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

Sauquet et al.

Anonymous Referee #3

The objective of the study is to develop a risk-based framework to simulate water restrictions (WRs) under climate change in Rhone-Mediterranean district in order to evaluate the vulnerability of current Drought Management Plans (DMPs) to future climate conditions. The proposed framework is based on the assessment of three components: sensitivity of WRs to changes in different climate factors, sustainability of WRs for users and exposure in terms of climate response surfaces. General comments The paper presents an interesting topic. Although the applied methodology seems appropriate to some extent, it is rather unclear in some parts. Overall, I believe that further details should be added to the paper in order to support the interpretations and conclusions drawn from the analyses carried out by the authors.

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

Major comments

Section 3 For the sake of better understanding, I suggest to report the equations of low flow indicators and regulatory thresholds used in the manuscript.

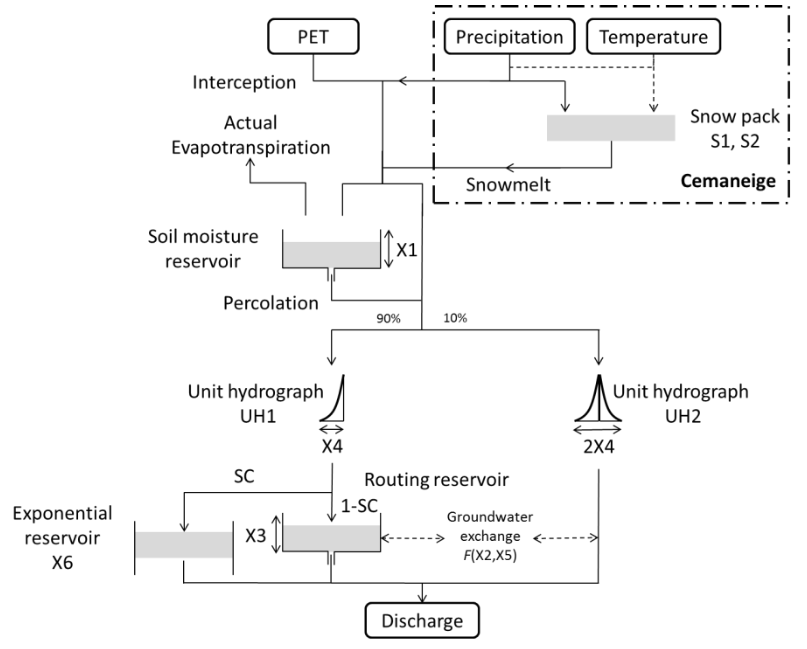
➔ Changes will be made in Section 3 to better define the variables of interest:

- "The low-flow monitoring indicators usually considered are: the daily discharge *Qdaily*, the *d*-day maximum discharge *QCd*, $QCd(t)=\max\left(Qdaily(t'), t' \in [t-d+1, t]\right)$ and the *d*-day mean discharge *VCd*, $VCd(t)=\int_{t-d+1}^{t} Qdaily(t')\,dt'$, with duration *d* associated with WR decision varying between 2 and 10 days depending on DMPs."
- "The threshold associated with WR also varies, generally associated with statistics derived from low-flow frequency analysis, but some being fixed to locally-defined ecological requirements. In the context of DMPs, series of minimum *QCd* or *VCd* are calculated by the block minima approach and thereafter fitted to the lognormal distribution. The block is not the year but the month or given by the division of the year into 10-day time-window. The regulatory thresholds are given by quantiles with four different recurrence intervals associated to the four restriction levels. For example, let us consider thresholds based on the annual monthly minima of *VCNd*. The block minima approach is carried out on the *N* years of records for each month *i*, i=1…,12 leading to twelves datasets $\{min\{VCNd(t), month(t)=i, year(t)=j\}$,

j=1,…,N}. The twelve fitted distribution allows the calculation of 48 values of thresholds (=12 months × 4 levels) with four T-year recurrence intervals. To enable comparison of results across all catchments, the same definitions for the monitoring variables and the regulatory thresholds have been adopted for all the catchments. *VC*3 was selected as the monitoring indicator and the regulatory thresholds are low flow quantiles 10*d*-*VCN*3 based on the minimum 3-day mean discharges extracted by the block minima approach considering the fixed 10-day time-windows spanning the year as blocks with return periods, as they are the most common single indicators used in the 28 DMPs of the RM district. Lastly return periods T of 2, 5, 10 and 20 years will be associated with the "vigilance", "alert", "reinforced alert" and "crisis" restriction levels, respectively, due to their prevalence in the DMPs"

Section 4.2 Details on the rainfall-runoff model should be added, with special reference to the way how the influence of reservoirs is taken into account.

