Response to reviewer#2 of: “Influence of the multidecadal hydroclimate variations on hydrological extremes: the case of the Seine basin” by R. Bonnet, J. Boé and F. Habets.

First, we would like to thank the second reviewer for his interest in our study and for his general comments, which helped us to improve the structure and the clarity of the manuscript. Major modifications were made in the new version of the manuscript, which is much clearer now. Please find below the point-by-point answers to the comments. For clarity, all reviewer comments are in **bold**.

Summary:
This paper details research conducted on the link between multidecadal hydroclimate variations and streamflow in the Seine basin in France. Unfortunately, I have found it difficult to determine the scientific contribution of this paper due to its flaws in language, layout and lack of detail in the methods. I am recommending that this paper be significantly revised, in order that we may better understand the outcomes of this research. The major points that I recommend are:

We significantly improved the abstract, the introduction and the conclusion in the new version of the manuscript in order to better highlight the scientific questions and contributions of this work. Important modifications were also made to the method section, which is now more detailed and is much more clearer. An in-depth proofreading was also carried out. These points are discussed with the comments below.

1. The paper’s motivation needs to be better set out. The abstract states that precipitation and groundwater modulate river flows, and that extreme events are influenced by hydroclimate variations. This is not news to the hydrological community. What is novel here? Why is this interesting, and to who, and why?

We modified the abstract in order to better highlight the motivations and the insights of this work. What is really new in this work is the focus on multi-decadal variations, which are poorly characterized and poorly understood currently.
Please find below the new version of the abstract:

“The multidecadal hydroclimate variations of the Seine basin since the 1850s are investigated. Given the scarcity of long-term hydrological observations, a hydrometeorological reconstruction is developed based on a method that combines the results of a downscaled long-term atmospheric reanalysis and local observations of precipitation and temperature. This method improves previous attempts and provides a realistic representation of daily and monthly river flows. This new hydrometeorological reconstruction, available over a period longer than 150 years while maintaining fine spatial and temporal resolutions, provides an interesting tool to improve our understanding of the multidecadal hydrological variability in the Seine basin, as well as its influence on high and low flows. This long term reconstitution allows analyzing the strong multidecadal variations of the Seine river flows. The main hydrological mechanisms at the origin of these variations are highlighted. Spring precipitation plays a central role by directly influencing the multidecadal variability in spring flows, but also soil moisture and groundwater recharge, which then modulate summer river flows. These multidecadal hydroclimate variations in the Seine basin are driven by anomalies in large scale atmospheric circulation, which themselves appear to be influenced by sea surface temperature anomalies over of the North Atlantic and the North Pacific. The multidecadal hydroclimate variations seem also to influence high flows and low flows over the last 150 years. The analysis of two particularly severe historical droughts, the 1921 and the 1949 events, illustrates how long-term hydroclimate variations may impact short-term drought events, with in particular an important role of groundwater-river exchanges. The multidecadal hydroclimate variations described in this study, probably of internal origin, could play an important role in the evolution of water resources in the Seine basin in the coming decades. The way in which the associated uncertainties are correctly accounted for in future projections remains to be addressed."

2. You are missing significant volumes of literature surrounding connections between atmospheric circulation patterns and streamflow across wider Europe. e.g. (among many others)
• Folland, C. K., Hannaford, J., Bloomfield, J. P., Kendon, M., Svensson, C., Marchant, B. P., Prior, J., and Wallace, E.: Multi-annual droughts in the English Lowlands: a review of their characteristics and climate drivers in the winter half-year, Hydrol. Earth Syst. Sci., 19, 2353-2375,https://doi.org/10.5194/hess-19-2353-2015, 2015.
• Eva Steirou, Lars Gerlitz, Heiko Apel, Bruno Merz, Links between large-scale circulation patterns and streamflow in Central Europe: A review https://doi.org/10.1016/j.jhydrol.2017.04.003

We acknowledge that a vast body of work exists regarding the connections between atmospheric circulation and streamflows across Europe at the interannual time scales, but it is not the real subject of the paper. Our paper is specifically focused on multidecadal variations, and the problematics are different. We have also largely improved the presentation of the current literature, with a specific focus on multidecadal time scales.

3. You need to explain your method much more clearly. Can you summarize the method in Bonnet et al (2017) as it is mentioned over and over, and is seemingly critical to the understanding of your method here.

The method used to develop the hydrometeorological reconstruction is now presented in a dedicated section. The description was entirely rewritten in order to improve its clarity. The method used in Bonnet et al., (2017) is now explained in more details and we improved the description of the method used in this study. We also added a diagram in order to make the understanding of the method easier.

