton, *Chemometrics for Pattern Recognition*, John Wiley and Sons Ltd., Bristol, UK, 2009.
- [23] M. Defernez, E.K. Kemsley, The use and misuse of chemometrics for treating classification problems, *Trends Anal. Chem.* 16 (1997) 216–221.
- [24] B.D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge University Press, Cambridge, 1996.
- [25] A.R. Webb, *Statistical Pattern Recognition*, 2nd ed. Wiley, Chichester, 2002.
- [26] S.R. Goode, S.L. Morgan, R. Hoskins, A. Oxsher, Identifying alloys by laser-induced breakdown spectroscopy with a time-resolved high resolution Echelle spectrometer, *J. Anal. At. Spectrom.* 15 (2000) 1133–1138.
- [27] 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* 65 (2010) 688–694.
- [28] W. Wu, D.L. Massart, S. de Jong, Kernel-PCA algorithms for wide data part II: fast cross-validation and application in classification of NIR data, *Chemom. Intell. Lab. Syst.* 37 (1997) 271–280.
- [29] E.K. Kemsley, S. Ruault, R.H. Wilson, Discrimination between *Coffea arabica* and *Coffea canephora* variant *robusta* beans using infrared spectroscopy, *Food Chem.* 54 (1995) 321–326.
- [30] 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* 73 (2012) 1–6.
- [31] A. Hrdlička, L. Prokeš, A. Staňková, K. Novotný, A. Vitešnická, V. Kanický, V. Otruba, J. Kaiser, J. Novotný, R. Malina, K. Páleníková, Development of a remote laser-induced breakdown spectroscopy system for investigation of calcified tissue samples, *Appl. Opt.* 49 (2010) C16–C20.
- [32] J. Feng, Z. Wang, Z. Li, W. Ni, Study to reduce laser-induced breakdown spectroscopy measurement uncertainty using plasma characteristic parameters, *Spectrochim. Acta Part B* 65 (2010) 549–556.
- [33] <http://cran.r-project.org/>.
- [34] Z.A. Abdel-Salam, Z. Nanjing, D. Anglos, Effect of experimental conditions on surface hardness measurements of calcified tissues via LIBS, *Appl. Phys. B* 94 (2009) 141–147, <http://dx.doi.org/10.1007/s00340-008-3304-z>.
- [35] K. Varmuza, P. Filzmoser, *Introduction to Multivariate Statistical Analysis in Chemometrics*, Taylor & Francis – CRC Press, Boca Raton, FL, 2009.

## Article 5

P. Pořízka, A. Demidov, J. Kaiser, J. Keivanian, I. Gornushkin, U. Panne, and J. Riedel. Laser-induced breakdown spectroscopy for in situ qualitative and quantitative analysis of mineral ores. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 101:155–163, 2014.

DOI: 10.1016/j.sab.2014.08.027; 24 citations

This publication delivered consideration on the necessity of utilization of multivariate algorithms and their supremacy over univariate ones. It was shown that judicious data processing enables more accurate quantitative analysis than simple, black-box, implementation of chemometrics. Such issue was discussed also in this thesis.

First, the sample set was measured with two slightly different instrumental setups, differing in the collection optics (side view *versus* top-view). This proved that minor changes in the experimental settings can induce major changes in the data matrix structure, which was observed when visualizing the data in the PC space. Interestingly, the top-view arrangement resulted in higher pulse-to-pulse fluctuation, expressed as the standard deviation.

Second, sample set of igneous rock distributed in three different matrices were measured and the calibration of both systems to Cu content was constructed. Differences in sample matrices led to the matrix effect of the laser ablation and divided the calibration curve to three sections. Utilization of principal component regression (PCR) and partial least squares regression (PLSR) were not able to suppress the matrix effect and the coefficient of determination reached not-satisfying 79 %.

The data set was classified in the PC space using Gaussian clustering algorithm. Then, individual data sets were subjected to univariate calibration. This methodological approach resulted in the improvement of bias by the order of magnitude.



## Laser-induced breakdown spectroscopy for *in situ* qualitative and quantitative analysis of mineral ores<sup>☆</sup>



P. Pořízka<sup>a,c</sup>, A. Demidov<sup>a</sup>, J. Kaiser<sup>c</sup>, J. Keivanian<sup>d</sup>, I. Gornushkin<sup>a</sup>, U. Panne<sup>a,b</sup>, J. Riedel<sup>a,\*</sup>

<sup>a</sup> BAM, Federal Institute for Materials Research and Testing, Richard Willstätter-Strasse 11, D-12489 Berlin, Germany

<sup>b</sup> Chemistry Department, Humboldt Universität zu Berlin, Brook-Taylor-Strasse 2, D-12489 Berlin, Germany

<sup>c</sup> Institute of Physical Engineering, Faculty of Mechanical Engineering, Brno University of Technology, Technická 2896/2, 61669 Brno, Czech Republic

<sup>d</sup> Institute for Mining, Technical University Clausthal, Erzstraße 18, 38678 Clausthal-Zellerfeld, Germany

### ARTICLE INFO

#### Article history:

Received 26 December 2013

Accepted 8 August 2014

Available online 27 August 2014

#### Keywords:

Laser-induced breakdown spectroscopy

LIBS

Chemometrics

Principal component analysis

Geochemical analysis

### ABSTRACT

In this work, the potential of laser-induced breakdown spectroscopy (LIBS) for discrimination and analysis of geological materials was examined. The research was focused on classification of mineral ores using their LIBS spectra prior to quantitative determination of copper. Quantitative analysis is not a trivial task in LIBS measurement because intensities of emission lines in laser-induced plasmas (LIP) are strongly affected by the sample matrix (matrix effect). To circumvent this effect, typically matrix-matched standards are used to obtain matrix-dependent calibration curves. If the sample set consists of a mixture of different matrices, even in this approach, the corresponding matrix has to be known *prior* to the downstream data analysis. For this categorization, the multielemental character of LIBS spectra can be of help. In this contribution, a principal component analysis (PCA) was employed on the measured data set to discriminate individual rocks as individual matrices against each other according to their overall elemental composition. Twenty-seven igneous rock samples were analyzed in the form of fine dust, classified and subsequently quantitatively analyzed. Two different LIBS setups in two laboratories were used to prove the reproducibility of classification and quantification. A superposition of partial calibration plots constructed from the individual clustered data displayed a large improvement in precision and accuracy compared to the calibration plot constructed from all ore samples. The classification of mineral samples with complex matrices can thus be recommended prior to LIBS system calibration and quantitative analysis.

© 2014 Elsevier B.V. All rights reserved.

### 1. Introduction

With a steadily increasing consumption of mineral resources, it is crucial to optimize the exploration of new ore deposits and the mining process itself with reliable and rapid analytical methods for identification of minerals. However, reliable *in situ* identification of minerals can be very challenging. Currently, in-field identification of minerals is typically based on visual examination of their physical properties by an experienced geologist, which, however, can result in false positives [1]. More advanced laboratory-based techniques, such as X-ray fluorescence (XRF) and inductively coupled plasma optical emission or mass spectrometry (ICP-OES/MS) after wet digestion of samples [2] are not always suitable due to their long turnaround times and analysis cost per sample.

Laser-induced breakdown spectroscopy (LIBS) can meet the challenge as a field technique for the identification and analysis of various minerals. LIBS became a popular technique in harsh environments due

to its fast and non-demanding measurement routine, minor need for sample preparation, and low-cost instrumentation compared to other atomic emission spectroscopic techniques. LIBS provides qualitative and quantitative information about samples under investigation in *real-time* and *in situ* with an inherent multielemental capability [3,4]. The observed LIBS spectrum reflects the entire elemental composition, i.e. chemical fingerprint, of the sample.

The composition of igneous rocks differs with their provenance as well as with the way of their alterations [5]. Alteration is the property of a rock that explains its chemical and mineralogical changes in the course of time. In geology, alteration is important because it may have an effect on the grades of included elements (e.g. copper). Therefore, rocks are classified in individual alteration types. Different mineral ores and their alterations can be discriminated by examining their chemical fingerprints, especially the relative abundance of matrix elements (e.g. Al, Ca, K, Na, and Si) [6]. Igneous rocks can be clustered into individual groups forming a QAPF (quartz, alkali feldspar, plagioclase, and feldspathoid) diagram [5]. Different areas in this diagram span up natural rock types; each rock type displays a certain continuous variation of a mineral content [5,7]. *Nota bene* the exact classification of mineral types does not rely on the knowledge of concentrations of only one or two characteristic elements but has to be derived using the

<sup>☆</sup> Selected paper from the 7th Euro-Mediterranean Symposium on Laser Induced Breakdown Spectroscopy (EMSLIBS 2013), Bari, Italy, 16–20 September 2013.

\* Corresponding author.

E-mail address: [jens.riedel@bam.de](mailto:jens.riedel@bam.de) (J. Riedel).

complete elemental fingerprint of the matrix. A well-established method for multidimensional classification of complex and large data sets is the principal components analysis (PCA). It is, thus, expected that PCA discrimination of samples represented by their LIBS spectra can emulate the distribution of the rock types in the QAPF diagram.

Earlier applications of LIBS in geology were comprehensively reviewed by Harmon *et al.* [8]. LIBS has already been used for quantitative analysis of mineral samples including field-portable devices [1, 9–11]. Multivariate statistical approaches for the identification of different kinds of rocks and minerals were employed using both laboratory bench-top and stand-off LIBS systems [12]. Harmon *et al.* [13] demonstrated the possibility of ascertaining the provenance of conflict minerals employing the partial least squares discriminant analysis (PLS-DA) of LIBS spectra. Bousquet *et al.* [14] tested the hypothesis that the most significant differences between soils comes from the varying amounts of matrix elements. Further successful applications of PCA to LIBS spectra can be found in [15–18].

In this work, the potential of LIBS for determination of copper in mineral ores is investigated based on a preliminary identification of ores according to their chemical compositions. This pre-discrimination into individual classes of minerals allows for a reduction of the strong influence of sample compositions on intensities of copper emission lines induced by the mineral matrix. The principal component analysis (PCA) is used for unsupervised classification of the mineral ores into different groups prior to the quantitative determination of the copper content. Partial calibration plots are constructed from these groups of samples and compared to the calibration plot constructed from all certified ore samples simultaneously.

## 2. Experimental

Representative samples were collected at the Sungung copper mine located in East Azerbaijan, Iran. The sample set covers 27 samples consisting of three igneous rock types (andesite, ANS; diorite, DIO, and monzonite to quartz monzonite, KP) each consisting of different alterations (Table 1). The samples were received from the Clausthal University of Technology in Clausthal, Germany, where they were analyzed by ICP-MS after four acid digestions. The amount of copper ranged

between 0.06 and 0.79 weight percent. Apart from the chemical analysis, the samples were classified into groups and arranged in the QAPF diagram by an experienced geologist who tested samples by visual inspection. After this analysis and classification, the samples were analyzed with two different LIBS systems to ensure the independence of the classification results upon the individual LIBS system and, thus, to prove the robustness of the selected multivariate technique for unambiguous sample classification. Samples were measured in the form of a fine dust pressed onto the surface of the double-sided sticky tape; this approach was earlier suggested by Gornushkin *et al.* [19,20]. Samples were placed on a motorized translational stage to provide a fresh spot for every laser pulse.

System 1 consists of a high-energy Nd:YAG laser (Continuum Spitlight-10, 10 Hz, 1064 nm, 10 ns) focused onto the sample with an irradiance of  $30 \text{ GW cm}^{-2}$  (the spot diameter is  $\approx 250 \mu\text{m}$ ) using a 100 mm focal length plano-convex lens. The radiation from the luminous plasma is collected using a large aperture collector-collimator (Andor CC52,  $f/2$ ) placed 250 mm above the interaction region at a  $30^\circ$  angle with respect to the laser axis. The collector is coupled to a 400  $\mu\text{m}$  optical fiber that delivers light at the entrance slit of an echelle spectrometer (LTB Aryelle Butterfly). The resolving power and spectral range of the spectrometer are 15,000 and 300–600 nm, respectively. The spectral information is recorded by an intensified CCD (Andor iStar 734,  $1024 \times 1024$  pixels with an effective pixel size of  $13 \times 13 \mu\text{m}$ ). The ICCD operates at  $120\times$  gain and a binning of  $2 \times 2$  pixels. The individual timings of the experiment were previously optimized and set to 2  $\mu\text{s}$  gate delay and 10  $\mu\text{s}$  gate width. The whole LIBS setup is triggered by a delay generator (DG535, Stanford Research Systems).

The complementary LIBS system 2 consists of a diode-pumped solid state (DPSS) laser (Quantel Ultra 100) operated at the fundamental wavelength of  $\lambda = 1064 \text{ nm}$  and 5 Hz repetition rate. A short laser pulse (8 ns) is focused using a 450 mm focal length plano-convex lens creating a spot diameter  $\approx 200 \mu\text{m}$  on the target surface with the irradiance  $20 \text{ GW cm}^{-2}$ . The plasma radiation is collected top-on (collinearly with the laser beam) by a toroidal mirror with 300 mm focal length and  $f/7$ . The radiation is detected by a CCD (Andor Newton,  $1024 \times 256$  pixels with an effective pixel size of  $26 \times 26 \mu\text{m}$ ) attached to an echelle spectrometer (LTB Aryelle 400) with a spectral working range of 200–600 nm, resolving power 15,000, and  $f/10$ . The spectrometer is equipped with an optomechanical chopper to cut off the plasma continuum radiation during the first 0.3  $\mu\text{s}$  of the plasma evolution, the gate width is also determined by the chopper to  $\sim 400 \mu\text{s}$ . Timings of the LIBS system are triggered by a delay generator (DG535), which is master triggered by the output of a light guard inside the optomechanical chopper.

Recorded with system 1, one spectrum consists of 10 accumulations, 20 spectra per sample; giving overall 200 laser pulses per sample. A similar measurement routine is employed with system 2, i.e. 20 accumulations and 10 spectra per a sample, also giving 200 pulses per a sample. The 200 pulses per sample were chosen for satisfying statistics and smoothing of effects caused by sample inhomogeneity.

Lines of copper as well as the rock forming matrix elements (Al, Ca, Na, Cu, and Si; listed in Table 2) were extracted from the spectra for further analysis. The custom software for spectral processing was written in MATLAB (version R2012a). On each obtained spectrum, a Z-test was carried out to remove possible outliers after which the spectra were averaged. The Z-test was utilized to ensure that no outlier spectra affected by local sample inhomogeneity are used in further data analysis. The averaged spectra were normalized to their integral intensities and mean-centered. The classification by PCA and multivariate regression by PLSR and PCR were performed on the preprocessed data using the customized LIBRA software (KU Leuven, Belgium) [21].

## 3. Multivariate analysis

PCA is commonly used for classification of LIBS data [12,14,18,22]. PCA is a mathematical coordinate transformation method that

**Table 1**  
list of samples.

Sample No.	Rock type	Alteration	Cu (wt.%)	
ANS1	Andesite	ANS	Phyllic	0.64
ANS2	Andesite	ANS	Phyllic	0.49
ANS3	Andesite	ANS	Phyllic	0.60
ANS4	Andesite	ANS	Phyllic	0.61
ANS5	Andesite	ANS	Phyllic	0.79
ANS6	Andesite	ANS	Phyllic	0.21
ANS7	Andesite	ANS	Phyllic	0.52
DIO1	Diorite	DIO	Potassic	0.40
DIO2	Diorite	DIO	Potassic	0.71
DIO3	Diorite	DIO	Potassic	0.73
DIO4	Diorite	DIO	Potassic	0.78
DIO5	Diorite	DIO	Potassic	0.32
DIO6	Diorite	DIO	Potassic	0.31
DIO7	Diorite	DIO	Potassic	0.30
DIO8	Diorite	DIO	Potassic	0.37
DIO9	Diorite	DIO	Potassic	0.73
KP1	Monzonite-quartz monzonite	KP	Potassic	0.24
KP2	Monzonite-quartz monzonite	KP	Potassic	0.18
KP3	Monzonite-quartz monzonite	KP	Potassic	0.21
KP4	Monzonite-quartz monzonite	KP	Phyllic	0.06
KP5	Monzonite-quartz monzonite	KP	Phyllic	0.51
KP6	Monzonite-quartz monzonite	KP	Phyllic	0.34
KP7	Monzonite-quartz monzonite	KP	Phyllic	0.45
KP8	Monzonite-quartz monzonite	KP	Potassic	0.48
KP9	Monzonite-quartz monzonite	KP	Potassic	0.71
KP10	Monzonite-quartz monzonite	KP	Potassic	0.56
KP11	Monzonite-quartz monzonite	KP	Potassic	0.30

**Table 2**  
List of elemental lines selected for analysis.

Elemental line	$\lambda$ (nm)	$E_i$ (eV)	$E_j$ (eV)	Einstein coefficient ( $10^8 \text{ s}^{-1}$ )
Al (I)	308.22	0	4.02	0.63
Al (I)	309.27	0.014	4.02	0.75
Al (I)	394.4	0	3.14	0.51
Al (I)	396.12	0.014	3.14	1.01
Ca (II)	393.37	0	3.15	1.47
Ca (II)	396.85	0	3.12	1.44
Ca (I)	422.7	0	2.93	2.18
Ca (I)	558.87	2.53	4.74	0.49
Na (I)	589	0	2.104	0.62
Na (I)	589.59	0	2.102	0.61
Si (I)*	251.43	0	4.93	0.74
Si (I)*	251.61	0.03	4.95	1.68
Si (I)*	251.92	0.01	4.93	0.55
Si (I)*	288.16	0.78	5.08	2.17
Si (I)	390.55	1.91	5.08	0.13
Cu (I)	324.75	0	3.82	1.37
Cu (I)	327.39	0	3.79	1.36
Cu (I)	515.32	3.79	6.19	0.6
Cu (I)	521.82	3.82	6.19	1.22

\* Elemental lines used only when data from system 2 were utilized for multivariate analysis and linearization.

emphasizes variance in the data set and reduces redundancy [23–25]. A data matrix  $X$  consists of individual spectra arranged in rows, with the columns denoting the wavelengths/variables. The PCA algorithm decomposes the data matrix into matrices of scores and loadings. The scores are analogous to coordinates of the data points (sample spectra) in a newly constructed PC space and the loadings are the transformation coefficients that are indicative of variances at specific wavelengths. It is expected that samples of the same rock type will form one cluster, based on the similarity among their spectra. Since the spectral resemblance originates in similar elemental composition, the distribution of points in the PC space should correlate with the distribution in a QAPF diagram.

Principal components are conventionally used for multivariate quantification in a principal component regression (PCR) algorithm [25]. A PCA algorithm is applied to the data matrix  $X$  to truncate redundant information and to form a new data matrix. Then the regression vector  $y$  representing the content of an analyte in the samples is regressed against the PC scores. In an alternative approach, a PCA is used to model both the data and the regression vector (or matrix when more variables are considered) simultaneously; such an algorithm is called the partial least squares regression (PLSR) [25].

#### 4. Results and discussion

No reliable calibration curve could be constructed from the complete data set (consisting of LIBS spectra of 27 igneous rocks) due to a weak overall correlation between intensities and concentrations, likely caused by matrix effects. Therefore, we attempted the construction of partial calibration curves for each individual rock type. First, all samples

were divided into three groups according to their class membership (ANS, DIO, and KP; see Table 1), as determined by a geologist. Second, PCA classification of the sample set based on the similarities in the composition of sample matrices (composition in the Al, Ca, Na, and Si) was utilized. Then, partial calibration curves for each individual cluster were constructed. The procedure was applied to the data obtained with both LIBS systems in order to assess the independence of measurements on a particular instrument.

As stated above, the presence of lines of matrix elements Al, Ca, K, Na, and Si in LIBS spectra can help with reliable classification. Those matrix elements are conventionally used in geochemistry to fully discriminate igneous rocks against rock types[5]. Fifteen spectral lines of these matrix elements were selected (see Table 1) for multivariate analysis. The experimental system 1 does not cover a spectral range of 200–300 nm, where several silicon lines are located; only the Si 390.55 nm line was used to account for Si concentration. It was therefore expected that system 2, covering both the UV and VIS spectral ranges, would yield the more reliable classification. No strong potassium lines (at 766 and 769 nm) could be detected with either system (Fig. 1).

##### 4.1. Calibration without PCA classification

The calibration curves for copper are constructed using the Cu(I) line at 521.8 nm for both LIBS systems data. This line is unlikely to get self-absorbed because of the low concentration range of Cu (<1%) in analyzed samples and because of its high lower energy level of  $E_i = 3.82$  eV. Prior to the analysis, spectra of each sample were first vector normalized to the spectra total intensities and subsequently averaged to obtain one spectrum per sample. Spectral lines selected for uni- and multivariate analysis were fitted with pseudo-Voigt profile and the area under such fitted line was adopted as the line intensity.

It has been proposed to use individual calibration curves for analysis of individual rock types. Fig. 2 shows the combined (for all rock types) and partial/individual (for each rock type) calibration plots. The linear regression performed on the whole data set (the combined plot) yields in a coefficient of determination  $R^2 = 0.49$ ; for measurements with system 1 and only  $R^2 = 0.23$  for measurements with system 2. These low  $R^2$  values clearly indicate high uncertainty and low reliability of the calibration.

On the contrary, the partial calibration plots (according to the rock type), by showing different slopes, reveal higher values for coefficients of determination, between 0.57 and 0.69 and between 0.68 and 0.87 for systems 1 and 2, correspondingly. Partial calibration curves for DIO and KP rock types result only in a moderate improvement of the coefficient of determination, i.e.  $R^2 \sim 0.6$ , for both LIBS systems. The coefficient of determination for ANS rock type is higher with system 2,  $R^2 = 0.87$ , than with system 1,  $R^2 = 0.57$ . Even though the partitioning of calibration curves in accordance with rock types led to an improved linear regression, the accuracy of analysis with such low  $R^2$  partial calibration curves is expected to be moderate. Also, it is known from linear

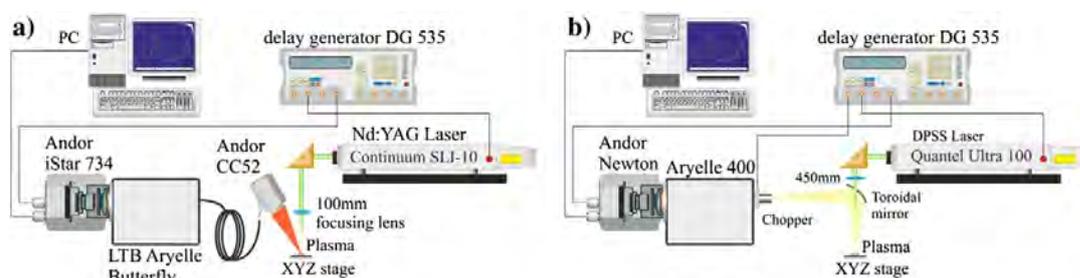


Fig. 1. Experimental LIBS: (a) setup 1 and (b) setup 2.

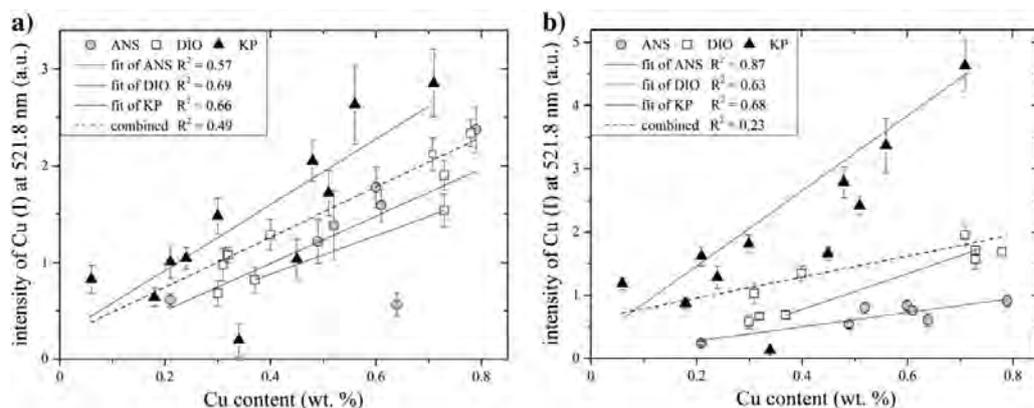


Fig. 2. Calibration curves for selected copper line 521.8 nm: (a) system 1 data set and (b) system 2 data set. The dashed line refers to calibration plot for complete sample set (assigned as “combined” in the legend of the plot).

regression that a reduction of data points usually leads to an increase in  $R^2$ . It is obvious from Fig. 2 that, even though assigned to the same rock type by a trained geologist, some points scatter strongly from the obtained linear behavior. Especially the samples KP6 and ANS1 (see Table 1 for assignment), which clearly appear misclassified in both experiments (Fig. 2a and b), respectively.

The preliminary results of the multivariate quantitative analysis of the whole data set using the partial least squares regression (PLSR) and the principal components regression (PCR) are shown in Fig. 3. The data matrix  $X$  was constructed from all copper lines plus selected lines of matrix elements, listed in Table 2. Again, spectra from each sample were first normalized to the sum of their total intensity and then averaged to obtain one spectrum per sample. The spectra were organized as rows and individual variables (lines of copper and selected matrix elements) were assigned to columns of the data matrix. The vector  $y$  for regression consists of certified copper contents (Table 1). The whole data set (all samples) was used simultaneously. Copper lines were not expected to suffer from self-absorption in the limited range of Cu concentrations in the samples (from 0.06 wt.% to 0.79 wt.%). Therefore, the variations of interest were assumed to be linear and were analyzed by linear multivariate regressions, PCR and PLSR [26]. In every step of the multivariate regression, all spectra of one sample were left out from the PCR and PLSR model estimation and then the content of the analyte (Cu) in the left-out sample was predicted using the model built from the remaining samples. The approach follows a

suggestion of Death *et al.* [27]. In their study, they proposed to introduce matrix lines into the computation of latent variables to enhance inter-group dissimilarities and improve the cost function.

PCR and PLSR regressions were applied on data sets obtained with both LIBS systems. To avoid under- or overfitting of the model, up to 10 PCs were gradually added to the model and related mean square errors of prediction (MSEP) were calculated. The optimal number of PCs used for each particular model is given in the box in Fig. 3. In general, a lower number of PCs is needed in PLSR caused by simultaneously modeling both the data matrix  $X$  and the regression vector  $y$ . A slightly lower number of PCs was needed for satisfactory regression models of system 2 data. This may be a consequence of more variables reflecting the Si content in the samples. In both cases, PLSR provided higher values for  $R^2$ , namely  $R^2 = 0.76$  for system 1 and  $R^2 = 0.79$  for system 2, respectively, compared to  $R^2 = 0.66$  and  $R^2 = 0.7$  obtained with PCR. Nevertheless, such low coefficients of determination can still not guarantee a reliable quantitative analysis. Thus, multivariate calibration using the whole data set has to be considered unsuccessful for such a wide variation in the composition of the sample set matrices as well as the narrow range of copper concentrations.

#### 4.2. Calibration after PCA classification

We attempted to classify the samples based on the similarities in their matrix compositions. It was expected that calibration curves

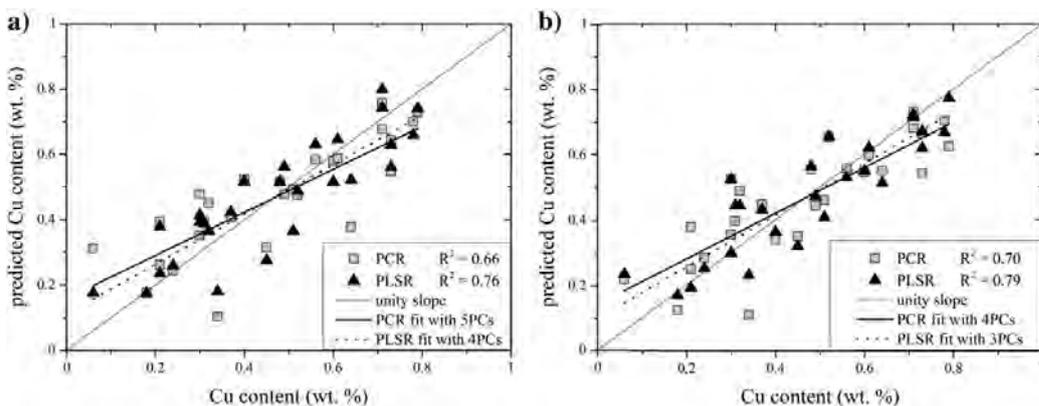


Fig. 3. Quantification of Cu content using PCR and PLSR for (a) system 1 data and (b) system 2 data.

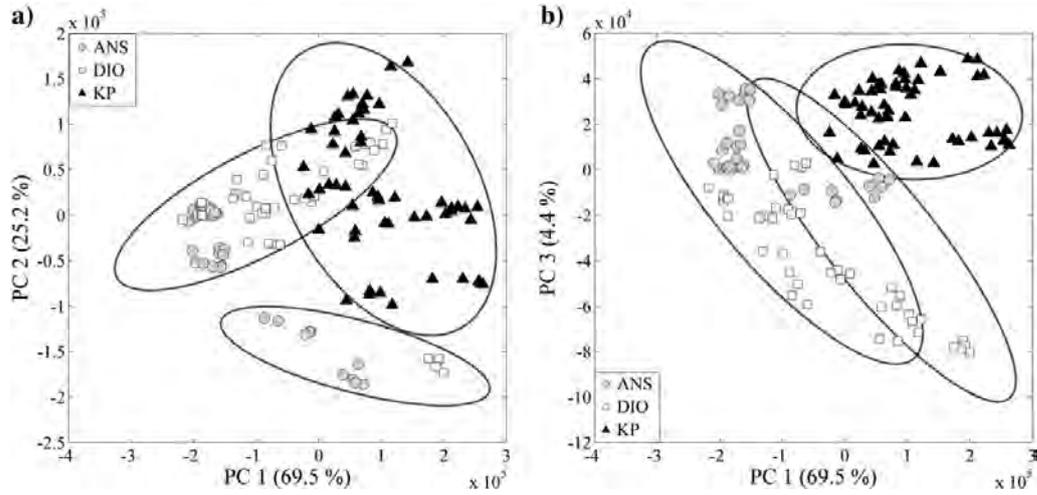


Fig. 4. PCA applied on the system 1 data: (a) first and second PC scores and (b) first and third PC scores.

constructed for each rock type individually would lead to improved linearity and related accuracy/bias of quantitative analysis. The data matrix was constructed from the lines of matrix elements listed in Table 2, and pretreated in the same manner as in the foregoing regression case. The spectra of each sample measured by system 1 were averaged to give 5 data points per sample. Respectively, the individual spectra of each sample from system 2 were also averaged to give 5 data points per sample. As stated before, the samples were classified by a geologist who inspected their phenomenological properties. This approach relies entirely on the experience and qualification of the geologist; in some difficult cases, it might yield erroneous results. It is, thus, highly desirable to have a more robust and unsupervised classification technique, for example, multielement PCA. The analysis can be utilized for matrix assignment or identification of misassignment.

The results of the PCA analysis are shown in Figs. 4 and 5 as the 2D projections of the data scores on the planes of the first three principal components taken pair-wisely. The first three PCs covered 97.1% (69.5% + 23.2% + 4.4%) of the total variance in the data obtained

with system 1 and 97.5% (60.1% + 32.9% + 4.5%) in the data obtained with system 2. The Gaussian clustering algorithm, built in the MATLAB Statistics Toolbox, was used to discriminate individual groups of samples according to their distributions in the PC space.

The clustering was done in the space spanning over the first three PCs obtained. The Gaussian clustering algorithm requires the number of modes to be set initially, those modes are simply  $n$ -dimensional multivariate Gaussian functions: if the  $k$ -th function is denoted as  $G_k$ ,  $n$  corresponds to the dimension of analysis space. The total number of modes was set in accordance with the number of groups given by a geologist. Furthermore,  $G_k$  is completely defined by the mean vector  $\bar{\mu}_k$  and covariance matrix  $\bar{\Sigma}_k$ . Initial values of  $\bar{\mu}_k$  and  $\bar{\Sigma}_k$  were randomly distributed in the analysis domain. The algorithm afterwards using maximal likelihood iteratively optimizes values of  $\bar{\mu}_k$  and  $\bar{\Sigma}_k$ , parameters of a multimodal Gaussian distribution that is applied to the analyzed set of points. To assure convergence, the optimization was performed 1000 times, each time with new random values of  $\bar{\mu}_k$  and  $\bar{\Sigma}_k$ , the distribution with the best likelihood was taken as a final answer. The ellipses in the

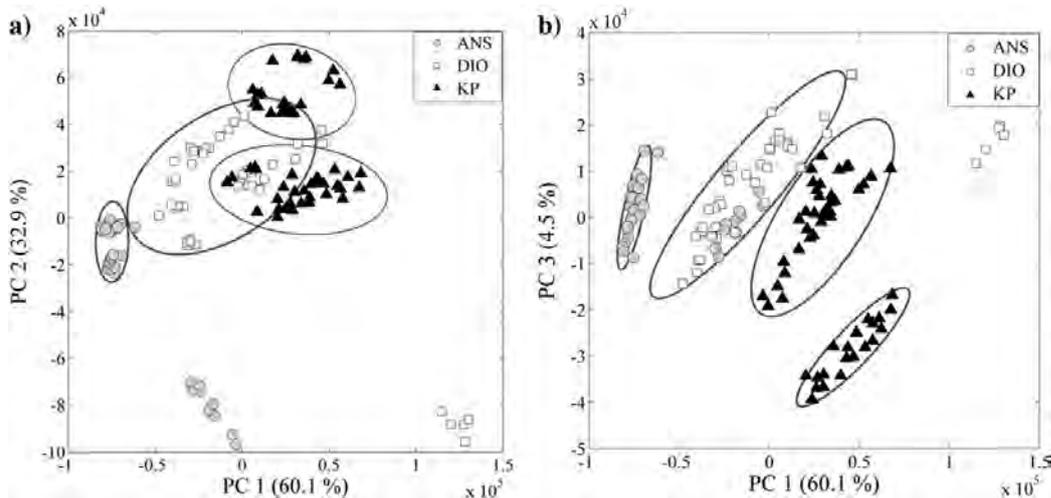


Fig. 5. PCA applied on the system 2 data: (a) first and second PC scores and (b) first and third PC scores.

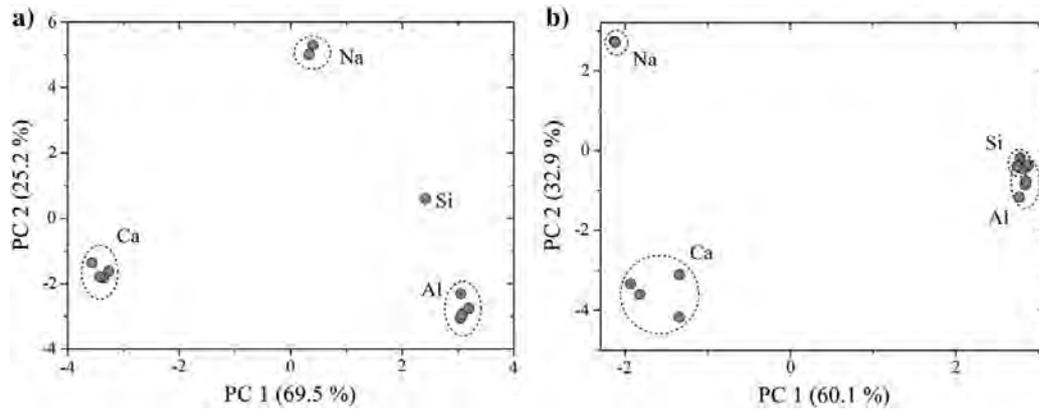


Fig. 6. First two PC loadings: (a) system 1 data and (b) system 2 data.

clustering plots denote the projection of the volume that covers 99% of the “mass” of the particular Gaussian mode,  $G_k$ , around its individual center of mass. Some clusters can overlap for samples with similar compositions and further analysis is needed. One should note that Gaussian clustering was used only for enclosing the points into separate groups. It is worth mentioning that data points do not have to necessarily obey a Gauss distribution in the PC space. Moreover, the number of PCs used for the clustering was determined experimentally. In this case study, we used only three principal components to classify the data set, while four or more principal components did not lead to any improvement to the clustering of data points. Note that the fourth principal component carries only ~1% of total variation in the samples’ composition measured using both systems.

When PCA is applied to the system 1 data set, three distinct outliers are observed (see Fig. 4). Those outliers (ANS1, ANS7, and DIO1) are enclosed within one cluster and may be discarded from further classification since the composition of their matrices does not correspond to any other cluster in the data set. We may assume that such samples were accidentally (wrongly) assigned by the geologist. Under this assumption, the rest of the data set is clustered into two distinct groups. The higher content of Si and Al in the KP rock type is responsible for its separation from the cluster of ANS and DIO rock types, as indicated

in the loadings plot in Fig. 6a. In this figure, loading values for each matrix element line of the first two principal components were cross-plotted. In this way, the figure depicts how each element (via its elemental lines) contributes to the discrimination of the samples using PCA. In a parallel analysis, PCA was applied to the system 2 data set and resulted in similar distribution of the data in newly constructed PC space, Fig. 5, as in the foregoing case. However, in the system 2 case, the outliers (ANS1, ANS7, and DIO1) are more distant from the rest of the data set. Except for several significantly outlying samples, rock types ANS, DIO and KP are distinctly separated. This is most probably the consequence of more variables (Si lines from the 200 to 300 nm spectral region) introduced to the computation of PCs. Furthermore, the KP rock type is separated into two clusters according to the alteration of the soil samples. Therefore, it can be concluded that one Si line (390.55 nm) may not be sufficient for the direct classification of samples in the system 1 data set. Nevertheless, the loadings plot in Fig. 6a reveals similar grouping of matrix lines of Al, Ca, and Si that are distributed along the first principal component as in the system 2’s case as depicted in Fig. 6b.

It is seen, while comparing the scores plots in Figs. 4 and 5 with the loadings plots in Fig. 6, that the data tends to cluster according to spectral lines of matrix elements. For instance, ANS rock type is on the left-

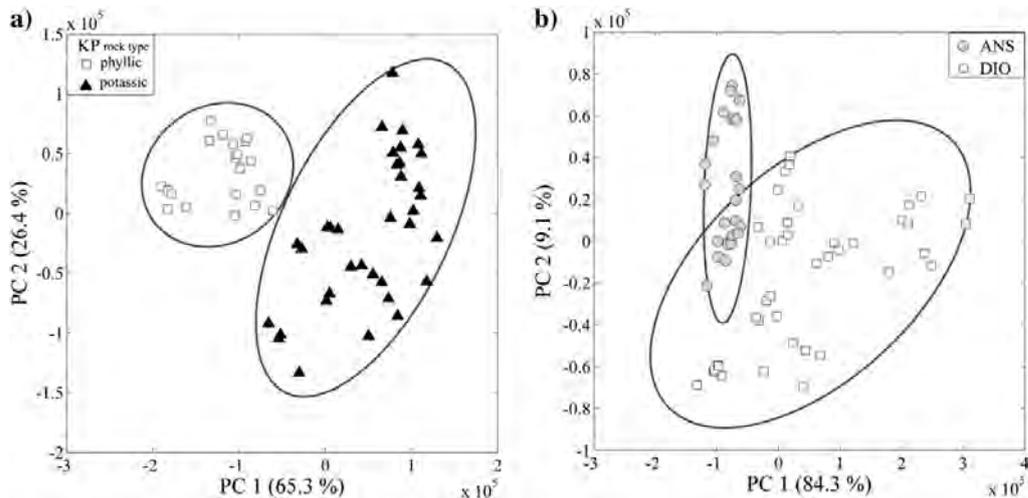


Fig. 7. Further classification of the system 1 data set: (a) classification of KP alterations and (b) discrimination of ANS and DIO.

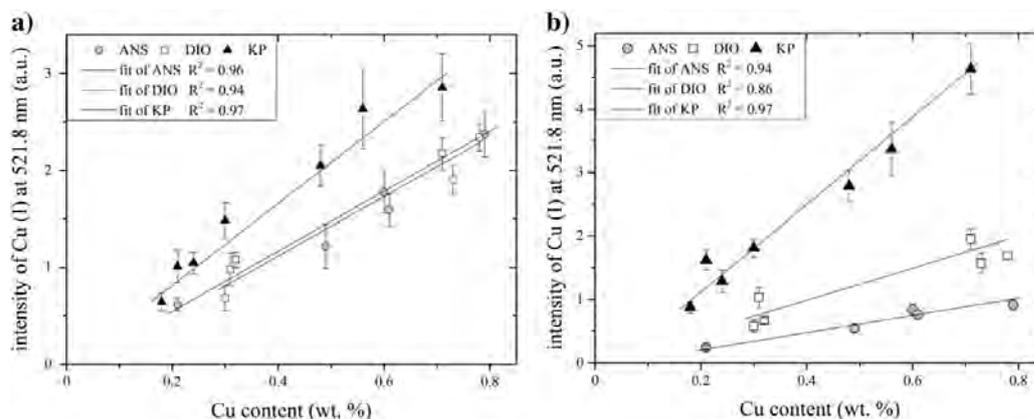


Fig. 8. Separate calibration curves formed after PCA analysis (a) for system 1 data set and (b) for system 2 data set.

hand side of the PC scores plots, which suggests that this rock type has higher content of Ca in its matrix compared to DIO and KP rock type. Spectral lines of matrix elements determine clustering patterns in subspaces of principal components and subsequent clustering of the data is possible. The main reason to use only the lines of matrix elements for classification is the reduction of data sets. The reduction of data matrices additionally accelerates PCA calculations that can be helpful for *online* analysis of data during the mining process. As suggested by Death *et al.* [27], data points of individual samples are satisfactorily distributed in more distinct groups in the PC space while only the important information is used, i.e. lines of matrix elements. Moreover, the PCA results for data sets from systems 1 and 2 are comparable. Those results support the reliability and reproducibility of measurements with different LIBS setups.

The data set of system 1 was not separated in distinct groups as that of system 2 (Fig. 4). Therefore, we additionally used an advanced approach for data classification suggested by Multari *et al.* [28]. They suggested discarding well-discriminated (tightly clustered) data points, i.e. samples/measurements, from the data matrix and then applying the PCA and consequent clustering algorithms again on the reduced data matrix. This step in the data analysis should emphasize the differences

between groups, which were formerly interfered. This procedure could be repeated until only well discriminated groups remain in the PC space. Based on the algorithm suggested by Multari *et al.* [28], Fig. 7a shows the clustering of the KP rock type, which was well discriminated from the rest of the data in previous classifications. Therefore, we applied PCA to this rock type solely. Data points were then clustered according to the relevant rock alteration. Similarly, the PCA was applied to the reduced data matrices of ANS and DIO rock types (when the outliers were discarded). Fig. 7b depicts two clusters that are distinguished from each other, i.e. enveloped with the according ellipses. The clusters are projected onto the first two PCs and are oriented randomly over each other in PC space. Furthermore, DIO rock type was treated with the PCA once more to reveal two more outliers (DIO8 and DIO9; not shown in Fig. 7.)

Thus, the PCA helps in the identification of outliers among the whole data set. The outliers (ANS1, ANS7, DIO1, DIO8, and DIO9) were discarded from further quantitative analysis. After this, the partial calibration curves were re-plotted and are shown in Fig. 8. According to the classification of rocks in geochemistry, the rock type and alteration should correspond to each other. Therefore, a calibration curve was constructed only for the potassic alteration of the KP rock type while the

Table 3  
Copper content calculated utilizing individual and combined calibration curves.

Sample	Cu content/wt.%, certified	System 1				System 2			
		Combined calibration	Bias	Partial calibration	Bias	Combined calibration	Bias	Partial calibration	Bias
ANS2	0.49	0.38 ± 0.09	−21.9%	0.45 ± 0.08	−8.1%	0.96 ± 0.11	96.1%	0.43 ± 0.04	−11.9%
ANS3	0.6	0.6 ± 0.08	−0.3%	0.63 ± 0.07	5.7%	0.06 ± 0.06	−90.0%	0.71 ± 0.08	19.0%
ANS4	0.61	0.52 ± 0.06	−14.3%	0.55 ± 0.05	−9.2%	0.01 ± 0.05	−98.1%	0.62 ± 0.06	2.4%
ANS5	0.79	0.84 ± 0.09	6.0%	0.84 ± 0.08	6.4%	0.15 ± 0.05	−81.0%	0.7 ± 0.06	−10.8%
ANS6	0.21	0.14 ± 0.03	−32.7%	0.23 ± 0.02	10.2%	0.5 ± 0.01	136.8%	0.19 ± 0.03	−9.9%
DIO2	0.71	1.13 ± 0.07	59.5%	0.86 ± 0.05	21.0%	0.9 ± 0.12	26.1%	0.92 ± 0.08	29.2%
DIO3	0.73	0.64 ± 0.06	−12.4%	0.68 ± 0.05	−7.2%	0.58 ± 0.11	−20.1%	0.65 ± 0.07	−11.2%
DIO4	0.78	0.82 ± 0.06	5.4%	0.73 ± 0.04	−6.3%	0.67 ± 0.05	−13.9%	0.69 ± 0.03	−11.1%
DIO5	0.32	0.33 ± 0.03	2.9%	0.34 ± 0.02	6.6%	0.65 ± 0.02	104.2%	0.31 ± 0.03	−2.0%
DIO6	0.31	0.29 ± 0.07	−7.0%	0.33 ± 0.05	6.6%	0.18 ± 0.12	−40.6%	0.37 ± 0.05	19.4%
DIO7	0.3	0.16 ± 0.05	−45.3%	0.26 ± 0.05	−14.5%	0.77 ± 0.06	157.9%	0.26 ± 0.04	−13.9%
KP1	0.24	0.32 ± 0.04	34.1%	0.23 ± 0.03	−4.4%	0.51 ± 0.05	110.5%	0.26 ± 0.01	6.4%
KP2	0.18	0.16 ± 0.04	−12.8%	0.17 ± 0.04	−7.4%	0.38 ± 0.08	112.8%	0.16 ± 0.02	−11.8%
KP3	0.21	0.31 ± 0.06	47.2%	0.23 ± 0.04	8.2%	0.31 ± 0.03	49.1%	0.23 ± 0.01	11.7%
KP8	0.48	0.71 ± 0.08	48.7%	0.47 ± 0.05	−1.8%	0.91 ± 0.11	89.3%	0.44 ± 0.04	−9.2%
KP9	0.71	1.07 ± 0.15	51.3%	0.77 ± 0.09	8.5%	1.15 ± 0.13	62.0%	0.77 ± 0.07	7.9%
KP10	0.56	0.96 ± 0.16	71.7%	0.61 ± 0.1	9.4%	1.14 ± 0.19	104.2%	0.53 ± 0.06	−6.2%
KP11	0.3	0.49 ± 0.07	63.8%	0.33 ± 0.04	9.8%	0.49 ± 0.07	64.1%	0.29 ± 0.02	−3.8%
		+ max	63.3%	+ max	21.0%	+ max	136.8	+ max	29.2
		− max	−45.3%	− max	14.5%	− max	−98.1	− max	−13.9%

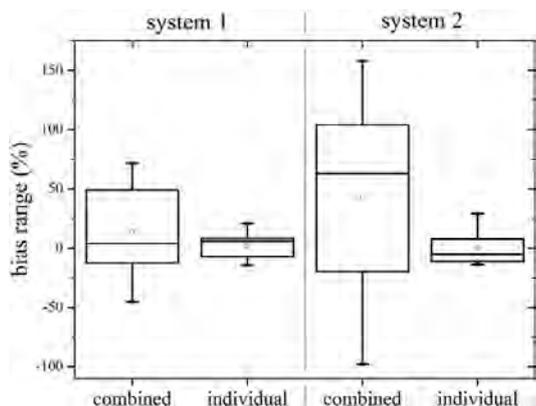


Fig. 9. The range bias depicted in box plots for the combined and partial calibration curves constructed for both systems.

phyllitic alteration was excluded from the calibration. The difference in the composition of phyllic and potassic alteration of the KP rock type was as well revealed in the PC space, as it is seen in Fig. 7a.

Significant improvements in the linearity of the partial calibration curves are obtained for all the rock types (ANS, DIO, and KP) as can be found in the comparison of Figs. 2 and 8. As a consequence of such data pretreatment, the matrix effect may be avoided to such an extent that it does not significantly affect the accuracy/bias of partial calibration curves. The precision of the prediction can be estimated from bias:

$$\text{bias (\%)} = \frac{\bar{x} - \mu_0}{\mu_0} \times 100,$$

where  $\bar{x}$  is the measured/predicted result and  $\mu_0$  is the reference value.

The improvement in prediction of copper concentrations using partial calibration curves was additionally proved by the leave-one-sample-out (LOSO) method described above. Table 3 provides the comparative results of LOSO using combined and partial calibration curves. The reference (certified) values of Cu concentrations were obtained from ICP-MS measurements. The bias was calculated for two methods (using combined and partial calibrations) and for both LIBS systems. As one sees from Table 3, the lower bias values, i.e. more accurate estimation of the copper content, are observed when partial calibration curves are used for the quantification. We used box plots to

schematically depict computed biases (Fig. 9) where the central solid line stands for the median of each data range and the box shows the central 50% of the data range. This figure visualizes what we expected, the biases computed from partial calibration curves are much lower compared to those computed from combined calibration curves. Based on these results, it is recommended to discriminate the data set based on the similarity in the sample matrices prior to the quantitative analysis. Fig. 10a and b show the correlation plots of the predicted versus certified concentrations for the combined (a) and partial (b) calibration curve methods. This visualization of the predicted copper contents strengthens the recommendation to use the partial calibration curves to obtain more reliable quantitative analysis, when the coefficient of determination for the measurement is  $R^2 = 0.955$  for system 1 and  $R^2 = 0.949$  for system 2, respectively.

Results from system 2 are less reliable/accurate, i.e. results are observed with higher biases, compared to system 1. This difference may arise from lower energy per pulse utilized to ablate material and to create plasma. However, in both cases, the results prove that partial calibration curves are more apt for reliable quantitative analysis. The biases of predicted copper contents were ranging from  $-14.5\%$  to  $21\%$  for system 1 and from  $-13.9\%$  to  $29.2\%$  for system 2, respectively. As seen from the Figs. 9 and 10, the results for systems 1 and 2 are similar when partial calibration curves are utilized. It may therefore be stated that improved quantitative analysis using partial calibration plots takes advantage of the discrimination of the samples based on dissimilarities in their matrices, i.e. in the composition of matrix elements (Al, Ca, Na, and Si). Moreover, in both LIBS systems, different lasers (high-energy Nd:YAG laser in system 1 and high-rate DPSS laser in system 2) were used as ablation sources as well as different collection optics (side view collector-collimator in system 1 and toroid mirror in system 2). Those results prove the reproducibility of the LIBS measurement on various systems, although similar spectrometers (LTB Aryelle with ICCD detector in system 1 and CCD detector with chopper in system 2) were utilized for the detection of plasma radiation.

## 5. Conclusion

Based on the presented study, LIBS can satisfy the needs of the mining industry when LIBS measurements are processed with PCA. In our study, two different LIBS systems were utilized to compare the compatibility of results obtained using partial calibration curves from suggested processes for data analysis. It was demonstrated that PCA can be employed for the classification of samples and the detection of outliers prior to the construction of partial calibration curves. Although reducing the overall number of samples, discarding identified outlier spectra is an

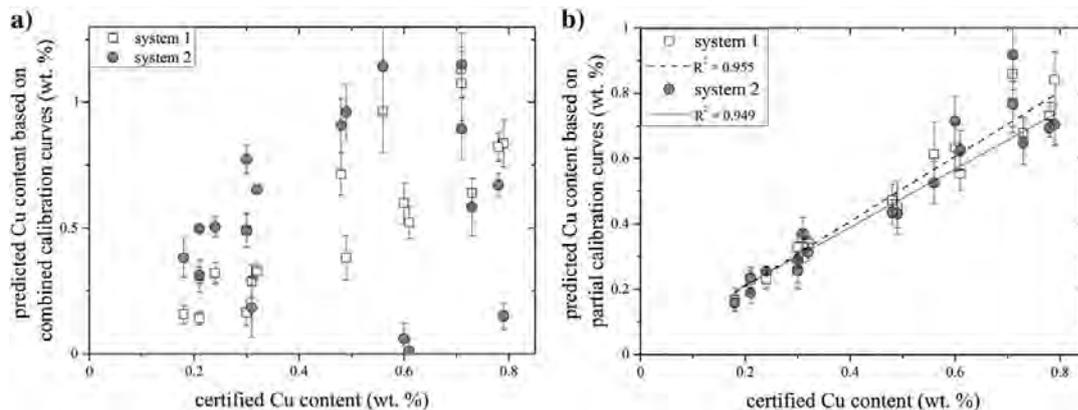


Fig. 10. Predicted Cu content for data of both LIBS systems from (a) combined calibration curves and (b) partial calibration curves.

approach that can be afforded in cases where enough sample material is available (as it is the case on mining sites). It is shown that discrimination of the samples into classes based on the differences in the content of their matrix elements (Al, Ca, Na, and Si) improves the quality of quantitative analysis. The bias in concentrations predicted from partial calibration curves (constructed for individual rock types) was lower than that obtained with the calibration curve constructed from the whole data set. However, further measurements should be conducted in order to create more robust data libraries. We strongly recommend discriminating the igneous rocks based on the similarities in their matrices, i.e. the composition of matrix elements Al, Ca, Na, and Si, prior to any further quantitative analysis. A reliable quantitative analysis may be provided only when the matrix effect affecting the LIBS measurement is suppressed.

#### Acknowledgment

P.P. gratefully acknowledges prior and ongoing support from the AiF ZIM and the support of Brno University of Technology on the frame of grant FSI-S-11-22 (Application of Advanced Optical Methods).

