Geophysical Research Abstracts
Vol. 15, EGU2013-8571, 2013
EGU General Assembly 2013

Never trust straightforward intuition when choosing the number of Monte Carlo simulations

Philipp Leube (1), Wolfgang Nowak (1), Felipe de Barros (2), and Ram Rajagopal (3)

(1) Institute for Modelling Hydraulic and Environmental Systems, University of Stuttgart, Stuttgart, Germany (philipp.leube@iws.uni-stuttgart.de), (2) Sonny Astani Dept. of Civil and Environmental Engineering, University of Southern California, Los Angeles, USA, (3) Stanford Sustainable Systems Lab, Dept. of Civil and Environmental Engineering, Stanford University, Stanford, USA

Uncertainty quantification for predicting flow- and transport in heterogeneous aquifers often entails Monte Carlo simulations executed on top of random field generation. Typically, the number of Monte Carlo simulations ranges between 500-1000, sometimes even higher. In many cases, this choice is based on restricted available computational time, or on convergence analysis of the Monte Carlo simulations. The spatial resolution is most frequently fixed to experience values from the literature, independent of the number of Monte Carlo realizations. Sometimes, a compromise is found between spatial resolution, Monte Carlo resolution and available computational time. We question this practice, because it does not look at the trade-off between the individual resolution, individual errors, total errors and computational time.

Our goal is to show that what models really want is neither poor statistics of good physics, nor good statistics of pour physics. Instead, one should look for an overall optimum choice in both decisions. To this end, we assess an optimum for the number of Monte Carlo simulations together with the spatial resolution of computational models. Our analysis is based on the idea to jointly consider the discretization errors and computational costs of all individual model dimensions (physical space, time, parameter space). This yields a cost-to-error surface which serves to aid modelers in finding an optimal allocation of the computational resources. The optimal allocation yields highest accuracy associated with a given prediction goal for a given computational budget.

We illustrate our approach with two examples from subsurface hydrogeology. The examples are taken from wetland management and from a remediation design problem. When comparing the two different optimum allocation patterns among each other and to typical values found in the literature, we make counterintuitive observations. For example, a realistic number of Monte Carlo realizations should be aligned with spatial resolutions of up to 16 or 32 grid cells per correlation length. Also, the number of Monte Carlo necessary to achieve a minimal error is much smaller compared to those found in the literature. Allocating computational budgets wisely leads, for our specific examples, to a speedup in the order of 3-10. Our results indicate that the optimum strongly depends on the underlying physical problem and prediction task. We conclude that this type of optimum should be performed for every individual task, and in a goal oriented manner.