Optimal Site Investigation: What data to collect so that calibrated models have best prognistic power?

WOLFGANG NOWAK[1], FELIPE P.J. DE BARROS[2], ANDREAS GEIGES[1], PHILIPP LEUBE[1]

[1] Uni Stuttgart, FB Bau und Umwelt, Pfaffenwaldring 7a, 70569 Stuttgart., vor.nach@iws.uni-stuttgart.de
[2] Uni Stuttgart, FB Mathematik, Pfaffenwaldring 7a, 70569 Stuttgart., felipe.debarros@simtech.uni-stuttgart.de

Subsurface variability and the scarcity of site-specific data suggest to use stochastic approaches, such as Monte-Carlo simulation with random conductivity fields, for predicting groundwater flow and contaminant transport. This acknowledges the uncertainty and limited prognostic power of modeling under incomplete knowledge. After collecting site-specific data (e.g., from well or tracer tests), these data can be used to calibrate the stochastic model, for example with the Pilot-Point method.

The remaining practical question is what data to collect in order to minimize the uncertainty and hence maximize the prognostic power of a model: what types of measurements (conductivity, heads, concentrations,...) at what scale (core sampling versus large-scale well tests,...) at which locations should be collected? Most of all: how are the financial resources available for data collection invested best? The combination of locations, numbers and data types are called a sampling design. So-called optimal experimental design techniques can be used to find the best design for a given site. While experimental design is a well-known technique in many other disciplines (PUKELSHEIM, 2006) and in geostatistics (MÜLLER, 2007), the needs of hydrogeological modeling cast a new light onto it (NOWAK, 2009).

Traditionally, designs are optimized to yield the best confidence in model parameters. We argue that this is not suitable for most hydrogeological applications. In practice, hydrogeologists will only seldom seek for an all-purpose model, but set up and calibrate their models towards a certain task-specific prediction goal. A model that is suitable to predict groundwater quantity may be insufficient to predict groundwater quality, and will have to be upgraded accordingly. The same holds for designs: they should not be all-purpose designs, but rather be optimized for the actual modeling and prediction task at hand.

Therefore, the objective in optimal design should be a most close representation of the data utility towards the actual modeling purpose. For example, if the model purpose is to decide on the safety of a drinking water well, site investigation should be optimized to allow a maximum confidence level of a hypothesis test for well contamination or for human health risk (NOWAK, 2008). In catchment delineation problems, the optimization should look at the confidence in drawing the catchment outline, or even at the confidence in specifying a certain critical location to be inside or outside.

The current study will provide an overview over recent research activities and paradigms in optimal design of site investigation. Different modeling objectives will be addressed, and it will be demonstrated how different sources of uncertainty pose different information needs, resulting in fundamentally different site investigation strategies and sampling patterns (NOWAK et al., 2009). The different approaches will be explained on synthetic scenarios for the sake of illustration.

Literature

PUKELSHEIM, F. (2006): Optimal Design of Experiments. – SIAM, Philadelphia.

MÜLLER, W.G. (2007): Collecting Spatial Data. Optimum Design of Experiments For Random Fields, 3rd Ed. – Springer, Berlin.

NOWAK, W. (2009): Measures of Parameter Uncertainty in Geostatistical Estimation and Geostatistical Optimal Design. – Math. Geosc., doi:10.1007/s11004-009-9245-1 (in press)

NOWAK, W. (2008): A Hypothesis-Driven Approach to Optimal Site Investigation. – In: Eos Trans. AGU, 89(53): H43A-0984.

NOWAK, W. & DE BARROS, F.P.J. & RUBIN, Y. (2009): Bayesian Geostatistical Design: Task-Driven Optimal Site Investigation when the Geostatistical Model is Uncertain. – Water Resour. Res., doi:10.1029/2009WR008312 (in press).