THE EUROPEAN
PHYSICAL JOURNAL D

Regular Article – Plasma Physics

Determination of austenitic steel alloys composition using laser-induced breakdown spectroscopy (LIBS) and machine learning algorithms

Ivan Traparić[a] and Milivoje Ivković[b]

Institute of Physics Belgrade, Pregrevica 118, Belgrade 11080, Serbia

Abstract. In this paper, the determination of composition of certified samples of austenitic steel alloys was done by combining laser-induced breakdown spectroscopy (LIBS) technique with machine learning algorithms. Isolation forest algorithm was applied to the MinMax scaled LIBS spectra in the spectral range form (200–500) nm to detect and eject possible outliers. Training dataset was then fitted with random forest regressor (RFR) and Gini importance criterion was used to identify the features that contribute the most to the final prediction. Optimal model parameters were found by using grid search cross-validation algorithm. This was followed by final RFR training. Results of RFR model were compared to the results obtained from linear regression with $\mathcal{L}^2$ norm and deep neural network (DNN) by means of R^2 metrics and root-mean-square error. DNN showed the best predictive power, whereas random forest had good prediction results in the case of Cr, Mn and Ni, but in the case of Mo, it showed limited performance.

1 Introduction

The structural materials of fusion reactors are subjected to thermal, mechanical, chemical, and radiation loads. Due to their excellent manufacturability, good mechanical properties, welding ability, and corrosion resistance, austenitic stainless steels were chosen as structural reference material for ITER [1]. In addition, entire vacuum vessel of LHD stellarator in Japan is made of austenitic steel [2], and to diagnose the composition of the deposits on the fusion reactor's first wall, test targets made of austenitic steel (AISI 316 L) were settled at ten positions on the first wall [3]. Laser-induced breakdown spectroscopy (LIBS) is one of the emerging analytical technique that is non-destructive, easy to use and requires little to no preparation of the sample [4]. Therefore, it represents a great tool for the analysis of the composition of austenitic steel samples. There are two main approaches to the LIBS analysis, namely standard calibration method and calibration—free method [5]. In the method where calibration curve

is constructed, a connection between one integrated line intensity and known concentrations is established, thus enabling the determination of unknown concentration. This method is by far the most used one. Alternatively, one can assume local thermodynamic equilibrium (LTE) in plasma and use Saha–Boltzmann equation to obtain plasma temperature and density, and from this the unknown concentrations regardless of the matrix effect. Machine learning algorithms have been successfully applied in analysis of Raman spectra, NIR and THz spectroscopy, vibrational spectroscopy, fusion plasma spectroscopy, etc., just to name a few [6–10]. In recent years, to speed up the analysis of LIBS spectra, machine learning methods are being used intensively [11–14]. These methods involve the usage of principal component analysis (PCA) for dimensionality reduction, support vector machine (SVM) for classification purposes and partial least squares regression (PLS) for multivariate regression problems [15]. Also, for classification or regression problems, many authors applied back propagation neural networks (BPNN) or convolutional neural networks (CNN) to the LIBS spectra in order to perform quantitative analysis of different samples [16–20]. Other regression algorithms, like random forest regression (RFR), have also been widely used [21–24]. Random Forest was constructed and reported by Breiman [25], and it is based on the ensemble of decision threes, where the decision or prediction is made

[a] e-mail: traparic@ipb.ac.rs (corresponding author)

[b] e-mail: ivke@ipb.ac.rs

by the majority prediction. This algorithm was previously applied on steel spectra by Zhang et al. [26] where they showed that this regression could be applied for the determination of composition of steel alloys. Later, Zhang with his collaborators used BPNN combined with SelectKBest algorithm for feature selection to trace minor elements in steel samples [27]. Liu and his coworkers also used random forest, combined with permutation importance feature selection to train and predict the composition of steel alloys [28]. Gini importance criteria was also used previously in combination with random forest on classification problems [29,30], but here we are applying it to regression problem.

