Supplementary information

Transfer learning based on dynamic time warping algorithms to improve qualitative analysis and quantitative prediction of rocks over multiple LIBS instruments

Yu Rao,[a,b] Lingwei Zeng,[a] Mengfan Wu,[a] Weiheng Kong,[a] Wenxin Ren,[a] Sha Chen,[a] Qinwen Fan,[a] Yixiang Duan,[a] Xu Wang[1,a] and Jie Wang[2,b]

[a] *Research Center of Analytical Instrumentation, School of Mechanical Engineering, Sichuan University, Chengdu 610065, China*
[b] *School of Mechanical Engineering, Sichuan University, Chengdu 610065, China*

[1,2] Corresponding author.
E-mail address: wangxu028@scu.edu.cn (Xu Wang), wangjie@scu.edu.cn (Jie Wang).

Table of Contents

Table S1. The elemental content of all samples. The content of Si, Al, Fe, Ca and Mg are high as major elements and Ti, Mn, K, Na and Ba are less as trace elements in the rocks. In this work, the main purpose is to validate the application of transfer learning in LIBS, therefore, only the major elements are considered in the subsequent experiments. The elemental content in the table is converted by the percentage of mass fraction.

Samples		Concentration (%)									
		Si	Al	Fe	Ca	Mg	Ti	Mn	K	Na	Ba
Dolomite	GBW07114	0.289	0.053	0.028	21.443	13.080	0.009	0.008	0.032	0.022	0.004
	GBW07136	3.850	0.053	0.040	23.621	10.800	0.002	0.021	0.008	0.019	0.003
	GBW070157	3.929	0.635	0.333	20.521	11.856	0.022	0.015	0.032	0.024	0.288
	GBW070158	0.873	0.109	0.171	21.586	12.510	0.005	0.012	0.015	0.009	0.234
	GBW070159	1.008	0.132	0.174	21.536	12.546	0.007	0.012	0.022	0.008	0.223
	GBW070160	2.436	0.374	0.250	21.071	12.258	0.013	0.014	0.027	0.017	0.250
Igneous rock	GBW07103	33.987	7.094	2.291	1.107	0.252	0.172	0.046	4.157	2.322	0.034
	GBW07104	28.289	8.561	5.289	3.714	1.032	0.309	0.060	1.568	2.864	0.102
	GBW07105	20.832	7.322	15.291	6.293	4.662	1.420	0.131	1.925	2.508	0.053
	GBW07109	25.424	9.381	5.185	0.993	0.390	0.288	0.093	6.207	5.312	0.025
	GBW07110	29.428	8.524	3.305	1.764	0.504	0.480	0.069	4.290	2.270	0.105
	GBW07111	27.851	8.767	4.244	3.371	1.686	0.462	0.073	2.904	3.005	0.190
	GBW07112	16.655	7.486	17.321	7.043	3.150	4.614	0.150	0.124	1.565	0.009
	GBW07113	33.964	6.861	2.245	0.421	0.096	0.180	0.108	4.506	1.907	0.051
	GBW07121	30.926	8.645	3.428	1.900	0.978	0.18	0.043	2.157	3.932	0.114
	GBW07122	23.156	7.285	18.760	6.857	4.320	0.551	0.160	0.398	1.536	0.062
Mudstone	GBW03101a	23.324	13.908	7.385	0.093	0.276	0.420	0.040	0.656	0.045	--
	GBW03102a	25.046	16.581	0.231	1.286	0.050	0.018	0.015	0.954	1.892	--
	GBW03103	31.099	7.031	3.870	2.307	1.104	0.396	0.068	2.074	1.343	--
	GBW03104	32.494	7.846	3.969	0.157	0.402	0.408	0.019	3.120	0.148	0.040
	GBW03115	27.953	15.125	0.602	0.500	0.180	0.726	--	1.278	1.291	--
	GBWE070146	21.126	7.264	3.388	6.986	1.044	0.331	0.185	3.568	0.190	--
	GBW07107	27.641	9.964	5.320	0.429	1.206	0.395	0.017	3.452	0.260	0.045
Gypsum	GBW03109	0.784	0.180	0.112	28.029	1.044	0.010	--	0.078	0.048	--
	GBW03111a	0.294	0.074	0.077	23.071	1.482	0.006	--	0.022	0.010	0.001
Limestone	GBW03105a	0.509	0.127	0.077	38.593	0.486	0.006	0.005	0.070	0.013	0.001
	GBW03106a	0.975	0.175	0.119	36.864	1.350	0.009	0.007	0.141	0.013	0.003
	GBW03107a	1.890	0.498	0.406	35.779	1.074	0.031	0.011	0.349	0.020	0.003
	GBW03108a	1.050	0.318	0.266	33.621	3.486	0.018	0.009	0.166	0.012	0.002
	GBWE070147	1.129	0.198	0.175	33.364	3.930	0.014	0.004	0.041	0.008	0.001
	GBWE070148	1.097	0.287	0.292	29.864	6.222	0.019	0.005	0.061	0.008	0.001
	GBWE070149	1.409	0.329	0.202	36.15	1.728	0.019	0.004	0.136	0.019	0.002

