JAAS

ROYAL SOCIETY OF CHEMISTRY

TECHNICAL NOTE

View Article Online
View Journal | View Issue

Cite this: *J. Anal. At. Spectrom.*, 2026, **41**, 593

Many-to-many transfer of LIBS spectra across multiple experimental conditions

Antoni Pawłowicz, *[ac] Jakub Vrábel, [ab] Jakub Buday, [ab] Erik Képeš, [d] Pavel Pořízka *[ab] and Jozef Kaiser [ab]

Laser-Induced Breakdown Spectroscopy (LIBS) is a powerful analytical technique widely used for extraterrestrial remote analysis. However useful, its primary limitation is its sensitivity to measurement conditions, making direct data transfer (DT) between LIBS systems with different analytical systems impractical. Addressing this challenge directly would require costly studies, extensive sample analysis or simulations of plasma formation in different atmospheres, moreover this approach would demand extensive calibration across various LIBS systems. Previous studies have demonstrated that machine learning models can facilitate DT across different instruments and conditions. However, existing approaches either rely on one-to-one spectral pairs or are limited to predefined condition pairs. We propose an alternative solution: a single machine learning model capable of many-to-many transfer across multiple conditions without requiring both one-to-one spectral representations and huge amounts of data. Our model has been trained on regolith LIBS spectra, measured in-house across two simulated atmospheres (Earth, Moon/vacuum) and with two laser energies (30 and 15 mJ). The model evaluation focuses on the Root Mean Square Error (RMSE) of predicted elemental concentrations from transformed spectra, serving as the primary metric for the transfer quality. The proposed model for which task outperforms Piecewise Direct Standardization (PDS) based baseline approaches by around 10% in terms of RMSE.

Received 19th October 2025
Accepted 12th December 2025

DOI: 10.1039/d5ja00401b

rsc.li/jaas

1 Introduction

Laser-Induced Breakdown Spectroscopy[1] (LIBS) is a versatile analytical technique used to determine the elemental qualitative and semi-quantitative composition of a sample in various phases. Due to its capability for *in situ* analysis, it is commonly used for remote sensing applications, namely on Mars rovers such as Curiosity with its ChemCam[2] LIBS instrument or SuperCam[3] on the Perseverance rover. Additionally, LIBS can quantify a wide range of major and minor elements,[4–6] it allows for real-time and stand-off sensing, multielemental detection with only a single measurement and producing high-resolution spectral images.

One of the main limitations of LIBS is its limited repeatability and a relatively high sensitivity to the sample matrix, surface topography and changes in experimental conditions.[7–9] This is largely due to the nature of the laser-induced plasma, which is a rapidly evolving, in both space and time, non-equilibrium system. In essence, LIBS is highly sensitive to atmospheric changes, and the consistent calibration is critical for quantitative analysis under a given atmosphere and sample matrix (besides other factors), as even slight variations in experimental parameters can lead to significant differences in the spectra. To overcome these limitations and to enable the effective reuse of LIBS data across varying conditions, various strategies have been developed to address instrumental and environmental variability.

In LIBS, data transfer[10] (DT) involves transforming or aligning spectral data acquired under varying conditions or from different instruments, so that it can be effectively used in different settings. This allows for the reuse of existing datasets and reduces the need for repeated calibration.

Calibration transfer[11–14] (CT), in contrast, refers to the process of adapting a calibration model developed on one (source) spectrometer for use on another (target) spectrometer, ensuring the consistent performance despite the differences in instrument response or experimental setup. This is crucial for extending model applicability across systems and reusing the knowledge learned under varying conditions.

While both DT and CT aim to address instrumental and conditional variability, they differ in approach and complexity.[15] DT typically places higher demands on the model, as it involves

[a]*CEITEC, Brno University of Technology, Purkynova 123, Brno, 612 00, Czech Republic. E-mail: pavel.porizka@ceitec.vutbr.cz*

[b]*Institute of Physical Engineering, Brno University of Technology, Technicka 2, Brno, 616 00, Czech Republic*

[c]*Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, 11/12 Gabriela Narutowicza, 80-233 Gdańsk, Poland. E-mail: antonipawlowicz123@gmail.com*

[d]*Independent Researcher, Brno, Czech Republic*

aligning or modifying entire datasets, which may include irrelevant or noisy information not essential to the target task. CT, on the other hand, focuses on transferring the predictive model itself, which may internally rely on simpler spectral transformations to account for variability between instruments. A classical approach in this context is Direct Standardization (DS), which is a DT technique that maps spectra from the source instrument to the target domain using a single global transformation. While conceptually simple, DS often fails when the spectral differences are highly nonlinear or vary across the wavelength range.

To overcome this limitation, Piecewise Direct Standardization (PDS) was proposed.[11,16,17] PDS[17] is a nonlinear technique that aims to transform a given spectrum to match a target spectrum. This is achieved by segmenting the data into smaller intervals and applying linear transformations within each segment. Although locally linear, the method is globally nonlinear due to its segmented approach.

