Journal Pre-proof

Quantification of alloying elements in steel targets: The LIBS 2022 regression contest

Erik Képeš, Jakub Vrábel, Panagiotis Siozos, Victor Pinon, Pavlos Pavlidis, Demetrios Anglos, Tong Chen, Lanxiang Sun, Guanghui Lu, Dillam Jossue Díaz-Romero, Simon Van den Eynde, Isiah Zaplana, Jef Peeters, Václav Kaňa, Antonín Záděra, Vincenzo Palleschi, Alessandro De Giacomo, Pavel Pořízka, Jozef Kaiser

PII: S0584-8547(23)00097-6

DOI: https://doi.org/10.1016/j.sab.2023.106710

Reference: SAB 106710

To appear in: *Spectrochimica Acta Part B: Atomic Spectroscopy*

Received date: 16 January 2023

Revised date: 12 May 2023

Accepted date: 19 May 2023

Please cite this article as: E. Képeš, J. Vrábel, P. Siozos, et al., Quantification of alloying elements in steel targets: The LIBS 2022 regression contest, *Spectrochimica Acta Part B: Atomic Spectroscopy* (2023), https://doi.org/10.1016/j.sab.2023.106710

Selected Papers from the XII International Conference on Laser Induced Breakdown Spectroscopy (5-9 September 2022, Bari, Italy

Quantification of alloying elements in steel targets: The LIBS 2022 regression contest

Erik Képeš[1,2,], Jakub Vrábel[1,2], Panagiotis Siozos[3,*], Victor Pinon[3], Pavlos Pavlidis[4], Demetrios Anglos[3,5], Tong Chen[6,7,8,9], Lanxiang Sun[6,7,8,*], Guanghui Lu[6,7,8,9], Dillam Jossue Díaz-Romero[10,11,12,*], Simon Van den Eynde[10,11], Isiah Zaplana[13], Jef Peeters[10,11], Václav Kaňa[14], Antonín Záděra[14], Vincenzo Palleschi[15], Alessandro De Giacomo[16,17], Pavel Pořízka[1,2,18], Jozef Kaiser[1,2,18]*

[1]Central European Institute of Technology, Brno University of Technology, Purkyňova 656/123, CZ 612 00, Brno, Czech Republic

[2]Institute of Physical Engineering, Faculty of Mechanical Engineering, Brno University of Technology, Technická 2, CZ 616 69, Brno, Czech Republic.

[3]Institute of Electronic Structure and Laser, Foundation for Research and Technology-Hellas (IESL-FORTH), P.O. Box 1385, GR 711 10, Heraklion, Crete, Greece

[4]Institute of Computer Sciences, Foundation for Research and Technology Hellas, GR 711 10, Heraklion, Crete, Greece

[5]Department of Chemistry, University of Crete, P.O. Box 2208, GR 710 03, Heraklion, Crete, Greece

[6]State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China

[7]Key Laboratory of Networked Control Systems, Chinese Academy of Sciences, Shenyang 110016, China

[8]Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, China

[9]University of Chinese Academy of Sciences, Beijing 100049, China

[10]Department of Mechanical Engineering, KU Leuven, Celestijnenlaan 300, Leuven 3000, Belgium

[11]Core Lab Value, Cost & Circularity in Manufacturing, Flanders Make, Leuven 3000, Belgium

[12]PSI-EAVISE, KU Leuven, Jan Pieter de Nayerlaan 5, Sint-Katelijne-Waver 2860, Belgium

[13]Institut d'Organització i Control de Sistemes Industrials Universitat Politècnica de Catalunya · BarcelonaTech (UPC) Diagonal, 647, planta 11 (Edifici ETSEIB) 08028 Barcelona, Spain

[14]Institute of Manufacturing Technology, Faculty of Mechanical Engineering, Brno University of Technology, Technická 2, CZ61669 Brno, Czech Republic

[15]Applied and Laser Spectroscopy Laboratory, ICCOM-CNR, 56124 Pisa, Italy

[16]Department of Chemistry, University of Bari, Via Orabona 4, 70126 Bari, Italy

[17]CNR-NANOTEC, Istituto di Nanotecnologia, Via Amendola 122/d, 70126 Bari, Italy

[18]Lightigo Space s.r.o., Renneská třída 329/13, CZ63900 Brno, Czech Republic

[*]Corresponding authors: kepes@vutbr.cz, psiozos@iesl.forth.gr, sunlanxiang@sia.cn, dillam.diaz@kuleuven.be

Abstract

We present the results of the regression contest organized for the LIBS 2022 conference. While the motivation and design of the contest are briefly presented, the work focuses on the methodologies of the three best-performing teams. The employed spectral preprocessing strategies, choice of regression models and its optimization are detailed for each team separately. The aim of the contest reflects the long-term challenges faced by quantitative laser-induced breakdown spectroscopy (LIBS) analysis. Thus, the contest was designed with the purpose of providing a transparent platform for comparing and evaluating the large range of data processing tools available in the LIBS literature. Namely, the contest consisted of the quantification of two major (Cr, Ni) and two minor (Mn, Mo) elements in 15 steel targets. For constructing an appropriate regression model, spectra of 42 targets were provided. The spectra were collected using a commercially available laboratory-based LIBS system and made publicly available. The contest lasted 53 days during which the teams did not receive feedback. In total, 21 teams participated out of which the three best-performing methodologies are presented here. A single linear partial least squares model and two artificial neural network regression models are presented. The corresponding feature selection strategies included emission line selection, spectral range selection, and automatized wavelength selection. Various spectral normalization strategies and data augmentation strategies are also presented.

Keywords: laser-induced breakdown spectroscopy, quantification of alloying elements, regression analysis, artificial neural networks, partial least squares regression, feature selection

1. Introduction

Laser-induced breakdown spectroscopy (LIBS) [1,2] is a prominent elemental spectroscopic technique with rapid sensing capabilities and versatile instrumentation. As such, LIBS boasts a rich portfolio of applications, e.g., in the industrial analysis of metals [3], mining [4], forensic analysis [5], medical [6,7], analysis of geological materials [8], and space exploration [9–11]. In most of these applications, LIBS is used in a qualitative manner, i.e., material classification. Meanwhile, quantitative analysis is often regarded as the Achilles's heel of LIBS. In fact, LIBS, at best, offers semi-quantitative [12] results. This is, in most practical applications, done based on a regression model—either univariate or multivariate [13,14].

At a glance, quantitative analysis using LIBS is straightforward. In an ideal case, increasing the analyte concentration corresponds to a linear increase of its emission signal in the recorded spectra [15,16]. However, altering the material composition (i.e., changes in the analytes' concentrations) generally corresponds to changes in the material properties as well. Consequently, the laser-matter interactions and the resulting plasma formation and emission are also affected [17]. As such, the stoichiometric assumptions are often perturbed [18,19]. Changing plasma properties (e.g., the presence of LTE, temperature, electron number density, homogeneity) introduce additional deviation from the optimal linear signal response [20–22]. Moreover, spectral interference in emission spectra rich in emission lines presents further challenges.

These limitations are commonly addressed by various spectral normalization strategies, e.g., the application of an internal standard [23]. More advanced techniques attempt the normalization of spectra to common temperature and electron number density [24]. These approaches generally require considerable expertise, effort, and the optimization of the detection parameters. Consequently, multivariate approaches have been increasingly adopted with the aim of finding correlations between spectral data and the analyte concentration [25,26]. These approaches range from linear partial least squares regression [27–29] and support vector machines (with a linear kernel) [30,31], decision trees [32,33] and random forests [34], to deep learning [35–38]. There are countless models to choose from as well as countless pre-processing strategies to be paired with. Hence, the various methodologies of quantitative LIBS should be frequently systematically reviewed and compared. The latter falls within the scope of the present work which presents details of such an effort.

Similar efforts have been made in the past. Namely, two public contests addressing the quantification of the elemental composition of steel alloys by LIBS were organized by Dr. Igor Gornushkin and BAM (Berlin, GE). As part of these two contests, steel targets were distributed among the contesting laboratories that carried out their measurements data analysis. The results of these contests were presented at conferences but yielded no publications. A different approach was taken by the French LIBS community which distributed only spectra among participating laboratories [39]. This approach (of sharing only data) has been repeated in the context of classification as part of a contest organized for the EMSLIBS 2019 conference [40,41]. More recently, another regression contest was launched by BAM (Berlin, Ge), which addressed the quantification of chlorine in concrete targets [42].

In this work, we present the details of a regression contest organized for the LIBS 2022 conference together with the methodologies used by the three top performing teams. We strongly believe that the presented results will aid the choice of regression methodologies applied to LIBS data. In addition, presenting the successes and limitations of each method, we hope to inspire new research directions addressing the analytical performance of LIBS. The next section presents the design of the contest, including the target preparation, data collection, data preparation, and evaluation. Subsequently, Section 3 presents the methodologies of the teams achieving the top three ranks. Then, the methodologies and results are discussed.

2. Contest description

The aim of the contest was to compare the various calibration-based quantification methodologies available in the literature and to inspire the development of novel regression strategies, including the implementation of state-of-the-art data modeling techniques devised in other fields.

2.1. Motivation

A commonly considered limitation of LIBS is its relatively poor quantitative performance compared to other spectroscopic techniques [43]. Thus, this regression analysis contest was motivated by a desire to improve the accuracy and predictive power of statistical models applied to LIBS data. As regression analysis is a statistical technique that models the relationship between a dependent variable and one or more independent variables, we hoped to not only compare different regression models, but also different approaches to feature selection and engineering strategies. By holding this contest, we hoped to encourage the development of new and innovative approaches to regression analysis, as well as compare and evaluate different methods in a systematic way. The goal of the contest was to identify the best-performing methodologies. Lastly, the contest was also aimed to serve as a learning opportunity for participants, as it allowed them to hone their skills in statistical modeling and gain experience working with LIBS data that might differ from what they can usually access.

2.2. Design of the contest

The contest played out in a straightforward fashion. LIBS data were collected and made public by the Group of Laser Spectroscopy of CEITEC BUT. The online data repository Figshare was used to host the datasets [44–47]. The contest was announced through the conference website and mailing list. To motivate the application of multivariate techniques, the initially published datasets (training and testing) contained spectra normalized to unit maxima. This proved to be an oversight, as it made univariate regression analysis impossible. As such, 2 weeks after the contest's launch, raw datasets were also published. The deadline of the submissions was postponed accordingly.

The predicted concentration values (1 value per each of the 4 chosen alloying elements Mn, Mo, Ni, and Cr) were submitted for each of the 15 test targets. The submission was made possible in various forms to facilitate the participation. No additional information was collected regarding the applied methodology at this stage. This turned out to be another potential oversight as the contest was concluded and the final rankings were announced, the participants had little incentive to provide the requested information. Consequently, we have limited information about the methodologies used by the teams not reaching the top three spots.

2.3. Targets and data

Target preparation

The dataset was collected from steel targets which were designed and created for the regression contest. The target materials were prepared by melting in an electric induction furnace. The composition of the materials was adjusted by alloying with pure metals or ferroalloys. After melting, the melt was homogenized by stirring it in a magnetic field (using the Lorentz force) for a minimum of 5 minutes at a temperature of 1550 to 1600 °C. Subsequently, the melt was cast into metal molds. The sampling was carried out in such a way as not to affect the concentration of the basic metal elements according to standards commonly used in metallurgical plants. Targets for chemical composition analysis were designed so that the chemical composition of the steels ranged from non-alloyed to low alloyed to high alloyed.

Reference measurement

The certified concentration values of the created targets were determined using spark-discharge optical emission spectrometry on a commercial Q4 Tasman (Bruker, GE) instrument. The instrument was calibrated using a reference calibration set prior to the analysis of the target set. The analysis was performed under Argon 5.0 atmosphere. The chemical homogeneity of the measured targets was ensured both by the preparation of the melt itself (mixing) and by the cooling and solidification conditions of the targets in the metal mold of the sampler.

LIBS spectra acquisition

Each target's surface was homogenized using 800 grid sandpaper and cleaned with isopropyl alcohol using paper towels. The targets were sampled at 200 distinct spots in single-shot mode, collecting a single spectrum from each individual spot. The spectra were then outlier filtered using principal component analysis (PCA) [48,49]. Namely, PCA was applied to the spectra obtained from each target individually and the distance of the spectra's PCA scores from their mean was evaluated. Spectra with a larger-than-average distance from the center were removed [50]. The remaining spectra were randomly sampled. Consequently, there was a total of 50 spectra made available for each target in both the training and testing datasets. A 1064 nm Nd:YAG laser (Litron TRLi D-DP, Litron, UK) was used for ablation with a pulse energy of 95 mJ, 10 ns pulse width. The laser pulse was focused into a spot with a diameter of 0.2 mm. The collected emission was resolved with an echelle spectrograph (EMU 65, Catalina Scientific, US; resolving power 6000) in the 240—1000 nm spectral range (40002 resolved wavelength values). The resolved emission was recorded using an EMCCD camera (Falcon Blue, Raptor Photonic, IR) with a 1.5 μs delay and 50 μs gate width (exposition time). The spectra were not intensity calibrated. Thus, in combination with the long gate width, calibration-free approaches were discouraged.

Datasets

Two training and two testing datasets were made available. The contest was launched with spectra normalized to unit maxima [44,45]. This normalization was used to motivate the application of multivariate techniques. On the contrary, normalization made the use of univariate methodologies (and the necessary normalization techniques) impossible to apply. As such, after a two-weeks delay, raw spectra were also made available [46,47]. Both training datasets contained 2100 spectra (50 spectra of 42 targets). The corresponding concentration values (Figure 1, Table S1 (Appendix A)) were appended to the intensity values as additional columns. The test datasets contained 750 spectra (50 spectra of 15 targets). The test datasets contained no concentration values.

