Response to Referee #1 Comments:

We thank the referee for their valuable comments, and here are our point-by-point responses. Our responses are in red, and note that RC = referee comments, and AC = author comments. However, before the point-by-point response for each reviewer, we would like to provide a General Response, which is identical for both reviewers, that should help to address some of the overarching and overlapping concerns that were brought up by both Reviewers 1 and 2.

AC: General Response

In reading through the comments from both reviewers, we saw that there was one main issue that weaved throughout, namely the need for a more thorough discussion of the study's contribution to drought feedback loop, along with the associated limitations and caveats. To address this, we added a more detailed discussion in the Conclusions section, as well as a new figure to the manuscript, now "Figure 9", below.

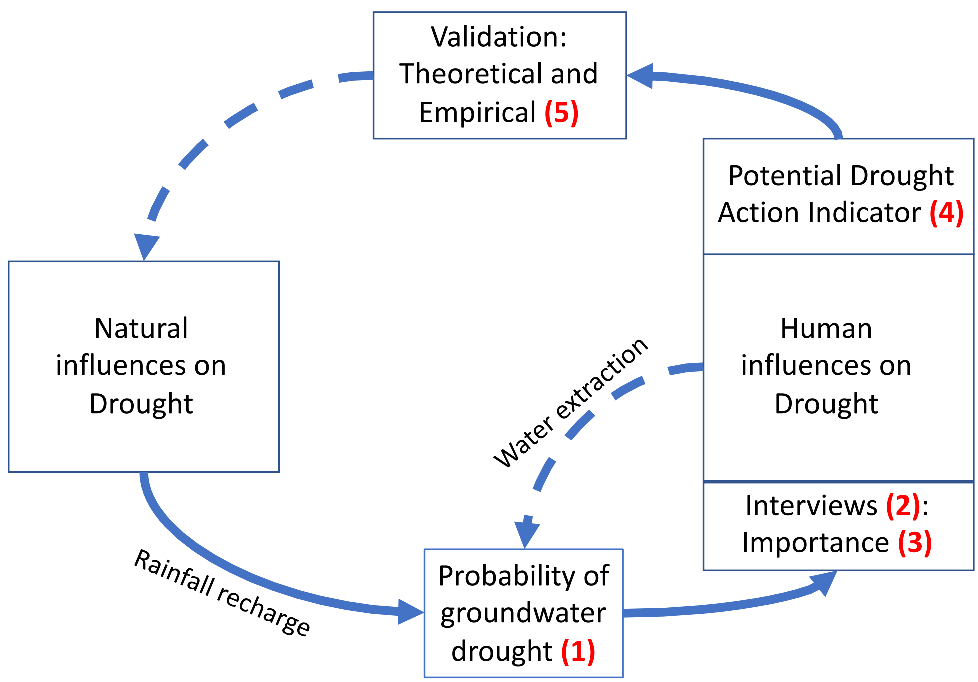

Figure 9. Conceptual map of drought feedback loop components addressed in this study (blue solid lines) and remaining gaps (blue dashed lines). Numbers correspond to discussion points in Conclusions.

Here we provide the excerpt that will be added to the Conclusions section of the manuscript to describe the new Figure 9 and the study limitations:

"We develop and demonstrate this methodology as a step towards closing the drought feedback loop, but note that there are caveats and limitations that warrant discussion. A conceptual overview of the contribution of our study to the drought feedback loop is shown in Figure 9, and we use this figure to identify five places where there is scope for future enhancements; each number below corresponds to a place in the drought feedback loop in Figure 9:

