Obliquity Feature Extraction for Fossil Data Analysis: The Stickleback Fish Case

Rolf Ergon

rolf.ergon@usn.no

University of South-Eastern Norway

February 1, 2026

Abstract

A moving average smoothing method for extraction of cycles in time series data is described, with focus on obliquity cycles and fossil data. The proposed method works satisfactory when applied on stickleback fish data from around 10 million years ago. The extraction of obliquity cycles will however depend on the dynamics of the time series, and the method is thus not universally applicable. It may however be possible to adapt the size of the moving window to problems under study.

Keywords: adaptive peak tracking, feature extraction, fossil record, moving average smoothing, stickleback fish

1 Introduction

Obliquity refers to the angle of tilt of a planet's rotational axis relative to its orbital plane, and on Earth this angle varies a few degrees with a period of 41,000 years (Wikipedia contributors, 2025). Information on obliquity cycles is utilized in climate and oceanographic research (Tian et

al., 2013; Zeebe and Lourens, 2022), but until recently they are apparently not referred to in fossil data studies. In Ergon (2025b), however, I showed that the evolution of a trait in a stickleback fish lineage for around 10 million years ago was partly driven by obliquity cycles. This is an example of adaptative peak tracking, as described in Ergon (2025a), with oxygen isotope $\partial^{18}O(t)$ values used as environmental driver. Since these cycles affect marine life, they have left traces in the $\partial^{18}O(t)$ values over time, as found in deep sea drilling projects (Westerhold et al., 2020). The underlying connection between deep sea drilling data and the evolution of stickleback fish traits as described in Bell et al. (1985) is that the obliquity cycles also affected these fishes.

The extraction of obliquity cycles information from $\partial^{18}O(t)$ data is described in Appendix B in Ergon (2025b), and the purpose of the present article is to give a more thorough explanation. In Appendix B in Ergon (2025b), the feature extraction method was also illustrated by power spectrum analysis, which unfortunately had some serious errors. This is corrected here. As will be shown, the power spectrum of the $\partial^{18}O(t)$ data gives in fact the fundamental explanation of the stickleback fish results in Ergon (2025b).

The theoretical background for the feature extraction method is given in Section 2, while the detailed explanation is given in Section 3. A summary with discussion follows in Section 4. MATLAB code and oxygen isotope data are archived at bioRxiv.

2 Theory and Methods

2.1 Extraction of obliquity information

The obliquity cycles have a period of 0.041 million years, and thus a frequency $f = 24.4$ periods per million year. For extraction of obliquity information, the $\partial^{18}O(t)$ time series as found in Westerhold et al. (2020) can be modeled as

$$u(t) = Asin(2\pi ft + \varphi) + u_s(t) + e(t), \tag{1}$$

where $u_s(t)$ has a structure that is important for prediction of the mean traits under study, while $e(t)$ is unimportant high frequence noise. Also assume a sampling frequency $f_s \gg f$, centered moving average smoothing with window sample size n, and n large enough to make the smoothed high frequency noise $e_{smooth}(t) \approx 0$.

For an integer number of periods in the moving average time window, the average of the sinusoidal component in Eq. (1) is zero. For a non-integer number of periods we thus find $\sum_{\tau=t-0.5n/f_s}^{\tau=t+0.5n/f_s} sin(2\pi f\tau + \varphi) = \sum_{\tau=0}^{n_p/f_s} sin(2\pi f\tau + \varphi)$, where $0 < n_p < f_s/f$ (where f_s/f is the number of samples per period). From Eq. (1) we thus find

$$u_{smooth}(t) = \frac{1}{n} A \sum_{\tau=o}^{n_p/f_s} sin(2\pi f\tau + \varphi) + \frac{1}{n} \sum_{\tau=t-0.5n/f_s}^{\tau=t+0.5n/f_s} u_s(\tau). \tag{2}$$

For a non-integer number of periods in the moving average time window, the sinusoidal component will thus appear in $u_{smooth}(t)$ with the frequency f, but with an amplitude depending on n and n_p according to Eq. (2), while the structured information will be smoothed. The sampling window must thus be wide enough to filter out the noise component $e(t)$, but small enough to retain some scaled information on both the sinusoidal component and the features in the structured information in $u_s(t)$. The first term in Eq. (2) will have local maximum values for time window sizes $n = (i + 0.5)/f$, where i is an integer number, but with decreasing values for increasing window sizes, and the optimal value of i for best predictions can be found by tuning.

