**Referee #1:**

This manuscript introduces a methodology to assess the effect of climate change on water resources systems by using a model chain connecting climate model results, a semi-distributed rainfall-runoff model and a water management simulation model. While I think the themes dealt with in this work are relevant and interesting for the scientific community, I am afraid the manuscript requires a severe degree of revision before it can be considered for publication in HESS.

As you consider this work is interesting for the scientific community, we will carry out an in-depth review of the manuscript in order to solve all its weak points, both linguistic and structural issues, as the clarification of results.

To be accepted for publication in any journal, the first thing the authors should do is a thorough language and structure revision. At present, the text is clumsy and difficult to read.

We consider that the manuscript may be difficult to read since English is not our mother tongue and the process of writing in detail all the steps of the methodology is very complicated. We understand that some expressions or meanings that we think are correct may not be right and difficult to understand for other readers. For this reason we will make an in-depth language review to make it more understandable and to ensure that grammatical structures and the vocabulary used are correct.

In any case, one of the strengths of the HESS journal is the English language copy-editing, so if after our review of structure, grammar, etc., it needs to be improved, we are sure that this problem will be solved before its publication.

In addition, I found some relevant information out of place, e.g., the mention to the improvement introduced in the study in page 4, which should appear in the conclusions section too (I will come back to the added value of the study later), or the description of the drought risk indicator calculation in section 4.3 that should be in methods (I will come back to this too).

We will respond in more detail later, as you will come back to them in other comments.

Anyway, in the first case you refer to this statement (page 4): "The improvement developed in this study lies in the characterization of future natural inflows and the combination of the management and risk assessments." You are right, we will integrate the main improvement in the conclusions as there we focused in the results obtained in each part and what conclusion we extract from them.

Regarding the description of the drought risk indicator in section 4.3 (Results section), we know it is out of the place, but we considered important to guide the reader in the process in order to know where this results come from, as this is a complicated process and no easy to understand. As we said before, we will come back to this later.

Regarding the content of the manuscript, my first major concern is its absolute lack of focus.

Perhaps we did not guide the text to the key points or results, as there are many points to be covered. As you saw, the study is very complete with different options and we did not want to leave any point out or unexplained. We will solve this problem in the following comments and their responses.

Section 1 introduces a series of concepts mostly disconnected from each other and somehow irrelevant for the rest of the manuscript. Half of the introduction talks about climate services and how to deal with the data they provide, yet no further mention is made to them later in the text either in the discussion of the conclusions section.

We consider that highlight the increasing amount of climate services and the huge available data to work with is important, as well as the lack of a clear rule for handling them. As we say in the introduction, some authors choose to use the ensemble, others select only those that fit with the observed data from the reference period, then they correct them or not, etc. With this content, our idea was to communicate to the reader the complexity of working with climate change projections. Not only due to their large number, but also due to the amount of work involved in their selection and subsequent treatment, including the inconvenience that in Mediterranean areas the skill of these data may not be sufficient (Suárez-Almiñana et al., 2017; Barranco et al., 2018; Collados-Lara et al., 2018), thus obtaining disperse results with great uncertainty, as it happens in our case.

Furthermore, this information is used to justify the decisions taken during the first steps of the methodology, such as in the choice of SWICCA data as inputs (due to the easy downloading, handling and confidence in the selection of models adapted to all Europe made by SMHI, an institute of recognized prestige in these issues). In addition to the need to correct the data and use all the members of the ensemble for the study. All this was later named or justified in section 2 (lines 128-134), sections 3.1 to 3.3, section 4 (lines 304-308) and section 5 (lines 483-495, lines 504-506) of the manuscript.

Therefore, we do not believe that everything introduced here about climate services and the way to work with them is irrelevant. It is, after all, a way of exposing what we use and why in this study.

However, we will improve the introduction as discussed below.

The last 20 lines of the introduction more or less describe the problem the authors want to study and one could discern what the research objective is. However, it is not the task of the reader to guess the objectives of research work. The authors must explicit what they want to achieve with their work and communicate this to the reader in an efficient and straightforward way.

Based on the previous comment and this one, we realize that we have not reached the objective we wanted explaining all this in the introduction, therefore, the reformulation of the introduction using a clearer and more direct language it is necessary. There, we will indicate why we present all this information, making more references to it in the other sections of the manuscript and specifying what is the main objective of this methodology developed to help decision makers to face an uncertain future.

I would like to add that developing a methodology is not an objective in itself but rather a tool to pursue the answer to a research question. Concluding that the methodology is general enough to be applied in other case studies would be an acceptable conclusion though.

