Hydrol. Earth Syst. Sci. Discuss., 7, C2184–C2191, 2010
www.hydrol-earth-syst-sci-discuss.net/7/C2184/2010/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Bayesian approach for three-dimensional aquifer characterization at the hanford 300 area" *by* H. Murakami et al.

H. Murakami et al.

harukom@nuc.berkeley.edu

Received and published: 9 September 2010

We appreciate the constructive comments. The responses to the comments are as follows. Three figures (Fig. C3-1 to Fig. C3-3) were added at the end of this document.

First, we would like to mention several issues common to many of the comments:
(1) Although the MAD framework itself is general, we could customize each component depending on the specific problem. For example, we may choose parametric or nonparametric likelihood estimation, depending on the data. The modular structure of MAD makes such customization possible.
(2) The inverse modeling always comes after analyzing the forward simulations, using the same numerical code for simulating a physical process (e.g., injection tests in our

case). In the forward simulations, we generate random fields using a given parameter set, and simulate the physical process in each field. By analyzing the ensemble of the simulated observations (measurements), we may analyze the distribution of, or correlation between, the possible observations.
(3) In order to reduce the computational burden, some approximations may be necessary. We may make such assumptions based on analyzing the forward simulations.

The responses to each of the individual comments are as follows.

Response to Comment Number 1:

For the likelihood estimation, we used a parametric approach (multivariate Gaussian distribution). We have confirmed that 250 realizations per parameter set are sufficient. Please see Response to Comment Number 4(4) and Fig. C2-1 in Responses to Comments by Anonymous Referee #2.

Response to Comment Number 2:

The distribution of the parameters does not have to be zero at the boundaries, since available data is sometimes insufficient for narrowing down the distribution. You may find such cases in Diggle and Ribeiro (2002) and Hou and Rubin (2005). We set reasonable bounds for the prior distribution based on the information from geologically similar sites.

To support this claim, we conducted a numerical experiment. We generated a set of synthetic log$K$ measurements at the 283 EBF data locations and inferred the 3-D structural parameters. As a set of synthetic measurements, we generated ln$K$

at 283 locations from a multivariate Gaussian distribution with $\{\beta, \eta^2, \lambda_h, \lambda_v, \nu\} = \{-2.0, 2.0, 30, 3.0, 0.020\}$. Note that this process does not involve MAD, but only uses the model-based Bayesian geostatistics summarized in Appendix B in the paper (Chapter 6 in Diggle and Ribeiro, 2006). In the synthetic study, we have one set of ln$K$ values, in contrast to the real case where we have the distributions of ln$K$ due to the uncertainty in ln$T$ from the injection test inversion.

Figure C3-1 shows the inferred marginal distributions of the parameters. For the horizontal scale, vertical scale, and nugget, the upper and lower boundaries of the x-axis correspond to the bounds of the prior distributions. The distribution of the horizontal scale is not zero at the boundary. We can conclude that 283 data points are not enough to narrow down the distribution in a 3-D space. Especially, spacing between the wells restricts variation in horizontal lags, whereas we have many different lags for the vertical scale.

Response to Comment Number 3:

In order to reduce the computational burden, it is more advantageous to use a parametric approach whenever possible. While analyzing the simulation results in forward modeling, we found that the distribution of the possible observations ($\tilde{\boldsymbol{z}}$) is approximately multivariate Gaussian. For example, we generated 500 fields with a geostatistical parameter set $\{\mu, \sigma^2, \phi\} = \{2.21, 2.99, 27.2\}$, and simulated the zeroth moments in the injection tests. We chose this parameter set, since the variance is high among the ones generated from the prior. Figure C3-2 shows the histograms and pair plot of the simulated observations at two observations wells in the same injection test. We observe that the Gaussian assumption is quite reasonable. We would also like to note that even when using a nonparametric approach for the density estimation, some approximations and assumptions are unavoidable, such as choosing kernels and bandwidth.

