Alkenone isotopes show evidence of active carbon concentrating mechanisms in coccolithophores as aqueous carbon dioxide concentrations fall below 7 mol L\(^{-1}\).
Alkenone isotopes show evidence of active carbon concentrating mechanisms in coccolithophores as aqueous carbon dioxide concentrations fall below 7 µmol L\(^{-1}\)

Marcus P. S. Badger
School of Environment, Earth & Ecosystem Sciences, The Open University, Milton Keynes, MK7 6AA, UK

Correspondence: Marcus P. S. Badger (marcus.badger@open.ac.uk)

Received: 25 September 2020 – Discussion started: 2 October 2020
Revised: 6 January 2021 – Accepted: 8 January 2021 – Published: 15 February 2021

Abstract. Coccolithophores and other haptophyte algae acquire the carbon required for metabolic processes from the water in which they live. Whether carbon is actively moved across the cell membrane via a carbon concentrating mechanism, or passively through diffusion, is important for haptophyte biochemistry. The possible utilization of carbon concentrating mechanisms also has the potential to over-print one proxy method by which ancient atmospheric CO\(_2\) concentration is reconstructed using alkenone isotopes. Here I show that carbon concentrating mechanisms are likely used when aqueous carbon dioxide concentrations are below 7 µmol L\(^{-1}\). I compile published alkenone-based CO\(_2\) reconstructions from multiple sites over the Pleistocene and recalculate them using a common methodology, which allows comparison to be made with ice core CO\(_2\) records. Interrogating these records reveals that the relationship between proxy CO\(_2\) and ice core CO\(_2\) breaks down when local aqueous CO\(_2\) concentration falls below 7 µmol L\(^{-1}\). The recognition of this threshold explains why many alkenone-based CO\(_2\) records fail to accurately replicate ice core CO\(_2\) records, and it suggests the alkenone proxy is likely robust for much of the Cenozoic when this threshold was unlikely to be reached in much of the global ocean.

1 Introduction

Alkenones are long-chain (C\(_{37−39}\)) ethyl and methyl ketones (Fig. 1; Brassell et al., 1986; Rechka and Maxwell, 1987) produced by a restricted group of photosynthetic haptophyte algae (Conte et al., 1994). Produced by a narrow group of organisms which live exclusively in the photic zone, alkenones allow probing of algal biogeochemistry, and as alkenones are often preserved in the sedimentary record, alkenones can also provide information about past environmental conditions.

Two main proxy systems based on alkenone geochemistry exist: one allows reconstruction of sea surface temperature (SST) and relies on the changing degree of unsaturation of the C\(_{37}\) alkenone (U\(_K^{37}\)) (Brassell et al., 1986), whilst a second for atmospheric CO\(_2\) concentration is based on reconstructing the isotopic fractionation which takes place during photosynthesis (\(\varepsilon_p\)) (Laws et al., 1995; Bidigare et al., 1997). It is the second system using the stable carbon isotopic composition of the preserved alkenones for reconstructing atmospheric CO\(_2\) concentration (referred to throughout as CO\(_2(\varepsilon_{p-alk})\)) which is the focus of this study.

In the modern ocean, alkenones are produced primarily by two dominant coccolithophore species: *Emiliania huxleyi* and *Gephyrocapsa oceanica*. *E. huxleyi* first appeared 290 kyr ago and began to dominate over *G. oceanica* around 82 kyr ago (Gradstein et al., 2012; Raffi et al., 2006). However alkenones are commonly found in sediments throughout the Cenozoic, with the oldest reported detections from mid-Albian-aged black shales (Farrimond et al., 1986). Prior to the evolution of *G. oceanica*, alkenones were most likely produced by other closely related species from the Noelaerhabdaceae family (Marlowe et al., 1990; Volkmann, 2000). Micropalaeontological and molecular data split the coccolith-bearing haptophytes into two distinct phylogenetic clades: the Isochrysidales and Coccolithales. The Isochrysidales contain the modern alkenone-producing taxa, including *E. huxleyi* and *G. oceanica*, and fossil reticulofenestrids. Meanwhile the non-alkenone-producers are separated into
the order Coccolithales, which includes *Coccolithus pelagicus* and *Calcidiscus leptoporus* along with most other coccolithophores.

