Authors' response to reviewers' comments on: "Reduced Carbon Cycle Resilience across the Palaeocene-Eocene Thermal Maximum" (Ms. Ref. No.: cp-2018-57)

Dear Yannick Donnadieu and Reviewers,

5   Thank you for the timely reviews of our manuscript "Reduced Carbon Cycle Resilience across the Palaeocene-Eocene Thermal Maximum". Please find below our detailed responses to the reviews. We include the original comments and our response under each point in **bold** text and with line numbers referencing the revised manuscript with changes marked.

We hope we have sufficiently answered your queries in our response.

10   Yours sincerely,

David I. Armstrong McKay & Tim M. Lenton

Response to Reviewer 1 (R1) Valerie Livina:

The paper by McKay and Lenton "Reduced carbon cycle resilience across the Paleoscene-Eocene thermal maximum" investigates paleo data for early warning signals prior to known paleo events and discusses the system resilience in geochemical context. The paper is well-written, applies previously developed and tested techniques, and can he published after a revision.

I suggest the following modifications.

The authors could add definitions and references to the ideas of tipping points and resilience.

The definitions and descriptions of tipping points and resilience have been clarified and additional references added (page 9 line 25 & page 10 line 1).

In Figure 1, in the upper panel the time series have local minima at 55Ma, whereas in the enlarged bottom panel these minima are located at 56Ma. Why?

The data in the top panel is from the Zachos compilation (2008), which uses an older age model for the data in this section, whereas the bottom panel uses higher-resolution data with an updated age model (Westerhold et al 2015) across just the study period. To maintain consistency with published records we have left them as-is, but we have clarified the difference in age models and PETM date in the caption.

The paper does not include the reference [Held and Kleinen, GRL 2004], which was the first to apply the lag-1 autocorrelation as an early warning indicator of a climatic bifurcation. The authors should cite it in page 2, line 25; in page 4, line 6, and discuss the H&K contribution as pioneers of this technique in climatology.

We indeed should have cited Held & Kleinen (2004), and have now included and discussed this reference where required (page 9 line 26, page 11 line 24)

Variance does not necessarily increase prior to a bifurcation, see the counter-example in [Livina et al, Physica A 2012]. Changes in skewness and curtosis are often consequences of the asymmetric effects due to the appearing or disappearing potential wells. This can be discussed in more detail in page 2.

We have added discussion of counter-examples to increasing variance, skewness, and kurtosis (page 9 line 31; page 10 line 6).

It is not clear to me why the authors separate the cases of rolling windows for indicators and binned indicators. As I understand it, the binned indicator is a particular case of a rolling window, when the shift of the window is equal to the size of the window. Why do authors stress that, for instance, DFA is a binned indicator? It can be estimated in rolling windows just the same. Accordingly, I do not understand the comment in brackets in page 5 , line 21, which separates rolling window and "metric-based".

The binned metrics are indeed a special case of the rolling window, although with differing sized bins for pre-PETM, PETM-ETM2, and post-ETM2. The key aspect is that in contrast to the rolling window run over the whole dataset for the binned analysis the data from the events themselves are removed and so not included

in these bins, avoiding the indicators being biased by the events themselves. However, we agree that the way this section is currently phased is a bit confusing, so we have rephrased this section to make it clearer (page 13 lines 3-6). The comment on page 5 line 21 has also been clarified (page 13 line 17).

You are also right to say that DFA can also be done on a rolling window along with the other EWS indicators.
5 **This analysis was originally excluded due to time constraints in the initial project (with AR1 used as the main 'memory' indicator and DFA added later to just the binned section), but we have now run and included this additional analysis in the revised manuscript. These analyses strongly support our original AR1 results and interpretation, and discussion of these results have been included where necessary.**

When using several sizes of windows for an indicator (the figures in the supplement), the authors could do this
10 not just for three sizes but for a range of windows, with estimation of uncertainty in the indicator curve - see the example of such indicator curves in [Livina et al, JCSHM 2013].

A continuous range of rolling window length rather than just a comparison of 25 / 50 / 75% windows is also advantageous, and so we have also included a more robust sensitivity analysis of rolling window size (presented as plots and histograms of kendall tau values for each 1% increase in rolling window sizes
15 **between 25% and 75%) in the Supplementary Material (Supp. Figures S3 & S4) and mentioned in the text (page 12 line 16).**

In page 4, line 6, "the methodology was first outlined in [Held and Kleinen, GRL 2004], then used in [Livina and Lenton, GRL 2007]", etc.

Clarification incorporated (page 11 line 24)

20 In page 4, lines 17-18: while trends may be not the focus of the analysis, their removal reduces the value of the lag-1 autocorrelation, and this should be kept in mind in analysing the indicators.

We have added a mention of this to the methods discussion (page 12 line 6)

For the DFA reference in page 5, line 14, please use instead of [Lenton et al 2012b] the original DFA reference [Peng et al, PRE 1994].

25 **Citation incorporated (page 12 line 23)**

In page 1, line 17, what do the authors mean by "differing"?

By "differing carbon cycle dynamics preceding the PETM and ETM2" we refer to the drivers of variance being different for each event, with an increase in jump intensity (and overall conditional variance) in the d13C prior to the PETM versus an increase in diffusion (and a decrease in jump intensity) prior to ETM2. This
30 **difference suggests potentially differing carbon cycle dynamics prior to each event, which we discuss in Section 3.3. However, this could be phrased more clearly in the abstract and so will clarify this line (page 8 line 19) along with Section3.3 (page 16 line 18) in the revised manuscript.**

In page 8, lines 2-3, to avoid line break between 2 and My, use LaTeX command~

This manuscript was composed in Microsoft Word, but we have adjusted this line to avoid the line break (page 16 line 7).

Response to Reviewer 2 (R2) Pierre Maffre:

5First, I have to admit that this study uses statistical methods for time-series analysis which I am not very familiar with. So my remarks may be naive. Nevertheless, care has been taken to explain how to interpret the results given by these methods, as well as the limitations in the eponymous section. This is appreciable for a unfamiliar reader.

As far as I know, this methodology of early warning signals for detecting tipping points is relatively new in
10paleoclimate studies. It makes this study all the more interesting and the results worth to be broadly communicated. The use of different indicators (distribution moments, autoregressive model, detrended fluctuation analysis, drif-diffusion-jump model) and the sensitivity analysis contribute to strengthen this study. Moreover, this methodology has not been applied to paleorecords earlier than Quaternary according to the cited literature (Dakos et al., 2008, Lenton, 2011, Lenton et al., 2012a, 2012b). Despite the difficulties of using
15data as far in the past of the Earth ˘ AˇTthat are clearly mentioned in the manuscriptˆ AˇT this study constitutes therefore a significant advance. I would suggest that this pioneering application should be highlighted in the main text.

Here follows my comments.

General remarks:

20This study deals with the reduction of climate and carbon cycle resilience with very little mention of the processes responsible for this (lack of) resilience. It is acknowledge Page 6 lines 4-6 that "most indicators do not reveal exact information about the nature of the transition itself" and I understand it is not the aim of the study. Nonetheless, for periods of time as extended as the one covered by this study (5 and 8 million years, which is substantially more than the previous EWS studies mentioned), it is commonly hypothesized than silicate
25weathering is the feedback that stabilize Earth climate (Walker, Hays and Kasting (1981) J Geophys Res 86, 9776; Berner and Caldeira (1997), Geol, 25, 955; François and Goddéris (1998), Chem Geol, 145, 177-212). The authors should explicitly say if they aim at tracking carbon-climate resilience due to silicate weathering feedback (or more generally feedbacks in geological carbon cycle), or only "shorter-term" climate resilience (for instance, the processes mentioned Page 3, lines 21–23), or both.

30**You are right to say that it is important to be clearer about which long-term carbon cycle or climate processes may be implicated in our analysis, and the silicate weathering feedback is indeed the most important of these for the geological carbon cycle (along with long-term changes in the burial rates of organic carbon and ocean carbonate). As mentioned in page 11 line 6, due to the temporal resolution of our data we cannot resolve short-term processes (i.e. that take place over less then ~10 kyr), but in our revised manuscript we have**
35**made this section clearer and be more explicit about which processes of the geological carbon cycle and long-**

term climate system we expect to play a role (page 10 line 29) and include our interpretation that warming-driven changes in organic carbon burial are a likely pre-PETM mechanism in the Conclusions (page 17 line 14).

It is of particular importance because of the timescale of considered perturbations of d13C and d18O: It is specified Page 4, lines 16-19 that the data are detrended in order to remove any long-term trends. This is indeed essential to get stationary time-series. But though it is explicitly said that bandwidth "is an important consideration", the only given information is that it is "adjusted heuristically for each dataset". More precision should be given on bandwidth value (is it constant along one given time-serie?) and above all because if it is less than the response time of carbon cycle (~100–200ky, François and Goddéris, 1998), then the indicators are not (or only partially) measuring the resilience due to its feedbacks. From the timescale of variations of smoothed records shown in Fig. 3, 4 S1 and S2 (red lines), I guess the bandwidth is actually in the order of 100ky. The choice of the timescale of the fluctuations to study and the feedbacks to investigate is up to the authors, but it should be specified, and conclusions can be drawn only for the focused feedbacks.

You are also right that bandwidth choice is important in the context of process timescales as well, and that the original manuscript failed to mention the bandwidth choices for the main rolling window metrics – we found that the optimal Gaussian kernel bandwidth was 0.1 and have made this clear in the revised manuscript (page 12 line 8). This does not directly translate to a frequency limit, but as shown by the red line in Figures 2 and 3 this removes all of the secular trends and the long-term orbital-scale (~100+ kyr) variability. While this at first inspection would seem like we're filtering out what we're interested in (the long-term carbon cycle / climate system), with this methodology it is in fact the short-term noise that we are interested in as this short-term noise reveals the resilience of the longer-term processes. But as mentioned above, the data's temporal resolution places a lower limit on what timescale processes this method can reveal (in this case anything shorter than 10s of kyr). As a result, although we filter out the direct signal of longer-term processes, with this method this does not exclude these processes from affecting the results and interpretation. We have now added discussion of this to the background (page 10 line 29 onwards) and methodology (page 12 line 8).

The section 3.2 (Binned Metrics) and Table 1 show that most of the indicators exhibit significant variations before/after the PETM and the ETM2. This suggest (as mentioned Page 7, lines 26–28) that the hyperthermal events are partly responsible for the loss of resilience of the carbon-climate system, or at least that these events are not simple perturbations followed by a relaxation towards the same "initial state": they induce some permanent or irreversible changes. Even if evidences of tipping points are lacking (as said Page 7, lines 27–28), I think it is an interesting results and should be highlighted in the conclusion, where there is no mention of this fact. I also wonder which component of carbon-climate system can be expected to undergo irreversible changes. If there is any "good candidate", it may be interesting to precise it.

Our results do indeed clearly show increased and persistent long-term destabilisation following both hyperthermal events in our record, and this has implications for the early Eocene (as a destabilised carbon-climate system may have played a role in the subsequent repeated hyperthermals). In our original manuscript we mostly focused on the implications of destabilisation for the presence of tipping points preceding the two hyperthermal events in our record, but you are right to point out that the subsequent

5

persistence of the loss of resilience is important as well and we have made this clearer in the discussion (page 15 line 13 & line 30) and conclusions (page 16 line 28 & page 17 line 18) (as well as the processes likely to be implicated in this persistent loss of resilience, such as a dampened silicate weathering feedback in the Eocene relative to the Palaeocene).

