Supporting Information for

Bayesian estimation of past astronomical frequencies, lunar distance, and length of day from sediment cycles

A. Malinverno[1] and S. R. Meyers[2]

[1]Lamont-Doherty Earth Observersity of Columbia University, Palisades, New York, USA

[2]Department of Geoscience, University of Wisconsin-Madison, Madison, Wisconsin, USA

Contents of this file

Text S1. Calculation of the Likelihood for the Spectral Fit

To calculate the data $\mathbf{d}_{\text{pred}}$ predicted by parameters in $\mathbf{m}$, the first step is to transform the stratigraphic depth in the data to an age vector $\mathbf{t}$ (which is zero at the top) using the sedimentation rate u in $\mathbf{m}$. (Symbols and acronyms used here are listed in Table S1.) To compute the predicted data, we then construct a matrix $\mathbf{G}$ whose columns contain sine and cosine signals as a function of ages in $\mathbf{t}$ with the frequencies of eccentricity, obliquity, and climatic precession listed in Table 1 of the main text calculated for the values of g_i, s_i, and k in $\mathbf{m}$.

We then obtain by least squares the amplitudes of the sine and cosine terms in a vector $\mathbf{y}$ fitted to the data as

$$\mathbf{y} = (\mathbf{G}^{\mathsf{T}}\mathbf{G})^{-1}\mathbf{G}^{\mathsf{T}}\mathbf{d}, \tag{S1}$$

where the data in $\mathbf{d}$ were detrended and standardized to zero mean and unit variance. The data predicted by the astronomical frequencies are

$$\mathbf{d}_{\text{pred}} = \mathbf{G}\,\mathbf{y}. \tag{S2}$$

The calculation of the spectral likelihood is based on fitting an AR(2) process to the vector $\mathbf{e}$ of residuals that are the difference between observed and predicted data as in

$$\mathbf{e} = \mathbf{d} - \mathbf{d}_{\text{pred}}. \tag{S3}$$

Following an empirical Bayes strategy, we estimate the AR(2) coefficients ϕ_1 and ϕ_2 from $\mathbf{e}$ with a method originally due to Burg (1967) that is based on minimizing the prediction error variance computed both in the forward and backward direction (Andersen, 1974; Ulrych & Bishop, 1975). Using the fitted AR(2) coefficients, we then compute the vector $\mathbf{w}$ of the driving noise from

$$w_i = e_i - \phi_1 e_{i-1} - \phi_2 e_{i-2}. \tag{S4}$$

If $\mathbf{e}$ can be successfully modeled as an AR(2) process, the vector $\mathbf{w}$ should be close to uncorrelated white noise. This can be checked from the sample autocorrelation of $\mathbf{w}$, shown in Figures S1b to S3b for the data sets examined here.

If residuals $\mathbf{e}$ are modeled as an AR process, the spectral fit likelihood can be written as the multivariate normal PDF of the noise $\mathbf{w}$ (Dettmer et al., 2012), which is

$$p(\mathbf{d} \mid \mathbf{m}) = p(\mathbf{w} \mid \mathbf{m}) = \frac{1}{(2\pi)^{N/2} \, \sigma_w^N} \exp\left[-\frac{\mathbf{w}^\top \mathbf{w}}{2\sigma_w^2}\right], \tag{S5}$$

where N is the number of data points in $\mathbf{d}$ and $\mathbf{w}$. Applying again empirical Bayes, we estimate the variance σ_w^2 from $\mathbf{w}$ as

$$\sigma_w^2 = \frac{\mathbf{w}^\top \mathbf{w}}{N}. \tag{S6}$$

Substituting this estimate of σ_w^2 in Equation S5 we obtain a final expression for the spectral fit likelihood that is

$$p(\mathbf{d} \mid \mathbf{m}) = p(\mathbf{w} \mid \mathbf{m}) = \frac{1}{(2\pi)^{N/2} \, \sigma_w^N} \exp\left[-\frac{N}{2}\right], \tag{S7}$$

Text S2. Calculation of the Likelihood for the Envelope Fit

In the envelope fit, the observed data vector $\mathbf{d}$ is the amplitude envelope of the climatic precession signal in the data and $\mathbf{d}_{\text{pred}}$ is the envelope predicted by eccentricity signals with the frequencies given by the values of g_i in $\mathbf{m}$. The precession envelope of the data is calculated by first extracting the signal in the climatic precession frequency band applying a Taner bandpass filter with a roll-off rate of 10^7 (Zeeden et al., 2018). The cutoff frequencies equal the minimum climatic precession frequency minus 0.005 cycles/kyr and the maximum climatic precession frequency plus 0.005 cycles/kyr. The envelope of the climatic precession signal is then computed from its Hilbert transform and is standardized to have zero mean and unit variance.

The predicted precession envelope is calculated by fitting the eccentricity frequencies given by the g_i values in $\mathbf{m}$ to the precession envelope of the data; the result is then standardized to zero mean and unit variance. The procedure is the same as that described for the spectral likelihood, except that the matrix $\mathbf{G}$ contains only sine and cosine terms for the eccentricity frequencies. The residual vector $\mathbf{e}$ in the envelope fit is the difference between the precession envelope extracted from the data by bandpass filtering and that predicted by the eccentricity frequencies defined by the values in $\mathbf{m}$.

The envelope likelihood calculation is based on an effective number of observations

$$N_{\text{eff}} = N/\tau < N, \tag{S8}$$

where $\tau > 1$ is the lag where the autocorrelation of the envelope residuals decays to zero (see Section 2.2 in the main text). Consequently, if we were to subsample the vector of residuals $\mathbf{e}$ by

taking one every τ values, we would obtain a vector of uncorrelated residuals $\mathbf{e}_{sub}$ of length N_{eff}. The likelihood for these subsampled uncorrelated residuals would be a multivariate normal PDF as in

$$p(\mathbf{d} \mid \mathbf{m}) = p(\mathbf{e}_{sub} \mid \mathbf{m}) = \frac{1}{(2\pi)^{N_{eff}/2} \, \sigma_e^{N_{eff}}} \exp\left[-\frac{\mathbf{e}_{sub}^{\mathsf{T}}\mathbf{e}_{sub}}{2\sigma_e^2}\right]. \tag{S9}$$