Proxies for atmospheric CO\(_2\) concentration – including CO\(_2\)(\(\varepsilon_p-alk\)), those based on \(\delta^{13}\)B of planktic foraminifera, geochemical modelling, and stomatal density – broadly agree that over the Cenozoic atmospheric \(p\text{CO}_2\) declined from high levels (> 1000 \(\mu\text{atm}\)) in the "greenhouse" worlds of the Palaeocene and Eocene to close to modern-day values (around 400 \(\mu\text{atm}\)) in the Pliocene (Pagani et al., 2005, 2011; Pearson et al., 2009; Anagnostou et al., 2016; Foster et al., 2017; Sosdian et al., 2018; Super et al., 2018; Zhang et al., 2013; Beerling and Royer, 2011). However, recent discrepancies have emerged between CO\(_2\)(\(\varepsilon_p-alk\)) and other CO\(_2\) proxies at the < 400 \(\mu\text{atm}\) atmospheric CO\(_2\) concentrations of the Pleistocene (i.e. Badger et al., 2019, 2013a, and compare Badger et al., 2013b, and Pagani et al., 2009, with Martínez-Botí et al., 2015). Whilst the long-standing differences between alkenone (Pagani et al., 1999), \(\delta^{13}\)B (Foster et al., 2012), and stomatal proxies (Kürschner et al., 2008) in the Miocene CO\(_2\) reconstructions have been partially resolved with new SST records (Super et al., 2018), differences remain in the Pliocene (Pagani et al., 2009; Badger et al., 2013b; Martínez-Botí et al., 2015) and Pleistocene (Badger et al., 2019).

**Carbon concentrating mechanisms**

One plausible reason for the discrepancies between CO\(_2\)(\(\varepsilon_p-alk\)) and other proxies for atmospheric CO\(_2\) is the operation of active carbon concentrating mechanisms (CCMs) in haptophytes. These are potentially important as CO\(_2\)(\(\varepsilon_p-alk\)) assumes purely passive uptake of carbon into the haptophyte cell purely via diffusion (Laws et al., 1995; Bidigare et al., 1997). The potential for CCMs to affect CO\(_2\)(\(\varepsilon_p-alk\)) has long been known (Laws et al., 1997, 2002; Cassar et al., 2006), and recent work has refocussed efforts on understanding CCMs in CO\(_2\)(\(\varepsilon_p-alk\)) (Bolton et al., 2012; Bolton and Stoll, 2013; Stoll et al., 2019; Zhang et al., 2019, 2020). Coccolithophores are thought to have low-efficiency CCMs – especially compared to diatoms, dinoflagellates, and *Phaeocystis* – with evidence that CCMs play a minor role in coccolithophore biochemistry in the CO\(_2\)-replete worlds of the early Cenozoic (Bolton et al., 2012; Reinfelder, 2011). Direct evidence from experimentation with the marine diatom *Phaeodactylum tricornutum* suggests that both passive diffusive uptake and active CCMs operate at the same time, with active uptake used to moderate internal cell CO\(_2\) concentrations to minimize energy use during transport to carboxylation sites (Laws et al., 1997). CO\(_2\), unlike some other nutrients, is abundant within the water column, especially when considering the dissolved inorganic carbon (DIC) reservoir which includes bicarbonate (HCO\(_3^-\)), carbonate (CO\(_3^{2-}\)), and dissolved CO\(_2\)(\(\text{CO}_2\)(aq)). However, due to the relatively slow diffusion of dissolved [CO\(_2\)](aq) through water and the slow kinetics of the bicarbonate-to-[CO\(_2\)](aq) transformation, surface water [CO\(_2\)](aq) can still be depleted by photosynthetic activity. This can become particularly problematic in species which form blooms and at the cell boundary of species with limited motility. It should be no surprise therefore that many marine photosynthetic organisms have evolved with mechanisms to concentrate carbon within the cell.

The enzyme carbonic anhydrase (CA) can catalyse the dehydration of HCO\(_3^-\) to [CO\(_2\)](aq) to speed up availability of carbon if the [CO\(_2\)](aq) reservoir is depleted and has been observed in several haptophytes, including coccolithophores (Rost et al., 2003; Riebesell et al., 2007). The exact contribution of CA remains unclear, but two possible mechanisms for CCMs have been postulated (Reinfelder, 2011): (1) CA catalyses dehydration of HCO\(_3^-\) at the cell surface, which then allows increased CO\(_2\) to diffuse into the cell passively, and (2) HCO\(_3^-\) is transported into the cell and then converted by CA. Both of these options will likely impact the CO\(_2\)(\(\varepsilon_p-alk\)) proxy, firstly by changing the effective [CO\(_2\)](aq) within the cell (and so impacting \(\varepsilon_p\)) and secondly by imparting another carbon isotopic fractionation during CA catalysis which is not considered by the CO\(_2\)(\(\varepsilon_p-alk\)) proxy system. However CA activity in coccolithophores does not appear to be regulated by CO\(_2\) as it is in diatoms and *Phaeocystis* (Rost et al., 2003), which may indicate a less-well-developed CCM in coccolithophores.

Calciifying coccolithophores (which include alkenone producers *E. huxleyi* and *G. oceanica*) may be able to utilize HCO\(_3^-\) directly as a carbon source (Trimborn et al., 2007), with precipitation of CaCO\(_3\) providing an acid for the dehydration of HCO\(_3^-\), but this still requires sufficient HCO\(_3^-\) entering the cell, and it is unclear whether calcification aids DIC acquisition (Riebesell et al., 2000; Zondervan et al., 2002). The light-dependent leak of carbon (as CO\(_2\) and DIC) back from haptophyte cells (including the coccolithophore *E. huxleyi*) to seawater (Tchernov et al., 2003) suggests that CCMs are energy intensive and can concentrate DIC within the cell. Even with active CCMs, it appears that in the ocean...
cocolithophores are CO₂ limited under some circumstances (Riebesell et al., 2007).

