A Decomposition Approach to Cyclostratigraphic Signal Processing

Sébastien Wouters[1,2], Michel Crucifix[3], Matthias Sinnesael[4], Anne-Christine Da Silva[1], Christian Zeeden[5], Miroslav Zivanovic[6], Frédéric Boulvain[1], Xavier Devleeschouwer[2]

1 Sedimentary Petrology, Liege University, Allée du Six Août 12, 4000 Liège, Belgium
2 O.D. Earth and History of Life, Royal Belgian Institute of Natural Sciences, 29 rue Vautier, 1000 Brussels, Belgium
3 Earth and Life Institute, UCLouvain, Place Louis Pasteur 3, Louvain-La-Neuve, 1348, Belgium
4 Department of Earth Sciences, Mountjoy Site, Durham University, South Road, Durham DH1 3LE, UK
5 LIAG – Leibniz Institute for Applied Geophysics, Geozentrum Hannover, Stilleweg 2, 30655, Hannover, Germany
6 Electrical Engineering Department, Public University of Navarra, Campus Arrosadía s/n, 31006 Pamplona, Spain

Correspondence to: Sébastien Wouters (wouterseb@gmail.com)

Abstract

Sedimentary rocks can record signals produced by highly complex processes. These signals are generated by a progressive deposition of sediments which can be affected, mainly through the climate system, by regular astronomical cycles (i.e. Milankovitch cycles), and by irregular oscillations like the El Niño-Southern Oscillation. Also, usually through biological, chemical and/or physical post-depositional processes, the sedimentary records can be affected by pattern-creating heterogeneous processes. The noise in the signals further complicates the records, and the deposition rate (or sedimentation rate) can fluctuate, which greatly reduces the effectiveness of the classical stationary time-series analysis methods commonly used in cyclostratigraphy (i.e. the study of the cycles found in the sedimentary records).

Faced with this multiplicity of processes, a common approach used in cyclostratigraphy is to reduce each signal to more manageable sub-signals, either over a given range of frequencies (e.g., by filtering), or by considering a continuum of constant frequencies (e.g., using transforms). This makes it possible to focus on the features of interest, commonly astronomical cycles. However, working with sub-signals is not trivial. Firstly, sub-signals have a certain amount of cross-cancellation when they are summed back to reconstruct the initial signal. This means that in filters and in transforms, wiggles that are not present in the initial signal can appear in the sub-signals. Secondly, the sub-signals considered often cannot be summed to reconstruct the initial signal: this means that there are processes affecting the signal which remain unstudied.

It is possible to take cross-cancellation into account and to consider the entire content of a signal by dividing the signal into a decomposition: a set of sub-signals that can be added back together to reconstruct the original signal. We discuss here how to reframe commonly used time-series analysis techniques in the context of decomposition, how they are affected by cross-cancellation, and how adequate they are for comprehending the whole signals. We also show that decomposition can be carried out by non-stationary time-series methods, which can minimise cross-cancellation, and have now reached sufficient maturity to tackle sedimentary records signals.

We present novel tools to adapt non-stationary decomposition for cyclostratigraphic purposes, based on the concepts of Empirical Mode Decomposition (EMD) and Instantaneous Frequency (IF), mainly: (1) a fast Ensemble Empirical Mode Decomposition (EEMD) algorithm, (2) quality metrics for decomposition, and (3) plots to visualise instantaneous frequency, amplitude and frequency ratio.

We illustrate the use of these tools by applying them on a greyscale signal from the site 926 of the Ocean Drilling Program, at Ceara Rise (western equatorial Atlantic), especially to identify and characterise the expression of astronomical cycles. The main goal is to show that by minimising cross-cancellation, we can apply in real signals what we call the wiggle-in-signal approach: making the sub-signals in the decomposition more representative of the expression, wiggle by wiggle, of all the processes affecting the signal (e.g., astronomical cycles). We finally argue that decomposition could be used as a practical standard output for time-series analysis interpretation of cyclostratigraphic signals.

Keywords Cyclostratigraphy, Astrochronology, Signal Decomposition, Ensemble Empirical Mode Decomposition (EEMD)

1 Introduction

Sedimentary cycles are repetitions of patterns in sedimentary geological records; they are the object of study of cyclostratigraphy (Weedon, 2003). These cycles bear critical information on the dynamics of the Earth's system and can be viewed as the expression of either **oscillators** or **heterogeneous processes** (Authors' note: the abbreviations and concepts in bold are recapped in Box 1).

An oscillator is mathematically defined as a **dynamical system**, i.e. a system whose state evolves with time, where variables alternate between high and low values. In idealised conditions, the variables may fluctuate at given frequencies (i.e. periodically), or as a finite sum of periodic components (i.e. quasi-periodically).

Oscillators can be conservative, i.e. non-dissipative. This means that energy and angular momentum in the system are constant as a function of time; in simpler words, there are no hindrance, dampening or irregularity in the fluctuation of variables. Accordingly, some conservative oscillators can display clockwork-like regularity.

The orbital and rotational motions of planets are very well approximated as regular conservative oscillators, which allowed their use throughout human history to define the passage of time: the 24-hour day cycle follows the rotation of the Earth around its axis, the months were historically defined by the orbit of the Moon around the Earth, and the year follows the revolution of the Earth around the Sun. Because of the virtually conservative and regular character of celestial mechanics, astronomical cycles unfolding at the time scale of geological processes can be modelled computationally.

The exact way that the Earth is oriented towards the Sun can be modelled with good accuracy up to ca. 10 million of years ago (Lourens et al., 2001; Laskar et al. 2004), by calculating the **climatic precession** and **obliquity** parameters. This is, however, difficult to model further back in time in a reliable way, due to dissipative effects: typically the tides caused by the gravitational influences of the Earth, Moon and Sun on each other. Accordingly, when going further than 10 million of years ago, the expression of astronomical cycles recorded in geological records shows increasingly important discrepancy with astronomical models computed purely from measurements of the current state of the solar system (Zeeden et al. 2014). The computation of the parameters of the orbits of the planets in the solar system (i.e. simply obtaining the position of the planets, not their orientation, e.g., the **eccentricity** of their orbits), is less affected by dissipative effects such as tides, and can be done reliably up to ca. 50 million of years ago (Laskar et al. 2011; Zeebe and Lourens 2019). Beyond 50 million of years ago, the **deterministic chaotic** behaviour of the Solar system prevents the precise computing of the orbits (Laskar et al. 2004).

Climatic precession: change in the angle between the perihelion, i.e. the closest point to the Sun of the Earth's orbit, and the vernal point, i.e. the position of Earth in its orbit around the Sun at the March equinox.
Component: sub-signal that is part of a decomposition.
Cyclically-connoted residual: monotonic function having at most one extremum, and assumed to be centred on either zero (symmetric cyclically connoted residual) or another value (asymmetric cyclically connoted residual) if the data set was of infinite length. If symmetric, the cyclically connoted residual can be studied using instantaneous frequency, and if asymmetric it can be compared to reference curves of the interpreted cycle.
Decomposition: set of sub-signals that can be summed to reconstitute a given signal, or a procedure to obtain such a set (e.g., Fig 1).
Demodulation: extraction of the variations of amplitude of a signal or sub-signal (e.g., Fig. 7).
Deterministic Chaos: unpredictable behaviour of deterministic dynamic systems having high sensitivity to initial conditions.
Diagenesis: ensemble of processes happening after the deposition of a sediment.
DQ: direct quadrature (see section 3 for more details).
Dyadic: property of doubling the mean period of successive filters, in a collection of filters used to perform decomposition (Franzke, 2009; Wu and Huang 2009).
Dynamical system: system whose state evolves with time.
Eccentricity: measure of the shape of an ellipse (ranging from 0, for a perfect circle, to 1, for a parabola). In cyclostratigraphy this refers to the shape of Earth's orbit.
EEMD: Ensemble Empirical Mode Decomposition; application of the empirical mode decomposition algorithm to an ensemble of white-noise added realisations, which are averaged to obtain a physically meaningful decomposition.
EMD: Empirical Mode Decomposition; decomposition algorithm based on the iterative removal of a smoothed version of the signal, through iterated sifting.
FFT: Fast Fourier Transform.
Filtering: extraction of a part of a signal based on specific properties of that signal part, usually frequency.
GZC: Generalised Zero-Crossing (see section 3 for more details).
Heterogeneous process: process generating and reinforcing heterogeneities.
HT: Hilbert transform (see section 3 for more details).
Instantaneous: referring to an infinitesimally small point in time, height or depth.
Instantaneous Frequency: extension of the definition of frequency, to make it a continuous value.
Intra-wave frequency modulation: fluctuations of the instantaneous frequency at a scale lower than the period of the cycle. It is typically observed in oscillations deviating from a sinusoidal form.
IMF (Intrinsic Mode Function): sub-signal (also called function) for which (I) zero-crossing and extrema follow each other in a minima / zero-crossing / maxima / zero-crossing pattern and (II) the upper and lower envelopes, respectively defined by the maxima on one side and the minima on the other, cancel each other at any point, meaning that they are symmetrical.
ka: thousands of years *ago*.
kyr: thousands of years.
Linear trend: linear function, typically completing a decomposition made otherwise of several IMFs, and of a cyclically connoted residual.
Local: referring to a given interval of time, height or depth.
Meaningful: character of a sub-signal to be fully representative of the expression of one single process in the signal.
Modally meaningful: the character of a sub-signal to be fully representative of all the cyclical processes in a signal that, due to their neighbouring frequency range, are interfering with the expression of one another when summed, generating amplitude and frequency modulation.
Mode mixing: widely disparate scales for the oscillations in a component of a decomposition.
NHT: Normalised Hilbert transform (see section 3 for more details).
Nonlinear trend: monotonic function in which there can be at most one extremum, and on which no assumption is made for its mean. This can be understood as the sum of a monotonic residual that cannot be interpreted as cyclical, and the linear trend.
Nonlinearity: property of a system to have an unproportioned response to a change of input.
Non-stationarity: property of time series to be changing; in practice, having changing mean and variance.
Normalisation: separation of an IMF into a frequency carrier of amplitude 1 and its amplitude modulated envelope obtained through spline fitting.
Obliquity: the angle between the Earth's equator and the ecliptic, i.e. the plane of the Earth's orbit.
Oscillator: dynamical system where variables fluctuate between high and low values.

Box 1: Abbreviations and concepts (part 1)

Parsimonious: character of a decomposition to display as little cross-cancellation as possible.
Reconstruction: the action of reconstituting a signal by the addition of a set of sub-signals.
Reversibility: providing the output of any kind of data processing in a data form that allows reconstructing the data in the form it had in the input.
Riding wave: oscillations for which each neighbouring extrema are not separated by a zero-crossing. This is a characteristic that prevents a sub-signal to be a defined as an IMF (e.g., Fig. 6).
Sedimentation rate: thickness of sediment being deposited as a function of time.
Sifting: the process applied to extract each component in an EMD: a smoothed version of a signal is removed from the signal to obtain the extracted component. The smoothed component is obtained by averaging the upper and lower envelopes of the signal, respectively defined by the minima on one side and maxima on the other (see section 3 for further details).
Sifting number: the number of times sifting is applied on each component in EMD. This iteration is performed to eliminate riding waves and to smooth uneven amplitudes.
Stratigraphic quanta: smallest depth or height interval having stratigraphical relevance.
Tuning: procedure of adjusting the time model of a sediment record, to match the cycles found in the record with time-constrained astronomical cycles; it can also be aided by other kind of stratigraphic constraints (e.g., radioisotopic dating, palaeontology, or magnetostratigraphy, etc.).
Unity Frequency Modulation: frequency modulation (or similarly, height, depth or time modulation) of a frequency carrier to make it into a sinusoid of period 1.
White noise: noise from random and uncorrelated fluctuations, e.g., sampled from a Gaussian distribution. In Fourier spectra it has similar power over the full frequency spectrum, like white light.

Box 1: Abbreviations and concepts (part 2)

Astronomical cycles have the potential to affect climate and other environmental factors at both regional and global scales, making them an important part of the Earth's system. Indeed, they influence the seasonal and spatial distributions of incoming solar energy over the globe. At geological timescales, typically from thousands to millions of years, a few parameters are usually relevant in this respect. Firstly, climatic precession is defined as the change of the angle measured between the perihelion, i.e. the closest point to the Sun of the Earth's elliptic orbit, and the vernal point, which is the position of Earth in its orbit around the Sun at the March equinox. Climatic precession affects the Earth-Sun distance at any particular moment in the year, and therefore controls the seasonality of the amount of energy received by the Earth from the Sun. The climatic precession is quasi-periodic, with a main period of ca. 20 **kyr** (thousands of years) in present times (see Berger 1978, Laskar et al. 2004 and Waltham 2015). The second parameter of interest is obliquity, which is the angle between the equator and the ecliptic (i.e. the plane of the Earth's orbit). Obliquity is the main driving force of the Earth's seasons. Obliquity varies periodically between ca. 22° and 24.5° with a period of about 40 kyr in contemporaneous times (see again Berger 1978, Laskar et al. 2004 and Waltham 2015). The effect of climatic precession on climate is modulated by the elliptic character of the Earth's orbit, which is measured by **eccentricity**. Indeed, if the orbit of the Earth is less elliptic (closer to a perfect circle), the location of the Earth in its own orbit, under the effect of climatic precession, will have less impact on the climate. Eccentricity also influences the total amount of solar energy received by the Earth annually, but to a very small extend (Milankovitch 1941; Laskar et al. 2004). The main

periods of variations of eccentricity are of ca. 100 kyr for the so-called short eccentricity cycle, and of 405 kyr for the long eccentricity cycle (e.g., Berger 1978; Laskar et al. 2004). Remarkably, the 405 kyr long eccentricity cycle ($g_2 - g_5$ cycle; see Laskar et al. 2004), along with a 173 kyr obliquity amplitude modulation cycle (s_3 - s_6 cycle; see Boulila et al. 2018 and Charbonnier et al. 2018) are part of a few peculiar cycles for which the period is considered stable beyond 50 million of years ago. Other modulators of climatic precession and obliquity are detailed in Hinnov (2000).

In contrast to the virtually conservative astronomical cycles, the oscillators in the ocean and atmosphere systems are dominated by non-conservative processes such as friction and heat exchanges. In that case, oscillators are called dissipative. For such dissipative oscillators to be maintained through time, the energy lost to heat exchanges and friction needs to be compensated by an external source of energy. Typically, it is the energy from the Sun that sustains the dissipative oscillators found in the climate system, such as the El Niño-Southern Oscillation (ENSO) and the millennial oscillations of the ocean circulation (e.g., Dijkstra and Ghil 2005).

Oscillations produced by dissipative oscillators can theoretically be periodic, or close to it (a good example being gas engines). However, Earth system processes, including dissipative climatic oscillators, influence and perturb one another. This effect is heightened by the fact that, contrary to conservative oscillators which preserve their initial regular oscillations, dissipative oscillators create oscillations continuously anew. Regular oscillations occur in dissipative climatic oscillators, but these can then be drowned by perturbations, only for new regular oscillation to be regenerated without preserving the continuity of the previous regular oscillations. In other words, dissipative oscillators can forget and skip their own oscillations; this is a phenomenon called finite-memory effect, which makes dissipative oscillators unreliable to keep track of time.

As just shown, oscillators affecting the climate can either be dissipative, or almost conservative. The expression of these oscillators on the climate can be traced back in sedimentary records through the use of climatic proxies, i.e. climate-sensitive physical, chemical, mineralogical or paleontological markers. In other words, climatic oscillations can be found as patterns in the sedimentary records. But patterns in sediments can also be created by generation and reinforcement of heterogeneities, i.e. heterogeneous processes. Heterogeneous processes can occur during the transformation of soft sediments into rocks (i.e. **diagenesis**). Vertical spatial heterogeneity, especially, can originate from the combined effects of small depositional changes and pressure-related processes, which can lead to stratified lithological alternations. It has been shown that post-depositional heterogeneities, induced from a sediment that was virtually homogeneous to begin with, can create strictly diagenetic lithological alternations (see Westphal et al., 2015, especially their Fig. 2 for a visual demonstration in calcareous sediments).

Even the crystallisation of diagenetic minerals itself can be heterogeneous (e.g., in cherts; Maliva and Siever 1989), and thus lead to lithological alternations. Heterogeneous processes are space-time processes which can be modelled by a class of dissipative dynamical systems called "reaction-diffusion systems" (Nicolis and Wit 2007; Murray 2013).

This all suggests that repetitive patterns in sediments can have different origins. More importantly, astronomical cycles, dissipative oscillators, and heterogeneous processes may end up entangled in the sedimentary records. Climatic cycles can match the pace of astronomical cycles, with the astronomical cycles acting as a memory for the climatic ones. But due to perturbations in the Earth's system, astronomical and climatic cycles may not necessarily synchronise flawlessly. Furthermore, diagenetic processes have the potential to either undermine or reinforce the expression of climatic cycles. Still, as the astronomical cycles are fairly regular and predictable, it was suggested that they could be discriminated from other processes using time-series analysis techniques, such as Fourier-based spectral estimation. With these techniques, Hays et al. (1976) famously first identified patterns with periods corresponding to climatic precession and obliquity in deep-sea sediments cored from the southern Indian Ocean, and confirmed that the last four Quaternary ice ages each had a duration of approximately 100 kyr, which suggested a link to eccentricity (e.g., Raymo 1997).

Since the work of Hays et al. in 1976, the field of cyclostratigraphy has been expanding considerably. Among others, cyclostratigraphy took advantage of the unique quasi-periodic properties of astronomical cycles for duration estimates. The regularity and predictability of astronomical cycles indeed allows them to be used for dating stratigraphic intervals, and determining the evolution of the **sedimentation rate** (i.e. the thickness of sediment being deposited as a function of time) within these intervals. Dating geological intervals based on the astronomical cycles is the object of study of astrochronology. A unique advantage of astrochronology is its potential to provide a continuous record of time. This contrasts with most other dating methods, which give time estimations that are discretely positioned (i.e. **instantaneous**) or restricted to a given interval (i.e. **local,** such as the interval defined by two points on both sides of a magnetic inversion) in the time scale (or in the depth/height scale for stratigraphic records where time is related to the spatial accumulation of sediments).

In this article, we will focus on the ubiquitous **decomposition** aspect of the time-series analysis methods. Decomposition is defined as any method used to obtain a set of sub-signals called **components**, which can be summed to reconstitute the signal they are derived from. We also use the term decomposition to refer to the ensemble of components that results from this procedure. The main advantage of decomposition is its potential to

provide a comprehensive interpretation of a given signal, by attributing a given process (e.g., periodic or quasi-periodic cycle, heterogeneous process, noise, etc.) to each component.

We will here review how decomposition can be used to better understand sedimentary cycles. Firstly, we will use concepts related to decomposition to assess the time-series methods commonly used for the purpose of finding the cycles of astronomical origin. Then, we will review some of the available tools related to signal decomposition, which have not been used regularly for cyclostratigraphic analysis so far. To improve upon existing tools, and adapt them specifically to cyclostratigraphy, we present a new algorithm, "extricate", to perform decomposition, which runs in a matter of seconds. We demonstrate that through decomposition a new type of frequency analysis can be performed, using the concept of **instantaneous frequency**. Further computational tools are provided; among others, quantified metrics to assess the quality of decompositions. All of the computational tools are implemented in R, a free open-source software environment (R Core Team 2019), through the DecomposeR package (https://CRAN.R-project.org/package=DecomposeR).

Our approach generally pertains to the goal of documenting the quality of cyclostratigraphic results. Indeed, there has been a recent effort of the cyclostratigraphic community to provide flexible guidelines for effective cyclostratigraphic studies (Sinnesael et al. 2019). These guidelines are flexible for the reason that each combination of proxy, archive, geological age and environmental setting requires a tailored approach, which limits the applicability of a strict set of rules (Sinnesael et al. 2019). With our method, we intent to show that cyclostratigraphic studies can be documented through decomposition, which can be performed by various methods. This would allow to further develop quality metrics specific to decompositions of cyclostratigraphic signals, and would provide a standard reference for scientific discussions.

The discussion on how decomposition relates to other time-series analysis methods is provided in section 2. In section 3 we discuss on the techniques available to perform a quantitative instantaneous frequency analysis on decompositions, and on how instantaneous frequency can be interpreted in terms of physical and geological processes. In section 4 we present practical aspects of our novel tools to perform decompositions, assess their quality, and apply instantaneous frequency analysis on them.

We present and analyse, in section 5, a case study meant to illustrate the practical aspects of the application of decomposition and of related concepts to a real geological signal. This case study is made of a rock colour signal that covers an interval between 9 and 5 millions of years ago. The signal comes from site 926 of the Ocean Drilling Program, which is located at Ceara Rise (western equatorial Atlantic). As attempts to make a comprehensive analysis of a decomposition in the context of cyclostratigraphy are scarce and recent (see for instance Zeng et al.

2019 and Sun et al. 2021), some perspective remains to be gained on the subject. Therefore, we also give in the supplementary material (sections 9 and 10) further examples of comprehensive analysis of decomposition; on the magnetic susceptibility of site 926 in the same age interval, and on a decomposition from an artificial signal (the first case of the Cyclostratigraphic Intercomparison Project; Sinnesael et al. 2019). We refer to those additional analyses during our interpretation of the greyscale signal of site 926. We also provide an illustration of how a climatic noise model behaves in a decomposition (section 8 in the supplementary materials). Overall, the case studies in the supplementary materials provide insight on how to apply and interpret decompositions performed on various time-series.

We finally discuss (section 6) and conclude (section 7) on the potential of decomposition. Ultimately, we want to show that decomposition can provide a fundamental understanding on all aspects of signal processing, and therefore makes for a perfect tool for tackling cyclostratigraphic signals.

2 Meaningfulness, parsimony, and invertibility

Among the time-series analysis methods used to identify and characterise oscillations in a signal, decomposition is a simple way to overview the properties of a signal, which we illustrate in Figure 1.

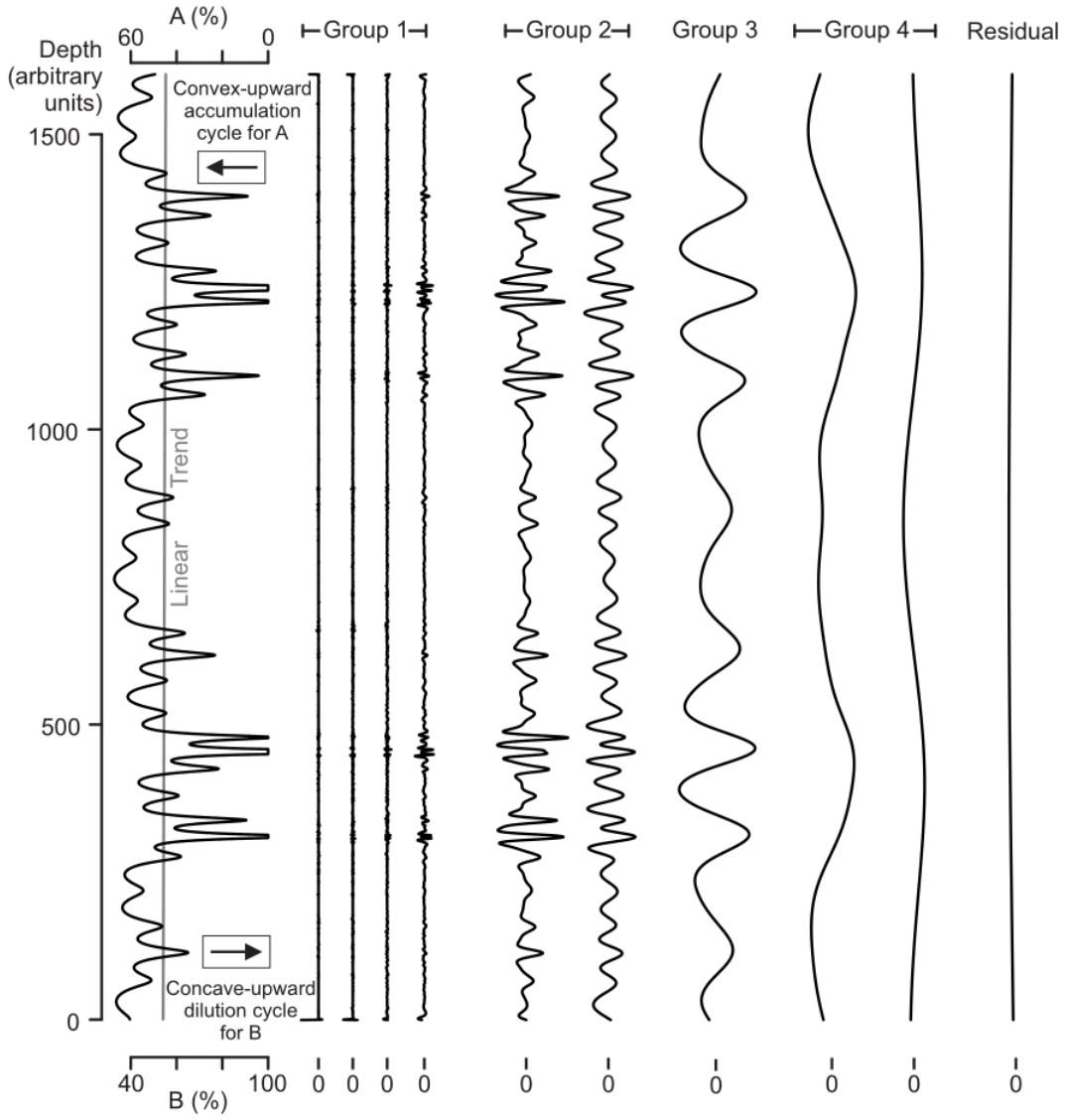


Figure 1: Decomposition of a depth-distorted artificial signal. The amplitude of the components is scaled to that of the initial signal. The linear trend is part of the decomposition. The signal is generated based on a model suggested by Herbert (1994): two components, A and B, having different sensibilities to a simplified astronomical forcing, are accumulating at variable rates and at variable proportions with one another. The simplified astronomical forcing is made of 3 sinusoids having periods of 23, 95, and 405 kyr. The model modulates this forcing onto a distorted stratigraphic depth. In our case, the maxima of accumulation of the element A expands the sedimentation rate, creating for the element A patterns called "accumulation cycles" (convex-upward when considering the axis of A), and by contrast, diluting the element B in a pattern of "dilution cycles" (concave-upward when considering the axis of B). The first 4 components (group 1) take in small-scale high-frequency variations that are best explained as artefacts of the EEMD-based decomposition method, but also locally as the expression of abrupt features in the signal. Groups 2, 3 and 4 are respectively representative of the 23, 95, and 405 kyr forcing sinusoids. The residual and trend are basically flat, and do not carry information in this example. The decomposition is performed using the "extricate" EEMD method explained in section 4.1. The related R script (generating the signal, its decomposition, and this figure) is available in the SC1 supplementary material. [NOTE TO REVIEWERS: 2 column fitting image]

To be intuitively understandable, a decomposition should be **meaningful** (Wu and Huang 2009), by which we mean that the components must *fully* represent the expression of single physical or geological processes in the signal. We emphasize the "fully", meaning that the entire expression of a single process needs to be included into a meaningful component. Depending on the decomposition, components can be merged to be meaningful, e.g., the components in the groups of Figure 1. Whether a decomposition can be called meaningful is a subjective matter; but it can be justified by several quantified criteria, such as parsimony and invertibility that we will present below, or any metric that can establish a clear correspondence of particular components with given processes.