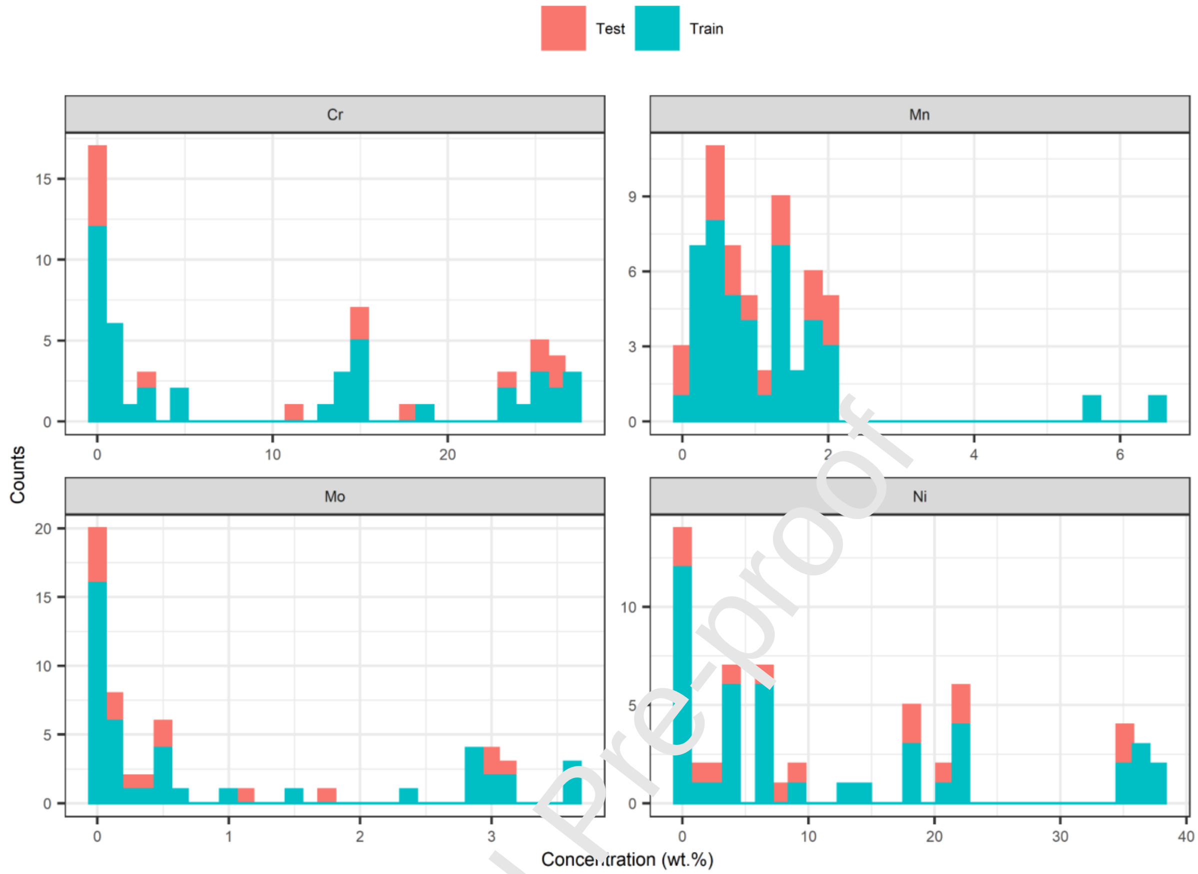


Figure 1 Histograms representing the concentration distributions in the training and test datasets for the four alloying elements of interest.

2.4. Rules

The contest lasted 53 days. During this time, the contestants were allowed to submit their predictions repeatedly. However, they received no feedback during the duration of the contest and only their latest submission was evaluated.

The evaluation of the contest presented a challenge. The four alloying elements covered concentration ranges which differed by an order of magnitude. Thus, when comparing the performance of various predictions, the following questions had to be addressed:

1) How to correctly aggregate the errors across the full range of the predicted values: does a 0.1 wt.% error at a nominal concentration value of 0.2 wt.% have the same impact as the same 0.1 wt.% error at a nominal concentration of 20 wt.%?

2) How to compare the prediction errors across distinct elements? This is an especially important question if the nominal concentration ranges of the distinct elements differ widely.

Naturally, while answering these questions, one must take into account the target application of the regression models. However, this regression contest was not designed with a practical application in mind but rather to provide a platform to compare various distinct regression methodologies.

Consequently, we opted for evaluating the predictions in terms of root-mean-square error (RMSE) values:

$$\mathrm{RMSE_e} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (1)$$

where the subscript e denotes the element, $\hat{y_i}$ is the concentration predicted for target i, y_i denotes the corresponding ground truth, and N is the number of targets. In addition, the mean absolute error (MAE) was also considered, as it is less sensitive to outlier values:

$$\mathrm{MAE_e} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \quad (2)$$

Nevertheless, the final ranking was made based on the RMSE metric. The contestants' performance was evaluated separately for the individual elements. To determine the final ranking, the contestants' ranks across the 4 elements was averaged. The aim was to avoid overwhelming the final ranking with the performance on major alloying elements of Ni and Cr whose concentrations covered the 0—35 and 0—25 wt.% ranges, respectively. As reference, the concentration ranges of the minor alloying element Mn and Mo were 0—2 and 0—3.5 wt.% ranges. This is, however, far from being the only possible approach for comparing the performance of the methodologies. A more exhaustive comparison (but still incomplete) is presented in Appendix B.

3. Methodologies of the top-performing participants

The following sections describe the methodologies devised by the three top-performing teams. These teams and their ranking have been determined according to the rules described above. While the unique aspects of each methodology are detailed below, a general introduction to the chosen regression models (artificial neural networks and partial least-squares regression) is not provided, as there is ample literature available for the interested reader. These are referenced in the text below.

3.1. Panagiotis Siozos, Victor Pinon, Pavlos Pavlidis, Demetrios Anglos (FORTH team)

The first place was attained by the FORTH team. Their methodology used for calculating the concentration of Ni, Cr, Mn, and Mo from the LIBS spectra of the steel targets has already been presented by their group (IESL-FORTH) for the discrimination of individuals' archeological human bones based on LIBS spectra combined with artificial neural networks (ANNs) [51]. The methodology was established using existing open-source machine-learning algorithms rather than developing or adjusting such algorithms to LIBS applications. In addition, a significant effort has been applied to the preparation and preprocessing of the data to achieve optimum results.

The open-source machine learning algorithms have the potential to be a reliable tool for LIBS applications. These algorithms have several advantages as they continue to evolve, have reduced associated cost, and exhibit high adaptability and interoperability concerning existing requirements. Therefore, using open-source machine learning software and tools can boost the application of advanced machine learning methods, particularly in the field of LIBS analysis. Several machine-learning algorithms (K-Nearest Neighbors, Random Forests, and Support Vector Machines) have been tested for bone discrimination by our team [52]. It was concluded that the ANN method provided the most accurate results.

3.1.1. Spectral pre-processing

Logit transformation of element concentration values

The ANN model may predict negative values of elemental concentration, particularly for elements with low concentration. Thus, the concentration values were transformed using the logit function to eliminate this effect by forcing the ANN model to predict only positive values. The logit function transforms a finite range of values, which in our case is from 0 to 100%, to real numbers:

$$\sigma = logit(p) = ln\frac{p}{100 - p}, \;\; p \in (0, 100), \;\; \sigma \in (-\infty, +\infty) \quad (3)$$

where p is the concentration value of the element and σ is the output of the transformation. The σ values calculated from equation (1) are used as the ANN model's output instead of the elemental concentration values. Finally, the predicted σ values are transformed back to concentration p values by the inverse transformation using the standard logistic function:

$$p = 100\frac{1}{1 + e^{-\sigma}} \quad (4)$$

Figure 2a displays the predicted values of nine targets from the dataset with known concentrations versus the actual Mo concentration percentage. The model was built without applying any transformation to the datasets and predicted values were derived by training the ANN algorithm on the dataset of the remaining 32 known targets. A wide distribution of the predicted values was observed for the lowest concentration values. Moreover, two of the predicted values were negative and thus were replaced by zero to calculate the root mean squared error (RMSE) and root mean squared relative error (RMSRE) values.

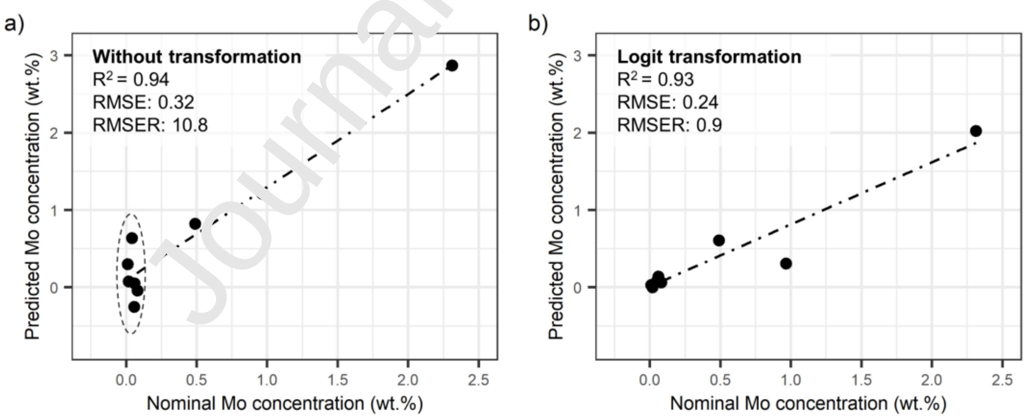


Figure 2 The predicted values of the ANN model of nine targets versus the known Mo concentration values a) without applying any transformation on the Mo concentration values and b) applying logit transformation to the Mo concentration values.

Figure 2b displays the predicted Mo concentration values of the same nine targets with the logit transformation. A significant improvement in the model's accuracy is observed using logit transformation. The distribution of the predicted values of the low concentration targets is smaller if logit transformation is applied than the distribution of the predicted values by training using the actual concentration values (without applying transformation). The R^2 values are similar in both cases. However, the RMSE values decreased from 0.34 to 0.24 and the RMRSE values significantly fell from 10.8 to 0.9 by applying the logit transformation.

Dataset normalization

Four different normalization methods were applied for standardizing the spectral data in both the training and the test sets and their effect on the robustness of the model was examined. These were: 1) normalization by the sum of the intensity values, 2) normalization by the square root of the sum of the squared intensity values, 3) normalization by the maximum intensity value and 4) normalization by the variance of the spectral part used as the input. The RMSE values from the model's predicted values of all four elements were calculated and the computer processing time required for the ANN model to converge was measured for each normalization method. All normalization methods provided comparable RMSE values. The two most accurate ones were the normalization by the sum of intensities (2) and the unit vector normalization method (3). However, there was a significant difference in the processing time with the unit vector requiring almost half of the time to converge than the sum of intensities method. Thus, we selected the unit vector method to be used for the ANN analysis.

Lastly, an attempt was made to reduce the variance of the provided spectra. Namely, the accuracy of the trained ANN models was evaluated using different datasets. The first one used the whole available dataset, i.e., 50 LIBS spectra for each target. The second approach used 25 spectra calculated by averaging 2 spectra from the initial dataset, the third used 10 spectra calculated by averaging 5 spectra, and the fourth used 5 spectra calculated by averaging 10 spectra. The ANN model built using all the whole dataset provided the least accurate concentration predictions. The accuracy of the model improved significantly by averaging two spectra. The models built from the datasets produced by averaging 5 and 10 spectra had similar accuracies and were better than those from the first two datasets. Thus, we concluded that the ANN's performance was improved by improving the signal-to-noise ratio of the dataset. Therefore, we selected the ANN model built using the 10 spectra/target (i.e., by averaging 5 of the initial spectra) to calculate the concentrations of the unknown targets.

3.1.2. Model selection and optimization

The open-access R-package "neuralnet" (version 1.44.2) was used to perform the ANN analysis [53]. R is a programming language appropriate for statistical computing and graphics that provides a wide range of statistical and graphical tools [54]. The advantage of R is its availability as free, open software under the terms of the Free Software Foundation GNU General Public License in source code form. The 'neuralnet' package was obtained from the CRAN repository and applied to develop and construct the corresponding ANN models in R (Version 4.1.3). The globally convergent algorithm of 'neuralnet' is based on resilient back-propagation without weight back-tracking [53]. This particular package was selected based on its frequent use by experts in biology and other scientific disciplines and its continued maintenance and testing.

Table 1 Atomic emission lines in the LIBS spectra used as an input in the ANN model.

Element[a]	Wavelength (nm)[b]	Element[a]	Wavelength (nm)[b]
Cr I	357.869	Ni I	336.957
	359.349		337.199
	360.533		337.422
	520.452		337.464
	520.604		337.897
	520.844		338.057
Mo I	550.649		338.085
	553.305		339.105
Mn I	403.076		339.299

	403.307	340.105
	403.449	340.958
	403.573	341.348
	403.873	341.394
	404.136	341.476
	475.404	342.371
	476.153	343.356
	476.238	343.728
	476.586	344.626
	476.643	345.289
	478.342	345.847
	482.352	346.165
Fe I	282-288 range	351.505
		352.454

[a] I: neutral atom.

[b] The emission lines and spectral ranges were identified based on the NIST atomic spectra database [55] and have been used as input to the ANN model.

The wavelengths of the LIBS spectra that correspond to emission lines of one of the four elements of interest (Cr, Mn, Mo, and Ni) were selected as the input of the ANN model. Moreover, a wavelength range corresponding to multiple emission lines of the matrix element (Fe) was also included as input. The considered emission lines are summarized in Table 1. The input neurons corresponded to the wavelength values covering the emission lines' whole profile (Figure 3). Thus, the number of input neurons was reduced from 40000 points (the number of wavelengths in the raw spectra) to approximately 1300 points (≈ 3% of the initial spectral range). Previous work of our group concluded that selecting all the data points of the peak profiles as input neurons provides the highest degree of accuracy of the model [52]. Other works reaffirmed this conclusion [56]. A single hidden layer with 10 neurons was selected. The output layer consisted of a single neuron predicting the concentration value of one of the four elements. In other words, a separate model was trained for each element.

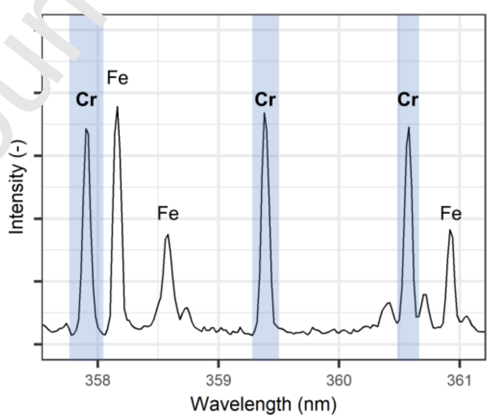


Figure 3 Segment of the average LIBS spectrum of target 22 with a known composition. The wavelength ranges corresponding to 3 Cr emission lines used as input for the ANN model for predicting the Cr concentrations are highlighted in pale blue.

The ANN model was optimized by training the model using the spectra of 32 of the total 42 known targets (training dataset). The performance of the model was evaluated by measuring the accuracy of the predicted values of the dataset of the nine remaining targets (validation dataset). The spectra of target 35 consistently provided inaccurate results and thus were excluded from both the training and validation datasets as outliers. The optimized ANN models were then used to predict the composition of the 15 unknown targets. For each element, 16 models were trained. Each model yielded concentration prediction for a single chemical element. The final concentration prediction was then calculated as the mean of the 16 predicted values. Correspondingly, the prediction error was determined as the 2σ error of the predicted values.

3.2. Tong Chen, Lanxiang Sun, Guanghui Lu (SIA team)

By observing the given element concentrations, we found that the elemental contents of the 42 different targets fluctuate greatly. Taking Cr element as an example, its content ranged from 0.05 to 27.16 wt.% in the training set, which certainly leads to the variance of matrix effect and self-absorption effect. As a result, plenty of experience and time was required to establish a univariate model with appropriate spectral line input, so we first established full-spectrum model based on partial least squares regression (PLSR). Due to the limited training set size, we evaluated the performance of the different models using leave-one-out cross-validation. With two different input forms and two different output forms (more details below), we have established 4 different models for each element.

