Supplement of

Optimizing sampling strategies in high-resolution paleoclimate records

Niels J. de Winter et al.

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

The copyright of individual parts of the supplement might differ from the article licence.
This information accompanies and completes the manuscript titled “Optimizing sampling strategies in high-resolution paleoclimate records” submitted for publication in Climate of the Past. The document below contains a guide to the Supplementary Data files accompanying this submission, lists several additional result plots not included in the main submission and ends with some supplementary discussion on the applications of this work in future studies. All supplementary files are also stored in the online open-source database Zenodo (www.doi.org/10.5281/zenodo.3899926).
<table>
<thead>
<tr>
<th>File name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1: Overview scenarios</td>
<td>Detailed overview of the specifics of all scenarios tested within this study. Lists variability in SST, growth rate, δ\textsubscript{18}O\textsubscript{w}, sampling density, record, length and age model uncertainty along with a short description of what type of natural variability each case is meant to represent/test. An abridged version of this table is included in the main text as Table 1.</td>
</tr>
<tr>
<td>S2: Texel hourly SST/SSS</td>
<td>Hourly SST and SSS measurements made between 2001 and 2016 at the NIOZ measuring station in the harbor on the south side of Texel island (NW the Netherlands, see case 30, Fig. 3 in main text), obtained with help from Sonja van Leeuwen and Eric Wagemaakers (pers. comm.; see Fig. S1). This dataset was used as a basis for defining variability on the seasonal vs. weather scale, with weather-scale variability being added to the virtual cases (cases 1-29) to more closely approximate natural circumstances. δ\textsubscript{18}O\textsubscript{w} values are obtained from SSS measurements using the mass balance presented in the main text.</td>
</tr>
<tr>
<td>S3: Normality test (R)</td>
<td>Script for testing the normality of “weather” residuals of SST and SSS after subtracting the seasonal cycle through a Kolmogorov-Smirnov test (see Fig. S1).</td>
</tr>
<tr>
<td>S4: Data overview</td>
<td>Overview of true and reconstructed SST and δ\textsubscript{18}O\textsubscript{w} data for all reconstruction approaches. Includes calculations of accuracy and precision for all cases, at every sampling density and using all reconstruction approaches. Also includes separate sheets in which the effect of a certain parameter (e.g. sampling density) is isolated, which serve as basis for Figures 8-11 in the main manuscript.</td>
</tr>
<tr>
<td>S5: Generate virtual data (R)</td>
<td>Script used to generate virtual δ\textsubscript{18}O\textsubscript{c} and Δ\textsubscript{47} data from SST and SSS curves of all virtual cases. The functions used to generate these “virtual carbonate archives” were compiled in the documented R package seasonalclumped which is available on the open-source online repository CRAN (de Winter, 2021).</td>
</tr>
<tr>
<td>S6: Virtual data overview</td>
<td>Overview of virtual data generated for all cases (1-33) using the script in S5. All virtual datasets were also included as internal datasets in the documented R package seasonalclumped which is available on the open-source online repository CRAN (de Winter, 2021).</td>
</tr>
<tr>
<td>S7: Combined data processing routine (R)</td>
<td>Script listing the formulae used to calculate monthly SST and δ\textsubscript{18}O\textsubscript{w} reconstructions from the virtual data sets provided in S6. The functions used to do these reconstructions were compiled in the documented R package seasonalclumped which is available on the open-source online repository CRAN (de Winter, 2021).</td>
</tr>
<tr>
<td>S8: Raw data results and figures</td>
<td>Large folder containing all raw exported datasheets from the script in S7, listing all in between steps of the SST and δ\textsubscript{18}O\textsubscript{w} reconstructions as well as figures comparing the results of all these reconstructions on all cases for every sampling density.</td>
</tr>
</tbody>
</table>
**S9: Accuracy and precision table**

Color-coded table showing the differences in accuracy and precision of mean annual and seasonal SST and δ¹⁸Ow reconstructions on all cases using the four different reconstruction methods (basis for Figures 6 and 7; see also Figures XXX and XXX).

**S10: Optimal sample size results**

Overview of the median optimal sample sizes required for successful separation of summer and winter sample groups for seasonality reconstructions in all cases through the optimization approach.