	GBWE070150	2.147	0.150	0.131	32.921	3.588	0.009	0.004	0.032	0.007	0.005
	GBWE070151	1.549	0.180	0.127	36.586	1.458	0.012	0.003	0.078	0.005	0.001
	GBWE070152	0.616	0.119	0.106	38.421	0.702	0.008	0.003	0.055	0.004	0.001
	GBWE070153	1.862	0.348	0.211	34.686	2.586	0.031	0.004	0.153	0.021	0.001
	GBWE070154	1.101	0.429	0.207	37.107	0.738	0.022	0.006	0.137	0.007	0.002
	GBWE070155	0.448	0.174	0.102	38.714	0.492	0.010	0.004	0.070	0.004	0.001
	GBWE070156	0.390	0.128	0.100	38.400	0.900	0.007	0.004	0.040	0.004	0.001
	GBW07108	7.280	2.663	3.040	25.479	3.114	0.196	0.043	0.647	0.059	0.012
	GBW07120	3.103	0.360	0.147	36.500	0.426	0.023	0.003	0.124	0.022	0.001
	GBW07127	0.257	0.090	0.252	34.207	4.056	0.007	0.007	0.036	0.016	0.001
	GBW07128	0.336	0.116	0.268	29.964	6.972	0.013	0.007	0.043	0.022	0.001
	GBW07129	0.140	0.079	0.054	39.636	0.144	0.004	0.023	0.010	0.010	0.001
	GBW07130	0.504	0.095	0.225	38.629	0.852	0.004	0.003	0.036	0.011	--
Sandstone	GBW03112	45.971	0.445	0.065	0.055	0.040	0.012	0.001	0.051	0.016	0.032
	GBW03113	44.679	1.249	0.147	0.121	0.059	0.022	0.003	0.556	0.185	0.055
	GBW03114	41.809	2.901	0.336	0.243	0.096	0.061	0.008	1.718	0.809	0.067
	GBW07106	42.168	1.864	2.254	0.214	0.049	0.158	0.016	0.539	0.045	0.014

Considering the LIBS spectral data as a time series, the DTW algorithm can be used to calculate the best relationship of spectral between the primary and the secondary instrument. Then, a transfer model is constructed according to this relationship, so that the same model can be shared among multi-instruments of LIBS.

In the first step, the relationship between the two serials is calculated. Set the primary instrument spectral series as $M = \{M_i,\ i = 1,...,n\}$, the secondary instrument spectral series as $S = \{S_j,\ j = 1,...,k\}$. The DTW algorithm optimally matches channels to each other by calculating the sum of minimum distance between the channels on the two serials. The serial and mapping relationships are shown in **Figure S1** The correlation coefficient between the data of the i_{th} spectral channel of the primary instrument and the j_{th} spectral channel of the secondary instrument can be calculated by equation (1).

$$r(M,S) = \frac{Cov(M(:,i), S(:,j))}{\sqrt{Var(M(:,i)}\sqrt{Var(S(:,j)}} \tag{1}$$

Where $Cov(M(:,i), S(:,j))$ is the covariance and Var is the variance between the two channels. Based on the correlation coefficients between the primary and secondary instruments, the related distances between the spectral series can be found as equation (2).

$$D(M(:,i), S(:,j)) = 1 - r(M,S) \tag{2}$$

Then a cost matrix C is constructed to store the related distances of each channel.