2.2 Other methods

Weighted least squares estimation and Akaike information criterion (AIC) computations will be used as described in Ergon (2025b).

3 The stickleback fish case used as example

Data for the stickleback fish case in Bell et al. (1985) are summarized in Ergon (2025b), while $\partial^{18}O(t)$ from 9,000 to 11,200 million years ago are found in Westerhold et al. (2020), Table S34, column marked *ISOBENbinned_d18Ointerp*. The zero time point is not given in Bell et al. (1985), but since $t = 0$ in the data is approximately 10 million years ago (personal information) it was possible to align the data for optimal predictions, and the best tracking result was obtained when $t = 0$ in the data corresponds to $t_0 = -9.994$ million years with present time as zero point.

A segment of the true $\partial^{18}O(t)$ time series is shown in Fig. 1, upper panel, together with a moving mean time series $\partial^{18}O_{mean}(t)$ using a centered moving average time window of 0.080 million years. Fig. 1, lower panel, shows simulated data

$$u_{sim}(t) = \partial^{18}O_{mean}(t) + 0.15\sin\left(\frac{2\pi t}{0.041} - \pi/2\right) + 0.08e(t), \tag{3}$$

where $e(t)$ are normally distributed random numbers with variance one. The phase shift $\varphi = -\pi/2$ was chosen for alignment of the simulated cycles with the true cycles (see Fig. 3 below).

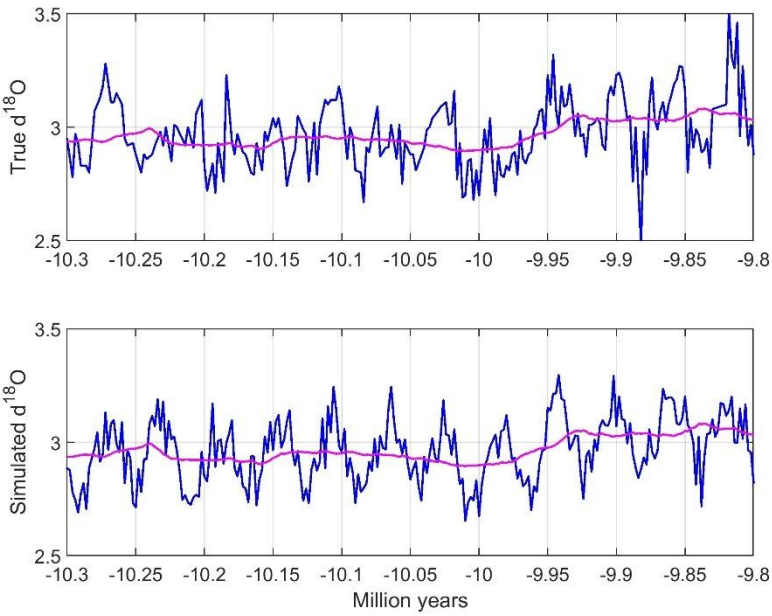

Figure 1. Upper panel shows true and moving mean $\partial^{18}O(t)$ data (blue and red lines), while lower panel shows simulated data according to Eq. (3).

