Hydrol. Earth Syst. Sci. Discuss., 2, S1229–S1231, 2005
www.copernicus.org/EGU/hess/hessd/2/S1229/
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: 5 January 2006

We highly appreciate the detailed and thoughtful review provided by Referee # 3. The final submitted manuscript will incorporate the suggested minor technical corrections numbered 1-20. Referee # 3 was correct in identifying a typo on Figure (6). The horizontal axes should read "Function Evaluations (x10000)". This typo will be corrected in the final manuscript. Additionally, the final manuscript's notation for the interpretation function (see p. 2485) will be clarified using Referee # 3's suggestion in their second specific comment.

The remainder of this comment will address Referee # 3's questions regarding performance differences between epsilon-NSGAII and SPEA2.

Reviewer Specific Comment (1) discussing p. 2471, lines 23-26 in the manuscript:

The authors mention that epsilon-NSGAII out performs SPEA2 (in terms of ease of use, reliability, and the provision of more diverse representations of tradeoffs) on a four-objective groundwater monitoring application presented in Kollat and Reed [2005b]. Given the audience of HESS, and the fact that SPEA2 appears to outperform epsilon-NSGAII on the test function suite and the two watershed calibration applications presented in this study, I believe readers of HESS would be interested in the reasons that might account for this. Can the authors provide suggestions (or a few sentences in the discussion section) to provide further explanation? What are the key differences between the applications provided in these two papers?

Authors' Response:

Some clarification is necessary on the reviewer's interpretation of epsilon-NSGAII's performance on the test function suite and two watershed calibration applications. For the test function suite (see p. 2487-2488) epsilon-NSGAII's convergence and diversity measures were proven to be superior to SPEA2's at the 99-percent confidence level. Additionally, epsilon-NSGAII attained these results using one-third of the evaluations. SPEA2 had superior performance on the Leaf River case study. Epsilon-NSGAII was competitive with SPEA2 when solving the Shale Hills case study. On p. 2491 it states that the two algorithms' mean performance metrics' distributions where statistically similar. In the same discussion, the results show that SPEA2 had a slightly higher reliability but epsilon-NSGAII generated 94-percent of the best known solutions.

The performance differences between epsilon-NSGAII and SPEA2 for this study as well as Kollat and Reed [2005b] result from two primary algorithmic differences (1) epsilon-NSGAII's dynamic population sizing and epsilon dominance archiving and (2) SPEA2's use of a fixed archive size that has to be specified a priori by the user. As we discuss in lines 10-14 on page 2493 the biggest challenge in using SPEA2 is specifying an appropriate archive size without prior information on the Pareto optimal set.

These issues do not arise for the epsilon-NSGAII since the algorithm's epsilon dominance archiving will automatically generate a set size based on the user's precision requirements for each objective. SPEA2's performance is heavily impacted by the user specified archive size because this parameter controls the degree of clustering used in its diversity operators.

Performance differences between the calibration applications in this paper and the long-term monitoring application addressed in Kollat and Reed [2005b] are related to the problems' structures and SPEA2's archiving. Kollat and Reed [2005b] tested the algorithms' performances in solving a discrete optimization problem where the 4-objective Pareto optimal set was known from enumeration. In this case, SPEA2's archive was set using the known size of the Pareto optimal set whereas epsilon-NSGAII was simply given precision requirements for each of the objectives without any additional information. The results show that SPEA2's clustering operators where not as effective in producing diverse representations of the known Pareto optimal set. Additionally, these studies and prior work have shown SPEA2 often will not converge to the exact Pareto optimal set.

In the calibration applications (see p. 2483-2484) the Pareto optimal sets for these continuous search spaces are unknown and SPEA2's performance was maximized by setting its archive size to the average archive size attained by epsilon-NSGAII. This allowed SPEA2 to maximize its clustering-based diversity operators and maximized the algorithms' performance metrics. It should be noted that in our study we have maximized SPEA2's performance using information generated automatically from epsilon-NSGAII's runs. We structured our study to maximize SPEA2's performance to ensure our algorithm was rigorously tested.

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

HESSD

2, S1229–S1231, 2005

Interactive Comment