We define here the **parsimonious** character of a decomposition as the property of having as little cross-cancellation as possible. The level of cross-cancellation is linked to the level of excess absolute intensity that a decomposition has relative to its original signal (Fig. 2). Cross-cancellation occurs in intervals (in time, or in stratigraphic depth or height) where components of a decomposition are of opposite signs (Fig. 2A). As components are summed, they cancel each other out in those intervals (Fig. 2B). The excess of absolute intensity can be quantified: we use it to define a *parsimony* metric (see section 4.2.3).

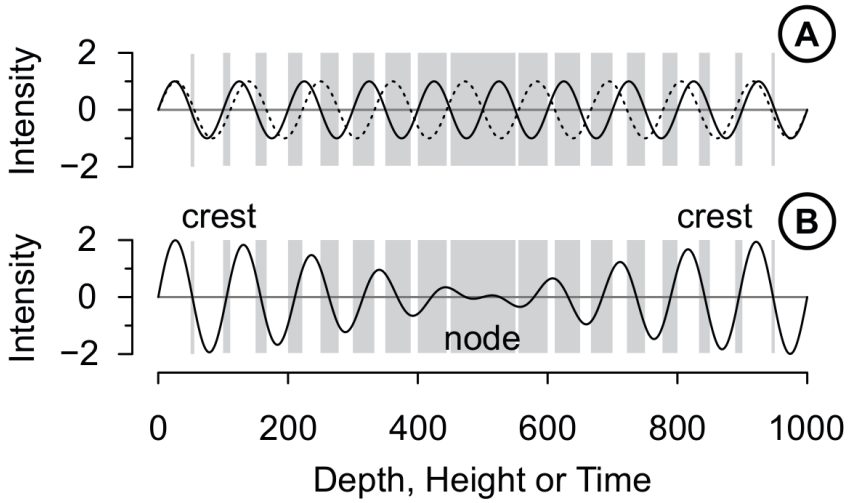


Figure 2: Cross-cancellation of two sinusoidal components of nearby frequencies (A), occurring when the two sinusoids are summed to reconstruct a given signal (B). Cross-cancellation occurs in the intervals where components are of opposite signs; they are in indicated by the light-grey rectangles. The cross-cancellation occurs more prominently near the nodes, creating an amplitude modulation called "beat". The absolute intensity of the two sinusoids is mostly preserved and cumulated at the crests, but lost at the nodes, resulting in an overall loss of absolute intensity. [NOTE TO REVIEWERS: 1.5 column fitting image]

A parsimonious decomposition remains faithful to the original signal. Contrastingly, a poorly parsimonious decomposition tends to create predicted wiggles, i.e. wiggles that are not present in the initial signal. The interpretation of predicted wiggles is a challenge; it can rely on *a priori* knowledge, or on underlying assumptions on their origin. Critically, such assumptions must commonly relate to the frequency and regularity of the predicted wiggles. Therefore, it is difficult to have reliable interpretations on predicted wiggles when the

signals are affected by erratic frequency modulations, or more typically in geological records, when they are affected by strong sedimentation rate changes. Furthermore, signal disturbances such as complex noise processes make the interpretation of the wiggles observably present in the raw signal already challenging; for predicted wiggles, finding a physical meaning is all the more uncertain. We suggest that by having a more parsimonious decomposition, one can minimise the amount of predicted wiggles; accordingly, the wiggles that remain in the parsimonious decomposition will be as close as possible as those found in the original signal. In turn, the need for assumptions and the risk of incorrect interpretations will be reduced. Considering the complex nature of cyclostratigraphic signals, we suggest that the concept of parsimony to minimise cross-cancellation can be useful to perform meaningful time-series analyses. Ideally, one can obtain decompositions in which the wiggles studied are all observably present in the signal. We call this the wiggle-in-signal approach.

A decomposition is mathematically invertible: from the addition of all components, the initial signal can be reconstituted flawlessly (an action called **reconstruction**). This is necessary to characterise parsimony, and implies that from a decomposition the entirety of a signal could be interpreted in the form of meaningful components.

The concept of meaningfulness (e.g., Wu and Huang 2009) is rarely addressed when using methods that do not generate decompositions as an output, and the concept of parsimony is introduced only in this paper. However, those two concepts can provide insight on how several time-series analysis methods are performed; especially because decomposition is often an integral part -albeit sometimes hidden- of the time-series analysis techniques used in cyclostratigraphy. Therefore, we will review in this section the methods commonly used to process cyclostratigraphic signals, through the lenses of meaningfulness and parsimony. To do this, we will reframe these methods in the context of decomposition, which implies that such methods can be invertible.

A familiar mathematical application of decomposition is the Fourier decomposition, which decomposes a signal into a sum of sines and cosines. As sines and cosines are periodic functions, all components of the Fourier decomposition are periodic as well. However, for many signals, it makes poorly parsimonious decompositions; this can be illustrated by a signal consisting of an amplitude-modulated carrier. In simple amplitude-modulated carriers, the Fourier decomposition can be dominated by two components having nearby frequencies (Fig. 2A). When added together these components produce a pattern called a "beat": they interfere constructively near the crests, and destructively at the nodes (Fig. 2B). If the modulation changes over time, or if the frequency of the carrier evolves, then the number of components grows quickly, destructive interferences increase, which then worsen the parsimony of the Fourier decomposition.

In sedimentary records, Fourier decomposition often produces poorly parsimonious and poorly meaningful decompositions, especially when sedimentation rates vary strongly, distorting the signal (see also discussion in Hilgen et al. 2014). Indeed, sines and cosines are an appropriate conceptualisation of a periodic or quasi-periodic signal in the geological record only when sedimentation rates are close to being constant. This can sometimes be the case; when the signal does contain a strong periodic component (caused, for example, by a periodic forcing), and when the sedimentation rate is constant, a power spectrum (the energy or power that is attributed to each frequency) obtained by Fourier decomposition will be characterised by a narrow peak of energy or power at the frequency of the periodic component (similar to what is observed in Fig. 3A). Specific statistical tests have therefore been devised, based on power spectra, to detect astronomical cycles (e.g., Percival and Walden 1993, Mann and Lees 1996, Ghil et al. 2002, Meyers 2012, Lenoir and Crucifix 2018). These tests allow to probe for the presence of periodic components against given null hypotheses, which are defined by the noise processes expected to occur. Suppose, however, that the climatic signal contains a strong periodic component, but that the sedimentation rate varies. As a result, the initially periodic signal will be frequency-modulated in the sedimentary record. Instead of the sharp spectral peaks of the astronomical forcing, which are expected to occur in the power spectra, the energy or power is spread in larger peaks for which the width depend on the intensity of the variations of the sedimentation rate, or in separate peaks (e.g., Fig. 3B; see also Schiffelbein and Dorman, 1986, and Herbert, 1994). Variable sedimentation rates do not necessarily harm the entirety of a power spectra, notably when the sedimentation rate is changing at a frequency virtually similar to that of a forcing cycle (e.g., the 405 kyr cycle in Fig. 3). However, this is only one of the many possible scenarios occurring during sedimentation. Similarly to a decomposition not being parsimonious enough, the loss of spectral sharpness is a process that can make the Fourier decomposition less representative of the signal (see also Kemp 2016).

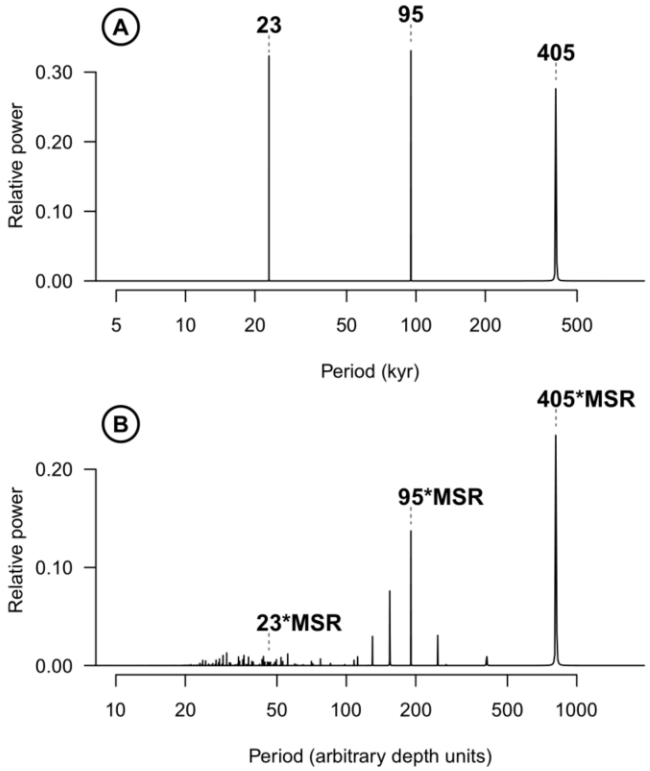


Figure 3: periodograms of (A) the sedimentation rate of the element B (i.e. the forcing used to make the signal in Figure 1), and (B) of the depth-distorted signal of the proportion of the same element B (i.e. the signal used in Fig. 1). MSR stands for Mean Sedimentation Rate. The depth distortion creates strong loss of spectral sharpness at high frequencies, making the 23 kyr cycle hardly decipherable by Fourier transform, and the 95 kyr forcing partially spread out in separate peaks. The 405 kyr peak is mostly unscathed, as the depth-distortion occurs virtually at the same 1/405 kyr^{-1} frequency. The signals for Figures 3A and 3B were extended 50 times compared to the one of Figure 1 in order to improve the frequency resolution. The R script generating the forcing and depth-distorted signals, their periodograms and this figure is available in the SC1 supplementary material. [NOTE TO REVIEWERS: 1.5 column fitting image]

To account for changing sedimentation rate, spectral estimation can be adapted to provide a windowed estimation of frequency, based most commonly on wavelet analysis and evolutionary spectra (e.g., Meyers et al., 2001; Hinnov 2013; Kodama and Hinnov 2014; and Martinez et al. 2015, for evolutive spectra, and Torrence and Compo, 1998, for wavelet analysis). Essentially, those two methods can be thought of as decompositions where components are representative of a given frequency *and* a given interval in time, depth or height (i.e. the window); this makes for a substantial number of components that are very local, and in theory reduces parsimony problems caused by sedimentation rate changes. Meaningful decompositions could be obtained with these techniques by combining, throughout the signal's length, all of the local components that are relevant for distinct processes. However, this requires methodological developments on continuous wavelet transform using ridge extraction (e.g., Qin et al., 2016) or on discrete wavelet transforms (Percival and Walden 2000); these are outside the scope of this paper.

Having frequency estimated locally allows to study the local correlation and the frequency ratios of different scale cycles. Indeed, when the sedimentation rate locally increases by a factor *k*, the periods of each cycle, as recorded in the geological strata, increase proportionally to *k*. Therefore, the period (or frequency) ratios between regular astronomical cycles remain locally constant even with varying sedimentation rate. The values of the ratios can be used to identify the cycles, especially if multiple independent ratios can be found (Meyers et al., 2012; Waltham 2015). The presence of stable frequency ratios respecting the general astronomical hierarchy and behaviour of cycles in relation with each other, should be reliable indicators of the unique regularity of astronomical cycles, and could serve as a criteria for a meaningful extraction of astronomical cycles. We include in these stable frequency ratios, the ratios related to the cycles driving amplitude modulation. Identification of these properties unique to astronomical cycles is the basis for several astronomical cyclicity identification techniques (Meyers 2015; 2019; Zeeden et al. 2015; 2019). In the context of decomposition, extracting components to test such ratios can be done using **filtering**.

Once the parts of a signal attributable to an astronomical origin are identified, for instance with windowed estimation of frequency, they can be extracted by filtering. Filtering is the procedure that extracts a part of a signal based on certain properties, usually frequency. One then defines a **filter bank** as a set of filters that can be used to decompose a signal, typically by having each filter cover a complementary part of the entire frequency range of the signal. Fourier decomposition can provide the basis to make a filter: the latter is obtained by adding up all the Fourier components belonging to a specific range of frequencies. Doing so allows to extract the part of the signal making up that specific frequency range (see for instance the "Band pass filter" section of appendix 2 in Muller and MacDonald 2000). Filtering is highly customizable; it is therefore difficult to theorise how meaningfulness and parsimony behave with them. When used to extract climatic precession or obliquity, and study their amplitude modulation, a common rule is to use a wide frequency range around the target frequency, which reduces the generation of unwanted spurious amplitude variations (Zeeden et al. 2015; 2019, and references therein). In general, including all the frequency bands related to a single process in a single filter would bring it closer to meaningfulness, and merging as many frequency bands as possible would indeed improve the parsimony.

Other methods exist that are somewhat similar to the filtering approach, whereby a signal is subdivided into a set of sub-signals and a residual: this is a kind of decomposition. Among such techniques, the most popular in cyclostratigraphy is the Singular Spectrum Analysis method, or SSA (Vautard and Ghil 1989). Essentially, SSA first finds eigenmodes from the signal, which in practice appear as sub-signals shorter than the signal's entire length. Then, these eigenmodes are repeated to each reconstruct a component, with a weighting by coefficients

that vary throughout the length of the signal, which reconstitutes amplitude features. Moreover, the components can be merged by the user, generally by pairs, which allows SSA to produce sub-signals with a certain degree of frequency modulation. Still, SSA is limited by the fact that the eigenmodes themselves are not stretched or compressed in the reconstruction process, thereby limiting the frequency modulation (or depth distortion) that can be reconstituted. We illustrate this in Figure 4 with a depth-distorted signal, where a substantial part of the amplitude of the most distorted forcing (the 23 kyr one) remains in the residual part, thereby limiting the meaningfulness of this type of decomposition.

Both SSA and filters accommodate only a limited amount of frequency modulation (in other words depth-distortion); they can thus be inadequate when the sedimentation rate varies too much. This problem can be solved by determining the sedimentation rate through **tuning**. Tuning, as we refer to it in this study, is the manipulation by which the time model of a sediment record is adjusted so to match the cycles found in the record with time-constrained astronomical cycles; it can also be aided by other kind of stratigraphic constraints (e.g., radioisotopic dating, palaeontology, magnetostratigraphy, etc.).

Sedimentation rate estimates based on the identification of the regular astronomical cycles in climatic can be made for a complete record, considering a uniformly stable sedimentation rate, using methods such as the average spectral misfit (ASM) of Meyers and Sageman (2007) and TimeOpt of Meyers (2015; 2019). More importantly for our purpose, methods have also even been established to provide a local or instantaneous estimate of how the sedimentation rate evolves as the record unwinds (Malinverno et al. 2010; Meyers et al., 2012; Sinnesael et al. 2016; Li et al. 2018b; Sinnesael et al. 2018; Meyers 2019).

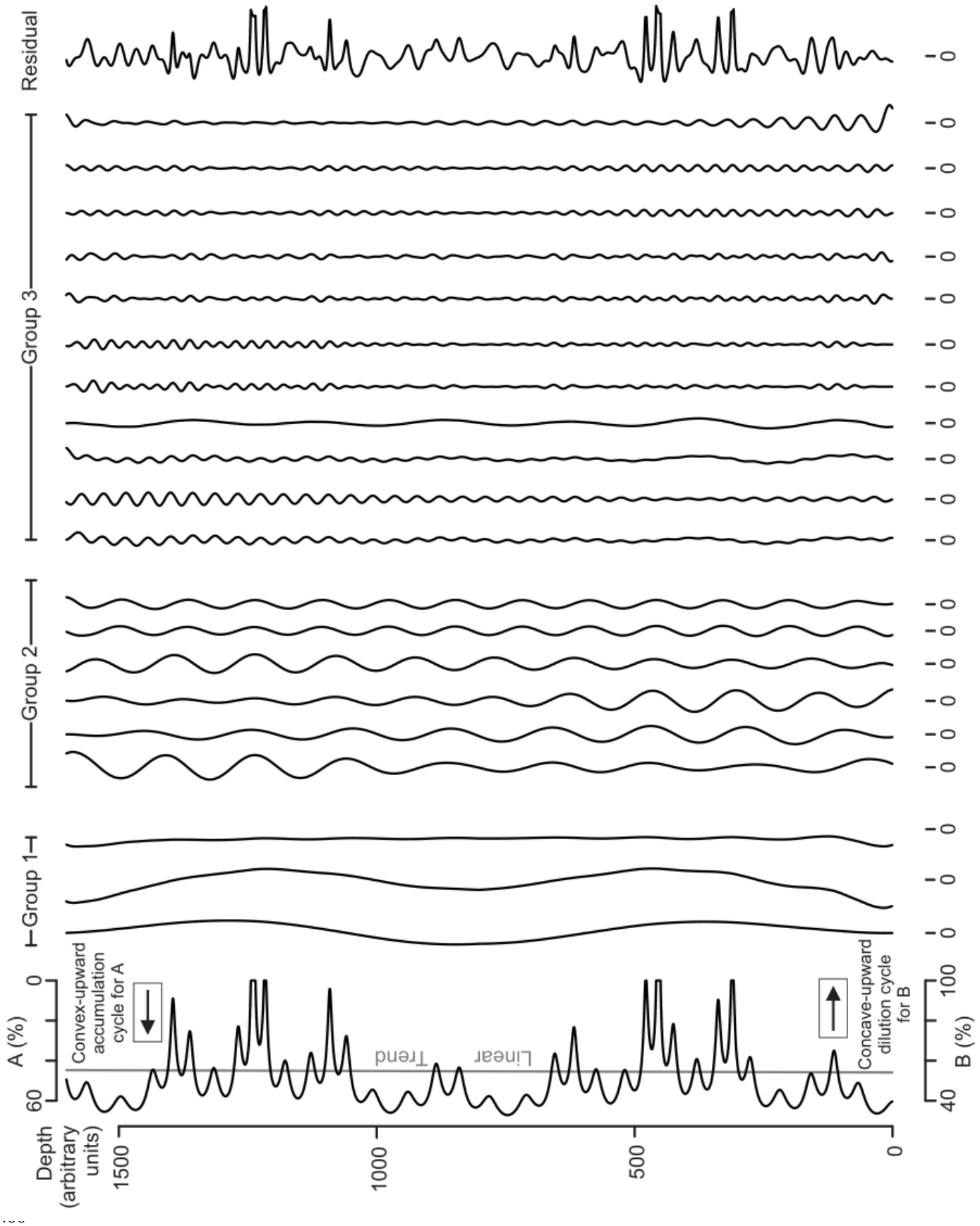


Figure 4

Figure 4: Decomposition of a depth-distorted stratigraphic signal (identical to Figure 1), using Singular Spectrum Analysis (SSA). The amplitude of the components is scaled to that of the initial signal. The

linear trend is part of the decomposition. The first 3 components (group 1) take in low-frequency components corresponding roughly the 405 kyr forcing, but some of the 95 kyr forcing contaminates the 2nd component. Group 2 and 3 roughly correspond to the 95 and 23 kyr forcing cycles respectively, with some spurious cycles not corresponding to any forcing cycle. The residual still contains significant amplitude related mainly to the 23 kyr forcing cycle. This parallels the leakage observed in the periodogram in Figure 3B, affecting mostly the 23 kyr cycle. The related R script (generating the signal, the decomposition and this figure) is available in the SC1 supplementary material. [NOTE TO REVIEWERS: 2 columns fitting image]

What tuning does in practice to the signal is stretching or compressing it in given intervals; in other words a depth distortion (or a frequency modulation) is applied on the initial signal. The aim is that this depth distortion counteracts the effects of sedimentation rate changes, but this is not necessarily the case; for instance when the interpretation of the signal's cycles, in terms of astronomical origin, is erroneous. Generally speaking, tuning refers to the *intention* of matching a signal to the astronomical cycles, without guaranteeing that it is done correctly. Therefore, we will in this paper speak of depth distortion of a signal for tuning purposes, or of a tuning interpretation, rather than merely "tuning". We introduce this distinction to underline that a tuning interpretation applied to a signal should be tested thoroughly, and even further, that it should be tested regularly as new data becomes available.

However, testing whether a tuning interpretation is correct can be plagued by several interconnected issues, chiefly: overtuning, amplitude artefacts, and circular reasoning. Overtuning occurs when the sedimentation rates estimated via tuning are strongly and/or abruptly changing, to finely fit a signal to a target curve (typically an astronomical curve), potentially without physical or geological justification. Amplitude artefacts can then occur in the tuned signal due to overtuning. Indeed, because the signal was adapted to finely fit an astronomical target curve, some of the properties of the target curve get transferred onto the tuned signal. When the tuned signal is subsequently filtered in a narrow frequency range matching the tuning target to extract and test amplitude variations, those properties can be expressed as artificial amplitude modulations that mimic the ones found in the target astronomical curve (e.g., section 5.7 in Muller and MacDonald 2000; see also Hilgen et al. 2014 and Zeeden et al. 2015). This leads us to the issue of possible circular reasoning, which can occur when the steps used to perform tuning, by themselves, impose the detection of astronomical properties in the tests applied on the tuned signal.

Solutions to tuning issues have usually focused on the overtuning part of the problem, by taking precautions on how the tuning is performed (see section 5.8 Muller and MacDonald 2000). Typically, avoiding overtuning consists in avoiding using too much data to parametrise the age model (by carrying out the tuning using only a small number of tie-points; by tuning on one target astronomical cycle such as precession, and testing on another like eccentricity; or by using other stratigraphic constraints that were not used to set the tuning interpretation for

testing, such as radioisotopic dating, palaeontology, or magnetostratigraphy, etc.). These precautions can prevent circular reasoning, typically by separating the data used for tuning from the data used for testing. Importantly, the correspondence of the tuned signal to the data used for testing needs to surpass any correspondence from random (or unrelated) signals that are tuned similarly than the signal being tested. This is commonly verified using Monte Carlo testing (e.g. section 3.13 in Muller and MacDonald, 2000). But avoiding overtuning means avoiding the exploration of all the possible age models. What if the sedimentation rate indeed follows abrupt variations, even if they may seem unlikely? What if a dataset tuned in high detail is as close as the correct one as can be? Rather than limiting the tuning possibilities, one would like to be able to use all the information to have the truest sedimentation rate estimation, and find a way to disqualify overtuned interpretations regardless of the artefacts they produce. This is especially relevant for geological records, for which age information is scarce, and should be exploited to its fullest extent. However, such practice has trade-offs. If the depth distortion to perform tuning is allowed to rely on the entirety of astronomical and stratigraphical data, any frequency (or time, depth, height) criteria to test the tuning will be subject to circular reasoning. Alternatively, if we could address the problem of amplitude artefacts, tuned signals could be tested based solely on the amplitude correspondence between a signal and astronomical curves, without the preoccupation of circular reasoning.