2 Materials and methods

Calculating CO₂ from alkenone δ¹³C values: the CO₂(δ₁³C – alk) proxy

In this study I use the now large number of published CO₂(δ₁³C – alk) records which overlap with ice core records of atmospheric CO₂ concentration (Tables 1 and 2) to explore the relationship between CO₂(δ₁³C – alk) and CCMs in the Pleistocene, where our understanding of atmospheric CO₂ concentration is best.

Multiple records of CO₂(δ₁³C – alk) have been published for the Pleistocene (Fig. 2, Table 1), allowing direct comparison with ice-core-based CO₂ records (Table 2). These records are globally distributed in longitude but are concentrated at low-latitude sites, largely as there is a general preference for sites which have (in the modern ocean) surface waters close to equilibrium with the atmosphere (Fig. 2, Table 1). In longer-term palaeoclimate studies there has also been a preference for low-latitude gyre sites in the belief that these sites are more likely to be oceanographically stable over long time intervals (Pagani et al., 1999). Most of the records included here (Table 1, Fig. 2) were generated with the aim to reconstruct atmospheric CO₂ concentration; however one, the MANOP Site C of Jasper et al. (1994), was used to explicitly reconstruct changing disequilibrium due to oceanographic frontal changes over time and so is excluded from the following analysis.

Whilst these sites do only span a relatively small latitudinal extent, the diversity of settings does allow for investigation of any secondary controls on alkenone δ¹³C values (δ¹³Calkalene) – in particular, differences in oceanographic setting and SST to test the hypothesis that low [CO₂]aq breaks the relationship between δ¹³Calkalene and atmospheric CO₂ concentration, as might be expected if haptophytes are able to actively take up carbon from seawater to meet metabolic demand (i.e. activate CCMs).

To facilitate fair comparison between sites and consistent comparison with the ice core records, all CO₂(δ₁³C – alk) records were recalculated using a consistent approach. The approach is based on Bidigare et al. (1997), which updated the initial approach of Jasper and Hayes (1990) to CO₂(δ₁³C – alk). This approach removes some additional corrections used in the original publication of the records (such as growth rate adjustment for NiOPG 464; Palmer et al., 2010) but does allow for direct comparison to be made. For all sites the “b” term was estimated using modern-day surface [PO₄³⁻] (Bidigare et al., 1997; Pagani et al., 2009).

An overview of how CO₂(δ₁³C – alk) data are typically generated is given in Badger et al. (2013b). Briefly, to calculate εₚ requires the stable carbon isotopic composition of the dissolved CO₂ (δ¹³CCO₂(aq)) and haptophyte biomass (δ¹³Corg). The isotopic fractionation between δ¹³Calkalene and δ¹³Corg is first corrected assuming a constant fractionation (εalkalene) of 4.2‰ (Garcia et al., 2013; Popp et al., 1998; Bidigare et al., 1997):

εalkalene = \frac{\delta^{13}C_{alkalene} + 1000}{\delta^{13}C_{org} + 1000} - 1. \quad (1)

The isotopic composition of DIC is estimated using (ideally) the δ¹³C value of planktic foraminifera and the temperature-dependent fractionation between calcite and CO₂(g) experimentally determined by Romanek et al. (1992), where T is sea surface temperature in degrees Celsius (SST):

εcalcite–CO₂(g) = 11.98 - 0.12T. \quad (2)

The value of the carbon isotopic composition of CO₂(g) (δ¹³CCO₂(g)) can then be calculated:

δ¹³CCO₂(g) = \frac{\delta^{13}C_{carbonate} + 1000}{\varepsilon_{\text{calcite–CO}_2(g)} + 1000} - 1000. \quad (3)

From this δ¹³CCO₂(aq) can be calculated using the relationship experimentally determined by Mook et al. (1974),

εCO₂(aq)–CO₂(g) = \frac{-373}{T + 273.15} + 0.19, \quad (4)
Table 1. Sites with Pleistocene $\text{CO}_2(\varepsilon_p-\text{alk})$ records. Note that the MANOP Site C record was generated to track changes in surface water-atmosphere equilibrium, not atmospheric $p$CO$_2$, so, although it is included here for completeness, it is not included in the analysis. Distance from the coast is calculated from the intermediate-resolution version of GSHHG and computed using Generic Mapping Tools (Wessel and Smith, 1996; Wessel et al., 2019).

