Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-331-RC1, 2019

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "On the potential of variational calibration for a fully distributed hydrological model: application on a Mediterranean catchment" *by* Maxime Jay-Allemand et al.

Anonymous Referee #1

Received and published: 28 November 2019

I read with attention the paper proposed by Jay-Allemand et al. The issue of calibrating distributed hydrological models is of high interest for the readers of HESS. The authors apply a variational algorithm to estimate the parameters of a conceptual model distributed over a 1-km grid. Despite this very interesting and important topic, the authors failed in my opinion to produce an article publishable in HESS. Namely, I will point out in my review issues regarding the introduction and scientific objective of this work, the lack of in-depth analysis and discussion of the results, and the ill-posed mathematical problem that is studied.

Printer-friendly version

Discussion paper

Major remarks:

The paper lacks clear scientific questions. Three points are mentioned by the authors at the end of the introduction:

- upgrading the GRD model,

- calibrating GRD with a variational approach, and finally,

- upgrading the variational approach.

These three objectives are more similar to a model development process that is described in a project report than to science objectives tackled to fill a gap in knowledge in a research article. They are very specific to the chosen model and very general in terms of objectives (upgrading and calibrating are broad terms). They do not introduce valuable objectives for a scientific paper. Moreover, the proposed scientific objectives are not supported by a proper experiment protocol: the new GRD model is not compared to any benchmark (i.e. the former GRD model or another model), the calibration with the variational approach is not compared to a calibration with another approach, only with overly simple homogeneous parameter sets, and the improved variational algorithm is not compared with the classical variational approach. As a consequence, we don't know if the developments of this work are valuable for other research works, we just know their performance, with no landmark.

This lack of clear scientific questions comes with (in my opinion) a deficient introduction. Actually, an introduction aims at clarifying the work presented, through an introduction of the problem and proposed solution, a relevant literature review about what has been done and what is the gap of knowledge the authors want to address. These aspects are barely present in the introduction proposed by the authors. They start by presenting pros and cons of distributed models (raising the issue of equifinality, I acknowledge it, but not saying they are going to work on that somehow in this study), then they introduce DA methods used in hydrology, later on they introduce the concept of variational approaches and some DA methods applied to hydrology. Then we have a small paragraph about the issue of calibration for distributed models, with only two references and none over the last 10 years. Finally the issue for flash floods forecasting and the goal of the paper are mentioned. There is no clear continuity in this sequence of topics not really connected together or to this study. For example, do we need 15 lines about DA methods used in hydrology for state updating while this study is not about that? I would recommend presenting a deep analysis of the literature regarding the calibration of distributed or semi-distributed hydrological models. If this is a challenge, then explain why, explain what has been tried before and to which extent what you propose in this article is new. As mentioned in the introduction, other distributed models exist. As a consequence, they are calibrated, some of them with sophisticated methods. Recent examples include Rakovec et al. (2016), Piniewski et al. (2017), among others. However, as they are not mentioned, we do not know how the work compares to these previous studies.

The abstract is written in a very unusual way. Usually, an abstract should contain some context, a description of the methodology, the most important results and finally one or two sentences about perspectives, in terms of further research or improvements. Here, the context is given, and then the rest of the abstract is about the methodology implemented. Only the last sentence provides some elements of results ("encouraging results") with no much detail, and perspectives are never provided. In my opinion, the abstract should be entirely rewritten.

Variational methods are powerful tools for data assimilation and parameter estimation, which are used for quite some time already in the fields of meteorology or oceanography. Their use in the field of hydrology is clearly less developed, especially compared to the EnKF or particle filter, but I highly doubt that "In hydrology, the variational estimation method as described above (i.e. including the adjoint model) has not been reported so far." (page 3, line 12 ; see also the conclusions "To the best of our knowledge, this is the first time when the variational estimation involving the adjoint sensitivities has been applied in the field of hydrology."). For instance, in this journal, HESS, Castaings et al. (2009) seem to have developed a similar method. In addition, in their paper Castaings et al. cite some other works (see "Early applications of the adjoint state method to hydrological systems have been carried out in groundwater hydrology (Chavent, 1974; Carrera and Neuman, 1986; Sun and Yeh, 1990)" and the following sentences). Nguyen et al. (2016) might also be relevant. I encourage the authors to make a proper literature review on this specific aspect, which is necessary for the HESS audience to identify the added value of this specific study. In addition, the use of variational methods for data assimilation (i.e. state updating) is quite common in 'hydrology' in a broad sense.

The assessment of the performance is not developed. Only NSE values are calculated, while the authors specifically want to address flash floods. No criterion about peak-over-threshold, timing, intensity, is used. Since the study is already limited to a single watershed, limiting the analysis to a single criterion is out of the standards of nowadays hydrological studies.

The presentation of scores is poor. In figures 3, 4 and 5, the stations are ranked by their performance. It is a pity that we cannot identify anymore the stations. What a hydrologist would like would be to analyze whether there is a difference between experiments for a specific station, whether there are links between performances over upstream / downstream stations, etc. This kind of analyses is impossible to perform from the presented graphs. In addition, the fact that scores are mixed between the two periods (P1 and P2), if I understand well, is even more confusing.

