The manuscript "Early warnings and missed alarms for abrupt monsoon transitions" by Thomas et al. poses the question whether there were bifurcation induced abrupt changes of East Asian monsoon intensity during the penultimate glacial cycle. They address this question by analysing trends in autocorrelation and variability in speleothem records from Chinese caves because linear stochastic theory suggests an increase of these properties before a bifurcation.

**General comments**

I think that this question and approach are very interesting and reasonable and within the scope of Climate of the Past. The analysis of different potential Tipping Elements using models, reconstructions and observations is an important issue in earth system science, and the authors' approach is a step in this direction. However, I also find it difficult to understand how and why different statistical methods are applied throughout the paper, what the results are, and how the authors interpret these results. I would suggest to explain these things more explicitly using a clearer structure and wording to make the results more transparent for readers unfamiliar with the technical details.

My main concern in terms of contents is whether the interpretation of the authors is fully justified by the results of the study. I get the impression that the record the authors analyse does not show "early warnings" (except before one of the abrupt transitions). Nonetheless, the authors maintain the interpretation of these abrupt shifts as bifurcations with the argument that the data is too scarce to see any signal. I would assume that other explanations are equally possible and I suggest to highlight such alternatives more clearly in the paper. The authors also fit a simple stochastic model to the data (which they call non-stationary potential analysis), whose parameters are coupled to the solar insolation at 30 N and which features bifurcations. In artificial time series from this model, significant early warning signals appear.

I may have misunderstood the logic of the paper but I get the impression that this approach is flawed by circular reasoning. Is the model the authors fit to the data not built in a way that it must show such signals? In this case the question arises whether the original record is adequately described by the model. What is needed in my eyes is some kind of statistical test which allows to falsify a model or an hypothesis, e.g. that the framework of bifurcation theory (or some alternative explanation) is inconsistent with the data. Although the authors perform an ensemble of time series with their stochastic model and get significant results, the original data incorporated in the model was too badly resolved to see early warnings. I wonder why the fit of model parameters can be more precise than the scanning of the original record. Why is the potential stochastic model needed in the paper at all? I do have the feeling that the authors are somehow aware of this and follow a certain logic, but in order to understand and assess this logic it should be made much more transparent in my opinion.

It would also be interesting to see physical arguments for the author's interpretation of the abrupt monsoon shifts, although I understand that this is not meant to be the focus of their paper. If there are bifurcations, what physical properties are involved, and what could be the different timescales the authors mention in the introduction? As the atmosphere adjusts very quickly to its boundary conditions, what is the element in the monsoon system that would show a memory in such a way that the authors expect to see it in palaeorecords? Another aspect in this context is the reference to the concept model by Leverman et al. (2009) and Schewe et al. (2012). The authors introduce this model as consistent with the bifurcation hypothesis and state that "It has been hypothesised that ... the EASM exhibits two stable states with bifurcation-type tipping points between them (Schewe et al., 2012)". However, I take it from these publications that the monsoon is a "switch" in their model, where the on or off state is determined by a moisture threshold. I wonder why such a threshold should be consistent with the bifurcation hypothesis and why the authors expect early warning signals. It seems to me that the whole "off" state and the small hysteresis which exists in the model has been artificially built in at the threshold (Schewe et al., 2012) and is not an emergent result of the moisture advection feedback. Furthermore, the model only describes equilibrium solutions, but involves no timescales. I therefore do not find it compelling that the concept model is really in agreement with the bifurcation hypothesis, at least not without additional arguments.

**Specific comments**

Abstract

- I suggest not to cite other papers in the abstract, at least it is not very common.

- The abstract mentions the conceptual Levermann/Schewe model, "model simulations" (referring to the author's stochastic model), and the detection of critical slowing down. It should be clarified that the Levermann/Schewe model is not the one the authors performed simulations with, and the early warnings are found in their model, not in the data itself. Also, what is "consistent with long-term orbital forcing", and why is it a result rather than an ingredient to the stochastic model? These aspects are examples why I find the paper hard to read and suggest to use a more precise wording throughout the paper.

Methods

- I wonder whether the paper would be easier to understand if the details of each method would be explained directly when it is applied. In the introduction or methods section one could instead explain the general logic of the methods and their role in the paper more generally and briefly.

- Is the relation between the d18O record and monsoon intensity not time dependent? What are the uncertainties in this regard? Is there a quantitative reasoning behind the authors' statement that dating uncertainties do not affect the results?

- p. 1317, line 2: "we use an insolation latitude". At this point in the paper, it is not clear at all why and how the authors use the insolation.

