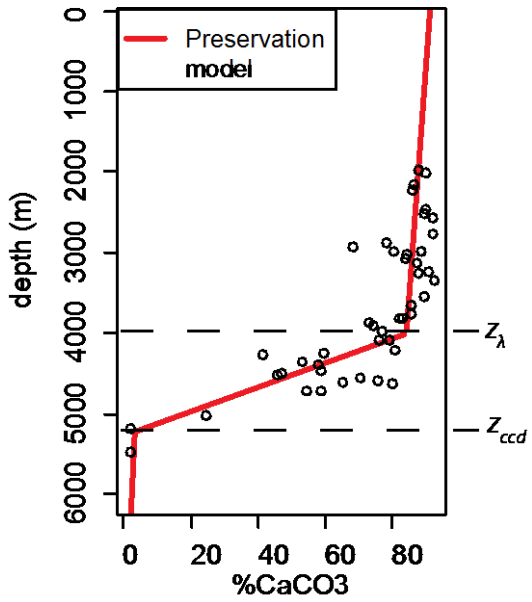
The Probable Datum Method (PDM): A technique for estimating the age of origination or extinction of nannoplankton

Jonathan D. Schueth, Klaus Keller, Timothy J. Bralower, and Mark E. Patzkowsky

Supplementary Materials

Nannofossil Preservation in the Water Column

As a basic test of the applicability of our piecewise water column preservation model, we parameterized the model (eqs. 2-4) with modern values for the Indian Ocean (Petersen and Prell 1985). Petersen and Prell (1985) also measured the % carbonate in the sediment at depth transects for the same area. By plotting both their % carbonate values and our piecewise linear model, it is apparent our model approximates the trends in the data (Supplementary Figure 1). Because the model provides a realistic representation of the true data we conclude that it is an adequate representation of pelagic carbonate preservation.



Supplementary Figure 1. The model of water-column preservation as a function of depth (eqs. 2-4) fit to modern values of carbonate preservation from the Indian Ocean (Petersen and Prell 1985). The depth of the lysocline, z_λ, and the CCD, z_{ccd} are marked.

Model-Data Residual Analysis

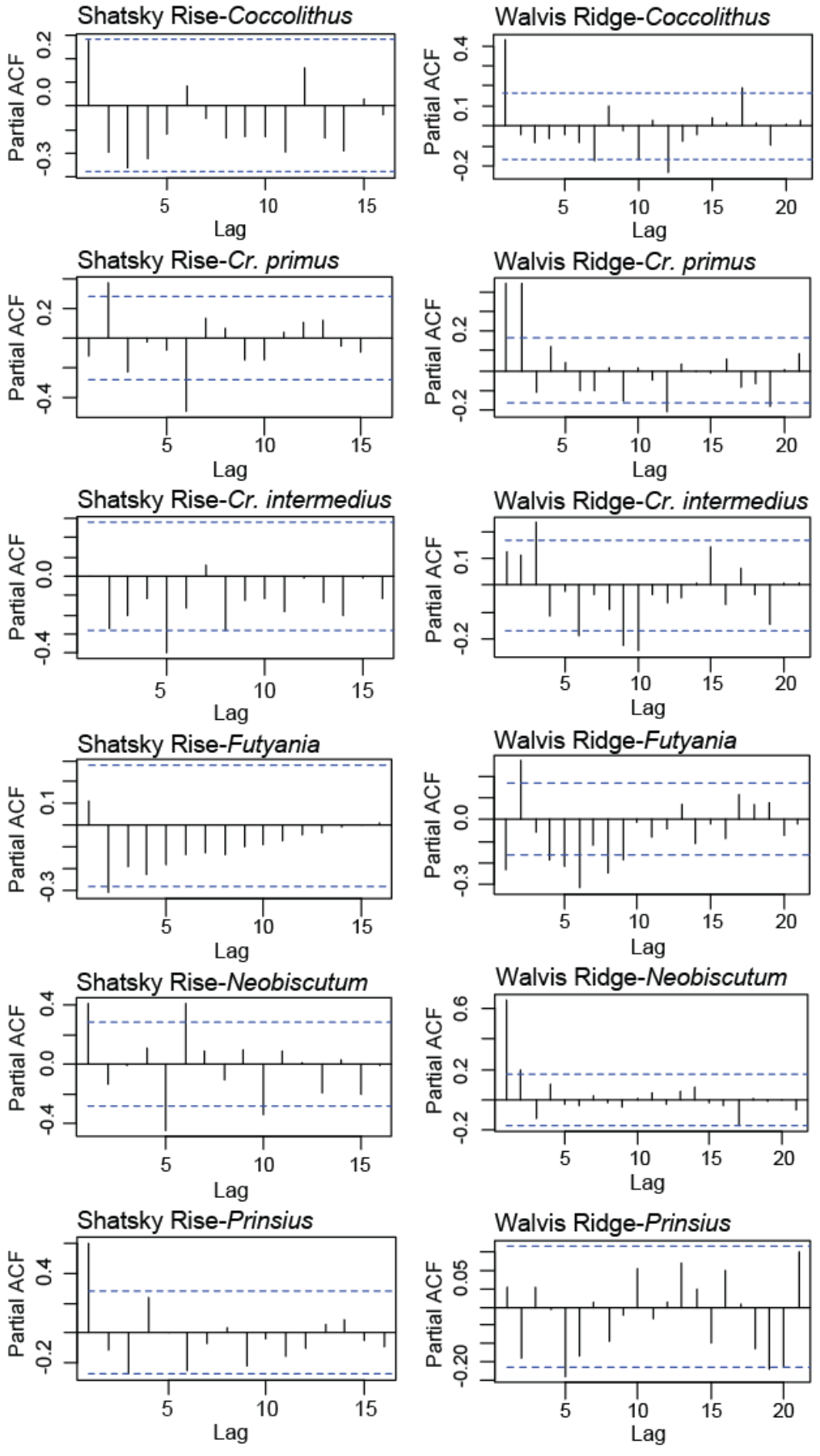
We fit our logistic model to abundance data by minimizing the sum of squares between the model and observations. Fitting a model with sum of squares relies on the assumption that the residuals are uncorrelated and that the variance of the data is homoscedastic. We observed autocorrelation in the residuals in some model fits, shown with a partial autocorrelation function of the residuals constructed with the with the "pacf" command in R (Supplementary Figure 2). In some cases our abundance data also exhibits heteroscedasticity in which the variance in the data increases as the magnitude of the data increases (e.g., Hageman 1992). The heteroscedasticity becomes apparent as a linear correlation between the absolute value of the residuals (data – model; eq. 11) and the count abundance (Supplementary Figure 3). We correct for the heteroscedasticity as described in the main text.

