Answers to comments from Christian Zeeden

Christian Zeeden (CZ): the La Thure series shows both precession and obliquity. Could you exemplary discuss what the result from your test means for this example record, and how it aids the interpretation?
>> The authors: This example is indeed interesting because both obliquity and precession have been observed (De Vleeschouwer et al., 2015), with obliquity having higher powers than precession. Several studies suggested that a dominance of obliquity in tropical sediments reflect cooling or icehouse conditions, while dominance of precession would be associated to greenhouse conditions (Zachos et al., 2001; Westerhold and Röhl, 2009; Boulila et al., 2011). With the implementation of our test we show that precession nearly vanishes at 15% of uncertainty. As a result, if one does not take into account this sampling bias on power spectrum one can misleadingly interpret a dominance of obliquity in sediments, which impacts in turn on the climatic interpretations.

CZ: explain what the Nyquist frequency represents.
>> The authors: The Nyquist frequency is the highest frequency (or smallest period) that can be detected. It corresponds to the inverse of twice the sample step. This information will be added in the next version of the manuscript.

>> Done. A short definition of the Nyquist frequency is provided lines 224-225 (line numbers from the marked-up manuscript below).

CZ: you suggest uncertainty to be fully random. I propose to briefly discuss why you assume this – and what effect(s) systematic uncertainty may have.
>> The authors: We assume a fully random error, based on comparisons with actual data of sample distances repeatedly measured on the La Charce series. This comparison went as follows: In a first step, the thicknesses of the individual beds were measured and a lithologic log was drawn based on these measurements. In a second step, samples were taken from the studied section every 20 centimetres and the sample positions were indicated on the lithologic log. After this second step, we observed that the distances between two successive samples was not exactly 20 cm, but rather ranged from 10 cm to 30 cm, with an average of 19.7 cm and a standard deviation of 2.5 cm. This observation was made by comparing the expected stratigraphic position of the n^{th} sample (n x 20 cm) with the stratigraphic position of the bed the sample comes from in the lithologic log. The mismatch in sample position between the lithologic log and the bed from which the sample was taken can be quantified for every sample within the studied stratigraphic interval. We evaluated the distribution of every sample´s mismatch and observed a log-normal distribution. This observation is the basis for our suggestion to consider the stratigraphic uncertainty to be fully random.

In addition, the total thickness of the series was measured at 109,33 m. With this thickness we expected to take 547 samples. Instead, we took 555 samples. We thus have an error of 8 samples out of 555 samples, either 1.4% difference. This is of course much lower than the actual thickness measurement error for individual sample distances, which implies no systematic error.

We will briefly discuss the absence of systematic error in the revised manuscript

>> Done. Lines 94-102, we add a new condition for justifying the fact that no systematic error is made when measuring each sample distance independently from the previous measurements

CZ: Lines 117-120: 106-116m is the overall spread in section thickness. From a conceptual point of view I think that this spread can hardly directly be used to estimate uncertainty in sample distance, because you see a result of ~550 (gamma distributed) sample distances summed up. Several of these will be shorter and longer than 20 cm – so your relative uncertainty will probably be higher – or fully systematic.
>> The authors: The reviewer is right. On average, the error made to measure the entire section is lower than the error made to measure sample distances. On entire sections, systematic errors will have for consequence overestimate the thickness of certain parts of the section while other parts will be have underestimated thicknesses. Thus, the error made to measure the total thickness of the section will be lower than the distance between two successive points. Here, the error made to measure the total thickness of a section is rather used to provide a minimum amount of thickness uncertainty. It will be indeed very hard to do better on short distances than what is done on a long, average distance.

>> This answer does not need any change in the manuscript

CZ: Lines 143, 223: Do I understand correct that you interpolate all time series (also with spacing of ~0.2 m and ~0.38 m) at 0.01 m intervals? Is this necessary and useful, and does this oversampling influence your results?
>> The authors: When linearly interpolating at the average sample step of the original series, we can reduce the amplitude of the high frequencies, independently of the error made on measuring the sample distance (Hinnov et al., 2002). So we overinterpolated at 0.01m to not create this bias in the analysis. However, we acknowledge that this procedure results in an inflation of the AR-1 coefficient of the red-noise fit. In the revised version of the manuscript, we will linearly interpolate the series at the median sample distance, as also suggested by Linda Hinnov (the other referee). To limit the loss of power in the high frequencies, we designed an optimized interpolation scheme, that will be applied in the revised version of the manuscript. This optimized strategy will be based on the minimal average offset between the original sample positions and the interpolated sample positions.

>> Done. The optimized linear interpolation is now based of best-fit curve between the original time series and the time series that has been resampled at the mean sample distance of the original time series. Changes are now observed in the last four figures of the manuscript. An appendix has been added to detail the optimized linear interpolation, and the method has been updated in lines 227-232.

CZ: Lines 159-164: Your approach is good, but personally I would propose to also determine 95% confidence intervals of power by considering not only the average power spectrum from simulations. This may facilitate to compare (integrated) precession and obliquity power for paleoclimate studies.
>> The authors: This is a great idea! That will be applied in the next version of the manuscript

>> Done. The grey areas in the last four figures represent the 95% confidence intervals of power (Figs. 3 and 8-11).

CZ: 175ff: a table summarizing the results presented may be helpful in addition.
>> The authors: Another great idea to make the results clearer and present them in a concise form that will help the readers
>> Done. See Table 2

CZ: In Fig. 4 the confidence levels of the MTM and Lomb-Scargle spectra are different. I would propose to mention this in the figure caption.
>> The authors: That's true! We will explicitly mention that in the next version of the manuscript

>> Differences exist in the confidence levels between the MTM method and the Lomb-Scargle method due to different degrees of freedom between the two approaches. The 2π-MTM analysis has a degree of freedom of ~6 (Mann and Lees, 1996), while the Lomb-Scargle method has a degree of freedom of 2 (Schulz and Stattegger, 1997). Differences are also observed in the La Thure series, which exhibits a high autocorrelation coefficient value. Addition al information is added

CZ: 10, 13: maybe express Nyquist frequency as sampling interval to be clearer
>> The authors: This is another good idea to make the things clearer! Knowing the fact that the Nyquist frequency is twice the sample step, it is very easy to convert the percentage of the Nyquist frequency to number of sample steps. For instance, 20% of the Nyquist Frequency represents 10 times the sample step.
>> Done. See notably Figs. 3, 8-11 and Table 2

CZ: 15-17: "In addition, the simulations indicate that taking at least 6-10 samples per precession cycle should allow calculation of robust power spectra estimates in the Milankovitch band." – This is not limited to precession I think, what about a more general statement as 'In addition, the simulations indicate that taking at least 6-10 samples per cycle should allow calculation of robust power spectra estimates in the respective cycle band'?
>> The authors: The reviewer is right. This requirement is actually valid for shortest cycle to be analysed, whatever its origin and period (obliquity, eccentricity or solar cycles).

>> In the manuscript, we replaced 4-10 samples per precession to 4-10 samples per thinnest cycles of interest. The case of the precession is maintained as example for the case of studies focused on the Milankovitch band (e.g., lines 30, 34, 517, 543 in the abstract and conclusion).

CZ: 28-29: "In core sediments, uncertainties in the sample position are also observed when performing physical sampling at very high resolution or because of core expansion phenomena (Hagelberg et al., 1995)" – suggestion:

‘In cored sediments, uncertainties in the sample position are also observed when performing physical sampling at very high resolution or because of core expansion phenomena (Hagelberg et al., 1995) or imperfect coring (Ruddiman et al., 1987).’
>> The authors: We would like to thank the reviewer for this suggestion. It is indeed very important to say that core sections are not devoid of bias. We will rephrase as suggested.

>> Done (line 58).

CZ: 37-38: “In this study, we address this problem by quantifying the impact of such errors on the frequency, as well as the power of higher-frequency cycles.” ! the second part of this sentence (“the frequency, as well as the power of higher-frequency cycles”) may be ‘the frequency and power distributions’?
>> The authors: We thank the reviewer for this suggestion, which makes the sentence much clear. We will rephrase as suggested.

>> Done (lines 69-70)

CZ: 42-44: This sentence seems in contradiction to the last sentence of the abstract, more consistent phrasing may solve this.
>> The authors: I think the reviewer refers to this sentence: **“**Based on our results, we suggest that one should take at least ~10 measurements per high-frequency cycle in order to provide robust estimates of the power of the high-frequency cycles.” And that is in contradiction with the last sentence of the abstract in which we said 6-10 samples per thinnest cycle targeted are necessary to identify all necessary cycles in the band we wish to explore. The authors apologize for this inconsistency and we will change “~10 measurements” by “6-10 samples per highest-frequency cycle…” for more consistency with the abstract.

>> This answer does not need any change in the manuscript

CZ: 48: delete ‘correctly’?
64: remove ‘easily’
>> The authors: OK for both

>>Done

CZ: 98/99: could you mention that these are Devonian, and give a rough age as for the La Charce section?
>> The authors: OK for precising the ages of the sections. The ages of the La Thure section (Givetian, middle Devonian) are around 380 Ma (De Vleeschouwer and Parnell, 2014).

>> Done (lines 171-172)

CZ: 108: are the two brackets necessary?
>> The authors: Sorry for that misspelling. We will remove one of the brackets in the next version

>> Done (line 188)

CZ: 119/120: “with an average of 110.3 ± 5.1 m, and a relative uncertainty of 4.6%” I would propose to mention that the “5.1 m” and “uncertainty of 4.6%” are estimated from only three experiments, and that these are regarded as representative, but may not be actually.
>> The authors: The comment from Linda Hinnov in page C2, bullet point 4, perfectly illustrates your comment: their team measured the La Charce section twice and found 112 m and 132 m thicknesses, either an average of 122 ± 10 m. The uncertainty is (10/122*100) 8.2% of the total thickness of the series. So our estimate is based on published data, but according to the personal comments from Linda Hinnov, available online in the second referee comments, this can be larger from a team to another.

CZ: 146: maybe give also reference to the R package used (‘dplR’)
>> The authors: We will mention that Lomb-Scargle analyses have been done with the dplR in the next version.

>> Done (lines 238-239)

CZ: 155: "The confidence levels of the datasets were calculated before randomisation and directly plotted to the simulated spectra." I am unsure how this is meant, and I would suggest phrasing this more clearly.
>> The authors: The authors apologize for this unclear phrasing. That only means that we plotted in the 1,000 randomized spectra the AR1-confidence levels calculated in the original series to make easier the comparison of powers.

>> A comment from Linda Hinnov requested to calculate the confidence levels for each simulation, so this part has been updated to take into account the comment from the other reviewer (lines 252-253)

CZ: 160-164: "Pori: the power spectrum before randomization" – as you calculate this for individual frequencies, following may be more clear: 'Pori: power before randomization for a specific frequency', same for Pave (if I understand this correct).
>> The authors: The reviewer is right. The rigorous phrasing should be: Pori(*f*): power before randomization at frequency *f*. The same for the others. We will rephrase in the next version of the manuscript

>> Done (lines 264-265)

CZ: 172 "with 5% uncertainty" – maybe clarify as 'with 5% stratigraphic uncertainty'
>> The authors: This change will be done

>> Done (line 276)

CZ: 200: with "first frequency" 'lowest frequency' is meant I assume – could you clarify this?
>> The authors: The reviewer is right, and we now see that our phrasing was ambiguous because it depends if we read the spectrum from the left or from the right. The phrasing suggested by the reviewer should eliminate our ambiguous phrasing.

>> Done (throughout sections 6 and 7)

CZ: 205-211: Please make clearer that geological data usually have no precise frequencies, but frequency ranges. You mention this, but I am not sure if everyone will understand this easily.
>> The authors: We suggest to mention after the sentence line 210: "For instance, because of variations of the sedimentation rates, the sedimentary expression of the orbital cycles is not focalised on specific frequencies but rather expressed on ranges of frequencies"

>> Done (lines 327-329)

CZ: 217: I am not sure if you need to mention "that the stratigraphic order of the samples in the raw series is preserved after randomisation" again. You develop this earlier in the manuscript.
>> The authors: We agree that this statement has been repeated and is superfluous in this line. We will delete this statement at 217 in the next version of the manuscript.

>> Done

CZ: 220f: "This difference realistically simulates small thickness errors, which accumulate when measuring successive sample steps." – this can in my opinion be formulated better, and should highlight that errors may accumulate, or may also not accumulate but level out.-
>> The authors: We suggest the following sentence to rephrase: "this difference is interpreted as the simulation of small thickness measurement errors, which accumulate when measuring successive sample steps".

>> Done. We finally do not need to add this piece of information since we precised that error model in random and not systematic (lines 94-102).

CZ: 241: "above 40% of the Nyquist frequency", I would suggest to also mention the frequency, maybe in brackets after this statement. Maybe bring these ratios in direct reference to precession (e.g. _1/3rd of precession frequency/wavelength), so that this is more clear for readers not so familiar with time series analysis.
>> The authors: ok for both suggest, and we will also precise the equivalent in terms of number of sample steps, as suggested in a previous comment from the reviewer.

CZ: 258/59: "As in the case of the La Charce series, the stratigraphic order of the samples is preserved in the randomised series" – In my opinion this is clear by now in the manuscript, and does not need to be repeated.

>> The authors: OK for removing this piece of information

>> Done

CZ: 304: replace "powers" by "power"
>> The authors: OK

>> Done

CZ: 310: "result suggest" – one of these need an "s" in the end
>> The authors: result needs an "s" in the end

>> Done

CZ: 312/13: "This requires that more than 6 samples per precession cycle have to be taken" - samples or measurements?
>> The authors: The authors see the ambiguity and regret it. We are talking about number of samples to take per precession cycles. This will be clarified in the manuscript

CZ: 355: 356: maybe also refer to (Meyers, 2015; Shackleton et al., 1995)
>> The authors: OK for adding the references

>> Done (line 598)

CZ: 396: "on the field" – in the field?
>> The authors: The correct expression is "in the field". Will be corrected in the next version

>> Done (lines 585, 632)

References :

Boulila, S., Galbrun, B., Miller, K.G., Pekar, S.F., Browning, J.V., Laskar, J., Wright, J.D., 2011. On the origin of Cenozoic and Mesozoic "third-order" eustatic sequences. Earth-Science Reviews 109, 94-112.

De Vleeschouwer, D., Parnell, A.C., 2014. Reducing time-scale uncertainty for the Devonian by integrating astrochronology and Bayesian statistics. Geology 42 (6), 491-494.

De Vleeschouwer, D., Boulvain, F., Da Silva, A.C., Labaye, C., Claeys, P., 2015. The astronomical calibration of the Givetian (Middle Devonian) timescale (Dinant Synclinorium, Belgium), from : Da Silva, A.-C., Whalen, M.T., Hladil, J., Chadimowa, L., Chen, D., Spassov, S., Boulvain, F., Devleeschouwer, X. (Eds.): Magnetic Susceptibility Application: A Window onto Ancient Environments and Climatic Variations. Geological Society, London, Special Publications 414, 245 – 256.

Hinnov, L.A., Schulz, M., Yiou, P., 2002. Interhemispheric space–time attributes of the Dansgaard–Oeschger oscillations between 100 and 0 ka. Quaternary Science Reviews 21, 1213-1228.

Westerhold, T., Röhl, U., 2009. High resolution cyclostratigraphy of the early Eocene – new insights into the origin of the Cenozoic cooling trend. Climate of the Past 5, 309-327.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present. Science 292, 686-693.

Answers to comments from Linda Hinnov

2. The error model

Linda Hinnov (LA): The authors call on the gamma probability density distribution to characterize stratigraphic sampling. Here there could be more explanation, e.g., a simple illustration of the problem, i.e., in Figure 1 add a diagram of a hypothetical stratigraphic section, different sampling sequences, and their histograms – perhaps the same ones as presented in Figure 2);

>> The authors: We thank the reviewer for this interesting suggestion that will help the reader to understand the problem. We actually have prepared a figure to illustrate the problem showing a hypothetical series with positions of samples obviously non-equally spaced. The diagrams used in real examples will be reused here, as suggested

>> Done: new Figure 1 illustrates the problem

LA: in Figure 1 caption indicate "gampdf(x, k, θ)" and label horizontal axis as "x". The models presented in Figure 2 displayed in F, G and H: what values of k and θ do these correspond to?

