JAAS

PAPER

View Article Online
View Journal | View Issue

Cite this: *J. Anal. At. Spectrom.*, 2025, **40**, 2526

Towards a chemical fingerprint of graphite by laser-induced breakdown spectroscopy†

Róbert Arató, *[a] Derrick Quarles, Jr, [b] Gabriella Obbágy, [a] Zsolt Dallos, [a] Miklós Arató, [c] Phillip Gopon [a] and Frank Melcher [a]

Graphite is a critical raw material for sustainable energy technologies, and establishing its traceability is crucial for ensuring responsible sourcing in the future. This study presents maps acquired on a comprehensive set of natural graphite concentrates *via* Laser-induced Breakdown Spectroscopy (LIBS). LIBS generates multi-elemental data at an unprecedented speed even from samples with non-ideal ablation characteristics, such as pressed graphite pellets. The generated data is used for constructing elemental maps to shed light on the chemical distribution of elements as well as for multivariate classification. Natural graphite concentrates exhibit inhomogeneous chemical composition. As such, the graphite concentrate LIBS-fingerprint is a heterogeneous mixture of LIBS signals from pure graphite and mineral impurities, which either represent crystal intergrowth with graphite, or they are adsorbed on graphite flakes as a result of natural or artificial processes. The observed chemical heterogeneity serves as a prominent fingerprint of individual deposits, although the heterogeneity is also omnipresent between different samples of the same deposit. The generated multivariate dataset is well suited for multivariate data analysis. Random forest classifiers show a robust performance across a broad range of hyperparameters, achieving over 90% classification accuracy. The heterogeneity of the concentrates presents a significant challenge for classification, regardless of the analytical and classification approach used. The addition of chemically different samples to the same classification group (*i.e.*, graphite deposit) does not necessarily hinder correct classification and renders the routine application of the method possible.

Received 10th February 2025
Accepted 3rd July 2025

DOI: 10.1039/d5ja00053j

rsc.li/jaas

Introduction

Natural graphite is classified as both a critical and strategic raw material in major economies,[1] serving as a vital component in battery anodes and finding broad application in the refractories, lubricants, and foundry industries. Currently, *ca.* half of the global graphite production is attributed to natural graphite, while the global share of natural graphite in battery production is over 80%.[2] Natural graphite is mostly mined in China, while other major producing countries include Brazil, Madagascar, Mozambique, Norway and Ukraine.[3] Due to growing energy storage demand and transitions in the energy sector, graphite demand is expected to increase by a factor of two by 2050.[4] While synthetic graphite has notable advantages over natural graphite in terms of purity, natural graphite is expected to gain a larger market share due to its smaller environmental footprint, lower cost and excellent processability.[5] The EU's Critical Raw Materials Act reinforces the importance of traceability and sustainable sourcing by setting binding targets for imports from a single third country and domestic extraction of critical raw materials—including natural graphite—to secure resilient and transparent supply chains. In order to ensure the stable supply and responsible sourcing of natural graphite to meet international sustainability, transparency and society-related requirements, it is necessary to enable its traceability.[6] However, the practical implementation of such traceability mechanisms remains undefined, and techniques to accurately determine material provenance are still under development. The EU-funded MaDiTrace (Material & Digital Traceability for CRM Certification) project aims at reinforcing the transparency, reliability and sustainability of critical raw material supply chains, with a special focus on key commodities for battery and magnet production, including natural graphite.

Material fingerprinting refers to the use of analytical techniques to capture unique, intrinsic chemical or physical characteristics of a material, enabling its identification and traceability. For a technique/method to be routinely used in material traceability based on intrinsic material fingerprint, it needs to fulfil most/all of the following criteria: (i) high

[a]*Montanuniversität Leoben, Chair of Geology and Economic Geology, Peter-Tunner-Straße 5, 8700 Leoben, Austria. E-mail: robert.arato@unileoben.ac.at*
[b]*Elemental Scientific, 7277 World Communications Drive, Omaha, NE 68122, USA*
[c]*Department of Probability Theory and Statistics, Eötvös Loránd University, Pázmány Péter Sétány 1/C, 1117 Budapest, Hungary*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d5ja00053j

differentiating power (ii) simple usage (iii) high sample throughput (iv) simple accessibility and (v) low cost. For (i), multi-parameter output and a large number of repeated measurements (*e.g.*, spectrometry data) provides notable advantages over the very precise measurement of few parameters (*e.g.*, isotopic ratios). (ii) Includes the complexity of sample preparation and analysis, which is inherently interconnected with (iii), where the time needed for sample preparation and analysis is critical. (ii) Is again connected to (iv) and (v), where the versatility and availability of different instrumental setups and trained personnel is decisive.

Due to its chemical and thermal resistance, the chemical analysis of graphite is particularly challenging. Complete dissolution of graphite is nearly impossible under standard laboratory conditions, and *in situ* analyses (*e.g.*, by LA-ICP-MS) are challenging due to its flakey/brittle structure. Accordingly, the knowledge of the chemical composition of graphite is currently limited. There is a need to determine which impurities—such as other mineral phases, chemical substitutions, or adsorbed elements—are present in natural graphite concentrates and how these are spatially distributed. Key questions include the nature and origin of these impurities and how effectively they can be used to fingerprint graphite deposits for traceability. Currently, no routine methodology exists to distinguish between natural graphite deposits and address these needs. The present study takes a step toward filling this knowledge gap.

In this study, we present laser induced breakdown spectroscopy (LIBS) data, acquired on a series of pressed natural graphite concentrate pellets from important graphite producers worldwide. We apply LIBS-mapping on these pellets, to understand the spatial and multivariate distribution of elements.

The chemical heterogeneity of graphite concentrates is evaluated by considering both local elemental anomalies and universally present elements. Given that the samples originate from natural sources but may have been altered during processing steps such as grinding and flotation, potential sources of chemical inhomogeneity are also examined. Individual elemental maps are used to assess the spatial distribution and grouping of elements within the concentrates. Using this comprehensive dataset, strategies for extracting the most informative sections of the acquired data are presented, and the challenges associated with within-deposit heterogeneity are discussed.

Table 1 Samples analyzed and their main characteristics. Note that various number of samples were obtained from individual deposits

Sample code	Country	Region/city (deposit/s)	Type	Processing method	Metamorphic facies[e]	Ref.
3B	Brazil	Minas Gerais (3 deposits[a])	Flake	Grinding & flotation	Amp–gran	7
21C	Brazil	Minas Gerais (3 deposits[a])	Flake	Grinding & flotation	Amp–gran	
21N	Brazil	Minas Gerais (3 deposits[a])	Flake	Grinding & flotation	Amp–gran	
22C	Brazil	Minas Gerais (3 deposits[a])	Flake	Grinding & flotation	Amp–gran	
6B	China	Inner Mongolia	Flake	Grinding & flotation + ct[d]	Granulite	8
7B	China	Heilongjian	Flake	Grinding & flotation	Granulite	
4A	China	Hunan (Lutang deposit)	mc[c]	Grinding	Contact	9
23A	China	Hunan (Lutang deposit)	mc	Grinding	Contact	
21A	China	Shandong	Flake	Grinding & flotation	Gsch–amp–gran	10
22A	China	Shandong	Flake	Grinding & flotation	Gsch–amp–gran	
13B	Germany	Passau (Kropfmühl)	Flake	Grinding & flotation	Amp–gran	11
5A	Korea	Uncertain	mc	Grinding	Contact	
9B	Korea	Uncertain	mc	Grinding & flotation + ct	Contact	
11B	Madagascar	Brickaville (2 deposits[b])	Flake	Grinding & flotation	Granulite	12
21D	Madagascar	Brickaville (2 deposits[b])	Flake	Grinding & flotation	Granulite	
22D	Madagascar	Brickaville (2 deposits[b])	Flake	Grinding & flotation	Granulite	
23D	Madagascar	Brickaville (2 deposits[b])	Flake	Grinding & flotation	Granulite	
1B	Mozambique	Cabo Delgado (Balama)	Flake	Grinding & flotation	Amp–gran	13
21B	Mozambique	Cabo Delgado (Balama)	Flake	Grinding & flotation	Amp–gran	
21J	Mozambique	Cabo Delgado (Balama)	Flake	Grinding & flotation	Amp–gran	
22B	Mozambique	Cabo Delgado (Balama)	Flake	Grinding & flotation	Amp–gran	
23B	Mozambique	Cabo Delgado (Balama)	Flake	Grinding & flotation	Amp–gran	
23E	Namibia	Karas (Aukam)	Vein	Grinding & flotation	Hydrothermal	14
2B	Norway	Skaland (Traelen)	Flake	Grinding & flotation	Amphibolite	15
15B	Norway	Skaland (Traelen)	Flake	Grinding & flotation	Amphibolite	
22G	Norway	Skaland (Traelen)	Flake	Grinding & flotation	Amphibolite	
23G	Norway	Skaland (Traelen)	Flake	Grinding & flotation	Amphibolite	
5B	Russia	Chelyabinsk (Taiginka)	Flake	Grinding & flotation	Amp–gran	16
21E	Russia	Chelyabinsk (Taiginka)	Flake	Grinding & flotation	Amp–gran	
22E	Russia	Chelyabinsk (Taiginka)	Flake	Grinding & flotation	Amp–gran	
4B	Ukraine	Kirovograd (Zavallia)	Flake	Grinding & flotation	Granulite	17
22F	Ukraine	Kirovograd (Zavallia)	Flake	Grinding & flotation	Granulite	
23F	Ukraine	Kirovograd (Zavallia)	Flake	Grinding & flotation	Granulite	

