

1 **Optimizing sampling strategies in high-resolution paleoclimate records**

2

3 Niels J. de Winter<sup>1,2</sup> \*, Tobias Agterhuis<sup>1</sup>, Martin Ziegler<sup>1</sup>

4

5 <sup>1</sup>Department of Earth Sciences, Utrecht University, Princetonlaan 8a, 3584 CB Utrecht, the Netherlands

6 <sup>2</sup>AMGC research group, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

7

8 Correspondence to: Niels J. de Winter (n.j.dewinter@uu.nl)

9

10 **Abstract**

11 The aim of paleoclimate studies to resolve climate variability from noisy proxy records can in essence be  
12 reduced to a statistical problem. The challenge is to extract meaningful information about climate variability  
13 from these records by reducing measurement uncertainty through a combination of proxy data while  
14 retaining the temporal resolution needed to assess the timing and duration of variations in climate  
15 parameters. In this study, we explore the limits of this compromise by testing different methods for  
16 combining proxy data (smoothing, binning and sample size optimization) on a particularly challenging  
17 paleoclimate problem: resolving seasonal variability in stable isotope records. We test and evaluate the  
18 effects of changes in the seasonal temperature and the hydrological cycle as well as changes in accretion  
19 rate of the archive and parameters such as sampling resolution and age model uncertainty on the reliability  
20 of seasonality reconstructions based on clumped and oxygen isotope analyses in 33 real and virtual  
21 datasets. Our results show that strategic combinations of clumped isotope analyses can significantly  
22 improve the accuracy of seasonality reconstructions compared to conventional stable oxygen isotope  
23 analyses, especially in settings where the isotopic composition of the water is poorly constrained.  
24 Smoothing data using a moving average often leads to an apparent dampening of the seasonal cycle,  
25 significantly reducing the accuracy of reconstructions. A statistical sample size optimization protocol yields  
26 more precise results than smoothing. However, the most accurate results are obtained through monthly  
27 binning of proxy data, especially in cases where growth rate or water composition cycles obscure the  
28 seasonal temperature cycle. Our analysis of a wide range of natural situations reveals that the effect of  
29 temperature seasonality on oxygen isotope records almost invariably exceeds that of changes in water  
30 composition. Thus, in most cases, oxygen isotope records allow reliable identification of growth seasonality  
31 as a basis for age modelling in absence of independent chronological markers in the record. These specific  
32 findings allow us to formulate general recommendations for sampling and combining data in paleoclimate  
33 research and have implications beyond the reconstruction of seasonality. We briefly discuss the  
34 implications of our results for solving common problems in paleoclimatology and stratigraphy.

35

## 36 1. Introduction

37 Improving the resolution of climate reconstructions is a key objective in paleoclimate studies because it  
38 allows climate variability to be studied on different timescales and sheds light on the continuum of climate  
39 variability (Huybers and Curry, 2006). However, the temporal resolution of climate records is limited by the  
40 accretion rate (growth or sedimentation rate) of the archive and the spatial resolution of sampling for climate  
41 reconstructions, which is a function of the sample size required for a given climate proxy. This tradeoff  
42 between sample size and sampling resolution is especially prevalent when using state-of-the-art climate  
43 proxies which require large sample sizes, such as the carbonate clumped isotope paleothermometer ( $\Delta_{47}$ ;  
44 see applications in Rodríguez-Sanz et al., 2017; Briard et al., 2020; Caldarescu et al., 2021) or stable  
45 isotope ratios in specific compounds or of rare isotopes (e.g. phosphate-oxygen isotopes in tooth apatite,  
46 triple oxygen isotopes in speleothems or carbon isotopes of CO<sub>2</sub> in ice cores; Jones et al., 1999; Schmitt  
47 et al., 2012; Sha et al., 2020). The challenge of sampling resolution persists on a wide range of timescales:  
48 from attempts to resolve geologically short-lived (kyr-scale) climate events from deep sea cores with low  
49 sedimentation rates (e.g. Stap et al., 2010; Rodríguez-Sanz et al., 2017) to efforts to characterize tidal or  
50 daily variability in accretionary carbonate archives (e.g. Warter and Müller, 2017; de Winter et al., 2020a).  
51 What constitutes “high-resolution” is therefore largely dependent on the specifics of the climate archive.

52 Sample size limitations are especially important in paleoseasonality reconstructions. Reliable archives for  
53 seasonality (e.g. corals, mollusks and speleothem records) are in high demand in the paleoclimate  
54 community, because the seasonal cycle is one of the most important cycles in Earth’s climate and  
55 seasonality reconstructions complement more common long-term (kyr to -Myr) records of past climate  
56 variability (e.g. Morgan and van Ommen, 1997; Tudhope et al., 2001; Steuber et al., 2005; Steffensen et  
57 al., 2008; Denton et al., 2005; Huyghe et al., 2015; Vansteenberghe et al., 2019). A more detailed  
58 understanding of climate dynamics at the human timescale is increasingly relevant for improving climate  
59 projections (IPCC, 2013). Unfortunately, the growth and mineralization rates of archives that capture high-  
60 resolution variability (rarely exceeding 10 mm/yr) limit the number and size of samples that can be obtained  
61 at high temporal resolutions (e.g. Mosley-Thompson et al., 1993; Passey and Cerling, 2002; Treble et al.,  
62 2003; Goodwin et al., 2003). This problem is exacerbated by the fact that accurate methods for climate

63 reconstructions often require comparatively large sample sizes while methods relying on smaller sample  
64 sizes rely on uncertain assumptions. A case in point is the popular carbonate stable oxygen isotope  
65 temperature proxy ( $\delta^{18}\text{O}_c$ ) which relies on assumptions of the water composition ( $\delta^{18}\text{O}_w$ ) that become  
66 progressively more uncertain further back in geological history (e.g. Veizer and Prokoph, 2015). Contrarily,  
67 the clumped isotope proxy ( $\Delta_{47}$ ) does not rely on this assumption but requires larger amounts of sample  
68 (e.g. Müller et al., 2017)

69 A promising technique for circumventing sample size limitations is to analyze larger numbers of small  
70 aliquots from the same sample or from similar parts of the climate archive. These smaller aliquots typically  
71 have poor precision but averaging multiple aliquots into one estimate while propagating the measurement  
72 uncertainty leads to a more reliable estimate of the climate variable (Dattalo, 2008; Meckler et al., 2014;  
73 Müller et al., 2017; Fernandez et al., 2017). This approach yields improved sampling flexibility since aliquots  
74 can be combined in various ways after measurement. It also allows outlier detection at the level of individual  
75 aliquots, thereby spreading the risk of instrumental failure and providing improved control on changes in  
76 measurement conditions that may bias results.

77 Previous studies have applied several different methods for combining data from paleoclimate records to  
78 reduce analytical noise or higher order variability, and extract variability with a specific frequency (e.g. a  
79 specific orbital cycle or seasonality; e.g. Lisiecki and Raymo, 2004; Cramer et al., 2009). These data  
80 reduction approaches can in general be categorized into **smoothing** techniques, in which a sliding window  
81 or range of neighboring datapoints is used to smooth high resolution records (see e.g. Cramer et al., 2009)  
82 or **binning** techniques, in which the record is divided into equal bins in sampling direction (e.g. time, depth  
83 or length in growth direction; e.g. Lisiecki and Raymo, 2004; Rodríguez-Sanz et al., 2017). In addition, a  
84 third approach is proposed here based on **optimization** of sample size for dynamic binning of data along  
85 the climate cycle using a moving window in the domain of the climate variable (as opposed to the sampling  
86 domain) combined with a T-test routine (see section 2.1). All three approaches have advantages and  
87 caveats.

88 In this study, we explore the (dis)advantages of these three data reduction approaches by testing their  
89 reliability in resolving seasonal variability in sea surface temperature (SST) and water stable oxygen isotope

90 composition ( $\delta^{18}\text{O}_w$ ), both highly sought-after variables in paleoclimate research. We compare  
91 reconstructions of SST and  $\delta^{18}\text{O}_w$  in real and virtual datasets from accretionary carbonate archives (e.g.  
92 shells, corals and speleothems) using the clumped isotope thermometer ( $\Delta_{47}$ ) combined with stable oxygen  
93 isotope ratios of the carbonate ( $\delta^{18}\text{O}_c$ ).

94

## 95 **2. Methods**

### 96 **2.1 Reconstruction approaches**

97 Throughout the remainder of this work, the three approaches for combining data for reconstructions are  
98 defined as follows (see also **Fig. 1**):

99 **Smoothing** refers to the reconstruction of SST and  $\delta^{18}\text{O}_w$  based on **moving averages** of  $\Delta_{47}$  and  $\delta^{18}\text{O}_c$   
100 records (**Fig. 1B**). For every dataset, the full possible range of moving window sizes (from **one sample to the**  
101 **full length of the record**) for SST and  $\delta^{18}\text{O}_w$  reconstructions was explored. The window size that resulted in  
102 the most significant difference between maximum and minimum  $\Delta_{47}$  values (based on a student's T-test)  
103 was applied to reconstruct SST and  $\delta^{18}\text{O}_w$  from  $\Delta_{47}$  and  $\delta^{18}\text{O}_c$  records. SST and  $\delta^{18}\text{O}_w$  were calculated for  
104 all case studies using a combination of empirical temperature relationships by Kim and O'Neil (1997;  $\delta^{18}\text{O}_c$ -  
105  $\delta^{18}\text{O}_w$ -temperature relationship) and Bernasconi et al. (2018;  $\Delta_{47}$ -temperature relationship). Here and in  
106 other approaches, a typical analytical uncertainty on measurements of  $\Delta_{47}$  (one standard deviation of  
107 0.04‰) and  $\delta^{18}\text{O}_c$  (one standard deviation of 0.05‰) was used to include uncertainty due to measurement  
108 precision. These analytical uncertainties were chosen based on typical uncertainties reported for these  
109 measurements in the literature (e.g. Schöne et al., 2005; Huyghe et al., 2015; Vansteenberghe et al., 2016)  
110 and long-term precision uncertainties obtained by measuring in-house standards using the MAT253+ with  
111 Kiel IV setup in the clumped isotope laboratory at Utrecht University (e.g. Kocken et al., 2019). The  
112 measurement uncertainty was propagated through all calculations using a Monte Carlo simulation (N =  
113 1000) in which  $\Delta_{47}$  and  $\delta^{18}\text{O}_c$  records were randomly sampled from a normal distribution with the virtual  $\Delta_{47}$   
114 and  $\delta^{18}\text{O}_c$  values as means and analytical uncertainties as standard deviations.

115 **Binning** refers to reconstructions of SST and  $\delta^{18}\text{O}_w$  based on binning of  $\Delta_{47}$  and  $\delta^{18}\text{O}_c$  records into monthly  
116 time bins (**Fig. 1C**). The  $\Delta_{47}$  and  $\delta^{18}\text{O}_c$  data from each case study were grouped into monthly time bins and  
117 converted to SST and  $\delta^{18}\text{O}_w$  using the Kim and O'Neil (1997) and Bernasconi et al. (2018) formulae. Here  
118 too, Monte Carlo simulation was applied to propagate measurement uncertainties onto monthly SST and  
119  $\delta^{18}\text{O}_w$  reconstructions. Note that the prerequisite for this method is that the data is aligned using a (floating)  
120 age model accurate enough to allow samples to be placed in the right bin. The age of virtual samples in  
121 this study is known so this prerequisite poses no problems in this case. However, in the fossil record this  
122 alignment might be less certain in absence of accurate chronologies within the archive (e.g. through daily  
123 growth increments in mollusk shells; e.g. Schöne et al., 2008; Huyghe et al., 2019; see 4.1.3).

124 **Optimization** refers to reconstructions of SST and  $\delta^{18}\text{O}_w$  based on sample size optimization in  $\Delta_{47}$  records  
125 (**Fig. 1D**). In this approach aliquots of each virtual dataset are ordered from warm (low  $\delta^{18}\text{O}_c$ ) to cold (high  
126  $\delta^{18}\text{O}_c$  data) samples, regardless of their position relative to the seasonal cycle. From this ordered dataset,  
127 increasingly large samples of multiple aliquots (from 2 aliquots to half the length of the record) are taken  
128 from both the warm ("summer") and the cold ("winter") side of the distribution. Summer and winter samples  
129 were kept equal (symmetrical grouping) to reduce the number of possible sample size combinations and  
130 allow for more efficient computation. However, asymmetrical grouping with differing sample sizes on the  
131 summer and winter ends of the  $\delta^{18}\text{O}_c$ -spectrum are possible (see 4.1.3 and 4.2.2). Sample sizes with  
132 significant difference in  $\Delta_{47}$  value between summer and winter groups ( $p \leq 0.05$  based on a student's T-  
133 test) were selected as optimal sample sizes. The moving window T-test in the proxy domain ensures that  
134 an optimal compromise is reached between high precision and resolving differences between seasonal  
135 extremes. For each successful sample size, SST and  $\delta^{18}\text{O}_w$  values were calculated from  $\Delta_{47}$  and  $\delta^{18}\text{O}_c$  data  
136 according to Kim and O'Neil (1997) and Bernasconi et al. (2018) formulae. The relationship between SST  
137 and  $\delta^{18}\text{O}_w$  obtained from these reconstructions was used to convert all  $\Delta_{47}$  and  $\delta^{18}\text{O}_c$  data to SST and  
138  $\delta^{18}\text{O}_w$ , which are then grouped into monthly SST and  $\delta^{18}\text{O}_w$  reconstructions. Measurement uncertainties  
139 were propagated through the entire approach by Monte Carlo simulation (N = 1000).

140 For comparison, we also include reconstructions based solely on  $\delta^{18}\text{O}_c$  measurements with an (often  
141 inaccurate) assumption of a constant  $\delta^{18}\text{O}_w$  (equal to the modern ocean value of 0‰ VSMOW), which form

142 the most common method for carbonate-based temperature reconstructions in paleoclimate research (see  
143 e.g. Schöne et al., 2005; Westerhold et al., 2020; **Fig. 1A**; hereafter:  $\delta^{18}\text{O}$ ). For these reconstructions,  $\delta^{18}\text{O}_c$   
144 records were grouped into monthly time bins with analytical uncertainties propagated using the Monte Carlo  
145 approach ( $N = 1000$ ) and were directly converted to SST using the Kim and O'Neil (1997) temperature  
146 relationship.

147 For each reconstruction, SST and  $\delta^{18}\text{O}_w$  results were aggregated into monthly averages, medians, standard  
148 deviations, and standard errors. Step by step documentation of calculations made for the three  $\Delta_{47}$ -based  
149 reconstruction approaches and the  $\delta^{18}\text{O}_c$  reconstructions are given in **S7** and in the complementary R  
150 package (de Winter, 2021a).

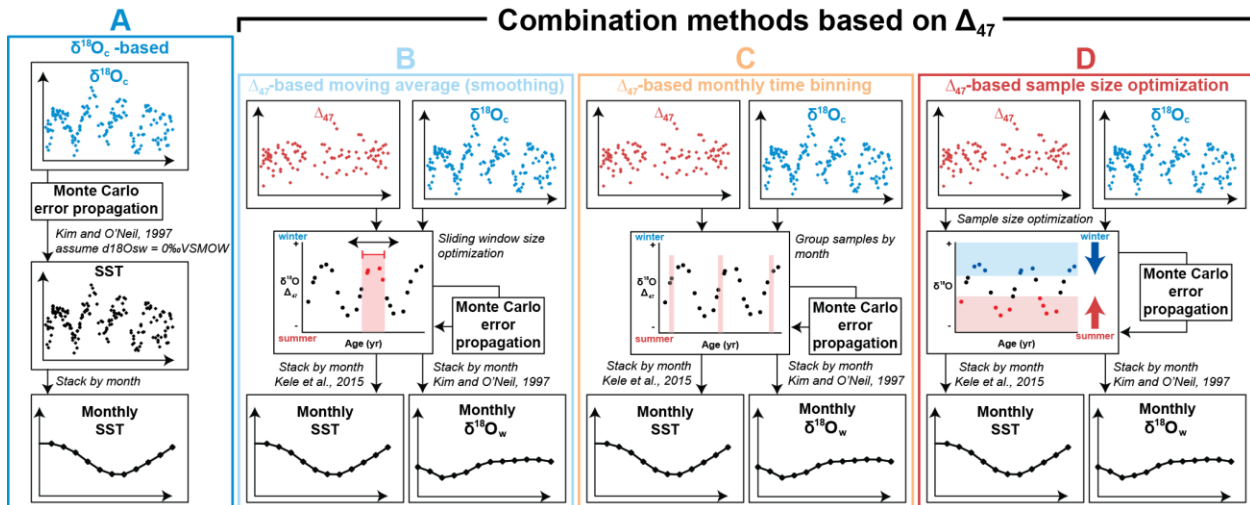
## 151 **2.2 Benchmarks for accuracy and precision**

152 Accuracy and precision of reconstructions were evaluated against official USGS definitions of climate  
153 parameters (O'Donnell et al., 2012):

- 154 1. mean annual SST (MAT), defined as the average of all 12 monthly temperature reconstructions.
- 155 2. seasonal range in SST, defined as the temperature difference between warmest and coldest  
156 month.
- 157 3. mean annual  $\delta^{18}\text{O}_w$ , defined as the average of all 12 monthly  $\delta^{18}\text{O}_w$  reconstructions.
- 158 4. seasonal range in  $\delta^{18}\text{O}_w$ , defined as the  $\delta^{18}\text{O}_w$  difference between most enriched (highest  $\delta^{18}\text{O}_w$ )  
159 and most depleted (lowest  $\delta^{18}\text{O}_w$ ) monthly reconstruction.

160 Accuracy was defined as the absolute offset of the reconstructed climate parameter from the “true” value.

161 Precision was defined as the (relative) standard deviation of the reconstruction, as calculated from the  
162 variability within monthly time bins resulting from Monte Carlo error propagation (see **2.1**). An overview of  
163 monthly SST and  $\delta^{18}\text{O}_w$  reconstructions using the four approaches in all cases is given in **S4**. Raw data  
164 and figures of reconstructions of all cases using all sampling resolutions are compiled in **S8**.



165

166 **Figure 1:** Schematic overview of the four approaches for seasonality reconstructions: (A)  $\delta^{18}\text{O}$ -based  
 167 reconstructions, assuming constant  $\delta^{18}\text{O}_w$ . (B) Reconstructions based on **smoothing**  $\delta^{18}\text{O}_c$  and  $\Delta_{47}$  data  
 168 using a moving average. (C) Reconstructions based on binning  $\delta^{18}\text{O}_c$  and  $\Delta_{47}$  data in monthly time bins.  
 169 (D) Reconstructions based on **optimization** of the sample size for combining  $\delta^{18}\text{O}_c$  and  $\Delta_{47}$  data (see  
 170 description in 2.1). Colored **curves** represent virtual  $\delta^{18}\text{O}_c$  (blue) and  $\Delta_{47}$  (red) series in sampling domain.  
 171 Black curves represent reconstructed monthly SST and  $\delta^{18}\text{O}_w$  averages.

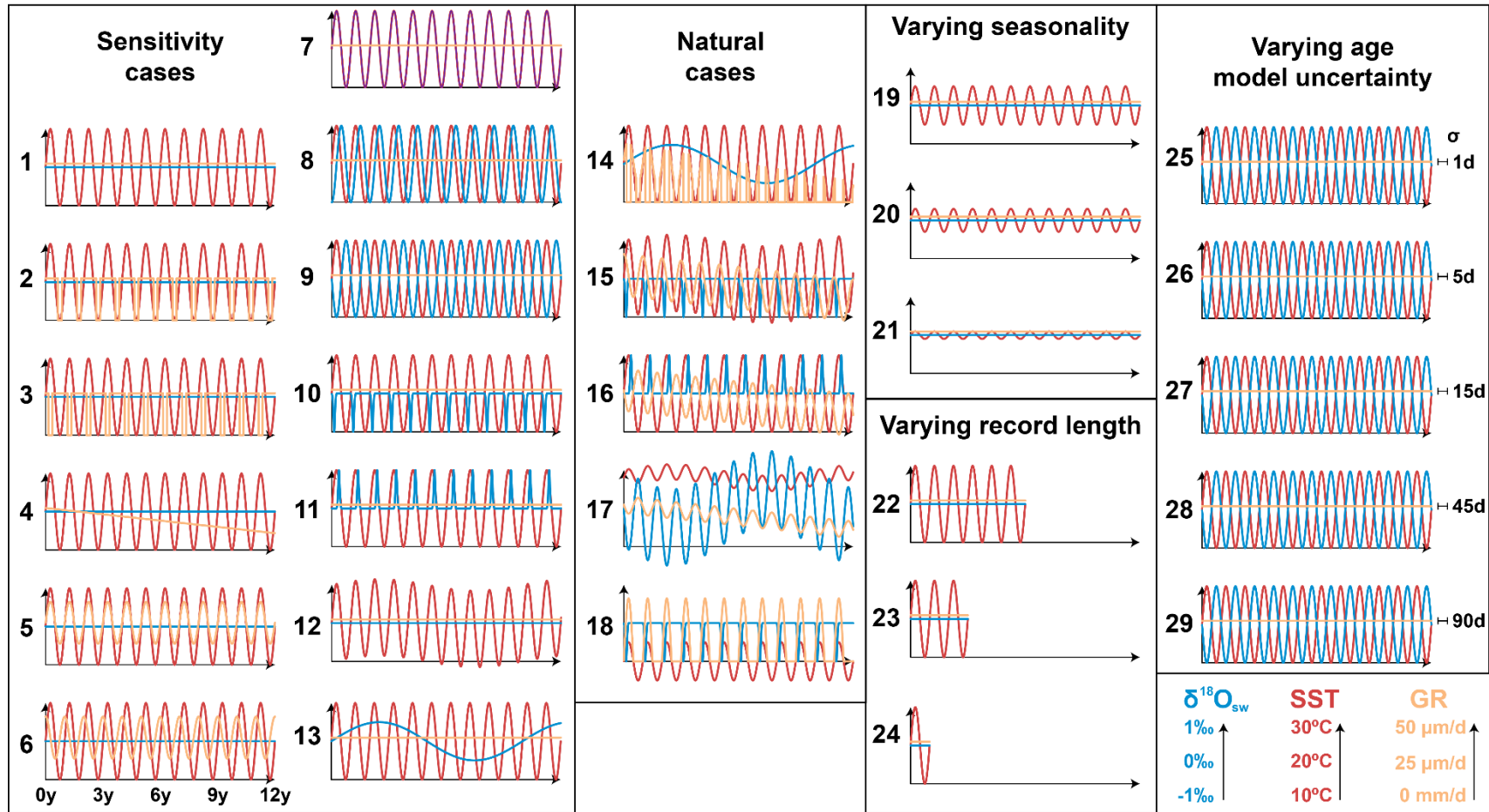
172

### 173 2.3 SST and $\delta^{18}\text{O}_w$ datasets

174 The three reconstruction approaches were tested and compared based on three types of data: Firstly, data  
 175 from a real specimen of a Pacific oyster (*Crassostrea gigas*, syn. *Magallana gigas*) reported in Ullmann et  
 176 al. (2010). Secondly, data based on actual measurements of natural variability in SST and sea surface  
 177 salinity (SSS; case 30-33) converted to virtual  $\Delta_{47}$  and  $\delta^{18}\text{O}_c$  records. Thirdly, a set of datasets based on  
 178 fully artificial SST and  $\delta^{18}\text{O}_w$  data (case 1-29; see Fig. 2) converted to virtual  $\Delta_{47}$  and  $\delta^{18}\text{O}_c$  records.



