

UCLA

UCLA Electronic Theses and Dissertations

Title

The Mantle Crux: Solid-Earth Controls on the Long-Term Carbon Cycle

Permalink

<https://escholarship.org/uc/item/7kg7x7h7>

ISBN

9798190911572

Author

Bogumil, Matthew Walter

Publication Date

2026-09-04

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

The Mantle Crux: Solid-Earth Controls on the Long-Term Carbon Cycle

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Geology

by

Matthew Bogumil

© Copyright by
Matthew Bogumil

2026

ABSTRACT OF THE DISSERTATION

The Mantle Crux: Solid-Earth Controls on the Long-Term Carbon Cycle

by

Matthew Bogumil

Doctor of Philosophy in Geology

University of California, Los Angeles, 2026

Professor Carolina Lithgow-Bertelloni, Chair

The long-term carbon cycle on million-year timescales is shaped by planetary interior evolution, which governs CO_2 outgassing, volatile recycling, topography, and ocean-floor bathymetry. Because marine sediments and lithified carbonates represent the largest hydrosphere-accessible carbon sink through much of Earth history, changes in bathymetry, continental flooding, and ocean basin geometry may strongly influence ocean-atmosphere chemistry. This dissertation investigates how solid-Earth evolution controls bathymetry, ocean basins, and the long-term carbon cycle across Earth history and rocky ocean-world analogs.

First, I reconstruct bathymetry over the last 80 Myr and evaluate its carbon-cycle consequences using an Earth system model of intermediate-complexity (EMIC). These results show that evolving seafloor geometry significantly affects ocean chemistry and carbon partitioning. Faster tectonics produces shallower pelagic oceans and greater continental flooding, reducing deep-ocean carbon sequestration. However, pelagic bathymetry and shallow shelf area can evolve independently, producing competing effects on carbon storage under different climatic and sea-level states.

Second, I develop *slabmodellib*, a kinematic subduction framework that reconstructs slab histories from plate-motion and seafloor-age models, incorporates slab deflection at the ringwoodite–bridgmanite phase transition, and computes the resulting non-hydrostatic geoid and dynamic topography. Applied to the last 80 Myr, the model shows that mantle-driven

topography and geoid anomalies reorganize coastline length and seafloor area without a simple linear relationship. Dynamic topography generally supports shallow shelves, reduces intermediate-depth seafloor, and has smaller effects on deep-ocean area, implying that mantle flow can alter carbon-cycle boundary conditions by redistributing shelf and pelagic carbon sinks.

Finally, I develop *ExoCycle*, a graph-theory community-detection framework for objectively defining ocean basins from single or stacked scalar fields on a sphere. By applying Louvain, Leiden, and Girvan-Newman methods to bathymetry and related fields, ExoCycle identifies consistent basin boundaries for paleo reconstructions and statistically constrained exoplanet topographies. This enables basin-scale data aggregation, model assimilation, and objective spatial domains for reduced-order carbon-cycle and volatile-cycle models.

Together, these results show that the solid Earth influences the long-term carbon cycle not only through volcanic forcing and recycling, but also by shaping the geometry, depth distribution, and connectivity of ocean basins on Earth and other rocky ocean worlds.

The dissertation of Matthew Bogumil is approved.

Lars P. Stixrude

Caroline Beghein

Craig E. Manning

Carolina Lithgow-Bertelloni, Committee Chair

University of California, Los Angeles

2026

To my parents, Lucie, Micheal, David, Elizabeth, Salt, and Pepper

Contents

Abstract	ii
List of Figures	ix
Acknowledgements	xxi
1 Introduction	1
1.1 Carbon Reservoirs	2
1.2 Long-Term Carbon Cycle Processes	5
1.3 Carbon Cycle Proxies and Their Spatial Interpretation	8
1.4 Planetary Cooling Regimes	10
1.5 Plate Tectonics on Earth	11
1.6 Solid Earth Controls on Ocean Basin Geometry	13
1.7 Research questions and dissertation structure	16
2 The effects of bathymetry on the long-term carbon cycle and CCD	18
2.1 Significance	18
2.2 Abstract	18
2.3 Introduction	19
2.4 Results	24
2.4.1 Synthetic Bathymetry	24
2.4.2 Reconstructed Bathymetry	28
2.4.3 Early Paleogene Case Study	30
2.5 Discussion	33
2.6 Materials and Methods	37
2.6.1 Bathymetry Reconstructions	37
2.6.2 C Cycle Modeling	37
2.7 Data, Materials, and Software Availability	38
2.8 Acknowledgments	39
3 The Influence of Dynamic Topography and Non-hydrostatic Geoid on the Long-Term Carbon Cycle	40
3.1 Introduction	40
3.2 Mantle and Surface Simulations	42
3.2.1 Subduction, Mantle Flow, and Surface Expression	44
3.2.2 Subduction History Parameter Choice	47

3.3	Results	48
3.3.1	Subduction history Dynamic Topography & Geoid	48
3.3.2	Dynamic Topography & Geoid's Contribution to Topography	50
3.4	Discussion	53
4	ExoCycle v1.0.0: A Generalized Framework for Spherical Community Detection and its Application to Defining Global Ocean Basins from Multi-Field Data	57
4.1	Abstract	57
4.2	Introduction	58
4.3	Methods - Basin Community Detection	65
4.3.1	Quality Functions	66
4.3.2	Community Detection Algorithms	68
4.3.2.1	Agglomerative Algorithms - Louvain & Leiden	69
4.3.2.2	Top-down Algorithms - Girvan-Newman	72
4.3.3	Spatial Graph & Weight Construction	72
4.3.4	Composite Algorithm	77
4.3.5	Post-processing Algorithms	78
4.3.6	Computational Efficiency	78
4.3.7	Evaluation Metrics	79
4.4	Method Results	81
4.4.1	Accuracy of our approach - Stability of communities	81
4.4.2	Controls on Basin Definition - Role of Seafloor Bathymetry roughness	83
4.4.3	Parameter choices in the clustering algorithm	86
4.4.4	Reduction Method - Leiden vs Louvain	88
4.4.5	Multifield Community Detection	90
4.4.6	Field Statistics: Measured Bathymetry	90
4.5	Scientific Results & Discussion	93
4.5.1	Resolving Field Feature Wavelength: Seafloor Bathymetry Resolution	94
4.5.2	Shelf and Pelagic Ocean Basins: Seafloor Bathymetry	96
4.5.3	Paleo-Ocean Basins: Reconstructed Bathymetry Change	101
4.5.4	Surface Ocean Basins: Ocean Temperature - surface vs deep	103
4.6	Discussion	107
4.6.1	Thermochemical Fields	107
4.6.1.1	Surface Ocean Basin Structure	107
4.6.1.2	Carbon Cycle Diagnostics and Carbonate Compensation Depth (CCD)	110
4.6.1.3	Basin Reconstructions and High Resolution Model Integration	113
4.6.2	Bathymetry Field	114
4.6.2.1	Global Bathymetric Modeling: Gateways, Continental Shelves, & Rugosity	115
4.6.2.2	Role of Oceanic Plateaus in Basin Evolution	115
4.6.3	Toward a Comprehensive Planetary Carbon Cycle	116
4.6.3.1	Earth applications	116
4.6.3.2	Exoplanet long-term habitability	119

4.6.4	Applications to other domains	120
4.7	Conclusion	120
4.8	Code & Data Availability	121
4.9	Acknowledgements	122
5	ExoCycle Applications to Data Aggregation and Exoplanet Habitability	123
5.1	Introduction	123
5.2	Data Aggregation	124
5.2.1	Results	124
5.2.2	Implications	126
5.3	Exoplanets - exobasins & habitability	129
5.3.1	Results	130
5.3.2	Implications	133
5.4	Applications to other domains	134
6	Conclusion	137
	Appendices	138
A	Supporting Information from Chapter 2	139
A.1	Methods	139
A.1.1	(A) LOSCAR model description	139
A.1.2	(B) Bathymetry basin volume corrections	140
A.1.3	(C) Ocean volume calculations and changing initial conditions	141
A.1.4	(D) Source to sink equilibrium	142
B	Supporting Information from Chapter 3	156
B.1	Kinematic Subduction Model	156
B.1.1	Characterizing Slab Deflection	156
B.2	Comparisons with Various Dynamic Topography Models	159

List of Figures

1.1	The long-term carbon cycle system. A) Pre-industrial surface carbon reservoirs (red text) and exchanges (black text) between them. Carbon reservoir magnitudes are estimated from (1; 2) and processes are described in (1). Notably, deep ocean water contains ~ 8 times the ocean sediment carbon within the bioperturbed layer. However, the lithified carbon reservoir —that is often merged with land and ocean lithospheric reservoirs —is much larger. At pre-industrial accumulation rates, approximately 150,000 GtC of sediment would lithify into oceanic lithosphere every 1 Myr. B) Periodicity vs relative variance of long-term evolution of Earth’s surface temperature (3). Top color bars show internal and external forcings. Orbital parameters are labeled by letters: eccentricity (e), obliquity (o), and precession (p). Volcanism and impacts are aperiodic forcings of highly variable strength. Responses to these forcings and variability arising from Earth internal dynamics shown by filled color curves. These events determine multi-scale planetary climate and habitability.	3
1.2	Cenozoic CCD evolution from regional (Pälike et al. (4)) and global (Delaney & Boyle (5) and Boss & Wilkinson (6)). The curves are constructed from measurement of wt% $CaCO_3$ in ocean drill-site cores. Variations between the curves are partially due to the spatial extent and quantity of data available and selected for reconstructions. The use of proxies such as the CCD is integral for understanding the past climate and tuning Earth system models.	8
1.3	A planetary cooling regime diagram from Lyu et al. (2025) (7), showing mobile lid, episodic lid, stagnant lid, episodic squishy lid, plutonic-squishy lid, and sluggish lid as a function of internal heating, core mantle boundary temperature, and surface yield stress. Regimes are defined by a simulation’s period in mobility (blue), sluggish/squishy (green), and stagnant (red) base regimes.	10
1.4	Summary of suggested times for the onset of plate tectonics based on metamorphic, geomorphic, and geochemical lines of evidence, from the Korenaga 2013 annual review (8).	12
1.5	Global topography map highlighted with the major 1st and 2nd order bathymetry features and processes that shaped modern and paleo bathymetry. The topography model shown is ETOPO1 (9) and features are hand-drawn from high-relief features.	14

2.1	Evolution of 1st and 2nd order bathymetry. (<i>A</i> and <i>B</i>) Cartoon of seafloor morphology. Deep ocean bathymetry is represented by the oceanic lithosphere (rust red), while shallow bathymetry is represented by the continent (yellow) water interface. The brown line denotes the CCD, below which no carbon is sequestered. The light and dark blue interfaces represent eustatic sea-level (ESL). These cartoons show how slow and fast plate motions lead to increased or decreased thermal subsidence, respectively, before subduction, resulting in 1st order changes in ocean floor depth, sea-level, and the CCD. Bathymetry reconstructions at present-day (<i>C</i>) and 80 Ma (<i>D</i>) represent slow (<i>C</i>) and fast plate (<i>D</i>) motion regimes. (E) Kernel depth distributions for (<i>C</i> and <i>D</i>) denote two major components of bathymetry evolution: 1) changes to shallow seafloor (≤ 600 m) area and 2) changes to the distribution and mean deep bathymetry ($>1,000$ m).	22
2.2	Response of CCD given synthetic bathymetric distributions representing 1st order changes in deep bathymetry and shallow seafloor area. (<i>A-C</i>) LOSCAR bathymetry inputs shown in (<i>D</i>): 1) deep bathymetry represented through Gaussian distributions, and 2) shallow bathymetry areal extent ranging from 96 to 120% of its present value. Synthetics produce similar areal extents and average deep marine depths to those within the last 80 Myr. (<i>A-C</i>) columns and rows show different ocean basins and Earth system model setups, respectively. (<i>A</i>) Base model for the steady-state CCD (contours) obtained with LOSCAR modified to include our preferred bathymetry. No silicate/carbonate weathering feedback included. (<i>B</i> and <i>C</i>) Models with decreased and increased biological pump efficiency, respectively. Decreased biological pump efficiency reduces calcite rain, shallow sequestered alkalinity, and calcite erosion in the deep marine. It allows the Pacific and Indian basins to persist as depositional basins for greater shallow shelf areas.	26
2.3	Response of CCD given reconstructed bathymetry. Contours in (<i>A</i> and <i>B</i>) show the CCD for synthetic bathymetry models and dots show realistic reconstructions over the last 80 Myr for the Pacific and global ocean(s), respectively. Synthetic bathymetry modeled CCDs are plotted within reconstruction ranges. The saturation within the synthetic models indicates that CCDs calculated from realistic bathymetry reconstructions are subdued, but still lead to ~ 500 m variations which are exclusively bathymetric dependent. CCD evolution for the Pacific (<i>C</i>) and globally (<i>D</i>). We show (<i>C</i> and <i>D</i>) the total synthetic range shown throughout the reconstruction period. The blue and red curves represent the Lyle et al. (10) global CCD and Pälike et al. (4) equatorial CCD, respectively. Notably, reconstructions show the CCD shoaling between 65 and 50 Ma and deepening between 30 Ma and present-day. The major CCD deepening between 50 and 40 Ma occurred concurrently with bathymetric changes resulting from the India-Asia collision.	29

2.4	Discrepancy of bathymetry influenced CCD as a function of riverine flux and ocean alkalinity. (A) Default LOSCAR bathymetry model (Left) and updated 60 Ma bathymetry reconstruction (Right). Perturbed steady-state CCD and ocean alkalinity using our updated bathymetry model and a shelf-deep marine sediment partition coefficient half its default LOSCAR value ($FSHLF = 2.25$). (B) Global CCD with updated bathymetry and varying the degree of initial $CaCO_3$ riverine flux. Blue, red, and black stars represent initial riverine fluxes corresponding to the default LOSCAR values, LOSCAR model CCD reproduction, and calibrated LOSCAR model alkalinity reproduction. The red line is a sediment core-derived global CCD with no specific riverine flux [Lyle et al. (10)]. (C) Corresponding ocean alkalinity throughout basin water columns (Fig. A.1) with the addition of LT and DT denoting the Tethys Ocean. The symbols as in B except for blue dots showing LOSCAR's default alkalinity.	31
3.1	Diagram of kinematic subduction from <i>slabmodellib</i> . <i>slabmodellib</i> uses plate motion modeled surface velocities, the age of subducted seafloor, the mantle viscosity structure, and parameters constraining the dynamics of common and uncommon slab geometries (penetrating, deflecting, and orphaning). Outputs include slab location and lithospheric properties that can be used in post processing calculations of the non-hydrostatic geoid anomaly and dynamic topography.	43
3.2	Full mantle volume rendering of subducted slablets at present-day with regional highlights under A) West Pacific, B) East Pacific, and C) ancient Tethys ocean. The slablets are shown as spheres scaled to their modeled volumes.	48
3.3	Amplitude of spherical harmonic expansions of the non-hydrostatic geoid and dynamic topography over the last 80 Myr and initiated 250 Ma. Panels A) and B) show the amplitudes of dynamic topography and the non-hydrostatic geoid anomaly, respectively, while panels C) and D) show those amplitudes normalized to their present-day values. Black lines show the zero value contour. Past dynamic topography had larger amplitudes at wavelengths greater than continental scales (≥ 6000 km). The geoid was generally higher amplitude down to 2000 km length scales. An exception occurred from 5-40 Ma, when subcontinental-scale amplitudes were smaller.	49
3.4	Top) Coastline length calculated with paleoDEM with and without the contribution of dynamic topography and the non-hydrostatic geoid anomaly. Bottom) Percent change in coastline length resulting from dynamic topography and non-hydrostatic geoid anomaly contributions. Solid Earth influences on coastline length tend to decrease coastline length during most of the past 70 Ma.	51

- 3.5 The effect of the non-hydrostatic geoid anomaly and dynamic topography on low topography and bathymetry throughout the last 80 Myr and binned in the following ranges [200,0), [0,-200), [-200, -1000), and [-1000, $depth_{max}$). **Top)** Binned global surface area in intervals as calculated from topography/bathymetry reconstructions. **Middle)** Changes in global surface area within each bin due to the contribution of the non-hydrostatic geoid anomaly and dynamic topography. If the change in surface area of all shown bins are summed then they will represent the change in surface above the low lands (>200 m). **Bottom)** Results from the middle panel shown as a function of age. 52
- 4.1 (a) Present-day and (b) 60 Ma reconstructed bathymetry from (11). Black lines indicate basin boundaries while question marks indicate unknown and assumed boundaries. (c) Example of intermediate volatile cycling box models (the Long-term Ocean-atmosphere-Sediment CARbon Reservoir model, LOSCAR) used to reconstruct carbon and sulfur perturbation history for the present-day 3 ocean basin system (12; 13). Atlantic, Pacific, and Indian ocean basins are defined as separate boxes with their own ocean chemistry, input and transport of chemical properties, and depositional carbonate zones. A basin's ocean area, volume, and shape of depositional surface are defined by the interior of a basin while communication between basins is a function of bathymetry at the boundaries. 60
- 4.2 Present-day defined ocean basins from (a) Limits of Oceans and Seas (14), (b) Longhurst 2010 (15), (c) Dalvand et al. 2025 (16), and (d) Dutkiewicz et al. 2021 (17) 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*. 62
- 4.3 Diagram of the main 3 components required for spatial community detections: 1) graph construction, 2) adjacent node connection strength calculation, 3) community detection algorithm. Each sub-category relies on user input that must be appropriate for their final application. ExoCycle is also packaged with some general post-processing capabilities. 65
- 4.4 Diagram of (a) *Louvain* and (b) *Girvan-Newman* community detection algorithms. The Louvain algorithm is a two stage process that assumes each node forms a community. In Stage 1 the algorithm groups communities (i.e., nodes at this stage) if they improve the networks quality function. Stage 2 merges nodes into new communities before Stage 1 iterates. Girvan-Newman community detection assumes one contiguous community and then iteratively cuts edges with the highest edge betweenness centrality metric until a desired community count is achieved. The resulting community structure can still be evaluated with a quality function. 70

4.5	Shows the construction of node edge weights using ETOPO1 seafloor bathymetry (9) sampled at 1 degree resolution. Distribution of (a) absolute value difference of adjacent nodes mirrored about 0 value, and (b) interquartile range filtered absolute value difference of adjacent nodes mirrored about 0 value. (c) shows (b) after a normal distribution quantile transformation. (d) Shows (c) distribution, shortened and shifted distribution, and the 1 - cumulative density functions (1-CDF) of each distribution. Weights are inversely proportional to radian separation and proportional to 1-CDF on the red curve for some given absolute value difference of adjacent nodes.	73
4.6	Illustration of our multi-field node edge weight construction for graphs. (a) Diagram of nodes and connections for 2 fields (red and blue) that do not necessarily overlap. Red and blue S1's represent 1 - the cumulative density function (CDF) that is calculated separately for fields 1 and 2, respectively. (b) 1 - CDF for the shortened and shifted weight distribution as calculated with the difference quartile transformation CDF weighting method. Options for merging fields include minimum, maximum, mean, and product of 1-CDF.	74
4.7	Composite basin detection using measured bathymetry. (a) ETOPO1 bathymetry (9) sampled at 1 degree resolution and used to construct equidistantly spaced spatial network (b). (c) graph network constructed using communities from an ensemble of Leiden community detections (d). (d-e) detected communities for Leiden and Louvain-Girvan-Newman composite algorithms. (f) detected communities after small basins post-process step. (g) Detected communities after Girvan-Newman and small basins post-process step. Detected communities capture basins as regions with high water column connectivity. The basin boundaries often occur near mid-ocean ridge axes, ocean plateaus and seamounts, continental rises, and at ocean gateways.	82
4.8	Figure shows one method to analyze community detection stability using ensemble averaged silhouette values. Basins were detected using Leiden SB-Reduction ($\gamma = 0.01$, ensemble size = varied) and DQT-CDF weighting (shortened = 5σ , shifted = 0.5σ) on 1 degree spatially resolved present-day measured bathymetry. (a) Sample of global node-wise calculated silhouette scores for an ensemble size of 50 and with No, Weak, Medium, and Strong structure divided at silhouette scores of 0.25, 0.5, and 0.7. (b) ensemble averaged silhouette values (red line) with minimum ensemble silhouette value and σ lower boundary (black lines) using 20 ensembles for ensemble sizes ranging from 1 to 50. Strong, good, and weak clustering regions are highlighted. (c) a close-up of the ensemble averaged silhouette values and when they stabilize at ensemble sizes of 30 (magenta)	84

4.9	Basin detection on (a) reconstructed present-day bathymetry (11), and (b) measured bathymetry sampled at 1 degree (9). (c) and (d) are transformed fields of (a) and (b) used to calculate node connection weights. (e) and (f) are basins detected using SB-Reduction ($\gamma = 0.01$, ensemble size = 50) and DQT-CDF weighting (shortened = 5σ , shifted = 0.5σ) on 1 degree spatially resolved (c) reconstructed bathymetry and (d) measured bathymetry present-day, respectively. Models have modularity (.88) and (.91) and silhouette (.80) and (.85) values for (e) and (f), respectively. Seafloor volcanism represented in sampled ETOPO1 helps to define the edge of basins. Basins from reconstructed bathymetry are more likely to share boundaries along mid-ocean-ridges. . . .	85
4.10	Basin partitioning evaluation metrics for Leiden (γ = varied), not SB-Reduction, detection on ETOPO1 (9) sampled at 1 degree resolution and using DQT-CDF weighting method (shortened 5σ , shifted 0.5σ). (a) Newman Girvan modularity that describes better community divisions with values closer to 1. (b) Number of basins that represent more than 0.5% of the total seafloor area. Basins divisions (c-e) and (f-h) correspond to ensemble and γ value in (a-b). Blue hatched regions define ensemble sizes where increases in ensemble size lead to visually different community structure (or potentially global average silhouette). Red hatched regions define community models with larger γ and many smaller communities as shown in (f-h).	87
4.11	Node reduction using Louvain and Leiden community detection algorithms. Input surface (a) (≤ 200 m), (d) intermediate (> 200 m and ≤ 1000 m), and (g) deep (> 1000 m) ocean potential temperatures averaged over 1994 (GLORYS12V1; (18)). (b,e,h) are Leiden and (c,f,i) are Louvain community detections for (a,d,g) inputs. Models have modularity (.85, .81, .84) and (.85, .81, .83) and silhouette (.96, .94, .92) and (.98, .97, .87) values for (b,e,h) and (c,f,i), respectively. Communities were detected using SB-Reduction ($\gamma = 0.01$, ensemble size = 50) and DQT-CDF weighting (shortened = 3σ , shifted = σ). Both Louvain and Leiden produce similar community boundaries for given ocean depths.	89
4.12	Multifield community detection. (a-b) Input surface (≤ 200 m depth) salinity and potential temperature averaged over 1994 (GLORYS12V1; (18)). (c-d) Difference quantile transformed fields used to determine spatial network weight connections. Communities detected using (e) minimum (strongest connections) and (f) mean of weights from DQT-CDF method (shortened = 3σ , shifted = σ) with a Leiden SB-Reduction ($\gamma = 0.01$, ensembleSize = 50). Models have modularity (.85) and (.79) and silhouette (.95) and (.97) values for (e) and (f), respectively.	91
4.13	Basin-community detection and bathymetry distributions using ETOPO1 (9) sampled at 1 degree resolution and with Leiden GN-SB-Reduction ($\gamma = 0.01$, ensemble size = 50, basin count = 4) and DQT-CDF weighting (shortened 5σ , shifted 0.5σ). (a) Atlantic, Indian, Mediterranean Sea, and Pacific basins and (c) their bathymetry distributions. (b) Connecting regions defined as seafloor within 444 km of boundaries, and (d) their distributions. The model has a modularity of .91 and silhouette value of .85.	92

4.14	Basin detection when using the Leiden ($\gamma=0.01$) and small basin merger community detection on ETOPO1 (9) with the DQT-CDF weighting (shortened 5σ , shifted 0.5σ) for sampled bathymetry and node grid spatial resolutions of (a) 8, (b) 4, (c) 2, and (d) 1 degree. Models (a-d) have a modularity of .70, .76, .84, .91, respectively. Higher spatial resolution models lead to more community detections with γ held constant. Spatial degree resolution 2 and 1 models capture enough bathymetry features that are important for representing basin-community structure.	95
4.15	Basin detection on ETOPO1 (9) deep bathymetry sampled at 1 degree resolution. (a) Deep bathymetry (>1000 m), and (b) basin-communities using Leiden SB-Reduction ($\gamma=0.01$, ensemble size = 50) with DQT-CDF weighting (shortened 5σ , shifted 0.5σ). The model has a modularity of .89 and silhouette value of .84. Basin boundaries largely run along shallow regions (e.g., MORs and seafloor volcanism), similarly to basin boundaries calculated from the full Etopo bathymetry field.	97
4.16	Basin detection on ETOPO1 (9) shallow bathymetry (<600 m) sampled at 0.3 degree resolution. (a) Shelf bathymetry (<600 m), and (b) global shelf partitions using Leiden SB-Reduction ($\gamma = 0.15$, ensemble size = 50) with DQT-CDF weighting (shortened 5σ , shifted 0.5σ) at 0.3 degree spatial resolution. γ is chosen to maximize global modularity to 0.98 and results in a silhouette value of 0.94. The small basin merger step is modified such that it merges communities of relative similar scale when compared with full ocean detections. (c) South East Asia shelf partitions identifying isolated regions corresponding to present-day and past river sediment supply from now submerged river systems that span throughout the Sunda Shelf (19; 20; 21).	98
4.17	Basin detection on reconstructed bathymetry. (a-c) Reconstructed present-day, 60 Ma, and 80 Ma bathymetry (Bogumil et al. 2024), respectively. (d-f) Basin detection for (a-c) fields when using the ensemble (50) Leiden ($\gamma=0.01$) and small basin merger community detection with the DQT-CDF method (shortened 5σ , shifted 0.5σ) at 1 degree spatial resolution. Models have modularity (.88, .87, .91) and silhouette (.80, .85, .82) values for (d-f), respectively. Changes to continental distribution (e.g., opening Atlantic), shelf space (e.g., epicontinental seas), mid-ocean ridge distribution directly influence the number of ocean basins that existed during CO_2 forcing events since the Late Cretaceous.	102
4.18	Ocean basins detected throughout the water column. Ocean potential temperature averaged over 1994 (GLORYS12V1; (18)) and averaged over the a) shallow (<200 m) and B) deep (>1000 m) water column. Basin detection (c-d) for (a-b) respectively. Basin detection used Leiden SB-Reduction ($\gamma = 0.01$, ensemble size = 50) with DQT-CDF weighting (shortened 3σ , shifted σ). Models have modularity (.85) and (.84) and silhouette (.96) and (.92) values for (c) and (d), respectively. Shallow basins reflect latitudinal changes in temperature near the equator. Deep ocean basins are primarily defined by continental division and larger-scale changes in temperature driven by ocean circulation.	104

4.19	A simple example of an alternative intermediate-complexity Earth system model constructed with modular frameworks (e.g., GEOCLIM7 (22) or ES-BMTK (23)). Multiple shelf regions feed into shallow and intermediate-deep boxes through surface currents and turbidity currents, respectively. The n^{th} box illustrates the 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. 4.13), sill depths (Table 4.4), and silhouette scores (Table 4.4 and Fig. 4.8a) can be used to heuristically determine the number of basins and strength of mixing/circulation for modular frameworks.	108
5.1	Basin detection on deep (>1000 m) reconstructed bathymetry. Ocean basins detected from reconstructed present-day bathymetry to 80 Ma at 10 Myr intervals (Bogumil et al. 2024, (11)) using the ensemble (50) Leiden ($\gamma=0.01$) and small-basin-merger community-detection method with DQT-CDF weighting (shortened 5σ , shifted 0.5σ) at 1-degree spatial resolution. All models have modularity and silhouette scores above 0.86 and 0.69, respectively. Table 5.1 reports all values, including the total number of basins, which varies from 15-23 throughout the reconstruction period. Symbols on reconstructed basins represent DSDP (red), ODP (black), and IODP (yellow) drill sites with present-day basin basin divisions: Pacific (dots), Atlantic (triangles), and Indian Ocean (X) from Bogumil et al. 2024 (11).	125
5.2	Proxy confidence (C) map for the detected basin shown in Figure 5.1. Confidence is calculated using equation 5.1. Symbols on reconstructed basins represent the collection of DSDP, ODP, and IODP drill sites. Unique symbol and color combinations distinguish sites in each basin cluster. White basins contain no drill sites. Darker basins (high C) generally have both more of the global drill site record and higher silhouette scores. Simple data assimilation schemes can overfit these regions when they do not account for the spatial nonuniformity of the record.	127
5.3	Planetary bathymetry created from (A) Earth, (B) Venus, (C) Mars, and (D) Moon when flooded with 2.6 (modern), 0.3, 1, and 1 GEL, respectively. Blue lines indicate coastlines and black lines indicate basin boundaries.	130
5.4	(A-B) Deep and shallow ocean diagrams, highlighting the expected change in atmospheric pCO_2 and ocean alkalinity. (C) Atmospheric pCO_2 from Bogumil et al. (2024) as a function of shallow seafloor area (< 600 m) and average deep-marine depth (> 1 km). (D) Changing bathymetry for increasingly flooded solar-system terrestrial-like planets, where larger symbols indicate larger GEL.	132
5.5	Stacked histogram of ocean basin sizes for planets flooded with various amounts of GEL.	133

A.1	Long-term Ocean-Sediment-Carbon cycle Reservoir (LOSCAR) model present-day setup. Pacific, Indian, and Atlantic oceans are represented with individual bathymetry (sediment layers) and consist of three reservoirs: Low latitude (≤ 100 m, L), interMediate (> 100 m and ≤ 1000 m, M), and deep (> 1000 m, D) boxes. The brown surface depicting sediment represents basin specific bathymetry in the lower latitudes (i.e., 90% of the lowest-latitude seafloor area). For this study we define < 600 m and > 600 m as shallow and deep bathymetry, respectively. The Early Paleogene setup shown in Figure 2.4 includes an additional 3 reservoirs for the Tethys ocean.	143
A.2	Percent change in input ocean volume vs. ocean volume used in LOSCAR calculations. There is effectively no change in volume for changes to seafloor area or average depth of deep marine. However, the distribution of ocean volume from intermediate and deep ocean boxes does change and is not shown here.	144
A.3	Base model (nuwd = 0.31 and FBIOL = 0.8) response to alkalinity given synthetic bathymetric distributions representing first order changes in deep bathymetry and shallow area extent. Deep bathymetry is represented through Gaussian distributions within an Earth system model, with area averaged values plotted above. Shallow bathymetry areal extent is increased incrementally from 96 to 120% of its present value, producing similar areal extents to those within the last 80 Myr. Figure columns and rows represent ocean basins and basin box depths (L,M,D) from LOSCAR, respectively for base model setup.	145
A.4	Base model (nuwd = 0.31 and FBIOL = 0.8) response to calcite saturation state given synthetic bathymetric distributions representing first order changes in deep bathymetry and shallow area extent. Axes as described in Figure A.3.	146
A.5	MinS, PD, and MaxS represent LOSCAR model steady state runs with average deep bathymetry of 4 km and minimum ($3e6$ km ²), present-day ($\sim 30e6$ km ²), and maximum ($\sim 78e6$ km ²) shelf areas for base model (nuwd = 0.31 and FBIOL = 0.8). The 3.5 km model represents average deep bathymetry of 3.5 km with present-day ($\sim 30e6$ km ²) shelf area. Lines represent rain, dissolution, and burial rates in m/y while values below represent rates in kg CaCO ₃ /yr. For bottom values, negative erosion represents burial while negative dissolution represents dissolution occurring at the sediment layer interface. Shallow and deep values represent rates summed < 600 m and ≥ 600 m, respectively. Note that total burial (15.3×10^{12} kg/y) represents source rates from both silicate (4.5×10^{12} molC/yr) and carbonate riverine (10.8×10^{12} molC/yr) fluxes; summing fluxes and converting to kg CaCO ₃ /yr results in a flux of 1.53×10^{12} kg CaCO ₃ /y). The MaxS bathymetry model forces the Pacific into a dissolution only regime.	147
A.6	Alkalinity response for the model with a 10% decrease in biological pump efficiency (nuwd = 0.31 and FBIOL = 0.72) and synthetic bathymetric distributions described in the main text. Axes as described in Figure A.3.	148
A.7	Calcite saturation state response for the model with 10% decrease in biological pump efficiency (nuwd = 0.31 and FBIOL = 0.72) and synthetic bathymetric distributions described in the mian text. Axes as described in Figure A.3.	148

A.8	MinS, PD, MaxS, and the 3.5 km LOSCAR model steady state runs for bathymetry described in Figure A.5. LOSCAR setups are similar to the base model (Fig. A.3, A.4, A.5) except for a decrease in the biological efficiency of 10% ($\text{nuwd} = 0.31$ and $\text{FBIOL} = 0.72$). Bottom rows of values are similar to the base models (Fig. A.5) with the exception of decreased calcite rain and deep marine dissolution. The MaxS bathymetry model forces the Pacific into a dissolution only regime, leading to high CCD sensitivity for changing shallow bathymetry area (Fig. 2.2).	149
A.9	Alkalinity response for the model with increased water column dissolution ($\text{nuwd} = 0.5$ and $\text{FBIOL} = 0.8$) and synthetic bathymetric distributions described in the main text. Axes as described in Figure A.3.	150
A.10	Calcite saturation state response for the model with increased water column dissolution ($\text{nuwd} = 0.5$ and $\text{FBIOL} = 0.8$) and synthetic bathymetric distributions described in the main text. Axes as described in Figure A.3.	150
A.11	MinS, PD, MaxS, and the 3.5 km LOSCAR model steady state runs for bathymetry described in Figure A.5. LOSCAR setups are similar to the base model (Fig. A.3, A.4, A.5) except for an increase in water column dissolution ($\text{nuwd} = 0.5$ and $\text{FBIOL} = 0.8$). Bottom rows of values are similar to the base models (Fig. A.5) with the exception of decreased calcite rain and deep marine dissolution greater than that in the decreased biological pump efficiency models ($\text{nuwd} = 0.31$ and $\text{FBIO} = 0.72$). MaxS does not enter into a dissolution regime in any ocean basin as it does in the other two ocean systems (Fig. A.5 and A.8).	151
A.12	Global ocean volume corrections applied to bathymetry distributions. A vector of correction factors for depths is calculated by taking the difference between Etopo and our present-day reconstructed bathymetry model distribution. Correction factors for each depth bin are proportionally applied to past and present reconstructions such that reconstructed distributions will result in realistic ocean basin volumes. Red lines represent the original distribution and blue lines the ocean basin volume corrected distribution. (Left) present-day. (Right) 60 Ma. (Top) Result of redistributing bathymetry sampled at ~ 20 m bin intervals. (Bottom) Same as (Top) using the LOSCAR sediment bin intervals. Bathymetry reconstructions tend to overestimate the global volume of ocean basins. Applied corrections shallow deep bathymetry while barely affecting shallow bathymetry.	152
A.13	Cumulative reconstructed low-latitude bathymetry area evolution throughout the last 80 Myr. Low-latitude refers to the lowest 90% of seafloor area and describes the area where bathymetry distributions inform LOSCAR. While the area of deep bathymetry (>600 m depth) is relatively constant, shallow area (<600 m depth) changes significantly due to continental flooding.	153
A.14	Evolving distribution of global basin system low-latitude seafloor area for LOSCAR depth boxes (L, M, and D, see Fig. A.1). Calculations as in A.13.	153

A.15	Evolution of Atlantic and Pacific basin ocean chemistry of base model (Fig. A.3, A.4, A.5) with (Left) and without (Right) weathering feedbacks. Blue and Red dots and lines represent models of smaller shelf area values for average deep and shallow deep marine, respectively. Green and black dots and lines represent models of larger shelf area for average deep and shallow deep marine, respectively. Models with disabled weathering feedback all experience an increase in atmospheric carbon and decrease in shallow and deep marine alkalinity and loss of deep Pacific carbonates. Weathering feedback enabled models experienced the same loss in deep Pacific carbonates. Initially, deep and shallow marine alkalinity increased and $p\text{CO}_2$ climbed. As $p\text{CO}_2$ relaxed to present-day values (280 ppm) through temporary increased weathering, silicate weathering captured additional carbon, reducing deep and shallow ocean alkalinity further.	154
A.16	Seafloor carbonate concentration steady-states for default LOSCAR sediment layer depths and basins (see Fig. A.1). Subplots axes as in Figure A.3. Contours in each subplot represent low (blue) and high (red) CaCO_3 sediment concentrations for varied bathymetry. Deep bathymetry with a mean of 4.5 km only has seafloor and the potential for carbonate sequestration within the shallow marine and between 3.5-5.5 km depth. For this deep bathymetry model, the CCD kink (red and black arrows), described in Figure A.5 is clear. Increasing shallow seafloor leads to an increased shelf carbonate supply and a lack of deep Pacific carbon.	155
A.17	Basin boundary evolution at 0, 20, 60, and 80 Ma. We found no community defined systematic method for defining paleo basin boundaries. Therefore, we drew basin boundaries along bathymetry informed chokepoints, which communicate ocean waters between basins. While the Tethys ocean basin is described during its existence as a separate basin, the 3 basin models merge the Tethys and Indian oceans to describe a single Indian ocean basin (Fig. 2.2 and 2.3).	155
B.1	Example of slab deflection look-up table for slabs that deflect off the ringwoodite-bridgmanite phase transition at ~ 660 km depth. Left) period of slab deflection as a function of the radial rate of subduction in the upper mantle and age of the subducted ocean lithosphere. The red line indicates the boundary between deflected and non-deflected slabs. The black lines correspond to the profiles plotted in the right panel. Right) Slab deflection-period profiles as a function of radial subduction rate in the upper mantle.	157
B.2	The effect of dynamic topography on low topography and bathymetry throughout the last 80 Myr, using the bins and axes described in Figure 3.5. Top) Topography/bathymetry reconstructions provide the global surface area within each bin. Middle) Change in global surface area within each bin due to the contribution from the M1 dynamic topography model of (24). Bottom) The bottom panel plots the middle-panel results as a function of age.	160

B.3	The effect of dynamic topography on low topography and bathymetry throughout the last 80 Myr, using the bins and axes described in Figure 3.5. Top) Topography/bathymetry reconstructions provide the global surface area within each bin. Middle) Change in global surface area within each bin due to the contribution from the M7 dynamic topography model of (24). Bottom) The bottom panel plots the middle-panel results as a function of age.	161
-----	---	-----

Acknowledgements

Chapter 2 and 4 and Appendix 2 are versions of my published and in-review works with Carolina Lithgow-Bertelloni and Tushar Mittal (11; 25).

First and foremost, I want to express my gratitude for my advisor Carolina Lithgow-Bertelloni. Her experience and dedication to her students has been invaluable to my growth as a scientist. She has helped me to develop a balance between scientific rigor and conveying research ideas which has fostered my writing an presentation style to be both impactful and welcoming. Furthermore, she has helped me to navigate toward a potential future career in academia. I will miss our chats, group journal clubs, and mock presentations as a graduate student. But I, excitedly, look forward to our future collaborations and discoveries!

I would also like to thank my committee members. *Lars Stixrude* was my first instructor in this program. His clarity in conveying complex concepts in a measured and methodical way has been a valuable lesson in "less is more." His expertise in mineral physics and extension of experimental to theoretical mantle phase assemblages and proportions has directly helped to improve the subduction simulation framework presented in this dissertation. Craig Manning's experience in subduction system and the carbon cycle as well as Caroline Beghein's experience with seismic tomography provides a diverse set of feedback on the wide-breadth of applications for this and future work.

I also sincerely acknowledge my collaborator, Tushar Mittal, for helping to foster my knowledge of the carbon cycle, maintain a health dose of skepticism when evaluating my and others' work, and in navigating a potential future career in academia.

Throughout my PhD program I have met many friends within my lab and department. These friends have provided stability and academic advice throughout my time at UCLA. I would like to thank my lab mates: Xiyuan Bao, Boontigan Kuhasubpasin, Xuesong Ding, Joy Lewis-Merrill, Lan Hu, Lingxi (Shania) Liu, Madelenie Kerr. And I would like to thank the friends that I have made within the department: Leslie Insixiengmay, David James (DJ), Travis Gilmore, Francis Dragulet, Bing Hong Chua, Kyle Webster, Lizet Casillas, Sarah Preston, Jade Knighton, and Zijin Zhang.

Rock climbing, which I have taken-up in the last two years, has been an invaluable mental outlet and has greatly improved my focus while conducting this research. Instead of taking a walk to clear my head and tackle a problem, I climb. I also want to graciously thank my rock climbing community for creating a supportive environment throughout the final two years of my program. Specifically, I want to thank Paul Rickman, Andrew Morris, Shania (Lingxi) Liu, Sarah Preston, Zijin Zhang, Kyle Webster, Emma Landry, Tim Lawrence, Cody England. I also thank each instant friend I make when tackling a new climbing problem, and I hope that I can extend such collaboration into my scientific community.

Throughout my program I had experienced administrative hurdles that were easily resolved by a collection of invaluable EPSS department staff. I would like to specifically acknowledge Laurie Holbrook, Tasha Taylor, Eric Wessenauer, Evgenia Grigorova, Sean Silverstein, Bill Harris, Nick Baerg, James Diaz, Kelli Yang, Jessica Ban, Elyse Chou, Carlene Brown, and Nanette Capulong for the support they have offered throughout my roles at UCLA: student, student organization committee member, and teaching assistant.

I would also like to acknowledge my teachers and advisors that prepared me for or supported me throughout my PhD: Paul Davis (advisor/teacher), Mackenzie Day (advisor), and Bill Schopf (mentor), Lars Stixrude (teacher), Xiaolin Zhong (teacher), Seulgi Moon (teacher), Kunihiro "Sam" Taira (teacher), Tina Treude (teacher), Pirouz Kavehpour (teacher), Jacob Bortnik (teacher), and Ed Young (teacher).

I also thank the students who I was able to mentor throughout my program which allowed me to grow as a mentor. Specifically, I thank Emily Welsh, Raj Hamlai, and Veronica Chandler.

My parents and siblings (all older) have significantly supported higher education throughout my upbringing and paved a path for my success. I thank them, and I further thank my mother who encouraged me to keep going when bridging the gap between homeschooling and my higher education that started at Santa Monica Community College.

Finally, I want to thank my partner Lucie Bosler for her love, patience, and support. Without her, life would not sparkle and the quality of this body of work and my enjoyment creating it would be less.

Vita

Education

2016 - 2019 B.S. Geology, University of California, Los Angeles.

2019 - 2022 M.S. Geology, University of California, Los Angeles.

Grants & Awards

NASA F.5 FINESST 23-PLANET23-0141, Multiscale Planetary Habitability: Analyzing the Role of Large Volcanic Eruptions and Seafloor Bathymetry, 2024.

Graduate Research Award, UCLA Earth Planetary and Space Sciences Department, 2024.

Publications

M. Bogumil, T. Mittal, and C. Lithgow-Bertelloni, 2024, The effects of bathymetry on the long-term carbon cycle and CCD, Proceedings of the National Academy of Sciences. DOI: 10.1073/pnas.2400232121.

M. Bogumil, C. Lithgow-Bertelloni, T. Mittal, 2025, ExoCcycle v1.0.0: A Generalized Framework for Spherical Community Detection and its Application to Defining Global Ocean Basins from Multi-Field Data, Geoscientific Model Development, DOI: 10.5194/egusphere-2025-5581 (In Review).

M. Bogumil, C. Lithgow-Bertelloni, 2026, The Influences of Dynamic Topography and Geoid Evolution on the Long-Term Carbon Cycle, (Manuscript in preparation).

M. Bogumil, T. Mittal, C. Lithgow-Bertelloni, 2026, Objective Ocean Basin and their Chemical Footprint Throughout Cenozoic, (Manuscript in preparation).

M. Bogumil, C. Lithgow-Bertelloni, T. Mittal, 2026, Exo-ocean basin systems, their properties, and influence on long-term carbon cycling, (Manuscript in preparation).

Chapter 1

Introduction

Long-term planetary habitability is deeply tied to the evolution and dynamics of the solid planet, from the emergence of land to the recycling of volatiles via plate tectonics and subduction. Yet most studies of long-term habitability and climate ignore or underplay the role of the solid Earth. Two justifications are typically motivate this omission: one scientific, that the timescales of change are vastly different, and one technical, that coupling models across those timescales is difficult. Neither is fully valid. Solid Earth evolution, while slower than climatic processes, couples with ocean and atmospheric chemistry to produce profound impacts on the system. The technical challenge can also be addressed by identifying the most important coupling pathways and working through reduced (or intermediate) order models.

Earth's long-term carbon cycle (LTCC), the climatic thermostat that regulates planetary temperature on million-year timescales, is a case in point. Solid Earth evolution shapes the LTCC through at least three pathways: it controls the depth distribution of the ocean floor and thus where carbon can be stored in the deep marine realm (2); it modifies coastline length and shallow seafloor area, setting the extent of carbonate platform environments (3); and it determines circulation pathways between chemically distinct ocean masses, which in turn governs how proxy data are aggregated and interpreted across Earth's history 4, 5).

This dissertation leverages reduced-order modeling to characterize and interrogate these solid-Earth controls on the ocean-atmosphere system. Such models are invaluable for paleo-climate science because they can simulate the LTCC while integrating carbon cycle processes whose timescales span multiple orders of magnitude. Recent advances in plate motion reconstructions, bathymetric modeling, graph-based community detection, and modular Earth system Models of Intermediate Complexity (EMIC) frameworks for both paleo-Earth and exoplanet applications make this work both technically feasible and timely.

1.1. Carbon Reservoirs | Planetary habitability, requiring at least the presence of stable liquid surface water (3; 26; 27; 28), is not necessarily a persistent state (29). The amount and evolution of surface water, together with topography-forming processes, have shaped Earth’s oceans and potentially those of Mars and exoplanets. Earth’s ability to sustain habitability is often attributed to the carbon-silicate weathering feedback that stabilizes atmospheric CO_2 , a greenhouse gas. Major greenhouse gases (CO_2 , methane (CH_4), and water vapor) collectively increase atmospheric temperatures, with CO_2 and methane playing particularly important climate-driving roles on Earth. A runaway greenhouse effect may result from a positive heating feedback: increased atmospheric CO_2 heating $\rightarrow$ increased ocean evaporation (H_2O) $\rightarrow$ decreased H_2O and CO_2 stability. Conversely, sulfur aerosols contribute to planetary cooling that could stabilize atmospheric temperatures depending on coupled ocean-atmospheric and eruption characteristics (i.e., rate, volume, source, and submarine or aerial flux). One of the major components for understanding the atmospheric exchange of carbon, sulfur, and methane is the ability of ocean systems to take up and retain them.

The division of carbon reservoirs in the pre-industrial Earth provides insight into carbon reservoir balances throughout geological time. These reservoirs fall into two major groups as shown in Figure 1.1A: 1) *surface carbon* (ocean water, atmosphere, and land biosphere) and 2) *geologic carbon* (rock formations and sediments). Figure 1.1B highlights the timescales for

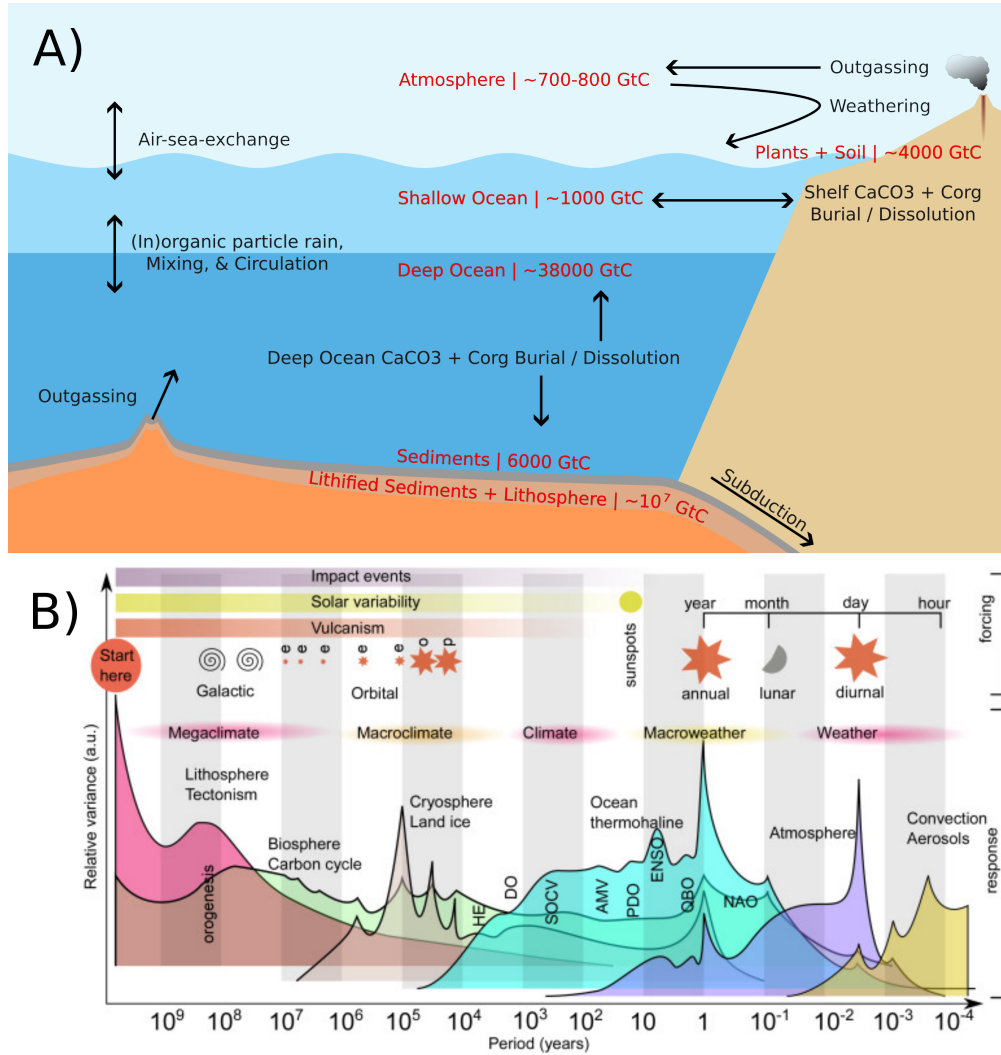


Figure 1.1: The long-term carbon cycle system. A) Pre-industrial surface carbon reservoirs (red text) and exchanges (black text) between them. Carbon reservoir magnitudes are estimated from (1; 2) and processes are described in (1). Notably, deep ocean water contains ~ 8 times the ocean sediment carbon within the bioperturbed layer. However, the lithified carbon reservoir—that is often merged with land and ocean lithospheric reservoirs—is much larger. At pre-industrial accumulation rates, approximately 150,000 GtC of sediment would lithify into oceanic lithosphere every 1 Myr. B) Periodicity vs relative variance of long-term evolution of Earth's surface temperature (3). Top color bars show internal and external forcings. Orbital parameters are labeled by letters: eccentricity (e), obliquity (o), and precession (p). Vulcanism and impacts are aperiodic forcings of highly variable strength. Responses to these forcings and variability arising from Earth internal dynamics shown by filled color curves. These events determine multi-scale planetary climate and habitability.

various climate forcing processes and reservoirs. A few specific reservoirs and processes that affect the long-term carbon cycle are highlighted here for reference. Surface reservoirs are in quasi-equilibrium on timescales of $1 - 10^2$ yr (e.g., Ekman transport and air-sea exchange), $10^2 - 10^3$ yr (e.g., (in)organic particle rain, mixing, and circulation). Geologic reservoirs are in quasi-equilibrium on longer timescales $10^2 - 10^4$ yr (carbonate weathering), $10^5 - 10^6$ yr (silicate weathering), and $10^6 - 10^7$ yr (subduction). The largest carbon reservoir accessible to the hydrosphere on these timescales comprises seafloor sediments and the underlying lithified carbonates, primarily calcite and aragonite.

At present-day, modeled and measured ocean chemistry (18) and carbonate sediment erosion (30; 31) vary across and within major ocean basins (Atlantic, Indian, Pacific). These different conditions reflect the efficiency of geological carbon storage (i.e., in seafloor carbonate rocks and sediment) and the susceptibility of stored carbon to erosion and dissolution during quasi-steady-state conditions and external forcing (e.g., volcanism). We know that carbon cycle parameters have evolved at the basin scale throughout Earth’s history (e.g., (32) and references therein; and more recently (16; 17; 33; 34)), and that the basins themselves have evolved significantly. For Earth, the primary process that evolves the ocean’s surface area distribution —not necessarily entire basin structure —is the supercontinent cycle (35), partially a consequence of planetary cooling. There is now a plethora of evidence for basin evolution through the destruction and formation of entire oceans (e.g., the Western Interior Seaway (36) and the North and South Atlantic Ocean, Tethys Ocean, Panthalassa Ocean summarized by plate motion reconstruction (e.g., (37; 38; 39))), as well as through the merging or connection of basins (e.g., Central American Seaway and Drake Passage). Although it is well-accepted that ocean basins, their chemistry, and seafloor have evolved significantly throughout Earth’s history, the consequences of their evolution on the long-term carbon cycle have not been quantified.

Solid Earth evolution changes 1) the depth distribution available for deep ocean carbon storage, 2) the coastline length and shallow seafloor area that describe the extent of carbonate

platform environments, and 3) circulation pathways between chemically distinct ocean masses thus how data are aggregated to interrogate carbon cycle processes throughout Earth’s history.

In this thesis, the goal is to elucidate the role of solid planetary evolution on the LTCC. I do this by investigating the role that Earth’s largest hydrosphere-accessible carbon reservoir (seafloor and seafloor sediment) plays in the LTCC. I quantify the effects of evolving seafloor depth distributions and the dynamic support of shallow seas and coastlines on the LTCC and develop an objective basin detection framework for proxy aggregation and LTCC model development. My primary focus is on Earth’s recent past (the last ~ 80 myr), for which well-established and robust diagnostics (models and data) exists. However, the theory developed in this thesis has implications for interrogating volatile cycling in deep time and on terrestrial exoplanets and for identifying geological conditions conducive for long-term habitability. In the remainder of this introduction, I briefly describe the processes that govern the LTCC and carbon cycle proxies and geological records used to reconstruct it, Earth in the context of planetary cooling regimes, the evolution of plate tectonics, and the processes that influence topographic and bathymetric evolution.

1.2. Long-Term Carbon Cycle | Processes | The long-term carbon cycle (Fig. 1.1A) is a balance between carbon sources (i.e., volcanic emissions from continental and oceanic arcs, mid-ocean ridges, and hotspot volcanism) and sinks (i.e., carbon burial by carbonate-bearing minerals on the seafloor, e.g., (40; 41; 42)).

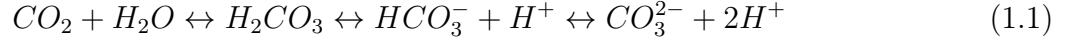
The transfer of carbon between the solid Earth, ocean, and atmosphere is governed by a set of chemical reactions (summarized with equations 1.1-1.5). Ocean-atmospheric carbon exchange will transfer inorganic carbon between the atmosphere and the surface ocean (i.e. Ekman transport layer) on shorter timescales than that of the deep ocean with a circulation timescale of 10^3 yr (Fig. 1.1). Within the ocean, both globally and locally (e.g., shallow versus deep environments), carbon system speciation (eq. 1.1) equilibrates in response to

seawater pH, which is partly controlled by the atmospheric partial pressure of CO_2 (pCO_2). The organic carbon storage rate in the ocean and on land reflects the balance between 1) net primary productivity by photosynthesizing organisms, which converts inorganic carbon into solid organic carbon (CH_2O), and 2) respiration by heterotrophic organisms, which returns inorganic CO_2 to the ocean-atmosphere system 1.2. Carbonate particles ($CaCO_3$) can form from precipitation in saturated ocean water or through biological production of organisms' hard parts. Both inorganic ($CaCO_3$) and organic (CH_2O) particulate carbon sink through the water column toward the seafloor. These processes constitute the ocean's biological pump, whose efficiency reflects the balance between solid-carbon export and preservation at depth and the return of carbon to surface waters.

The total flux of cations (e.g., Ca^{2+} , Mg^{2+}) provided by silicate weathering (eq. 1.4 and (42; 43; 44)) is the primary negative feedback that forms a quasi-steady state with volcanic degassing fluxes on a geological timescale. The Earth's ocean chemistry and dynamics (e.g., ocean mixing and circulation) determine how rapidly the system reaches steady state. On timescales shorter than the long-term steady-state timescale ($< \tau_{10-100kyr}$), the silicate weathering feedback is ineffective at drawing down CO_2 . Instead, *ocean carbonate chemistry & dynamics and associated feedbacks* will primarily control atmospheric CO_2 and thus global climate (45; 46). Ultimately, silicate weathering draws down pCO_2 through the production of inorganic carbonate solids (e.g., calcite, aragonite, and magnesium carbonate depending on ocean conditions) that reside on the seafloor. Deep-ocean carbon-bearing solids can then enter a subduction zone if they survive dissolution during perturbations such as volcanic forcing and subsequent ocean acidification, as well as alteration associated with reverse weathering (47). Beyond this point, the pathways and ultimate fate of stored carbon become increasingly uncertain: metamorphism in accretionary wedges can release CO_2 back to the surface (e.g., eq. 1.5) or mantle storage can occur with the formation of graphite and diamonds (48). Subduction characteristics (e.g., ocean lithosphere hydration, subduction dip/curvature, and convergence speed (43) and references therein) determine the efficacy of carbon transfer into

the mantle and continental crust versus outgassing to the atmosphere. At present, the carbon entering subduction zones varies both in its organic-to-inorganic composition and spatially along subduction zones, even when normalized by trench length (48).

