Response to editor decision

We thank the editor and the reviewers for their positive evaluation of our manuscript. We also thank for their final comments that we carefully addressed in the final version. Please find our responses below in italic text.

Kind regards,

Gregor Laaha

Editor Decision: Publish subject to technical corrections (26 Aug 2016) by Kerstin Stahl
Comments to the Author:
Dear authors,

The reviewers found their comments fully or mostly addressed and I am glad to be able to accept the manuscript for publication on that basis. Two of the reviewers made some final comments they suggest should be considered when finalizing the paper. I trust that these more technical/editorial comments can easily be addressed before uploading the final manuscript files.

Reviewer 3 notes that the requested information on the groundwater discharge modeling is still insufficient and couldn't be found in the given reference. Please add enough information here to make this important aspect for low flow modelling transparent to the readership.

Additional information has been added.

Reviewer 2 found the three-pillar framing improved but still notes a bias between the introduction and the way the actual results are then used&discussed. Perhaps some small adjustments could be done to improve this.

We acknowledge that the data structure of the three-pillar synthesis is somewhat different to classical ensembles in climatology, and suggest that learning from agreement/disagreement is crucial in addition ensemble mean and uncertainty. This concept is consistently followed throughout the paper. We are now relating the framing to the broader concept of consilience, which suggests that, if multiple sources of independent evidence are in agreement, the conclusion can be very strong even if the individual sources do not provide strong evidence on their own (Wilson, 1998).

Using the extrapolated trends I think R2 illustrates well the danger based own failure - perhaps the paper would benefit from some acknowledgement of other studies' experience of failing (using this or any other example of failure along the lines of the argument made by the reviewer – would mean to at least revise pg 13 lines 11,12).

We extended the paragraph by the following text: "Luce et al. (2013) pointed out that in their study initial interpretations of apparently consistent trends would have been misleading, partly due to artifacts in data, missing information and overextrapolation of trends, which triggered additional analyses leading to a differing perception of hydrological change. This example illustrates the importance of careful process reasoning in every step of the analysis."

I also agree with Reviewer 2 about the weakness of the analogy between (most) multi-model ensemble experiments and the three approaches here and would suggest to revisit the respective paragraphs in light of this comment.

We agree that many application in climatology apply the concept to deal with the epistemic uncertainty of multi-model ensembles that are similar in basic design and only differ in hypothesis of particular processes, but it is also stated in Knutti et al. (2010, p. 3, fourth paragraph) that in CMIP5 the multi-model ensembles are getting more heterogeneous, meaning that ensembles will include additional sources of uncertainty that require methods that go beyond ensemble expectation and uncertainty: "The reliability of projections might be improved if models are weighted according to some measure of skill and if their interdependencies are taken into account, or if only subsets of models are considered (Knutti et al., 2010, p.3, fifth paragraph). All these concepts perfectly fit to the model comparison put forward in this paper. To be clearer about the stated analogy, we acknowledge the difference between classical multi-model ensembles and the three approaches combined in this paper by adding the following text (p. 2, L. 47): "While the climate models Knutti et al. (2010) are referring to are similar in their basic design and only differ in specific process representations, the notion of inferring predictive reliability from model consistency builds on the broader principle of consilience, which suggests that, if multiple sources of independent evidence are in agreement, the conclusion can be very strong even if the individual sources do not provide strong evidence on their own (Wilson, 1998)."

A few minor issues from an editorial point of view that I stumbled upon on my last read are:
- To draw more attention, in the abstract perhaps the flow could be improved a bit and the important link found between low flow process and pillar agreement could be more clearly highlighted (essentially filling the statement "processes are discussed" with some more concrete results) .

We updated the abstract accordingly.

- Pg 7 line 10 which suggests that rivers become "ephemeral". This means they will flow only during rare events. Can this be concluded from a trend in Q95? I suggest to check whether the resulting flow regime would not rather classify as "intermittent river" instead.

We replaced "ephemeral" by "intermittent"

- Pg 11 line 41ff: Uncertainty from model structure and from model parameters are two different uncertainty sources – only the latter is considered. The sentence is not entirely clear in this respect.

We amended the sentence to make it clearer.

I am looking forward to seeing the paper in print.

Best regards,

Kerstin