Hydrol. Earth Syst. Sci. Discuss., 2, S1295–S1300, 2005
www.copernicus.org/EGU/hess/hessd/2/S1295/
European Geosciences Union

Hydrology and Earth System Sciences Discussions

Interactive comment on "How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration?" *by* Y. Tang et al.

Y. Tang et al.

Received and published: 17 January 2006

Our discussion below first lists Referee #2's specific comments and then our response to each comment.

1. The one issue which is somewhat disappointing is that the authors use only streamflow data at the outlet. Of course, by looking on different aspects of the hydrograph this can be seen as multi-objective problem, but, at least to me, the more interesting aspect of multi-objective calibration is the problem when a model has to be fitted to different kinds of data, such as runoff and soil moisture. Given that such data seems to be available for the Shale Hills watershed I wonder why these data were not used?

Authors' Response: The focus of this paper is on comparing the relative strengths and limitations of the evolutionary multiobjective optimization algorithms performances for real world test cases. The Shale Hills test is computationally intensive and requires the quantification of a 3-objective Pareto surface. Current research efforts are being invested in updating the Shale Hills data to consider real-time groundwater levels and soil moisture data. These series will be analyzed in future work.

2. Another issue which seems not be addressed is the issue that there might be different solutions (i.e. parameter sets) for the same point on the pareto front. This should be discussed.

Authors' Response: We agree that complex parameter sensitivities and correlations could yield non-unique parameter sets that could theoretically have similar Pareto optimal points. With such a large number of parameters and a single output variable, it is to be expected that some of the parameters are insensitive to the objective function(s) selected. Different optimization runs will therefore result in different combinations of parameter sets that provide equal objective function value. It should be noted that the focus of this study is testing the relative performances of the three evolutionary multi-objective algorithms assuming an initial problem formulation. Each of the algorithms faced the same parameter correlations and sensitivities in their search for Pareto optimal points. The fact that different parameter combinations are obtained is not a problem in this context. It is, however, a considerable problem if the actual parameter values should not only be within a feasible range, but they should for example be used within a regionalization study. This issue is discussed at length for example in Wagener and Wheater (Journal of Hydrology, in Press, available on-line), but beyond the scope of this paper.

We should also mention from the convergence theory for evolutionary algorithms (Thierens et al. 1998) that the rate of convergence for individual hydrologic model parameters will be tied directly to their "salience" or impact on the application objectives (i.e., the most important parameters converge fastest). So the Pareto optimal points found in the hydrologic calibration applications used in this study represent the

HESSD
2, S1295–S1300, 2005

solutions that had the largest impact on the specified objectives and the algorithms' selection operators.

3. The authors could consider providing a little more information about e-NSGAII and SPEA2 as the typical HESS reader might not be familiar with those. A schematic figure might be helpful. Also, some more explanation on the test function suite would be helpful. Could the metrics described on p.2485 be clarified using a schematic figure? Figure 1, on the other hand, could be omitted.

Authors' Response: As discussed in our response to Referee #4, we will add an additional figure illustrating the two unary metrics. Figure 1 will be removed from the final manuscript. Given the manuscript's current length, readers interested in the details of SPEA2 or Epsilon-NSGAII have been given the original citations that provide full descriptions of these algorithms (see p. 2471 lines 13-14).

4. P2473, l.1: "publishable precision or error tolerances for their objectives to avoid wasting computational resources on unjustifiably precise results", usually it is not possible to a-priori know what goodness-of-fit can be achieved for a certain catchment (data quality, ...)!

Authors' Response: We agree that it is not possible to know a priori the goodness-of-fit (i.e., the objectives' values after search). This information is not necessary when defining the epsilon tolerances for each objective. Any numerical model requires that the user define the precision of their calculations via its data structures (6-digits, 10-digits, 16-digits, etc.) Our point is that often users will accept the highly precise data structures of an optimization source code that are often defined to be double precision (16-digit precision) without consideration of the computational costs of ranking and selecting solutions at such a high level of precision. For example, epsilons can be defined so that 16-digits of precision in the objectives' values are not used in nondomination sorting when a precision of 6-digits captures RMSE sufficiently given the accuracy of a model and its data.