#### References

- [1] N.J. McMillan, R.S. Harmon, F.C. DeLucia, A.W. Miziolek, Laser-induced breakdown spectroscopy analysis of minerals: carbonate and silicates, *Spectrochim. Acta B* 62 (2007) 1528–1536.
- [2] L. Arenas, M. Ortega, M.J. García-Martínez, E. Querol, J.F. Llamas, Geochemical characterization of the mining district of Linares (Jaen, Spain) by means of XRF and ICP-AES, *J. Geochem. Explor.* 108 (2011) 21–26.
- [3] D.A. Cremers, L.J. Radziemski, *Handbook of Laser-Induced Breakdown Spectroscopy*, John Wiley & Sons, Ltd., New York, USA, 2006.
- [4] A.W. Miziolek, V. Palleschi, I. Schechter, *Laser-Induced Breakdown Spectroscopy (LIBS) Fundamentals and Applications*, Cambridge University Press, Cambridge, UK, 2006.
- [5] A. Streckeisen, To each plutonic rock its proper name, *Earth Sci. Rev.* 12 (1976) 1–33.
- [6] K. Novotný, J. Kaiser, M. Galiová, V. Konečná, J. Novotný, R. Malina, M. Liška, V. Kanický, V. Otruba, Mapping of different structures on large area of granite samples using laser-ablation based analytical techniques, an exploratory study, *Spectrochim. Acta B* 63 (2008) 1139–1144.
- [7] J. Lameyre, P. Bowen, Plutonic rock types series: discrimination of various granitoid and related rocks, *J. Volcanol. Geotherm. Res.* 14 (1982) 169–186.
- [8] R.S. Harmon, R.E. Russo, R.R. Hark, Applications of laser-induced breakdown spectroscopy for geochemical and environmental analysis: a comprehensive review, *Spectrochim. Acta B* 87 (2013) 11–26.
- [9] R.S. Harmon, F.C. DeLucia, C.E. McManus, N.J. McMillan, T.F. Jenkins, M.E. Walsh, A. W. 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.
- [10] R.S. Harmon, J. Remus, N.J. McMillan, C. McManus, L. Collins, J.L. Gottfried Jr., F.C. DeLucia, A.W. Miziolek, LIBS analysis of geomaterials: geochemical fingerprint for the rapid analysis and discrimination of minerals, *Appl. Geochem.* 24 (2009) 1125–1141.
- [11] R.T. Wainer, R.S. Harmon, A.W. Miziolek, K.L. McNesky, P.D. French, Analysis of environmental lead contamination: comparison of LIBS field and laboratory instruments, *Spectrochim. Acta B* 56 (2001) 777–793.
- [12] J.L. Gottfried, R.S. Harmon, F.C. DeLucia Jr., A.W. Miziolek, Multivariate analysis of laser-induced breakdown spectroscopy chemical signatures for geomaterial classification, *Spectrochim. Acta B* 64 (2009) 1009–1019.
- [13] R.S. Harmon, K.M. Shughue, J.J. Remus, M.A. Wise, L.J. East, R.R. Hark, Can the provenance of the conflict minerals columbite and tantalite be ascertained by laser-induced breakdown spectroscopy? *Anal. Bioanal. Chem.* 400 (2011) 3377–3382.
- [14] B. Bousquet, J.B. Sirven, L. Canioni, Towards quantitative laser-induced breakdown spectroscopy analysis of soil samples, *Spectrochim. Acta B* 62 (2007) 1582–1589.
- [15] G. Vítková, K. Novotný, L. Prokeš, A. Hrdlička, J. Kaiser, J. Novotný, R. Malina, D. Procházka, Fast identification of biominerals by means of stand-off laser-induced breakdown spectroscopy using linear discriminant analysis and artificial neural networks, *Spectrochim. Acta B* 73 (2012) 1–6.
- [16] C.A. Munson, F.C. DeLucia Jr., T. Piehler, K.L. McNesby, A.W. Miziolek, Investigation of statistic strategies for improving the discriminating power of laser-induced breakdown spectroscopy for chemical and biological warfare agent stimulants, *Spectrochim. Acta B* 60 (2005) 1217–1224.
- [17] M.Z. Martin, S. Allman, D.J. Brice, R.C. Martin, N.O. Andre, Exploring laser-induced breakdown spectroscopy for nuclear analysis and in-situ applications, *Spectrochim. Acta B* 74 (2012) 177–183.
- [18] M.Z. Martin, N. Labbé, T.G. Rials, S.D. Wullschlegler, Analysis of preservative-treated wood by multivariate analysis of laser-induced breakdown spectroscopy spectra, *Spectrochim. Acta B* 60 (2005) 1179–1185.
- [19] 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.
- [20] I.B. Gornushkin, A. Ruiz-Medina, J.M. Anzano, B.W. Smith, J.D. Winefordner, Identification of particulate minerals by correlation analysis using a microscope laser induced breakdown spectrometer, *J. Anal. At. Spectrom.* 15 (2000) 581–586.
- [21] S. Verboven, LIBRA: a MATLAB Library for Robust Analysis, 2 Aug 2012, [Online]. Available <http://wis.kuleuven.be/stat/robust/Programs/LIBRA> [Accessed 9 Oct 2013].
- [22] A.M. Ollila, J. Lasue, H.E. Newson, R.A. Multari, R.C. Wiens, S.M. Clegg, Comparison of two partial least squares-discriminant analysis algorithms for identifying geological samples with the ChemCam laser-induced breakdown spectroscopy instrument, *Appl. Opt.* 51 (2012) 130–142.
- [23] R.G. Brereton, *Chemometrics for Pattern Recognition*, John Wiley and Sons, Ltd., Bristol, UK, 2009.
- [24] L. Eriksson, E. Johansson, N. Kettaneh-Wold, J. Trygg, C. Wikström, S. Wold, *Multi- and megavariable data analysis, Part I: Basic Principles and Applications*, Umetrics AB, Umeå, Sweden, 2006.
- [25] H. Martens, T. Naes, *Multivariate Calibration*, John Wiley & Sons Ltd, Chichester, UK, 1989.
- [26] S. Laville, M. Sabsabi, F. Doucet, Multi-elemental analysis of solidified mineral melt samples by laser-induced breakdown spectroscopy coupled with a linear multivariate calibration, *Spectrochim. Acta B* 62 (2007) 1557–1566.
- [27] D. Death, A.P. Cunningham, L.J. Pollard, Multi-element and mineralogical analysis of mineral ores using laser induced breakdown spectroscopy and chemometric analysis, *Spectrochim. Acta B* 64 (2009) 1048–1058.
- [28] R.A. Multari, D.A. Cremers, J.M. Dupre, J.E. Gustafson, The use of laser-induced breakdown spectroscopy for distinguishing between bacterial pathogen species and strains, *Appl. Spectrosc.* 64 (2010) 750–759.

## Article 7

D. Prochazka, M. Bilík, P. Prochazková, J. Klus, P. Pořízka, J. Novotný, K. Novotný, B. Ticová, A. Bradáč, M. Semela, and J. Kaiser. Detection of tire tread particles using laser-induced breakdown spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 108:1–7, 2015.

DOI: 10.1016/j.sab.2015.03.011; 12 citations

Our research focus is also on the transfer of LIBS to industry and different applications, developing adaptations of LIBS systems and methodological protocols. One of such applications is utilization of LIBS in forensics.

This publication brought a feasibility study of detection of visually unrecognizable braking tracks. The idea behind the detection is in the fact that ABS (active braking system) improves the braking process of a vehicle by avoiding sliding on the road and thus there is no visible braking track. In spite of that, hard braking leaves traces of tyre particles on the road. They can be observed *via* the detection of zinc which is added to the tyre composition during the production process.

This is a completely novel application in the LIBS community and successful feasibility study led to a joint grant application and development of a unique robot LIBS system.



Contents lists available at ScienceDirect

## Spectrochimica Acta Part B

journal homepage: [www.elsevier.com/locate/sab](http://www.elsevier.com/locate/sab)

## Detection of tire tread particles using laser-induced breakdown spectroscopy<sup>☆</sup>



David Prochazka<sup>a,d,\*</sup>, Martin Bilík<sup>b</sup>, Petra Prochazková<sup>c</sup>, Jakub Klus<sup>d</sup>, Pavel Pořízka<sup>d</sup>, Jan Novotný<sup>d</sup>, Karel Novotný<sup>c,d</sup>, Barbora Ticová<sup>c</sup>, Albert Bradáč<sup>b</sup>, Marek Semela<sup>b</sup>, Jozef Kaiser<sup>a,d</sup>

<sup>a</sup> Brno University of Technology, Institute of Physical Engineering, Technická 2, 616 00 Brno, Czech Republic

<sup>b</sup> Brno University of Technology, Institute of Forensic Engineering, Údolní 244/53, 602 00 Brno, Czech Republic

<sup>c</sup> Masaryk University, Faculty of Science, Department of Chemistry, Kamenice 735/5, 625 00 Brno, Czech Republic

<sup>d</sup> Brno University of Technology, Central European Institute of Technology, Technická 3058/10, CZ-616 00 Brno, Czech Republic

### ARTICLE INFO

#### Article history:

Received 26 June 2014

Accepted 12 March 2015

Available online 25 March 2015

#### Keywords:

LIBS

Tire tread

Optimization

### ABSTRACT

The objective of this paper is a study of the potential of laser induced breakdown spectroscopy (LIBS) for detection of tire tread particles. Tire tread particles may represent pollutants; simultaneously, it is potentially possible to exploit detection of tire tread particles for identification of optically imperceptible braking tracks at locations of road accidents. The paper describes the general composition of tire treads and selection of an element suitable for detection using the LIBS method. Subsequently, the applicable spectral line is selected considering interferences with lines of elements that might be present together with the detected particles, and optimization of measurement parameters such as incident laser energy, gate delay and gate width is performed. In order to eliminate the matrix effect, measurements were performed using 4 types of tires manufactured by 3 different producers. An adhesive tape was used as a sample carrier. The most suitable adhesive tape was selected from 5 commonly available tapes, on the basis of their respective LIBS spectra. Calibration standards, i.e. an adhesive tape with different area content of tire tread particles, were prepared for the selected tire. A calibration line was created on the basis of the aforementioned calibration standards. The linear section of this line was used for determination of the detection limit value applicable to the selected tire. Considering the insignificant influence of matrix of various types of tires, it is possible to make a simple recalculation of the detection limit value on the basis of zinc content in a specific tire.

© 2015 Elsevier B.V. All rights reserved.

### 1. Introduction

The ability to detect tire tread particles in real time and in-situ has importance to a number of real world applications. Wide attention was recently dedicated to a tire tread particles as a source of pollution in environment. It was shown that tire treads contain heavy metals such as Mn, Fe, Co, Ni, Cu, Zn, Cd and Pb [1]. Concurrently with the increase of road traffic, the level of environmental pollution increases not only due to exhaust from combustion engines but also in connection with tire tread particles. Tire tread particles are released mainly in connection with vehicle speed changes; to a lesser extent, however, they are released even in the course of continuous driving. The respective particles consequently pollute air, soil and, subsequently, water sources [1–4].

Another possibility of exploitation of fast detection of tire tread particles might pertain to detection of optically imperceptible braking tracks. This thesis is based on the prerequisite that intense braking results in a higher level of abrasion of a tire tread than in the case of standard driving; therefore the average concentration of tire tread particles on a road surface would be statistically higher at places where braking was performed – in comparison with places where braking was not performed. Introduction of modern braking assistance systems (Anti-lock Braking System, Electronic Stability Control) results in low levels of tire sliding during braking, and therefore braking tracks prove to be visually very badly identifiable. This fact significantly impedes analyses of accidents so it is essential to seek methods and procedures of identification of even such braking tracks.

At present there are two different approaches to detection of tire tread particles. One of them is their detection on the basis of identification of polymers. For example, Gueissaz and Massonnet used pyrolysis in combination with gas chromatography and mass spectrometry (Py-GC/MS) [5] for detection of tire tread particles. Another example of a method used for detection of tire tread particles on the basis of typical polymers is, for example, infrared spectrometry (IR) [6].

<sup>☆</sup> Selected paper from the European Symposium on Atomic Spectrometry ESAS 2014 & 15th Czech-Slovak Spectroscopic Conference, Prague, Czech Republic, 16–21 March 2014.

\* Corresponding author at: Brno University of Technology, Institute of Physical Engineering, Technická 2, 616 00 Brno, Czech Republic. Tel.: +420 5 4114 2828.

E-mail address: [prochazka.d@fme.vutbr.cz](mailto:prochazka.d@fme.vutbr.cz) (D. Prochazka).

Another approach exploits detection of extractable organic zinc, which is added to tires in the form of ZnO<sub>2</sub> as a vulcanizing agent. This approach was selected by numerous investigators. On the basis of preparation of a sample and methods such as AAS or ICP-OES it is possible to detect extractable organic zinc, which is present solely in tire treads. An overview of methods focused on detection of tire tread particles including the respective references is listed in Table 1.

This paper suggests detection of tire tread particles using the LIBS method, on the basis of zinc identification. The fundamental benefits of this method include its speed and zero requirements as regards preparation of samples. Simultaneously, it is possible to perform an in-situ analysis at a distance of several meters. For an overview of properties of the LIBS method – in comparison with methods mentioned in Table 1 and some other – see [8,9].

The following sections describe the process of optimization of measurement parameters such as incident laser energy, gate width and gate delay with the objective of obtaining the maximum possible signal to noise ratio applicable to the selected zinc spectral line. Furthermore, the detection limit of the selected zinc line is determined for optimum parameters using the created calibration standards. ICP-OES resp. ICP-MS was selected as the reference method used for determination of zinc content in the selected tires.

## 2. Experimental section

### 2.1. Physical and instrumental background of the laser-induced breakdown spectroscopy (LIBS) method

LIBS is a quasi non-destructive method based on the principles of atomic emission spectrometry. The scope of fundamental instrumentation for LIBS comprises a pulsed laser, a lens focusing the laser pulse on the sample surface, optics collecting plasma radiation, a wavelength analyzer and a detector. Detailed description of LIBS principle, methodology and instrumentation is plentiful in the literature [8–12]. Here we only summarize the basic principle.

A short laser pulse (from dozens of fs to units of ns) featuring a major irradiance (several GW·cm<sup>-2</sup>) focused on the sample surface heats up, melts, atomizes and ionizes a small amount of the sample. This complex process as a whole leads to creation of radiating micro-plasma featuring a high temperature. Even though ablation of the material continues solely for several nanoseconds, the created plasma is specific for its high electron density (10<sup>17</sup>–10<sup>19</sup> cm<sup>-3</sup>), high temperature (6000–20,000 K) and rate of expansion of approx. 10<sup>6</sup> cm·s<sup>-1</sup>. Radiation of the micro-plasma may be subdivided into two parts – radiation originating from atomic emissions, and thermic and recombination continuum. The first component of the radiation, i.e. emission radiation, is essential for analyses of samples using the LIBS method as it contains chemical “fingerprints” of each element in the vaporized volume of the sample. Therefore, scanning of the plasma is important following elapse of hundreds on nanoseconds from the laser beam shot. Following elapse of this time we can observe sharp emission lines caused by transitions of electrons from higher energy levels to lower ones while continuous emissions caused by braking emissions and recombination

already prove to be negligible. Plasma radiation is, using reception optics and a fiber optic cable, transported to the input of the spectrometer. The obtained spectrum is recorded using a detector (e.g. CCD) and displayed using a computer.

One of the advantages of this method is the possibility of creation of chemical or element maps. The term of “chemical or element map” refers to the spatial arrangement and relative concentrations of individual elements on a heterogeneous sample [13,14]. LIBS is, considering its chemical mapping properties, frequently used and, in principle, it allows – following completion of a spatially differentiated measurement – for creation of two-dimensional chemical maps for all elements included in the sample. Spatial differentiation of this method is limited solely by the size of the ablation crater featuring the size of dozens of μm.

### 2.2. Applied instrumentation

Fig. 1 shows the experimental LIBS setup available at the Brno University of Technology. The whole setup consists of dual pulse Nd:YAG lasers, an interaction chamber equipped with two reception optic components and a camera used for sampleview, and two spectrometers. However, solely the primary laser LQ 529a (SOLAR) shown on Fig. 1 no. 1 and spectrometer in the Czerny–Turner (Lot Oriol 260I) arrangement were used for the experiments described below. The laser operates on the frequency of 10 Hz and wavelength of 532 nm, with the pulse length of ~10 ns. The laser beam is led, using mirrors, into the interaction chamber perpendicularly to the sample surface and focused on the sample using a glass triplet (focal distance: 32 mm).

The sample was placed on a handling device inside the interaction chamber (Fig. 1 no. 7) and its position and location of analysis were checked using a preview CCD camera (Fig. 1 no. 2). Plasma radiation was collected using a doublet of lenses ( $f = 90$  mm) and, using a fiber optic cable (diameter: 1000 μm), it was led to the entrance of the spectrometer (Lot Oriol 260I) in the Czerny–Turner arrangement (Fig. 1 no. 8). An iCCD camera (iStar 734i, ANDOR) was used as a detector.

The gate delay ( $t_d$ ) and gate width ( $t_w$ ) were controlled using a pulse generator (DG535, Stanford Research System, US) and special electronics developed in the laboratory of the Brno University of Technology. All the aforementioned devices prove to be computer-controlled. All measurements were performed in ambient atmosphere, with normal atmospheric pressure.

### 2.3. Samples

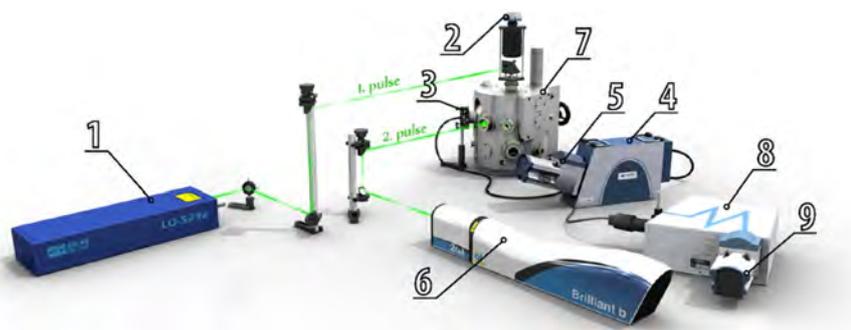
#### 2.3.1. Tire tread

Chemical compositions of tire treads prove to differ in the case of various manufacturers; however, in general it may be stated that it is based on natural or synthetic rubber, resp. a combination of them. During the process of vulcanization, other substances such as zinc oxide, sulfur, filling compounds, reinforcing agents, softeners, antioxidants and antiozonants are added into the rubber structure. Types and dosage of individual additives are trade secrets of individual tire manufacturers.

**Table 1**  
Some analytical methods for tire tread particle detection.

Method	Detection approach	Sample preparation	In-situ detection	Reference
AAS	Zn detection	HNO <sub>3</sub> to pH 2, sonicate, filter	No	[1]
XRF	Zn detection	No	Yes	[2]
EDX	Multielem. analysis	No	Yes	[3]
ICP-OES	Zn detection	HNO <sub>3</sub> digestion, microwave digestion	No	[4]
Py-GC/MS	Polymer detection	No	No	[5]
NA	Zn detection	No	No	[6]
IR	Polymer detection	No	No	[7]

AAS = atomic absorption spectrometry; XRF = X-ray fluorescence; ICP-OES = inductively coupled plasma optically emission spectrophotometry; GC-MS = gas chromatography–mass spectrometry; NA = neutron activation; IR = infrared spectrometry.



**Fig. 1.** Table top LIBS experimental setup, 1—Nd:YAG laser LQ-529, SOLAR, 2—CCD camera for sampleview, 3—Collecting optics CC52, Andor, 4—spectrometer in echelle arrangement, ME5000, Andor, 5—iCCD camera iStar 734i, Andor, 6—Nd:YAG laserBrilliant b, Quantel, 7—Interaction chamber, Tescan, 8—Spectrometer in Czerny–Turner arrangement, LOT Oriel 260I, 9—iCCD camera iStar 734i, Andor.

Various combinations and ratios of carbon particles or silica (silicon monoxide) are used as filling compounds.

In order to detect tire tread particles on the road, it was essential to determine a suitable element which complies with the following criteria:

- It is present in the tire in an amount detectable using the selected method;
- It is improbable for it to be naturally present on the road in a detectable amount.

Zinc was selected as the most suitable element. Zinc oxide is added into tire treads as an agent simplifying the process of vulcanization. Zinc content in tire proves to range, as regards various patents and expert analyses, from 0.04 to 2.5% by weight; however, the arithmetic mean and modulus of these values is approximately 1% by weight [4]. The overview of Zn contents in different tires obtained by different analytical methods published in literature is shown in Table 2. LIBS is generally capable of detection of heavy metals (for example Zn) under the level of  $1 \text{ mg} \cdot \text{kg}^{-1}$ .

To control the influence of matrix effect dependent on different tire compositions five different tires from different manufacturers were measured using LIBS and the zinc content was controlled using reference technique. The tire brands and types and zinc contents are listed in Table 3. As a reference technique to determine Zn content in tire tread was employed inductively coupled plasma mass spectrometry (ICP-MS) and inductively coupled plasma optical emission spectrometry (ICP-OES). Prior to the chemical analysis the tire tread particles were weighted on analytical weights (63.9–70.9 mg) and placed in plastic tubes together with 4 ml nitric acid (Sigma-Aldrich, puriss. p.a.,

$\geq 69\%$ ). The plastic tubes were exposed to microwave digestion for 25 min at the temperature of  $180^\circ \text{C}$ .

For the analysis of Matador Prima F, 185/60 R14 was employed ICP spectrometer 7700x Agilent ICP-MS with collision cell (He) for the remaining set of tires was employed ICP-OES spectrometer iCAP 6500 Duo Thermo Scientific.

Fig. 2 shows that the signal intensity depends on the zinc content in different tires linearly so it can be assumed that the matrix effect is neglectable. A Matador Prima F, 185/60 R 14 tire was selected for experiments described in this paper.

### 2.3.2. Adhesive tapes

A standard adhesive tape, resp. a dactyloscopic foil, is highly suitable for collection of tire tread particles. The use of an adhesive tape for handling or collecting the powder samples was described, for example, in [19–21] and the use of a dactyloscopic foil for collection of tire tread particles from the road surface was described, for example in [5].

In order to select the most suitable tape that is commonly available, several types of tapes were tested. The principal tape-related requirement had been that such a tape could not contain Zn or an element in the case of which the spectral line might interfere with selected Zn lines.

Four commonly available adhesive tapes and a dactyloscopic foil were tested and a transparent adhesive tape for office use was selected as the most suitable one, based on the above-mentioned requirements. For a comparison of the spectrum of a transparent adhesive tape and spectrum of the tire see Fig. 3.

### 2.3.3. Calibration standards

Samples containing the known weight of tire tread particles on a specifically defined area were prepared for determination of the theoretical limit of detection and quantification of surface content of tire tread particles.

Tire tread particles were made by the means of mechanical grinding of the tire tread using a steel file. The size of particles was measured using an optical microscope and ranged between  $10 \mu\text{m}$  and  $100 \mu\text{m}$  (Fig. 4). The particle size and morphology described in literature proofs differ depending on the particles origin from  $10 \mu\text{m}$  to  $300 \mu\text{m}$  [3,5]. Particles made in the above-mentioned manner were weighed and spread on an adhesive tape (dimensions:  $20 \times 15 \text{ mm}^2$ ). In order to preserve the same area of all samples, the adhesive tape was placed on a paper mask featuring exact dimensions.

**Table 2**

Zn content in different tire treads measured with different analytical methods from literature.

Zn Content ( $\text{mg} \cdot \text{kg}^{-1}$ )	Analytical method	References
8400	NA	[6]
8400	AAS	[11]
10,000	NA	[15]
10,000	NA	[16]
10,200	AAS	[17]
12,600	ICP-OES	[4]
15,500	ICP-OES	[18]

NA = neutron activation; ICP-OES = inductively coupled plasma optically emission spectrophotometry; AAS = atomic absorption spectrometry.

**Table 3**

Tire brands and types of tire treads used for experiments. In the table are presented zinc contents and LIBS emission signals (for LIBS signal definition see Section 2.4) including standard deviations. All samples were measured under same condition described below.

Tire brand and type	Zn content (mg·kg <sup>-1</sup> )	Standard deviation of Zn content (mg·kg <sup>-1</sup> )	LIBS emission signal of Zn I 330.25 nm spectral line (a.u.)	SD of LIBS emission signal (a.u.)
Matador Prima F, 185/60 R14	3500	24	5689	211
Michellin Primacy HP 215/55 R16	4954	28	7139	303
Goodyear EfficientGrip 205/55 R16	1249	26	1983	353
Matador MP 55 Plus 165/80 R13	11,949	39	23,163	411
Michellin Primacy HP 235/55 R17	7384	46	11,780	231

#### 2.4. Optimization of measurement parameters

As it was specified above, the used experimental setup is equipped with two spectrometers in different arrangements. One of the spectrometers features the echelle (Mechelle 5000, Andor) arrangement while the other one features the Czerny–Turner (2601, Lot Oriel) arrangement. The spectrometer featuring the echelle arrangement offers – in comparison with the spectrometer featuring the Czerny–Turner arrangement – a more extensive range of measured wavelengths (from 200 to 850 nm) but a lower sensitivity. Since the objective of the measurement was to obtain the lowest possible detection limits for a specific element, resp. a specific emission line of an element, the Mechelle 5000 spectrometer was used solely for primary measurements – for the purpose of determination of suitable emission lines. Additional optimizations and measurements were performed using the Lot Oriel 2601 spectrometer.

In order to be able to consider practical implementation of this method in the future, it is essential to strive for the simplest possible instrumentation and maintaining acquisition and operating costs on the lowest possible levels. Due to this reason, solely the single-pulse LIBS method (even though the device is equipped with two lasers) was used for the analysis.

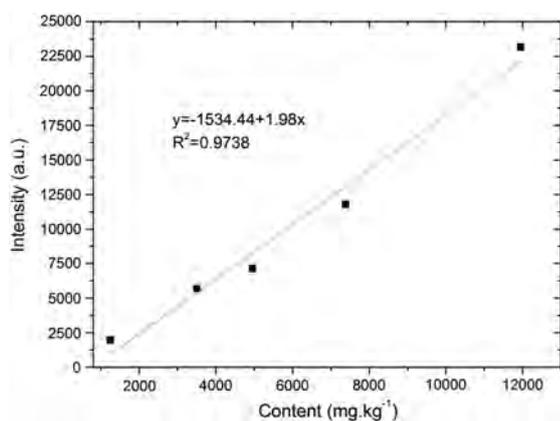
In the case of measuring LIBS spectra, the overall response of the detector represents the aggregate of the sample emission signal and continuous plasma radiation caused by recombination and bremsstrahlung (for the purpose of simplification, the detector noise is not considered). As the continuous radiation does not carry any information pertaining to the content of the sample, the effort is focused on its minimisation by the means of timing detector readings. As described in [12] and the respective list of references, the most important parameter describing sensitivity of the LIBS

method is the signal/noise (S/N) ratio and signal/background (S/B) ratio. In this paper, the element emission signal refers to any point on the selected emission line which correlates with its area and from the value of which the arithmetic mean of background measured in the close surrounding of the emission line is deducted. The maximum value of the emission line (Zn I 330.25 nm) was selected as the element emission signal and the surrounding of the emission line selected for calculation of the background and noise is marked in Fig. 5. Considering the fact that the sample was measured repeatedly for each setting, the emission signal represents the arithmetic mean of the maximum emission line value across all the measured spectra, for a specific setup. Background refers to the arithmetic mean of surrounding of the emission line across all the measured spectra, for a specific setup. Noise represents the standard deviation from the background average value across all the measured spectra, for a specific setup. Emission line Zn I 330.25 nm was selected in order to prevent possible interference with emission lines of adhesive tape (see Fig. 3). In order to prevent self-absorption the optical opacity of plasma was controlled using a mirror behind plasma as described in [22].

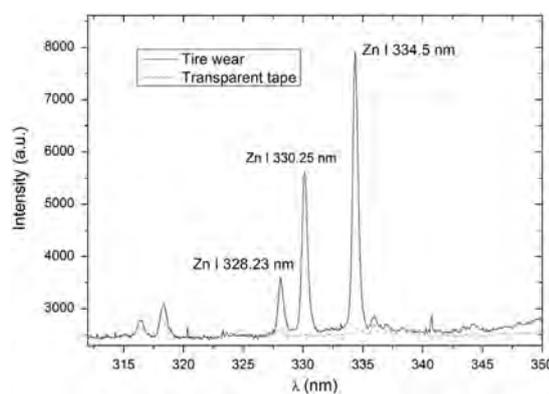
As many as 100 measurements were performed for each specific setup in order to achieve statistically significant, robust results.

The first optimized measurement parameter was laser energy pulse reaching the sample. The scope of measurement was selected in the range of 20–120 mJ per pulse. The gate delay and gate width were constant for all laser power settings and the respective values were selected considering the past experience with LIBS measurements ( $t_d = 1.5 \mu\text{s}$  and  $t_w = 10 \mu\text{s}$ ).

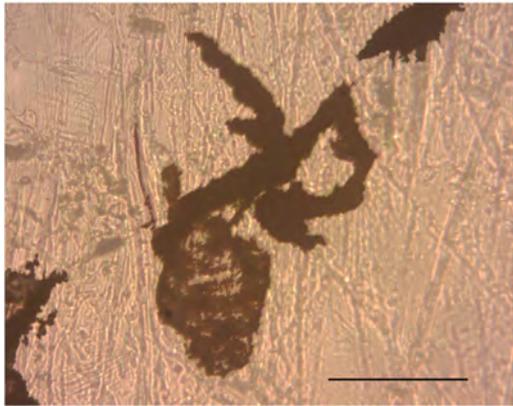
For the final charts depicting the signal/noise, resp. signal/background, ratio depending on the laser power see Fig. 6, resp. Fig. 7. Based on both the charts it may be deduced that the optimum laser power ranges between 80 and 100 mJ per pulse. The power of 90 mJ



**Fig. 2.** LIBS signal intensity of Zn I 330.25 nm vs zinc content in different tire types – see Table 3. Each point is average from 5 measurements. Errors are presented in Table 3. The red line in the graph is the linear fit. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 3.** Comparison of LIBS spectrums from tire tread and transparent adhesive tape. The measurement parameters were kept constant on both samples. Laser energy was 91 mJ per pulse, gate delay 1.6  $\mu\text{s}$ , gate width 9  $\mu\text{s}$ . It is clearly visible that transparent tape has much lower signal in the area of interest than tire tread and Zn respectively.

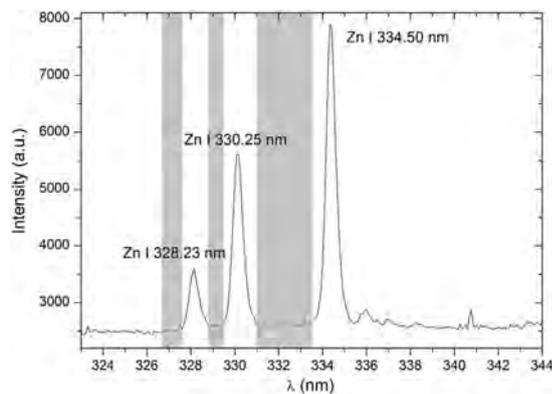


**Fig. 4.** Tire tread particles made by mechanical grinding on the adhesive tape. The size of bar is 100  $\mu\text{m}$ .

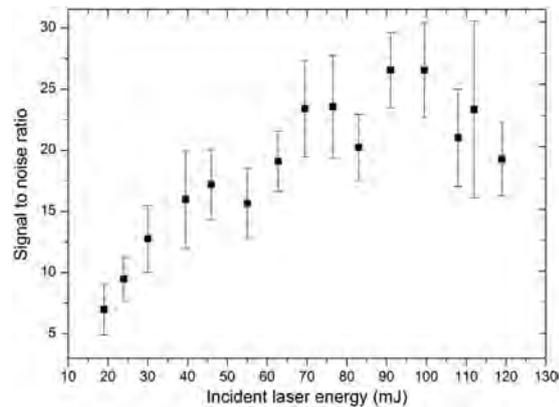
per pulse was selected for the measurement. For this power level, diameter of the ablation crater measured using an optical microscope was approximately 60  $\mu\text{m}$ .

Another optimized parameter was the gate delay. For the signal/noise, resp. signal/background, ratio pertaining to gate delay see Fig. 8, resp. Fig. 9. The time range and sampling rate were selected in manner that best mapped the course of dependence. The increment of 10 ns was selected in the range between 310 ns and 500 ns as accelerated cooling of plasma and decreasing of intensity of continuous radiation occur in this range, and emission lines become visible. Between 500 ns and 800 ns the increment was increased to 50 ns and from 800 ns to 7400 ns the increment was further increased to 200 ns. The last measurement was performed separately, for the value of 8000 ns. Charts in Figs. 8 and 9 clearly show that the optimum gate delay values range between 1000 ns and 2000 ns. The selected (optimum) time value was 1600 ns, which was used for all subsequent measurements.

The last optimized value was the time period during which plasma radiation is detected (gate width). For the signal/noise, resp. signal/background, ratio pertaining to gate width see Figs. 10 and 11. The increment of used values was 0.5  $\mu\text{s}$  in the range between 1  $\mu\text{s}$  and 15  $\mu\text{s}$ , and 1  $\mu\text{s}$  in the range between 15  $\mu\text{s}$  and 23  $\mu\text{s}$ . The charts show that the gate width influences the signal/noise ratio a hundred times less than gate delay. The difference between the minimum and



**Fig. 5.** Typical LIBS spectrum of tire tread, from spectrometer in Czerny–Turner arrangement. The used background range is marked with gray. The emission signal is the maximum value of selected emission spectral line intensity. Spectral line of Zn I 330.25 nm was used in this case.



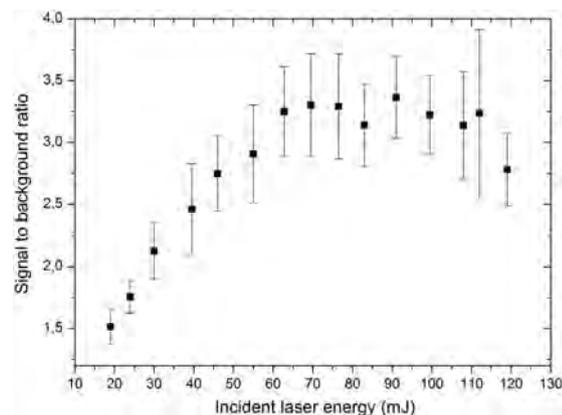
**Fig. 6.** Signal to noise ratio for selected spectral line Zn I 330.25 nm as a function of incident laser energy for tire tread. The parameters as gate delay and gate width were selected regarding to our experience with LIBS measurements and set as gate delay 1.5  $\mu\text{s}$  and gate width 10  $\mu\text{s}$ . Each point in the graph is averaged from (an average of) 100 measurements. The error bars represent standard deviation of signal to noise ratio.

maximum value of signal/noise ratio for gate delay was more than 100 times bigger while for gate width it was less than a double. This fact is caused by gradual extinction of plasma as the time elapses from the time of impact of the laser beam. After some time the increment caused by atomic emissions proves to be negligible and solely the level of noise increases. From charts in Figs. 10 and 11 it may be deduced that the respective time is in the range of 8–10  $\mu\text{s}$ . The time of 9  $\mu\text{s}$  was therefore selected for subsequent measurements.

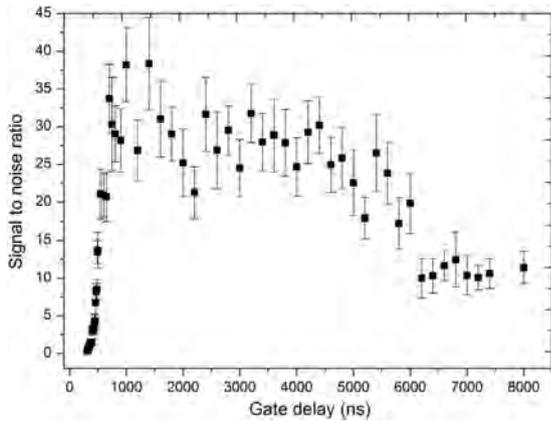
### 3. Results and discussion

#### 3.1. Calibration curve and theoretical detection limit value

The LIBS allows for determination of relative content of selected elements in various samples, resp. in various parts of a heterogeneous sample. In order to determine absolute content of the selected element, it is necessary to create a calibration curve using standards containing the known content of the selected element. This actually represents the



**Fig. 7.** Signal to background ratio vs incident laser energy for tire tread. The parameters as gate delay and gate width were selected regarding to our experience with LIBS measurements and set as gate delay 1.5  $\mu\text{s}$  and gate width 10  $\mu\text{s}$ . Each point in the graph is average from 100 measurements. The error bars represent standard deviation of signal to background ratio.



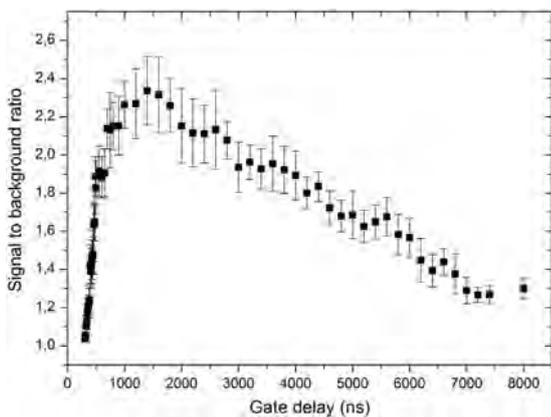
**Fig. 8.** Signal to noise ratio vs gate delay for tire tread and Zn I 330.25 nm emission spectral line respectively. The laser energy was set on 91 mJ and gate width on 10  $\mu$ s throughout all measurements. Each point in the graph is average from 100 measurements. The error bars represent standard deviation of signal to noise ratio.

ratio of intensity of the selected spectral line on content of the selected element in the sample.

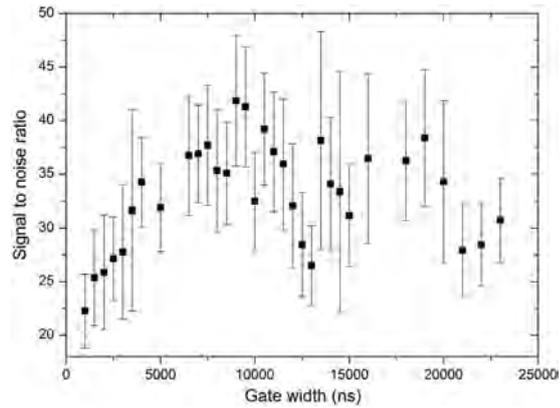
The aforementioned calibration standards were created for the purpose of determination of theoretical detection limits. Five samples were created for each area density and each sample was measured across its whole area of  $20 \times 15 \text{ mm}^2$  with an increment of 1 mm in both directions. 1 mm step was selected with regard to the microscope observation in order to avoid mutual influence of ablation craters. The Zn I average emission line signal (330.25 nm) was subsequently regarded as the value of intensity. For the ratio between the Zn I emission line intensity (330.25 nm) and area density of tire tread particles see Fig. 12.

The theoretical detection limit was calculated using the formula (2) described in [23] and recommended for calculation of the detection limit value in [12]

$$LOD = k \left( \frac{s_B}{\bar{X}_B} \right) \left( \frac{\bar{X}_B}{b} \right), \quad (2)$$



**Fig. 9.** Signal to background ratio vs gate delay for tire tread and Zn I 330.25 nm emission spectral line respectively. The laser energy was set on 91 mJ and gate width on 10  $\mu$ s throughout all measurements. Each point in the graph is average from 100 measurements. The error bars represent standard deviation of signal to background ratio.



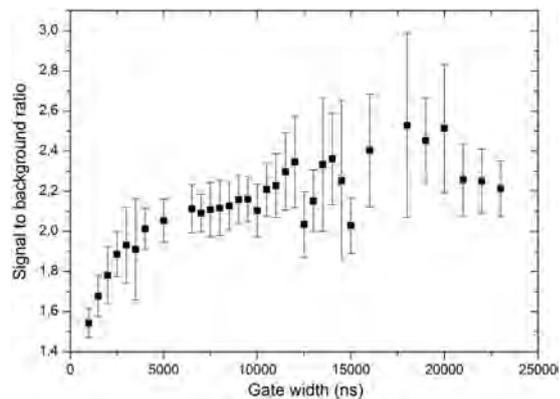
**Fig. 10.** Signal to noise ratio vs gate width for tire tread and Zn I 330.25 nm emission spectral line respectively. The laser energy was set on 91 mJ and gate delay on 1.6  $\mu$ s throughout all measurements. Each point in the graph is average from 100 measurements. The error bars represent standard deviation of signal to noise ratio.

whereas  $k = 3$ ,  $s_B$  represent noise of the background or blank,  $\bar{X}_B$  represents the average background intensity and  $b$  represents the slope of linear section of the calibration curve. The element in the first bracket describes the relative standard deviation (RSD) while the element in the second bracket describes the background element concentration (BEC). For linear section of the calibration curve see Fig. 12. The slope of this line is  $b = 2476.44$  while the blank noise value is  $s_B = 151.03$ . Thus, the detection limit value calculated using the aforementioned values is approximately  $0.18 \text{ mg} \cdot \text{cm}^{-2}$ .

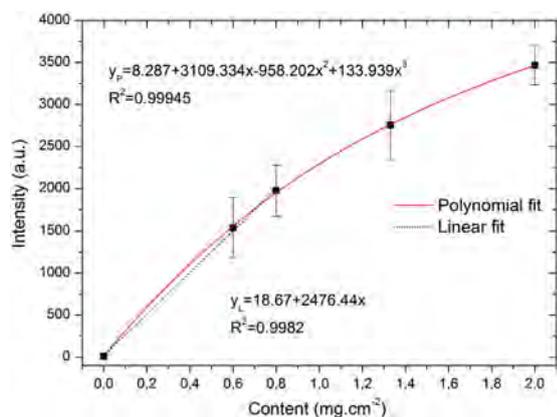
Since it has been shown that for different tire types and brands respectively the Zn I (330.25 nm) spectral line intensity changes linearly depending only on the zinc content (see Fig. 2). Thus, it is possible to determine the LOD for any tire only from the linear part of calibration curve, as it is shown in Fig. 12, and the zinc content in appropriate tire.

#### 4. Ongoing work

The ongoing work focuses on measurement on tire tread particles on the road surface – braking tracks. These measurements are based on a



**Fig. 11.** Signal to background ratio vs gate width for tire tread and Zn I 330.25 nm emission spectral line respectively. The laser energy was set on 91 mJ and gate delay on 1.6  $\mu$ s throughout all measurements. Each point in the graph is average from 100 measurements. The error bars represent standard deviation of signal to background ratio.



**Fig. 12.** LIBS signal intensity of Zn I 330.25 nm vs area content of tire tread particles on the transparent tape. The concentration was selected from 0 (blank), 0.6 mg/cm<sup>2</sup>, 0.8 mg/cm<sup>2</sup>, 1.4 mg/cm<sup>2</sup>, and 2 mg/cm<sup>2</sup>. Each point is average from 5 measurements and each measurement is composed of 300 single shots each from fresh point on area 20 × 15 mm<sup>2</sup> with equidistant step 1 mm in both directions.  $y_p$  refers to the polynomial fit and  $y_1$  to the linear fit. The error bars represent standard deviation of intensity.

unique mobile device allowing for performance of analyses outside a laboratory – directly at the location of a road traffic accident. It is legitimate to anticipate that the surface content of tire tread particles present directly on the road on the spots of braking tracks will be higher than on the rest of the road. Nevertheless, there are still a lot of challenges to overcome in further experimental work, for instance the estimation of the actual content of tire tread particles on the road surface for braking from different speeds, temporal degradation of braking tracks, etc.

This research is supposed to lead to the development of a device which is capable of automatic identification of braking tracks (transported to a laboratory or directly on a road surface) and, based on the respective data, determination of, for example, the initial speed and position of a vehicle.

#### Acknowledgment

We acknowledge the support of Brno University of Technology on the frame of grant FSI-S-14-2494 (Application of Advanced Optical Methods) and ÚSI-J-14-2181 (Modern methods for marks-detection on the road).

This work was also supported by the project “CEITEC – Central European Institute of Technology” (CZ.1.05/1.1.00/02.0068) from European Regional Development Fund.

#### References

- [1] P. Fauser, J.C. Tjell, H. Mosbaek, K. Pilegaard, Quantification of tire-tread particles using extractable organic zinc as tracer, *Rubber Chem. Technol.* 72 (1999) 969–977, <http://dx.doi.org/10.5254/1.3538846>.
- [2] L.M. Hildemann, Markowski Gregory, C. Glen, Chemical composition of emissions from urban sources of fine organic aerosol, *Environ. Sci.* 759 (1991) 744–759 (<http://pubs.acs.org/doi/abs/10.1021/es00016a021>) (accessed July 17, 2014).
- [3] K. Adachi, Y. Tainosho, Characterization of heavy metal particles embedded in tire dust, *Environ. Int.* 30 (2004) 1009–1017, <http://dx.doi.org/10.1016/j.envint.2004.04.004>.
- [4] T.B. Cuncell, K.U. Duckenfield, E.R. Landa, E. Callender, Tire-wear particles as a source of zinc to the environment, *Environ. Sci. Technol.* 38 (2004) 4206–4214 (<http://www.ncbi.nlm.nih.gov/pubmed/15352462>).
- [5] L. Gueissaz, G. Massonnet, Tire traces – discrimination and classification of pyrolysis-GC/MS profiles, *Forensic Sci. Int.* 230 (2013) 46–57, <http://dx.doi.org/10.1016/j.forsciint.2012.10.013>.
- [6] N. Fukuzaki, T. Yanaka, Y. Urushiyama, Effects of studded tires on roadside airborne dust pollution in Niigata, Japan, *Atmos. Environ.* 20 (1986) 377–386, [http://dx.doi.org/10.1016/0004-6981\(86\)90041-7](http://dx.doi.org/10.1016/0004-6981(86)90041-7).
- [7] D. Gross, Degradation of vulcanizates and their identification by IR spectrometry, *Rubber Chem. Technol.* 48 (1975) 289–300, <http://dx.doi.org/10.5254/1.3547454>.
- [8] F.J. Fortes, J.J. Laserna, The development of fieldable laser-induced breakdown spectrometer: no limits on the horizon, *Spectrochim. Acta B* 65 (2010) 975–990, <http://dx.doi.org/10.1016/j.sab.2010.11.009>.
- [9] D.A. Cremers, L.J. Radziemski, *Handbook of Laser-Induced Breakdown Spectroscopy*, 2nd ed. Wiley, 2013.
- [10] R. Noll, *Laser-Induced Breakdown Spectroscopy: Fundamentals and Applications*, Springer, 2012. (<http://www.amazon.com/Laser-Induced-Breakdown-Spectroscopy-Fundamentals-Applications/dp/3642206670>) (accessed December 16, 2013).
- [11] 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) 335–366, <http://dx.doi.org/10.1366/000370210793561691>.
- [12] 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, <http://dx.doi.org/10.1366/11-06574>.
- [13] V. Piñón, M.P. Mateo, G. Nicolas, Laser-induced breakdown spectroscopy for chemical mapping of materials, *Appl. Spectrosc. Rev.* 48 (2013) 357–383, <http://dx.doi.org/10.1080/05704928.2012.717569>.
- [14] J. Kaiser, K. Novotný, M.Z. Martin, A. Hrdlička, R. Malina, M. Hartl, et al., Trace elemental analysis by laser-induced breakdown spectroscopy – biological applications, *Surf. Sci. Rep.* 1–11 (2012) <http://dx.doi.org/10.1016/j.surfrep.2012.09.001>.
- [15] W.R. Pierson, W.W. Brachaczek, Airborne particulate debris from rubber tires, *Rubber Chem. Technol.* 47 (1974) 1275–1299, <http://dx.doi.org/10.5254/1.3540499>.
- [16] J.M. Ondov, A.S. Wexler, Where do particulate toxins reside? An improved paradigm for the structure and dynamics of the urban mid-Atlantic aerosol, *Environ. Sci. Technol.* 32 (1998) 2547–2555, <http://dx.doi.org/10.1021/es971067y>.
- [17] M. Legret, C. Pagotto, Evaluation of pollutant loadings in the runoff waters from a major rural highway, *Sci. Total Environ.* 235 (1999) 143–150, [http://dx.doi.org/10.1016/S0048-9697\(99\)00207-7](http://dx.doi.org/10.1016/S0048-9697(99)00207-7).
- [18] K.A. Handreck, Zinc toxicity from tire rubber in soilless potting media, *Commun. Soil Sci. Plant Anal.* 27 (1996) 2615–2623, <http://dx.doi.org/10.1080/00103629609369726>.
- [19] M. Pouzar, T. Kratochvíl, L. Capek, L. Smoláková, T. Cernohorský, A. Krejčová, et al., Quantitative LIBS analysis of vanadium in samples of hexagonal mesoporous silica catalysts, *Talanta* 83 (2011) 1659–1664, <http://dx.doi.org/10.1016/j.talanta.2010.11.047>.
- [20] M.B. Rosenberg, C.R. Dockery, Determining the lifetime of detectable amounts of gunshot residue on the hands of a shooter using laser-induced breakdown spectroscopy, *Appl. Spectrosc.* 62 (2008) 1238–1241, <http://dx.doi.org/10.1366/000370208786401473>.
- [21] S.I.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 (<http://as.osa.org/abstract.cfm?URI=as-56-4-433>) (accessed August 25, 2014).
- [22] G. Bekefi, *Principles of Laser Plasmas*, 1976.
- [23] E. Voigtman, Limits of detection and decision. Part 4, *Spectrochim. Acta B* 63 (2008) 154–165, <http://dx.doi.org/10.1016/j.sab.2007.11.014>.

## Article 10

J. Klus, P. Mikysek, D. Prochazka, P. Pořízka, P. Prochazková, J. Novotný, T. Trojek, K. Novotný, M. Slobodník, and J. Kaiser. Multivariate approach to the chemical mapping of uranium in sandstone-hosted uranium ores analyzed using double pulse Laser-Induced Breakdown Spectroscopy. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 123:143–149, 2016.

DOI: 10.1016/j.sab.2016.08.014; 17 citations

We stood at the beginning of the multivariate mapping. This manuscript was among the first in the LIBS community to discuss potential utilization of chemometrics to reveal hidden causalities in the data matrix obtained from the LIBS mapping of heterogeneous sample.

Principal component analysis (PCA) was used to visualize large data matrix, with 42 000 variables, on a lower dimensional scale. Individual PCs were then printed not in typical cross-plots but in the map, each score was attributed with a position in the map where it was obtained. In our consequent study, Kohonen's self organizing maps were used for similar purposes. Both algorithms provide advanced capabilities in visualization and clustering which can be used also in a non-conventional way.

The uranium bearing ore sample was mapped with LIBS. Spectral line response was correlated with the first principal component, revealing a simple fact that the variance of uranium content and its spectral fingerprint was so significant that it overshadowed other elements, such as calcium, potassium, zirconium, *et al.*



## Multivariate approach to the chemical mapping of uranium in sandstone-hosted uranium ores analyzed using double pulse Laser-Induced Breakdown Spectroscopy



Jakub Klus<sup>a,b,c,\*</sup>, Petr Mikysek<sup>d</sup>, David Prochazka<sup>a,c</sup>, Pavel Pořízka<sup>a,b,c</sup>, Petra Prochazková<sup>e</sup>, Jan Novotný<sup>a,b,c</sup>, Tomáš Trojek<sup>f</sup>, Karel Novotný<sup>c,e</sup>, Marek Slobodník<sup>d</sup>, Jozef Kaiser<sup>a,b,c</sup>

<sup>a</sup> Central European Institute of Technology, Brno University of Technology, Purkyňova 123, CZ-612 00 Brno, Czech Republic

<sup>b</sup> Faculty of Mechanical Engineering, Brno University of Technology, Technická 2896/2, 616 69 Brno, Czech Republic

<sup>c</sup> AtomTrace s.r.o., Kolejní 9, 612 00 Brno, Czech Republic

<sup>d</sup> Masaryk University, Faculty of Science, Institute of Geological Sciences, Kotlářská 2, 611 37 Brno, Czech Republic

<sup>e</sup> Masaryk University, Faculty of Science, Department of Chemistry, Kamenice 735/5, 625 00 Brno, Czech Republic

<sup>f</sup> Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering, Department of Dosimetry and Application of Ionizing Radiation, Břehová 7, 115 19 Praha 1, Czech Republic

### ARTICLE INFO

#### Article history:

Received 14 March 2016

Received in revised 10 August 2016

Accepted 10 August 2016

Available online 12 August 2016

#### Keywords:

Laser-Induced Breakdown Spectroscopy

Uranium, principal component analysis

Chemical mapping

Sandstone-hosted deposit

X-ray Fluorescence

### ABSTRACT

The goal of this work is to provide high resolution mapping of uranium in sandstone-hosted uranium ores using Laser-Induced Breakdown Spectroscopy (LIBS) technique. In order to obtain chemical image with highest possible spatial resolution, LIBS system in orthogonal double pulse (DP LIBS) arrangement was employed. Owing to this experimental arrangement the spot size of 50  $\mu\text{m}$  in diameter resulting in lateral resolution of 100  $\mu\text{m}$  was reached. Despite the increase in signal intensity in DP LIBS modification, the detection of uranium is challenging. The main cause is the high density of uranium spectral lines, which together with broadening of LIBS spectral lines overreaches the resolution of commonly used spectrometers. It results in increased overall background radiation with only few distinguishable uranium lines. Three different approaches in the LIBS data treatment for the uranium detection were utilized: i) spectral line intensity, ii) region of apparent background and iii) multivariate data analysis. By utilizing multivariate statistical methods, a specific specimen features (in our case uranium content) were revealed by processing complete spectral information obtained from broadband echelle spectrograph. Our results are in a good agreement with conventional approaches such as line fitting and show new possibilities of processing spectral data in mapping. As a reference technique to LIBS was employed X-ray Fluorescence (XRF). The XRF chemical images used in this paper have lower resolution (approximately 1–2 mm per image point), nevertheless the elemental distribution is apparent and corresponds to presented LIBS experiments.

© 2016 Elsevier B.V. All rights reserved.

### 1. Introduction

The global demand for energy is rising and the future generation will need a range of options if greenhouse gas emissions are to be reduced [1]. One of the basic source of energy comes from nuclear processes, especially uranium fission. Sufficiency of uranium reserves found on the Earth can contribute to energy sustainability.

Uranium accumulations are formed in various ways, the most important include magmatic, hydrothermal and sedimentary processes, which create a wide range of deposit types and formations. The widespread hydrothermal types are sandstone-hosted uranium deposits constituting approximately 18% of the world's known reserves. It is distributed in more than six hundred deposits [2,3] and contributes to

more than half of a current uranium production. Orebodies of this type are commonly low to medium grade (0.05–0.35% U) and the number of operating deposits (50) considerably exceeds other deposit types. Rather than conventional methods (underground and opencast mining), the chemical extraction (*in-situ* leaching) is chiefly used for economic valuation [4].

Czech Republic does not belong any more to the world's most valuable uranium provinces, but especially in European and regional scale it has considerable ore reserves. These are represented by peneconcordant sandstone-hosted uranium deposits in the North Bohemian Cretaceous Basin. These deposits are defined as U-Zr-P-Ti type and represent a remarkable elemental as well as mineral associations, that makes this area worldwide unique.

The study of ore mineralization is difficult for the size of mineral phases (in  $\mu\text{m}$ ), gel nature of components *etc.* and hence application of traditional geological instruments such as optical microscopy, electron microprobe, scanning electron microscopy or x-ray diffraction is often

\* Corresponding author at: Faculty of Mechanical Engineering, Brno University of Technology, Technická 2896/2, 616 69 Brno, Czech Republic.  
E-mail address: [jakub.klus@ceitec.vutbr.cz](mailto:jakub.klus@ceitec.vutbr.cz) (J. Klus).

**Table 1**  
Some analytical methods utilized for uranium detection.

Method	Stand-off detection	Chemical mapping	Reference
ICP-MS	No	No	[6]
ICP-AES	No	No	[6]
XRF	No	Yes	[7]
Raman spectroscopy	Yes	Yes	[8]
LIF	Yes	Yes	[9,10]
Gamma ray spectroscopy	Yes	No	[11]
LIBS	Yes	Yes	[12–16]

ICP-MS – Inductively Coupled Plasma Mass Spectrometry; ICP-AES – Inductively Coupled Plasma Atomic Emission Spectrometry; XRF – X-ray Fluorescence; LIF – Laser-Induced Fluorescence, LIBS – Laser-Induced Breakdown Spectroscopy.

limited due to the low-grade ore, particle size or colloidal nature of uranium and uranium-bearing phases [5]. These methods need extensive sample preparation prior the analysis. Some analytical methods typically used for uranium detection and their capabilities are listed in Table 1. However, for direct *in-situ* and/or *stand-off* detection, the aforementioned methods are inconvenient. Analytical method with great potential to satisfy above discussed demands is Laser-Induced Breakdown Spectroscopy (LIBS). Thus this paper aims to study the possibilities of LIBS concerning the detection of uranium precipitated in sandstones, and consequently selects and also suggests new approaches enabling fast and accurate chemical mapping of elements of interest.

The scope of advantages of LIBS in comparison to other methods routinely used for ore mineralization studies comprises mainly the relatively simple and robust equipment, possibility to detect broad spectrum of elements, no necessity of sample preparation, relatively high sensitivity and possibility of spatially resolved measurements [17–19]. Moreover LIBS is capable of *in-situ* and *stand-off* measurements [13,20]. Despite unquestionable advantages of LIBS, the detection of some elements is challenging due to various reasons. One of these difficult-to-detect elements is uranium. The main challenge in uranium detection is its complex emission spectrum with very high density of spectral lines. As is stated in ref. [12] the number of uranium lines measured using a hollow cathode discharge lamp exceeds 5000 in the range 384.8 to 908.4 nm. Considering the spectral line broadening such line density may cause interference of spectral lines. In this paper it is suggested to use this feature of uranium emission spectrum to determine the relative content of uranium in the studied specimen.