Supplementary Figure 2 (page S4). Partial autocorrelation functions (pacfs) of the residuals (data – model fit) for the K/Pg data at both Shatsky Rise and Walvis Ridge. The blue dashed lines represent the levels above which correlations are statistically significant. Some residual timeseries exhibit strong lag-1 autocorrelations (AR1).

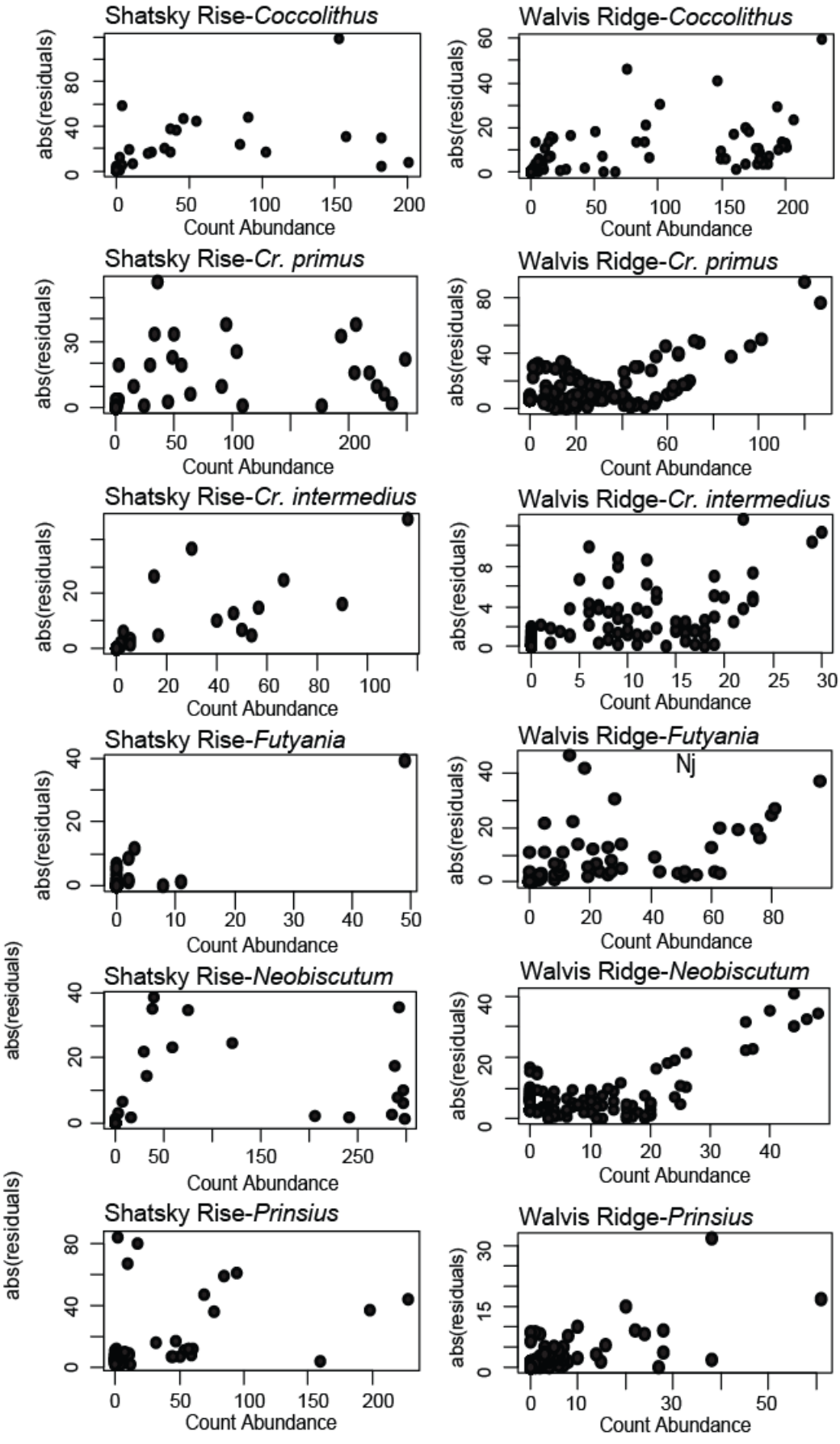
Supplementary Figure 3 (page S5). Plots of count abundance vs. the absolute value of the residuals (data - model fit) for the K/Pg data at both Shatsky Rise and Walvis Ridge. The original data is from Bown (2005), Bernaola and Monechi (2007), and Jiang et al. (2010). Strong linear relationships between the count abundance and the absolute value of the residuals suggest strong heteroscedasticity in the data. Some data exhibit this more clearly than others.