5. P2476, l.13: "Euclidean norms for measuring distance from neighbour solutions", what is meant, the distance in the parameter space or the distance in the objective-function space?

Authors' Response: All distance calculations where in objective function space. We will clarify this in the final manuscript.

6. P2478, l.5: the effect of the transformation depends on the size/unit of the runoff. If y-values are larger 1 the transformation actually will increase the importance of high flows. Please provide information o y-values (units and average runoff)

Authors' Response: The manuscript had a typo in our description of the Box-Cox transformation. The actual lambda value used was 0.3 (not 3 as listed on p. 2478 line 5). The typo will be corrected in the final manuscript.

7. P2479: the model used for the Shale Hill watershed should be better described, it is not well-known. What are the 4 parameters for each spatial zone? How are these zones delineated?

Authors' Response: The parameters defined for each spatial zone or river section are defined in detail in Table 2 on p. 2503. Given the manuscript's length and focus it should be sufficient for readers to reference the citations recommended in the manuscript on p. 2479 (Duffy 1996, 2004, Lin et al. In Press).

8. Table 4: there are too many digits.

Authors' Response: These results are published in the standard precision used in the computer science literature [e.g., see (Kollat and Reed 2005)].

9. The information provided in fig 5b and fig 9 does not require a figure.

Authors' Response: The metric results presented in the paper are based on the best reference sets generated across the algorithms. These figures concisely show how the algorithms contributed to these reference sets. We feel these results warrant figures.

HESSD

2, S1295–S1300, 2005

Full Screen / Esc

Print Version

Interactive Discussion

Discussion Paper

EGU

10. For single-objective calibration the SCE-UA algorithm is often seen as the best algorithm. Given the results presented in this paper I wonder whether this still holds. I would like the authors to comment on this. Can the SCE-UA still be claimed to be best or is a similar study for single-objective calibration needed?

Authors' Response: MOSCEM and SCE are similar in name and use some of the same underlying ideas. However, the ways these ideas are implemented are very different and the SCE has been shown to outperform other algorithms in many published studies. It would be inappropriate to make a generalized conclusion on an algorithm that was not tested in this study. The conclusions made here only apply to the MOSCEM algorithm.

References Used in Response: Duffy, C. J. 1996. A two-state integral-balance model for soil moisture and groundwater dynamics in complex terrain. Water Resources Research 32:2421-2434.

Duffy, C. J. 2004. A Distributed-Dynamical Model for Mountain-Front Recharge & Water Balance Estimation: The Rio Grande of Southern Colorado and New Mexico. in H. a. Scanlon, editor. AGU Monograph: Recharge in Semi-Arid Regions: State of the Art.

Kollat, J. B., and P. Reed. 2005. The Value of Online Adaptive Search: A comparison of NSGA-II, ε-NSGAII, and εMOEA. in C. Coello Coello, A. Hernandez, and E. Zitzler, editors. Evolutionary Multi Criterion Optimization: Third International Conference (EMO 2005). Lecture Notes in Computer Science Springer-Verlag, Guanajuato, Mexico.

Lin, H. S., W. Kogelmann, C. Walker, and M. A. Bruns. In Press. Soil moisture patterns in a forested catchment: A hydropedological perspective. Geoderma.

Thierens, D., D. E. Goldberg, and A. G. Pereira. 1998. Domino Convergence, Drift, and the Temporal-Salience Structure of Problems. Pages 535-540 in The 1998 IEEE International Conference on Evolutionary Computation. IEEE Press.

Wagener, T. and Wheater, H.S. In Press. Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty. Journal of Hydrology, In Press, Corrected Proof, Available online 2 September 2005

Interactive comment on Hydrology and Earth System Sciences Discussions, 2, 2465, 2005.