1) For the natural influence on drought, we examine the probability of groundwater drought. In our case, the groundwater levels are closely related to rainfall recharge, which is a function of natural climate variability. We recognize that this is not the case for many groundwater aquifers, where human activities, such as groundwater extraction, may trump the natural climate signal (e.g., Tarhule and Bergey 2006), often leading to water scarcity, rather than a natural phenomenon of temporary water deficiency. In many systems a full water balance would need to be examined to understand the relative contributions of extraction versus moisture deficit to the likelihood of going below a relevant hydrologic threshold.
2) Our interviews were conducted following a drought event, and we recognize that the timing of the interviews will likely affect the responses. For instance, interviews conducted during wet or average conditions might elicit less polarized responses, since drought impacts haven't been recently experienced. We note that our approach of applying the interview responses across different climate conditions (i.e., wettest to driest) makes the assumption that the importance of water uses and management preferences are stationary. We acknowledge that different climate conditions, as well as cultural change, technological innovation, climate adaptation, and other processes are likely to influence the cultural factors we investigated here and may mediate how people interact with their environments. Future work could investigate how responses change with different climate conditions over time, and the subsequent implications for drought action. However, hazards and disasters research is almost always conducted immediately after an event, so this is a wise-spread epistemological issue with both pros and cons in terms of what we learn from post-disaster research.
3) We use stakeholders' importance ratings as a proxy for their willingness to take action in relation to particular water uses, where by "action", we generally mean some effort towards drought mitigation. The interviews included questions about the importance of different water uses to test the application of the Cultural Theory of Risk (usually applied in a more global sense) to a specific water management issue, which had not been done before (Lazrus 2016). For the purposes in this paper, multiplying the importance ratings by the probability served as a way to make an objective characterization of drought subjective, that is, we wanted to modulate the groundwater drought probability by each individual stakeholder's lens.
4) The formulation of the PDAI strongly affects the conclusions drawn. Our formulation of the PDAI follows from other precedents in risk management that take the product of the likelihood of an event and its importance (Jones and Preston 2011; Oppenheimer et al. 2014). However, the functional form of the PDAI is flexible, allowing it to be tailored to other locations. As such, we note that the PDAI, as well as the best data to use to calculate it, will depend on the needs of the community, as well as the water system context.
5) We use social science theory to interpret our results, and to better understand the theoretical underpinnings of how and why people take action in response to drought. However, we note that empirical validation is important for indicator development and refinement. We recommend that future project designs include a validation component in the methodology. This could take the form of follow-up interviews, such as direct

feedback from stakeholders on if the indicator reflects their willingness to take action for certain water uses at certain drought levels. Methods, including stakeholder processes, for developing and evaluating drought indicator effectiveness have been put forth in the drought community (Steinemann and Hayes 2005 Steinemann and Cavalcanti 2006, Steinemann et al. 2015). Other options for validation can be indirect, such as looking at historical data, like government and local reports, media, and/or other collected response information, e.g., in the United States the US Drought Impact Reporter (Wilhite et al. 2007). Promising methods for mining social media, such as Twitter, have also been developed (Demuth et al. 2018) and could be adapted for evaluative purposes.

Although the methodology to develop the PDAI is experimental, we posit that explicit efforts to combine natural and human perspectives is critical to gaining a deeper and more nuanced understanding of drought feedbacks, and this paper provides a novel contribution to this end."

We will refer to this General Response in our point-by-point responses, below.

Interactive comment on "Characterizing the Potential for Drought Action from Combined Hydrological and Societal Perspectives" by Erin Towler et al.
A.F. Van Loon (Referee) a.f.vanloon@bham.ac.uk
Received and published: 10 May 2018

General comments:
RC0. This is an interesting and important paper that shows an approach to bring together natural and social drought processes. The authors propose an index that is composed of groundwater drought probability and stakeholders ratings of the importance of a water use, with the aim of showing the potential for drought action. I think that this is a very important step in drought research. The literature review and framing of the research are excellent. The results are good, but I do find the work quite "thin". The authors use groundwater levels from one well and interview results from a previous study, multiply these into an index, show only part of the results, and then draw conclusions, some of which are quite obvious. I do think that there is potential for the paper to be published and I give a few suggestions to improve the paper below.

AC0. We thank the referee for this constructive feedback, and address the specific suggestions in the responses, below. We also point out a few subtle changes based on feedback from Reviewer #2 to ensure readability of this response: (i) we are changing the "I" in PDAI from "Index" to "Indicator", and (ii) notate GW as "Z".

Specific comments:
RC1) Only the results for drinking water and recreation are shown in Figures 5-7. I would encourage the authors to find a way to show all results. I would suggest combining the plots of the empirical cumulative density functions into one 6-panel figure. I think the different lines will still be visible if you make the figures slightly smaller. Alternatively, the figures of the four remaining water uses can be placed in the appendix / supplementary material and be referred to for more information. Similarly, I suggest to include all decades in Figure 6. It is unclear why some decades have been left out.

AC1. The reviewer raises a good point about providing the ecdf results for all the decades, for all the water uses. To address this, first we update Figure 5 to be a two-panel plot with drinking water and recreation, including all the decades.

