Matthew Bogumil
Corresponding author for manuscript - EGUSPHERE-2025-5581

20 April 2026

Dear Editor and RC1,
Re: Edits of manuscript reference No. EGUSPHERE-2025-5581

*We thank the referee for a thorough and constructive review. By addressing it, we feel the manuscript has improved significantly. We have identified three core concerns, which we address first before proceeding to point-by-point responses. Throughout this response we have highlighted (**in bold text**) revised manuscript changes by section, page, and line, when relevant. We have also included some of our revised manuscript's major text blocks in our point-by-point responses when our blue text explanations need further details. We are happy to share the revised manuscript and difference file during the discussion period. However, we are instructed by the discussion reply platform to "NOT submit your revised manuscript here as supplement." We now proceed with the identified three core concerns:*

How does ExoCcycle differ from existing methods? *ExoCcycle is the first objective, automated framework for defining ocean basins from scalar fields on a sphere. Unlike hand-drawn definitions (e.g., IHB 1953; Dalvand, Dutkiewicz, Wright, Mather, et al. 2025; Dutkiewicz and Müller 2021) or flow-based methods that require high-resolution GCM output (e.g., Froyland, Stuart, and Sebille 2014), ExoCcycle operates on data-limited scalar fields (e.g., bathymetry alone), produces contiguous communities, handles multiple fields simultaneously, and resolves depth-dependent layered structure. We have added a systematic comparison to existing methods in the revised Discussion (**Section 5.1.1, Line 657**; see also **Table 3** in the manuscript).*

What new information does it provide? *Our analysis demonstrates three results that, while qualitatively expected, have not been quantified before: (1) a spatial resolution of 1–2 degrees is required to capture basin-defining features; (2) basin boundaries from reconstructed bathymetry differ from those using measured bathymetry due to unresolved volcanic features; and (3) the ocean's community structure is layered, with surface and deep basins governed by distinct physics, directly challenging the vertical-wall assumption in intermediate-complexity models. The silhouette analysis further shows that representing the ocean with fewer than ∼15 communities loses significant thermal and bathymetric structure (Figs. 8, 10b), providing the first quantitative constraint on box-model resolution. We now make this clear in **Section 5.3.1, lines 786-830 and Table 4**.*

Are the results predetermined by plate models? *Basin boundaries from ExoCcycle are not equivalent to plate boundaries. Transform faults, volcanic plateaus, seamount chains, and continental gateways, all of which control basin geometry, are second-order features relative to the kinematic plate boundary framework. We demonstrate this directly: basins detected from measured bathymetry (ETOPO1) differ from those detected from plate-model-derived reconstructed bathymetry (Fig. 9e-f), precisely because the former captures volcanic roughness that reconstructions smooth over. We have expanded a discussion of* *this distinction in Section 4.3 (**Page 32 Line 581-584**).*

We have also updated figures 7, 8, 9f, 10, 13, 14, 15, their captions, and corresponding values in Table 2 and the text (**Page 17 Line 363 and Page 24 Line 449**) to reflect a minor debugging on ETOPO1-based calculations. These updates do not meaningfully change our results or their interpretation.

1 General Comment: Discussion Too Long, No Comparison to Other Methods

RC1: The discussion is too long and very broad… I also miss a comparison of the workflow to other similar studies… the manuscript is not clear on how it provides any information that is not already known, or you could detect using other methods.

Response: We have shortened the Applications to other domains discussion section to a single paragraph (**Section 5.4, Page 42-43, Line 848-855**). We have also added a table (**Table 3, Section 5.1.1, Page 36**), text (**Section 5.1.1, Page 36, Line 657-674**) and a line (**Section 4.3 Page 32, Line 581-584**) comparing basin identification methods. These additions expand on the limitations and domain of existing basin identification methods. We have reproduced **Lines 657-674** and **Table 3** for reference.

Lines 657-674:

Our framework also allows for a more flexible basin detection analysis compared to current methods (Table 3). Flow-based approaches (e.g., Froyland, Stuart, and Sebille 2014) define basins from simulated surface transport, providing dynamically meaningful boundaries for the present-day ocean, but require high-resolution GCM output that is unavailable for most paleo-time slices. Geomorphic classification (e.g., Harris et al. 2014) partitions the seafloor by morphological type but does not produce basin-scale communities from the data's own covariance structure. Ecological provinces (e.g., Longhurst 2010) capture biologically meaningful surface domains but are defined from modern observations with heuristic thresholds, limiting applicability to deep time. Hand-drawn definitions (e.g., Dutkiewicz and Müller 2022; Dalvand, Dutkiewicz, Wright, Mather, et al. 2025) are based on expert judgment and are neither reproducible nor formally objective. ExoCcycle occupies a distinct niche: it operates on any scalar field (or combination), requires no flow-field input, produces contiguous communities with quantified stability, resolves depth structure, and generates boundary statistics needed for box-model parameterization. No existing method combines these capabilities.

Bathymetry-defined basins (explicitly defined by ExoCcycle methods) and circulation-defined basins serve complementary roles. The former captures the time-stable geometric container set by seafloor topography; the latter captures the time-varying flow regime operating within that container. For paleo-applications, bathymetry is the most reliably reconstructable field, making it the natural first-order constraint on basin geometry. Circulation changes within these basins, driven by orbital forcing, ice-sheet dynamics, or gateway evolution, can then be assessed by overlaying proxy-derived fields (e.g., ϵ-Nd and $\delta^{13}C$) onto the bathymetric basin framework. This two-stage approach — defining containers from bathymetry and characterizing their contents from proxies — separates the geometric problem (which ExoCcycle solves) from the dynamical problem (which requires additional data or modeling).