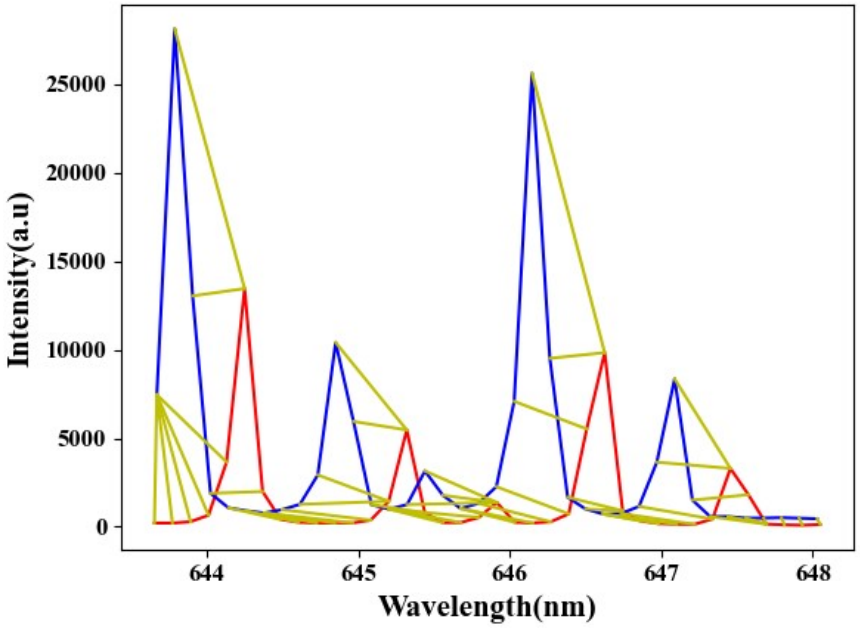

Figure.S1 Spectral serial and mapping relationship

In the second step, optimal route mapping. With the cost matrix C, the DTW algorithm can find the related of minimum distance between each spectral channel on

the primary and the secondary instrument, which is the best correspondence between the two instruments. The route mapping must cover three conditions. Firstly, the route mapping is constrained by the boundary, which must start at point $(1,1)$ and end at point (i,j). Then, the route mapping is limited by the order and cannot skip or cross for matching. Finally, the route mapping is restricted by monotonicity to move only from the right or top or upper right side of a point to avoid entering a loop.

In the cost matrix C, suppose the current point is (i,j), then the next point can only be compared among the three points $(i+1,j),(i,j+1),(i+1,j+1)$. The dynamic route mapping is used to solve for the route that sum of minimum distance in the whole spectral channel, and the calculation is shown in equation (3).

$$D_{min}(i,j) = \min\{D(i-1,j),D(i,j-1),D(i-1,j-1)\} + C(i,j) \tag{3}$$

where i, j are the maximum spectral channels of the primary and secondary instruments, respectively; the route that sum of minimum distance is the best correspondence.

As shown in **Figure S2**, if the optimal route is calculated directly for the full-spectrum channel, it needs to be calculated once for each point. In order to save computation time, in this experiment, the spectra are divided into three segments with 2048 channels in each segment and calculate the optimal dynamic time regularization route together. Compared with the direct calculation of the full spectrum channel, the calculation time is reduced by two-thirds.

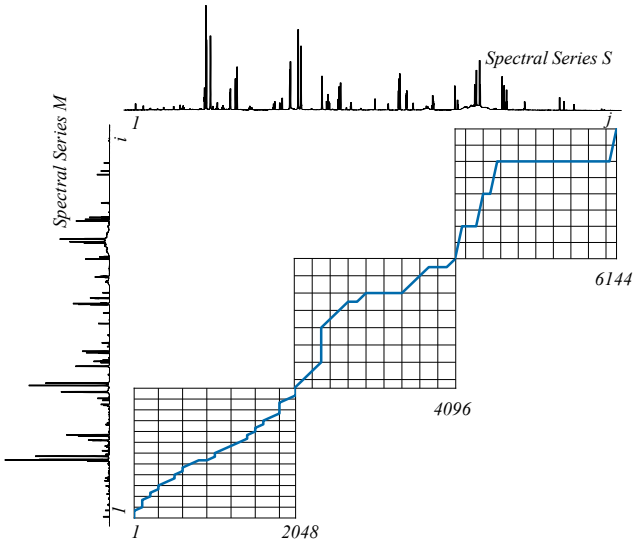

Figure.S2 The best route calculation by DTW

Once we get the optimal correspondence, the third step, training the calibration model. The optimal correspondence for each channel between the primary and secondary instruments is obtained by DTW algorithm, and this correspondence includes one by one and one to many.

