



# VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

## FAKULTA STROJNÍHO INŽENÝRSTVÍ

FACULTY OF MECHANICAL ENGINEERING

## ÚSTAV FYZIKÁLNÍHO INŽENÝRSTVÍ

INSTITUTE OF PHYSICAL ENGINEERING

### VÝVOJ ALGORITMU PRO AUTOMATICKOU CHARAKTERIZACI VZORKU NA ZÁKLADĚ DAT ZÍSKANÝCH SPEKTROSKOPIÍ LASEREM INDUKOVANÉHO PLAZMATU (LIBS)

DEVELOPMENT OF AN ALGORITHM FOR AUTOMATIC CHARACTERIZATION OF A SAMPLE BASED ON  
THE DATA RECEIVED BY MEANS OF THE LASER INDUCED BREAKDOWN SPECTROSCOPY (LIBS)

#### ZKRÁCENÁ VERZE DIZERTAČNÍ PRÁCE

DOCTORAL THESIS, SHORT VERSION

#### AUTOR PRÁCE

AUTHOR

Ing. JAKUB KLUS

#### VEDOUCÍ PRÁCE

SUPERVISOR

prof. Ing. JOZEF KAISER, Ph.D.

BRNO 2017

**Klíčová slova**

spektroskopie laserem indukovaného plazmatu, multivariační analýza

**Keywords**

Laser-Induced Breakdown Spectroscopy, Multivariate Data Analysis

© Jakub KLUS, 2017

**Místo uložení práce:**

Vysoké učení technické v Brně

Fakulta strojního inženýrství

Technická 2896/2

616 69 Brno

Česká Republika

ISSN 1213-4198

ISBN 80-214-??

# Obsah

Úvod	5
Cíle práce	7
<b>1 Konvenční metody analýzy</b>	<b>8</b>
1.1 Kvantitativní analýza . . . . .	8
1.2 Kalibrační křivka . . . . .	8
1.3 Bezkalibrační LIBS . . . . .	9
1.4 Inverzní bezkalibrační LIBS . . . . .	9
1.5 Metoda jednobodové kalibrace . . . . .	10
1.6 CSigma LIBS . . . . .	10
<b>2 Současný stav automatické charakterizace</b>	<b>12</b>
2.1 Korelační analýza . . . . .	12
2.2 Software pro semiautomatické zpracování dat . . . . .	12
2.3 Metoda normalizovaných souřadnic . . . . .	13
2.4 Vektorový model zpracování textu . . . . .	13
2.5 Korelace s teoretickým modelem . . . . .	14
2.6 Záměna spektrálních čar automatickými algoritmy . . . . .	14
<b>3 Vybrané multivariační statistické metody</b>	<b>15</b>
3.1 Metoda hlavních komponent (PCA) . . . . .	15
3.2 Algoritmy podpůrných vektorů (SVM) . . . . .	16
3.3 Lineárně oddělitelná data . . . . .	16
3.4 Samouspořádané mapy (SOM) . . . . .	17
3.5 Základní algoritmus . . . . .	18
<b>4 Vybrané aplikace automatické analýzy</b>	<b>19</b>
4.1 Vliv normalizace na přesnost klasifikace . . . . .	19
4.2 Charakterizace uranonosné rudy pomocí PCA . . . . .	19
4.3 Charakterizace uranonosné rudy pomocí SOM . . . . .	20
4.4 Klasifikace hornin z echellogramů . . . . .	21
<b>5 Nový koncept automatické identifikace prvků ve spektru</b>	<b>22</b>
5.1 Postup automatické identifikace . . . . .	22
5.2 Vyhodnocování výsledků . . . . .	23
<b>6 Závěr</b>	<b>24</b>
Použitá literatura	25
Curriculum Vitae	29
Abstrakt	31



# Úvod

Metoda spektroskopie laserem indukovaného plazmatu (LIBS) patří mezi skupinu analytických metod zabývajících se atomovou emisní spektroskopií. Zdrojem záření pro LIBS je horké plazma, které vzniká těsně nad povrchem studovaného vzorku při dopadu elektromagnetického záření o vysoké plošné hustotě výkonu (řádově v  $\text{GW} \cdot \text{cm}^{-2}$ ). Potřebná plošná hustota výkonu je zpravidla dodávána pevnolátkovými lasery v pulzech, které trvají jen zlomky sekund. Vybuzené plazma má tvar rotačně symetrické chocholky o velikosti několika milimetrů. Záření plazmatu tvoří především zářivé přechody elektronů vázaných v atomech a iontech jednotlivých prvků tvořících vzorek a okolní atmosféru. Ionty a elektrony v plazmatu jsou v lokální termodynamické rovnováze, a tak je plazma navenek neutrální a vykazuje kolektivní chování. Takový systém částic je nestabilní, plazma září řádově desítky mikrosekund a vyzařované spektrum se během této doby dramaticky mění. Dopad časového vývoje plazmatu lze zmírnit pomocí speciálních detektorů, které umožňují měření vysokých intenzit za krátkou dobu, tj. časově rozlišená spektroskopie.

Přestože kladou popsané jevy vysoké požadavky na instrumentaci potřebnou v LIBS, nachází tato metoda své uplatnění v mnoha průmyslových odvětvích a to díky nízkým provozním nákladům a nízkým nárokům na provozní plyny či chemikálie (v nejjednodušším případě je potřeba pouze deionizovaná voda pro chlazení laseru a inertní plyn pro vytvoření atmosféry se stabilním složením nad povrchem vzorku). Srovnatelné metody atomové emisní spektroskopie, jako jsou optická nebo hmotnostní spektroskopie indukčně vázaného plazmatu, vyžadují složitou přípravu vzorku spočívající v jeho rozpouštění a následném ředění. To je hlavním know-how potřebným pro úspěšné zvládnutí výše zmíněných metod. Metoda LIBS umožňuje studium vzorků ve všech skupenstvích a principiálně bez specifické přípravy. To otvírá možnosti četným průmyslovým aplikacím vyžadujícím analýzu v reálném čase nebo ve ztížených podmínkách. Mezi nejúspěšnější aplikace LIBS ve ztížených podmínkách lze určitě zařadit aparaturu ChemCam, která je umístěna na robotickém vozítku Curiosity, které slouží k analýze povrchu planety Mars.

Analýza LIBS spektra se většinou zabývá prvkovým složením vzorku. Výsledky prvkové analýzy lze prezentovat dvojím způsobem. První z nich je kvalitativní analýza, která se zabývá pouze přítomností prvků ve vzorku. Výsledky kvalitativní analýzy lze uplatnit například při hloubkovém profilování nebo při třídění kovů. Druhým způsobem je kvantitativní analýza, která se zabývá množstvím prvků ve studovaném vzorku. Kupříkladu studium koncentrace stopových prvků v oceli může být zásadním příspěvkem při kontrole jakosti materiálů nejen vstupujících do výrobního procesu, ale i těch, jež výrobní linky opouští. Přísně vzato je kvantitativní analýza LIBS schopna odhalit pouze složení prvků v plazmatu, které nemusí přesně korespondovat se složením vzorku. K těmto nepříznivým jevům však dochází jen zřídka a lze je v mnoha případech zanedbat. Kvantitativní analýzu lze provádět jednak přímo vytvořením kalibrační křivky pomocí intenzit spektrálních čar a známých koncentrací prvků ve vzorku, anebo sofistikovanou rozvahou nad kvantovými přechody jednotlivých prvků, jež tvoří elektromagnetické spektrum detekované detektorem. V takových případech hovoříme o bezkalibrační metodě. V posledních letech přibýly do repertoáru LIBS také izotopová analýza a analýza molekulových spekter. První z nich představuje slibnou možnost studia těžkých radioaktivních prvků nebo určování stáří z poměru izotopů boru (borové datování). Možnosti analýzy molekulových spekter, jak je tom například u C–N přechodů, se prozatím nacházejí spíše v teoretické oblasti.

Způsoby uplatnění LIBS jsou nepřehledné a díky výkonnější, rychlejší, levnější a přesnější instrumentaci se metodě LIBS pomalu dostává značné pozornosti a to nejen ve vědeckém výzkumu ale i v průmyslu. S rostoucí opakovací frekvencí měření a se zvětšujícím se počtem měřených spektrálních proměnných dosahují datové soubory, které jsou výstupem LIBS, miliard proměnných, a při jejich zpracování již nelze použít konvenční techniky zakládající se na manuálním přiřazování spektrálních čar k píkům ve spektru. Řešením problémů velkých datových souborů o značném počtu problémů se zabývá multivariační statistická analýza. Multivariační analýza se často spojuje s dalšími obory. V případě chemie tak vzniká obor chemometrie, při analýze obrazu jde o strojové učení a u rozsáhlých nestrukturovaných dat to je „Big-Data“ analýza. Přínos chemometrie, nebo obecně multivariační analýzy do LIBS, lze spatřit především v aplikaci klasifikační analýzy, kdy lze jednotlivá měření zařadit do tříd podle jejich chemické podobnosti. Další způsob multivariační analýzy je přehledová analýza, která slouží ke hledání struktury v datech, v případě spektroskopie se jedná o hledání vlnových délek, které reprezentují významné rozdíly mezi jednotlivými měřeními. Tyto obecné poznatky, společně s chemickým nebo fyzikálním popisem vzorků nebo experimentálních podmínek, lze využít pro detailnější analýzu vzorků nebo prvků v nich. Díky poznatkům multivariační analýzy je navíc možné zjednodušit nebo zpřesnit konvenční metody zpracování LIBS dat. Právě možnostmi a důsledky propojení multivariační analýzy a metody LIBS se předkládaná disertační práce zabývá.

Charakterizaci dat můžeme vnímat jako automatickou identifikaci významných částí spektra, klasifikaci na základě předem stanovených tříd nebo přiřazení spojité veličiny naměřeným spektrům. V jednotlivých kapitolách jsou pak popsány různé nástroje statistické analýzy, jejich možné uplatnění a také diskuze nad získanými výsledky.

