Palaeogeography, Palaeoclimatology, Palaeoecology 461 (2016) 475–480

Contents lists available at ScienceDirect

Palaeogeography, Palaeoclimatology, Palaeoecology

journal homepage: www.elsevier.com/locate/palaeo

Reply to the comment on “Geologic evidence for chaotic behavior of the planets and its constraints on the third-order eustatic sequences at the end of the Late Paleozoic Ice Age” by Qiang Fang, Huaichun Wu, Linda A. Hinnov, Xiuchun Jing, Xunlian Wang, and Qingchun Jiang [*Palaeogeography Palaeoclimatology Palaeoecology 400* (2015) 848–859]

Linda A. Hinnov[a,*], Huaichun Wu[b,c,*], Qiang Fang[b,c]

[a] *Department of Atmospheric, Oceanic and Earth Sciences, George Mason University, Fairfax, VA 22030, USA*
[b] *State Key Laboratory of Biogeology and Environmental Geology, China University of Geosciences, Beijing 100083, China*
[c] *School of Ocean Sciences, China University of Geosciences, Beijing 100083, China*

ARTICLE INFO

Article history:
Received 1 June 2016
Received in revised form 26 June 2016
Accepted 21 July 2016
Available online 22 July 2016

Keywords:
Cyclostratigraphy
MTM spectrum estimation
Significance thresholds
Spectral noise models
Bonferroni correction

ABSTRACT

In their comment, Smith et al. (2016) discount the astronomical cycle identifications made from Middle Permian cyclostratigraphy in our recent paper (Fang et al. 2015), declaring that we presented defective null models and improper hypothesis tests, and greatly overestimated the statistical significance of cycles. Here we respond in detail to clarify the decisions that were made in our work, and to correct errors and omissions and other misunderstandings arising in their comment. We also discuss the advent of objective methodologies that promise to improve research in cyclostratigraphy in the near future.

1. Introduction

Smith et al. (2016–in this issue), henceforth Smith et al., maintain that the spectral analysis presented in Fang et al. (2015) on Middle Permian cyclostratigraphy greatly overestimates the significance of cyclostratigraphic periodicities, and that the interpretations that flow from the analysis should be rejected. They point to a multitude of perceived improper procedures in our paper, and suggest that other studies based on a similar approach have “serious flaws” and must be “regarded with suspicion.”

Smith et al.’s condemnation stems from a series of misunderstandings about multiple taper method spectrum estimation, conventional and robust autoregressive models, significance thresholds and Bonferroni corrections. Below, we review these concepts in detail in order to correct the record. We also discuss the inadequacies of their synthetic example, the unique problem of uncertain time scales in cyclostratigraphy, and the recent move by the American Statistical Association to redirect the use of *P*-values in science that is particularly relevant to this discussion. We conclude with a review of objective methodologies that represent a significant step forward in cyclostratigraphic research.

2. MTM spectrum estimation in cyclostratigraphy

An important goal in cyclostratigraphy is to determine whether Milankovitch (astronomical) forcing has been preserved in the form of stratigraphic cycles. This has led to widespread use of the periodogram, or spectrum, in particular the multiple (Slepian) taper method (MTM) spectrum estimator (Thomson, 1982). The MTM estimator is unique in its “eigen-expansion” of the spectrum to obtain significantly more independent degrees of freedom and a narrower averaging bandwidth than is possible with other spectral estimators (Park et al., 1987; Percival and Walden, 1993). These among other special features of the MTM estimator are essential for the effective analysis of short, noisy time series such as those encountered in cyclostratigraphy. Smith et al.’s repeated reference to the unsmoothed periodogram is therefore surprising, as it is well known to be unreliable (Thomson, 1977a, 1977b, 2009).

The SSA-MTM Toolkit (Ghil et al., 2002) was one of the first freeware packages to compute MTM power spectra; another earlier very influential package was Analyseries (Paillard et al., 1996). Notably, the

---

* Corresponding authors at: State Key Laboratory of Biogeology and Environmental Geology, China University of Geosciences, Beijing, 100083, China.
E-mail addresses: lhinnov@gmu.edu (L.A. Hinnov), whcgeo@cugb.edu.cn (H. Wu).

http://dx.doi.org/10.1016/j.palaeo.2016.07.030

SSA-MTM Toolkit carries out adaptive weighting of the MTM spectrum, and provides the means for hypothesis testing with a choice of spectral noise models. Recently *astrochron* (Meyers, 2014) was launched with many of the same tools, and with significantly more options and flexibility than the earlier packages.