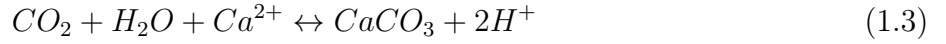
Atmospheric-Ocean Exchange & Speciation:



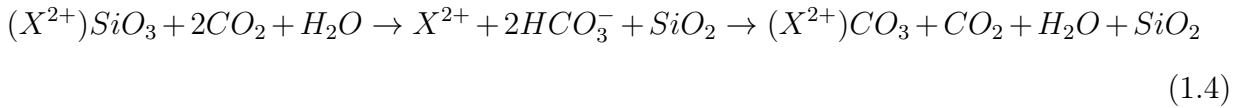
Photosynthesis & Respiration:



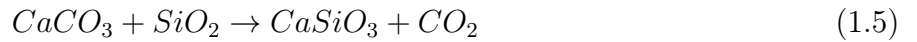
Carbonate Precipitation & Dissolution:



Silicate Weathering:



Metamorphism:



The coupling of processes (eqs. 1.1-1.5) acting on the atmosphere, ocean, and solid Earth shapes the modern carbon reservoirs (Fig. 1.1A). The distribution and flux of carbon between reservoirs in the past is not always directly measurable, forcing interpretations to rely on Earth system models that track measurable carbon cycle proxies.

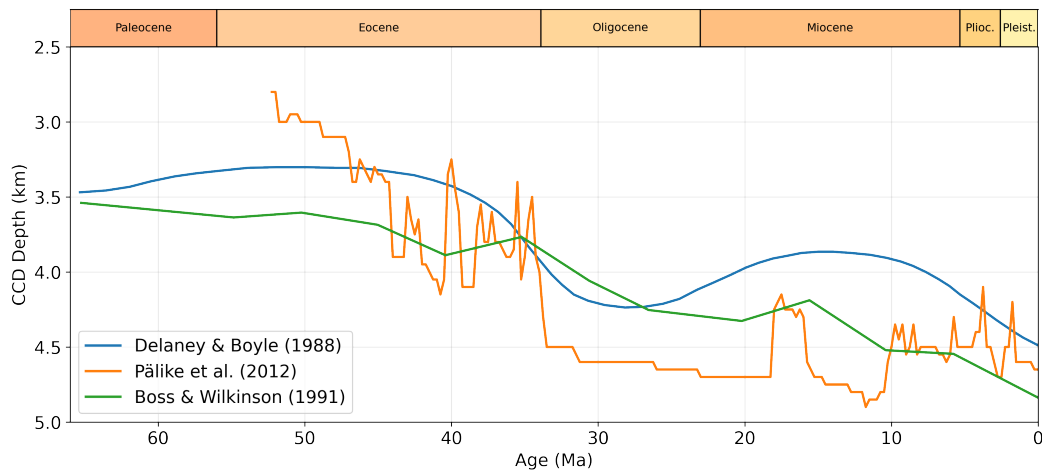


Figure 1.2: Cenozoic CCD evolution from regional (Pälke et al. (4)) and global (Delaney & Boyle (5) and Boss & Wilkinson (6)). The curves are constructed from measurement of wt% $CaCO_3$ in ocean drill-site cores. Variations between the curves are partially due to the spatial extent and quantity of data available and selected for reconstructions. The use of proxies such as the CCD is integral for understanding the past climate and tuning Earth system models.

1.3. Carbon Cycle Proxies and Their Spatial Interpretation | It is useful to divide carbon-cycle proxies into two categories: directly measured quantities (e.g., isotopic and carbonate measurements) and reconstructed quantities (e.g., pCO_2) inferred by inverting relationships between measurements and the carbonate system or biological responses to CO_2 . Although the amount of atmospheric C is not always a directly measurable proxy, it is often used this way in Earth system model simulations. Therefore, I will refer to both measured and inverted quantities as proxies for simplicity.

A non-comprehensive list of carbon-cycle proxies, their source measurements, and their diagnostic contributions to the LTCC is described below. This list also provides a sense of the common proxies used to constrain carbon cycle analyses (e.g., model simulations). $\delta^{13}C$

measures organic versus inorganic preservation calculated from carbonate and organic matter samples. $\delta^{18}O$ is used as a temperature proxy (defining global mean surface temperature and glaciation) and is calculated from carbonate samples. The Sr-K-Mg-Li-Os isotope systems are used to reconstruct continental rock weathering rates from samples including igneous, metamorphic, carbonate, and sedimentary rocks (49). The $^{143}Nd/^{144}Nd$ ratio, often reported as ϵ_{Nd} when compared to a standard composition, describes the ocean circulation and the age of seawater (e.g., (50; 51)). It can be calculated from ferromanganese crusts and sediment samples. Boron can be used to estimate ocean CO_3^{2-} ions (1) and is determined from carbonate (e.g., foraminifera) samples. Ocean alkalinity (ALK) is a measure of alkalinity ions and DIC is a measure of dissolved inorganic carbon (1; 45). Both can be inferred from carbonate proxy systems. pH is a measure of ocean acidity and can be calculated from carbonate (e.g., boron isotope) samples. pCO_2 is a proxy for weathering and global mean surface temperatures and can be determined from a set of proxy systems (e.g., alkenones, boron, isotopes, and stomatal indices). The carbonate compensation depth (CCD) indicates the maximum depth of carbonate deposition in the ocean and is determined by wt% carbonate in seafloor sediments. Some of these proxies are representative of global averages (e.g., pCO_2) while others leverage spatial variations (e.g., ϵ_{Nd} and CCD) to describe the inhomogeneous ocean chemical system that moderates the global LTCC. Proxies that sample the spatial inhomogeneity of the ocean are crucial for understanding the ocean carbon distribution and climate vulnerability.

Figure 1.2 shows Cenozoic CCD proxies calculated from a collection of seafloor sediment cores distributed throughout all modern-day major ocean basins (Atlantic, Indian, and Pacific). Climate proxies capture a fragmented record of paleoclimate evolution that modelers link to simulated processes; consequently, how these records are spatially aggregated strongly influences interpretations of the LTCC. However, both the seafloor that stores the vast majority of surface carbon and the geological records of catastrophic and long-term changes in volcanic forcings ultimately reflect planetary cooling expressed through plate tectonics.

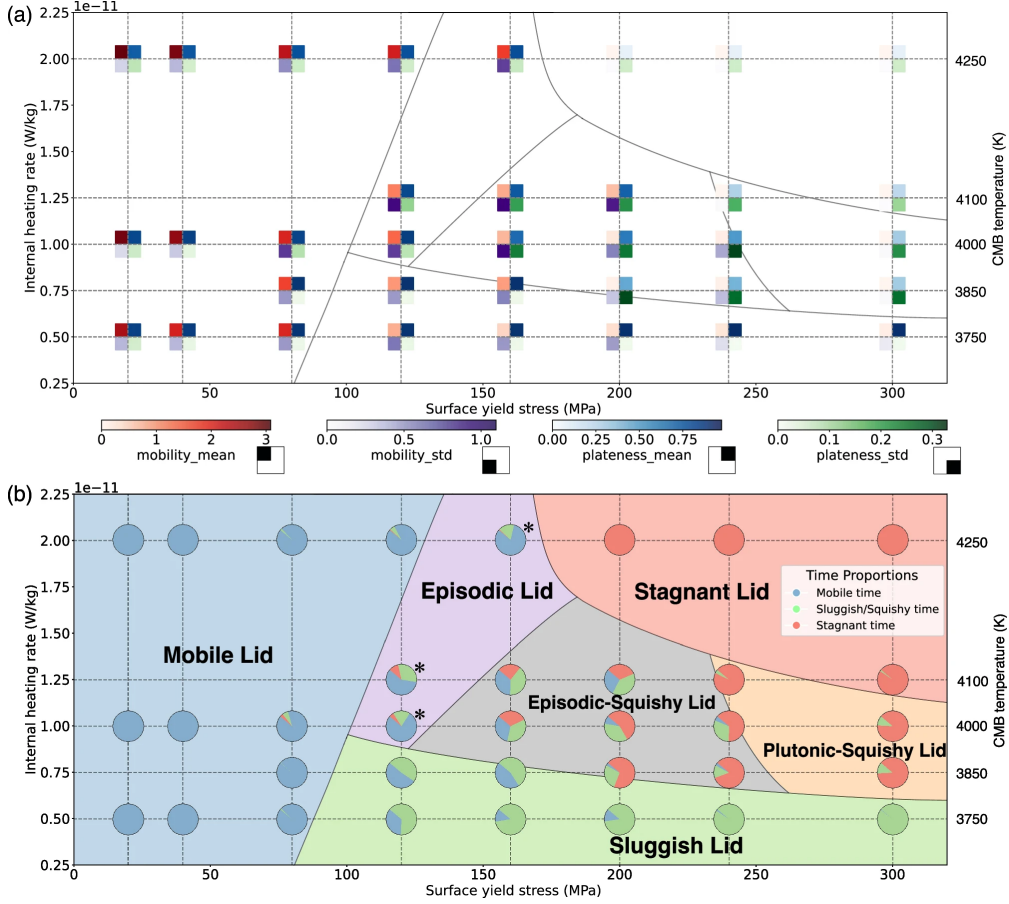


Figure 1.3: A planetary cooling regime diagram from Lyu et al. (2025) (7), showing mobile lid, episodic lid, stagnant lid, episodic squishy lid, plutonic-squishy lid, and sluggish lid as a function of internal heating, core mantle boundary temperature, and surface yield stress. Regimes are defined by a simulation’s period in mobility (blue), sluggish/squishy (green), and stagnant (red) base regimes.

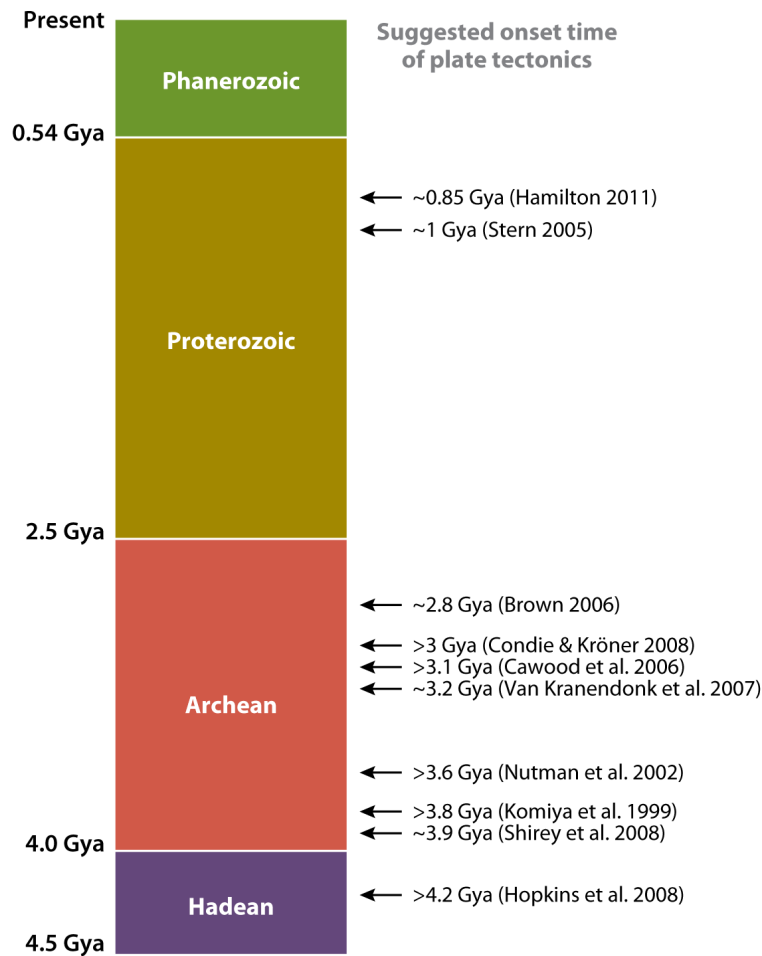
1.4. Planetary Cooling Regimes | Processes governed by planetary cooling are the primary control on volatile cycling in rocky planets such as Earth. Internal heating rate, surface yield stress, and core-mantle boundary temperature control these cooling regimes (e.g., (7)). The initial heat budget of a rocky planet provides a starting point for understanding the evolution of these parameters.

The initial heat budget of a rocky planet is the result of accretion, radioactive decay, differentiation through core formation, and external factors (e.g., tidal heating, induction

heating, and stellar irradiation). Accretion, through either small (pebble) or more likely large (planetary embryos and planetesimal) impactors (52), transfers kinetic energy to potential and thermal energy. Throughout this process, radioactive isotopes contribute to short and long-term heating. ^{26}Al and ^{60}Fe contribute to the early solar system heating, with half lives of 0.72 and 2.6 Myr, respectively (53). ^{235}U , ^{40}K , ^{238}U , and ^{232}Th have half-lives approximately three orders of magnitude longer: 0.71, 1.25, 4.47, and 14.0 Gyr, respectively. Thus, they survive the early stages of planetary formation and can contribute to the internally distributed mantle heat budget throughout Earth history. Planetary differentiation, although facilitated by short-lived radioactive isotope decay and accretion (e.g., (53)), determines the final major independent component of the initial heat budget. Core segregation, the primary product of planetary differentiation, results from the gravity separation of thermodynamically immiscible phases. Under conditions characteristic of terrestrial planets in the Solar System, heat loss can occur through several regimes (e.g., mobile lid, episodic lid, stagnant lid, episodic squishy lid, plutonic-squishy lid, sluggish lid, and heat pipe) based on the evolution on internal heating and surface yield stress (e.g., (54) and Fig. 1.3).

Each of these regimes sets the stage for surface topography and volatile evolution. Earth’s mobile-lid plate tectonic regime enables persistent volatile cycling and substantial reworking of surface topography and seafloor bathymetry.

1.5. Plate Tectonics on Earth | The evolution of Plate Tectonics on Earth is well studied but not well understood. Estimates for the operation of modern plate tectonics extend back to ~ 1 Ga and are marked by the observed preservation of ultrahigh-pressure metamorphism and ophiolites (55). Evidence for plate tectonics in the Achaean is marked by the preservation of orogens, accretionary prisms, and paired metamorphic belts (8; 56, and references therein). An onset of plate tectonics as early as the Hadean has been suggested from analyses of Hadean zircons that record regional- to global-scale heat flow indicative of modern subduction zones (e.g., (57)).



 Korenaga J. 2013.
Annu. Rev. Earth Planet. Sci. 41:117–51

Figure 1.4: Summary of suggested times for the onset of plate tectonics based on metamorphic, geomorphic, and geochemical lines of evidence, from the Korenaga 2013 annual review (8).

The operation of plate tectonics within the mobile-lid regime was likely different on the early Earth. However, both early (~ 4 Ga) and long-term continental extraction and/or mafic-to-felsic continental crust refinement (58) require a mechanism resembling plate tectonics to produce lithospheric overturn through ocean-crust formation and subduction. Furthermore, the extraction and/or refinement of continental crust produces a topographic dichotomy that redistributes surface water, facilitating both the emergence of land and the formation of shallow-marine environments (e.g., (59; 60)). Thus, plate tectonics establishes a mechanism for cycling geologically stored volatiles (e.g., CO_2) through the solid Earth and strongly controls the surfaces on which atmospheric carbon can ultimately be stored.

1.6. Solid Earth Controls on Ocean Basin Geometry | At present, two depth regimes of the ocean seafloor are governed by distinct dominant processes: 1) shallow seafloor (< 600 m) and 2) intermediate-to-deep seafloor (> 600 m).

Deep seafloor evolution is dominated by decompression melting and cooling by thermal conduction as seafloor spreads away from mid-ocean ridges ((61) and Fig. 1.5). At mid-ocean ridges, about 7 km of oceanic crust forms from the partial melting of upper mantle peridotite. The residual mantle forms the underlying lithosphere and has a composition similar to that of upper mantle peridotite (e.g., dunite ($> 90\%$ olivine) and lherzolite ($> 40\%$ olivine)). The base of the lithosphere is formally defined by mechanical rigidity (i.e., strength) of the rock; practically, geodynamicists define the lithospheric boundary at the ~ 1600 K contour (i.e., the modern surface potential temperature). The base of the lithosphere grows as the near surface mantle (> 100 km) cools by thermal conduction. Consequently, older oceanic lithosphere is colder, less thermally expanded, denser, and therefore subsides until it reaches isostatic equilibrium. A semi-infinite half-space cooling model provides a solution for the thermal and subsidence evolution of oceanic lithosphere under prescribed mantle and surface conditions (eq. 1.6 from (61) for mantle density ρ_m , water density ρ_w , mantle temperature T_m , surface temperature T_s , thermal expansion α_v , thermal diffusivity κ , ridge distance x , half-spreading

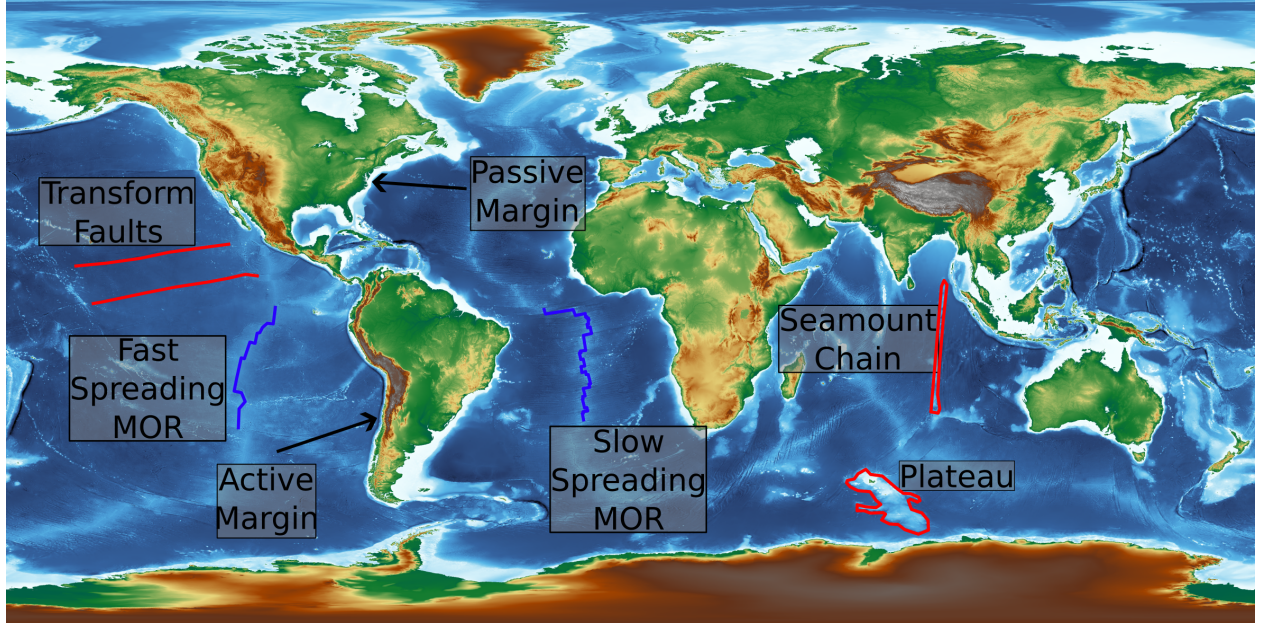


Figure 1.5: Global topography map highlighted with the major 1st and 2nd order bathymetry features and processes that shaped modern and paleo bathymetry. The topography model shown is ETOPO1 (9) and features are hand-drawn from high-relief features.

rate u_0 , and lithospheric age t). Actual measurements of seafloor subsidence versus age will deviate from equation 1.6, flattening once the seafloor exceeds ~ 70 Myr. This flattening could be the result of 1) sub-lithospheric Rayleigh-Taylor instabilities inducing small-scale convection on older ocean lithosphere, impeding cooling and additional thermal subsidence (62) or 2) thermal expansion from integrated mantle thermal insulation under older ocean lithosphere (63). Regardless of the underlying mechanism, plate-cooling models reproduce observed ocean-lithosphere subsidence as a function of seafloor age (e.g., eq. 1.7 from (64) or other formulations (65; 66)). However, their reliability for deep time (on Earth) or simulations (of exoplanet bathymetry) is unclear since they have built-in assumptions, i.e., constant equations for the depth-age relationship that do not capture the potential evolution of the mechanism that causes seafloor flattening.

$$d = 2500 + \frac{2\rho_m\alpha_v(T_m - T_s)}{\rho_m - \rho_w} \left(\frac{\kappa x}{\pi u_0} \right)^{1/2} \quad m \propto \left(\frac{x}{u_0} \right)^{1/2} = t^{1/2} \quad (1.6)$$

$$d(t) = \begin{cases} 2652 + 324\sqrt{t} & m, t \leq 75 \text{ Myr} \\ 5028 + 5.26t - 250\sin\left(\frac{t-75}{30}\right) & m, 75 < t \leq 160 \text{ Myr} \\ 5750 & m, t > 160 \text{ Myr} \end{cases} \quad (1.7)$$

Both shallow and deep seafloor are also shaped by additional volcanic and MOR spreading related features (e.g., plateaus, near-MOR-axis forming seamounts, off-MOR-axis seamount chains, submarine volcanic seamount arcs, and MOR-axis parallel normal faulting and perpendicular transform faulting) and pelagic and terrigenous sedimentation that contribute to 2^{nd} order variations in bathymetry (Fig. 1.5). Volcanic and MOR spreading features have amplitudes and wavelengths that range from 10-1000s m and 1-100s km, respectively. While statistical models exist to predict the bathymetry of some of these features, predicting the *spatial distribution* of some uniquely forming seamounts and plateaus is not trivial. Seafloor sediment layers have amplitudes and wavelengths that range from 100-1000s m and 100-1000s km, respectively. Seafloor sediment thickness models have been developed for young ocean lithosphere (i.e., further from the terrigenous sediment supply, (67)) and for passive margins (68). Predictions for sediment layers can also be derived from source-to-sink sedimentation models when paleo-conditions (e.g., topography, precipitation, lithology) can be constrained (69; 70; 71). Changes in eustatic and relative sea-level result from numerous processes that also influence seafloor bathymetry. While sea-level changes are most noticeable near coastlines and drastically influence the extent and distribution of shallow seafloor, they also contribute to deep ocean bathymetry. Solid Earth processes and their features (e.g., glacial/aquifer water storage, dynamically supported topography, geoid, volume of marine sediments, changes in MOR-ridge volume, and seafloor volcanism volume) drive long-term (myr timescale) changes

in eustatic and relative sea-level (72; 73), and thus seafloor bathymetry.

Bathymetry forming processes, whose expression varies even throughout Earth’s recent history, govern the distribution of seafloor and the oceanic chemical reservoirs that hold the vast majority of carbon accessible to the hydrosphere on timescales from 10s kyr to Myr. Together, planetary cooling, tectonic regimes, and bathymetric evolution define the geometry and dynamics of ocean basins, which in turn control the spatial distribution and efficiency of carbon storage and cycling.

1.7. Research questions and dissertation structure | This thesis aims to elucidate the role of solid planetary evolution on the LTCC. Three key components of solid planetary evolution (seafloor depth distributions, dynamic support of shallow seas coastlines, and basins’ extent for proxy aggregation/assimilation and terrestrial exoplanet volatile model development) are addressed in Chapters 2, 3, 4, and 5.

First, most carbon accessible to the hydrosphere on geological (i.e., LTCC) timescales is sunk into seafloor sediment and lithified sedimentary rocks, yet the influence of this evolving depositional surface on the LTCC remains poorly understood. Chapter 2 quantifies how variations in seafloor depth distributions control carbon partitioning between 1) the atmosphere and ocean basins and 2) within ocean basins. Second, shallow seafloor is an important carbon sink throughout much of Earth’s history and its extent is strongly influenced by the evolution of subduction and mantle flow. Chapter 3 quantifies the shallow seafloor area and coastline length supported by mantle flow and subduction related dynamic topography and geoid anomalies throughout the last 80 Myr, with implications for the sensitivity of shallow marine carbon preservation to mantle evolution before the Late Triassic. Lastly, ocean basins define the spatial distribution of chemical reservoirs that drive the LTCC, yet their boundaries are poorly constrained, leading to divergent interpretations when proxy records are calculated or compared with model simulations. Chapters 4 and 5 address this problem by developing a systematic, data-limited approach to basin identification that leverages graph

theory community detection. This approach also provides a diagnostic framework that can be used to systematically interrogate the geological surface conditions (driven by planetary cooling regimes) under which LTCC feedbacks remain stable and are conducive to long-term habitability.

A central theme of this thesis is the use of reduced-order modeling to characterize and interrogate coupled solid-planet and ocean-atmosphere systems. Reduced-order models are invaluable tools for paleoclimate science because they enable reconstructions and simulations of the LTCC while integrating carbon-cycle processes that operate on timescales spanning multiple orders of magnitude. In this vein, I develop reduced-order modeling tools that, among other applications, characterize the spatial domains used in EMICs. Recent advances in plate motion reconstructions, bathymetric modeling, graph-based community detection theory, reapplication of EMICs to terrestrial exoplanet science, and state-of-the-art modular EMIC frameworks for paleo-Earth applications make the focus of this dissertation technically feasible and timely.

Chapter 2

The effects of bathymetry on the long-term carbon cycle and CCD

2.1. Significance | Our study illustrates that seafloor bathymetry—both mean depth as well as bathymetry distribution—has a significant influence on the global carbon cycle by controlling the location of seafloor carbonate sedimentation. Considering the plausible range of bathymetric configurations over the last 80 Myr of Earth’s history, bathymetry cannot be ignored when studying the carbon cycle and inferring changes in ocean chemistry and productivity.

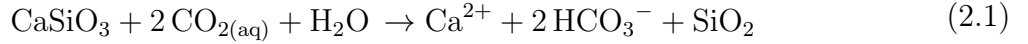
2.2. Abstract | The shape of the ocean floor (bathymetry) and the overlying sediments provide the largest carbon sink throughout Earth’s history, supporting $\sim$ one to two orders of magnitude more carbon storage than the oceans and atmosphere combined. While accumulation and erosion of these sediments are bathymetry dependent (e.g., due to pressure, temperature, salinity, ion concentration, and available productivity), no systemic study has quantified how global and basin scale bathymetry, controlled by the evolution of tectonics and mantle convection, affects the long-term carbon cycle. We reconstruct bathymetry spanning

the last 80 Myr to describe steady-state changes in ocean chemistry within the Earth system model LOSCAR. We find that both bathymetry reconstructions and representative synthetic tests show that ocean alkalinity, calcite saturation state, and the carbonate compensation depth (CCD) are strongly dependent on changes in shallow bathymetry (ocean floor ≤ 600 m) and on the distribution of the deep marine regions ($>1,000$ m). Limiting Cenozoic evolution to bathymetry alone leads to predicted CCD variations spanning 500 m, 33 to 50% of the total observed variations in the paleoproxy records. Our results suggest that neglecting bathymetric changes leads to significant misattribution to uncertain carbon cycle parameters (e.g., atmospheric CO_2 and water column temperature) and processes (e.g., biological pump efficiency and silicate-carbonate riverine flux). To illustrate this point, we use our updated bathymetry for an Early Paleogene C cycle case study. We obtain carbonate riverine flux estimates that suggest a reversal of the weathering trend with respect to present-day, contrasting with previous studies, but consistent with proxy records and tectonic reconstructions.

2.3. Introduction | The long-term surface carbon cycle is a balance of carbon fluxes between sources (e.g., volcanic emissions) and sinks (carbon burial by carbonate-bearing minerals and organic-rich sediments on the seafloor, which are then recycled via subduction). Silicate weathering on land or seafloor is a source of a) bicarbonate ions, acting as a short-term carbon sink through the removal of atmospheric carbon and b) cations (Ca^{2+} , Mg^{2+} , $\text{Fe}^{2,3+}$, Eq. 2.1). These are physically transported to ocean basins by rivers and eventually deposited in ocean sediments. Thus, weathering generates ocean alkalinity, which leads to long-term carbon removal from the ocean-atmosphere system. On timescales of 10s of Myr, studies suggest that the carbon cycle reaches a steady state, wherein the ocean bicarbonate concentration is largely modulated by dissolved continental rocks (Eq. 2.1) and dissolution of atmospheric CO_2 into the ocean’s surface (Eq. 2.2). However, this is an end-member perspective focused on carbon (C) sources and sinks without detailed consideration of how the steady state

is achieved. Instead, it implicitly assumes that ocean chemistry and carbonate deposition always adapt to reach the required C sink to match the C sources. Most studies of long-term carbon cycle dynamics for Earth and other planetary bodies make these assumptions.

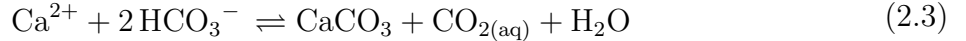
Silicate weathering :



Ocean carbonate chemistry :



Carbonate precipitation :



A different perspective for the C cycle focuses on the dynamics of marine carbonate precipitation, dissolution, and deposition: the net rate of carbonate deposition determines the evolution toward steady state. This perspective highlights the important role of the ocean as the dominant surface C reservoir as well as the main buffer for short-term C cycle perturbations, due to the high solubility of CO_2 in water (74). Marine carbonate chemistry (Eq. 2.2) allows the ocean to take up far more CO_2 than expected based on CO_2 solubility alone, by transforming added CO_2 into HCO_3^- and CO_3^{2-} ions [together termed dissolved inorganic carbon (DIC)]. The dissolved carbonate ions in seawater are eventually biotically and abiotically precipitated as carbonate-bearing minerals (calcite, aragonite, magnesite, and siderite) through ocean productivity, primarily in surface waters (Eq. 2.3, long-term C sink). After precipitation, CaCO_3 particles sink to the deeper ocean via the biological pump, where they are either deposited or dissolved. Dissolution of CaCO_3 (calcite and aragonite) in the deep ocean is driven by increasing solubility with increasing pressure and decreasing temperature (75). The lysocline marks the depth interval in the ocean where $\sim 90\%$ CaCO_3 dissolves, while the carbonate compensation depth (CCD) is the depth below which all newly formed CaCO_3 is dissolved. However, the exact depths (pressure) of the lysocline and CCD

in an ocean basin are dependent on its seawater chemistry: primarily the DIC and alkalinity. In a deposition/precipitation focused perspective of the C cycle, the depth and the spatial distribution of the seafloor depositional area (i.e., the seafloor bathymetry) directly determine the area of seafloor above the lysocline and consequently the carbonate burial efficiency. This, in turn, controls seafloor carbonate coverage, ocean carbonate chemistry, and hence atmospheric CO₂ concentration, as well as the ocean’s ability to buffer any fast and large CO₂ releases from volcanic eruptions.

Therefore, we expect bathymetry to strongly influence the net rate of ocean C burial and the carbonate saturation in the oceans as it adjusts to reach the long-term steady state. In this study, we highlight the critical role of seafloor bathymetry on the evolution of the marine C cycle, using idealized marine carbonate cycle models that modulate the available shallow (≤ 600 m) vs deep ($>1,000$ m) ocean floor area. Because the mean depth and relative depth distribution of the seafloor are directly controlled by the speed of plate tectonics, mantle flow, and relative sea level, our results highlight a hitherto unappreciated strong coupling between planetary interiors and surface climate through chemical processes. This coupling is important because it evolves significantly through time as ocean basin structure changes in response to large-scale tectonic and convective processes, such as super-continent formation and breakup (37; 38). For instance, as plate motions slowed throughout the last 100 Myr, more thermal subsidence has deepened ocean basins (Fig. 2.1) (76; 77).

Our study focuses explicitly on the chemical effects of bathymetry on Earth’s ocean-atmosphere system. This is distinct and complementary to the effect of bathymetry on the physical dynamics of the ocean through changes in ocean circulation, mixing, and heat flux. For instance, ocean circulation is strongly influenced by the opening and closing of ocean basins (e.g., closure of the Tethys and opening of the Atlantic) and the development of gateways between them (e.g., Drake’s Passage, Tasman opening, and closure of the Central American Seaway, e.g., refs. (78; 79; 80)). Both global circulation and vertical mixing rates (also sensitive to seafloor bathymetry) contribute to varied regional/global water column

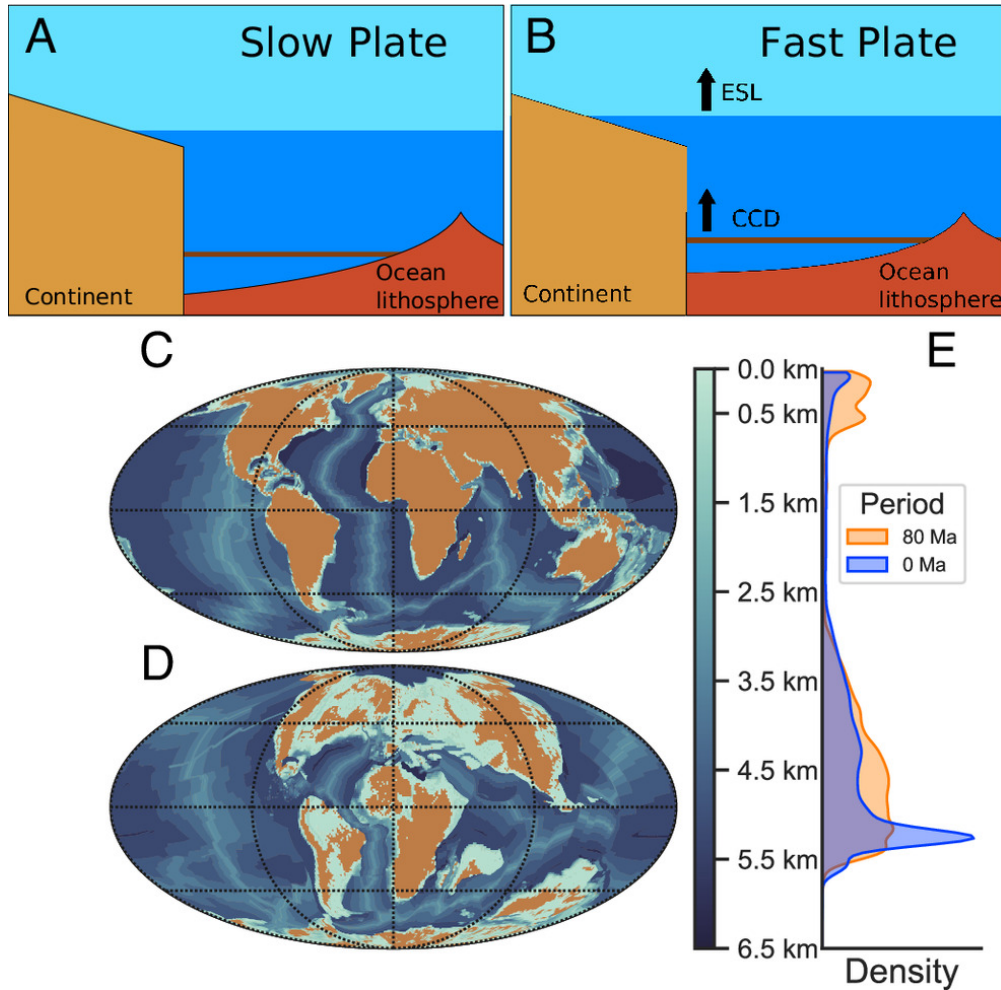


Figure 2.1: Evolution of 1st and 2nd order bathymetry. (*A* and *B*) Cartoon of seafloor morphology. Deep ocean bathymetry is represented by the oceanic lithosphere (rust red), while shallow bathymetry is represented by the continent (yellow) water interface. The brown line denotes the CCD, below which no carbon is sequestered. The light and dark blue interfaces represent eustatic sea-level (ESL). These cartoons show how slow and fast plate motions lead to increased or decreased thermal subsidence, respectively, before subduction, resulting in 1st order changes in ocean floor depth, sea-level, and the CCD. Bathymetry reconstructions at present-day (*C*) and 80 Ma (*D*) represent slow (*C*) and fast plate (*D*) motion regimes. (*E*) Kernel depth distributions for (*C* and *D*) denote two major components of bathymetry evolution: 1) changes to shallow seafloor (≤ 600 m) area and 2) changes to the distribution and mean deep bathymetry ($>1,000$ m).

ocean chemistry and nutrient availability (81). However, these studies do not analyze the effects of changing bathymetry on ocean carbonate deposition/dissolution.

The limited work on the chemical effects of bathymetric changes has focused on using bathymetry as a boundary condition to determine solid earth carbon fluxes, the impact of continental shelf evolution on biotic productivity, or the evolution of continental shelves on the long-term carbon cycle. For instance, atmospheric CO₂ concentrations were found to have been impacted by shallow carbonate burial during the Quaternary large glacial-interglacial sea-level changes (82; 83). On timescales of 1 to 10s of Myr, Komar and Zeebe proposed that changes in ocean chemistry, previously attributed to long-term global carbonate weathering trends, could instead be linked to shelf productivity evolution coinciding with continental flooding (84). Mesozoic-Cenozoic sedimentary carbon subduction and degassing fluxes have been computed using depositional zones, defined by the intersection of a reconstructed CCD with evolving bathymetry (43; 85). While these previous studies highlight integral roles of the bathymetric effect on the carbon cycle, bathymetry itself is not integrated into the C cycle dynamics. Instead, they focus on bathymetry as a carbon transport facilitator. They do not explore how bathymetry influences when, where, and how much carbon-bearing sediment is deposited or distinguish between the role of mean depth vs. bathymetric distribution.

Here, we address this critical knowledge gap. We utilize a commonly implemented intermediate complexity Earth system model—Long-term Ocean-Sediment-Carbon cycle Reservoir (LOSCAR) (Fig. A.1) that allows for the analysis of basin-scale ocean chemical responses in the water column and seafloor sediments (12). The Earth system is spatially resolved to first order—laterally with different ocean basin boxes to simplify dynamics, and vertically to represent depositional surfaces. We use LOSCAR for this study, instead of more complex Earth system models such as CGenie (86) because of ease of interpretation and modification. With LOSCAR, we can easily tease apart causation, by independently focusing on the chemical response of the ocean separately from the physical.

To gain mechanistic insights into the carbon cycle’s relevant feedbacks, we first explore

the cycle’s sensitivity to idealized bathymetric models that use present-day basin sizes but synthetic bathymetry distributions. We then create and use a unique reconstruction of global bathymetry for the last 80 Myr (in 5 Ma increments) to put our results in the context of Earth’s evolution. This reconstruction is based on recent ocean lithosphere age reconstructions paired with a lithospheric cooling relationship (38; 64). We account for the effect of oceanic plateaus and other intermediate depth bathymetry and eustatic sea-level changes (c.f. Materials and Methods). Finally, we illustrate the dramatic effect of selecting accurate seafloor bathymetry for inferring riverine flux, using the late Paleocene as a case study.

2.4. Results |

2.4.1. Synthetic Bathymetry | We constructed synthetic bathymetry with changes in the total area of the shallow ocean within the range of estimates for the last 80 Myr (between 96 and 120% of present-day area, Fig. 2.2D) and changes in the average depth of deep bathymetry. For deep bathymetry, we used Gaussian distributions centered between 3.5 and 4.5 km depths (Fig. 2.2D). We use our synthetic bathymetry models directly in LOSCAR’s modern three ocean system (Fig. A.1) to calculate ocean and atmosphere steady states. We keep surface weathering and alkalinity fluxes constant as we change bathymetry. To track changes in ocean chemistry, we analyze variations in the CCD and first-order ($>99\%$, ref. (45)) ocean total alkalinity ($\text{TA} = \text{HCO}_3^- + 2\text{CO}_3^{2-} + \text{B}(\text{OH})_4^- + \text{HS}^- + 2\text{S}^{2-} + \text{OH}^- - \text{H}^+ - 2\text{SO}_4^{2-}$). In our work, changes in TA are primarily driven by carbonate chemistry ($\text{HCO}_3^- + 2\text{CO}_3^{2-}$). TA measures the charge difference between the ocean’s major conservative cations and anions (e.g., Cl and sulfate) and is a metric for the ocean’s ability to precipitate carbonate and buffer against pH variations. From here on out TA and alkalinity are used interchangeably. Changes in ocean alkalinity are good proxies for the carbonate saturation state of the ocean. For example, a higher ocean TA is required to prevent dissolution and deposit carbonate in the deep ocean as opposed to shallow water continental shelves. CCD variations represent

1) changes to the depositional area in carbonate-saturated waters and 2) solid carbonate dissolution or deposition above the CCD. The CCD is a useful marine C cycle proxy because it can be observationally inferred over recent Earth history. Throughout the Cenozoic, the CCD varied between ~ 4.5 and 3 km depth, with perturbations marked by CO_2 forcing events (4).

First, we isolated the effects of changing the average depth of the deep ocean by choosing a constant shallow ocean area for the present-day ($\sim 30 \times 10^6 \text{ km}^2$, black point in Fig. 2.2A). We find that shallowing the average deep marine depth by 500 m decreases the CCD by ~ 500 m (Fig. 2.2A). Physically, this can be understood as follows: as the depth of the deep ocean decreases (Fig. 2.2D), there is a reduction in the pressure-dependent calcite solubility that supports increased calcite burial rates, leading to a reduction in shallow and deep ocean alkalinity (Fig. A.3) and saturation state (Fig. A.4) when burial equilibrates with riverine alkalinity input. This leads to a drop in ocean CO_2 storage capacity, increased atmospheric CO_2 , and a shoaled CCD (Fig. A.15).

Second, we isolated the effects of changing shelf bathymetry given a constant average deep bathymetry similar to present-day (4 km). Doubling the present-day shelf area ($\sim 60 \times 10^6 \text{ km}^2$) leads to the CCD being shoaled by ~ 250 m (Fig. 2.2A). Physically, this can be understood as follows: Given constant global productivity, increased shelf area leads to increased proportion of calcite rain and burial on the shelf and decreased open ocean rain. The change in rain partitioning reduces the flux of bicarbonate to the deep ocean, decreasing the deep marine alkalinity (Fig. A.3) and calcite saturation (Fig. A.4), leading to increased deep marine dissolution and a shoaled CCD. This process is analogous to what has been proposed for Quaternary glacial changes in sea-level and the marine C cycle (coral reef hypothesis, ref. (87)).

In contrast to changing mean deep ocean depths, we see two CCD shoaling trends with increasing shallow seafloor area, both stemming from shelf vs. open ocean productivity partitioning and consequent changes in carbonate burial vs dissolution. The initially weaker

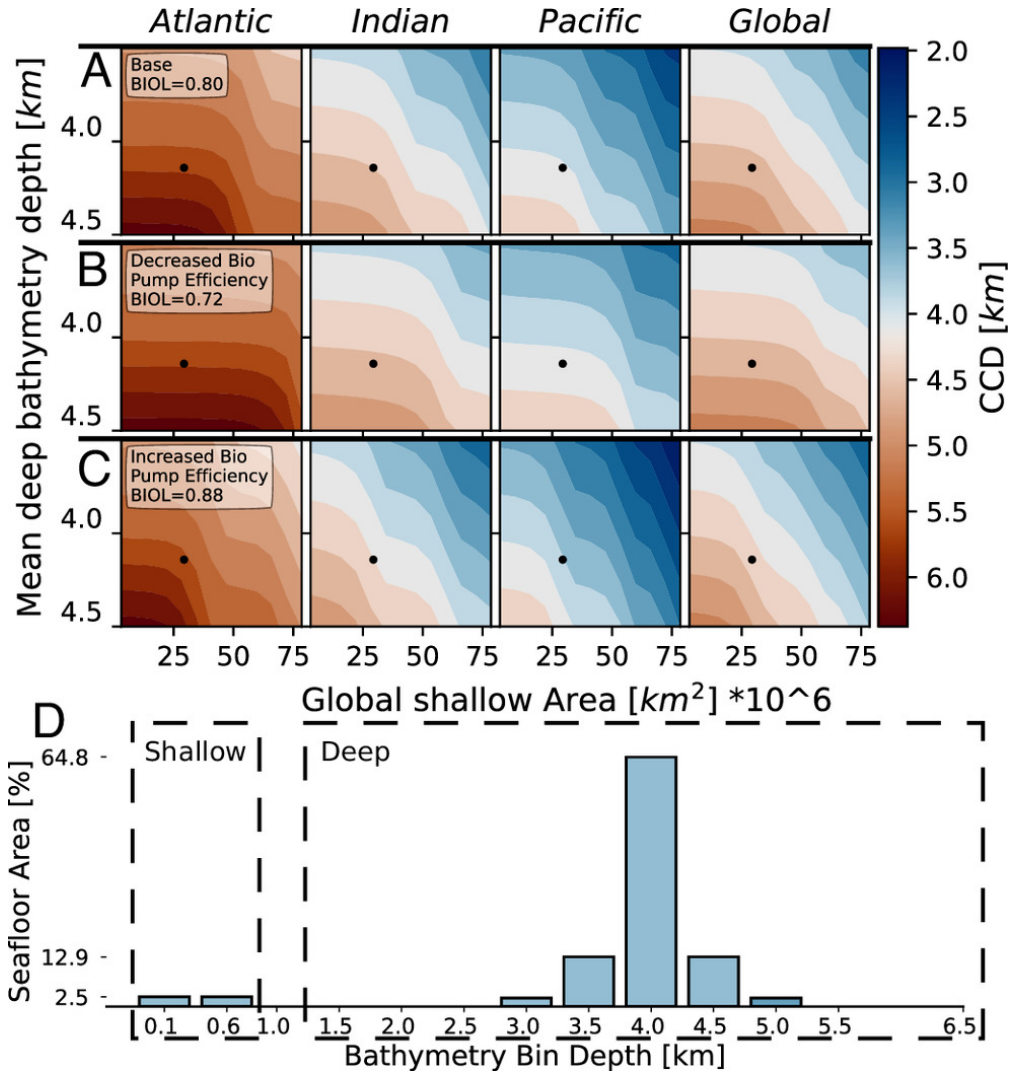


Figure 2.2: Response of CCD given synthetic bathymetric distributions representing 1st order changes in deep bathymetry and shallow seafloor area. (A-C) LOSCAR bathymetry inputs shown in (D): 1) deep bathymetry represented through Gaussian distributions, and 2) shallow bathymetry areal extent ranging from 96 to 120% of its present value. Synthetics produce similar areal extents and average deep marine depths to those within the last 80 Myr. (A-C) columns and rows show different ocean basins and Earth system model setups, respectively. (A) Base model for the steady-state CCD (contours) obtained with LOSCAR modified to include our preferred bathymetry. No silicate/carbonate weathering feedback included. (B and C) Models with decreased and increased biological pump efficiency, respectively. Decreased biological pump efficiency reduces calcite rain, shallow sequestered alkalinity, and calcite erosion in the deep marine. It allows the Pacific and Indian basins to persist as depositional basins for greater shallow shelf areas.

effect on the CCD, most evident in the Atlantic, is a result of stripping away more and more calcite deposition from the deep Pacific and Indian ocean basins as shelf area increases. This leads to large amounts of alkalinity being supplied to the shelf through deep marine calcite erosion before the earth system model reaches a steady state. However, the stronger effect (kinks at shallow area $> 50 \times 10^6 \text{ km}^2$) appears once the Pacific and Indian oceans transition from a weak-deposition to a nondeposition regime since the CCD is above any available seafloor (Fig. A.5). Open ocean calcite rain is not deposited at all in the deep Indian and Pacific regions and significant changes in Atlantic Ocean alkalinity (and hence CCD) are needed to reach equilibrium (Fig. A.3). Subsequently, increased shelf area leads to stripping away alkalinity from the deep marine water column in the Atlantic, causing rapid decreases in the CCD. We note that the exact location of the regime transition (kink) depends on the assumed modern ocean chemistry and interbasin differences. Presently, the Atlantic Ocean has high alkalinity and a deeper CCD due to the pattern of precipitation, river influx, and ocean circulation. However, during Earth history, there is no reason to expect a priori that each ocean basin would have very similar alkalinity. We expect instead a transition in the shallow shelf CCD shoaling trend between the two end-members described above at different shallow shelf areas and mean deep ocean depths.

Overall, we find that the CCD (as a proxy for the ocean carbonate system) is strongly sensitive to bathymetry—both the global bathymetry distribution (shallow area as well as deep ocean depth) as well as the distribution within different ocean basins. However, the final extent of sensitivity will be dependent on coupled bio-geochemical-ocean dynamics processes that are not fully included. For example, while increased seafloor area led to increased shelf vs. open ocean productivity, nutrients (PO_4) upwelled from the deep ocean could become restricted and mitigate shelf productivity growth. One way to parameterize this effect on the Earth system is by decreasing the efficiency of the biological pump. Reducing the pump’s efficiency by 10% with respect to the base model results in less surface water alkalinity being sequestered in the shelf area (Fig. 2.2B). Ultimately, this reduces deep ocean dissolution,

allowing for calcite burial in the Pacific and Indian ocean basins, making the CCD less sensitive to changes in shelf area. On the other hand, increases in the biological pump efficiency produce the opposite effect (Fig. 2.2C). We find that our results remain essentially the same even when we include the atmospheric CO₂-dependent weathering feedback. This feedback relaxes the atmospheric CO₂ back to the initial state but not the ocean alkalinity and CCD (Fig. A.15). We find that bathymetry alone can lead to as much as 2.5 km CCD variations both within basins and globally, with accompanying shifts in calcite saturation, alkalinity, and ocean and atmospheric CO₂ (Fig. 2.2 and Figs. A.3 and A.4).

2.4.2. Reconstructed Bathymetry | One of the limitations of the synthetic results is that they don't provide a direct assessment of how CCDs will change with realistic bathymetry evolution over the Cenozoic, which saw changes in the areal extent and depth distribution of individual basins due to the growth and decay of oceans. To mitigate this limitation and directly compare with Pacific & global CCD inferences, we generated unique bathymetry reconstructions, corrected for basin volume (c.f. Materials and Methods and Fig. A.12) over the last 80 Myr. We find significant bathymetric evolution over this time, culminating in a reduction to $\sim 33\%$ of the shallow ocean area at present-day with respect to 80 Ma. In addition, the average depth of the oceans deepened by ~ 500 m (Fig. 2.1). We find two primary reasons for this behavior: a) changes to shallow bathymetry (≤ 600 m) associated with continental topography and flooding, and changes in eustatic and relative sea-level reconstructions since the Cretaceous; b) changes in deep bathymetry (>1 km depths) occur between mid-ocean ridge depths (~ 2.5 km) and abyssal plain depths (~ 5 km) and are largely a consequence of the disappearance of the Tethys and older global seafloor compared to over the last ~ 80 Myr. These factors combine to generate a narrow and strongly peaked deep seafloor distribution today.

In addition to global changes, we find relative differences in the Atlantic, Indian, and Pacific Ocean basin surface areas over time (c.f. Materials and Methods for details regarding

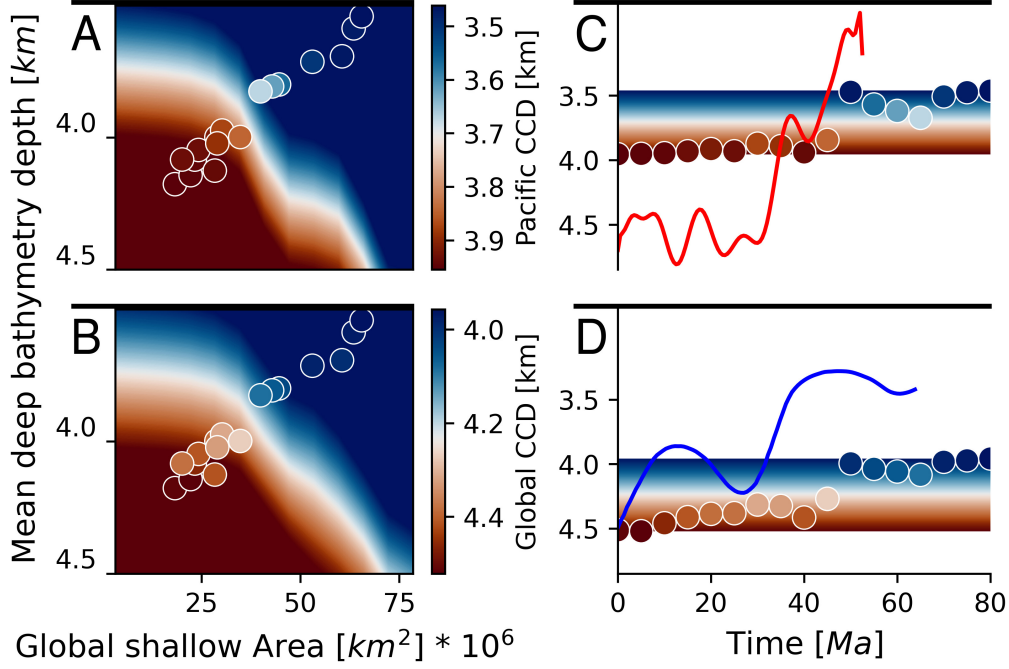


Figure 2.3: Response of CCD given reconstructed bathymetry. Contours in (A and B) show the CCD for synthetic bathymetry models and dots show realistic reconstructions over the last 80 Myr for the Pacific and global ocean(s), respectively. Synthetic bathymetry modeled CCDs are plotted within reconstruction ranges. The saturation within the synthetic models indicates that CCDs calculated from realistic bathymetry reconstructions are subdued, but still lead to ~ 500 m variations which are exclusively bathymetric dependent. CCD evolution for the Pacific (C) and globally (D). We show (C and D) the total synthetic range shown throughout the reconstruction period. The blue and red curves represent the Lyle et al. (10) global CCD and Pälke et al. (4) equatorial CCD, respectively. Notably, reconstructions show the CCD shoaling between 65 and 50 Ma and deepening between 30 Ma and present-day. The major CCD deepening between 50 and 40 Ma occurred concurrently with bathymetric changes resulting from the India-Asia collision.

basin separation, Figs. A.14 and A.17). Today, pelagic carbonate accumulation varies by ocean basin and is highest in the Atlantic Ocean due to the nature of ocean circulation and weathering fluxes (88; 89). Interocean basin changes hence add to the bathymetric controls on the carbon cycle. Our synthetic results (Fig. 2.2) demonstrate the importance of basinal bathymetric distributions through the CCD and the seafloor carbonate distribution above it (Fig. A.16).

Using our bathymetry reconstructions, we modified the bathymetry distribution in the LOSCAR present-day model setup while keeping the other ocean (e.g., mixing), atmosphere (e.g., pCO_2), global productivity, and weathering components constant at present-day values. Similar to synthetic tests, the CCDs calculated from actual bathymetry evolution deepened over the last 80 Myr with an increased average depth of deep marine seafloor and decreased shallow seafloor area. Notably, CCD variations calculated from the last 80 Myr of bathymetry evolution can describe ~ 33 to 50% (~ 500 m) of the total Cenozoic variation derived from seafloor sediment cores (Fig. 2.3 B and D) (4; 10). The CCD has a significant dependence on the evolution of deep bathymetry, which is tied to plate tectonics and mantle dynamics. The current lack of bathymetric representation in many C cycle models leads to a misunderstanding of the contribution of uncertain carbon cycle parameters (e.g., atmospheric CO_2 , water column temperature, carbonate-silicate weathering, productivity) over 10s of Myr. We readily acknowledge that we have simplified the processes at play, since other C cycle parameters (e.g., volcanic outgassing, pCO_2 , carbonate weathering, ocean Ca^{2+} and Mg^{2+} concentrations) are also changing during this time period. However, although a full model inversion and sensitivity analysis is beyond the scope of the present study, it is clear from the magnitude of the effect that neglecting well-known changes to the shape of the marine container imperils interpretation.

