Review of “Spatially contrasted response of Devonian anoxia to astronomical forcing”. Justin Gérard et al. 2025, for Climate of the Past

First Reviewer: Anonymous Referee #1

In this manuscript, Gérard et al. explore the impact of astronomical parameters on marine oxygen levels in a Devonian context using cGENIE model experiments with an updated, emulator-based, weathering scheme. The 1.1 Myr transient cGENIE simulations are forced with prescribed surface albedo and wind stress from a specific HadCM3 simulation of Valdes et al. (2021) and variable continental phosphorus inputs that evolve following a 1.1 Myr synthetic Devonian astronomical forcing that includes transitions between 405-kyr-modulated 100-kyr cycles and a 2.4-Myr eccentricity node without 100-kyr cycles. Following the methods described in Sablon et al. (2025), the phosphorus inputs to the ocean are emulated based on GEOCLIM simulations forced with climate conditions obtained from HadSM3 simulations with various combinations of orbital parameters and CO_2. 3 simulations are evaluated with either the full orbital parameters evolving together or either obliquity or eccentricity/precession fixed to their mean values across the 1.1 Myr astronomical solution. The main result is that astronomy weakly influence the global mean concentration in O_2 and PO_4 but may have a greater impact on regional water mass characteristics, via inputs of weathered PO_4 to the proximal ocean. The results support the hypothesis that Late Devonian anoxic events occurred close to eccentricity maxima as the latter forcing yields the largest anoxic volume in the simulations discussed in this manuscript.

I think this manuscript is worthy of publication after revisions. I have particularly liked the key message that the astronomical signal may have different regional expression on oceanic PO_4 and O_2. The links between the astronomical forcing and the simulated PO_4 and O_2 ocean concentrations are adequately illustrated, even if they are somehow expected from the start given the simplicity of the nutrient cycle – marine productivity is function of the concentration of PO_4 – and that the astronomically-varying input of PO_4 to the ocean is the only transient forcing across the simulations. The regional analysis is interesting but the impact of astronomy on O_2 variability is in my opinion a bit oversold in some of the regions discussed.

We thank the Reviewer for their overall positive evaluation of the manuscript and constructive comments.

My major concern is about the absence of astronomically-varying climate change in the simulations. If I understand correctly, the surface albedo and wind stress field are prescribed using outputs from a Devonian HadCM3 simulation with the same CO_2 as Valdes et al. (2021) – which is what exactly? because Valdes et al. present 370 Ma simulations with 2 different pCO_2 of 926 and 811 ppmv – and orbital values of 0 and 23° for eccentricity and obliquity and a longitude of perihelion of 0°. Now with very different orbital parameters, e.g. high eccentricity and/or obliquity, albedo and wind stress will change, perhaps significantly, and I suspect that it would drive ocean circulation changes leading to changes in marine O_2 concentrations of at least similar, if not much larger, amplitude.

First, we acknowledge that additional clarification should have been provided regarding the exact experiment from Valdes et al. (2021) used to prescribe the surface boundary conditions and we thank the reviewer for point this out. The atmospheric fields used as boundary conditions in GENIE are derived from the simulation with a pCO_2 of 926 ppm. We therefore revised the manuscript as follows:

Lines 107-110: "The surface albedo and zonally averaged wind stress profiles are prescribed from the HadCM3L simulation of Sablon et al. (2025), conducted with the same pCO_2 (926 ppm) as in Valdes et al. (2021) and an orbital configuration characterized by an eccentricity of 0, a longitude of perihelion of 0°, and an obliquity of 23°."

Second, the Reviewer's concern regarding the absence of astronomically-varying climate change in our simulations, particularly the lack of orbital modulation of wind stress and albedo, is entirely sounded. However, we kept these fields constant for two main reasons. (1) These GCM-derived fields already represent a substantial improvement over the idealized fields generated by default in muffingen (the GENIE boundary condition generator – https://zenodo.org/records/5500687) and used in previous Devonian studies (Gérard et al., 2025; their Fig. 1) and other work (Cermeño et al., 2022). These idealized fields consist of a simplified zonally averaged wind stress profile and an albedo field that depends only on latitude. (2) More importantly, our intent in this contribution is to test the hypothesis proposed in the data community, that the main mechanism by which astronomical forcing influences ocean anoxia is through variations in nutrient fluxes from the continents to the ocean. This "top-down" mechanism has been proposed and supported in several studies, including Carmichael et al. (2019) and De Vleeschouwer et al. (2017) for the Frasnian-Famennian boundary, Monteiro et al. (2012) for Oceanic Anoxic Event 2, and Hülse et al. (2021) for the Permian–Triassic boundary. In this context, holding albedo and wind stress invariant through time allows us to better isolate and quantify the exact impact of this mechanism without the confounding influence of circulation-induced feedbacks. The focus on the "top-down" hypothesis is already made explicit in the manuscript (lines 3-4): "Here, we explore how astronomical forcing influences ocean oxygenation by modulating the continental weathering flux of phosphate within a Late Devonian climate framework."

We do realize that keeping these fields constant constitutes an important assumption of our current setup, and therefore acknowledge this point in the manuscript (lines 380-382). To best address the Reviewer's concern, we revised the introduction to summarise the main mechanism through which astronomical forcing could influence anoxia and expanded the discussion to better remind the goal of the study and further motivate the use of dynamic albedo and wind-stress profiles in future work.

Lines 26-29: "There is also growing evidence that astronomical forcing could modulate at least some of these events, primarily through a "top-down" mechanism (Carmichael et al., 2019) whereby cyclic changes in temperature and precipitation alter continental nutrient fluxes to the ocean (De Vleeschouwer et al., 2014; Martinez and Dera, 2015; De Vleeschouwer et al., 2017; Wichern et al., 2024; Huygh et al., 2025)."

Lines 387-389: "...better connect orbital forcing with ocean anoxia. For instance, subsequent work would benefit from incorporating dynamic albedo and wind-stress profiles when exploring the full range of climate-circulation–biogeochemistry interactions."

This being said, sensitivity tests previously conducted using an Ordovician continental configuration suggest that the global ocean circulation in GENIE is poorly sensitive to moderate changes in albedo and wind forcing (Fig. R1). This suggests that employing time-varying wind and albedo forcings would not drastically alter our results.

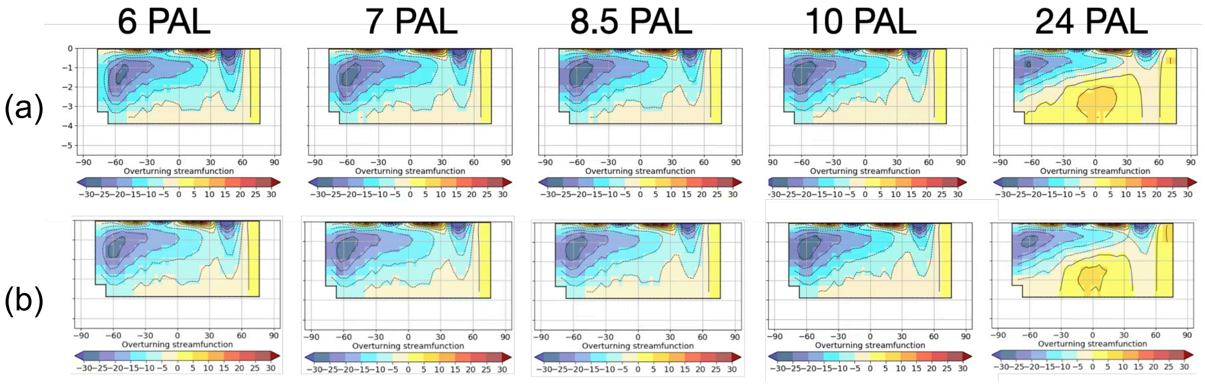

Figure R1: Sensitivity test to wind and albedo forcing. Global meridional overturning circulation (expressed in Sv; 1 Sv ≡ 10^6 m^3 s^{-1}) simulated for the Late Ordovician using cGENIE (a) with albedo and wind forcing derived at each atmospheric pCO_2 level (6 to 24 PAL; 1 PAL = 280 ppm) from general circulation model (FOAM) simulations conducted for matching pCO_2 values (hence adapting wind and albedo forcing in each GENIE simulation) vs. (b) with albedo and winds derived from a single FOAM simulation run at 10 PAL pCO_2 (hence only adapting pCO_2 in GENIE).

Cermeño, P., García-Comas, C., Pohl, A. et al. Post-extinction recovery of the Phanerozoic oceans and biodiversity hotspots. Nature 607, 507–511 (2022). https://doi.org/10.1038/s41586-022-04932-6

Hülse, D., Lau, K.V., van de Velde, S.J. et al. End-Permian marine extinction due to temperature-driven nutrient recycling and euxinia. Nat. Geosci. 14, 862–867 (2021). https://doi.org/10.1038/s41561-021-00829-7

Monteiro, F.M., Pancost, R.D., Ridgwell, A., Donnadieu, Y., 2012. Nutrients as the dominant control on the spread of anoxia and euxinia across the Cenomanian- Turonian oceanic anoxic event (OAE2): model-data comparison. Paleoceanography 27 (4), PA4209

Given the numerous HadCM3 simulations with different orbital parameters that I suppose available to the authors from what is written in this manuscript and in Sablon et al. (2025), I suggest to perform at least one cGENIE simulation with prescribed wind stress and albedo from a simulation with an 'extreme' orbital configuration among those available in the HadCM3 ensemble to evaluate how dynamical changes in ocean circulation affect PO_4 and O_2 in the regions defined in Fig. 4, as well as globally.

We thank Reviewer 1 for this interesting suggestion. We performed an additional experiment using the wind stress and albedo fields from a HadCM3L simulation forced with an extreme orbital configuration characterized by an eccentricity set to 0.08, an obliquity of 25° and a precession of 45°, to be compared with respectively 0, 23° and 0° in our original simulation. The consideration for such extreme orbital configurations is required for designing the emulator of Sablon et al., 2025, now accepted.

In this new simulation, only these fields were changed compared with the baseline simulations presented in this study, allowing us to specifically capture their impact on [O_2]. On the global scale, the difference in mean [O_2] between the two simulations is 3 μmol.kg^{-1}. At the regional scale, we find variations of 16, 3, 9, 8, and 10 μmol.kg^{-1} for regions SA, WL, Si, NP, and LG, respectively. These values can be directly compared to the [O_2] variability induced by astronomically driven weathering changes in our main experiments, where global-scale variations reach 5 μmol.kg^{-1} and regional amplitudes reach 45, 9, 6, 7, and 19 μmol.kg^{-1} for the same five regions.