Amplitude artefacts can be avoided using parsimonious decomposition methodologies. As the wiggles in a parsimonious decomposition must be observably present in the original signal, they cannot be artefactual wiggles (or, as we called them earlier, predicted wiggles). Alternatively, in the case where a decomposition is not perfectly parsimonious, artefactual wiggles can be tracked down and accounted for, using among others the parsimony metric (section 4.2.3). Therefore, if parsimony is taken into account, even methodologies that are prone to overtuning, such as optimised matching of the observed cycles (e.g., by wiggle matching, such as in Lisiecki and Lisiecki 2002), could be used for tuning. This is all the more important for methodological purposes, to apply tests based on amplitudes, because checking the correspondence in amplitude of a signal with a reference curve requires them to be optimally aligned.

The need for a parsimonious decomposition, to avoid amplitude artefacts, can conflict with making the decomposition meaningful. This is best illustrated in cycles having neighbouring frequency ranges; for instance, the four major components of climatic precession (which at the present day have periods of ca. 23.7, 22.4, 18.9 and 19.1 kyr; Berger 1978; Laskar et al. 2004; Waltham 2015). If one defines each of these four components as separate processes, in a meaningful decomposition they should be separated as four distinct components. However, as similar frequency range cycles, they tend to interfere with one another when combined, cross-cancelling

themselves into frequency and amplitude modulated components. Going the other way around, from the frequency and amplitude modulated wiggles found in signals, to well-isolated and regular sinusoids, can only be achieved through the use of predicted wiggles. In a cyclostratigraphic signal, predicted wiggles could be attributed to the 23.7, 22.4, 18.9 and 19.1 kyr components of precession, but could equally well be artefacts. Because of complex noise processes, sedimentation rate changes, and other mechanisms such as heterogeneous processes, we even argue that interpreting the predicted wiggles found in the decomposition of a cyclostratigraphic signal is hazardous. This all means that meaningfully applying the wiggle-in-signal approach will require some compromises.

To reconcile meaningfulness and parsimony, in the case of cycles having neighbouring frequency ranges, we introduce the concept of **modally meaningful** component extraction. We define modal meaningfulness in an extracted component as the characteristic to be fully representative of all the cyclical processes that, due to their neighbouring frequency range, are interfering with the expression of one another when summed, generating amplitude and frequency modulation, or interferences. For instance, the expression of the four major components of climatic precession would be bundled into a single modally meaningful component including the expression of other similar-frequency-scale processes (among which noise, but not only), that are interfering with the expression of climatic precession. Inversely, the expression of the different frequency scale sub-components of eccentricity (usually, short eccentricity –grouping the main 95 and 125 kyr period components into one modally meaningful component- and long eccentricity -405 kyr-) would be expressed as separate modally meaningful components. In essence, modally meaningful and parsimonious decompositions are a pragmatic implementation of the wiggle-in-signal approach.

To summarize; if the tuning process is performed correctly (which needs to be thoroughly and regularly tested), perturbations induced by sedimentation rate changes on the signal are reduced. On this corrected signal, filters or SSA can be applied to extract sub-signals representative of the astronomical cycles without the issues caused by sedimentation rate changes. This can be done in a parsimonious and modally meaningful way, by regrouping similar components in SSA, or by applying filters with wide bandwidths, thereby avoiding the issue of the interpretation of predicted wiggles. To make decompositions as meaningful as possible, the residuals should be minimised as much as possible. This all illustrates that parsimonious and modally meaningful decomposition can be implemented using tools widely used in the framework of cyclostratigraphy, in a way that can then intuitively be interpreted, as geological and/or physical processes.

3 Frequency analysis of decompositions

A parsimonious and modally meaningful decomposition allows to have an overview of the processes at the origin of the signal. However, this is not enough: to study sedimentary cycles in the context of decomposition, an estimation of the behaviour of their frequency is crucial.

Decomposition exists in the time domain, which means that the decomposition shows how the components of a signal change with time. Decomposition can also be applied on geological time-series expressed as a function of height in an outcrop or of depth in a core, i.e. the stratigraphic domain. Decomposition is here set in contrast with transforms (Fourier or wavelet) that describe a signal into the frequency domain. In the frequency domain, the signal is described as the amount of energy or power attributed to given frequencies. The duality between the time/stratigraphic domain on one hand and the frequency domain on the other seems to imply that frequency analysis of a decomposition is impossible without using a transform, which is not the case. The frequency of signals, when they are filtered or decomposed, may still be characterised quantitatively using the concept of **instantaneous frequency**, i.e. a frequency that continuously evolves in the time, depth or height scale.

Traditionally, frequency is defined as the inverse of the period of a cycle, and the simplest way to determine it is by measuring the length between consecutive zero-crossings (or between consecutive corresponding points of the phase) in successive waves (Huang et al. 2009). By that definition, determining the frequency can only be performed locally, with a resolution of a complete wavelength. However, the concept of frequency can be stretched to become instantaneous. Based on the traditional definition of frequency, the temporal resolution of the frequency computation can actually be lowered to a quarter of the period using the generalised zero-crossing (**GZC**) method. This method is illustrated in Figure 5: for a given point, we compute frequency through a weighted average of the quarter of the cycle period, two half periods, and four full periods based on the extrema and zero-crossings (e.g., Huang et al. 2009). GZC is local, but not yet instantaneous; still, it can serve as a useful reference to test methods providing estimates for the instantaneous frequency.

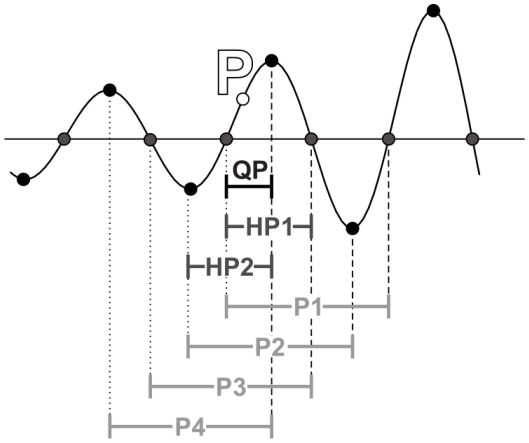


Figure 5: Illustration of the frequency computation by the generalised zero-crossing method. For each point P, the frequency can be computed by making a weighted average of the quarter of the cycle period (QP), two half periods (HP1 and HP2), and four full periods (P1 to P4) that are defined by the extrema and the zero-crossings. [NOTE TO REVIEWERS: 1 column fitting image]

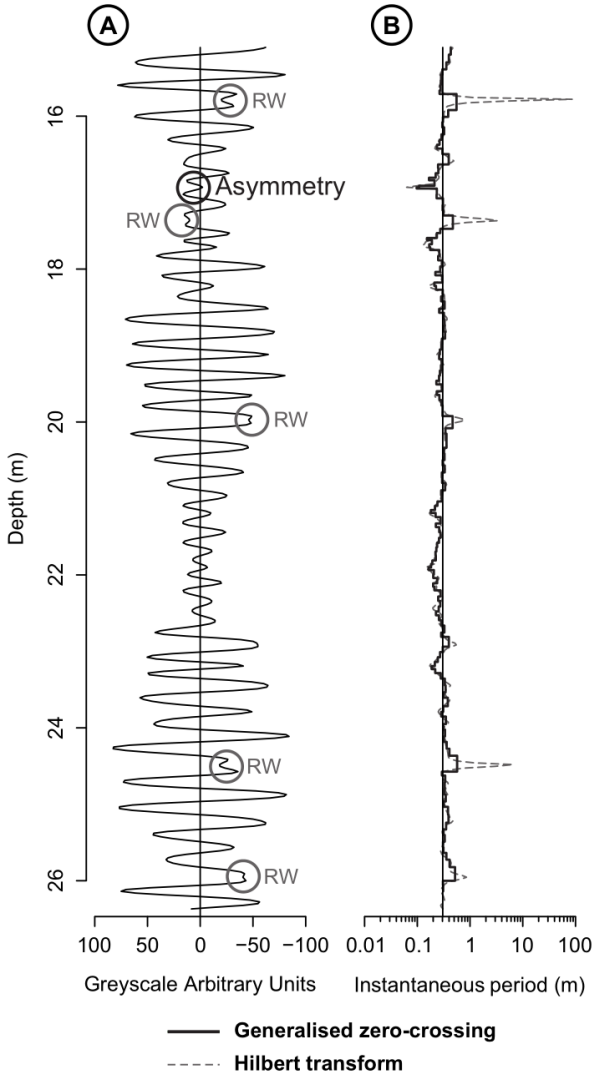


Figure 6

Figure 6: Illustration of instantaneous frequency. (A) Intrinsic mode function (IMF) (excerpt of comp. 2 of the CIP1 decomposition, see supplementary material), for which instantaneous frequency is computed. (B) Two instantaneous frequency computation methods are compared: the GZC (generalised zero-crossing) and the HT (Hilbert transform) methods (see Huang et al. 2009 for details). The straight line stand for the expected frequency of climatic precession (23 kyr divided by 75.81 kyr.m^{-1}, see supplementary material). Note that the largest discrepancies (deviating from the expected frequency by several orders of magnitude) between GZC and HT, and between the instantaneous frequency and the expected frequency, happen at the location of riding waves (RW) and where the component is strongly asymmetric, i.e. in the places where the component deviates from the definition of IMF, which means that the instantaneous frequency computation is locally invalid (see text). [NOTE TO REVIEWERS: 1 column fitting image]

Indeed, there are mathematical tools, such as the Hilbert transform (**HT**; e.g., Hinnov et al. 2002; Huang et al. 2009), that can be used to provide instantaneous values that mimic the behaviour of frequency obtained by generalized zero-crossing (see Fig. 6, and Huang et al. 2009). Such approximations are considered adequate to be used as instantaneous frequency: they have been explored previously on filters and SSA (Ghil et al. 2002; Hinnov et al., 2002), and are reviewed by Huang et al. (2009). We discuss here the main requirements to perform sound instantaneous frequency analysis, especially in the case of decomposition, focusing on the meaning it can have for cyclostratigraphy.

Instantaneous frequency only applies in the case of a monocomponent sub-signal (Boashash 1992; Huang et al. 2009), which can be understood as a single sinusoid, modulated in amplitude and frequency at each point in the time or stratigraphic domain. Coincidentally, we expect a modally meaningful component to be interpretable as a monocomponent signal, and thus have an instantaneous frequency. A more practical definition for sub-signals apt for instantaneous frequency determination is the one of intrinsic mode function (**IMF**). The IMF, as defined by Huang et al. (1998), is a component for which (I) zero-crossing and extrema follow each other in a minima / zero-crossing / maxima / zero-crossing pattern and (II) the upper and lower envelopes, respectively defined by the maxima on one side and the minima on the other, cancel each other at any point, meaning that they are symmetrical (e.g., Fig 7A or 8B). Deviations from this definition, especially **riding waves** (i.e. oscillations for which each neighbouring extremum is not separated by a zero-crossing, e.g., Fig. 6) make it so that instantaneous frequency is no longer interpretable. The way that instantaneous frequency behaves in asymmetric parts of the signal and in riding waves differs depending on how it is computed.

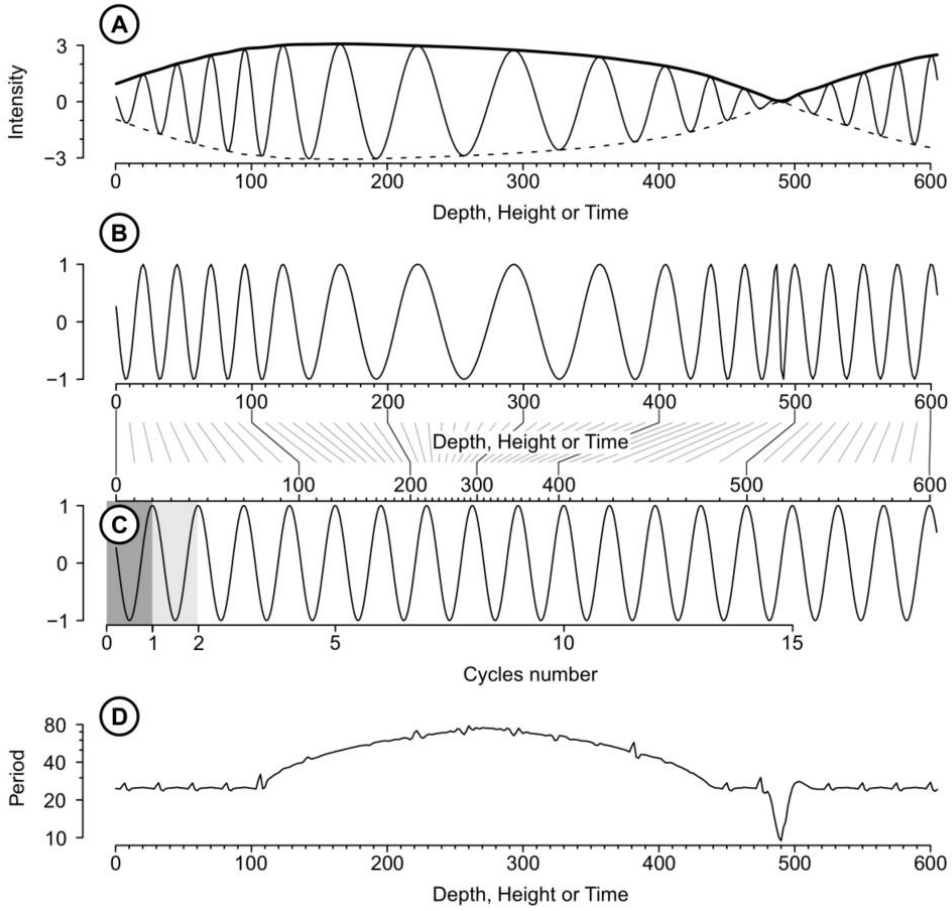


Figure 7: Steps to change an intrinsic mode function (IMF, see Box 1) into a sine function of period 1 and amplitude 1. (A) The signal, for which amplitude is computed as the envelope of the signal. (B) The frequency carrier is computed as the signal divided by the envelope, a procedure called normalisation. (C) A period 1 sine can be obtained by modulating the depth, height or time via unity frequency modulation (see text). (D) The instantaneous frequency of frequency carriers can be computed through the direct quadrature (DQ) method (see text). Note that a similar figure has been presented by other authors to discuss the performance and physical meaning of the Hilbert transform (Figure S3 in Hinnov et al., 2002)
[NOTE TO REVIEWERS: 2 column fitting image.]

Using IMFs offers the possibility to model **nonlinear** and **non-stationary** processes such as the ones that can lead to sedimentary cycles. Therefore, decomposing a signal into IMFs is expected to be better than using transforms to explore the full expression of processes in the signal. Furthermore, IMFs can be understood in a simple mathematical context. Estimates of instantaneous amplitude and frequency can be explained through the steps allowing to change a monocomponent signal into a sinusoid of amplitude 1 and period 1 (see Fig. 7). An IMF can first be changed into a frequency carrier of amplitude 1 by dividing it by its envelope (Fig. 7A and B, also see Huang et al. 2009). From this frequency carrier, the phase angle can be determined for each point as if it belonged to a sine, using the equation from Huang et al. (2009):

$$\phi(t) = arctan \frac{F(t)}{\sqrt{1 - F^2(t)}} \qquad (1)$$

where ϕ is the phase angle (in radians), and F the frequency carrier (that needs to be strictly between -1 and 1) which is a function of the time t, or alternatively, of the height or depth. In plain language, this equation converts the F(t) intensity (between -1 and 1, e.g., Fig 7B) into $\phi(t)$, which is the corresponding angle in a sine; for instance an intensity of 0.5 corresponds to an angle of $\pi/6$ radians, or 30°, as $\sin(\pi/6) = 0.5$. By readjusting the intervals between points to the difference of phase angle between each successive point, divided by 2π to account for radians, the frequency carrier can be changed into a pure sinusoid of period 1 (Fig. 7C), in a procedure we define as **unity frequency modulation**. On the same principle, instantaneous frequency can be computed by differentiation of the phase angle between successive points (Fig 7D): this is the direct quadrature instantaneous frequency computation method (**DQ**; Huang et al. 2009).

As instantaneous frequency can be determined reliably only on components complying with the IMF definition, any given signal that needs to be analysed in instantaneous frequency first has to be decomposed into IMFs. To this end, Huang et al. (1998) first suggested the empirical mode decomposition (**EMD**) algorithm. EMD smooths a record, and extracts components by subtracting the smoothed signal from the initial signal; the process is then iterated on the smoothed signal. In Huang's (1998) algorithm, the smoothing was performed through a process called **sifting** (Fig. 8). In sifting, maxima and minima of the signal are inter- and extrapolated to define an upper and lower envelope, using a smooth interpolation method like spline fitting (Fig. 8A). Then the two envelopes are averaged, forming the smoothed sub-signal that is subsequently subtracted from the signal to obtain a first component. As that component is defined as the average of the envelopes of the signal, it is expected to have symmetrical envelopes (Fig. 8B). Having the algorithm based on local extrema allows the resulting decomposition to be molded after the wiggles as they are in the signal. This makes EMD, in its fundamental form, a wiggle-in-signal approach.

In practice, components extracted by sifting do not necessarily comply with the definition of an IMF: they can still present asymmetry and riding waves. To solve this problem, Huang (1998) introduced an inner loop, which iteratively repeats the sifting algorithm on each extracted component (Fig. 8C). The number of iterations is called the **sifting number**. The sifting iteration eliminates riding waves to bring each component closer to what an IMF should be, and smooths uneven amplitudes (Huang et al. 1998). After the inner loop, a component is extracted (Fig. 8D), and a remainder is defined as the signal minus that extracted component (Fig. 8E). Then the entire process is iterated on that remainder to extract the next components (i.e. the outer loop in Fig. 8F). All the components are extracted when the remaining smoothed sub-signal contains one extremum at most.

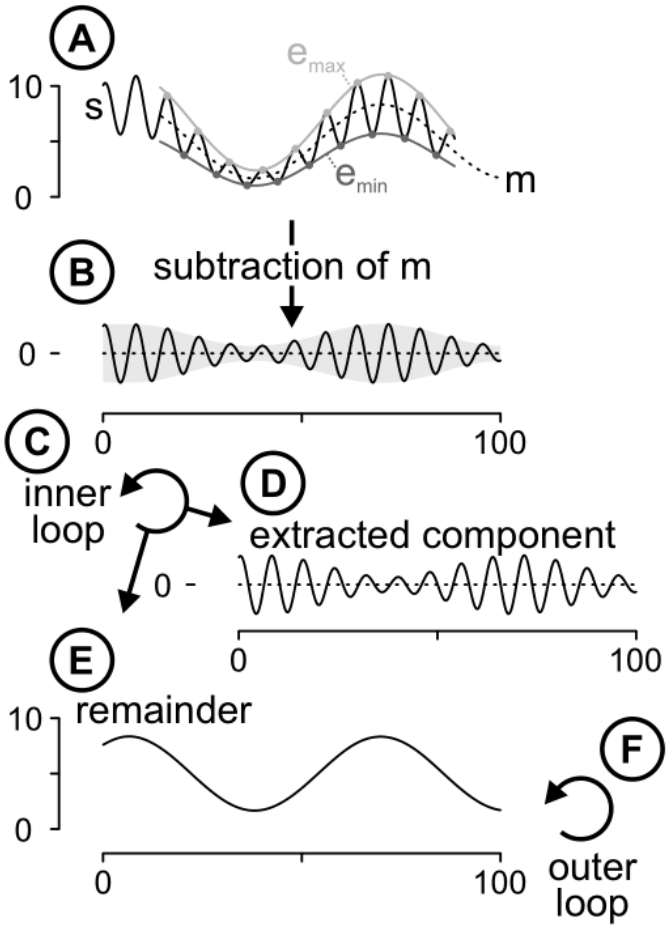


Figure 8: Schematisation of the sifting process in EMD. (A) From the signal [s], the extrema are used to define upper [e_{max}] and lower [e_{min}] envelopes, and the mean of these envelopes [m]. (B) The mean [m] is subtracted from the signal [s] to obtain a first prototype of component. (C) The process is iterated on the prototype component in the inner loop, in order to smooth the component and eliminate riding waves. (D) From the inner loop of sifting, a component is extracted. (E) A remainder is defined as the signal minus the extracted component. (F) The remainder is used to perform a new round of sifting and define the next components. [NOTE TO REVIEWERS: 1.5 column fitting image]

The EMD algorithm defined by Huang et al. (1998) is not perfect. Among others, there can be (I) instabilities near the ends of the signal, (II) asymmetric wave forms, (III) and **mode mixing** (i.e. widely disparate scales for the oscillations of each extracted IMF). Because of these problems EMD has been adapted and improved into a wide variety of algorithms (e.g., Lin et al., 2009; Wu and Huang 2009; Kim and Oh 2016; see Bowman and Lees 2013, Sprovieri et al. 2013 and Cicone et al. 2016 for geological applications). A simple yet substantial improvement was provided by Wu and Huang in 2009 in the form of the ensemble empirical mode decomposition (**EEMD**). As the name suggests, EEMD makes an ensemble of versions (also called realisations) of the signal, each of them having a different **white-noise** realisation added to them, then EMD is performed on each version, and the versions are averaged, component by component, to obtain the final decomposition. EEMD reduces mode mixing, considers multiple possible solutions (that are then all averaged) and minimises subjective intervention

(see Wu and Huang 2009 for more details on how this is achieved). However, the original EEMD algorithm does not solve the instability occurring near the ends of the signal, and its results do not necessarily satisfy the IMF definition (i.e. asymmetry and riding waves may still appear). Furthermore, the added random white noise does not cancel entirely when averaging, and the implementation of an ensemble large enough (i.e. performing hundreds of EMDs) has a high computational cost (tens of minutes to hours on a standard workstation). Still, in practice, EEMD remains a powerful tool (e.g., Bowman and Lees 2013 use it for seismology), especially to minimise mode mixing.

To illustrate how frequency behaves in estimations from decompositions obtained via EMD or EEMD, we compare EMD results to frequency estimations obtained by Fourier transform, in the case of **white noise** (i.e. random and uncorrelated fluctuations, typically sampled from a Gaussian distribution). For the Fourier transform, that reconstructs the random white noise using only rigid sines and cosines, amplitude has to be distributed evenly to the entire range of frequencies (Fig. 9).

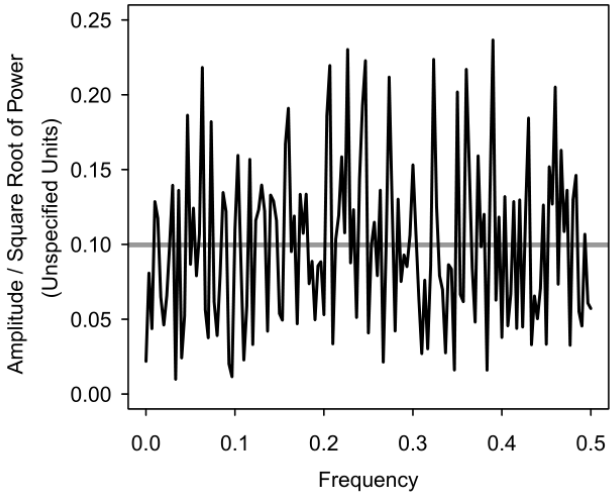


Figure 9: Amplitude against frequency of a direct Fourier transform of Gaussian white noise (standard deviation = 1, N [number of data points] = 300). The horizontal grey line marks the mean of amplitudes. There is no general tendency of amplitude towards any specific range of frequency. [NOTE TO REVIEWERS: 1 column fitting image]

On the other hand, EMD commonly acts on white noise like a **dyadic** filter bank. A filter bank is called dyadic when the mean period approximately doubles for each consecutive filter (Franzke 2009). This is what is observed in Figure 10: the white noise gets distributed into frequency-modulated components, for which amplitude decreases towards lower frequency, with ratios of mean frequencies between successive components of 2 or less (Wu and Huang 2010 and references therein). The behaviour of EMD is actually similar to what is observed in wavelet analysis. If the wavelet transform is parametrised so that power is directly proportional to amplitude, the power of white noise (and thus its amplitude) decreases towards lower frequencies (see Maraun and Kurths 2004,

especially their Fig. 2). Furthermore a dyadic behaviour can be obtained in wavelet analysis by using a scale factor of 2 (Wu and Huang 2010).