"3 Development of the Seine hydrometeorological reconstruction

A new hydrological reconstruction, based on hydrological modelling, is developed over the Seine basin, improving the method presented in Bonnet et al. (2017), with two main objectives: (i) to extend the study period to the 1850s, in order to characterize more robustly multidecadal hydroclimate variations, and (ii) to improve the representation of river flows, particularly at the daily time scale, in order to obtain a better representation of high and low flows and study their multidecadal variations. Figure 2 describes the main steps of the method developed in the present study and highlights the improvements over the one used in Bonnet et al. (2017).
To obtain the meteorological forcing necessary for hydrological modelling the main idea of the Bonnet et al. (2017) method is to use the analog method (Lorenz, 1969), a stochastic statistical downscaling method, to downscale a long-term atmospheric reanalysis such as NOAA 20CRv2c and produce an ensemble of trajectories of precipitation and temperature over France (Step 1, Figure 2). Then, local long-term monthly precipitation and temperature observations are used to select the best trajectory.
The analog method is based on the hypothesis that two days with similar large scale atmospheric states (e.g. large scale atmospheric circulation over the North Atlantic) are characterized by similar local weather conditions. In its most basic form, for each day D of the reanalysis, the day Da (the so-called analog day), with the closest large scale atmospheric state is searched in the learning period, defined as the common period between the reanalysis and the observational database with the local variables necessary for

hydrological modelling, e.g. here the Safran analysis. The local variables of interest of the day Da in the observational database are selected as an estimate of the local weather conditions for the day D. To quantify the similarity between large scale atmospheric states, four predictors are used in the present work: precipitation, surface temperature, sea level pressure and specific humidity at 850 hPa. An Euclidean distance is computed for each predictor, except for sea level pressure, for which the Teweles and Wobus score (Teweles Jr and Wobus, 1954; Obled et al., 2002) is calculated. The distances and the score are then combined after standardization to give the same weight to each predictor. Two domains of analogy are used. The domain for sea level pressure is delimited by the following coordinates: 44°N, 56°N, -11°E, 16°E. The domain for the three other predictors is defined by 46°N, 51°N, -2°E, 7°E.

In Bonnet et al. (2017), instead of searching only for the best analog day Da for each day D of the reanalysis, the N best analog days were selected Da1, Da2, ... DaN, with N = 10. The corresponding maps of precipitation Pr(Da1), Pr(Da2), ... Pr(DaN) and temperature Tas(Da1), Tas(Da2), ... Tas(DaN) from Safran constituted different estimates of precipitation and temperature for the day D. Multiple trajectories of precipitation and temperature over the domain of interest were then created by repeatedly selecting randomly one of the 10 analog days for each day D of the reconstruction. In practice, 5000 trajectories were created. The monthly averages of precipitation (temperature) for these trajectories were computed. From the 10 different maps of precipitation (temperature) over France obtained with the analog method (as N = 10) for each day D of the reanalysis, 5000 different monthly maps of precipitation (temperature) were obtained with this procedure (Bonnet et al., 2017).

For each month of the reconstruction, the 5000 maps of precipitation (temperature) obtained on the Safran grid were regridded and compared to the actual observed precipitation (temperature) map, using the long-term homogenized precipitation (temperature) series over France (see section 2) as reference. Regridding simply consisted in selecting the Safran grid point the closest to the long-term homogenized precipitation (temperature) stations. The spatial root mean square errors (RMSE) were computed for temperature and precipitation. The sum of the RMSEs corresponding to precipitation and to temperature was then computed, after the temporal standardization of the series of RMSEs in order to give the same weight to each variable. In the end, for each month of the reconstruction, the daily series of analog days among the 5,000 ones that leads to the lowest sum of RMSEs was selected. The term "monthly constraint" used in this study refers to this last step (it corresponds to Step 3 in Figure 2).

This approach benefits from the advantages of the analog statistical downscaling method. From the analog days, all the meteorological variables from Safran necessary to force the Surfex-AquiFR hydrological model were obtained. The spatial and inter-variable consistencies were maintained after this procedure, because for each day of the reconstruction the entire map of precipitation (and temperature, humidity etc.) over France from Safran was selected based on a single analog day. Compared to a basic statistical downscaling method, this approach allows additionally taking into account local observations in the downscaling process and not simply large scale information. This approach is therefore more accurate, as shown in Bonnet et al. (2017). Note that the temporal consistency of the meteorological forcing is ensured by both the temporal consistency of the predictors and of the local observations.