This approach enables obtaining and isolating optimal areas of the sample with ore-bearing material. Moreover, elemental mapping brings additional information as it gives us a basic idea about the distribution of ore-phases and their possible associations (based on the correlation of elements). Furthermore, it is an interesting mean of monitoring of the ore-forming solutions fluxes, which create metallic accumulation within the sandstones. This study is focused on multiple ways of data processing concerning monitoring of spatial distribution of uranium only.

## 2. Experimental

### 2.1. Samples and sample preparation

Břevniště deposit is one of the uranium deposits of sandstone type situated within geological structure called the Stráž block (in the area of northern part of the Bohemian Cretaceous Basin). Unique association of elements U-Zr-P-Ti characterizes the ore field and forms mostly fine-grained mineralization bonded to the Cenomanian sediments (especially sandstones and siltstones) [21]. These host rocks are composed mainly from quartz, clay matrix and accessory minerals (such as oxides, sulfides or carbonates). Mineralization formed in these rocks is a result of infiltration of ore-forming solutions into the open or clay matrix filled pores between quartz grains. Specific nature of this deposit consist in the occurrence of uraniferous hydrozircon and leucogenes *s. l.*

(alteration products and mixtures of Fe-Ti oxides) and partial colloidal to meta-colloidal nature of ore minerals [22].

Preparation of samples included several specific steps, which are necessary for application of different methods. Ten samples collected at the site were cut, dried, and cemented with Araldite epoxy. For XRF elemental mapping it was required to create a flat surface and circumference of the sample. This was achieved by grinding an excess of epoxy from the rock surface and cropping edges.

The final step comprised of the selection of most interesting sample and its region respectively for the consecutive LIBS analysis, based on the uranium content.

### 2.2. X-ray Fluorescence (XRF) analysis

Results described in this article were obtained with the XRF system designed and operated at the Czech Technical University in Prague. XRF system consists of an exchangeable X-ray source, a spectrometric X-ray detector (Si-PIN), and a positioning stage for sample investigation. The Si-PIN detector (Amptek Inc.) provided X-ray spectrometry in the energy range approximately from 2 keV up to 30 keV. System utilizes a beryllium window 25.4  $\mu\text{m}$  in thickness and with a sensitive volume of 6 mm<sup>2</sup>  $\times$  0.5 mm. Window distance from the measured sample was less than 1 cm. Sample stage is movable in both horizontal directions with a step down to 10  $\mu\text{m}$  and the maximum range of operation is approximately 50 cm.

This XRF analyzer was equipped with Au anode Mini-X X-ray tube that was operated at the voltage of 35 kV and the current of 0.1 mA. The X-ray beam was collimated with a narrow collimator and filtered with 1 mm of pure aluminum. The aperture diameter of the shipped collimator is 2 mm and the lateral resolution in scanning mode is 3 mm (FWHM determined with knife-edge scan). Resulting lateral resolution was achieved by adding copper collimators.

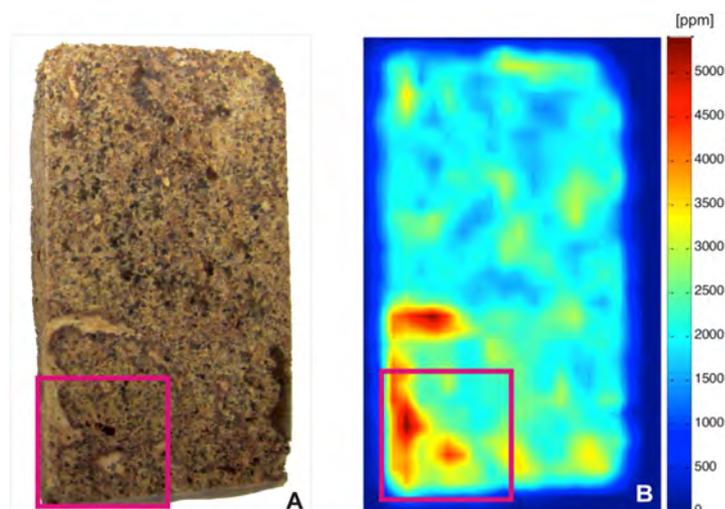
This technique was used for mapping of the largest possible surface of ore samples to determine the most abundant spheres of occurrence the elements of interest. At a lateral resolution 1–2 mm per image point and dwell time from 8 to 20 s, it has been, in a relatively short time, possible to examine the sample surfaces of roughly 2500–4000 mm<sup>2</sup>.

### 2.3. Laser-Induced Breakdown Spectroscopy (LIBS)

The experimental double-pulse (DP) LIBS setup was utilized in order to reach high sensitivity, maintaining low crater diameter, subsequently enabling high spatial resolution chemical mapping. High energy Nd:YAG laser LQ 529A (Solar LS, Belarus), operating on its second harmonic (532 nm, 12 ns pulse duration), was introduced into the LIBS interaction chamber (AtomTrace, CZ) by a series of mirrors and then focused by a 20 mm focal length glass triplet (Sill Optics, Germany) collinearly with the sample surface normal. Secondary laser, Brilliant b (Quantel, FR) operated at the fundamental wavelength (1064 nm, 8 ns pulse duration), was introduced into the chamber perpendicularly with respect to the first laser pulse axis using mirrors and then focused into the emerging plasma with 40 mm focal length lens. Such arrangement resulted in the orthogonal DP LIBS, where the first laser pulse

**Table 2**  
Experimental parameters of DP LIBS measurements.

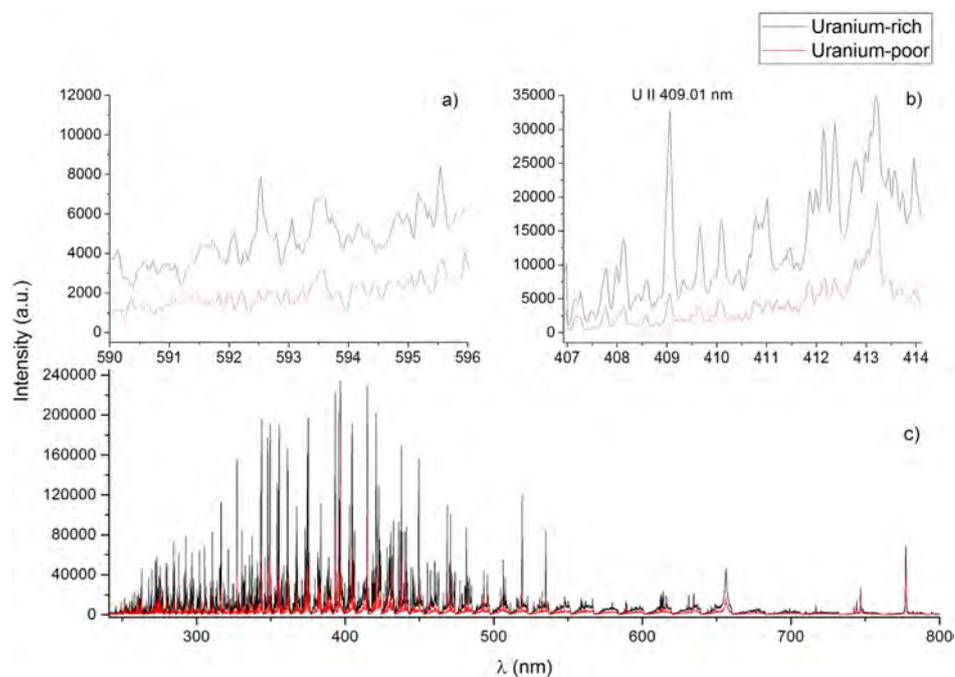
Parameter	Value
Ablation laser energy (mJ)	30
Secondary laser energy (mJ)	80
Spot size ( $\mu\text{m}$ )	50
Interpulse delay ( $\mu\text{s}$ )	0.5
Gate delay ( $\mu\text{s}$ )	1.5
Gate width ( $\mu\text{s}$ )	20
Spatial resolution ( $\mu\text{m}$ )	100



**Fig. 1.** Maps of spatial distribution of selected elements in the sample of uranium ore provided using XRF analysis. A) photograph of scanned sample (sized 70 × 44 mm); B) distribution of uranium content within the sample surface. Red square highlights the region for further LIBS analysis (15 × 15 mm). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

served for ablation of the material and LIP creation and the second laser pulse served for LIP signal enhancement. The ablation crater diameter was measured using optical microscope. Radiation of luminous laser-induced plasma (LIP) was collected with mirror optics CC52 (Andor, UK) and via optical fibre (Ø40 µm, Thorlabs, USA) fed to an echelle spectrometer Mechelle 5000 (Andor, UK; 200–975 nm, F/7, 6000  $\lambda/\Delta\lambda$ ).

Spectrally resolved radiation was then recorded using ICCD camera iStar 734i (Andor, UK; 1024 × 1024 pixels, effective pixel size 19.5 × 19.5 µm). The gate delay and gate width were controlled using a pulse generator DG535 (Stanford Research System, US) and special electronics developed in the laboratory of Brno University of Technology.



**Fig. 2.** a) The difference in spectra for studied background range (590–595 nm). b) Region of emission spectrum with uranium ionic line (409.01 nm) for both uranium-rich and uranium-poor measurements. c) Full range LIBS emission spectrum of uranium-rich and uranium-poor part of sample.

All the aforementioned devices were computer-controlled via AtomTrace Suite software. Thanks to this software it was possible to automatically provide high resolution chemical maps [23] by synchronizing the ablation and the reheating lasers, the precision movements of the sample holder and the detection system. This software allowed us setting the number and the distance of the ablation craters in the x and y direction. Moreover, via autofocus algorithm even samples with 3D topology can be scanned [24]. Chemical maps of selected elements could be monitored online.

The main advantage of the DP LIBS technique is retaining low limits of detection with small size of ablation craters [25]. The sample was moved to provide fresh spot before consecutive DP LIBS measurements, i.e. single pulse of the first ablation laser was introduced to each spot. Experiment parameters were optimized to obtain highest signal to background ratio (the optimization procedure was performed analogously as it is described in [26], the selected spectral line was U II 409.01 nm and the background region was in the close proximity of this line). All experimental parameters are listed in Table 2. Sample of sandstone was mapped with spatial resolution 0.1 mm and crater diameter 50  $\mu\text{m}$ . The mapped area was 15  $\times$  15 mm, which resulted in a chemical map of 150  $\times$  150 locations.

#### 2.4. Multivariate analysis

Typical modern experiments provide vast amount of data and LIBS is not an exception; utilizing echelle spectrograph one can easily obtain thousands of spectral variables within a single laser shot. Multiplied by the map size the amount of information is enormous. Nevertheless, bulky data sets can be statistically analyzed by means of multivariate data analysis (MVDA) [27]. Fundamental methods of MVDA are projection methods, namely: principal component analysis (PCA) and partial least squares projection to latent structures (PLS). The input for PCA is matrix of sizes  $N \times K$ , where  $N$  stands for number of objects (in our case laser pulses/measurements) and  $K$  stands for variables (spectra recorded on detector). Statistically, PCA finds projections called principal components (PC) that approximate the data in the least squares sense. There are as many PCs as is the rank of input matrix.

To each PC a loadings vector is assigned, which defines the linear combination of variables in the original  $K$ -space that PC represents. In words of our application loadings vector represents information about which spectral variable contributes the most to the formation of chosen PC. PCA computes a scores vector representing projections of objects into newly constructed space given by PCs. Moreover, each principal component stands for certain portion of variance within the original data set. First principal component describes the greatest portion of variance within the data, and so on. When cross-plotting scores of selected PCs, one can reveal hidden structures in data, such as clusters or outliers. The sorting using PCA can be very useful for instance in the application like chemical mapping of any inhomogeneous sample with unknown composition. In order to select appropriate spectral line to estimate the spatial distribution of certain element the data can be sorted in different clusters and only the representative of each cluster may be analyzed.

In ideal case, some of the PCs or their combination can be associated with physical or chemical property of the sample, so called hidden variable, which is not directly measured. An example of these hidden variables can be the different mineralization of geological samples or ores. This paper shows that the first PC is associated with the presence of uranium in ore.

We have utilized software R [28], namely the *prcomp* function from package *stats* for purposes of LIBS data analysis and MVDA. As an option of the function *prcomp* the data were mean centered and not scaled. A set of functions from package *ggplot2* was used for map visualizations.

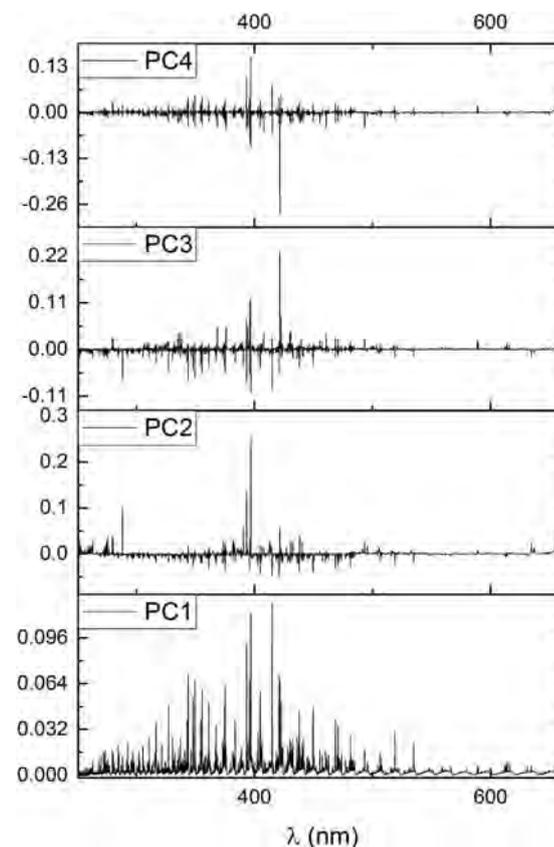
**Table 3**  
Pair-wise correlation of individual LIBS and XRF chemical maps (depicted in Fig. 1 and Fig. 4).

	U II line	Sum of background	PC 1	Predicted PC1	XRF
U II line	1	0.910312	0.94653	0.94644	0.80039
Sum of background	0.910312	1	0.95479	0.95440	0.77935
PC 1	0.94653	0.95479	1	0.99997	0.79633
Predicted PC1	0.94644	0.95440	0.99997	1	0.79634
XRF	0.80039	0.77935	0.79633	0.79634	1

### 3. Results and discussion

In the first step of the analysis, XRF was utilized in order to get information about spatial distribution of selected anomalous elements. In Fig. 1 is depicted spatial distribution of uranium. Based on this measurement, the most analytically interesting part of the sample (with highest content of uranium), was selected for further LIBS analysis. This region is highlighted in the Fig. 1A and B with red square. Selected region was then measured by the means of DP LIBS. Typical DP LIBS emission spectra of uranium-rich and uranium-poor part of sample are depicted in the Fig. 2a–c.

The elemental mapping of the sample with fine spatial resolution is enabled using DP LIBS system. Afterwards, obtained data set, consisting of 150  $\times$  150 measurements, was analyzed using four different approaches. In the first instance, the intensity of selected spectral ionic



**Fig. 3.** Loadings of first four PCs. In PC1 is apparent contribution from seeming background - uranium.

line U II detected at 409.01 nm (depicted on Fig. 2b) was fitted with pseudo-Voigt profile and the integral intensity under the fitted line was considered as analytical signal, with no further normalization. This was done on each spectra in the whole data set and responses for individual spectra were distributed in a map depicted in Fig. 4a. This spectral line was selected considering the obtained limits of detection of uranium in ores using different uranium spectral lines published in ref. [12]. During optimization measurements it was noticed that the U II line detected at 409.01 nm was the only well-resolved uranium line found throughout the whole LIBS spectrum (ranging from 200 to 900 nm).

Chinni et al. [12] proved that typical uranium spectrum consists of more than 5000 lines in the range 384.8 to 908.4 nm. Considering the spectral line broadening and resolution of commonly used spectrometers uranium lines are rather not spectrally resolved but interfere with one another. This results in seeming increase in the overall background (Fig. 2a). This effect was described in detail by Chinni et al. [12] comparing high resolution spectra of a uranium hollow cathode lamp with LIBS spectra of uranium using Czerny–Turner and Echelle spectrographs. Thus analysis of uranium employing typical LIBS setup is challenging. That means that the presence of uranium in the sample could be associated with the background intensity increase (Fig. 2a). Based on this assumption, short region of spectra (590–595 nm) without any interfering spectral lines of other elements present in the sample throughout the whole data set was identified. This spectral region was selected with respect to XRF measurements and known sample composition respectively. In this region it was expected only the “background” intensity enhancement when certain content of uranium was present.

Therefore in the whole data set, background intensity was integrated over selected spectral region and considered as new analytical signal. Analyzed background signal was then plotted as can be seen in Fig. 4b. Both figures given so far have good visual and numerical correlation see Table 3.

Concerning MVDA, the PCA was applied on whole data matrix (where rows were individual measurements and columns spectral variables) and new space of coordinates was established. Naturally, the greatest variability was attributed to the first principal component, in this case over 82%. Then, scores of this first principal component (PC1) were plotted in a map similarly as in previous cases, Fig. 4c. It is apparent that the resulting map describes the spatial distribution of sample components regarding their spectral response. As it is stated in the study of Chinni et al. [12], uranium affects spectral response of LIP in the whole spectral range. From this assumption we derived the approach of spectral data treatment which to the best of our knowledge was not used before, at least in LIBS community. It is expected that this variability is mostly associated with the occurrence of uranium in region of interest, i.e. ablation spot. Considering the PC1 loadings (Fig. 3) this assumption seems to be legitimate (see the typical echelle background in PC1 in comparison to PC2, PC3 or PC4). As can be seen, the distribution correlates with both previous cases (see Table 3). Thus, it may be stated that the meaning of first latent variable, i.e. first PC, corresponds to the uranium distribution on the sample surface.

It is computationally extensive to apply PCA on such bulky data matrix of dimensions  $22500 \times 26000$  (locations  $\times$  spectral variables), demanding more than 33 GB of RAM. In order to simplify the computing process it was tested a simplified procedure of PCA. Part of the original

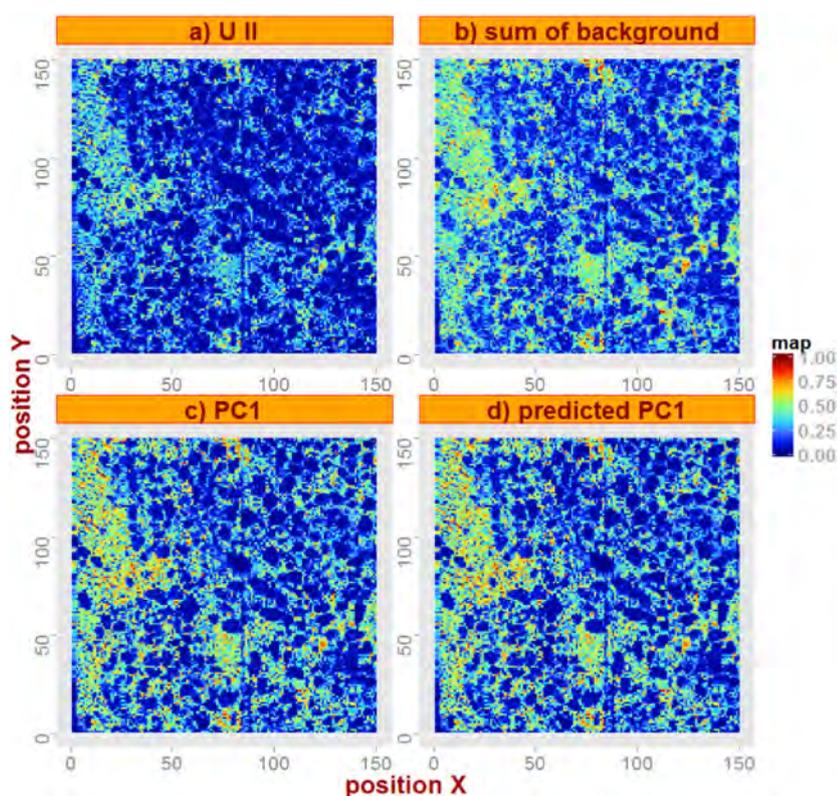


Fig. 4. Maps of uranium spatial distribution depicted as a) intensity of U II @ 409.01 nm line, b) sum of background intensity in the region from 590 to 595 nm, c) scores of PC1 and d) scores of predicted PC1 respectively.

data, the matrix was used to train the PC loading and scores. It was estimated that first  $50 \times 50$  data points are representative for the variability within the rest of the data matrix based on inspection of Fig. 3a–c. Therefore, PCA was applied on reduced data matrix ( $50 \times 50$  data points) and a new system of coordinates was obtained. Variation in the first principal component constructed from reduced data set was above 85%. Afterwards, whole original data matrix was introduced into the computation with pre-estimated loading vector and new scores matrix was predicted for the whole data set. This approach resulted in less extensive computation (operation memory and turn-around time). Finally, predicted scores of first principal component were plotted in Fig. 4d.

From the loadings of first principal component it is clear that the scores are not associated solely to the uranium content. As well determination of spatial distribution of selected element based on detection of the single spectral line and background intensity respectively can't be accepted unquestioningly. Therefore the pair-wise correlation of each LIBS chemical map with XRF map was performed. Since the lateral resolution of XRF chemical map is only 1–2 mm (in comparison to 0.1 mm for LIBS) the resolution of LIBS chemical maps was decreased accordingly. The lower resolution was achieved by summing of appropriate points in LIBS chemical map. The results of correlation for particular chemical maps are presented in Table 3.

Each of the previous computation algorithms was done using lab-made R script. To prove the mutual visual correlation of all presented maps in Fig. 4, numerical correlation was computed. It is clearly given in Table 3 that overall pair-wise correlation of presented LIBS maps is over 90% and LIBS to XRF correlation is approximately 80%. It is not surprising that maps constructed using scores of principal component and predicted scores based on trained principal components correlate at the level reaching almost unity.

#### 4. Conclusion and ongoing work

In this study we showed the possibility of high dimensional mapping by means of MVDA. The measurements were performed using orthogonal DP LIBS. By utilizing multivariate methods we were able to process the complete spectral information acquired by broadband echelle spectrograph and obtain the scores of PCs from all spectra. The scores of PC1 were plotted considering their position on the sample. The resulting map describes the spatial distribution of the sample components regarding the spectral response. Considering the papers dealing with uranium detection we expected that strongest variations in spectral response are associated with presence of uranium. Results obtained using conventional ways of LIBS data analysis (map of surface selected spectral line), and those suggested for the purposes of uranium monitoring (plotting sum of selected region of background intensity as well as scores of principal components) are in a good agreement, pair-wise correlation over 90%. XRF method was used as a reference technique to provide reliable analysis of spatial distribution of uranium in the sample. The pair-wise correlations of above described LIBS maps and uranium map obtained via XRF are approximately 80% (see Table 3). Considering the differences between LIBS and XRF the correlation seems to be satisfactory to prove our assumption about PC1 and background association to uranium distribution.

Moreover, as a new approach stated in this article, multivariate analysis was found viable for large datasets of raw spectral information, where appropriately chosen part of data covered enough variation to predict the whole map information and brought a new insight into the problem. MVDA and conventional results are in good mutual agreement. It can be seen in Table 3 that the sum of background spectral region is slightly better correlated to MVDA analysis than the intensity of uranium ionic line, this suggests that spectrometer resolution is low and the background is more sensitive to uranium abundance than the intensity of ionic line.

Ongoing work will be focused on the study of principal components themselves, their mutual dependence or connection to the physical or chemical properties of the sample. Possibly, the information about element association within the specimen can be extracted or devised using MVDA.

#### Acknowledgement

This research has been financially supported by the Ministry of Education, Youth and Sports of the Czech Republic under the project CEITEC 2020 (LQ1601). P.P. and J. Ka. would like to acknowledge that this work is also an output of GACR project 15-23274S: Design of advanced materials using selective laser melting. J. Ka. would like to acknowledge the support by the project "CEITEC – Central European Institute of Technology" on the frame of grant STI-S-14-2523 (Advanced nanotechnologies and materials). P. Veselý and DIAMO s.p. are gratefully acknowledged for providing samples.

#### References

- [1] Electricity generation - what are the options? World Nuclear Association, 2015.
- [2] Geology of uranium deposits, in, World Nuclear Association.
- [3] World distribution of uranium deposits: original resources of uranium deposits, in, International Atomic Energy Agency.
- [4] World distribution of uranium deposits: numbers of uranium deposits, International Atomic Energy Agency, 2012.
- [5] B. Schram, P. Kühn, P. Sulovský, Výzkum mineralogicko-geochemického hodnocení uranových rud: závěrečná zpráva za období leden 1977 - říjen 1980, in, Stráž pod Ralskem, 1980.
- [6] L.G. Woodruff, W.F. Cannon, C.L. Dicken, S. Pimley, Bedrock and Soil Geochemistry from Voyageurs National Park, US Department of the Interior, US Geological Survey, Minnesota, 2002.
- [7] R.L. Johnson, L.A. Durham, C.R. Rieman, J.E. Hummel, Triad case study: Rattlesnake Creek, Remediat. J. 15 (2004) 69–77.
- [8] M.L. Palacios, S.H. Taylor, Characterization of uranium oxides using in situ micro-Raman spectroscopy, Appl. Spectrosc. 54 (2000) 1372–1378.
- [9] C. Moulin, P. Decambox, P. Mauchien, D. Pouyat, L. Couston, Direct uranium (VI) and nitrate determinations in nuclear reprocessing by time-resolved laser-induced fluorescence, Anal. Chem. 68 (1996) 3204–3209.
- [10] C. Moulin, P. Decambox, V. Moulin, J.G. Decaillon, Uranium speciation in solution by time-resolved laser-induced fluorescence, Anal. Chem. 67 (1995) 348–353.
- [11] V. Piskalnik, G. Pshakin, V. Roshchenko, Review of methods and instruments for determining undeclared nuclear materials and activities, Sci. Glob. Secur. 14 (2006) 49–72.
- [12] R.C. Chinni, D.A. Cremers, L.J. Radziemski, M. Bostian, C. Navarro-Northrup, Detection of uranium using laser-induced breakdown spectroscopy, Appl. Spectrosc. 63 (2009) 1238–1250.
- [13] D.A. Cremers, A. Beddingfield, R. Smithwick, R.C. Chinni, C.R. Jones, B. Beardsley, L. Karch, Monitoring uranium, hydrogen, and lithium and their isotopes using a compact laser-induced breakdown spectroscopy (LIBS) probe and high-resolution spectrometer, Appl. Spectrosc. 66 (2012) 250–261.
- [14] E.J. Judge, J.E. Barefield II, J.M. Berg, S.M. Clegg, G.J. Havrilla, V.M. Montoya, L.A. Le, L.N. Lopez, Laser-induced breakdown spectroscopy measurements of uranium and thorium powders and uranium ore, Spectrochim. Acta B At. Spectrosc. 83–84 (2013) 28–36.
- [15] Y.S. Kim, B.Y. Han, H.S. Shin, H.D. Kim, E.C. Jung, J.H. Jung, S.H. Na, Determination of uranium concentration in an ore sample using laser-induced breakdown spectroscopy, Spectrochim. Acta B At. Spectrosc. 74–75 (2012) 190–193.
- [16] X.K. Shen, Y.F. Lu, Detection of uranium in solids by using laser-induced breakdown spectroscopy combined with laser-induced fluorescence, Appl. Opt. 47 (2008) 1810–1815.
- [17] D.A. Cremers, L.J. Radziemski, Handbook of Laser-Induced Breakdown Spectroscopy, 2013 426.
- [18] 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.
- [19] 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.
- [20] F.J. Fortes, J.J. Laserna, The development of fieldable laser-induced breakdown spectrometer: no limits on the horizon, Spectrochim. Acta B At. Spectrosc. 65 (2010) 975–990.
- [21] B. Schram, Zirkonium v sedimentech Českého masívu a jeho anomální koncentrace v severočeské křídě, Sborník geologických Věd, ÚÚG ČSAV, Praha 1984, pp. 9–80.
- [22] M. Schramová, B. Schram, L. Novák, Mineralogická charakteristika západní části ložiska Břevniště: "Pásmo strážského zlomu", in, VZUP, Stráž pod Ralskem, 1986.
- [23] V. Piňón, M.P. Mateo, G. Nicolas, Laser-induced breakdown spectroscopy for chemical mapping of materials, Appl. Spectrosc. Rev. 48 (2013) 357–383.
- [24] J. Novotny, R. Malina, J. Kaiser, M. Liska, M. Galiova, K. Novotny, Implementation of an autofocus algorithm based on searching the best in-focus image into a

- table-top laser-induced breakdown spectroscopy setup, *Opt. Eng.* 48 (2009), 103604.
- [25] V.I. Babushok, F.C. DeLucia, J.L. Gottfried, C.A. Munson, A.W. Miziolek, Double pulse laser ablation and plasma: laser induced breakdown spectroscopy signal enhancement, *Spectrochim. Acta B At. Spectrosc.* 61 (2006) 999–1014.
- [26] D. Prochazka, M. Bilik, P. Prochazkova, J. Klus, P. Porizka, J. Novotny, K. Novotny, B. Ticova, A. Bradac, M. Semela, J. Kaiser, Detection of tire tread particles using laser-induced breakdown spectroscopy, *Spectrochim. Acta B At. Spectrosc.* 108 (2015) 1–7.
- [27] L. Eriksson, T. Byrne, E. Johansson, J. Trygg, C. Vikström, *Multi- and Megavariate Data Analysis Basic Principles and Applications*, Umetrics Academy, 2013.
- [28] R. Ihaka, R. Gentleman, R: a language for data analysis and graphics, *J. Comput. Graph. Stat.* 5 (1996) 299–314.

## Article 11

P. Pořízka, J. Klus, D. Prochazka, E. Képeš, A. Hrdlička, J. Novotný, K. Novotný, and J. Kaiser. Laser-Induced Breakdown Spectroscopy coupled with chemometrics for the analysis of steel: The issue of spectral outliers filtering. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 123:114–120, 2016.

DOI: 10.1016/j.sab.2016.08.008; 20 citations

Another publication from the body of work related to the rigorous investigation of data preprocessing approaches. From my point of view, the outliers filtering is the very first step that should be done. Naturally, it is necessary to filter gross errors from the data set and mitigate the bias that they would induce.

In this publication, a set of steel standards was selected for their busy spectra which complicates the data processing itself. Obtained data were processed in their full format without any extraction of intensities of spectral lines. Nonlinear Sammon's maps were used for visualization.

Three different approaches were selected to detect outliers within obtained data:

- spectra of each sample were project to the principal component analysis (PCA) space (first three components) and their distance to origin was calculated and ordered from lowest to highest;
- Pearson's correlation coefficient was calculated, 50 % of spectra with highest mutually correlated spectra were left for further consideration;
- total intensity of each spectrum was estimated and then ordered, marginal 50 % (25 % of lowest and 25 % of highest total intensities) were discarded.

In this case, 50 % of all data were considered as outliers. Such extreme case led to more significant changes in the data matrix. It was shown, that each approach selected different data points as outliers. Thus, I recommend to use more alternative approaches to compare their results.

My personal favourite is, without any surprise, the utilization of PCA.



## Laser-Induced Breakdown Spectroscopy coupled with chemometrics for the analysis of steel: The issue of spectral outliers filtering



Pavel Pořízka<sup>a,\*</sup>, Jakub Klus<sup>a</sup>, David Prochazka<sup>a</sup>, Erik Képeš<sup>b</sup>, Aleš Hrdlička<sup>a,c</sup>, Jan Novotný<sup>a</sup>, Karel Novotný<sup>a,c</sup>, Jozef Kaiser<sup>a</sup>

<sup>a</sup> Brno University of Technology, Central European Institute of Technology, Technická 3058/10, CZ-616 00 Brno, Czech Republic

<sup>b</sup> Brno University of Technology, Faculty of Mechanical Engineering, Technická 2896/2, CZ-616 69 Brno, Czech Republic

<sup>c</sup> Masaryk University, Central European Institute of Technology, Kamenice 753/5, CZ-625 00 Brno, Czech Republic

### ARTICLE INFO

#### Article history:

Received 1 January 2016

Received in revised form 4 August 2016

Accepted 5 August 2016

Available online 5 August 2016

#### Keywords:

Laser-Induced Breakdown Spectroscopy, LIBS

Outlier filtering

Principal Component Analysis, PCA

Linear correlation

Total spectral intensity

Soft Independent Modelling of Class Analogies,

SIMCA

### ABSTRACT

In this manuscript we highlight the necessity of outlier filtering prior the multivariate classification in Laser-Induced Breakdown Spectroscopy (LIBS) analyses. For the purpose of classification we chose to analyse BAM steel standards that possess similar composition of major and trace elements. To assess the improvement in figures of merit we compared the performance of three outlier filtering approaches (based on Principal Component Analysis, linear correlation and total spectral intensity) already separately discussed in the LIBS literature. The truncated data set was classified using Soft Independent Modelling of Class Analogies (SIMCA). Yielded results showed significant improvement in the performance of multivariate classification coupled to filtered data. The best performance was observed for the total spectral intensity filtering approach gaining the analytical figures of merit (overall accuracy, sensitivity, and specificity) over 98%. It is noteworthy that the results showed relatively low sensitivity and high specificity of the SIMCA algorithm regardless of the presence of outliers in the data sets. Moreover, it was shown that the variance in the data topology of training and testing data sets has a great impact on the consequent data classification.

© 2016 Elsevier B.V. All rights reserved.

### 1. Introduction

Laser-Induced Breakdown Spectroscopy (LIBS) [1–4] is a technique based on Atomic Emission Spectroscopy gaining its popularity due to its advantages (e.g. multi-elemental capability, fast turn-around time, no need for sample preparation) as can be seen from comprehensive review articles across applications [5–11]. This technique is based on spectral analysis of collected Laser-Induced Plasma (LIP) radiation.

LIBS system is capable of a spectrochemical analysis of fast repetitive measurements. Each sample is then represented by a vast number of collected LIP spectra. The assumption is given that all spectra represent the sample they were obtained from. Each LIP spectrum carries characteristic information, i.e. chemical fingerprint, about the elemental composition of the investigated interaction spot on the sample surface. Yet similarly to other laser-ablation based techniques, LIBS is significantly affected by changes in properties of laser-matter interaction, i.e. the matrix effect. It is therefore necessary either to filter the spectral outliers from the data set, or to normalize the LIP signal *a priori*.

Significant changes in the laser-matter conditions are assigned to variations in chemical composition and physical properties of the sample

matrix. Local inhomogeneity together with fluctuating performance of the analytical system influence the laser-matter interaction and, in turn, the LIP dynamics and its properties. Consecutive classification and quantification would be then biased, giving lower accuracy, sensitivity, etc. Thus, some detected spectra are not representative of the rest of the data set and should be regarded as outliers. El Haddad [6] et al. highlights the necessity of thorough data pre-treatment prior to any further data analysis. This data pre-treatment dictates among others to detect and omit outlying data. Several works dealt with the spectral outliers filtering and consequently with the improvement of the analytical figures of merits. Various metrics are utilized in order to filter out the outliers giving non-unified protocol for the data pre-treatment.

Mermet et al. [12] studied the influence of sample inhomogeneity and other sources of signal fluctuation on the precision of LIBS analysis at the micro scale. They concluded that the outliers found by Hampel test had no impact on the performance of their LIBS system, since the internal standardization was efficient enough to mitigate any sources of fluctuation. Numerous studies were undertaken to reduce or to avoid matrix effect occurrence due to fluctuation in the laser-matter conditions in a typical LIBS experiment. To reduce fluctuations in LIP signal, normalization of the spectrum to its total intensity, internal standard, or unity may be utilized [11,13]. Those approaches are mainly effective in those cases when samples of the same matrix are of interest.

\* Corresponding author.

E-mail address: [pavel.porizka@ceitec.vutbr.cz](mailto:pavel.porizka@ceitec.vutbr.cz) (P. Pořízka).

Yaroshchik et al. [14] introduced the algorithm based on variations in the total sum of spectral intensity (Total Energy, TE), where spectra with the value of TE higher than two standard deviations from the mean value were discarded, giving the reduction of data set by maximum 10%. Carranza and Hahn [15] established a filtering algorithm on selected Fe II lines to reduce the shot-to-shot variability in the LIBS analysis of aerosols. Due to the phenomenon of moving breakdown, they discarded 60–70% of spectra. In the research introduced by Lazic et al. [16] the biggest improvement in the analytical performance of LIBS applied to various states of matter was achieved when 50% of data were filtered out.

In the articles concerning multivariate classification and quantification the multivariate algorithms were adapted also for outliers filtering. It is generally accepted in the LIBS literature that multivariate algorithms may compensate for the matrix effect to a certain extent [6]. This can lead to mitigation of fluctuations in LIBS data set. However, the outliers persist even after multivariate data treatment as proved by Braga et al. [17]. Gornushkin et al. [18] noted that linear correlation and Principal Component Analysis (PCA) are the most frequently utilized algorithms in LIBS data processing. Each sample may be represented by a signature spectrum and all spectra from the same sample must mutually correlate nearly perfectly. For that reason, they estimated the correlation threshold at 99% for outliers filtering, when the whole spectral set was regarded as the input for linear correlation only. This approach gave the possibility to reflect mutual differences of all spectral variables, which can be beneficial for filtering. The authors also stressed out, that spectra truly representing each sample can be obtained via a great number of repetitions giving a solid basis for statistics.

Sirven et al. [19] listed aspects of outliers filtering in their manuscript. In basic cases the selected line ratios can be used to classify the samples, but not to detect potential outliers. Therefore also in complex cases, when diverse sample matrices are under investigation, more sophisticated algorithms are needed. The authors proposed utilization of multivariate algorithm, PCA, which takes into account mainly the distinct variations among variables describing data points (spectra). Also, it provides visualization of the data set in newly constructed space describing the biggest amount of latent variation, i.e. principal components (PCs). This approach readily shows the spectra mutually similar as they group to a cluster as well as outliers, i.e. abnormally different spectra. It is then straightforward to discard the outlying data points from the data set. In their work, 30% of spectral outliers per each matrix were filtered out based on visual inspection of loadings and PC scores plot. PCA was utilized for outlier filtering also in other studies [20–22].

In this manuscript we aim to compare the effect of outliers filtering approaches on the consequent classification of a truncated data set. For those purposes we selected the PCA, linear correlation and total spectral intensity approaches. The spectra were obtained from LIBS measurement of BAM steel standards. After filtering, Soft Independent Modelling of Class Analogies (SIMCA) algorithm was applied on the data set (constructed from the whole spectra with no masking or normalization). The improvement in analytical performance of LIBS coupled with SIMCA after outliers filtering was observed and compared via yielded figures of merits values.

## 2. Experimental

### 2.1. Samples

The sample set consisted of ten certified steel standards, provided in the frame of LIBS 2008 contest by BAM (Federal Institute for Materials Research and Testing, DE). The compositions of BAM standards are listed in Table 1. The samples surfaces were finely polished by 3  $\mu\text{m}$  and 1  $\mu\text{m}$  diamond pastes (Urdiamant, CZ) prior to LIBS analysis. The sample surface was wiped with ethanol after the diamond paste polishing. The sample

**Table 1**  
composition of BAM steel standards in wt.%, content of iron is a supplement to 100 wt.%.

BAM	Ni	Mn	Cr	C	Si	Mo	Co
C1	12.55	0.74	12.35	0.092	0.46	–	–
C2	6.124	0.686	14.727	0.0103	0.374	0.0138	–
C3	12.85	0.722	11.888	0.0345	0.463	0.0304	–
C4	10.2	1.4	18.46	0.019	0.27	0.265	0.116
C5	20.05	0.791	25.39	0.086	0.57	–	0.054
C6	9.24	1.38	17.31	0.066	0.405	0.092	0.053
C7	10.2	1.311	17.84	0.0141	0.48	2.776	0.0184
C8	8.9	1.7	17.96	0.143	1.41	–	0.018
C9	5.66	0.89	14.14	0.05	0.21	1.59	0.22
C10	10.72	1.745	16.811	0.0201	0.537	2.111	0.0525

surface was polished in order to reduce any potential fluctuations in LIP signal due to the surface roughness.

### 2.2. LIBS system

Measurements were performed utilizing the Sci-Trace device including the LIBS Interaction Chamber (both by AtomTrace, CZ) under the atmospheric pressure. A laser pulse (LF 121, Sol Instrument, BY; flashlamp pumped Nd:YAG, 1064 nm, 9 ns pulse duration) was introduced into the chamber by a set of mirrors (NB1-K13, Thorlabs, USA). The laser pulse of 90 mJ was focused on the sample surface into a tight spot ( $\varnothing \sim 100 \mu\text{m}$ ) with a 40 mm anti-reflex coated focal length lens (Thorlabs, USA). Radiation of luminous LIP was collected with the reflective optics CC52 (Andor, UK), the axis of collection optics was levelled from the focusing optics axis by the angle of  $70^\circ$ . Then the collected radiation was introduced into an echelle spectrometer Mechelle 5000 (Andor, UK; 200–975 nm, F/7, 6000  $\lambda/\Delta\lambda$ ) via optical fibre ( $\varnothing 40 \mu\text{m}$ , Thorlabs, USA). Spectrally resolved radiation was recorded using ICCD detector iStar 734i (Andor, UK; 1024  $\times$  1024 pixels, effective pixel size 19.5  $\times$  19.5  $\mu\text{m}$ ) with temporally gated detection. Timing of the experiment was controlled using a pulse generator DG535 (Stanford Research System, US) and a signal inhibitor developed in the Laboratory of laser spectroscopy (Brno University of Technology, CZ). The whole system was operated using AtomChamber software (AtomTrace, CZ).

Optimization of the system to the best signal-to-background ratio resulted in the values of temporal gating of the ICCD as follows: gate delay of 2.5  $\mu\text{s}$  and gate width of 5  $\mu\text{s}$ . The samples in the form of chipped cylinders were analysed by a series of 100 consecutive laser pulses with the irradiance of 130 GW/cm<sup>2</sup>. The sample was moved before each measurement by the step of 200  $\mu\text{m}$  to provide fresh spot for each laser pulse.

### 2.3. Data processing

Especially in the analysis of iron and its alloys, the broadband LIBS emission spectra offer a high number of variables that can be beneficially utilized for multi-variate classification and quantification. Thus, we decided to work with the complete spectral range and not to select any particular regions or spectral lines of iron and other minor and trace elements. It is also beneficial to couple the whole spectral range into the multivariate algorithm, since also minor spectral lines could contribute significantly to the analysis. The obtained data set was organized into a data matrix  $\mathbf{X}$ , where rows represented individual measurements and columns were assigned to variables/wavelengths. The spectra were not normalized, though it was proved in several studies that spectra normalization can compensate for fluctuations in LIP signal. This is actually an unwanted feature in our study. The data matrix  $\mathbf{X}$  was just mean centred along the columns.

In the filtering step, three approaches based on PCA, linear correlation and total spectral intensity were selected. In all cases we intended to show extreme example and thus 50% of measurements per sample

were discarded, i.e. 50% of data were considered as outliers. It is a well-known fact that the number of repetitions/measurements and their fluctuation inserted in the multivariate algorithm affect its performance. Thus, in order to be consistent regarding also the multivariate analysis of the non-treated original data set we randomly selected 50 of the original 100 measurements per sample. So the dimensions of data matrices (filtered and non-filtered) as well as the results assessed from SIMCA models are more comparable throughout the whole study. Moreover, 50 spectra were randomly divided into training and testing sets by the ratio 3:2, in spite of the advice by El Haddad et al. [6] to divide the data set to training, validating and testing sets. We judged the optimal number of principal components for the SIMCA classification from the scree plots<sup>1</sup> of the individual PCA models, as commented below. Thus, no validation data set was assembled in this study.

PCA as rather visualization technique cannot be solely used for the classification. However, it considers linear relationship within the data and thus it is sensitive to outliers, data points strongly deviating from the rest of the data set. The goal of PCA [25] is to reduce the dimensionality of the data matrix  $\mathbf{X}$  and to visualise the highest amount of variability carried by the original data in low-dimensional space. For filtering, PCA algorithm was applied individually on spectra of each sample. In other publications [19–22], PCA filtering is based either on visual inspection of loadings and scores plots, or on clustering of the data in the PC space using unsupervised algorithms.

In this work, while assuming the Gaussian distribution of data, we utilized a metric given by the Euclidean distance of a data point to the centre of PC space constructed by the first three principal components. The number of three PCs was estimated from the scree plot and was valid for all partial data sets. Then, the spectra with higher distance than the selected threshold (the median of all Euclidean distances) were regarded as outliers and omitted from further computation.

A simple hypothesis exists that spectra/measurements from the same sample have to correlate almost perfectly [18], assuming that the sample is homogeneous. Therefore the threshold value for mutual linear correlation should be over 99%. Nevertheless, this threshold may change according to the targeted application or after inspection of data set fluctuations. In the case of linear correlation approach we also stood with the median rule, assuming the Gaussian distribution in the data. The linear correlation was estimated for each sample individually. A stepwise correlation of spectrum to the rest of the spectra set was estimated and the correlation matrix was constructed. Medians were estimated for all columns of this correlation matrix. Those values were ordered and quartiles were calculated. Those spectra with medians in first the two quartiles (having lower correlation coefficients) were considered as outliers and thus discarded.

The last utilized filtering algorithm was suggested by Yaroshchik et al. [14] and is based on summing the total intensities of spectra (Total Energy, TE). The TEs of all spectra per sample were computed and ordered according to their values. Then, spectra with the TE distributed in the first and the fourth quartile were discarded from the data set as being potential outliers. This was repeated for each sample. As it is obvious from above mentioned statistics, the data sets were truncated and only 50% of measurements were left for further computation after utilization of any outliers filtering method. All the above mentioned approaches utilized for outliers filtering are graphically depicted in the Appendix.

For the classification purposes we chose unsupervised SIMCA algorithm [19,24,25], which is one of the most popular pattern recognition algorithms in the LIBS community. In SIMCA, each sample set is individually modelled using PCA algorithm. Then, the distance of the unknown/test

spectrum to the centre of each PCA model is computed. This spectrum is then assigned to the class with the minimal distance. Since 50% of data were omitted from the classification, each sample was represented only by 50 spectra, giving total number of 500 spectra. For the training set, 30 spectra per sample were randomly selected (using a generator of random numbers) and assessed into the training data matrix  $\mathbf{X}$ . We estimated the number of principal components from the scree plots. This judgement was assessed from PCA models of 1000 randomly selected data sets. Hence, the obtained number of three principal components is valid regardless of the data set selection. Afterwards, testing data set constructed from 20 spectra per sample was inserted into the PCA models and their class memberships were determined.

The confusion matrix was built from obtained classes of unknown spectra and the overall accuracy, sensitivity, and specificity were estimated. The detailed definitions of confusion matrix and individual figures of merit can be found elsewhere [6]. The overall accuracy is calculated as the ratio of the sum of true positives (all testing spectra correctly classified) to total number of testing spectra. The sensitivity is given by the ratio of true positive to condition positive (the sum of true positive and false negative) for the considered class. The specificity is the ratio of true negative to the condition negative (true negative summed with false positive). Both sensitivity and specificity were estimated step by step solely for each class (samples). Then the total sensitivity and specificity were presented as the arithmetic mean of all sensitivities and specificities, respectively, for the whole data set. Finally, the performances of filtering approaches were compared based on the classification figures of merit.

The multivariate analysis of the data and visualization was assessed using R software packages [23], namely 'caret' for PCA, 'rrcovHD' for SIMCA and 'MASS' for Sammon's map.

### 3. Results and discussion

The sample set was investigated using PCA prior to any classification analysis. With this approach the data points (spectra) can be visualized in more simplified manner and only the most important spectral information is used. As can be seen in Fig. 1, the samples C1 and C5 are completely distinguished from the rest of the dataset. The rest of the samples partly overlap what will confuse their further accurate classification with no particular data pre-treatment. There is also notable fluctuation in data points of samples from C2 to C7, which are dispersed mainly across the first principal component. This elongated shape of clusters is assigned to random fluctuation of possible laser-matter interaction conditions. Fluctuations of total intensities of spectra per sample were reaching up to 30%. Regardless of the obstacles visible in the visualized data, this preliminary partial clustering in PC space (covering only 28.1% of overall variation in the data) is a promising starting point for further multivariate data analysis.

The scattering of data points observed in Fig. 1 is attributed to wavelength variables fluctuation and is visualized *via* principal component loadings, depicted in Fig. 2. The original wavelength range (from 200 to 900 nm) was cut for easier visualization and depicted only from 250 to 600 nm. The sensitivity of the spectrometer and camera is rapidly decreasing below 250 nm, showing low intense and thus less significant spectral lines. The range over 600 nm does not contain many significant lines contributing to the clustering of the data points, as proved by PCA loadings. The raw spectra set is composed from vast number of variables present namely in the depicted range and contains the most redundant information on its boundaries. This redundancy was not truncated from the data matrix for further SIMCA classification. As expected, lines with major contribution to scattering of the data in the first three principal components belong to Fe, Cr, and Ni. The majority of lines was attributed namely to Fe; the ionic lines below 280 nm and the atomic lines in the range from 340 to 400 nm. The lines of Co II at 241.63 nm (not highlighted in the figure), Mn I at 403.08 nm, Mo I at 550.65, and Si I lines at 288.15 nm and 390.55 nm were also detected, but with lower loadings

<sup>1</sup> Scree plot [24] is a bar chart where x-axis represents individual principal components ordered in increasing order, y-axis then represents the total variability described by respective principal component, i.e. the eigenvalue. The function is naturally exponentially decreasing, having maximum in the first principal component. Steep part of this curve represents the most important principal components. At certain point the curve bands and continues in more flat manner. This point is generally accepted as a good estimate for principal components number used for modelling and classification of the data.

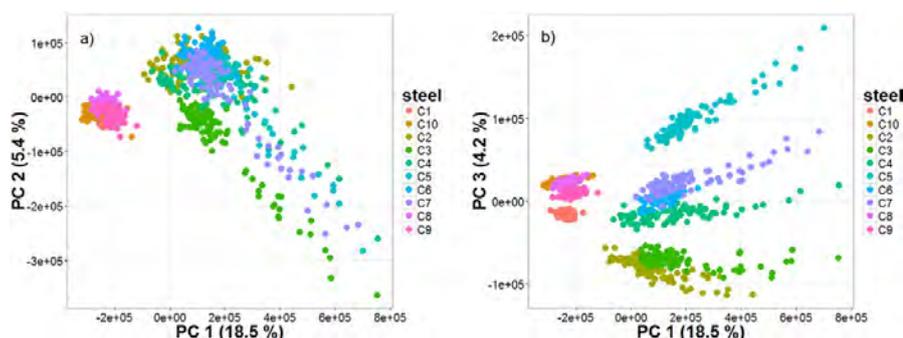


Fig. 1. Visualization of the whole data set in the space of a) the first and second principal component, b) the first and third principal component.

values due to lower concentration ranges in the samples. The observed lines were cross-checked with the NIST online database [26].

The topology of the original data, total number of measurements and samples, and number of outliers significantly influence the performance of any utilized multivariate algorithm in an exploratory data analysis. To demonstrate this phenomenon we randomly selected 50 spectra of each sample from the complete data matrix. Then SIMCA was coupled to truncated data and figures of merit were estimated. This protocol was run 1000 times to achieve robust statistics. Clearly, *via* such protocol it is possible to randomly pick up data set consisting only of outliers in the worst case. In the best case, no outliers will be selected leading to perfect classification and class membership assignment. This is reflected in figures of merit depicted in Fig. 3. The random selection of the spectra results in the worst case in poor overall accuracy of 69.5% suggesting the data set consists mostly from outliers. On the other hand the best case reaches 96% of overall accuracy concluding that well filtered data set can give convincing class membership assignment. Note that the median value of total accuracy is 84.5%. The histograms in Fig. 3 b) and c) indicate moderate sensitivity (*i.e.* assigning the unknown spectra to correct class) and good specificity (*i.e.* rate of assignment of spectra to false class) of SIMCA classification. This stands regardless of the presence and number of outliers in the truncated data set.

Prior to any further multivariate analysis, each sample spectra were independently treated to filter the outliers using three different filtering

algorithms. At this point we intended to highlight the distribution of outliers in the original data space marked by each approach. For example, the data of sample C5 is plotted in the two dimensional space assessed by Sammon's map, see Fig. 4. We utilized Sammon's map, as it was found to be the best representation of LIBS data in a 2D space by Lasue et al. [27]. This non-linear algorithm finds the global minimum of the stress function by iteration [28], while conserving the original topology of raw data to the maximum extent. Note that we firstly intended to plot the original data in a PC space. However, this step could be regarded as data manipulation since we use PCA also for data filtering.

Naturally, each algorithm filters different data points from the set as potential outliers, as it is depicted in Fig. 4. Thus, it is reasonable to expect that consequent multivariate analyses will essentially differ. The closest circle of data points is highlighted by PCA approach suggesting the best classification after such filtering. On the other hand, the data points highlighted by TE approach are distributed in the tightest cloud around the centre of the first component in Sammon's map. Therefore, the outlier filtering based on linear correlation may seem as the least feasible one. The threshold considered for data filtering is set to filter 50% of data points regardless of the outliers filtering approach. This approach was selected only for demonstration purposes and to supply the same amount of data using each filtering algorithm into the SIMCA classification. Clearly, such threshold can be too strict for the utilization in real

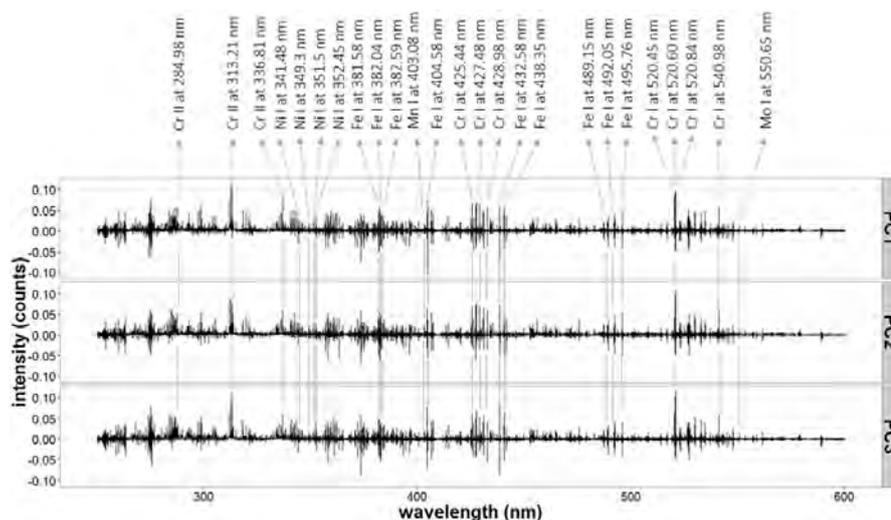


Fig. 2. Loadings of first two principal components showing the most significant lines responsible for clustering of individual data points.

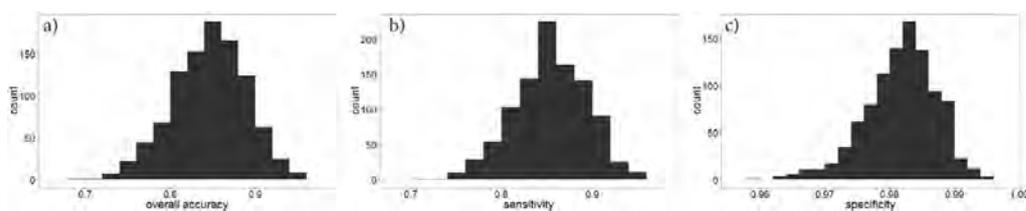


Fig. 3. Figures of merit: a) overall accuracy, b) sensitivity, and c) specificity obtained from the SIMCA classification of 1000 times randomly selected data sets.