As the vector $\mathbf{e}_{sub}$ consists of N_{eff} elements, the sum of their squared values will be approximately the same as the sum of squares of $\mathbf{e}$ times the ratio N_{eff}/N as in

$$\frac{\mathbf{e}_{sub}^{\mathsf{T}}\mathbf{e}_{sub}}{\sigma_e^2} \approx \frac{\mathbf{e}^{\mathsf{T}}\mathbf{e}}{\sigma_e^2} \frac{N_{eff}}{N} = \frac{\mathbf{e}^{\mathsf{T}}\mathbf{e}}{\tau\sigma_e^2}, \tag{S10}$$

which shows that correlations in the vector $\mathbf{e}$ result in an increase of the variance σ_e^2 by a factor τ. If we follow an empirical Bayes strategy and estimate σ_e^2 from the sample variance of $\mathbf{e}$ as in

$$\sigma_e^2 = \frac{\mathbf{e}^{\mathsf{T}}\mathbf{e}}{N}, \tag{S11}$$

the likelihood of the subsampled vector of residuals can be written as

$$p(\mathbf{d} \mid \mathbf{m}) = p(\mathbf{e}_{sub} \mid \mathbf{m}) = \frac{1}{(2\pi)^{N_{eff}/2} \, \sigma_e^{N_{eff}}} \exp\left[-\frac{N_{eff}}{2}\right]. \tag{S12}$$

(Note that the calculation of the envelope fit likelihood never requires the hypothetical vector $\mathbf{e}_{sub}$.)

We estimate the lag τ from a simple model of the autocorrelation of envelope fit residuals that contain periodic components with eccentricity frequencies. The autocorrelation will first cross zero at a lag that is approximately a quarter of the wavelength λ_{ecc} of the shortest eccentricity cycle in the data, so that

$$\tau \approx \frac{\lambda_{ecc}}{4\Delta z}, \tag{S13}$$

where Δz is the data sampling interval and λ_{ecc} will equal uT_{ecc}, where u is the sedimentation rate and T_{ecc} is the period of the shortest eccentricity cycle (e.g., 100 kyr), so that

$$\tau = \frac{uT_{ecc}}{4\Delta z}. \tag{S14}$$

However, making τ proportional to a variable sedimentation rate has the effect of inducing a systematic bias that makes the envelope likelihood in Equation S12 substantially larger at lower sedimentation rates where τ is lower and N_{eff} is larger. (The measured and predicted envelope data are standardized to unit variance, so that the residual $\sigma_e < 1$; for a given value of σ_e, the likelihood will be greater for a greater value of N_{eff}.) Numerical experiments show that this sedimentation rate bias overwhelms the effect of differences in σ_e, and the highest likelihood for the envelope fit is invariably at the lowest sedimentation rate considered. To avoid this bias, the value of τ is calculated from Equation S14 at a reference value of u, set to the average sedimentation rate in the prior range considered.

Text S3. Lunar Distance and LOD from the Axial Precession Frequency

In this section, we describe our approach and provide easy-to-use expressions to obtain lunar distance and LOD from estimates of the axial precession frequency. To simplify the notation, henceforth we denote the ratios between past and present values of axial precession frequency k, lunar distance a, and Earth spin rate ω as follows:

$$k_r = \frac{k(t)}{k(0)}, \qquad a_r = \frac{a(t)}{a(0)}, \qquad \omega_r = \frac{\omega(t)}{\omega(0)}, \tag{S15}$$

where t is age and 0 denotes present day. With this notation, Equation 5 in the main text for the K-curve is

$$\omega_r = k_r \frac{K+1}{K + a_r^{-3}} \tag{S16}$$

and Equation 6 for the AM-curve is

$$\omega_r = 1 + A - A \sqrt{a_r} \tag{S17}$$

The values of the constants K and A in Equation S16 and S17 were adjusted from the original values in Walker & Zahnle (1986) to better fit the relationship between lunar distance, Earth spin rate, and axial precession frequency in the tidal dissipation model results of Farhat et al. (2022), which take into account the long-term increase in the obliquity ε of the Earth's axis and the effect of solar ocean tides in slowing down Earth's rotation (see Section 7 in the the main text).

   The value of k that can be predicted for given values of a and ω by rearranging Equation S16 and the original value of $K = 0.465$ in Walker & Zahnle (1986) results in a difference with the k computed by Farhat et al. (2002) that increases with increasing age, reaching 3.3% at 3.3 Ga. This misfit matches the predicted effect of the secular trend in obliquity: as the full expression for the axial precession frequency contains a $\cos \varepsilon$ term (e.g., Berger & Loutre, 1994; Laskar, 2020), the ratio $k(t)/k(0)$ at age t will be multiplied by a factor $\cos \varepsilon(t)/\cos \varepsilon(0)$, which equals 1.032 at 3.3 Ga in the results of Farhat et al. (2022). By adjusting the value of K to 0.358, the difference between the k predicted from Equation S16 and that computed by Farhat et al. (2022) remains within ±0.14% between the present and 3 Ga, increasing to 0.28% at 3.3 Ga.

   We also compared the LOD predicted by Equation S17 for a given lunar distance a to the values computed by Farhat et al. (2022). The original value of $A = 4.87$ in Walker & Zahnle (1986) results in a difference in LOD with the Farhat et al. (2022) values that increases with increasing age, reaching 0.55% at 3.3 Ga. With an adjusted value of $A = 4.81$, this difference remains between ±0.03% from the present to 3 Ga, reaching −0.07% at 3.3 Ga.

   At the intersection of the K-curve and of the AM-curve, the value of ω_r in Equation S16 and S17 must be the same, and we obtain an expression for the ratio k_r that conserves Earth-Moon angular momentum and is a function of a_r only:

$$k_r = \left(1 + A - A \sqrt{a_r}\right) \frac{K + a_r^{-3}}{K + 1}. \tag{S18}$$

This nonlinear equation cannot be solved directly for a_r. However, an accurate value can be obtained by computing the value of k_r from Equation S18 for a_r between 1 and 0.77 (the value at 3.3 Ga in the results of Farhat et al. 2022) and fitting a polynomial to a_r as a function of k_r. The relationship

between a_r and the natural logarithm of k_r is almost linear, and the third order polynomial in the following equation fits the values of a_r to within $\pm 3 \times 10^{-5}$:

$$a_r = 1 - 0.217194 \log(k_r) - 0.00060922 \left[\log(k_r)\right]^2 + 0.00621404 \left[\log(k_r)\right]^3. \qquad \text{(S19)}$$

Once the ratio a_r is obtained from k_r using Equation S19, ω_r can be computed from Equation S17.

