Response to the comments of Reviewer #2:

This paper focuses on critical transitions in precipitation and potential evapotranspiration (PET) both globally and in U.S. urban areas, based on monthly and annual datasets. The authors use various network and correlation measures to identify how system properties change leading up to critical transitions, which are defined as abrupt changes in behavior. They find that autocorrelation and standard deviation computed on moving time windows tend to increase before a defined critical transition point, indicating the potential use as early warning indicators. In an extension to a spatial network of precipitation in urban regions, the authors introduce network connectivity measures and similarly consider how these measures predict critical transitions in precipitation anomalies.

The paper was interesting and relevant to the journal and I think it will make a valuable contribution. The addition of a spatial network perspective on critical transitions in precipitation was particularly interesting to me. However, I have several major and minor comments on the structure of the paper and the methods as detailed below. Mainly, in the methods I would like to question the identification of a critical transition in general, and the possibility for trends in indicators without any critical transition occurring. In terms of the writing, there is a lot going on and several of the sections could be more clearly explained and tied together.

**Note, after writing this initial review, I notice that some comments were addressed based on previous reviewer comments, but have not removed them, so they may now be redundant.*

We'd like to thank the reviewer for the constructive feedback and help in improving the quality of this manuscript. Below are detailed responses to the comments. All changes and clarifications were included in the revised manuscript and highlighted (yellow to Reviewer #1, and green Reviewer #2).

Major comments:

Writing:
The introduction would benefit from some restructuring. For example, there are 3 separate places where different "research gaps" are established, and these could be better tied together Specifically: line 69 where "early warning signals remain obscure", line 83 where "hydrological processes remain un-explored", and line 87 "few studies have examined climate similarity". This makes it hard to follow what is actually being addressed. These could be combined into one more specific statement about what the literature has not fully addressed, that directly leads in to how you address it.

Thanks for the comment. We have removed the line 69 and line 87 from the original manuscript. And we have rephrased line 83 to make the statement concise and easier to follow.

From the methods, it is clear that many different datasets and metrics are used in this study, and some sort of illustrative figure or flow chart would be really useful here. For example, you use 3 different precipitation datasets at different scales, have a temporal analysis and a spatial analysis, yearly and monthly data, and several statistical measures. It would be good to have an overview of this at the beginning of the methods section (and/or a figure) to tie *these different parts of the study together.*

We did try to summarize the use of different datasets in a tabulated form in Section 2.1, and found it not much more informative than the current text summary with links to each dataset. We believe the tie of subsequent section of results to the corresponding dataset is made self-clear when we refer to results of "global", "regional", or "city" scale, respectively. We are open to have an additional table for dataset summary if the reviewer finds it more convenient.

Figure 1: I like that you have included an illustrative example, but it comes very suddenly (I was surprised by "harvest" and thought it was somehow linked to precipitation) and is not fully explained. This example could use its own subsection and then some linkages to exactly what we are looking for in the following precipitation-based results.

Thanks for the comments. As same with the feedback provided by the reviewer #1, the harvest model is a *benchmark* problem used to illustrate the concept of *early warning* signals of critical transition, particularly the increasing trends of s.d. and AR_1. We added some transitional phrase in the context in the hope that it will not come up as a "surprise".

Related to the above, the introduction and methods section seem a lot longer than the actual results and discussion of the study. The results section would benefit from more discussion, and ties between sections. For example, many studies are brought up in the introduction, and some could be moved here to compare with your specific results. Also since the methods are heavy on different metrics, the reader could use reminders of what some of these metrics mean from a physical standpoint within the results.

The methods section is heavy because the major novelty of the current study is the introduction of new network-based metrics, while the results turned out to be a natural "proof-of-concept" and did not involve extensive discussion.
Per your suggestion, we removed some studies in the introduction into the results section to make the whole structure more balanced. For the network-based metrics, we added concise reminders of the physical meanings of them in the context to make the reading smoother.

