Posted on 30 Sep 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/essoar.169603519.93470287/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Reconstruction of Cenozoic δ11Bsw Using A Gaussian Process

Ross Whiteford[1], Timothy J Heaton[2], Michael James Henehan[3], Eleni Anagnostou[4], Hana Jurikova[1], Gavin L Foster[5], and James William Buchanan Rae[1]

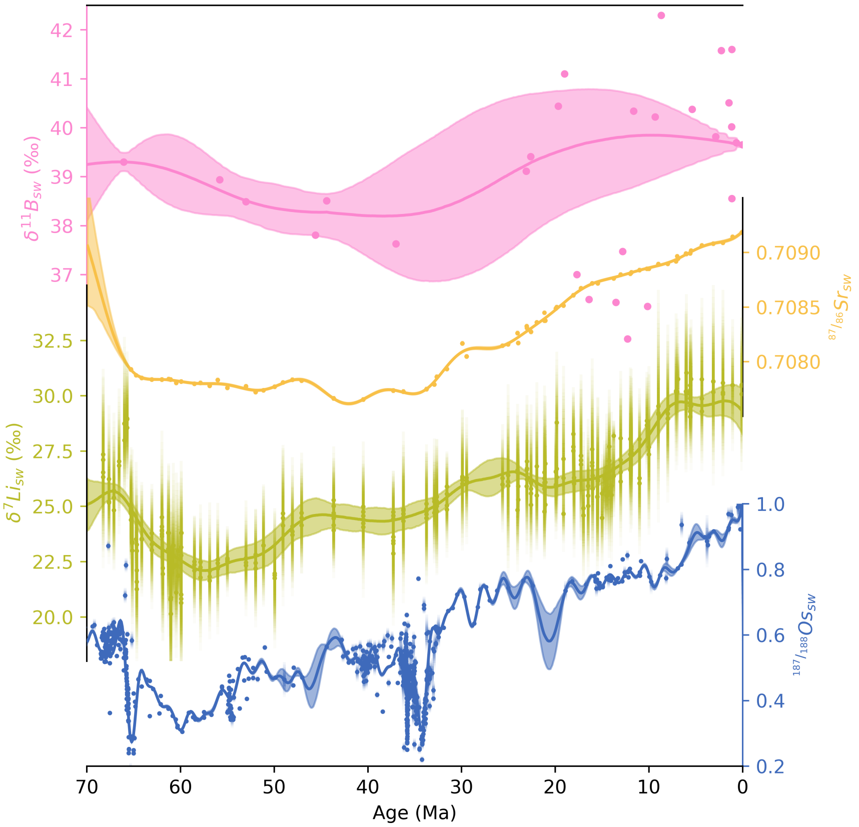
[1]University of St Andrews
[2]University of Leeds
[3]University of Bristol
[4]GEOMAR, Helmholtz Centre for Ocean Research Kiel
[5]University of Southampton

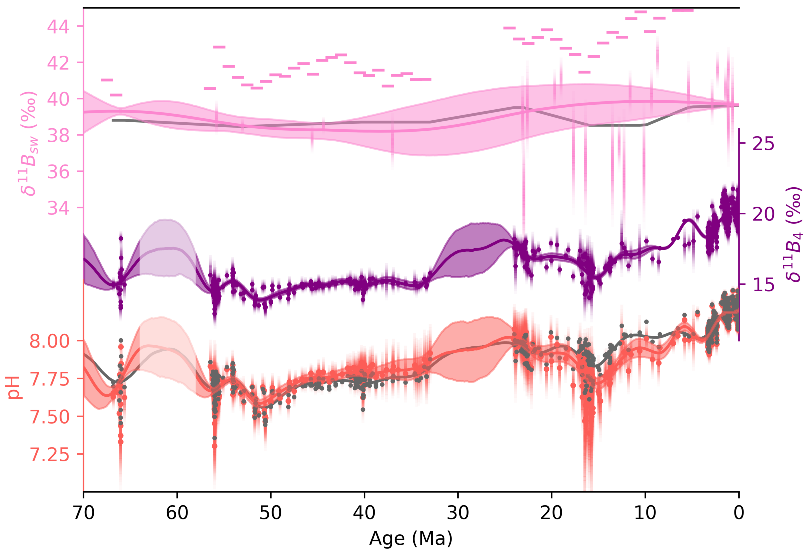
September 30, 2023

Abstract

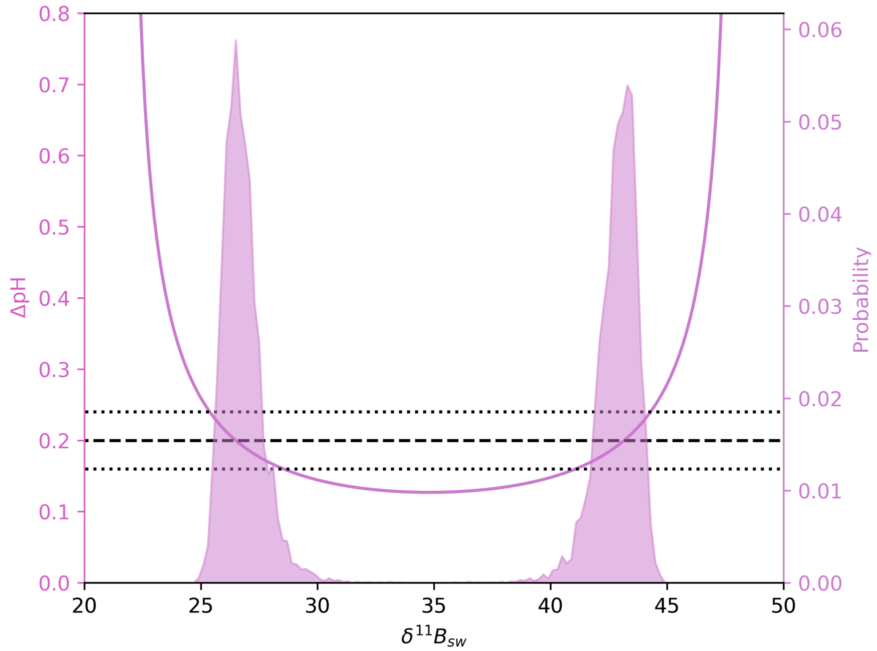
The boron isotope ratio of seawater (δ11Bsw) is a parameter which must be known to reconstruct palaeo pH and CO2 from boron isotope measurements of marine carbonates. Beyond a few million years ago, δ11Bsw is likely to have been different to modern. Palaeo δ11Bsw can be estimated by simultaneously constraining the vertical gradients in foraminiferal δ11B ([?]δ11B) and pH ([?]pH). A number of subtly different techniques have been used to estimate [?]pH 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 estimates alongside limitations on the rate of change of δ11Bsw from modelling, and combine these into 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 δ11Bsw to generate random plausible evolutions of δ11Bsw. We reconstruct δ11Bsw, and by extension palaeo pH, across the last 65Myr using this novel methodology. Reconstructed δ11Bsw is compared to other seawater isotope ratios, namely 87/86Sr, 187/188Os, and δ7Li, which we also reconstruct with Gaussian Processes. Our method provides a template for incorporation of future δ11Bsw constraints, and a mechanism for propagation of uncertainty in δ11Bsw into future studies.

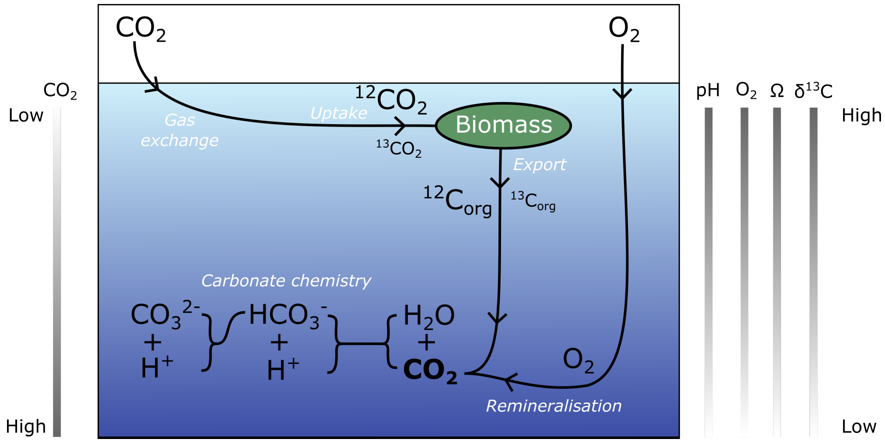
Posted on 30 Sep 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/essoar.169603519.93470287/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.



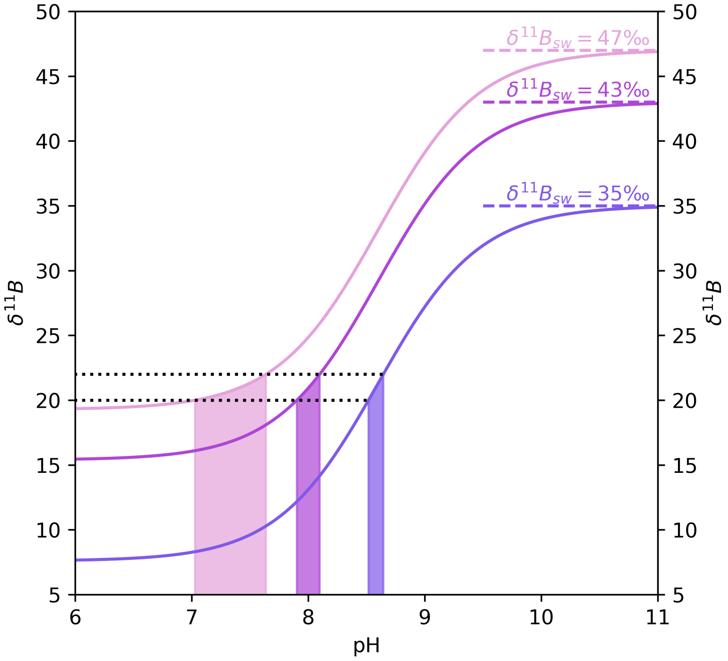


Posted on 30 Sep 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/essoar.169603519.93470287/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

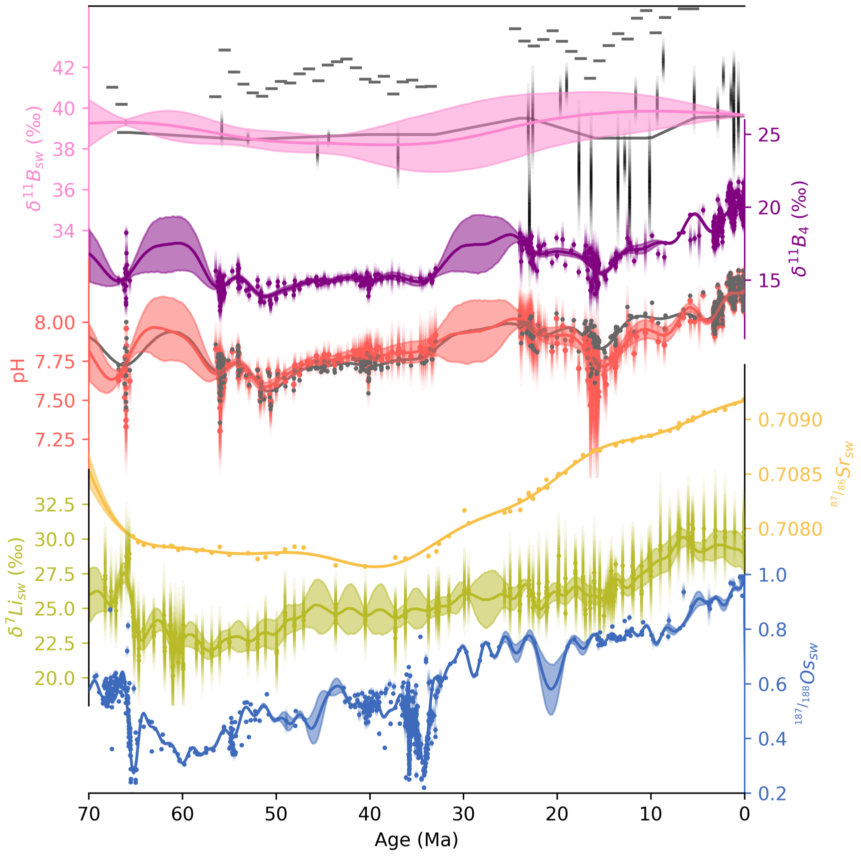


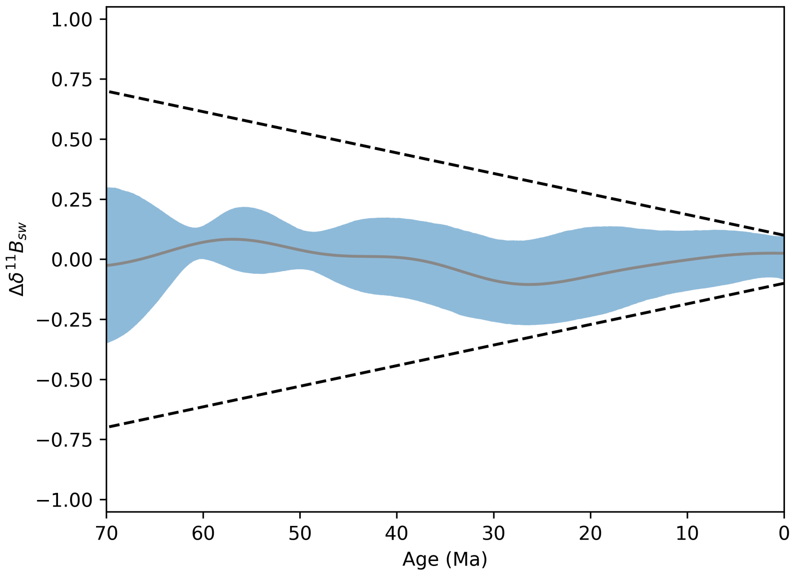

CO_2
O_2
CO_2
Low
High
Gas exchange
Uptake
$^{12}CO_2$
$^{13}CO_2$
Biomass
Export
$^{12}C_{org}$
$^{13}C_{org}$
Carbonate chemistry
CO_3^{2-} + H^+
HCO_3^- + H^+
H_2O + $\mathbf{CO_2}$
O_2
Remineralisation
pH
O_2
Ω
$\delta^{13}C$
High
Low

Posted on 30 Sep 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/essoar.169603519.93470287/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

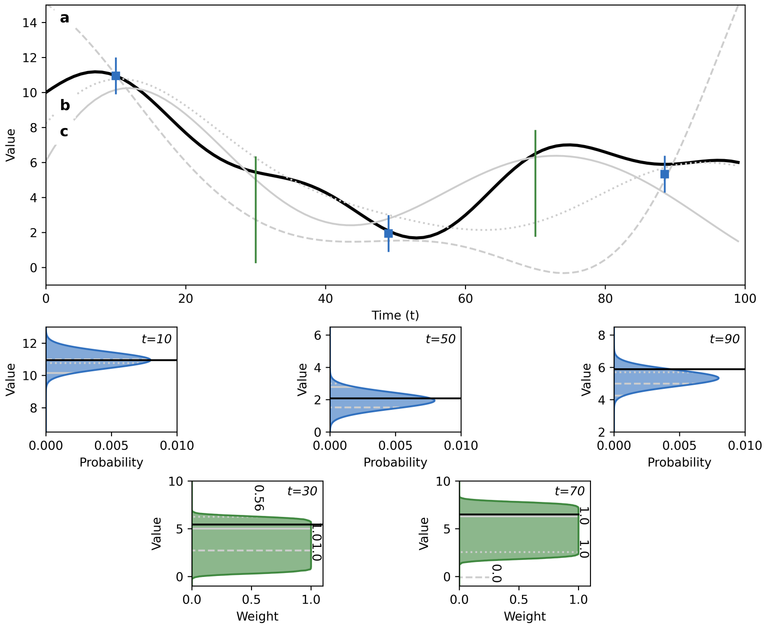


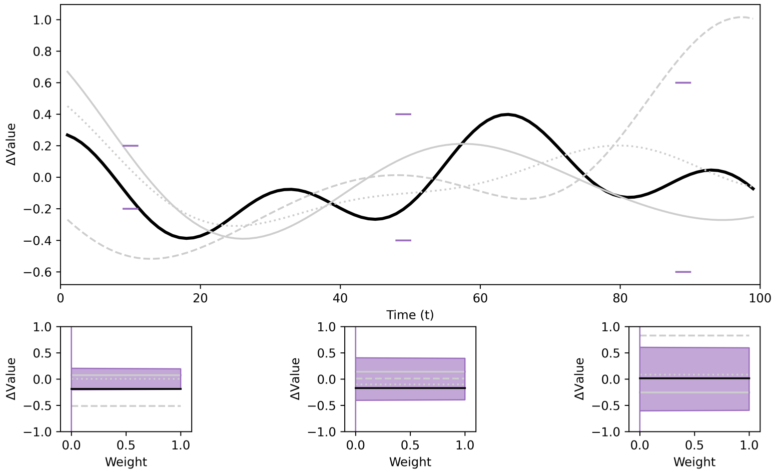
Posted on 30 Sep 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/essoar.169603519.93470287/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.