Response to Comment Number 4:

While Dr. Zhang is focused on a general concept, we are focused on a case-specific approach. In the same way as the previous comments, in the forward simulations, we simulated the possible observations and checked the density of the ensemble. We found that the correlation among the observations in the different tests is quite small.

For example, based on the same ensemble as the previous comment, Figure C3-3 shows the histograms and pair plot of the simulated observations at two observations wells in the different injection tests. In the scatter plot, the contour lines of the estimated distribution are close to a circle, which suggests that the correlation between the two wells is insignificant.

Response to Comment Number a:

We agree that our goal is to characterize the whole field. However, we consider that the transect is the most effective way to check predictive ability, since we can see whether the true field is contained in the confidence interval and how wide the confidence interval is. We picked one transect along the centerline of the field, which also corresponds to the line passing through large variability, from high near the top to low in the middle of the well plot.

Response to Comment Number b:

Please see Response to Specific Comment Number 4 by Anonymous Referee #1.

Complete Captions for Figures

* Due to the limited space, the caption below each figure was truncated.

Figure C3-1. Marginal posterior distributions of 3-D geostatistical structural parameters of logK values inferred from the synthetic measurements. The red lines are the true parameter values.

Figure C3-2. Histograms and pair plot of the simulated observations at two observations wells (Well 2-22 and Well 3-29) in the same injection test (injection at Well 2-18). The red lines represent the estimated multivariate Gaussian distribution (contour lines in the pair plot).

Figure C3-3. Histograms and pair plot of the simulated observations at two observations wells (Well 2-22 and Well 3-29) in the different injection test (injection at Well 2-18 and Well 2-24). The red lines represent the estimated multivariate Gaussian distribution (contour lines in the pair plot).

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 7, 2017, 2010.

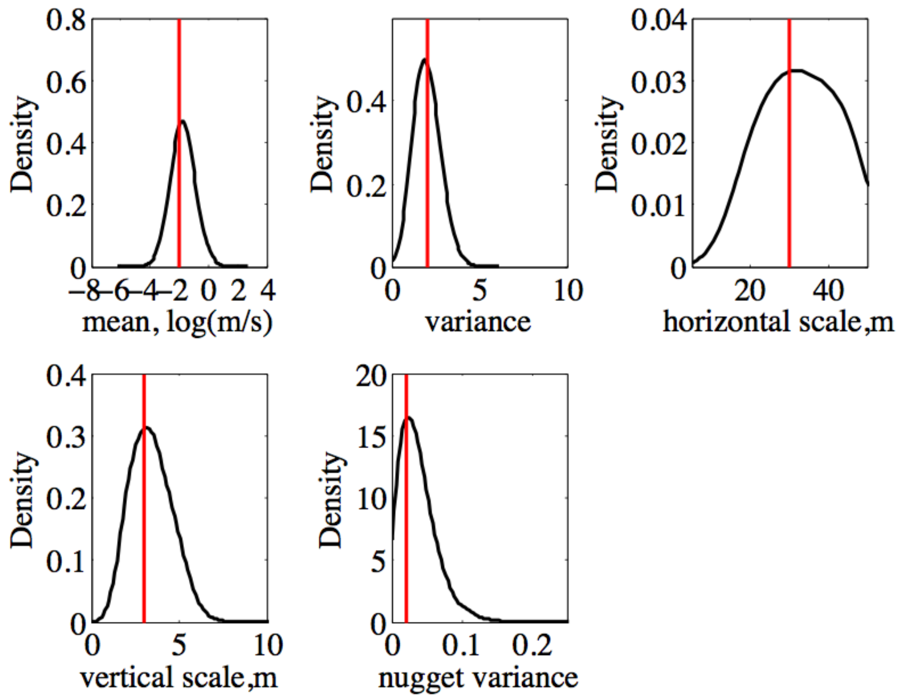

Fig. 1. (Figure C3-1). Marginal posterior distributions of 3-D geostatistical structural parameters of logK values inferred from the synthetic measurements. The red lines are the true parameter values.