Taken together, these results show that changes in wind stress and albedo do exert an influence on [O_2], but their effect is generally smaller than that associated with orbitally driven changes in continental phosphate weathering, particularly in regions exhibiting the largest [O_2] variability (SA and LG). These results are in line with the moderate impact of wind and albedo forcing on the global ocean circulation previously noted in Fig. R1. Most importantly, the numbers obtained here are uppermost end-members, as they were obtained under extreme orbital conditions that typically never occurred during the Devonian (Zeebe and Lantink, 2024). Although this analysis falls beyond the scope of the present study, we revised the manuscript as follows:

Lines 359-365: "To evaluate the impact of this assumption, we performed a sensitivity experiment using wind stress and albedo fields from an available HadCM3L simulation forced with an extreme orbital configuration (eccentricity of 0.08, obliquity of 25°, and precession of 45°), following the emulator design (Sablon et al., 2025). Even under these conditions, the resulting [O_2] variations (3 μmol.kg^{-1} globally and 16, 3, 9, 8, and 10 μmol.kg^{-1} for regions SA, WL, Si, NP, and LG) remain generally smaller than those driven by orbitally paced phosphate weathering in our main experiments (5 μmol.kg^{-1} globally and up to 45, 9, 6, 7, and 19 μmol.kg^{-1} regionally). More generally,..."

Zeebe, R. E., and Lantink, M. L. (2024). Milanković forcing in deep time. Paleoceanography and Paleoclimatology, 39(5), e2024PA004861, https://doi.org/10.1029/2024PA004861.

A related question is how does the transient astronomical changes affect the ocean circulation via local insolation and freshwater fluxes, as I guess the cGENIE atmospheric EMB still operates across the model runs? With the wind stress field prescribed, do these heat and freshwater fluxes lead to changes in ocean convection zones? If yes, does this play a role in the O_2 variability in the SA region? More generally, I think some ocean diagnostics could be useful to get a basic idea about the ocean circulation of this Devonian simulation.

Reviewer 1 is right, the EMBM of cGENIE remains active throughout the transient simulations, and as a result, ocean circulation evolves in response to the changing astronomical parameters. However, these variations remain minor, with a maximum change in overturning strength of 0.7 Sv when all astronomical parameters vary together (0.56 Sv with eccentricity + precession only, and 0.32 Sv with obliquity alone). This limited sensitivity is consistent with the results reported in our previous, dedicated investigation of the role of astronomical forcing on Devonian ocean circulation and anoxia (Gérard et al., 2025; their Fig. 7c). Importantly, [O_2] variability in region SA shows a strong correlation ($R^2 = 0.81$ and p-value $< 10^{-3}$) with regional biological productivity. These results indicate that while ocean circulation likely contributes to regional [O_2] changes, its influence is secondary compared to nutrient-driven processes. Isolating the specific role of circulation variability would require a dedicated analysis and is outside the primary objectives of the present study, which is to test the sensitivity of Devonian anoxia to astronomically-paced nutrient fluxes. This point has been clarified in the revised version of the manuscript in response to the Reviewer's previous points.

Nevertheless, the large-scale circulation patterns described in Gérard et al. (2025, see their series of simulations "Scot370M") remain directly applicable here, given the continental configuration used. To clarify this point in the revised version of our manuscript, we added the following statements:

Lines 102-104: "... 370 Ma (see Fig. 1). As this continental reconstruction was also used in Gérard et al. (2025), the large-scale ocean circulation patterns and associated diagnostics described there remain valid here; in particular, the overturning circulation shows only limited sensitivity to temperature changes (see their Fig. 7c)."

We also added the spatial distribution of deep water formation regions in the new Annex A. Lines 425-429: "Changes in ocean circulation associated with astronomical forcing are small in our simulations, with a maximum variation in overturning strength of 0.7 Sv (Sverdrup) when all orbital parameters vary together (0.56 Sv for eccentricity–precession forcing and 0.32 Sv for obliquity alone). Figure A2 further indicates that the main regions of deep-water formation remain largely unchanged between extreme orbital configurations (extracted from the 1.1 Myr transient simulation). Overall, the spatial distribution of deep-water formation explains the global overturning circulation shown in Fig. A1e."

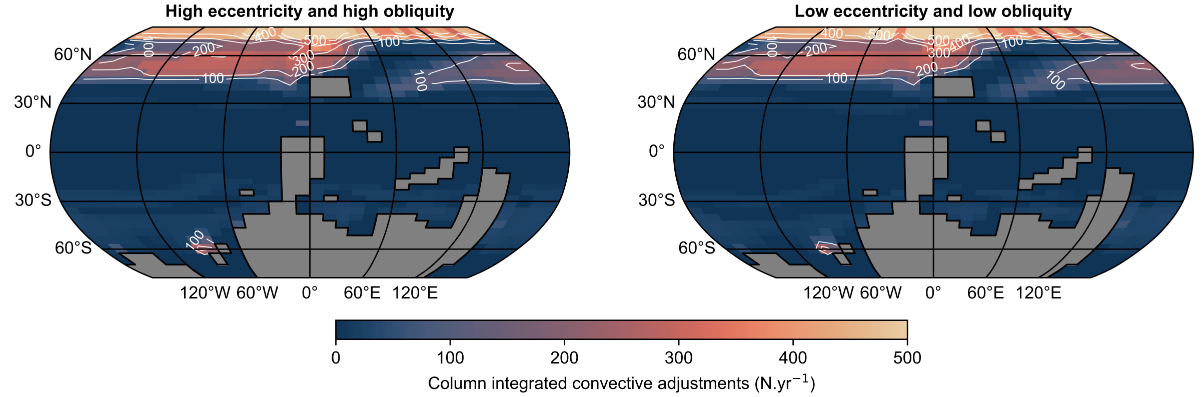

Figure R2: **Deep water formation.** Convective adjustments (expressed as the number of convective adjustments integrated across the water column per year) for **(a)** high eccentricity-high obliquity and **(b)** low eccentricity-low obliquity astronomical configurations.

Other comments.

1. Methods section.

The climate-weathering emulator construction is not very clear. The authors refer to Sablon et al. 2025 for the detailed description, which is fine, but the information laid out in this manuscript are ambiguous and/or not sufficiently detailed. Was HadSM3 or HadCM3, or both, used to construct the surface climate? If the methodology follows Sablon then it is probably the two of them but this should explicitly appear in the manuscript. If not, then how the procedure differs from Sablon must be clarified.

In the revised manuscript, we clarified and streamlined the description of the emulator, focusing specifically on the weathering fluxes it provides within our modelling framework and explicitly stating its reliance on the three models (HadCM3L, HadSM3, and GEOCLIM). All methodological and coupling details remain fully described in Sablon et al. (2025), ensuring a concise yet transparent presentation in the present study. The revised text read as follows:

Lines 77-88: "Specifically, we used the framework described in Sablon et al. (2025) to couple cGENIE with a climate–weathering emulator. This emulator computes weathering fluxes under varying boundary conditions (orbital parameters and pCO_2) based on statistical relationships derived from outputs of HadCM3L, HadSM3 (the latter being the slab-ocean version of HadCM3) and GEOCLIM (used with fixed present-day average lithological fractions), explicitly accounting for the coupled contributions of chemical weathering and physical erosion to continental weathering. Both HadCM3L and HadSM3 were used, as the 81 HadSM3 simulations used to train the emulator required heat convergence flux data obtained from HadCM3L. The weathering fluxes are instantaneously routed from continental source cells to the closest ocean coastal grid cell via a nearest-neighbour scheme, which remains suitable given the similar grid layouts, both adapted from Scotese and Wright (2018), and the limited knowledge of the Late Devonian hydrological network. A detailed overview of GEOCLIM's continental weathering module is presented in Maffre et al. (2025), while a complete description of the emulator's architecture, training process and topography is documented in Sablon et al. (2025) and repeated here for clarity (see Fig. 2)."

If the same HadCM3-HadSM3 methodology as Sablon was applied, then I find a bit problematic that HadCM3 is only run for 250 years before completing the simulations with HadSM3 (Sablon et al. 2025). How can the response of the ocean circulation to different orbital parameters and CO_2 be adequately captured? I agree that 250 years is probably sufficient for near surface equilibrium but what about the deeper ocean? This can probably be alleviated, at least partly, if the simulations start from previous fully equilibrated HadCM3 simulations but this information is nowhere to be found, and even in this case, the longer-term equilibrium of the HadCM3 simulations should ideally be checked. If the simulations are initialized with idealized conditions, I have a hard time buying that surface ocean is in near-equilibrium without additional diagnostics.

The Reviewer's concern regarding the relatively short 250-year HadCM3L simulations is entirely valid, given that the deep ocean typically requires several millennia to reach thermal equilibrium. The 250-year spin-up duration used in Sablon et al. (2025), similarly to Araya-Melo (2015), was adopted to ensure that the upper ocean reached near-equilibrium. Considering that the focus of both Sablon et al. (2025) and the present study lies in the continental climate response rather than in the deep ocean, this trade-off was considered acceptable.

Incidentally, under transient astronomical forcing, achieving a fully equilibrated state would in itself be idealized and physically inconsistent with the evolving boundary conditions.

Araya-Melo, P. A., Crucifix, M., and Bounceur, N.: Global sensitivity analysis of the Indian monsoon during the Pleistocene, Clim. Past, 11, 45–61, https://doi.org/10.5194/cp-11-45-2015, 2015.

Btw, it should be mentioned somewhere in the text that HadSM3 is the slab-ocean version of the HadCM3 GCM and therefore not a GCM as generally understood.

We thank the Reviewer for pointing out this oversight. This clarification has now been included in the revised manuscript, where HadSM3 is explicitly described as the slab-ocean version of the HadCM3 GCM (see updated description of the emulator just above).

2. As it stands, Figure 2 is rather useless, unless I missed something.

We agree that Figure 2 does not provide essential information in its current form. We therefore replaced it with the Late Devonian topography presented in Sablon et al. (2025).

3. Figure 3.

In Gérard et al. (2025) in Clim. Past, the global mean O_2 concentration at 370 Ma with preindustrial pCO_2 and pO_2 and mean ocean PO_4 is about 210 mmol/m3 (Fig. 3 of Gérard et al. 2025). In this manuscript, at $\sim$ 2x preindustrial pCO_2 (550 ppmv), 0.8x preindustrial pO_2 and preindustrial mean ocean PO_4, the global mean O_2 concentration is 55 mmol/m3 (Fig. 3). Is it simply the effect of decreased atmospheric O_2 concentration? I find this surprising.

