*****bold indicates our response.**

Reviewer 3

<u>Our general response</u>

The original manuscript submission did a poor job delivering a number of messages and providing additional details. We have tried to address these in an updated version of the manuscript. The main comments raised by all three reviewers can be summarized in the following three major concerns:

1. **The manuscript lacks a strong motivation why one would be interested to focus on drought and wet spells within a semi-arid environment that is predominantly dry throughout the year.**
2. **A good description of the newly presented combined identification method including a logical figure.**
3. **An incorrect focus where ephemeral rivers were compared with perennial rivers, without providing the necessary relationships.**

We have tried to address these major issues in a new version of the manuscript and we believe that this has strongly improved the manuscript readability and clarifies the above mentioned topics. This original submission was poorly presented and we would like to thank the reviewers for identifying this. However, we also believe that the updated version provides considerable interesting insights into the occurrence and variability of semi-arid hydrological anomalies that strongly reflect the ecohydrological functioning of the channel bed and riparian zone, which we believe are of interest to readers of HESS.

<u>Major concern 1:</u>

To the introduction we have added the following:

"This combined procedure was able to identify hydrological anomalies specifically within transitional regions, where zero flow conditions are common but not the standard. However, as the current work will show, this method has difficulty correctly identifying drought and wet spell occurrences within ephemeral rivers within semi-arid regions with a strong seasonal precipitation signal (e.g. Monsoon). The strength and occurrence of seasonal runoff within these rivers strongly impacts groundwater recharge, and the ecohydrological state of the channel bed and riparian zone (Goodrich et al., 2004; Scott et al., 2008). As a result, these reaches become hotspots for biodiversity, especially during the dry season (e.g. Moreno-de las Heras et al., 2012; Cleverly et al., 2016). Correct identification of a hydrological anomaly is therefore important. As a result, this paper presents an updated version of the combined approach originally developed by Van Huijgevoort et al. (2012), improving the identification and continuation of a drought or wet spell, when transitioning from the wet into the dry season. This newly combined approach provides information on the hydrological status of the river including the occurrence of anamolies. For semi-arid ephemeral channels this information serves as a proxy for the moisture state of the channel bed and riparian zone, and its ecohydrological functioning (Scott et al., 2008; Moreno de las Heras et al., 2012). Specifically, this paper focuses on the semi-arid southwest US, where considerable amounts of long-term observations are available, which allows for a detailed analysis."

To Section 2.1 we have added the following:

"Within the San Pedro, upslope headwaters have low permeability bedrock relatively close to the surface, resulting in infiltrating rainfall and snow melt quickly to reach the groundwater system and river network (Fan et al., 2007; Kampf et al., 2016). These perennial streams become ephemeral downslope through transmission losses from evaporation and infiltration into the dryer ambient subsurface (Cataldo et al., 2010; Blasch et al., 2013). Within the semi-arid San Pedro basin, stream infiltration has been shown to account for 10-40% of total groundwater recharge (Goodrich et al., 2004). However, for the majority of precipitation generating flow events the depth of infiltration is relatively shallow, with only the biggest events infiltrating deeper (>1 meter below the channel bed). Transmission losses therefore strongly impact root zone moisture availability. Lower elevation riverbeds and their

surrounding areas are therefore favorable for biodiversity (e.g. Moreno-de las Heras et al., 2012; Cleverly et al., 2016).”

To Section 3.2 we have added the following:

“As indicated in the introduction, flow events replenish rootzone moisture below the channel bed, which is used by the riparian zone vegetation during the dry season for transpiration. Therefore, the occurrence of a drought or wet spell during the positive flow season, directly impacts moisture availability afterwards. As such, it is important that a given hydrological state continuous from the wet into the dry season. For a hydrological drought to prolong from the wet into the dry season, drought should also be identified at lower consecutive drought/dry day numbers, which currently is impossible.”

To the discussion we added the following:

“It should be noted that for situations where there is generally no flow, a small flow event can result in a sudden increase of the discharge percentile, as shown in Fig. 6 for the early summer of 2007. This can result in a wet spell anomaly, which due to a 30-day MA window size, can last for 30 days even though the total flow amount during this period was very small. As the channel flow ceases, the discharge percentile becomes more indicative for the moisture state of the channel bed, which can be expected to be relatively wet due to the strong impact of transmission losses.”