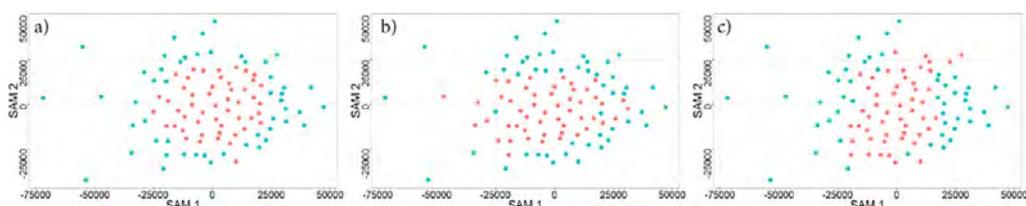


Fig. 4. Data of sample C5 depicted in first two components with marked data points left for analysis (in salmon-pink) and outliers (in cyan) filtered using a) PCA, b) linear correlation, and c) total intensity approach. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

samples analysis. Other metrics can be used and their thresholds varied according to the needs of particular application.

In the last step of our exploratory data analysis, the original data matrix was treated by selected filtering algorithms and truncated as commented above. The selection of training and testing data sets after outliers filtering was repeated 1000 times to achieve robust statistics. Then, we compared the improvement in classification figures of merit yielded from filtered data with respect to data from raw spectra. To estimate the improvement factor we regarded only the medians from the repeated classification of raw and filtered data. The significant improvement namely in overall accuracy and sensitivity is apparent from the results presented in Fig. 5. The overall accuracy and sensitivity were improved from moderate 85% to more than 95% for all filtering algorithms. As already stated above, the specificity was coherently very good, reaching 98% regardless of the data set topology. Comparing the improvement yielded utilizing individual filtering algorithms, the best performance was observed in the case of total spectral intensity. In that case, all figures of merit reached more than 98% in median. On the other hand, the relatively worst performance yielded the PCA filtering algorithm reaching only 95% for overall accuracy and sensitivity. It is worth highlighting that maximum values for accuracy, sensitivity and specificity are equal to 1, therefore they are overlapped in the Fig. 5.

These results are in a slight disagreement with the results observed in Fig. 4, where better performance of PCA filtering approaches is suggested. It was also expected that the performance of the linear correlation and total intensity approach would distinctly differ. The Euclidean distance was computed from three PCs, as optimized from the scree plots, hence the higher number of PCs could compensate better for the peak intensity fluctuation. Regarding the metric used for linear correlation filtering, in the first step each spectrum of the sample is correlated with the others and then the median of mutual correlations is estimated. Median of each spectrum is considered as the metric for outliers filtering. However, potentially outlying spectrum may yield relatively higher median value and thus nominate itself among representative spectra used for classification. This can lead to a biased data matrix and relatively lower figures of merit for linear correlation. On the other hand, total intensity is to a certain level insensitive to the fluctuation of the individual peaks and for this reason it is more applicable. Moreover, it is noteworthy that relatively high fluctuation in total intensity in spectra was observed, up to 30%. This leads to the conclusion that normalization of the spectra could improve the statistics of multivariate classification. However, the thorough study of this phenomenon is behind the scope of this paper. This we aim to

investigate in our following work, where we intend to compare the influence of various normalization techniques on the multivariate classification. We also aim on the comparison of results gained from full spectra and data matrices composed only from prominent peak intensities of major and minor elements.

#### 4. Conclusion

Firstly, we assessed the impact of the fluctuation of the original topology in LIBS data (i.e. presence of different number of outliers) on the performance of classification. 50 spectra per sample were randomly selected and truncated matrices were coupled to SIMCA. It was shown that for instance the overall accuracy may vary from 69.5% to 96%, judged from 1000 repetitions. That suggested the necessity of prior data treatment and outliers filtering. Moreover, the figures of merit significantly depend on the data set and its fluctuation. The fluctuation can cause an overlap in data points and thus confuse the classification itself. For this reason it is necessary to carefully evaluate the data set and set the threshold for outliers accordingly.

The main aim of this work was to study the influence of various outliers filtering approaches on the performance of multivariate classification algorithm coupled with LIBS data. Naturally, each approach treats the data and filters outliers in its own way. This results in different compositions

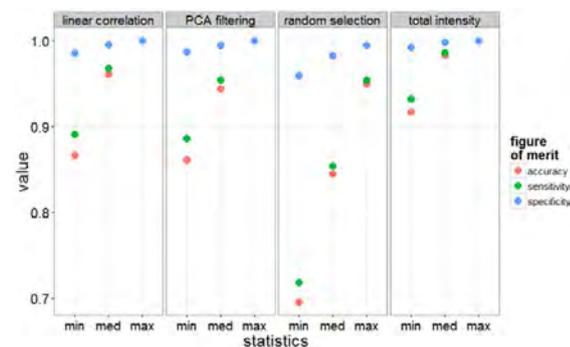


Fig. 5. Figures of merits (overall accuracy, sensitivity, and specificity) yielded for data after applying individual filtering approaches compared to those obtained for the randomly selected data.

(topologies) of truncated data sets. Since the composition of training and testing matrices affects the performance of multivariate algorithm, the training and testing sets were randomly selected for 1000 times from the truncated data sets. The presented results suggest the utilization of total spectral intensity for effective outliers filtering. The obtained figures of merit (overall accuracy, sensitivity, and specificity) reached in their median values more than 98%. The poorest performance yielded the PCA filtering approach. Yet still the total improvement after filtering is almost comparable varying from 95 to 98% in median. However, the improvement using any filtering approach is significant compared to randomly selected data. Thus, it is advisable to always filter the outliers prior to any multivariate analyses. Our results also show that SIMCA performance gained low sensitivity (median at 85%) and high specificity (median at 98%) regardless of the presence of the outliers in the data.

It is also reasonable to suggest obtaining higher repetition of LIBS measurements per sample in order to achieve robust statistics. The data has to be treated to filter the outliers. However certain fluctuation in the data has to be present to increase the probability of correct classification of the unknown spectra/data. This means that it is not possible to classify the unknown sample without any natural variation in training data. Those contradictory conditions put emphasis on the right number of data points per sample or the filtering threshold value. In our case, the filtering of outliers was done in an extreme case and 50% of the data were discarded from further computation. Such threshold may be too strict for other applications, where its value should be selected cautiously.

The outliers filtering approach based on total spectral intensity yielded the best performance due to vast fluctuations in LIP total intensity, *i.e.* data with relatively high instability. Therefore, it has to be noted that it was not our goal to investigate the sources of fluctuations. The measured spectra were regarded strictly as exemplary data and served only as a tool to make a point about the difference in outliers filtering approaches and the consequent improvement in the classification performance. The selection of normalization procedures and their impact on outliers as well as on the classification performance is the goal of the ongoing work. Also, the prominent peaks selection for further normalization will be investigated.

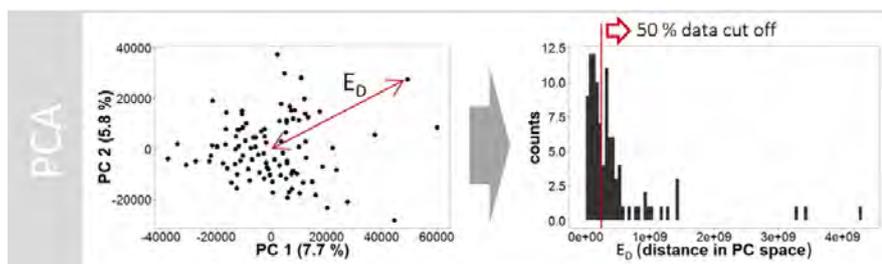
#### Acknowledgement

This research was carried out under the project CEITEC 2020 (LQ1601) with financial support from the Ministry of Education, Youth and Sports of the Czech Republic under the National Sustainability Programme II. P.P. and J.Ka. would like to acknowledge that this work is also an output of GACR project 15-23274S: Design of advanced materials using selective laser melting. We acknowledge the support from the project “CEITEC – Central European Institute of Technology” (CZ.1.05/1.1.00/02.0068) from the European Regional Development Fund. J.Ka. would like to acknowledge the support by the project “CEITEC – Central European Institute of Technology” on the frame of grant STI-S-14-2523 (Advanced nanotechnologies and materials).

#### Appendix A. Appendix

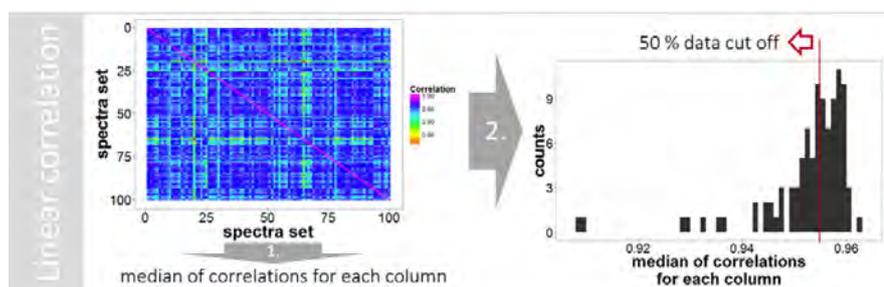
Graphical representation of individual approaches in outliers filtering. The figures are plotted for the C10 steel CRM.

##### 1) Principal Component Analysis

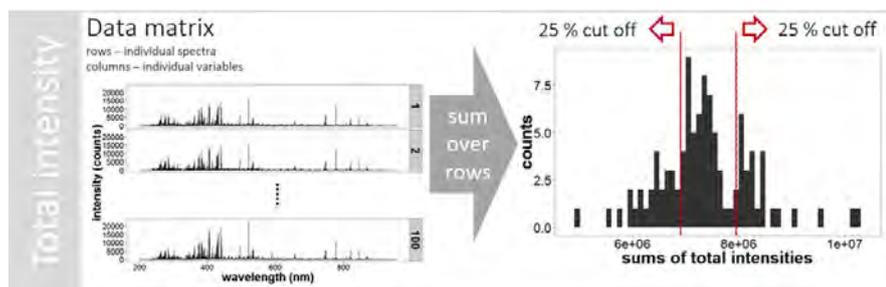


$E_D$  refers to the Euclidean distance

##### 2) Linear Correlation



## 3) Total spectral intensity



## References

- [1] A. Miziolek, V. Palleschi, I. Schechter, *Laser Induced Breakdown Spectroscopy*, Cambridge University Press, Cambridge, 2006.
- [2] R. Noll, *Laser-Induced Breakdown Spectroscopy Fundamentals and Applications*, 2012, Springer-Verlag Berlin Heidelberg, Heidelberg, 2012.
- [3] D. Hahn, N. Omenetto, *Laser-Induced Breakdown Spectroscopy (LIBS)*, part I: review of basic diagnostics and plasma-particle interactions, *Appl. Spectrosc.* 64 (12) (2010) 335–366.
- [4] C. Aragón, J. Aguilera, Characterization of laser induced plasmas by optical emission spectroscopy: a review of experiments and methods, *Spectrochim. Acta B At. Spectrosc.* 63 (9) (2008) 893–916.
- [5] D. 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 (4) (2012) 347–419.
- [6] J. El Haddad, L. Canioni, B. Bousquet, Good practices in LIBS analysis: review and advices, *Spectrochim. Acta B At. Spectrosc.* 101 (2014) 171–182.
- [7] J. Kaiser, K. Novotný, M. Martin, A. Hrdlička, R. Malina, M. Hartl, V. Adam, R. Kizek, Trace elemental analysis by Laser-Induced Breakdown Spectroscopy—biological applications, *Surf. Sci. Rep.* 67 (11–12) (2012) 233–243.
- [8] D. Santos, L. Nunes, G. de Carvalho, M. Gomes, P. de Souza, F. Leme, L. dos Santos, F. Krug, *Laser-Induced Breakdown Spectroscopy for analysis of plant materials: a review*, *Spectrochim. Acta B At. Spectrosc.* 71–72 (2012) 3–13.
- [9] P. Pořízka, P. Prochazková, D. Prochazka, L. Sládková, J. Novotný, M. Petrilak, M. Brada, O. Samek, Z. Pilát, P. Zemánek, V. Adam, R. Kizek, K. Novotný, J. Kaiser, *Algal biomass analysis by laser-based analytical techniques—a review*, *Sensors* 14 (9) (2014) 17725–17752.
- [10] R. Gaudiuso, M. Dell'Aglio, O. Pascale, G. Senesi, A. De Giacomo, *Laser induced breakdown spectroscopy for elemental analysis in environmental, cultural heritage and space applications: a review of methods and results*, *Sensors* 10 (8) (2010) 7434–7468.
- [11] R. Harmon, R. Russo, R. Hark, *Applications of Laser-Induced Breakdown Spectroscopy for geochemical and environmental analysis: a comprehensive review*, *Spectrochim. Acta B At. Spectrosc.* 87 (2013) 11–26.
- [12] J. Mermet, P. Mauchien, J. Lacour, *Processing of shot-to-shot raw data to improve precision in laser-induced breakdown spectrometry microprobe*, *Spectrochim. Acta B At. Spectrosc.* 63 (10) (2008) 999–1005.
- [13] N. Zorov, A. Gorbatenko, T. Labutin, A. Popov, *A review of normalization techniques in analytical atomic spectrometry with laser sampling: from single to multivariate correction*, *Spectrochim. Acta B At. Spectrosc.* 65 (8) (2010) 642–657.
- [14] P. Yaroshchik, D. Death, S. Spencer, *Comparison of principal components regression, partial least squares regression, multi-block partial least squares regression, and serial partial least squares regression algorithms for the analysis of Fe in iron ore using LIBS*, *J. Anal. At. Spectrom.* 27 (1) (2012) 92–98.
- [15] J. Carranza, D. Hahn, *Sampling statistics and considerations for single-shot analysis using Laser-Induced Breakdown Spectroscopy*, *Spectrochim. Acta B At. Spectrosc.* 57 (4) (2002) 779–790.
- [16] V. Lazić, F. Colao, R. Fantoni, V. Spizzichino, *Laser-Induced Breakdown Spectroscopy in water: improvement of the detection threshold by signal processing*, *Spectrochim. Acta B At. Spectrosc.* 60 (7–8) (2005) 1002–1013.
- [17] J. Braga, L. Trevizan, L. Nunes, I. Rufini, D. Santos, F. Krug, *Comparison of univariate and multivariate calibration for the determination of micronutrients in pellets of plant materials by laser induced breakdown spectrometry*, *Spectrochim. Acta B At. Spectrosc.* 65 (1) (2010) 66–74.
- [18] I. Gornushkin, U. Panne, J. Winefordner, *Linear correlation for identification of materials by laser induced breakdown spectroscopy: improvement via spectral filtering and masking*, *Spectrochim. Acta B At. Spectrosc.* 64 (10) (2009) 1040–1047.
- [19] J. Sirven, B. Sallé, P. Mauchien, J. Lacour, S. Maurice, G. Manhès, *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 (12) (2007) 1471.
- [20] F. Yueh, H. Zheng, J. Singh, S. Burgess, *Preliminary evaluation of Laser-Induced Breakdown Spectroscopy for tissue classification*, *Spectrochim. Acta B At. Spectrosc.* 64 (10) (2009) 1059–1067.
- [21] T. Sahoo, A. Negi, M. Gundawar, *Study of Preprocessing Sensitivity on Laser Induced Breakdown Spectroscopy (LIBS) Spectral Classification 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI) 2015*, pp. 137–143.
- [22] P. Pořízka, A. Demidov, J. Kaiser, J. Keivanian, I. Gornushkin, U. Panne, J. Riedel, *Laser-Induced Breakdown Spectroscopy for in situ qualitative and quantitative analysis of mineral ores*, *Spectrochim. Acta B At. Spectrosc.* 101 (2014) 155–163.
- [23] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2006 (ISBN 3-900051-07-0) <http://www.R-project.org>.
- [24] L. Eriksson, T. Byrne, E. Johansson, J. Trygg and C. Vikström, *Multi- and Megavariate Data Analysis: Basic Principles and Applications*, 3rd, rev. ed. Malmö: Umetrics Academy, 2013.
- [25] K. Vanden Branden, M. Hubert, *Robust classification in high dimensions based on the SIMCA method*, *Chemom. Intell. Lab. Syst.* 79 (1–2) (2005) 10–21.
- [26] Kramida, A., Ralchenko, Yu., Reader, J., and NIST ASD Team (2015). *NIST Atomic Spectra Database (ver. 5.3)*, [Online]. Available: <http://physics.nist.gov/asd> [2016, July 21]. National Institute of Standards and Technology, Gaithersburg, MD.
- [27] J. Lasue, R. Wiens, T. Stepinski, O. Forni, S. Clegg, S. Maurice, *Nonlinear mapping technique for data visualization and clustering assessment of LIBS data: application to ChemCam data*, *Anal. Bioanal. Chem.* 400 (10) (2011) 3247–3260.
- [28] J. Sammon, *A nonlinear mapping for data structure analysis*, *IEEE Trans. Comput.* vol. -18 (5) (1969) 401–409.

## Article 15

P. Pořízka, J. Klus, A. Hrdlička, J. Vrábel, P. Škarková, D. Prochazka, J. Novotný, K. Novotný, and J. Kaiser. Impact of Laser-Induced Breakdown Spectroscopy data normalization on multivariate classification accuracy. *Journal of Analytical Atomic Spectrometry*, 32:277–288, 2017.

DOI: 10.1039/c6ja00322b; 32 citations

This is the most significant contribution from the series of publications oriented on more rigorous preprocessing of spectroscopic signal. Its scientific scope dwells in the investigation of different standardization approaches on the consecutive classification accuracy. The scientific merit of this work was enhanced by the analysis of three sample matrices (steel, aluminum alloys, and sedimentary ores) having different composition (sparsity) of typical spectra.

First, the impact of column-wise data preprocessing (mean centering, scaling, and their combination) was shown. Note that obtained spectra were manipulated in their raw form without extraction of intensity of any spectral line. As it was expected in this approach, scaling of raw spectra showed unwanted influence on the data when the significance of the background noise was unified with the significance of informative spectral variables. Therefore, it was suggested to avoid scaling of complete spectra.

Second, row-wise data preprocessing was investigated. Seven different approaches were suggested, reflecting the most typical ways of signal standardization. Naturally, the results shown that the best possible standardization has to be selected according to the sample matrix and, thus, there is no universal approach.

CrossMark  
click for updates

Cite this: DOI: 10.1039/c6ja00322b

## Impact of Laser-Induced Breakdown Spectroscopy data normalization on multivariate classification accuracy

P. Pořízka,<sup>\*abc</sup> J. Klus,<sup>abc</sup> A. Hrdlička,<sup>ac</sup> J. Vrábel,<sup>b</sup> P. Škarková,<sup>a</sup> D. Prochazka,<sup>abc</sup> J. Novotný,<sup>abc</sup> K. Novotný<sup>ac</sup> and J. Kaiser<sup>abc</sup>

Multivariate data analysis (MVDA) is getting popular across the spectroscopic community. To assess accurate results, the obtained data should be preprocessed prior to utilization of any MVDA algorithm. The process of data normalization or "internal standardization" is widely used across a broad range of applications. In this manuscript we investigate the utilization of Laser-Induced Breakdown Spectroscopy (LIBS) coupled with MVDA. However, many articles regarding the use of MVDA on data from LIBS do not provide any information about the data pretreatment. This work describes the impact of LIBS data normalization approaches on MVDA classification accuracy. Also, the impact of classical data preprocessing (mean centering and scaling) exploiting the prior utilization of MVDA was studied. This issue was investigated exploiting simple soft independent modelling of class analogies algorithm. The findings were generalized for three sample matrices (steel, Al alloys, and sedimentary ores). Furthermore, the selection of an appropriate normalization algorithm is not trivial since the spectrum of each sample matrix is composed of a different number of elements and corresponding elemental lines.

Received 30th August 2016  
Accepted 18th October 2016

DOI: 10.1039/c6ja00322b

www.rsc.org/jaas

### Introduction

Laser-Induced Breakdown Spectroscopy (LIBS)<sup>1–4</sup> is gaining its position among other analytical techniques owing to its advantages such as instrumental simplicity, capability of real-time analysis of samples in any state of matter, ability to detect a broad range of elements, and namely the possibility of *in situ* stand-off measurements. In LIBS analysis, the laser pulse is focused onto a tight spot on the sample surface (solids and liquids) or on its bulk (liquids and gases), and so a small amount of material is ablated and luminous Laser-Induced Plasma (LIP) is created. LIP radiation is then detected and spectrally resolved. The obtained LIP spectrum provides qualitative and quantitative information about the chemical composition of the sample. Each element is represented by a unique set of spectral lines and thus the detected LIP spectrum is considered to be the so-called chemical fingerprint of the investigated sample.

LIBS has already been utilized in various applications as it is reflected in a number of review articles.<sup>5–14</sup> However as any other laser-ablation based analytical technique, LIBS is strongly

affected by the matrix effect; efforts to mitigate its influence have been undertaken.<sup>15</sup> The matrix effects can be divided into processes taking place during (i) the ablation of the material and LIP formation and (ii) the spatial and temporal evolution of the LIP.<sup>4</sup> This makes the quantitative analysis challenging. On the other hand, the matrix effect is beneficial for classification of samples based on LIBS spectra differences. The issue of classification has recently gained particular attention in the LIBS community.<sup>4–6,8,11,12</sup>

Improvements in the instrumentation have led to the increased repetition rate of LIBS analysis. Laser sampling has certain advantages (no sample pretreatment, spatially and depth resolved analysis, no loss of volatile components or contamination during digestion, *etc.*) over conventional techniques, such as inductively coupled plasma after acid digestion.<sup>16</sup> However, during the pulse to pulse LIBS analysis, it is necessary to reduce any possible signal fluctuations. Several studies have already aimed at the investigation of fluctuation sources. Hahn and Omenetto<sup>4</sup> proposed the variation in laser-matter interaction (local inhomogeneity due to the chemical composition and/or physical parameters) as the main source. This leads to changes in the ablation mechanism and its efficiency influencing in turn the temperature and number density of particles in LIP. Nonetheless, the aim of this manuscript does not target deeper investigation of the origin of signal fluctuation, but only its mitigation.

The normalization strategies of LIBS signal fluctuations are of concern mainly for quantitative analysis. The individual

<sup>a</sup>CEITEC BUT, Central European Institute of Technology, Brno University of Technology, Purkyňova 123, 612 00 Brno, Czech Republic. E-mail: pavel.porizka@ceitec.vutbr.cz

<sup>b</sup>Faculty of Mechanical Engineering, Brno University of Technology, Technická 2896/2, 616 69 Brno, Czech Republic

<sup>c</sup>AtomTrace, s.r.o., Kolejní 9, 612 00 Brno, Czech Republic

approaches and their efficiency were thoroughly reviewed in other studies.<sup>4,16,17</sup> The most common normalization strategies utilize the total emission signal or the line of the matrix element. Zorov *et al.*<sup>16</sup> also dealt with an acoustic signal, emission detected by the photodiode, electric current and Mie-scattering for signal normalization. Moreover, they also proposed the utilization of multivariate normalization regarding the complex process of laser sampling. In this approach more than one reference signal is used as a correction factor for mitigation of the signal fluctuation. The correlation between the LIP signal and its acoustic signal was also studied in another study.<sup>18</sup> However, in this work, we target only straightforward and univariate approaches applicable to *in situ* analysis. This includes working with the detected optical signal and its features.

The presented work is aimed only at the classification of materials exploiting LIBS data coupled with chemometric algorithms. Therefore, the issue of multivariate quantification is not of interest due to its different nature of data processing. Despite that, the recently published manuscript by Castro and Pereira-Filho<sup>17</sup> regarding LIBS data normalization prior to univariate and multivariate quantitative analyses is noteworthy. Twelve normalization strategies were established, based namely on spectral intensity summation, matrix line standardization, or their adaptations. Their work concerns the quantification of the analyte content in steel and metallic alloys. The feasibility of multivariate classification of LIBS data was also introduced, however missed any reference to signal normalization. Similar normalization strategies were also investigated in the work of Sarkar *et al.*<sup>19</sup> on the analysis of boron isotopic composition.

The boom of chemometrics in the field of LIBS lasted over the past decade, albeit it led to the introduction of various overall methodologies in multivariate data processing. Different spectra preprocessing algorithms (normalization, outliers filtering, variable selection, *etc.*) and methods (linear and non-linear) were exploited. This has resulted in a non-unified way of chemometric utilization across the LIBS community.

LIBS technique is becoming very popular for sample classification. Broad range LIBS spectra contain a vast number of characteristic variables describing individual samples. Multivariate data analysis (MVDA) algorithms are utilized for the purpose of robust data processing. Even though MVDA algorithms are able to compensate for the signal fluctuation to a certain extent; normalization and outlier filtering strategies are still utilized *a priori*. Outlier filtering approaches include the utilization of Principal Component Analysis (PCA), linear correlation, and total energy. The effectivity of individual approaches was investigated together with comprehensive literature research in our recent publication.<sup>20</sup>

Therefore, outlier filtering was not concerned in this study. It is worth mentioning that the multivariate classification yielded the best performance after outlier filtering using the total energy approach.

Gornushkin *et al.*<sup>21</sup> found out that raw data is much more suitable for correlation (MVDA outlier filtering) than treated ones, since any artificial manipulation with data leads to loss of certain crucial information and, as a result, poorer correlation.

Other authors also utilized raw spectra that were not treated with any particular normalization or preprocessing<sup>22,23</sup> or used only averaging.<sup>24</sup> However prior to that, the data normalization approach on multivariate classification was partially provided in several studies with a positive impact on the figures of merit.

Forni *et al.*<sup>25</sup> normalized the selected and truncated spectral regions to zero mean and unit variance. This normalization enhances the relative importance of low signal spectra and allows better discrimination between the emission lines and any noise or apparatus response. Munson *et al.*<sup>26</sup> used total spectra, intensity of selected peaks or intensity ratios for the classification of bacterial specimens. Mean-centering was used prior to PCA or SIMCA in the case of total spectra; scaling was used in the case of selected spectral regions, discrete wavelengths, or line intensity ratios. The variable weighting algorithm was provided, the first principal component loading was multiplied by the original spectra to enhance the mutual differences/variability within the dataset. Weighting schemes were also exploited in the work of Hark *et al.*<sup>27</sup> Gottfried *et al.*<sup>28</sup> normalized spectra of geological origin, however the normalization approach was not specified. Afterwards the spectra were mean-centered.

To mitigate the fluctuation in the signal intensity due to the changes in laser pulse energy, the two most often utilized approaches are (i) normalization to total energy and (ii) normalization to matrix line. Larsson *et al.*<sup>29</sup> concluded that normalization of LIBS data resulted only in visual effects in the PCA score-plots with no significant effect on predictability by the PLS-DA models. However, normalization improved the predictability, provided that the amount of variables was considerably reduced. Despite that, normalization to total energy (total intensity of the particular spectrum) resulted in improvement of analytical figures of merit.<sup>30–35</sup> Normalization to matrix line was exploited in studies of Moncayo *et al.*<sup>36</sup> (Ca II 393 nm in bone samples), Koulejev<sup>37</sup> (Al I line 396 nm in the case of Al alloys and O in the case of minerals), and Yueh<sup>38</sup> (Ca I line 422 nm for normalization of tissues sample spectra). Sirven *et al.*<sup>39</sup> normalized rock spectra to intensity of O I 777 nm, since the content/line intensity is the most abundant in any rock sample/spectrum.

The overall literature research gives a non-unified methodology with diverse normalization approaches of LIBS spectra prior to MVDA analysis across different applications (sample matrices). Certain conclusions derived in particular articles are contradictory to each other; this can result in misapplied normalization strategies of LIBS data.

In this work we aim at comparing the influence of several normalization strategies on the improvement of classification performance. The selected figure of merit is the overall accuracy; as suggested by El Haddad *et al.*;<sup>5</sup> which refers to the ratio of the sum of true positives (all testing spectra correctly classified) to the total number of testing spectra. For the purpose of thorough investigation of normalization phenomenon we chose three different sample matrices (steel, Al alloys, and sedimentary ores). Those samples significantly differ in the composition of matrix and minor elements, total number of spectral variables and number of spectral variables per each element.

## Experimental

### Samples

Three sample matrices with distinct composition in spectral features were selected for the analysis; steel and cast-iron, Al alloys and sedimentary ore certified reference materials (CRMs). Each sample matrix is represented by significantly different composition of characteristic spectral features, their number and spectroscopic properties. This approach will enable complex investigation of the impact of normalization strategies on the performance of classification accuracies. Detailed experimental settings are listed in Table 1.

**(A) Steel and cast-iron samples.** Steel and cast-iron samples were collected from various producers. Sample PT-24-6 was obtained from SPL Bohumin (CZ). Various grades and types of steel (C45, E335GC, i11SMn30, i16MnCrS5, MU 100, MU 128, MU 52-60, MU 54-51, MU 57-71, MU 85-133, MU A53-81, MU B82-253, S235JRC, and S355J2C) were provided by Královopolská steel, s.r.o. (CZ). The aforementioned steel samples were measured 100 times. Steel and cast-iron samples from the first and the second LIBS round robin test organized by Federal Institute for Materials Research and Testing (BAM, GE; BAM1 C1, BAM1 C10, BAM1 C2, BAM1 C3, BAM1 C4, BAM1 C5, BAM1 C6, BAM1 C7, BAM1 C8, BAM1 C9, BAM1 S1, BAM1 S2a, BAM1 S2b, BAM2 C1, BAM2 C2, BAM2 C3, BAM2 C4, BAM2 C5, BAM2 C6, BAM2 C7, BAM2 UNKN1, BAM2 UNKN2, and BAM2 UNKN3) were measured 200 times. Unknown BAM samples were analyzed using ICP-OES after acid digestion in an external laboratory (Lithea, s.r.o., CZ), and the content of minor elements was estimated.

All steel sample surfaces were finely polished by using 3  $\mu\text{m}$  and 1  $\mu\text{m}$  diamond paste (Urdiamant, CZ) prior to LIBS analysis. The sample surface was wiped with ethanol after the diamond paste polishing. The sample surface was polished in order to reduce any potential fluctuations in the LIP signal due to the surface roughness. The samples from the second round robin test were delivered in the form of powders and chunks and thus analyzed as deposited on the sticky tape. The samples were moved prior to each consecutive laser pulse to provide a fresh spot for analysis.

**(B) Al alloys CRMs.** Gleich aluminium standards (AW2017, AW2030, AW6082, and AW7075), BAM standards (GE; BAM 308,

BAM 310, BAM 311, EB313, EB316, and EB317), and MBH standards (MBH Analytical Ltd., UK; MBH E2 7, MBH E5 7, and MBH G12 H5) were measured. The sample surfaces were finely polished by using 3  $\mu\text{m}$  and 1  $\mu\text{m}$  diamond paste (Urdiamant, CZ) prior to LIBS analysis. The sample surface was wiped with ethanol after the diamond paste polishing. The sample surface was polished in order to reduce any potential fluctuations in the LIP signal due to the surface roughness. Each sample was measured 200 times, 1 laser pulse per spot.

**(C) Sedimentary ores CMRs.** 50 sedimentary ores samples (OREAS, AU; OREAS 100a, OREAS 101a, OREAS 132a, OREAS 132b, OREAS 133a, OREAS 133b, OREAS 134a, OREAS 134b, OREAS 13b, OREAS 140, OREAS 141, OREAS 142, OREAS 14P, OREAS 151b, OREAS 152b, OREAS 153b, OREAS 170a, OREAS 170b, OREAS 172, OREAS 201, OREAS 202, OREAS 203, OREAS 250, OREAS 251, OREAS 252, OREAS 36, OREAS 38, OREAS 401, OREAS 402, OREAS 405, OREAS 406, OREAS 45d, OREAS 45e, OREAS 601, OREAS 602, OREAS 603, OREAS 604, OREAS 605, OREAS 700, and OREAS 701) were delivered in the form of fine powder and then pressed into pellets with a manual hydraulic press. The matrices of this sample set differ and thus also the composition of individual matrix elements varies. Each sample was represented by a single pellet, which was measured 100 times, 1 laser pulse per spot and 100 spots per sample giving 100 representative spectra per sample.

### LIBS system

The samples were analyzed using the Sci-Trace LIBS system (AtomTrace, CZ) including the LIBS interaction chamber (AtomTrace, CZ), the chamber was thoroughly described elsewhere.<sup>40</sup> In the Sci-Trace device, the laser pulse (10 ns, 532 nm; CFR 400, Quantel, FR) is led to the interaction chamber by using a series of mirrors (NB1-K13, ThorLabs, US) and focused onto a tight spot in the interaction region on the sample surface by using a triplet lens (25.4 mm focal length, Sill Optics, GE). The laser is delivered collinearly with the surface normal, the collection optics axis is positioned under the angle of 60° to the surface normal. The LIP radiation is collected by using an objective (100 mm CaF<sub>2</sub> and 75 mm UVFS lenses; ThorLabs, US). Collected light is introduced *via* the optical fiber (400 microns core diameter; Thorlabs, US) on the entrance slit of an echelle spectrometer (200–900 nm wavelength range; EMU 65, Catalina Scientific, US). The light is spectrally resolved and imaged with an EMCCD detector (Falcon Blue, Raptor Photonic, IR). Measurement settings for each matrix are given in Table 1, these were kept constant throughout the whole analysis of each sample set. The AtomChamber software (AtomTrace, CZ) was used for system operation. The LIBS system was in each case preliminary optimized to the best Signal-to-Noise Ratio (SNR). However, it was already preliminary proved that there is no particular need for system optimization in the case of multivariate classification.

### Data processing and methods

Detected spectra were preliminary processed using AtomAnalyzer (AtomTrace, CZ) software, then for multivariate

Table 1 LIBS system parameters

Sample matrix	Steel CRMs	Al alloy CRMs	Sedimentary ore CRMs
Laser energy [mJ]	50	50	50
Spot size [ $\mu\text{m}$ ]	~100	~100	~100
Irradiance [ $\text{GW cm}^{-2}$ ]	~64	~64	~64
Step [ $\mu\text{m}$ ]	200	200	200
Gate delay [ns]	1000	1500	500
Gate width [ $\mu\text{s}$ ]	50	50	50
Accumulations per spot	1	1	1
No. of spectra per sample	100 or 200	200	100
Number of samples	38	13	40
Total no. of spectra	7100	2600	4000

analysis and plotting the figures the free software R was used, namely libraries 'rccovHD', 'caret', and 'ggplot'. The computation was run on a standard personal computer with 16 GB RAM.

The spectrometer and the detector provide wide range spectra (from 200 to 900 nm). However the intensity detected under 250 nm (spectral efficiency) rapidly vanishes and over 700 nm there are no significant spectral features usable for classification, except potassium atomic lines at 766.49 and 769.89 nm. Beyond this limit, there are mainly oxygen and nitrogen lines but they were also sacrificed regardless of the sample matrix. Therefore, the spectra were truncated prior to any further processing to a final range of 250 to 700 nm. Another reason was also lower dimensions of data matrices and thus less demanding computation.

Samples of each matrix were measured separately and then each case study (matrix) was naturally processed individually. Each sample was represented by a set of 100 or 200 spectra. Half of the spectra per sample were randomly selected (using a random number generator) and assigned for training set, the rest of the spectra per sample were saved for testing purposes. Two data matrices (training and testing) were organized, where rows were individual measurements/spectra and columns were variables (intensities on respective wavelengths). Samples in both data matrices were organized under each other and named, as given above in the Samples section, for classification purposes. El Haddad *et al.*<sup>5</sup> suggested to divide the original dataset into three sub-data matrices (training, validation, and testing). However in our work we do not aim at the validation of optimal number of Principal Components (PCs) prior to classification. On the other hand, we varied the number of PCs from 1 up to 20 (or to 10 respectively, see further in the text) to visualize the trend in yielded classification accuracy.

The whole spectra sets of each matrix were visualized using the PCA.<sup>41</sup> This MVDA algorithm is based on the least squares approach, where only the most important information is used. PCA is effective for data visualization where multivariate information is projected to newly constructed, low-dimensional space described by PCs. Each PC consists of scores (data points representing original spectra) and loadings (rotation of original data space to newly established one). PCA describes the total variability of the entire dataset, it does not distinguish between group-to-group variability and within-group variability.<sup>6</sup> It works well when the variability within the group or sample is much smaller than the variability among the samples.

Soft Independent Modelling of Class Analogies (SIMCA)<sup>41</sup> was used for supervised classification of the sample sets. In SIMCA, PCA is applied to training data of each sample separately. Model is then considered as series of PCA models. Spectra from the test set are one-by-one applied to the model and distance to each PCA model is estimated. The testing spectrum is assigned to the samples with the smallest distance to the center of its PCA model. The number of principal components used for establishing the model was varied. SIMCA was selected because of its simplicity, thus its performance is more influenced by the topology and fluctuation in the data than in the case of other linear or even non-linear MVDA algorithms. It is also expectable that the utilization of another more

advanced algorithm (based for instance on neural networks) could lead to higher accuracies.<sup>5</sup>

Raw spectra organized in the data matrices were treated with various approaches in row-wise (over spectra individually) or column-wise (over variables individually) manner. Firstly, column-wise (applied on each variable/wavelength individually) data preprocessing (mean centering and/or scaling) was studied, since they are recommended data preprocessing approaches prior to MVDA in the literature.<sup>41</sup> In mean centering, the mean value is subtracted from each column. This moves the center of mass to the origin of the coordinate system. In scaling, values in each column are divided by their respective standard deviation.

Secondly, row-wise (applied on each spectrum individually) signal normalization was provided. The normalization strategies were selected with respect to *in situ* applications. We have selected several most often utilized normalization strategies. In the total energy case, spectral variables were divided by the sum of all intensities detected in the respective spectrum. A similar approach is in the maximum peak intensity when the spectral features are divided by the maximum peak value, the intensity integral in the whole range of the peak was not considered. [0,1] data scaling approach was obtained when the minimal value was subtracted from the spectrum and then this spectrum was divided by its maximum peak intensity. For matrix line normalization, several matrix lines per spectrum were selected and the respective spectra were divided by this value. We also utilized row-wise mean centering and scaling, provided in the same manner as in the column-wise case. Row-wise scaling equalizes to a certain extent the differences in total intensities of individual spectra.

The methodology for spectral line intensity estimation, as described elsewhere,<sup>42</sup> was utilized, if the intensity of any spectral line was necessary. The intensity of the line was considered as the sum of intensities in the range of this spectral line with background subtraction. The level of background was estimated in a close proximity to the selected spectral line.

## Results and discussion

### Data projection

In this work we investigated the influence of (i) column-wise variable mean centering/scaling and (ii) row-wise spectra normalization. However, firstly the topology of LIBS data of individual matrices was checked. For these purposes training spectra (column-wise mean centered) of all samples were investigated using PCA scores and loadings. Processed data were visualized in a truncated PC space using first two principal components. This simple step enables visualization of similarities and differences among individual sample sets and shows if they overlap or if they are discriminated from each other. The distribution of steel and cast-iron samples in the PC space (PC 1 = 20.2%, PC 2 = 9.2%, and PC 3 = 3.9% of total variability in the data) shows distinct separation of certain samples, Fig. 1a. The data points of each sample are ordered in compact clusters. Despite the bigger number of samples and more complicated resolution of the color scale, the samples are well distinguished

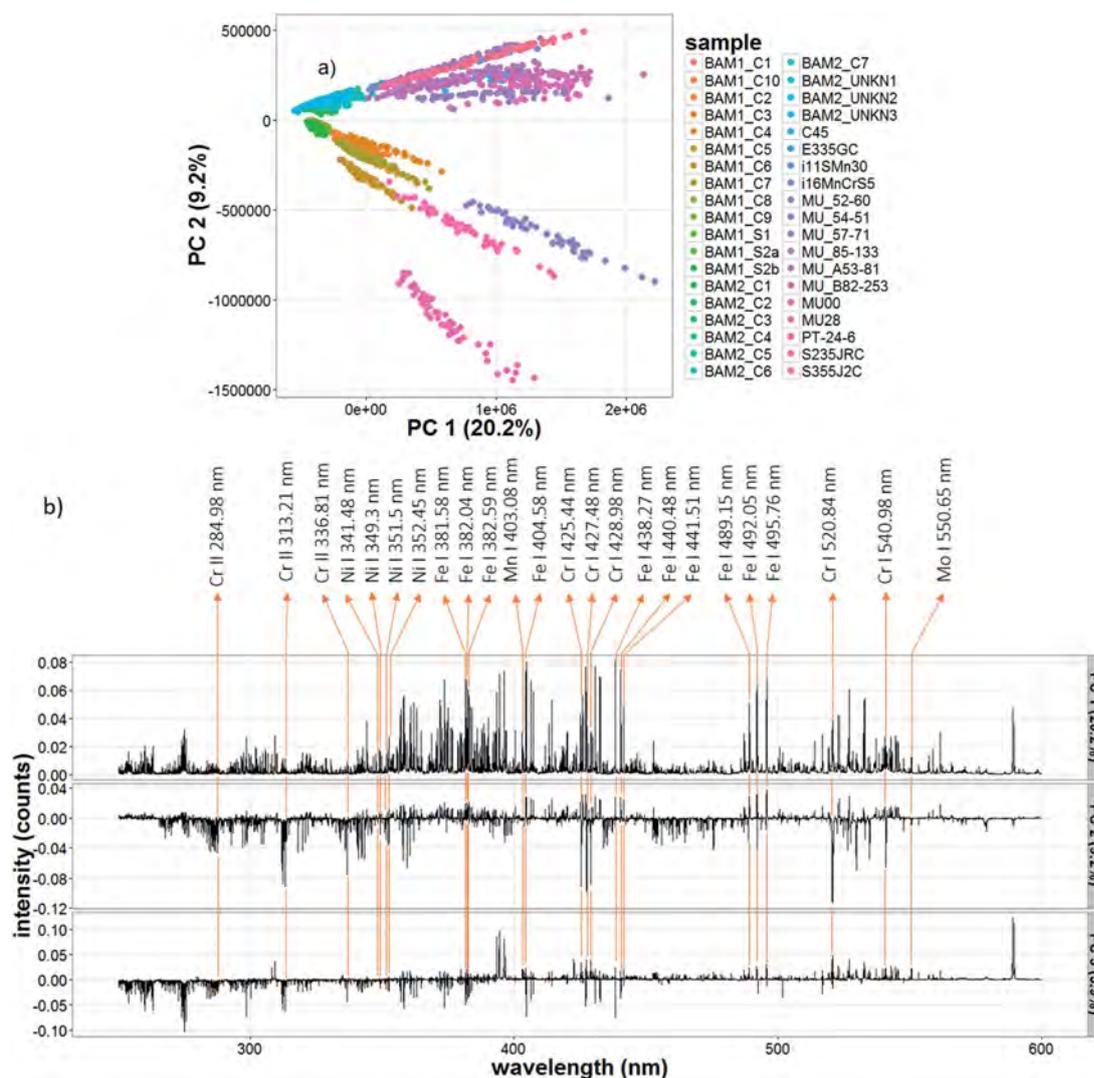


Fig. 1 PCA analysis yielded from steel and cast-iron samples LIBS data: (a) scores and (b) loadings.

in the figure. The elongated shape of sample clusters is attributed to the fluctuation in spectral intensities, *i.e.* fluctuation of LIP and its properties. Samples are rather distinguished what could suggest easier classification. In the case of steels and cast-irons, a high number of detected lines are apparent. PCA loadings are composed mainly from lines of major elements (Fe, Cr, and Ni), Fig. 1b. The majority of lines was attributed namely to Fe; ionic lines below 280 nm and atomic lines in the range of 340 to 400 nm. Lines of Co II at 241.63 nm (not highlighted in the figure), Mn I 403.08 nm, Mo I 550.65, and Si I lines 288.15 nm and 390.55 nm were also detected, but with lower loading values due to lower concentration ranges in the samples. The detailed list of lines is given in the loading figure.

The PCA scores of Al alloy CRMs are plotted in Fig. 2a showing more overlapping clusters in the newly constructed PC space (PC 1 = 13.7%, PC 2 = 4.5%, and PC 3 = 1.8% of total variability in the data). The individual PCs describe a less amount of variability in the data as in the case of steel and cast-iron samples. This could lead to more confusing classification and moderate figures of merit. PC loadings of Al alloys, Fig. 2b, are dominated by the Al doublet at 394.4 and 396.15 nm, which overshadows the rest of the loadings. However, lines of minor and trace elements (Cu, Fe, Mg, Pb, Si, Ti, and Zn) were found contributing to the classification less significantly. Lines of Cr and Ni were not detected. The list of lines is given in Fig. 2b.

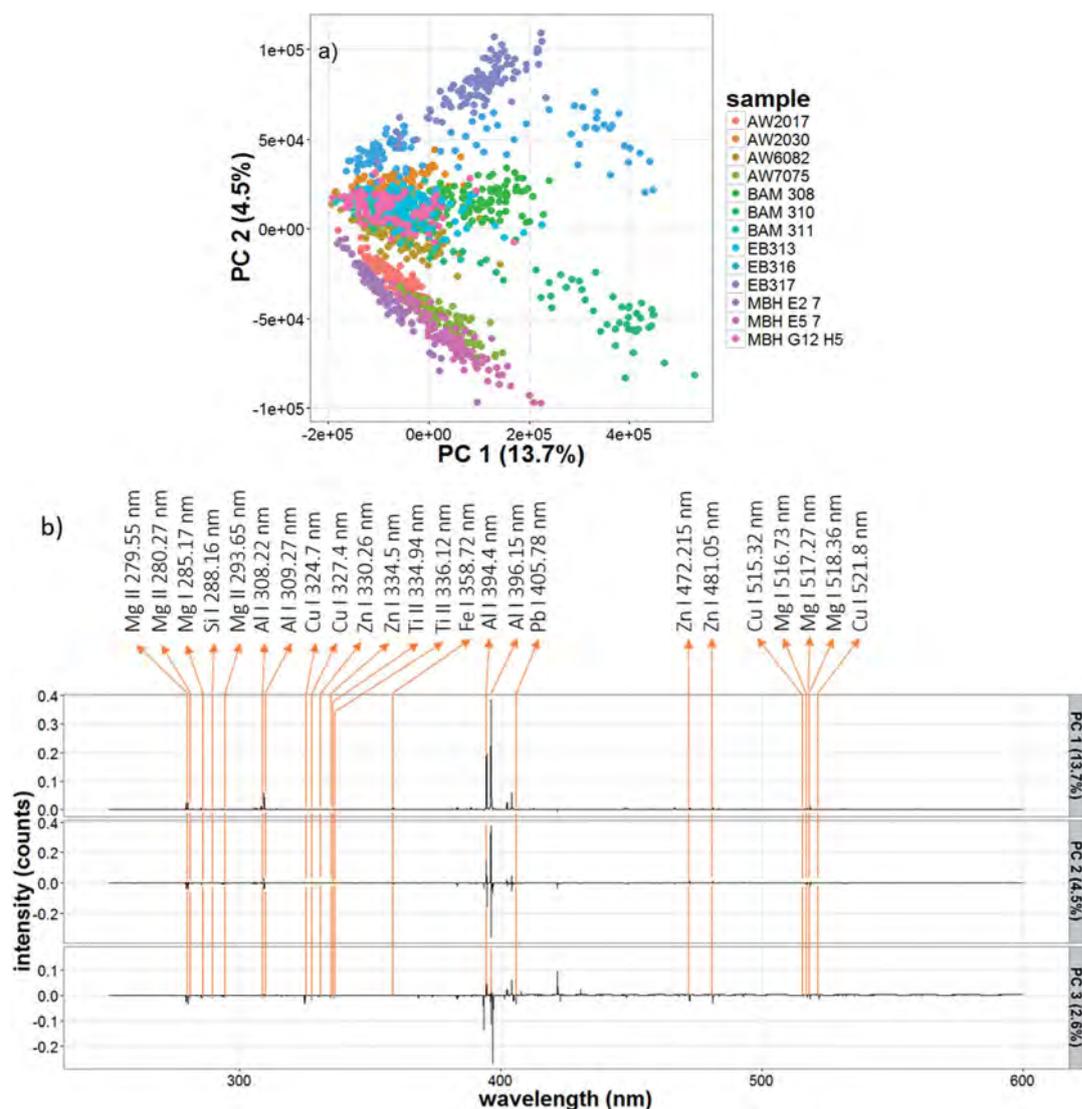


Fig. 2 PCA analysis yielded from Al alloy CRM LIBS data: (a) scores and (b) loadings.

The clustering of sedimentary ore CRMs is depicted in the space of first two PCs, see Fig. 3a. The elongated shape of clusters and outliers are also present as in the previous cases. The color scale is hard to read due to the high number of samples. However, distinct clustering of certain samples from the rest of the dataset is clearly visible. First three principal components carry 21.2% of total variability (PC 1 = 8.4%, PC 2 = 6.6%, and PC 3 = 6.2%). Lower differences in percentages of individual PCs suggest a more complex dependence of loadings on various variables (elemental lines). Loadings of first three PCs are shown in Fig. 3b, giving also the list of most significant lines (Na, Ca, Al, Fe, Mn, and Cr) or lines of matrix elements (Ca, Na, Al, Mg, and

Si). The highest loading values possess the sodium doublet at 589.00 and 589.59 nm and the doublet of calcium ionic lines at 393.37 and 396.85 nm. Lines with considerable high loading values were also highlighted in the figure; Mn I 403.08, 403.31, 403.45, and 403.57, Fe I lines 427.18, 430.79, 432.58, and 438.35 nm (more lines of Fe are present in the spectra but not highlighted), and Cu I t 324.7 and 327.4 nm. Lines of Mg I 382.94, 383.23, and 383.83 nm, Zn I 472.215 and 481.05 nm, and Cr I 425.44 and 427.48 nm were less significant. Potassium, another matrix element, having an intense doublet at 766.49 and 769.89 nm was not utilized in classification due to the limited spectral window truncated for less demanding computation. In the case

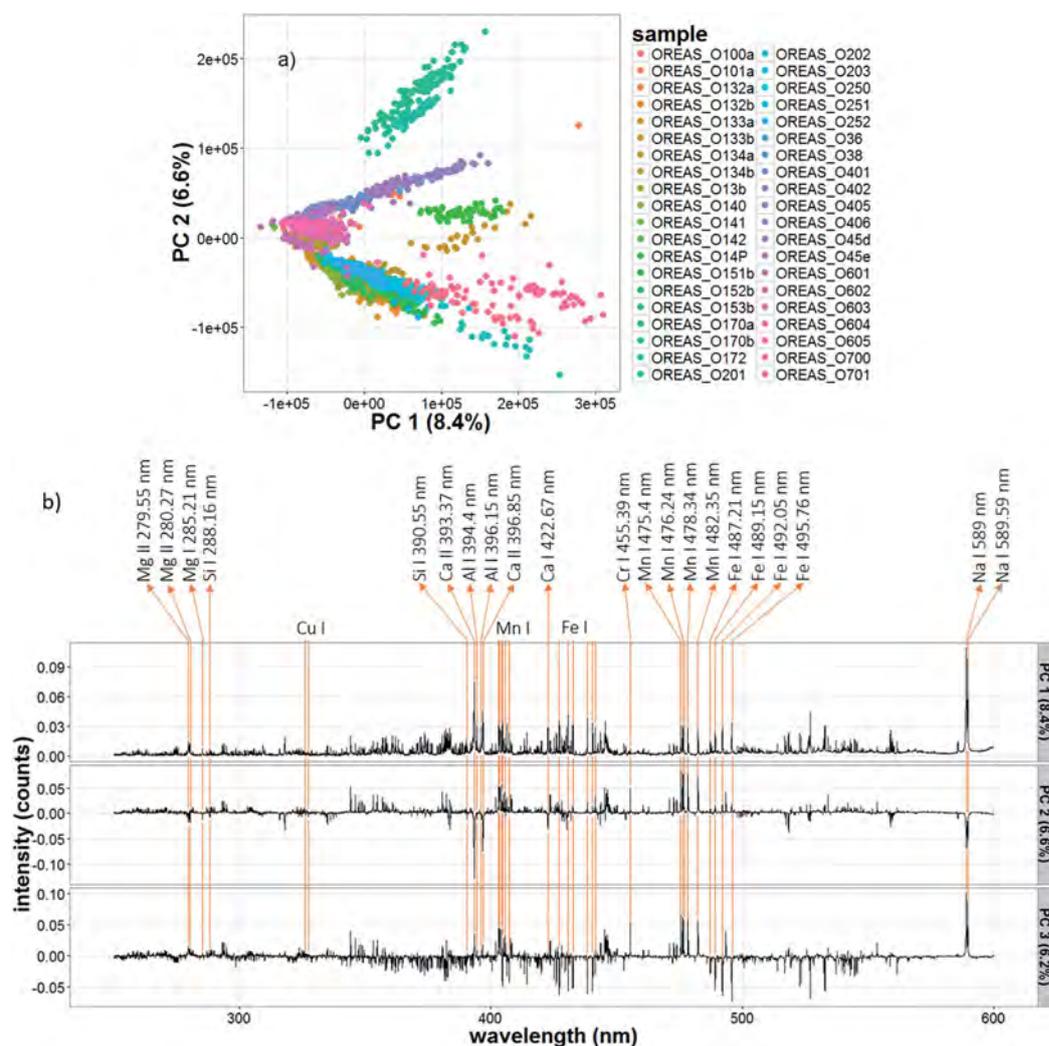


Fig. 3 PCA analysis yielded from sedimentary ore CRM LIBS data: (a) scores and (b) loadings.

of sedimentary ores, no clear matrix element is present, except for sodium and calcium thanks to their relatively high loading values. The introduction of more matrix elements (with a number of elemental lines and strongly varying quantum properties) makes the data normalization challenging.

The wavelength range depicted in loading images was truncated up to 600 nm for better readability. It is worth mentioning that lines of  $H_z$  656 nm, O I 777 nm, nitrogen triplet (Ni I 742.364, 744.229, and 746.831 nm), Ca (ionic 393.37 and 396.85 nm, and atomic at 422.67 nm) and Na I lines (589 and 589.59 nm) are present in each spectrum with certain intensity, regardless of their content in the sample composition. The atmospheric elements are detected due to ablation of ambient air surrounding the sample surface. Ca and Na signals can be observed in any spectrum,

resulting from any possible surface contamination during the handling of the sample. All observed and abovementioned lines were cross-checked with the NIST online database.<sup>43</sup>

Each sample, regardless of the matrix, contains several data points outlying from the rest of the respective cluster in the PC space. This strengthens the need for outlier filtering prior to any further data processing. However, this was not provided since omitting outliers from the dataset would lead to significant improvement in the classification accuracy, as proved in our recent manuscript.<sup>20</sup> Filtering of the outliers changes the topology of the data space and creates more distinct and less overlapping clusters. This was unwanted in our present research since potentially outlying data points were welcomed for the possibility of classification confusion.

Individual sample matrices are representing various compositions of major, minor, and trace elements. Those elements are characterized by different numbers of emission lines. In general, the variation in elemental composition characterized by a lower number of spectral lines is overshadowed by the variation of elements with more spectral lines. The amount of variation described by each variable has a great impact on the clustering/classification. As presented in loading figures of all samples, the major/matrix elements possess the highest loading values. Thus the changes in matrix lines are chief contributors to the classification of data.

### Column-wise data preprocessing

At first, the impact of column-wise (over each variable individually) data preprocessing (mean centering and/or scaling) of raw data (without any normalization) on the classification accuracy was investigated, see Fig. 4. In general, the accuracy is increasing with the number of PCs, varied from 1 to 20. This trend is valid to a certain point where the accuracy levels off and the increase is insignificant for higher number of PCs. Regardless of the sample matrix there is an optimal number of 3 to 5 principal components. It is an interesting phenomenon that classification of raw data, mean centered data and scaled data yields similar classification accuracies over the whole range of PCs. However, the utilization of mean centering and scaling results in significantly lower accuracies. The worst performance was obtained in the case of Al alloy CRMs.

