Review of 'Progressive water deficits using multi-year droughts in Central-Southern Chile' by Alvarez-Garreton et al.

This paper tackles an interesting topic of multi-year droughts in Chile using a large-sample dataset CAMELS-CL. It first analyses observed streamflow and precipitation data to identify shifts in rainfall-runoff behaviour during the megadrought and then explores the role of hydrological memory in controlling drought propagation intensity using simulations from HBV.

Overall I enjoyed reading this paper – the figures are very well presented, it tackles an interesting topic and there are a range of analyses to support the conclusion. However, the paper needs to discuss in more detail the limitations of the study, greater clarity is needed in the methods section and the key messages of the paper could be better highlighted. Below are my suggestions and comments, I have a couple of general (more major) comments and then a series of minor comments.

General Comments

Limitations. I was missing a broader discussion of the limitations of the study in the discussion section. The authors mention the absence of physical factors in the analysis (L390-392) but the authors should also discuss

(1) *human impacts in these catchments* – I appreciate this study focuses on physical processes but humans can also play a large part in the intensification of drought propagation. Consequently, it is important to state how human-impacted these catchments are and whether human activities increased during the mega drought (particularly if groundwater abstractions increased because of deficits in streamflow).

(2) *evapotranspiration* – there is a brief analysis of ET in Section 4.4, but a limitation of this study is the absence of any detailed analyses on ET or temperature. Given their role for snow processes and intensification of drought, this is an important missing process from the analysis. The figure in the supplementary info showed high ET anomalies for the semi-arid catchments which had large RR shifts.

HBV Modelling. Given the richness of the CAMELS-CL dataset, I was a little surprised that the authors went down the route of extending the analysis with a model (as great as HBV is!) rather than further exploring the meteorological (e.g. ET), physical (e.g. soils, geology, land cover) and human impact characteristics of these catchments to explain the R-R shift in the megadrought. You could have used BFI from CAMELS-CL for Figure 6 and calculated baseflow contributions from the observed flow time series for the analysis in Figure 7. The authors need to:

(1) better clarify the contribution of the modelling to the paper

(2) better clarify where outputs from HBV are used in the analysis, particularly where *observed* streamflow or *modelled* streamflow is used throughout the text.

(3) comment more broadly on the uncertainties and limitations of HBV in the discussion

Classifications. There are lots of different classifications of catchments (snow-dominated/pluvial basins in Figure 5, semi-arid/temperate in Figure 8) which is very confusing for the reader. I would use one classification and keep it consistent throughout the paper.

Minor Comments (in order of appearance in paper)

Abstract L16. "mediated by groundwater flows" – not just groundwater flows but surely storage in snow packs too?

Section 2 L89-90. There are multiple rainfall and PET products available in CAMELS-CL, it would be helpful to specify exactly which hydrometeorological data products you used from the dataset.

Section 2 L92-97. The infilling of the flow time series is currently not clear in the paper and I have a number of questions related to this.

CAMELS-CL provides daily flow timeseries (as far as I am aware), which I then assume you aggregate up to monthly and then annual values for the rest of the analysis. So how do you define a month where there is no streamflow data – does this mean all days from that month are missing or a threshold of xx days? Do you calculate a sum of the daily flow in that month or take a mean (a mean would be less sensitive to missing data)?

Why did you select gauges with 15 years of data, is this really sufficient coverage of the period 1979-2018? What was the highest number of months that needed to be gap filled for a single station?

Section 2 L108 change to **"**from less than 100 mm to the north **to more than 3000 mm** in the .."

Figure 1. Given the focus on groundwater dynamics you need a map of geology in Figure 1 and a description of the geology of Chile in the study description.

Figure 1 Caption. You need to add the source of the precipitation and temperature value either into the figure caption or the text.

Section 3.1 L128. Was there a reason for choosing 8 years to calculate the mean flux?

Figure 3. The red dots are quite hard to see – it may be worth increasing the size of the red dot or perhaps trying a red cross instead?

Figure 6. I was a little confused how Figure 6 was created – is this an average from all the snow-dominated and pluvial basins? Are Q, BF and snowmelt derived from HBV for this plot?

Section 4.3. You used HBV to calculate the baseflow index. Exactly how did you calculate the baseflow index and how does this value of baseflow index differ to the baseflow index calculated from observed streamflow and provided in CAMELS-CL.

Figure 7. I found the overlapping bars in Figure 7 quite confusing to interpret – have you thought about an alternative way to visualise these results as currently Figure 7a and 7b are quite difficult to interpret. Also is this seasonal runoff and baseflow analysed over the whole time period (i.e. from 1979-2010)?

Section 4.4 L329. "Figure 8 presents drought propagation for 25 years with negative anomalies over the last four decades". Do you mean negative anomalies in P and/or Q deficits? Or negative anomalies in the difference between runoff and precipitation deficits? If it is the latter then not all those years have negative anomalies (i.e. there are some years where the red dot is above 0%).

Figure 9. The text in L358-362 is essentially just a description of the figure and there should be some analysis of the results of Figure 9 here (currently, the analysis of the results is confusingly mixed into the discussion). The figure caption also needs to be more informative i.e. it refers to panel a and b

but there are subplots a-d in the plot, not really clear what Q-Pred R-R is or the number of cases or what ‘1’, ‘2’, ‘3’ and ‘4’ relate to on the x-axis of Figure 9b and 9d.

Discussion L370. Do you mean baseflow contribution to runoff – rather than low flow?

Discussion L397. What do you mean by “higher resistance”?

Conclusions. There is a lot of different analyses in this paper but the key message and results tend to get a little lost. I suggest to significantly shorten the conclusions – there is a lot of repetition in this section and removing it would better highlight the core results.

Supplementary Material. Figure S1 needs a figure caption.