Minor specific remarks:

Page 4, line 15 and line 22: With the definition of autoregressive model as 'x(t+dt) = a*x(t) + e(t)', a constant timestep 'dt' is inherently necessary for the autoregressive coefficient 'a' is directly link to 'dt'. I wonder then how to fit an autoregressive model to the non-interpolated data? Just dividing 'a' by the local timestep is enough? It is not precised in the description of the 'generic_ews' function in R. Perhaps it is a "routine analysis" and is not worth to be precised, I can't really judge it.

For the AR model time-step the R function simply assumes a constant time-step throughout the time-series (i.e. it takes the time-series length and divides by N to get the constant dt). Of course without interpolation the actual data points will slightly differ from this dt which will introduce some error (and is why the interpolated run is considered the default, despite the alternative problem discussed in the text that this introduces instead), but there is no systematic bias in the distribution real data time-steps which limits any systematic error as a result. The assumption of a constant time-step has been clarified (page 11 line 30 & page 12 line 14) and a plot of data time-steps provided in the Supplementary Material.

Page 6, line 21: The word "divergence" (of standard deviation) may be misleading because the reader would firstly expect to find an "increase", which is in contradiction with the rest of the sentence (and the Figure). Perhaps it is preferable to substitute it for "decrease of standard deviation" or "reducing standard deviation". In addition, this decrease of SD is likely to be due to earlier "extreme" events (between 59.5 and 58.5Ma) that slip out of the rolling window when it reaches 57-56Ma. Indeed, with a 75% rolling window (Fig. S1), SD does not exhibit such a decrease. It does with a 25% rolling window probably because there are "extreme" events up to 57.5Ma. Therefore, this decrease of SD could be not linked to what happen immediately before the PETM, and not contradicting with the increase of AR1.

We have clarified that d18O SD declines just before the PETM and discuss the likely reason why (page 14 line 19).

Page 6, line 29: How to interpret a decrease in kurtosis? Shouldn't we except while approaching a tipping point more frequent extreme deviations, and then a higher kurtosis?

We agree that one would expect kurtosis to increase prior to a tipping point as extreme data values become more common, but this is not considered to be as strong an indicator of resilience as other measures like variance or AR1. In this case it seems that the section in the d13C record between ~58-59.5 Ma has an especially high kurtosis (as shown in supplementary fig 1 especially by the 25% window) which then gradually falls as this period leaves the rolling window, which is likely masking any pre-PETM signal. Discussion of this feature has been added to the caption of Supp. Fig. 1.

Page7, lines10–11: "but that the increase of AR(1) for both d18O and d13Carehighly significant (p=0)". Please add "for non-interpolated data" to be explicit.

Clarified (page 15 line 11)

Page 7, lines 13-15: There is another evidence that the values of d13C during the events (at least for the PETM) are partly responsible for the upward steps in the indicators: with a 25% rolling window (Fig. S2), both AR1 and SD show a downward step at 54Ma, exactly when the PETM leaves the rolling window (and after the ETM2 has come into the rolling window). However, it is true that the binned metrics clearly show steps that are not due to what happen during the events.

The within-event data in the whole-dataset rolling window analysis does affect the results, but the binned data metrics that exclude the events confirm the shifts are present anyway – further detail on this contrast has been added (page 15 line 27).

Page 8, lines 2-3: Same remark as for the decrease of SD before PETM: the decrease of jump intensity can be due to the "extreme" events between 59.5Ma and 57.5Ma.

Unlike with the previous comment, the within-event data doesn't affect the DDJ model results as this method does not use a rolling window and all data are used and relevant.

Supplementary Figures 1 and 2: How come than skewness is systematically positive? At least during PETM and ETM2, it seems there are more points beneath the red line than above. Moreover, in the binned metrics analysis (Table 1), skewness in systematically negative.

There is indeed a mismatch between the d13C skewness as calculated on a rolling window (>~0 at all times and gradually increasing) and in the bins (<0 and decreasing for every bin). This is because the generic_ews function of the earlywarnings R package actually calculates and plots the absolute of the skewness – this is because skewness can increase or decrease when approaching a critical transition, and so the package plots the absolute in order to maintain consistency with the other indicators that should always increase. This means that the plotted skewness is in fact axis inverted, and matches the negative and declining values in Table 1. The events themselves skew positive (shown as a decrease in the plot) due to a detrending artefact that results in a temporary increase. We have updated the text, supplementary figures, and captions to make it clear the skewness plot is of the absolute values.

Additional Changes

In addition to the changes described above, we have also updated the p-values on the AR(1) rolling window and binned DFA-h in order to use arima model-generated surrogate records rather than bootstrapped datasets as a more appropriate null model choice (this led to only very minor changes in p-values though, with no impact on our results or discussion).

Reduced Carbon Cycle Resilience across the Palaeocene-Eocene Thermal Maximum

David I. Armstrong M[c]Kay*[1,2 a] & Timothy M. Lenton[32]

[1] Ocean and Earth Science, University of Southampton, National Oceanography Centre Southampton, Southampton, SO14 3ZY, UK (work undertaken here)
[2] Stockholm Resilience Centre, Stockholm University, Kräftriket 2B, SE-10691 Stockholm, Sweden (current address)
[23] Earth System Science group, College of Life and Environmental Sciences, University of Exeter, Exeter, EX4 4QE, UK
[a] Stockholm Resilience Centre, Stockholm University, Kräftriket 2B, SE-10691 Stockholm, Sweden (current address)

Correspondence to: David I. Armstrong McKay (david.armstrongmckay@su.se)

Abstract. Several past episodes of rapid carbon cycle and climate change are hypothesised to be the result of the Earth system reaching a tipping point beyond which an abrupt transition to a new state occurs. At the Palaeocene-Eocene Thermal Maximum (PETM) ~56 Ma, and at subsequent hyperthermal events, hypothesised tipping points involve the abrupt transfer of carbon from surface reservoirs to the atmosphere. Theory suggests that tipping points in complex dynamical systems should be preceded by critical slowing down of their dynamics, including increasing temporal autocorrelation and variability. However, reliably detecting these indicators in palaeorecords is challenging, with issues of data quality, false positives, and parameter selection potentially affecting reliability. Here we show that in a sufficiently long, high-resolution palaeorecord there is consistent evidence of destabilisation of the carbon cycle in the ~1.5 Myr prior to the PETM, elevated carbon cycle and climate instability following both the PETM and Eocene Thermal Maximum 2 (ETM2), and differing different drivers of carbon cycle dynamics preceding the PETM and ETM2 events. Our results indicate a loss of 'resilience' (weakened stabilising negative feedbacks and greater sensitivity to small shocks) in the carbon cycle before the PETM, and in the carbon-climate system following it. This pre-PETM carbon cycle destabilisation may reflect gradual forcing by the contemporaneous North Atlantic Volcanic Province eruptions, with volcanism-driven warming potentially weakening the organic carbon burial feedback. Our results are consistent with but cannot prove the existence of a tipping point for abrupt carbon release, e.g. from methane hydrate or terrestrial organic carbon reservoirs, whereas we find no support for a tipping point in deep ocean temperature.

1.   Background

The Palaeocene-Eocene Thermal Maximum (PETM) at ~56 Ma is considered a potential example of passing a tipping point in the carbon-climate system where a smooth change in forcing triggered a large response (Lenton, 2013). Palaeorecords across the PETM indicate that an abrupt release of isotopically-light carbon (between 2000 and 13000 Pg C, best estimate ~3000 Pg C) into the ocean-atmosphere system occurred in under ~5 kyr, accompanied by global warming of ~5 °C, a 2.5 to 3.0 ‰ benthic $\delta^{13}C$ excursion, and significant ocean acidification (Dickens, 2011; Dickens et al., 1995; Dunkley Jones et al., 2013;

8

Frieling et al., 2017; Kirtland Turner et al., 2017; Littler et al., 2014; McInerney and Wing, 2011; Sluijs et al., 2007b; Zachos et al., 2005, 2008, Zeebe et al., 2009, 2016). It has been hypothesised that gradual warming during the late Palaeocene (Figure 1Figure 1) eventually crossed a tipping point, either through an internal process or an external perturbation such as volcanism (Svensen et al., 2004), which triggered the extensive dissociation of a carbon cycle 'capacitor' such as methane hydrates in ocean sediments (Dickens, 2011; Dickens et al., 1995; Minshull et al., 2016), permafrost soil carbon (DeConto et al., 2012) or organic carbon from a source such as peat (Cui et al., 2011; Kurtz et al., 2003) that benthic $\delta^{13}C$ records and modelling indicate accumulated earlier in the Palaeocene (Dickens, 2011; Komar et al., 2013; Kurtz et al., 2003). This in turn led to a rapid increase in atmospheric CO_2 (pCO_2) and the subsequent amplification of global warming and carbon release in a positive feedback loop that shifted the Earth System to a warmer state for ~100 kyr. An alternative hypothesis is that a very large external perturbation of volcanic carbon caused the PETM (Gutjahr et al., 2017) with little a smaller role for amplifying feedbacks within the carbon cycle and therefore no significant role for a tipping point.

The PETM was followed by the Early Eocene Climatic Optimum (EECO; Figure 1Figure 1) containing subsequent hyperthermal events such as Eocene Thermal Maximum 2 (ETM2) at ~54 Ma and ETM3 at ~53 Ma, which are potentially paced by orbital eccentricity forcing interacting with long-term warming and discharging methane hydrate deposits to produce threshold responses past repeated tipping points (Archer et al., 2009; Kirtland Turner et al., 2014; Komar et al., 2013; Littler et al., 2014; Lourens et al., 2005; Lunt et al., 2011; Stap et al., 2010; Westerhold et al., 2007; Westerhold and Rohl, 2009). However, the PETM occurred in a different orbital setting to the later events, suggesting that the PETM required an additional external "push" while the latter hyperthermals were eccentricity-paced tipping points (Littler et al., 2014). This push could have come from the emissions of the contemporaneous North Atlantic Volcanic Province (NAVP) eruptions both before and during the PETM (Frieling et al., 2016; Gutjahr et al., 2017; Storey et al., 2007; Svensen et al., 2004). Methane release from hydrate dissociation may also have been significantly limited or delayed by sediment transport processes, potentially limiting its role as a positive feedback (Minshull et al., 2016).

Many complex systems have been found to include tipping points, thresholds beyond which even small changes in condition theycan trigger the system to abruptly transition into a new equilibrium state (Dakos et al., 2015; Held and Kleinen, 2004; Lenton, 2013; Lenton et al., 2008; van Nes et al., 2016; Scheffer et al., 2001, 2009). Theory suggests that, prior to reaching such a tipping point, a system will exhibit 'critical slowing down' of its dynamics – meaning a slowing recovery rate in response to perturbations – which can be detected as increasing trends in autocorrelation and variability in time-series data (Carpenter and Brock, 2006; Dakos et al., 2008, 2012a; Kéfi et al., 2013; Lenton, 2011; Lenton et al., 2012a; Scheffer et al., 2009). Changes in skewness and kurtosis may also occur (although skewness can both increase or decrease depending on the nature of the alternative stable state and potential wells), and if internal variability is high, a system can 'flicker' between different states before undergoing a more permanent shift between them (Dakos et al., 2012a, 2013; Scheffer et al., 2009; Wang et al., 2012). Previous work has suggested that these indicators, which can be used as 'early warning signals' (EWS) or

more generally as metrics of resilience ~~metrics~~(the capacity of a system to recover from disturbance and return to its reference state (Grimm and Wissel, 1997; Holling, 1973; Scheffer et al., 2001)), may be detectable prior to some abrupt climate transitions in the palaeorecord (Dakos et al., 2008; Lenton, 2011), including the Eocene-Oligocene Transition and during several Pleistocene climate shifts (Dakos et al., 2008; Lenton, 2011; Lenton et al., 2012b, 2012a). However, autocorrelation

5   and variance can also increase prior to non-catastrophic transitions, with or without bifurcations in phase space (Kéfi et al., 2013), or under some specific circumstances even decrease (Boettiger and Hastings, 2012b; Dakos et al., 2012b; Livina et al., 2012). Hence here increasing autocorrelation and variance is viewed more broadly as indicating declining resilience of a system (i.e. weakening negative feedbacks and greater sensitivity to small shocks), whether or not a critical transition is imminent. Other potential issues with detecting changing system resilience in palaeorecords include infrequent sampling rate, dating

10   uncertainties, the possibility of producing false positives or negatives, and the extent to which these methods are dependent on subjective parameter choices (Boettiger et al., 2013; Boettiger and Hastings, 2012a; Lenton, 2011) (see Methods for further discussion).

