

## Review #1

**Reply:** We thank the reviewer for their careful reading of the manuscript and their constructive remarks. We have taken the comments on board to improve and clarify the manuscript. Please find below a detailed point-by-point response to all comments (reviewers' comments in black, our replies in blue). N.B Since the reordering and restructuring of the manuscript was substantial, we have written bullet points of our major changes to the manuscript, rather than including a 'track changes' document. Line numbering refers to the revised manuscript, attached as a supplement to the Editor Response.

### Major changes:

- Altered the abstract to reflect the new structure of the manuscript.
- Added a clearer 'Aims' section.
- Provided more detail on the potential model.
- Altered the structure, separating the Results and Discussion sections and ensuring a consistent structure in the added sub-sections throughout the manuscript.
- The manuscript now follows a more logical format, with tipping point analysis of the entire speleothem sequence followed by the potential analysis results.
- Reordered the figures and added a paragraph of text to explain each figure sequentially.

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.

**Reply:** We have added further detail, and a better structure, to explain our choice of methods and make these explanations more easily understandable for those readers unfamiliar with the techniques used. Specifically, we apply two main methods to help determine whether the East Asian Summer Monsoon is characterised by a bistable system, with bifurcations between the strong and weak monsoon regimes, and whether 'early warning signals' of these bifurcations can be detected. Tipping point analysis uses techniques from Kleinen et al. 2003, Held and Kleinen 2004, Dakos et al. 2008, and many others, to identify the characteristic fluctuations that occur in data prior to a bifurcation, caused by a phenomenon called critical slowing down. Non-stationary potential analysis is used to create a simple model of the monsoon based on the speleothem data. We believe our substantial restructuring has helped make the aims, methods and results more transparent.

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.

**Reply:** An alternative interpretation of the speleothem isotopic data is a linear response to orbital forcing. However, the abrupt nature of the data in comparison to the sinusoidal forcing argues against this, strongly implying a non-linear threshold. This has been emphasised in the revised manuscript (lines 47-51). By examining both palaeoclimate data as well as model-derived data, the presence of a bifurcation within the monsoon system appears to be the most plausible. We have however explained our aims more clearly (lines 78-83), and justified our conclusions.

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.

**Reply:** The potential model is not built in a way that it necessarily exhibits bifurcations and hysteresis behaviour. It also has the possibility of two states always being available and the transitions being noise-induced with or without stochastic resonance. The parameter estimation reveals which mechanism is better supported by the data.

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?

**Reply:** As acknowledged by the reviewer, a thorough discussion of the physical mechanisms behind these abrupt monsoon shifts is really beyond the scope of this paper. However, we agree that the manuscript would benefit from a brief discussion of the physical mechanism(s). We speculate on possible mechanisms and focus on one likely contender: the moisture-advection feedback, as described by Schewe et al. (2012). We also direct the reader to other papers such as Zickfeld et al. 2005 that cover this in further detail. See line 62-73 "*A minimum conceptual model of the East Asian Summer Monsoon developed by (Zickfeld et al., 2005), stripped down by Levermann et al. (2009) and updated by Schewe et al. (2012), shows a non-linear solution structure...*"

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.

**Reply:** Several papers refer to the Asian monsoon system being an important 'tipping element' in the Earth's climate system (e.g. Lenton et al. 2008; Zickfeld et al. 2005; Donges et al. 2015). The Schewe et al. (2012) paper directly refers to a 'critical threshold', and 'threshold behaviour', which is directly

relevant to our bifurcation hypothesis. Indeed, Figures 5 and 6 in this paper show a bifurcation structure in the conceptual model, which implicitly infers a bifurcation. In this paper, the critical point refers to the bifurcation point. Zickfeld et al. 2005 also describes the Indian Summer Monsoon (which is closely linked to the East Asian Summer Monsoon) as being a multistable system, with saddle-node bifurcations. We have thus amended the revised manuscript to explain more fully the development of this conceptual model, and moisture advection feedback: *'A minimum conceptual model of the East Asian Summer Monsoon developed by Zickfeld et al. (2005), stripped down by Levermann et al. (2009) and updated by Schewe et al. (2012), shows a non-linear solution structure with thresholds for switching a monsoon system between 'on' or 'off' states that can be defined in terms of atmospheric humidity – in particular, atmospheric specific humidity over the adjacent ocean (Schewe et al., 2012). Critically, if specific humidity levels pass below a certain threshold, for instance, as a result of reduced sea surface temperatures, insufficient latent heat is produced in the atmospheric column and the monsoon fails. This moisture-advection feedback allows for the existence of two stable states, separated by a saddle-node bifurcation (Zickfeld et al., 2005) (although interestingly, the conceptual models of Levermann et al. (2009) and Schewe et al. (2012) are characterised by a single bifurcation point for switching 'off' the monsoon and an arbitrary threshold to switch it back 'on'). Crucially, the presence of a critical threshold at the transition between the strong and weak regimes of the EASM means that early warning signals related to 'critical slowing down' (Dakos et al., 2008; Lenton et al., 2012) could be detectable in suitable proxy records.'* (lines 62-76).