In this paper, we will consider three algorithms, random forest, linear regression with $\mathcal{L}^2$ norm and deep neural network (DNN) to predict steel samples composition. Instead of making our own database, we will use the dataset published at the LIBS 2022 conference site [31] and record our own test dataset under similar conditions to check how much these small differences affect the final model performance. Idea to use RF algorithm is twofold. On the one end, it is able to catch nonlinear phenomena in the data, on the other end to see to what extent we can use already implemented Gini importance criteria within RF to make good regression model. Although simple neural networks have yielded good analytical prediction in the past, in general, they are hard to train (better said, it is not easy to find most favorable architecture), so we wanted to see how close RF predictions are going to be with respect to DNN.

The paper is organized as follows: In the first section, a brief introduction and overview of previous results is given. In Sect. 2, the experimental setup and sample preparation is described. Section 3 gives the detail description of applied methodology and data preprocessing, while the results are given in Sect. 4. Finally, we gave the conclusion of this work in Sect. 5.

2 Experimental setup and sample preparation

Experimental setup is shown in Fig. 1.

The setup is a classical LIBS setup consisting of Quantel Q switched neodymium-doped yttrium aluminum garnet (Nd:YAG) laser having pulse width of 6 ns, repetition rate of 10 Hz, pulse energy of 96 mJ and operated at fundamental wavelength $\lambda = 1064$ nm. Laser beam was reflected from 45° angle mirror M and focused via lens L onto a target mounted on a $x-y$ micrometric moving stage by a lens of focal length $f = 11$ cm. Light emitted from plasma was collected using a fiber optic cable with collimator having a focal length of $f_{fc} = 4.4$ cm and directed onto the 50 µm width entrance slit of Mechelle 5000 spectrograph that can record spectra from 200 to 950 nm. As a detector, we used Andor iStar ICCD camera (model DH734, 1024 × 1024 pixels) cooled to −15 °C. Camera was triggered with a photodiode and gated by usage of Stan-

ford Research digital delay unit (model DG535). Delay from laser pulse was set to 0.6 µs and the gate width was set to 50 µs.

Steel samples used in this work were AISI steels with certified composition from National Bureau of Standards (NBS, today NIST), whose elemental composition is given in Table 1.

Sample mentioned above, austenitic steel AISI 316 L lies in between these tested models (concentrations of main elements: Cr 17%, Ni 12%, Mo 2% and Mn 2%). Each sample was firstly polished by sandpaper 200, followed by polishing it with sandpaper 600. In front of laser beam, external shutter was placed, coupled with laser pulse counter. Counter was set to 16 counts, as it is a binary counter, and after 16 pulses, the shutter is closed for another 16 pulses. This represents one acquisition of the spectra. For each sample, we recorded 22 spectra from different places on the target, and each spectra is a result of averaging 20 acquisitions on the same place (this gives 320 individual laser shots per place on the target). To further improve and increase signal, electrical gain of the camera was set to 80 (on the scale of 0–255).

3 Methodology and data preprocessing

Database used in this paper was downloaded from LIBS 2022 website [31]. This database consists of a spectra of 42 different steel samples, and for each sample, a 50 single-shot spectra were taken. This gives in total a database of 2100 spectra samples divided into 40,002
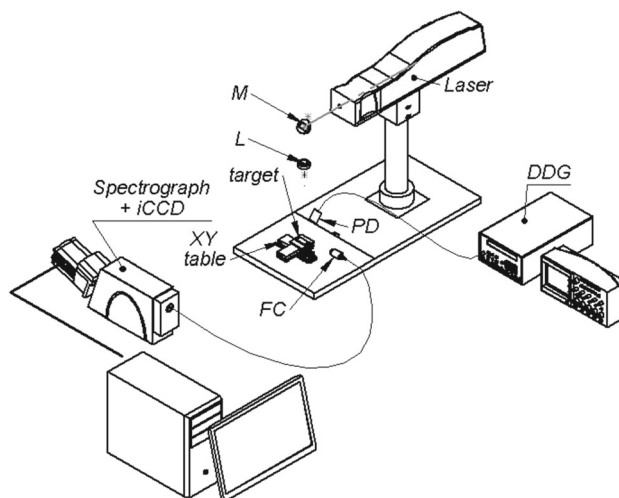