When the primary instrument $M(:,i)$ corresponds one by one with the secondary instrument $S(:,j)$, the univariate regression model is constructed by Equation (4).

$$M(:,i) = a_0 + a_1 S(:,j) \tag{4}$$

When the primary instrument $M(:,i)$ corresponds with the secondary instrument from $S(:,j-k)$ to $S(:,j+m)$, the multivariate regression model is constructed by Equation (5).

$$M(:,i) = b_0 + b_1 S(:,j-k) + ... + b_{k+m} S(:,j+m) \quad (5)$$

where a linear regression algorithm was used to calculate the equation coefficients for the univariate regression model and the PLS algorithm was used to calculate the equation coefficients for the multivariate regression model. The transfer model coefficient matrix F is obtained by calculating for the full spectral channel, then the relationship between the primary and secondary instrument spectra can be seen in equation (6).

$$M(:,i) = S(:,j) * F \quad (6)$$

It is necessary to make an accurate evaluation of the accuracy of the transfer model, a concept is introduced here called the correction rate of spectrum (TCRS). Where the average difference in spectra (ARMS) is calculated as:

$$ARMS = \frac{1}{m}\sum_{i=1}^{m}\sqrt{\frac{1}{k}\sum_{j=1}^{k}(M(:,j) - S(:,j))^2} \tag{7}$$

where k is the number of spectral channels and m is the number of samples, $M(:,j)$ is the spectral intensity at the j_{th} channel of the primary instrument, and $S(:,j)$ is the spectral intensity at the j_{th} channel of the secondary instrument. Then the TCRS can be calculated by Equation (8).

$$TCRS(\%) = \frac{ARMS_{uncorrected}^{2} - ARMS_{corrected}^{2}}{ARMS_{uncorrected}^{2}} \times 100\% \tag{8}$$

Where $ARMS_{uncorrected}^{2}$ is average difference in spectra between primary and secondary instrument before corrected, $ARMS_{corrected}^{2}$ is average difference in spectra between primary and secondary instrument after corrected.