The Al alloy case was selected, because of its poorest performance, to describe the significant classification confusion after mean centering with scaling preprocessing. Loadings estimated from the PCA analyses of preprocessed LIBS data are depicted in Fig. 5; (a) raw data, (b) mean centered data, (c) scaled data, and (d) mean centered and scaled data. The total amount of variation described by the first three principal components is (a) 33.9%, (b) 20.8%, (c) 5.4%, and (d) 2.5%. However as presented above, the decreasing trend has no impact on the classification, except the case of mean centering together with scaling. Scaling of data (raw or mean centered) introduces an unwanted increase in significance of loadings attributed to noise, Fig. 5c and d. After this data preprocessing approach the orders of echelle spectrograph are highlighted in the loadings, as depicted in Fig. 5d. After mean centering and scaling the persisting valuable information vanishes and LIBS spectra are useless for any classification.

Therefore, we do not recommend using scaling for data processing regardless of the matrix to be analyzed when typical full range LIBS spectra are of interest. This algorithm simply equalizes the importance of all variables (including noise). Spectral lines should be firstly fitted and treated separately. Data matrix composed only from the line emission signal should be constructed prior to scaling.

### Row-wise data normalization

Based on the aforementioned results, no column-wise data preprocessing (raw data) was chosen for further analysis. Raw data of each sample matrix were normalized with *a priori*

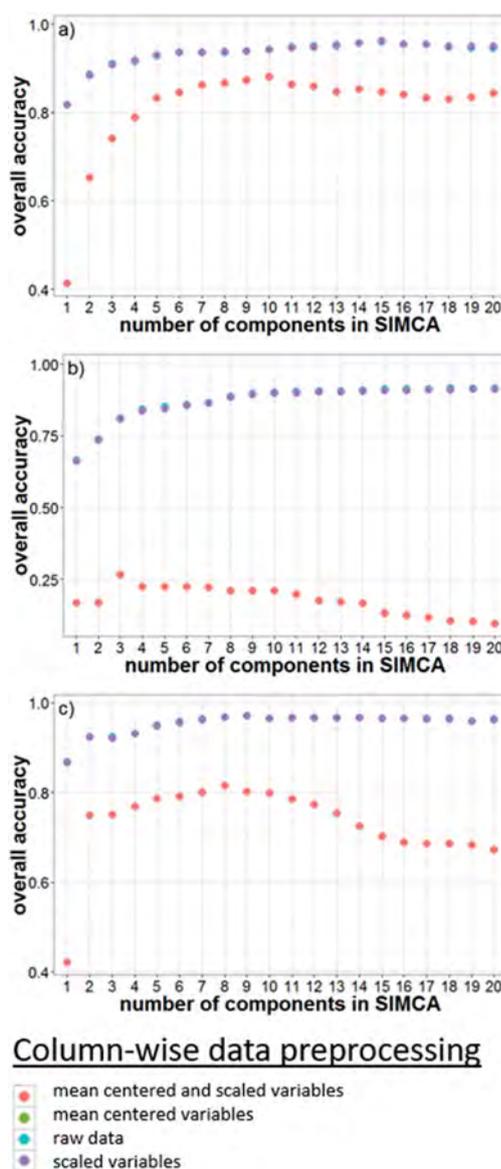


Fig. 4 Impact of column-wise data preprocessing (scaling and mean centering) on the classification accuracy of (a) steels and cast-irons, (b) Al alloy CRMs, and (c) sedimentary ore CRM LIBS data. Accuracies obtained from raw, scaled, and mean centered data are partly overlapping in the figure.

selected normalization approaches: (i) total intensity, (ii) normalization to maximum peak intensity, (iii) [0,1] normalization, (iv) mean centering, (v) scaling, and (vi) normalization to intensity of selected matrix element spectral line. SIMCA was applied on the data after row-wise normalization. Overall accuracy estimated from the obtained confusion matrix was

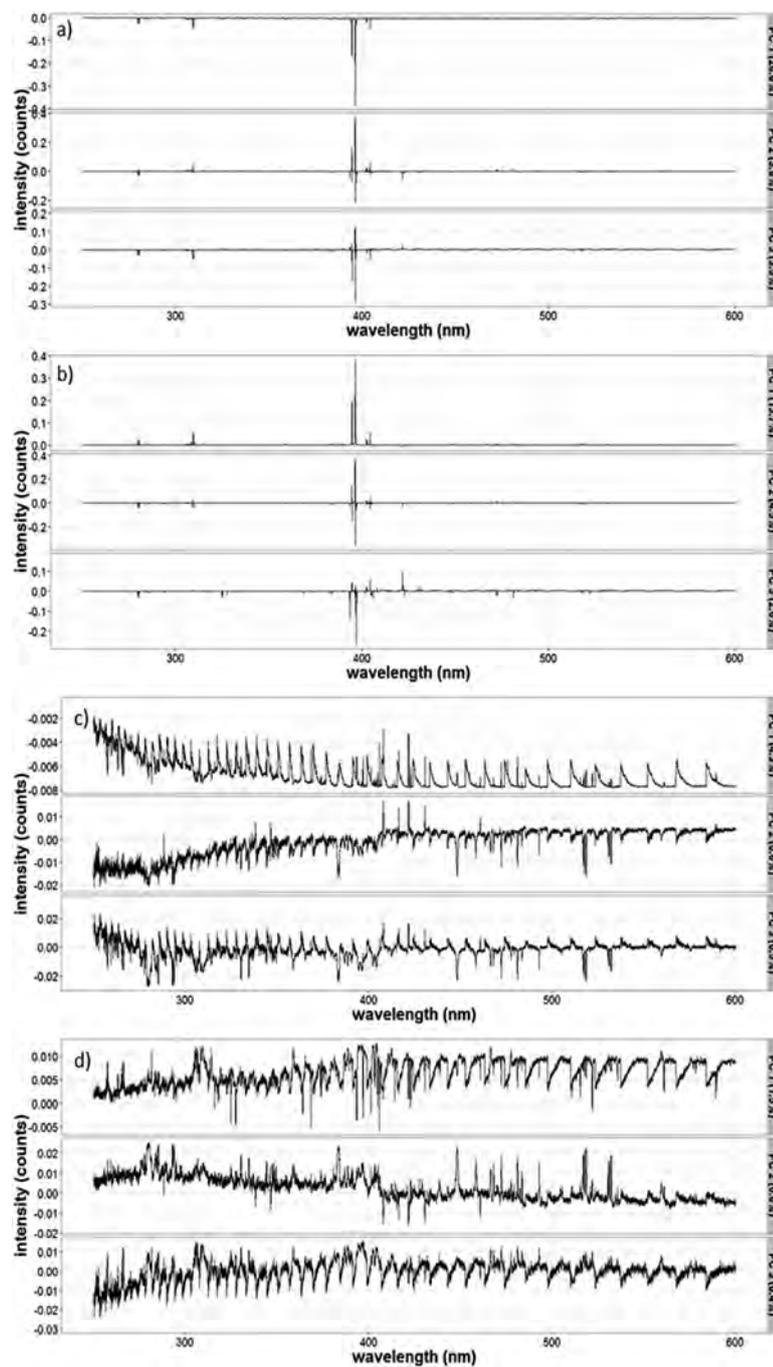


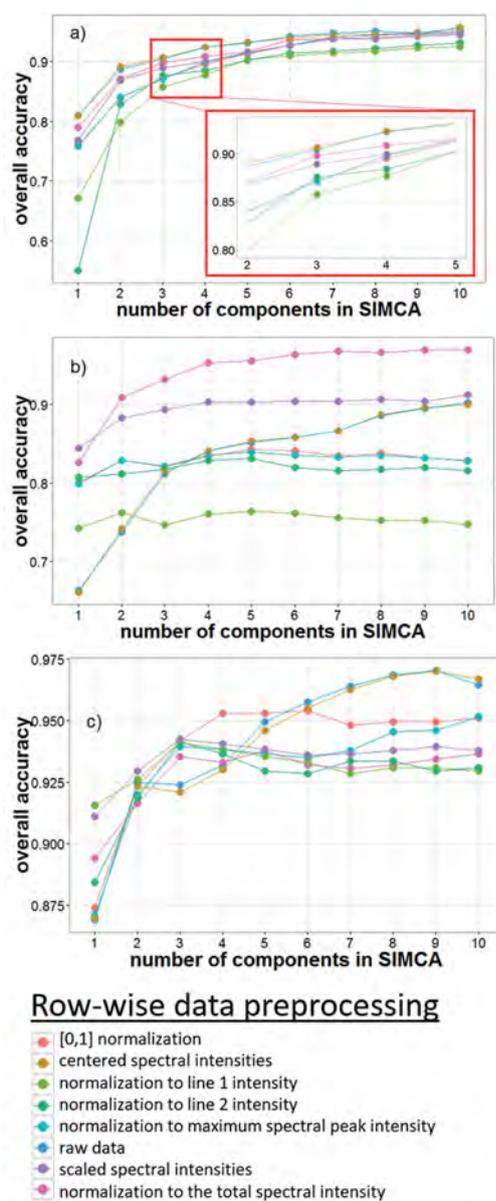
Fig. 5 PCA loadings yielded from Al alloy CRM LIBS data after (a) no data preprocessing, (b) mean centering, (c) scaling and (d) mean centering and scaling.

considered for comparison of data preprocessing impact on classification. The range of PCs used in SIMCA was limited to 10. From this point further no significant increase was obtained in previous analyses, Fig. 4. Thus no increase was expected even after row-wise normalization (each spectrum treated individually).

In the case of steel and cast-iron samples, good classification accuracy was predicted from the distinct clustering of data in PC space, Fig. 1a. Only a minimal difference between individual normalization approaches is evident, see Fig. 6a. The trend is also similar for all approaches, suggesting the optimal number of 3 or 4 PCs. This region is highlighted in the figure. The best performance was yielded for classification of row-wise mean centered and raw data. Normalization to the total intensity of the spectrum is also worth mentioning showing reasonable accuracies with respect to the other matrices. The worst performance was obtained for classification of data normalized to selected spectral line intensities. This result is in disagreement with the general approach used for the improvement of conventional quantitative analysis. In spite of the fact that iron is the major element, such normalization demands higher quantum states of both spectral transitions (matrix line used for normalization and line to be normalized) to be similar. However, this condition cannot be naturally met for all the lines present in the spectra.

Classification accuracies of aluminum alloy spectra after individual normalization approaches are presented in Fig. 6b. Significantly different impact of each normalization approach was obtained. This is a result of the nature of Al alloy spectra, where a low number of elemental lines (signal) *versus* total number of variables (including noise) are present. Moreover, the data points are more overlapped in the PC space, shown in Fig. 2a, suggesting more challenging classification. This sample matrix is therefore more sensitive to data treatment. The best performance was yielded for classification of data after total intensity normalization. Considerably high classification accuracy is also shown in row-wise scaling normalization, what is in contrast to column-wise scaling results. Once more the worst results were obtained for the normalization using matrix line intensities, attributed to the phenomenon described above.

Classification accuracies after individual normalization preprocessing of sedimentary ore LIBS data are depicted in Fig. 6c. For most of the cases, the trend levels off between three to four PCs. However, the trend for raw data and mean centered data levels off firstly with 2 PCs and then rises and reaches the maximum accuracy for 9 PCs. Omitting those normalization methods from the investigation, the best performance is obtained for [0,1] normalization. The rest of selected normalization methods show moderate figures of merit, overreaching 92.5% of overall accuracy. The matrix lines selected were Ca I line 422.7 nm and Na I 589 nm. Those atomic lines showed high loading values and thus were considered as important for classification purposes. Their performance in data normalization shows good accuracy when three PCs are considered. For more PCs, the accuracy decreases and normalization to line intensity gets ineffective.



	a) steel	b) Al alloys	c) ores
line 1	Fe I at 372 nm	Al I at 309 nm	Ca I at 422.7 nm
line 2	Fe I 404.58 nm	Al I at 396 nm	Na I at 589 nm

Fig. 6 Impact of row-wise LIBS data normalization on overall classification accuracy estimated using SIMCA for (a) steels and cast-irons, (b) Al alloys CRMs, and (c) sedimentary ore CRMs.

Regardless of the sample matrix and number of PCs, utilization of raw data and row-wise mean centered data gives almost the same accuracies. For LIBS users it can be work saving

to know that the normalization to any single line does not lead to better classification efficiency. Taking more lines or total spectral intensity can be beneficial or comparable with analysis of raw data.

## Conclusion and ongoing work

In this manuscript we assessed the impact of row-wise and column-wise preprocessing of the data matrix prior to multivariate classification, utilizing SIMCA. Different normalization approaches were chosen, with the aim of their exploitation for LIBS *in situ* analysis. Data were provided by the LIBS analysis of various sample matrices (steel and cast-iron, Al alloy, and sedimentary ores) representing different numbers of matrix and minor elements as well as significantly different numbers of detected emission lines. Thus, the presented results may be accepted in a general manner, since they are not aimed only at one case study. The estimation of number of PCs used in MVDA algorithms is also of great issue. The presented results show the best performance of SIMCA with 3 to 5 PCs in spite of the sample matrix and data preprocessing approach.

Column-wise scaling of raw or mean centered data is not advised when classical LIBS spectra (including signal and noise) are under investigation. This approach equalizes the impact of noise with the signal on classification accuracy, as proved by loading plots. For scaling, spectral line intensities should be firstly estimated and organized into a data matrix without contribution of variables carrying background noise. Spectral line selection and their utilization in MVDA algorithms (for instance impact of differing the number of lines per sample) are behind the scope of this manuscript and will be addressed in the ongoing work.

Row-wise data preprocessing, *i.e.* normalization of individual spectra, has to be selected according to the sample matrix. Spectra are composed from elements represented by various numbers of spectral lines. Moreover, each spectral line has a different impact on classification, as shown in loadings plots. In general, normalization to matrix line intensity brings no improvement and is not recommended regardless of the sample matrix.

## Acknowledgements

We acknowledge that this research was carried out under the project CEITEC 2020 (LQ1601) with financial support from the Ministry of Education, Youth and Sports of the Czech Republic under the National Sustainability Programme II. P. P. and J. Ka. would like to acknowledge that this work is also an output of GACR project 15-23274S: Design of advanced materials using selective laser melting. J. Ka. acknowledges the support of Brno University of Technology on the frame of grant FSI-S-14-2494 (Application of Advanced Optical Methods).

## References

- 1 A. Miziolek, V. Palleschi and I. Schechter, *Laser Induced Breakdown Spectroscopy*, Cambridge University Press, Cambridge, 2006.
- 2 R. Noll, *Laser-induced Breakdown Spectroscopy Fundamentals and Applications*, Springer-Verlag, Berlin, Heidelberg, 2012.
- 3 D. Hahn and N. Omenetto, Laser-Induced Breakdown Spectroscopy (LIBS), Part I: Review of Basic Diagnostics and Plasma-Particle Interactions, *Appl. Spectrosc.*, 2010, **64**(12), 335–366.
- 4 D. Hahn and 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.*, 2012, **66**(4), 347–419.
- 5 J. El Haddad, L. Canioni and B. Bousquet, Good practices in LIBS analysis: Review and advices, *Spectrochim. Acta, Part B*, 2014, **101**, 171–182.
- 6 J. L. Gottfried, F. C. De Lucia, C. A. Munson and A. W. Miziolek, Laser-induced breakdown spectroscopy for detection of explosives residues: a review of recent advances, challenges, and future prospects, *Anal. Bioanal. Chem.*, 2009, **395**(2), 283–300.
- 7 J. J. Fortes and J. J. Laserna, The development of fieldable laser-induced breakdown spectrometer: No limits on the horizon, *Spectrochim. Acta, Part B*, 2010, **65**(12), 975–990.
- 8 S. J. Rehse, H. Salimnia and A. W. Miziolek, Laser-induced breakdown spectroscopy (LIBS): an overview of recent progress and future potential for biomedical applications, *J. Med. Eng. Technol.*, 2012, **36**(2), 77–89.
- 9 D. Santos, L. C. Nunes, G. G. A. De Carvalho, M. d. S. Gomes, P. F. De Souza, F. d. O. Leme, L. G. C. Dos Santos and F. J. Krug, Laser-induced breakdown spectroscopy for analysis of plant materials: A review, *Spectrochim. Acta, Part B*, 2012, **71**–72, 3–13.
- 10 J. Kaiser, K. Nnovotný, M. Z. Martin, A. Hrdlička, R. Malina, M. Hartl, V. Adam and R. Kizek, Trace elemental analysis by laser-induced breakdown spectroscopy—Biological applications, *Surf. Sci. Rep.*, 2012, **67**(11–12), 233–243.
- 11 R. S. Harmon, R. E. Russo and R. R. Hark, Applications of laser-induced breakdown spectroscopy for geochemical and environmental analysis: A comprehensive review, *Spectrochim. Acta, Part B*, 2013, **87**, 11–26.
- 12 P. Pořízka, P. Prochazková, D. Prochazka, *et al.*, Algal Biomass Analysis by Laser-Based Analytical Techniques—A Review, *Sensors*, 2014, **14**(9), 17725–17752.
- 13 R. Noll, C. Fricke-Begemann, M. Brunk, *et al.*, Laser-induced breakdown spectroscopy expands into industrial applications, *Spectrochim. Acta, Part B*, 2014, **93**, 41–51.
- 14 J. Rakovský, P. Čermák, O. Musset and P. Veis, A review of the development of portable laser induced breakdown spectroscopy and its applications, *Spectrochim. Acta, Part B*, 2014, **101**, 269–287.
- 15 B. C. Windom and D. W. Hahn, Laser ablation—laser induced breakdown spectroscopy (LA-LIBS): A means for overcoming matrix effects leading to improved analyte response, *J. Anal. At. Spectrom.*, 2009, **24**(12), 1665–1675.
- 16 N. B. Zorov, A. A. Gobratenko, T. A. Labutin and A. M. Popov, A review of normalization techniques in analytical atomic spectrometry with laser sampling: From single to multivariate correction, *Spectrochim. Acta, Part B*, 2010, **65**(8), 642–657.

- 17 J. P. Castro and E. R. 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.*, 2016, **31**, 2005–2014.
- 18 A. Hrdlička, L. Zaorálková, M. Galiová, *et al.*, Correlation of acoustic and optical emission signals produced at 1064 and 532 nm laser-induced breakdown spectroscopy (LIBS) of glazed wall tiles, *Spectrochim. Acta, Part B*, 2009, **64**(1), 74–78.
- 19 A. Sarkar, X. Mao and R. E. Russo, Advancing the analytical capabilities of laser ablation molecular isotopic spectrometry for boron isotopic analysis, *Spectrochim. Acta, Part B*, 2014, **92**, 42–50.
- 20 P. Pořízka, J. Klus, D. Prochazka, E. Képeš, A. Hrdlička, J. Novotný, K. Novotný and J. Kaiser, Laser-Induced Breakdown Spectroscopy coupled with chemometrics for the analysis of steel: The issue of spectral outliers filtering, *Spectrochim. Acta, Part B*, 2016, **123**, 114–120.
- 21 I. B. Gornushkin, B. W. Smith, H. Nasajpour and J. D. Winefordner, Identification of Solid Materials by Correlation Analysis Using a Microscopic Laser-Induced Plasma Spectrometer, *Anal. Chem.*, 1999, **71**(22), 5157–5164.
- 22 R. A. Multari, D. A. Cremers, J. M. Dupre and J. E. Gustafson, The Use of Laser-Induced Breakdown Spectroscopy for Distinguishing Between Bacterial Pathogen Species and Strains, *Appl. Spectrosc.*, 2010, **64**, 750–759.
- 23 X. Zhu, T. Xu, Q. Lin, *et al.*, Advanced statistical analysis of laser-induced breakdown spectroscopy data to discriminate sedimentary rocks based on Czerny–Turner and Echelle spectrometers, *Spectrochim. Acta, Part B*, 2014, **93**, 8–13.
- 24 M. Z. Martin, N. Labbé, T. G. Rials and S. D. Wullschleger, Analysis of preservative-treated wood by multivariate analysis of laser-induced breakdown spectroscopy spectra, *Spectrochim. Acta, Part B*, 2005, **60**(7–8), 1179–1185.
- 25 O. Forni, S. Maurice, O. Gasnault, R. C. Wiens, A. Cousin, S. M. Clegg, J.-B. Sirven and J. Lasue, Independent component analysis classification of laser induced breakdown spectroscopy spectra, *Spectrochim. Acta, Part B*, 2013, **86**, 31–41.
- 26 C. A. Munson, F. C. De Lucia, T. Piehler, K. L. McNesby and 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*, 2005, **60**(7–8), 1217–1224.
- 27 R. R. Hark, J. J. Remus, L. J. East, *et al.*, Geographical analysis of “conflict minerals” utilizing laser-induced breakdown spectroscopy, *Spectrochim. Acta, Part B*, 2012, **74**–75, 131–136.
- 28 J. L. Gottfried, R. S. Harmon, F. C. De Lucia and A. W. Miziolek, Multivariate analysis of laser-induced breakdown spectroscopy chemical signatures for geomaterial classification, *Spectrochim. Acta, Part B*, 2009, **64**(10), 1009–1019.
- 29 A. Larsson, H. Andersson and L. Landström, Impact of data reduction on multivariate classification models built on spectral data from bio-samples, *J. Anal. At. Spectrom.*, 2015, **30**(5), 1117–1127.
- 30 J. D. Hybl, G. A. Lithgow and S. G. Buckley, Laser-Induced Breakdown Spectroscopy Detection and Classification of Biological Aerosols, *Appl. Spectrosc.*, 2003, **57**(10), 1207–1215.
- 31 F. C. De Lucia and J. L. Gottfried, Influence of variable selection on partial least squares discriminant analysis models for explosive residue classification, *Spectrochim. Acta, Part B*, 2011, **66**(2), 122–128.
- 32 J. Lasue, R. C. Wiens, T. F. Stepinski, O. Forni, S. M. Clegg and S. Maurice, Nonlinear mapping technique for data visualization and clustering assessment of LIBS data: application to ChemCam data, *Anal. Bioanal. Chem.*, 2011, **400**(10), 3247–3260.
- 33 R. B. Anderson, J. F. Bell, R. C. Wiens, R. V. Morris and S. M. Clegg, Clustering and training set selection methods for improving the accuracy of quantitative laser induced breakdown spectroscopy, *Spectrochim. Acta, Part B*, 2012, **70**, 24–32.
- 34 J. Cisewski, E. Snyder, J. Hanning and L. Oudejans, Support vector machine classification of suspect powders using laser-induced breakdown spectroscopy (LIBS) spectral data, *J. Chemom.*, 2012, **26**(5), 143–149.
- 35 M. Hoehse, A. Paul, I. Gornushkin and U. Panne, Multivariate classification of pigments and inks using combined Raman spectroscopy and LIBS, *Anal. Bioanal. Chem.*, 2012, **402**(4), 1443–1450.
- 36 S. Moncayo, S. Manzoor, F. Navarro-Villoslada and J. O. Caceres, Evaluation of supervised chemometric methods for sample classification by Laser Induced Breakdown Spectroscopy, *Chemom. Intell. Lab. Syst.*, 2015, **146**, 354–364.
- 37 A. Koulejev, M. Sabsabi, V. Motto-Ros, S. Laville and S. L. Lui, Laser-induced breakdown spectroscopy with artificial neural network processing for material identification, *Planet. Space Sci.*, 2010, **58**(4), 682–690.
- 38 F. Y. Yueh, H. Zheng, J. P. Singh and S. Burgess, Preliminary evaluation of laser-induced breakdown spectroscopy for tissue classification, *Spectrochim. Acta, Part B*, 2009, **64**(10), 1059–1067.
- 39 J. B. Sirven, B. Sallé, P. Muchien, J. L. Lacour, S. Maurice and G. Manhès, Feasibility study of rock identification at the surface of Mars by remote laser-induced breakdown spectroscopy and three chemometric methods, *J. Anal. At. Spectrom.*, 2007, **22**(12), 1471–1480.
- 40 J. Novotný, M. Brada, M. Petrilak, D. Prochazka, K. Novotný, A. Hrdlička and J. Kaiser, A versatile interaction chamber for laser-based spectroscopic applications, with the emphasis on Laser-Induced Breakdown Spectroscopy, *Spectrochim. Acta, Part B*, 2014, **101**, 149–154.
- 41 R. G. Brereton, *Applied Chemometrics for Scientists*, John Wiley, Hoboken, 2007.
- 42 D. Prochazka, M. Bilik, P. Prochazková, *et al.*, Detection of tire tread particles using laser-induced breakdown spectroscopy, *Spectrochim. Acta, Part B*, 2015, **108**, 1–7.
- 43 A. Kramida, Y. Ralchenko, J. Reader and NIST ASD Team, *NIST Atomic Spectra Database: (Ver. 5.2)*, 2014, <http://physics.nist.gov/asd>, accessed: 30-08-2016.

## Article 21

P. Modlitbová, K. Novotný, P. Pořízka, J. Klus, P. Lubal, H. Zlámalová-Gargošová, and J. Kaiser. Comparative investigation of toxicity and bioaccumulation of Cd-based quantum dots and Cd salt in freshwater plant *Lemna minor* L. *Ecotoxicology and Environmental Safety*, 147:334–341, 2018.

DOI: 10.1016/j.ecoenv.2017.08.053; 17 citations

This publication is from the body of work dedicated to study the toxicity of nanoparticles. Toxicology is well established field where standard procedures are set to monitor the processes induced in plant tissues with varying contents of selected chemical compounds. The interest lies in the mechanisms beyond uptake, transport, distribution, and storage of novel materials and their influence on the growth, development and nutrition of model plant tissues.

LIBS opens new perspective in toxicology when providing a platform for large-scale mapping of whole plants or its parts (roots, stems, leaves). This is not possible by using standard analytical techniques; most often, ICP-OES is used for assessment of bulk chemical composition. Map of distribution of selected analyte within the plant tissue is then a result of LIBS analysis.

Our work has shown that LIBS is fully matured technique that yields important information about the investigated sample - toxicity of nanoparticles on living organisms.



Contents lists available at ScienceDirect

## Ecotoxicology and Environmental Safety

journal homepage: [www.elsevier.com/locate/ecoenv](http://www.elsevier.com/locate/ecoenv)Comparative investigation of toxicity and bioaccumulation of Cd-based quantum dots and Cd salt in freshwater plant *Lemna minor* L.Pavλίna Modlitbová<sup>a,\*</sup>, Karel Novotný<sup>b,c</sup>, Pavel Pořízka<sup>a</sup>, Jakub Klus<sup>a</sup>, Přemysl Lubal<sup>b,c</sup>, Helena Zlámalová-Gargošová<sup>d</sup>, Jozef Kaiser<sup>a</sup><sup>a</sup> Central European Institute of Technology (CEITEC) Brno University of Technology, Technická 3058/10, 616 00 Brno, Czech Republic<sup>b</sup> Central European Institute of Technology (CEITEC) Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic<sup>c</sup> Department of Chemistry, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic<sup>d</sup> Faculty of Chemistry - The Institute of Chemistry and Technology of Environmental Protection, Brno University of Technology, Purkyňova 118, 612 00 Brno, Czech Republic

## ARTICLE INFO

## Keywords:

Bioaccumulation  
Cadmium  
Quantum dots  
Laser-Induced Breakdown Spectroscopy  
*Lemna minor* L.  
Toxicity

## ABSTRACT

The purpose of this study was to determine the toxicity of two different sources of cadmium, i.e. CdCl<sub>2</sub> and Cd-based Quantum Dots (QDs), for freshwater model plant *Lemna minor* L. Cadmium telluride QDs were capped with two coating ligands: glutathione (GSH) or 3-mercaptopropionic acid (MPA). Growth rate inhibition and final biomass inhibition of *L. minor* after 168-h exposure were monitored as toxicity endpoints. Dose-response curves for Cd toxicity and EC<sub>50,168h</sub> values were statistically evaluated for all sources of Cd to uncover possible differences among the toxicities of tested compounds. Total Cd content and its bioaccumulation factors (BAFs) in *L. minor* after the exposure period were also determined to distinguish Cd bioaccumulation patterns with respect to different test compounds. Laser-Induced Breakdown Spectroscopy (LIBS) with lateral resolution of 200 μm was employed in order to obtain two-dimensional maps of Cd spatial distribution in *L. minor* fronds. Our results show that GSH- and MPA-capped Cd-based QDs have similar toxicity for *L. minor*, but are significantly less toxic than CdCl<sub>2</sub>. However, both sources of Cd lead to similar patterns of Cd bioaccumulation and distribution in *L. minor* fronds. Our results are in line with previous reports that the main mediators of Cd toxicity and bioaccumulation in aquatic plants are Cd<sup>2+</sup> ions dissolved from Cd-based QDs.

## 1. Introduction

Quantum dots (QDs) are fluorescent semiconductor nanocrystals, commonly made up by a 3–6 nm diameter core of CdS, CdSe, PbSe, CdTe or a range of other metals, and coated by an organic polymer (Chan et al., 2002). The use of QDs has been increasing because of their great potential to replace traditional organic dyes as labels for tagging and imaging in biological systems (Jamieson et al., 2007). The main advantage of QDs in comparison to organic dyes or fluorescent proteins is that QDs are brighter, more stable against photobleaching, and can be excited for multicolor emission with a single light source (Bailey et al., 2004; Resch-Genger et al., 2008). However, analogously to other classes of nanomaterials, QDs may eventually find their way into the environment. In contact with aqueous media, Cd-based QDs have been shown to leach ionic Cd (Xu et al., 2010) which has been ranked the 7th out of 275 compounds, including organic chemicals, in the 2015 Priority List of Hazardous Substances (Agency for Toxic Substances and Disease Registry, 2015). Therefore, it is of a high importance to assess

the toxicity of Cd-based QDs to environmental organisms, which may come into contact with QD-containing products when they are discarded.

Quantum dot toxicity is ascribed either to the induction of reactive oxygen species (ROS) formation or to the direct release of Cd ions; in most cells, these reactions cause cellular changes culminating in DNA damage (Gomes et al., 2011). Several studies have researched toxic effects of Cd-based QDs and/or possibility for their bioaccumulation in freshwater organisms, such as microorganisms (Gomes et al., 2011); algae – *Chlamydomonas reinhardtii* Dangeard (Domingos et al., 2011) and *Phaeodactylum tricorutum* Bohlin (Xu et al., 2010); and invertebrates – *Hydra vulgaris* (Ambrosone et al., 2012), *Leptocheirus plumulosus* (Jackson et al., 2012), *Daphnia magna* (Lee et al., 2009) and *Elliptio complanata* (Gagné et al., 2008; Peyrot et al., 2009). Free Cd released from QDs was shown to bioaccumulate in algae (Domingos et al., 2011) and amphipods (Jackson et al., 2012); to alter the synthesis of metallothioneins and trigger oxidative stress and DNA damage in mussels (Gagné et al., 2008; Peyrot et al., 2009); to cause cytotoxicity in

\* Corresponding author.

E-mail address: [Pavlina.Skarkova@ceitec.vutbr.cz](mailto:Pavlina.Skarkova@ceitec.vutbr.cz) (P. Modlitbová).<http://dx.doi.org/10.1016/j.ecoenv.2017.08.053>Received 6 March 2017; Received in revised form 17 August 2017; Accepted 21 August 2017  
0147-6513/© 2017 Published by Elsevier Inc.

algae (Xu et al., 2010); and to generate ROS in daphnids (Lee et al., 2009). Although studies on the distribution of Cd-based QDs in plants (core/shell QDs: CdSe/CdZnS and CdSe/ZnS) by detection of their fluorescence in different parts of plants have also been conducted (Navarro et al., 2012; Wang et al., 2014; Koo et al., 2015), they focused on the potential of core/shell QDs for uptake, translocation, or transformation, but not on the evaluation of toxicity values. To the best of our knowledge, no toxicity studies with QDs have been conducted in aquatic plants, except in green algae (Xu et al., 2010; Domingos et al., 2011). Given the hazardous potential of QDs, it is very important to address these knowledge gaps in plant ecotoxicology.

The free floating macrophyte *Lemna minor* L. (Lemnoideae, Araceae) is a common aquatic ecotoxicity test organism: it is a bioindicator species for the detection and monitoring of metal pollution (Garczarska and Ratajczak, 2000) and also metal bioaccumulator. *L. minor* has already been used as a model organism in toxicity and bioaccumulation studies of several types of nanoparticles (NPs), e.g. Ag (Gubbins et al., 2011; Jiang et al., 2012; Oukarroum et al., 2013; Üçüncü et al., 2014), CuO (Shi et al., 2011; Perreault et al., 2014), C<sub>60</sub> (Santos et al., 2013), Al<sub>2</sub>O<sub>3</sub> (Juhel et al., 2011), TiO<sub>2</sub> (Song et al., 2012; Li et al., 2013), and ZnO (Hu et al., 2013) NPs, but not yet Cd-based QDs. The effects of Cd<sup>2+</sup> on *L. minor* have been studied with the use of Cd salts: CdCl<sub>2</sub> (Razinger et al., 2008; Tkalec et al., 2008; Balen et al., 2011) and less typically CdSO<sub>4</sub> (Drost et al., 2007) or Cd(NO<sub>3</sub>)<sub>2</sub> (Kwan and Smith, 1991). These studies have shown that Cd accumulates in *L. minor* and causes adverse effects, such as decreased growth, reduced levels of photosynthetic pigments, impaired chloroplast ultrastructure, increased activities of antioxidant enzymes, increased lipid peroxidation and decreased chlorophyll and protein contents (Razinger et al., 2008; Tkalec et al., 2008; Balen et al., 2011). As Cd is known to leach from Cd-based QDs (Xu et al., 2010) and dissolved ions are the main mediators of nanoparticle toxicity to organisms (Ivask et al., 2015), similar Cd bioaccumulation and toxicity as in the case of Cd can be expected upon exposure of *L. minor* to Cd-based QDs.

In this study we focused on several major goals. The first objective was to measure the toxic effects of different Cd compounds in *L. minor*, i.e. Cd salt and two types of Cd-based QDs (QDs capped by glutathione, GSH-QDs, or by 3-mercaptopropionic acid, MPA-QDs), where CdCl<sub>2</sub> served as a positive control for Cd toxicity to *L. minor*. After 168-h exposure two toxicity endpoints were monitored in *L. minor*: growth rate inhibition and final biomass inhibition according to the OECD 221 norm (OECD, 2002). The second objective was to determine the total content of Cd in plants and to distinguish how different sources of Cd are accumulated in *L. minor* fronds by using a conventional method for metal detection, ICP-OES. In this part, CdCl<sub>2</sub> served as a reference compound to test the hypothesis that *L. minor* bioaccumulates Cd ions that leach from Cd-based QDs (Xu et al., 2010). The third objective was to demonstrate that Laser-Induced Breakdown Spectroscopy (LIBS) is a useful tool for mapping the elemental distribution in *L. minor* fronds, as well as to monitor whether there are differences in the bioaccumulation patterns of Cd salt or Cd-based QDs. In the recent years, LIBS has been developed as an alternative and fast method for investigation of spatial distribution of elements; its applicability in plant samples has been summarized in three extensive reviews (Kaiser et al., 2012; Pořízka et al., 2012; Santos et al., 2012). However, LIBS has so far been used only for the detection of nano Ag in root tissues of *Vicia faba* L. (Krajcarová et al., 2017), therefore further studies of its applicability for the detection of NPs in plants are needed. The fourth objective was to inspect *L. minor* fronds by transmission electron microscopy (TEM) to determine if QDs are able to penetrate cell walls and become accumulated in plant tissue or they are only adsorbed on frond surface. We discuss possible hazardous potential of QDs in aquatic environment for *L. minor* as a bioindicator species. We also show that LIBS technique is a successful alternative to conventional analytical methods due to the ability to map large areas of samples in short time with sufficient resolution.

## 2. Materials and methods

### 2.1. Quantum dot synthesis and analysis

Two types of QDs were synthesized – GSH-QDs and MPA-QDs. The methods of preparation were based upon extensive instructions available in Lišková et al. (2011) and Řezáčová et al. (2015). Quantum dot properties were analyzed as follows: the nominal average particle size was determined by FEI Tecnai F20 electron microscope (Thermo Fisher Scientific, Waltham, Massachusetts, USA); zeta potential was measured with Zetasizer Nano ZS (Malvern Instruments, Malvern, UK) and Cd content in QDs was measured by ICP-OES spectrometer iCAP 6500 Duo (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

### 2.2. Toxicity experiment with *L. minor*

#### 2.2.1. Experimental setup

Laboratory stock culture of *L. minor* was used. Seven days before testing, sufficient colonies were transferred aseptically into fresh sterile modified Steinberg medium (ISO, 2005) and cultured under the test conditions. No contaminating organisms (such as algae) were present.

*Lemna* test was performed according to the OECD (Organization for Economic Co-operation and Development) Test No. 221: *Lemna* sp. Growth Inhibition Test using Steinberg medium (OECD, 2002). Toxicity tests were carried out in 200 mL beakers filled with 150 mL solution which consisted of the dilution series of test compounds in the Steinberg medium. Presence of EDTA in Steinberg medium was proven not to influence the metal uptake (Drost et al., 2007). *L. minor* plants were exposed to three Cd-containing compounds: GSH-QDs, MPA-QDs and CdCl<sub>2</sub>·2.5H<sub>2</sub>O (hereafter referred to as CdCl<sub>2</sub>). For both GSH-QDs and MPA-QDs, the nominal test concentrations were 0, 0.01, 0.05, 0.1, 0.5, 1.0, 2.5, 5.0, 7.5, 10, and 15 mg compound/L; for CdCl<sub>2</sub>, the nominal test concentrations were 0, 0.01, 0.05, 0.1, 0.5, 1.0, 2.5, 5.0, 7.5, and 10 mg compound/L.

One exposure group consisted of seven test replicates (beakers) for each test compound at each exposure concentration. There was one control group for each test compound. Before the exposure, the beakers were inoculated with four plants with three fronds each, which resulted in 12 fronds per beaker. The test was carried out for seven days (168 h) at the temperature of 24 ± 2 °C and light intensity of 8000 lx. pH of the Steinberg medium was 6.8 ± 0.1. The test was considered valid if the number of fronds in controls had grown eightfold.

Immediately after the exposure, the number of all *L. minor* fronds (both healthy and necrotic) was counted in each vessel as “all fronds” and “green fronds” (fronds without any evidence of necrosis or damage). The number of “green fronds” was used for the calculation of *L. minor* growth rate ( $\mu$ ) and growth rate inhibition (%  $I_r$ ) for six replicates of each exposure group. Then each exposure group was divided into three parts for various analyses. Three replicates per exposure group were dedicated to the assessment of Cd bioaccumulation, another three replicates to the assessment of final biomass inhibition, and one replicate to the investigation of Cd spatial distribution in fronds and QD adsorption on the plant surface.

Plants from the seventh replicate of each exposure group were thoroughly washed in deionized water. Three or four fronds per beaker were used for LIBS spatial distribution mapping experiments and two fronds per beaker for TEM analyses. For LIBS measurements, fronds were carefully dried, molded and glued by epoxide glue onto a glass slide. For TEM analyses, frond cross-sections were prepared and photographed.

#### 2.2.2. Toxicity parameters

After the exposure period, the number of *L. minor* fronds in each beaker was counted. For all the tests, the *L. minor* growth rate  $\mu$  was used as the first toxicity endpoint. The plant growth rate was calculated on the basis of frond numbers as:

$$\mu = \frac{\ln(F_{t_7}/F_{t_0})}{t_7 - t_0} \quad (1)$$

where  $F_{t_0}$  is the number of fronds at the start of the exposure and  $F_{t_7}$  is the number of fronds at the end of the exposure (OECD, 2002). From a mathematical perspective,  $\mu$  is based on the assumption of exponential growth and gives an average of the growth during the time period from  $t_0$  to  $t_7$ . Growth rate inhibition (%  $I_r$ ) was then calculated in relation to the control mean:

$$\%I_r = 100 \times \frac{(\mu_{\text{control}} - \mu_{\text{sample}})}{\mu_{\text{control}}} \quad (2)$$

The final biomass was determined on the basis of fresh biomass present in each test vessel at the start and at the end of the exposure. The starting biomass is determined on the basis of a sample of fronds taken from the same batch used to inoculate the test vessels. The mean percent inhibition in final biomass (%  $I_b$ ) was calculated for controls and each treatment group as follows:

$$\%I_b = 100 \times \frac{(b_c - b_t)}{b_c} \quad (3)$$

%  $I_b$  is percent reduction in biomass;  $b_c$  was calculated as  $\{\ln(\text{final biomass}/\text{starting biomass})\}$  for the control group and  $b_t$  as  $\{\ln(\text{final biomass}/\text{starting biomass})\}$  in the exposure group (OECD, 2002). These calculations were done for three replicates from each test concentration. Final fresh biomass was determined immediately after the exposure. Before weighing, plants were centrifuged at  $2349 \times g$  for 10 min (NF 800 & 800R Multi-Purpose Centrifuge –Nüve A.Ş., Ankara, Turkey) separately for each test concentration.

### 2.3. Analyses of QD adsorption on *L. minor* surface and Cd content and distribution in fronds

#### 2.3.1. Cd bioaccumulation in *L. minor* (ICP-OES)

Cd bioaccumulation was determined from three replicates at each tested concentration. All the plants from each replicate (beaker) were separately and adequately washed in deionized water, and dried at 60 °C to constant mass.

To quantify the uptake of Cd by *L. minor*, the bioaccumulation factor (BAF) was calculated from the amount of Cd in the biomass and the concentration of Cd in the medium after the exposure:

$$BAF = \frac{m_{Cd}}{C_{Cd}} \quad (4)$$

$m_{Cd}$  is the mass of Cd in dried *L. minor* mass (mg/kg) and  $C_{Cd}$  is the concentration of Cd in the test medium (mg/kg) (Kalčíková et al., 2016).

The total Cd content in *L. minor* was measured by ICP-OES spectrometer iCAP 6500 Duo (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Before the analysis, the samples were decomposed in a microwave lab station (Ethos One, Milestone, Bergamo, Italy) under the following time and temperature regime: 15 min of heating to 200 °C; 20 min of decomposition at 200 °C; and 10 min of cooling to 110 °C. All plants from one replicate were pooled and digested as a single sample in 7 mL of conc.  $\text{HNO}_3$  ( $\geq 65\%$ , p.a. purity, Sigma Aldrich, Steinheim, Germany) and 1 mL of conc.  $\text{H}_2\text{O}_2$  ( $\geq 30\%$ , p.a. purity, Sigma Aldrich, Steinheim, Germany).

#### 2.3.2. Spatial distribution of Cd in *L. minor* fronds (DP-LIBS)

The modified laser system UP-266 MACRO (New Wave Research, Fremont, CA, USA) equipped with software-controlled movement in  $x$  and  $y$  directions was used. Glass slides with glued plant samples were placed into a lab-made holder. Two Nd: YAG lasers were arranged in the orthogonal reheating configuration. The first laser (UP-266 MACRO) that operated at the fourth harmonic frequency (266 nm) with the energy 10 mJ per pulse irradiated the sample surface and created a

micro plasma. The second laser (Q-Smart, Quantel Laser, Les Ulis, France) at fundamental wavelength 1064 nm with energy 100 mJ per pulse was focused with 80 mm focal length glass lens to intersect the path of the first laser beam 0.5 mm above the sample surface and finally to create the coincident spark. Both lasers operate with the pulse length of ~5 ns. The energy of the laser was measured with laser power/energy meter (Nova, Ophir Optronics Ltd., Jerusalem, Israel). The plasma emission was collected and transported with 3 m long optical fiber into the Czerny-Turner spectrometer (TRIAx 320, Horiba Jobin Yvon Inc., Edison, NJ, USA) with 0.1 mm entrance slit and optical grating with 2400 grooves/mm. The spectral resolution of the spectrometer with this grating was 0.050 nm. The LIBS spectrum was detected with ICCD detector (PI MAX 3, Princeton Instruments, Trenton, NJ, USA). Both lasers were externally triggered by two delay generators (DG 645, Stanford Research System, Sunnyvale, CA, USA) which allowed the time synchronization of flash lamps and Q-switches of both laser and also setting of ICCD detection delay.

For our samples the optimal inter-pulse delay of 500 ns (the delay of the second laser pulse after the first one), detection delay of 1000 ns and detection integration time of 10  $\mu\text{s}$  were established. The raster covered an area from  $14 \times 15$  points (the smallest fronds) to  $29 \times 18$  points (the largest fronds) in size with lateral resolution of 200  $\mu\text{m}$ . Each laser pulse reached through the whole “depth” of the frond. The intensity of analytical emission line Cd I at 508.58 nm was evaluated as the maximum line intensity after appropriate background subtraction. The background was estimated as the mean of the detected signal (range 507.8–508.2 nm) in close proximity to the selected analytical line. No further intensity normalization or correction was used. The measured data allowed the construction of 2D maps showing the spatial distribution of Cd element in selected fronds. All evaluations were conducted in the R statistical package (“ggplot2” and “signal” library; “ggplot2” and “signal” library; R Development Core Team, 2015 Development Core Team, 2015).

#### 2.3.3. Adsorption of QDs on the *L. minor* surface (TEM)

Frond samples were prepared as described in Li et al. (2013). Fronds were fixed in 2.5% glutaraldehyde (Grade I, 25% in  $\text{H}_2\text{O}$ , specially purified for use as an electron microscopy fixative; Sigma Aldrich, Steinheim, Germany) – 0.1 M sodium cacodylate trihydrate buffer ( $\geq 98\%$ ; Sigma Aldrich, Steinheim, Germany) for 3 h. After three washing steps in 0.1 M cacodylate buffer, samples were post-fixed in 1%  $\text{OsO}_4$  solution ( $\geq 98\%$ ; Sigma Aldrich, Steinheim, Germany) in the same buffer for 2 h. The specimens were then dehydrated in ethanol, treated with acetone and embedded in epoxy resin Durcupan™ ACM (Sigma Aldrich, Steinheim, Germany). Ultrathin sections were cut at 60 nm on an ultramicrotome Leica EM UC6 (Leica Microsystems Inc., Buffalo Grove, US) and placed on 50 mesh formvar-coated nickel grids. Selected ultrathin sections were contrasted with 2.5% uranyl acetate (p.a., Lachema, Brno, Czech Republic) for 10 min and alkaline Reynolds lead citrate solution (p.a., Penta, Chrudim, Czech Republic) for 6 min.

4  $\mu\text{L}$  of freshly prepared QDs solutions were applied onto glow discharged Electron Microscope (EM) grids and left to dry in ambient atmosphere. Cu 400 mesh EM grids coated by homemade 20 nm continuous carbon film were glow discharged in nitrogen atmosphere at 12 nA.

Samples of QDs as well as frond cross-sections were transferred to an FEI Tecnai F20 electron microscope (Thermo Fisher Scientific, Waltham, Massachusetts, USA) operating at 200 kV using basic room temperature sample holder. Micrographs were recorded using Eagle 4 K  $\times$  4 K CCD under low-dose mode ( $\sim 20 \text{ e}^-/\text{\AA}^2$ ). Focus values for micrograph acquisition ranged from 3 to 4  $\mu\text{m}$ . Nominal magnifications were 29,000; 50,000 and 100,000, which resulted in pixel size of 3.78; 2.22; 1.12  $\text{\AA}/\text{pixel}$ .

## 2.4. Data analysis

Toxicity data were fitted by the Hill equation to curves of dose (Cd concentration  $c_{Cd}$ ) vs. response (growth rate inhibition and final biomass inhibition). The fitted curves were compared by ANOVA using R statistical package (“nlme” and “car” library; “nlme” and “car” library; R Development Core Team, 2015 Development Core Team, 2015). Calculated  $F$ -values were used for the evaluation of significant differences ( $F$ -values > 0.05; \*). For the data from bioaccumulation experiment (BAF and  $m_{Cd}$ ), only the linear approximations of BAF and  $m_{Cd}$  dependencies on Cd concentration ( $c_{Cd}$ ) were performed to obtain these data at the same x axis positions (approximately same concentrations). Those were then compared by Mann-Whitney T-test using R statistical package to evaluate significant differences ( $p$ -values < 0.05; \*).

## 3. Results and discussion

### 3.1. Quantum dot properties

MPA-QDs contained ( $23.19 \pm 0.06$ ) % m/m of Cd (ICP-OES). Their nominal average particle size was 4.0–5.4 nm (TEM) and zeta-potential was ( $-45 \pm 10$ ) mV while GSH-QDs of the nominal average particle size 4.0–4.4 nm (TEM) contained ( $22.99 \pm 0.05$ ) % m/m of Cd (ICP-OES) and exhibited zeta-potential of ( $-45 \pm 14$ ) mV.

Aggregation of both QD types at different pH values was studied in our previous paper (Škarková et al., 2017). In the pH range from 9.8 to 5, the hydrodynamic diameter and zeta potential did not significantly change for MPA-QDs. In the pH range from 5 to 2, MPA-QDs started to agglomerate and form clusters of 500 nm diameter at pH 4 and even of 1.500 nm at pH 3 (Škarková et al., 2017). The aggregation was not investigated in detail for GSH-QDs, but their behavior is expected to be the same as that of MPA-QDs. In our toxicity test described here, pH values of Steinberg medium before and after the test were in the range from 6.6 to 6.9, so the presence of any QD aggregates or clusters were not expected.

Cd concentrations in test medium before and after the exposure are listed in Table 1. Additionally, recoveries for measured versus nominal Cd concentrations before the toxicity tests were calculated. Recoveries, except at the two lowest concentrations, were within the range from 97% to 133%. Only the measured Cd concentrations (Table 1) were used in our calculations and in the figures.

### 3.2. Toxicity of Cd-based QDs and CdCl<sub>2</sub> to *L. minor*

Fig. 1 shows the dependence of growth rate inhibition and final biomass inhibition (calculated from fresh biomass) on Cd concentration ( $c_{Cd}$ ) plotted for each test compound: GSH-QDs (Fig. 1a), MPA-QDs (Fig. 1b), and CdCl<sub>2</sub> (Fig. 1c). To evaluate the differences in the toxicity of tested compounds, our data were first fitted by the Hill equation (Gadagkar and Call, 2015) to the dose-response curves of growth rate inhibition and final biomass inhibition. Then the fitted curves were compared to each other by ANOVA. No significant differences were obtained between the dose-response curves of *L. minor* growth rate inhibition or final biomass inhibition parameters to both types of QDs ( $F \ll 1$  in all four cases). However, both toxicity endpoints showed that CdCl<sub>2</sub> was significantly more toxic than GSH-QDs ( $F = 47.9^*$  for growth rate inhibition and  $F = 4.7^*$  for final biomass inhibition) as well as MPA-QDs ( $F = 23.2^*$  for growth rate inhibition and  $F = 22.8^*$  for final biomass inhibition).

The mechanism of in vivo toxicity of Cd-based QDs is still unclear, but recently published studies have inclined that the main responsible agents are Cd<sup>2+</sup> ions, which can be (a) released from the particle core in the exposure medium, (b) adsorbed on the particle surface, or (c) released inside the organisms upon internalization into cells (Ambrosone et al., 2012; Chen et al., 2012). Positive relationship between the Cd<sup>2+</sup> release rate from QDs and cytotoxicity has already been

confirmed in alga *P. tricornutum* (Xu et al., 2010). Our results are in line with these predictions since (a) no “particle-specific” toxic effects of QDs were detected, and (b) QDs were less toxic than CdCl<sub>2</sub> at approximately the same nominal Cd exposure concentrations (Table 1), probably because the concentration of Cd<sup>2+</sup> leached into the test medium from QDs (Xu et al., 2010) was smaller than the concentration of Cd<sup>2+</sup> in the test media spiked with CdCl<sub>2</sub>. At the same time, the assumption of Cd<sup>2+</sup>-mediated QD toxicity implies that GSH and MPA ligands did not protect QDs from dissolution in the test medium, similarly as already shown by Kirchner et al. (2005). The (cyto)toxicity of QDs can be decreased by using stable and effective coating material such as ZnS and SiO<sub>2</sub> shells, which prevent Cd<sup>2+</sup> release from QD core (Kirchner et al., 2005; Xu et al., 2010; Chen et al., 2012). However, a study on alga *C. reinhardtii* (Domingos et al., 2011) has found that toxicity of CdTe-QDs and free Cd<sup>2+</sup> ions may be mediated by different physiological mechanisms, which was corroborated by different whole transcriptome profiles using RNA-Seq analysis. Further experiments are needed to elucidate whether the toxicity mechanism of QDs is different than that of CdCl<sub>2</sub> also for *L. minor*.

Based on the dependences from Fig. 1, EC50<sub>168h</sub> values for all test compounds and for both toxicity endpoints (growth rate inhibition and final biomass inhibition) were calculated (Table 2). The observed EC50<sub>168h</sub> values for growth rate inhibition were in the range from 0.31 to 0.53 mg/L for all sources of Cd. Our results are comparable to previous studies with *Lemna*, where EC50<sub>168h</sub> for CdSO<sub>4</sub> were determined to be 0.43 mg/L by Drost et al. (2007) and for CdCl<sub>2</sub> to be 0.21 mg/L (Basile et al., 2012), while EC50<sub>240h</sub> for Cd(NO<sub>3</sub>)<sub>2</sub> was reported to be 0.32 mg/L (Kwan and Smith, 1991). However, it can be seen in Table 2 that the EC50<sub>168h</sub> values with the 95% confidence interval did not overlap. Based on the EC50<sub>168h</sub> values only, one could conclude that all three compounds have different toxicity to *L. minor*. However, Hill equation utilized for experimental data fitting showed no significant differences between the dose-response toxicity curves of both types of Cd-based QDs. This finding reveals that a single data point could be inconclusive; it is thus necessary to evaluate the response of *L. minor* to compounds throughout the whole exposure range in order to reliably estimate the toxicity trends.

### 3.3. Bioaccumulation of Cd in *L. minor*

The dependences of BAF and  $m_{Cd}$  on Cd concentration ( $c_{Cd}$ ) are shown in Fig. 2. Mann-Whitney T-test for approximate linear dependence curves was used to evaluate the differences among the test compounds. No significant differences were found between both types of QDs, as well as between CdCl<sub>2</sub> and GSH-QDs or CdCl<sub>2</sub> and MPA-QDs ( $p$ -values > 0.05 in all cases). These results indicate that the Cd bioaccumulation pattern by *L. minor* was similar regardless of the test compound.

Bioaccumulation is known to be a nonlinear process, so BAF is generally the highest at low concentrations and decreases with increasing metal concentrations (Gregor, 2004). BAF results were recalculated for EC50<sub>168h</sub> growth rate inhibition (Table 2) to compare them better with published values for Cd salts. For the tested compounds, BAF values for EC50<sub>168h</sub> growth rate inhibition ranged from 4463 to 6607 (Table 2) and the uptake of Cd was inversely proportional to its concentration in the solution; this trend was observed for both sources of Cd (Table 1). The same trend as in our study was reported also by Basile et al. (2012):  $m_{Cd}$  values in *L. minor* increased with increasing Cd concentrations (from CdCl<sub>2</sub>) while BAF values decreased with increasing Cd concentrations and ranged from 717 at the Cd exposure concentration 0.02 mg/L to 13,991 at the Cd exposure concentration of 0.0002 mg/L. The dependence of BAF values on the Cd exposure concentration were confirmed also in a study on *L. trisulca* L. treated with CdSO<sub>4</sub> (Prasad et al., 2001). However, after 3-day exposure of *L. minor* to CdSO<sub>4</sub>, BAF for Cd was reported to be 1371 for EC50 value of growth rate inhibition (Drost et al., 2007); although this value

Table 1

Nominal concentrations of tested compounds (mg compound/L), nominal concentrations of Cd in tested compounds (mg Cd/L), and measured concentrations of Cd before and after the test (mg Cd/L). Recovery (in %) was calculated for measured concentration before the test in comparison to calculated Cd nominal concentration before the test. All data with standard deviations are averages of three replicates for each concentration. Mark (-) stands for “below LOD” and (/) for “not calculated”.