#### Specific comments

##### Abstract

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

**Reply:** We have removed the citations from the abstract (lines 15-32).

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.

**Reply:** We have amended the revised manuscript to be clearer that our model simulations are separate to the Levermann/Schewe model, both in the abstract and in the introductory section. We have also ensured that we explain which data we find the early warning signals in i.e. whether it is from the palaeoclimate data or the model output. We have also revised the abstract to clarify the wording.

##### 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.

**Reply:** We understand the reviewer's viewpoint here, and we do agree that a brief introduction to the general logic of the methods could be advantageous in the introduction to provide more context. We have added small section (lines 78-83) to this effect. Further subheadings (*'Detecting early warning signals'*, *'Missed alarms'*, and *'Using speleothem  $\delta^{18}O$  data as a proxy of past monsoon strength'*; lines 85, 106 and 116 respectively) are used to explain what our intended aims of the paper are.

We have restructured the methods section to clarify the approaches taken, which now reflects the structure of the results and discussion.

Is the relation between the  $\delta^{18}\text{O}$  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?

**Reply:** The relationship between  $\delta^{18}\text{O}$  and monsoon intensity has a proven relationship on centennial to millennial timescales within speleothems in southeast China (e.g. Wang et al. 2008; 2012; Li et al 2013). As a result we do not investigate this aspect here. We are unsure where the author is referring to about dating uncertainties not affecting the results. The U-Th ages provide a robust chronological framework for the speleothem sequences investigated.

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.

**Reply:** We thank the reviewer for highlighting this point. We have reordered the methods section and to ensure that the relevance of the insolation latitude is explained in the appropriate place (now line 295). We have also added subheadings to improve the structure of the methods section.

#### Data selection

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

**Reply:** We have provided additional information to explain 'tipping point analysis' (*'This analysis aims to find early warning signs of impending tipping points that are characterised by a bifurcation (rather than a noise-induced or rate-induced tipping e.g. Ashwin et al. (2012)). These tipping points can be mathematically detected by looking at the pattern of fluctuations in the short-term trends of a time-series before the transition takes place'*; line 86-91). As mentioned above, we have restructured the methods section to ensure that each method is introduced in the appropriate context, and added subheadings to help this structure.

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

**Reply:** We have replaced the phrase 'clear climate proxy' with '*a measure of climate*' (line 140). By 'adequate length', we mean of sufficient length to enable a robust analysis over a sliding window (line 141-142); however we cannot be much more specific since the exact length inevitably depends on each record.

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.

**Reply:** We apologise for any confusion. We should have been referring to Figures 5 and 6 here; the figures have been reordered in the revised manuscript and are now are Figures 6 and 7. The density of the data points specifically refers to panel c) – this has been amended; Figure 6 (now Figure 7) panels are now also labelled a to e. Figures 6c and 7c shows the density of the data over time; this depicts how unequally spaced the data are. If the data were equally spaced, Figures 6c and 7c would depict a straight horizontal line. Figures 6c and 7c therefore show that although the data is not exactly evenly spaced, the density of the data points (how equally they are spaced) does not change significantly along the record; this is now better explained (*'In addition, since time series analysis methods require interpolation to equidistant data points, a relative constant density of data points is important, so that the interpolation does not skew the data. The speleothem  $\delta^{18}\text{O}$  records that we have selected fulfil these criteria, as described in more detail in section 2.1'*; lines 142-145; line 154).

#### Tipping point analysis

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

**Reply:** This did indeed refer to Figure 7; we have updated the revised manuscript and re-ordered the figures; this is now referred in the manuscript as Figure 5 (line 359).

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

**Reply:** As suggested by Reviewer #3, we have reordered the methods section and instead separate a general explanation of tipping points and the detailed method that we use in our paper. We believe that this makes the manuscript substantially clearer. The introduction now includes a sub-section on '*Detecting early warning signals*' (lines 85-104).

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

**Reply:** We feel that it is important and relevant to at least briefly highlight that there is some dispute regarding whether autocorrelation and variance should increase together or not. Our substantial restructuring provides additional context for this particular point (now within the '*Detecting early warning signals*' sub-section; lines 96-99). This technical discussion also provides context for the discussion of the proportional positive trends in autocorrelation and variance in both autocorrelation and variance in the Results section (lines 383-385).

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.

**Reply:** The parameters of the potential model are estimated according to maximum likelihood. The procedure is now described in more detail in the revised manuscript (lines 221-290). The parameter estimates and the noise level are unique and the uncertainties are very small.

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.

**Reply:** The different sampling steps are necessary to investigate their effect on the indicators of critical slowing down. There has been little discussion of the importance of the sampling steps of palaeoclimate data; this paper presents the first examples of how autocorrelation and variance are affected by changing the sampling step. In particular, the memory of the system (measured by autocorrelation) is less represented under sparser sampling. The fact that the same iteration of the model is used is important since this eliminates noise as a factor in the sampling step changes. We have changed "we manipulated the noise level" to "*we manipulated the noise level of the model by altering the amplitude of the stochastic forcing ( $\sigma$  in Equation 1)*" (lines 318-319). In terms of the linearization of the solar forcing, we have added a sentence to explain this: "*This approach was preferred rather than using a sinusoidal forcing since early warning signals are known to work most effectively when there is a constant increase in the forcing*" (lines 313-315).