It is not easy to see from Fig. 1 that the true $\partial^{18}O(t)$ signal includes obliquity cycles. This is, however, found by means of spectral analysis using the function *periodogram* in MATLAB, and for that purpose variations around moving mean values as shown by red lines in Fig. 1 were used. Note, however, that the spectral analysis made use of all 1101 samples in the raw data, which made Fourier analysis based on 1024 samples possible. Fig. 2 shows that the true $\partial^{18}O(t)$ signal has a significant spectral density peak with a frequency of 24.7 periods per million years, corresponding to the obliquity cycles, while the simulated signal has a narrow single peak at 24.4 periods per million years.

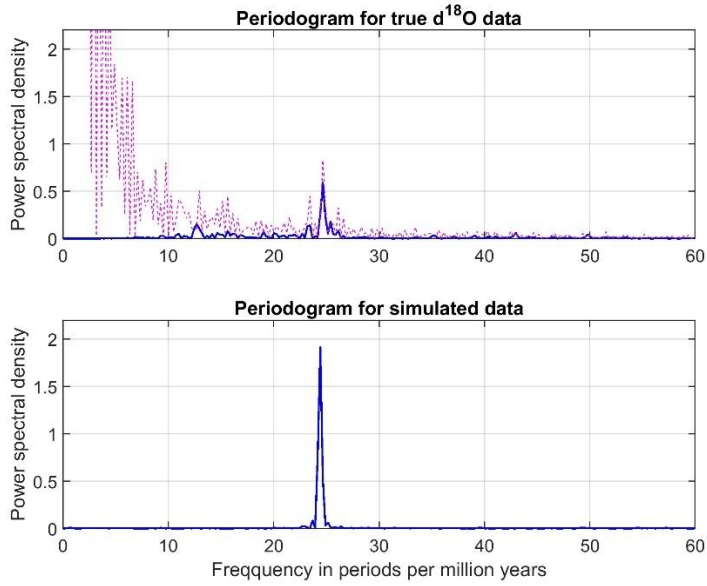

Figure 2. Periodograms for true and simulated $\partial^{18}O(t)$ data after subtraction of moving mean values (blue lines). Dashed red lines show results for the raw data, without subtraction of the moving mean values.

The period of the obliquity cycles is 0.041 million years, and the optimal moving average time window size for stickleback fish predictions and according to $n = (i + 0.5)/f$ was found to be 4.5 periods. Since the $\partial^{18}O(t)$ data has 500 samples per million years, the time window size

was thus chosen to be 0.184 million years. The smoothed true and simulated data are shown in Fig. 3, upper panel. Note that the obliquity cycles are clearly visible only in some time periods, and that the stickleback fish data falls within such a period around 10 million years ago. As shown in Fig. 3, lower panel, a window size of $3.5 \cdot 0.041 = 0.140$ million years could possibly be a good choice for other cases. Window size $2.5 \cdot 0.041 = 0.103$ is in fact a useful choice in the stickleback fish case ($AIC_c = -117$) , but as shown below $4.5 \cdot 0.041 = 0.184$ isa better choice ($AIC_c = -128$).