As I mentioned at the beginning of my review, the analysis and discussion of the results is insufficient. The description of the results consists in a one-page long text and the discussion in less than 20 lines. Only NSE is presented to assess the performance, as well as the maps of parameters. This is a pity, as there would be a lot to say. First, it is clear from figures 6 and 7 that the parameter values are highly different when calibrated over P1 or P2. For many grid points, Cp and Ct can reach the lower bound for a period and the upper bound for the other. The authors blame the change of precipitation between the periods or the chosen model. In my opinion, it simply indicates equifinality. Indeed, the performances do not decrease a lot between the periods but the parameter values are completely opposite. It comes from the fact that not enough information is provided to the algorithm to calibrate 540 x 3 parameter values. In other words, the optimization problem is ill posed. It indicates that calibrating these three parameters over each grid cell is not possible with only discharge time series and the variational algorithm. While the presentation of negative results is interesting and should be encouraged in my opinion, it has to come with a proper experiment protocol and in-depth analyses. If different precipitation patterns are the key factor for explaining these results, then the reader has no element to assess that: no mean or extreme precipitation values or even maps are provided.

The discussion of the results comes with several rude and coarse judgments, not supported by evidence. For instance, the authors blame "a structural deficiency of the chosen model" (page 16, line 1), saying it is not surprising since "the model is conceptual". Then, they state that "the hydrological modeling at the cell scale is very primitive". These two elements may explain why the parameter values are so different according to the authors. First, these assertions are very surprising, as the authors did develop the model they used, if I understand well. Second, being simple or conceptual is not necessarily a deficiency. On the opposite, it is often considered as being an advantage, as such models are easier to run or to understand. Third, if the authors identified a structural deficiency, then it has to be shown and analyzed, and solutions for improvement must be discussed. It is true that some processes are not modeled, but what shows that this is the reason of the poor results? In addition, the authors also suspect the "routing scheme" (line 3). I am surprised by that, as figure 7 and page 14, lines 29 to 30 indicate that the routing velocity is the best determined parameter.

Minor remarks:

The quality of the English used in this work is sometimes rather poor. The manuscript contains a high number of formulation that are more typical from oral English than from written scientific English (“Let us note”, “As we already said”, etc.). I strongly recommend that a native English speaker reads and corrects the manuscript.

Page 1, lines 4 to 6: it is not clear whether AIGA is a forecasting system (as said on line 4) or not (as we learn on line 6 that it uses radar observations, not forecasts).

Page 1, line 8: “greater”: do you mean “higher”?

Page 1, line 10: “have also” must be changed into “also have”.

Page 1, line 10: “This must be larger enough”: what do you mean? Do you mean “large enough”?

Page 3, lines 23 to 29: local methods indeed sometimes fail to identify global optima. However, between local methods and DA approaches, one can find the global optimization algorithms, which prove to be sometimes more efficient than local methods. DA is an option, not necessarily the only option!

Page 6; line4: I would say that the soil reaches its “minimal” absorption capacity when all rainfall contributes to runoff, rather than “maximal”.

What is the meaning of the word “scalable” used in the introduction and in section 2.2?

Page 9, lines 14 to 19: this paragraph and equation is introduced, to finish by saying that it is not going to be used. This can be deleted not to confuse readers.

Page 10, line 2: this equation is not numbered.

Page 11, line 31: what is an “active” cell? Why mentioning the rectangular 1600 km^2 grid? GRD is not capable of running over catchments irregular shapes?

Page 12, line 1: this is a classical split sample test as defined by Klemes 30 years ago, it is worth mentioning it.

Page 12, line 8 to 9: it seems to me it is a proxy-basin test, as also defined by Klemes.

Printer-friendly version

Discussion paper

Please confirm.

Page 14, line 14: "One can see that the model spatial predictive performance is also better if the distributed calibration (red) is used, with one exception.": is that true? Since the stations are ranked, it might not be true, the reader cannot know.

Table 2 and 3 show a wrong unit for Ct. Indeed, if we check equation 8, then Ct must have the same unit as h and q, i.e. mm.

Page 15, the authors state that "For a chosen observation period and the associated test signal (rainfall) one can get a relatively stable set of calibrated parameters.". This statement is not supported by any kind of evidence in the manuscript and is not very clear. I guess that stability stands for temporal stability, but then how can it be assessed from a single period?

Page 17, line 23: is the code available under a GPL license? The proposed website requires a username but there is no possibility to register.

References:

Castaings, W., Dartus, D., Le Dimet, F.-X., and Saulnier, G.-M.: Sensitivity analysis and parameter estimation for distributed hydrological modeling: potential of variational methods, Hydrol. Earth Syst. Sci., 13, 503–517, https://doi.org/10.5194/hess-13-503-2009, 2009.

V. KLEMEŠ (1986) Operational testing of hydrological simulation models, Hydrological Sciences Journal, 31:1, 13-24, DOI: 10.1080/02626668609491024

Piniewski, M., Szcześniak, M., Kardel, I., Berezowski, T., Okruszko, T., Srinivasan, R., Schuler, D. V., & Kundzewicz, Z. W. (2017). Hydrological modelling of the Vistula and Odra river basins using SWAT. Hydrological Sciences Journal, 62(8), 1266–1289. https://doi.org/10.1080/02626667.2017.1321842

Rakovec, O., Kumar, R., Attinger, S., and Samaniego, L. (2016), Improving the realism of hydrologic model functioning through multivariate parameter estimation, Water Resour. Res., 52, 7779– 7792, doi:10.1002/2016WR019430.

Van Tri Nguyen, Didier Georges, Gildas Besançon, Adjoint-based state and distributed parameter estimation in a switched hyperbolic overland flow model, IFAC-PapersOnLine, Volume 49, Issue 18, 2016, Pages 205-210, ISSN 2405-8963, https://doi.org/10.1016/j.ifacol.2016.10.164.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-331, 2019.

Printer-friendly version

Discussion paper