The lower mean $[O_2]$ obtained in this study compared to Gérard et al. (2025) indeed partly reflects the lower atmospheric pO_2 level used here. However, several additional factors also contribute. First, the mean oceanic $[PO_4]$ in the present simulations is about 25 % higher (2.656 μmol.kg^{-1} vs. modern 2.159 μmol.kg^{-1}; Gérard et al., 2025, Fig. 3), which enhances biological productivity and, consequently, O_2 consumption. Second, the higher pCO_2 tends to reduce global mean $[O_2]$, a relationship also reported in Gérard et al. (2025) using the same paleogeography.

4. Looking at the bathymetric differences between the 370 Ma configuration of Scotese and Verard in Gérard et al. (2025) makes me wonder if a similar sensitivity of O_2 variability would be found with the Verard paleogeography? There are indeed a lot fewer shallow seas in Verard. It could be nice to add some lines about this in the discussion.

The Scotese and Vérard continental reconstructions indeed differ substantially, resulting in distinct ocean circulation patterns and climatic responses (as shown in Gérard et al., 2025). Consequently, it remains uncertain how directly the results of the present study would translate if the same experiments were performed using the Vérard paleogeography, as stated in our original submission on lines 369-371: "Our results could depend on the selected paleogeographic reconstruction, given the substantial uncertainties that persist in deep-time topography and, by extension, in slope, erosion, and weathering patterns".

A robust assessment would require targeted experiments, which unfortunately lie beyond the scope of the present contribution, since they would require a new emulator, hence a brand-new ensemble of HadCM3 simulations. Nevertheless, exploring alternative continental configurations such as Vérard's would therefore represent an interesting avenue for future work. To acknowledge this, we propose adding a short note in the discussion highlighting that testing the sensitivity of our results to alternative paleogeographies would be a valuable step for future studies.

Lines 371-373: "... and weathering patterns. It would therefore be valuable for future studies to assess how alternative reconstructions, such as those from Vérard (2019) or Cermeño et al. (2022), might influence these processes and the resulting oceanic response."

Cermeño, P., García-Comas, C., Pohl, A. et al. Post-extinction recovery of the Phanerozoic oceans and biodiversity hotspots. Nature 607, 507–511 (2022). https://doi.org/10.1038/s41586-022-04932-6

5. l. 237.

I don't get what is remarkable in that the NP region shows no anoxia. It is instead rather expected if this a deep-water formation area (l. 223). It is more remarkable to me that region SA exhibits anoxia if it is indeed a deep-water formation zone, though I guess that the convection area must be quite localized, at the highest latitudes.

Yes, the absence of anoxia in the NP region is indeed expected given its role as a deep-water formation area. The term “remarkable” was therefore inappropriate and will be removed in the revised manuscript to more accurately reflect this point.

Lines 249-253: "Region NP exhibits no anoxic conditions over the period, because of its consistently high [O_2] and the limited magnitude of [O_2] variations, which are insufficient for any grid cell in the region to cross the anoxic threshold. These high [O_2] values over region NP are expected given the presence of major deep water formation in this region, whereas the relatively weaker deep water formation in region SA explains why this region can exhibit anoxia (see Fig. A1e and Fig. A2)."

6. Figure 4 and l. 249 and on.

To me, the very weak astronomical effect on O_2 and anoxia in regions NP, Si and WL is a bit overemphasized. At any rate, the purported insolation-driven changes invoked to explain the small variations in regions NP and Si should be shown in a figure, which could replace Fig. 2 for instance.

Regarding the first part of the Reviewer's comment on the "overemphasis" of astronomical forcing in regions NP, Si, and WL, we believe the text does not overstate the effect. At this stage (line 249 and on), the paragraph only aims to describe the mechanisms driving the simulated regional [O_2] variations and to link them to the only varying boundary condition in the experiments: astronomical forcing. Moreover, the manuscript already notes (lines 246-247 and 314-315) that some regions, such as Si, are largely unaffected by astronomical forcing, making it clear that not all regional [O_2] variations are substantial.

That being said, we agree with the Reviewer that some parts of the current paragraph could be clarified. Specifically, in region Si, [O_2] variability is not primarily controlled by regional biological productivity ($R^2 = 0.22$ versus > 0.74 in all other regions), indicating a dominant role of large-scale physical transport and mixing. This is further supported by the strong correlation between [O_2] in region Si and the northern hemisphere mean ($R^2 = 0.79$), highlighting that circulation rather than regional biological productivity drives the variability. In region NP, the insolation-driven [O_2] variations are supported by a robust correlation ($R^2 = 0.66$) between the insolation component of biological productivity and [O_2], while the nutrient-driven component shows no significant relationship ($R^2 = 0.01$, p-value = 0.63). This is also clarified in the methods that primary productivity in our setup is influenced by both the [PO_4] and the insolation. Unless otherwise noted, all correlations are significant with p-value $< 10^{-3}$. We propose to revise the paragraph to make these points clearer, explicitly distinguishing the mechanisms controlling [O_2] variability in regions Si and NP together with the correct related mechanisms.

Lines 93-94: "... control on biological productivity (Ridgwell et al., 2007). In this configuration, biological productivity is driven by both [PO_4] availability and local insolation."

Lines 259-271: "By definition, variations in anoxic volume arise from changes in regional [O_2]. Typically, these [O_2] fluctuations are closely tied to changes in biological productivity and subsequent remineralization. In regions SA, WL, and LG, [O_2] variability is almost entirely driven by variations in the [PO_4] component of biological productivity (regional $R^2 \geq 0.75$ and p-values $< 10^{-3}$). In region NP, [O_2] variability aligns closely with the insolation component of biological productivity ($R^2 = 0.66$ and p-value $< 10^{-3}$), whereas the [PO_4] component shows no meaningful relationship ($R^2 = 0.01$, p-value = 0.63). This indicates that insolation-forced biological productivity, rather than [PO_4], controls [O_2] dynamics in this region. Region Si exhibits a distinct behaviour compared with all others. It is the only region where [O_2] variability is not primarily driven by biological productivity ($R^2 = 0.22$ and p-value $< 10^{-3}$), indicating the dominance of another mechanism. The strong correspondence between [O_2] in region Si and the northern hemisphere mean ($R^2 = 0.79$, p-value $< 10^{-3}$) suggests that large-scale physical transport and mixing, consistent with strong currents and circulation in this region (see Fig.

7), drive these regional [O_2] fluctuations. Overall, our results indicate that when regional [O_2] variability is high, it is primarily driven by the [PO_4] component of biological productivity. However, when [PO_4] fluctuations are limited, other processes, such as insolation-forced changes in biological productivity or large-scale advection and mixing, become more critical."

We also revised the last paragraph of the section to improve the general flow and consistency of the text.

Lines 286-292: "In region Si, the relationship between [PO_4] and PO_4 weathering differs fundamentally from that of all other regions. Unlike SA, WL, and LG, where regional PO_4 weathering clearly imprints on [PO_4], no meaningful correspondence emerges in region Si. The regional PO_4 weathering flux is comparatively weak (about half that of region LG), and strong currents promote rapid advection and mixing, preventing local accumulation of weathered PO_4 (see Fig. 7). As a result, regional [PO_4] variations do not reflect local weathering input. Instead, [PO_4] in region Si closely follows the broader Northern Hemisphere signal ($R^2 = 0.98$ and p-value $< 10^{-3}$), indicating that large-scale transport governs its temporal evolution. This behaviour is consistent with the mechanism identified earlier: circulation and mixing dominate both [O_2] and [PO_4] variability in region Si."

Finally, we propose to add two small clarifications in the discussion section:

Lines 322-323: "... obliquity-driven changes, particularly through their influence on the insolation component of biological productivity."

Lines 325-327: "However, physical processes such as ocean circulation and tracer advection can override this control, as exemplified in region Si, where local [O_2] and [PO_4] dynamics are decoupled from regional weathering due to strong transport."

7. l. 256-268.
What is the reason for the different lag values in regions SA, WL, LG?

The lag arises because when PO_4 weathering begins to decline, while still relatively high compared to its mean value, oceanic [PO_4] continues to increase. In other words, the PO_4 input to a given region temporarily exceeds the rate at which it is consumed, leading to a further rise in average [PO_4]. Since both the magnitude of the PO_4 weathering peak and the rate of PO_4 consumption through biological productivity vary across regions, different lag values are therefore expected among them. We propose to add this precision to the revised manuscript to improve clarity.

Lines 277-279: "... [PO_4] continues to rise as long as the PO_4 input to a given region temporarily exceeds the rate at which it is consumed, producing a lag of several thousand years between the weathering signal and the [PO_4] response (see Fig. 6)."

Lines 281-282: "... precession cycle). Because both the magnitude of the PO_4 weathering peak and the rate at which PO_4 is consumed through biological productivity differ across regions, the resulting lag is region-specific."

8. l. 269-281.
What explains the [PO_4] lead in the time series then?

In the manuscript, we showed that [PO_4] in region Si correlates strongly with northern hemisphere [PO_4], whereas no clear relationship exists with local PO_4 weathering, indicating that [PO_4] in region Si is mainly driven by external weathering inputs (out of region Si). Because these external sources can exhibit different phase relationships with the astronomical forcing, no consistent lead–lag relationship between [PO_4] and PO_4 weathering is expected in region Si.

We believe that the changes made following previous comments 6 and 7 improve the clarity of this point as well, hence requiring no further changes.

It is written that the regional [PO_4] signature is strongly influenced by NH weathering outside of the Si region, but there are not a lot more land masses in the NH.

We thank the Reviewer for noting this small inconsistency in our wording. As mentioned above, there is a clear relationship between [PO_4] in region Si and that of the northern hemisphere. However, this does not necessarily imply that the evolution of [PO_4] in region Si is predominantly driven by weathering fluxes originating from the northern hemisphere itself. Rather, the temporal evolution of the northern hemisphere [PO_4] reflects a complex interplay between advection, driven by ocean circulation, and diffusion processes, making it difficult to identify the precise weathering original locations.

Again, modifications following comments 6 and 7 have also improved the clarity of the text regarding this point.

Also, region Si is located right poleward of significant deep upwellings (see the MOC on Fig. 8 of Gérard et al. 2025). Perhaps worth checking whether these upwellings may mask continental PO_4 inputs from weathering.

While it can be stated that large-scale ocean circulation, including both surface and deep currents, clearly contributes to the mixing of [PO_4] across the northern hemisphere, and hence influences [PO_4] in region Si, it is not possible within our current experimental design to isolate the exact contribution of local upwellings in masking continental PO_4 inputs.

Nevertheless, this clarification aligns with one of the general messages of the paper (lines 325-327): "However, physical processes such as ocean circulation and tracer advection can override this control, as exemplified in region Si, where local [O_2] and [PO_4] dynamics are decoupled from regional weathering due to strong transport."

9. There is no map of the prescribed topography, even though I suppose that steep slopes strongly affect the simulated weathering fluxes. This could be easily added to Fig. 1.