>> The authors: *k* and θ can be easily calculated using equations (3) and (4) of the manuscript. The mean sample distance is 1 unit in this case, and we performed the gamma test using setting the standard deviation at 0.05, 010 and 0.15 units respectively. In the 3 cases, k and θ values are as follows:

- Sd=0.05 units: θ=0.0025 and k=400
- Sd=0.10 units: θ=0.01 and k=100
- Sd=0.15 units: θ=0.0225 and k=44

This piece of information will be added in the revised version of the manuscript

>> Done (new Figure 3)

4. Implementation of the models in the stratigraphic-uncertainty tests

LA: This reviewer can personally attest to the difficulty in measuring a consistent thickness for the same outcrop by different researchers - in my experience in one case: 112 m vs. 132 m! For overturned sections, any dip error committed will contribute to a positive bias in stratigraphic thickness measurements. There is undoubtedly such a problem in the steeply dipping Cretaceous section at La Charce examined in this paper.

>> The authors: We thank the reviewer for this comment that was reused in the answer to referee 1. This example supports our idea that reaching a constant sample step on geological data is not trivial.

>> This discussion does not need revision in the manuscript

LA: On issues concerning methods, it is important to restrict interpolation to mean or median rate when applying AR noise models with MTM spectra (such as used in SSA-MTM Toolkit). The Devonian section has a mean sample rate of 0.38 m – not clear what the median rate is – and this is much larger than the interpolation to 0.01 m. The Cretaceous section has a mean sample rate of 0.20 m, so has a similar problem. The authors should recalculate the MTM analysis with interpolation to the median sample spacing of the two sections. (The red noise spectra will be significantly different because of the way the autocorrelation lag-1 coefficient is calculated.) The other parameter that requires reporting is whether "log" or "linear" fitting was enabled in the calculation of robust red noise for the MTM spectra.

>> The authors: The other reviewer (Christian Zeeden) has also commented on the overinterpolation procedure. Basically, we will provide a new method of interpolation, in order to optimize this step and limiting the loss of power in the high frequencies, that naturally occurs when resampling at the mean sample distance (see Hinnov et al., 2003).

>> Correction done, interpolation is now optimized in order to limit the loss of powers in the high frequencies. Data are linearly interpolated at the average sample distance of the original dataset (lines 227-232 + Appendix A)

As for the comment on the linear or log-fit, we employed a linear fit, from Meyers' astrochron ML96 function. In this function, the method for calculating the background median smooth fit has been modified by entering a Tukey's robust end point rule for the very low frequencies, which allows the level of lag-1 coefficient to be increased. This is below what the help of mtm.ML96 function says:

"This function conducts the Mann and Lees (1996; ML96) "robust red noise" analysis, with an improved median smoothing approach. The original Mann and Lees (1996) approach applies a truncation of the median smoothing window to include fewer frequencies near the edges of the spectrum; while truncation is required, its implementation in the original method often results in an "edge effect" that can produce excess false positive rates at low frequencies, commonly within the eccentricity-band (Meyers, 2012).

To help address this issue, an alternative median smoothing approach is applied that implements Tukey's robust end-point rule and symmetrical medians (see the function runmed for details)."

>> The median-fit of the red-noise background of La Thure was previously based on a "linear" comparison of powers. We now used a comparison of log-powers, much consistent with the red-noise background of La Thure with the REDFIT analysis (new Figures 6c, 10).

5. Application to a sum of sinusoids

LA: This section quantifies the loss of power at high frequencies with increasing uncertainty of (variability in) the sample step sequence for a simulated sum-of-sinusoids series. The absence of windowing in the Lomb-Scargle (LS) spectra would be expected to result in higher spectral variance compared to multitaper-windowed MTM spectra, and may account for the elevated grey spectra from the LS Monte Carlo simulations (compared to those of the MTM spectra). Interestingly, for 10% and 15% σ, loss of power occurs at practically the same frequencies in both MTM and LS spectra. Would it be possible to indicate the expected variance in Nyquist frequency for the 3 cases (5%, 10%, 15%) in order to understand the accuracy of the MTM and LS spectra? A new order of the graphs in Figures 2 and 3 might benefit the presentation:

- **New Figure 2: display Figs. 2F, G, H only, and explain how these relate to k and θ (or put them into a Figure 1B).**
- **New Figure 3: in top row, display Fig. 2A, B, C, D; bottom row display Fig. 3A, B, C, D.**
- **Figs. 2E and 3E could be placed into a new figure.**

>> The authors: We will modify the figures as suggested and ask the reviewer to further clarification about the variance question.
>> Figures modified (new Figures 3 and 4)
>> As for the variance question, the reviewer is right; the grey-spectra level is much elevated in the Lomb-Scargle analysis than in the MTM analysis. Following the recommendation of Christian Zeeden, we now display in Figure 3 and from Figures 8 to 11 the 95% zone of the 1000 simulations, which we assume to represent the 2σ-interval of the power spectrum estimates. We think than displaying the curve is much meaningful than the value of variance at the Nyquist frequency.

LA: What did we learn from this exercise and how will it help with the interpretation of the geological datasets to follow?

>> The authors: This exercise is performed on a pure sinusoid signal, not related to any geological data, and having an arbitrary sample step. It shows the general pattern of disturbing the sampling interval on the power spectrum,

independently of the nature of the geological data (finite length, noisy and non-strictly periodic). In this case, the power spectrum can be controlled and be fixed as equal for all spectral peaks, which helps to examine the relative change in power throughout the spectrum.

>> In the sum of sinusoid case, we now emphasise the loss of power in the high frequencies, since the power spectrum of spectral peak can be controlled. In the real geological example, we rather emphasise on the loss of significance level in the high frequencies, which is a direct consequence of the loss of power in the high frequencies, and which has the most implications for matching sedimentary series to insolation series.

6. Application to geological datasets

LA: The MTM spectrum of the Devonian series (Figure 4D) shows a robust red noise model with extremely elevated low frequencies, implying that a "log" fit was calculated in SSAMTM Toolkit, and that the model suffers additionally from the 0.01 m interpolation (see comments for Section 4). Some of the text in this section about differences in red noise calculations (which by the way are not meaningfully explained) may not be needed once the interpolation problem is addressed.

>> The authors: To calculate the spectrum of the La Thure section with the confidence levels, we used the mtm.ML96 function from astrochron package. A linear model of background fit was used (please find below the code line we applied:

```
ML96_1    = mtmML96(dat_pad1,tbw=2,ntap=3,padfac=1,demean=T,detrend=T,medsmooth=0.2,
            opt=3,linLog=1,siglevel=0.95,output=1,CLpwr=T,xmin=0,xmax=1/(2*dtmoy),
            sigID=F,pl=2,genplot=F,verbose=F)
```

linLog=1 means we used a linear fit model.

Definitely, we will re-fix the resample step at the median sample distance, which is 0.30 m.

>> The spectral background of the La Thure section is now calculated based on a comparison of log power, instead of linear powers previously. The results are much consistent with the red-noise fit calculated with the REDFIT method.
>> We decided to fix the sample step at 0.38 m for La Thure, which corresponds to the average sample distance of the series. This choice is motivated by the fact that only 30% of the series is sampled at a density equal or higher than the median sample distance, while 43% of the series is sampled at a density equal or higher than the average sample distance.

7. Discussion

LA: The main point of this study is that sampling is the critical decision that must be made when evaluating a stratigraphic sequence for paleoclimate signals. Almost all problems can be controlled with high-density sampling, e.g., 6-10 samples per putative precession cycle. It appears that one can easily expect 5% errors in stratigraphic position measurements, which combined with sedimentation rate variations, will mix the highest frequencies of a sampled sequence. Thus we are always alarmed at how low in power – and misaligned – precession cycles are in stratigraphic spectra. In the end, one never knows if a sample that has been collected has been assigned to its true stratigraphic position. This is an important limitation that is under-appreciated by the geological community and the authors should be commended for tackling this problem.

>> The authors: We thank the reviewer for this very positive comment, which will probably feed the discussion of the revised version of the manuscript.
>> This was implemented in the discussion from lines 601-624.

LA: A number of issues have been left unexplored: (1) how does systematic sample position error, such as can occur with receded marls alternating with prominent limestones in outcrops, affect stratigraphic spectra; (2) can astronomical tuning bypass the positional uncertainty problem (notwithstanding the recent approach described in Zeeden et al., 2015); and (3) how does the positional uncertainty problem affect the red noise model estimates?

>> The authors: For question (1), we have shown to the other reviewer that the error is not systematic but fully random, even in the case of alternating sedimentation.
>> For question (2), the error in the sample position acts on average like variations of the sedimentation rate: it decreases the power spectrum in the high frequencies, and distributes the power of the obliquity and precession over a large range of frequencies. The approach of Zeeden et al (2015) applies a very wide bandpass filter which should limit the effect of such error, because a large of frequencies are taken into account in the filter. However, we acknowledge that in spectra of sedimentary series, it is common to observe a band of frequencies between the obliquity and precession for which we don't know if they are related to one or the other cycle. A combination of methods involving wide filters and evolutive spectral analysis should help in resolving this issue.
>> For question (3), this is an interesting question, and actually at this point, we do not have the answer to this question. However, this could be the topic of a follow-up study.
>> For question (3), we have included in the new manuscript the calculations of the red-noise confidence levels for each simulation. After 1000 simulations, the red-noise confidence levels are on average very similar to the confidence levels before randomisation and show very narrow dispersion. This stability is probably due to the fact that the sample distance randomisation implies a dispersion of the power spectrum on a broad spectral band. The fit of the spectral background being calculated on a broad band, the randomisation procedure does not change the average power calculated on a broad band. Implemented in Figs. 8-11, Table 1, and discussed from lines 601-608.

Other comments

LA: Lines 23-24: The Multi-Taper Method (Thomson, 1982) might be more accurately characterized as a spectrum estimator that is based on the Fourier Transform – not as a derivative of the Fourier Transform.
>> The authors: The reviewer is right. The multi-taper method is roughly the average of Fourier Transforms of the series studied weighted by windows called Slepian sequences.

LA: Line 27: A recent massive improvement to the Jacob's staff in outcrop studies is terrestrial laser scanning with precision positioning at the mm level (Franceschi et al., 2011; Franceschi et al., 2015).
>> The authors: We thank the reviewer for having provided us with this reference.

LA: Line 101: Change to "Pas et al., 2015".
>> The authors: OK

LA: Lines 264- 265: what does the output of the "long-term trend of the variance" look like, and what was used to compute the "LOWESS regression with a 10% coefficient"?
>> The authors: The figure of the detrend procedure will be added in the next version of the manuscript, and

LA: Line 350: For monotonous stratigraphy yielding Milankovitch signal see also Latta et al., 2006.
>> The authors: This reference is elder than the one we have cited in the original manuscript. OK to add this citation.
>> Done (line 639).

LA: Line 355: Change "require" to "requires"
>> The authors: OK

LA: Line 361: Delete "Note than".
>> The authors: OK

LA: Supplementary File: R package dplR appears to be used but not referenced in the main text. R is used to calculate REDFIT– is it provided in the dplR package?
>> The authors: The authors are really sorry for having forgotten to cite dplR package. The other referee, Christian Zeeden, has also noticed that. As we cited the astrochron package from Stephen Meyers, we also have to cite dplR package, which will be done in the revised version of the manuscript.
>> Done (lines 238-239).

Testing the impact of stratigraphic uncertainty on spectral analyses of sedimentary series

Mathieu Martinez[1], Sergey Kotov[1], David De Vleeschouwer[1], Damien Pas[2], Heiko Pälike[1]

1: MARUM – Centrum for Marine Environmental Sciences, Leobenerstr., Universität Bremen, D-28359, Germany.
2: Pétrologie sédimentaire, B20, Géologie, Université de Liège, Sart Tilman, B-4000 Liège, Belgium.

Abstract

Spectral analysis is a key tool for identifying periodic patterns in sedimentary sequences, including astronomically related orbital signals. While most spectral analysis methods require equally-spaced samples, this condition is rarely achieved either in the field or when sampling sediment core. Here, we propose a method to assess the impact of the uncertainty or error made on the measurement of the sample stratigraphic position on the resulting power spectra. We apply a Monte-Carlo procedure to randomise the sample steps of depth series using a gamma distribution. Such a distribution preserves the stratigraphic order of samples, and allows controlling the average and the variance of the distribution of sample distances after randomisation. We apply the Monte-Carlo procedure on two geological datasets and find that gamma distribution of sample distances completely smooths the spectrum at high frequencies and decreases the power and significance levels of the spectral peaks in an important proportion of the spectrum. At 5% of stratigraphic uncertainty, few portion of the spectrum is completely smoothed. Taking at least 3 samples per thinnest cycle of interest should allow this cycle to be still observed in the spectrum, while taking at least 4 samples per thinnest cycle of interest should allow its significance levels to be preserved in the spectrum. At 10 and 15% uncertainty, these

thresholds increase, and taking at least 4 samples per thinnest cycle of interest should allow the targeted cycles to be still observed in the spectrum. In addition, taking at least 10 samples per thinnest cycle of interest should allow their significance levels to be preserved. For robust applications of the power spectrum in further studies, we suggest to provide a strong control of the measurement of the sample position. A density of 10 samples per putative precession cycle is a safe sampling density for preserving spectral power and significance level in the Milankovitch band. For lower density sampling, the use of gamma-law simulations should help in assessing the impact of stratigraphic uncertainty in the power spectrum in the Milankovitch band. Gamma-law simulations can also model the distortions of the Milankovitch record in sedimentary series due to variations in the sedimentation rate.

1. Introduction

Spectral analysis methods have become a key tool for identifying Milankovitch cycles in sedimentary series and are a crucial tool in the construction of robust astronomical time scales (Hinnov, 2013). The climatic or environmental proxy series that form the subject of spectral analyses are generally the result of measurements, performed on rock samples collected from a sedimentary sequence, either in cores or in outcrop. Most of spectral analysis methods (Fourier Transforms and derivatives such as Multi-Taper Method) require equally-spaced depth- or time-series, which implies that samples need to be taken at a constant sample step (Fig. 1a). Unfortunately, this is rarely achieved, especially for sedimentary sequences sampled in outcrops (e.g., Figs. 1b-c and e). Often, an uncertainty of ~5-15% is observed in the thickness or distance measurements, even when using a Jacob's staff (Weedon and Jenkyns, 1999). In core sediments, uncertainties in the sample position are also observed when performing physical sampling at very high resolution or because of core expansion phenomena (Hagelberg et al., 1995) or imperfect coring (Ruddiman et al., 1987).

Although uncertainties exist on the actual position of samples, few case studies document their effect on the identification of periodic patterns. Moore and Thomson (1991) recognised that perturbations of the regular sampling scheme (i.e. jittered sampling) impact the power spectrum by reducing spectral power in the high frequencies. Huybers and Wunsch (2004) and Martinez and Dera (2015) address an analogous problem, by assessing the effect of the uncertainty on the age model on a calibrated time series that is plotted against numerical age. However, none of these studies explicitly addresses the impact of errors in the measurement of the sample position on uncertainties in the power spectrum amplitudes. In this study, we address this problem by quantifying the impact of such errors on the frequency and power distribution. Therefore, we provide a new procedure that is based on a Monte-Carlo approach for randomising the distance between two successive samples in a sedimentary series. The resulting simulated series are subsequently used to assess the impact of the sample-position error on spectral analyses. We first apply the procedure to a theoretical example, and then to two previously published geological datasets, one as-regularly-as-possible sampled and another irregularly sampled.