Text S4. Uncertainty in Estimated Lunar Distance and LOD

The K-curve and AM-curve will intersect at a point of coordinates $\hat{a}_r$ and $\hat{\omega}_r$, which define the lunar distance and Earth spin rate at a past time from an estimate of the axial precession frequency (Figure 14a in the main text). This section describes how to obtain the uncertainty in $\hat{a}_r$ and $\hat{\omega}_r$ from the uncertainty in an estimated value of axial precession frequency ratio $\hat{k}_r$. There are three sources of uncertainty to take into account in this problem:

- Uncertainty in the precession frequency ratio $\hat{k}_r$ estimated from stratigraphic data, quantified by a standard deviation $\sigma_{\hat{k}_r}$;

- Uncertainty in the value of ω_r predicted by the K-curve in Equation S16 for a given a_r and k_r, quantified by a standard deviation σ_{K};

- Uncertainty in the value of ω_r predicted by the AM-curve in Equation S17 for a given a_r, quantified by a standard deviation σ_{AM}.

The standard deviation $\sigma_{\hat{k}_r}$ equals the posterior standard deviation of k divided by the present day axial precession frequency $k(0)$. A thorough determination of the standard deviations σ_{K} and σ_{AM} requires quantifying and propagating uncertainties in the constants that define the K-curve and AM-curve (in Equation S16 and S17), and is beyond the scope of the present study. As a first approximation, we use here a value of 0.01 for both σ_{K} and σ_{AM}, meaning that the K-curve and AM-curve predict the value of the ratio ω_r with a standard deviation of 1%. This value is well above the misfit in fitting the values calculated by Farhat et al. (2022) with the adjusted values of K in Equation S16 and of A in Equation S17 (see Text S3), thus we take it as providing a conservative measure of the the uncertainty.

The uncertainty of the AM-curve in the region around the intersection at $\hat{a}_r$ and $\hat{\omega}_r$, where we approximate the AM-curve by a straight line, can be represented by a bivariate normal PDF that is centered on the intersection and has a covariance matrix

$$\mathbf{C}_{\text{AM}} = \sigma_0^2 \begin{bmatrix} 1 & b_{\text{AM}} \\ b_{\text{AM}} & b_{\text{AM}}^2 + \frac{\sigma_{\text{AM}}^2}{\sigma_0^2} \end{bmatrix}, \qquad \text{(S20)}$$

where b_{AM} is the slope of the AM-curve (Equation S17) at the intersection

$$b_{\text{AM}} = \left. \frac{d\omega_r}{da_r} \right|_{a_r = \hat{a}_r} = -\frac{A}{2\sqrt{\hat{a}_r}} \qquad \text{(S21)}$$

and σ_0 spans a range of a_r that is large compared to the overlap between the uncertain AM-curve and K-curve; see Figure 14b in the main text for an illustration. (As shown below, σ_0 will be eliminated from our final expressions.)

The bivariate normal PDF that represents the uncertainty of the K-curve in the region around the intersection is

$$\mathbf{C_K} = \sigma_0^2 \begin{bmatrix} 1 & b_K \\ b_K & b_K^2 + \frac{\sigma_K^2 + \sigma_{\hat{k}_r}^2}{\sigma_0^2} \end{bmatrix}, \tag{S22}$$

where b_K is the slope of the K-curve (Equation S16) at the intersection

$$b_K = \frac{d\omega_r}{da_r}\bigg|_{a_r=\hat{a}_r} = \hat{k}_r \frac{3(K+1)\hat{a}_r^2}{(K\hat{a}_r^3 + 1)^2} \tag{S23}$$

and the sum $\sigma_K^2 + \sigma_{\hat{k}_r}^2$ accounts for the uncertainty of both the K-curve and of the estimated k_r. (Equation S16 shows that the uncertainty in the value of ω_r predicted by k_r is the same as the uncertainty of k_r.)

The uncertainty of the intersection will be defined by the product of the two PDFs that quantify the uncertainty of the AM-curve and of the K-curve. This product is another bivariate normal PDF that is centered on the intersection point and that has a covariance matrix

$$\mathbf{C} = \left[\mathbf{C_K^{-1}} + \mathbf{C_{AM}^{-1}}\right]^{-1}, \tag{S24}$$

and an expression for $\mathbf{C}$ can be obtained letting $\sigma_0^2 \to \infty$ as

$$\mathbf{C} = \frac{1}{(b_K - b_{AM})^2} \begin{bmatrix} \sigma_{AM}^2 + \sigma_K^2 + \sigma_{\hat{k}_r}^2 & b_{AM}(\sigma_K^2 + \sigma_{\hat{k}_r}^2)^2 + b_K\sigma_{AM}^2 \\ b_{AM}(\sigma_K^2 + \sigma_{\hat{k}_r}^2) + b_K\sigma_{AM}^2 & b_{AM}^2(\sigma_K^2 + \sigma_{\hat{k}_r}^2) + b_K^2\sigma_{AM}^2 \end{bmatrix}. \tag{S25}$$

The diagonal elements of $\mathbf{C}$ contain the variances of the values of $\hat{a}_r$ (C_{11}) and $\hat{\omega}_r$ (C_{22}) at the intersection of the AM-curve and K-curve.

Table S1. List of symbols and acronyms.