# Cíle práce

Disertační práce se zabývá možnostmi automatické charakterizace vzorků pomocí multivariačních metod. Cílem práce je uplatnit znalosti multivariační statistické analýzy zejména na tyto tři problematiky v LIBS:

- I. Přehledová analýza chemických map** Přední výhodou LIBS je možnost mapování heterogenních vzorků, kterými mohou být geologické vzorky rud, archeologické vzorky zubů, biologické vzorky tkání a jiné. V takovém případě je důležitá informace ukryta pouze ve zlomku naměřených dat a je obtížné měření nesoucí tuto specifickou informaci najít. Přehledová analýza umožňuje nalézt vlnové délky nebo spektrální čáry, které souvisí s největšími změnami ve spektru. Kvalitativní analýza nalezených čar najde uplatnění například v geologických studiích a v těžařském průmyslu. Užití přehledové analýzy je demonstrováno v kapitole 4.2.
- II. Redukce dimenzionality dat** Multivariační statistické metody je možné použít i pro výběr reprezentativních měření v rozsáhlém souborů měření. S redukováním souborem dat lze pracovat jako s původními daty a zkoumat tak, které prvky se v jednotlivých měřeních vyskytují. Tento přístup byl zvolen v případě mapování heterogenního vzorku popsaného v kapitole 4.3.
- III. Klasifikace vzorků** Klasifikace neboli třídění vzorků je přístup zpracování dat, který nezapadá mezi konvenční metody analýzy v LIBS. Klasifikační metody umožňují rozdělit jednotlivé vzorky do předem stanovených skupin. Uplatnění klasifikace lze nalézt například při zkoumání geografického původu vzorků, rozlišování jednotlivých chemických matic nebo analýze kvality. Kapitola 4.1 se zabývá předúpravou spektroskopických dat pro klasifikaci a dále kapitola 4.4 se zabývá novou možností klasifikace vzorků na základě surových dat z echelle spektrometrů.

# 1. Konvenční metody analýzy

## 1.1. Kvantitativní analýza

Kvantitativní analýza spočívá ve zjištění neznámé koncentrace zkoumané látky ve studovaném vzorku. Hlavní výhoda LIBS (nízké nároky na přípravu vzorku) se v kvantitativní analýze jeví jako jedna z hlavních nevýhod. Aby bylo dosaženo dostatečné přesnosti měření, je nutné dodržet co nejlépe shodnost experimentálních podmínek při kalibraci zařízení i při měření neznámých vzorků.

Podle [1] patří mezi hlavní faktory ovlivňující přesnost kvantitativní analýzy tyto parametry experimentů LIBS:

- Energie, vlnová délka a opakovací frekvence laseru.
- Citlivost a linearita detektoru, případně přesnost časování.
- Ohnisková vzdálenost fokusační čočky, nastavení sběru záření, složení a tlak okolního plynu.
- Homogenita složení vzorku, matriční jevy.

## 1.2. Kalibrační křivka

Základní způsob kvantitativní analýzy spočívá v sestrojení kalibrační křivky (anglicky „calibration curve“<sup>1</sup>). To je závislost změny velikosti detekovaného signálu na koncentraci studovaného prvku.

Kalibrační křivky se používají i v ostatních analytických metodách, nicméně v LIBS je obtížné získat podobné vzorky s různými koncentracemi studovaného prvku (takzvané matričně odpovídající vzorky (anglicky „matrix-matched standards“)). To platí zejména v případě pevných látek, kdy nejsou dostupné kvalitní kalibrační referenční materiály a změny koncentrace jsou zajišťovány při peletizaci práškových standardů s různým poměrem přídavného plniva. Například u ICP je vzorek rozpuštěn a změny koncentrace je následně dosaženo přidáním nosného média v různých poměrech. Tento fakt ukazuje, že za všech okolností je kvantitativní analýza limitována na daný typ vzorku a na omezenou škálu koncentrací. Detekovaným signálem je v jednodušším případě intenzita vyzařované čáry daného prvku nebo ve složitějších případech poměr intenzit dvou a více čar. Komplexnější modely zahrnují do kalibrační křivky i koncentrace dalších prvků (princip používaný v rentgenové fluorescenci).

Kvalitu kalibrační křivky, respektive kvantitativní analýzy, lze vyjádřit pomocí limitů detekce (LOD z anglického „limit of detection“). LOD je nejmenší množství nebo koncentrace hledané látky, které spolehlivě vyvolá signál na detektoru. Podle [1] můžeme LOD vypočítat z:

$$c_L = k_{sbl}m, \quad (1.1)$$

<sup>1</sup>V některých pracích se kalibrační křivce říká křivka růstu (COG z anglického „curve of growth“).



kde  $c_L$  je limit detekce,  $s_{bl}$  je standardní odchylka měření s nulovou koncentrací (blank měření),  $m$  je směrnice kalibrační křivky a  $k_c$  je předem zvolené celé číslo vyjadřující spolehlivost limitu detekce, pro  $k_c = 3$  můžeme uvažovat hladinu spolehlivosti 90 %.

Výše zmíněné nevýhody sestavování kalibračních křivek byly motivací pro odvození bezkalibračních metod založených na teoretickém popisu záření plazmatu.

### 1.3. Bezkalibrační LIBS

V roce 1999 Ciucci a kol. [2] provedli podrobný rozbor rovnic popisujících záření plazmatu a odvodili metodu, která umožňuje vypočítat absolutní koncentrace zářičů v plazmatu úplnou analýzou spektra. Později se tato metoda začala označovat zkratkou CF-LIBS z anglického názvu „calibration-free LIBS“.

Předpoklady použitelnosti CF-LIBS jsou lokální termodynamický rovnováha (LTE) a opticky tenké plazma. Základním krokem CF-LIBS je stanovení elektronové teploty. Intenzity čar jednotlivých prvků jsou pak zaneseny do Boltzmannova grafu, kde každý prvek a každý stupeň ionizace je reprezentován přímkou. Za předpokladu LTE jsou všechny přímky rovnoběžné a úseky vytyčené každou přímkou na ose  $y$  vyjadřují obsazenost stavů. Samotné stanovení absolutní koncentrace pak je jen převedení této obsazenosti na základě poměru partiční funkce a součtu partičních funkcí všech prvků v plazmatu.

Poslední krok znamená v některých případech zásadní obtíž. V případech, kdy se jedná o vzorky s podobným chemickým složením (například oceli se známým počtem možných legujících prvků), lze CF-LIBS považovat za velice efektivní. Avšak jedná-li se o neznámé vzorky, může kvantitativní analýze bránit nemožnost zařazení všech prvků pouze na základě přítomnosti jejich spektrální informace, a to kvůli spektrální interferenci nebo matričnímu jevu.

Z výše uvedeného popisu pak vyplývají další nepřímé předpoklady: vyzařování plazmatu musí být homogenní, stechiometrické poměry prvků v plazmatu musí být shodné s těmi ve vzorku a spektroskopické informace o studovaných přechodech musí být dostupné. Všechny předpoklady kladou zvýšené požadavky na stanovení přesných experimentálních podmínek.

### 1.4. Inverzní bezkalibrační LIBS

Inverzní bezkalibrační metoda (CFI-LIBS z anglického „calibration-free inverse LIBS“) byla popsána v roce 2012 v [3]. Ve zmiňované práci je zpochybněna přesnost metody CF-LIBS, protože dostupnost spektrometrických konstant a jejich přesnost jsou velice diskutabilní. V CFI je odvozeno, že základním parametrem ovlivňujícím přesnost kvantifikace měření je elektronová teplota, nebo v jejich případě nestechiometrického plazmatu, termodynamická teplota jednotlivých prvků. Pomocí Boltzmannových grafů jednotlivých konstituentů plazmatu spočítali, že každý prvek vykazuje jinou termodynamickou teplotu a koncentrace spočítané pomocí CF-LIBS nesouhlasí s certifikovanými hodnotami. Na základě těchto zjištění navrhuje stanovit teplotu pomocí minimalizace odchylek certifikovaných a naměřených koncentrací. Tato teplota je pak jakousi referenční hodnotou teploty pro dané nastavení aparatury a druh vzorku.

Pro určení koncentrace neznámých vzorků musí být jejich chemické složení alespoň srovnatelné se složením vzorku a také experimentální podmínky se musí přibližně shodovat s těmi při měření známého vzorku. Spektrum neznámého vzorku je pak zpracováno metodou CF-LIBS za předpokladu LTE s termodynamickou teplotou shodnou s referenční teplotou naměřenou při kalibraci.

Je vidět, že se jedná spíše o kvazi-bezkalibrační metodu, jejíž použití je omezeno na úzký rozsah vzorků a nepřináší zásadní pokrok oproti CF-LIBS. Hlavní výhodou je, že minimalizace odchylek naměřených koncentrací souvisí se spektrálními čarami, které byly pro konstrukci Boltzmannových grafů použity. Protože samoabsorbované a špatně identifikované čáry budou vykazovat velké nelinearity, napovídá rozbor odchylek v závislosti na použitých čarách o správnosti použití té či oné spektrální čáry. Sami autoři naznačují, že tato metoda může napomoci při budování spektrálních databází pro CF-LIBS a bezkalibrační LIBS metody obecně.

## 1.5. Metoda jednobodové kalibrace

Metoda jednobodová kalibrace (OPC z anglického „One-Point Calibration method“) byla publikována v roce 2013 [4]. Principiálně se tato metoda opírá o stejné předpoklady jako CFI: opticky tenké plazma, lokální termodynamická rovnováha a měření matričně shodného vzorku. Prakticky se OPC liší od CF-LIBS tím, že kromě spektrální odezvy systému se snaží kompenzovat i nejistoty ve spektrometrických konstantách  $A_{ij}$  a  $g_j$ . Naopak proti CFI-LIBS se vymezuje tím, že není nutné dodržet všechny experimentální podmínky, stačí pouze matričně shodný vzorek.

Základem metody je spočítání CF-LIBS pro vzorek se známou koncentrací jednotlivých složek. Analýzou rozdílu koncentrací jsou pak vyvozeny kalibrační koeficienty pro úpravy intenzit každého prvku zvlášť. Tyto koeficienty definují nové souřadnice Boltzmannova grafu (nebo dokonce i Sahova-Boltzmannova grafu) a umožňují tak počítat koncentrace s mnohem větší přesností.