2. The error model

In this paper, the term "stratigraphic uncertainty" refers to the uncertainty on the sample positions. Testing the impact of the stratigraphic uncertainty on the spectral analyses requires a randomisation procedure that reflects typical errors made during measurements of the stratigraphic position of samples. Figures 1c to e illustrates the consequences of the stratigraphic uncertainties on a geological series (here the La Charce series, see section 3.1). Fig. 1c compares the real sampling made on this series (in red) to an ideal sampling in which samples are taken at a strictly even sample distance (in black). Errors in the sample positions distort the sedimentary series: some intervals are compressed while some others are dilated. Ideally, all sample distances should be strictly the same, so that the distribution of sample distances should be concentrated on only one value (Fig. 1d). In reality, as uncertainties exist on the sample positions, the sample distances show a distribution over a certain range of values, which depends on the accuracy with which the distance measurements have been taken (Fig. 1e). In the case of the La Charce series, the standard deviation of the sample distances is assessed at 12.5% of the average sample distance (the method to estimate this standard deviation is provided in section 4). If the error in the distance measurement was systematic, one should expect the same level of error in the total length of the series. However in total, the difference of the length of the series between the ideal case (all sample distances strictly the same) and the real case is only 1.4% of the total length of the series (Fig. 1c). Each sample distance is measured independently from the other sample distances, so that each measurement can overestimate or underestimate the real distance between two successive samples. The errors thus compensate each others, implying that the process at the origin of the error measurements is not systematic but random.

Three conditions must be respected to design the error model: (i) the stratigraphic order of sample is hard set and must not be changed by the randomisation process (e.g., Fig. 1c), (ii) the average and standard deviation of sample steps must be maintained during the randomisation process, (iii) the error model must be random. These conditions can be achieved if the error model randomises the sample distances rather than the sample positions. In that case, the probability density function should have a positive and continuous distribution (*i.e.* values obtained after randomisation are continuous and positive). In addition, the average sample step and the standard deviation of the distance between two successive samples are known and should be parameterized.

The gamma distribution respects all these conditions. The gamma distribution is continuous and has a positive support. Two parameters are used to define the shape of the distribution (k) and its range of values (Θ□. The mean (E) of the density of probability is defined as (Burgin, 1975):

$$E = k * \Theta \qquad (1)$$

and its variance (σ^2) as:

$$\sigma^2 = k * \Theta^2 = E * \Theta \qquad (2)$$

Both the mean (E) and the variance (σ^2) are known, as they correspond to the mean and variance of the sample steps, and they can be quantified in the field (see Section 4 for a discussion on the variance of sample steps). Therefore, k and Θ can be parameterized using the following relations:

$$\Theta = \frac{\sigma^2}{E} \qquad (3)$$

$$k = \frac{E}{\Theta} \qquad (4)$$

Various versions of gamma probability density functions are shown in Fig. 2. A high variance-to-mean ratio corresponds to a high Θ-parameter value compared to the value of the k-parameter. The resulting density probability function corresponds to an exponential probability function in the most severe and spectrum-destructive case. This distribution corresponds to sampling conditions during which no control was exerted on the stratigraphic position of samples, so that the uncertainty on the sample position is at a maximum. Obviously, this situation is not a realistic case to reflect geologic practice.

In the opposite case, a low variance-to-average ratio corresponds to a low Θ-parameter value compared to the value of the *k*-parameter. The resulting density probability function is close to a Gaussian curve, although bound on one side to 0, so that the curve has a positive support. This case corresponds to geologic sampling during which the position of each sample was carefully measured and reported with respect to the stratigraphic column. Nevertheless, even in this case, stratigraphic uncertainties exist, mainly because of outcrop or core conditions. Interestingly, this latter case has a similar distribution to the distribution of sample distances in the La Charce series (Fig. 1e). This illustrates that the gamma model is well adapted for simulating the errors made on the measurement of the sample distances.

3. The geological datasets

Two geological datasets from previously published papers were used here to assess the effect of stratigraphic uncertainty on power spectra.

3.1. Gamma-ray spectrometry from La Charce (Valanginian, Early Cretaceous)

A total of 555 gamma-ray spectrometry measurements were performed *in situ* on the La Charce section (Department of Drôme, SE France; Martinez et al., 2013, 2015). The section is composed of marl-limestone alternations that were deposited in a hemipelagic environment during the Valanginian and Hauterivian stages (~134-132

Ma ago, Early Cretaceous; Martinez et al., 2015). Detailed analyses of their clay mineralogical, geochemical, faunal contents allowed these alternations to be attributed to orbital climate forcing. Gamma-ray spectrometry measurements have been used to discriminate the precession, obliquity and 405-kyr eccentricity cycles (see Martinez et al., 2015).

Gamma-ray spectrometry measurements have been performed directly in the field with an as-regular-as-possible sample step of 0.20 m. Before each measurement, rock surfaces have first been cleaned from the reworked material and flattened to prevent any border effects that could affect the measurement value. Each measurement was performed using a SatisGeo GS-512 spectrometer, with a constant acquisition time of 60 seconds (more details are provided in Martinez et al., 2013).

3.2. Magnetic susceptibility from La Thure section (Givetian, Middle Devonian)

The second case-study consists of the 184-m-thick continuous early-Givetian to early-Frasnian sequence of the La Thure section (~383-380 Ma, Middle-Late Devonian; De Vleeschouwer and Parnell, 2014; De Vleeschouwer et al., 2015; Pas et al., 2016). The Givetian sequence is composed of bedded limestone, mainly deposited in a shallow-water rimmed-shelf characterised by a large set of internal and external rimmed-shelf environments (Pas et al. 2016). The overlying early Frasnian sequence is dominated by shale deposited in a siliciclastic drowned platform (Pas et al., 2015). The magnetic susceptibility data from the La Thure section, in combination with three other MS data sets from the Dinant Syncline in southern Belgium and northern France were used by De Vleeschouwer et al. (2015) to make an estimate of the duration of the Givetian Stage, and subsequently to calibrate the Devonian time scale (De

Vleeschouwer and Parnell, 2014). Spectral analysis of the MS data from the La Thure section revealed the imprint of different Milankovitch astronomical parameters, including eccentricity, obliquity and precession (Fig. 3c in De Vleeschouwer et al., 2015). A total of 484 samples were taken along the 184-m thick sequence, with an irregular sample step that varied between 20-45 cm, depending on outcrop conditions (average sample step: 38 cm). Magnetic susceptibility measurements were performed using a KLY-3S instrument (AGICO, noise level 2 x 10^{-8} SI) at the University of Liège (Belgium) (more details are provided in De Vleeschouwer et al., 2015).

4. Implementation of sample step uncertainty models for the stratigraphic-uncertainty tests

Weedon and Jenkyns (1999) estimated the error on the stratigraphic position of a sample at 5.3% by measuring the thickness of the same sequence twice. The La Charce section, one of the datasets treated here, has been measured multiple times in the framework of different publications. The thickness of the studied section was assessed at 106 m, 109 m and 116 m (Bulot et al., 1992; Martinez et al., 2013; Reboulet and Atrops, 1999) ~~either~~ an average of 110.3 ± 5.1 m, ~~or~~ an error of 4.6%. In the field, the distance between two successive samples was measured independently from the construction of the log, providing an independent assessment of the dispersion of the actual distance between two successive samples. The average sample step is 20 cm, with a standard deviation of the sample steps of 2.5 cm, ~~either~~ a level of uncertainty of 12.5% of the average sample step (Fig. 1e).

Based on the assessments summarised in the previous paragraph, we tested three different levels for the error on the measurement of sample steps (5%, 10% and 15%), which we consider realistic scenarios for geologic sampling during fieldwork. We applied our Monte-Carlo based procedure for randomising sample steps to a sinusoidal series, as well as to the two previously published geologic datasets described in section 3 (De Vleeschouwer et al., 2015; Martinez et al., 2013, 2015), with three different error levels. During every Monte-Carlo simulation, the distance between two points is randomised according to a gamma distribution, of which the mean corresponds to the distance between two points measured in the field, and of which the standard deviation corresponds to 5%, 10% or 15% of the measured distance. Each test consists of 1000 Monte-Carlo simulations, leading to 1000 different time series, each with a different distortion of the stratigraphic positions of samples.

Spectral analyses were performed using the Multi-Taper Method (MTM; Thomson, 1982, 1990), using three 2π-tapers (2π-MTM analysis) and with the Lomb-Scargle method (Lomb, 1976; Scargle, 1982). For the 2π-MTM analysis, confidence levels of the spectra of the original geological datasets tested have been calculated using the Mann and Lees (1996) approach (ML96), with median-smoothing calculated with the method of the Tukey's end point rule, as suggested by Meyers (2014). The window width for the median-smoothing was fixed at 20% of the Nyquist Frequency (the highest frequency which can be detected in a time series), as evaluated empirically by Mann and Lees (1996). MTM analysis requires strictly regular sample steps to be performed, so that geological datasets were linearly interpolated at the average sample distance of the original series before and after randomisation. We limit the loss of amplitude in the high-frequency fluctuations due to resampling by applying an optimized procedure to find the best starting point of the interpolated series. To our knowledge, this procedure is new, and we therefore describe it in Appendix 1. We provide the corresponding R-function in the supplementary material. The sum of sinusoid series is generated with a regular sample step of 1 arbitrary unit. After randomisation, the depth-randomised series was linearly interpolated at 1 arbitrary unit.

Lomb-Scargle spectra were calculated with the REDFIT algorithm (Schulz and Mudelsee, 2002) available in the R-package dplR (Bunn, 2008, 2010; Bunn et al., 2015). The Lomb-Scargle method calculates the spectrum of unevenly-sampled series. Lomb-Scargle power spectra can be biased in the high frequencies due to the non-independency of the frequencies (Lomb, 1976; Scargle, 1982), so that the REDFIT algorithm has been provided to correct the power spectrum by fitting a red-noise model to the spectrum (Mudelsee, 2002; Schulz and Mudelsee, 2002). Here, we applied no segmentation to the series and a rectangular window. This parameterization maximises the effect of sample step randomisation on the spectrum.

During each test, both MTM and REDFIT Lomb-Scargle power spectra were calculated for each of the 1000 Monte-Carlo distorted series. Subsequently, the average power spectra and the range of powers covered by 95% of the simulations were calculated for the MTM and Lomb-Scargle analyses. The confidence levels of the datasets deduced from the red-noise fit of the spectral background were calculated after each simulation. The average power of the confidence levels and the range of powers of the confidence levels covered by 95% of the simulations were calculated and directly plotted to the simulated spectra. The sum of sinusoids series does not need correction to red noise and the raw Lomb-Scargle spectra are shown. The two geological datasets show a red-noise background and the REDFIT-corrected Lomb-Scargle spectra were shown.

We finally provide a quantification of the relative change in spectral power, using the following criterion:

$$E_r(f) = abs\left(\frac{P_{ori}(f) - P_{ave}(f)}{P_{ave}(f)}\right) \qquad (5)$$

With f: the frequencies explored in the spectral analyses

E_r: the relative change of power

$P_{ori}(f)$: the power spectrum before randomisation at frequency f.

$P_{ave}(f)$: the average power spectrum of the 1000 simulations at frequency f.

5. Application to a sum of sinusoids

The effect of randomising the sample position on the section is first tested on a sum of pure sinusoids. A dataset of 600 points is generated with a sample step of 1 arbitrary unit. The series is a sum of 24 sinusoids, having equal amplitudes and different frequencies: frequencies range from 0.02 to 0.48 cycles/arbitrary unit and increase with increments of 0.02 cycles/arbitrary unit (Fig. 3a, b). Fig. 3 shows the 2π-MTM and Lomb-Scargle spectra of the sum of sinusoids before and after applying 1000 Monte-Carlo simulations of distorted sample distances. The grey zones indicate the interval covering 95% of the power in the 1000 simulations. The average spectrum of these simulations is shown in orange for the test with 5% stratigraphic uncertainty (Figs. 3c, d), red for 10% uncertainty (Figs. 3e, f), and brown for 15% uncertainty (Figs. 3g, h). The most striking feature after gamma-model randomisation is the progressive and strong decrease of the power spectrum towards the high frequencies, even when the lowest level of uncertainty (5%) is considered.

Fig. 4 notably shows the relative change in power of the average spectrum after applying the 1000 simulations. At 5% uncertainty, a decrease of 50% in the power spectrum is observed in the 2π-MTM spectrum at 57% of the Nyquist frequency, equivalent to 3.5x the average sample distance. The level of 50% of decrease in the power spectrum is rather observed in the Lomb-Scargle spectrum at 80% of the Nyquist frequency, *i.e.* 2.5x the average sample distance. This implies that even for a very low level of noise, the values of the power spectrum can be largely underestimated in the upper half of the spectrum. At 10% uncertainty, a decrease of power spectrum is observed at 38-39% of the Nyquist frequency, both in the Lomb-Scargle and the 2π-MTM spectra, which is equivalent to 5.2x the average sample distance. Finally, at 15% uncertainty, both Lomb-Scargle and 2π-MTM indicate that 50% of decrease in the power spectrum has occurred at 27% of the Nyquist frequency, which is equivalent to 7.4x the average sample distance. This example shows the worse ~~is~~ the control of the sample position in the sedimentary series the more ~~one needs to take~~ sample per cycle to limit the loss of power of the cycles targeted.

Stratigraphic uncertainty does not only trigger loss of power of the spectral peaks, it also increases the power spectral background (Fig. 3). At 5% and 10% uncertainties, the average and background spectrum still preserve the structure of individual peaks in both 2π-MTM and Lomb-Scargle analyses (Figs. 3c-f). Indeed, spectra for individual Monte-Carlo simulations still exhibit spectral peaks at these frequencies although they are characterized by variable power and deviations in the frequencies at which the peaks are localised. However, at 15% uncertainty, the average power at the highest frequencies is flattened and the structure of the peaks is not distinguishable anymore (Figs. 3e-h). This zone of the spectrum cannot be regarded as reliably interpretable.

These analyses from a sum of pure sinusoids show that the higher ~~is~~ the stratigraphic uncertainty, the higher is the loss in power of the spectral peaks and the more the low frequencies are affected by this loss of power. At 15% uncertainty, the spectrum is flattened in the highest frequency and cannot be interpreted in this part of the spectrum. Because of its higher frequency resolution, the Lomb-Scargle analysis, ~~as we computed, here~~ displays higher spectrum background levels than ~~in~~ the 2π-MTM analysis. It however changes very ~~few~~ the highest frequency that can be interpreted, even at 15% uncertainty.

It should be noticed that in the case of pure sinusoids, the signal is only composed of pure harmonics concentrating the spectral power at specific frequencies. This ~~makes~~ that a small shift in the sample position triggers a strong decrease of the average power spectrum at these specific frequencies. In addition, in this theoretical example, the sample before randomisation procedure was strictly constant (1 arbitrary unit). More realistically, spectra of geological datasets are rather composed of a mixture of harmonics, narrow-band and background components, and the sample step is not strictly constant. For instance, because of variations of the sedimentation rates, the sedimentary expression of the orbital cycles is not ~~focalised~~ on specific frequencies but rather expressed on ranges of frequencies (e.g., Weedon, 2003, p. 132). This can add some noise in the high frequencies, and blur the spectra even more than in the case of pure sinusoids. In the following, the results of the application of the test on two geological datasets are shown.

6. Application to geological datasets

6.1. Spectral analysis prior to randomisation

6.1.1. The La Charce series

Prior to performing 2π-MTM analyses, the gamma-ray series was detrended using a best-fit linear regression, linearly interpolated ~~each~~ 0.20 m, and standardised to zero average and unit variance (Fig. 5). Prior to REDFIT Lomb-Scargle analysis, the datasets (raw and randomised) were simply linearly detrended using a best-fit linear regression and standardised.

The 2π-MTM analysis of the La Charce section shows two main significant bands (>99% Confidence Level, hereafter abbreviated CL) at 20 m and from 1.3 to 0.8 m (Fig. 6a). The peak ~~of~~ 20 m has been interpreted as the imprint of 405-kyr eccentricity forcing, while the group of peak ~~of~~ 1.3 to 0.8 m has been dominantly related to precession (Boulila et al., 2015; Martinez et al., 2013, 2015). The REDFIT spectrum shows two bands of periods exceeding the 99% CL at 18 m and from 1.4 to 0.8 m (Fig. 6b). These periods are similar to the periods observed in the 2π-MTM spectrum. The small differences in periodicity observed in the lowest frequencies are likely to be related to the difference in frequencies explored between both methods. In addition, the REDFIT spectrum as parameterised here produces narrower peaks than the multi-taper spectrum, so that the lowest frequencies in the REDFIT spectrum are composed of a group of narrow peaks, rather than a single broad peak observed in the 2π-MTM spectrum.