One of the possible applications of DT is the process of transferring data from one instrument or setting to another allowing us to use a calibration curve measured in one condition for samples measured in another. This problem commonly arises when comparing spectra acquired in extraterrestrial environments, such as those measured by space rovers, with spectral libraries developed under Earth's atmospheric conditions.[16,18,19] Beyond extraterrestrial applications, DT is also necessary to address the gradual instrument deterioration, replacement of components, and variability between nominally identical instruments, such as lasers, which may introduce systematic differences despite being designed to the same specifications.[20] Addressing this challenge is crucial for enabling robust data and calibration transfer across laboratories, creating the possibility of shared spectral databases, and enabling remote missions where models must incorporate standards or reference materials not present during initial training.[14]

Therefore, our main evaluation metric is the root mean square error (RMSE) of regolith oxide content predictions based on transferred spectra. This choice reflects how well the transferred spectra can support downstream quantitative analysis tasks.

A DT model we propose is based on an auxiliary classifier variational autoencoder[21] (ACVAE), which maps spectra into a latent representation and then reconstructs them for a specified condition, the auxiliary classifier enforces reconstruction to be condition dependent, as illustrated in Fig. 1.

We demonstrate that our methodology can transform spectra across many-to-many conditions, meaning that a single model is trained to handle transfers between multiple experimental settings in both directions, rather than being restricted to a fixed pair of source and target domains. The performance is evaluated on samples excluded from training, and the proposed model is compared against baselines such as piecewise direct standardization (PDS) and piecewise direct standardization-partial least squares (PDS-PLS),[17] consistently outperforming them across all studied conditions.

Our contributions in this paper are threefold. First, we introduce a single machine learning model that enables many-to-many transfer across multiple conditions. Second, unlike earlier works (*e.g.*, Vrábel *et al.*[14]), our method does not rely on one-to-one spectral correspondences between samples from different conditions, allowing it to be applied to independently collected datasets. Finally, we demonstrate that our approach not only outperforms conventional PDS-based and other baseline methods but also maintains computational efficiency through its fully convolutional architecture.

2 Related work

DT in the context of LIBS was studied by Vrábel *et al.*,[14] where the transfer relied on simultaneously measured spectra of the same samples under different systems. This approach provided a one-to-one correspondence between source and target spectra but required specialized synchronized acquisition setups and paired datasets. In contrast, our method trains on LIBS spectra library collected independently under different conditions, enabling many-to-many transfer without paired samples.

Concretely, one-directional transfer methods such as PDS are designed for a fixed mapping (*e.g.*, from a spectrometer operating at 30 mJ in Earth atmosphere to another operating at 15 mJ). Many-to-many transfer, by contrast, enables a single model to handle several conditions simultaneously and in both directions—for instance, transferring between Earth-atmosphere and vacuum spectra without training a separate model for each case.

Earlier works on DT[16,22] mainly focus on using simple previously mentioned techniques like PDS or PDS-PLS that offer results allowing for a reliable sample quantification (see Table 3). However, as already mentioned, they only allow for a one-way transfer between two predefined conditions. A more recent DT method proposed by Ding *et al.*[23] projects the source and target spectra from their original feature spaces into a new, shared feature space where the two domains become more similar.

Zhang *et al.*[24] have also presented a generative adversarial network (GAN)-based[25] solution for DT. It aimed at transforming spectra taken from low-resolution instruments into high-resolution spectra, effectively reconstructing full-spectrum information from sparse multi-spectral measurements. This approach also significantly differs from our proposed method, as we focus primarily on the aforementioned many-to-many condition transfer. Although GAN-based methods often require very large datasets to train effectively, our approach is designed to perform well even without massive data volumes—though, as with most deep learning models, additional data could further improve performance.

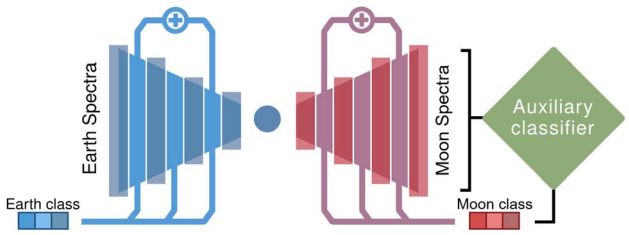

Fig. 1 Model architecture overview.

3 Methodology

In this section, we define the key components of our approach, including the data setup, model architecture, and training procedure. We then describe our customized model and the modified loss function developed to improve spectral transfer performance.

The proposed approach focuses on developing a primary-and-secondary-condition-agnostic model, capable of transferring spectra bidirectionally between any two conditions denoted as c:

$$c_i \rightleftarrows c_j \,;\, i, j \in \{(i, j) \mid i \neq j\} \tag{1}$$

In our case we consider a total of three conditions: Earth atmosphere (atm) with the laser energy set to 30 mJ, Earth atm with 15 mJ laser energy and Moon/vacuum atm with 30 mJ laser energy. Given the clear physical differences between these conditions, the resulting spectra vary substantially in the total emissivity, line width, and profile, as well as in the presence and relative intensities of spectral features Fig. 2.

