Review HESSD 9, 7441-7474, 2012: Velazquez et al., An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources

General comments

The manuscript addresses a very relevant issue, namely assessment of the uncertainty caused by hydrological model structures in climate change impacts. The author performs modeling using ensembles of climate models, bias correction/scaling methods and hydrological models for two catchments in Canada and Germany. The results are interesting and contain sufficient new results to justify publication in HESS. However, there are a number of places in the manuscript where the description is so short that it is not possible to assess the soundness of the applied methodologies. Furthermore, there are important weaknesses in the manuscript that need to be improved. Therefore, I recommend that the authors are requested to address the comments given below, before the manuscript is accepted for publication in HESS.

Specific comments

1. A serious weakness in the manuscript is the lack of emphasis on scientific novelty of the findings of the study. The authors must document the scientific novelties through discussions with reference to state-of-the-art.

2. The Introduction has very good references to state-of-the-art literature. However, it has some serious shortcomings that should be improved:
   - There is no identification of science gaps in state-of-the-art. In order for the present manuscript to document scientific novelty it should address science gap(s).
   - The objectives of the paper/study should be stated towards the end of the Introduction.
   - Lines 12-25, p 7445: This paragraph is a summary of approach and comprises duplication with text in Chapter 2. It does not belong in the Introduction and may be deleted.
   - The last paragraph, p 7445 line 26 - p 7446 line 2, is not required as the manuscript follow a standard format for a scientific paper.

3. Two different emission scenarios are used for the Canadian (A2) and the German case study (A1B). I am concerned that this may influence the results of the study. This is not even reflected upon in the manuscript.

4. Please provide just a little more information about the basis for the precipitation and climate input in the two catchments – e.g. the number of stations and whether daily or hourly values.

5. It appears that three methods are used for bias-correction/downscaling: (i) monthly correction factors for temperature; (ii) LOCI for precipitation; and (iii) SCALMET for the remaining meteorological

variables. The two last methods are not explained and the reader has to study literature. It would be useful if the authors provide just a brief description with key characteristics of the methods, so that an otherwise informed reader would be able to assess them.

6. I am concerned that three different calibration techniques are used for the models including automatic SCE for two models, manual trial-and-error for one model and no calibration for the third model. This is far from ideal and may affect the results significantly.

7. The objective functions for HSAMI and HYDROTEL are the sum of squared errors and the root mean square error. Is that not essentially two functions which give the same result, with one being the square root multiplied by a factor of the other?

8. The models have been subject to split-sample tests. This is fine, but information on the results are lacking.

9. P 7452 lines 8-14: The test procedure is not well explained and hence not transparent. It is not clear to me how the Wilcoxon rank sum test was done. Which variables were tested? Is it the series of annual values of runoff with the hypothesis that the median values are identical? Or is it the series of five or three values of average runoff originating from the five/three GCMs?

10. Table 2. The information in this table is not transparent. I recommend that the authors provide some basic statistics (e.g. sample size, mean, standard deviation for the series being compared and p-values to evaluate the level of significance instead of just crosses.

11. Figs 7-8: I wonder if there is any need to include the absolute errors. I do not think they provide interesting results here, and they complicate the figures.

12. P 7458, lines19-20. The authors recommend using a hydrological model ensemble to fully assess the uncertainty in the climate change signal. This is unclearly phrased, because the hydrological models cannot assess the uncertainty in the climate change signal, but rather the uncertainty on hydrology due to climate change.

13. P 7458, lines22-23: For the indicators where null hypotheses have not been rejected for any pairs of models the authors recommend that a single conceptual model can be used with some certainty. I think that use of single hydrological models instead of an ensemble of models will, also in these cases, result in underestimation of the uncertainty.

14. The authors recommend (p7460 lines 4-10) that uncertainty in projections added by hydrological models should be included in climate change impact studies. This may well be correct, but it is nog documented in the manuscript. The present study only includes natural (GCM initial conditions) climate uncertainty, but not the uncertainty on GCMs nor the uncertainty on downscaling (RCMs + statistical downscaling/bias correction). To draw such conclusion you need to compare the uncertainty generated by hydrological models to the other sources of uncertainty.

15. Chapter 4 conclusions: This chapter should emphasize the novel findings of the study compared to state-of-the-art. This is done to a too limited extend, and should be improved substantially.

16. Table 2, caption: It is not the test that is rejected but the hypothesis.

17. Fig. 1: The quality of this figure could be improved. The site location map showing Quebec and Bavaria could be reduced (maybe to inserts) and more pace given to the catchment maps. Then it would be interesting to see the location of the precipitation stations on the catchment maps.

18. Fig 1 caption uses the term watershed, while the term catchment is used in the text otherwise. Please be consistent.