Nominal C <sub>COMPOUNDS</sub> (mg/L)	Nominal C <sub>Cd</sub> (mg/L)	Real C <sub>Cd</sub>		Recovery before the test (%)
		Before the test (mg/L)	After the test (mg/L)	
<b>GSH-QDs</b>				
15	3.449	3.743 ± 0.003	3.59 ± 0.05	112
10	2.299	2.569 ± 0.003	2.44 ± 0.04	115
7.5	1.724	1.927 ± 0.003	1.82 ± 0.01	115
5	1.150	1.284 ± 0.002	1.22 ± 0.02	115
2.5	0.575	0.742 ± 0.001	0.68 ± 0.02	133
1	0.230	0.281 ± 0.002	0.21 ± 0.02	126
0.5	0.115	0.128 ± 0.001	0.073 ± 0.004	115
0.1	0.023	0.0257 ± 0.0001	0.0046 ± 0.0001	115
0.05	0.011	0.0129 ± 0.0003	–	116
0.01	0.002	0.0026 ± 0.0005	–	117
0	0	–	–	/
<b>MPA-QDs</b>				
15	3.479	4.116 ± 0.004	3.34 ± 0.03	118
10	2.319	2.613 ± 0.003	2.21 ± 0.03	113
7.5	1.739	1.988 ± 0.003	1.68 ± 0.02	114
5	1.160	1.277 ± 0.002	1.20 ± 0.08	110
2.5	0.580	0.606 ± 0.001	0.55 ± 0.005	104
1	0.232	0.248 ± 0.001	0.22 ± 0.003	107
0.5	0.116	0.1129 ± 0.001	0.078 ± 0.002	97
0.1	0.023	0.0234 ± 0.0004	0.0107 ± 0.0003	102
0.05	0.012	0.0129 ± 0.0003	0.0043 ± 0.0002	108
0.01	0.002	0.0033 ± 0.0001	–	165
0	–	–	–	/
<b>CdCl<sub>2</sub></b>				
10	4.923	5.528 ± 0.005	5.50 ± 0.05	112
7.5	3.692	4.103 ± 0.002	4.06 ± 0.03	111
5	2.461	2.769 ± 0.006	2.75 ± 0.03	113
2.5	1.231	1.337 ± 0.004	1.30 ± 0.02	109
1	0.492	0.543 ± 0.005	0.49 ± 0.01	110
0.5	0.246	0.269 ± 0.002	0.21 ± 0.01	109
0.1	0.049	0.060 ± 0.001	0.027 ± 0.002	122
0.05	0.025	0.031 ± 0.000	0.008 ± 0.001	124
0.01	0.005	0.012 ± 0.000	–	240
0	–	–	–	/

is lower than the one found in our study, it may be ascribed to the shorter exposure time.

The similarities of BAF values in our study to those reported in the literature for various Cd salts during the same exposure duration, and no significant differences between BAF values for CdCl<sub>2</sub> and QDs strongly suggest that only Cd<sup>2+</sup> ions and not whole QDs were bioaccumulated by *L. minor*. Our results are in line with the study of Koo et al. (2015) where Cd from QDs with unstable surface coating (polyethylenimine, PEI) was bioaccumulated in *Arabidopsis thaliana* (L.) Heynh due to QD destabilization in the hydroponic medium and direct uptake of free Cd<sup>2+</sup> ions. Our quantitative bioaccumulation data were also corroborated with Cd spatial distribution analyses by LIBS, which are described in the next chapter.

#### 3.4. Distribution of Cd in *L. minor* and adsorption of QDs on *L. minor* surface

Two-dimensional maps of Cd element spatial distribution were constructed for each frond after the analysis with LIBS. Fig. 3 comprises photographs of fronds before the measurements and LIBS maps for all tested Cd compounds at the nominal exposure concentrations 0.1, 1 and 10 mg/L. Our findings were as follows: (a) no differences in the distribution of Cd in the fronds of *L. minor* exposed to CdCl<sub>2</sub> and those exposed to both types of Cd-QDs were found; (b) spatial distribution of Cd was independent of the tested concentration; (c) the intensity of the LIBS signal increased with increasing nominal concentrations of Cd for all test compounds; (d) the only part of *L. minor* where Cd accumulation was notably higher than in other tissues was the node of the plant; any

other preferences (lateral or middle veins, apex or daughter frond) for the accumulation of Cd were not observed. TEM photographs of frond cross-sections (Fig. 4) demonstrate no visible surface attachment of QDs or their aggregates as well as no presence of QDs inside plant tissues.

The LIBS and TEM findings collectively support the common presumption (Xu et al., 2010) that Cd bioaccumulation in *L. minor* is due to uptake of free Cd<sup>2+</sup> ions released from QDs core into test medium and not due to internalization of intact QDs. Our results are in good agreement with studies on *C. reinhardtii* (Domingos et al., 2011) as well as on *A. thaliana* (Navarro et al., 2012), where it was shown that polymer-coated (carboxylate-terminated) core/shell CdSe/ZnS-QDs were not internalized and translocated as intact QDs within 7 days of exposure. Unlike its constituent ions, the QDs were evidently not taken up and were generally adsorbed onto the plant root surfaces (Navarro et al., 2012).

Laser-Induced Breakdown Spectroscopy was already successfully used for element mapping in plants with similar lateral resolution and demonstrated to be a good choice for this type of analysis. However, only one investigation (Kaiser et al., 2007) has so far focused on spatial Cd mapping in dried sunflower (*Helianthus annuus* L.) leaves with similar spatial resolution (crater diameter after laser pulse was approximately 100 μm), but with different LIBS instrumentation (femtosecond laser, Ti: sapphire, single pulse mode, wavelength 795 nm). Our present study is therefore the first successful demonstration of the use of Nd: YAG laser LIBS (nanosecond lasers, Nd: YAG, double pulse mode, wavelength 266 nm and 1064 nm) for mapping Cd distribution in dried aquatic plant tissues.

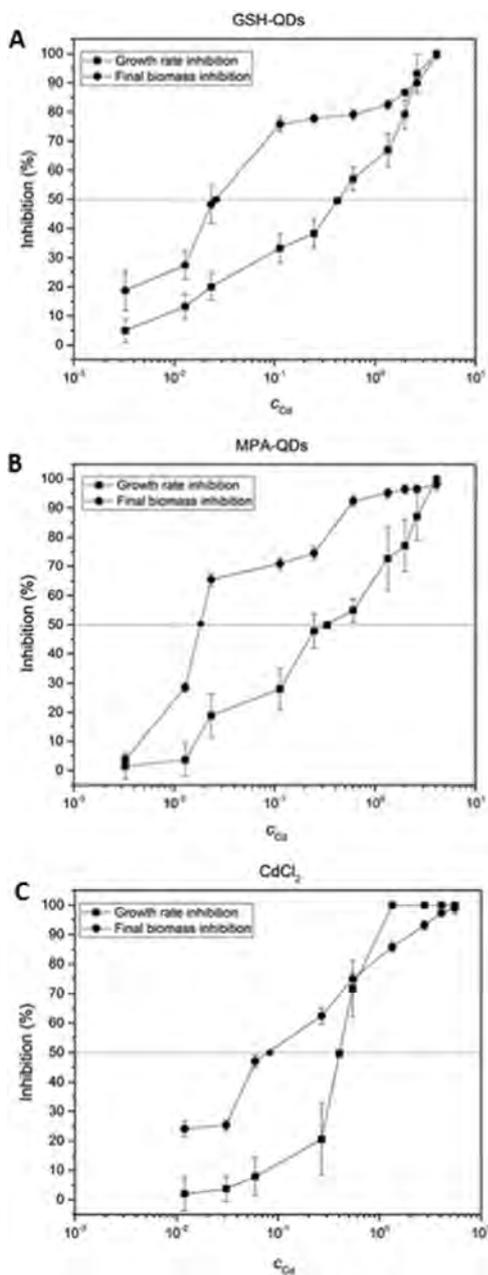


Fig. 1. The dependence of growth rate inhibition and final biomass inhibition on Cd concentration ( $c_{Cd}$ ) for A) GSH-QDs, B) MPA-QDs, and C)  $CdCl_2$ . Growth rate inhibition values were calculated from six test replicates and final biomass inhibition values from three replicates. Standard deviations are shown as error bars, while  $EC_{50_{168h}}$  values from both inhibition parameters are marked by symbols  $\bullet$  and  $\blacksquare$  and by dotted lines.

4. Conclusion

In summary, toxicity/bioaccumulation/spatial element distribution of two different sources of Cd ( $CdCl_2$  and Cd-based QDs) in the aquatic model organism *L. minor* was tested. Several measured endpoints did

Table 2  
Estimated  $EC_{50_{168h}}$  values and their 95% confidence intervals for two endpoints: final biomass inhibition and growth rate inhibition; concentrations of Cd in the test medium before test (mg Cd/L); calculated bioaccumulation factor (BAF) for growth rate inhibition  $EC_{50_{168h}}$  values.

	$EC_{50_{168h}}$	95% Confidence interval	BAF
<b>GSH-QDs</b>			
Final biomass inhibition	0.03	0.03–0.03	/
Growth rate inhibition	0.51	0.48–0.53	4463
<b>MPA-QDs</b>			
Final biomass inhibition	0.02	0.02–0.02	/
Growth rate inhibition	0.33	0.31–0.34	5211
<b><math>CdCl_2</math></b>			
Final biomass inhibition	0.08	0.08–0.09	/
Growth rate inhibition	0.39	0.37–0.41	6607

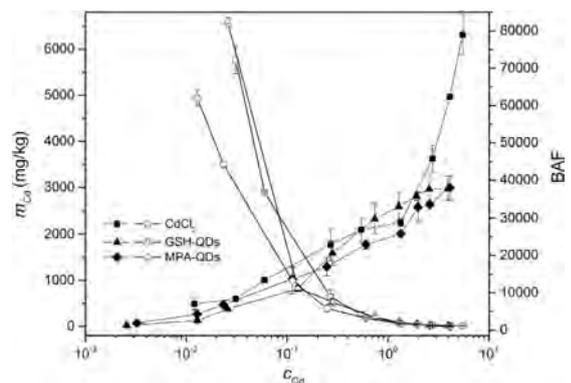


Fig. 2. The dependence of BAF and  $m_{Cd}$  on Cd concentration ( $c_{Cd}$ ) for all sources of Cd; all values are averages of three replicates. BAF – contour marks,  $m_{Cd}$  – full marks, standard deviations are shown as error bars.

not differ between the plants exposed to different types of Cd compounds. As observed by LIBS and ICP-OES analyses,  $CdCl_2$ , MPA-QDs and GSH-QDs showed the same bioaccumulation pattern in *L. minor* fronds, which quantitatively depends on Cd concentration but spatially independent both on Cd concentration and its source. Inspecting *L. minor* fronds with TEM revealed no visible surface attachment of QDs or their aggregates as well as no presence of QDs inside plant tissues. These results strongly point to the conclusion that Cd accumulated by *L. minor* originated from dissolved  $Cd^{2+}$  ions that were released into test medium from QD core; they also suggest that employed coatings did not stabilize QDs against dissolution. Our results are in line with the current findings of other researchers studying the response of algae and plants to Cd-based QDs (Xu et al., 2010; Domingos et al., 2011; Navarro et al., 2012; Koo et al., 2015).

However, comparing the dose-response dependences growth rate inhibition and final biomass inhibition of *L. minor* to Cd concentration for different Cd compounds revealed that toxicity of QDs to *L. minor* was comparable, but significantly lower than that of the  $CdCl_2$ . It was presumed that the concentration of dissolved  $Cd^{2+}$  ions was lower than the concentration of  $Cd^{2+}$  in the  $CdCl_2$ -spiked test media. Our results show that MPA-QDs and GSH-QDs may be toxic to *L. minor* due to unstable coatings that do not prevent leaching of  $Cd^{2+}$  from QD core, but less toxic than  $CdCl_2$  at approximately the same total Cd exposure concentrations.

Acknowledgments

This research has been financially supported by the Ministry of Education, Youth and Sports of the Czech Republic under the project

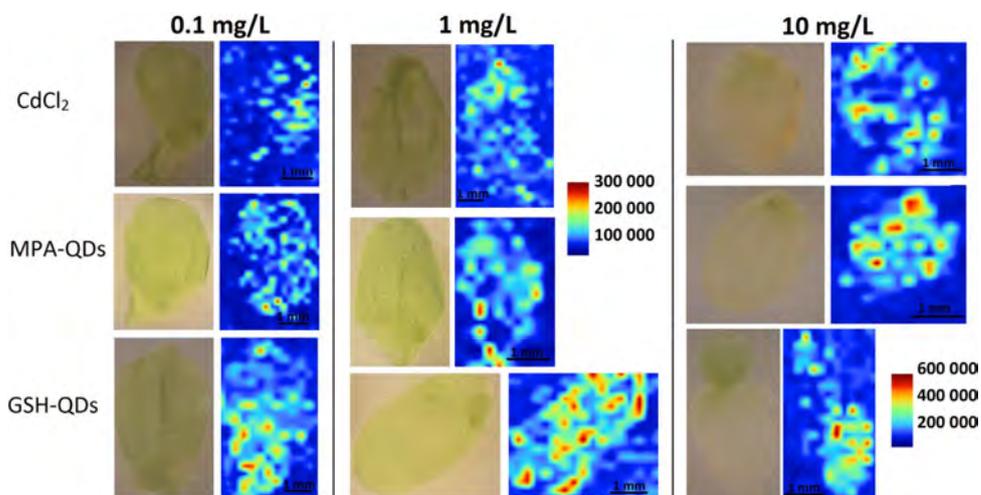


Fig. 3. Photographs of *L. minor* fronds exposed to nominal concentration 0.1, 1 and 10 mg/L CdCl<sub>2</sub>, MPA-QDs and GSH-QDs before LIBS measurements (left column). LIBS maps constructed for Cd (I) 508.58 nm emission line of *L. minor* fronds exposed to the nominal concentration 0.1, 1 and 10 mg/L CdCl<sub>2</sub>, MPA-QDs and GSH-QDs (right column). The first scale with the maximum value 300,000 is related to the nominal concentrations 0.1 and 1 mg/L, the second scale with the maximum value 600,000 is related to the nominal concentration 10 mg/L.

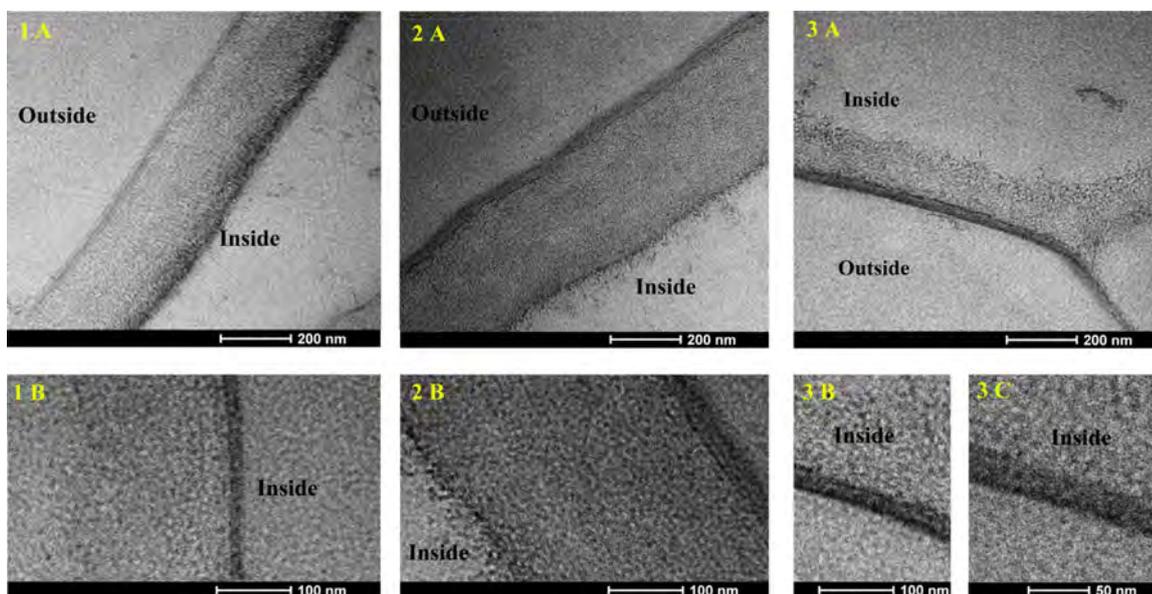


Fig. 4. TEM photographs of *L. minor* frond cross-sections after the test: control group (1A-1B), exposure to nominal concentration 1 mg/L GSH-QDs (2A-2B), and exposure to nominal concentration 1 mg/L MPA-QDs (3A-3B-3C).

CEITEC 2020 (LQ1601) and under the project FCH-S-16–3364; and by Masaryk University under the project MUNI/A/1500/2015. CIHSB research infrastructure project LM2015043 funded by MEYS CR is gratefully acknowledged for the financial support of the measurements at the Core Facility CEITEC Cryo-electron Microscopy and Tomography. And we would like to gratefully thank to Dr. Tea Romih for her precious advice, to Assoc. Prof. Aleš Hampl for preparing cross-sections of samples and to Zdeněk Farka for assistance in preparing samples for LIBS analysis.

References

Agency for Toxic Substances and Disease Registry (ATSDR), 2015. Priority List of Hazardous Substances. U.S. Department of Health and Human Services, Public Health Service, Washington, D.C., USA. <https://www.atsdr.cdc.gov/spl/>.

Ambrosone, A., Mattered, L., Marchesano, V., Quarta, A., Susa, A.S., Tino, A., Rogach, A.L., Tortiglione, C., 2012. Mechanisms underlying toxicity induced by CdTe quantum dots determined in an invertebrate model organism. *Biomaterials* 33, 1991–2000. <http://dx.doi.org/10.1016/j.biomaterials.2011.11.041>.

Bailey, R.E., Smith, A.M., Shuming, N., 2004. Quantum dots in biology and medicine. *Physica* 25, 1–12. <http://dx.doi.org/10.1016/j.physe.2004.07.013>.

Balen, B., Tkalec, M., Šikic, S., Tolic, S., Cvjeto, P., Pavlica, M., Vidakovic-Cifrek, Ž., 2011. Biochemical responses of *Lemma minor* experimentally exposed to Cd and zinc.

- Ecotoxicology 20, 815–826. <http://dx.doi.org/10.1007/s10646-011-0633-1>.
- Basile, A., Sorbo, S., Conte, B., Gobianchi, R.C., Trinchella, F., Carginale, V., Sorbo, S., Conte, B., Gobianchi, R.C., Trinchella, F., 2012. Toxicity, accumulation, and removal of heavy metals by three aquatic macrophytes. *Int. J. Phytoremediat.* 14, 374–387. <http://dx.doi.org/10.1080/15226514.2011.620653>.
- Chan, W.C.W., Maxwell, D.J., Gao, X., Bailey, R.E., Han, M., Nie, S., 2002. Luminescent quantum dots for multiplexed biological detection and imaging. *Curr. Opin. Biotechnol.* 13, 40–46. [http://dx.doi.org/10.1016/S0958-1659\(02\)00282-3](http://dx.doi.org/10.1016/S0958-1659(02)00282-3).
- Chen, N., He, Y., Su, Y., Li, X., Huang, Q., Wang, H., Zhang, X., Tai, R., Fan, C., 2012. The cytotoxicity of Cd-based quantum dots. *Biomaterials* 33, 1238–1244. <http://dx.doi.org/10.1016/j.biomaterials.2011.10.070>.
- Domingos, R.F., Simon, D.F., Hauser, C., Wilkinson, K.J., 2011. Bioaccumulation and effects of CdTe/CdS quantum dots on *Chlamydomonas reinhardtii* – nanoparticles or the free ions? *Environ. Sci. Technol.* 45, 7664–7669. <http://dx.doi.org/10.1021/es201193s>.
- Drost, W., Matzke, M., Backhaus, T., 2007. Heavy metal toxicity to *Lemma minor* studies on the time dependence of growth inhibition and the recovery after exposure. *Chemosphere* 67, 36–43. <http://dx.doi.org/10.1016/j.chemosphere.2006.10.018>.
- Gadagkar, S.R., Call, G.B., 2015. Computational tools for fitting the Hill equation to dose – response curves. *J. Pharmacol. Toxicol. Methods* 71, 68–76. <http://dx.doi.org/10.1016/j.vascn.2014.08.006>.
- Gagné, F., Auclair, J., Turcotte, P., Fournier, M., Gagnon, C., Sauvé, S., Blaise, C., 2008. Ecotoxicity of CdTe quantum dots to freshwater mussels: impacts on immune system, oxidative stress and genotoxicity. *Aqua Toxic.* 86, 333–340. <http://dx.doi.org/10.1016/j.aquatox.2007.11.013>.
- Garczarska, M., Ratajczak, L., 2000. Metabolic responses of *Lemma minor* to lead ions I. Growth, chlorophyll level and activity of fermentative enzymes. *Acta Physiol. Plant* 22, 423–427. <http://dx.doi.org/10.1007/s11738-000-0083-5>.
- Gomes, S.A.O., Vieira, C.S., Almeida, D.B., Santos-Mallet, J.R., Menna-Barreto, R.F.S., Cesar, C.L., Feder, D., 2011. CdTe and CdSe quantum dots cytotoxicity: a comparative study on microorganisms. *Sensors* 11, 11664–11678. <http://dx.doi.org/10.3390/s111211664>.
- Gregor, M., 2004. Metal Availability, Uptake, Transport and Accumulation in Plants. Heavy Metal Stress in Plants – From Biomolecules to Ecosystems. Springer-verlag, Berlin, pp. 1–27.
- Gubbins, E.J., Batty, L.C., Lead, J.R., 2011. Phytotoxicity of silver nanoparticles to *Lemma minor*. *L. Environ. Pollut.* 159, 1551–1559. <http://dx.doi.org/10.1016/j.envpol.2011.03.002>.
- Hu, C., Liu, Y., Li, X., Li, M., 2013. Biochemical responses of duckweed (*Spirodela polyrrhiza*) to zinc oxide nanoparticles. *Arch. Environ. Contam. Toxicol.* 64, 643–651. <https://doi.org/10.1007/s00244-012-9859-z>.
- International Organisation for Standardisation. ISO DIS 20079, 2005. Water Quality – Determination of the Toxic Effect of Water Constituents and Waste Water to Duckweed (*Lemma minor*) – Duckweed Growth Inhibition Test.
- Ivask, A., Titma, T., Visnapuu, M., Vija, H., Käkinen, A., Sihtmäe, M., Pokhrel, S., Mädlar, L., Heinlaan, M., Kisand, V., Shimmo, R., Kahru, A., 2015. Toxicity of 11 metal oxide nanoparticles to three mammalian cell types in vitro. *Curr. Top. Med. Chem.* 15, 1914–1929.
- Jackson, B.P., Bugge, D., Ranville, J.F., Chen, C.Y., 2012. Toxicity, and bioaccumulation of quantum dot nanoparticles to the amphipod *Leptocheirus plumulosus*. *Environ. Sci. Technol.* 46. <http://dx.doi.org/10.1021/es202864r>. (5550-5536).
- Jamieson, T., Bakshi, R., Petrova, D., Poocek, R., Imani, M., Seifalian, A.M., 2007. Biological applications of quantum dots. *Biomaterials* 28, 4717–4732. <http://dx.doi.org/10.1016/j.biomaterials.2007.07.014>.
- Jiang, H.S., Li, M., Chang, F.L., Li, W., Yin, L.Y., 2012. Physiological analysis of silver nanoparticles and AgNO<sub>3</sub> toxicity to *Spirodela polyrrhiza*. *Environ. Toxicol. Chem.* 31, 1880–1886. <http://dx.doi.org/10.1002/etc.1899>.
- Juhel, G., Batisse, E., Hugues, Q., Daly, D., Pelt, F.N.A.M., Van, Halloran, J.O., Jansen, M.A.K., 2011. Alumina nanoparticles enhance growth of *Lemma minor*. *Aquat. Toxicol.* 105, 328–336. <http://dx.doi.org/10.1016/j.aquatox.2011.06.019>.
- Kaiser, J., Samek, O., Reale, L., Malina, R., Ritucci, A., Poma, A., Tromba, G., 2007. Monitoring of the heavy-metal hyperaccumulation in vegetal tissues by X-ray radiography and by femto-second laser induced breakdown spectroscopy. *Microsc. Res. Tech.* 70, 147–153. <http://dx.doi.org/10.1002/jemt.20394>.
- Kaiser, J., Novotný, K., Martin, M.Z., Hrdlička, A., Malina, R., Hartl, M., Adam, V., Kizek, R., 2012. Trace elemental analysis by laser-induced breakdown spectroscopy – Biological applications. *Surf. Sci. Rep.* 67, 233–243. <http://dx.doi.org/10.1016/j.surfrep.2012.09.001>.
- Kalčíková, G., Županič, M., Jemec, A., Žgajnar Gotvajn, A., 2016. The impact of humic acid on chromium phytoextraction by aquatic macrophyte *Lemma minor*. *Chemosphere* 147, 311–317. <http://dx.doi.org/10.1016/j.chemosphere.2015.12.090>.
- Kirchner, C., Liedl, T., Kudara, S., Pellegrino, T., Javier, A.M., Gaub, H.E., Stolze, S., Fertig, N., Parak, W.J., 2005. Cytotoxicity of colloidal CdSe and CdSe/ZnS nanoparticles. *Nano Lett.* 5, 331–338. <http://dx.doi.org/10.1021/nl047996m>.
- Koo, Y., Wang, J., Zhang, Q., Zhu, H., Chehab, E.W., Colvin, V.L., Alvarez, P.J.J., Braam, J., 2015. Fluorescence reports intact quantum dot uptake into roots and translocation to leaves of *Arabidopsis thaliana* and subsequent ingestion by insect herbivores. *Environ. Sci. Technol.* 49, 626–632. <http://dx.doi.org/10.1021/es5050562>.
- Krajcarová, L., Novotný, K., Kummerová, M., Dubová, J., Glöser, V., Kaiser, J., 2017. Mapping of the spatial distribution of silver nanoparticles in root tissues of *Vicia faba* by the laser-induced breakdown spectroscopy (LIBS). *Talanta* 173, 28–35. <http://dx.doi.org/10.1016/j.talanta.2017.05.055>.
- Kwan, K.H.M., Smith, S., 1991. Some aspects of the kinetics of Cd and thallium uptake by fronds of *Lemma minor* L. *New Phytol.* 117, 91–102. <http://dx.doi.org/10.1111/j.1469-8137.1991.tb00948.x>.
- Lee, J., Ji, K., Kim, J., Park, C., Lim, K.H., Yoon, T.H., Choi, K., 2009. Acute toxicity of two CdSe/ZnSe quantum dots with different surface coating in *Daphnia magna* under various light conditions. *Environ. Toxicol.* 25, 593–600. <http://dx.doi.org/10.1002/tox.20520>.
- Li, L., Sillanpää, M., Tuominen, M., Lounatmaa, K., Schultz, E., 2013. Behavior of titanium dioxide nanoparticles in *Lemma minor* growth test conditions. *Ecotox. Environ. Safe.* 88, 89–94. <http://dx.doi.org/10.1016/j.ecoenv.2012.10.024>.
- Lišková, M., Voráčková, I., Klepárník, K., Hezinová, V., Prikryl, J., Foret, F., 2011. Conjugation reactions in the preparations of quantum dot-based immunoluminescent probes for analysis of proteins by capillary electrophoresis. *Anal. Bioanal. Chem.* 400, 369–379. <http://dx.doi.org/10.1007/s00216-011-4700-5>.
- Navarro, D.A., Bisson, M.A., Aga, D.S., 2012. Investigating uptake of water-dispersible CdSe/ZnS quantum dot nanoparticles by *Arabidopsis thaliana* plants. *J. Hazard. Mater.* 211–212, 427–435. <http://dx.doi.org/10.1016/j.jhazmat.2011.12.012>.
- OECD guidelines for the testing of chemicals, 2002. *Lemma* sp. *Growth Inhibition Test* 221.
- Oukarroum, A., Barhoumi, L., Pirastu, L., Dewez, D., 2013. Silver nanoparticles toxicity effect on growth and cellular viability of the aquatic plant *Lemma gibba*. *Environ. Toxicol. Chem.* 32, 902–907. <http://dx.doi.org/10.1002/etc.2131>.
- Perreault, F., Popovic, R., Dewez, D., 2014. Different toxicity mechanisms between bare and polymer-coated copper oxide nanoparticles in *Lemma gibba*. *Environ. Pollut.* 185, 219–227. <http://dx.doi.org/10.1016/j.envpol.2013.10.027>.
- Peyrot, C., Gagnon, C., Gagné, F., Wilkinson, K.J., Turcotte, P., Sauvé, S., 2009. Effects of Cd telluride quantum dots on Cd bioaccumulation and metallothionein production of the freshwater mussel, *Elliptio complanata*. *Comp. Biochem. Physiol. Part C* 150, 246–251. <http://dx.doi.org/10.1016/j.cbpc.2009.05.002>.
- Požizka, P., Prochazka, D., Pilát, Z., Krajcarová, L., Kaiser, J., Malina, R., Novotný, J., Zemánek, P., Ježek, J., Šerý, M., Bernatová, S., Krzyžánek, V., Dobranská, K., Novotný, K., Trtílek, M., Samek, O., 2012. Application of laser-induced breakdown spectroscopy to the analysis of algal biomass for industrial biotechnology. *Spectrochim. Acta Part B Spectrosc.* 74–75, 169–176. <http://dx.doi.org/10.1016/j.sab.2012.06.014>.
- Prasad, M.N.V., Malec, P., Waloszek, A., Bojko, M., Strzałka, K., 2001. Physiological responses of *Lemma trisulca* L. (duckweed) to Cd and copper bioaccumulation. *Plant Sci.* 161, 881–889. [https://doi.org/10.1016/S0168-9452\(01\)S](https://doi.org/10.1016/S0168-9452(01)S).
- R\_Core\_Team, 2015. R: A Language and Environment for Statistical Computing. From. <http://www.R-project.org/>.
- Razinger, J., Dermastia, M., Dolenc, J., Zrimec, A., 2008. Oxidative stress in duckweed (*Lemma minor* L.) caused by short-term Cd exposure. *Environ. Pollut.* 153, 687–694. <http://dx.doi.org/10.1016/j.envpol.2007.08.018>.
- Resch-Genger, U., Grabolle, M., Cavaliere-Jaricot, S., Nitschke, R., Nann, T., 2008. Quantum dots versus organic dyes as fluorescent labels. *Nat. Methods* 5, 763–775. <http://dx.doi.org/10.1038/nmeth.1248>.
- Řezáčová, L., Rybáková, S., Lubal, P., 2015. Quantum dot-based fluorosensor for determination of Cu (II) and Pb (II) ions. In: Proceedings of the NANOCON: 7TH International Conference on Nanomaterials – Research & Application, pp. 650–654.
- Santos, D., Nunes, L.C., de Carvalho, G.G.A., Gomes, M. da S., de Souza, P.F., Leme, F., de O., dos Santos, L.G.C., Krug, F.J., 2012. Laser-induced breakdown spectroscopy for analysis of plant materials: a review. *Spectrochim. Acta Part B Spectrosc.* 71–72, 3–13. <http://dx.doi.org/10.1016/j.sab.2012.05.005>.
- Santos, S.M.A., Dinis, A.M., Rodrigues, D.M.F., Peixoto, F., Videira, R.A., Jurado, A.S., 2013. Studies on the toxicity of an aqueous suspension of C<sub>60</sub> nanoparticles using a bacterium (*gen. Bacillus*) and an aquatic plant (*Lemma gibba*) as in vitro model systems. *Aquat. Toxicol.* 142–143, 347–354. <http://dx.doi.org/10.1016/j.aquatox.2013.09.001>.
- Shi, J., Abid, D.A., Kennedy, I.M., Hristova, K.R., Silk, W.K., 2011. To duckweeds (*Landoltia punctata*), nanoparticulate copper oxide is more inhibitory than the soluble copper in the bulk solution. *Environ. Pollut.* 159, 1277–1282. <https://doi.org/10.1016/j.envpol.2011.01.028>.
- Škarková, P., Novotný, K., Lubal, P., Jebavá, A., Požizka, P., Klus, K., Farka, Z., Hrdlička, A., Kaiser, J., 2017. 2D distribution mapping of quantum dots injected onto filtration paper by laser-induced breakdown spectroscopy. *Acta Part B Spectrosc.* 131, 107–114. <http://dx.doi.org/10.1016/j.sab.2017.03.016>.
- Song, G., Gao, Y., Wu, H., Hou, W., Zhang, C., Ma, H., 2012. Physiological effects of anatase TiO<sub>2</sub> nanoparticles on *Lemma minor*. *Environ. Toxicol. Chem.* 31, 2147–2152. <http://dx.doi.org/10.1002/etc.1933>.
- Tkalec, M., Prebeg, T., Roje, V., Pevalek-Kozlina, B., Ljubec, N., 2008. Cd-induced responses in duckweed *Lemma minor* L. *Acta Physiol. Plant* 30, 881–890. <http://dx.doi.org/10.1007/s11738-008-0194-y>.
- Üçüncü, E., Özkan, A.D., Kursungoz, C., Ülger, Z.E., Ölmez, T.T., Tekinay, T., Ortac, B., Tunca, E., 2014. Effects of laser ablated silver nanoparticles on *Lemma minor*. *Chemosphere* 108, 251–257. <http://dx.doi.org/10.1016/j.chemosphere.2014.01.049>.
- Wang, J., Yang, Y., Zhu, H., Braam, J., Schnoor, J.L., Alvarez, P.J.J., 2014. Uptake, translocation, and transformation of quantum dots with cationic versus anionic coatings by *Populus deltoides* × *nigra* cuttings. *Environ. Sci. Technol.* 48, 6754–6762. <http://dx.doi.org/10.1021/es501425r>.
- Xu, M., Deng, G., Liu, S., Chen, S., Cui, D., Wang, Q., 2010. Free Cd ions released from CdTe-based nanoparticles and their cytotoxicity on *Phaeodactylum tricornutum*. *Metallomics* 2, 469–473. <http://dx.doi.org/10.1039/c005387m>.

## Article 26

P. Pořízka, J. Klus, E. Képeš, D. Prochazka, D.W. Hahn, and J. Kaiser. On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review. *Spectrochimica Acta - Part B Atomic Spectroscopy*, 148:65–82, 2018.

DOI: 10.1016/j.sab.2018.05.030; 16 citations

I have reviewed the utilization of the Principal Component Analysis (PCA) for the processing of Laser-Induced Breakdown Spectroscopy data. The article reflects the utilization of PCA and its variations (PCR and SIMCA) across applications and also gives overall comparison of the performance of selected PCA-based algorithms with their counterparts. The scope covers the data preprocessing, visualization, dimensionality reduction, model building, classification, quantification and non-conventional multivariate mapping. A list of recommendations is given as a conclusion to this article; namely:

- having a solid theoretical background in laser spectroscopy, statistics (incl. MVDA algorithms),
- optimizing experimental design and obtaining statistically valuable dataset,
- applying information-sensitive preprocessing steps,
- implementing sophisticated MVDA algorithms with awareness of over-training and providing detailed report on their implementation.

Those recommendations may be transferred also to data obtained using other analytical techniques.



## Review article

## On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review

Pavel Pořízka<sup>a,b,c,\*</sup>, Jakub Klus<sup>a,b</sup>, Erik Képeš<sup>a</sup>, David Prochazka<sup>a,b</sup>, David W. Hahn<sup>c</sup>, Jozef Kaiser<sup>a,b</sup>

<sup>a</sup> Central European Institute of Technology, Brno University of Technology, Purkyňova 656/123, CZ-61200 Brno, Czech Republic

<sup>b</sup> AtomTrace a.s., Koleční 9, CZ-61200 Brno, Czech Republic

<sup>c</sup> Department of Mechanical and Aerospace Engineering, University of Florida, 939 Sweetwater Drive, FL-32611 Gainesville, FL, USA



### ABSTRACT

An implementation of a fast, robust, and effective algorithm is inevitable in modern multivariate data analysis (MVDA). The principal component analysis (PCA) algorithm is becoming popular not only in the spectroscopic community because it complies with the qualities mentioned above. PCA is, therefore, often used for the processing of detected multivariate signal (characteristic spectra). Over the past decade, PCA has been adopted by the Laser-Induced Breakdown Spectroscopy (LIBS) community and the number of scientific articles referring to PCA steadily increases. The interest in PCA is not caused only by the basic need to obtain a fast data visualization on a lower dimensional scale and to inspect the most prominent variables. Most recently, PCA has also been applied to yield unconventional data analyses, i.e. processing of large scale LIBS maps. However, a rapid development of LIBS-related instrumentation and applications has led to some non-uniform methodologies in the implementation and utilization of MVDA, including PCA. Thus, in this work, we critically assess and elaborate on the approaches to utilize PCA in LIBS data processing. The aim of this article is also to derive some implications and to suggest advice in data preprocessing, visualization, dimensionality reduction, model building, classification, quantification and non-conventional multivariate mapping. This review reflects also other MVDA algorithms than PCA and consequently, presented conclusions and recommendations can be generalized.

### 1. Introduction

Sample characterization using Laser-Induced Breakdown Spectroscopy (LIBS) technique has been dynamically advancing in recent years. The parameters of conventionally utilized analytical instrumentation (lasers, spectrometers, and detectors) are being constantly improved. Moreover, the complicated or basic lab-built systems have been transformed to the sophisticated and commercially available systems, which enable an effortless and fast spectroscopic analysis. Contemporary state-of-the-art LIBS systems are capable of a high-end performance analysis (repetition rate, resolution, sensitivity). The high-end performance of LIBS is in certain cases superior to the performance of its analytical counterparts or reference techniques, such as Laser-Ablation Inductively Coupled Plasma (LA-ICP) based techniques, X-ray Fluorescence (XRF), etc.

LIBS is a well-established technique in many different applications, such as biology [1–4], geology [5], and industry [6]. The reason is the simplicity and robustness of the LIBS instrumentation together with its

capability of a fast-throughput multielemental analysis. Its potential has been repeatedly demonstrated by its high-end lab-based [7], *in-situ* and stand-off [8,9], and even extraterrestrial [10,11] utilization.

LIBS is one of the atomic emission spectroscopic techniques [6,12,13] based on the laser ablation sampling. Thorough articles were published with the aim to review the basic theory of the Laser-Induced Plasma (LIP) formation [14–16] and LIBS in general [17–20].

The introduction covering the basic theory about LIBS technique was brief because this review article targets namely the aspects of data processing. The reader should follow referenced books and review articles for more detailed background of LIBS theory prior to any further data processing through MVDA algorithms. As it was emphasized by Hahn and Omenetto [17]: “*advanced chemometric algorithms must be used with knowledge of what emission features (e.g. atomic or molecular emission peaks) are providing the associated discrimination.*”

A typical LIBS system is able to provide a high number of measurements (given by its repetition rate) when each measurement is described by a high number of variables (especially in the case of

\* Corresponding author at: Central European Institute of Technology, Brno University of Technology, Purkyňova 656/123, CZ-61200 Brno, Czech Republic.  
E-mail address: [pavel.porizka@ceitec.vutbr.cz](mailto:pavel.porizka@ceitec.vutbr.cz) (P. Pořízka).

echelle spectrometers). Note that the high repetition rate systems are mentioned for their leading edge in the LIBS applications, however, obtaining large number of measurements is not strictly related to LIBS systems with high repetition rate. The collected LIP spectrum is rich in information and represents the sample from which it originated, *i.e.* the chemical/spectral fingerprint of the sample [21,22]. The processing of large scale datasets is a demanding task which can be accomplished by using the so-called multivariate data analysis (MVDA; often related to as chemometric, exploratory data analysis or pattern recognition). It is noteworthy that unique LIP spectra are strongly affected by the matrix effect [19] which requires special attention when it comes to the conventional univariate calibration and quantitative analysis. On the contrary, the relation to the sample matrix enables a classification of samples according to their spectral fingerprints using simple MVDA algorithms. When processing large datasets, there are two more requirements to be met, namely, to process the data in the least possible time and in the most efficient manner. Efficiency can be measured by the conservation of variance during the dimensionality reduction, the sensitivity to outliers and the specificity to discriminate between individual matrices of analytes.

MVDA algorithms are massively spread throughout the LIBS community and are used in a number of applications. It may be stated that the future of the LIBS data analysis lies in the implementation of MVDA algorithms. The use of multivariate algorithms for processing of spectroscopic data has already been well documented [23–26]. Moreover, several review articles [5,17,27,28] dealt solely with the multivariate processing of LIBS data. A full chapter in the LIBS book by Cremers and Radziemski [12] was also dedicated to this topic. Based on the literature survey, the most popular MVDA algorithm in the LIBS community is the Principal Component Analysis (PCA). This simple linear algorithm provides powerful means of data visualization and pattern recognition on a lower-dimensional scale.

Based on our thorough literature research, the methodological approaches in the processing of LIBS data through MVDA algorithms significantly differ. This is given *i)* by the needs of a particular application, *ii)* by the uniqueness of the data acquisition and data size, *iii)* by the data topology, *iv)* by the variety of MVDA algorithms and also *v)* by the internal methodology of each research group. Consequently, there is not a unified approach and it might not exist in the future. Moreover, a wide range of MVDA algorithms together with the available software for the processing of data creates an option to perform a reachable and easy-to-use analysis. This might lead to the misguided implementation of these algorithms and software, *i.e.* when their use leads to aesthetic improvement of low-quality data (high fluctuation, low sensitivity, *etc.*) [29]. Nevertheless, it has to be stressed that a stable and optimized analytical system providing a reproducible high-performance analysis (high-quality data) should be the cornerstone of any experimental work. The same is valid for the understanding of the theory of *i)* LIBS (e.g. laser-ablation and plasma dynamics and its properties) and *ii)* MVDA algorithms and their considerate and judicious implementation in the data analysis process [17].

In this work we bring a summary of the most common approaches in the implementation of PCA in LIBS data analysis for: low-dimensional visualization, clustering, outliers filtering, variable selection, quantification, classification, and non-conventional multivariate mapping. Additionally, general suggestions for the data preprocessing and the model building, as well as a comparison with the performance of other MVDA counterparts, are given.

## 2. Multivariate data analysis

### 2.1. Data preprocessing

Prior to an implementation of any MVDA algorithm such as PCA and its variations, it is strongly advised to preprocess the obtained data [28,30]. Detected multivariate signal in its raw state is burdened with

unwanted background signal, fluctuation in the experimental parameters, *etc.* It has to be kept in mind that the data structure is changing during the data handling. This leads to consecutive changes in the performance of MVDA algorithm applied to the final data [31]. In general, there is a whole list of data preprocessing algorithms that should be considered prior to any MVDA. The order of the implementation of individual preprocessing steps is not given, therefore we propose the following one based on our experience and data processing algorithms found throughout the literature research. For the sake of brevity, individual steps are reviewed only tangentially. References to relevant, more detailed literature are given.

First of all, we advise to collect a high number of representative spectra per each sample to get robust statistics. Obtained data (the sample from the parent distribution) are organized into the data matrix  $X$  with  $n$  rows ( $n$  is the total number of spectra from all investigated samples) and  $p$  columns (variables related to wavelengths or even processed intensities of spectral lines). Harrington [32] also suggests checking the structure of the matrix and its rank.

Background subtraction mitigates the contribution from the background noise and any non-gated background continuum radiation. Van Veen and de Loos-Vollebregt [33] reviewed and compared several algorithms for the background correction of optical emission signal in ICP applications. Smoothing the spectrum is a possible way of correction [34]. Both preprocessing steps truncate certain amount of information that could be sacrificed for the sake of improved performance [35]. Most often, time-resolved detectors are used in order to avoid unwanted background from the continuum emission. Contrary to that, Myakalwar et al. [36] proved that it is possible to classify samples based on non-gated spectra detected using low-cost Czerny-Turner spectrometers and time-integrated detectors.

Row-wise spectra normalization, *i.e.* internal standardization [37–39] might reduce the influence of any potential fluctuations and dependence on the sample matrix. Most common ways of internal standardization (to matrix line intensity, to total spectral intensity, *etc.*) are described and compared in the aforementioned publications [40]. However, Colao et al. [41], with reference to Sirven et al. [42], suggested avoiding the utilization of matrix element for internal standardization, especially in cases when the content of the matrix element significantly changes throughout the sample set. Sirven et al. [43] proposed to use the background intensity as an alternative to the conventionally used standardization procedures. This approach was also applied in other works [41,44]. To mention another possibility for signal normalization, Death et al. [45] normalized spectra to regions of minimal variance. By contrast, they used the total spectral intensity in their further work [46].

The row-wise standardization is often followed by the column-wise mean centering and scaling. Column-wise mean centering is required to fulfill the underlying mathematical assumptions used to derive certain MVDA algorithms. In that case, mean value of each column (variable) is subtracted and thus moving the center of mass of data points (so-called scores) towards the beginning of a coordinate system. In the column-wise scaling, each variable is divided by its standard deviation and thus provides a sort of standardization of variables, *i.e.* the significance of variables is then unified. The process of mean centering and scaling is also referred to as standard normal variate (SNV). The whole process assumes that data follow the Gaussian distribution. However, this is not the case of LIBS data where data are also found under extreme value distribution [47,48]. Doucet et al. [49] compared impact of mean-centering and scaling on the performance of the principal component regression (PCR) and the partial least squares regression (PLSR) quantification. Porizka et al. [39] investigated the impact of column-wise mean centering and scaling in the case of steel, Al alloys and sedimentary ores data. They concluded that column-wise scaling should be avoided in the case of raw, unprocessed spectra that include background noise.

Detected spectra contain raw information described as the intensity

versus wavelength. In the most ideal case, the individual spectral lines should be assigned to elements and their relative intensities should be quantitatively processed prior to any exploratory data analysis. That leads to the significantly lower dimension of the variables based on the sample composition (*i.e.* content of individual analytes). The intensity estimation of the spectral line may be done in several ways [6,12,13]. Deriving the intensity of the analyte is more convenient from the perspective of further processing of the data matrix **X** by MVDA algorithms, *i.e.* the lower the rank of **X**, the faster the computation response. This approach is not so straightforward due to the necessary calibration of the system to individual analytes and related sample matrices. To make the preprocessing simple and convenient, the spectra are mostly kept in their raw state when introduced to MVDA, as it is evident in many publications, see Section 3.3.

Spectral lines are usually manually assigned by comparing the wavelengths to spectral libraries [50,51]. Those libraries contain a high number of spectral lines and the decision making is mostly up to the researcher who has an incomplete list of parameters (ionization, energy levels, Einstein coefficients or degeneracy levels). Apart from that, libraries specially dedicated to the LIBS purposes with limited selected spectral lines also exist [52,53]. The automatic approach in line assignment was also demonstrated by Amato et al. [54].

To deliver a complete set of steps in the data processing, we introduce also a brief discussion on the subsetting of the data set and an estimation of figure of merits. This is, however, not directly related to the implementation of PCA itself, but leads to PCA-based algorithms used in the classification (Soft Independent Modelling of Class Analogies - SIMCA) and quantification (Principal Component Regression - PCR). In the process of quantitative analysis and classification, the model is built using a training subset of the original data. Then the robustness of this model is cross-validated by a testing subset of data, usually complementary to the training subset. The dissection of the original data set into the subsets demands special attention. On one hand, the data set should be split into the set for modelling and for testing as suggested by Gottfried in the chapter dedicated to chemometrics [12], *i.e.* the “split-sample design” [32]. On the other hand, El Haddad [28] recommends to use another independent subset of original data - a validation subset - for the validation of the model in order to avoid its under- or over-fitting. Naturally, a model performs better when similar data are in both the training and the test sets. Nevertheless, Anderson et al. [55] avoided the over-optimistic assessment of results by assigning the similar spectra only to either the training or the test set. They tested five different ways to select the training set. The testing part should be followed by an estimation of figures of merit.

Pořízka et al. [56] repeated the random selection of the model and test data subsets and showed slight differences in resulting figures of merit. This study supported the urgent need for an elaborate selection of data in order to cover the natural fluctuation within the original dataset. The figures of merit serve mainly as a metric for comparison of LIBS systems and approaches suggested by individual research groups. Individual figures of merit are summarized elsewhere [17,28]. The original body of work by Voigtman [57] should be studied prior to the estimation of LODs.

## 2.2. Principal component analysis

If we consider investigating  $p$ -dimensional space by the 2D projections of pairs of variables then we will have to depict  $\frac{1}{2}p(p-1)$  scatter plots. However, that is a tedious and misleading process. As a response to this, PCA was introduced by Pearson in 1901 [58] with the aim to find lines and planes that could fit the set of points in the  $p$ -dimensional space. Later on, in 1933, Hotelling [59] derived the algebraic form of PCA similar to the one of Factor Analysis (FA). In both cases, PCA can be generalized to an algorithm as it is known today, *i.e.* visualizing complex multivariate data on a low-dimensional scale. Moreover, the suggested dimensionality reduction retains the most of the variance

carried by the original data. From that essence, PCA is used as an effective tool for the exploratory data analysis and pattern recognition. PCA provides a dimensionality reduction of the original data set by the generalization of the original variance. This is done by a transformation of the original high-dimensional space (wavelength as variables) into a smaller set of independent variables, *i.e.* principal components (PCs). The data projection by PCA concerns only the variances between the objects and their structure and it is unsupervised, *i.e.* there is no direct relation to, for instance, a sample class or the content of investigated analyte. From the geometric point of view [60], there is a line (the first principal component) going through the center of gravity of  $n$  points in the  $p$ -dimensional space and carrying the information of the highest variance within the respective data set.

The PCA algorithm is based on the covariance or correlation matrix, depending on the utilized algorithm, and therefore the data should satisfy certain assumptions [61]. Two main assumptions are *i)* **linearity**: all variables should be linearly correlated and *ii)* **normal distribution**: each variable should follow the normal distribution. However, both main assumption are not met in the LIBS experiment. The linearity of spectral response (intensity of analytical line versus analyte content) breaks at certain point. This issue is, for instance, well described in the review by Hahn and Omenetto [17]. The normality of data distribution was disproved by Klus et al. [48]. However, this fact is not considered throughout the LIBS literature. Thus, we will proceed with our review regardless of validity of mentioned assumptions.

PCA essentially provides three outputs: *i)* variance, *ii)* loadings and *iii)* scores. It is possible to create a PCA model of the original data matrix in several ways: the power method, Singular Value Decomposition (SVD), Nonlinear Iterative Partial Least Squares (NIPALS), *etc.* To overcome the problems of non-linear data, the kernel [62] or robust [63,64] PCA algorithms may be used, showing improved figures of merit in comparison with typical PCA. Other algorithms, for instance the Discriminant Function Analysis (DFA) [65–67], Independent Component Analysis (ICA) [68] and Self-Organizing Maps (or Kohonen Maps) [69] present an alternative to PCA and might be applied for the projection and visualization on a low-dimensional scale. Topological Data Analysis (TDA) is also promising in the data analysis, yet, to the best of our knowledge, still not applied for the visualization of LIBS data.

### 2.2.1. Variance and selecting the number of PCs

The data points are distributed in the  $k$ -dimensional ( $k$  being the number of constructed PCs) space and the data set may be described by a parameter called variance [70]. Considering the sample of points from a given population, the variance is given as a sum of standard deviations from the sample mean divided by the total number of objects minus one. Thus, as it is summarized by Kendal [60], the sum of eigenvalues may be estimated as the sum of squares of the distances of all the objects (data points, scores) from their mean. This sum may be assigned  $S$ , then the  $i$ th PC carries a portion of  $\lambda_i/S$  of the original data set's total variation.

PCs are arranged in descending order with respect to the variation that they carry. Therefore, only a first few of them are capable to satisfactorily describe the original data set. However, the judicious estimation of the optimal number of PCs depends on many factors, such as the structure and topology of the original data, the application in which they are going to be used and mainly the experience and expertise of the researcher. For that reason, the records of the number of PCs in published articles differ. Most often the optimal number is found to be in units of PCs but also extreme cases in tens of PCs exist, which is rather doubtful. In general, finding the optimal number of PCs leads to the tradeoff between the robustness and over-training of the model.

There are several ways where to cut-off the non-informative PCs [71]. The first one is the *variance explained criteria* [72] which gives a threshold to cumulative variance. Most often, the threshold is set to 90% or 95%. This threshold is set by the operator and can be varied at

his will. The *Cattell's scree plot* [73] depicts the explained variance versus the PCs. The plot is in descending manner and the cut-off is advised to be done in the elbow of the diagram, *i.e.* in the place where the steep part of the diagram levels-off and flattens. The *Kaiser's criterion* [74] dictates to cut-off PCs with the eigenvalue lower than 1. This is enabled by the fact that the average of all eigenvalues is equal to unity.

It is important that the loadings of retained PCs are investigated after the cut-off in order to check the information that they carry. It is reasonable to expect that the loadings of higher orders' PCs will describe only noise. Defernez and Kemsley [75] stress a potential misuse of MVDA algorithms by the overfitting of the model. In other words, the higher the number of components the better the MVDA analysis. But from a certain number of components the model gets over-trained and fails in predictions of unseen data. It has higher, yet illusive performance, giving the results that may be deceptive.

### 2.2.2. Loadings

The relation between the newly constructed PCs and the original data are represented by loadings values. Geometrically, the loadings describe a projection of the original space in the direction of the highest variation. Thus, the composition of each PC loading describes the most significant lines contributing to each latent variable. This information is fundamentally important in the line selection and related dimensionality reduction, see Section 3.3.

The loadings matrix  $\mathbf{P}$  has  $(m, p)$  dimensions. Loadings are usually plotted against the wavelengths. But in some articles [44,76,77] loadings were presented in the form of cross-plots, showing the relation of each wavelength variable in respective PCs.

### 2.2.3. Scores

When a PCA model of the original data is constructed, a certain number of PCs is selected giving a  $k$ -dimensional space. The scores are coordinates of the original objects in this newly constructed low-dimensional space. In contrast to the original space, which is organized in arbitrary order (determined by the experiment), PCs are organized according to the structure of the dataset (described by the variance). Thus, plotting of the most significant PCs quickly shows the structure of the dataset. The information from scores can be used in many applications as described in following sections.

The pattern recognition approach is based on the fact that the spectra which are mutually similar group together and form a joint cluster. Of course, it is supposed that, prior to the multivariate visualization, the spectra of each sample are highly correlated, having similar spectral fingerprint, and thus should share the same cluster. The size of the cluster is related to the fluctuation in the measurement (analytical system instability, local sample inhomogeneity, *etc.*). It may be stated that the higher the fluctuation, the broader the distribution of points in the cluster. Yet still, this projection is unsupervised and any conclusion drawn from this subspace (such as clustering or classification) reflects the real structure of the data, regardless of their origin, *i.e.* relation to individual sample.

## 3. PCA in LIBS

Advances in instrumentation development enable measurements with higher repetition rates, broader spectral ranges and better resolutions. Nowadays, an analysis results in datasets with thousands of variables [78] and millions of spectra [7]. Thus, the state-of-the-art LIBS system routinely provides big datasets (high number of spectra and variables) and so it is crucial to manage an effective and fast-response data processing. The MVDA algorithms must be applied into the analytical data correctly and with information-sensitivity.

When it comes to the exploratory analysis of spectroscopic data, PCA is the most often used algorithm. The first goal is to find similarities and patterns in the data that can lead to more complex conclusions about the sample set under investigation, *i.e.* clustering. However,

PCA is an unsupervised learning algorithm and thus the projection created by PCA has to be carefully evaluated. Moreover, PCA deals with the overall variation in the data and may be overloaded with the fluctuation or outliers present in the data set. Consequently, PCA might not be able to accurately cluster individual samples [27].