46 Supplementary Figure 2

47



Supplementary Figure 3



Justification for the Frequentist Approach of the PDM

The PDM utilizes a frequentist approach to determine the likely age of true origination or extinction. We acknowledge that a Bayesian method would be more consistent and likely more powerful, For example, our frequentist approach does not well account for prior information about the parameters (e.g., Box and Tiao 1992: pg 72**;** Jaynes 2003: pg. 550). However, Bayesian methods are typically computationally and conceptually more complex compared to a frequentist approach (e.g., D'Agostini 2003: pg. 169). Therefore, while a Bayesian approach would provide more robust and provide more statistically sound results, our sampling method more easily obtains an estimate of the probability distribution for an unknown fixed parameter, such as the datum age t_o. Frequentist approaches can provide adequate approximations in problems with few parameters, in large datasets, or if probability distributions for parameters are normal or close to normal (e.g., Box and Tiao 1992: pg. 80-86; Jaynes: 2003 pg. 550; D'Agostini 2003: pg. 39). We are careful, however, to describe our parameter histograms as approximating the probability distribution *for* a fixed, unknown parameter, rather than probability distribution *of* the parameter, as was suggested by Jaynes (2003: pg. 108).

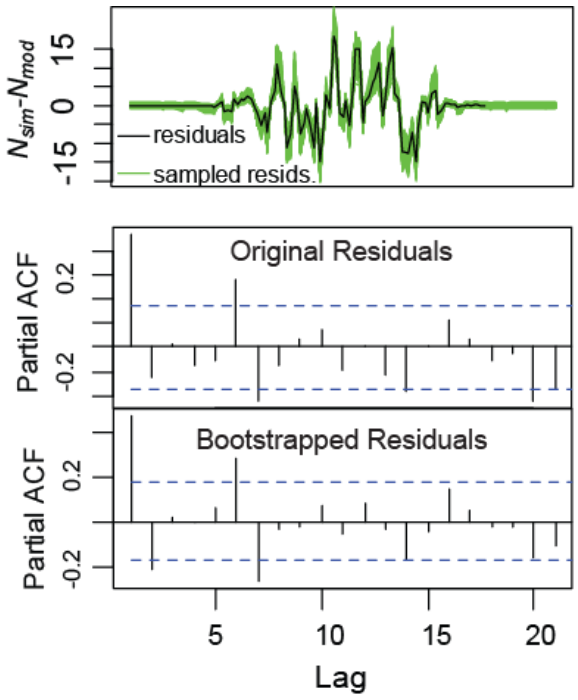
Positive Control Appendix

Our positive control experiment is described in the main text. We set all parameters to approximate a real nannofossil dataset (Supplementary Table 1) then generated a simulated timeseries from these parameters by adding autocorrelated residual noise to the model (eq. 11). We first tested whether the bootstrap sampling of model residuals retained the autocorrelated structure. Residuals were bootstrap sampled with the Maximum Entropy bootstrap with the meboot package in R (Vinod and López-de-Lacalle 2009), and we determined the partial

autocorrelation function (pacf) in R for both the original and resampled residuals. The resampled residuals have a similar pacf, showing the structure is retained in the bootstrap (Supplementary Figure 4).

Supplementary Table 1: Logistic model (eq. 11) parameters chosen for the positive control experiment. These values are set and assumed true, but should approximate real nannofossil count abundance data.

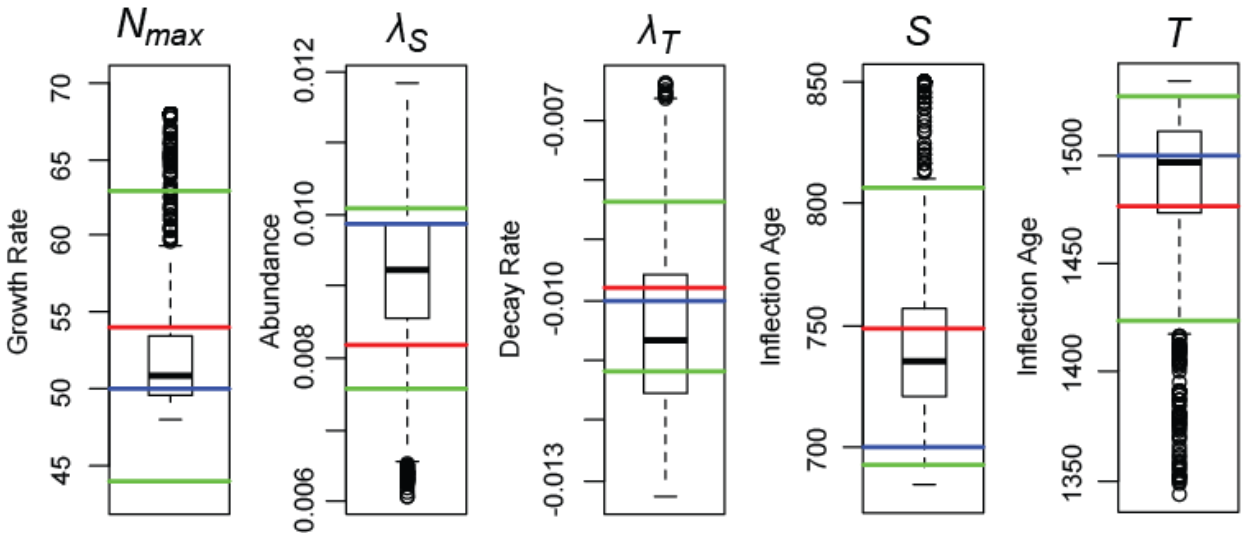
Parameter	**Given Value**
t_o	0 Kyr
N_{max}	50
λ_S	0.01
λ_T	-0.01
S	700 Kyr
T	1500 Kyr



Supplementary Figure 4. Residual bootstrapping positive control. Here we show the result of 1000 bootstrapped samples of the residuals done with the meboot package in R (Vinod and López-de-Lacalle 2009). The pacf of the original residuals is compared to the pacf of the bootstrapped residuals to show the autocorrelated structure is retained in the bootstrap.