Fang et al. (2015) used the SSA-MTM Toolkit to compute MTM power spectra for their cyclostratigraphic data, with hypothesis testing using conventional AR1 noise modeling. The latter step is the focus of Smith et al.'s strong disapproval, and is discussed further below.

3. Spectral noise (null) models: conventional and robust AR1 models

A key goal is to discriminate between "signal" and "noise" in cyclostratigraphic data series. Noise tends to occur at all frequencies with random phases at a power level that is generally–but not exclusively–lower than that of the signal, which is confined to narrow frequency bands and lines (single frequencies) with constant or slowly varying phase. In cyclostratigraphy one would expect to detect significant cycles as lines associated with Milankovitch forcing (eccentricity, obliquity, precession), together with noise from concurrent processes nonresponsive to Milankovitch forcing. The number of detectable lines depends on the Milankovitch forcing, climate response, stability of the recording medium (e.g., sediment accumulation rate), data sampling, and the noise. Since Milankovitch forcing is detected as a response, there is also the potential for line damping and generation of nonlinearities.

For detection of significant cycles in a cyclostratigraphic data series, the practice of hypothesis testing has become routine in which a "null hypothesis" H_0 assumes that the data series is a random process represented by a null (noise) model. The "alternative hypothesis" H_A is that the data series represents a combination of a nonrandom process and random chance. Frequencies at which the power spectrum of the data exceeds that of the noise model by a statistically significant margin are marshaled to reject H_0 and accept H_A (see Sections 3 and 4). This approach cannot detect lines with power that does not exceed that of the noise, which can be addressed by the MTM harmonic F-test (Thomson, 1990, 2009).

The autoregressive order 1 ("AR1") red noise model has long been favored as a suitable spectral "null" model for climate processes (e.g., Gilman et al., 1963). In this respect, Smith et al. are right to question the rather uncritical use of the AR1 noise model in cyclostratigraphy: there are many opportunities for Earth surface processes to filter Milankovitch-forced climate and "redden" the stratigraphic response beyond any simple AR1 process that may have characterized the original climate (see Section 7). Anticipating this problem, Schulz and Mudelsee (2002) devised a non-parametric runs test to evaluate the fit of the conventional AR1 model spectrum to the data spectrum in their REDFIT freeware. Vaughan et al. (2011) also raise the null model fit problem, suggesting the use of alternative (power-law) null models.

The robust AR1 model is a modification of the conventional AR1 model in which lines (with high power) assumed to represent signal are statistically excluded from the model (Mann and Lees, 1996). It is an attractive alternative to the conventional AR1 model, which overestimates spectral noise as it assumes the data series to be completely random even if it includes nonrandom signal. The SSA-MTM Toolkit computes both conventional and robust AR1 models ("Raw" and "Robust" in the MTM Options panel), the latter with two adjustable parameters. Unfortunately, it is now known that the robust AR1 model tends to underestimate noise at low frequencies (Meyers, 2012). Therefore Fang et al. (2015) elected not to use the robust AR1 model.

Fig. 1A compares conventional and robust AR1 model spectra to the 2π MTM spectrum of a synthetic AR1 series $N_0(t)$ for $n = 1340$ points with $\rho = 0.7$ to conform with the example in Smith et al. (See Supplementary Information, Fig. S1 for a display.) Assuming that they selected the SSA-MTM Toolbox default parameters (as we have), in their Fig. 1 a 2π MTM power spectrum was computed for their synthetic series. Smith et al. computed the default robust AR1 model, which is log-fitted with a median smoothing window set to 20% of the Nyquist frequency range. Our Fig. 1A shows that at low frequencies the conventional AR1 model has higher values than the robust AR1 model; this results in a lower apparent statistical significance of the spectrum at frequencies <1 cycles/m.