PCA and its alterations (PCR and SIMCA) have already been successfully implemented to the analysis of LIBS data in many applications. De Lucia and Gottfried [27] brought a brief review on the utilization of chemometrics in a LIBS data analysis of hazardous and geological materials, especially focused on the endeavors of the US Army Research Laboratory. Harmon et al. [5] reviewed a combination of LIBS and MVDA algorithms in geology. Some more general reviews on processing of LIBS data through MVDA algorithms [17,28] were also published.

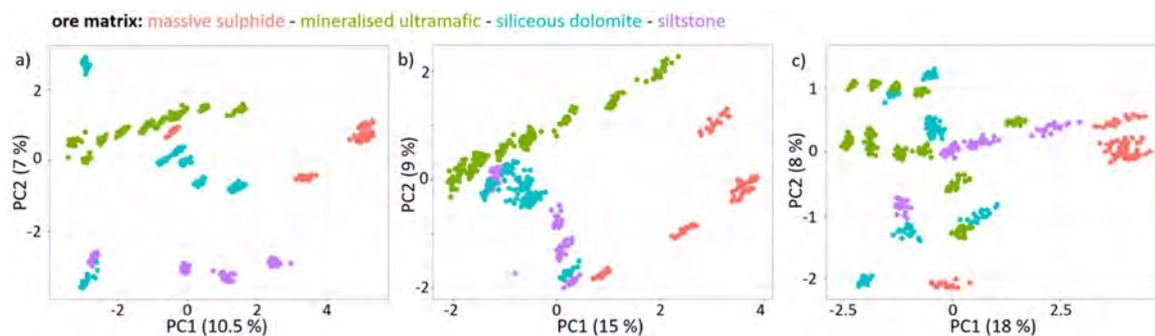
### 3.1. Data visualization

Structure, topology and quality of the data should be of interest prior to any further analysis. Therefore, a simple step of data visualization in a lower dimensional space is of great advantage. Most often, PCA is used solely for the visualization purposes and understanding of the fundamental relationship within the data, prior to the classification and quantification by other algorithms. However, De Lucia et al. [79] concluded a necessary condition for obtaining a meaningful clustering of LIBS data via PCA. This is that PCA is convenient when the intraclass variability is lower than the interclass variability, *i.e.* the discrimination gets worse when the intraclass fluctuation increases. Thus, other MVDA algorithms different from the PCA were also used for the low-dimensional visualization. Sirven et al. [80] showed that data in the PC space, *i.e.* their distances to the center of their respective cluster, follow a normal distribution. This is, however, in contrast to what was published in aforementioned articles [47,81] assuming the data to follow rather Extreme Value Distribution (EVD) over the normal distribution. Lazić et al. [82] suggested improving the linearity of data prior the creation of multivariate models.

The clustering of data points in the space given by selected original variables or PCs reflects the differences in unique and characteristic spectral fingerprints of individual sample matrices. In certain studies, the loading values were depicted directly in the PC space to show the significant variables in each PC [83,84]. It is obvious that a clustering of samples with distinct differences in their composition of matrix elements is quite effortless. Goode et al. [85] used PCA and the hierarchical cluster analysis (HCA) to show the possibility of clustering of LIBS data from various alloys and steel samples. Erdem et al. [86] used PCA to study archaeological pottery from eastern Turkey, suggesting some correlations between the composition of pottery and local sources. Gregoire et al. [87] used PCA to visualize polymer samples prior their classification by PLS. Vitkova et al. [88] explored clustering of data from various solid samples with PCA and then applied LDA and SVM for classification.

When LIBS is used in geological applications, it results in a large amount of data for the consequent MVDA analysis. Gottfried et al. [83] studied the topology of the minerals spectra in the PC space prior their classification using PLS-DA. Their spectra showed significant differences in detected spectral lines and relative intensities. A similar phenomenon was shown by Porizka et al. [89]. They presented a study describing, among other things, the differences in the visualization of minerals by PC space based on the broadband echelle spectra (200–1000 nm) and narrowband Czerny-Turner spectra (60 nm centered around 305 nm and 405 nm), see Fig. 1. The distinct separation and compactness of data clusters was retained. But the eigensystem, *i.e.* structure, of the data matrices changed accordingly to the composition of variables in the selected wavelength ranges. This is natural for data obtained by using different LIBS systems and/or under varying analytical settings.

Clustering and classification of material based on already established data libraries would be of great interest to many research groups.



**Fig. 1.** This figure depicts a difference in PCA projection based on data from a) broad-range echelle spectra (200–1000 nm), and short-band (60 nm window) Czerny-Turner spectra at b) 305 nm central wavelength and c) 405 nm central wavelength. Data represent 28 sedimentary rocks in four different matrices. Obtained from [89] with permission provided by Elsevier and Copyright Clearance Center, license number: 4347640902794. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Lanza et al. [90] provided a discrimination of carbonate minerals in the PCA space; data were measured using the MSL Curiosity. Moreover, such analysis naturally leads to a provenance study. McManus et al. [91] used PCA for a provenance study of gemstones from 3 US states. The major concern of their provenance study was a diverse database covering varieties of geological materials including complex metadata (including original GPS locations of collected samples). That was, however, a very bold idea that would be hardly met in reality. Alvey et al. [92] executed an extensive provenance study of 157 garnet minerals from 92 world-wide locations and arranged them in a reduced PC space.

Studying organic material and polymers is also of considerable interest, however, the discrimination of related data depends only on the ratios and relative intensities of major elements (C, H, N, and O) and molecular bands (CN and C<sub>2</sub>). This aspect makes the classification of such samples very challenging, moreover, some more sophisticated and non-linear algorithms are needed. Despite that, De Lucia and Gottfried [93] showed a possibility to discriminate spectra of explosives samples in PC space (3 PCs accounting for 92% of variance). Labbe et al. [94] presented a PCA of biological samples where the first two principal components stood for 99% of variance; the first PC represented 98% and the second only 1% of total variance. The influence of the substrate on the bacterial spectra was also studied [84,95]. It was shown that PCA was capable to discriminate bacteria of different strains, even when the samples were analyzed on different surfaces. During the sampling of thin films (such as bacteria strains spread on the substrates) a certain amount of the substrate itself was also ablated and its characteristic radiation contributed to the detected spectra. Therefore, it was necessary to filter this contribution out because it had no particular analytical meaning and could be confusing for the consecutive MVDA.

Another interesting conclusion concerns the analysis of biological samples and polymers by LIBS in general. Spectra of those samples reflect compositions that form complex molecules namely in C, H, N and O in the samples. Thus, the techniques providing molecular information (such as Raman spectroscopy) are more suitable for this kind of analysis. Despite that, LIBS proved to be a powerful alternative. Landström et al. [96] applied PCA directly on echellograms (CCD images of plasma characteristic radiation spectrally resolved via an echelle spectrometer) to discriminate biological and chemical warfare agents.

Tandem systems (such as the complementary LIBS and Raman spectroscopy techniques) are often used to gain more information (elemental and molecular composition) from the sample [97]. This, in turn, improves the MVDA algorithms' performance. Hoehse et al. [98] used PCA to visualize inks and pigments measured by LIBS and Raman spectroscopy, and a similar analysis was done by Prochazka et al. [99]

on bacteria samples. In both cases, the LIBS and Raman spectra were simply merged together. Naturally, the combination of LIBS and Raman spectra improved the visualization and classification of samples.

It is also worth mentioning that PCA was used for a visualization of biological warfare agents based on the typical 1D spectra (intensity vs. wavelength) and 2D echellograms (images downloaded directly from the CCD detector on an echelle spectrometer) [100]. Porizka et al. [78] further developed the concept of echellograms proposed by Larsson et al. [100]. They proved the possibility to use directly the truncated information obtained from a CCD detector and to accelerate the data processing. In this kind of analysis it is not necessary to slow down the computing process by downloading and transforming the echellograms to spectra. Transforming a large amount of data is the bottleneck of the analysis when it comes to a potential utilization of echelle spectrometers in the high-repetition rate experiments. In fact, only a limited number of pixels from the whole megapixel image is necessary. Such an approach may lead to a significant enhancement of the repetition rate with an intensified detection of broadband spectra (using echelle spectrometers and sCMOS detectors).

Lasue et al. [101] compared PCA, ICA and Sammon's map (SM) for a visualization of ChemCam data. Of course, non-linear SM proved its supremacy over the other algorithms but this supremacy appeared only because of inconsiderably selected figures of merit. They used the stress function by applying the distances between points of the same cluster (the stress value of 26% for PCA, 17% for ICA and 4% for SM). Nevertheless, this indicator (stress value) is unfair because the SM model of the original data space is provided by finding the global minimum of the stress function. Moreover, no optimization of the number of PCs or ICs is done and the methods are compared strictly in the 2D space. Harmon et al. [102] used PLS-DA latent variable to study the clustering distribution of minerals spectra. Klus et al. [103] used SOMs in a similar analysis to show clustering of spectra of similar minerals in the grid of a self-organized map.

### 3.2. Multivariate elemental mapping

Mapping of elements and their distribution on the sample surface by using LIBS is a well-known method [1]. In the elemental mapping, a location on the sample surface, *i.e.* on the map, is represented by a relative intensity of selected spectral line of the element or even by the content of this element. But, most recently, the application of PCA has transformed the data analysis from a basic visualization in a reduced dimensional space to more sophisticated approaches. When the scores of selected PCs are depicted in the map (hereinafter referred to as multivariate map) it can represent a distribution of elements' combinations or even of individual matrices. This approach gives an

additional insight into a sample composition and a relation of individual phases (depending on the sample and the structure of original data). The meaning of PCs is then described as the original data by respective loadings and their spectral structure. The surface mapping of heterogeneous samples surfaces is in the spotlight of geological analysis, while it enables an elaborate understanding of individual phases and their distribution on the sample surface. This is due to the convenience in sample form (hard solid matrix) and its fast, straightforward preparation (cutting and polishing). Nevertheless, this approach in multivariate mapping of geological-sample surfaces may spread also to other applications.

Amador-Hernandez et al. [104] used LIBS to analyze printed electrodes in 3D. The mapping of the surface was provided in  $25 \times 13$  spots and 10 consecutive layers. Intensities of Ag, Al, Au, C, Pd, Pt, and Ti spectral lines were estimated and proceeded to a PCA for further dimensionality reduction of data from layers 1, 3, 6 and 10. Cluster analysis (k-means clustering) was then applied on selected PCs, separately on data from each individual layer. From each layer, up to 3 PCs covering joint variance from 88.6% to 94.1% were selected. Data points were then assigned to a defined number of classes and depicted in the map showing the distribution of materials rather than elements in the investigated layers. In their following study [105], authors further worked with the same data set showing only PC scores biplots with distinct clusters.

In the work by Klus et al. [81], the distribution of uranium content in a sandstone-hosted uranium ore was investigated using maps of uranium line distribution (U II 409.01 nm), variation in the intensity of background (find out more about this in [106]), and values of the first principal component. All the maps showed a high degree of mutual numeric correlation. Thus it was found out that the first principal component (with 82% of variance) is namely representative of the uranium content variation. Another step in the MVDA analysis was done in order to reduce the original data set for PCA modelling and in turn to reduce demands on the computing power and computing time. The original data set was truncated to ~11% (randomly selected spectra) which resulted in creation of a model. The PC scores of the rest of the data set were then predicted based on the PCA model. A predicted map which was consequently created correlated well with other maps. This approach turned out to be a promising way of data set truncation with preserving the information carried by the original data.

Romppanen et al. [107] employed the PCA algorithm (SVD in fact) in their study of the rare earth elements distribution on the surface of yttrium-bearing ore minerals. They depicted maps of first 4 PCs (covering together 99.5% of variance). Individual PCs were related via their loadings to four sets of elements (Al and Ca; Ca, Y, Fe, and Al; Si, Fe, and Y; Fe and Y respectively to each PC). In the next step, data points were plotted in the PC space given by 3 components (PC 2–4) showing the clusters of individual matrices. It was assumed that spectra would cluster in this 3D space based on their similarities. Data points were then colored by their position in this artificial directional cube and replotted in the map. Finally, the individual clusters were assigned to minerals present in the sample. Despite a rather demanding data processing, the obtained results represented well individual minerals with a focus on yttrium distribution.

Moncayo et al. [108] performed a study similar to those already mentioned. They tested the performance of PCA in an analysis of megapixel elemental maps - a data set compiled of more than 1 million LIP spectra, see Fig. 2. The sample was highly heterogeneous and contained matrices of turquoise, pyrite and silica-bearing mineral. It was once again shown that the stand-alone PCA is capable to dissect the multidimensional signal and to assign various minerals (matrices) to individual PCs. The contribution of Si and Ti-bearing minerals was found. Those minerals were present in low amounts and distributed on the sample surface. They were present in high order PCs.

The presented PCA approach, or the mapping using MVDA algorithms in general, may be further extended by using other linear (ICA,

TDA) and even non-linear (SOM) MVDA algorithms. Special attention will be given to SOM because they provide a powerful alternative to PCA [69]. Klus et al. used SOM for multivariate mapping of U-Zr-Ti-Nb in sandstones-hosted uranium ores [103]. The total number of 22,500 spectra was modeled, based on their similarities, into 900 neurons. Neural responses were then correlated with individual spectra and depicted in the original map. That showed the isolation of individual matrices (Si-, Ti- and U-based minerals) and also the affiliation of different elements (U, Zr, and Nb). A similar work was done by Pagnotta et al. [109] in the study of ancient mortars. LIBS maps of selected elements were processed through the SOM algorithm to achieve a segmentation of different materials present on the sample surface.

Another non-conventional utilization of PCA was presented by Lanza et al. [110]. They used a combination of depth-profiling LIBS data with PCA for the study of weathered surfaces of basaltic rocks. The PC scores showed a clear dependence on the depth-profile along the first PC. The investigation of PC loadings suggested that the amount of Ca, Na and Mg was increasing when going through the weathered layer and *vice-versa* the content of Mn and Ba was decreasing. Such an application proves the versatility and flexibility of the data analysis done by using PCA.

The additional value of multivariate mapping implemented to the LIBS applications is the fast response and robust processing of MVDA algorithm. Multivariate mapping may be a new, progressive way in the exploratory data analysis, enabling to process large data sets without prior knowledge. Inferred information is clear and can be directly used for deciphering of the contribution of different matrices and their mutual relations.

### 3.3. Dimensionality reduction

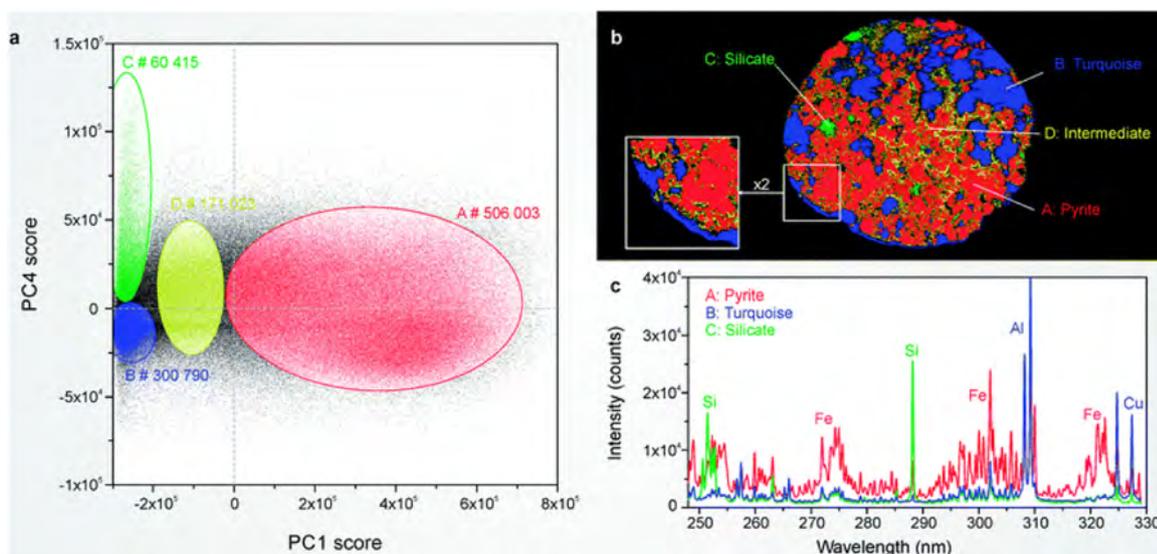
A typical size of obtained data matrix ( $n \times p$ , i.e. number of objects vs. number of variables) varies from experiment to experiment. First of all, the total spectra set is given by the number of samples and the number of detected spectra per sample and are usually collected *ad hoc* for feasibility studies. Total number of spectra per sample set is rather low, in many cases not even reaching 100 samples and 1000 spectra, as it is demonstrated by data collected from reviewed articles and depicted in the Fig. 3. The reason for that is that the experiment is usually done using low repetition rate LIBS systems and obtaining large data sets is thus tedious. From a statistical point of view, large data sets are more convenient. With higher repetition rates the number of objects (obtained spectra) can grow over thousands. In the case of aforementioned multivariate elemental mapping, Section 3.2, the number of objects was in the range of  $10^5$  to  $10^6$ ; giving millions of pixels in the elemental map.

The total number of variables depends on the spectrometer used during the experiment. Going to extremes, the number of variables can reach over 40,000 with echelle spectrometers. Overall, the improvement in the instrumentation results in higher number of objects and variables. This is a shift in the paradigm and a new challenge for data processing.

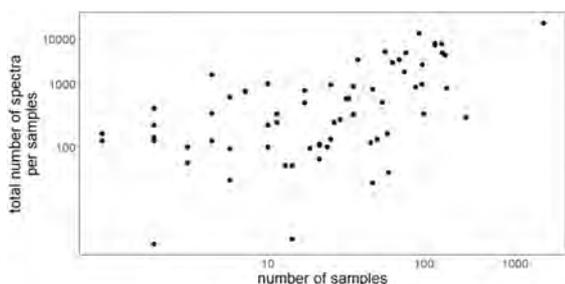
In the most extreme case, the size of the data matrix obtained during a LIBS experiment can be in the range of  $10^6$  objects  $\times$   $10^4$  variables. Processing such data matrix is then demanding on computing power and improvement in this aspect is beneficiary. Therefore, in following chapters, we are going to dissect individual approaches in data processing with special attention to a dimensionality reduction in objects and variables.

#### 3.3.1. Dimensionality reduction in variables

The literature research does not reveal any decisive viewpoint which would indicate how the number of variables should be reduced, or if the reduction should be done at all. Keeping the whole spectral range without any selection of spectral lines and processing of peak intensities can lead to increased computation demands. It also



**Fig. 2.** This figure depicts the processing of a megapixel elemental map. Data were collected throughout the analysis of a heterogeneous turquoise mineral. a) Scatter plot of PC4 vs. PC1 with highlighted clusters of data points; clusters represent a similarity in spectral fingerprint of the highlighted spectra. b) Clustered data are further depicted in the pseudo-colored map to respective location demonstrating the distribution of individual matrices within the heterogeneous sample. c) Spectral fingerprints of selected matrices: pyrite, turquoise and silicate. Obtained from [108] with permission provided by Royal Society of Chemistry and Copyright Clearance Center, license number: 4347680572130. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 3.** This figure depicts the total number of spectra over total number of samples collected from 70 articles reviewed in this paper.

introduces a lot of noise and some correlated, redundant variables into the model. It is indisputable that the background information may be beneficial in certain extreme cases. Klus et al. [81] proved the importance of background signal in the multivariate mapping of uranium. The detection of uranium is challenging due to the high density of lines that exceed the resolution of a typical spectrometer. Thus, its content in the sample may be correlated with systematic variations of background signal.

The utilization of echelle spectrometer gives the possibility to detect abundant broad-range spectra. Such analysis can be provided also using broad-range Czerny-Turner spectrometers, however with significantly lower resolution compared to their echelle counterparts. In general, such spectrometers are beneficially utilized in the LIBS experiments and provide extensive data sets for MVDA. On the contrary, Colao et al. [41] suggested selecting the most significant lines and a related proper wavelength range when using a short-band Czerny-Turner spectrometer. A similar study was introduced by Pořizka et al. [89]. It was shown that the utilization of short-band Czerny-Turner spectrometer instead of broad-range echelle spectrometer leads to the same classification power

when the spectral window is judiciously selected. Moreover, limits of detection of selected analytes improved in this experiment. Sirven et al. [111] used PCA to discriminate between three types of soil prior further classification and a quantitative analysis by ANN. They showed that a short spectral window, 10 nm, is sufficient for the discrimination of data in the PC space. However, shorter band widths led to a low performance.

Another possible way is to filter out the background noise by thresholding. The level of threshold may be set by the  $3\sigma$  limit of detection. This approach is not universal and does not consider low intensities of the trace and minor elemental lines. Moreover, detected intensity below  $3\sigma$  level is not accepted by the spectroscopic community as an analytical signal and should not be used in further data processing. Thus, there is no reason to associate this under limit intensity as a contribution from any particular element. Sahoo et al. [112] set manually the threshold to filter the noise from the spectra. Kong et al. [113] compared the classification performance after dimensionality reduction by a selection of spectral lines and a reduction into PCs. A similar comparison was provided by Lee et al. in the case of classification of salts using LIBS [114] and tandem LIBS and LA-ICP-MS techniques [115].

Especially in the case of broad-range echelle spectra, the number of the informative lines might be overloaded by the number of variables representing only the background noise. In the worst scenario, such data could lead to a confusion of the learning method, such as PCA, and consequently to its poor performance and to biased results [116]. Bousquet et al. [76] applied PCA on the full spectra, regardless of any information carried by the background and considering good SBR and SNR. This fact may be acceptable assuming that the systematic change in strong spectral lines intensities will simply overshadow the contribution from randomly fluctuating background noise. On the contrary, in their former study [44], they recommended to select only a small part of the spectra, i.e. relevant information from matrix element lines (68 lines in total; 4 for Al, 10 for Ca, 31 for Fe, 3 for K, 4 for Mg, 2 for

Mn, 11 for Ti, and 3 for Si). Darby Dyar et al. [117] showed that PLSR was unsuccessful in the quantification of sulfur while using the whole spectral range. This was then confirmed by the investigation of PCA loadings where the contribution of sulfur was overshadowed by other more significant lines. However, from the perspective of this review, this could be avoided or mitigated by the selection of spectral lines and their scaling prior to MVDA. The scaling could then lead to the enhancement of the sulfur significance in the eigensystem of the data matrix.

A rigorous wavelength selection and manual elemental line assignment is time consuming and tedious. Automatic algorithms designed especially for LIBS applications are still rather lab-made and commercially not available. Processed line intensities combined with their ratios can be used but this process demands prior knowledge of the sample composition. However, the qualitative analysis concerns the “usual suspects” and an experienced spectroscopist can predict the composition in major elements and identify the most common spectral lines typically detected in LIP spectra. Sirven et al. [80] also suggested selecting only spectral lines of major elements because they are supposed to dominate the MVDA performance. The suggestion was then not to consider the spectral lines from minor and trace elements. On the contrary, in their other work [42], they recommended to include spectral lines of trace elements in the data set. Moreover, they compared the classification power of the system based on spectra with the artificially lowered resolving power (from the original 10,000 to final 2800). After a variable down selection, the physical interpretation of the loadings was much easier. They reduced the number of variables to 41 lines of 10 elements based on the PC loadings in consecutive steps of a global PCA approach. Then the number of variables was limited even to 10 lines (one line per each element) with a high classification accuracy proving that a low number of variables is not limiting when they are judiciously selected.

Gottfried et al. [83] used PCA for a significant feature identification of mineral spectra, naturally lines of Ca, Si, Na, and K together with Sr, Li, and Mg were identified. Lewis et al. [84] cross-plotted loadings of PCs to see the impact of individual variables on the variance describing the data set. Then, variables with a low impact were discarded from further computation. The authors showed that the contribution of Ca, Mg and trace elements had a significant impact on the analysis of bacteria. Pereira et al. [118] used PCA loadings to select major constituents, macro- and micro-nutrients responsible for the discrimination of spectra from LIBS analysis of plants, citrus leaves.

Munson et al. [119] investigated various approaches in dimensionality reduction in order to improve the classification capability between bacterial and chemical warfare agents. They used a) six lines (P, C, H, and O) and combination of their ratios and b) simple masks omitting the whole spectral range except the regions around selected lines. The approach using various ratios of line intensities brings a way to artificially increase the number of information-carrying variables. This may in turn improve the performance of MVDA algorithms. Moreover, the selection and processing of spectral lines enables to scale the variables and thus to unify their impact on the performance of MVDA. Anderson et al. [55] used five different approaches to improve the quantitative performance of PLSR. This was achieved *via* pre-selection of training and test subsets using PCA prior to k-means and hierarchical clustering, and SIMCA. They have simulated the Mars conditions (ChemCam system, pressure, distance) for over than 220 ore samples. They used the full spectral ranges despite the statement which claims the improved performance after variable selection. Anabitarte et al. [120] used kernel PCA to reduce dimensions prior to SVM in order to detect Al impurities in steel coatings. A relatively high number of 100 PCs was used in this study, then this number was further reduced to 8, 5, 3 and 1 PC showing that at least 3 PCs were necessary to achieve accurate results.

Selecting the right number of PCs is also of great interest in this case. This has to be done in order to avoid any overtraining and biased

results. Typically, the high number PCs cover mostly only noise and, therefore, are usually omitted from computation. Marcos-Martinez et al. [121] stated that PCA for dimensionality reduction actually led to a decrease of the ANN's analytical performance. However, this is in contrast with aforementioned papers and should be accepted with care.

Variable Importance in Projection (VIP) presents a powerful way to identify the most prominent lines. The VIP is connected to factors provided by PLS models of the original data [122]. De Lucia et al. [123] studied the influence of the number of variables *versus* the whole spectra approach in the classification process of explosives spectra through PLS-DA. The set of variables was chosen by VIP having a score higher than 1.0. Naturally the contribution of C, H, N, O and molecules of C<sub>2</sub> and CN were the most significant in the analysis of explosives. After this selection, the number of selected variables was extended by their ratios, sums and combinations. It is noteworthy that this process is not fully correct from the mathematical point of view and should be approached cautiously. Regardless of our comment, this process was done twice, in the first case they got 40 variables and in the second case 177 variables. In the latter case, the ratios and combinations of variables followed stoichiometry and relative amounts of C, H, N, and O in the molecular bonds of analyzed samples. It was shown that the latter model had performed better than the first one, which proves the necessity to incorporate a higher degree of variability. The whole spectra model gave the best results, however, its performance was doubted by the authors due to the fact that the VIP procedure indicated a possible classification by factor not related to the contribution of C, H, N and O. Fink et al. [124] used a genetic algorithm to select variables prior to PCR and PLSR of analytes (such as Sb, Sn, and Ti) in plastic samples. They selected from 9 to 87 variables or the whole spectral ranges (up to 200 nm) for the PCR and PLSR models. The performance of the models based on the whole spectral ranges was significantly inferior. This once again proved the necessity to preprocess the data appropriately prior to any MVDA.

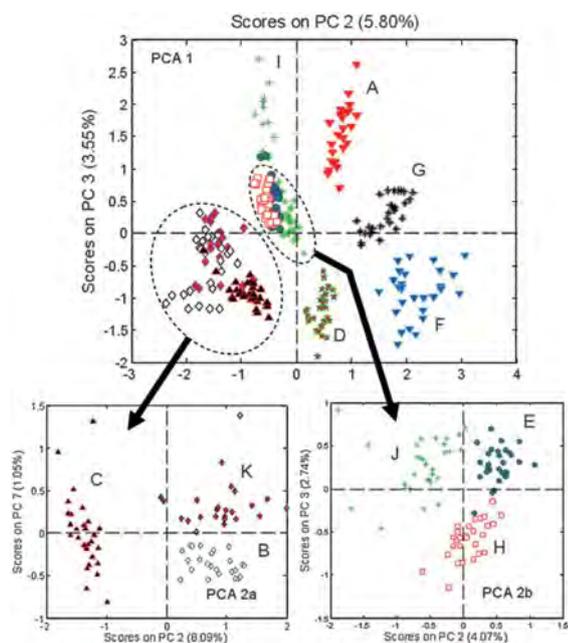
Corsi et al. [125] introduced an interesting way of the dimensionality reduction, similarly it was suggested in Section 3.1. The content of 11 elements within the archaeological samples (copper-based artefacts) was estimated by the Calibration Free (CF) LIBS [126]. The reduced data set (13 objects described by 11 variables) was then introduced into the PCA (66.3% of joint variance in the first two PCs). This is in contrast to typical analysis of raw LIP spectra.

PCA provides the dimensionality reduction of the original data set to a smaller set of independent variables but preserving the most of the variance. This essential feature of PCA may then be used primarily for data processing while the lower number of variables can fully represent the original data set. It is necessary to overcome the curse of dimensionality and select a right number of latent variables [127]. Then the selected number of PCs can be further used for classification and quantification purposes. A review of articles in which the authors implemented the reduced PC space for an exploratory data analysis and consecutive unsupervised classification is given in Section 4.

### 3.3.2. Dimensionality reduction in objects

The distinct separation of individual data points in PCA space suggests that samples can be classified based on characteristic LIBS spectra. However, PCA focuses on variances among data. Thus, any significantly different data overload the discrimination power of PCA and some smaller differences among the rest of the data are mitigated.

Multari et al. [128] suggested an approach to overcome this problem. Any well-discriminated and distinctly outlying cluster of data (spectra of a sample) is removed from the data set. Consequently, PCA is applied once again on the truncated data set when the differences between the formerly overlapped data points in the PC space are now revealed. This step-by-step approach may be regarded as a sort of dimensionality reduction in the sense of number of objects. This in turn reduces the variation in the data and changes significantly the data structure and topology. A US patent was issued [129] claiming



**Fig. 4.** This figure shows the capability to provide a provenance study. The distribution of yellow cake spectra in PC space suggests a possibility to classify samples according to their location of origin. Moreover, PCA was calculated three times; first for the complete data set (top figure) and then separately for both highlighted regions (bottom figures). This approach demonstrates the increased performance of PCA when the number of objects is reduced. Thus, the PCA algorithm is not overloaded with variance from a high number of sample matrices. Obtained from [42] with permission provided by Royal Society of Chemistry and Copyright Clearance Center, license number: 4347711143996. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

mentioned data processing steps. Before this patent, Sirven et al. [42] suggested a so-called global PCA that works in step-by-step mode for estimation of geographical origin of uranium ores, see Fig. 4. This algorithm was used for a mitigation of the number of data and a consequent implementation of MVDA to reduced space; this was done similarly to the approach patented by Multari. This successive approach was then adopted by Colao et al. [41] for the identification of historical building materials.

Anderson et al. [55] studied a possibility to reduce the number of samples in order to reduce the computation burden. Only samples lying furthest from a cluster center were selected and used for a furthest-neighbor clustering tree. Other samples were discarded from modelling. Thus, this approach retains the biggest amount of variation in the data and at the same time avoids redundancy. The authors also reduced the original data set by PCA prior to the construction of a dendrogram. Despite their efforts, the most reliable results were obtained when they used the full training sets, covering quite a large amount of variation and diversity. Fink et al. [124] claimed to use PCA to select the samples with the highest statistical leverage for the calibration purposes. However, this process was not further commented.

### 3.3.3. Outliers filtering

It is crucial to preprocess the data prior to any further implementation of MVDA algorithms, Section 3.1. Outliers filtering should be also considered; especially when the internal standardization is not sufficiently effective in mitigating the fluctuations of analytical signal.

In that respect PCA is a powerful candidate because its essence is being sensitive to high order of variation, which might lead to detection of outliers.

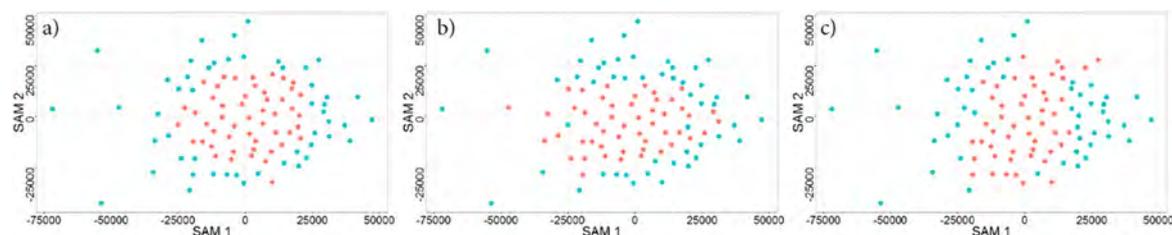
In general, it could be assumed that spectra detected from a shot-to-shot analysis of the homogeneous sample should possess a high degree of mutual correlation. Each spectrum is described by a unique set of values, *i.e.* descriptors, (correlation coefficient, Euclidean distance to the center of PC space, total energy of the spectrum, intensity of matrix lines, etc.). Selected descriptors represent analyzed spectra and relate them to the rest of the data set. The descriptors also enable a comparison of analyzed spectra with the rest of the data set. Those values are then compared and thresholded (either visually or numerically) which leads to a truncated data set with more similar spectra. This high mutual correlation may be demonstrated by a compact cluster of points in the PC space. The threshold level should be set carefully. Harder filtering should be avoided because it significantly influences the topology of the data, especially in the case of modelling. Fluctuation is natural to certain extent and describes the response of analytical system [32]. Therefore, only the most outlying points should be filtered.

Signal fluctuation is of great concern in LIBS and is a result of many independent parameters influencing the laser-matter interaction, which is usually concluded as the matrix-effect. The source of fluctuation is laser temporal instability, local heterogeneities on the sample surface, etc. The process of material ablation and consecutive plasma formation is therefore not stable from shot-to-shot perspective. The properties of laser-induced plasma (described by its temperature and electron density) are significantly affected together with emitted analytical signal. Those aspects increase the standard deviation (variance) of data points in the  $p$ -dimensional space and decrease the performance of the MVDA technique, which is then reflected in the resulting figure of merits.

Outliers filtering is a possible preprocessing step mitigating the analytical signal fluctuation. But, as it is typical in LIBS applications, the experimental parameters and protocols in data processing are optimized *ad hoc* and various approaches and algorithms are used. This leads also to non-unified ways of outliers filtering such as linear correlation [130], dendrogram [112], Student's  $t$ -test after the computation of the Mahalanobis distances of individual spectra in the original data space [131], total energy of each spectrum [132], or intensity of matrix lines [133,134]. PCA for outliers filtering was introduced into the processing of LIBS data by Sirven et al. [80]. They used basic visual thresholding and omitted up to 30% of spectra prior to a rocks classification in a preflight ChemCam testing. In their further work [42] they judged outliers based on the abnormal value of their residual fit. Myakalwar et al. [36] used a dendrogram after the PCA dimensionality reduction in order to detect outliers in the analysis of pharmaceuticals. Porizka et al. [77] improved the quantitative analysis of Cu in igneous rocks by filtering based on the distribution of data points in the PCA space. Prochazka et al. [99] applied PCA on the set of measurements of the same sample. In the space given by PC1 and PC2, the most outlying points (those with the biggest distance from the weighted center) were omitted and PCA was recalculated. This procedure was repeated until the selected number of the most similar measurements remained.

Porizka et al. [56] compared impact of three different approaches for outliers filtering (PCA, linear correlation, and total energy) on classification based on the selected figures of merit (overall accuracy, sensitivity, and specificity). Any of those approaches leads to an improvement in the classification accuracy. The total energy proved its supremacy over PCA and the linear correlation. It is, however, necessary to mention that the case study involved only steel samples. It was also found that each approach led to filtering of different data points (spectra), see Fig. 5. That, in turn, affected the topology of filtered data and also the differences in resulting figures of merit. Therefore, it should be advised to compare different approaches when samples of other matrices (other than steel) are subjected to LIBS.

It needs to be stressed that outliers filtering should be applied only to the classification or quantification (calibration and regression) of



**Fig. 5.** This figure depicts a distribution of data points filtered using a) PCA, b) linear correlation, and c) total intensity approach. Depicted in the first two components of Sammon's map, data left for further data processing are marked in salmon-pink and outliers are marked in cyan. The Figure was obtained from [56] with permission from Elsevier and Copyright Clearance Center, license number: 4347640902794. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

homogeneous sets of samples. In the case of multivariate elemental mapping, it is naturally counterproductive to omit outliers from a data set which describes heterogeneous samples. In that particular case, outliers are mostly the subject of the analysis, for instance the detection of gold in ores [135].

### 3.4. Regression

A problem of modelling a continuous variable is called a regression analysis. The extensive theory and examples of univariate calibration may be found in books dedicated to LIBS [13]. The Principal Component Regression (PCR) is an adaptation of the PCA for the purposes of a multivariate regression [30].

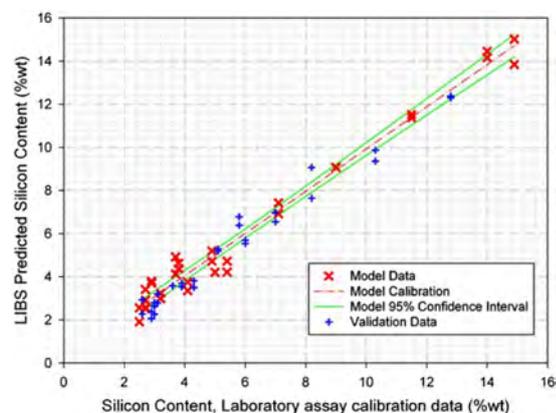
In regression analysis, or calibration, two variables are correlated via an equation describing this relationship. The most basic concept deals with the relation of the elemental line's intensity and the content of the analyte. Having an option to incorporate the information from more variables at once opens new possibilities in the regression analysis, *i.e.* the multivariate regression. This suggests that more than one variable contribute to the response related to the content of the analyte. In the regression, the data matrix  $X$  is supplied with a vector  $c$  describing the known concentration of an analyte in each object. Note that this vector may be related to any parameter characterizing each sample and in turn may affect the quality and structure of LIP and related emission spectra (such as sample hardness, roughness, *etc.*).

Bousquet et al. [44] concluded that the univariate calibration of Cr in soil samples does not provide good results due to a vast variation of matrices. Instead of a single univariate calibration, several calibration curves have to be constructed. For the classification, they suggest to use PCA directly and to separate samples according to their location in the PC space. Choi et al. [136] analyzed 21 samples with distinctly different matrix and used PCA for visualization prior to the univariate calibration. Of course, a significant impact of the matrix effect was evident in the PC space and also in the structure of univariate calibration that was split according to each individual sample matrix. Thus, for the regression analysis, the samples should be of the same matrix in order to provide the best possible results. Pořizka et al. [77] used PCA to filter the outliers and then to classify samples according to their matrix. The univariate calibration of Cu in soil samples was then assessed giving lower bias. In fact, the matrix effect was not mitigated but completely avoided using this kind of data processing. They also compared PCR and PLSR to the univariate calibration after classification proving that a multivariate data analysis was not a necessary step in data analysis. Regardless of this finding, it is generally accepted and proved in the LIBS literature that multivariate algorithms may compensate the matrix effect to certain extent and may influence the LIBS analysis itself [43,137].

Wisbrun et al. [138] brought a pioneering study on calibration of heavy metals (Cd, Cr, Cu, Ni, Pb, and Zn) in soils, sand, and sewage sludge using PCR. They selected and weighed 50 spectral lines of

analytes prior to a PCR computation; only two PCs were used. They also estimated limits of detection of tens of ppm. Death et al. [45] performed a quantitative analysis of six elements (Al, Fe, K, Mn, P, and Si) in iron ores by applying PCR based on a maximum number of ten PCs forming a model of short band spectral region (centered at 250, 400 or 750 nm). This analysis showed results in good agreement with referenced values. An example of PCR used for the modelling and prediction in the quantification process of Si is shown in Fig. 6. In their further work, Death et al. [46] used PCR for quantification of elements in three sets of samples from Australia and West Africa showing good calibration results. Tripathi et al. [139] used PCR to predict the content of plutonium oxide surrogates ( $CeO_2$ ,  $Fe_2O_3$ ,  $Cr_2O_3$ ,  $MoO_3$ , NiO).

Most often, the results of PCR are compared with those of PLSR, the latter is generally considered to have better performance. Both algorithms are collated across various applications of LIBS. Fink et al. [124] analyzed a set of polymers. They also showed the improvement of one order of magnitude in the detection limits when comparing the performance of multivariate and univariate algorithms. Doucet et al. [140] compared PCR and PLSR to find out which one performs better in a quantitative analysis of elements (Cu, Fe, Mg, Mn, Si, and Ti) in 260 aluminum alloys. Data were obtained using a LIBS system with Paschen-Runge spectrometer focused on 11 elemental lines (no matrix line



**Fig. 6.** This figure shows a PCR calibration model constructed from 16 samples and validated by 15 samples unknown to the model. The sample set was collected from iron ores and the quantification study was in this case focused on the Si content. This figure also demonstrates a good performance of PCR when the sample set is well chosen; the coefficient of determination was estimated to 0.99 and average relative errors were estimated to 2.5% for model and 5.6% for prediction. Obtained from [45] with permission provided by Elsevier and Copyright Clearance Center, license number: 4347691418259. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of aluminum). They stressed that it is necessary to linearize the data prior the implementation of MVDA algorithms. This data pretreatment resulted in a dramatic improvement of figures of merit. In their following study [49] they compared PCR and PLSR used for a quantification of pharmaceutical tablets and their composition through the emission of molecular bands (CN, CH, C<sub>2</sub>). Yaroshchuk et al. [132] provided a calibration and quantification of Fe in ore pellets using PCR, PLSR and variations of PLSR (serial PLSR and multi-block PLSR). The data were obtained using a LIBS system with two spectrometers covering two separate spectral ranges in UV and VIS, resulting in two different datasets. In the case of PCR and PLSR, the considered data set was either from UV range, VIS range, or merged UV and VIS ranges. S-PLSR and MB-PLSR were able to work with both ranges (UV and VIS) at once without merging. Thus, S-PLSR and MB-PLSR present another way to process data from different spectrometers or even techniques, considering hyphenated systems (such as LIBS and Raman). Devandag et al. [141] compared the performance of univariate (6 various spectral lines) and multivariate (PCR and PLSR) calibration for the purpose of a quantitative analysis of Mn in glass matrix. One order of magnitude improvement in detection limits was obtained with the utilization of MVDA.

#### 3.4.1. PCR counterparts

PCR is not so massively utilized due to its relatively low performance compared to other algorithms. Despite that we deliver articles where PCA was used as the first step of an analysis for visualization prior to any MVDA regression. Sirven et al. [43] investigated the influence of matrix effect on the performance in a quantitative analysis of Cr by PLSR and ANN. In another study [111] they used PCA prior to ANN for a quantification of Cr in soil. El Haddad et al. [137] visually inspected the ICP-OES data through PCA in order to check their clustering. As a result, no clusters in PC space were observed which directly implies that the chemical composition of samples obtained through the ICP-OES analysis does not significantly differ. It might be stated that on a certain level the samples collected from various places of the mine could be considered similar. Martin et al. [142] used PCA for discrimination of four wood species. PLSR was then used for a quantitative analysis of Cu and Zn. Martin et al. used PCA (scores and loadings) prior to a quantitative study using PLSR of the forensic samples [143] and nuclear materials (Ce, Cs, and Sr) [144]. Anderson et al. [145] used PCA for a visualization and loadings inspection prior the comparison of PLSR and ANN performance in quantitative analysis of soils for the ChemCam instrument. Bacterial contamination of food was assessed by a combination of PCA and PLSR [146]. Bilge et al. used PCA for a visualization of LIBS data from meat [147] and milk [148] samples prior to a quantitative analysis by PLSR. Moncayo et al. [149] visualized milk samples prior to their quantitative analysis done by ANN.

It is worth mentioning that PLSR was used in other applications, such as a quantification of components in geo-samples under Martian conditions [150], of micronutrients in pellets of plant material [151], of loss on ignition in iron ore [152] and of Cl in concrete [153].

#### 3.5. Clustering and classification

The PCA model transforms the original data space and provides a low-dimensional visualization of the data. This projection provides only the unsupervised pattern recognition. Despite that, general relationships between data may be observed. The most basic kind of analysis is the clustering of data points according to their mutual similarities. The level of similarity is determined using a metric set by the Pearson correlation coefficient or by the distance (Euclidean, Manhattan or Mahalanobis) between data points in the  $k$ -dimensional space ( $k$  is the number of PCs,  $k < m$ ). Resulting values are often organized in a

dendrogram (a result of hierarchical clustering) and utilized for the discrimination of objects based on their characteristic spectra.

In contrast to the unsupervised pattern recognition (clustering), there is the supervised approach of classification. In the case of the latter, data matrix  $X$  is extended with a vector describing the class membership of individual objects. Afterward, a model is constructed using a MVDA algorithm and then unknown data are classified. For those purposes, the PCA-based algorithm may be used - Soft Independent Modelling of Class Analogies (SIMCA). The concept of *soft modelling* enables to classify one object to more groups at the same time or to leave the object unclassified. At this point we would like to emphasize that MVDA algorithms classify the obtained spectra, not the samples themselves. Only after a careful validation of obtained statistical results, a classification of the real samples might be considered. Therefore, it is important to include figures of merit describing the performance of the classification process.

This section overviews articles that deal with a classification of LIBS data through i) the implementation of classification algorithms after the dimensionality reduction by PCA or ii) the direct utilization of the SIMCA algorithm. In contrast to quantification, an application of LIBS technique benefits from the strong matrix-dependence when it comes to classification. Possession of an extensive data library, *i.e.* collection of unique spectra covering the diversity of materials, is absolutely crucial in order to provide a reliable classification. No data library is currently available commercially, and even if it was, the transfer between different LIBS systems is not possible yet. Therefore, each research group is building its own database. Bohling et al. [154] suggested an approach where it would be possible to further extend an already existing model with new measurements.

The most challenging-to-build database seems to be the collection of various minerals and ores. Though, the application of LIBS to geology is one of the most promising in terms of LIBS potential (*in-situ* analysis). For instance, it is expected that MVDA discrimination of samples can emulate the distribution of the igneous rock types in the so-called QAPF and TAS diagrams and therefore can provide a routine provenance study. The provenance study of minerals was claimed in the patent [155] utilizing the PCA and other MVDA algorithms. The feasibility of LIBS combined with MVDA for the classification of geological samples is also demonstrated by the Curiosity rover equipped with the ChemCam device [156].

PCA is usually utilized only for a brief inspection of the spectra and a visualization of hidden patterns within the data set. Therefore, the preliminary exploratory data analysis using PCA is followed by an application of linear and non-linear MVDA algorithms for classification and quantification purposes. Nevertheless, distinct clusters in the PCA scores plot suggest that it is possible to use simple algorithms and that the use of advanced non-linear methods is in certain cases (low number of samples) unnecessary. Thus, in several articles the classification of samples was demonstrated only by a visual inspection of the PCA scores plot [44,76,77,111]. Samuels et al. [116] presented a discrimination of biological samples, such as bacterial spores, molds, pollens and the protein ovalbumin. They proved that it is possible to separate bio-samples by applying unsupervised multivariate method (PCA) on their typical LIP spectra (full echelle range). PCA can also suggest the class of the new object (spectrum, set of spectra representing a sample) in a visual way by projecting it into a model given by PCs subspace [80]. Yet they concluded that in this approach there were insufficiencies, such as lacking any supervised process and related figures of merit.

A rather pioneering work by Fink et al. [124] claimed a clustering of spectra from plastic samples in the PC space (3 PCs covering 96% of variance) and suggested their potential classification. Hybl et al. [157] presented a discrimination of biological aerosols from natural background by PCA. 12 samples in four classes were analyzed by a broad-

band LIBS system and 30 most prominent lines were selected. Data points in the PC space indicated possible inter-class but no intra-class discrimination, *i.e.* the system showed a low specificity. However, as the authors stated, this work did not push the limits of chemometrics due to the simplicity of the presented data, *i.e.* the number of sample classes was quite low. Baudelet et al. [158] demonstrated the possibility to discriminate between individual bacteria using their Ca, Na and K signals obtained from a fs ablation. De Lucia et al. [159] showed a classification between explosive and non-explosive organic materials in the PCA space. Only first PC (70%) was used and a threshold was applied in order to discriminate between samples. Xia and Bakker [160] compared the performance of PCA-Adaboost (a hybrid PCA designed for classification) and PLS-DA in a real-life application of LIBS providing a sorting of materials. Considering the significantly different matrices that are subject to classification, the estimated optimal number of 100 PCs for PCA-Adaboost and 80 latent variables for PLS-DA is rather disturbing - especially when a good separation of clusters was shown in the space of 3 PCs.

### 3.5.1. Classification after dimensionality reduction by PCA

PCA is first used to reduce the number of dimensions, which can get over 30,000 of variables (*i.e.* wavelengths in the detected spectrum) in the case of LIBS data. Then, only the most descriptive PCs are kept in the model. This step preserves only the most valuable information and mitigates the contribution of fluctuations and noise within the data. Thus, this preprocessing step increases the prediction power of the model on which other algorithms are applied to yield a sample classification. This approach was successfully used also in other techniques, such as Raman spectroscopy [161], Fourier Transform Infrared [162], and X-ray fluorescence [163]. Yet, it has to be stressed that this approach is unsupervised and the classification results may be biased.

Samek et al. [164] applied the Mahalanobis distance on the reduced PC space as a metric to classify teeth caries. In another work [165] they used the same approach to classify solid standards. Yueh et al. [166] used PCA (11 PCs) followed by HCA to classify murine soft tissue samples (brain, kidney, liver, lungs, muscle, and spleen). PCA was used to reduce the original multivariate data to a lower number of variables while preserving the latent variance at the same time. Obtained results were compared with the performance of the PLS-DA and ANN applied on the original dataset. Moderate results were found regardless of the utilized algorithm. A higher number of measurements encompassing the system fluctuation and improved outliers filtering was suggested.

Gottfried et al. [167] used PCA for a visualization of several hazardous materials (explosives and bacteria). Three PCs constituted 99.39% of total variance in the model based on 6 ratios of the most significant elemental lines (C, O, H and N). Then simple classification was done directly in the PC space by drawing ellipses and counting positive and false hits.

Unnikrishnan et al. [168] used PC scores for a clustering based on Mahalanobis distance and also used spectral residuals for the classification of plastics where the diagnostic threshold of classification was estimated using Youden's index plot and Receiver Operating Characteristics (ROC). Spectra were masked except for three 15 nm spectral regions including lines from C, H, and CN. Poor SNR and SBR which were due to the composition of the samples (C, H, N, and O) spoiled the classification accuracy. Pokrajac et al. [169] introduced PCA to LIBS spectra of four proteins. After the PCA dimensionality reduction they implemented various classifying algorithms (kNN, LDA, SVM, and ANN). They found the best classification performance for a relatively high number of PCs (from 12 to 41). That is rather extreme taking into consideration the total number of samples, spectra per sample, and spectral lines per spectrum. This fact might be caused by a low number

of informative spectral lines with poor SBR and SNR despite the broad wavelength range (200–950 nm) of echelle spectrometer. On the contrary to this finding, they claimed that the optimal number was 6 PCs, as they estimated from the Cattell's scree plot.

Porizka et al. [77] used PCA for the dimensionality reduction and following classification in the space of first three PCs (covering up to 97.5% of total variance) applying the Gaussian clustering algorithm. The Gaussian distribution of points in each cluster was assumed but not statistically tested. After the clustering, they did a univariate quantitative analysis. Vitkova et al. [170] presented a classification of 29 brick samples using LDA after PCA reduction. They used the table-top and stand-off LIBS systems. In PCA scores plots, the firing temperature of bricks was evident. This leads to the matrix effect influencing the quality of LIP spectra. Merk et al. [171] tested the PCA and PLS-DA capability to provide a fast turnaround time in metal scrap sorting, and they reached a repetition rate 25 samples per second.

Amador-Hernandez et al. [104,105] used PCA prior to unsupervised k-means clustering for class assignment and consecutive multivariate mapping of printed electrodes. In other articles, PCA was used for a dimensionality reduction prior to hazardous material analysis through neural networks [154], classification of proteins by SVM [172], classification of geological samples by k-means clustering [55], classification of wood samples [173], classification of bricks utilizing stand-off system and LDA [88], classification of soils by PLS-DA [174], salts by PLS-DA [175], and classification of explosives by kNN [112]. Pontes et al. [176] used the successive projection algorithm (SPA), genetic algorithm (GA), and stepwise formulation (SF) for a dimensionality reduction prior to LDA classification. The findings were then compared to SIMCA.

### 3.5.2. Supervised classification using SIMCA

SIMCA is one of the simplest classification algorithms and its concept was developed in 1970s by Wold and Sjöström [177]. In the modelling step, each class of objects is individually modeled using PCA; each PCA model can be described by a different number of PCs. Then, unknown or test data (for cross-validation) are projected in the multi-dimensional spaces represented by individual PCA models and their distances to the center of gravity of individual models are calculated. The geometric distance can be also converted to statistical probabilities. Finally, membership of the unknown or test data is assigned to a class which has the smallest (or under threshold) distance to the model.

Sirven et al. [42] presented a classification and a provenance study of uranium ores using SIMCA. They did such a supervised classification twice, first time, they used 41 lines and second time only 10 lines (one spectral line per element). In both cases, good figures of merit were obtained. They also suggested that supervised classification algorithms could provide more accurate results than the unsupervised ones, such as hierarchical clustering, k-means clustering, *etc.* Clegg et al. [178] concluded that both PCA and SIMCA algorithms will be beneficial for the Mars rover and related exploration. SIMCA was employed with a satisfactory accuracy of 88.1%. Colao et al. [41] used SIMCA to classify 35 unknown samples based on 10 reference samples. In other works, they used SIMCA to classify citrus leaves [118] and pharmaceutical tablets [131].

SIMCA is popular for its simplicity but it is generally accepted that it has lower performance than other methods. Thus it is usual to accompany SIMCA with other methods and compare their performance. Sirven et al. [80] showed that PLS-DA (with a classification rate of 85.9%) was more sensitive than SIMCA (with a classification rate of 77.5%) in the classification of rocks. However, SIMCA proved to be more robust and effective in the classification of a few similar samples. Therefore, the authors suggested using SIMCA at the beginning of the

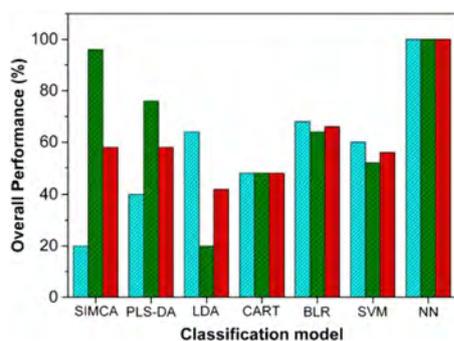


Fig. 7. This figure shows the overall performance of selected MVDA algorithms in classification of 25 bone samples; selected figures of merit are success rate (cyan), robustness (green) and accuracy (red). It is obvious that simple SIMCA is outperformed by other, sophisticated (even non-linear) algorithms. Abbreviations: PLS-DA – Partial Least Squares Discriminant Analysis, LDA – Linear Discriminant Analysis, CART – Classification and Regression Tree, BLR – Binary Logistic Regression, SVM – Support Vector Machines, NN – Neural Networks. Obtained from [183] with permission provided by Elsevier and Copyright Clearance Center, license number: 4347691418259. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ChemCam operation and the utilization of PLS-DA was planned for later stages when the library has grown. Moreover, they improved the classification rate to 100% when the SIMCA and PLS-DA algorithms were combined. Duchene et al. [179] showed similar results in the classification of archaeological material. Gottfried et al. [180] used PCA (99.27% of joint variance in 3 PCs) to visualize spectra of explosives based on 6 elemental line ratios. Then PCA of 29 variables (9 intensities of spectral lines and molecular bands and their ratios) was done. Classification of four samples was provided via SIMCA models using 5 to 8 PCs and via PLS-DA model using 17 latent variables. However, no figure of merits were provided which made the comparison of both MVDA algorithms difficult. Anderson et al. [55] compared SIMCA with other techniques in the classification of igneous rocks and showed SIMCA's lower performance.