Here we test the hypothesis that the PETM and ETM2 are examples of tipping points being reached in the carbon-climate

15   system following long-term destabilisation (e.g. of a sensitive carbon cycle capacitor rich in isotopically-light carbon), by looking for declining resilience preceding them using published methodologies (Dakos et al., 2008, 2012a, Lenton et al., 2012a, 2012b). Whilst a signal of declining resilience cannot prove the existence of a tipping point, its absence would tend to falsify the tipping point hypothesis. Palaeorecords suffer from greater dating uncertainties and a less frequent sampling rate than is possible with modern climate data, making robust time-series analysis more challenging. Hence sufficiently long and high-

20   resolution palaeorecords available across the late Palaeocene and early Eocene were required in order to enable significant results to be obtained. To this end we use the ~7.7 Myr-long benthic $\delta^{18}O$ and $\delta^{13}C$ palaeorecords from ODP Site 1262 in the South Atlantic (Littler et al., 2014), and sub-divide the datasets into pre-PETM and post-PETM bins, as well as sub-dividing the post-PETM bin into pre-ETM2 and post-ETM2 bins, for separate analyses. These isotope records track the long-term global state of high latitude climate and the carbon cycle respectively (Zachos et al., 2001, 2008) and are therefore appropriate data

25   choices for detecting the resilience of the global carbon-climate system, which in turn determines the long-term resilience of the whole Earth System as its key slow-timescale components.

A major limitation of the available palaeorecords is that their resolution is of the order of ~3 kyr, which only allows us to monitor changes in the dynamics of the slowest parts of the carbon cycle and climate system (i.e. >10 kyr). For the carbon

30   cycle these could include the silicate weathering feedback, which is hypothesised to act as the main long-term negative feedback on atmospheric CO_2 and therefore climate change (Berner, 1991; Berner et al., 1983; Kump and Arthur, 1997; Urey, 1952; Walker et al., 1981; Walker and Kasting, 1992), the strength of the biological pump and carbon burial rates in the ocean (Boscolo-Galazzo et al., 2018; Chamberlin, 1898; Derry and France-Lanord, 1996; France-Lanord and Derry, 1997; Hay, 1985; John et al., 2014), and medium-term fluctuations in the storage capacity of carbon reservoirs such as the deep ocean,

methane hydrates, permafrost, or soil carbon (Batjes, 1996; Buffett and Archer, 2004; Cui et al., 2011; DeConto et al., 2012; Dickens et al., 1995; Klinger et al., 1996; Tarnocai et al., 2009). For the climate system, slow processes could include substantial reorganisations of ocean circulation (Hofmann and Rahmstorf, 2009; Rahmstorf, 2002; Stocker and Wright, 1991; Stommel, 1961) and the growth or collapse of large ice sheets (although no substantial ice sheets existed at this time) (DeConto et al., 2008; DeConto and Pollard, 2003; Pagani et al., 2011; Pollard and DeConto, 2009)~~potentially including large carbon reservoirs in the ocean, methane hydrates, permafrost, or soil carbon, and the ocean thermohaline circulation~~). Any shorter-term drivers of instability closer to the event, for example changes in ocean and atmospheric dynamics or precursor warming on millennial timescales (Secord et al., 2010; Sluijs et al., 2007a), will be missed and thus could constitute 'missed alarms'. As a result, in this study we focus only on the long-term ~~trends~~ processes in the global carbon-climate system prior to and across the PETM and ETM2.

We use multiple indicators – including autoregressive coefficient at lag 1 (AR(1)) and detrended fluctuation analysis h-value (DFA-h~~; binned metrics only~~) (Lenton et al., 2012b; Livina and Lenton, 2007) to reveal slowing down, and standard deviation (SD) and non-parametric drift-diffusion-jump (DDJ) model function metrics (Dakos et al., 2012a) to reveal increasing variability. An overall increasing trend in AR(1) or DFA-h would show the slow parts of the climate or carbon systems were recovering more slowly from regular perturbations, while increasing SD or variance as measured by the DDJ model would show each system was being perturbed further from their current state. Together they indicate a system being destabilised and becoming less resilient to being knocked into a new state. Skewness and kurtosis are also measured to provide further context (see Supplementary Material) as both may change in the presence of more extreme values. Sensitivity analyses are conducted in order to ensure detected signals are robust across different methodologies and parameter choices (see Methods and Supplementary Material).

2.   Methods

2.1.   Rolling Window Metrics

For the rolling window metrics we follow the methodology first outlined in a climate context by (Held and Kleinen, 2004), and subsequently used by other studies including (Dakos et al., 2008, 2012a, Lenton et al., 2012a, 2012b; Livina and Lenton, 2007), and the 'Early Warning Signals Toolbox' developed based on this work (documented at www.early-warning-signals.org and available in as the '*earlywarnings*' package in R (R Foundation for Statistical Computing, 2016)). After selecting the dataset and for the pre-PETM analysis terminating it just prior to the hypothesised transition to avoid biasing the analysis, the data are first interpolated (using linear interpolation by default with the '*interp1*' function in Matlab (The MathWorks Inc., 2016)) to provide the equidistant data-points required for rigorous statistical analysis and assumed by the AR(1) model. However, interpolation itself can introduce statistical artefacts into the analysis as, by definition, the addition of interpolated data-points increases self-similarity and thus autocorrelation in the dataset. In palaeorecords this tends to result in an artificial

increase in autocorrelation in parts of the dataset with either sparser data-points or complete gaps in the data, but in this record there is no marked systematic shift in data time-steps (Supplementary Figure S5). As a result we also analyse non-interpolated data in order to assess the sensitivity of our results to interpolation. Following this, the data are then detrended by subtracting the smoothed dataset, estimated with a Gaussian kernel smoothing function (using the '*ksmooth*' function in R), in order to

5   remove any long-term trends because as these are not the focus of the analysis. This makes the dataset stationary – a necessary prerequisite for time-series analysis – but this also somewhat reduces the value of lag 1 autocorrelation in the results. Bandwidth is an important consideration in this process and is adjusted heuristically for each the datasets in order to best remove long-term trends but leaving short-term fluctuations, in this case giving a Gaussian kernel bandwidth of 0.1. This removes long-term secular trends and orbital cyclicity (>100 kyr) (Figure 2 & Figure 3, top panels), leaving only the short-

10   term noise that reveals the resilience of the underlying longer-term (>10 kyr) processes.

An autoregressive model of order 1 (AR(1)) is fitted to the data within a rolling window (using the '*generic_ews*' function of the '*earlywarnings*' package in R). The AR(1) model is of the form: $x_{t+1} = \alpha_1 x_t + \varepsilon_t$, fitted by an ordinary least-squares method and with a Gaussian random error and a constant time-step. Following previous studies the default window size is set

15   at half the length of the dataset, but as part of our sensitivity testing we also repeat our analyses for window sizes of between 25% and 75% (Supplementary Figures S1-S4). The choice of window length is a trade-off between dataset resolution and the reliability of the estimate of the indicator, with a short window allowing shorter-term changes in indicators to be tracked at the cost of lower estimate reliability and vice versa. On the same rolling window the absolute skewness, kurtosis, and standard deviation of the dataset are also calculated (also using the '*generic_ews*' function of the '*earlywarnings*' toolbox in R).

20   Detrended fluctuation analysis h-value (DFA-h) was also used as an alternative measure to AR(1) for short-term memory and critical slowing down in the dataset bins (performed using the '*DFA*' function of the '*fractal*' package in R). DFA extracts the fluctuation function over a window s, and if the data is long-term power-law correlated, the fluctuation function $F(s)$ increases as a power law: $F(s) \propto s^h$, where h is the DFA fluctuation exponent (Peng et al., 1994) and reaches value 1.5 at a critical transition (Lenton et al., 2012b; Livina and Lenton, 2007).

25

Finally, the likelihood of there being a real trend in the results is calculated by estimating the nonparametric Kendall rank-correlation statistic (τ), which measures the strength of an indicator's tendency to increase (>>0) or decrease (<<0) against the null hypothesis of randomness (~0) (also using the '*generic_ews*' function of the '*earlywarnings*' toolbox in R). However, this statistic is most robust when the trend is consistent over a long period, while increasing but oscillating trends or trends only at

30   the very end of the record can produce weak or even negative values despite a clearly visible trend (Dakos et al., 2012a). We also calculate a p-value for each metric by bootstrapping our detrended datasets to generate 1000 surrogate records (or arima model-generated for AR(1) and DFA-h) with equivalent mean and variance, re-calculating the metric and Kendall τ value for each, and finding the proportion of Kendall τ values equal or greater than that of the original palaeorecord (Dakos et al., 2008).

2.2. Binned metrics

As well as performing rolling window time-series analysis, we also measure AR(1), SD, skewness, and kurtosis on ~~binned~~ data (detrended but not interpolated) binned into pre- and post-event bins and excluding the events themselves to provide simple before/after comparisons of changes across the events. To this end the datasets (n=2302) were binned into pre-PETM (n=1331), post-PETM (n=921), PETM to ETM2 (n=593), and post-ETM2 (n=240) bins, excluding data-points from during each event so as to avoid biasing by extreme or missing data. ~~Detrended fluctuation analysis (DFA) was also used as an alternative measure to AR(1) for short-term memory and critical slowing down in the dataset bins (performed using the '*DFA*' function of the '*fractal*' package in R). DFA extracts the fluctuation function over a window s, and if the data is long-term power law correlated, the fluctuation function $F(s)$ increases as a power law: $F(s) \propto s^h$, where h is the DFA fluctuation exponent and reaches value 1.5 at a critical transition .~~ A p-value is calculated for each metric using a permutation test (i.e. by reshuffling and repartitioning the before/after event data into the same sized bins 1000 times and comparing the metrics' resultant before/after differences with the observed metric before/after difference), except for AR(1) and DFA-h for which we instead use AR(1) model-derived surrogate data to compare against (i.e. by generating 1000 surrogate datasets with the same AR(1) value, mean, and variance as the before-event bin over the length of the after-event bin and compare this distribution to the observed after-event AR(1) value).