Observing different spectra of the same target we found relatively large fluctuations, which lead to the large uncertainty in the regression model's predictions. Moreover, the provided full spectra (with over 40000 wavelengths, i.e., dimensions) contained plenty of redundant and correlated information. Consequently, the PLS model could not perform effective feature extraction, and the predictive ability of the model was found to be limited. To address these challenges and to improve the accuracy of the regression model, we employed the feature selection method named spectral distance variable selection (SDVS), which has been proposed by our team for the analysis of iron ore slurry samples [57]. With the thousands of selected features, we built 4 new models. Finally, the performance of 8 models was compared via leave-one-out cross-validation and the best-performing model for each element was selected.

3.2.1. Spectral pre-processing

We only applied feature selection as a pre-processing step. Our feature selection method utilized ideas of linear discriminant analysis (LDA). The training dataset consisted of 2100 spectra from 42 iron alloy targets. The spectra from the same target were regarded as the same category (or class). Subsequently, the ratio of between-class and within-class scatter was used to measure the importance of each variable. Thus, the ratio vector (feature importance) is defined as:

$$\boldsymbol{d} = \frac{S_b}{S_w} \quad (5)$$

where $\boldsymbol{S}_b$ is the between-class scatter vector defined as:

$$\boldsymbol{S}_b = N \sum_{c=1}^{C} (\boldsymbol{\mu}^c - \boldsymbol{\mu}) \odot (\boldsymbol{\mu}^c - \boldsymbol{\mu}) \quad (6)$$

and $\boldsymbol{S}_w$ is the within-class scatter vector given by:

$$\boldsymbol{S}_w = \sum_{c=1}^{C} \sum_{n=1}^{N} (\boldsymbol{x}_n^c - \boldsymbol{\mu}^c) \odot (\boldsymbol{x}_n^c - \boldsymbol{\mu}^c) \quad (7)$$

where $\odot$ represents the Hadamard product of two vectors, $\boldsymbol{x}_n^c$ is the spectrum of the nth spectrum of the cth target, $\boldsymbol{\mu}^c$ is the mean spectrum of the cth target, and $\boldsymbol{\mu}$ is the mean spectrum of the whole training dataset.

A large d value (element of the vector $\boldsymbol{d}$) corresponds to a low within-class scatter and a high between-class scatter, which is beneficial for improving the accuracy of the regression model. When the SDVS method is employed for spectral feature selection, d values are arranged in a descending order. Commonly, the optimal number of features is determined by evaluating the model's performance on the validation dataset. However, considering that the provided training dataset contained spectra of only 42 targets, the number of feature variables was directly set as 1000. In addition, in order to rank the importance of each feature dimension more comprehensively, the top 1000 features with the highest correlation between the spectrum with each element were calculated separately. Finally, the selected important spectral features were further filtered by taking the intersection of the four sets of features (one for each element of interest). The spectral range obtained in this fashion was used as the input of the PLS model. The complete methodology is shown as a flowchart in Figure 4.

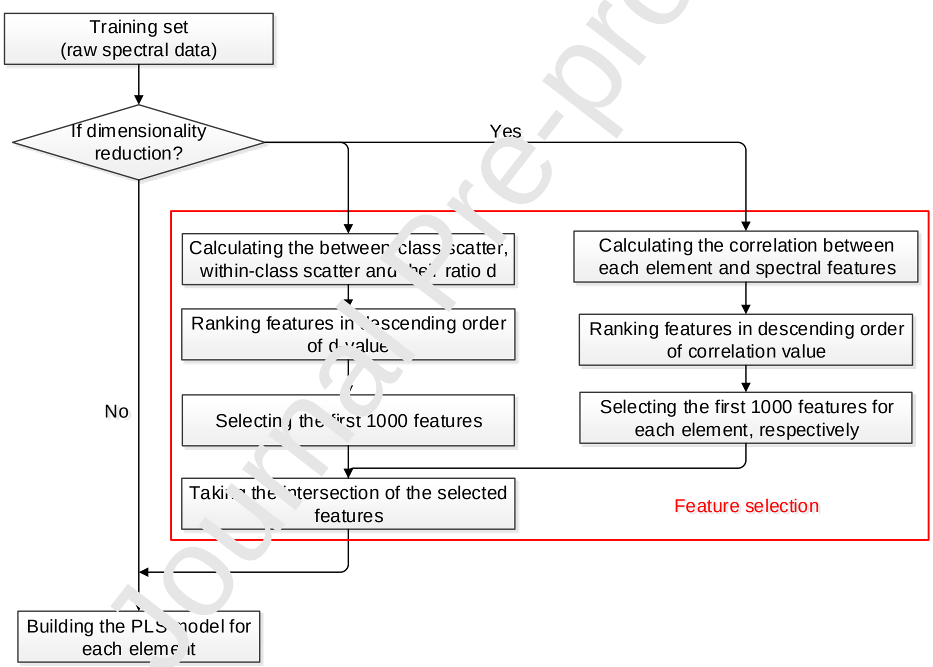


Figure 4 Flowchart representing the SIA team's methodology.

3.2.2. Model selection and optimization

We first fitted 4 PLSR models on the full spectra. As shown in Figure 5, observing the variation of the training RMSE and the validation RMSE with the changing number of PLS latent variables, using more than 15 latent variables does no longer improve the PLS model's performance. Considering the limited size of the training dataset, we did not optimize the number of latent variables individually for each model (with one model used to predict one element's concentration). Instead, the number of latent variables was set to 15 for every PLS model.

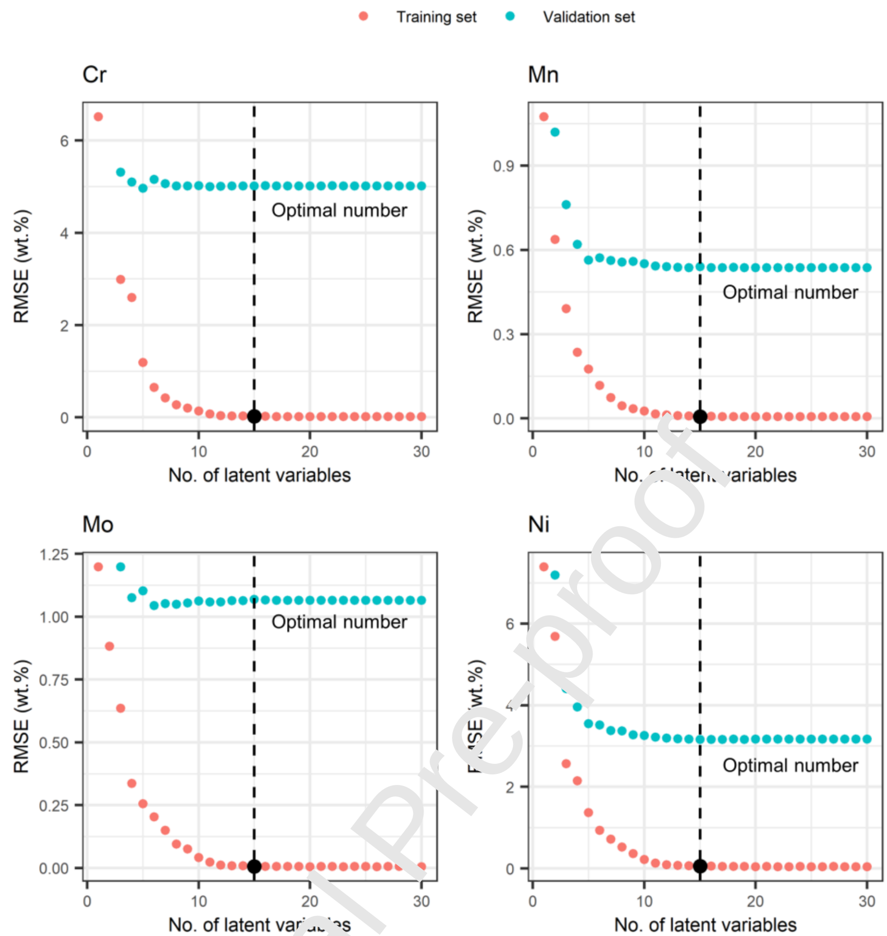


Figure 5 Selecting the optimal number of latent variables based on training and validation errors.

With two different input forms (raw spectra and spectra normalize to unit maxima) and two different output forms (single element prediction and combined elemental prediction), we could establish 4 different models for each element. The optimal model for each element was selected based on the RMSE performance of leave-one-out cross-validation, as shown in Figure 6.

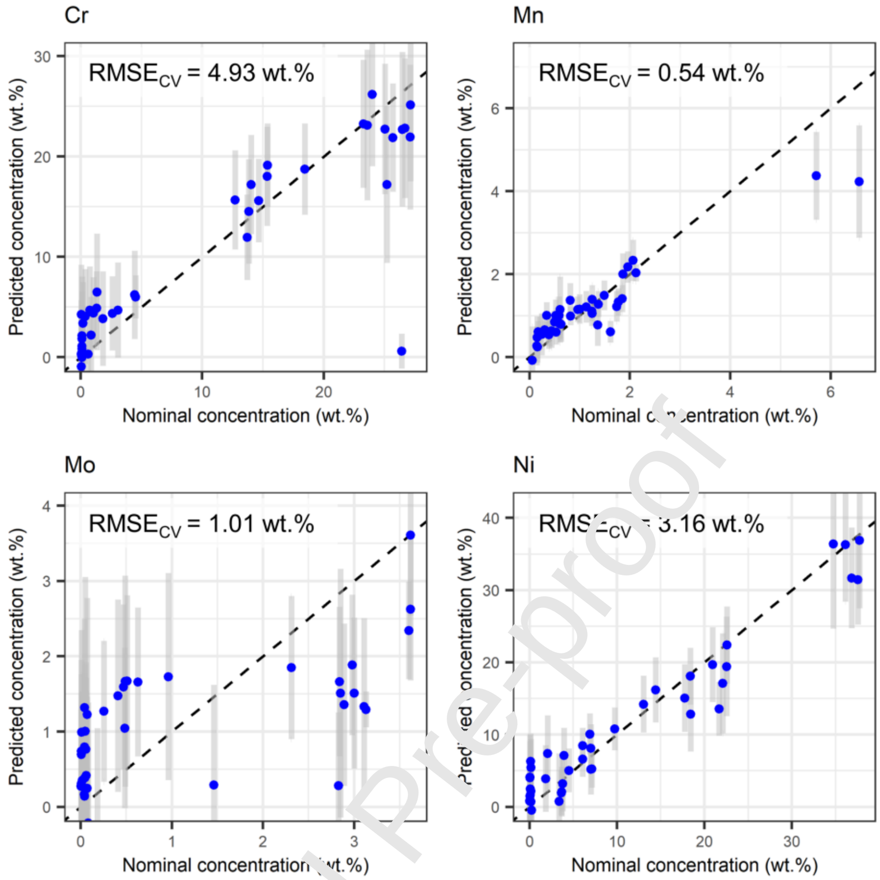


Figure 6 The optimal leave-one-out cross-validation performance for each element predicted based on the full-spectra. The error bars represent the standard deviation of the values obtained via cross-validation.

Considering the regression results in Figure 6, the models based on the full spectra were found to underperform. Exploring the relationship between the normalized intensity of Mn I 279.48 nm and Ni I 341.48 nm emission lines and the predicted Mn and Ni concentration, respectively (Figure 7), we generally found large fluctuations among the spectra of the same targets. This is consistent with the large uncertainty of the predicted concentration values. In addition, owing to the number of features exceeding 40000 wavelength values, the PLS model cannot effectively extract features, thus affecting the model's prediction accuracy.

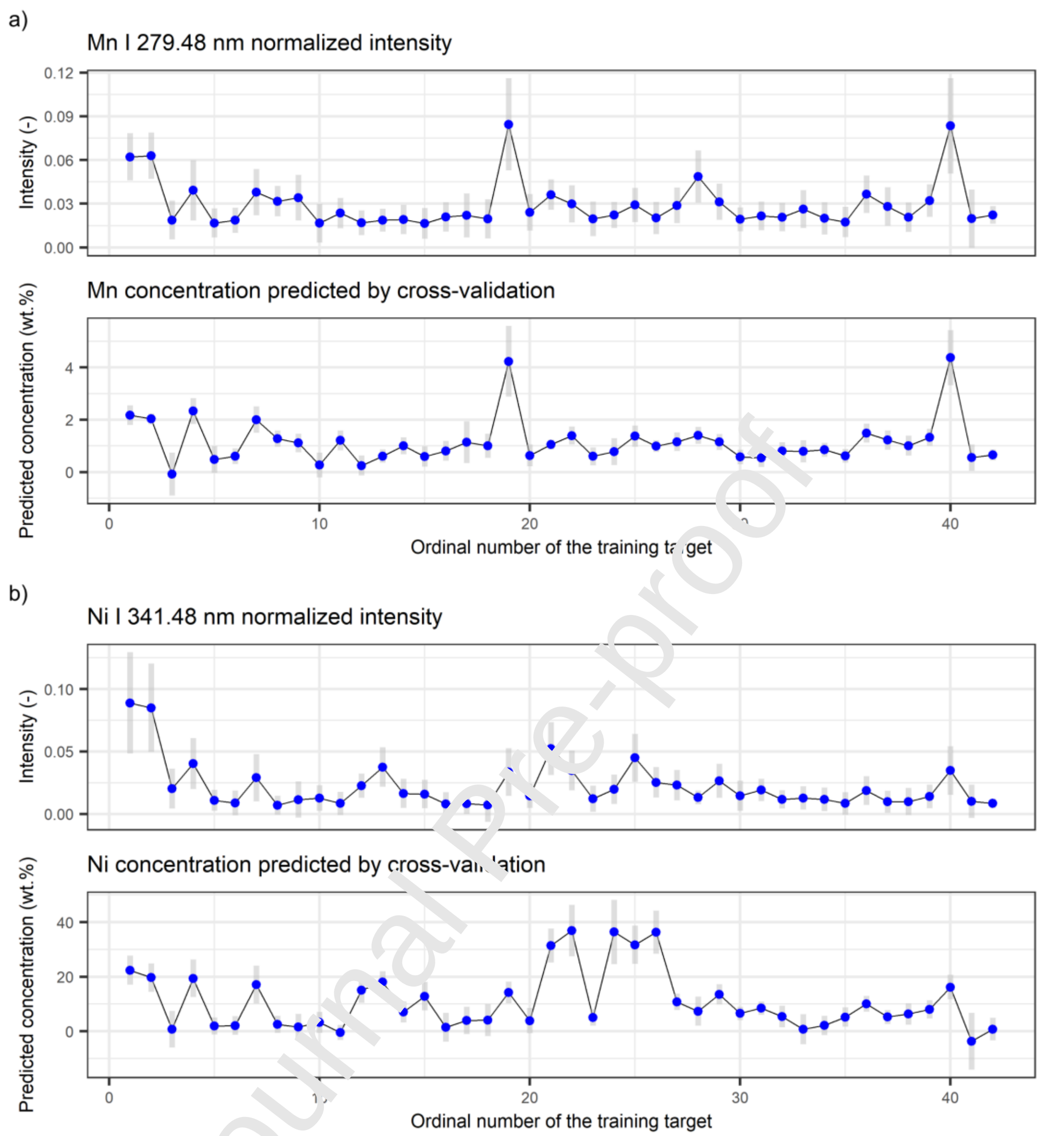


Figure 7 Relationship between the normalized intensity of selected elemental emission lines and their corresponding predicted concentration. (a) Mn 279.48 nm. (b) Ni 341.48 nm. The error bars represent the standard deviation of the values obtained from the 50 distinct spectra.

