Hydrol. Earth Syst. Sci. Discuss., 8, C2891–C2893, 2011
www.hydrol-earth-syst-sci-discuss.net/8/C2891/2011/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Identification of hydrological model parameters for flood forecasting using data depth measures" *by* T. Krauße and J. Cullmann

T. Krauße and J. Cullmann

thomas.krausse@tu-dresden.de

Received and published: 19 July 2011

Dear Referee,

The paper introduces a new algorithm that aims to be more efficient and yield more robust parameter sets than existing algorithms for robust parameter estimation (ROPE) based on Monte Carlo sampling. Case studies with a complex hydrological model are used to demonstrate the benefits of the new algorithm.

>My two main comments relate to how results from the case studies demonstrate improved efficiency and >robustness of the new algorithm.

>1. Efficiency >As discussed on page 2377, the main rationale for developing the new algorithm (replacing Monte Carlo sampling >with Particle Swarm Optimization PSO) is the need for computational efficiency. Hence, the focus here should be on >a comparison of computational efficiency between the various versions of the ROPE algorithm. Such a comparison >is done for the first synthetic case study, showing improved efficiency of the new algorithm. However, it would be >more interesting and convincing to do this for the two real-world case studies. The synthetic case study is actually >not that interesting and should perhaps be omitted.

According to your and the editor's comment we omitted the synthetic case study. The computational efficiency of a calibration algorithm is highly depend on the number of model runs required in order to obtain a stable solution. We inserted a brief discussion on this issue (refer to page 14 and 17) to both case studies. The PSO based algorithm requires a tremendously smaller number of model runs while providing better results.

>2. Robustness >Using performance in validation as the main robustness criterion, the two real-world case studies show improved >robustness with the new algorithm compared to the existing ROPE algorithm (figs. 6 and 14). However, the last case >study shows very similar results between the new algorithm and PSO without deep parameter generation. This >seems to suggest that deep generation, designed to increase robustness, is not that important in this case. This >should be more extensively discussed; when is the new ROPE algorithm developed here expected to improve >robustness above a method that does not perform deep parameter generation?

The deep parameter sampling cannot generate parameter vectors with higher model performance. However it can remove parameters with bad model performance from the previously (by PSO-GA_u) estimated set. The ROPE-PSO results therefore cannot outperform the PSO-GA_u results in terms of the best achieved model performance. However it can remove outliers and improve the mean performance. Refer to Figure 17 in the new manuscript. This issue is also discussed while presenting these results (page 21). Consider that the results slightly changed because we redefined the Flood-

Skill criterion according to your advice (see below)

>3. Other comments: >-Definition of the floodskill score is counter-intuitive, as one expects “skill” something that is to be maximized, yet >here it is minimized.

We agree to your comment. Therefore redefined the FloodSkill score and repeated the calibration with the new defined criterion.

-The word “representative” in the title is quite vague; what is meant by a representative parameterization?

In this context the word representative is related to the term robust and transferable. Thus, a representative parameterisation leads to a model that corresponds to a sufficient performance on all or most validation time periods the model was calibrated for. Within the scope of this paper that means that a representative model shows a sufficient model performance on the various flood events in the validation set.

-Good parameter sets are defined by a threshold parameter tolf – how was its value determined? And how does it compare to the 10% best parameters criterion in the other ROPE algorithms? To what extent do these settings directly affect the spread in the derived parameter populations (comparing parameter histograms in figs 9 and 10)?

The uncertainty tolerance is estimated according to a method given in the paper of Bardossy and Singh (2008). We briefly introduced this method and also inserted a discussion on this issue. The 10% criterion in the AROPE is just a criterion for every Monte Carlo iteration. However in the stopping criterion the tolerance is also considered (Figure 5 in the new manuscript).

Kind regards,

Johannes Cullmann and Thomas Krauße

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 2423, 2011.