Hydrol. Earth Syst. Sci. Discuss., 10, C6774–C6777, 2013
www.hydrol-earth-syst-sci-discuss.net/10/C6774/2013/

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Challenges in conditioning a stochastic geological model of a heterogeneous glacial aquifer to a comprehensive soft dataset" *by* J. Koch et al.

Anonymous Referee #1

Received and published: 17 December 2013

Review of "Challenges in conditioning a stochastic geological model of a heterogeneous glacial aquifer to a comprehensive soft dataset"

This manuscript essentially describes a practical approach to integrate auxiliary information in the transition probability approach (Tprogs). Although the approach is quite relevant for this particular application, the existing literature on this problem is largely missed. Integration of exhaustive or dense data in geostatistics has been dealt with from the early days of geostatistics (cosimulaltion, cokriging, kriging/simulation with external drift) until the most recent developments (integration of non-stationarity in multiple-point geostatistics). This study entirely misses this. In the limited context

of Tprogs however, it may be true that this problem has not been addressed. The approach proposed consists of using a decimated version of the auxiliary variable as conditioning data for the primary variable. It is quite a crude way of proceeding since it ignores the differences of support in the different variables, as well as the differences in spatial variability. I think this is entirely missed in the discussions and is not tested throughout. The title and abstract have a much too general focus. The paper is essentially only focusing on the transition probability method, and more particularly on the tprogs software. In my opinion, this manuscript would be more appropriate in a journal focusing specifically on the application of geostatistical methods, such as Mathematical Geosciences. It may also be fit for Hydrogeology Journal because of the emphasis on the case study. In any case I would recommend major revisions before resubmission.

Specific comments: p.15222, top: To be a bit more exhaustive, references to truncated Gaussian and Truncated pluriGaussian methods could be added. Object-based models could also be mentioned.

p.15222, l.8: What is meant by compatible?

p.15222, l.24-25: I don't agree with this statement: there is a large number of papers that include soft data with geostatistical methods (although maybe not specifically with tprogs). Please have a look at the works on collocated simulation, probability aggregation and tau models. It has been heavily used with indicator geostatistics, multiple-point geostatistics and object-based methods. Soft data is sometimes called "soft probability". A few examples of references: In the context of variogram-based methods: DEUTSCH, C. V. & WEN, X. H. 2000. Integrating large-scale soft data by simulated annealing and probability constraints. Mathematical Geology, 32, 49-67. MARIETHOZ, G., RENARD, P. & FROIDEVAUX, R. 2009. Integrating collocated auxiliary parameters in geostatistical simulations using joint probability distributions and probability aggregation. Water Resources Research, 45. In the context of multiple-point simulations: CAERS, J. 2003. History matching under training-image-based geological model constraints. SPE Journal, 8, 218-226. CHUGUNOVA, T. & HU, L. 2008. Multiple-Point

Simulations Constrained by Continuous Auxiliary Data. Mathematical Geosciences, 40, 133-146. These are only a few examples, you will find many other references regarding other geostatistical methods. More generally, the references in the introduction used are often outdated.

p. 15223,l.12-14: There are numerous studies where exhaustive geophysics is used and validated. Below is a very recent one, however it is a problem that has been heavily studied in the last decade. EMERY, X. & PARRA, J. 2013. Integration of crosswell seismic data for simulating porosity in a heterogeneous carbonate aquifer. Journal of Applied Geophysics, 98, 254-264.

p. 15223,l.15-16: This statement shows again that an entire side of the literature is missed on geophysical inversion with geostatistics. See for example: HANSEN, T. M., CORDUA, K. S., LOOMS, M. C. & MOSEGAARD, K. 2013. SIPPI: A Matlab toolbox for sampling the solution to inverse problems with complex prior information: Part 2-Application to crosshole GPR tomography. Computers and Geosciences, 52, 481-492.

p. 15223,l.19-22: These are very vague statements. For geophysical data integration, the usual criterion is a forward problem that calculates the geophysical response given a certain geological model. Valid models are those that reproduce the measures data when such a forward model is applied.

p. 15228,l.13-15: Such an exhaustive conditioning is never applied. Instead, a general approach is the one of probability aggregation where the probability distribution coming from the spatial model (tprogs or any other method) is combined with the prior probability coming from geophysics. See: ALLARD, D., COMUNIAN, A. & RENARD, P. 2012. Probability Aggregation Methods in Geoscience. Mathematical Geosciences, 44, 545-581. For an example of synthetic application, see LIU, Y. 2006. Using the Snesim program for multiple-point statistical simulation. Computers & Geosciences, 23, 1544-1563.

Section 4.3: The question of support size is not addressed here, however it is critical

to know what is the support size of the geophysical data. Is a data point representative of 1m2 or 1km2? This should be a major factor when deciding whether to use subsampling or moving averaging approaches. The description of the moving sampling is not clear enough. Is the same sampling used for each realization? Are the data resampled or interpolated? Why "moving" sampling - is there a moving window defined?

p. 15229, l.5: such an "optimal" combined knowledge should be formulated in a Bayesian framework, which is completely not done here. Therefore the word optimal is not correct in this context.

Section 5.3: Figure 5 is good and justifies the approach. However what is proposed is a "fix" that has no generality. I have nothing against such fixes that work in practice, but they should be acknowledged as such and their limitations should be clearly stated.

p. 15234, l.26: The terminology of more/less deterministic is not correct. A model is deterministic or it is not. I would rather speak of higher/lower variability. This is throughout the manuscript.

Section 5.4: In my opinion the approach of comparing the realizations with the probability maps derived from geophysics is flawed. Geophysics does not provide a facies distribution, but a smoothed and coarsened version of it. It is a different variable, indirectly related to the facies. Therefore they are not expected to have the same spatial features and cannot be compared in terms of connectivity or correlation scales. It is clear from figure 8 that the borehole data and the SkyTEM data present different spatial properties.

Section 6: This section essentially repeats material that was discussed earlier, and therefore can be shortened or removed.

p. 15241, l.3-5: I'd recommend reading some textbooks on kriging with external drift (in particular the book of Chiles and Delfiner).

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 15219, 2013.