MULTIMODEL BAYESIAN ANALYSIS OF DATA-WORTH: THEORY AND APPLICATIONS

Ming Ye, Florida State University, 850-644-4587, mingye@scs.fsu.edu

1. Ming Ye, Florida State University
2. Dan Lu, Florida State University
3. Liang Xue, University of Arizona
4. Shlomo P. Neuman, University of Arizona

The management of contaminated groundwater systems requires an understanding of their response to alternative remediation strategies. Such understanding requires the collection of suitable data to help characterize the system and monitor its response to existing and future cleanup and/or containment options. It also requires incorporating such data in suitable models of water flow and contaminant transport. As the acquisition of subsurface characterization and monitoring data is costly, it is imperative that the design of corresponding data collection schemes be cost-effective, i.e., that the expected benefit of new information exceed its cost. A major benefit of new data is its potential to help improve one's understanding of the system, in large part through a reduction in model predictive uncertainty. Traditionally, value-of-information or data-worth analyses have relied on a single conceptual-mathematical model of site hydrology. Yet there is a growing recognition that analyses and predictions based on a single hydrologic concept are prone to statistical bias and underestimation of uncertainty. This has led to a recent emphasis on conducting hydrologic analyses and rendering corresponding predictions by means of multiple models. Here we adopt a multimodel approach to optimum value-of-information or data-worth analyses based on model averaging within a Bayesian framework (Neuman et al., 2011), and develop three approximations to reduce computational cost of implementing the approach using the Monte Carlo methods. The approach and its approximations are applied to pneumatic permeability data from an unsaturated fractured tuff site near Superior, Arizona, USA. The application includes a cross-validation study for proof of concept and evaluation of the accuracy of the approximations. It also includes an optimization method that uses the differential evolution techniques to locate optimum sample locations for maximum uncertainty reduction. Applications of the Bayesian method of data-worth analysis to groundwater reactive transport modeling have also been explored.