2.3. Nonparametric Drift-Diffusion-Jump Model

A model-based alternative to the time-series analysis methods (whether rolling window or ~~metric~~bin-based) above is to fit a general nonparametric drift-diffusion-jump model to the dataset with as a surrogate for an unknown data-generating process (Carpenter and Brock, 2011; Cox and Ross, 1976; Dakos et al., 2012a; Johannes, 2004). In this model functions are estimated for drift, diffusion, and jump processes using nonparametric regression, where drift measures the local rate of change, diffusion measures the standard deviation of the relatively small shocks that occur at each time step, and jumps are large intermittent shocks. The conditional variance of the data is also estimated from the nonparametric regression, and represents the variance of the data from its conditional expectation estimated using kernel regression. We use the '*ddjnonparam_ews*' function in the '*earlywarnings*' package on R, using the default options of a bandwidth of 0.6 and 500 points for computing the kernel. We use raw data for this analysis, with no detrending or interpolation and without log transforming the data first. In interpreting the results we focus on the general long-term trends in the estimated terms as many of the shorter-term fluctuations potentially represent model over-fitting.

2.4. Limitations

Despite positive results in some palaeoclimate EWS studies (Lenton, 2011), there are several potential issues with searching for resilience indicators in palaeorecords. Palaeorecords suffer from greater dating uncertainties and a less frequent sampling rate than is possible with modern climate data, making robust time-series analysis more challenging. Most indicators also do

not reveal exact information about the nature of the transition itself, with increasing slowing down and variability detected prior to both catastrophic and non-catastrophic transitions featuring a bifurcation in phase space and even before non-catastrophic transitions without a bifurcation (Kéfi et al., 2013). Concerns have also been raised over the likelihood of producing false positives (where EWS appear to indicate an impending transition which never occurs) or false negatives (i.e. a "missed alarm", when EWS may be entirely absent prior to a known critical transition), and the extent to which these methods are dependent on subjective parameter choices (Boettiger et al., 2013; Lenton, 2011). There is a risk that selecting and analysing known or suspected critical transitions in the palaeoclimate record is particularly liable to false positives, as positive indicators at the transitions could potentially have occurred purely by chance rather than due to systemic instability (Boettiger and Hastings, 2012a). However, it has been argued that EWS can be reliably detected if both increasing autocorrelation and variance are seen prior to the transition rather than one of these indicators alone (Ditlevsen and Johnsen, 2010). Detecting multiple, consistent, and robust signals from the indicators can be indicative of decreasing system resilience even if a catastrophic transition is not reached or is instead triggered by an external perturbation rather than internal processes (Dakos et al., 2015).

3.   Results and Discussion

3.1.   Rolling Window Metrics

Rolling window metrics prior to the PETM reveal a sudden increase in AR(1) and SD after ~5~~8.2~~7.7 Ma in the interpolated benthic $\delta^{18}O$ record associated with a~~n~~ step ~~excursion~~ in the data (and the benthic $\delta^{13}C$ record peaking), which despite a temporary drop in DFA-h suggests some degree of destabilisation of the slow climate system prior to the PETM (Figure 2~~Figure 2~~ & Supporting Figure S1). However, the subsequent ~~divergence~~ decline of standard deviation after ~5~~7~~6.5 Ma (likely to partially be the result of earlier extreme data-points from ~58.5-59.5 Ma leaving the rolling window) does not support a tipping point involving deep ocean temperature at the PETM, which is also indicated by the non-significant bootstrapped p-values for the metrics of both interpolated and non-interpolated data. The non-interpolated DFA-h metric does show a significant increase (p=0.015), which would suggest systemic slowing down, but this does not match the non-interpolated AR(1) metric. Alternatively, th~~is~~ese results~~is~~ could represent a 'missed alarm' as the shorter-term climate dynamics that might be critical to the dynamics of the tipping point are not sufficiently resolved by the available data.

The benthic $\delta^{13}C$ record shows clearer evidence of declining resilience in the slow components of the carbon cycle, with long-term increases in AR(1), DFA-h, and SD in the run-up to the PETM with steps at ~5~~8.2~~7.7 Ma and ~5~~7.3~~6.8 Ma which are consistent across the sensitivity analyses (absolute skewness also increases, while kurtosis declines up to the PETM; Supporting Figure S1). Bootstrapped p-values indicate that the $\delta^{13}C$ SD trend is significant (p=0.002) for the interpolated data while the $\delta^{13}C$ AR(1) and DFA-h trend~~s~~ ~~are~~ is significant (p=0.02~~4~~2 & p=0.021 respectively) for the non-interpolated data.

14

This supports a long-term slowing down in benthic δ^{13}C in the late Palaeocene, which may reflect a gradually-forced destabilisation of the global carbon cycle prior to the PETM.

Rolling window analysis across the whole of the late Palaeocene / early Eocene (LPEE) interval suggest but cannot prove
5 systemic changes in carbon cycle and climate (in)stability across both the PETM and ETM2 (Figure 3~~Figure 3~~ & Supporting Figure S2). Between the PETM and ETM2 δ^{18}O AR(1) and DFA-h increase~~s~~ up until ~200 kyr before ETM2 and SD experiences a small temporary increase followed by a larger decrease. In contrast, all metrics for δ^{13}C experience a rapid jump during the PETM and then remain relatively stable until ETM2. Following ETM2, δ^{18}O AR(1) and DFA-h increase~~s~~ significantly while SD increases slightly, whereas for δ^{13}C ~~both~~ all metrics (as well as absolute skewness and kurtosis;
10 Supporting Figure S2) consistently increase. However, the bootstrapped p-values indicate that none of these trends are significant for the interpolated data, but that for the non-interpolated data the increase in AR(1) and DFA-h for both δ^{18}O and δ^{13}C are highly significant (p=0 to 0.001). This indicates that there is some evidence for slowing down – but not for increased variability – in both the δ^{18}O and δ^{13}C data and therefore the long-term climate system and carbon cycle across the LPEE interval into the Eocene, but this is dependent on not interpolating the data prior to the analysis. It should also be recognised
15 that the abrupt shifts in δ^{13}C at the PETM and ETM2 are not fully removed by detrending prior to the analysis, hence they are at least partly responsible for the upward steps in the indicators at the events.

3.2.   Binned Metrics

To address the issue of large excursions failing to be removed by detrending for the rolling window metrics, as well as the issue of data gaps caused by dissolution at the peak of each event (Littler et al., 2014), we calculate aggregate metrics (i.e. no
20 rolling windows) on the binned data (excluding data from within the events) (Table 1~~Table 1~~). The binned metrics show significant increases in AR(1) across both the PETM and ETM2 for δ^{18}O and δ^{13}C, with DFA-h also significantly increasing across the PETM and ETM2 for δ^{18}O and across ETM2 for δ^{13}C. SD significantly falls in the interval between the PETM and ETM2 and then significantly increases across ETM2 for both datasets, but the significant increase for δ^{13}C for the all data after the PETM ~~the overall increase is only significant for δ^{13}C~~is likely to be biased by the inclusion of extreme data-points within
25 ETM2. These results are consistent with the rolling window metrics, except for δ^{13}C SD which in the rolling window metrics is higher between PETM and ETM2 than before the PETM, and is likely to be the result of excluding the extreme data values during the PETM itself in the binned analysis (with SD beginning to drop just before the ETM2 in the rolling window analysis once the PETM leaves the window). These results indicate both the long-term climate system and carbon cycle slowed down to some extent after both the PETM and ETM2 (but became less variable following the PETM itself until ETM2), providing
30 support for the slow parts of global carbon-climate system being progressively destabilised through the LPEE interval and into the Eocene by both of the hyperthermal events but not for any tipping points.

3.3. Nonparametric Drift-Diffusion Jump Model

Fitting a non-parametric drift-diffusion-jump model to the datasets provides independent model-based metrics to compare to the rolling window metrics, with terms for the conditional variance measuring variance from dataset's conditional mean (estimated by kernel regression), diffusion measuring the standard deviation of regular small shocks at every time-step, and jump intensity measuring either irregular large shocks or flickering (Carpenter and Brock, 2011; Dakos et al., 2012a) (Figure 4Figure 4). For benthic $\delta^{18}O$ this model reveals an overall increase in conditional variance and diffusion and a reduction in jump intensity ~2 Myr before the PETM, followed by a slight decrease in conditional variance after the PETM and intermittent spikes in jump intensity and conditional variance during and following ETM2. This suggests the climate shifted to a state with higher variability featuring regular small shocks ~2 Myr prior to the PETM, became slightly less variable following the PETM, but featured larger irregular shocks during and after ETM2. While this suggests some degree of climate instability in the ~2 Myr before the PETM and following ETM2, there is no evidence of a critical transition in the climate system at the PETM or ETM2 themselves. In contrast, the benthic $\delta^{13}C$ model reveals decreasing diffusion and increasing conditional variance and jump intensity in the 1.5 Myr run-up to the PETM, indicating increasing total variability driven by large irregular shocks and consistent with a critical transition being approached in the carbon cycle at the PETM (Dakos et al., 2012a). Conditional variance and jump intensity remain high and diffusion remains low for ~1 Myr after the PETM, before reversing ~300 kyr before ETM2 except for brief spikes in both diffusion and jump intensity during and after ETM2. This indicates that variability in the carbon cycle remained high and driven by large shocks for ~1 Myr after the PETM, but that variability mostly shifted towards smaller regular shocks prior to and after ETM2. This difference in the sources of variance prior to each event (the PETM is preceded by elevated jump intensity and overall conditional variance, whereas ETM2 is preceded by increased diffusion and decreased jump intensity) suggests potentially differing carbon cycle dynamics and drivers prior to each event. The shift in variability before ETM2 also slightly precedes the biotic turnover detected in both marine and terrestrial records in the ~200 kyr prior to ETM2, despite there being no obvious shift in the palaeorecords that may have driven this turnover (Westerhold et al., 2018). Overall these $\delta^{13}C$ DDJ results are consistent with elevated carbon cycle instability following the PETM, but suggests that ETM2 was not preceded by the same dynamics as the PETM.

4. Conclusion

In summary, both rolling-window metrics before and across the PETM, binned metrics, and nonparametric drift-diffusion-jump models suggest indicate that there was a decline in variability – suggesting a loss of resilience – in the slow components of the carbon cycle before and following the PETM and ETM2. Following both the PETM and ETM2 there is also evidence of slowing down in both the long-term carbon cycle and climate system, indicating that both events led to a longstanding destabilisation of the carbon-climate system. In contrast, while there is some evidence for destabilisation in the $\delta^{18}O$ data prior to and after the PETM, there is no clear evidence of a critical transition in the climate system at this time. Minimal lag between $\delta^{18}O$ and $\delta^{13}C$ in the late Palaeocene indicates close coupling between climate and the carbon cycle prior to the PETM (Littler

