The authors would like to thank Yi He for her thorough review of our manuscript; we think it really helped to improve the text. In the following we reply to her comments in detail.

Köplin et al.: Relating climate change signals and physiographic catchment properties to clustered hydrological response types, HESSD, 2012.

Y. He (Referee), yi.he@uea.ac.uk

General comments

1. It is an interesting application of catchment classification and regionalisation in impact studies. It aims to provide estimation of impacts on clusters of catchments that exhibit similarities rather than each individual catchment. The advantages are clear: it can reduce work load in a typical model cascade set up in impact studies; it can help identify catchments that are more sensitive to environmental changes and hence should be give more attention to; regionalisation can handle ungauged catchments.

2. Introduction section should include references, and provide background information and limitations about using cluster analysis in catchment classification and regionalisation, because the Ward's minimum variance clustering method is the main method applied in this study. Cluster analysis is a purely data mining exercise. The key issue is that it is hard to draw conclusion on the relationship between behaviour/response similarity and physiographic-climatic similarity based on clustering methods. Ultimately, people would be interested in similarities in functional responses of catchments. See discussions in e.g. Wagener et al (2007) Catchment classification and hydrologic similarity; He et al (2011) A review of regionalisation for continuous stream flow simulation. How would such methods really help, if the clusters identified do not necessarily mean a functional similarity in catchments' hydrological response?

Response: We clustered the catchments based on their mean monthly temperature-, precipitation- *and* runoff-change signals (P3174 L5). That is, we clustered the response (the runoff) together with the causing climate change signals. Thereby, the resulting clusters constitute groups of catchments exhibiting functional similarity, per definition. We clustered the catchments like this because the same climate change signals might provoke different runoff change signals (P3174 L6), which appeared to be true (P3178 L14f.). Afterwards, we studied the catchments of a cluster for their physiographic similarities in a descriptive manner (Fig. 4 is the corresponding visualization) to test if there are catchment properties that clearly determine the catchments' functional response to climate change. It is true that we only applied the Ward's minimum variance clustering and did not test other cluster methods. We did that because several other studies, which assessed different clustering methods for classification of hydrological data, found that the Ward's method yields the physically most plausible groups of catchments (e.g. Bower et al., 2004; Gobena and Gan, 2006; Kingston et al., 2011). Moreover, it is well known that there is no best clustering method (e.g. Hannah et al., 2005). The referee is right, of course, that we did not cover possible alternatives to the chosen method and we integrated the following text passage in the methods section (at the end of Sect. 3.1):

"*It has to be stated that there are various clustering methods (see e.g. Borcard et al., 2011 for an overview) and none of them can be objectively rated as the best method (Hannah et al., 2005). A number of studies that tested different clustering methods for their use in catchment classification found that Ward's minimum variance method yields robust and physically meaningful results (see e.g.*

Bower et al., 2004; Gobena and Gan, 2006; Kingston et al., 2011). Therefore, we chose this method for our study."

Moreover, we already mentioned in Sect. 5.4 ("Further research") that a possible extension of our study would be indeed to apply different clustering methods and compare their results (P3187 L9–11). In addition, the redundancy analysis (RDA), which was performed both on the entire set of catchments (no clustering) and on each cluster alone, showed for the entire set of catchments a similar grouping as produced by cluster analysis. This we rate as a confirmation that the chosen clustering method yields robust results for our study set up (P3179 L16f.).

3. Page 3167 L14-25, such climate impact studies use the so-called top-down approach. But this is not the only way to study impacts and also not the best way. Drawbacks associate with the top-down approach has been discussed by other people already. There have been an increasing number of papers that adopt the so-called bottom-up approach, see for example: Prudhomme et al (2010) Scenario-neutral approach to climate change impact studies: Application to flood risk; Wetterhall et al (2011) Using ensemble climate projections to assess probabilistic hydrological change in the Nordic region; van Pelt and Swart (2011) Climate Change Risk Management in Transnational River Basins: The Rhine. To put this study in perspective, it is necessary to cover the state-of-the-art of impact studies in the intro section.

