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

Reviewer 2

<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.**

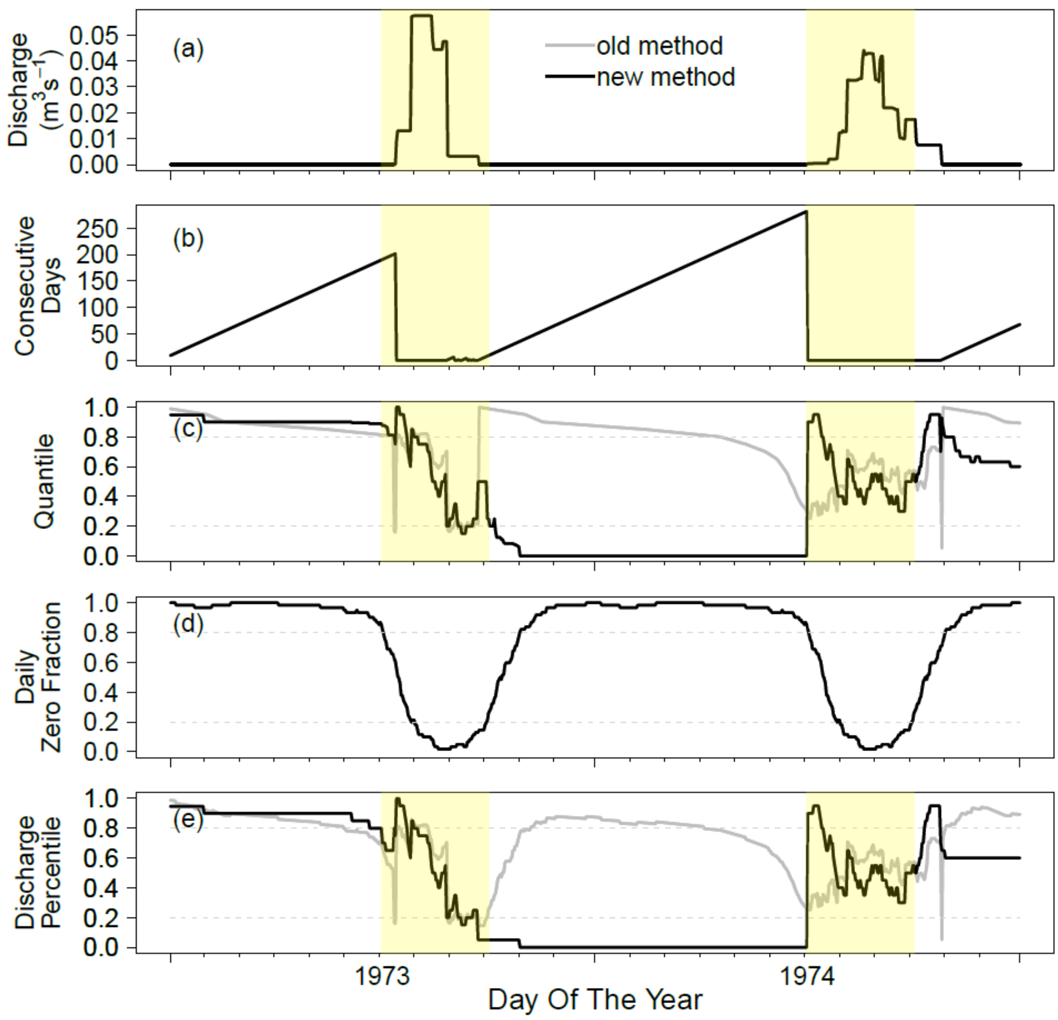
We are convinced that by resolving these concerns and addressing the various issues raised by the reviewers the quality of the manuscript has improved consirably.

The authors deal with a modification on the definition of hydrological drought that is supposed to be more robust for ephemeral rivers with prolonged zero-flow periods. While I understand the reasons behind the need of such modification of the van Huijgevoort et al. method, I find the message not well communicated and the proposed solution not fully on point. Both the tackled issue and the proposed solution need, in my opinion, to be better described, maybe with the help of a visual representation of a real study case (for a specific event/year) or even an artificial case (that highlight the key drawback of the current method) before showing the performance on all the available data.