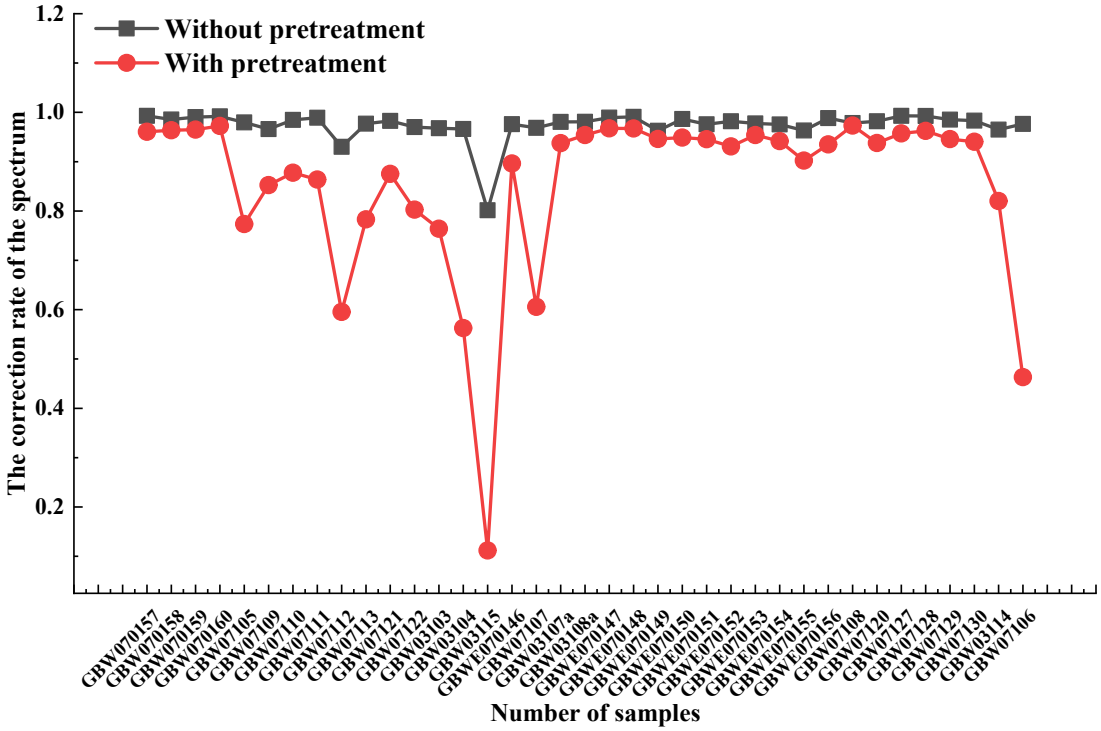

Figure S3. The comparison of the spectral correction rates before and after pre-treatment showed that the average correction rate of the spectra was 97.38% without pre-treatment, and 85.29% with pre-treatment. It may be that some information was lost in the small-scale spectra, which caused the spectrum correction was poor. Therefore, the spectral data without pretreatment were used in the subsequent experiments.

It is necessary to make an accurate evaluation of the accuracy of the transfer model, we have introduced a concept called spectral transfer error rate (STER). The smaller the value of STER, the smaller the difference between the two spectra collected from primary and secondary instrument. For the same batch of samples, set the spectral matrix collected by the primary instrument as $\{M_{ij}\ ,\ i=1,...,N\ ;\ j=1,...,K\}$, set the spectral matrix collected by the secondary instrument as $\{\ S_{ij}\ ,\ i=1,...,N\ ;\ j=1,...,K\ \}$, where N is the number of samples and K is the number of wavelength points.

The spectra transfer error rate for the i_{th} sample is calculated by the formula:

$$STER_i = \frac{\frac{1}{K}\sum_{j=1}^{K}|M_{ij} - S_{ij}|}{\frac{1}{K}\sum_{j=1}^{K}|Z_{ij}|} \tag{1}$$

Where, $Z_{ij} = \frac{M_{ij} + S_{ij}}{2}$, $j=1,...,K$ is the average spectrum of the ith sample.

For the entire sample set of data $i=1, ...,N$ there are the average and maximum transfer error rates:

$$STER_{ave} = \frac{1}{N}\sum_{i=1}^{N} STER_i \tag{2}$$

$$STER_{max} = max(STER_i) \tag{3}$$

The difference of their transfer error rates between half-width windows of 10 and 12 is not much, the spectral transfer error rate of average is 10.67% and 10.64%, the spectral transfer error rate of max is 25% and 24.91%, respectively. However, the calculation time of the half-width window of 12 is slower than that of the half-width window of 10, 36.68 seconds and 29.57 seconds, respectively, so the half-width window of 10 is chosen.