Methods and interpretation:
I have a question on the selection of a critical transition for a given time-series: Can there be a time-series with no critical transition? Currently I get the idea that this time point is selected in every dataset as the maximum rate of change, or "abrupt change of slopes"...which does not necessarily indicate a critical transition, but adjacent years with high variability. I thought it would make more sense to define a critical transition as a step change, where magnitudes or statistical properties "before" and "after" are maximally different. In general, the definition and reasoning to identify a critical transition should be more clear. As it is, referring to changes in precip between two years as a "catastrophic transition" seems tenuous.

Firstly, yes, there are time-series with no critical transition. But then the time series will be of no use to our purpose of illustrating early-warning signals of critical transitions; thus they are naturally excluded.

The process of detecting "critical transitions" in time series of observational datasets is illustrated below in Fig. R1, using the precipitation climatology of the city of Miami as an example, where the year of transition is found to be in 1998. Other critical transition detection follows exactly the same procedure. As will be made clear, the “catastrophic transitions”, though it is determined in a specific year, does not refer to changes in precipitation between two particular **years** before and after this point (i.e. Year 1998 in Miami), but really to the transition of the *longterm historical evolution* of precipitation climatology prior to that particular year into a new trend of precipitation pattern after it.

To begin with, we bisected the cumulative precipitation data (scatters) and fitted each segment using linear regression (thus each being *quasi-stationary* with a constant slope). The year of transition is determined as the intersection of two trend lines (solid red: prior to transition and dashed red: after transition) with different slopes. In addition, Fig.R1b shows the annual (not cumulative) precipitation, where the two different means (prior to and after the transition year) are subtracted. The precipitation anomalies are then used for subsequent statistical analysis to determine the two statistical metrics (AR_1 and s.d.). We clarified the meaning of quasi-stationary in the revision.

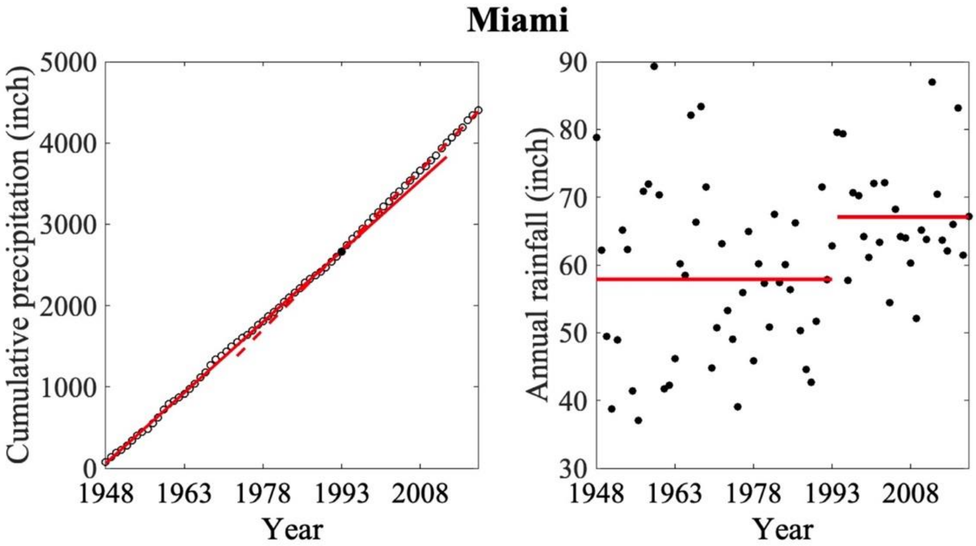

Figure R1. The statistics of precipitation in Miami (1948-2019): (a) the solid and dashed red line denote the fitted lines before and after the critical transition year (solid black dot) (b) the two different solid red lines are the mean values for each part spilt by the critical transition year.

The time windows (or data lengths) over which statistical measures are calculated is very small, e.g. 13 or 7 annual data points in all these cases. This leads to some question of statistical significance of trends in the early warning indicators. For example, you highlight several cities out of a much larger pool of data that show these increases in AR and stdev before a critical transition, but is that actually typical? Or are there many cases where these indicators are increasing where there is no critical transition (false positives), or cases where they do not well predict a transition (false negatives)?