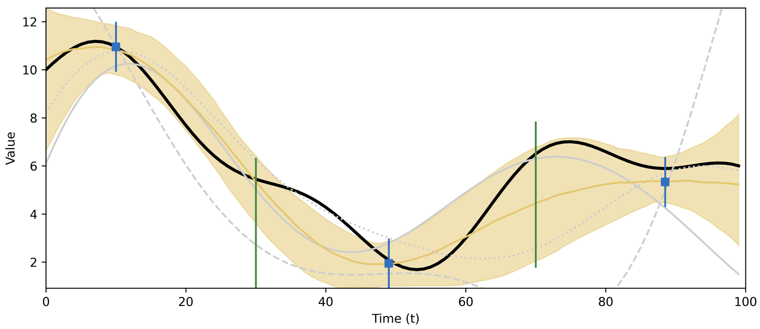


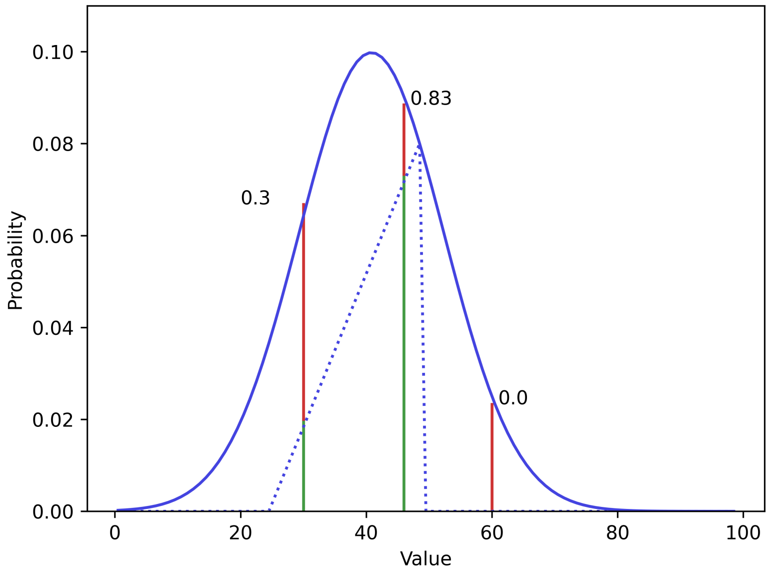
Posted on 30 Sep 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/essoar.169603519.93470287/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.





Posted on 30 Sep 2023 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/essoar.169603519.93470287/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.





Reconstruction of Cenozoic $\delta^{11}B_{sw}$ Using A Gaussian Process

Ross Whiteford[1], Timothy J. Heaton[2], Michael J. Henehan[3], Eleni Anagnostou[4], Hana Jurikova[1], Gavin L. Foster[5], and James W.B. Rae[1]

[1]School of Earth and Environmental Sciences, University of St Andrews, UK
[2]Department of Statistics, School of Mathematics, University of Leeds, UK.
[3]School of Earth Sciences, University of Bristol, Wills Memorial Building, Queens Road, UK
[4]GEOMAR Helmholtz Centre for Ocean Research Kiel
[5]School of Ocean Earth Science, University of 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

Corresponding author: Ross Whiteford, `rdmw1@st-andrews.ac.uk`

Abstract

The boron isotope ratio of seawater ($\delta^{11}B_{sw}$) is a parameter which must be known to reconstruct palaeo pH and CO_2 from boron isotope measurements of marine carbonates. Beyond a few million years ago, $\delta^{11}B_{sw}$ is likely to have been different to modern. Palaeo $\delta^{11}B_{sw}$ can be estimated by simultaneously constraining the vertical gradients in foraminiferal $\delta^{11}B$ ($\Delta\delta^{11}B$) 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 estimates alongside limitations on the rate of change of $\delta^{11}B_{sw}$ from modelling, and combine these into 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 $\delta^{11}B_{sw}$ to generate random plausible evolutions of $\delta^{11}B_{sw}$. We reconstruct $\delta^{11}B_{sw}$, and by extension palaeo pH, across the last 65Myr using this novel methodology. Reconstructed $\delta^{11}B_{sw}$ is compared to other seawater isotope ratios, namely $^{87/86}$Sr, $^{187/188}$Os, and δ^{7}Li, which we also reconstruct with Gaussian Processes. Our method provides a template for incorporation of future $\delta^{11}B_{sw}$ constraints, and a mechanism for propagation of uncertainty in $\delta^{11}B_{sw}$ into future studies.

Plain Language Summary

Boron naturally exists in two forms - ^{11}B and ^{10}B. Measuring the ratio of these two forms of the element 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 $\delta^{11}B_{sw}$). Preexisting studies have estimated $\delta^{11}B_{sw}$ 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 $\delta^{11}B_{sw}$ can change, based on model simulations. We provide statistical samples which can be used in future work when calculating past ocean pH.

1 Introduction

Boron, and its two isotopes (^{10}B and ^{11}B), 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: $\delta^{11}B_4$ (the isotopic composition of borate), $\delta^{11}B_{sw}$ (the isotope composition of seawater), ε (the fractionation factor), pK^*_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 palaeoclimatological application, we wish to determine palaeo pH, which is done by measuring $\delta^{11}B_{calcite}$ and translating this to to $\delta^{11}B_4$ (potentially requiring a species specific calibration). $\delta^{11}B_4$ 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 (pK^*_B - which is estimated from temperature, pressure, and seawater composition (Dickson & Goyet, 1994)), and the boron isotope ratio of seawater ($\delta^{11}B_{sw}$). 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 $\delta^{11}B_{sw}$. Modern day seawater has a $\delta^{11}B_{sw}$ of 39.61±0.04 ‰ (Foster et al., 2010) - however $\delta^{11}B_{sw}$ 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, Tsung-Hung, 1982) - that is, the average amount of time an atom of

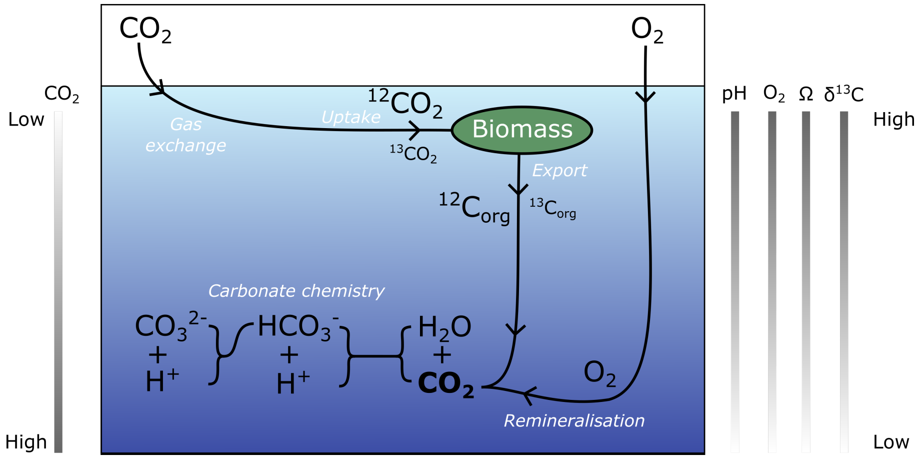


Figure 1. Schematic of the processes which cause vertical gradients in oceanic properties as harnessed in reconstructions of $\delta^{11}B_{sw}$. Oxygen and carbon dioxide concentrations are set in surface waters by the atmospheric concentrations and Henry's Law (which is temperature dependent). Surface 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 remineralised, organic carbon is oxidised (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 $\delta^{13}C$. As $\delta^{11}B_{sw}$ is a well mixed signal, differences in $\delta^{11}B_4$ between the surface and subsurface are primarily driven by the gradient in pH.

boron spends in the ocean. It is estimated from geochemical box modelling of the boron cycle that $\delta^{11}B_{sw}$ might have changed by up to 0.1 ‰/Myr (Lemarchand et al., 2000). This means that beyond a few million years ago, ocean $\delta^{11}B_{sw}$ might have been appreciably different to modern day seawater, which must be taken into account when calculating palaeo pH. While halites show some promise as a direct proxy for $\delta^{11}B_{sw}$ (Paris et al., 2010), uncertainty remains about the observed variability in halite $\delta^{11}B$ reconstructions, so $\delta^{11}B_{sw}$ is currently estimated using more indirect techniques.

The primary technique for estimating, or placing limits on, palaeo $\delta^{11}B_{sw}$ is to exploit vertical gradients in the ocean (Palmer et al., 1998; Greenop et al., 2017; Anagnostou et al., 2016; Henehan et al., 2019). This works due to the sigmoidal relationship between $\delta^{11}B_4$ and pH (as shown in Figure 2). The non-linearity of this relationship makes it possible to calculate $\delta^{11}B_{sw}$ given two estimates of $\delta^{11}B_4$ and an estimate of the change in pH across this $\delta^{11}B_4$ gradient. The vertical gradient in pH (ΔpH) is correlated to the vertical gradients in carbon isotopes ($\Delta\delta^{13}C$), carbon concentration (ΔDIC), and oxygen concentration ($\Delta[O_2]$, linked with Apparent Oxygen Utilisation - AOU) due to their shared controlling processes (see Figure 1). Vertical gradients in all of these parameters can therefore be related to $\delta^{11}B_{sw}$.

The simplest option for reconstructing palaeo $\delta^{11}B_{sw}$ from measured $\delta^{11}B$ in carbonates is to assume a constant ΔpH through time equal to modern, an approach used in Raitzsch and Hönisch (2013). By combining this with the sigmoidal relationship of $\delta^{11}B_4$ and pH, it is possible to estimate $\delta^{11}B_{sw}$ (see Figure 2). To improve upon the assumption of a constant ΔpH through time, one can use foraminiferal $\Delta\delta^{18}O$ and $\Delta\delta^{13}C$ to guide identification of palaeo depth habitats of particular species, and to place reasonable limits on ΔpH (Palmer et al., 1998; Anagnostou et al., 2016; Greenop et al., 2017). For $\delta^{13}C$, this works because $\Delta\delta^{13}C$ and ΔpH are jointly controlled by subsurface remineralisation of organic carbon, which releases isotopically light carbon into the water and causes acidification (decreasing pH) - as illustrated in Figure 1. 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 Rugenstein et al., 2019; Derry, 2022) our understanding of Cenozoic carbon cycle fluxes, and 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 sulphate 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 $\delta^{11}B_{sw}$. Despite these challenges, using model predicted gradients in $\delta^{13}C$ and pH with paired planktic-benthic (or surface-subsurface) foraminferal $\delta^{11}B$ 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 $\delta^{13}C$) and pH.

$\delta^{11}B_{sw}$ 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 $\delta^{11}B$ 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 $\delta^{13}C$, the vertical gradient in oxygen concentration is primarily a function of how much remineralisation of organic carbon has occurred (see Figure 1). Remineralisation 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. 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 $\Delta\delta^{13}C$ 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; Henehan et al., 2019), which simultaneously informs $\delta^{11}B_{sw}$, and palaeo CO_2 reconstructed from boron isotope derived pH.

The requirement of a second carbonate system parameter to predict palaeo CO_2 from boron isotope derived pH presents a challenge due to various limitations in carbonate system reconstructions. For instance, although alkenone carbon isotopes can be used to reconstruct surface water CO_2 (Stoll et al., 2019; Zhang et al., 2020), the pairing of pH and CO_2 to determine the carbonate system tends to lead to large uncertainties due to their close covariance (J. W. B. Rae, 2018). While boron/calcium ratios show promising relationships with carbonate system parameters (including DIC, ΔCO_3^{2-}, and HCO_3^- (Yu & Elderfield, 2007; Haynes et al., 2017; Sosdian et al., 2020)), there is still some debate as to which of these parameters is most faithfully recorded, and Cenozoic-scale records remain fragmentory. Current approaches tend to centre on assumptions linked to the saturation state for calcite (denoted $\Omega_{calcite}$). Saturation state can not have been too low, or carbonate would not have been preserved, nor can it have been too high, or abiotic carbonate would have precipitated. The precise limits on reasonable values of near-surface seawater $\Omega_{calcite}$ are somewhat subjective and are both context and site dependent, but a reasonable broad range is from 1 to 10. In some circumstances, such as when producing estimates for the subtropical surface ocean or targeting background Cenozoic climates,

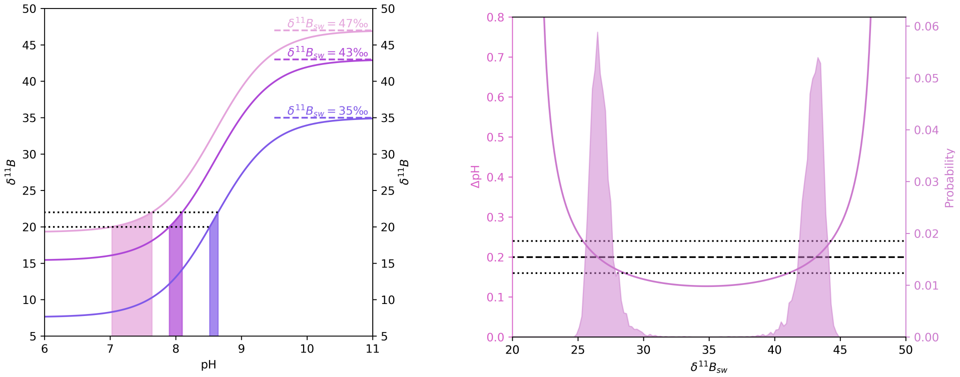


Figure 2. The left panel shows a graph of pH vs $\delta^{11}B_4$ for three different $\delta^{11}B_{sw}$'s (indicated by pink shades). A hypothetical 2‰ excursion from 22‰ to 20‰ is depicted by the black dotted lines. The equivalent pH change for the same excursion at variable $\delta^{11}B_{sw}$ is 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 shows (for the same 22‰ to 20‰ excursion as shown in the left panel) ΔpH as a function of $\delta^{11}B_{sw}$ (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 $\delta^{11}B_{sw}$'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 $\delta^{11}B_{sw}$, aligns with the emplaced constraint. Due to the symmetry of the curves shown in the left panel, there are often two $\delta^{11}B_{sw}$'s compatible with a ΔpH constraint. Typically the lower window is rejected as it would resolve to an unreasonably low pH.

one can be more prescriptive - using a one or two unit range around the modern average surface ocean saturation state of 6 (Ridgwell & Zeebe, 2005; Anagnostou et al., 2016; Boudreau et al., 2019). We note, however, that this assumption may not be valid during transient periods of acidification (e.g. PETM (Penman et al., 2014)) or alkalinity imbalance (e.g. early Danian (Henehan et al., 2019)). Crucially, quantification of carbonate system conditions using the saturation state approach depend on quantification of $[Ca^{2+}]_{sw}$, which is used in combination with the saturation state assumption to derive $[CO_3{}^{2-}]$. For a given assumption or constraint on saturation state, any change in $[Ca^{2+}]_{sw}$ thus translates almost linearly into a change in $[CO_3{}^{2-}]$), meaning that accurately quantifying the concentration of $[Ca^{2+}]_{sw}$, and its uncertainty, is of paramount importance to saturation state based constraints on the past ocean carbonate system.

Overall, it is therefore necessary to balance a number of requirements when reconstructing $\delta^{11}B_{sw}$. The estimated $\delta^{11}B_{sw}$ must produce a reasonable: ΔpH, $\Delta\delta^{13}C$, saturation state and AOU. Finding the space in which all these parameters are viable allows quantification of $\delta^{11}B_{sw}$ with attendant uncertainty. The boron isotope proxy is fortunate however, in that almost all data and assumptions required to estimate $\delta^{11}B_{sw}$ are already a part of the $\delta^{11}B_4$ to CO_2 calculation.

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 $\delta^{11}B_{sw}$ evolved throughout the Cenozoic. Our study represents an advance compared to two particularly relevant previous works that provided Cenozoic timescale estimates of $\delta^{11}B_{sw}$. Raitzsch and Hönisch (2013) predicted Cenozoic $\delta^{11}B_{sw}$ by assuming a constant ΔpH and a linear trend in deep ocean pH, meaning reconstructions of pH using this $\delta^{11}B_{sw}$ are only able to reconstruct the initial assumption (J. W. B. Rae, 2018) (further impacts of which are discussed in Section 5.3). Rae et al. (2021) compiled and renanalysed marine palaeo CO_2 proxy data (including boron isotopes), and therefore required a curve for $\delta^{11}B_{sw}$. However, generation of a Cenozoic $\delta^{11}B_{sw}$ 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 $\delta^{11}B_{sw}$ through time by bringing existing and updated $\delta^{11}B_{sw}$ constraints together into a statistical framework called a Gaussian Process. The Gaussian Process is able to integrate both the expected smoothness of the $\delta^{11}B_{sw}$ signal from modelling work and various forms of constraint from the available data.

2 Data

Data constraints on $\delta^{11}B_{sw}$ are varied. Some previous works presented central estimates of $\delta^{11}B_{sw}$ (Gutjahr et al., 2017; Henehan et al., 2019, 2020), while another provided lower or upper limits on $\delta^{11}B_{sw}$ (Anagnostou et al., 2016), and another provided full probability distributions for possible values of $\delta^{11}B_{sw}$ (Greenop et al., 2017). Two of those studies reconstructed $\delta^{11}B_{sw}$ for a particular event (or short timeslice) (Henehan et al., 2019; Gutjahr et al., 2017), and two reconstructed $\delta^{11}B_{sw}$ over a wider time window, within which $\delta^{11}B_{sw}$ might have evolved (Anagnostou et al., 2016; Greenop et al., 2017). Here we use $\delta^{11}B_{sw}$ estimates from Greenop et al. (2017) exactly as presented 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, and so here we update carbonate system calculations, and hence $\delta^{11}B_{sw}$ 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 the Supplementary Materials, and are plotted in Henehan and Witts (2023). Anagnostou et al. (2016) is updated by modifying the calculation of vital effects to account for changing seawater chemistry. With these recalculations,Henehan

Age	Central Estimate	95% interval	Form	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)*

Table 1. Table of data constraints with representative central values and standard deviations as an indicator of uncertainty. * symbol indicates the values has been updated from the original publication for this work.

et al. (2019)'s $\delta^{11}\mathrm{B}_{\mathrm{sw}}$ estimate is updated from the originally published value of 39.05 - 39.85 ‰ to 39.05 - 39.55 ‰, while estimates from Anagnostou et al. (2016) are integrated as full probability distributions instead of the lower/upper limits presented in the original study (see Table 1 and Supplement S1).