Our method is based on a modified Auxiliary Classifier Variational Autoencoder (ACVAE) which maps spectra into a condition-dependent latent space and transforms them according to a desired secondary system. Unlike PDS-based approaches, it does not require predefined primary and secondary conditions and operates in a non-parallel manner, enabling the model to capture complex, non-local and non-linear correlations across widely separated spectral regions. Furthermore, the approach does not require one-to-one correspondence between spectra.

The model we draw inspiration from was originally designed for voice conversion (VC). However, the problem formulation in both VC and spectral transfer is similar; both involve transforming data from one condition to another, such as converting a man's voice to a woman's voice or transforming LIBS spectra from Earth to Moon atmospheric conditions. The ACVAE itself is based on the Conditional Variational Autoencoder[26] (CVAE). CVAE for a given dataset $\mathcal{D} = (x_i, c_i)_{i=1}^{N}$ learns to encode the input data x_i and condition denoted as c_i into a normally distributed latent space z *via* f_{enc}:$(x_i; c_i) \rightarrow z_i$, and decode the latent representation and condition back into the reconstructed data *via* $f_{\text{dec}} : (z_i; c_i) \rightarrow \hat{x}_i$, where $\hat{x}_i$ indicates reconstructed data. This approach allows the model to encode and decode information while taking the given condition into account. However, standard CVAEs impose no restriction on how the encoder and decoder utilize the attribute class and are free to ignore it by learning a mapping, such that $f_{\text{enc}}(x_i; c_i) = f_{\text{enc}}(x_i)$ and decoding as $f_{\text{dec}}(z_i; c_i) = f_{\text{dec}}(z_i)$, thereby reducing the model to a vanilla Variational Autoencoder[27] (VAE). In contrast to a standard CVAE, our method explicitly encourages the model to learn and apply the characteristics of each condition during transfer. Without this constraint, the model could optimize only for reconstruction accuracy while neglecting the target condition. To address this, we use a modified ACVAE model that originally satisfies the following criterion

$$L(\varphi, \theta, \psi) = \mathcal{J}(\varphi, \theta) + \lambda_{\mathcal{L}} \cdot \mathcal{L}(\varphi, \theta, \psi) + \lambda_I \cdot I(\psi) \tag{2}$$

With $\mathcal{J}$ denoting the variational lower bound for CVAE, where φ are the encoder network parameters, θ decoder network parameters, and ψ auxiliary classifier network parameters. $\mathcal{L}$ represents the regularization term with respect to φ, θ and ψ, and I represents cross entropy between the predicted and true conditions, which captures the dependence of the latent variable z on the condition c through the auxiliary classifier parameters, λ_L and λ_I being regularization parameters. For clarity, we omit the data dependence in the loss expression. This objective can be considerably simplified to

$$\begin{aligned} L = \big(-\log p\big(x|\mu_{\text{dec}}, {\sigma_{\text{dec}}}^2\big) + D_{\text{KL}}\big(q_\varphi(z)\|p(z)\big)\big) \\ +\lambda_{\mathcal{L}} \cdot \mathcal{L}_{\text{cls}}\big(c|f_\theta(z)\big) + \lambda_I \cdot I(\psi; c) \end{aligned} \tag{3}$$

which consists of the negative Gaussian log likelihood[28] between the input x and its reconstruction from the decoder, parameterized by μ_{dec} and ${\sigma_{\text{dec}}}^2$, the Kullback–Leibler divergence[29] D_{KL} between parametrized $q_\varphi(z)$ and the assumed Gaussian prior $p(z) = \mathcal{N}(z|0, I)$, classification loss $\mathcal{L}_{\text{cls}}$, which encourages accurate prediction of the condition c from the decoder output $f_\theta(z)$, mapping the latent representation z to its reconstruction $\hat{x}$, and the cross-entropy term $I(\psi; c)$ serving as the objective for the auxiliary classifier.

We then extend this with additional physics-informed regularization terms

$$\begin{aligned} L(\varphi, \theta, \psi) = \mathcal{J}(\varphi, \theta) + \lambda_{\mathcal{L}} \cdot \mathcal{L}(\varphi, \theta, \psi) + \lambda_I \cdot I(\psi) \\ +\lambda_T \left(1 - \frac{1^T F(x, c, c')}{1^T x'}\right)^2 \\ +\lambda_P \left(1 - \frac{\hat{P}(x)}{P(x')}\right)^2 \\ +\lambda_N \left(\frac{1^T \min\big(0, F(x, c, c')\big)}{1^T |F(x, c, c')|}\right)^2 \end{aligned} \tag{4}$$

That extends the original one with physics-based objectives.[30] The model is defined as a composition $F = f_{\text{encoder}} \bigcirc f_{\text{decoder}}$ with