2.4.3. Early Paleogene Case Study | As a specific example, we use Early Paleogene (pre-Paleocene-Eocene Thermal Maximum, or PETM) as a case study. This is an extensively

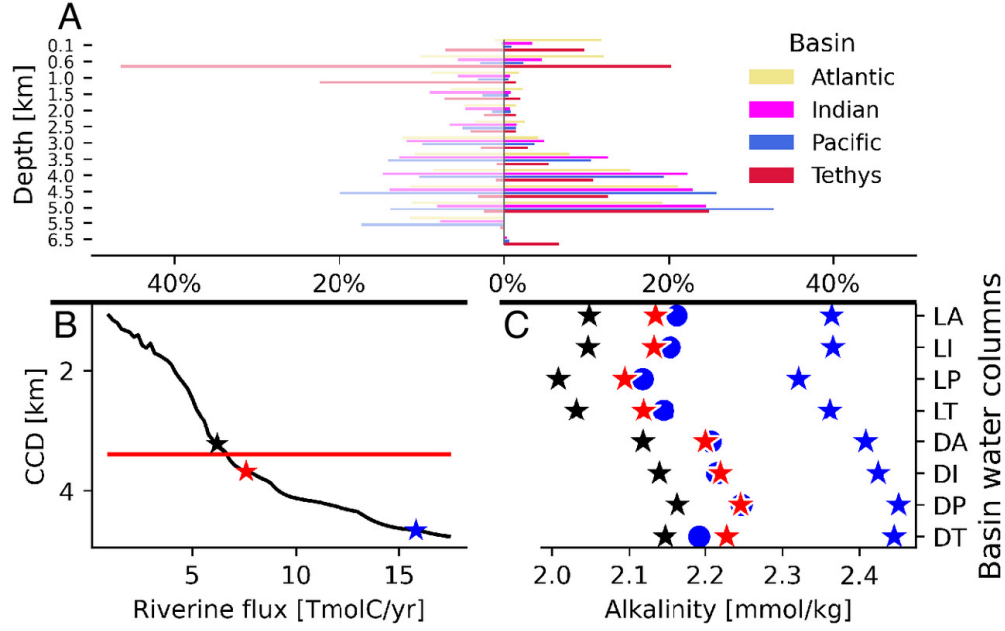


Figure 2.4: Discrepancy of bathymetry influenced CCD as a function of riverine flux and ocean alkalinity. (A) Default LOSCAR bathymetry model (Left) and updated 60 Ma bathymetry reconstruction (Right). Perturbed steady-state CCD and ocean alkalinity using our updated bathymetry model and a shelf-deep marine sediment partition coefficient half its default LOSCAR value ($FSHLF = 2.25$). (B) Global CCD with updated bathymetry and varying the degree of initial CaCO_3 riverine flux. Blue, red, and black stars represent initial riverine fluxes corresponding to the default LOSCAR values, LOSCAR model CCD reproduction, and calibrated LOSCAR model alkalinity reproduction. The red line is a sediment core-derived global CCD with no specific riverine flux [Lyle et al. (10)]. (C) Corresponding ocean alkalinity throughout basin water columns (Fig. A.1) with the addition of LT and DT denoting the Tethys Ocean. The symbols as in B except for blue dots showing LOSCAR's default alkalinity.

studied and well-constrained period with a large wealth of ocean paleoclimate data (ref. (12), and references therein). We used LOSCAR’s paleo setup which represents conditions before the PETM, including paleobathymetry for the four major ocean basins (Fig. 2.4A). First, we updated the bathymetry model with our 60 Ma reconstruction (Fig. 2.4A) and reduced the shelf vs open ocean production partitioning factor by 50% (to numerically stabilize the model). In steady-state, the C system responded to updated bathymetry by deepening the global CCD by ~ 2 km and increasing ocean alkalinity by $\sim 10\%$ throughout the water column (Fig. 2.4 B and C). This large change in the CCD is inconsistent with available data (4; 10) and the previously tuned model (ref. (12), and references therein), suggesting that other carbon cycle model parameters need to be reassessed. We were then able to reproduce the measured CCD and alkalinity values by varying one poorly constrained geological variable—the riverine alkalinity flux (Fig. 2.4 B and C) (84; 90). Tuning for either the CCD or deep ocean alkalinity resulted in a riverine alkalinity flux that is lower than the estimated present-day value of 12×10^{12} molC/y (4; 91) or the Early Paleogene value of $\sim 15.8 \times 10^{12}$ molC/y (12) previously tuned by LOSCAR with an outdated bathymetry reconstruction. The obtained value (6.2 to 7.6×10^{12} molC/y) is consistent with recent surface process modeling estimates (71). While we have not exhaustively explored the Early Paleogene Earth system parameter space considering other model parameters (e.g., ocean mixing and ocean productivity), the strong sensitivity of riverine flux to bathymetric changes clearly illustrates that model calibrations must account for the bathymetry effect (mean depth and crucially distribution). Consider a gedanken experiment where we are trying to estimate the temporal evolution of global carbonate-silicate weathering fluxes by fitting CCD observations from ocean core data. The CCD is sensitive to weathering fluxes since reduced bicarbonate riverine flux limits ocean bicarbonate concentration, therefore, reducing ocean alkalinity, calcite saturation, ocean C sequestration, and shoaling the CCD. However, our results show that ocean bathymetry is another previously ignored strong sensitivity parameter to the CCD. Since ocean bathymetry is, to first order, reconstructable via standard geodynamical methods, it is a critical climate

forcing that must be correctly accounted for.

2.5. Discussion | Our analysis clearly illustrates that the bathymetric effect—both mean depth and bathymetry distribution—on the C cycle is significant. Considering the plausible range of bathymetric configurations over the last 80 Myr of Earth’s history, it cannot be ignored when studying the carbon cycle and inferring changes in ocean chemistry and productivity. Remarkably, up to 50% of the carbon cycle evolution throughout the Cenozoic can potentially be explained through bathymetry evolution alone. This evolution is important for describing the carbonate reservoir distribution, ocean alkalinity, and therefore the buffering capacity of the oceans. These effects are purely bathymetry driven, prescribing no change in ocean salinity or temperature yet producing similar trends in DIC and ocean alkalinity.

More broadly, the results highlight the importance of thinking about the long-term C cycle from the perspective of marine carbonate precipitation, dissolution, and deposition rather than just a balance between sources and sinks. Why is this the case? Because the ocean needs to reach saturation in a reasonable fraction of the oceans to deposit carbonates (C removal) and balance volcanic flux. If there is a significant amount of shallow shelf area with ocean productivity, deposition can occur with a lower ocean alkalinity. Else, carbonate saturation in deeper ocean basins is needed for carbonate deposition, requiring a higher ocean alkalinity. Conceptually, the net impact of changing bathymetry is similar to the classic carbonate compensation feedback due to changes in ocean alkalinity following changes in ocean temperature, salinity, or productivity (89). Here, the captured C cycle feedback is driven by seafloor bathymetry changes and is thus especially relevant for long timescales. Since the ocean-dissolved carbon is the dominant C reservoir in the ocean-atmosphere system, changes in ocean carbonate chemistry directly influence the atmospheric C values on timescales of 10s of kyr. On longer timescales (100s of kyr), changes to atmospheric C (e.g., arising from tectonic evolution) are accommodated by silicate weathering feedbacks. Bathymetry

evolution further informs changes in C cycle processes across multiple timescales through other feedbacks. For instance, changes in mid-ocean ridge spreading rate vary CO₂ outgassing and long-term weathering trends (100s of kyr). Additionally, eustatic sea-level change due to bathymetric evolution (1 to 10s of Myr) can lead to continental flooding, potentially increasing seafloor biomass, CO₂ from respiration, and dissolution of seafloor CaCO₃ sediment within calcite over-saturated bottom waters (92) on shorter timescales ranging 10s of y to 100s of kyr (ref. (93), and references therein). The strength of these influences is partially dependent on the spatio-temporal pattern of bathymetry changes and can be either net positive or negative feedbacks. Furthermore, a deposition-focused perspective illustrates that the distribution of sediment (partly bathymetry controlled) determines the ocean buffering capacity (through carbonate dissolution) to external CO₂ perturbations and how the system evolves/relaxes back to equilibrium. Finally, the sediment and seafloor carbonate deposition and subduction closes the long-term surface-interior C cycle loop. In other words, bathymetry is not just a boundary condition (let alone constant!). It is also the driver of change, controlling depositional areas, consumption of riverine flux, and the chemical state of individual ocean basins. Seafloor bathymetry truly represents the interface along which the solid earth and ocean lie and actively interact.

While our reconstructions are restricted to the past 80 Myr, large variations to bathymetry can be inferred from reconstructed seafloor ages throughout the last 230 Myr (94) and even back to 410 Ma when approximating seafloor age from past plate configurations (95). From a climate modeling perspective, future reconstructions will benefit from careful modeling of processes affecting shallow (0 to 600 m), intermediate (600 to ~2,500 m), and deep (>2,500 m) bathymetry evolution not associated with the classical oceanic plate evolution: syn-rift extension and sediment deposition, continental shelf extent, mantle flow supported bathymetry (e.g., the Western Interior Seaway), seafloor volcanism (e.g., LIP, seamounts, and paleo island arcs), and glacier eustasy (96; 97; 98). Our model results highlight that the shallow bathymetry component leads to large CCD shifts. Despite the various challenges in

reconstructing shallow seafloor area, we have identified it as a high priority direction for future work and strongly recommend that geochemical models calibrate poorly constrained Earth system parameters (e.g., weathering rates, riverine flux to deep marine, conveyor transport, biological pump misinterpretation of productivity, and ocean chemistry are clearly illustrated by our Early Paleogene case study) while using period-appropriate bathymetry.

Although the focus of this study is the chemical processes associated with ocean bathymetry, chemical and physical processes are also coupled, through the interior-driven bathymetry evolution shaping Earth’s biological ecosystem (e.g., productivity) and the long-term carbon cycle. The combination of gateways, continental shelves, magmatic emplacements, island arcs, and instances of seafloor roughness informs bathymetry evolution that can influence mixing rates, upwelling deep water nutrients, and therefore the spatial distribution and amount of biological production in surface waters (99). Changes in continental-ocean configuration directly affect the amount of riverine flux associated nutrients and alkalinity consumed in the shelves and unable to reach the deeper part of ocean basins. There is a rich set of coupled chemical-physical ocean dynamics feedbacks associated with changes in seafloor bathymetry driven by Earth’s deep interior. Our results highlight that these should be studied as first-order effects to develop a better understanding of Earth’s carbon cycle evolution and when comparing to the geological record at global and especially regional scales.

Some of these bathymetry-associated couplings become increasingly important for the Early Cretaceous with the advent of calcareous phytoplankton (ref. (100), and references therein) and significant change in the ratio of deep ocean vs shallow water sediments as global carbon sinks (43). On Earth today, about half of the global carbonate burial occurs at shallow depths (< 200 m water depth) on continental shelves (Neritic carbonate) while the rest occurs on the deeper ocean floor (above the CCD). Prior to ~ 250 Ma, carbonate precipitation in Earth’s oceans was caused only by neritic organisms such as corals (101), and carbonates were deposited only in shallow shelf regions. Furthermore, the Early Cretaceous exhibits a rich diversity of tectonic events: the formation and destruction of carbonate depositional

zones with the formation of the south Atlantic basin, continued breakoff of Greater India from Gondwana, and emplacement of large-scale continental and oceanic flood basalts. These lead to strong variations in ocean circulation patterns, mixing rates, and the spatial pattern of upwelling zones (associated with high productivity). Previously, the carbon cycle has been shown to have a strong dependence on such changes in ocean circulation and mixing (ref. (80), and references therein) since the ocean is the main C reservoir. Here, we showed bathymetry also independently modulates the long-term carbon cycle’s chemical processes. Therefore, the long-term C cycle on Earth is better understood when evolving bathymetry is self-consistently coupled with chemical processes and ocean dynamics changes.

The significant coupling between the evolution of the carbon sequestration depth and seafloor evolution suggests that the vast differences in seafloor age distributions spanning the last 410 Myr (95; 102) will affect global bathymetry, the depth of carbon sequestration, and therefore carbon cycling to the solid Earth and the subsequent rates of volcanic outgassing from subduction zones and mid-ocean ridges. Mid-ocean ridges, island arcs, and oceanic plateau formation contribute to bathymetry evolution, are nonunique to the recent past, or even Earth, and could be statistically constrained (96; 103) where direct reconstruction methods are inapplicable. Since the planetary bathymetry distribution is controlled by the planet’s interior geodynamics and tectonics, the bathymetric influence on carbon cycling represents another important coupling between the solid planet and its surface-atmosphere environment. While present models analyzing the habitability of Earth and Earth-like planets to remain habitable have extensively focused on the role of continental and seafloor silicate weathering, the role of bathymetry-controlled carbonate deposition in modulating atmospheric CO₂ remains unstudied (104). Given the large diversity in geodynamics and ocean volumes for terrestrial exoplanets, analysis of bathymetry-associated effects will provide different constraints on planetary habitability and how it evolves following a large internal (e.g., volcanic eruption) or external (e.g., orbital forcings) perturbation.

2.6. Materials and Methods |

2.6.1. Bathymetry Reconstructions | Bathymetry reconstructions were assembled over the last 80 Myr, consisting of the major evolving deep and shallow marine components. Similar to previous bathymetry reconstructions (85), we use the age-depth relationship of Crosby and McKenzie (38; 64) to convert the age of the oceanic lithosphere from paleo-isochrons to compute the thermal evolution and subsidence of the oceanic lithosphere. These reconstructions were further modified by adding isostatically compensated global paleodeep marine sediment distributions using a relationship based on present-day sediment thickness and ocean crust age (67; 105). Continental shelves, flooded continental crust, and deep marine not represented in the plate motion model reconstructions were supplemented by paleo digital elevation maps (106). Bathymetry was then corrected for eustatic sea-level using the long-term Haq87 sea level curve (107), which is broadly consistent with paleoDEMs shoreline reconstructions. The evolved global bathymetry distributions were deformed based on present-day model uncertainty to reproduce accurate global ocean container volumes throughout the reconstruction period (Fig. A.12).

In consideration of reconstruction complexities and covarying shallow and deep bathymetry, we also produced a set of synthetic bathymetry models to compare with reconstructions. These consist of a roughly Gaussian type distribution ($\sigma = 500$ m) of deep bathymetry and varying uniform shallow bathymetry (≤ 600 m) areal extents within realistic ranges predicted over the reconstruction period. Synthetic bathymetry has no basin distribution uniqueness (i.e., same depths in each basin) in contrast to geologic reconstructions.

2.6.2. C Cycle Modeling | The bathymetric effect on the carbon cycle was evaluated by a narrowly modified Earth system model: long-term ocean-atmosphere-sediment carbon cycle reservoir (LOSCAR). LOSCAR models ocean-atmospheric chemistry through ocean circulation, mixing, air-sea gas exchange, degassing, weathering, and a biological pump within

a box model representing the atmosphere, a high latitude ocean, and ocean basins with 3 water column boxes (A.1.1). Since LOSCAR and variants have been used in numerous studies for the evaluation of paleo climates (84; 87; 108; 109), we use it to examine the influences of bathymetric components on ocean chemistry. We expect broadly similar behavior in other intermediate complexity ESMs, namely GEOCARBSULF variants (110), COPSE (111), GEOCLIM (81), and DCESS (112).

The default present-day LOSCAR model setup, parameterizing the Atlantic, Indian, and Pacific oceans, was used to calculate our synthetic and evolving bathymetry steady-state parameters (Figs. 2.2 and 2.3). Since LOSCAR describes ~ 500 m sediment box intervals from 100 m to 6,500 m, we parameterize our bathymetry distributions likewise for accurate comparison with the existing present-day LOSCAR bathymetry and associated parameter calibrations. Ocean floor area and volume were also evolved in LOSCAR with each reconstruction (A.1.2). The area of the high-latitude box represents cold surface waters globally, not purely high-latitude ocean floor. However, for simplicity, we describe the high-latitude bathymetry as the 10% (by area) of ocean floor at the highest latitudes following the original LOSCAR setup. For the majority of our calculations, we turned off calcite and silicate weathering feedbacks to determine the steady state determined primarily by ocean carbonate chemistry without coupled $p\text{CO}_2$ —weathering flux changes.

The default Paleo LOSCAR model setup, parameterizing the present-day and Tethys oceans as four ocean basin boxes, was used to calculate 60 Ma bathymetry steady-state parameters (Fig. 2.4). The model setup is the same as described for synthetic and evolving bathymetry runs except the silicate and carbonate weathering feedbacks are included and riverine shelf partitioning is reduced by 50%. We impose the riverine shelf partitioning condition to stabilize the Earth system model.

2.7. Data, Materials, and Software Availability | The source codes, models, and Jupyter notebooks for reproducing figures and results in this paper are available at

<https://doi.org/10.5281/zenodo.10460278> (113).

2.8. Acknowledgments | We would like to thank Richard Zeebe for providing the Earth System model code, LOSCAR. This project is supported by the NSF under grant EAR-1900633 to C.L.-B. C.L.-B. was also supported by the Louis B. and Martha B. Slichter Endowed Chair in Geosciences fund.

The contents of this chapter appeared in The effects of bathymetry on the long-term carbon cycle and CCD (11)

Chapter 3

The Influence of Dynamic Topography and Non-hydrostatic Geoid on the Long-Term Carbon Cycle

3.1. Introduction | Are dynamic topography and the non-hydrostatic geoid anomaly key components of the long-term carbon cycle (LTCC)? The LTCC modulates greenhouse gases, influencing atmospheric temperature and consequently global-scale weathering (via increased precipitation and rain acidity), ocean nutrient supply, and biodiversity throughout geologic history (71). The majority of carbon accessible to the hydrosphere is sequestered in seafloor carbonate sediments and lithified rocks (primarily calcite, aragonite, and magnesium carbonate). The evolution and stability of these environments (seafloor bathymetry) and thus sequestered carbon set the stage for the LTCC within which other long-term evolution and catastrophic (e.g., large igneous province forming volcanism) events occur (e.g., (114; 115; 116)).

Solid Earth processes are undoubtedly the primary drivers of topographic evolution. Nearly all seafloor surface area is continuously reshaped as a result of ocean lithosphere formation, cooling, and age-dependent subsidence (see chapter 1). The nonuniform distribution of ocean lithospheric age—globally and over Myr timescales as a result of mid-ocean-ridge

(MOR) spreading and subduction—is directly expressed in the nonconstant and nonuniform distributions of global bathymetry. Changes in the distribution of seafloor age directly influence deep ocean bathymetry (≥ 1000 m) and indirectly contribute to shallow bathymetry (< 1000 m) through sea level change. Subducting ocean lithosphere is also the primary driver of mantle flow and evolution of the solid Earth over timescales of Myr to 100s of Myr (e.g., (117; 118)). Dynamically supported topography and the geoid are consequences of subducted slabs (high density mantle anomalies) and their induced mantle flow (119). Both the geoid and dynamic topography induced by mantle flow are long-wavelength (>100 km) features, with geoid amplitudes often roughly half to one order of magnitude smaller than dynamic topography (119). The combination of dynamic topography (changes in actual topography) and the geoid (changes in relative sea-level) acts across Earth’s surface, influencing the location and extent of shallow seafloor and coastal margins receiving terrigenous sediment (e.g., (120; 121; 122)).

A simple argument illustrates the potential influences that dynamic topography and the geoid anomaly have on carbon sequestration: If dynamic topography and the geoid anomaly reduce the coastline length of a circular continent from L to L_{new} , then, given the same sediment supply and export to the deep marine per unit coastline length (i.e., the same shelf width $a \ll L$), shallow seafloor area is reduced proportionally to the change in coastline length ΔL (Eq. 3.1). Thus, the location and magnitude of the geoid anomaly and dynamic topography over continental shelves will influence the total available shallow seafloor area for carbonate sequestration. Although changes in sedimentation and erosion also shape coastlines and bathymetry, they are often responses to solid Earth controls on accommodation space and tectonic uplift. For simplicity, I do not consider how these processes vary with geoid anomaly and dynamic topography. However, surface process methods and simulations can quantify the effect of dynamic topography on margin sedimentation (69).

$$\Delta A = \left(\frac{L_{new}}{L} - 1 \right) A \quad (3.1)$$

How does bathymetry evolution affect the long-term carbon cycle? Studies of the effects of global- and basin-scale bathymetry on the long-term carbon cycle can be summarized by two major findings: (1) bathymetry controls horizontal and vertical water column communication through, e.g., ocean gateways, continental margins, and seafloor rugosity, and (2) bathymetry defines basin-specific chemical sinks; both influence global ocean chemistry. The opening of the Drake Passage and Tasmanian Gateway provides a prominent example of the influence of evolving bathymetry on ocean circulation and the long-term carbon cycle. The opening of these gateways and the formation of Antarctic circumpolar circulation have been linked to Paleocene global cooling (e.g., (123)). Furthermore, Komar and Zeebe found that changes in shelf-basin (shallow vs deep bathymetry) carbonate burial, a function of shallow seafloor area and the transport of riverine flux from shelves to the pelagic ocean, could explain million year timescale changes in the carbon compensation depth (CCD) (84). Bogumil et al. (2024) elucidated the influence of the distribution of bathymetry and basins on the long-term carbon cycle (113). I found that increases in the shallow shelf area leads to increased shallow (≤ 600 m) versus deep marine (> 1000 m) carbonate storage, resulting in a CCD shoaling. Additionally, a decrease in the average depth of the deep marine (> 1000 m), raises the seafloor carbon sink, reducing the deep marine alkalinity buffer, and shoals the CCD. An important result of these studies is that bathymetric effects on the carbon cycle are closely linked to, but independent of, the evolution of marine carbonate productivity.

Here, I elucidate the effects of the solid Earth (i.e., dynamic topography and the geoid) on the long-term carbon cycle by quantifying their contribution to shelf area and average deep-marine depth. I focus on reconstructing these influences over the past 80 Myr, when constraints (e.g., shoreline location and shelf area) are better known and simulations of dynamic topography and the geoid are more robust. I then qualitatively extend my interpretations to consider the effects of solid Earth processes in deep time.

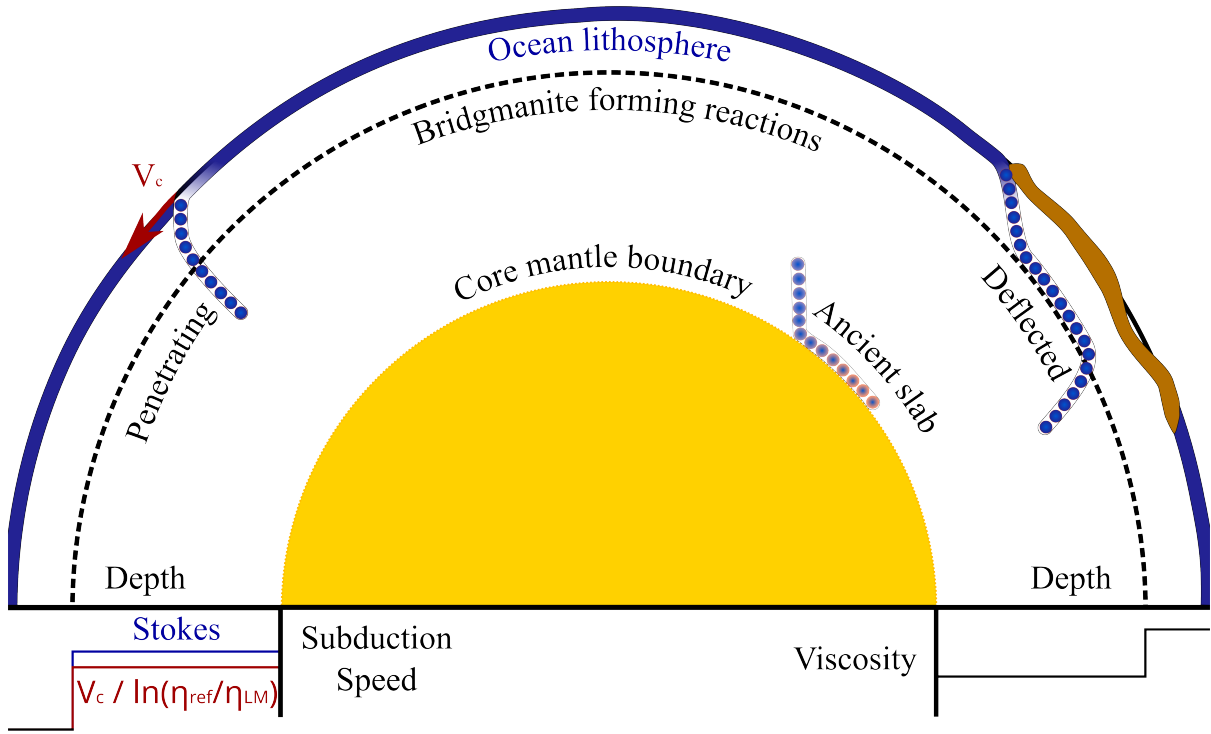


Figure 3.1: Diagram of kinematic subduction from *slabmodellib*. *slabmodellib* uses plate motion modeled surface velocities, the age of subducted seafloor, the mantle viscosity structure, and parameters constraining the dynamics of common and uncommon slab geometries (penetrating, deflecting, and orphaning). Outputs include slab location and lithospheric properties that can be used in post processing calculations of the non-hydrostatic geoid anomaly and dynamic topography.

3.2. Mantle and Surface Simulations | I created a new framework for reconstructing subduction history to evaluate its surface expressions (Fig. 3.1) and their effect on the LTCC. I use direct constraints from recent plate motion and seafloor age reconstructions spanning the last 250 Myr (38), along with trench locations, to construct the framework for my Python library (*slabmodellib*) that reconstructs the kinematic evolution for the fate of slabs. I built on previous work (119; 124) by incorporating (i) new plate reconstructions, (ii) the effects of phase transitions on physical properties (see A.1.4), and (iii) changes in slab buoyancy due to thermal assimilation during mantle sinking. Next, I use the reconstructed subduction histories to calculate the evolution of dynamic topography and regional relative sea

level through the non-hydrostatic geoid. Finally, I validate the present-day mantle structure by comparing the modeled and observed present-day non-hydrostatic geoid.

In this section, I describe my subduction model framework and parameter choices for reconstructing subduction history.

3.2.1. Subduction, Mantle Flow, and Surface Expression | My kinematic reconstructions use locations and rates of subduction from a well-accepted plate motion reconstruction that are supported by the most recent measurements and reconstructions of ocean crustal ages (38). Although I use this reconstruction to describe my methodology, alternative plate motion reconstructions can also be implemented, including models extending back to 1 Ga (37; 38; 125). I interface these plate motion reconstructions with pyGPlates (126), a Python package that facilitates the organization of and calculations using formatted plate motion models. Using these inputs, I calculate subduction history.

slabmodellib treats slabs in a fluid dynamical limit (119), meaning that individual slablets do not interact with one another. The most basic construction is simple: (1) discrete slablets drop vertically into the mantle at the trench every 1 Myr and sink at their convergence speed (the overriding plate’s trench normal velocity) in the upper mantle. Any slab dip between slablets is the result of trench migration. (2) At viscosity interfaces the speed of a slablet is modified by a factor proportional to the natural logarithm of the viscosity ratio between the upper mantle and current layer (Eq. 3.2), following (119; 124). However, users can select additional processes that modify slablet trajectories and consequently the total slab geometry.

slabmodellib can impose a slab morphology to represent penetrating, deflected (see A.1.4), or orphaned slabs. The upper mantle slab dip angle parameter defines partitioning of the convergence velocity to radially and laterally subduction. An alternative lower mantle slablet sinking rate can be selected too; it is calculated using Stokes terminal velocity (Eq. 3.3), with non-thermally assimilated lithosphere density ($\Delta\rho$), constant gravity ($g = 9.81$ m/s), lithospheric half-thickness (R), and depth dependent viscosity (ν_{layer}). Users can also merge

multiple slab models, for example, by assigning penetrating slabs to some subduction zones and deflected slabs to others. Regardless of the chosen processes, the *slabmodellib* outputs retain surface properties relevant to mantle heterogeneity calculations (e.g., seafloor age and subducted lithosphere volume (Eq. 3.4, 3.5), where L_{trench} is sampled trench length (50 km) and T_{lith} is lithospheric thickness calculated as in previous work (124)). Finally, *slabmodellib* uses slablets to produce a 3-D density field for computing induced viscous mantle flow, associated geoid anomalies and dynamic topography, and surface plate velocities.

$$v_{layer} = \left(\ln \left(\frac{\nu_{layer}}{\nu_{ref}} \right) \right)^{-1} * v_{convergence} \left[\frac{km}{myr} \right] \quad (3.2)$$

$$v_{stokes} = \frac{\Delta \rho g R^2}{3 \nu_{layer}} \quad (3.3)$$

$$v_{layer} = (L_{trench} * T_{lith}) * (\Delta t * v_{convergence}) [km], \Delta t = 1 Myr \quad (3.4)$$

$$T_{lith} = \sqrt{\pi \tau * 3.154 * 10^{13}} * 10^{-3} [km], \tau \in [20, 100] Myr \quad (3.5)$$

Subducted density heterogeneities drive mantle flow and, consequently, dynamic topography and the non-hydrostatic geoid. The densities of the slablets are determined using a two-layer lithosphere (Eq. 3.6), where $\Delta \rho_{crust} = -400 kg/m^3$ and $\Delta \rho_{lith} = 80 kg/m^3$. When the age of a subducted slablet is unknown, *slabmodellib* assigns it the global average age of seafloor subducted during the current timestep. *slabmodellib* excludes slablets that reach the core-mantle boundary or are temporarily deflected at the ringwoodite-bridgmanite phase transition from post-processing flow calculations, assuming that they are dynamically supported and do not contribute to mantle flow. In post-processing, *slabmodellib* also applies an additional linear reduction in slablet density as a function of subduction depth to simulate the loss of negative buoyancy from slablet thermal assimilation, ranging from 0% (total buoyancy

reduction) to 100% (no buoyancy reduction). This post processing slablet density adjustment affects only the instantaneous flow calculation, not previously calculated subduction history. Additionally, users can write more sophisticated functions into the *slabmodellib* python library, making use of the main processing outputs, to model this process.

$$\Delta\rho_{slab\text{let}} = 7km * \Delta\rho_{crust} + (T_{lith} - 7km) * \Delta\rho_{lithosphere} \quad (3.6)$$

I can now compute instantaneous mantle flow every 1 Myr (resolution of most modern plate motion models) given subduction history represented through layers of mantle density anomaly fields. I treat the mantle as Newtonian and incompressible, solving for conservation of mass and momentum (Eq. 3.7, 3.8). I follow well-established methods of solving layered and spherical harmonically expanded velocity fields (see (127) for derivations) which incorporate mantle viscosity and density layering. The model obtains velocity and stress fields by decoupling toroidal and poloidal fields into a system of first order differential equations. These equations are solved using an analytical formulation: propagator matrices that communicate flow between spherical shell layers. Users can select the maximum spherical-harmonic degree and order, as well as the degree and order at which tapering begins, to reduce overfitting and the expression of Gibbs phenomena (i.e., “ringing” artifacts). The model then calculates the non-hydrostatic geoid anomaly from mantle density heterogeneities and surface deformation produced by the calculated mantle flow (Eq. 3.9) (128). Then Earth’s surface and core dynamic topography are calculated given the radial stress field σ , gravity g , and the density contrast across the free surface (e.g., $\Delta\rho_{surface} = 2500kg/m^3$)(Eq. 3.10). *slabmodellib* makes use of a modified implementation of these methods from previous work (124).

$$\nabla \cdot \vec{v} = 0 \quad (3.7)$$

$$0 = -\nabla \vec{p} + \nu \nabla^2 \vec{v} + \delta \rho g \quad (3.8)$$

$$\nabla^2 \vec{U} = 4\pi\rho G \quad (3.9)$$

$$h = \frac{\sigma_r}{\Delta\rho g} \quad (3.10)$$

With surface velocity, the non-hydrostatic geoid anomaly, and density fields expanded in spherical harmonics, users can calculate correlations with measured or modeled plate velocities, geoid anomalies (EGM84), and seismic tomography fields converted to density anomalies. I focus on optimizing and evaluating the model against geoid anomalies because this study primarily concerns dynamic topography and non-hydrostatic geoid anomaly calculations. I compute the present-day correlation between the spherical-harmonic expansions of modeled (T) and measured (S) geoid anomalies using Eq. 3.11, where l and m represent spherical-harmonic degree and order.

$$C = \frac{\Sigma_{l,m} T_{lm} S_{lm}^*}{\sqrt{\Sigma_{l,m} |T_{lm}|^2 \times \Sigma_{l,m} |S_{lm}|^2}} \quad (3.11)$$

3.2.2. Subduction History | Parameter Choice | I calculated the subduction, non-hydrostatic geoid anomaly, and dynamic topography history over the past 250 Myr using the Müller et al. (2019) (38) plate motion model and a viscosity ratio of 50:1:100 corresponding to the crust, upper mantle, and lower mantle. All subducted slabs deflect due to the bridgmanite-ringwoodite phase transition. I determined the deflection periods for individual slablets from my look-up table (see A.1.4 and Figure B.1). I applied a linear reduction in slablet density after calculating the position of each slablet; I reduced slablets to 80% of their initial density anomaly throughout the mantle and assigned a density of 0 to slablets that reach the CMB, thereby isostatically compensating them. Finally, I calculated the non-hydrostatic geoid anomaly and surface dynamic topography over the past 80 Myr in 1 Myr increments, to spherical-harmonic degree and order 20, with tapering beginning at degree 15.

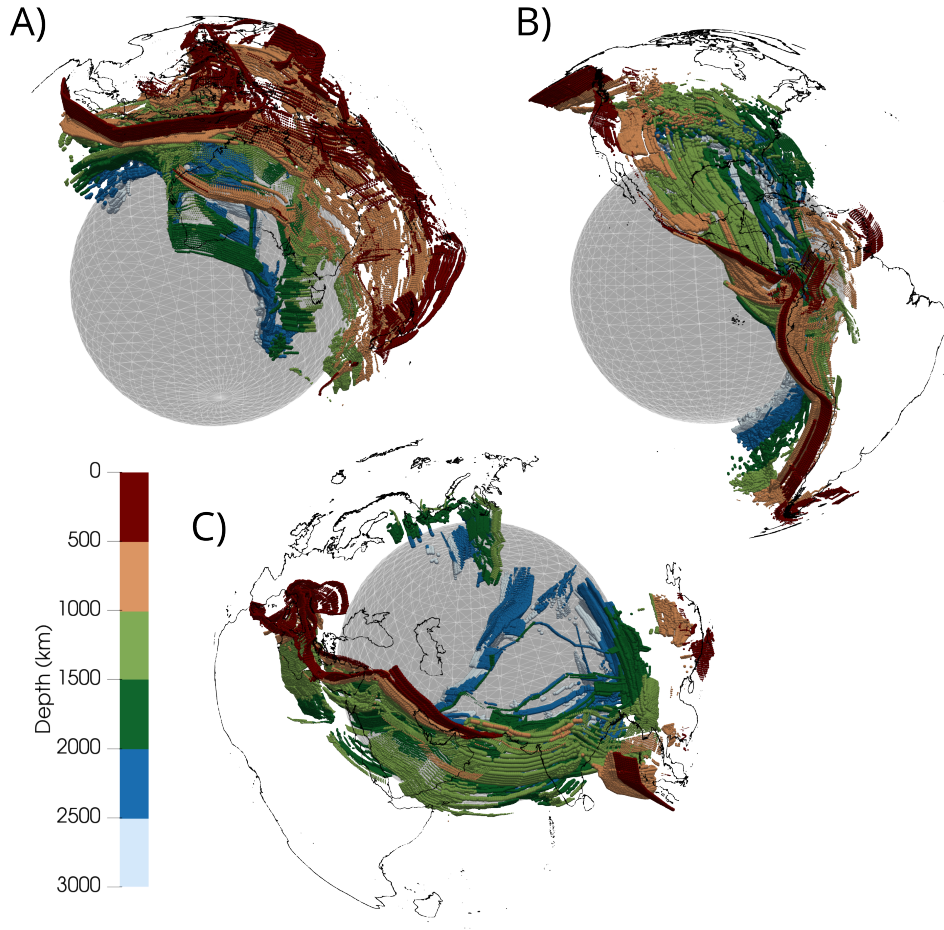


Figure 3.2: Full mantle volume rendering of subducted slabs at present-day with regional highlights under **A)** West Pacific, **B)** East Pacific, and **C)** ancient Tethys ocean. The slabs are shown as spheres scaled to their modeled volumes.

3.3. Results | Figure 3.2 shows the present-day mantle slab geometry given my kinematically defined subduction history. The rendering shows slabs at their modeled volumes beneath all major subduction zones active during the past 80 Myr. Each of these density heterogeneities contributes to the history of mantle flow and surface deformation.

3.3.1. Subduction history | Dynamic Topography & Geoid | At present-day, my model predicts a correlation coefficient of 0.68 for the non-hydrostatic geoid anomaly. This is lower than in previous kinematic subduction frameworks (e.g., (119)) and likely stems

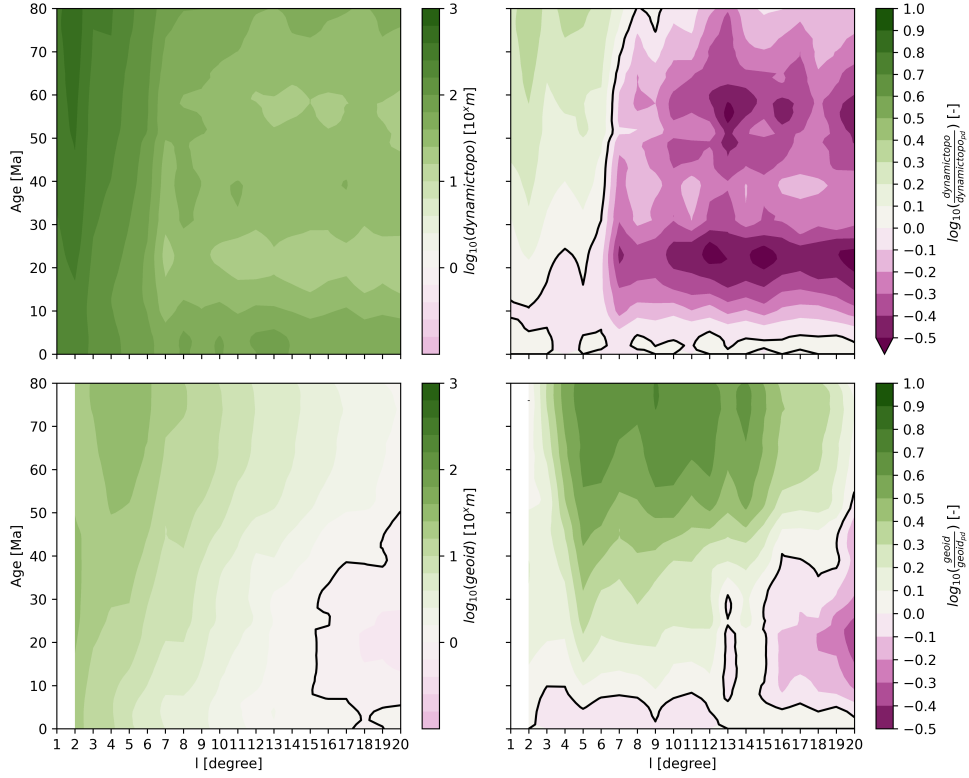


Figure 3.3: Amplitude of spherical harmonic expansions of the non-hydrostatic geoid and dynamic topography over the last 80 Myr and initiated 250 Ma. Panels **A)** and **B)** show the amplitudes of dynamic topography and the non-hydrostatic geoid anomaly, respectively, while panels **C)** and **D)** show those amplitudes normalized to their present-day values. Black lines show the zero value contour. Past dynamic topography had larger amplitudes at wavelengths greater than continental scales (≥ 6000 km). The geoid was generally higher amplitude down to 2000 km length scales. An exception occurred from 5-40 Ma, when subcontinental-scale amplitudes were smaller.

from the combination of an updated plate motion model (i.e., changes in the position of subduction zones and convergence rates), seafloor age reconstructions resulting from new data and magnetic reversal picks, and changes in the kinematic framework (see methods). I do not attribute the lower correlation primarily to the updated subduction framework because models with similar mantle characteristics also fail to reproduce correlations greater than 0.8.

The ranges of the geoid anomaly and dynamic topography are on the order of 100 m and 1 km, respectively. With a viscosity structure that reproduces a realistic geoid magnitude, the model overestimates dynamic topography amplitude relative to residual topography (± 1

km, e.g., (129) and references therein). Therefore, I reduce the dynamic topography field by 50% when analyzing its contribution to topography resulting in a conservative peak-to-peak amplitude of 1.5 km.

A primary result of my subduction-history reconstruction is that the structure and magnitude of both the non-hydrostatic geoid anomaly and dynamic topography have evolved substantially over the past 80 Myr and have evolved differently from one another. The histories of the non-hydrostatic geoid anomaly and dynamic topography are more readily analyzed when collapsed into spherical harmonic degree magnitudes (Fig. 3.3). My subduction history indicates decreasing amplitudes in the low-degree geoid anomaly and dynamic topography harmonics ($l \leq 8$, continental scale) since 80 Ma. The non-hydrostatic geoid anomaly also decreases in amplitude at subcontinental scale harmonics ($l \geq 9$), except that high-degree harmonics ($l \geq 15$) increase in amplitude toward the present over the past 40 Myr.

3.3.2. Dynamic Topography & Geoid’s Contribution to Topography | I estimated the contributions of dynamic topography and the non-hydrostatic geoid anomaly to reconstructed topography (106) by subtracting both fields from the paleoDEMs. Then, I calculated the mantle contribution over the past 80 Myr at 5 Myr increments.

First, I calculate the global coastline lengths of 1) paleoDEMs, which inherently include mantle influences to topography, with 2) paleoDEMs after subtracting the non-hydrostatic geoid anomaly and dynamic topography fields (Fig. 3.4). I find that mantle influences have tended to decrease global coastline lengths by $\sim 5\text{-}15\%$ since 80 Ma. However, global coastline lengths increased due to mantle influences at 50 Ma (by $\sim 1\%$) and from 75-80 Ma (by $\sim 5\text{-}10\%$). Following the simple argument described in the introduction, I would expect to see similar percentage decreases in total shelf area. I can evaluate this argument by analyzing the area versus depth distribution structure of the global topography.

I found that dynamic topography and the non-hydrostatic geoid influence key components of seafloor area (i.e., the shallow shelf) throughout the past 80 Myr. I calculated global

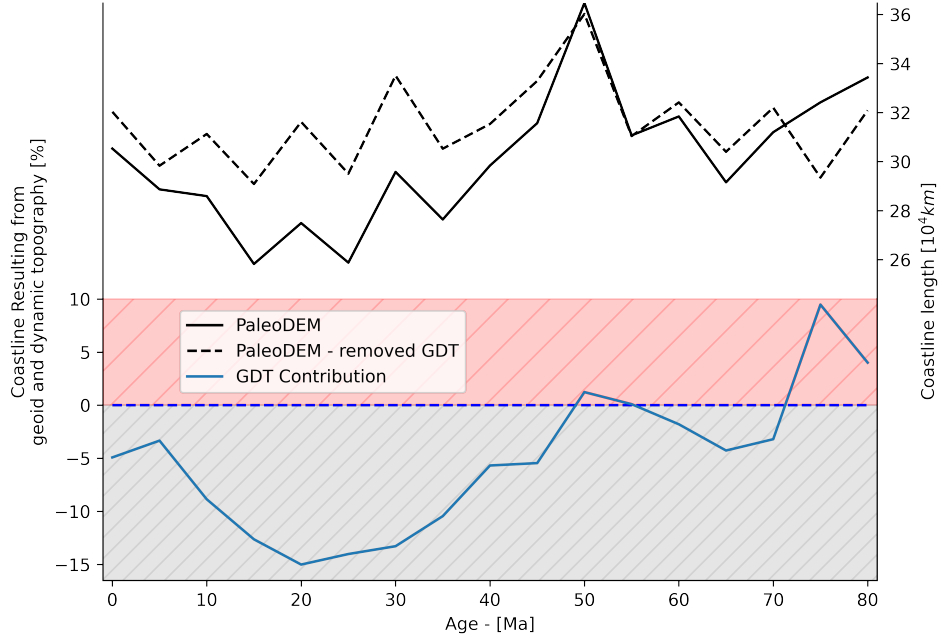


Figure 3.4: **Top)** Coastline length calculated with paleoDEM with and without the contribution of dynamic topography and the non-hydrostatic geoid anomaly. **Bottom)** Percent change in coastline length resulting from dynamic topography and non-hydrostatic geoid anomaly contributions. Solid Earth influences on coastline length tend to decrease coastline length during most of the past 70 Ma.

surface area within the depth intervals $[200,0)$, $[0,-200)$, $[-200, -1000)$, and $[-1000, depth_{max})$ for the paleoDEMs (top of Fig. 3.5), as well as changes within these bins after subtracting dynamic topography and the non-hydrostatic geoid (middle and bottom of 3.5). I also found that mantle influences tended to decrease lowlands (by 0 to 2%), intermediate marine area (by 2 to 8%), and deep marine area (by 0 to 2%). However, the shelf area was increased (by 0 to 4%). Furthermore, the period of maximum mantle influence on coastlines (5 – 40 Ma) does not coincide with the period of greatest mantle contribution to shelf area. Larger contributions to shelf area are most prominent between 40 and 80 Ma. Additionally, periods of major mantle influence on the global bathymetry distribution are not temporally aligned with major mantle influences on global coastline lengths (Fig. 3.5). The nonlinear response of global coastline length and shelf area to mantle influences was not predicted by my simple

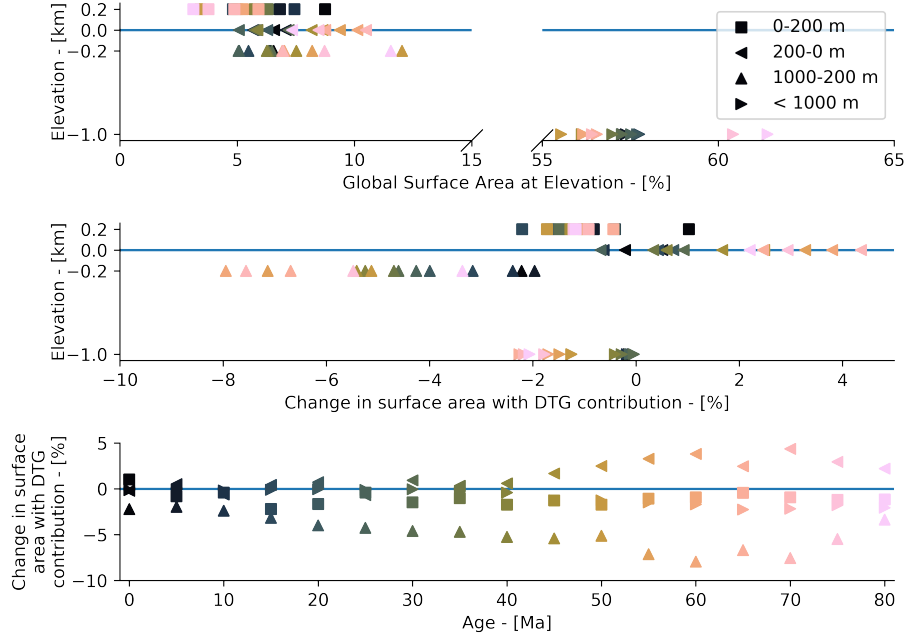


Figure 3.5: The effect of the non-hydrostatic geoid anomaly and dynamic topography on low topography and bathymetry throughout the last 80 Myr and binned in the following ranges [200,0), [0,-200), [-200, -1000), and [-1000, $depth_{max}$). **Top)** Binned global surface area in intervals as calculated from topography/bathymetry reconstructions. **Middle)** Changes in global surface area within each bin due to the contribution of the non-hydrostatic geoid anomaly and dynamic topography. If the change in surface area of all shown bins are summed then they will represent the change in surface above the low lands (>200 m). **Bottom)** Results from the middle panel shown as a function of age.

analytical expression (Eq. 3.1).

In my analysis, the shelf area also decreases more ($\sim 0-4\%$ of global seafloor) than the analytical solution predicts. This discrepancy is partially due to (1) coastline complexity (i.e., real continents are not circular), and (2) changes in shelf width between the reconstructed topography and topography with the dynamic topography and the geoid anomaly contribution removed. If shelf widths remain equal when removing the contribution of dynamic topography and the geoid anomaly then we can estimate the change in shallow seafloor area with the analytical expression. Given the modeled changes in coastline length, dynamic topography and the geoid anomaly contribute to a $\pm 1\%$ (i.e., $\pm 5 \times 10^6 km^2$) change in shallow seafloor

area over the past 80 myr.

I performed additional calculations using M1 and M7 dynamic topography models (24) to evaluate the robustness of my seafloor-area results (Appendix B; Fig. B.2, B.3). These models agree with my kinematic framework and predict a similar magnitude of dynamic-topography effects on shallow seafloor area. All dynamic topography models show similar trends in their contribution to shallow seafloor (200-0 m) and intermediate bathymetric depths (1000-200 m) from 80 to 40 Ma. The M1 and M7 models predict an increasingly strong contribution to supporting low lands (0-200 m) and reducing intermediate (1000-200 m) to deep (<1000 m) bathymetry from 30 Ma to present-day.

Ultimately the contribution of all dynamic topography simulations are similar and can be summarized as follows: dynamic topography (1) can either increase or decrease lowland surface area, (2) supports shallow seafloor area by $\sim 0-4\%$, (3) reduces intermediate seafloor area by 2-10%, and (4) reduces deep seafloor area by 0-3%.

3.4. Discussion | I have shown that the evolution of dynamic topography and the non-hydrostatic geoid has substantially changed their contributions to shallow shelf area and coastline length, with consequent effects on the LTCC over the past 80 Myr.

Given that the magnitude of deposition on shallow shelves influences deep marine saturation state and deposition (e.g., (84)), I can infer how these topographic contributions influence ocean chemistry. The analytical solution predicts contributions to shallow shelf area of $\pm 1\%$ of the total Earth's surface area (i.e., $\pm 5 \times 10^6 km^2$) throughout the last 80 Myr. The field calculated changes in surface area are approximately two times greater and are comparable in magnitude to the total area of present-day shelves. Based on Bogumil et al. (2024) (11), changes of this magnitude correspond to substantial variations in carbon-cycle parameters relative to present-day values, including the carbonate compensation depth (hundreds of meters), biological-pump efficiency ($\sim 10\%$), and weathering fluxes ($\sim Tmol\ C/yr$). These changes may also influence other carbon-cycle parameters, including primary productivity

and pCO_2 .

Changes in the deep marine area (0-2%) that result from dynamic topography and the geoid anomaly may initially appear inconsequential. However, this component of solid Earth evolution also contributes to eustatic sea level evolution because the integrated dynamic topography of oceanic and continental regions changes through time (72; 130; 131). Recent efforts have been made to extend these interpretations to the phases of the supercontinent cycle, finding that dynamic topography causes sea-level to increase during continental assembly, dispersion, and transition and decrease during supercontinent stability (132). A similar supercontinent-cycle trend may characterize dynamic topography and geoid anomaly effects on shallow seafloor, either directly through changes across shelf regions or indirectly through eustatic sea-level change (24). If the contribution on shallow and deep seafloor area trend inversely (as suggested by Fig. 3.5), shallow seafloor area may have become less sensitive to changes in dynamic topography since the initial breaking of Pangea. Even with constant sensitivity of shallow seafloor area to the geoid anomaly and dynamic topography, I expect carbon cycle sensitivity to have been greater before the advent of calcite bearing pelagic plankton species (i.e., pre-Late Triassic).

Since the Late Triassic, pelagic calcifiers (open-ocean) plankton species (e.g., *Coccolithophorids*) have increased the export of carbon to the deep marine (133). As a consequence, since the Late Triassic the shallow seafloor has transformed from the primary direct carbon sink (i.e., disregarding remineralization and turbidity currents) to one of two major marine carbon sinks. Thus, the LTCC was likely far more sensitive to solid Earth processes acting on continental shelves before the Late Triassic.

Previous studies have quantified the effects of dynamic topography on continental flooding and sedimentation, providing constraints on LTCC components complementary to my analysis. Many studies investigating the role of dynamic topography in continental flooding and sedimentation have a regional focus; e.g., the formation of the North America's Western Interior Seaway (36; 122) during early Late Cretaceous and north-south tilting of the Australia

since the Eocene (134). These works exemplify strong cases of dynamic topography that, when compounded with smaller global trends, influence the shallow seafloor evolution shown in Figure 3.4 and 3.5. The evolution of dynamic topography has also been used to correct for basin CCD reconstructions (e.g., (4; 135)). These types of analyses use dynamic topography to adjust the water column depths for paleo drill site measurements of carbonate concentrations before estimating the CCD. A key difference in my analysis is that I estimate contributions to the shallow shelf area. This, in turn, influences the partitioning of riverine fluxes between shelves and the deep marine, thereby affecting deep marine calcite saturation state and changes in the CCD. I also account for the influence, albeit smaller, of the non-hydrostatic geoid anomaly. If the geoid anomaly were included in CCD reconstructions, as dynamic topography has been (e.g., (4; 135)), I would expect corrections on the order of 10s of meters. For example, South Atlantic Ocean CCD reconstructions throughout the Cenozoic may be particularly sensitive to geoid anomaly variations of 10s of m over 10s of Myr (e.g., Fig. 2 of (136)).

The choice of dynamic topography and topography model used will influence the estimated extent of flooding, especially in extreme cases such as the conditions contributing to the formation of the Western Interior Seaway. For my study, I developed a kinematic framework for describing subduction history, dynamic topography, and the geoid anomaly. Since these frameworks produce magnitudes and long-wavelength structures similar to those of numerical convection models (24; 137), I expect the choice of dynamic topography history to have little influence on global estimates of its contribution to shallow seafloor area.

Topography, dynamic topography, and the geoid anomaly are poorly constrained before the Late Triassic, particularly during intervals of supercontinent assembly (e.g., Pangea during the Carboniferous). However, the contributions of dynamic topography and the geoid anomaly to shallow shelf area have significantly affected carbon cycle parameters over the past 80 Myr. Additionally, changes to both carbon cycle parameters and shallow seafloor area will influence the extent of marine habitats and, thus, biodiversity. Future work could explore the

effects of evolving dynamic topography and the geoid anomaly, paired with surface evolution modeling during supercontinent assembly and dispersal, to predict when these processes support the growth or destruction of shallow seafloor and the consequent reorganization of marine habitats.

Chapter 4

ExoCycle v1.0.0: A Generalized Framework for Spherical Community Detection and its Application to Defining Global Ocean Basins from Multi-Field Data

4.1. Abstract | Ocean basins are fundamental units for modeling the Earth system, paleoceanography, and the global carbon cycle. However, their boundaries are often defined heuristically, limiting the robustness of reduced-order models and the interpretation of paleo-proxy data, especially in data-limited paleo- or planetary contexts. We present ExoCycle, an open-source Python library for objective, automated community detection on spherical grids. This framework implements novel composite algorithms (e.g., SB-Reduction) that couple efficient partitioning (Leiden/Louvain) with ensemble-based agglomerative clustering for robust boundary detection. A key technical innovation is our Difference Quantile Transformation Cumulative Density Function (DQT-CDF) edge-weighting scheme, enabling the principled analysis of single or multiple, non-normally distributed scalar fields in a large spherical domain. We validate the method using modern bathymetry and temperature/salinity fields, demonstrating that (1) a spatial resolution of 1-2 degrees is necessary to capture critical basin-defining features such as ridges and plateaus; (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 shallow shelf partitions (shaped by sedimentation, sea-level changes, and riverine fluxes) and surface basins (driven by wind and temperature/precipitation). ExoCycle provides a systematic tool for generating physically-grounded, time-evolving basin definitions, enabling the development of next-generation modular intermediate-complexity models for Earth and exoplanet habitability. As a generalized spherical community detection tool, our new framework is also broadly applicable to other non-ocean related domains, including ecology and land processes, atmospheric science, solid-Earth geophysics, and planetary science.

4.2. Introduction | Ocean basins represent the key building blocks in our conceptual understanding of the dynamics of the Earth system, especially when considering the response of the ocean-atmosphere system to changes in the atmospheric & radiative forcings (e.g., volcanic eruptions, orbital parameters). For example, the distribution of bathymetry and the positioning of continents, i.e., the shape (area, volume, location) of ocean basins, are key controls on thermohaline circulation and mixing. This overturning circulation, in turn, connects deep-ocean properties to the atmosphere and determines the evolution of Earth's climate over timescales of 1000 yr (3). To simplify the complex interactions within the 3D Earth system dynamics, many studies treat each "ocean basin" as a cohesive unit, chemically and dynamically, to analyze long-timescale changes in the Solid-Earth-atmosphere interaction. Ocean basins are consequently central to paleoceanographic reconstructions and planetary carbon cycle modeling (e.g., (11) and LOSCAR (12), CANOPS-GRB (138), GEOCLIM7 (22), DCESS II (139), ESBMTK (23)). The seafloor is the largest carbon sink on Earth accessible to the hydrosphere. Changes in basins and their bathymetry distributions thus influence ocean chemical cycling on both basin and global scales (Fig. 4.1c; (11)). This makes

ocean basins crucial for studying past ocean conditions and modeling the carbon cycle.

