MIST: AN ONLINE TOOL AUTOMATING MINERAL IDENTIFICATION BY STOICHIOMETRY IN GEOCHEMICAL DATASETS K. L. Siebach[1], E. L. Moreland[1], G. Costin[1], and Y. Jiang[1], [1]Department of Earth, Environmental and Planetary Sciences, Rice University, Houston, TX (corresponding: ksiebach@rice.edu).

Overview: Identifying minerals is a prerequisite to interpreting geologic history from samples. A variety of techniques for mineral identification are available, but one method commonly used for individual grains or crystals in a sample is to obtain high-resolution geochemical information of the grain and compare element ratios to known minerals, i.e., stoichiometry.

The Problem: Classically, the process of identifying minerals by stoichiometry requires some understanding of expected minerals so that geochemical measurements can be compared with previous examples of those minerals. In datasets with very little information about likely minerals, a series of binary and ternary diagrams can be used to identify geochemical endmembers and then check their stoichiometry. When minerals are not expected or are very complex, they may not be identified at all.

Our Solution: MIST (Mineral Identification by Stoichiometry) is a first principles based computational algorithm available for use online that identifies geochemical observations with stoichiometric elemental ratios that match real mineral compositions [1]. MIST is a decision-tree model that filters the geochemical analyses based on elemental ratios and normalized oxide percentages. It follows mineral classification rules based on real mineral structures and stoichiometries, including typical elemental substitutions. Mixtures of phases or chemical ratios that do not fit a mineral structure are identified as mixtures.

Datasets for MIST: Identifying minerals by stoichiometry works when the geochemical measurement spot size is smaller than the size of the mineral. Ideally, the observation would be of a single crystal, but a monomineralic grain will also have averaged mineral stoichiometry. After this relative size requirement, stoichiometric rules are agnostic to the source of geochemical data. Analysis tools that could be used include EPMA (Figure 1), XRF (e.g., PIXL instrument on Perseverance), EDS, ICP-MS, and LIBS (e.g., ChemCam and SuperCam instruments).

In-situ Spacecraft Instruments: The MIST tool was developed for the PIXL instrument on Mars 2020, which is a high resolution XRF mapper with a spot size of ~120 µm, and MIST has been very successful identifying minerals in coarse-grained in-situ samples on Mars [see Moreland et al., this meeting]. In particularly coarse samples, like igneous phenocrysts, geochemical techniques with even lower spatial resolution may be used (e.g., [2]).

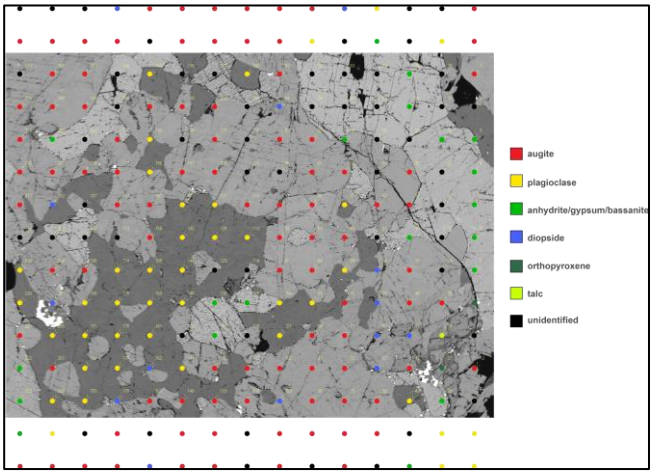

Figure 1. Example MIST results from a grid of EPMA measurements 200 µm apart on a thin section from the Bushveld igneous complex. Augite, plagioclase, and anhydrite grains make up most of the sample. Analyses on fractures or mineral mixtures were unidentified.

MIST Approach: MIST accepts geochemical analyses as wt% oxides. The algorithm normalizes the input to 100 wt% and uses the normalized oxides to identify the mineral class (e.g., silicates, oxides, carbonates, etc, see Figure 2). The cations are grouped based on geochemical affinities and are converted into atoms per formula unit given the number of oxygens for the mineral group (or sub-group). The compositions must pass through a series of conditions to be classified into mineral groups and sub-groups until the 5th order species level classification. At each stage, the filters match the requirements imposed by mineral structures within the mineral group or sub-group; mineral structures allow some substitutions and some solid solutions, but others are excluded because a change in cations would require a different mineral structure or

MIST mist.rice.edu

We present MIST as an online tool for rapid identification of rock-forming minerals and mineral crystal chemistry in any high-resolution chemical dataset on Earth or other worlds. This tool both simplifies and standardizes the process of identifying and reporting mineral stoichiometry and crystal chemistry, making it particularly useful for high-value shared astromaterial samples or in-situ robotic observations.