The different behaviours of Fourier and EMD with white noise is well reflected by the amount of cross-cancellation. As Fourier decomposition needs to create sinusoids over the entire frequency range available, its sinusoids will cancel each other heavily in order to recreate the random behaviour of noise. Therefore the parsimony of a Fourier decomposition of white noise is very poor. The cross-cancellation in the decomposition obtained via EMD-based methodology is considerably lower (see section 4.2.3 for more details).

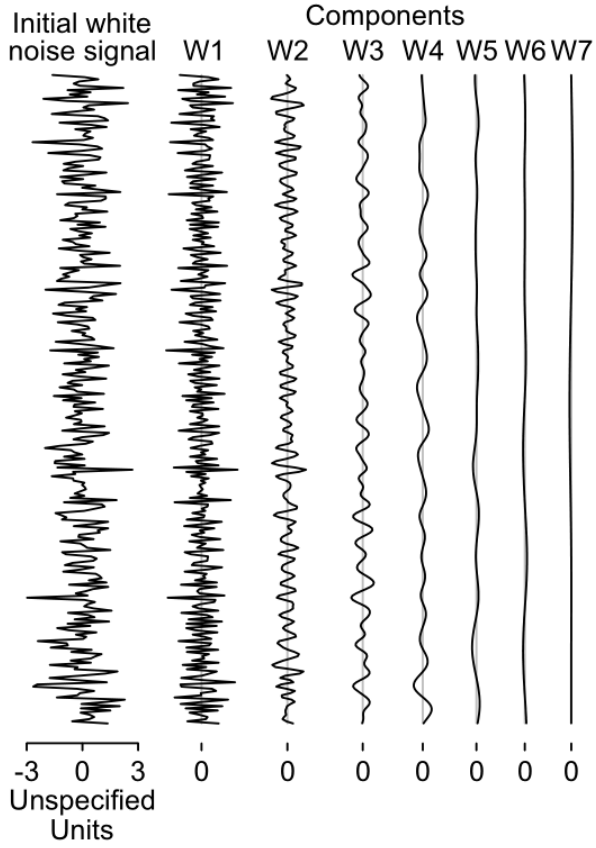


Figure 10: Gaussian white noise (standard deviation = 1, N [number of data points] = 300, identical to the one used in Fig. 9), and its decomposition using the "extricate" algorithm. The intensity scale of the components is identical to that of the initial signal. Note the decrease of amplitude in lower frequency components. The decomposition has less than $\log_2(N)$ components [$\log_2(300) = 8.22$]. [NOTE TO REVIEWERS: 1.5 column fitting image]

EMD, however, does not always act as a dyadic filter bank. By dramatically increasing the sifting number in the EMD algorithm, white noise can be decomposed closer to how a Fourier decomposition would work. Getting EMD to act like that arises from the fact that iterating the sifting smooths the uneven amplitudes of components (Huang et al. 1998). Therefore, the preservation of amplitude features declines as the number of sifting iterations increases (see Fig. 7 of Wu and Huang 2010). With high sifting numbers, EMD can even separate two sinusoidal waves of slightly different frequencies (Huang et al. 1998; Wu and Huang 2010). In that extreme parametrisation

the value of the ratios of mean frequencies between successive components in white noise approaches 1 (Wu and Huang 2010), but such parametrisation defeats the purpose of obtaining a parsimonious and modally meaningful decomposition, with amplitude and frequency modulation intact. In fact, similar results would be more readily obtained via Fourier decomposition.

To further illustrate the behaviour of EMD in mainstream applications (i.e. with low sifting numbers), let us consider harmonics. In Fourier and wavelet analysis, when periodic oscillations diverge from a pure sine wave (or wavelet) form, harmonic peaks appear at frequencies that are multiples of the frequency of the distorted oscillations. This is the way that the Fourier and wavelet transforms account for the non-sinusoidal or non-wavelet-like form of the oscillations (Huang et al. 2009). However, when EMD is parametrised to generate decompositions made of modally meaningful IMFs, harmonics are theoretically irrelevant. In the IMFs of an EMD, the original shape of oscillations can preserved (to a certain degree), and is translated in instantaneous frequency as **intra-wave frequency modulation** (i.e. fluctuations of the instantaneous frequency at a scale lower than the period of the cycle; Huang et al. 2009). The preservation of the shape of the oscillations in the data allows to provide a few modally meaningful components. Typically, the EMD algorithm produces no more than $\log_2(N)$ components, N being the number of data points (Wu and Huang 2010), which is illustrated in Figure 10. Actually, with simpler processes, where the signals are made of sinusoids, and where perturbation is limited to white noise, the simplest Fourier based algorithm (i.e. the periodogram) is enough to meaningfully identify periodic components (Tukey 1967). However, the complexity of geological processes makes real sedimentary records quite more challenging. In most real cyclostratigraphic cases, the intra-wave frequency modulation of astronomical origin would be expressed together with sedimentation rates changes adding their own frequency modulation. This further justifies extracting modally meaningful components and focus on their amplitude features to identify astronomical processes.

The concept of instantaneous frequency further opens the question on how to characterise low frequency cycles, which span the length of the entire record with only a few oscillations. In Fourier transform, such cycles are generally left aside, but as instantaneous frequency provides a detailed overview of how the frequency evolves within the cycles themselves, they could be taken into account and studied in a decomposition. This is exemplified by Wu et al. (2007), which uses EMD to extract and interpret trends and low-frequency components of short-scale data. This highlights that even trends can be recognised and interpreted in a decomposition. But as a trend is the consequence of the limited length of time series, it should be kept in mind that "one [person]'s 'trend' can be

another's 'cycle'" (Stock and Watson 1988). To address the problem in the case of cyclostratigraphy, we introduce a few definitions:

- **Cyclically-connoted residual**: a monotonic function having at most one extremum, and assumed to be centred on either zero (symmetric cyclically-connoted residual) or on a non-zero value (asymmetric cyclically-connoted residual) if the data set was of infinite length. If symmetric, the cyclically connoted residual can be studied using instantaneous frequency, and if asymmetric it can still be compared to reference curves of the interpreted cycle.

- **Linear trend**: a linear function, completing a decomposition made otherwise of several IMFs, and of a cyclically connoted residual.

- **Nonlinear trend**: a monotonic function in which there can be at most one extremum, and on which no assumption is made about its mean. This can be obtained as the sum of an EMD's residual that cannot be interpreted as cyclical, and a linear trend removed before the EMD, or more simply as the residual of an EMD on a signal that was not detrended.

The choice of having a nonlinear trend to put away the awkwardly-interpretable remainders of an EMD, or decomposing it into cyclically-connoted residual and linear trend, is a matter interpretation.

A final problem for EMD is that even if its concept allows getting closer to parsimonious and modally meaningful decompositions, it can fail to fit certain realities of geology. Hiatuses (i.e. losses of sedimentation that result in a part of the signal simply being erased), typically, cannot be conceptualised without sudden phase changes. The issue is that such phase changes could be translated in a decomposition as riding waves, and that the EMD process is specifically designed to remove the riding waves as best as possible, thereby removing the indications of hiatuses. However, this is a problem of the EMD algorithm, and not of the entire idea of parsimonious and modally meaningful decomposition, in which gaps can be included.

4 Methods

We present in this section the practical details of the tools we developed. The software used in this study to generate and handle decompositions is implemented in the newly developed R package DecomposeR (https://CRAN.R-project.org/package=DecomposeR). The code available in the supplementary material for the processing of data is based on DecomposeR version 1.0.4, working with R 4.0 or above.

4.1 The "extricate" algorithm

To use the EMD concept in cyclostratigraphy, ideally the algorithm would need to produce parsimonious and modally meaningful components, while having to specify as few parameters as possible. This can be done using EEMD (Wu and Huang 2009), but it is a slow algorithm. To overcome this problem, we present "extricate", a version of EEMD optimised for short to intermediary length time-series (typically less than 10 000 data points).

EEMD is an ensemble: it iterates EMDs on different white-noise-added realisations of a signal. These iterations usually mean that there are a lot of nested loops in the algorithm: the sifting inner loop is repeated in the sifting outer loop (Fig. 6), which is itself repeated for each realisation in the EEMD. As a consequence, the typical EEMD algorithm is computationally inefficient, and runs slowly. If somehow the ensemble of noise-added signals were to be treated in one single EMD, the process could be made to be more efficient. This can be achieved by joining the realisations of white-noise-added signals, uniting them into one large signal. Then, there would only be the matter of optimizing the algorithm for a large data set. However, the joining step is problematic in itself: it creates spurious oscillations at the location of the junction of repeated signals, oscillations which will affect the resulting decomposition. This can be circumvented using an accepted method for joining repeated signals, which was initially designed to answer the problem of instability near the ends of the signal: even-odd extension (Zeng and He 2004). Even extension (also called mirror extension) is extending data with the mirror of the data (i.e. using axial symmetry; Fig. 11). Odd extension (or anti-mirror extension) is identical, but extends the data with a negative intensity version of the mirror of the data (i.e. using central symmetry; Fig. 11). The principle of even-odd extension is to perform EMD on the signals with even and odd extension separately, and to average the resulting decomposition in order to reduce instability at the end of the signal. This principle can be used to join any ensemble in two sets: even and odd (Fig. 11). Based on these concepts, we can now explain the "extricate" algorithm: the linear trend in the signal is first removed prior to joining the signals, in order to avoid extreme behaviours at the junction between anti-mirrored signals. Then, EMD is performed on the two sets, joined using even and odd extension respectively. Next, each individual signal in the resulting decompositions of the two sets is rotated back as it was originally, and the two sets are averaged into one set. Finally, the joined signals in the resulting set (which

is an average of the even-odd sets) are disjoined and averaged. To have all the signals undergo similar procedures, the first and last signals of the sets are not considered.

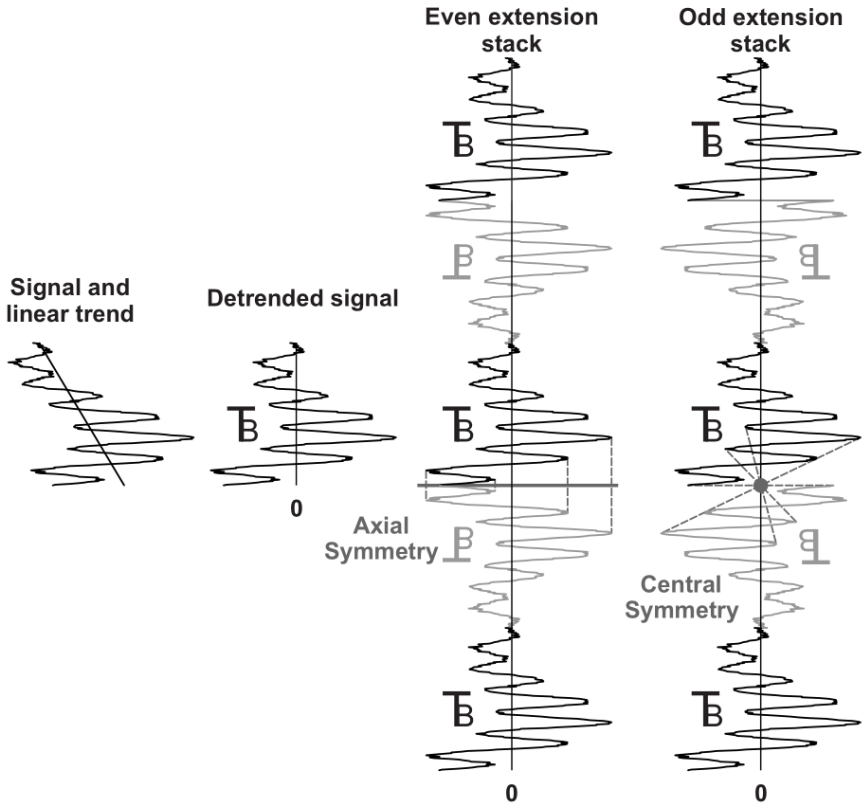


Figure 11: Illustration of the even and odd pasting used in "extricate". The signal is detrended, and then joined using even and odd extension to respectively make an even set and an odd set. EMD is performed on the two sets. The resulting decompositions of each repeated signals are put back in the original orientation of the data (cancelling the even odd extension) and the two sets are averaged into one. Then all the different realisations in the averaged set are themselves averaged to make unique decompositions.
[NOTE TO REVIEWERS: 1.5 column fitting image]

In essence, even- odd extension is a way of predicting the behaviour of data outside of its studied range. The method uses the data itself for the prediction in order to resolve instability at the signal end. However, there is no way to accurately predict inexistent data, so by nature the even-odd extension is not failproof. This was underlined by Wu and Huang (2009), who have proposed a method reducing instability at the ends without relying on prediction, which would be analytically better than the even-odd extension used in "extricate". On the other hand, the even-odd extension is what makes extricate efficient. For all intents, the "extricate" algorithm should be thought of as a trick allowing easy experimentation on the EEMD concept, not as a panacea. To make "extricate" efficient, all the computational elements used to perform the sifting process have been optimised for large data sets. Another feature of "extricate" is that the white noise that is added to the even set can be itself mirrored and used in the odd set in such a way that when the two sets are averaged, the noise cancels exactly, except for rounding errors inherent to the representation of real numbers on a computer.

The "extricate" algorithm can be set to average the ensemble by groups of signals, producing different replicates of a decomposition. The only difference between these replicates is the white-noise realisations used in their respective ensemble, but as they are all sampled from the same process, the white noise is expected to have the same variance in every replicate. Because the decomposition parameters are similar for all replicates, we can expect them to have similar properties; Wu and Huang (2009) highlighted this, showing that signals being different only in the first 10% of data make similar decompositions for the 90% of data that is similar, in the case where parameters were identical (especially the number of sifting iterations for each component). To have a comprehensive exploration of the different decompositions that can be obtained via "extricate", the sensitivity of the decomposition results to the parametrisation needs to be evaluated. The parameters that affect how the decomposition is performed, and influence the results, are:

- The number of realisations combined to be averaged into one replicate.
- The sifting number applied for each component.
- The intensity of the white noise added to the signals.

A comprehensive study of the influence of these parameters is beyond the scope of this paper. There are common recommendations in the literature to adjust these parameters (Wu and Huang 2009; Stallone et al., 2020), but these recommendations can stem from the need of each author to give certain characteristics to a given decomposition. For instance, Wu and Huang (2009; 2010) recommend fixing the sifting number to a low number in the EMD (typically 10) to optimise the dyadic behaviour for noise while keeping the upper and lower envelopes of IMFs almost symmetric around zero. However, even a sifting number as low as 10 can smooth the amplitude variations compared to the ones observed in EMD performed with a sifting number of 1 (i.e. avoiding the repetition of the inner loop in the EMD). Amplitude variations are a major characteristic of the astronomical signal, especially in the form of modally meaningful components, therefore such smoothing can be problematic. We therefore recommend a trial-and-error approach for exploring the potential of EMD in cyclostratigraphy. To this end, we suggest that quality criterions for decompositions should first be defined and tested for the specific case of cyclostratigraphy, before setting any guideline.

4.2 Quality assessment of a decomposition

Decomposition offers the opportunity to interpret any wiggle present in a signal in a format that can be checked *a posteriori*. Decomposition allows the user to visualise the impact of noise and other non-astronomical processes, in contrast with what is interpreted to be of astronomical origin. On the other hand, the components linked to astronomical processes can be validated by comparison to the astronomical models or to

contemporaneous geological sections, or by putting them back in their geological context (e.g., other sedimentological observations or proxies in the same section). A decomposition in itself should be checked for quality, and we offer the concepts explained in the following sections to this end. An automated testing using these concepts is provided in the DecomposeR package, through the check.emd() function.

4.2.1 Reversibility

To preserve a way of comparing a decomposition to the signal it originates from, we encourage **reversibility**, which we define here as providing the output of the signal processing in a data form that allows to reconstruct the data in the form it had in the input. This concept is a generalisation of the mathematical invertibility, which applies to mathematical functions like the decomposition, where for instance the reconstruction process allows to mathematically reconstitute the initial signal from the components of its decomposition. Reversibility extends invertibility to the practical implementation of the mathematical functions: for decomposition, reversibility typically means preserving the original data points in the signal processing, even after linear interpolation and depth distortion applied for tuning purposes.

Linear interpolation (along the time-depth-height axis) is used to have regular sampling, which is necessary when computing the Fourier transform via the fast Fourier transform algorithm (**FFT**, used among others in filtering procedures or in the Hilbert transform to compute instantaneous frequency). Indeed, sampling geological sections and cores in a perfectly regular manner is highly difficult; linear interpolation is therefore often a necessity in cyclostratigraphic data analysis.

To make a regular linear interpolation and still include all the original data points, the interval for resampling can be set as the lowest interval for depth or height that was considered in practice (i.e. the interval marking the spatial precision level for sampling and logging, typically 1 centimetre for high-resolution field logging). We will refer to this minimal interval on the height/depth measurement as **stratigraphic quanta**. Even when performing depth distortion for tuning purposes, which can change the size of the relative intervals between samples, every sample position can be rounded to the scale of a temporal quanta interval, which would be chosen to minimise information loss when used as a sampling interval for linear interpolation. In practice, the stratigraphic quanta can be computed as the greatest common rational divisor of the dataset depth or height values, allowing to perform regular linear interpolation without the user having to manually set the sampling interval (see Fig 12A). This is the purpose of the "respace" function in DecomposeR: it linearly interpolates by the greatest common rational divisor or a divisor thereof (which can be defined by the user), and identifies the original data points against the interpolated ones. The greatest common rational divisor as interval for interpolation is only used here

to preserve the original data points for decomposition, filtering and instantaneous frequency computation. For other time-series methods where interpolation is subject to other problems, other procedures may be considered (see for instance section 2.2 in Laepple and Huybers 2013, and appendix A of Martinez et al. 2016). Noteworthily, obtaining the greatest common rational divisor remains, to date, a computational challenge (Wouters et al. 2021).

4.2.2 Integrity

For a set of time-series (e.g., filtered signals) to qualify as a decomposition, the sum of all these time-series should flawlessly reconstruct the signal. To quantify this property, we define integrity (Fig. 12B), which is the mean of the absolute difference between (I) the signal, and (II) the summed components of the decomposition:

$$\text{integrity} = \frac{1}{N}\sum_{n=1}^{N}\left|S(dt) - \sum_{k=1}^{K} C_k(dt)\right| \qquad (2)$$

where S(dt) is the initial signal, C_k(dt) represents each K^{th} component, and N is the number of data points. This should be as close to zero as floating-decimal arithmetic permits (typically 15 orders of scale lower than the highest order of scale of intensity). In classical EMD and with "extricate", which are both designed to allow virtually perfect integrity, this is easily achievable. However, when considering other algorithms and methods, which can fail to have integrity by design (such as classical EEMD), or when checking for mistakes, this is a useful metric.

4.2.3 Parsimony

To quantify the cross-cancellation phenomenon in decompositions, we introduce a second quantified quality metric: parsimony (Fig. 12C). Parsimony is defined as the ratio between (I) the cumulated absolute intensities of each component (except the linear trend), and (II) the cumulated absolute intensities of the signal (minus the linear trend):

$$\text{parsimony} = \frac{\sum_{k=1}^{K}\sum_{n=1}^{N}|C_k(dt)|}{\sum_{n=1}^{N}|S(dt) - T(dt)|} \qquad (3)$$

where S(dt) is the initial signal, T(dt) is its linear trend, C_k(dt) represents each K^{th} component except the linear trend, and N is the number of data points. The linear trend should be subtracted in the calculation, because an added trend decreases the parsimony estimation of a similar decomposition. Our choice of the linear trend is arbitrary, to allow for uniformity, but a nonlinear trend or the mean could be used as well.

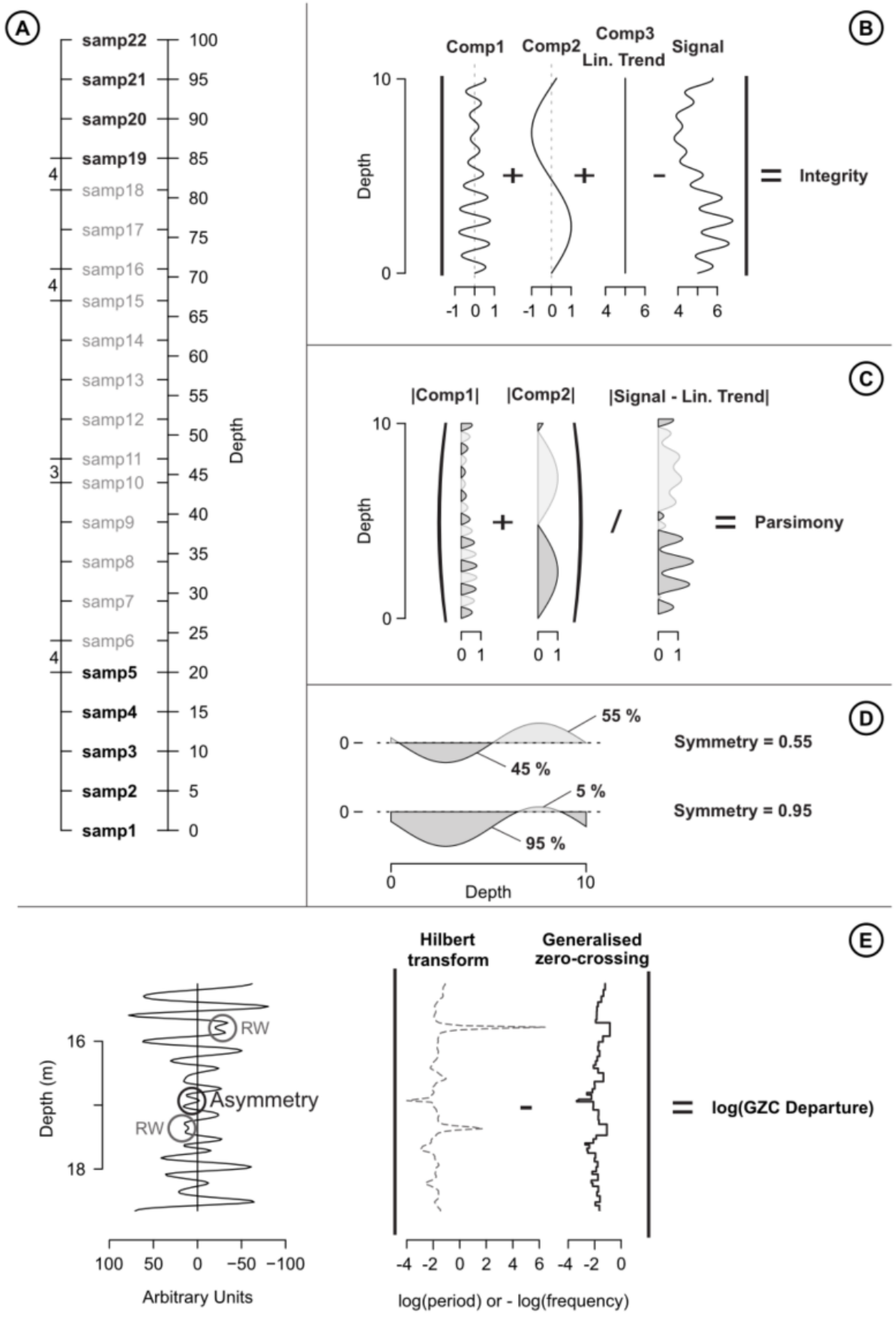

A
samp22
samp21
samp20
samp19
samp18
samp17
samp16
samp15
samp14
samp13
samp12
samp11
samp10
samp9
samp8
samp7
samp6
samp5
samp4
samp3
samp2
samp1
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
Depth
4
4
3
4
B
Comp1
Comp2
Comp3
Lin. Trend
Signal
Depth
10
0
Integrity
-1 0 1
-1 0 1
4 6
4 6
C
|Comp1|
|Comp2|
|Signal - Lin. Trend|
Depth
10
0
Parsimony
0 1
0 1
0 1
D
55 %
45 %
5 %
95 %
Symmetry = 0.55
Symmetry = 0.95
0
0
Depth
10
E
Hilbert transform
Generalised zero-crossing
log(GZC Departure)
-4 -2 0 2 4 6
-4 -2 0
log(period) or - log(frequency)
Depth (m)
16
18
RW
Asymmetry
RW
100 50 0 −50 −100
Arbitrary Units

Figure 12

Figure 12: Illustration of metrics and concepts for assessing the quality of a decomposition. (A) Due to uneven sampling rate, a linear interpolation with a sampling rate of 5 depth units will exclude some of the original data points. To have a reversible resampling including all the original samples, the sampling rate has to be set as a divisor of the greatest common rational divisor, in this case 1 depth unit. (B) Illustration of integrity: all components are summed, and subtracted by the initial signal, which should ideally cancel out. The difference is in absolute values. (C) Illustration of parsimony: the same decomposition than (B) is used. The components except the trend are put in absolute values and summed, and divided by the absolute values of the initial signal minus its trend. (D) Illustration of the symmetry metric, which is the maximum amplitude proportion of the sides of a component. (E) Illustration of GZC departure, which is based on the absolute difference between the logarithm of the real instantaneous frequency (obtained via Hilbert transform for instance) and the logarithm of the generalised zero-crossing instantaneous frequency, which is used as reference. The computation of GZC is done by simplifying any component into a suite of extrema and zero-crossings, where each set of successive extrema not separated by a zero-crossing are grouped into one extremum. (B, C and D) Integrity, parsimony and GZC departure are computed by default for the entire record or for a complete component, but can be made to be instantaneous. [NOTE TO REVIEWERS: 2 column fitting image]

If each sinusoid obtained via Fourier transform is treated like a component of a decomposition, high parsimony values are observed, because (I) the presence of noise or interferences, which necessitate a lot of cross-cancellation to model via sinusoids, and (II) the size of time-series, which is equal to the number of sinusoids used to model it; numerous sinusoids increase the opportunities for cross-cancellation to take place. For the 300-samples-long white-noise example of Fig. 9 and 10, the parsimony of decomposition via Fourier transform is of 12.26 (i.e. there is an excess of 1126 % of absolute intensity in the decomposition, compared to the original signal), whereas the parsimony of the decomposition via "extricate" is of 1.74 (74% excess of absolute intensity). This is an important aspect of decomposition, as high values of excess absolute intensity indicate a proportionally high likelihood of having added wiggles not present in the signal to the decomposition.