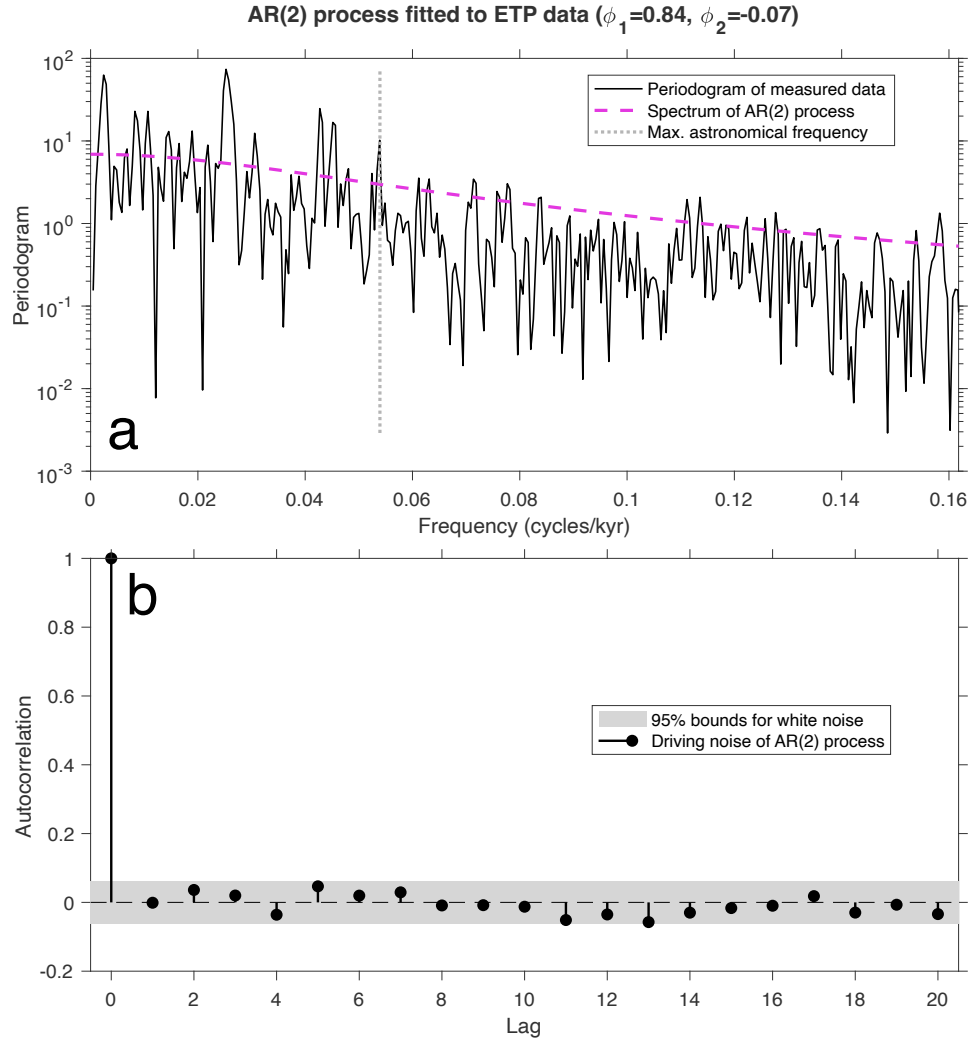|	
a	Semi-major axis of lunar orbit
a_r	Ratio $a(t)/a(0)$ of a at age t over the present day value
A	Constant in equation for conservation of angular momentum
$\mathbf{C}$	Covariance matrix
$\mathbf{d}$	Vector of observed data
$\mathbf{d}_{\mathrm{pred}}$	Vector of data predicted by a given value of $\mathbf{m}$
$\mathbf{e}$	Vector of data residuals $\mathbf{d} - \mathbf{d}_{\mathrm{pred}}$
$\mathbf{e}_{\mathrm{sub}}$	Subsampled vector of data residuals
$\mathbf{G}$	Matrix of sine and cosine terms with astronomical frequencies
g_i	Fundamental Solar system frequencies for the rotation of the planetary perihelia
k	Precession frequency of the Earth's spin axis
k_r	Ratio $k(t)/k(0)$ of k at age t over the present day value
$\mathbf{m}$	Vector of parameters (g_i, s_i, k, u)
N	Number of data points in $\mathbf{d}$
N_{eff}	Effective number of independent observations in $\mathbf{d}$
$p(x \mid y)$	Probability density function (PDF) of x given y
P	Order of an AR(P) process
r_i	Autocorrelation of data residuals $\mathbf{e}$ at lag i
s_i	Fundamental Solar system frequencies for the rotation of the ascending nodes of the orbital planes
$\mathbf{t}$	Vector of ages in the stratigraphic data
T_{ecc}	Period of shortest eccentricity cycle in the data
u	Sedimentation rate
$\mathbf{w}$	Vector of uncorrelated (white) noise driving an autoregressive (AR) process
λ_{ecc}	Wavelength of shortest eccentricity cycle in the data
ϕ_i	i-th coefficient of an AR(P) process ($1 \leq i \leq P$)
σ_e^2	Variance of data residuals $\mathbf{e}$
σ_w^2	Variance of uncorrelated noise $\mathbf{w}$
τ	Minimum lag where the autocorrelation function crosses zero
ω	Earth spin rate
ω_r	Ratio $\omega(t)/\omega(0)$ of ω at age t over the present day value
AR	Autoregressive process
LOD	Length of day
MAP	Maximum a posteriori
MCMC	Markov chain Monte Carlo
PDF	Probability density function