Fig. 1 Experimental setup. Laser (Quantel, $\lambda = 1064$ nm, pulse width 6 ns, peak energy 96 mJ) was focused via lens L onto the movable target and plasma spectrum was recorded by Andor iStar iCCD camera mounted on Echelle spectrograph. Camera gating was done by Stanford Research Digital Delay Generator (DDG, model 535) and triggered by photodiode (PD). Mirror M and lens L are integrated within a laser head, which was not drawn on this figure

Table 1 Steel alloy certified composition

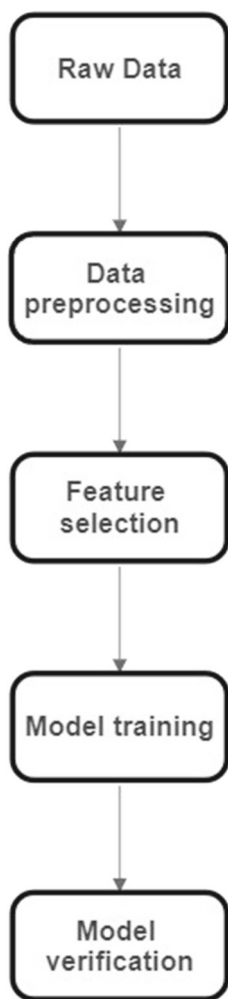| Steel number	Steel type	Cr	Mn	Mo	Ni
443	Cr18.5–Ni9.5	18.5	3.38	0.12	9.4
445	AISI 410	13.31	0.77	0.92	0.28
446	AISI 321	18.35	0.53	0.43	9.11
447	AISI 309	23.72	0.23	0.053	13.26

Fig. 2 Flowchart of procedures taken in this work

columns (each column corresponds to one wavelength). The flow diagram of our methodology is given in Fig. 2. For machine learning part of this work, we used python public repository scikit-learn.

3.1 Data preprocessing

Firstly, we restricted our dataset to the spectral range between 200 and 500 nm, as this is the spectral area where the most emission lines of metals of interest can be found. It is worth mentioning that all training dataset spectra were not intensity corrected. Therefore, no intensity correction was done on the test dataset.

In the spectra normalization step, two normalizations were tried to later adopt the best one, and those were total spectral area normalization, and standard normal variate (SNV) normalization. First one is clear, whereas SNV normalization represents a spectral normalization tool that mean centers the spectra and then divide each mean-centered intensity with its standard deviation [32]:

$$I_{\text{new}} = \frac{I_{\text{old}} - I_{\text{mean}}}{\sigma} \qquad (1)$$

where I_{new} is the new intensity, I_{old} is the intensity that is being mean centered, I_{mean} is the mean intensity and σ is standard deviation of intensities. Besides these two, MinMax data scaling was also tried. MinMax scaling represents procedure where for each feature, we scale the values according to the formula below, so we have feature values between zero and one:

$$I_{\text{scaled}} = \frac{I - I_{\text{min}}}{I_{\text{max}} - I_{\text{min}}} \qquad (2)$$

Proceeding further, we detected and ejected outliers with the help of Isolation Forest algorithm implemented in sci-kit learn. After the outliers have been removed, we fitted Random Forest regressor with aim to find features that give the most contribution to the final result. To achieve this, we actually trained four random forest models, one for each element, to have features that contribute to the each element prediction separately. Feature importances were calculated within random forest algorithm by usage of Gini importance. The higher the value, the more valuable this feature is to the final prediction.