In the present study, to extend the study period, the long-term NOAA 20CRv2c atmospheric reanalysis (Compo et al., 2011), which begins in 1851, is used. This reanalysis is based on a global atmospheric model, using observed sea ice and sea surface temperature as boundary conditions, and with the assimilation of surface and sea level pressure observations. 56 members, sampling the reanalysis uncertainties, are available. Compared to Bonnet et al. (2017) where only one member of the long term reanalysis is downscaled, here we statistically downscale with the same analog method as described above, the 56 members of NOAA 20CRv2c. It leads for each day D of the reconstruction period to a much larger pool of analog days, which allows adding a new step: a daily constraint with local observations (Step

2, Figure 2). The objective of this additional daily constraint is to obtain a better representation of the daily variations of the meteorological forcing.
As previously, for each day D of the reconstruction period (1852-2008) of a given member, the N best analog days Da1, Da2 ... DaN in the learning period (1958-2008, limited by the availability of Safran) i.e. with the most similar large-scale atmospheric states are searched. In the present method, N = 50. As the 56 members of NOAA 20CRv2c are downscaled, in the end 2800 potential analog days are obtained for each day D of the reconstruction period (with potentially similar analog days for the different members). As each analog day corresponds to a day of the learning period, the corresponding daily maps of precipitation and temperature from Safran are selected and compared to the daily station observations (SQR, see section 2.1) after regridding. Regridding consists in selecting the Safran grid point the closest to each observation station over the Seine basin. Note that the number of stations varies on the 1852-2008 period. The comparison is therefore done each day of the reconstruction with the available stations.
The daily comparison is based on the following approach:
(i) The average daily bias in mean precipitation averaged over the Seine basin is calculated for the 2800 analog days, and the 60 analog days with the lowest bias are selected.
(ii) The spatial root mean square errors for the 60 analog days are calculated for temperature. For precipitation, the error to the cubic power rather than to the square power is used, in order to give more weight to strong values of precipitation, and the absolute value is used.
(iii) The daily series of spatial errors obtained for precipitation and temperature are then standardized based on the statistics of the entire period and added, with a weight of 1 for precipitation and 0.5 for temperature.
(iv) Finally, each day of the reconstruction period, the 3 best analog days (out of 60), i.e. with the smallest errors, are selected.
Based on these 3 selected analog days, a monthly constraint is then applied as in Bonnet et al. (2017) and described above, except that the number of analog days is different (3 versus 10) (Step 3, Figure 2). Multiple trajectories are created by repeatedly randomly selecting one of the 3 analog days for each day D of the reconstruction. The monthly averages are computed over the Seine basin and then long-term monthly homogenized precipitation and temperature series are used to select the best overall trajectory. The interest of using a monthly constraint after the daily constraint is that monthly data are homogenized contrary to daily data (Section 2) and therefore it allows for a better representation of low-frequency variations.
Multiple tests have been conducted to set-up the different ad-hoc aspects of the method, trying to obtain the best overall hydrometeorological reconstruction. These tests concern, for example, the best combination of weights given to precipitation and temperature errors, the number of analogs selected at each steps etc. For example, selecting only the 3 best analog days leads to best overall performance in capturing daily and monthly variations. Using more analog days may allow for a better representation of monthly variations but degrade the representation of daily variations.
To sum up, the hydrometeorological reconstruction developed on the Seine basin is constrained on a daily basis over the period 1885-2003 by observations of precipitation and temperature (SQR), on a monthly basis over the period 1885-2005 by homogenized observations of precipitation and temperature (SMR), and over the 1852-1884 and 2005-2008 periods by the monthly series of precipitation at Paris (Slonosky, 2002) (see section 2.1 for more details). The results, especially at the daily time scale have therefore to be interpreted with more caution over the period only constrained by the monthly series of precipitation.
During the development of the reconstruction, mean climatological biases were found on reconstructed precipitation and incoming shortwave radiation with comparison to Safran on their common period. These mean climatological biases are simply corrected based on Safran as reference before forcing the hydrological model.

The meteorological forcing obtained on the 1852-2008 period with the approach described in this section is finally used to force the Surfex-AquiFR hydrogeological model to obtain the hydrological reconstruction over the Seine basin (Step 4, Figure 2).

