Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-259-AC3, 2019

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Technical note: the caRamel R package for Automatic Calibration by Evolutionary Multi Objective Algorithm" *by* Céline Monteil et al.

Céline Monteil et al.

celine-c.monteil@edf.fr

Received and published: 29 August 2019

We would like to thank Andreas Efstratiadis for constructive comments on our manuscript. The comments are helpful and will certainly improve the quality of the paper. Below we provide the Reviewer's comments verbatim in black italic text, and our responses are immediately below each comment in blue text.

My overall opinion about this article is positive, yet in its current form cannot stand neither as a technical note not as research paper. Actually, it rather resembles to an extended abstract of a clearly hard and long research, which may be useful as a brief

documentation for the R community, but is not suitable for a top hydrological journal such as HESS. First of all, the authors have to decide the orientation and the objectives of this article. There are several alternatives, i.e. (a) a state-of-the-art discussion of the multiobjective calibration problem in hydrology; (b) a comprehensive description and justification of the algorithm and its technological advances, accompanied by extended tests of its performance against problems of varying complexity and against other well established methodologies, and (c) a more synoptic description of the algorithm, with emphasis to its application to few (not only one) representative hydrological calibration problems of varying difficulty, to be presented and discussed in detail.

We intend to clarify the orientation of the article with the alternative (c), by giving more details on the algorithm and by adding a new application in hydrology with the code GR4J (Coron et al., 2017; Coron et al., 2019). We also suggest to change the title as "Multiobjective calibration by combination of stochastic and gradient-like parameter generation rules: the caRamel algorithm".

Specific comments

Page 1, lines 13-14: "The main function of the package, caRamel(), requires to define a multi-objective calibration function as well as bounds on the variation of the underlying parameters to optimize". Too obvious technical detail to be referred in the abstract.

Thank you, we will correct it.

Page 1, lines 24-25: ". . . it is well-know that errors in a simulated discharge time series are not normally distributed, and do not have constant variance and autocorrelation." This statement is true (a reference would be helpful, e.g. Sorooshian and Dracup, 1980), but is not so much evidently linked with the need for multiobjective calibration. Actually, the multiobjective approach in hydrological modelling covers much more cases, including fitting to multivariable and multisite data, as well as soft information

(cf., Madsen, 2003; Efstratiadis Koutsoyiannis, 2010).
We will add the references you mentioned.

Page 1, lines 28-29: "Evolutionary algorithms have become widely used to explore the Pareto-optimal front in multi-objective optimization problems that are too complex to be solved by descent methods". Do they exist descent methods for multiobjective optimization? Maybe you refer to classical aggregation approaches (e.g. weighting of criteria) that have to be solved multiple times with different weighting values, in contrast to evolutionary approaches that only require "a single optimization run", as correctly mentioned just after (page 2, line 2).
Yes, we refer to descent methods with aggregation approaches.

Page 2, lines 5-6: "The caRamel optimizer has been developed to meet the need of an automatic calibration procedure that delivers not only one but a family of parameters sets that are optimal regarding a multi-objective target". There do exist many algorithms covering this general objective. Is there any specific objective for the development of caRamel? Which shortcomings of the existing algorithms have you detected before deciding building a new method?
Our feeling is that most of MOEA rely mainly on stochastic generation rules, with few deterministic aspects. The idea in caRamel is of course to keep these stochastic, "global" mechanisms (such as recombination or multivariate sampling using the covariance) but also to make place for more "local" mechanisms, such as extrapolation along vectors in the parameter space which are associated with an improvement in all objective functions (a "gradient-like" approach extended to the set of objective functions, in a qualitative way). A shared feature between caRamel and MEAS is the use of the simplexes in which generational rules are applied. However, in MEAS these simplexes are randomly chosen, with the sole constraint that at least one vertex is in the approximated Pareto front; conversely, in caRamel the choice of the simplexes

is entirely deterministic since they are the result of the Delaunay triangulation of the individuals in the objective space (with each objective scaled by the specified precision), and the probability of using a given simplex for generating new individuals is proportional to the volume of this simplex in the scaled objective space. The same kind of geometrical rationale applies for the selection of edges along which an "expansion" is tested (see the description of the rules when added).

Page 2, line 17: Terms "flood objective" and "low flow objective" are unclear (at least for a non-expert).
We will give more detail.