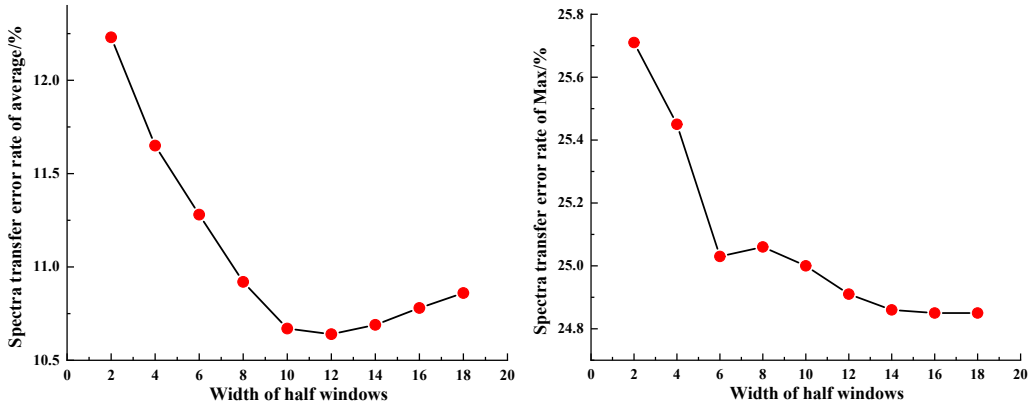

Figure S4. The parameter optimization process of PDS algorithm

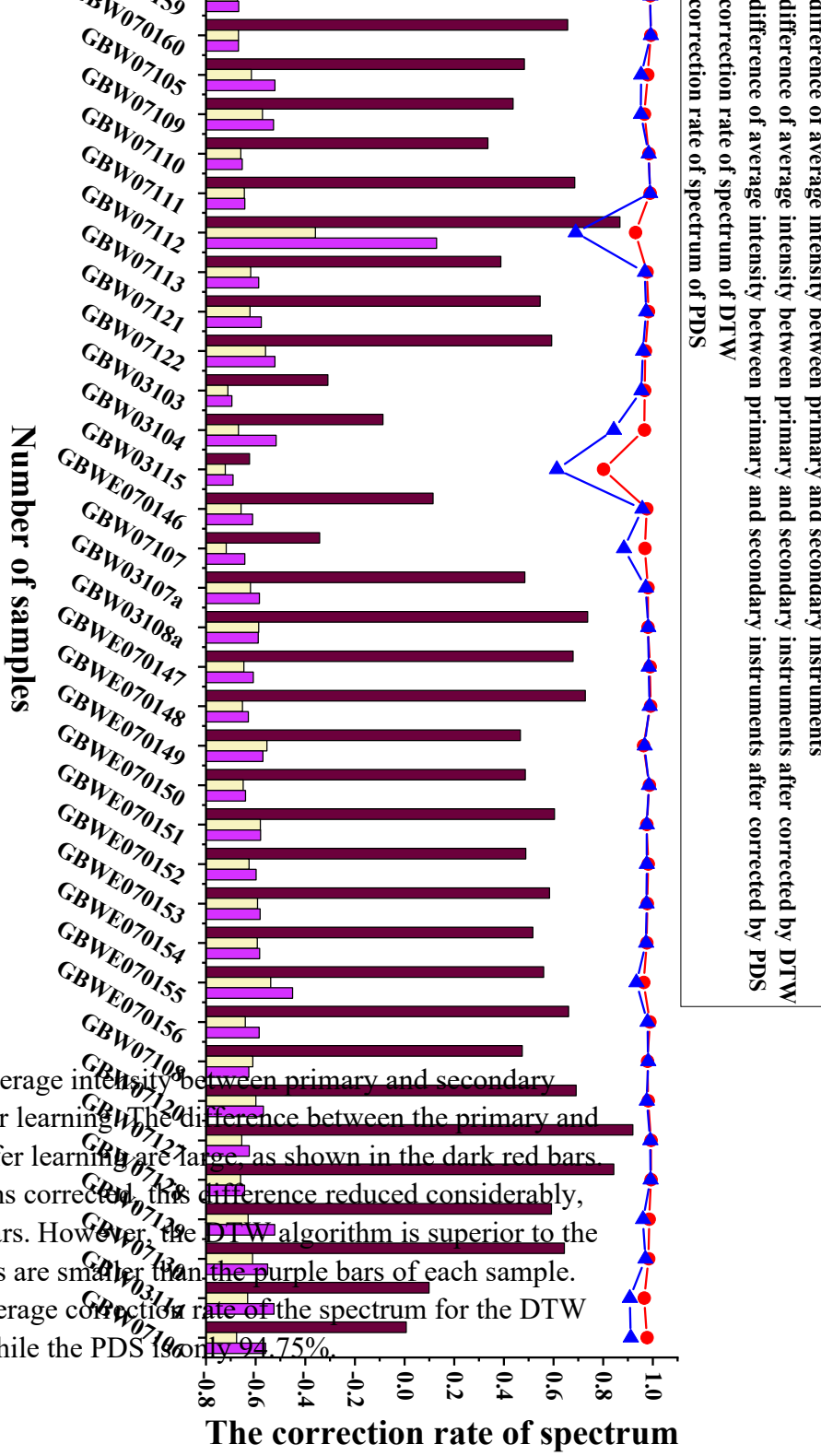

Figure S5. The difference of average intensity between primary and secondary instruments before and after transfer learning. The difference between the primary and secondary instruments before transfer learning are large, as shown in the dark red bars. With the DTW and PDS algorithms corrected, this difference reduced considerably, as shown in the beige and purple bars. However, the DTW algorithm is superior to the PDS algorithm, and the beige bars are smaller than the purple bars of each sample. And it can also be seen that the average correction rate of the spectrum for the DTW is 97.38%, while the PDS is only 94.75%.