3.2 Hyperparameters tuning and model selection

To find the optimal parameters of the model, we performed GridSearch cross-validation. This validation technique takes the given model parameters and initializes the model of interest with these parameters, splits provided dataset into training and test datasets, fits the model and reports the accuracy of the model through R^2 coefficient. This procedure is done five times in a row for each set of model parameters, where, at the end, for each model algorithm reports the best performance and with which parameters they were obtained. Used metrics to assess the predictive performance of the models were coefficient of determination R^2 and root-mean-square error (RMSE). With optimal parameters
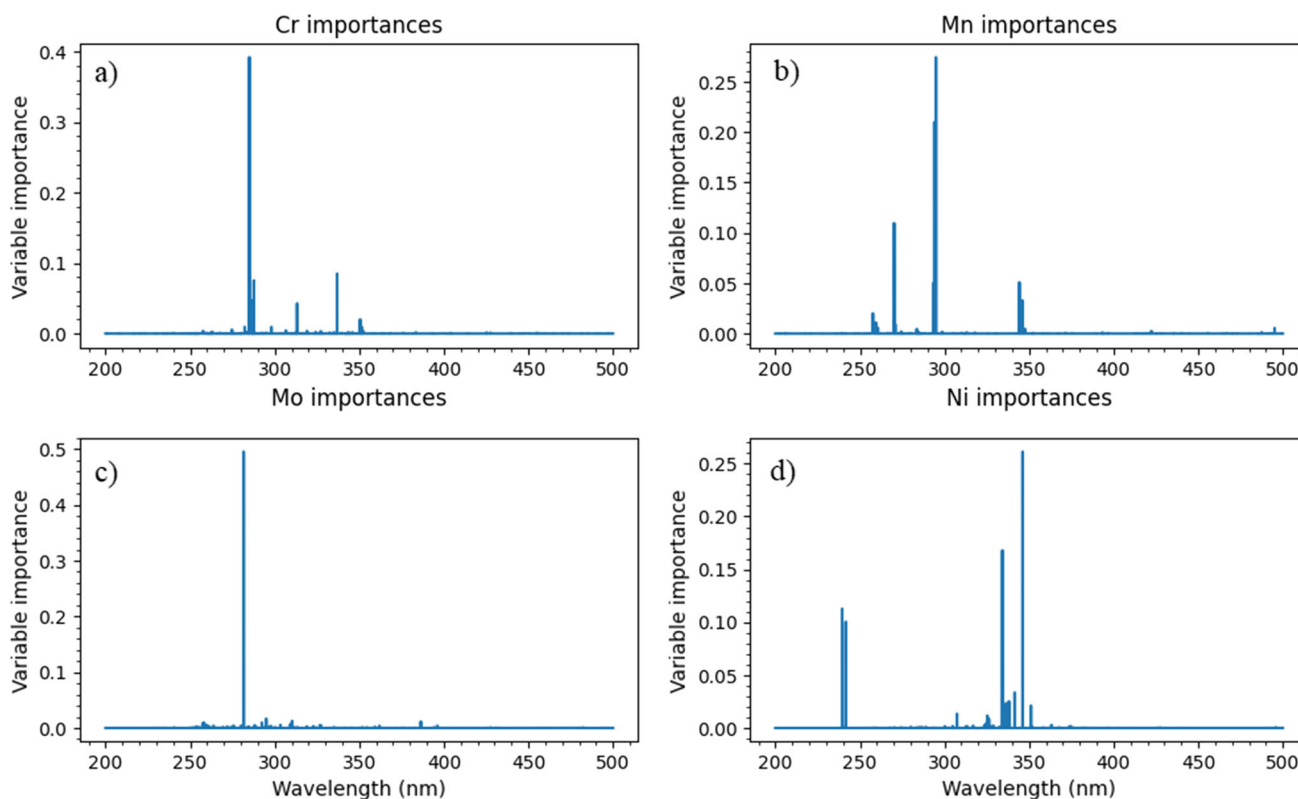
Fig. 3 Results for feature importance analysis for Cr and Mn (**a**, **b**) and Mo and Ni (**c**, **d**)

found, we proceeded to final model training and finally the prediction of steel samples composition.

4 Results

The results of feature importance analysis is given in Fig. 3. It is evident that the algorithm successfully recognized and selected persistent line of Mo II at 281.61 nm (see Fig. 3c)). Also in Fig. 3d), lines of Ni II at 239.45 nm and 241.6 nm were successfully identified. Great importance was also given for Cr II lines around 285, 286, 287, 313 nm, as well as to Cr II line at 336 nm (see Fig. 3a)). Unimportant features have value of zero or close to zero, so the condition threshold was set to $10^{-4}, 2 \times 10^{-4}$ and 5×10^{-4}, while the best results were obtained for threshold 2×10^{-4}. Hence, the final dimensionality of dataset used to train the final model is given in Table 2.