## 1.6. CSigma LIBS

Další bezkalibrační techniku popsali v roce 2014 Aragon a Aguilera [5]. Metoda, autory pojmenovaná CSigma, staví v základu na stejných rovnicích jako CF-LIBS, avšak raději než poměry intenzit jednotlivých čar sleduje předpokládané změny intenzit v závislosti na koncentraci prvků. První krok spočívá v charakterizaci plazmatu na daném vzorku. Plazma je charakterizováno teplotou, hustotou interagujících částic a dalšími třemi parametry závislými na nastavení měřicí aparatury (přístrojová funkce). Druhým krokem je odvození takzvaných „čarových průřezů“ (z anglického „line cross-section“). Ty jsou funkcí spektroskopických parametrů čáry, uvažované teploty a hustoty částic. Vynesením závislosti poměru intenzity vyzařování studované čáry a intenzity Planckova záření (na střední vlnové délce čáry při známé teplotě plazmatu) vůči součinu koncentrace a čarového průřezu získáme tzv. „obecnou kalibrační křivku“ (z anglického „generalized curve of growth“).

Sestrojením obecné kalibrační křivky na základě vlastnosti plazmatu a experimentálního systému a následným dosazením naměřené intenzity (normované intenzitou Planc-

kova záření) lze určit koncentraci emitujících částic. Kromě zvýšených požadavků na teoretické znalosti umožňuje metoda CSigma také definovat obor koncentrací, ve kterém je pro daný problém použitelná, navíc metoda umožňuje zahrnout spektra, jež nejsou vhodná ke konstrukci CSigma grafu (neboli obecné kalibrační křivky). Předpoklady o LTE, stechiometrické ablaci a opticky tenkém plazmatu však musí být nadále dodrženy.

Kritickému shrnutí úspěšnosti CF-LIBS zhruba 10 let od jejího publikování se věnuje [6]. Poslední tři zde zmíněné přístupy k bezkalibrační kvantitativní analýze jsou srovnány v nedávné studii [7], která se zaměřuje na robustnost, použitelnost a přesnost jednotlivých metod.

## 2. Současný stav automatické charakterizace

Dosud není uznán žádný obecný postup přiřazení spektrálních čar píkům ve spektru. Kvantitativní analýzu nebo klasifikaci lze v určitém rozsahu automatizovat pomocí pokročilých statistických metod, které jsou popsány ve druhé části předkládané práce. V následujících odstavcích jsou chronologicky seřazené vybrané přístupy k automatizaci charakterizace měření pomocí LIBS.

### 2.1. Korelační analýza

Automatická charakterizace je předmětem značného zájmu. Gornushkin a kol. [8] navrhli jednoduchou metodu identifikace (klasifikace) vzorků založenou na korelaci naměřených spekter se spektry uloženými v databázi. Porovnáváním lineární nebo pořadové korelace (anglicky „linear“ a „rank correlation“) neznámého vzorku se všemi vzorky v databázi a následnou statistickou analýzou výsledků lze tato spektra přiřazovat. Navrhovanou metodu autoři dále rozšířili o maskování a filtrování spekter [9].

### 2.2. Software pro semiautomatické zpracování dat

Jeden z prvních software zaměřený na semiautomatickou analýzu spekter vzniknuvších z LIBS je popsán v [10]. Software se skládá ze tří modulů, z nichž první slouží k prohlížení databáze čar a k vyhledávání čar podle kritérií. Druhý modul umožňuje vytváření umělých spekter na základě zvolených parametrů, jmenovitě to jsou: výběr čar, jejich ionizací a difrakčních řádů, tvar funkce, lorentzovské rozšíření, gaussovo rozšíření nebo jejich součet.

Aby byly postiženy koncentrace různých prvků ve spektru, umožňuje software určitý druh proporcionality a to jak mezi prvky, tak mezi jednotlivými přechody/stavy. Posledním modulem je modul pro analýzu spekter, který dovoluje odečtení pozadí, pomocí metody popsané pro rentgenovou fluorescenci (XRF) [11]. Další funkcí je hledání spektrálních čar, které lze provádět buď manuálně nebo automaticky. Při automatickém hledání spektrálních čar musí uživatel zvolit, jaké procentuální množství spektrálních čar z databáze musí být nalezeno ve spektru (při uvažovaném spektrálním okně) a dále, jaká musí být minimální relativní intenzita těchto čar. Automatický vyhledávač pak postupně projde všechna nastavení a odhadne přítomnost prvků.

Popsaný software dává uživateli řadu funkcí, avšak prezentované principy nepředstavují všechny možnosti analýzy v LIBS. Především práce s interními intenzitami namísto partičních funkcí, elektronových hustot a termodynamických teplot vnáší určitý stín na zvládnutí problematiky LIBS. Lze se domnívat, že automatická metoda vyhledávání čar bude fungovat značně hůře u širokospektrálních analyzátorů, které umožňují detekovat téměř všechny čary, avšak absolutní intenzity mohou být značně zkreslené.

### 2.3. Metoda normalizovaných souřadnic

Ferrero a kol. [12] zvolili odlišný způsob. Místo přiřazování jednotlivých spektrálních čar vytvořili databázi měření a hodnotí míru souvztažnosti mezi naměřeným spektrem a spektry v databázi. Ve své práci se zaměřili na identifikaci plastů (konkrétně polyamid, polyethylen, polystyren, polypropylen a polyethylentereftalát (PET)). Metoda nazvaná „Metoda normalizovaných souřadnic“ (MNC z anglického „Method of Normalized Coordinates“) se zakládá na jednoduché úvaze z lineární algebry. Uvažujme neznámý vektor  $\vec{p}$ , vyjadřující spektrum v jednotlivých souřadnicích  $j$ , pak je našim úkolem najít průmět  $a_i$  do báze tvořené jednotlivými spektry v databázi  $\{v\}_i$ , což lze vektorově zapsat jako:

$$p_j = \sum_{i=1}^N a_i v_{ji}, \quad (2.1)$$

kde  $N$  je počet vektorů (spekter) v bázi (databázi),  $p_j$  je  $j$ -tá souřadnice neznámého vektoru (neznámé spektrum v bodě  $j$ ),  $v_{ji}$  je  $j$ -tá souřadnice  $i$ -tého spektra v databázi a  $a_i$  je  $i$ -tá souřadnice  $p_j$  v nových souřadnicích. Tato úvaha vede na systém lineárních rovnic, kde je potřeba vyřešit  $N$  neznámých. Maticově lze tento systém vyjádřit jako:

$$\vec{p} = \mathbf{V}\vec{a}, \quad (2.2)$$

kde  $\vec{p}$  je sloupcový vektor o velikosti  $(M \times 1)$ , kde  $M$  odpovídá počtu vlnových délek měření,  $\vec{a}$  je sloupcový vektor nových souřadnic  $(N \times 1)$  a  $\mathbf{V}$  je matice reprezentující novou souřadnou soustavu  $(M \times N)$ . Protože  $\mathbf{V}$  není obecně čtvercová matice a řešení řešení systému z (2.2) nelze přímo odvodit, navrhli autoři odvodit řešení násobením transponovanou  $\mathbf{V}^T$  zprava:

$$\mathbf{V}^T \vec{p} = \mathbf{V}^T \mathbf{V} \vec{a}, \quad (2.3)$$

řešení souřadnic  $\vec{a}$  pak je:

$$\vec{a} = (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{p} \quad (2.4)$$

Autoři se vymezují především vůči použití korelací pro identifikaci spekter [8], jež díky použití velkého množství vlnových délek dávají falešně pozitivní výsledky. Jako výhodu použití navrhovaného postupu lící možnost studia kombinací spekter, čili možnost identifikace například směsi PET a polystyrenu.

### 2.4. Vektorový model zpracování textu

Protože prostředky lineární algebry nejsou dostatečně robustní, Amato a kol. [13] hledali inspiraci v analýze textu. Navržená metoda je složitá na popis a nelze ji vyjádřit matematickými rovnicemi. Jedná se spíše o algoritmický popis přiřazování jednotlivých spektrálních čar. Základem metody je vytvoření čárové databáze, která vyjadřuje selektivitu a relevanci jednotlivých spektrálních čar každého prvku. Selektivita čáry vyjadřuje unikátnost dané střední vlnové délky na základě počtu čar ostatních prvků v databázi nalezených v jejím malém okolí. Relevantnost čáry je spočítána jako poměr intenzity záření čáry zvoleného prvku a součtu intenzit ostatních čar zvoleného prvku.

Každý prvek, který je databáze schopna rozpoznat, je pak vyjádřen vektorem vah, které vyjadřují příslušnost dané čáry k prvku, případně nula pro nesouvisející čáru. Nevýhodou je, že tento vektor vah má stejný počet prvků, jako je počet čar v databázi. Pro identifikace je pak nutné naměřené spektrum rozdělit na spektrální čáry a každé spektrální čáře přiřadit malý rozsah podobných čar v databázi, případně nulu pro všechny odlehle čáry. Takto vznikne opět vektor se stejným počtem prvků, jako je počet čar v databázi. Identifikace poté spočívá ve skalárním součinu vah neznámého spektra a postupně všech prvků v databázi. Tímto postupem dojde k výpočtům skóre každého prvku pro danou čáru, prvek s nejvyšším skóre lze přiřadit studované čáře.

## 2.5. Korelace s teoretickým modelem

Stav automatického přiřazování spektrálních čar k prvkům vedl vědeckou skupinu z Moskevské státní univerzity k odvození dalšího principu [14]. Nejprve byla vytvořena modelová spektra, kde v sedmi krocích postupně simulovali jednotlivé jevy tvarující spektrum: chemické složení plazmatu, kvantové vlastnosti atomů, rozšíření čar, Starkův jev, šířku štěrbinu detektoru, velikost pixelu, digitalizace signálu. Takto bylo vytvořeno nejprve 105 modelových spekter s různou elektronovou hustotou a teplotou. Pro tato spektra byla spočítána Pearsonova korelace. Pro elektronovou hustotu a teplotu s největším koeficientem korelace bylo vytvořeno dalších 99 modelových spekter s menším krokem. Takto byla nalezena nejvíce odpovídající elektronová teplota 0,675 eV a elektronová hustota  $10^{16,7} \text{ cm}^{-3}$ .

