Paleobiology

cambridge.org/pab

The Fractional MacroEvolution Model: a simple quantitative scaling macroevolution model

Shaun Lovejoy[1] and Andrej Spiridonov[2]

[1]Physics Department, McGill University, Montreal, Quebec H3A 2T8, Canada
[2]Department of Geology and Mineralogy, Faculty of Chemistry and Geosciences, Vilnius University, Vilnius 03101, Lithuania

Article

Cite this article: Lovejoy S, Spiridonov A (2024). The Fractional MacroEvolution Model: a simple quantitative scaling macroevolution model. *Paleobiology* **50**, 376–400. https://doi.org/10.1017/pab.2023.38

Received: 6 February 2023
Revised: 19 December 2023
Accepted: 29 December 2023

Corresponding author:
Shaun Lovejoy;
Email: lovejoy@physics.mcgill.ca

Non-technical Summary

Is it environment or life that drives macroevolution? A recent analysis of the massive paleobiology database argues that the answer depends on the timescale. At short timescales, less than 40 million years, it is the environment, at longer timescales, life can effectively adapt. Both the environment and life are scaling—they fluctuate over the full range of scales from millions to hundreds of millions of years (the megaclimate regime). In this paper, we present a simple model of this scaling "crossover" phenomenon. The model has some unusual features: it is fully random and is based on fractional (rather than classical integer-ordered) differential equations.

The model is driven by temperature (a proxy for the environment) and the turnover rate (a proxy for life); it has two exponents, a cross-over time and two correlations, yet it is able to reproduce not only the statistics of the temperature, diversity, extinction, origination, and turnover rates, but it also effectively reproduces the pairwise correlations between them, and this over the whole range of timescales. If forced deterministically, it gives the response to bolide impact or other sharp forcing events.

Abstract

Scaling fluctuation analyses of marine animal diversity and extinction and origination rates based on the Paleobiology Database occurrence data have opened new perspectives on macroevolution, supporting the hypothesis that the environment (climate proxies) and life (extinction and origination rates) are scaling over the "megaclimate" biogeological regime (from ≈1 Myr to at least 400 Myr). In the emerging picture, biodiversity is a scaling "crossover" phenomenon being dominated by the environment at short timescales and by life at long timescales with a crossover at ≈40 Myr. These findings provide the empirical basis for constructing the Fractional MacroEvolution Model (FMEM), a simple stochastic model combining destabilizing and stabilizing tendencies in macroevolutionary dynamics, driven by two scaling processes: temperature and turnover rates.

Macroevolution models are typically deterministic (albeit sometimes perturbed by random noises) and are based on integer-ordered differential equations. In contrast, the FMEM is stochastic and based on fractional-ordered equations. Stochastic models are natural for systems with large numbers of degrees of freedom, and fractional equations naturally give rise to scaling processes.

The basic FMEM drivers are fractional Brownian motions (temperature, T) and fractional Gaussian noises (turnover rates, E_+) and the responses (solutions), are fractionally integrated fractional relaxation noises (diversity [D], extinction [E], origination [O], and $E_- = O - E$). We discuss the impulse response (itself representing the model response to a bolide impact) and derive the model's full statistical properties. By numerically solving the model, we verified the mathematical analysis and compared both uniformly and irregularly sampled model outputs with paleobiology series.

© The Author(s), 2024. Published by Cambridge University Press on behalf of Paleontological Society. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Introduction

Several centuries of paleontological research revealed that the evolution of life on Earth is characterized by high temporal complexity characterized by periods of sluggish and predictable evolution (Jablonski 1986; Casey et al. 2021) with mass extinctions characterized by selectivity that is low or different in kind than in "background intervals" (Raup 1992a, 1994; Payne and Finnegan 2007). There are also mass evolutionary radiations that are sometimes contemporaneous with mass extinctions (Cuthill et al. 2020). Moreover, the factors and modes of macroevolution apparently vary with time—for example, the Cambrian explosion or Ediacaran–Cambrian radiation and post-Cambrian evolution (Gould 1990; Erwin 2011; Mitchell et al. 2019); environment (Jablonski et al. 2006; Miller and Foote 2009; Kiessling et al. 2010; Boyle et al. 2013; Spiridonov et al. 2015; Tomašových et al. 2015); and timescales

