

Paleoceanography and Paleoclimatology*



RESEARCH ARTICLE

10.1029/2023PA004769

Reconstruction of Cenozoic $\delta^{11}\text{B}_{\text{sw}}$ Using a Gaussian Process

Ross Whiteford¹ , Timothy J. Heaton² , Michael J. Henehan³ , Eleni Anagnostou⁴ , Hana Jurikova¹, Gavin L. Foster⁵, and James W. B. Rae¹ 

Special Collection:

Advances and Best Practices in Boron-based paleo-CO₂ reconstruction

¹School of Earth and Environmental Sciences, University of St Andrews, St Andrews, UK, ²Department of Statistics, School of Mathematics, University of Leeds, Leeds, UK, ³School of Earth Sciences, University of Bristol, Bristol, UK, ⁴GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany, ⁵School of Ocean Earth Science, University of Southampton, Southampton, UK

Key Points:

- We reconstruct the temporal evolution of seawater isotope ratios of boron, strontium, lithium, and osmium over the last 65 million years
- The evolution of seawater boron isotope ratio shows similarity to the evolution of strontium, lithium and osmium isotope ratios
- Randomly drawn, smooth time series are provided for use in uncertainty propagation in calculation of palaeo pH

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

R. Whiteford,
rdmw1@st-andrews.ac.uk

Citation:

Whiteford, R., Heaton, T. J., Henehan, M. J., Anagnostou, E., Jurikova, H., Foster, G. L., & Rae, J. W. B. (2024). Reconstruction of Cenozoic $\delta^{11}\text{B}_{\text{sw}}$ using A Gaussian process. *Paleoceanography and Paleoclimatology*, 39, e2023PA004769. <https://doi.org/10.1029/2023PA004769>

Received 25 SEP 2023
Accepted 27 APR 2024

Author Contributions:

Conceptualization: Ross Whiteford, Timothy J. Heaton, James W. B. Rae
Data curation: Ross Whiteford, Michael J. Henehan, Eleni Anagnostou, Hana Jurikova, Gavin L. Foster
Formal analysis: Ross Whiteford, Timothy J. Heaton
Methodology: Ross Whiteford, Timothy J. Heaton, James W. B. Rae
Software: Ross Whiteford, Timothy J. Heaton
Supervision: James W. B. Rae

© 2024. The Author(s).

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Abstract The boron isotope ratio of seawater ($\delta^{11}\text{B}_{\text{sw}}$) is a parameter which must be known to reconstruct palaeo pH and CO₂ from boron isotope measurements of marine carbonates. Beyond a few million years ago, $\delta^{11}\text{B}_{\text{sw}}$ is likely to have been different to modern. Palaeo $\delta^{11}\text{B}_{\text{sw}}$ can be estimated by simultaneously constraining the vertical gradients in foraminiferal $\delta^{11}\text{B}$ ($\Delta\delta^{11}\text{B}$) and pH (ΔpH). A number of subtly different techniques have been used to estimate ΔpH in the past, all broadly based on assumptions about vertical gradients in oxygen, and/or carbon, or other carbonate system constraints. In this work we pull together existing data from previous studies, alongside a constraint on the rate of change of $\delta^{11}\text{B}_{\text{sw}}$ from modeling. We combine this information in an overarching statistical framework called a Gaussian Process. The Gaussian Process technique allows us to bring together data and constraints on the rate of change in $\delta^{11}\text{B}_{\text{sw}}$ to generate random plausible evolutions of $\delta^{11}\text{B}_{\text{sw}}$. We reconstruct $\delta^{11}\text{B}_{\text{sw}}$, and by extension palaeo pH, across the last 65Myr using this novel methodology. Reconstructed $\delta^{11}\text{B}_{\text{sw}}$ is compared to other seawater isotope ratios, namely $^{87/86}\text{Sr}$, $^{187/188}\text{Os}$, and $\delta^7\text{Li}$, which we also reconstruct with Gaussian Processes. Our method provides a template for incorporation of future $\delta^{11}\text{B}_{\text{sw}}$ constraints, and a mechanism for propagation of uncertainty in $\delta^{11}\text{B}_{\text{sw}}$ into future studies.

Plain Language Summary Boron naturally exists in two forms—¹¹B and ¹⁰B. Measuring the ratio of these two forms of boron within marine shells allows us to estimate how alkaline the ocean was in the past, which is related to how much carbon dioxide is in the atmosphere. Before we can do this calculation though, we need to know some other parameters, one of which is the relative abundance of the two forms of boron in the ocean at the time (which we call $\delta^{11}\text{B}_{\text{sw}}$). Preexisting studies have estimated $\delta^{11}\text{B}_{\text{sw}}$ at particular times, and here we combine them to generate a full reconstruction across the last 65 million years, accounting for uncertainties. Our reconstruction is informed by limiting the rate at which $\delta^{11}\text{B}_{\text{sw}}$ can change, based on model simulations. We provide a set of plausible evolutions of $\delta^{11}\text{B}_{\text{sw}}$ which can be used in future work when calculating past ocean pH.

1. Introduction

Boron and its two isotopes (¹⁰B and ¹¹B) exist in seawater with speciation governed by acid-base equilibrium. There are four degrees of freedom in the boron isotope system, which we typically express with the following five parameters: $\delta^{11}\text{B}_4$ (the isotopic composition of borate), $\delta^{11}\text{B}_{\text{sw}}$ (the isotope composition of seawater), ϵ (the fractionation factor), $\text{p}K_B^*$ (the apparent equilibrium constant for boron in solution), and pH. Knowing any four of these parameters allows us to calculate the fifth (Zeebe & Wolf-Gladrow, 2001). For a typical paleoclimatological application, we wish to determine palaeo pH, which is done by measuring $\delta^{11}\text{B}_{\text{calcite}}$ and translating this to $\delta^{11}\text{B}_4$ (potentially requiring a species specific calibration). $\delta^{11}\text{B}_4$ is then combined with the fractionation factor (ϵ —known from Klochko et al. (2006); Nir et al. (2015)), the apparent equilibrium constant for boron in seawater ($\text{p}K_B^*$ —which is estimated from temperature, pressure, and seawater composition (Dickson & Goyet, 1994)), and the boron isotope ratio of seawater ($\delta^{11}\text{B}_{\text{sw}}$). We refer the interested reader to Marschall and Foster (2018) for a full description of boron systematics in seawater, but for the purposes of this work it is sufficient to say that one of the key parameters that must be established in order to reconstruct palaeo ocean pH is $\delta^{11}\text{B}_{\text{sw}}$. Modern day seawater has a $\delta^{11}\text{B}_{\text{sw}}$ of $39.61 \pm 0.04 \text{‰}$ (Foster et al., 2010) - however $\delta^{11}\text{B}_{\text{sw}}$ is very likely to have been different in the past (Lemarchand et al., 2000).

Based on assessments of the relevant input and output fluxes to the ocean, boron is thought to have a relatively long residence time in seawater of roughly 10 million years (Broecker & Peng, 1982)—that is, the average amount of time an atom of boron spends in the ocean. It is estimated from geochemical box modeling of the boron cycle

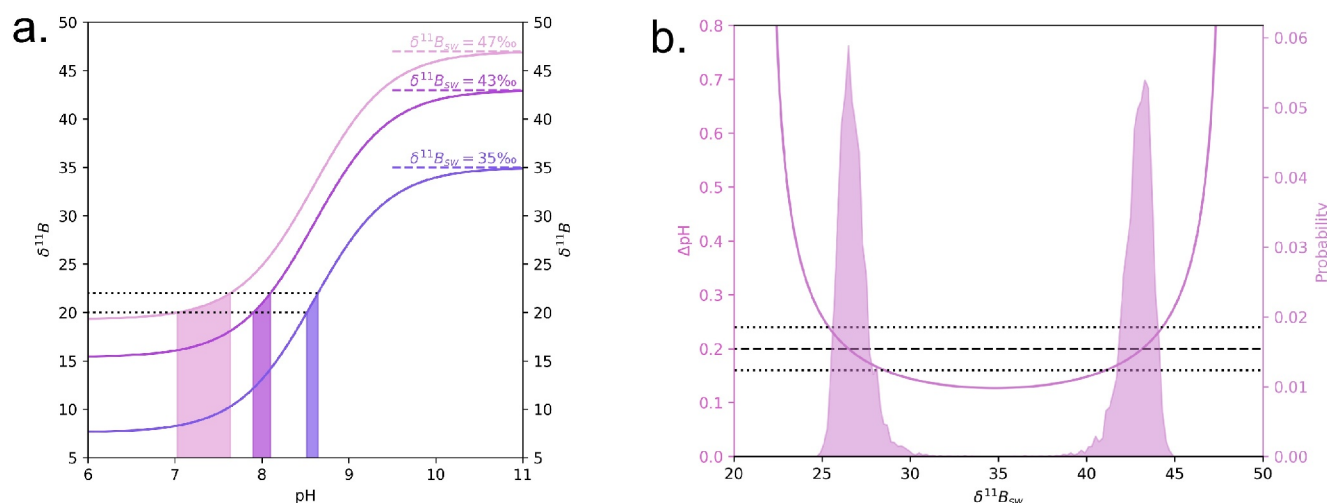


Figure 1. The left panel (a) shows a graph of pH versus $\delta^{11}\text{B}_4$ for three different $\delta^{11}\text{B}_{\text{sw}}$'s (indicated by lines and corresponding shaded windows). The y axis here can be used for both $\delta^{11}\text{B}_4$ and $\delta^{11}\text{B}_{\text{sw}}$. A hypothetical 2‰ excursion from 22 to 20‰ is depicted by the black dotted lines. The equivalent pH change for the same excursion at variable $\delta^{11}\text{B}_{\text{sw}}$ is displayed in the shaded region. Note this curve is symmetrical about the value of pK_B^* (roughly 8.6 for standard modern open ocean conditions). The right panel (b.) shows (for the same 22–20‰ excursion as shown in the left panel) ΔpH as a function of $\delta^{11}\text{B}_{\text{sw}}$ (pink line). A hypothetical constraint on ΔpH of 0.2 ± 0.04 (at 2σ) is shown by the black dashed lines, with uncertainty in the black dotted lines. Where this hypothetical constraint intersects with the pink line gives the region of possible $\delta^{11}\text{B}_{\text{sw}}$'s. We illustrate that a Monte Carlo approach, which samples possible ΔpH from the hypothetical constraint to give probability distributions (shown in pink shaded regions) for $\delta^{11}\text{B}_{\text{sw}}$, aligns with the emplaced constraint. Due to the symmetry of the curves shown in the left panel, there are often two $\delta^{11}\text{B}_{\text{sw}}$'s compatible with a ΔpH constraint. Typically the lower window is rejected as it would resolve to an unreasonably high pH.

Validation: Ross Whiteford, James W. B. Rae
Visualization: Ross Whiteford, Michael J. Henehan
Writing – original draft: Ross Whiteford, Timothy J. Heaton
Writing – review & editing: Ross Whiteford, Michael J. Henehan, Eleni Anagnostou, Hana Jurikova, Gavin L. Foster, James W. B. Rae

that $\delta^{11}\text{B}_{\text{sw}}$ might have changed by up to 0.1 ‰/Myr (Lemarchand et al., 2000). This means that beyond a few million years ago, ocean $\delta^{11}\text{B}_{\text{sw}}$ might have been appreciably different to modern day seawater, which must be taken into account when calculating palaeo pH. Unfortunately we currently have no direct archives of $\delta^{11}\text{B}_{\text{sw}}$. Halites show some promise as a direct proxy for $\delta^{11}\text{B}_{\text{sw}}$ (Paris et al., 2010), but uncertainty remains about the observed variability in halite $\delta^{11}\text{B}$ reconstructions, so $\delta^{11}\text{B}_{\text{sw}}$ is currently calculated using more indirect techniques.

The primary technique for estimating, or placing limits on, palaeo $\delta^{11}\text{B}_{\text{sw}}$ is to exploit vertical gradients in the ocean (Anagnostou et al., 2016; Foster et al., 2012; Greenop et al., 2017; Henehan et al., 2019). This works due to the sigmoidal relationship between $\delta^{11}\text{B}_4$ and pH (as shown in Figure 1a). The non-linearity of this relationship makes it possible to calculate $\delta^{11}\text{B}_{\text{sw}}$ given two $\delta^{11}\text{B}_4$ datapoints and an estimate of the change in pH across this $\delta^{11}\text{B}_4$ gradient. The vertical gradient in pH (ΔpH) is correlated to the vertical gradients in carbon isotopes ($\Delta\delta^{13}\text{C}$), carbon concentration (ΔDIC), and oxygen concentration ($\Delta[\text{O}_2]$), linked with Apparent Oxygen Utilisation—AOU due to their shared controlling processes (see Figure 2). Vertical gradients in all of these parameters can therefore be related to $\delta^{11}\text{B}_{\text{sw}}$. Paired measurements which allow the vertical gradient in $\delta^{11}\text{B}_4$ to be established are currently only available at select times within the Cenozoic—shown below in Table 1.