K přiřazení jednotlivých spektrálních čar pak dochází analýzou původního spektra a modelového spektra před přidáním rozšíření čar (tj. pouze seznam čar a jejich příslušných intenzit). Ke každému lokálnímu maximu nalezenému v naměřeném spektru je dosazen prvek nebo prvky ze spektra modelového. V případě, že lokálnímu maximu v rozsahu k nejbližším lokálním minimům náleží více spektrálních čar, za tu správnou je považována ta s největší relativní intenzitou (spočítanou vůči ostatním čarám modelového spektra). Popsaná metoda je vyzkoušena na vzorcích oceli na spektrálním rozsahu 393,34–413,04 nm, což lze považovat za náročný úkol kvůli četnosti čar železa a dalších stopových prvků v této spektrální oblasti.

## 2.6. Záměna spektrálních čar automatickými algoritmy

V [15] jsou podrobně rozebrány jevy vedoucí ke špatnému přiřazení prvků ke spektrálním čarám. Na příkladu iontové čáry hliníku Al II 281,619 nm je ukázána její zaměnitelnost s čarami Mo II 281,744 nm a Cs II 281,692 nm při různých experimentálních parametrech. Hlavní příčinou záměny je Starkův posun, který pro stanovené experimentální parametry činí až 130 pm. Autoři navíc dodávají, že Starkův jev lze významně potlačit, volíme-li větší časové zpoždění mezi detektorem a laserovým pulzem (gate delay).

# 3. Vybrané multivariační statistické metody

## 3.1. Metoda hlavních komponent (PCA)

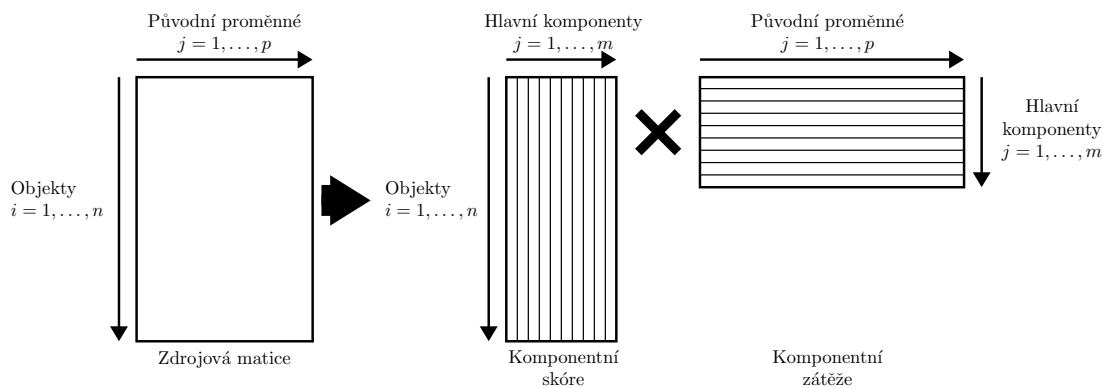
Jednou z nejpoužívanějších metod přehledové analýzy je metoda hlavních komponent (PCA z anglického „Principal Component Analysis“), ta je založena na systematickém odebrání variace v datovém souboru. Historicky vzato byla PCA zavedena již v roce 1901 Pearsonem [16], který hledal roviny a přímky nejlépe vystihující systém bodů v prostoru [17]. Hlavním cílem PCA je rozklad zdrojových dat do strukturní a šumové matice. Zjednodušeně se jedná o transformaci původních znaků na nové proměnné, které nejsou korelované (jinými slovy jsou ortogonální). Nové proměnné se nazývají hlavní komponenty a jsou to lineární kombinace původních znaků. Hlavní komponenty jsou řazeny podle toho, kolik odebírají variability ze souboru. Odebraná variabilita je základní parametr každé komponenty a používá se pro sestavení indexového grafu úpatí.

Vzhledem k tomu, že nejvíce variability odebere prvních pár komponent, slouží PCA k redukcí dimenzionality dat. Místo analýzy stovek a tisíců znaků každého objektu lze použít komponentní skóre těchto objektů například pro první tři komponenty (konstrukce rozptylového diagramu komponentního skóre).

Zjednodušeně lze matematickou podstatu PCA vyjádřit pomocí rovnice:

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E}, \quad (3.1)$$

kde  $\mathbf{X}$  reprezentuje datový soubor,  $\mathbf{T}$  je matice komponentního skóre,  $\mathbf{P}^T$  je transponovaná matice komponentních vah a  $\mathbf{E}$  je matice reziduí. Vybráním počtu komponent pak definujeme statistický model reprezentující vstupní data a matici reziduí, která vyjadřuje odchylky modelu od skutečnosti. Maximální počet hlavních komponent je roven menšímu z  $n - 1$  (o jednu méně než počtu objektů) nebo  $m$  (počet znaků souboru). Rozklad PCA je schematicky zobrazen na Obrázku 3.1, matematické odvození vah jednotlivých komponent je provedeno například v [16, 18].



Obrázek 3.1: Schematický popis rozkladu PCA na skóre a váhy jednotlivých hlavních komponent.

V podstatě existují dva způsoby výpočtu hlavních komponent. Jedním je rozklad kovarianční matice na vlastní čísla a vlastní vektory (k rozkladu se používá metoda rozkladu na singulární hodnoty (SVD z anglického „Singular Value Decomposition“)), druhým je vyjádření komponent komponentu po komponentě. V prvním případě odpovídají vlastní vektory přímo koeficientům hlavních komponent a takto získaný model neobsahuje rezidua. Následně musí být určen optimální počet hlavních komponent. Sestavení kovarianční matice velkých datových souboru může být výpočetně náročné a v konečném důsledku nechceme všechny hlavní komponenty, pouze model vystihující soubor s dostatečnou přesností. V druhém případě lze jednotlivé komponenty vyjádřit krok po kroku:

1. Začneme s datovým souborem  $\mathbf{X}$  a vypočítáme rozklad metodou SVD<sup>2</sup> na součin dvou ortogonálních matic  $\mathbf{t}_1\mathbf{p}_1^T$ . Rozdíl  $\mathbf{E}_1 = \mathbf{X} - \mathbf{t}_1\mathbf{p}_1^T$  je vlastně reziduem vyjádření  $\mathbf{X}$  a součin  $\mathbf{t}_1\mathbf{p}_1^T$  vyjadřuje první hlavní komponentu.
2. V dalším kroku pokračujeme vyjádřením druhé hlavní komponenty  $\mathbf{t}_2\mathbf{p}_2^T$  aplikací postupu z bodu 1 na matici  $\mathbf{E}_1$ , výsledkem je  $\mathbf{E}_2 = \mathbf{E}_1 - \mathbf{t}_2\mathbf{p}_2^T$ . Dále platí, že  $\mathbf{X} = \mathbf{t}_1\mathbf{p}_1^T + \mathbf{t}_2\mathbf{p}_2^T + \mathbf{E}_2$ .
3. Postup opakujeme, dokud hodnoty  $\mathbf{E}_i$  nejsou dostatečně malé.

Výše popsaný postup zaručuje automatický výběr optimálního počtu hlavních komponent.

## 3.2. Algoritmy podpurných vektorů (SVM)

Pro klasifikaci lineárně neoddělitelných dat, kde se v prostoru znaků jednotlivé třídy překrývají, lze použít metodu podpurných vektorů (SVM z anglického „Support Vector Machines“) [20]. Problematika SVM je definována jako hledání nadroviny, nejlépe dělící sadu dat na třídy. Matematický popis je podrobně popsán jinde [21, 22], my se zde omezíme na jednodušší variantu vysvětlení odvozenou v [20].

## 3.3. Lineárně oddělitelná data

Pro zjednodušení uvažujme pouze jednoduchý klasifikátor, který určuje, zda data do třídy patří či nikoli. V tom případě data vypadají jako  $N$  uspořádaných dvojic  $(\vec{x}_1, y_1), (\vec{x}_2, y_2), \dots, (\vec{x}_N, y_N)$ , kde  $\vec{x}_i \in \mathbb{R}^p$  je  $i$ -tý objekt vyjádřený  $p$  znaky a  $y_i \in \{-1, 1\}$  je binární informace o příslušnosti k dané třídě. Potom lze říct, že všechny vektory  $\vec{x}$  náležejí nadrovině platí-li pro ně:

$$f(\vec{x}) = \vec{x}^T \vec{\beta} + \beta_0 = 0, \quad (3.2)$$

kde  $\beta$  je jednotkový vektor, funkce  $f(\vec{x})$  dělí prostor na dvě třídy vyjádřené pravidlem:

$$G(\vec{x}) = \text{sign}(\vec{x}^T \vec{\beta} + \beta_0) \quad (3.3)$$

<sup>2</sup>Metoda SVD hledá rozklad matice  $\mathbf{M}$  o rozměrech  $m \times n$  na součin tří matic  $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$ , kde  $\mathbf{\Sigma}$  je diagonální matice o rozměrech  $m \times n$ , jejíž diagonála vyjadřuje singulární hodnoty,  $\mathbf{U}$  je ortogonální matice o rozměrech  $m \times m$  a  $\mathbf{V}$  je ortogonální matice o rozměrech  $n \times n$ . Matice odvozené pomocí SVD mají mnoho využití, například lze odvodit, že součin  $\mathbf{UV}^*$  tvoří nejbližší ortogonální matici k  $\mathbf{M}$ , nebo při hledání nejlepší rotace prostoru pro srovnání dvou množin souřadnic [19].