Table 3: Basin identification method comparison. ExoCcycle leverages scalar fields and graph-based community detection to algorithmically identify ocean basins for data-rich to data-poor systems. Our model is also capable of interrogating depth dependent ocean structure and providing basin and boundary statistics useful for structure analysis and carbon cycle box model constructions.

Feature	ExoCcycle	Froyland et al. (2014)	Harris et al (2014)	Longhurst (2010)	IHB 1953	Hand-drawn
Input Data	Any scalar field(s)	Ocean flow (GCM)	Seafloor Geomorphology	Ecology & SST	Geography	Expert judgement
Requires GCM?	No	Yes	No	Partially	No	No
Paleo-applicable?	Yes	No (needs paleo-GCM)	partially	No	No	Yes (by hand)
Objective?	Yes (algorithmic)	Yes (algorithmic)	Yes (classification)	Partially (heuristic)	No	No
Multi-field?	Yes (arbitrary)	No (single flow)	No	Partially	No	No
Depth-resolved?	Yes (any interval)	No	No	No	No	No
Contiguous?	Yes (graph-based)	No (may overlap)	Yes	Yes	Yes	Yes
Boundary Stats?	Yes (sill, width, distributions)	Imposed low flow condition	No	No	No	No

2 Plate Model Circularity: "Basin Boundaries Are Controlled by the Plate Model"

RC1: The basin boundaries are controlled mostly by the specific plate model, you can see this directly from the plate boundaries.

Response: We appreciate this concern but note a fundamental distinction, assuming we have understood the reviewer's point correctly: plate boundaries and basin boundaries are different physical entities. Transform faults offset ridges within the plate interior. Oceanic plateaus, seamount chains, and continental rises, all of which ExoCcycle identifies as basin boundaries, are second-order volcanic and sedimentary features unrelated to the kinematic plate boundary framework. If plate boundaries solely determined basin boundaries, we could divide the Atlantic into E and W, but that's not the case.

We demonstrate this distinction directly. Basin communities detected from measured present-day bathymetry (ETOPO1; Fig. 9f, 14d) differ from those detected from plate-model-derived reconstructed bathymetry (Fig. 9e), because measured bathymetry captures volcanic roughness (seamounts, plateaus, and faulting) that reconstructions smooth over. For example, the Ontong Java Plateau, Caribbean Plateau, and Kerguelen Plateau all generate basin boundaries in the measured field that are absent or weakened in reconstructions (Section 3.2, Fig. 9). This comparison demonstrates that ExoCcycle extracts information beyond what the plate model encodes.

We have added text to clarify the differences between plate model and bathymetry defined basins (**Page 32 Line 581-584**).

3 Lines 35-40: Box Model Design - "Determine the Least Number of Boxes"

RC1: I am wondering if there is a way one can use your workflow, taking a specific box model, and through time determine the least number of boxes needed for the specific problem the model is trying to solve?

Response: We agree with the review that this is an interesting question and a task that ExoCcycle has the capability to partially solve. The ideal total number of basins and magnitude of communication are not explicitly defined by ExoCcycle. However, outputs (e.g., connective bathymetry distributions, silhouette scores, and estimated sill depths) can be used to heuristically determine the number of basins needed to capture the seafloor (i.e., carbon sink) topology of ocean systems throughout geological time. Within the same vein, ExoCcycle is capable of defining the basin structure that is *smoothed over* in box-limited reduced-order models (e.g., LOSCAR). To address these concerns we have added 1) a **Table 4** (**Page 41**) that summarizes the number detected basins, number of strongly independent basins, ocean structure missed by 3-basin aggregation, and estimated sill depths useful for estimating basin mixing/separation, and 2) supporting text to our new "Earth applications" subsection (**Section 5.3.1, Page 40-41, Line 786-809**). We have reproduced the entire "Earth applications" section **Lines 786-830** and **Table 4** for reference here and in other point-by-point responses.

Lines 786-830: 5.3.1 Earth applications

Our basin detection algorithm also shows significant potential to describe a new class of Earth system models of intermediate-complexity and to analyze higher resolution GCMs. Given appropriate bathymetry reconstructions, our basin detection algorithm determines a first order approximation of regions that are expected to be chemically distinct. For reduced order model setups, users may construct ocean boxes and connections using a modular Earth system box model framework (e.g., LOSCAR and GEOCLIM7). Setups can leverage edge and internal basin bathymetry distributions to either impose or inform changes in circulation (e.g., by a clustering analysis of proxy circulation constraints such as ϵ_{Nd} *Batenburg et al. 2018; Tachikawa et al. 2020). Model constraining, with additional paleo proxies, and tuning, for poorly constrained volatile cycle parameters, will be application dependent.*

To illustrate how ExoCcycle output maps to box-model configurations, we compare detected basin structure across three time slices (Table 4). For the present day, SB-Reduction with Leiden (γ *= 0.01, ensemble = 50) identifies 22 deep-ocean basins (Fig. 17d), of which 17 maintain silhouette scores above*

0.7 (Table 4). The standard LOSCAR configuration assumes 3 deep basins (Atlantic, Pacific, Indian). The transition from 17 to 3 basins requires merging the Mediterranean, Arctic, equatorial Pacific, and Southern Ocean into their neighboring major basins, losing information about their distinct bathymetric distributions and gateway connections.

At 60 Ma, the opening Atlantic is narrower and the Tethys Ocean occupies a significant fraction of low-latitude ocean area. ExoCcycle identifies 12 deep basins, including distinct domains for the proto-North Atlantic, Tethys, Indian, and multiple Pacific sub-basins. A 3-box model merging these into Atlantic, Pacific, and Indian loses the Tethyan domain entirely, a basin with fundamentally different carbonate chemistry (Dutkiewicz and Müller 2022). At 80 Ma, the Western Interior Seaway and expanded Tethys generate additional distinct basins, and the minimum meaningful configuration increases to at least 23 basins.