The simplest option for reconstructing palaeo $\delta^{11}\text{B}_{\text{sw}}$ from measured $\delta^{11}\text{B}$ in carbonates is to assume a constant ΔpH through time equal to modern, an approach used in Raitzsch and Hönisch (2013) (which assumed a ΔpH of 0.3 units). By combining this with the sigmoidal relationship of $\delta^{11}\text{B}_4$ and pH, it is possible to estimate $\delta^{11}\text{B}_{\text{sw}}$ (see Figure 1). To improve upon the assumption of a constant ΔpH through time, one can use foraminiferal $\Delta\delta^{18}\text{O}$ and $\Delta\delta^{13}\text{C}$ to guide identification of palaeo depth habitats of particular species, and to place reasonable limits on ΔpH (Anagnostou et al., 2016; Greenop et al., 2017; Palmer et al., 1998). For $\delta^{13}\text{C}$, this works because $\Delta\delta^{13}\text{C}$ and ΔpH are jointly controlled by subsurface remineralization of organic carbon, which releases isotopically light carbon into the water and causes acidification (decreasing pH)—as illustrated in Figure 2. The rationale for such constraints is broadly justified by the difficulty of reversing the sign of the gradients (which would require a fundamental change in biogeochemical dynamics), demonstrated by exploring a wide swath of model parameter space (Greenop et al., 2017) in carbon cycle models such as CYCLOPS (Hain et al., 2010), LOSCAR (Zeebe, 2012), and cGENIE (Ridgwell et al., 2007). That said, it is possible that models may not yet be capable of representing the full plethora of possible carbonate chemistry states that the ocean can truly inhabit. Despite recent progress (Caves Rügenstein et al., 2019; Derry, 2022) our understanding of Cenozoic carbon cycle fluxes, and

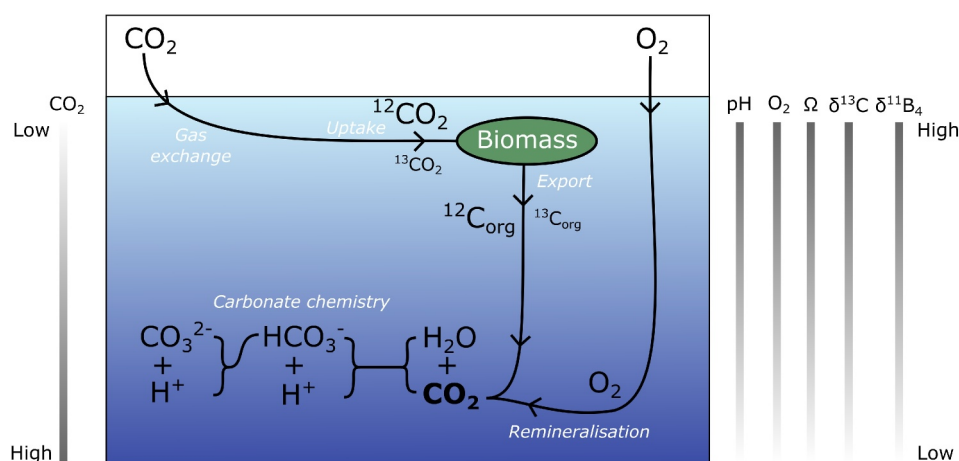


Figure 2. Schematic of the processes which cause vertical gradients in oceanic properties as harnessed in reconstructions of $\delta^{11}\text{B}_{\text{sw}}$. Oxygen and carbon dioxide concentrations are set in surface waters by their atmospheric concentrations and Henry's Law (which is temperature dependent). Surface ocean carbon is taken up by biomass to form organic tissues, which are preferentially enriched in the lighter carbon isotope. When that biomass is exported to depth and remineralized, organic carbon is oxidized (consuming oxygen) and releasing CO_2 , which drives acidification. These processes cause correlated gradients in pH, aqueous CO_2 , aqueous O_2 , saturation state, and $\delta^{13}\text{C}$. As $\delta^{11}\text{B}_{\text{sw}}$ is a well mixed signal, differences in $\delta^{11}\text{B}_4$ between the surface and subsurface are primarily driven by the gradient in pH. $\delta^{11}\text{B}_{\text{sw}}$ is homogenous throughout the water column due to the long 10 Myr residence time of boron (Foster et al., 2010; Lemarchand et al., 2000).

their drivers, remains incomplete, as does our understanding of their relationship to oceanic properties such as DIC and pH, and by extension the vertical gradients in these ocean properties. There are circumstances which would result in the “normal” relationships between these parameters breaking down, for instance if organic carbon is predominantly respired by sulfate reducers (which would alter local alkalinity and pH). While these complications are rare, it is prudent to keep in mind their potential to impact estimates of $\delta^{11}\text{B}_{\text{sw}}$. Despite these challenges, using model predicted gradients in $\delta^{13}\text{C}$ and pH with paired planktic-benthic (or surface-subsurface) foraminiferal $\delta^{11}\text{B}$ data remains attractive, as it requires only stable carbon isotope data (analyses which are routinely performed), and a model capable of providing estimates of feasible ocean profiles of $\delta^{13}\text{C}$ and pH.

$\delta^{11}\text{B}_{\text{sw}}$ from paired surface-deep samples can also be refined by estimation of AOU (Anagnostou et al., 2016). Surface ocean water oxygen saturation is mainly controlled by temperature (which is already a requirement for the $\delta^{11}\text{B}$ to pH calculation) and it is assumed based on modern observations that samples from deeper dwelling planktic foraminifera can not have inhabited anoxic waters (Hull et al., 2011). Similar to pH and $\delta^{13}\text{C}$, the vertical gradient in oxygen concentration is primarily a function of how much remineralization of organic carbon has occurred (see Figure 2). Remineralization adds CO_2 to the water, increasing the DIC concentration. A Redfield Ratio can be used to convert constraints on ΔO_2 to ΔDIC , however the impact on ΔpH is complicated by the buffering influence of alkalinity (Subhas et al., 2022). The calculation therefore requires that we know (or assume) a second carbonate system parameter (in addition to pH) to make ΔDIC calculable. This is not a major imposition, as quantification of palaeo CO_2 from boron isotope derived pH already requires assumption or knowledge of a second carbonate system parameter. Such an estimate may itself be derived from the same type of carbon cycle models used to estimate $\Delta\delta^{13}\text{C}$ and ΔpH . Carbon cycle models have been used to provide an estimate of saturation state for specific sites or times in the past, or to provide estimates of alkalinity (Anagnostou et al., 2016; Henahan et al., 2019), which simultaneously informs $\delta^{11}\text{B}_{\text{sw}}$ and palaeo CO_2 reconstructed from boron isotope derived pH.

Overall, it is therefore necessary to balance a number of requirements when reconstructing $\delta^{11}\text{B}_{\text{sw}}$. The established $\delta^{11}\text{B}_{\text{sw}}$ must produce a reasonable: ΔpH , $\Delta\delta^{13}\text{C}$, saturation state and AOU. Finding the space in which all these parameters are viable allows quantification of $\delta^{11}\text{B}_{\text{sw}}$ with attendant uncertainty. The boron isotope proxy is fortunate however, in that almost all data and assumptions required to estimate $\delta^{11}\text{B}_{\text{sw}}$ are already a part of the $\delta^{11}\text{B}_4$ to CO_2 calculation.

Table 1
Table of Data Constraints With Representative Central Values and 95% Confidence Intervals as an Indicator of Uncertainty

Age (Myr)	Central estimate (‰)	95% interval (‰)	Uncertainty structure	Source
0.0	39.61	0.04	Gaussian	Foster et al. (2010)
0.68	39.69	4.28	Non-Gaussian	Greenop et al. (2017)
1.14	38.55	3.74	Non-Gaussian	Greenop et al. (2017)
1.16	41.60	3.27	Non-Gaussian	Greenop et al. (2017)
1.45	40.51	2.00	Non-Gaussian	Greenop et al. (2017)
2.27	41.57	1.66	Non-Gaussian	Greenop et al. (2017)
2.87	39.82	2.05	Non-Gaussian	Greenop et al. (2017)
5.37	40.38	2.78	Non-Gaussian	Greenop et al. (2017)
8.67	42.30	1.88	Non-Gaussian	Greenop et al. (2017)
9.33	40.22	2.28	Non-Gaussian	Greenop et al. (2017)
10.14	36.35	5.98	Non-Gaussian	Greenop et al. (2017)
11.62	40.34	2.05	Non-Gaussian	Greenop et al. (2017)
12.27	35.69	5.99	Non-Gaussian	Greenop et al. (2017)
12.80	37.47	2.53	Non-Gaussian	Greenop et al. (2017)
13.53	36.43	5.89	Non-Gaussian	Greenop et al. (2017)
16.39	36.49	5.76	Non-Gaussian	Greenop et al. (2017)
17.69	37.00	5.66	Non-Gaussian	Greenop et al. (2017)
19.00	41.10	1.95	Non-Gaussian	Greenop et al. (2017)
19.67	40.44	1.93	Non-Gaussian	Greenop et al. (2017)
22.62	39.41	4.47	Non-Gaussian	Greenop et al. (2017)
22.98	34.65	6.10	Non-Gaussian	Greenop et al. (2017)
23.08	39.11	3.50	Non-Gaussian	Greenop et al. (2017)
37.00	37.63	2.36	Non-Gaussian	Anagnostou et al. (2016)*
44.40	38.51	0.75	Non-Gaussian	Anagnostou et al. (2016)*
45.60	37.81	1.18	Non-Gaussian	Anagnostou et al. (2016)*
53.00	38.49	0.70	Non-Gaussian	Anagnostou et al. (2016)*
55.80	38.94	0.41	Gaussian	Gutjahr et al. (2017)
66.04	39.30	0.50	Uniform	Henehan et al. (2019)*

Note. * symbol indicates the values has been updated from the original publication for this work.

Most previous applications of the techniques described above have focused on individual time slices of the Cenozoic (as described in Section 2 and Table 1). By combining each of these individual studies, we are able to reach a critical mass of information, allowing us to produce an estimate for how $\delta^{11}\text{B}_{\text{sw}}$ evolved throughout the Cenozoic. Our study represents an advance compared to two particularly relevant previous works that provided Cenozoic timescale estimates of $\delta^{11}\text{B}_{\text{sw}}$. Raitzsch and Hönisch (2013) predicted Cenozoic $\delta^{11}\text{B}_{\text{sw}}$ by assuming a constant ΔpH and a linear trend in deep ocean pH, meaning reconstructions of pH using this $\delta^{11}\text{B}_{\text{sw}}$ are only able to reconstruct the initial assumption (J. W. B. Rae, 2018) (further impacts of which are discussed in Section 5.3). J. W. Rae et al. (2021) compiled and reanalyzed marine palaeo CO_2 proxy data (including boron isotopes), and therefore required a curve for $\delta^{11}\text{B}_{\text{sw}}$. However, generation of a Cenozoic $\delta^{11}\text{B}_{\text{sw}}$ curve was not the primary focus of that study and the authors noted the need to improve interpolation and constrain uncertainty. Here we provide a more robust estimate of the evolution of $\delta^{11}\text{B}_{\text{sw}}$ through time by bringing existing and updated $\delta^{11}\text{B}_{\text{sw}}$ constraints together into a rigorous statistical framework called a Gaussian Process.