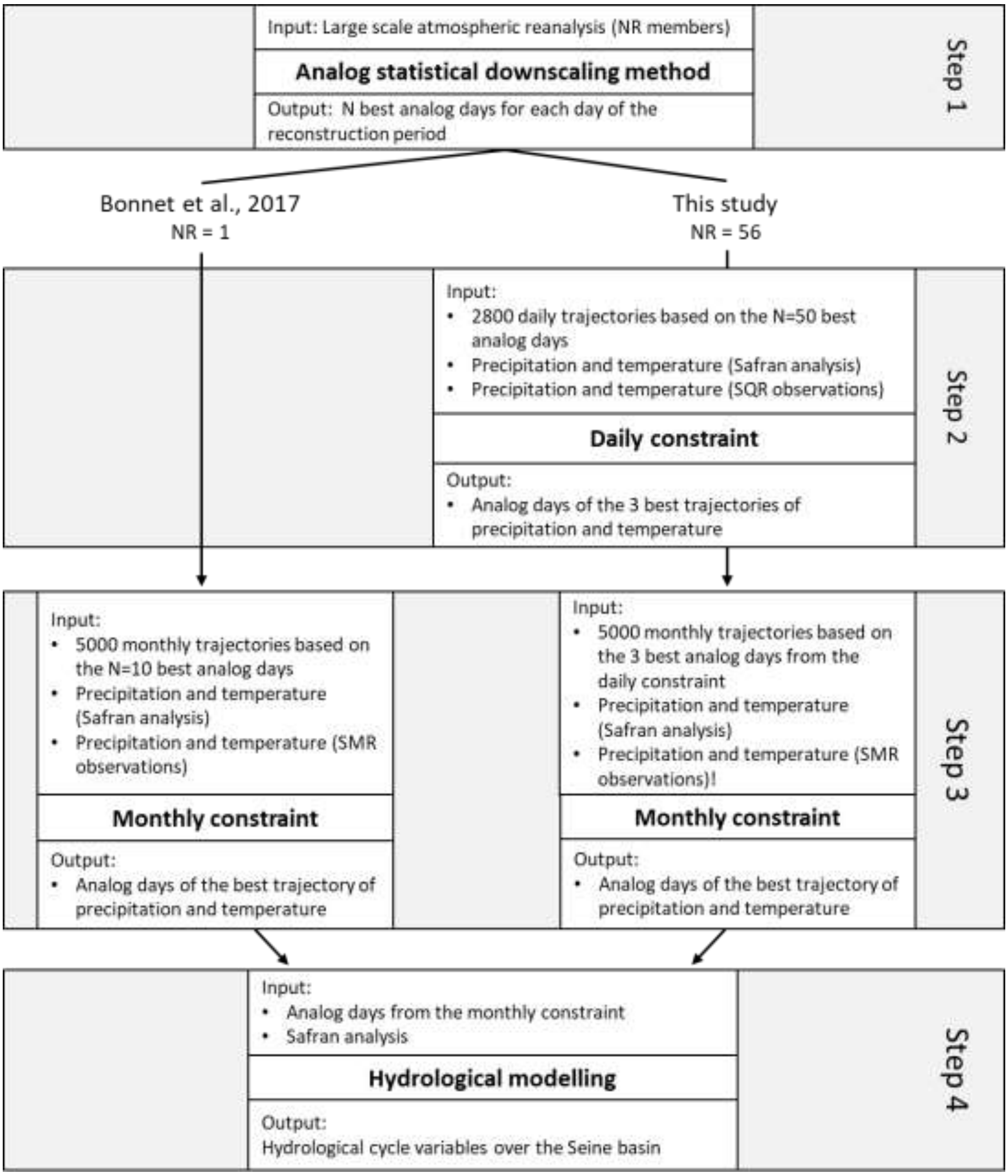

Figure 2. Schematic representation of the methods used to obtain the hydrological reconstruction, with on the right the method used in this study and on the left the method used in Bonnet et al. (2017), which doesn't include the daily constraint."

4. The paper's language be reviewed by a confident English speaker. The sentence structures are often incorrect, and the tenses are jumbled, which has made the paper very difficult to read. The English appears to improve later in the paper, so I think a little more effort is required.

We apologize for these language errors; a more in-depth proofreading was carried out in the new version of the manuscript.

5. The papers structure and headings need to be amended, especially in the methods section.

We modified the structure in the revised manuscript. As suggested by the third reviewer, we switched the 5th and the 6th section. The subsection about the development of the Seine hydrometeorological reconstruction, originally in the “Data, models and methods” section, was moved in a separate section. These changes improve the clarity of the paper.

6. Most of your figures have no legends, and many do not have appropriately descriptive axis labels. Please correct this to aid interpretation

We added legends into the figures and more descriptive axis labels in the new version of the manuscript.

7. Nearly all of your figures rely on red/green differentiation. 1 in 10 of the male readers will not be able to see this due to common colour-blindness. Please amend your color schemes to avoid this issue.

Done.

I have felt unable to go into the detail of the research for these reasons, and would be happy to re-review the paper once these issues have been addressed.