Hydrol. Earth Syst. Sci. Discuss., 8, C3231–C3236, 2011
www.hydrol-earth-syst-sci-discuss.net/8/C3231/2011/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Evaluation of the transferability of hydrological model parameters for simulations under changed climatic conditions" *by* S. Bastola et al.

G. Heuvelmans (Referee)

g.heuvelmans@vmm.be

Received and published: 4 August 2011

Bastola et al. discuss the validity of the 'standard way' of using conceptual rainfall-runoff models for climate change impact assessment: calibrate the model parameters using historical stream flow records and use these calibrated parameter values when modelling future climate scenarios. Doing so, one implicitly assumes that models calibrated for current weather conditions deliver reliable predictions for the future climate scenario. Is this assumption valid? Are model parameters really independent of the climatic situation? That is the main question of this paper, a question that surely is

significant on both a practical and scientific level. The question is of practical relevance, because it needs to be answered before rainfall-runoff models can be applied in a reliable way for predicting the impact of climate change. Its scientific relevance could consist of generating insight in (1) how model parameters relate to hydrological processes, and (2) how these processes and parameters are affected by climatic variables.

The practical and scientific question that I just outlined, cannot be separated: if one wants to formulate a substantiated answer to the practical question, one needs to have insight in the theoretical one. Unfortunately, the paper deals almost exclusively with the practical issue, questioning how parameter transfers affect model performance, but tends to ignore the question why this effect occurs. This is regrettable, because it limits the relevance of the findings of this study for other cases.

Before I can elaborate on this problem, I should first give a quick overview of the modelling experiment conducted by the authors. This experiment roughly consists of three steps: 1. split the available data records into several subdatasets with varying climatic conditions; 2. establish a set of behavioural parameter sets – using GLUE – for every subdataset and for the dataset as a whole. This results in several groups of behavioural parameter sets, called 'wet', 'dry', 'average' and 'conventional' sets. 3. Compare the model performances obtained with the wet, dry, average and conventional parameter sets for wet and dry periods, and for the entire record of streamflow observations.

The results of step 3 are then used to assess the transferability of model parameters to future climatic conditions. To my opinion, the criterion for splitting up the data (step 1 of the 3-step overview) in 'wet' 'average' and 'dry' conditions is the crucial point of the method. The authors used the 'annual average rainfall' and compared this rainfall volume with two thresholds, above and underneath which a year is considered wet / dry. First of all, it should be explained (1) what these threshold values are; and (2) the reason(s) why the thresholds were set at these values . This is an essential element of the method, but I can't find any information about this issue in the paper. Furthermore,

HESSD

8, C3231–C3236, 2011

I have two questions concerning the appropriateness of the splitting criterion used by the authors.

First: How does the splitting criterion relate to runoff generating processes and model parameters?

Which processes are the model parameters supposed to represent? How are these processes related to climate? What are the differences between the behavioural parameter sets derived for wet and dry conditions? Can these differences be understood as differences in runoff generating mechanisms? And can you, based on the answers to all these questions, formulate some hypotheses about the climatic conditions to which parameters can be transferred without running into trouble? Ideally, the splitting criterion used in the modelling experiment should take into account the answer to this last point. But it is still useful to do the analysis now, after all runs have been done, to better understand why parameter transfers were found to be (un) problematic. This theoretical insight is necessary to apply the findings of the study to other cases, especially to climate change impact studies in other catchments.

Second: Is the splitting criterion representative for the expected future climate change?

If this would not be the case, the transfer of parameters between wet and dry subdatasets would say nothing about the transfer of parameters from present to future climatic conditions. However, the paper does not provide any info on the way that climate is changing according to the scenario under study. I'm not an expert in this field, but I thought that some climate change scenarios do not so much predict a change of the yearly total rainfall, but rather expect a more frequent occurrence of extreme events, intense storms and prolonged dry periods. If this is the case, then the splitting criterion used by the authors might be inappropriate because it does not reflect future climate changes. The splitting should then be based on the occurrence of extreme events, not on the average annual rainfall. Or it should at least be explained what the frequency of extreme events is in the wet and dry subsets, how these frequencies compare with the

HESSD

8, C3231–C3236, 2011

occurrence of extreme events in the climate change scenario.

To address these questions, more background information should be provided on the models and model parameters, and on the climate scenario. The expected impact of climate change on runoff generating processes, and the representation of these processes by the model parameters should be discussed. Based on this discussion, the authors should either explain why the splitting criterion that was used in their study is appropriate, re-interpret the results of the study or revise the splitting criterion and redo the experiment. The last option would of course entail a huge amount of work. The paper does not contain the necessary information (e.g. on the nature of the climate change scenario) to judge whether such a drastic revision is really required or whether a re-interpretation of the results will suffice.

Minor comments:

- p. 5893, l 26-28: it is stated that geographical transferability is more problematic than temporal transferability. This is true if the spatial variation of catchment and rainfall properties that are relevant for runoff generation, is larger than the temporal variation of these properties. But if you are simulating the impact of future land use or climate change scenarios, then the temporal variation might become larger than spatial variation, so that temporal parameter transfers might become more problematic than spatial transfers.

- p. 5894 line 10-11: '...only relatively few studies have looked into the temporal transferability of model parameters'. I don't think this is the case: in a 'standard' study, only split-sample tests are performed subdividing a dataset in a calibration period followed by a validation period. This is a kind of temporal transferability.

- p. 5898 line 5-9: when did you consider a parameter set as behavioural? Which NSE, volume error etc. did you use as threshold value to distinguish between behavioural and non-behavioural sets? Moreover, I do not understand what is meant with 'count efficiency' and with the last criterion.

HESSD

8, C3231–C3236, 2011

- p. 5898, line 16-19: here, only 1 criterion is mentioned, whereas on line 5-9, 4 criteria are presented. Why this difference?

- p. 5899, line 7-8: If I get it right, there is not always a loss in NSE; for HYMOD, for example, the calibration for the dry period seems to result in a higher NSE for the wet than for the dry period. . The same applies to the NAM model.

- p. 5899, line 16-19: can this difference in model structure explain the observed differences in model performance? Can this be explained in terms of runoff generating mechanisms?

- p. 5900 (bottom): figure 5 is referred to on line 24, but the lines following this reference seem to refer to figure 4. I can't find any discussion or conclusion about the results presented on figure 5.

- p. 5901 line 9-10: It is stated that the predictions are similar with and without parameter updates. From Table 2, it can also be concluded that using different parameter sets for different climatic conditions, does not really improve the model predictions. Sometimes, the performance for the time variant parameters is even worse than the performance obtained with the time invariant parameters. This might mean two things: 1. That parameter updates are not really needed, because the same parameter set delivers predictions with a similar degree of reliability in varying climatic conditions. 2. That the updating scheme is inadequate (for example, because you ignore the impact of extreme events). What is the most probable explanation in this study?

The manuscript is understandable but contains lots of mistakes with respect to language and sentence construction. Attached is a list of the most important linguistic errors. This list also contains some remarks to increase the understandability of the figures.

Please also note the supplement to this comment:
http://www.hydrol-earth-syst-sci-discuss.net/8/C3231/2011/hessd-8-C3231-2011-

HESSD

8, C3231–C3236, 2011

supplement.pdf

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 5891, 2011.