[a] Pedra Azul, Itapecerica, Salto da Divisa. [b] Vatomina & Sahamamy Sahasoa. [c] Microcrystalline. [d] Chemical treatment. [e] Amp = amphibolite, gran = granulite, gsch = greenschist; ref = relevant publication about the deposit or region.

Experimental

Samples

Commercially available graphite concentrates were obtained from various sources for this study (Table 1). The majority of the graphite from the studied deposits, are of organic origin and classified as flake-type graphite. Most of the concentrates were obtained from the raw ore after grinding and flotation. Depending on availability, up to five samples were collected from the same deposit. We also obtained materials from different years of production, to represent different parts of the same deposit or changes in processing. It must be pointed out that for these samples, there is no information over the homogeneity and representativity of the samples. On-site mixing from different mines within the same deposit or even from different deposits cannot be excluded. As such, the goal of our study is not the construction of a comprehensive geochemical dataset for individual deposits, but rather a proof of concept.

Samples of several 100 grams were obtained from each deposit. These were thoroughly mixed, and *ca.* 250 mg of flakes were extracted from different parts of the bags by a small spatula and loaded into a conventional hydraulic pellet press without a binding matter. The 13 mm dies were loaded by 4 tons for two minutes for each pellet. The pellets were then mounted on glass slides, carefully levelled with the top of the sample drawer and placed in the laser-ablation chamber. Altogether, 33 pressed natural flake graphite pellets were analyzed in the study.

LIBS-mapping

LIBS is a fast and efficient spectroscopic technique for detecting a series of elements simultaneously in various types of materials.[18,19] In LIBS analysis, the incident laser creates a high-temperature plasma environment locally, which brings the constituting elements' electrons to an excited state. In a fraction of a second after laser incidence, the temperature drops, which results in plasma breakdown and the transition of electrons back to their normal state. The element-specific energy difference between excited and normal state are thereby emitted and can be detected. Notably, LIBS is extremely sensitive to light elements, and records a full energy spectrum at every laser shot. LIBS has been successfully applied to biological samples,[20] and is starting to be widely applied to geological materials as well. Among others, LIBS has been used for mineral identification,[21,22] classification of Li-bearing pegmatites[23] and the detection of minor compounds in complex mineral mixtures.[24] Its capabilities were also demonstrated at a broad range of space applications,[25] the on-site analysis of Li-bearing tailing-slurries,[26] on fluorine distribution in shark teeth[27] and the elemental mapping of organic-rich shales,[28] just to name a few examples. The versatility of the LIBS technique and its successful application to such a broad range of materials and research questions, makes it a promising candidate for studying such a notoriously-hard-to-analyze material as graphite.

In this study, we present results collected with an ESLumen LIBS coupled to an ImageGEO193 laser-ablation system, comprising a 193 nm excimer laser and a two-volume (TwoVol3) ablation chamber (Fig. 1). The laser is delivered to the sample through an XYR beam aperture and for LIBS analysis, the emitted light from the laser-induced plasma is collected by an optical fiber inserted directly into the laser-ablation chamber (see Manard *et al.*, 2022)[27] for details. The fiber optics consist of a 5-channel spectrometer with fixed-grating Czerny-Turner design, covering a spectral range from 188 nm to 1099 nm *via* a Complementary Metal-Oxide Semiconductor (CMOS) detector.

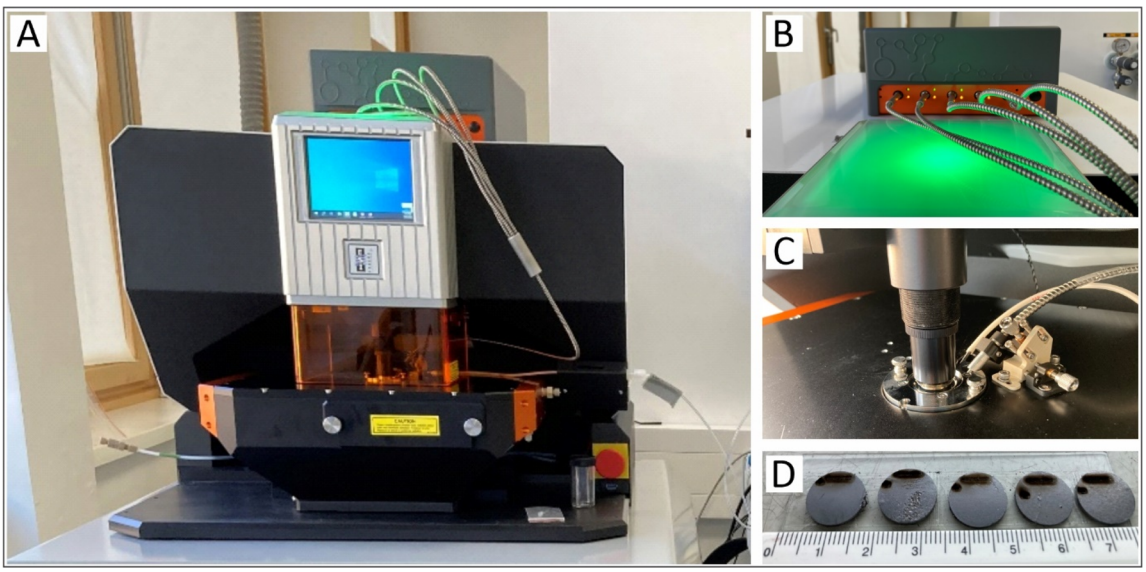

Fig. 1 (A) ImageGEO193 laser-ablation system. (B) ESL Lumen LIBS unit. Each fiber is responsible for a wavelength range, adding up to a spectrum from 188 to 1099 nm. (C) LIBS optical fiber entering the laser-ablation-chamber with a horizontal and a vertical adjustment screw for fiber alignment (D) pressed pellets of graphite concentrates used for analysis.

The LIBS fiber was calibrated by using NIST SRM 612 glass for maximum intensity across the LIBS spectrum. Alignment consists of an *X*–*Y* adjustment of the LIBS fiber *via* adjustment screws (Fig. 1C) and the *Z*-position of the sample is fixed in order to always keep the same depth into the analytical cup within the ablation chamber.

The LIBS maps were acquired over an area of 1 mm × 1 mm on one to five samples from each deposit, by ablating with 100 μm × 100 μm rectangular spots and 90 μm overlap in *X* and *Y* directions, corresponding to a nominal 10 μm × 10 μm *X*–*Y* resolution. By this approach, high resolution 2D maps could be constructed, while maintaining high signal intensity due to larger spot size. 100 Hz laser frequency and *ca.* 13 J cm^{-2} laser fluence was applied in He atmosphere, flushed at a 1000 ml min^{-1} rate. The described settings resulted in a scan speed of 1000 μm s^{-1}, which resulted in the acquisition of a 1 mm × 1 mm map within 100 seconds. All measurements were acquired within a single day.

