Reply to 'Review comments for hess-2018-173', Anonymous Referee #1

Comments: This paper provides very interesting results, and topic of the research is within the scope of this journal. The manuscript is well written and organized, but I have several minor comments as follows.
[General Comments] In the dynamic threshold method, three parameters (windows width: w, cut-off frequency: y and confidence level: alpha) should be specified, and sensitivity of w is discussed in this study. However, how about the sensitivity of y and alpha on extracting run length? # If authors could provide some comments/ideas on that, it would be helpful for readers. # (but do not have to conduct additional sensitivity analysis.)

Reply: We thank the reviewer for the comments and appreciate the opportunity to clarify aspects of the manuscript. Below we present our responses and indicate the changes made to the original manuscript.

The cut-off frequency parameter in the dynamic threshold method is used to filter out centennial trends that may be mixed with decadal variability. We therefore believe it is appropriate to use a cut-off frequency of 1/100 years. This parameter is aimed at filtering centennial trends and has little influence on the inferred runs which have significantly higher frequencies.

We have replaced Line 12 Page 10 in the original manuscript with “The cut-off frequency parameter in the dynamic threshold method is used to filter out centennial trends that may be mixed with decadal variability and should have little influence on the statistical characteristics of the inferred runs. A cut-off frequency of 1/100 years is considered to adequately meet these requirements.”

The Mann-Whitney test is used to determine whether two independent sets of data come from the same distribution. There is a (100-α)% percent chance of the test statistic falling outside the α% confidence limits under the null hypothesis. If α is selected close to 100 and the test statistic falls outside the confidence limits, we are confident that the detected change point is actually a change point. The value of α is typically set by the researcher and involves consideration of type 1 and 2 errors (Zar, 1999). In this study, we follow the practice used in other reconstruction studies (Gedalof and Smith, 2001; Shen et al., 2006; McGregor et al., 2010; Pent et al., 2015) that adopt a confidence level of 90% so that the chance of making a type 1 error (that is, rejecting a changepoint when in fact there is one) is low, namely 10%.

We have added the following lines in the original manuscript at Line 14 Page 10 “In this study, we set the confidence level to be 90%, a value that is consistent with other reconstruction studies (eg Gedalof and Smith, 2001; Shen et al., 2006; McGregor et al., 2010; Pent et al., 2015)”.

[Specific comments]
P13, P24-25: How did you set the value of “confidence level alpha (=90%)” and “cutting-off frequency (=1/11)” for this analysis? A little more explanation is expected (if possible). # P10, L14: y=100 (?)

Reply: See the above reply for our response on setting of the confidence level α.

The cut-off frequency of 1/11 is only used in Figure 7 to plot filtered PDV reconstructions. It is not used in the dynamic threshold method. The cut-off frequency of 1/11 was selected to be consistent with the cut-off frequency used in the Parker instrumental IPO time series.

P15, L13- “The absence of a consistent difference ... may be a consequence of sampling variability ... ” –> The meaning of this part is a little unclear. # What is “sampling variability”?

Reply: Random samples from a given population are used to estimate target statistics for this population. Sampling variability refers to the variability in the target statistics that arises when random sampling is repeated. For example, two samples of 10 runs from the true run population will give different estimates of the run standard deviation. In this research, there are a limited number of PDV run samples. Therefore, it is important to recognize that differences in run statistics may be due to sampling variability. It is only when differences are bigger than what would be expected due to sampling variability that we would conclude there is a significant difference.

We have added the following text at Line 15 Page 15 to clarify this important issue:” Sampling variability refers to the variability in the target statistics that arises when random sampling is repeated. Therefore, it is important to recognize that differences in run statistics may be due to sampling variability. It is only when differences are bigger than what would be expected due to sampling variability that we would conclude there is a significant difference.”

P28, Fig.1 What blue broken lines in (a4), (b4) and (c4) of Fig.1 are representing? #

95% confidence bands for zero autocorrelation (as described P7, L8)?

Reply: The dashed lines present the 95% confidence bands for zero autocorrelation. If the true autocorrelation were zero, there is 95% chance that the sample autocorrelation coefficient will lie between the dashed lines.

We have added the following text in Line 11 Page 28 Figure 1 caption: " The dashed lines in (a4)-(c4) present the 95% confidence bands for zero autocorrelation."

Reference:

Gedalof, Z., and Smith, D. J.: Interdecadal climate variability and regime-scale shifts in Pacific North 32 America, Geophysical Research Letters, 28, 1515-1518, 10.1029/2000GL011779, 2001.

McGregor, S., Timmermann, A., and Timm, O.: A unified proxy for ENSO and PDO variability since 20 1650, Climate of the Past, 6, 1-17, 10.5194/cp-6-1-2010, 2010.

Peng, Y., Shen, C., Cheng, H., and Xu, Y.: (2015): Simulation of the Interdecadal Pacific Oscillation and its impacts on the climate over eastern China during the last millennium. J. Geophys. Res. Atmos., 120, 7573–7585. doi: 10.1002/2015JD023104.

Shen, C., Wang, W. C., Gong, W., and Hao, Z.: A Pacific Decadal Oscillation record since 1470 AD 46 reconstructed from proxy data of summer rainfall over eastern China, Geophysical Research Letters, 47 33, L03702, 10.1029/2005GL024804, 2006.

Zar, J. H. (1999). Biostatistical analysis, 663 pp: Prentice Hall, Englewood Cliffs.