In GridSearch cross-validation, parameters for random forest that were supplied to the algorithm were number of estimators (number of threes in forest) which was changed from 200 to 350 in the step of 50, and maximal depth of the individual three which was varied from none to 4. None here means that the three is going to expand until all leaves are pure. In the case of linear regression, the only parameter that could be changed is $\mathcal{L}^2$ norm penalization coefficient α, and we have chosen the values of 0.5, 0.8 and 1. For DNN, considered

architectures were ones with one, two and three hidden layers [(100), (100, 100) and (100, 150, 50)]. Numbers in parentheses represent number of neurons in each hidden layer. Activation function was ReLU (Rectified Linear Unit). Best results reported for all models were ones with MinMax scaling. For RF, best results were the ones where number of threes was equal to 350 and maximum depth that was set to none. Best results with linear regression were reached for the α parameter equal to 0.8. Finally, for DNN, architecture with three hidden layers showed best performance. After dimensionality reduction via Gini importance, resulting dataset was divided into training and test datasets, keeping 20% of the data for testing. Validation of the models was done by using R^2 metrics and RMSE, and it is given in Table 3.

With model training finished, judging by the R^2 score, best overall performance is showed by deep neural network. The prediction precision for each element goes above 0.9, whereas the predicted values in the case of RF are little less. Results for linear regression are not given, since they are significantly worse than these, thus they were omitted. Prediction on recorded test dataset was done with RF as well as with DNN, and the predicted results are summed in Fig. 4. From Fig. 4a–d, it can be seen that DNN showed good performance on all elements, while the predictions made using RF are quite good for the case of Cr, Mn and Ni, but it showed bad overall performance regarding the prediction of Mo, see Fig. 4d. There was no difference when we tried to

Table 2 Dimensionality and number of samples in training dataset used for model training

Element	Number features	Number of samples
Cr	273	1608
Mn	129	1608
Mo	317	1608
Ni	120	1608

All useful information is contained in these selected features

predict Mo concentration with all features, where unimportant features were not removed.

5 Conclusion and future development

In this paper, the prediction of austenitic steel alloy samples was done using the random forest algorithm and deep neural network. Data preprocessing consisted of applying MinMax scaler on the raw data, followed by outliers removal with isolation forest algorithm. Feature