The silhouette score provides an objective criterion for determining the coarsening threshold: the number of basins below which internal consistency degrades measurably. For volatile cycling models, this threshold represents the minimum basin count required to preserve the spatial structure of seafloor carbon reservoirs. The connective bathymetry distributions (Fig. 13d) provide the sill depths and gateway characteristics needed to parameterize exchange between these basins.

The geometric parameters extracted by ExoCcycle from basin boundaries — including sill depth, gateway width, and cross-sectional interface area between adjacent basins — provide the physical constraints required to parameterize inter-basin exchange in box models. For each pair of adjacent basins, ExoCcycle outputs the connective bathymetry distribution, from which the effective sill depth H_{sill} *and gateway width* W *are extracted. These quantities constrain the maximum density-driven volume exchange* Q *through hydraulic control (Whitehead, Leetmaa, and Knox 1974):* $Q = C_d W H_{sill} \sqrt{g' H_{sill}}$*, where* g' *is the reduced gravity across the interface. For gateways wider than the local Rossby deformation radius, the effective width is limited by rotation. Basin volumes are computed directly from the detected community masks and the bathymetric field. These geometric parameters, combined with density contrasts estimated from proxy data or GCM output, provide the exchange coefficients* $k_{ij} = \alpha Q / V_i$ *needed to parameterize inter-basin transport in box models. This approach replaces the hand-tuned exchange parameters currently used in models like LOSCAR and GEOCLIM7 with objective, bathymetry-derived values that can be recalculated for any time slice.*

For model-proxy comparison, basin detections allow for unambiguous clustering of paleo proxies and will enforce spatial self-consistency of model constraints. Basin detections for GCMs provide an additional regional analysis tool for both 1) explaining large-scale ocean chemistry evolution and 2) comparing to paleo proxies. Modelers can achieve comparisons by averaging GCM outputs over bathymetry defined basins and comparing with proxies (e.g., tracers) within the same basins. A related application arises in paleoclimate data assimilation. Ensemble-based methods (e.g., Ensemble Kalman Filtering) require covariance localization to suppress spurious long-range correlations from small ensembles. Standard approaches use isotropic distance tapers, but in the ocean, physical proximity does not guarantee dynamical connectivity: a distance-based taper will bleed proxy signals across mid-ocean ridges and continental barriers. ExoCcycle's basin definitions provide a topologically informed alternative, restricting covariance to within detected basin domains and preventing unphysical cross-basin signal propagation.

This approach replaces arbitrary distance parameters with objective, bathymetry-derived boundaries.

Table 4: Basin system parameters from 0, 60, and 80 Ma basin detections (Fig. 17) compared to the LOSCAR 3-box model basin system (R. E. Zeebe 2012). The gateway sills are estimated as the median depth of all shared basin boundary nodes. All values are calculated from basin detections in shown figure 17 except for modern gateway sills that are taken from Dalrymple 2023. These diagnostic characteristics can be used to determine the minimum number of basins to capture the topological structure of ocean systems throughout geological time.

Property	Present-day (reconstruction)	60 Ma	80 Ma	LOSCAR (3-box)
Deep basins Detected	22	14	26	3 (Atl, Pac, Ind)
Basins with silhouette > 0.7	17	12	23	3
Basins missed by 3-box model	Med, Arctic, eq. Pac, S. Ocean	Tethys, proto-S. Ocean	WIS, Tethys, multi Pac	—
Major gateway sills (km)	*Indonesia: 0.2-0.6, Gib: 0.3*	Drake: 0.5, N Atlantic-Pac: 3.2, Tethys-N Atlantic: 2.6	Drake: 0.3, WIS: 0.23, N Atlantic-Pac: 4.5, Tethys-N Atlantic: 0.5	N/A

4 Lines 73-76: Water Masses Differ Under Same Bathymetry

RC1: You could have completely different deep water in the major ocean basins depending on if the deep water is sourced from the south or from the North Atlantic… one can have quite different structures under the same bathymetric boundary conditions.

Response: The reviewer identifies a physically important point. Under the same bathymetric boundary conditions, fundamentally different water mass distributions can exist, for example, depending on whether deep water is sourced from the North Atlantic or from the Southern Ocean. This is precisely why ExoCcycle is designed to analyze multiple fields. Bathymetry-defined basins capture the first-order physical container: the geometry of the seafloor that constrains where water can flow. Temperature or salinity-defined basins capture the circulation regime operating within that container. Our framework allows users to compare these definitions and identify when and where they diverge, as demonstrated in Figures 15 (bathymetry-only) versus 18 (temperature-based). For paleo-applications where temperature and salinity fields are unavailable, bathymetry provides the most robust and reconstructable constraint on basin geometry. Changes in circulation within these containers can then be assessed using proxy data such as ϵ-Nd (Batenburg et al. 2018; Tachikawa et al. 2020), mapped onto ExoCcycle-defined domains. We have added text that expands on the connection between bathymetry defined basins, circulation, and integration to reduced-order models (**Section 5.1.1, Page 36, Line 667-674 - see Point 1** and our new "Earth applications" subsection **Section 5.3.1, Page 41-42, Line 810-830 - see Point 3**)

5 Lines 112-117: Similar Approaches, Wall Geometry

RC1: Line 112-117: " Interesting question but the manuscript do not go much into other or similar approached to trace/group water masses at basin scale. Also, box models generally have vertical walls, no? And does the geometry of the walls of the model even matter for most box models? To properly solve for e.g. bathymetry GCMs might be the option"

Response: We agree that the scientific question can be rewritten to more clearly describe the results shown in this manuscript. We have revised the question (**Page 6, Line 114 & Page 26, Line 470**) and further addressed 1) method comparisons (**in Point 1**) and 2) box-model design (**in Point 3**).