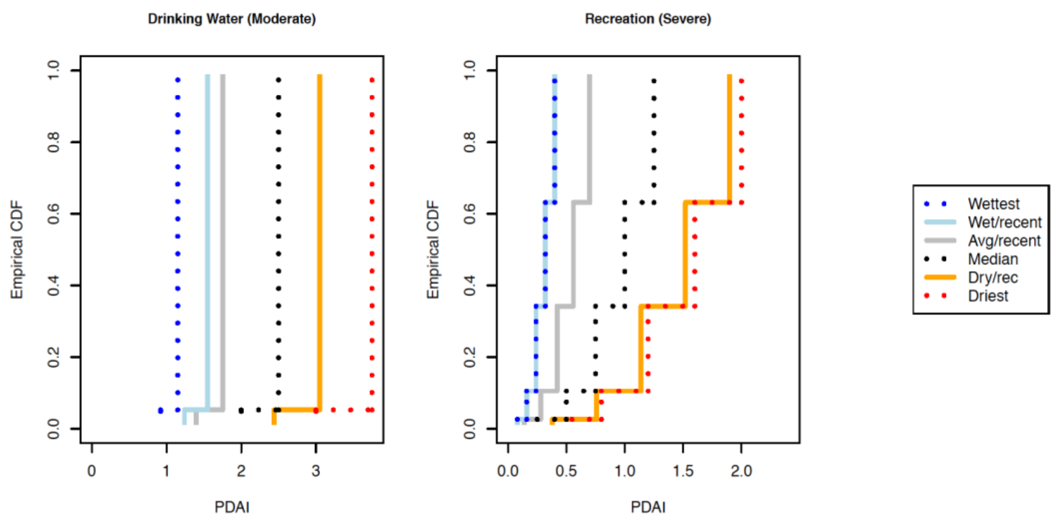

Fig. 5. Empirical cumulative density functions (eCDFs) for the PDAI (Potential Drought Action Indicator) for drinking water using the moderate threshold (left) and for recreation using the severe threshold (right), for their respective wettest, wet/recent, average/recent, dry/recent, and driest historical decades.

Second, we include a 6-panel plot (one for each water use) to the supplemental information. The Supplemental figure (Figure S1) is included below. As a result, the manuscript text has been updated to say (changes in **bold**):

"To demonstrate the PDAI, we examine two different water uses: drinking water and recreation, **although results for all of the water uses can be found in the Supplemental information**."

First, we focus on drinking water, which is an example of a water use which exhibited more consensus among interviewees. For drinking water, to calculate the PDAI, we use the moderate threshold, since this is the threshold at which municipal supply is monitored (see Section 3.1). Figure 5 (**left**) shows the PDAI for drinking water for the different drought conditions (e.g., wet/recent, dry/recent, etc) from Table 1. Results are shown as empirical Cumulative Density Functions (eCDFs) to reflect the discrete nature of the importance ratings. In the eCDFs, the vertical lines represent the PDAI values, and the horizontal lines represent the percentage of data that are equal or less than that value. In Figure 5 (**left**), as the eCDF moves across drought conditions from very wet to very dry, the PDAI shifts towards higher values, reflecting the increased potential for action under drier conditions. Specifically, the very wet decade has an average PDAI value of 1.1, and the very dry decade has an average PDAI value of 3.7. Given the stakeholder consensus on the importance for drinking water, for each drought condition there is very little range – that is, the eCDFs are fairly vertical. **Results are similar when the Moderate threshold is used for the other two water uses, habitat and livelihood, that showed strong consensus (see Figure S1).**

Next, we focus on the PDAI for Recreation, a water use that shows diverse importance ratings from stakeholders (Figure 5, right). For recreation, to calculate the PDAI, we use the severe threshold, since that is the threshold at which artesian springs no longer flow (see Section 3.1). **Figure 5 (right) shows the PDAI for recreation for the select decadal drought conditions, using the severe threshold likelihoods from Table 1**. Similar to drinking water, we see that as we move from wetter to drier, the PDAI also increases; for example, from wet/recent to dry/recent, the average PDAI values are 0.3 and 1.5, respectively. However, given the stakeholder diversity in importance ratings, as we move towards drier conditions, the PDAI becomes more diffuse, spanning a great range of values: in the wet/recent, the PDAI spans from .08 to .4, or for 0.32 units of the PDAI scale, and in the dry/recent it spans from 0.4 to 1.9, or 1.5 units on the PDAI scale, indicating a wide range in stakeholder appetite for potential action. **Interestingly, the wet/recent decade (1993-2002) was also the wettest decade on record, with the groundwater threshold only being exceeded 8% of the time. Results are similar when the Severe threshold is used for the other two water uses that showed diverse ratings, i.e., cultural practices and spiritual fulfillment (see Figure S1).**

