Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-393-RC4, 2020

Hydrology and Earth System Sciences
Discussions

Interactive comment on "The role and value of distributed precipitation data for hydrological models" *by* Ralf Loritz et al.

Anna E. Sikorska-Senoner (Referee)

anna.senoner@geo.uzh.ch

Received and published: 16 September 2020

This paper proposed an adaptive modelling as an alternative to a distributed model for representing spatial variability of the catchment and forcing input (precipitation). Such an adaptive modelling should be able to run faster than a distributed model but should provide a similar model performance as its fully distributed version. The manuscript is generally well written and it is easy to follow. The idea of a spatially adaptive model that dynamically adjusts its spatial structure during runtime is indeed very interesting and has a potential for being applied in many (hydrologic) modelling approaches. Yet, I have few major issues that should be addressed before considering this manuscript for a publication in HESS. Thus, I recommend a major revision.

Printer-friendly version

Discussion paper

General comments

1. The adaptive model (model c) is tested here only on two rainfall events, which I see as the major weakness of this manuscript. As the strength of this approach should lie in the possibility to apply it to a continuous modelling and not to an event-based modelling. Thus, I think it would be important to demonstrate how the model c works on continuous time series. As this is missing in the current manuscript, we still do not know at the end whether it is a good or a bad option to be used.

2. The performance metrics of the calibrated (tuned) models should be provided so that the model ability to predict rainfall events could be assessed.

3. A model set-up between the model a and b could be very didactical, i.e. having a structure as the model b but using the precipitation input as the model a (the same for each grid cell).

4. It is not quite clear how the switch between different model setups (i.e. the number of model run in the model c) affect the setup of initial conditions for next runs, which is important to be considered for continuous model simulations but also for simulations of events. More details should be provided on that.

5. It would be also very didactical to see the comparison of the precipitation records from the ground station with the precipitation fields obtained from the gridded data. This is never done in the manuscript and no reason for not doing that is given.

6. The (rather) poor model performance of all tested models' set-ups for two selected events requires some discussion. It appears that none of these model can really capture the dynamics of these two events even if using the distributed

model and the distributed rainfall information (with KGE$<$0.3). Hence, it is even more important to verify the model performance (pkt. 2). An addition of other metrics that focus entirely on the flood event such as peak or time to peak could be here very informative. Given a rather poor models' performance, it is difficult to justify the need of developing the adaptive model based on the distributed model if the latter does not provide acceptable simulation results.

7. A fair comparison of all presented models should involve the same metrics, i.e. computation over the same time at a continuous time scale. In this study, different model setups are compared at different scales that makes it difficult to get an overview of their performance.

Detailed comments

1. Abstract: 'a mesoscale catchment'; a 20-km^2 catchment appears rather small to me than meso-scale.
2. Abstract: 'three hydrological models', the model is actually the same but different set-ups are used that span from the averaged model until the distributed model. Please clarify that here.
3. L. 20-21 p. 3: It is not always possible and justified to switch from a continuous model to an event based model. Hence, continuous modelling is often required in many applications.
4. L. 19 p. 4: The tested catchment appears rather small to me. How do you define the cut here for a small/meso-scale catchment?
5. L. 7-9 p. 5: consider restructuring this sentence.

6. L. 11-13 p. 7: could you add the location of these meteorological stations to the map in fig 1?

7. L. 15 p. 7: it should be 'and measures...'

8. L. 16 p.7: is there any weighting applied here and what kind of?

9. L. 28 p. 7: could the locations of radars be also placed in the fig. 1?

10. L. 12-14 p. 8: Why do you compare these values with literature and not with the ground station records from your catchment? Is there any reason that you are not using the precipitation records from the ground station?

11. Fig. 2: One would expect that the radar values would be compared with the ground station values as a corresponding mean (Fig. 2b). Could you add these values to the figure?

12. L. 10-17 p. 9 till 21 p. 10: I am not quite sure if this text is really helpful. After reading these lines, we still do not know how the reference model and other models look like. Maybe you could merge these lines with the sections 3.1-3.3.

13. L. 19-20 p. 10: I would say that the main goal is to test or verify whether similar model performance can be achieved with the adaptive model as compared to the model b. However, by the comparison that you did we still do not know the answer to this question as the comparison is done only based on two pre-selected events both having rather a poor model performance. Please comment on that and also state why these events were chosen for the comparison (and not others)?