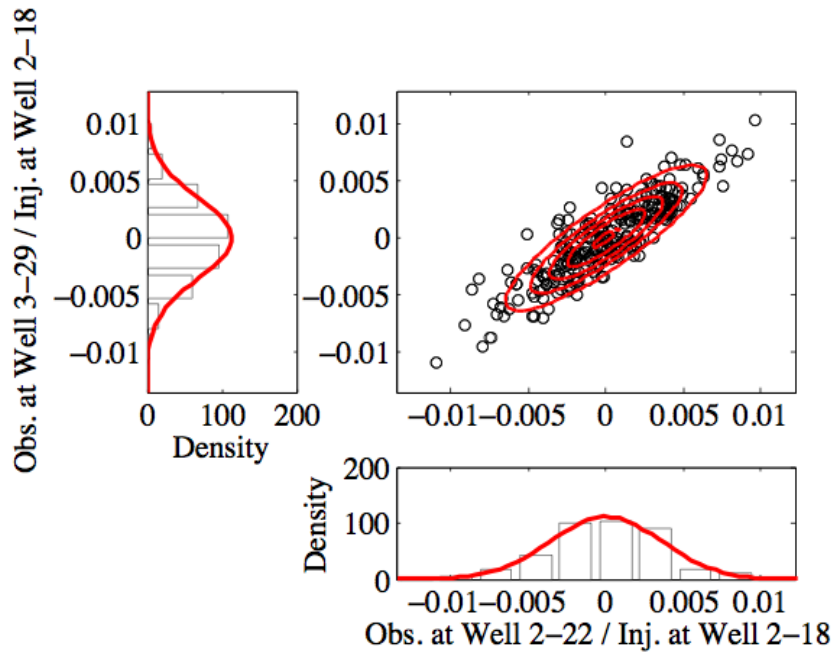

Fig. 2. (Figure C3-2). Histograms and pair plot of the simulated observations at two observations wells (Well 2-22 and Well 3-29) in the same injection test (injection at Well 2-18). The red lines represent t

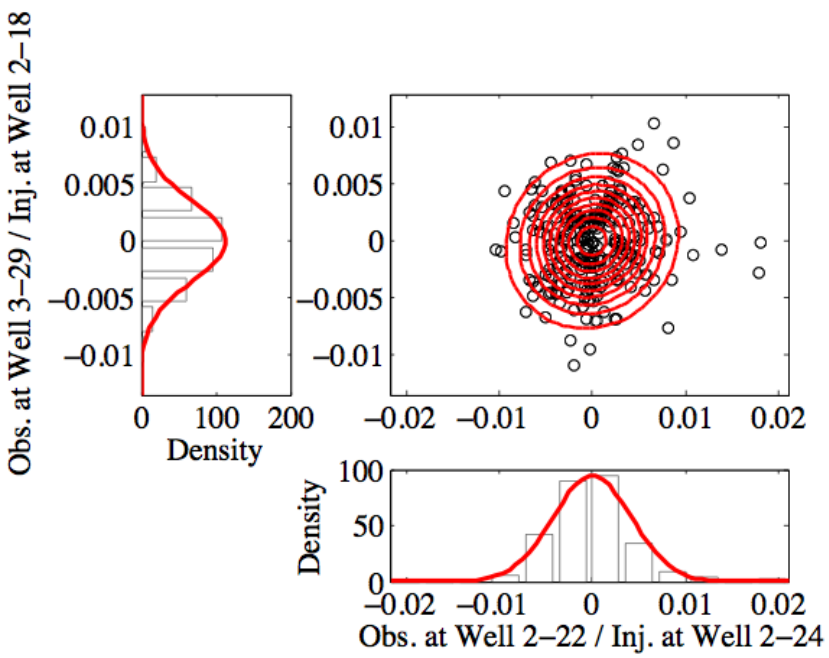

Fig. 3. (Figure C3-3). Histograms and pair plot of the simulated observations at two observations wells (Well 2-22 and Well 3-29) in the different injection test (injection at Well 2-18 and Well 2-24). The re