selection was performed by Gini importance criterion within random forest algorithm. It successfully isolated most important features, thus enabling the dimensionality reduction while keeping all the necessary information. This was preceded by final training of three models: random forest, linear regression with $\mathcal{L}^2$ norm and deep neural network. Random forest and neural network showed better predictive power than linear regression; hence, they were used as selected models for prediction of the steel alloy composition. Trained random forest model showed good predictive power for
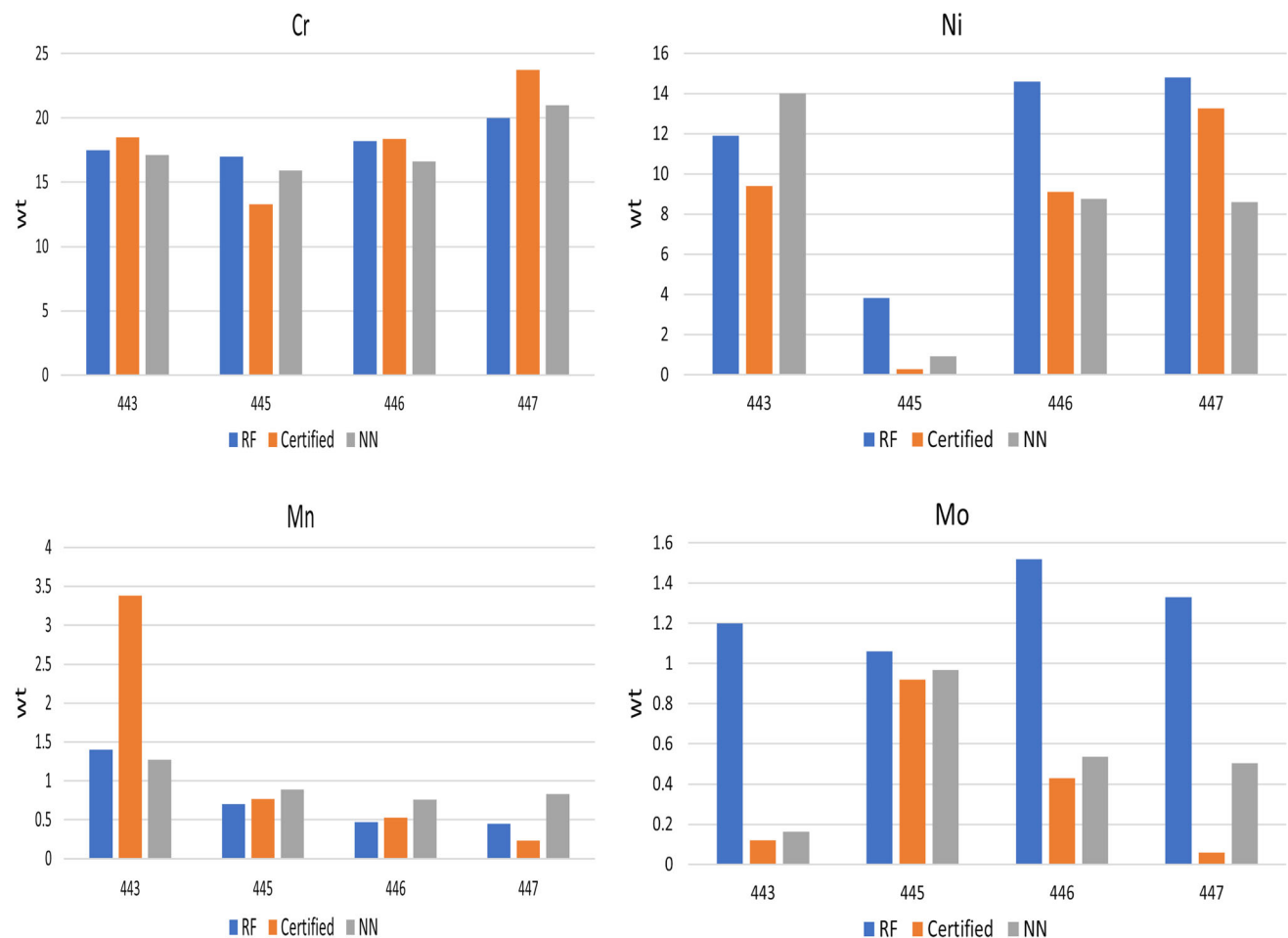

Fig. 4 RFR and DNN predicted results (denoted with RF and NN on the figure) and comparison with certified values. Numbers on x-axis denote the steel sample number given in Table 1. The figure indicates that the models learned and yielded good results in the case of Cr (**a**), Mn (**c**) and Ni (**b**), but on the other hand, RFR had rather poor performance in the case of Mo (**d**)

Table 3 R^2 and RMSE values for validation dataset

Element	R^2_{RFR}	R^2_{DNN}	$\mathrm{RMSE}_{\mathrm{RF}}$	$\mathrm{RMSE}_{\mathrm{NN}}$
Cr	0.88	0.97	3.68	1.84
Mn	0.89	0.93	0.397	0.313
Mo	0.85	0.96	0.511	0.263
Ni	0.97	0.98	1.77	1.21

Cr, Ni and Mn, but rather poor performance in the case of Mo. On the other hand, neural network showed good overall predictive power. Nevertheless, random forest algorithm, combined with the data preprocessing techniques, shows a good potential for application in austenitic steel alloy composition prediction, which was also confirmed by results from other authors. For future work, we tend to write a better feature extraction software that should improve the feature selection and hence the predictive power of a used regressor. Also, good overall results are obtained, although the training and test datasets were not intensity corrected. This work shows how useful it can be, to build a unique steel dataset for later usage by different authors, as they not need to every time record their own datasets. These results can be further improved, if one performs calibration transfer, as these spectra were recorded on different instruments. Here, this was not performed as we have not had any identical standard that was used on primary instrument.

Author contributions

IT involved in methodology, software and original draft writing; MI involved in conceptualization, supervision and original draft editing.

Data Availability Statement This manuscript has associated data in a data repository. [Authors' comment: As we recorded the spectra of only four steel standard samples, we are of the opinion that scientific community will not have much use of this data, hence it is decided not to deposit it.]