We tested the size of moving windows continuously between 5 to 25 years as suggested by a prior climatology studies from Tsonis et al. (2007). And the window sizes of 13 years (for the global dataset) and 7 years (for the CONUS dataset) were determined because these window sizes yield the most statistically manifest results, while other window sizes give similar trends of evolution of early-warning signals but not as manifest. The same procedure was performed in an earlier work (Wang et al., 2020) and was proven using more rigorous

statistical test, such as sensitivity of Kendall's τ to window sizes, which we did carried out in this study but did not report. See Fig. 2 in
Wang, C., Wang, Z.H., & Sun, L. (2020). Early warning signals for critical temperature transition. Geophysical Research Letters, 47, e2020GL088503.

It is a very insightful question if "are there many cases where these indicators are increasing where there is no critical transition (false positives), or cases where they do not well predict a transition (false negatives)". To illustrate, we plotted the trend of changes before and after the transition for CONUS precipitation, using s.d. and clustering coefficient (as they appear more reliable than other measures). The results are shown below in Fig. R2. There does not seem to be any *false positives*, but the identified critical transitions are the only manifest positives over this sufficiently long period of time, if we consider the global trends (not the local crests or troughs). In addition, both trends relaxed after the transitions, and there does not seem to be *positive negatives* neither. Yet, there are time lags (potential hysteresis) for different indicators (e.g. the clustering coefficient plateaued slightly after the transition year and gradually relaxed). This is somehow expected as the network parameters represented the "concatenated" system behavior, and should experience some lag in response and relaxation to the critical transition. We really appreciate your insightful comments and hope this clarifies.

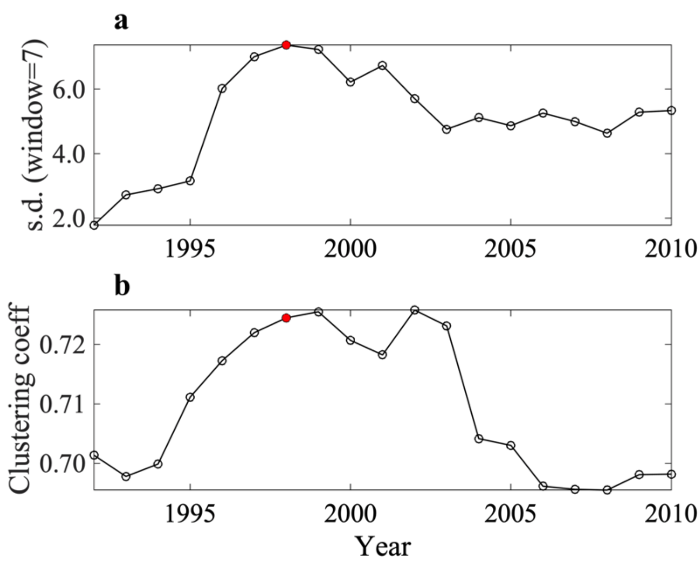

Figure R2. Two different metrices of CONUS precipitation: (a) the conventional s.d., and (b) the network clustering coefficient.

I liked the network analysis, but it was hard to go between the table of the regions and Figure 4 – could the table with the regions be made into a colored map that goes into Figure 4? This would tie these regions into the results in more directly and make them easier to discuss.

Thanks for the advice. As it was also suggested by the other reviewer, we have modified the network map and incorporated different regions corresponding to the table with different colors. The revised Fig. 4 is shown below in Fig. R3. We have revised it in the revision.