The autoregressive coefficient, a measure for the redness of the spectrum, is assessed at 0.440 in the 2π-MTM analysis, while it is assessed at 0.468 in the REDFIT analysis (Table 1). The S0-value, the average power of the red-noise process within the entire spectrum, is 3.54×10^{-4} in the MTM analysis, while it is 0.398 in the REDFIT analysis (Table 1). This difference in the S0 value is due to the difference of signal treatment when calculating the MTM or the REDFIT spectrum.

6.1.2. The La Thure series

Prior to performing 2π-MTM analyses, the magnetic susceptibility series was detrended by subtracting a piecewise best-fit linear regression (Fig. 7a). The series was then linearly interpolated ~~each~~ 0.38 m, and the trend of the variance was removed by dividing the series by its instantaneous amplitude smoothed with a LOWESS regression with a 10% coefficient (Fig. 7b). Such approach allows the series to have a stationary mean and variance (Fig. 7c). The series was subsequently standardised (average=0; standard deviation=1). Prior to the REDFIT analysis, the identical procedure was applied, except for interpolation at an even sample step, as this is not required by the Lomb-Scargle method.

The 2π-MTM analysis of the La Thure section shows significant periods at 39 m (>99% CL) interpreted as the manifestation of the 405-kyr eccentricity cycle (De Vleeschouwer et al., 2015), at 7.8 m (>95% CL) interpreted as 100-kyr eccentricity cycles, a group of significant periods from 2.8 m to 2.2 m (99% CL) interpreted as obliquity, and a group of significant periods from 1.6 to 1.1 m (>95% and >99% CL) interpreted as precession (Fig. 6C). In the lowest frequencies, the REDFIT spectrum (Fig. 4F) shows a group of peaks centred on 30-40 m (>99% CL), a peak at 13 m (>95% CL), which is not significant in the 2π-MTM spectrum. Conversely, the period at 7.9 m observed in the 2π-MTM spectrum does not reach the 90% CL in the REDFIT spectrum. These differences are likely related to the difference in the frequencies explored between both methods, and to the fact that REDFIT spectra as parameterised here produce narrower peaks than the 2π-MTM spectra. In the REDFIT spectrum, the obliquity band shows two periods at 3.3 m (95% CL), and 2.3 m (>95% CL). The precession band shows periods at 1.5 m (>90% CL), 1.1 m (>99% CL) and at 0.9 m (>95% CL).

The autoregressive coefficient of the red-noise background level is assessed at 0.657 in the 2π-MTM analysis, and at 0.407 in the REDFIT analysis (Table 1). The difference in the autoregressive coefficient is due to the method of calculation of the red-noise background (from the spectrum in the MTM analysis, from the time series in the REDFIT analysis; Mann and Lees, 1996; Meyers, 2014; Mudelsee, 2002). The Lomb-Scargle analysis also tends to produce higher powers in the high frequencies, thus reducing the autoregressive coefficient estimate in the REDFIT analysis (Schulz and Mudelsee, 2002). This difference also illustrates the difficulty in calculating the autoregressive coefficient when the redness of the spectrum increases (see Meyers, 2012). Finally, the S0-value is assessed at 1.67×10^{-3} in the 2π-MTM analysis, and at 0.890 in the REDFIT analysis (Table 1).

6.2. Impact on the power spectrum of randomising the sample distances

6.2.1. The La Charce series

At 5% uncertainty, the average 2π-MTM spectrum of the La Charce still shows periods at 20.5 m as well as several periods around 1 m exceeding the 99% CL (Fig. 8a). At 10% uncertainty, the peak at 0.8 m does not exceed the 95% CL (Fig. 8b), and it is completely smoothed at 15% uncertainty (Fig. 8c). The increasing level of stratigraphic uncertainty progressively smooths the average spectrum, with the highest frequencies most affected (Figs. 8d-f). Notably at 5% uncertainty, fluctuations of the spectrum at frequencies higher than 81% of the Nyquist frequency are suppressed (Table 2). At 10% and 15% uncertainty, this threshold decrease to respectively 58 and 43% of the Nyquist frequency (Figs. 8d-f). Increasing levels of uncertainty also tend to reduce the power of the spectral peaks in an increasing portion of the spectrum. At 5% uncertainty, the average spectrum of the simulations is practically identical to the spectrum of the original series from frequency 0 to 27% of the Nyquist frequency (Fig. 8d). This range is reduced to 0 - 19% of the Nyquist frequency at 10% uncertainty (Fig. 8e) and to 0 - 18% of the Nyquist frequency at 15% uncertainty (Fig. 8f).

In the REDFIT spectrum with 5% of stratigraphic uncertainty, the periods at 20.5 m and around 1 m still exceed the 99% CL (Fig. 9a). Like in the 2π-MTM analyses, the period at 0.8 m does not exceed the 99% CL at 10% uncertainty, while it is completely smoothed at 15% uncertainty (Figs. 9b-c). The tendency of the Lomb-Scargle analysis to produce high-power peaks in the high frequencies limits the effect of the smoothing of the spectrum at 5% uncertainty (Fig. 9d). However, at 10 and 15% uncertainties, fluctuations in the spectrum at frequencies higher than respectively

58 and 42% of the Nyquist frequency are completely smoothed (Figs. 9e-f; Table 2). At 5% uncertainty, the average spectrum of the simulations cannot be distinguished from the spectrum of the original series from frequency 0 to 29% of the Nyquist frequency (Fig. 9d), while at 10 and 15% uncertainties, this range is restricted to 0 - 19% of the Nyquist frequency (Fig. 9e-f).

The average autoregressive coefficients of the 1000 simulations (with ± the interval covering 95% of the simulations) are ~~respectively~~ assessed for 5, 10, and 15% of stratigraphic uncertainties at 0.433 ± 0.025, 0.432 ± 0.037, 0.434 ± 0.048 in the 2π-MTM analyses, and at 0.468 ± 0.002, 0.467 ± 0.003, 0.467 ± 0.006 in the REDFIT analyses (Table 1). The average S0-values of the 1000 simulations are ~~respectively~~ assessed for 5, 10, and 15% of stratigraphic uncertainties at $3.55 \times 10^{-4} \pm 0.13 \times 10^{-4}$, $3.58 \times 10^{-4} \pm 0.20 \times 10^{-4}$, $3.61 \times 10^{-4} \pm 0.25 \times 10^{-4}$ in the 2π-MTM analyses, and at 0.399 ± 0.003, 0.402 ± 0.005, 0.407 ± 0.008 in the REDFIT analyses.

6.2.2. The La Thure series

At 5% uncertainty, the 2π-MTM spectrum of the La Thure series still exhibits significant frequencies at 39 m, 1.5 m and 1.1 m exceeding the 99% CL, and at 7.5 m, 2.9 m, 2.2 m and 1.6 m exceeding the 95% CL (Fig. 10a). At 10% uncertainty, the 1.1-m peak is much smoother, centred on a period of 1.2 m and only exceeds the 95% CL (Fig. 10b). The other periods of the precession at 1.5 and 1.6 m, only exceed the 90 and 95% CL, respectively. The significant periods of the obliquity bands, at 2.2 and 2.9 m show weaker powers than in the spectrum of the original series, but still exceed the 95% CL. At 15% uncertainty, the band of periods at 1.2 m is nearly entirely flattened and hardly distinguishable from the spectral background (Fig. 10c). In addition, all frequencies from the obliquity and the precession do not exceed the 95% CL. The reduction in the significance levels in the precession and obliquity bands is the consequence of increasing loss in power of the spectral peaks ~~in~~ high frequencies. At 5% uncertainty, the average spectrum of the simulations is ~~confounded~~ to the spectrum of the original series from frequency 0 to 52% of the Nyquist frequency (Fig. 10d), while at 10 and 15% uncertainties, this range is restricted to 0 - 20% of the Nyquist frequency (Figs. 10e-f; Table 2).

At 5% uncertainty, the REDFIT analysis still displays a significant period at 30-40 m exceeding the 99% CL, and a period at 2.3 m exceeding the 95% CL (Fig. 11a). The peak at 1.5 m does not exceed anymore the 90% CL, while the peaks at 1.1 m and 0.9 m do not exceed anymore the 95% CL. At 10% uncertainty and 15% uncertainties, spectral peaks in the precession and the obliquity bands do~~es~~ not reach the 95% CL anymore. The tendency of the Lomb-Scargle analysis to produce high-power peaks in the high frequencies prevents ~~from~~ strong smoothing of the power spectrum at 5% uncertainty. However, at 10 and 15% uncertainties, all fluctuations of the power spectrum at frequencies higher than 53% Nyquist frequency are flattened and not distinguishable (Table 2). The significance level in the eccentricity band is still preserved in the average spectrum. At 10 and 15% uncertainty, the power spectrum displays spectral peaks with reduced powers compared to the spectrum of the original series, which impacts the significance levels at the obliquity and precession bands (Figs. 11d-f). At 5% uncertainty the REDFIT spectrum of the La Thure series remains practically unchanged compared to the spectrum of the original series from 0 to 58%

Nyquist (Fig. 11d), while at 10 and 15% uncertainty this range is respectively restricted to 0 - 22% and 0 - 19% Nyquist frequency (Figs. 11e-f).

The average autoregressive coefficients of the 1000 simulations are ~~respectively~~ assessed for 5, 10, and 15% of stratigraphic uncertainties at 0.658 ± 0.025, 0.653 ± 0.029, 0.651 ± 0.033 in the 2π-MTM analyses, and at 0.406 ± 0.004, 0.405 ± 0.008, 0.404 ± 0.013 in the REDFIT analyses (Table 1). The average S0-values of the 1000 simulations are ~~respectively~~ assessed for 5, 10, and 15% of stratigraphic uncertainties at $1.67x10^{-3} \pm 0.04x10^{-3}$, $1.67x10^{-3} \pm 0.05x10^{-3}$, $1.68x10^{-3} \pm 0.07x10^{-3}$ in the 2π-MTM analyses, and at 0.894 ± 0.011, 0.900 ± 0.019, 0.904 ± 0.008 in the REDFIT analyses.

7. Discussion

7.1. Comparison of the results between the two geological datasets

In the 2π-MTM simulations, the spectral peaks tend to be smoothen at 5% of stratigraphic uncertainty from ~80% Nyquist frequency to the Nyquist frequency, which implies that taking at least 3 samples per cycle of interest should not smooth the spectral peaks in the frequency band targeted (e.g., the Milankovitch cycles) (Table 2). In the REDFIT simulations, the tendency of the spectrum to produce high-power spectra in high frequencies even makes all the spectral peaks of the original spectrum still identifiable at 5% uncertainty. If a low level of stratigraphic uncertainty is maintained, practically all spectral peaks at frequencies below 80% Nyquist frequencies will be preserved. These thresholds dramatically decrease to 53% to 66% of Nyquist frequencies at 10% of stratigraphic uncertainty in all simulations, while it decreases to 42% to 53% of Nyquist frequency at 15% uncertainty. Thus, a medium level of stratigraphic uncertainty implies taking at least 4 samples per cycles of interest, while a high level of uncertainty implies taking at least 5 samples per cycle of interest.

Comparisons between original and average simulated spectra show that at 5% uncertainty, both ~~could be confounded~~ from 0 to 27% of Nyquist frequency in the La Charce series and from 0 to 52% of Nyquist frequency in the La Thure series. At 10 and 15% uncertainties, these range dramatically shift from 0 to 20-22% Nyquist frequency. Although differences exist in the variance of the average spectrum and in the frequency resolution between the 2π-MTM and the REDFIT analyses, both analyses show, for each series, the same range of frequencies in which simulated and original spectra ~~could be confounded~~. These thresholds imply that taking 4-8 samples per cycle of interest should limit loss of power of the spectral peaks in the targeted bands at 5% uncertainty. At 10 and 15% uncertainty, taking at least 10 samples per cycle of interest should limit the loss of power in the targeted band. Limiting the loss of power in the frequencies of interest appears to be crucial because the average power of the confidence levels remain unchanged after applying the simulations. Simulations of distortions of the geological series smoothes the spectrum by distributing the power spectrum from the spectral peaks to the surrounding frequencies. The calculation of confidence levels in the MTM analyses is based on a moving median of the power spectrum performed over a broad range of frequencies (usually 1/5 of the total spectrum; Mann and Lees, 1996). Thus, when distorting the time series, the distribution of the power spectrum over a narrow range of frequencies does not change the overall median of the power spectrum calculated over 1/5 of the total spectrum, and thus does not change the average level of confidence levels after simulations. The effect of time-series distortions on the power of confidence levels is even smaller in the REDFIT analysis, in which the confidence levels are directly calculated on the time series itself and not on the spectrum (Mudelsee, 2002). The decrease of the power of the spectral peaks due to distortions of the geological series thus implies a decrease in the significance levels of the main cycles of the series. In case of low level of red noise, like in the La Charce series (Figs. 8-9), spectral smoothing and decrease in power in the precession band does not strongly impact the interpretations, since the significance level in the precession band still exceed the 99% CL, even after implementation of a level of 15% of stratigraphic uncertainty. However, in case of strong red noise, like in the La Thure series, the decrease of power in high frequencies have a strong impact on the significance levels after implementation of the simulations. At a medium level of stratigraphic uncertainty (10%), taking 10 samples per cycle of interest is needed to limit the loss of power in the cycles of interest and thus to limit the decrease in the level of significance of these targeted cycles.

As an example, if the targeted range of frequencies are the Milankovitch cycles, the shortest period of interest are the precession cycles. A density of 1 sample 4 kyr should allow the detection of the spectral peaks in the precession band. A density of sampling of 1 sample per 2 kyr should then ensure the detection of significant peaks in the precession band, even in case of strong red noise and medium-to-high levels of stratigraphic uncertainy. The minimum density of sampling being dependant of the level of red noise and stratigraphic uncertainty, we deeply recommend to apply the simulations developed here to assess the impact of stratigraphic uncertainty to the identification of significant spectral peaks in the sedimentary record.

7.2. In which case to apply this test?

Uncertainties in the measurement of sample position can practically not be avoidable in outcrop conditions. The similarity between the topographic slope and the sedimentary dip, the absence or scarcity of marker beds, or the need to move laterally in a section trigger disturbances in the regularity of the sampling. In core sedimentary sequences, non-destructive automation measurements such as X-ray fluorescence, gamma-ray spectrometry or magnetic susceptibility should ~~prevent from~~ errors in the sample position. However, physical samplings (e.g. for geochemistry or mineralogy) are subject to small uncertainties, especially when the sampling resolution is very thin. Core sedimentary series can in addition be affected by expansion of sediment caused by release of gas or release of overburden pressure (Hagelberg et al., 1995). This test is thus useful for geologists who wish to run spectral analyses on sedimentary depth-series generated from outcropping sections or core samples. All analyses in this paper show that with higher uncertainty on the sample step, the low frequencies are increasingly affected. The relative change in power between the various tests all showed different patterns, and no general model could be deduced. The relative change in power at a given frequency depends on the dispersion of the sample step, on the method of spectral analysis, but also on the original sedimentary sequence to study. Each depth-series generated from this sampling can be seen as one of the 1000 random simulations. The test randomises the sample position from the original series, and produces a smooth version of the spectrum of the raw series. The generation of the raw series impacts on the test at frequencies having low powers (a small change in a weak power can trigger high values of relative change in power), and at high frequencies. The relative change in power does not depend on the size of the sample step itself, as the same proportion of the spectrum is affected for a given level of uncertainty. However, a control on the dispersion of the sample steps and the application of the test proposed here are needed to assess the dispersion of the sample step during the sampling procedure and the impact of this dispersion on the spectrum. The question is how to assess the dispersion of the sample step in the field? If the section is well bedded, we suggest applying the same procedure as we did for La Charce, i.e. sample position measured independently from the bed thickness measurements, and precise report of the sample positions on the sedimentary log of the series. Orbital forcing can also be detected in a monotonous thick marly section, showing no apparent bedding (e.g., Ghirardi et al., 2014; Matys Grygar et al., 2014). In that case, we rather suggest measuring several times the total thickness of the sequence to assess the potential dispersion of the sample steps.