6 Section 4.4: Depth Threshold Justification (200 and 1000 m)

RC1: On the choice of what is deep and shallow ocean I would expect some rationale for the choices. E.g. why exclude intermediate waters?

Response: The 200 m and 1000 m thresholds correspond to established oceanographic regime transitions. The 200 m boundary approximates the base of the wind-driven Ekman layer and the photic zone in most of the global ocean; above this depth, lateral property gradients are dominated by wind-driven and buoyancy-forced surface currents. The 1000 m boundary corresponds approximately to the base of the main thermocline, below which lateral property gradients are set by thermohaline circulation rather than wind forcing (e.g., Talley 2011). These are not arbitrary; they capture the two primary transitions in the physics governing lateral ocean structure. We agree that the upper portion of the deep field (1000–2000 m) can differ from the abyssal ocean, but this variation is secondary to the first-order transition at 1̃000 m. ExoCcycle can analyze any depth interval, as demonstrated for intermediate depths in Figure 11d–f.

We have made adjustments to the text **(Section 1, Page 6, Line 121-123)** and **(Section 4.4, Page 32-33, Line 586-597)** explaining the choice of depth intervals.

7 Lines 602-610: "How Many Boxes Are Needed Through Time?"

RC1: How many boxes are needed through time in order gain something for specific models.

Response: This is an important question. Our silhouette analysis provides the first quantitative constraint on the minimum number of basins required to capture the ocean's spatial structure. For present-day bathymetry and temperature fields, community stability (area-averaged silhouette score) requires representation by at least ∼15 basins (Figs. 8a, 10b). Below this threshold, merging basins incurs measurable loss of internal consistency.

For a model like LOSCAR (currently 3 deep basins: Atlantic, Pacific, Indian), our results indicate that this simplification loses structure in the Mediterranean, Arctic, equatorial Pacific, and Southern Ocean (Fig. 15b). Whether this lost structure matters for a given research question — say, reconstructing the PETM CCD excursion — depends on whether these sub-basins respond differently to the forcing. ExoCcycle provides the tools to assess this: the silhouette score quantifies how much structure is lost at each level of coarsening, and the connective bathymetry distributions (Fig. 13d) quantify how strongly adjacent basins communicate.

Text added in response to **Point 3** addresses how ExoCcycle can be used to determine the number of basins needed to represent the topological structure (e.g., of bathymetry) and the implications of simplified reduced order models (e.g., LOSCAR's modern or paleo setup)

8 Mixing: "You Mention Mixing Throughout but Say Little"

RC1: You mention mixing throughout the manuscript but says little on why that is important and how you can improve estimations.

Response: We agree that mixing parameterization is central to the utility of basin definitions for volatile cycling models. While a full mixing parameterization is beyond the scope of this methodology paper, our framework provides the geometric inputs that mixing parameterizations require. We added text in our new "Earth applications" subsection **(Section 5.3.1, Page 41, Line 810-820 - see Point 3)** that clarifies what these parameters are and how they can be used. We also briefly explain the physical exchange flux between two ocean boxes is fundamentally constrained by the geometry of their shared boundary: the sill depth, the boundary length, and the cross-sectional area of the gateway. ExoCcycle calculates these quantities directly from the detected basin structure and the connective bathymetry distributions (Fig. 13d). Polzin et al. 1997 demonstrated that diapycnal mixing intensity correlates with seafloor roughness, and our DQT-CDF weighting scheme captures this roughness signal in the edge weights. The boundary statistics ExoCcycle generates (sill depths, boundary widths, connective bathymetry distributions) are precisely the geometric inputs needed to parameterize inter-basin exchange in box models, even for paleo-configurations where direct observations are unavailable.

9 Figure 19: "Minimum Number of Boxes for GEOCLIM7?"

RC1: I would like to see a more detailed description of this figure in the text.

Response: We agree with the reviewer and have appended a more detailed description to figure 19's caption **(Section 4.4, Page 34, Figure 19)** and text and a table (**Table 4**) expanding on the procedure using ExoCcycle to determine the number of ocean basins needed to represent a system (**Section 5.1.1 Page 36, Line 667-674 - see Point 1** and **Section 5.3.1, Page 40-41, Line 795-809 - see Point 3**). Figure 19's caption is reproduced here for reference.

Figure 19 caption:

A simple example of alternative intermediate-complexity Earth system model constructed with modular frameworks (e.g., GEOCLIM7 (Maffre et al. 2025) or ESBMTK (Wortmann et al. 2025)). Multiple shelf regions feed into shallow and intermediate-deep boxes through surface currents and turbidity currents, respectively. The n^{th} *box illustrates characterization of non-uniform layered basin structure. ExoCcycle's ocean and shelf basin detection framework can be used to systematically define box domains* *and communication pathways between modular frameworks as discriminated by circulation impedance fields (i.e., bathymetry) or chemical fields (e.g., salinity). The ideal total number of basins and magnitude of communication are not explicitly defined by ExoCcycle. However, outputs, e.g., connective bathymetry distributions (Fig. 13), sill depths (Table 4), and silhouette scores (Table 4 and Fig. 8a) can be used to heuristically determine the number of basins and strength of mixing/circulation for modular frameworks.*

10 Lines 578-580: "Please Be More Specific"

RC1: Please be more specific and elaborate on how this is/can be done.

Response: We have added a subsubsection named "Earth applications" (**Section 5.3.1, Page 40-42, Line 786-830 - see Point 3**) to the discussion that expands on how detected basins can be used to construct box models from modular volatile cycling frameworks.

11 Section 5.1.2: "You Have Not Shown Any of This"

RC1: How does ExoCcycle's workflow provides a systematic and objective method for grouping seafloor sediment core data. . . ? You have not shown any of this? How does it work with proxies, drill sites, geological information, sediments etc?

