Integrating Multivariate Statistical & Geospatial Analyses to Predict Processes Governing Geogenic Contaminant Release in Groundwater

ALANDRA M. LOPEZ[1], JEF CAERS[2], AND SCOTT FENDORF[1]

[1]Department of Earth System Science, Stanford University, Stanford, CA-94305, USA
[2]Department of Geological Sciences, Stanford University, Stanford, CA-94305, USA

With the increasing global need for groundwater resources to fulfill domestic and agricultural needs, we face the threat of enhancing concentrations of naturally occurring contaminants (e.g. As, Cr(VI), U) in water sources. Groundwater is already a critical resource for semi-arid agricultural regions, like the Central Valley, California, USA, and it is crucial that we fully understand how climate-driven changes in land use and land management practices alter groundwater quality.

This multicomponent statistical study applies multivariate and geospatial methods to assess underlying biogeochemical processes and the probability of groundwater arsenic (As), uranium (U), and hexavalent chromium [Cr(VI)] contamination across the Central Valley of California. Sequential factor and cluster analyses are used on a statewide geochemical database to identify groundwater chemistry patterns and their relationships with elevated As, U, or Cr(VI). Anthropogenic and natural influences on local As, U and Cr(VI) release are discerned through factor loading interpretations. For example, groundwater signatures characterized by nitrate, sulfate, Ca, and Cl are governed by irrigation recharge and distributed in shallower wells proximal to agricultural lands; those influenced by Be, F and nitrite are associated with urban centers. This method also suggests geogenic sources (e.g. granitic, sedimentary, mafic) of each contaminant and the prevailing redox and pH conditions within the aquifer. Factor scores are spatially interpolated using ordinary kriging and geostatistical simulation to generate maps of each governing factor and their spatial uncertainty, respectively. Finally, logistic regression is applied to exploit the relationship between groundwater clusters and predictor variables (land use, crop type, aquifer parameters, geomorphology, geology, surface hydrology) to build predictive models and improve our understanding about groundwater contamination throughout the Central Valley.