7.3. Implications for astronomical time scale and palaeoclimate reconstructions

Linking sedimentary cycles to orbital cycles or assessing the quality of an orbital tuning procedure often require a good matching between the sedimentary period ratios and the orbital period ratios (Huang et al., 1993; Martinez et al., 2012; Meyers and Sageman, 2007) and/or the determination of the amplitude modulation of the orbital cycles (Meyers, 2015; Moiroud et al., 2012; Shackleton et al., 1995; Zeeden et al., 2015). On average, stratigraphic uncertainties trigger a decrease of the power spectrum of the main significant frequencies while distributing the power spectrum to the surrounding frequencies. In the studied geological data, stratigraphic uncertainties mostly impact the precession band, by decreasing the power and significance levels of the spectral peaks and multiplying the main frequencies for each individual runs. The occurrence of low-power spectral peaks in the precession bands, and the fact that frequency ratios between the precession and lower frequencies does not match the orbital frequency ratios are quite common in the geological data (e.g., Ghirardi et al., 2014; Huang et al., 2011; Thibault et al., 2016), and can be a consequence of stratigraphic uncertainties.

Variations in the sedimentation rate produces a similar effect as stratigraphic uncertainties and can be modelled with the Monte-Carlo simulations applied in this study. As sedimentation rates always vary within a sedimentary series, any particular astronomical cycle can be recorded on various thicknesses of sediments, which in turn decreases the power of this astronomical cycle and distributes its power over a large range of frequencies (Weedon, 2003). Stratigraphic uncertainties add additional noise which blurs the spectra of sedimentary series at high frequencies.

Astronomical tuning can help in removing the effects of stratigraphic uncertainties and variations in sedimentation rates (e.g., Hays et al., 1976; Huang et al., 2011; Zeeden et al., 2013). The identification of the repetition of any astronomical cycle and their attribution to the same duration removes the effects of distortion of the sedimentary series, and concentrates the variance of the power over several frequencies. Filtering a band of frequencies of interest can help in identifying the repetition of the cycle used for the astronomical calibration (e.g., Westerhold et al., 2008; Thibault et al., 2016; De Vleeschouwer et al., 2015). Because of distortions of the sedimentary series, a filter, if designed very narrowly, can lead to a distortion of the actual amplitude and number of repetitions of the filtered frequency. This is particularly critical for the precession band, which has been proven to be sensitive to stratigraphic uncertainty (Figs. 8 to 11), and for which amplitude modulation is governed by eccentricity. The use of a wide- band filter, such as in the procedure of Zeeden et al. (2015), limits these biases and helps in a better reconstruction of the short wavelengths. Otherwise, a robust reconstruction of the amplitude modulation of the precession band requires limited biases of the power spectrum in the precession, which requires a good control on the sample position in the field. In addition, the simulations indicate that taking at least 4-10 samples per cycle should allow calculation of robust power spectra estimates in the respective cycle band (Table 1; Figs. 8-11).

Also in the evaluation of the relative contribution of precession and obliquity-related climatic forcing, an accurate assessment of the respective spectral power is essential (Ghirardi et al., 2014; Latta et al., 2006; Martinez et al., 2013; Weedon et al., 2004). Notably, whenever obliquity cycles are expressed more ~~manifestly~~ compared to precession cycles, this has been interpreted as a reflection of important climate dynamics and feedback mechanisms at high latitudes (Ruddiman and McIntyre, 1984), the build-up and decay of quasi-stable carbon reservoirs (Laurin et al., 2015), or direct obliquity forcing at tropical latitudes (Bosmans et al., 2015; Park and Oglesby, 1991). A robust evaluation of the relative contribution of precession and obliquity requires at least that no bias occurs from the generation of the depth-series, which includes the sampling procedure. This is particularly crucial in the case where the autoregressive coefficient of the red-noise background is high as in the La Thure series. Because of their low powers in the spectrum of the raw series, the spectral peaks related to the precession cycles become not significant at 10 to 15% uncertainties (Figs. 9-10). In that case, one can misleadingly interpret the absence of the record of the precession cycles in the sedimentary series while, the absence of significant high frequencies can simply be the consequence of spectral smoothing when increasing the level of stratigraphic uncertainty. Once again, a good control of the sample position accompanied by a high density of sampling will importantly improve interpretations of the relative contributions of the precession and obliquity on the spectrum, which will in turn help making accurate palaeoclimatic interpretations.

8. Conclusion

Errors made during the measurement of the stratigraphic position of a sample significantly affect the power spectrum of depth series. We present a method to assess the impact of such errors that is compatible with different techniques for spectral analysis. Our method is based on a Monte-Carlo procedure that randomises the sample steps of the time series, using a gamma distribution. Such a distribution preserves the stratigraphic order of samples, and allows controlling the average and the variance of the distribution of sample steps after randomisation. The simulations presented in this paper show that the gamma distribution of sample steps realistically simulates errors that are generally made during the measurement of sample positions. The three case studies presented in this paper all show a strong decrease in the power spectrum at high frequencies. Simulations indicate that the power spectrum can be completely smoothed for periods ~~samples~~ less than 3-4 times the average sample distance. Thus taking at least 3-4 samples per thinnest cycle of interest (e.g., precession cycles for the Milankovitch band) should preserve spectral peaks of this cycle. However, the decrease of power observed in a large portion of the spectrum implies a decrease in the significance level of the spectral peaks. Taking at least 4-10 samples per thinnest cycle of interest should allow their significance level to be preserved, depending on the level of stratigraphic uncertainty and depending on the redness of the power spectrum. Robust reconstruction of the power spectrum in the entire Milankovitch band requires a robust control of the sample step in the field, and requires a high density of sampling. To avoid any dispersion of the power spectrum in the precession band, taking 10 samples per precession cycles appears to be a safe density of sampling. For lower resolution of sampling, it is recommended to apply gamma-law simulations to ensure that stratigraphic uncertainty few impact powers and significance levels of the targeted cycles. Gamma-law simulations can also be used to simulate the effect of variations in the sedimentation rate on insolation series, which should help in modelling the trasnfert from insolation series to sedimentary series.

Appendix A: optimized linear interpolation

When interpolating an unevenly sampled time-series to an even sample distance, part of amplitude is lost in the high frequencies because the position of the samples in the interpolated series do not necessarily correspond to the position of the maximums and minimums of the original time-series (Figs. A1a and b). Oversampling has been suggested to limit the loss of amplitude during the interpolation process (Hinnov et al., 2002). However, oversampling impacts the autoregressive coefficient when estimating the level of red noise in the spectrum background (Hinnov, 2016). The optimized linear interpolation is here designed to limit the loss of amplitude of high-frequency cycles by minimizing the average misfit (*M*) of the MS values between the original and the resampled time series (Fig. A1c, Eq. A1):

$$M = \frac{1}{n} * \sum_{i=1}^{n} \left| s_{ori}[i] - s_{interp}[i] \right| \qquad \text{Eq. A1}$$

With: M is the average misfit between the values of the two curves

n is the number of points compared

s_{ori} is the original signal

s_{interp} is the resampled signal at the average sample distance of the original series

This comparison is only possible if the depths (or ages) of $s_{ori}[i]$ and $s_{interp}[i]$ are the same. This is of course not the case between the original and the resampled time series (Fig. A1b), otherwise interpolation would not be necessary. To circumvent this problem, the original and the resampled time series are both linearly interpolated with a sample step equal to the maximum resolution by which the depths (or ages) are provided. For instance, in the case of the La Thure series, the depths are given with a resolution of 0.01 m, so that s_{ori} and s_{interp} are linearly interpolated at 0.01 m. This procedure does not change the shape of neither the original time series nor the time series resampled at the average sample distance (Fig. A1c).

To test which resampled time series fits ~~the~~ best with the original time series, various depths are tested as starting points to resample the time at the average sample distance (Fig. A1d). The various scenarios of starting points tested increment by dx and have the following range:

$$T_{st.test} = T_{st.ori} : dx : (T_{st.ori} + dmoy - dx) \qquad \text{Eq. A2}$$

With, $T_{st.test}$, the tested starting points of time series resampled at the average sample distance

$T_{st.ori}$, the starting point of the original time series

dmoy, the average sample distance of the original time series

dx, the resolution with which the depths (or ages) are given

The best-fit curve is the one for which *M* is minimized.

Acknowledgements

ERC Consolidator Grant "EarthSequencing" (Grant Agreement No. 617462) funded this project. We acknowledge Hubertus Fischer for having editing the manuscript, Christian Zeeden and Linda Hinnov for having reviewed it and Anna-Joy Drury for English-proof reading.

References

Bosmans, J. H. C., Hilgen, F. J., Tuenter, E., and Lourens, L. J.: Obliquity forcing of low-latitude climate, Clim Past, 11, 1335-1346, doi:10.5194/cp-11-1335-2015, 2015.

Boulila, S., Charbonnier, G., Galbrun, B., and Gardin, S.: Climatic precession is the main driver of Early Cretaceous sedimentation in the Vocontian Basin (France): Evidence from the Valanginian Orpierre succession, Sedimentary Geology, 324, 1-11, doi:10.1016/j.sedgeo.2015.04.014, 2015.

Bulot, L. G., Thieuloy, J.-P., Eric, B., and Klein, J.: Le cadre stratigraphique du Valanginien supérieur et de l'Hauterivien du Sud-Est de la France: définition des biochronozones et caractérisation de nouveaux biohorizons, Géologie Alpine, 68, 13-56, 1992.

Bunn, A. G.: A dendrochronology program library in R (dplR), Dendrochronologia, 26, 115-124, doi:10.1016/j.dendro.2008.01.002, 2008.

Bunn, A. G.: Statistical and visual crossdating in R using the dplR library, Dendrochronologia, 28, 251-258, doi:10.1016/j.dendro.2009.12.001, 2010.

Bunn, A. G., Korpela, M., Biondi, F., Campelo, F., Mérian, P., Qeadan, F., and Zang, C., dplR: Dendrochronology Program Library in R. R package version 1.6.2. http://CRAN.R-project.org/package=dplR, 2015.

Burgin, T. A.: The Gamma Distribution and Inventory Control, Operational Research Quarterly, 26, 507-525, 1975.

De Vleeschouwer, D., Boulvain, F., Da Silva, A. C., Pas, D., Labaye, C., and Claeys, P.: The astronomical calibration of the Givetian (Middle Devonian) timescale (Dinant

Synclinorium, Belgium), Geological Society, London, Special Publications, 414, 245-256, doi:10.1144/sp414.3, 2015.

De Vleeschouwer, D. and Parnell, A. C.: Reducing time-scale uncertainty for the Devonian by integrating astrochronology and Bayesian statistics, Geology, 42, 491-494, doi:10.1130/g35618.1, 2014.

Ghirardi, J., Deconinck, J.-F., Pellenard, P., Martinez, M., Bruneau, L., Amiotte-Suchet, P., and Pucéat, E.: Multi-proxy orbital chronology in the aftermath of the Aptian Oceanic Anoxic Event 1a: Palaeoceanographic implications (Serre Chaitieu section, Vocontian Basin, SE France), Newsl Stratigr, 47, 247-262, doi:10.1127/0078-0421/2014/0046, 2014.

Hagelberg, T. K., Pisias, N. G., Shackleton, N. J., Mix, A. C., and Harris, S. E.: Refinement of a high-resolution, continuous sedimentary section for studying equatorial Pacific Ocean paleoceanography. Pisias, N. G., Mayer, L. A., Janecek, T. R., Palmer-Julson, A., and van Andel, T. H. (Eds.), Proceedings of the Ocean Drilling Program, Scientific Results, 138, Texas A & M University, College Station, Texas, 1995.

Hinnov, L. A.: Cyclostratigraphy and its revolutionizing applications in the earth and planetary sciences, Geological Society of America Bulletin, 125, 1703-1734, doi:10.1130/b30934.1, 2013.

Hinnov, L. A.: Interactive comment on "Testing the impact of stratigraphic uncertainty on spectral analyses of sedimentary series" by Mathieu Martinez et al. Climate of the Past Discussion, doi: 10.5194/cp-2015-188-RC2, 2016.

Hinnov, L. A., Schulz, M., and Yiou, P.: Interhemispheric space-time attributes of the Dansgaard-Oeschger oscillations between 0-100 ka, Quaternary Science Reviews, 21, 1213-1228, 2002.

Huang, Z., Ogg, J. G., and Gradstein, F. M.: A quantitative study of lower Cretaceous cyclic sequences from the Atlantic Ocean and the Vocontian Basin (SE France), Paleoceanography, 8, 275-291, 1993.

Huang, C., Hesselbo, S. P., and Hinnov, L.: Astrochronology of the late Jurassic Kimmeridge Clay (Dorset, England) and implications for Earth system processes, Earth and Planetary Science Letters, 289, 242-255, doi:10.1016/j.epsl.2009.11.013, 2011.

Huybers, P. and Wunsch, C.: A depth-derived Pleistocene age model: Uncertainty estimates, sedimentation variability, and nonlinear climate change, Paleoceanography, 19, PA1028, doi:10.1029/2002pa000857, 2004.

Latta, D. K., Anastasio, D. J., Hinnov, L. A., Elrick, M., and Kodama, K. P.: Magnetic record of Milankovitch rhythms in lithologically noncyclic marine carbonates, Geology, 34, 29-32, doi: 10.1130/G21918.1, 2006.

Laurin, J., Meyers, S. R., Uličný, D., Jarvis, I., and Sageman, B. B.: Axial obliquity control on the greenhouse carbon budget through middle- to high-latitude reservoirs, Paleoceanography, 30, 133-149, doi:10.1002/2014pa002736, 2015.

Lomb, N. R.: Least-squares frequency analysis of unequally spaced data, Astrophysics and space science, 39, 447-462, 1976.

Mann, M. E. and Lees, J. M.: Robust estimation of background noise and signal detection in climatic time series, Climatic Change, 33, 409-445, 1996.

Martinez, M., Pellenard, P., Deconinck, J.-F., Monna, F., Riquier, L., Boulila, S., Moiroud, M., and Company, M.: An orbital floating time scale of the Hauterivian/Barremian GSSP from a magnetic susceptibility signal (Río Argos, Spain), Cretaceous Research, 36, 106-115, doi:10.1016/j.cretres.2012.02.015, 2012.

Martinez, M., Deconinck, J.-F., Pellenard, P., Reboulet, S., and Riquier, L.: Astrochronology of the Valanginian Stage from reference sections (Vocontian Basin, France) and palaeoenvironmental implications for the Weissert Event, Palaeogeography, Palaeoclimatology, Palaeoecology, 376, 91-102, doi:10.1016/j.palaeo.2013.02.021, 2013.

Martinez, M., Deconinck, J.-F., Pellenard, P., Riquier, L., Company, M., Reboulet, S., and Moiroud, M.: 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, 158-173, doi:10.1016/j.gloplacha.2015.06.001, 2015.

Martinez, M. and Dera, G.: Orbital pacing of carbon fluxes by a ~9-My eccentricity cycle during the Mesozoic, Proceedings of the National Academy of Sciences of the United States of America, 112, 12604-12609, 2015.

Matys Grygar, T., Mach, K., Schnabl, P., Pruner, P., Laurin, J., and Martinez, M.: A lacustrine record of the early stage of the Miocene Climatic Optimum in Central Europe from the Most Basin, Ohře (Eger) Graben, Czech Republic, Geological Magazine, 151, 1013-1033, doi:10.1017/s0016756813001052, 2014.

Meyers, S. R.: Seeing red in cyclic stratigraphy: Spectral noise estimation for astrochronology, Paleoceanography, 27, PA3228, doi:10.1029/2012PA002307, 2012

Meyers, S. R.: Astrochron: An R package for Astrochronology. 2014.

Meyers, S. R.: The evaluation of eccentricity-related amplitude modulation and bundling in paleoclimate data: An inverse approach for astrochronologic testing and time scale optimization, Paleoceanography, 30, 1625-1640, doi: 10.1002/2015PA002850, 2015.

Meyers, S. R. and Sageman, B. B.: Quantification of deep-time orbital forcing by a verage spectral misfit, American Journal of Science, 307, 773-792, doi: 10.2475/05.2007.01, 2007.