To supplement these existing $\delta^{11}\mathrm{B}_{\mathrm{sw}}$ constraints, we take planktic $\delta^{11}\mathrm{B}_4$ as presented in Rae et al. (2021), based on original studies by Anagnostou et al. (2016, 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); Hönisch et al. (2009); Lemarchand et al. (2000); Martínez-Botí et al. (2015); Paris et al. (2010); Pearson et al. (2009); Penman et al. (2014); Raitzsch and Hönisch (2013); Sosdian et al. (2018), to calculate a range of valid $\delta^{11}\mathrm{B}_{\mathrm{sw}}$ by exploiting the sigmoidal shape of the relationship between pH and $\delta^{11}\mathrm{B}_4$ (as seen in Figure 2). The maximum offset between $\delta^{11}\mathrm{B}_4$ and $\delta^{11}\mathrm{B}_{\mathrm{sw}}$ is seen at low pH, and is described by the fractionation factor ε (or α). For any given estimate of $\delta^{11}\mathrm{B}_4$, the minimum valid $\delta^{11}\mathrm{B}_{\mathrm{sw}}$ is equal to $\delta^{11}\mathrm{B}_4$, and the maximum $\delta^{11}\mathrm{B}_{\mathrm{sw}}$ can be calculated by combination with the fractionation factor. We first bin data into 1 Myr intervals, then take the maximum measured $\delta^{11}\mathrm{B}_4$ as the lower limit for $\delta^{11}\mathrm{B}_{\mathrm{sw}}$, and use the minimum measured $\delta^{11}B_4$ to calculate the maximum $\delta^{11}B_{sw}$. Uncertainty in measured $\delta^{11}B_{foram}$ is propagated to uncertainty in the possible $\delta^{11}B_{sw}$, and the upper 99% quantile is used to estimate the maximum permissible $\delta^{11}B_{sw}$. 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 3.

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 $\delta^{11}B_{sw}$ from modelling.

3 Methods

Having assembled the data described in Section 2, we reconstruct $\delta^{11}B_{sw}$ using a Gaussian Process (GP). The Gaussian Process is a statistical technique which allows us to generate smooth time series conditioned to match data constraints (for a fuller, more rigorous description of the Gaussian Process - see Supplement S1). In the case of $\delta^{11}B_{sw}$, we have an expectation of smoothness from the work of Lemarchand et al. (2000), which we combine with data constraints as described in Section 2 and shown in Figure 3. The Gaussian Process works by using a kernel function, which encodes structure into the reconstruction, and hyperparameters which tune the behaviour. Uncorrelated Gaussian noise is transformed into autocorrelated values with the covariance structure prescribed by the kernel function and hyperparameters. 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 correlation, 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 a time 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, 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. 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 Supplement S1 for a demonstration of this behaviour). 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. 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 $\delta^{11}B_{sw}$, limitations on the rate of change in $\delta^{11}B_{sw}$, and data constraints with a non Gaussian uncertainty structure. Limitations (either on the value of $\delta^{11}B_{sw}$ or the rate of change in $\delta^{11}B_{sw}$) 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 be 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 Supplement 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 $\delta^{11}B_{sw}$ 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 $\delta^{11}B_{sw}$ by conditioning a Gaussian Process on the few available central estimates of $\delta^{11}B_{sw}$ (shown in Table 1). Additional data constraints are integrated by filtering the generated statistical samples (as described above and in Supplement S1). Any samples which fall outside the lower or upper limits are rejected, and some samples are rejected to adjust the Gaussian Process such that it appropriately incorporates non Gaussian constraints. We use hyperparameters of 2 ‰ noise scale, and 10 Myr length scale. The Gaussian Process does not enforce a specific rate of change, but the chosen noise scale and length scale give an approximation of the rate of change, in this case roughly equivalent to a 0.2 ‰/Myr rate of change, which is on the faster end of agreement with the residence time and maximum rate of change from box modelling of the geochemical cycle of boron (Lemarchand et al., 2000). Lemarchand et al. (2000) suggest an upper limit to the rate of change of $\delta^{11}B_{sw}$ 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 $\delta^{11}B_{sw}$ suggests rather muted change across the Cenozoic relative to the change in $\delta^{11}B_4$. $\delta^{11}B_4$ has changed by roughly 8 ‰, and we find that $\delta^{11}B_{sw}$ has been responsible for a perhaps 2 ‰ of that change, leaving a 6 ‰ change driven predominantly by pH. The pattern of change of $\delta^{11}B_{sw}$ is non-linear, with a decline of roughly 1 ‰ from the early to mid Cenozoic, followed by an increase of approximately 1 ‰ from 30 Ma to 10 Ma, and a relatively stable value between those time periods and from 10 Ma until the present day. Our reconstruction generally matches the evolution curve used by 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 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 favours the higher $\delta^{11}B$ values in this interval due to their lower reported uncertainty (relative to the lower $\delta^{11}B_{sw}$ estimates from Greenop et al. (2017)). This results in an overall higher estimate of $\delta^{11}B_{sw}$ during the Miocene, and implies pH was potentially lower than previously calculated at this time. However, we note the uncertainties

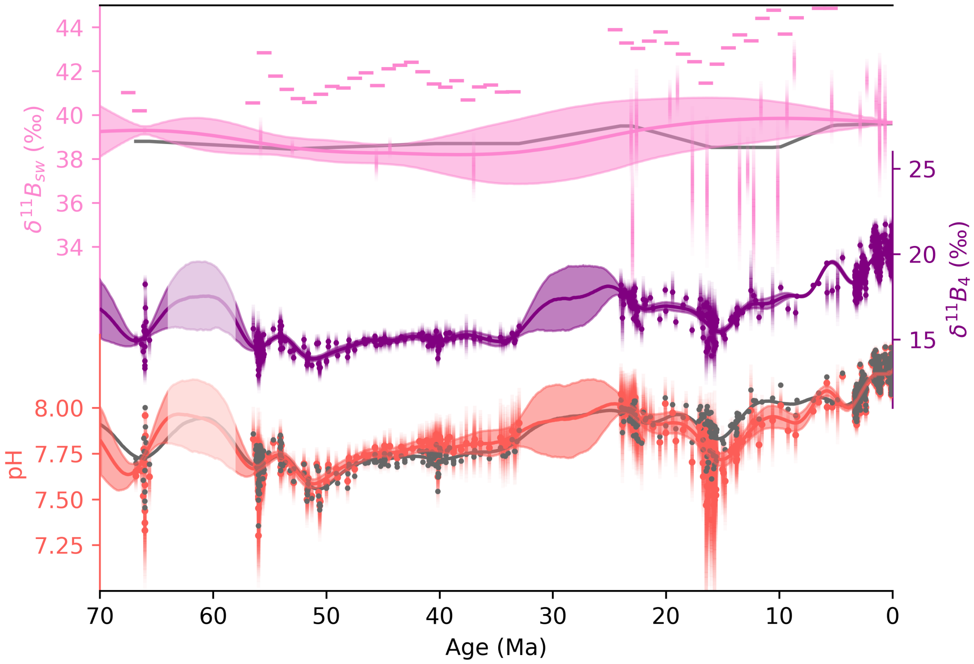


Figure 3. Our reconstruction of $\delta^{11}B_{sw}$ is shown pink, with a central line depicting the median, and a window representing the 95 % confidence interval. Central data constraints are shown by vertical pink bars representative of 95 % confidence intervals, and upper limits are shown by horizontal pink bars. The grey line depicts $\delta^{11}B_{sw}$ reconstructed by Rae et al. (2021) for comparison. $\delta^{11}B_4$ datapoints from 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 $\delta^{11}B_{sw}$ (red points, line, and shaded region) and the $\delta^{11}B_{sw}$ from Rae et al. (2021) (grey points and line) for comparison. We note that our Gaussian Process interpolation of $\delta^{11}B_4$ (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 $\delta^{11}B_4$, and our reconstruction would be tempered by filling in this region with $\delta^{11}B_4$ data.

in both our reconstruction and the individual $\delta^{11}B_{sw}$ constraints are high in the Neogene, and that some of the data constraints suggest rates of change in $\delta^{11}B_{sw}$ exceeding those compatible with geochemical box model predictions (as discussed further in Section 5.2). Uncertainty on $\delta^{11}B_{sw}$ is also particularly high in data gaps where $\delta^{11}B_4$ 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 $\delta^{11}B_{sw}$.

We provide illustrative curves for $\delta^{11}B_4$ and pH (see Figure 3) by using a Gaussian Process to interpolate $\delta^{11}B_4$, then combining this with our predicted $\delta^{11}B_{sw}$ and other ancillary parameters as in Rae et al. (2021). The Gaussian Process used to reconstruct $\delta^{11}B_4$ uses a length scale of 2 Myr to target relatively long term changes in pH, rather than individual palaeoclimatic events.

The pattern of change we find in $\delta^{11}B_{sw}$ is very similar to temporal trends in a number of other related biogeochemical signals, such as oceanic $^{87/86}Sr$, $^{187/188}Os$, and δ^7Li. Here we take published $^{87/86}Sr$ and δ^7Li from Misra and Froelich (2012), and compile a new record of $^{187/188}Os$ 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, 2020); van der Ploeg et al. (2018); Ravizza (1993); Ravizza and Turekian (1992); Ravizza and Peucker-Ehrenbrink (2003); Ravizza et al. (2001); 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 $^{87/86}Sr$ 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 ($^{87/86}Sr$ - 5.1 Myr (Broecker & Peng, Tsung-Hung, 1982), δ^7Li - 2.8 Myr (Stoffynegli & Mackenzie, 1984), $^{187/188}Os$ - 1 Myr). Interpolated data products for both $^{87/86}Sr$ and δ^7Li have been previously published (for instance in Misra and Froelich (2012)), but to our knowledge this is the first interpolated data product available for $^{187/188}Os$. 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 $\delta^{11}B_{sw}$, shows a long term increase over the Cenozoic. Oceanic strontium isotopes appear to show most similarity with our reconstructed $\delta^{11}B_{sw}$, while we note resemblance between the evolutions of pH, lithium isotopes, and osmium isotopes (as shown in Figure 4). We are able to provide an extremely narrow uncertainty window on our $^{87/86}Sr$ reconstruction due to strontium's long residence time and relatively high data density (and data quality) over the Cenozoic. By contrast, $^{187/188}Os$ 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. δ^7Li 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 δ^7Li 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.

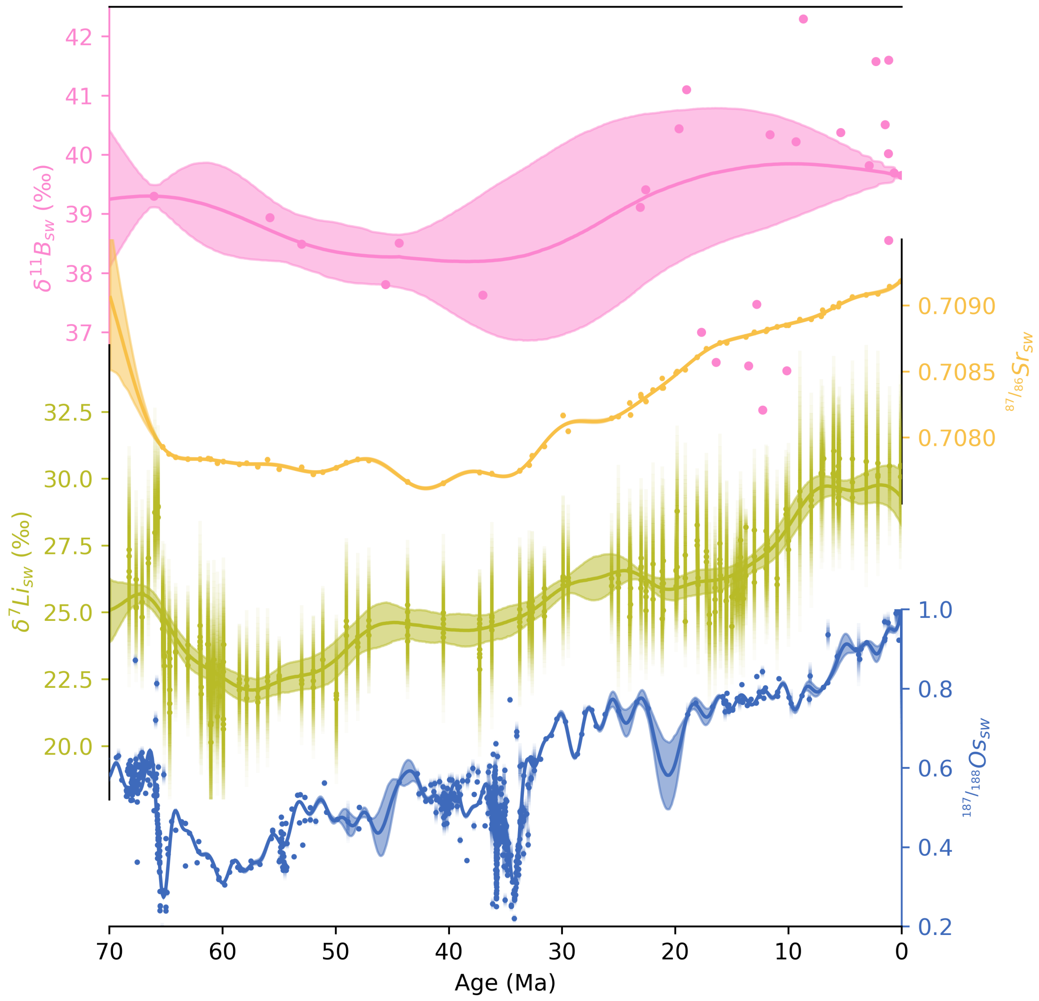


Figure 4. Our reconstruction of $\delta^{11}B_{sw}$ is shown pink, with a central line depicting the median, and a window representing the 95 % confidence interval. Data constraints are shown by the pink dots (here displayed without uncertainty for clearer comparison of signal trends). Other related geochemical signals δ^{7}Li, $^{87/86}$Sr, and $^{187/188}$Os are shown in yellow, green, and blue respectively after Misra and Froelich (2012). Data constraining $^{87/86}$Sr and δ^{7}Li is sourced from Misra and Froelich (2012), while data constraining $^{187/188}$Os is compiled from a number of sources (listed in the main text). We intepolate each of these signals with Gaussian Process with hyperparameters guided by the residence time of each signal.

5 Discussion

5.1 Decomposition of $\delta^{11}B_4$

$\delta^{11}B_4$ is an integrated signal, combining the effects of $\delta^{11}B_{sw}$, pH, pK_B^*, and ε. Of these, only a few are time-variable ($\delta^{11}B_{sw}$, pH, and some of the factors which influence pK_B^* - temperature, and seawater elemental composition). The sensitivity of $\delta^{11}B_4$ to these factors is such that only changes in $\delta^{11}B_{sw}$ or pH would be of sufficient magnitude to drive the observed 8 ‰ change in $\delta^{11}B_4$ across the Cenozoic, though temperature could have a non-negligible impact. To first order we might therefore simplify the $\delta^{11}B_4$ record into two components, $\delta^{11}B_{sw}$ and pH. Any change in $\delta^{11}B_4$ not explained by changes in $\delta^{11}B_{sw}$ must be the result of pH, and vice versa. Given the observed 8 ‰ change in $\delta^{11}B_4$, and our reconstructed 2 ‰ change in $\delta^{11}B_{sw}$, this suggests roughly 6 ‰ 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 $[H^+]$ perturbation is therefore approximately two thirds the magnitude of long term change seen within the last 65 million years.

5.2 $\delta^{11}B_{sw}$ Rate of Change

A key feature of the Gaussian Process methodology used here is the ability to draw smooth potential evolutions of $\delta^{11}B_{sw}$, which we augment by filtering samples to reject those with an unfeasibly high temporal gradient. The feasibility of the rate of change in $\delta^{11}B_{sw}$ (and the hyperparameters which tune the Gaussian Process smoothness) are primarily based on the geochemical box modelling work of Lemarchand et al. (2000). Lemarchand et al. (2000) suggest that the likely maximum rate of change in $\delta^{11}B_{sw}$ 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 $\delta^{11}B_{sw}$. 0.7 ‰/Myr is large enough that it does not result in rejection of potential evolutions in the earlier half of the Cenozoic, 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 $\delta^{11}B_{sw}$ than previously recognised are present in the dataset of Greenop et al. (2017). The surface-deep $\delta^{11}B_4$ pairs of Greenop et al. (2017) (as described in Section 2) appear to record a bimodal distribution, with some indicating $\delta^{11}B_{sw}$ similar to modern, and others (mostly in the middle Miocene climatic optimum) suggesting $\delta^{11}B_{sw}$ roughly 2 ‰ lower. These would appear to suggest oscillations in $\delta^{11}B_{sw}$ 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 3.3Myr). Instead, we draw evolutions consistent with the rate of change suggested by (Lemarchand et al., 2000) - which 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 centre of these constraints, drawn slightly high by the purportedly lower uncertainties in the higher $\delta^{11}B_{sw}$ 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 $\delta^{11}B_{sw}$. However, we acknowledge the possibility that the rate of change in $\delta^{11}B_{sw}$ is faster than currently appreciated, and look to further modelling 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 $\delta^{11}B_{sw}$. Indeed our work underscores the value of, and need for more, data-derived $\delta^{11}B_{sw}$ constraints outside of individual palaeoclimatic events.

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

Figure 4 shows that there is a broad scale similarity between the temporal evolution of $\delta^{11}B$, $\delta^{7}Li$, $^{87/86}$Sr and $^{187/188}$Os. In particular though, we observe greater similarity between $\delta^{11}B_{sw}$ and $^{87/86}$Sr, and also between pH, $\delta^{7}Li$ and $^{187/188}$Os. The similarity between $^{87/86}$Sr and $\delta^{11}B_{sw}$ is present in the overall shape of both signals, and in both showing an inflection point at approximately the same time, around 35Ma. However we note that $^{87/86}$Sr has been increasing continuously since 40 Ma, while our reconstruction of $\delta^{11}B_{sw}$ begins to decline at approximately 10 Ma. Given $^{87/86}$Sr and $\delta^{11}B_{sw}$ share many controls (such as weathering and hydrothermal influx), it is not surprising that their signals share some similarities.