et al., 2014), and so the observed instability in the climate system is likely to have been induced by the contemporaneous destabilisation of the carbon cycle. Furthermore, ETM2 appears to be preceded by different carbon cycle dynamics to the PETM, which fits with the suggestion that the PETM required an extra "push" unlike the later eccentricity-paced hyperthermals which might represent more classical tipping points (Littler et al., 2014). These results are consistent with the hypothesis of a gradual destabilisation of the long-term carbon cycle ~~by increasing pCO$_2$ concentrations~~ in the ~1.5 My**r** preceding the PETM (starting at ~57.5 Ma, intensifying after ~56.5 Ma) associated with increasing pCO$_2$ concentrations and the warming seen in the benthic δ^{18}O record from ~58.2 Ma.~~, which~~ This coincides with the intensification of North Atlantic Volcanic Province (NAVP) eruptions between 61 and 57 Ma and its associated volcanic and thermogenic CO$_2$ and methane emissions, with subsequent large-scale ~~emissions~~ eruptions ~~during the PETM~~from ~56.1 Ma (likely preceded by some degree of cryptic degassing (e.g. Armstrong McKay et al., 2014)) potentially triggering or ~~helping to~~ prolonging ~~its~~ the PETM's duration (Frieling et al., 2016; Storey et al., 2007; Svensen et al., 2004). This time also coincides with a dramatic long-term decrease in organic carbon storage following a large build-up as indicated by the ~~sharp~~downturn in benthic δ^{13}C from ~58**7** Ma (Figure 1~~Figure 1~~), either as a result of large-scale methane hydrate or peat dissociation and oxidation in response to warming (Dickens, 2011; Komar et al., 2013; Kurtz et al., 2003) or a reduction in marine biological pump strength as higher temperatures lead to increased respiration rates of particulate organic carbon (Boscolo-Galazzo et al., 2018; John et al., 2014). This weakening of the organic carbon burial feedback in response to volcanism-driven warming could be the primary driver of the observed geological carbon cycle destabilisation in the ~1.5 Myr prior to the PETM, and may have in turn prolonged the duration of the PETM itself. Reconstructed silicate weathering feedback strength fell by ~40% in the ~3 Myr after the PETM (due to reduced continental weatherability) (Caves et al., 2016; van der Ploeg et al., 2018), potentially allowing the carbon-climate system to remain destabilised and susceptible to further shocks long after the PETM and ETM2 as observed. The hypothesis of a carbon cycle tipping point at the PETM survives our tests (although they cannot directly confirm it or rule out an external trigger). In contrast the hypothesis of a tipping point in deep ocean temperature (as recorded by the δ^{18}O record considered) is not supported. A large external perturbation, e.g. a massive, abrupt injection of volcanic carbon from the NAVP during the PETM (Gutjahr et al., 2017; Storey et al., 2007) or the ~~recent~~ suggestion of a meteorite strike (Schaller et al., 2016), could have played a role in triggering the PETM, but we find clear evidence that the carbon cycle had already been getting progressively more unstable and thus more vulnerable to being pushed beyond a tipping point, and remained so in its aftermath.

Author Contributions

DIAM & TML designed the study; DIAM performed the analyses; DIAM & TML wrote the paper.

Data availability

The data used for the plots and analysis of the global Palaeocene and Eocene benthic isotope stack (Zachos et al., 2001, 2008) and the ODP Site 1262 benthic isotopes (Littler et al., 2014, with ages adjusted as per Westerhold et al., 2015) are available via their published sources.

5 **Conflicts of Interest**

The authors declare that they have no conflict of interest.

References

Archer, D., Buffett, B. and Brovkin, V.: Ocean methane hydrates as a slow tipping point in the global carbon cycle., Proc. Natl. Acad. Sci. U. S. A., 106(49), 20596–20601, doi:10.1073/pnas.0800885105, 2009.

15 Armstrong McKay, D. I., Tyrrell, T., Wilson, P. A. and Foster, G. L.: Estimating the impact of the cryptic degassing of Large Igneous Provinces: A mid-Miocene case-study, Earth Planet. Sci. Lett., 403, 254–262, doi:10.1016/j.epsl.2014.06.040, 2014.

Batjes, N. H.: Total carbon and nitrogen in the soils of the world, Eur. J. Soil Sci., 47(2), 151–163, doi:10.1111/j.1365-2389.1996.tb01386.x, 1996.

Berner, R. A.: A model for atmospheric CO2 over Phanerozoic time, Am. J. Sci., 291(4), 339–376, doi:10.2475/ajs.291.4.339,
20 1991.

Berner, R. A., Lasaga, A. C. and Garrels, R. M.: The carbonate-silicate geochemical cycle and its effect on atmospheric carbon dioxide over the past 100 million years, Am. J. Sci., 283, 641–683, 1983.

Boettiger, C. and Hastings, A.: Early warning signals and the prosecutor's fallacy., Proc. Biol. Sci., 279(1748), 4734–9, doi:10.1098/rspb.2012.2085, 2012a.

25 Boettiger, C. and Hastings, A.: Quantifying limits to detection of early warning for critical transitions., J. R. Soc. Interface, 9(75), 2527–39, doi:10.1098/rsif.2012.0125, 2012b.

Boettiger, C., Ross, N. and Hastings, A.: Early warning signals: The charted and uncharted territories, Theor. Ecol., 6(3), 255–264, doi:10.1007/s12080-013-0192-6, 2013.

Boscolo-Galazzo, F., Crichton, K. A., Barker, S. and Pearson, P. N.: Temperature dependency of metabolic rates in the upper ocean: A positive feedback to global climate change?, Glob. Planet. Change, 170(March), 201–212, doi:10.1016/j.gloplacha.2018.08.017, 2018.

Buffett, B. and Archer, D.: Global inventory of methane clathrate: sensitivity to changes in the deep ocean, Earth Planet. Sci. Lett., 227(3–4), 185–199, doi:10.1016/j.epsl.2004.09.005, 2004.

Carpenter, S. R. and Brock, W. A.: Rising variance: a leading indicator of ecological transition., Ecol. Lett., 9(3), 311–318, doi:10.1111/j.1461-0248.2005.00877.x, 2006.

Carpenter, S. R. and Brock, W. A.: Early warnings of unknown nonlinear shifts: A nonparametric approach, Ecology, 92(12), 2196–2201, doi:10.1890/11-0716.1, 2011.

Caves, J. K., Jost, A. B., Lau, K. V. and Maher, K.: Cenozoic carbon cycle imbalances and a variable weathering feedback, Earth Planet. Sci. Lett., 450, 152–163, doi:10.1016/j.epsl.2016.06.035, 2016.

Chamberlin, T. C.: The Influence of Great Epochs of Limestone Formation upon the Constitution of the Atmosphere, J. Geol., 6(6), 609–621 [online] Available from: http://www.jstor.org/stable/30054731 Accessed:, 1898.

Cox, J. C. and Ross, S. A.: The valuation of options for alternative stochastic processes, J. financ. econ., 3(1–2), 145–166, doi:10.1016/0304-405X(76)90023-4, 1976.

Cui, Y., Kump, L. R., Ridgwell, A. J., Charles, A. J., Junium, C. K., Diefendorf, A. F., Freeman, K. H., Urban, N. M. and Harding, I. C.: Slow release of fossil carbon during the Palaeocene–Eocene Thermal Maximum, Nat. Geosci., 4(7), 481–485, doi:10.1038/ngeo1179, 2011.

Dakos, V., Scheffer, M., van Nes, E. H., Brovkin, V., Petoukhov, V. and Held, H.: Slowing down as an early warning signal for abrupt climate change., Proc. Natl. Acad. Sci. U. S. A., 105(38), 14308–12, doi:10.1073/pnas.0802430105, 2008.

Dakos, V., Carpenter, S. R., Brock, W. a., Ellison, A. M., Guttal, V., Ives, A. R., Kéfi, S., Livina, V., Seekell, D. a., van Nes, E. H. and Scheffer, M.: Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data, PLoS One, 7(7), doi:10.1371/journal.pone.0041010, 2012a.

Dakos, V., van Nes, E. H., D'Odorico, P. and Scheffer, M.: Robustness of variance and autocorrelation as indicators of critical slowing down, Ecology, 93(2), 264–271, doi:10.1890/11-0889.1, 2012b.

Dakos, V., Nes, E. H. and Scheffer, M.: Flickering as an early warning signal, Theor. Ecol., 6(3), 309–317, doi:10.1007/s12080-013-0186-4, 2013.

Dakos, V., Carpenter, S. R., van Nes, E. H. and Scheffer, M.: Resilience indicators: prospects and limitations for early warnings of regime shifts, Philos. Trans. R. Soc. B Biol. Sci., 370(1659), 20130263–20130263, doi:10.1098/rstb.2013.0263, 2015.

DeConto, R. M. and Pollard, D.: Rapid Cenozoic glaciation of Antarctica induced by declining atmospheric CO2., Nature, 421(6920), 245–9, doi:10.1038/nature01290, 2003.

DeConto, R. M., Pollard, D., Wilson, P. A., Pälike, H., Lear, C. H. and Pagani, M.: Thresholds for Cenozoic bipolar glaciation, Nature, 455(7213), 652–6, doi:10.1038/nature07337, 2008.

DeConto, R. M., Galeotti, S., Pagani, M., Tracy, D., Schaefer, K., Zhang, T., Pollard, D. and Beerling, D. J.: Past extreme

warming events linked to massive carbon release from thawing permafrost., Nature, 484(7392), 87–91, doi:10.1038/nature10929, 2012.

Derry, L. A. and France-Lanord, C.: Neogene growth of the sedimentary organic carbon reservoir, Paleoceanography, 11(3), 267, doi:10.1029/95PA03839, 1996.

5 Dickens, G. R.: Down the Rabbit Hole: toward appropriate discussion of methane release from gas hydrate systems during the Paleocene-Eocene thermal maximum and other past hyperthermal events, Clim. Past, 7(3), 831–846, doi:10.5194/cp-7-831-2011, 2011.

Dickens, G. R., Neil, J. R. O., Rea, D. K. and Owen, R. M.: Dissociation of oceanic methane hydrate as a cause of the carbon isotope excursion at the end of the Paleocene, Paleoceanography, 10(6), 965–971, 1995.

10 Ditlevsen, P. D. and Johnsen, S. J.: Tipping points: Early warning and wishful thinking, Geophys. Res. Lett., 37(L19703), 1–4, doi:10.1029/2010GL044486, 2010.

Dunkley Jones, T., Lunt, D. J., Schmidt, D. N., Ridgwell, A., Sluijs, A., Valdes, P. J. and Maslin, M.: Climate model and proxy data constraints on ocean warming across the Paleocene-Eocene Thermal Maximum, Earth-Science Rev., 125, 123–145, doi:10.1016/j.earscirev.2013.07.004, 2013.

15 France-Lanord, C. and Derry, L. A.: Organic carbon burial forcing of the carbon cycle from Himalayan erosion, Nature, 390(November), 65–67, 1997.

Frieling, J., Svensen, H. H., Planke, S., Cramwinckel, M. J., Selnes, H. and Sluijs, A.: Thermogenic methane release as a cause for the long duration of the PETM, , 1–6, doi:10.1073/pnas.1603348113, 2016.

Frieling, J., Gebhardt, H., Huber, M., Adekeye, O. A., Akande, S. O., Reichart, G.-J., Middelburg, J. J., Schouten, S. and Sluijs,
20 A.: Extreme warmth and heat-stressed plankton in the tropics during the Paleocene-Eocene Thermal Maximum, Sci. Adv., 3(3), e1600891, doi:10.1126/sciadv.1600891, 2017.

Grimm, V. and Wissel, C.: Babel, or the ecological stability discussions: an inventory and analysis of terminology and a guide for avoiding confusion, Oecologia, 109(3), 323–334, doi:10.1007/s004420050090, 1997.

Gutjahr, M., Ridgwell, A., Sexton, P. F., Anagnostou, E., Pearson, P. N., Pälike, H., Norris, R. D., Thomas, E. and Foster, G.
25 L.: Very large release of mostly volcanic carbon during the Palaeocene–Eocene Thermal Maximum, Nature, 548(7669), 573–577, doi:10.1038/nature23646, 2017.

Hay, W. W.: Potential Errors in Estimates of Carbonate Rock Accumulating Through Geologic Time, in The Carbon Cycle and Atmospheric CO2: Natural Variations Archean to Present, vol. 32, edited by E. T. Sundquist and W. S. Broecker, pp. 573–583, American Geophysical Union, Washington, D. C., 1985.