We would like to thank reviewer 2 for the constructive comments. This really helped us to evaluate the quality of our manuscript an assess the current bottle necks. Besides making these major changes to the paper, we have tried to carefully address the major and minor comments as given by reviewer 2, which are presented below.

At the current state, it is difficult to gasp how the suggested modification actually works, since the full description is based only on text, and the examples in Fig. 4 is really hard to read. Also, since the method is needed for zero-flow rivers, the authors should focus only on such stations rather than all the available stations. If I understand correctly, the authors state that the method of van Huijigevoort et al. may “break” a drought event in two events if an event start during the wet period and continue during the dry one, but this should be highlighted in a figure that show one of such case and how the method solve the issue.
My understanding is that the author define a cdf of number of (antecedent) dry days that is different for each day of the year, rather than the same for all the day based on the total length of the zero-flow. This seems a solution to avoid “breaks” for events that started during the wet period, but can lead to difficulties for events starting at the beginning of a dry period. If my understanding is correct, I suggest to the authors to consider, first of all, if their goal is to produce an indicator that can be update in near real time (while the event is developing) or that defines the drought on past data. In the second case, better solutions can be found than the one proposed. Since a “true” definition of drought/wet spell start and length is not available (for obvious reasons), the authors need to clearly highlight that their outcomes are at least more reasonable that the one obtained with the previous method and not just as arbitrary.

We agree with the reviewer that the original manuscript was not very clear in explaining the different steps and indicating the benefit of the newly presented approach. This was partly due to the fact that it was our interest to compare the different flow domains throughout the region, as indicate our general response above. As a result we did not want to put too much emphasis on the newly presented approach, but did feel it was important present these changes wrt the original method. As indicated in major concern 2 given above, we have updated the manuscript and better describe the newly presented approach. Furthermore, in line we major concern 1, we now motivate why we want a drought that starts in during the NAM to continue during the no-flow season afterwards. We have updated figure 5 of the manuscript which now shows the year 1973. This figure shows that for new approach, this “catching up” is not occurring anymore. We agree with the reviewer that the original manuscript did not emphasize this aspect. Therefore, we have decided to stress this aspect more by adding additional text in Section 3.2 of the manuscript as well as within a newly created section 4.1.



A second point of contention for me is the need for a better explanation on the reasoning behind this kind of definition of hydrological drought and (especially) wet spells in zero-flow rivers. As the authors stated, they are looking at an issue that arise for specific rivers, with zero-flow during most of the year and flow only during monsoon, but it is not clear what is the goal of having hydrological drought (or wet spells) defined for such rivers in such a way. While the classic analysis of dry spells (length of periods of zero-flow) or wet spells (length of period with positive flow) in such rivers is relevant, I do not understand, for instance, the reasoning to define a day as part of a "wet spell" even if the flow is zero (e.g., first half of 2007 in Fig. 6a).

We agree that the original version of the manuscript did not do a good job in indicating why we chose to focus on the three different flow domains, and why we opted to focus on drought and wet spells for rivers with zero-flow conditions occurring throughout the year. We would like to refer to our response to major concern 1. We believe the updated version of the manuscript better addresses this issue, as it was clearly lacking in the original work.

Finally, the title of the paper is ambiguous, since the focus in on a redefinition of drought events in zero-flow rivers but the title seems to imply that the effects of elevation and flow dynamics in semi-arid rivers will be discussed.

The original manuscript was indeed not very clear on why the focus was on elevation and flow dynamics. We believe that by adding this information, as stated in the previous point, the focus of the paper on both rivers with and without baseflow within eastern Arizona becomes

clear. We will update the title accordingly. Furthermore, we will thoroughly address the major concerns identified above. By doing so, we believe the paper becomes less ambiguous.

Follow some minor comments:
P2L30. Why wet spells on rivers are defined based on precipitation here?

For southern Arizona we believe that precipitation is the main driver of hydrological drought. At shorter timescales, groundwater releases can cause a wet spell but there impact is only visible for the perennially flowing rivers. We have stated this in the introduction.

P2L39. This is not true for all the cases. There are plenty of evidences that over some regions less extreme drought are expected (e.g., northern Europe).