Consequently, to improve the accuracy of the regression model, we applied the SDVS method and calculated the correlation among the spectral features and the corresponding elemental compositions. After ranking the importance of each feature, the first 1000 features were extracted. The final feature subset was obtained by taking the intersection. Figure 8 shows the position of the selected features in the LIBS spectrum of an iron alloy sample.

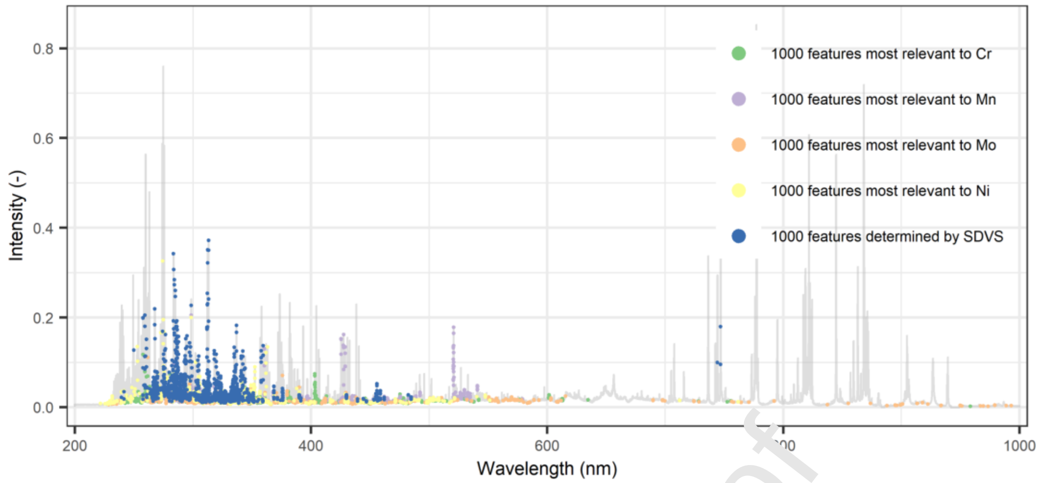


Figure 8 Position of the selected features

Based on the selected features, we built 4 new PLS models for each element. Finally, 8 regression models were established for each element. The leave-one out cross-validation model performance in terms of RMSE of every model is summarized in Table 2. Based on these results, the optimal model for each element was selected and their performance is shown in Figure 9.

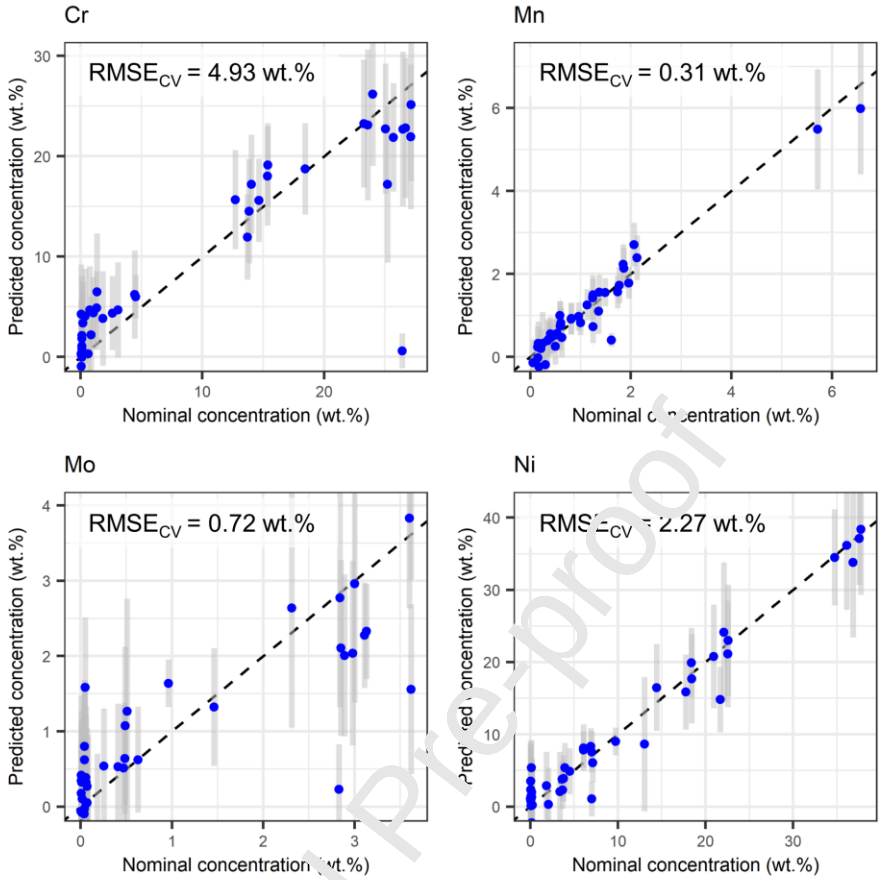


Figure 9 The optimal models' leave-one-out cross-validation performance for each element. The error bars represent the standard deviation of the values obtained via cross-validation.

Table 2 Leave-one-out cross-validation performance in terms of RMSE of every model for each of the considered elements The best-performing models are highlighted for each element.

Input Features	Input Spectra	Output Model	Mn	Cr	Mo	Ni
Full Spectra	Raw Spectra	Single-output	0.56	4.93	1.02	3.84
		Multiple-output	0.57	4.94	1.01	3.98
	Normalized Spectra	Single-output	0.54	5.02	1.07	3.16
		Multiple-output	0.56	5.08	1.05	3.39
Feature Selection	Raw Spectra	Single-output	0.32	5.83	0.72	2.43
		Multiple-output	0.31	5.70	0.72	2.53
	Normalized Spectra	Single-output	0.32	6.05	0.86	2.35
		Multiple-output	0.33	6.20	0.75	2.27

Lastly, we found that for Cr, using the full spectra yielded better performance than using a selected feature subset. Upon further investigation, we found an outlier target in the training dataset in terms of Cr elemental content. The spectral comparison between this target and another target with similar Cr content is shown in Figure 10. An obvious difference between these spectra can be observed, which explains the abnormal predictions yielded by the model. Nevertheless, as there were only 42 training targets available, and this target was not found to be an outlier considering other elements, we did not remove it from the training.

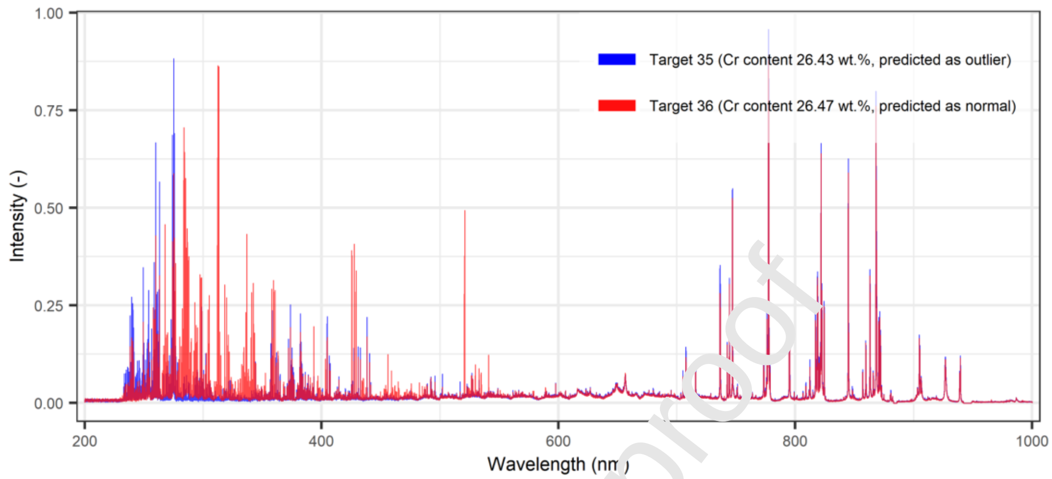


Figure 10 Spectral comparison of the outlier target considering its provided Cr content.

3.3. Dillam Jossue Díaz-Romero, Simon Van den Eynde, Isiah Zaplana, Jef Peeters (DD team)

To estimate the concentrations of the four alloying elements (Cr, Mn, Mo, Ni) in the test targets, we used a deep learning regression model. The most novel aspect of the method is the application of a multi-loss function in the deep learning model, first proposed for quantitative LIBS analysis by our team [58]. The method was inspired by the loss-conditional training [59] which is motivated by the idea of training a single model that covers all coefficients of the loss terms instead of training a separate model for each set of coefficients. The loss term is the error function [60] quantifying the difference between the output of the DL model and the specified target value. The advantage of using a multi-loss function over a separate loss function for each element is that the concentration of multiple elements is learned simultaneously when training the model. As such, the model can learn the relation between the spectral features and the concentrations of the different alloying elements so that improving the performance on one element does not necessarily compromise the performance on the other elements.

3.3.1. Spectral pre-processing

The proposed method does not require the application of filtering, normalization, baseline correction, or dimensional reduction. The only applied pre-processing step is reducing the signal from 40002 to 8190 data points to reduce the computational cost for the DL algorithms. The 8190 wavelengths used by the model, defined as the region of interest (ROI), span the wavelength range between 340 nm and 503.8 nm. As such, only 20.4% of the provided data is used for the regression, as shown in Figure 11.

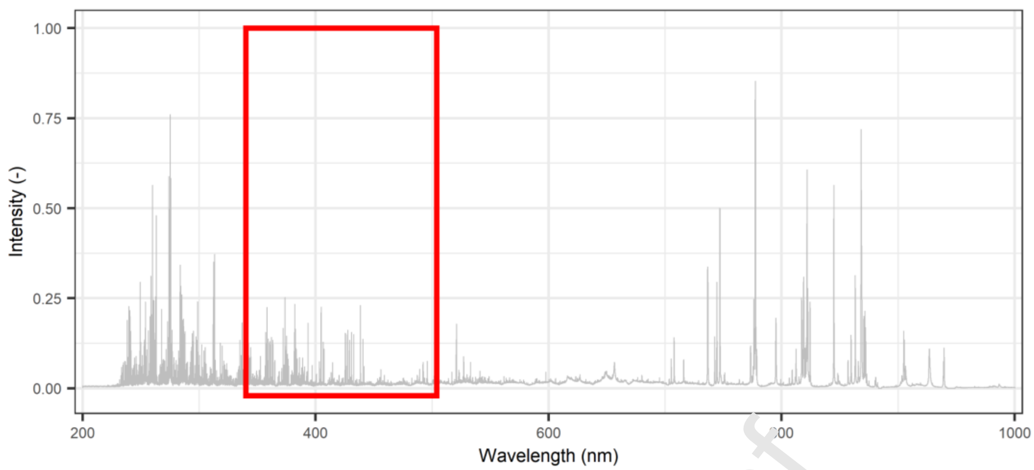


Figure 11 Example of the raw spectra with the proposed ROI.

The ROI was chosen to include the significant atomic and ionic emission lines corresponding to the alloying elements Cr, Mn, Mo, and Ni. The wavelengths of the emission lines per element that are considered critical for the quantitative analysis are shown in Table 3. The wavelengths in Table 3 have been successfully used in previous research as well as to predict the composition of certified steel targets [61–63]. The selected ROI is fed entirely into the DL model, allowing the regression model to learn from all data points in the selected wavelength range rather than just the specific emission lines. This approach makes the system more robust against matrix effects that typically complicate the application of regression methods. In addition, the reduced ROI size (compared to the originally provided data) allows the use of larger batches for training the DL model while keeping the computational cost reasonable.

Table 3 Relevant spectral lines to determine the region of interest (ROI).

Element	Wavelength (nm)[a]	Element	Wavelength (nm)[a]
Cr	374.59	Ni	440.115
	397.6		471.442
	426.135		503.3
	[illegible]		
	438.4	Mn	403.307
Mo	386.4		403.449

[a] The emission lines and spectral ranges were identified based on the NIST atomic spectra database [55].

3.3.2. Model selection and optimization

In order to avoid overfitting the proposed DL model, and thus, to improve the prediction performance and accuracy, we performed data augmentation as a first step. Our data augmentation strategy was the same as the one proposed in our previous works [58,64,65], and it consisted of creating synthetic data and increasing the number of training epochs. For the creation of synthetic data from existing data, the methods *random intensity* and *random wavelength masking*, inspired by the Pytorch audio feature augmentation [66] were used. Random intensity changes the intensity of the original signal between -35% and $+35\%$ to mimic the change in spectral intensity caused by imperfect laser-object interaction, e.g., due to surface contamination. Meanwhile, random wavelength masking is adopted to produce a mask with a width between 0 and 50 nm somewhere in the range between 340 and 503.8 nm, reducing the intensity value to zero in the masked region.

The chosen DL model is a fully connected ANN and was initially proposed in [64] for the purpose of aluminium scrap classification. Figure 12 depicts the schematic representation of the proposed ANN architecture, which consists of three fully connected (FC) layers (which are the same as in the classification model proposed in [58,64]) followed by four FC layers in parallel. The dimensions of each one of the layers are (4096,2048), (2048,1024), (1024,2048), and (2048,1)×4, respectively (where the first integer denotes the input's size, while the second integer denotes the output's size). The Rectified Linear Unit (ReLU) activation function was used after every layer.

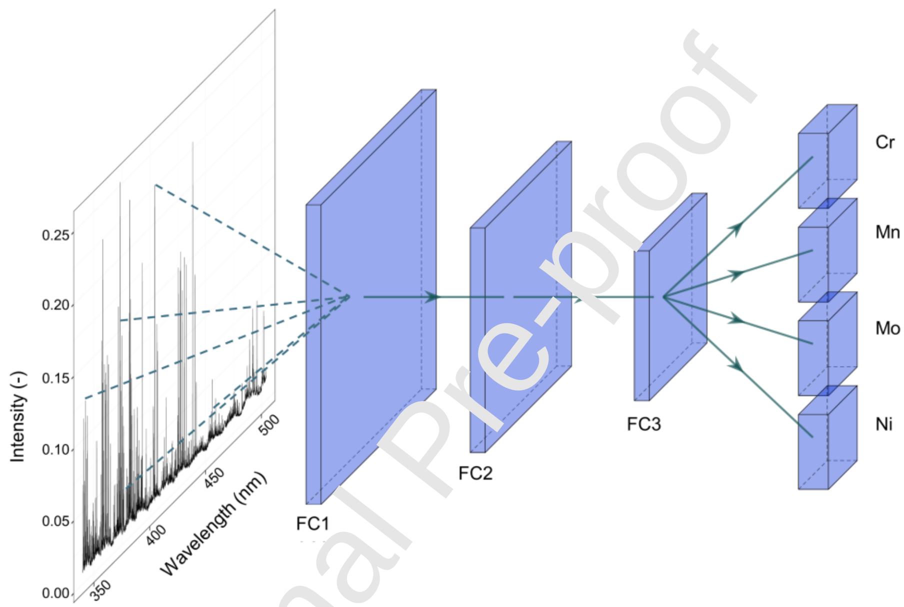


Figure 12 Schematic representation of the ANN architecture used by the DD team.