<table>
<thead>
<tr>
<th>Site</th>
<th>Age interval (kyr)</th>
<th>Water depth (m)</th>
<th>Distance from coast (km)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>05PC-21</td>
<td>0.5–188</td>
<td>1721</td>
<td>108</td>
<td>Bae et al. (2015)</td>
</tr>
<tr>
<td>DSDP 619</td>
<td>3–92</td>
<td>2259</td>
<td>489</td>
<td>Jasper and Hayes (1990)</td>
</tr>
<tr>
<td>NIOP 464</td>
<td>7.8–29</td>
<td>1470</td>
<td>333</td>
<td>Palmer et al. (2010)</td>
</tr>
<tr>
<td>ODP 999</td>
<td>111–258</td>
<td>2839</td>
<td>249</td>
<td>Badger et al. (2019)</td>
</tr>
<tr>
<td>ODP 925</td>
<td>20–580</td>
<td>3042</td>
<td>626</td>
<td>Zhang et al. (2013)</td>
</tr>
<tr>
<td>MANOP Site C</td>
<td>0.8–253</td>
<td>4287</td>
<td>998</td>
<td>Jasper et al. (1994)</td>
</tr>
<tr>
<td>GeoB 1016-3</td>
<td>1.3–196</td>
<td>3410</td>
<td>185</td>
<td>Andersen et al. (1999)</td>
</tr>
</tbody>
</table>

The full record for ODP Site 925 extends to 38.62 Ma.

Table 2. Sources of ice core data, as compiled by Bereiter et al. (2015). WAIS – West Antarctic Ice Sheet; TALDICE – TALos Dome Ice CorE; and EDML – EPICA Dronning Maud Land. Age given as gas age relative to 1950.

<table>
<thead>
<tr>
<th>Age interval (kyr)</th>
<th>Ice core location</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>−0.051 to 1.8</td>
<td>Law Dome</td>
<td>Rubino et al. (2013)</td>
</tr>
<tr>
<td>1.8–2</td>
<td>Law Dome</td>
<td>MacFarling Meure et al. (2006)</td>
</tr>
<tr>
<td>11–22</td>
<td>WAIS</td>
<td>Marcott et al. (2014)</td>
</tr>
<tr>
<td>22–40</td>
<td>Siple Dome</td>
<td>Ahn and Brook (2014)</td>
</tr>
<tr>
<td>40–60</td>
<td>TALDICE</td>
<td>Bereiter et al. (2012)</td>
</tr>
<tr>
<td>60–115</td>
<td>EDML</td>
<td>Bereiter et al. (2012)</td>
</tr>
<tr>
<td>105–155</td>
<td>Dome C</td>
<td>Schneider et al. (2013)</td>
</tr>
<tr>
<td>155–393</td>
<td>Vostok</td>
<td>Petit et al. (1999)</td>
</tr>
</tbody>
</table>

\[
\delta^{13}\text{CO}_2(\text{aq}) = \left( \frac{\varepsilon_\text{CO}_2(\text{aq}) - \varepsilon_\text{CO}_2(\text{g})}{1000} + 1 \right) \times \left( \frac{\delta^{13}\text{C}_{\text{CO}_2(\text{aq})}}{\delta^{13}\text{C}_{\text{CO}_2(\text{g})}} + 1000 \right) - 1000. \tag{5}
\]

Finally $\varepsilon_p$ can be calculated:

\[
\varepsilon_p = \left( \frac{\delta^{13}\text{C}_{\text{CO}_2(\text{aq})} + 1000}{\delta^{13}\text{C}_{\text{CO}_2(\text{g})} + 1000} \right) \times 1000. \tag{6}
\]

and from that $[\text{CO}_2]_{(\text{aq})}$ is calculated using the isotopic fractionation during carbon fixation ($\varepsilon_f$) and $b$, which represents the summation of physiological factors:

\[
[\text{CO}_2]_{(\text{aq})} = \frac{b}{\varepsilon_f - \varepsilon_p}. \tag{7}
\]

Here $\varepsilon_f$ is assumed to be a constant 25%o (Bidigare et al., 1997). In the modern ocean the $b$ term, which accounts for physiological factors such as cell size and growth rate, shows a close correlation with $[\text{PO}_4^{3–}]$ (Bidigare et al., 1997; Pagani et al., 2009). However, the relationship between $b$, growth rate, and $[\text{PO}_4^{3–}]$ has recently been questioned (Zhang et al., 2019, 2020) but for the purposes of this analysis is assumed to hold. This is discussed further below. Values for SST, $\delta^{13}\text{C}_{\text{alkenone}}$, $\delta^{13}\text{C}_{\text{carbonate}}$, salinity, and $[\text{PO}_4^{3–}]$ are either taken from the original publications or estimated from modern ocean estimates (Takahashi et al., 2009; Antonov et al., 2010; Garcia et al., 2013; Locarnini et al., 2013).