Despite their importance, there is still no clear and objective way to define and identify ocean basin boundaries, especially when data are limited. Developing an objective approach to define ocean basin boundaries is critical for several important scientific questions and applications, namely:

- *Ensuring the validity and robustness of reduced order models of planetary carbon cycles and ocean-atmosphere dynamics:* Despite the increasing potential of global climate models (GCMs and models like CGENIE; (140)) for understanding paleoclimate and the Earth system’s response to orbital and volcanic forcings, we still need simplified single- or multiple-box models for multi-Myr time evolution and integration. A number of models of planetary evolution and exoplanets still assume a single ocean box, i.e., a single ocean basin (e.g., (111; 141)). Is this a reasonable assumption for a range of bathymetric distributions?
- *Developing tools to analyze multi-scale (spatio-temporal) planetary habitability:* The number, location, and connectivity of ocean basins on Earth and exoplanets determine the global steady-state ocean-atmosphere composition (e.g., seafloor carbon reservoirs), rate and timescales of overturning circulation, and the rate of basin (local) versus global responses due to external climate forcings. These processes directly influence, e.g., pCO_2 and global surface temperatures and consequentially, a planet’s ability to maintain long-term habitability (142). Thus, to determine the probability of maintaining long-term habitability across multiple process timescales (yr-Myr), we need to systematically define ocean parameterizations at intermediate complexity to increase generalization, validity, and computational efficiency of systems. This type of analysis requires an objective, data-limited ocean basin characterization method to define the spatial domain of reduced order volatile cycling models.
- *Comparison with paleoclimate observations:* How do we objectively decide which paleocli-

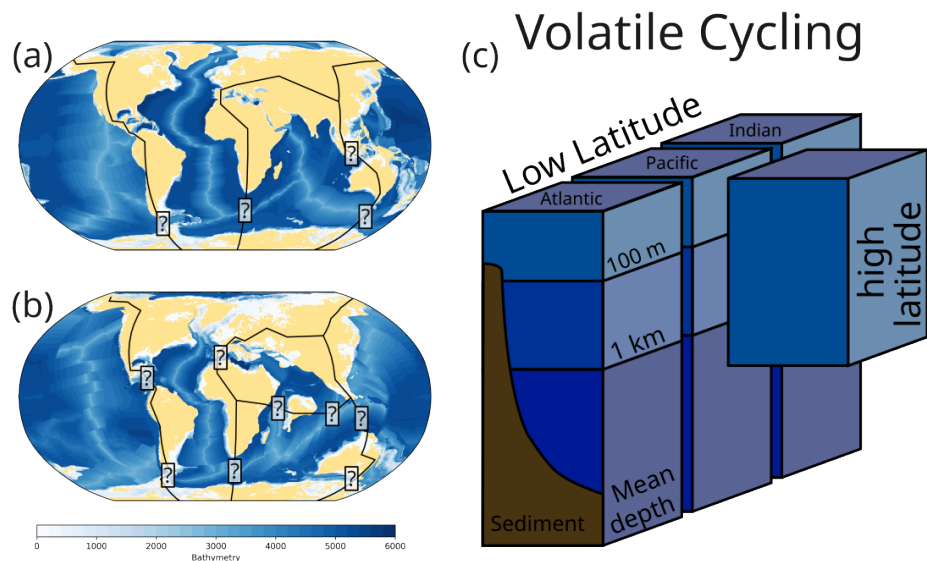


Figure 4.1: (a) Present-day and (b) 60 Ma reconstructed bathymetry from (11). Black lines indicate basin boundaries while question marks indicate unknown and assumed boundaries. (c) Example of intermediate volatile cycling box models (the Long-term Ocean-atmosphere-Sediment CARbon Reservoir model, LOSCAR) used to reconstruct carbon and sulfur perturbation history for the present-day 3 ocean basin system (12; 13). Atlantic, Pacific, and Indian ocean basins are defined as separate boxes with their own ocean chemistry, input and transport of chemical properties, and depositional carbonate zones. A basin's ocean area, volume, and shape of depositional surface are defined by the interior of a basin while communication between basins is a function of bathymetry at the boundaries.

mate observations to aggregate into a single set by moving beyond simple spatial aggregation? Reconstructions and simulations of paleo-oceanic-atmospheric systems heavily rely on a few chemical constraints that need to be averaged over basins that maintain similar properties during climate perturbations. Can basin definitions be left unchanged for 10s of myr as shown in Figure 4.2b-d and numerous studies (33; 34; 143; 144)? The sparse paleoproxy data and the inferences they yield about the global-scale climate state form the basis of our understanding of planetary system evolution on Earth and other Earth-like terrestrial bodies (e.g., Mars). The existing literature lacks clear consensus on how and under what conditions one can reasonably average paleo-ocean chemical measurements to obtain basin-scale estimates of ocean chemistry.

The first step to automate objectively determined ocean basins is having a precise

definition for the question: **What is an ocean basin?** They are **communities** of ocean water that describe a host of similarities in ocean properties, e.g., salinity, temperature, pH, alkalinity, nutrients, trace isotopes, non-locomotive marine life, or seawater origin over ‘relevant’ timescales. Earth has ocean basins since the spatial distribution of ocean properties is neither uniform within a water column (in depth) nor spatially. In practice, basins are often defined heuristically and/or semi-arbitrarily based on our understanding of modern ocean circulation and/or historical definitions of ocean bodies. How do we know where ocean basins are at present, i.e., the boundaries of the Indian, Pacific, Arctic, and Atlantic Oceans? In 1919, a classical definition for present-day ocean basins was formed at the International Hydrographic Conference. It was most recently updated in the 1953 edition of *Limits of Oceans and Seas* (Fig. 4.2a; (14)), wherein oceans (Arctic, Indian, and North and South Atlantic and Pacific) were described as regions of *surface* seawater primarily divided by the equator and continents. These boundaries have been used as a standard for present-day oceans. However, they are of limited value (at best) for paleo- and planetary applications in which continental configurations were very different. Basin definitions primarily evolve on Earth and other terrestrial planets as a consequence of planetary cooling via plate tectonics, leading to a constant redistribution of continents and the seafloor on timescales of 10s of Myr (Fig. 4.1a-b). Thus, we expect that ocean basin boundaries will be time-evolving, to some extent probabilistic entities, indicating regions where there is significant exchange between water masses and changes in water chemistry, temperature, and salinity. The modern interpretation of ocean basins from the seafloor geomorphic features map (GSFM) captures seafloor basins with the isolating effect of bathymetry (145). However, this approach does not necessarily capture the ease of overlaying deep-water communication between these basins. Another basin definition proposed to capture the coupled effect of shallow and deep water used simulated surface ocean dynamics to find regions of attraction and boundaries with low seawater flux (i.e., basin boundaries; (146)). While this method most closely captures our definition of an ocean basin, it becomes increasingly uncertain for describing ocean basins in paleo and

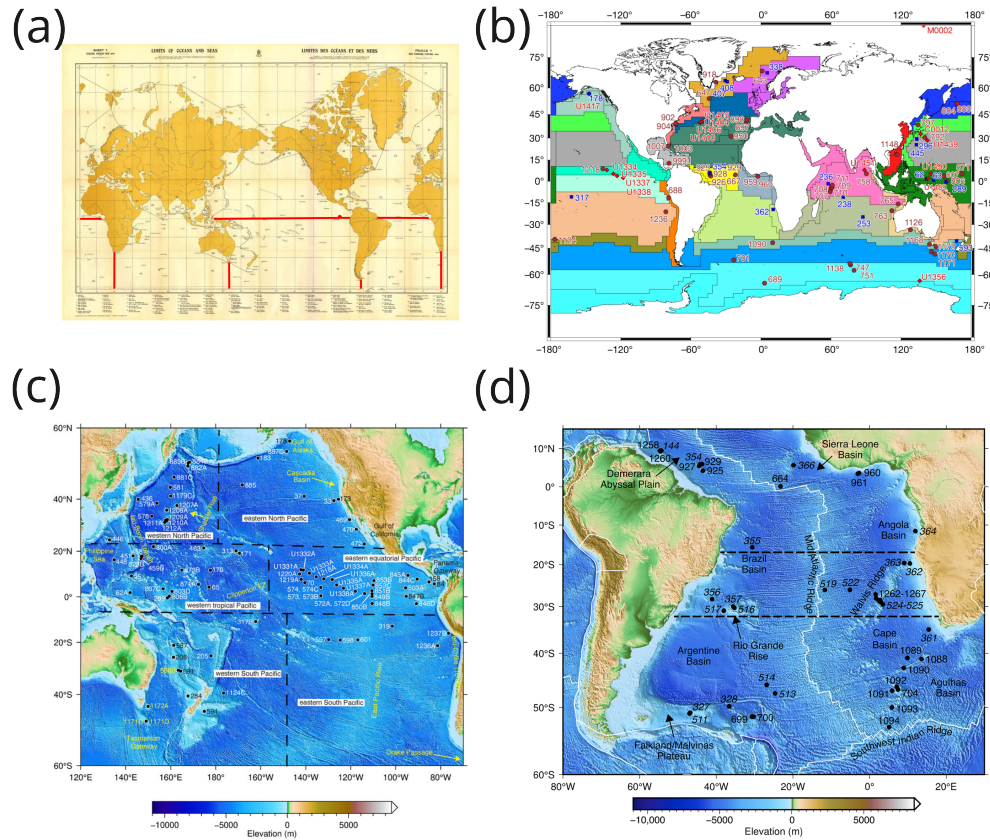


Figure 4.2: Present-day defined ocean basins from (a) Limits of Oceans and Seas (14), (b) Longhurst 2010 (15), (c) Dalvand et al. 2025 (16), and (d) Dutkiewicz et al. 2021 (17) used for paleo-proxy data aggregation. Many of these boundaries differ significantly and have been used to reconstruct regional 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*.

planetary applications, where the nature of ocean circulation is highly uncertain, especially in the deep ocean. Another approach would involve using well-established watershed calculation methods that are currently used to determine the boundaries of land catchment basins (e.g., (147; 148)). Ocean and land basins differ intrinsically in their intra-basin communication. Land basins primarily communicate downslope, while ocean basins can communicate through the overlying water column, potentially in any direction. However, watershed methods rely on the downslope transport (discharge) of properties. An analogous intra-basin transport between neighboring water columns occurs through turbidity currents and solid submarine landslides. However, this mode of material transport does not constitute the majority of intra-basin water communication. At intra-basin scales, network-theory-based methods have been used with great success to determine communities from time-averaged surface water transport (i.e., eddy currents; (149)). However, their implementation of a spatial network theory approach relies on the probability of communication derived from a timeseries of eddy-current surface flows—something that requires high-resolution and accurate global circulation models.

In this study, we propose a novel and general method for objective, automated ocean basin detection applicable to paleo- and planetary studies based on observed scalar fields. For instance, seafloor bathymetry, a scalar field, is one of the few reliable and data-constrained global ocean fields that can be reconstructed throughout the last ~ 100 Myr, extending back to Byr timescales (150) or paleo-planetary fields, albeit with increasing uncertainty (e.g., Mars, exoplanets, (151)). **Our method leverages computationally efficient community detection of scalar fields from spatial network theory to describe ocean basins and shelf partitions for a variety of conditions.** Spatial networks are used to describe the shapes of spaces (with nodes) and their adjacency and connectivity (with weighted edges connecting nodes). A data field such as bathymetry can be defined on a spatial network by creating nodes where bathymetry is available and weighted edges (connections) between adjacent nodes. Scalar fields can be represented on spatial networks with non-directed and

weighted edges that describe node spacing and changes in the scalar field. The detection of a spatial network’s communities can then be conducted with a variety of algorithms (152; 153; 154; 155) that optimize a community quality function to determine regions of similarity (Fig. 4.4). **In summary, our new method incorporates the following novel theoretical components:**

- A spatial community detection algorithm that acts on the surface of a sphere (i.e., can be used to identify long wavelength features), whereas most existing network methods assume a local domain (flat-Earth approximation).
- A node edge weighting scheme allowing for community detection across single or stacked non-normally distributed fields such as bathymetry, temperature, and salinity.
- An open-source pipeline for community (e.g., basin) detection from single or multiple fields on the surface of a sphere.

We have packaged these methods into our Python library (ExoCcycle), which objectively and consistently defines ocean basins for data-limited systems by leveraging spatial community detection methods from graph theory (Fig. 4.3). *With this new objective approach for the definition of an ocean basin, this study then addresses the following scientific questions:*

- What spatial resolution and accompanying features (e.g., oceanic plateaus) are needed to capture the primary structure of ocean basins?
- How to capture the layered structure of the water column at the ocean basin scale and its effects on chemical transport (i.e., continent to ocean) through measurement aggregation? Frequently, in box models, ocean basins are assumed to have vertical walls that extend through the water column (Fig. 4.1c). Is there a better way to capture the non-perfect union between surface, intermediate, and deep ocean basins?
- How different are our estimated ocean basin boundaries for paleo-bathymetry reconstructions for the last ~ 100 Myr? Is it reasonable to use the modern definition of the

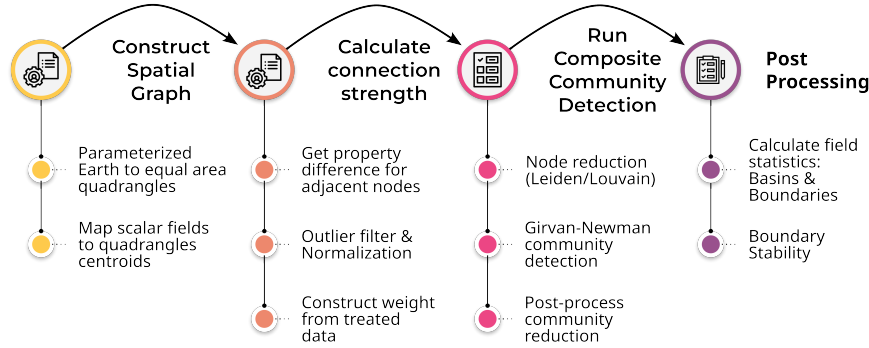


Figure 4.3: Diagram of the main 3 components required for spatial community detections: 1) graph construction, 2) adjacent node connection strength calculation, 3) community detection algorithm. Each sub-category relies on user input that must be appropriate for their final application. ExoCycle is also packaged with some general post-processing capabilities.

ocean basins in intermediate-complexity Earth system modeling?

We address these questions by using our method to define ocean basins and their statistics for 1) present-day and modern reconstructed bathymetry, 2) measured present-day bathymetry below both shallow ocean (neritic, 0-200 m) and deep ocean (bathypelagic zone and deeper, > 1000 m) water columns that capture significant differences in temperature, pressure, light, and water age, 3) paleo-reconstructed bathymetry, and 4) individual and coupled present-day ocean thermochemical fields representative of ocean mixing and circulation across timescales of months to hundreds of years. In Section 2, we describe the main components of our new ocean basin community detection framework. In Section 3, we use the results of our new approach to illustrate the validity of our methodology and results and the specific hyper-parameter choices. In Section 4, we use our new approach to answer the science questions outlined above, and thus highlight the utility of our method for various applications.

4.3. Methods - Basin Community Detection | The spatial community detection method includes the following four primary components: 1) the construction of a

spatial node network, 2) a method to calculate the strength of adjacent node connections, 3) a node clustering algorithm, and 4) post-process operations. In Figure 4.3, we illustrate these components schematically for an overview of the algorithm for ocean basin community detection from the bottom up. In subsequent subsections, we review each of these components, starting by defining the quality functions used in node clustering algorithms.

4.3.1. Quality Functions | Graphs, or networks, are a common approach to model connections in systems ranging from social media to protein interactions (156; 157). Within these complex networks, "communities" are groups of nodes that are much more densely connected to one another than to the rest of the network. Community detection algorithms primarily find these groups by partitioning the graph in a way that optimizes a quality function (158). This function assigns a score to any proposed partition, effectively rewarding solutions that maximize the density of connections within communities while minimizing connections between them. Conceptually, all graph community detection methods aim to optimize a chosen quality function when forming communities. The most popular quality function for these algorithms is the modularity metric, or Reichardt and Bornholdt's Potts model, (Q_m , eq. 4.1) which represents the density of intracommunity to intercommunity connections (159; 160). The equation for modularity compares the number of actual internal links in communities against the number of links that would be expected in a random network with the same degrees. Q_m is defined as:

$$Q_m = \sum_{ij}^N \left[A_{ij} w_{ij} - \gamma \frac{k_i k_j}{2m} \right] \delta(\sigma_i, \sigma_j), \quad (4.1)$$

Here m is the sum of weights in a graph, N is the number of nodes in a graph, A_{ij} is the adjacency matrix representing edges connecting nodes i and j , w_{ij} is a connection's weight, k_i and k_j are the sum of their node's edge weights, δ is the Kronecker delta function, σ_i and σ_j are the nodes' community index, and γ is a resolution parameter. $\gamma < 1$ and $\gamma > 1$ produces a Q_m that favors detection of larger and smaller communities, respectively. A higher Q_m

indicates better community divisions. However, the modularity metric has some inherent theoretical limitations, particularly the resolution limit. The modularity function has an intrinsic scale that depends on the total number of links (L) in the network. This causes the modularity optimization to overlook small communities, even if they are very dense and clearly separated from the rest of the network. This can happen for two communities (let's call them A and B) that are connected by only a few links. If the total number of links in the entire network (L) is large enough, the modularity function may score a higher value by merging communities A and B into a single, larger community than by keeping them separate. This causes the detection algorithm to "miss" the true, smaller community structure. Thus, the resolution-limit wherein modularity fails to capture small node-count communities within a graph, has been a significant hurdle for the implementation of Q_m when $\gamma = 1$ (161).

Another less common, but potentially more reliable quality function exists for community detection - Constant Potts Model (CPM) (162). Modularity takes a global view, comparing the observed internal connections to a null model based on the entire network's properties. In contrast, CPM defines a community as a set of nodes where the number of connections within the community is maximized relative to the possible connections within that specific community, without reference to a global null model. It uses a resolution parameter (γ_{CMP}) to control the granularity of detected communities. In contrast to modularity, CPM focuses on local properties of the network structure. CPM is also designed to be resolution-limit-free, meaning it can detect communities of varying sizes, including small ones, without the inherent bias of modularity towards larger communities (i.e., defined with a constant γ).

$$Q_{CMP} = \sum_{i,j} (A_{ij}w_{ij} - \gamma_{CMP}) \delta(\sigma_i, \sigma_j), \quad (4.2)$$

CPM quality (Q_{CMP}) is calculated in equation 4.2, where the resolution parameter (γ_{CMP}) represents a threshold for community division or merger. Higher quality (Q_{CMP}) is achieved when equation 4.3 is satisfied, where p_i is the internal link density of a community and p_{ij} is

the external link density of a community and are calculated with the internal edge weight of the i^{th} community (m_i), the edge weights between the i^{th} and j^{th} community (m_{ij}), and n_i and n_j are the sum of nodes in community i and j . Consequently, $n_i * n_j$ represents the maximum number of connections between i^{th} and j^{th} communities.

$$p_{ij} = \frac{m_{ij}}{n_i n_j} \leq \gamma_{CPM} \leq \frac{2m_i}{n_i * (n_i - 1)} = p_i, \quad (4.3)$$

Practically, lower and higher γ_{CMP} will favor loosely and tightly bound communities respectively. Overall, we can select the resolution parameters (γ and γ_{CPM}) as a hyper parameter to have some control over the desired community size in the optimized model.

Although not strictly a quality function, edge-betweenness centrality ($B_{max}(m)$) is another metric commonly used in community detection algorithms to optimize network modularity (154; 163). The metric ($B_{max}(m)$) can be used to define the most influential edges for node communication within a graph, and is defined as the number of shortest paths that pass through an edge for the set of all node pairings within a graph. It is calculated by equation 4.4, where V is the set of all nodes, $\sigma(n_i, n_j)$ is the number of shortest (n_i, n_j) -paths, and $\sigma(n_i, n_j|m)$ are the number of paths that pass through the edge m . Edge-betweenness centrality's primary advantage is that it can be used with community detection algorithms to produce deterministic communities.

$$B_{max}(m) = \sum_{n_i, n_j \in V} \frac{\sigma(n_i, n_j|m)}{\sigma(n_i, n_j)} \quad (4.4)$$

4.3.2. Community Detection Algorithms | The other critical component needed to determine ocean basin boundaries is to use community detection algorithms that make use of modularity, CPM, or other quality functions to form an optimized community structure. Broadly, these algorithms follow two main strategies: (A) an agglomerative (bottom-up) approach (Louvain and Leiden algorithms) or (B) a divisive (top-down) approach (e.g.,

Girvan-Newman algorithm). The bottom-up approach starts with each node as its own separate community and then greedily and iteratively merges nodes or groups together, choosing the merge at each step that provides the largest immediate increase to a quality function. Thus, communities are built from the ground up. The opposite top-down method starts with the entire network as a single community and systematically removes the most "between" edges—the critical bridges that connect different groups. This effectively causes the graph to progressively disaggregate and reveal its underlying community structure. However, in practice, there are specific computational choices in each of the class of algorithms for efficient community detection in graphs with a large number of nodes. In our analysis, we use the three most commonly used algorithms for community detection: *Louvain*, *Leiden*, and *Girvan-Newman*.

4.3.2.1. Agglomerative Algorithms - Louvain & Leiden | The *Louvain algorithm* initially assigns each node to a unique community and then operates in two steps to optimize the quality function (Fig. 4.4; (152)). *Stage 1 (local moving of nodes)*: all graph nodes are iterated in a random order and moved to neighboring communities if the quality function is maximized ($\max(\Delta Q)$). These iterations are repeated throughout all nodes until an iteration yields little or no increase in the quality function (e.g., $\Delta Q \leq \Delta Q_{threshold}$). Then Stage 1 is terminated. *Stage 2 (network aggregation)*: each group of nodes comprising a community are collapsed into aggregated nodes (i.e., a singleton partition), and then Stage 1 is repeated at this higher level. The algorithm is terminated once $\Delta Q \leq \Delta Q_{threshold}$ and Stage 2 does not lead to further node reduction, or Stage 1 has been repeated N_{max} times. *The Louvain algorithm is ideal for large networks, as the algorithm's computational scaling is $O(n \log(n))$ for n nodes.*

The *Leiden algorithm* differs mainly from the Louvain algorithm by introducing a refinement stage before node aggregation and by improving the local node-movement algorithm (153). Leiden's fast local moving of nodes algorithm does not iterate through all nodes

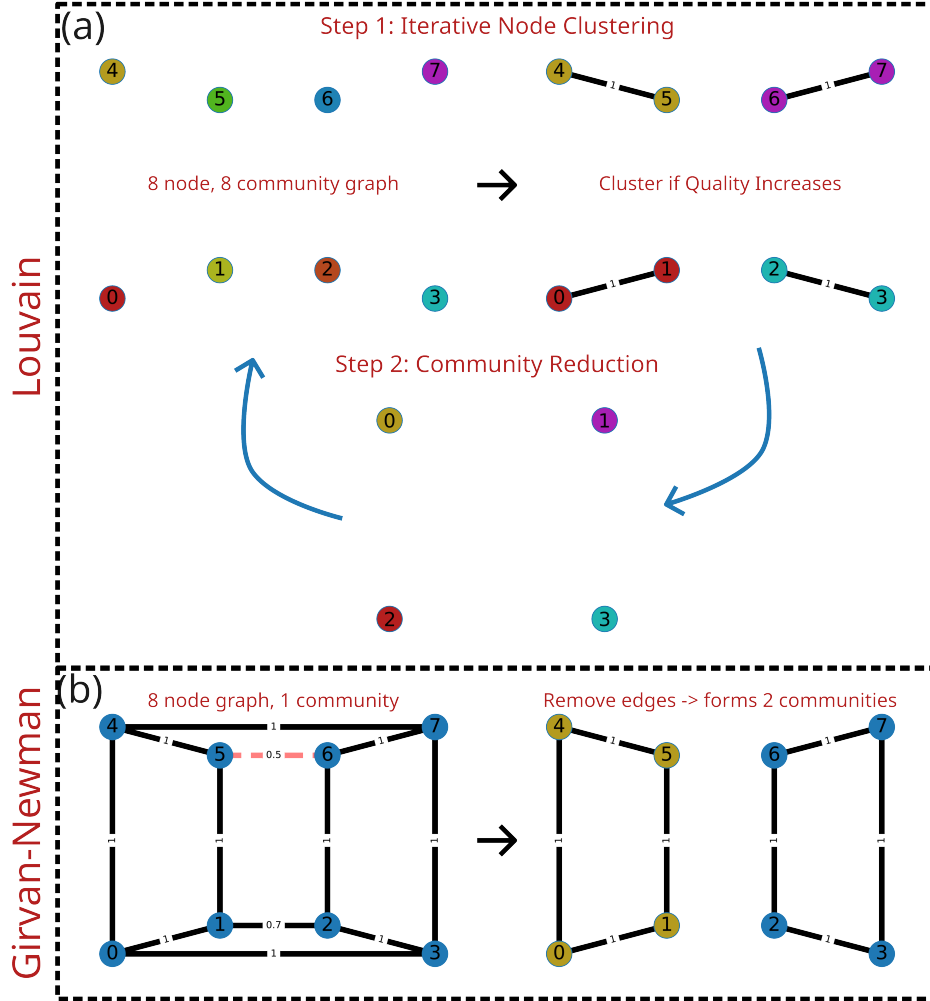


Figure 4.4: Diagram of (a) *Louvain* and (b) *Girvan-Newman* community detection algorithms. The Louvain algorithm is a two stage process that assumes each node forms a community. In Stage 1 the algorithm groups communities (i.e., nodes at this stage) if they improve the networks quality function. Stage 2 merges nodes into new communities before Stage 1 iterates. Girvan-Newman community detection assumes one contiguous community and then iteratively cuts edges with the highest edge betweenness centrality metric until a desired community count is achieved. The resulting community structure can still be evaluated with a quality function.

multiple times. Instead, fast local moving first iterates through all nodes. If a node is moved to a new partition, all neighboring nodes that are not queued within the moved node’s previous partition are added to the end of the iteration queue. The refinement stage occurs when Stage 1 no longer produces mergers that improve the quality function result (e.g., $\Delta Q \leq \Delta Q_{threshold}$). Then, before aggregation, each partition’s well-connected nodes are iteratively clustered into refined partitions. Node mergers in this step are not performed by optimizing the quality function. Instead, they are based on a probability that scales with strictly increasing quality function values (i.e., higher increases in $\Delta Q \rightarrow$ higher probability of merger). The Leiden algorithm is more complex than the Louvain algorithm but does not incur any increase in computational expense, since it uses the fast local moving of nodes. The primary benefit of the Leiden algorithm and its fast local moving algorithm is that they guarantee that detected communities are well connected. For example, consider the Louvain community detection scenario, where the last node movement in Stage 1 improves the quality function but removes the sole connection between two portions of a community. In this case, the Louvain algorithm would lead to poorly connected communities after the node is moved to a new community, and nodes are aggregated.

Both Louvain and Leiden algorithms use a greedy local clustering scheme in Stage 1 to group nodes (Fig. 4.4a). This scheme does not enforce that the highest global modularity gain mergers be commenced first, leading to some stochasticity in convergence. As a consequence, running either algorithm multiple times with different local-clustering sequences can yield different community structures, even in fields with well-defined communities. This scheme may appear to be a problem and an avoidable and negative feature, but it provides insight into the quality of community structure. Modularity optimization is a non-convex problem, meaning that comparing community structures across multiple algorithm runs provides a better description of community boundaries.

To address this practical challenge, agglomerative clustering of Louvain and Leiden community partition ensembles can be used to find stable ocean basins. *We have built*

agglomerative clustering, an unsupervised hierarchical clustering method, into ExoCycle to find the optimal node clustering from an ensemble of communities. Agglomerative clusters are calculated using an average-distance minimization linkage criterion, where distance is defined as 1 minus the normalized co-association matrix of nodes within similar communities across the ensemble. We set the distance threshold to 0.3 for all our calculations. Physically, this can be thought of as follows: clustering terminates when all final clusters have an average cluster similarity of 70% across the ensemble. A lower distance threshold favors an increase in division (more and smaller clusters). An appropriate concern would be that many communities might be identified if the distance threshold was set arbitrarily low (e.g., 0) when the community structure is poor across a large ensemble. However, if few communities are found at low distance thresholds, the boundaries of communities are stable despite the inherent stochasticity of the Leiden and Louvain algorithms.

4.3.2.2. Top-down Algorithms - Girvan-Newman | The *Girvan-Newman algorithm* does community detection by assuming that all nodes are connected within a single community (154). The edge connections are iteratively broken at the highest edge betweenness centrality value (eq. 4.4) until X communities are obtained (4.4b). For example, Figure 4.4b shows a graph that has the highest betweenness centrality for edges between nodes $(n_5 - n_6)$, $(n_1 - n_2)$, and then $(n_0 - n_3)$ and $(n_4 - n_7)$. Edge connections are iteratively broken until 2 communities are obtained. This scheme results in a computational scaling with network size of $\mathbf{O}nm^2$ for n nodes and m edges. Uniquely, this algorithm does not use a quality function to cluster communities. However, the final communities can still be evaluated with modularity, CPM, or other quality functions. We make the common choice of modularity as an evaluation metric when using the Girvan-Newman algorithm.

4.3.3. Spatial Graph & Weight Construction | To apply the algorithms described in the previous sections, we first need to construct a spatial graph representing a dataset(s) on the surface of a sphere in a grid with equal area. We do this by expanding a data field to

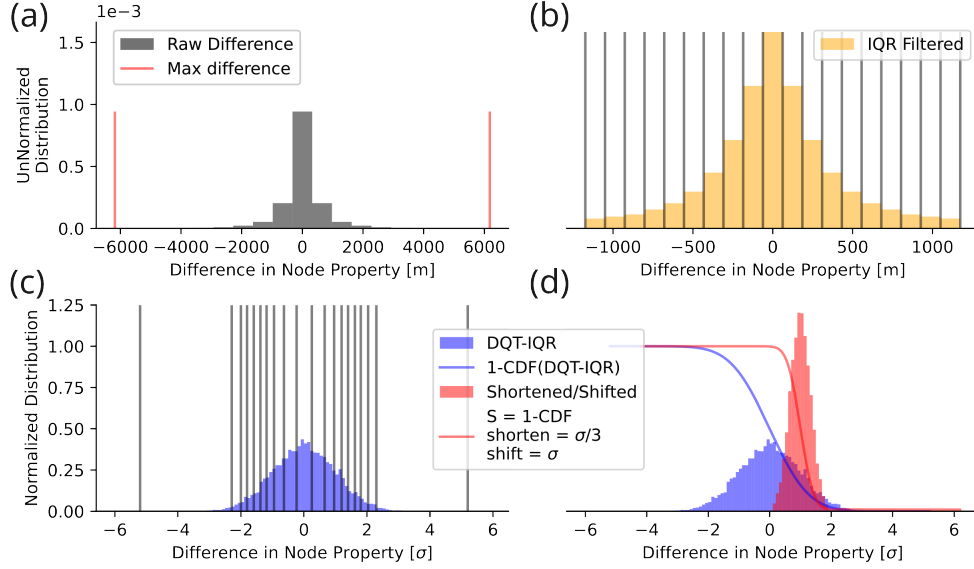
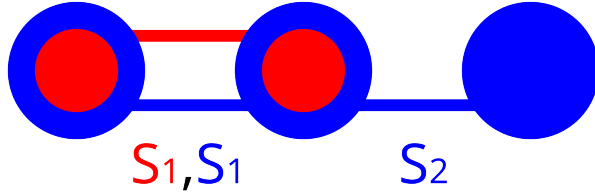


Figure 4.5: Shows the construction of node edge weights using ETOPO1 seafloor bathymetry (9) sampled at 1 degree resolution. Distribution of (a) absolute value difference of adjacent nodes mirrored about 0 value, and (b) interquartile range filtered absolute value difference of adjacent nodes mirrored about 0 value. (c) shows (b) after a normal distribution quantile transformation. (d) Shows (c) distribution, shortened and shifted distribution, and the 1 - cumulative density functions (1-CDF) of each distribution. Weights are inversely proportional to radian separation and proportional to 1-CDF on the red curve for some given absolute value difference of adjacent nodes.

nodes (n) centered in a set of quadrangles of equal area on the surface of a sphere, following the methods from (164). We calculate node positions by incremental rotations between the 4 edges of a dodecahedron face on the surface of a sphere. Then the face is rotated to form the 12 sides of a spherical dodecahedron. Finally, repeated points are removed, edges (m) are drawn between the neighboring nodes, and the data field(s) are interpolated to the node positions using gmt's `grdtrack` function with bicubic interpolation (Fig. 4.7a-b; (165)). The edge weights are then calculated on the basis of the field properties.

In the following, we describe four bidirectional edge weighting schemes and their potential drawbacks to emphasize a best practice for defining edge weights. What should define a strong edge weight reward for a spatial network? The 1) proximity to a neighboring node and 2) similarity in node properties. Correspondingly, the edge weights have two components – weighting the relative node proximity and (dis)similarity of node properties. The common

(a)



(b)

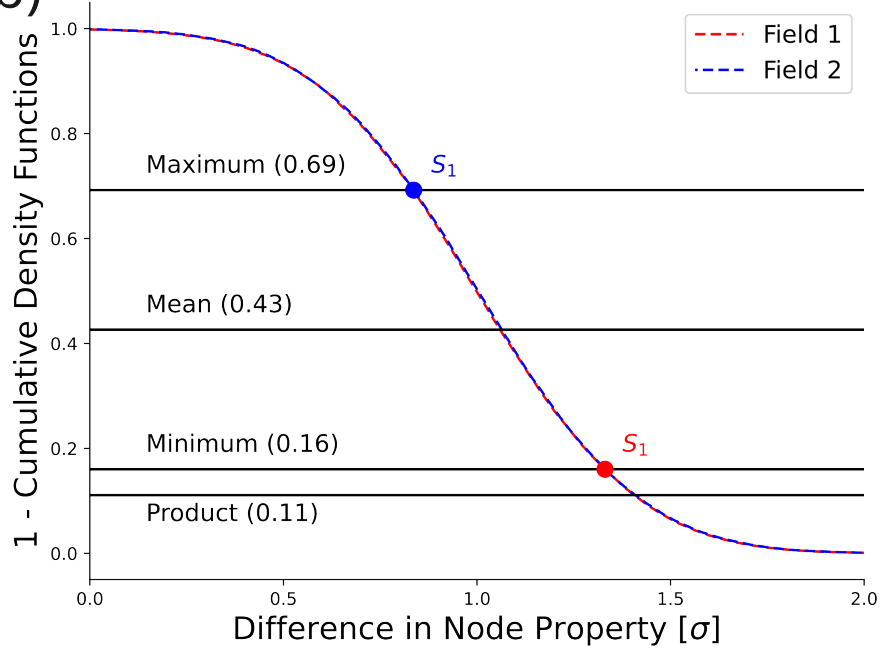


Figure 4.6: Illustration of our multi-field node edge weight construction for graphs. (a) Diagram of nodes and connections for 2 fields (red and blue) that do not necessarily overlap. Red and blue S_1 's represent 1 - the cumulative density function (CDF) that is calculated separately for fields 1 and 2, respectively. (b) 1 - CDF for the shortened and shifted weight distribution as calculated with the difference quartile transformation CDF weighting method. Options for merging fields include minimum, maximum, mean, and product of 1-CDF.

approach to weighting node or data proximity in clustering analysis is to use the inverse or squared-inverse distance weighting of the node (data) separation distance (e.g., (166)). In all weighting methods we adopt for our analysis, we use the inverse separation distance unless noted. For equidistantly spaced nodes, this portion of the edge weight does not change across the network. The (dis)similarity of neighboring node properties is the primary control on edge weights for network analysis.

A common edge weighting scheme in community cluster analysis is gravity centrality (eq. 4.5, (167; 168)). Its original form is based on Newton’s law of universal gravitation and estimates the influence of node i on nodes within some distance (or a specified number of nodes) away from it. We can use an adapted version (eq. 4.5) to determine the gravity edge centrality - interaction ($w_{i,j}$) - between two nodes i and j with property values: P_i & P_j , and square-inverse node separation $\frac{1}{d_{i,j}^2}$. When used in a community detection algorithm (e.g., Louvain and Leiden) with a spatial field (e.g., bathymetry), the high magnitude values (i.e., thick water columns) will preferentially attract lower magnitude values (i.e., shallow and intermediate ocean depths) to their cluster. A major practical pitfall with this approach is that the weighting scheme would fail to capture independent shelf regions that do not communicate well with the pelagic ocean (e.g., flooded continental seas). To avert this, we can impose weights that punish dissimilarity between neighboring node properties.

$$w_{i,j}(n_i, n_j) = \frac{P_i * P_j}{d_{i,j}^2}, \quad (4.5)$$

Equation 4.6 is the simplest bidirectional edge weight description that rewards node proximity and punishes dissimilarity. For example, this weighting scheme would produce strong weights across neighboring abyssal plains or shelf regions and weak weights near mid-ocean ridges and continental rises. We remedy the common appearance of a singularity ($P_i = P_j$) in data fields with a slightly more complex logistic representation (eq. 4.7), where

k and s are defined to pin the minimum and maximum node differences to $w_{min} = 0.1$ and $w_{max} = 0.9$. This new formulation can avoid the singularity and allow users to emphasize node weighting at specified difference values, e.g., those associated with small differences (e.g., shelves and abyssal plains) or large differences (e.g., mid-ocean ridge flanks and continental rises). Another benefit of this scheme is that node property weights are normalized so that multiple fields can be readily merged. A primary drawback of the logistic weighting scheme is the difficulty of anticipating and interpreting how choices of k and s influence the weights across fields, which can be uniform, bimodal, or trimodal.

$$w_{i,j}(n_i, n_j) = \frac{1}{|P_i - P_j|} \frac{1}{d_{i,j}}, \quad (4.6)$$

$$w_{i,j}(n_i, n_j) = \frac{1}{1 + e^{-k|P_i - P_j| + s}} \frac{1}{d_{i,j}} \quad (4.7)$$

We have developed a new generalized difference weighting scheme that avoids the pitfalls of the previous methods: the Difference Quantile Transformation Cumulative Density Function (DQT-CDF) edge weighting method. Weights for this method are calculated as in the following steps: A distribution (D) of the absolute values of all adjacent node differences is created (Fig. 4.5a). Then a filtered distribution (D_{filter}) is created by removing outliers from D , where outliers are less than $Q_1 - (1.5 * IQR)$ and greater than $Q_3 + (1.5 * IQR)$ for the 25th percentile value Q_1 , the 75th percentile value Q_3 , and $IQR = Q_3 - Q_1$ (Fig. 4.5b). A normal distribution (D_{norm}) is created using a quantile transformation (QT) on D_{filter} (Fig. 4.5c). Edge weights (W) are partially constructed from the cumulative density function (CDF) of a shortened (by $\frac{\sigma}{a}$) and shifted (by $b\sigma$) D_{norm} distribution, where σ refers to the standard deviation of (D_{norm}) and a and b are user defined scalars (Fig. 4.5d). The final edge weight W is proportional to a function of the data field and inversely proportional to node spacing in

radians (eq. 4.8). In our proposed approach, multiple fields can be readily merged to create a single edge weight. This merging is done by constructing edge weights for each individual field using equation 4.8. The weights can then be combined using the average, minimum, maximum, or product of weights (Fig. 4.6). The benefits of using the DQT-CDF over the logistic method are the ease of interpreting the shortened and shifted values relative to k and s , and the avoidance of carefully choosing different values for $w_{min} = 0.1$ and $w_{max} = 0.9$. Therefore, we adopt DQT-CDF as the preferred weighting scheme for all of our calculations.

$$W(n_i, n_j) = \frac{1 - CDF(QT(|P_i - P_j|))}{d_{i,j}}, \quad (4.8)$$

4.3.4. Composite Algorithm | We have applied the concepts described in the previous sections to develop two new composite community detection algorithms. Both start with constructing a graph representing a dataset on the surface of a sphere (e.g., seafloor depth, chemical composition, temperature). Second, node edge weights are calculated using a weighting scheme. Our preferred weighting scheme (**DQT-CDF**) is used to produce the weighted graph (Fig. 4.7b). Lastly, we apply our composite community detection algorithm to the graph. Our first composite algorithm (**GN-SB-Reduction**) has the following steps: 1) an ensemble of N Louvain or Leiden algorithms is run, reducing the node and edge count (Fig. 4.7c-d). 2) A new graph is constructed using the reduced community structure detected as the basis for node and edge connections (Fig. 4.7c). The edge weights for this new graph are the sum of edge weights between two connected communities (Fig. 4.7e). 3) The Girvan-Newman algorithm is applied to reduce the community structure further. 4) Post-processing steps for additional community merger are applied (Fig. 4.7f). Our second composite algorithm (**SB-Reduction**) is identical to GN-SB-Reduction, but omits the Girvan-Newman community detection step.

A summary of parameter choices for our community detection algorithm is shown in Table

Table 4.1: Summary of parameters that must be defined for community detection models.

Model Component	Choice	Additional Choices
Inputs	Global Data Field(s)	Averaging method; spatial resolution
Reduction Method	Leiden, Louvain	Quality function $\rightarrow \gamma, \gamma_{CPM}$
	Girvan–Newman	Basin count
	Agglomerative Clustering	Ensemble size; linkage criterion; distance threshold; distance
	SB-Reduction	Set of χ
Edge Weighting	Gravity Centrality	
	Inverse Difference	
	Logistic	$k, s, w_{\min}, w_{\max}$
	DQT–CDF	shortened; shift; minimum weight

4.1.

4.3.5. Post-processing Algorithms | All results shown here use post-processing that merges small basins with neighboring large basins. Small basins are defined as node clusters smaller than a threshold χ . We use all thresholds as percentages of the total field’s coverage. For example, if $\chi = 0.5\%$ for present-day bathymetry, then any communities with a surface area less than $1.8 \times 10^6 \text{ km}^2$ would be merged with the closest community with a surface area greater than $1.8 \times 10^6 \text{ km}^2$. Practically, this allows for merger of small basins (e.g., small mid-ocean-ridge axial segments as well as small coastal basins) that typically do not influence global ocean chemistry. This method can also be applied iteratively with a set of small basin thresholds to generate a sequential set of basins. Unless otherwise noted, all the results shown later include an iterative small basin merger with $\chi = 0.1\%$ and $\chi = 0.5\%$.

4.3.6. Computational Efficiency | The Louvain and Leiden algorithms are ideal for large networks, as their computational scaling is approximately $\mathcal{O}n \log(n)$ for n nodes. The Girvan–Newman algorithm’s computational scaling is $\mathcal{O}nm^2$ for n nodes and m edges; therefore, it is impractical for global community reduction at 1 degree resolution ($n = \mathcal{O}(1e4), m = \mathcal{O}(1e4)$). However, it can be used on a network that is reduced first by the Louvain or Leiden community clustering algorithms ($n = \mathcal{O}(1e2), m = \mathcal{O}(1e2)$).

Table 4.2 shows the computational scaling for different community detection algorithms given variable spatial resolutions on an Intel® Core™ i9 processor 14900HX. As the aforementioned scalings indicate, the Louvain and Leiden algorithms should scale similarly with increased resolution, and they do within our implementation using the `leidenalg` (v0.10.2) and `louvain` (v0.8.2) python package. Girvan-Newman community detection with the `networkx` (v3.3) python package has poor computational scaling and becomes impractical for basin detection at resolutions greater than 4 degrees $n = \mathcal{O}(1e3)$. For example, the Girvan-Newman algorithm will take about 60 days to complete a global basin detection at 1 degree resolution even for single-core clock speeds that generally outperform supercomputers. The alternative Leiden-Girvan-Newman composite method will complete in less than 30 minutes for about 1600 communities detected in the Leiden step. The Girvan-Newman and Leiden-Girvan-Newman results will not be identical, but Leiden-Girvan-Newman provides a consistent community detection approach to identify an *a priori* number of important basins. Additionally, the agglomerative clustering of Louvain and Leiden based ensemble runs has been parallelized to reduce computational times of large ensembles. The total memory demand is ~ 20 GB for single field community detection at 1 degree resolution and $\sim 70\%$ surface coverage. Memory usage scales by about 100 MB per parallel process.

4.3.7. Evaluation Metrics | We use modularity for node clustering within the Louvain and Leiden community detection methods. However, additional metrics can be used after detection to determine the quality of all or individual communities outputted from ExoCycle (e.g., average embeddedness). Which evaluation metrics to use and how they can be interpreted will depend on the user’s application. ExoCycle uses the python package `cdlib` (0.4.0) to calculate some of these clustering metric outputs. However, we note that one needs to be cautious in implementing evaluation metrics that rely on comparing to network null models with random node connections (e.g., the traditional ‘surprise’ metric). Our networks are *spatially gridded* with no data sparsity (i.e., inherently well-structured). Consequently, these

Table 4.2: Computational times and basin count for basin detection given Louvain, Leiden, Girvan-Newman, and Leiden-Girvan-Newman community detection on present-day seafloor bathymetry using non-parallel processes on an Intel® Core™ i9 processor 14900HX. For all models the ensemble size is one. For all Louvain and Leiden based models $\gamma = 0.01$. The Girvan-Newman based community detection restricts the maximum basin count to less than 10. Computation times exclude post-processing steps (i.e., SB-Reduction, interpolation, plotting, calculation of evaluation metrics). Girvan-Newman results are not created for resolutions greater than 4 degrees due to restrictive computational costs. However, theoretical computational times with respect to the 10 degree resolution runs are included (π -values). Basin count values show basins greater than 0.5% of the field’s surface area. Note that computational times for Louvain and Leiden algorithms have some variability due to the random initial conditions.

Model	10° ($n=284 \simeq \frac{m}{2}$)	8° ($n=527 \simeq \frac{m}{2}$)	4° ($n=1603 \simeq \frac{m}{2}$)	1° ($n=24441 \simeq \frac{m}{2}$)
Louvain	~ 0.8 s, 6 basins, $\pi = \text{N/A}$	~ 2 s, 6 basins, $\pi = 1.6$ s	~ 3.0 s, 5 basins, $\pi = 5.9$ s	~ 47.1 s, 16 basins, $\pi = 123.1$ s
Leiden	~ 0.7 s, 6 basins, $\pi = \text{N/A}$	~ 2.1 s, 6 basins, $\pi = 1.4$ s	~ 3.1 s, 6 basins, $\pi = 5.2$ s	~ 49.6 s, 15 basins, $\pi = 107.8$ s
Girvan-Newman	~ 8.3 s, 4 basins, $\pi = \text{N/A}$	~ 36.9 s, 3 basins, $\pi = 53.0$ s	~ 35.8 min, 2 basins, $\pi = 24.9$ min	N/A, N/A basins, $\pi = 61.2$ days
Leiden-GN	~ 0.7 s, 5 basins	~ 2.0 s, 5 basins	~ 3.0 s, 5 basins	~ 48.5 s, 5 basins

metrics will tend to overestimate the quality of community structure.

The **stability of community boundaries** generated from the agglomerative cluster of an ensemble can be assessed with a cohesion metric, e.g., a silhouette score. ExoCycle can calculate node-by-node silhouette scores (eq. 4.9) using a distance value of 1-coassociation of nodes within similar communities from the ensemble, where a is the mean distance of the evaluated node to each node within its community and b is the minimum of the set of average distances to every other community. Silhouette scores range from -1 to 1, where values closer to 1 represent nodes that are almost always found within the agglomerative cluster community. Typically, silhouette values in the intervals $[-1, 0]$, $(0, 0.5]$, $(0.5, 0.7]$, and $(0.7, 1]$ correspond to bad, weak, good, and strong clustering, respectively.

Global-area-averaged silhouette values can be used to assess the overall boundary stability of an ensemble run (i.e., to determine which ensemble size yields a stable result given the input field(s) and weighting scheme). For example, it is expected that the global silhouette value will be 1 for the Leiden community detection of a field with an ensemble size of 1 and concave-shaped communities. As the ensemble size increases, the stochastic nature of Leiden detection will reduce the global silhouette value. Furthermore, averaging multiple ensemble runs of the same size will provide an estimate of the ensemble’s stability. This calculation can be done for increasing ensemble sizes until the decrease in *ensemble-averaged silhouette values* becomes asymptotic (for well-structured fields) or shows a continuous decline (for weak or unstructured fields).

$$S = \frac{b - a}{\max(b, a)} \quad (4.9)$$

4.4. Method Results |

4.4.1. Accuracy of our approach - Stability of communities | Before applying our method to different problems, we first assess the validity of the results. To first order, the

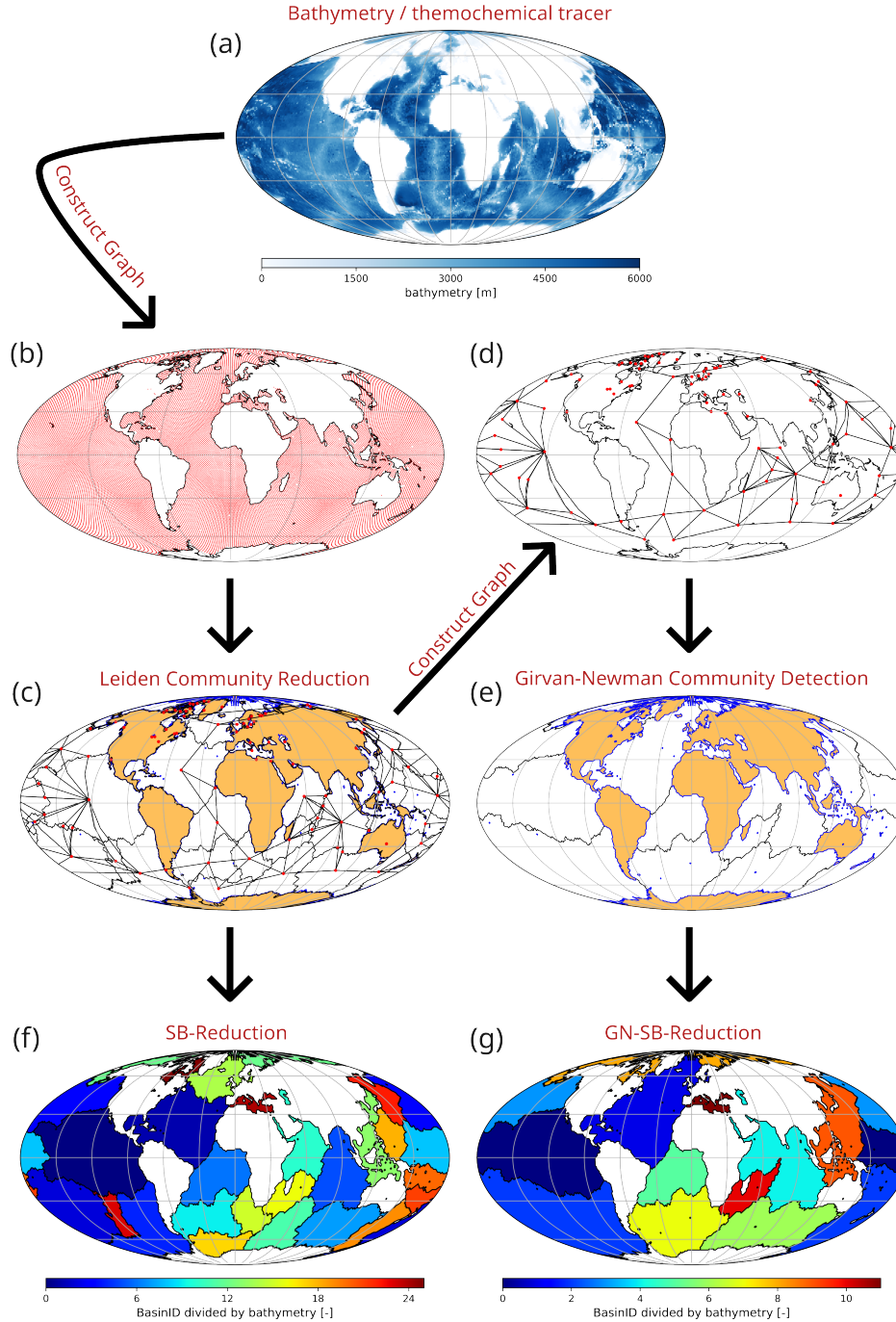


Figure 4.7: Composite basin detection using measured bathymetry. (a) ETOPO1 bathymetry (9) sampled at 1 degree resolution and used to construct equidistantly spaced spatial network (b). (c) graph network constructed using communities from an ensemble of Leiden community detections (d). (d-e) detected communities for Leiden and Louvain-Girvan-Newman composite algorithms. (f) detected communities after small basins post-process step. (g) Detected communities after Girvan-Newman and small basins post-process step. Detected communities capture basins as regions with high water column connectivity. The basin boundaries often occur near mid-ocean ridge axes, ocean plateaus and seamounts, continental rises, and at ocean gateways.

basins detected in Figure 4.7f, g capture ocean basins we are familiar with at present-day (e.g., the North, equatorial, and South Pacific, the North and South Atlantic, and the Indian Ocean). However, given the ensembles used in the clustering, are the basins (i.e., communities) drawn consistently across the ensemble that define them or would they change with increased ensemble size? We directly address this by evaluating the change in global area averaged silhouette values for ensembles of similar sizes when using Leiden SB-Reduction ($\gamma = 0.01$) with DQT-CDF weighting (shortened = 5σ , shifted = 0.5σ) on measured seafloor bathymetry at 1 degree spatial resolution (Fig. 4.8). We found that the global area averaged silhouette values do not appreciably decrease with ensemble sizes greater than 30. Additionally, the minimum global silhouette value for all ensemble sizes indicates strong stability within the ensemble. Finally, the global silhouette almost always indicates at least good community stability within uncertainty. These results suggest that ensembles sizes greater than 30 will produce numerically robust clustering.

4.4.2. Controls on Basin Definition - Role of Seafloor Bathymetry roughness |

In order to provide further confidence in the detected ocean basins, we need to understand what dominant field features control boundary descriptions. Based on a visual analysis of the basins detected in the previous section (Fig. 4.7), we find that seafloor volcanic features like submarine ridges and oceanic plateaus play a critical role in determining basin boundaries. We test this observation by using a bathymetry field devoid of those features.