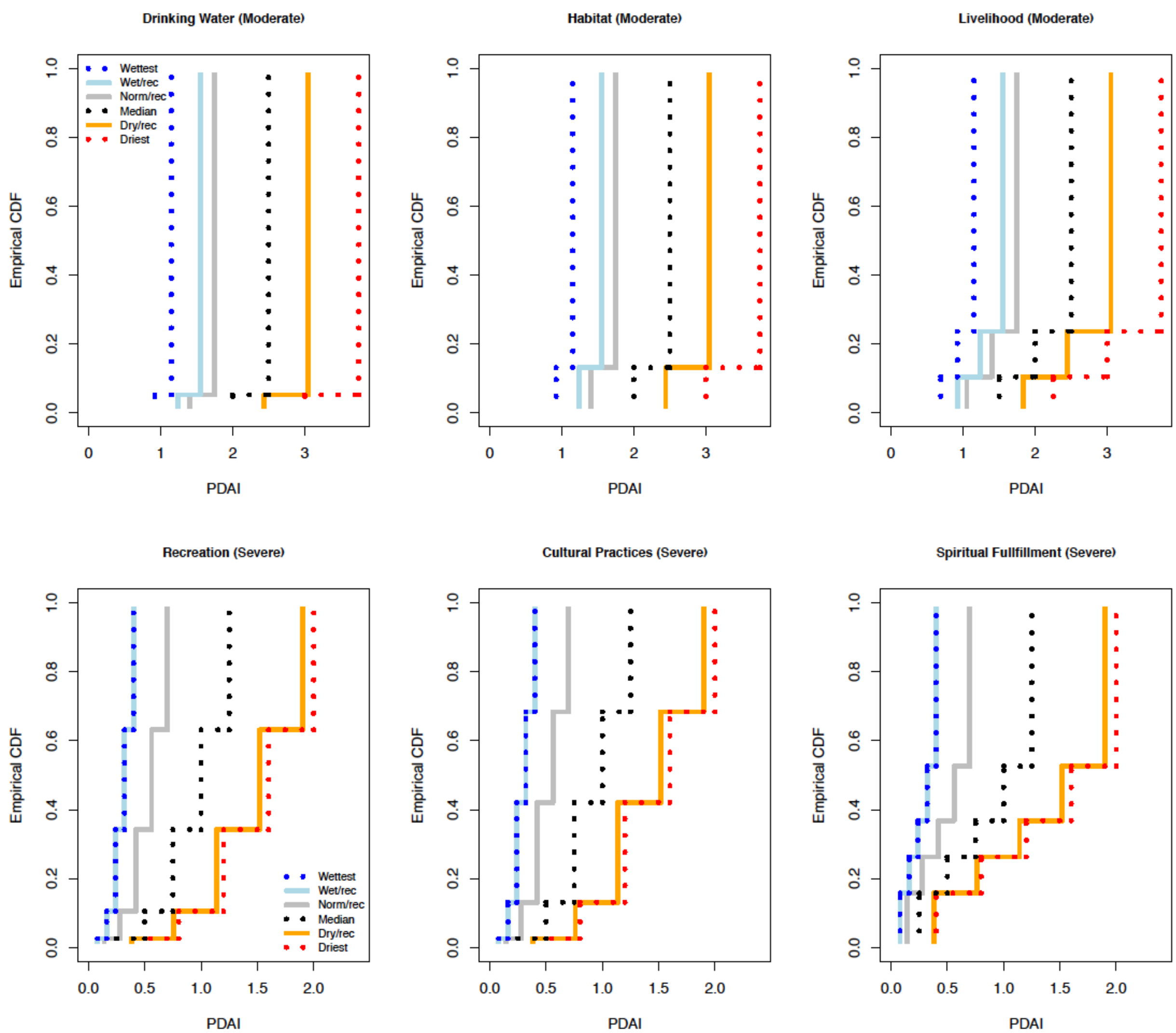

Figure S1. Empirical cumulative density functions (eCDFs) for the PDAI (Potential Drought Action Indicator) for drinking water, habitat, and livelihood using the moderate threshold (top row) and for recreation, cultural practices, and spiritual fulfillment *using the severe threshold (bottom row) under the wettest, wet/recent, normal/recent, median, dry/recent, and driest historical decades (see Table 2 for corresponding years).*

