Reviewer 2

Review of Grant et al., Historical trends of seasonal droughts in Australia

The manuscript presents a statistical assessment of the historical trends of drought in Australia. The methodology and the analysis are well designed and contain elements of novelty, while the results are interesting, relevant and generally well discussed. There are, however, area where the manuscript could be improved, in particular with regards to the explanation of some of the methodologies and their contribution to the final results, as well as discussion of the implication of the statistical choices made.
Overall, the manuscript is of good quality and I recommend its publication in HESS with minor amendments.

We would like to sincerely thank the reviewer for taking the time to review the manuscript and for their positive assessment.

Scientific significance:
Does the manuscript represent a substantial contribution to scientific progress within the scope of Hydrology and Earth System Sciences (substantial new concepts, ideas, methods, or data)? **YES**

Despite presenting what is in practice a statistical analysis, the manuscript represents a significant contribution both methodologically and, to a lesser extent, because of the results presented. Methodologically, the manuscript presents a sound and thorough methodology for historical drought assessment, taking into consideration different aspects contributing to drought hazard, as well as potential impacts. The application of ML to model drought impact data is particularly interesting in this regard. Most of the results presented, additionally, highlight the complexity of drought as a natural disaster and present some interesting novel insight.

While the implications for agriculture and water supply of the results are outlined in the discussion section, the relevance of the manuscript could benefit from a deeper discussion of the relationship between the results obtained with the impact-based metric and the traditional ones and their implication. e.g. what meteorological variables seem to be the most relevant drivers of impact? What does this entail with regards to future climate scenarios?

We thank the reviewer for their suggestion. We agree that it would be helpful to more explicitly connect the trends in the impacts metric to the trends seen in the traditional metrics. We have shown that traditional drought metrics are able to capture historical trends in drought that we see in the impact-based drought metric (see Section 3.1), suggesting the traditional metrics are able to capture trends in drought impacts. We will expand further on the implications of the trends in the impacts-based metric on our understanding of the traditional metric trends.

Scientific quality:
Are the scientific approach and applied methods valid? Are the results discussed in an appropriate and balanced way (consideration of related work, including appropriate references)? **YES**

The methodology is well designed and generally well explained and referenced. Assumptions associated with the various methods are also generally well explained and reasoned. However, a few methodological choices could benefit for better explanation, reasoning and discussion of their implications:

- The choice of using the 15th percentile threshold as opposed to more traditional metrics, especially for the precipitation (SPI).

The 15th percentile is approximately equivalent to an SPI of -1 (Mckee et al., 1993), but using this method does not require any assumption of the data's distribution, and so can be applied consistently across the multiple drought types. We will revise the methods section to further clarify this.

- The trend emergence section (2,4) seems to have an implicit assumption of normality in the calculation of the Cis, is this correct? This approach feels like it needs better justification, the text just says that it is "used in previous studies".

We will revise the manuscript to include assumptions and limitations of the statistical methods used, where we have not already done so. Additionally, we will provide a more detailed explanation and justification of the decadal confidence intervals method used for assessing the stationarity of the area under drought trends.

- Also with regards to this section, I had to read through several times to understand how this was was differently for "time in drought" and "area in drought", could be more explicit.

The decadal confidence intervals were only applied to the area under drought trends, whereas the signal-to-noise and Kolmogorov-Smirnoff tests were applied to both the time and area under drought trends. We will review this section and, where necessary, make the text more explicit.

- Section 2.5 about the contributions from mean and variability seem quite simplistic and is not clear to be if this is a method that the authors came up with or is rooted in literature (there are no references). If the latter, a discussion of the limitations is warranted.

The method was developed by the authors and used here as there was no method in the literature which attributed drought trends to the mean and variability changes of the underlying hydrological variable. We will further clarify the method and its motivation and methodological justification in the revised manuscript.

- In general, in section 4.2 on the limitations, the focus is on the data, however a better discussion of the limitations and the assumptions of the statistical methods and tests performed should be included.

We have included limitations of the statistical methods within the results sections relevant to these methods. For example, we discuss the limitations and assumptions of the feature importance of the Random Forest models in Section 3.5 (L396-398). We will add information on assumptions and limitations of statistical tests where we have not already done so.

The results are also clear and well presented and their implications and limitations are well discussed, however the results from the RF model could benefit from additional synthesis and discussion.

As stated above, we will expand the discussion to include the implications of the (Random Forest) impact-based metric.

Presentation quality:
Are the scientific results and conclusions presented in a clear, concise, and well-structured way (number and quality of figures/tables, appropriate use of English language)? **YES**

Except for the few examples already mentioned above, the presentation is clear and easy to follow. Tables and Figures are clear and relevant.

References

Mckee, T. B., Doesken, N. J., and Kleist, J.: The Relationship of Drought Frequency and Duration to Time Scales, 8th Conf. Appl. Climatol., 179–184, 1993.