Hydrol. Earth Syst. Sci. Discuss., 6, S665–S672, 2009
www.hydrol-earth-syst-sci-discuss.net/6/S665/2009/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Multi-objective calibration of a distributed hydrological model (WetSpa) using a genetic algorithm" *by* M. Shafii and F. De Smedt

M. Shafii and F. De Smedt

Received and published: 4 June 2009

Responses to comments of the Anonymous Referee 1

We would like to thank this reviewer for the provided comments. To frame the response, the Authors, within this text, provide the comments in their original format, followed by corresponding responses. Furthermore, as there was a need to re-organize and rectify the paper, a revised version of the paper will be uploaded for consideration of the reviewer. In order to better trace the corrections, the edited/added parts will appear in red in the revised paper.

1. ”Significant progress has been made in the development of multi-objective global search algorithms in the past few years and highly efficient algorithms are available today. It is unfortunate that the authors did not choose from these more powerful tools. NSGAII has also experienced significant development since it was first presented in 2002, for example by the introduction of enhanced Epsilon Dominance (e-NSGA-II). The e-NSGA-II has been proven superior to its parent algorithm as shown by Kollat and Reed (2005) and Tang et al. (2006).”

Response: We agree that there has been a huge progress during past years in development of optimization algorithms, especially the population-based evolutionary algorithms, aiming at improving the efficiency of optimization models, though sometimes at the cost of introducing further complexities into the process. However, our experience in this respect is limited. In this paper, the main goal is to adopt a multi-objective formulation for calibration of the WetSpa model, along with performing identifiability analysis of the parameters. Although achieving the highest efficiency of the search optimization routine is crucially important, we selected NSGA-II as a proper trade-off between efficiency and complexity in multi-objective evolutionary optimization.

2. ”The comparison of the NSGA-II and PEST results in the form it is presented in the paper is problematic because significant differences between the two algorithms are not clearly described and accounted for in the paper. This has also resulted in misleading conclusions about the performance of the algorithms. NSGA-II is a non-linear, global search algorithm with the capacity to explore the entire parameter space. On the other hand, PEST is a linear, local search algorithm and therefore the results are dependent on starting values and on the complexity of the search problem. PEST is naturally placed in a single-objective context. Therefore it can be expected that the final solution might not compare favorably with other objectives, especially if trade-off exists between them. But in the paper parameter solutions obtained from PEST are compared to Pareto efficient solutions determined by NSGA-II. Note that the PEST solutions can be compared with the ”Pareto extreme” of objective 2. It is no surprise

HESSD

6, S665–S672, 2009

Interactive Comment

that PEST found a similar solution when started with parameter values from Pareto solution 18".

Response: More detailed information about PEST, NSGA-II, and their differences, along with re-interpretation of the results has been provided in the revised paper (i.e. Section 2.4; and Section 3.1, paragraphs 4 to 6). The main purpose of comparison between NSGA-II and PEST in this paper is not to compare them in terms of efficiency or technical aspects, but to verify the performance of the multi-objective evolutionary algorithm to locate optimum results. The reasons for applying PEST are emphasized in Section 3.1., paragraph 4. The results and conclusions have been re-interpreted accordingly, in the same section paragraphs 5, and 6.

3. "The fact that the solution is actually "better" (a higher CR2 function value) than the best NSGA-II for this objective suggests that the NSGA-II has not found the Pareto extreme for objective 2."

Response: The final solution obtained by PEST is only slightly better than those of NSGA-II in terms of CR2. Moreover, (i) NSGA-II deals with parameter sets which are optimum for multiple objective functions, and (ii) PEST is a local search technique which requires an appropriate starting point. Hence, the results obtained with NSGA-II enabled PEST to find the "best" solution. We have changed the conclusions on this accordingly (see Section 3.1, paragraph 6).

4. "The study confirms existing knowledge that linear search algorithms are not as efficient when dealing with non-linear optimization problems. But it is essential for the comparison of the NSGA-II and PEST algorithms to use the same objective function. This has not been done in this study and therefore the results need careful interpretation. Multiple starting values for linear search algorithms such as PEST can lead to better solutions but it is not known how many are needed to find the global solution for a particular objective function."

HESSD

6, S665–S672, 2009

Response: As previously explained, the goal of this study is just to adopt a new approach for calibration of WetSpa model which was traditionally calibrated only with PEST. This has been corrected in the paper and conclusions have been re-interpreted accordingly.

5. "The three different criteria to be optimized by NSGA-II are using the same data (stream flow). It is a desirable feature in multi-objective optimization that the objectives are contrasting. If no trade-off exist between different objectives, then a single-objective aggregate would probably be more efficient to use."

Response: We fully agree that objective functions within a multi-objective formulation are usually contrasting which is the essence of multi-objective optimization. This seems really not to be the case in this study, but this is not quite true, because CR2 puts more emphasis on high flows while CR3 puts more emphasis on low flows, which are contrasting objective functions. Therefore, the objective functions are expected to serve well in a multi-objective formulation of WetSpa calibration. Aggregation of the objective functions compared to Pareto-based optimization techniques would require a higher number of function evaluations, consequently leading to higher computational costs.

6. "The first criterion, CR1, is actually not the mass balance of the WetSpa model but one component, namely the stream flow. The CR1 values are one order of magnitude smaller compared to the CR2 and CR3 values. Therefore the objective function values should be normalized in the optimization scheme."