Our parsimony metric increases with the number of components that are needed to describe the entire signal. Parsimony will also vary depending on how wiggles of different frequency ranges populate the signal. This makes parsimony in its current form irrelevant to compare decompositions of different signals. Additionally, parsimony can widely change depending on the trend chosen for the decomposition. Therefore, the parsimony metric is only suited to compare different decompositions of a same signal, with an identical trend.

4.2.4 Symmetry

This metric quantifies roughly whether the positive and negative parts of a components are symmetrical around zero: it is defined as the highest proportion of integrated intensity (i.e. the area between the component's curve and the zero line) on either of the two sides (see Fig. 12D). It varies between 0.5 (perfectly symmetric) and 1 (either purely negative or purely positive). This can be used as a rough estimate whether instantaneous frequency can be determined reliably, especially on low-frequency components such as cyclically-connoted residuals.

4.2.5 Generalized Zero-Crossing (GZC) departure

To compute the instantaneous frequency of a decomposition, the components should be IMFs. However, current algorithms optimised to generate meaningful decompositions sometimes generate components that locally deviate from the IMF definition. To still be able to work with such algorithms, despite their shortcomings, we suggest having a way to quantify the reliability of the instantaneous frequency computation of each component. Such a metric allows to identify potential problems and their location. To that effect we introduce a third quantified quality metric: GZC departure (Fig. 12E). We define GZC departure as the exponential of the mean of the absolute differences of the logarithms of frequencies obtained using (I) a robust generalised zero-crossing method and (II) a truly instantaneous method such as the Hilbert Transform (or any other method used to compute instantaneous frequency):

$$\text{GZC departure} = \exp\left(\frac{1}{\mathrm{N}}\sum_{\mathrm{n}=1}^{\mathrm{N}}\left|\log\left(\mathrm{IF}_{\mathrm{GZC}}(dt)\right) - \log\left(\mathrm{IF}_{\mathrm{inst}}(dt)\right)\right|\right) \qquad (4)$$

where $\mathrm{IF}_{\mathrm{GZC}}$ is the frequency obtained by GZC, $\mathrm{IF}_{\mathrm{inst}}$ the one obtained by an instantaneous method such as HT, and N is the number of data points. To ensure that the GZC computation is as robust as possible, the GZC method should be coupled with an algorithm which simplifies any component into a suite of extrema and zero-crossings; any ensemble of successive extrema not separated by a zero-crossing should be simplified into one single extremum. In practice GZC departure is a factor of the mean error of frequency. For a component having a mean period of 23 kyr, a GZC departure of 1.13 defines an average error interval of 13% above and below 23 kyr, i.e. between 20.4 and 26 kyr.

4.3 &Instantaneous spectral analysis for cyclostratigraphy

We present novel visualisation tools, which are meant to provide a general overview of the behaviour of instantaneous frequency and amplitude. They are based on the distribution of instantaneous frequency, amplitude, and/or frequency ratio of the original data points for each component. The distributions are presented in the form of density plots, or of histograms, and can integrate several realisations of identically parametrised decompositions, which correspond to the replicates computed by "extricate".

Instantaneous frequency is independent of amplitude. Therefore, instantaneous frequency can be visualised alone in a histogram to show the distribution of frequency or of period (e.g., Fig. 13A, based on the signal of Fig. 1). This allows to estimate the variability of frequency in a given component, but can also help to identify potentially spurious frequencies coming from predicted wiggles. Indeed, even if the decomposition methods are optimised to reduce cross-cancellation, predicted wiggles are still a possibility. Typical predicted

wiggles coming from EEMD methodologies are wiggles of very low amplitude (e.g., in group 1 of Fig. 1 and 13). A first check for predicted wiggles can be done in the period distribution histogram: if the period distribution of two separate components overlap, it may turn out that the two components can be merged into a single meaningful or modally meaningful component. Amplitude should also be inspected to detect typical EEMD predicted wiggles, by their low amplitude. This can be done in the spectral population plot (e.g., Fig 13B), which we present here as a bivariate distribution plot of instantaneous amplitude against instantaneous frequency.

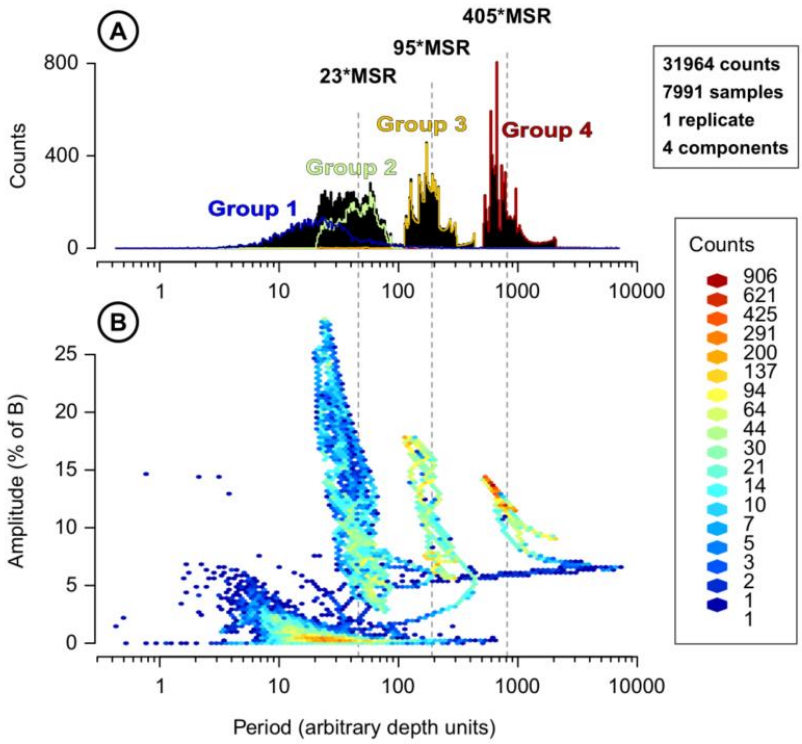


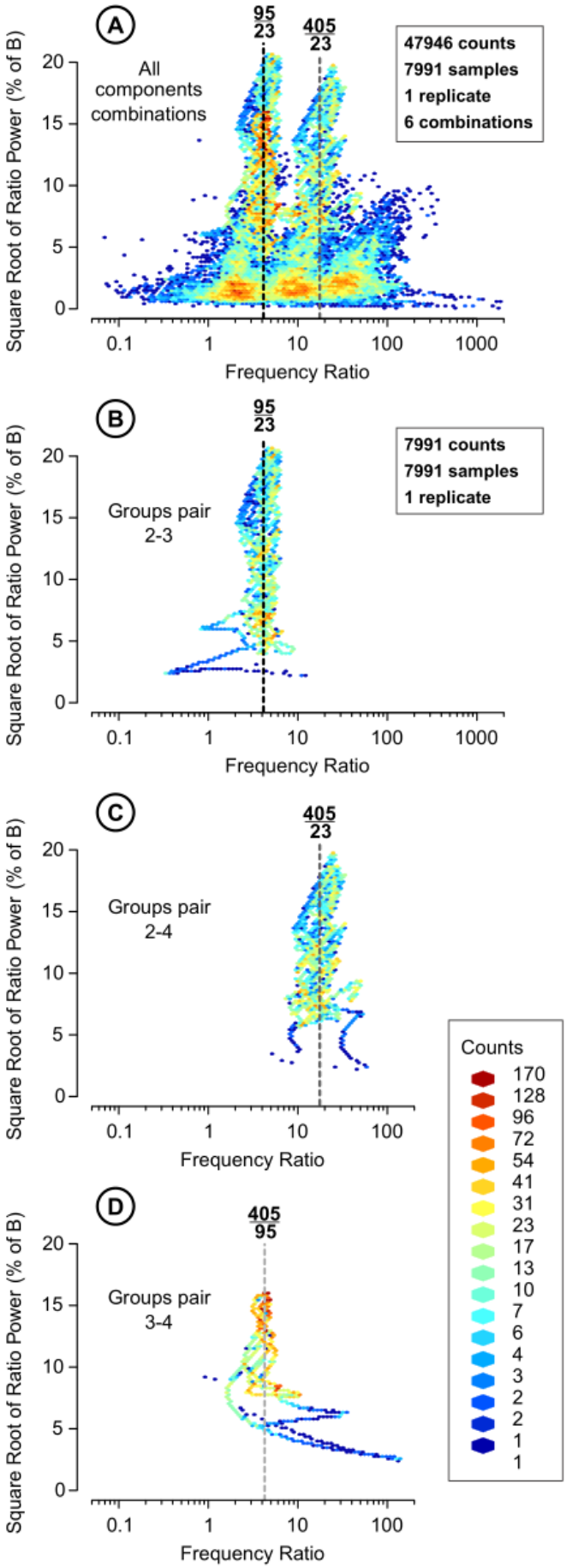
Figure 13: (A) Period distribution and (B) spectral population of the grouped components of the depth-distorted artificial stratigraphic signal used in Figure 1. The inverse correlation between the amplitude and the logarithm of the period comes from the decreasing sedimentation rate where the element B making up the signal is highest (see Fig. 1). The residual and the trend are considered as nonlinear trends, i.e. not usable in spectral analysis [NOTE TO REVIEWERS: colour online only; 1.5 column fitting image]

Observing the behaviour of frequency in relation to other parameters, for instance amplitude (Fig. 13B, see also caption of Fig. 13), is useful to gather information on how the sedimentation rate changes (see also the discussion related to Fig. SF7 in the supplementary materials). But to further interpret plots based on frequency in terms of astronomical cycles, one needs to provide assumptions about the sedimentation rate; this, we have shown, is not necessarily a valid strategy. Assumptions on sedimentation rate changes can be bypassed, through the identification and characterisation of specific frequency ratios potentially linked to astronomical cycles, either from IMFs or from their demodulation. Modally meaningful decompositions made of IMFs seem to be an ideal framework to study frequency ratios, because instantaneous frequency ratio can be determined for each pair of IMF of a decomposition, which would be representative of the entire behaviour of the frequency ratios in a given

signal. Instantaneous frequency ratio (Fig. 14) enables this: the ratios are obtained by taking the instantaneous frequency of each combination of pairs of components, data point by data point. This generates, for a decomposition having an M number of components, $\sum_{k=1}^{M-1}(k)$ pairs of components (or, in another formulation, $m(m-1)/2$ pairs of components).

In contrast to period distribution and spectral population plots where components are naturally distributed from high-frequency to low-frequency, the distribution of frequency ratio by pairs of components is less intuitive. This, in turn, makes it difficult to discern the patterns of ratios of frequencies in a mere histogram. Therefore, we rather visualise frequency ratio as a bivariate density plot of the square root of ratio power (in other words the square root of the instantaneous amplitudes of the two components multiplied together) against the instantaneous frequency ratio; we define this plot as a ratio population plot (e.g., Fig 14A). The square root of ratio power is a way to discriminate the relevant ratios based on amplitude. Selected component pairs can further be isolated in separate ratio population plots for better visualisation (e.g., Fig 14 B, C and D).

If frequency ratios of untuned data were associated with an appropriate testing of the noise continuum, they could help determine whether a stable frequency ratio is present in the signal. A stable frequency ratio can serve as proof for the existence of astronomical cycles of frequencies not known *a priori* (e.g. in geological times too ancient to have an accurate estimation of all the astronomical frequencies). If the value of the stable frequency ratio uniquely corresponds to a known pair of astronomical cycles, it can serve to connect the cycles in the signal to their astronomical counterparts (see also propositions 1 and 2 of Meyers, 2019).



970 **Figure 14**

Figure 14: (A) Ratio population for all pairs of components of the grouped components of the depth-distorted artificial stratigraphic signal used in Figure 1, and (B, C and D) for selected grouped components pairs, with the expected ratios of frequencies indicated by vertical lines. [NOTE TO REVIEWERS: colour online only; 1.5 column fitting image]

Although period distribution, spectral population and ratio population are all distribution plots, their statistical interpretation has a few peculiarities. The first of these peculiarities is that amplitude modulation will be expressed in the spectral and ratio population plots by spreading the distribution of the data points along amplitude or square root of ratio power (e.g., Fig. 13B), which comes from the fact that every component has the same number of data points. For the ratio analysis, the decomposition method used should be tested whether it has any bias towards generating specific frequency ratios between components. As mentioned earlier, the EMD and EEMD algorithms tend to work as dyadic bank filters: each pair of successive components will approach a frequency ratio of 2. White noise is characterised by the ratios in pairs of successive components (pairs of components 1-2, 2-3, 3-4, 4-5 and 5-6 of the white noise used in Figures 9 and 10) are roughly centred around 2 (Fig. 15B). The behaviour of the whole ratio spectrum of white noise (Fig. 15A) is to have the square root of ratio power peaking around a ratio of two and steadily decreasing towards higher ratios. Even when various environmental and sedimentary processes enter the signal, the dyadic behaviour of modally meaningful decomposition methods such as EMD and EEMD tends to persist. We further exemplify this in section 8 of the supplementary materials, where the modally meaningful decomposition of low-frequency dominated noise processes, which are more typical of climatic noise, behaves with the same centring on a frequency ratio of two. This causes a risks of overinterpretation, and could be unfortunate considering that the frequency ratio of precession (ca. 20 kyr period) and obliquity (ca. 40 kyr) is close to two. Such frequency ratio should only be considered relevant when there are palpable cycles to begin with, with amplitudes strong enough to be distinguished from the continuum. This is the rationale behind using the square root of ratio power: it is a good indicator that the cycles with a specific ratio are evident in the signal.

We finally note that in the spectral population and ratio population plots, the units of amplitude and square root of ratio power are identical to the units of intensity of the time series (i.e. the physical units of the proxies used). This should ease the intuitive understanding of the plots.

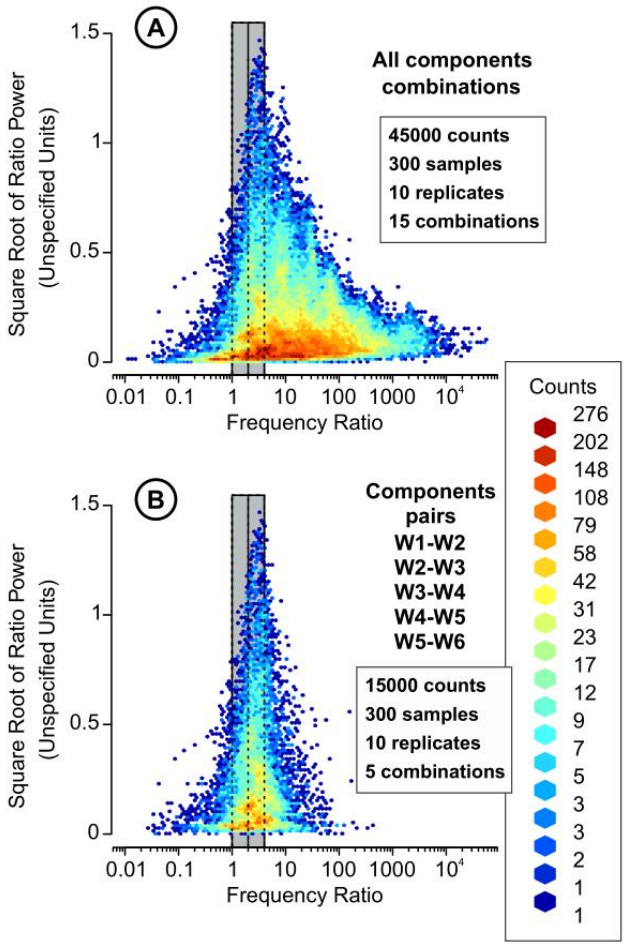


Figure 15: (A) Ratio population for all pairs of components of white noise and (B) for only the successive components. The white noise is the same that the one used for Fig. 9 and 10 (ensemble of 10 decompositions with identical parameters but with different white noise realisations). The grey rectangle delimits frequency ratios between 1 and 4, and is centred on a ratio of 2. [NOTE TO REVIEWERS: colour online only; 1.5 column fitting image]

4.4 Instantaneous frequency computation

The normalisation scheme of Huang et al. (2009) has been implemented in the DecomposeR package to perform instantaneous frequency computation. With the normalisation procedure, the amplitude is determined via spline-fitting of the maxima of the absolute values of the signal (Huang et al. 2009). The amplitude is then used to compute a frequency carrier; this is the basis of instantaneous frequency computation by normalised Hilbert transform (NHT) and direct quadrature (DQ) (see Fig. 7 and Huang et al., 2009). However we discovered that the spline-fitting algorithm provided by Huang et al. (2009) for normalisation can be unstable; in highly amplitude-modulated components, the spline fitting can give negative amplitude values or dramatic undershoots, making the algorithm fail or give unrealistic values. Therefore, we prefer to use the Hilbert transform (HT; see section 3 and Huang et al., 2009) for routine and automated instantaneous frequency determination, as the algorithm is currently sturdier.

5 Case study

5.1 Presentation

The ODP 926 cores were obtained from the Ceara Rise (Shipboard Scientific Party 1995a), a bathymetric high situated in the western equatorial Atlantic, ca. 900 kilometres North-East of the Amazon river mouth (see Fig. 1 in Zeeden et al., 2013). The cores are mainly composed of calcium carbonate that has a marine pelagic origin, and of terrigenous material, attributed to the Amazon river outflow (Shipboard Scientific Party 1995b). We will focus here on proxy data for terrigenous material against calcium carbonate (colour reflectance and magnetic susceptibility), obtained from the interval estimated to span between 9000 and 5000 **ka** (thousands of years *ago*) (Fig. 16).

Three sites, 926 A, B and C, were drilled during the Ocean Drilling Program (ODP) Leg 154 (Shipboard Scientific Party 1995a). Magnetic susceptibility and colour reflectance were measured on board, and used to splice the cores of all three sites onto an identical composite depth scale (Shipboard Scientific Party 1995b). This first depth scale preserves the original length of all cores, and only depth offsets to each core are used to overlap the magnetic susceptibility and colour reflectance features; there is no stretching or compression of the depth scales of each separate core (Shipboard Scientific Party 1995b).

A first cyclostratigraphic study was performed by Shackleton and Crowhurst in 1997; they provided a tuning interpretation, and on the tuned signal, they applied complex demodulation on the magnetic susceptibility signal to extract the amplitude modulation of climatic precession. New, higher-resolution measurements of greyscale and magnetic susceptibility allowed Zeeden et al. (2013) to revisit the work of Shackleton and Crowhurst (1997). Zeeden et al. (2013) provided a revised composite depth scale (which did not change too much in the interval that concerns this study), an updated tuning interpretation, and an amplitude demodulation of the colour reflectance (greyscale) data. This amplitude demodulation was performed on the signal tuned to the new tuning interpretation to extract the amplitude of climatic precession.

By applying our "extricate" EEMD algorithm on the greyscale data of ODP 926 between 9000 and 5000 ka, we intent to show that a modally meaningful and parsimonious decomposition can be obtained via our methodology. We will further attempt to give an interpretation for all the components in the decomposition. To give some further perspective, the analysis of the magnetic susceptibility of site 926, in the same 9000 and 5000 ka interval, is provided in the supplementary material. Colour reflectance and magnetic susceptibility are in general inversely correlated to each other (Fig. 16; see also Shipboard Scientific Party 1995b). For the magnetic susceptibility signal, the tuning interpretation of Wilkens et al. (2017) will be used. We mention that the depth

scales for the greyscale signal (presented by Zeeden et al. 2013) and the depth scales for the magnetic susceptibility (presented by Wilkens et al. 2017), are offset by 0.385 meters in average: this value was obtained using correlation coefficients of incrementally lagged records.

