Statistically leveraging spectroscopy to link sediment properties and groundwater composition

ATHENA NGHIEM[1*], MASON STAHL[2], ALEXANDER VAN GEEN[1], BENJAMIN BOSTICK[1]

[1]Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York, 10964, USA, (*correspondence: anghiem@ldeo.columbia.edu)
[2]Union College, Schenectady, New York, 12308, USA

Naturally-occurring contamination of groundwater by arsenic threatens the health of millions of people who depend on groundwater for drinking water[1]. Arsenic release to groundwater is controlled by adsorption to iron minerals; microbially-mediated reductive dissolution of iron minerals to dissolved iron releases arsenic to groundwater[2]. Despite the critical role of iron minerals in this biogeochemical process, direct characterization of iron mineralogy, associated sedimentology, and weathering and redox transformations in aquifers is not well characterized in most environments. Characterization is particularly needed in redox transitions where sediment and aqueous composition vary spatially and temporally. Recent advancements in X-ray absorption spectroscopy (XAS) allow detailed descriptions of mineralogy with increased sample throughput. Here, we (1) characterize a relatively large data set of iron XAS spectra across a 2-km scale natural redox gradient in Vietnam with paired aqueous compositions. From hierarchical cluster analysis and principal component analysis on these spectra, we find distinct mineralogical redox transformations that can be linked to distinct solution compositions. Then, we (2) scale up this relationship by creating a supervised multinomial logistic regression model for classification of sediment by depositional history and post-depositional reduction. We run this model on a large (n=16294) compilation of aqueous arsenic and related aqueous redox-sensitive element concentrations compiled through an ongoing USGS Powell Center effort. This classification model is able to predict sedimentological profiles and maps based solely on aqueous concentrations well beyond the study area; it effectively identifies known Pleistocene terraces and Holocene valleys in the Bangladeshi aquifers susceptible to contamination. We also explore the application of this modeling effort using sediment composition and other available data. From statistically leveraging spectroscopy, these classifications distinguish depositional history, weathering, and redox transformations in sediments linked to groundwater quality.

[1] Fendorf *et al*. (2010), Science 328(5982), 1123–7. [2] Nickson *et al*. (2000) Appl. Geochemistry 15, 403–413.