Providing that the atmosphere is in equilibrium with surface water, the concentration of atmospheric $p$CO$_2$ can be calculated from $[\text{CO}_2]_{(\text{aq})}$ (and vice versa if atmospheric CO$_2$ concentration is known) using Henry’s law:

\[
p\text{CO}_2 = \frac{[\text{CO}_2]_{(\text{aq})}}{K_H}. \tag{8}
\]

The solubility coefficient ($K_H$) is dependent on salinity and SST, and here it is calculated following the parameterization of Weiss (1970, 1974).
concentration of dissolved carbon (not only atmospheric CO$_2$ sphere, obtaining their carbon from dissolved carbon. As it is important in controlling the accuracy of CO$_2$ concentration, it is therefore likely that the proxy controlling the equilibrium state between surface waters at the temperature, alkalinity, and other oceanographic factors which control the equilibrium between CO$_2$ and CO$_2$ at some sites and at some times, but not throughout. Sites 05-PC21 (Baee et al., 2015) and DSDP Site 619 (Jasper and Hayes, 1990) perform quite well throughout, whilst ODP Site 999 (Badger et al., 2019) and NIOP 464 (Palmer et al., 2010) only appear to agree at higher values of CO$_2$, and at ODP Site 925 (Zhang et al., 2013) and GeoB 1016-3 (Andersen et al., 1999) there is very little overlap between the two methods of reconstructing atmospheric CO$_2$ concentration.

To explore whether [CO$_2$(aq)] is an important influence on CO$_2$(alk) -- I calculate predicted [CO$_2$(aq)] –[CO$_2$(alk)] for each of the samples. To calculate [CO$_2$(aq)]–predicted, the time-equivalent value of atmospheric CO$_2$ concentration from the ice core record is used in combination with Eq. (8) to calculate [CO$_2$(aq)] at the time of alkenone production for each sample. Reconstructed estimates of SST and salinity are used as for CO$_2$(alk) above, along with any estimated surface water–atmosphere disequilibrium. Points in Fig. 4 are then coloured by [CO$_2$(aq)]–predicted.

Inspection of Fig. 4 suggests a connection between [CO$_2$(aq)]–predicted and the skill of CO$_2$(alk) to reconstruct atmospheric CO$_2$ concentration. The points clustering around the 1:1 line are lighter in colour (so with higher [CO$_2$(aq)]–predicted), whilst points falling away from the 1:1 line have lower [CO$_2$(aq)]–predicted. To explore this relationship, I progressively restricted the included samples on the basis of [CO$_2$(aq)]–predicted and at each stage calculated a Pearson correlation coefficient (r) and coefficient of determination (r$^2$) for each subset. Under this analysis the correlation progressively increased as more of the low [CO$_2$(aq)]–predicted samples were excluded (Fig. 5). All analyses were performed in R (R Core Team, 2020) using RStudio (RStudio Team, 2020). This suggests that the fidelity of the CO$_2$(alk) depends on the concentration of [CO$_2$(aq)], improving at higher levels of [CO$_2$(aq)].

To further investigate this potential relationship, I progressively exclude samples based on [CO$_2$(aq)]–predicted with a step size of 0.05 µmol L$^{-1}$, again calculating Pearson correlation coefficients and coefficients of determination between ice core and CO$_2$(alk) for each subsample of the population. The result is shown in Fig. 6. Here the analysis shows, similar to Fig. 5, that, as the samples with lowest [CO$_2$(aq)]–predicted are progressively removed, the correlation between ice core and CO$_2$(alk) increases. Furthermore, this continues only up until [CO$_2$(aq)]–predicted reaches

Figure 3. Compiled CO$_2$(alk)–based estimates of atmospheric CO$_2$ concentration over the past 260 kyr (blue circles), with the ice core compilation of Bereiter et al. (2015) shown as the solid black line. Full sources for the ice core and CO$_2$(alk) records are in Tables 1 and 2.

3 Results

3.1 Multi-site comparisons between CO$_2$(alk) and the ice core records

Across the six sites included in this analysis, there are 217 CO$_2$(alk)–based estimates of atmospheric CO$_2$ concentration over the past 260 kyr for comparison with the ice core records (Table 2; Bereiter et al., 2015). When all CO$_2$(alk) estimates are considered together over 260 kyr, this compilation of proxy-based records fails to replicate the ice core record (Fig. 3). This has already been noted at specific sites (e.g. Site 999 in the Caribbean; Badger et al., 2019), but this is the first time that all available records coincide with the Pleistocene ice core records have been compiled using a common methodology. Notably the CO$_2$(alk)–based estimates are rarely lower than time-equivalent ice core estimate, but frequently higher. Given that haptophytes require carbon to satisfy metabolic demand, this is perhaps unsurprising; if at times of low carbon availability haptophytes can switch from passive to active uptake to satisfy metabolic demand, this is perhaps unsurprising; if at times of low carbon availability haptophytes can switch from passive to active uptake to satisfy metabolic demand, it would be times of low atmospheric CO$_2$ concentration (and so lower [CO$_2$(aq)]) when the active uptake is most likely to be needed. As CO$_2$(alk)–based estimates of atmospheric CO$_2$ concentration rely on the assumption of a purely diffusive uptake of carbon, it is therefore likely that the proxy would perform worse at times of low atmospheric CO$_2$ concentration.