Where we felt needed, we refer to these sections throughout the rest of the manuscript. By providing these details we believe the motivation and focus behind this work become clear to the reader.

<u>Major concern 2:</u>

To address this point, in Section 3.2 we have tried to identify the current limitation of the original combined method of Van Huijgevoort et al. (2012). Furthermore, in our original submission we did not highlight that this method was only developed for the identification of drought. Also, the updated version provides now a link to point 1 raised above, to highlight that a given hydrological state as well as the occurrent of a drought and wet spell, here reflects that available root zone moisture.

Therefore, Section 3.2 was completely rewritten. The first part now contains the following:

“Even though the combined method proposed by Van Huijgevoort et al. (2012) has been a major advancement to identify hydrological anomalies within transition regions (e.g. Breyer et al., 2018; Heinke et al., 2019), the method has three potential issues. First, the final discharge percentile distribution for a given day of the year does not have to be perfectly uniform between 0–100. This is the result of combining the original TLM, where discharge percentiles are obtained for a given day, with the CDPM, which defines the cumulative density function (cdf) on the bases of all dry days. Even though consecutive drought/dry numbers are rescaled (see step 3 of Section 3.1), this does not guarantee a perfect uniform distribution between 0 and 100 for a given day.
Second, during zero flow situations, drought only occurs in case a given day has a high consecutive drought/dry day number. For the San Pedro these occur in spring at the end of the dry season. Therefore, in case a drought starts during the positive flow season in summer, it will generally stop identifying the drought once flow ceases. If zero flow conditions continue to occur, the combined method will then identify a drought again later in the dry season in spring. As such, the original combined method therefore identifies two droughts (i.e. one in summer during the NAM and one in spring after a prolonged period without flow, see also Section 4.1). As indicated in the introduction, flow events replenish rootzone moisture below the channel bed, which is used by the riparian zone vegetation during the dry season for transpiration. Therefore, the occurrence of a drought or wet spell during the positive flow season, directly impacts moisture availability afterwards. As such, it is important that a given

hydrological state continuous from the wet into the dry season. In order for a hydrological drought to prolong from the wet into the dry season, drought should also be identified at lower consecutive drought/dry day numbers, which currently is impossible.
Third, originally the method was not developed to identify wet spells for zero flow conditions, as these would always occur during situations of low consecutive drought/dry day number. For the San Pedro, this would result in a wet spell to occur immediately at the start of the dry season, even if a drought was observed during the positive flow season (see also Section 4.1)."

We also extended the explanation of the new combined procedure presented, following an approach similar to the one presented in the original work of Van Huijgevoort et al (2012) making use of both mathematical symbols and enumerated lists. Here we use a figure to indicate the different steps.

Furthermore, in line with suggestion raised by the reviewer, we moved the figure that showed the difference between the old and new combined method to a newly created Section 4.1, which contains the following:
"4.1 Comparison of the original and modified combined method
Figure 5 shows the difference in identified discharge percentile between the combined method as proposed by Van Huijgevoort et al. (2012) (old method) and the modified combined method proposed here (new method). As indicated in Section 3.2 the old method was unable to identify drought during zero flow conditions with a low consecutive drought/dry day number. This can clearly be observed from Fig. 5b where these days, that generally occur at the beginning of the dry season during the fall, show high corresponding discharge percentiles. Furthermore, for the years 2004–2005 and 2009–2010, the drought observed during the positive flow season, first ceases and return the following spring at the end of the dry season for high consecutive drought/dry day number. This situation does not occur for the new method presented here, where the consecutive drought/dry day number of a given zero flow day is compared with its number observed during other years. As such, discharge percentiles in the fall can be small enough to indicate a drought. As a result, for both 2004–2005 and 2009–2010 the drought continued from the wet season into the fall, as an indication of the moisture state of the riparian zone. The fact that the observed hydrological drought continues after precipitation ceases results in a considerable increase in the total number of drought days for 2002–2005 and 2009–2011 (Fig. 5c).
Section 3.2 also mentioned that, the original method of Van Huijgevoort et al. (2012) would lead to high discharge percentiles at low consecutive drought/dry day number and as such was not specifically developed for this. For the fall, Fig. 5b shows for station 11 this would indicate the occurrence of a wet spell during the fall of almost each year, irrespective of the strength of the NAM. The newly combined method presented here, does not show this strong dependency. In stead, high discharge percentiles in the fall only occur during years where already during the NAM a wet spell was observed (i.e. 2000–2001 and 2006–2007). During the other years discharge percentiles in the fall do not indicate the occurrence of a wet spell."