Response: The reviewer raises an important point. Our method provides the spatial framework for clustering paleo-proxy sites into objectively defined basin domains. Specifically, once basins are defined from bathymetry (or any scalar field), discrete drill-site locations (e.g., DSDP/ODP/IODP sites) are assigned to basins by spatial membership. This enables rigorous calculation of intra-basin variance versus inter-basin variance for any proxy quantity, replacing hand-drawn regional groupings with objective, reproducible domains.

We acknowledge that the present manuscript demonstrates the basin-detection methodology and validates it on modern data, but does not include a full worked example with actual proxy data. A dedicated study applying ExoCcycle basin definitions to CCD reconstructions with DSDP/ODP/IODP drill sites is in preparation (see Bogumil, Mittal, and Lithgow-Bertelloni n.d. for preliminary work). In the revised manuscript, we have clarified the proxy-integration workflow in Section 5.1.2 (**Page 37-38, Line 701-723**) and the new "Earth applications" subsection (**5.3.1, Page 41-42, Line 821-830 - see Point 3**). We have reproduced Line 701-723 of the revised manuscript for reference here.

Line 701-723:

... More fundamentally, ExoCcycle basin detections allow for for proxy-data clustering with evolving basins even if calculations are restricted to vertically averaged basins, as in previous work.

The proxy-integration workflow proceeds in three steps. First, ExoCcycle defines basin domains Ω_1,,Ω_2, ..., Ω_k *from paleo-bathymetry for a given time slice. Second, discrete proxy measurements (e.g.,* $\delta^{13}C$ *from sediment cores at locations* x_m, y_m*) are assigned to the basin containing their paleo-coordinates, producing basin-specific proxy populations (eq. 10), where* m *is the proxy site index,* k *is the basin index, and* B *is a function that maps* m *to* k*.*

$$B(m) = k, \text{ for } (x_m, y_m) \in \Omega_k \tag{10}$$

Sites within a user-defined distance $d_{threshold}$ *of a basin boundary are flagged as ambiguous and can be excluded or shared between adjacent basins. Third, intra-basin statistics and inter-basin contrasts are computed (eq. 11 12, and 13), where* N_k *is the number of sites in the* k^{th} *basin and* P_m *is a proxy site value.*

Intra-basin mean:

$$\bar{P}_k = \frac{1}{N_k} \sum_{m \in \Omega_k} P_m \tag{11}$$

Intra-basin variance:

$$\sigma^2_{intra,k} = \frac{1}{N_k} \sum_{m \in \Omega_k} \left(P_m - \bar{P}_k\right)^2 \tag{12}$$

Inter-basin variance:

$$\sigma^2_{inter} = \frac{1}{N_k} \sum_{m \in \Omega_k} \left(\bar{P}_k - \bar{P}_{global}\right)^2 \tag{13}$$

The ratio $\sigma^2_{inter}/\sigma^2_{intra}$ *quantifies whether basins are chemically distinct. A high ratio means the basin definitions capture real spatial structure in the proxy field; a ratio near 1 means the proxy field has no basin-scale structure (or the basin definitions are wrong). This spatial-join approach replaces qualitative regional groupings with objective, reproducible domains whose boundaries are derived from the physical geography of the seafloor.*

This provides a far more physically-grounded basis for regional CCD reconstructions and other proxy averaging. Additionally, a vertically averaged or layered-basin intersection map can also be used strategically to identify critical gaps in the existing seafloor core network, guiding future sampling campaigns to the most diagnostic locations.

12 Section 5.1.3: "Says Nothing About Proxies!"

RC1: Yet (although the title suggests so), it says nothing about proxies! Output form deep time GCM simulations are not proxies.

Response: The reviewer is correct that the original title was misleading. We have changed it to “Basin Reconstructions and High Resolution Model Integration” **(Page 38, Line 724)**. We have also added text clarifying how ExoCcycle basin detections provide the spatial framework for model-data comparison **(Section 5.1.3, Page 38-39, Line 730-736)**: GCM output fields are aggregated within detected basins and compared to basin-averaged proxy observations using the added workflow described in **Section 5.1.2 (Page 37-38, Line 701-723 - see Point 11)**. This connects the high-resolution model integration discussion to the concrete proxy-handling methodology. The new “Earth applications” subsection (**Section 5.3.1, Page 42, Line 821-830 - see Point 3**) further develops this connection. We have reproduced Line 730-736 of the revised manuscript for reference here.

Line 730-736:

Using paleo-reconstructed fields from high-resolution GCM simulations, ExoCcycle basin detections *provide the spatial framework for model-data comparison: GCM output fields can be aggregated within detected basin domains and compared directly to basin-averaged proxy observations (Section 5.1.2). This two-step approach — defining basins from bathymetry, then evaluating model fidelity within those basins — separates the geometric question (basin structure) from the dynamical question (whether the model reproduces observed tracer distributions within each basin). The intra-basin versus inter-basin variance diagnostic described in Section 5.1.2 applies equally to model output and to proxy data, providing a consistent metric for assessing whether a given model resolution captures the spatial heterogeneity that the proxy record demands.*

13 Final Section 5 Comment: Proxy Integration as Key Contribution

RC1: You might be missing what could be the key contribution from your algorithm, which is quantifiable/reproducible basin definitions though time based on paleogeography and paleo proxies… if you can combine that with paleo ocean proxy reconstructions... you might do something quite novel with wide applications for paleo-oceanography/climatology.