We note an excellent correspondence between $\delta^{11}B_4$ and $\delta^{7}Li$ (shown in Figures 3 and 4, or together in Figure S1), in agreement with previous findings (Raitzsch & Hönisch, 2013; Greenop et al., 2017) as well as similarity between $\delta^{11}B_4$ and $^{187/188}$Os. Previously the correspondence between $\delta^{11}B_4$ and $\delta^{7}Li$ has been attributed to the overlap in drivers of $\delta^{11}B_{sw}$ and $\delta^{7}Li$. Under the two component model described in Section 5.1, we propose that similarity of $\delta^{11}B_4$ and $\delta^{7}Li$ could either be due to a relationship between $\delta^{7}Li$ and $\delta^{11}B_{sw}$ (as previously suggested by Raitzsch and Hönisch (2013)), or due to a relationship between $\delta^{7}Li$ and pH, or a combination of both. $\delta^{11}B_{sw}$ and $\delta^{7}Li$ share drivers, as do pH and $\delta^{7}Li$, which are linked to weathering, clay formation, and seafloor spreading. These confounding factors make it difficult to ascribe this correlation to either the $\delta^{11}B_{sw}$ or pH signals unambiguously (see Figure S1). If pH is prescribed to change slowly and linearly (as in Raitzsch and Hönisch (2013)), then, by necessity, $\delta^{11}B_{sw}$ will reflect $\delta^{11}B_4$ - requiring relatively rapid fluctuations in $\delta^{11}B_{sw}$. However, as discussed in Section 5.2, we proceed here under the assumption that the rate of change in $\delta^{11}B_{sw}$ is as calculated in Lemarchand et al. (2000), removing the possibility of fast changes in $\delta^{11}B_{sw}$, and necessitating that $\delta^{11}B_4$ and pH are tightly correlated. Therefore here, it is pH and $\delta^{7}Li$ which have similar trajectories across the Cenozoic.

Notwithstanding the broad scale correlation between $\delta^{7}Li$ and $\delta^{11}B_4$, we note an interesting divergence between these signals at the K-Pg boundary. At this time, there is a large excursion in $\delta^{7}Li$ 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 $^{187/188}$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. $\delta^{11}B_4$ undergoes an excursion but rapidly recovers to near pre-perturbation levels, whereas $\delta^{7}Li$ values stay low for the next ∼15 million years. This indicates that the drivers of pH and $\delta^{7}Li$ can be decoupled. However, the recorded change in $\delta^{7}Li$ 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 $\delta^{7}Li$ 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 $\delta^{11}B_4$ and $\delta^{7}Li$ more broadly during the Cenozoic. At present, however, studies disagree as to the nature of carbonate system control on foraminiferal $\delta^{7}Li$ (Vigier et al., 2015; Roberts et al., 2018).

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 Rae et al. (2021) and our pH (shown in Figure 3) 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 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 Rae et al. (2021). These suggestions of increased CO_2 relative to previous work come with the important caveat that, as mentioned above, the constraints on $\delta^{11}B_{sw}$ in this interval are highly variable, perhaps suggesting additional complicating factors which make $\delta^{11}B_{sw}$ difficult to estimate at this time. Our reconstruction of $\delta^{11}B_{sw}$ (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.

Secondly, 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 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 1200 μmol/kg to 2500 μmol/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 $\delta^{11}B_{sw}$ from modelling, 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 $\delta^{11}B_{sw}$ - and the other isotope systems reconstructed here (Figure 4).

While the standard 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 $\delta^{11}B_{sw}$ over the last 65 Myr. The 10,000 possible evolutions can be summarised 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 $\delta^{11}B_{sw}$ 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 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 $\delta^{11}B_{sw}$, we believe that it can not have changed substantially. Combining the $\delta^{11}B_{sw}$ 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 $\delta^{11}B_{sw}$ while each sample preserves a reasonable $\Delta\delta^{11}B_{sw}$ (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 $\delta^{11}B_{sw}$ that we provide are considered equally likely, it also allows further filtering. For instance if looking at a particular time period, or new information comes to light, which means we are able to be more certain about the rate of change in $\delta^{11}B_{sw}$, 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 $\delta^{11}B_{sw}$ from data and modelling, but also in terms of producing results which make 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 input data constraints are expected to be Gaussian distributions. As discussed above and in Supplement S1, we adjust the standard Gaussian Process methodology by using a rejection sampling strategy to incorporate other types of constraint.

The second limitation of the Gaussian Process 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 $\delta^{11}B_{sw}$, 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 $\delta^{11}B_{sw}$ 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 change 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. 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 $\delta^{11}B_{sw}$ is also limited by our understanding of past ocean conditions. As discussed in Section 1, most constraints on $\delta^{11}B_{sw}$ are at some level dependent on models. Typically, a wide range of model conditions were used to predict vertical gradients in $\delta^{13}C$ 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 ocean occupied a different mode in the past where the vertical gradients in $\delta^{13}C$ 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 $\delta^{13}C$ vs 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 modelling of the boron cycle.

8 Conclusions

We provide Cenozoic reconstructions of $\delta^{11}B_{sw}$ 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 $\delta^{11}B_{sw}$ was slightly higher than previously thought during the Miocene, but are generally in agreement with previous estimates of $\delta^{11}B_{sw}$ during the remainder of the Cenozoic. Generally speaking, uncertainties on $\delta^{11}B_{sw}$ are approximately 1 ‰, apart from during the Neogene where large uncertainties on data constraints propagate to large uncertainties in our reconstruction. Our higher estimated $\delta^{11}B_{sw}$ is indicative of lower Miocene pH than previously thought, though we acknowledge high uncertainties during this time.

We see an notable correspondence between $\delta^{11}B_4$ and δ^7Li, something which has previously been used to infer that $\delta^{11}B_{sw}$ likely followed a similar trajectory to δ^7Li (Raitzsch & Hönisch, 2013). However the evolution of $\delta^{11}B_{sw}$ constrained here shows that the majority of the Cenozoic change in $\delta^{11}B_4$ was driven by pH, indicating that links between the controls on pH and δ^7Li are perhaps the more important control on the similar trajectories of $\delta^{11}B_4$ and δ^7Li.

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 $\delta^{11}B_{sw}$, and to improve constraints on the various models which guide both the data constraints, and limits on the rate of change in $\delta^{11}B_{sw}$. However, using this method we are able to constrain $\delta^{11}B_{sw}$ to a range of ± 1 ‰ across most of the Cenozoic, improving current estimates. Results from this study can be used to propagate uncertainties in $\delta^{11}B_{sw}$ 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 $\delta^{11}B_{sw}$ (the median, and 95% confidence interval), but also provide 10,000 statistical samples of possible evolutions of $\delta^{11}B_{sw}$. Uncertainty in $\delta^{11}B_{sw}$ 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 ($\delta^{11}B_4$, pH, $^{87/86}Sr$, $^{187/188}Os$, and δ^7Li).

9 Open Research

9.1 Data Availability

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 a final version will be archived on Zenodo on publication. Data files associated with this project are currently privately shared at the following link: `https://figshare.com/s/043a054532aea2348125`, and by the time of publication will be permanently and publically archived. Three data files are provided which include:

- Data output for $\delta^{11}B_{sw}$ in the form of a .xlsx file, which contains both summary metrics and all 10,000 time series realisations.
- Original data, metrics summary of the reconstruction, and 10,000 individual statistical samples for $^{87/86}Sr$, $^{187/188}Os$, and $\delta^{7}Li$ in a .xlsx file.
- Metrics summary of the reconstruction, and individual statistical samples for $\delta^{11}B_4$ and pH in a .xlsx file.

Two forms of output for $\delta^{11}B_{sw}$ 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 $\delta^{11}B_{sw}$ and also the rate of change of $\delta^{11}B_{sw}$ as described in Section 6.

9.2 Software Availability

Software used in the project is written in Python, and is available on our GitHub repository during peer review: `https://github.com/St-Andrews-Isotope-Geochemistry/d11Bsw-Gaussian-Process`, and will be permanently archived on Zenodo on publication. 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, analyse the output, and display the results are also included.

References

Anagnostou, E., John, E. H., Babila, T. L., Sexton, P. F., Ridgwell, A., Lunt, D. J., ... Foster, G. L. (2020, September). Proxy evidence for state-dependence of climate sensitivity in the Eocene greenhouse. *Nature Communications*, *11*(1), 4436. doi: 10.1038/s41467-020-17887-x

Anagnostou, E., John, E. H., Edgar, K. M., Foster, G. L., Ridgwell, A., Inglis, G. N., ... Pearson, P. N. (2016, May). Changing atmospheric CO2 concentration was the primary driver of early Cenozoic climate. *Nature*, *533*(7603), 380–384. doi: 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. doi: 10.1002/palo.20015

Boudreau, B. P., Middelburg, J. J., Sluijs, A., & van der Ploeg, R. (2019, April). Secular variations in the carbonate chemistry of the oceans over the Cenozoic. *Earth and Planetary Science Letters*, *512*, 194–206. doi: 10.1016/j.epsl.2019.02.004

Broecker, W. S., & Peng, Tsung-Hung. (1982). *Tracers in the Sea.* Lamont-Doherty Geological Observatory, Columbia University.

Caves Rugenstein, J. K., Ibarra, D. E., & von Blanckenburg, F. (2019, July). Neogene cooling driven by land surface reactivity rather than increased weathering fluxes. *Nature*, *571*(7763), 99–102. doi: 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., … Wilson, P. A. (2017, December). Causes of ice age intensification across the Mid-Pleistocene Transition. *Proceedings of the National Academy of Sciences*, *114*(50), 13114–13119. doi: 10.1073/pnas.1702143114

de la Vega, E., Chalk, T. B., Wilson, P. A., Bysani, R. P., & Foster, G. L. (2020, July). Atmospheric CO2 during the Mid-Piacenzian Warm Period and the M2 glaciation. *Scientific Reports*, *10*(1), 11002. doi: 10.1038/s41598-020-67154-8

Derry, L. A. (2022, November). Carbonate weathering, CO2 redistribution, and Neogene CCD and pCO2 evolution. *Earth and Planetary Science Letters*, *597*, 117801. doi: 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. In *DOE*.

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. doi: 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 pages. doi: 10.14465/2022.REU03.OAC

Foster, G. L. (2008, July). 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. doi: 10.1016/j.epsl.2008.04.015

Foster, G. L., Lear, C. H., & Rae, J. W. B. (2012, August). The evolution of pCO2, ice volume and climate during the middle Miocene. *Earth and Planetary Science Letters*, *341–344*, 243–254. doi: 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). doi: 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. doi: 10.1002/2014PA002653

Greenop, R., Hain, M. P., Sosdian, S. M., Oliver, K. I. C., Goodwin, P., Chalk, T. B., … Foster, G. L. (2017, February). A record of Neogene seawater δ^{11}B reconstructed from paired δ^{11}B analyses on benthic and planktic foraminifera. *Climate of the Past*, *13*(2), 149–170. doi: 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. doi: 10.1029/2018PA003420

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

Hain, M. P., Foster, G. L., & Chalk, T. (2018, October). Robust Constraints on Past CO $_2$ Climate Forcing From the Boron Isotope Proxy. *Paleoceanography and Paleoclimatology*, *33*(10), 1099–1115. doi: 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). doi: 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. doi: 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. doi: 10.1029/2019PA003699

Haynes, L. L., Hönisch, B., Dyez, K. A., Holland, K., Rosenthal, Y., Fish, C. R., … Rae, J. W. B. (2017). Calibration of the B/Ca proxy in the planktic foraminifer Orbulina universa to Paleocene seawater conditions. *Paleoceanography*, *32*(6), 580–599. doi: 10.1002/2016PA003069

Henehan, M. J., Edgar, K. M., Foster, G. L., Penman, D. E., Hull, P. M., Greenop, R., … Pearson, P. N. (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. doi: 10.1029/2019PA003713

Henehan, M. J., Ridgwell, A., Thomas, E., Zhang, S., Alegret, L., Schmidt, D. N., … Hull, P. M. (2019, November). Rapid ocean acidification and protracted Earth system recovery followed the end-Cretaceous Chicxulub impact. *Proceedings of the National Academy of Sciences*, *116*(45), 22500–22504. doi: 10.1073/pnas.1905989116

Henehan, M. J., & Witts, J. D. (2023, May). Continental flood basalts do not drive later Phanerozoic extinctions. *Proceedings of the National Academy of Sciences*, *120*(21), e2303700120. doi: 10.1073/pnas.2303700120

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.

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 and Oceanography*, *56*(2), 562–576. doi: 10.4319/lo.2011.56.2.0562

Josso, P., Parkinson, I., Horstwood, M., Lusty, P., Chenery, S., & Murton, B. (2019, May). Improving confidence in ferromanganese crust age models: A composite geochemical approach. *Chemical Geology*, *513*, 108–119. doi: 10.1016/j.chemgeo.2019.03.003

Klemm, V., Levasseur, S., Frank, M., Hein, J. R., & Halliday, A. N. (2005, September). Osmium isotope stratigraphy of a marine ferromanganese crust. *Earth and Planetary Science Letters*, *238*(1), 42–48. doi: 10.1016/j.epsl.2005.07.016

Klochko, K., Kaufman, A. J., Yao, W., Byrne, R. H., & Tossell, J. A. (2006, August). Experimental measurement of boron isotope fractionation in seawater. *Earth and Planetary Science Letters*, *248*(1), 276–285. doi: 10.1016/j.epsl.2006.05.034

Lemarchand, D., Gaillardet, J., Lewin, É., & Allègre, C. J. (2000, December). The influence of rivers on marine boron isotopes and implications for reconstructing past ocean pH. *Nature*, *408*(6815), 951–954. doi: 10.1038/35050058

Marschall, H., & Foster, G. (Eds.). (2018). *Boron Isotopes*. Cham: Springer International Publishing. doi: 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., … Schmidt, D. N. (2015, February). Plio-Pleistocene climate sensitivity evaluated using high-resolution CO2 records. *Nature*, *518*(7537), 49–54. doi: 10.1038/nature14145

Misra, S., & Froelich, P. N. (2012, February). Lithium Isotope History of Cenozoic Seawater: Changes in Silicate Weathering and Reverse Weathering. *Science*, *335*(6070), 818–823. doi: 10.1126/science.1214697

Nir, O., Vengosh, A., Harkness, J. S., Dwyer, G. S., & Lahav, O. (2015, March). Direct measurement of the boron isotope fractionation factor: Reducing the uncertainty in reconstructing ocean paleo-pH. *Earth and Planetary Science Letters*, *414*, 1–5. doi: 10.1016/j.epsl.2015.01.006

Oxburgh, R. (1998, June). Variations in the osmium isotope composition of sea water over the past 200,000 years. *Earth and Planetary Science Letters*, *159*(3), 183–191. doi: 10.1016/S0012-821X(98)00057-0

Oxburgh, R. (2001). Residence time of osmium in the oceans. *Geochemistry, Geophysics, Geosystems*, *2*(6). doi: 10.1029/2000GC000104

Oxburgh, R., Pierson-Wickmann, A.-C., Reisberg, L., & Hemming, S. (2007, November). 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. doi: 10.1016/j.epsl.2007.08.033

Palmer, M. R., Pearson, P. N., & Cobb, S. J. (1998, November). Reconstructing Past Ocean pH-Depth Profiles. *Science*, *282*(5393), 1468–1471. doi: 10.1126/science.282.5393.1468

Paquay, F. S., Ravizza, G., & Coccioni, R. (2014, November). The influence of extraterrestrial material on the late Eocene marine Os isotope record. *Geochimica et Cosmochimica Acta*, *144*, 238–257. doi: 10.1016/j.gca.2014.08.024

Paquay, F. S., Ravizza, G. E., Dalai, T. K., & Peucker-Ehrenbrink, B. (2008, April). Determining Chondritic Impactor Size from the Marine Osmium Isotope Record. *Science*, *320*(5873), 214–218. doi: 10.1126/science.1152860

Paris, G., Gaillardet, J., & Louvat, P. (2010, November). Geological evolution of seawater boron isotopic composition recorded in evaporites. *Geology*, *38*(11), 1035–1038. doi: 10.1130/G31321.1

Park, H., & Schlesinger, W. H. (2002). Global biogeochemical cycle of boron. *Global Biogeochemical Cycles*, *16*(4), 20-1-20-11. doi: 10.1029/2001GB001766

Pearson, P. N., Foster, G. L., & Wade, B. S. (2009, October). Atmospheric carbon dioxide through the Eocene–Oligocene climate transition. *Nature*, *461*(7267), 1110–1113. doi: 10.1038/nature08447

Pegram, W. J., & Turekian, K. K. (1999, December). 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. doi: 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. doi: 10.1002/2014PA002621

Peucker-Ehrenbrink, B., & Ravizza, G. (2000, June). 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. doi: 10.1016/S0016-7037(99)00429-9

Peucker-Ehrenbrink, B., & Ravizza, G. E. (2020, January). Chapter 8 - Osmium Isotope Stratigraphy. In F. M. Gradstein, J. G. Ogg, M. D. Schmitz, & G. M. Ogg (Eds.), *Geologic Time Scale 2020* (pp. 239–257). Elsevier. doi: 10.1016/B978-0-12-824360-2.00008-5

Rae, Zhang, Y. G., Liu, X., Foster, G. L., Stoll, H. M., & Whiteford, R. D. (2021, May). Atmospheric CO $_2$ over the Past 66 Million Years from Marine Archives. *Annual Review of Earth and Planetary Sciences*, *49*(1), 609–641. doi: 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). Cham: Springer International Publishing. doi: 10.1007/978-3-319-64666-4_5

Raitzsch, M., Hain, M., Henehan, M., & Gattuso, J.-P. (2022, January). *Seacarbx* -

seacarb extension for deep-time carbonate system calculations. Zenodo. doi: 10.5281/zenodo.5909811

Raitzsch, M., & Hönisch, B. (2013, May). Cenozoic boron isotope variations in benthic foraminifers. *Geology*, *41*(5), 591–594. doi: 10.1130/G34031.1

Ravizza, G. (1993, July). Variations of the187Os/186Os ratio of seawater over the past 28 million years as inferred from metalliferous carbonates. *Earth and Planetary Science Letters*, *118*(1), 335–348. doi: 10.1016/0012-821X(93)90177-B

Ravizza, G., Norris, R. N., Blusztajn, J., & Aubry, M. P. (2001). An osmium isotope excursion associated with the Late Paleocene thermal maximum: Evidence of intensified chemical weathering. *Paleoceanography*, *16*(2), 155–163. doi: 10.1029/2000PA000541

Ravizza, G., & Peucker-Ehrenbrink, B. (2003, May). 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. doi: 10.1016/S0012-821X(03)00137-7