30 Held, H. and Kleinen, T.: Detection of climate system bifurcations by degenerate fingerprinting, Geophys. Res. Lett., 31(23), 1–4, doi:10.1029/2004GL020972, 2004.

Hofmann, M. and Rahmstorf, S.: On the stability of the Atlantic meridional overturning circulation, Proc. Natl. Acad. Sci., 106(49), 20584–20589, doi:10.1073/pnas.0909146106, 2009.

Holling, C. S.: Resilience and Stability of Ecological Systems, Annu. Rev. Ecol. Syst., 4, 1–23 [online] Available from:

http://pure.iiasa.ac.at/id/eprint/26/1/RP-73-003.pdf, 1973.

Johannes, M.: The Statistical and Economic Role of Jumps in Continuous-Time Interest Rate Models, J. Finance, 59(1), 227–260, doi:10.1111/j.1540-6321.2004.00632.x, 2004.

John, E. H., Wilson, J. D., Pearson, P. N. and Ridgwell, A.: Temperature-dependent remineralization and carbon cycling in the warm Eocene oceans, Palaeogeogr. Palaeoclimatol. Palaeoecol., 413, 158–166, doi:10.1016/j.palaeo.2014.05.019, 2014.

Kéfi, S., Dakos, V., Scheffer, M., Van Nes, E. H. and Rietkerk, M.: Early warning signals also precede non-catastrophic transitions, Oikos, 122(5), 641–648, doi:10.1111/j.1600-0706.2012.20838.x, 2013.

Kirtland Turner, S., Sexton, P. F., Charles, C. D. and Norris, R. D.: Persistence of carbon release events through the peak of early Eocene global warmth, Nat. Geosci., 7(10), 748–751, doi:10.1038/ngeo2240, 2014.

Kirtland Turner, S., Hull, P. M., Kump, L. R. and Ridgwell, A.: A probabilistic assessment of the rapidity of PETM onset, Nat. Commun., 8(1), 353, doi:10.1038/s41467-017-00292-2, 2017.

Klinger, L. F., Taylor, J. A. and Franzen, L. G.: The Potential Role of Peatland Dynamics in Ice-Age Initiation, Quat. Res., 45(1), 89–92, doi:10.1006/qres.1996.0008, 1996.

Komar, N., Zeebe, R. E. and Dickens, G. R.: Understanding long-term carbon cycle trends: The late Paleocene through the early Eocene, Paleoceanography, 28(4), 650–662, doi:10.1002/palo.20060, 2013.

Kump, L. R. and Arthur, M. A.: Global Chemical Erosion during the Cenozoic: Weatherability Balances the Budgets, in Tectonic uplift and climate change, edited by W. F. Ruddiman, pp. 399–426, Springer US, Boston, MA., 1997.

Kurtz, A. C., Kump, L. R., Arthur, M. A., Zachos, J. C. and Paytan, A.: Early Cenozoic decoupling of the global carbon and sulfur cycles, Paleoceanography, 18(4), doi:10.1029/2003PA000908, 2003.

Lenton, T. M.: Early warning of climate tipping points, Nat. Clim. Chang., 1(4), 201–209, doi:10.1038/nclimate1143, 2011.

Lenton, T. M.: Environmental Tipping Points, Annu. Rev. Environ. Resour., 38(1), 1–29, doi:10.1146/annurev-environ-102511-084654, 2013.

Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S. and Schellnhuber, H. J.: Tipping elements in the Earth's climate system, Proc. Natl. Acad. Sci., 105(6), 1786–1793, doi:10.1073/pnas.0705414105, 2008.

Lenton, T. M., Livina, V. N., Dakos, V. and Scheffer, M.: Climate bifurcation during the last deglaciation?, Clim. Past, 8(4), 1127–1139, doi:10.5194/cp-8-1127-2012, 2012a.

Lenton, T. M., Livina, V. N., Dakos, V., van Nes, E. H. and Scheffer, M.: Early warning of climate tipping points from critical slowing down: comparing methods to improve robustness., Philos. Trans. A. Math. Phys. Eng. Sci., 370(1962), 1185–204, doi:10.1098/rsta.2011.0304, 2012b.

Littler, K., Röhl, U., Westerhold, T. and Zachos, J. C.: A high-resolution benthic stable-isotope record for the South Atlantic: Implications for orbital-scale changes in Late Paleocene–Early Eocene climate and carbon cycling, Earth Planet. Sci. Lett., 401, 18–30, doi:10.1016/j.epsl.2014.05.054, 2014.

Livina, V. N. and Lenton, T. M.: A modified method for detecting incipient bifurcations in a dynamical system, Geophys. Res. Lett., 34(3), 1–5, doi:10.1029/2006GL028672, 2007.

Livina, V. N., Ditlevsen, P. D. and Lenton, T. M.: An independent test of methods of detecting system states and bifurcations in time-series data, Phys. A Stat. Mech. its Appl., 391(3), 485–496, doi:10.1016/j.physa.2011.08.025, 2012.

Lourens, L. J., Sluijs, A., Kroon, D., Zachos, J. C., Thomas, E., Röhl, U., Bowles, J. and Raffi, I.: Astronomical pacing of late Palaeocene to early Eocene global warming events., Nature, 435(7045), 1083–7, doi:10.1038/nature03814, 2005.

5   Lunt, D. J., Ridgwell, A., Sluijs, A., Zachos, J., Hunter, S. and Haywood, A.: A model for orbital pacing of methane hydrate destabilization during the Palaeogene, Nat. Geosci., 4(11), 775–778, doi:10.1038/ngeo1266, 2011.

McInerney, F. A. and Wing, S. L.: The Paleocene-Eocene Thermal Maximum: A Perturbation of Carbon Cycle, Climate, and Biosphere with Implications for the Future, Annu. Rev. Earth Planet. Sci., 39(1), 489–516, doi:10.1146/annurev-earth-040610-133431, 2011.

10   Minshull, T. A., Marín-Moreno, H., Armstrong McKay, D. I. and Wilson, P. A.: Mechanistic insights into a hydrate contribution to the Paleocene-Eocene carbon cycle perturbation from coupled thermohydraulic simulations, Geophys. Res. Lett., 43(16), 8637–8644, doi:10.1002/2016GL069676, 2016.

van Nes, E. H., Arani, B. M. S., Staal, A., van der Bolt, B., Flores, B. M., Bathiany, S. and Scheffer, M.: What Do You Mean, 'Tipping Point'?, Trends Ecol. Evol., 31(12), 902–904, doi:10.1016/j.tree.2016.09.011, 2016.

15   Pagani, M., Huber, M., Liu, Z., Bohaty, S. M., Henderiks, J., Sijp, W., Krishnan, S. and DeConto, R. M.: The role of carbon dioxide during the onset of Antarctic glaciation., Science, 334(6060), 1261–4, doi:10.1126/science.1203909, 2011.

Peng, C. K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E. and Goldberger, A. L.: Mosaic organization of DNA nucleotides, Phys. Rev. E, 49(2), 1685–1689, doi:10.1103/PhysRevE.49.1685, 1994.

van der Ploeg, R., Selby, D., Cramwinckel, M. J., Li, Y., Bohaty, S. M., Middelburg, J. J. and Sluijs, A.: Middle Eocene

20   greenhouse warming facilitated by diminished weathering feedback, Nat. Commun., 9(1), 2877, doi:10.1038/s41467-018-05104-9, 2018.

Pollard, D. and DeConto, R. M.: Modelling West Antarctic ice sheet growth and collapse through the past five million years, Nature, 458(7236), 329–332, doi:10.1038/nature07809, 2009.

R Foundation for Statistical Computing: R: A language and environment for statistical computing, [online] Available from:

25   https://www.r-project.org/, 2016.

Rahmstorf, S.: Ocean circulation and climate during the past 120,000 years, Nature, 419(6903), 207–214, doi:10.1038/nature01090, 2002.

Schaller, M. F., Fung, M. K., Wright, J. D., Katz, M. E. and Kent, D. V: Impact ejecta at the Paleocene-Eocene boundary, Science (80-.)., 354(6309), 225–229, doi:10.1126/science.aaf5466, 2016.

30   Scheffer, M., Carpenter, S., Foley, J. A., Folke, C. and Walker, B.: Catastrophic shifts in ecosystems, Nature, 413(6856), 591–596, doi:10.1038/35098000, 2001.

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., van Nes, E. H., Rietkerk, M. and Sugihara, G.: Early-warning signals for critical transitions., Nature, 461(7260), 53–9, doi:10.1038/nature08227, 2009.

Secord, R., Gingerich, P. D., Lohmann, K. C. and MacLeod, K. G.: Continental warming preceding the Palaeocene-Eocene

thermal maximum, Nature, 467(7318), 955–958, doi:10.1038/nature09441, 2010.

Sluijs, A., Brinkhuis, H., Schouten, S., Bohaty, S. M., John, C. M., Zachos, J. C., Reichart, G. J., Sinninghe Damsté, J. S., Crouch, E. M. and Dickens, G. R.: Environmental precursors to rapid light carbon injection at the Palaeocene/Eocene boundary, Nature, 450(7173), 1218–1221, doi:10.1038/nature06400, 2007a.

5   Sluijs, A., Bowen, G. J., Brinkhuis, H., Lourens, L. J. and Thomas, E.: The Palaeocene-Eocene Thermal Maximum super greenhouse: Biotic and geochemical signatures, age models and mechanisms of global change, Geol. Soc. Spec. Publ., 1, 323–349, doi:1747-602X/07/$15.00, 2007b.

Stap, L., Lourens, L. J., Thomas, E., Sluijs, A., Bohaty, S. and Zachos, J. C.: High-resolution deep-sea carbon and oxygen isotope records of Eocene Thermal Maximum 2 and H2, Geology, 38(7), 607–610, doi:10.1130/G30777.1, 2010.

10   Stocker, T. F. and Wright, D. G.: Rapid transitions of the ocean's deep circulation induced by changes in surface water fluxes, Nature, 351(6329), 729–732, doi:10.1038/351729a0, 1991.

Stommel, H.: Thermohaline Convection with Two Stable Regimes, Tellus, 13(2), 224–230, 1961.

Storey, M., Duncan, R. A. and Tegner, C.: Timing and duration of volcanism in the North Atlantic Igneous Province: Implications for geodynamics and links to the Iceland hotspot, Chem. Geol., 241(3–4), 264–281,
15   doi:10.1016/j.chemgeo.2007.01.016, 2007.

Svensen, H., Planke, S., Malthe-Sørenssen, A., Jamtveit, B., Myklebust, R., Rasmussen Eidem, T. and Rey, S. S.: Release of methane from a volcanic basin as a mechanism for initial Eocene global warming, Nature, 429(6991), 542–545, doi:10.1038/nature02566, 2004.

Tarnocai, C., Canadell, J. G., Schuur, E. A. G., Kuhry, P., Mazhitova, G. and Zimov, S.: Soil organic carbon pools in the
20   northern circumpolar permafrost region, Global Biogeochem. Cycles, 23(GB2023), 1–11, doi:10.1029/2008GB003327, 2009.

The MathWorks Inc.: Matlab Release 2016a, 2016.

Urey, H. C.: The planets, their origin and development, Mrs. Hepsa Ely Silliman Meml. Lect. Yale Univ. [online] Available from: http://adsabs.harvard.edu/abs/1952ptod.conf.....U, 1952.