The Gaussian Process is a flexible Bayesian non-parametric regression technique. We provide further details in the methods, a schematic illustration in Figure 3, and a full description in Supporting Information S1. Intuitively, we can consider the Gaussian Process as a method for creating stochastic smooth curves influenced to go through

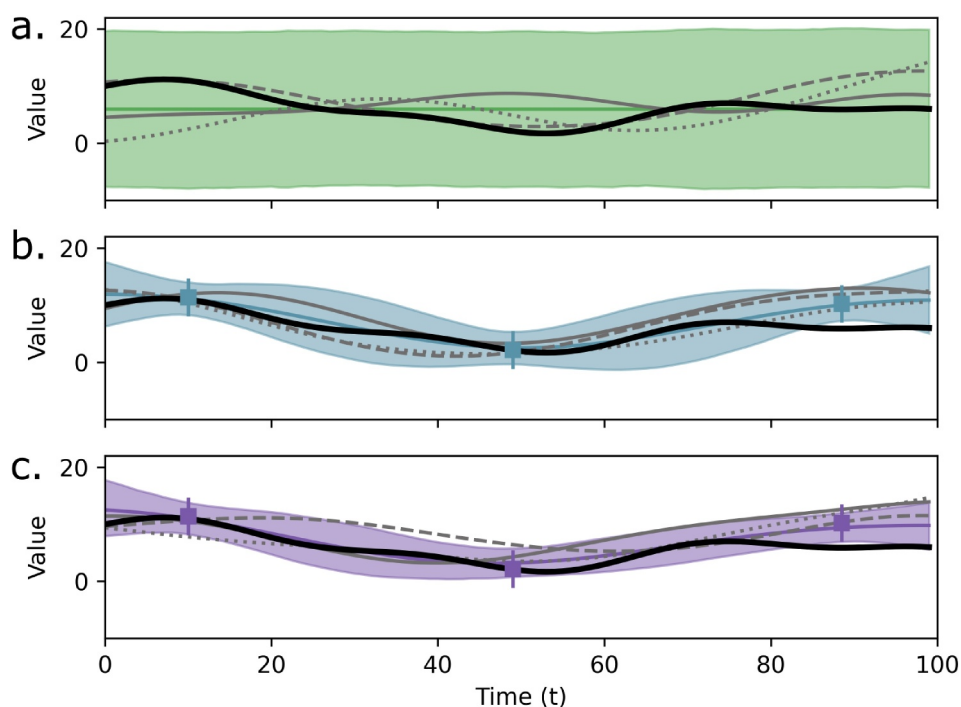


Figure 3. An illustrative example of the Gaussian Process approach applied to reconstruct a hypothetical time series (shown in black). Each subplot shows the mean reconstruction (colored line), 95% confidence interval (colored window), and three random Gaussian Process samples (gray lines). Panel (a) shows a Gaussian Process prior with set mean, kernel function, and hyperparameters but without any data constraints. Panel (b) shows a Gaussian Process posterior with the same prescribed mean and hyperparameters as panel a., but with three noisy data constraints (shown with 2σ uncertainties in blue). Panel (c) shows the same Gaussian Process posterior as panel (b), where, in addition to the three noisy data constraints, we have applied additional constraints on the acceptable rate of change of the value through time.

datapoints. For $\delta^{11}\text{B}_{\text{sw}}$, we provide data constraints and prescribe the smoothness based on mechanistic modeling, which allows us to can draw many potential evolutions of $\delta^{11}\text{B}_{\text{sw}}$. We are able to impose additional constraints (such as on the rate of change) by filtering the potential evolutions of $\delta^{11}\text{B}_{\text{sw}}$ to find those which are compatible with all available information—thus providing an improved estimate of $\delta^{11}\text{B}_{\text{sw}}$. The realizations can be summarized to provide an estimate of $\delta^{11}\text{B}_{\text{sw}}$ at any specific time or used as inputs to other models as a method for propagating uncertainty in $\delta^{11}\text{B}_{\text{sw}}$.

2. Data

To create our curves, we combine observational data regarding levels of $\delta^{11}\text{B}_{\text{sw}}$ at specific time points from a range of previous studies. These observational estimates come with various forms of uncertainty, primarily due to how the raw (direct) measurements were processed/transformed into values for $\delta^{11}\text{B}_{\text{sw}}$. Some of these previous studies report observed values that relate to central estimates of the underlying $\delta^{11}\text{B}_{\text{sw}}$ at the specific time of the sample (Gutjahr et al., 2017; Henehan et al., 2019, 2020). Others provide lower- and upper-bounds for the possible value of $\delta^{11}\text{B}_{\text{sw}}$ (Anagnostou et al., 2016). A further study provides complete probability distributions for possible values of $\delta^{11}\text{B}_{\text{sw}}$ given their observational constraints (Greenop et al., 2017).

Two of those studies reconstructed $\delta^{11}\text{B}_{\text{sw}}$ for a particular event (or short timeslice) (Gutjahr et al., 2017; Henehan et al., 2019), and two reconstructed $\delta^{11}\text{B}_{\text{sw}}$ over a wider time window, within which $\delta^{11}\text{B}_{\text{sw}}$ might have evolved (Anagnostou et al., 2016; Greenop et al., 2017). Data constraints on $\delta^{11}\text{B}_{\text{sw}}$ are varied. Some previous works presented central constraints of $\delta^{11}\text{B}_{\text{sw}}$ (Gutjahr et al., 2017; Henehan et al., 2019, 2020), while another provided lower or upper limits on $\delta^{11}\text{B}_{\text{sw}}$ (Anagnostou et al., 2016), and another provided full probability distributions for possible values of $\delta^{11}\text{B}_{\text{sw}}$ (Greenop et al., 2017). Two of those studies reconstructed $\delta^{11}\text{B}_{\text{sw}}$ for a particular event (or short timeslice) (Gutjahr et al., 2017; Henehan et al., 2019), and two reconstructed $\delta^{11}\text{B}_{\text{sw}}$ over a wider time window, within which $\delta^{11}\text{B}_{\text{sw}}$ might have evolved (Anagnostou et al., 2016; Greenop et al., 2017).

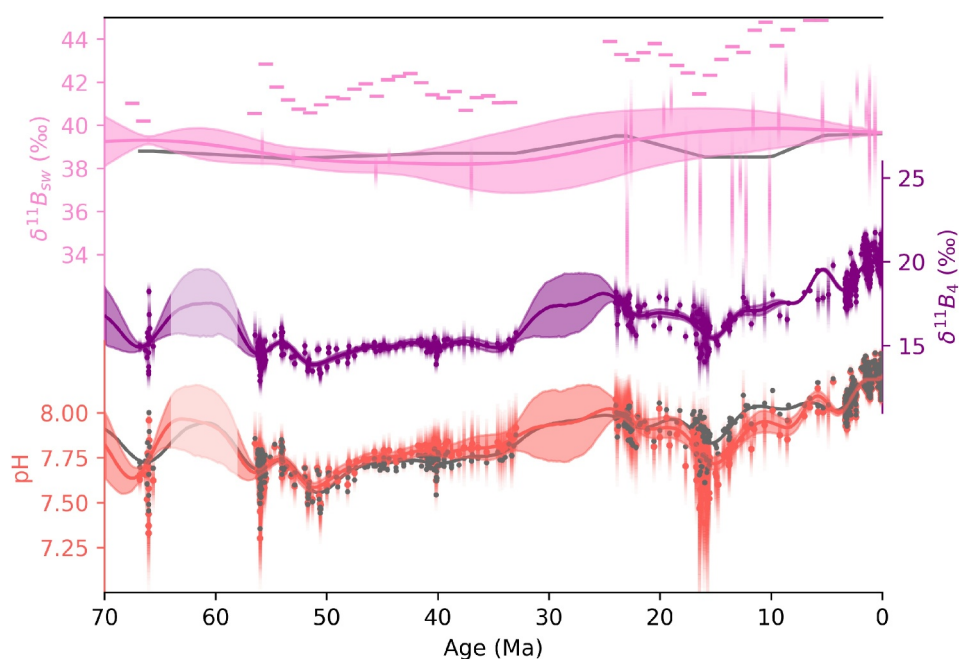


Figure 4. Our reconstruction of $\delta^{11}\text{B}_{\text{sw}}$ is shown in pink, with a central line depicting the median, and a window representing the 95% confidence interval. Central data constraints are shown by vertical pink bars representative of 95% confidence intervals, and upper limits (obtained based upon the planktic $\delta^{11}\text{B}_4$ measurements) are shown by horizontal pink bars. The gray line depicts $\delta^{11}\text{B}_{\text{sw}}$ reconstructed by J. W. Rae et al. (2021) for comparison. $\delta^{11}\text{B}_4$ datapoints from J. W. Rae et al. (2021) are shown as purple dots (interpolated in the purple shaded region using a Gaussian Process), which are converted to pH using our $\delta^{11}\text{B}_{\text{sw}}$ (red points, line, and shaded region) and the $\delta^{11}\text{B}_{\text{sw}}$ from J. W. Rae et al. (2021) (gray points and line) for comparison. We note that our Gaussian Process interpolation of $\delta^{11}\text{B}_4$ (and therefore also pH) has difficulty during large data gaps. In particular during the Palaeogene (which we have faded out), the large data gap is bounded by events with rapid changes in $\delta^{11}\text{B}_4$, and our reconstruction would be tempered by filling in this region with $\delta^{11}\text{B}_4$ data.

Table 1 contains the compilation of datapoints used to create our estimate of $\delta^{11}\text{B}_{\text{sw}}$ using a Gaussian Process. Some of these data are as originally published, and others have been updated and refined. We use the $\delta^{11}\text{B}_{\text{sw}}$ values from Greenop et al. (2017) exactly as reported in the original study. The original published estimate from Henehan et al. (2019), however, is based on carbonate system calculations whose equilibrium constants use fitting parameters from the supplementary tables of Hain et al. (2015). These tables have since been found to contain inaccuracies (CenCO2PIP Consortium, 2023), and so here we updated carbonate system calculations, and hence $\delta^{11}\text{B}_{\text{sw}}$ constraints, based on corrected equilibrium constants packaged with Raitzsch et al. (2022). Updated carbonate system calculations and CO_2 estimates for this time are provided in Supporting Information S1, and are plotted in Henehan and Witts (2023). Datapoints from Anagnostou et al. (2016) are updated by modifying the calculation of vital effects to account for changing seawater chemistry. With these recalculations, Henehan et al. (2019)'s estimate of $\delta^{11}\text{B}_{\text{sw}}$ is updated from the originally published value of 39.05–39.85 ‰ to 39.05–39.55 ‰. Finally, estimates from Anagnostou et al. (2016) are integrated as full probability distributions instead of the lower/upper limits that were presented in the original study (see Table 1 and Data Set S1 in Supporting Information S2).

To supplement these existing $\delta^{11}\text{B}_{\text{sw}}$ constraints, we take planktic $\delta^{11}\text{B}_4$ data as presented in J. W. Rae et al. (2021), and shown in Figure 4, based on original studies by Anagnostou et al. (2016); Anagnostou et al. (2020); Badger et al. (2013); Chalk et al. (2017); de la Vega et al. (2020); Dyez et al. (2018); Foster (2008); Foster et al. (2012); Greenop et al. (2014, 2017, 2019); Gutjahr et al. (2017); Harper et al. (2020); Henehan et al. (2019, 2020); Hönlisch et al. (2009); Martínez-Botí et al. (2015); Pearson et al. (2009); Penman et al. (2014); Sosdian et al. (2018), to calculate a range of valid $\delta^{11}\text{B}_{\text{sw}}$ by exploiting the sigmoidal shape of the relationship between pH and $\delta^{11}\text{B}_4$ (as seen in Figure 1). The maximum offset between $\delta^{11}\text{B}_4$ and $\delta^{11}\text{B}_{\text{sw}}$ is seen at low pH, and is described by the fractionation factor ϵ (or α). For any given estimate of $\delta^{11}\text{B}_4$, the minimum valid $\delta^{11}\text{B}_{\text{sw}}$ is equal to $\delta^{11}\text{B}_4$, and the maximum $\delta^{11}\text{B}_{\text{sw}}$ can be calculated by combination with the fractionation factor. The nature of the relationship between $\delta^{11}\text{B}_4$ and $\delta^{11}\text{B}_{\text{sw}}$ is examined in Figure 1a. For any given $\delta^{11}\text{B}_4$, in order to sit

somewhere on the sigmoidal curve, $\delta^{11}\text{B}_{\text{sw}}$ must be equal to or greater than $\delta^{11}\text{B}_4$, but no more than $\delta^{11}\text{B}_4$ and the fractionation factor (the height of the sigmoid shape—strictly given by $\delta^{11}\text{B}_{\text{sw}} = \delta^{11}\text{B}_4 \times \alpha + \epsilon$). To create these additional constraints, we first bin the planktic $\delta^{11}\text{B}_4$ data into 1 Myr intervals, then take the maximum measured $\delta^{11}\text{B}_4$ as the lower limit for $\delta^{11}\text{B}_{\text{sw}}$, and use the minimum measured $\delta^{11}\text{B}_4$ to calculate the maximum $\delta^{11}\text{B}_{\text{sw}}$. Uncertainty in measured $\delta^{11}\text{B}_{\text{foram}}$ is propagated to uncertainty in the possible $\delta^{11}\text{B}_{\text{sw}}$, and the upper 99% quantile is used to estimate the maximum permissible $\delta^{11}\text{B}_{\text{sw}}$. As described, lower limits can be calculated using this method, however these values are so low as to be uninformative across the Cenozoic. Upper limits created using this method are shown in Figure 4 (the horizontal pink bars).

In this work, we therefore have four potential categories of constraint: constraints where the uncertainty is well represented by a Gaussian distribution (Gaussian constraints), constraints where the uncertainty is not well represented by a Gaussian distribution (non Gaussian constraints), lower and upper limits, and limitations on the rate of change in $\delta^{11}\text{B}_{\text{sw}}$ from modeling.

