Tensor Component Analysis for the Investigation of Depth Trends in ChemCam LIBS Data from Gale Crater, Mars

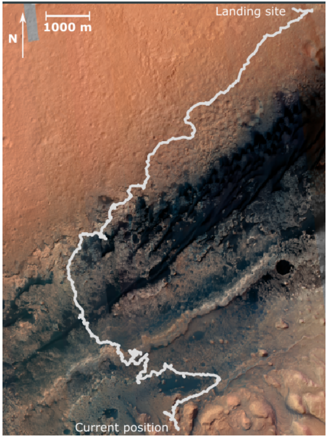
Kristin Rammelkamp, Olivier Gasnault, Olivier Forni, Erwin Dehouck, Candice C. Bedford, Agnès Cousin, Susanne Schröder, Jeremie Lasue, Roger C. Wiens, Nina Lanza

06-21-2022

Mars Science Laboratory

- NASA Mars mission landed in Gale crater 6th of August 2012
- Curiosity has traveled more than 28 km.

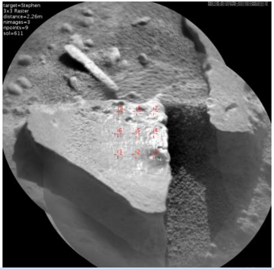
Latest selfie taken on sol 3303 (11-20-2021)
NASA/JPL-Caltech/MSSS

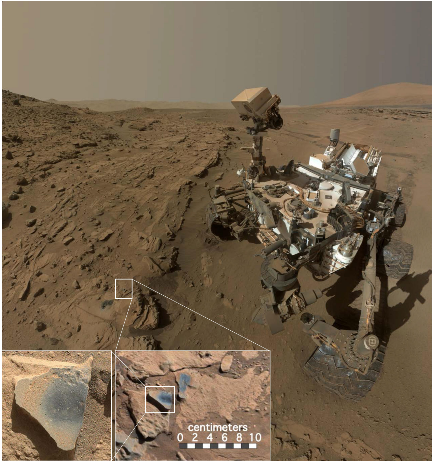


Position on sol 3502 (06-13-2022)
https://mars.nasa.gov/msl/mission/where-is-the-rover/

ChemCam on Curiosity

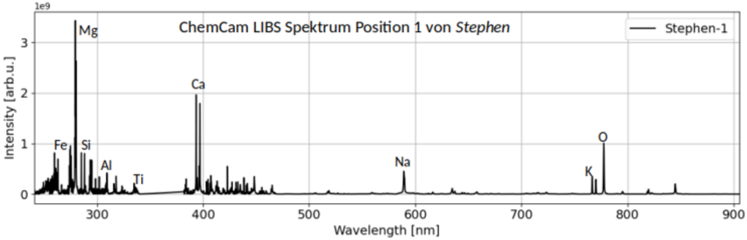
- Chemistry composition at remote distances and at small scales (300-500 microns)
- Technique: Laser-induced breakdown spectroscopy (LIBS)
- Remote micro imager (RMI)
- Along traverse: More than 3 000 targets analyzed with 30 000 LIBS spots and 900k laser shots





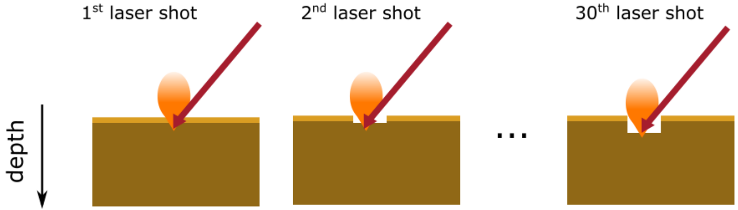
MSSS/JPL/NASA (PIA18390)

ChemCam Data



- One spectrum: 6144 channels
- Usually: average spectrum from one position discarding the first five shots (dust contamination)
- Recent procedure for chemical composition: Balanced combination of Independent Component Analysis (ICA) and Partial Least-Squares Regression (PLS-R) for prediction of geological major elemental abundances (SiO_2, TiO_2, Al_2O_3, FeO_T, MgO, CaO, Na_2O, K_2O)