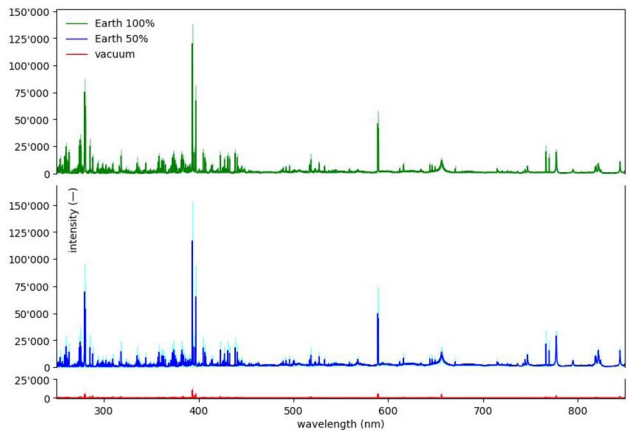

Fig. 2 Comparison of LIBS spectra acquired under different conditions for the same sample. The solid line indicates the mean spectrum across repeated measurements for the respective condition, while the shaded region represents its standard deviation. The visualized sample is a synthetic lunar material rich in aluminum and aluminum oxide.

transfer denoted as $x' = F(x, c, c')$, where x is the input spectrum, c its associated condition and x' is the corresponding spectrum of the same sample under a different condition c'. We use $\hat{P}$ to denote the achieved peak coverage, with P representing the desired peak coverage within the optimization objectives. Peak coverage is computed by first identifying significant peaks in the target spectra and based on them creating a mask. This mask is then applied to both the transferred spectra and target spectra, and the summed peak intensities are compared. This approach focuses specifically on reproducing key spectral features rather than on full-spectrum reconstruction, ensuring that the optimization emphasizes physically meaningful peaks.

The proposed changes account for three key physical properties of spectra—total emissivity, peak structure, and non-negativity—respectively regularized by λ_T, λ_P, and λ_N. Total emissivity reflects the material's overall energy emission and must remain within realistic bounds. Spectral peaks encode elemental and molecular signatures; distorting them hinders reliable identification or quantification. Negative values are physically meaningless in this context, as spectral intensity represents emitted energy. Incorporating such physics-informed constraints in neural network training has been shown to improve model generalization and maintain scientific plausibility without compromising performance.[31,32] Enforcing these constraints helps ensure that the model produces outputs that are both accurate and physically plausible. Additionally, we expand the model architecture by including more layers, skip connections, and learnable attribute class embeddings.

3.1 Baseline

We assess the model based on how well the transferred spectra can be used for regression tasks, using PDS and PDS-PLS models as baselines. Convolutional Neural Networks[33] (CNNs) are trained on the known sample compositions $\mathcal{D}_c = (x_i, y_i)_{i=1}^{N_c}$ for each condition c, where x_i denotes the spectra of sample i under condition c and y_i denotes the corresponding target property or composition. A separate CNN is trained for each condition, but for convenience we omit the condition-specific notation in the following. Predictions on the transferred spectra obtained with the CNNs are compared against the ground-truth labels y as well as against predictions made directly from the original spectra. Let $y_{\text{transf}} = f_{\text{CNN}_c}(F(x, c, c'))$ where x denotes the original spectra and F is the transfer model. We then define: n_c denotes the total number of conditions, and N_c denotes the number of samples for a given condition.

$$E_{\text{true}} = \frac{1}{n_c(n_c-1)} \sum_{i=1}^{n_c} \sum_{j=1 j\neq i}^{n_c} \sqrt{\frac{\sum_{k=1}^{N_c} \left(y_k - y_{\text{transf},\ k}^{(i,j)}\right)^2}{N_c}} \quad (5)$$

$$E_{\text{delta}} = \frac{1}{n_c(n_c-1)} \sum_{i=1}^{n_c} \sum_{j=1 j\neq i}^{n_c} \sqrt{\frac{\sum_{k=1}^{N_c} \left(f_{\text{CNN}}(x_k) - y_{\text{transf},\ k}^{(i,j)}\right)^2}{N_c}} \quad (6)$$

We propose the latter evaluation function as there are noticeable spectral differences between individual measurements. This approach helps reduce the number of factors that affect the evaluation, making the results more consistent and easier to interpret.

Further details regarding the CNN and its training procedure are provided in the Appendix in the SI.

3.2 Experiments and data

3.2.1 Samples. A set of commercially available simulant samples from Exolith lab (Mars samples: Jez-1, MGS-1C, MGS-1S, MGS-1; Moon samples: LHS-1D, LHS-1E, LHS-1, LMS-1D, LMS-1, Dusty)[34] were mixed to create additional samples for robustness. Each new sample was created by mixing two original samples for the respective type (Mars and Moon). Their resulting composition is listed in the SI. The original samples were weighed with 0.001 g precision, and based on the weight ratio, the composition of the mixed samples was calculated. As a result of this process, a total of 42 samples were analyzed for both Martian (21 samples) and Lunar (21 samples) regolith simulants. All the samples were compressed into pellets of 5 g, 30 mm diameter, 15 ton pressure for 1 minute. The average grain size was 70 μm.

The 42 synthetic Lunar and Martian regolith samples were measured under Earth and simulated Lunar (Earth-based vacuum) atmosphere using two laser energies (30 mJ and 15 mJ), with 100 spectra collected for each configuration. This resulted in four measurement conditions. However, due to the low spectral information content in the simulated lunar atmosphere with 15 mJ laser energy, that condition was excluded, leaving three conditions in total. All data were collected on the same instrument, the CEITEC Discovery[35,36] equipped with an atmospheric chamber.

