Reply to comments by Chris Onof

February 29, 2012

Reviewer comment: A question of clarification: on what data was the Kruskal-Wallis test performed? Is it on simulated and observed data, or on some set of statistics thereof?

Reply: The Kruskal-Wallis test is performed on the obtained objective function values, for the different objective functions and repetitions, by the different optimization methods. The Kruskal-Wallis test is thus used to check whether these repetitions have the same median. This way, an objective comparison of the overall performances of the optimization methods is enabled.

Reviewer comment: it might have been interesting to include an analysis of how these algorithms and objective functions perform when simulations from a Bartlett-Lewis model are substituted for observed data. Although this would not be sufficient as it would not tell us how well algorithms cope with the difficulty of calibrating a model to real "messy" data, it would have helped evaluate the algorithms and objective functions' ability to find the known "true" values of the parameters.

Reply: This is a valid observation. Such analysis might indeed bring added value to the paper. Therefore, a parameter set was taken from Verhoest et al. (1997) to perform a total of 400 simulations with the MBL model. The length of each simulation is equal to 105 year. These simulations were used to replace the observed data during calibration. For each objective function, and for each optimization method, 400 calibrations were carried out. Ideally, each of these calibrations should result in the retrieval of the known parameter set. However, since this is very unlikely, the distribution of estimation errors will highlight the algorithms and objective functions' ability to find the known parameters. It must be noted that multiple starting points were used for the DSM, 30 to be precise (see previous comment by anonymous referee).
Figure 1 visualises the distribution of the estimation errors on the six MBL model parameters. It can be seen that SIMPSA, PSO, and SCE-UA perform better in identifying the true parameter set in comparison with the DSM with mutiple starting points. Significant differences between the former optimization methods, or between the objective functions, are not clearly visible. To facilitate this, the calibration results are grouped according to the objective function with which they were obtained, regardless of the used optimization method (see Fig. 2). Similarly, Fig. 3 displays the distribution of the estimation error in function of the used optimization method, regardless of the used objective function.
Figure 2 indicates that the use OF3 might lead to better identifiability (this is especially visible for α), however differences are very small.
As for the ability of the optimization methods to identify the true parameters, Fig. 3 confirms the DSM's inability to do so. SIMPSA seems to lead to very large estimation errors on several occasions. PSO seems to be the most consistent in identifying the true parameter, however, its results are comparable to those of SCE-UA, apart from a few outliers.

Reviewer comment: p. 9711, line 26. The description of the models as "deterministic" is misleading. The authors probably meant something like "with fully identifiable parameters". But the meaning of the sentence is also not clear.

Reply: What we are trying to bring across is that usually only one parameter set will be used for practical purposes. Therefore, it is treated as a fully identifiable model. We will rephrase this sentence in the new version as follows: "*ldots* as from a practical point of view the models are usually treated as being fully identifiable, i.e. only one parameter set is used for simulation."

Reviewer comment: The authors could also add that it is not possible to obtain a likelihood function in a closed form, so maximum likelihood is not available as a parameter estimation method.

Reply: We fully agree to this comment. This will be added to section 2.

Reviewer comment: p. 9717, line 24. "constrains" would be better than "enforces"

Reply: We are not sure how the addition of the Simulated Annealing framework to the DSM would constrain the latter's global search. On the contrary, it will strengthen its ability to perform a global search, by allowing occasional 'wrong' moves. Therefore, it will remain phrased as such in the article.

Reviewer comment: p. 9718, line 5 'process'

Reply: Adjusted.

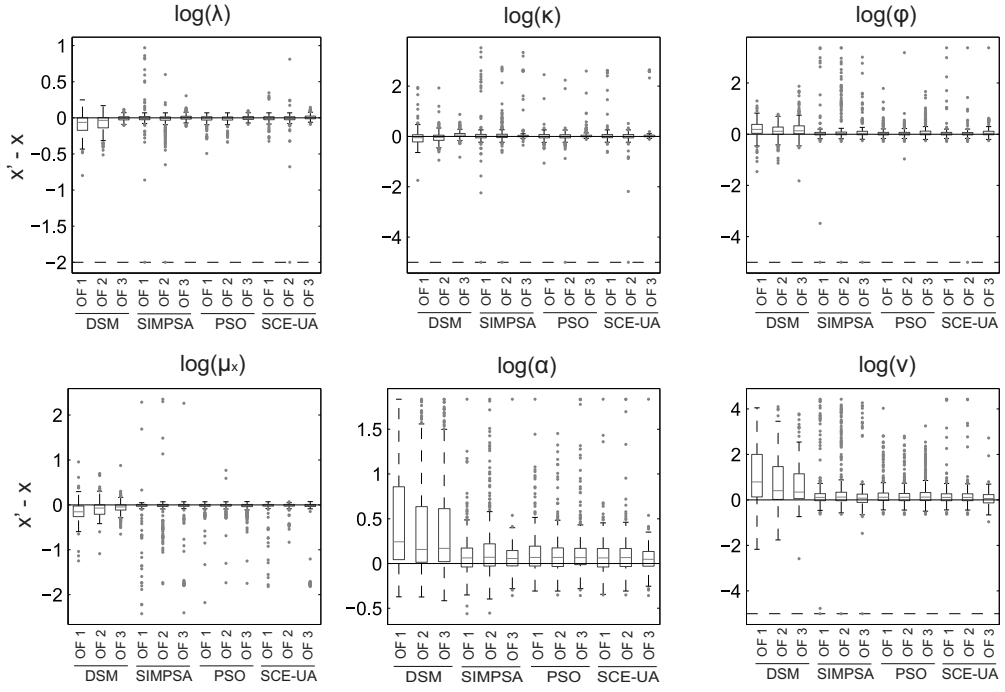

Figure 1: Distributions of estimation error for each parameter under different optimization methods and objective functions.

