**UNSUPERVISED CLUSTERING OF MARS ROCK SOUNDS.** U. Alam\(^1\), J. Lasue\(^2\), S. Maurice\(^2\), B. Chide\(^2\), R.C. Wiens\(^3\)

\(^1\)E4R, Thoughtworks, Pune, India (ujjaini.alam@thoughtworks.com), \(^2\)IRAP-CNRS, Toulouse, France, \(^3\)EAPS, Purdue University, West Lafayette, IN, USA

**Introduction:** The NASA Perseverance rover has been exploring Jezero crater on Mars for three years, coring and sealing a diversity of rocks for a possible sample return to Earth. The rocks analyzed include crater floor deposits of igneous origin with local aqueous alteration (Máaz formation), an olivine-rich cumulate (Séítah formation) and alluvial sedimentary deposits of the Jezero delta fan [1, 2]. Amongst the payload used to analyze the geological samples, the SuperCam instrument includes a microphone to complement the Laser-Induced Breakdown Spectroscopy (LIBS) analysis of the geological targets [3, 4, 5]. In this study, we describe how clustering the microphone acoustic data can provide complementary information likely related to rock surface properties.

**Datasets:** SuperCam is a multi-technique remote sensing instrument able to acquire high resolution color images, obtain chemical composition with LIBS, and infer mineralogical information from passive visible/near-infrared and Raman spectroscopy. It also includes a microphone, mostly used to characterize the atmosphere dynamics in conjunction with LIBS, as shown hereafter [3, 4]. The microphone is a Knowles EK-23132 electret sensor using a charged membrane to detect pressure fluctuations with a sensitivity of 0.6 V/Pa and bandwidth from 20 Hz to ~50 kHz. Two sampling frequencies, 25 kHz & 100 kHz, are used [5].

In a typical LIBS analysis, a series of 30 powerful laser pulses ablates the surface of targets at a distance of a few meters, inducing a plasma. The spectroscopy of the plasma light gives information on the elemental composition of the target. As the plasma expands, it creates a shock wave that propagates at the local speed of sound (~240 m/s). This acoustic wave, recorded by the microphone, has been shown to be related to properties of ablation, crater morphology, rock hardness and possible mineral properties [6, 7, 8].

From the online Planetary and Data System (PDS), we used all LIBS shot-to-shot data acquired since the landing of Perseverance, sol 0 till sol 777 (a “sol” is a day on Mars). Each laser shot creates an acoustic shock wave that is recorded at 100 kHz as a time series containing 6000 data points. There are 10 location analyses per target and each contains typically 30 shots, of which the first 29 acoustic signals are recorded.

**Method:** We first cleaned the data by applying a band-pass frequency filter between (2, 25) kHz, which corresponds to the shock wave signal [9], on the time series. We then identified the first peak of the time series as the moment of hitting the target and cropped the time series from 100 steps before to 1500 steps after the peak. This avoids explicit dependence on the focal distance and also eliminates noise from the latter part of the data. These cropped time series were then Fourier transformed to amplitude spectra in frequency space. To reduce interference from wind noise, the spectra were cut-off at 15 kHz. They were also normalized to the total energy under the spectrum. The spectra for acoustic shock waves tend to show strong characteristic behavior which is easier to classify into different groups as opposed to the time series.

**Clustering:** A total of 122717 cropped and cleaned amplitude spectra were now available for further analysis. Since previous information on these acoustic waves was sparse, unsupervised machine learning methods seemed appropriate to study them. A recursive clustering pipeline was created using FAISS (a library for efficient and fast similarity search of dense vectors [10]). The optimal number of clusters was calculated using measures of inter and intra-cluster distances e.g. Davies-Bouldin Index. This pipeline was very successful in consistently separating the spectra into 9 distinct groups. On the first clustering, an outlier group consisting mostly of noisy, out of focus spectra was isolated. On the next clustering, after removal of the out-of-focus spectra, we were able to separate out eight distinct groups, one with more soil-like behavior, and others with rock-like behaviors, but with clear differences in the amount of energy under the second peak of the spectra. Examples of some typical spectra are shown in fig. 1.

![Figure 1: Examples of typical frequency spectra: Noisy (left), Soil-like (middle) and Rock (right).](image-url)

**Results:** We could not find any clear direct correlation between these clusters and various physical properties of the targets and the environment. The features checked for correlation include: the chemical composition of the targets (Major Oxide Composition values of SiO\(_2\), TiO\(_2\), Al\(_2\)O\(_3\), FeOT, MgO, CaO, Na2O, K2O in the target), the focal distance, the local solar time and solar longitude, the pointing elevation and azimuth. However, we noted that most of the shots in a single location analysis tended to belong to the same cluster, and of the 10 location analyses of each target, most belonged to the same cluster. Our analysis therefore...
suggests a correlation between the nature of the target, its surroundings, and the clustering.

Figure 2 shows a visualization of characteristic targets in the different clusters. In the central figure, the spectra are represented as points on a 2-D projection space [11] using t-SNE (t-distributed stochastic neighbor embedding, an unsupervised, non-linear, dimensionality reduction technique). Each color corresponds to a particular cluster. A representative target for each cluster is illustrated using the RMI mosaic images from the SuperCam camera giving the geological context at the sub-millimeter scale. The column to the left of each target image represents the 10 LIBS locations analyses on 10 vertical lines, each colored dot on the line corresponding to the laser shots from shot 1 on the left to shot 29 on the right.

One immediately notices that each set of shots at a given location tend to belong to a single cluster, and that most shots on a target also belong to a single cluster. To the right of the plot, the green targets correspond to soils or fine materials (like core tailings). The mainly yellow target Snowy Mountain is a mixture of pebbles and fine soils and this shows in the analysis color distribution where some locations analyze single pebbles while others analyze a mixture of the two. In the center of the map and at the bottom are located rocky targets, the red, orange and pink targets corresponding to rocks with apparent decreasing grain sizes. On the left are located apparently sedimentary rocks, large-grained in gray, and fine-grained in brown. Finally on top left, the black cluster seems to correspond to rocks with apparent coatings on their surface.

Conclusions: Using the first two years of Perseverance data, we have demonstrated that acoustic signal from LIBS shots contains information complementary to the chemistry of the rock. Unsupervised clustering of the LIBS shots frequency spectra appears correlated with the surface properties of the targets analyzed as illustrated in the figure 2 projection.

Future work will include analysis of the correlations between the surface properties of the targets and the clustering to quantify the relation between the two and to examine whether these properties can be predicted based on the clustering.

Acknowledgments: The engineers and scientists of the Perseverance rover are thanked for their generous contributions to this work. This work was supported by CNES in France and by the Mars Exploration Program in the US. We acknowledge interesting discussions and the use of hardware at Engineering for Research (E4R), Thoughtworks, India.

References:

Figure 2: Nonlinear projection map of the acoustic data points highlighting in color the 8 clusters obtained by the analysis. Each cluster is associated with a representative example target presenting different surface properties as detailed in the text.