14. L. 21-22 p. 10: Why do you compare the adaptive model with model b using only these two events and not the entire simulation period? In my opinion, the greatest potential of the adaptive modelling lies in continuous modelling and not in the event-based.

15. L. 31 p. 10 – l. 1 p. 11: why do you test the model only based on the annual assessment and not on hourly simulations? It is quite surprising because you use the model for assessing the model performance at an event-based scale in the second step, i.e. when comparing different models. I think it is important to report here how the model behaves at an hourly time scale so that one knows what can be expected from the model.

16. L. 3 p. 11: Which metrics were used here for assessing that the model performance agreed well with the dynamics of observed values? can you give some more details on that?

17. L. 7-8 p. 11: It is not surprising that the model performs poorly at time series scale if it was evaluated only on an annual basis. Some insights should be given here; why was the model tested at an annual basis if its intention is to predict events?

18. L. 3 p. 12: similar to what?

19. L. 12. p. 12: the model analysis should go after the introduction of all models.

20. L. 16-19 p. 12: an additional model between the models a and b would be here very useful, i.e. a model that has a structure as the model b but uses precipitation as in the model a (so it uses the same precipitation for each grid cell). This inclusion could nicely show the added value (or no value) of including a spatial distribution of i) the model and ii) of the precipitation input data.

21. L. 8-10 p. 13: why exactly? In my opinion, the strength of this approach lies in the possibility to apply it to a continuous modelling and not to an event-based modelling. Thus, I think it would be nice to demonstrate how the model works on continuous time series in terms of the model performance and computational efforts. As such a test is missing in the current manuscript, we still do not know

at the end whether it is a good or a bad option to use the adaptive modelling approach... Based on the two events selected, we cannot say much about the value of the adaptive approach as the model performance remains poor for these events (as seen from the Table 1 and fig. 7). If the full simulation is not possible (could you give more details why exactly?), already a simple test with shorter but continuous time series of few months or weeks could provide some more insights on how this approach is really working.

22. Tab. 1: as the initial idea is to improve the model performance for rainfall events, it appears from the table that the model c and model b have still rather poor model performance for the event I and II. In addition, all models perform poor for these events. Yet, an inclusion of the spatial variability does not improve much the model performance that is still not so good. Thus, it calls a question of the need of such an adaptive inclusion to this spatially distributed model which performance is rather low.... Could you comment on that? A decomposition of KGE into its components would bring more insights on the models' behaviour.

23. L. 9-10 p. 15: How many grid cells need to have a difference higher than this threshold to use the model c?

24. L. 30-31 p. 15 fig.3: is the re-arranged model running with the same initial conditions of the original model or how do you decide on these initial conditions if you want to increase or decrease the number of M in the subsequent time intervals particularly if a continuous simulation is performed?

25. L. 4-6 p. 18: for a fair comparison of different models, you should use the same metrics and the same time periods for evaluation. It is not clear why this is not the case here.

26. Fig. 4: could you add simulations with the model a?

27. L 8. P. 20: The reference Knoben et al. (2019) is missing in the literature list.

28. L. 6-7 p. 21: the performance of KGE below 0 is still rather very poor, which requires some further explanations. According to Knoben et al. (2019), simulations can be considered as behavioural if KGE$>$0.3 (with KGE ≈ -0.41 for a mean flow benchmark).

29. Table 1: the performances (KGE) of all models is rather poor for the events here selected (with KGE between -0.41 and 0.29). As already the model b (distributed) cannot simulate the two events in a good way (as also seen from the fig.7), why would you spend time on developing the adaptive model based on the model b instead if improving the model b or testing different models here? Could you comment or justify that?

30. Fig. 7: the simulations with the model a and the reference model should also be added here. Moreover, for both events, all models largely underestimate the events. Could you comment on that?

31. L 10 p. 26 – l. 2 p. 27: do you have any idea where this large underestimation may come from and how it could be improved?

32. Discussion: I missed some recommendations for other works. When and how would such an adaptive modelling be recommended? How one can set up the adaptive process? And why it is really needed to implement such an adaptive modelling?

Reference

Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores, Hydrol. Earth Syst. Sci., 23, 4323–4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-393, 2020.