## Virtual cases



179

180 **Figure 2:** Overview of time series of all virtual test cases. Colored curves represent time series of SST (red),  $\delta^{18}O_w$  (blue), and growth rate (orange, abbreviated as “GR”). Horizontal axes in all plots are 12 years long (see legend below case 6). Vertical axis of all plots has the same scale (SST: 10 to 30°C;  $\delta^{18}O_w$ : -1 to +1‰; Growth rate: 0–50  $\mu\text{m}/\text{day}$ ; see legend in bottom right corner). Horizontal error bars and labels on the right side of cases 25-29 represent standard errors introduced on the age model (bars not to scale). The  $\delta^{18}O_c$  and  $\Delta_{47}$  records resulting from these virtual datasets are provided in **S6** (see also **Fig. 3** for natural examples).

181

182

183

184

Sensitivity cases		Natural cases	Varying seasonality	Varying age model uncertainty		
1. Control	7. $\delta^{18}O_w$ seasonality in phase with SST	14. Full marine case with ontogenetic GR trend	19. Control case with reduced SST amplitude ( $\sim 5^\circ C$ )	25. Case 9 with $\pm 1$ day age model uncertainty		
	8. $\delta^{18}O_w$ seasonality in antiphase with SST		20. Control case with reduced SST amplitude ( $\sim 3^\circ C$ )			
	2. Growth stops $< 12^\circ C$		9. $\delta^{18}O_w$ seasonality lags SST by $\frac{1}{4}$ year		21. Control case with reduced SST amplitude ( $\sim 1^\circ C$ )	26. Case 9 with $\pm 5$ days age model uncertainty
	3. Growth stops $> 28^\circ C$		10. Negative $\delta^{18}O_w$ in spring		<b>Varying record length</b>	
	4. Linear decrease in GR		11. Positive $\delta^{18}O_w$ in summer		22. Control case shortened to 6 yr	27. Case 9 with $\pm 15$ days age model uncertainty
	5. GR seasonality in phase with SST		12. Multi-annual (5 yr) SST cycle		23. Control case shortened to 3 yr	28. Case 9 with $\pm 45$ days age model uncertainty
6. GR seasonality lags SST by $\frac{1}{4}$ year	13. Multi-annual (5 yr) $\delta^{18}O_w$ cycle	15. Coastal case with spring $\delta^{18}O_w$ decrease and decreasing GR trend	29. Case 9 with $\pm 90$ days age model uncertainty			
		16. Lagoonal case with summer $\delta^{18}O_w$ increase	24. Control case shortened to 1 yr			
		17. Tropical monsoon case with confined SST seasonality and strong multi-annual SST cycle				
		18. Worst-case scenario with growth limited to summer half of the year				

185 **Table 1:** Overview of virtual cases 1-29 used to test the reconstruction methods. Case descriptions are  
186 abbreviated. Details on the SST, growth rate and  $\delta^{18}O_w$  included in each case are described in detail in **S1**.  
187 SST, growth rate and  $\delta^{18}O_w$  records of all cases are shown in **Fig. 2**. "GR" = growth rate.

188

### 189 2.3.1 Modern oyster data

190 Environmental SST and  $\delta^{18}O_w$  data from the List Basin in Denmark ( $54^\circ 59.25'N$ ,  $8^\circ 23.51'E$ ), where the  
191 modern oyster specimen lived, were obtained from local *in situ* measurements of SST and SSS described  
192 in Ullmann et al. (2010). Since direct, *in situ* measurements of  $\delta^{18}O_w$  variability at a high temporal resolution  
193 were not available,  $\delta^{18}O_w$  was estimated from more widely available SSS data using a mass balance  
194 (equation 1 and 2; following e.g. Ullmann et al., 2010):

$$195 \delta^{18}O_{sw} = \delta^{18}O_{w,freshwater} * f + \delta^{18}O_{w,ocean} * (1 - f) \quad (1)$$

196 
$$f = \frac{SSS_{sample} - SSS_{ocean}}{SSS_{freshwater} - SSS_{ocean}} \quad (2)$$

197 Here, we assume salinity ( $SSS_{sample}$ ) results from a mixture of a fraction ( $f$ ) isotopically light and low-salinity  
 198 ( $\delta^{18}O_{w,freshwater} = -8.5\text{‰}$ ;  $SSS_{freshwater} = 0$ ) freshwater and a fraction ( $1-f$ ) ocean water ( $\delta^{18}O_{w,ocean} = 0\text{‰}$ ;  
 199  $SSS_{ocean} = 35$ ), with negative amounts of freshwater contribution ( $f < 0$ ) representing net evaporation  
 200 ( $SSS_{sample} > SSS_{ocean}$ ). The value for  $\delta^{18}O_{w,freshwater}$  was based on the discharge weighted average  $\delta^{18}O_w$  of  
 201 water in the nearby Elbe and Weser rivers (see Ullmann et al., 2010). All  $\delta^{18}O_w$  values throughout the text  
 202 are with reference to the VSMOW scale. Contrary to the virtual datasets (cases 1-33; see **2.3.2** and **2.3.3**),  
 203 the Ullmann et al. (2010) data was already available in the sampling domain, hence no subsampling was  
 204 required.

205 **2.3.2 Cases based on real climate data**

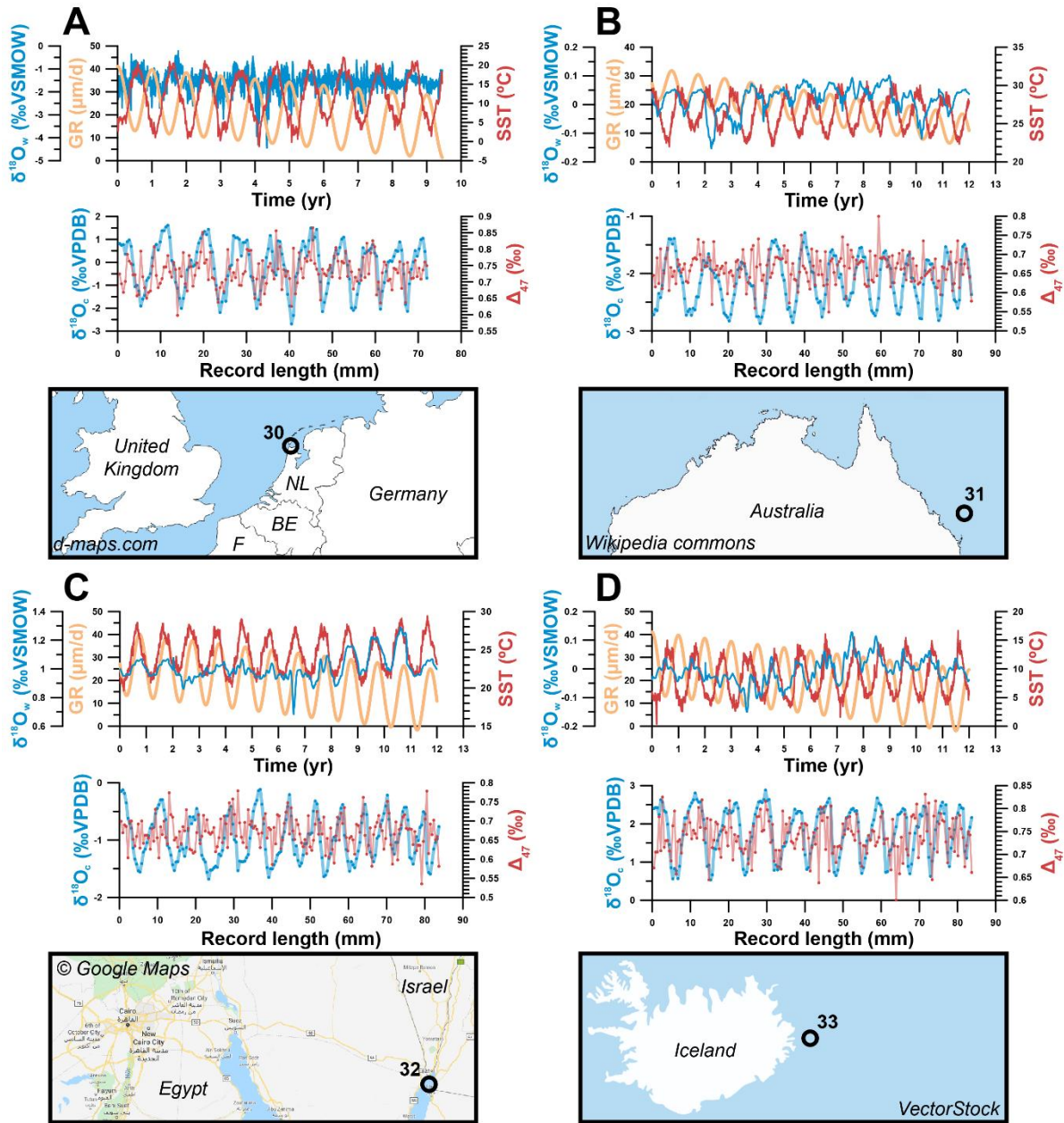
206 Four test cases were based on time series of real SST and SSS data from four different locations, selected  
 207 to capture a variety of environments with different SST and SSS variability (see **Fig. 3**):

- 208 1. Tidal flats of the Wadden Sea near Texel, the Netherlands (case 30)
- 209 2. Great Barrier Reef in Australia (case 31)
- 210 3. Gulf of Aqaba between Egypt and Saudi Arabia (case 32)
- 211 4. Northern Atlantic Ocean east of Iceland (case 33).

212 Daily measurements of SST and SSS for case 31-33 were obtained from worldwide open-access datasets  
 213 of the National Oceanic and Atmospheric Administration (NOAA, 2020) and European Space Agency (ESA,  
 214 2020) respectively. Hourly SST and SSS measured *in situ* in the Wadden Sea (case 30) were obtained  
 215 from the Dutch Institute for Sea Research (NIOZ, Texel, the Netherlands). Since direct, *in situ*  
 216 measurements of  $\delta^{18}O_w$  variability at a high temporal resolution are scarce,  $\delta^{18}O_w$  was estimated from (more  
 217 widely available) SSS data using the same mass balance described in **2.3.1**. The value for  $\delta^{18}O_{w,freshwater}$   
 218 was based on the  $\delta^{18}O_w$  of rain in the Netherlands ( $-8\text{‰}$ ; Mook, 1970; Bowen, 2020). Applying this mass  
 219 balance on the SSS record of the Wadden Sea tidal flats (case 30) results in  $\delta^{18}O_w$  values and a SSS- $\delta^{18}O_w$   
 220 relationship in agreement with measurements in this region (Harwood et al., 2008). SST and  $\delta^{18}O_w$  time  
 221 series for all cases are given in **S4** and natural cases are plotted in **Fig. 3**.

222 For all virtual datasets (cases 1-33), records of SST and  $\delta^{18}\text{O}_w$  were converted to the sampling domain  
223 (along the length of the record) by defining a virtual growth rate in the sampling direction. Adding this growth  
224 rate as a variable allowed us to test the sensitivity of approaches to changes in the extension rate of the  
225 archive, including hiatuses (growth rate = 0). This is important, because fluctuations in linear extension rate  
226 and periods in which no mineralization occurs (hiatuses or growth cessations) are common in all climate  
227 archives (e.g. Treble et al., 2003; Ivany, 2012). After conversion to the sampling domain, virtual aliquots  
228 were subsampled at equal distance from the SST and  $\delta^{18}\text{O}_w$  series of all cases using six sampling intervals:  
229 0.1 mm, 0.2 mm, 0.45 mm, 0.75 mm, 1.55 mm and 3.25 mm. The four largest sampling intervals were  
230 chosen such that the standard growth rate (10 mm/yr) was not an integer multiple of the sampling interval  
231 (e.g. 0.45 mm instead of 0.5 mm, and 3.25 mm instead of 3 mm). This decision prevents sampling the same  
232 parts of the seasonal cycle (e.g. same months) every year, which biases both the mean value and the  
233 precision of monthly SST and  $\delta^{18}\text{O}_w$  reconstructions. This bias towards certain parts of the seasonal cycle  
234 is much stronger at low sample sizes (large sampling intervals) and is illustrated in the **Supplementary**  
235 **Information**.

## Real data based cases



236

237 **Figure 3:** Overview of the four cases of virtual data based on natural SST and SSS measurements explored  
 238 in this study. (A) Case 30: Tidal flats on the Wadden Sea, Texel, the Netherlands. (B) Case 31 Great Barrier  
 239 Reef, Australia). (C) Case 32: Gulf of Aqaba between Egypt and Saudi Arabia. (D) Case 33: Atlantic Ocean  
 240 east of Iceland. For all cases, graphs on top show environmental data, with SST plotted in red,  $\delta^{18}\text{O}_w$  in  
 241 blue and growth rate (abbreviated as “GR”) in orange (as in Fig. 2). The graph below shows virtual  $\delta^{18}\text{O}_c$   
 242 (blue) and  $\Delta_{47}$  (red) records created from these data series using a sampling interval of 0.45 mm and  
 243 including analytical noise (see 3.3). Note that the scale of vertical axes varies between plots.

244

245 2.3.3 Virtual cases

246 Virtual SST and  $\delta^{18}\text{O}_w$  time series were artificially constructed to test the effect of various SST and  $\delta^{18}\text{O}_w$   
247 scenarios on the effectivity of the reconstruction methods. The default test case (case 1) contained an ideal,  
248 12-year sinusoidal SST curve with a period of 1 year (seasonality), a mean value of 20°C and a seasonal  
249 amplitude of 10°C, a constant  $\delta^{18}\text{O}_w$  value of 0‰ and a constant growth rate of 10 mm/yr. Other cases  
250 contain various deviations from this ideal case (see also **Fig. 2**, **Table 1** and **S1**):

- 251 • Linear and/or seasonal changes in growth rate, including growth stops (cases 2-6, 14-18)
- 252 • Seasonal and/or multi-annual changes in  $\delta^{18}\text{O}_w$  (cases 7-11, 13-18)
- 253 • Multi-annual trends in SST superimposed on the seasonality (cases 12, 15 and 17)
- 254 • Variations in the seasonal SST amplitude (cases 19-21)
- 255 • Change in the total length of the time series (cases 22-24).
- 256 • Variation in uncertainty on the age of each virtual datapoint (cases 25-29)

257 Comparison of the virtual time series (case 1-29; **Fig. 2**) with the natural variability (case 30-33; **Fig. 3**)  
258 shows that the virtual cases are not realistic approximations of natural variability in SST and  $\delta^{18}\text{O}_w$ . Natural  
259 SST and  $\delta^{18}\text{O}_w$  variability are not limited to the seasonal or multi-annual scale but contain a fair amount of  
260 higher order (daily to weekly scale) variability. To simulate this natural variability, we extracted the seasonal  
261 component of SST and  $\delta^{18}\text{O}_w$  variability from our highest resolution record of measured natural SST and  
262 SSS data (case 30: data from Texel, the Netherlands, see **2.3.2** and **Fig. 3**). The standard deviation of  
263 residual variability of this data after subtraction of the seasonal cycle was used to add random high-  
264 frequency noise to the SST and  $\delta^{18}\text{O}_w$  variability in virtual cases. Note that while sub-annual environmental  
265 variability can be approximated by Gaussian noise (Wilkinson and Ivany, 2002), this representation is an  
266 oversimplification of reality. In the case of our Texel data, the SST and SSS residuals are not normally  
267 distributed (Kolmogorov-Smirnov test:  $D = 0.010$ ;  $p = 7.2 \cdot 10^{-14}$  and  $D = 0.039$ ;  $p < 2.2 \cdot 10^{-16}$  for SST and  
268 SSS residuals respectively; see **S2-4**). SST and  $\delta^{18}\text{O}_w$  data from cases 1-29 was converted to the sampling  
269 domain and subsampled at a range of sampling resolutions following the same procedure applied to cases  
270 30-33 (see **2.3.2**).

271

272 **2.4 Conversion to  $\delta^{18}\text{O}_c$  and  $\Delta_{47}$  data**

273 After subsampling, SST and  $\delta^{18}\text{O}_w$  series (cases 1-33) were converted to  $\delta^{18}\text{O}_c$  and  $\Delta_{47}$  using a carbonate  
274 model based on empirical relationships between  $\Delta_{47}$  and  $\delta^{18}\text{O}_c$  with and SST and  $\delta^{18}\text{O}_w$  (equation 3 and 4;  
275 Kim and O'Neil, 1997; Kele et al., 2015; Bernasconi et al., 2018) and the conversion of  $\delta^{18}\text{O}$  values from  
276 VSMOW to VPDB scale (equation 5; Brand et al., 2014).

277 
$$\Delta_{47} = \frac{0.0449 \cdot 10^6}{(SST+273.15)^2} + 0.167 \quad (3)$$

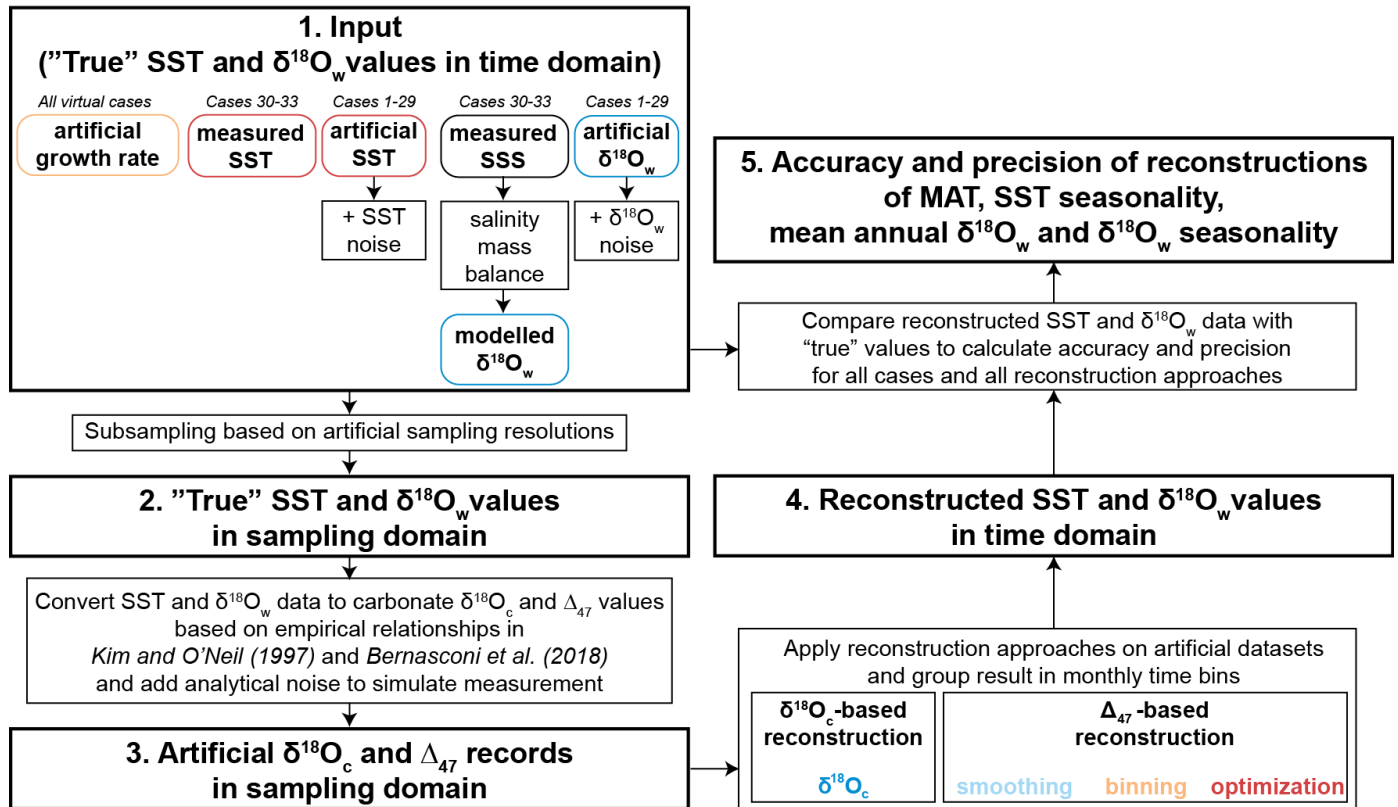
278 
$$1000 * \ln \frac{\left(\frac{^{18}\text{O}}{^{16}\text{O}}\right)_{\text{CaCO}_3}}{\left(\frac{^{18}\text{O}}{^{16}\text{O}}\right)_{\text{H}_2\text{O}}} = 18.03 * \left(\frac{10^3}{(SST+273.15)}\right) - 32.42 \quad (4)$$

279 
$$\delta^{18}\text{O}_{VPDB} = 0.97002 * \delta^{18}\text{O}_{VSMOW} - 29.98 \quad (5)$$

280 For the real oyster data (Ullmann et al., 2010; see **2.3.1**), only the  $\Delta_{47}$  data needed to be created because  
281  $\delta^{18}\text{O}_c$  was directly measured. As a result, each case study yielded records of  $\Delta_{47}$  and  $\delta^{18}\text{O}_c$  in the sampling  
282 domain and corresponding “true” SST and  $\delta^{18}\text{O}_w$  records in the time domain, allowing assessment of the  
283 reliability of the reconstruction approaches in different scenarios. (**Figure 4**). The result of applying these  
284 steps is illustrated on case 31 (Great Barrier reef data, **Fig. 5**). All calculations for creating  $\Delta_{47}$  and  $\delta^{18}\text{O}_c$   
285 series in sampling domain were carried out using the open-source computational software R (R core team,  
286 2013), and scripts for these calculations are given in **S7** and compiled in the documented R package  
287 “seasonalclumped” (de Winter, 2021a). All  $\Delta_{47}$  and  $\delta^{18}\text{O}_c$  datasets are provided in **S6**.