(Van Dam et al. 2006; Spiridonov et al. 2017b; Crampton et al. 2018; Beaufort et al. 2022). Furthermore, macroevolution is strongly influenced by Earth's systems: geological, climatic, and paleoceanographic factors (Marshall et al. 1982; Lieberman and Eldredge 1996; Lieberman 2003; Saupe et al. 2019; Carrillo et al. 2020; Halliday et al. 2020), but also by biotic interactions, which can translate into patterns that are apparent on extremely long timescales of tens to hundreds of millions of years (Vermeij 1977, 2019; Jablonski 2008; Erwin 2012). There are also questions on the role of general stochasticity and path dependence/memory in evolutionary dynamics (Schopf 1979; Hoffman 1987; Gould 2001, 2002; Cornette and Lieberman 2004; Erwin 2011, 2016). The question is: can we reconcile in a single simple model this multitude of hierarchically organized and causally heterogenous processes producing macroevolutionary dynamics, while maintaining simplicity and conceptual clarity? Here, we argue that we can.

The development of large, high temporal resolution databases—both of past climate indicators (Veizer et al. 1999; O'Brian et al 2017, Song et al. 2019; Grossman and Joachimski 2022) and of paleobiological information, such as the Paleobiology Database (PBDB; Alroy et al. 2001, 2008) or NOW (Jernvall and Fortelius 2002; Žliobaitė et al. 2017; Žliobaitė 2022)—is transforming our understanding of macroevolution. Time series are frequently long enough that they can be studied systematically, not just as chronologies to be compared with other chronologies, but as functions of temporal *scale*, that is, the behavior of their fluctuations as functions of duration (or equivalently, their behavior as functions of frequency).

Before attempting to understand processes at specific timescales, it is important to understand their context, that is, the dynamical regime in which they operate. Dynamical regimes are objectively defined by scaling; they are regimes over which fluctuations are *scaling* (see the review by Lovejoy [2023]). By definition, a scaling regime is one in which fluctuations ΔT (in some quantity such as temperature) are of the form $\Delta T(\Delta t) \propto \Delta t^H$, where Δt is duration, or "lag," scale, and H is an exponent. If such power law relationships hold (typically for various statistics), then long- and short-duration events/fluctuations only differ quantitatively; they are qualitatively the same. This is because over such a regime, long-duration fluctuations, $\Delta T(\lambda\Delta t)$, at scale, $\lambda\ \Delta t$ ($\lambda > 1$), are related to the shorter-duration fluctuations, $\Delta T(\Delta t)$, by: $\Delta T(\lambda\Delta t) = \lambda^H \Delta T(\Delta t)$, that is, the fluctuations at different timescales differ only in their amplitudes, λ^H (with the equality understood in an appropriate stochastic sense). In addition, we can already distinguish the qualitatively different types of regime by the sign of the exponent H. $H > 0$ implies that fluctuations increase with scale and can *appear* nonstationary, whereas $H < 0$ implies that they decrease with scale, they *appear* to converge.

An important consequence for our understanding of deep-time biogeodynamics—here understood as the joint Earth–life systems—is the robustness of the "megaclimate" regime. The megaclimate regime was first discovered in benthic paleotemperatures (Lovejoy 2015), and it was characterized by "positive scaling" (a shorthand for $H > 0$) on the basis of long paleotemperature data from ocean core stacks (Veizer et al. 2000; Zachos et al. 2001), see also Lovejoy (2013) for the shorter timescale weather, macroweather, and climate regimes (up to Milankovitch scales). This implies that the difference between temperatures typically becomes larger and larger at epochs separated by longer and longer intervals of time. Theoretically, megaclimate is the hypothesis that there is a unique (presumably highly nonlinear) biogeological dynamical regime that operates over timescales spanning the range ≈1 Myr to (at least) several hundred millions of years. This would be the consequence of a unique (albeit complex, nonlinear) underlying dynamic that is valid over this wide range of scales; presumably it involves a scaling (hence hierarchical) mechanism that operates from long to short durations. A consequence is the existence of a statistical scaling regimes (notably of paleotemperatures), empirically verified throughout the Phanerozoic. While its inner scale appears to be fairly robust at around 1 Myr, its outer scale (the longest duration over which it is valid) is not known, although it appears to be at least 300 Myr. The megaclimate regime implies that the underlying biology–climate dynamics are essentially the same over these timescales: that is, at long enough time scales the statistics are stationary (although up to the outer, longest, limiting scale of the regime, they may *appear* to be nonstationary).