As for the training, a weighted multiple-component loss function (henceforth referred to as multi-loss function) has been used. This multi-loss function, inspired by [59], has been used by the authors for the purpose of aluminium classification and plastics detection, and plastic mass estimation [58,65,67]. The global loss $\mathcal{L}$ used for the stochastic gradient descent algorithm is calculated as

$$\mathcal{L} = \mathcal{L}_{\mathrm{Cr}} + \mathcal{L}_{\mathrm{Mn}} + \mathcal{L}_{\mathrm{Mo}} + \mathcal{L}_{\mathrm{Ni}} \quad (8)$$

where the four sub-loss functions $\mathcal{L}_e$ (with their subscripts e referring to the four elements of interest) are of the L_1 type loss function (i.e., the sum of the absolute values of the differences between the model's prediction and the corresponding ground truth value). In addition, we weighted the sub-loss functions to prioritize the learning of one or more of the outputs. We used two different sets of weights, namely $\boldsymbol{\omega} = (\omega_{Cr}, \omega_{Mn}, \omega_{Mo}, \omega_{Ni})$ and $\boldsymbol{w} = (w_{Cr}, w_{Mn}, w_{Mo}, w_{Ni})$, which are applied as according to:

$$\mathcal{L} = \begin{cases} \omega_{Cr}\mathcal{L}_{\mathrm{Cr}} + \omega_{Mn}\mathcal{L}_{\mathrm{Mn}} + \omega_{Mo}\mathcal{L}_{\mathrm{Mo}} + \omega_{Ni}\mathcal{L}_{\mathrm{Ni}} & \text{if epoch} \leq \text{steps} \\ w_{Cr}\mathcal{L}_{\mathrm{Cr}} + w_{Mn}\mathcal{L}_{\mathrm{Mn}} + w_{Mo}\mathcal{L}_{\mathrm{Mo}} + w_{Ni}\mathcal{L}_{\mathrm{Ni}} & \text{if epoch} > \text{steps} \end{cases} \quad (9)$$

The first set of weights is assigned to the multi-loss function during the first epochs (until reaching a predefined limit defined by the input parameter steps), while the latter is assigned in the last epochs of the training. For the present task, the used weighted vectors were $\boldsymbol{\omega} = (1.25, 1, 1.25, 1)$, $\boldsymbol{w} = (1, 1, 1, 1)$, with a value for the parameter steps equal to 250. The first weight vector, $\boldsymbol{\omega}$, prioritizes the learning of Cr and Mo which are more challenging to predict than Mn and Ni. Meanwhile, the second weight vector, $\boldsymbol{w}$, does not prioritize learning any of the elements. Finally, the Adam optimizer [68] was used with an initial learning rate of 0.0001 with weightless decay and a learning schedule of $[50, 80, 100, 150, 800]$ and a scheduler gamma of 0.5. Batch sizes of 10 and over 60 epochs are used. The models were implemented using the Pytorch DL framework [66]. The provided training dataset was randomly divided into two datasets, namely 80% for training and 20% for validation. The experiments were carried out using a NVIDIA RTX3090 24 GB GPU, an AMD Ryzen Thread ripper 3990x 64-core CPU, and 128 GB DDR4 RDIMM memory.

Finally, the performance of the proposed regression method has been evaluated with the following three metrics: the root-mean-square error (RMSE, (1)), mean absolute error (MAE, (2)), and coefficient of determination (R^2, (10)) metrics, where:

$$R^2 = \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (10)$$

Here, $\bar{y}$ is the mean concentration across the N targets. RMSE is the standard deviation of the residuals (prediction errors), MAE is the arithmetic average of the absolute errors on the predictions, regardless of the direction of over- or under-prediction, and R^2 is the proportion of variation in the dependent variable that is predictable from the independent variable. Therefore, R^2 represents how well the model fits the dependent variables.

Final concentration prediction

With the developed DL regression model, an estimate of the alloying element concentrations for the 15 test targets can be made. The regression model provides alloying element concentration prediction from a single spectrum. Since 50 spectra were provided for each of the 15 test targets, the regression model yielded 50 different sets of estimates (1 set consisting of the 4 concentration values) for the alloying element concentrations of each target. Since the contest's goal was to provide a single number for the concentration of each of the four elements in each target, the 50 sets of concentrations per target were first compared to each other. Most estimates were (relatively) close to each other. However, several outliers were observed. Thus, the corresponding spectra were manually inspected. If a spectrum corresponding to one of the outliers had a significantly weaker overall intensity than the other spectra of the target, the outlier was deleted from the set of estimates since that estimate would most likely be less accurate than those corresponding to the other spectra. The distribution of the estimated alloying element concentrations of the 15 target targets after filtering the outliers is shown in Figure 13 in the form of box plots. For Cr and Ni, most of the estimates based on the spectra that passed the filtering step were very close to each other. More variation in the estimated concentrations was noticed for Mn and Mo, specifically for targets 6, 9, and 10.

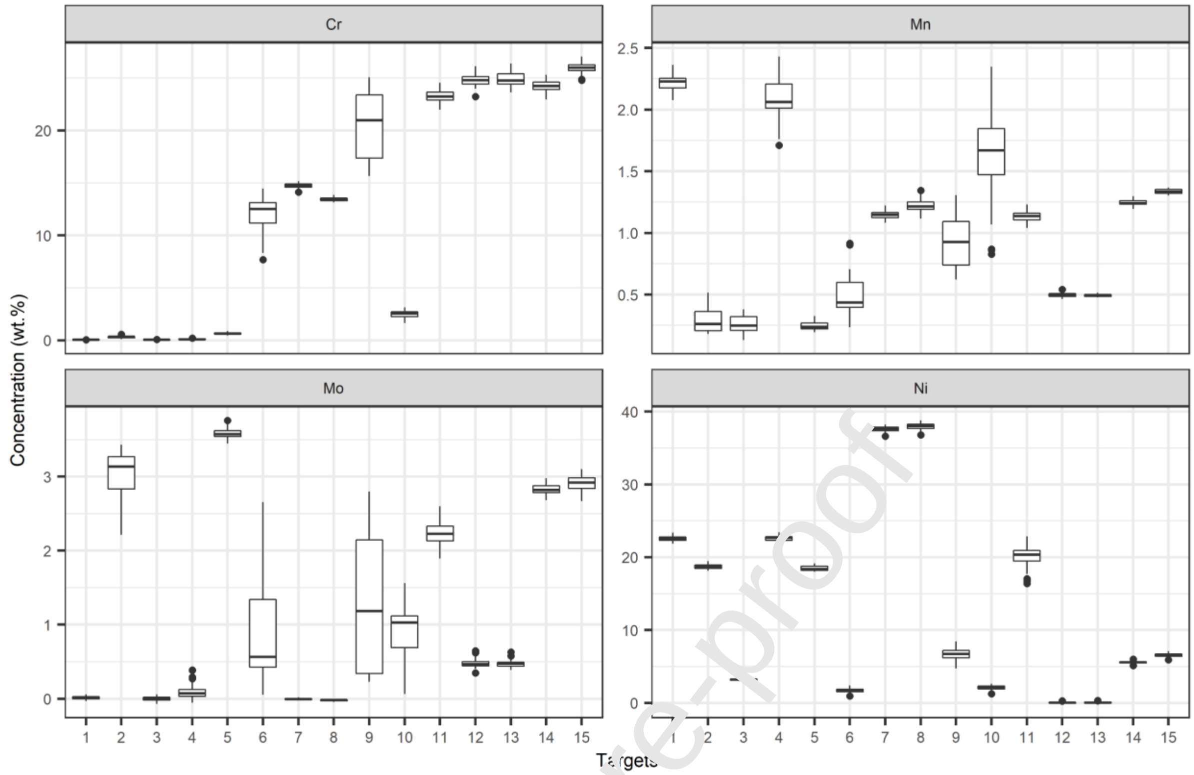


Figure 13 Distribution of the estimated alloying element concentration predictions of the 15 test targets after filtering outliers. The boxes summarize the values obtained from the individual spectra.

The same data was also visualized in the form of histograms to facilitate their interpretation (an example is shown in Figure 14). The final estimate was selected based on the shape of these histograms. When the distribution was approximately centered around the median of the estimates corresponding to the spectra that passed the filtering step, the median of these estimates was used as the final estimate. However, if the distribution was heavily skewed in one direction, the final estimate was chosen not as the median value but the quartile value of the distribution in the direction in which the distribution was skewed. This was expected to give a more accurate result than the median value in these cases.

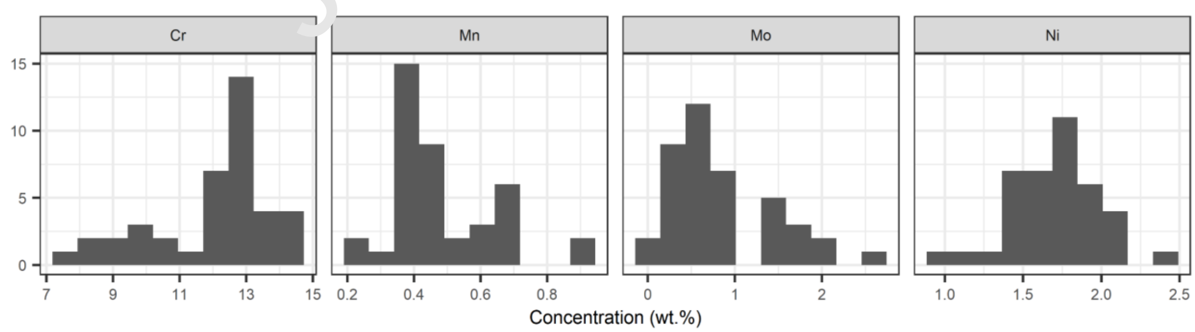


Figure 14 Histograms of the estimated alloying element concentrations of the test target 6.

4. Results and Discussion

This section presents several observations that could be made from the contest. Note that only the mean ranking determined by RMSE is considered here. Additional approaches for ranking the contestants are presented in Appendix B.

Two out of the three approaches were centered around ANNs, while the third approach combined an extensive feature selection optimization with a linear PLS regression model. All three teams addressed the sparsity and redundancy of the LIBS spectra [69], although the three approaches varied significantly. Namely, the FORTH team manually selected full emission lines. This allowed the employed ANN to potentially consider changes in the line shape, even if the intensity did not necessarily differ. This could have arisen as a consequence of self-absorption or spectral intensity with less intense emission lines. However, the simplicity of the FORTH team's ANN model (5 hidden neurons) limits the ANN's encoding capabilities. Hence, whether the ANN is truly capable of detecting changes in the line shape should be further explored. Nevertheless, this is beyond the scope of the present work. Meanwhile, the SIA team applied an automatic wavelength selection methodology, which did not ensure that whole emission lines are available for the subsequent analysis. Nevertheless, the SIA team relied on a linear PLS algorithm which considers the individual wavelengths separately. Lastly, the DD team used a subset of the available spectral range. This approach potentially left plenty of redundant information in the datasets.

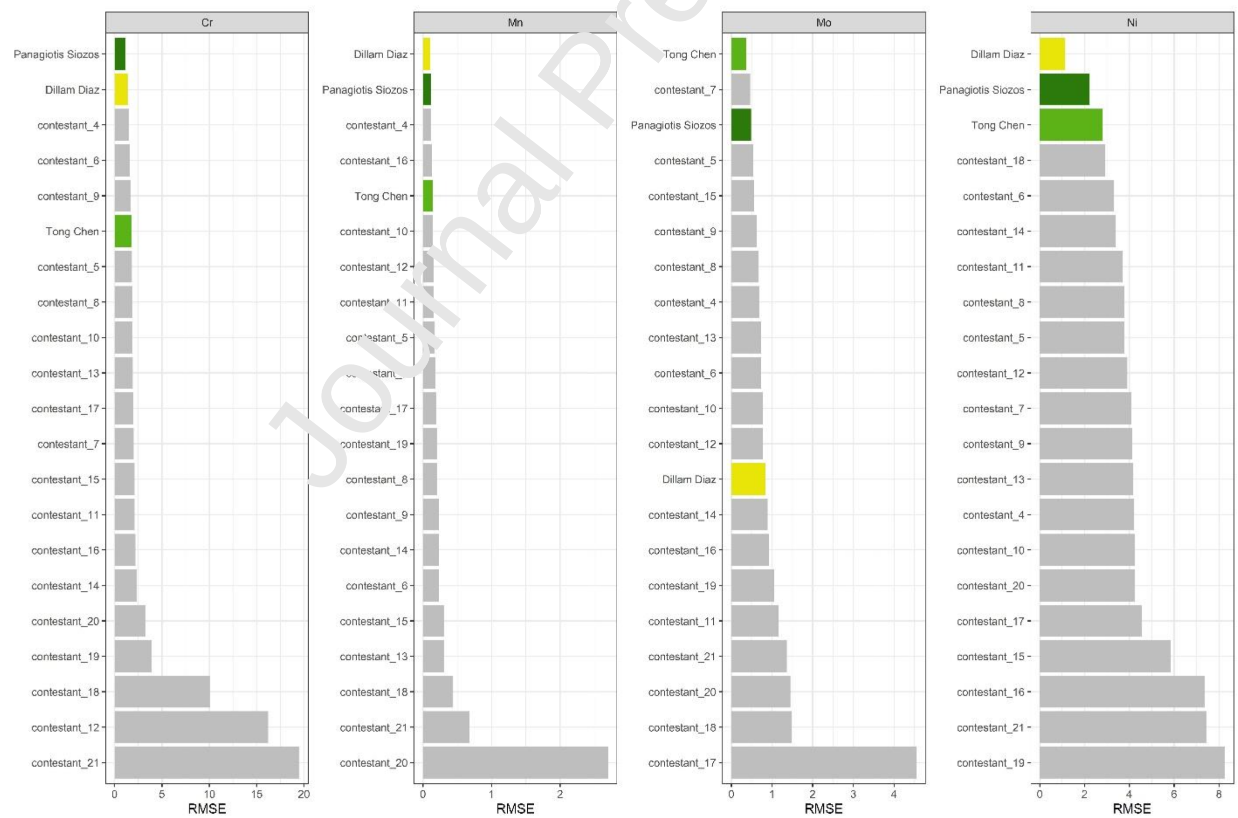


Figure 15 Ranking of the participants on the individual elements as determined by RMSE. A smaller value (at a higher position) represents better performance. Here, Panagiotis Siozos represents the FORTH team, Tong Chen represents the SIA team, and Dillam Diaz represents the DD team.