Ravizza, G., & Turekian, K. K. (1992, May). The osmium isotopic composition of organic-rich marine sediments. *Earth and Planetary Science Letters*, *110*(1), 1–6. doi: 10.1016/0012-821X(92)90034-S

Reusch, D. N., Ravizza, G., Maasch, K. A., & Wright, J. D. (1998, July). Miocene seawater 187Os/188Os ratios inferred from metalliferous carbonates. *Earth and Planetary Science Letters*, *160*(1), 163–178. doi: 10.1016/S0012-821X(98)00082-X

Ridgwell, A., Hargreaves, J. C., Edwards, N. R., Annan, J. D., Lenton, T. M., Marsh, R., … Watson, A. (2007, January). Marine geochemical data assimilation in an efficient Earth System Model of global biogeochemical cycling. *Biogeosciences*, *4*(1), 87–104. doi: 10.5194/bg-4-87-2007

Ridgwell, A., & Zeebe, R. (2005, June). 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. doi: 10.1016/j.epsl.2005.03.006

Roberts, J., Kaczmarek, K., Langer, G., Skinner, L. C., Bijma, J., Bradbury, H., … Misra, S. (2018, September). Lithium isotopic composition of benthic foraminifera: A new proxy for paleo-pH reconstruction. *Geochimica et Cosmochimica Acta*, *236*, 336–350. doi: 10.1016/j.gca.2018.02.038

Robinson, N., Ravizza, G., Coccioni, R., Peucker-Ehrenbrink, B., & Norris, R. (2009, May). 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. doi: 10.1016/j.epsl.2009.02.019

Sosdian, S. M., Babila, T. L., Greenop, R., Foster, G. L., & Lear, C. H. (2020, January). Ocean Carbon Storage across the middle Miocene: A new interpretation for the Monterey Event. *Nature Communications*, *11*(1), 134. doi: 10.1038/s41467-019-13792-0

Sosdian, S. M., Greenop, R., Hain, M. P., Foster, G. L., Pearson, P. N., & Lear, C. H. (2018, September). Constraining the evolution of Neogene ocean carbonate chemistry using the boron isotope pH proxy. *Earth and Planetary Science Letters*, *498*, 362–376. doi: 10.1016/j.epsl.2018.06.017

Stoffynegli, P., & Mackenzie, F. T. (1984, April). Mass balance of dissolved lithium in the oceans. *Geochimica et Cosmochimica Acta*, *48*(4), 859–872. doi: 10.1016/0016-7037(84)90107-8

Stoll, H. M., Guitian, J., Hernandez-Almeida, I., Mejia, L. M., Phelps, S., Polissar, P., … Ziveri, P. (2019, March). Upregulation of phytoplankton carbon concentrating mechanisms during low CO2 glacial periods and implications for the phytoplankton pCO2 proxy. *Quaternary Science Reviews*, *208*, 1–20. doi: 10.1016/j.quascirev.2019.01.012

Tierney, J. E., Zhu, J., Li, M., Ridgwell, A., Hakim, G. J., Poulsen, C. J., … Kump, L. R. (2022, October). Spatial patterns of climate change across the Paleocene–Eocene Thermal Maximum. *Proceedings of the National Academy of Sciences*, *119*(42), e2205326119. doi: 10.1073/pnas.2205326119

van der Ploeg, R., Selby, D., Cramwinckel, M. J., Li, Y., Bohaty, S. M., Middelburg, J. J., & Sluijs, A. (2018, July). Middle Eocene greenhouse warming facilitated by diminished weathering feedback. *Nature Communications*, *9*(1), 2877. doi: 10.1038/s41467-018-05104-9

Vigier, N., Rollion-Bard, C., Levenson, Y., & Erez, J. (2015, January). Lithium isotopes in foraminifera shells as a novel proxy for the ocean dissolved inorganic carbon (DIC). *Comptes Rendus Geoscience*, *347*(1), 43–51. doi: 10.1016/j.crte.2014.12.001

Yu, J., & Elderfield, H. (2007, June). Benthic foraminiferal B/Ca ratios reflect deep water carbonate saturation state. *Earth and Planetary Science Letters*, *258*(1), 73–86. doi: 10.1016/j.epsl.2007.03.025

Zeebe, R. E. (2012, January). LOSCAR: Long-term Ocean-atmosphere-Sediment CArbon cycle Reservoir Model v2.0.4. *Geoscientific Model Development*, *5*(1), 149–166. doi: 10.5194/gmd-5-149-2012

Zeebe, R. E., & Wolf-Gladrow, D. A. (2001). *CO_2 in seawater: Equilibrium, kinetics, isotopes* (No. 65). Amsterdam ; New York: Elsevier.

Zhang, Y. G., Henderiks, J., & Liu, X. (2020, July). Refining the alkenone-pCO2 method II: Towards resolving the physiological parameter 'b'. *Geochimica et Cosmochimica Acta*, *281*, 118–134. doi: 10.1016/j.gca.2020.05.002

Figure 1.

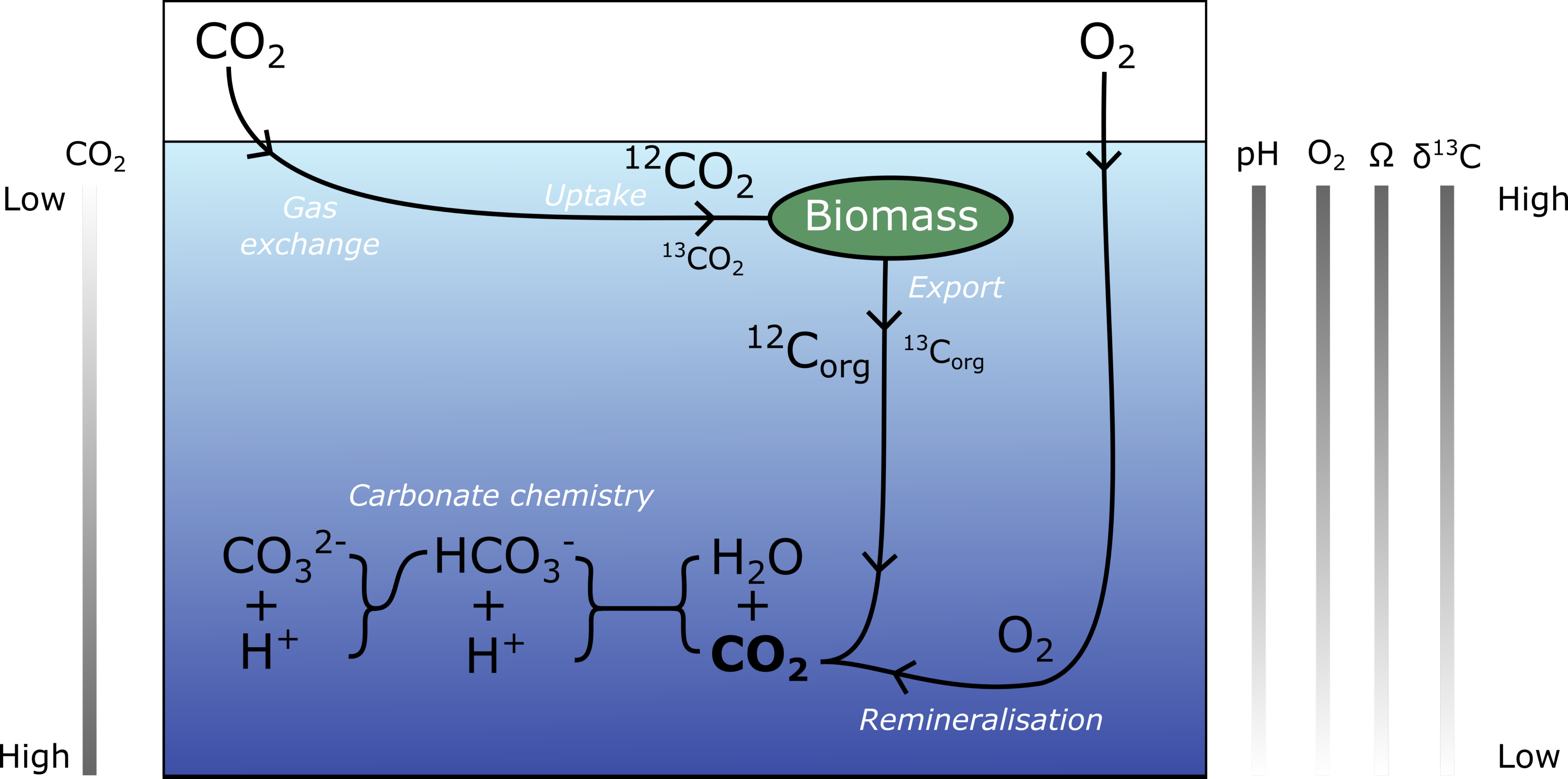

CO2
O2
CO2
Low
High
Gas exchange
Uptake
12CO2
13CO2
Biomass
Export
12Corg
13Corg
Carbonate chemistry
CO3 2- + H+
HCO3 - + H+
H2O + CO2
O2
Remineralisation
pH
O2
Ω
δ13C
High
Low

Figure 2a.

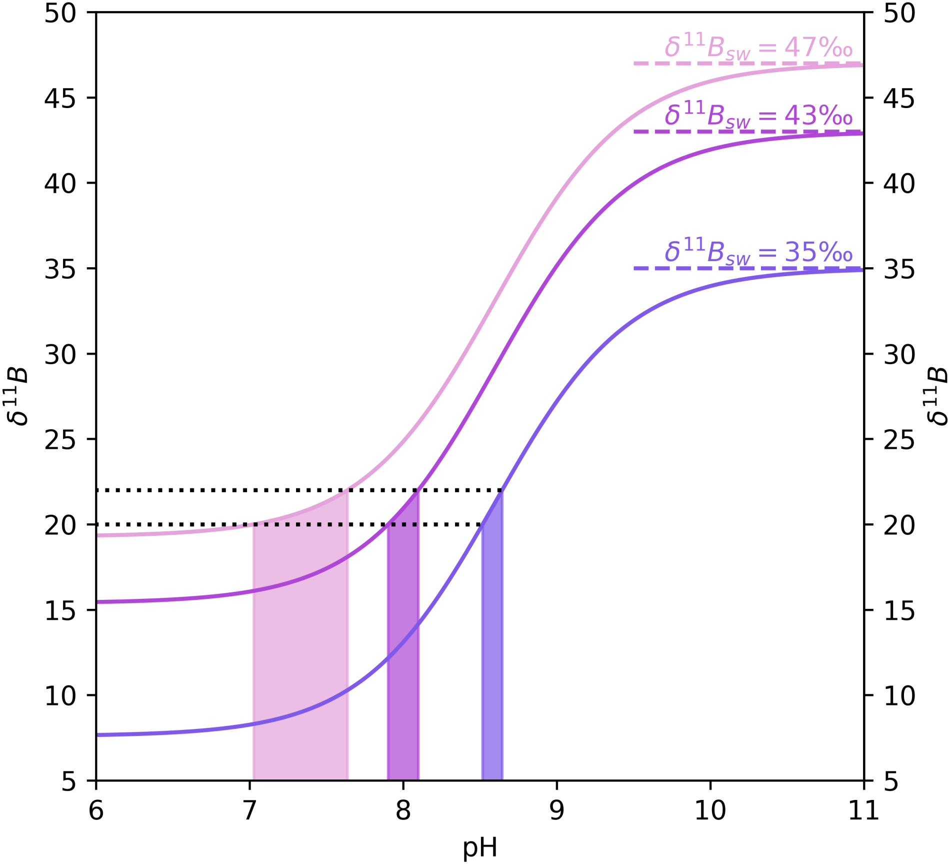

$\delta^{11}B_{sw} = 47‰$
$\delta^{11}B_{sw} = 43‰$
$\delta^{11}B_{sw} = 35‰$
$\delta^{11}B$
$\delta^{11}B$
pH
50
45
40
35
30
25
20
15
10
5
6
7
8
9
10
11

Figure 2b.

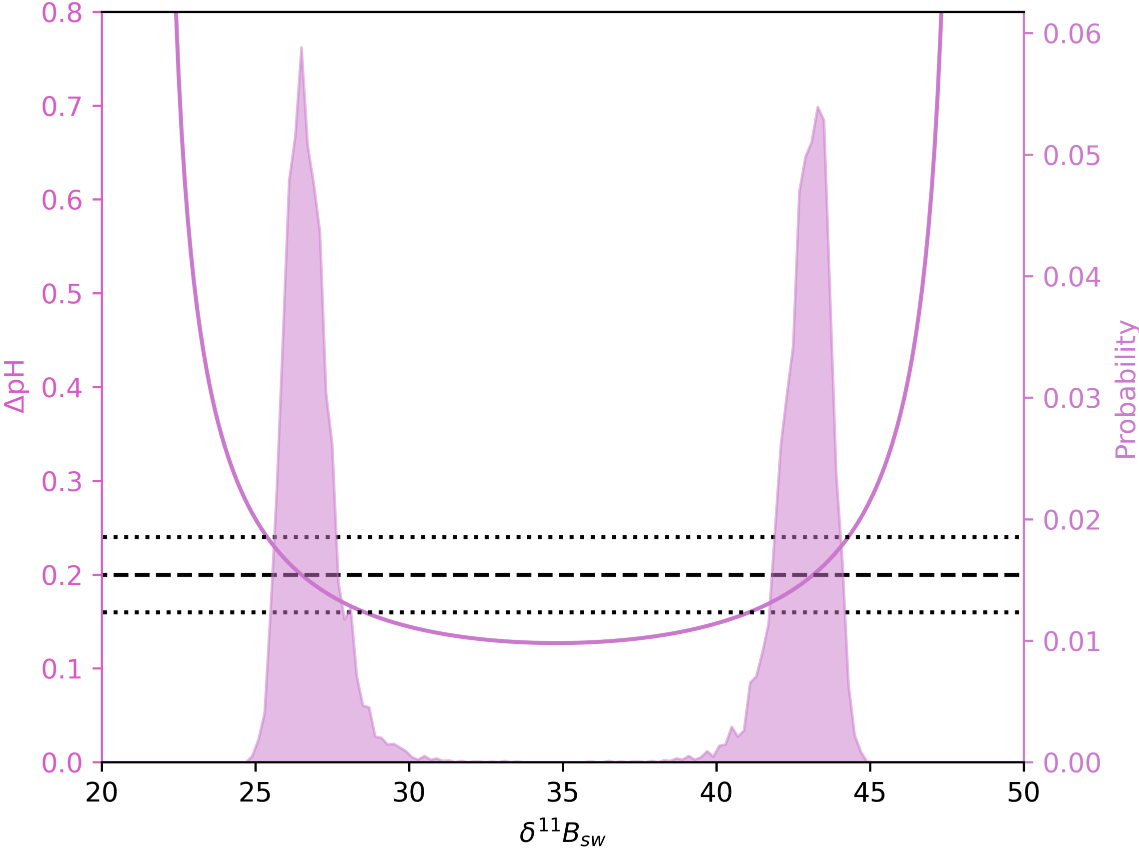
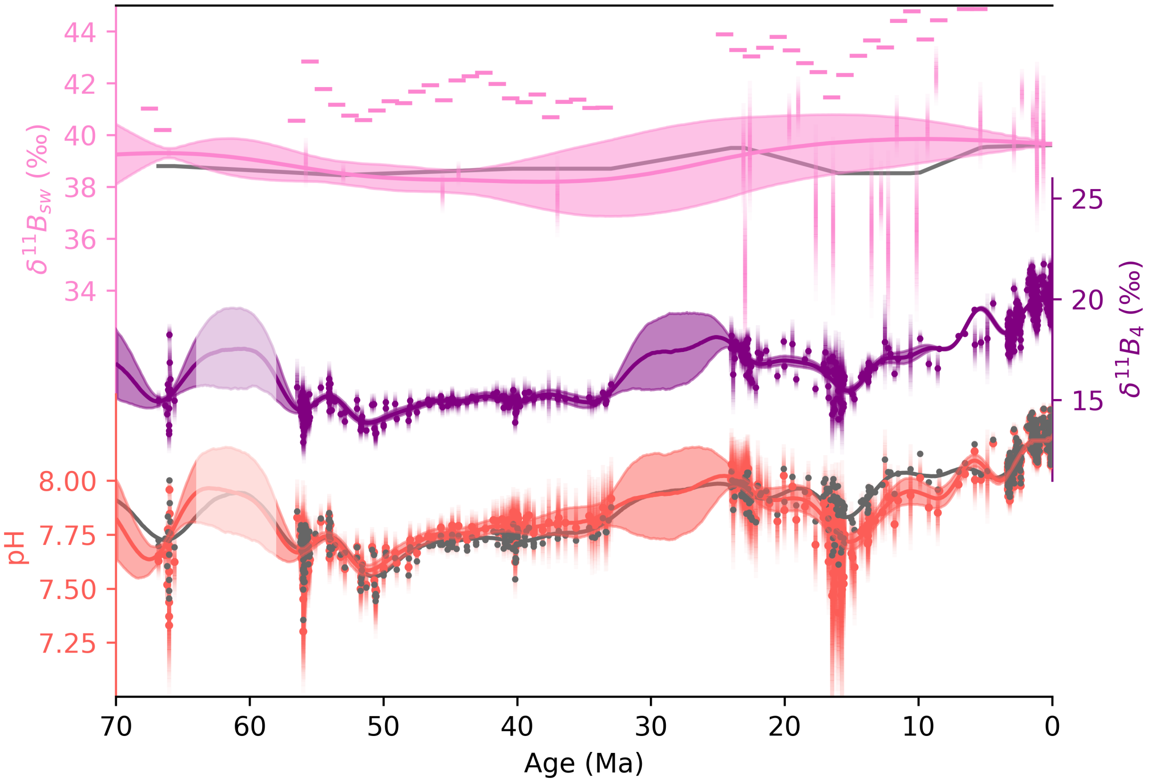
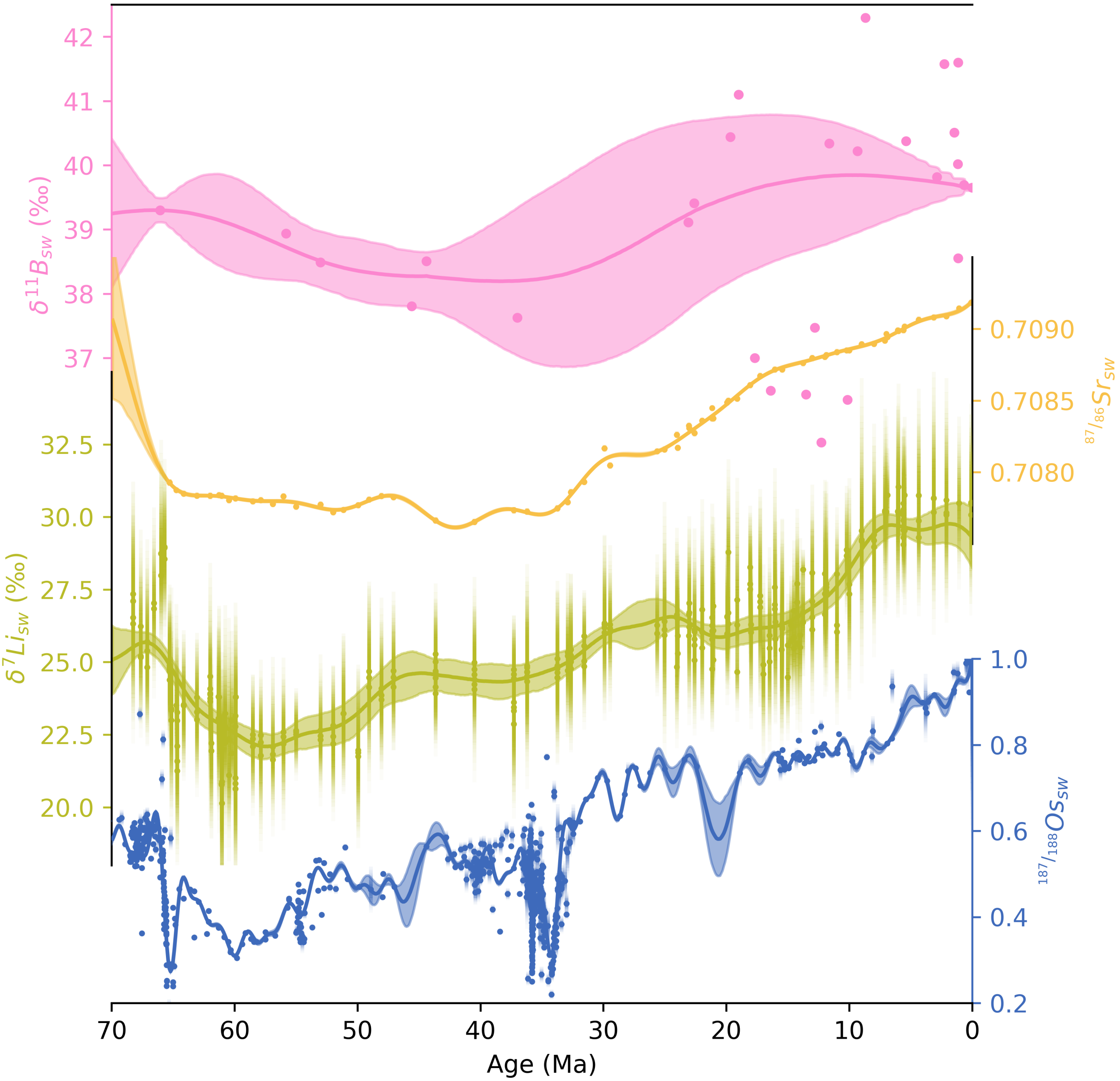

