Geophysical Research Abstracts,
Vol. 11, EGU2009-12187, 2009
EGU General Assembly 2009

The direct sampling: a new way of performing multiple-points simulations

P. Renard, G. Mariethoz, and J. Straubhaar
University of Neuchatel, Centre d'Hydrogeologie, Neuchatel, Switzerland (philippe.renard@unine.ch)

Stochastic hydrogeology is based on the premise that heterogeneity is a key factor controlling groundwater flow and transport processes and that because of the lack of data it is necessary to model it and its impact in order to provide not only reliable forecasts but also reliable error bounds. During the last 50 years massive progresses have been made in this field but mostly in the context of relatively simple random field models (binary, multi-Gaussian). We believe however that there are a significant number of applications in which the underlying assumptions made when applying those models are not valid.

This is why we defend the idea that a richer model should be used to describe the heterogeneity and integrate as much as possible geological, hydrological, and geophysical observations. Toward this end we will present the last developments of the direct sampling multiple point algorithm.

Indeed, multiple-points statistics has become a prevalent (and sometimes controversial) subject in geostatistics. Despite the existence of certain shortcomings, it is by now the most flexible method to integrate a conceptual geological model in a stochastic framework. The concept behind the method is to use a training image to represent the desired spatial structure of the field. The training image is scanned and all pixels configurations of a certain size (the template size) are stored in a catalogue of data events, most often in a tree structure. This structure is then used to compute the conditional probabilities at each simulated node. Because of limited memory usage, this approach can only deal with categorical variables. Note that the memory load is directly proportional to the size of the template and the number of facies. It can be prohibitive for large 3D grids. In general, the template is not large enough to capture large-scale structures such as channels. Multi-grids have been introduced to palliate this problem by simulating the large-scale structures first, and later the small-scale features.

We argue in this work that there is no need to store and count the configurations found in the training image. Instead, it is more convenient to sample randomly the training image for a given data event. For each of the successive samples, the mismatch between the data event observed in the simulation and the one sampled from the training image is calculated. If there is no mismatch or if it is lower than a given threshold, the sampling process is stopped and the value at the central node of the data event in the training image is directly used for the simulation.

This method presents several advantages. First it reduces considerably memory usage. Second, the data event can have any geometry (no need for data template) and can change for each simulated node. The Multiple-grids are replaced by a continuous variation of the data events size during the simulation process. Third, the approach allows dealing with continuous variables and can easily handle secondary variables either to account for real secondary variables or non stationarity. Finally, the comparison of execution times show that this approach can be faster than the traditional one depending on the size of the training image and on the parameters that are chosen to control the simulation. Last but not least, the approach is easy to parallelize.