288

## Workflow for creating virtual datasets and testing reconstruction approaches

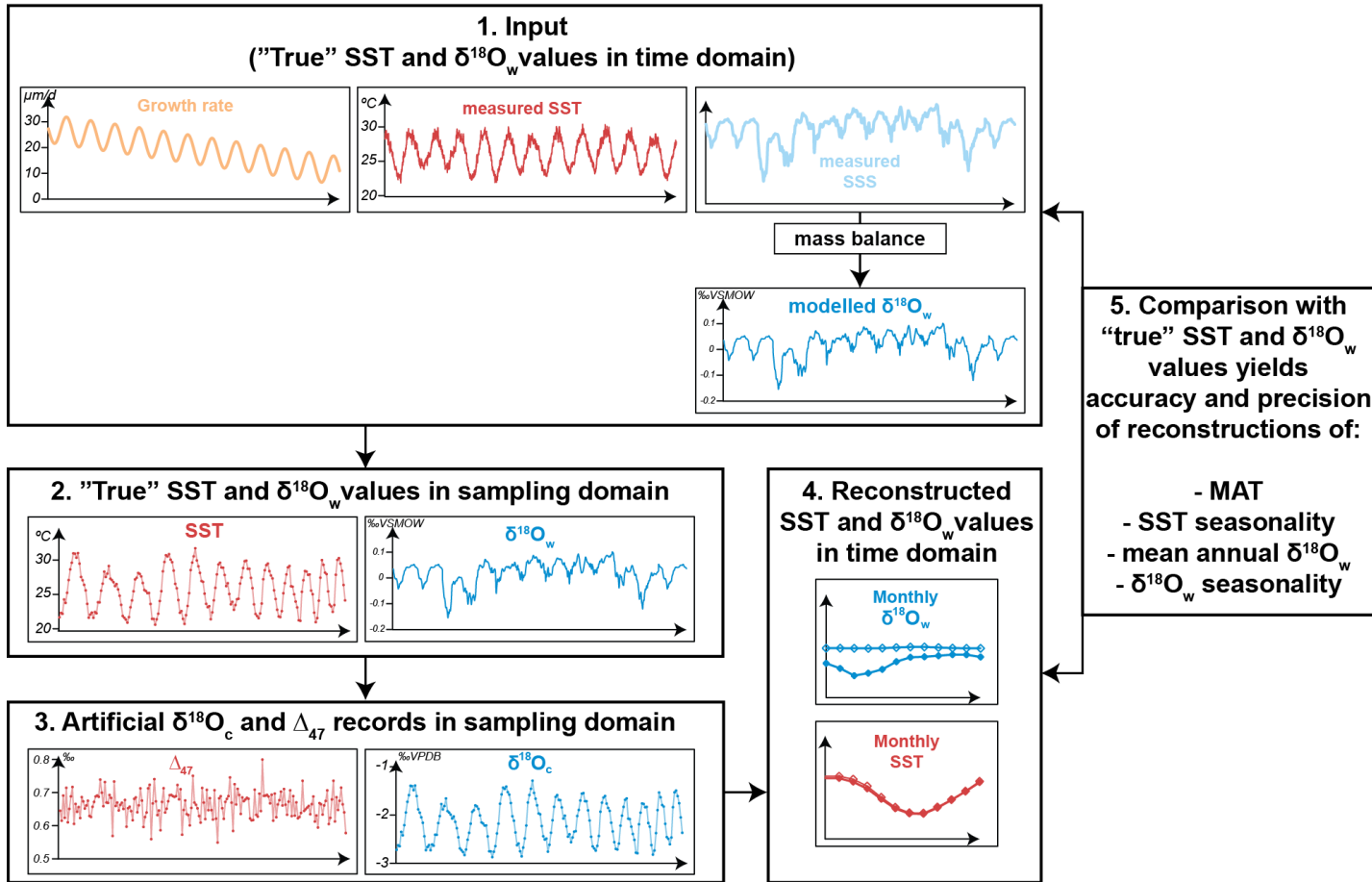


289

290 **Figure 4:** Flow diagram showing the steps taken to create virtual data ( $\Delta_{47}$  and  $\delta^{18}O_c$ ; cases 1-33) and  
 291 compare results of SST and  $\delta^{18}O_w$  reconstructions with the actual SST and  $\delta^{18}O_w$  data the record was  
 292 based on (counterclockwise direction). Steps 1-3 outline the procedure for creating virtual  $\Delta_{47}$  and  $\delta^{18}O_c$   
 293 datasets (see sections 2.3 and 2.4), step 4 shows the application of the different reconstruction methods  
 294 on this virtual data (see Fig. 2 for details) and step 5 illustrates how the reconstructions are compared with  
 295 the original ("true") SST and  $\delta^{18}O_w$  data to calculate accuracy and precision of the reconstruction  
 296 approaches. Note that step 1 is different for cases 1-29 (based on fully artificial SST and  $\delta^{18}O_w$  records;  
 297 2.3.3) than for cases 30-33 (SST and  $\delta^{18}O_w$  records based on real SST and SSS data; see 2.3.2).



Workflow for creating virtual datasets and testing reconstruction approaches:  
 Example for case 31 (Great Barrier Reef satellite data)



298

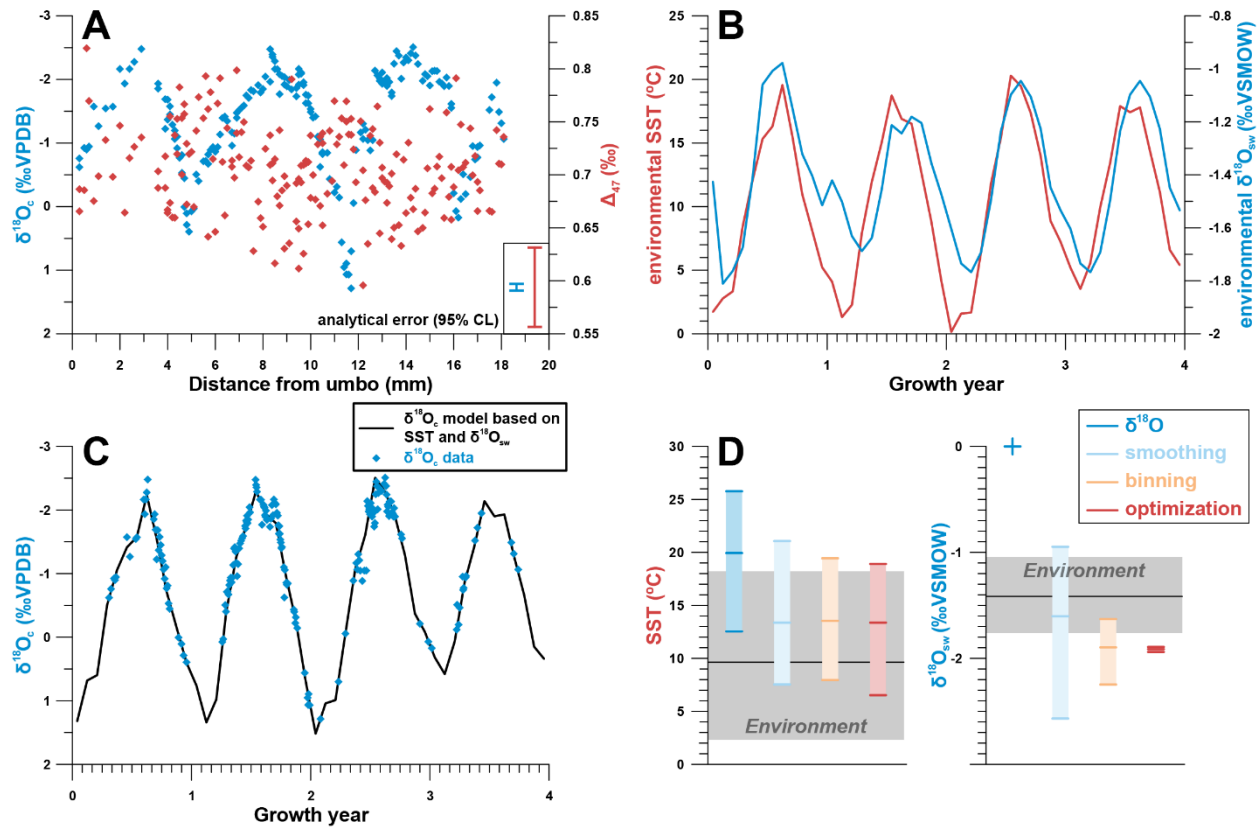
299 **Figure 5:** An example of the steps highlighted in Fig. 4 using case 31 (Great Barrier Reef data) to illustrate  
 300 the data processing steps. Virtual data plots include normally distributed measurement uncertainty on  $\Delta_{47}$   
 301 and  $\delta^{18}\text{O}_c$

302

### 303 3. Results

#### 304 3.1 Real example

305 Measured ( $\delta^{18}\text{O}_c$ ) and simulated ( $\Delta_{47}$ ) data from the Pacific oyster from the Danish List Basin yielded  
306 estimates of SST and  $\delta^{18}\text{O}_w$  seasonality using all reconstruction approaches (**Fig. 6**). While a model of shell  
307  $\delta^{18}\text{O}_c$  based on SST and SSS data closely approximates the measured  $\delta^{18}\text{O}_c$  record (**Fig. 6C**), basing SST  
308 reconstructions solely on  $\delta^{18}\text{O}_c$  data without any *a priori* knowledge of  $\delta^{18}\text{O}_w$  variability (assuming constant  
309  $\delta^{18}\text{O}_w$  equal to the global marine value) leads to **high inaccuracy in SST seasonality and mean annual SST**  
310 **(Fig. 6D)**. The in-phase relationship between SST and SSS (**Fig. 6B**) dampens the seasonal  $\delta^{18}\text{O}_c$  cycle,  
311 causing underestimation of temperature seasonality, while a negative mean annual  $\delta^{18}\text{O}_w$  value in the List  
312 Basin biases SST reconstructions towards higher temperatures. In terms of SST reconstructions, the  
313 **smoothing, binning** and **optimization** approaches based on  $\Delta_{47}$  and  $\delta^{18}\text{O}_c$  data yield more accurate  
314 reconstructions, albeit with a reduced seasonality and a bias towards the summer season. The latter is a  
315 result of severely reduced growth rates in the winter season, which was therefore undersampled (see **Fig.**  
316 **6A** and **6C**). Approaches including  $\Delta_{47}$  data also yield far more accurate  $\delta^{18}\text{O}_w$  estimates than the  **$\delta^{18}\text{O}$**   
317 approach. However, the accuracy of  $\delta^{18}\text{O}_w$  seasonality and mean annual  $\delta^{18}\text{O}_w$  estimates is low in these  
318 approaches too, largely because of the limited sampling resolution, especially in winter. The **optimization**  
319 approach suffers from the strong in-phase relationship between SST and SSS, which obscures the  
320 difference between the  $\delta^{18}\text{O}_w$  effect and the temperature effect on shell carbonate. Yet, disentangling SST  
321 from  $\delta^{18}\text{O}_w$  seasonality is central to the success of the approach (see **3.4**). **Fig. 6D** does not show the  
322 reproducibility error on SST and  $\delta^{18}\text{O}_w$  estimates, which is much larger for the **smoothing** approach than  
323 for the **binning** and **optimization** approaches due to the limited data in the winter seasons (see **S5**). These  
324 results show that several properties of carbonate archives, such as growth rate variability, phase  
325 relationships between SST and  $\delta^{18}\text{O}_w$  seasonality and sampling resolution, can impact the reliability of  
326 paleoseasonality reconstructions. The virtual and real data cases in this study were tailored to test the  
327 effects of these archive properties more thoroughly.



328

329 **Figure 6:** (A) Plot of  $\delta^{18}O_c$  and (virtual)  $\Delta_{47}$  data from a modern Pacific oyster (*Crassostrea gigas*; see  
 330 Ullmann et al., 2010). (B) shows SST and  $\delta^{18}O_w$  data from the List Basin (Denmark) in which the oyster  
 331 grew. (C) shows the fit between  $\delta^{18}O_c$  data and modelled  $\delta^{18}O_c$  calculated from SST and  $\delta^{18}O_w$  on which  
 332 the shell age model was based. (D) Shows a summary of the results of different approaches for  
 333 reconstructing SST and  $\delta^{18}O_w$  from the  $\delta^{18}O_c$  and  $\Delta_{47}$  data. The vertical colored bars show the reconstructed  
 334 seasonal variability using all methods with ticks indicating warmest month, coldest month, and annual  
 335 mean. The grey horizontal bars show the actual seasonal variability in the environment. Precision errors on  
 336 monthly reconstructions are not shown but are given in S4.

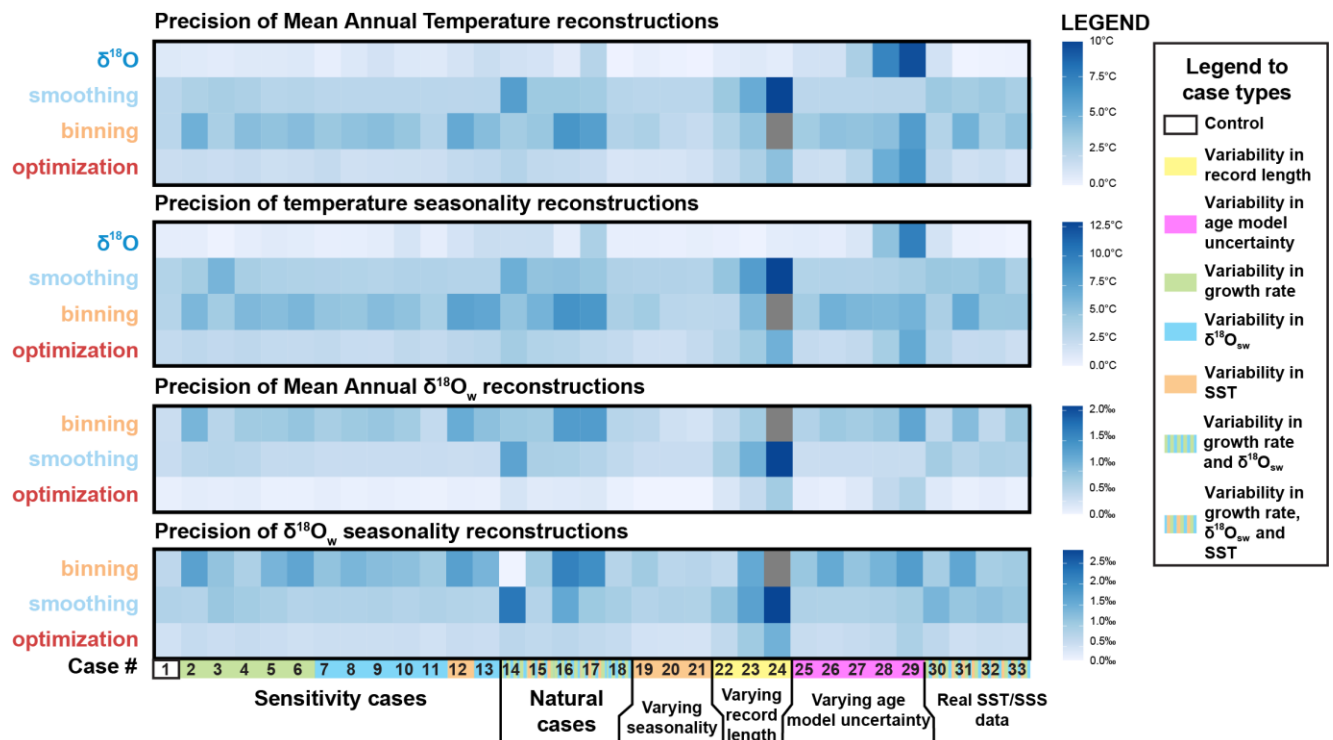
337

### 338 3.2 Case-specific results

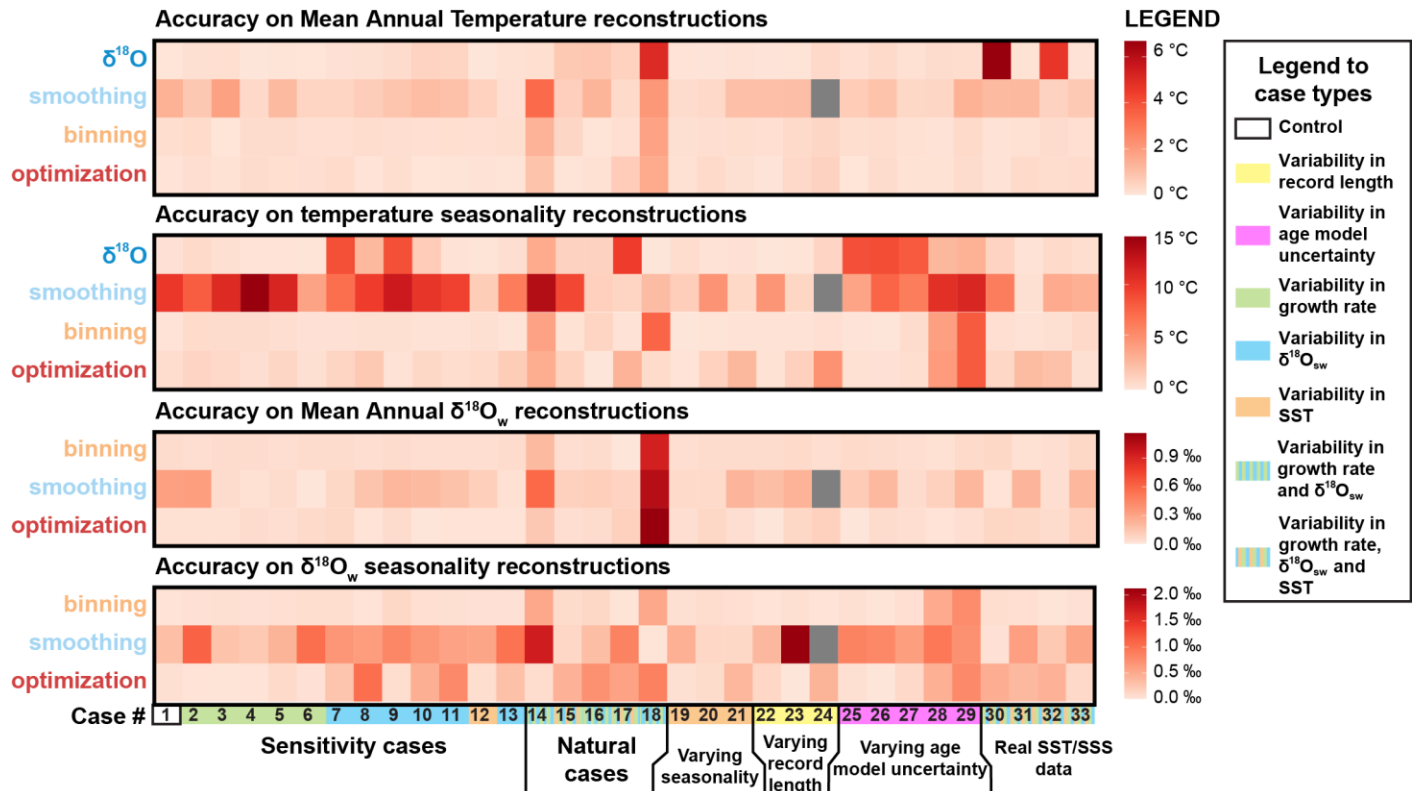
339 A case-by-case breakdown of the precision (Fig. 7) and accuracy (Fig. 8) of reconstructions using the four  
 340 approaches shows that reliability of reconstructions varies significantly between approaches and is highly  
 341 case-specific. In general, precision is highest in  $\delta^{18}O$  reconstructions, followed by **optimization** and  
 342 **binning** with **smoothing** generally yielding the worst precision. Average precision standard deviations of  
 343 the underperforming methods (**binning** and **smoothing**) are up to 2-3 times larger than those of  $\delta^{18}O$  (e.g.  
 344 respectively 3.9°C and 3.5°C vs. 1.3°C for  $\delta^{18}O$  MAT reconstructions; see also **Supplementary**  
 345 **Information**). It is worth noting that precision on  $\delta^{18}O$ -based estimates is mainly driven by measurement

346 precision (which is better for  $\delta^{18}\text{O}_c$  than for  $\Delta_{47}$  measurements, see section 4.1.1).  $\Delta_{47}$ -based reconstructions  
347 lose precision due to the higher measurement error on  $\Delta_{47}$  measurements and the method used for  
348 combining measurements for seasonality reconstructions. On a case-by-case basis, the hierarchy of  
349 approaches can vary, especially if strong variability in growth rate is introduced, such as in case 14, where  
350 the size of hiatuses in the record increases progressively, or in case 18, in which half of the year is missing  
351 due to growth hiatuses (see Table 1, S1 and S4). Of the  $\Delta_{47}$ -based methods (**smoothing**, **binning** and  
352 **optimization**), **optimization** is rarely outcompeted in terms of precision in both SST and  $\delta^{18}\text{O}_w$   
353 reconstructions.

