Water restrictions under climate change: a Rhone-Mediterranean perspective combining 'bottom up' and 'top-down' approaches"

Sauquet et al.

Anonymous Referee #1

The paper "Water restrictions under climate change: a Rhone-Mediterranean perspective combining 'bottom-up' and 'top-down' approaches" presents a study that uses decision scaling (Borwn et al., 2012) to evaluate future water restrictions pattern in southeast France. I think the topic fit the scope of this journal, but I have several major concerns on this manuscript.

First, the novelty of the paper is questionable. Applying a bottom-up approach such as decision scaling to evaluate climate change impact uncertainty is not a new topic in the field. Although authors might argue that presenting the climate response surface in WR (not streamflow) is relatively new, I do not see any additional information regarding policy inform that can be generated from this result. Some visualization techniques used in this paper could be attractive (such as using color to represent mean value and size to represent the s.d.) but I failed to understand the overall scientific contribution of this paper.

Second, the modeling framework is extremely unclear. Authors use Section 4 to explain their method but they spent a lot of space to explain decision scaling which is other people's work. They briefly mention the rainfall-runoff model they used but no details about the actual water restriction level modeling framework (Section 4.3). They explain their concept of computing WR in fair details (which is helpful) but I still do not understand how they build the WRL model. What are the input and output of this model? What parameters can be calibrated in this model? How to authors link this model with the rainfall-runoff model? Information about these is partly provided in Section 4.3 but hard to follow from a reader's perspective.

➔ Authors agree with this remark and the method needs to be more explained.

Inputs of the WRL model are daily discharges and precipitation. Outputs are WRL for each of the 21 10-day periods for each year spanning the April-to-October period. VC3(t) is first computed from daily discharge Q(t) every day t, WRL(t) is then deduced by comparing VC3(t) to the four regulatory thresholds and finally a unique representative WR level is assigned to each of the 21 10-day periods, as the median of WRL(t) observed or simulated within that 10-day period. To best match the whole monitoring process stated in most of the DMPs, a simple precipitation correction was applied ("Pcorr", in Fig. 5). It consists to give a 'no alert' when precipitation during the preceding 10 days exceeds 70% of inter-annual precipitation average, regardless of the WR simulation results. The WRL framework is applied to observed and simulated data of both discharge and precipitation. To assess the performance of the WRL model under current condition against stated WR decisions, the WRL model is run with observed daily discharges extracted from the HYDRO database (named "HYDRO" in the text) and with daily discharges simulated by the rainfall-runoff model GR6J forced by the SAFRAN reanalysis (named "GR6J" in the text). In the context of climate change the WRL model is run with daily discharges obtained with GR6J forced by one of the 1350 sets of perturbed precipitation, temperature and PET time series. In this later case the regulatory thresholds are calculated on the simulated discharge time series to limit the possible effect of bias in rainfall-runoff modeling.

Third, lack of in-depth discussion on the policy implementation. Given that authors use WR in the climate response surface, one can expect that authors should use a lot of space to link their results to drought policy implementation or some information about the adaptation action. However, only a short discussion of WR has been provided at Section 5.5. Given that this is not a methodological paper, these in-depth discussions become the critical point to prove that this paper is worth to be published because readers around the world can learn from this study and apply it to their own drought management policy.

➔ Discussing the policy implementation is out of the scope of this article. This paper presents a first attempt to simulate water restrictions over a large area in France. This paper aims at promoting the approaches developed in parallel by Brown (named 'Decision Tree Framework") and Prudhomme (named "Scenario neutral approach") and one of the challenges was to define critical thresholds of unacceptable number of days with legally-binding WR for irrigation use. This paper suggests using information provided by insurance (here from a national system of compensation) at the regional scale.

Finally, the structure of the manuscript and English is extremely difficult for readers to follow. The general outline of the paper follows a typical modeling paper while authors introduce their study area and data than their model. However, as I mentioned above, the modeling framework especially for the "Water restriction level modeling framework" is not clear at all. Also, there are general equations list in the results section (Section 5) and irrelevant results (Line 432-474) presented in the result section.

➔ "Results at both local and regional scales are presented and discussed in Sect. 5 before drawing general conclusions in Sect. 6." ➔ "The scenario-neutral approach is applied at both local and regional scales and results discussed in Sect. 5 before drawing general conclusions in Sect. 6."