0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
ΔpH
0.06
0.05
0.04
0.03
0.02
0.01
0.00
Probability
20
25
30
35
40
45
50
$\delta^{11}B_{sw}$

Figure 3.



44
42
40
38
36
34
$\delta^{11}B_{sw}$ (‰)
25
20
15
$\delta^{11}B_4$ (‰)
8.00
7.75
7.50
7.25
pH
70
60
50
40
30
20
10
0
Age (Ma)

Figure 4.


$\delta^{11}B_{sw}$ (‰)
42
41
40
39
38
37
$^{87}/_{86}Sr_{sw}$
0.7090
0.7085
0.7080
$\delta^{7}Li_{sw}$ (‰)
32.5
30.0
27.5
25.0
22.5
20.0
$^{187}/_{188}Os_{sw}$
1.0
0.8
0.6
0.4
0.2
70
60
50
40
30
20
10
0
Age (Ma)

Ross Whiteford[1], Timothy J. Heaton[2], Michael J. Henehan[3], Eleni Anagnostou[4], Hana Jurikova[1], Gavin L. Foster[5], and James W.B. Rae[1]

[1]University of St Andrews

[2]Department of Statistics, School of Mathematics, University of Leeds, Leeds LS2 9JT, UK

[3]School of Earth Sciences, University of Bristol, Wills Memorial Building, Queens Road, Bristol, UK

[4]GEOMAR Helmholtz Centre for Ocean Research Kiel

[5]School of Ocean Earth Science, University of Southampton, UK

Contents of this file

1. Text S1
2. Figures S1 to S4
3. Data files S1 to S3

Additional Supporting Information (Files uploaded separately)

1. Data file S1 - Reconstructed $\delta^{11}B_{sw}$ and summary metrics
2. Data file S2 - Reconstructed pH summary metrics
3. Data file S3 - Data and reconstructed $^{87/86}$Sr, δ^{7}Li, and $^{187/188}$Os.

---

1. Gaussian Processes With Non-Gaussian Constraint Noise

1.1. Introduction

In the main text, we gave a brief overview of the Gaussian Process methodology as it related to reconstruction of $\delta^{11}\mathrm{B}_{\mathrm{sw}}$. Here we give a fuller, more generalised and statistically rigorous description of the Gaussian Process methodology, including how it was adapted to incorporate each style of constraint mentioned in the main text. We then illustrate how this technique works in practice by testing it against a hypothetical signal with noisy constraints, with a step by step walkthrough of integrating various forms of information analogous to the types of constraint we have on $\delta^{11}\mathrm{B}_{\mathrm{sw}}$.

Suppose that we observe a function $f(\cdot)$ subject to (potentially non-Gaussian) noise in constraints at a set of known times t_i, i.e.,

$$y_i = f(t_i) + \epsilon_i \quad \text{for } i = 1, \dots, M.$$

Here, ϵ_i can be a general probability distribution and is not required to be a standard Gaussian. Furthermore, suppose that we may have some further constraints on the value of $f(t)$ for certain values of t (e.g. lower/upper bounds), or additional non-standard information on the values of $f(t)$ (such as constraints on the change in value over time $\frac{df}{dt}$).

We wish to obtain a non-parametric posterior estimate of the function $f(\cdot)$ modelled as a Gaussian Process (GP) given both the potentially complex observations y_i and any additional non-standard information. In the standard GP setting, the function $f(\cdot)$ is assumed to be observed subject to normally distributed noise. As a consequence, the exact posterior for f given $\mathbf{y}$ can be easily calculated directly. However, when the available constraints ϵ are non-Gaussian, the GP's posterior can not be written down exactly and is much more challenging to calculate, as it does not take a standard form. To obtain posterior samples under such a non-Gaussian constraint model, we therefore implement a rejection sampling approach. Specifically, we aim to draw from a *nearby* distribution (from which it is possible to sample directly) and then reject/accept these samples using rejection sampling principles to obtain the correct GP posterior under our non-Gaussian constraint model.

To explain our approach in this Supplementary Information, we will first provide a brief background to Gaussian Processes and explain how they are usually fitted in the context of non-parametric regression given a set of constraints with Gaussian noise. We then introduce the idea of rejection sampling before going on to show how this idea can be used to sample from the posterior of a Gaussian Process in the presence of constraints with non-Gaussian uncertainty (or when additional, non-standard information is available).

1.2. Definition of a Gaussian Process (GP) Prior

A (one-dimensional input) zero-mean *Gaussian Process* $f(z) \sim \mathcal{GP}(0, k(t,t'))$ is a collection of random variables, any finite number of which have a joint Gaussian distribution (Rasmussen & Williams, 2006). It is completely specified by its covariance function:

$$k(t,t') = \mathbf{E}[f(t)f(t')].$$

When using a Gaussian Process to perform regression, the random variables represent the values of the function $f(t)$ at time t. For a set of N times $\mathbf{t}_\star = (t_1^\star, t_2^\star, \ldots, t_N^\star)^T$, our prior specifies

$$\mathbf{f}_\star = \mathbf{f}(\mathbf{t}_\star) \sim \mathcal{N}(0, K_{\mathbf{t}_\star,\mathbf{t}_\star}),$$

where $K_{\mathbf{t}_\star,\mathbf{t}_\star}$ denotes the matrix of the covariances evaluated at all pairs of the times $t_i^\star$.

1.3. Updating the GP Prior under a Normal Observational Model

Typically, when performing non-parametric regression, we assume that we observe the function $f(t)$ subject to normally-distributed noise, i.e.,

$$\mathbf{y} = \mathbf{f}(\mathbf{t}) + \boldsymbol{\eta}$$

where the noise $\boldsymbol{\eta} \sim \mathcal{N}(0, \Sigma)$. This observational noise may have dependence encoded in the covariance matrix Σ but critically is assumed to be normally distributed. In such a situation, we can use the standard properties of the multivariate normal distribution to derive the posterior distribution for our function values at our times of interest $\mathbf{t}_\star$ exactly:

$$\mathbf{f}_\star | \mathbf{t}, \mathbf{y}, \mathbf{t}_\star \sim \mathcal{N}(\bar{\mathbf{f}}_\star, cov(\mathbf{f}_\star)),$$

where

$$\begin{aligned} \bar{\mathbf{f}}_\star &= K^T_{\mathbf{t},\mathbf{t}_\star}[K_{\mathbf{t},\mathbf{t}} + \Sigma]^{-1}\mathbf{y}, \\ cov(\mathbf{f}_\star) &= K_{\mathbf{t}_\star,\mathbf{t}_\star} - K^T_{\mathbf{t},\mathbf{t}_\star}[K_{\mathbf{t},\mathbf{t}} + \Sigma]^{-1}K_{\mathbf{t},\mathbf{t}_\star}. \end{aligned}$$

See Rasmussen and Williams (2006) for full details. However, as soon as the observational model becomes non-normal, i.e., it is no longer the case that $y_i | f(t_i) \sim N(f(t_i), \sigma_i^2)$ then the GP posterior becomes much more complex and will no longer take the form of a simple multivariate normal. To estimate the posterior distribution in such instances, we will therefore take a different approach based upon rejection sampling.

1.4. Rejection Sampling

Rejection sampling is a general purpose method that enables sampling from non-standard distributions. Suppose we wish to sample $\mathbf{X}$ from a particular target probability density $f_X(\mathbf{x})$ but, for some reason, we cannot do so directly. However, suppose there exists an alternative *envelope* density function $g_Z(\mathbf{z})$ from which we can sample that satisfies the condition $\frac{f_X(\mathbf{x})}{g_Z(\mathbf{x})}$ bounded $\forall\mathbf{x}$. For any constant $c \geq \sup_{\mathbf{x}} \frac{f_X(\mathbf{x})}{g_Y(\mathbf{x})}$, we can then obtain samples from our desired target $f_X(\mathbf{x})$ using the following rejection method:

1. Sample $\mathbf{z}$ from an envelope density that is proportional to $g_Z(\mathbf{z})$, and a uniform u from $U[0,1]$.

2. If $u \leq \frac{f_X(\mathbf{z})}{c\, g_Y(\mathbf{z})}$, state $\mathbf{X} = \mathbf{z}$, otherwise return to step 1.

For maximum efficiency we therefore want c, i.e., $\sup_{\mathbf{x}} \frac{f_X(\mathbf{x})}{g_Z(\mathbf{x})}$ as small as possible. We therefore aim to find an *envelope* density function g_Z that is both easy to sample from and mimics the target f_X as closely as possible. The concept of rejection sampling is shown graphically in Figure S3.

2. Rejection Sampling GPs with Non-Gaussian Noise

Returning to our specific non-parametric regression, suppose that we observe a function subject to (potentially non-Gaussian) noise at a set of known times t_i,

$$y_i = f(t_i) + \epsilon_i \quad \text{for } i = 1, \ldots, M.$$

We wish to place a Gaussian Process prior on the values of $f(t)$ and then sample from the posterior under the (potentially non-Gaussian) observational model ϵ,

$$p_\epsilon(\mathbf{f}|\mathbf{y}) = \frac{p_\epsilon(\mathbf{y}|\mathbf{f})\pi(\mathbf{f})}{p_\epsilon(\mathbf{y})} \propto p_\epsilon(\mathbf{y}|\mathbf{f})\pi(\mathbf{f}). \tag{$\dagger$}$$

Due to the non-Gaussian nature of ϵ, we cannot directly sample from this posterior distribution. However we can sample from an alternative distribution of our choosing and then use rejection sampling principles. We will typically use the GP posterior under a normally-distributed error model for this envelope distribution. This η-error model as discussed in Section 1.3 can be calculated precisely and, in general, will hopefully be close to the true target posterior. Having chosen a suitable envelope density, proportional to $g(\mathbf{f}|\mathbf{y})$, the rejection algorithm thus becomes:

1. Calculate $c^{\ddagger} = \sup_{\mathbf{f}} \frac{p_{\epsilon}(\mathbf{y}|\mathbf{f})\pi(\mathbf{f})}{g(\mathbf{f}|\mathbf{y})}$

2. Sample from the envelope density $g(\mathbf{f}^{\ddagger}|\mathbf{y})$ a potential $\mathbf{f}^{\ddagger}$ at both times of interest $\mathbf{t}^{\star}$ and the times $\mathbf{t}$ at which we have observations $\mathbf{y}$

3. Sample $u \sim U[0,1]$, if $u \leq \frac{p_{\epsilon}(\mathbf{y}|\mathbf{f}^{\ddagger})\pi(\mathbf{f}^{\ddagger})}{c^{\ddagger} g(\mathbf{f}^{\ddagger}|\mathbf{y})}$ then accept $\mathbf{f} = \mathbf{f}^{\ddagger}$ as a draw from the correct posterior, otherwise return to step 2.

The calculation of both $c^{\ddagger}$ and the acceptance criteria in step 3 will generally only depend upon the sampled values of $\mathbf{f}^{\ddagger}$ at the times $\mathbf{t}$ with observations. This will reduce calculation. Furthermore, we note that $p_{\eta}(\mathbf{y}|\mathbf{f}) = \prod_{i=1}^{N} p_{\eta_i}(y_i|f(t_i))$ if the observations are independent. Exceptions may however occur if we have additional, non-standard, constraints such as on the range or variation of the function $f(\cdot)$.

We can repeat this rejection sampling technique until we obtain a large number of posterior realisations $\mathbf{f}$ from the target distribution (that corresponds to the general observational noise model). These can then be summarised by Monte Carlo to provide posterior means and variances for any $f(t)$.

2.1. Rejection Sampling Implementation

Our rejection sampling algorithm to sample from the correct posterior under a general observational error model ϵ then becomes (after cancelling common terms):

1. Sample $\mathbf{f}^\eta$ at times of interest $\mathbf{t}^\star$ and also times $\mathbf{t}$ at which we have observations $\mathbf{y}$ from GP posterior under normally-distributed η error model.

2. Sample $u \sim U[0,1]$, if $u \leq \frac{p_\epsilon(\mathbf{y}|\mathbf{f}^\eta)}{c^\star\, p_\eta(\mathbf{y}|\mathbf{f})}$ (where, as defined above, $c^\star = \sup_{\mathbf{f}} \frac{p_\epsilon(\mathbf{y}|\mathbf{f}^\eta)}{p_\eta(\mathbf{y}|\mathbf{f})}$) then accept $\mathbf{f} = \mathbf{f}^\eta$ as a draw from the correct posterior, otherwise return to step 1. Again, we note that this only depends upon the sampled values of $\mathbf{f}^\eta$ at the times $\mathbf{t}$ with observations and that, e.g., $p_\eta(\mathbf{y}|\mathbf{f}) = \prod_{i=1}^N p_{\eta_i}(y_i|f(t_i))$ if those observations are independent.

We repeat this sampling technique until we obtain a large number of posterior realisations $\mathbf{f}$ from the target distribution (that correspond to the general observational noise model) which can then be summarised by Monte-Carlo.

3. Specific Examples

While the rejection sampling approach may appear complicated, in many instances it will simplify considerably. We discuss some specific examples below.

3.1. Incorporating Upper and Lower Bounds

Suppose that we have a set of normally-distributed observations $\mathbf{y}$ but, in addition, a further set of values $\mathbf{z} = (z_1, \ldots, z_K)^T$ that operate as upper bounds on the unknown function, i.e., it is the case that $f(t_j^b) < z_j$ for given times $t_1^b, \ldots, t_K^b$. In this case, we consider that these K additional values are entirely uninformative about the value of $f(t)$ beyond providing such a bound. Consequently, the target posterior is:

$$p(\mathbf{f}|\mathbf{y}, \mathbf{z}) \propto \left\{ \prod_{j=1}^{M} \mathbb{1}_{[f(t_j^b) < z_j]} \right\} p_\eta(\mathbf{y}|\mathbf{f}) \pi(\mathbf{f}),$$

where $p_\eta(\mathbf{y}|\mathbf{f})$ is the usual normal likelihood function for the observations $\mathbf{y}$. For our envelope function, we can sample directly from the GP posterior considering just the regular, normally-distributed, observations $\mathbf{y}$, i.e., $g(\mathbf{f}|\mathbf{y}) \propto p_\eta(\mathbf{y}|\mathbf{f})\pi(\mathbf{f})$ so that our $c^\ddagger = 1$. Our algorithm then becomes simply:

1. Sample $\mathbf{f}^\ddagger$ from the standard GP posterior based upon normally-distributed observations $\mathbf{y}$ at both times of interest $\mathbf{t}^\star$ and the times $\mathbf{t}^b$ at which there are upper bounds. This can be done as described in Section 1.3
2. Accept $\mathbf{f}^\ddagger$ as a draw from the true target posterior if it satisfies all the constraints $\mathbf{z}$; otherwise reject and return to step 1.

This has a straightforward analogue when we have combinations of upper and lower bounds.

