




# Algorithms for SEM-EDS Mineral Dust Classification

Austin M. Weber  <sup>1,2</sup>

<sup>1</sup> Byrd Polar and Climate Research Center, Columbus, Ohio <sup>2</sup> School of Earth Sciences, The Ohio State University, Columbus, Ohio

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

## Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Open Journals](#) 

## Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

## License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

## Summary

A repository of MATLAB functions has been assembled to work with scanning electron microscope (SEM) and energy dispersive x-ray spectrometry (EDS) data, with the objective of providing researchers with tools to quickly and accurately determine the mineral compositions of dust particles. The repository includes a novel machine learning classifier and three additional sorting algorithms that have been transcribed from the peer-reviewed literature, as well as functions for importing and visualizing x-ray energy spectra, and functions for extracting and working with hidden metadata in SEM images.

## Statement of need

The mineral classification of single particles (i.e., dust) has major implications for atmospheric and glaciological research. For instance, there is a great need for building mineral dust datasets that can be used to improve the predictive ability of climate models ([Gieré, 2010](#)). Mineral dust observations are also severely limited in many glaciated mountain regions, meaning that the dust's radiative impact on glacier melt rates cannot be easily determined ([Gilardoni et al., 2022](#)).

One of the most common methods for reconstructing the mineral dust history of a location or region has been the use of SEM-EDS to analyze single particles extracted from ice cores ([Donarummo Jr et al., 2003](#); [Nagatsuka et al., 2021](#); [Wu et al., 2016](#)). However, the process of classifying individual particles based on their EDS data is highly laborious and prone to error due to the complexity of certain minerals' elemental signatures. The adoption of numerous classification techniques is typically required to improve the certainty of the classifications before finalizing the results. Many have used the dichotomous key in [Severin \(2004\)](#) to compare EDS spectra to that of mineral standards, but this protocol requires access to a physical or digital copy of the book and cannot be automated. Others have used published sorting algorithms like the one by [Donarummo Jr et al. \(2003\)](#), but these algorithms are limited in the number of minerals that they can identify and, to the author's knowledge, each of these algorithms is only available in print and not as a functional computer program.

Therefore, to satisfy the needs of Earth and atmospheric scientists seeking to rapidly classify mineral particles from EDS data with a high degree of certainty, an open-source repository of relevant scripting algorithms has been developed for MATLAB (along with several homologous functions written in the Julia programming language). This library of functions gives users the tools to work with SEM-EDS data and to perform mineral classification with ease.

## Methods

### Mineral Dust Classification

The function library introduces a novel supervised machine learning model (the “Weber algorithm”) that is trained to recognize 18 common minerals from EDS net intensity data and exhibits an overall accuracy exceeding 99% (see [Supplemental Information](#)). The library also includes functions for three non-machine learning algorithms that have been painstakingly transcribed from the literature into MATLAB and Julia code. This includes the sorting algorithm from Donarummo Jr et al. (2003) which can be used to classify 16 unique aluminosilicate minerals based on EDS net intensity data, the Panta algorithm (Panta et al., 2023) which can be used to classify 17 unique minerals and 6 mineral groups based on EDS atom percent data, and the Kandler algorithm (Kandler et al., 2011) which can be used to categorize EDS atom percent data into broad mineral groups, classes, and refractive indexes.

The four mineral classification algorithms are available as individual functions (`weber_classification`, `donarummo_classification`, `panta_classification`, and `kandler_classification`) as well as packaged together in a single function (`eds_classification`). Each function has clearly written documentation that can be viewed using the MATLAB help or doc commands. The general syntax, however, is very basic:

```
minerals = eds_classification(data, Algorithm="Name")
```

where `data` is a table of EDS data and “Name” specifies the desired algorithm; i.e., “Weber” (default), “Donarummo”, “Kandler”, or “Panta”. The output `minerals` is a categorical vector containing the mineral assignments for each row in the input table.

### Visualizing EDS Spectra

For individuals who want to compare the EDS spectra of their single particle measurements to the standard spectra in Severin (2004), it will be necessary to visualize each energy spectrum along with labels for the major characteristic x-rays. Most EDS software save x-ray energy data in the EMSA (.msa) file format which is compiled differently depending on the software and therefore makes it difficult to simply import the data using a base MATLAB function such as `readtable`. While methods for visualizing EDS spectra are available for the Julia programming language (Ritchie, 2022) and as a desktop Java application (see [DTSA-II](#)), there are currently no simple alternatives to perform these tasks in MATLAB.

To address this limitation, the EDS mineral dust classification library also provides functions for importing .msa files into MATLAB and for visualizing the EDS spectrum contained in a file. Just use the `xray_plot` function to view the spectrum and the `xray_peak_label` function to label the characteristic x-rays with the most likely element.

```
plt = xray_plot('spectrum.msa');  
xray_peak_label(plt)
```