The haptophytes do not directly interact with the atmosphere, obtaining their carbon from dissolved carbon. As it is not only atmospheric CO$_2$ concentration which controls the concentration of dissolved carbon ([CO$_2$(aq)]) but also temperature, alkalinity, and other oceanographic factors which control the equilibrium state between surface waters at the atmosphere (Fig. 2), the multiple sites in different settings now give the opportunity to test whether other factors are important in controlling the accuracy of CO$_2$(alk).

To produce time-equivalent estimates of atmospheric CO$_2$ concentration for comparison with the ice core records, a simple linear interpolation of the Bereiter et al. (2015) compilation was initially used (Fig. 4). This assumes that both the age model of the ice core and the published age models of the sites are correct and equivalent. This is almost certainly not the case, and so for the calculations below, a ±3000 year uncertainty is included for ages of both the ice core and CO$_2$(alk) values. Figure 4 shows that CO$_2$(alk)–based atmospheric CO$_2$ concentration agree with ice core CO$_2$ at some sites and at some times, but not throughout.
7 μmol L$^{-1}$. Above this, the coefficient of determination plateaus, until the subsample reaches such a small size that spurious correlations become important (Fig. 6b).

### 3.2 Sensitivity and uncertainty tests

It is possible that the pattern seen in Fig. 6b could emerge from a dataset shaped with increasing density surrounding the 1 : 1 correlation line without being driven by changes in [CO$_2$]$^{aq}_{(p-alk)}$-predicted. To explore this possibility, I ran a series of sensitivity experiments. In these, rather than reducing the sample by filtering by [CO$_2$]$^{aq}_{(p-alk)}$-predicted, the whole dataset (Table 1) was randomly ordered and then stepwise subsampled. To make this equivalent to the [CO$_2$]$^{aq}_{(p-alk)}$-predicted analysis above, I set the size of each subsample to be equal to each step in the original analysis. This produces a randomly selected but same-sized subsample such that the size of the subsample reduces in the same way as shown in Fig. 6b. Pearson correlation coefficients and coefficients of determination were calculated for each subsample as above, and I repeated this 1000 times, with the order of each sample randomized each time.

To allow for possible age model uncertainties, a 3000-year (1σ) uncertainty was also applied to each sample. This uncertainty was applied to the age of each sample prior to sampling of the ice core record, and it is applied as a normally distributed uncertainty. Uncertainty in CO$_2$$(ε_{p-alk})$ measurements is typically calculated using Monte Carlo modelling of all the parameters (i.e. Pagani et al., 1999; Badger et al., 2013a, b); however this was not done in all the published work (Table 1), and some differences in approach were found across the published work. Therefore to create CO$_2$$(ε_{p-alk})$ uncertainty estimates for each value in this study, I emulate the uncertainties based on the CO$_2$$(ε_{p-alk})$ value. I built a simple emulator (Fig. 7) by running Monte Carlo uncertainty estimates for all of the included datasets (Table 1) using the same estimates of uncertainty for each variable in the CO$_2$$(ε_{p-alk})$ calculation as applied in Badger et al. (2013a, b). This then allows the uncertainty to be included in the [CO$_2$]$^{aq}_{(p-alk)}$-predicted calculation as well as CO$_2$$(ε_{p-alk})$, and it allowed for uncertainty estimates to be site-ambivalent.

The result is shown in Fig. 6c and d, and it suggests that the 7 μmol L$^{-1}$ break point remains valid. The absolute value of $r^2$ is reduced, even at higher [CO$_2$]$^{aq}_{(p-alk)}$-predicted, but this would be expected given the addition of uncertainty in the age model, as the published age is most likely to align with the ice core. Given the rapid rate of change at deglaciations, this effect is likely to be particularly pronounced in this dataset as many records have high temporal resolution around deglaciations in order to attempt to resolve them.

---

**Figure 4.** Crossplots of CO$_2$$(ε_{p-alk})$-based atmospheric CO$_2$ concentration (y axes) vs. the time-equivalent estimate from ice core records (x axes; Bereiter et al., 2015; Table 2). The large panel compiles all sites, with the exception of MANOP Site C, as explained in the text. Symbols are coloured by predicted [CO$_2$]$^{aq}_{(p-alk)}$ for each site and time as explained in the text. Full sources for alkenone data are shown in Table 1. A 1 : 1 line is included in all plots for comparison.
Any small age model offset introduced by the error modelling in these intervals also clearly has the potential to induce large differences between the CO$_2$($\varepsilon$$_{\text{p}}$-$\text{alk}$) and ice core values. Figure 6c and d clearly demonstrate that it is the filtering by [CO$_2$]$_{\text{aq}}$-predicted rather than any spurious correlations which determines the shape of the data in Fig. 6a.