Moiroud, M., Martinez, M., Deconinck, J.-F., Monna, F., Pellenard, P., Riquier, L., and Company, M.: High-resolution clay mineralogy as a proxy for orbital tuning: Example of the Hauterivian–Barremian transition in the Betic Cordillera (SE Spain), Sedimentary Geology, 282, 336-346, doi:10.1016/j.sedgeo.2012.10.004, 2012.

Moore, M. I. and Thomson, P. J.: Impact of jittered sampling on conventional spectral estimates, Journal of Geophysical Research, 96, 18519, doi:10.1029/91jc01623, 1991.

Mudelsee, M.: TAUEST: a computer program for estimating persistence in unvenly spaced weather/climate time series, Computers & Geosciences, 28, 69-72, 2002.

Park, J. and Oglesby, R. J.: Milankovitch rhythms in the Cretaceous: a GCM modelling study, Palaeogeography, Palaeoclimatology, Palaeoecology, 90, 329-355, 1991.

Pas, D., Da Silva, A.-C., Devleeschouwer, X., De Vleeschouwer, D., Labaye, C., Cornet, P., Michel, J., and Boulvain, F.: Sedimentary development and magnetic susceptibility evolution of the Frasnian in Western Belgium (Dinant Synclinorium, La Thure section), Special Publication-Geological Society of London, 414, 15-36, 2015.

Reboulet, S. and Atrops, F.: Comments and proposals about the Valanginian-Lower Hauterivian ammonite zonation of south-eastern France, Eclogae Geologicae Helvetiae, 92, 183-198, 1999.

Ruddiman, W. and McIntyre, A.: Ice-age thermal response and climatic role of the surface Atlantic Ocean, 40 N to 63 N, Geological Society of America Bulletin, 95, 381-396, 1984.

Scargle, J. D.: Studies in astronomical time series analysis. II-Statistical aspects of spectral analysis of unevenly spaced data, The Astrophysical Journal, 263, 835-853, 1982.

Schulz, M. and Mudelsee, M.: REDFIT: estimating red-noise spectra directly from unevenly spaced paleoclimatic time series, Computers & Geosciences, 28, 421-426, 2002.

Shackleton, N., Hagelberg, T., and Crowhurst, S.: Evaluating the success of astronomical tuning: Pitfalls of using coherence as a criterion for assessing pre-Pleistocene timescales, Paleoceanography, 10, 693–697, doi:10.1029/95PA01454, 1995.Thibault, N., Jarvis, I., Voigt, S., Gale, A. S., Attree, K., and Jenkyns, H. C.: Astronomical calibration and global correlation of the Santonian (Cretaceous) based on marine carbon isotope record, Paleoceanography, 31, 847–865, doi:10.1002/2016PA002941, 2016.

Thomson, D. J.: Quadratic-inverse spectrum estimates: applications to palaeoclimatology, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 332, 539-597, 1990.

Thomson, D. J.: Spectrum estimation and harmonic analysis, Proceedings of the IEEE, 70, 1055-1096, 1982.

Weedon, G. P.: Time-Series Analysis and Cyclostratigraphy, Cambridge University Press, Cambridge, UK, 2003.

Weedon, G. P., Coe, A. L., and Gallois, R. W.: Cyclostratigraphy, orbital tuning and inferred productivity for the type Kimmeridge Clay (Late Jurassic), Southern England, Journal of the Geological Society, 161, 655-666, 2004.

Weedon, G. P. and Jenkyns, H. C.: Cyclostratigraphy and the Early Jurassic timescale: Data from the Belemnite Marls, Dorset, southern England, Geological Society of America Bulletin, 111, 1823-1840, doi:10.1130/0016-7606(1999)111<1823:CATEJT>2.3.CO;2, 1999.

Westerhold, T., Röhl, U., Raffi, I., Fornaciari, E., Monechi, S., Reale, V., Bowles, J., and Evans, H. F.: Astronomical calibration of the Paleocene time, Palaeogeography, Palaeoclimatology, Palaeoecology, 257, 377-403, doi:10.1016/j.palaeo.2007.09.016, 2008.

Zeeden, C., Meyers, S. R., Lourens, L. J., and Hilgen, F. J.: Testing astronomically tuned age models, Paleoceanography, 30, 369-383, doi:10.1002/2014pa002762, 2015.

Figure caption

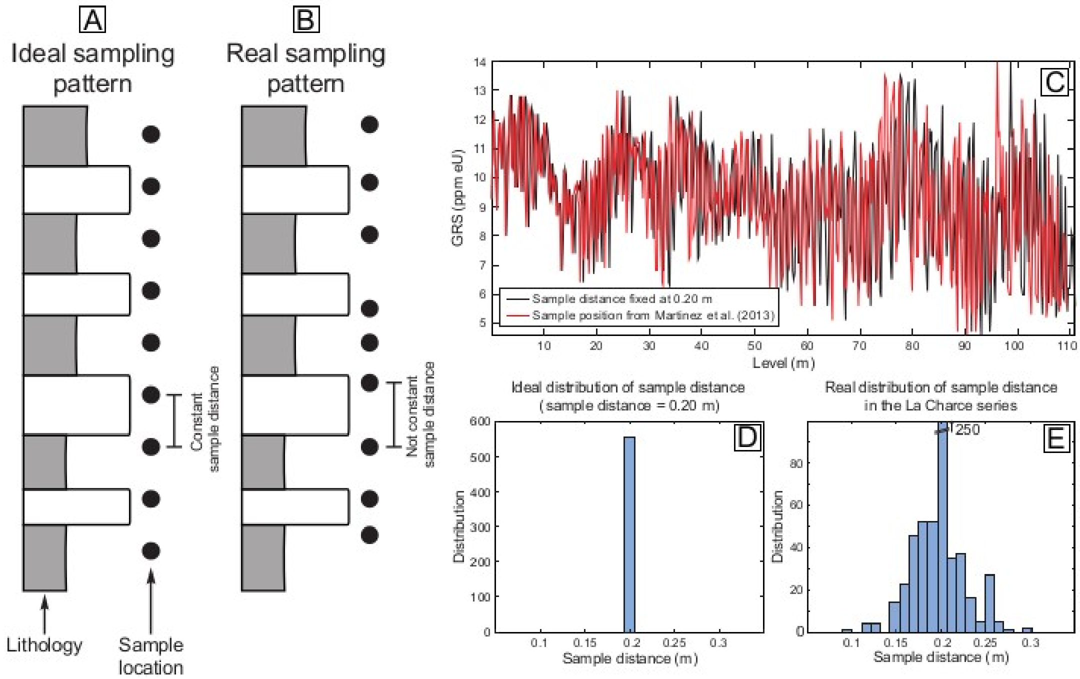

Fig. 1. Illustration of the problem. (a) Theoretical sedimentary log with position of samples in an ideal case where the samples are strictly equally distant. (b) Theoretical sedimentary log with position of samples in a common sampling pattern where all samples are not strictly equally distant. Here the error in the sample position is exaggerated for the purpose of the example. (c) The gamma-ray series from La Charce shown as if all samples were strictly equidistant (black curve), and as they are positioned in Martinez et al. (2013) (red curve). (d) Distribution of sample distances in case of ideal sampling of the La Charce series (all sample distances are fixed at 0.20 m). (e) Distribution of sample distances in case of the La Charce series as published in Martinez et al. (2013).

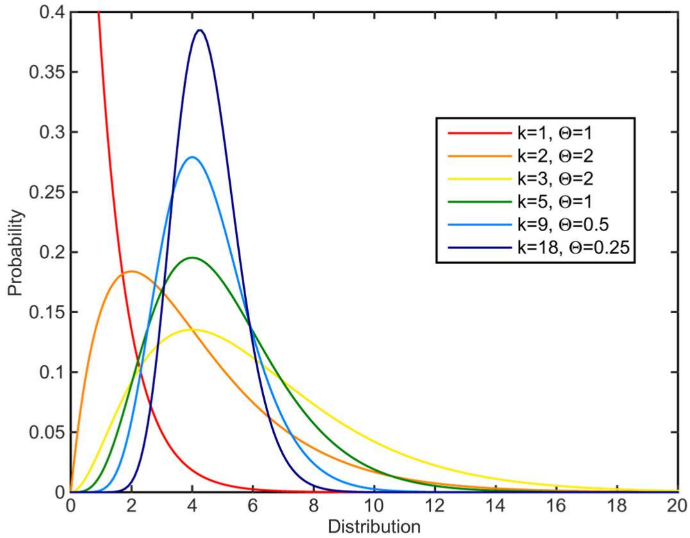

Fig. 2. Gamma probability density functions (PDF). All Gamma PDF's have a positive support, which is a crucial characteristic to realistically simulate sample steps. The gamma density probability functions were generated with the Matlab *gampdf* function.

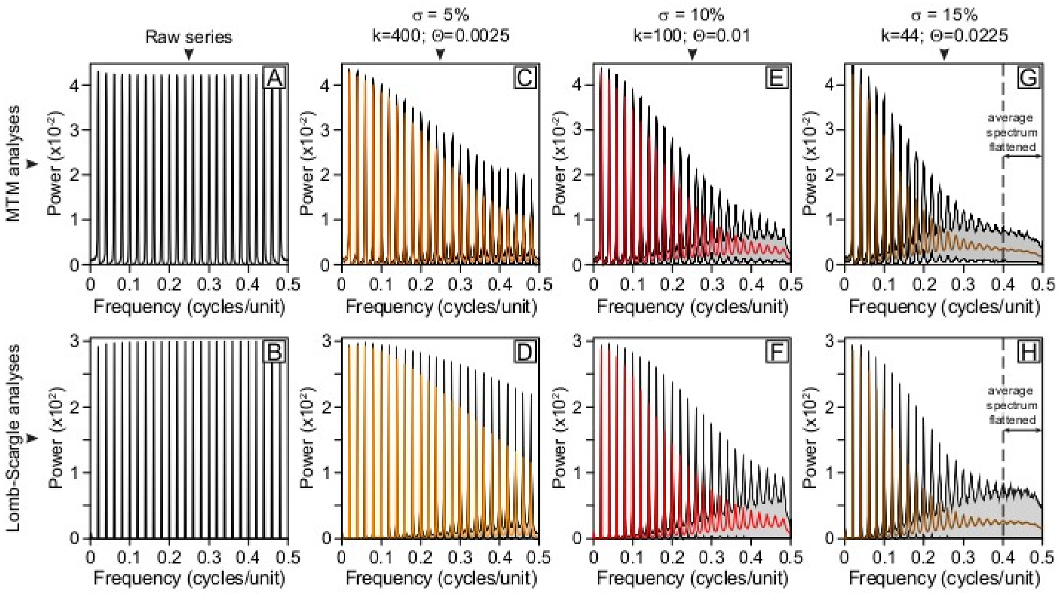

Fig. 3. Effect of the gamma-law randomised sample distances on the 2π-MTM and Lomb-Scargle spectra of the series of sum of pure sinusoids. (a) and (b) Spectra of the series without sample step randomisation. (c) and (d) with 5% of stratigraphic uncertainty. (e) and (f) with 10% of stratigraphic uncertainty. (g) and (h) with 15% of stratigraphic uncertainty. For each simulation shown from (c) to (h), the grey areas represent the interval covering 95% of the simulations, while the red, orange and brown curves represent the average spectrum.

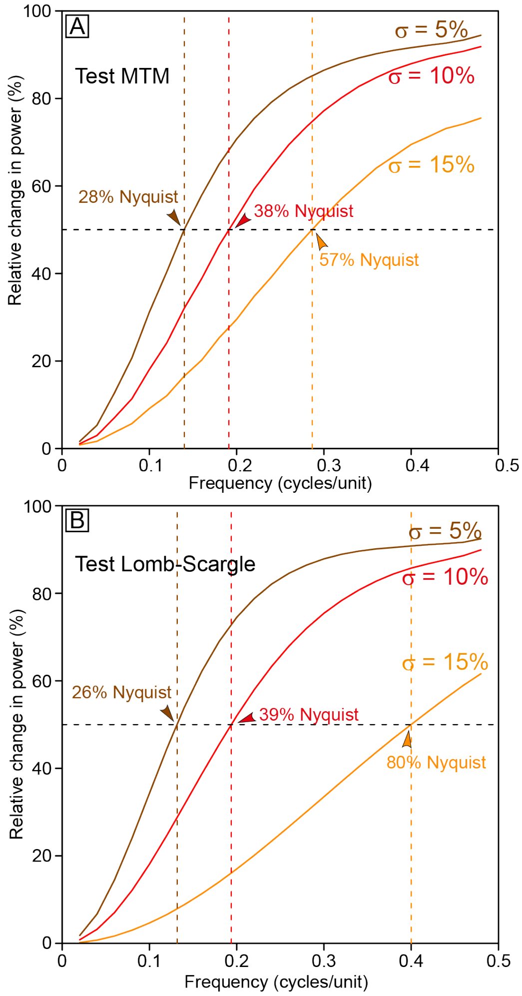

Fig. 4. Relative change in power in the (a) 2π-MTM spectra, and (b) Lomb-Scargle spectra after applying the gamma-law simulations of distortion of the time series. The arrows indicate at which frequency (relatively to the Nyquist frequency) the change in power exceeds 50%.

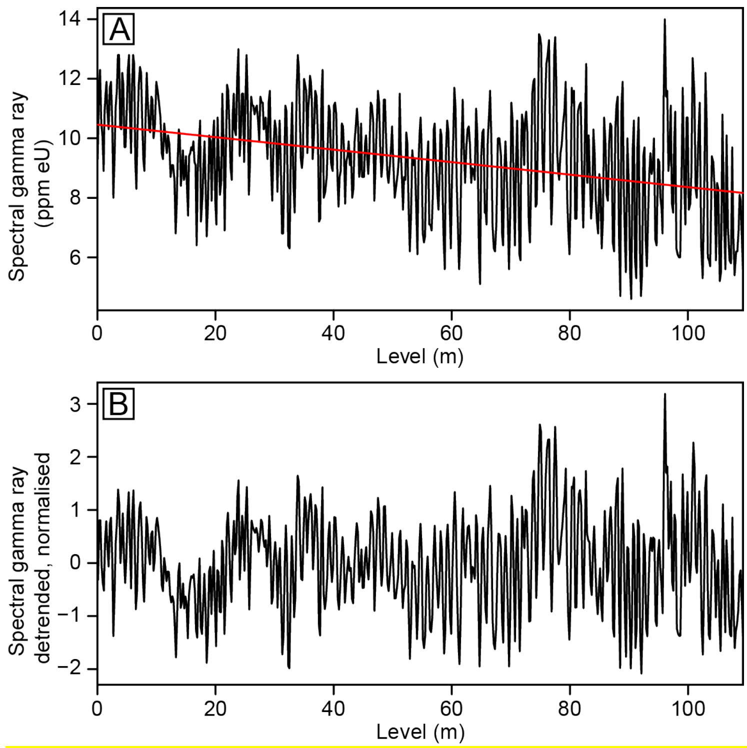

Fig. 5. Detrending procedure of the gamma-ray series from the La Charce section. (a) Raw gamma-ray signal (black curve) with best-fit linear trend (red curve). (b) Gamma-ray series after subtraction of the linear trend and standardisation (average = 0; standard deviation = 1).

