Hydrology and Earth System Sciences Discussions

Interactive
Comment

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

Anonymous Referee #1

Received and published: 13 March 2009

The study compares the efficiency of two model calibration algorithms, NSGA-II and PEST, to calibrate the hydrological model WetSpa using stream flow data from the Hornad River in Slovakia.

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. NSGA-II 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

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

(2005) and Tang et al. (2006).

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

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.

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. 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.

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.

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 2.1. WetSpa Model 2.3. NSGA-II algorithm (including the formulation of optimization problem) 2.4. PEST

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.

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

Specific comments

Replace the term "Evaluation criteria" with "objective function".

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.

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 compared to the calibration period?

"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.

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.

NSGA-II should find the same Pareto solutions independently from the starting values. Consider omitting Fig. 8 and summarizing the results in the text.

Please comment how the confidence intervals for the PEST solutions (Table 2) were determined.

Please note for the discussion on parameter uniqueness that the results are possibly 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.

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