354 The comparison based on precision alone is misleading, as the most precise approach ( **$\delta^{18}\text{O}$** ) runs the risk  
355 of being highly inaccurate (offsets exceeding  $4^\circ\text{C}$  on some MAT reconstructions; see Fig. 7C), especially  
356 in cases based on natural SST and SSS (case 30-33). The **smoothing** approach also often yields highly  
357 inaccurate results, especially in cases with substantial variability in  $\delta^{18}\text{O}_w$  (e.g. case 9-11). Accuracy of  
358 **optimization** and **binning** outcompete the other methods in most circumstances. **Binning** outperforms  
359 **optimization** in reconstructions of  $\delta^{18}\text{O}_w$  seasonality, making it overall the most accurate approach.  
360 Interestingly, **optimization** is less accurate specifically in cases with sharp changes in growth rate in  
361 summer (e.g. cases 11, 14, 16 and 17), while **binning** performs better in these cases. Reconstructions of  
362 mean annual SST and  $\delta^{18}\text{O}_w$  in case 18 are especially inaccurate regardless of which method is applied.  
363 This extreme case with growth only during one half of the year combined with seasonal fluctuations in both  
364 SST and  $\delta^{18}\text{O}_w$  presents a worst-case scenario for seasonality reconstructions leading to strong biases in  
365 mean annual temperature reconstructions. In situations like case 18, the **optimization** approach is most  
366 accurate in MAT and SST seasonality reconstructions, but  $\delta^{18}\text{O}_w$  is more accurately reconstructed using  
367 the **binning** approach. Finally, it is worth noting that in natural situations (Fig. 3), variability in SST almost  
368 invariably has a larger influence on  $\delta^{18}\text{O}_c$  and  $\Delta_{47}$  records than  $\delta^{18}\text{O}_w$ , such that fluctuations in  $\delta^{18}\text{O}_c$  records  
369 closely follow the SST seasonality even in cases with relatively large  $\delta^{18}\text{O}_w$  variability (e.g. case 30).  
370 Chronologies based on these  $\delta^{18}\text{O}_c$  fluctuations are therefore generally accurate.



372 **Figure 7:** Overview of precision (propagated standard deviation of variability within reconstructions, see  
 373 **2.2**) of reconstructions of mean annual temperature (**A**), seasonal temperature range (**B**), mean annual  
 374  $\delta^{18}O_w$  (**C**) and seasonal range in  $\delta^{18}O_w$  (**D**), with higher values (darker colors) indicating lower precision  
 375 (more variability between reconstructions) based on average sampling resolution (sampling interval of 0.45  
 376 mm). The different cases on the horizontal axis are color coded by their difference from the control case  
 377 (case 1; see legend on the right-hand side). Grey boxes indicate cases for which reconstructions were not  
 378 successful. All data on precision (standard deviation values) is provided in **S4**.

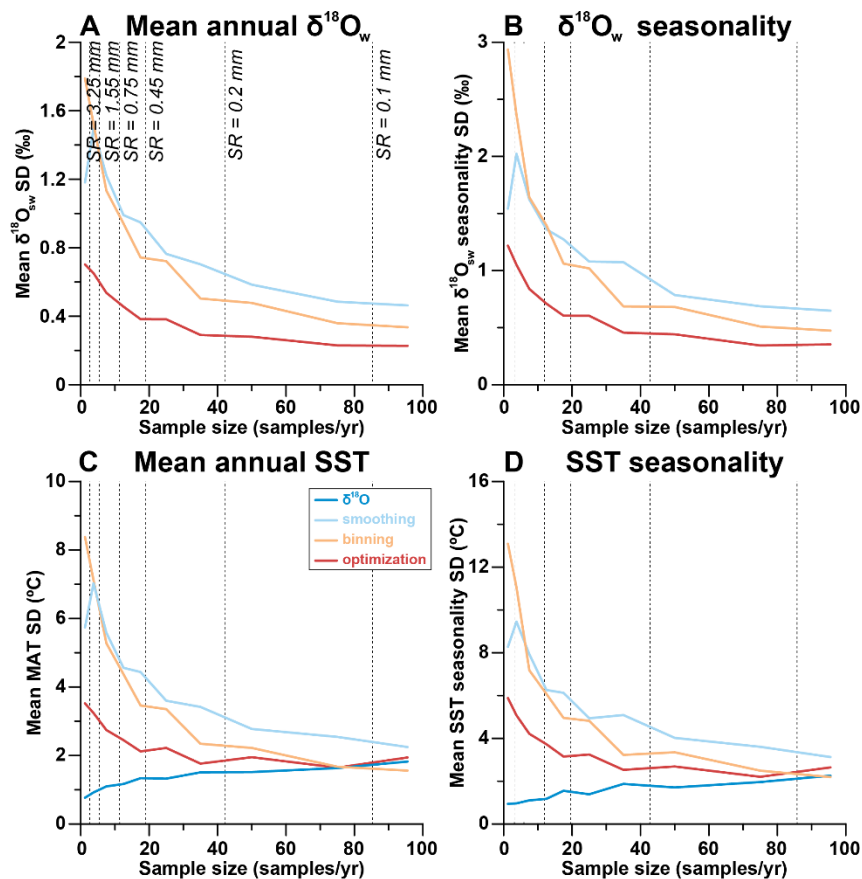


379

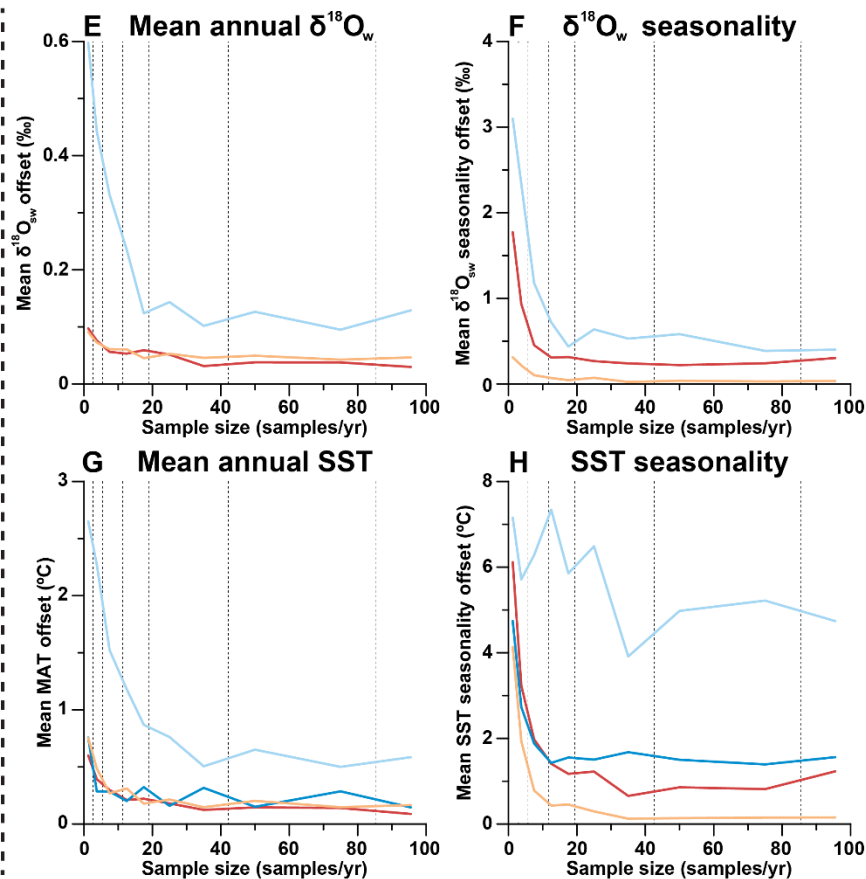
380 **Figure 8:** Overview of accuracy (absolute offset from “true” values) of reconstructions of mean annual  
 381 temperature (A), seasonal temperature range (B), mean annual  $\delta^{18}O_w$  (C) and seasonal range in  $\delta^{18}O_w$  (D),  
 382 with higher values (darker colors) indicating lower accuracy (higher offsets) based on average sampling  
 383 resolution (sampling interval of 0.45 mm). The different cases on the horizontal axis are color coded by  
 384 their difference from the control case (case 1; see legend on the right-hand side). Grey boxes indicate  
 385 cases for which reconstructions were not successful. All data on accuracy (difference between  
 386 reconstructed and “true” values) is provided in S4.

387

## Precision



## Accuracy



388

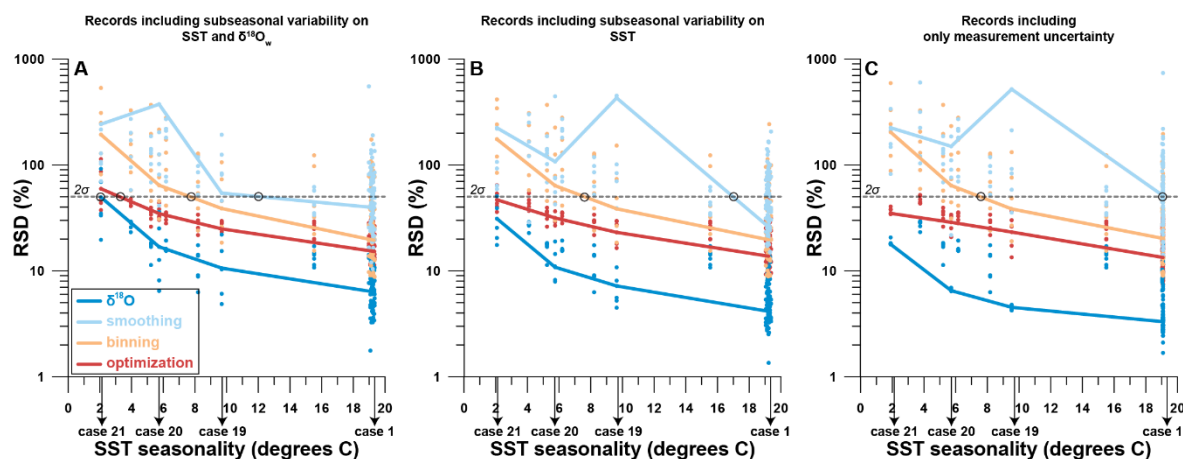
389 **Figure 9:** Effect of sampling resolution (in samples per year, see **S5**) on the precision (one standard deviation) of results of reconstructions of mean  
 390 annual  $\delta^{18}\text{O}_w$  (**A**), seasonal range in  $\delta^{18}\text{O}_w$  (**B**), mean annual SST (**C**) and seasonal range in SST (**D**). Effect on the accuracy (absolute offset from  
 391 actual value) of results of reconstructions of mean annual  $\delta^{18}\text{O}_w$  (**E**) and seasonal range in  $\delta^{18}\text{O}_w$  (**F**), mean annual SST (**G**) and seasonal range in  
 392 SST (**H**). Color coding follows the scheme in **Fig. 1** and **Fig. 4**.



### 393 3.3 Effect of sampling resolution

394 As expected, increasing the temporal sampling resolution (i.e. number of samples per year) almost  
 395 invariably increases the precision and accuracy (Fig. 9) of reconstructions using all methods. An exception  
 396 to this rule is the precision of  $\delta^{18}\text{O}$  reconstructions, which decreases with increasing sampling resolution.  
 397 Precision errors of all  $\Delta_{47}$ -based approaches eventually converge with the initially much lower precision  
 398 error of  $\delta^{18}\text{O}$  reconstructions when sampling resolution increases. However, the sampling resolution  
 399 required for  $\Delta_{47}$ -based reconstructions to rival or outcompete the  $\delta^{18}\text{O}$  reconstructions differs, with  
 400 optimization requiring lower sampling resolutions than the other methods (e.g. 20-40 samples/year  
 401 compared to 40-80 samples/year for smoothing and binning; Fig. 9A-D). Accuracy also improves with  
 402 sampling resolution (Fig. 9E-H). When grouping all cases together, it becomes clear that  $\delta^{18}\text{O}$   
 403 reconstructions can only approach the accuracy of  $\Delta_{47}$ -based approaches for reconstructions of MAT.  
 404 Seasonality in both SST and  $\delta^{18}\text{O}_w$  is most accurately reconstructed using binning, and the smoothing  
 405 approach once again performs worst.

406



407 **Figure 10:** Effect of SST seasonality range (difference between warmest and coldest month) in the record  
 408 on the relative precision of SST seasonality reconstructions (“RSD”, defined as one standard deviation  
 409 divided by the mean value). **Panel A** shows precision results if random variability (“weather patterns”) in  
 410 both SST and  $\delta^{18}\text{O}_w$  as well as measurement uncertainty is added to the records (see 2.3.3 and S1). Panel  
 411 **B** shows precision of records with random variability in SST and measurement uncertainty only. Panel **C**  
 412 shows precision if only measurement uncertainty is considered. Color coding follows the scheme in Fig. 1  
 413 and Fig. 4. Shaded dots represent results at various sampling resolutions, while bold lines are averages  
 414 for all reconstruction approaches. Black circles highlight the places where curves cross the threshold of two  
 415 standard deviations, which indicates the minimum SST seasonality that can be resolved within 2 standard  
 416 deviations (~95% confidence level) using the reconstruction approach.

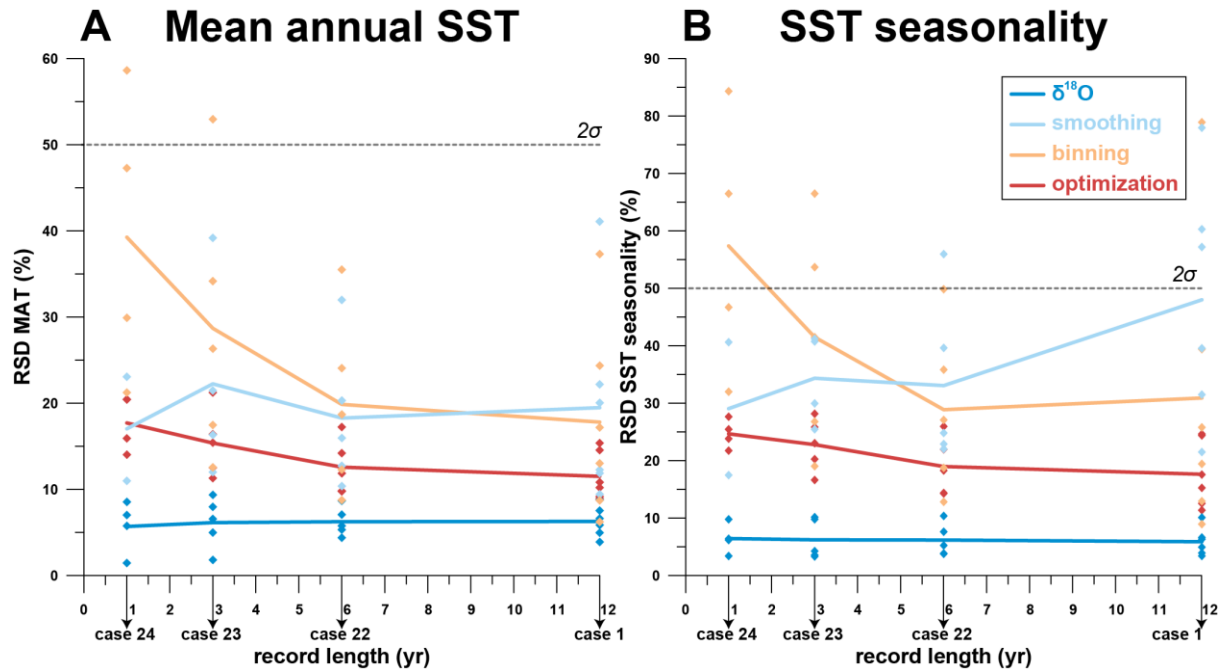
417



### 418 3.4 Resolving SST seasonality

419 Comparison of cases 19, 20 and 21 (SST seasonality of 9.7°C, 5.7°C and 2.1°C respectively) with control  
420 case 1 (SST seasonality of 19.3°C) shows how changes in the seasonal SST range affect the precision of  
421 measurements (**Fig. 10**; see also **Table 1** and **S1**). The data reconfirms that  $\delta^{18}\text{O}$  reconstructions are most  
422 precise; a deceptive statistic given the risk of highly inaccurate results this approach yields (see **Fig. 8**).  
423 Taking into consideration only analytical uncertainty, all approaches except for **smoothing** can confidently  
424 resolve at least the highest SST seasonality within a significance level of two standard deviations (~95%)  
425 using a moderate sampling resolution (mean of all resolutions shown in **Fig. 10**). Increasing sampling  
426 resolution improves the precision of  $\Delta_{47}$ -based reconstructions (see **Fig. 9D**), so high sampling resolutions  
427 (0.1 or 0.2 mm) allow smaller seasonal differences to be resolved. When random sub-annual variability is  
428 added to the SST and  $\delta^{18}\text{O}_w$  records (see **2.3.3**), the minimum seasonal SST extent that can be resolved  
429 decreases for all approaches (**Fig. 10B** and **10C**). Nevertheless,  $\delta^{18}\text{O}$  and **optimization** reconstructions  
430 remain able to resolve a relatively small SST seasonality of 2-4°C.

431



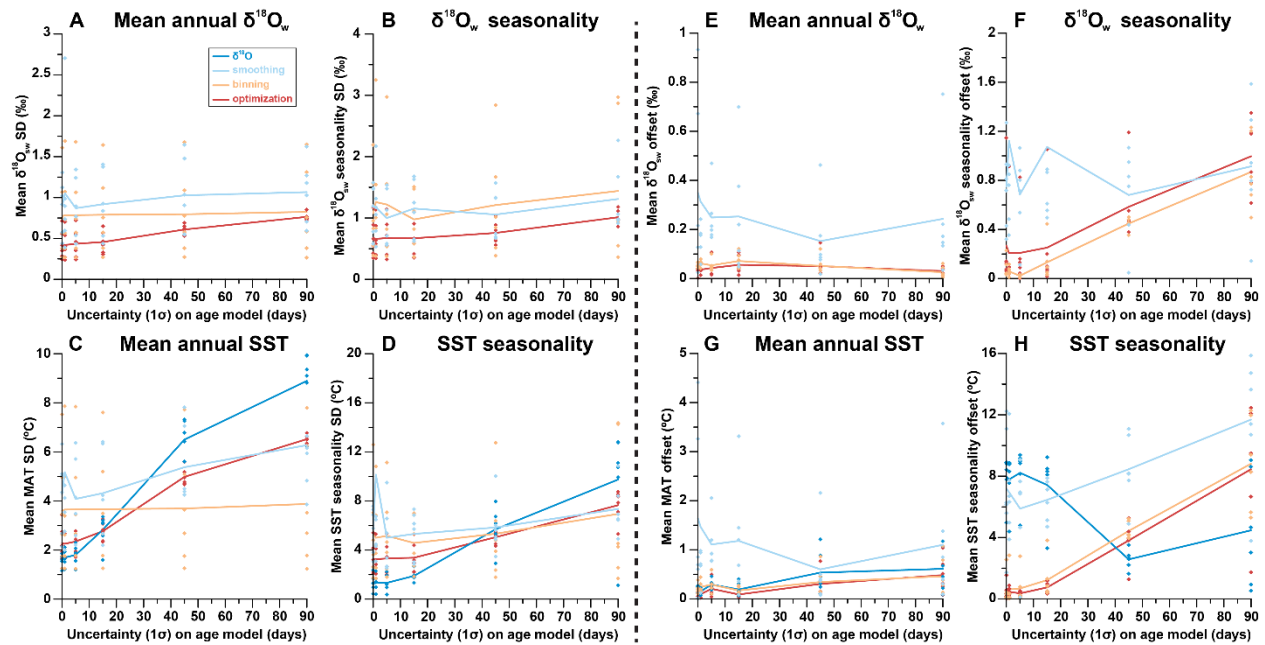
432

433 **Figure 11:** Effect of record length (in years) on the relative precision (one standard deviation as fraction of  
 434 the mean value) of results of reconstructions of mean annual SST (A) and SST seasonality (B). Shaded  
 435 dots represent results for the six different sampling resolutions. Solid lines connect averages for cases 1,  
 436 22, 23 and 24 for each reconstruction approach.