Data selection

- p. 1318, line 1 (and elsewhere): What is "tipping point analysis"?

- p. 1318, line 2, 3: what is meant with "clear climate proxy" and "adequate length"?

- p. 1318, line 5: "Fig. 4 and 5 show that density of data points do not change" (sic). How do I see this in the figures? I find it hard to understand them.

Tipping point analysis

- p. 1318, line 18: "A sensitivity analysis was undertaken...". Is this Fig. 7? Then why not refer to it?

- p. 1318, line 20-27: I suggest to move such general explanations to the introduction.

- p. 1319, line 1-4: Why is this technical discussion relevant in this context?

Non-stationary potential analysis

- I don't clearly see from the paper how the parameters of the model are estimated. Is this estimate unique (including the noise level), and what are the uncertainties? It could also become clearer here why the potential model is used at all.

- p. 1321, line 1-15: These steps are not easy to follow and I find them too vague. For example, "we manipulated the noise level", "we linearized the solar insolation", ''the same iteration of the model was used", ... I also cannot follow the argument why different sampling steps of the data are necessary.

Results and discussion

- p. 1321, line 22: "a ... potential model was fitted". How? And how was it "modulated by the solar forcing"?

- p. 1322, line 1-5: Do these clear trends in autocorrelation and variance concern the artificial time series or the record? I suggest to make this distinction clear every time such trends are mentioned because I consider it important for the conclusions we can draw from this study.

- p. 1322, line 27-29: "To help interpret these results we applied the potential model...", "explaining the high degree of synchroneity between the transitions and solar forcing". I find it impossible to judge if this is really a confirmation of a hypothesis or just the result of how the model was tuned, especially because not much details are provided on the tuning. How hard would it be for the potential model to clearly contradict the bifurcation hypothesis? I think that these aspects are probably the most important to interpret the results of the study and should be made much clearer.

- p. 1323, line 3-4: "There are instances when bifurcations are not preceded by slowing down". This should be explained more precisely as it seems in conflict with what is stated in the introduction.

- p. 1324, line 3-4: The fact that palaeodata often has insufficient resolution for statistics like "early warnings" is a somewhat trivial remark and in my eyes no specific result of this paper.

- p. 1324, line 15 - end of section: It would be interesting to know how these hypotheses relate to the bifurcation hypothesis? Do they exclude each other, i.e. could this represent an alternative hypothesis to the authors' bifurcation scenario? I think these possibilities could be explained right away in the introduction instead in the very end of the paper. How should we proceed to eliminate some of the possible explanations and do the authors suggest that early warnings can play a role?

Conclusions

- "We detect a fold bifurcation structure... in data". I do not agree that this is what the authors do. As I understand their paper, they look for (but hardly find) indicators of slowing down in the data. If there were such indicators, how do the authors know they result from slowing down, and why must it be due to a fold bifurcation?

- "Our results have important implications..." Which implications?

- "a failure to identify slowing down does not preclude a bifurcation". Given the low resolution of the data this is a somewhat trivial statement. I suggest to highlight in the conclusions what the results mean for the potential mechanism of the abrupt shifts.

Figures and References

- The Figures do not seem to be cited in order.

- I suggest to reduce the number of figures. For example I wonder if all panels in Fig. 5 and 6 are needed. Also, it is not always clear to me what they show. What does the density data in Fig. 5 and 6 show and mean?

- How are the p-values in Fig. 5 and 6 calculated? This seems to be some kind of test result (implicitly mentioned on p. 1319, line 5-6?; p. 1322, line 15-17?), though at odds with the approach of the histograms in Fig. 8 and 9. As the analysis is about autocorrelation in the data, it seems contradictory to use a test which assumes independent data points, but the authors do not comment on this.

- The references mostly consist of very recent papers but sometimes ignore the original work. I suggest to also give credit to the more original papers. For example, the Levermann (2009) model seems to be identical to the more often cited Schewe et al. (2012) model. Also, the effect of slowing down was first introduced to climate research by Kleinen et al. (2003) and Held and Kleinen (2004). However, only the more recent work by Dakos, Lenton and Scheffer is cited.

Held, H.; Kleinen, T. (2004): Detection of climate system bifurcations by degenerate fingerprinting. Geophysical Research Letters, 31, L23207.

Kleinen, T., H. Held, and G. Petschel-Held (2003), The potential role of spectral properties in detecting thresholds in the Earth System: Application to the thermohaline circulation, Ocean Dyn., 53, 53– 63.