Goldschmidt 2025 Abstract
https://doi.org/10.7185/gold2025.31021

Automating spectroscopic data fitting and interpretation for enhancing (bio)geochemical analysis in complex environmental samples

JAMES M. BYRNE

University of Bristol

The continuing and growing relevance of earth and environmental sciences is resulting in an explosion of experimental data. Analysis of this data seeks to address specific environmental questions related to (bio)geochemistry, especially at the nanoscale, such as how soil bacteria are connected to the release of greenhouse gases, or how changes in geochemistry influence the release of pollutants into drinking water. The accumulation of this information raises challenging questions about how to efficiently store, analyze, and share the data, especially in the era of open science. Spectroscopy data is often analyzed manually, with domain expertise used to interpret and eventually publish results. While this approach works for specialists or experienced researchers, it creates an invisible barrier for new adopters or more general users, such as those working in multidisciplinary contexts. To overcome this barrier and advocate for an open culture in data analysis, this project considers ways to implement machine learning to provide semi- and fully automated fitting routines for spectroscopic data, particularly those relevant to environmental sciences.

This presentation focuses on the development of a web application designed to read, fit, and interpret spectroscopy data. This cutting-edge project aims to deliver a framework that accelerates the analysis and interpretation of data through a combination of supervised and unsupervised machine learning methods. This will make data analysis tools available to researchers worldwide. We are currently working on three key aspects: (a) automated data interpretation tools to assist with analysis; (b) validation of the concept through the analysis of complex environmental samples as part of a case study; and (c) simplification of learning approaches used for spectral analysis to encourage more widespread uptake of different techniques. The project has initially focused on a few techniques, including Mössbauer spectroscopy and FTIR, but it is planned to expand into a range of analytical methods. The resulting platform will cater to various research fields, including (bio)geochemistry, geophysics, paleomagnetism, geomicrobiology, and astrobiology, among countless others. This presentation will discuss the development of this project, as well as ongoing challenges and new opportunities.