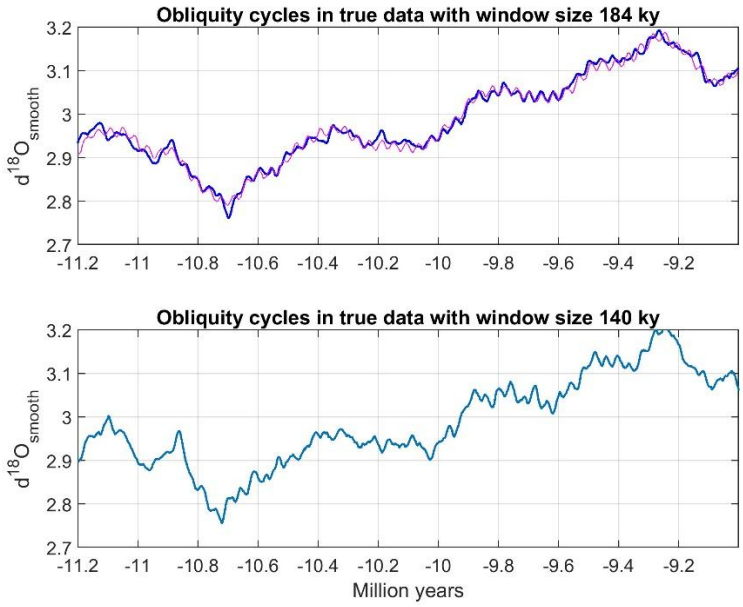

Figure 3. Extracted obliquity cycles using moving average smoothing with two different window sizes (blue lines). Upper panel also shows simulated cycles (red line).

Results of the stickleback fish adaptive peak estimation and mean trait predictions as discussed in more detail in Ergon (2025b) are shown in Figures 4 and 5. Repeated simulations with different realizations of $e(t)$ in Eq. 3 show that the Akaike weights are approximately the same for predictions based on true and simulated $\partial^{18}O(t)$ data. Predictions based on true data and moving window size 0.184 million years resulted in $AIC_c = -128$, which should be compared with $AIC_c = -131$ as found with window size 0.192 (Ergon, 2025b). With given mean

trait data it is thus possible to find improved predictions by tuning, using $n = (i + 0.5)/f$ as a starting point.

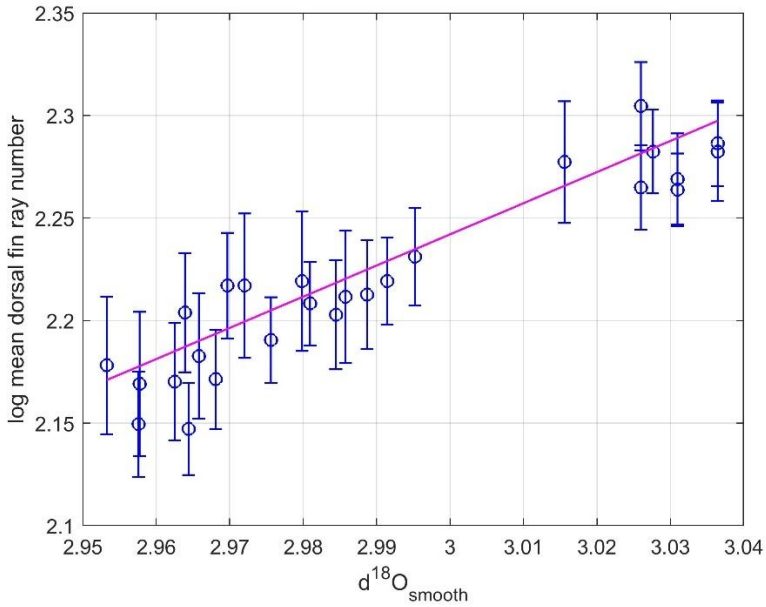

Figure 4. Mean trait values $\bar{y}(i)$ with standard error bars versus environmental driver values $u(i) = \partial^{18}O_{smooth}(i)$, and estimated linear adaptive peak function (red line).