Declarations

Conflict of interest Hereby, we want to state that these results should not be in any case compared to ones obtained by various authors at the LIBS 2022 conference benchmarking competition, where this training dataset was first used. Best results from that competition will be published separately in special issue of Spectrochimica Acta B, and results from this paper have nothing to do with that competition. This paper was an extension of the work presented at SPIG 2022 conference.

References

1. R.J. Konings, R.E. Stoller (eds.), Comprehensive Nuclear Materials. (Elsevier, Amsterdam, 2020)
2. N. Inoue, A. Komori, H. Hayashi, H. Yonezu, M. Iima, R. Sakamoto, Y. Kubota, A. Sagara, K. Akaishi, N. Noda, N. Ohyabu, O. Motojima, Design and construction of the LHD plasma vacuum vessel. Fus. Eng. Des. **41**(1), 331–336 (1998). https://doi.org/10.1016/S0920-3796(98)00248-8
3. V. Alimov, M. Yajima, S. Masuzaki, M. Tokitani, Analysis of mixed-material layers deposited on the toroidal array probes during the FY 2012 LHD plasma campaign. Fus. Eng. Des. **147**, 111228 (2019). https://doi.org/10.1016/j.fusengdes.2019.06.001
4. D.C.L.J. Radziemski, *Spectrochemical analysis using laser plasma excitation*, edited by D.C.L.J. Radziemski (Marcel Dekker Inc, New York, 1989)
5. E. Tognoni, G. Cristoforetti, S. Legnaioli, V. Palleschi, Calibration-free laser-induced breakdown spectroscopy: state of the art. Spectrochim. Acta Part B At. Spectrosc. **65**(1), 1–14 (2010). https://doi.org/10.1016/j.sab.2009.11.006
6. C.A.M. Ramirez, M. Greenop, L. Ashton, I. ur Rehman, Applications of machine learning in spectroscopy. Appl. Spectrosc. Rev **56**(810), 733–763 (2021). https://doi.org/10.1080/05704928.2020.1859525
7. N.M. Ralbovsky, I.K. Lednev, Towards development of a novel universal medical diagnostic method: Raman spectroscopy and machine learning. Chem. Soc. Rev. **49**, 7428–7453 (2020). https://doi.org/10.1039/D0CS01019G
8. W. Fu, W.S. Hopkins, Applying machine learning to vibrational spectroscopy. J. Phys. Chem. A **122**, 167–171 (2017). https://doi.org/10.1021/acs.jpca.7b10303
9. H. Park, J.-H. Son, Machine learning techniques for THz imaging and time-domain spectroscopy. Sensors (2021). https://doi.org/10.3390/s21041186
10. M. Koubiti, M. Kerebel, Application of deep learning to spectroscopic features of the Balmer-Alpha line for hydrogen isotopic ratio determination in tokamaks. Appl. Sci. (2022). https://doi.org/10.3390/app12199891
11. T. Chen, T. Zhang, H. Li, Applications of laser-induced breakdown spectroscopy (LIBS) combined with machine

learning in geochemical and environmental resources exploration. Trends Anal. Chem. **133**, 116113 (2020). https://doi.org/10.1016/j.trac.2020.116113

12. C. Sun, Y. Tilan, L. Gao et al., Machine learning allows calibration models to predict trace element concentration in soils with generalized LIBS spectra. Sci. Rep. **9**, 11363 (2019). https://doi.org/10.1038/s41598-019-47751-y

13. X. Zhang, F. Zhang, H.-T. Kung, P. Shi, A. Yushanjiang, S. Zhu, Estimation of the Fe and Cu contents of the surface water in the Ebinur lake basin based on LIBS and a machine learning algorithm. Int. J. Environ. Res. Public Health (2018). https://doi.org/10.3390/ijerph15112390

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

15. Y. Tian, Q. Chen, Y. Lin, Y. Lu, Y. Li, H. Lin, Quantitative determination of phosphorus in seafood using laser-induced breakdown spectroscopy combined with machine learning. Spectrochim. Acta Part B At. Spectrosc. **175**, 106027 (2021). https://doi.org/10.1016/j.sab.2020.106027