4 Discussion

The plateau in $r^2$ in Fig. 6a and c suggests that below a [CO$_2$]$_{\text{aq}}$-predicted of $\sim$7 µmol L$^{-1}$ CO$_2$($\varepsilon$$_{\text{p}}$-$\text{alk}$) is no longer as good a predictor of ice core CO$_2$ as when [CO$_2$]$_{\text{aq}}$-predicted $>$ 7 µmol L$^{-1}$. This is clear from comparing the relationship between samples where [CO$_2$]$_{\text{aq}}$-predicted $<$ 7 µmol L$^{-1}$ with those where [CO$_2$]$_{\text{aq}}$-predicted $>$ 7 µmol L$^{-1}$ in Fig. 8. Here the $r^2$ for the former of 0.15 is substantially less than the latter of 0.55. I suggest that this is because below this threshold the fundamental assumption of CO$_2$($\varepsilon$$_{\text{p}}$-$\text{alk}$), that carbon is passively taken up by haptophytes, no longer holds true. One obvious explanation for why this would be the case is that at low levels of [CO$_2$]$_{\text{aq}}$ haptophytes have to rely more on active uptake of carbon via CCMs in order to satisfy metabolic demand. Similar behaviour has been recognized in some culture studies (Laws et al., 1997, 2002; Cassar et al., 2006), with some evidence that the diatom *Phaeodactylum tricornutum* has a similar CCM threshold of 7 µmol L$^{-1}$ (Laws et al., 1997). Whilst the evidence for the mechanism of CCM is poorer for coccolithophores than it is for diatoms, any CCM would be expected to compromise the CO$_2$($\varepsilon$$_{\text{p}}$-$\text{alk}$) proxy, either by increased supply of [CO$_2$]$_{\text{aq}}$ or by further carbon isotopic fractionation effects during carbon transport, or both (Stoll et al., 2019).

By applying a threshold value for [CO$_2$]$_{\text{aq}}$-predicted of 7 µmol L$^{-1}$ to the published records (Table 1), values of CO$_2$($\varepsilon$$_{\text{p}}$-$\text{alk}$) which are influenced by active CCMs can be eliminated. Recognition of this new threshold value of [CO$_2$]$_{\text{aq}}$-predicted allows for a new record of Pleistocene CO$_2$($\varepsilon$$_{\text{p}}$-$\text{alk}$) to be compiled. This compilation then much better replicates the glacial–interglacial pattern of CO$_2$ change over the last 260 kyr (Fig. 9). Whilst this present compilation does rely on ice core CO$_2$ records to estimate [CO$_2$]$_{\text{aq}}$-predicted, and therefore has little direct utility as a CO$_2$ record, it does demonstrate that recognition of a threshold response allows accurate CO$_2$ reconstruction using CO$_2$($\varepsilon$$_{\text{p}}$-$\text{alk}$). This may represent the point at which isotopic effects of CCMs (plausibly through increased CA activity or HCO$_3^-$ dehydration to meet C demand) overwhelm the assumptions of the CO$_2$($\varepsilon$$_{\text{p}}$-$\text{alk}$) proxy. This, as well as
Figure 6. Coefficient of determination (a) of a reducing sample of all compiled CO$_2$(ε$\_p$−alk) (Table 1) vs. the time-equivalent estimate from ice core records (Bereiter et al., 2015; Table 2). The sample reduces stepwise by 0.05 µmol L$^{-1}$, and the number of records in each subsample is shown in panel (b). Panel (c) shows a 1000-member Monte Carlo analysis, whereby uncertainty in CO$_2$(ε$\_p$−alk) and age is considered, as detailed in the text. Panel (d) shows a similar 1000-member Monte Carlo analysis, but with random sampling of the whole CO$_2$(ε$\_p$−alk) population so that the number of samples is equivalent to the dataset shown in panel (c); i.e. the size of the sample follows that shown in panel (b). Means and 1σ uncertainties are shown as the bold lines.

Figure 7. Emulated uncertainty in CO$_2$(ε$\_p$−alk), generated by running Monte Carlo uncertainty models for all sites in Table 1, applying the same approach to uncertainty as Badger et al. (2013a, b). Estimates used in this study are highlighted in blue.

the behaviour shown in Fig. 6a and c, suggests that from the standpoint of the CO$_2$(ε$\_p$−alk) proxy CCMs may effectively be considered either active or not, and that when [CO$_2$]$_{aq}$ is plentiful passive uptake dominates, at least sufficiently in most oceanographic settings that CO$_2$(ε$\_p$−alk) can accurately record atmospheric CO$_2$ concentration. This implies that, if areas of the ocean (or intervals of time) with low [CO$_2$]$_{aq}$ can be avoided, accurate reconstructions of atmospheric CO$_2$ concentration can be acquired using CO$_2$(ε$\_p$−alk).