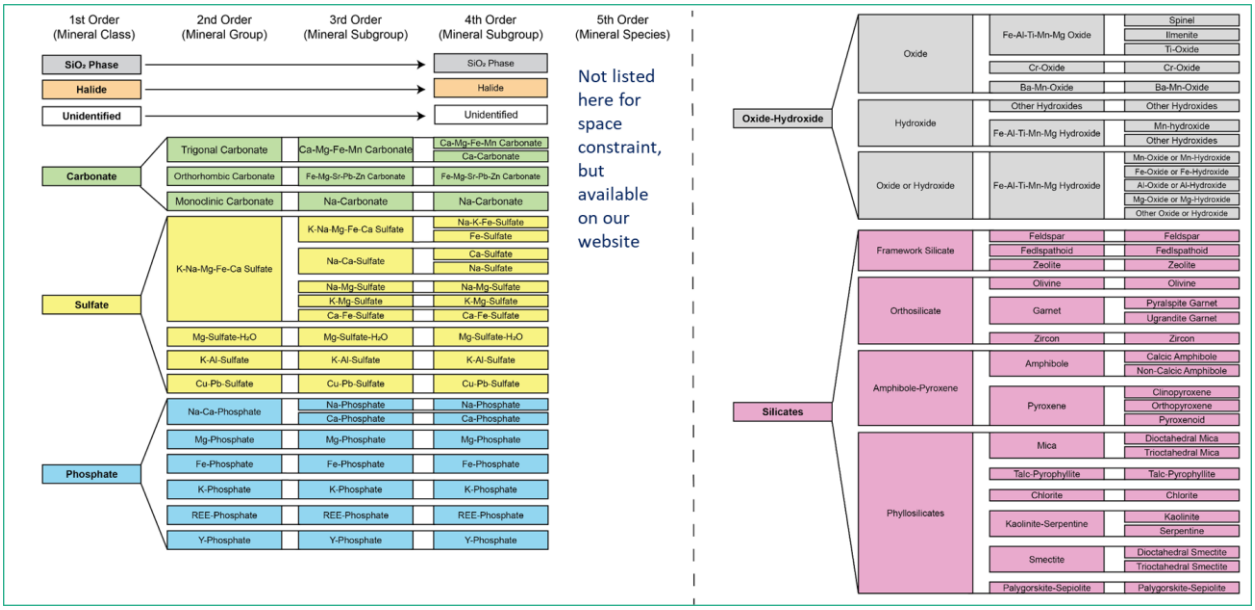

Figure 2. Diagram showing MIST model hierarchical classification approach; classification of mineral species is based on constraints defined by real mineral structures and traditional mineral classification.

group. This hierarchical classification approach allows recognition of near-misses or potentially of mineral species that are not yet coded into the algorithm but are in a mineral subgroup (e.g., trioctahedral smectites).

Importantly, species classification is based on a well-calibrated set of terrestrial analyses of the minerals, and expected ratios between elements are allowed to vary within typically 95-105% of expected values so that the algorithm recognizes real, imperfect mineral compositions rather than ideal mineral compositions. Each identified mineral species is checked against two conditions: (1) total number of cations (ΔABCDT) for the oxygen number in the ideal formula of the mineral as reported by IMA in the RRUFF database [3], and (2) the number of cations in the tetrahedral position (ΔT) of the ideal formula (e.g. Si+Al(IV) for silicates, or other cation ratio checks are designed for salts and oxides). Mineral species names are reported only when both checks are passed.

MIST development data set: As of writing this abstract, MIST can recognize 249 mineral species containing oxygen in the formula (the full list is available on our website). Minerals that cannot be distinguished by stoichiometry alone are grouped as one mineral (e.g., antigorite/lizardite/chrysotile). We have not included reduced mineral species at this time.

MIST output: Model output includes the mineral group and three sub-group classifications (as shown in figure 2), and mineral species identification. If a composition meets some classifications (e.g., silicate class), but then does not fit into any groups, then it is just reported as "silicate" and we interpret this as a mixture of minerals dominated by silicates.

The mineral output also includes the actual stoichiometric formula derived from the measurement and relevant solid solution endmembers. For example, in Fig. 1, the identified plagioclase was An76, the augite was $Wo_{45}Fs_{10}En_{45}$, and the diopside was $Wo_{46}Fs_{9}En_{45}$.

MIST as a tool for Database Standardization: Data quality testing and standardization is a significant problem in geochemical databases and compilations of geochemical results [4]. MIST can help standardize the procedures for recognizing when a geochemical measurement stoichiometrically matches a mineral. For example, we tested 222,543 geochemical compositions classified as olivine in the GEOROC chemical database [5] and found that only 21.5% matched an olivine composition; 38% were chemically consistent with serpentine minerals and 33.1% were classified as trioctahedral smectites. This is a known issue with a database with records spanning decades, but if each individual user has to filter the datasets themselves, they likely select different compositions based on the individual's perception of the tolerance of real mineral variations. MIST can standardize this filtering process.

References: [1] Siebach, K.L., et al., *IMA* 2022. Lyon, France. [2] Payré, V., et al., *JGR*, 2020. 125(8). [3] Lafuente, B., et al., RRUFF, in *Highlights in Mineralogical Crystallography*, T. Armbruster and R.M. Danisi, Ed. 2015: Berlin, Germany. p. 1-30. [4] Lehnert, K., et al., OneGeochemistry in *CODATA2019*. 2019: Beijing, China. [5] GEOROC Compilation: Minerals. 2022, GRO.data