16. M.S. Babu, T. Imai, R. Sarathi, Classification of aged epoxy micro-nanocomposites through PCA- and ANN-adopted LIBS analysis. IEEE Trans. Plasma Sci. **49**(3), 1088–1096 (2021). https://doi.org/10.1109/TPS.2021.3061410

17. X. Cui, Q. Wang, Y. Zhao et al., Laser-induced breakdown spectroscopy (LIBS) for classification of wood species integrated with artificial neural network (ANN). Appl. Phys. B **125**, 12556 (2019). https://doi.org/10.1007/s00340-019-7166-3

18. R. Junjuri, M.K. Gundawar, A low-cost LIBS detection system combined with chemometrics for rapid identification of plastic waste. Waste Manag. **117**, 48–57 (2020). https://doi.org/10.1016/j.wasman.2020.07.046

19. 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**, 106183 (2021). https://doi.org/10.1016/j.sab.2021.106183

20. F. Poggialini, B. Campanella, S. Legnaioli, S. Pagnotta, S. Raneri, V. Palleschi, Improvement of the performances of a commercial hand-held laser-induced breakdown spectroscopy instrument for steel analysis using multiple artificial neural networks. Rev. Sci. Instrum. **91**(7), 073111 (2020). https://doi.org/10.1063/5.0012669

21. H. Tang, T. Zhang, X. Yang, H. Li, Classification of different types of slag samples by laser-induced breakdown spectroscopy (LIBS) coupled with random forest based on variable importance (VIRF). J. Anal. At. Spectrom. **32**, 2194–2199 (2017). https://doi.org/10.1039/C7JA00231A

22. F. Ruan, J. Qi, C. Yan, H. Tang, T. Zhang, H. Li, Quantitative detection of harmful elements in alloy steel by LIBS technique and sequential backward selection-random forest (SBS-RF). Anal. Methods **7**, 9171–9176 (2015). https://doi.org/10.1039/C5AY02208H

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

24. G. Yang et al., The basicity analysis of sintered ore using laser-induced breakdown spectroscopy (LIBS) combined with random forest regression (RFR). Anal. Methods **9**, 5365–5370 (2017). https://doi.org/10.1039/C7AY01389B

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

26. T. Zhang et al., A novel approach for the quantitative analysis of multiple elements in steel based on laser-induced breakdown spectroscopy (LIBS) and random forest regression (RFR). J. Anal. At. Spectrom. **29**, 2323 (2014). https://doi.org/10.1039/c4ja00217b

27. Y. Zhang, C. Sun, L. Gao, Z. Yue, S. Shabbir, W. Xu, M. Wu, J. Yu, Determination of minor metal elements in steel using laser-induced breakdown spectroscopy combined with machine learning algorithms. Spectrochim. Acta Part B At. Spectrosc. **166**, 105802 (2020). https://doi.org/10.1016/j.sab.2020.105802

28. K. Liu et al., Quantitative analysis of toxic elements in polypropylene (PP) via laser-induced breakdown spectroscopy (LIBS) coupled with random forest regression based on variable importance (VI-RFR). Anal. Methods **11**, 4769 (2019). https://doi.org/10.1039/c9ay01796h

29. K. Wei, Q. Wang, G. Teng, X. Xu, Z. Zhao, G. Chen, Application of laser-induced breakdown spectroscopy combined with chemometrics for identification of penicillin manufacturers. Appl. Sci. (2022). https://doi.org/10.3390/app12104981

30. X. Jin, G. Yang, X. Sun, D. Qu, S. Li, G. Chen, C. Li, D. Tian, L. Yao, Discrimination of rocks by laser-induced breakdown spectroscopy combined with random forest (RF). J. Anal. At. Spectrom. **38**, 243–252 (2023). https://doi.org/10.1039/D2JA00290F

31. E. Kepes. (2022) LIBS 2022 quantification contest. https://figshare.com/projects/LIBS2022_Quantification_Contest/142250

32. T.W. Randolph, Scale-based normalization of spectral data. Cancer Biomark. **2**, 135–144 (2006). https://doi.org/10.3233/CBM-2006-23-405