The performance of the three approaches on the individual elements is summarized in Figure 15. Most participants were kept anonymous (the numbering of these participants is done based on the final ranking, Figure 16). Note that only the FORTH team ranked among the top three contestants consistently across all four elements. Nevertheless, no team managed to claim first place across all four elements. This is in line with the common observation that the regression methodologies must generally be tailored to the considered material constituent [70]. Overall, comparable prediction errors were obtained by the three teams on Mn. On the contrary, while the three top teams provided the most accurate predictions for Ni, the DD team significantly outperformed the other two teams. In turn, Mo presented a limitation for the DD team's methodology.

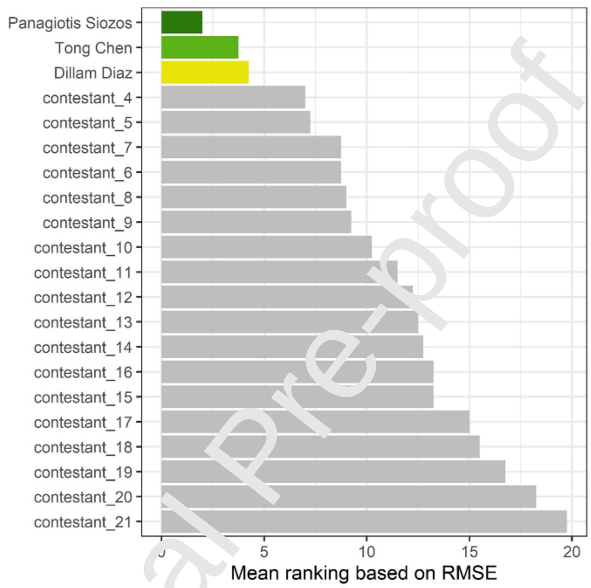


Figure 16 Overall performance of the participants as determined by averaging their obtained ranks across the 4 elements. Here, Panagiotis Siozos represents the FORTH team, Tong Chen represents the SIA team, and Dillam Diaz represents the DD team.

The prediction errors obtained by the 3 teams are discussed in more detail in the following paragraphs. Let us begin with the FORTH team, who combined the careful selection of spectroscopic features with data normalization. In addition, the team applied a bootstrapping approach [71] to reduce the dataset's variance. However, we believe that this step could have benefited the team further, by producing larger training (and testing) datasets. In this respect the FORTH team was unique. In contrast with the DD team (who also used an ANN regression model), the FORTH team kept only the profiles of selected emission lines for further analysis. This minimized the redundancy in the dataset and enabled the application of a relatively simple ANN regression model, especially compared to the elaborate architecture used by the DD team. This is in line with earlier works, which show that relatively simple ANNs can perform well in the context of LIBS analysis [72].

The FORTH team's performance is shown in Figure 17, which shows the nominal vs. predicted concentration values. Consistent prediction accuracy can be observed for Mn, Ni, and Cr. On Mo, the FORTH team's ANN based methodology provided the most consistent predictions in the low concentration range. This is an obvious outcome of the employed logit transformation which was used exactly for this reason. On the contrary, for higher Mo concentrations, the proposed methodology considerably overestimated the element's concentration.

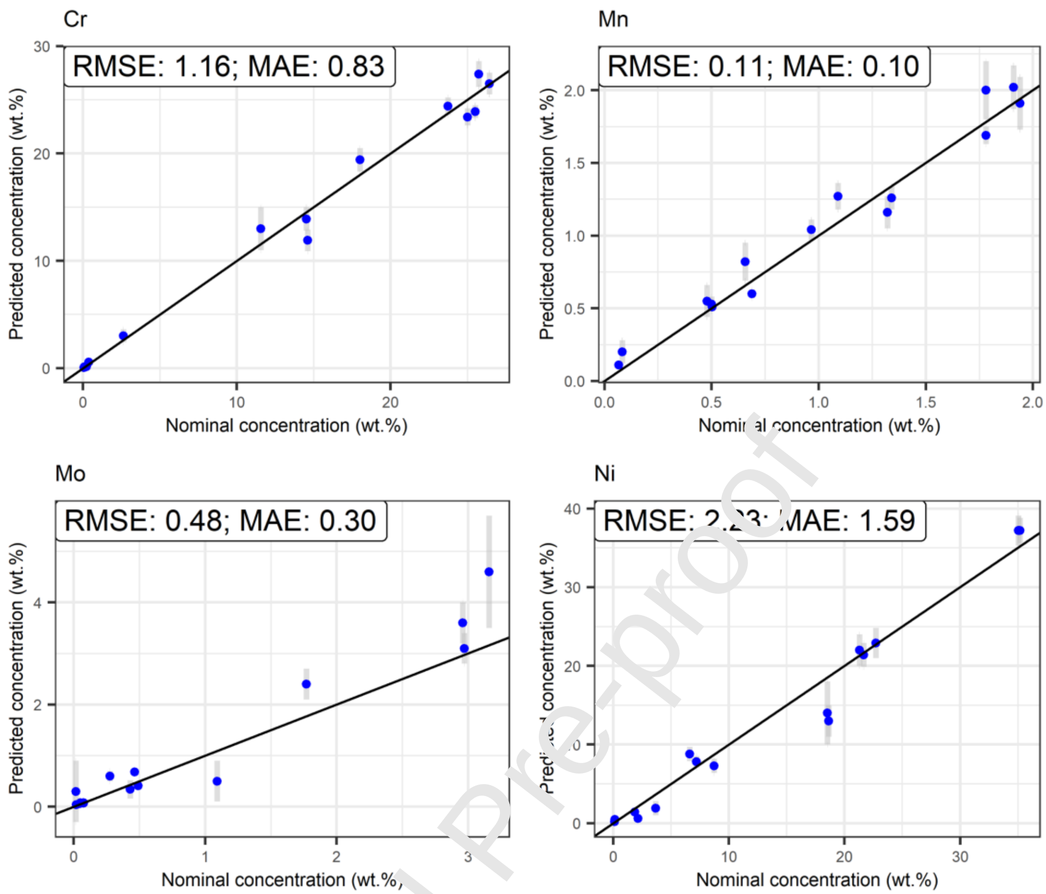


Figure 17 Nominal vs. predicted concentration values of the FORTH team. The error bars represent the standard deviation of the 16 concentration values obtained using the 16 distinct models, each of which is the average of the 50 concentration values predicted for the 50 spectra available for each target.

The SIA team was the only one among the top three with a linear regression model. The achieved performance is not surprising, as it has been shown in the context of support vector machine classifiers that linear and non-linear models tend to learn the same features provided a wide enough spectral range [73]. Nevertheless, the PLS regression model proved to be the least robust: The spectral pre-processing (feature selection and normalization) had to be optimized for each element separately. On the contrary, the FORTH and DD teams employed the same spectral preprocessing steps for every element. It is noteworthy that the linear PLS model of the SIA team yielded the best performance on Mo: the approach performed consistently on concentrations above 0.5 wt.%. However, lower concentrations were overestimated in most cases, suggesting a relatively high limit of quantification via the approach.

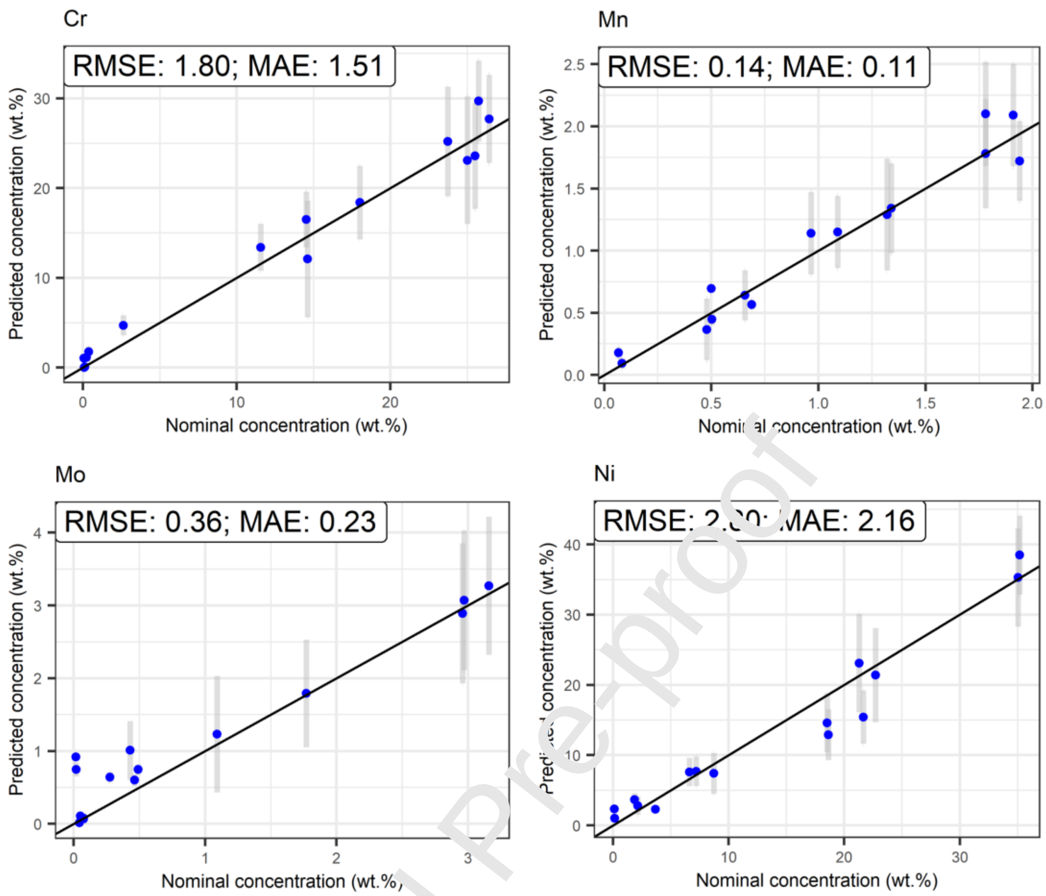


Figure 18 Nominal vs. predicted concentration values of the SIA team. The error bars represent the standard deviation of the 50 predicted concentration values from the 50 spectra available for each target.

Lastly, Figure 19 represents the prediction errors of the DD team. The methodology proposed by the DD team relied on a relatively simple feature selection approach combined with robust data augmentation to improve the generalization capabilities of their model. Note that the predictions yielded by the DD team exhibit the largest differences between RMSE and MAE values. As MAE is less sensitive to outliers, this suggests that while the proposed methodology is less prone to overfitting over most of the concentration ranges, it is not immune to it. Consequently, a few outlier predictions cost the DD team the first place, since their proposed DL approach based on the multi-loss function performed remarkable well on the two major elements and on most of the minor elements: The Ni and Cr concentration were predicted almost perfectly. Nevertheless, a single target's Cr concentration was overestimated quite significantly.

Interestingly, the two worst predictions yielded by the DD team originate from the low concentration range of Mo, where training data were abundant. A potential reason could be the small number of Mo emission lines present in the considered spectral range. At low concentrations, these emission lines are expected to be weak. Consequently, the model is more likely to overtrain on features unrelated to Mo and is, hence, more prone to overfitting this concentration range. An alternative could be exploring the combination of the logit transformation proposed by the FORTH team with the multi-loss approach of the DD team could offer potential benefits. Another aspect of the multi-loss approach is the weighting of the sub-loss functions. Namely, these weights (and the step parameter used to change the weights) introduced additional parameters to be optimized. As the impact of these parameters is currently not well known, the study of their impact and their optimization could lead to further improvements.

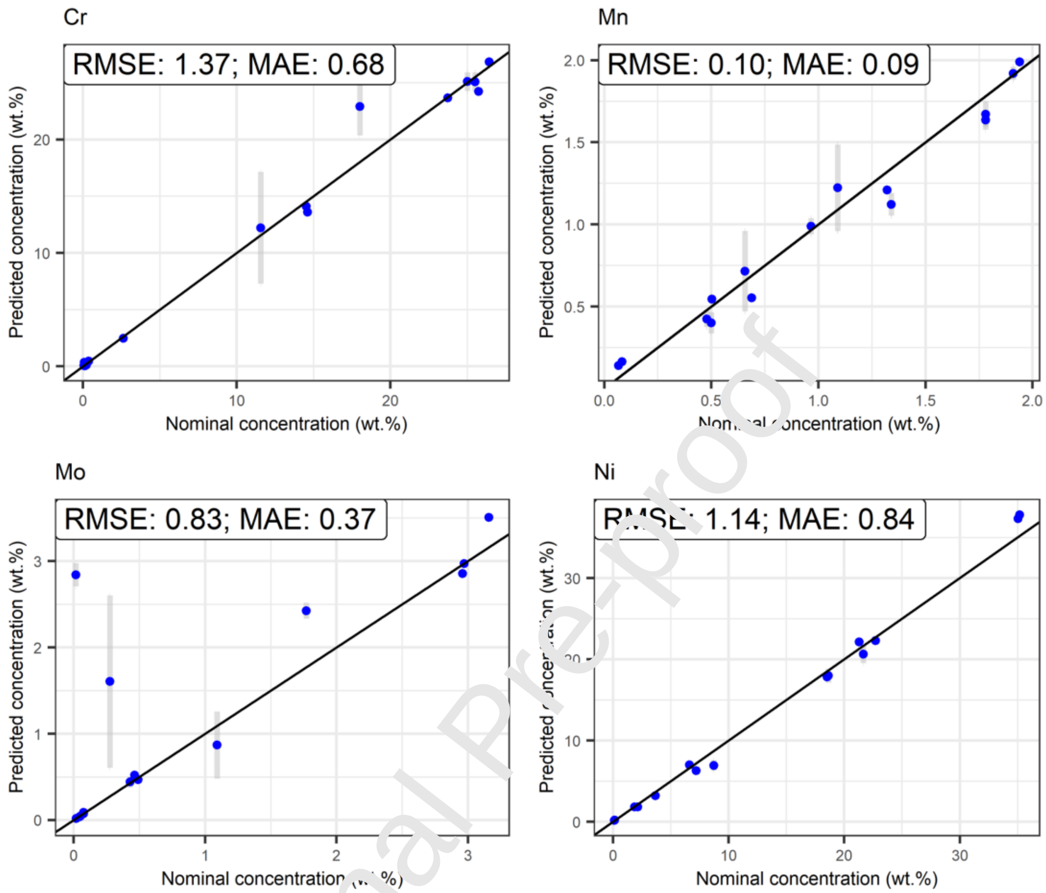


Figure 19 Nominal vs. predicted concentration values of the DD team. The error bars represent the standard deviation of the predicted concentration values after the filtering step described in the text.