3. Methods

Having assembled the data described in Section 2, we reconstruct $\delta^{11}\text{B}_{\text{sw}}$ using a Gaussian Process (GP). The Gaussian Process is a statistical technique that allows us to generate smooth time series conditioned to match data constraints (see Figure 3, or for a fuller, more rigorous description of a Gaussian Process—see Supporting Information S1). In the case of $\delta^{11}\text{B}_{\text{sw}}$, we have an expectation of smoothness from the modeling work of Lemarchand et al. (2000), which we combine with data constraints as described in Section 2 and shown in Figure 4. A Gaussian Process works by using a kernel function, which encodes structure into the reconstruction, and hyperparameters which tune the behavior. Here we use the squared exponential kernel (also known as the Radial Basis Function), which expresses that nearby points are more likely to be similar to each other than distant points. There are two controlling hyperparameters—one expresses the length scale over which there is significant covariance, which is related to the expected rate of change in a signal. The other parameter is the noise scale—which manifests as how much uncertainty is expected at times where no data constraints are available. The Gaussian Process can be used to generate random smooth lines with the prescribed characteristics even in the absence of data constraints (Figure 3a.), however data can be incorporated by adapting the covariance structure such that the statistical samples it generates will go through datapoints where permitted by the chosen hyperparameters (Figure 3b.). Uncertainty in data constraints can be incorporated directly into the Gaussian Process if uncertainty in the estimates is Gaussian in nature, and other types of constraint can be incorporated by adapting the approach.

The Gaussian Process is able to provide a number of equally likely, independent, stochastic time series, which are useful in reconstructing time series of palaeo data with an estimate of the uncertainty. Each obeys the smoothness constraint encapsulated by the kernel, while simultaneously attempting to match any available data. The result is that where a data constraint with low uncertainty is available, time series will be strongly influenced to go through this datapoint. Where data are sparse, or data constraints are uncertain, statistical replicates diverge to represent increasing uncertainty (see Figure 3b or Supporting Information S1 for a demonstration of this behavior). Because each sample drawn from the Gaussian process is independent and equally likely, we can straightforwardly apply filters to reject any sample that has undesirable properties (Figure 3c). As mentioned above, data constraints where uncertainties can be well represented by a Gaussian distribution can be directly assimilated into the Gaussian Process. Here we also wish to enforce three other types of constraint, lower/upper limitations on $\delta^{11}\text{B}_{\text{sw}}$, limitations on the rate of change in $\delta^{11}\text{B}_{\text{sw}}$, and data constraints with a non Gaussian uncertainty structure. Limitations (either on the value of $\delta^{11}\text{B}_{\text{sw}}$ or the rate of change in $\delta^{11}\text{B}_{\text{sw}}$) are relatively straightforward to enforce—we can compare each Gaussian Process sample to the limitation and simply reject those which are outside the established limits. Non-Gaussian constraints are slightly more difficult to integrate, but can also be done within a rejection framework by approximating each non-Gaussian constraint with a Gaussian distribution. Here we do this by creating Gaussian constraints with the same mean as their non-Gaussian counterpart, but with higher standard deviation. Samples can then be strategically rejected to tweak the Gaussian Process such that it has effectively sampled from non-Gaussian constraints (the mechanics of this type of rejection sampling are clarified in Supporting Information S1).

One difficulty of rejection sampling, however, is dealing with simultaneous but mutually exclusive data constraints. Gaussian distributions are nonzero across their domain, meaning no value (however unlikely) is truly

impossible. The nature of the non-Gaussian constraints in this work though, suggests that at any given time some $\delta^{11}\text{B}_{\text{sw}}$ values are impossible to reconcile with the data estimates. Given multiple constraints of this type, in close proximity to one another, it becomes exceedingly unlikely to draw samples from the Gaussian Process which have the requisite smoothness and pass through the possible region of every datapoint. We overcome this difficulty by giving each non-Gaussian constraint the possibility of being an outlier value. In summary, Gaussian Process samples are not always rejected completely where in disagreement with data, but the probability of acceptance is highest where in agreement with the data constraints, and lower (but non zero) further from the data constraints.

We reconstruct $\delta^{11}\text{B}_{\text{sw}}$ by conditioning a Gaussian Process on the few available central estimates of $\delta^{11}\text{B}_{\text{sw}}$ (shown in Table 1). Additional data constraints are integrated by filtering the generated statistical samples using rejection sampling (as described above and in Supporting Information S1). Any curves which fall outside the lower or upper limits are rejected, and some curves are rejected to adjust the Gaussian Process posterior such that it appropriately incorporates non Gaussian constraints. We use hyperparameters of 2 ‰ noise scale, and 10 Myr length scale. The Gaussian Process methodology does not enforce a specific rate of change, but the chosen noise scale and length scale enable an approximation of the rate of change, in this case roughly equivalent to a 0.2 ‰/Myr rate of change. This is on the faster end of agreement with the residence time and maximum rate of change from box modeling of the geochemical cycle of boron (Lemarchand et al., 2000). Lemarchand et al. (2000) suggest an upper limit to the rate of change of $\delta^{11}\text{B}_{\text{sw}}$ of 0.1 ‰/Myr. This estimate is based on modern day fluxes, so we allow greater uncertainty in the past when the boron fluxes in and out of the ocean may have been different, increasing the permissible maximum rate of change linearly to 0.7 ‰/Myr at 70 Ma (discussed further in Section 5.2). Output samples are filtered such that those with a rate of change greater than the values described above are rejected. This, alongside the rejection strategy used to integrate non-Gaussian constraints, leaves a subset of the originally generated samples. Approximately 10 in each 10,000 are accepted, and we run the algorithm until 10,000 samples have been accepted.

4. Results

Our reconstruction of $\delta^{11}\text{B}_{\text{sw}}$ suggests rather muted change across the Cenozoic relative to the change in $\delta^{11}\text{B}_4$. $\delta^{11}\text{B}_4$ has changed by roughly 8 ‰, and we find that $\delta^{11}\text{B}_{\text{sw}}$ has been responsible for perhaps 2 ‰ of that change, leaving a 6 ‰ change driven predominantly by pH. The pattern of change in $\delta^{11}\text{B}_{\text{sw}}$ is non-linear, with a decline of roughly 1 ‰ from the early to mid Cenozoic, followed by an increase of approximately 1.5 ‰ from 30 Ma to 10 Ma, and a relatively stable value between those time periods and from 10 Ma until the present day. At times there is large uncertainty in the absolute value of $\delta^{11}\text{B}_{\text{sw}}$, in particular around 30 Ma, however the relative change in $\delta^{11}\text{B}_{\text{sw}}$ is more certain and the rise in $\delta^{11}\text{B}_{\text{sw}}$ between 40 and 10 Ma—also illustrated in Figure 4—appears robust. Our reconstruction generally matches the evolution curve used by J. W. Rae et al. (2021), except in the interval between 20 Ma and 10 Ma. Within this interval both curves are guided predominantly by the data from Greenop et al. (2017), but J. W. Rae et al. (2021) used Greenop et al. (2017)'s binned values whereas our approach uses the probability distribution of each individual estimate. Our Gaussian Process methodology favors the higher $\delta^{11}\text{B}$ values in this interval due to their lower reported uncertainty (relative to the lower $\delta^{11}\text{B}_{\text{sw}}$ estimates from Greenop et al. (2017)). This results in an overall higher estimate of $\delta^{11}\text{B}_{\text{sw}}$ during the Miocene, and implies pH was potentially lower than previously calculated at this time. However, we note the uncertainties in both our reconstruction and the individual $\delta^{11}\text{B}_{\text{sw}}$ constraints are high in the Neogene, and that some of the data constraints suggest rates of change in $\delta^{11}\text{B}_{\text{sw}}$ exceeding those compatible with geochemical box model predictions (as discussed further in Section 5.2). Uncertainty on $\delta^{11}\text{B}_{\text{sw}}$ is also particularly high in data gaps where $\delta^{11}\text{B}_4$ has yet to be measured. In particular we note a large window during the Oligocene within which no boron isotope data has yet been published, and which correspondingly has high uncertainty in $\delta^{11}\text{B}_{\text{sw}}$.

We provide illustrative curves for $\delta^{11}\text{B}_4$ and pH (see Figure 4) by using a Gaussian Process to interpolate $\delta^{11}\text{B}_4$, then combining this with our predicted $\delta^{11}\text{B}_{\text{sw}}$ and other ancillary parameters as in J. W. Rae et al. (2021). The Gaussian Process used to reconstruct $\delta^{11}\text{B}_4$ uses a length scale of 2 Myr to target relatively long term changes in pH, rather than individual paleoclimatic events.

The pattern of change we find in $\delta^{11}\text{B}_{\text{sw}}$ is broadly similar to temporal trends in a number of other related biogeochemical signals, such as oceanic $^{87/86}\text{Sr}$, $^{187/188}\text{Os}$, and $\delta^7\text{Li}$. Here we take published $^{87/86}\text{Sr}$ and $\delta^7\text{Li}$ from Misra and Froelich (2012), and compile a new record of $^{187/188}\text{Os}$ from Josso et al. (2019); Klemm et al. (2005); Oxburgh (1998); Oxburgh et al. (2007); Paquay et al. (2008, 2014); Pegram and Turekian (1999); Peucker-

Ehrenbrink and Ravizza (2000); van der Ploeg et al. (2018); Ravizza (1993); Ravizza and Turekian (1992); Ravizza and Peucker-Ehrenbrink (2003b, 2003a); Reusch et al. (1998); Robinson et al. (2009). For each signal, we assume that data constraints have Gaussian uncertainty (with magnitude given with original estimates, except in the case of $^{87/86}\text{Sr}$ where an illustrative uncertainty is used), which allows us to perform a straightforward Gaussian Process reconstruction, with the residence time of each element informing the length scale used ($^{87/86}\text{Sr}$ —5.1 Myr (Broecker & Peng, 1982), $\delta^7\text{Li}$ —2.8 Myr (Stoffynegli & Mackenzie, 1984), $^{187/188}\text{Os}$ - 1 Myr). Interpolated data products for both $^{87/86}\text{Sr}$ and $\delta^7\text{Li}$ have been previously published (for instance in Misra and Froelich (2012)), but to our knowledge this is the first interpolated data product available for $^{187/188}\text{Os}$ (though an interpolation is shown in Torres et al. (2014). There remains uncertainty in the residence time of osmium, but it is thought to be extremely short relative to the other signals shown here (all estimates are <100kyr—see e.g. Oxburgh (2001)). We use a length scale of 1 Myr in the reconstruction to balance the available data density against the short residence time. Overall, our reconstructions use a more sophisticated fitting strategy than previous incarnations, which is guided primarily by the data and integrates information on residence time of each element. This allows us to provide a robust quantification of the uncertainty in each signal, within the limitations of currently available data density.

Each of the seawater isotope signals we examine, like $\delta^{11}\text{B}_{\text{sw}}$, shows a long term increase over the Cenozoic. Oceanic strontium isotopes appear to show most similarity with our reconstructed $\delta^{11}\text{B}_{\text{sw}}$, while we note resemblance between the evolutions of pH, lithium isotopes, and osmium isotopes (as shown in Figure 5). We are able to provide an extremely narrow uncertainty window on our $^{87/86}\text{Sr}$ reconstruction due to strontium's long residence time and relatively high data density (and data quality) over the Cenozoic. By contrast, $^{187/188}\text{Os}$ which has greater data density and comparable scale of uncertainty in datapoints, has more uncertainty in our reconstruction. This is due to osmium's extremely short residence time, and persists despite us using an artificially long length scale to reconstruct this signal. $\delta^7\text{Li}$ by contrast has quite large uncertainties, due in large part to the uncertainty in the individual data constraints. We are unable to reconstruct the rapid shift in $\delta^7\text{Li}$ at the K-Pg boundary, which occurs faster than the modern day residence time of lithium (see Section 5.2 for further discussion). Despite these challenges we believe reconstruction of these signals benefits from the Gaussian Process approach, in that uncertainties on our reconstruction are more representative, and the timescale of change in each signal aligns more closely to our expectations based on their residence times.

