

The authors present a hydrometeorological reconstruction established using a combination between a statistical downscaling method and data assimilation. This reconstruction is then used to study the multidecadal hydroclimate variability of the Seine basin, as well as the influence of the multidecadal variations on extreme events. The paper is well constructed and address relevant questions about the mechanisms of hydrological variability on the Seine basin. Figures are clear, and results and conclusion are relevant. I mostly have questions about some details in the methodological part.

Comments that should be addressed in the paper

Introduction

p.2 l.22-32 – After reading this paragraph, we think that statistical downscaling method are really not appropriate and that these methods should be discard. But in fact, you use this method with a second step combining observations with the results of a statistical downscaling method to improve them. Maybe it would be worth to reformulate some sentences of this paragraph to explain that these methods are not enough to well characterize climate variability, and that we should use an ‘add on’ or an ‘evolution’ to take into account observations. This couldn’t be considered as a completely different method.

p.2 l.33 to p.3 l.3 – Indeed, using observations could improve the results of a statistical downscaling method. But observations are also inhomogeneous so at the end, isn’t it the same disadvantage than statistical downscaling methods? Inhomogeneous trends? Combining observations does not cancel the disadvantages of the downscaling (so we have inhomogeneous trend from the reanalysis + inhomogeneous trend from the observations?).

Why only talking about statistical downscaling methods? What other methods could be used to reconstruct the past (dynamical downscaling, weather generators, ...) and why is it better to choose statistical downscaling + data assimilation?

p.2 l.6 Why the Seine basin? Can you add some explanations?

P2 l.8 Extending until 1850 leads to using 20CR between 1850 and 1900 (quality - and very large dispersion over the 56 members) + using very few observations (so inhomogeneous trends). How can we drive conclusions over this period given the poor data quality?

p.2 l.16 – Are the observations independent for the evaluation?

Data, models and methods

p.4 l.3 – The SMR developed by Moisselin, 2002, show significant inhomogeneous trends. It could be worth to add this fact the in text.

p.5 l.1-2 – Is Safran really independent from the reconstruction as the same observations are certainly used in Safran and data assimilation?

p.5 l.16 – At this point two questions:

- What about the spatial and temporal coherence? I suppose the spatial is respected as the same domain is used for the entire basin, but the temporal one?

- Isn’t it possible, for each day, to constrain the results of the downscaling to create trajectories in function of observations, instead of creating all the 56 trajectories independently and them choose the closest to the observations? In the first case, the final trajectories would be different than the trajectories from the downscaling method without data assimilation. -> I believe this is done by the

process explained l.17-32, and with only 3 trajectories instead of 56. Maybe reverse the explanations in the text, explaining the daily constrain before talking about the monthly one (not mention the monthly one before).

p.5 l.25 – Indeed, all the variables from Safran are obtained but the predictors for the downscaling method are only optimized for precipitation and temperature, isn't it? So what about the quality of the reconstruction for other variables, did you assess it? Is it good enough to used them in the hydrological model? It is for example possible to compare the reconstructed signals to Safran on the recent period.

p.6 l.7 – Is it possible to give a little more details about these tests? Comparison to what?

p.6 l.8 – Why 3 analogue days?

p.6 l. 10 – Is it possible to sum up the method in Bonnet, 2017 in a few lines?

p.6 l.15 – What is the influence of using different types of observations for the data assimilation on different periods?

p.6 l.22 – Is the spatial (temporal?) coherence conserved after the correction of these biases?

p.6 l.22 – Maybe change the title “Method” in “Extraction of multidecadal variability” as the previous paragraph was also talking about methodological facts?

Evaluation of the Seine reconstruction

p.7 l.20 – On Figure 3, we see that before 1900, reconstructions are not close to observations. This is not a surprise as 20CR has a poor quality before 1900 and the network of observations is less dense. The signal at Paris seems more “flat” than observations, with under and over estimations.

p.8 l.4 Which type of correlation is used?

Multidecadal hydroclimate variations

p.9 – Isn't it difficult to drive conclusions about other variables than P, T or Q as they are difficult to model in hydrological models?

Conclusion

Wouldn't it be more logical to reverse the conclusion part (for now, first) and the discussion part (second) in the conclusion?

Orthographic corrections

P.1 l.4 – Reformulate the sentence “This method improves the representation of daily flow as well as at longer time step”

P.1 l.4 – Provides

P.1 l.8 – Maybe “regulate” instead of “modulate” would be better?

p.1 l.12 – to influence the drought intensities

p.1 l.16 – Missing a “,” after “for example”

p.1 l.22, p.2 l.11 – Same remark for modulate / regulate

p.2 l.1 – Verify the expression “internal variability in climate and/or...”

p.2 l.20 – “It” not necessary in “which makes ...”

p.2 l.25 – “.” In the middle of the sentence, between “downscaling” and “of”

p.6 l.2 – Maybe add “spatial” before “error”. The sentence is not really clear, maybe there is a way to rephrase it, talking about spatial errors for both precipitation and temperature.

p.12 l.27 – The SSTs ~~there~~ are