Response: We agree that this is an exciting and natural extension of the framework. ExoCcycle is designed to accept multiple scalar fields simultaneously (Section 3.5, Fig. 12), so combining reconstructed bathymetry with proxy-derived fields (e.g., paleo-temperature, salinity, or isotopic reconstructions) is a direct application of the existing methodology. The current sparsity of deep-time proxy data makes direct basin definition from proxies alone challenging; however, ExoCcycle's bathymetry-derived basins provide the objective spatial framework within which sparse proxy data can be rigorously aggregated and tested for internal consistency (see our new proxy-integration workflow description in **Section 5.1.2** (**see Point 11 and 12**)).

We are actively pursuing this direction. A study applying ExoCcycle basin definitions to regional CCD reconstructions using DSDP/ODP/IODP drill-site data is in preparation Bogumil, Mittal, and Lithgow-Bertelloni n.d. In the revised manuscript, we have strengthened the discussion of proxy integration in Section 5.1.2 (**see Point 11 and 12**), the "Earth applications" subsection 5.3.1 (**see Point 3, 4, 7-12**), and the updated **Figure 2** and introduction text (**Page 2, Line 53–54**) to emphasize the importance of time-evolving, proxy-informed basin definitions.

Regarding the comparison to plate-model-derived basins: we have demonstrated that ExoCcycle basins from measured bathymetry differ from those derived from plate-model reconstructions due to second-order volcanic features (Fig. 9e–f and **see Point 2**). This difference is not a limitation but a feature: it quantifies the information content of volcanic roughness in defining basin geometry. Even for reconstructed bathymetry, ExoCcycle captures gateways, restricted epicontinental seas, and volcanic barriers that simple plate-boundary polygons do not resolve (e.g., Western Interior Seaway, Tethyan gateway region; Fig. 17f).

We have reproduced Line 53–54 and Figure 2 of the revised manuscript for reference here.

Line 53–54:

Can basin definitions be left unchanged for 10s of myr as shown in Figure 2b-d and numerous studies (Dutkiewicz and Müller 2022; Li et al. 2023; Wallmann et al. 2019; Dalvand, Dutkiewicz, Wright, and

Müller 2025)?

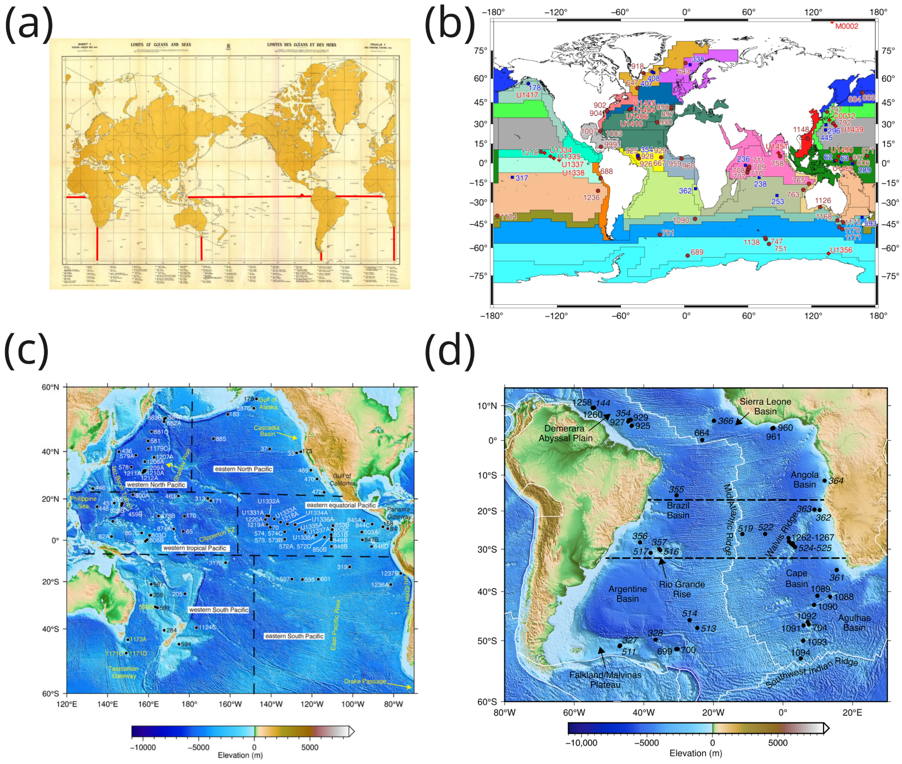

Figure 2: **Present-day defined ocean basins from (a) Limits of Oceans and Seas (IHB 1953), (b) Longhurst 2010, (c) Dalvand, Dutkiewicz, Wright, Mather, et al. 2025, and (d) Dutkiewicz and Müller 2021 used for paleo-proxy data aggregation. Many of these boundaries differ significantly and have been used to reconstruct regional carbonate compensation depths and explore basin redox conditions that extend to global changes in atmospheric CO_2. Generally, these boundaries are hand-drawn and defined by *surface ocean properties that do not necessarily stay constant throughout 10s of Myr*.**

14 A 10 Myr Systematic Basin Sequence Analysis

RC1: I would rather think that if you were to run every 10 Myr of the last 200 Ma, that an algorithm like this (fast, transparent, reproducible) would be nice addition to the plate models.

Response: We agree that extending this analysis to a systematic sequence of reconstructions at 10 Myr intervals throughout the Mesozoic and Cenozoic is a natural next step. We show a partial application of this in figure 17. The computational efficiency of the Leiden-based composite algorithm (Table 2) makes an entire Mesozoic and Cenozoic basin evolution computation feasible: basin detection at 1-degree resolution completes in approximately 1 minute per time slice, enabling a 200 Myr sequence in under an hour. We have not included such an analysis in this methodology paper since we focus on breadth of applications for the ExoCcycle community detection framework.

15 Minor Fixes

RC1: Line 10-15 "2 & 3 seem obvious."