5. Discussion

5.1. Decomposition of $\delta^{11}\text{B}_4$

$\delta^{11}\text{B}_4$ is an integrated signal, combining the effects of $\delta^{11}\text{B}_{\text{sw}}$, pH, pK_b^* , and ϵ . Of these, only a few are time-variable: $\delta^{11}\text{B}_{\text{sw}}$, pH, and some of the factors which influence pK_b^* —in particular temperature, and seawater elemental composition. The sensitivity of $\delta^{11}\text{B}_4$ to these factors (illustrated in Figure S2 in Supporting Information S1) is such that only changes in $\delta^{11}\text{B}_{\text{sw}}$ or pH would be of sufficient magnitude to drive the observed 8 ‰ change in $\delta^{11}\text{B}_4$ across the Cenozoic. Temperature may also have driven some portion of the change, though estimating how much is complicated because individual sample sites may deviate significantly from global average temperature. All other factors being equal, a temperature change of +10°C drives approximately -1.5 ‰ change in $\delta^{11}\text{B}_4$. Temperature change is taken into account in all calculations presented here., and if we adjust for any secular evolution of temperature we might therefore simplify the $\delta^{11}\text{B}_4$ record into two components, $\delta^{11}\text{B}_{\text{sw}}$ and pH. Any change in $\delta^{11}\text{B}_4$ not explained by changes in $\delta^{11}\text{B}_{\text{sw}}$ must be the result of pH, and vice versa. Given the observed 8 ‰ change in $\delta^{11}\text{B}_4$, our reconstructed 2 ‰ change in $\delta^{11}\text{B}_{\text{sw}}$, and temperature change across the Cenozoic of approximately 12°C driving approximately 1.8 ‰ of change, this suggests roughly 4 ‰ of change has been driven by changing pH. That is roughly equivalent to an increase in pH of 0.4 units across the Cenozoic, or a decrease in hydrogen ion availability by 60%. By comparison, anthropogenic CO_2 release has driven a surface ocean pH change of roughly 0.15 units, equivalent to a 41% increase in hydrogen ion availability (Findlay et al., 2022) relative to preindustrial conditions. Relative to their respective initial conditions, the anthropogenic surface ocean $[\text{H}^+]$ perturbation is therefore approximately two thirds the magnitude of long term change seen within the last 65 million years.

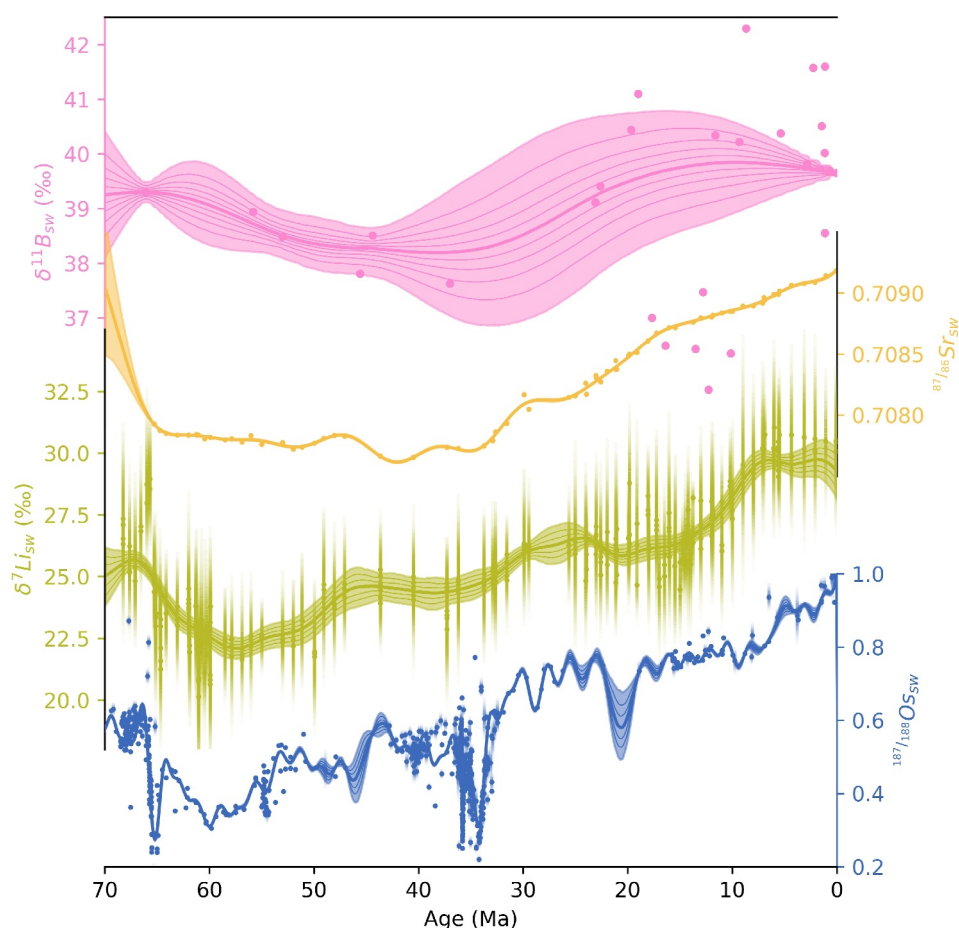


Figure 5. Our reconstruction of $\delta^{11}\text{B}_{\text{sw}}$ is shown pink, with a thick central line depicting the median, and a window representing the 95% confidence interval. Thinner pink lines give a sense of the covariance structure in our posterior estimate of the $\delta^{11}\text{B}_{\text{sw}}$. Data constraints are shown by the pink dots (here displayed without uncertainty for clearer comparison of signal trends). Other related geochemical signals $\delta^7\text{Li}$, $^{87/86}\text{Sr}$, and $^{187/188}\text{Os}$ are shown in yellow, green, and blue respectively after Misra and Froelich (2012). Data constraining $^{87/86}\text{Sr}$ and $\delta^7\text{Li}$ is sourced from Misra and Froelich (2012), while data constraining $^{187/188}\text{Os}$ is compiled from a number of sources (listed in the main text). We interpolate each of these signals with a Gaussian Process with hyperparameters guided by the residence time of each signal.

5.2. $\delta^{11}\text{B}_{\text{sw}}$ Rate of Change

A key feature of the Gaussian Process methodology used here is the ability to draw smooth potential evolutions of $\delta^{11}\text{B}_{\text{sw}}$, which we augment by filtering samples to reject those with an unfeasibly high temporal gradient. The feasibility of the rate of change in $\delta^{11}\text{B}_{\text{sw}}$ (and the hyperparameters which tune the Gaussian Process smoothness) are primarily based on the geochemical box modeling work of Lemarchand et al. (2000). Lemarchand et al. (2000) suggest that the likely maximum rate of change in $\delta^{11}\text{B}_{\text{sw}}$ is 0.1 ‰/Myr. We acknowledge that this is based on modern fluxes—fluxes which are not exceptionally well constrained (Park & Schlesinger, 2002)—so (as described in Section 3) we allow increasing maximum rates of change further back into the past - up to 0.7 ‰/Myr at 70 Ma. We choose the value of 0.7 ‰/Myr at 70 Ma because it allows us to enforce a rate limit near 0.1 ‰/Myr when we are most sure of that value (close to modern), without enforcing a strong limit earlier in the Cenozoic when we are much less certain about the residence time of boron and potential rate of change in $\delta^{11}\text{B}_{\text{sw}}$. 0.7 ‰/Myr is large enough that it does not result in rejection of potential evolutions in the earlier half of the Cenozoic (see Figure S3 in Supporting Information S1), meaning the influence of the rate of change limit is mostly constrained to the rejection of samples as a result of their temporal gradient during the Neogene.

Suggestions of a higher rate of change in $\delta^{11}\text{B}_{\text{sw}}$ than previously recognized are present in the data set of Greenop et al. (2017). The surface-deep $\delta^{11}\text{B}_4$ pairs of Greenop et al. (2017) (as described in Section 2) appear to record a

bimodal distribution, with some indicating $\delta^{11}\text{B}_{\text{sw}}$ similar to modern, and others (mostly in the middle Miocene climatic optimum) suggesting $\delta^{11}\text{B}_{\text{sw}}$ roughly 2 ‰ lower. These would appear to suggest oscillations in $\delta^{11}\text{B}_{\text{sw}}$ more rapid than would be consistent with a rate of change of 0.1 ‰/Myr. We find that in order to match these data constraints, our Gaussian Process would need to have a length scale of approximately 3 Myr (which is very close to Park and Schlesinger (2002)'s estimate of the residence time of boron at 3.3 Myr). Our understanding of the primary difference between the estimate of Park and Schlesinger (2002) and Lemarchand et al. (2000)'s is that is driven by different estimate of the atmospheric fluxes of boron, which are particularly difficult to constrain. We draw evolutions consistent with the rate of change suggested by Lemarchand et al. (2000), which is the current paradigm for palaeo pH estimation from boron isotopes—however we note the possibility for future changes in our understanding of the rate at which $\delta^{11}\text{B}_{\text{sw}}$ may have changed. Our method requires that we give each non-Gaussian constraint (shown in Table 1) the possibility of being an outlier. The Gaussian Process then forges a path roughly through the center of these constraints, drawn slightly high by the purportedly lower uncertainties in the higher $\delta^{11}\text{B}_{\text{sw}}$ estimates.

In summary, we proceed here under the existing paradigm of a long residence time for boron in seawater, and a slow rate of change in $\delta^{11}\text{B}_{\text{sw}}$. However, we acknowledge the possibility that the rate of change in $\delta^{11}\text{B}_{\text{sw}}$ is faster than currently appreciated, and look to further modeling efforts and an improved understanding of the boron cycle to guide future interpretation of the data collated here. Faster rates of change may be explored with the algorithm presented here, but this increases the requirement for higher resolution data constraining $\delta^{11}\text{B}_{\text{sw}}$. Indeed our work underscores the value of, and need for more, data-derived $\delta^{11}\text{B}_{\text{sw}}$ constraints outside of individual paleoclimatic events.

5.3. Comparison to $^{87/86}\text{Sr}$, $^{187/188}\text{Os}$, and $\delta^7\text{Li}_{\text{sw}}$

Figure 5 shows that there is a broad scale similarity between the temporal evolution of $\delta^{11}\text{B}$, $\delta^7\text{Li}$, $^{87/86}\text{Sr}$ and $^{187/188}\text{Os}$. In particular though, we observe greater similarity between $\delta^{11}\text{B}_{\text{sw}}$ and $^{87/86}\text{Sr}$, and also between pH, $\delta^7\text{Li}$ and $^{187/188}\text{Os}$. The similarity between $^{87/86}\text{Sr}$ and $\delta^{11}\text{B}_{\text{sw}}$ is present in the overall shape of both signals, and in both showing an inflection point at approximately the same time, around 35 Ma. Given $^{87/86}\text{Sr}$ and $\delta^{11}\text{B}_{\text{sw}}$ share many controls (such as weathering and hydrothermal influx), it is not surprising that their signals share some similarities. However we note that $^{87/86}\text{Sr}$ has been increasing continuously since 40 Ma, while our reconstruction of $\delta^{11}\text{B}_{\text{sw}}$ begins to decline at approximately 10 Ma.

We note an striking correspondence between $\delta^{11}\text{B}_4$ and $\delta^7\text{Li}$ (shown in Figures 4 and 5, or together in Figure S1 in Supporting Information S1), in agreement with previous findings (Greenop et al., 2017; Raitzsch & Hönisch, 2013) as well as similarity between $\delta^{11}\text{B}_4$ and $^{187/188}\text{Os}$. Previously the correspondence between $\delta^{11}\text{B}_4$ and $\delta^7\text{Li}$ has been attributed to the overlap in drivers of $\delta^{11}\text{B}_{\text{sw}}$ and $\delta^7\text{Li}$. Under the two component model described in Section 5.1, we propose that similarity of $\delta^{11}\text{B}_4$ and $\delta^7\text{Li}$ could either be due to a relationship between $\delta^7\text{Li}$ and $\delta^{11}\text{B}_{\text{sw}}$ (as previously suggested by Raitzsch and Hönisch (2013)), or due to a relationship between $\delta^7\text{Li}$ and pH, or a combination of both. $\delta^{11}\text{B}_{\text{sw}}$ and $\delta^7\text{Li}$ share drivers, as do pH and $\delta^7\text{Li}$, which are linked to weathering, clay formation, and seafloor spreading. These confounding factors make it difficult to ascribe this correlation to either the $\delta^{11}\text{B}_{\text{sw}}$ or pH signals unambiguously (see Figure S1 in Supporting Information S1). If pH is prescribed to change slowly and linearly (as in Raitzsch and Hönisch (2013)), then, by necessity, $\delta^{11}\text{B}_{\text{sw}}$ will reflect $\delta^{11}\text{B}_4$ —requiring relatively rapid fluctuations in $\delta^{11}\text{B}_{\text{sw}}$. However, as discussed in Section 5.2, we proceed here under the assumption that the rate of change in $\delta^{11}\text{B}_{\text{sw}}$ is as calculated in Lemarchand et al. (2000), removing the possibility of fast changes in $\delta^{11}\text{B}_{\text{sw}}$, and necessitating that $\delta^{11}\text{B}_4$ and pH are tightly correlated. Therefore here, it is pH and $\delta^7\text{Li}$ which have similar trajectories across the Cenozoic.