➔ The GR6J model has six parameters to be fitted (see Figure below): the capacity of soil moisture reservoir (X1) and of the routing reservoir (X3), the time base of a unit hydrograph (X4), two parameters of the groundwater exchange function F (X2 and X5) and a coefficient for emptying exponential store (X6). GR6J is combined with the daily snow module Cemaneige. The catchment is divided into five altitudinal bands of equal area on which snowmelt and snow accumulation processes are represented. For each band, daily meteorological inputs – including solid fractions of precipitation - are extrapolated using elevation as covariate and the snow routine is calculated separately. Finally, its outputs are then aggregated at the catchment scale to feed GR6J. The two parameters of Cemaneige are: the parameter controlling snowpack inertia (X1) and the degree-day coefficient controlling snowmelt (X2). No routine to simulate water management (e.g. reservoir) was considered here since discharges of the 106 gauging stations are weakly altered by human actions or naturalized discharges.



Section 4.3 The description of the water restriction level modelling is unclear in some parts. For instance, I would expect that the comparison between simulated WRLs driven by GR6J data and historical WRs will provide a lower sensitivity score than the comparison with simulated WRLs driven by HYDRO data (considered as benchmark), but it's not (see Lines 287-290). Could it be a consequence of the fact that the model disregards socio-political aspects of the decision making-process?

➔ Inputs of the WRL model are daily discharges and precipitation. Outputs are WRL for each of the 21 10-day periods defined between the 1st April and the 31st October. 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.

➔ The discharges simulated by GR6J introduced in the WRL model lead to higher *Sensitivity* scores than those obtained with observed discharges extracted in the HYDRO database. The reasons for this unexpected result have been investigated. In particular we have compared the observed and simulated temporal variability in the time series VCN3. A "smoothing" effect in the GR6J simulations compared to observations was initially suspected. Finally no obvious difference in autocorrelation functions was found between observed and simulated time series. One reason could that the period of interest 2005-2013 – with for some basins only three years with stated water restrictions – may be too short to analyse accurately the relative performance of WRL obtained with OBS and with HYDRO, respectively.

The two scores gives a global insight on the performance of the WRL modelling framework and too much weight should not been given to the differences between scores. The developed WRL modelling framework leads to similar results (moderate performance in detecting stated legally-binding water restrictions during the period 2005-2013) with both data sources HYDRO and GR6J. The WRL modelling framework provides an overview of the on-going drought and the drought committees are partly free to account for this information, i.e. to state or to postpone water restrictions. The developed framework is a tool to predict water restrictions with no interference of lobbies, i.e. only based on the physical processes.

These sentences must be better explained:

- lines 295-296- "Furthermore, GR6J performance under low-flow conditions show no statistical link with its WRL modelling performance" ➔ "Furthermore, there is no significant link between the GR6J efficiency in simulating low flows (NSE_{LOG}) and the performance of the WRL (*Sensitivity* and *Specificity* scores), since the determination coefficients between NSE_{LOG} and *Sensitivity*, and between NSE_{LOG} and *Sensitivity* are lower than 7%."
- lines 300-301 "possible biases in rainfall-runoff modeling does not affect much the ability of the WR modeling framework to simulate correctly declared or not declared WRs" It seems that despite the difficulties of GR6J model in simulating low-flows accurately, the results of WRL modelling driven by GR6J data are good anyway. How do the authors explain that? ➔ The WRL framework is applied to observed and simulated discharge data available before 31st December 2013. 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 modelling. The possible reasons of comparable performance between GR6J and OBS is that the WRL framework is carried out using regulatory thresholds derived from GR6J outputs and that even if the discharge data are not exactly reproduced by GR6J, their ranking and their relative position to the regulatory thresholds is correctly reproduced.

Section 5.3 Vulnerability is computed against a critical threshold. The latter is defined as the difference between the number of WRs simulated by the WR GR6J modelling framework for 2011 and over the baseline period. On the other hand, the Vulnerability Index is computed as the proportion (frequency) of RCM-based simulations that fail above the critical threshold. It sounds like a frequency is compared to a number. I believe that this step must be described in details.

➔ Indeed there are two measures of vulnerability. Given one specific climate change projection, a catchment could be judged vulnerable if on average the critical threshold is exceeded. The Vulnerability Index is a proportion reflecting the fraction of RCM leading to critical situations on average. This index is introduced here to account for the uncertainty in climate projections in vulnerability assessment. It should be interpreted as conditional probability (risk) with respect to a set of possible future climates and only used as a relative measure to rank the regions, from the less to the most likely impacted regions.

For the same reason, it is not clear how the black dotted lines representing the critical threshold are drawn in Figures 10 and 14.

➔ The dotted black lines are isopleths connecting points of the response surface with $\Delta WR^*= \Delta WR^*(2011)$. Their location in the response surface depends on the shape of the response surface; this is why the dotted lines differ from one catchment to another in Figure 10, and later from one class to another in Figure 14.

Section 5.4 With regard to the hierarchical cluster analysis for catchment classification at regional scale, the authors should specify the catchment characteristics considered to investigate similarity through the Euclidean distance (see line 421-424). Details on the CART model and its implementation should be added.

➔ CART methods perform successive binary splittings of a given dataset according to decision variables. The algorithm identifies automatically through a set of "if-then" logical conditions the best possible predictors, starting from the most discriminating decision variable to the less important factors, to predict the membership to the one of the four groups. The optimal choices are fixed recursively by increasing the homogeneity within the two resulting clusters. At each step one of the clusters (node) is divided into two nonoverlapping parts.