Thank you for pointing this out. We believe that specifying the question to which we want to give an answer with this methodology is necessary. In this case, the aim is offering a tool to the decision makers that provides both, deterministic and probabilistic intuitive results, for different future periods and thus have the opportunity to apply measures, in order to avoid or mitigate the possible adverse effects of climate change. Using this tool it is also possible to test whether the proposed measures provide an improvement in the future state of the system or not. This is possible by modifying the conditions of the system when applying the measures and re-simulating the risk assessment model to obtain results about greater or lesser probabilities of having certain deficits in the demands, or certain volumes of water resources in the basin.

All this information clarifies why the general methodology was developed and its applicability to other case studies, even though the step of applying the measures and testing them is beyond our scope in this work.

I do not have major concerns about the methodological approach of the research: climate-hydrology-management, but there is an evident need to improve the clarity of exposition in terms of sections arrangement and description of methods.

You are right, it seems that the problem of this manuscript is the way of telling the whole process, which is not clear and a little unstructured, so we will make a great effort to change all these aspects. We detail them below.

With regard to sections arrangement, I have the impression that sections 2 and 3 are not separated appropriately. I would suggest that the case study was presented first. The Jucar River Basin is an extensively studied catchment in literature, especially from this research group. Still, I think the system deserves having its own section. Afterward, the whole section 2 and subsections 3.1 through to 3.3 should be merged in a single “materials and methods” section.

Perfect, we will follow your suggestion, in this way we give more importance to the case study and then we can develop further certain points mentioned above by putting together section 2 and 3.1 to 3.3. This change will allow us to reduce the length of the manuscript and be clearer and more concise by explaining everything in the same place. Despite this restructuring, we also have to clarify that this is a general methodology that has to be adapted to each case study, in this case it was applied to the Jucar River Basin.

Coming now to the description of methods, I think the part between lines 89 through to 115 requires a better explanation, including justification of figures 1 and 2. Line 89 reads: “In this section, a distinction between the current assessment in the management of water resources and the analysis of risks was made, despite of being intimately related”. Disregarding the quality of this sentence, there is nothing in section 2 that actually deals with that distinction unless the reader is imaginative to say the least. My assumption is that the authors call one thing (current way) to management made on the basis of current/past climate analysis and they call another thing (risk analysis way?) to the assessment of management under future climate, and they argue that the two approaches should be integrated. I do not understand the reasons for such extravagant differentiation of a traditional present versus future analysis that does not add anything conceptually new to the current state-of-the-art.

We made this differentiation by the way water resources are managed in the case study. Currently, the water allocation or management model is used for water resources planning for a horizon of 6 to 18 years, and then, for real time management and droughts events the managers use the risk assessment, which is normally used for a horizon of 1 to 12 or 24 months. All this is stated in the Júcar River Basin District Management Plan (CHJ, 2015) and the Drought Management Plan (CHJ, 2018).

What we are trying to say with these figures is that we can take advantage of both methods for the future by inserting climate change projections into them and this methodology has not yet been integrated into the River Basin Management Plans design. Thus, Figure 2 shows the steps to do this process and throughout the text, we explain it step by step. At the end, we can extract the results of the evolution of the basin's water resources in a deterministic (Figure 12) and probabilistic (Figure 10) ways to help decision-makers to take decisions for the future.

However, we consider that we did not introduce this information in a clear and understandable way, so we will remove Figure 1 and then we can introduce this information in the introduction section or in the new section dedicated to the case study (commented on later), as it is the background or the current state-of-the-art of this area.

Now is when things get spicy. The authors mention that the novelty introduced in this research “lies in the characterization of future natural inflows and the combination of the management and risk assessments”. But, what is new about determining streamflow under future climate conditions and compare it against present conditions? I want to think this must be a writing error from the authors as I do see more value to the results they present further than just comparing two situations.

The point here is the process related to the characterisation of the future river flows and the planning and risk assessments. With this statement, we mean all the process involved in the treatment of climate change projections in order to adapt them to the basin features and then the model chain developed to extract some results that complement each other. All this means the bias correction (before or after the hydrological model), the future simulation of the water management model, the integration of the statistical properties of each ensemble member into the stochastic model to generate multiple and equiprobable series, their integration in the risk assessment model, and finally the extraction of the risk indicator and the evolution of water resources of the basin in the future. We do not refer only to the change rates, which we know are very common in this type of studies. Anyway, as we said in the discussion section (lines 517-526), this change rates for the entire basin may be more reliable than those from other studies because we used a reference period adapted to the current situation of the basin. This is another new income that has to be considered as an improvement of the technique.

We will include all this information in the text.