Results and discussion

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

**Reply:** The potential model is now explained in more detail in the revised manuscript (lines 221-290).

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.

**Reply:** We agree that it is important to make this distinction clear; in this particular case the sentence directly refers to the Sanbao Cave record. When we use the phrase the Sanbao cave record, this means that we are referring to the palaeoclimate speleothem data. When we refer to the model simulations, we are referring to the data derived from our model. We have ensured that we are clear in this regard in the revised manuscript (helped also by the restructuring of the manuscript).

p. 1322, line 27-29: "To help interpret these results we applied the potential model...", "explaining the high degree of synchronicity 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.

**Reply:** The rationale, construction and estimation of the potential model are now explained in more detail. The model does provide a test between alternative mechanisms.

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.

**Reply:** We agree that this wording could be unclear. We have amended this to '*there are instances when critical slowing down cannot be detected/recorded prior to a bifurcation*' (lines 491-492). As with most statistical techniques there are a number of circumstances when the theory does not always tie with reality. Although as stated in the introduction, critical slowing down theoretically precedes a bifurcation, there are indeed some occasions when this critical slowing down is not recorded in the data. There can be many reasons for this, including a high noise level, and an insufficient sampling resolution. We have explained this more precisely in the revised manuscript (lines 491-496).

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.

**Reply:** The tipping point literature rarely discusses data resolution as an important aspect of early warning signals, perhaps largely because the majority of the tipping point literature analyses data from models or observational data, which is generally high resolution. We feel that although to a degree this is to be expected, our results actually illustrate how data resolution changes will affect the indicators of critical slowing down, rather than an arbitrary discussion of the limitations of resolution. We believe that highlighting this point will help to inform discussions of the limitations of the early warning signals, particularly when working with palaeoclimate data.

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?

**Reply:** We fear the reviewer may have misunderstood us here. These hypotheses relate to possible reasons why the bifurcation is detected during termination II. They are not alternative hypotheses to bifurcation. The key point here is during earlier bifurcations no early warning signals were detected and we discuss here possible reasons why this might be so.

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?

**Reply:** The fold bifurcation structure is detected by means of the potential model. We agree that slowing down is not necessarily linked to a fold bifurcation; it may also be associated to other bifurcation types.

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

**Reply:** We have elaborated on this sentence to be clearer about the implications: "*Our results have important implications for identifying early warning signals in other natural archives, such as the importance of sampling resolution and the background state of the climate system (full glacial versus termination).*" (line 543-545). At present, these aspects are overlooked, and we feel that this is an important aspect to highlight in the conclusions.

- "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.

**Reply:** We agree with the reviewer here and have revised this section to change the implications (lines 543-549), as described above to highlight the significance of the background climate state; the significance of which previous work has not identified.

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

**Reply:** We have ensured that the figures are cited in order in the revised manuscript.

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?

**Reply:** Whilst we agree that there are a large number of figures, we believe that these help to take the reader through our results. We have reordered the figures, which we believe now vastly improves the clarity, and explained each figure in more detail (e.g. lines 356-361, lines 379-387, lines 398-402, lines 414-422, lines 435-439, lines 452-456, lines 463-470). The density data in Figure 5 and 6 (now 6 and 7) demonstrate the comparable sampling resolution across the sequences, stressing the absence of early warning signals is not an artefact of the data. However, to help slightly with the reduction of figures we have removed the lower panels from Figure 11, as these results are easily explained in the text, and merged Figures 8 and 9 (now Figure 5) to remove unnecessary duplication.

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.

**Reply:** The p-values are calculated as was discussed on p.1319 and p.1322; however we did not refer explicitly to p-values in this description; we have amended this in the revised manuscript to increase clarity (now lines 200-211). The p-values themselves do not refer to autocorrelation directly; they are based on the Kendall tau value of the trend in autocorrelation over 1000 realisations, as displayed in the histograms in Figures 8 and 9 (now Figure 5). This method has been used in several papers; we also include citations to these papers (e.g. Dakos et al. 2012) in our further explanation: '*This method for assessing significance of the results is based on Dakos et al. (2012a)...*' (lines 202-203).

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.

**Reply:** We have included citations to the Held and Kleinen (2004) and Kleinen et al. (2003) papers to acknowledge this original work in addition to the more recent papers (line 93). We do cite Levermann et al. (2009); however we tend to cite Schewe et al. (2012) more often due to the advances that this paper made to Levermann (2009) in terms of the application of their model to speleothem data, and the notion of a 'critical humidity threshold'. However, we have added a sentence to increase clarity about the Levermann/Zickfeld/Schewe papers: '*A minimum conceptual model of the East Asian Summer Monsoon developed by Zickfeld et al. (2005), stripped down by Levermann et al. (2009) and updated by Schewe et al. (2012), shows a non-linear solution structure...*' (lines 62-66).