The list of the potential decision variables by type is:

- Severity:
  - Flow exceeded 95% of the time (*Q*95);
  - Annual minimum 10-day daily mean low flows with a 5-year recurrence interval;
  - Annual maximum deficit below threshold *Q*95 exceeded 20% of time;
- Duration:
  - Annual maximum maximal duration of the continuous sequence of zero flow within the year, exceeded on average every five years (*D*80). Maximum duration of consecutive zero flows (*D*) are sampled by block maxima approach and *D*80 is defined as the empirical 80th percentile of cumulative distribution function of *D*;
  - Seasonal recession time scales (*DT* and *Drec*). This duration based on the hydrograph defined by the 1-day and 30-day moving average of the 365 long term mean daily discharges, *d*= 1,…, 365 (*Qd* and *Q*30*d*,

respectively). *Drec* is defined by the time lapse between the median $Qd50$ and the 90th quantile $Qd90$ of Qd on the falling limb of the hydrograph defined by $Q30d$ and $DT = \ln(Qd50/Qd90)/Drec$;

- Rate of change:
    - Ratio $Q95/Q50$;
    - Concavity index derived from flow duration curve $(Q10 - Q99)/(Q1 - Q99)$ (Sauquet and Catalogne, 2011). This descriptor is a dimensionless measure of the contrast between low-flow and high-flow regimes derived from quantiles of the Flow Duration Curve;
    - Baseflow index (*BFI*). *BFI* is a measure of the proportion of the baseflow component to the total river flow, calculated by the separation algorithm separation suggested by Lyne and Hollick (1979);
    - Class of river flow regime based on average monthly runoff pattern defined by Sauquet *et al*. (2008) (between 1 and 12)
    - Seasonality ratio (*SR*) $SR = Q95_{AMJJASON}/Q95_{DJFM}$ ($SR > 1$ for mountainous catchment);
- Frequency:
    - Proportion of years with at least one value $Q < Q95$;
- Timing:
    - Mean day of first occurrence of low flow $< Q95$;
    - Mean and dispersion of the occurrence of flows $< Q95$ within the year (θ and r, $r\sin(\theta)$ and $r\cos(\theta)$. These two variables are circular statistics. Each day i with zero flow is converted into an angular (ti) and represented by a unit vector with rectangular coordinates ($\cos(ti)$; $\sin(ti)$). The mean of the cosines and sines defines a representative vector. The value for θ is obtained by calculating the inverse tangent of the angle of the mean vector and the norm of the mean vector provides a measure of the regularity in the dates (a value close to one indicates a high concentration around θ while a value close to zero indicates no seasonality).

We will include this list in a table.

Technical comments

In Line 260, “VC3 is with 10d-VCN3(T) each day . . .”, something is missing.

➔ “VC3 is compared with 10d-VCN3(T) each day . . .”

In lines 273-274 “OBS WRLs are correctly reproduced by both GR6J and HYDRO simulations, but also can be consistent with OBS” (???). This sentence is rather misleading, I wonder if “OBS” at the beginning of the sentence could be a mistake and could be deleted.

➔ There is a problem with the phrasing on these lines. “Both GR6J and HYDRO simulations are globally consistent with observed WRLs (OBS). However GR6J and HYDRO results may differ from OBS (*e.g.* basins 9 to 11 in the Lozère department during the year 2005).”

In line 420: “. . . a classification (of what?) was conducted on to define typical response surfaces, . . .”. Please specify.

➔ “a classification of the 106 gauging stations based on the 1350 values of ΔWR* was conducted on to define typical response surfaces”

In line 482: “come catchment” to be replaced by “some catchments”.

➔ “are found for ~~come~~ some catchments”.

In line 540: replace “precipitations” with “precipitation”.

➔ “more accurately temperature and/or precipitation~~s~~”

Missing references:

Brekke et al., 2009: Brekke L.D., Maurer E.P., Anderson J.D., Dettinger M.D., Townsley E.S., Harrison A., and Pruitt T.: Assessing reservoir operations risk under climate change. Water Resour. Res., 45, W04411, doi:10.1029/2008WR006941, 2009.

Gupta et al., 2009: Gupta H. V., Kling H., Yilmaz K., and Martinez G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol., 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Kay et al., 2014: Kay A. L., Crooks S. M., and Reynard N. S.: Using response surfaces to estimate impacts of climate change on flood peaks: assessment of uncertainty. Hydrol. Process., 28, 5273–5287, https://doi.org/10.1002/hyp.10000, 2014.

Schlef et al., 2018: Schlef K.E., Steinschneider S., and Brown C.M.: Spatiotemporal Impacts of Climate and Demand on Water Supply in the Apalachicola-Chattahoochee-Flint Basin. J. Water Resour. Plann. Manage., 2018, 144(2): 05017020, 2018.

Weib, 2011: Weiß M.: Future water availability in selected European catchments: a probabilistic assessment of seasonal flows under the IPCC A1B emission scenario using response surfaces. Nat Hazards Earth Syst Sci 11:2163–2171, 2011.

⇨ The authors would like to thank Reviewer3 for his helpful comments.