Continuing down the line, I think section 3.3 is the core description of the methodology. I suggest it appears earlier in the text and that it uses a more generalized language only mentioning that modules from the Aquatool software will be used (substitute Aquatool modules' names by generic names, e.g., EVALHID rainfall-runoff model, MASHWIN stochastic streamflow series generator, SIMGES water management simulation model).

Yes, this section corresponds with Figure 2 and Figure 3, and there is where the main part of the methodology is detailed. In this case, as we agreed to join sections 2 and 31-3.3, this part may be the core of that material and methods section with the help of Figure 2 and a new version of Figure 3, as you demand below. Of course, we can use a more generalized language substituting the names of the Aquatool modules by the name of the models. Then, we will add a new section to introduce this software and its modules as you suggested below.

Next, I suggest sections 2.1 and 3.1 are merged into a single 'current and future climate (or better name)' section. Section 3.2 could be a subsection of the new merged section.

We agree with you, the parts of sections 2 and 3 with the same information will be merged in single sections, as we said before.

By the way, the words precipitation and temperature do not appear so often in the text and are not that long to require using an acronym. I suggest you revise this.

We agree with you, so we will remove the acronyms and we will write precipitation and temperature throughout the text and figures.

Finally, all the models that are actually part of the Aquatool software would be better together under an 'Aquatool modeling package (or better name)' section.

Yes, that makes sense if we name the Aquatool modules with a more generalized language as we agreed before. The name of the section may be Aquatool Decision Support System Shell or similar, where we will explain in detail the features of this software and the modules used for building the different models.

I would like the authors to clarify their position with regard to bias correction. By the way, figure 3 is an absolute mess, it should be revised for clarity. The authors claim in the discussion that "working with the raw data would lead to unfavorable results for the future since the underestimation of flows in the headwaters is notable, this fact may also lead to alarming conclusions about the future hydrology in this basin". But, in figure 6, we see that bias correction actually changes one problem for another, especially at the resource generating catchments of Alarcon and Contreras. While the uncorrected data fits visually well precipitation between March and September (dry months), underestimating it during the wet months in the winter, the bias-corrected data overestimates spring and summer precipitation while still underestimating winter precipitation. This is potentially a problem if the extra amount of water in the summer introduced with bias correction exceeds the winter deficit of the uncorrected data. I think this might be explored.

We thought that Figure 3 was a good way to introduce the reader to the two options considered in the characterization of future natural flows. There are all the steps we followed in both options and we think that explaining them belter in the text will be enough to understand it properly. However, we can remove it or make it simpler to the reader, as we have to restructure the manuscript, we will think about it.

Regarding the bias correction, looking at Figure 8 (Non corrected flows) we decided that it was necessary to correct the data because the underestimation of flows in the Alarcon and Contreras basins was huge. This is a problem from the point of view of water management as there is where the main reservoirs of the system are placed. This means that if we use this data, we are accepting that the resources we are taking as inputs are much lower than those that are really there, which is not acceptable in this field.

Thus, once we decided to correct the data, the most recommended method to do it was the quantile mapping, which tries to keep the mean and standard deviation of the reference series (Collados-Lara et al., 2018). Then, we applied it but the results were not convincing because precipitations and flows were overestimated in spring and summer months. Despite this, we accepted them as the differences between the averages were minimised and the flows of the headwaters basins were better fitted to the observed values (Figure 7).

However, we tried other techniques for bias correction in previous studies, as month-specific correction factors (Suárez-Almiñana et al., 2017), which results were not so different from those of this study. Hence, we think that the currently available methods of bias correction may not provide satisfactory fittings to the observed series of this area. Despite this, a future consideration can be the application of seasonal corrections or one method that be capable of catch the drought statistics of the observed data, as we know that the RCMs are not able to reproduce them in some cases (Collados-Lara et al., 2018; Cook et al., 2008; Seager et al., 2008). We referred to all this in the discussion section, lines 491-499.

Anyway, in the next comments we refer to the bias correction issue in more detail.

The descriptions in 4.2.1 and 4.2.2 correspond to the methods section. The authors should limit to describe the results in these sections.

We thought that repeating the process in these sections was necessary to guide the reader through them and refer to how they were obtained. Following the whole process without any guidance may be complicated and can lead the reader into confusion.

Anyway, when we restructure the manuscript we will consider removing or reducing these descriptions from these sections, as this is not the place for them.

Moreover, I think the section would benefit from merging figures 7 and 8, although I am not sure whether the first column in figure 8 should be maintained for what I will mention next.

We are not sure about putting these two figures together as they may be too stacked, the text may be illegible. However, we will consider their combination when we restructure the text.

Figure 9 shows the mean rates of streamflow change for the three future periods with regard to the reference period. Did the authors check the rate of change between reference and future periods of non-corrected flows from option B? I think this might be revealing. Also, the size difference between graph A and graph B in figure 9 should be revised.