The hypothesis that biology and the climate are linked and that climate is a crucial and defining variable in ecological and evolutionary turnovers (Vrba 1985, 1993; Eldredge 2003; Lieberman et al. 2007; Hannisdal and Peters 2011; Mayhew et al. 2012; Crampton et al. 2016; Spiridonov et al. 2016, 2017a, 2020a,b; Mathes et al. 2021) is hardly controversial. However, the scope and utility of the megaclimate notion would increase if it could be backed up by direct analysis of paleobiological series, particularly of extinction and origination rates. This has now been done. A recent paper (Spiridonov and Lovejoy 2022), hereafter SL, found that genus-level extinction and origination rates exhibited scaling statistics over roughly the same range as the paleotemperatures confirming that the megaclimate includes these key macroevolutionary parameters (see Fig. 9, the left-hand side of the figure in the section "The Statistics of the Simulated Series Resampled at the Data Sampling Times," for a plot of the data used in SL and modeled in this paper).

The shortest scale of SL's paleobiological time series was closer to ≈ 3 Myr (the average age resolution was 5.9 Myr), which corresponds to the durations of the shortest PBDB stages—a standard shortest time resolution for Phanerozoic-scale global biodiversity analyses (e.g., Alroy et al. 2008; Alroy 2010b), although note that some sub-age data are available at resolutions closer to 1 Myr). Systematic reviews and multiple case studies revealed that even variously defined (molecular, morphological, phylogenetic, and taxic) evolutionary rates universally exhibit negative temporal scaling ($H < 0$) behavior (Gingerich 1993, 2001, 2009; Roopnarine 2003; Harmon et al. 2021; Spiridonov and Lovejoy 2022), which suggests the universality of the temporal scaling—hence hierarchical—evolutionary dynamics. An inner megaclimate scale of ≈1 Myr was also proposed in Lovejoy (2015) and is discussed in the nonspecialist book *Weather, Macroweather and Climate* (Lovejoy 2019). The scaling, and thus by implication dominance of timescale symmetric hierarchical interactions, was also detected on multimillion-year timescales in sedimentation rates/stratigraphic architecture (Sadler 1981), sea level (Spiridonov and Lovejoy 2022), and dynamics of continental fragmentation (Spiridonov et al. 2022), which shows universality of the pattern in major Earth systems as well. Therefore, the time-scaling patterns of evolution and megaclimate overlap at the very wide range of temporal scales (from $\approx 10^6$ to $> 4 \times 10^8$ yr), which motivates the development of quantitative models that explicitly tackle and integrate these timescale symmetries.

If macroevolution and climate respect wide range scaling, then it may be possible to resolve a long-standing debate in macroevolution. In terms first posed by Van Valen (1973), we may ask: are

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

evolutionary processes dominated by external factors—especially climate, the "Court Jester" (Barnosky 2001; Benton 2009)—or by life itself—the "Red Queen" (Van Valen 1973; Finnegan et al. 2008)? SL proposed a scaling resolution of the debate in which at scales below a critical transition time τ of ≈40 Myr, the climate process is dominant, but there is a crossover beyond which life (self-regulating by means of geodispersal and competition) is dominant. SL thus quantitatively concluded that at long enough timescales, the Red Queen ultimately overcomes the Court Jester. The scaling processes of the Earth system here are playing a double role (thus the Geo-Red Queen theory)—climate fluctuations growing with timescale cause perturbations in diversity to grow in size, but at the same time, at longer and longer timescales, fluctuating climates and plate tectonics cause the mixing and competitive matching of biota, thus effecting global synchronization that results in a crossover when an unstable and wandering diversity regime changes to a longer timescale fluctuation canceling or stabilizing regime (Spiridonov and Lovejoy 2022).