We believe these changes have improved the description of the new combined procedure presented here and its applicability to identify hydrological anomalies within semi-arid ephemeral channels.

<u>Major concerning 3:</u>
As the current work focuses on hydrological anomalies within semi-arid environments and more specifically the San Pedro basin, we feel that only focusing on the ephemeral channels as suggested by reviewer 2 would not provide a complete picture of hydrological anomalies occurring throughout this region. Instead we assess their occurrence of three typical cases: 1) dry lands channel beds that only show runoff from intense precipitation events during the NAM and from remnants from hurricanes in the fall (NB), 2) the upland regions that have perennially flow conditions (BU), and 3) low land regions with perennially flowing conditions received from baseflow upslope (BD). To introduce this we added the following to the introduction:

"Results obtained for ephemeral channels as derived using the new combined procedure presented here, will be compared with observations from perennial rivers at higher elevation

as well as for downslope locations receiving continuous flow from upslope. By analyzing the occurrence of drought and wet spell anomalies across these locations, this work will provide a detailed overview of their occurrence and variability within the San Pedro, as well as the role of climate and local geographical location."

For case a) dry lands channels, the occurrence of a single event can generate a wet spell even after runoff ceases. This might be counterintuitive, as indicated by all three reviewers, as "How can zero flow conditions correspond to a wet spell". We hope that the comments provided by major point 1 and 2 as given above, have resolved this aspect.

Reviewer 2 indicated that one option would be to only focus on case a). As we did a poor job describing the different system, we understand this suggestion. However, we hope by correctly adding these details throughout the paper, it becomes clear why it is important also to address the uplands.

For case b) the hydrological response of these the upslope regions is much more similar to temperate environment. Hydrological drought and wet spells have a direct link to the amount of water available in the river network. Furthermore, as flow conditions are always positive, we added the following to Section 3.2:

"for locations within continuously flowing condition (BU and BD category) the TLM approach was used solely."

For case c), at shorter time scales (30-day moving average window), observed hydrological drought and wet spell characteristics resemble those observed for the upland reaches, since during the dry season all water originates from these uplands. However, also in these environments, transmission losses form an important source of moisture for the surrounding riparian zone. Since these channels transport the majority of their water during the NAM, when focusing on a longer timescale (e.g. MA-window of 1 year), their hydrological drought and wet spell characteristics resemble more case a) the NB category. As such, at these timescales a hydrological anomaly is more representative for the state of the riparian zone. Again, the paper did a poor job in explaining this.

Because of this difference in behavior across MA-window scales, we felt it was important to assess this. Therefore, we presented two figures for 30-day MA windows (Figs 6-8), various MA-window sizes (Fig. 10) and a one-year MA-window size (Figs. 9 and 11). However, we did not properly motivate these choices and have tried to include this in the updated version of the paper, by addressing these in both the Results section as well in within the Discussion. The latter now states:
"The mean duration of a wet spell last longer for upslope domains (BU) as compared to the lower elevation categories (BD and NB). For the upslope regions, shallow groundwater is expected to have a stronger control on the observed amount of baseflow for a longer period of time, increasing the mean duration (see also Section 4.2)."

Besides addressing these major concerns we have:

- **Changed the title as suggested by the reviewers**
- **Include more up to date references in introduction and section 3**
- **We propose to alter the setup of the manuscript by creating a new paragraph 4.1, before the original sections 4.1-4-3 which will become section 4.2-4.4.**
- **For the one-year MA-window analyses in Figures 11 and 13, we have added the details of why we feel this is interesting both in the Introduction as well as in Section 3.4 and in the Discussion. We will also highlight that this situation effectively corresponds to applying the TLM only.**
- **To decrease the length of the discussion and not to present new results we have moved the figures of the discussion of the original manuscript to a newly created section 4.5.**
- **Adjust the original figures where suggested.**

We are convinced that by resolving these concerns and addressing the various issues raised by the reviewers the quality of the manuscript has improved consirably.