Depth Trends with ChemCam

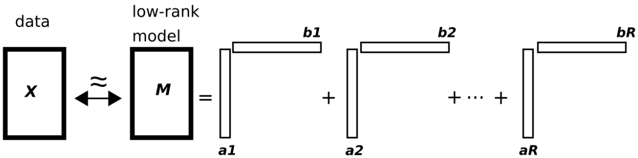


Schematic representation: scaling not realistic

- Usually 30 shots at one position
- How much material is ablated with each shot depends on sample properties (rock hardness, laser coupling...). But the final depth of a LIBS crater after 30 shots can be estimated to be in the 100 μm range (*Maurice et al., 2016, JAAS*)
- Different correlations of elemental emission lines with depth for different mineral phases

Tensor Component Analysis (TCA)

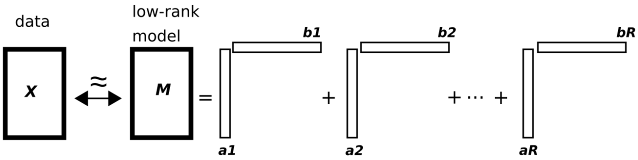
Concept is similar to matrix decomposition → detecting low-rank structure in data



$$\mathcal{X} \approx \mathcal{M} = \sum_{i=1}^{R} \mathbf{a_i} \circ \mathbf{b_i} = \mathbf{AB^T}$$

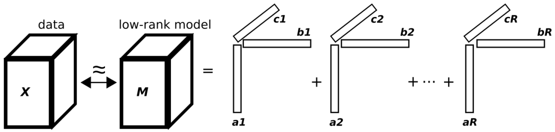
Tensor Component Analysis (TCA)

Concept is similar to matrix decomposition → detecting low-rank structure in data



$$\mathcal{X} \approx \mathcal{M} = \sum_{i=1}^{R} \mathbf{a_i} \circ \mathbf{b_i} = \mathbf{AB^T}$$

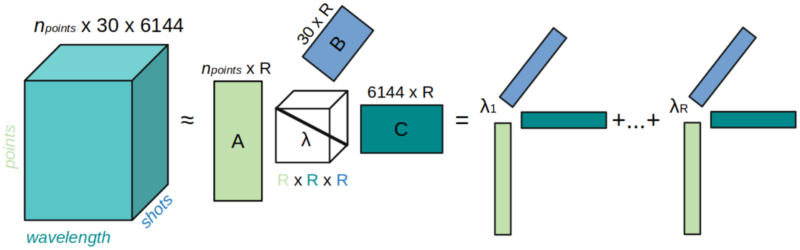
Going to 3 dimensions → tensor component analysis



$$\mathcal{X} \approx \mathcal{M} = \sum_{i=1}^{R} \mathbf{a_i} \circ \mathbf{b_i} \circ \mathbf{c_i} = [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$$

- Different names: PARAFAC/CANDECOMP/CP all names for same Canonical Polyadic
- Great overview paper by Kolda and Bader (2009): *Tensor Decompositions and Applications*

TCA for ChemCam shot-to-shot data



- A-matrix represents the analysis points → observation dimension
- B-matrix represents the consecutive shots at one analysis point → depth dimension
- C-matrix represents the wavelength → spectral dimension

Dataset

Table 1
Mean MOCs and Sum of Oxides (Totals) for Each Cluster Together With the Median and the Standard Deviation, all Values are in wt%

		Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
SiO_2	mean	75.77	56.16	54.70	48.92	44.42	28.32
	median	76.0	55.30	54.90	49.90	44.10	32.80
	stddev	5.48	4.27	2.63	6.14	3.39	16.65
TiO_2	mean	2.47	0.74	1.00	1.23	0.92	0.59
	median	2.32	0.75	0.97	1.03	0.90	0.65
	stddev	1.11	0.21	0.15	0.84	0.17	0.34
Al_2O_3	mean	5.54	19.30	11.95	10.94	9.42	5.32
	median	5.20	19.30	11.80	10.90	9.10	5.90
	stddev	3.33	2.59	1.45	2.50	2.03	3.01
FeO_T	mean	5.43	7.44	18.78	21.15	18.96	10.74
	median	4.90	7.40	18.90	20.10	19.10	11.80
	stddev	3.61	4.91	1.59	4.40	1.88	5.74
MgO	mean	2.43	1.50	5.79	3.43	7.91	2.69
	median	2.30	1.40	5.50	3.50	7.60	2.70
	stddev	0.85	0.77	1.32	0.92	2.30	1.11
CaO	mean	1.60	6.44	1.76	3.32	6.68	23.08
	median	1.10	6.80	1.70	2.80	6.70	21.50
	stddev	1.59	2.66	0.56	1.87	1.61	8.55
Na_2O	mean	1.69	5.70	2.72	2.91	2.28	1.01
	median	1.50	5.60	2.65	2.80	2.15	0.97
	stddev	0.86	1.26	0.63	0.77	0.66	0.67
K_2O	mean	0.66	1.54	1.43	1.16	0.62	0.30
	median	0.54	1.14	1.39	1.10	0.43	0.26
	stddev	0.67	1.40	0.48	0.58	0.63	0.28
Totals	mean	95.59	98.82	98.13	93.07	91.19	72.05
	median	95.45	98.79	98.18	93.01	91.13	77.88
	stddev	2.72	2.40	2.75	4.85	4.31	17.27

Table from *Rammelkamp et al., 2021, Earth and Space Science*

- Previous study about classification of ChemCam LIBS data with unsupervised clustering

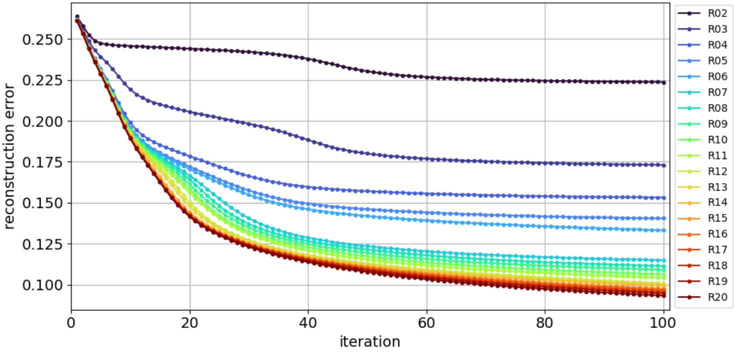
→ 6 clusters of different sizes

- **This study**: Only data from cluster 2 observations with **felsic** compositions (typical elements: Si, Al and alkalis Na and K)
- Rather small cluster with 485 analysis points

→ The data tensor $\mathcal{X}$ in this study has dimensions: $485 \times 30 \times 6144$

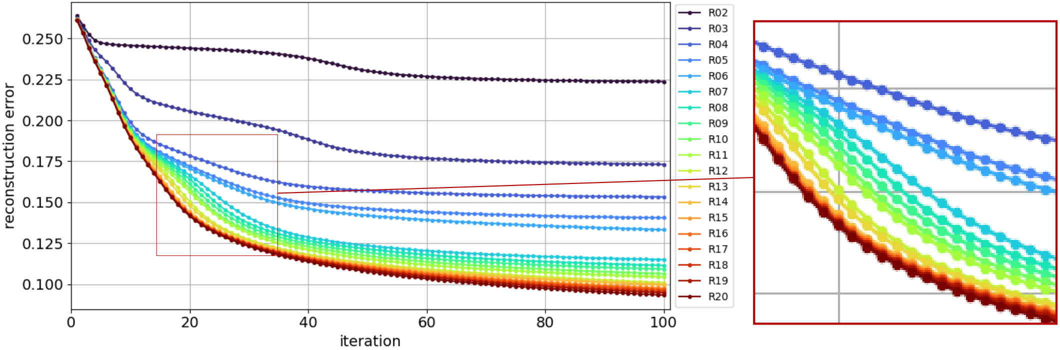
Selection number of ranks

- Basic metric to compare models: normalized reconstruction error $\frac{\|\mathcal{X}-\mathcal{M}\|_F^2}{\|\mathcal{X}\|}$