Page 2, lines 17-19: "Multi-objective calibration is also a way to add some constraints to an underconstrained problem when many parameters have to be quantified. This can help to reduce the equifinality of parameters sets". More discussion should be made here (for 30 years, equifinality remains a hot topic in hydrology), including some representative references, e.g. Her and Seong (2018).
We will add the reference.

Page 2, lines 20-21: "Equifinality may be caused by the model structure, when two sets of parameters give similar results. Another kind of equifinality is related to the calibration objectives, when two different model results give similar objective values." Term "result" is unclear - probably you refer to the model outputs, by means of response time series. In this respect, two different parameter sets, except if they are very close, cannot provide the same outputs (i.e., similar individual values), they can provide outputs with similar statistical characteristics, and thus similar performance metrics, as correctly stated in the second phrase.
We propose to rephrase: "Equifinality may be caused by the model structure, when two sets of parameters give similar model outputs due to interactions between model

parameters."

Page 2, line 28: Please, also cite the more detailed and peer-reviewed paper by Efstratiadis and Koutsoyiannis (2008), published as book chapter.
We will add the reference.

Page 3, section 2.1.1 (Generation rules): The description of the algorithm is very poor and only provides a very general idea about the generation mechanisms. How are these rules associated with the ones used in MEAS? I see quite many differences and very interesting ideas implemented here, but the text is too short to allow understanding and evaluating the methodology (and its potential novelties). Figure 1 is also little helpful; for instance, green and blue points, indicating new sets, are missing, although they are referred in the legend.
We will expand this section and try to clarify it.

Page 3, line 25: Why you keep points of the lower level? Aren't they dominated by points of the upper one?
The non-dominated level number is 1, so points of upper level are dominated by points of lower levels.

Pages 4-5, section 2.2 (The caRamel R package): This section is very technical and not so much relevant with the broader philosophy of HESS.
We propose to move it to an appendix.

Page 5, line 15: "The diversity which can be described with two aspects: the spread of the set . . ." Diversity may refer both to the parameter and the objective space. Which of the two sets are used here?

We refer to diversity in the objective space.

Page 5, line 25: Please, cite Deb et al. (2002) who developed NSGA-II.
We will add the reference.

Page 6, lines 12-13: "Comparison with MCO (NSGA-II only) shows that the use of MEAS makes the optimization process converge more rapidly but with a lower diversity". Can you explain the reason of this behavior? Is this an inherent drawback of caRamel, or is due to the algorithmic inputs used in this experiment? As shown in Table 1 (and similarly to all hybrid optimization schemes), caRamel uses quite a large number of input arguments that need manual tuning. Did you run the algorithm by testing alternative set-ups? Do you have any recommendations for the users, regarding the selection of these inputs?
The user may choose to give more weight to some of the rules in the input arguments. We tested different combination but our feeling is that it is better to have a "balanced" approach with the same number of parameter sets generated for each rule.

Page 7, section 4.2 (Hydrological modeling): Your case study does not allow extracting safe conclusions about the performance of your method and its comparison against NSGA II. The key reason is that the use of a single overall metric, i.e. KGE, ensures almost perfect fitting to observations (KGE = 95).
The KGE had been computed on 3 different flow signatures, so we consider them as 3 metrics. For the other example we plan to add, we will also use Nash-Sutcliffe Efficiency as a metric.

Page 8, line 8: Please, better explain criteria (2) and (3) and the associated signatures. Have been these criteria used elsewhere? If yes, please also provide the associated

references.
These criteria are also used in Rouhier et al (2017). We will add the reference.

Page 9, lines 3-4: "MCO has been used with crossover probability set to 0.5 and mutation probability to 0.3". Have you made any preliminary tests before selective these values? Which are the values applied to the input arguments of caRamel?
Please find below (Fig. 1) a heat map of the Hypervolume metric regarding cprob and mprob for the hydrological model calibration with mco, that helps us to chose the values. We will add the values of the input arguments for caRamel.

Page 9, Figure 5: I find your figure a little bit misleading. In the vertical axis, the spread of solutions is very small, and within the anticipated range of uncertainty induced in any hydrological calibration exercise. For instance, the lower value of KGEamd is 0.83, while the higher is 0.86. From my point-of-view, such differences do not make sense in the real world.
For the additional hydrological example, we will not use the KGEamd objective.