We conclude this section by providing a comparative summary of the three presented approaches. The highlights of the three approaches are summarized in Table 4. The strengths and weaknesses of each method (from the organizing team's perspective) are further highlighted in Table 5. Overall, there is a vast difference between the number of trainable parameters of the models chosen the three teams; namely, the small ANN used by the FORTH team contains approximately $1.3 \cdot 10^4$ trainable parameters (although the final prediction was obtained by an ensemble model consisting of 16 such ANNs, essentially increasing the total parameter count by an order of magnitude). On the contrary, the SIA team's PLS model contained $1.5 \cdot 10^4$ to $6 \cdot 10^5$ parameters, depending on the input's type (features selected by the proposed algorithm versus using the full spectra). Nevertheless, the parameters in the PLS model essentially perform a linear transformation of the input data. In contrast, an ANN with a non-linear activation function (such as the piece-wise linear rectified linear unit, ReLU) can fit non-linear functions. Nevertheless, both models' parameter counts are dwarfed by the approximately $2.4 \cdot 10^7$ parameters contained in team DD's model, out of which the majority are shared for predicting the 4 analytes' concentrations simultaneously. Such a large number of parameters (4 orders of magnitude higher than the number of predictive variables) is generally susceptible to overfitting. Nevertheless, team DD battled this challenge by implementing robust data augmentation into their model; in fact, team DD was the only team who performed data augmentation. As a result of the vast difference in model complexity (in terms of both architecture and number of trainable parameters), the training of the DD team's ANN required more machine learning expertise, which practically replaced the spectroscopic expertise applied by the SIA team; namely, the team employed advanced machine learning techniques (for spectroscopic standards), such as learning rate scheduling, longer training (more epochs), and carefully choosing the specialized loss function's weights. A popular approach of reducing the model complexity (in terms of parameter count) while maintaining or even improving the model's capabilities is the use of more sophisticated neural networks, such as ANNs with convolutional layers [74]. These layers perform weight sharing, essentially adding a negligible number of trainable parameters while performing powerful automatized feature selection. Nevertheless, convolutional neural networks have been designed for tasks which exhibit translational invariance, which is not directly the case in spectroscopic analysis. As such, despite the many successful applications of convolutional neural networks in LIBS (and other spectroscopic fields), they do not always yield superior results [75].

Apart from the number of parameters, model complexity can be expressed as the number of steps taken. In the present context, this would entail feature selection, data normalization, architecture optimization, etc. In this regard, the methodologies used by teams FORTH and DD are relatively simple; DD incorporated the whole data processing workflow into a single ANN model, whose architecture does not need to be tailored to the present challenge. Team FORTH optimized their approach to a slightly larger extent, e.g., by selecting the optimal spectral normalization method. In contrast, team SIA optimized their whole workflow for each analyte. Nevertheless, all three approaches are automatable, making the optimization procedures straightforward, i.e., not requiring extensive human intervention.

Table 4 Summary table of the three presented methodologies.

Methodology	*FORTH*	*SIA*	*DD*
Spectral preprocessing	Logit transformation of the analyte concentrations; outlier removal.	None.	None.
Feature selection	Selection of spectral ranges corresponding to emission lines of the analytes and matrix element.	Wavelength selection based on the ratio of the within-class and between-class intensity variances.	Region of interest selection based on the presence of the analytes' emission lines.
Normalization	Normalization to unit maxima.	Normalization to unit maxima.†	None.
Data augmentation	Bootstrapping (averaging 10 spectra).	None.	Random wavelength masking; random intensity modulation.
Regression model	Fully connected artificial neural network (final prediction calculated as the mean of 16	Partial least squares regression without regularization.	Fully connected artificial neural network with 4 output branches.

	independent models' predictions).		
Optimization strategy	75%-25% split of the training data.	Leave-one-out cross validation.	80%-20% split of the training data.
Software Tools	R programming language (neuralnet package).	Python.	Python (PyTorch).

†Not applied for every analyte; please see the text for details.

Table 5 Pros and cons of each presented methodology.

TEAM	PROS	CONS
FORTH	Relatively small model enables the construction of an ensemble method consisting of several small models.	There are no striking limitations of the methodology. However, in the context of this work (i.e., compared to the other ANN-based approach), a separate ensemble model has to be trained separately for each analyte.
SIA	Interpretable models (PLS).	The methodology had to be tailored to each element: both the input and output were optimized for each analyte. In addition, the approach lacked regularization. Lastly, PLS is a linear method, which fails at modeling non-linear relationships.
DD	Single model for all analytes; simple feature selection; end-to-end methodology with robust regularization.	The additional hyperparameters (weights in the loss function, learning rate schedule, etc.) introduced might require extended tuning. However, the impact of these hyperparameters is not straightforward and the tuning might be negligible. The large number of trainable parameters could be viewed as limitation (due to the risk of overtraining) but the authors coupled it with robust data augmentation and regularization strategies.

Considering the dataset, none of the teams attempted to optimize the training dataset. While this is not common practice–beyond outlier removal—selecting an optimal subset of the training dataset has shown promise in transfer learning approaches. In the present context, reducing the training dataset's size might have improved the prediction accuracy of team DD in the low concentration region of Mo, where the team's approach exhibited considerable overtraining. Nevertheless, considering the training dataset's limited size, the further trimming of the dataset would be challenging. Indeed, increasing the training dataset, especially in the underrepresented concentration ranges, would yield more predictable performance enhancements. This could be attempted without obtaining additional targets by data augmentation strategies, specifically generative artificial neural networks. Nevertheless, robust generative models have not yet been demonstrated in the context of LIBS.

5. Conclusion

This work presents three unique approaches applied to a regression task, namely the determination of the concentrations of four alloying elements (Mn, Mo, Ni, Cr) in steel targets based on LIBS spectra. The task was presented as a contest during the 2 months foregoing the LIBS 2022 conference. As the contestants had no access to the testing data, the outcome of the regression challenge is aimed at providing an objective overview of the state of model-based quantitative LIBS analysis. Without a well-defined target application and corresponding requirements on prediction accuracy it is impossible to judge whether LIBS combined with the presented methodologies may finally be considered a reliable semi-quantitative spectroscopic method. Nevertheless, the three methodologies represent considerable advances in semi-quantitative LIBS analysis. More importantly, several potential improvements of the neural network-based methodologies were identified for further research. Meanwhile, similarly to the EMSLIBS 2019 classification challenge, a well-established PLS-based model provided one of the best results. However, this time, a neural network-based approach took the top spot. Overall, the three top-performing teams demonstrated unique approaches, which could be further improved by borrowing from each other. In addition, none of the applied methodologies is specific to the analyzed dataset.

Contributions

EK, JV, PP, and JK were responsible for dataset design, contest organization, and paper writing. VK and AZ provided the targets and the corresponding reference measurements. ADG and VP contributed to the review of the manuscript and provided the platform for the contest. The rest of the authors contributed to the “Methodologies of the top-performing participants” section.

Declaration of Competing Interests

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

Acknowledgments

JV acknowledges the financial support provided through the grant TACR TREND 6 - FW06010042 (Research and development of an advanced interaction vacuum system for laser spectroscopy). PP and EK acknowledge the financial support provided through the grant NCK II - TN02000020 (CAEPO: Center for advanced electronics and photon optics).

Appendix A. Summary table of the training dataset’s concentrations

Table S1 Summary of the analyte concentrations of the targets whose spectra were provided for training. All values are wt.%.

	Mn	Cr	Mo	Ni
mean	1.14	10.06	1.03	10.78
std	1.27	10.65	1.33	12.02
min	0.06	0.05	0.00	0.02
25%	0.39	0.18	0.04	0.18
50%	0.72	3.76	0.22	6.08
75%	1.49	23.26	2.83	18.41
max	6.57	27.16	3.62	37.72

std – standard deviation.

Appendix B. Supplementary data

Supplementary data to this article can be found online at doi

References

[1] R. Noll, Laser-Induced Breakdown Spectroscopy, Springer, Berlin, Heidelberg, 2012. https://doi.org/10.1007/978-3-642-20668-9.

[2] S. Musazzi, U. Perini, Laser-Induced Breakdown Spectroscopy, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014. https://doi.org/10.1007/978-3-642-45085-3.

[3] R. Noll, C. Fricke-Begemann, S. Connemann, C. Meinhardt, V. Sturm, LIBS analyses for industrial applications – an overview of developments from 2014 to 2018, J Anal At Spectrom. 33 (2018) 945–956. https://doi.org/10.1039/C8JA00076J.

[4] A. Harhira, P. Bouchard, K. Rifai, J. el Haddad, M. Sabsabi, A. Blouin, M. Laflamme, Advanced laser-induced breakdown spectroscopy (LIBS) sensor for gold mining, in: Proceedings of the 56th Conference of Metallurgists, Vancouver, BC, Canada, 2017: pp. 27–30.

[5] J. Naozuka, A.P. Oliveira, CHAPTER 4: Laser-induced Breakdown Spectroscopy (LIBS) in Forensic Sensing, in: RSC Detection Science, Royal Society of Chemistry, 2019: pp. 48–83. https://doi.org/10.1039/9781788016117-00048.

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

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

[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 Part B At Spectrosc. 87 (2013) 11–26. https://doi.org/10.1016/j.sab.2013.05.017.

[9] 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. Quertier, B. Sallé, C. Sotin, V. Sautter, H. Séran, J.J. Simmonds, J.-B. Sirven, R. Stiglich, N. Striebig, J.-J. Thocaven, M.J. Toplis, D. Vaniman, The ChemCam Instrument Suite on the Mars Science Laboratory (MSL) Rover: Science Objectives and Mast Unit Description, Space Sci Rev. 170 (2012) 95–166. https://doi.org/10.1007/s11214-012-9912-2.

[10] T. Nelson, R. Wiens, S. Clegg, R. Newell, S. Robinson, S. Storms, J. Michel, M. Caffrey, J. Deming, B. Sandoval, S. Maurice, P. Bernardi, P. Cais, F. Rull, The SuperCam Instrument for the Mars 2020

Rover, in: 2020 IEEE Aerospace Conference, IEEE, 2020: pp. 1–12. https://doi.org/10.1109/AERO47225.2020.9172661.

[11] S. Maurice, R.C. Wiens, P. Bernardi, P. Caïs, S. Robinson, T. Nelson, O. Gasnault, J.-M. Reess, M. Deleuze, F. Rull, J.-A. Manrique, S. Abbaki, R.B. Anderson, Y. André, S.M. Angel, G. Arana, T. Battault, P. Beck, K. Benzerara, S. Bernard, J.-P. Berthias, O. Beyssac, M. Bonafous, B. Bousquet, M. Boutillier, A. Cadu, K. Castro, F. Chapron, B. Chide, K. Clark, E. Clavé, S. Clegg, E. Cloutis, C. Collin, E.C. Cordoba, A. Cousin, J.-C. Dameury, W. D'Anna, Y. Daydou, A. Debus, L. Deflores, E. Dehouck, D. Delapp, G. de Los Santos, C. Donny, A. Doressoundiram, G. Dromart, B. Dubois, A. Dufour, M. Dupieux, M. Egan, J. Ervin, C. Fabre, A. Fau, W. Fischer, O. Forni, T. Fouchet, J. Frydenvang, S. Gauffre, M. Gauthier, V. Gharakanian, O. Gilard, I. Gontijo, R. Gonzalez, D. Granena, J. Grotzinger, R. Hassen-Khodja, M. Heim, Y. Hello, G. Hervet, O. Humeau, X. Jacob, S. Jacquinod, J.R. Johnson, D. Kouach, G. Lacombe, N. Lanza, L. Lapauw, J. Laserna, J. Lasue, L. le Deit, S. le Mouélic, E. le Comte, Q.-M. Lee, C. Legett, R. Leveille, E. Lewin, C. Leyrat, G. Lopez-Reyes, R. Lorenz, B. Lucero, J.M. Madariaga, S. Madsen, M. Madsen, N. Mangold, F. Manni, J.-F. Mariscal, J. Martinez-Frias, K. Mathieu, R. Mathon, K.P. McCabe, T. McConnochie, S.M. McLennan, J. Mekki, N. Melikechi, P.-Y. Meslin, Y. Micheau, Y. Michel, J.M. Michel, D. Mimoun, A. Misra, G. Montagnac, C. Montaron, F. Montmessin, J. Moros, V. Mousset, Y. Morizet, N. Murdoch, R.T. Newell, H. Newsom, N. Nguyen Tuong, A.M. Ollila, G. Orttner, L. Oudda, L. Pares, J. Parisot, Y. Parot, R. Pérez, D. Pheav, L. Picot, P. Pilleri, C. Pilorget, P. Pinet, G. Pont, F. Poulet, C. Quantin-Nataf, B. Quertier, D. Rambaud, W. Rapin, P. Romano, L. Roucayrol, C. Royer, M. Ruellan, B.F. Sandoval, V. Sautter, M.J. Schoppers, S. Schröder, H.-C. Seran, S.K. Sharma, P. Sobron, M. Sodki, A. Sournac, V. Sridhar, D. Standarovsky, S. Storms, N. Striebig, M. Tatat, M. Toplis, I. Torre-Fdez, N. Toulemont, C. Velasco, M. Veneranda, D. Venhaus, C. Virmontois, M. Viso, P. Willis, K.W. Wong, The SuperCam Instrument Suite on the Mars 2020 Rover: Science Objectives and Mast-Unit Description, Space Sci Rev. 217 (2021) 47. https://doi.org/10.1007/s11214-021-00807-w.

[12] E.P. Bertin, Qualitative and Semiquantitative Analysis, in: Introduction to X-Ray Spectrometric Analysis, Springer US, Boston, MA, 1978: pp. 255–278. https://doi.org/10.1007/978-1-4899-2204-5_7.

[13] T. Zhang, H. Tang, H. Li, Chemometrics in laser-induced breakdown spectroscopy, J Chemom. 32 (2018) e2983. https://doi.org/10.1002/cem.2983.

[14] T.-L. ZHANG, S. WU, H.-S. TANG, K. WANG, Y.-X. DUAN, H. LI, Progress of Chemometrics in Laser-induced Breakdown Spectroscopy Analysis, Chinese Journal of Analytical Chemistry. 43 (2015) 939–948. https://doi.org/10.1016/S1872-2040(15)60832-5.

[15] 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-336A. https://doi.org/10.1366/000370210793561691.

[16] A.P. Thorne, Spectrophysics, Springer Netherlands, Dordrecht, 1988. https://doi.org/10.1007/978-94-009-1193-2.

[17] M. Capitelli, G. Colonna, G. D'Ammando, R. Gaudiuso, L.D. Pietanza, Physical Processes in Optical Emission Spectroscopy, in: 2014: pp. 31–57. https://doi.org/10.1007/978-3-642-45085-3_2.

[18] D. Autrique, I. Gornushkin, V. Alexiades, Z. Chen, A. Bogaerts, B. Rethfeld, Revisiting the interplay between ablation, collisional, and radiative processes during ns-laser ablation, Appl Phys Lett. 103 (2013) 174102. https://doi.org/10.1063/1.4826505.

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

[20] D.W. Hahn, N. Omenetto, Laser-Induced Breakdown Spectroscopy (LIBS), Part II: Review of Instrumental and Methodological Approaches to Material Analysis and Applications to Different Fields, Appl Spectrosc. 66 (2012) 347–419. https://doi.org/10.1366/11-06574.

[21] S.V. Shabanov, I.B. Gornushkin, Geometrical effects in data collection and processing for calibration-free laser-induced breakdown spectroscopy, J Quant Spectrosc Radiat Transf. 204 (2018) 190–205. https://doi.org/10.1016/J.JQSRT.2017.09.018.

[22] I.B. Gornushkin, J.M. Anzano, L.A. King, B.W. Smith, N. Omenetto, J.D. Winefordner, Curve of growth methodology applied to laser-induced plasma emission spectroscopy, Spectrochim Acta Part B At Spectrosc. 54 (1999) 491–503. https://doi.org/10.1016/S0584-8547(99)00004-X.