Figure S1. Fit of an AR(2) process to the residuals for the synthetic ETP test data set. (a) Comparison of data periodogram with the spectrum of an AR(2) process with the coefficients ϕ_1 and ϕ_2 fitted for the MAP value of the parameters. The vertical dotted line marks the maximum climatic precession frequency in the data. (b) Sample autocorrelation of the driving noise **w** of the residuals obtained for the MAP values of the parameters.
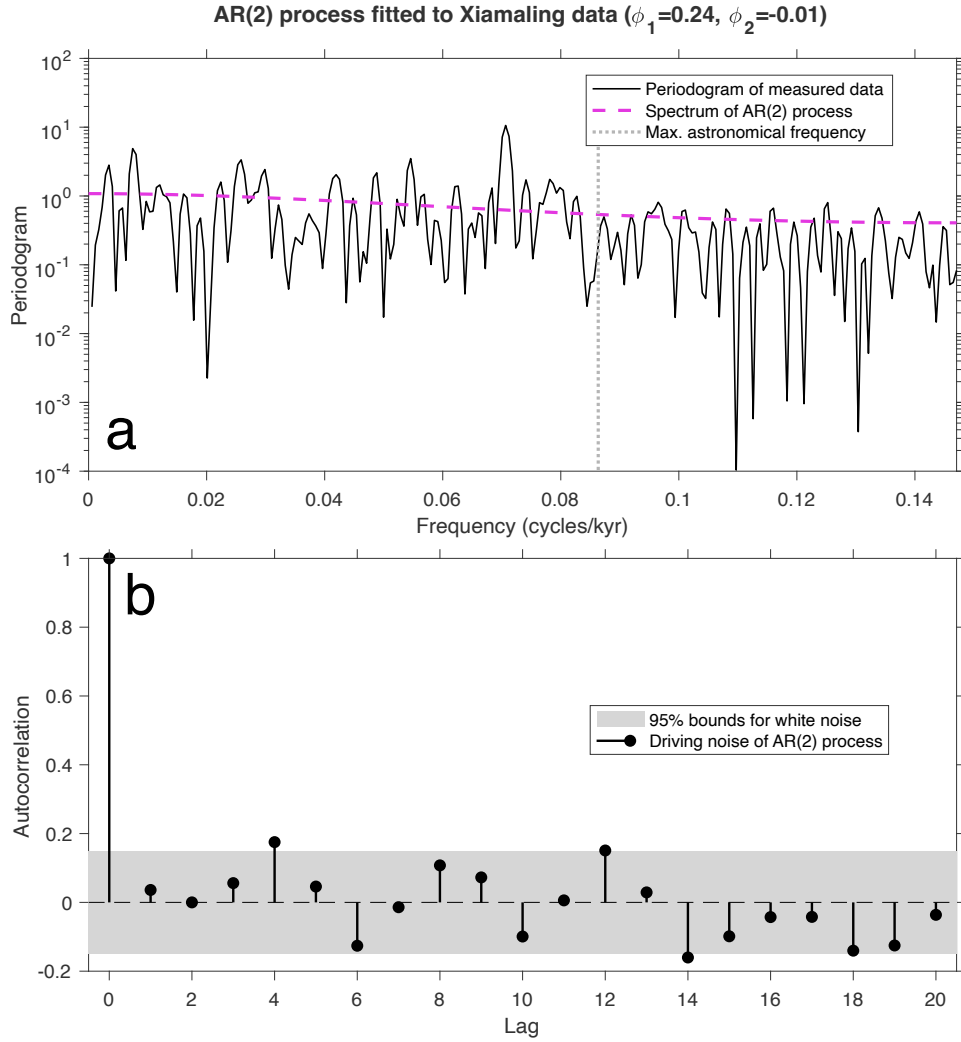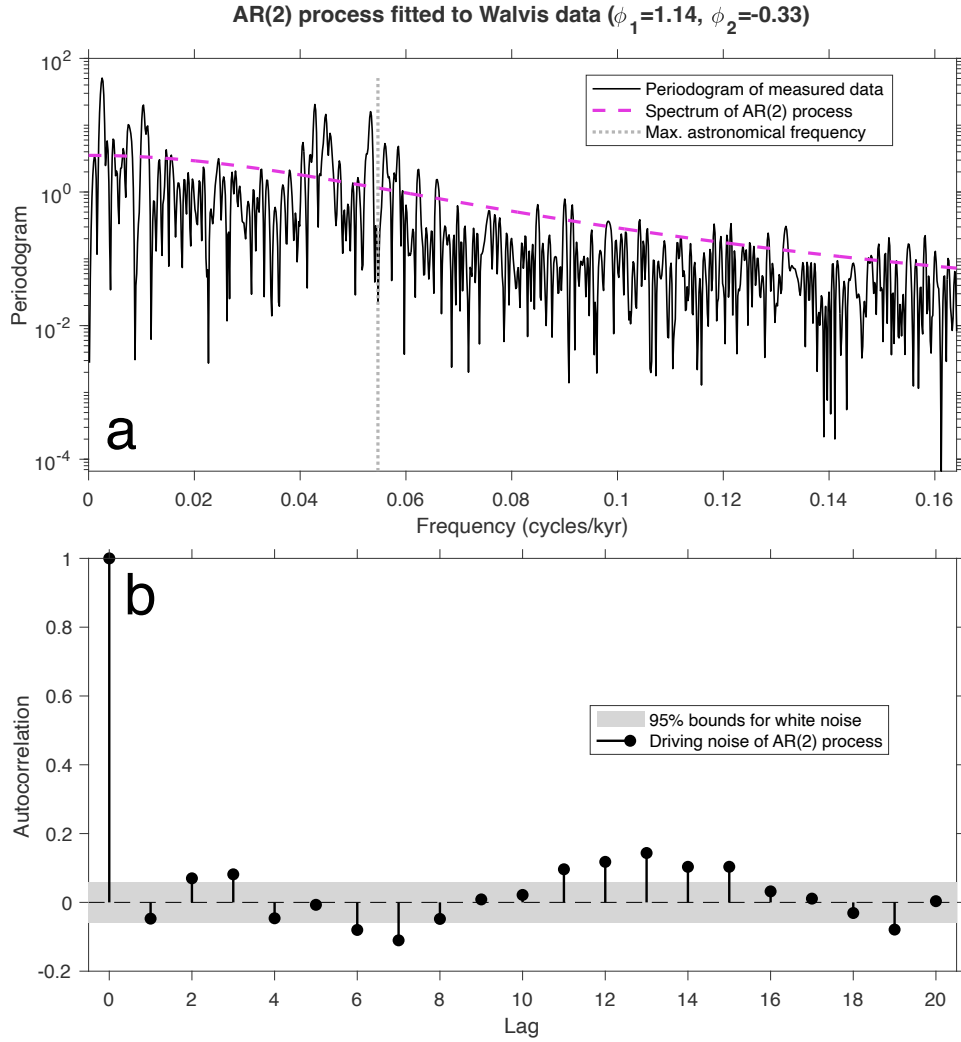
Figure S2. Fit of an AR(2) process to the residuals for the Xiamaling formation Cu/Al data set. (a) Comparison of data periodogram with the spectrum of an AR(2) process with the coefficients ϕ_1 and ϕ_2 fitted for the MAP value of the parameters. The vertical dotted line marks the maximum climatic precession frequency in the data. (b) Sample autocorrelation of the driving noise **w** of the residuals obtained for the MAP values of the parameters.