This manuscript deals with a new methodology for assessing hydrological drought with a variable threshold approach in semi-arid catchments and its application on case study catchments in Arizona.

We would like to thank reviewer 3 for the constructive comments. We realize that reviewer 3 was very critical with respect to the quality of the original submission. Based on the suggestion raised by the reviewer we believe we have improved the focus and quality of this work considerably.

Main comments
The comments will clearly overlap with those of the other referees, emphasizing the main issues found in the manuscript. But before that, it has to be noted that the study is reasonably well conducted and the manuscript generally well written and organised.

However, several major issues can be identified:
1 The study extends some previous work on a generic method for defining a hydrological drought with a variable threshold (Van Huijgevoort et al., 2012), a method that would be applicable to semi-arid catchments with long periods of zero flows. First, the methodological extension presented L165-L198 is not clearly described, and would deserve some more didactic illustration and/or pseudo-algorithm.

We agree with the reviewer that the original manuscript did not do a good job in explaining the updated method. We refer to major concern number 2 as defined above. We have increased the description of the methon, making it inline with the original method of Van Huijgevoort et al (2012) as well as including an additional figure we hope to have addressed this accordingly.

2. Second, one wonders about the usefulness of proposing an extension of the variable-threshold hydrological drought definition. Indeed, the basic reasoning of using a variable threshold is to detect streamflow lower than usual. When streamflow is usually zero, what is lower than usual? The conceptual limits of the variable threshold method are here clearly reached. Indeed, what is the point of by all means trying to compute/define a streamflow deficit when there is no streamflow? I guess that defining drought in a non-perennial river may be useful, for example for aquatic biodiversity, but streamflow drought is here not relevant. What could be relevant for example in this case is defining an edaphic drought in the river bed to assess the decline of soil moisture during the dry season, and the state of the corresponding habitat for invertebrates.

The reviewer is indeed correct! We did a very poor job in addressing this aspect in the original manuscript (see also our general response above. We refer to our major concern 1 where this comment was treated explicitly. By updating the different components throughout the manuscript, we feel we have properly addressed this aspect.

3. Third, and in relation to the above comment, the objective of the study is clearly ambiguous. Indeed, the title does not even mention the proposition of a new methodology. Second, the current title suggests some relationships between catchment characteristics and drought characteristics, and the way results are presented – including the poor map of Figure 1 that don't even show the delineation of catchments – do not allow to even extrapolate results in other basins with similar characteristics.

Agreed. Though we feel that it is not so much Figure 1 that should address this, but in stead, additional information in the text providing more detailed information about the three different hydrological categories and there locations within the landscape. Furthermore, it should be noted that the title will be altered in line with the suggestion raised by the different reviewers.

4. Fourth, the interpretation of results in terms of hydrological droughts as computed with the newly proposed method is in my view just that: an interpretation. Indeed, this is when results are presented

that the reader realises that there is no possible assessment of whether the new method is more relevant than another one, because precisely of the questionable relevance of using a variable threshold for zero flows. This method indeed comes down to assess how much the river is not flowing...

As a conclusion, a new methodology to compute hydrological drought – understood as anomalies with respect to a variable threshold – for semi-arid basins has to show a continuity with perennial basins in order to be taken as a serious candidate. This would in my opinion be the only way to assess the underlying computation assumptions. This is what the authors have tried to present, but results are unfortunately unconvincing. And even with such a continuity, I definitely question the overall approach of hydrological drought understood as anomalies when considering non-perennial basins. This study has indeed one merit: to exemplify the irrelevance of considering variable threshold approaches for characterizing hydrological droughts. As an example on the other side of the hydrological spectrum, I am not sure why a slightly-shorter-than-usual severalmonth-long period with a dry riverbed would be named a wet spell for a given basin.

As indicated, we would like to refer to our general response stated above. We believe this addresses the different concerns raised by the reviewer.