---

**Supplementary Figures**

Below we list several supplementary figures illustrating additional aspects of the results of this study which were not shown in the figures in the main text.
Figure S1: Seasonality and “weather-scale” variability in measured SST and SSS and modelled $\delta^{18}O_w$ values from the Texel locality (case 30; see S2). With A showing hourly $\delta^{18}O_w$ variability (in blue) over time (years 2001-2016) modelled from hourly SSS measurements (B; in purple) with the
black line showing a sinusoidal fit through the data to represent the seasonal component. C shows the same for hourly SST measurements (in red). The histograms and Q-Q plots in D-G show the distribution of the residuals in SSS (D and E) and SST (F and G) obtained by subtracting the seasonality (black lines in A-C) from the SST and SSS records. Note that the Kolmogorov-Smirnov test (E and G) demonstrates that both SST and SSS residuals ("weather-scale" variability) are not normally distributed (P << 0.05).
Fig. S2: Comparison of the same reconstructions using the 4 different approaches on case 1 using (A) an even sampling interval of 2 mm (an integer fraction of the 10 mm/yr growth rate) or (B) and uneven sampling interval of 1.55 mm (not an integer fraction of the growth rate). Note how even sampling intervals cause some months to be oversampled while some other months never contain samples. Consequently, no reconstruction can be done for these months and the resulting seasonal and mean annual reconstructions run the risk of being biased. On the contrary, uneven sampling intervals (B) result in different months being sampled in different growth years and yield equal sample distributions across the months, leading to more accurate mean annual and seasonal SST and $\delta^{18}O_w$ reconstructions from monthly data. Bars on the right-hand side show mean annual and seasonality SST reconstructions in both cases.
**Fig. S3**: Overview of precision (one standard deviation) of reconstructions of mean annual $\delta^{18}O_w$ ($\delta^{18}O_{sw}$) (A), seasonal range in $\delta^{18}O_w$ ($\delta^{18}O_{sw}$) (B), mean annual SST (C) and seasonal range in SST (D), with higher values indicating lower precision (higher precision errors) based on average sampling resolution (sampling interval of 0.45 mm). The horizontal axis displays the different cases, color coded by their difference from the control case (case 1; see legend on the left-hand side). Colored lines indicate the different data treatment approaches. Box-whisker plots to the right show medians and distributions of precision on all cases using different reconstruction approaches (outliers are identified as black dots based on 2x interquartile distance).
Fig. S4: Overview of accuracy (absolute difference between reconstruction and “true” value) of reconstructions of mean annual δ¹⁸Oₘᵢₜ (A), seasonal range in δ¹⁸Oₘᵢₜ (B), mean annual SST (C) and seasonal range in SST (D), with higher values indicating lower accuracy (higher offset with “true” value) based on average sampling resolution (sampling interval of 0.45 mm). The horizontal axis displays the different cases, color coded by their difference from the control case (case 1; see legend on the left-hand side). Colored lines indicate the different data treatment approaches. Box-whisker plots to the right show medians and distributions of accuracy on all cases using different reconstruction approaches (outliers are identified as black dots based on 2x interquartile distance).
**Fig. S5**: Comparison of the precision (MAT: A and SST seasonality: B) and accuracy (MAT: C and SST seasonality: D) of the clumped isotope reconstruction approaches (smoothing, binning and optimization) with the δ₁⁸O_c-based reconstructions (which do not require sample combination). In all figures, higher values (directed upwards) represent less precise/accurate results, with “0” representing equal accurate/precise results as the δ₁⁸O_c-based reconstructions. Only fully artificial cases are shown and case 18 is omitted.
Supplementary discussion: Implications for other sample size problems

While the discussion above focuses on optimizing approaches for combining samples for clumped isotope analyses in paleoseasonality reconstructions, the problem of combining samples to lower uncertainty and isolate variation in datasets is very common (e.g. Zhang et al., 2004; Merz and Thieken, 2005; Tsukakoshi, 2011). Therefore, the approaches outlined and tested in this study have applications beyond paleoseasonality reconstructions. Below, we briefly highlight four examples of problems that could benefit from applying similar approaches for lowering the uncertainty of estimates of target variables or reducing the number of analyses required to meet analytical requirements.

