Decision support tools for planetary exploration using bespoke databases and machine learning.

F. Sauro[1,2], I. Drozdovskiy[1], S.J. Payler[1, 3], L. Turchi[1], L. Bessone[1]; [1]European Space Agency (EAC, Linder Hoehe, 51147 Köln, Germany), [2] Department of Biological, Geological and Environmental Sciences, Italian Institute of Speleology - Bologna University, Italy, [3]Agenzia Spaziale Italiana, Rome, Italy,

Introduction: During the Artemis programme, astronauts will once again explore the surface of the Moon to collect geological samples for return to Earth. During these EVAs, limited supplies will mean rapid decisions need to be taken in order to identify and select the best samples. For the astronauts and science support teams, this leaves little time to perform detailed analysis on new information and make quality informed science decisions. The Artemis III science definition team report [1] highlights the need for "*real-time transmission of data from in situ science instrumentation that provides documentation for site characteristics and enables a science support team (backroom, operations center, etc.) to support EVA operations with (near) real-time feedback to the crew when necessary on science decision-making*". In order to assist with this decision making, during the PANGAEA astronaut training we have developed a set of tools to provide geological decision support to both the science backroom and astronaut teams involved in an EVA. Here we present two of the main parts: an analytical database of all known minerals on the Moon, Mars and asteroids for planetary exploration; and a set of machine learning algorithms that take advantage of this database to identify minerals from spectra and provide relevant information. These tools have been tested by astronauts, planetary scientists and operations engineers in the context of ESA's analogue testing programme, PANGAEA-X [2].

Database repository for geological decision support: The PANGAEA Mineralogical Database (MDB) [3] is focused on providing information for scientific decision support to space and ground teams during geological exploration. It achieves this through two core products: a catalogue of petrographic information and an analytical library. The catalogue contains information on all currently known minerals identified on Moon, Mars, and found primarily, or exclusively, within meteorites. The catalogue is envisioned to provide essential analytical in-field information for each mineral to assist in rapid identification and understanding of significance in real time geological exploration. Each mineral entry includes: IMA recognised Name, Chemical Formula, Mineral Group, surface abundance on planetary bodies, geological significance in context of planetary exploration (occurrence, environmental conditions, marker for important processes, etc.), number of collected VNIR and Raman spectra and their spectral discoverability and the possible spectral features. In addition, supplementary characteristics for each mineral that may help with its identification are included, such as chemical abundances calculated from the known empirical chemical formula, as well as basic physical properties such as hardness, specific gravity, and crystal system. The database was compiled through literature research and cross-validation ("out-of-sample" testing) of all characteristic mineral information (including the flagging of potentially erroneous data). Including metadata in the MDB, such as the significance of particular minerals, is key to it providing information relevant to geological decision making.

In order to match this catalogue of information to relevant minerals detected during an EVA, it was seen as important to include the spectral information associated to the different minerals in the database, so they can be used in conjugation with handheld analytical tools. Therefore, the PANGAEA Mineralogical Database also includes a customised library of analytical data. This covers four analytical methods: reflective Visual-to-Near- & Shortwave-Infrared (VNIR), Raman vibrational (molecular) spectroscopy, Laser-Induced-Breakdown (LIBS), and X-Rays Fluorescence (XRF) atomic spectroscopy. This library also includes a set of standard spectra, which is used for evaluating the detectability of minerals with different analytical methods. Part of this archive consists of spectra collected from available open access on-line catalogues, such as experimental Raman (RRUFF) and VNIR (USGS, RELAB, ECOSTRESS), as well as from our own and our collaborators bespoke spectroscopic measurements (VNIR, Raman, LIBS & XRF) of planetary analogue minerals taken from different collections, and synthetic spectral libraries, such as LIBS NIST; see [3] and references therein. Only the high-quality spectra of confirmed mineral samples are included as determined by the quality flag in the original database, or by our statistical evaluation of the within-class spectra. We removed outliers from all available single-method spectra of each mineral by finding the average spectrum for each class and subsequently removing spectra that significantly deviated from the average spectrum [3]. This ensures that the set is not skewed by extremely divergent spectra due to the random instrumental artefacts or misclassification of the sample.

Machine Learning for matching field measurements accurately to database entries: To properly leverage the information contained in the MDB during an EVA, spectra collected in the field by astronaut teams must be rapidly matched to entries in the database to identify minerals and notify the backroom and astronaut of their significance. To do this accurately, we developed identification methods that combine two types of material characteristics: mineral structure (obtained with VNIR and Raman spectra) and chemical composition (derived from XRF and LIBS spectra) using a ML based approach. ML was chosen for several reasons: it is fast and accurate when developed properly, it can handle multimethod datasets, and the accuracy can be progressively improved by adding new training data without losing the recognition speed.