Response: Granted, though one could argue that the top-down approach is state-of-the-art, too. At least, the top-down approach applied here is "[…] widely used and accepted" (van Pelt and Swart 2011). In fact, both approaches should be used complementary as van Pelt and Swart (2011) suggested. We added a paragraph to Sect. 5.4 "Further research" to cover possible alternatives to the applied top-down approach. We think it fits best there because applying a bottom-up approach, too, would actually be a possible next step. We did not include it in the intro section, where we already extensively covered the full spread of uncertainties related to impact studies using the applied top-down approach (P3167 L16–23). The additional paragraph in Sect. 5.4 reads as follows: "*A different approach to assess hydrological change would be to test a range of plausible changes in climate (derived from climate models) for their threshold exceedance of safety margins, for example (Prudhomme et al. 2010, van Pelt and Swart, 2011; Wetterhall et al., 2011). That is, this is a sensitivity analysis of a hydrological system to different climate changes rather than an impact analysis of climate change on hydrology. This approach has only recently gained attention for use in adaptation strategies (van Pelt and Swart, 2011) and offers an addition to the top-down approach applied in this study. However, it is not a substitute to climate impact studies, as it cannot account for the complete storyline of climate change an ensemble of GCM-RCMs depicts.*"

4. The current generation of climate models have shown very limited skill at predicting (in hindcast even) changes in climate statistics on regional and local scales, as discussed in many papers, see e.g. Stephens et al (2010) Dreary state of precipitation in global models; Anagnostopoulos et al (2010) A comparison of local and aggregated climate model outputs with observed data. The concern here is using these climate signals in impact studies does not prevent people from falling in the problems of uncertainties in climate model predictions. In other words, the clusters of catchments identified in this study could be subject to 'false alarm' if climate model outputs turned out to be wrong. Why

these climatic variables are not divorced from catchment physiographic properties in the cluster analysis and RDA? The future climate signals could actually bias the clustering results.

Response: In the cluster analysis, the climate change signals are used together with the hydrological response, and not together with the physiographic properties (see above, Comment 2), to study the functional similarity of the catchments with respect to climate change. In the RDA, climate change signals and physiographic properties are both used to study their constraining effects on the runoff change. The constraints in an RDA are, however, always independent from each other, which is one of the advantages of redundancy analysis. We are aware of the fact that the climate change signal is always biased. It is therefore imperative to incorporate a bias-correction in impact studies. In the present study, the bias-correction was done implicitly in the downscaling step (P3172 L15f): The delta change method derives the climate change signals within the modelled time series of precipitation and temperature. Because not the direct climate model output but only the inherent change signals are used, a bias correction is not necessary, or in other words, it is implicitly incorporated. The delta change method is a conservative and robust method and its limitations are well known: Because the observed frequency and variability, i.e. the observed time series, are scaled with the climate change signals, one cannot assess changes in frequency or variability of the output variables. One can, however, assess changes in the mean and the annual cycle of hydrological variables which we did in this study and which are important information for water resources managers, for example. An advantage of the delta method: By scaling the observed time series with the derived deltas, one can avoid the problematic wet-day-frequency of direct climate model output, which Stephens et al. (2010) mentioned in their study. Only the measured and physically plausible frequency is used to run the hydrological model. By using an ensemble of 10 RCMs and by evaluating their ensemble mean, that part of the climate model uncertainty that cannot be bias-corrected is partly accounted for. Moreover, applying an ensemble of climate scenarios to represent the climate model uncertainty is state of the art in impact studies. We covered that in the first paragraph of Sect. 5.3 “Sources of uncertainty”. One result that makes us confident the clustering and the RDA are not severely biased through the climate change signals is the fact that temperature is by far the most important driver for hydrological change. As Anagnostopoulos et al. (2010) showed, the temperature signal is much more certain than the precipitation change signal is.

5. Can authors explain why hourly simulations were performed using the PREVAH model but the cluster analysis only used monthly average values?

Response: The hydrological model internally runs on hourly time steps. The calibration and regionalisation was done with an hourly resolution of the data to correctly represent daily cycles of temperature, radiation and, hence, melt processes, for example. In a following study, we will examine high flow conditions, which is another reason why we ran the model in an hourly resolution. For the analysis in the present study, however, the focus is more on general changes and not on peak flow. Therefore, we aggregated the data to monthly values. We missed to point this out in the manuscript, specifically, and we added this important information to Sect. 2.2.3 which reads as (P3173 L10ff.): *We aggregated the simulated hourly time series to mean annual cycles of monthly values because the focus is on changes in the hydrological regime, here, and not on peak flows, for example.*

6. P3174, it is not clear what exactly ‘hydro-climatological’ variables are used in the cluster analysis. Table 1 lists physiographic properties, with the exception of ‘gl_ctrl_rel’, ‘gl_near_rel’, and

'gl_far_rel', which are climate related. Table 2 lists some climatic variables, but they are used in the RDA. A bit lost here...

