

## ***Interactive comment on “Sensitivity and identifiability of hydraulic and geophysical parameters from streaming potential signals in unsaturated porous media” by Anis Younes et al.***

**Anis Younes et al.**

fahs@unistra.fr

Received and published: 10 April 2018

We thank the reviewer for his/her thoughtful and detailed comments that definitely helped us clarify the manuscript and avoid misinterpretations.

The paper addresses the important topic of assessing the influence of hydraulic and geophysical parameters on streaming potential in the unsaturated zone. I think that the topic would be interesting to the journals audience. Furthermore the paper is generally well written and includes novel results. We thank the reviewer for his/her positive appraisal of our work.

C1

Comments:

(1) What is the notion behind replacing the numerical model with data-driven PCE in global sensitivity analysis? I mean PCE is often used when the computational cost is prohibitive, and off course, at the cost of reduced accuracy. However, the computational cost does not seem to be high in the current study. So justification is needed on this issue. The Sobol' indices are introduced in the framework of ANalysis Of VAriance (ANOVA) which aims at decomposing the variance of the output as a sum of contributions of each input variable, or combinations between them. It can be stated that the PCE method is a surrogate-based approach. However, we argue that this method employs ANOVA-like decomposition and hence can be considered as a spectral method (such as the Fourier amplitude sensitivity test, Cukier et al., 1973; Saltelli et al., 1999). Indeed, the Sobol' indices are directly obtained from the PCE coefficients without needing to run the surrogate model. This point is specified in the revised version (2) Why does the study use two different methods for parameter estimation? Is the objective (a) comparison of the two methods, or (b) double checking the results (or possibly both)? I suggest that the authors clearly define the objective. Moreover, since the two methods are conceptually very different (e.g. sampling vs. optimization), if the objective includes (a), then attention must be given to other issues such as the very different and superior information content of the results in MCMC, and its ability to effectively employ prior information. According to this comment, we specify in the revised version that (i) FOA and MCMC are conceptually very different and contrarily to FOA, the MCMC method is robust since no assumptions of model linearity or differentiability are required. Furthermore, MCMC can include prior information available for the parameters and yields not only the optimal point estimate of the parameters but also a quantification of the entire parameter space. (ii) The comparison between FOA and MCMC methods is performed to check whether the popular FOA approach can provide reliable estimation of parameters and associated uncertainties for the investigated highly nonlinear hydro-geophysical problem both in the case of abundant data (small uncertainty regions) and in the case of scarcity of data (large uncertainty regions).

C2

(3) MCMC is not a method but a general class of strategies that may not necessarily be based on the steps described in page 17. Hence, I suggest that before describing the steps, the authors mention the specific algorithm used in the study (i.e. DREAM) and then focus their discussions on how DREAM works. Please also add reference for the MCMC steps. We agree and according to this comment (i) we specify before describing the steps that several MCMC strategies have been developed for Bayesian sampling of the parameter space (Gallagher and Doherty, 2007; Vrugt, 2016). In groundwater and vadose zone modeling context, the most widely used of these strategies is the Metropolis Hastings algorithm (Metropolis et al., 1953; Hastings, 1970) described in the paper. (ii) We add the reference Gelman et al. (1996) for the Metropolis Hastings algorithm steps (iii) More details are provided concerning DREAM and DREAM(ZS). Indeed, in the revised version, we specify that Vrugt et al. (2009b, 2009c) developed the DREAM MCMC sampler based on the differential evolution–Markov Chain method of ter Braak (2006) to improve sampling efficiency. DREAM runs multiple Markov chains in parallel and uses subspace sampling and outlier chain correction to speed up MCMC convergence (Vrugt, 2016). Laloy and Vrugt (2012) developed the DREAM(ZS) MCMC sampler in which a candidate for each chain is drawn from an archive of past states denoted Z which plays the role of the generator q. The interested readers are referred to Vrugt (2016) for more details about properties and implementation of DREAM and DREAM(ZS). In the current study, the DREAM(ZS) software is used for the MCMC estimation of the hydrogeophysical parameters. These points are specified in the revised version (4) Some other choices in the paper also require further explanations and justification. These include: (a) choice of the uniform priors (line 193), (b) choice of parameters of the Gaussian distribution for  $\theta$  (lines 366-367), and (c) the logic behind the definition of scenarios in section 4.2. We agree and specify in the revised version that (i) The non-informative uniform distributions are used here to express the absence of prior information which makes all possible values of the parameter equally likely. (ii) The saturated water content is assumed accurately measured by weighing the saturated soil to be  $\theta_{\text{sat}}=0.43\text{cm}^3.\text{cm}^{-3}$ . The error measurements are assumed to be

C3

independently and normally distributed with a zero mean and a standard deviation  $0.01\text{cm}^3.\text{cm}^{-3}$ . (iii) The two scenarios are considered to check whether the FOA approach can provide reliable estimation of parameters and associated uncertainties for the investigated highly nonlinear hydrogeophysical problem both in the case of abundant data (small uncertainty regions) and in the case of scarcity of data (large uncertainty regions).

Minor Comments: All the comments have been taken into account in the revised version

(1) The numbering of some sections requires correction (e.g. after section 4.2 we have section 3.1) We agree and correct the numbering in the revised version (2) Line 83: Move “parameters” before the parenthesis. Done in the revised version (3) Section 2.2: I also suggest adding a schematic figure and possibly a photo of the laboratory setup. A schematic figure with the experimental device has been added.

(4) Line 166: Data are generated “from” or “for”? laboratory setup. The sentence has been changed.

(5) Line 186: Delete additional “a”. Done in the revised version

(6) Line 193: Why use the word “prior” and not simply “distribution”, since the authors are discussing GSA in section 3.1, and not a Bayesian inference method. We agree and change the word “prior” by “distribution” in the revised version

(7) Line 230: Correct typing error for “blue”. Done in the revised version

(8) Line 289: Correct typo to “shows that:” Done in the revised version (9) Line 357: Add reference for the DREAM software. Done in the revised version (10) Line 378: Not all sub-plots of figure 3 are symmetric though (e.g. plot for  $\alpha$ ) and so some cannot be considered bell-shaped. We agree and change the sentence to nearly bell-shaped

(11) Fig. 2: The colors are not so distinct (e.g. blue and blue-green) and so the difference between the plots may be hard to conceive for some readers. Done in the

C4

revised version

(12) Fig 3 and 4: Purely as a suggestion, the histograms can be replaced with PDFs, as I think this would provide better visualization. We obtain a better visualization with histograms that we keep in the Figures 3 and 4.

Please find attached the revised version of the paper.

Please also note the supplement to this comment:

<https://www.hydrol-earth-syst-sci-discuss.net/hess-2017-730/hess-2017-730-AC2-supplement.pdf>

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2017-730>, 2018.