Response: We agree that the language should be clarified. We have changed the text on these lines **(Page 1, Line 10-15)** to the following: "(2) basin boundaries evolve significantly over geological time, underscoring the inadequacy of using static and time inappropriate boundaries for past climate simulations and data aggregation; and (3) contiguous ocean community *structure* is fundamentally layered and changes with depth - deep basins (defined by bathymetry) are distinct from shelf partitions (shaped by sedimentation, sea-level changes, and riverine fluxes) and surface basins (driven by wind and temperature/precipitation)."

RC1: Line 40-45 "Very general, please be more specific. How will the tools analyze/help quantify conditions for planetary habitability?"

Response: We agree that clarification is needed in this paragraph and have updated the summary/application lines (**Page 2, Line 44-49**).

RC1: Figure 1 "I do not see any question marks"

Response: We have updated figure 1 to include the question marks that were supposed to be located along the black basin boundary lines. We have also updated the caption which clarifies the bathymetry characteristics (from figure 1a-b) that influence model setups (figure 1c) (**Page 3, Figure 1**).

RC1: Line 487 "«continental choke points and gateways». What is the difference? What do you mean by continental choke points?"

Response: Gateways are narrow gaps that are used to describe pathways that connect large bodies of water (e.g.., basins). We used the term continental choke points so as to not imply the, a priori, existence of a gateway. In our model, ocean gateways can be an emerging feature even if the bathymetry model was not explicitly designed with a particular gateway in mind. All ocean gateways are emergent in the case of basin detection with measured topography. However, for simplicity, we change each reference of "continental choke points" to "gateways" (**Page 26, Line 489; Page 29, Line 520; Page 32, Line 576, last line of Figure 7's caption**).

RC1: Line 549-551 "Not very specific, please elaborate on how EcoCcycle is able to solve this."

Response: We changed the wording to make the application more clear. The wording changed to "ExoCcycle's ability to capture shelf structure based on shelf growth from nearby rivers and decay into neighboring deep basins— based on the very topography shaped by these processes— can help quantify the partitioning of riverine and turbidity fluxes into multiple shallow and deep basins. This type of analysis may elucidate local spatial variation in deep-basin chemistry and nearby sediment cores." (**Page 30, Line 551-554**).

RC1: Line 567 "Fig 17? Cannot see this in 16."

Response: This is correct, Figure 17 is supposed to be cited here. We have updated the text to reflect this change (**Page 31, Line 570**).

RC1: Line 575-576: "Is it assumed that box structure in the ocean should be static?"
Response: Yes, box structure has been assumed to be constant in some paleoclimate analyzes. Then only bathymetry might be updated as in Komar and Richard E. Zeebe 2021 or not at all over 10s of myr as in Papadomanolaki et al. 2022. However, the box structure is not being determined with a systematic approach as we have described in our manuscript. We have changed the sentence to clarify the statement and provided an examples (Papadomanolaki et al. 2022; Komar and Richard E. Zeebe 2021) (**Page 32, Line 577-579**).

RC1: "Why would basin change matter? Are you talking about continued habitability of ecosystems within specific basins?"
Response: The expansion of section 5.1.1, indicated in **Point 1 and 4**, also supports the paragraph discussing basin stability. Additionally we have changed the introductory sentence to clarify ExoCcycle's role in measuring regional field stability for these applications (**Section 5.1.1, Page 36, Line 675-676**).

RC1: "«corresponding to present-day and past river sediment supply from now submerged river systems that span throughout the Sunda Shelf» Please show this correlation with a figure and provide a references!"
Response: References (Voris 2000; P. D. Clift 2006; P. Clift, Layne, and Blusztajn 2004) were cited in section 4.2 when we described our result. We did not initial repeat the citations in the figure caption, but we have now included them (**last sentence of Figure 16**) such that readers can more easily compare our results with previous published analyses.

We thank RC1 for their comments and made the necessary changes to convey ExoCcycle's methodology and applications in an easily digestible manner.

Matthew Bogumil

References

Batenburg, S. J. et al. (2018). “Major intensification of Atlantic overturning circulation at the onset of Paleogene greenhouse warmth”. In: *Nature Communications* 9.1, p. 4954. ISSN: 2041-1723. DOI: 10.1038/s41467-018-07457-7. URL: https://doi.org/10.1038/s41467-018-07457-7.

Bogumil, Matthew, Tushar Mittal, and Carolina R. Lithgow-Bertelloni (n.d.). “The evolving boundaries and number of ocean basins since the Late Cretaceous”. In: *AGU Fall Meeting 2025*.

Clift, Peter, Graham Layne, and J. Blusztajn (2004). “The erosional record of Tibetan uplift in the East Asian marginal seas”. In: vol. 149, pp. 255–282.

Clift, Peter D. (2006). “Controls on the erosion of Cenozoic Asia and the flux of clastic sediment to the ocean”. In: *Earth and Planetary Science Letters* 241.3, pp. 571–580. ISSN: 0012-821X. DOI: https://doi.org/10.1016/j.epsl.2005.11.028. URL: https://www.sciencedirect.com/science/article/pii/S0012821X05008009.

Dalrymple, Robert W. (2023). “A review of the morphology, physical processes and deposits of modern straits”. In: *Straits and Seaways: Controls, Processes and Implications in Modern and Ancient Systems*. Ed. by V. M. Rossi et al. Geological Society of London, p. 0. ISBN: 9781786205704. DOI: 10.1144/sp523-2021-76. URL: https://doi.org/10.1144/SP523-2021-76.

Dalvand, Faranak, Adriana Dutkiewicz, Nicky M. Wright, Ben R. Mather, et al. (2025). “Regional carbonate compensation depth variability in the Pacific Ocean since the Oligocene”. In: *Frontiers in Earth Science* Volume 13 - 2025. ISSN: 2296-6463. DOI: 10.3389/feart.2025.1605906. URL: https://www.frontiersin.org/journals/earth-science/articles/10.3389/feart.2025.1605906.