We compare the basin detection of present-day bathymetry for measured and process-limited reconstructions (11). Our seafloor reconstruction (Fig. 4.9) does not capture rugosity resulting from area-averaged small-scale seafloor features (i.e., near-MOR-axis forming seamounts, seamount chains, submarine volcanic seamount arcs, and normal faulting) or large igneous province (e.g., the Ontong Java plateau). The basin boundaries detected with the reconstructed bathymetry tend to favor delineation along the mid-ocean ridge axes and their transform faults (with large depth differences), whereas the boundaries from measured

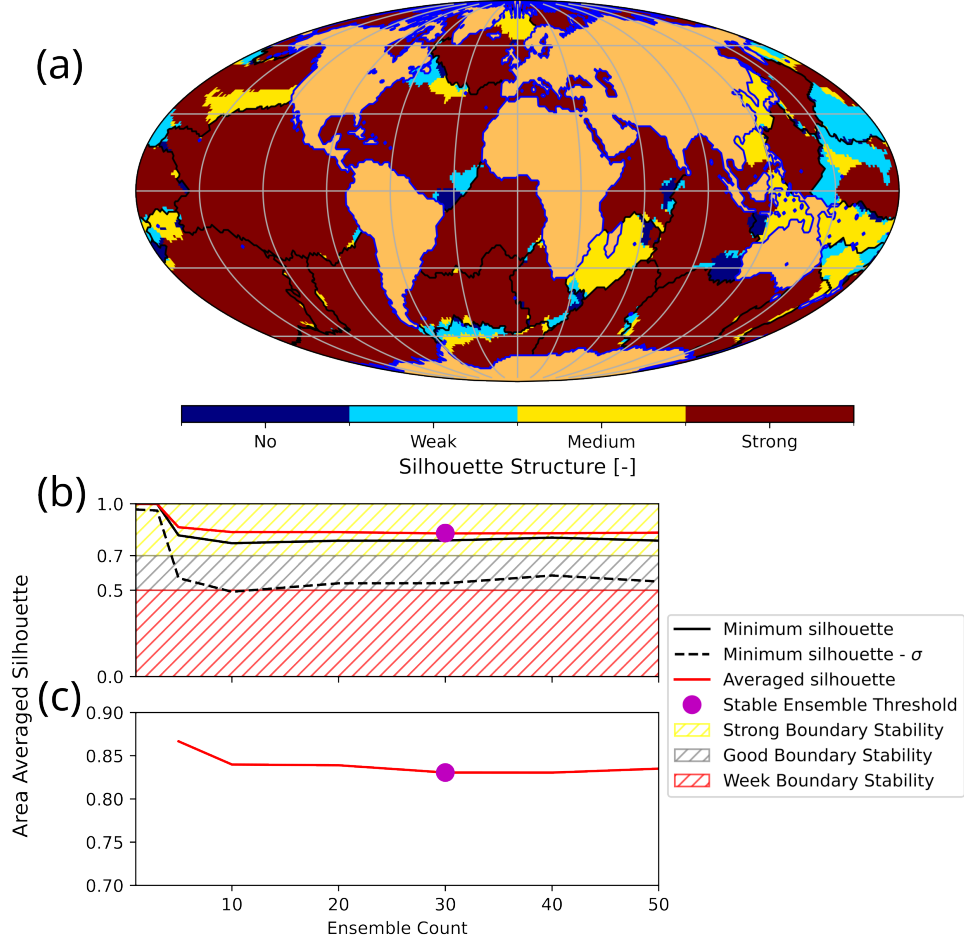


Figure 4.8: Figure shows one method to analyze community detection stability using ensemble averaged silhouette values. Basins were detected using Leiden SB-Reduction ($\gamma = 0.01$, ensemble size = varied) and DQT-CDF weighting (shortened = 5σ , shifted = 0.5σ) on 1 degree spatially resolved present-day measured bathymetry. (a) Sample of global node-wise calculated silhouette scores for an ensemble size of 50 and with No, Weak, Medium, and Strong structure divided at silhouette scores of 0.25, 0.5, and 0.7. (b) ensemble averaged silhouette values (red line) with minimum ensemble silhouette value and σ lower boundary (black lines) using 20 ensembles for ensemble sizes ranging from 1 to 50. Strong, good, and weak clustering regions are highlighted. (c) a close-up of the ensemble averaged silhouette values and when they stabilize at ensemble sizes of 30 (magenta)

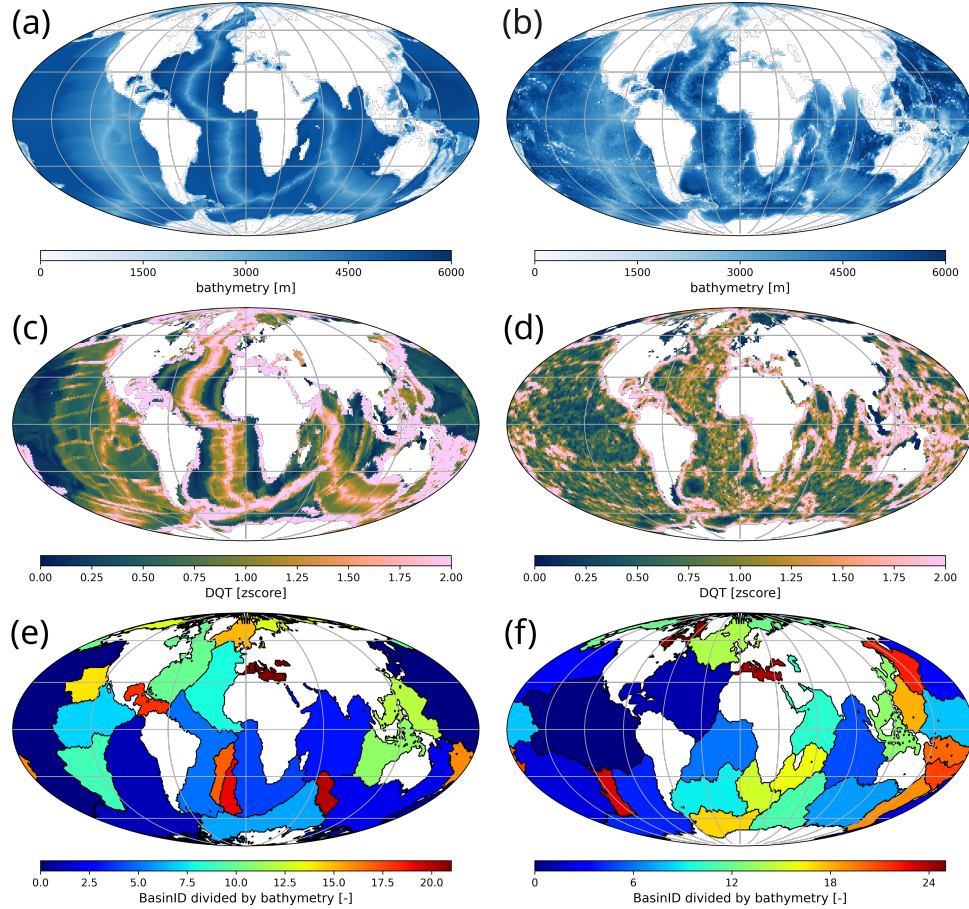


Figure 4.9: Basin detection on (a) reconstructed present-day bathymetry (11), and (b) measured bathymetry sampled at 1 degree (9). (c) and (d) are transformed fields of (a) and (b) used to calculate node connection weights. (e) and (f) are basins detected using SB-Reduction ($\gamma = 0.01$, ensemble size = 50) and DQT-CDF weighting (shortened = 5σ , shifted = 0.5σ) on 1 degree spatially resolved (c) reconstructed bathymetry and (d) measured bathymetry present-day, respectively. Models have modularity (.88) and (.91) and silhouette (.80) and (.85) values for (e) and (f), respectively. Seafloor volcanism represented in sampled ETOPO1 helps to define the edge of basins. Basins from reconstructed bathymetry are more likely to share boundaries along mid-ocean-ridges.

bathymetry do not (Fig. 4.9e-f).

The different quartile transformation values of reconstructed and measured seafloor (Fig. 4.9c-d) indicate that reconstructions are smooth in the normal direction to ridges and abyssal plains but rough at MORs, across MOR transform faults, at continental rises (e.g., Patagonian shelf), and near tectonically active regions (e.g., South China Sea and Banda Sea). In contrast, the actual present-day seafloor has much less pronounced MORs and MOR transforms in the difference quartile space (Fig. 4.9d). However, slow spreading ridges (e.g., mid-Atlantic ridge) are more pronounced compared to fast spreading ridges (e.g., Nazca ridge). This is a direct result of faster versus slower spreading ridges undergoing less thermal subsidence for equal distance from the MOR axis (117). Atlantic ocean rugosity is also more pronounced than near the Nazca ridge, likely resulting from the morphology of normal faults and averaging of seamounts and their regional differences in formation. Fast spreading ridges tend to support uniform topography, symmetric high-angle normal faults while slow ones show irregular topography with asymmetric low-angle normal faults (169). On slower spreading ridges, seamounts form from pillow lavas, less often, and are irregularly distributed along the MOR axis (170; 171, and references therein). Seamounts on faster spreading ridges form from sheet and lobate lava flows, more often, and are more regularly distributed along the MOR axis. Overall, the rugosity of the actual seafloor topography reduces the power of mid-ocean ridges and their transform faults to dictate ocean basin boundaries. Instead, the 2nd order volcanic features play a larger role. The contrast in detected basin boundaries of reconstructed versus measured bathymetry highlights the need to reconstruct the location of past seafloor volcanic features for detecting paleo deep ocean basins, as well as the uncertainty in basin detection when the location or existence of these features is unknown.

4.4.3. Parameter choices in the clustering algorithm | We next evaluate how different parameters in our composite algorithms affect the final quality of the results. We systematically vary resolution (γ) and ensemble sizes for the Leiden SB-Reduction algorithm using DQT-

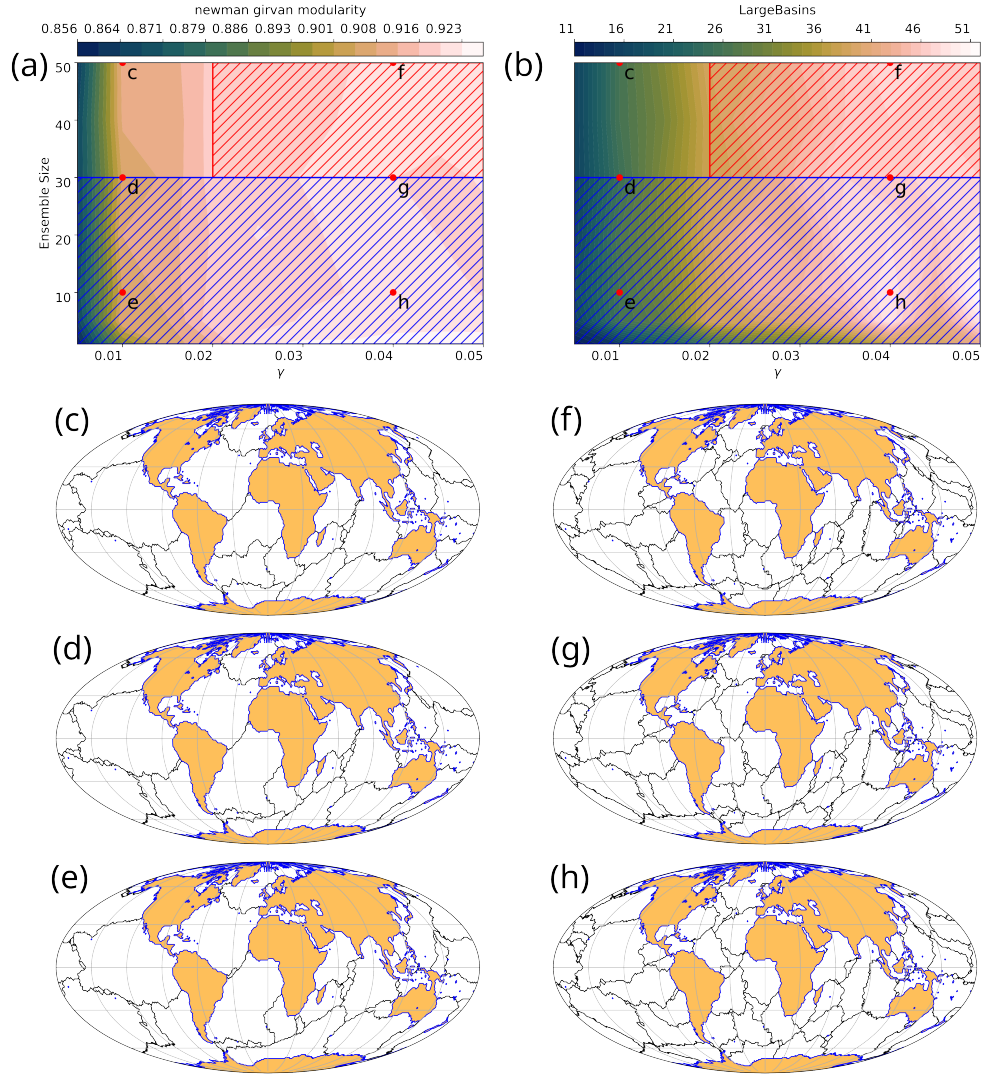


Figure 4.10: Basin partitioning evaluation metrics for Leiden ($\gamma = \text{varied}$), not SB-Reduction, detection on ETOPO1 (9) sampled at 1 degree resolution and using DQT-CDF weighting method (shortened 5σ , shifted 0.5σ). (a) Newman Girvan modularity that describes better community divisions with values closer to 1. (b) Number of basins that represent more than 0.5% of the total seafloor area. Basins divisions (c-e) and (f-h) correspond to ensemble and γ value in (a-b). Blue hatched regions define ensemble sizes where increases in ensemble size lead to visually different community structure (or potentially global average silhouette). Red hatched regions define community models with larger γ and many smaller communities as shown in (f-h).

CDF weighting (shortened 5σ , shifted 0.5σ) - See (Fig. 4.7a, f). Figure 4.10 shows the modularity and large basin count of the entire graph for varying γ and ensemble sizes, along with examples of the detected communities for select end-members. In general, a higher modularity indicates that nodes fit well within their communities. Large basins indicate the number of communities preserved after SB-Reduction. We find that ensembles larger than 1 will almost always reduce the metrics that define the quality of community structure. While these metrics cannot tell us the exact best choice for ensemble sizes, ensembles are beneficial for reducing the variance in community boundaries that results from the initialization of the greedy Louvain and Leiden detection methods. *We consider community structure to be of good quality when an ensemble of size X and $X + 5$ show little visual difference between the community structures.* For a range of settings, we find that this is almost always achieved for ensemble sizes above ~ 30 in our measured topography example (Fig. 4.10) and is consistent with our analysis of community stability (Fig. 4.8). However, other systems may require exploration to determine the appropriate ensemble size. All models with $0.01 \leq \gamma \leq 0.02$ have high modularity (≥ 0.9) and between ~ 15 -30 large basins. For our typical composite model, we choose the resolution ($\gamma = 0.01$) that leads to quality communities (Fig. 4.10) that capture large-scale basin structure (i.e., ≤ 26 Large basins).

4.4.4. Reduction Method - Leiden vs Louvain | Both of our composite community detection algorithms require a node reduction step that uses either the Louvain or Leiden algorithm. The Leiden algorithm is theoretically considered an improvement over the Louvain algorithm, but this may not hold for the network graphs of interest to us. We thus evaluated the performance of both algorithms for community detection on modern surface (≤ 200 m), intermediate (> 200 m and ≤ 1000 m), and deep (> 1000 m) ocean potential temperature fields averaged over 1994 (GLORYS12V1; (18)). Both node-reduction community-detection algorithms produce similar ocean basins within each water-column interval (Fig. 4.11b-c, e-f, and h-i). Both algorithms also produce the same number of basins per ocean layer.

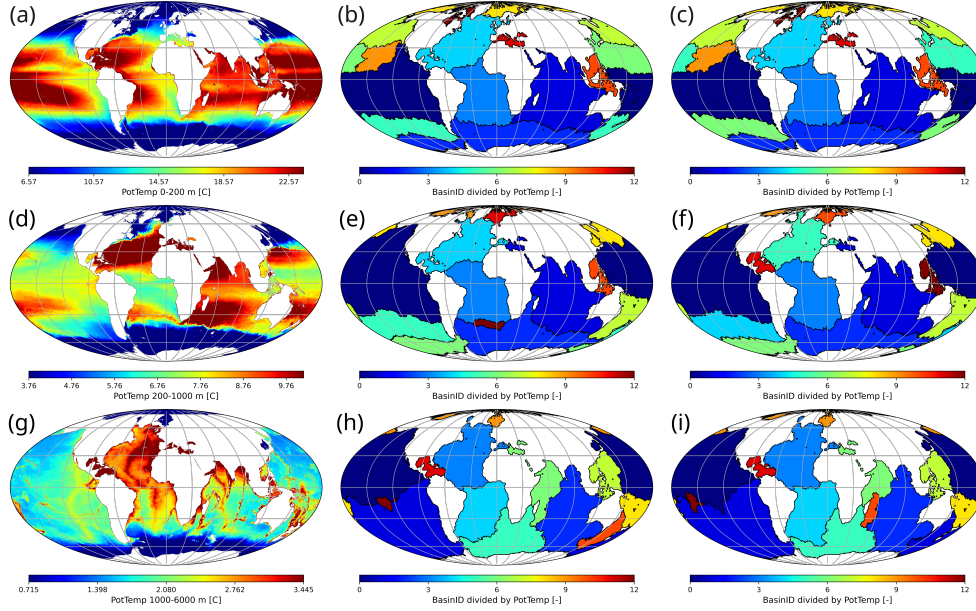


Figure 4.11: Node reduction using Louvain and Leiden community detection algorithms. Input surface (a) (≤ 200 m), (d) intermediate (> 200 m and ≤ 1000 m), and (g) deep (> 1000 m) ocean potential temperatures averaged over 1994 (GLORYS12V1; (18)). (b,e,h) are Leiden and (c,f,i) are Louvain community detections for (a,d,g) inputs. Models have modularity (.85, .81, .84) and (.85, .81, .83) and silhouette (.96, .94, .92) and (.98, .97, .87) values for (b,e,h) and (c,f,i), respectively. Communities were detected using SB-Reduction ($\gamma = 0.01$, ensemble size = 50) and DQT-CDF weighting (shortened = 3σ , shifted = σ). Both Louvain and Leiden produce similar community boundaries for given ocean depths.

However, there are some minor differences in the location of the basins and their boundaries (Fig. 4.11b-c). The primary benefit of the Leiden algorithm in guaranteeing well-connected communities is not necessarily maintained during ensemble runs with more than 1 iteration and the subsequent agglomerative clustering step. Still, the location of boundaries will be more consistent for each run that comprises the ensemble. Neither choice of reduction step leads to significant computational improvements for these examples. Therefore, we establish a preference for using the Leiden algorithm for both composite community-detection algorithms in subsequent analysis.

4.4.5. Multifield Community Detection | In complex systems like the ocean, defining a "community" or "basin" using just one variable, such as bathymetry or temperature alone, can be overly simplistic and limiting. A truly effective basin is often a region defined by a combination of properties, including its physical dynamics, water temperature and salinity, and geochemical signatures. Therefore, a more powerful and realistic approach is to use multifield community detection. This requires defining the graph's connections based on the similarity across multiple, simultaneous data fields, allowing for the discovery of more holistic and functionally relevant basins that any single-field analysis would overlook. Thus, ExoCycle was designed to handle multifield community detection analysis.

We show an example of such an analysis using present-day surface ocean fields (Fig. 4.12). First, we initialize two or more fields, for example, surface ocean (≤ 200 m depth) salinity and potential temperature. Then node edge weights are calculated per individual field (Fig. 4.12c-d) and combined (Fig. 4.6) according to the user's preference and application. Here we show detected communities when weights are assigned to the minimum (Fig. 4.12e) and mean (Fig. 4.12d) of overlapping field weights. Minimum and mean weight combinations favor either more or less structure and, therefore, either more or less community division, respectively.

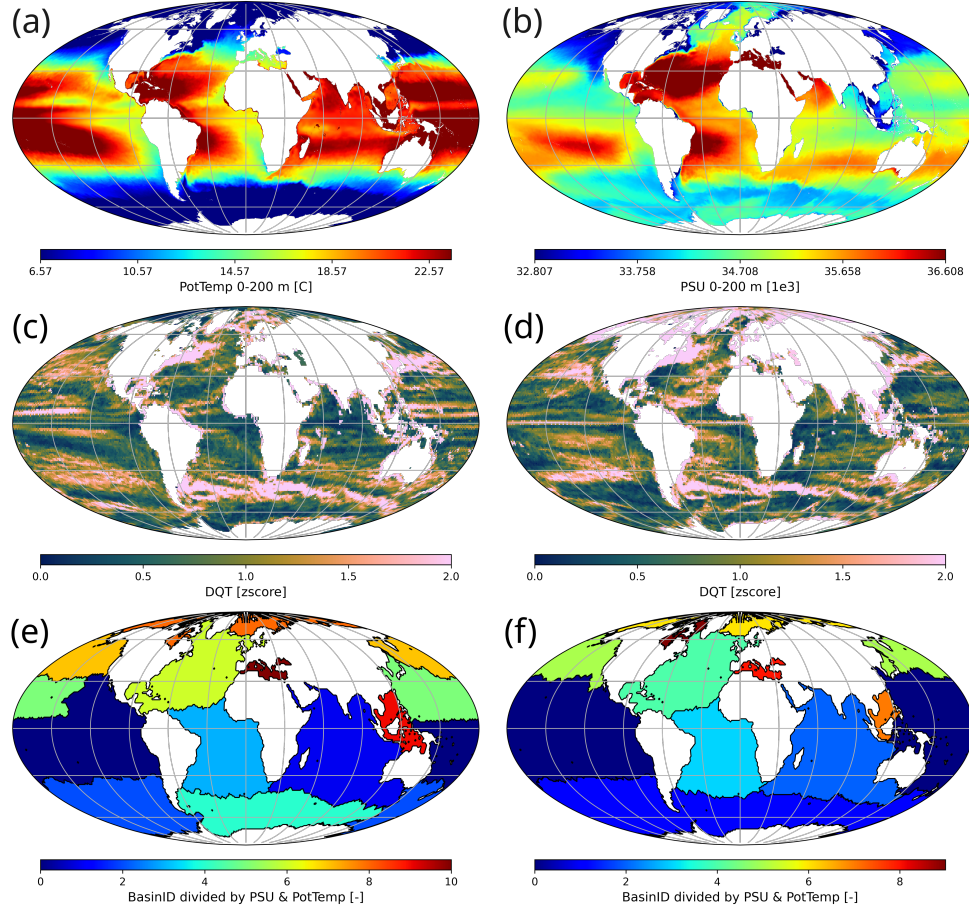


Figure 4.12: Multifield community detection. (a-b) Input surface (≤ 200 m depth) salinity and potential temperature averaged over 1994 (GLORYS12V1; (18)). (c-d) Difference quantile transformed fields used to determine spatial network weight connections. Communities detected using (e) minimum (strongest connections) and (f) mean of weights from DQT-CDF method (shortened = 3σ , shifted = σ) with a Leiden SB-Reduction ($\gamma = 0.01$, ensembleSize = 50). Models have modularity (.85) and (.79) and silhouette (.95) and (.97) values for (e) and (f), respectively.

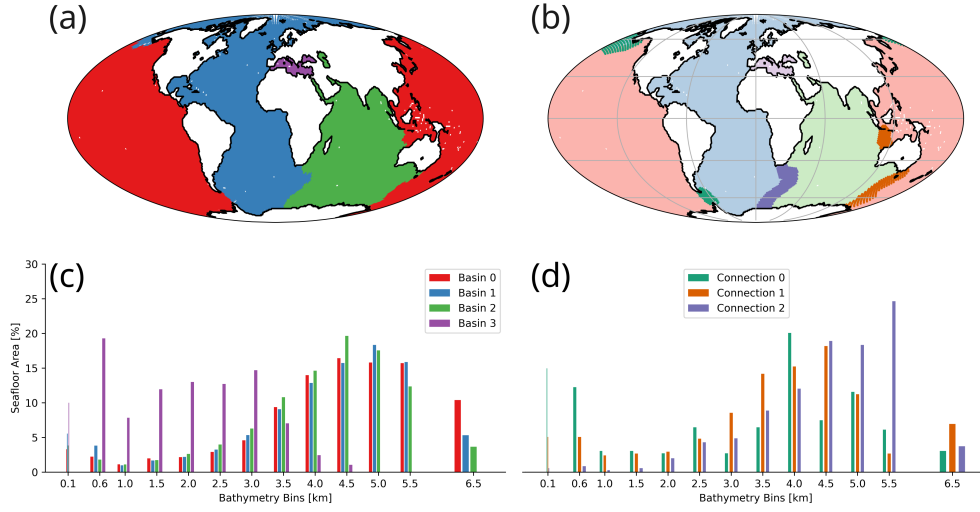


Figure 4.13: Basin-community detection and bathymetry distributions using ETOPO1 (9) sampled at 1 degree resolution and with Leiden GN-SB-Reduction ($\gamma = 0.01$, ensemble size = 50, basin count = 4) and DQT-CDF weighting (shortened 5σ , shifted 0.5σ). (a) Atlantic, Indian, Mediterranean Sea, and Pacific basins and (c) their bathymetry distributions. (b) Connecting regions defined as seafloor within 444 km of boundaries, and (d) their distributions. The model has a modularity of .91 and silhouette value of .85.

4.4.6. Field Statistics: Measured Bathymetry | Once communities are detected, our analysis in ExoCycle shifts from identification to characterization. Understanding how the global ocean functions requires moving beyond just drawing ocean boundaries on a map. ExoCycle is packaged with methods for calculating the field characteristics within and in the vicinity of basin-community boundaries to quantify the physical and geochemical properties of these basins (Fig. 4.3). Typically, the inter-basin boundaries are not uniform; some are deep, broad gateways facilitating large-scale exchange, while others are shallow, narrow sills that act as critical chokepoints, strongly regulating ocean circulation and the transport of heat and nutrients.

We show these methods with Leiden GN-SB-Reduction ($\gamma=0.01$, ensemble size = 50, basin count = 4) detected ocean basin-communities of present-day measured seafloor bathymetry (Fig. 4.13). The Pacific, Atlantic, and Indian Oceans and the Mediterranean Sea are identified as single basins (Fig. 4.13a). The Indian and Atlantic Oceans are shallower since they have

experienced less lithospheric thermal subsidence (Fig. 4.13c). The Mediterranean Sea is shallow and uniform in depth, reflecting the uplift and increased sedimentation initiated by the closure of the meso-Tethys ocean (Fig. 4.13c; (172)). The strength of connection between these basins can be evaluated using boundary seafloor distributions (Fig. 4.13b,d). The ~ 14 km thick Strait of Gibraltar that connects the Atlantic Ocean to the Mediterranean Sea is not captured with our 1 degree spatial resolution. The connections that facilitate larger water communication between the South Pacific, Atlantic, and Indian Oceans are dominated by seafloor greater than 3.0 km depth. However, North Pacific and Indian Ocean boundaries from Vietnam to East Indonesia are dominated by shallower water transport (≤ 3.0 km depth). These observations may seem obvious within the context of major present-day ocean basins. However, our framework allows for the calculation of basin statistics from an objective basin detection algorithm that can operate across multiple scales, times, and data fields. The calculated basin statistics can also be directly used in paleoclimate ocean box models (e.g., LOSCAR and GEOCLIM variants) to calculate the long-term marine carbon cycle as well as ocean overturning circulation.

4.5. Scientific Results & Discussion | In the previous results section, we illustrated and validated ExoCycle’s capability to automatically and objectively define ocean basins by highlighting case examples of present-day, paleo-bathymetry, and using ocean data fields that capture laterally variations in bathymetry and temperature. We next leverage the generalized framework of ExoCycle community detection to address the following scientific questions described in the introduction.

- *Question:* What spatial resolution and accompanying features (e.g., oceanic plateaus) are needed to capture the primary structure of ocean basins? *Result:* Bathymetry spatially resolved to at least 2-degree resolution is able to capture global basin structure which also strongly depends on 2^{nd} order seafloor volcanic features.
- *Question:* How to capture the layered structure of the water column at the ocean

basin scale and its effects on chemical transport (i.e., continent to ocean) through measurement aggregation? *Result:* Shallow ocean bathymetry (< 600 m) community detections capture differences in major present-day and paleo- river outputs (i.e., sediment transport to oceans). ExoCycle has a pipeline for depth dependent basin definitions - separately for surface, intermediate, and deep ocean basins - that can be used with traditional intermediate-complexity Earth system models. We find that deep (> 1000 m) and shallow ocean chemistry (≤ 200 m) basin detections are not equal, indicating a layered structure at scales smaller than, e.g., the Atlantic, Pacific, and Indian Oceans.

- *Question:* How different are our estimated ocean basin boundaries for paleo-bathymetry reconstructions for the last ~ 100 Myr? *Result:* We find that ocean basin boundaries can strongly change over time even over the last 100 Myr, with changes in continental structure.

In the following sections, we describe each of these results in detail.

4.5.1. Resolving Field Feature Wavelength: Seafloor Bathymetry Resolution |

A feature of spatial community detection is that it divides a space into individual cluster communities. However, this naturally depends on the resolution of the underlying data used in the algorithm. Thus, a key question is: what is the right spatial resolution that can capture the appropriate seafloor features? Addressing this question is critical for the choice of spatial resolution that is needed to capture the community (e.g., basin) structure while balancing the computational trade-offs (see subsection **Computational Efficiency**).

We address this question using the present-day measured seafloor bathymetry (sampled ETOPO1; (9)) to evaluate the quality of basin-community detection for spatial resolutions ranging from 8 to 1 degree (Fig. 4.14). Generally, similar boundaries are found near gateways at all spatial resolutions. Bathymetric features (e.g., mid-ocean ridge systems, oceanic plateaus, seamount chains) contribute to additional basin division when spatial resolution is

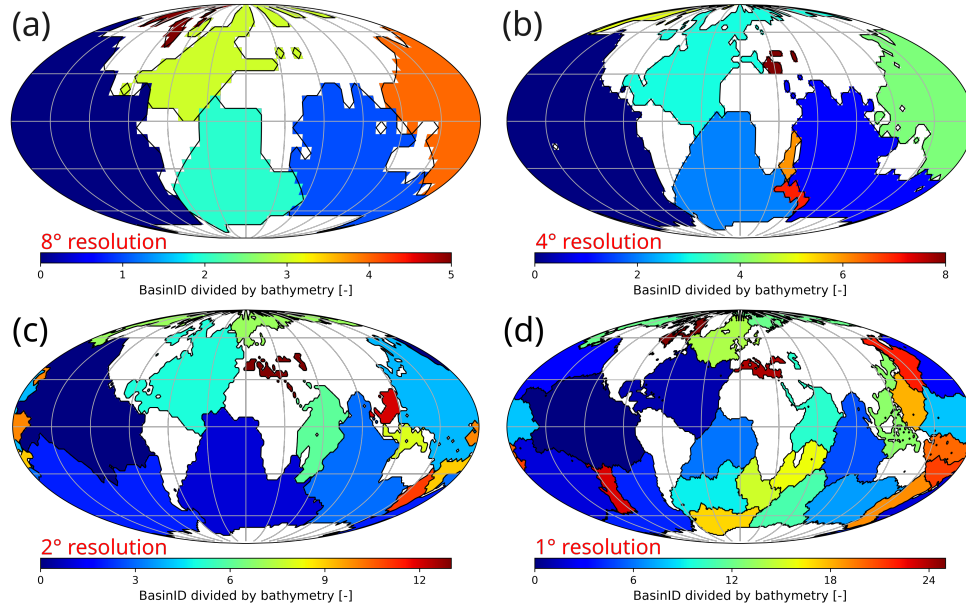


Figure 4.14: Basin detection when using the Leiden ($\gamma=0.01$) and small basin merger community detection on ETOPO1 (9) with the DQT-CDF weighting (shortened 5σ , shifted 0.5σ) for sampled bathymetry and node grid spatial resolutions of (a) 8, (b) 4, (c) 2, and (d) 1 degree. Models (a-d) have a modularity of .70, .76, .84, .91, respectively. Higher spatial resolution models lead to more community detections with γ held constant. Spatial degree resolution 2 and 1 models capture enough bathymetry features that are important for representing basin-community structure.

increased to 2-1 degrees (Fig. 4.14c-d). In some cases, higher-frequency features at 2-1 degree resolution dictate the boundary location between larger-scale basins (e.g., the Atlantic and Indian Oceans). This is evident even with a constant Leiden resolution (γ) that leads to an increased number of large basins (i.e., basin area percent $> \max(\chi)$) with increasing spatial resolution. Ultimately, there is little difference between 1- and 2-degree spatial-resolution community structure for modern measured bathymetry on a global scale. Therefore, we expect that 1-degree-resolution graphs capture the global basin-community structure of seafloor reconstructions. This finding has significant implications for the GCMs mentioned earlier. If a model operating at a much coarser resolution, such as 10 degrees, cannot even correctly resolve the fundamental boundaries, gateways, and community structure of the major ocean basins, the model’s ability to simulate realistic ocean dynamics—like large-scale circulation and inter-basin exchange—is inherently compromised from the outset.

4.5.2. Shelf and Pelagic Ocean Basins: Seafloor Bathymetry | The seafloor (and planetary hypsometry) distribution is bimodal on Earth (Fig. 4.13C), clearly delineating two primary domains: the shallow continental shelves and the deep abyssal plains. These domains are the product of fundamentally different processes. The distribution of shallow ocean (primarily passive margin shelf) regions is a consequence of continental rifting, stretching and faulting that form the transition from continental to ocean lithosphere and stages in sedimentation (pre-, syn, post-rift) and erosion (e.g., by turbidity currents) (173). Continental sedimentation and changes in sea-level are the major drivers of the shallow ocean in post-rift regions. Conversely, the deep ocean’s topography is primarily a function of: (1) syn-rifting dynamics; and (2) mid-ocean ridge spreading rates (which control lithospheric thermal structure and subsidence; (61)). Both regions can also be affected by additional mantle process confounders, e.g., 2nd order volcanism that forms oceanic plateaus and seamounts, changes in topography from mantle flow supported topography (dynamic topography; e.g., (174)), relative sea-level from changes in the geoid, and eustatic sea-level from changes in,

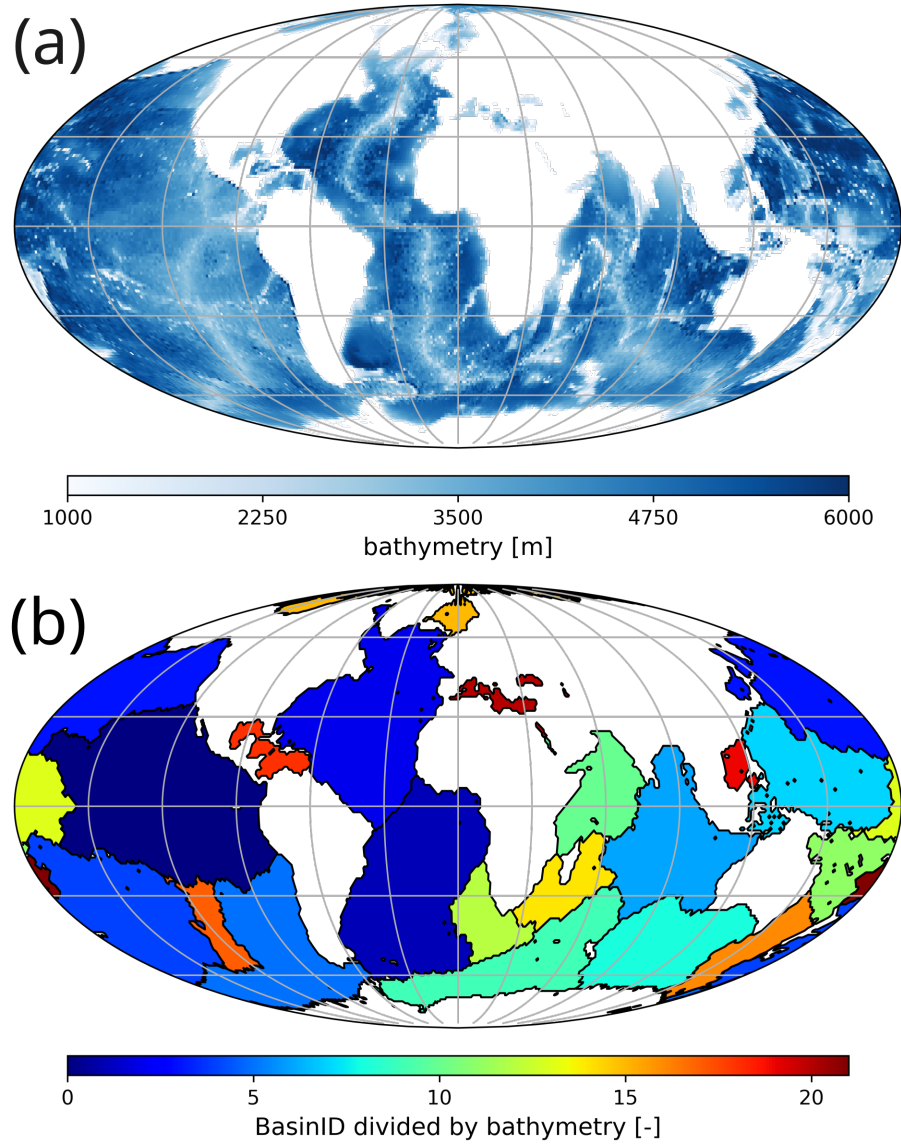


Figure 4.15: Basin detection on ETOPO1 (9) deep bathymetry sampled at 1 degree resolution. (a) Deep bathymetry (>1000 m), and (b) basin-communities using Leiden SB-Reduction ($\gamma=0.01$, ensemble size = 50) with DQT-CDF weighting (shortened 5σ , shifted 0.5σ). The model has a modularity of .89 and silhouette value of .84. Basin boundaries largely run along shallow regions (e.g., MORs and seafloor volcanism), similarly to basin boundaries calculated from the full Etopo bathymetry field.

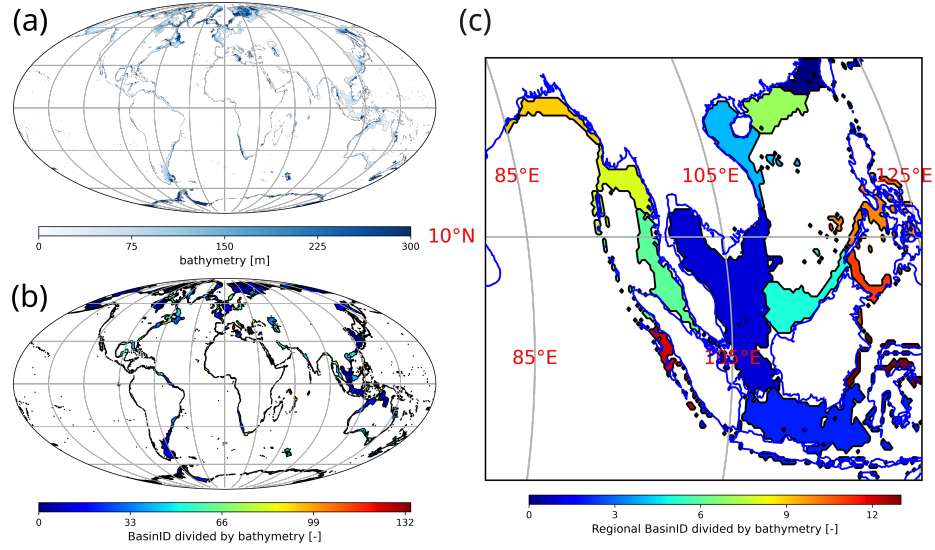


Figure 4.16: Basin detection on ETOPO1 (9) shallow bathymetry (<600 m) sampled at 0.3 degree resolution. (a) Shelf bathymetry (<600 m), and (b) global shelf partitions using Leiden SB-Reduction ($\gamma = 0.15$, ensemble size = 50) with DQT-CDF weighting (shortened 5σ , shifted 0.5σ) at 0.3 degree spatial resolution. γ is chosen to maximize global modularity to 0.98 and results in a silhouette value of 0.94. The small basin merger step is modified such that it merges communities of relative similar scale when compared with full ocean detections. (c) South East Asia shelf partitions identifying isolated regions corresponding to present-day and past river sediment supply from now submerged river systems that span throughout the Sunda Shelf (19; 20; 21).

e.g., global mid-ocean-spreading rates, (de)glaciation, aquifer-storage, and the distribution of continental versus ocean dynamic topography and geoid contributions (72; 73). While processes in both domains span long timescales, their characteristic stability differs critically. Shallow bathymetry is highly dynamic, responding to sedimentation and sea-level changes on short timescales (100s to 10s of kyr). In contrast, the deep ocean’s structure is stable over much longer geological periods (100s of kyr to 10s of Myr). This fundamental difference in stability and governing processes motivates analyzing them separately. It is therefore of interest to know how these two domains are individually subdivided into basins.

We start by identifying ocean basins in the deep marine (≥ 1000 m) beyond continental shelves, as they will represent the most temporally stable ocean basins. Figure 4.15 shows basin-communities detected for deep marine measured topography by using Leiden SB-Reduction ($\gamma=0.01$, ensemble size = 50) with DQT-CDF weighting (shortened 5σ , shifted 0.5σ). The resulting basin-communities (Fig. 4.15) are similar to those found using the full seafloor sampled ETOPO1 field (Fig. 4.14d), with boundaries following mid-ocean ridge axes, extensive seafloor volcanism, and major gateways. This deep-only analysis tends to identify fewer, larger, and more robust basins because the input field is inherently smoother and more normally distributed, having removed the high-frequency, bimodal influence of the continental shelves.

What about the continental shelves and shallow bathymetry? Are the dynamic physical processes that shape these regions detectable with our community detection framework? The continental shelf and rise, islands, and tectonically uplifted regions represent the reciprocal shallow bathymetry. These regions are critical components of the Earth system, as they host a plethora of biological activity in shallow areas. Furthermore, the shallow shelf is the primary area of carbonate production and burial and thus marks the exchange between the land and marine systems (the boundary layers). Additionally, the shelf-community structure for the present and past is necessary to determine where and when ocean chemical measurements are spatially independent or comparable and aggregatable for use with intermediate volatile

cycling models. We illustrate the power of ExoCycle for this application by automatically defining the shelf (<600 m) community structure on a global scale and at a finer ~ 0.3 degree spatial-resolution (Fig. 4.16a-b). Shelf-community detection was carried out with Leiden SB-Reduction (ensemble size = 50) with DQT-CDF weighting (shortened 5σ , shifted 0.5σ) on shelf bathymetry (Fig. 4.16). However, we tuned the resolution parameter ($\gamma=0.15$) for this finer-scale application. The SB-Reduction χ -values are modified such that smaller merged communities consist of similar node counts (~ 80 , ~ 400) as we had for global ocean community detection. This process maintains a consistent organizational scale despite the different input data.

Overall, this shelf-focused analysis yields powerful and physically realistic results. *Shelf partitions, particularly in South East Asia (Fig. 4.16c), identify isolated regions that correspond with both A) present sediment supply from the i) Brahmaputra, ii) Irrawaddy and Salween, iii) Chao Phraya and Mekong, iv) Rouge, and v) Pearl rivers and B) past sediment supply since 200 kya from now submerged river systems that span the present-day Sunda Shelf (19).* Remarkably, our shelf partitions, which rely solely on coarse (0.3° resolution) seafloor topography, reproduce larger-scale basin definitions similar to those used in basin sediment stratigraphy analysis (20; 21).

Although global-scale shelf community detection at a spatial resolution greater than 0.3 degree (≥ 30000 nodes) becomes computationally impractical on standard desktop lab computers, additional ExoCycle methods allow *regional-scale analysis* at higher resolution if needed. Applying this method at the basin scale will enable the systematic characterization of shelf inputs (from continents) and outputs (to pelagic basins). This directly influences the interpretation of data from IODP, ODP, and future sediment records (e.g., from D/V Chikyū or Meng Xiang) and large-scale compilations (e.g., PAGES2K (175) and PetDB Databases).

Ultimately, the growth and evolution of continental shelves are a function of inputs (sediment supply, organic matter) and outputs (turbidity currents to the deep marine). ExoCycle’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. Such an analysis can inform reconstructions of major paleogeographic events, such as the erosion of Greater India into the Indian and Tethys oceans, or Eurasian and Australian erosion into multiple oceans during the Early Cretaceous and leading into the Paleocene-Eocene Thermal Maximum (PETM).

4.5.3. Paleo-Ocean Basins: Reconstructed Bathymetry Change | Averaged seafloor sediment records are critical constraints for intermediate-complexity Earth system models, which are often used to tune poorly resolved or unrecorded volatile cycling parameters (12). These models, by necessity, address data paucity by reducing the 'fine-scale' (~ 1 -5 degree) Earth system to a coarse, basin-scale resolution (e.g., a "box model"). This reduction creates a critical dependency: the location of the defined paleo-ocean boundaries dictates which chemical and bathymetric records are averaged together. This subjective choice directly influences the model's simulation of volatile (carbon) transport within and between basins (11; 12). Therefore, an objective and physically meaningful method is required to "draw" these boundaries, ensuring we are correctly delineating regions that genuinely interacted as a cohesive system during a paleoclimate event.

ExoCycle provides a systematic framework for solving this exact problem: it objectively defines basin boundaries, determining where seafloor sediment records can be aggregated and to which ocean box they can be appropriately applied. We demonstrate this process using reconstructed bathymetry for the present-day, 60 Ma, and 80 Ma (Fig. 4.17; (11)). The resulting paleo basin-communities are geologically meaningful, with boundaries clearly delineated along mid-ocean ridges, transform faults, and major gateways (Fig. 4.17d-f). The method is also sensitive enough to capture restricted epicontinental seas as separate hydrographically isolated basins, such as the Western Interior Seaway and Western Eurasia during the Late

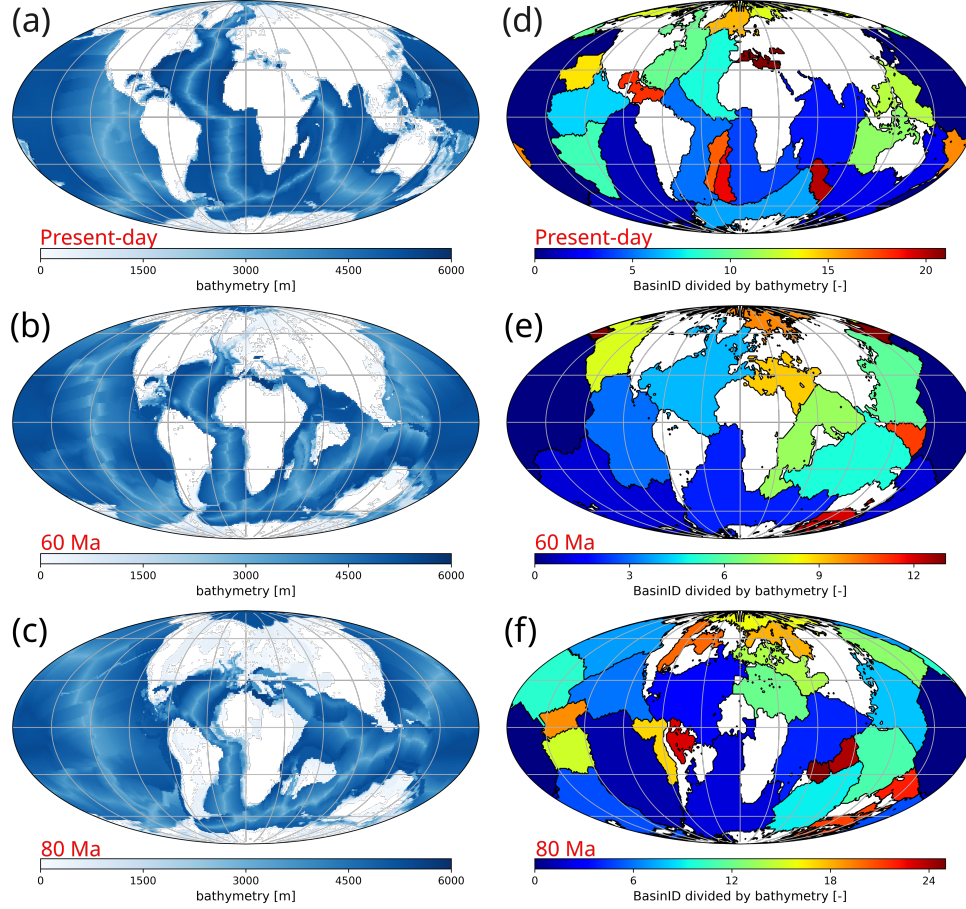


Figure 4.17: Basin detection on reconstructed bathymetry. (a-c) Reconstructed present-day, 60 Ma, and 80 Ma bathymetry (Bogumil et al. 2024), respectively. (d-f) Basin detection for (a-c) fields when using the ensemble (50) Leiden ($\gamma=0.01$) and small basin merger community detection with the DQT-CDF method (shortened 5σ , shifted 0.5σ) at 1 degree spatial resolution. Models have modularity (.88, .87, .91) and silhouette (.80, .85, .82) values for (d-f), respectively. Changes to continental distribution (e.g., opening Atlantic), shelf space (e.g., epicontinental seas), mid-ocean ridge distribution directly influence the number of ocean basins that existed during CO_2 forcing events since the Late Cretaceous.

Cretaceous (Fig. 4.17f). Depending on the computational needs, ExoCycle allows for simplification of these basins into major basin systems using the GN-SB-Reduction method and basin statistics generation. Basin descriptions can then be used with intermediate-complexity Earth system models to evaluate paleoclimate events (e.g., Aptian-Albian Boundary Interval, Cretaceous-Paleogene transition, Maastrichtian Cool Period, PETM, Middle Eocene Climatic Optimum).

A key conclusion from this test is that the ocean basin boundaries change significantly over time, especially the evolving area/depth distribution of the basins even if there is a consistent "Pacific" Ocean basin throughout the last 100 Myr. The critical implication of our result for models like LOSCAR is that the fundamental 'box' structure of the ocean is not static, as is often assumed (e.g., (84; 176)), but evolves significantly through time. Our objective, time-dependent framework allows modelers to dynamically update their model's configuration to reflect the actual paleogeography of a given time-slice, rather than imposing a fixed approximation over 10s of Myr. Our basin detection method also improves on a simple plate model basin classification that is limited by the lack of representation of ocean and seafloor in tectonically complex areas and key basin delineating features (e.g., oceanic plateaus, seamounts, position of coastlines, and sedimentation). Ultimately, this provides a necessary pathway to move beyond static, subjectively-defined box models, enabling more physically realistic simulations of global biogeochemical cycles during periods of major tectonic reorganization.

4.5.4. Surface Ocean Basins: Ocean Temperature - surface vs deep | Carbon cycle modeling and data aggregation often assume that ocean basin boundaries are vertical walls, independent of depth (Fig. 4.2). However, the physical drivers of ocean structure are fundamentally different at different depths, so there should be no expectation that basin boundaries are consistent throughout the water column. We analyze the ocean in two primary depth domains: the surface ocean (less than 200 m) and the deep ocean (greater than 1000

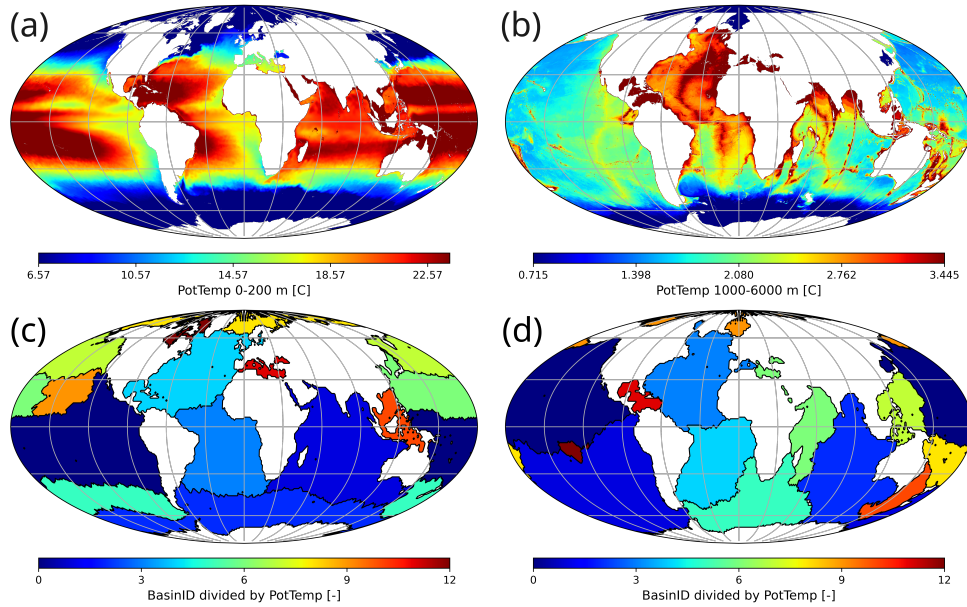


Figure 4.18: Ocean basins detected throughout the water column. Ocean potential temperature averaged over 1994 (GLORYS12V1; (18)) and averaged over the a) shallow (<200 m) and B) deep (>1000 m) water column. Basin detection (c-d) for (a-b) respectively. Basin detection used Leiden SB-Reduction ($\gamma = 0.01$, ensemble size = 50) with DQT-CDF weighting (shortened 3σ , shifted σ). Models have modularity (.85) and (.84) and silhouette (.96) and (.92) values for (c) and (d), respectively. Shallow basins reflect latitudinal changes in temperature near the equator. Deep ocean basins are primarily defined by continental division and larger-scale changes in temperature driven by ocean circulation.

m). These thresholds capture the two principal transitions in the physics governing lateral ocean structure. Above 200 m, surface property gradients are dominated by wind-driven circulation and solar heating (the Ekman layer and photic zone). Below 1000 m, lateral gradients are set by thermohaline circulation and constrained by seafloor topography (e.g., (177)). The intermediate ocean (200-1000 m) spans the main thermocline and is analyzed separately in Figure 4.11d-f. ExoCycle can operate on any user-defined depth interval; we focus on the surface and deep domains because they represent the two end-member physical regimes most relevant to volatile cycling models. We also expect that the basin structures vary significantly between these primary depth domains due to the contrast in processes that govern them.

To illustrate this, we applied our detection framework (see Fig. 4.11) to potential temperature fields from both the shallow and deep oceans (Fig. 4.18). As a methodological note, the deep-marine field was averaged from 1000 m down to the maximum GLORYS12V1 profile depth (5900 m). This averaging choice captures significant lateral temperature variations in the upper deep ocean (>1000 m and <2000 m), which contribute to the detected basin divisions (e.g., between the east and west Indian Ocean). The basin division between the east and west Indian Ocean and the North and South Pacific and Atlantic is partially a consequence of the shallower ocean (>1000 m and <2000 m) lateral temperature variation. However, at greater depths (~ 3000 m), temperatures are up to $\sim 1^\circ\text{C}$ cooler in the Eastern Indian, Northern Pacific, and South Atlantic oceans than in their conjugate oceans. We found that these differences are captured in the larger-scale community structure of deep and shallow marine environments (Fig. 4.18c-d). Our analysis confirms that these distinct thermal differences are captured by the algorithm, resulting in visibly different large-scale community structures for the shallow and deep domains (Fig. 4.18c-d). This analysis opens the door for more sophisticated studies. Further analysis of deep-marine basin-community structure may benefit from higher layer-detection resolution or an alternative layer-averaging scheme, e.g. applying a quantile transformation to each ocean layer before averaging GLORYS12V1 over

some interval (e.g., 0-200 or 1000-5900 m). Such a detailed study is beyond the scope of this paper, but the methods we describe now make it possible.

Our results confirm that deep and shallow ocean structures differ significantly within the major modern oceans (Pacific, Atlantic, and Indian). Yet, intermediate-complexity volatile cycling box models (e.g., LOSCAR and GEOCLIM variants; Fig. 4.1) implicitly assume that these structures are vertically aligned. They do this by, for example, defining a single "Pacific" box and then modeling vertical upwelling and mixing (deep $\rightarrow$ intermediate $\rightarrow$ surface) within that same geographic column.

This discrepancy raises a critical question: is it appropriate to horizontally average deep-ocean interactions at scales that force the global system into just 3 or 4 basins? If it is not, then paleoclimate reconstructions using simple box models with vertical walls are fundamentally insufficient to explain layered (shallow, intermediate, deep) basin fluxes and thermochemical properties. Instead, a new modeling process is needed. Intermediate-complexity models, especially those with a modular design (e.g., GEOCLIM7 or ESBMTK), may need to be custom-constructed for particular paleo- or exoplanetary ocean systems (e.g., Fig. 4.19). Such an analysis would require (1) updating the descriptions of the surface basins based on objective detection, and (2) incorporating multiple, non-vertical upwelling and mixing pathways between distinct deep, intermediate, and shallow basins. The parameters for such a model—mixing and upwelling rates—can potentially be estimated from physical data (bathymetry, SST, salinity, and atmospheric forcing, e.g., (178)). As has been shown, changing the spatial distribution and area of ocean basins (179), as well as corresponding riverine fluxes (180), can significantly affect the stability of ocean circulation. This new framework would allow these mixing and circulation estimates to be tuned for paleo-Earth models based on objective ocean-basin reconstructions and averaged chemical records, as in (12) and (138).

In summary, the common assumption of vertically-aligned, depth-independent basin boundaries is a critical simplification that may introduce significant errors into volatile cycle

modeling. The physical drivers at the surface (wind) and in the deep (thermohaline) create distinct 2D community structures at different depths. The methods presented here provide a clear, data-driven pathway to move beyond static, subjective box definitions. This allows for the construction of more physically-grounded, depth-resolved, and time-evolving intermediate-complexity models that objectively reflect the ocean’s true, complex 3D structure.

4.6. Discussion | Overall, our methodological and scientific results clearly demonstrate that ExoCycle enables the detection of ocean basins with consistent and objective methods for data-limited conditions. ExoCycle is fundamentally a tool capable of determining communities for any number of scalar fields on the surface of a sphere (or regionally) - with a range of post-processing tools that can be adapted for different use cases. In the following, we discuss some implications of our results in the context of Earth system modeling.

4.6.1. Thermochemical Fields |

4.6.1.1. Surface Ocean Basin Structure | Our analysis reveals that the global ocean’s structure is fundamentally layered, with shallow and deep domains governed by distinct physical (wind/eddy vs. thermohaline) and temporal scales - and thus the consequent ocean "basins" are different at different water depths. This detection of a layered structure, using an objective and statistically robust approach, is not merely a classification exercise; it establishes our framework as a powerful diagnostic tool for understanding Earth system processes.

Using this framework, output fields from high-resolution models or reanalysis products—such as temperature, pH, salinity, or dissolved oxygen—can be processed to objectively delineate coherent biogeochemical provinces. This provides a systematic method for characterizing how these distinct provinces, and the connections between them, respond to perturbations across a range of timescales. For example, one could trace the differential impact of a rapid CO_2 increase, the opening/closing of an ocean gateway, or the expansion

Layered Multi-Basin Interaction Box Model

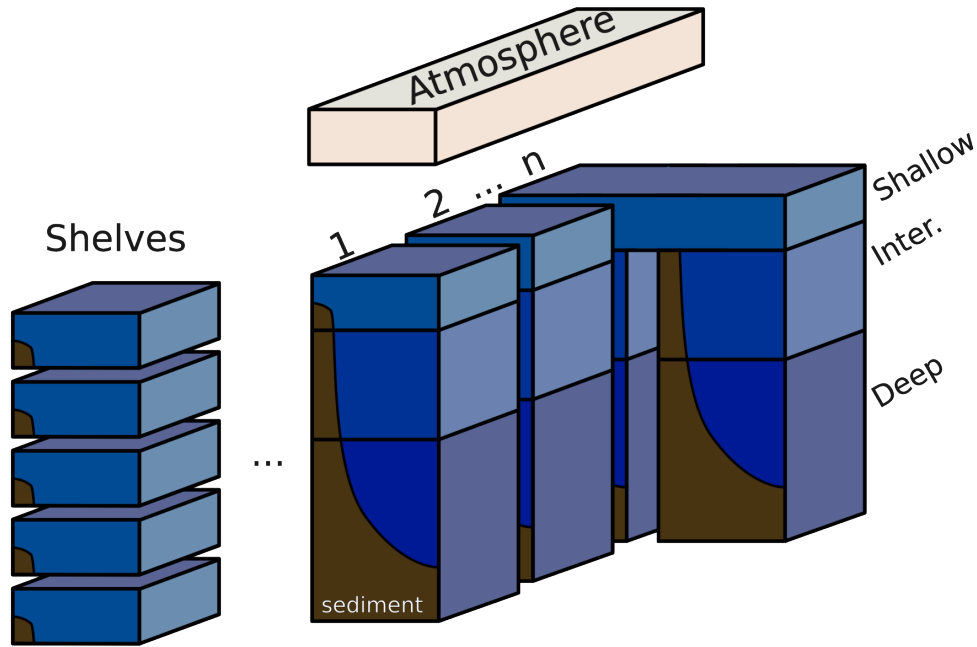


Figure 4.19: A simple example of an alternative intermediate-complexity Earth system model constructed with modular frameworks (e.g., GEOCLIM7 (22) or ESBMTK (23)). Multiple shelf regions feed into shallow and intermediate-deep boxes through surface currents and turbidity currents, respectively. The n^{th} box illustrates the characterization of non-uniform layered basin structure. ExoCycle’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 ExoCycle. However, outputs, e.g., connective bathymetry distributions (Fig. 4.13), sill depths (Table 4.4), and silhouette scores (Table 4.4 and Fig. 4.8a) can be used to heuristically determine the number of basins and strength of mixing/circulation for modular frameworks.

of anoxic zones across these objectively-defined basins. On shorter timescales, analysis of the modern annual drift of surface basins from changes in sea surface temperatures can elucidate the stability of surface ocean basins during rapid increases in atmospheric CO_2 (i.e., volcanic eruptions). **Our basin detection framework provides a systematic approach to characterize regions** influenced by perturbations to shallow ocean properties on either week-to-month timescales and longer with coupled atmospheric-ocean dynamics (e.g., multi-year - El Nino and La Nina, and decadal - post-industrial revolution (18) warming).