3.2.2 Measurement parameters. LIBS Discovery vacuum chamber was used for the measurements to ensure controlled experimental conditions. The setup included an ablation laser operating at 1064 nm with a pulse energy of 30/15 mJ and a pulse duration of 8 ns. An Echelle spectrometer with an EMCCD detector, covering a spectral range of 200–1000 nm with a resolution of 0.035 nm, was used in combination with wide-range collection optics and a 400 μm optical fiber. The spectrometer was configured with a gate delay of 0.15 μs after the ablation and a gate width of 50 μs for all examined conditions. The vacuum chamber was equipped with a rotary pump to simulate the required experimental conditions: Earth's atmosphere and low vacuum (0.1 mbar). Each condition experiment consisted of 100 laser pulses per sample.

3.3 Spectral differences

There are significant spectral differences between measurements performed under given experimental conditions. As mentioned in Section 1, these differences arise from variations in plasma evolution, emission intensities, and line broadening effects associated with pressure and composition of the given atmospheres.

To illustrate these effects, Fig. 2 and 3 shows example spectra acquired under all conditions for the same sample, highlighting shifts in peak intensities and the presence or absence of specific

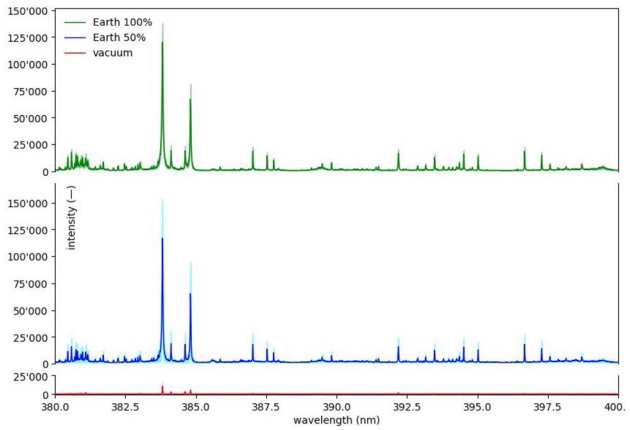

Fig. 3 Comparison of LIBS spectra for 380–400 nm spectral range, acquired under different conditions for the same sample. The solid line indicates the mean spectrum across repeated measurements for the respective condition, while the shaded region represents its standard deviation. The visualized sample is a synthetic lunar material rich in aluminum and aluminum oxide.

emission lines. The main difference lies in the enormous intensity disparity between low-vacuum and Earth spectra, whereas spectra acquired at 15 mJ and 30 mJ energy under Earth's atmospheric conditions exhibit significantly different variance with only subtle differences in peak intensity.

4 Results and discussion

In this section, we present the performance of the proposed method compared to the baseline PDS-based approaches. The evaluation results are summarized in Tables 1 and 2, as well as visualized in Fig. 4. We achieved noticeably better results in terms of metrics and spectra fidelity, as visualized in Fig. 5–8, in comparison to the baseline models in our Table 1 and similar work by Lepore *et al.*[22] Table 3.

Table 1 Performance of models under different test spectra. The first column indicates the source spectra used for model training, while the second column shows the target spectra to which the data were transferred and on which the models were evaluated. The results are reported as average RMSE, with their respective standard deviation values, across samples when comparing ground truth concentration labels y and predicted concentrations of transferred spectra $y_{transf} = f_{CNN_c}(F(x, c, c'))$ (see Section 3.1)[a]

	Ave. RMSE	Std. dev.
Test on Earth 30 mJ spectra		
Earth 15 mJ_{true}	2.72	0.05
Earth 15 mJ_{delta}	2.32	0.12
Vacuum 30 mJ_{true}	2.68	0.04
Vacuum 30 mJ_{delta}	1.76	0.03
Test on Earth 15 mJ spectra		
Earth 30 mJ_{true}	3.30	0.07
Earth 30 mJ_{delta}	2.69	0.10
Vacuum 30 mJ_{true}	3.27	0.01
Vacuum 30 mJ_{delta}	1.68	0.02
Test on vacuum 30 mJ spectra		
Earth 15 mJ_{true}	3.09	0.10
Earth 15 mJ_{delta}	1.97	0.20
Earth 30 mJ_{true}	3.17	0.10
Earth 30 mJ_{delta}	2.53	0.10

[a] Each spectra was transformed to the condition it was tested on.

Table 2 Comparison of models in terms of performance metrics across multiple seeds. Results are reported as the mean value ± standard deviation to reflect variability across different random initializations of parameters

Metric	ACVAE [ours]	PDS	PDS-PLS
E_{true} (eqn (5))	3.08 ± 0.04	3.53 ± 0.01	5.62 ± 0.03
E_{delta} (eqn (6))	2.24 ± 0.07	2.44 ± 0.01	2.44 ± 0.01
R^2	0.93 ± 0.01	0.90 ± 0.01	0.72 ± 0.01