Selection number of ranks

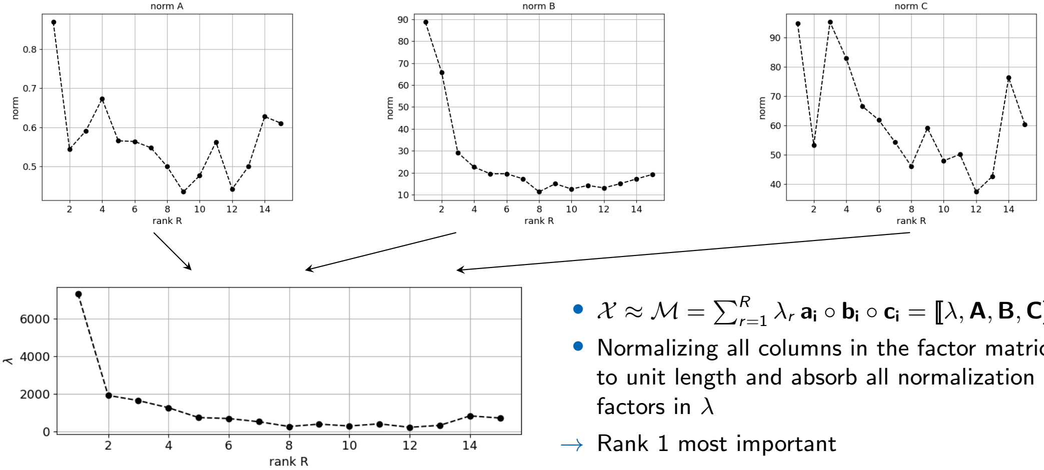
- Basic metric to compare models: normalized reconstruction error $\frac{\|\mathcal{X}-\mathcal{M}\|_F^2}{\|\mathcal{X}\|}$



- Convergence for all number of ranks R
- Reconstruction error $< 10\,\%$ for all ranks R > 13

$\rightarrow$ Decision for R=15 model

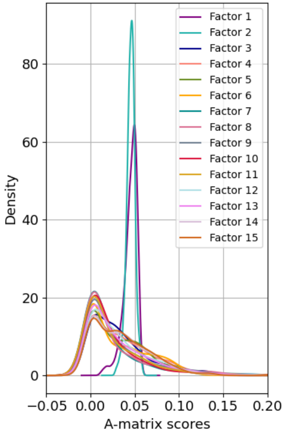
Normalization factors



- $\mathcal{X} \approx \mathcal{M} = \sum_{r=1}^{R} \lambda_r \, \mathbf{a_i} \circ \mathbf{b_i} \circ \mathbf{c_i} = [\![\lambda, \mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$
- Normalizing all columns in the factor matrices to unit length and absorb all normalization factors in λ