Our framework also allows for a more flexible basin detection analysis compared to current methods (Table 4.3). Flow-based approaches (e.g., (146)) 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., (145)) partitions the seafloor by morphological type but does not produce basin-scale communities from the data’s own covariance structure. Ecological provinces (e.g., (15)) capture biologically meaningful surface domains but are defined from modern observations with heuristic thresholds, limiting applicability to deep time. Hand-drawn definitions (e.g., (16; 33)) are based on expert judgment and are neither reproducible nor formally objective. ExoCycle 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 ExoCycle 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., $\epsilon\text{-Nd}$ and $\delta^{13}\text{C}$) onto the bathymetric basin

framework. This two-stage approach—defining containers from bathymetry and characterizing their contents from proxies—separates the geometric problem (which ExoCycle solves) from the dynamical problem (which requires additional data or modeling).

On the dynamical side, the concept of basin, or more generally, regional field stability is also central to assessing basin to planetary scale habitat stability. Many organisms, from phytoplankton assemblages to coral reef ecosystems, are adapted to specific environmental conditions localized within these basins (e.g., (181)). In this context, a "basin" effectively defines the spatial extent of a primary ecological niche. As environmental parameters shift, our method can be used to quantify how these niches shrink, expand, or drift in response. Analyzing the seasonal, inter-annual (e.g., ENSO-driven), or decadal drift of surface basins can reveal the system's sensitivity and resilience, offering a tangible way to measure the stability of marine ecosystems during rapid climate change.

4.6.1.2. Carbon Cycle Diagnostics and Carbonate Compensation Depth (CCD) |

One of the most immediate applications of our framework is in the generation of objective, regional-scale carbon cycle diagnostics, such as the Carbonate Compensation Depth (CCD). The interpretation of the carbon cycle over time relies on such diagnostics, which are known to have significant regional variability. However, these "regions" are often defined qualitatively or based on static, present-day geography (16; 33), which can spur inconsistencies when data are sparse (182). ExoCycle's workflow provides a systematic and objective method for grouping seafloor sediment core data based on the integrated water column chemistry that influences carbonate solubility.

Furthermore, our layered analysis directly addresses a key complication: the CCD is set by deep-water chemistry but is also heavily influenced by overlying surface carbonate productivity. Surface ocean productivity is an input to intermediate and deep ocean water chemistry, which together with regional seafloor distributions, determines the conditions for carbonate solubility throughout the region's water column (75; 183) and subsequently

Table 4.3: Basin identification method comparison. ExoCycle 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	ExoCycle	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

the regional CCD. The CCD has been found to have regional evolution within major ocean basins. For example, using temporally static surface basin descriptions, (16) found that the CCD in the western equatorial Pacific was ~ 1 km shallower and then ~ 250 m deeper than the eastern tropical Pacific at the start of the Oligocene and 0.5 Ma, respectively. (33) also determined regional scale CCDs for the northern North Atlantic, central North Atlantic, and South Atlantic. The use of static surface basin boundaries for the Pacific stems from the assumption that the thermocline depth and surface ocean conditions that drive carbonate productivity of different plankton species are constant throughout time (182). However, these qualitative descriptions can spur inconsistencies when describing 'statistically' independent regions (16; 182) which are actually evolving over time. Our method can, for the first time, objectively define the geometric intersections of distinct surface (productivity driven) and deep (chemistry driven) basins. More fundamentally, ExoCycle basin detections allow 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, ExoCycle defines basin domains $\Omega_1, \Omega_2, \dots, \Omega_k$ from paleo-bathymetry for a given time slice. Second, discrete proxy measurements (e.g., $\delta^{13}\text{C}$ from sediment cores at locations x_m, y_m) are assigned to the basin containing their paleo-coordinates, producing basin-specific proxy populations (eq. 4.10), where m is the proxy site index, k is the basin index, and B is a function that maps m to k .

$$B(m) = k, \text{ for } (x_m, y_m) \in \Omega_k \quad (4.10)$$

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. 4.11 4.12, and 4.13), where N_k is the number of sites in the k^{th} basin and P_m is a proxy site value.

Intra-basin mean:

$$\bar{P}_k = \frac{1}{N_k} \sum_{m \in \Omega_k} P_m \quad (4.11)$$

Intra-basin variance:

$$\sigma_{intra,k}^2 = \frac{1}{N_k} \sum_{m \in \Omega_k} (P_m - \bar{P}_k)^2 \quad (4.12)$$

Inter-basin variance:

$$\sigma_{inter}^2 = \frac{1}{N_k} \sum_{m \in \Omega_k} (\bar{P}_k - \bar{P}_{global})^2 \quad (4.13)$$

The ratio $\sigma_{inter}^2/\sigma_{intra}^2$ 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.

4.6.1.3. Basin Reconstructions and High Resolution Model Integration | A key challenge is extending this layered analysis into the deep past, as coupled model outputs like GLORYS12V1 are not available for paleo-reconstructions. However, increasingly high resolution global circulation models of sea-surface temperature, and other ocean field properties, run for specific paleo-time slices have become available (184). With these model runs, we can use outputs (e.g., sea surface temperature) for paleo surface ocean basin detection.

Deep ocean basin detection can also be more easily performed using reconstructed seafloor bathymetry.

Using paleo-reconstructed fields from high-resolution GCM simulations, ExoCycle 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.

Interestingly, we found a high degree of agreement between deep ocean basins detected using only bathymetry (Fig. 4.15) and those detected using modern thermochemical fields (Fig. 4.18). This agreement is a critical finding: it strongly suggests that major bathymetric features (mid-ocean ridges, transform faults, and continental gateways) are the first-order control on deep thermohaline circulation and effectively define the boundaries of deep-water masses. This validates the use of bathymetric reconstructions as a robust proxy for deep ocean basin structure. However, it also underscores the need for paleo-bathymetry reconstructions to be carefully reviewed for their representation, uncertainty, and inclusion of these critical basin-defining features. Appropriate bathymetry, and therefore basin, reconstructions will be integral for proxy comparison outlined in section 5.1.2.

4.6.2. Bathymetry Field | We found that volcanic seafloor features (seamounts, volcanic arcs, and plateaus), and continental gateways are responsible for delineating deep ocean basins (Fig. 4.15 and 4.17). This is expected since these features either impede or facilitate ocean water communication. Consequently, paleo-basin detections require a careful reconstruction

of these features and an understanding of their uncertainty.

4.6.2.1. Global Bathymetric Modeling: Gateways, Continental Shelves, & Ru-

gosity | Many global bathymetry models exist within the literature (e.g., (11; 33; 96)). Most modern bathymetry reconstructions include thermal subsidence of ocean lithosphere since it describes the first-order changes to ocean bathymetry. Some focus on the timing or morphology of gateway reconstructions (e.g., (79; 80)) beyond what is defined by major plate motions. The difficulty of shelf reconstructions often leads to the adoption of simple constant depth of low vertical resolution for shelves (106). However, methods have been developed to reconstruct shelf-slope-rise structure given the seafloor depth trends from modern passive and active margin tectonic settings (96). Furthermore, modern source-to-sink sedimentation modeling approaches can be used for predicting the shelf-slope-rise structure when climate parameters are better understood (69; 185; 186). These approaches are viable for global shelf-slope-rise reconstructions. However, we caution users that high quality shelf reconstructions require detailed stratigraphic analysis, modeling, and comparison to the fossil record of depth dependent species (e.g., corals) (181; 185; 187). Within the deep ocean, the rate of MOR spreading influences both the style of normal faulting and seamounts forming volcanism along mid-ocean ridges that contribute to deep ocean rugosity. Deep bathymetry rugosity gives rise to both the fuzziness of basin boundaries (this study) and vertical mixing (99). Reconstructions of seafloor rugosity for the recent past can be done directly with backtracking present-day measured seafloor. Furthermore, rugosity can also be statistically determined for reconstructions which have large areas of the seafloor covered with previously subducted MOR flanks by, e.g., using the MOR spreading rate as a proxy abyssal hill root-mean-square height (188).

4.6.2.2. Role of Oceanic Plateaus in Basin Evolution | Oceanic plateaus are one of the most influential features for delineating ocean basins. While there are bathymetry reconstructions that include plateaus (e.g., (33; 79; 96)), the effect of their formation on basin

boundaries and ocean water communication is poorly known within global reconstructions. The timescales of oceanic plateau formation are $\sim 1\text{-}10$ Myr, during which there are records of significant environmental changes such as ocean anoxia events (114). Since we have found that the boundaries and connectivity of basins are closely tied to the location of plateaus, we expect the global basin structure and connectivity to evolve throughout the series of plateau forming eruptions. Furthermore, oceanic plateau formation is temporally correlated with the establishment of seafloor spreading and have formed within each present-day major ocean basin - e.g., the Ontong Java Nui ($\sim 93.0 - 127.2$ Ma), North Atlantic Magmatic Province ($\sim 46.2 - 64.3$ Ma), Kerguelen ($\sim 89.9 - 127.9$ Ma), Caribbean ($\sim 69.6\text{-}142.4$ Ma), Shatsky Rise ($\sim 124\text{-}147$ Ma) large igneous provinces (114; 189). Their temporal emplacement into the open ocean and especially within forming oceanic gateways will change the timescale of ocean overturn (e.g., similar to (179)) and volatile cycling following an eruption at the start, during, and end of plateau formation. ExoCycle provides a framework to investigate changes in volatile cycling throughout plateau formation by parameterizing the basin and connectivity evolution. These sharp changes in ocean basin locations/areas can have, in turn, significant impacts on the vertical and lateral mixing in the ocean as well as the onset and duration of ocean anoxia events with confounding processes that lead to increased *surface ocean temperatures and organic carbon burial and decomposition*.

4.6.3. Toward a Comprehensive Planetary Carbon Cycle |

4.6.3.1. 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

Table 4.4: Basin system parameters from 0, 60, and 80 Ma basin detections (Fig. 4.17) compared to the LOSCAR 3-box model basin system (12). The gateway sills are estimated as the median depth of all shared basin boundary nodes. All values are calculated from basin detections in shown the Figure 4.17 except for modern gateway sills that are taken from (190). 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)	<i>Indonesia: 0.2-0.6, Gib: 0.3</i>	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

impose or inform changes in circulation (e.g., by a clustering analysis of proxy circulation constraints such as ϵ_{Nd} (50; 51)). Model constraints from additional paleo proxies and tuning of poorly constrained volatile cycle parameters will be application dependent.

To illustrate how ExoCycle output maps to box-model configurations, we compare detected basin structure across three time slices (Table 4.4). For the present day, SB-Reduction with Leiden ($\gamma = 0.01$, ensemble = 50) identifies 22 deep-ocean basins (Fig. 4.17d), of which 17 maintain silhouette scores above 0.7 (Table 4.4). 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. ExoCycle 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 (33). 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. 4.13d) provide the sill depths and gateway characteristics needed to parameterize exchange between these basins.

The geometric parameters extracted by ExoCycle 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, ExoCycle 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 (191): $Q = C_d W H_{sill} \sqrt{g' H_{sill}}$, 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. ExoCycle’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.

4.6.3.2. Exoplanet long-term habitability | Our basin detection algorithm also shows significant potential for studying the long-term habitability of exoplanets. Volatile cycling is a ubiquitous planetary process. Similar to Earth, planets like early Mars or those orbiting M-type stars may host stable oceans for extended periods (192; 193; 194). There is a growing effort to characterize which planetary conditions are capable of long-term habitability, partly a function of the carbon cycle’s ability to maintain stable surface oceans (27; 178).

Whether on a lifeless or a microbe-bearing world, these oceans are capable of carbon sequestration on the seafloor as well as within the water column as dissolved inorganic carbon (195). The style of this sequestration (e.g., shallow seafloor preference, similar to pre-plankton Earth) and the entire cycle’s stability have been shown to be sensitive to the ocean’s basin structure (11). However, it is currently unknown how to systematically characterize the ocean basin systems of Earth-like exoplanets based simply on their topography.

Our framework provides the missing link. Planetary topography, which is a direct

outcome of its governing tectonic regime (plate tectonics, semi-mobile-lid, stagnant-lid, etc; (54; 196; 197)), can be used as a direct input for our algorithm. This workflow can produce a suite of realistic global "exo-ocean" systems that can then be analyzed with intermediate-complexity carbon cycle models. These modeling efforts will facilitate the first systematic study of how different planetary tectonic regimes influence long-term carbon cycle stability and habitability through their direct control on ocean basin geometry.

4.6.4. Applications to other domains | A key strength of ExoCycle is its generality: the framework operates on any number of scalar fields defined on a sphere or regional subset thereof. *Beyond oceanography*, potential applications include objective climate classification from multi-field analysis of temperature, precipitation, and seasonality (replacing heuristic Köppen thresholds); identification of coherent geophysical domains from seismic tomography models (e.g., delineating Large Low-Shear-Velocity Provinces); delineation of ice-sheet drainage basins from surface velocity and topography fields; and characterization of multi-field ecological niches from remotely sensed vegetation and climate data. In each case, the algorithm provides objective, reproducible domain definitions that can evolve through time as the input fields change.

4.7. Conclusion | In this study, we developed and validated ExoCycle, a novel open-source computational framework for the objective, automated detection of communities from scalar fields on a sphere. Our method introduces several key technical innovations to overcome the limitations of heuristic definitions in data-limited systems. We implement a flexible pipeline that combines efficient graph-based community detection algorithms, such as Leiden and Louvain, with composite algorithms (e.g., SB-Reduction) and a novel, robust Difference Quantile Transformation Cumulative Density Function (DQT-CDF) edge-weighting scheme. This DQT-CDF method is a primary contribution, enabling the principled generation of spatial graphs from single or multiple non-normally distributed fields (e.g., bathymetry and temperature). Furthermore, our framework integrates ensemble-based agglomerative

clustering to rigorously manage the stochasticity inherent in these detection algorithms, using metrics like the silhouette score to assess the stability and objectivity of the resulting boundaries.

Applying this technically-grounded framework, our analysis demonstrates that the ocean’s basin structure is fundamentally layered—with deep, intermediate, and surface-ocean communities defined by distinct physical drivers—and that these structures evolve significantly over geological time. This finding, which directly challenges the "vertical wall" assumption in intermediate-complexity Earth system models, highlights the power of our multi-field and ensemble-based approach and suggests that this simplification may introduce significant errors in simulating global biogeochemical cycles. Furthermore, we find these basin boundaries evolve significantly over geological time and a spatial resolution of at least 1-2 degrees is required to capture the critical bathymetric features that define them.

The immediate application of our ExoCcycle framework is towards the development of more physically realistic, time-evolving, and layered box models for paleoceanography and carbon cycle analysis. Our method provides a data-driven method for aggregating sparse paleoproxy data, such as for regional CCD reconstructions. Beyond oceanography, the generality of ExoCcycle makes it a powerful tool for other scientific domains. It can be applied to characterize atmospheric weather regimes, define geodynamic reservoirs within the Earth’s mantle, delineate ice-sheet drainage basins, or systematically study the impact of tectonic regimes on ocean basin geometry and long-term habitability on exoplanets.

4.8. Code & Data Availability | The ExoCcycle python library is available on github under the GMD-release branch (<https://github.com/Bogumil-Matthew/ExoCcycle/tree/GMD-release>) under a GPL-3.0 license. The version (v1.0.0) used to produce results shown throughout this paper is archived on Zenodo ([10.5281/zenodo.17576467](https://doi.org/10.5281/zenodo.17576467), (198)). Both github and Zenodo repositories contain a jupyter notebook that shows the analysis and creation of each Figure and can serve as a guide for future analyses described

throughout the paper. Some datasets (e.g., ETOPO1 and GLORYS12V1) can be automatically downloaded from their source repositories using ExoCycle. However, all datasets used for our analysis can readily be accessed manually through their respective repositories: ETOPO1 at <http://doi.org/10.7289/V5C8276M> (9), reconstructed bathymetry at <https://doi.org/10.5281/zenodo.10460278> (113), and GLORYS12V1 at <https://doi.org/10.48670/moi-00021> (18). We have created a repository with exact input models used in our analysis that is stored within a separate Zenodo repository ([10.5281/zenodo.18664604](https://doi.org/10.5281/zenodo.18664604), (199)).

4.9. Acknowledgements | This project was supported by the National Aeronautics and Space Administration (NASA) under the Future Investigators in NASA Earth and Space Science and Technology (FINESST) award [80NSSC24K1716] to C.L-B. (PI) and M.B. (FI) were also supported by the Louis B. and Martha B. Slichter Endowed Chair in Geosciences fund.

The contents of this chapter appeared in a manuscript in review (25)

Chapter 5

ExoCycle Applications to Data Aggregation and Exoplanet Habitability

5.1. Introduction | Chapter 4 described applications of ocean basin detection to data-limited systems. Two high impact applications are (1) aggregating ocean data for use in paleo-climate studies and (2) describing how geological conditions, including planetary cooling regimes, produce differences in ocean basin system characteristics. Both contribute to a better understanding of the role of solid-planet processes in the long-term carbon cycle (LTCC).

In this chapter, I first briefly restate the current issues with ocean data aggregation when reconstructing paleoproxies to validate carbon cycle interpretations and models. I then (1) calculate the deep ($>1000\text{m}$) ocean basin systems for Earth reconstructions throughout the last 80 myr, (2) partition drill-sites among those basins, (3) evaluate change in basin-scale drill-site coverage throughout the reconstruction period, and (4) describe how changes in proxy basin-clustering can bias interpretations of carbon cycle simulations. These evolving basin partitions also allow me to evaluate intra-basin proxy averages and inter-basin difference. Second, I briefly restate the advantages and potential applications of ocean basin detection

for studying the LTCCs of terrestrial planets. I then (1) detect ocean basin systems for a set of proxy exoplanets topographies that are flooded to varying extents, (2) characterize those ocean basin systems, and (3) describe implications for LTCCs on terrestrial exoplanets and for future reduced-order models that use ExoCycle basins as spatial domains. Finally, I touch on the potential future applications of ExoCycle’s detection algorithm across the broader scientific community.

5.2. Data Aggregation | The limited representation of climate throughout Earth’s history constrains climate models and our interpretations of paleoclimate forcing events. Since the geological record is incomplete (e.g., spatially inconsistent and indirect in measurement), the community often clusters multiple data to represent stable ocean systems over some shorter timescale (see chapter 4). How those data are clustered significantly changes interpretations of where carbon is sequestered throughout the ocean (e.g., (16; 17; 143)) when compared global averages (6; 182) or regional reconstructions that are often used as global or beyond their spatial scope (4) (e.g., equatorial Pacific carbon cycle parameters mapped to entire Pacific basin). Figure 4.2 illustrates several basin divisions based on the present-day ocean basin configuration and the assumption of largely invariant ocean chemistry. However, Chapter 4 clearly shows that these assumptions are invalid over million year timescales. Therefore, I use ExoCycle’s basin detection framework to produce a set of period accurate ocean basins and characterize drill-site clustering for paleoclimate proxy data aggregation.

5.2.1. Results | I identified paleo-ocean basins using the procedure described in Chapter 4 (section 4.5.3 and Figure 4.17). I detected ocean basins from present-day to 80 Ma using reconstructed deep (>1000 m) bathymetry (Bogumil et al. 2024, (11)) in 5 Myr increments and ExoCycle’s ensemble (50) Leiden ($\gamma=0.01$) and small basin merger community detection with the DQT-CDF method (shortened 5σ , shifted 0.5σ) at 1 degree spatial resolution (see Chapter 4). Figure 5.1 shows a sample of evolving basin configurations at 10 Myr intervals, and Table 5.1 reports the corresponding detection and quality metrics. I use deep ocean

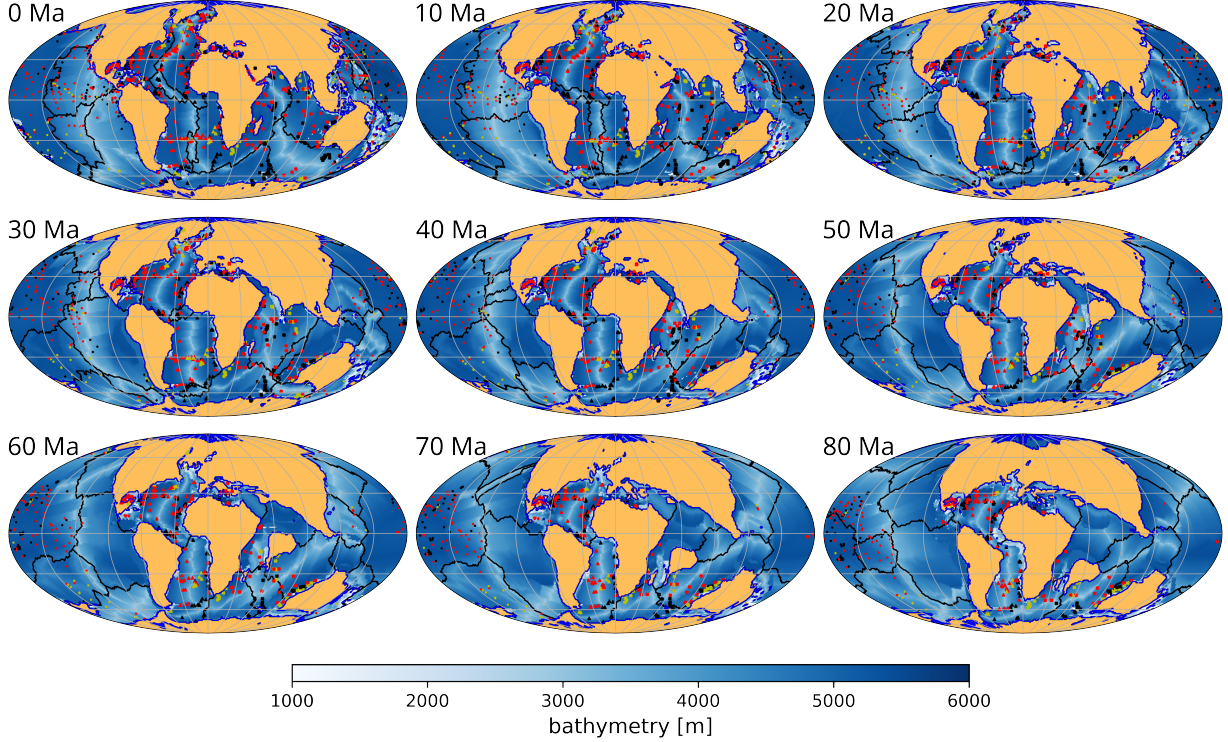


Figure 5.1: Basin detection on deep (>1000 m) reconstructed bathymetry. Ocean basins detected from reconstructed present-day bathymetry to 80 Ma at 10 Myr intervals (Bogumil et al. 2024, (11)) using the ensemble (50) Leiden ($\gamma=0.01$) and small-basin-merger community-detection method with DQT-CDF weighting (shortened 5σ , shifted 0.5σ) at 1-degree spatial resolution. All models have modularity and silhouette scores above 0.86 and 0.69, respectively. Table 5.1 reports all values, including the total number of basins, which varies from 15-23 throughout the reconstruction period. Symbols on reconstructed basins represent DSDP (red), ODP (black), and IODP (yellow) drill sites with present-day basin divisions: Pacific (dots), Atlantic (triangles), and Indian Ocean (X) from Bogumil et al. 2024 (11).

Table 5.1: Characteristics of basin detection on deep (>1000 m) reconstructed bathymetry (Figure 5.1). Modularity and silhouette scores are above 0.86 and 0.69, respectively, throughout the entire reconstruction period. The total number of large basins ranges from 15-23. The < 5 Sites row reports the number of basins with fewer than five DSDP, ODP, and IODP sites, while the underrepresented seafloor row reports the summed ocean surface area of those basins. Representation of the global ocean system drops sharply between 45-50 Ma.

Age (Ma)	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80
Modularity (%)	87	87	87	88	87	89	89	88	86	86	86	87	87	86	87	89	90
Silhouette (%)	78	69	77	88	88	87	86	80	91	82	77	85	88	89	78	78	84
Large Basins	19	20	17	18	17	23	23	21	18	18	17	16	17	15	15	20	18
< 5 Sites	1	1	0	0	2	3	6	3	4	3	6	7	5	7	5	11	6
Unrepresented Seafloor (%)	1	1	0	0	2	2	5	3	5	3	11	20	18	29	19	26	16

basin structure for my ocean drill site cluster analysis since these basins capture the relatively conservative divisions within which global ocean circulation operates. Globally averaged modularity scores and silhouette values are above 0.86 and 0.69, respectively, throughout the entire reconstruction period (Table 4.17), indicating high-quality basin detections (see Chapter 4). I also evaluated the number of drill sites (i.e., Ocean Drilling Program, International Ocean Drilling Program, and Deep Sea Drilling Project) within the evolving ocean basin system. The total number of basins range between 15 and 23 throughout the last 80 Myr. Within the last 45 Myr, only 0-26% of ocean basins (1-5% seafloor area) have fewer than five drill sites to represent them. This indicates that only a few small basins are underrepresented in the paleo seafloor record. Between 50-80 Ma, 29-55% of ocean basins (11-29% seafloor area) have fewer than five drill sites to represent their ocean chemistry. This indicates a switch toward many large, underrepresented basins in the paleo seafloor record.

5.2.2. Implications | The sparsity of ocean basin chemistry records clearly affects how the community interpret and model past climates. Before 50 Ma, seafloor core data are sparse

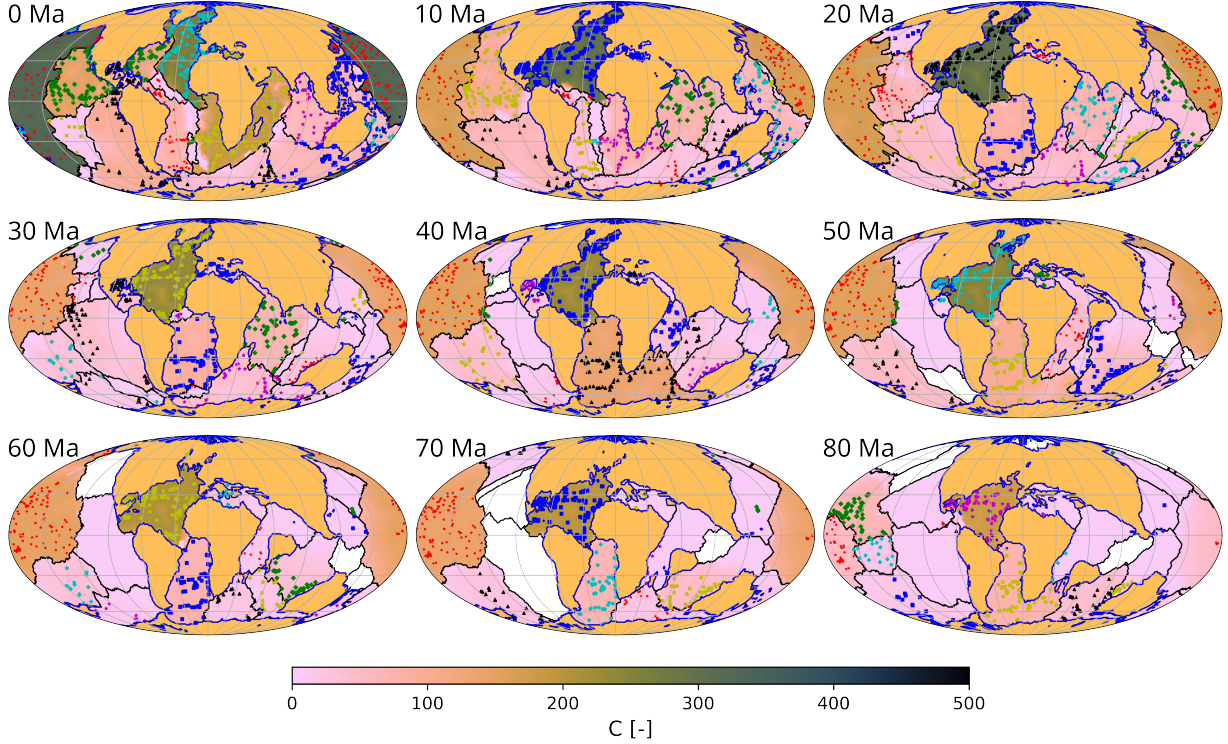


Figure 5.2: Proxy confidence (C) map for the detected basin shown in Figure 5.1. Confidence is calculated using equation 5.1. Symbols on reconstructed basins represent the collection of DSDP, ODP, and IODP drill sites. Unique symbol and color combinations distinguish sites in each basin cluster. White basins contain no drill sites. Darker basins (high C) generally have both more of the global drill site record and higher silhouette scores. Simple data assimilation schemes can overfit these regions when they do not account for the spatial nonuniformity of the record.

primarily across the young eastern Pacific and parts of the Tethys ocean because records from now-subducted seafloor are missing. However, even long-lived proxies (e.g., in modern East Pacific and Southern Ocean) cannot be used to represent neighboring independent ocean basins (Paleo- East/West Pacific and Southern Tethys oceans) without potentially missing the inter basin variations I observe at present-day and in the past few Myr (16; 31; 34). Overall, these results (Figure 5.1) clearly illustrate reconstructions of paleoclimate proxies, e.g. the carbonate compensation depth (CCD), are limited by both (1) reworking of Earth’s surface which destroys the geological record, and (2) the spatial expression of proxies throughout a basin system.

The evolving basin-scale drill site clustering shown in Figure 5.1 provides a framework for describing regional changes in ocean chemistry record. For example, well-established methods for the calculation of regional (basin-scale) carbon compensation depths (16; 31; 34) and basin-weighted organic carbonate burial history (143) could be repeated with evolving, not constant present-day, proxy clusters.

Additionally, I use missing basin records to inform interpretations of climate simulations. A simple approach to apply this for global circulation models, e.g. cGENIE, is to produce a post-processing confidence map from detected ocean basins and ocean system data. Equation 5.1 defines a confidence map in which confidence C is a function of the number of sites within a basin N_b , $\min_{k \in b} d(x, x_k)$ is the minimum distance between grid point x and drill sites x_k , $s(x)$ is the spatially resolved silhouette score (see Chapter 4), and $r = \sqrt{\frac{A_b}{\pi}}$ is the effective radius of the basin. Tying the length scale r to basin size increases the radius of influence of drill sites in larger basins. This choice is appropriate when basin definitions capture chemical variability over a similar length scale. Overall, the function responds to drill site clustering in three key ways: a point within a basin receives higher confidence when it is (1) closer to a drill site, (2) located in a basin containing a larger proportion of the global proxy dataset, and (3) associated with a higher silhouette score (basin stability metric). Figure 5.2 shows corresponding examples of the confidence map for Figure 5.1. Darker basins (high C) generally have both more of the global drill site record and higher silhouette scores. This analysis yields four key takeaways: (1) the drill-site record resolves Atlantic basins, especially the North Atlantic, better than nearly all other ocean basins over the last 80 Myr; (2) present-day drill sites near the Arabian Sea constrain central Tethys Ocean chemistry; (3) the drill sites used here do not constrain western Tethys Ocean chemistry; and (4) restructuring of Pacific Ocean basins between 60 and 80 Ma shifts where ocean chemistry can be reconstructed, from the northeastern Pacific toward the eastern and southeastern Pacific. This analysis can help evaluate the spatial confidence and bias of cGENIE-like outputs. Furthermore, I can extend this analysis to evaluate how proposed ocean drill site locations improve the spatial coverage

of the ocean-chemistry record and reduce clustering bias over a chosen time interval. The framework can also address the inverse problem: determining a series of drill-site locations from gridded modern seafloor topography that would best constrain the paleo-ocean chemistry record.

$$C(x) = \frac{N_b}{500} \exp \left[- \left(\frac{\min_{k \in b} d(x, x_k)}{s(x) \sqrt{\frac{A_b}{\pi}}} \right)^2 \right] \quad (5.1)$$

This analysis provides a framework for extending the sparse ocean proxy drill-site record to the global ocean system. I highlight two components directly applicable to carbon-cycle studies: (1) data aggregation for (i) creating basin averaged carbon cycle parameter fields that directly constrain ocean box models and (ii) calculating carbon cycle parameters that require spatial aggregation (e.g., CCD), and (2) confidence maps based on basin-proxy aggregation for evaluating (i) spatial sampling bias and (ii) the usefulness of proposed sampling locations for improving ocean chemistry coverage and confidence over a chosen time interval.

5.3. Exoplanets - exobasins & habitability | To first order, a balance between volcanic carbon emissions to the atmosphere and ocean carbon deposition on the seafloor, followed by recycling into the planetary interior, governs the long-term planetary carbon cycle. These processes are not unique to Earth and play a critical role in maintaining long-term planetary habitability. Chapter 2 demonstrated that seafloor bathymetry has a significant influence on the partitioning of carbon between the ocean-atmosphere system. Here, I address two related questions: (a) how seafloor-area distributions vary among measured terrestrial topographies (e.g., Mars, Venus, and the Moon) and statistically constrained exoplanet topographies that represent stages of planetary evolution (see Chapter 1), and (b) how to determine ocean-basin structure systematically and objectively, including where basin boundaries existed on Mars or would theoretically exist on exoplanets. I leverage ExoCycle (Chapter 4) to perform both analyses for different topographies and automatically define ocean

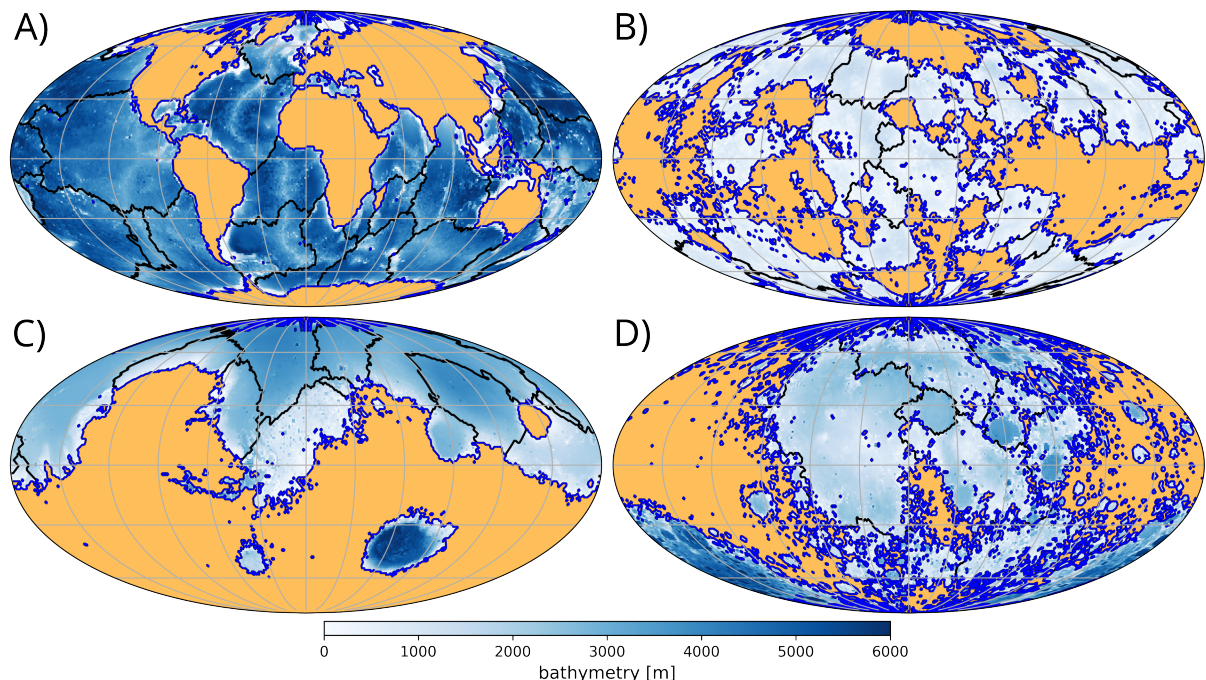


Figure 5.3: Planetary bathymetry created from (A) Earth, (B) Venus, (C) Mars, and (D) Moon when flooded with 2.6 (modern), 0.3, 1, and 1 GEL, respectively. Blue lines indicate coastlines and black lines indicate basin boundaries.

basins based on bathymetry fields. First, I apply this method to Mars and other terrestrial bodies to characterize variability in the spatial distributions of basins and bathymetry. I then use Earth system carbon cycle model results to describe how basin geometry and seafloor depth distributions influence seafloor, water column, and consequently atmospheric carbon storage across a range of conditions. Overall, my results help elucidate the key role that ocean basins play in determining and maintaining long-term planetary habitability.

Together, these analyses connect the geometry of flooded planetary surfaces to basin-scale ocean connectivity and the potential efficiency of volatile storage.

5.3.1. Results | I use flooded solar-system terrestrial topographies as analogs for exoplanet ocean systems whose surfaces may have formed under different planetary cooling, tectonic, and volcanic regimes. Specifically, I treat Earth, Mars, Venus, and the Moon as end-member examples of terrestrial topography and flood each surface with increasing Global Equivalent

Layers (GEL), where GEL is the volume of ocean water expressed as an equivalent globally uniform water layer. I then calculate ocean-basin divisions using the ExoCycle graph-theory-based community-detection framework applied to global bathymetry scalar fields. In these calculations, I construct ocean node connections from the unsigned spatial gradient of flooded topography and weight them using the spatially and data-field-normalized formulation outlined in Chapter 4.

The resulting flooded bathymetries show that planets with equal GEL can have substantially different amounts of exposed land and ocean-covered area (Fig. 5.3). These differences arise from the amplitude and wavelength of topography-forming processes on each planetary surface. Consequently, equivalent water inventories do not necessarily imply equivalent ocean geometries, basin structures, or shallow-to-deep seafloor area distributions. The spatial distribution of ocean basins (Fig. 5.3) and their connectivity are also sensitive to the magnitude of flooding, indicating that even modest differences in water inventory can alter ocean communication pathways between basins.

The flooded topographies also differ in their shallow and deep ocean structure. All topographies require extensive flooding to reduce the global fraction of shallow ocean seafloor, but Mars- and Earth-like topographic dichotomies are particularly favorable for producing deep oceans (Fig. 5.4). When interpreted using the carbon cycle sensitivity explored in Chapter 2, these bathymetric differences imply that many planet/GEL combinations would favor atmospheric pCO_2 values higher than those expected for Earth-like conditions. In this framework, increased shallow carbon burial lowers alkalinity and tends to raise atmospheric pCO_2 , whereas deeper ocean configurations can favor lower atmospheric pCO_2 depending on the balance between shallow seafloor area, deep marine depth, and carbonate preservation.

The basin-size distributions further show that no single dominant basin necessarily forms with increased flooding (Fig. 5.5). Instead, different planetary topographies produce different basin-growth regimes. On Earth, a small number of basins grow jointly as flooding increases. In contrast, additional basins can form on Mars- and Moon-like topographies as flooding

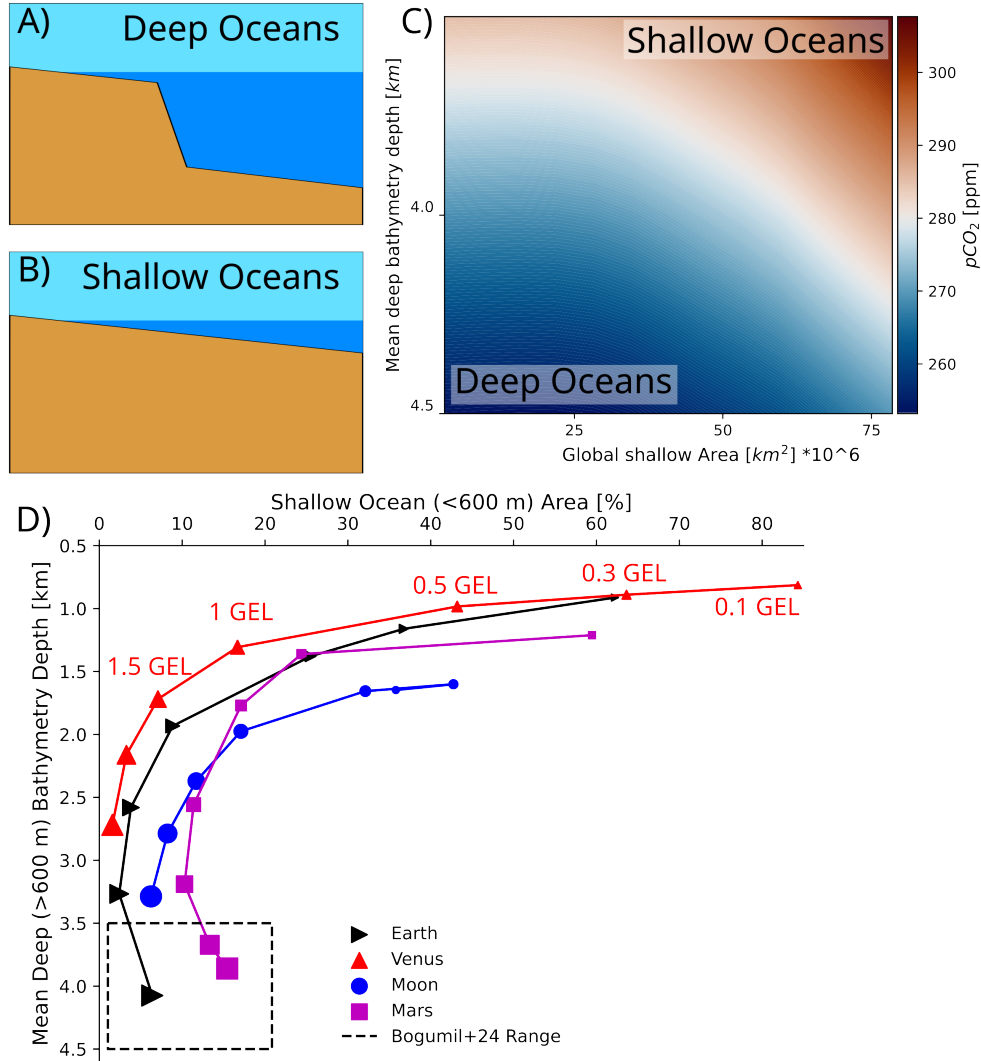


Figure 5.4: (A-B) Deep and shallow ocean diagrams, highlighting the expected change in atmospheric pCO_2 and ocean alkalinity. (C) Atmospheric pCO_2 from Bogumil et al. (2024) as a function of shallow seafloor area (< 600 m) and average deep-marine depth (> 1 km). (D) Changing bathymetry for increasingly flooded solar-system terrestrial-like planets, where larger symbols indicate larger GEL.

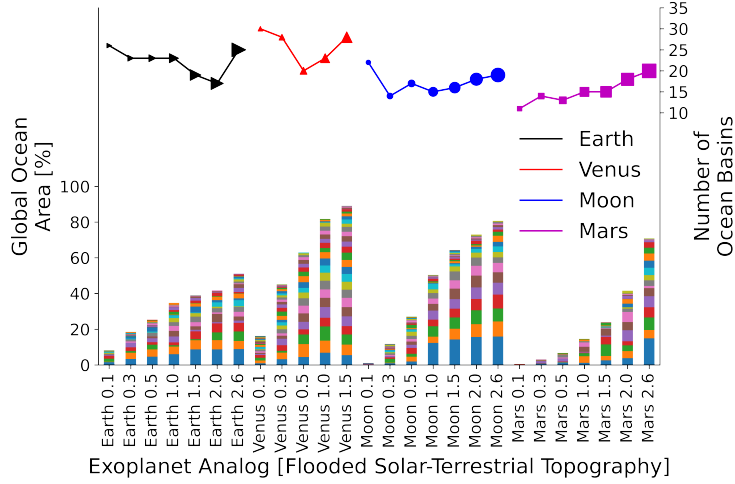


Figure 5.5: Stacked histogram of ocean basin sizes for planets flooded with various amounts of GEL.

connects new low-lying regions to the greater ocean system. These differences indicate that basin connectivity is not controlled by water volume alone. Planetary surfaces with large topographic dichotomies, such as Earth- and Mars-like surfaces, do not necessarily produce basins with similar connectivity or comparable inter basin water exchange.

5.3.2. Implications | These results suggest that exoplanet ocean basins could span a wide range of geometries and connectivity regimes, even for planets with similar total water inventories. Because basin-scale bathymetry influences carbon sequestration efficiency, ocean alkalinity, and atmospheric pCO_2 , the geometry of ocean basins may be an important and underappreciated control on long-term volatile cycling and long-term planetary habitability. Complex basin connectivity could produce hierarchical ocean chemistry, in which apparently large ocean basins are internally partitioned into smaller chemical reservoirs, analogous to the distinction between the North and South Atlantic within the broader Atlantic basin on Earth.

ExoCycle therefore provides a way to define the spatial ocean domains required for reduced-order volatile cycle models of Earth, solar-system terrestrial analogs, and potential exoplanet ocean worlds. Rather than prescribing ocean basins subjectively, the method

identifies basin boundaries directly from the geometry and scalar-field structure of the flooded planetary surface. This allows volatile cycling models to account for how different surface topographies and water inventories partition ocean area, shallow seafloor, deep seafloor, and basin connectivity.

The use of statistically constrained exoplanet topographies provides a path toward generalizing these results beyond the small number of terrestrial bodies for which topography is directly measured (151). By sampling plausible exoplanet topographic regimes, future work can identify the range of basin discretizations expected for rocky ocean worlds and evaluate how those discretizations influence carbon storage, atmospheric composition, and climate stability. In this sense, ExoCycle can help bridge future atmospheric measurements of terrestrial ocean worlds to the unobserved basin-scale surface conditions that regulate their volatile cycles.

More broadly, this application illustrates that ExoCycle is not limited to ocean basin detection; it is general spherical community detection framework for defining coherent spatial domains in planetary-scale datasets, motivating broader applications.

5.4. Applications to other domains | A key strength of ExoCycle is the generality of its basin detection algorithms: the framework is not ocean-centric but can operate on any number of scalar fields on a sphere, a set of spherical shells, or even a large regional grid. This capability opens applications far beyond oceanography and provides an objective approach to defining "basins," "domains," or "provinces" in data rich spatial fields. Below, I briefly describe several potential use cases in atmospheric science and climatology, solid-Earth geophysics and geodynamics, cryospheric science, and ecology and land-surface processes.

I can apply the same principles used to identify ocean basins to the atmosphere, a fluid shell with its own complex structures. In particular, the basin detection methods can support an objective climate classification scheme. The classic Köppen climate classification is a heuristic

system based on pre-defined thresholds of temperature and precipitation (200; 201; 202). My framework instead provides a path to an objective, data-driven climate classification. By feeding multi-field data (e.g., mean annual temperature, precipitation, and seasonality) into the algorithm, I can identify the "natural" boundaries of coherent climate zones on Earth, or any planet, based on the covariance structure of the data. On shorter timescales, the tool could analyze fields such as sea-level pressure (SLP) or geopotential height for the objective identification and spatial-temporal tracking of teleconnections and weather regimes.

I can also extend my layered basin analysis to characterize 3D communities within mantle simulations and geophysical models such as seismic tomography. For example, using a global seismic tomography model as a 3D input field (V_p or V_s anomalies), the algorithm could detect globally consistent, geophysically-coherent 3D volumes. This provides an objective method for delineating the boundaries of Large Low-Shear-Velocity Provinces (LLSVPs), identifying the 3D geometry of subducted slabs, and potentially finding the "roots" of mantle plumes or detached slab fragments in the mantle transition zone. As with ocean-atmosphere box models, my approach could also help develop updated multibox models of planetary evolution. Currently, these models use simplified domains/"boxes" (e.g., "depleted upper mantle," "primordial reservoir/LLSVP") that are often defined a priori (e.g., (203)). My framework could substantially improve this representation. By applying my methods to the 3D output of a mantle-dynamics simulation (e.g., temperature, composition, and/or viscosity fields), I could objectively define the boundaries of emergent dynamic reservoirs such as slabs, plumes, and other dynamically distinct mantle regions. This would enable next-generation, data driven multi-box models of mantle processing and volatile cycling.

ExoCycle operates at planetary scale and can analyze spatial data on any spherical body, thus, it has many applications in planetary science. For instance, the algorithm could be applied to fields of cloud-top temperature, composition, and wind velocity for giant planets like Jupiter. This would provide a quantitative, objective method for defining the boundaries of Jupiter's zones and belts and tracking their stability and interactions over time. Similarly,

the method could be used with input fields of ice surface velocity, bed topography, and surface mass balance to objectively define the major ice drainage basins—the catchment areas for outlet glaciers—which is a critical prerequisite for ice sheet stability modeling (204).

Finally, my newly developed methods also support many applications in ecology and land-surface processes. By integrating relatively static fields like topography, slope, aspect, and soil type with dynamic, temporal features such as mean temperature, precipitation patterns, and remotely-sensed vegetation indices (e.g., NDVI), the algorithm can move far beyond simple biome definitions to find regions of cohesive interaction. Instead of just mapping "desert" or "forest", my approach can objectively identify holistic, multi-field ecological domains—regions that share a cohesive set of environmental and biological characteristics. These data-driven "basins" can be viewed as the spatial extent of comprehensive ecological niches. I can then analyze how these niches and their boundaries shift, shrink, or expand in response to seasonal cycles or long-term climate change, providing a dynamic way to assess habitat stability. Furthermore, this approach can illuminate biogeography by analyzing animal migration pathways and determining if they are constrained by, or preferentially follow, the boundaries of these multi-field domains, revealing how animal species and vegetation interact with the full, integrated landscape.

Chapter 6

Conclusion

This thesis investigates how the tectonic evolution of rocky planets reorganizes ocean geometry and circulation structure, and how these changes influence long-term carbon and volatile cycling. I show that plate tectonic evolution influences the long-term carbon cycle (LTCC) not only through degassing and weathering fluxes, but also through the evolving bathymetric structure, connectivity, and spatial partitioning of the global ocean. Using coupled analyses of bathymetric evolution, mantle-flow-driven dynamic topography and geoid, and graph-based ocean basin detection, I developed reduced-order frameworks for quantifying how tectonic processes reorganize ocean structure and marine carbon storage through time.

Appendices

Appendix A

Supporting Information from Chapter 2

A.1. Methods |

A.1.1. (A) LOSCAR model description | The Long-term Ocean-atmosphere-Sediment Carbon cycle Reservoir (LOSCAR) models ocean-atmospheric chemistry through ocean circulation, mixing, air-sea gas exchange, degassing, weathering, and a biological pump within a box model representing the atmosphere, a high latitude ocean, and ocean basins with 3 water column boxes (Fig. A.1). Equation A.1 shows how LOSCAR represents transfer of tracers (TCO_2 , TA, PO_4 , etc) between boxes (e.g., basins' water column, high-latitude oceans, and the atmosphere, see Fig. A.1) as a set of first-order differential equations. F_{thm} , F_{gas} , F_{bio} , F_{in} , and F_{sed} represent thermohaline circulation and mixing, air-sea exchange, biological pump, riverine and weathering, and sediment deposition/dissolution fluxes while V_k and y_k represent box volume and tracer values.

$$V_k \frac{dy_k}{dt} = F_{thm} + F_{gas} + F_{bio} + F_{in} + F_{sed} \quad (\text{A.1})$$

The contents of appendix A appeared in The effects of bathymetry on the long-term carbon cycle and CCD (11)

Deposition and erosion (F_{sed}) of seafloor carbonates and water column dissolution are governed by 1) calcite and clay rain and 2) dissolution at the sediment interface. Calcite rain is primarily a function of atmospheric dependent CO_2 riverine flux input and its partitioning between basins and shelf vs deep marine (a constant parameter that is physically tied to transport) and the biological pump efficiency. Seafloor CaCO_3 dissolution (erosion) is not tied to benthic metabolic processes, but the local CO_3 and CO_3 saturation. CO_3 is calculated (in LOSCAR's `csys.c`) and depends on the H^+ tracer, depth dependent solubility coefficients, and dissolved organic carbon (DIC). Dissolution throughout the water column is represented at sediment layers by the change in depth dependent ocean chemistry that results from the above-described processes. For example, at 3000 m depth we might find 60% dissolution and 40% deposition occur, while at 3500 m there is 70% dissolution and 30% deposition instead. The other major component describing carbonate rain is the fraction of CaCO_3 remineralized in the deep marine (31% in the original LOSCAR paper (12), a tuned value) which is consistent with present-day observations of deep marine alkalinity. All of our models make no changes to equations that govern ocean chemistry. Therefore, for further details on LOSCAR's calculations, we direct the readers to the author of LOSCAR (12).

A.1.2. (B) Bathymetry basin volume corrections | Throughout 0-80 Ma and on 1-10s myr timescales, ocean basin volume changes with all the modeled processes (due to bathymetric changes, see main text methods) while ocean seawater volume changes only via (de)glaciation. At each reconstruction time step, we need to ensure that the ocean basin and seawater volume are equal. To balance seawater and basin volume, we first estimate seawater volume change resulting from glaciation by adding $\sim 1\%$ (for $5 \text{ Ma} \leq \text{age} < 40 \text{ Ma}$) and $\sim 2\%$ (for $40 \text{ Ma} \leq \text{age}$) to LOSCAR's initial present day ocean volume ($\text{VOC} = 1.29\text{e}18 \text{ m}^3$), simplifying Cenozoic trends (205; 206; 207; 208). We calculate ocean basin volume from the reconstructed bathymetry and paleo-DEMs (see main text methods) by determining the ocean-land boundary at each time slice while including global sea-level changes that flood

the paleo-topography (107). In our study, as well as others (79; 85; 106), globally averaged bathymetry methods (e.g., seafloor thermal subsidence relationships, simplified sediment reconstructions at passive/active margins) tend to result in reconstructions that overestimate basin volume (up to 10%) compared to the seawater volume. This is likely due to insufficiently modeled/unrepresented components of the seafloor bathymetry - anomalous and sparsely occurring thick seafloor sediments, seamounts, and submarine large igneous provinces that are not easily reconstructed throughout history.

Here, we correct for these biases by applying a correction based on present-day bathymetry. We determine a bias correction factor by first taking the difference between present-day reconstructed and measured global bathymetry distributions (9). We then apply this bias to each reconstruction, scaling the bias by a constant factor up or down until the ocean basin and seawater volume are exactly equal (Fig. A.12).

A.1.3. (C) Ocean volume calculations and changing initial conditions | LOSCAR (12) model calculations of ocean basin boxes assumes that ocean area (AOC) is equal at 100, 600, and 1000 m depth. However, this is not realistic since sediment layers at 600 m depth do not have any ocean volume laying between 600 and 1000 m depth. Regardless, the intermediate basin box volumes are calculated as $(1000 \text{ m} - 100 \text{ m}) \times \text{AOC}$. This leads to LOSCAR overestimating the volume of water partitioned to the intermediate ocean depths and therefore underestimates the deep basin boxes' volume. Figure A.2 shows that although the intermediate bin volume is overestimated, the total ocean volume input and value used in calculations (vb) are equal. A consequence of underestimating the deep box volume is to reduce the average depth of basin boxes in the hb variable and potentially change initial conditions in the deep box.

Affected Initial conditions can be identified with small changes in the values of deep marine CO_3 , pH, pCO_2 , and calcite saturation. For example, while alkalinity appears to be unchanged, it changes in response to the initial conditions. Therefore, alkalinity changes can

only be observed after an initial timestep. It might be tempting to suggest that these changes to initial conditions are what drive the CCD perturbations with changing seafloor area, mean depth, and distribution of deep marine depths. However, this is not the case. Consider the steady states of calcite saturation state for increasing seafloor area and decreasing average depth of deep bathymetry. Both bathymetry changes lead to a lower steady state calcite saturation state while the initial conditions trend in the opposite direction. Therefore, while initial conditions are not the cause of the steady state trends, they might mitigate their amplitude.