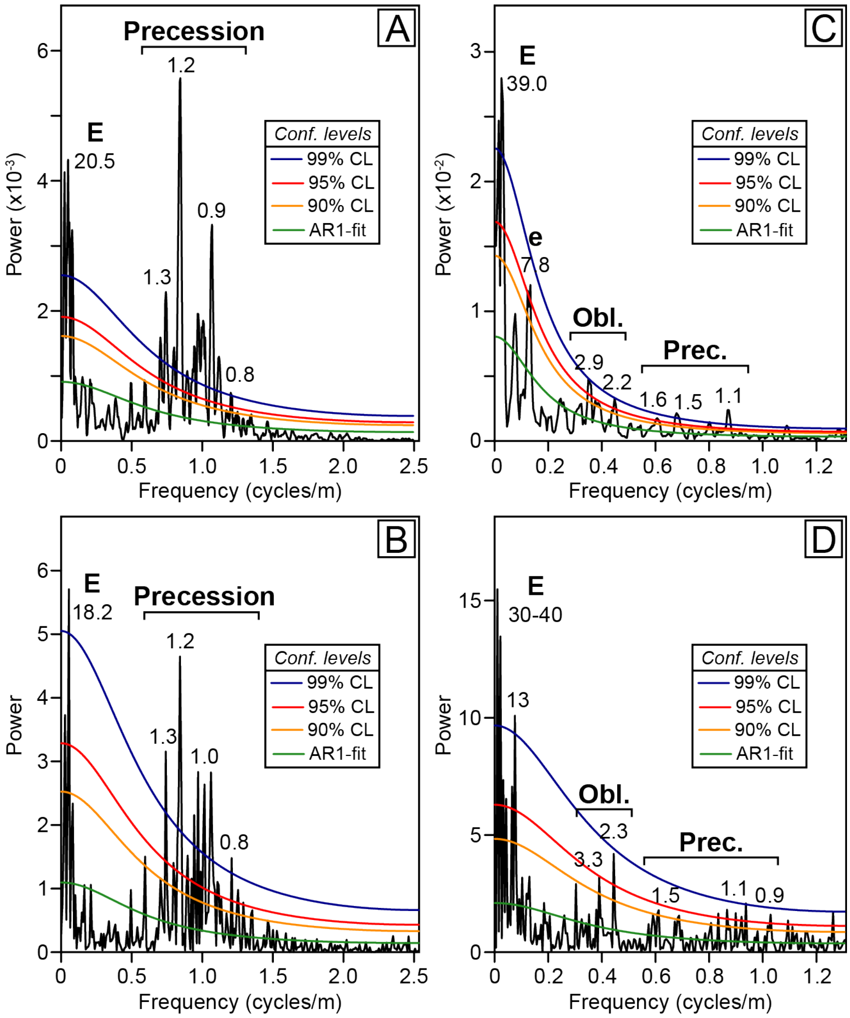

Fig. 6. Spectra of the La Charce and La Thure series before Monte-Carlo simulations of the sample distances. (a) 2π-MTM spectrum of the La Charce series. (b) REDFIT spectrum of the La Charce series. (c) 2π-MTM spectrum of the La Thure series. (d) REDFIT spectrum of the La Thure series. The main significant frequencies are given in meters.

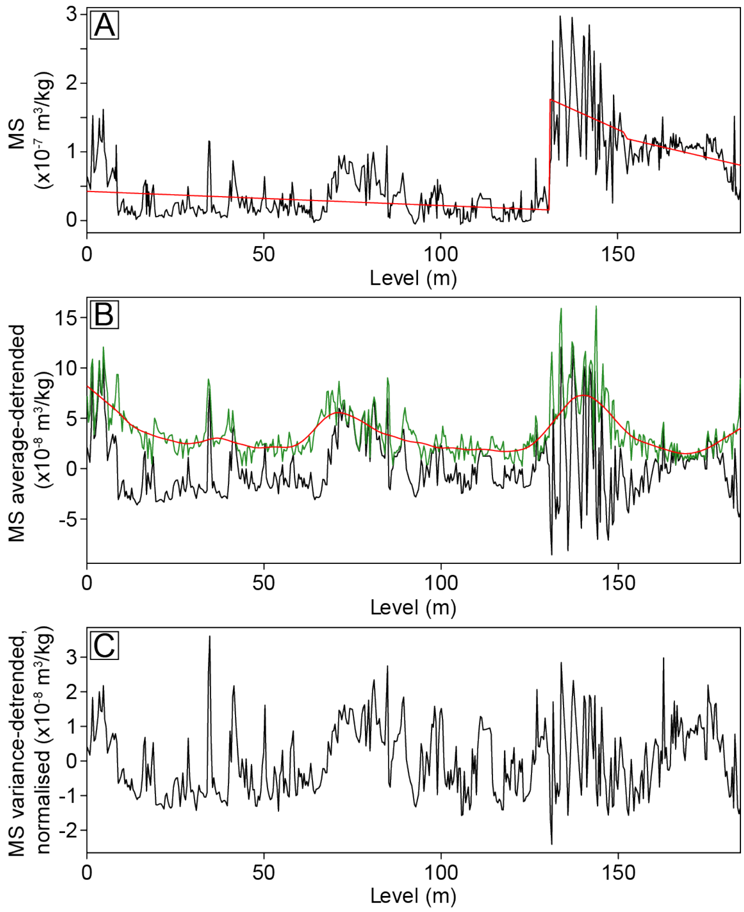

Fig. 7. Detrending procedure of the magnetic susceptibility (MS) series from the La Thure section. (a) Raw MS signal (black curve) with piecewise best-fit linear trend of the average (red curve). (b) MS series after subtraction of the piecewise linear trend (black curve), with instantaneous amplitude (green curve) and LOWESS regression of the instantaneous amplitude applied with a coefficient of 10% (red curve). (c) MS curve after dividing the MS series "average-detrended" by the LOWESS regression of the instantaneous amplitude, and after standardisation.

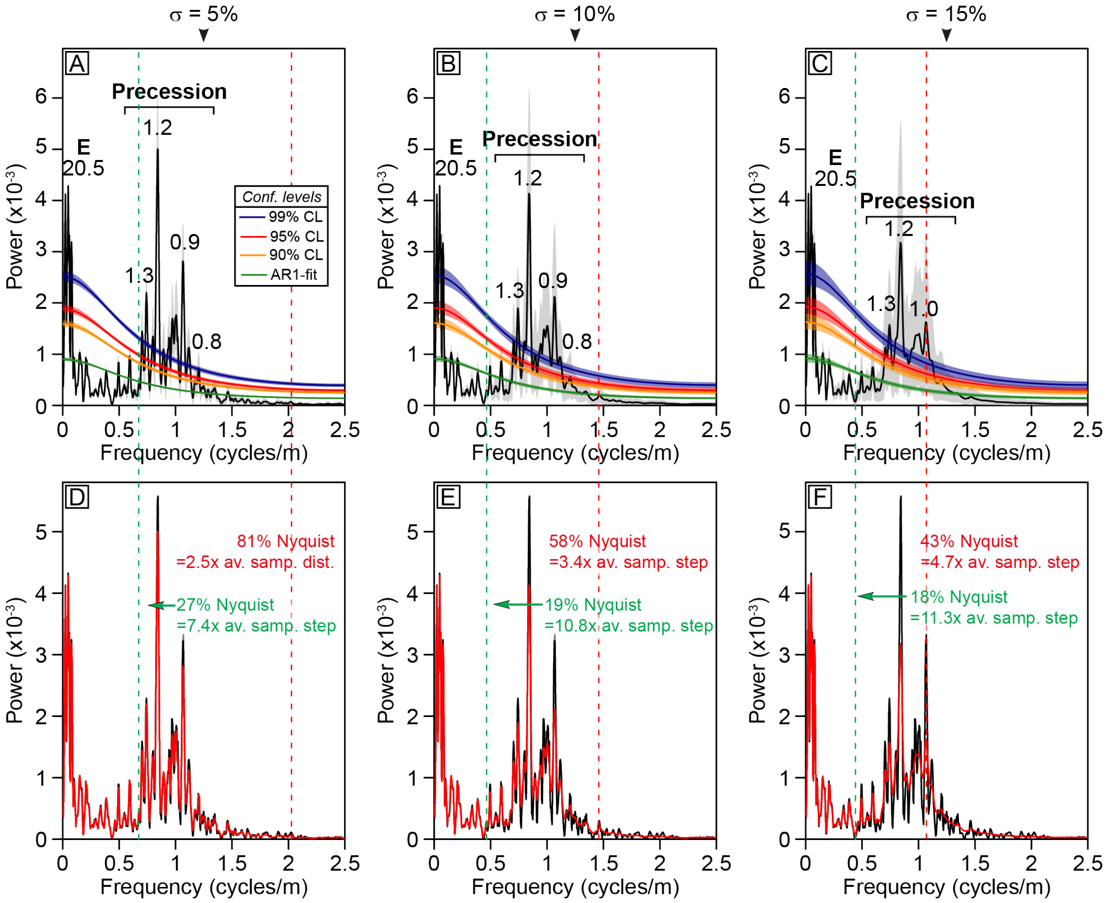

Fig. 8. Effect of the gamma-law randomisation of the sample distances on the 2π-MTM spectrum of the La Charce series. (a to c) 2π-MTM spectra with a level of stratigraphic uncertainty ~~respectively~~ fixed 5%, 10% and 15% of the average sample distance of the series. The grey area represents the interval covering 95% of the simulations. The average confidence levels are reported on the spectra with their respective areas covering 95% of the simulations. Main significant periods are indicated in meters with, in bold, their corresponding orbital cycles. E: 405-kyr eccentricity. (d to f) Superposition of the 2π-MTM spectra before randomisation (*in black*) and the average spectrum after the 1000 simulations (*in red*). The red dashed bands indicate the lowest frequency from which the spectrum is completely smoothen, so that no more frequency can be identified. The green dashed band represents the highest frequency in which the spectrum of the series before randomisation appears practically confounded to the spectrum after randomisation.

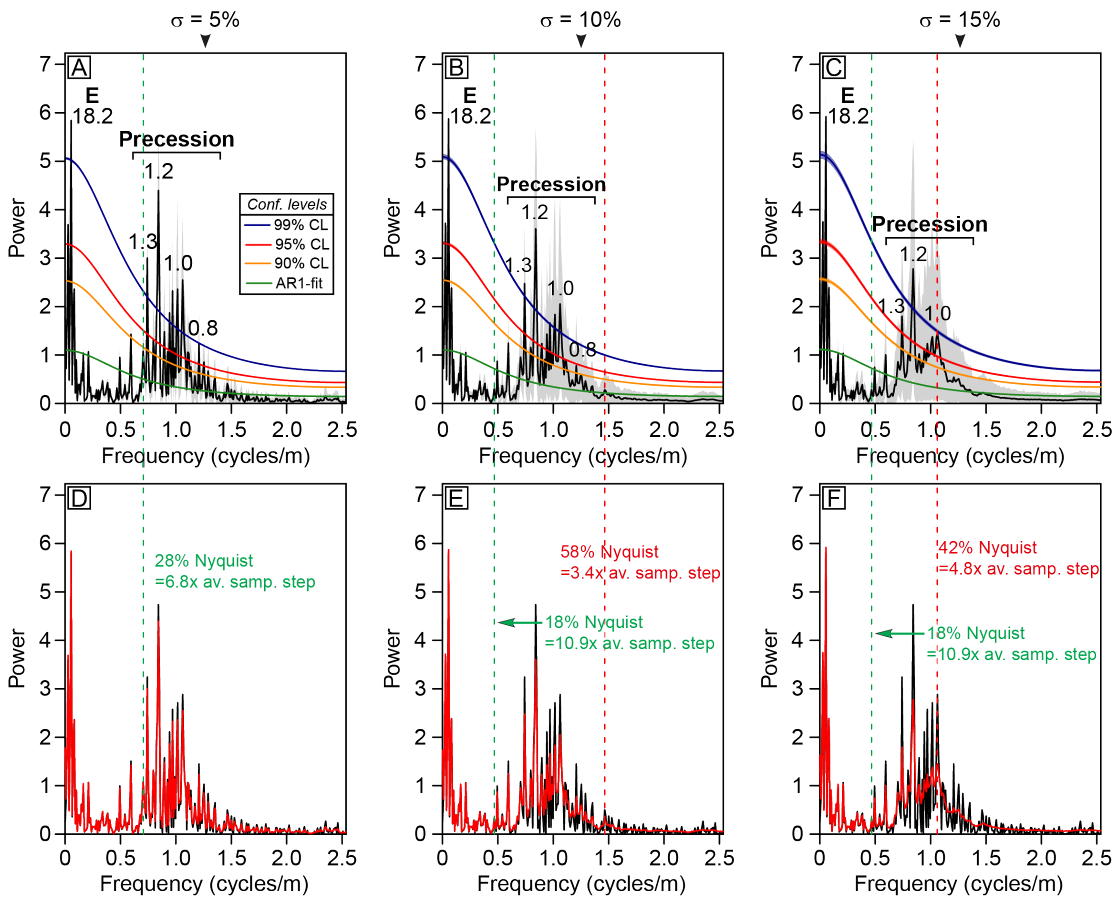

Fig. 9. Effect of the gamma-law randomisation of the sample distances on the REDFIT spectrum of the La Charce series. (a to c) REDFIT spectra with a level of stratigraphic uncertainty ~~respectively~~ fixed 5%, 10% and 15% of the average sample distance of the series. The grey area represents the interval covering 95% of the simulations. The average confidence levels are reported on the spectra with their respective areas covering 95% of the simulations. Main significant periods are indicated in meters with, in bold, their corresponding orbital cycles. E: 405-kyr eccentricity. (d to f) Superposition of the REDFIT spectra before randomisation (*in black*) and the average spectrum after the 1000 simulations (*in red*). The red dashed ~~bands~~ indicate the lowest frequency ~~from~~ which the spectrum is completely smoothed, so that no more frequency can be identified. The green dashed ~~band~~ represents the

highest frequency ~~in~~ which the spectrum of the series before randomisation appears practically ~~confounded~~ to the spectrum after randomisation.

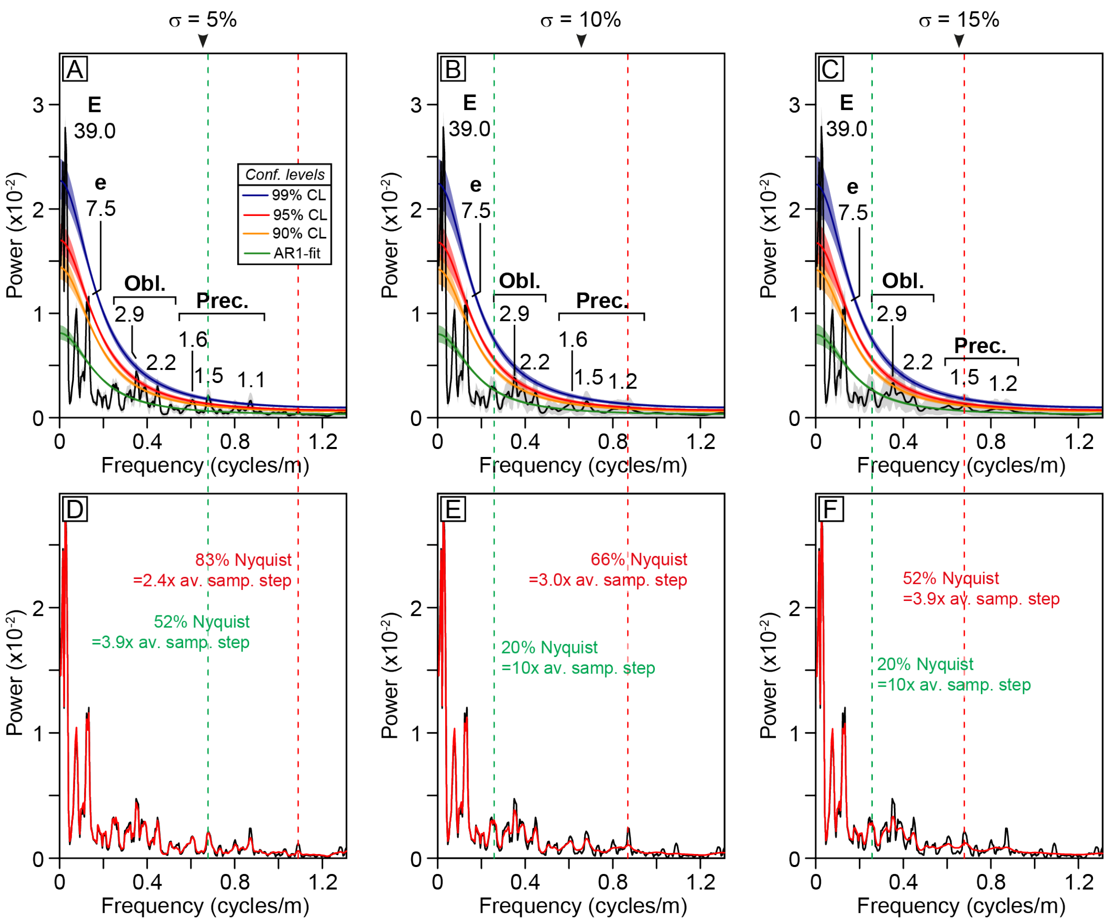

Fig. 10. Effect of the gamma-law randomisation of the sample distances on the 2π-MTM spectrum of the La Thure series. (a to c) 2π-MTM spectra with a level of stratigraphic uncertainty ~~respectively~~ fixed 5%, 10% and 15% of the average sample distance of the series. The grey area represents the interval covering 95% of the simulations. The average confidence levels are reported on the spectra with their respective areas covering 95% of the simulations. Main significant periods are indicated in meters with, in bold, their corresponding orbital cycles. E: 405-kyr eccentricity; e: 100-kyr eccentricity. (d to f) Superposition of the 2π-MTM spectra before randomisation (*in black*) and the average spectrum after the 1000 simulations (*in red*). The red dashed ~~bands~~ indicate the lowest frequency ~~from~~ which the spectrum

is completely smoothed, so that no more frequency can be identified. The green dashed band represents the highest frequency in which the spectrum of the series before randomisation appears practically confounded to the spectrum after randomisation.

