Referee #2:

First, and from the point of view of the structure of the paper, its title is too long and inaccurate. The work contains, in addition to the methodology, a case study and results obtained after applying the developed methodology.

In the title we wanted to highlight the importance of the methodology, as it integrates the water planning, the risk assessment and the possibility of making the bias correction in different ways that leads to the characterisation of natural flows. We did not include the case study in the title since it is a methodology that can be developed in other basins, taking into account its features. Thus, we can say that it is a generalist methodology and within its main framework (Fig. 2) decisions to the basin involved are taken. In this case, we chose the Júcar River Basin due to its hydrological features described in section 3, such as the high hydrological variability that lead to recurrent multiannual droughts and the exploitation rate of the water resources (≈ 90%), among others.

However, if you find it convenient, we can give a more concise format to the title integrating also the case study, these are some options we are considering:

- Characterisation of natural flows and modelling chain methodologies for risk assessment in water planning under climate change at Jucar River Basin.
- Methodology for risk assessment in water resource planning under climate change at Jucar River Basin.

Second, there is confusion between sections 2, Material and methods, and 3, Case study, since subsections 3.1 and 3.2, and perhaps 3.3, would be better classified as Material and methods.

You are right, now we think that developing the general methodology in section 2 and then describe the details of the methodology applied to this case study in the subsections of section 3 was a mistake. To resolve this, we will include an individual section to the case study, as it is a very problematic and interesting basin from the point of view of water management. Then, after this new section, we will develop the material and methods by joining section 2 and sub-sections 3.1, 3.2 and 3.3. In this way, we could reduce the length of the manuscript by bringing everything together in the same section, giving more sense to the structure of the document.

As for figures, figure 1 does not seem necessary and figure 2 is difficult to understand.

Figure 1 was introduced as a small clarification that water management and risk assessment are closely related and how each one works, then in Figure 2 we highlighted the key points of each process. However, we believe that removing Figure 1 and its explanation in this section is necessary. We can include something about this relation in the introduction and link it to the objective of the work, as a background for the way of working in this area up to now. Another option is to include this differentiation in the case study section, as it is part of a specialisation related to water resource management in problematic basins such as the Jucar River Basin, which are stressed due to high exploitation rates (demands/resources ≈ 1).

Regarding Figure 2, it is a general diagram of the methodology, which together with the surrounding text gives a general and quite explicit idea of what the study is about, detailing the main points and how to reach them. We decided to build it in this way in order to not overburden the reader with too much information in a single figure, as the process can be confusing and tedious for outsiders in this field.

Therefore, Figure 1 will be removed from the manuscript and we will try to redo Figure 2 to make it more understandable, for example by including the hydrological model and the bias correction in the item "Characterization of future natural flows". In the case that the way of doing it does not convince us, we will develop in the text its peculiarities to clarify the key points and guide the reader through the following sections, where each point will be developed in more detail.

The introduction lacks the reference to similar works that have incorporated climate change projections in decision-making processes in other basins, not just those in the Mediterranean environment. And this is important since the results of the work show a great dispersion (see figure 12).

In this part we decided to focus on studies carried out in this area due to the great dispersion of our results. In this way we justify them since most of the authors agree that the skill of climate change projections of this area is very low and usually they are not capable of representing the characteristics of historical droughts (Collados-Lara et al., 2018; Cook et al., 2008; Seager et al., 2008).

However, if we name similar studies developed in other areas, we could highlight the differences in RCM skills and the dispersion of results depending on the geographical area where they are applied. These differences could be named in both the introduction and the discussion.

As an example, we can name the studies developed in Sweden by Teutschbein and Seibert (2013) and in Germany by Hattermann et al. (2014).

The Material and methods section is quite robust since this work group has implemented numerous modules, already contrasted, in the Aquatool Decision Support System and now used (hydrological model; management model; water allocation model; stochastic model and risk assessment model). This paper provides the integration of climate projections into the model and its impact on future flows in the basin and on the storage of water in the system. In this sense, it uses nine Ensemble members (table 1) that cause a great dispersion of results, as already mentioned, and an inaccuracy in the conclusions. Would it be possible to use only those that have given better results in the Mediterranean region?

As we say in the text, we decided to use the ensemble provided by the SWICCA portal (Table 1), where they selected the members that were most suitable for all Europe. However, none of them provides a good enough fit in this area, both uncorrected and corrected, as they are not able to fit perfectly with the observed data and they are not capable of reproducing the statistical characteristics or trends in the average year (Figures 6 and 8) or over the whole period, neither the characteristics of the historical droughts. For this reason we thought that the use of the whole ensemble would provide us with more options and more robustness to the study, considering more options, since increasing the number of ensemble members reduces the sampling uncertainty (Collados-Lara et al., 2018; Thompson et al., 2017).