Notwithstanding the broad scale correlation between $\delta^7\text{Li}$ and $\delta^{11}\text{B}_4$, we note an interesting divergence between these signals at the K-Pg boundary. At this time, there is a large excursion in $\delta^7\text{Li}$ of approximately 5 ‰ in scale (almost as large as all the change which occurs during the rest of the Cenozoic). This excursion occurs at the same time as a large excursion in $^{187/188}\text{Os}$. While there is also a large perturbation in foraminiferal boron isotopes at approximately this time (Henehan et al., 2019), the temporal agreement is poor, and the nature of the two signals is different. $\delta^{11}\text{B}_4$ undergoes an excursion but rapidly recovers to near pre-perturbation levels, whereas $\delta^7\text{Li}$ values stay low for the next ~15 million years. This indicates that the drivers of pH and $\delta^7\text{Li}$ can be decoupled. However, the recorded change in $\delta^7\text{Li}$ at the K-Pg is extremely rapid, faster even than the modern residence time of lithium

in seawater (approximately 1.2 Myr (Misra & Froelich, 2012)). It seems equally plausible that there is alternative, non-seawater, driver of foraminiferal $\delta^7\text{Li}$ at this time. If this alternative control is linked to the carbonate system (as suggested by Vigier et al. (2015); Roberts et al. (2018)), similar effects could conceivably be influencing the correlation between $\delta^{11}\text{B}_4$ and $\delta^7\text{Li}$ more broadly during the Cenozoic. At present, however, studies disagree as to the nature of carbonate system control on foraminiferal $\delta^7\text{Li}$ (Roberts et al., 2018; Vigier et al., 2015).

5.4. Palaeo pH and CO_2

pH is linked to the ocean carbonate system, with a particularly close relationship to atmospheric CO_2 concentrations (Hain et al., 2018). The maximum offset between central estimates of the long term trajectory of the Cenozoic pH reconstruction of J. W. Rae et al. (2021) and our pH (shown in Figure 4) occurs during the Miocene (at approximately 12 Ma) and is 0.15 units in scale. All other factors being equal, a fall in pH of 0.15 units would suggest an increase in atmospheric CO_2 concentration of roughly 50%. Naïvely scaling Miocene CO_2 estimates from J. W. Rae et al. (2021) results in CO_2 concentrations of approximately 750 ppm after the Miocene Climatic Optimum (MCO). During the MCO, the offset between our predicted pH and previous work is slightly smaller, which would suggest (again, assuming all other factors are equal) an increase in atmospheric CO_2 of 45% relative to J. W. Rae et al. (2021) - suggesting an MCO CO_2 of approximately 860ppm. These suggestions of increased CO_2 relative to previous work come with the important caveat that, as mentioned above, the constraints on $\delta^{11}\text{B}_{\text{sw}}$ in this interval are highly variable, perhaps suggesting additional complicating factors which make $\delta^{11}\text{B}_{\text{sw}}$ difficult to estimate at this time. We note that some recent paleotemperature and modeling studies (e.g., Steinhilber et al. (2021) and Herbert et al. (2022)) have made the suggestion that Miocene CO_2 might be higher than estimated in earlier reconstructions, with Burls et al. (2021) finding aspects of improved model-data agreement at CO_2 of around 850 ppm. Our reconstruction of $\delta^{11}\text{B}_{\text{sw}}$ (and consequently pH) has high overall uncertainties during the Neogene to reflect this. We encourage future work to provide additional constraints in this interval and across the Cenozoic.

Second, uncertainties remain in estimates of the second carbonate system parameter in the Miocene and throughout the Cenozoic. The difference in estimated CO_2 from this work and J. W. Rae et al. (2021) may be partially or wholly ameliorated by changing our expectations of the second carbonate system parameter. In this case, the suggested higher atmospheric CO_2 could be averted by a reduction in DIC relative to previous estimates—or the reality could lie somewhere in between, with atmospheric CO_2 concentration mildly elevated compared to previous reconstructions, and DIC mildly lowered. Understanding of possible ocean DIC at this time is mostly derived from carbon cycle box models, supplemented by suggestions from the B/Ca proxy (Sosdian et al., 2018). The range in DIC estimates is from roughly 1,200 to 2500 $\mu\text{mol/kg}$, meaning we are unable to disambiguate whether this record is truly indicative of higher CO_2 concentrations at this time, or lower DIC concentrations, or a combination of both.

6. Benefits of the Gaussian Process Approach

The Gaussian Process methodology allows us to integrate data constraints with limitations on the rate of change in $\delta^{11}\text{B}_{\text{sw}}$ from modeling, and (as described in Section 3) by tweaking the standard approach we are able to incorporate constraints with non-Gaussian uncertainty structures. Uncertainty in the reconstruction itself behaves intuitively, such that the spread in the reconstruction is guided by uncertainties in the data where available, and grows larger with increasing separation from data constraints. The shape of the reconstruction is not specified a priori, as would be the case with parametric fits. Instead, the reconstruction can take on almost any shape as guided by the data constraints and chosen hyperparameters, allowing us to model arbitrary shapes in the evolution of $\delta^{11}\text{B}_{\text{sw}}$ —and the other isotope systems reconstructed here (Figure 5).

While the standard approach to fitting a Gaussian process (constrained by data with Gaussian uncertainties) can directly predict the mean and variance of the signal being reconstructed, our approach is reliant on drawing samples from the Gaussian Process and then filtering them to adjust the posterior prediction. Once this process is complete we have 10,000 possible time series which are plausible evolutions in $\delta^{11}\text{B}_{\text{sw}}$ over the last 65 Myr. The 10,000 possible evolutions can be summarized by their mean, median, and/or 95% confidence interval, but retaining each of the possibilities makes it possible to propagate uncertainties into future data products. For instance, uncertainty in palaeo pH can be propagated by using a Monte Carlo approach whereby each of these evolutions is sampled alongside other required parameters to provide 10,000 possible evolutions of pH. Keeping

each possible evolution maintains the embedded covariance structure, meaning that it is possible to calculate derivative properties (such as the change in $\delta^{11}\text{B}_{\text{sw}}$ or pH between one time and another), which would not be possible to do from metrics such as the mean and standard deviation. This is particularly beneficial in the context of reconstructing the palaeo carbonate system, as trends in parameters are often more robust or important than their absolute value. For instance, for short time windows (relative to the residence time of boron), although we may not know the absolute value of $\delta^{11}\text{B}_{\text{sw}}$, we believe that it can not have changed substantially. Combining the $\delta^{11}\text{B}_{\text{sw}}$ statistical samples generated here with a Monte Carlo approach allows uncertainty to be propagated in such a way as to explore the full range of absolute values for $\delta^{11}\text{B}_{\text{sw}}$ while each sample preserves a reasonable $\Delta\delta^{11}\text{B}_{\text{sw}}$ (see Tierney et al. (2022) for an example of how an analogous approach was used to constrain change in atmospheric CO_2 concentration). As the 10,000 possible evolutions of $\delta^{11}\text{B}_{\text{sw}}$ that we provide are considered equally likely, it also allows further filtering. For instance if looking at a particular time period, or when new information comes to light, which means we are able to be more certain about the rate of change in $\delta^{11}\text{B}_{\text{sw}}$, then the samples provided here can be refiltered to enforce the more restrictive condition—though naturally this will result in a decreased number of valid samples and weakened statistical power.

In summary, we believe the Gaussian Process approach provides benefits both in being able to provide a holistic representation of our understanding of $\delta^{11}\text{B}_{\text{sw}}$ from data and modeling, and also in terms of producing results which facilitate more sophisticated forms of onward uncertainty propagation.

7. Limitations

The Gaussian Process methodology used in this work has many properties that make it ideal for geochemical data interpolation, but also has a few caveats. The first is that, in the standard approach, input data constraints are expected to be Gaussian distributions. As discussed above and in Supporting Information S1, we adjust the standard Gaussian Process fitting methodology by using a rejection sampling strategy to incorporate other types of constraint.

The second limitation of the Gaussian Process approach is more fundamental. The two hyperparameters which tune the fit describe the length scale and noise scale, as described above (Section 3). Here, we use the residence time of each element as the length scale of the Gaussian Process, however those concepts are not identical. In particular, as most signals we are reconstructing here are isotope ratios, the concept of elemental residence time is not necessarily directly applicable. Nonetheless, we believe that using the residence time as a guide for the rate of change in these signals is an improvement over using parametric methods, or non-parametric methods with arbitrary smoothing parameters. In particular, for $\delta^{11}\text{B}_{\text{sw}}$, we are not strongly reliant on the assumption that the Gaussian Process length scale is equivalent to the residence time because Lemarchand et al. (2000) provide a direct rate of change estimate which we use as a constraint. Given a residence time for boron of 10 Myr, and a rate of change of $\delta^{11}\text{B}_{\text{sw}}$ of 0.1 ‰/Myr, this implies a noise scale of 1 ‰. However, 1 ‰ is a small range for uncertainty where there are no data constraints. Increasing the noise scale results in rates of change incompatible with rate of change presented in Lemarchand et al. (2000) unless the length scale is commensurately increased—however this is then inconsistent with our understanding of the residence time of boron in seawater. Our solution to this is to use values which permit slightly faster changes than suggested by Lemarchand et al. (2000)—a length scale of 10 Myr, but a noise scale of 2 ‰, then filtering out results which are incompatibly fast.

Using the Gaussian Process with the methodology described here is highly computationally intensive. Each generated sample has only a small chance of being accepted, and it is necessary to try tens of millions of possibilities to achieve 10,000 viable statistical samples. It is inherent in the rejection sampling methodology to be inefficient in this way, and the more criteria that are used (or the more restrictive those criteria are) the less efficient this method becomes. For this use case, our approach takes approximately 40 hr of computational time on a standard desktop machine to generate the requisite 10,000 samples. In future, we may look to alternative statistical techniques to increase our efficiency and allow us to explore a greater range of possibilities, in particular with respect to gradient limitations and signal smoothness.

Our reconstruction of $\delta^{11}\text{B}_{\text{sw}}$ is also limited by our understanding of past ocean conditions. As discussed in Section 1, most constraints on $\delta^{11}\text{B}_{\text{sw}}$ are at some level dependent on models. Typically, a wide range of model conditions were used to predict vertical gradients in $\delta^{13}\text{C}$ and pH, and an even wider range was used in uncertainty propagation to calculate these estimates (see for instance Greenop et al. (2017)). Nonetheless, it would be remiss not to acknowledge that models are a simplification of reality, meaning it is possible that the ocean occupied a

different mode in the past where the vertical gradients in $\delta^{13}\text{C}$ and pH were decoupled or otherwise difficult to predict. Overall, the data constraints in this work are contingent upon carbon cycle model simulations producing realistic ranges for the $\delta^{13}\text{C}$ versus pH gradient, and the validity of assumptions which determine AOU, while our reconstruction itself is dependent on these data and the rate of change determined from modeling of the boron cycle.

8. Conclusions

We provide Cenozoic reconstructions of $\delta^{11}\text{B}_{\text{sw}}$ by collating existing data constraints and integrating these into a Gaussian Process based statistical approach. This allows us to bring together varying types of constraint (including central estimates, lower and upper limits, and other forms of distribution) while rigorously propagating uncertainties. Our results suggest that $\delta^{11}\text{B}_{\text{sw}}$ was slightly higher than previously thought during the Miocene, but are generally in agreement with previous estimates of $\delta^{11}\text{B}_{\text{sw}}$ during the remainder of the Cenozoic. Generally speaking, uncertainties on $\delta^{11}\text{B}_{\text{sw}}$ are approximately 1 ‰, apart from during the Neogene where large uncertainties on data constraints propagate to large uncertainties in our reconstruction. Our higher estimated $\delta^{11}\text{B}_{\text{sw}}$ is indicative of lower Miocene pH than previously thought, though we acknowledge high uncertainties during this time.

We see a notable correspondence between $\delta^{11}\text{B}_4$ and $\delta^7\text{Li}$, something which has previously been used to infer that the evolution of $\delta^{11}\text{B}_{\text{sw}}$ was likely similar to that of $\delta^7\text{Li}$ (Raitzsch & Hönisch, 2013). However the evolution of $\delta^{11}\text{B}_{\text{sw}}$ constrained here shows that the majority of the Cenozoic change in $\delta^{11}\text{B}_4$ was driven by pH, indicating that links between the controls on pH and $\delta^7\text{Li}$ are perhaps the more important control on the similar trajectories of $\delta^{11}\text{B}_4$ and $\delta^7\text{Li}$.