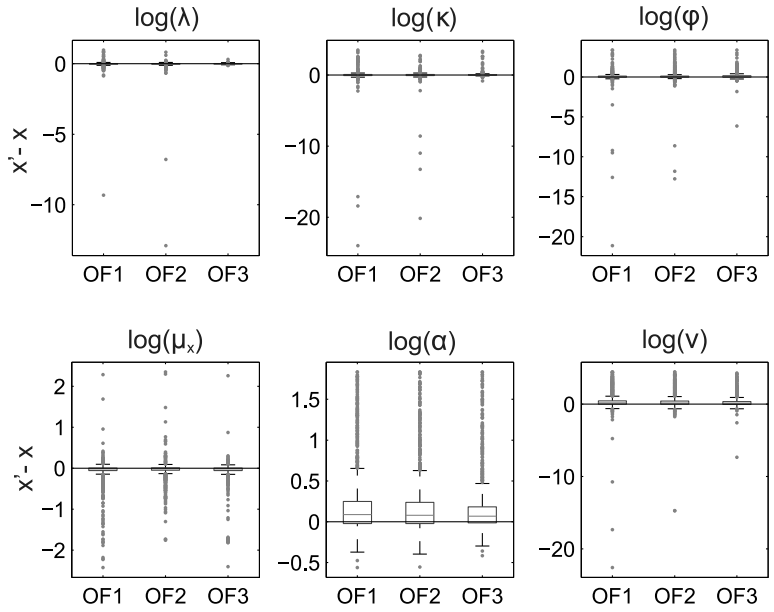

Figure 2: Distributions of estimation error for each parameter, grouped according to the used objective function.

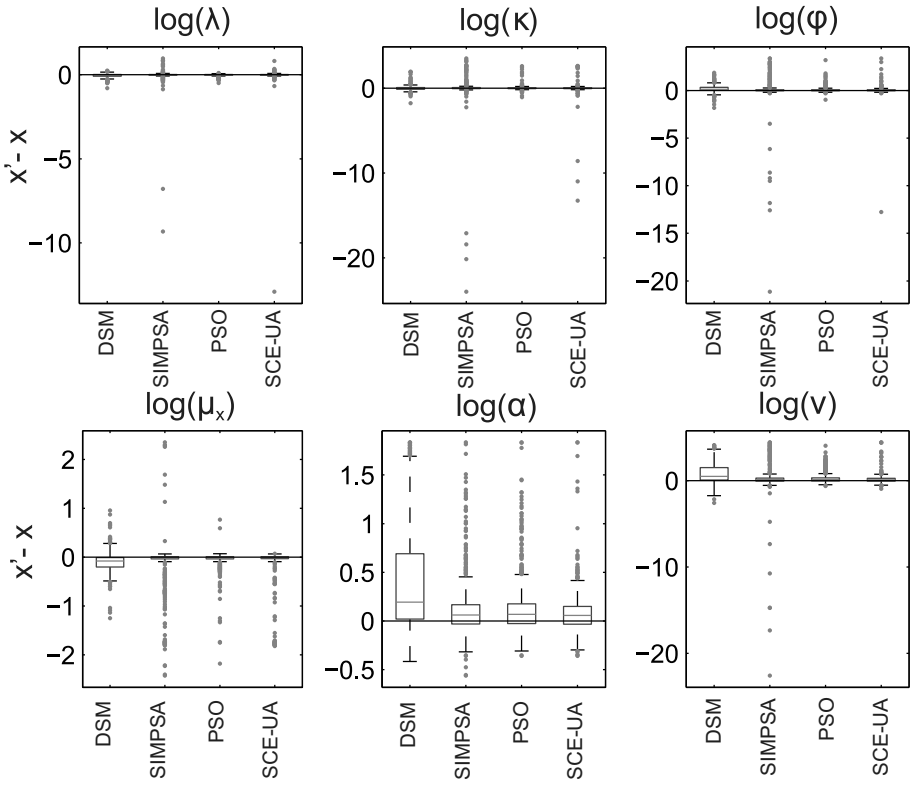

Figure 3: Distributions of estimation error for each parameter, grouped according to the used optimization method.

Bibliography

Verhoest, N. E. C., Troch, P. A., and De Troch, F. P.: On the applicability of Bartlett-Lewis rectangular pulses models in the modeling of design storms at a point, Journal of Hydrology, 202, 108–120, 1997.