Page 10, line 12: How did you selected the best compromise parameter set? What do you mean by term "observed set"?
The best-compromise set has been selected regarding to the distance to the point (1,1,1) in the objective space. The "observed one" stand for measured streamflow. We will rephrase it.

Page 11, section 5 (Conclusions): This section is poorly developed. It has to be written from scratch, to highlight the advantages and weaknesses of the methodology and also discuss ideas for future research.
We will rewrite it, and also add conclusions from the additional example.

Minor editorial comments
Page 2, line 10: In which of the aforementioned papers do you describe the algorithm? It is not clear here.
These papers refer to research work using caRamel but the algorithm itself was not described.

Page 2, line 17: Term "Hydrology" should not start with capital.
We will correct it.

Page 2, line 18: Please, change "underconstrained" to read "unconstrained".
We will change it.

Page 3, line 3: Please, change to read "with respect to".
We will change it.

Page 8, line 16: Please, change to read "parameter sets".
We will change it.

Page 9, line 7: Please, change to read "Pareto fronts".
We will change it.

Page 9, lines 16-17: Please, change to read "The GS metric exhibits a larger variability, thus a larger envelope for both optimizers".
We will change it.

References

Coron, L., Thirel, G., Delaigue, O., Perrin, C. and Andréassian, V. The Suite of Lumped GR Hydrological Models in an R package. Environmental Modelling and Software, 94, 166-171. DOI: 10.1016/j.envsoft.2017.05.002, 2017.

Coron, L., Delaigue, O., Thirel, G., Perrin, C. and Michel, C. airGR: Suite of GR Hydrological Models for Precipitation-Runoff Modelling. R package version 1.3.2.23. https://CRAN.R-project.org/package=airGR, 2019.

Rouhier, L., Le Lay, M., Garavaglia, F., Le Moine, N., Hendrickx, F., Monteil, C., and Ribstein, P.: Impact of mesoscale spatial variability of climatic inputs and parameters on the hydrological response. Journal of Hydrology 553, 13-25. http://dx.doi.org/10.1016/j.jhydrol.2017.07.037, 2017.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-259, 2019.

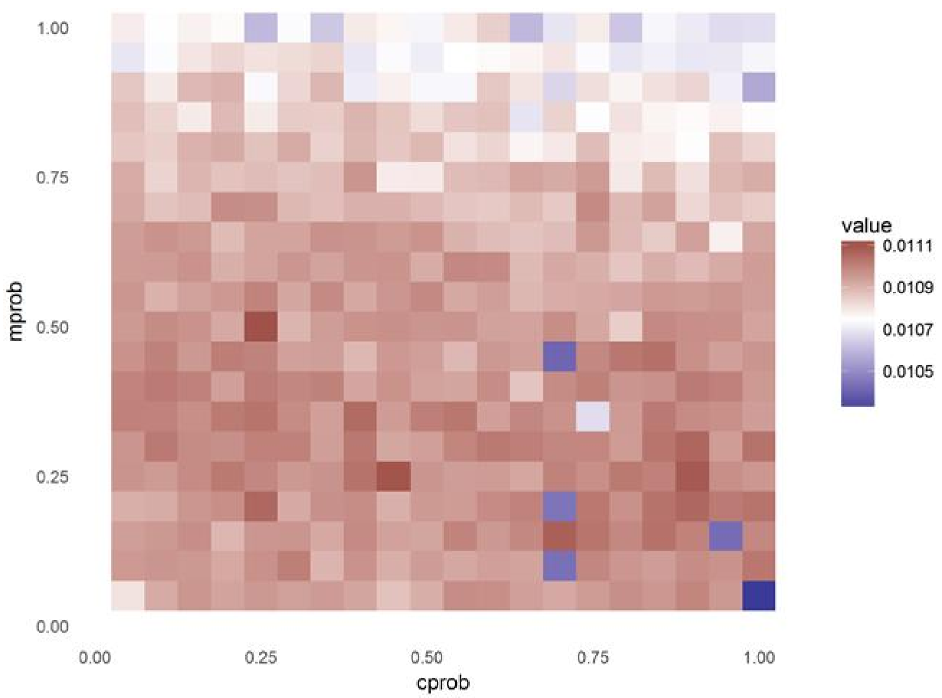

Fig. 1. Heat map of Hypervolume for mco regarding cprob and mprob (calibration of Mordor-SD, catchment Tarn at Millau).