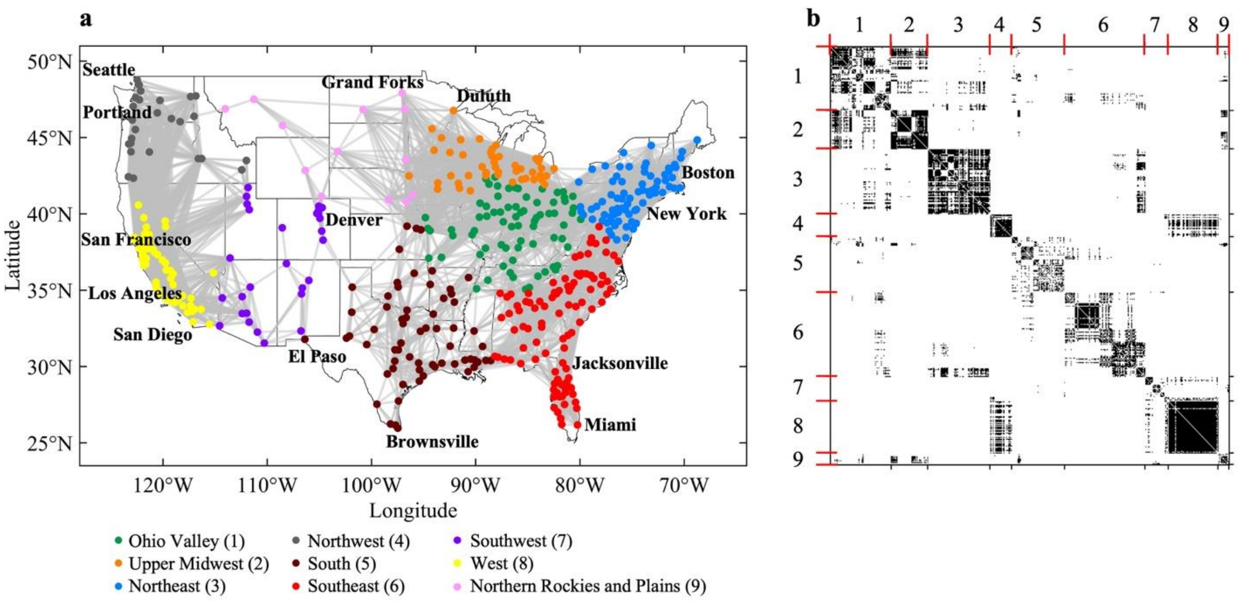

Figure R3. The precipitation network of CONUS cities: (a) The geographic map of connectivity and (b) the adjacency matrix, with A_{ij}=1 in black (connected), A_{ij}=0 in white, and red lines marking the division of nine geographic regions as shown in (a)

Minor comments:

Line 7: The first sentence of the abstract could be restructured to not start with "In this study…" as that is apparent.
We rephrase the first sentence as suggested.

Line 17: shed new light
Typo corrected.

Line 55: get rid of "aka" and explain fully. Similarly, in various places, recommend getting rid of term "viz" and explaining fully.
We replaced the "aka" and all other "viz" with the full expressions.

Line 84: cites = cities
Typo corrected.

Line 100: I don't think PET has been defined, or could use re-defining here
We defined potential evapotranspiration (PET) as it first appears in line 97.

For the AR1 as a measure (e.g. on the y-axis of several figures) – is the measure itself actually the alpha term in Equation 1? Or This was not completely clear to me at first, since the label is just "AR1". Actually, it seems like Equation 1 lines up with Equation 4, and 3 goes with 5, so perhaps this subsection could be better re-organized and less repetitive.
The alpha term in Equation 1 is the autocorrelation coefficient in the simple autoregression model, whereas the lag-1 autocorrelation AR1 is defined in Equation 4. We now deleted the name of the alpha term in Eq. 1 as the autocorrelation as it causes confusion with AR1, and the alpha term is an intermediate variable which was not used again in subsequent sections.

Hope this helps to make better clarity.

Line 172: governing dynamics are
Typo corrected.

Line 188: another statement of a research gap, not needed here really
We removed this sentence to avoid the repetition.

Line 189: emerge
The sentence was removed.

Line 205: measure
Corrected.

Line 265: You had over 100 cities in this analysis according to the methods but only introduce and discuss these 4, would be good to rationalize that small election (as they are exemplary, show the largest trends in early warning indicators, etc).
We pre-selected these 4 cities out of all 481 CONUS cities/towns as representative to their distinct geographic and climatic conditions, *before* we carried out the actual early-warning signal analysis. To verify, we conducted the same analysis to other cities: either they exhibit similar indicators when transitions exist, or no manifest signals detected. The first case was dismissed as repetitive, and the second due to insignificance. We added rationale to explain the choice in the revision.

Line 288: "highly assortative with large modularity" needs more expansion
We added physical explanation to this phrase in the context.

Line 297: responds, presages
Typos corrected.