437

### 438 3.5 Effect of record length

439 The effect of variation in the length of the record was investigated by comparing cases 22, 23 and 24 (record  
 440 lengths of 6 years, 3 years and 1 year, respectively) with the control case (record length of 12 years; see  
 441 Fig. 11 and Table 1). Precision of MAT and SST seasonality reconstructions slightly increase in larger  
 442 datasets (longer records) for **optimization** and **binning**, but not for **smoothing** and **δ<sup>18</sup>O** reconstructions.  
 443 Differences between reconstruction approaches remain relatively constant regardless of the length of the  
 444 record, with **general precision hierarchy** remaining intact (**δ<sup>18</sup>O** > **optimization** > **binning** > **smoothing**).  
 445 However, in very short records (1-2 years) **smoothing** gains an advantage over other Δ<sub>47</sub>-based methods  
 446 due to its lack of sensitivity to changes in the record length and **binning** reconstructions are not precise  
 447 enough to resolve **MAT** and SST seasonality within two standard deviations (~95% confidence level).  
 448 Variation in precision is largely driven by very high precision errors of reconstructions in records with low  
 449 sampling resolutions (sampling intervals of 1.55 mm or 3.25 mm; see also Fig. 9A-D). As a result, most of  
 450 the reduction in precision in shorter records can be mitigated by denser sampling.



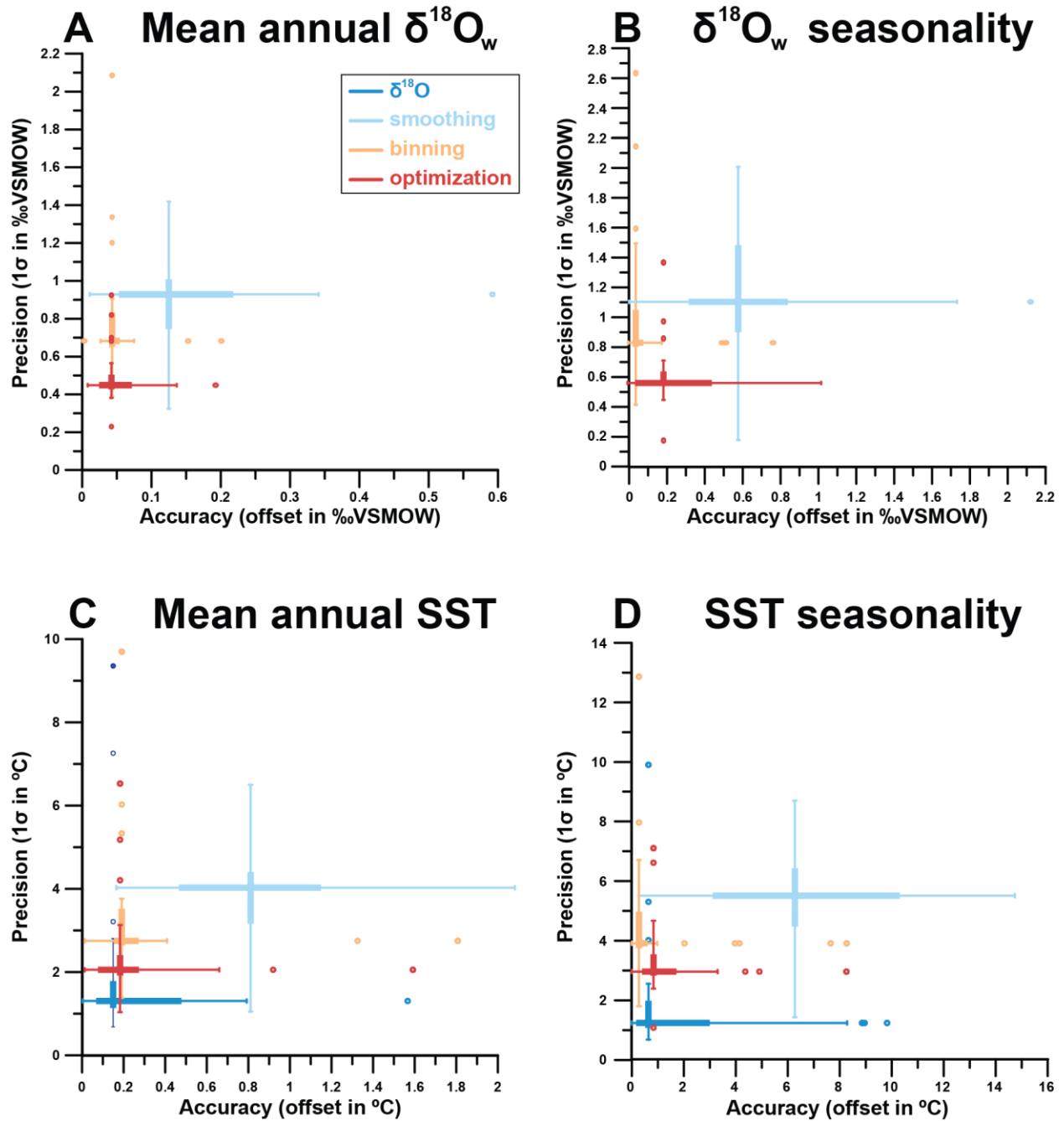
451

452 **Figure 12:** Effect of uncertainty in age model on the reproducibility (standard deviation on estimate) of  
 453 results of reconstructions of mean annual  $\delta^{18}\text{O}_w$  (A) and seasonal range in  $\delta^{18}\text{O}_w$  (B), mean annual SST  
 454 (C) and seasonal range in SST (D). Effect of uncertainty in age model on the accuracy (offset from true  
 455 value) of results of reconstructions of mean annual  $\delta^{18}\text{O}_w$  (E) and seasonal range in  $\delta^{18}\text{O}_w$  (F), mean annual  
 456 SST (G) and seasonal range in SST (H). Color coding follows the scheme in Fig. 1 and Fig. 4.

457

### 458 3.6 Effect of age model uncertainty

459 Uncertainty on the age model has a significant effect on both the precision and the accuracy (Fig. 12) of  
 460 reconstructions using all approaches. The  $\delta^{18}\text{O}$  reconstructions are most strongly affected by  
 461 uncertainties in the age model and suffer from a large decrease in precision with increasing age model  
 462 uncertainty (Fig. 12C-D). The high reproducibility of the  $\delta^{18}\text{O}$  approach in comparison with the  $\Delta_{47}$   
 463 approaches quickly disappears when age model uncertainty increases beyond 20-30 days. Accuracy of  
 464  $\delta^{18}\text{O}_c$ -based SST seasonality reconstructions initially improves with age model uncertainty (Fig. 12H).  
 465 However, this observation is likely caused by the fact that age model uncertainty was compared based on  
 466 conditions in case 9, which features a phase offset between SST and  $\delta^{18}\text{O}_w$  seasonality causing the  $\delta^{18}\text{O}$   
 467 method to be highly inaccurate even without age model uncertainty. The precision of **smoothing** and  
 468 **optimization** approaches also decreases with increasing age model uncertainty (Fig 12A-D), and the  
 469 **optimization** approach loses its precision advantage over the **binning** and **smoothing** approaches when  
 470 age model uncertainty increases beyond 30 days. The monthly **binning** approach is most resilient  
 471 against increasing age model uncertainty. Seasonality reconstructions through both the **binning** and  
 472 **optimization** approach quickly lose accuracy when age model uncertainty increases but the accuracy of  
 473 the **smoothing** approach remains the worst of all approaches in regardless of age model uncertainty  
 474 (Fig. 12E-H).



475

476 **Figure 13:** Overview of averages and ranges of accuracy (absolute offset from real value) and precision  
 477 (one standard deviation from the mean) on mean annual  $\delta^{18}\text{O}_w$  (A) and seasonal range in  $\delta^{18}\text{O}_w$  (B), mean  
 478 annual SST (C) and seasonal range in SST (D) within all cases using the four different reconstruction  
 479 approaches. Color coding follows the scheme in Fig. 1 and Fig. 4. Box-whisker plots for precision and  
 480 accuracy cross at their median values and outliers (colored symbols) are identified based on 2x the  
 481 interquartile difference.

482

## 483 4. Discussion

### 484 4.1 Performance of reconstruction approaches

#### 485 4.1.1 $\delta^{18}\text{O}_c$ vs $\Delta_{47}$ -based reconstructions

486 **Figure 13** summarizes the general reliability of the four approaches.  $\delta^{18}\text{O}$  reconstructions are generally  
487 less accurate than  $\Delta_{47}$ -based reconstructions (especially **binning** and **optimization**; see also **S9**). This is  
488 a consequence of the assumption that  $\delta^{18}\text{O}_w$  remains constant year-round, and that one knows its true  
489 value. Both these assumptions are problematic in absence of independent evidence of the value of  $\delta^{18}\text{O}_w$ ,  
490 especially in deep time settings (see e.g. Veizer and Prokoph, 2015; Henkes et al., 2018). The risk of this  
491 assumption is made clear when comparing cases in which  $\delta^{18}\text{O}_w$  is indeed constant year-round at the  
492 assumed value (0‰; e.g. cases 1-6 and 19-24) with cases in which shifts in  $\delta^{18}\text{O}_w$  occur, especially when  
493 these shifts are out of phase with respect to the SST seasonality (e.g. cases 9-11, 18 and 25-33; **Fig. 8C-**  
494 **D**). Cases mimicking or based on natural SST and SSS variability (cases 14-18 and 30-33) as well as the  
495 modern oyster data (**Fig. 6**) yield stronger inaccuracies in MAT and seasonality reconstructions, showing  
496 that even in many modern natural circumstances the assumption of constant  $\delta^{18}\text{O}_w$  is problematic.

497 It is important to consider that the value of mean annual  $\delta^{18}\text{O}_w$  remained very close to the assumed value  
498 of 0‰ (within 0.15‰) in all cases except for natural data cases 30 (-1.55‰), 32 (1.01‰; see **S5**) and the  
499 real oyster data (-1.42‰; **Fig. 5**). The SST values of these cases reconstructed using  $\delta^{18}\text{O}_c$  data show  
500 large offsets from their actual values (+6.7°C, -4.7°C and +10.3°C for case 30, case 32 and the real oyster  
501 data respectively; see **Fig. 6 and 8** and **S5**). These offsets are equivalent to the temperature offset one  
502 might expect from inaccurately estimating  $\delta^{18}\text{O}_w$  (~-4.6 °C/‰; Kim and O’Neil, 1997) and are only rivaled  
503 by the offset in MAT reconstructions of case 18 (+5.0°C), which has growth hiatuses obscuring the coldest  
504 half of the seasonal cycle. The fact that such differences in  $\delta^{18}\text{O}_w$  exist even in modern environments should  
505 not come as a surprise, given the available data on variability of  $\delta^{18}\text{O}_w$  (at least -3‰ to +2‰; e.g. LeGrande  
506 and Schmidt, 2006) and SSS (30 to 40; ESA, 2020) in modern ocean basins. However, it should warrant  
507 caution in using  $\delta^{18}\text{O}_c$  data for SST reconstructions even in modern settings. Implications for deep time  
508 reconstructions are even greater, given the uncertainty on and variability in global average (let alone local)  
509  $\delta^{18}\text{O}_w$  values (Jaffrés et al., 2007; Veizer and Prokoph, 2015). The complications of using  $\delta^{18}\text{O}_c$  as a proxy

510 for marine temperatures in deep time are discussed in detail in O'Brien et al. (2017), and Tagliavento et al.  
511 (2019).

512 The analytical uncertainty of individual  $\delta^{18}\text{O}_c$  aliquots (typically 1 S.D. of 0.05‰; e.g. de Winter et al., 2018)  
513 represents only ~1.1% of the variability in  $\delta^{18}\text{O}_c$  over the seasonal cycle (~4.3‰ for the default 20°C  
514 seasonality in case 1, following Kim and O'Neil, 1997). This is much smaller than the analytical uncertainty  
515 of  $\Delta_{47}$  (typically 1 S.D. of 0.02-0.04‰; e.g. Fernandez et al., 2018; de Winter et al., 2020b), which equates  
516 to 25-50% of the seasonal variability in  $\Delta_{47}$  (~0.08‰ for 20°C seasonality, following Bernasconi et al., 2018;  
517 see **S7**). This roughly 20-fold difference in relative precision causes  $\delta^{18}\text{O}_c$  based SST reconstructions to be  
518 much more precise (see **Figs 7, 9-12**) than those based on  $\Delta_{47}$ , and forces the necessity for grouping  $\Delta_{47}$   
519 data in reconstructions. However, as discussed above, the high precision of  $\delta^{18}\text{O}$  reconstructions is a  
520 misleading statistic if they are highly inaccurate.

521 Our results show that paleoseasonality reconstructions based on  $\delta^{18}\text{O}_c$  can only be relied upon if there is  
522 strong independent evidence of the value of  $\delta^{18}\text{O}_w$  and if significant sub-annual variability in  $\delta^{18}\text{O}_w$  (>0.3‰,  
523 equivalent to a 2-3°C SST variability; see **Fig. 9-10**; Kim and O'Neil, 1997) can be excluded with confidence.  
524 Examples of such cases include fully marine environments unaffected by influxes of (isotopically light)  
525 freshwater or evaporation (increasing  $\delta^{18}\text{O}_w$ ; Rohling, 2013). Carbonate records from environments with  
526 more stable  $\delta^{18}\text{O}_w$  conditions include, for example, the *A. islandica* bivalves from considerable depth (30-  
527 50m) in the open marine Northern Atlantic (e.g. Schöne et al., 2005, on which case 33 is based). However,  
528 even here variability in  $\delta^{18}\text{O}_{sw}$  due to, for example, shifting influence of different bottom water masses  
529 cannot be fully excluded. Previous reconstruction studies show that  $\delta^{18}\text{O}_w$  in smaller basins are heavily  
530 influenced by the processes affecting  $\delta^{18}\text{O}_w$  on smaller scales, such as local evaporation and freshwater  
531 influx from nearby rivers (e.g. Surge et al., 2001; Petersen et al., 2016). Consequently, accurate quantitative  
532 reconstructions of seasonal range in shallow marine environments with extreme seasonality may not be  
533 feasible using the  $\delta^{18}\text{O}$  approach, because these environments are invariably characterized by significant  
534 fluctuations in  $\delta^{18}\text{O}_w$  and growth rate.

535 While variability in  $\delta^{18}\text{O}_w$  compromises accurate  $\delta^{18}\text{O}$ -based seasonality reconstructions, the compilation  
536 in **Fig. 3** shows that its influence on the  $\delta^{18}\text{O}$  records is too small to affect the shape of the record to such

537 a degree that seasonality is fully obscured. While natural situations with  $\delta^{18}\text{O}_w$  fluctuations large enough to  
538 totally counterbalance the effect of temperature seasonality on  $\delta^{18}\text{O}$  records are imaginable, these cases  
539 are likely rare. This means that chronologies based on  $\delta^{18}\text{O}$  seasonality, which are a useful tool to anchor  
540 seasonal variability in absence of independent growth markers (e.g. Judd et al., 2018; de Winter, 2021b),  
541 are reliable in most natural cases.

#### 542 *4.1.2 Seasonality reconstructions using moving averages (**smoothing**)*

543 Of the three methods for combining  $\Delta_{47}$  data, the **smoothing** approach clearly performs worst in all four  
544 reconstructed parameters (MAT, SST seasonality, mean annual  $\delta^{18}\text{O}_w$  and  $\delta^{18}\text{O}_w$  seasonality), both in  
545 terms of accuracy and precision (**Fig. 13**). While applying a moving average may be a good strategy for  
546 lowering the uncertainty of  $\Delta_{47}$ -based temperature reconstructions in a long time series (e.g. Rodríguez-  
547 Sanz et al., 2017), the method underperforms in cases where baseline and amplitude of a periodic  
548 component (e.g. MAT and SST seasonality) are extracted from a record. This is likely due to the smoothing  
549 effect of the moving average, which reduces the seasonal cycle and causes highly inaccurate seasonality  
550 reconstructions (offsets mounting to  $>6^\circ\text{C}$ ; **Fig. 13**). This bias is especially detrimental in cases where the  
551 seasonal cycle is obscured by seasonal growth halts (e.g. case 18), multi-annual trends in growth (e.g.  
552 case 4, 14 and 17) and multi-annual trends in SST (e.g. case 15 and 17; see **Fig. 7 and 8**). The poor  
553 performance of the **smoothing** approach can be slightly mitigated by increasing sampling resolution (**Fig**  
554 **9**), but even at high sampling resolutions (every 0.1 or 0.2 mm) the method still fails to reliably resolve  
555 seasonal SST ranges below  $5^\circ\text{C}$  even in idealized cases (case 19-21; **Fig. 10**). Increasing the number of  
556 samples by analyzing longer records does not improve the result, because smoothing of the seasonal cycle  
557 by a moving average window introduces the same dampening bias if the temporal sampling resolution  
558 (number of samples per year) remains equal (**Fig. 11**).

559 More critically, employing the **smoothing** method may give the illusion that seasonality is more reduced,  
560 and severely bias reconstructions. This bias highlights the importance of using the official meteorological  
561 definition of seasonality as the difference between the averages of warmest and coldest month in  
562 paleoseasonality work (O'Donnell et al., 2012). This definition is much more robust than the “annual range”  
563 often cited based on maxima and minima in  $\delta^{18}\text{O}_c$  records. This “annual range” strongly depends on

564 sampling resolution, which is typically <12 samples/yr (Goodwin et al., 2003), equivalent to the third lowest  
565 sampling interval (0.75 mm) simulated in this study. Therefore, we strongly recommend future studies to  
566 adhere to the monthly definition of seasonality to foster comparison between studies. While inter-annual  
567 variability is lost by combining data from multiple years into monthly averages, this approach increases  
568 precision, accuracy and comparability of paleoseasonality results. Inter-annual variability can still be  
569 discussed from plots of raw data plotted in time or sampling domain.

#### 570 4.1.3 Monthly **binning**, sample size **optimization** and age model uncertainty

571 Overall, the most reliable paleoseasonality reconstructions can be obtained from either **binning** or  
572 **optimization** (Fig. 13). In general, **optimization** is slightly more precise, while **binning** yields more  
573 accurate estimates of seasonal range in SST and  $\delta^{18}\text{O}_w$  (Fig. 13B and D). The more flexible combination  
574 of aliquots in the **optimization** routine yields improved precision (especially on mean annual averages) in  
575 cases where parts of the record are undersampled or affected by hiatuses and simultaneous fluctuations  
576 in both SST and  $\delta^{18}\text{O}_w$  (e.g. case 3-6, 14-18, 30-33). The downside of this flexibility is that in case of larger  
577 sample sizes, the seasonal variability may be dampened, like in the **smoothing** approach (see 4.1.2). This  
578 apparent dampening effect may be reduced by allowing the sample size of summer and winter samples to  
579 vary independently in the **optimization** routine, at the cost of higher computational intensity due to the  
580 larger number of sample size combinations (see 2.1 and 4.2.2). The rigid grouping of data in monthly bins  
581 in **binning** prevents this dampening and therefore yields slightly more accurate estimates of seasonal  
582 ranges in SST and  $\delta^{18}\text{O}_w$ . A caveat of **binning** is that it requires a very reliable age model of the record, at  
583 least on a monthly scale. If the age model has a large uncertainty, there is a risk that samples are grouped  
584 in the wrong month, which compromises the accuracy of **binning** reconstructions, especially for  
585 reconstructions of seasonal range (Fig 12H). This problem is exacerbated by potential phase shifts between  
586 seasonality in paleoclimate variables (SST and  $\delta^{18}\text{O}_w$ ) and calendar dates, which may occur in the presence  
587 of a reliable age model.

588 Previous authors attempted to circumvent the dating problem by analyzing high-resolution  $\delta^{18}\text{O}_c$  transects  
589 and subsequently sampling the seasonal extremes for clumped isotope analyses (Keating-Bitonti et al.,  
590 2011; Briard et al., 2020). While this approach does not require sub-annual age models, it has several



591 disadvantages compared with the **binning** and **optimization** approaches: Firstly, it requires separate  
592 sampling for  $\delta^{18}\text{O}_c$  and  $\Delta_{47}$ , which may not be possible in high-resolution carbonate archives due to sample  
593 size limitations. Analyzing small aliquots for combined  $\delta^{18}\text{O}_c$  and  $\Delta_{47}$  analyses consumes less material.  
594 Secondly, individual summer and winter temperature reconstructions require large (> 1.5 mg; e.g.  
595 Fernandez et al., 2017)  $\Delta_{47}$  samples from seasonal extremes, which causes more time-averaging than the  
596 approaches combining small aliquots. Finally, the position of seasonal extremes estimated from the  $\delta^{18}\text{O}_c$   
597 record may not reflect the true seasonal extent if seasonal SST and  $\delta^{18}\text{O}_w$  cycles are not in phase (as in  
598 case 9), causing the seasonal  $\Delta_{47}$ -based SST reconstructions to underestimate the temperature  
599 seasonality. In such cases,  $\delta^{18}\text{O}_c$  and  $\Delta_{47}$  analyses on small aliquots allow the seasonality in SST and  $\delta^{18}\text{O}_w$   
600 to be disentangled, yielding more accurate seasonality reconstructions.