LIBS spectra were processed in the 'iolite' v4 software (Elemental Scientific Lasers).[29,30] Data processing within iolite included atomic emission line identification, intensity integration within selected spectral windows, normalization and the creation of 2D elemental images. Spectral windows for the characteristic optical emission lines of individual elements were manually set. Finally, to avoid missing important emission lines, the 'Interesting Feature Finder' was applied and further lines were added to the integration set.[31] LIBS spectra were integrated by selecting the left and right integration windows around the targeted emission lines. All integrated data was background-corrected *via* the 'rolling ball' background subtraction approach.[32]

The presented maps show raw spectra, while for data analysis, normalized spectra were used. To normalize each spectrum, each individual integrated emission line was divided by the total emission intensity over the entire spectral range between 188 nm and 1099 nm. Each map comprises all acquired data except the upper and lower 99th percentile of relative intensities. The 'Polar' colormap is used for all maps without smoothing/filter.

Data analysis

Brunnbauer *et al.* (2023)[33] listed a series of opportunities to extract useful information from LIBS datasets, many of which are relevant to our dataset and therefore this work largely follows their recommendations. All steps related to data analysis were carried out in the sklearn (scikit-learn) package of Python.[34] For classification, the random forest (RF) classifier is applied, which is a commonly used classification method in chemometrics.[35] In brief, random forest classification is based on multiple decision trees, each of which is built from a random subset of features (*i.e.*, elemental emission lines) and data.[36] Each tree starts with a root node. Internal nodes are any additional forks or branches, which represent decisions splitting from the root node and other internal nodes. Leaf nodes are the final outputs of each decision tree. As a result, each constructed tree results in a prediction for one of the classes. Predictions for the class are made on the majority of the votes on the individual trees. In this work, individual deposits represent the classes. For simplicity, sample labels include the country name and in the case of Chinese deposits also the abbreviation of the deposit/region (Table 1).

There are two general approaches when applying multivariate methods to LIBS data (see Brunnbauer *et al.*, 2023 for a review):[33] (i) classification based on selected emission lines or (ii) classification based on entire spectra. While (i) has the advantage of being computationally less intensive, (ii) is capable of capturing the variance of the entire multivariate dataset. In this study, we tested both approaches.

Multivariate classification was applied on the same dataset, as is used for constructing the elemental maps. Raw intensities (variables) normalized to the total intensity were considered in each shot. Shot data were averaged for each line, and outliers exceeding two standard deviations from the mean were excluded. Data was not scaled, whereas labels were numerically encoded when necessary. The classifiers were trained on 80% of all individual shots, whereas 20% of the shots were held out of training for testing classification accuracy. To test for group-wise classification accuracy and potential misclassification phenomena, a confusion matrix is used.

The model's sensitivity was tested with several hyperparameters, which are important to consider when optimizing the performance of the algorithm. Most importantly, these parameters include the number of trees in the forest, the maximum depth of each tree as well as the minimum number of observations to split an internal node in the tree (min_samples_split) and the minimum number of samples to be at a leaf node (min_samples_leaf). First, randomized large intervals are searched for each hyperparameter. Then a grid search on a narrower range is based on the results of the previous step. For randomized search and grid search, a ten-fold cross-validation was applied. To test the model's robustness on unknown samples, the classifier's performance is tested by leaving out one sample from each group (deposit), and using all other samples for training and iterating through all such combinations. This approach is essentially similar to GroupKFold cross-validation in scikit-learn with the advantage that the generalization potential of the classifier is tested by leaving out entire maps from the training, thereby treating them completely unknown for the classifier.

Results

Each of the 33 acquired maps consists of 100 lines, each line comprising 100 individual shots. As such, a total of 10 000 shots per map are made, while a complete LIBS spectrum between 188 and 1098 nm is recorded at every shot (Fig. 2).

Anomalously high intensities at the edges of the acquired maps were removed to avoid eventual cross-contamination effects. As a result, over 300 000 spectra were evaluated in this work. In these spectra 32 characteristic emission lines of H, Li, B, C, O, Na, Mg, Al, Si, K, Ca, V, Cr, Mn, Fe, Cu, Zn, Rb, Sr and Ba were identified. Most concentrates showed a number of elements besides C (Fig. 2).

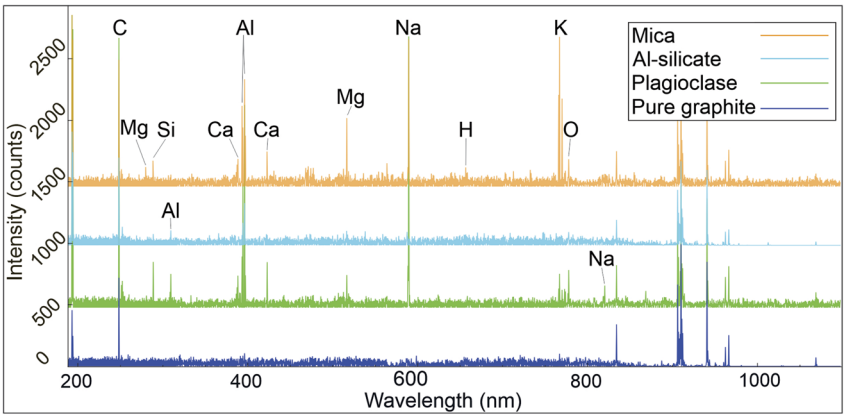

Fig. 2 Representative LIBS spectrum of a graphite concentrate containing the signal of multiple mineral species. Chemical symbols mark characteristic emission lines of abundant elements. Intensity values refer to the lowermost spectrum, while the other spectra are shifted for better visibility.

From the maps, it is evident that the elements are not uniformly distributed, and instead show anomalies of spatially correlated elements (Fig. 3). While certain elements show a very narrow intensity range close to background level, the most abundant elements are present in almost all individual spectra (Fig. 2 and 3).

Based on the spatial correlation of silicon with oxygen and other elements, it is apparent that graphite flakes are intermixed with abundant silicate minerals (Fig. 2 and 3). Exceptions in terms of mineral contents include samples 6B and 9B where silicate minerals are virtually absent.

By extracting absolute intensities underlying the elemental maps, common and highly variable elements can be detected, as they are generally characterized by high mean and % RSD (Table 2). As these measures refer to the entire population, they do not necessarily mean that the same elements can be best used for separating between different groups (*e.g.*, deposits). For this, the ratio of the between-group variance and the within-group variance (*i.e.*, 'Separation') is used.

'Separation' is calculated from the relative intensities to avoid artefacts introduced by the different total sum intensity between individual maps. From these values, the largest differences between deposits are given by their sodium, calcium and potassium contents (Fig. 4). Microcrystalline graphite samples from China (Lutang deposit, samples 4A and 23A) and Korea (sample 5A), which were not subjected to flotation, show the highest degree of impurity. The effect of chemical treatment is also evident. Samples from Inner Mongolia (China) (6B) and Korea (9B) of this category are moderate to low in Ca and Na content, while they are among the lowest in terms of potassium and aluminum. Brazilian samples show a fairly uniform composition with low calcium, sodium and potassium content, but high aluminum content. The group of samples from Mozambique and Madagascar are similar with respect to Ca and Na, but show a much larger scatter in K compared to Brazil. Notably, samples from Madagascar show the largest within-sample and between-sample variability, especially evidenced by the aluminum and potassium concentrations (Fig. 4B).