Response: Traditionally, CR1 is considered to be error in the mass balance, though it is called model bias in the original paper. Because, given the precipitation as being correct, WetSpa estimates the evapotranspiration, and as a result, river flows are obtained. This criterion has been corrected in the revised paper (e.g. Section 2.3, one paragraph before the last one, line 3). Because CR1 is a relative measure, it is already

HESSD

6, S665–S672, 2009

Interactive Comment

normalized, thus located between 0 and 1, while the optimum value for this objective function is 0. Likewise, CR2 and CR3 are also normalized with values between 0 and 1, but with an optimum value of 1. Therefore, there is consistency between the values of these three objective functions, and there is no need for any more normalization.

7. "The bi-criterion plots in Figure 3 need more discussion, e.g. on the trade-off between objectives, the sampling density along the Pareto fronts, etc. The small number of Pareto efficient solutions suggests that the Pareto surface has not converged at the time the search was terminated. 4000 model evaluations appear to be too less for the 11-dimensional problem and 27 Pareto points seem to be a too small number to represent the Pareto surface in the 3D objective space."

Response: More detailed information about this Figure has been provided in the revised paper (Section 3.1, paragraph 2). Obviously, a higher number of function evaluations would lead to better results with a higher density of Pareto solutions. But the computational cost would also be much higher. Hence, this is a trade-off between cost and accuracy, and nevertheless, we obtained good results that show performance of the multi-objective optimization routine to simulate the hydrologic behavior of the catchment, along with giving an insight into the identifiability of the parameters. Furthermore, in the revised paper (Section 3.1, paragraph 1), we show that convergence of the algorithm has been achieved with this number of Pareto front solutions using the C-Function of Zitzler and Thiele (1999).

8. "The methodology can and should be presented in a more clear and concise way. At several occasions paragraphs from the Results and Discussion section should be better attributed to the Methods section. I would propose the following change in structure: 2.1. Study area, 2.2. WetSpa Model, 2.3. NSGA-II algorithm (including the formulation of optimization problem), 2.4. PEST."

Response: We agree with the reviewer that the paper should be reorganized. Accord-

HESSD

6, S665–S672, 2009

Interactive Comment

ingly, it has been done so in the revised paper.

9. "The description of NSGA-II could be more concise (e.g. list individual steps of the algorithm) and should be specific to the optimization problem at hand (i.e. chromosome = parameter set, etc.). List values for all algorithmic parameters such as crossover and mutation probabilities, crossover rate, termination criteria, etc. and explain their meaning. Include appropriate references."

Response: Step-by-step explanation of NSGA-II with the problem-specific meaning of the parameters and corresponding references has been provided in the revised paper according to the advice of the reviewer (Section 2.3, after the paragraph 2).

10. "Based on the comments above, the discussion and conclusions need to be revised carefully, especially for the comparison of NSGA-II and PEST results."

Response: The discussion and concluding remarks have been rephrased in the new paper, as suggested.

Specific comments

a. *Replace the term "Evaluation criteria" with "objective function":* This has been corrected in the revised paper.

b. *The third objective function is rather the coefficient of efficiency (Ce) of log transformed flow values than the log transformed Ce. No bar over the ln in Eq. 3.:* It is correct that CR3 is the Ce of log-transformed discharges. Therefore, we must deal with the average of log-transformed values (i.e. average of ln(discharge)). In this way, the bar over the ln is mandatory..

c. *What is the reason that the values of the 3rd objective performs are better during the evaluation period? Are there less low-flow events occurring during 1996-2000 as*

HESSD

6, S665–S672, 2009

compared to the calibration period?: Indeed, as suggested by the reviewer, the validation period is much dryer; hence, there are more and even smaller low-flows. This results in errors of smaller magnitude and thus better CR3 values. This been added to the revised paper, Section 3.1, paragraph 1, the last four lines.

d. *"Similar" CR1 values are reported for the calibration and the evaluation periods. However, in most cases the values are larger during the evaluation period. This needs clarification and discussion.:* As discussed before, the optimum value for CR1 is zero. Hence, the objective function values, obtained in calibration process, are expected to be more in the evaluation period. So, it is logical that the CR1 objective function values are larger for the validation period, due to the uncertainties associated with the calibration process.

e. *What is the significance of plotting min/max/average objective function values in Fig. 4? The convergence of the NSGA-II algorithm should be measured by the change of the shape of the Pareto surface and the sampling density along Pareto fronts.:* In the revised paper, we show the convergence of the algorithm with C-Function (Zitzler and Thiele, 1999). Information on this index is provided in the revised paper, Section 2.3, paragraphs 3 to 5. The old Figure 4 has been omitted from the revised paper, and instead, the variation of C-Function values over iterations is shown (Fig. 3 in the new paper).

f. *NSGA-II should find the same Pareto solutions independently from the starting values. Consider omitting Fig. 8 and summarizing the results in the text.:* Figure 8 has been left out according to the suggestion of the referee, and the text is provided in the revised paper, Section 3.2, last paragraph.

g. *Please comment how the confidence intervals for the PEST solutions (Table 2) were determined.:* This information has been provided in the revised paper, Section 2.4, paragraph 2.

h. *Please note for the discussion on parameter uniqueness that the results are possi-*

HESSD

6, S665–S672, 2009

Interactive Comment

bly a reflection of the small number of Pareto efficient points and the formulation of the objective functions. It would be desirable to adapt a more robust method to quantify parameter uncertainty.: We have added a general discussion on parameter uncertainty in Section 3.2, paragraph 3. Parameter uncertainty assessment is now ongoing by the authors and results will be presented in a future paper. This is also mentioned in the last paragraph of the conclusions in the revised paper.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 6, 243, 2009.