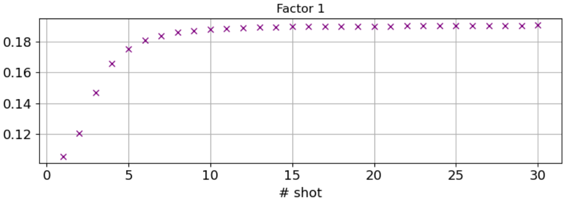
$\rightarrow$ Rank 1 most important

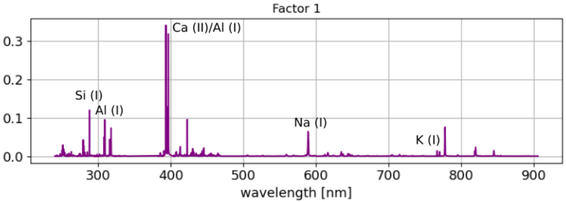
Factors interpretation



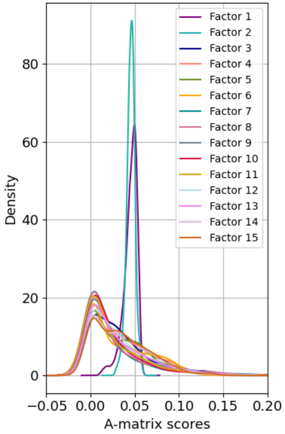
→ All analysis points have non-zero A-matrix scores on factor 1 and 2

Factors interpretation



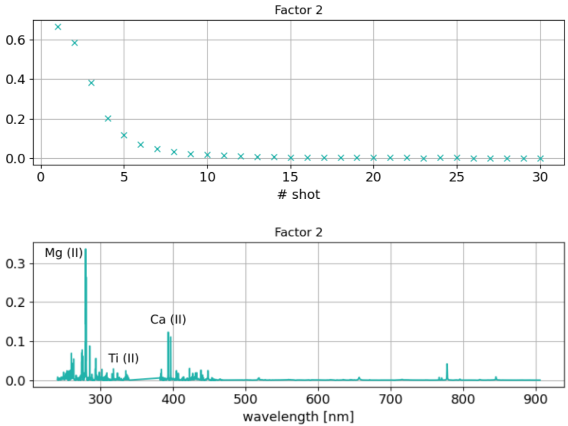


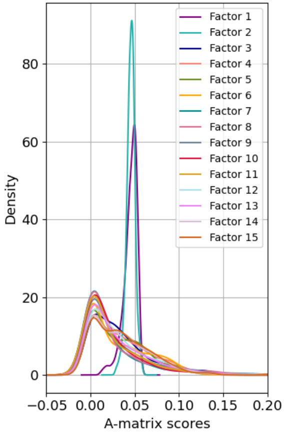
→ Baseline felsic composition



→ All analysis points have non-zero A-matrix scores on factor 1 and 2

Factors interpretation





→ Dust contribution

→ All analysis points have non-zero A-matrix scores on factor 1 and 2

Example: Factor 2

A-matrix scores factor 2

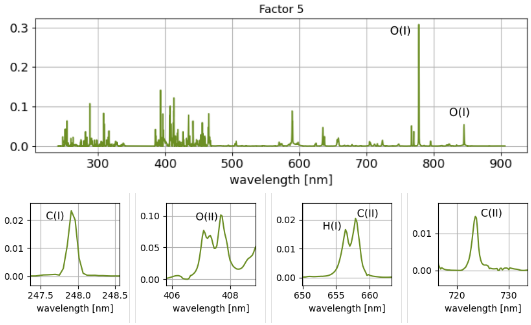



→ Targets with a stronger dust coverage have higher A-matrix scores on factor 2

→ Interpretation of factor 2 as the "dust contribution" factor seems correct

Influence of experimental conditions

Factor 5: Atmospheric contribution

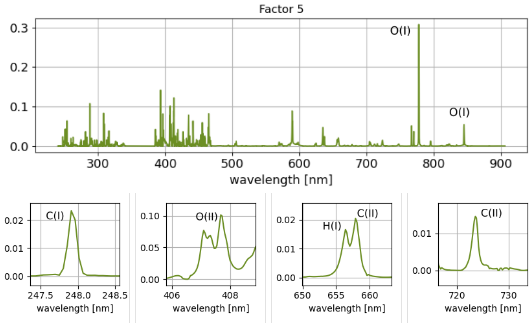


- Relatively strong emission lines of C and O
- Martian atmosphere is dominated by CO_2