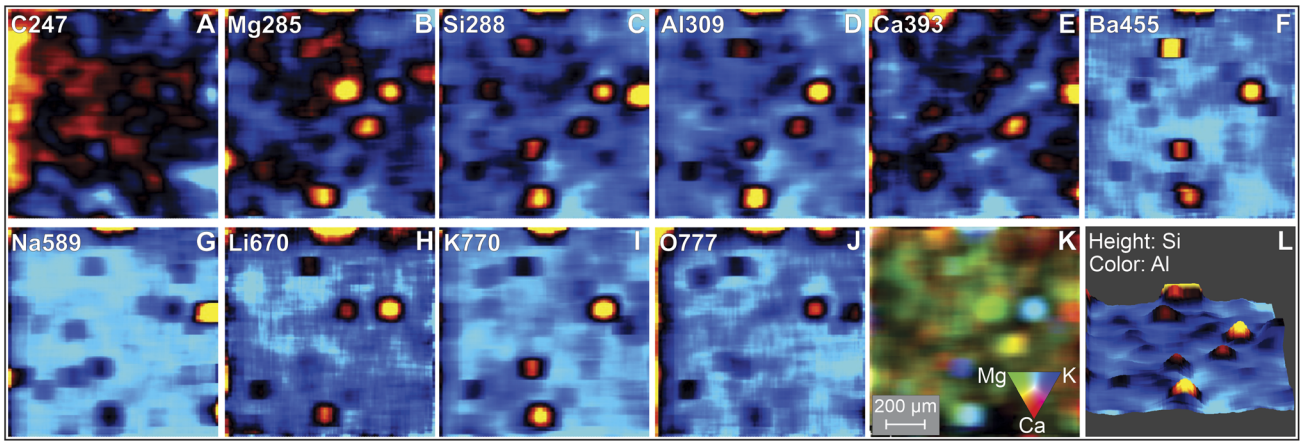

Fig. 3 (A–J) Maps of simultaneously acquired, selected wavelengths of the same area on a graphite pellet from sample 2B. Intensities are relative over the intensity range of each individual map. The corresponding element and its selected wavelength (in nm) are noted in the top left corner. (K) Overlap of three selected elements on an RGB scale to represent variability in silicate composition. (L) Three-dimensional map of the same are on the example of Si (height of the map) and Al (color of the map). Refer to ESI SI1† for all other maps.

Table 2 List of elements and wavelengths extracted from the spectra and their basic statistical properties. Units in counts. The five highest values in the last column are highlighted by bold[a]

Variable	Mean	Std	% RSD	Median	Min.	Max.	Separation
A1308	878	613	70	691	87	8994	5522
A1309	1263	1125	89	874	130	15 121	6719
A1394	1207	1376	114	697	49	19 440	8533
A1396	2105	2395	114	1238	59	32 020	**9319**
Ba455	285	206	72	270	0	11 849	1444
C247	5616	2535	45	5004	385	16 334	5554
Ca393	1299	1718	132	677	83	40 914	**15 413**
Cr357	216	89	41	203	23	2998	2732
Cr427	208	99	48	192	22	3637	1428
Cu324	102	42	41	98	0	1839	3137
Fe253	273	99	36	261	43	4537	3967
H656	426	162	38	419	0	4233	4385
K766	2752	3779	137	1183	0	74 142	**15 540**
K770	1704	2423	142	739	0	48 847	15 069
Li610	267	208	78	279	0	6835	6489
Li670	462	332	72	435	0	8114	**9195**
Mg285	1128	534	47	1012	202	19 818	4622
Mn259	712	574	81	563	119	16 115	3795
Na589	1325	2348	177	524	0	77 780	**16 043**
O777	981	623	64	841	0	17 129	2790
Rb780	277	94	34	272	0	1854	5812
Si288	990	921	93	675	31	18 852	4860
Sr407	194	86	44	181	19	4433	6061
V437	222	75	34	214	30	5334	2387
Zn202	676	135	20	665	219	1652	6642
Zn206	188	59	31	182	23	993	4365
B249	269	90	33	258	49	3159	4027
Ca854	666	516	77	628	0	19 300	5710
Ca866	650	373	57	647	0	12 098	2996
Cl837	168	74	44	168	0	1400	3434
Li812	152	60	39	151	0	597	4308
Mg518	1548	1931	125	1245	314	108 185	3238
N743	136	65	48	135	0	684	2468
Na818	1045	506	48	1019	0	32 205	4936
O844	286	150	53	276	0	3785	2047
Zn213	247	70	29	239	23	815	4866
Zn330	245	65	27	240	41	1047	4841
AllLight	741 123	247 080	33	687 803	242 331	2 077 900	

[a] AllLight = total sum intensity.

Samples from Russia, Ukraine and Norway show a relative enrichment in all four elements (Na, Ca, K, Al), but due to severe within-sample and between-sample variations at each deposit, no characteristic fingerprint is described for any of the deposits solely based on four elements. This underlines the limitations of traditional discrimination plots in high-dimensional datasets,[37] and shows the necessity of applying multivariate techniques for classification.

Despite the fact that the majority of the concentrates consists of graphitic carbon (Fig. 3A), the coexistence of non-C elements in these spectra shows that other minerals are present in the concentrates (Table 3). It also has to be considered that by overlapping ablation in both *X* and *Y* direction, a 100-shot deep 3D column of each pellet is presented as 2D maps. There are several instances where the chemical signature of multiple minerals is present at the same *X*–*Y* coordinate. This is due to, for example, quartz (SiO_2) and/or K-feldspar ($KAlSi_3O_8$) underlying muscovite ($KAl_2AlSi_3O_{10}(OH)_2$) which would have a spectrum quite similar to pure muscovite. On the other hand, the set of co-detected elements helps to restrict the options to a small set of possible minerals based on straightforward exclusion criteria. The overarching majority of the minerals analyzed here (Table 3) are potassium and hydroxide-bearing silicate minerals, presumably sheet silicates. These commonly include K, Al, Si, Na, Mg, O and H as well as less commonly Ca and Li.

Discussion

Chemical heterogeneities as seen in LIBS-maps

Analytical fingerprinting does not necessarily require the detailed characterization of the fingerprinted materials, especially if single-phase/homogeneous materials are studied. The question of homogeneity and classification generalizability in

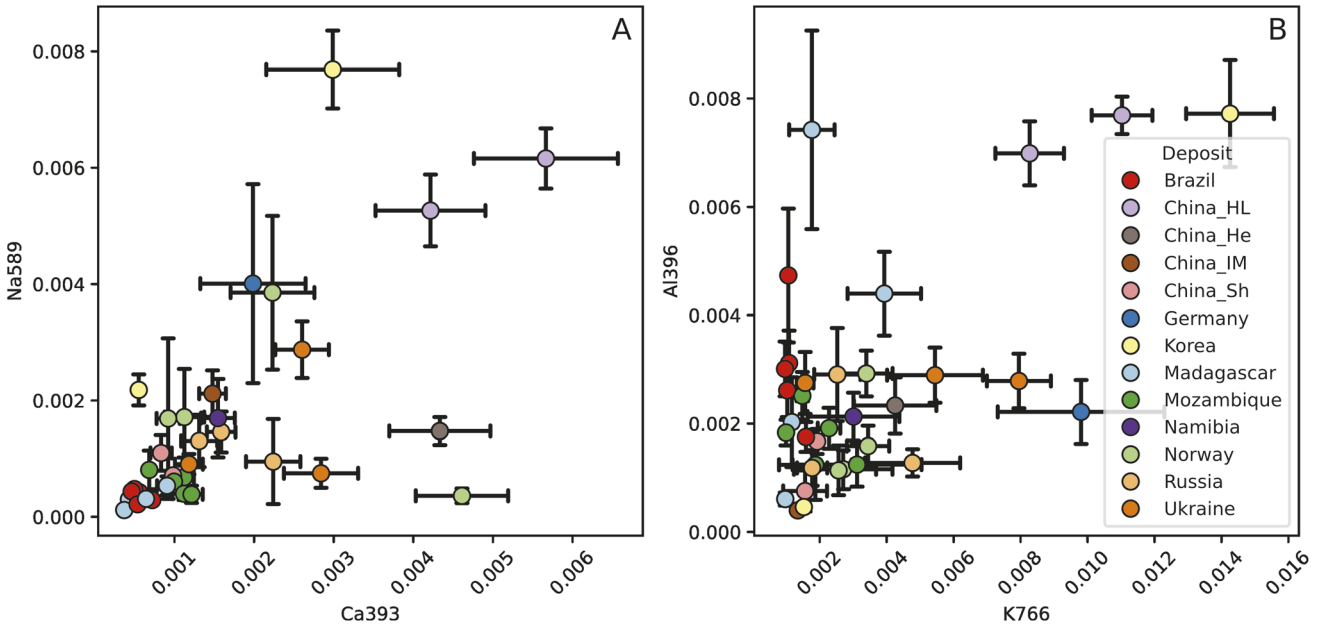

Fig. 4 Scatterplot of highly variable elements. (A) Sodium and calcium (B) aluminum and potassium. Each datapoint is calculated by averaging 100 lines, each comprising 100 individual shots. Error bars indicate 2 sigma standard deviation.

many cases is not discussed accordingly. Our graphite concentrates underscore the importance of first addressing the most fundamental question: what is the material that we are fingerprinting? Natural graphite is mined from natural environments, mainly embedded in high temperature metamorphic terranes.[38] It is then then pre-concentrated by wet grinding and flotation under a variety of different conditions.[39] Therefore, the fingerprint of natural graphite concentrates potentially includes a modified/diluted natural signature and industrial contribution at the same time. This transitional state of samples between natural and industrial products might explain why there is currently little to no literature about the chemical and mineralogical composition of natural graphite concentrates.