There are A LOT of grammar errors and typos that make the manuscript hard to read. This is surprising that one of the co-authors is from the UK.

➔ The text will be screened to correct grammar errors and typos.

Line 34 - What do you mean by "changes" Climate or human activities?

➔ We mean both since we are dealing with global change issues.

Line 35 -What kind of drought? Climatic? Hydrologic? Or economic?

➔ "The intensity of the changes is still uncertain, however, climate models agree on significant future increase in frequency and intensity of meteorological, agricultural and hydrological droughts in Southern Europe"

Line 86 and 88 - You are arguing with yourself. In Line 86 you said water is abundant globally but Line 88 you said water resources are under high stress.

➔ At the regional scale, water resources are abundant. However during low flow periods there is an intense competition for water between different users and needs—agricultural, municipal, industrial, the environment—resulting in tradeoffs between human demands and environmental needs. The French RM Water Agency has identified areas with persistent imbalance between water supply and water demand (around 40% of the RM district, http://www.rhone-mediterranee.eaufrance.fr/gestion/gestion-quanti/problematique.php).

Line 90 - Why 43% is high proportion? It is less than the half.

➔ Water abstraction for irrigation needs is ranked first in terms of volume; that is why we have considered that 43 is high proportion. We will make changes in Section 2.1. "The total net water withdrawal is around 6 billion of m^3 in the period 2008-2013 (water abstraction for cooling nuclear plants and hydropower is excluded) with a high proportion of them to support irrigation needs (3.4 billion of m^3, including 2 billion of m^3 for channel conveyance). Total annual abstracted volumes for drinking water and for water for industrial uses represent 1.6 and 1 billion of m^3, respectively."

Line 96 - You never explain what is "Drought management plan?"

➔ Drought management plans define specific actions to be undertaken to enhance preparedness and increase resilience to drought. This definition will be included in Section 2.2.

Line 111 and 115 - You are arguing with yourself agian. If water restriction decisions are frequent (Line 115), why these catchments are with minor human influence? Water restriction decisions are human influence.

➔ The 15 catchments are benchmark catchments where near natural drought event can be observed. Water can be taken in another nearby catchment. This will be added in section 2.3 to avoid misinterpretation.

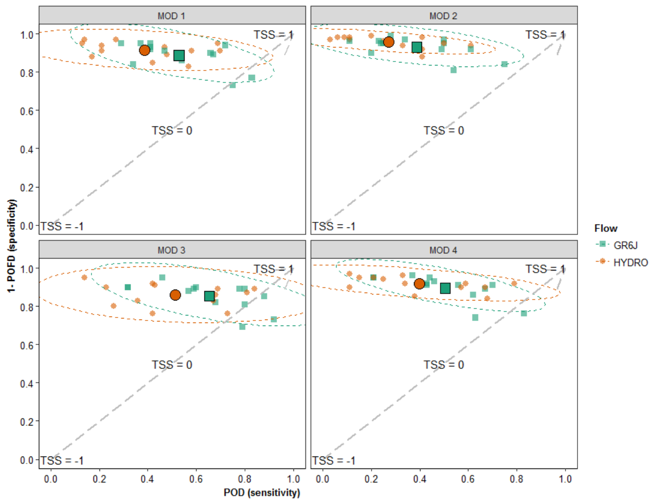
Line 173 to 174 - Will the selection of index affects all of your results? You should discuss this in the discussion section.

➔ Indeed choosing the same definitions for the monitoring indicator and regulatory thresholds may partly explain the deviations to the stated WR. This was a simplification assumption. Before stating for VC3 and 10d-VCN3 the four prevalent modalities have been chosen to implement WR simulations:

Modality	Monitoring variable	Threshold variable	Benchmark period
MOD1	QC7	10d-VCN3(T)	1958 – 2013
MOD2		m-VCN3(T)	

MOD3	VC3	10d-VCN3(T)
MOD4		m-VCN3(T)

where m-VCN3(T) refers to quantiles defined by monthly. Results show a weak sensitivity to the choice of these variables. In terms of sensitivity, MOD3 on average performed best than other modalities for both HYDRO and GR6J simulations on the 15 catchments. MOD3 was finally considered for all the catchments.



Skill scores obtained for the WR level model over the period 2005-2013. Large dots show the mean values of the skill scores. Colored dotted lines are confidence interval.