Tooth bioapatite

Enamel from vertebrate teeth constitute a useful archive for paleoenvironmental and paleoecological change in the terrestrial realm, complementing the carbonate records discussed in this work (e.g. Luz and Kolodny, 1985; Fricke et al., 1996; Balasse, 2002; Van Dam and Reichart, 2009; de Winter et al., 2016). However, the tooth bioapatite archive suffers from similar limitations of sample size and resolution as carbonate archives when it comes to reconstructing high-resolution variability (see discussion in Passey and Cerling, 2002 and Kohn, 2004). Oxygen and carbon isotopes of carbonate and phosphate in tooth enamel contain valuable information about the animal's life cycle and environment (e.g. Fricke et al., 1996; Balasse et al., 2002; Van Dam and Reichart, 2009). However, structurally bound carbonate constitutes a mere 2-5% of tooth enamel (LeGeros et al., 1986), and enamel samples need to be pretreated to remove labile components, so analyses of $\delta^{18}O$ in these archives require comparatively large sample sizes (0.5-1 mg; Fricke et al., 1998; Balasse, 2002; Pellegrini and Snoeck, 2016). Phosphate-bound $\delta^{18}O$ is less susceptible to diagenesis, but requires a more complicated procedure to analyze, resulting in similar sample size limitations (Joachimski et al., 2004; Lecuyer et al., 2007). Most applications of isotope profiles from teeth rely on precise determination of both the phase and amplitude of the seasonal cycle, and therefore suffer from the same complications as isotope records in carbonate archives (e.g. Balasse et al., 2002; Straight et al., 2004). The binning and optimization approaches discussed here could help reduce uncertainty and provide a basis for better comparison of seasonal profiles in tooth enamel.

Cyclostratigraphy
Within the field of cyclostratigraphy, a multitude of stratigraphical approaches have been developed for signal processing, with the aim to use regular orbital cycles expressed in stratigraphic time series as tools for dating rock sequences (e.g. Paillard et al., 1996; Meyers, 2014; Sinnesael et al., 2016). However, the focus on timing has caused many methods for extracting the climatic impact of these orbital cycles from stratigraphic records (e.g. bandpass filtering; Hilgen, 1991) to remain qualitative. This is unfortunate, because the magnitude of the effect of this cyclicity on climate and environmental change is of major interest in paleoclimatology studies (e.g. Berger, 1992; Shackleton, 2000; Zachos et al., 2001; Lourens et al., 2005; De Vleeschouwer et al., 2017a). The problem of quantitatively extracting the impact of orbital cycles is very similar to the problem of paleoseasonality reconstructions central to this study, and the same approaches can therefore be used in the orbital time domain. The time binning approach is probably most robust for this purpose, since cyclostratigraphic records are often longer (record length >> period of the cycle) and sampling resolutions (samples/cycle) are often lower than in seasonal records (see above; e.g. De Vleeschouwer et al., 2017b). Quantitative analyses of the contribution of orbital cyclicity to rhythmic changes in paleoclimate can help separate variability in records caused by external forcing from autogenic behavior or (positive or negative) feedback of the climate system itself (Lourens et al., 2010; Noorbergen et al., 2017; Nohl et al., 2018).

**Strontium isotope dating**

Another type of analysis that could benefit from smart combination of measurement results is strontium isotope dating. The strontium isotope composition ($^{87}$Sr/$^{86}$Sr) of the ocean has evolved over time, and the isotopic composition of marine carbonates can therefore be used to estimate the age of the sample by comparing it with a composite strontium isotope curve (Elderfield, 1986; McArthur et al., 2012). In time intervals where the global marine strontium isotope curve is steep, strontium isotope dating ranks among the most precise methods for absolute dating in stratigraphy (Wagreich et al., 2012). However, accurate dating based on the strontium isotope curve requires propagation of errors on the composite curve and the sample. Doing so results in asymmetric errors due to the non-linear character of the strontium isotope curve, which require complex error propagation (see Barlow, 2003; 2004; Wan et al., 2019). The state-of-the-art uncertainty of individual strontium isotope analyses ranges between 2-10 ppm (1 standard deviation;
Yobregat et al., 2017), which translates to an age uncertainty of 100-200 kyr, (1 standard deviation) depending strongly on the slope of the global strontium isotope curve at the time interval under study. Combining multiple strontium isotope analyses from the same stratigraphic unit can reduce the uncertainty on these composite ages (Korte and Ullmann, 2016; de Winter et al., 2020b), allowing the dating method to be combined with cyclostratigraphy to produce for orbital scale age models (see above). In stratigraphy studies that use this dating method, the need arises to compromise between the resolution of the age model and the precision and accuracy of dating, analogous to the tradeoff that occurs when combining Δ47 analyses for paleoseasonality reconstructions outlined in this study. In this case, the smoothing approach with a dynamic moving window discussed in this study is likely the best candidate for combining data to improve these age models. Such an approach can be seen as a more flexible adaptation of the Δ47-based approach for SST reconstruction outlined in Rodríguez-Sanz et al. (2017) that provides the flexibility to adapt the sample window depending on the available data and the slope of the global strontium curve. At the same time, the shape of the global composite strontium isotope curve itself can be refined by using a similar protocol on well-dated samples. The approaches discussed in this study are more adaptable to changes in sampling density over time and can in theory achieve higher precision than the LOWESS fit approach currently employed for constructing the global composite (McArthur et al., 2012). Similarly, techniques for compromising between sampling resolution and accuracy and precision can be applied to improve other dating methods based on matching curves such as radiocarbon dating (Ramsay and Lee, 2013), carbon isotope stratigraphy (Salzman and Thomas, 2012) and dendrochronology (Cook and Kairiukstis, 2013).