### 3.5.3. SIMCA counterparts

Some more sophisticated and even non-linear algorithms were used.

Table 1

This table summarizes overview of literature related to implementation of PCA in LIBS.

The way of PCA implementation	
Visualization	[4,43,44,78,79,83–88,90–96,99–101,110,111,116–118,131,136,137,141–43,145,147–149,157,168,170,173,183]
Classification	[36,39,41,42,55,76–78,80,89,98,112,118,119,124,125,131,136,154,160,164–166,171–174,176,178,179,180–183]
Regression (PCR)	[45,46,49,77,124,132,138,139–141]
Variable down-selection	[42,111,166]
Dimensionality reduction	[55,77,88,95,112,114,120,154,164–166,169,170,174,175,183]
Outliers filtering	[42,56,77,80]
Multivariate mapping	[81,104,105,107]
Application of LIBS	
Archaeology and forensics	[41,86,88,125,143,170,179]
Miscellaneous applications	[98,104,105,114,115,141,160,175]
Biology	[4,76,84,94,95,99,100,116,118,119,142,146,147,148,149,157,164,166,169,172,173]
Geology	[39,42–45,46,55,76–78,80,81,83,89,90–92,101,107,110,111,117,132,137–139,143,145,174,176,178]
Explosives and chemicals, nuclear materials	[79,93,96,112,119,144,154,159,167,180,186]
Pharmaceutics	[36,49,131,182]
Polymers	[87,124,168,181]
metals	[39,56,85,113,120,140,171]

Godoi et al. [181] compared SIMCA, PLS-DA and KNN for a classification of toys, however their goal was to estimate the toxicity, *i.e.* the amount of Cd and Pb. Dingari et al. [182] compared the performance of SIMCA, PLS-DA and SVM, showing that a non-linear SVM algorithm provides higher classification performance than its linear counterparts. Myakalwar et al. [36] used SIMCA-derived models to provide an average correct classification rate of 88% of non-gated pharmaceuticals spectra, whereas PLS-DA and ANN reached 100%. Moncayo et al. [183] compared 7 MVDA algorithms (including SIMCA and LDA done after PCA dimensionality reduction) for a classification of 55 bone samples of 5 individuals, see Fig. 7. They proved that ANN, due to its non-linear nature, provides the most sensitive and robust results.

PLS-DA is a very frequently used alternative to SIMCA for the classification of LIBS data [83,156]. Remus et al. [184] recommended using PLS-DA rather than PCA related algorithms because the latter were not capable to fully discriminate samples coming from different locations, *i.e.* to provide a provenance study. This obvious superiority of PLS-DA over PCA was summarized by Barker and Rayens [185].

Classification and feasibility of the provenance of conflict minerals were proved based on the conflict minerals' characteristic spectral fingerprints using PLS-DA [102]. Explosive residues were analyzed using LIBS as a thin films spread on some organic and inorganic substrates, consequently the collected spectra were classified by PLS-DA [186]. It is worth mentioning that other MVDA algorithms were also used for classification, such as linear and rank correlation [187], Linear Discriminant Analysis (LDA) [88], ICA [68], ANN [188–190], DFA [67,191,192], Graph Theory Method [193].

An interesting application of SIMCA arises due to its rather low performance (and in turn quite a high sensitivity to changes in the topology and structure of the data). Thus, SIMCA tends to be more biased than other MVDA algorithms (PLS-DA, SVM, SOM, ANN, *etc.*). Because of that, SIMCA can be utilized to study the effects of preprocessing (normalization, outliers filtering, *etc.*). It compares the changes in the data structure by resulting classification figures of merit [39,56].

For the sake of completeness, the combination of LIBS and Raman data aids to get a correct classification. Three data sets were created: i) LIBS, ii) Raman, iii) LIBS and Raman. Hoehse et al. [98] compared performance of SIMCA, PLS-DA and SVM in the classification inks and pigments. Prochazka et al. [99] used SOM to classify 6 bacterial strains.

## 4. Summary of publications (Table 1)

## 5. Conclusion and future prospects

Based on the literature survey, LIBS combined with MVDA algorithms proved the capability to classify unknown samples and quantify analytes in many applications. However, the majority of reviewed articles represented only feasibility and preliminary studies. The impact of presented alterations in data pre-processing and MVDA algorithms on the resulting figures of merit was demonstrated on a limited number of samples, with a low number of spectra per sample, etc.

Generally, LIBS is on its rise and the hand-held as well as the table-top commercial systems attract more and more attention. The hand-held LIBS systems are used namely for the classification of steels and alloys, and for quantitative analyses. MVDA algorithms also penetrate this sector and will be decisive in terms of the performance and the resulting figures of merit.

Processing of LIBS data has undergone several dynamic developments in recent years. An increasing number of samples and detected representative spectra demands the implementation of MVDA algorithms in a more sophisticated way. The goal is to provide fast and robust data processing with desired selectivity and specificity, significantly reducing the dimensionality and information redundancy, yet retaining the most valuable information within the data matrix. Considering the trends and historical evolution, it seems that a progress in the data analysis will be continually made in the future. The implementation of MVDA algorithms will become a vital part of a routine LIBS analysis.

Historically, PCA and its variations were abundantly applied for multivariate analysis of LIBS data, such as visualization, clustering, classification, regression and recently also for multivariate surface mapping. It was also shown that both basic and advanced analyses of LIBS data can be provided using only PCA and its variations (SIMCA, PCR). Thus, when judicious experimental design and data preprocessing is provided, there is no need for switching to a more sophisticated MVDA algorithm (such as PLSR, PLS-DA, SVM, SOM, ANN, etc.). Therefore, it may be stated that PCA is essentially rooted in the LIBS community. Articles summarized in this review reflect the crucial issues that have to be considered when MVDA is of interest. Even though this review article was primarily aimed at PCA algorithm and its implementation for processing of LIBS data, the majority of presented suggestions, recommendations and conclusions can be considered in implementation of any other MVDA algorithm.

The reviewed articles represent mostly feasibility studies and preliminary results proving some capacities of LIBS technique. This was caused by the fact that there are no available data libraries, researches therefore established *ad hoc* libraries which were based most often only on limited number of collected samples. The analytical approach is different in every research group due to different applications, data structure, and experience of each group in relation to MVDA and the associated algorithm implementation. There are many contradictory suggestions and recommendations in the literature, which implies that there is no established receipt for a correct data analysis and it may not exist. Despite that, the data preprocessing and application of MVDA algorithms should follow several guidelines:

- **Theoretical background** - firm knowledge of the basic processes occurring in LIP and affecting the characteristic signal is the key. The MVDA algorithm should not be used as a “black box”.
- **Experimental design** - an optimized LIBS system and a coherent experimental design should deliver high-quality data, both repeatable and reproducible.
- **Robust dataset** - a statistically significant number of objects (*i.e.* spectra) describing the selected number of samples is important to cover a possible variation in the studied sample matrix and the fluctuation in the analysis itself.
- **Preprocessing** - well balanced data preprocessing should improve the data. At the same time, retaining the most of the variability

within the dataset is desirable.

- **MVDA** – an optimization of the selected algorithm (avoiding overfitting) and an estimation of figures of merit should be done by using a model, validation and test subsets.
- **Detailed report** - the whole process (starting with samples, going through LIBS analysis and data collection, preprocessing and finally MVDA) should be described in detail. Figures of merit have to be estimated; this enables the means for a potential comparison with results of other research groups.

Looking at the issue from a future perspective, it will be necessary to deliver an optimized data processing algorithm. In the ideal case, this algorithm should enable to reduce the number of objects, which would decrease response time of the modelling and prediction step. All the same, it would be sensitive to the data and would conserve the information from matrix, minor, and trace elements. To summarize, the algorithms should be sensitive to outliers and should provide a high degree of specificity.

## Acknowledgement

Authors affiliated with CEITEC would like to acknowledge the financial support obtained from the National Sustainability program - CEITEC NPU II (LQ1061) and supported also from the ERDFund-Project CEITEC Nano+ (CZ.02.1.01/0.0/0.0/16\_013/0001728). PP is grateful to the Fulbright commission for supporting his research at the University of Florida (E0583833).

## References

- [1] J. Kaiser, K. Novotný, M.Z. Martin, A. Hrdlička, R. Malina, M. Hartl, R. Kizek, Trace elemental analysis by laser-induced breakdown spectroscopy—biological applications, *Surf. Sci. Rep.* 67 (11–12) (2012) 233–243.
- [2] S.J. Rehse, H. Salimnia, A.W. Miziolek, Laser-induced breakdown spectroscopy (LIBS): an overview of recent progress and future potential for biomedical applications, *J. Med. Eng. Technol.* 36 (2) (2012) 77–89.
- [3] D.J. Santos, L.C. Nunes, G.G.A. de Carvalho, M.d.S. Gomes, P.F. de Souza, F.d.O. Leme, F.J. Krug, Laser-induced breakdown spectroscopy for analysis of plant materials: a review, *Spectrochim. Acta B At. Spectrosc.* 71–72 (2012) 3–13.
- [4] P. Pořízka, P. Prochazková, D. Prochazka, L. Sládková, J. Novotný, M. Petrilak, J. Kaiser, Algal biomass analysis by laser-based analytical techniques—a review, *Sensors* 14 (12) (2014) 17725–17752.
- [5] R.S. Harmon, R.E. Russo, R.R. Hark, Applications of laser-induced breakdown spectroscopy for geochemical and environmental analysis: a comprehensive review, *Spectrochim. Acta B At. Spectrosc.* 87 (2013) 11–26.
- [6] R. Noll, *Laser-Induced Breakdown Spectroscopy Fundamentals and Applications*, Springer-Verlag Berlin Heidelberg, Heidelberg New York, 2012.
- [7] J.O. Cáceres, F. Pelascini, V. Motto-Ros, S. Moncayo, F. Trichard, G. Panczer, J. Martin-Chivelet, Megapixel multi-elemental imaging by laser-induced breakdown spectroscopy, a technology with considerable potential for paleoclimate studies, *Sci. Rep.* 7 (1) (2017).
- [8] F.J. Fortes, J.J. Laserna, The development of fieldable laser-induced breakdown spectrometer: no limits on the horizon, *Spectrochim. Acta B At. Spectrosc.* 65 (12) (2010) 975–990.
- [9] J. Rakovský, P. Čermák, O. Musset, P. Veis, A review of the development of portable laser induced breakdown spectroscopy and its applications, *Spectrochim. Acta B At. Spectrosc.* 101 (2014) 269–287.
- [10] S. Maurice, S.M. Clegg, R.C. Wiens, O. Gasnault, W. Rapin, O. Forni, A.R. Vasavada, ChemCam activities and discoveries during the nominal mission of the Mars Science Laboratory in Gale crater, Mars, *J. Anal. At. Spectrom.* 31 (4) (2016) 863–889.
- [11] CHEMCAM Team, “ChemCam on Mars,” 2 February 2018. [Online], Available: <http://www.msl-chemcam.com/index.php#.WnRdEahKswE>.
- [12] D.A. Cremers, L.J. Radziemski, *Handbook of Laser-Induced Breakdown Spectroscopy*, 2nd Edition, John Wiley & Sons, Ltd, Chichester, West Sussex, 2013.
- [13] A. Miziolek, V. Palleschi, I. Schechter, *Laser Induced Breakdown Spectroscopy*, Cambridge University Press, Cambridge, 2006.
- [14] E. Tognoni, V. Palleschi, M. Corsi, G. Cristoforetti, Quantitative micro-analysis by laser-induced breakdown spectroscopy: a review of the experimental approaches, *Spectrochim. Acta B At. Spectrosc.* 57 (7) (2002) 1115–1130.
- [15] C. Aragón, J.A. Aguilera, Characterization of laser induced plasmas by optical emission spectroscopy: a review of experiments and methods, *Spectrochim. Acta B At. Spectrosc.* 63 (9) (2008) 893–916.
- [16] R.E. Russo, X. Mao, J.J. Gonzalez, V. Zorba, J. Yoo, Laser ablation in analytical chemistry, *Anal. Chem.* 85 (13) (2013) 6162–6177.
- [17] 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 (4) (2012) 347–419.
- [18] R. Gaudiuso, M. Dell'Aglio, O.D. Pascale, G.S. Senesi, A.D. Giacomo, Laser induced breakdown spectroscopy for elemental analysis in environmental, cultural heritage and space applications: a review of methods and results, *Sensors* 10 (12) (2010) 7434–7468.
- [19] 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 (12) (2010) 335A–336A.
- [20] A.P.M. Michel, Review: applications of single-shot laser-induced breakdown spectroscopy, *Spectrochim. Acta B At. Spectrosc.* 65 (3) (2010) 185–191.
- [21] T. Kim, Z.G. Specht, P.S. Vary, C.T. Lin, Spectral fingerprints of bacterial strains by laser-induced breakdown spectroscopy, *J. Phys. Chem. B* 108 (17) (2004) 5477–5482.
- [22] R.S. Harmon, J. Remus, N.J. McMillan, C. McManus, L. Collins, J.L.J. Gottfried, A.W. Miziolek, LIBS analysis of geomaterials: geochemical fingerprinting for the rapid analysis and discrimination of minerals, *Appl. Geochem.* 24 (6) (2009) 1125–1141.
- [23] P. Geladi, Chemometrics in spectroscopy. Part 1. Classical chemometrics, *Spectrochim. Acta B At. Spectrosc.* 58 (5) (2003) 767–782.
- [24] P. Geladi, B. Sethson, J. Nyström, T. Lilhonga, T. Lestander, J. Burger, Chemometrics in spectroscopy, *Spectrochim. Acta B At. Spectrosc.* 59 (9) (2004) 1347–1357.
- [25] J.M. Andrade, M.J. Cal-Prieto, M.P. Gómez-Carracedo, A. Carlosena, D. Prada, A tutorial on multivariate calibration in atomic spectrometry techniques, *J. Anal. At. Spectrom.* 23 (1) (2008) 15–28.
- [26] N. Kumar, A. Bansal, G.S. Sarma, R.K. Rawal, Chemometrics tools used in analytical chemistry: an overview, *Talanta* 123 (2014) 186–199.
- [27] F.C.J. De Lucia, J.L. Gottfried, Rapid analysis of energetic and geo-materials using LIBS, *Mater. Today* 14 (6) (2011) 274–281.
- [28] J. El Haddad, L. Canioni, B. Bousquet, Good practices in LIBS analysis: review and advices, *Spectrochim. Acta B At. Spectrosc.* 101 (2014) 171–182.
- [29] R.G. Brereton, G.R. Lloyd, Partial least squares discriminant analysis: taking the magic away, *J. Chemom.* 28 (4) (2014) 213–225.
- [30] R.G. Brereton, Chemometrics: Data Analysis for the Laboratory and Chemical Plant, Antony Rowe Ltd, Chippingham, Wiltshire, 2003.
- [31] R. Cangelosi, A. Goriely, Component retention in principal component analysis with application to cDNA microarray data, *Biol. Direct* 2 (1) (2007) 2.
- [32] P. de Boves Harrington, Statistical validation of classification and calibration models using bootstrapped Latin partitions, *Trends Anal. Chem.* 25 (11) (2006) 1112–1124.
- [33] E.H. van Veen, M.T.C. de Loos-Vollebregt, Application of mathematical procedures to background correction and multivariate analysis in inductively coupled plasma-optical emission spectrometry, *Spectrochim. Acta B At. Spectrosc.* 53 (5) (1998) 639–669.
- [34] A.A. Urbas, S.J. Choquette, Automated spectral smoothing with spatially adaptive penalized least squares, *Appl. Spectrosc.* 65 (6) (2011) 665–677.
- [35] P. Yaroshchik, J.E. Eberhardt, Automatic correction of continuum background in laser-induced breakdown spectroscopy using a model-free algorithm, *Spectrochim. Acta B* 99 (2014) 138–149.
- [36] A.K. Myakalwar, N.C. Dingari, R.R. Dasari, I. Barman, M.K. Gundawar, Non-gated laser induced breakdown spectroscopy provides a powerful segmentation tool on concomitant treatment of characteristic and continuum emission, *PLoS One* 9 (8) (2014) e103546.
- [37] 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 B At. Spectrosc.* 65 (8) (2010) 642–657.
- [38] J.P. Castro, E.R. 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 (10) (2016) 2005–2014.
- [39] P. Pořízka, J. Klus, A. Hrdlička, J. Vráběl, P. Škarková, D. Procházka, J. Kaiser, Impact of laser-induced breakdown spectroscopy data normalization on multivariate classification accuracy, *J. Anal. At. Spectrom.* 32 (2) (2017) 277–288.
- [40] T. Takahashi, B. Thornton, Quantitative methods for compensation of matrix effects and self-absorption in laser induced breakdown spectroscopy signals of solids, *Spectrochim. Acta B* 138 (2017) 31–42.
- [41] 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 B At. Spectrosc.* 65 (8) (2010) 688–694.
- [42] J.-B. Sirven, A. Pailloux, Y. M'Baye, N. Coulon, T. Alpettaz, S. Gossé, Towards the determination of the geographical origin of yellow cake samples by laser-induced breakdown spectroscopy and chemometrics, *J. Anal. At. Spectrom.* 24 (4) (2009) 451.
- [43] J.-B. Sirven, B. Bousquet, L. Canioni, L. Sarger, Laser-induced breakdown spectroscopy of composite samples: comparison of advanced Chemometrics methods, *Anal. Chem.* 78 (5) (2006) 1462–1469.
- [44] B. Bousquet, J.-B. Sirven, L. Canioni, Towards quantitative laser-induced breakdown spectroscopy analysis of soil samples, *Spectrochim. Acta B At. Spectrosc.* 62 (12) (2007) 1582–1589.
- [45] D.L. Death, A.P. Cunningham, L.J. Pollard, Multi-element analysis of iron ore pellets by laser-induced breakdown spectroscopy and principal components regression, *Spectrochim. Acta B At. Spectrosc.* 63 (7) (2008) 763–769.
- [46] D.L. Death, A.P. Cunningham, L.J. Pollard, Multi-element and mineralogical analysis of mineral ores using laser induced breakdown spectroscopy and chemometric analysis, *Spectrochim. Acta B At. Spectrosc.* 64 (10) (2009) 1048–1058.
- [47] A.P.M. Michel, A.D. Chave, Analysis of laser-induced breakdown spectroscopy spectra: the case for extreme value statistics, *Spectrochim. Acta B At. Spectrosc.* 62 (12) (2007) 1370–1378.
- [48] J. Klus, P. Pořízka, D. Procházka, J. Novotný, K. Novotný, J. Kaiser, Effect of experimental parameters and resulting analytical signal statistics in laser-induced breakdown spectroscopy, *Spectrochim. Acta B At. Spectrosc.* 126 (2016) 6–10.
- [49] F.R. Doucet, P.J. Faustino, M. Sabsabi, R.C. Lyon, Quantitative molecular analysis with molecular bands emission using laser-induced breakdown spectroscopy and chemometrics, *J. Anal. At. Spectrom.* 23 (5) (2008) 694.
- [50] National Institute of Standards and Technology, "National Institute of Standards and Technology," 2 February 2018. [Online], Available: <https://www.nist.gov/>.
- [51] C. Heise, "Atomic spectral line database," 2 February 2018. [Online], Available: <https://www.cta.harvard.edu/amp/ampdata/kurucz23/sekur.html>.
- [52] Applied Photonics, "Applied Photonics," 2 February 2018. [Online], Available: <http://www.appliedphotonics.co.uk/lib/capabilities/lib.htm>.
- [53] AtomTrace\*, "AtomTrace," 2 February 2018. [Online], Available: <https://www.atomtrace.com/>.
- [54] G. Amato, G. Cristoforetti, S. Legnaioli, G. Lorenzetti, V. Palleschi, F. Sorrentino, E. Tognoni, Progress towards an unassisted element identification from laser induced breakdown spectra with automatic ranking techniques inspired by text retrieval, *Spectrochim. Acta B At. Spectrosc.* 65 (8) (2010) 664–670.
- [55] R.B. Anderson, J.F. Bell, R.C. Wiens, R.V. Morris, S.M. Clegg, Clustering and training set selection methods for improving the accuracy of quantitative laser induced breakdown spectroscopy, *Spectrochim. Acta B At. Spectrosc.* 70 (2012) 24–32.
- [56] P. Pořízka, J. Klus, D. Procházka, E. Képeš, A. Hrdlička, J. Novotný, J. Kaiser, Laser-induced breakdown spectroscopy coupled with chemometrics for the analysis of steel: the issue of spectral outliers filtering, *Spectrochim. Acta B At. Spectrosc.* 123 (2016) 114–120.
- [57] J. Carlson, A. Wysoczanski, E. Voigtman, Limits of quantitation — yet another suggestion, *Spectrochim. Acta B At. Spectrosc.* 96 (2014) 69–73.
- [58] K. Pearson, On lines and planes of closest fit to systems of points in space, *Philos. Mag.* 2 (11) (1901) 559–572.
- [59] H. Hotelling, Analysis of a complex of statistical variables into principal components, *J. Educ. Psychol.* 24 (6) (1933) 417–441.
- [60] M. Kendall, *Multivariate Analysis*, MacMillan, New York, 1980.
- [61] L. Hatcher, N. Rourke, A Step-by-step Approach to Using SAS® for Factor Analysis and Structural Equation Modeling, SAS Institute, Cary, North Carolina, 2013.
- [62] W. Wu, D.L. Massart, S. de Jong, Kernel-PCA algorithms for wide data part II: fast cross-validation and application in classification of NIR data, *Chemom. Intell. Lab. Syst. Syst.* 37 (2) (1997) 271–280.
- [63] M. Hubert, P.J. Rousseeuw, S. Verboven, A fast method for robust principal components with applications to chemometrics, *Chemom. Intell. Lab. Syst. Syst.* 60 (1–2) (2002) 101–111.
- [64] M. Hubert, P.J. Rousseeuw, K. Vanden Branden, ROBPCA: a new approach to robust principal component analysis, *Technometrics* 47 (1) (2005) 64–79.
- [65] J. Diedrich, S.J. Rehse, S. Palchadhuri, Escherichia coli identification and strain discrimination using nanosecond laser-induced breakdown spectroscopy, *Appl. Phys. Lett.* 90 (16) (2007) 163901.
- [66] J. Diedrich, S.J. Rehse, S. Palchadhuri, Pathogenic Escherichia coli strain discrimination using laser-induced breakdown spectroscopy, *J. Appl. Phys.* 102 (1) (2007) 14702.
- [67] R.A. Putnam, Q.I. Mohaidat, A. Daabous, S.J. Rehse, A comparison of multivariate analysis techniques and variable selection strategies in a laser-induced breakdown spectroscopy bacterial classification, *Spectrochim. Acta B At. Spectrosc.* 87 (2013) 161–167.
- [68] O. Forni, S. Maurice, O. Gasnault, R.C. Wiens, A. Cousin, S.M. Clegg, J. Lasue, Independent component analysis classification of laser induced breakdown spectroscopy spectra, *Spectrochim. Acta B At. Spectrosc.* 86 (2013) 31–41.
- [69] T. Kohonen, Self-organized formation of topologically correct feature maps, *Biol. Cybern.* 43 (1) (1982) 59–69.
- [70] S. Kachigan, *Multivariate Statistical Analysis: A Conceptual Introduction*, New York: Radius Press, 1991.
- [71] A.B. Costello, J.W. Osborne, Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis, *Pract. Assess. Res. Eval.* 10 (2005) 173–178.
- [72] I. Jolliffe, *Principal Component Analysis*, Second Edition, Springer, New York, 2002.
- [73] R.B. Cattell, The scree test for the number of factors, *Multivar. Behav. Res.* 1 (2) (1966) 245–276.
- [74] H.F. Kaiser, The application of electronic computers to factor analysis, *Educ. Psychol. Meas.* 20 (1) (1960) 141–151.
- [75] M. Defernez, E.K. Kemsley, The use and misuse of chemometrics for treating classification problems, *Trends Anal. Chem.* 16 (4) (1997) 216–221.
- [76] B. Bousquet, G. Travaillé, A. Ismaël, L. Canioni, K. Michel-Le Pière, E. Brasseur, S. Belbèze, Development of a mobile system based on laser-induced breakdown spectroscopy and dedicated to in situ analysis of polluted soils, *Spectrochim. Acta B At. Spectrosc.* 63 (10) (2008) 1085–1090.
- [77] P. Pořízka, A. Demidov, J. Kaiser, J. Keivanian, I. Gornushkin, U. Panne, J. Riedel, Laser-induced breakdown spectroscopy for in situ qualitative and quantitative analysis of mineral ores, *Spectrochim. Acta B At. Spectrosc.* 101 (2014) 155–163.
- [78] P. Pořízka, J. Klus, J. Mašek, M. Rajnoha, D. Procházka, P. Modlitbová, J. Kaiser, Multivariate classification of echellograms: a new perspective in laser-induced

- breakdown spectroscopy analysis, *Sci. Rep.* 7 (1) (2017).
- [79] J.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 (31) (2008) G112.
- [80] J.-B. Sirven, B. Sallé, P. Mauchien, J.-L. Lacour, S. Maurice, G. Manhès, 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 (12) (2007) 1471.
- [81] J. Klus, P. Mikysek, D. Prochazka, P. Pořízka, P. Prochazková, J. Novotný, J. Kaiser, Multivariate approach to the chemical mapping of uranium in sandstone-hosted uranium ores analyzed using double pulse laser-induced breakdown spectroscopy, *Spectrochim. Acta B At. Spectrosc.* 123 (2016) 143–149.
- [82] V. Lazic, A. Palucci, S. Jovicic, C. Poggi, E. Buono, Analysis of explosive and other organic residues by laser induced breakdown spectroscopy, *Spectrochim. Acta B At. Spectrosc.* 64 (10) (2009) 1028–1039.
- [83] J.L. Gottfried, R.S. Harmon, F.C.J. De Lucia, A.W. Miziolek, Multivariate analysis of laser-induced breakdown spectroscopy chemical signatures for geochemical classification, *Spectrochim. Acta B At. Spectrosc.* 64 (10) (2009) 1009–1019.
- [84] D.E. Lewis, J. Martinez, C.A. Akpovo, L. Johnson, A. Chauhan, M.D. Edington, Discrimination of bacteria from Jamaican bauxite soils using laser-induced breakdown spectroscopy, *Anal. Bioanal. Chem.* 401 (7) (2011) 2225–2236.
- [85] S.R. Goode, S.L. Morgan, R. Hoskins, A. Oxsher, Identifying alloys by laser-induced breakdown spectroscopy with a time-resolved high resolution echelle spectrometer, *J. Anal. At. Spectrom.* 15 (9) (2000) 1133–1138.
- [86] A. Erdem, A. Çilingiroğlu, A. Giakoumaki, M. Castany, E. Kartsonaki, C. Fotakis, D. Anglos, Characterization of Iron age pottery from eastern Turkey by laser-induced breakdown spectroscopy (LIBS), *J. Archaeol. Sci.* 35 (9) (2008) 2486–2494.
- [87] S. Grégoire, M. Boudinet, F. Pelascini, F. Surma, V. Detalle, Y. Holl, Laser-induced breakdown spectroscopy for polymer identification, *Anal. Bioanal. Chem.* 400 (10) (2011) 3331–3340.
- [88] G. Vítková, K. Novotný, L. Prokeš, A. Hrdlička, J. Kaiser, J. Novotný, 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 B At. Spectrosc.* 73 (2012) 1–6.
- [89] P. Pořízka, J. Klus, D. Prochazka, G. Vítková, M. Brada, J. Novotný, J. Kaiser, Assessment of the most effective part of echelle laser-induced plasma spectra for further classification using Czerny-Turner spectrometer, *Spectrochim. Acta B At. Spectrosc.* 124 (2016) 116–123.
- [90] N.L. Lanza, R.C. Wiens, S.M. Clegg, A.M. Ollila, S.D. Humphries, H.E. Newsom, J.E. Barefield, Calibrating the ChemCam laser-induced breakdown spectroscopy instrument for carbonate minerals on Mars, *Appl. Opt.* 49 (13) (2010) C211.
- [91] C.E. McManus, N.J. McMillan, R.S. Harmon, R.C. Whitmore, J.F.C. De Lucia, A.W. Miziolek, Use of laser induced breakdown spectroscopy in the determination of gem provenance: beryls, *Appl. Opt.* 47 (31) (2008) G72.
- [92] D.C. Alvey, K. Morton, R.S. Harmon, J.L. Gottfried, J.J. Remus, L.M. Collins, M.A. Wise, Laser-induced breakdown spectroscopy-based geochemical fingerprinting for the rapid analysis and discrimination of minerals: the example of garnet, *Appl. Opt.* 49 (2010) C168–C180.
- [93] F.C.J. De Lucia, J.L. Gottfried, Characterization of a series of nitrogen-rich molecules using laser induced breakdown spectroscopy, *Propellants, Explos., Pyrotech.* 35 (3) (2010) 268–277.
- [94] N. Labbé, I.M. Swamidoss, N. André, M.Z. Martin, T.M. Young, T.G. Rials, Extraction of information from laser-induced breakdown spectroscopy spectral data by multivariate analysis, *Appl. Opt.* 47 (31) (2008) G158.
- [95] D. Merdes, J. Suhan, J.M. Keay, D.M. Hadka, W.R. Bradley, The investigation of laser-induced breakdown spectroscopy for detection of biological contaminants on surfaces, *Spectroscopy* 22 (4) (2007) 28–38.
- [96] L. Landström, A. Larsson, P.-Å. Gradmark, L. Örebrand, P.O. Andersson, P. Wåsterby, T. Tjärnhage, Detection and Monitoring of CWA and BWA Using LIBS, Chemical, Biological, Radiological, Nuclear, and Explosives (CBRNE) Sensing XV, Baltimore, Maryland, United States, (2014).
- [97] P. Pořízka, S. Kaski, A. Hrdlička, P. Modlitbová, L. Sládková, H. Häkkinen, J. Kaiser, Detection of fluorine using laser-induced breakdown spectroscopy and Raman spectroscopy, *J. Anal. At. Spectrom.* 32 (10) (2017) 1966–1974.
- [98] M. Hoehse, A. Paul, I. Gornushkin, U. Panne, Multivariate classification of pigments and inks using combined Raman spectroscopy and LIBS, *Anal. Bioanal. Chem.* 402 (4) (2012) 1443–1450.
- [99] D. Prochazka, M. Mazura, O. Samek, K. Rebrošová, P. Pořízka, J. Klus, J. Kaiser, Combination of laser-induced breakdown spectroscopy and Raman spectroscopy for multivariate classification of bacteria, *Spectrochim. Acta B At. Spectrosc.* 139 (2018) 6–12.
- [100] A. Larsson, H. Andersson, L. Landström, Impact of data reduction on multivariate classification models built on spectral data from bio-samples, *J. Anal. At. Spectrom.* 30 (5) (2015) 1117–1127.
- [101] J. Lasue, R.C. Wiens, T.F. Stepinski, O. Forni, S.M. Clegg, S. Maurice, Nonlinear mapping technique for data visualization and clustering assessment of LIBS data: application to ChemCam data, *Anal. Bioanal. Chem.* 400 (10) (2011) 3247–3260.
- [102] R.S. Harmon, K.M. Shughrue, J.J. Remus, M.A. Wise, L.J. East, R.R. Hark, Can the provenance of the conflict minerals columbite and tantalite be ascertained by laser-induced breakdown spectroscopy? *Anal. Bioanal. Chem.* 400 (10) (2011) 3377–3382.
- [103] J. Klus, P. Pořízka, D. Prochazka, P. Mikysek, J. Novotný, K. Novotný, J. Kaiser, Application of self-organizing maps to the study of U-Zr-Ti-Nb distribution in sandstone-hosted uranium ores, *Spectrochim. Acta B At. Spectrosc.* 131 (2017) 66–73.
- [104] J. Amador-Hernández, J.M. Fernández-Romero, M.D. Luque de Castro, In-depth characterization of screen-printed electrodes by laser-induced breakdown spectroscopy and pattern recognition, *Surf. Interface Anal.* 31 (4) (2001) 313–320.
- [105] J. Amador-Hernández, J. Fernández-Romero, M. Luque de Castro, Three-dimensional analysis of screen-printed electrodes by laser induced breakdown spectroscopy and pattern recognition, *Anal. Chim. Acta* 435 (2) (2001) 227–238.
- [106] R.C. Chinni, D.A. Cremers, L.J. Radziemski, M. Bostian, C. Navarro-Northrup, Detection of uranium using laser-induced breakdown spectroscopy, *Appl. Spectrosc.* 63 (11) (2009) 1238–1250.
- [107] S. Romppanen, H. Häkkinen, S. Kaski, Singular value decomposition approach to the yttrium occurrence in mineral maps of rare earth element ores using laser-induced breakdown spectroscopy, *Spectrochim. Acta B At. Spectrosc.* 134 (2017) 69–74.
- [108] S. Moncayo, L. Duponchel, N. Mousavipak, G. Panczer, F. Trichard, B. Bousquet, V. Motto-Ros, Exploration of megapixel hyperspectral LIBS images using principal component analysis, *J. Anal. At. Spectrom.* 33 (2018) 210–220.
- [109] S. Pagnotta, M. Lezzerini, L. Ripoll-Sequer, M. Hidalgo, E. Grifoni, S. Legnaioli, G. Lorenzetti, F. Poggialini, V. Palleschi, Micro-laser-induced breakdown spectroscopy (micro-LIBS) study on ancient roman mortars, *Appl. Spectrosc.* 71 (4) (2017) 721–727.
- [110] N.L. Lanza, S.M. Clegg, R.C. Wiens, R.E. McInroy, H.E. Newsom, M.D. Deans, Examining natural rock varnish and weathering rinds with laser-induced breakdown spectroscopy for application to ChemCam on Mars, *Appl. Opt.* 51 (7) (2012) B74.
- [111] J.-B. Sirven, B. Bousquet, L. Canioni, L. Sarger, S. Tellier, M. Potin-Gautier, I.L. Hecho, Qualitative and quantitative investigation of chromium-polluted soils by laser-induced breakdown spectroscopy combined with neural networks analysis, *Anal. Bioanal. Chem.* 385 (2) (2006) 256–262.
- [112] T.K. Sahoo, A. Negi, M.K. Gundawar, Study of Preprocessing Sensitivity on Laser Induced Breakdown Spectroscopy (LIBS) Spectral Classification, International Conference on Advances in Computing, Communications and Informatics (ICACCI), Kochi, India, (2015).
- [113] H.Y. Kong, L.X. Sun, J.T. Hu, Y. Xin, Z.B. Cong, A comparative study of two data reduction methods for steel classification based on LIBS, *Appl. Mech. Mater.* 644–650 (2014) 4722–4725.
- [114] Y. Lee, K.-S. Ham, S.-H. Han, J. Yoo, S. Jeong, Revealing discriminating power of the elements in edible sea salts: line-intensity correlation analysis from laser-induced plasma emission spectra, *Spectrochim. Acta B At. Spectrosc.* 101 (2014) 57–67.
- [115] Y. Lee, S.-H. Nam, K.-S. Ham, J. Gonzalez, D. Oropeza, D.J. Quarles, R.E. Russo, Multivariate classification of edible salts: simultaneous laser-induced breakdown spectroscopy and laser-ablation inductively coupled plasma mass spectrometry analysis, *Spectrochim. Acta B At. Spectrosc.* 118 (2016) 102–111.
- [116] A.C. Samuels, F.C. De Lucia, K.L. McNesby, A.W. Miziolek, Laser-induced breakdown spectroscopy of bacterial spores, molds, pollens, and protein: initial studies of discrimination potential, *Appl. Opt.* 42 (30) (2003) 6205.
- [117] M.D. Dyar, J.M. Tucker, S. Humphries, S.M. Clegg, R.C. Wiens, M.D. Lane, Strategies for Mars remote laser-induced breakdown spectroscopy analysis of sulfur in geological samples, *Spectrochim. Acta B At. Spectrosc.* 66 (1) (2011) 39–56.
- [118] F.M.V. Pereira, D.M.B.P. Milori, A.L. Venâncio, M.d.S.T. Russo, P.K. Martins, J. Freitas-Astúa, Evaluation of the effects of *Candidatus* *Liberibacter asiaticus* on inoculated citrus plants using laser-induced breakdown spectroscopy (LIBS) and chemometrics tools, *Talanta* 83 (2) (2010) 351–356.
- [119] C.A. Munson, F.C.J. 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 B At. Spectrosc.* 60 (7–8) (2005) 1217–1224.
- [120] F. Anabitarte, J. Mirapeix, O.M.C. Portilla, J.M. Lopez-Higuera, A. Cobo, Sensor for the detection of protective coating traces on boron steel with Aluminium-silicon covering by means of laser-induced breakdown spectroscopy and support vector machines, *IEEE Sensors J.* 12 (1) (2012) 64–70.
- [121] D. Marcos-Martinez, J.A. Ayala, R.C. Izquierdo-Hornillos, F.J.M. de Villena, J.O. Caeres, Identification and discrimination of bacterial strains by laser induced breakdown spectroscopy and neural networks, *Talanta* 84 (3) (2011) 730–737.
- [122] I.-G. Chong, C.-H. Jun, Performance of some variable selection methods when multicollinearity is present, *Chemom. Intell. Lab. Syst.* 78 (1–2) (2005) 103–112.
- [123] F.C.J. De Lucia, J.L. Gottfried, Influence of variable selection on partial least squares discriminant analysis models for explosive residue classification, *Spectrochim. Acta B At. Spectrosc.* 66 (2) (2011) 122–128.
- [124] H. Fink, U. Panne, R. Niessner, Process analysis of recycled thermoplasts from consumer electronics by laser-induced plasma spectroscopy, *Anal. Chem.* 74 (17) (2002) 4334–4342.
- [125] M. Corsi, G. Cristoforetti, M. Giuffrida, M. Hidalgo, S. Legnaioli, L. Masotti, V. Palleschi, A. Salvetti, E. Tognoni, C. Vellabona, A. Zanini, Archaeometric analysis of ancient copper artefacts by laser-induced breakdown spectroscopy technique, *Microchim. Acta* 152 (1–2) (2005) 105–111.
- [126] A. Ciucci, M. Corsi, V. Palleschi, S. Rastelli, A. Salvetti, E. Tognoni, New procedure for quantitative elemental analysis by laser-induced plasma spectroscopy, *Appl. Spectrosc.* 53 (8) (1999) 960–964.
- [127] A.K. Myakalwar, N. Spegazzini, C. Zhang, S. Anubham, R.R. Dasari, I. Barman, M.K. Gundawar, Less is more: avoiding the LIBS dimensionality curse through judicious feature selection for explosive detection, *Sci. Rep.* 5 (1) (2015).
- [128] R.A. Multari, D.A. Cremers, J.M. Dupre, J.E. Gustafson, The use of laser-induced breakdown spectroscopy for distinguishing between bacterial pathogen species and strains, *Appl. Spectrosc.* 64 (7) (2010) 750–759.
- [129] R. Multari and D. Cremers, "Methods For Forming Recognition Algorithms For

- Laser-Induced Breakdown Spectroscopy". Patent US20110246145 A1, 2012.
- [130] I.B. Gornushkin, U. Panne, J.D. Winefordner, Linear correlation for identification of materials by laser induced breakdown spectroscopy: improvement via spectral filtering and masking, *Spectrochim. Acta B At. Spectrosc.* 64 (10) (2009) 1040–1047.
- [131] A.K. Myakalwar, S. Sreedhar, I. Barman, N.C. Dingari, S. Venugopal Rao, P. Prem Kiran, G. Manoj Kumar, Laser-induced breakdown spectroscopy-based investigation and classification of pharmaceutical tablets using multivariate chemometric analysis, *Talanta* 87 (2011) 53–59.
- [132] P. Yaroshchik, D.L. Death, S.J. Spencer, Comparison of principal components regression, partial least squares regression, multi-block partial least squares regression, and serial partial least squares regression algorithms for the analysis of Fe in iron ore using LIBS, *J. Anal. At. Spectrom.* 27 (1) (2012) 92–98.
- [133] J.E. Carranza, D.W. Hahn, Sampling statistics and considerations for single-shot analysis using laser-induced breakdown spectroscopy, *Spectrochim. Acta B At. Spectrosc.* 57 (4) (2002) 779–790.
- [134] V. Lazic, F. Colao, R. Fantoni, V. Spizzicchio, Laser-induced breakdown spectroscopy in water: improvement of the detection threshold by signal processing, *Spectrochim. Acta B At. Spectrosc.* 60 (7–8) (2005) 1002–1013.
- [135] D. Diaz, D.W. Hahn, A. Molina, Quantification of gold and silver in minerals by laser-induced breakdown spectroscopy, *Spectrochim. Acta B At. Spectrosc.* 136 (2017) 106–115.
- [136] S.-J. Choi, K.-J. Lee, J.J. Yoh, Quantitative laser-induced breakdown spectroscopy of standard reference materials of various categories, *Appl. Phys. B Lasers Opt.* 113 (3) (2013) 379–388.
- [137] J. El Haddad, M. Villot-Kadri, A. Ismaël, G. Gallou, K. Michel, D. Bruyère, B. Bousquet, Artificial neural network for on-site quantitative analysis of soils using laser induced breakdown spectroscopy, *Spectrochim. Acta B At. Spectrosc.* 79–80 (2013) 51–57.
- [138] R. Wisbrun, I. Schechter, R. Niessner, H. Schroeder, K.L. Kompa, Detector for trace elemental analysis of solid environmental samples by laser plasma spectroscopy, *Anal. Chem.* 66 (18) (1994) 2964–2975.
- [139] M.M. Tripathi, K.E. Eseller, F.-Y. Yueh, J.P. Singh, Multivariate calibration of spectra obtained by laser induced breakdown spectroscopy of plutonium oxide surrogate residues, *Spectrochim. Acta B At. Spectrosc.* 64 (11–12) (2009) 1212–1218.
- [140] F.R. Doucet, T.F. Belliveau, J.-L. Fortier, J. Hubert, Use of chemometrics and laser-induced breakdown spectroscopy for quantitative analysis of major and minor elements in aluminum alloys, *Appl. Spectrosc.* 61 (3) (2007) 327–332.
- [141] P. Devangad, V.K. Unnikrishnan, M.M. Tamboli, K.M.M. Shameem, R. Nayak, K.S. Choudhari, C. Santhosh, Quantification of Mn in glass matrices using laser induced breakdown spectroscopy (LIBS) combined with chemometric approaches, *Anal. Methods* 8 (39) (2016) 7177–7184.
- [142] M.Z. Martin, N. Labbé, T.G. Rials, S.D. Wullschlegler, Analysis of preservative-treated wood by multivariate analysis of laser-induced breakdown spectroscopy spectra, *Spectrochim. Acta B At. Spectrosc.* 60 (7–8) (2005) 1179–1185.
- [143] M.Z. Martin, N. Labbé, N. André, R. Harris, M. Ebinger, S.D. Wullschlegler, A.A. Vass, High resolution applications of laser-induced breakdown spectroscopy for environmental and forensic applications, *Spectrochim. Acta B At. Spectrosc.* 62 (12) (2007) 1426–1432.
- [144] M.Z. Martin, S. Allman, D.J. Brice, R.C. Martin, N.O. Andre, Exploring laser-induced breakdown spectroscopy for nuclear materials analysis and in-situ applications, *Spectrochim. Acta B At. Spectrosc.* 74–75 (2012) 177–183.
- [145] R.B. Anderson, R.V. Morris, S.M. Clegg, J.F. Bell, R.C. Wiens, S.D. Humphries, R. McInroy, The influence of multivariate analysis methods and target grain size on the accuracy of remote quantitative chemical analysis of rocks using laser induced breakdown spectroscopy, *Icarus* 215 (2) (2011) 608–627.
- [146] R.A. Multari, D.A. Cremers, J.A.M. Dupre, J.E. Gustafson, Detection of biological contaminants on foods and food surfaces using laser-induced breakdown spectroscopy (LIBS), *J. Agric. Food Chem.* 61 (36) (2013) 8687–8694.
- [147] G. Bilge, H.M. Velioglu, B. Sezer, K.E. Eseller, I.H. Boyaci, Identification of meat species by using laser-induced breakdown spectroscopy, *Meat Sci.* 119 (2016) 118–122.
- [148] G. Bilge, B. Sezer, K.E. Eseller, H. Berberoglu, A. Topcu, I.H. Boyaci, Determination of whey adulteration in milk powder by using laser induced breakdown spectroscopy, *Food Chem.* 212 (2016) 183–188.
- [149] 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.
- [150] B. Sallé, J.-L. Lacour, P. Mauchien, P. Fichet, S. Maurice, G. Manhès, Comparative study of different methodologies for quantitative rock analysis by laser-induced breakdown spectroscopy in a simulated Martian atmosphere, *Spectrochim. Acta B At. Spectrosc.* 61 (3) (2006) 301–313.
- [151] J.W.B. Braga, L.C. Trevizan, L.C. Nunes, I.A. Rufini, D.J. Santos, F.J. Krug, Comparison of univariate and multivariate calibration for the determination of micronutrients in pellets of plant materials by laser induced breakdown spectroscopy, *Spectrochim. Acta B At. Spectrosc.* 65 (1) (2010) 66–74.
- [152] P. Yaroshchik, D.L. Death, S.J. Spencer, Quantitative measurements of loss on ignition in Iron ore using laser-induced breakdown spectroscopy and partial least squares regression analysis, *Appl. Spectrosc.* 64 (12) (2010) 1335–1341.
- [153] C. Gottlieb, S. Millar, T. Günther, G. Wilsch, Revealing hidden spectral information of chlorine and sulfur in data of a mobile laser-induced breakdown spectroscopy system using chemometrics, *Spectrochim. Acta B At. Spectrosc.* 132 (2017) 43–49.
- [154] C. Bohling, K. Hohmann, D. Scheel, C. Bauer, W. Schippers, J. Burgmeier, W. Schade, All-fiber-coupled laser-induced breakdown spectroscopy sensor for hazardous materials analysis, *Spectrochim. Acta B At. Spectrosc.* 62 (12) (2007) 1519–1527.
- [155] C. Mcmanus, J. Dowe, T. Likes, J. Dowe, P. Mcmanus, N. Winkless, A. Shelton, K. Baughn, R. Baughn and R. Orr, "Methods and systems for analyzing samples". Patent WO2014150696 A1, 2014.
- [156] A.M. Ollila, J. Lasue, H.E. Newsom, R.A. Multari, R.C. Wiens, S.M. Clegg, Comparison of two partial least squares-discriminant analysis algorithms for identifying geological samples with the ChemCam laser-induced breakdown spectroscopy instrument, *Appl. Opt.* 51 (7) (2012) B130.
- [157] J.D. Hybl, G.A. Lithgow, S.G. Buckley, Laser-induced breakdown spectroscopy detection and classification of biological aerosols, *Appl. Spectrosc.* 57 (10) (2003) 1207–1215.
- [158] M. Baudalet, J. Yu, M. Bossu, J. Jovelet, J.-P. Wolf, T. Amodeo, P. Laloi, Discrimination of microbiological samples using femtosecond laser-induced breakdown spectroscopy, *Appl. Phys. Lett.* 89 (16) (2006) 163903.
- [159] F.C.J. 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 B At. Spectrosc.* 62 (12) (2007) 1399–1404.
- [160] H. Xia, M.C.M. Bakker, Reliable classification of moving waste materials with LIBS in concrete recycling, *Talanta* 120 (2014) 239–247.
- [161] B.K. Lavine, C.E. Davidson, A.J. Moores, P.R. Griffiths, Raman spectroscopy and genetic algorithms for the classification of wood types, *Appl. Spectrosc.* 55 (8) (2001) 960–966.
- [162] F.L. Martin, M.J. German, E. Wit, T. Fearn, N. Ragavan, H.M. Pollock, Identifying variables responsible for clustering in discriminant analysis of data from infrared microspectroscopy of a biological sample, *J. Comput. Biol.* 14 (9) (2007) 1176–1184.
- [163] B. Vekemans, L. Vincze, F.E. Brenker, F. Adams, Processing of three-dimensional microscopic X-ray fluorescence data, *J. Anal. At. Spectrom.* 19 (10) (2004) 1302.
- [164] O. Samek, H.H. Telle, D.C. Beddows, Laser-induced breakdown spectroscopy: a tool for real-time, in vitro and in vivo identification of carious teeth, *BMC Oral Health* 1 (1) (2001).
- [165] O. Samek, V. Krzyžánek, D. Beddows, H. Telle, J. Kaiser, M. Liška, Material Identification Using Laser Spectroscopy, *Computer Analysis of Images and Patterns: 9th International Conference, LNCS 2124, Warsaw, Poland, (2001).*
- [166] F.-Y. Yueh, H. Zheng, J.P. Singh, S. Burgess, Preliminary evaluation of laser-induced breakdown spectroscopy for tissue classification, *Spectrochim. Acta B At. Spectrosc.* 64 (10) (2009) 1059–1067.
- [167] J.L. Gottfried, F.C.J. De Lucia, C.A. Munson, A.W. Miziolek, Double-pulse standoff laser-induced breakdown spectroscopy for versatile hazardous materials detection, *Spectrochim. Acta B At. Spectrosc.* 62 (12) (2007) 1405–1411.
- [168] 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 (48) (2013) 25872.
- [169] D. Pokrajac, A. Lazarevic, V. Kecman, A. Marciano, Y. Markushin, T. Vance, N. Melikechi, Automatic classification of laser-induced breakdown spectroscopy (LIBS) data of protein biomarker solutions, *Appl. Spectrosc.* 68 (9) (2014) 1067–1075.
- [170] G. Vítková, L. Prokeš, K. Novotný, P. Pořízka, J. Novotný, D. Všianský, 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 B At. Spectrosc.* 101 (2014) 191–199.
- [171] S. Merk, C. Scholz, S. Florek, D. Mory, Increased identification rate of scrap metal using laser induced breakdown spectroscopy echelle spectra, *Spectrochim. Acta B At. Spectrosc.* 112 (2015) 10–15.
- [172] T. Vance, N. Reljin, A. Lazarevic, D. Pokrajac, V. Kecman, N. Melikechi, S. McDaniel, Classification of LIBS Protein Spectra Using Support Vector Machines and Adaptive Local Hyperplanes, The 2010 International Joint Conference on Neural Networks (IJCNN), Barcelona, Spain, (2010).
- [173] J.R. Cordeiro, M.I.V. Martínez, R.W.C. Li, A.P. Cardoso, L.C. Nunes, F.J. Krug, J. Gruber, Identification of four wood species by an electronic nose and by LIBS, *Int. J. Electrochem.* (2012) 1–5.
- [174] G. Kim, J. Kwak, K.-R. Kim, H. Lee, K.-W. Kim, H. Yang, K. Park, Rapid detection of soils contaminated with heavy metals and oils by laser induced breakdown spectroscopy (LIBS), *J. Hazard. Mater.* 263 (2013) 754–760.
- [175] G. Park, H. Yoo, Y. Gong, S. Cui, S.-H. Nam, K.-S. Ham, Y. Lee, Feasibility of rapid classification of edible salts by a compact low-cost laser-induced breakdown spectroscopy device, *Bull. Kor. Chem. Soc.* 36 (1) (2015) 189–197.
- [176] M.J.C. Pontes, J. Cortez, R.K.H. Galvão, C. Pasquini, M.C.U. Araújo, R.M. Coelho, B.E. Madari, Classification of Brazilian soils by using LIBS and variable selection in the wavelet domain, *Anal. Chim. Acta* 642 (1–2) (2009) 12–18.
- [177] S. Wold, M. Sjöstöm, SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy, ACS Symposium Series, (1977), pp. 243–282.
- [178] S.M. Clegg, E. Sklute, M.D. Dyar, J.E. Barefield, R.C. Wiens, Multivariate analysis of remote laser-induced breakdown spectroscopy spectra using partial least squares, principal component analysis, and related techniques, *Spectrochim. Acta B At. Spectrosc.* 64 (1) (2009) 79–88.
- [179] S. Duchene, V. Detalle, R. Bruder, J. Sirven, Chemometrics and laser induced breakdown spectroscopy (LIBS) analyses for identification of wall paintings pigments, *Curr. Anal. Chem.* 6 (1) (2010) 60–65.
- [180] J.L. Gottfried, F.C.J. De Lucia, C.A. Munson, A.W. Miziolek, Strategies for residue explosives detection using laser-induced breakdown spectroscopy, *J. Anal. At. Spectrom.* 23 (2) (2008) 205–216.
- [181] Q. Godoi, F.O. Leme, L.C. Trevizan, E.R. Pereira Filho, I.A. Rufini, D.J. Santos, F.J. Krug, Laser-induced breakdown spectroscopy and chemometrics for

- classification of toys relying on toxic elements, *Spectrochim. Acta B At. Spectrosc.* 66 (2) (2011) 138–143.
- [182] N.C. Dingari, I. Barman, A.K. Myakalwar, S.P. Tewari, M. Kumar Gundawar, Incorporation of support vector machines in the LIBS toolbox for sensitive and robust classification amidst unexpected sample and system variability, *Anal. Chem.* 84 (6) (2012) 2686–2694.
- [183] S. Moncayo, S. Manzoor, F. Navarro-Villoslada, J.O. Caceres, Evaluation of supervised chemometric methods for sample classification by laser induced breakdown spectroscopy, *Chemom. Intell. Lab. Syst.* 146 (2015) 354–364.
- [184] J.J. Remus, J.L. Gottfried, R.S. Harmon, A. Draucker, D. Baron, R. Yohe, Archaeological applications of laser-induced breakdown spectroscopy: an example from the Coso Volcanic Field, California, using advanced statistical signal processing analysis, *Appl. Opt.* 49 (13) (2010) C120.
- [185] M. Barker, W. Rayens, Partial least squares for discrimination, *J. Chemom.* 17 (3) (2003) 166–173.
- [186] J.L. Gottfried, F.C. De Lucia Jr., A.W. Miziolek, Discrimination of explosive residues on organic and inorganic substrates using laser-induced breakdown spectroscopy, *J. Anal. At. Spectrom.* 24 (3) (2009) 288.
- [187] I.B. Gornushkin, B.W. Smith, H. Nasajpour, J.D. Winefordner, Identification of solid materials by correlation analysis using a microscopic laser-induced plasma spectrometer, *Anal. Chem.* 71 (22) (1999) 5157–5164.
- [188] R. Sattmann, I. Mönch, H. Krause, R. Noll, S. Couris, A. Hatzia Apostolou, R. Miguel, Laser-induced breakdown spectroscopy for polymer identification, *Appl. Spectrosc.* 52 (3) (1998) 456–461.
- [189] V. Motto-Ros, A.S. Koujelev, G.R. Osinski, A.E. Dudelzak, Quantitative multi-elemental laser-induced breakdown spectroscopy using artificial neural networks, *J. Eur. Opt. Soc.* 3 (2008).
- [190] M. Boueri, V. Motto-Ros, W. Lei, L. Qain-LiMa, Z.H. Zheng, J. Yu, Identification of polymer materials using laser-induced breakdown spectroscopy combined with artificial neural networks, *Appl. Spectrosc.* 65 (2011) 307–314.
- [191] S.J. Rehse, J. Diedrich, S. Palchaudhuri, Identification and discrimination of *Pseudomonas aeruginosa* bacteria grown in blood and bile by laser-induced breakdown spectroscopy, *Spectrochim. Acta B At. Spectrosc.* 62 (10) (2007) 1169–1176.
- [192] S.J. Rehse, Q.I. Mohaidat, S. Palchaudhuri, Towards the clinical application of laser-induced breakdown spectroscopy for rapid pathogen diagnosis: the effect of mixed cultures and sample dilution on bacterial identification, *Appl. Opt.* 49 (13) (2010) C27.
- [193] E. Grifoni, S. Legnaioli, G. Lorenzetti, S. Pagnotta, V. Palleschi, Application of graph theory to unsupervised classification of materials by laser-induced breakdown spectroscopy, *Spectrochim. Acta B At. Spectrosc.* 118 (1) (2016) 40–44.