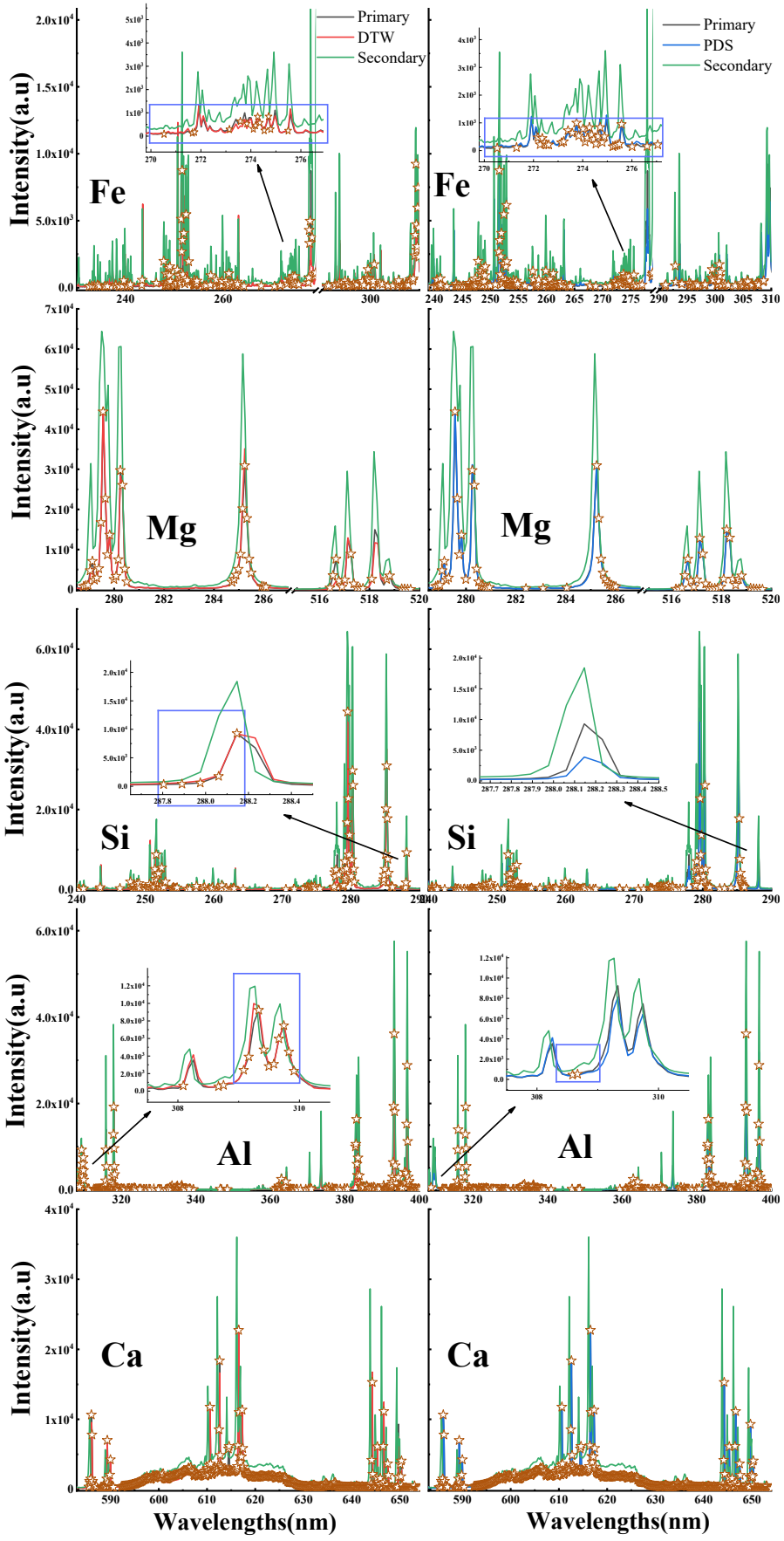

Figure S6. It shows the selection features of the major elements.