The logistic model was fit with surrogate data constructed with the bootstrapped residuals as described in the main text. We also fit the model with a generalized nonlinear least squares (gnls) algorithm provided in the "nmle" package in R (Pinheiro et al. 2013). The use of the gnls adds another source of positive control; if both methods provide the same or similar results, we can be more confident that our method is valid. The gnls estimated both parameter values and 95% confidence limits. We compare these to box plots of the model parameters determined with our bootstrap routine for the five-parameter growth and decay model (Supplementary Figure 5).

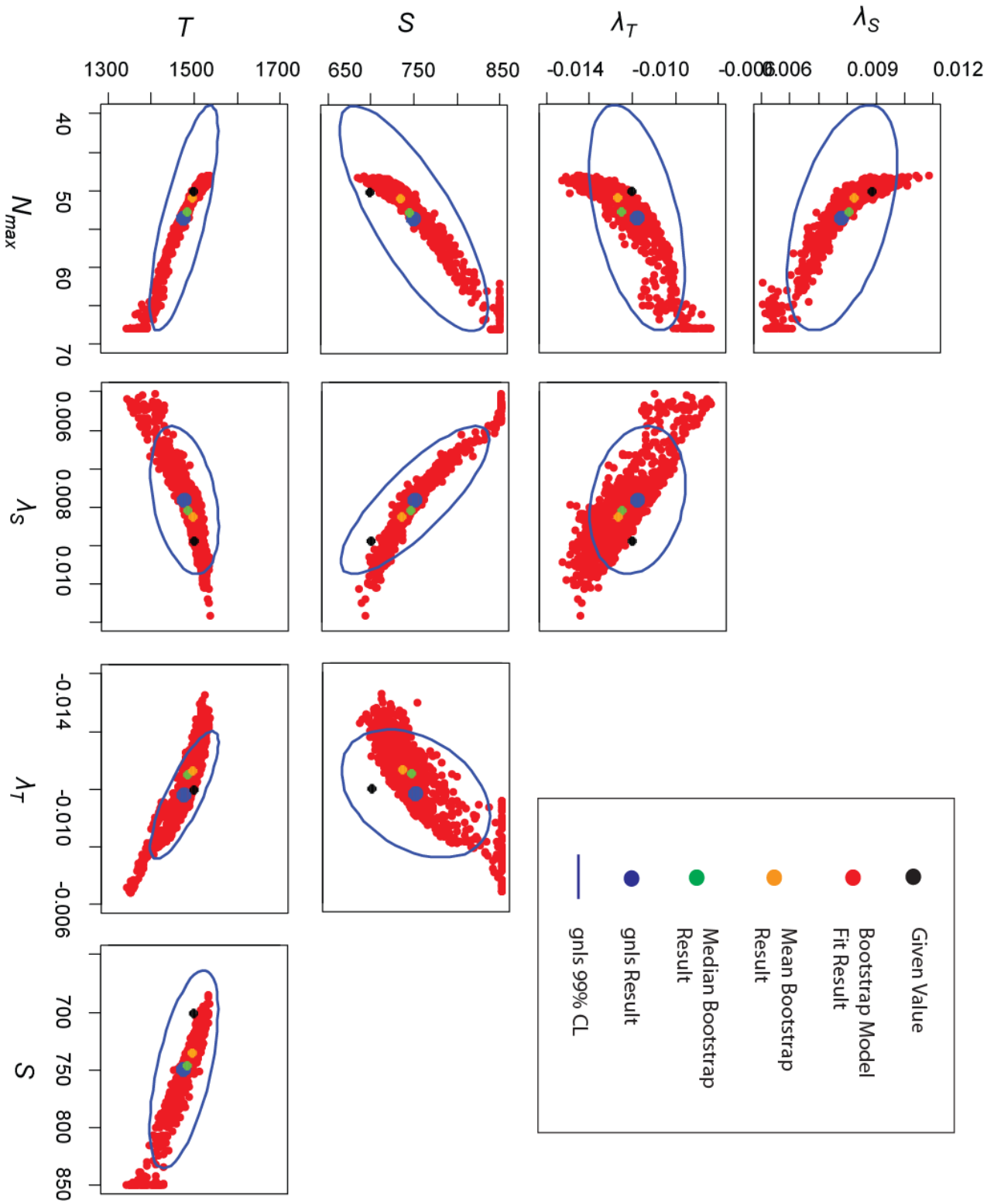
Both the gnls and our bootstrap method provide similar results for the model parameters, and both adequately approximate the given values (Supplementary Table 1).