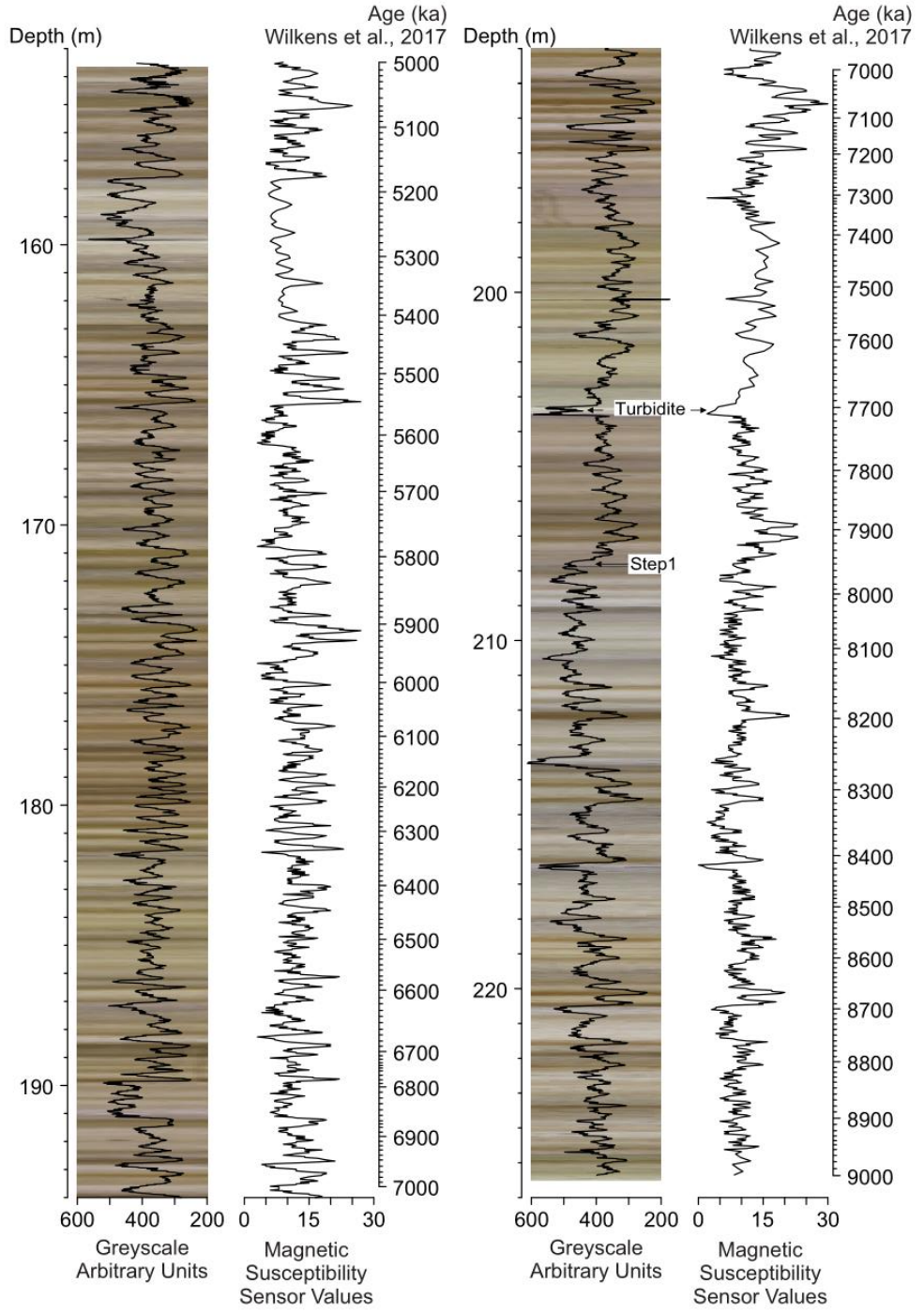


Figure 16: Greyscale (overlaying photographical core scans) and magnetic susceptibility signals between 9000 and 5000 ka in ODP 926. The greyscale axis is inverted, showing the inverse correlation with magnetic susceptibility. The data comes from Zeeden et al. (2013) and Wilkens et al. (2017), the time scale shown here is based on the tuning of magnetic susceptibility performed by Wilkens et al. (2017). [NOTE TO REVIEWERS: colour online only, 2 columns fitting image]

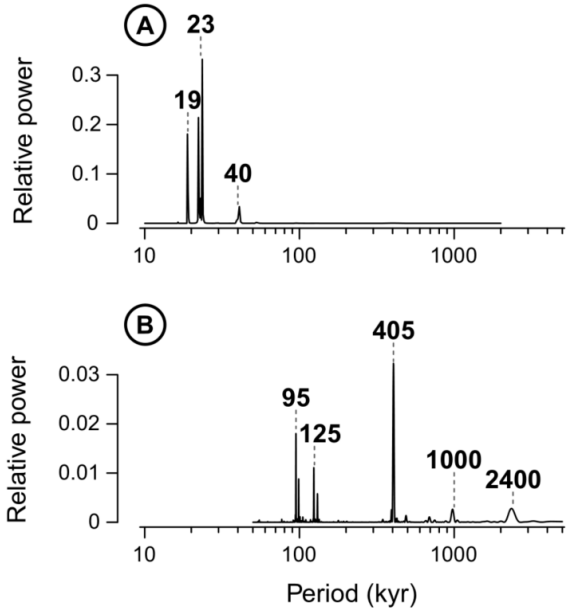


Figure 17: Estimation of the expected frequencies in the astronomical signal used to make the CIP1 artificial signal. (A) Periodogram of the 21 July insolation after Laskar et al. (2004) at 55°N between 8000 and 6000 ka used as basis for making the CIP1 artificial signal. The periodogram shows climatic precession-related peaks (main periods of 19 and 23 kyr) and a fainter influence of obliquity (period of 40 kyr) (Laskar et al. 2004). (B) Periodogram of the demodulation via direct normalisation (Huang et al., 2009) of the climatic precession solution (Laskar et al., 2004), showing the main climatic precession amplitude modulators, that are expected to be observed in the CIP1 signal: short eccentricity (main periods of 95 and 125 kyr), long eccentricity (period of 405 kyr), and 2 faint peaks of ca. 1000 kyr and 2400 kyr periods which are respectively attributed to the $g_1 - g_5$ and $g_4 - g_3$ component (Olsen and Kent 1999; Hinnov 2000; Laskar et al. 2004; 2011). The climatic precession solution was extended between 20 000 and 0 ka to better resolve the low frequency cycles. [NOTE TO REVIEWERS: 1 column fitting image]

We also analyse, still in the supplementary material, an artificial signal (the first case of the Cyclostratigraphic Intercomparison Project, CIP1; for more details see supplementary material, and Sinnesael et al. 2019). This signal is based on an insolation curve computed between 8000 and 6000 ka; an interval encompassed in the ODP 926 signals that we study here. With this artificial signal, we should be able to link, more easily than in a real signal, the processes used to generate the signal to their expression in the decomposition. Typically, the frequencies expected to occur are known from the astronomical solutions used to generate the signal (Fig. 17). These frequencies are also the ones expected for the signals at ODP 926, only with different relative power values. The data used is available in the DecomposerR package.

5.2 Results

The greyscale signal of ODP 926 (Fig. 16) is decomposed into 10 components using "extricate" (Figs. 18 and 19): GSC (for Grey Scale Component) 1 through 9, plus a linear trend. Running "extricate" on the signal took 10 seconds on a standard workstation. We make the assumption here that GSC9 is a cyclically-connoted residual. The decomposition results from averaging 500 realisations, all having different uniform white noise realisations added to them, contained within ± 75 greyscale arbitrary units. For comparison, the signal values are ranging

between ca. 170 and 610 greyscale arbitrary units. The amplitude of white noise was chosen to be of roughly equivalent amplitude compared to the variations in the signal, in order to induce a change in extrema in each realisation (see Wu and Huang, 2009). The adequate level of white noise is chosen based on the mode mixing that is visually observed in the signal decomposition. The sifting is not iterated (i.e. sifting number = 1) to reduce the smoothing of the amplitude of the components. This setting is meant to preserve the amplitude features. We only make one replicate of the decomposition. The decomposition is performed on the signal in the stratigraphic domain (i.e. the depth axis). We compute the instantaneous frequency in the time and in the stratigraphic domain (Fig. 20), and the instantaneous frequency ratio in the stratigraphic domain only (Fig. 21). The amplitude of GSC3 (Figs. 18 and 19) is determined using the normalisation scheme of Huang et al. (2009), and is used to further study the cyclicity in the amplitude modulation.

The parameters for the extricate algorithm were adjusted to reach two main objectives: optimising parsimony, and centring one component (GSC3) on an astronomical target curve that was found by Zeeden et al. (2013) to be well expressed in the signal; the expression of climatic precession with its interferences by obliquity. In practice, we took the normalised climatic precession minus half of the normalised tilt/obliquity (obtained from Laskar et al., 2004), referred to as p-0.5t (Zeeden et al. 2013). The use of the p-0.5t target as a reference to parametrise the decomposition opens the question of circular reasoning. Are predicted wiggles with artificial astronomical properties introduced in the decomposition because of our parametrisation, which was chosen to obtain a component that resembles the tuning target? Also, could we have inadvertently selected the wiggles in the signal that, by coincidence, happen to fit what we expect?

A few precautions should prevent such circular reasoning: firstly, the use of an EEMD-based algorithm, which is fully data-adaptive, and setting it with a sifting number of 1, which prevents smoothing of the amplitude variations. Furthermore, the parametrisation of the algorithm to optimise parsimony, within the set of parameters that allow to have a component fitting well with the target, provides a safety net against the creation of predicted wiggles. A further safeguard against circular reasoning is that the decomposition parameters are identical for the entirety of the signal; in this context, there is no possibility for a fine evolutive control of the decomposition. Such a control might have been provided by tuning the record before decomposition, but the decomposition is applied on the untuned record. Indeed, being parsimonious, the decomposition should not require strong assumptions on sedimentation rate, even to extract the amplitude modulation. The tuning interpretation of Zeeden et al. (2013) is applied on the decomposition after it is obtained, for comparison to the tuned amplitude demodulation of Zeeden et al. (2013) and to the astronomical solutions provided by Laskar et al. (2004). To the extent of the methodological

development we have currently reached, manual parametrisation and visual checks are the most practical way of working; more advanced automated optimisation is out of the scope of this article, although it might prove a promising opportunity to obtain standardised and reproducible decompositions in the future.

We show, in Figure 22, the components GSC 5 to 8, that have frequencies close to the 100 kyr short-eccentricity, 405 kyr long-eccentricity, 1000 kyr g_1 - g_5 and the 2400 kyr $g_4 - g_3$ cycles respectively (see Olsen and Kent 1999 and Hinnov 2000 for details on the g_1 - g_5 and the $g_4 - g_3$ cycles); they are compared to their equivalent in the demodulation of GSC3, in the demodulation of climatic precession solution obtained from Laskar et al. (2004), and in the demodulation performed by Zeeden et al. (2013). The amplitude demodulation of GSC3 is decomposed using "extricate", with 2 siftings applied to 500 white noise-added realisations (limited to ± 100 greyscale arbitrary units). The demodulation performed by Zeeden et al. (2013) is also decomposed using "extricate" to make the Fig. 22 (see R script in supplementary material SC2). The demodulation obtained by Zeeden et al. (2013) was provided by C. Zeeden, and is now available in the DecomposeR package. The frequency ratios of the decomposition of the amplitude demodulation of GSC3 are computed in the stratigraphic domain only (Fig. 23). The R script for all time-series analysis performed on the ODP 926 greyscale signal is available as supplementary material SC2.

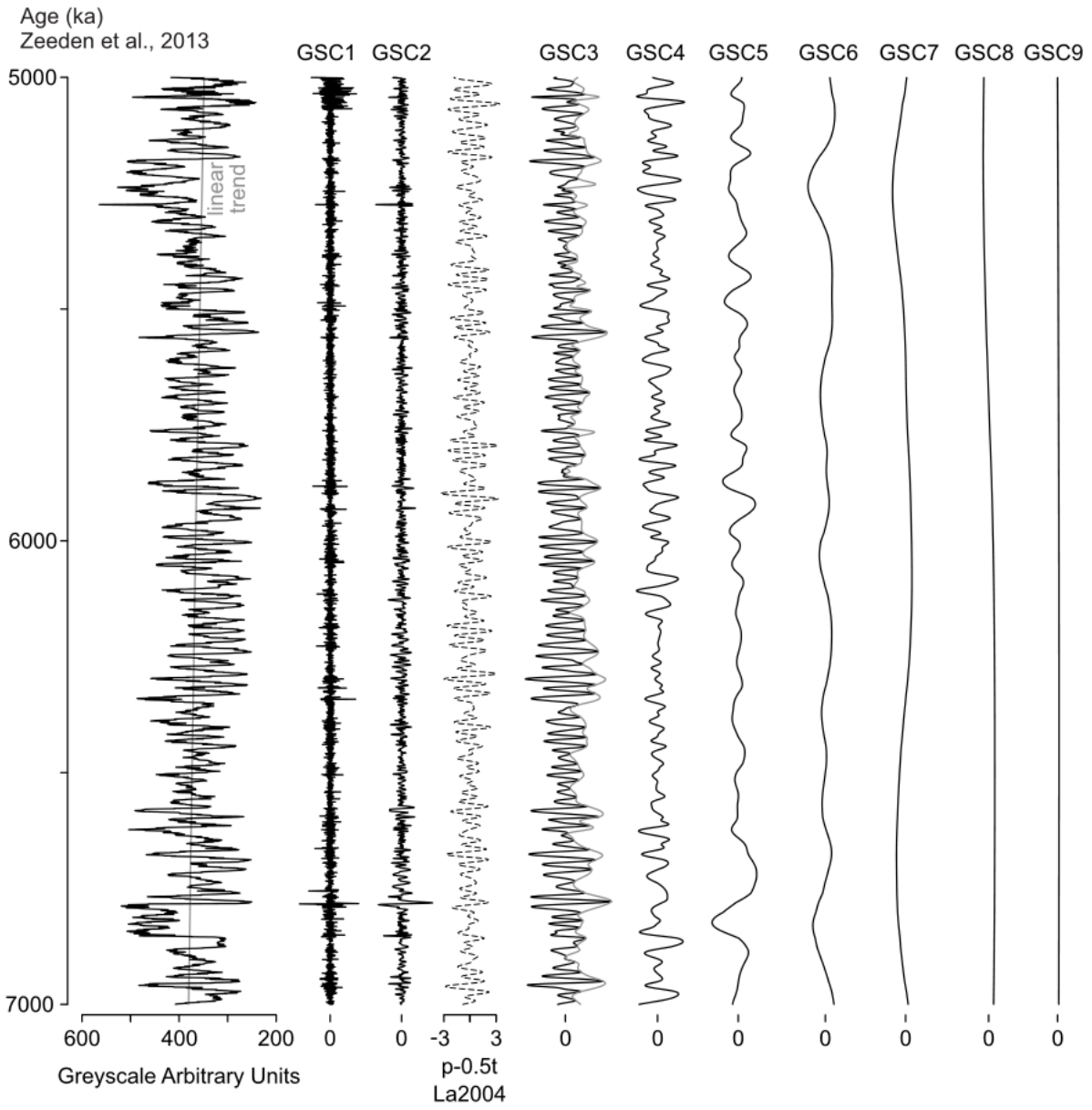


Figure 18: EEMD decomposition of the greyscale signal of ODP 926 (upper part, between 7000 and 5000 ka). The intensity scale for each component (GSC for Grey Scale Component) is identical to the one of the signal. The linear trend (grey line through the signal) is part of the decomposition. The decomposition was performed in the stratigraphic domain, and its output is here plotted in the time domain using the tuning interpretation of Zeeden et al. (2013). The tuning target, based on the climatic precession (p) and tilt (t) modelled by Laskar et al. (2004), is shown to the left of GSC3. The amplitude demodulation of GSC3 is shown in grey against GSC3. The amplitude features of GSC3 can be seen to parallel that of the tuning target (p-0.5t). The other components are discussed in the text. [NOTE TO REVIEWERS: 2 column fitting image]

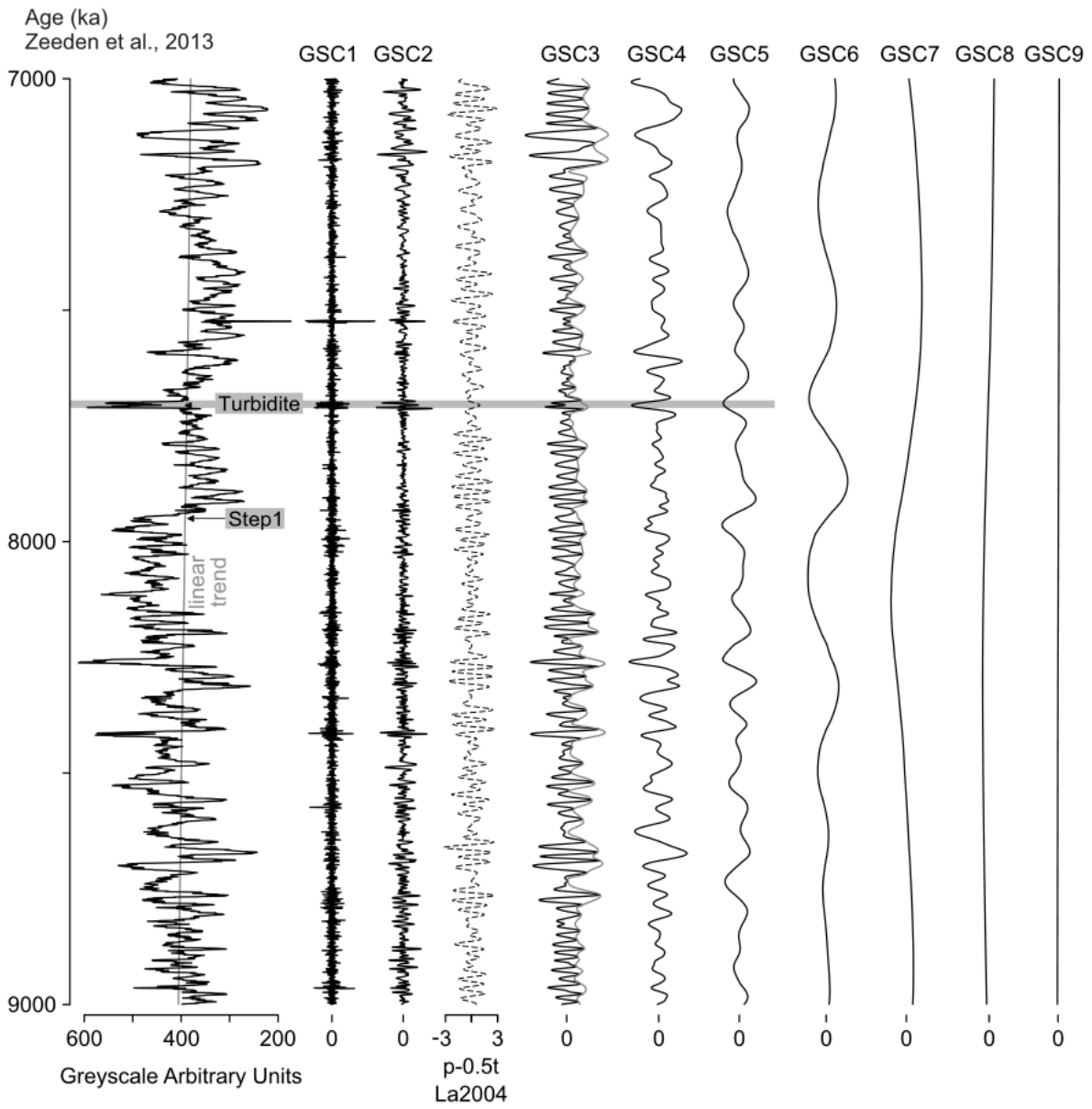
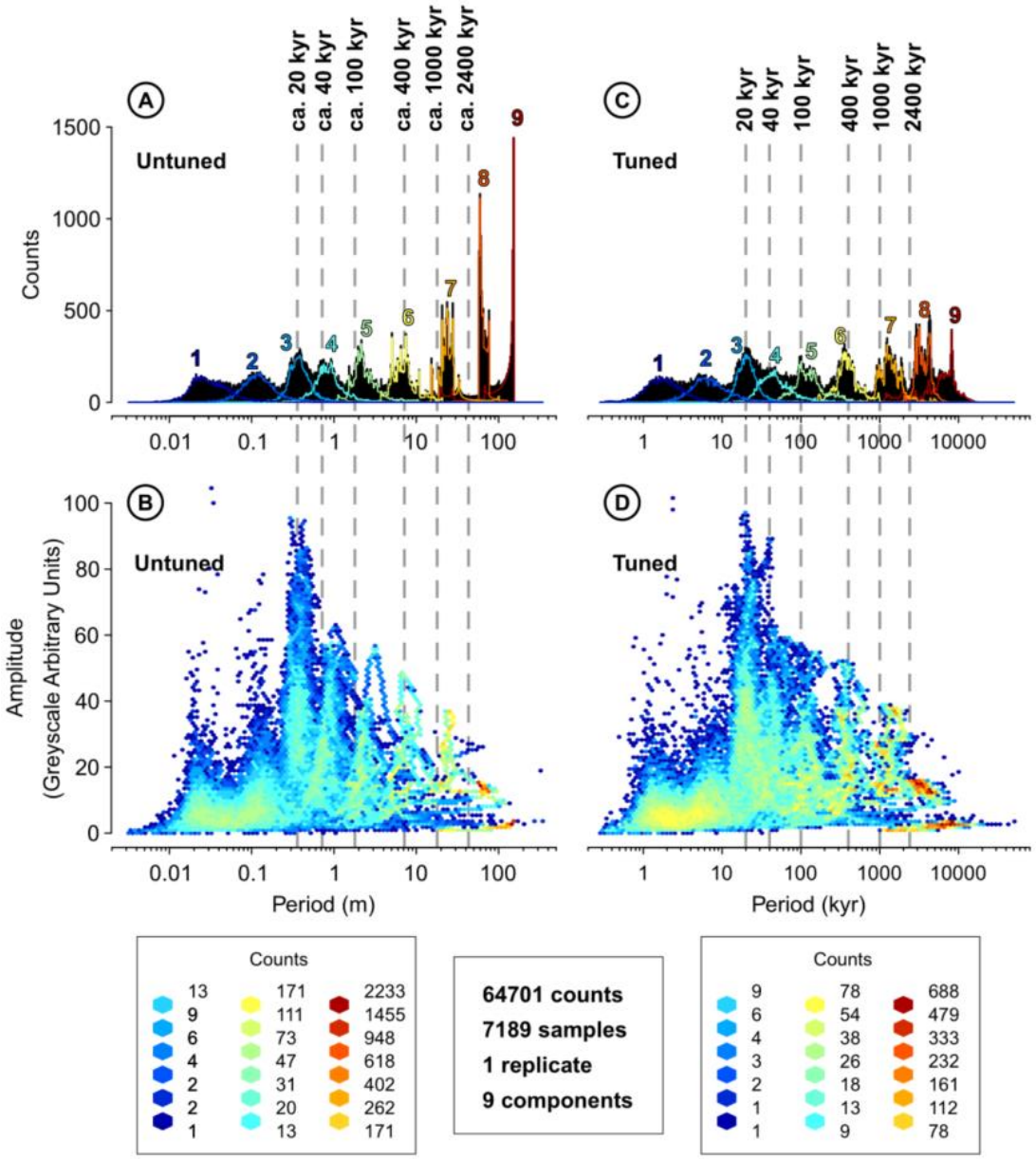


Figure 19: EEMD decomposition of the greyscale signal of ODP 926 (lower part, between 9000 and 7000 ka). The expression of the 405 kyr long eccentricity cycle is quite noticeable in GSC6. The turbidite, signalled by Zeeden et al. (2013), and shown in Fig. 16, introduces a lot of moderate to high amplitude oscillations in the 5 first components. We also signal a strong step-like value shift, Step1, whose effect on the decomposition is discussed in the text. See caption of Figure 18 for further information. [NOTE TO REVIEWERS: 2 column fitting image]



1150

Figure 20: comparison of the behaviour of instantaneous frequency in the stratigraphic domain (untuned) and in the time domain (tuned following the interpretation of Zeeden et al., 2013). (A) Period distribution, (B) spectral population of the untuned decomposition of the greyscale signal of ODP 926 (period estimations are based on the mean sedimentation rate estimated by Zeeden et al., 2013). (C and D) 1155 **Equivalent plots are provided for the same decomposition recontextualised in the time domain. The tuning was performed after the decomposition.** [NOTE TO REVIEWERS: colour online only; 2 column fitting image]

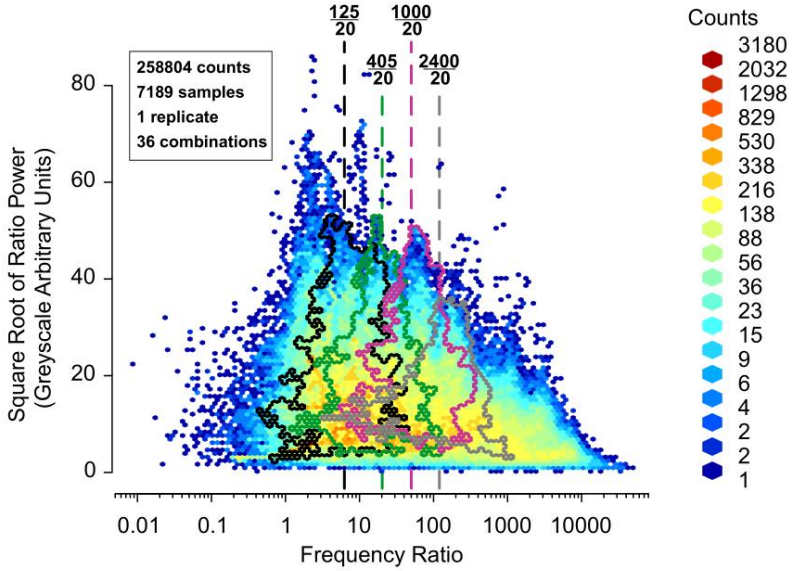


Figure 21: Ratio population for all pairs of components of the decomposition of the greyscale signal of ODP 926. The dotted lines stand for the expected frequency ratios. The black, green, purple and grey contours delineate the ratios of GSC3 with, respectively, GSC 5 to 8. [NOTE TO REVIEWERS: colour online only; 1.5 column fitting image]

The decomposition of the raw signal (Fig. 18 and 19) extracts 6 components that can potentially be attributed to astronomical cycles: GSC 3 to 8 have periods at the 20, 40, 100, 400, 1000 and 2400 kyr scale, respectively (we rounded them for simplicity), which can be visualised in Fig. 20. However, when looking more in details, the expression of the astronomical cycles is more subtle. GSC3 is centred on the 20 kyr period expected from climatic precession (Fig. 20). It further follows the p-0.5t tuning target quite closely (Fig. 18 and 19), therefore it should be expected that some expression of obliquity might be present. GSC4 is close to the expected 40 kyr cyclicity characterising obliquity. In the stratigraphic domain, GSC4 makes a distinct peak in the spectral population plot (Fig. 20B), but it displays period values higher than expected for obliquity (considering the mean sedimentation rate) (Fig. 20A). Contrastingly, in the time domain, the period distribution (Fig. 20C) is closer to 40 kyr, but the spectral population is merged with the expression of climatic precession (Fig. 20D). This might be due to the close interaction of climatic precession and obliquity, which might complicate the extraction of obliquity. In the frame of this article, we only characterise climatic precession and its amplitude modulators, therefore we will limit ourselves to mentioning the strongly suggested presence of obliquity in GSC4.

Out of the four components potentially related to eccentricity (GSC 5, 6, 7, and 8, respectively displaying periods close to the 100, 400, 1000 and 2400 kyr cycles), only GSC6 is consistently centred on a 400 kyr period, in both the stratigraphic and time domains (Fig. 20). When compared with the demodulation of climatic precession from Laskar et al.'s solution (Fig. 22), GSC6 is the only one of the four (GSC 5 to 8) to be of similar frequency and in similar phasing with its related 400 kyr cycle, but only between 9000 and 7000 ka.