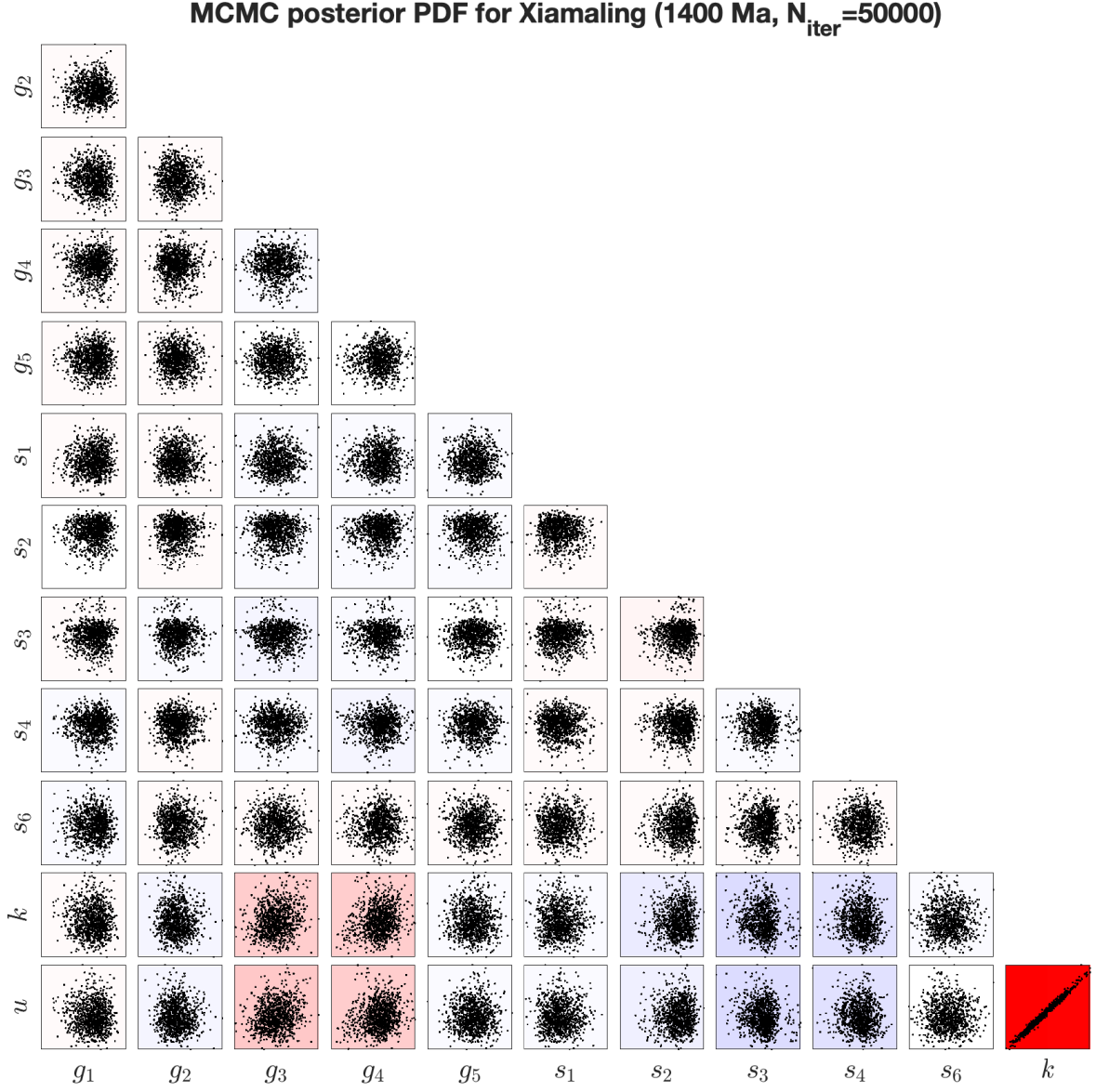**Figure S3.** Fit of an AR(2) process to the residuals for the Walvis Ridge a* data set. (a) Comparison of data periodogram with the spectrum of an AR(2) process with the coefficients ϕ_1 and ϕ_2 fitted for the MAP value of the parameters. The vertical dotted line marks the maximum climatic precession frequency in the data. (b) Sample autocorrelation of the driving noise **w** of the residuals obtained for the MAP values of the parameters.