A.1.4. (D) Source to sink equilibrium | While the depth of calcite burial may change between earth system models with different bathymetry distributions, the total CaCO_3 sources and sinks must be in equilibrium for results to represent a steady state system. The CaCO_3 sink is represented by sediment layer deposition: 1.53×10^{12} kg CaCO_3/yr and is equal throughout all model results (Fig. A.5, A.8, and A.11). The CaCO_3 sources stem from 1) weathering CaCO_3 riverine flux (finc in LOSCAR) and 2) a combination of silicate weathering CaSiO_3 (fsi = 5×10^{12} molC/yr in LOSCAR) and volcanic degassing CO_2 (fvc0 = 12×10^{12} molC/yr in LOSCAR). Summing the sources supplied over the low latitude region (i.e., 90% of seafloor) results in a total of 1.53×10^{12} kg CaCO_3/yr .

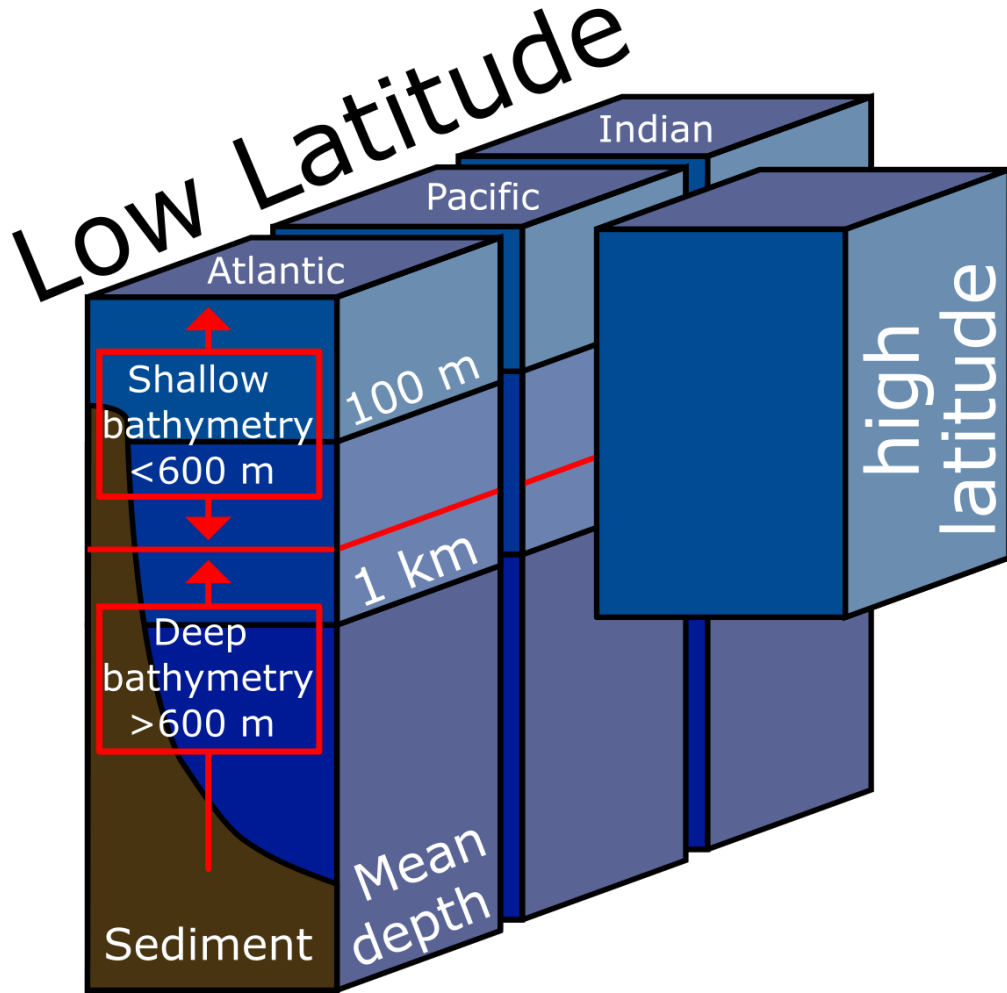


Figure A.1: Long-term Ocean-Sediment-Carbon cycle Reservoir (LOSCAR) model present-day setup. Pacific, Indian, and Atlantic oceans are represented with individual bathymetry (sediment layers) and consist of three reservoirs: Low latitude (≤ 100 m, L), interMediate (> 100 m and ≤ 1000 m, M), and deep (> 1000 m, D) boxes. The brown surface depicting sediment represents basin specific bathymetry in the lower latitudes (i.e., 90% of the lowest-latitude seafloor area). For this study we define < 600 m and > 600 m as shallow and deep bathymetry, respectively. The Early Paleogene setup shown in Figure 2.4 includes an additional 3 reservoirs for the Tethys ocean.

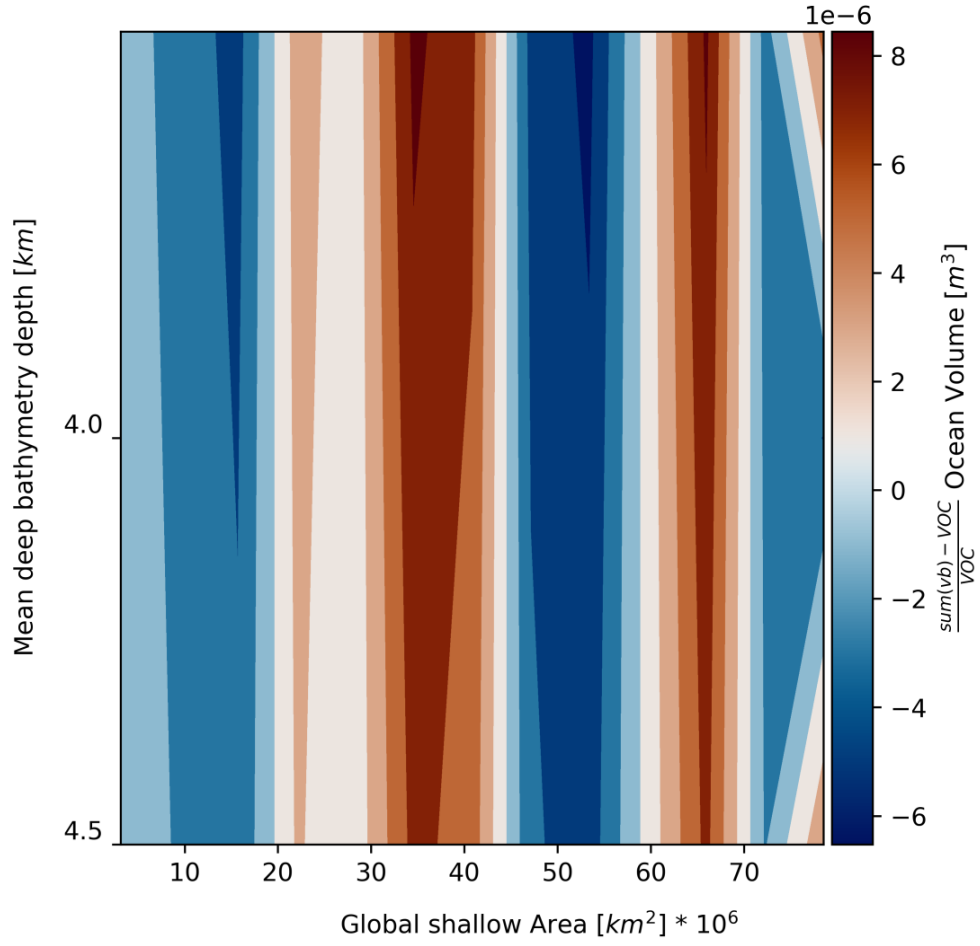


Figure A.2: Percent change in input ocean volume vs. ocean volume used in LOSCAR calculations. There is effectively no change in volume for changes to seafloor area or average depth of deep marine. However, the distribution of ocean volume from intermediate and deep ocean boxes does change and is not shown here.

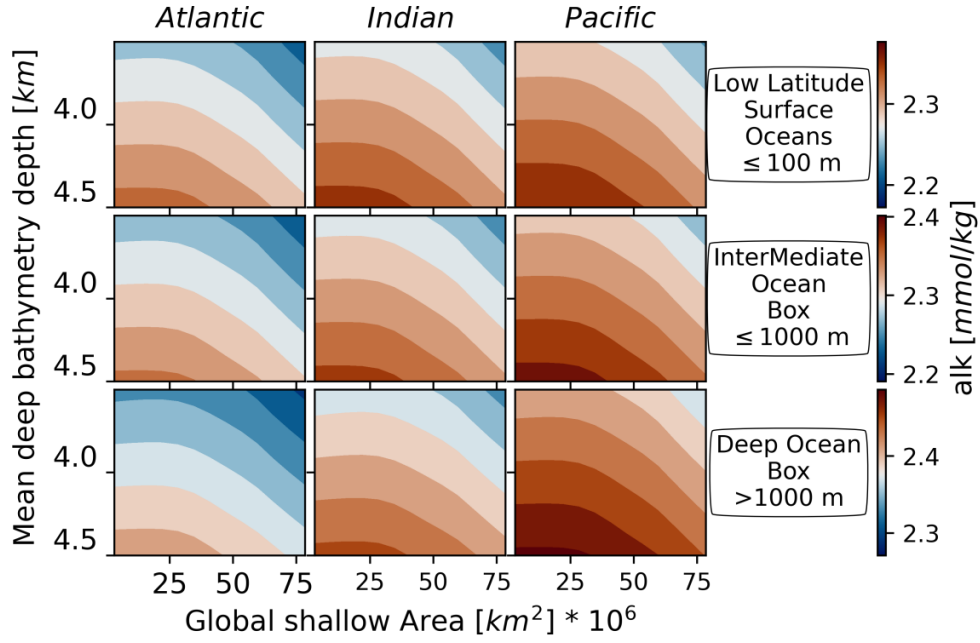


Figure A.3: Base model ($\text{nuwd} = 0.31$ and $\text{FBIOL} = 0.8$) response to alkalinity given synthetic bathymetric distributions representing first order changes in deep bathymetry and shallow area extent. Deep bathymetry is represented through Gaussian distributions within an Earth system model, with area averaged values plotted above. Shallow bathymetry areal extent is increased incrementally from 96 to 120% of its present value, producing similar areal extents to those within the last 80 Myr. Figure columns and rows represent ocean basins and basin box depths (L,M,D) from LOSCAR, respectively for base model setup.

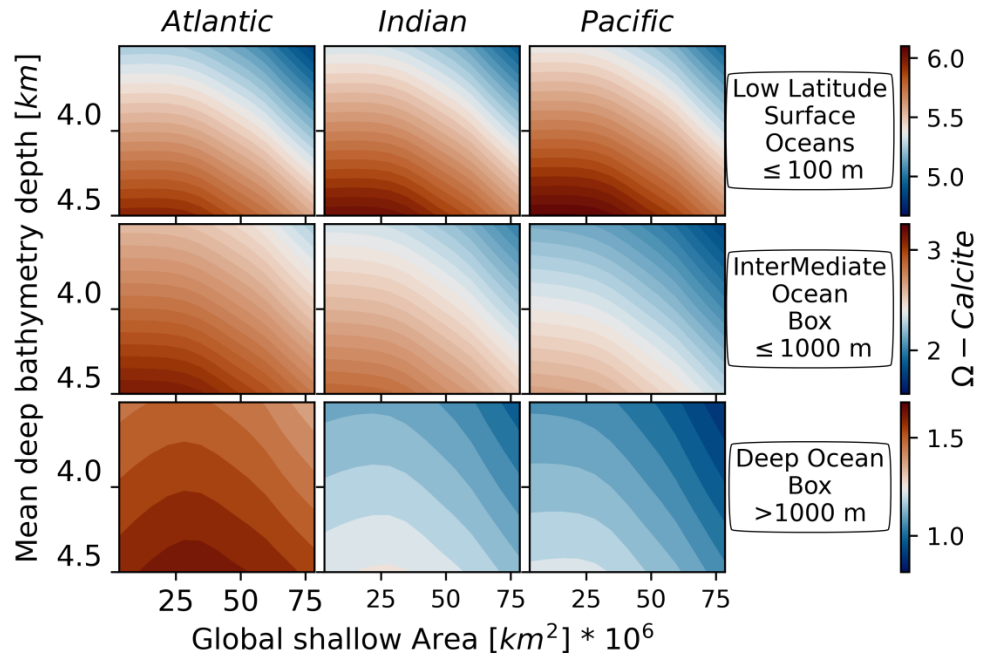


Figure A.4: Base model ($\text{nuwd} = 0.31$ and $\text{FBIOL} = 0.8$) response to calcite saturation state given synthetic bathymetric distributions representing first order changes in deep bathymetry and shallow area extent. Axes as described in Figure A.3.

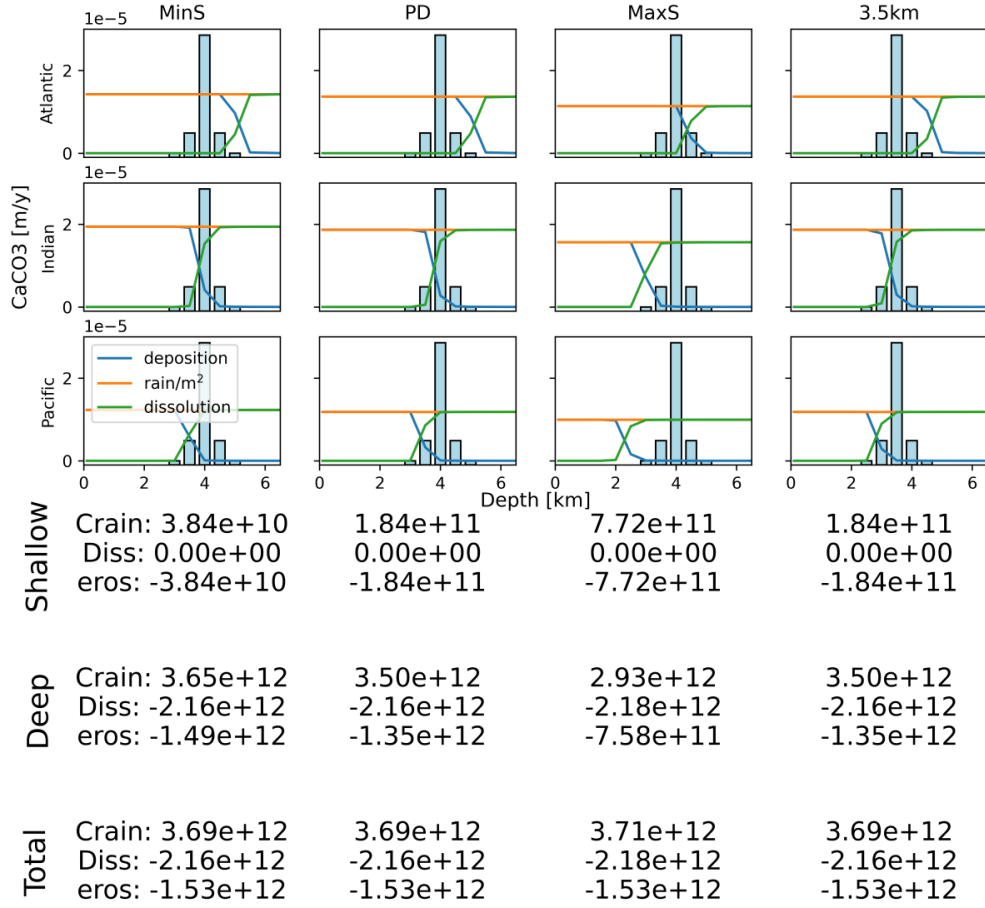


Figure A.5: MinS, PD, and MaxS represent LOSCAR model steady state runs with average deep bathymetry of 4 km and minimum ($3 \times 10^6 \text{ km}^2$), present-day ($\sim 30 \times 10^6 \text{ km}^2$), and maximum ($\sim 78 \times 10^6 \text{ km}^2$) shelf areas for base model ($\text{nuwd} = 0.31$ and $\text{FBIOL} = 0.8$). The 3.5 km model represents average deep bathymetry of 3.5 km with present-day ($\sim 30 \times 10^6 \text{ km}^2$) shelf area. Lines represent rain, dissolution, and burial rates in m/y while values below represent rates in $\text{kg CaCO}_3/\text{yr}$. For bottom values, negative erosion represents burial while negative dissolution represents dissolution occurring at the sediment layer interface. Shallow and deep values represent rates summed $< 600 \text{ m}$ and $\geq 600 \text{ m}$, respectively. Note that total burial ($15.3 \times 10^{12} \text{ kg/y}$) represents source rates from both silicate ($4.5 \times 10^{12} \text{ molC/yr}$) and carbonate riverine ($10.8 \times 10^{12} \text{ molC/yr}$) fluxes; summing fluxes and converting to $\text{kg CaCO}_3/\text{yr}$ results in a flux of $1.53 \times 10^{12} \text{ kg CaCO}_3/\text{y}$. The MaxS bathymetry model forces the Pacific into a dissolution only regime.

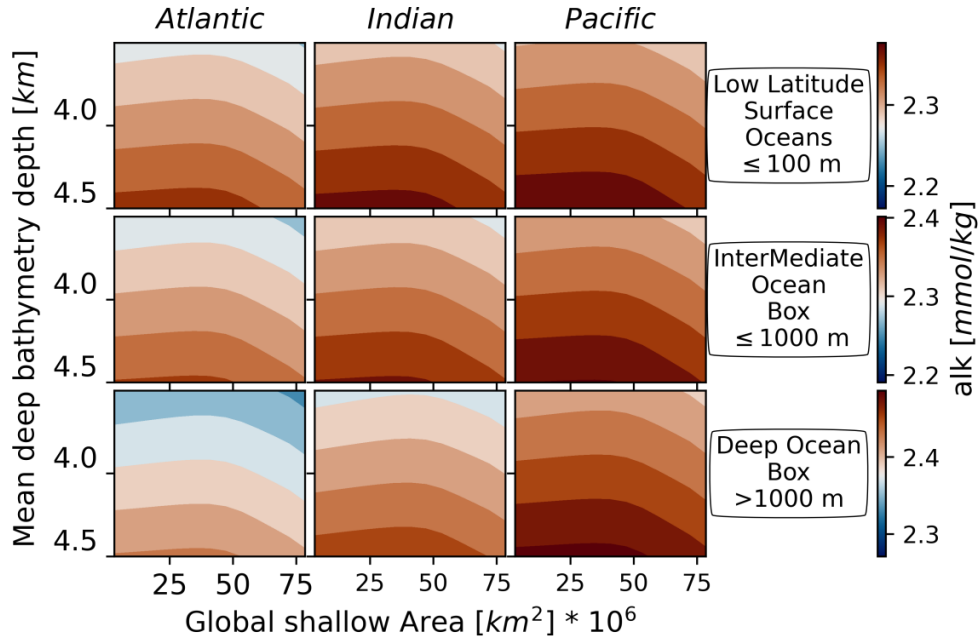


Figure A.6: Alkalinity response for the model with a 10% decrease in biological pump efficiency ($\text{nuwd} = 0.31$ and $\text{FBIOL} = 0.72$) and synthetic bathymetric distributions described in the main text. Axes as described in Figure A.3.

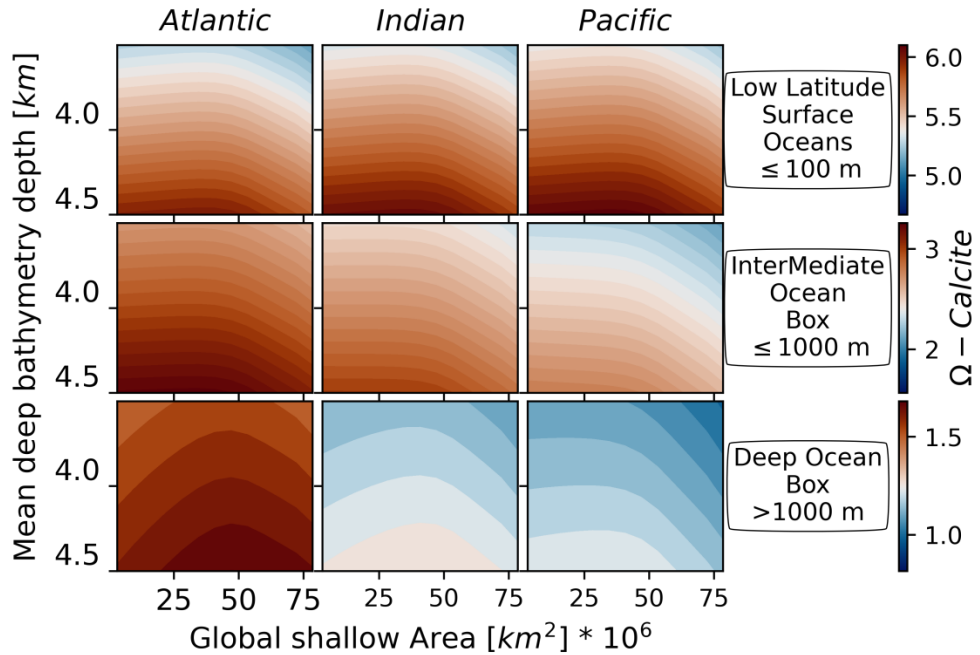


Figure A.7: Calcite saturation state response for the model with 10% decrease in biological pump efficiency ($\text{nuwd} = 0.31$ and $\text{FBIOL} = 0.72$) and synthetic bathymetric distributions described in the main text. Axes as described in Figure A.3.

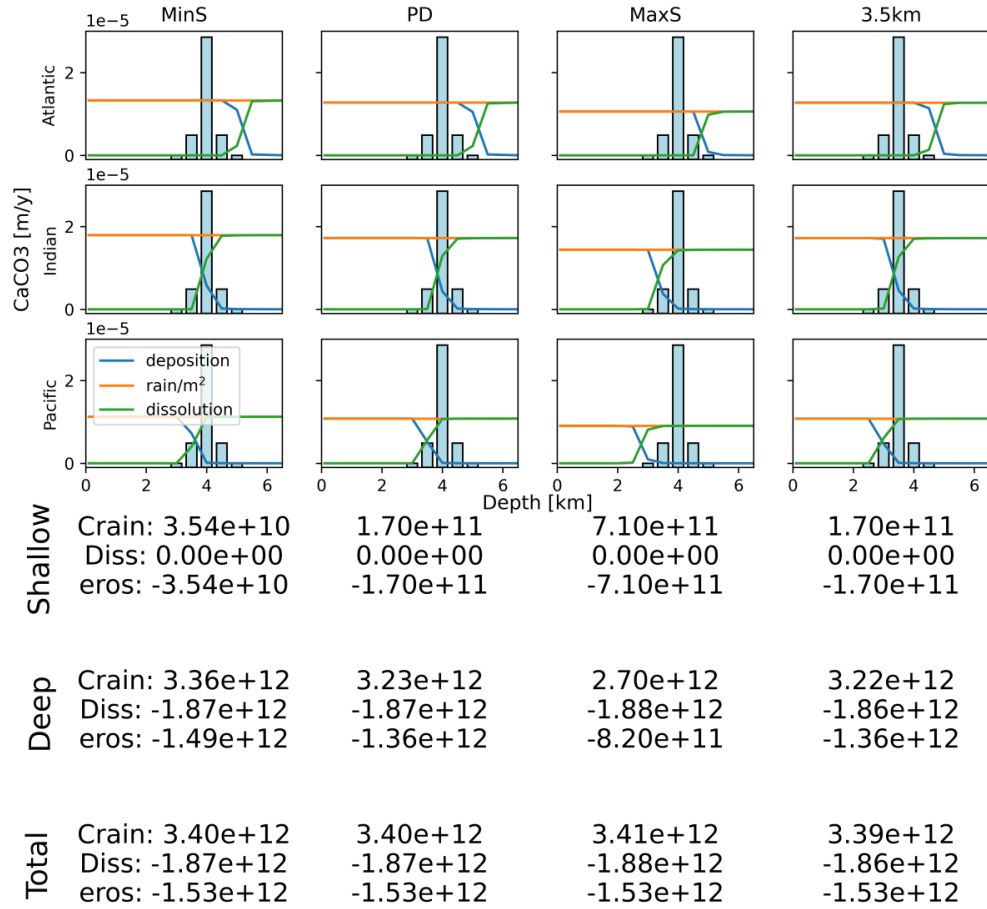


Figure A.8: MinS, PD, MaxS, and the 3.5 km LOSCAR model steady state runs for bathymetry described in Figure A.5. LOSCAR setups are similar to the base model (Fig. A.3, A.4, A.5) except for a decrease in the biological efficiency of 10% ($\text{nuwd} = 0.31$ and $\text{FBIOL} = 0.72$). Bottom rows of values are similar to the base models (Fig. A.5) with the exception of decreased calcite rain and deep marine dissolution. The MaxS bathymetry model forces the Pacific into a dissolution only regime, leading to high CCD sensitivity for changing shallow bathymetry area (Fig. 2.2).

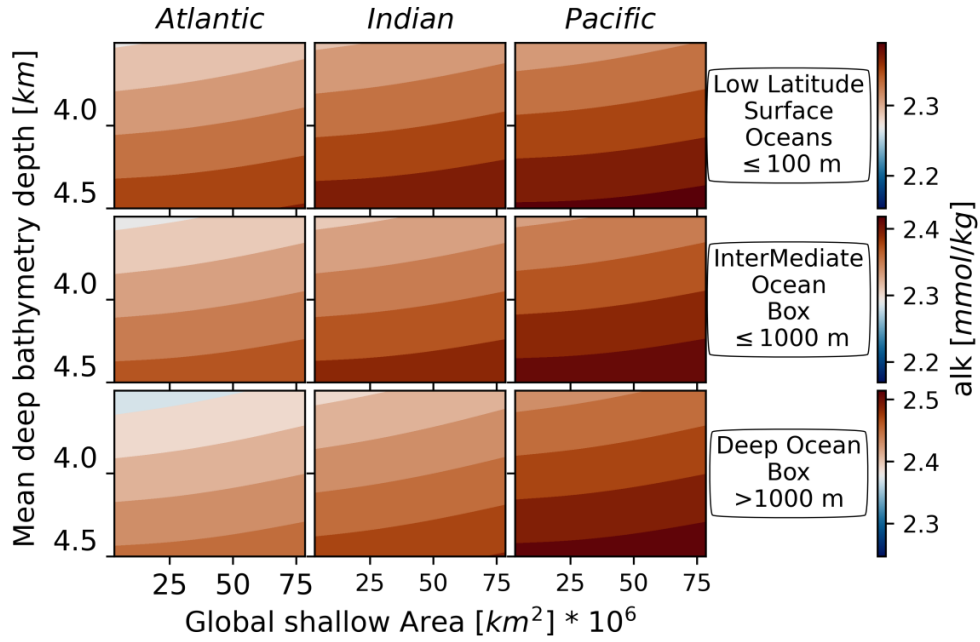


Figure A.9: Alkalinity response for the model with increased water column dissolution ($\text{nuwd} = 0.5$ and $\text{FBIOL} = 0.8$) and synthetic bathymetric distributions described in the main text. Axes as described in Figure A.3.

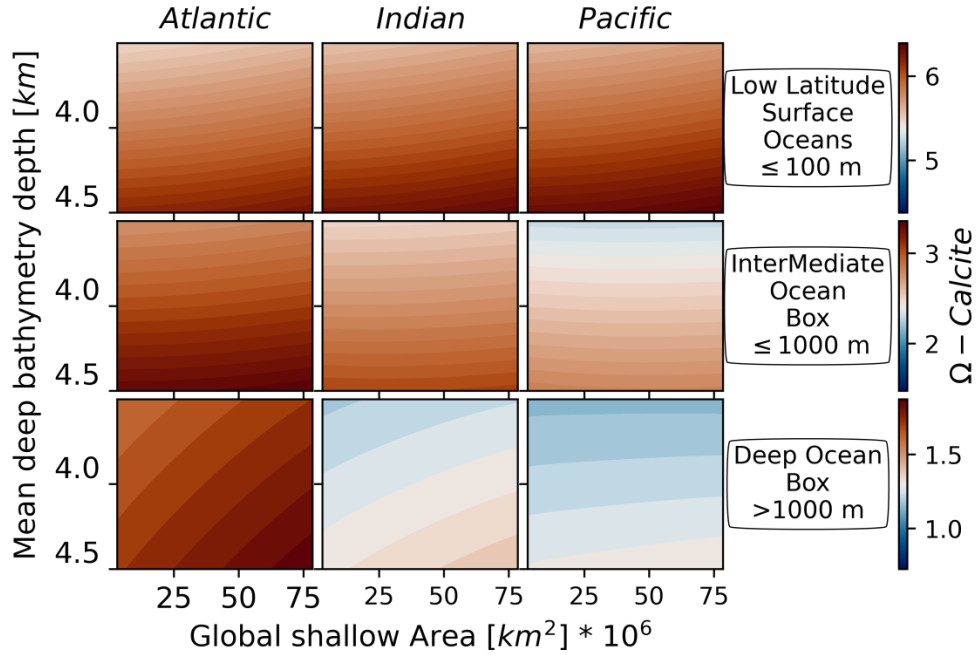


Figure A.10: Calcite saturation state response for the model with increased water column dissolution ($\text{nuwd} = 0.5$ and $\text{FBIOL} = 0.8$) and synthetic bathymetric distributions described in the main text. Axes as described in Figure A.3.

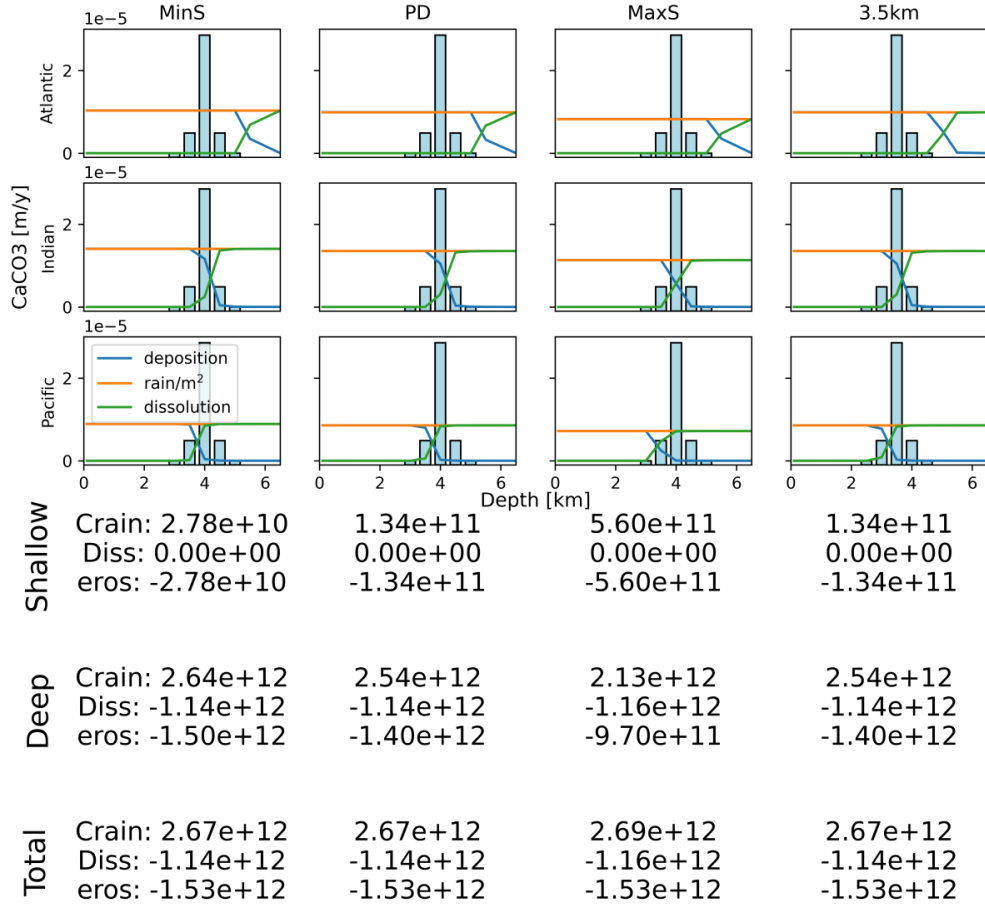


Figure A.11: MinS, PD, MaxS, and the 3.5 km LOSCAR model steady state runs for bathymetry described in Figure A.5. LOSCAR setups are similar to the base model (Fig. A.3, A.4, A.5) except for an increase in water column dissolution ($\text{nuwd} = 0.5$ and $\text{FBIOL} = 0.8$). Bottom rows of values are similar to the base models (Fig. A.5) with the exception of decreased calcite rain and deep marine dissolution greater than that in the decreased biological pump efficiency models ($\text{nuwd} = 0.31$ and $\text{FBIOL} = 0.72$). MaxS does not enter into a dissolution regime in any ocean basin as it does in the other two ocean systems (Fig. A.5 and A.8).

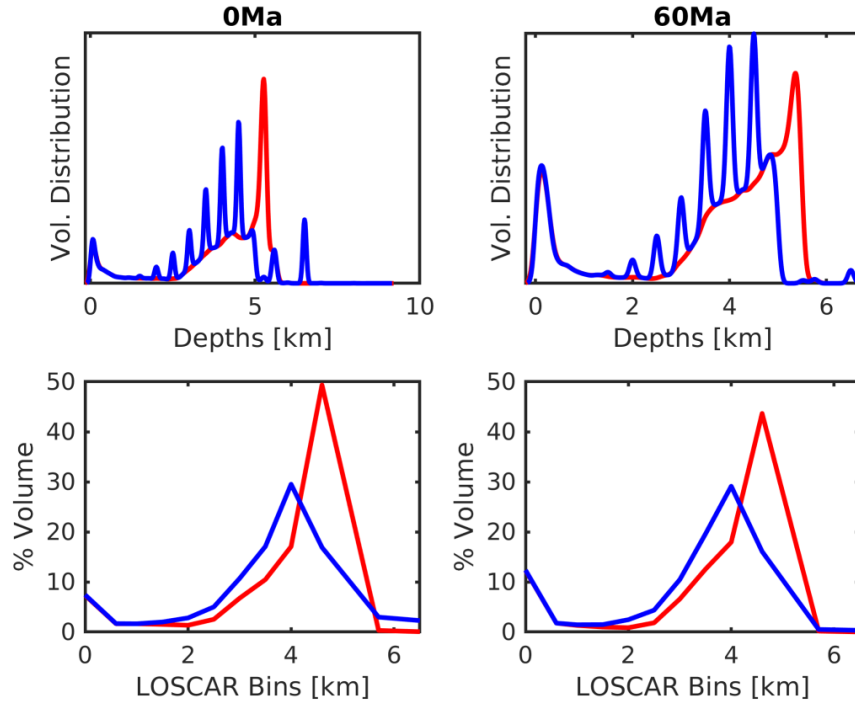


Figure A.12: Global ocean volume corrections applied to bathymetry distributions. A vector of correction factors for depths is calculated by taking the difference between Etopo and our present-day reconstructed bathymetry model distribution. Correction factors for each depth bin are proportionally applied to past and present reconstructions such that reconstructed distributions will result in realistic ocean basin volumes. Red lines represent the original distribution and blue lines the ocean basin volume corrected distribution. (Left) present-day. (Right) 60 Ma. (Top) Result of redistributing bathymetry sampled at ~ 20 m bin intervals. (Bottom) Same as (Top) using the LOSCAR sediment bin intervals. Bathymetry reconstructions tend to overestimate the global volume of ocean basins. Applied corrections shallow deep bathymetry while barely affecting shallow bathymetry.

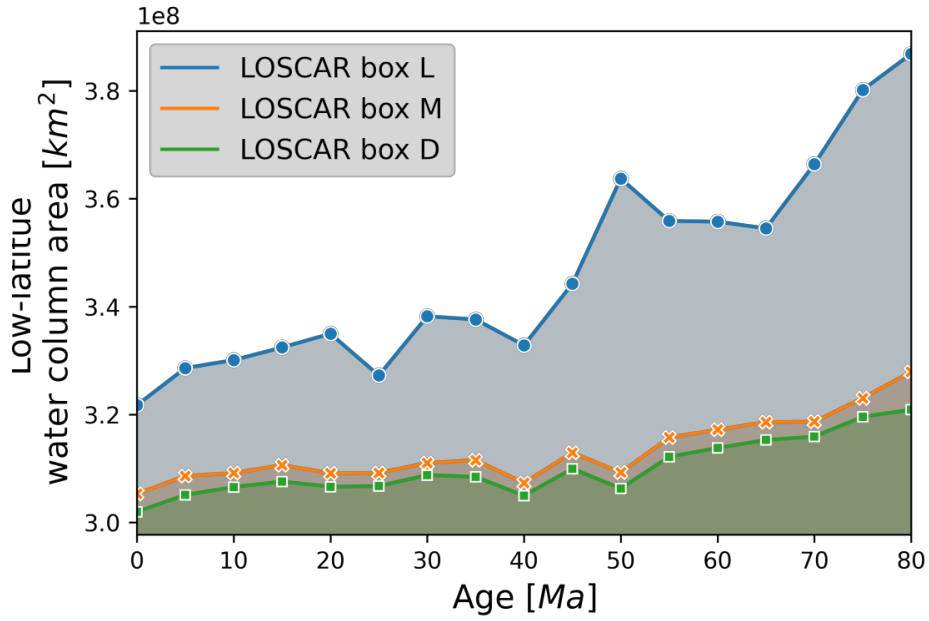


Figure A.13: Cumulative reconstructed low-latitude bathymetry area evolution throughout the last 80 Myr. Low-latitude refers to the lowest 90% of seafloor area and describes the area where bathymetry distributions inform LOSCAR. While the area of deep bathymetry (>600 m depth) is relatively constant, shallow area (<600 m depth) changes significantly due to continental flooding.

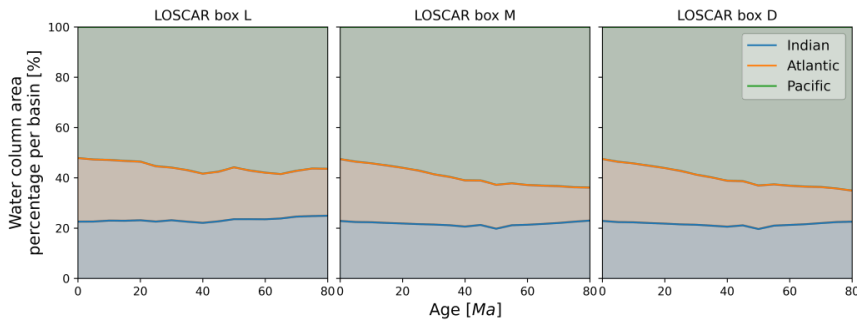


Figure A.14: Evolving distribution of global basin system low-latitude seafloor area for LOSCAR depth boxes (L, M, and D, see Fig. A.1). Calculations as in A.13.

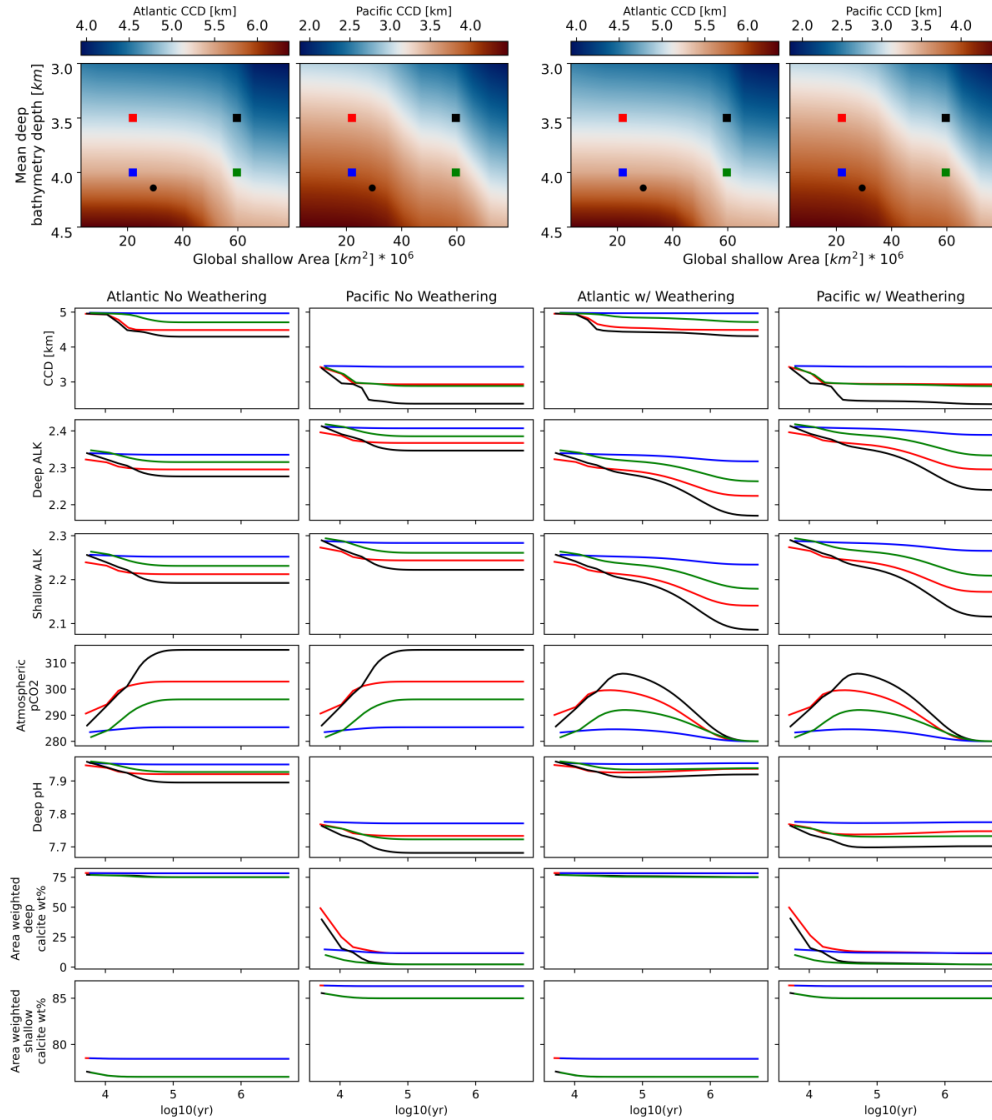


Figure A.15: Evolution of Atlantic and Pacific basin ocean chemistry of base model (Fig. A.3, A.4, A.5) with (Left) and without (Right) weathering feedbacks. Blue and Red dots and lines represent models of smaller shelf area values for average deep and shallow deep marine, respectively. Green and black dots and lines represent models of larger shelf area for average deep and shallow deep marine, respectively. Models with disabled weathering feedback all experience an increase in atmospheric carbon and decrease in shallow and deep marine alkalinity and loss of deep Pacific carbonates. Weathering feedback enabled models experienced the same loss in deep Pacific carbonates. Initially, deep and shallow marine alkalinity increased and $p\text{CO}_2$ climbed. As $p\text{CO}_2$ relaxed to present-day values (280 ppm) through temporary increased weathering, silicate weathering captured additional carbon, reducing deep and shallow ocean alkalinity further.

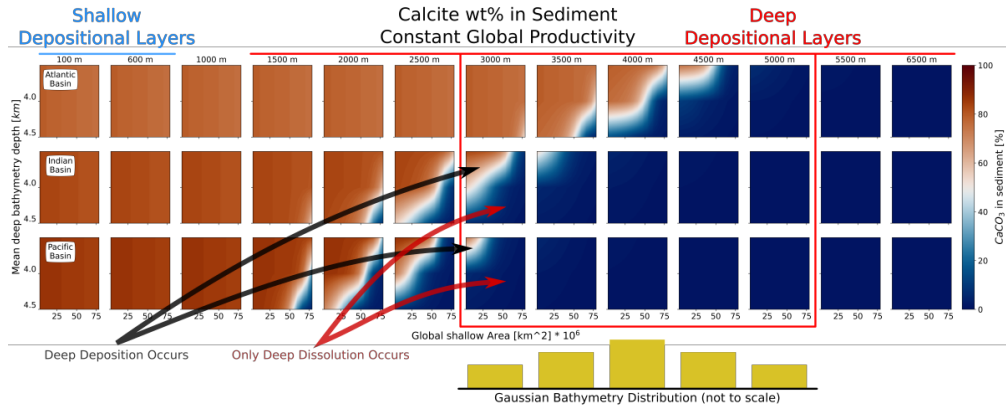


Figure A.16: Seafloor carbonate concentration steady-states for default LOSCAR sediment layer depths and basins (see Fig. A.1). Subplots axes as in Figure A.3. Contours in each subplot represent low (blue) and high (red) CaCO_3 sediment concentrations for varied bathymetry. Deep bathymetry with a mean of 4.5 km only has seafloor and the potential for carbonate sequestration within the shallow marine and between 3.5-5.5 km depth. For this deep bathymetry model, the CCD kink (red and black arrows), described in Figure A.5 is clear. Increasing shallow seafloor leads to an increased shelf carbonate supply and a lack of deep Pacific carbon.

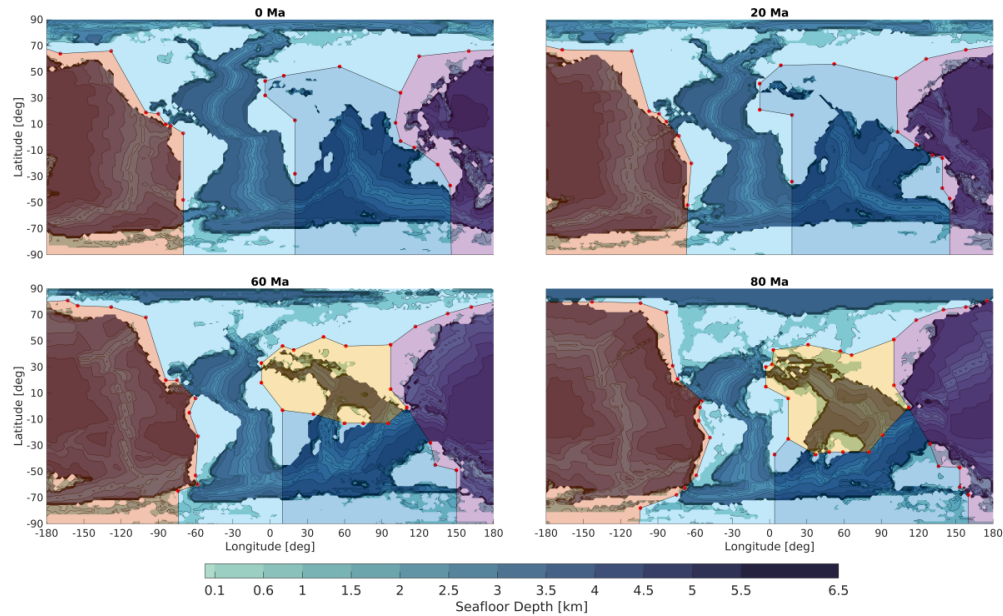


Figure A.17: Basin boundary evolution at 0, 20, 60, and 80 Ma. We found no community defined systematic method for defining paleo basin boundaries. Therefore, we drew basin boundaries along bathymetry informed chokepoints, which communicate ocean waters between basins. While the Tethys ocean basin is described during its existence as a separate basin, the 3 basin models merge the Tethys and Indian oceans to describe a single Indian ocean basin (Fig. 2.2 and 2.3).

Appendix B

Supporting Information from Chapter 3

B.1. Kinematic Subduction Model |

B.1.1. Characterizing Slab Deflection | Slab morphology controls the majority of the mantle’s excess mass distribution. Although seismic tomography reveals many types of slab morphology (209), penetrating and deflected slabs appear ubiquitous and readily identifiable in the recent past (<200 Myr). What processes produce these slab morphologies?

Subducting cold, differentiated ocean lithosphere experiences increasing pressure (P) and heating (ΔT) from the surrounding mantle. Both the chemical make-up of ocean lithosphere and changes to pressure and temperature lead to ocean lithospheric layers undergoing different phase transitions. **One major phase transition** that both the mantle and slabs experience is the ringwoodite-bridgmanite (ri-bg) phase transition. The ri-bg transition occurs at ~ 660 km depth within both the mantle (pyrolite) and the peridotite layer of ocean lithosphere. This phase transition increases mantle and slab density below 660 km and has a negative Clapeyron slope ($\frac{dP}{dT} = \frac{dP}{dz} \frac{dz}{dT} < 0$). When cold ocean lithosphere subducts, it reaches the phase transition but may not be hot enough to transition to the denser bridgmanite phase. If this occurs, the ri-bg phase transition impedes further sinking of the ocean lithosphere until

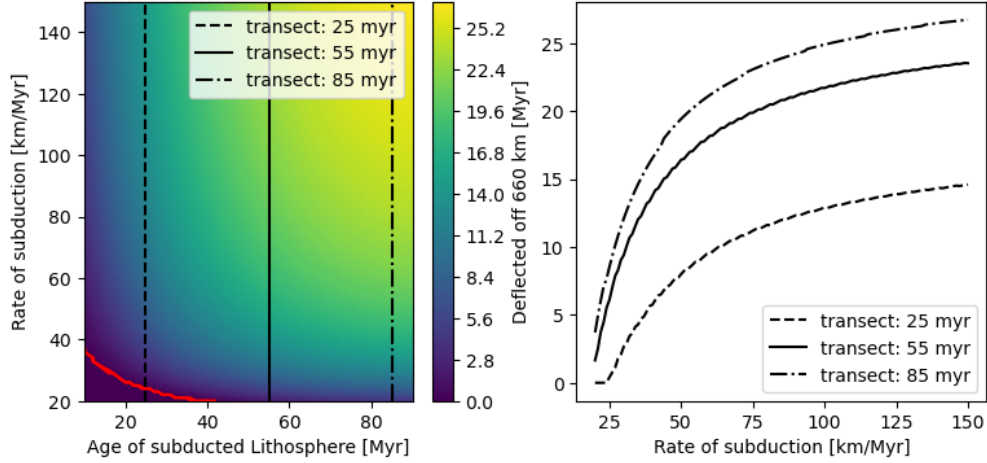


Figure B.1: Example of slab deflection look-up table for slabs that deflect off the ringwoodite-bridgmanite phase transition at ~ 660 km depth. **Left)** period of slab deflection as a function of the radial rate of subduction in the upper mantle and age of the subducted ocean lithosphere. The red line indicates the boundary between deflected and non-deflected slabs. The black lines correspond to the profiles plotted in the right panel. **Right)** Slab deflection-period profiles as a function of radial subduction rate in the upper mantle.

it warms sufficiently to transition to bridgmanite.

This presents a challenge for simulating the history of mantle heterogeneities. `slabmodellib` provides two approaches for modeling slab deflection at the ri-bg phase transition. **The first approach** applies a constant deflection period to all subducted slablets regardless of their thickness, initial thermal profile, or subsequent heating. This approach is useful for comparison with seismic tomography models and for determining which deflection period best reproduces present-day mantle observations. **The second approach** determines the deflection period based on the thermal evolution of ocean lithosphere (slablets) and the properties of the mineral phase transition. Ultimately, the subduction rate, age of the ocean lithosphere, and ri-bg mineral and phase-transition properties determine the slablet deflection period at the transition depth (~ 660 km), allowing construction of a lookup table.

The balance between gravitational force per unit area, arising from the slablet's excess density relative to the ambient mantle, and positive buoyancy at the ri-bg phase transition, arising from the higher density of the underlying ambient mantle relative to the overlying slablet, determines slab deflection.

The gravity force per area for a slablet is calculated using the surface density contrast of a slab per unit area (Eq. B.1) with respect to the mantle from Ricard et al. 1993 and Lithgow-Bertelloni et al. (1998) (119; 124).

$$F_g = g\Delta\sigma = g\Delta\rho\sqrt{\frac{age}{90}} \quad (\text{B.1})$$

The relationship derived by Weidner and Wang (1998) (210) gives the counteracting positive buoyancy force per unit area at the ri-bg phase transition (Eq. B.2), where v_i is the volume decimal percentage of phase i, $\Delta\rho_i$ is the i^{th} phase density change, $\left(\frac{\partial z}{\partial P}\right)_i$ is a depth-to-pressure conversion, and $\left(\frac{\partial P}{\partial T}\right)_i$ the Clapeyron slope of the i^{th} phase. $\Delta T = T_{ref} - T_{lith}$ is the excess temperature associated with the density anomaly, where T_{ref} is the temperature above the phase transition and T_{lith} evolves by a 1-D thermal evolution model with evolving ambient mantle boundary conditions as the slablet sinks through the upper mantle. HeFESTo calculates the phase-transition parameters for a harzburgitic composition used in my analysis (211; 212; 213; 214). I allow only the ri-bg volume of the slablet to contribute to F_b (i.e., v_i represents the total bridgmanite fraction formed below 660 km depth). Complete harzburgitic density, pressure, and depth profiles across the ri-bg phase transition determine $\Delta\rho_i$, $\left(\frac{\partial z}{\partial P}\right)_i$, and $\left(\frac{\partial P}{\partial T}\right)_i$. The decrease in ringwoodite and increase in perovskite (i.e., theoretical bridgmanite mineral) in the HeFESTo fort.67 output define the phase-transition depth interval.

$$F_b = \delta * g = \int (\Delta\rho) dz = - \sum v_i (\Delta\rho_i) \left(\frac{\partial z}{\partial P}\right)_i \left(\frac{\partial P}{\partial T}\right)_i \Delta T \quad (\text{B.2})$$

When $F_g < F_b$ at 660 km depth, positive phase-boundary buoyancy deflects the slab. Once T_{lith} increases sufficiently that $F_b < F_g$, gravitational forcing exceeds phase-boundary buoyancy and the slab resumes sinking. By simulating slablets that subduct after X Myr of surface cooling and sink radially through the upper mantle at Y km/Myr, I construct a lookup table for deflection period as a function of subducted ocean-lithosphere age and upper-mantle subduction rate. Figure B.1 shows an example of this table.

B.2. Comparisons with Various Dynamic Topography Models |

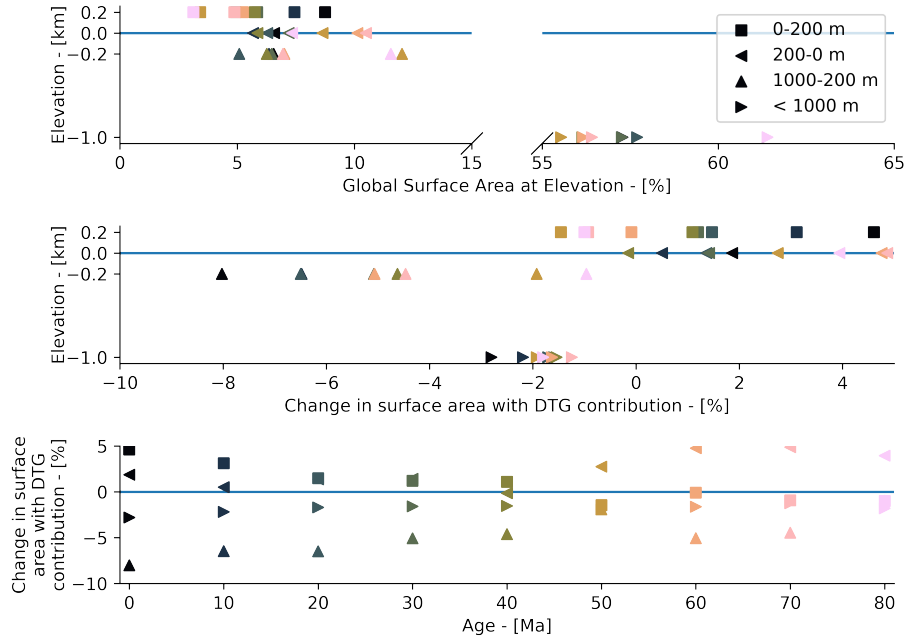


Figure B.2: The effect of dynamic topography on low topography and bathymetry throughout the last 80 Myr, using the bins and axes described in Figure 3.5. **Top)** Topography/bathymetry reconstructions provide the global surface area within each bin. **Middle)** Change in global surface area within each bin due to the contribution from the M1 dynamic topography model of (24). **Bottom)** The bottom panel plots the middle-panel results as a function of age.