Physicists use the term "crossover" as shorthand to describe analogous phenomena involving processes that are subdominant over one scale range but eventually become dominant at longer scales. However, such transitions are typically modeled by Markov processes such that the autocorrelations are exponential, implying that at the critical timescale, the transition between two regimes is fairly sharp. In SL, on the contrary, in keeping with the basic megaclimate scaling dynamics, the crossover was postulated to be a consequence of the interaction of two scaling processes, that is, the transition is a slow, power law decay of one and the slow emergence of another. An analogous scaling crossover phenomenon was found in phytoplankton, in which the competing scaling processes were phytoplankton growth (with turbulence) and a predator–prey process of zooplankton grazing (Lovejoy et al. 2001).

SL argued that both macroevolution and climate respect wide-range statistical scaling, but that nevertheless, their quantitative and qualitative differences are significant, and these differences were the key to understanding the diversity (D) statistics that appeared to be involve a crossover between two different power laws. While temperature (T) fluctuations vary with timescale, Δt as $\Delta T(\Delta t) \approx \Delta t^{H_T}$ with $H_T \approx 0.25$, the corresponding laws for extinction (E) and origination (O) have $H_E \approx H_O \approx -0.25$. When $H > 0$, fluctuations grow with scale, such that the corresponding series tend to "wander" without any tendency to return to a well-defined value, and they appear "unstable." On the contrary, when $H < 0$, successive fluctuations tend to have opposite signs, such that they increasingly cancel over longer and longer timescales, and they fluctuate around a long-term value, thus appearing "stable."

To deepen our understanding, it is necessary to build a quantitative model of the interaction of climate and life. In recognition of the strongly nonlinear nature of evolutionary dynamics, numerous deterministic chaos models such as predator−prey models (e.g., Huisman and Weissing 1999; Caraballoa et al. 2016) have been developed. Although extensions with some stochastic forcing exist, for example, in Vakulenko et al. (2018), the stochasticity is a perturbing noise on an otherwise deterministic system. In paleontology, the model of exponential (unconstrained) proportional growth of diversity was historically popular (Stanley 1979; Benton 1995) or expanded to include possible accelerations due to niche construction effects (a second-order positive feedback)—a hyperbolic model (Markov and Korotayev 2007). These simple models of expansion were contrasted with single or coupled logistic models of resource-constrained competitive macroevolutionary dynamics, sometimes also supplemented with random perturbations that account for effects of mass extinctions (Sepkoski 1984, 1996); or implicitly hierarchical, and also competition-constrained, Gompertz models (Brayard et al. 2009). However, such models assume that only a few degrees of freedom are important (typically fewer than 10), whereas the true number is likely to be astronomical. It is therefore logical to model the process in a stochastic framework (involving infinite dimensional probability spaces), where the primary dynamics are stochastic, using the scaling symmetry as a dynamical constraint. Therefore, there is growing recognition of stochastic models as essential tools for understanding macroevolutionary dynamics. Actually, some of the first models that tried to explain complexities of macroevolutionary dynamics were stochastic Markovian birth and death models (Gould et al. 1977; Raup and Valentine 1983; Raup 1985, 1992b; Nee 2006). Several recent applications of linear stochastic differential equations were used in causal inference of macroevolutionary drivers and competitive interactions between clades (Liow et al. 2015; Reitan and Liow 2017; Lidgard et al. 2021).

