Response to Anonymous Referee #1

This manuscript explores the question of whether recent extreme hurricane seasons can be attributed to changes in circulation and SST. To explore this question, the authors use a new statistical model to explore weather patterns and find that increases in Atlantic SSTs have led to an increase in the probability of extremely active hurricane seasons. While I think that this is an interesting result worthy of publication, I believe that in its current form the manuscript has two major issues:

1. There are several inconsistencies throughout the manuscript (see examples below), as it pertains to the time periods and regions used for the analysis. Some additional clarity is needed and urge the authors to check the full manuscript for consistency.

We agree with the reviewer and are thankful for them pointing out the above mentioned inconsistencies. Practical choices that made sense during the process of our explorative analysis led to inconsistencies in the current version of the study and we are more than happy to provide a fully consistent methodology in a revised version of the manuscript. Furthermore, we think that clearing out the rightfully mentioned inconsistencies would not affect the results and conclusions of the paper.

2. In its current form, I find the organization and utility of the supplementary information figures difficult to follow. While there are references to the figures throughout the main manuscript, the order - and need for 28 additional figures - is not clear and distracts from the main results of the manuscript. I would encourage the authors to include all necessary figures for interpreting the results in the main manuscript and use supplementary information plots to contextualize statements. In addition, all Figures should be able to stand on their own and be interpretable. I find the lack of axis markers and labels limit the effectiveness of some figures.

We agree that the supplementary information could be organized and structured more meaningfully. In a revised version we would leave out a few figures in the supplement and improve the remaining ones. We are however hesitant to include more figures in the manuscript. We carefully selected the content that we want to show in the main paper as we don't want to overload the paper with figures that might distract the reader. We will, however, think this through again and see whether one or two figures could be included in the manuscript in a helpful way.

Once these major comments, as well as the list of minor comments below, are taken into consideration, I believe the manuscript would be suitable for publication in Weather and Climate Dynamics.

Additional Comments:

L49: Why is the 1979-2018 period used? In doing so, the authors are leaving out the 2019 and 2020 seasons, which were destructive. Does the signal become larger? If so, this would be very interesting to include.

We thank the reviewer for pointing out this inconsistency. When starting the clustering exercise for the weather patterns, we were working on the available ERA5 data at that time which was 1979-2018. Later on we decided to use the DOISST data set from NOAA, which only starts in 1982. In a revised

manuscript we would consistently use the period 1982-2020 (or even 2021) for each part of the analysis.

L53: What is the original frequency that is averaged to daily?

The original data is sampled on a 6-hourly frequency. We will clarify this in the revised manuscript.

L55: Why is the SST averaging area (20W-90W) different than the area used to classify the weather patterns (10W-90W)? Is the analysis sensitive to this approach?

Our weather patterns are supposed to represent the large scale atmospheric circulation in the region while the region over which we average SSTs is the hurricane main development region (MDR) that is commonly used in the literature. These two regions serve a different purpose for the analysis. We think that it makes sense to stick to the used region for the SST averaging to facilitate comparability with other studies. The region for the weather pattern classification could be altered, but we do not expect this to considerably affect the weather patterns.

L57: IBTrACKS should be IBTrACS.

We will change this.

L58: How is this region in Figure S3 selected? Why not all of the Atlantic? Is the results sensitive to this selection?

We initially used this region to filter out extratropical cyclones. We now checked again and found out, that the type of storm is documented in IBTrACS. In the revised manuscript we will use this storm classification from IBTrACS and won't need this region any more.

L60: When calculating ACE, do the authors follow Bell et. al, 2000 and only include storms while they are tropical storm strength or great? That should be clarified since the NOAA classification for above normal and extremely active seasons. Are the results sensitive to this methodology?

We thank the reviewer on pointing out this inconsistency! In the initial manuscript we calculated ACE including all storms, also those which are weaker than tropical storm strength. In the revised manuscript we will change that to be consistent with Bell et al. 2000. As the seasonal ACE is calculated in the same way for observations and emulations, we don't expect to see any major effects on the results.

L93: Missing "and" before iii)

Well noted.

Figure 2: States SST areas is 85W-20W, 10N-20N which is different than the in-text description.

Apologies for this typo! The used region is 90W-20W, 10N-20N. We will update this figure accordingly.

L150: How is major hurricanes defined?

The definition can be found in L62:
"TCs are classified according to the Saffir-Simpson hurricane wind scale according to which TCs with

sustained winds of more than 64 knots are named hurricanes and TCs with sustained winds above 96 knots are major hurricanes."

Figure S1-3, S5 and S7: Should include axis (lat-lon) markers.

Will be included in the revised manuscript.

L152: Why is 1982-2018 used here, when it was stated earlier that 1979-2018 is used for classification of weather patterns?

We have answered this question in our reply to the comment about L49 above.

L195: Why is 1900 used, if the long-term SST trend calculation starts with 1850?

This is a good question. Looking at figure S9, we would argue that the SST levels would be quite similar for 1900 and 1850. This however points to more important flaw in our way of constructing counterfactual SST scenarios: applying a linear trend from 1850 to 2020 is too simplistic and does not reflect our understanding of global warming. Global mean temperatures did not increase linearly over the period 1850-2020. The increase after the 1970s was substantially more pronounced than before. In a revised manuscript we therefore want to detrend the 1982-2020 period using the trend estimated over the the same period. The counterfactual scenarios would consist of this detrended SST data shifted to the values of the linear fit (over the period 1982-2020) at the years 1982 and 2020. For the pre-industrial scenario we would shift the detrended SST time series to the levels corresponding to the average over 1850-1880.

L230-234: What about all years?

The probability of finding 225 ACE irrespective for the weather patterns and SST variations of all observed seasons also increases by roughly 30% as a result of the SST warming since the 1980s (see dashed cyan line in fig. 6b). We will an interpretation of this result in the revised manuscript.

L240-241: What about dust impacts on these weather patterns and potential trends?

This is a good question and we would be interested in investigating this further. Changes in atmospheric circulation patterns are beyond the scope of this paper.