25   Walker, J. C. G. and Kasting, J. F.: Effects of fuel and forest conservation on future levels of atmospheric carbon dioxide., Palaeogeogr. Palaeoclimatol. Palaeoecol., 97(1902), 151–89 [online] Available from: http://www.ncbi.nlm.nih.gov/pubmed/11537854, 1992.

Walker, J. C. G., Hays, P. B. and Kasting, J. F.: A negative feedback mechanism for the long-term stabilization of Earth's surface temperature, J. Geophys. Res., 86(C10), 9776–9782, 1981.

30   Wang, R., Dearing, J. A., Langdon, P. G., Zhang, E., Yang, X., Dakos, V. and Scheffer, M.: Flickering gives early warning signals of a critical transition to a eutrophic lake state., Nature, 492(7429), 419–22, doi:10.1038/nature11655, 2012.

Westerhold, T. and Rohl, U.: High resolution cyclostratigraphy of the early Eocene – new insights into the origin of the Cenozoic cooling trend, Clim. Past, 5, 309–327, 2009.

Westerhold, T., Röhl, U., Laskar, J., Raffi, I., Bowles, J., Lourens, L. J. and Zachos, J. C.: On the duration of magnetochrons

C24r and C25n and the timing of early Eocene global warming events: Implications from the Ocean Drilling Program Leg 208 Walvis Ridge depth transect, Paleoceanography, 22(2), 1–19, doi:10.1029/2006PA001322, 2007.

Westerhold, T., Röhl, U., Frederichs, T., Bohaty, S. M. and Zachos, J. C.: Astronomical calibration of the geological timescale: Closing the middle Eocene gap, Clim. Past, 11(9), 1181–1195, doi:10.5194/cp-11-1181-2015, 2015.

5   Westerhold, T., Röhl, U., Wilkens, R. H., Gingerich, P. D., Clyde, W. C., Wing, S. L., Bowen, G. J. and Kraus, M. J.: Synchronizing early Eocene deep-sea and continental records – cyclostratigraphic age models for the Bighorn Basin Coring Project drill cores, Clim. Past, 14(3), 303–319, doi:10.5194/cp-14-303-2018, 2018.

Zachos, J. C., Pagani, M., Sloan, L., Thomas, E. and Billups, K.: Trends, rhythms, and aberrations in global climate 65 Ma to present., Science, 292(5517), 686–93, doi:10.1126/science.1059412, 2001.

10 Zachos, J. C., Röhl, U., Schellenberg, S. A., Sluijs, A., Hodell, D. A., Kelly, D. C., Thomas, E., Nicolo, M., Raffi, I., Lourens, L. J., McCarren, H. and Kroon, D.: Rapid acidification of the ocean during the Paleocene-Eocene thermal maximum., Science, 308(5728), 1611–5, doi:10.1126/science.1109004, 2005.

Zachos, J. C., Dickens, G. R. and Zeebe, R. E.: An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics., Nature, 451(7176), 279–83, doi:10.1038/nature06588, 2008.

15 Zeebe, R. E., Zachos, J. C. and Dickens, G. R.: Carbon dioxide forcing alone insufficient to explain Palaeocene–Eocene Thermal Maximum warming, Nat. Geosci., 2(8), 576–580, doi:10.1038/ngeo578, 2009.

Zeebe, R. E., Ridgwell, A. and Zachos, J. C.: Anthropogenic carbon release rate unprecedented during the past 66 million years, Nat. Geosci., (March), 1–5, doi:10.1038/ngeo2681, 2016.
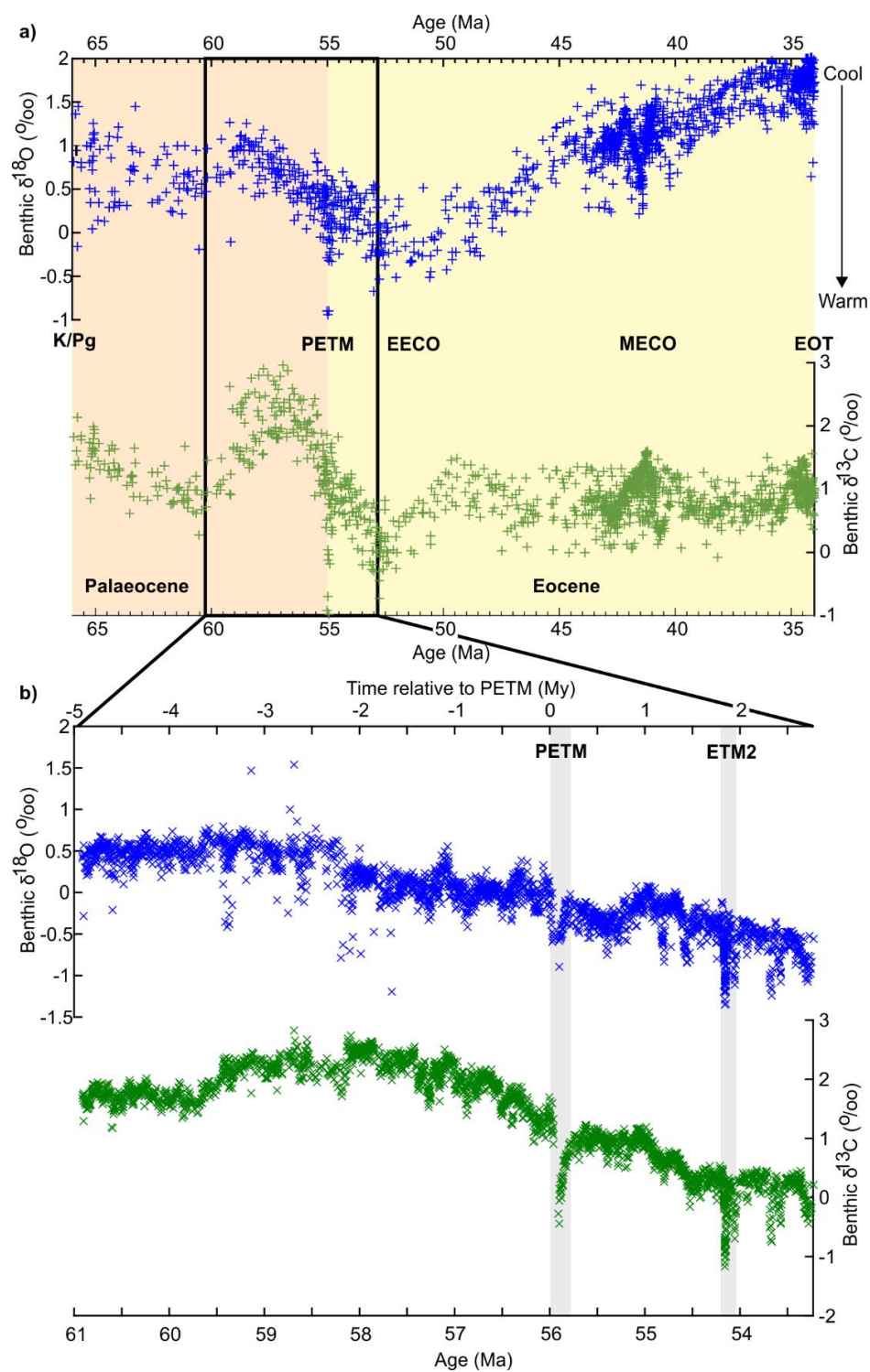
Figure 1: Palaeorecords of benthic δ¹⁸O (blue) and δ¹³C (green) across a) the Palaeocene and Eocene (data from global stack (Zachos et al., 2001, 2008)) and b) the Late Palaeocene-early Eocene (LPEE) study interval (data from ODP Site 1262 (Littler et al., 2014) with ages adjusted as per (Westerhold et al., 2015)). Significant climate and carbon cycle events are labelled, including the Cretaceous/Palaeogene boundary (K/Pg), Palaeocene-Eocene Thermal Maximum (PETM), Eocene Thermal Maximum 2 (ETM2), Early Eocene Climatic Optimum (EECO), the Mid-Eocene Climatic Optimum (MECO), and the Eocene-Oligocene Transition (EOT), while the black box marks the LPEE interval. The mismatch in PETM date between the two datasets is a result of the updated age model of (Westerhold et al., 2015) applied to the record of (Littler et al., 2014), which we have left as-is to maintain consistency with published records.
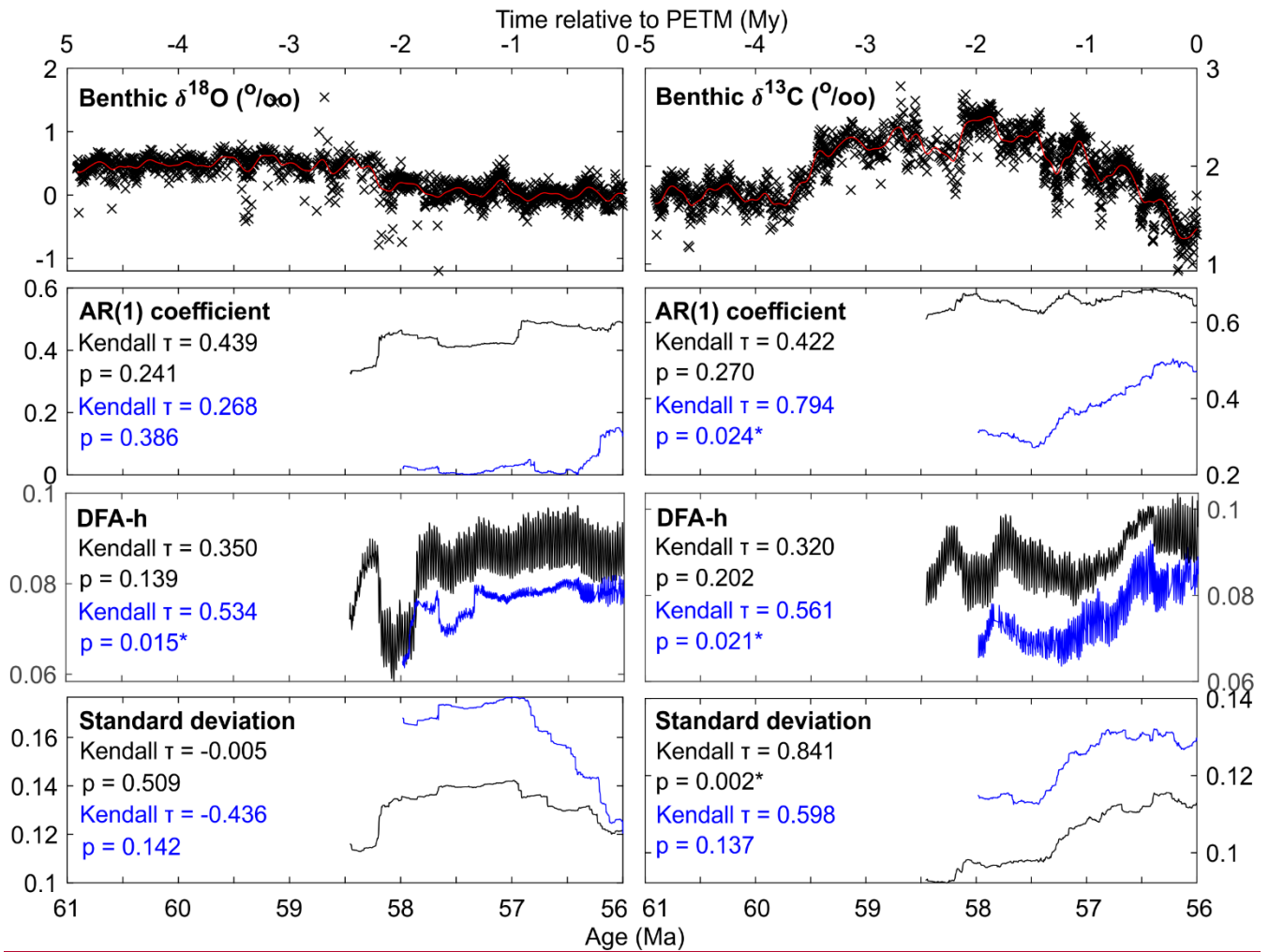
Figure 2: Rolling window resilience analysis of benthic δ¹⁸O (left) and δ¹³C (right) in the run-up to the PETM. The top panels illustrate the palaeorecord (black crosses) and the detrending applied to the data (red line), with the panels below illustrating the results of the analysis for AR(1) coefficient, detrended fluctuation analysis h-value, and standard deviation calculated in a 50% rolling window across each time-series for both interpolated (black line) and non-interpolated (blue line) data with the Kendall τ rank-correlation and boostrapped p-value for each. Results for skewness, kurtosis, and sensitivity analyses for all metrics can be found in the Supplementary Material.
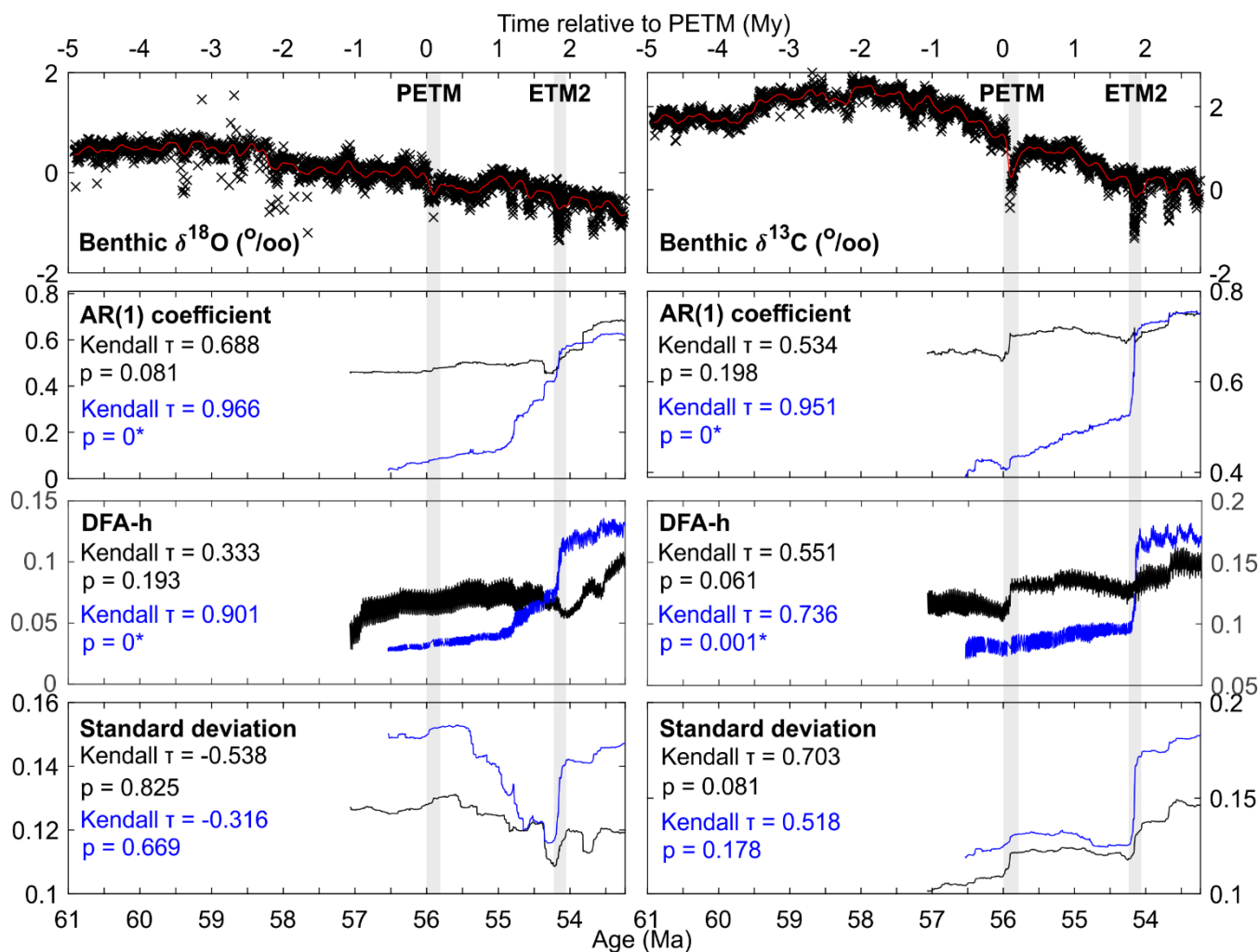
Figure 3: Rolling window resilience analysis of benthic δ¹⁸O (left) and δ¹³C (right) across the PETM and ETM2. The top panels illustrate the palaeorecords (black crosses) and the smoothed record used to detrending the data (red line), with the panels below illustrating the results of the analysis for AR(1) coefficient, detrended fluctuation analysis h-value, and standard deviation calculated in a 50% rolling window across each time-series for both interpolated (black line) and non-interpolated (blue line) data with the Kendall τ rank-correlation and boostrapped p-value for each. The PETM and ETM2 are marked by the grey bars. Results for skewness, kurtosis, and sensitivity analyses for all metrics can be found in the Supplementary Material.