We did not check the average change rate of non-corrected flows because the flows of the reference period are hugely underestimated in the reference period compared to the observed data (Figure 8), mostly in the headwaters basins.

Anyway, the change rates of river flow are available in the SWICCA portal for each future period, but the results without skill in the re-forecast analysis are almost in the whole basin. Thus, we thought that using these change rates would not give us realistic results for the future. However, we can test it with our flow data, and if they are revealing we will include them in the study.

We will review and fix the size differences in Figure 9.

Regarding the final step of the methodology, relative to water management simulation, I think the authors should justify better the added value of using stochastic modeling when they already have a reasonable amount of data (several 30 years series from various climate models) to perform the statistics relative to water storage.

We think that only 9 series (one for each ensemble member) are not enough to apply the risk assessment process, even if we divided it in periods of 30 years. The more equiprobable series we generate with the statistical properties of each ensemble member, more agreement between them, which means more reliable results. This statement is not shown in the ensemble risk indicator (Figure 10) as the ensemble members are quite disperse (dry or wet periods) and they complement each other in the ensemble resulting in almost the same probabilities of being in any volume interval of the total capacity of the system.

I think the results from using the water management simulation model should appear before the ones from the stochastic simulation and, in any case, both results should be comparable (e.g., calculating the drought indicator, or the exceedance probability in September).

We can place the results of water management before those of the risk management. Actually, they are already comparable because in both cases we are showing the evolution of the water resources of the system, one in form of risk indicator with probabilities for each future period and in the other with mean volumes of the ensemble and the range covered by it for the entire period. Thus, they complement each other.

The exceedance probability at the beginning of the irrigation season might be a relevant result too.

You are right, in fact it is possible to extract the exceedance probability for each month, but we focused on September because it is the end of the irrigation season and the end of the hydrological year. Thus, this result is probably the one that best summarizes the final state of each campaign. In addition, this is a meaningful data for the stakeholders because it is better understood by them.

Anyway, we will think about including the exceedance probability at the beginning of the irrigation season (March) because it can be a good way of informing the irrigation associations about the possibilities of having shortages and take measures to avoid them. Thus, we may show both results in the reviewed manuscript.

Anyway, the results of this section show how the relevance of the bias correction is dampened through the modeling chain. Considering limitations and additional uncertainty introduced by bias correction highlighted by Ehret et al. (2013) and the results of studies like Muerth et al. (2013) who argue about the utility of bias correction in model chains, I think the authors lost a good opportunity to contribute to the existing debate on the added value of bias correction in the modeling of climate change impacts on water resources.

You are right, the debate on the added value of bias correction in the modelling of climate change impacts on water resources is very interesting, but that is not one of the purposes of this study.

What is important here is that the raw data were very inappropriate in this case, and we applied the most recommended method to correct them, the quantile mapping (QM). We know that it has its pros and cons, so we decided to test it on the meteorological data and on the flows, just in case there was a notable difference between them and be able to recommend the better option. In this case, the difference was not significant.

However, the papers of  Ehret et al. (2012) and Muerth et al. (2013) are very interesting, such as the one by Teutschbein and Seibert (2013) which applies the QM in different seasons and the one by Switanek et al. (2017) in which the QM method for climate change applications is improved. We plan to mention all these works and go a bit further into the subject of bias correction in the discussion, but as you will understand, we cannot develop in detail this part since the important thing in the paper is the methodology and its adaptation to the case study.

By this, we mean that the methodology will already be developed by the time we are able to obtain better climate change projections adapted to each area or correct the data to a better fitting to those observed in every sense.

I would not like to finalize my review mentioning that the authors make the wrong use of the term tendency throughout the whole text to my understanding. Mostly because the authors do not show whether their results really follow any trend and whether this is significant.

In this case, the term tendency refers to the decrease of flows as we approach 2100, which is shown in the average change rate of the Jucar River Basin (Figure 9) and in the average ensemble of the system's resource after applying the water allocation model (Figure 12).

Perhaps the trend is not so evident because the reduction is relatively low (from 1% to 12%), but considering that we are working with the Jucar River system, these decreases can lead to major problems of water deficits and huge economic losses.

Therefore, these small decreases are significant, as the basin is already stressed (demands/resources ≈ 1) and this presents a big challenge for decision makers during extreme events such as droughts, which will be more intense and longer according to authors such as CEDEX, 2017 and Marcos-Garcia et al., 2017.

We will try to provide more information about this topic in order to the reader understand why this small decrease is relevant. In addition, we will not use the word "tendency", we will directly refer to the "decrease" in order to avoid any misunderstanding.

**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.**