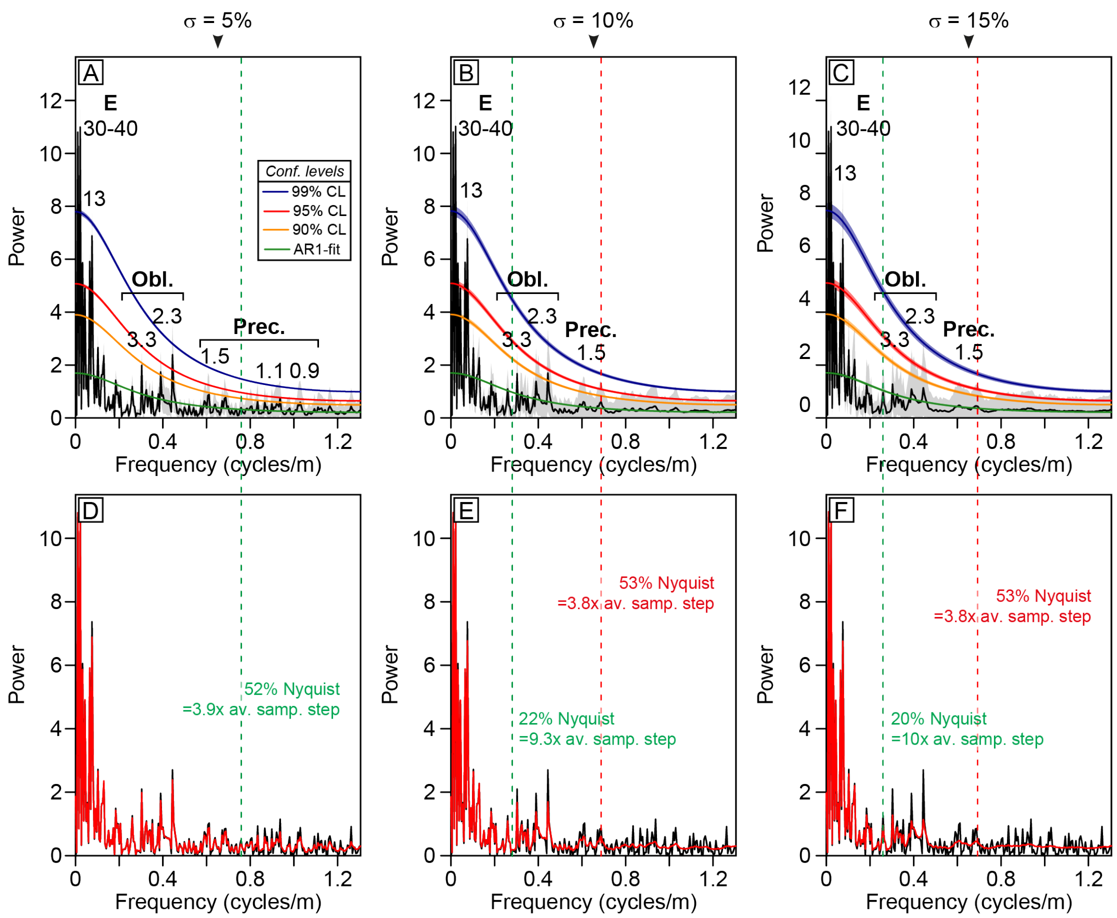

Fig. 11. Effect of the gamma-law randomisation of the sample distances on the REDFIT spectrum of the La Charce series. (a to c) REDFIT spectra with a level of stratigraphic uncertainty ~~respectively~~ fixed 5%, 10% and 15% of the average sample distance of the series. The grey area represents the interval covering 95% of the simulations. The average confidence levels are reported on the spectra with their respective areas covering 95% of the simulations. Main significant periods are indicated in meters with, in bold, their corresponding orbital cycles. E: 405-kyr

eccentricity; e: 100-kyr eccentricity. (d to f) Superposition of the REDFIT spectra before randomisation (*in black*) and the average spectrum after the 1000 simulations (*in red*). The red dashed ~~bands~~ indicate the lowest frequency ~~from~~ which the spectrum is completely smoothed, so that no more frequency can be identified. The green dashed ~~band~~ represents the highest frequency ~~in~~ which the spectrum of the series before randomisation appears practically ~~confounded~~ to the spectrum after randomisation.

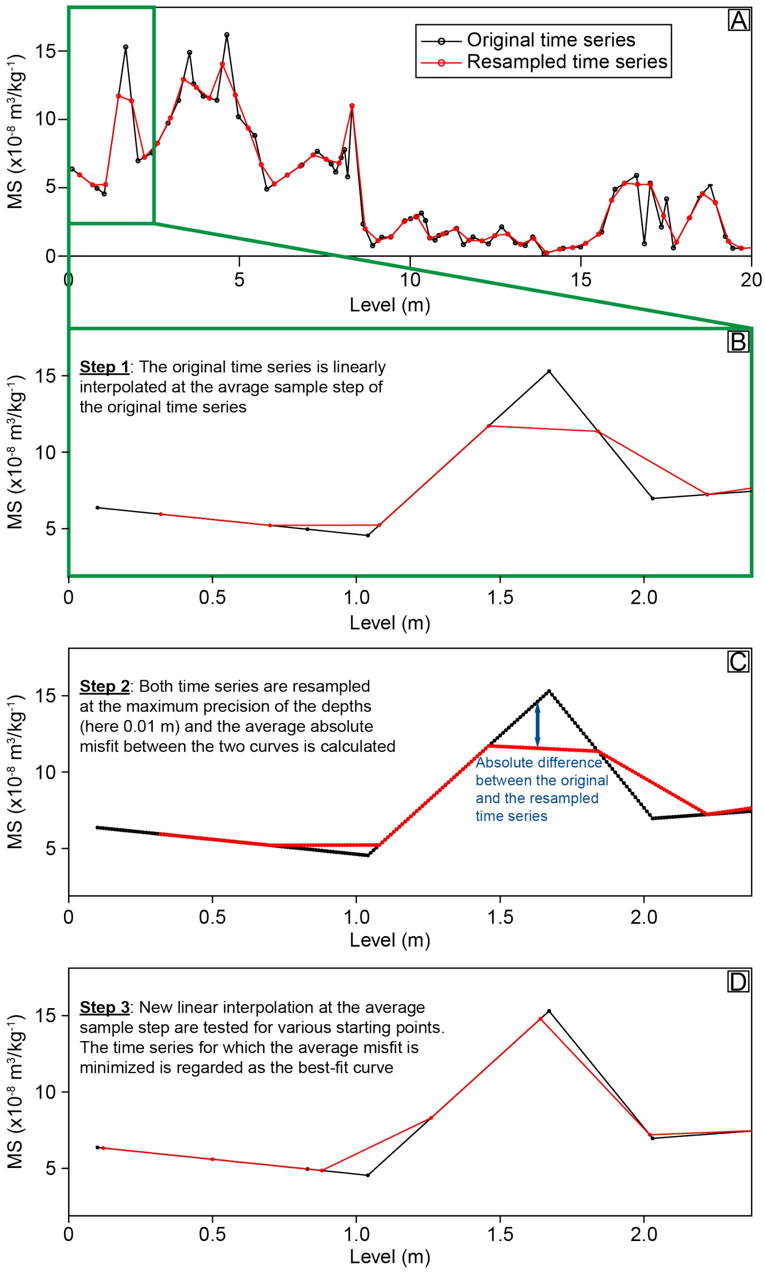

Fig. A1. Scheme of the procedure of the optimized linear interpolation of time series. An example of application is shown for the La Thure section in Fig. A2. Differences in the resulting spectrum between the best-fit and the worst-fit resampled time series are displayed in this figure. Main differences in the spectra of the two cases are observed in the middle and high frequencies. Compared to the worst-fit resampling, the spectra of the best-fit resampling show decreased power and confidence levels in the middle frequencies (from 0.2 to 0.7 cycles.m^{-1}), while increased power and confidence levels rather occur in the high frequencies (from 0.7 cycles.m^{-1} to the Nyquist frequency). Fitting the best curve to the original time series thus impacts on the calculation of the power spectrum and the confidence levels of the spectral peaks.

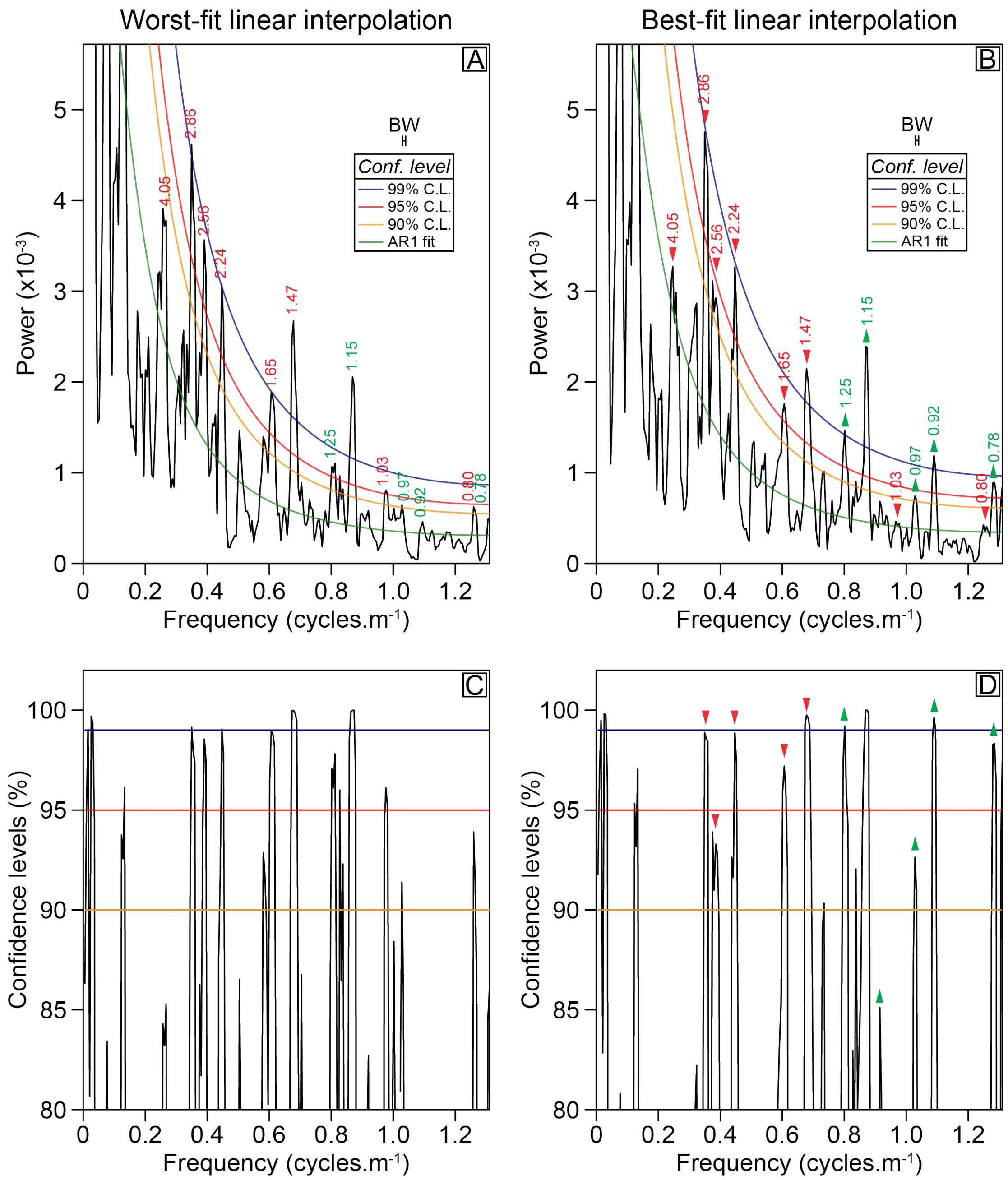

Fig. A2. Comparison of spectra of the resampled time series for the worst-fit case (a and c), and for the best-fit case (b and d). Spectra (a) and (b) are calculated using the 2π -multi-taper method with confidence levels calculated using the method of Mann and Lees (1996) with a Tukey's end-point rule (Meyers, 2014). (c) and (d) show the confidence levels compared to a red noise. Red arrows indicate the frequency at which powers and confidence levels decrease from the worst-fit case to the best-fit case. Green arrows display the opposite case.

		σ=0%	σ=5%	σ=10%	σ=15%
La Charce - MTM	Autoregressive coefficient	0.440	0.433 ± 0.025	0.432 ± 0.037	0.434 ± 0.048
	Average power (x10^{-4})	3.54	3.55 ± 0.13	3.58 ± 0.20	3.61 ± 0.25
La Charce - redfit	Autoregressive coefficient	0.468	0.468 ± 0.002	0.467 ± 0.003	0.467 ± 0.006
	Average power	0.398	0.399 ± 0.003	0.402 ± 0.005	0.407 ± 0.008
La Thure - MTM	Autoregressive coefficient	0.657	0.658 ± 0.025	0.653 ± 0.029	0.651 ± 0.033
	Average power (x10^{-3})	1.67	1.67 ± 0.04	1.67 ± 0.05	1.68 ± 0.07
La Thure - redfit	Autoregressive coefficient	0.407	0.406 ± 0.004	0.405 ± 0.008	0.404 ± 0.013
	Average power	0.890	0.894 ± 0.011	0.900 ± 0.019	0.904 ± 0.027

Table 1. Results of red-noise background estimates from the La Charce and the La Thure series with the 2π-MTM and the REDFIT analyses.

		Level of stratigraphic uncertainty		
		5%	**10%**	**15%**
La Charce MTM	Highest frequency before smoothing	81% Nyquist	58% Nyquist	43% Nyquist
	equivalent number sample steps	2.5x	3.4x	4.7x
	Highest frequency confounded spectra	27% Nyquist	19% Nyquist	18% Nyquist
	equivalent number sample steps	7.4x	10.8x	11.3x
La Charce REDFIT	Highest frequency before smoothing	/	58% Nyquist	42% Nyquist
	equivalent number sample steps	/	3.4x	4.8x
	Highest frequency confounded spectra	28% Nyquist	18% Nyquist	18% Nyquist
	equivalent number sample steps	6.8x	10.9x	10.9x
La Thure MTM	Highest frequency before smoothing	83% Nyquist	66% Nyquist	52% Nyquist
	equivalent number sample steps	2.4x	3.0x	3.9x
	Highest frequency confounded spectra	52% Nyquist	20% Nyquist	20% Nyquist
	equivalent number sample steps	3.9x	10x	10x
La Thure REDFIT	Highest frequency before smoothing	/	53% Nyquist	53% Nyquist
	equivalent number sample steps	/	3.8x	3.8x
	Highest frequency confounded spectra	52% Nyquist	22% Nyquist	20% Nyquist
	equivalent number sample steps	3.9x	9.3x	10x%

Table 2. Synthesis of the results of highest frequencies before smoothing of the spectra when applying the Monte-Carlo simulations, and of highest frequency in which the spectra before and after simulation can be confounded.