601 Techniques for establishing independent age models for climate archives range from counting of growth  
602 layers or increments (Schöne et al., 2008; Huyghe et al., 2019), modelling and extracting of rhythmic  
603 variability in climate proxies through statistical approaches (e.g. De Ridder et al., 2007; Goodwin et al.,  
604 2009; Judd et al., 2018; de Winter, 2021b) and interpolation of uncertainty on absolute dates (e.g. Scholz  
605 and Hoffman, 2011; Meyers, 2019; Sinnesael et al., 2019). While propagating uncertainty in the data on  
606 which age models are based onto the age model is relatively straightforward, errors on underlying *a priori*  
607 assumptions such as linear growth rate between dated intervals, (quasi-)sinusoidal forcing of climate cycles  
608 and the uncertainty on human-generated data such as layer counting are very difficult to quantify (e.g.  
609 Comboul et al., 2014) and may not be normally distributed. Results of cases 25-29 show that uncertainties  
610 in the age domain can significantly compromise reconstructions (**Fig. 12**). Within the scope of this study,  
611 only the effect of symmetrical, normally distributed uncertainties on an artificial case with phase decoupled  
612 SST and  $\delta^{18}\text{O}_w$  seasonality (case 9) was tested. The effects of other types of uncertainties on the  
613 reconstructions remain unknown, highlighting an unknown uncertainty in paleoseasonality and other high-  
614 resolution paleoclimate studies that may introduce bias or lead to over-optimistic uncertainties on  
615 reconstructions. Future research could quantify this unknown uncertainty by propagating estimates of  
616 various types of uncertainty on depth values of samples and on the conversion from sampling to time  
617 domain in age models.

618 **4.2 Conditions influencing success of reconstructions**

619 The reliability (accuracy and precision) of SST and  $\delta^{18}\text{O}_w$  reconstructions depend on case-specific  
620 conditions. The range of case studies tested in this study allowed us to evaluate the effect of variability in  
621 SST, growth rate,  $\delta^{18}\text{O}_w$ , sampling resolution and record length relative to the control case (case 1; see  
622 **S1**). A summary of the effects of these changes is given in **Table 2**.

623

Variable	cases	Metric	Effect on reconstructions			
			$\delta^{18}\text{O}$	smoothing	binning	optimization
SST	12	Precision	0	+++	+	0
	15					
	17					
	19-21 30-33	Accuracy	+	+	0	+
Growth rate	2-6	Precision	+	++	++	+
	14-18					
	30-33	Accuracy	+	++	0	+
$\delta^{18}\text{O}_w$	7-11	Precision	+	++	0	0
	13-18					
	30-33	Accuracy	+++	+++	+	++
Sampling resolution	1-33	Precision	0	+++	++	++
		Accuracy	+	+	+++	+
Record length	22-24	Precision	0	0	+++	++
		Accuracy	+	0	++	++
Age model uncertainty?	25-29	Precision	+++	++	0	++
		Accuracy	+	+	++	++

624 **Table 2:** Qualitative summary of the effects of changes in variables relative from the ideal case on  
625 reconstructions using the four approaches. The “cases” column lists cases in which the changes in the  
626 respective variable relative to the control case (case 1) were represented (see **Table 1** and **S1**). “0” =  
627 negligible effect, “+” = weak increase in uncertainty, “++” = moderate increase in uncertainty, “+++” = strong  
628 increase in uncertainty. Precision and accuracy of all tests is given in **S9**.

629

#### 630 4.2.1 SST variability

631 Variability in water temperature most directly affects the proxies under study. By default (case 1), SST  
632 varies sinusoidally around a MAT of 20°C with an amplitude of 10°C (see **2.3.3**, **Fig. 2** and **S1**). In cases in  
633 which multi-annual variability in SST is simulated (e.g. case 15 and 17), the accuracy of SST reconstructions  
634 using  $\delta^{18}\text{O}$  and **optimization** are reduced, while the **binning** approach is less strongly affected. Examples  
635 of such multi-annual cyclicity are El-Niño Southern Oscillation (ENSO; Philander, 1983) or North Atlantic  
636 Oscillation (NOA; Hurrell, 1995). The effect is especially large in case 17, which simulates a tropical  
637 environment with reduced SST seasonality and a strong multi-annual cyclicity. This type of environment is  
638 analogous to the environment of tropical shallow water corals, which are often used as archives for ENSO  
639 variability (e.g. Charles et al., 1997; Fairbanks et al., 1997) and is similar to tropical cases from the  
640 Australian Great Barrier Reef (case 31) and Red Sea (case 32; see **Fig. 3**). We therefore recommend using  
641 the **binning** approach on carbonate records where multi-annual cyclicity is prevalent and if a reliable age  
642 model can be established for these records (as in e.g. Sato, 1999; Scourse et al., 2006; Miyaji et al., 2010).

643 4.2.2 Growth rate variability and hiatuses

644 **Figures 7 and 8** show that variations in the growth rate of records, including the occurrence of hiatuses,  
645 have a strong effect on reconstructions, especially using the **smoothing** approach. In general, hiatuses  
646 and slower growth reduce precision of monthly SST and  $\delta^{18}\text{O}_w$  reconstructions by reducing mean temporal  
647 sampling resolution (samples/yr; see **Fig. 9**), and because parts of the record are undersampled. The effect  
648 on accuracy depends strongly on the timing of changes in growth rate or the occurrence of hiatuses. Cases  
649 2-6 simulate specific growth rate effects and can be used to test these differences. The **smoothing** method  
650 is especially sensitive to changes in growth rate that take place in specific seasons, such as hiatuses in  
651 winter (case 2) or summer (case 3) and growth peaks in summer (case 5) or spring (case 6). The other  
652 reconstruction approaches are less affected by this bias, because they generally do not mix samples from  
653 different seasons. The  $\delta^{18}\text{O}$  method is especially well suited to deal with changes in growth rate because  
654 it does not require combining different aliquots for accurate SST reconstructions. The **binning** and  
655 **optimization** approaches are slightly less reliable in cases where growth rate decreases linearly or  
656 seasonally along the entire record (cases 4-6; **Fig. 2**). Because these two methods consider all samples in  
657 the records at once, they are more sensitive to changes in temporal sampling resolution along the record.  
658 It is worth noting that **optimization** is especially sensitive to sharp changes in growth rate in summer (e.g.  
659 cases 11, 14, 16 and 17) because those conditions force the **optimization** routine to use larger sample  
660 sizes or include samples outside the warmest month for summer temperature estimates. A potential solution  
661 to this problem could be to allow sample sizes of summer and winter groups to vary independently in the  
662 **optimization** routine (see 2.1). This would allow sample size in the undersampled season (in this case:  
663 summer) to become larger than that at the other end of the  $\delta^{18}\text{O}_c$  spectrum, reducing uncertainty on the  
664 more densely sampled season and therefore improving the entire seasonality reconstruction.

665 A worst-case scenario is represented by case 18, where the cold half of the year is not recorded. Such  
666 cases result in strong biases in reconstructions of mean annual and seasonal ranges in SST and  $\delta^{18}\text{O}_w$ ,  
667 regardless of which method is used. In such extreme cases the record simply contains insufficient  
668 information to reconstruct variability in growth rate, SST and  $\delta^{18}\text{O}_w$ , and it seems that no statistical method  
669 would enable this missing information to be recovered. The solution for these reconstructions would be to

670 establish reliable age models, independent of  $\delta^{18}\text{O}$  or  $\Delta_{47}$  data, which show that a large part of the seasonal  
671 cycle is missing. All methods used in this study rely on a conversion of SST and  $\delta^{18}\text{O}_w$  reconstructions to  
672 the time domain to define monthly time bins. This conversion breaks down in fossil examples when the  
673 seasonal cycle cannot be extracted from the archive, which happens when half of the seasonal cycle or  
674 more is obscured by growth hiatuses, as exemplified in case 18.

675 While hiatuses encompassing half of the seasonal cycle are uncommon, changes in growth rate are  
676 common in accretionary carbonate archives because conditions for (biotic or abiotic) carbonate  
677 mineralization often vary over time. This variability is either driven by biological constraints, such as  
678 senescence (e.g. Schöne, 2008; Hendriks et al., 2012), the reproductive cycle (Gaspar et al., 1999) or  
679 stress (Surge et al., 2001; Compton et al., 2007) or by variations in the environment that promote or inhibit  
680 carbonate production, such as seasonal variations in temperature (Crossland, 1984; Bahr et al., 2017) or  
681 precipitation (Dayem et al., 2010; Van Rampelbergh et al., 2014). In general, such conditions occur more  
682 frequently in mid- to high-latitude environments than in low-latitudes, and in more coastal environments  
683 rather than in open marine settings, because these environments contain stronger variations in the factors  
684 that influence growth rates (e.g. temperature, precipitation or freshwater influx; e.g. Surge et al., 2001;  
685 Ullmann et al., 2010). This difference was simulated in the cases representing natural variability (case 14-  
686 18 and 30-33). Accuracy in the coastal high-latitude settings (cases 16, 18 and 29) are indeed more strongly  
687 affected by changes in growth rate. Because in such highly variable environments growth rate variability  
688 often co-occurs with variability in  $\delta^{18}\text{O}_w$ , using  $\delta^{18}\text{O}_c$ -based reconstructions is not advised, unless  $\delta^{18}\text{O}_w$   
689 variability can be constrained or neglected (which is rare in these environments).

690 Additional complications include **that** the lack of constraint on growth rate variability because of  
691 uncertainties in the record's age model (see 4.1.3) and the effect of growth rate variability on the sampling  
692 resolution. The effect of growth rate on time-averaging within samples was not specifically tested in this  
693 study but introduces uncertainty in practice when archives with variable growth rate are sampled at a  
694 constant sampling resolution in the depth domain. In this case, parts of the archive with a lower growth rate  
695 yield more time-averaged samples, potentially dampening one extreme of the seasonal cycle (e.g. Goodwin  
696 et al., 2003). In highly dynamic environments it is challenging to isolate all variables that introduce bias, and

697 irregular variability in growth rate and  $\delta^{18}\text{O}_w$  will invariably introduce uncertainty in SST reconstructions,  
698 even when applying the best  $\Delta_{47}$ -based approaches (e.g. **binning** and **optimization**). In such examples,  
699 the results of natural variability cases (14-18 and 30-33) and of the real oyster data (**Fig. 6**) serve as  
700 benchmarks for the degree of uncertainty that may remain unexplained in these records.

#### 701 *4.2.3 Variability in $\delta^{18}\text{O}_w$*

702 As discussed in **4.1.1**, these variations in  $\delta^{18}\text{O}_w$  have a large effect on the accuracy of  $\delta^{18}\text{O}_c$ -based  
703 reconstructions, and their occurrence constitutes the main advantage of applying the  $\Delta_{47}$  thermometer  
704 (Eiler, 2011). However, results of cases 7-11 in **Fig. 8** and **Table 2** show that  $\delta^{18}\text{O}_w$  variations can also bias  
705  $\Delta_{47}$ -based reconstructions, especially those of seasonal ranges and using the **smoothing** approach.  
706 **Smoothing** reconstructions are biased by these  $\delta^{18}\text{O}_w$  shifts in much the same way as they are affected  
707 by shifts in growth rate (see **4.2.2**). The **optimization** approach is sensitive to seasonal changes in  $\delta^{18}\text{O}_w$   
708 in antiphase with SST seasonality and by increases in  $\delta^{18}\text{O}_w$  in summer (e.g. due to excess evaporation;  
709 e.g. case 11), especially when used for reconstructions of  $\delta^{18}\text{O}_w$  seasonality. This effect arises because  
710 the **optimization** approach orders data based on  $\delta^{18}\text{O}_c$  and  $\Delta_{47}$  seasonality to isolate the  $\delta^{18}\text{O}_w$ -SST  
711 relationship. Both antiphase  $\delta^{18}\text{O}_w$  seasonality and summer evaporation dampen the seasonal  $\delta^{18}\text{O}_c$  cycle  
712 and therefore influences the reconstruction of the  $\delta^{18}\text{O}_w$ -SST relationship. A good example of this is seen  
713 in the real oyster data (**Fig. 6**), where  $\delta^{18}\text{O}_w$  and SST vary in phase and  $\delta^{18}\text{O}_w$  dampens the SST  
714 seasonality. The **binning** approach is more robust against  $\delta^{18}\text{O}_w$  variability that dampens the seasonal  
715 cycle and is therefore a better choice for absolute SST reconstructions in environments where summer  
716 evaporation or other  $\delta^{18}\text{O}_w$  variability in phase with SST seasonality is expected to occur, if the age model  
717 is reliable enough to allow monthly binning of raw data (see **4.1.3**). Indeed, reconstructions from the  
718 lagoonal environment (case 16) and Red Sea case (case 32 which is characterized by strong summer  
719 evaporation; e.g. Titschack et al., 2010) show that **binning** is the most reliable choice in these  
720 environments.

#### 721 *4.2.4 Variability in sampling resolution and record length*

722 Other factors influencing the **effectivity** of reconstructions are the sampling resolution and the length of the  
723 record. Many of the cases discussed in this study represent idealized cases with comparatively high

724 sampling resolutions over comparatively long (12 yr) paleoseasonality records, which yield large sample  
725 sizes. By comparison, the typical age of mollusks, which are often used as paleoseasonality archives, is 2-  
726 5 years (Ivany, 2012). Records with the highest sampling resolutions (0.1 and 0.2 mm) contain up to 1200  
727 samples. Generating such records is not impossible, but it is highly unlikely to be applied in paleoclimate  
728 studies given the limitation of resources (e.g. instrument time) and the desire to analyze multiple records  
729 from different specimens, species, localities or ages to gain a better understanding of the variability in  
730 paleoseasonality (e.g. Goodwin et al., 2003; Schöne et al., 2006; Petersen et al., 2016). In some cases  
731 large datasets are meticulously collected from single carbonate records (e.g. Schöne et al., 2005;  
732 Vansteenberghe et al., 2016; de Winter et al., 2020a; Shao et al., 2020). However, in such studies, the aim  
733 is often to investigate variability at a higher (e.g. daily; de Winter et al., 2020a) resolution or longer  
734 timescales (e.g. decadal to millennial; Schöne et al., 2005; Vansteenberghe et al., 2016; Shao et al., 2020)  
735 in addition to the seasonal cycle, rather than to improve the reliability of reconstructing one type of variability  
736 (e.g. seasonality) alone.

737 **Fig. 9** shows that increasing temporal sampling resolution (samples/yr) improves both the accuracy and  
738 precision of all  $\Delta_{47}$ -based reconstructions. This occurs because  $\Delta_{47}$  samples have a large analytical  
739 uncertainty (see 4.1.2) and grouping of data therefore improves reconstructions. The decrease in precision  
740 of  $\delta^{18}\text{O}_c$ -based reconstructions (**Fig. 9C-D**) is explained by the fact that the analytical uncertainty of  $\delta^{18}\text{O}_c$   
741 measurements is much smaller than the variability introduced by natural sub-annual variability in SST and  
742  $\delta^{18}\text{O}_w$  unrelated to the seasonal cycle (see **S4**). Therefore, higher sampling resolutions allow  $\delta^{18}\text{O}_c$  records  
743 to better capture this sub-seasonal variability, which introduces more noise on the seasonal cycle (reducing  
744 precision) but causes monthly mean SST and  $\delta^{18}\text{O}_w$  to be more accurately reconstructed. Towards higher  
745 sampling resolutions, the gap in precision between  $\delta^{18}\text{O}_c$ - and  $\Delta_{47}$ -based reconstructions closes, eventually  
746 (in an ideal case) diminishing the advantage of high analytical precision in  $\delta^{18}\text{O}_c$  measurements (**Fig. 9C-**  
747 **D**).

748 An optimum sample resolution can be defined for each method after which improving sampling resolution  
749 does not significantly improve the reliability of the reconstruction (as in de Winter et al., 2017). **Figure 9**  
750 shows that this optimum varies depending on which variable (MAT, SST seasonality, mean annual  $\delta^{18}\text{O}_w$

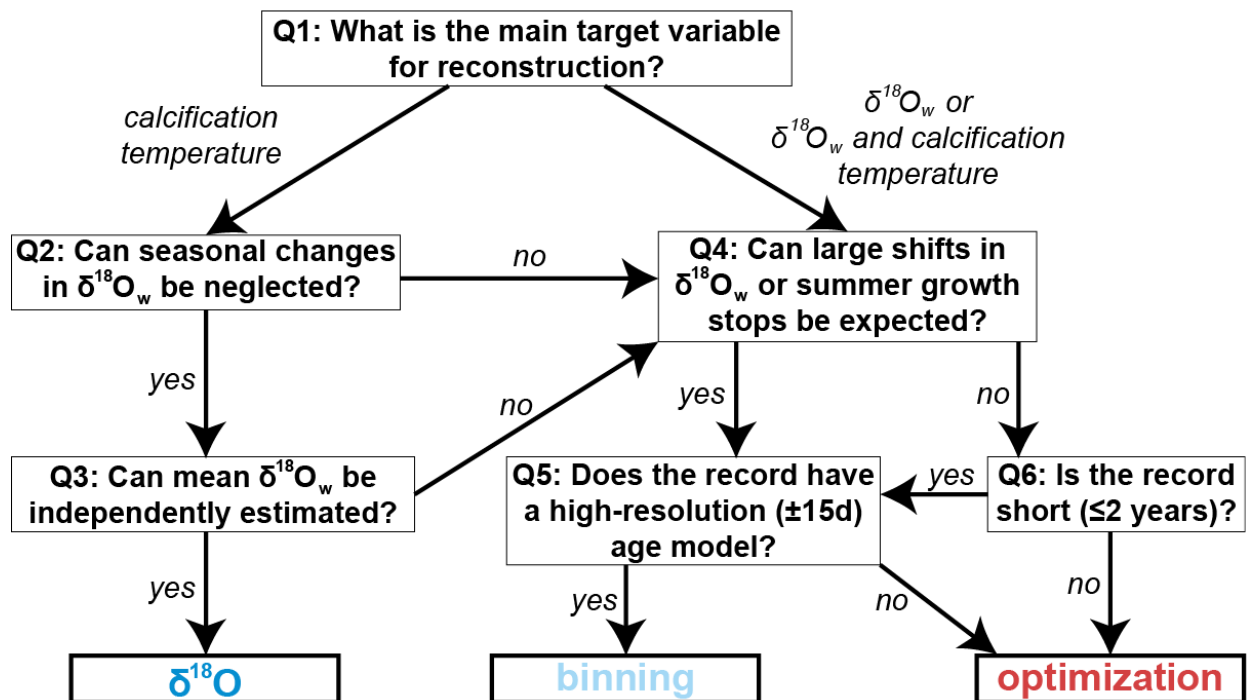
751 or  $\delta^{18}\text{O}_w$  seasonality) is reconstructed. Therefore, **Fig. 9** will allow future researchers to determine the  
752 sampling resolution that is tailored to their purpose. In general, the improvement after a sample size of 20-  
753 30 samples per year is negligible for the **binning** and **optimization** methods if the total number of samples  
754 (depending on both sampling resolution and record length) is sufficient for monthly temperature  
755 reconstructions. Our data show that 200-250 paired  $\delta^{18}\text{O}_c$  and  $\Delta_{47}$  measurements are in general sufficient  
756 for a standard deviation of 2-3°C on monthly SST reconstructions using the **binning** or **optimization**  
757 approach (**Fig. 10; S5**).

758 Record length only has a minimal influence on the **optimization** method but for very short records ( $\leq 2$   
759 years) **binning** becomes very imprecise, especially at low sampling resolutions (**Fig. 11**). The reason is  
760 that the sample size within monthly time bins becomes too small in these cases, while the more flexible  
761 sample size window of the optimization routine circumvents this problem. The choice between these two  
762 approaches should therefore be based on a tradeoff between the length of the record (in time) and the  
763 number of samples that can be retrieved from it. As a result, shorter-lived, fast-growing climate archives,  
764 such as large or fast-growing (e.g. juvenile) mollusk shells, are best sampled using a high temporal  
765 resolution ( $>30$  samples/yr) sampling strategy with the **optimization** approach. Longer lived archives with  
766 a lower mineralization rate, such as annually laminated speleothems, corals and gerontic mollusks, are  
767 best sampled using long time series at monthly resolution using the **binning** approach.

768 A simplified decision tree that could guide sampling strategies for future paleoseasonality studies is shown  
769 in **Figure 14**. Note that choices and tradeoffs for these reconstructions may differ depending on the archive  
770 and environment in which it formed (see discussion above).



## Schematic guide to reconstructing SST and $\delta^{18}\text{O}_w$ from accretionary carbonate archives



771

772 **Figure 15:** Schematic guide to choosing the right approach for reconstructing annual mean or seasonality  
 773 in SST and  $\delta^{18}\text{O}_w$  from accretionary carbonate archives. Recommendations are based on the results of  
 774 testing all four approaches on the entire range of cases. Researchers can follow the six steps (questions  
 775 Q1-6) to decide on the right approach for reconstructing the target variable. Guidelines are based on  
 776 **minimizing** both accuracy and precision (see details in **S9**). Note that the **smoothing** approach is never the  
 777 best choice. The choice between the two remaining  $\Delta_{47}$ -based approaches (**binning** and **optimization**)  
 778 relies heavily on the situation and may be driven by a preference for more accurate or more precise results.