To maximise the accuracy of the ML matching to the MDB, we evaluated various Machine Learning approaches used to identify mineral species from single analytical methods (Raman, VNIR or LIBS) and developed a flexible and modular algorithm that can classify minerals either from standalone or combinations of spectroscopic methods. The flow diagram detailing this methodology is shown in Figure 1.

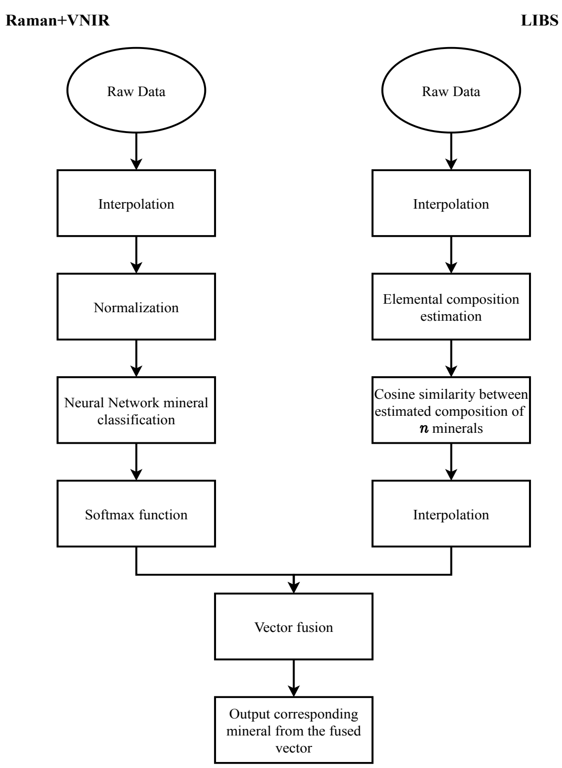

Figure 1: Simplified flow diagram showing our method for recognising minerals from combined Raman/VNIR and LIBS spectra.

Our new approach was evaluated using our own internal archive of analytical data, as well as in some cases, on publically available spectroscopic datasets. Our cross-validation tests show that using our ML methods with multi-method spectroscopy paired with ML paves the way towards rapid and accurate characterisation of minerals [4] (see Figure 2), as well as improving the quantification of mineral abundances in rocks and soils using ML-based spectral unmixing.

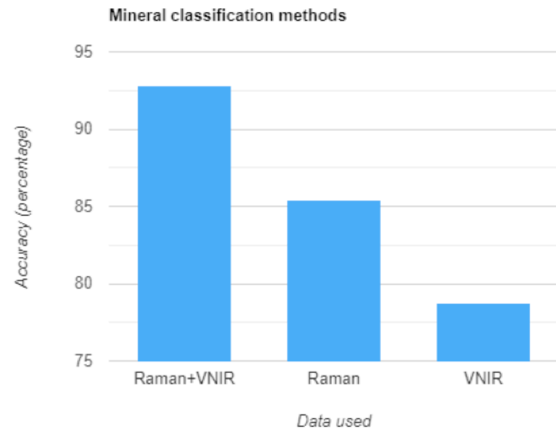

Figure 2: Mineral predictions rates from combined Raman and VNIR spectra compared to single-method spectroscopy using ML techniques.

Making decision support tools deployable: Although able to be operated standalone, the true power of the MDB and ML tools will come from them being integrated into a deployable system used by the backroom science teams and the astronauts in the field to centralise and analyse information collected on the lunar surface. For this reason, the MDB and ML tools are being integrated into the PANGAEA Electronic Fieldbook (EFB) [5]. The EFB is a deployable system enabling scientific documentation of field traverses and geological sampling. It provides situational interaction with science and mission support teams through exchange of contextual data. The EFB can interface with handheld instrumentation intended for planetary exploration, feeding the ML classification tools with mineral spectra, which are then matched to entries in the MDB to provide relevant geological information in real-time. The EFB Tool Suite, a combined set of handheld analytical tools, and the instrument agnostic nature of the PANGAEA MDB and ML tools, will enable fast and reliable in-situ recognition of rocks and minerals, crucial for decision support during future human and robotic planetary surface exploration missions.

References: [1] Artemis III Science Definition Team Report, 2020, NASA. [2] ISECG – Global Exploration Roadmap, 2020. [3] Drozdovskiy, I. et al. 2020, *Data in Brief*, 31, 105985 https://doi.org/10.1016/j.dib.2020.105985. [4] Jahoda, P. et al. 2020, *The Analyst*, May. https://doi.org/10.1039/D0AN01483D. [5] Turchi L. et al. *Planetary Space Science*.