Dosazením libovolného vektoru  $\vec{x}$  do rovnice (3.2) zjistíme jeho vzdálenost k nadrovině. Uvažujeme-li oddělitelné třídy, pak lze najít funkci  $f(\vec{x})$ , pro kterou platí  $y_i f(\vec{x}_i) > 0$  pro všechny objekty  $\vec{x}_i$ , které chceme klasifikovat. Hledání funkce  $f(\vec{x})$  lze převést na problém hledání maximální hodnoty  $M$  změnou hodnot  $\vec{\beta}$  a  $\beta_0$ , za podmínky  $\|\vec{\beta}\| = 1$ , pro kterou platí:

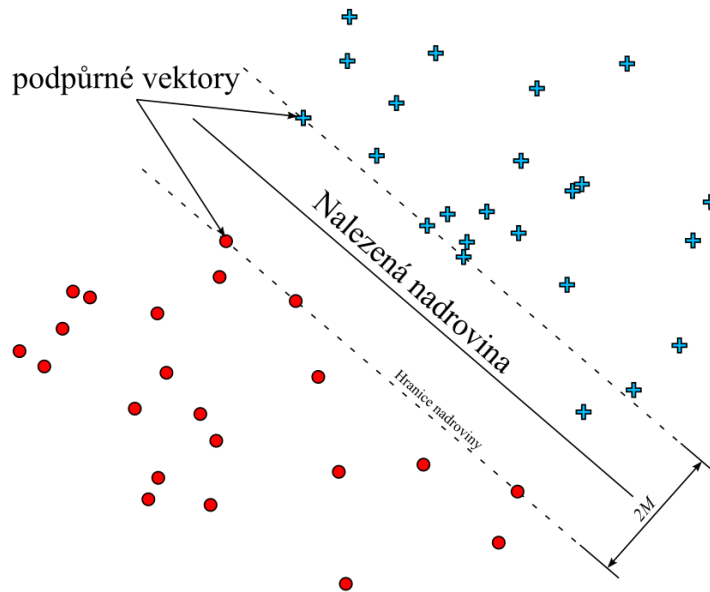
$$y_i \left( \vec{x}_i^T \vec{\beta} + \beta_0 \right) \geq M, i = 1, \dots, N. \quad (3.4)$$

Hodnota  $M$  se nazývá hranice (anglicky „margin“) a definuje minimální vzdálenost bodů  $\vec{x}_i$  od nadroviny. Maximalizaci popsanou rovnicí (3.4) lze převést na minimalizaci ve tvaru:

$$\min_{\vec{\beta}, \beta_0} \|\vec{\beta}\| \quad (3.5)$$

$$\text{s ohledem na } y_i \left( \vec{x}_i^T \vec{\beta} + \beta_0 \right) \geq 1, i = 1, \dots, N.$$

Rovnice (3.5) je obvyklý způsob zápisu oddělitelnosti dat v SVM. V [21] je pak popsáno, jak tuto rovnici řešit. Schematický popis problému je na Obrázku 3.2, pro hranici  $M$  v minimalizační úloze platí  $M = 1/\|\vec{\beta}\|$ . Nalezená nadrovina společně s hodnotou  $M$  vymezují v prostoru  $\mathbb{R}^p$  oblast o šířce  $2M$ . Natrénovaný model SVM *de facto* tvoří ty vektory  $\vec{x}_i$ , které leží na hranici vymezené oblasti, tyto vektory se nazývají „podpůrné vektory“.



Obrázek 3.2: Schématické znázornění metody SVM. Nadrovina společně s hranicí  $M$  vymezují podprostor, ve kterém se nesmí nacházet žádné objekty. K zaznamenání tohoto modelu jsou potřeba pouze souřadnice objektů na hranici, tyto se nazývají podpůrné vektory.

### 3.4. Samouspořádané mapy (SOM)

Samouspořádané mapy (SOM z anglického „Self-Organized Maps“) tvoří celou kategorii algoritmů, jejíž základ tvoří metoda odvozená v roce 1981 Tuevo Kohonenem [23, 24],

tzv. Kohonenovo mapování (anglicky „Kohonen maps“). SOM lze řadit mezi algoritmy umělých neuronových sítí (ANN z anglického „Artificial Neural Networks“), umožňujících učení bez učitele. Díky své jednoduchosti a nízkému počtu parametrů se SOM těší velké popularitě a stávají se jednou z často používaných metod pro vizualizaci a hledání hierarchie dat. Dalším výhodným aspektem SOM je, že učení může probíhat postupně po jednom objektu a šetřit tím výpočetní výkon nebo dokonce umožnit online analýzu.

### 3.5. Základní algoritmus

Vstupní data představují stejně jako v předchozích případech objekty reprezentované vektory  $\vec{x} \in \mathbb{R}^p$ , kde  $p$  je počet vlastností daného objektu. Modelem (a vlastně i výsledkem SOM) je pak uspořádaná  $k$ -tice neuronů  $\vec{m}_1, \vec{m}_2, \dots, \vec{m}_k, m_i \in \mathbb{R}^p, \forall i$ , přičemž jednotlivé souřadnice každého vektoru  $\vec{m}_i$  nazýváme váhy. Neurony jsou navíc uspořádány do dvourozměrné mřížky nejčastěji pravoúhlého nebo hexagonálního typu. Takto jsou každému neuronu  $\vec{m}_i$  přiřazeny celočíselné koordináty  $l_j \in Q_1 \times Q_2$ , kde  $Q_1 = \{1, 2, \dots, q_1\}$ ,  $Q_2 = \{1, 2, \dots, q_2\}$ , kde pro  $q_1$  a  $q_2$  platí  $K = q_1 \times q_2$ .

Parametry modelu jsou počet iteračních kroků, míra ponaučení  $\alpha$  (anglicky „learning rate“) a práh vzdálenosti (anglicky „threshold distance“). Na začátku učení jsou váhy jednotlivých neuronů iniciovány na výchozí hodnoty (buď náhodně nebo pomocí rozkladu (PCA, SVD)). V každé iteraci učení je proveden následující sled kroků:

1. Na začátku iterace je vybrán jeden objekt  $\vec{x}_i$  z datového souboru a je spočítána jeho vzdálenost ke všem neuronům v síti.
2. Neuron, k němuž je vybraný objekt nejbližší, se nazývá vítězný neuron  $\vec{m}_v$  (anglicky „winning neuron“) a jeho váhy jsou upraveny podle míry ponaučení:

$$\vec{m}_v = \vec{m}_v + \alpha (\vec{x}_i - \vec{m}_v) \quad (3.6)$$

3. Váhy neuronů sousedících s vítězným neuronem jsou upraveny podle míry ponaučení a funkce okolí  $h$ :

$$\vec{m}_j = \vec{m}_j + h (|l_j - l_v|) \alpha (\vec{x}_i - \vec{m}_j) \quad (3.7)$$

4. Míra ponaučení je proporcionálně zmenšena vzhledem k počtu zbývajících kroků.
5. Postup se opakuje od bodu 1.

Procesu učení SOM se také říká kompetitivně kooperativní učení, kdy kompetitivní část učení představuje hledání vítězného neuronu  $\vec{m}_v$  a kooperativní část představuje upravení neuronů v sousedství  $\vec{m}_v$ . Naučená síť neuronů zjednodušuje původní data a s jednotlivými neurony lze pracovat stejně jako s původními objekty. Tento přístup je aplikován v kapitole 4.4.

## 4. Vybrané aplikace automatické analýzy

### 4.1. Vliv normalizace na přesnost klasifikace

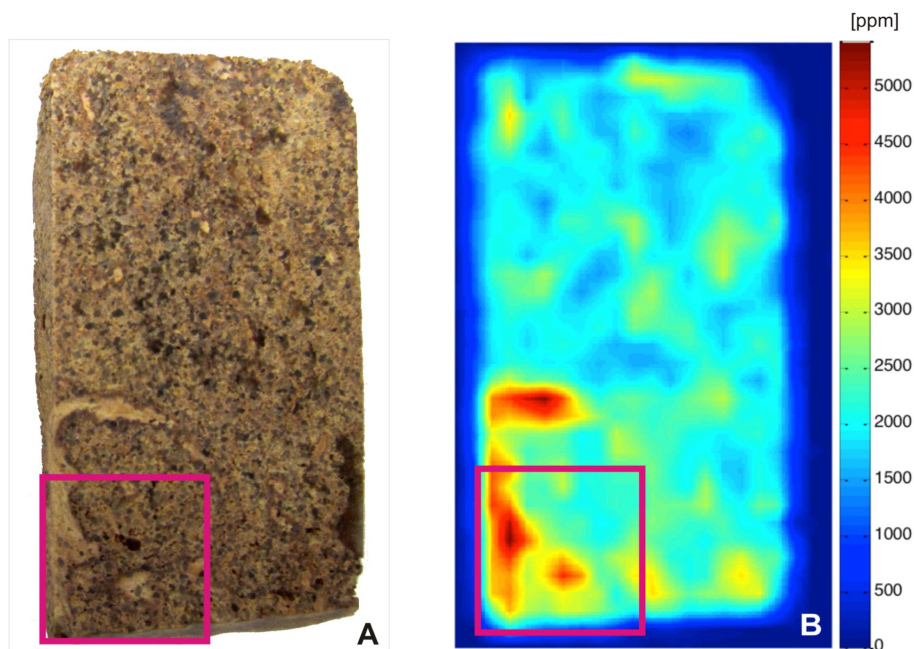
První zde popsaná práce [25] úzce související s přípravou dat před samotnou multivariační analýzou nese název „Impact of Laser-Induced Breakdown Spectroscopy data normalization on multivariate classification accuracy“. Normování dat, nebo lépe interní standardizace měření, je další z technik vedoucích ke zpřesnění a zvýšení stability měření metodou LIBS. Přehled způsobů normování signálu je dostupný například v [26], kde se zabývají nejen možnostmi interní standardizace, ale také normováním k externím signálům jako je zvuk doprovázející výstřel laseru, ablatovaná hmotnost nebo elektrický proud způsobený pohybem nabitých částic v plazmatu. Práce je motivována především absencí informace o předúpravě dat v řadě publikací zabývajících se multivariační analýzou dat z LIBS. Význam takové studie se zakládá na rostoucím počtu publikací spojujících multivariační analýzu a LIBS, nebo LIBS a chemometrii. Kromě vědecké a publikační aktivity ve zmíněné vědecké oblasti upozorňuje na problematiku popisu zpracování dat i El Haddad a kol. [27]. Akceptovaná práce pak na třech sadách materiálových standardů ukazuje vliv různých způsobů předzpracování na přesnost klasifikace jednotlivých druhů vzorků.