Figure S4. Posterior correlation of parameters sampled by TimeOptBMCMC for the Xiamaling formation Cu/Al data set. The color background for each pair of parameters is proportional to the correlation coefficient (blue for −1, white for zero, and red for +1). Posterior correlations are near zero, with the exception of a strong positive correlation between *u* and *k*.
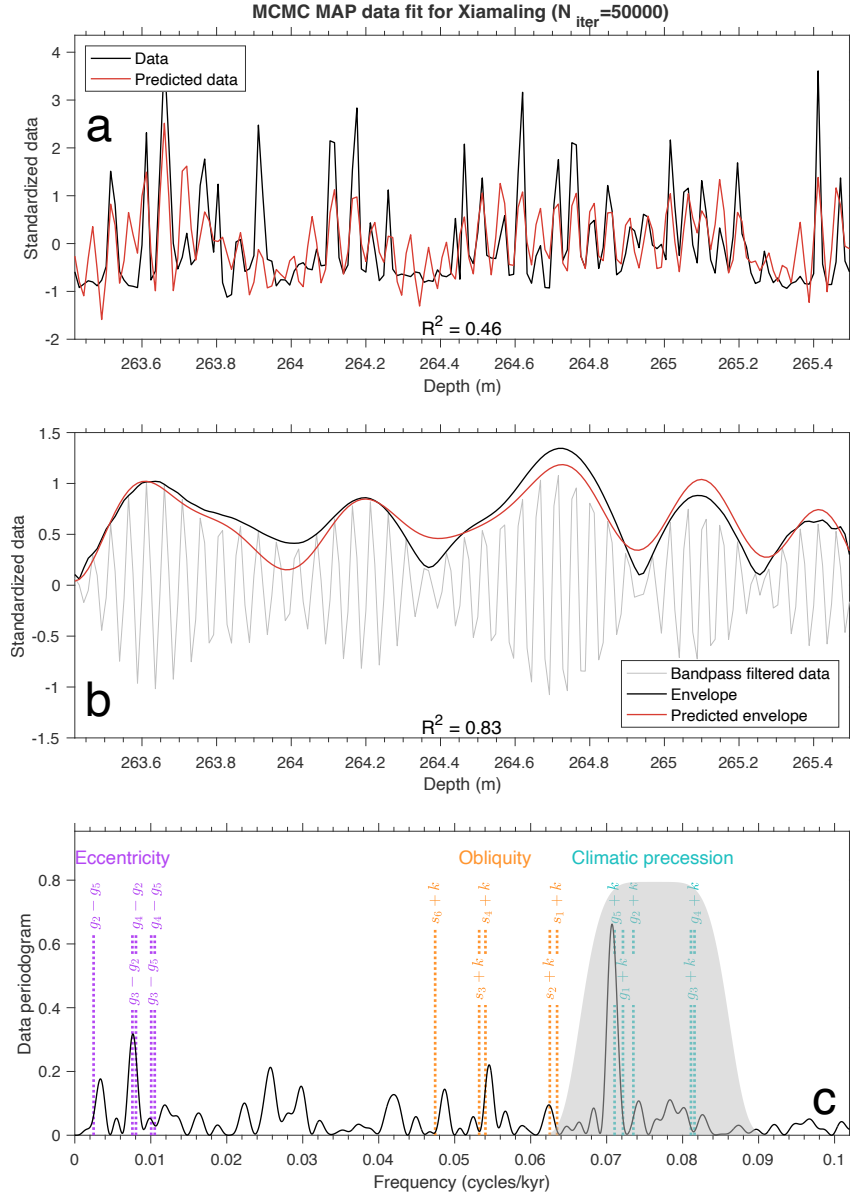
Figure S5. Fit to the Xiamaling formation Cu/Al data obtained by TimeOptBMCMC for the MAP value of sedimentation rate u and astronomical frequencies g_i, s_i, and k. (a) Fit between measured and predicted stratigraphic data (spectral fit). (b) Fit of the bandpassed climatic precession signal (envelope fit). (c) Data periodogram (black continuous line) and frequencies of the reconstructed astronomical cycles in the data (dotted vertical lines). The gray shaded area shows the frequency response of the filter used to compute the bandpassed climatic precession signal in the data (gray curve in (b)).
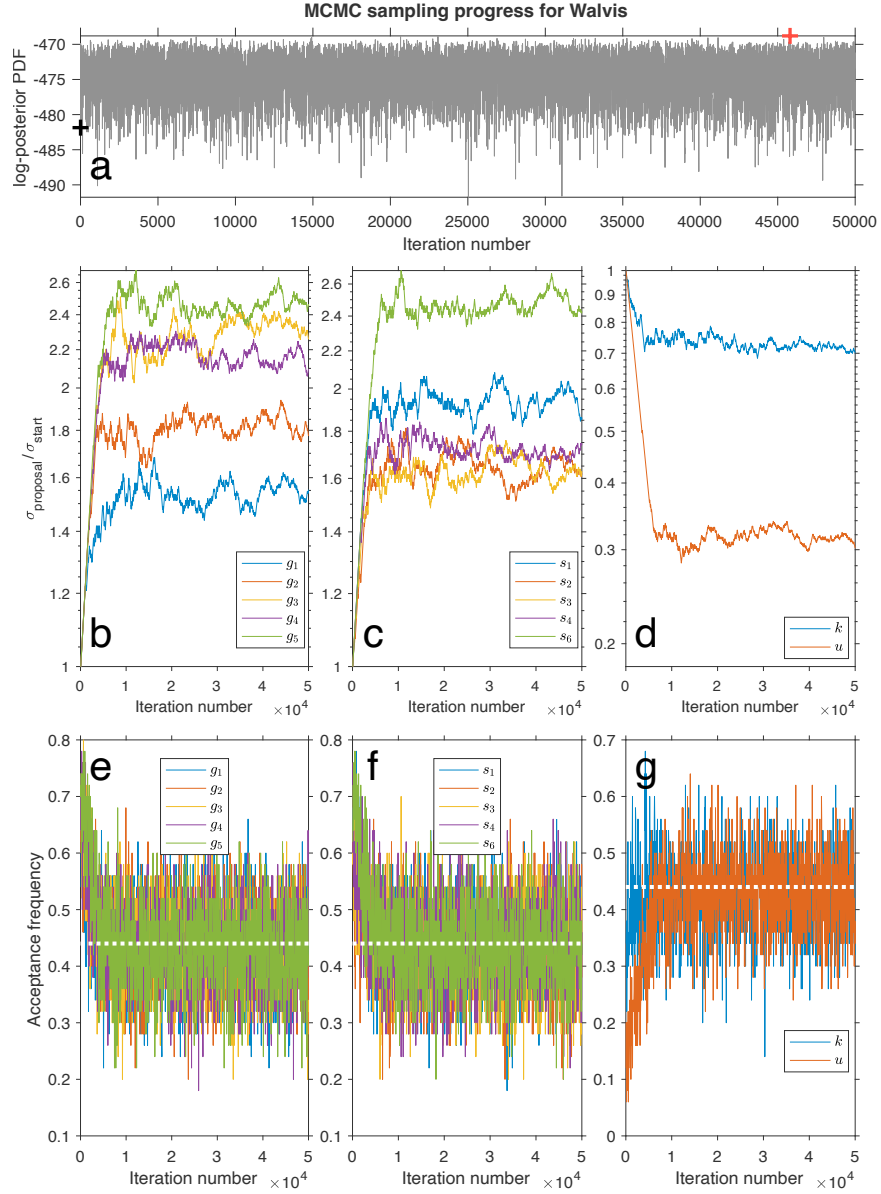
Figure S6. Progress of TimeOptBMCMC sampling for the Walvis Ridge a* data set over 50,000 iterations. (a) Value of the log-posterior PDF for the sampled model parameter vectors. The black cross is the starting value and the red cross the MAP. (b, c, d) Standard deviation of the proposal PDF (as a ratio over the starting value) for each model parameter. (e, f, g) Frequency of acceptance of the proposed steps in the MCMC random walk. The adaptive Metropolis algorithm used in TimeOptBMCMC adjusts the standard deviations of the proposal PDF to keep the frequency of acceptance around the optimal value of 0.44 for all model parameters (white horizontal dotted line).
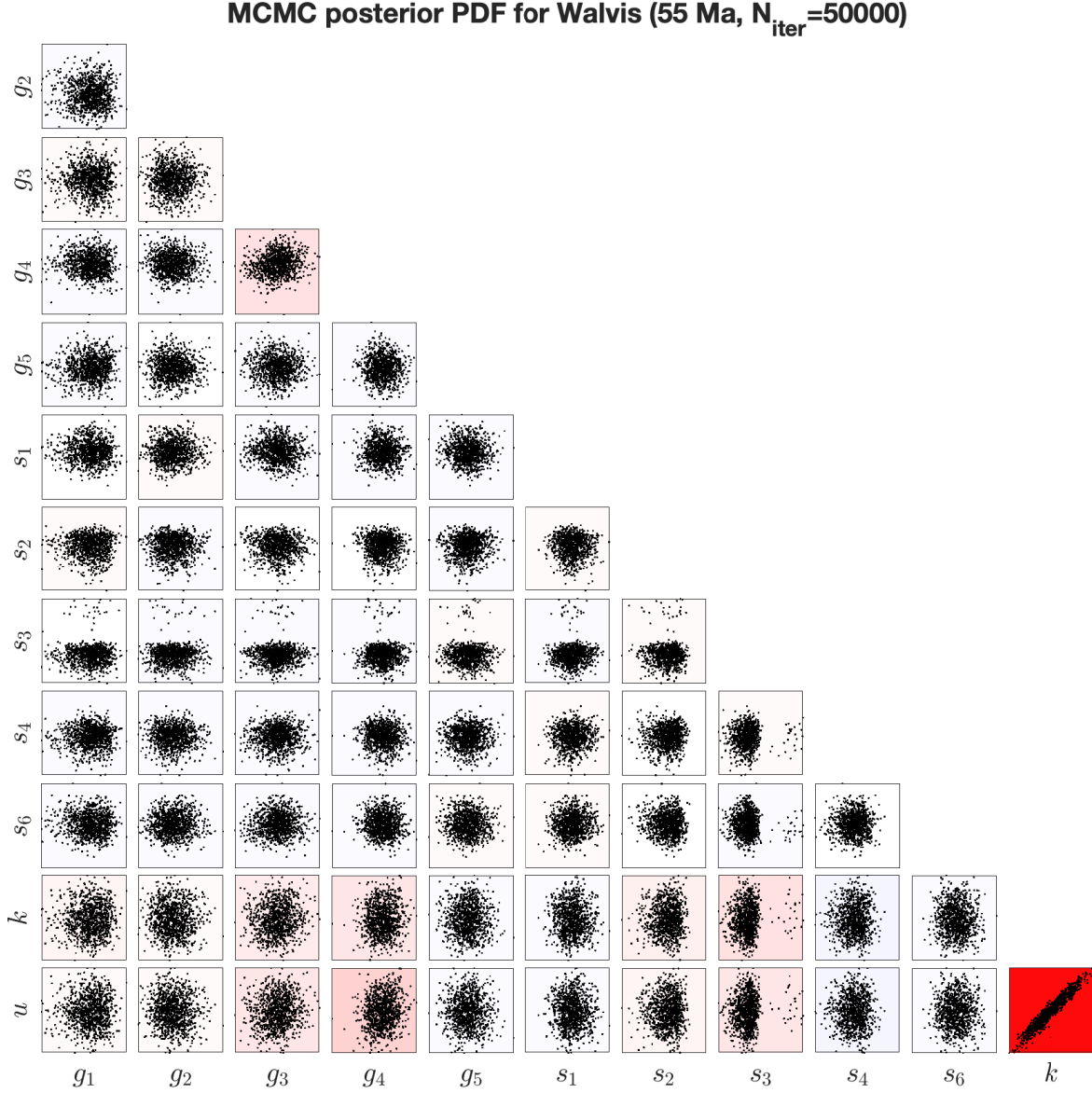
13