The presence of the 400 kyr eccentricity cycle, not only as an amplitude modulator, but directly in the geological record, can be explained by several mechanisms. First of all, eccentricity controls the globally averaged annual insolation of the Earth, but this effect on insolation is generally considered negligible compared to the effect of precession and/or obliquity (Laskar et al. 2004). The amplitude modulation of precession and obliquity may however undergo partial rectification due to non-linear processes. Rectification makes the amplitude modulation asymmetrical and allows it to be expressed as direct cyclicity in the signal (see section 5.3.1 in Weedon 2003). Such asymmetry can further be amplified by smoothing processes like bioturbation (e.g., section 5.3.2 in Weedon 2003).

The GSC 5, 7 and 8 do not follow the expected cycles obtained from Laskar et al.'s solution (Fig. 22). The frequency ratios of GSC 5, 6, 7 and 8 with GSC 3 are at the expected values for all four components (Fig. 21), however the square root of ratio power are (slightly) above the continuum only for the ratios of GSC3 with GSC6 and of GSC3 with GSC7. The fact that the frequency ratios are at the expected ratios comes from the fact that GSC 5, 6, 7 and 8 are of the same scale than the expected eccentricity cycles, but are not necessarily representative of them. The fact that the square root of ratio power of GSC6 (with GSC3) can exceed the continuum, on the opposite, might indicate that it has significant amplitude, and is representative of the 400 kyr cycle. The high square root of ratio power of GSC7 (with GSC3), however, does not seem to mean that it is representative of the 1000 kyr cycle. Indeed, we rather interpret the high amplitude of GSC7 to come from a step-like increase of greyscale values beginning around 7950 ka (Step1 in Figs. 16 and 19). This increase is coinciding with the inception of a large amplitude wave starting at 7950 ka in both GSC6 and GSC7 (Fig. 19 and 22). We substantiate this by comparing the greyscale signal to the magnetic susceptibility one (see point 10.2 in the supplementary material); in the latter, the step is absent (Fig. 16), no particularly large oscillation occurs at 7950 ka (Fig. SF12 and SF14 in the supplementary material), and some expression of the 1000 kyr astronomical cycle seems to be recoverable (Fig. SF14).

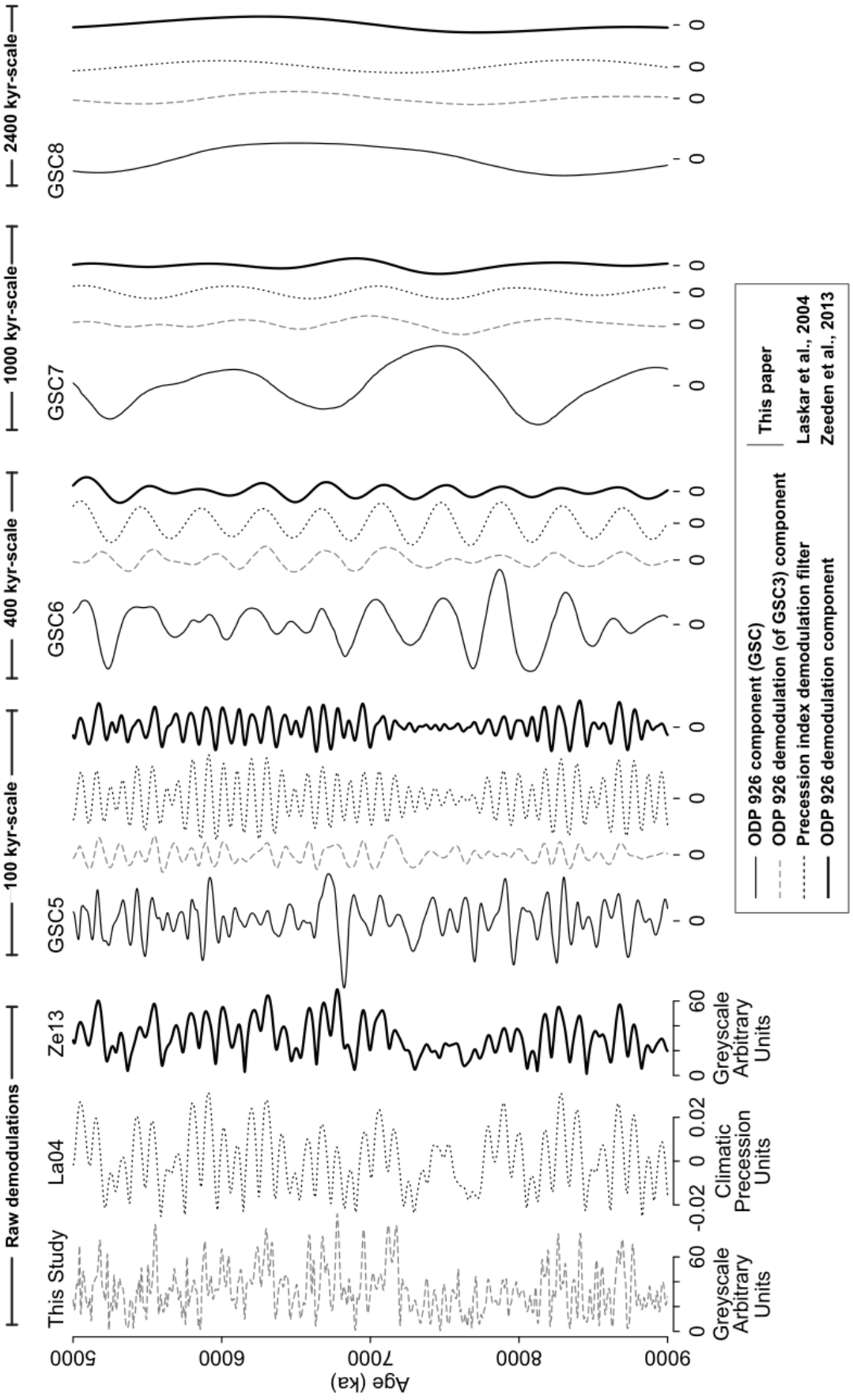

Raw demodulations
100 kyr-scale
400 kyr-scale
1000 kyr-scale
2400 kyr-scale
This Study
La04
Ze13
GSC5
GSC6
GSC7
GSC8
Greyscale Arbitrary Units
Climatic Precession Units
-0.02 0 0.02
0 60
5000
6000
7000
8000
9000
Age (ka)
ODP 926 component (GSC)
ODP 926 demodulation (of GSC3) component
Precession index demodulation filter
ODP 926 demodulation component
This paper
Laskar et al., 2004
Zeeden et al., 2013

Figure 22

Figure 22: Comparison of demodulations of the ODP 926 greyscale signal. The raw demodulations from this study and from Zeeden et al. (2013) are decomposed, the raw demodulation of the climatic precession (Laskar et al., 2004) is filtered. They are and compared in groups, respectively representing of the ca. 100, 405, 1000 and 2400 kyr-scale cycles. The corresponding GSC are provided for comparison. The intensity scales are similar to the raw demodulations (in arbitrary units for the decomposition and demodulations of the ODP 926 greyscale signal, and in climatic precession values for the Laskar et al.'s solution). [NOTE TO REVIEWERS: 2 column fitting image]

The GSC 1, 2 and 9 remain to be explained. GSC 1 and 2, with their high frequency, can be attributed to white noise, either instrumental or natural. GSC9 was treated like a cyclically-connoted residual to observe its instantaneous frequency and amplitude. Its amplitude is very small (Fig. 20), which could actually mean that GSC9 is actually a residual, and our assumption of treating it as a cyclically-connoted residual was erroneous. In this case, we could merge it with the linear trend to make a nonlinear trend.

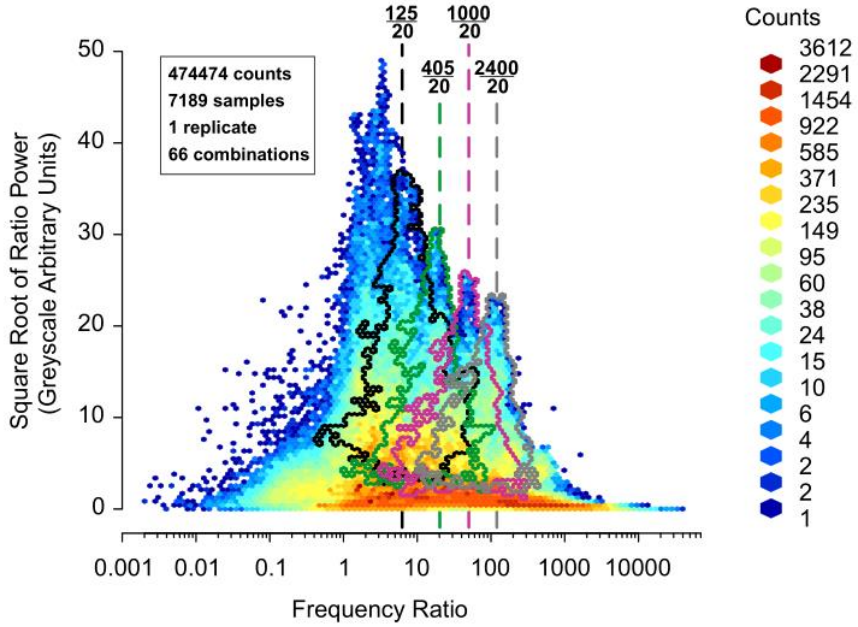


Figure 23: Ratio population for all pairs of components of the amplitude demodulation of one realisation of GSC3 of the greyscale signal of ODP 926 (see Figs. 18 and 19). The dotted lines stand for the expected frequency ratios. The black, green, purple and grey contours delineate the ratios of GSC3 with, respectively, the demodulation components corresponding to the 100, 400, 1000 and 2400 kyr (see fig. 22). [NOTE TO REVIEWERS: colour online only; 1.5 column fitting image]

From the demodulation of GSC3, we performed a decomposition to extract the relevant eccentricity cycles that came to modulate climatic precession; we compare this to the demodulation performed by Zeeden et al. (2013). The 400 kyr long-eccentricity is found with near perfect phasing with both methods (Fig. 22). The 1000 kyr cycle is found with decent phasing in both methods, but our method seem to uncover the cycles more clearly (Fig. 22). The 2400 kyr cycle is dubious for both methods. The 100 kyr component is found by Zeeden et al.'s method; some cycles may be lost, but it generally gets a good fit with Laskar et al.'s solution (Fig. 22). Even the

amplitude modulation is well identified. Our method recovers some of the correct 100 kyr cycles in GSC5, but skips most of the peaks; moreover the amplitude modulation is not as clear (Fig. 22). The frequency ratios are well identified in the demodulation of GSC3 (Fig. 23). The component used for demodulation (GSC3) is taken in consideration for the ratios, and as it has higher amplitude than the demodulated components. Consequently, when GSC3 is part of a pair of components, it amplifies the square root of ratio power (similarly to what we observe in the CIP1 analysis, see point 10.1 in the supplementary material). The frequency ratios are better centred when obtained from the demodulation (Fig. 23) than from the raw signal (Fig. 21): as the expression of eccentricity cycles is in our case better in the demodulation that the decomposition of the raw signal, this better centring could be a criteria of a higher quality of astronomical cycles' extraction. Therefore, a good centring, along with relatively high square root of ratio power values, might be an important aspect in understanding the ratio population plot.

The integrity of the greyscale signal raw decomposition is under 10^{-13} arbitrary units, which falls into acceptable limits for our intensity values. Its parsimony is of 2.06, i.e. 106% more absolute intensity in the decomposition than in the signal. We highlight that this is higher than the 1.65 parsimony value obtained for the CIP1 decomposition (see point 10.1 in the supplementary material). This difference could be due to the fact that the ODP 926 signal is more complicated than the CIP1 signal, but also due to the greater number of components in the ODP 926 signal (9 plus the linear trend for the greyscale signal of ODP 926, against 7 plus the linear trend for the CIP1 signal). We highlight, however, that parsimony is better suited to compare different decompositions of a same signal, rather than decompositions of different signals.

We show in Table 1 that the GZC departure is low for well-expressed cycles, typically GSC3 and GSC6 who stand for the 20 kyr climatic precession and the 400 kyr eccentricity cycles. The lowest value however comes from GSC7, which is not representative of the 1000 kyr astronomical cycle (although of similar periodicity), but still is expressed as clear cycles (Fig. 22). This highlights that GZC departure is a metric of how well a component can be expressed in instantaneous frequency, but that it does not necessarily mean that a component can be interpreted as a periodic or quasi-periodic cycle. We computed the GZC departure for both the tuned and untuned versions, and the values are roughly comparable: this means that tuning does not seem to affect the relevance of the calculation of instantaneous frequency. The highest GZC departure values are found for the two highest frequency components, which could be explained by the abrupt nature of oscillations of periods close to the sampling interval, but also by the difficulty of avoiding riding waves in components representative of the high frequency part of white noise. For GSC 8 and 9, who do not have enough cycles for computing the GZC, we mention a symmetry of 0.53 and 0.94 respectively. Therefore GSC8 is highly symmetric, and its instantaneous

frequency computation is considered to be correct. On the other hand, the dissymmetry of GSC9 confirms that it might be better interpreted as a nonlinear trend (when combined with the linear trend).

			20 kyr	**ca. 40 kyr**	**ca. 100 kyr**	**400 kyr**	
	GSC1	GSC2	GSC3	GSC4	GSC5	GSC6	GSC7
Untuned	1.55	1.42	1.18	1.28	1.23	1.15	1.09
Tuned	1.45	1.38	1.18	1.30	1.28	1.20	1.15

Table 1: GZC departure for the Grey Scale Components (GSC; see Figs. 18 and 19). The astronomical interpretation is provided (first line) when available. We compare it for both untuned (stratigraphic domain) and tuned (time domain, following the tuning interpretation of Zeeden et al., 2013) frequency computation.

We estimate in Fig. 20 the effect, on the frequency distribution, of applying the tuning interpretation of Zeeden et al. (2013) on the decomposition. GSC3 and GSC6 are equally well centred on their expected periods of respectively 20 kyr and 400 kyr. This may come from the fact that the sedimentation rate changes are not excessive. The most striking change, however, is that low-frequency components (GSC 7 to 9) are more spread out in the tuned version. This is explained by the EEMD process, which produces components that are smooth, especially when they are of low amplitude and low frequency. This smoothness means that intra-wave frequency modulation is minimal. As there are only a few of these smooth oscillations present in the low-frequency components, and that each of them are made out of a large number of points, the instantaneous frequency distribution is limited to a narrow range of values. The tuning process reduces the smoothness of these components, thus spreading their frequency range.

We finally mention that the greyscale signal is affected by a turbidite (an instantaneous deposit produced by a sediment gravity flow) which disturbs the otherwise continuous recording climatic signal (Fig. 16). A single turbidite was indeed mentioned by Zeeden et al. (2013), and we study here its effect on the signal decomposition (Fig. 19). The turbidite adds thickness to the wiggles including it. Specifically, the turbidite locally increases the period of GSC6 which represents the 400 kyr cycle. For the cycles that have similar or lower frequency, it adds wiggles that have moderate to high amplitude, relative to the other cycles. This is mainly because this specific turbidite acts as an outlier, i.e. a relatively high amplitude oscillation. In the magnetic susceptibility signal, the turbidite has less high-frequency wiggles, and it is therefore expressed differently in the decomposition (see section 10.2 in supplementary material). This warns us that individual features in a signal can impact its decomposition in different ways. However, the decompositions are only affected locally by this event deposit/outlier.

6 Discussion

6.1 Concepts synthesis

We have shown that cyclostratigraphic signals display the expressions of several processes, entangled with one another. Synthetically, we could refer to these expressions as "information", which concretely refers to the behaviour of the wiggles found in the signal. Also, we have shown that wiggles not present in the signal can appear in a decomposition. Such wiggles can give predictions on the processes having affected the signal, which is why we refer to them as predicted wiggles. Similarly, the behaviour of predicted wiggles can be referred to as "predicted information". On this basis, we reviewed and devised key concepts and tools, with three goals in mind:

- Disentangling the information pertaining to distinct processes, using time-series analysis.
- Characterising and/or reducing the predicted information that is added to the signal in the time-series analysis process.
- Checking that no information from the original signal is discarded in the time-series analysis process.

The classical cyclostratigraphic tools for disentangling the information contained in a signal are transforms, filters, and decompositions. The extent to which the disentanglement is performed depends on the methodology; in idealised conditions, one would want perfect meaningfulness, i.e. the full separation of all processes involved in creating the signal. But such an objective is challenging in signals as complex as cyclostratigraphic ones, and we especially stressed that changes in the sedimentation rate are perhaps the most challenging circumstance for meaningful signal analysis. The main risk, in wanting to push meaningfulness, especially in the context of changing sedimentation rate, is producing wiggles not present in the signal.

Therefore, the characterisation and/or reduction of predicted information is a necessity when applying time-series analysis to disentangle cyclostratigraphic signals. Indeed, the multiplicity and complexity of processes affecting sedimentary records, among which sedimentation rate changes, make the interpretation of predicted information hazardous. However, we highlight that in other contexts, by using reasonable assumptions, predicted information can judiciously be attributed to given processes. Specifically, a signal that is modulated in amplitude may as well be decomposed into a sum of non-modulated signals that exhibit heavy cross-cancellation. Still, although it is mathematically equivalent, we argue that in order to best characterise how a cyclostratigraphic signal is expressing itself, sticking as closely as possible to the observable wiggles carries less risks of mistakes. And to the extent that this cannot be accomplished, one should at least characterise the information that is added in the decomposition.

Additionally, at the end of the time-series analysis process, no information should be lost. This is here addressed by two concepts: integrity, which quantifies whether the components of a decomposition can be summed to reconstruct the signal flawlessly, and more generally, reversibility, which conceptualises the ability to retrieve the input, as it was initially, from the output of any data processing method.

To address those three aspects of signal analysis, we should strive to obtain modally meaningful, parsimonious, and reversible decompositions. We provided the concept of parsimony to guide decompositions methods into minimising the cross-cancellation of wiggles, thereby reducing the occurrence of predicted information. We refer to this as the wiggle-in-signal approach. Modal meaningfulness was introduced as a practical way to approach meaningfulness as closely as possible without compromising parsimony; effectively, this means that sub-signals are representative of all the same frequency scale processes interfering and modulating the expression of one another. This is the furthest that information disentanglement can be brought while minimising predicted wiggles. Moreover, a reversible decomposition, including all the original data points, makes the flawless reconstruction of the initial signal possible, which in turn serves as guarantee against any loss of information.

We showed that EMD-based methodologies provide a practical way of obtaining modally meaningful, parsimonious, and reversible decompositions. They are data-adaptive, which effectively reduces the introduction of predicted information (when the decomposition algorithms are adequately parametrised). Such parsimonious decompositions open the possibility to systematically test tuning interpretations, even when they are overtuned, by preventing the creation of amplitude artefacts. EMD-based methodologies can also be optimised to reduce mode mixing, making the decompositions practically dyadic, which is similar to modal meaningfulness. Coincidentally, modally meaningful (or dyadic) components can be treated like IMFs (Intrinsic Mode Functions), from which instantaneous frequency, amplitude, and ratio of frequency can be obtained. Therefore, pushing the disentanglement of cyclostratigraphic signals "only" to modal meaningfulness is enough to have a quantified spectral analysis of cyclostratigraphic signals. Noteworthily, the computation of instantaneous frequency ratio opens new opportunities to detect astronomical cycles in deep geological time. Finally, all the data processing methodologies can be made to be reversible in every practical way. Specifically, the methodologies needing regularly sampled time-series can use the greatest rational common divisor for a linearly interpolated resampling that preserves the original data points.

6.2 Case study

We tested the concept of modally meaningful, parsimonious, and reversible decomposition with a real case study: the greyscale signal of ODP 926 (and its magnetic susceptibility equivalent, in the supplementary

material). The main challenge that we needed to overcome was the extraction of modally meaningful components that could be interpreted in terms of astronomical cycles. The identification of climatic precession, with its eccentricity modulation (and interferences from obliquity), along with a partial 400 kyr eccentricity cycle directly in the signal, confirms that our method can indeed obtain such modally meaningful components. This is further confirmed by the good correspondence between the demodulation obtained by our method (using our EEMD algorithm and the normalisation procedure of Huang et al., 2009) and demodulations performed by Zeeden et al. (2013), despite the wide difference between the techniques (among others, the absence of tuning on our part).

We attempted to provide an interpretation for all components of the decomposition of ODP 926. However, we did not explicitly account for the impact of continuum climatic variability, i.e. climatic noise. Typically, climatic noise is described in power spectra by power-law scaling with a dominance of low-frequencies (see Huybers and Curry, 2006 and Ditlevsen et al., 2020 for more details). This variability is well understood in power spectra, but the extraction and characterisation of noise in the context of decomposition (especially in climatic signals that are a mix of noise, oscillatory processes, etc.; see, e.g., Mann and Lees 1996 for more details) is out of the scope of this paper. We, however, mention the works of Flandrin et al. (2004) and Rilling et al. (2005) on characterising the behaviour of different noise processes in decomposition, and of Franzke (2009) on the extraction of low-frequency noise.

Nonetheless, our method allows for a discussion on the processes recording the cycles. This was more easily done in the CIP1 artificial signal, because the way the signal was generated is well known (see supplementary material). For the ODP 926 signals, we observed that the 400 kyr long-eccentricity cycle was clearly recorded directly in the signal between 9000 and 7000 ka, which can be explained by non-linear processes such as rectification and smoothing. Similarly, other processes which were not investigated more in details in this study can be investigated in our decomposition; we mentioned for instance the suspected expression of obliquity in the ODP 926 signals. This confirms our point that in a meaningful decomposition, the entirety of the processes affecting the signal can be perceived, outlining the processes that remain to be studied more thoroughly.

6.3 Perspective for the future

In a general sense, the problem of signal extraction is deciding how to deal with all the wiggles found in a signal; how they can best be singled out and grouped into well-defined components. In that regard, the EMD-based techniques is only one of the possible answers to that problem.

We suggest that decompositions could be improved upon by locally exchanging intensity between different components of a first prototype of decomposition, to further refine it. Such a process of iterative

improvement of decompositions could be computer-driven or user-driven (or a mix of both), set to preserve integrity, and optimised to search for specific properties of processes such as astronomical cycles (e.g., stable frequency ratios, similarity to an astronomical solution or another decomposition) while preserving reasonably good parsimony.

The exchange of intensity between components could also be especially useful to study the shape of wiggles: we showed for instance that the sharply angular shaped oscillations found in our Figure 1 example (mainly due to the 23 kyr forcing) were distributed into several components of distinct scale. These different scale wiggles can be regrouped, based on their relative position with one another. Indeed, there can be a phase relationship between different frequency range oscillations in a decomposition. Typically, we can observe in the Figure 1 decomposition that high-frequency wiggles occur exactly at the angular features of the 23 kyr forcing peaks. This is a phenomena similar to harmonics, and merging all wiggles having a phase relationship with one another would better reconstruct the actual form of the wiggles as they are in the signals. Similarly, we have observed in section 8 (see supplementary materials) that problems at the ends of the signals can be expressed by intensity being exchanged between components. Therefore, setting a procedure to exchange intensity back into the appropriate components might provide a solution for such problems.

The study of the shape of oscillations is only one of the opportunities that the interpretable nature of decomposition opens. In the example of the dilution/accumulation cycle of Fig. 1, the link between the relative sedimentation rate and the intensity of the cycles was shown in a spectral population plot (Fig. 13). This was also observed in the CIP1 signal (see supplementary material fig. SF7). This implies that decomposition can be coupled with quantitative estimation of the sedimentary input, which would allow to test records against geological models for the sensitivity of sediment deposition to astronomical cycles (e.g., Herbert 1994).

7 Conclusions

Decomposition provides a comprehensive understanding of a signal, and can be used to study the interactions of conservative (astronomical) cycles, dissipative oscillators, heterogeneous processes or noise in cyclostratigraphic signals. Furthermore, decomposition highlights possible problems of interpretation or processing, typically coming from the loss of information, or the adding of predicted information. We argue that this makes decomposition particularly suitable to study cyclostratigraphic signals, as there is no way of unscrambling with absolute certainty and precision the blended influences of all the processes having left an impact on the sedimentary records.

To perform modally meaningful decompositions we implemented an Ensemble Empirical Mode Decomposition (EEMD) method to be used in R: the "extricate" algorithm. The components of decompositions so obtained can directly be compared to the astronomical solutions or to decompositions of other geological signals. This can be applied similarly on the amplitude demodulation of filters or components, or directly on the components of a decomposition. We also provide tools to perform instantaneous spectral analysis, which allows identifying frequencies and frequency ratios expected from astronomical cycles directly from the decompositions. To further provide a quantified basis to discuss the representativeness of a decomposition and of its instantaneous spectral analysis to a given signal, we present the integrity, parsimony, GZC (generalized zero-crossing) departure and symmetry quality metrics.