### 4.2. Charakterizace uranonosné rudy pomocí PCA

Druhá práce [28] se zabývá charakterizací heterogenního vzorku uranonosné rudy. Práce byla přijata na podzim 2016 a v originále se nazývá „Multivariate approach to the chemical mapping of uranium in sandstone-hosted uranium ores analyzed using double pulse Laser-Induced Breakdown Spectroscopy“. Hlavním cílem práce bylo analyzovat prostorové rozložení uranu v pískovci v rozměrném datovém souboru více než dvaceti tisíc měření pořízeném pomocí širokospektrálního Echelle analyzátoru na přibližně dvaceti pěti tisících vlnových délkách. Datový soubor pochází z rastrování části uranonosného pískovce zobrazeného na Obrázku 4.1. Analýza takového množství dat byla ztížena chemickým složením vzorku, který se skládal z těžkých prvků, jež jsou reprezentovány enormním množstvím spektrálních čar. Tabulka 4.1 uvádí koncentrace prvků naměřené z části pískovce pomocí referenční metody (hmotnostní spektroskopie indukčně vázaného plazmatu (ICP-MS)) a počty spektrálních čar těchto prvků. Navíc podle [31] lze při experimentech s výbojovou

Tabulka 4.1: Koncentrace prvků (hmotnostní procenta) v uranonosném pískovci a odpovídající počet spektrálních čar (určen z počtu záznamů pro atomové a jednou ionizované spektrální čáry na rozsahu 250–900 nm v databázi Kurucz [29]). Převzato a upraveno z [30].

	Zr	Fe <sub>2</sub> O <sub>3</sub>	U	Nb	P <sub>2</sub> O <sub>5</sub>	TiO <sub>2</sub>
Koncentrace	>5,00	1,41	0,38	0,34	0,10	0,91
Počet čar	1164	1603	1151	1306	6	3418



Obrázek 4.1: **A** fotografie vzorku, **B** koncentrace uranu zjištěná referenční metodou (rentgenová fluorescence), růžový čtverec zvýrazňuje oblast vyřezanou a připravenou pro LIBS měření. Převzato z [28].

trubicí<sup>3</sup> pozorovat více než 5000 spektrálních čar uranu na rozsahu 384,8–908,4 nm. Problém je tedy následující: při desetitisících spekter, kde každé může obsahovat tisíce čar různých prvků, je obtížné provádět běžnou analýzu zejména proto, že:

1. nevíme, zda hledaný prvek nelze ve spektru nalézt nebo
2. měříme v místě, kde neprochází uranonosná žíla nebo
3. jeho koncentrace v daném bodě byla pod limity detekce.

### 4.3. Charakterizace uranonosné rudy pomocí SOM

Poznatky a analýzy z [28] jsou dále rozvedeny v práci [30] na téma „Application of self-organizing maps to the study of U-Zr-Ti-Nb distribution in sandstone-hosted uranium ores“ přijaté v březnu 2017. Charakterizace rozdělení uranu je dále rozšířena o analýzu substituentů a mineralizací uranonosného pískovce. Jak je uvedeno v diskuzi předchozí kapitoly PCA analýza umožnila „pouze“ automaticky diskriminovat měření na základě variability. Spektra s vysokými skóre první hlavní komponenty byla úspěšně korelovaná s uranonosnou žílou.

Nvazující studie se zabývá přítomností ostatních prvků. Pomocí SOM (kapitola 3.4) se snaží vysvětlit, které prvky se ve spektrech vyskytují společně (prvkové asociace), a které nikoliv (izolace prvku). Znalost asociací prvků, případně jejich záměn, je hlavním záměrem studia mineralizace hornin. Studium mineralizace hornin má zásadní význam pro těžbu. Asociace některých prvků přímo definují kvalitu a výtěžnost horniny.

<sup>3</sup>anglicky „Hollow-Cathode Lamp“

## 4.4. Klasifikace hornin z echellogramů

Poslední zde uváděná práce [32] se zabývá možnostmi klasifikace vzorků na základě echellogramů - surových dat z detektoru připevněném na Echelle spektroskopu. Práce byla přijata v dubnu 2017 a nese název „Multivariate classification of echellograms: a new perspective in Laser-Induced Breakdown Spectroscopy analysis“. Hlavní motivací práce bylo zrychlení opakovací frekvence Echelle spektrometrů při zachování přesnosti klasifikace použitím metod zpracování obrazu přímo na echellogramech. Ačkoli jsou Echelle spektrometry hojně používané, oproti ostatním typům spektrometrů mají nižší opakovací frekvenci. To je způsobeno především nutností vyčítat celý obraz z detektoru a ten následně převádět na spektrum.

O převod se stará ovládací software dodaný výrobcem, protože je potřeba znát optické vlastnosti spektrometru. Z difrakční účinnosti jednotlivých optických prvků je matematicky vypočítán ideální kompromis mezi spektrálním rozlišením a spektrální účinností. V závislosti na výrobci může tento proces trvat desítky až stovky milisekund, což je výrazným limitem pro chemické mapování s vysokou opakovací frekvencí. Možnost vyčítání jenom části obrazu by vedla k rapidnímu zrychlení měření. V ideálním případě může být tato oblast vybrána pomocí multivariačních metod zcela automaticky v závislosti na nárocích na klasifikaci, která má být dosažena.

## 5. Nový koncept automatické identifikace prvků ve spektru

Z výše uvedených informací vyplývá, že nejdůležitějším krokem automatické charakterizace je automatické přiřazení spektrálních čar jednotlivým píkům ve spektru. V rámci studia byl vypracován nový koncept, kterým lze spektrální čáry automaticky přiřadit. Tento koncept byl prezentován na konferenci EMSLIBS 2017 v italské Pise.

### 5.1. Postup automatické identifikace

Koncept automatické identifikace je založen na sérii klasifikátorů, z nichž každý vyjadřuje jednoduchou informaci o právě zpracovávaném bodu spektra. Klasifikátory je možno rozdělit na dva typy. První z nich se zabývá hledáním píků ve spektru. Jeho výstupem je seznam středních vlnových délek odpovídajících jednotlivým nalezeným píkům. Druhý typ klasifikátoru zpracovává každý pík v seznamu středních vlnových délek a vytváří rozhodovací tabulku.

Klasifikátory prvního typu musí pro každý bod ve spektru rozhodnout, do jaké míry se jedná o pík. V našem případě jsme použili dva klasifikátory:

1. Vybraný bod spektra je označen za pík, jsou-li všechny body levého a pravého okolí o velikosti  $m$  menší.
2. Vybraný bod spektra je označen za pík, je-li jeho intenzita větší než prahová hodnota  $I_{\min}$ .

Pro klasifikátory prvního typu platí, že jejich výstupem je vždy jednička nebo nula.

Klasifikátor lze použít vícekrát, pokaždé s jiným nastavením. Tím jsou započítány i píky vyhovující více pravidlům. V případě použití více klasifikátorů je za pík považován takový bod ve spektru, který vyhovuje většině z nich.

Klasifikátory druhého typu přiřazují spektrální čáry píkům pomocí rozhodovacích tabulek. Každý řádek tabulky tvoří jiný klasifikátor, sloupce pak tvoří jednotlivé prvky, které se algoritmus snaží přiřadit. Pro přiřazování spektrálních čar byly navrženy tři klasifikátory:

**Pozice čáry**, z naměřené střední vlnové délky píku je pro každý uvažovaný prvek stanovena pravděpodobnost identifikace na základě rozdílu s nejbližším tabelovaným přechodem ze spektroskopických tabulek;

**Intenzita čáry**, poměr maximální intenzity studované čáry je porovnán s ostatními nalezenými čarami přiřaditelnými uvažovanému prvku. Vyhovuje-li poměr intenzit pravděpodobnostem přechodu, jsou tyto hodnoty mezi sebou vynásobeny. Takto jsou vyloučeny identifikace, které porušují rovnováhu;

**Počet čar**, na základě nejbližší čáry uvažovaného prvku je spočítáno množství spektrálních čar s podobnými vlastnostmi (energie přechodu, stupeň ionizace). Vybrané čáry jsou pak porovnány s píky v tabulce. Čím více podobných čar je nalezeno, tím je skóre vyšší.

## 5.2. Vyhodnocování výsledků

Výsledky dosažené pomocí rozhodovacích tabulek lze vyhodnocovat různými způsoby. Jedním ze způsobů je považovat za správně přiřazený prvek ten, jehož součet klasifikačních skóre je největší. Dalším způsobem je rozhodovat na základě většiny, čili prvek, který má pro většinu klasifikátorů největší skóre, je správně přiřazen. Nejlepším způsobem je rozhodovat na základě maxima součinu jednotlivých skóre. Tento postup však vyžaduje, aby skóre byly omezeny pouze na interval  $\langle 0, 1 \rangle$ .

Způsob automatické identifikace je navržen s ohledem na známé postupy, jimiž se výzkumníci sami řídí. První krok, hledání píků, je z oboru zpracování signálu a se spektroskopii příliš nesouvisí, nicméně i do této fáze vyhodnocování lze spektroskopické poučky zapracovat. Jedná se například o návrh klasifikátorů, které budou píky hodnotit na základě šířky nebo proložení Gaussovým, Lorentzovým či Voigtovým profilem. Druhý krok, vytvoření rozhodovacích tabulek, představuje hlavní přínos metody.

Rozhodovací tabulky umožňují zpětně vystopovat, jakým způsobem se algoritmus rozhoduje a určit, které klasifikátory přispívají přesnosti vyhodnocení nejvíce. Lze navrhnout i další klasifikátory, například klasifikátor zohledňující termodynamickou teplotu bude rozhodovat na základě poměrů intenzit jednotlivých čar pro vybranou teplotu. Takový klasifikátor lze použít i vícekrát, s různým parametrem teploty, to umožňuje dosáhnout lepších výsledků, jelikož je pravděpodobné, že správná teplota a správný prvek budou mít největší skóre.

Navrhovaný koncept lze rozšířit pomocí vah. Uvažujeme-li váhy jednotlivých sloupců, zohledňujeme koncentraci nebo pravděpodobnost, s jakou se prvek ve vzorku vyskytuje. Váhy sloupců tedy umožňují vyloučit některé prvky pro identifikaci. Uvážíme-li váhy jednotlivých řádků, vyjadřujeme tím míru, do jaké jsou jednotlivé klasifikátory spolehlivé.

