Hydrol. Earth Syst. Sci. Discuss., 8, C2365–C2370, 2011
www.hydrol-earth-syst-sci-discuss.net/8/C2365/2011/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Robust multi-objective calibration strategies – chances for improving flood forecasting" *by* T. Krauße et al.

A. Efstratiadis (Referee)

andreas@itia.ntua.gr

Received and published: 23 June 2011

General comments

1. The paper presents a framework for model calibration, in which data depth measures are used within a new multiobjective optimization algorithm, to identify robust non-dominated solutions. The effectiveness and efficiency of the algorithm are tested on the basis of two typical benchmark problems, while the entire framework is employed in a real-world case study, involving the calibration of a hydrological model (WaSiM) against a number of flood events, in a small experimental catchment in Switzerland.

2. The topic of the paper is interesting and the manuscript is well-organized and well-

written. However, its originality and novelty are questionable. For, there are two papers that have been recently submitted to HESSD dealing with a very similar subject, where the same algorithms, the same model and the same study area seem to be recycled (Krauße and Cullmann, 2011a, b). Parts of the text are verbatim reproduction, while some of the tables and figures are repeated. In order to be suitable for publication, a substantial review is essential, to remove the already published components of the paper and provide really original material.

3. Despite the very promising title ("chances for improving flood forecasting") and some important hints that are discussed mainly in the first two sections (and have been also discussed in the two aforementioned papers), my final impression was rather about "another calibration exercise". Specifically, the very challenging task of identifying "robust" (realistic? behavioural?) parameters, which is of major interest in hydrological modelling, is addressed just as an algorithmic issue that is handled through the so-called "Robust Parameter Estimation" (ROPE) approaches. Naming a computational procedure, even the most sophisticated one, as "robust" is, to my point-of-view, not useful, impossible to understand and even misleading. One can find a large number of alternative calibration methods and strategies in the hydrological literature – which of them are robust and under which premises? Are the SCE and GLUE methods, with thousands of applications (and citations) worldwide, robust or not? Who is able to identify the most robust solution, an expert hydrologist or an "expert" algorithm?

4. In general, the parameter estimation problem is satisfactory posed, although some of its aspects may required more development (e.g. the concepts of uncertainty and parsimony are rather poorly explained). The authors, quoting Bardossy and Singh (2008), rightly state that a key goal of model calibration is to find parameters that perform well both in calibration and validation, and at the same time ensure "hydrologically reasonable representation of the corresponding processes" (p. 3699, line 22). The data depth technique, which was initially proposed by Bardossy and Singh (2008) for single-objective functions and now generalized for multiobjective calibration, is next

introduced as “a possible approach to achieve this goal” (p. 3696, line 14). However, there is until now little experience with this strategy, to justify such an imperative statement. In addition, it is very difficult to trust any automatic method not accounting for the role of knowledge, in terms of hydrological experience and understanding (cf. Boyle et al., 2000). There are also some practical disadvantages, which are revealed in the case study. Why implementing a computationally expensive technique with negligible physical interpretation, just for rejecting part of the non-dominated solutions that lie in the extremes of the Pareto front? As the authors claim “the tails of the Pareto front estimated in the calibration are not required” (p. 3711, line 20). However, this is not a surprising conclusion: even an elementary approach, based on subjective yet realistic cut-off thresholds, could easily distinguish such “non-behavioural” solutions with negligible effort (cf. Efstratiadis and Koutsoyiannis, 2010).

5. Regarding the presentation of the methodologies, I am afraid that the authors deal with too many issues, thus failing to adequately develop their ideas and highlight their effort. This mainly involves the MO-PSO-GA algorithm, which is presented in a too synoptic manner (section 3.1) that makes difficult to understand the procedures and, especially, the innovations (the search scheme is not fully original, but it is based on an effective combination of various techniques). The same problem exists with the GenDeep function (section 3.2), which was very hard to understand, without referring to the literature. Since this is a relatively new method, I would suggest spending sufficient effort on explaining the details, and, at the same time, drastically eliminating (or even removing) section 2.1, since the concepts and definitions of Pareto optimization are rather trivial.

6. The testing framework for evaluating the performance of the MO-PSO-GA algorithm, on the basis of two rather simple benchmark problems, is insufficient. To make sense, this test should involve a representative sample of multiobjective functions, including high-dimensional problems (in terms of both the number of control variables and the number of objectives), and different levels of complexity, regarding the geometry of the

Pareto front.