As [CO$_2$]$_{aq}$ is affected by both SST via the temperature dependance of the Henry’s law constant and atmospheric CO$_2$ concentration, for CO$_2$(ε$\_p$−alk) to be effective in reconstructing atmospheric CO$_2$ concentration, areas of warm wa-
ter (i.e. tropical or shallow shelf regions) under relatively low atmospheric CO$_2$ concentration must be avoided. However, as the atmospheric CO$_2$ control renders the global surface ocean sufficiently replete with [CO$_2$]$_{aq}$ at Pliocene-like levels of atmospheric CO$_2$ concentration and above (Martínez-Botí et al., 2015) at all but the warmest surface ocean temperatures, CO$_2$(ε$_{p}$−alk) is likely to be a reliable system for most of the Cenozoic. It is only in the Pleistocene that atmospheric CO$_2$ concentration is low enough for CCMs to be widely active across the surface ocean, with the low-CO$_2$ glacial areas providing the most difficulty (Badger et al., 2019). This finding aligns well with evidence that CCMs developed in coccolithophores as a response to declining atmospheric CO$_2$ concentration through the Cenozoic and were developing in CO$_2$(ε$_{p}$−alk)$^+$ limited parts of the ocean in the late Miocene at the earliest, and likely not widespread until the Plio-Pleistocene (Bolton et al., 2012; Bolton and Stoll, 2013).

There have been recent attempts to correct for CCMs in CO$_2$(ε$_{p}$−alk)$^+$-based reconstructions of atmospheric CO$_2$ concentrations (Zhang et al., 2019; Stoll et al., 2019; Zhang et al., 2020). However, these assume that CCMs are always active and crucially do not fundamentally break the relationship between ε$_p$ values and atmospheric CO$_2$ concentration. However if this is not the case, and the relationship between ε$_p$ values and atmospheric CO$_2$ concentration fails at Pleistocene levels of atmospheric CO$_2$, then Pleistocene records cannot be used to develop corrections of CO$_2$(ε$_{p}$−alk)$^+$ to be applied throughout the Cenozoic. If, as suggested by the analyses presented here, CCMs only act at low [CO$_2$]$_{aq}$, and largely only in conditions prevalent throughout the late Pliocene and Pleistocene, it is plausible that corrections based on Pleistocene records could overcompensate for CCMs in the rest of the Cenozoic, when the assumption of passive carbon uptake inherent in CO$_2$(ε$_{p}$−alk) as traditionally applied may still be valid.

## 5 Conclusions

Reconstructions of past atmospheric CO$_2$ concentration with proxy tools like CO$_2$(ε$_{p}$−alk)$^+$ are critical for understanding how the Earth’s climate system operates, as long as the tools used can be relied upon to be accurate and precise. This re-analysis of existing Pleistocene CO$_2$(ε$_{p}$−alk)$^+$ records reveals that below a critical threshold of [CO$_2$]$_{aq}$ of 7 µmol$^{-1}$ the relationship between δ$^{13}$C$_{alkenone}$ and atmospheric CO$_2$ concentration breaks down, plausibly because below this threshold haptophytes are able to actively take up carbon using CCMs in order to satisfy metabolic demand.

Although reconstructing the low levels of atmospheric CO$_2$ concentration in the Pleistocene glacial and areas of the global ocean where [CO$_2$]$_{aq}$ is less than 7 µmol$^{-1}$ will be impossible, for much of the Cenozoic the CO$_2$(ε$_{p}$−alk)$^+$ proxy retains utility. If care is taken to avoid regions and oceanographic settings where [CO$_2$]$_{aq}$ is expected to be abnormally low, CO$_2$(ε$_{p}$−alk)$^+$ remains an important and useful proxy to understand the Earth system.

### Code and data availability

This paper relies exclusively on previously published data, available with the original papers and in publicly available repositories. An R notebook supplement is available alongside this paper, with a data file, which allow full replication of all analyses performed.

### Supplement

The supplement related to this article is available online at: https://doi.org/10.5194/bg-18-1149-2021-supplement.

### Competing interests

The author declares there is no conflict of interest.

### Acknowledgements

I am grateful to Gavin Foster and Tom Chalk for frequent and stimulating discussions on alkenone paleoceanography. I thank all authors who made full datasets available online. I thank Kirsty Edgar for comments on various drafts and the two anonymous reviewers, whose comments greatly improved this paper.

### Financial support

Financial support for this work was provided by the School of Environment, Earth and Ecosystem Sciences, The Open University.

### Review statement

This paper was edited by Jack Middelburg and reviewed by two anonymous referees.
References


M. P. S. Badger: CCMs in coccolithophores at low CO$_2$
Laws, E. A., Popp, B. N., Cassar, N., and Tanimoto, J.: 


R Core Team: A language and environment for statistical computing, available at: https://www.r-project.org/ (last access: 9 October 2020), 2020.