Beyond the realism of implicitly involving larger numbers of degrees of freedom, stochastic models have the advantage that they may be linear, even though the corresponding deterministic models may be highly nonlinear. Also, by the simple expedient of using fractional-ordered differential equations rather than the classical integer-ordered ones, stochastic models can readily handle scaling, which is rarely explicitly considered in macroevolutionary analyses. Fractional differential equations provide a natural way of defining processes that vary over a wide range of timescales. Although, we forced the model with a nonintermittent Gaussian white noise in this paper, in future, this can easily be replaced by strongly intermittent (multifractal) processes that are presumably necessary to realistically model the extremely intermittent behavior, implied, for example, by mass extinctions or thermal climatic events, which are observed in macroevolution and megaclimate respectively. Scaling processes are characterized by a slow (power law) decay of the memory, much slower than an exponential rate, such that values of a time series are nontrivially correlated, even if they are separated by long time periods. Therefore, the dynamics of the system are potentially conditioned not only on the current state of the system but also on its distant past. This is exactly the property that is exploited in constructing state-of-the-art descriptive and predictive models of long-memory phenomena at shorter timescales (weeks to years), namely macroweather forecasts, based on the Scaling LInear Macroweather Model (SLIMM; Lovejoy et al. 2015) and StocSIPS (Del Rio Amador and Lovejoy 2019, 2021; Lovejoy and Del Rio Amador 2023). The same basic model using deterministic climate forcings produces climate projections to 2100, notably with much lower uncertainty than classical general circulation model (GCM) approaches (Hébert et al. 2022; Lovejoy 2022b; Procyk et al. 2022).

The useful scaling property of fractional equations arises because they have impulse response functions (Green's functions)—and hence solutions—that are based on scaling (power laws) rather than the exponential Green's functions associated with integer-ordered differential equations. In general, fractional derivatives are simply convolutions with power laws of different orders, and convolutions with different domains of integration define different types of fractional derivatives. In the fractional equations discussed in this paper, the particularly simple Weyl fractional derivative is used; in the frequency domain, it simply corresponds

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

to a power law filter. Finally, it could be noted that although fractional derivatives were discovered several centuries ago, there has been an explosion of interest in them in the last decade or so, and many books on the topic are now available (e.g., Oldham and Spanier 1974; Miller and Ross 1993; Hilfer 2000; West et al. 2003; Petras 2011; Baleanu et al. 2012; Klafter et al. 2012; Owolabi and Atangana 2019).

In this paper, we therefore build a simple model for biodiversity (D) that can reproduce and explain SL's findings. The model is parsimonious: it has only two scaling drivers—the climate and life—and by construction it reproduces the observed scaling crossover at 40 Myr. Although the model has two basic exponents (climate and life) and two coupling coefficients between temperature and turnover and between turnover and diversity, a total of four basic parameters, it satisfactorily reproduces the fluctuation statistics of D, T, E, and O as well as the turnover ($E_+ = O + E$) and difference $E_- = O - E$ over the range of ≈3 Myr to several hundred millions of years (the longest scales available). The six variables imply 15 pairwise correlations, and the correlations that are implied by the model are not single values between pairs of variables at a unique timescale, but each correlation is a function of the timescale indicating an agreement between the model and data over a wide range of scales.

In the model, the driver of the life processes is turnover. As with many other properties of groups of biological individuals, turnover can be defined at many levels. Here we use the taxic approach to modeling and analysis, such that we consider macroevolutionary turnover at the level of genera. Similarly, as in the case of organisms from different species in populations, turnover of genera in the biota shows total perturbations to a given diversity state. Namely perturbation by subtraction (extinction) and addition (origination), which represent cases of creative destruction and the destructive creation respectively, which could work as stabilizing or destabilizing factors of the global diversity depending on the circumstances (Cuthill et al. 2020). Beyond this, it explains the 15 pairs of scale by scale fluctuation correlations over the same observed range. The data are from the SL paper; they represent stage-level time series of Phanerozoic marine animal genera O and E (second-for-third origination and extinction proportion [Alroy 2015; Kocsis et al. 2019] not standardized for the duration of stages), sample standardized using the shareholder quorum method (Alroy 2010a) genus diversity of Phanerozoic marine animals based on PBDB data (https://paleobiodb.org). The paleotemperatures (T) are also the same as in the SL paper, obtained from Song et al. (2019). The rates used in the analyses are proportions, per-lineage rates, which is a natural way of describing processes working on a per capita basis (such as evolution).

The inclusion of the maximal amount of data and the taxonomic precision are the essential trade-offs of any taxic macroevolutionary study. With sufficiently well-preserved fossils, it is often possible to make accurate genus-level identification (relatively high accuracy and robustness of identification). On the other hand, if the taxon can be identified only to the family level, as is often a case for multi-element taxa, probably very few remains were preserved, or in the case of single skeleton taxa, the preservation is not adequate. Therefore, the decision to use higher-rank taxonomic data far removed from the species level could increase the risk of including more noise to the data. We chose the genus level as a compromise, which is the closest taxonomic level to the species level, while on the other hand, covering more occurrences than the fossil record resolved to the species level.