Celý koncept lze použít pro modelování různých situací. Pomocí správného přiřazení a učícího algoritmu je možné váhy jednotlivých klasifikátorů automaticky trénovat a tím dosáhnout přesnějších výsledků, případně vyloučit ty klasifikátory, které přidávají falešnou informaci. Kombinací sloupcových a řádkových vah lze snížit množství chybných klasifikací pro užší skupinu vzorků.

Navrhovaný koncept představuje univerzální možnost rozložení problematiky identifikace čar do jednoduchých jednotek, se kterými lze následně pracovat pomocí automatických metod. Právě tyto metody mohou rozhodovat o jednotlivých vahách a tím ladit klasifikátory na různé aplikace.

## 6. Závěr

Předkládaná dizertační práce se zabývá možnostmi automatické charakterizace vzorků na základě měření spektroskopií laserem buzeného plazmatu.

Témata jednotlivých kapitol jsou volena s ohledem na propojení dvou vědních oborů, a sice matematické analýzy a laserové spektroskopie. Úvod do principů vzniku laserem buzeného plazmatu, jeho záření a snímání spektra přináší podrobnější vhled těm, jež se zabývají čistě statistickým zpracováním dat. Mluvíme-li o automatické charakterizaci, pak můžeme říci, že všechny jevy, které záření plazmatu ovlivňují, jsou potenciálním faktorem, který k automatické charakterizaci slouží. Teoretický popis jevů ukazuje na možná úskalí, se kterými se budou muset algoritmy automatické charakterizace vypořádat. Zmíněná úskalí pramení zejména z fyzikální a stochastické podstaty vzniku laserem buzeného plazmatu. Podrobné znalosti konstrukce detektorů a uspořádání experimentu mohou vést k novým přístupům v oblasti zpracování neupravených dat.

Popis principů multivariační analýzy usnadňuje fyzikům a spektroskopikům vstup do oboru multivariační analýzy. Vybrané metody jsou popsány nejen s důrazem na jejich použití ve spektroskopii, ale také s ohledem na autorem již publikované práce.

Na základě znalostí a zkušeností získaných během studia byl odvozen nový koncept popisu spektrálních dat. Tento koncept reaguje na potřebu automatické identifikace spektrálních čar ve spektrech LIBS. Automatická identifikace čar dosud nemá pevný rámec metod ani užívané paradigma, které by naznačovaly jednoznačný způsob přiřazení té či oné spektrální čáry vybranému píku ve spektru. Popsaný koncept se snaží vystihnout empirii identifikace, se kterou se autor během práce setkal, a položit tak základ systematickému zpracování spekter. Navrhovaný koncept lze rozšířit nejen o další empirické metody, ale i o možnosti automatizace přiřazování pomocí neuronových sítí nebo genetických algoritmů. Za hlavní výhodu nového přístupu lze považovat jednoduchost vyjadřování nových metod a možnost zpětně trasovat rozhodování automatické identifikace v případě shody nebo nejednoznačnosti přiřazení.



## Použitá literatura

- [1] CREMERS, David A.; RADZIEMSKI, Leon J. *Handbook of Laser-Induced Breakdown Spectroscopy*. 1. vyd. Chichester: John Wiley, 2006. ISBN 978-047-0093-009 (cit. na s. 8).
- [2] CIUCCI, A.; CORSI, M.; PALLESCHI, V.; RASTELLI, S.; SALVETTI, A.; TOGNONI, E. New Procedure for Quantitative Elemental Analysis by Laser-Induced Plasma Spectroscopy. *Applied Spectroscopy*. 1999, roč. 53, č. 8, s. 960–964. ISSN 00037028. Dostupné z DOI: 10.1366/0003702991947612 (cit. na s. 9).
- [3] GAUDIUSO, R.; DELL’AGLIO, M.; PASCALE, O. De; SANTAGATA, A.; GIACOMO, A. De. Laser-induced plasma analysis of copper alloys based on Local Thermodynamic Equilibrium: An alternative approach to plasma temperature determination and archeometric applications. *Spectrochimica Acta Part B: Atomic Spectroscopy*. 2012, roč. 74-75, s. 38–45. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2012.06.034 (cit. na s. 9).
- [4] CAVALCANTI, G.H.; TEIXEIRA, D.V.; LEGNAIOLI, S.; LORENZETTI, G.; PARDINI, L.; PALLESCHI, V. One-point calibration for calibration-free laser-induced breakdown spectroscopy quantitative analysis. *Spectrochimica Acta Part B: Atomic Spectroscopy*. 2013, roč. 87, s. 51–56. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2013.05.016 (cit. na s. 10).
- [5] ARAGÓN, C.; AGUILERA, J.A. CSigma graphs: A new approach for plasma characterization in laser-induced breakdown spectroscopy. *Journal of Quantitative Spectroscopy and Radiative Transfer*. 2014, roč. 149, s. 90–102. ISSN 00224073. Dostupné z DOI: 10.1016/j.jqsrt.2014.07.026 (cit. na s. 10).
- [6] TOGNONI, E.; CRISTOFORETTI, G.; LEGNAIOLI, S.; PALLESCHI, V. Calibration-Free Laser-Induced Breakdown Spectroscopy: State of the art. *Spectrochimica Acta Part B: Atomic Spectroscopy*. 2010, roč. 65, č. 1, s. 1–14. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2009.11.006 (cit. na s. 11).
- [7] GRIFONI, E.; LEGNAIOLI, S.; LORENZETTI, G.; PAGNOTTA, S.; POGGIALINI, F.; PALLESCHI, V. From Calibration-Free to Fundamental Parameters Analysis: A comparison of three recently proposed approaches. *Spectrochimica Acta Part B: Atomic Spectroscopy*. 2016, roč. 124, s. 40–46. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2016.08.022 (cit. na s. 11).
- [8] GORNUSHKIN, I. B.; SMITH, B. W.; NASAJPOUR, H.; WINEFORDNER, J. D. Identification of Solid Materials by Correlation Analysis Using a Microscopic Laser-Induced Plasma Spectrometer. *Analytical Chemistry*. 1999, roč. 71, č. 22, s. 5157–5164. Dostupné z DOI: 10.1021/ac9905524 (cit. na s. 12, 13).
- [9] GORNUSHKIN, I.B.; PANNE, U.; WINEFORDNER, J.D. Linear correlation for identification of materials by laser induced breakdown spectroscopy: Improvement via spectral filtering and masking. *Spectrochimica Acta Part B: Atomic Spectroscopy*. 2009, roč. 64, č. 10, s. 1040–1047. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2009.07.038 (cit. na s. 12).

- [10] MATEO, M.P.; NICOLÁS, G.; PIÑÓN, V.; ALVAREZ, J.C.; RAMIL, A.; YÁÑEZ, A. Versatile software for semiautomatic analysis and processing of laser-induced plasma spectra. *Spectrochimica Acta Part B: Atomic Spectroscopy*. 2005, roč. 60, č. 7-8, s. 1202–1210. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2005.05.010 (cit. na s. 12).
- [11] TORRES, E. L.; FUENTES, M. V.; GREAVES, E. D. SAX, software for the analysis of x-ray fluorescence spectra. *X-Ray Spectrometry*. 1998, roč. 27, č. 3, s. 161–165. ISSN 1097-4539. Dostupné z DOI: 10.1002/(SICI)1097-4539(199805/06)27:3<161::AID-XRS260>3.0.CO;2-D (cit. na s. 12).
- [12] FERRERO, A.; LUCENA, P.; HERRERA, R.G.; DOÑA, A.; FERNÁNDEZ-REYES, R.; LASERNA, J.J. Libraries for spectrum identification: Method of normalized coordinates versus linear correlation. *Spectrochimica Acta Part B: Atomic Spectroscopy*. 2008, roč. 63, č. 3, s. 383–388. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2007.11.040 (cit. na s. 13).
- [13] AMATO, G.; CRISTOFORETTI, G.; LEGNAIOLI, S.; LORENZETTI, G.; PALLESCI, V.; SORRENTINO, F.; TOGNONI, E. 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*. 2010, roč. 65, č. 8, s. 664–670. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2010.04.019 (cit. na s. 13).
- [14] LABUTIN, Timur A.; ZAYTSEV, Sergey M.; POPOV, Andrey M. Automatic Identification of Emission Lines in Laser-Induced Plasma by Correlation of Model and Experimental Spectra. *Analytical Chemistry*. 2013, roč. 85, č. 4, s. 1985–1990. ISSN 00032700. Dostupné z DOI: 10.1021/ac303270q (cit. na s. 14).
- [15] HÜBERT, Waldemar; ANKERHOLD, Georg. Elemental misinterpretation in automated analysis of LIBS spectra. *Analytical and Bioanalytical Chemistry*. 2011, roč. 400, č. 10, s. 3273–3278. ISSN 16182642. Dostupné z DOI: 10.1007/s00216-011-4793-x (cit. na s. 14).
- [16] MELOUN, Milan; MILITKÝ, Jiří. *Statistická analýza experimentálních dat*. Vyd. 2., upr. a rozš. Praha: Academia, 2004. ISBN 8020012540 (cit. na s. 15).
- [17] *Multi- and megavariable data analysis: basic principles and applications*. 3rd, rev. ed. Malmö: Umetrics Academy, 2013. ISBN 9789197373050 (cit. na s. 15).
- [18] MELOUN, Milan; MILITKÝ, Jiří. *Kompendium statistického zpracování dat: metody a řešené úlohy*. Vyd. 2., přeprac. a rozš. Praha: Academia, 2006. ISBN 80-200-1396-2 (cit. na s. 15).
- [19] KABSCH, W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A*. 1976, roč. 32, č. 5, s. 922–923. ISSN 05677394. Dostupné z DOI: 10.1107/S0567739476001873 (cit. na s. 16).
- [20] HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, J. H. *The elements of statistical learning: data mining, inference, and prediction*. 2. vyd. New York: Springer, 2003. ISBN 0387952845 (cit. na s. 16).