It will be discussed in section 4.3.

Line 186 - Why cross a threshold is unsustainable? How do you know it won’t come back? Quantify sustainability is a difficult challenge and if you don’t know what it is, you should not use the word. Otherwise, you should define sustainability.

➔ Sustainability - like vulnerability - has no universal definition. Sustainability assessment is based on the analysis of failures or unacceptable conditions that lead to low crop yield and quality, and consequently to economic losses at such a level that the national system of compensation is initiated. In this application,

- we assumed that irrigated farming is not sustainable if restrictions during drought periods are , on average, too severe - i.e. duration with limited or suspended abstraction for irrigation above a critical threshold - to ensure enough water for crops;
- since it was not possible to compute the effect of water restrictions on crop yield and quality (no crop modelling was considered here) and on economic losses, we used ‘agricultural disaster’ notifications as proxies to identify the conditions that would be unacceptable/damaging for farmers activities.

This sustainability is thus indirectly related to agricultural economy (not directly related to losses expressed in euros).

Line 190 - What do you mean by intersection?

→ Each component contributes to the vulnerability assessment of the system (including its management) to systematic climatic deviation.

Line 215 - I don't see any calculation related to irrigation water use in your 4.3? How you do this?

→ Only the impact of WR on irrigation has been examined here. Irrigation is selected since it is the sector which consumes most water at the regional scale. Water needs are not computed. The impact of WR is highlighted by the 'agricultural disaster' status notified at the department scale and we have assumed that when the total number of days with legally-binding water restrictions exceeds a fixed threshold (defined using data from the year 2011), the situation for farmers is unacceptable (significant losses) and as a consequence the national system of compensation is initiated.

Line 254 - Don't understand what you mean.

→ The physical components (drought severity) that lead to WR decisions are only considered in the WRL model and no socio-political factor was taken into account to reproduce water restrictions.

Line 262 - Why not use the worst WRL as indicators? And also why not just use daily time step as your rainfall-runoff model? Why change it to 10-day?

→ Water restrictions are decided after consulting drought committees that convene irregularly. The time-step for modelling WRL was chosen to be compatible with the frequency of drought committees estimated from the analysis of the water restriction orders: WRL is thus computed at a regular time step of ten days.

Line 302 to 303 - I do not understand what is your point here. If you know this, then why don't you model that? This means you understand that just a hydrologic model is not enough to do this type of modeling but you still do it and write a paper about it. This just implies that your model is not only WRONG (as all models are) but also not very USEFUL.

→ The model is not totally wrong. The WRL framework is able to reproduce the physical bases in the making-decision process and thus can simulate 68% of the stated restrictions over the period 2005-2013 (performance obtained with "GR6J", section 4.3); reaching 68% is not so bad.

Results of our study (conclusions based on the 15 catchments) show that: (i) surprisingly there are noticeable deviations between the drought severity perceived on discharge data and the final decisions to order restrictions but the decisions are not totally uncorrelated with drought conditions, and (ii) most of the catchments are subject to deviations. The performance is judged acceptable to be applied in the scenario-neutral approach. We are aware that the WRL model is far to be perfect and we are convinced that the WR framework will be improved if relevant socio-economical controls are introduced and it will be certainly a challenging task. Just keep in mind that this study is a first attempt to simulate WR decisions at the regional scale.

Line 357 - What drivers? I thought in climate change studies, T and P changes are drivers.

➔ Indeed temperature and precipitation are the main physical drivers. Here we wanted to assess if WR is more sensitive to P and T over a specific period.

Line 358 to 359 - I don't understand your English.

➔ Response surfaces have been displayed for different pairs of potential climate drivers (X and Y related to temperature and to precipitation, respectively). Their shapes were first examined (a flat response surface is an evidence of no link between WR and (X, Y)). In addition we have used metrics which measure globally the dispersion around the response surface: the median and the maximum of Sd (small values means small deviation and thus strong links between WR and (X, Y)). The drivers are (X, Y) for which the dispersion is minimal and with the most contrasted surface response.

Line 402 - Typo.

➔ "tree" should read "three": "Finally the vulnerability resulting from the combination of the three components sensitivity, sustainability and exposure"

Line 432 to 474 - I do not understand why you have these results here which are not related to WR.