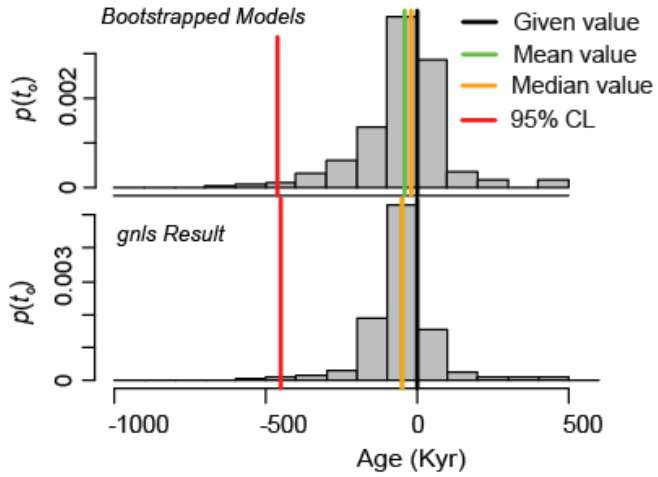
Supplementary Figure5. Boxplots of parameter estimation of the five-parameter model (eq. 11) fit with the bootstrapped residual routine. Overlaid on the boxplots are: blue line—given value (Supplementary Table 1), red line—gnls result, green lines—gnls 95% confidence limits. Note that the bootstrapped residual technique provides similar results to the gnls and both approximate the given value.

We also show the correlation of parameters estimated from the bootstrapped model fits (Supplementary Figure 6). These scatterplots are superimposed with parameter 99% confidence limits generated from the gnls algorithm using the "car" package in R (Fox and Weisberg 2011). Once again both methods provide similar results, although in some cases the bootstrap technique results in a wider range of possible parameter values. This could be a result of the model being fit to bootstrapped surrogate data rather than the original data.

Supplementary Figure 6 (page S11). Correlation of parameters in the five-parameter model (Eq. 11). Each red point represents one bootstrapped model fit out of the total 1000 bootstrapped samples. The green dot shows the mean value of the bootstrapped model fits, and the orange dot shows the median result. The blue dot represents the gnls result with the blue ellipse representing the gnls 99% confidence limit for the parameter. The given value (Supplementary Table 1) is shown with the black dot.


T
S
λ_T
λ_S
1300
1500
1700
650
750
850
-0.014
-0.010
-0.006
0.006
0.009
0.012
N_{max}
40
50
60
70
λ_S
0.006
0.008
0.010
λ_T
-0.014
-0.010
-0.006
S
700
750
800
850
Given Value
Bootstrap Model Fit Result
Mean Bootstrap Result
Median Bootstrap Result
gnls Result
gnls 99% CL

Finally, we compare the end result of the PDM between models fit with the bootstrap technique and the gnls by presenting the probability distributions for the age of origination, t_o (Supplementary Figure 7). The histograms are very similar between results generated with bootstrapped technique and the gnls routine, and both methods provide histograms with median values near to the given value of $t_o = 0$. The PDM accurately returns both the model parameters and the true age of origination or extinction for the positive control case. This strengthens our confidence in the method applied to real data.

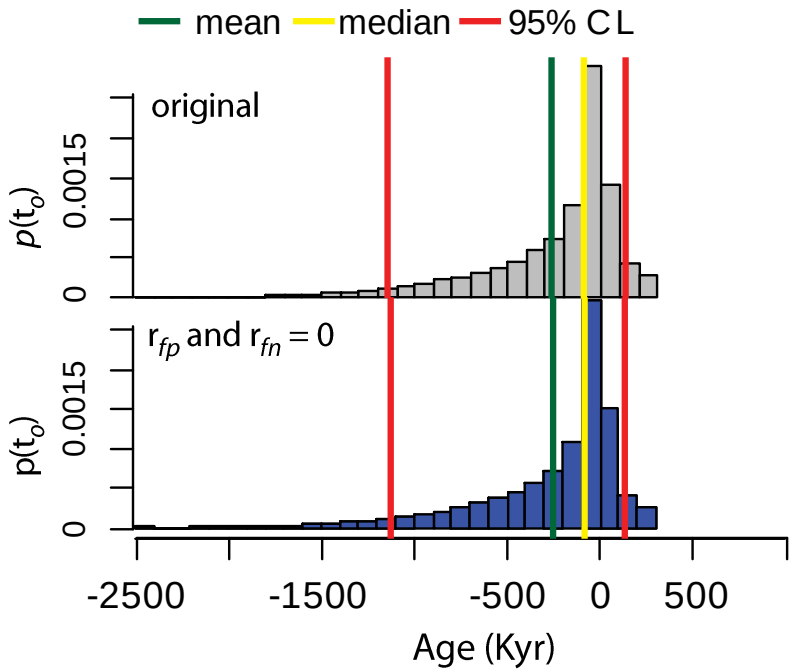


Supplementary Figure 7. The histograms of the calculated true age of origination, t_o, in the positive control experiment for both the bootstrapped routine and the gnls model fit. In all cases the results are returning results close to the given value of zero. Both the bootstrap technique and gnls provide similar results. This result shows the PDM adequately returns given values in a positive control experiment.