The reviewer is indeed correct.

P2L50. What is the relevance for water managers in rivers with zero-flow? Are those rivers under any water managing?

We refer to our response to major concern 1. These anomalies are not directly relevant for water managers, but instead indicate the rootzone moisture availability within the riparian zone.

P4L88. Why the river is here identified as perennial, whereas is defined ephemeral in the rest of the text (see e.g., P3L74) or just partially perennial (P4L103). Please be consistent.

We have changed this into:
"As flow conditions at lower elevations for the majority of channels throughout this domain are ephemeral,"

P4L120. You should focus only on the NB rivers, and eventually show that your method works for the other rivers too (if this is the case). P6L142. This should read as: "The TLM has a problem for locations with zero flow (in a specific period) for a considerable amount of years". Please reword.
We refer to our major concern 1. The original manuscript did a poor job in describing why we wanted to differentiate between these three domains and not just present an updated version of the paper by Van Huijgevoort et al. (2012).

P7L181. It is really difficult to extrapolate how the two methods work from fig. 4b. If 2003-2004 is a good example year, please make a specific figure that highlight the key differences between the two methods.

We have added Fig. 5 (see above) to highlight this aspect and are currently creating a schematization figure where both methods are compared.

P8. Sensitivity Analysis. This section does not seem well thought-out in my opinion. The range of values adopted in this analysis need to be better supported by some reasoning (e.g., pooling up to 180 days? moving windows of 5 years?).

W.r.t to the length of te MA windows, we refer to our general response given above. Concerning the pooling, we agree with the reviewer that pooling up to 180 days is extreme. The reason why we decided to do this is to show that results do not very much, with these long pooling windows. However, we did not address this properly in the first submission of the manuscript. We will ensure that this will be treated in an updated version of the discussion (as also raised by the reviewer).

P9L265. This description of the behavior of drought is a consequence of your definition of theevents rather a fact. You need "independent" evidences on the behavior of drought to support that your reconstruction is more adherent to the reality than the one obtained with the previous method.

Correct. However, we feel that for semiarid regions the method presented here is able to represent the expected behavior as indicated above. However, we agree that for other locations this does not hold.

P10L290-300. This analysis on the long-term variations is out of topic and not well supported by formal trend tests and analyses.

We do not agree as we believe there is scientific literature to motivate these assumptions. Without accounting for long-term variations, the San Pedro would show a continuous drought starting in the 1990's which is solely the result of human behavior. Although this is technically correct, we believe that it does not provide the real message concerning the observed hydrological state. However, this was not clearly defined in the original submission. We added the following statement to Section 3.3:

"Although, the occurrence of these type of changes was rare for the San Pedro, it was observed for a few locations including station 3 (see also Fig. 2), where both a gradual decrease and shift in the late fifties was observed (see also Fig. 2). By not identifying and correcting for trend and step changes, results would show a drying out of the system, with a wet spell being the predominant condition during the first twenty years, and a drought being dominant during the last twenty years. Although, technically this is correct, we believe this would only reflect the impact of human behavior with increased amount in water uptake throughout time and added modifications throughout the catchment. As stated before, it is the interest of the current work to understand the behavior of hydrological anomalies and how these vary between locations. To be able to assess these, the identification and correction of shifts and trends is necessary, which subsequently allows for the identification and analyses of anomalies within the hydrological system (e.g. Villarini and Smith, 2010; Sadri et al., 2016)."

P11. Fig. 8. This figure is rather confusion, and, in my opinion, not the best way to convey the key finding that the authors want to show here.
In line with the suggestion raised by reviewer 1, we have changed this figure into 6 separate panels. We believe that this improve the readability, shows the dominant behavior of the North American Monsoon and enables differentiation between the three different flow domains representative for eastern Arizona.

P12. Section 4.3 is this on only one river station or all the stations combined? This is not clear.
This is for all stations combined and has been added to the text.

P13L374. If a backward moving window is used (rather than a most common centered one) this need to be clarified and justified in the methodology.
This is indeed an important point raised by the reviewer! Given that the hydrological state shown here is indicative for the state of the channel bed including vegetation, a backward looking window is justifyable. We include this statement in Section 3.4.