In creating this figure, we recalled that the decades listed in Table 2 for the "very wet", "median", and "very dry" were all based on the decades of wettest, median, and driest exceedance likelihoods (%s) that corresponded to the *Moderate* Threshold. These did correspond to very wet, normal, and very dry likelihoods for the Severe threshold, but they were not the "wettest", "median", and "driest" likelihoods for the Severe threshold. So, in creating plots that include all the decades for the ecdfs for the water uses that use the Severe threshold (i.e., Recreation, Cultural Practices, and Spiritual Fulfillment), it is more illustrative to show the wettest and driest decades for the severe threshold. Hence, we have updated Table 2, which results in no changes to the P(Z<Mod) column, but slight changes to the P(Z<Sev) column. Specifically, P(Z<Sev) for "very dry" was 38% in 1959-1968 – where 1959-68 was the driest category for the Moderate threshold – but the *driest* decade for the Severe threshold was actually 40%, which occurred in 1964-1973 and 1972-1981. The table has been updated to this effect. Similarly, the "very wet" decade P(Z<Sev) was 13% for 1985-1994, but the wettest decade was 8% in 1993-2002, which is the same as the wet/recent. These changes do not affect the conclusions drawn. We have now updated the table Comment column and figure legends to say "wettest" and "driest" rather than "very wet" and "very dry". These changes are easiest seen by looking at both tables, which are included below, where changes from the old Table 2 to the updated Table 2 are shown in Red here:

Table 2 (UPDATED). Decadal Likelihood (P) of Groundwater (Z) Level Going Below Moderate (Mod) and Severe (Sev) Thresholds for Recent Decades, as Well as Driest, Median, and Wettest Decades.

P(Z<Mod) (%)	P(Z <Sev) (%)	Comment	Decade
61	38	Dry/recent decade; most recent decade to interviews	2003-2012
35	14	Average/recent decade; third most recent decade	1983-1992
31	8	Wet/recent decade; second most recent decade	1993-2002
75	40	Driest decade; highest exceedance likelihood	1959-1968(Mod); 1964-1973(Sev); 1972-1981(Sev)
50	25	Median decade; median exceedance likelihood	1999-2008(Mod); 1979-1988(Mod); 1969-1978(Sev); 1980-1989(Sev)
23	8	Wettest decade; lowest exceedance likelihood	1985-1994(Mod); 1993-2002(Sev)

Table 2 (OLD). Decadal Likelihood (P) of Groundwater (GW) Level Going Below Moderate (Mod) and Severe (Sev) Thresholds for Recent Decades, as Well as Very Dry, Median, and Very Wet Decades.

Decade	P(GW<Mod) (%)	P(GW<Sev) (%)	Comment
2003-2012	61	38	Dry/recent decade; most recent decade to interviews
1983-1992	35	14	Average/recent decade; third most recent decade
1993-2002	31	8	Wet/recent decade; second most recent decade
1959-1968	75	38	Very dry decade; highest exceedance likelihood
1999-2008	50	24	Median decade; median exceedance likelihood
1985-1994	23	13	Very wet decade; lowest exceedance likelihood

Figure 6 remains the same, but now the text focuses on the concept of the "new normal", and the text has been updated: "In Figure 6, we also looked at recreation under the possibility of a new "normal" drought baseline (Van Loon 2016b)." We further develop this, as can be seen in our response to a later reviewer comment, see RC4 and AC4, below. Figure 7 remains unchanged, as its purpose was focus on the most recent decade (2003-2012) with drinking water and recreation to illustrate the effect of the different threshold.

RC2) I think some form of validation is needed. Either by checking back with the stakeholders whether they agree that the index shows more willingness to take action for certain water uses at certain drought levels, or by using historic information. Are you completely sure that there is no information on drought measures being taken in this area (or a comparable area)? Even not for the most recent drought? Did you consider looking at government reports, (social) media or the US Drought Impact Reporter (which includes lots of information on responses as well)? If you do a bit of this, it gives more backing to the statements like on l. 385-386 that the potential for drought action is diverse because the water use values are diverse (this is obvious because that is what you have put in the equation).

