

Comments to the Author:

I thank the authors for addressing the initial review comments thoughtfully and thoroughly. As noted by all of the reviews, the manuscript is also very well written. I agree with the review comments that this manuscript will provide a useful technical note to aid in the detection of change in environmental phenomena. Approaches and techniques to identify these shifts are critical to the future of hydrology research and I commend the authors for taking on such a difficult yet highly important topic.

I also had the opportunity to carefully look over the revised version of the manuscript and have some additional comments noted below. I hope the authors do not find these comments too onerous to address and I look forward to the revised manuscript.

R: We would like to thank the Associated Editor for these additional comments that have improved the clarity of the message in our revised MS now entitled “Higher-order statistical moments and a procedure that detects potentially anomalous years as two alternative methods describing alterations in continuous environmental data”.

1. Please provide additional description of the N-MDS procedure. This is not a common statistical technique applied to hydrologic problems. Please add 2-3 more sentences about this method and, if possible, include an example or equations so that a reader can gain a basic understanding of this technique.

R: We now include additional information about the N-MDS procedure and the corresponding algorithm used here. In addition, we added the equation used to estimate the Kruskal’s stress (see lines 206-219).

2. Clarify the time series used. In line 183, it is stated that “365 days of daily minimum stream temperature” were used. Does this mean you had sub-daily temperature data from which the daily minimum was obtained? Describe in more detail how the data were broken up and standardized by season. Was this a result of the technique or did you make this classification prior to applying the technique (this was particularly confusing in lines 216-223).

R: We clarified and reorganize this in the text (starting line 112). Time series characteristics including time periods and seasons were defined in the section named Study Site and Time Series (see also Table 1). In this section we described the availability of daily time series as “...(2) complete information for time series of daily minimum (min), mean (mean), and maximum (max) stream temperature...”. In addition, we mentioned “...To ensure enough observations to adequately represent the tails of the respective distributions at a seasonal scale for analyses of higher-order statistical moments (i.e., winter: December-February; spring: March-May; summer: June-August; fall: September-November), we grouped and compared daily stream temperature data at each site among the three decades 1980-1989, 1990-1999, and 2000-2009.” In the new version of the MS we also added the sentence “...see study sites and time series section above...” after the text mentioned in line 183 (old version; line 201 in the revised MS).

3. Consider another naming convention rather than “unregulated site2” (line 230). Maybe a shorter abbreviation or site name?

R: We changed “unregulated site2” to “site2”. In addition, we revised this convention throughout the text for the rest of the sites.

4. How does your approach take into account sample size when identifying a shift? Can the method determine with some confidence that observed shifts are simply not due to chance versus a true shift?

R: Based on the reviewer’s suggestion, we clarified that the procedure to detect potentially anomalous years does not identify shifts in long-term data. In addition, the time scale used to illustrate the utility of higher-order statistical moments in detecting shifts of distributions was determined *a priori* as decadal (using daily values). Unfortunately, the question whether the length of time series may affect the detection of shifts was outside of the scope of this technical note, but certainly it merits further attention. In the discussion section, we added a short paragraph about sample size and other additional considerations related to detection of changes using time series of environmental data. We added (starting at line 338) the following sentences at the end of the discussion “There are some considerations when detecting potential changes in continuous environmental phenomena that are inherent to time series analysis including the length, timing, and quality of the time series as well as the type of the driver that is investigated as responsible for such change. Often, the detection of shifts in time series of environmental data is affected by the amount of censored data that limits the length and timing of the time series (e.g., Arismendi et al. 2012). There are uncertainties regarding the importance of regional drivers and the representativeness of sites (e.g., complex mountain terrain) and periods of record (e.g., ENSO, and PDO climatic oscillations). Lastly, the type of climatic influences may affect the magnitude and duration of the responses resulting in short-term abrupt shifts (e.g., extreme climatic events), persistent long-term shifts (e.g., climate change), or a more complex combination of them (e.g., regime shifts - Brock & Carpenter, 2012).”

5. In lines 303-306, you state that you used longer time periods to remove the effects of serial dependence. But if the time resolution of those longer periods were still on the order of days, using longer sequences would not result in less serial correlation. Aggregating to seasonal or annual series would, however. Could you clarify your statements here?

R: Thanks for pointing this out to us. In the new version of the MS we clarified as follows (starting line 190) “In this study, we reduced the potential for serial correlation by using higher-order statistical moments aggregated over longer time periods that allowed for a contrast among decades.”