I would appreciate some comments on this.

Yes, topography plays a major role in the simulation of weathering fluxes, by modulating erosion. The reason why Fig. 1 does not display topography is that, in cGENIE, the continental surface is represented as flat. Only a runoff map is prescribed, specifying the direction of continental runoff.

However, the PO_4 weathering fluxes simulated here do effectively account for topographic slopes. This is because they are computed by the weathering emulator, which is built upon the Scotese paleogeographic reconstruction and therefore incorporates its full topographic information. Consequently, the dependence of weathering fluxes on surface slopes is resolved within the coupled cGENIE–emulator framework. The topography used in the emulator can be found in Sablon et al. (2025, Fig. 1), and has now replaced the old Fig. 2 in the manuscript (see Fig. R3).

This precision was also explicitly added to the new description of the emulator in the methods section.

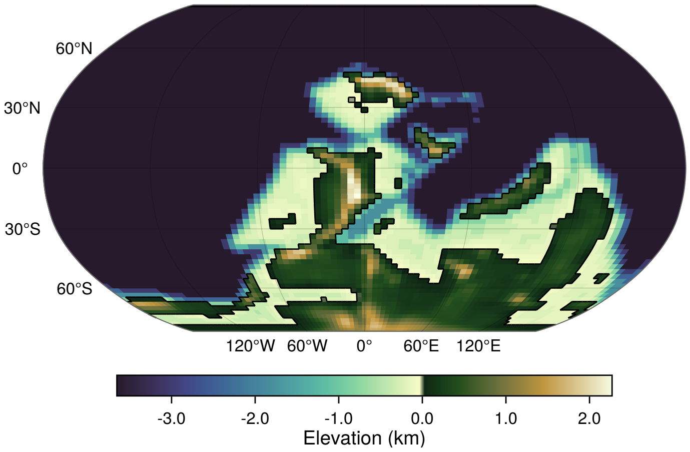

Figure R3: **Topography of the Late Devonian (370 Ma)**. Topography according to Scotese and Wright (2018), reproduced from Sablon et al. (2025). The bathymetry is identical to the one shown in Fig. 1b.

10. l. 374.

I am not convinced that the results presented in this manuscript support this. Obliquity-driven changes of a few mmol/m3 around of mean of more than 100 mmol/m3 in regions NP and Si would probably barely affect the sedimentary record.

Our goal in highlighting the obliquity-driven productivity response is not to claim strong sedimentary expression, but to motivate further investigation of the insolation component of biological productivity. We propose to rephrase this sentence to clarify our point.

Lines 400-403: "While eccentricity and precession primarily affect [O_2] through their control on PO_4 weathering fluxes, obliquity exerts its influence by modulating the insolation component of biological productivity at high latitudes, highlighting that even in a greenhouse world, obliquity can leave an imprint on ocean biogeochemistry, hence possibly sedimentary records."

Furthermore, we also understand that the corresponding sentence in the abstract could also benefit from a more nuanced phrasing:

Lines 8-10: "Additionally, global analysis reveals that obliquity variations can leave an imprint on global ocean oxygen levels via their influence on biological productivity, suggesting a potential pathway for obliquity-driven anoxia under greenhouse conditions."

Second Reviewer: Alexander Farnsworth

Gérard et al. investigate how orbital forcing may have modulated ocean oxygenation during the Late Devonian by altering continental phosphorus (PO_4) weathering fluxes and, in turn, marine productivity and redox state. Specifically, trying to assess the impact of a 2.4 Myr eccentricity cycle is implicated to drive the pacing of anoxic events using cGENIE, an Earth system intermediate complexity model in conjunction with HadCM3Bl temperature and hydrology, GEOCLIM7 and an emulator approach to produce more accurate fluxes containing orbital variation. This allows spatially explicit phosphorus (PO4) values instead of a single global value, theoretically allowing greater accuracy (both globally and regionally).

Transient and sensitivity simulations show that eccentricity drives a 40 ppm change in CO2 showing that PO4 is anti-correlated leading to smaller changes in dissolved oxygen on a global level, but much larger regional changes.

This work suggests that orbital forcing alone is not responsible for the Devonian anoxic events (if we believe they were truly global events), but could lead to anoxic states in some regions if the background state was nearer its tipping point.

There is good novelty throughout this manuscript, it is well written and generally well argued. I applaud the authors limitations section, and while there are limitations, it does show what the current state of the art is and ways forward for future modelling of deep-time orbital biogeochemical studies. More can be done (see below) but this open discussion on the matter in the ms is very welcome.

We thank the Reviewer for the constructive and encouraging feedback. We are pleased that the manuscript's novelty, clarity, and discussion of model limitations were appreciated.

Major comments

Given that the hydrological signal determines a large part of the amount of weathering occuing, can anything be said as to how well the model (HadCM3BL) is representing the hydrological cycle? I suspect there isn't much proxy-data to do this. Likewise, can we constrain the rate of weathering from ROKGEM somehow? Presumably a higher or lower rate of weathering could substantially change these results? Or if another weathering model were used? Some further discussion about the implications of this (as well as the type of rock type being eroded – can we constrain this too?) might be warranted.

Regarding the representation of the hydrological cycle, we acknowledge that assessing the accuracy of the HadSM3 hydrology is challenging given the limited availability of proxy data for direct comparison. Nevertheless, because HadSM3 explicitly resolves key atmospheric processes and has been widely applied in palaeoclimate studies, its use represents a clear and substantial improvement over the simplified hydrological formulation implemented in cGENIE.

Constraining weathering rates in ROKGEM is equally challenging. It is therefore possible that different weathering rates, or the use of an alternative weathering scheme, could affect the quantitative outcomes. However, assessing the magnitude or direction of such differences would require running dedicated simulations. In the present framework, absolute weathering fluxes are prescribed in ROKGEM to close the carbon cycle, based on burial fluxes diagnosed in SPIN1. The resulting silicate weathering fluxes are within 25 % of independent Devonian estimates (Maffre et al., 2022). Moreover, the temperature sensitivity of silicate weathering in ROKGEM follows a standard formulation, consistent with the behaviour documented by Kaufhold et al. (2025). Similarly, Late Devonian lithology is poorly constrained (as testified by the little information available for the Devonian in MacroStrat; Fig. R4), which is why we rely on present-day average lithological fractions (this information was added to the new description of the emulator, lines 77-88). Overall, we designed our modelling framework using the most suitable tools currently available for this problem, but we fully agree that comparing these results with other weathering models would be an important next step. We propose to add the following sentence to the revised manuscript to acknowledge this:

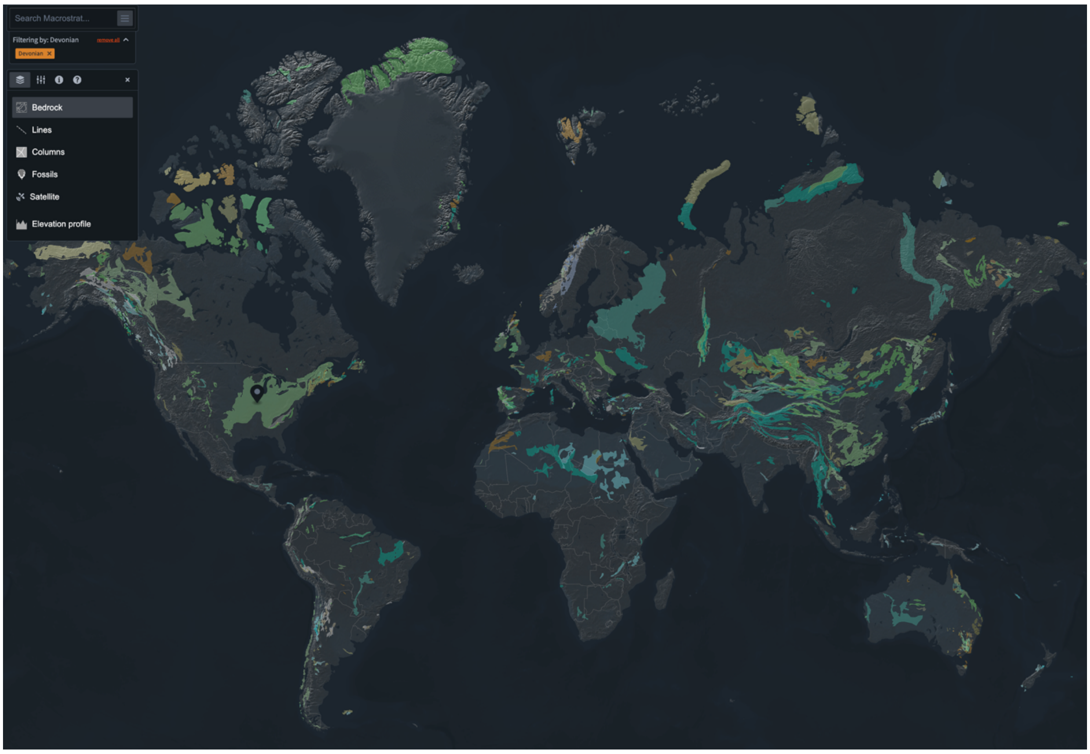

Figure R4: Devonian extraction of the MacroStrat Database (https://macrostrat.org/map/loc/-89.77/39.62#intervals=94&x=47.02&y=44.34&z=2&show=bedrock&hide=labels).

Lines 375-377: "...astronomical forcing. Similarly, the use of HadCM3L, HadSM3 and GEOCLIM necessarily reflects one possible representation of climate and weathering dynamics; alternative emulator models or hydrological schemes could yield quantitatively different outcomes, although assessing the magnitude of such differences would require dedicated sensitivity experiments."

Maffre, P., Godderis, Y., Pohl, A., Donnadieu, Y., Carretier, S., and Le Hir, G. (2022). The complex response of continental silicate rock weathering to the colonization of the continents by vascular plants in the Devonian. American journal of science, 322(3), 461-492.

Kaufhold, C., Willeit, M., Liu, B., and Ganopolski, A. (2025). Assessing the lifetime of anthropogenic CO_2 and its sensitivity to different carbon cycle processes. Biogeosciences, 22(12), 2767-2801.

Likewise, can we constrain at all the predicted background state of the Late Devonian O2? Whether the model becomes anoxic/how sensitivity it might be to any change in CO2/PO4 will have an impact on whether it becomes anoxic. I.e. if the background state was already near the threshold, then even small changes might be important.