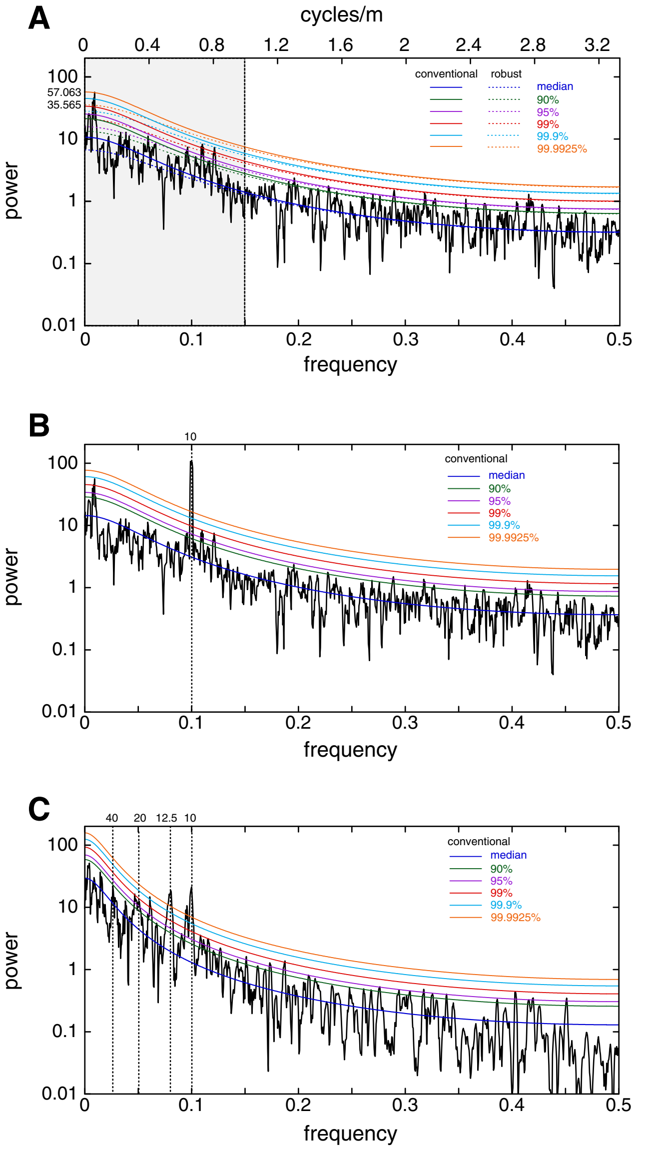

Fig. 1. 2π MTM power spectral analysis and AR1 modeling of synthetic time series (available in the Supplementary Information). All analysis was performed using SSA-MTM Toolbox, MTM Analysis tools, output files "mtmspec-raw.out", "mtmspec-conf.out" and "mtm-spec.inf". In the toolbox, the input data of length $n = 1340$ is zero-padded to 2048 points (next power of 2); the output spectra have 1024 sampled frequencies. The SSA-MTM Toolbox error (see text) is corrected for all significance thresholds. A. Power spectrum of $N_0(t)$ with conventional (solid lines) and robust with log fitting and 20% median smoothing (dashed lines) AR1 models (estimated $\rho = 0.7052$) and significance thresholds. Bonferroni corrections discussed in the text are shown as 99.9% and 99.9925% thresholds. The values of the 99.9925% threshold at $f = 0$ for the conventional (57.063) and robust (35.565) models are indicated on the y-axis. The two models differ significantly over $f = [0,1$ cycles/m$]$ (shaded region). The top x-axis shows frequency in "cycles/m" assuming $\Delta t = 0.15$ m as in Smith et al. Spectral peaks exceeding the 99% threshold of the conventional AR1 model are at $f = 0.0636$, 0.7342, 0.8094, and 2.763. B. Power spectrum of $N_0(t) + s_0(t)$ with the conventional AR1 model (estimated $\rho = 0.7245$). The periodicity of $s_0(t)$ is clearly visible, and denoted by the vertical dashed line. C. Power spectrum of $N_1(t) + s_1(t)$ with the conventional AR1 model (estimated $\rho = 0.8756$). The four vertical dashed lines indicate the different signal frequencies arising from the four sedimentation rates (see Fig. S1C).