Response: The hydro-climatological change signals are defined on the preceding page P3173 L20–23; we do not refer to Tables 1 or 2 in the context of the cluster analysis. The hydro-climatological change signals consist of the mean annual cycles of monthly runoff change signals ('hydro') and temperature and precipitation change signals ('climatological'). The glacier extents mentioned in the comment ('gl_ctrl_rel', 'gl_near_rel', 'gl_far_rel') are indeed catchment properties that change with climate. Those are used to describe the clusters with respect to their physiographic properties (see Fig. 4), and they are used in the RDA as the referee mentioned correctly.

7. I struggle to understand at times what the authors really wanted to say, due to for example long sentences or some unfamiliar terms used. It would help if a native English speaker can proofread and improve the text from the language point of view.

Response: We worked through the manuscript and shortened or split up some long sentences. As none of the authors is a native speaker, we would like to leave it to the editor to decide whether the manuscript should be proofread or not.

8. In general, the manuscript is well structured, methods are well explained, and results are properly analysed. It is a worthwhile study that contributes to climate impact studies. The results have particular values in other catchments in alpine regions, and the approach can be generally applied elsewhere. I recommend the manuscript be accepted for publication after some moderate revision.

Specific comments

P3168 L14: 'characteristic catchment properties' → Delete 'characteristic'

Response: We changed that.

P3168 L20: 'modifying catchment characteristics' → Change to modified catchment properties

Response: We really mean 'modifying' as an adjective to 'catchment characteristics'. It refers to the citation in the previous sentence, and we added 'such' to highlight this. The sentences now read: *Sawicz et al. (2011) found, for example, "[. . .] that soil properties will modify the impacts of climate change on hydrologic regimes, which means that changes in precipitation and temperature will not impact the streamflow response equally". If there are* such *modifying catchment characteristics like soil properties or other physiographic signatures in our study region, then …*

P3168 L24: 'classify the catchments' → Delete 'the'

Response: We changed that.

P3169 L5: 'facilitate to designate prioritized regions which adaptation measures should be applied to, first.' → Change to 'facilitate identification of important regions where adaptation measures should be applied to with priority.'

Response: We changed that to: "*[…] will facilitate identification of regions where adaptation measures should be applied to with priority*."

P3169 L8: 'Switzerland provides a variety' → Change to: Switzerland has a variety

Response: We changed that.

P3169 L9: 'over the flat and hilly Swiss Plateau' → Quite confusing here, flat or hilly?

Response: Actually, the Swiss Plateau is both. But to avoid confusion and compared to the Jura mountain ranges and the Alps, the Swiss Plateau is flat. We changed the sentence to: "*[…] rather flat Swiss Plateau […]*".

P3169 L15: 'water balance basins' → What are 'water balance basins'?

Response: We decided to delete that part of the sentence and, thereby, the Swiss-specific technical term "water balance basins" to avoid any confusion. It is not relevant for understanding.

P3170 L1: 'Model forcing is mainly based on the 76 highly resolved meteorological stations and is complemented with the less highly resolved data.' → Change to 'Model forcing is mainly based on the 76 meteorological stations with hourly data and complemented by other stations with data at lower temporal resolutions.'

Response: We changed that.

P3170 L13: 'we arrived at a comprehensive set' → Change to 'we collated at a comprehensive set'

Response: We think "collated" is not the right term either. We changed the sentence to: "*we obtained a comprehensive set*"

P3171 L1: 'based on one distinct' → Change to 'based on a single set'

Response: We changed that.

P3171 L8: 'cross validation approach that proved good model' → Change to: 'cross validation approach that produced good model'

Response: It is not the cross validation that produces a result but the regionalized parameter sets. We replaced 'proved' with 'showed', here.

P3173 L1: 'were superimposed' what does this mean exactly?

Response: We replaced 'superimposed' with 'scaled'; the sentence now reads: "*The observed time series of precipitation and temperature were* scaled *with the annual cycles of climate change signals. Thereby, we generated a set of climate scenarios with which the catchment models were run.*"

P3174 L24: what do you mean by 'ratio scale'?

Response: We refer to the ratio scale (and the nominal scale, too) as levels of measurements. We added the information on the data's scale types because the different types are treated differently within the RDA (the nominally scaled variables are displayed as centroids, the variables of ratio scale are shown as vectors).

P3175 L3-5: 'Mean slope, mean available field capacity and mean soil depth are highly correlated among each other. Because available field capacity and soil depth are variables derived from the soil map that is spatially less highly resolved.' → Change to: 'Mean slope, mean field capacity and mean soil depth are highly correlated among each other. Because field capacity and soil depth are variables derived from the soil map that has a relatively low spatial resolution.'