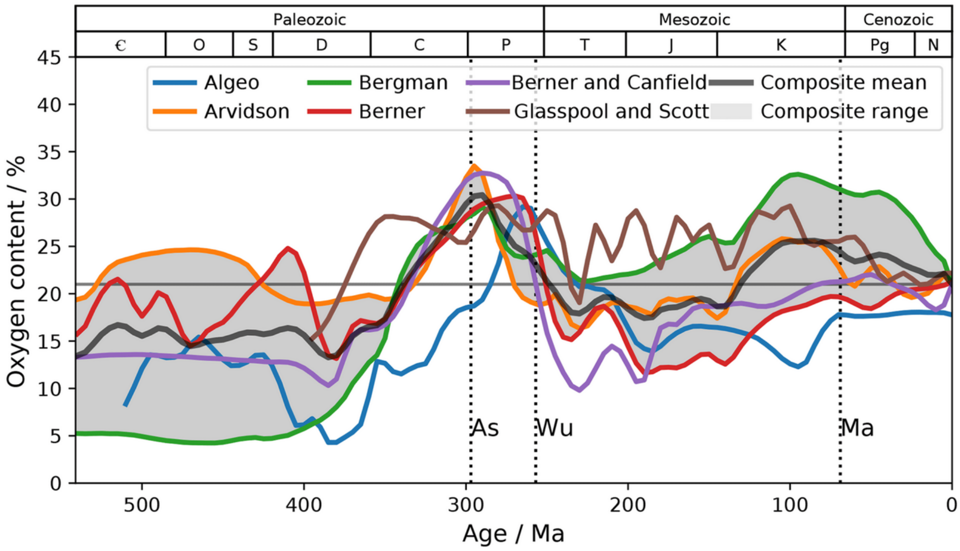

Figure R5: Compilation of Phanerozoic atmospheric oxygen concentration reconstructions (Wade et al., 2019).

Constraining the background oxygenation state of the Late Devonian is challenging (see Fig. R5). In this study, we relied on the most up-to-date estimates of key boundary conditions, including atmospheric pCO_2 (Chen et al., 2021; Dahl et al., 2022), mean PO_4 supply (Sharoni and Halevy, 2023), and atmospheric O_2 (Krause et al., 2018; Mills et al., 2023). Under these constraints, cGENIE simulates a background ocean state that is already strongly deoxygenated. We agree with the Reviewer that such a background state inherently increases the system's sensitivity to perturbations. When the ocean is already near threshold, even modest changes can push regions into anoxia. We see this as consistent with the broader hypothesis that the Late Devonian ocean was predisposed to anoxia, and our results provide modelling support for this interpretation. We propose to add the following sentence to the discussion:

Lines 296-298: "... Devonian ocean oxygenation. Our simulations indicate that the Late Devonian background state was already characterised by low oceanic $[O_2]$, which likely contributed to the increased frequency of anoxic events observed throughout this period (Becker et al., 2020)."

Finally, we acknowledge that the background state and its sensitivity to forcing reflect the behaviour of a single model configuration. Although we refined several aspects of our setup (PO_4 weathering, wind stress, albedo), cGENIE inevitably suffers from internal limitations. We therefore view this work as a first, internally consistent estimate of Late Devonian oxygenation. We propose to explicitly highlight this in the limitations section:

Lines 391-392: "... of climate-circulation–biogeochemistry interactions. They could also explicitly test the robustness of the Late Devonian background state and its sensitivity to evolving external forcing across multiple models."

Wade, D. C., Abraham, N. L., Farnsworth, A., Valdes, P. J., Bragg, F., and Archibald, A. T.: Simulating the climate response to atmospheric oxygen variability in the Phanerozoic: a focus on the Holocene, Cretaceous and Permian, Clim. Past, 15, 1463–1483, https://doi.org/10.5194/cp-15-1463-2019, 2019.

Was an ice sheet prescribed from the paleogeography (Famennian) in cGENIE? It is suggested that the Late Devonian may have had one present. Scotese and Wright do not explicitly have one, however one was artificially added into the Valdes et al. simulations (change in orography and albedo). Is there a potential disconnect here? If so, what impact could this have on fluxes and ocean circulation in cGENIE and in turn anoxia on a global and regional level?

All boundary-condition files used in cGENIE for albedo, orography, and wind stress were taken directly from the HadCM3L output of Valdes et al. (2021). This ensures smooth continuity between the two models regarding prescribed orography and albedo profiles.

cGENIE does not incorporate dynamic ice sheets or sea-level changes. We therefore acknowledge that this model is not the optimal tool for assessing ice-sheet feedbacks or their climatic consequences. This limitation is already noted in the manuscript on lines 382–384: "The exclusion of ice sheets and sea-level variations removes critical mechanisms through which obliquity can strongly impact [O_2], although the accurate representation of sea-level changes would remain challenging at the cGENIE grid resolution."

How well does cGENIE represent modern climate and ocean biogeochemistry? How sensitive is the model to changes in nutrient fluxes and/or CO2. I may have missed it, but some references or discussion to justify that cGENIE and GEOCLIM adequately produces

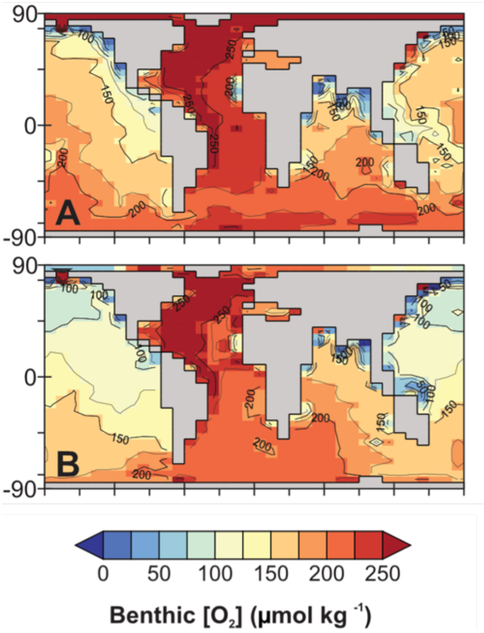

Figure R6: World Ocean Atlas (top panel) vs. GENIE-derived (bottom panel) benthic [O_2].

We are confident in cGENIE's ability to reasonably simulate paleoclimate conditions, as the model performs well in reproducing key aspects of modern climate and ocean biogeochemistry. In terms of the global ocean circulation, cGENIE realistically captures large-scale variations in the Atlantic meridional overturning circulation (AMOC), as recently shown in Gérard and Crucifix (2024), where simulated AMOC changes closely follow those produced by models of the Coupled Model Intercomparison Project phase 5 generation. cGENIE also satisfactorily represents modern dissolved oxygen concentrations and spatial patterns (Ridgwell et al., 2007), supporting its suitability for studies of ocean redox dynamics. As an illustration, Fig. R6 (modified after Pohl et al., 2022) displays the benthic oxygen concentrations after the World Ocean Atlas 2018 (re-gridded to the model resolution; top panel) and after the standard modern cGENIE simulation (bottom panel). Benthic oxygen concentrations are underestimated in the

Southern Ocean and North Pacific as a result of too weak an Antarctic circumpolar current, which results from the difficulty in accurately representing the Drake Passage at the model resolution. Except for this bias, cGENIE captures well first-order spatial patterns and values. This ability of cGENIE to capture essential climate and biogeochemical processes was already highlighted in the manuscript on lines 351–356. In response to the Reviewer's point, we propose to further clarify the text as follows:

Lines 356-357: "... Stappard et al., 2025). These studies collectively demonstrate the model's reliability in simulating paleoclimate conditions and associated biogeochemical responses."

cGENIE sensitivity to changes in PO_4 supply and atmospheric pCO_2 has been extensively assessed in Gérard et al. (2025). There, we show, for instance, that doubling pCO_2 produces a radiative forcing of approximately 4 $W\,m^{-2}$, resulting in a global mean surface warming of about 3 °C in our simulations. We incorporate the climate sensitivity of cGENIE into the manuscript

Lines 113 - 115: "The estimated climate sensitivity of cGENIE is approximately 3°C per doubling of pCO_2, which is close to the mean value from the Coupled Model Intercomparison Project (CMIP) 5 and slightly lower than that from CMIP6 Zelinka et al. (2020)."

The same reasoning applies to material related to GEOCLIM. Although the present study builds directly on Sablon et al. (2025), we chose to avoid repeating content that is already extensively covered there, including the description, performance, and validation of GEOCLIM. All relevant information can therefore be found in Maffre et al. (2025) and Sablon et al. (2025), now accepted. This preserves a clear distinction between the two studies and limits redundancy.

How appropriate is the emulator at representing Deep time periods?

- Was the emulator trained on Quaternary time periods?

  This type of climate emulator has indeed been applied to Quaternary studies using HadSM3, for example, in Araya-Melo et al. (2015), Sun et al. (2019), and Lord et al. (2017). However, the emulator used in the present work was trained on a Devonian-specific ensemble of experiments. An emulator is fundamentally a statistical model calibrated on a given training dataset; consequently, the one developed here cannot be applied to Quaternary conditions.

  Importantly, the emulator employed in this study successfully passes standard validation tests, demonstrating that it is an appropriate surrogate for the corresponding Hadley Centre model under Devonian boundary conditions. Compared with the earlier Quaternary-focused implementations, the emulator framework includes several methodological improvements (Sablon et al., 2025). These enhancements have not yet been implemented in a Quaternary context.

- Is the La11a-based orbital solution (123.3–122.2 Ma in the original chronology) applicable to your time frame? Granted the Devonian-specific solution (e.g. Zeebe & Lantink 2024) came out while your study was already running, but do you think your results would be different if it were used?

  Zeebe & Lantink (2024) do not provide a single orbital solution for the Devonian but rather a suite of solutions, reflecting the intrinsic chaotic nature of Solar System dynamics over such long timescales. This chaotic behaviour makes it impossible to reconstruct a unique, precise evolution of eccentricity or inclination beyond a certain temporal horizon.

That said, the astronomical solution employed here shares the same frequency content and amplitude-modulation structure as those found across the Zeebe–Lantink ensemble. In particular, the solution we adopted exhibits a resonance between inclination (s_3 - s_4) and eccentricity (g_3 - g_4), producing an obliquity amplitude-modulation pattern that aligns with the 2.4-Myr eccentricity nodes. This behaviour is common across many plausible Devonian solutions and is also present in a majority of the Zeebe–Lantink solutions. Nonetheless, we acknowledge that alternative resonance regimes are possible for the Devonian.

We added the following sentence to the manuscript to improve clarity. Lines 127-129: "...that publication appeared. Nevertheless, the astronomical solution employed here shares the same frequency content and amplitude-modulation structure as those found across the Zeebe–Lantink ensemble. We adopted the same approach..."

Regarding the potential impact of using a different orbital solution, we mentioned in the manuscript that cGENIE responds relatively directly to the astronomical forcing imposed (Lines 336-337). Therefore, we believe that the core results would remain largely the same. The most pronounced deoxygenation would still occur near eccentricity maxima, and the system would still exhibit strong spatial heterogeneity in its response. However, explicitly running simulations using different astronomical solutions would be needed to further demonstrate this point.