28
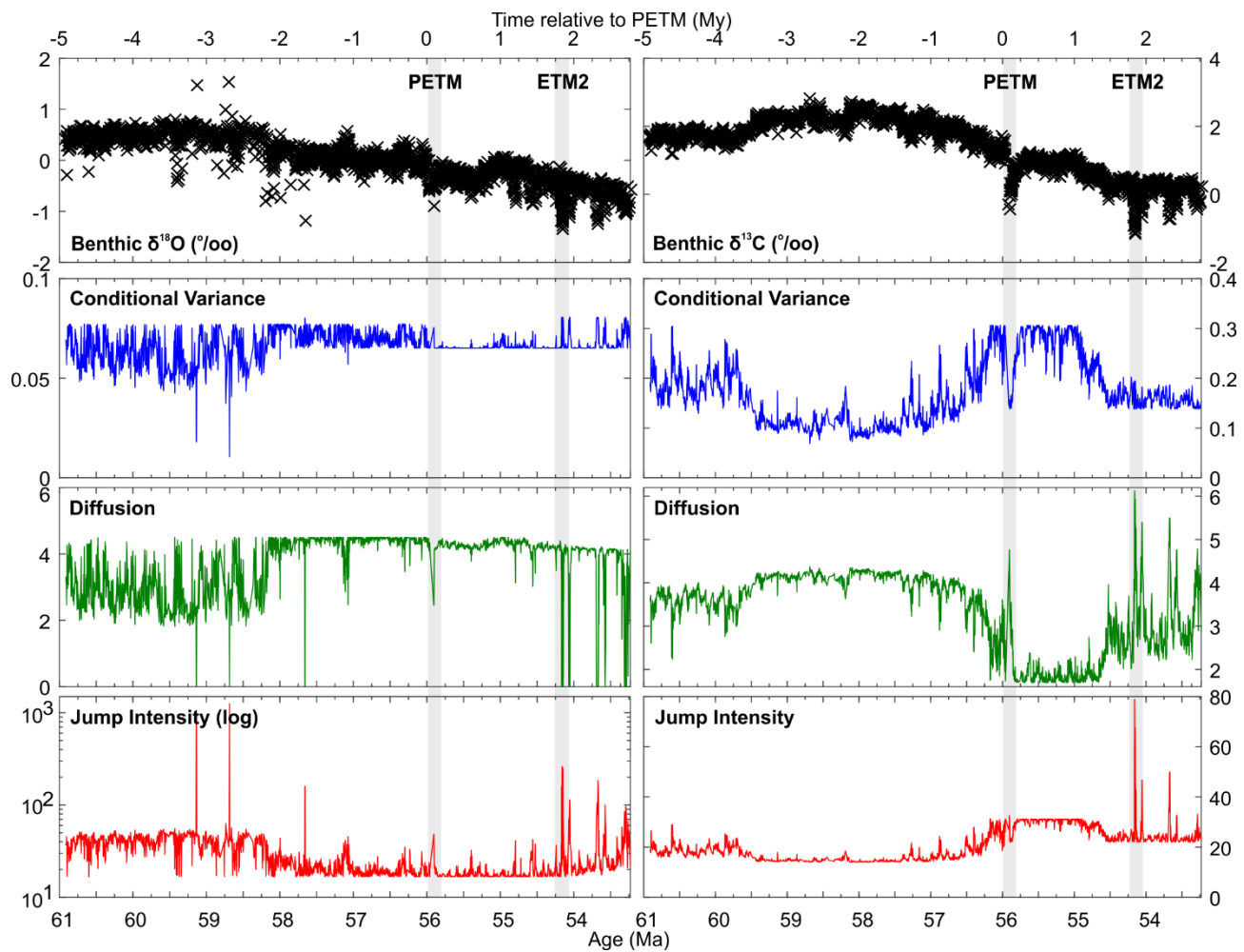
Figure 4: Non-parametric drift-diffusion-jump model functions of benthic δ¹⁸O (left) and δ¹³C (right) across the PETM and ETM2.
The top panels illustrate the palaeorecords (black crosses), with the panels below illustrating the model functions for conditional variance (blue line), diffusion (green line), and jump intensity (red line) for each palaeorecord (drift is not plotted). The PETM and ETM2 are marked by the grey bars (left-most and right-most respectively).

Tables

Table 1: Values of binned metrics for both benthic δ¹⁸O (top) and δ¹³C (bottom). Bins are of all data (detrended but not interpolated) before the PETM, all data after the PETM (including data from within ETM2), all data between the PETM and ETM2, and all data after ETM2. The indicators are AR(1) coefficient, detrended fluctuation analysis h-value (DFA-h), standard deviation (SD), skewness (SKEW), and kurtosis (KURT). Green shading indicates the indicator has increased relative to before the event (either the PETM or ETM2), red shading indicates a decrease in value. Each value is followed by a p-value (*in italics within parentheses preceded by a * if significant*) computed using a permutation test, except for AR(1) for which we instead use AR(1) model-derived surrogate data to compare against (see Methods for details).

Metric	Before PETM	After PETM (all)	After PETM (to ETM2)	After ETM2
δ¹⁸O				
AR(1)	0.0611	0.659 *(0)	0.509 *(0)	0.657 *(0)
DFA-h	0.035	0.140 *(0)	0.098 *(0)	0.187 *(0)
SD	0.146	0.152 (0.279)	0.119 *(0.991)	0.145 *(0.002)
SKEW	-1.447	-0.635 *(0.098)	-0.927 (0.281)	-0.713 (0.211)
KURT	14.621	4.905 *(1)	4.490 *(0.999)	3.884 (0.741)
δ¹³C				
AR(1)	0.402	0.784 *(0)	0.500 *(0.001)	0.743 *(0)
DFA-h	0.071	0.189 *(0)	0.088* (0.504)	0.203 *(0)
SD	0.122	0.189 *(0)	0.105 *(0.999)	0.187 *(0)
SKEW	-0.541	-1.165 *(0.98)	-0.700 (0.845)	-1.056 (0.828)
KURT	4.046	7.634 *(0.005)	4.073 (0.415)	4.728 (0.307)