7. Although the title focuses on flood forecasting, little attention is given to the specific aspects, challenges and peculiarities of this problem. The authors could also take advantage from related applications (e.g., Pappenberger et al., 2007; Moussa and Chahinian, 2009), thus providing a much more attractive paper.

Minor comments and technical corrections

1. Page 3697, line 2: “The developed approach is tested on synthetical data.” I do not agree characterizing the benchmarks problems as “synthetic”. The term is used when contrasting to actual or historical conditions.

2. Page 3697, line 20: “. . . where θ = (θ1, . . ., θ2) is a d-dimensional vector” Use bold fonts for and vectors and change θ2 by θd.

3. Page 3698, line 11: “Often both terms [Pareto set and Pareto front] are used synonymously.” The authors have right, but they should further emphasize on the negative impacts of this practice, which often leads to misleading conclusions.

4. Page 3700, lines 5-8: “One starting point which recently attracted rising scientific interest is a more intelligent selection of the calibration data . . ., another one is the development of advanced methods for the identification of parsimonious model parameters” Parsimony is associated to the model structure, not the parameters. It is a key concept in modelling, asking to represent a model structure with as few parameters as possible, where the essential number of parameters depends on the available information. Please, see the related discussion and the literature provided by Efstratiadis and Koutsoyiannis (2010).

5. Page 3700, line 29: It is preferable using “simple” instead of “small” (example).

6. Page 3703, line 3: Change to read “population-based”.

7. Page 3705, eq. 2: It is very hard to understand this equation. What is the vector u?

What is the symbol T?

8. Page 3705, lines 13-17: “The developed solution addresses some of the drawbacks of existing multi-objective and robust single-objective calibration procedures. It provides a good possibility for the identification of robust model parameter vectors with respect to multiple calibration objectives.” This statement is not justified and should be removed.

9. Page 3706, section 3.3.1: How are the constraints of test function 1 handled?

10. Page 3709, line 3: Change to read “Wolpert and Macready (1997)”.

11. Page 3710, lines 15-16: “In a first case study we calibrated WaSiM with the MO-ROPE algorithm in terms of two objective functions: rPD and NS. Additionally we applied the single-objective robust parameter estimation algorithm ROPEPSO to this problem using rPD, NS and their aggregate FloodSkill as objective functions.” The three functions (PD, NS, FloodSkill) are not defined.

12. Page 3715, lines 21-25: “This underlines that robust parameter estimation can identify the most robust solutions within the given constraints. However, a good selection of appropriate calibration objectives and a suitable model structure are as important as a reliable and robust model parametrisation.” This explains why the use of term “robust” should be done more carefully. Robustness is a combination of all the aforementioned aspects, i.e. the model structure, which should be as parsimonious as possible, the data, which should be as representative as possible, and the calibration.

References

Bardossy, A. and Singh, S. K.: Robust estimation of hydrological model parameters, Hydrol. 15 Earth Syst. Sci., 12, 1273–1283, doi:10.5194/hess-12-1273-2008, 2008.

Boyle, D. P., Gupta, H. V., and Sorooshian, S.: Toward improved calibration of hydrologic models: combining the strengths of manual and automatic methods, Water Resour. Res., 36(12), 3663-3674, 2000.

Efstratiadis, A., and Koutsoyiannis, D.: One decade of multiobjective calibration approaches in hydrological modelling: a review, Hydrol. Sci. J., 55(1), 58-78, 2010.

Krauße, T. and Cullmann, J.: Identification of hydrological model parameters for flood forecasting using data depth measures, Hydrol. Earth Syst. Sci. Discuss., 8, 2423–2476, doi:10.5194/hessd-8-2423-2011, 2011a.

Krauße, T. and Cullmann, J.: Towards a more representative parametrisation of hydrological models via synthesizing the strengths of particle swarm optimisation and robust parameter estimation, Hydrol. Earth Syst. Sci. Discuss., 8, 2373–2422, doi:10.5194/hessd-8-2373- 2011, 2011b.

Moussa, R. and Chahinian, N.: Comparison of different multi-objective calibration criteria using a conceptual rainfall-runoff model of flood events, Hydrol. Earth Syst. Sci., 13, 519-535, doi:10.5194/hess-13-519-2009.

Pappenberger, F., Beven, K. J., Frodsham, K., Romanovicz, R. and Matgen, P.: Grasping the unavoidable subjectivity in calibration of flood inundation models: a vulnerability weighted approach. J. Hydrol. 333, 275-287, 2007.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 3693, 2011.