The multi-elemental maps of this study clearly show that certain elemental impurities are spatially restricted (Fig. 5). These mineral impurities, which were not removed from graphite during preconcentration, are characteristic for the geological environment where the sample was mined from. Table 3 shows the minerals that were identified in the various concentrates. The most common minerals include various sheet silicates, probably mica or clay minerals. Also common are feldspars and quartz. Graphite-mica intergrowths are described in the literature,[40] however, little was known about their abundance in concentrates. The observations of this study imply that the accumulation process of organic material always involves the input of anorganic material, which develops to minerals – to a great part mica – during metamorphism. Importantly, due to the large size (100 × 100 μm) of the ablation spots, we cannot assess the actual size of most ablated minerals. We can however say that there is almost no 100 × 100 μm domain, which does not contain non-carbon elements. A more detailed understanding of the mineral impurities, and as if there are also other types of chemical heterogeneities in graphite, would require a technique with higher spatial resolution (*e.g.*, transmission electron microscopy or atom probe tomography), or the separation of minerals from graphite, which is beyond the scope of this paper.

The described mineralogical observations can be used for fingerprinting purposes. Certain types of minerals, and the combination of minerals, define the 'fingerprint' of individual deposits. While K–Na–Mg-bearing micas dominate in most of the studied concentrates (Fig. 5A–D), including deposits from China, Korea, Madagascar, Mozambique, Namibia, Norway and Russia, there are marked differences in their elemental composition and accompanying minerals. K, Mg and Fe-rich mica (s.l. biotite) is far less abundant, but well-recognizable by its intense K and Mg lines as well as a series of transition metal lines below 500 nm. They predominate in the concentrates from Ukraine, but are also abundant in the samples from Norway, Germany and Korea, as well as a minor phase in Madagascar graphite. Clay minerals of the kaolin group, in turn are a uniform mineralogical fingerprint of samples from Brazil (Fig. 5E–H) with occasional higher abundance in the samples from Shandong-China (sample 21A), Madagascar, Mozambique, Namibia and Ukraine. Quartz shows high abundance only in samples from Mozambique and/or its signal is intermixed with the previously mentioned phases in other samples. The remaining minerals do not appear to be diagnostic for any deposit. Overall, the mineralogical signature of the studied concentrates is mostly dominated by sheet silicates, which are harder to separate from graphite during the grinding and flotation steps, due to their similar habit and hydraulic properties to graphite flakes.

Apart from obvious mineral impurities, elemental inhomogeneities are apparent throughout the majority of all maps. This is especially apparent in the case of aluminum, silicon, oxygen

Table 3 Characteristic minerals in the analyzed samples. Diagnostic features in individual samples are highlighted in bold

Sample code	Country	Type	Proc. method[b]	K-Na-Ca-Mg-mica	K-Mg-Fe-mica	Kln	Qz	Kfs	Pl	Cal	Other phases	Mineral amount	Peculiarity
3B	Brazil	flake	gr & fl	x	-	xxx	-	-	-	-	Sps	xxx	kaolin-gr.
21C	Brazil	flake	gr & fl	x	-	xxx	-	-	-	-	Cls	xxx	kaolin-gr.
21N	Brazil	flake	gr & fl	x	-	xxx	-	-	1		Act-Tr	xxx	kaolin-gr.
22C	Brazil	flake	gr & fl	x	-	xxx	-	-	-	-	Gln	xxx	kaolin-gr.
6B	China	flake	gr & fl + ct	-	-	-	-	-	-	-	NaOH, Ca(OH)2	-	no minerals
7B	China	flake	gr & fl	xxx	-	-	-	-	-	-	-	xx	Mg-Na-Li
4A	China	mc[a]	gr	xxx	-	-	-	-	-	-	-	xxx	Na-Ca-Mg
23A	China	mc	gr	xxx	-	x	-	-	-	-	-	xxx	Na-Ca
21A	China	flake	gr & fl	xx	-	xxx	-	-	-	-	-	xx	low int.
22A	China	flake	gr & fl	xx	-	-	-	-	-	-	-	x	low int., Na-(Li)
13B	Germany	flake	gr & fl	xxx	xx				x	x		xxx	
5A	Korea	mc	gr	xxx	xx	-	-	-	-	-	-	xxx	Na-Ca-Li
9B	Korea	mc	gr & fl + ct	-	-	-	-	-	-	-	NaOH, $Ca(OH)_2$	-	no minerals
11B	Madagascar	flake	gr & fl	xx	xx	x	-	-	1	-		xx	
21D	Madagascar	flake	gr & fl	xx	-	xx	1	-	1	1	Hbl	xx	low K
22D	Madagascar	flake	gr & fl	xxx	xx	-	-	-	-	-	-	xxx	
23D	Madagascar	flake	gr & fl	xx	-	x	-	-	-	-	-	xx	low int.
1B	Mozambique	flake	gr & fl	xx	-	-	xx	1	x	x	x	xx	Qz, low K
21B	Mozambique	flake	gr & fl	xx	-	x	xx	-	x	-	-	x	Qz
21J	Mozambique	flake	gr & fl	xx	1	xx	-	1	x	-	Ca-Al-silicate	xx	low K
22B	Mozambique	flake	gr & fl	xxx	-	x	x	-	x	-	Sps	xxx	low K
23B	Mozambique	flake	gr & fl	x	-	xx	x	-	x	-	Hbl	xxx	low K, high Mn
23E	Namibia	vein	gr & fl	xxx	-	xx	-	-	-	-	-	xxx	Na
2B	Norway	flake	gr & fl	xxx	xx	-	-	-	x	x	-	xxx	Mg
15B	Norway	flake	gr & fl	xxx	xx	-	-	-	x	x	-	xxx	Mg-Ca
22G	Norway	flake	gr & fl	xxx	x	-	-	-	x	x	-	xxx	Na-Ca
23G	Norway	flake	gr & fl	xxx	xx	-	-	1	x	-	-	xx	Na-Mg-Ca
5B	Russia	flake	gr & fl	xxx	-	x	-	-	-	-	-	xxx	Na
21E	Russia	flake	gr & fl	xxx	-	-	-	-	-	x	-	xxx	Na-Ca-Li
22E	Russia	flake	gr & fl	xxx	x	x	-	-	-	-	-	xxx	Na-Mg-Ca-(Li)
4B	Ukraine	flake	gr & fl	xx	x	xx	-	-	x	-	-	xxx	
22F	Ukraine	flake	gr & fl	xxx	xxx	x	x	-	-	1	-	xxx	Mg-Ca
23F	Ukraine	flake	gr & fl	xx	xxx	xx	x	-	-	x	Cpx	xxx	Ca-Na-Mg-Mn

[a] Microcrystalline. [b] gr = grinding, fl = flotation, ct = chemical treatment, xxx = very abundant; xx = abundant; x = multiple grains identified; 1 = one grain identified; - = not identified; Kln = kaolin-group minerals; Qz = quartz; Kfs = K-feldspar; Pl = plagioclase; Cal = calcite; Sps = spessartine; Cls = celsian; Act-Tr = actinolite-tremolite; Gln = glaucophane; Hbl = hornblende; Ca-Al-sil. = Ca-Al-silicate; Cpx = clinopyroxene; low int. = low intensity.

and potassium. Two elements, aluminum and oxygen are detectable at almost every single shot, except for samples 6B and 9B. Those two samples have been subjected to chemical treatment during ore processing and only contain Ca, Na, O and H (with only very rare traces of Si), corresponding to remnants of Na and Ca-hydroxides, commonly used for removing silicate minerals from natural graphite concentrate *via* alkali autoclave-acid leaching.[41] The absence of silicates and the presence of abundant hydroxides in those samples serve as a distinct fingerprint of chemically treated graphite and these features make them easily distinguishable from untreated products (Fig. 6A–D). On the other hand, microcrystalline samples, which were not subjected to flotation, mark the other end of the spectrum with the highest abundance of silicates (Fig. 6E–H) and an accordingly definite fingerprint.