3.2. Incorporating Non-Gaussian Observations

When our observations $\mathbf{y}$ are subject to non-normal noise (which we have denoted by η) then an appropriate envelope density to use for rejection sampling might be the GP posterior for $\mathbf{f}$ had the noise been normally distributed (see Figure S3). In other words, we use might use a GP conditioned on observations with normal noise as our initial estimate of the posterior, then refine this through rejection sampling, i.e., the posterior for $f(\cdot)$ under the model:

$$y_i = f(t_i) + \eta_i \quad \text{for } i = 1, \dots, n.$$

where $\eta \sim N(0, \sigma_i^2)$. In this case the envelope function is $g(\mathbf{f}|\mathbf{y}) = p_\eta(\mathbf{f}|\mathbf{y}) = p_\eta(\mathbf{y}|\mathbf{f})\pi(\mathbf{f})$. This distribution is known, see Section 1.3, and it is easy to sample from it directly. To perform rejection sampling, we are required to calculate

$$c^\dagger = \sup_{\mathbf{f}} \frac{p_\epsilon(\mathbf{y}|\mathbf{f})\pi(\mathbf{f})}{p_\eta(\mathbf{y}|\mathbf{f})\pi(\mathbf{f})} = \sup_{\mathbf{f}} \frac{p_\epsilon(\mathbf{y}|\mathbf{f})}{p_\eta(\mathbf{y}|\mathbf{f})}.$$

Since $p_\eta(\mathbf{y}|\mathbf{f})$ is a normal distribution with infinite support, this supremum will exist for almost all alternative errors models (unless they have different tail behaviour). The calculation of $c^\dagger$ only depends upon the sampled values of $\mathbf{f}$ at the times $\mathbf{t}$ for which we have observations $\mathbf{y}$. Furthermore, if the observations y_i are independent, the numerator and denominator in the supremum can be calculated as independent products since, e.g., $p_\eta(\mathbf{y}|\mathbf{f}) = \prod_{i=1}^{N} p_{\eta_i}(y_i|f(t_i))$. Our rejection sampling algorithm to sample from the correct GP posterior under a general observational error model ϵ then becomes:

1. Sample $\mathbf{f}^\ddagger$ at times of interest $\mathbf{t}^\star$ and also times $\mathbf{t}$ at which we have observations $\mathbf{y}$ from GP posterior under normally-distributed η error model.

2. Sample $u \sim U[0,1]$, if $u \leq \frac{p_\epsilon(\mathbf{y}|\mathbf{f}^\ddagger)}{c^\dagger\, p_\eta(\mathbf{y}|\mathbf{f}^\ddagger)}$, then accept $\mathbf{f} = \mathbf{f}^\ddagger$ as a draw from the correct posterior. Otherwise return to step 1.

Again, the acceptance criteria in step 2 only depends upon the sampled values of $\mathbf{f}^\ddagger$ at the times $\mathbf{t}$ corresponding to the observations $\mathbf{y}$. Also, if the observations $\mathbf{y}$ are independent, then the likelihood terms reduce to products, e.g., $p_\eta(\mathbf{y}|\mathbf{f}) = \prod_{i=1}^{N} p_{\eta_i}(y_i|f(t_i))$.

Modelling Outlying Observations: A specific instance where we may wish to consider non-normal noise occurs if we believe that some of the observations $\mathbf{y}$ may be outliers. In such situations, we are required to select both the probability w of an observation y_i being an outlier and, when it is an outlier, its specific distribution. We will denote the observational noise in an outlier model as ζ. Our likelihood for the observed y_i given $f(t_i)$ then becomes a mixture:

$$p_\zeta(y_i|f_i) = (1-w)p_\epsilon(y_i|f_i) + w p_{\mathrm{o}}(y_i|f_i)$$

Here, $p_\epsilon(y_i|f_i)$ is the quoted *non-outlier* likelihood; and $p_\mathrm{o}(x)$ the selected outlier likelihood. A natural choice for $p_\mathrm{o}(y_i|f_i)$ may be a uniform distribution $U[f_i - a, f_i + b]$ where a and b are chosen suitably (or even simply $U[a, b]$). We can then proceed as above using the mixture $p_\zeta(x)$ as our observational model.

3.3. Additional Constraints

We are also able to incorporate additional types of constraints on the value of the function beyond simply direct observations of the function at individual times. Examples might include additional prior information, or observed information that might depend upon the value of the function at multiple times. For example, suppose that we have a belief that the gradient of the pH function should not change by more than x per million years. This can be encapsulated by modifying $\pi(\mathbf{f})$, the standard GP prior, to instead be $\pi'(\mathbf{f}) \propto \mathbb{1}_{[\max \text{gradient} < x]}\pi(\mathbf{f})$. To include this additional belief, we can simply sample from the standard (non-gradient-constrained) GP posterior, and then reject those realisations for which the maximum gradient is greater than x per million years. We note that, in practice, we estimate the maximum gradient of the function by sampling the GP extremely densely in time.

Aside: We can build up our posterior by using the GP posterior from a subset of the observations; and then use rejection sampling to adjust/update this for the full set of observations. Suppose that we observe $y_i = f(t_i) + \epsilon_i$ for $i = 1, \ldots, M$. We can sample from a reduced posterior considering all the observations excluding one, without loss of generality we suppose this is y_M:

$$g(\mathbf{f}|y_1, \ldots, y_{M-1}) \propto \prod_{i=1}^{M-1} p_\epsilon(y_i|f_i)\pi(\mathbf{f}).$$

To update this preliminary distribution to the full target posterior using all the observations, we sample from the reduced posterior $g(\mathbf{f}|y_1, \ldots, y_{M-1})$ ensuring we include the value at $f_M = f(t_M)$. We then accept this draw with probability $\frac{p_\epsilon(y_M|f_M)}{\max_f p_\epsilon(y_M|f_M)}$. Otherwise we sample from the reduced envelope again. This reduced-to-full approach is however likely to be much less efficient than sampling from an appropriate envelope based upon all the samples.

4. Rejection Sampling for $\delta^{11}\mathrm{B}_{\mathrm{sw}}$ with Diverse Constraints

When reconstructing $\delta^{11}\mathrm{B}_{\mathrm{sw}}$ we have multiple types of constraints: non-Gaussian observations, upper/lower bounds, and restrictions on the maximum rate of change over time. We are required to integrate all these varied constraints into our GP posterior. We illustrate how this is achieved in Figure S4. The large upper panel shows a hypothetical signal (in the thick black line) which is assumed unknown. We wish to reconstruct this function using a Gaussian Process and five noisy observations. Three observations, shown in blue, are subject to Gaussian noise (displayed at 1 standard deviation uncertainty). The other two observations, shown in green, are subject to non Gaussian noise reminiscent of a Tukey window - derived from a uniform distribution with Gaussian noise in the end members.

Here a Gaussian Process with prescribed hyperparameters (length scale of 15, noise scale of 30) assimilates the three observations with Gaussian uncertainty. This will be our envelope density. Three proposed samples are drawn from the Gaussian Process, shown in grey and labelled: **a** (dashed line), **b** (dotted line), and **c** (solid line). Looking at each sample in the upper panel, we see that:

- Gaussian Process sample **a** is completely inconsistent with the non-Gaussian observation around $t = 70$.
- Gaussian Process sample **b** is potentially consistent with both non-Gaussian observations, but is right at the limits of the possible outcomes for the constraint around $t = 30$.
- Gaussian Process sample **c** is consistent with both non-Gaussian observations - passing through a high probability region of both.

Each of the five observations is shown in a separate subpanel beneath the main time series. We might consider these as time slices through the main panel, displaying each observation probabilistically. The true value of the signal at these times is shown by the black horizontal line in each panel. The value of each Gaussian Process sample at the time slices is shown in the panels with the corresponding line style.

To assimilate observations with non-Gaussian uncertainties (shown in green), we use a rejection sampling strategy described above. To calculate a probability of acceptance, first each non-Gaussian distribution is scaled such the the maximum is equal to one (as shown in the lower panels). Then for each sample drawn from the Gaussian Process, the relative likelihood of the sample is calculated (this is shown numerically for each sample in the lower panels). The relative likelihood of each sample from the Gaussian Process is the product of the likelihoods at each of these individual timeslices.

- Gaussian Process sample **a** has a $1.0 \times 0.0 = 0$ probability of acceptance.
- Gaussian Process sample **b** has a $0.56 \times 1.0 = 0.56$ probability of acceptance.
- Gaussian Process sample **c** has a $1.0 \times 1.0 = 1.0$ probability of acceptance.

This strategy allows us to draw samples which are consistent with different types of observation - though we note there are potential failure conditions. If one of the non-Gaussian observations were much higher than the Gaussian observations, every sample would receive a 0 probability of acceptance. We mitigate against this failure condition by giving non-Gaussian constraints the possibility of being an outlier (as described above in Section 3.2).

In addition to the types of constraints shown above, we also place limitations on the rate of change in $\delta^{11}\mathrm{B}_{\mathrm{sw}}$. The same technique as shown for the non-Gaussian constraints is used to enforce these constraints. Using the same synthetic example, this would appear as in Figure S5.

Here the gradient in each of the samples is calculated using the first difference, and a weight for each sample can be determined by comparing this gradient to the constraints. In both the synthetic data example and the $\delta^{11}\mathrm{B}_{\mathrm{sw}}$ reconstruction we place a uniform prior on the gradient, which effectively describes the maximum rate of change (either in a negative or positive direction). This is displayed using horizontal bars in the large panel, within which the signal must fall, and each uniform window is plotted in individual subpanels underneath. We see that two samples remain within the imposed constraints, whereas the sample **a** is incompatible with both the earliest and latest constraint.

The prescribed maximum rate of change in both the synthetic example and the reconstruction of $\delta^{11}\mathrm{B}_{\mathrm{sw}}$ depends on time. In the synthetic example, the gradient is constrained in three places, with increasing acceptable range from ±0.2 units in the earliest constraint to ±0.6 units in the latest. For $\delta^{11}\mathrm{B}_{\mathrm{sw}}$, the maximum rate of change is constrained for each discretised age window (at a resolution of 0.1Myr), and grows linearly from 0.1‰/Myr in the modern day to 1‰/Myr at 100Ma to account for increasing uncertainty in this limit.

If we run the algorithm described above for 10,000 Gaussian Process samples, accepting and rejecting the proposed samples according to the rejection algorithm, we can obtain a set of realisations from the complete posterior that incorporate all the various forms of information we have on its value: the three Gaussian observations, the two non-Gaussian, and the gradient constraints. We can then summarise these using a median and 95% pointwise posterior probability window and compare agreement to the underlying original signal as shown in Figure S6. We see a good match, within the limitations imposed by not having many observations on which to base our reconstruction, and considering that each of these observations has substantial uncertainty.

5. Data File Description

5.1. Data Supplement S1

Data Supplement S1 is an excel file containing two worksheets. The first has every accepted reconstructed $\delta^{11}B_{sw}$ time series, with age in each column, and an independent statistical sample in each row. The second worksheet contains summary metrics, specifically the median and 5% and 95% quantiles of the time series. These quantiles give a sense of uncertainty at any individual time, and can be used to propagate uncertainties when targeting absolute pH reconstructions from $\delta^{11}B_4$ within a narrow time window. When looking at longer term trends, or robustly assessing uncertainty in *change* in pH, the full time series should be integrated by sampling from the time series presented in the former tab.

5.2. Data Supplement S2

Data Supplement S3 is an excel file containing three triples of worksheets (nine in total) which contain the data, summary metrics for fits, and 10,000 Gaussian Process samples for the evolution of $^{87/86}$Sr, δ^7Li, and $^{187/188}$Os (as shown in Figure 4 and Figure S1).

Strontium and lithium signals are taken from Misra and Froelich (2012). Osmium deserves special mention here as no Cenozoic compilation was found in an accessible format. Our Cenozoic $^{187/188}$Os record was constructed from previously published data in Josso et al. (2019); Klemm, Levasseur, Frank, Hein, and Halliday (2005); Oxburgh (1998); Oxburgh, Pierson-Wickmann, Reisberg, and Hemming (2007); Paquay, Ravizza, Dalai, and Peucker-Ehrenbrink (2008); Paquay, Ravizza, and Coccioni (2014); Pegram and Turekian (1999); Peucker-Ehrenbrink and Ravizza (2000, 2020); van der Ploeg et al.

(2018); Ravizza (1993); Ravizza and Turekian (1992); Ravizza and Peucker-Ehrenbrink (2003); Ravizza, Norris, Blusztajn, and Aubry (2001); Reusch, Ravizza, Maasch, and Wright (1998); Robinson, Ravizza, Coccioni, Peucker-Ehrenbrink, and Norris (2009). Ages of the data from Paquay et al. (2008) were adjusted to match the age model of the record of Paquay et al. (2014). Our $^{187/188}$Os compilation integrates data from pelagic sediments and Fe-Mn crusts into a single record. The trends are broadly consistent with those previously reported in Peucker-Ehrenbrink and Ravizza (2020), however we are able to produce a representative curve with propagated uncertainties using a Gaussian Process. For most signals in this work we have used the residence time of the element in question to determine the length scale of the Gaussian process, however in the case of osmium the residence time is too short (Oxburgh, 2001) for this to be viable given the current data density. Instead we choose a low (1 Myr) but still inflated value which bridges the gaps between data without overly smoothing the signal, in order to produce the curve we believe to be most representative.

5.3. Data Supplement S3

Data Supplement S3 is an excel file containing four worksheets describing summary metrics and 10,000 possible evolutions of $\delta^{11}B_4$ and pH (as shown in Figure 3 and Figure S1).

References

Josso, P., Parkinson, I., Horstwood, M., Lusty, P., Chenery, S., & Murton, B. (2019, May). Improving confidence in ferromanganese crust age models: A composite geochemical approach. *Chemical Geology*, *513*, 108–119. doi: 10.1016/j.chemgeo.2019

.03.003

Klemm, V., Levasseur, S., Frank, M., Hein, J. R., & Halliday, A. N. (2005, September). Osmium isotope stratigraphy of a marine ferromanganese crust. *Earth and Planetary Science Letters*, *238*(1), 42–48. doi: 10.1016/j.epsl.2005.07.016

Misra, S., & Froelich, P. N. (2012, February). Lithium Isotope History of Cenozoic Seawater: Changes in Silicate Weathering and Reverse Weathering. *Science*, *335*(6070), 818–823. doi: 10.1126/science.1214697

Oxburgh, R. (1998, June). Variations in the osmium isotope composition of sea water over the past 200,000 years. *Earth and Planetary Science Letters*, *159*(3), 183–191. doi: 10.1016/S0012-821X(98)00057-0

Oxburgh, R. (2001). Residence time of osmium in the oceans. *Geochemistry, Geophysics, Geosystems*, *2*(6). doi: 10.1029/2000GC000104

Oxburgh, R., Pierson-Wickmann, A.-C., Reisberg, L., & Hemming, S. (2007, November). 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. doi: 10.1016/j.epsl.2007.08.033

Paquay, F. S., Ravizza, G., & Coccioni, R. (2014, November). The influence of extraterrestrial material on the late Eocene marine Os isotope record. *Geochimica et Cosmochimica Acta*, *144*, 238–257. doi: 10.1016/j.gca.2014.08.024

Paquay, F. S., Ravizza, G. E., Dalai, T. K., & Peucker-Ehrenbrink, B. (2008, April). Determining Chondritic Impactor Size from the Marine Osmium Isotope Record. *Science*, *320*(5873), 214–218. doi: 10.1126/science.1152860

Pegram, W. J., & Turekian, K. K. (1999, December). 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. doi: 10.1016/S0016-7037(99)00308-7

Peucker-Ehrenbrink, B., & Ravizza, G. (2000, June). 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. doi: 10.1016/S0016-7037(99)00429-9

Peucker-Ehrenbrink, B., & Ravizza, G. E. (2020, January). Chapter 8 - Osmium Isotope Stratigraphy. In F. M. Gradstein, J. G. Ogg, M. D. Schmitz, & G. M. Ogg (Eds.), *Geologic Time Scale 2020* (pp. 239–257). Elsevier. doi: 10.1016/B978-0-12-824360-2.00008-5

Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian processes for machine learning.* Cambridge, Mass: MIT Press.