AC2. We agree that empirical validation is an important component, and is a limitation of our current study. Unfortunately, at this point we do not have the resources to check back with stakeholders about the indicator nor to thoroughly investigate historical information. To this point, we tried to be as up-front as we could in the paper, noting in the Introduction (changes in bold): “We use the term “potential”, since in this study, we do not have the data to validate whether or not human actions were actually taken as a result of these natural drought influences…” and “Though we do not directly validate the PDAI, we are able to interpret the findings to provide insights to water management policy using additional interview data on stakeholder worldviews and social science theory; **this is unique in that it allows us to investigate the theoretical underpinnings that are not typically explored in drought risk studies or stakeholder processes**.” However, we do recognize that we should add more on this point in the manuscript limitations, and further, we appreciate your suggestions of historical information that could be examined, and include them in our discussion. As noted in the General Response, we add Figure 9 and point 5 in the Conclusions (see “5” in General Response, above).

“5. We use social science theory to interpret our results, and to better understand the theoretical underpinnings of how and why people take action in response to drought. However, we note that empirical validation is important for indicator development and refinement. We recommend that future project designs include a validation component in the methodology. This could take the form of follow-up interviews, such as direct feedback from stakeholders on if the indicator reflects their willingness to take action for certain water uses at certain drought levels. Methods, including stakeholder processes, for developing and evaluating drought indicator effectiveness have been put forth in the drought community (Steinemann and Hayes 2005 Steinemann and Cavalcanti 2006, Steinemann et al. 2015). Other options for validation can be indirect, such as looking at historical data, like government and local reports, media, and/or other collected response information, e.g., in the United States the US Drought Impact Reporter (Wilhite et al. 2007). Promising methods for mining social media, such as Twitter, have also been developed (Demuth et al. 2018) and could be adapted for evaluative purposes

RC3) The interviews were done just after a drought event, which might have influenced the outcomes. Especially since, according to the social memory concept, people might be very aware and willing to save water during and just after a drought, but this awareness and willingness might fade over time when conditions return to normal. I understand that it would take a lot of time to go back into the field and redo the interviews in wet and normal years, but this issue should at least be mentioned in the discussion.”

AC3) We agree that the timing of the interviews will influence the outcomes. Here, our purpose was to investigate stakeholder responses after experiencing a drought – which is likely to elicit the most polarized response as compared to non-drought times – and then took the approach of applying that snapshot across different climate conditions. As noted in the General Response, we add Figure 9 and point 2 in the Conclusions (see "2" in General Response, above).:

"2. Our interviews were conducted following a drought event, and we recognize that the timing of the interviews will likely affect the responses. For instance, interviews conducted during wet or average conditions might elicit less polarized responses, since drought impacts haven't been recently experienced. We note that our approach of applying the interview responses across different climate conditions (i.e., wettest to driest) makes the assumption that the importance of water uses and management preferences are stationary. We acknowledge that different climate conditions, as well as cultural change, technological innovation, climate adaptation, and other processes are likely to influence the cultural factors we investigated here and may mediate how people interact with their environments. Future work could investigate how responses change with different climate conditions over time, and the subsequent implications for drought action. However, hazards and disasters research is almost always conducted immediately after an event, so this is a wise-spread epistemological issue with both pros and cons in terms of what we learn from post-disaster research."

RC4) The test of the "new normal" is interesting, but not well developed. Why is it only done for recreation? It would be also interesting for the other water uses, and especially for drinking water. How is the value of 120 feet chosen? Is that a relevant level, because if springs and streams have dried up at 117 feet already, what would be the difference of a level of 120 feet to recreation uses? The paper would be much stronger if you would have a rationale why groundwater levels would go down that much in the future or why different water uses are adapting to lower water availability.

AC4) We wanted to illustrate this concept, but agree that we should provide more rationale for the 120 feet level, which was only available for recreation. This is based on a previous analysis (Towler and Lazrus 2016), the figure I include here for the benefit of the reviewer:

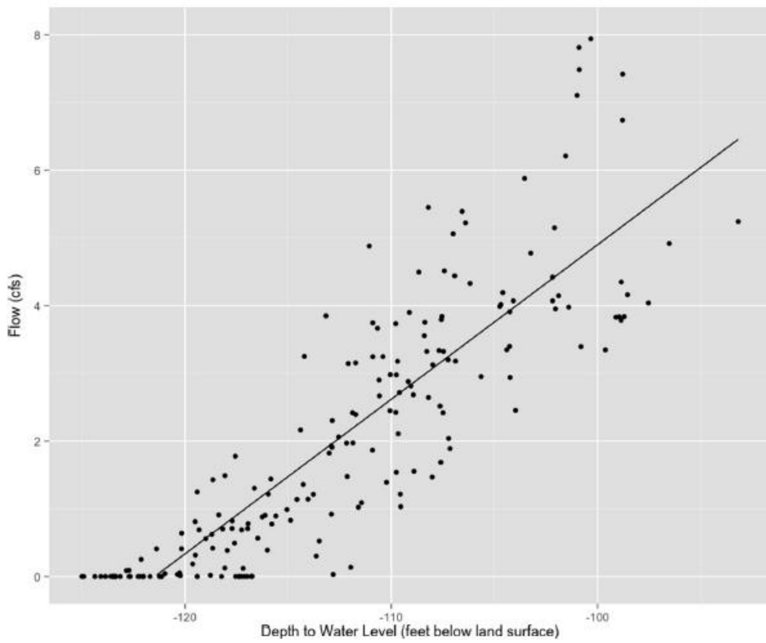

Fig. 2. Scatterplot of monthly Fittstown groundwater depths and Antelope Springs flows ($r = 0.87$).

The x-axis is the same groundwater level data used here, and this is related to spring flows in a Recreation area. From the graph, between -117 and -120 feet below the surface, there is variability in the springs flow, but the linear fit shows a monotonic decrease between the two levels. To clarify this in the paper: "The second threshold is the "severe" threshold: this is when the groundwater level lowers further, to 117 feet below the surface, which is the level at which artesian springs in the area have **minimal flow or stop flowing altogether**, affecting uses such as wildlife and recreation. **For illustrative purposes, we also look at an "extreme" threshold of groundwater levels to 120 feet below the surface, which have been experienced in the aquifer and further the likelihood of minimal or stopped spring flows (see Figure 2 in Towler and Lazrus 2016**)."

RC5) Textual comments: - There is quite a lot of repetition in the paper. For example the sentence that this study combines hydrological and social perspectives of drought comes back a few times. Maybe you do not need to mention it again in the first paragraph of the Methods? Also, the calculation of the PDAI is mentioned in the Methods and the Results (l.326-328). And the results of the differences in PDAI range are mentioned on lines 342-343, 353- 356, and 380-385. It would be good if this could minimised.

AC5) Thanks for pointing these out. We have removed mention of the hydrological and societal perspectives in the first paragraph of the methods, which now reads: "Figure 1 provides the conceptual overview of the study methodology, which is detailed in the subsequent sections." Similarly, we remove the calculation of the PDAI in the Results (formerly l.326-328 that you pointed out), so now it reads: "To demonstrate the PDAI, we examine two different water uses: drinking water and recreation, although results for all of the water uses can be found in the Supplemental information." We opted to leave the specific values of the PDAI range results the first time they are mentioned, but do remove one repetition of the specific range numbers at former lines 380-385, which now reads: “Another key point from Figure 7 is that drinking water spans a smaller range on the PDAI scale than recreation, which is more diffuse.”

RC6) - L.106: can’t > cannot

AC6) This has been changed.

RC7) - L.124: Climate Division 8 > please explain or give a reference for readers who do not know what Climate Divisions are

AC7) We have added the reference (Karl and Koss 1984):

Karl, T. R., and Koss, W.J.: *Regional and National Monthly, Seasonal, and Annual Temperature Weighted by Area, 1895-1983.* Historical Climatology Series 4-3, National Climatic Data Center, Asheville, NC, 38 pp, 1984.

RC8)- L.180: is there a word missing here?

AC8) We have fixed this sentence to read (change in bold): We use data from the beginning of the GW monitoring record through the year the interviews were conducted**, which corresponds to** 1959-2012.

RC9) - There are a lot of references to Figures and Tables in the Conclusions. It would be better is if the Conclusions could be read on a more standalone basis.

AC9) We have removed the references to Figures and Tables in the conclusions.

RC10) More discussion is needed on the limitations of this study. There is a bit now in the last paragraph of the Conclusions, but this could be developed more.

AC10) Based on your feedback, we explicitly address several limitations, outlined in the General Response, above.

RC11)- Figure 2 & 3: please make the axis labels and legend text a bit bigger (maybe also for Figures 5-7)

AC11) Thank you for this input, we have made a note and will do this when we submit the final revised manuscript.