779

### 780 4.3 Implications for clumped isotope sample size

781 The **optimization** technique for grouping  $\Delta_{47}$  aliquots for accurate SST and  $\delta^{18}\text{O}_w$  reconstructions allows  
 782 us to assess the limitations of the clumped isotope thermometer for temperature reconstructions from high-  
 783 resolution carbonate archives. The optimal sample size given by the approach is different for different cases  
 784 and depends on the temporal sampling resolution and the characteristics of the record (see **S4**). As  
 785 expected, in cases more like the ideal case (case 1), optimal sample sizes are low (~14-24), while sample  
 786 sizes increase in more complicated cases based on simulated natural environments (case 14-18) or cases  
 787 based on actual SST and SSS data (cases 30-33). More confined SST seasonality (cases 19-21) also  
 788 requires larger samples to reconstruct (up to 100 samples in some cases). This is not surprising, because

789 variability within samples will increase in records in which the seasonality is smaller or more obscured by  
790 other environmental variability. The optimal sample size between cases and sampling resolutions is not  
791 normally distributed but tails towards high sample sizes with some extreme outliers (Shapiro Wilk test  $p <<$   
792 0.05; **S10**). The median sample size of all our simulations is 17 aliquots. This number lies between the  
793 minimum number of 14 ~100  $\mu\text{g}$  replicates of standards calculated by Fernandez et al. (2017) and the  
794 minimum of 20-40 ~100  $\mu\text{g}$  aliquots required for optimal paleoseasonality reconstruction from fossil bivalves  
795 by de Winter et al. (2020b). This is to be expected since many of the cases explored in this study represent  
796 ideal cases compared with the natural situation. However, in these virtual cases a measure of random sub-  
797 annual variability in SST and  $\delta^{18}\text{O}_w$  was added (see **Fig. 4** and **S2**), simulating a more realistic environment  
798 and resulting in poorer precision than replicates of a carbonate standard (as in Fernandez et al., 2017). Our  
799 simulations show that the optimum number of samples to be combined in seasonality studies depends on  
800 both the analytical uncertainty of  $\Delta_{47}$  measurements (as represented by the estimate in Fernandez et al.,  
801 2017) and the variability between aliquots pooled within a sample that is attributed to actual variability within  
802 the record (as represented by our simulations and the estimate in de Winter et al. 2020b). The optimal  
803 sample size is therefore a good measure for the limitations of temperature variability that can be resolved  
804 in a record and can help researchers decide which strategy to apply for combining measurements to obtain  
805 the most reliable paleoseasonality estimates, or to decide whether extra sampling is required, even if the  
806 chosen approach is not to use the **optimization** routine itself. Note that the optimum sample size is kept  
807 equal for summer and winter samples in this study, and that the **optimization** approach can likely achieve  
808 better performance by considering unequal sample sizes in opposite seasons (see 4.1.3 and 4.2.2). While  
809 this added flexibility comes at a higher computational cost due to the increased number of possible sample  
810 size combinations to be considered, future studies should investigate whether this updated **optimization**  
811 approach could yield more reliable seasonality reconstructions.

#### 812 **4.4 Implications for other sample size problems**

813 While the discussion above focuses on optimizing approaches for combining samples for clumped  
814 isotope analyses in paleoseasonality reconstructions, the problem of combining samples to reduce  
815 uncertainty and isolate variation in datasets is very common (e.g. Zhang et al., 2004; Merz and Thieken,

816 2005; Tsukakoshi, 2011). Therefore, the approaches outlined and tested in this study have applications  
817 beyond paleoseasonality reconstructions. Examples of other problems that could benefit from applying  
818 similar approaches for reducing the uncertainty of estimates of target variables while minimizing the  
819 number of analyses required to meet analytical requirements include: (1) reconstructing  
820 paleoenvironmental variability in the terrestrial realm from tooth bioapatite (e.g. Passey and Cerling,  
821 2002; Kohn, 2004; Van Dam and Reichart, 2009; de Winter et al., 2016), (2) quantitative time series  
822 analysis of orbital cycles in stratigraphic records (e.g. Lourens et al., 2010; de Vleeschouwer et al., 2017;  
823 Noorbergen et al., 2017; Westerhold et al., 2020), (3) strontium isotope dating (e.g. McArthur et al., 2012;  
824 de Winter et al., 2020c), (4) reconstructing sub-seasonal variability from ultra-high-resolution records (e.g.  
825 from fast-growing mollusks and gastropods; e.g. Sano et al., 2012; Warter and Müller, 2017, de Winter et  
826 al., 2020d; Yan et al., 2020), and (5) reconstructing sea surface and deep-sea temperatures across short-  
827 lived (10–100 kyr) episodes of climate change or climate shifts from deep marine archives characterized  
828 by low sedimentation rates (e.g. Lear et al., 2008; Jenkyns, 2010; Stap et al., 2010; Lauretano et al.,  
829 2018). A more detailed discussion of the implications for other sample size problems is provided in the  
830 **Supplementary Information.**

831

## 832 5. Conclusions and recommendations

833 The performance of three  $\Delta_{47}$ -based approaches to reconstruct seasonality from accretionary carbonate  
834 archives was evaluated in comparison with conventional  $\delta^{18}\text{O}_c$ -based reconstructions in a wide range of  
835 case studies. From the results, we conclude that while  $\delta^{18}\text{O}_c$ -based reconstructions ( **$\delta^{18}\text{O}$** ) yield superior  
836 precision for SST reconstructions, this method runs a high risk of yielding inaccurate results due to innate  
837 assumptions about the value of  $\delta^{18}\text{O}_w$ , which must be estimated and assumed constant year-round. Unless  
838  $\delta^{18}\text{O}_w$  can be independently constrained or variability in  $\delta^{18}\text{O}_w$  can be neglected,  $\Delta_{47}$ -based reconstructions  
839 should be the method of choice for absolute mean annual temperature and SST seasonality  
840 reconstructions. Various techniques for combining  $\Delta_{47}$  data were evaluated. Our findings suggest that  
841 smoothing  $\Delta_{47}$  data using a moving average almost **always cases** in a dampening of the seasonal cycle  
842 which severely hampers recovery of seasonality. Applying the **smoothing** approach results in inaccuracies  
843 in reconstructions of MAT as well, especially in cases where part of the seasonal cycle is obscured by  
844 variability in growth rate or multi-annual trends. More reliable seasonality reconstructions are achieved with  
845 two approaches for combining  $\Delta_{47}$  data using time binning (**binning**) or applying a flexible sample size  
846 optimization (**optimization**) approach. Of these two approaches, **optimization** achieves better precision  
847 and can resolve smaller seasonal temperature differences with confidence. However, **binning** is often more  
848 accurate, and outperforms **optimization** as the most reliable approach. This is especially true in cases with  
849 growth stops or  $\delta^{18}\text{O}_w$  changes in phase with temperature seasonality (e.g. strong seasonal evaporation or  
850 freshwater influx) and in longer multi-annual time series with a reliable age model. **Optimization** is the  
851 better choice for shorter (<3 years) records, especially if the sampling resolution can be increased, such as  
852 in short, fast growing climate archives.

853 Despite the focus on the problem of resolving seasonality in carbonate archives, the findings in this study  
854 have applications for other problems in earth science where sample size and sampling resolution put limits  
855 on the ability to resolve specific trends, events, and cycles from time series. While the above-mentioned  
856 recommendations of the **optimization** and **binning** methods are likely valid for most studies aiming to  
857 quantify the mean and amplitude of a specific cycle or event (equivalent to MAT and SST seasonality),

858 (dynamic) moving averages (**smoothing**) are expected to yield the best results in studies quantifying  
859 aperiodic trends from longer data series.

860

#### 861 **Code availability**

862 All scripts used to make the calculations described in this study are compiled in the documented R package  
863 “seasonalclumped”, which is freely available on the open-source online R-database CRAN (de Winter,  
864 2021a; <https://cran.r-project.org/web/packages/seasonalclumped>). Annotated R scripts used to make  
865 calculations for this study are available in the digital supplement uploaded to the open-source online  
866 repository Zenodo ([www.doi.org/10.5281/zenodo.3899926](http://www.doi.org/10.5281/zenodo.3899926)).

867

#### 868 **Data availability**

869 Supplementary data, figures and tables as well as all scripts used to do the calculations and create the  
870 virtual datasets used in this study are deposited in the open-source online repository Zenodo  
871 ([www.doi.org/10.5281/zenodo.3899926](http://www.doi.org/10.5281/zenodo.3899926)). Virtual datasets generated within the context of this study are also  
872 made available as datafiles within the R package that contains the scripts used for this study  
873 (“seasonalclumped”; de Winter, 2021a; see <https://cran.r-project.org/web/packages/seasonalclumped>).

874

#### 875 **Author contributions**

876 NJW designed the study, wrote the scripts for all calculations, and created a first draft of the manuscript  
877 text and figures. MZ, TA and NJW worked together from the first draft towards the final manuscript. All  
878 authors contributed to the representation of the data and methods in figures and to the discussion of the  
879 implications of the data in the discussion.

880

#### 881 **Competing Interests**

882 The authors have no potential conflicts of interest to declare with regards to this study.

883

#### 884 **Acknowledgements**

885 The authors would like to thank all members of the Clumped Isotope research group of Utrecht University,  
886 most notably Ilja Kocken and dr. Inigo Müller, for their comments and recommendations on a presentation  
887 of the initial results of this study.

888

#### 889 **Financial support**

890 NJW is funded by the European Commission through a Marie Skłodowska Curie Individual Fellowship  
891 (UNBIAS, grant # 843011) and by the Flemish Research Council (FWO) through a Junior Postdoctoral  
892 Fellowship (12ZB220N).

893

#### 894 **References**

895 Bahr, K. D., Jokieli, P. L. and Rodgers, K. S.: Seasonal and annual calcification rates of the Hawaiian reef  
896 coral, *Montipora capitata*, under present and future climate change scenarios, *ICES J Mar Sci*, 74(4),  
897 1083–1091, <https://doi.org/10.1093/icesjms/fsw078>, 2017.

898 Bernasconi, S. M., Müller, I. A., Bergmann, K. D., Breitenbach, S. F., Fernandez, A., Hodell, D. A., Jaggi,  
899 M., Meckler, A. N., Millan, I. and Ziegler, M.: Reducing uncertainties in carbonate clumped isotope  
900 analysis through consistent carbonate-based standardization, *Geochemistry, Geophysics, Geosystems*,  
901 19(9), 2895–2914, 2018.

902 Brand, W. A., Coplen, T. B., Vogl, J., Rosner, M. and Prohaska, T.: Assessment of international reference  
903 materials for isotope-ratio analysis (IUPAC Technical Report), *Pure and Applied Chemistry*, 86(3), 425–  
904 467, <https://doi.org/10.1515/pac-2013-1023>, 2014.

905 Briard, J., Pucéat, E., Vennin, E., Daëron, M., Chavagnac, V., Jaillet, R., Merle, D. and de Rafélis, M.:  
906 Seawater paleotemperature and paleosalinity evolution in neritic environments of the Mediterranean  
907 margin: Insights from isotope analysis of bivalve shells, *Palaeogeography, Palaeoclimatology,*  
908 *Palaeoecology*, 543, 109582, <https://doi.org/10.1016/j.palaeo.2019.109582>, 2020.

909 Caldarescu, D. E., Sadatzki, H., Andersson, C., Schäfer, P., Fortunato, H. and Meckler, A. N.: Clumped  
910 isotope thermometry in bivalve shells: A tool for reconstructing seasonal upwelling, *Geochimica et*  
911 *Cosmochimica Acta*, 294, 174–191, <https://doi.org/10.1016/j.gca.2020.11.019>, 2021.

- 912 Charles, C. D., Hunter, D. E. and Fairbanks, R. G.: Interaction between the ENSO and the Asian monsoon  
913 in a coral record of tropical climate, *Science*, 277(5328), 925–928, 1997.
- 914 Comboul, M., Emile-Geay, J., Evans, M. N., Mirnategui, N., Cobb, K. M. and Thompson, D. M.: A  
915 probabilistic model of chronological errors in layer-counted climate proxies: applications to annually  
916 banded coral archives, *Climate of the Past*, 10(2), 825–841, 2014.
- 917 Compton, T. J., Rijkenberg, M. J. A., Drent, J. and Piersma, T.: Thermal tolerance ranges and climate  
918 variability: A comparison between bivalves from differing climates, *Journal of Experimental Marine  
919 Biology and Ecology*, 352(1), 200–211, <https://doi.org/10.1016/j.jembe.2007.07.010>, 2007.
- 920 Cook, E. R. and Kairiukstis, L. A.: *Methods of dendrochronology: applications in the environmental  
921 sciences*, Springer Science & Business Media., 2013.
- 922 Cramer, B. S., Toggweiler, J. R., Wright, J. D., Katz, M. E. and Miller, K. G.: Ocean overturning since the  
923 Late Cretaceous: Inferences from a new benthic foraminiferal isotope compilation, *Paleoceanography*,  
924 24(4), <https://doi.org/10.1029/2008PA001683>, 2009.
- 925 Crossland, C.: Seasonal variations in the rates of calcification and productivity in the coral *Acropora formosa*  
926 on a high-latitude reef, *Marine Ecology Progress Series*, 15, 135–140,  
927 <https://doi.org/10.3354/meps015135>, 1984.
- 928 van Dam, J. A. and Reichert, G. J.: Oxygen and carbon isotope signatures in late Neogene horse teeth  
929 from Spain and application as temperature and seasonality proxies, *Palaeogeography,  
930 Palaeoclimatology, Palaeoecology*, 274(1–2), 64–81, <https://doi.org/10.1016/j.palaeo.2008.12.022>,  
931 2009.
- 932 Dattalo, P.: *Determining Sample Size: Balancing Power, Precision, and Practicality*, Oxford University  
933 Press, USA., 2008.
- 934 Dayem, K. E., Molnar, P., Battisti, D. S. and Roe, G. H.: Lessons learned from oxygen isotopes in modern  
935 precipitation applied to interpretation of speleothem records of paleoclimate from eastern Asia, *Earth and  
936 Planetary Science Letters*, 295(1–2), 219–230, 2010.
- 937 De Ridder, F., de Brauwere, A., Pintelon, R., Schoukens, J., Dehairs, F., Baeyens, W. and Wilkinson, B.  
938 H.: Comment on: Paleoclimatic inference from stable isotope profiles of accretionary biogenic hardparts—  
939 a quantitative approach to the evaluation of incomplete data, by Wilkinson, B.H., Ivany, L.C., 2002.  
940 *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 185, 95–114, *Palaeogeography, Palaeoclimatology, Palaeoecology*,  
941 248(3–4), 473–476, <https://doi.org/10.1016/j.palaeo.2006.08.004>, 2007.
- 942 De Vleeschouwer, D., Vahlenkamp, M., Crucifix, M. and Pälike, H.: Alternating Southern and Northern  
943 Hemisphere climate response to astronomical forcing during the past 35 my, *Geology*, 45(4), 375–378,  
944 2017.
- 945 de Winter, N. J., Snoeck, C. and Claeys, P.: Seasonal Cyclicity in Trace Elements and Stable Isotopes of  
946 Modern Horse Enamel, *PLoS one*, 11(11), e0166678, 2016.
- 947 de Winter, N., Sinnesael, M., Makarona, C., Vansteenberge, S. and Claeys, P.: Trace element analyses of  
948 carbonates using portable and micro-X-ray fluorescence: Performance and optimization of measurement  
949 parameters and strategies., *Journal of Analytical Atomic Spectrometry*, 32(6), 1211–1223,  
950 <https://doi.org/10.1039/C6JA00361C>, 2017.

- 951 de Winter, N. J., Vellekoop, J., Vorsselmans, R., Golreihan, A., Soete, J., Petersen, S. V., Meyer, K. W.,  
952 Casadio, S., Speijer, R. P. and Claeys, P.: An assessment of latest Cretaceous Pycnodonte vesicularis  
953 (Lamarck, 1806) shells as records for palaeoseasonality: a multi-proxy investigation, *Climate of the Past*,  
954 14(6), 725–749, 2018.
- 955 de Winter, N. J., Müller, I. A., Kocken, I. J., Thibault, N., Ullmann, C. V., Farnsworth, A., Lunt, D. J., Claeys,  
956 P. and Ziegler, M.: First absolute seasonal temperature estimates for greenhouse climate from clumped  
957 isotopes in bivalve shells, *Nature Communications*, in review, <https://doi.org/10.21203/rs.3.rs-39203/v1>,  
958 2020a.
- 959 de Winter, N. J., Ullmann, C. V., Sørensen, A. M., Thibault, N., Goderis, S., Van Malderen, S. J. M., Snoeck,  
960 C., Goolaerts, S., Vanhaecke, F. and Claeys, P.: Shell chemistry of the boreal Campanian bivalve  
961 *Rastellum diluvianum* (Linnaeus, 1767) reveals temperature seasonality, growth rates  
962 and life cycle of an extinct Cretaceous oyster, *Biogeosciences*, 17(11), 2897–2922,  
963 <https://doi.org/10.5194/bg-17-2897-2020>, 2020b.
- 964 de Winter, N. J., Goderis, S., Malderen, S. J. M. V., Sinnesael, M., Vansteenberge, S., Snoeck, C., Belza,  
965 J., Vanhaecke, F. and Claeys, P.: Subdaily-Scale Chemical Variability in a *Torreites Sanchezi* Rudist  
966 Shell: Implications for Rudist Paleobiology and the Cretaceous Day-Night Cycle, *Paleoceanography and*  
967 *Paleoclimatology*, 35(2), e2019PA003723, <https://doi.org/10.1029/2019PA003723>, 2020c.
- 968 de Winter, N. J., Vellekoop, J., Clark, A. J., Stassen, P., Speijer, R. P. and Claeys, P.: The giant marine  
969 gastropod *Campanile giganteum* (Lamarck, 1804) as a high-resolution archive of seasonality in the  
970 Eocene greenhouse world, *Geochemistry, Geophysics, Geosystems*, 21(n/a), e2019GC008794,  
971 <https://doi.org/10.1029/2019GC008794>, 2020d.
- 972 de Winter, N. J.: seasonalclumped: Toolbox for Clumped Isotope Seasonality Reconstructions.  
973 <https://CRAN.R-project.org/package=seasonalclumped>, last access: 4 February 2021, 2021a.
- 974 de Winter, N. J.: ShellChron 0.2.8: A new tool for constructing chronologies in accretionary carbonate  
975 archives from stable oxygen isotope profiles, *Geoscientific Model Development Discussions*, 1–37,  
976 <https://doi.org/10.5194/gmd-2020-401>, 2021b.
- 977 Denton, G. H., Alley, R. B., Comer, G. C. and Broecker, W. S.: The role of seasonality in abrupt climate  
978 change, *Quaternary Science Reviews*, 24(10), 1159–1182,  
979 <https://doi.org/10.1016/j.quascirev.2004.12.002>, 2005.
- 980 Fairbanks, R. G., Evans, M. N., Rubenstone, J. L., Mortlock, R. A., Broad, K., Moore, M. D. and Charles,  
981 C. D.: Evaluating climate indices and their geochemical proxies measured in corals, *Coral Reefs*, 16(1),  
982 S93–S100, <https://doi.org/10.1007/s003380050245>, 1997.
- 983 Fernandez, A., Müller, I. A., Rodríguez-Sanz, L., van Dijk, J., Looser, N. and Bernasconi, S. M.: A  
984 reassessment of the precision of carbonate clumped isotope measurements: implications for calibrations  
985 and paleoclimate reconstructions, *Geochemistry, Geophysics, Geosystems*, 18(12), 4375–4386, 2017.
- 986 Gaspar, M. B., Ferreira, R. and Monteiro, C. C.: Growth and reproductive cycle of *Donax trunculus* L.,  
987 (Mollusca: Bivalvia) off Faro, southern Portugal, *Fisheries Research*, 41(3), 309–316,  
988 [https://doi.org/10.1016/S0165-7836\(99\)00017-X](https://doi.org/10.1016/S0165-7836(99)00017-X), 1999.
- 989 Goodwin, D. H., Schöne, B. R. and Dettman, D. L.: Resolution and fidelity of oxygen isotopes as  
990 paleotemperature proxies in bivalve mollusk shells: models and observations, *Palaios*, 18(2), 110–125,  
991 2003.