Uncertainties in our reconstruction are largest where boron isotope data are sparse, such as during the Oligocene, or where datapoints are in disagreement with one another, such as during the Neogene, and we encourage the generation of future records to target these intervals. Looking forward, it would undoubtedly be helpful to find more direct proxies for $\delta^{11}\text{B}_{\text{sw}}$, and to improve constraints on the various models which guide both the data constraints and limits on the rate of change in $\delta^{11}\text{B}_{\text{sw}}$. However, using this method we are able to constrain $\delta^{11}\text{B}_{\text{sw}}$ to a range of ± 1 ‰ across most of the Cenozoic, improving current estimates. Results from this study can be used to propagate uncertainties in $\delta^{11}\text{B}_{\text{sw}}$ into future reconstructions of palaeo pH from boron isotopes and, by extension, palaeo CO_2 . We provide both the metrics which describe our reconstruction of $\delta^{11}\text{B}_{\text{sw}}$ (the median, and 95% confidence interval), and also 10,000 statistical samples of possible evolutions of $\delta^{11}\text{B}_{\text{sw}}$. Uncertainty in $\delta^{11}\text{B}_{\text{sw}}$ can then be propagated into future data products using a Monte Carlo approach as described in Section 6. In addition we provide analogous information for all signals reconstructed here using the Gaussian Process ($\delta^{11}\text{B}_4$, pH, $^{87/86}\text{Sr}$, $^{187/188}\text{Os}$, and $\delta^7\text{Li}$).

Data Availability Statement

All code used in and produced by this project is stored within our GitHub Repository: <https://github.com/St-Andrews-Isotope-Geochemistry/d11Bsw-Gaussian-Process> and archived using Zenodo (Whiteford, 2024b). Data files associated with this project are archived using figshare (Whiteford, 2024a). Three data files are provided which include:

- Data output for $\delta^{11}\text{B}_{\text{sw}}$ in the form of a .xlsx file, which contains both summary metrics and all 10,000 time series realizations.
- Original data, metrics summary of the reconstruction, and 10,000 individual statistical samples for $^{87/86}\text{Sr}$, $^{187/188}\text{Os}$, and $\delta^7\text{Li}$ in a .xlsx file.
- $\delta^{11}\text{B}_4$ data input to the Gaussian Process, metrics summary of the reconstruction, and individual statistical samples for $\delta^{11}\text{B}_4$ and pH in a .xlsx file.

Two forms of output for $\delta^{11}\text{B}_{\text{sw}}$ are given, because while the median and 95% confidence interval give a sense of reasonable values and allow easy plotting, they are unable to convey the covariance embedded in each Gaussian Process sample. Thus for uncertainty propagation in future calculations, we recommend using the time series contained within the .xlsx file. This methodology allows propagation of both uncertainties in $\delta^{11}\text{B}_{\text{sw}}$ and also the rate of change of $\delta^{11}\text{B}_{\text{sw}}$ as described in Section 6. Data for $\delta^{11}\text{B}_4$ can be found as a supplementary table to J. W.

Rae et al. (2021). Software used in the project is written in Python, and is available on our GitHub repository: <https://github.com/St-Andrews-Isotope-Geochemistry/d1IBsw-Gaussian-Process> (Whiteford, 2024b). It uses a software package we've written to perform the statistical calculations, in particular representing distributions, drawing samples, and performing the Gaussian Process interpolation. Scripts to perform the calculation, analyze the output, and display the results are also included.

Acknowledgments

We are grateful to PAGES for their support of boron isotope workshops which facilitated the collaboration resulting in this paper. We also thank the organizers of this special issue.

References

- Anagnostou, E., John, E. H., Babila, T. L., Sexton, P. F., Ridgwell, A., Lunt, D. J., et al. (2020). Proxy evidence for state-dependence of climate sensitivity in the Eocene greenhouse. *Nature Communications*, *11*(1), 4436. <https://doi.org/10.1038/s41467-020-17887-x>
- Anagnostou, E., John, E. H., Edgar, K. M., Foster, G. L., Ridgwell, A., Inglis, G. N., et al. (2016). Changing atmospheric CO₂ concentration was the primary driver of early Cenozoic climate. *Nature*, *533*(7603), 380–384. <https://doi.org/10.1038/nature17423>
- Badger, M. P. S., Lear, C. H., Pancost, R. D., Foster, G. L., Bailey, T. R., Leng, M. J., & Abels, H. A. (2013). CO₂ drawdown following the middle Miocene expansion of the Antarctic Ice Sheet. *Paleoceanography*, *28*(1), 42–53. <https://doi.org/10.1002/palo.20015>
- Broecker, W. S., & Peng, T.-H. (1982). *Tracers in the Sea*. Lamont-Doherty Geological Observatory, Columbia University.
- Burls, N. J., Bradshaw, C. D., De Boer, A. M., Herold, N., Huber, M., Pound, M., et al. (2021). Simulating Miocene warmth: Insights from an opportunistic multi-model ensemble (MioMIP1). *Paleoceanography and Paleoclimatology*, *36*(5), e2020PA004054. <https://doi.org/10.1029/2020PA004054>
- Caves Rugenstein, J. K., Ibarra, D. E., & von Blanckenburg, F. (2019). Neogene cooling driven by land surface reactivity rather than increased weathering fluxes. *Nature*, *571*(7763), 99–102. <https://doi.org/10.1038/s41586-019-1332-y>
- Chalk, T. B., Hain, M. P., Foster, G. L., Rohling, E. J., Sexton, P. F., Badger, M. P. S., et al. (2017). Causes of ice age intensification across the Mid-Pleistocene Transition. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(50), 13114–13119. <https://doi.org/10.1073/pnas.1702143114>
- de la Vega, E., Chalk, T. B., Wilson, P. A., Bysani, R. P., & Foster, G. L. (2020). Atmospheric CO₂ during the Mid-Piacenzian warm period and the M2 glaciation. *Scientific Reports*, *10*(1), 11002. <https://doi.org/10.1038/s41598-020-67154-8>
- Derry, L. A. (2022). Carbonate weathering, CO₂ redistribution, and Neogene CCD and pCO₂ evolution. *Earth and Planetary Science Letters*, *597*, 117801. <https://doi.org/10.1016/j.epsl.2022.117801>
- Dickson, A., & Goyet, C. (1994). Handbook of methods for the analysis of the various parameters of the carbon dioxide system in sea water; version 2. *DOE*. <https://doi.org/10.2172/10107773>
- Dyez, K. A., Hönisch, B., & Schmidt, G. A. (2018). Early Pleistocene obliquity-scale pCO₂ variability at 1.5 million years ago. *Paleoceanography and Paleoclimatology*, *33*(11), 1270–1291. <https://doi.org/10.1029/2018PA003349>
- Findlay, H., Artioli, Y., Birchenough, S., Hartman, S., Leon, P., & Stiasny, M. (2022). Ocean acidification around the UK and Ireland. *MCCIP Rolling Evidence Updates*, *24*. <https://doi.org/10.14465/2022.REU03.OAC>
- Foster, G. L. (2008). Seawater pH, pCO₂ and [CO₂⁻³] variations in the Caribbean Sea over the last 130 kyr: A boron isotope and B/Ca study of planktic foraminifera. *Earth and Planetary Science Letters*, *271*(1), 254–266. <https://doi.org/10.1016/j.epsl.2008.04.015>
- Foster, G. L., Lear, C. H., & Rae, J. W. B. (2012). The evolution of pCO₂, ice volume and climate during the middle Miocene. *Earth and Planetary Science Letters*, *341–344*, 243–254. <https://doi.org/10.1016/j.epsl.2012.06.007>
- Foster, G. L., Pogge von Strandmann, P. A. E., & Rae, J. W. B. (2010). Boron and magnesium isotopic composition of seawater. *Geochemistry, Geophysics, Geosystems*, *11*(8), Q08015. <https://doi.org/10.1029/2010GC003201>
- Greenop, R., Foster, G. L., Wilson, P. A., & Lear, C. H. (2014). Middle Miocene climate instability associated with high-amplitude CO₂ variability. *Paleoceanography*, *29*(9), 845–853. <https://doi.org/10.1002/2014PA002653>
- Greenop, R., Hain, M. P., Sosdian, S. M., Oliver, K. I. C., Goodwin, P., Chalk, T. B., et al. (2017). A record of Neogene seawater δ¹¹B reconstructed from paired δ¹¹B analyses on benthic and planktic foraminifera. *Climate of the Past*, *13*(2), 149–170. <https://doi.org/10.5194/cp-13-149-2017>
- Greenop, R., Sosdian, S. M., Henehan, M. J., Wilson, P. A., Lear, C. H., & Foster, G. L. (2019). Orbital forcing, ice volume, and CO₂ across the Oligocene-Miocene transition. *Paleoceanography and Paleoclimatology*, *34*(3), 316–328. <https://doi.org/10.1029/2018PA003420>
- Gutjahr, M., Ridgwell, A., Sexton, P. F., Anagnostou, E., Pearson, P. N., Pälike, H., et al. (2017). Very large release of mostly volcanic carbon during the Palaeocene–Eocene Thermal Maximum. *Nature*, *548*(7669), 573–577. <https://doi.org/10.1038/nature23646>
- Hain, M. P., Foster, G. L., & Chalk, T. (2018). Robust constraints on past CO₂ climate forcing from the boron isotope proxy. *Paleoceanography and Paleoclimatology*, *33*(10), 1099–1115. <https://doi.org/10.1029/2018PA003362>
- Hain, M. P., Sigman, D. M., & Haug, G. H. (2010). Carbon dioxide effects of Antarctic stratification, North Atlantic Intermediate Water formation, and subantarctic nutrient drawdown during the last ice age: Diagnosis and synthesis in a geochemical box model. *Global Biogeochemical Cycles*, *24*(4), GB4023. <https://doi.org/10.1029/2010GB003790>
- Hain, M. P., Sigman, D. M., Higgins, J. A., & Haug, G. H. (2015). The effects of secular calcium and magnesium concentration changes on the thermodynamics of seawater acid/base chemistry: Implications for Eocene and Cretaceous ocean carbon chemistry and buffering. *Global Biogeochemical Cycles*, *29*(5), 517–533. <https://doi.org/10.1002/2014GB004986>
- Harper, D. T., Hönisch, B., Zeebe, R. E., Shaffer, G., Haynes, L. L., Thomas, E., & Zachos, J. C. (2020). The magnitude of surface ocean acidification and carbon release during Eocene thermal maximum 2 (ETM-2) and the Paleocene-Eocene thermal maximum (PETM). *Paleoceanography and Paleoclimatology*, *35*(2), e2019PA003699. <https://doi.org/10.1029/2019PA003699>
- Henehan, M. J., Edgar, K. M., Foster, G. L., Penman, D. E., Hull, P. M., Greenop, R., et al. (2020). Revisiting the middle Eocene climatic optimum “carbon cycle conundrum” with new estimates of atmospheric pCO₂ from boron isotopes. *Paleoceanography and Paleoclimatology*, *35*(6), e2019PA003713. <https://doi.org/10.1029/2019PA003713>
- Henehan, M. J., Ridgwell, A., Thomas, E., Zhang, S., Alegret, L., Schmidt, D. N., et al. (2019). Rapid ocean acidification and protracted Earth system recovery followed the end-Cretaceous Chicxulub impact. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(45), 22500–22504. <https://doi.org/10.1073/pnas.1905989116>
- Henehan, M. J., & Witts, J. D. (2023). Continental flood basalts do not drive later Phanerozoic extinctions. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(21), e2303700120. <https://doi.org/10.1073/pnas.2303700120>
- Herbert, T. D., Dalton, C. A., Liu, Z., Salazar, A., Si, W., & Wilson, D. S. (2022). Tectonic degassing drove global temperature trends since 20 Ma. *Science*, *377*(6601), 116–119. <https://doi.org/10.1126/science.abl4353>