- [21] BURGESS, Christopher J.C. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*. 1998, roč. 2, č. 2, s. 121–167. ISSN 1573-756X. Dostupné z DOI: 10.1023/A:1009715923555 (cit. na s. 16, 17).
- [22] CRISTIANINI, Nello.; SHAWE-TAYLOR, John. *An introduction to support vector machines: and other kernel-based learning methods*. 1. vyd. New York: Cambridge University Press, 2000. ISBN 0521780195 (cit. na s. 16).
- [23] KOHONEN, Teuvo. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*. 1982, roč. 43, č. 1, s. 59–69. ISSN 03401200. Dostupné z DOI: 10.1007/BF00337288 (cit. na s. 17).
- [24] *Self-organizing neural networks: recent advances and applications*. Softcover reprint of the original 1st ed. 2002. S.l.: Physica Springer, 2014. ISBN 9783662003435 (cit. na s. 17).
- [25] POŘÍZKA, P.; KLUS, J.; HRDLIČKA, A.; VRÁBEL, J.; ŠKARKOVÁ, P.; PROCHAZKA, D.; NOVOTNÝ, J.; NOVOTNÝ, K.; KAISER, J. Impact of Laser-Induced Breakdown Spectroscopy data normalization on multivariate classification accuracy. *J. Anal. At. Spectrom.* 2017, roč. 32, č. 2, s. 277–288. ISSN 02679477. Dostupné z DOI: 10.1039/C6JA00322B (cit. na s. 19, 30).
- [26] ZOROV, Nikita B.; GORBATENKO, Alexander A.; LABUTIN, Timur A.; POPOV, Andrey M. A review of normalization techniques in analytical atomic spectrometry with laser sampling: From single to multivariate correction. *Spectrochimica Acta Part B: Atomic Spectroscopy*. 2010, roč. 65, č. 8, s. 642–657. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2010.04.009 (cit. na s. 19).
- [27] HADDAD, J. El; CANIONI, L.; BOUSQUET, B. Good practices in LIBS analysis: Review and advices. *Spectrochimica Acta Part B: Atomic Spectroscopy*. 2014, roč. 101, s. 171–182. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2014.08.039 (cit. na s. 19).
- [28] KLUS, Jakub et al. 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*. 2016, roč. 123, s. 143–149. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2016.08.014 (cit. na s. 19, 20, 30).
- [29] KURUCZ, R.L.; BELL, B. *Atomic Line Data Kurucz CD-ROM No. 23*. Cambridge, Mass.: Smithsonian Astrophysical Observatory, 1995 (cit. na s. 19).
- [30] KLUS, Jakub; POŘÍZKA, Pavel; PROCHAZKA, David; MIKYSEK, Petr; NOVOTNÝ, Jan; NOVOTNÝ, Karel; SLOBODNÍK, Marek; KAISER, Jozef. 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*. 2017, roč. 131, s. 66–73. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2017.03.008 (cit. na s. 19, 20).

- [31] CHINNI, Rosemarie C.; CREMERS, David A.; RADZIEMSKI, Leon J.; BOSTIAN, Melissa; NAVARRO-NORTHRUP, Claudia. Detection of Uranium Using Laser-Induced Breakdown Spectroscopy. *Applied Spectroscopy*. 2009, roč. 63, č. 11, s. 1238–1250. ISSN 00037028. Dostupné z DOI: 10.1366/000370209789806867 (cit. na s. 19).
- [32] POŘÍZKA, Pavel et al. Multivariate classification of echellograms: a new perspective in Laser-Induced Breakdown Spectroscopy analysis. *Scientific Reports*. 2017, roč. 7, č. 1, s. 1–7. ISSN 20452322. Dostupné z DOI: 10.1038/s41598-017-03426-0 (cit. na s. 21, 30).
- [33] KLUS, Jakub; POŘÍZKA, Pavel; PROCHAZKA, David; MIKYSEK, Petr; NOVOTNÝ, Jan; NOVOTNÝ, Karel; SLOBODNÍK, Marek; KAISER, Jozef. 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*. 2017, roč. 131, s. 66–73. ISSN 0584-8547. Dostupné z DOI: 10.1016/j.sab.2017.03.008 (cit. na s. 30).
- [34] KLUS, Jakub; POŘÍZKA, Pavel; PROCHAZKA, David; NOVOTNÝ, Jan; NOVOTNÝ, Karel; KAISER, Jozef. Effect of experimental parameters and resulting analytical signal statistics in laser-induced breakdown spectroscopy. *Spectrochimica Acta Part B: Atomic Spectroscopy*. 2016, roč. 126, s. 6–10. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2016.10.002 (cit. na s. 30).

# Curriculum Vitae

JAKUB KLUS

## OSOBNÍ INFORMACE

*Narozen 30. ledna 1989 v Novém Jičíně*

*E-mail* jakub.klus@gmail.com

## VZDĚLÁNÍ

*2013–2017* Fyzikální a materiálové inženýrství (Ph.D.)

Vysoké učení technické v Brně, Fakulta strojního inženýrství, Technická 2896/2, 616 69 Brno.  
Téma dizertační práce: Vývoj algoritmu pro automatickou charakterizaci vzorku na základě dat získaných spektroskopii laserem indukovaného plazmatu (LIBS).  
Školitel: prof. Ing. Jozef KAISER, Ph.D. (Předpokládaná doba ukončení: 2017)

*2011–2013* Fyzikální inženýrství (Ing.)

Vysoké učení technické v Brně, Fakulta strojního inženýrství, Technická 2896/2, 616 69 Brno.  
Téma diplomové práce: Optimalizace procedury měření optického záření rozptýleného pevnými tělesy prováděného scattermeterem.  
Školitel: prof. RNDr. Miloslav OHLÍDAL, CSc.

*2008–2011* Fyzikální inženýrství (Bc.)

Vysoké učení technické v Brně, Fakulta strojního inženýrství, Technická 2896/2, 616 69 Brno.  
Téma bakalářské práce: Kalibrace únavového stroje MZGS-200.  
Školitel: Ing. Karel SLÁMEČKA, Ph.D.

*2004–2008* Gymnázium Mikuláše Koperníka

Gymnázium Mikuláše Koperníka 17. listopadu 526, 743 01 Bílovec.  
Zaměření: matematika.

## PŘEDCHOZÍ ZAMĚSTNÁNÍ

*2014-dosud* Technický pracovník, CEITEC

Středoevropský technologický institut, zpracování spekter, vývoj zakázkového software, spolupráce na vývoji zakázkové elektroniky pro vědecká zařízení.

*2012-2014* Technický pracovník, VUT

Vysoké učení technické v Brně, zpracování dat, práce na zobrazovacím spektrofotometru, vývoj elektroniky a řídicího software pro Scattermeter.

*2010* Technický pracovník, VUT

Vysoké učení technické v Brně, implementace experimentu nabíjení a vybíjení kondenzátoru.

2009 Programátor, Amenit s.r.o  
Vývoj zakázkového software pro meteorologické stanice.

#### STÁŽE

2011 Fraunhoferův institut jemné mechaniky a aplikované optiky,  
Německo  
Téma stáže: Návrh širokopásmového detektoru pro *in-situ* optickou analýzu tenkých vrstev.

#### ÚČAST NA PROJEKTECH

2016 Fond Vědy  
Název projektu: Klasifikace hutních materiálů pomocí spektrometrie laserem indukovaného plazmatu za použití hlubokých neuronových sítí

2014 IMPI  
Název projektu: Zařízení pro vyčítání intenzity elektromagnetického záření

#### NEJVÝZNAMNĚJŠÍ PUBLIKACE

Pořízka, Pavel et al. Multivariate classification of echellograms: a new perspective in Laser-Induced Breakdown Spectroscopy analysis. *Scientific Reports*. 2017, roč. 7, č. 1, s. 1–7. ISSN 20452322. Dostupné z DOI: 10.1038/s41598-017-03426-0

Klus, Jakub; Pořízka, Pavel; Prochazka, David; Mikysek, Petr; Novotný, Jan; Novotný, Karel; Slobodník, Marek; Kaiser, Jozef. 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*. 2017, roč. 131, s. 66–73. ISSN 0584-8547. Dostupné z DOI: 10.1016/j.sab.2017.03.008

Klus, Jakub et al. 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*. 2016, roč. 123, s. 143–149. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2016.08.014

Klus, Jakub; Pořízka, Pavel; Prochazka, David; Novotný, Jan; Novotný, Karel; Kaiser, Jozef. Effect of experimental parameters and resulting analytical signal statistics in laser-induced breakdown spectroscopy. *Spectrochimica Acta Part B: Atomic Spectroscopy*. 2016, roč. 126, s. 6–10. ISSN 05848547. Dostupné z DOI: 10.1016/j.sab.2016.10.002

Pořízka, P.; Klus, J.; Hrdlička, A.; Vrábek, J.; Škarková, P.; Prochazka, D.; Novotný, J.; Novotný, K.; Kaiser, J. Impact of Laser-Induced Breakdown Spectroscopy data normalization on multivariate classification accuracy. *J. Anal. At. Spectrom.* 2017, roč. 32, č. 2, s. 277–288. ISSN 02679477. Dostupné z DOI: 10.1039/C6JA00322B

# Abstrakt

Předkládaná práce se zabývá teoretickými a praktickými předpoklady pro automatickou charakterizaci vzorků pomocí spektroskopie laserem indukovaného plazmatu (LIBS). V práci jsou popsány základní aspekty interakce laseru a látky, rozpínání plazmatu a podstata jeho záření. Popis záření plazmatu je dán do kontextu detekce spektra elektromagnetického záření a statistických veličin popisujících plazma. Princip automatické charakterizace se zakládá na znalostech multivariační statistické analýzy, která, v souladu se současnými trendy, představuje stěžejní směr automatické analýzy spekter v LIBS. Náznačené úvahy jsou zúročeny na šesti aplikacích, které jsou reprezentovány komentáři k přijatým vědeckým publikacím.

## **Abstract**

Submitted work concerns with the theoretical and practical requirements for an automatic characterization of samples by means of Laser-Induced Breakdown Spectroscopy (LIBS). Theoretical aspects of laser-matter interaction, plasma expansion, and plasma emission are described theoretically within this work. The description of the plasma emission is enhanced with the spectral detection systems and statistical properties of the plasma. The principle of the automatic characterization is based on the multivariate data analysis theoretical background, which presents recent trend and fundamental approach for automatic spectra analysis in LIBS. Theoretical knowledge is manifested in six applications, which are presented as a comment to published manuscripts. These publications push the frontiers of automatic spectra processing in LIBS.