- 992 Goodwin, D. H., Paul, P. and Wissink, C. L.: MoGroFunGen: A numerical model for reconstructing intra-  
 993 annual growth rates of bivalve molluscs, *Palaeogeography, Palaeoclimatology, Palaeoecology*, 276(1),  
 994 47–55, <https://doi.org/10.1016/j.palaeo.2009.02.026>, 2009.
- 995 Harwood, A. J. P., Dennis, P. F., Marca, A. D., Pilling, G. M. and Millner, R. S.: The oxygen isotope  
 996 composition of water masses within the North Sea, *Estuarine, Coastal and Shelf Science*, 78(2), 353–  
 997 359, <https://doi.org/10.1016/j.ecss.2007.12.010>, 2008.
- 998 Hendriks, I. E., Basso, L., Deudero, S., Cabanellas-Reboredo, M. and Álvarez, E.: Relative growth rates of  
 999 the noble pen shell *Pinna nobilis* throughout ontogeny around the Balearic Islands (Western  
 1000 Mediterranean, Spain), *Journal of Shellfish Research*, 31(3), 749–756, 2012.
- 1001 Henkes, G. A., Passey, B. H., Grossman, E. L., Shenton, B. J., Yancey, T. E. and Pérez-Huerta, A.:  
 1002 Temperature evolution and the oxygen isotope composition of Phanerozoic oceans from carbonate  
 1003 clumped isotope thermometry, *Earth and Planetary Science Letters*, 490, 40–50,  
 1004 <https://doi.org/10.1016/j.epsl.2018.02.001>, 2018.
- 1005 Hurrell, J. W.: Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation,  
 1006 *Science*, 269(5224), 676–679, 1995.
- 1007 Huybers, P. and Curry, W.: Links between annual, Milankovitch and continuum temperature variability,  
 1008 *Nature*, 441(7091), 329, 2006.
- 1009 Huyghe, D., Lartaud, F., Emmanuel, L., Merle, D. and Renard, M.: Palaeogene climate evolution in the  
 1010 Paris Basin from oxygen stable isotope ( $\delta^{18}\text{O}$ ) compositions of marine molluscs, *Journal of the*  
 1011 *Geological Society*, 172(5), 576–587, 2015.
- 1012 Huyghe, D., de Rafélis, M., Ropert, M., Mouchi, V., Emmanuel, L., Renard, M. and Lartaud, F.: New insights  
 1013 into oyster high-resolution hinge growth patterns, *Marine biology*, 166(4), 48, 2019.
- 1014 IPCC: IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to  
 1015 the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, 1535 pp, Cambridge  
 1016 Univ. Press, Cambridge, UK, and New York., 2013.
- 1017 Ivany, L. C.: Reconstructing paleoseasonality from accretionary skeletal carbonates—challenges and  
 1018 opportunities, *The Paleontological Society Papers*, 18, 133–166, 2012.
- 1019 Jenkyns, H. C.: Geochemistry of oceanic anoxic events, *Geochemistry, Geophysics, Geosystems*, 11(3),  
 1020 <https://doi.org/10.1029/2009GC002788>, 2010.
- 1021 Jones, A. M., Iacumin, P. and Young, E. D.: High-resolution  $\delta^{18}\text{O}$  analysis of tooth enamel phosphate by  
 1022 isotope ratio monitoring gas chromatography mass spectrometry and ultraviolet laser fluorination, , 8,  
 1023 1999.
- 1024 Judd, E. J., Wilkinson, B. H. and Ivany, L. C.: The life and time of clams: Derivation of intra-annual growth  
 1025 rates from high-resolution oxygen isotope profiles, *Palaeogeography, Palaeoclimatology, Palaeoecology*,  
 1026 490, 70–83, 2018.
- 1027 Keating-Bitonti, C. R., Ivany, L. C., Affek, H. P., Douglas, P. and Samson, S. D.: Warm, not super-hot,  
 1028 temperatures in the early Eocene subtropics, *Geology*, 39(8), 771–774,  
 1029 <https://doi.org/10.1130/G32054.1>, 2011.

- 1030 Kele, S., Breitenbach, S. F., Capezzuoli, E., Meckler, A. N., Ziegler, M., Millan, I. M., Kluge, T., Deák, J.,  
1031 Hanselmann, K. and John, C. M.: Temperature dependence of oxygen-and clumped isotope fractionation  
1032 in carbonates: a study of travertines and tufas in the 6–95 C temperature range, *Geochimica et*  
1033 *Cosmochimica Acta*, 168, 172–192, 2015.
- 1034 Kim, S.-T. and O’Neil, J. R.: Equilibrium and nonequilibrium oxygen isotope effects in synthetic carbonates,  
1035 *Geochimica et Cosmochimica Acta*, 61(16), 3461–3475, [https://doi.org/10.1016/S0016-7037\(97\)00169-](https://doi.org/10.1016/S0016-7037(97)00169-5)  
1036 5, 1997.
- 1037 Kocken, I. J., Müller, I. A. and Ziegler, M.: Optimizing the Use of Carbonate Standards to Minimize  
1038 Uncertainties in Clumped Isotope Data, *Geochemistry, Geophysics, Geosystems*, 20(11), 5565–5577,  
1039 <https://doi.org/10.1029/2019GC008545>, 2019.
- 1040 Kohn, M. J.: Comment: tooth enamel mineralization in ungulates: implications for recovering a primary  
1041 isotopic time-series, by BH Passey and TE Cerling (2002), *Geochimica et Cosmochimica Acta*, 68(2),  
1042 403–405, 2004.
- 1043 Lauretano, V., Zachos, J. C. and Lourens, L. J.: Orbitally Paced Carbon and Deep-Sea Temperature  
1044 Changes at the Peak of the Early Eocene Climatic Optimum, *Paleoceanography and Paleoclimatology*,  
1045 33(10), 1050–1065, <https://doi.org/10.1029/2018PA003422>, 2018.
- 1046 Lear, C. H., Bailey, T. R., Pearson, P. N., Coxall, H. K. and Rosenthal, Y.: Cooling and ice growth across  
1047 the Eocene-Oligocene transition, *Geology*, 36(3), 251–254, 2008.
- 1048 LeGrande, A. N. and Schmidt, G. A.: Global gridded data set of the oxygen isotopic composition in  
1049 seawater, *Geophysical research letters*, 33(12), 2006.
- 1050 Lisiecki, L. E. and Raymo, M. E.: A Pliocene-Pleistocene stack of 57 globally distributed benthic  $\delta^{18}O$   
1051 records, *Paleoceanography*, 20(1), <https://doi.org/10.1029/2004PA001071>, 2005.
- 1052 Lourens, L. J., Sluijs, A., Kroon, D., Zachos, J. C., Thomas, E., Röhl, U., Bowles, J. and Raffi, I.:  
1053 Astronomical pacing of late Palaeocene to early Eocene global warming events, *Nature*, 435(7045),  
1054 1083–1087, <https://doi.org/10.1038/nature03814>, 2005.
- 1055 Lourens, L. J., Becker, J., Bintanja, R., Hilgen, F. J., Tuenter, E., Van de Wal, R. S. and Ziegler, M.: Linear  
1056 and non-linear response of late Neogene glacial cycles to obliquity forcing and implications for the  
1057 Milankovitch theory, *Quaternary Science Reviews*, 29(1–2), 352–365, 2010.
- 1058 McArthur, J. M., Howarth, R. J. and Shields, G. A.: Strontium isotope stratigraphy, *The geologic time scale*,  
1059 1, 127–144, 2012.
- 1060 Meckler, A. N., Ziegler, M., Millán, M. I., Breitenbach, S. F. and Bernasconi, S. M.: Long-term performance  
1061 of the Kiel carbonate device with a new correction scheme for clumped isotope measurements, *Rapid*  
1062 *Communications in Mass Spectrometry*, 28(15), 1705–1715, 2014.
- 1063 Merz, B. and Thieken, A. H.: Separating natural and epistemic uncertainty in flood frequency analysis,  
1064 *Journal of Hydrology*, 309(1–4), 114–132, 2005.
- 1065 Meyers, S. R.: Astrochron: An R package for astrochronology, [http://cran.r-](http://cran.r-project.org/package=astrochron)  
1066 [project.org/package=astrochron](http://cran.r-project.org/package=astrochron).  
1067 <http://scholar.google.com/scholar?cluster=14876361610707754388&hl=en&oi=scholar>, last access: 30  
1068 May 2017, 2014.

- 1069 Meyers, S. R.: Cyclostratigraphy and the problem of astrochronologic testing, *Earth-Science Reviews*, 190,  
1070 190–223, <https://doi.org/10.1016/j.earscirev.2018.11.015>, 2019.
- 1071 Miyaji, T., Tanabe, K., Matsushima, Y., Sato, S., Yokoyama, Y. and Matsuzaki, H.: Response of daily and  
1072 annual shell growth patterns of the intertidal bivalve *Phacosoma japonicum* to Holocene coastal climate  
1073 change in Japan, *Palaeogeography, Palaeoclimatology, Palaeoecology*, 286(3), 107–120,  
1074 <https://doi.org/10.1016/j.palaeo.2009.11.032>, 2010.
- 1075 Mook, W. G.: Stable carbon and oxygen isotopes of natural waters in the Netherlands, *Isotope hydrology*,  
1076 1970, 163–190, 1970.
- 1077 Morgan, V. and van Ommen, T. D.: Seasonality in late-Holocene climate from ice-core records, *The  
1078 Holocene*, 7(3), 351–354, <https://doi.org/10.1177/095968369700700312>, 1997.
- 1079 Mosley-Thompson, E., Thompson, L. G., Dai, J., Davis, M. and Lin, P. N.: Climate of the last 500 years:  
1080 High resolution ice core records, *Quaternary Science Reviews*, 12(6), 419–430,  
1081 [https://doi.org/10.1016/S0277-3791\(05\)80006-X](https://doi.org/10.1016/S0277-3791(05)80006-X), 1993.
- 1082 Müller, I. A., Fernandez, A., Radke, J., van Dijk, J., Bowen, D., Schwieters, J. and Bernasconi, S. M.:  
1083 Carbonate clumped isotope analyses with the long-integration dual-inlet (LIDI) workflow: scratching at  
1084 the lower sample weight boundaries: LIDI as key for more precise analyses on much less carbonate  
1085 material, *Rapid Communications in Mass Spectrometry*, 31(12), 1057–1066,  
1086 <https://doi.org/10.1002/rcm.7878>, 2017.
- 1087 Noorbergen, L. J., Abels, H. A., Hilgen, F. J., Robson, B. E., Jong, E. de, Dekkers, M. J., Krijgsman, W.,  
1088 Smit, J., Collinson, M. E. and Kuiper, K. F.: Conceptual models for short-eccentricity-scale climate control  
1089 on peat formation in a lower Palaeocene fluvial system, north-eastern Montana (USA), *Sedimentology*,  
1090 65(3), 775–808, <https://doi.org/10.1111/sed.12405>, 2018.
- 1091 O'Donnell, M. S. and Ignizio, D. A.: Bioclimatic predictors for supporting ecological applications in the  
1092 conterminous United States, *US Geological Survey Data Series*, 691(10), 2012.
- 1093 Passey, B. H. and Cerling, T. E.: Tooth enamel mineralization in ungulates: implications for recovering a  
1094 primary isotopic time-series, *Geochimica et Cosmochimica Acta*, 66(18), 3225–3234, 2002.
- 1095 Petersen, S. V., Tabor, C. R., Lohmann, K. C., Poulsen, C. J., Meyer, K. W., Carpenter, S. J., Erickson, J.  
1096 M., Matsunaga, K. K., Smith, S. Y. and Sheldon, N. D.: Temperature and salinity of the Late Cretaceous  
1097 western interior seaway, *Geology*, 44(11), 903–906, 2016.
- 1098 Philander, S. G. H.: El Nino southern oscillation phenomena, *Nature*, 302(5906), 295–301, 1983.
- 1099 R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical  
1100 Computing, Vienna, Austria. <http://www.R-project.org/>, 2013.
- 1101 Rodríguez-Sanz, L., Bernasconi, S. M., Marino, G., Heslop, D., Müller, I. A., Fernandez, A., Grant, K. M.  
1102 and Rohling, E. J.: Penultimate deglacial warming across the Mediterranean Sea revealed by clumped  
1103 isotopes in foraminifera, *Scientific Reports*, 7(1), 1–11, <https://doi.org/10.1038/s41598-017-16528-6>,  
1104 2017.
- 1105 Rohling, E. J.: Oxygen isotope composition of seawater, *The Encyclopedia of Quaternary Science*.  
1106 Amsterdam: Elsevier, 2, 915–922, 2013.

- 1107 Sano, Y., Kobayashi, S., Shirai, K., Takahata, N., Matsumoto, K., Watanabe, T., Sowa, K. and Iwai, K.:  
 1108 Past daily light cycle recorded in the strontium/calcium ratios of giant clam shells, *Nature*  
 1109 *Communications*, 3, 761, 2012.
- 1110 Sato, S.: Temporal change of life-history traits in fossil bivalves: an example of *Phacosoma japonicum* from  
 1111 the Pleistocene of Japan, *Palaeogeography, Palaeoclimatology, Palaeoecology*, 154(4), 313–323,  
 1112 [https://doi.org/10.1016/S0031-0182\(99\)00106-6](https://doi.org/10.1016/S0031-0182(99)00106-6), 1999.
- 1113 Schmitt, J., Schneider, R., Elsig, J., Leuenberger, D., Lourantou, A., Chappellaz, J., Kohler, P., Joos, F.,  
 1114 Stocker, T. F., Leuenberger, M. and Fischer, H.: Carbon Isotope Constraints on the Deglacial CO<sub>2</sub> Rise  
 1115 from Ice Cores, *Science*, 336(6082), 711–714, <https://doi.org/10.1126/science.1217161>, 2012.
- 1116 Scholz, D. and Hoffmann, D. L.: StalAge—An algorithm designed for construction of speleothem age models,  
 1117 *Quaternary Geochronology*, 6(3–4), 369–382, 2011.
- 1118 Schöne, B. R.: The curse of physiology—challenges and opportunities in the interpretation of geochemical  
 1119 data from mollusk shells, *Geo-Marine Letters*, 28(5–6), 269–285, 2008.
- 1120 Schöne, B. R., Fiebig, J., Pfeiffer, M., Gleß, R., Hickson, J., Johnson, A. L., Dreyer, W. and Oschmann, W.:  
 1121 Climate records from a bivalved Methuselah (*Arctica islandica*, Mollusca; Iceland), *Palaeogeography,*  
 1122 *Palaeoclimatology, Palaeoecology*, 228(1–2), 130–148, 2005.
- 1123 Schöne, B. R., Rodland, D. L., Fiebig, J., Oschmann, W., Goodwin, D., Flessa, K. W. and Dettman, D.:  
 1124 Reliability of multitaxon, multiproxy reconstructions of environmental conditions from accretionary  
 1125 biogenic skeletons, *The Journal of geology*, 114(3), 267–285, 2006.
- 1126 Scourse, J., Richardson, C., Forsythe, G., Harris, I., Heinemeier, J., Fraser, N., Briffa, K. and Jones, P.:  
 1127 First cross-matched floating chronology from the marine fossil record: data from growth lines of the long-  
 1128 lived bivalve mollusc *Arctica islandica*, *The Holocene*, 16(7), 967–974,  
 1129 <https://doi.org/10.1177/0959683606h1987rp>, 2006.
- 1130 Sha, L., Mahata, S., Duan, P., Luz, B., Zhang, P., Baker, J., Zong, B., Ning, Y., Brahim, Y. A., Zhang, H.,  
 1131 Edwards, R. L. and Cheng, H.: A novel application of triple oxygen isotope ratios of speleothems,  
 1132 *Geochimica et Cosmochimica Acta*, 270, 360–378, <https://doi.org/10.1016/j.gca.2019.12.003>, 2020.
- 1133 Shao, D., Mei, Y., Yang, Z., Wang, Y., Yang, W., Gao, Y., Yang, L. and Sun, L.: Holocene ENSO variability  
 1134 in the South China Sea recorded by high-resolution oxygen isotope records from the shells of *Tridacna*  
 1135 *spp.*, *Scientific Reports*, 10(1), 3921, <https://doi.org/10.1038/s41598-020-61013-2>, 2020.
- 1136 Sinnesael, M., De Vleeschouwer, D., Zeeden, C., Batenburg, S. J., Da Silva, A.-C., de Winter, N. J.,  
 1137 Dinarès-Turell, J., Drury, A. J., Gambacorta, G. and Hilgen, F. J.: The Cyclostratigraphy Intercomparison  
 1138 Project (CIP): consistency, merits and pitfalls, *Earth-Science Reviews*, 102965, 2019.
- 1139 Stap, L., Lourens, L. J., Thomas, E., Sluijs, A., Bohaty, S. and Zachos, J. C.: High-resolution deep-sea  
 1140 carbon and oxygen isotope records of Eocene Thermal Maximum 2 and H2, *Geology*, 38(7), 607–610,  
 1141 2010.
- 1142 Steffensen, J. P., Andersen, K. K., Bigler, M., Clausen, H. B., Dahl-Jensen, D., Fischer, H., Goto-Azuma,  
 1143 K., Hansson, M., Johnsen, S. J. and Jouzel, J.: High-resolution Greenland ice core data show abrupt  
 1144 climate change happens in few years, *Science*, 321(5889), 680–684, 2008.

- 1145 Steuber, T., Rauch, M., Masse, J.-P., Graaf, J. and Malkoč, M.: Low-latitude seasonality of Cretaceous  
 1146 temperatures in warm and cold episodes, *Nature*, 437(7063), 1341–1344,  
 1147 <https://doi.org/10.1038/nature04096>, 2005.
- 1148 Surge, D., Lohmann, K. C. and Dettman, D. L.: Controls on isotopic chemistry of the American oyster,  
 1149 *Crassostrea virginica*: implications for growth patterns, *Palaeogeography, Palaeoclimatology,*  
 1150 *Palaeoecology*, 172(3), 283–296, 2001.
- 1151 Titschack, J., Zuschin, M., Spötl, C. and Baal, C.: The giant oyster *Hytotissa hyotis* from the northern Red  
 1152 Sea as a decadal-scale archive for seasonal environmental fluctuations in coral reef habitats, *Coral*  
 1153 *Reefs*, 29(4), 1061–1075, 2010.
- 1154 Treble, P., Shelley, J. M. G. and Chappell, J.: Comparison of high resolution sub-annual records of trace  
 1155 elements in a modern (1911–1992) speleothem with instrumental climate data from southwest Australia,  
 1156 *Earth and Planetary Science Letters*, 216(1), 141–153, [https://doi.org/10.1016/S0012-821X\(03\)00504-1](https://doi.org/10.1016/S0012-821X(03)00504-1),  
 1157 2003.
- 1158 Tsukakoshi, Y.: Sampling variability and uncertainty in total diet studies, *Analyst*, 136(3), 533–539,  
 1159 <https://doi.org/10.1039/C0AN00397B>, 2011.
- 1160 Tudhope, A. W.: Variability in the El Niño-Southern Oscillation Through a Glacial-Interglacial Cycle,  
 1161 *Science*, 291(5508), 1511–1517, <https://doi.org/10.1126/science.1057969>, 2001.
- 1162 Ullmann, C. V., Wiechert, U. and Korte, C.: Oxygen isotope fluctuations in a modern North Sea oyster  
 1163 (*Crassostrea gigas*) compared with annual variations in seawater temperature: Implications for  
 1164 palaeoclimate studies, *Chemical Geology*, 277(1), 160–166, 2010.
- 1165 Van Rampelbergh, M., Verheyden, S., Allan, M., Quinif, Y., Keppens, E. and Claeys, P.: Seasonal variations  
 1166 recorded in cave monitoring results and a 10 year monthly resolved speleothem  $\delta^{18}\text{O}$  and  $\delta^{13}\text{C}$  record  
 1167 from the Han-sur-Lesse cave, Belgium, *Climate of the Past Discussions*, 10, 1821–1856, 2014.
- 1168 Vansteenberge, S., Verheyden, S., Cheng, H., Edwards, R. L., Keppens, E. and Claeys, P.: Paleoclimate  
 1169 in continental northwestern Europe during the Eemian and early Weichselian (125–97 ka): insights from  
 1170 a Belgian speleothem, *Clim. Past*, 12(7), 1445–1458, <https://doi.org/10.5194/cp-12-1445-2016>, 2016.
- 1171 Vansteenberge, S., Winter, N. de, Sinnesael, M., Verheyden, S., Goderis, S., Malderen, S. J. M. V.,  
 1172 Vanhaecke, F. and Claeys, P.: Reconstructing seasonality through stable isotope and trace element  
 1173 analysis of the Proserpine stalagmite, Han-sur-Lesse Cave, Belgium: indications for climate-driven  
 1174 changes during the last 400 years, *Climate of the Past Discussions*, 1–32, [https://doi.org/10.5194/cp-](https://doi.org/10.5194/cp-2019-78)  
 1175 2019-78, 2019.
- 1176 Veizer, J. and Prokoph, A.: Temperatures and oxygen isotopic composition of Phanerozoic oceans, *Earth-*  
 1177 *Science Reviews*, 146, 92–104, <https://doi.org/10.1016/j.earscirev.2015.03.008>, 2015.
- 1178 Vleeschouwer, D. D., Vahlenkamp, M., Crucifix, M. and Pälike, H.: Alternating Southern and Northern  
 1179 Hemisphere climate response to astronomical forcing during the past 35 m.y., *Geology*, 45(4), 375–378,  
 1180 <https://doi.org/10.1130/G38663.1>, 2017.
- 1181 Warter, V. and Müller, W.: Daily growth and tidal rhythms in Miocene and modern giant clams revealed via  
 1182 ultra-high resolution LA-ICPMS analysis—A novel methodological approach towards improved  
 1183 sclerochemistry, *Palaeogeography, Palaeoclimatology, Palaeoecology*, 465, 362–375, 2017.

- 1184 Westerhold, T., Marwan, N., Drury, A. J., Liebrand, D., Agnini, C., Anagnostou, E., Barnet, J. S., Bohaty,  
1185 S. M., De Vleeschouwer, D. and Florindo, F.: An astronomically dated record of Earth's climate and its  
1186 predictability over the last 66 million years, *Science*, 369(6509), 1383–1387, 2020.
- 1187 Wilkinson, B. H. and Ivany, L. C.: Paleoclimatic inference from stable isotope profiles of accretionary  
1188 biogenic hardparts – a quantitative approach to the evaluation of incomplete data, *Palaeogeography,*  
1189 *Palaeoclimatology, Palaeoecology*, 185(1), 95–114, [https://doi.org/10.1016/S0031-0182\(02\)00279-1](https://doi.org/10.1016/S0031-0182(02)00279-1),  
1190 2002.
- 1191 Yan, H., Liu, C., An, Z., Yang, W., Yang, Y., Huang, P., Qiu, S., Zhou, P., Zhao, N. and Fei, H.: Extreme  
1192 weather events recorded by daily to hourly resolution biogeochemical proxies of marine giant clam shells,  
1193 *Proceedings of the National Academy of Sciences*, 2020.