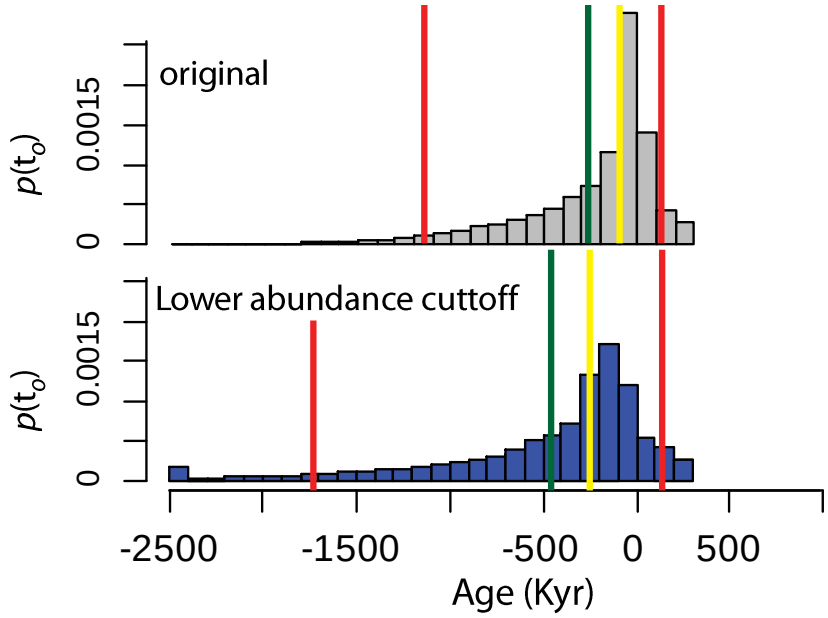
PDM Sensitivity to Parameter Change

We assessed the sensitivity of PDM results to changes in several key parameters. This was to both investigate the robustness of the method and to determine if some parameters that would be labor intensive to assess, such as the false positive/negative rates, were absolutely

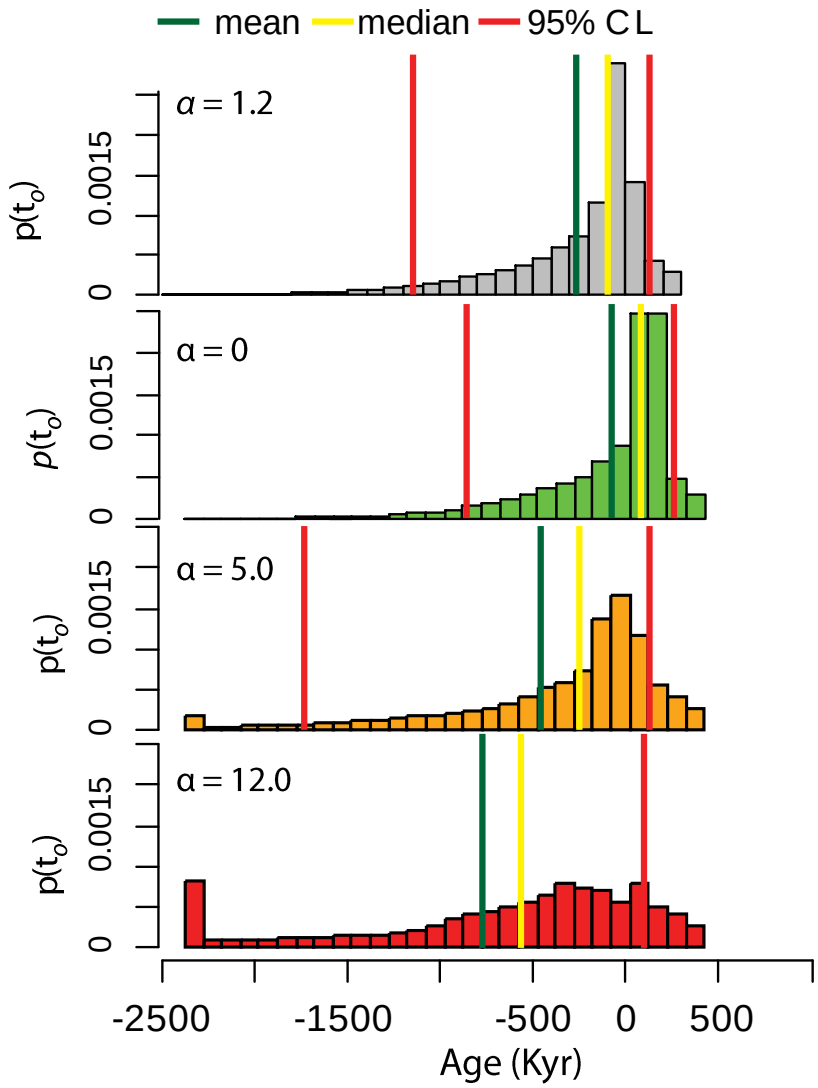
necessary to obtain meaningful results. Each test was run on the origination of *Coccolithus* at Shatsky Rise (ODP Site 1210). All results are shown and explained in Supplemental Figures 8-10.



Supplemental Figure 8. Histograms of true datum age showing the sensitivity of the PDM to changing the rates of false positive and false negative errors. Note that removing completely changes the results only slightly. It is therefore possible to safely assume these rates are equal to zero and still obtain meaningful results from the PDM.



Supplemental Figure 9. Histograms of true datum age showing the sensitivity of the PDM the sensitivity of the PDM to changing the "cutoff" vector from $< 2 \times 10^{-6}, 4 \times 10^{-6}>$ to a lower value, $< 2 \times 10^{-5}, 4 \times 10^{-5}>$. This makes the result more uncertain and provides a larger 95% confidences interval. We explain why the original vector was chosen in the main text.