Response: We changed that.

P3175 L6: 'are dispensed from' → Change to 'are excluded from'

Response: We changed that.

P3177 L24-25: 'The variable catchment area is not discriminating between the clusters; the clusters mean values vary around the overall mean.' → Change to: 'The variable catchment area does not differ much amongst the clusters. In other words, the clusters mean values vary around the overall mean.'

Response: We changed that, thank you.

P3179 L15: ‘because the catchments of a cluster’ → Change to: ‘because the same catchments of each cluster’

Response: We changed that to: “*[…] because the same catchments are grouped together in the RDA-biplot, too*”.

P3180 L8: ‘here: dominant aspect and dominant land use’ → Why ‘Centroids aspect’ and ‘Centroids land use’ are used in Fig. 6? Keep terms consistent.

Response: We deleted ‘dominant’ in the text. Thank you, this was misleading.

P3181 L7: ‘compared to the hydrological change signals of the far future because’ → Change to ‘compared to those of the far future because’

Response: We changed that.

P3185 L18: ‘modelling is the climate model itself’ → Change to ‘modelling comes from climate models’

Response: We changed that.

P3186 L19: ‘will be foci of another study’ → Change to ‘is the foci of another study’

Response: We changed it to “*[…] are foci of another study*”.

P3191 L15-19: can this article that is still in preparation be cited?

Response: The article is printed by now.

P3201-3202 Fig.6 and 7: fonts are too small to read.

Response: We increased the size of the fonts.

References:

Anagnostopoulos, G. G., Koutsoyiannis, D., Christofides, A., Efstratiadis, A. and Mamassis, N.: A comparison of local and aggregated climate model outputs with observed data, Hydrolog. Sci. J., 55, 1094–1110, doi: 10.1080/02626667.2010.513518, 2010.

Borcard, D., Gillet, F. and Legendre, P.: Numerical Ecology with R, Springer New York, New York, NY, ISBN: 978-1-4419-7975-9, 2011.

Bower, D., Hannah, D. M. and McGregor, G. R.: Techniques for assessing the climatic sensitivity of river flow regimes, Hydrol. Process., 18, 2515–2543, doi: 10.1002/hyp.1479, 2004.

Gobena, A. K. and Gan, T. Y.: Low-frequency variability in Southwestern Canadian stream flow: links with large-scale climate anomalies, Int. J. Climatol., 26, 1843–1869, doi: 10.1002/joc.1336, 2006.

Hannah, D. M., Kansakar, S. R., Gerrard, A. and Rees, G.: Flow regimes of Himalayan rivers of Nepal: nature and spatial patterns, J. Hydrol., 308, 18–32, doi: 10.1016/j.jhydrol.2004.10.018, 2005.

Kingston, D. G., Hannah, D. M., Lawler, D. M. and McGregor, G. R.: Regional classification, variability, and trends of northern North Atlantic river flow, Hydrol. Process., 25, 1021–1033, doi: 10.1002/hyp.7655, 2011.

Prudhomme, C., Wilby, R., Crooks, S., Kay, A. and Reynard, N.: Scenario-neutral approach to climate change impact studies: Application to flood risk, J. Hydrol., 390, 198–209, doi: 10.1016/j.jhydrol.2010.06.043, 2010.

Sawicz, K., Wagener, T., Sivapalan, M., Troch, P. A. and Carrillo, G.: Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA, Hydrol. Earth Syst. Sci., 15, 2895–2911, doi: 10.5194/hess-15-2895-2011, 2011.

Stephens, G. L., L'Ecuyer, T., Forbes, R., Gettlemen, A., Golaz, J.-C., Bodas-Salcedo, A., Suzuki, K., Gabriel, P. and Haynes, J.: Dreary state of precipitation in global models, J. Geophys. Res., 115, doi: 10.1029/2010JD014532, 2010.

van Pelt, S. C. and Swart, R. J.: Climate Change Risk Management in Transnational River Basins: The Rhine, Water Resour Manage, 25, 3837–3861, doi: 10.1007/s11269-011-9891-1, 2011.

Wetterhall, F., Graham, L. P., Andréasson, J., Rosberg, J. and Yang, W.: Using ensemble climate projections to assess probabilistic hydrological change in the Nordic region, Nat. Hazards Earth Syst. Sci., 11, 2295–2306, doi: 10.5194/nhess-11-2295-2011, 2011.