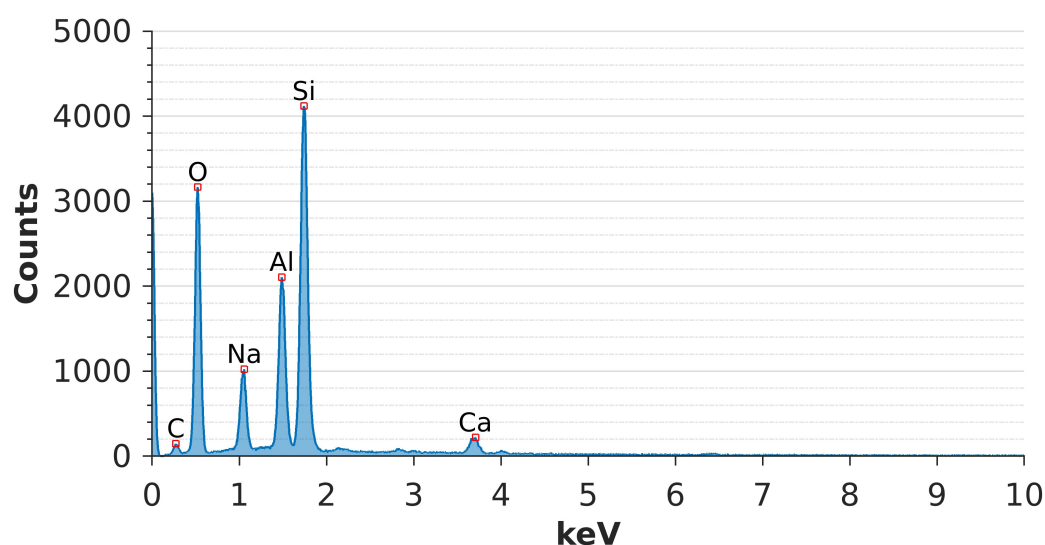


Figure 1: xray plot visualization

Users may also import x-ray energy data directly into the MATLAB workspace along with the corresponding metadata using `read_msa`.

```
[data,metadata] = read_msa('spectrum.msa');
```

## Extract metadata from SEM images

Images collected with SEM software are usually saved as 16-bit TIFFs and contain valuable metadata regarding the image itself as well as the operating conditions of the instrument at the time the image was taken. But while MATLAB users with access to the [Image Processing Toolbox](#) can typically view image metadata with the `imfinfo` function, the metadata stored in an SEM image are not so easily accessed. This is because the image metadata become compressed into a single character vector that is nested within several structure objects. Many string operations or regular expressions are therefore needed to access the relevant metadata. To combat this, a function called `get_sem_metadata` has been written to automatically extract the useful information from a SEM image.

```
md = get_sem_metadata('filename.tif');
```

## Acknowledgements

I thank Paul Pohwat and the Smithsonian Institution for contributing the mineral standards used in this project. Thanks are also extended to Lonnie Thompson, Dan Veghte, and Julia Sheets for their various contributions in both advising and teaching. The electron microscopy work used to train the machine learning algorithm was performed at the Center for Electron Microscopy and Analysis (CEMAS) in the College of Engineering at The Ohio State University. Additional electron microscopy was carried out at the Subsurface Energy Materials Characterization and Analysis Laboratory (SEMCAL) in the School of Earth Sciences at The Ohio State University.

## Copyrights

The software repository is made available under a MIT license, meaning that users are free to modify and distribute the software without restriction. However, the *intellectual copyrights* for the Donarummo, Kandler, and Panta algorithms belong to their original creators. If you use any of these algorithms in your research, **please** also cite the original references as appropriate.

## References

- Donarummo Jr, J., Ram, M., & Stoermer, E. F. (2003). Possible deposit of soil dust from the 1930's US dust bowl identified in Greenland ice. *Geophysical Research Letters*, 30(6). <https://doi.org/10.1029/2002GL016641>
- Gieré, R. (Ed.). (2010). Atmospheric Particles. *Elements*, 6(4). <https://www.elementsmagazine.org/atmospheric-particles/>
- Gilardoni, S., Di Mauro, B., & Bonasoni, P. (2022). Black carbon, organic carbon, and mineral dust in South American tropical glaciers: A review. *Global and Planetary Change*, 213, 103837. <https://doi.org/10.1016/j.gloplacha.2022.103837>
- Kandler, K., Lieke, K., Benker, N., Emmel, C., Küpper, M., Müller-Ebert, D., Ebert, M., Scheuvs, D., Schladitz, A., Schütz, L., & others. (2011). Electron microscopy of particles collected at Praia, Cape Verde, during the Saharan Mineral Dust Experiment: Particle chemistry, shape, mixing state and complex refractive index. *Tellus B: Chemical and Physical Meteorology*, 63(4), 475–496. <https://doi.org/10.1111/j.1600-0889.2011.00550.x>
- Nagatsuka, N., Goto-Azuma, K., Tsushima, A., Fujita, K., Matoba, S., Onuma, Y., Dallmayr, R., Kadota, M., Hirabayashi, M., Ogata, J., & others. (2021). Variations in mineralogy of dust in an ice core obtained from northwestern Greenland over the past 100 years. *Climate of the Past*, 17(3), 1341–1362. <https://doi.org/10.5194/cp-17-1341-2021>
- Panta, A., Kandler, K., Alastuey, A., González-Flórez, C., González-Romero, A., Klose, M., Querol, X., Reche, C., Yus-Díez, J., & Pérez García-Pando, C. (2023). Insights into the single particle composition, size, mixing state and aspect ratio of freshly emitted mineral dust from field measurements in the Moroccan Sahara using electron microscopy. *Atmospheric Chemistry and Physics*, 23(6), 3861–3885. <https://doi.org/10.5194/acp-23-3861-2023>
- Ritchie, N. W. M. (2022). Reproducible Spectrum and Hyperspectrum Data Analysis Using NeXL. *Microscopy and Microanalysis*, 28(2), 478–495. <https://doi.org/10.1017/S143192762200023X>
- Severin, K. P. (2004). *Energy Dispersive Spectrometry of Common Rock Forming Minerals*. Kluwer Academic Publishers. <https://doi.org/10.1007/978-1-4020-2841-0>
- Wu, G., Zhang, X., Zhang, C., & Xu, T. (2016). Mineralogical and morphological properties of individual dust particles in ice cores from the Tibetan Plateau. *Journal of Glaciology*, 62(231), 46–53. <https://doi.org/10.1017/jog.2016.8>