- Hönisch, B., Hemming, N. G., Archer, D., Siddall, M., & McManus, J. F. (2009). Atmospheric carbon dioxide concentration across the Mid-Pleistocene transition. *Science*, *324*(5934), 1551–1554. <https://doi.org/10.1126/science.1171477>
- Hönisch, B., Royer, D. L., Breecker, D. O., Polissar, P. J., Bowen, G. J., Henehan, M. J., et al. (2023). Toward a Cenozoic history of atmospheric CO₂. *Science*, *382*(6675), eadi5177. <https://doi.org/10.1126/science.adi5177>
- Hull, P. M., Osborn, K. J., Norris, R. D., & Robison, B. H. (2011). Seasonality and depth distribution of a mesopelagic foraminifer, *Hastigerinella digitata*, in Monterey Bay, California. *Limnology & Oceanography*, *56*(2), 562–576. <https://doi.org/10.4319/lo.2011.56.2.0562>
- Josso, P., Parkinson, I., Horstwood, M., Lusty, P., Chenery, S., & Murton, B. (2019). Improving confidence in ferromanganese crust age models: A composite geochemical approach. *Chemical Geology*, *513*, 108–119. <https://doi.org/10.1016/j.chemgeo.2019.03.003>
- Klemm, V., Levasseur, S., Frank, M., Hein, J. R., & Halliday, A. N. (2005). Osmium isotope stratigraphy of a marine ferromanganese crust. *Earth and Planetary Science Letters*, *238*(1), 42–48. <https://doi.org/10.1016/j.epsl.2005.07.016>
- Klochko, K., Kaufman, A. J., Yao, W., Byrne, R. H., & Tossell, J. A. (2006). Experimental measurement of boron isotope fractionation in seawater. *Earth and Planetary Science Letters*, *248*(1), 276–285. <https://doi.org/10.1016/j.epsl.2006.05.034>
- Lemarchand, D., Gaillardet, J., Lewin, É., & Allègre, C. J. (2000). The influence of rivers on marine boron isotopes and implications for reconstructing past ocean pH. *Nature*, *408*(6815), 951–954. <https://doi.org/10.1038/35050058>
- Marschall, H. & Foster, G. (Eds.) (2018). *Boron isotopes*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-64666-4>
- Martínez-Botí, M. A., Foster, G. L., Chalk, T. B., Rohling, E. J., Sexton, P. F., Lunt, D. J., et al. (2015). Plio-Pleistocene climate sensitivity evaluated using high-resolution CO₂ records. *Nature*, *518*(7537), 49–54. <https://doi.org/10.1038/nature14145>
- Misra, S., & Froelich, P. N. (2012). Lithium isotope history of Cenozoic seawater: Changes in silicate weathering and reverse weathering. *Science*, *335*(6070), 818–823. <https://doi.org/10.1126/science.1214697>
- Nir, O., Vengosh, A., Harkness, J. S., Dwyer, G. S., & Lahav, O. (2015). Direct measurement of the boron isotope fractionation factor: Reducing the uncertainty in reconstructing ocean paleo-pH. *Earth and Planetary Science Letters*, *414*, 1–5. <https://doi.org/10.1016/j.epsl.2015.01.006>
- Oxburgh, R. (1998). Variations in the osmium isotope composition of sea water over the past 200,000 years. *Earth and Planetary Science Letters*, *159*(3), 183–191. [https://doi.org/10.1016/S0012-821X\(98\)00057-0](https://doi.org/10.1016/S0012-821X(98)00057-0)
- Oxburgh, R. (2001). Residence time of osmium in the oceans. *Geochemistry, Geophysics, Geosystems*, *2*(6), 1018. <https://doi.org/10.1029/2000GC000104>
- Oxburgh, R., Pierson-Wickmann, A.-C., Reisberg, L., & Hemming, S. (2007). Climate-correlated variations in seawater ¹⁸⁷Os/¹⁸⁸Os over the past 200,000 yr: Evidence from the Cariaco Basin, Venezuela. *Earth and Planetary Science Letters*, *263*(3), 246–258. <https://doi.org/10.1016/j.epsl.2007.08.033>
- Palmer, M. R., Pearson, P. N., & Cobb, S. J. (1998). Reconstructing past ocean pH-depth profiles. *Science*, *282*(5393), 1468–1471. <https://doi.org/10.1126/science.282.5393.1468>
- Paquay, F. S., Ravizza, G., & Coccioni, R. (2014). The influence of extraterrestrial material on the late Eocene marine Os isotope record. *Geochimica et Cosmochimica Acta*, *144*, 238–257. <https://doi.org/10.1016/j.gca.2014.08.024>
- Paquay, F. S., Ravizza, G. E., Dalai, T. K., & Peucker-Ehrenbrink, B. (2008). Determining chondritic impactor size from the marine osmium isotope record. *Science*, *320*(5873), 214–218. <https://doi.org/10.1126/science.1152860>
- Paris, G., Gaillardet, J., & Louvat, P. (2010). Geological evolution of seawater boron isotopic composition recorded in evaporites. *Geology*, *38*(11), 1035–1038. <https://doi.org/10.1130/G31321.1>
- Park, H., & Schlesinger, W. H. (2002). Global biogeochemical cycle of boron. *Global Biogeochemical Cycles*, *16*(4), 20–1–20–11. <https://doi.org/10.1029/2001GB001766>
- Pearson, P. N., Foster, G. L., & Wade, B. S. (2009). Atmospheric carbon dioxide through the Eocene–Oligocene climate transition. *Nature*, *461*(7267), 1110–1113. <https://doi.org/10.1038/nature08447>
- Pegram, W. J., & Turekian, K. K. (1999). The osmium isotopic composition change of Cenozoic sea water as inferred from a deep-sea core corrected for meteoritic contributions. *Geochimica et Cosmochimica Acta*, *63*(23), 4053–4058. [https://doi.org/10.1016/S0016-7037\(99\)00308-7](https://doi.org/10.1016/S0016-7037(99)00308-7)
- Penman, D. E., Hönisch, B., Zeebe, R. E., Thomas, E., & Zachos, J. C. (2014). Rapid and sustained surface ocean acidification during the Paleocene-Eocene thermal maximum. *Paleoceanography*, *29*(5), 357–369. <https://doi.org/10.1002/2014PA002621>
- Peucker-Ehrenbrink, B., & Ravizza, G. (2000). The effects of sampling artifacts on cosmic dust flux estimates: A reevaluation of nonvolatile tracers (Os, Ir). *Geochimica et Cosmochimica Acta*, *64*(11), 1965–1970. [https://doi.org/10.1016/S0016-7037\(99\)00429-9](https://doi.org/10.1016/S0016-7037(99)00429-9)
- Rae, J. W., Zhang, Y. G., Liu, X., Foster, G. L., Stoll, H. M., & Whiteford, R. D. (2021). Atmospheric CO₂ over the Past 66 million years from marine archives. *Annual Review of Earth and Planetary Sciences*, *49*(1), 609–641. <https://doi.org/10.1146/annurev-earth-082420-063026>
- Rae, J. W. B. (2018). Boron isotopes in foraminifera: Systematics, biomineralisation, and CO₂ reconstruction. In H. Marschall & G. Foster (Eds.), *Boron isotopes: The fifth element* (pp. 107–143). Springer International Publishing. https://doi.org/10.1007/978-3-319-64666-4_5
- Raitzsch, M., Hain, M., Henehan, M., & Gattuso, J.-P. (2022). Seacarb x - Seacarb extension for deep-time carbonate system calculations. *Zenodo*. <https://doi.org/10.5281/zenodo.5909811>
- Raitzsch, M., & Hönisch, B. (2013). Cenozoic boron isotope variations in benthic foraminifers. *Geology*, *41*(5), 591–594. <https://doi.org/10.1130/G34031.1>
- Ravizza, G. (1993). Variations of the ¹⁸⁷Os/¹⁸⁶Os ratio of seawater over the past 28 million years as inferred from metalliferous carbonates. *Earth and Planetary Science Letters*, *118*(1), 335–348. [https://doi.org/10.1016/0012-821X\(93\)90177-B](https://doi.org/10.1016/0012-821X(93)90177-B)
- Ravizza, G., & Peucker-Ehrenbrink, B. (2003a). Chemostratigraphic evidence of Deccan volcanism from the marine osmium isotope record. *Science*, *302*(5649), 1392–1395. <https://doi.org/10.1126/science.1089209>
- Ravizza, G., & Peucker-Ehrenbrink, B. (2003b). The marine ¹⁸⁷Os/¹⁸⁸Os record of the Eocene–Oligocene transition: The interplay of weathering and glaciation. *Earth and Planetary Science Letters*, *210*(1), 151–165. [https://doi.org/10.1016/S0012-821X\(03\)00137-7](https://doi.org/10.1016/S0012-821X(03)00137-7)
- Ravizza, G., & Turekian, K. K. (1992). The osmium isotopic composition of organic-rich marine sediments. *Earth and Planetary Science Letters*, *110*(1), 1–6. [https://doi.org/10.1016/0012-821X\(92\)90034-S](https://doi.org/10.1016/0012-821X(92)90034-S)
- Reusch, D. N., Ravizza, G., Maasch, K. A., & Wright, J. D. (1998). Miocene seawater ¹⁸⁷Os/¹⁸⁸Os ratios inferred from metalliferous carbonates. *Earth and Planetary Science Letters*, *160*(1), 163–178. [https://doi.org/10.1016/S0012-821X\(98\)00082-X](https://doi.org/10.1016/S0012-821X(98)00082-X)
- Ridgwell, A., Hargreaves, J. C., Edwards, N. R., Annan, J. D., Lenton, T. M., Marsh, R., et al. (2007). Marine geochemical data assimilation in an efficient Earth System Model of global biogeochemical cycling. *Biogeosciences*, *4*(1), 87–104. <https://doi.org/10.5194/bg-4-87-2007>
- Roberts, J., Kaczmarek, K., Langer, G., Skinner, L. C., Bijma, J., Bradbury, H., et al. (2018). Lithium isotopic composition of benthic foraminifera: A new proxy for paleo-pH reconstruction. *Geochimica et Cosmochimica Acta*, *236*, 336–350. <https://doi.org/10.1016/j.gca.2018.02.038>
- Robinson, N., Ravizza, G., Coccioni, R., Peucker-Ehrenbrink, B., & Norris, R. (2009). A high-resolution marine ¹⁸⁷Os/¹⁸⁸Os record for the late Maastrichtian: Distinguishing the chemical fingerprints of Deccan volcanism and the KP impact event. *Earth and Planetary Science Letters*, *281*(3), 159–168. <https://doi.org/10.1016/j.epsl.2009.02.019>

- Sosdian, S. M., Greenop, R., Hain, M. P., Foster, G. L., Pearson, P. N., & Lear, C. H. (2018). Constraining the evolution of Neogene ocean carbonate chemistry using the boron isotope pH proxy. *Earth and Planetary Science Letters*, *498*, 362–376. <https://doi.org/10.1016/j.epsl.2018.06.017>
- Steinthorsdottir, M., Coxall, H. K., de Boer, A. M., Huber, M., Barbolini, N., Bradshaw, C. D., et al. (2021). The Miocene: The future of the past. *Paleoceanography and Paleoclimatology*, *36*(4), e2020PA004037. <https://doi.org/10.1029/2020PA004037>
- Stoffynegli, P., & Mackenzie, F. T. (1984). Mass balance of dissolved lithium in the oceans. *Geochimica et Cosmochimica Acta*, *48*(4), 859–872. [https://doi.org/10.1016/0016-7037\(84\)90107-8](https://doi.org/10.1016/0016-7037(84)90107-8)
- Subhas, A. V., Dong, S., Naviaux, J. D., Rollins, N. E., Ziveri, P., Gray, W., et al. (2022). Shallow calcium carbonate cycling in the North Pacific Ocean. *Global Biogeochemical Cycles*, *36*(5), e2022GB007388. <https://doi.org/10.1029/2022GB007388>
- Tierney, J. E., Zhu, J., Li, M., Ridgwell, A., Hakim, G. J., Poulsen, C. J., et al. (2022). Spatial patterns of climate change across the Paleocene–Eocene Thermal Maximum. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(42), e2205326119. <https://doi.org/10.1073/pnas.2205326119>
- Torres, M. A., West, A. J., & Li, G. (2014). Sulphide oxidation and carbonate dissolution as a source of CO₂ over geological timescales. *Nature*, *507*(7492), 346–349. <https://doi.org/10.1038/nature13030>
- van der Ploeg, R., Selby, D., Cramwinckel, M. J., Li, Y., Bohaty, S. M., Middelburg, J. J., & Sluijs, A. (2018). Middle Eocene greenhouse warming facilitated by diminished weathering feedback. *Nature Communications*, *9*(1), 2877. <https://doi.org/10.1038/s41467-018-05104-9>
- Vigier, N., Rollion-Bard, C., Levenson, Y., & Erez, J. (2015). Lithium isotopes in foraminifera shells as a novel proxy for the ocean dissolved inorganic carbon (DIC). *Comptes Rendus Geoscience*, *347*(1), 43–51. <https://doi.org/10.1016/j.crte.2014.12.001>
- Whiteford, R. (2024a). Data supplements [Dataset]. *Figshare*. <https://doi.org/10.6084/m9.figshare.24190347>
- Whiteford, R. (2024b). st-Andrews-Isotope-Geochemistry/d11Bsw-Gaussian-process: Publication release [Software]. *Zenodo*. <https://doi.org/10.5281/zenodo.11187691>
- Zeebe, R. E. (2012). LOSCAR: Long-term Ocean-atmosphere-Sediment CARbon cycle Reservoir Model v2.0.4. *Geoscientific Model Development*, *5*(1), 149–166. <https://doi.org/10.5194/gmd-5-149-2012>
- Zeebe, R. E., & Wolf-Gladrow, D. A. (2001). *CO₂ in seawater: Equilibrium, kinetics, isotopes* (No. 65). Elsevier.