All these concepts, applied to ODP 926's signals, show not only the expected expression of climatic precession, with its amplitude modulation by eccentricity, but also the expression of eccentricity as direct cycles in the signal, implying the action of non-linear processes in the geological system, which effectively rectify and smooth the initial insolation signal. The expression of obliquity and of climatic noise is hinted, but these were not studied in detail in this paper. However, we suggest that they can still be studied in the decomposition we performed.

Indeed, decomposition as a format for cyclostratigraphic interpretation provides the means for further re-examination of the signal. Facilitating such a re-examination of data is especially important for astrochronology, because its unmatched ability to provide a continuous estimation of the passage of time means that validating or updating cyclostratigraphic output can most effectively be performed using cyclostratigraphic analysis of new sections, with other stratigraphic methods providing tie-points between sections. The distortion of astronomical cycles, prevalence of noise and diagenetic processes, as well as the lack of astronomical solutions in the older records further means that comparing records to "fill in the gaps" is the most robust and universal way of finding a global signature of the expression of astronomical cycles.

Overall, decomposition allows to have a fundamental understanding of time-series analysis methods and results, regardless of their complexity. Indeed, the principle of decomposition should prove to remain firmly intuitive, whichever way decomposition will be performed in the future.

Code and data availability

The R script used to make and analyse the accumulation-dilution cycles signal (Figs. 1, 3, 4, 13 and 14) is given as supplementary material SC1. The R scripts used to process the ODP 926 greyscale and magnetic

susceptibility signals are provided as supplementary material SC2 and SC3 respectively. The R script to process the CIP1 is provided as supplementary material SC4.

Acknowledgments

We thank Slah Boulila, Linda Hinnov and Stephen Meyers for their crucial review of the manuscript. SW would like to thank the Belgian National Fund for Scientific Research (FNRS) for the FRIA funding provided for his thesis. MC is funded by the Belgian FNRS as research director. ACDS thanks the IGCP-652 project (Reading geologic time in Paleozoic sedimentary rocks), as well as the FNRS (grants T.0051.19-PDR and J.0037.21-CDR) for their support. Gratitude also goes to Renata Ferreira de Barros, Bruno Meyvis and Giorgia Stasi for discussions to make the examples and illustrations understandable.

References

Berger, A. 1978. 'Long-Term Variations of Daily Insolation and Quaternary Climatic Changes'. *Journal of the Atmospheric Sciences* 35 (12): 2362–67. https://doi.org/10.1175/1520-0469(1978)035<2362:LTVODI>2.0.CO;2.

Boashash, B. 1992. 'Estimating and Interpreting the Instantaneous Frequency of a Signal. I. Fundamentals'. *Proceedings of the IEEE* 80 (4): 520–38. https://doi.org/10.1109/5.135376.

Boulila, S., M. Vahlenkamp, D. De Vleeschouwer, J. Laskar, Yuhji Yamamoto, H. Pälike, S. Kirtland Turner, P. F. Sexton, T. Westerhold, and U. Röhl. 2018. 'Towards a Robust and Consistent Middle Eocene Astronomical Timescale'. *Earth and Planetary Science Letters* 486 (March): 94–107. https://doi.org/10.1016/j.epsl.2018.01.003.

Bowman, D. C., and J. M. Lees. 2013. 'The Hilbert–Huang Transform: A High Resolution Spectral Method for Nonlinear and Nonstationary Time Series'. *Seismological Research Letters* 84 (6): 1074–80. https://doi.org/10.1785/0220130025.

Charbonnier, G., S. Boulila, J. E. Spangenberg, T. Adatte, K. B. Föllmi, and J. Laskar. 2018. 'Obliquity Pacing of the Hydrological Cycle during the Oceanic Anoxic Event 2'. *Earth and Planetary Science Letters* 499 (October): 266–77. https://doi.org/10.1016/j.epsl.2018.07.029.

Cicone, A., J. Liu, and H. Zhou. 2016. 'Adaptive Local Iterative Filtering for Signal Decomposition and Instantaneous Frequency Analysis'. *Applied and Computational Harmonic Analysis*, Sparse Representations with Applications in Imaging Science, Data Analysis, and Beyond, Part II, 41 (2): 384–411. https://doi.org/10.1016/j.acha.2016.03.001.

Dijkstra, H. A., and M. Ghil. 2005. 'Low-Frequency Variability of the Large-Scale Ocean Circulation: A Dynamical Systems Approach'. *Reviews of Geophysics* 43 (3). https://doi.org/10.1029/2002RG000122.

Ditlevsen, P., T. Mitsui, and M. Crucifix. 2020. 'Crossover and Peaks in the Pleistocene Climate Spectrum; Understanding from Simple Ice Age Models'. *Climate Dynamics* 54 (3): 1801–18. https://doi.org/10.1007/s00382-019-05087-3.

Flandrin, P., G. Rilling, and P. Goncalves. 2004. 'Empirical Mode Decomposition as a Filter Bank'. *IEEE Signal Processing Letters* 11 (2): 112–14. https://doi.org/10.1109/LSP.2003.821662.

Franzke, C. 2009. 'Multi-Scale Analysis of Teleconnection Indices: Climate Noise and Nonlinear Trend Analysis'. *Nonlinear Processes in Geophysics* 16 (1): 65–76. https://doi.org/10.5194/npg-16-65-2009.

Ghil, M., M. R. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. W. Robertson, et al. 2002. 'Advanced Spectral Methods for Climatic Time Series'. *Reviews of Geophysics* 40 (1): 3-1-3–41. https://doi.org/10.1029/2000RG000092.

Hays, J. D., John Imbrie, and N. J. Shackleton. 1976. 'Variations in the Earth's Orbit: Pacemaker of the Ice Ages'. *Science* 194 (4270): 1121–32. https://doi.org/10.1126/science.194.4270.1121.

Herbert, T. D. 1994. 'Reading Orbital Signals Distorted by Sedimentation: Models and Examples'. In *Orbital Forcing and Cyclic Sequences*, 483–507. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781444304039.ch29.

Hilgen, F. J., L. A. Hinnov, H. A. Aziz, H. A. Abels, S. Batenburg, J. H. C. Bosmans, B. de Boer, et al. 2014. 'Stratigraphic Continuity and Fragmentary Sedimentation: The Success of Cyclostratigraphy as Part of Integrated Stratigraphy'. *Geological Society, London, Special Publications* 404 (October): SP404.12. https://doi.org/10.1144/SP404.12.

Hinnov, L. A. 2000. 'New Perspectives on Orbitally Forced Stratigraphy'. *Annual Review of Earth and Planetary Sciences* 28 (1): 419–75. https://doi.org/10.1146/annurev.earth.28.1.419.

Hinnov, L. A.. 2013. 'Cyclostratigraphy and Its Revolutionizing Applications in the Earth and Planetary Sciences'. *Geological Society of America Bulletin* 125 (November): 1703–34. https://doi.org/10.1130/B30934.1.

Hinnov, L. A., Michael Schulz, and Pascal Yiou. 2002. 'Interhemispheric Space–Time Attributes of the Dansgaard–Oeschger Oscillations between 100 and 0ka'. *Quaternary Science Reviews*, Decadal-to-Millennial-Scale Climate Variability, 21 (10): 1213–28. https://doi.org/10.1016/S0277-3791(01)00140-8.

Huang, N. E., Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu. 1998. 'The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis'. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 454 (1971): 903–95. https://doi.org/10.1098/rspa.1998.0193.

Huang, N. E., Z. Wu, S. R. Long, K. C. Arnold, X. Chen, and K. Blank. 2009. 'On Instantaneous Frequency'. *Advances in Adaptive Data Analysis* 01 (02): 177–229. https://doi.org/10.1142/S1793536909000096.

Huybers, P., and W. Curry. 2006. 'Links between Annual, Milankovitch and Continuum Temperature Variability'. *Nature* 441 (7091): 329–32. https://doi.org/10.1038/nature04745.

Kemp, D. B. 2016. 'Optimizing Significance Testing of Astronomical Forcing in Cyclostratigraphy'. *Paleoceanography* 31 (12): 1516–31. https://doi.org/10.1002/2016PA002963.

Kim, D., and H.-S. Oh. 2016. 'Empirical Mode Decomposition with Missing Values'. *SpringerPlus* 5 (1): 2016. https://doi.org/10.1186/s40064-016-3692-1.

Kodama, K. P., and L. A. Hinnov. 2014. *Rock Magnetic Cyclostratigraphy*. John Wiley & Sons.

Laepple, T., and P.Huybers. 2013. 'Reconciling Discrepancies between Uk37 and Mg/Ca Reconstructions of Holocene Marine Temperature Variability'. *Earth and Planetary Science Letters* 375 (August): 418–29. https://doi.org/10.1016/j.epsl.2013.06.006.

Laskar, J., A. Fienga, M. Gastineau, and H. Manche. 2011. 'La2010: A New Orbital Solution for the Long Term Motion of the Earth'. *Astronomy & Astrophysics* 532 (August): A89. https://doi.org/10.1051/0004-6361/201116836.

Laskar, J., P. Robutel, F. Joutel, M. Gastineau, A. C. M. Correia, and B. Levrard. 2004. 'A Long-Term Numerical Solution for the Insolation Quantities of the Earth'. *Astronomy and Astrophysics* 428 (December): 261–85. https://doi.org/10.1051/0004-6361:20041335.

Lenoir, G., and M. Crucifix. 2018. 'A general theory on frequency and time–frequency analysis of irregularly sampled time series based on projection methods – Part 1: Frequency analysis'. *Nonlinear Processes in Geophysics* 25: 145. https://doi.org/10.5194/npg-25-145-2018.

Li, M., L. R. Kump, L. A. Hinnov, and M. E. Mann. 2018. 'Tracking Variable Sedimentation Rates and Astronomical Forcing in Phanerozoic Paleoclimate Proxy Series with Evolutionary Correlation Coefficients and Hypothesis Testing'. *Earth and Planetary Science Letters* 501 (November): 165–79. https://doi.org/10.1016/j.epsl.2018.08.041.

Lin, L., Y. Wang, and H. Zhou. 2009. 'Iterative Filtering as an Alternative Algorithm for Empirical Mode Decomposition'. *Advances in Adaptive Data Analysis* 01 (04): 543–60. https://doi.org/10.1142/S179353690900028X.

Lisiecki, L. E., and P. A. Lisiecki. 2002. 'Application of Dynamic Programming to the Correlation of Paleoclimate Records'. *Paleoceanography* 17 (4): 1-1-1–12. https://doi.org/10.1029/2001PA000733.

Lourens, L. J., R. Wehausen, and H. J. Brumsack. 2001. 'Geological Constraints on Tidal Dissipation and Dynamical Ellipticity of the Earth over the Past Three Million Years'. *Nature* 409 (6823): 1029–33. https://doi.org/10.1038/35059062.

Malinverno, A., E. Erba, and T. D. Herbert. 2010. 'Orbital Tuning as an Inverse Problem: Chronology of the Early Aptian Oceanic Anoxic Event 1a (Selli Level) in the Cismon APTICORE'. *Paleoceanography* 25 (2). https://doi.org/10.1029/2009PA001769.

Maliva, R. G., and R. Siever. 1989. 'Nodular Chert Formation in Carbonate Rocks'. *The Journal of Geology* 97 (4): 421–33. https://doi.org/10.1086/629320.

Mann, M. E., and J. M. Lees. 1996. 'Robust Estimation of Background Noise and Signal Detection in Climatic Time Series'. *Climatic Change* 33 (3): 409–45. https://doi.org/10.1007/BF00142586.

Maraun, D., and J. Kurths. 2004. 'Cross Wavelet Analysis: Significance Testing and Pitfalls'. *Nonlinear Processes in Geophysics* 11 (November): 505–14. https://doi.org/10.5194/npg-11-505-2004.

Martinez, M., J.-F. Deconinck, P. Pellenard, L. Riquier, M. Company, S. Reboulet, and M. Moiroud. 2015. 'Astrochronology of the Valanginian–Hauterivian Stages (Early Cretaceous): Chronological

Relationships between the Paraná–Etendeka Large Igneous Province and the Weissert and the Faraoni Events'. *Global and Planetary Change* 131 (August): 158–73. https://doi.org/10.1016/j.gloplacha.2015.06.001.

Martinez, M., S. Kotov, D. De Vleeschouwer, D. Pas, and H. Pälike. 2016. 'Testing the Impact of Stratigraphic Uncertainty on Spectral Analyses of Sedimentary Series'. *Climate of the Past* 12 (9): 1765–83. https://doi.org/10.5194/cp-12-1765-2016.

Meyers, S. R. 2012. 'Seeing Red in Cyclic Stratigraphy: Spectral Noise Estimation for Astrochronology'. *Paleoceanography* 27 (September): PA3228. https://doi.org/10.1029/2012PA002307.

———. 2015. 'The Evaluation of Eccentricity-Related Amplitude Modulation and Bundling in Paleoclimate Data: An Inverse Approach for Astrochronologic Testing and Time Scale Optimization'. *Paleoceanography* 30 (12): 1625–40. https://doi.org/10.1002/2015PA002850.

———. 2019. 'Cyclostratigraphy and the Problem of Astrochronologic Testing'. *Earth-Science Reviews* 190 (March): 190–223. https://doi.org/10.1016/j.earscirev.2018.11.015.

Meyers, S. R., and B. B. Sageman. 2007. 'Quantification of Deep-Time Orbital Forcing by Average Spectral Misfit'. *American Journal of Science* 307 (May): 773–92. https://doi.org/10.2475/05.2007.01.

Meyers, S. R., B. B. Sageman, and M. A. Arthur. 2012. 'Obliquity Forcing of Organic Matter Accumulation during Oceanic Anoxic Event 2'. *Paleoceanography* 27 (3). https://doi.org/10.1029/2012PA002286.

Meyers, S. R., B. B. Sageman, and L. A. Hinnov. 2001. 'Integrated Quantitative Stratigraphy of the Cenomanian-Turonian Bridge Creek Limestone Member Using Evolutive Harmonic Analysis and Stratigraphic Modeling'. *Journal of Sedimentary Research* 71 (4): 628–44. https://doi.org/10.1306/012401710628.

Milankovitch, M. 1941. *Kanon der Erdbestrahlung und seine Anwendung auf das Eiszeitenproblem*. 33rd ed. Belgrade: Royal Serbian Academy, Section of Mathematical and Natural Sciences.

Muller, R. A., and G. J. MacDonald. 2000. *Ice Ages and Astronomical Causes: Data, Spectral Analysis and Mechanisms*. Environmental Sciences. Berlin Heidelberg: Springer-Verlag. https://www.springer.com/gp/book/9783540437796.

Murray, J. D. 2013. *Mathematical Biology*. Biomathematics. Springer Science & Business Media. https://doi.org/10.1007/978-3-662-08542-4.

Nicolis, G., and A. De Wit. 2007. 'Reaction-Diffusion Systems'. *Scholarpedia* 2 (9): 1475. https://doi.org/10.4249/scholarpedia.1475.

Olsen, P. E., and D. V. Kent. 1999. 'Long-Period Milankovitch Cycles from the Late Triassic and Early Jurassic of Eastern North America and Their Implications for the Calibration of the Early Mesozoic Time–Scale and the Long–Term Behaviour of the Planets'. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 357 (1757): 1761–86. https://doi.org/10.1098/rsta.1999.0400.

Percival, D. B., and A. T. Walden. 1993. *Spectral Analysis for Physical Applications*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511622762.

———, eds. 2000. 'Introduction to Wavelets'. In *Wavelet Methods for Time Series Analysis*, 1–19. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511841040.002.

Qin, Y., B. Tang, and Y. Mao. 2016. 'Adaptive Signal Decomposition Based on Wavelet Ridge and Its Application'. *Signal Processing* 120 (March): 480–94. https://doi.org/10.1016/j.sigpro.2015.09.032.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austrua: R Foundation for Statistical Computing. https://www.R-project.org/.

Raymo, M. E. 1997. 'The Timing of Major Climate Terminations'. *Paleoceanography* 12 (4): 577–85. https://doi.org/10.1029/97PA01169.

Rilling, G., P. Flandrin, and P. Gonçalves. 2005. 'Empirical Mode Decomposition, Fractional Gaussian Noise and Hurst Exponent Estimation'. In *IEEE Int. Conf. on Acoust. Speech and Sig. Proc.* Philadelphia, United States. https://hal.inria.fr/inria-00570581.

Schiffelbein, P., and L. Dorman. 1986. 'Spectral Effects of Time-Depth Nonlinearities in Deep Sea Sediment Records: A Demodulation Technique for Realigning Time and Depth Scales'. *Journal of Geophysical Research: Solid Earth* 91 (B3): 3821–35. https://doi.org/10.1029/JB091iB03p03821.

Shackleton, N. J., and S. Crowhurst. 1997. 'Sediment Fluxes Based on an Orbitally Tuned Time Scale 5 Ma to 14 Ma, Site 926'. In *Hackleton, N.J., Curry, W.B., Richter, C., and Bralower, T.J. (Eds.), Proceedings of the Ocean Drilling Program, Scientific Results, Vol. 154*.

Shipboard Scientific Party. 1995a. 'Leg 154 Introduction'. In *Curry, W., Shackleton, N., Richter, C., Bralower, T. (Eds.), Proc. ODP, Init. Repts, 154*, 5–10. College Station, TX.

———. 1995b. 'Site 926'. In *Curry, W., Shackleton, N., Richter, C., Bralower, T. (Eds.), Proc. ODP, Init. Repts, 154*, 153–232. College Station, TX.

Sinnesael, M., D. De Vleeschouwer, C. Zeeden, S. J. Batenburg, A.-C. Da Silva, N. J. de Winter, J. Dinarès-Turell, et al. 2019. 'The Cyclostratigraphy Intercomparison Project (CIP): Consistency, Merits and Pitfalls'. *Earth-Science Reviews* 199 (December): 102965. https://doi.org/10.1016/j.earscirev.2019.102965.

Sinnesael, M., M. Zivanovic, D. De Vleeschouwer, and P. Claeys. 2018. 'Spectral Moments in Cyclostratigraphy: Advantages and Disadvantages Compared to More Classic Approaches'. *Paleoceanography and Paleoclimatology* 33 (5): 493–510. https://doi.org/10.1029/2017PA003293.

Sinnesael, M., M. Zivanovic, D. De Vleeschouwer, P. Claeys, and J. Schoukens. 2016. 'Astronomical Component Estimation (ACE v.1) by Time-Variant Sinusoidal Modeling'. *Geosci. Model Dev.* 9 (10): 3517–31. https://doi.org/10.5194/gmd-9-3517-2016.

Sprovieri, M., N. Sabatino, N. Pelosi, S. J. Batenburg, R. Coccioni, M. Iavarone, and S. Mazzola. 2013. 'Late Cretaceous Orbitally-Paced Carbon Isotope Stratigraphy from the Bottaccione Gorge (Italy)'. *Palaeogeography, Palaeoclimatology, Palaeoecology* 379–380 (June): 81–94. https://doi.org/10.1016/j.palaeo.2013.04.006.

Stallone, A., A. Cicone, and M. Materassi. 2020. 'New Insights and Best Practices for the Successful Use of Empirical Mode Decomposition, Iterative Filtering and Derived Algorithms'. *Scientific Reports* 10 (1): 15161. https://doi.org/10.1038/s41598-020-72193-2.

Stock, J. H., and M. W. Watson. 1988. 'Variable Trends in Economic Time Series'. *Journal of Economic Perspectives* 2 (3): 147–74. https://doi.org/10.1257/jep.2.3.147.

Sun, Y., S. C. Clemens, F. Guo, X. Liu, Y.Wang, Y. Yan, and L. Liang. 2021. 'High-Sedimentation-Rate Loess Records: A New Window into Understanding Orbital- and Millennial-Scale Monsoon Variability'. *Earth-Science Reviews* 220 (September): 103731. https://doi.org/10.1016/j.earscirev.2021.103731.

Torrence, C., and G. P. Compo. 1998. 'A Practical Guide to Wavelet Analysis'. *Bulletin of the American Meteorological Society* 79 (1): 61–78. https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2.

Tukey, J. 1967. 'An Introduction to the Calculations of Numerical Spectral Analysis'. In *Advanced Seminar on Spectral Analysis of Time Series*. NY: John Wiley & Sons.

Vautard, R., and M. Ghil. 1989. 'Singular Spectrum Analysis in Nonlinear Dynamics, with Applications to Paleoclimatic Time Series'. *Physica D: Nonlinear Phenomena* 35 (3): 395–424. https://doi.org/10.1016/0167-2789(89)90077-8.

Waltham, D. 2015. 'Milankovitch Period Uncertainties and Their Impact On Cyclostratigraphy'. *Journal of Sedimentary Research* 85 (8): 990–98. https://doi.org/10.2110/jsr.2015.66.

Weedon, G. P. 2003. *Time-Series Analysis and Cyclostratigraphy: Examining Stratigraphic Records of Environmental Cycles*. Cambridge: Cambridge University Press.

Westphal, H., J. Lavi, and A. Munnecke. 2015. 'Diagenesis Makes the Impossible Come True: Intersecting Beds in Calcareous Turbidites'. *Facies* 61 (2): 3. https://doi.org/10.1007/s10347-015-0427-7.

Wilkens, R. H., T. Westerhold, A. J. Drury, M. Lyle, T. Gorgas, and J. Tian. 2017. 'Revisiting the Ceara Rise, Equatorial Atlantic Ocean: Isotope Stratigraphy of ODP Leg 154 from 0 to 5 Ma'. *Climate of the Past* 13 (7): 779–93. https://doi.org/10.5194/cp-13-779-2017.

Wouters, S., A.-C. Da Silva, F. Boulvain, and X. Devleeschouwer. 2021. 'StratigrapheR: Concepts for Litholog Generation in R'. *The R Journal*. https://journal.r-project.org/archive/2021/RJ-2021-039/index.html.

Wu, Z., and N. E. Huang. 2009. 'Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method'. *Advances in Adaptive Data Analysis* 01 (01): 1–41. https://doi.org/10.1142/S1793536909000047.

———. 2010. 'On the Filtering Properties of the Empirical Mode Decomposition'. *Advances in Adaptive Data Analysis* 02 (04): 397–414. https://doi.org/10.1142/S1793536910000604.

Wu, Z., N. E. Huang, S. R. Long, and C.-K. Peng. 2007. 'On the Trend, Detrending, and Variability of Nonlinear and Nonstationary Time Series'. *Proceedings of the National Academy of Sciences* 104 (38): 14889–94. https://doi.org/10.1073/pnas.0701020104.

Zeebe, R. E., and L. J. Lourens. 2019. 'Solar System Chaos and the Paleocene–Eocene Boundary Age Constrained by Geology and Astronomy'. *Science* 365 (6456): 926–29. https://doi.org/10.1126/science.aax0612.

Zeeden, C., F. J. Hilgen, S. K. Hüsing, and L. L. Lourens. 2014. 'The Miocene Astronomical Time Scale 9–12 Ma: New Constraints on Tidal Dissipation and Their Implications for Paleoclimatic Investigations'. *Paleoceanography* 29 (4): 296–307. https://doi.org/10.1002/2014PA002615.

Zeeden, C., S. R. Meyers, F. J. Hilgen, L. J. Lourens, and J. Laskar. 2019. 'Time Scale Evaluation and the Quantification of Obliquity Forcing'. *Quaternary Science Reviews* 209 (April): 100–113. https://doi.org/10.1016/j.quascirev.2019.01.018.

Zeeden, C., S. R. Meyers, L. J. Lourens, and F. J. Hilgen. 2015. 'Testing Astronomically Tuned Age Models'. *Paleoceanography* 30 (4): 369–83. https://doi.org/10.1002/2014PA002762.

Zeeden, C., F. Hilgen, T. Westerhold, L. Lourens, U. Röhl, and T. Bickert. 2013. 'Revised Miocene Splice, Astronomical Tuning and Calcareous Plankton Biochronology of ODP Site 926 between 5 and 14.4Ma'.

Palaeogeography, Palaeoclimatology, Palaeoecology 369 (January): 430–51. https://doi.org/10.1016/j.palaeo.2012.11.009.

Zeng, K., and M.-X. He. 2004. 'A Simple Boundary Process Technique for Empirical Mode Decomposition'. In . Anchorage, AK, USA: IEEE. https://doi.org/10.1109/IGARSS.2004.1370076.

Zeng, Y., S. Chen, S. Yang, Y. Liang, and Y. Wang. 2019. 'Multiscale Analysis of Asian Monsoon over the Past 640 Ka'. *Science China Earth Sciences* 62 (5): 843–52. https://doi.org/10.1007/s11430-018-9322-0.