The observed heterogeneity in individual maps contributes to the diagnostic geochemical fingerprint, capturing subtle differences even when the same minerals are present in varying quantities across different concentrates. This high degree of specificity enables detailed characterization at the sample level, in addition to distinguishing between deposits. As a result, the unique mineralogical composition and abundance reflected in each map (Table 3) allow for both inter-deposit and intra-deposit differentiation in practice.

Classification

The real power of LIBS datasets lies in the multivariate nature of the collected data and the large number of repeated measurements. Even if there is no obvious chemical difference between deposits in terms of certain elements, the entirety of the acquired dataset might be deposit-specific. To investigate this aspect, multivariate classification is applied as summarized in Fig. 7.

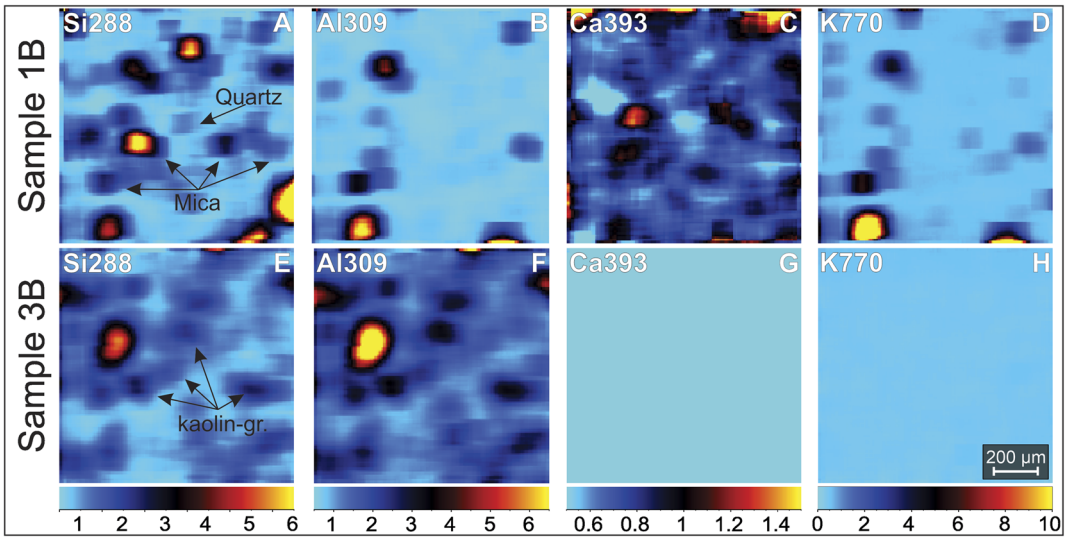

Fig. 5 (A–E) Representative elemental maps of a sample (1B) with abundant mica. (E–H) Representative elemental maps of a sample (3B) with abundant kaolin-group minerals. Units are in 1000 counts. Note the different intensity ranges.

The RF classifier shows a robust overall classification accuracy of ~93% (Fig. 8). The high classification accuracy is striking, especially considering the significant differences between individual samples within the same deposits. In order to assess the contribution of each selected elemental line (*i.e.*, feature) to the classification, feature importances were extracted. The most important features for classification are lines related to major elements, such as Ca, Al, K, Na, Si and Mg (ESI Table ST1†). Also, some minor elements, including Li, Ba, Zn, Sr and Mn contribute to the overall classification accuracy.

To ensure that the model's performance reflects actual patterns in the data rather than noise or overfitting to specific features, careful attention is paid to model optimization. Overfitting is avoided by not selecting unrealistic hyperparameters (*i.e.*, >100 trees, and <10 samples for splitting a tree) and by checking the 'area under curve,' which is used as an evaluation metric for each hyperparameter (ESI SI2†). These curves imply no overfitting. Furthermore, the performance of the classifier is robust over a broad range of hyperparameters, resulting in only a few percent reduction in the classification accuracy, if a very conservative hyperparameter-set is applied ESI2.† The effect of the number of spectra (*i.e.*, dataset size) was also tested and the results are summarized in the ESI SI3.† If entire spectra are used instead of manually selected wavelengths classification accuracy decreases notably (see details in ESI SI3†).

While the model numerically does not overfit the data (ESI SI2†), the extreme heterogeneity within and between samples, it

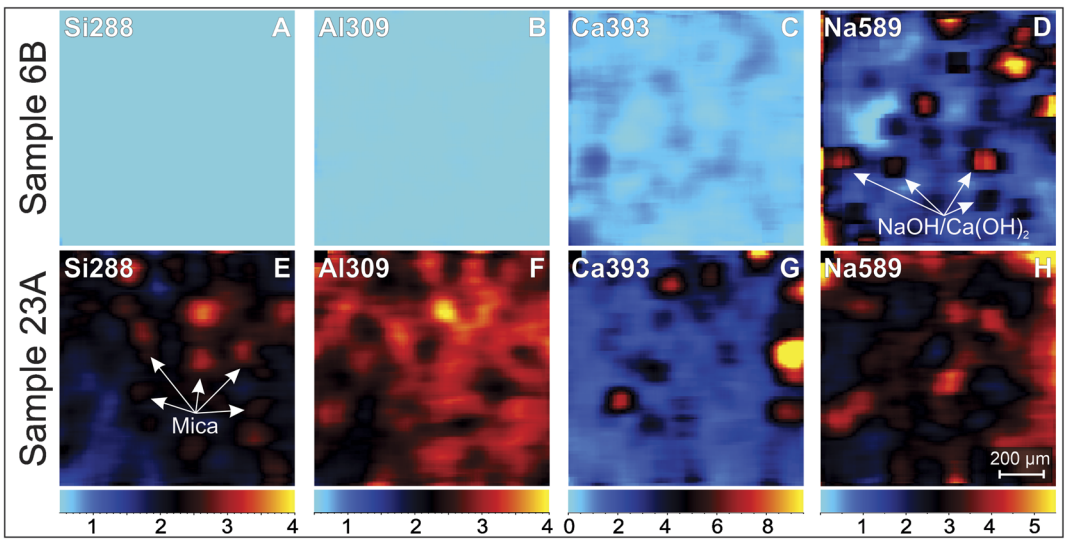

Fig. 6 (A–E) Representative elemental maps of a sample (6B), which was subjected to chemical treatment. (E–H) Representative elemental maps of a microcrystalline sample (23A) with no chemical treatment and flotation. Units are in 1000 counts.

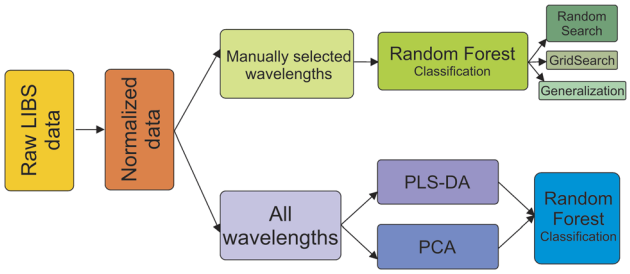

Fig. 7 Classification approach applied. Note that different approaches were tested to a multiple extent. Focus was put onto manually selected wavelengths and subsequent random forest classification. PLS-DA = partial least squares discriminant analysis; PCA = principal component analysis.

is difficult to argue that a 1 mm × 1 mm map is sufficient to classify an entire deposit. However, the observed high classification accuracy shows that the RF classifier is capable of learning the main features of the acquired multivariate LIBS dataset, independent of the representativity of the dataset.