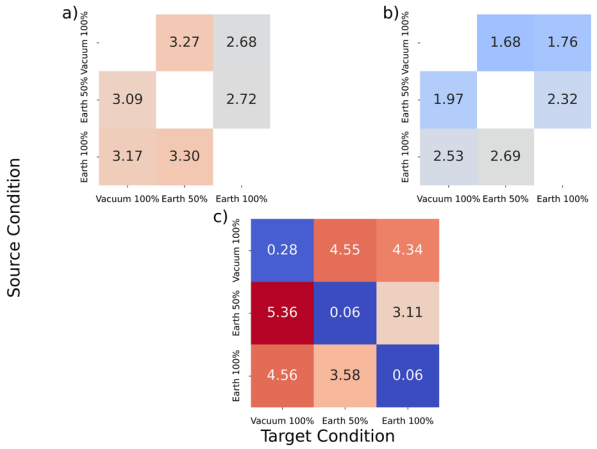

Fig. 4 Heatmaps illustrating spectral transfer performance across all pairs of measurement conditions. Panel (a) shows E_{true} (eqn (5)), representing the absolute prediction error obtained after transferring spectra from each source condition (rows) to each target condition (columns). Lower values correspond to more accurate transfer. Panel (b) displays E_{delta} (eqn (6)). Here lower values correspond to a more accurate transfer, as well. Together, the heatmaps provide a comprehensive overview of the transfer quality for every source–target combination. Panel (c) shows the baseline E_{true} values computed without performing any spectral transfer. In this case, spectra from each source condition (rows) are directly compared to each target condition (columns). This panel illustrates the magnitude of the prediction error when no transfer is applied and serves as a reference against which the improvements from Panels (a), (b) and Table 1 can be assessed.

Our method provides a substantial improvement over the baseline in terms of sample quantification. While PDS-based methods generally reproduce the overall spectral shape, they often distort peak profiles and introduce considerable variation within the transferred spectra. In contrast, the ACVAE-based model preserves key spectral features more consistently and, likely owing to its convolutional architecture, generates smoother and more believable spectra.

Across all transfer directions, the proposed method consistently yields lower RMSE values compared to the baselines. The magnitude of the improvement varies with the transfer scenario, but the trend remains stable across repeated runs,

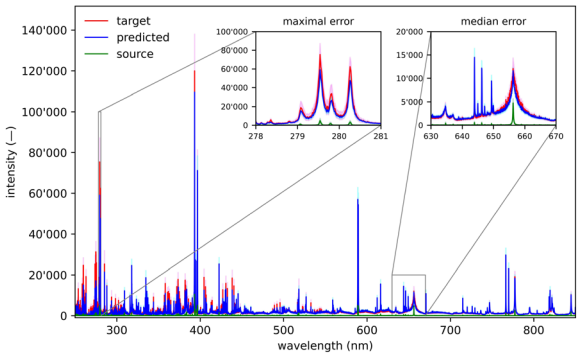

Fig. 5 Visualization of our approach, illustrating average transfer quality and effectiveness. Spectra are transferred from low-vacuum to Earth conditions at 30 mJ laser energy. The plot shows the mean transferred spectra alongside the target atmospheric spectra, with shaded regions indicating their respective standard deviations.

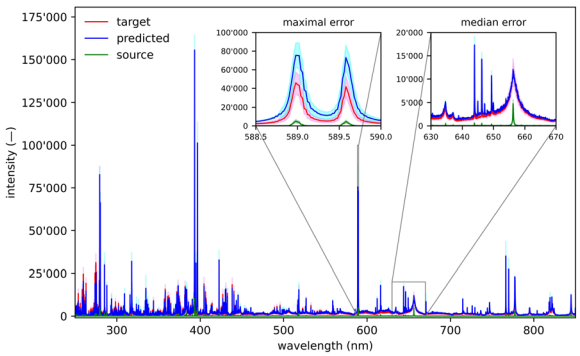

Fig. 6 Visualization of PDS approach, illustrating average to best transfer quality and effectiveness. Spectra are transferred from low-vacuum to Earth conditions at 30 mJ laser energy. The plot shows the mean transferred spectra alongside the target atmospheric spectra, with shaded regions indicating their respective standard deviations.

indicating that the proposed model is not overly sensitive to initialization or random sampling effects.

While the results consistently favor our approach, it should be noted that no standardized benchmarks for LIBS data

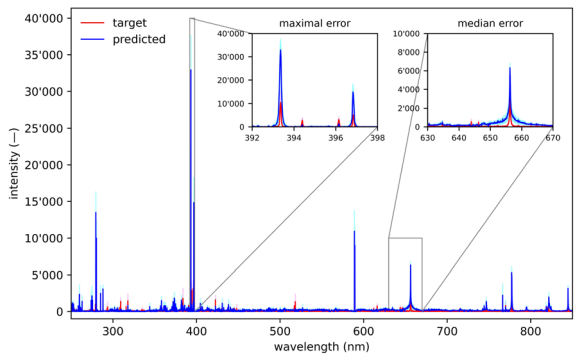

Fig. 7 Visualization of our approach, illustrating worst case transfer quality and effectiveness. Spectra are transferred from Earth at 15 mJ laser energy to low-vacuum. The plot shows the mean transferred spectra alongside the target atmospheric spectra, with shaded regions indicating their respective standard deviations.