Sub-seasonal variability

Ultra-high-resolution records from fast-growing archives (e.g. mollusks) are an emerging phenomenon in the field of high-resolution paleoclimatology (e.g. Sano et al., 2012; Warter and Müller, 2017, de Winter et al., 2020a). The emergence of such records allows new information to be obtained about the daily cycle (Warter et al., 2018; de Winter et al., 2020a) and extreme weather events (Yan et al., 2020) in the past, potentially bridging the gap between weather and climate reconstructions. The sampling resolution required to resolve variability at such a fine temporal scale warrants an even more careful consideration of the
tradeoff between sample size, sampling resolution and analytical uncertainty than the paleoseasonality examples considered here. If quantitative estimates of insolation, temperature and the frequency of extreme weather events are to be reconstructed from these novel records, a compromise will need to be found between analytical uncertainty and the temporal resolution of measurements (Sano et al., 2012; de Winter et al., 2020a; Yan et al., 2020). Applying the temporal (e.g. hourly) binning method (binning) discussed here on long, (sub-)daily resolved records could yield more accurate and precise records of ultra-high-resolution variability, given its reliability in extracting accurate cycle amplitude (e.g. seasonality) from long, less densely sampled records (see discussion in main text). Fast-growing bivalve and gastropod shells have already been marked as promising archives for such variability, while other fast-growing archives such as Acropora corals remain to be explored (Bak et al., 2009; Strauss et al., 2014; de Winter et al., 2020c). It must be noted that models for the timing of carbonate deposition in accretionary carbonate archives at the sub-daily scale are highly uncertain and that this may complicate the use of the binning approach (see discussion in main text), in which case optimization may be more appropriate.

**Event stratigraphy**

Accurate and precise temperature reconstructions of short-lived (10-100 kyr) episodes of climate change present a problem comparable to resolving seasonality in paleoclimate archives. Examples of such events include the Mesozoic ocean anoxic events (Hesselbo et al., 2000; Jenkyns, 2010), early Paleogene hyperthermals (Stap et al., 2010; Lauretano et al., 2015, 2018) and stepwise climate perturbations such as the Eocene-Oligocene transition (Dupont-Nivet et al., 2007; Lear et al., 2008) studied in deep-sea records. Currently, reconstructions of temperature variability in the deep-sea during such events are based on benthic foraminiferal δ¹⁸Oc (e.g. Erbacher et al., 2001; Lui et al., 2009; Stap et al., 2010; Lauretano et al., 2015, 2018), but may not be reliable due to assumptions made on δ¹⁸Osw. Deep-sea sedimentary environments are generally characterized by low sedimentation rates (~1 cm/kyr) as well as low abundance and small size of microfossils (e.g. foraminifera) which serve as archives of past marine conditions (e.g. Stap et al., 2010; Jennions et al., 2015). This limits the number of aliquots that can be obtained for Δ⁴⁷ and other analyses through these climate events. In these studies, a smoothing approach would probably underestimate the ‘true’ amplitude of temperature or geochemical change. With sufficient record length and
perhaps by combining multiple events, **binning** or **optimization** based on proxy data would be the most accurate and precise approach to resolve transient temperature change in the deep-sea during the geological past.

REFERENCES


De Vleeschouwer, D., Da Silva, A.-C., Sinnesael, M., Chen, D., Day, J. E., Whalen, M. T., Guo, Z. and Claeyts, P.: Timing and pacing of the Late Devonian mass extinction event regulated by eccentricity and obliquity, Nature Communications, 8(1), 1–11, [https://doi.org/10.1038/s41467-017-02407-1](https://doi.org/10.1038/s41467-017-02407-1), 2017b.