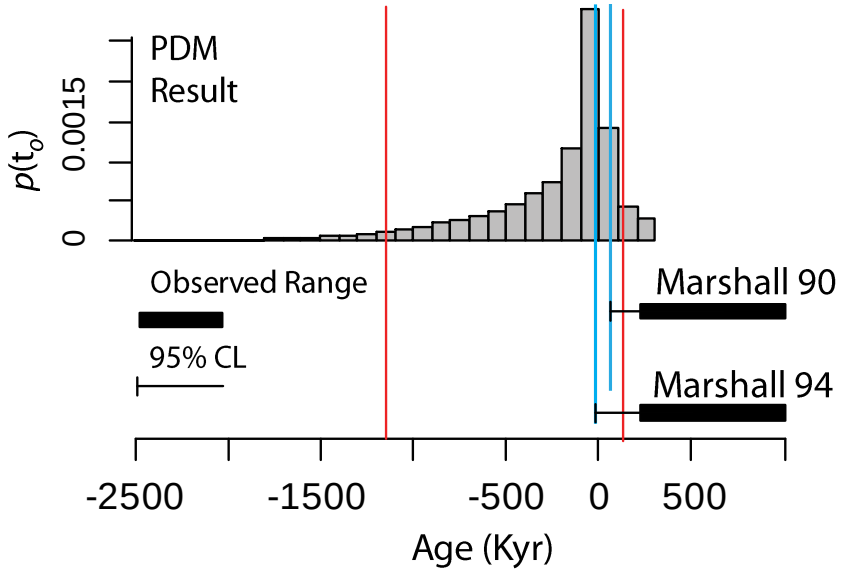


Supplemental Figure 10. Histograms of true datum age showing the sensitivity of the PDM the sensitivity of the PDM to changing the value of the taxon-specific dissolution factor α. Changes in this value do cause major changes in results. If α gets larger than 5, the results become very underconfident. We therefore recommend keeping the maximum value of α below 5.

PDM Comparison to SAS Method

We compared the PDM results generated from *Coccolithus* at ODP Site 1210 (see main text) to confidence limits generated by the SAS method. We ran the SAS both as described by

171 Marshall (1990) and Marshall (1994). The result is shown in Supplemental Figure 11. The
172 PDM generates much more conservative results and provides a larger range of uncertainty in the
173 position of the true datum age. We therefore feel this indicates the PDM provides a more
174 statistically sound result, at least for our data, than the SAS method.



175

176 Supplemental Figure 11. The PDM result, a histogram of determined values of the true
177 origination age of *Coccolithus* at Shatsky Rise compared to the confidence intervals produced by
178 the SAS method for both the technique described by Marshall (1990) and Marshall (1994). Note
179 that the confidence intervals actually more closely match the median results of the PDM, and the
180 PDM result is much less confident. This is likely due to our very small sample spacing and
181 relatively large dataset.

182

183 K/Pg Example Parameterization

184 We explain some of the model parameterization in the main text. Other parameters used
185 for the K/Pg dataset were taken from the literature (see main text). Here we list all of the
186 parameter values or ranges that were used (Supplementary Table 2). We also list the values for

187 α, the taxon-specific dissolution factor (Supplementary Table 3) derived from Thierstein (1980)
188 rescaled from 0 to a maximum of 2.6. We used the values of the experimental determination of
189 the false positive and negative error rates in this example.

190

191 Supplementary Table 2. Values of parameters used in the K/Pg example shown for Shatsky Rise
192 and Walvis Ridge. Data references are: (1) ODP Shipboard Scientific Party (2002); (2) ODP
193 Shipboard Scientific Party (2004); (3) Thierstein (1979).

Parameter	**Shatsky Rise value/range**	**Walvis Ridge value/range**
Paleodepth (km)	1.5-2.0[1]	2.0-2.3[2]
%CaCO3-lysocline	0.79[3]	0.80[3]
%CaCO3-CCD	0.02[3]	0.05[3]
Depth of lysocline (km)	1.9[3]	3.6[3]
Depth of CCD (km)	3.8[3]	4.2[3]

194

195 Supplementary Table 3. Values for the taxon-specific dissolution factor α. These values are
196 rescaled from original values given by Thierstein (1980) so that maximum susceptibility has a
197 value of $\alpha = 2.6$ and the minimum is $\alpha = 0$.

Taxon	**α value**
Coccolithus	0.4
Cr. intermedius	0.4
Cr. primus	2.6
Futyania	0.6
Neobiscutum	0.6
Prinsius	0.6