Despite all the efforts we made in the first part of the methodology to reduce the uncertainty provided by the RCMs, such as shortening the reference period to be more in line with the current situation of the basin or correcting both the meteorological data and the flows, this was not possible. Therefore the results are quite dispersed. This reveals the need to improve the skill of climate projections and the use of more sophisticated bias correction techniques, as we said in the discussion.

On the one hand, they work with flow data in the basin between 1980-2012 and, on the other hand, the reference period is reduced to the 1980-2000 period. However, as can be seen in Figure 5, there are differences in the average year inflows between the different periods. Can the use of these different periods have an influence on the results obtained?

In this basin, it is advisable to work with data from the period 1980-2012 (Reference), since using series with periods prior to 1980 can lead to an overestimation of water resources.This is due to the so-called "effect 80" (Pérez-Martín et al., 2013; Hernández Bedolla et al., 2019), mentioned and explained in the manuscript. It is a significant decrease in rainfall and inflows in the basin from the 1980s onwards. Thus, this is the reason why we decided to shorten the reference period provided by SWICCA (1971-2000) to 1980-2000, in order to try to better represent the current situation of the basin. In this way, we tried to decrease the uncertainty of the magnitude of future changes when we compare future flows with a reference period that represents the basin, such as the results of the average change rates of the whole basin (Figure 9).

Figure 5 shows how the inflows for both periods (1980-2012 and 1980-2000) can be considered equivalent since the difference between their averages is not significant. However, if we had used the period 1971-2000, we would have a different and unrealistic perception of the basin at this time, since water resources are greater than in reality and by correcting future data this would also be transferred to future periods.

Answering your question, yes, the use of the different periods influences the final results, but in this case the difference between using the period 1980-2012 and 1980-2000 is not significant. Indeed, we are in the process of publishing a paper (Suárez-Almiñana et al, *in press*) that compares the change rates for the whole basin (similar to Figure 9) using the three reference periods named before (1980-2012, 1980-2000 and 1971-2000), among others. In that paper we conclude that the average change rates of the basin are very similar when comparing the future flows of each period (2011-2040, 2041-2070, 2071-2098) with the flows of the reference periods 1980-2012 and 1980-2000. However, when those future periods are compared to the reference period 1971-2000, the average change rates are more drastic (up to -23% at the end of the century), which is logical since this period has more resources available, leading to a more extreme and alarmist conclusion than in this case where the average change rates are between -11% and -12% for the whole basin.

Thus, we think that the shortening of the period was a good decision and the differences between the reference periods 1980-2012 and 1980-2000 would not produce very different results.

The results obtained in figures 6, 7 and 8 are only visually compared. In the text it is written, for example, (lines 353-354): "There can be seen how both HBV models results are generally close to the observed flow values". Would it be possible to specify, from a statistical point of view, the term "close"?

In this case, with the term "close" we refer to the visual distance between their averages, but we can provide some table with basic statistics or Goodness-of-fit functions as the Mean Error (ME), the Root Mean Square Error (RMS), the Nash-Sutcliffe efficiency (NSE), etc. to give robustness to this statement.

The results of figure 9 show a great variability between options A and B, mainly in the two head reservoir, Alarcon and Contreras. In view of the results in Figure 12, could one option be recommended over another?

We cannot choose one option over another because the results obtained are not conclusive and very similar from both options. However, Figure 12 shows how the average of the ensemble is lower in option B and the shading area reaches much lower values than in option A. Therefore, despite the fact that option B is more dispersed, if we chose it we would be working from the side of security against future intense drought events, which seems to be more frequent and intense in the future (CEDEX, 2017 and Marcos-Garcia et al., 2017).

Some minor comments would be: - Figure 2: the acronyms of P and T have not been previously defined - Line 133: the acronym RCM is defined later (see line 236) - Lines 223-226: There are several references to geographical names such as the Albufera of Valencia that are not shown on the map in Figure 4.

We will define the names of the acronyms (P, T and RCM) prior to their appearance in the text or figures. Regarding the Albufera de Valencia, we will include it in the map of Figure 4.

We hope that our responses to the reviewers' comments and the changes we will make in the manuscript will be enough to be considered for publication in the HESS journal.