➔ This section details the result of the classification carried out on the 106 individual WR response surfaces, which is consistent with the title "5.4 Response surface analysis at the regional scale".

Line 788 - There is no need for Figure 2.

➔ Indeed this figure could be deleted.

Line 797 – The explanation of Figure 5 is unclear. This result in my second major comment regarding the modeling framework. A better explanation needed.

➔ Inputs of the WRL model are daily discharges and precipitation. Outputs are WRL for each of the 21 10-day periods for each year spanning the April-to-October period. VC3(*t*) is first computed from daily discharge Q(*t*) every day *t*, WRL(*t*) is then deduced by comparing VC3(*t*) to the four regulatory thresholds and finally a unique representative WR level is assigned to each of the 21 10-day periods, as the median of WRL(*t*) observed or simulated within that 10-day period. To best match the whole monitoring process stated in most of the DMPs, a simple precipitation correction was applied ("Pcorr", in Fig. 5). It consists to give a 'no alert' when precipitation during the preceding 10 days exceeds 70% of inter-annual precipitation average, regardless of the WR simulation results. The WRL framework is applied to observed and simulated data of both discharge and precipitation. To assess the performance of the WRL model under current condition against stated WR decisions, the WRL model is run with observed daily discharges extracted from the HYDRO database (named "GR6J" in the text) and with daily discharges simulated by the rainfall-runoff model GR6J forced by the SAFRAN reanalysis (named

"GR6J" in the text). In the context of climate change the WRL model is run with daily discharges obtained with GR6J forced by one of the 1350 sets of perturbed precipitation, temperature and PET time series. In this later case the regulatory thresholds are calculated on the simulated discharge time series to limit the possible effect of bias in rainfall-runoff modeling.

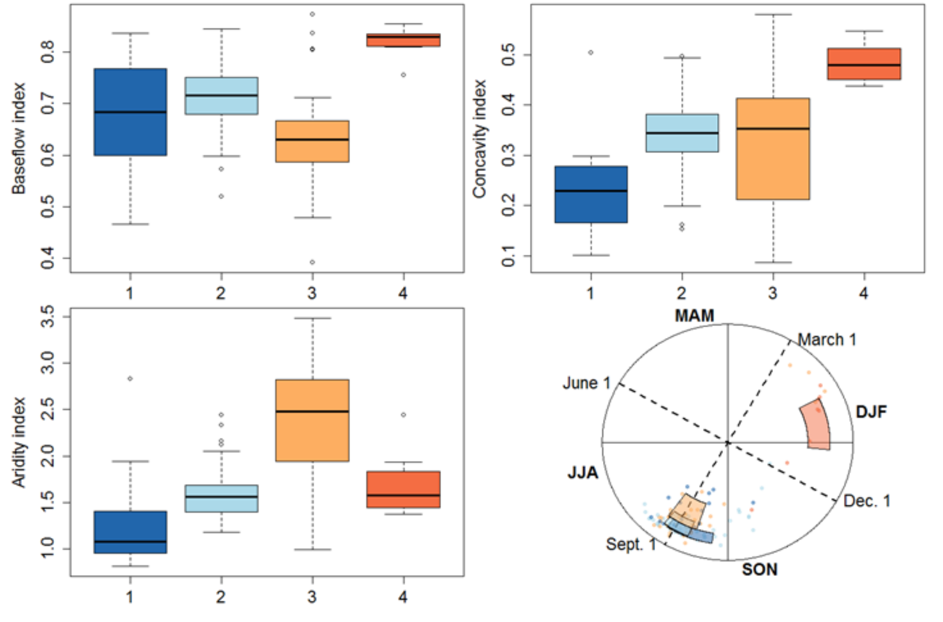
Line 801 - The results are weird here. If your GR67 model is good according to your NSE and Kling–Gupta efficiency, why GR67 and HYDRO show different results in a lot of place in this figure? Does not make sense.

➔ The GR6J model is not perfect (both criteria < 1). Small deviations to the observed discharge lead to difference in results obtained by the WR model.

Line 819 - If "2" and "3" are similar, why you need to separate them into two categories?

➔ This division into two classes have been suggested by the hierarchical clustering and the response surface representative of Class 2 is more contrasted than that of Class 3.

Line 822 - The figure at the lower-right corner is unreadable.



⇨ The authors would like to thank Reviewer1 for his helpful comments.