Because rates of macroevolution originations and extinctions can be estimated in many ways, and their properties (accuracy, and precision) can vary depending on the properties of the sampling process and the evolutionary process itself, we performed a sensitivity analysis using Sepkoski's genus-level marine animal data and Foote's per-lineage rates (see Appendix 2). Analysis shows that despite the differences in datasets, their completeness, differences in standardization (Sepkoski's data are not sample standardized), and differences in extinction and origination rates used, the basic results pertaining to the scaling are nearly identical in both cases. Therefore, in this paper, we only discuss the more complete and sample-standardized paleobiological time series derived from PBDB data. Currently, the PBDB data are the best source for multi-lineage global-scale analysis of evolutionary patterns, despite the presence of possible distortions related to the uneven spatial sampling, due to objective geological heterogeneities of the fossil record and various historical and socioeconomic factors that significantly impacted the study of ancient life (Raja et al. 2022; Ye and Peters 2023). Future work toward a more even representation of the global-scale data will certainly improve the accuracy of diversity estimates; this in itself would be an interesting test of the model.

As a final comment, we should note that the basic—simplest—stochastic crossover process is the fractionally integrated fractional relaxation noise (ffRn process), whose properties were only fully elucidated very recently (Lovejoy 2022a) in the context of long-term weather forecasts (Del Rio Amador and Lovejoy 2021) and climate projections (Procyk et al. 2022). The new model has conceptual commonalities with the environmental "stress model" of M. Newman that attempted to replicate the scaling statistics of extinction intensities of marine biota (Newman 1997; Newman and Palmer 2003). The model presented here is more sophisticated, as it ties the principal macroevolutionary variables—O and E—to a principal geophysical scaling process—the megaclimate—in producing realistic multi-timescale global dynamics of marine animal biodiversity, while keeping its conceptual simplicity in transparently using a few crucial parameters of time-scaling and correlations. The model is also implicitly hierarchical as implied by its scaling relations, and this is a desirable feature of a unified evolutionary theory (Eldredge 1985, 1989; Gould 2002; Lieberman et al. 2007).

The Model

The Equations

Wide Range Scaling. The SL picture is one where the extra-biological factors ("the climate") are scaling and drive biodiversity from ≈1 Myr to ≈40 Myr, where the crossover occurs, followed by the domination of biotic regulation at the longer timescales, which is also enabled by global homogenization of biota at long timescales caused by plate tectonics and changes in climate zones (Geo-Red Queen dynamics). The endemism and peculiarities or contingencies in evolutionary innovation occur at all times and places. Because evolutionary innovations are inherently unpredictable and can change carrying capacities of ecosystems and communities (Erwin 2012), their effects at the global scale of measurement of diversity are destabilizing: they act as random shocks. Because all innovations appear in certain times and places, it takes a certain time for the dispersion of innovations and the re-equilibration of the global-scale biosphere (adaptation of other taxa) following perturbations. The principal agent of mixing of biotas is plate tectonic motion (which also separates

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

biota at shorter timescales). Therefore, the only way biota can readjust and equilibrate at the global scale in the light of multiscale local and regional evolutionary innovations is by means of geodispersal mediated by competition and co-adaptation. Endemicity exists at all times, but the duration of endemicity of concrete faunas is limited by unstable Earth dynamics that work toward homogenization (the same faunas or floras cannot be endemic forever). Larger taxa either disperse to their maximal capacity and persist (and become effectively cosmopolitan) or they go extinct.

The Basic Diversity Equation. Based on this picture, we propose the following Fractional MacroEvolution Model (FMEM). First we describe the model, then we comment on it.