Araya-Melo, P. A., Crucifix, M., and Bounceur, N.: Global sensitivity analysis of the Indian monsoon during the Pleistocene, Clim. Past, 11, 45–61, https://doi.org/10.5194/cp-11-45-2015, 2015.

Sun, Y., Yin, Q., Crucifix, M. et al. Diverse manifestations of the mid-Pleistocene climate transition. Nat Commun 10, 352 (2019). https://doi.org/10.1038/s41467-018-08257-9

Lord, N. S., Crucifix, M., Lunt, D. J., Thorne, M. C., Bounceur, N., Dowsett, H., O'Brien, C. L., and Ridgwell, A.: Emulation of long-term changes in global climate: application to the late Pliocene and future, Clim. Past, 13, 1539–1571, https://doi.org/10.5194/cp-13-1539-2017, 2017.

Further questions relating to some of the studies limitations:

- Flat bottom bathymetry in the palaeogeography – what role could this play on ocean anoxia? Or indirectly through unrealistic ocean circulation?

  We fully agree with Reviewer 2 that uncertainties in Devonian paleogeography are substantial. This was already highlighted in our original submission (lines 349-351), and we expanded on this point in response to Reviewer #1 (Line 351: "... and weathering patterns. It would therefore be valuable for future studies to assess how alternative reconstructions, such as those from Vérard (2019) or Cermeño et al. (2022), might influence these processes and the resulting oceanic response").

  Regarding the deep-ocean bathymetry specifically, we anticipate the latter to largely impact benthic oxygen concentrations in the model (based on the sensitivity test to deep-ocean bathymetry conducted in Pohl et al. (2022) over the last 140 Myr, see their extended data fig. 6). It is also likely that resulting changes in the global ocean circulation may induce changes in the intensity of the return of the nutrients from the deep ocean to the upper ocean, hence modulating export production intensity at the global scale. Beyond these potential changes in ‘background’ oceanic conditions, we do not expect the deep-ocean bathymetry to largely impact our main results regarding the modulation of ocean anoxia by astronomical forcing. This is mainly because our study focuses on the upper ocean (since this is where deep-time sedimentary records are preserved).

- How appropriate is cGENIE in resolving the large scale ocean circulation, especially at orbital timescales? E.g. AMOC collapse on LGM timescales?

  In Gérard and Crucifix (2024), the cGENIE large-scale ocean circulation was validated against CMIP5 models (this is now explicitly mentioned in the revised manuscript). In the same study, transient hysteresis experiments of the AMOC were conducted over timescales of 20 000 to 40 000 years. Furthermore, another hysteresis of the large-scale ocean circulation was performed with cGENIE in Gérard et al. (2025) over a period of 150 000 years. Altogether, this provides confidence in cGENIE’s ability to represent large-scale ocean circulation at orbital timescales.

- Have any specific assumptions on nutrient and other co-varying fluxes been made?

  All assumptions relevant to how nutrients and associated fluxes are treated in our setup are explicitly stated in the manuscript. The key ones include: (1) the PO_4 weathering and delivery in an open system handled by the emulator of Sablon et al. (2025); (2) the single-nutrient (PO_4) control on marine productivity, which is standard in palaeo-biogeochemical modelling (Ridgwell et al., 2007); and (3) the sediment module options employed here, notably organic carbon burial and P recycling, following Hülse and Ridgwell (2025, now published).

- How much can we trust available proxy-date? Good enough and well constrained age model? Are there different interpretations for these Devonian deposits? Can we truly determine whether this was a global event from them alone?

  Geological proxy data provide crucial constraints on the duration, recurrence, and broad geographic extent of Devonian anoxic events, but they remain limited by sparse spatial coverage, dating uncertainties, and interpretative ambiguity. As a consequence, different stratigraphic records sometimes lead to contrasting reconstructions, and it is challenging to determine from data alone whether any given anoxic episode was truly global, synchronous, or uniform across the open ocean. This is precisely where climate–biogeochemical modelling becomes valuable: models allow us to test different scenarios suggested by the proxy record. Our results, for example, offer a mechanism explaining why sedimentary archives from distinct regions may record different expressions of an astronomical signal, despite being part of the same large-scale climatic context.

If HadCM3 circulation was used to initialised this work, how much memory of it persists through the orbital cycle? I suspect (not much once the orbital starts to change significantly) not much, which is good because you want ocean circulation to respond. However, does this then rely on cGENIE ocean circulation being appropriate? And if so, is it?

We agree that, even if the simulations were initialised using a circulation field derived from HadCM3L, the memory of this initial state would decay rapidly. Once the orbital forcing begins to vary, cGENIE’s own ocean circulation dynamics takes over, and the influence of any externally prescribed initial circulation would be lost well before completing a precession cycle. But is the ocean circulation of cGENIE appropriate to handle such a task? While cGENIE is not a full-complexity GCM, we have confidence in its ability to represent large-scale circulation and its variability. We elaborated on this point in response to a previous comment ("How well does cGENIE represent modern climate and ocean biogeochemistry?") and hope that the revised manuscript now adequately addresses the Reviewer's concern.

Minor comments

The introduction was nicely written highlighting why orbital variability is worthy to look at. However, I wonder if a sentence of two explaining how orbital forcing imprints itself onto changes in ocean biogeochemistry, e.g. through changes to ocean circulation, hydrological changes leading to change in nutrient input, etc. . .

Thank you for the suggestion. We agree that the text could benefit from several clarifications on how astronomical forcing leaves an imprint on ocean anoxia. We propose to revise the following sentence, which now also explicitly presents the working hypothesis of our study:

Lines 26-29: "There is also growing evidence that astronomical forcing could modulate at least some of these events, primarily through a "top-down" mechanism (Carmichael et al., 2019) whereby cyclic changes in temperature and precipitation alter continental nutrient fluxes to the ocean (De Vleeschouwer et al., 2014; De Vleeschouwer et al., 2017; Wichern et al., 2024; Huygh et al., 2025)."

It might be worth a small paragraph discussing the proxy-evidence for this time period? How reliable was it? Are they globally distributed? Etc. . .

We understand the Reviewer's suggestion to include a brief discussion of the proxy evidence. However, a challenge arises in that our study does not focus on any specific Late Devonian anoxic event. The reliability, spatial distribution, and interpretative uncertainties of proxy records vary substantially from one event to another. Providing a concise yet accurate overview for all of them would therefore require a level of detail that would shift the introduction away from its primary purpose.

Our goal within the introduction is simply to highlight that astronomical signals have been reported across several Late Devonian anoxic intervals, and to motivate a more general exploration of the link between astronomical forcing and anoxia using numerical modelling. Hence, we believe the introduction is clearer and more coherent in its current form, without additional event-specific proxy details.

When the PO4 flux was routed to coastal tiles, was the total amount for each watershed basin averaged between each coastal grid cell? Or where higher concentrations routed through large river systems (river exit node)? I suspect it's the later unless paleo-rivers were determined?

In our implementation, weathering from continental grid cells is not averaged over watershed basins nor routed through explicit river networks. Instead, we use a simple nearest-neighbour routing scheme: each continental cell is mapped to the geographically closest ocean-coastal cell (based on haversine distance), and its entire weathering flux is delivered directly to that coastal cell. This procedure does not make assumptions about paleo-topography or paleo-river pathways, and thus does not distinguish large river outlets from small ones.

HadSM3? Do you mean HadCM3? The 'S' implied a slab ocean model was used. However, I suspect the outputs from Valdes, 2021 was used as initialised this work, then it's the full AOGCM version (HadCM3BL).

We apologise for the confusion. In Sablon et al. (2025), both the fully coupled model and the slab-ocean model were used to build the emulator. We have double-checked the manuscript, and the distinction between HadSM3 and HadCM3L is used correctly throughout. We have now clarified in the revised text that HadSM3 refers specifically to the slab-ocean version. We also agree that referring to the coupled model as HadCM3L would improve precision, and we updated the manuscript accordingly. Although HadCM3BL could arguably be used in specific locations instead of HadCM3L, we find that referring only to HadCM3L is preferable to avoid any misinterpretation. Finally, to maintain consistency with the use of a slab model, we propose removing the term "fully coupled" from line 43.

What value of CO2 did the HadCM3BL late Devonian use?

The pCO_2 is 926 ppm, which is now clearly stated in the revised manuscript. We apologise for the oversight.

Line 44 – Change GEOCLIM to GEOCLIM7

The emulator used in this study was trained on GEOCLIM6.1.0. We cite Maffre et al. (2025) because the updates introduced in that version (GEOCLIM7) only concern ocean biogeochemistry, which is not involved in the emulator training. Furthermore, the paper provides a complete description of the model components relevant to our work (including DynSoil). To avoid any misunderstanding, we propose to keep the text as it, with all the details available in Sablon et al. (2025).

Line 188 – "This inverse relationship occurs because intensified weathering", would this not be dependent on the type of rock that is being eroded?

Although GEOCLIM allows for an explicit spatial distribution of lithological fractions, this option was not used in our study; instead we applied present-day average lithology values. While the magnitude and shape of the weathering response do depend on lithology, the direction does not: in GEOCLIM, any increase in silicate weathering necessarily produces an increase in phosphorus weathering, because the two fluxes are explicitly coupled in the model formulation.

Line 194:195 – "This outcome arises from the physical processes encoded in the emulator developed by Sablon et al. (2025), where the same feature was previously identified." Presumably driven by the hydrology? Can you say whether it was simply an enhancement of the hydrological cycle globally? Or was it changes in the location of precipitation maxima (e.g. changes in storm tracks) moving poleward/equatorward over regions with more erodible material (i.e. mountains vs flat regions?) leading to a greater PO4 flux? Is the model complex enough to shed any light on this?

The atmospheric forcing of weathering and erosion is supplied indirectly through the HadSM3-based emulator, which provides spatially explicit temperature and runoff fields. With DynSoil enabled, GEOCLIM is indeed complex enough to represent spatial variations in the hydrological cycle, so both a global intensification of runoff and regional shifts in precipitation can influence local erosion and PO_4 export. In our simulations, the dominant driver is the precession-sensitive monsoon system captured by the emulator: precession alters monsoon intensity and distribution, which directly modulates the spatial pattern of runoff and thus weathering and PO_4 delivery. We propose to add this in the revised manuscript to improve clarity:

Lines 204-208: "... modulation can reach 10.5%. In our simulations, the dominant driver is the precession-sensitive monsoon system captured by the emulator: precession alters monsoon intensity and distribution, which directly modulates the spatial pattern of runoff and thus weathering and PO_4 delivery. Beyond this, eccentricity is positively correlated with PO_4 weathering, such that periods of high eccentricity are associated with enhanced oceanic $[PO_4]$, a feature already identified and documented in Sablon et al. (2025)."