To test the generalization potential of the model, entire samples should be left out from training and used for testing. This can only partially be done with our dataset as four (or more) samples are only available from four deposits, where the train : test ratio can be kept close to the values used in common practice, usually between 90 : 10 and 70 : 30. If only deposits represented by four or more samples are included (Fig. 9), classification accuracy reaches ~65%. This is achieved by testing all combinations where one entire sample is left out for testing from each deposit and the other samples are used for training. This result shows the severe effect if we have no control over the geological heterogeneity of the deposits and the changes in processing practices, which can both have a large impact on the chemical signature of the traded material. Low classification accuracy in the case of Madagascar is also diagnostic, as it can reflect different provenance (*e.g.*, different mine from the same deposit), changing processing practices over years, on-site mixing or within-deposit heterogeneity. The results also showcase that, in this context, the amount and versatility of the available sample material is of greater importance than the choice of the classifier and its hyperparameters. Fig. 9B demonstrates that if the underlying training data describes the test population well (*i.e.*, the highly heterogeneous Madagascar deposit is not included), a correct prediction is made with a high probability. It is also important to note that these observations are not LIBS-specific: any chemical

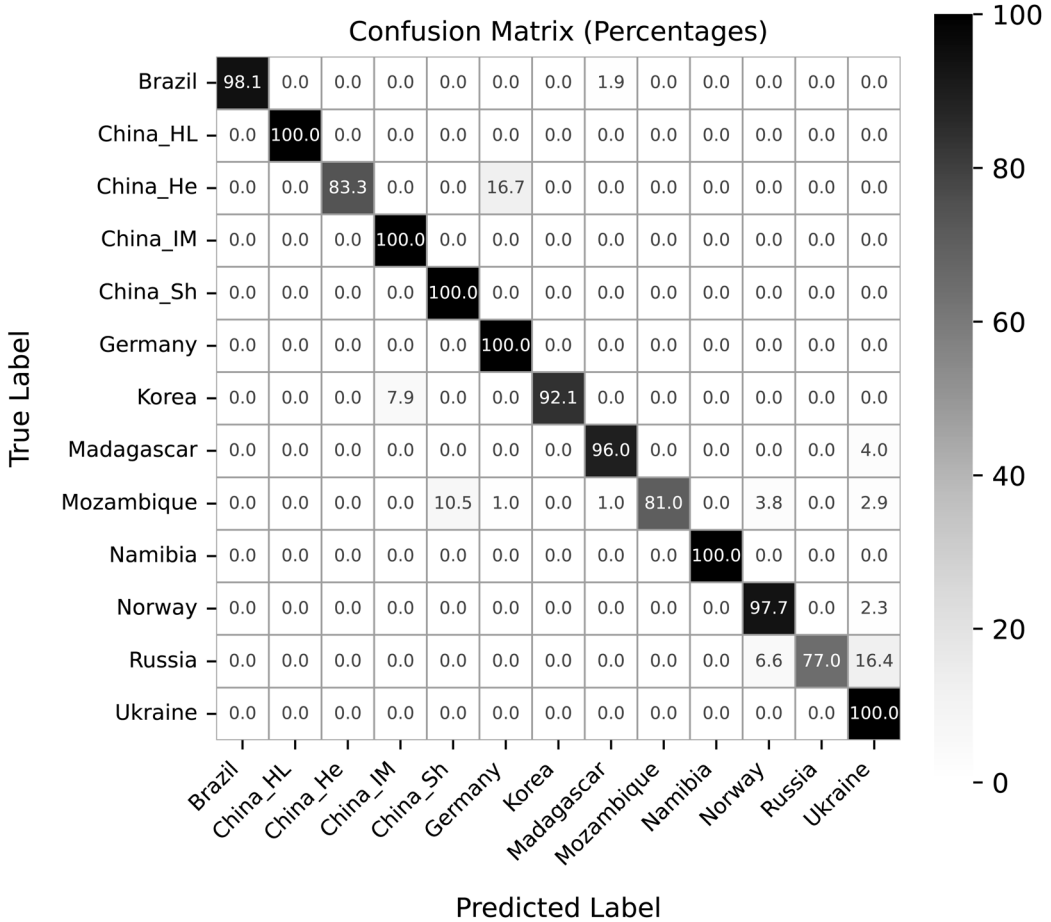

Fig. 8 Deposit-wise random forest classification accuracy based on selected emission lines. *X* axis shows the predicted labels while *Y* axis the true labels. "100" in the diagonal means perfect classification.

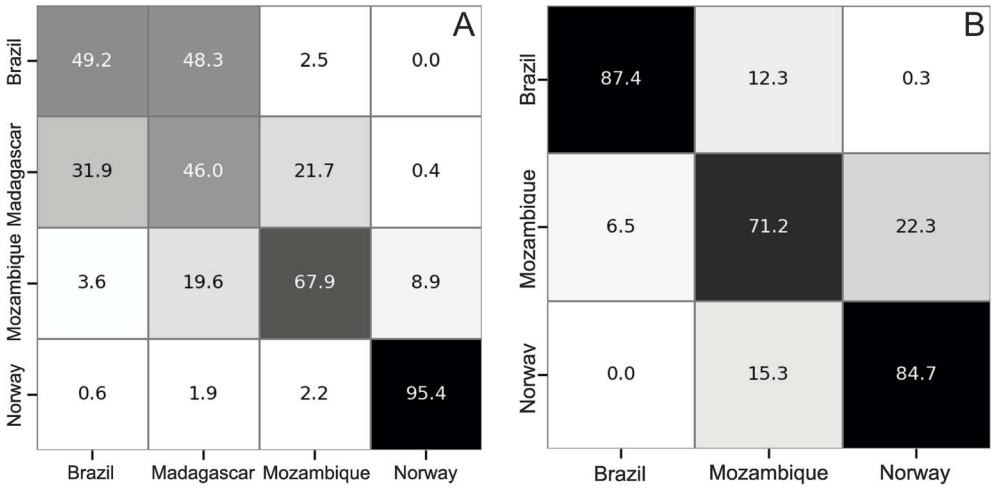

Fig. 9 Confusion matrix of graphite concentrates with at least four samples, obtained by leaving out an entire sample when training the random forest classifier and using that sample for test. The number of correctly classified spectra is seen in the diagonal in percent. (A) All combinations considered. (B) All combinations considered, excluding Madagascar samples.

fingerprinting method will face similar challenges with sample heterogeneity. Here the speed of the LIBS analysis is advantageous as it can analyze more sample within a short period of time, compared to most other techniques.

Practical aspects

The presented approach combines mineralogical interpretations of elemental maps from natural graphite concentrates and classification on the large multivariate datasets underlying those maps. This combination renders the differentiation and fingerprinting of graphite deposits possible. The extent to which the RF model can be used for global graphite traceability, without map interpretation, cannot be assessed based on the given dataset. However, even if several samples with very distinct geochemical characteristics are included in the training set, the correct classification of the graphite sample seems possible (Fig. 9). If a large-scale application is considered, a comprehensive database should be used including many individual samples from each deposit, which cover a broad range of heterogeneity in individual deposits. With that approach, the model can be trained to learn more generalized rather than individual-sample-specific patterns from the underlying dataset. If this is achieved, LIBS can be used as a routine instrument for graphite traceability, even considering quick on-site analysis *via* on-site or portable instruments.

Conclusions

One by one millimeter LIBS maps were obtained from graphite concentrates of worldwide graphite deposits, which are useful to understand the spatial distribution of elements within them. Chemical inhomogeneities are omnipresent in all studied concentrates, an aspect to be considered for all analytical methods applied for graphite analysis. Based on the spatial relationship of the detected elements, it is also obvious that most chemical impurity elements stem from silicate minerals. Due to the processing of the concentrates *via* grinding and flotation, the suite of associated minerals in the concentrates is restricted to a few mineral species with a similar hydraulic behavior to graphite, most notably sheet silicates. Despite this, there are notable differences in the mineralogical composition of the deposits. As a result, the chemical fingerprint of graphite concentrates is a signature controlled by graphite and silicate minerals. Accordingly, severe spatial heterogeneity is inherent to all deposits and concentrates. A Random Forest classifier successfully classifies graphite concentrates with over 90% probability based on individual LIBS spectra. If several samples are available from individual deposits covering a large range of variability, the correct classification of unknowns is possible, which is promising for potential routine applications in the future. This makes LIBS a strong candidate for becoming a standard tool in graphite traceability and material fingerprinting tasks.

