

## **Using big groundwater data to understand regional water chemistry**

S.L. BRANTLEY<sup>1\*</sup>, T. WEN<sup>1</sup>, X. NIU<sup>1</sup>, G. ZHENG<sup>2</sup>, M. GONZALES<sup>1</sup>, Z. LI<sup>2</sup>

<sup>1</sup>Earth and Environmental Systems Institute, Penn State Univ., University Park, PA 16802, USA

(\*correspondence: brantley@geosc.psu.edu)

<sup>2</sup>College of Information Sciences and Technology, Penn State Univ., University Park, PA 16802, USA

Geochemists generally do not have access to large volumes of groundwater chemistry data because such data are not collected, not available, or not publicized. The development of oil and gas in Pennsylvania has led to collection of large volumes of groundwater chemistry data that is now being published as part of Shale Network through the CUAHSI Hydrologic Information System. We are exploring what we can learn from these large volumes of data with respect to water quality issues around shale gas development, but also with respect to the distribution of geologic features, land use, and anthropogenic impacts on the subsurface. We have discovered that one of the best predictors of groundwater chemistry is bedrock composition. However, some analytes (e.g. Pb) do not show strong variations with bedrock composition. In some areas of shale gas development we see evidence that groundwater chemistry is improving but we also see evidence for localized areas of impact. Questions remain as to how to use big groundwater data to help communities make better decisions about their groundwater and land use choices. In addition, questions remain as to how to develop systems that can help communities safeguard their water as anthropogenic impacts increase.