Figure S7. Posterior correlation of parameters sampled by TimeOptBMCMC for the Walvis Ridge a* data set. The color background for each pair of parameters is proportional to the correlation coefficient (blue for −1, white for zero, and red for +1). Posterior correlations are near zero, with the exception of a strong positive correlation between u and k.
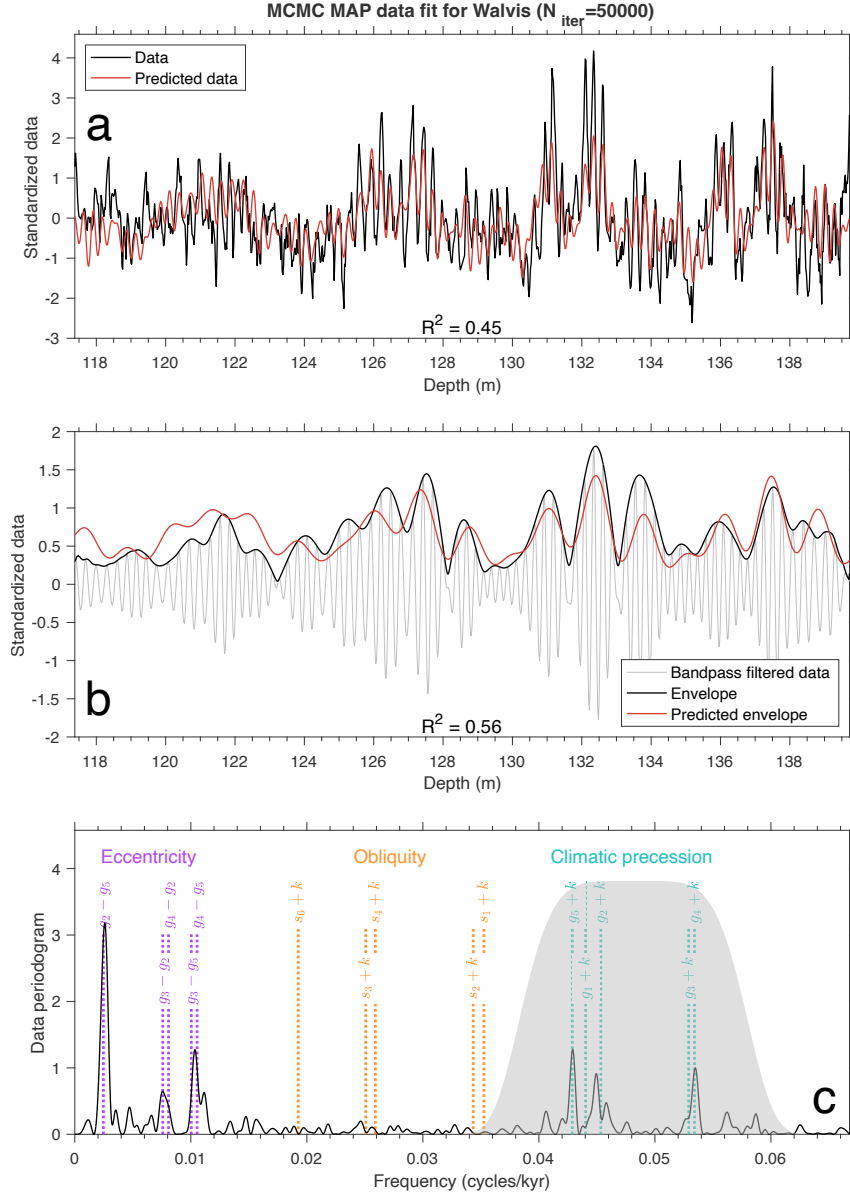
Figure S8. Fit to the Walvis Ridge a* data obtained by TimeOptBMCMC for the MAP value of sedimentation rate u and astronomical frequencies g_i, s_i, and k. (a) Fit between measured and predicted stratigraphic data (spectral fit). (b) Fit of the bandpassed climatic precession signal (envelope fit). (c) Data periodogram (black continuous line) and frequencies of the reconstructed astronomical cycles in the data (dotted vertical lines). The gray shaded area shows the frequency response of the filter used to compute the bandpassed climatic precession signal in the data (gray curve in (b)).