Response to Dr Chris Boulton (Reviewer#1)'s comments

on paper "Tipping point analysis helps identify sensor phenomena in humidity data"

by *V. N. Livina, K. Willett, and S. Bell*

We would like to thank the Reviewer for the positive opinion about novelty of our approach.

Comment 1. *My main suggestion would be to include variance analysis too as an EWS.*

Response. We have included variance indicators in the appendix (figures B1-B22) and corresponding text.

Comment 2. *With an expected increase in the signal-to-noise ratio as newer sensors are included, the AR(1) should increase as observed in your work, but the variance would decrease at the same time. It seems that the combination of these would aid your work as it could rule out any "natural" change in the system itself. Smith et al. (2023) shows an example of this.*

Response. We have added the suggested reference. Accordingly, we added discussion of variance and autocorrelations and their suitability for sensor change detection.

Comment 3. *I think it's important to highlight how these EWS are affected by the changes in measurement circumstances, from the viewpoint of the other stages in tipping point analysis such as prediction and the chance of false positives.*

Response. We have added the following text in the Conclusion: "We have demonstrated that autocorrelations are both sensitive and robust in detection of known sensor changes. While not all metadata of measurement circumstances may be available, autocorrelation indicators provide the tool for scanning datasets for such changes, whose rapid development may help destinghuish them from the false positive signals of climatic changes. This should be taken into account in development of predictive techniques based on EWS signatures."

Comment 4. *I also had some slight confusion about how and why ERA5 data is used at all. It would be good to explain that the large gaps in data (I assume) come from the station data and not ERA5. Also, why can't the station data measurement just be used since the method used to create the reanalysis will not have the same issues?*

Response. We have added the following explanation: "While there were station-level measurements of pressure (variable 'stnlp'), these records were often short or patchy and the data quality was quite poor. Instead of these station pressure measurements, we used a climatological surface pressure from the nearest gridbox of the ECMWF ERA5 reanalysis product."

Comment 5. *Lines 77-80: This section is slightly confusing. I think it's suggesting that the AR(1) has to reach 1 as a critical value but Kendall's tau is also mentioned. I would be wary of saying that AR(1)=1 is critical when detrending has occurred in the time series it is calculated on as this alters the absolute value of AR(1). I would also refer to a "time series" of the indicator throughout rather than "curve".*

Response. We have reformulated this sentence as follows: "If the indicator rises to a critical value of 1 (as one is the maximal value of normalised autocorrelation, provided no detrending or filtering is applied to the input data), this is a signature of early warning signal of critical behaviour. However, when any pre-processing is applied to the input data, this is likely to reduce autocorrelation values, and indicator may not reach the maximal value 1. In this case, the important property of the indicator is its monotonic increasing trend. Such a monotonic trend of the EWS indicator can be estimated, for instance, using Kendall rank correlation."

We have also replaced "curve" with "time series".

Comment 6. *Lines 91-94: This section may not be needed. The potential plots here have not been estimated, for example.*

Response. Because we mentioned detection and forecasting in the overview of tipping point analysis in lines 67-69, we prefer to keep their explanation in lines 91-94. To address the comment, we have added the following sentence at the end of the paragraph: "In the current paper, we do not perform detection and forecasting stages of tipping point analysis and focus on anticipating tipping only EWS)."

Comment 7. *Line 113: What was wrong with the station that wasn't used?*

Response. Data quality (large gaps) was the main issue with the dataset. We mentioned this, but to clarify it better, we reformulated it like this: "After removing data before the largest gap in each of them, 55 stations were selected (the 56th being too short after truncation)."

Comment 8. *Fig. 1: What window length is being used here? I would also centre the x-axis on -1 to 1 in both panels.*

Response. We have adjusted the x-axis of the left panel of Fig.1. The size of the sliding window was actually mentioned in the caption of Fig.3, but the reviewer is right, it should be stated earlier. We have added in the caption of Fig.1: "The window size for calculation of indicators was 10% of time series length (as the records had different lengths)."

Comment 9. *Page 6: I think it's important to say what the actual window length was that was used, and personally I would suggest trying a longer window length as well*

to see how the results contrast given the discussion on this page. Also, does the BCP analysis require any a priori input on the change of form that is searched for (e.g. looking for a certain number of changepoints)? If so, this should be stated.

Response. 10% of the dataset length is a compromise of sensitivity of indicators and sufficient aggregation of data within a window. Because we are interested in instrumental changes, the size of the window cannot be taken too large, as data aggregation in a large window would suppress the signal of such changes, and moreover, a large window would increase the uncertainty in timing the event of the instrument change.

We did not use any asumption on the number of change points, the BCP technique has no such parameter.

Comment 10. *Fig. 2: I feel like this could be better represented by a continuous blue to red scale rather than the size of the circles as I find it hard to distinguish between the sizes (except the blue ones look small).*

Response. We have replaced Fig.2 with a more clear version, with updated explanation in the caption.

Comment 11. *Fig. 3: The red box is not defined in the figure caption.*

Response. We have added the following explanation in the caption: "The red boxes denote the intervals of transitions: in the input time series the changes are not visible, whereas in the detrended series one can notice the change of pattern, which is then clearly detected by the EWS indicator."

Comment 12. *Line 180: There are no detections in the 1980s in Figures 7 or 8.*

Response. We have removed "1980s" in this line.

Comment 13. *Fig. 6-8: It would be good to add the red crosses on the bottom panels each time to see how they match up more clearly.*

Response. We have modified these three figures and the captions accordingly.

Comment 14. *Line 191: The shifts in the 1980s happen in the Appendix but not in the figures in the main paper.*

Response. We added the comment in brackets: "(see figures in the main text and in the Appendix)".

Comment 15. *Appendix Table: Is this for detections above 0.8 with the BCP analysis? If so, it should say in the caption.*

Response. We have added this in the caption.

We hope that the revised manuscript is now suitable for publication in the "Geoscientific Instrumentation, Methods and Data Systems" magazine.

Yours sincerely,
V. N. Livina, K. Willett, and S. Bell