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

[24] Z. Wang, L. Li, L. West, Z. Li, W. Ni, A spectrum standardization approach for laser-induced breakdown spectroscopy measurements, Spectrochim Acta Part B At Spectrosc. 68 (2012) 58–64. https://doi.org/10.1016/j.sab.2012.01.005.

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

[26] V. Palleschi, ed., Chemometrics and Numerical Methods in LIBS, Wiley, 2022. https://doi.org/10.1002/9781119759614.

[27] D. Chung, C. Hyonho, S. Keles, Sparse Partial Least Squares (SPLS) Regression and Classification, (2019). https://cran.r-project.org/web/packages/spls/spls.pdf (accessed April 9, 2021).

[28] M. Haenlein, A.M. Kaplan, A Beginner's Guide to Partial Least Squares Analysis, Understanding Statistics. 3 (2004) 283–297. https://doi.org/10.1207/s15328031us0304_4.

[29] J. El Haddad, A. Harhira, E. Képeš, J. Vrábel, J. Kaiser, P. Pořízka, Chemometric Processing of <scp>LIBS</scp> Data, in: Laser Induced Breakdown Spectroscopy (LIBS), Wiley, 2023: pp. 241–275. https://doi.org/10.1002/9781119758396.ch12.

[30] C.J.C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, Data Min Knowl Discov. 2 (1998) 121–167. https://doi.org/10.1023/A:1009715923555.

[31] P.-H. Chen, C.-J. Lin, B. Schölkopf, A Tutorial on ν-Support Vector Machines: Research Articles, Appl. Stoch. Model. Bus. Ind. 21 (2005) 111–136.

[32] A.D. Gordon, L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, Classification and Regression Trees., Biometrics. 40 (1984) 874. https://doi.org/10.2307/2530946.

[33] W.-Y. Loh, Classification and Regression Tree Methods, in: Encyclopedia of Statistics in Quality and Reliability, John Wiley & Sons, Ltd, Chichester, UK, 2008. https://doi.org/10.1002/9780470061572.eqr492.

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

[35] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature. 521 (2015) 436–444. https://doi.org/10.1038/nature14539.

[36] F. Marquardt, Machine Learning and Quantum Devices, in: M. Devoret, B. Huard, I. Pop (Eds.), Les Houches Summer School Lecture Notes 2019: Quantum Information Machines, Oxford University Press, 2019.

[37] P. Mehta, M. Bukov, C.-H. Wang, A.G.R.R. Day, C. Richardson, C.K. Fisher, D.J. Schwab, A high-bias, low-variance introduction to Machine Learning for physicists, Phys Rep. 810 (2019) 1–124. https://doi.org/10.1016/j.physrep.2019.03.001.

[38] J. Vrábel, E. Képeš, P. Pořízka, J. Kaiser, Artificial Neural Networks for Classification, in: Chemometrics and Numerical Methods in LIBS, Wiley, 2022: pp. 213–240. https://doi.org/10.1002/9781119759614.ch9.

[39] V. Motto-Ros, D. Syvilay, L. Bassel, E. Negre, F. Trichard, F. Pelascini, J. el Haddad, A. Harhira, S. Moncayo, J. Picard, D. Devismes, B. Bousquet, Critical aspects of data analysis for quantification in laser-induced breakdown spectroscopy, Spectrochim Acta Part B At Spectrosc. 140 (2018) 54–64. https://doi.org/10.1016/j.sab.2017.12.004.

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

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

[42] T. Völker, G. Wilsch, I.B. Gornushkin, L. Kratochvilová, P. Poř\'\izka, J. Kaiser, S. Millar, G. Galbács, D.J. Palásti, P.M. Janovszky, others, Interlaboratory Comparison for Quantitative Chlorine Analysis in Cement Pastes with Laser Induced Breakdown Spectroscopy, (2022).

[43] 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 super star, J Anal At Spectrom. 19 (2004) 1061. https://doi.org/10.1039/b400355c.

[44] E. Képeš, J. Vrábel, P. Pořízka, J. Kaiser, LIBS 2022 regression contest; Normalized test dataset, Figshare. (2022). https://doi.org/10.6084/m9.figshare.20141210.v1.

[45] E. Képeš, J. Vrábel, P. Pořízka, J. Kaiser, LIBS 2022 regression contest; Normalized training dataset., Figshare. (2022). https://doi.org/10.6084/m9.figshare.20141213.v1.

[46] E. Képeš, J. Vrábel, P. Pořízka, J. Kaiser, LIBS 2022 regression contest; Raw training dataset., Figshare. (2022). https://doi.org/10.6084/m9.figshare.20297118.v1.

[47] E. Képeš, J. Vrábel, P. Pořízka, J. Kaiser, LIBS 2022 regression contest; Raw test dataset., Figshare. (2022). https://doi.org/10.6084/m9.figshare.20297136.v1.

[48] I.T. Joliffe, Principal Component Analysis, Springer-Verlag, New York, 2002. https://doi.org/10.1007/b98835.

[49] P. Pořízka, J. Klus, E. Képeš, D. Prochazka, D.W. Hahn, J. Kaiser, On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review, Spectrochim. Acta - Part B At. Spectrosc. 148 (2018) 65–82. https://linkinghub.elsevier.com/retrieve/pii/S0584854718301526 (accessed February 5, 2019).

[50] P. Pořízka, J. Klus, D. Prochazka, E. Képeš, A. Hrdlička, J. Novotný, K. Novotný, J. Kaiser, Laser-Induced Breakdown Spectroscopy coupled with chemometrics for the analysis of steel: The issue of spectral outliers filtering, Spectrochim Acta Part B At Spectrosc. 123 (2016) 114–120. https://linkinghub.elsevier.com/retrieve/pii/S0584854716301367 (accessed June 4, 2018).

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

[52] L. Spanos, P. Siozos, A. Koropoulis, N. Hausmann, M. Holst, P. Pavlidis, D. Anglos, Laser-Induced Breakdown Spectroscopy (LIBS) Combined with Machine Learning Models and Neural Networks for the Classification of Hard Tissue Remains, in: 12th International Conference on Instrumental Methods of Analysis, Virtual Event, 2021.

[53] F. Günther, S. Fritsch, neuralnet: Training of Neural Networks, R J. 2 (2010) 30. https://doi.org/10.32614/RJ-2010-006.

[54] R Core Team, R: A Language and Environment for Statistical Computing, (2020). https://www.r-project.org.

[55] A. Kramida, Y. Ralchenko, J. Reader, others, NIST atomic spectra database (ver. 5.3), (2015).

[56] C. Sun, Y. Tian, L. Gao, Y. Niu, T. Zhang, H. Li, Y. Zhang, Z. Yue, N. Delepine-Gilon, J. Yu, Machine Learning Allows Calibration Models to Predict Trace Element Concentration in Soils with Generalized LIBS Spectra, Sci Rep. 9 (2019) 11363. https://doi.org/10.1038/s41598-019-47751-y.

[57] T. Chen, L. Sun, H. Yu, L. Qi, D. Shang, Y. Xie, Efficient weakly supervised LIBS feature selection method in quantitative analysis of iron ore slurry, Appl Opt. 61 (2022) D22. https://doi.org/10.1364/AO.441098.

[58] S. van den Eynde, D.J. Diaz-Romero, I. Zaplana, J. Peeters, Deep Learning Regression for Quantitative LIBS Analysis of Aluminium Scrap, SSRN Electronic Journal. (2022). https://doi.org/10.2139/ssrn.4284144.

[59] A. Dosovitskiy, J. Djolonga, You Only Train Once: Loss-Conditional Training of Deep Networks, in: International Conference on Learning Representations, 2020. https://openreview.net/forum?id=HyxY6JHKwr.

[60] A. Burkov, The hundred-page machine learning book, Andriy Burkov Quebec City, QC, Canada, 2019.

[61] S. Kashiwakura, K. Wagatsuma, Selection of Atomic Emission Lines on the Mutual Identification of Austenitic Stainless Steels with a Combination of Laser-induced Breakdown Spectroscopy (LIBS) and Partial-least-square Regression (PLSR), ISIJ International. 60 (2020) 1245–1253. https://doi.org/10.2355/isijinternational.ISIJINT-2019-549.

[62] J. Wu, H. Yu, Y. Qiu, Z. Zhang, T. Liu, F. Xue, W. Yu, X. Li, A. Qiu, Plasma characteristics and element analysis of steels from a nuclear power plant based on fiber-optic laser-induced breakdown spectroscopy, J Phys D Appl Phys. 52 (2019) 014006. https://doi.org/10.1088/1361-6463/aae7b4.

[63] X. LI, H. LU, J. YANG, F. CHANG, Semi-supervised LIBS quantitative analysis method based on co-training regression model with selection of effective unlabeled samples, Plasma Science and Technology. 21 (2019) 034015. https://doi.org/10.1088/2058-6272/aaee14.

[64] D.J. Díaz-Romero, S. van den Eynde, W. Sterkens, A. Eckert, I. Zaplana, T. Goedemé, J. Peeters, Real-time classification of aluminum metal scrap with laser-induced breakdown spectroscopy using deep and other machine learning approaches, Spectrochim Acta Part B At Spectrosc. 196 (2022) 106519. https://doi.org/10.1016/j.sab.2022.106519.

[65] D.J. Diaz-Romero, S. van den Eynde, I. Zaplana, C. Zhou, W. Sterkens, T. Goedemé, J. Peeters, Classification of Aluminum Scrap by Laser Induced Breakdown Spectrometry (Libs) and Rgb+D Image Fusion Using Deep Learning Approaches, SSRN Electronic Journal. (2022). https://doi.org/10.2139/ssrn.4272447.

[66] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, PyTorch: An Imperative Style, High-Performance Deep Learning Library, (2019).

[67] D.J.D. Romero, I. Zaplana, S. van den Eynde, W. Sterkens, T. Goedeme, J. Peeters, Enhanced Plastic Recycling using RGB+Depth Fusion with MassFaster and MassMask R-CNN, in: 2022 Fourth International Conference on Transdisciplinary AI (TransAI), IEEE, 2022: pp. 22–29. https://doi.org/10.1109/TransAI54797.2022.00010.

[68] D.P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, in: 3rd International Conference for Learning Representations, 2015: pp. 1–15.

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

[70] R.B. Anderson, O. Forni, A. Cousin, R.C. Wiens, S.M. Clegg, J. Frydenvang, T.S.J. Gabriel, A. Ollila, S. Schröder, O. Beyssac, E. Gibbons, D.S. Vogt, E. Clavé, J.-A. Manrique, C. Legett, P. Pilleri, R.T. Newell, J. Sarrao, S. Maurice, G. Arana, K. Benzerara, P. Bernardi, S. Bernard, B. Bousquet, A.J. Brown, C. Alvarez-Llamas, B. Chide, E. Cloutis, J. Comellas, S. Connell, E. Dehouck, D.M. Delapp, A. Essunfeld, C. Fabre, T. Fouchet, C. Garcia-Florentino, L. García-Gómez, P. Gasda, O. Gasnault, E.M. Hausrath, N.L. Lanza, J. Laserna, J. Lasue, G. Lopez, J.M. Madariaga, L. Mandon, N. Mangold, P.-Y. Meslin, A.E. Nelson, H. Newsom, A.L. Reyes-Newell, S. Robinson, F. Rull, S. Sharma, J.I. Simon, P. Sobron, I.T. Fernandez, A. Udry, D. Venhaus, S.M. McLennan, R.V. Morris, B. Ehlmann, Post-landing major element quantification using SuperCam laser induced breakdown spectroscopy, Spectrochim Acta Part B At Spectrosc. 188 (2022) 106347. https://doi.org/10.1016/j.sab.2021.106347.

[71] E. Képeš, P. Pořízka, J. Kaiser, On the application of bootstrapping to laser-induced breakdown spectroscopy data, J Anal At Spectrom. 34 (2019) 2411–2419. https://doi.org/10.1039/C9JA00304E.

[72] A. Safi, B. Campanella, E. Grifoni, S. Legnaioli, G. Lorenzetti, S. Pagnotta, F. Poggialini, L. Ripoll-Seguer, M. Hidalgo, V. Palleschi, Multivariate calibration in Laser-Induced Breakdown Spectroscopy quantitative analysis: The dangers of a ‘black box’ approach and how to avoid them, Spectrochim Acta Part B At Spectrosc. 144 (2018) 46–54. https://doi.org/10.1016/j.sab.2018.03.007.

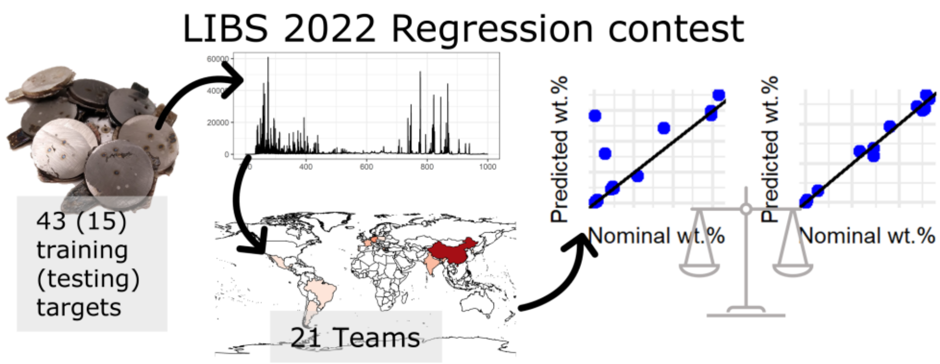
[73] E. Képeš, J. Vrábel, O. Adamovsky, S. Střítežská, P. Modlitbová, P. Pořízka, J. Kaiser, Interpreting support vector machines applied in laser-induced breakdown spectroscopy, Anal Chim Acta. (2021) 339352. https://doi.org/10.1016/j.aca.2021.339352.

[74] E. Képeš, J. Vrábel, J. El Haddad, A. Harhira, P. Pořízka, J. Kaiser, Machine Learning in the Context of <scp>Laser-Induced</scp> Breakdown Spectroscopy, in: Laser Induced Breakdown Spectroscopy (LIBS), Wiley, 2023: pp. 305–330. https://doi.org/10.1002/9781119758396.ch15.

[75] F. Poggialini, B. Campanella, S. Legnaioli, S. Raneri, V. Palleschi, Comparison of Convolutional and Conventional Artificial Neural Networks for Laser-Induced Breakdown Spectroscopy Quantitative Analysis, Appl Spectrosc. 76 (2022) 959–966. https://doi.org/10.1177/00037028221091300.

Declaration of Competing Interests

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


LIBS 2022 Regression contest
43 (15)
training
(testing)
targets
21 Teams
Predicted wt.%
Nominal wt.%
Predicted wt.%
Nominal wt.%

Highlights:

- World-wide LIBS quantification contest with 21 participating teams
- Highly optimized spectral preprocessing strategies, with focus on feature selection
- Potent data augmentation strategies for avoiding overtraining neural networks
- Linear multivariate models perform on par with artificial neural networks