

Machine Learning Approaches for Analysis of Multiscale Imaging Data for Atmospheric and Soil Particles

Swarup China^{1*}, Sean Colby¹, Anil Krishna Battu¹, Tamas Varga¹

¹ Pacific Northwest National Laboratory, Richland, WA, USA.

* Corresponding author: swarup.china@pnl.gov

Bioaerosol and soil organic particles are ubiquitous in the atmosphere, and they are multicomponent and complex in nature [1-4]. Biological particles contribute to a major fraction of coarse mode particles, especially in the tropical areas, such as the Amazon basin, contributing up to 80% of the particle mass concentration. The way components are mixed in an aerosol sample is referred to as its mixing state [5]. These particles display mixed inorganic (mineral) and organic (soil organic matter, root, microbial) components. The mixing state of particles is crucial to evaluate because it impacts several important environmental processes such as warm and cold cloud formation and radiation budget. Soils are also a mixture of inorganic (mineral) and organic (soil organic matter, SOM) components, which chemically associate with one another. Mixing states in aerosols are accompanied by chemical reactions across solid-liquid-gas interfaces. Full characterization of both classes of particles require a multiscale approach. Lower resolution (mm and micrometer scale) imaging provides larger-scale morphology of particles and the overall distribution of pores vs solid phases. When imaging using x-ray computed tomography (XCT), one learns about the 3D distribution of different density phases, including organic vs inorganic. High-resolution imaging is used next for semi-quantitative evaluation of mixing state in aerosol particles, and mineral-SOM associations in soil particles.

In this study we utilized elemental composition and microscopy images of thousands of atmospheric particles acquired by the computer-controlled scanning electron microscope equipped with an energy-dispersive X-ray spectrometer to compute the mixing state of atmospheric particles. We also used large amount of spectro-microscopy data such as Scanning Transmission X-ray Microscopy (STXM) and Near-Edge X-ray Absorption Fine Structure (NEXAFS) to analyze in a semi-automated way to extract spectroscopic information on specific particles/sub-particles we are interested in.

A 2D convolutional neural network (convnet) is used to model the relationship between low resolution imaging data and higher resolution spectroscopy data, with the former as training input and the latter as target output. For region of interest extraction and particle classification, separate 2D convnets are used to annotate parts of the input image for follow-up high resolution spectroscopy and to perform particle classification, respectively. Target images for ROI extraction and classification are utilized from previous, semi-automated efforts wherein ROIs and classes are already defined. Annotated images are verified, and successful cases are added to the respective training set. This study shows the use of deep learning to handle large amounts of imaging/chemical spectroscopy data efficiently and evaluate particle mixing state.

References:

- [1] Després et al., *Tellus B* **64** (2012).
- [2] China et al., *Environmental Science & Technology* **50** (22) (2016), p. 12179.
- [3] Wang et al., *Nature Geoscience* **9** (6) (2016), p. 433.

[4] Fröhlich-Nowoisky et al., Proceedings of the National Academy of Sciences **106 (31)** (2009), p. 12814.

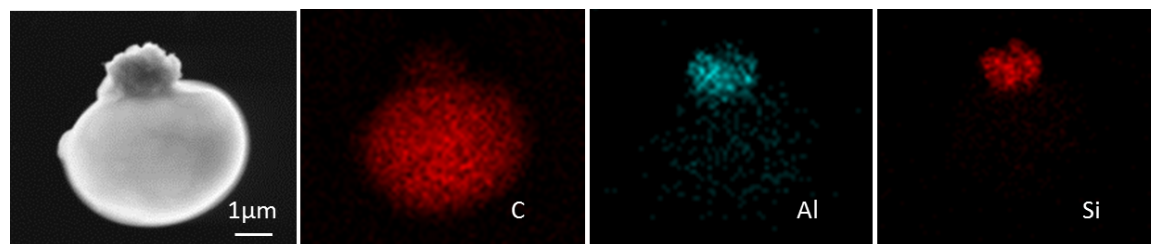


Figure 1. Biological particles mixed with dust particle. Elemental maps show presence of biological spore (high carbon content) and dust (Al and Si).

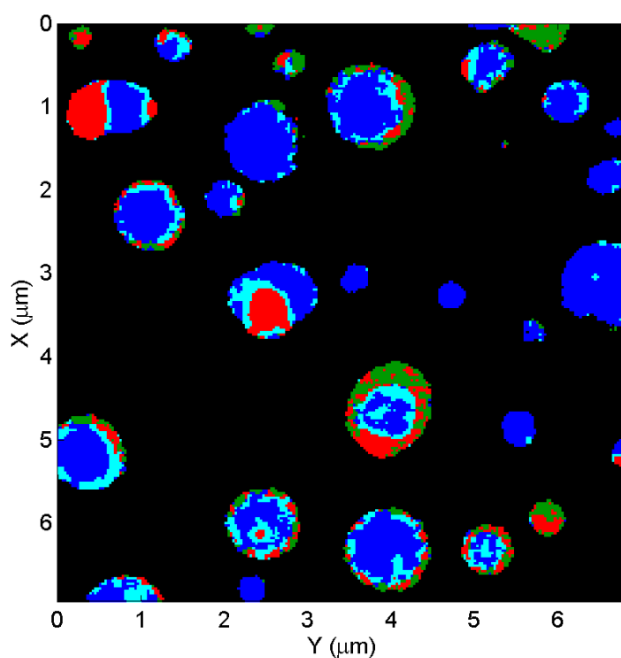


Figure 2. Chemical imaging and mixing state of individual atmospheric particles using STXM/NEXAFS. The different colors represent organic carbon (green), organics and inorganics (blue), organics and elemental carbon (red), and a mixture of organic, elemental carbon and inorganics (cyan).