Referee 2

1. Arismendi et al. have provided a significant revision on their original manuscript. Pending some minor revisions of the methodology and minor clarifications, I feel that this manuscript merits publication. They present tools to assess the distributional shifts of continuous climate variables and to identify potentially anomalous years. In the interest of a targeted response, I have provided line-by-line comments and suggestions. I am confident, that upon additional

revision, Arismendi et al. will have provided a good contribution to the analysis of environmental phenomena.

R: We would like to thank this reviewer for providing these additional comments that improved our revised MS.

2. Throughout: The authors refer to ‘higher statistical moments’ and ‘lower statistical moments’. Please consider revising to ‘higher-order statistical moments’ and ‘lower-order statistical moments’ throughout the manuscript. The point is not that the statistical moment is higher (a mean of 5 is higher than a mean of 3), but rather that the order of the moment is higher (a mean is a first-order moment, while a standard deviation is a second-order moment). Though not an exhaustive list, examples occur at lines 1, 28, 65, 72, 74, 96, 121, 125, 133, 226, etc.

3. Throughout: Central tendency statistics are not meant to capture characteristics of the variation. Furthermore, the paper discusses the shortcomings of statistics of central tendency and dispersion; this is not reflected in the current abstract or manuscript. I would suggest revising the first clause of line 25 to read “Statistics of central tendency and dispersion may not capture relevant or desired characteristics of the distribution of continuous...” This is also reflected in lines 47-50, line 331 and elsewhere.

R: We revised and incorporated these modifications throughout the text.

4. Throughout: I am uncomfortable with the term ‘outlier-detection’. As mentioned below, this technique will always identify 5% of the years as outliers. The term outlier carries a negative connotation, but all this method is doing is identifying years in the tails of the distribution. The authors proposed changing ‘outlier’ to ‘anomalous’ in the revised manuscript, but this completed, as can be seen throughout the manuscript and in the title. While I would argue that anomalous carries a similar connotation, I cannot think of a better term; perhaps something like ‘potentially anomalous’. In any case, please consider modulating the use of the terms ‘outlier’ or ‘anomalous’.

R: Although we have considered many alternative words in place of anomalous, as the reviewer mentions, there is not a better term that may properly describes the technique used here to identify those years that behave differently than the others. However, we revised throughout the text and modified the wording to include ‘potentially anomalous’ rather than ‘anomalous’.

5. Line 27-8: The author claims to present two methods that identify long-term changes. This suggests, along with comments in lines 62, 282-3, and elsewhere, that the outlier detection technique is capable of identifying distributional shifts. The identification of an extreme year, as this method will do for 5% of all years regardless, does not imply distributional shifts. Instead, it merely indicates a particularly extreme event. Instead, as the authors observe in lines 100-3, the outlier-detection methods merely identifies years that may prove interesting upon further investigation. The manuscript should be revised to clarify this point: outlier-detection does not imply distributional shifts.

R: We clarified throughout text that this procedure does not detect distributional shifts.

6. Line 47-50: This sentence states that most traditional statistical approaches are based on the detection of location changes. This implies that tests of variability or other parameters are not common. Considering that there are many tests of variability, the widely-used ANOVA test, for example, I do not think that this statement is accurate. I agree with the second clause, namely that testing locations over-simplifies the problem, but I do not think that the authors should dismiss the testing of higher-order statistics so quickly. Please revise or provide a citation that supports the claim.

7. Line 62-6: In the previous paragraph, the authors observe the limited value of central tendency metrics. Here they jump to a discussion of skewness and kurtosis without any discussion of the value of dispersion metrics (IQR, range, standard deviation).

R: We agree. In the revised version of the MS, we clarified this misunderstanding. We refer to the use of metrics of location and dispersion similar to traditional statistical approaches, and not only location as was previously stated.

8. Line 65: Revise ‘the shape of stream temperature empirical distributions’ to read ‘shape of the empirical distribution of stream temperature’.

9. Line 100-1: Insert ‘an’ between ‘Using’ and ‘outlier’. Remove the hanging parenthesis after technique. The revision should read: ‘Using an outlier detection technique, we...’

R: We modified the text accordingly.

10. Lines 116-7: The authors state that a stream was classified as regulated if reservoirs existed prior to 1978. Does this imply that an unregulated stream could have a reservoir on it that has been operational since 1980? The degree of abstraction allowable in the unregulated class is unclear. Please clarify the definitions of regulated and unregulated.