Line 220 – How well does the model represent shallow marine settings?

In cGENIE, this ability is primarily constrained by model resolution. Vertically, the use of the 16-layer logarithmically spaced grid provides higher resolution near the surface, which improves the representation of shallow-water processes. However, the resolution remains relatively low, with the first layer being 60 m thick. Horizontally, the 36×36 (equal-area) grid also limits cGENIE's ability to resolve highly localized environments.

However, the shallow-marine representation in cGENIE is often more strongly influenced by the underlying palaeogeographic reconstruction than by numerical resolution. As shown in Gérard et al. (2025), different continental configurations can lead to substantial differences in shelf extent, water depth, and connectivity. In this context, the resolution used here remains adequate for capturing and representing biogeochemical conditions in shallow marine settings. This is well illustrated by previous work using cGENIE to help interpret proxy data produced in coastal paleoenvironments. For instance, He et al. (2024) demonstrated that cGENIE permits explaining the contrasting redox changes that occurred in shallow-water environments – such as documented on I/Ca measured in carbonates – during the late Cambrian Steptoean positive carbon isotope excursion (SPICE). Pohl et al. (2021) similarly explained shallow-water I/Ca data during the Late Ordovician, and Hülse et al. (2021) simulated the spread and spatial patterns of euxinia onto the continental shelves in line with proxy records during the Permian-Triassic transition.

He, R., Pohl, A., Prow, A., Jiang, G., Huan, C. C., Saltzman, M. R., and Lu, Z. (2024). The dynamic ocean redox evolution during the late Cambrian SPICE: Evidence from the I/Ca proxy. Global and Planetary Change, 233, 104354.

Line 229:237 – Why does the Si region have such a high O2 background state? Any idea why?

This results from strong transport and mixing linking the Si region to the broader Northern Hemisphere circulation. As a consequence, dissolved O_2 from the atmosphere is efficiently advected into the Si region, leading to a comparatively high background O_2 state. This information has been added to the following revised paragraph:

Lines 259-271: "By definition, variations in anoxic volume arise from changes in regional $[O_2]$. Typically, these $[O_2]$ fluctuations are closely tied to changes in biological productivity and subsequent remineralization. In regions SA, WL, and LG, $[O_2]$ variability is almost entirely driven by variations in the $[PO_4]$ component of biological productivity (regional $R^2 \geq 0.75$ and p-values $< 10^{-3}$). In region NP, $[O_2]$ variability aligns closely with the insolation component of biological productivity ($R^2 = 0.66$ and p-value $< 10^{-3}$), whereas the $[PO_4]$ component shows no meaningful relationship ($R^2 = 0.01$, p-value $= 0.63$). This indicates that insolation-forced biological productivity, rather than $[PO_4]$, controls $[O_2]$ dynamics in this region. Region Si exhibits a distinct behaviour compared with all others. It is the only region where $[O_2]$ variability is not primarily driven by biological productivity ($R^2 = 0.22$ and p-value $< 10^{-3}$), indicating the dominance of another mechanism. The strong correspondence between $[O_2]$ in region Si and the northern hemisphere mean ($R^2 = 0.79$, p-value $< 10^{-3}$) suggests that large-scale physical transport and mixing, consistent with strong currents and circulation in this region (see Fig. 7), drive these regional [O_2] fluctuations. Overall, our results indicate that when regional [O_2] variability is high, it is primarily driven by the [PO_4] component of biological productivity. However, when [PO_4] fluctuations are limited, other processes, such as insolation-forced changes in biological productivity or large-scale advection and mixing, become more critical."

Line 327:329 – Assuming that the model is adequately capturing the background state correctly?

Yes, regarding anoxia, this assumes that the model adequately captures the background state, since anoxia is defined relative to a fixed [O_2] threshold. A different background state could therefore produce substantially different anoxic patterns. However, this does not apply to absolute [O_2] variations, which would remain largely unaffected by shifts in the background state. In the manuscript, we chose to highlight anoxic volume variations because they are widely used and understood across both the geological and modelling communities. Importantly, this choice is not misleading, as we have enough confidence in the model's ability to reproduce the relevant background state. This is because the model performs well in the modern and past, and because we relied on the most up-to-date estimates of key boundary conditions.

Line 329:330 – "Our results also suggest that anoxic events could manifest as a succession of smaller, transient episodes". How does this tally with the earlier suggestion that there isn't much memory in the system?

We apologise for the confusion. What we mean here is that the multiple anoxic peaks closely follow the precession cycles. This behaviour is therefore not related to any memory in the system but simply reflects the temporal structure of the astronomical solution itself. This point matters because anoxic events are often thought of as single, large, continuous episodes (e.g. the nutrient gun hypothesis from De Vleeschouwer et al., 2017), whereas our simulations show that they could instead occur as several short-lived peaks paced by the astronomical forcing. This result aligns with the idea, proposed in previous studies such as Hedhli et al. (2023), that certain Devonian anoxic events may in fact consist of multiple successive pulses. We thank the Reviewer for identifying the lack of clarity of our original phrasing and propose to revise the sentence as follows:

Lines: 342-344: "Our results also suggest that anoxic events could manifest as a series of smaller, short-lived, precession-paced episodes (see Fig. 4c and supplementary video) rather than a single, prolonged global event, similar to the scenario proposed by Hedhli et al. (2023)."

Will the simulations be made available to the community?

Yes. In line with cGENIE's open-science policy, we will provide all the files required to fully reproduce the simulations. Rather than distributing the raw model outputs, we will supply the complete set of configuration and input files, ensuring full transparency of the experimental design and allowing users to easily rerun or adapt the experiments for their own purposes. We also plan to deposit the three main simulations on Zenodo to further enhance accessibility of the results to the community.

Editor decision: Yannick Donnadieu

Dear authors, Your paper has now been reviewed by two experts in the field of paleoclimate modeling. Both recognize the importance of your work as well as the quality of the text and illustrations. They categorize your contribution as requiring moderate revisions (one recommending minor and the other major revisions). Both reviewers seem to have encountered difficulties with the methods section and the role played by ocean dynamics in your results. Here is a tentative summary (based on my own reading of the paper and the reviews):

First, even if a method has already been described in other previous papers, your paper must be readable as it is for the Climate of the Past audience. Hence, I'd like to see clear improvements regarding: 1) Which simulations were used and which components they forced (specifically regarding HadCM3 vs. HadSM3). You should add a table summarizing boundary/initial conditions, simulation durations, and the specific model names.

The following table featuring all information on the differences about forcing was added to the revised manuscript.

Name	Type	Emulator	Duration (kyr)	IC	Forcings	BC
SPIN1	Closed	Off	20	—	pCO_2; pO_2; surface flux of DIC and ALK	Constant wind fields and albedo profile (HadCM3).
SPIN2	Open	Off	500	SPIN1	pO_2	as in SPIN1
SPIN3	Open	On	100	SPIN2	pO_2	as in SPIN1
Full	Open	On	1100	SPIN3	pO_2	as in SPIN1
Obli	Open	On	1100	SPIN3	pO_2	as in SPIN1
Ecc-prec	Open	On	1100	SPIN3	pO_2	as in SPIN1

Table 1: **Experimental design.** The type indicates whether the model operates in a closed or open configuration, with the closed setup meaning that every chemical species lost through sedimentation is compensated by an equivalent input from weathering. The emulator is the one from Sablon et al. (2025) that is built on model outputs from HadCM3, HadSM3, and GEOCLIM. Initial conditions (IC) for SPIN1 are derived from GEOCLIM simulations and include PO_4, alkalinity (ALK), calcium, and sulfate. Forcings correspond to quantities that are explicitly prescribed during the simulations and are the same as in Hülse and Ridgwell (2025). Boundary conditions (BC) are derived from HadCM3 simulations as described in Valdes et al. (2021).

Including a reference in the text. Line 183: "...its average value (Ecc-prec). Simulations description is summarized in Table 1."

2) The description of Step 1 and Step 2, which requires more careful thought to be fully understandable. The reader needs to know exactly what the forcings are. For instance, I suspect you have to fix the global content of PO4, CO2, and O2. If I am correct, how do you handle the link with the atmosphere? Are you fixing a global content for O2 and CO2 (including both oceanic and atmospheric reservoirs), with GENIE then calculating pO2 and pCO2 as a result of ocean-atmosphere exchange? In that case, do you also need to fix the alkalinity value? Indeed, from my own reading, I was surprised that GENIE reached the "correct" values for CO2 and PO4 when compared to Devonian constraints. A related question: GENIE finds 560 ppm but uses climate forcing from a simulation at 930 ppm. Could this mismatch imply incorrect results?

We agree that a precise identification of what is forced and what is prognostic is necessary for understanding the experimental design. The addition of Table 1, which explicitly summarizes the imposed forcings and boundary conditions for each simulation, is intended to address this point. In SPIN1, the model is run in a closed configuration, with weathering and sedimentation enabled, every chemical species lost through sedimentation being compensated by an equivalent input from weathering. In this setup, the global oceanic inventory of PO_4 is thus fixed. Atmospheric concentrations of CO_2 and O_2 are imposed in the form of an atmospheric restoring flux. Ocean–atmosphere gas exchange is active, but because atmospheric compositions are restored to fixed values, the atmosphere effectively acts as a prescribed boundary condition. In SPIN2, In SPIN2, weathering fluxes are set to compensate burial fluxes based on the results of SPIN1, hence ensuring carbon cycle equilibrium, and the only quantity that is held constant through forcing is atmospheric O_2, as cGENIE currently lacks a fully operational O_2 cycle. All other quantities, including PO_4 and alkalinity, are free to evolve as part of the open carbon cycle. This forcing strategy is retained for all subsequent simulations presented in the manuscript, while a complete technical description of the SPIN1 and SPIN2 process is available in Hülse and Ridgwell (2025).

Unlike the reference configuration of Hülse and Ridgwell (2025), our SPIN2 simulations exhibit a significant drift away from their initial state. We explored this behaviour extensively and found that both the magnitude and direction of the drift depend strongly on model resolution and continental configuration. Despite substantial effort, we were unable to prevent this SPIN2 drift in our experimental design. As a pragmatic solution, we therefore heuristically adjusted the initial pCO_2 in SPIN1 so that the subsequent drift leads the system towards a geochemical state consistent with available Late Devonian constraints. As a result, it is not coincidental that cGENIE converges towards CO_2 and PO_4 values that are compatible with proxy evidence; the initial conditions (also including the initial PO_4 values following GEOCLIM simulations) were chosen explicitly to guide the model towards such a state. We acknowledge that this approach is not ideal, but in the absence of a more robust alternative, we opted for this transparent and controlled strategy, which is explicit in the manuscript. We propose to add the following sentence to further clarify this.