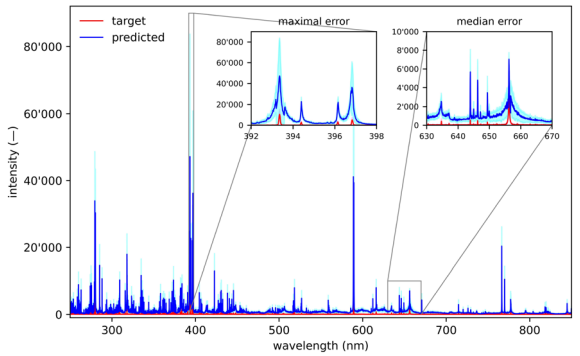

Fig. 8 Visualization of PDS approach, illustrating below-average to worst-case transfer quality and effectiveness. Spectra are transferred from Earth at 15 mJ laser energy to low-vacuum. The plot shows the mean transferred spectra alongside the target atmospheric spectra, with shaded regions indicating their respective standard deviations.

Table 3 Performance of PDS-PLS models under different test spectra from Lepore *et al.*[22]

	Ave. RMSE-*P*	Std. dev.	Ave. R^2
Test on Earth spectra			
Train on all atmospheres	1.49	1.10	0.83
Mars cal trans.[a] to Earth train	3.32	2.11	0.77
Vac. cal trans. to Earth train	2.56	2.41	0.75
Test on Mars spectra			
Train on all atmospheres	1.27	0.92	0.87
Earth cal trans. to Mars train	2.96	1.83	0.80
Vac. cal trans. to Mars train	2.43	2.02	0.80
Test on vacuum spectra			
Train on all atmospheres	1.26	0.94	0.81
Earth cal trans. to vac train	6.87	5.56	0.51
Mars cal trans. to vac train	2.88	2.35	0.74

[a] "cal trans." indicates use of the PDS-PLS calibration transfer algorithm.

transfer currently exist, which complicates direct comparisons with previously published work. Nevertheless, within our experimental setup, the proposed method demonstrates a clear and consistent advantage over both PDS and PDS-PLS.

5 Conclusion

Our research presents a method for accurate and computationally efficient transfer of LIBS spectra across many-to-many experimental conditions. The approach enables direct comparison and transformation of spectra collected under different configurations with minimal need for system recalibration or physical synchronization.

The methodology involves acquiring multiple measurements of the same set of samples under each condition and training a model to learn the mapping between them. The model is trained and validated using an in-house, multi-sample and multi-condition dataset comprised of synthetic Lunar and Martian regolith spectra, acquired on the same LIBS instrument

Section 3.2.1. The results show that the proposed method can successfully transfer spectra between conditions while preserving quantitative accuracy, achieving better performance than baseline PDS-based approaches. This makes it potentially relevant for scenarios such as planetary exploration, where measurement conditions are inherently variable and direct recalibration is impractical.

5.1 Future work

The dataset used in this study is modest in size and spans only a limited set of measurement conditions. Expanding it with more spectra and a broader range of conditions would improve the method's robustness and allow us to validate whether the model remains a viable approach across a wider set of conditions.

6 Computational resources

All experiments were conducted on a desktop with the following specifications:

- CPU: AMD Ryzen 5 7500F.
- GPU: NVIDIA RTX 4060 Ti 16 GiB.
- RAM: 32 GiB DDR5.

The code was executed using PyTorch 2.4.1 with CUDA 12.4 support where applicable.

Author contributions

Antoni Pawłowicz: methodology, software, visualization, writing – original draft and final version, formal analysis, validation, investigation, Jakub Vrábel: conceptualization, writing – review & editing, Jakub Buday: resources, data curation, writing – review & editing, Erik Képeš: writing – review, Pavel Pořízka: resources, supervision, writing – review, Jozef Kaiser: project administration.

Conflicts of interest

There are no conflicts to declare.

Data availability

The datasets generated and analyzed during the current study are available from the corresponding author upon request. The source code supporting this work is openly available at: https://github.com/AnthonyP57/Many-to-many-transfer-of-LIBS-spectra-across-multiple-experimental-conditions.

Supplementary information (SI) is available. See DOI: https://doi.org/10.1039/d5ja00401b.

Acknowledgements

PP gratefully acknowdge the support of Czech Science Foundation (GACR, no. 23-05186K). JK acknowledges the support of the Faculty of Mechanical Engineering, Brno University of Technology (FSI-S-23-8389).