The basic diversity equation is:

$$\tau^h \frac{d^h(D - s_T T)}{dt^h} + D = s_E E_+; \quad E_+ = O + E \tag{1}$$

where h is a (fractional) order of differentiation whose physical interpretation is given shortly, τ is the crossover timescale (≈40 Myr), and $E_+ = E + O$ is the turnover anomaly, that is, the deviation of the turnover rate from its long-term average (in the model, E_+ can therefore be either positive or negative). For reference, it is defined on the right, where O and E are the anomalies of the origination and extinction rates with respect to their long-term averages (the diversity D is a similarly defined anomaly with respect to its long-term average). Whereas D and E_+ are already nondimensional, the temperature anomaly T must be nondimensionalized, for example, by the standard deviation of its increments at some convenient reference scale, say 1 Myr. s_T and s_E are constants that are determined by the coupling between T and D (s_T) and E_+ and D (s_E; see eq. 17).

The Drivers. The basic drivers are the climate (T) and life (E_+), themselves driven by Gaussian white noises γ_T, γ_E:

$$\begin{aligned} \tau^{\alpha+h} \frac{d^{\alpha+h} T}{dt^{\alpha+h}} &= \gamma_T \\ \tau^{\alpha} \frac{d^{\alpha} E_+}{dt^{\alpha}} &= \gamma_E \end{aligned} \tag{2}$$

where α is the basic biology (extinction and origination rates) exponent (α ≈ 0.25 as deduced from SL's analysis), and h is the exponent difference (contrast) between the temperature and biology, from SL's analysis: $h = 0.75 - \alpha \approx 0.5$. As discussed later (eq. 7 or eq. 35), combined with the diversity equation (eq. 1), these determine D. Equations (1) and (2) specify the model dynamics; see Figure 1 for an overall schematic.

The derivatives are fractional; in this paper, we use the semi-infinite "Weyl" fractional derivatives. For the arbitrary function $W(t)$, the ζ-ordered Weyl fractional derivative is defined as:

$$\frac{d^\zeta W}{dt^\zeta} = \frac{1}{\Gamma(1-\zeta)} \frac{d}{dt} \int_{-\infty}^{t} (t-s)^{-\zeta} W(s) ds; \quad 0 < \zeta < 1 \tag{3}$$

where Γ is the usual gamma function, and s is an unimportant variable of integration. In this paper, the range of differentiation,