Ravizza, G. (1993, July). Variations of the187Os/186Os ratio of seawater over the past 28 million years as inferred from metalliferous carbonates. *Earth and Planetary Science Letters*, *118*(1), 335–348. doi: 10.1016/0012-821X(93)90177-B

Ravizza, G., Norris, R. N., Blusztajn, J., & Aubry, M. P. (2001). An osmium isotope excursion associated with the Late Paleocene thermal maximum: Evidence of intensified chemical weathering. *Paleoceanography*, *16*(2), 155–163. doi: 10.1029/2000PA000541

Ravizza, G., & Peucker-Ehrenbrink, B. (2003, May). 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. doi: 10.1016/S0012-821X(03)00137-7

Ravizza, G., & Turekian, K. K. (1992, May). The osmium isotopic composition of organic-rich marine sediments. *Earth and Planetary Science Letters*, *110*(1), 1–6. doi: 10.1016/0012-821X(92)90034-S

Reusch, D. N., Ravizza, G., Maasch, K. A., & Wright, J. D. (1998, July). Miocene seawater 187Os/188Os ratios inferred from metalliferous carbonates. *Earth and Planetary Science Letters*, *160*(1), 163–178. doi: 10.1016/S0012-821X(98)00082-X

Robinson, N., Ravizza, G., Coccioni, R., Peucker-Ehrenbrink, B., & Norris, R. (2009, May). 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. doi: 10.1016/j.epsl.2009.02.019

van der Ploeg, R., Selby, D., Cramwinckel, M. J., Li, Y., Bohaty, S. M., Middelburg, J. J., & Sluijs, A. (2018, July). Middle Eocene greenhouse warming facilitated by diminished weathering feedback. *Nature Communications*, *9*(1), 2877. doi: 10.1038/s41467-018-05104-9

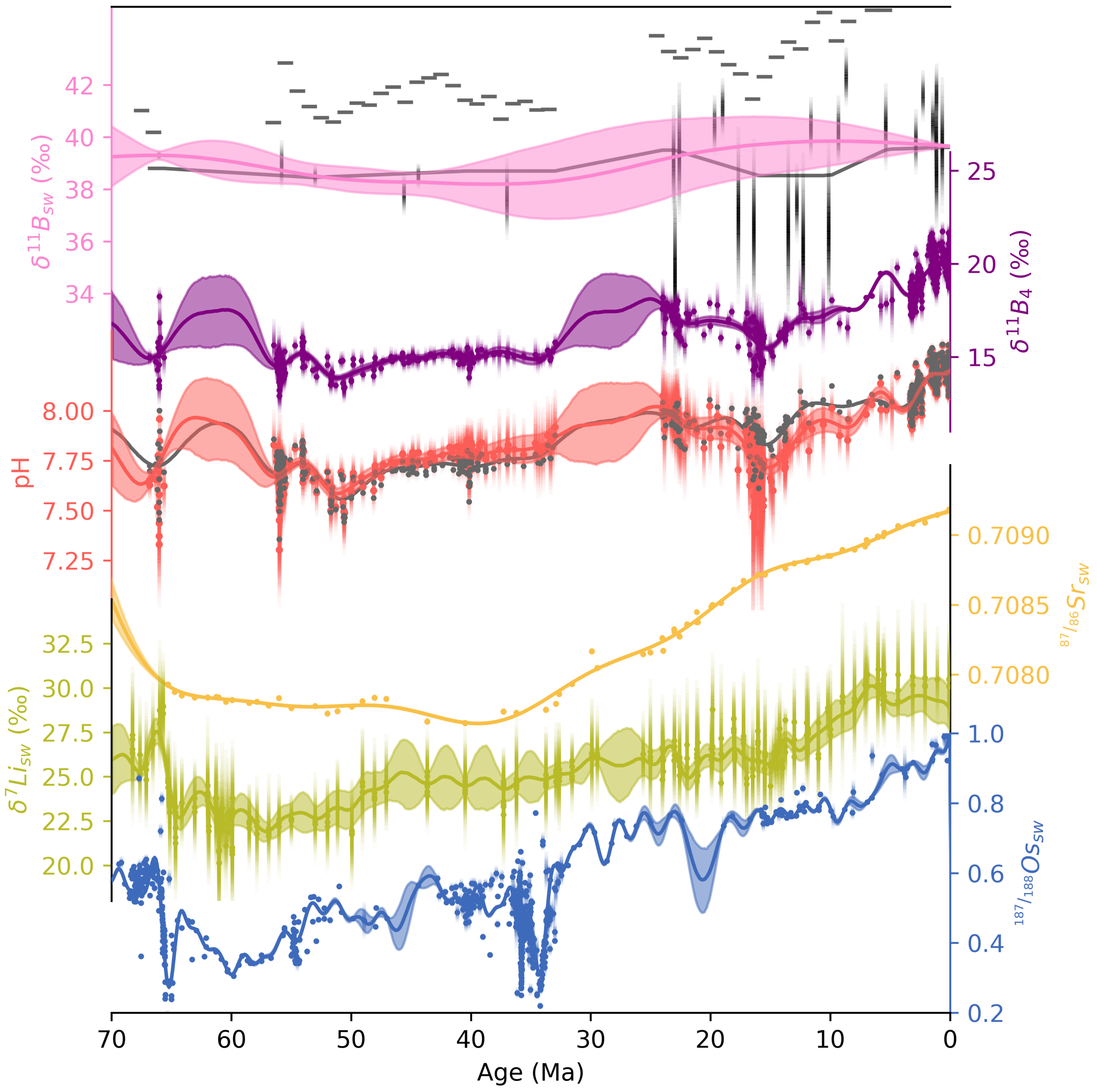


Figure S1: $\delta^{11}B_{sw}$ *(pink),* $\delta^{11}B_4$ *(purple), pH (red),* $^{87/86}Sr$ *(yellow),* δ^7Li *(green), and* $^{187/188}Os$ *(blue) are shown here with the same style as shown in the two separate plots in the main text. We provide an large summary figure here for easy comparison of the six signals.*

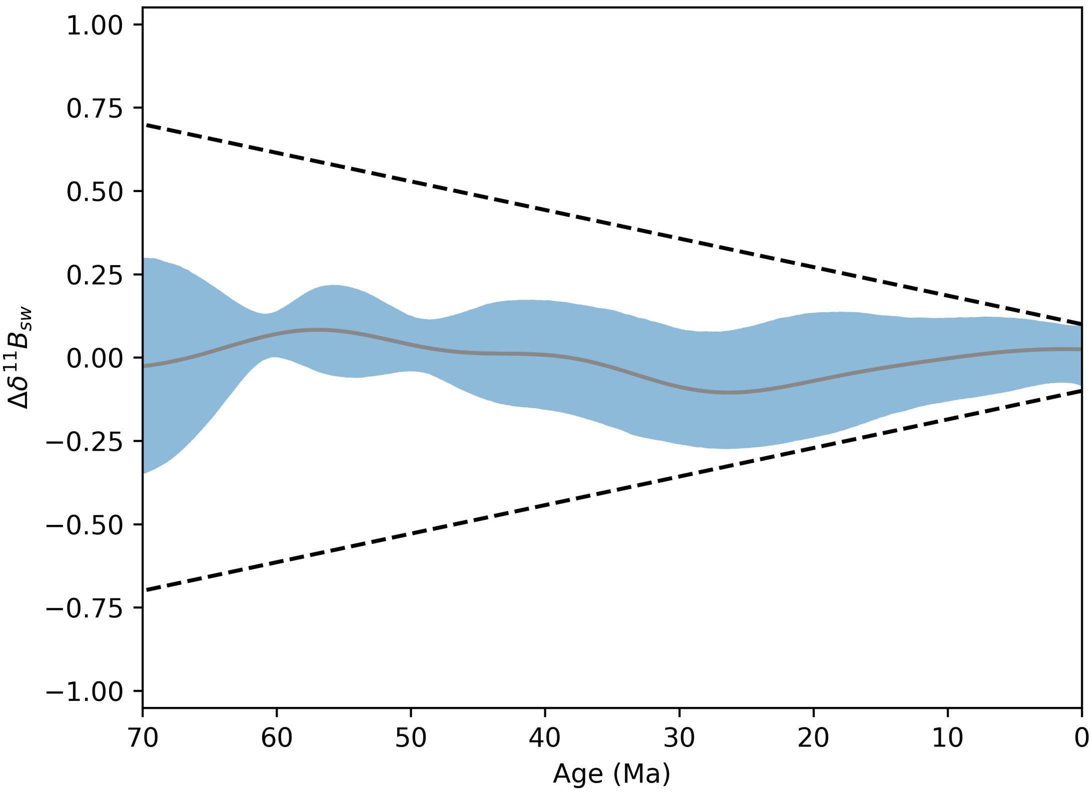


Figure S2: The temporal gradient in our $\delta^{11}B_{sw}$ *samples is shown by the blue window, with the mean average shown in grey. Our imposed limitation on the gradient of* $\delta^{11}B_{sw}$ *through time is shown by the dotted black lines. Any sample drawn outside of these bounds would be rejected. It can be seen that the limitations have most influence between the Neogene and modern, and further back do not result in rejection of any samples.*

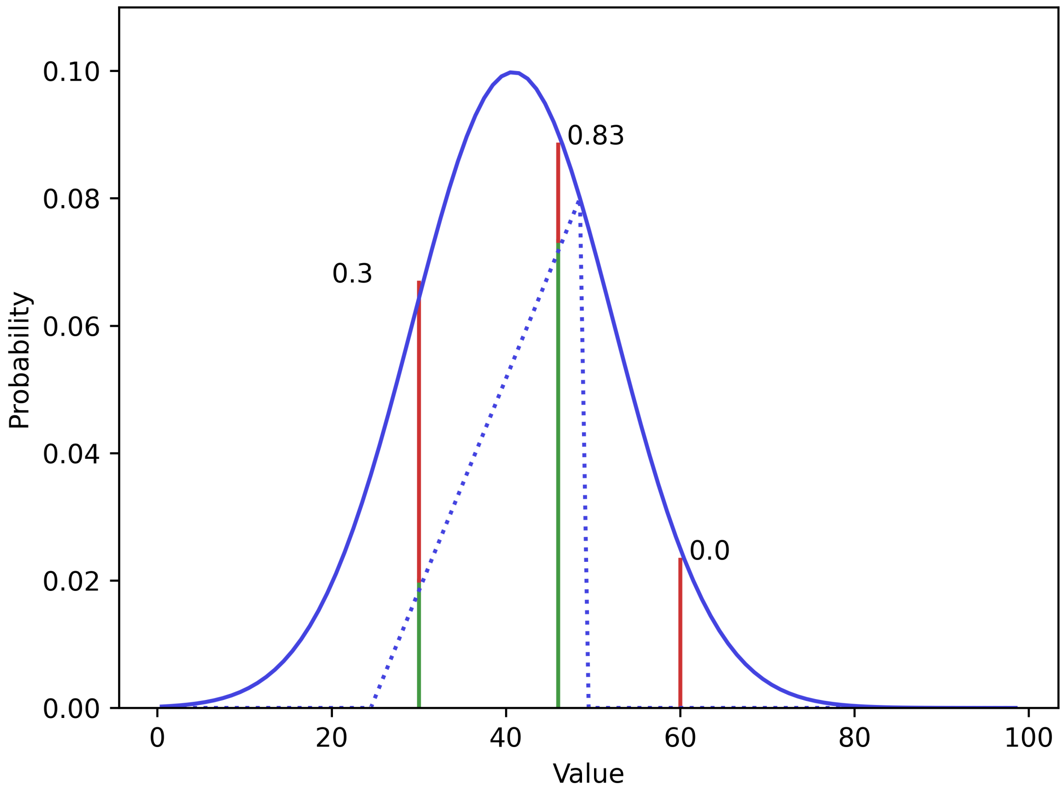


Figure S3: An illustration of rejection sampling. We aim to sample from the dotted blue saw-tooth density (shown as a dotted blue line) using a Gaussian distribution as the envolope (shown in solid blue). The Gaussian envelope has been rescaled from a standard Gaussian distribution so it encapsulates the target sawtooth density. To obtain a sample from the sawtooth distribution, we first sample a value envelope density. We show hypothetical three values $z_1 = 30$, $z_2 = 45$ *and* $z_3 = 60$. *The probability of accepting each sample* z_i *as a draw from the sawtooth distribution is then the ratio of the height of the target (dotted line) compared to the height of the envelope (solid blue line). These probabilities are 0.3, 0.83, and 0 respectively for our three hypothetical* z_i *samples.*

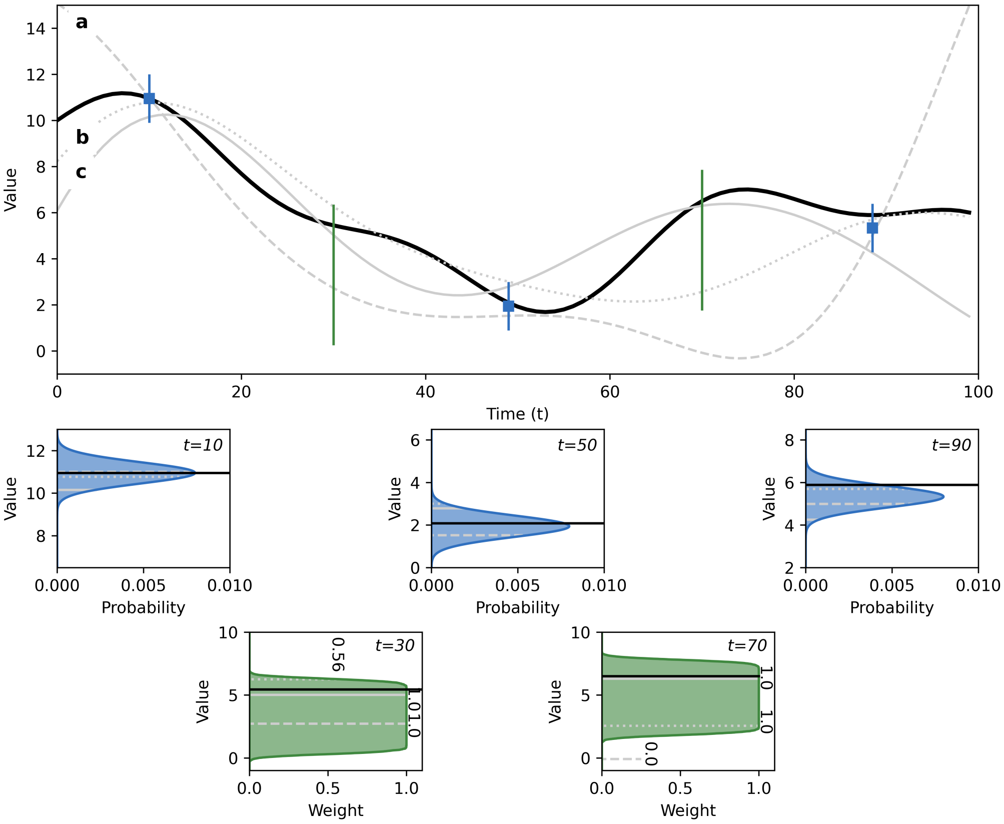


Figure S4: Calculating the GP posterior by rejection sampling that combines multiple constraints of both Gaussian and non-Gaussian types. Multiple samples are taken using a Gaussian Process conditioned on only the Gaussian constraints (samples are shown in the grey lines), and for each we quantify the probability of that sample at each data constraint (the blue and green windows). The probability of each sample is the product of the probabilities of that sample at each data constraint, meaning that sample a is rejected (it does not match the fourth data constraint), while other are likely to be accepted. This is described further in Section 3.3.

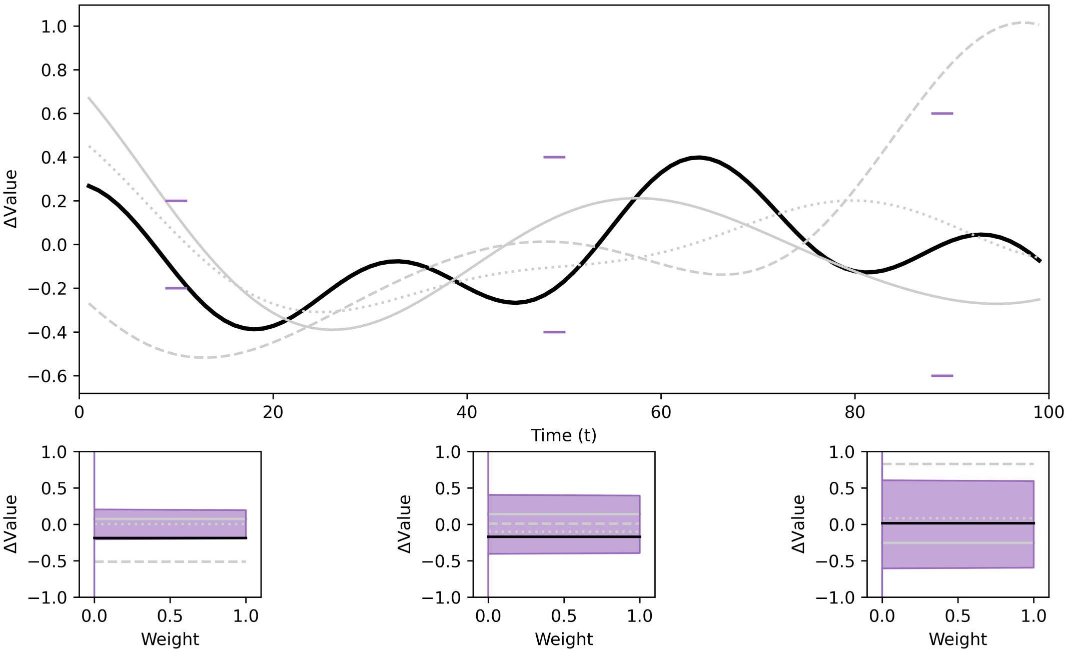


Figure S5: Incorporating constraints on the gradient into estimation of $\delta^{11}B_{sw}$*. The main plot shows the estimated gradient* Δ *Value of our function over time. The subplots shown the gradient constraint we impose upon the signal, and the probability of observing each statistical sample at that time. Note that here we impose gradient constraint only a three discrete locations, whereas in the main text we apply a continuous limitation on the rate of change in* $\delta^{11}B_{sw}$*.*

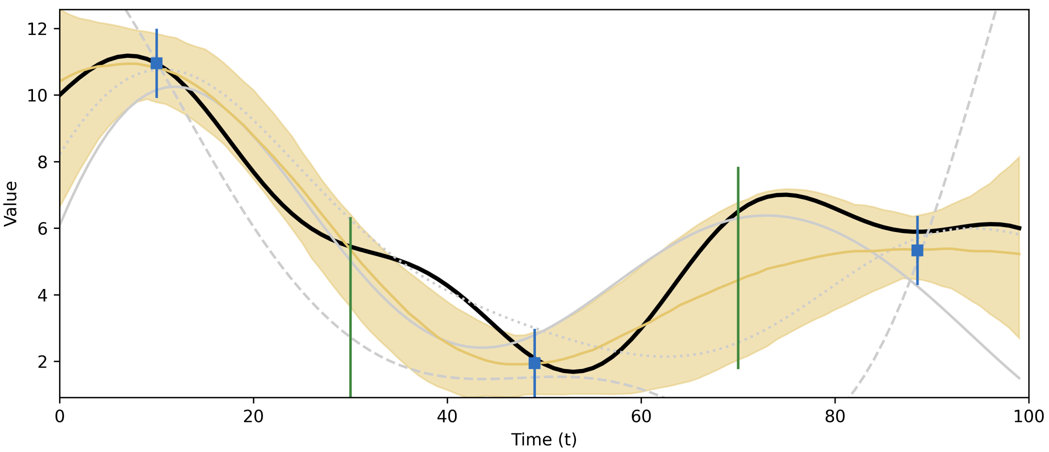


Figure S6: Reconstruction of the function shown in black incorporating information from noisy Gaussian observations (in blue) and non-Gaussian observations (in green) and gradient restrictions (in purple in Figure S5). The yellow line shows our central estimate, with a 95% confidence interval shown in the yellow shaded region.