References

1 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.
2 R. C. Wiens, S. Maurice, *et al.*, *Space Sci. Rev.*, 2012, **170**, 167–227.
3 R. C. Wiens, S. Maurice, *et al.*, *Space Sci. Rev.*, 2020, **217**, 4.
4 L. Jolivet, V. Motto-Ros, L. Sorbier, T. Sozinho and C.-P. Lienemann, *J. Anal. At. Spectrom.*, 2020, **35**, 896–903.
5 T. F. Boucher, M. V. Ozanne, M. L. Carmosino, M. D. Dyar, S. Mahadevan, E. A. Breves, K. H. Lepore and S. M. Clegg, *Spectrochim. Acta, Part B*, 2015, **107**, 1–10.
6 S. M. Clegg, E. Sklute, M. D. Dyar, J. E. Barefield and R. C. Wiens, *Spectrochim. Acta, Part B*, 2009, **64**, 79–88.
7 E. Képeš, I. Gornushkin, P. Pořízka and J. Kaiser, *Anal. Chim. Acta*, 2020, **1135**, 1–11.
8 E. Képeš, I. Gornushkin, P. Pořízka and J. Kaiser, *Analyst*, 2021, **146**, 920–929.
9 S. Shabanov and I. Gornushkin, *J. Quant. Spectrosc. Radiat. Transf.*, 2018, **204**, 190–205.
10 J. Chen, Y. Ding, A. Hu, W. Chen, Y. Wang, M. Zhao and Y. Shu, *Opt. Express*, 2023, **31**, 41129–41148.
11 Y. yuan Chen and Z. bin Wang, *Chemom. Intell. Lab. Syst.*, 2019, **192**, 103824.
12 E.-L. Bergman, H. Brage, M. Josefson, O. Svensson and A. Sparén, *J. Pharmaceut. Biomed. Anal.*, 2006, **41**, 89–98.
13 T. Boucher, C. Carey, S. Mahadevan and M. D. Dyar, *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015, pp. 2511–2517.
14 J. Vrábel, E. Képeš, P. Nedělník, J. Buday, J. Cempírek, P. Pořízka and J. Kaiser, *J. Anal. At. Spectrom.*, 2023, **38**, 841–853.
15 N. A. S. Suarin and K. S. Chia, *Control, Instrumentation and Mechatronics: Theory and Practice*, Singapore, 2022, pp. 707–716.
16 K. Lepore, M. Dyar and C. Ytsma, *Spectrochim. Acta, Part B*, 2024, **211**, 106839.
17 T. Boucher, M. D. Dyar and S. Mahadevan, *J. Chemom.*, 2017, **31**, e2877.
18 C. Sun, W. Xu, Y. Tan, Y. Zhang, Z. Yue, L. Zou, S. Shabbir, M. Wu, F. Chen and J. Yu, *Sci. Rep.*, 2021, **11**, 21379.
19 G. Jin, Z. Wu, Z. Ling, C. Liu, W. Liu, W. Chen and L. Zhang, *Remote Sens.*, 2022, **14**(16), 3960.
20 Z. Yue, C. Sun, L. Gao, Y. Zhang, S. Shabbir, W. Xu, M. Wu, L. Zou, Y. Tan, F. Chen and J. Yu, *Opt. Express*, 2020, **28**, 14345–14356.
21 H. Kameoka, T. Kaneko, K. Tanaka and N. Hojo, ACVAE-VC: Non-parallel many-to-many voice conversion with auxiliary classifier variational autoencoder, *arXiv*, 2020, preprint, arXiv:1808.05092, DOI: 10.48550/arXiv.1808.05092.
22 K. Lepore, I. Belkhodja, M. Dyar and C. Ytsma, *Spectrochim. Acta, Part B*, 2024, **217**, 106970.
23 Y. Ding, Q. Tan, J. Xu, A. Hu, M. Zhao, X. Li, Y. Shu and X. Liu, *Spectrochim. Acta, Part B*, 2025, **228**, 107171.

24 Z. Zhang, S. Zeng, T. Ji, M. Cao and W. Guo, *Comput. Electron. Agric.*, 2023, **210**, 107882.
25 I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio, Generative Adversarial Networks, *arXiv*, 2014, preprint, arXiv:1406.2661, DOI: 10.48550/arXiv.1406.2661.
26 K. Sohn, H. Lee and X. Yan, *Advances in Neural Information Processing Systems*, 2015.
27 D. P. Kingma and M. Welling, *arXiv*, 2013, preprint, arXiv:1312.6114, DOI: 10.48550/arXiv.1312.6114.
28 D. Nix and A. Weigend, *Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN'94)*, 1994, vol. 1, pp. 55–60.
29 S. Kullback and R. A. Leibler, *Ann. Math. Stat.*, 1951, **22**, 79–86.
30 G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang and L. Yang, *Nat. Rev. Phys.*, 2021, **3**, 422–440.
31 T. Beucler, M. Pritchard, S. Rasp, J. Ott, P. Baldi and P. Gentine, *Phys. Rev. Lett.*, 2021, **126**, 098302.
32 A. Engel and J. Strube, Evaluating Physically Motivated Loss Functions for Photometric Redshift Estimation, *arXiv*, 2023, preprint, arXiv:2310.13624, DOI: 10.48550/arXiv.2310.13624.
33 Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, *Proc. IEEE*, 1998, **86**, 2278–2324.
34 Exolith Lab, The Exolith Lab – Center for Lunar & Asteroid Surface Science, https://sciences.ucf.edu/class/exolithlab/.
35 J. Buday, P. Pořízka, M. Buchtová and J. Kaiser, *Spectrochim. Acta, Part B*, 2021, **182**, 106254.
36 J. Vrábel, E. Képeš, P. Nedělník, A. Záděra, P. Pořízka and J. Kaiser, *J. Anal. At. Spectrom.*, 2025, **40**, 1552–1565.