Data availability

The data supporting this article have been included as part of the ESI.†

Author contributions

RA was responsible for conceptualization, analysis, data analysis, and manuscript writing. DQ contributed to analysis and manuscript writing. GO provided mineralogical interpretation and contributed to manuscript writing. ZD contributed to mineralogical interpretation. MA was responsible for mathematical conceptualization and data analysis. PG contributed to manuscript writing. FM contributed to conceptualization and manuscript writing. All authors reviewed and approved the final manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This study was funded by the European Union under grant agreement no. 101091502, project MaDiTrace – Material and Digital Traceability for the Certification of Critical Raw Materials, coordinated by the French Geological Survey (BRGM). MA's contribution has been supported by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the ELTE TKP 2021-NKTA-62 funding scheme. We are grateful to an anonymous donator for the greatest part of our samples. Valentina Dietrich is thanked for her great efforts in sample acquisition. Jan Schönig's help with interpreting random forests is gratefully acknowledged. RA is thankful to Joe Petrus (ESI) for his invaluable help with Iolite and to Adam Douglas (ESI) for his support with the ImageGEO193 system. We are thankful for the editorial handling of Emma Stephen and for the reviews provided by two anonymous reviewers, which have significantly improved the quality of the manuscript.

References

1 L. Weber, *Mineral. Petrol.*, 2023, **117**, 387–399.
2 J. Zhang, C. Liang and J. B. Dunn, *Environ. Sci. Technol.*, 2023, **57**, 3402–3414.
3 SCRREEN2, European Union's Horizon 2020 research and innovation programme, Grant Agreement No 958211, 2023.
4 *International Energy Agency*, IEA, Paris, 2025.
5 L. Zhao, B. Ding, X. Y. Qin, Z. Wang, W. Lv, Y. B. He, Q. H. Yang and F. Kang, *Adv. Mater.*, 2022, **34**, 2106704.
6 S. Shannak, L. Cochrane and D. Bobarykina, *Energy Res. Soc. Sci.*, 2024, **118**, 103773.
7 D. A. Miranda, A. de Oliveira Chaves, M. S. Campello and S. L. L. M. Ramos, *Int. Geol. Rev.*, 2019, **61**, 1864–1875.
8 H. F. Zhang, M. G. Zhai, M. Santosh, H. Z. Wang, L. Zhao and Z. Y. Ni, *Ore Geol. Rev.*, 2014, **63**, 478–497.
9 Z. Zheng, J. Zhang and J. Y. Huang, *Int. J. Coal Geol.*, 1996, **30**, 277–284.
10 J. Wang, J. Liu, H. Zhang, H. Zhang and Y. Li, *Geol. J.*, 2020, **55**, 3748–3769.
11 G. Rantitsch and M. Linner, *Int. J. Coal Geol.*, 2021, **248**, 103864.
12 M. J. Yang, Y. J. Di, D. Zhang, D. T. Wu, Y. Fang and X. M. Zhang, *Minerals*, 2024, **14**, 484.
13 J. Feneyrol, G. Giuliani, D. Ohnenstetter, A. E. Fallick, J. E. Martelat, P. Monié, J. Dubessy, C. Rollion-Bard, E. Le Goff, E. Malisa and A. F. M. Rakotondrazafy, *Ore Geol. Rev.*, 2013, **53**, 1–25.
14 B. F. Walter, R. J. Giebel, P. Siegfried, S. Doggart, P. Macey, D. Schiebel and J. Kolb, *J. Geochem. Explor.*, 2023, **250**, 107229.
15 H. Gautneb, J. S. Rønning, A. K. Engvik, I. H. Henderson, B. E. Larsen, J. K. Solberg, F. Ofstad, J. Gellein, H. Elvebakk and B. Davidsen, *Minerals*, 2020, **10**, 626.
16 H. P. Echtler, K. S. Ivanov, Y. L. Ronkin, L. A. Karsten, R. Hetzel and A. G. Noskov, *Tectonophysics*, 1997, **276**, 229–251.
17 K. Shakina and L. Skakun, *Rev. Mineral.*, 2010, **60**, 70–85.
18 R. S. Harmon and G. S. Senesi, *Appl. Geochem.*, 2021, **128**, 104929.
19 R. S. Harmon, *Minerals*, 2024, **14**, 731.
20 M. Bonta, J. J. Gonzalez, C. D. Quarles, R. E. Russo, B. Hegedus and A. Limbeck, *J. Anal. At. Spectrom.*, 2016, **31**, 252–258.
21 P. Janovszky, K. Jancsek, D. J. Palásti, J. Kopniczky, B. Hopp, T. M. Tóth and G. Galbács, *J. Anal. At. Spectrom.*, 2021, **36**, 813–823.
22 D. Capela, M. F. Ferreira, A. Lima, F. Dias, T. Lopes, D. Guimarães, P. A. Jorge and N. A. Silva, *Spectrochim. Acta, Part B*, 2023, **206**, 106733.
23 S. Müller and J. A. Meima, *Spectrochim. Acta, Part B*, 2022, **189**, 106370.
24 A. Nardecchia, C. Fabre, J. Cauzid, F. Pelascini, V. Motto-Ros and L. Duponchel, *Anal. Chim. Acta*, 2020, **1114**, 66–73.
25 D. A. Cremers, *Laser-Induced Breakdown Spectroscopy: Theory and Applications*, ed. D. A. Cremers and L. J. Radziemski, Springer Berlin Heidelberg, Berlin, 2014, pp. 257–291.
26 N. Khajehzadeh, O. Haavisto and L. Koresaar, *Miner. Eng.*, 2016, **98**, 101–109.
27 B. T. Manard, C. J. Hintz, C. D. Quarles Jr, W. Burns, N. A. Zirakparvar, D. R. Dunlap, T. Beiswenger, A. M. Cruz-Uribe, J. A. Petrus and C. R. Hexel, *Metallomics*, 2022, **14**, mfac050.
28 C. D. Quarles Jr, T. Miao, L. Poirier, J. J. Gonzalez and F. Lopez-Linares, *Fuels*, 2022, **3**, 353–364.
29 C. Paton, J. Hellstrom, B. Paul, J. Woodhead and H. J. Iolite, *J. Anal. At. Spectrom.*, 2011, **26**, 2508–2518.
30 B. Paul, C. Paton, A. Norris, J. Woodhead, J. Hellstrom, J. Hergt and A. Greig, *J. Anal. At. Spectrom.*, 2012, **27**, 700–706.
31 Q. Wu, C. Marina-Montes, J. O. Cáceres, J. Anzano, V. Motto-Ros and L. Duponchel, *Spectrochim. Acta, Part B*, 2022, **195**, 106508.
32 S. R. Sternberg, *Computer*, 1983, **1**, 22–34.
33 L. Brunnbauer, Z. Gajarska, H. Lohninger and A. Limbeck, *TrAC, Trends Anal. Chem.*, 2023, **159**, 116859.
34 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.
35 A. Limbeck, L. Brunnbauer, H. Lohninger, P. Pořízka, P. Modlitbová, J. Kaiser, P. Janovszky, A. Kéri and G. Galbács, *Anal. Chim. Acta*, 2021, **1147**, 72–98.
36 L. Breiman, *Mach. Learn.*, 2001, **45**, 5–32.
37 M. Frenzel, *Ore Geol. Rev.*, 2023, **153**, 105566.
38 G. Rantitsch, W. Lämmerer, E. Fisslthaler, S. Mitsche and H. Kaltenböck, *Int. J. Coal Geol.*, 2016, **159**, 48–56.
39 T. Wakamatsu and Y. Numata, *Miner. Eng.*, 1991, **4**, 975–982.
40 J. H. Ahn, M. Cho and P. R. Buseck, *Am. Mineral.*, 1999, **84**, 1967–1970.
41 Z. Xiyue, S. Hongjuan, P. Tongjiang, Z. Li and L. Bo, *Phys. Chem. Miner.*, 2024, **51**, 25.