Dalvand, Faranak, Adriana Dutkiewicz, Nicky M. Wright, and R. Dietmar Müller (2025). “Indian Ocean carbonate compensation depth since the Late Oligocene”. In: *Geo-Marine Letters* 45.4, p. 38. ISSN: 1432-1157. DOI: 10.1007/s00367-025-00825-5. URL: https://doi.org/10.1007/s00367-025-00825-5.

Dutkiewicz, Adriana and R. Dietmar Müller (2021). “The carbonate compensation depth in the South Atlantic Ocean since the Late Cretaceous”. In: *Geology* 49.7, pp. 873–878. ISSN: 0091-7613. DOI: 10.1130/g48404.1. URL: https://doi.org/10.1130/G48404.1.

— (2022). “The History of Cenozoic Carbonate Flux in the Atlantic Ocean Constrained by Multiple Regional Carbonate Compensation Depth Reconstructions”. In: *Geochemistry, Geophysics, Geosystems* 23.11, e2022GC010667. DOI: https://doi.org/10.1029/2022GC010667. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022GC010667.

Froyland, Gary, Robyn M. Stuart, and Erik van Sebille (2014). “How well-connected is the surface of the global ocean?” In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 24.3. ISSN: 1054-1500. DOI: 10.1063/1.4892530. URL: https://doi.org/10.1063/1.4892530.

Harris, P. T. et al. (2014). “Geomorphology of the oceans”. In: *Marine Geology* 352, pp. 4–24. ISSN: 0025-3227. DOI: https://doi.org/10.1016/j.margeo.2014.01.011. URL: https://www.sciencedirect.com/science/article/pii/S0025322714000310.

IHB (1953/09/01 1953). *Limits of the Oceans and Seas*. Report.

Komar, Nemanja and Richard E. Zeebe (2021). “Reconciling atmospheric CO2, weathering, and calcite compensation depth across the Cenozoic”. In: *Science Advances* 7.4, eabd4876. DOI: doi:10.1126/sciadv.abd4876. URL: https://www.science.org/doi/abs/10.1126/sciadv.abd4876.

Li, Ziye et al. (2023). “Neogene burial of organic carbon in the global ocean”. In: *Nature* 613.7942, pp. 90–95. ISSN: 1476-4687. DOI: 10.1038/s41586-022-05413-6. URL: https://doi.org/10.1038/s41586-022-05413-6.

Longhurst, Alan R (2010). *Ecological geography of the sea*. Elsevier. ISBN: 0080465579.

Maffre, P. et al. (2025). “GEOCLIM7, an Earth system model for multi-million-year evolution of the geochemical cycles and climate”. In: *Geosci. Model Dev.* 18.18. GMD, pp. 6367–6413. ISSN: 1991-9603. DOI: 10.5194/gmd-18-6367-2025. URL: https://gmd.copernicus.org/articles/18/6367/2025/.

Papadomanolaki, Nina M et al. (2022). “Quantifying volcanism and organic carbon burial across Oceanic Anoxic Event 2”. In: *Geology* 50.4, pp. 511–515. ISSN: 0091-7613.

Polzin, K. L. et al. (1997). “Spatial Variability of Turbulent Mixing in the Abyssal Ocean”. In: *Science* 276.5309, pp. 93–96. DOI: doi:10.1126/science.276.5309.93. URL: https://www.science.org/doi/abs/10.1126/science.276.5309.93.

Tachikawa, Kazuyo et al. (2020). “RECONSTRUCTION OF OCEAN CIRCULATION BASED ON NEODYMIUM ISOTOPIC COMPOSITION Potential Limitations and Application to the Mid-Pleistocene Transition”. In: *Oceanography* 33.2, pp. 80–87. ISSN: 10428275, 2377617X. URL: https://www.jstor.org/stable/26937744.

Talley, Lynne D (2011). *Descriptive physical oceanography: an introduction*. Academic press. ISBN: 0080939112.

Voris, Harold K. (2000). “Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations”. In: *Journal of Biogeography* 27.5, pp. 1153–1167. ISSN: 0305-0270. DOI: https://doi.org/10.1046/j.1365-2699.2000.00489.x. URL: https://doi.org/10.1046/j.1365-2699.2000.00489.x.

Wallmann, Klaus et al. (2019). “Periodic changes in the Cretaceous ocean and climate caused by marine redox see-saw”. In: *Nature Geoscience* 12.6, pp. 456–461. ISSN: 1752-0908. DOI: 10.1038/s41561-019-0359-x. URL: https://doi.org/10.1038/s41561-019-0359-x.

Whitehead, J. A., A. Leetmaa, and R. A. Knox (1974). “Rotating hydraulics of strait and sill flows†”. In: *Geophysical Fluid Dynamics* 6.2. doi: 10.1080/03091927409365790, pp. 101–125. ISSN: 0016-7991. DOI: 10.1080/03091927409365790. URL: https://doi.org/10.1080/03091927409365790.

Wortmann, U. G. et al. (2025). “The Earth Science Box Modeling Toolkit (ESBMTK 0.14.0.11): a Python library for research and teaching”. In: *Geosci. Model Dev.* 18.4. GMD, pp. 1155–1167. ISSN: 1991-9603. DOI: 10.5194/gmd-18-1155-2025. URL: https://gmd.copernicus.org/articles/18/1155/2025/.

Zeebe, R. E. (2012). “LOSCAR: Long-term Ocean-atmosphere-Sediment CArbon cycle Reservoir Model v2.0.4”. In: *Geosci. Model Dev.* 5.1. GMD, pp. 149–166. ISSN: 1991-9603. DOI: 10.5194/gmd-5-149-2012. URL: https://gmd.copernicus.org/articles/5/149/2012/.