Lines 163-164: "... consistent Late Devonian state. Part of this consistency has already been achieved by estimating the initial mean PO_4 concentration based on outputs from GEOCLIM. Hence, by the end of SPIN2..."

Finally, it is correct that the wind stress and albedo fields are taken from a HadCM3 simulation at 930 ppm CO_2, whereas cGENIE converges towards an atmospheric pCO_2 of approximately 560 ppm. This mismatch could, in principle, affect the results. However, our motivation for using the fields from Valdes et al. (2021) was to improve upon the highly idealized default boundary conditions in cGENIE, which consist of zonally averaged wind stress and a purely latitudinal albedo profile. Even with a CO_2 mismatch, the HadCM3-derived fields provide a more realistic representation of atmospheric circulation. Moreover, given the large uncertainties in Late Devonian continental reconstructions (e.g., Gérard et al., 2024, their Fig. 1), one could argue that a climate simulation performed at a matching pCO_2 would not per se yield boundary conditions that are unequivocally more "correct". Finally, we note a weak impact of adapting or not wind and albedo forcing fields as a function of pCO2 in GENIE (Fig. R1).

Second, because you're using cGENIE, the readers will expect a more detailed description of the ocean dynamics. I'd like to see more details (keeping in mind that there is ample room for this in Clim Past). I suggest adding at least: 1) Lat-lon maps of O2 at 400-1000 meters (OMZ ?) and at seafloor + one vertical transect (depth vs. latitude). The seafloor map will help clarify where the model will be affected by positive feedbacks due to P release from sediments. 2) Lat-lon map of primary production to show where productive areas are located. You should explain their distribution in relation to ocean dynamics—for instance, areas with subtropical gyres will likely have low PP due to downwelling. Showing this clearly will allow you to go beyond simple regression descriptions. The rationale behind these requests is that both reviewers highlighted that the simulated Devonian ocean is positioned close to an anoxic tipping point. While you note that this increases sensitivity to astronomical forcing, it also raises the possibility that your results might not hold under a more oxygenated baseline.

Following the editor's request, we have produced a new figure (Fig. R7) that provides a more detailed description of the simulated Late Devonian background state and ocean dynamics using cGENIE. This figure includes lat–lon maps of $[O_2]$ at 600 m depth, $[O_2]$ at the seafloor, and primary productivity, together with a meridional mean section of $[O_2]$ (depth versus latitude) and the meridional overturning circulation (Fig. R7a–e, respectively). We agree that these diagnostics would improve the general understanding of the spatial structure of oxygenation, productivity, and circulation in our simulations, as well as the identification of regions where sedimentary P regeneration feedbacks are likely to be most active.

To preserve the current flow and focus of the manuscript, we propose to include this new material in a dedicated annex rather than in the main text. This additional figure and its associated description provide important dynamical context but are not required to follow the primary results and conclusions of the study, which focus on the response of ocean oxygenation to astronomical forcing through weathering-driven nutrient supply. We believe this offers the best balance between completeness and clarity.

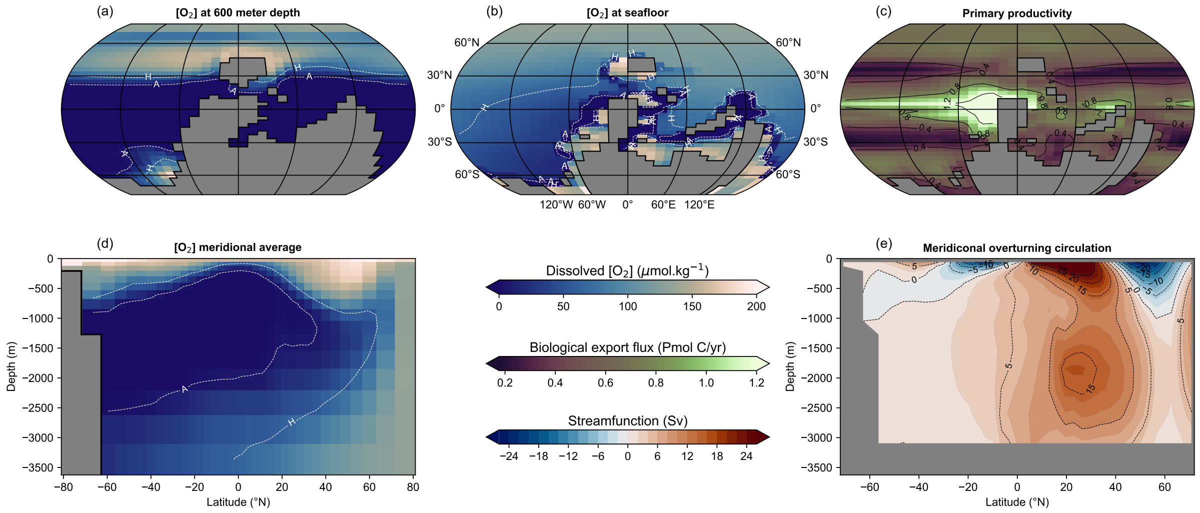

Figure R7: **Background state at the end of SPIN3.** $[O_2]$ at a depth of 600m **(a)** and at the seafloor **(b)**. **(c)** Primary productivity (Pmol C yr^{-1}). **(d)** Meridional average of $[O_2]$. **(e)** Meridional overturning streamfunction (1 Sv $\equiv 10^6$ $m^3\,s^{-1}$). The white dashed lines in **(a)**, **(b)**, and **(d)** indicate the boundaries for the anoxic (A) and hypoxic (H) regions, with threshold values taken after Sarr et al. (2022).

Additional annex, lines 415-424: "Figure A1a, b and d show a strong hemispheric asymmetry in oxygenation, with a substantially more oxygenated Northern Hemisphere compared to the Southern Hemisphere. This pattern primarily reflects large-scale ocean circulation (Fig. A1e), which allows well-oxygenated surface waters to ventilate the deep ocean in the Northern Hemisphere (Fig. A1d). In contrast, a pronounced oxygen minimum zone (OMZ) develops across much of the Southern Hemisphere, where meridional overturning circulation is nearly absent. The OMZ is shallower towards low latitudes, where primary productivity is high, consistent with enhanced biological oxygen demand (Fig. A1c). Primary productivity is also relatively elevated at high latitudes, while minimum values occur at mid latitudes, where downwelling associated with subtropical gyres prevents primary productivity. The seafloor [O_2] map further indicates that most anoxic conditions at the seafloor are restricted to shallow ocean regions, as confirmed by comparison with the model bathymetry (Fig. 1c). These shallow regions therefore correspond to areas where sedimentary P regeneration feedbacks are expected to be strongest."

We acknowledge that the background state used to initialize the transient simulations is strongly deoxygenated. Given the exceptionally high frequency of ocean anoxic events during the Late Devonian, such a state is to be expected. Nevertheless, we do not claim to have reconstructed a unique or exact Late Devonian background ocean state, as this interval spans a long duration and remains poorly constrained. Rather, our objective was to generate a climate and biogeochemical state that is broadly consistent with current understanding of Late Devonian continental configuration, boundary conditions, and commonly used modelling approaches.

Finally, while a highly deoxygenated background state does influence the extent of simulated anoxia, defined here using a fixed [O_2] threshold (same as Sarr et al., 2022), it does not affect the magnitude of the [O_2] changes induced by astronomical forcing. This distinction is crucial, as our conclusions rely on the amplitude and structure of oxygen variability rather than solely on changes in anoxic volume. Even in the absence of large variations in anoxic extent, the simulations demonstrate that astronomical forcing can induce substantial changes in oceanic [O_2], supporting its relevance for Late Devonian redox variability.

Third, your preliminary test regarding wind stress and albedo indicated that orbital-scale changes in circulation could account for a change in [O2] equivalent to roughly 35% of your primary weathering signal. In this light, dismissing physical feedbacks as "secondary" lacks sufficient quantitative support. You may discuss this further using results from the literature. For instance, the one of Sarr et al. (GRL) is particularly compelling showing that ocean dynamics changes in response to orbital parameters may induce changes in oceanic O2 by 20 to 40 micromoles.

To address a comment from Reviewer 1, we performed an additional sensitivity simulation in which the wind stress fields and albedo profile were modified for a specific astronomical configuration. Because of the requirements associated with the construction of the emulator, wind stress and albedo fields were only available for an extreme orbital configuration (eccentricity of 0.08). Under this configuration, we found that astronomically induced changes in circulation could lead to variations in [O_2] equivalent to approximately 35 % of the primary weathering-driven signal. This result indicates that physical circulation changes can contribute non-negligibly to oxygen variability, but within our experimental framework, they remain of second-order importance relative to the dominant effect of continental PO_4 weathering.

We insist that this result should not be interpreted as a claim that ventilation or physical feedbacks play a negligible role in ocean oxygenation. Rather, our experimental design is not optimized to robustly quantify the impact of circulation-driven oxygen changes, given the absence of dynamically consistent wind stress and albedo fields across orbital configurations. Importantly, the primary objective of this study is to test what is currently the dominant hypothesis regarding the orbital pacing of Late Devonian anoxic events, namely that astronomical forcing influences ocean redox conditions primarily through the modulation of continental weathering fluxes (De Vleeschouwer et al., 2017). Accordingly, our focus is on isolating this mechanism rather than on resolving the full spectrum of circulation–ventilation feedbacks on oxygenation.

We agree that explicitly acknowledging that physical feedbacks may exert a substantial influence on oxygenation, as demonstrated in previous studies, provides important context for our results. In this regard, studies such as Sarr et al. (2022) offer compelling evidence that astronomically driven changes in ocean circulation and ventilation can induce large regional oxygen anomalies (although this contribution may be largely dependent on the continental configuration, with the central Atlantic during the Late Cretaceous possibly providing an end-member case).

Lines 387-391: "...better connect orbital forcing with ocean anoxia. For instance, subsequent work would benefit from incorporating dynamic albedo and wind-stress profiles when exploring the full range of climate-circulation–biogeochemistry interactions. This is further encouraged by the results of Sarr et al. (2022), who showed that astronomically driven changes in ocean circulation and ventilation can exert a strong control on oxygenation, with regional variations reaching up to 40 μmol.kg^{-1} in a Late Cretaceous context."

A last comment, l. 177-190, what are the reasons for the CO2 changes? Is it related to RockGEM alk input or Corg burial?

Yes, as our experimental setup includes a functioning weathering module and the burial of C_{org}, both affecting the global carbon cycle, they induce variations in the atmospheric pCO_2. This clarification will be added to the revised manuscript (line 190).