R: We clarified in the text (lines 118-120) “Regulated streams were those with reservoirs constructed before 1978, whereas unregulated streams had no reservoirs upstream during the entire time period of the study (1979-2009)”.

11. Line 117-8: Add commas between ‘and’ and ‘for’ in line 117 and between ‘below)’ and ‘we’ in line 118. More importantly, how much of the record was interpolated: days, months, years?

R: The amount of days interpolated for each time series is included in Table 1. In the text of the revised version, we fixed this misunderstanding by reorganizing this section (see response 2 to the Associate Editor’s comments) and by adding a proper reference to Table 1 (see lines 112-127).

12. Lines 133-4: This first sentence, as it reads now, suggests that parametric statistics assumed normality. This is not the case. Parametric analysis implies a distributional assumption, not an assumption of normality. I believe the authors intend to point out that common estimators of skewness and kurtosis are unbiased only for normal distributions. Please revise this sentence accordingly.

R: We modified the text accordingly.

13. Line 134: Remove the phrase ‘In reality, however,’. The reality of the previous statement does not contradict the reality of this statement. Please revise the transition. Something like ‘Still,’ or ‘Regardless,’.

R: In the revised manuscript, we rephrased to (lines 137-140) “Although common estimators of skewness and kurtosis are unbiased only for normal distributions, these moments can be useful to describe changes in the shape of the distribution of environmental variables over long-term periods (see Shen et al., 2011; Donat & Alexander, 2012)”.

14. Lines 159-63: The authors cite the missing values as a motivation for the subsequent formulae. Weren’t these records already completed, per lines 117-8? Please explain.

R: We did not use interpolated data to estimate higher-order moments. To be clearer, we reorganized this section to improve clarity (see response 11 above and response 2 to the Associate Editor’s comments). We explained these details in the text (see lines 112-127 and 163-167).

15. Line 192: Strike the phrase ‘A rule of thumb, (‘ to begin the sentence ‘Clarke (1993) suggests...’

16. Line 219: Revise ‘occurred’ to ‘occurring’.

17. Line 224: Include the word ‘Possible’ to start this sentence. Because there is no way to reliably test the changes in skewness or kurtosis, it is inaccurate to trust changes explicitly.

18. Line 263: Add a comma and append an ‘s’ to ‘season’. The revision should read ‘regimes, showing which seasons were’.

19. Line 264: Revise the beginning of this sentence to read ‘The use of higher-order moments could help improve...’

R: We modified the text accordingly.

20. Line 264-6: This claim is not substantiated by the text. How do higher-order moments of temperature incorporate site-specific characteristics? Please clarify this statement and provide additional support or remove it.

R: We removed this section from the text.

21. Line 278: Strike the word ‘spatially’ to read ‘...sites located close...’

22. Line 279: Add a comma between ‘years’ and ‘suggesting’ to read ‘...anomalous years, suggesting that...’

R: We modified the text accordingly.

23. Line 306-11: This is a very important statement and should not be appended to the tail of the discussion. This is a limitation of the methodology. As such, it should be mentioned in the methods section and discussed in tandem with the relevant results. The interconnectedness of moments significantly alters the readers’ interpretation of the results presented in lines 224-66.

24. Line 311-6: As with the previous observation, this is an extremely important claim. Again, it is a limitation of the methodology and should be mentioned when the method is introduced. The ‘outlier-detection’ method will always identify 5% of the years as ‘outliers’. This limitation affects the interpretation of the results and should be included in the discussion.

R: In the new version of the MS we moved, from the discussion to the methods, the two respective paragraphs that address the potential limitations of the use of higher-order moments and the procedure to detect potentially anomalous years (see revised MS: lines 185-197 and lines 240-251 respectively). In light of the reviewer's suggestion, in the discussion we carefully reworded our statements and interpretation of our results. We believe the limitations of the methods are now in the right place and treated sufficiently in the text in a clear and concise manner (see also lines 338-348).

25. Lines 335-7: The authors have not shown the power of looking at the entire empirical distribution. Instead they have presented the merits of looking at both the empirical distributions and multiple moments. Please add the phrase ‘and multiple moments’ to read ‘By examining the whole empirical distribution and multiple moments, we can ...’

R: We modified the text accordingly.