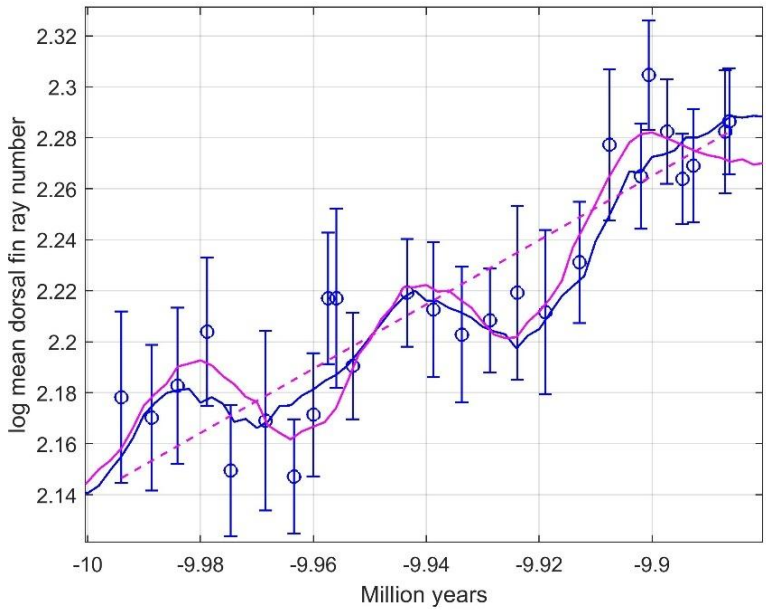

Figure 5. Predictions of stickleback fish mean trait by use of true data (blue line), and by use of typical simulated data (red line). The moving window size was 0.184 million years. Dashed red line shows weighted least squares predictions.

4 Summary and discussion

The feature extraction method used in Ergon (2025b) is described in some more detail, and it is also clarified by use of simulations. As apparent from Fig. 1, it is not obvious that the $\partial^{18}O(t)$ time series obtained from deep sea drilling samples contains obliquity cycles. This is, however, shown in the spectral analysis results in Fig. 2, with peaks corresponding to periods of 0.041 million years, although the true signal also has a clearly visible but unknown frequency component corresponding to periods of 0.077 million years. Centered moving average smoothing may extract the obliquity cycles for the purpose of mean trait predictions, but as shown in Fig. 3 this is not a universally reliable method. As also shown in Fig. 3, however, the size of the moving average window could possibly be adapted to prediction problems under study.

From Eq. (2) follows that the periodic and structured information in the smoothed $\partial^{18}O(t)$ data are scaled in the same way by the size of the smoothing window. This scaling is accounted for when the adaptive peak function is found, as shown in Fig. 4.

As shown in Fig. 5, the proposed feature extraction method is quite useful in the stickleback fish case, as also discussed in Ergon (2025b). The $\partial^{18}O(t)$ signal obtained through deep sea drilling can thus be used as if it was a temperature measurement signal at the field site in what is now Nevada, scaled in an appropriate way and presumably during a critical period of the year.

Acknowledgements

I thank Mike Bell for information regarding the time span of the stickleback fish data and for the observation of possible obliquity cycles. I also thank University of South-Eastern Norway for support.

Competing interests

There are no competing interests.

Data Availability Statement

MATLAB code and oxygen isotope data are archived at bioRxiv, https://doi.org/???

References

Ergon, R. (2025a). Adaptive Peak Tracking as Explanation of Sparse Fossil Data Across Fluctuating Ancient Environments. Ecology and Evolution. First published: 03 July 2025 https://doi.org/10.1002/ece3.71705

Ergon, R. (2025b). Examples of Adaptive Peak Tracking as Found in the Fossil Record, Including Obliquity Cycles Tracking. Modeling, Identification and Control, Vol. 46, No. 4, 2025, pp. 147–162, ISSN 1890–1328

Tian, J., Yang, M., Lyle, M.W., Wilkens, R., and Shackford, J.K., (2013). Obliquity and long eccentricity pacing of the Middle Miocene climate transition. Geochemistry, Geophysics, Geosystems, doi:10.1002/ggge.20108

Westerhold et al. (2020). An astronomically dated record of Earth's climate and its predictability over the last 66 million years. Science 369, 1383–1387. DOI: 10.1126/science.aba6853

Wikipedia contributors. Milankovitch cycles.

https://en.wikipedia.org/wiki/Milankovitch_ cycles, 2025.
[Online; accessed January -2026].

Zeebe, R.E, and Lourens, L.J. (2022). A Deep-Time Dating Tool for Paleo-Applications Utilizing Obliquity and Precession Cycles: The Role of Dynamical Ellipticity and Tidal Dissipation. Paleoceanography and Paleoclimatology, doi: 10.1029/2021PA004349