Two "Flavors" of Big Data Studies in the Remediation Field: Methods and Example Applications

Poonam Kulkarni, Charles Newell, Travis McGuire, Tom McHugh, Jim McDade, and Dave Adamson (GSI Environmental Inc. Houston, TX, USA)

Background/Objectives. Private industry and governments have invested billions of dollars in environmental restoration, with thousands of sites in the United States requiring some type of environmental decision-making. Despite this investment, challenges to remediating contaminated sites still remain.

In the process of remediating these sites, however, large amounts of monitoring data are collected. By compiling, processing, and then visualizing these "big data sets" new insights can emerge about how groundwater sites behave and what remediation can and cannot do.

Approach/Activities. The following peer-reviewed papers describe two major variants of big data studies in the remediation field: *Geotracker Studies:* "Big data" from the Geotracker database from California was mined to develop concentration versus time data from groundwater monitoring wells at 12,000 petroleum sites (McHugh et al., 2013). A similar project mined the same database and found higher source zone attenuation rates with higher temperatures in the California sites (Kulkarni et al., 2017). Finally, a study by Adamson et al. (2015) evaluated dioxane plumes and compared them to their comingled chlorinated solvent plumes.

Mining Remediation Reports: McGuire et al. (2016) mined remediation reports from 235 in-situ DNAPL source zone treatments and measured before- and after-remediation source zone concentrations of chlorinated solvents. A similar study involving mining remediation reports studied the fate of nine "exceptionally long" MTBE plumes of the past (McDade et al., 2014).

Results/Lessons Learned. *Hydrocarbon source attenuation:* 1. From 2001 to 2011, groundwater monitoring data from 12,000 sites show a large decrease in groundwater concentrations of gasoline constituents. For benzene, half of the sites showed a decrease in concentration of 85% or more. For methyl tert-butyl ether (MTBE), this decrease was 96%. 2. Statistically significant and positive relationships between temperature and source attenuation rates were established for benzene and toluene, indicating that temperature does impact hydrocarbon degradation.

DNAPL source zone remediation: The middle 50% of the remediation projects achieved between 0.5 and 2 OoMs reduction in the geometric mean of the parent compound (between 71% and 99% reduction in chlorinated solvent concentrations). Statistically the four major technology types generally achieve similar results.

Dioxane Plumes: The magnitude and prevalence of dioxane attenuation documented here suggest that natural attenuation may be used to manage some but not necessarily all dioxane-impacted sites. The results suggest that dioxane has not migrated beyond chlorinated solvent plumes and not beyond the existing monitoring networks at the majority of sites.

MTBE Plumes: Groundwater monitoring data compiled in our review demonstrate that these MTBE plumes have decreased in length over the past decade, with five of the nine plumes exhibiting decreases of 75% or more compared to their historical maximum lengths. MTBE

concentrations within these plumes have decreased by 93% to 100%, with two of the nine sites showing significant decreases (98% and 99%) such that the regulatory authority has subsequently designated the site as requiring no further action.