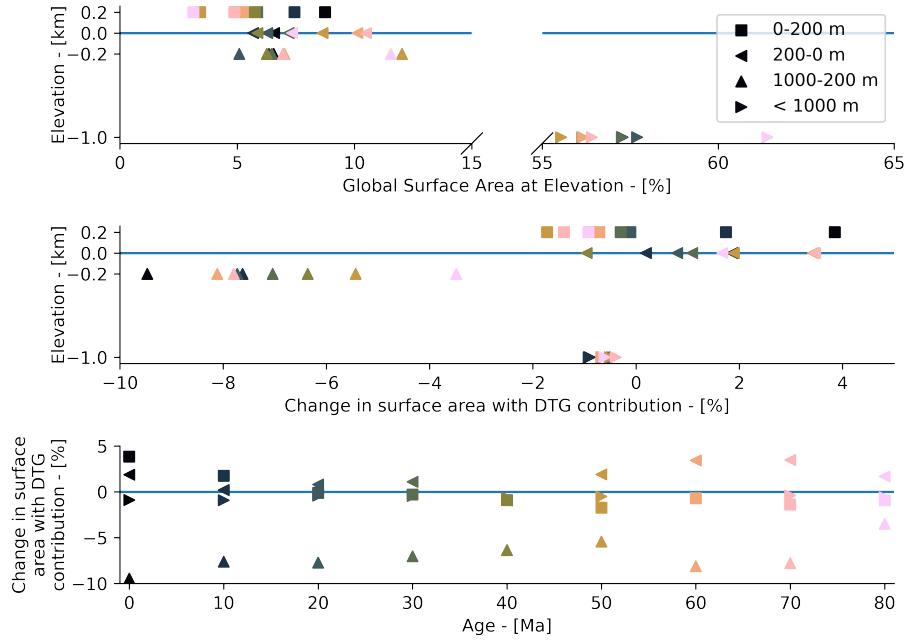


Figure B.3: The effect of dynamic topography on low topography and bathymetry throughout the last 80 Myr, using the bins and axes described in Figure 3.5. **Top)** Topography/bathymetry reconstructions provide the global surface area within each bin. **Middle)** Change in global surface area within each bin due to the contribution from the M7 dynamic topography model of (24). **Bottom)** The bottom panel plots the middle-panel results as a function of age.

Bibliography

- [1] M. P. Hain, K. A. Allen, S. Kirtland Turner. *Earth system carbon cycle dynamics through time*, pages 381–418. Elsevier, Oxford, 2025. DOI: <https://doi.org/https://doi.org/10.1016/B978-0-323-99762-1.00080-2>
- [2] S. Rackley. *Chapter 4 - The global carbon cycle*, pages 67–74. Elsevier, 2023. DOI: <https://doi.org/https://doi.org/10.1016/B978-0-12-819663-2.00002-2>
- [3] A. S. von der Heydt et al., Quantification and interpretation of the climate variability record. *Global and Planetary Change* **197**, 103399 (2021). DOI: <https://doi.org/https://doi.org/10.1016/j.gloplacha.2020.103399>
- [4] H. Pälike et al., A cenozoic record of the equatorial pacific carbonate compensation depth. *Nature* **488**(7413), 609–614 (2012). DOI: <https://doi.org/10.1038/nature11360>
- [5] M. L. Delaney, E. A. Boyle, Tertiary paleoceanic chemical variability: Unintended consequences of simple geochemical models. *Paleoceanography* **3**(2), 137–156 (1988). DOI: <https://doi.org/https://doi.org/10.1029/PA003i002p00137>
- [6] S. K. Boss, B. H. Wilkinson, Planktogenic/eustatic control on cratonic/oceanic carbonate accumulation. *The Journal of Geology* **99**(4), 497–513 (1991). DOI: <https://doi.org/10.1086/629513>

- [7] T. Lyu et al., Dissecting the puzzle of tectonic lid regimes in terrestrial planets. *Nature Communications* **16**(1), 10037 (2025). DOI: <https://doi.org/10.1038/s41467-025-65943-1>
- [8] J. Korenaga, Initiation and evolution of plate tectonics on earth: Theories and observations. *Annual Review of Earth and Planetary Sciences* **41**(Volume 41, 2013), 117–151 (2013). DOI: <https://doi.org/https://doi.org/10.1146/annurev-earth-050212-124208>
- [9] C. Amante, B. W. Eakins. Etopo1 1 arc-minute global relief model: Procedures, data sources and analysis, 2009. DOI: <https://doi.org/10.7289/V5C8276M>
- [10] M. Lyle et al., Pacific ocean and cenozoic evolution of climate. *Reviews of Geophysics* **46**(2) (2008). DOI: <https://doi.org/https://doi.org/10.1029/2005RG000190>
- [11] M. Bogumil, T. Mittal, C. Lithgow-Bertelloni, The effects of bathymetry on the long-term carbon cycle and ccd. *Proceedings of the National Academy of Sciences* **121**(21), e2400232121 (2024). DOI: <https://doi.org/doi:10.1073/pnas.2400232121>
- [12] R. E. Zeebe, Loscar: Long-term ocean-atmosphere-sediment carbon cycle reservoir model v2.0.4. *Geosci. Model Dev.* **5**(1), 149–166 (2012). GMD. DOI: <https://doi.org/10.5194/gmd-5-149-2012>
- [13] H. J. Cheong, T. Mittal, C. J. Sprain, I. M. Fendley, Large igneous province sulfur emissions have long-term (>1000 years) effects on the ocean carbon cycle. *Geochemistry, Geophysics, Geosystems* **26**(3), e2024GC011893 (2025). DOI: <https://doi.org/https://doi.org/10.1029/2024GC011893>
- [14] IHB. Limits of the oceans and seas. Report, 1953/09/01 1953.
- [15] A. R. Longhurst. *Ecological geography of the sea*. Elsevier, 2010.

- [16] F. Dalvand, A. Dutkiewicz, N. M. Wright, B. R. Mather, D. Müller, Regional carbonate compensation depth variability in the pacific ocean since the oligocene. *Frontiers in Earth Science* **Volume 13 - 2025** (2025). DOI: <https://doi.org/10.3389/feart.2025.1605906>
- [17] A. Dutkiewicz, R. D. Müller, The carbonate compensation depth in the south atlantic ocean since the late cretaceous. *Geology* **49**(7), 873–878 (2021). DOI: <https://doi.org/10.1130/g48404.1>
- [18] Copernicus. Global ocean physics reanalysis., 17 March 2025 2025. DOI: <https://doi.org/https://doi.org/10.48670/moi-00021>
- [19] H. K. Voris, Maps of pleistocene sea levels in southeast asia: shorelines, river systems and time durations. *Journal of Biogeography* **27**(5), 1153–1167 (2000). DOI: <https://doi.org/https://doi.org/10.1046/j.1365-2699.2000.00489.x>
- [20] P. D. Clift, Controls on the erosion of cenozoic asia and the flux of clastic sediment to the ocean. *Earth and Planetary Science Letters* **241**(3), 571–580 (2006). DOI: <https://doi.org/https://doi.org/10.1016/j.epsl.2005.11.028>
- [21] P. Clift, G. Layne, J. Blusztajn. *The erosional record of Tibetan uplift in the East Asian marginal seas*, volume 149, page 255–282. 2004.
- [22] P. Maffre et al., Geoclim7, an earth system model for multi-million-year evolution of the geochemical cycles and climate. *Geosci. Model Dev.* **18**(18), 6367–6413 (2025). GMD. DOI: <https://doi.org/10.5194/gmd-18-6367-2025>
- [23] U. G. Wortmann et al., The earth science box modeling toolkit (esbmtk 0.14.0.11): a python library for research and teaching. *Geosci. Model Dev.* **18**(4), 1155–1167 (2025). GMD. DOI: <https://doi.org/10.5194/gmd-18-1155-2025>

- [24] R. D. Müller, R. Hassan, M. Gurnis, N. Flament, S. E. Williams, Dynamic topography of passive continental margins and their hinterlands since the cretaceous. *Gondwana Research* **53**, 225–251 (2018). DOI: <https://doi.org/https://doi.org/10.1016/j.gr.2017.04.028>
- [25] M. Bogumil, C. Lithgow-Bertelloni, T. Mittal. Exocycle v1.0.0: A generalized framework for spherical community detection and its application to defining global ocean basins from multi-field data, 2026. DOI: <https://doi.org/https://doi.org/10.5194/egusphere-2025-5581>
- [26] R. Wordsworth, R. Pierrehumbert, Abiotic oxygen-dominated atmospheres on terrestrial habitable zone planets. *The Astrophysical Journal Letters* **785**(2), L20 (2014). DOI: <https://doi.org/https://doi.org/10.1088/2041-8205/785/2/L20>
- [27] C. S. Cockell et al., Sustained and comparative habitability beyond earth. *Nature Astronomy*, 1–9 (2024). DOI: <https://doi.org/https://doi.org/10.1038/s41550-023-02158-8>
- [28] A. H. Triaud et al., Atmospheric carbon depletion as a tracer of water oceans and biomass on temperate terrestrial exoplanets. *Nature Astronomy*, 1–13 (2023). DOI: <https://doi.org/https://doi.org/10.1038/s41550-023-02157-9>
- [29] A. Chopra, C. H. Lineweaver, The case for a gaian bottleneck: The biology of habitability. *Astrobiology* **16**(1), 7–22 (2016). doi: 10.1089/ast.2015.1387. DOI: <https://doi.org/https://doi.org/10.1089/ast.2015.1387>
- [30] O. Sulpis et al., Current caco3 dissolution at the seafloor caused by anthropogenic co2. *Proceedings of the National Academy of Sciences* **115**(46), 11700–11705 (2018). DOI: <https://doi.org/10.1073/pnas.1804250115>
- [31] O. Sulpis, E. Jeansson, A. Dinuer, S. K. Lauvset, J. J. Middelburg, Calcium carbonate

- dissolution patterns in the ocean. *Nature Geoscience* **14**(6), 423–428 (2021). DOI: <https://doi.org/10.1038/s41561-021-00743-y>
- [32] T. H. Van Andel, Mesozoic/cenozoic calcite compensation depth and the global distribution of calcareous sediments. *Earth and Planetary Science Letters* **26**(2), 187–194 (1975). DOI: [https://doi.org/https://doi.org/10.1016/0012-821X\(75\)90086-2](https://doi.org/https://doi.org/10.1016/0012-821X(75)90086-2)
- [33] A. Dutkiewicz, R. D. Müller, The history of cenozoic carbonate flux in the atlantic ocean constrained by multiple regional carbonate compensation depth reconstructions. *Geochemistry, Geophysics, Geosystems* **23**(11), e2022GC010667 (2022). DOI: <https://doi.org/https://doi.org/10.1029/2022GC010667>
- [34] F. Dalvand, A. Dutkiewicz, N. M. Wright, R. D. Müller, Indian ocean carbonate compensation depth since the late oligocene. *Geo-Marine Letters* **45**(4), 38 (2025). DOI: <https://doi.org/10.1007/s00367-025-00825-5>
- [35] R. S. Dietz, Continent and ocean basin evolution by spreading of the sea floor. *Nature* **190**(4779), 854–857 (1961). DOI: <https://doi.org/10.1038/190854a0>
- [36] J. X. Mitrovica, C. Beaumont, G. T. Jarvis, Tilting of continental interiors by the dynamical effects of subduction. *Tectonics* **8**(5), 1079–1094 (1989). DOI: <https://doi.org/https://doi.org/10.1029/TC008i005p01079>
- [37] R. D. Müller et al., A tectonic-rules-based mantle reference frame since 1 billion years ago – implications for supercontinent cycles and plate–mantle system evolution. *Solid Earth* **13**(7), 1127–1159 (2022). SE. DOI: <https://doi.org/10.5194/se-13-1127-2022>
- [38] R. D. Müller et al., A global plate model including lithospheric deformation along major rifts and orogens since the triassic. *Tectonics* **38**(6), 1884–1907 (2019). DOI: <https://doi.org/10.1029/2018TC005462>

- [39] T. H. Torsvik et al., Pacific-panthalassic reconstructions: Overview, errata and the way forward. *Geochemistry, Geophysics, Geosystems* **20**(7), 3659–3689 (2019). DOI: <https://doi.org/https://doi.org/10.1029/2019GC008402>
- [40] R. G. Hilton, A. J. West, Mountains, erosion and the carbon cycle. *Nature Reviews Earth Environment* **1**(6), 284–299 (2020). DOI: <https://doi.org/https://doi.org/10.1038/s43017-020-0058-6>
- [41] J. Krissansen-Totton, G. N. Arney, D. C. Catling, Constraining the climate and ocean pH of the early earth with a geological carbon cycle model. *Proceedings of the National Academy of Sciences* **115**(16), 4105–4110 (2018). DOI: <https://doi.org/https://doi.org/10.1073/pnas.1721296115>
- [42] A. Ridgwell, R. E. Zeebe, The role of the global carbonate cycle in the regulation and evolution of the earth system. *Earth and Planetary Science Letters* **234**(3-4), 299–315 (2005). DOI: <https://doi.org/https://doi.org/10.1016/j.epsl.2005.03.006>
- [43] R. D. Müller et al., Evolution of earth’s tectonic carbon conveyor belt. *Nature* **605**(7911), 629–639 (2022). DOI: <https://doi.org/10.1038/s41586-022-04420-x>
- [44] D. E. Penman, J. K. C. Rugenstein, D. E. Ibarra, M. J. Winnick, Silicate weathering as a feedback and forcing in earth’s climate and carbon cycle. *Earth-Science Reviews* **209**, 103298 (2020). DOI: <https://doi.org/https://doi.org/10.1016/j.earscirev.2020.103298>
- [45] J. J. Middelburg, K. Soetaert, M. Hagens, Ocean alkalinity, buffering and biogeochemical processes. *Reviews of Geophysics* **58**(3), e2019RG000681 (2020). DOI: <https://doi.org/https://doi.org/10.1029/2019RG000681>
- [46] T. DeVries, The ocean carbon cycle. *Annual Review of Environment and Resources* **47**, 317–341 (2022). DOI: <https://doi.org/https://doi.org/10.1146/annurev-environ-120920-111307>

- [47] T. T. Isson, N. J. Planavsky, Reverse weathering as a long-term stabilizer of marine ph and planetary climate. *Nature* **560**(7719), 471–475 (2018). DOI: <https://doi.org/10.1038/s41586-018-0408-4>
- [48] C.-T. A. Lee, H. Jiang, R. Dasgupta, M. Torres. *A framework for understanding whole-Earth carbon cycling*, pages 313–357. Cambridge University Press, 2019. DOI: <https://doi.org/https://doi.org/10.1017/9781108677950>
- [49] J. Zhang et al., A critical review of sr-k-mg-li-os isotopes for the study of continental weathering. *Journal of Asian Earth Sciences* **301**, 107004 (2026). DOI: <https://doi.org/https://doi.org/10.1016/j.jseaes.2026.107004>
- [50] S. J. Batenburg et al., Major intensification of atlantic overturning circulation at the onset of paleogene greenhouse warmth. *Nature Communications* **9**(1), 4954 (2018). DOI: <https://doi.org/10.1038/s41467-018-07457-7>
- [51] K. Tachikawa, W. Rapuc, Q. Dubois-Dauphin, A. Guihou, C. Skonieczny, Reconstruction of ocean circulation based on neodymium isotopic composition potential limitations and application to the mid-pleistocene transition. *Oceanography* **33**(2), 80–87 (2020). URL: <https://www.jstor.org/stable/26937744>
- [52] S. N. Raymond, T. Quinn, J. I. Lunine, High-resolution simulations of the final assembly of earth-like planets i. terrestrial accretion and dynamics. *Icarus* **183**(2), 265–282 (2006). DOI: <https://doi.org/https://doi.org/10.1016/j.icarus.2006.03.011>
- [53] T. Ruedas, Radioactive heat production of six geologically important nuclides. *Geochemistry, Geophysics, Geosystems* **18**(9), 3530–3541 (2017). DOI: <https://doi.org/https://doi.org/10.1002/2017GC006997>
- [54] R. Stern, T. Gerya, P. Tackley, A tectonic manifesto. *Perspectives of Earth and Space Scientists* **4**(1), e2023CN000214 (2023). DOI: <https://doi.org/https://doi.org/10.1029/2023CN000214>

- [55] R. J. Stern, Evidence from ophiolites, blueschists, and ultrahigh-pressure metamorphic terranes that the modern episode of subduction tectonics began in neoproterozoic time. *Geology* **33**(7), 557–560 (2005). DOI: <https://doi.org/10.1130/g21365.1>
- [56] P. A. Cawood et al., Geological archive of the onset of plate tectonics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **376**(2132) (2018). DOI: <https://doi.org/10.1098/rsta.2017.0405>
- [57] M. Hopkins, T. M. Harrison, C. E. Manning, Low heat flow inferred from >4gyr zircons suggests hadean plate boundary interactions. *Nature* **456**(7221), 493–496 (2008). DOI: <https://doi.org/10.1038/nature07465>
- [58] B. Dhuime, A. Wuestefeld, C. J. Hawkesworth, Emergence of modern continental crust about 3 billion years ago. *Nature Geoscience* **8**(7), 552–555 (2015). DOI: <https://doi.org/10.1038/ngeo2466>
- [59] N. Flament, N. Coltice, P. F. Rey, A case for late-archaeon continental emergence from thermal evolution models and hypsometry. *Earth and Planetary Science Letters* **275**(3), 326–336 (2008). DOI: <https://doi.org/https://doi.org/10.1016/j.epsl.2008.08.029>
- [60] K. Chotalia, N. Cagney, C. Lithgow-Bertelloni, J. Brodholt, The coupled effects of mantle mixing and a water-dependent viscosity on the surface ocean. *Earth and Planetary Science Letters* **530**, 115881 (2020). DOI: <https://doi.org/https://doi.org/10.1016/j.epsl.2019.115881>
- [61] D. L. Turcotte, G. Schubert. *Geodynamics*. Cambridge university press, 2002.
- [62] B. Parsons, D. McKenzie, Mantle convection and the thermal structure of the plates. *Journal of Geophysical Research: Solid Earth* **83**(B9), 4485–4496 (1978). DOI: <https://doi.org/https://doi.org/10.1029/JB083iB09p04485>

- [63] T. Korenaga, J. Korenaga, H. Kawakatsu, M. Yamano, A new reference model for the evolution of oceanic lithosphere in a cooling earth. *Journal of Geophysical Research: Solid Earth* **126** (2021). DOI: <https://doi.org/10.1029/2020JB021528>
- [64] A. G. Crosby, D. McKenzie, An analysis of young ocean depth, gravity and global residual topography. *Geophysical Journal International* **178**(3), 1198–1219 (2009). DOI: <https://doi.org/10.1111/j.1365-246X.2009.04224.x>
- [65] B. Parsons, J. G. Sclater, An analysis of the variation of ocean floor bathymetry and heat flow with age. *Journal of Geophysical Research (1896-1977)* **82**(5), 803–827 (1977). DOI: <https://doi.org/https://doi.org/10.1029/JB082i005p00803>
- [66] C. A. Stein, S. Stein, A model for the global variation in oceanic depth and heat flow with lithospheric age. *Nature* **359**(6391), 123–129 (1992). DOI: <https://doi.org/10.1038/359123a0>
- [67] E. O. Straume et al., Globsed: Updated total sediment thickness in the world's oceans. *Geochemistry, Geophysics, Geosystems* **20**(4), 1756–1772 (2019). DOI: <https://doi.org/https://doi.org/10.1029/2018GC008115>
- [68] A. Goswami, P. L. Olson, L. A. Hinnov, A. Gnanadesikan, Oesbathy version 1.0: a method for reconstructing ocean bathymetry with generalized continental shelf-slope-rise structures. *Geosci. Model Dev.* **8**(9), 2735–2748 (2015). GMD. DOI: <https://doi.org/10.5194/gmd-8-2735-2015>
- [69] T. Salles, X. Ding, G. Brocard, pybadlands: A framework to simulate sediment transport, landscape dynamics and basin stratigraphic evolution through space and time. *PloS one* **13**(4), e0195557 (2018). DOI: <https://doi.org/10.1371/journal.pone.0195557>
- [70] X. Ding, T. Salles, N. Flament, P. Rey, Quantitative stratigraphic analysis in a source-to-sink numerical framework. *Geoscientific Model Development* **12**(6), 2571–2585 (2019). DOI: <https://doi.org/10.5194/gmd-12-2571-2019>

- [71] T. Salles, L. Husson, M. Lorcery, B. Hadler Boggiani, Landscape dynamics and the phanerozoic diversification of the biosphere. *Nature* **624**(7990), 115–121 (2023). DOI: <https://doi.org/https://doi.org/10.1038/s41586-023-06777-z>
- [72] C. P. Conrad, The solid earth’s influence on sea level. *GSA Bulletin* **125**(7-8), 1027–1052 (2013). DOI: <https://doi.org/10.1130/b30764.1>
- [73] B. Sames, M. Wagreich, C. P. Conrad, S. Iqbal, Aquifer-eustasy as the main driver of short-term sea-level fluctuations during cretaceous hothouse climate phases. *Geological Society, London, Special Publications* **498**(1), 9–38 (2020). doi: 10.1144/SP498-2019-105. DOI: <https://doi.org/10.1144/SP498-2019-105>
- [74] H. C. Urey, On the early chemical history of the earth and the origin of life. *Proceedings of the National Academy of Sciences* **38**(4), 351–363 (1952). DOI: <https://doi.org/10.1073/pnas.38.4.351>
- [75] F. J. Millero, Thermodynamics of the carbon dioxide system in the oceans. *Geochimica et Cosmochimica Acta* **59**(4), 661–677 (1995). DOI: [https://doi.org/10.1016/0016-7037\(94\)00354-0](https://doi.org/10.1016/0016-7037(94)00354-0)
- [76] R. D. Müller, M. Sdrolias, C. Gaina, W. R. Roest, Age, spreading rates, and spreading asymmetry of the world’s ocean crust. *Geochemistry, Geophysics, Geosystems* **9**(4) (2008). DOI: <https://doi.org/https://doi.org/10.1029/2007GC001743>
- [77] D. C. Engebretson, A. Cox, R. G. Gordon. *Relative motions between oceanic and continental plates in the Pacific Basin / David C. Engebretson, Allan Cox, Richard G. Gordon*. Special paper / Geological Society of America ; 206. Geological Society of America, Boulder, Colo, 1985. Includes bibliographical references (pages 58-59).
- [78] J.-B. Ladant et al., Paleogeographic controls on the evolution of late cretaceous ocean circulation. *Climate of the Past* **16**(3), 973–1006 (2020). DOI: <https://doi.org/10.5194/cp-16-973-2020>

- [79] E. O. Straume, C. Gaina, S. Medvedev, K. H. Nisancioglu, Global cenozoic paleobathymetry with a focus on the northern hemisphere oceanic gateways. *Gondwana Research* **86**, 126–143 (2020). DOI: <https://doi.org/10.1016/j.gr.2020.05.011>
- [80] N. Herold et al., A suite of early eocene (55 ma) climate model boundary conditions. *Geosci. Model Dev.* **7**(5), 2077–2090 (2014). GMD. DOI: <https://doi.org/10.5194/gmd-7-2077-2014>
- [81] Y. Donnadieu et al., A geoclim simulation of climatic and biogeochemical consequences of pangea breakup. *Geochemistry, Geophysics, Geosystems* **7**(11) (2006). DOI: <https://doi.org/https://doi.org/10.1029/2006GC001278>
- [82] W. H. Berger, Increase of carbon dioxide in the atmosphere during deglaciation: the coral reef hypothesis. *Naturwissenschaften* **69**(2), 87–88 (1982). DOI: <https://doi.org/10.1007/BF00441228>
- [83] E. C. Geyman et al., The origin of carbonate mud and implications for global climate. *Proceedings of the National Academy of Sciences* **119**(43), e2210617119 (2022). DOI: <https://doi.org/doi:10.1073/pnas.2210617119>
- [84] N. Komar, R. E. Zeebe, Reconciling atmospheric co₂, weathering, and calcite compensation depth across the cenozoic. *Science Advances* **7**(4), eabd4876 (2021). DOI: <https://doi.org/doi:10.1126/sciadv.abd4876>
- [85] A. Dutkiewicz, R. D. Müller, J. Cannon, S. Vaughan, S. Zahirovic, Sequestration and subduction of deep-sea carbonate in the global ocean since the early cretaceous. *Geology* **47**(1), 91–94 (2018). DOI: <https://doi.org/10.1130/g45424.1>
- [86] A. Ridgwell, Interpreting transient carbonate compensation depth changes by marine sediment core modeling. *Paleoceanography* **22**(4) (2007). DOI: <https://doi.org/10.1029/2006PA001372>

- [87] D. M. Sigman, E. A. Boyle, Glacial/interglacial variations in atmospheric carbon dioxide. *Nature* **407**(6806), 859–869 (2000). DOI: <https://doi.org/10.1038/35038000>
- [88] C. T. Hayes et al., Global ocean sediment composition and burial flux in the deep sea. *Global Biogeochemical Cycles* **35**(4), e2020GB006769 (2021). DOI: <https://doi.org/https://doi.org/10.1029/2020GB006769>
- [89] D. A. Siegel, T. DeVries, I. Cetinić, K. M. Bisson, Quantifying the ocean’s biological pump and its carbon cycle impacts on global scales. *Annual Review of Marine Science* **15**(1), 329–356 (2023). DOI: <https://doi.org/10.1146/annurev-marine-040722-115226>
- [90] R. van der Ploeg, B. P. Boudreau, J. J. Middelburg, A. Sluijs, Cenozoic carbonate burial along continental margins. *Geology* **47**(11), 1025–1028 (2019). DOI: <https://doi.org/10.1130/g46418.1>
- [91] J. W. Morse, F. T. Mackenzie. *Geochemistry of sedimentary carbonates* / John W. Morse, Fred T. Mackenzie, volume 48 of *Developments in sedimentology*. Elsevier, Amsterdam, 1990. Includes bibliographical references (pages 609-679) and index.
- [92] D. E. Archer, An atlas of the distribution of calcium carbonate in sediments of the deep sea. *Global Biogeochemical Cycles* **10**(1), 159–174 (1996). DOI: <https://doi.org/https://doi.org/10.1029/95GB03016>
- [93] M. R. Raven et al., Biomass storage in anoxic marine basins: Initial estimates of geochemical impacts and co2 sequestration capacity. *AGU Advances* **5**(1), e2023AV000950 (2024). DOI: <https://doi.org/https://doi.org/10.1029/2023AV000950>
- [94] R. D. Müller et al., Ocean basin evolution and global-scale plate reorganization events since pangea breakup. *Annual Review of Earth and Planetary Sciences* **44**(1), 107–138 (2016). DOI: <https://doi.org/10.1146/annurev-earth-060115-012211>

- [95] S. Williams, N. M. Wright, J. Cannon, N. Flament, R. D. Müller, Reconstructing seafloor age distributions in lost ocean basins. *Geoscience Frontiers* **12**(2), 769–780 (2021). DOI: <https://doi.org/https://doi.org/10.1016/j.gsf.2020.06.004>
- [96] A. Goswami, L. Hinnov, A. Gnanadesikan, T. Young, Realistic paleobathymetry of the cenomanian–turonian (94 ma) boundary global ocean. *Geosciences* **8**(1), 21 (2018). URL: <https://www.mdpi.com/2076-3263/8/1/21>
- [97] A. M. Forte, D. B. Rowley, Earth’s isostatic and dynamic topography—a critical perspective. *Geochemistry, Geophysics, Geosystems* **23**(9), e2021GC009740 (2022). DOI: <https://doi.org/https://doi.org/10.1029/2021GC009740>
- [98] A. Young et al., Long-term phanerozoic sea level change from solid earth processes. *Earth and Planetary Science Letters* **584**, 117451 (2022). DOI: <https://doi.org/https://doi.org/10.1016/j.epsl.2022.117451>
- [99] K. L. Polzin, J. M. Toole, J. R. Ledwell, R. W. Schmitt, Spatial variability of turbulent mixing in the abyssal ocean. *Science* **276**(5309), 93–96 (1997). DOI: <https://doi.org/doi:10.1126/science.276.5309.93>
- [100] S. Rigby, C. V. Milsom, Origins, evolution, and diversification of zooplankton. *Annual Review of Ecology and Systematics* **31**(1), 293–313 (2000). DOI: <https://doi.org/10.1146/annurev.ecolsys.31.1.293>
- [101] R. E. Martin, Cyclic and secular variation in microfossil biomineralization: clues to the biogeochemical evolution of phanerozoic oceans. *Global and Planetary Change* **11**(1), 1–23 (1995). DOI: [https://doi.org/https://doi.org/10.1016/0921-8181\(94\)00011-2](https://doi.org/https://doi.org/10.1016/0921-8181(94)00011-2)
- [102] R. D. Müller, M. Seton. *Paleophysiography of Ocean Basins*, pages 1–15. Springer Netherlands, Dordrecht, 2014. DOI: https://doi.org/10.1007/978-94-007-6644-0_84-1

- [103] E. Cañón-Tapia, Seamount chains and hotspot tracks: Superficially similar, deeply different. *Geoscience Frontiers* **14**(6), 101659 (2023). DOI: <https://doi.org/https://doi.org/10.1016/j.gsf.2023.101659>
- [104] O. R. Lehmer, D. C. Catling, J. Krissansen-Totton, Carbonate-silicate cycle predictions of earth-like planetary climates and testing the habitable zone concept. *Nature Communications* **11**(1), 6153 (2020). DOI: <https://doi.org/10.1038/s41467-020-19896-2>
- [105] M. J. Hoggard, J. Winterbourne, K. Czarnota, N. White, Oceanic residual depth measurements, the plate cooling model, and global dynamic topography. *Journal of Geophysical Research: Solid Earth* **122**(3), 2328–2372 (2017). DOI: <https://doi.org/https://doi.org/10.1002/2016JB013457>
- [106] C. R. Scotese, N. M. Wright. Paleomap paleodigital elevation models (paleodems) for the phanerozoic, 2018. DOI: <https://doi.org/10.5281/zenodo.5460860>
- [107] B. U. Haq, J. A. N. Hardenbol, P. R. Vail, Chronology of fluctuating sea levels since the triassic. *Science* **235**(4793), 1156–1167 (1987). ObjectType-Article-1. DOI: <https://doi.org/10.1126/science.235.4793.1156>
- [108] T. H. Heimdal, M. T. Jones, H. H. Svensen, Thermogenic carbon release from the central atlantic magmatic province caused major end-triassic carbon cycle perturbations. *Proceedings of the National Academy of Sciences* **117**(22), 11968–11974 (2020). DOI: <https://doi.org/https://doi.org/10.1073/pnas.2000095117>
- [109] R. E. Zeebe, A. Ridgwell, J. C. Zachos, Anthropogenic carbon release rate unprecedented during the past 66 million years. *Nature Geoscience* **9**(4), 325–329 (2016). DOI: <https://doi.org/10.1038/ngeo2681>
- [110] A. B. Robert, Geocarbsulf: A combined model for phanerozoic atmospheric o₂ and co₂. *Geochimica et Cosmochimica Acta* **70**(23), 5653–5664 (2006). A Special Issue

Dedicated to Robert A. Berner. DOI: <https://doi.org/https://doi.org/10.1016/j.gca.2005.11.032>

- [111] T. M. Lenton, S. J. Daines, B. J. W. Mills, Copse reloaded: An improved model of biogeochemical cycling over phanerozoic time. *Earth-Science Reviews* **178**, 1–28 (2018). DOI: <https://doi.org/https://doi.org/10.1016/j.earscirev.2017.12.004>
- [112] G. Shaffer et al., Implementation of methane cycling for deep-time global warming simulations with the dcess earth system model (version 1.2). *Geosci. Model Dev.* **10**(11), 4081–4103 (2017). GMD. DOI: <https://doi.org/10.5194/gmd-10-4081-2017>
- [113] M. Bogumil, T. Mittal, C. Lithgow-Bertelloni. The effects of bathymetry on the ccd: Implications for the long-term carbon cycle , 2024. DOI: <https://doi.org/10.5281/zenodo.10460278>
- [114] Q. Jiang, F. Jourdan, H. K. H. Olierook, R. E. Merle, An appraisal of the ages of phanerozoic large igneous provinces. *Earth-Science Reviews* **237**, 104314 (2023). DOI: <https://doi.org/https://doi.org/10.1016/j.earscirev.2023.104314>
- [115] Y. Zhou et al., The role of lips in phanerozoic mass extinctions: An hg perspective. *Earth-Science Reviews* **249**, 104667 (2024). DOI: <https://doi.org/https://doi.org/10.1016/j.earscirev.2023.104667>
- [116] D. E. Penman, J. C. Zachos, New constraints on massive carbon release and recovery processes during the paleocene-eocene thermal maximum. *Environmental Research Letters* **13**(10), 105008 (2018). DOI: <https://doi.org/https://doi.org/10.1088/1748-9326/aae285>
- [117] D. Turcotte, E. Oxburgh, Finite amplitude convective cells and continental drift. *Journal of Fluid Mechanics* **28**(1), 29–42 (1967). DOI: <https://doi.org/10.1017/S0022112067001880>

- [118] D. Forsyth, S. Uyeda, On the relative importance of the driving forces of plate motion*. *Geophysical Journal International* **43**(1), 163–200 (1975). DOI: <https://doi.org/10.1111/j.1365-246X.1975.tb00631.x>
- [119] C. Lithgow-Bertelloni, M. A. Richards, The dynamics of cenozoic and mesozoic plate motions. *Reviews of Geophysics* **36**(1), 27–78 (1998). DOI: <https://doi.org/10.1029/97RG02282>
- [120] C. Chang, L. Liu, Investigating the formation of the cretaceous western interior seaway using landscape evolution simulations. *GSA Bulletin* **133**(1-2), 347–361 (2020). DOI: <https://doi.org/10.1130/b35653.1>
- [121] Z. Li, J. Aschoff, Shoreline evolution in the late cretaceous north american cordilleran foreland basin: An exemplar of the combined influence of tectonics, sea level, and sediment supply through time. *Earth-Science Reviews* **226**, 103947 (2022). DOI: <https://doi.org/https://doi.org/10.1016/j.earscirev.2022.103947>
- [122] L. Liu, S. Spasojević, M. Gurnis, Reconstructing farallon plate subduction beneath north america back to the late cretaceous. *Science* **322**(5903), 934–938 (2008). DOI: <https://doi.org/doi:10.1126/science.1162921>
- [123] L. Gao et al., The initial opening of the drake passage occurred during ca. 62-59 ma. *Geophysical Research Letters* **52**(7), e2024GL111455 (2025). DOI: <https://doi.org/https://doi.org/10.1029/2024GL111455>
- [124] Y. Ricard, M. Richards, C. Lithgow-Bertelloni, Y. Le Stunff, A geodynamic model of mantle density heterogeneity. *Journal of Geophysical Research: Solid Earth* **98**(B12), 21895–21909 (1993). DOI: <https://doi.org/10.1029/93JB02216>
- [125] E. J. Clennett et al., A quantitative tomotectonic plate reconstruction of western north america and the eastern pacific basin. *Geochemistry, Geophysics, Geosystems* **21**(8), e2020GC009117 (2020). DOI: <https://doi.org/10.1029/2020GC009117>

- [126] R. D. Müller et al., Gplates: building a virtual earth through deep time. *Geochemistry, Geophysics, Geosystems* **19**(7), 2243–2261 (2018). DOI: <https://doi.org/10.1029/2018GC007584>
- [127] B. H. Hager, R. J. O’Connell, Kinematic models of large-scale flow in the earth’s mantle. *Journal of Geophysical Research: Solid Earth* **84**(B3), 1031–1048 (1979). DOI: <https://doi.org/https://doi.org/10.1029/JB084iB03p01031>
- [128] M. A. Richards, B. H. Hager, Geoid anomalies in a dynamic earth. *Journal of Geophysical Research: Solid Earth* **89**(B7), 5987–6002 (1984). DOI: <https://doi.org/https://doi.org/10.1029/JB089iB07p05987>
- [129] M. J. Hoggard, N. White, D. Al-Attar, Global dynamic topography observations reveal limited influence of large-scale mantle flow. *Nature Geoscience* **9**(6), 456–463 (2016). DOI: <https://doi.org/10.1038/ngeo2709>
- [130] R. Moucha et al., Dynamic topography and long-term sea-level variations: There is no such thing as a stable continental platform. *Earth and Planetary Science Letters* **271**(1), 101–108 (2008). DOI: <https://doi.org/https://doi.org/10.1016/j.epsl.2008.03.056>
- [131] S. Spasojevic, M. Gurnis, Sea level and vertical motion of continents from dynamic earth models since the late cretaceous. *AAPG Bulletin* **96**(11), 2037–2064 (2012). DOI: <https://doi.org/10.1306/03261211121>
- [132] A. Yang, Long-term global sea-level change due to dynamic topography since 410 ma. *Deep Sea Research Part I: Oceanographic Research Papers* **191**, 103944 (2023). DOI: <https://doi.org/https://doi.org/10.1016/j.dsr.2022.103944>
- [133] P. G. Falkowski et al., The evolution of modern eukaryotic phytoplankton. *Science* **305**(5682), 354–360 (2004). DOI: <https://doi.org/doi:10.1126/science.1095964>

- [134] L. DiCaprio, M. Gurnis, R. D. Müller, Long-wavelength tilting of the australian continent since the late cretaceous. *Earth and Planetary Science Letters* **278**(3), 175–185 (2009). DOI: <https://doi.org/https://doi.org/10.1016/j.epsl.2008.11.030>
- [135] S. M. Campbell, R. Moucha, L. A. Derry, M. E. Raymo, Effects of dynamic topography on the cenozoic carbonate compensation depth. *Geochemistry, Geophysics, Geosystems* **19**(4), 1025–1034 (2018). DOI: <https://doi.org/https://doi.org/10.1002/2017GC007386>
- [136] P. Glišović, A. M. Forte, Cenozoic evolution of earth’s strongest geoid low illuminates mantle dynamics beneath antarctica. *Scientific Reports* **15**(1), 45749 (2025). DOI: <https://doi.org/10.1038/s41598-025-28606-1>
- [137] N. Flament, M. Gurnis, R. D. Müller, A review of observations and models of dynamic topography. *Lithosphere* **5**(2), 189–210 (2013). DOI: <https://doi.org/10.1130/1245.1>
- [138] K. Ozaki, D. B. Cole, C. T. Reinhard, E. Tajika, Canops-grb v1.0: a new earth system model for simulating the evolution of ocean–atmosphere chemistry over geologic timescales. *Geosci. Model Dev.* **15**(20), 7593–7639 (2022). GMD. DOI: <https://doi.org/10.5194/gmd-15-7593-2022>
- [139] E. Fernández Villanueva, G. Shaffer, Presentation, calibration and testing of the dcess ii earth system model of intermediate complexity (version 1.0). *Geosci. Model Dev.* **18**(7), 2161–2192 (2025). GMD. DOI: <https://doi.org/10.5194/gmd-18-2161-2025>
- [140] A. Ridgwell et al., Marine geochemical data assimilation in an efficient earth system model of global biogeochemical cycling. *Biogeosciences* **4**(1), 87–104 (2007). BG. DOI: <https://doi.org/10.5194/bg-4-87-2007>
- [141] B. J. Foley, P. E. Driscoll, Whole planet coupling between climate, mantle, and core: Im-

- plications for rocky planet evolution. *Geochemistry, Geophysics, Geosystems* **17**(5), 1885–1914 (2016). DOI: <https://doi.org/https://doi.org/10.1002/2015GC006210>
- [142] S. Olson, M. F. Jansen, D. S. Abbot, I. Halevy, C. Goldblatt, The effect of ocean salinity on climate and its implications for earth’s habitability. *Geophysical research letters* **49**(10), e2021GL095748 (2022). DOI: <https://doi.org/https://doi.org/10.1029/2021GL095748>
- [143] Z. Li, Y. G. Zhang, M. Torres, B. J. W. Mills, Neogene burial of organic carbon in the global ocean. *Nature* **613**(7942), 90–95 (2023). DOI: <https://doi.org/10.1038/s41586-022-05413-6>
- [144] K. Wallmann et al., Periodic changes in the cretaceous ocean and climate caused by marine redox see-saw. *Nature Geoscience* **12**(6), 456–461 (2019). DOI: <https://doi.org/10.1038/s41561-019-0359-x>
- [145] P. T. Harris, M. Macmillan-Lawler, J. Rupp, E. K. Baker, Geomorphology of the oceans. *Marine Geology* **352**, 4–24 (2014). DOI: <https://doi.org/https://doi.org/10.1016/j.margeo.2014.01.011>
- [146] G. Froyland, R. M. Stuart, E. van Sebille, How well-connected is the surface of the global ocean? *Chaos: An Interdisciplinary Journal of Nonlinear Science* **24**(3) (2014). DOI: <https://doi.org/10.1063/1.4892530>
- [147] J. Xie, X. Liu, P. Bai, C. Liu, Rapid watershed delineation using an automatic outlet relocation algorithm. *Water Resources Research* **58**(3), e2021WR031129 (2022). DOI: <https://doi.org/https://doi.org/10.1029/2021WR031129>
- [148] S. Haag, D. Schwartz, B. Shakibajahromi, M. Campagna, S. Ali, A fast algorithm to delineate watershed boundaries for simple geometries. *Environmental Modelling Software* **134**, 104842 (2020). DOI: <https://doi.org/https://doi.org/10.1016/j.envsoft.2020.104842>

- [149] V. Rossi, E. Ser-Giacomi, C. López, E. Hernández-García, Hydrodynamic provinces and oceanic connectivity from a transport network help designing marine reserves. *Geophysical Research Letters* **41**(8), 2883–2891 (2014). DOI: <https://doi.org/https://doi.org/10.1002/2014GL059540>
- [150] X. Cao et al., Earth’s tectonic and plate boundary evolution over 1.8 billion years. *Geoscience Frontiers* **15**(6), 101922 (2024). DOI: <https://doi.org/https://doi.org/10.1016/j.gsf.2024.101922>
- [151] F. Landais, F. Schmidt, S. Lovejoy, Topography of (exo)planets. *Monthly Notices of the Royal Astronomical Society* **484**(1), 787–793 (2019). DOI: <https://doi.org/https://doi.org/10.1093/mnras/sty3253>
- [152] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* **2008**(10), P10008 (2008). DOI: <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- [153] V. A. Traag, L. Waltman, N. J. van Eck, From louvain to leiden: guaranteeing well-connected communities. *Scientific Reports* **9**(1), 5233 (2019). DOI: <https://doi.org/10.1038/s41598-019-41695-z>
- [154] M. Girvan, M. E. J. Newman, Community structure in social and biological networks. *Proceedings of the National Academy of Sciences* **99**(12), 7821–7826 (2002). DOI: <https://doi.org/doi:10.1073/pnas.122653799>
- [155] M. Rosvall, C. T. Bergstrom, Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences* **105**(4), 1118–1123 (2008). DOI: <https://doi.org/doi:10.1073/pnas.0706851105>
- [156] A. Giuliani, A. Krishnan, J. P. Zbilut, M. Tomita, Proteins as networks: Usefulness

- of graph theory in protein science. *Current Protein and Peptide Science* **9**(1), 28–38 (2008). DOI: <https://doi.org/https://doi.org/10.2174/138920308783565705>
- [157] Y. Wu, D. Lian, Y. Xu, L. Wu, E. Chen, Graph convolutional networks with markov random field reasoning for social spammer detection. *Proceedings of the AAAI Conference on Artificial Intelligence* **34**(01), 1054–1061 (2020). DOI: <https://doi.org/10.1609/aaai.v34i01.5455>
- [158] M. Needham, A. E. Hodler. *Graph algorithms : practical examples in Apache Spark and Neo4j* / Mark Needham and Amy E. Hodler. O’Reilly, Beijing, first edition. edition, 2019.
- [159] J. Reichardt, S. Bornholdt, Statistical mechanics of community detection. *Physical Review E* **74**(1), 016110 (2006). PRE. DOI: <https://doi.org/10.1103/PhysRevE.74.016110>
- [160] E. A. Leicht, M. E. J. Newman, Community structure in directed networks. *Physical Review Letters* **100**(11), 118703 (2008). PRL. DOI: <https://doi.org/10.1103/PhysRevLett.100.118703>
- [161] S. Fortunato, M. Barthélemy, Resolution limit in community detection. *Proceedings of the National Academy of Sciences* **104**(1), 36–41 (2007). DOI: <https://doi.org/doi:10.1073/pnas.0605965104>
- [162] V. A. Traag, P. Van Dooren, Y. Nesterov, Narrow scope for resolution-limit-free community detection. *Physical Review E* **84**(1), 016114 (2011). PRE. DOI: <https://doi.org/10.1103/PhysRevE.84.016114>
- [163] L. C. Freeman, A set of measures of centrality based on betweenness. *Sociometry* **40**(1), 35–41 (1977). DOI: <https://doi.org/10.2307/3033543>

- [164] E. Tan, M. Gurnis, L. Armendariz, L. Strand, S. Kientz. *CitcomS User Manual, Version 3.3.0*. Computational Infrastructure for Geodynamics (CIG), Davis, CA, 2014. URL: <https://geodynamics.org/resources/1747/download/citcoms-3.3.0-manual.pdf>
- [165] P. Wessel et al., The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems* **20**(11), 5556–5564 (2019). DOI: <https://doi.org/https://doi.org/10.1029/2019GC008515>
- [166] I. Esri. Arcgis pro (version 3.x), 2024. URL: <https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview>
- [167] L.-l. Ma, C. Ma, H.-F. Zhang, B.-H. Wang, Identifying influential spreaders in complex networks based on gravity formula. *Physica A: Statistical Mechanics and its Applications* **451**, 205–212 (2016). DOI: <https://doi.org/https://doi.org/10.1016/j.physa.2015.12.162>
- [168] S. Zhu, J. Zhan, X. Li, Identifying influential nodes in complex networks using a gravity model based on the h-index method. *Scientific Reports* **13**(1), 16404 (2023). DOI: <https://doi.org/10.1038/s41598-023-43585-x>
- [169] M. Liu, T. Gerya, A. Rozel, The effect of brittle-ductile weakening on the formation of faulting patterns at mid-ocean ridges. *Tectonics* **44**(2), e2024TC008586 (2025). DOI: <https://doi.org/https://doi.org/10.1029/2024TC008586>
- [170] J. A. Goff, A global and regional stochastic analysis of near-ridge abyssal hill morphology. *Journal of Geophysical Research: Solid Earth* **96**(B13), 21713–21737 (1991). DOI: <https://doi.org/https://doi.org/10.1029/91JB02275>
- [171] M. D. Behn, J. M. Sinton, R. S. Detrick, Effect of the galápagos hotspot on seafloor volcanism along the galápagos spreading center (90.9–97.6°w). *Earth and Planetary*

- Science Letters* **217**(3), 331–347 (2004). DOI: [https://doi.org/10.1016/S0012-821X\(03\)00611-3](https://doi.org/10.1016/S0012-821X(03)00611-3)
- [172] A. E. Aksu, J. Hall, C. Yaltırak, Miocene–quaternary tectonic, kinematic and sedimentary evolution of the eastern mediterranean sea: A regional synthesis. *Earth-Science Reviews* **220**, 103719 (2021). DOI: <https://doi.org/10.1016/j.earscirev.2021.103719>
- [173] B. A. Van der Pluijm, S. Marshak. *Earth Structure: An Introduction to Structural Geology and Tectonics*. W.W. Norton, Incorporated, 2010. URL: <https://books.google.com/books?id=hzPUBwAACAAJ>
- [174] S. Liu, D. Nummedal, M. Gurnis, Dynamic versus flexural controls of late cretaceous western interior basin, usa. *Earth and Planetary Science Letters* **389**, 221–229 (2014). DOI: <https://doi.org/10.1016/j.epsl.2014.01.006>
- [175] J. Emile-Geay et al., A global multiproxy database for temperature reconstructions of the common era. *Scientific Data* **4**(1), 170088 (2017). DOI: <https://doi.org/10.1038/sdata.2017.88>
- [176] N. M. Papadomanolaki, N. A. Van Helmond, H. Pälke, A. Sluijs, C. P. Slomp, Quantifying volcanism and organic carbon burial across oceanic anoxic event 2. *Geology* **50**(4), 511–515 (2022). DOI: <https://doi.org/10.1130/G49649.1>
- [177] L. D. Talley. *Descriptive physical oceanography: an introduction*. Academic press, 2011.
- [178] S. L. Olson, M. Jansen, D. S. Abbot, Oceanographic considerations for exoplanet life detection. *The Astrophysical Journal* **895**(1), 19 (2020). DOI: <https://doi.org/10.3847/1538-4357/ab88c9>
- [179] Y. Zhang et al., Early eocene ocean meridional overturning circulation: The roles of atmospheric forcing and strait geometry. *Paleoceanography and Paleoclimatology*

- 37**(3), e2021PA004329 (2022). DOI: <https://doi.org/https://doi.org/10.1029/2021PA004329>
- [180] H. A. Dijkstra, W. Weijer, Stability of the global ocean circulation: Basic bifurcation diagrams. *Journal of Physical Oceanography* **35**(6), 933–948 (2005). DOI: <https://doi.org/https://doi.org/10.1175/JP02726.1>
- [181] V. Schoepf et al., Corals at the edge of environmental limits: A new conceptual framework to re-define marginal and extreme coral communities. *Science of The Total Environment* **884**, 163688 (2023). DOI: <https://doi.org/https://doi.org/10.1016/j.scitotenv.2023.163688>
- [182] M. Lyle, Neogene carbonate burial in the pacific ocean. *Paleoceanography* **18**(3) (2003). DOI: <https://doi.org/https://doi.org/10.1029/2002PA000777>
- [183] A. Mucci, The solubility of calcite and aragonite in seawater at various salinities, temperatures, and one atmosphere total pressure. *American Journal of Science* **283**(7), 780–799 (1983). DOI: <https://doi.org/10.2475/ajs.283.7.780>
- [184] P. J. Valdes, C. R. Scotese, D. J. Lunt, Deep ocean temperatures through time. *Clim. Past* **17**(4), 1483–1506 (2021). CP. DOI: <https://doi.org/10.5194/cp-17-1483-2021>
- [185] J. W. Shepherd, V. Paumard, T. Salles, S. Lang, A. D. George, Assessing the lateral and vertical variability of shelf-margin depositional systems and associated forcing mechanisms: A forward modeling approach. *GSA Bulletin* **137**(1-2), 203–219 (2024). DOI: <https://doi.org/10.1130/B37598.1>
- [186] T. Salles, C. Mallard, S. Zahirovic, gospl: global scalable paleo landscape evolution. *Journal of Open Source Software* **5**(56), 2804 (2020). DOI: <https://doi.org/10.21105/joss.02804>

- [187] T. Salles et al., A unified framework for modelling sediment fate from source to sink and its interactions with reef systems over geological times. *Scientific Reports* **8**(1), 5252 (2018). DOI: <https://doi.org/10.1038/s41598-018-23519-8>
- [188] J. A. Goff, Global prediction of abyssal hill root-mean-square heights from small-scale altimetric gravity variability. *Journal of Geophysical Research: Solid Earth* **115**(B12) (2010). DOI: <https://doi.org/https://doi.org/10.1029/2010JB007867>
- [189] W. W. Sager, What built shatsky rise, a mantle plume or ridge tectonics? *Special Paper of the Geological Society of America* **388**, 721–733 (2005). DOI: <https://doi.org/https://doi.org/10.1130/0-8137-2388-4.721>
- [190] R. W. Dalrymple. *A review of the morphology, physical processes and deposits of modern straits*, page 0. Geological Society of London, 2023. DOI: <https://doi.org/10.1144/sp523-2021-76>
- [191] J. A. Whitehead, A. Leetmaa, R. A. Knox, Rotating hydraulics of strait and sill flows†. *Geophysical Fluid Dynamics* **6**(2), 101–125 (1974). doi: 10.1080/03091927409365790. DOI: <https://doi.org/10.1080/03091927409365790>
- [192] S. M. Clifford, T. J. Parker, The evolution of the martian hydrosphere: Implications for the fate of a primordial ocean and the current state of the northern plains. *Icarus* **154**(1), 40–79 (2001). DOI: <https://doi.org/https://doi.org/10.1006/icar.2001.6671>
- [193] M. Turbet, F. Forget, The paradoxes of the late hesperian mars ocean. *Scientific Reports* **9**(1), 5717 (2019). DOI: <https://doi.org/10.1038/s41598-019-42030-2>
- [194] F. Schmidt et al., Circumpolar ocean stability on mars 3 gy ago. *Proceedings of the National Academy of Sciences* **119**(4), e2112930118 (2022). DOI: <https://doi.org/doi:10.1073/pnas.2112930118>

- [195] J. P. Grotzinger, N. P. James. Precambrian carbonates: Evolution of understanding, 2000. URL: <https://doi.org/10-2110/pec.00.67>
- [196] F. Landais, F. Schmidt, S. Lovejoy, Multifractal topography of several planetary bodies in the solar system. *Icarus* **319**, 14–20 (2019). DOI: <https://doi.org/https://doi.org/10.1016/j.icarus.2018.07.005>
- [197] J. Tian, P. J. Tackley, D. L. Lourenço, The tectonics and volcanism of venus: New modes facilitated by realistic crustal rheology and intrusive magmatism. *Icarus* **399**, 115539 (2023). DOI: <https://doi.org/https://doi.org/10.1016/j.icarus.2023.115539>
- [198] M. Bogumil, C. Lithgow-Bertelloni, T. Mittal. Exocycle: A generalized framework for spherical community detection and its application to defining global ocean basins from multi-field data - model and results, 2026. DOI: <https://doi.org/10.5281/zenodo.19671779>
- [199] M. Bogumil, C. Lithgow-Bertelloni, T. Mittal. Exocycle: A generalized framework for spherical community detection and its application to defining global ocean basins from multi-field data - input models, 2026. DOI: <https://doi.org/10.5281/zenodo.18664604>
- [200] W. P. Köppen, Die wärmezonen der erde, nach der dauer der heissen, gemässigten und kalten zeit und nach der wirkung der wärme auf die organische welt betrachtet: publisher not identified. *Meteorologische Zeitschrift* **1**, 215–226 (1884).
- [201] M. C. Peel, B. L. Finlayson, T. A. McMahon, Updated world map of the köppen-geiger climate classification. *Hydrol. Earth Syst. Sci.* **11**(5), 1633–1644 (2007). HESS. DOI: <https://doi.org/10.5194/hess-11-1633-2007>
- [202] H. E. Beck et al., Present and future köppen-geiger climate classification maps at 1-km resolution. *Scientific Data* **5**(1), 180214 (2018). DOI: <https://doi.org/10.1038/sdata.2018.214>

- [203] J. B. Kellogg, S. B. Jacobsen, R. J. O’Connell, Modeling the distribution of isotopic ratios in geochemical reservoirs. *Earth and Planetary Science Letters* **204**(1), 183–202 (2002). DOI: [https://doi.org/https://doi.org/10.1016/S0012-821X\(02\)00981-0](https://doi.org/https://doi.org/10.1016/S0012-821X(02)00981-0)
- [204] L. Li, A. R. A. Aitken, M. D. Lindsay, B. Kulesa, Sedimentary basins reduce stability of antarctic ice streams through groundwater feedbacks. *Nature Geoscience* **15**(8), 645–650 (2022). DOI: <https://doi.org/10.1038/s41561-022-00992-5>
- [205] Z. Liu et al., Global cooling during the eocene-oligocene climate transition. *Science* **323**(5918), 1187–1190 (2009). DOI: <https://doi.org/doi:10.1126/science.1166368>
- [206] A. J. Ward. *Constraining the large scale glaciation of Antarctica during the Eocene-Oligocene Transition using hydrogen isotopes of alkenones*. Thesis, 2016.
- [207] A. Grinsted, An estimate of global glacier volume. *The Cryosphere* **7**(1), 141–151 (2013). TC. DOI: <https://doi.org/10.5194/tc-7-141-2013>
- [208] B. De Boer, R. S. W. van de Wal, R. Bintanja, L. J. Lourens, E. Tuenter, Cenozoic global ice-volume and temperature simulations with 1-d ice-sheet models forced by benthic 18o records. *Annals of Glaciology* **51**(55), 23–33 (2010). DOI: <https://doi.org/10.3189/172756410791392736>
- [209] D. G. Van der Meer, D. J. Van Hinsbergen, W. Spakman, Atlas of the underworld: Slab remnants in the mantle, their sinking history, and a new outlook on lower mantle viscosity. *Tectonophysics* **723**, 309–448 (2018). DOI: <https://doi.org/10.1016/j.tecto.2017.10.004>
- [210] D. J. Weidner, Y. Wang, Chemical-and clapeyron-induced buoyancy at the 660 km discontinuity. *Journal of Geophysical Research: Solid Earth* **103**(B4), 7431–7441 (1998). DOI: <https://doi.org/10.1029/97JB03511>

- [211] L. Stixrude, C. Lithgow-Bertelloni, Thermodynamics of mantle minerals—i. physical properties. *Geophysical Journal International* **162**(2), 610–632 (2005). DOI: <https://doi.org/10.1111/j.1365-246X.2005.02642.x>
- [212] L. Stixrude, C. Lithgow-Bertelloni, Thermodynamics of mantle minerals-ii. phase equilibria. *Geophysical Journal International* **184**(3), 1180–1213 (2011). DOI: <https://doi.org/10.1111/j.1365-246X.2010.04890.x>
- [213] L. Stixrude, C. Lithgow-Bertelloni, Mineralogy and elasticity of the oceanic upper mantle: Origin of the low-velocity zone. *Journal of Geophysical Research: Solid Earth* **110**(B3) (2005). DOI: <https://doi.org/https://doi.org/10.1029/2004JB002965>
- [214] L. Stixrude, C. Lithgow-Bertelloni, Thermal expansivity, heat capacity and bulk modulus of the mantle. *Geophysical Journal International* **228**(2), 1119–1149 (2022). DOI: <https://doi.org/10.1093/gji/ggab394>