→ Possibly not optimal laser coupling or focus

Influence of experimental conditions

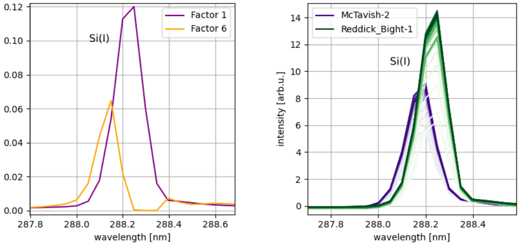
Factor 5: Atmospheric contribution



- Relatively strong emission lines of C and O
- Martian atmosphere is dominated by CO_2

→ Possibly not optimal laser coupling or focus

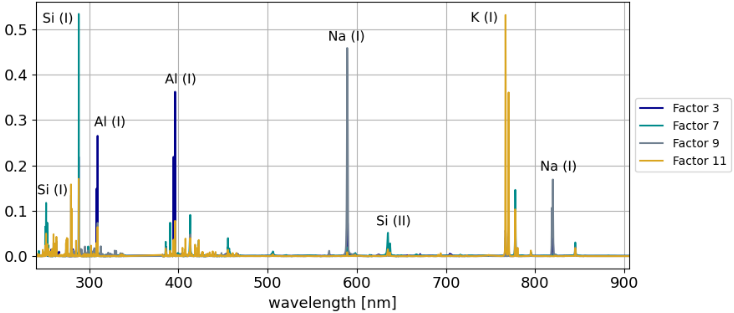
Factor 6: Wavelength calibration



- Line positions are shifted on factor 6 in comparison to other factors
- *McTravish2* has high and *Reddick_Bight1* low factor 6 score values

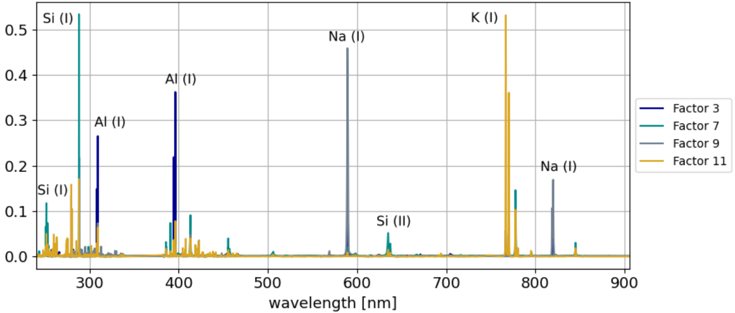
→ Slightly shifted wavelength calibration

Variations in felsic composition

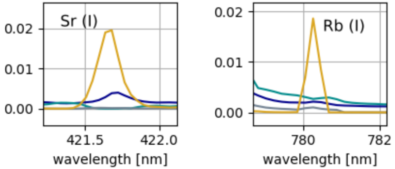


- Factor 3: Strong Al emission lines
- Factor 7: Strong Si emission lines
- Factor 9: Strong Na emission lines
- Factor 11: Strong K emission lines

Variations in felsic composition



- Factor 3: Strong Al emission lines
- Factor 7: Strong Si emission lines
- Factor 9: Strong Na emission lines
- Factor 11: Strong K emission lines



Summary and outlook

Summary

- Different types of factors were identified: experimental conditions or real compositional variations among the shots
- The main contribution comes from the baseline felsic composition in the selected dataset
- Starting from a group of similar compositions, the method can be used to observe finer effects

Outlook

- Apply to ChemCam datasets with more compositional variations
- Also interesting for hyperspectral data!

Summary and outlook

Summary

- Different types of factors were identified: experimental conditions or real compositional variations among the shots
- The main contribution comes from the baseline felsic composition in the selected dataset
- Starting from a group of similar compositions, the method can be used to observe finer effects

Outlook

- Apply to ChemCam datasets with more compositional variations
- Also interesting for hyperspectral data!

Thank you for your attention!