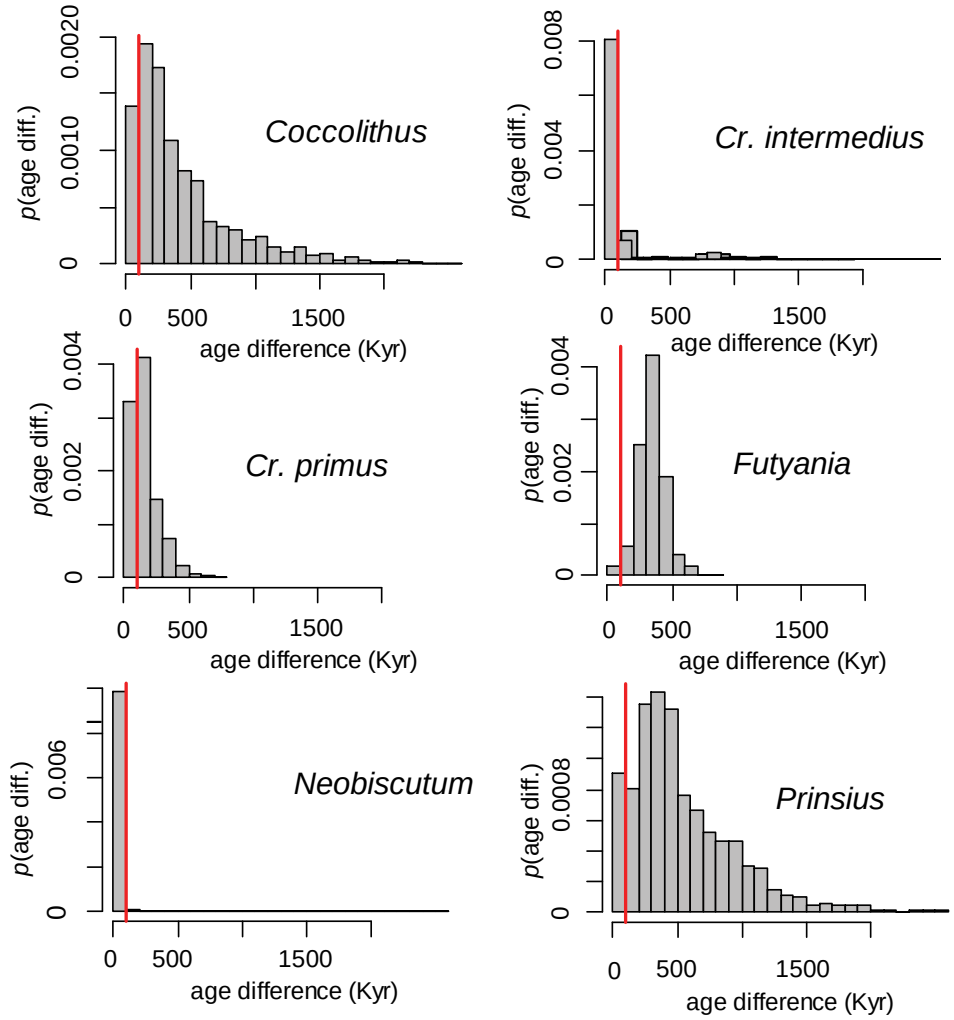
198

199

200 Supplement for the *p*(*synchronous*) Test between Shatsky Rise and Walvis Ridge.

201 The synchrony test results are given in Table 2 in the main text. These values were
202 determined by sampling the probability distributions for t_o for Walvis Ridge and Shatsky Rise

1000 times in a Monte Carlo approach. Differences were then taken between these samples, and the proportion of the sampled differences that are less than our minimum allowable synchronous age (100 Kyr, see main text). If a certain value of *p*(*synchronous*), say 5%, is reached then we could safely assess the datums as being truly diachronous between locations. The histograms of sampled differences for all six taxa studied in our K/Pg application are shown in Supplemental Figure 12.



Supplementary Figure 12. The sampled difference histograms that approximate the probability distribution for the true difference in the age of origination of the six taxa used in the K/Pg example between Shatsky Rise and Walvis Ridge (see Main Text). The red line represents the

213 age difference of 100 Kyr, the maximum allowable age difference for the origination to be
214 considered synchronous.

Literature Cited

Bernaola, G., and S. Monechi. 2007. Calcareous nannofossil extinction and survivorship across the Cretaceous-Paleogene boundary at Walvis Ridge (ODP Hole 1262C, South Atlantic Ocean). Palaeogeography, Palaeoclimatology, Palaeoecology 255:132-156.

Bown, P. 2005. Selective calcareous nannoplankton survivorship at the Cretaceous-Tertiary boundary. Geology 33:653-656.

Box, G. E. P., and G. C. Tiao. 1992. Bayesian Inference in Statistical Analysis. John Wiley and Sons, New York.

D'Agostini, G. 2003. Bayesian Reasoning in Data Analysis: A Critical Introduction. World Scientific, Singapore.

Fox, J., and S. Weisberg. 2011. An R Companion to Applied Regression, Second Edition. Thousand Oaks, CCA: Sage. URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Hageman, S. J. 1992. Alternative methods for dealing with nonnormality and heteroscedasticity in paleontological data. Journal of Paleontology 66:857-867.

Jaynes, E. T. 2003. Probability Theory: The Logic of Science. Cambridge University Press, Cambridge.

Jiang, S., T. J. Bralower, M. E. Patzkowsky, L. R. Kump, and J. D. Schueth. 2010. Geographic controls on nannoplankton extinction across the Cretaceous/Paleogene boundary. Nature Geosciences 3:280-285.

Marshall, C. R. 1990. Confidence intervals on stratigraphic ranges. Paleobiology 16:1-10.

———. 1994. Confidence intervals on stratigraphic ranges: Partial relaxation of the assumption of randomly distributed fossil horizons. Paleobiology 20:459-469.

Petersen, L. C., and W. L. Prell. 1985. Carbonate dissolution in recent sediments of the eastern equatorial Indian Ocean: Preservation patterns and carbonate loss above the lysocline. Marine Geology 64:259-290.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, and the R Development Core Team. 2013. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-111.

Shipboard Scientific Party. 2002. Site 1210. *In* T. J. Bralower, I. Premoli Silva, M. J. Malone, et al. Proceedings of the Ocean Drilling Program, Initial Reports 198:1-89.

Shipboard Scientific Party. 2004. Site 1262. *In* J. C. Zachos, D. Kroon,P. Blum, et al. Proceedings of the Ocean Drilling Program, Initial Reports 208:1-92.

Thierstein, H. R. 1979. Paleoceanographic implications of organic carbon and carbonate distribution in Mesozoic deepsea sediments. Pp. 249-274 i*n* M. Talwani, W. Hay, and W. B. F. Ryan, eds. Deep Drilling Results in the Atlantic Ocean: Continental Margins and Paleoenvironment, American Geophysical Union, Washington, D.C.

———. 1980. Selective dissolution of late Cretaceous and earliest Tertiary calcareous nannofossils: Experimental evidence. Cretaceous Research :165-176.

Vinod, H. D., and J. López-de-Lacalle. 2009. Maximum Entropy Bootstrap for Time Series: The meboot R Package. Journal of Statistical Software 29:1-19.