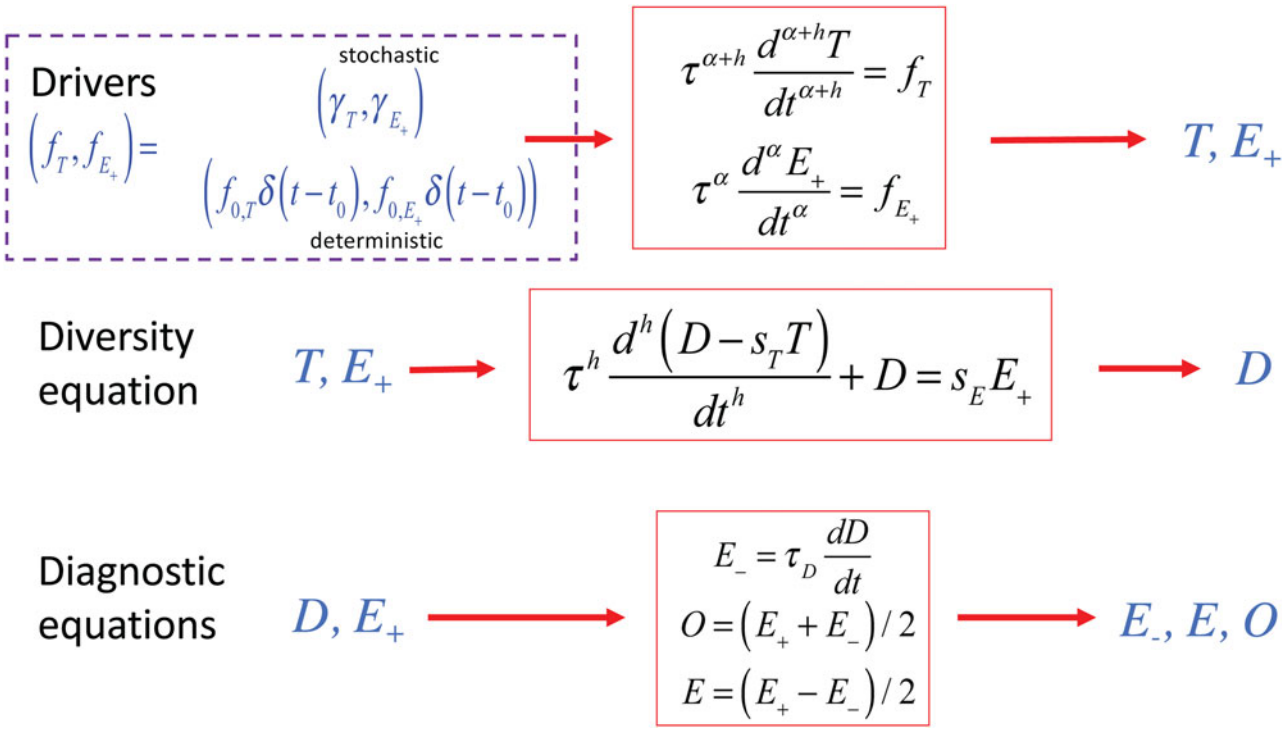


Figure 1. A schematic showing the way the various parts of the Fractional MacroEvolution Model (FMEM) fit together. The basic drivers are shown at the top; physical drivers are the temperature (T) and turnover rate (E_+). These are shown at the right, because they have nontrivial properties, such that they are best modeled as the responses to more elementary causes—the temperature and turnover rate forcings (f_T, f_{E+}). In the paper, we primarily discuss the simple case that reproduces the Paleobiology Database (PBDB) statistics, where these are Gaussian white noises ($f_T = \gamma_T$, $f_{E+} = \gamma_{E+}$), However, deterministic forcings such as bolide impacts are also discussed, shown here with both T and E_+ forced with a Dirac function of amplitude $f_{0,T}$, $f_{0,E+}$, respectively. More general forcings can be used and their responses can be obtained using the impulse response functions. The middle line shows how the T, E_+ drivers determine the diversity (D). Finally, to complete (close) the model, we need a diagnostic equation that enables us to determine E_-, E, O; this is shown in the bottom line.

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

$0 < \zeta < 1$, is adequate, but more generally, if ζ is outside the range, equation (3) can simply be combined with ordinary integer-ordered differentiation to obtain the required result. Because fractional derivatives (and their inverse, fractional integrals) are, as in equation (3)—generally convolutions, different fractional operators are defined on different ranges of integration for the convolutions. The semi-infinite Weyl derivatives are particularly easy to deal with, because they are simply power law filters in Fourier space, see equation (12) (for more information on fractional equations, see, e.g., Miller and Ross 1993; Podlubny 1999).

The γ's in equation (2) are Gaussian white noises; they are proportional to "unit" white noises, γ. Unit white noises have the properties:

$$\langle \gamma(t_1)\gamma(t_2)\rangle = \delta(t_1 - t_2); \quad \langle \gamma^2\rangle = 1; \quad \langle \gamma\rangle = 0 \tag{4}$$

where the angle brackets indicate ensemble (statistical) averaging, and δ is the Dirac function. Equation (2) therefore implies that T, E_+ are fractional integrals of white noises. Depending on the values of the exponents, these are fractional Gaussian noises (fGns) and fractional Brownian motions (fBms) (Mandelbrot and Van Ness 1968; see the later discussion on the small- and large-scale limits).

Completing the Model, the Diagnostic Equation. Equations (1) and (2) determine D, E_+, T. However, for the model to determine E and O, we need a final equation for E_-:

$$E_- = \tau_D \frac{dD}{dt}; \quad E_- = O - E \tag{5}$$

This is just the differential form of the usual discrete-time definition of diversity: $D_{j+1} = D_j(1 + O_j - E_j)$, where j is a time index. τ_D is the discretization time, the basic resolution of the series. Equation (5) plays no role in the dynamics, conventionally, it defines D (see eq. 34 for its Fourier expression). Mathematically, equation (5) is thus a "diagnostic equation," because it simply allows us to close (complete) the model by determining O, E:

$$\begin{aligned} O &= (E_+ + E_-)/2 \\ E &= (E_+ - E_-)/2 \end{aligned} \tag{6}$$

A schematic of the full model is given in Figure 1 showing its various parts, including the possibility of deterministic forcing discussed later (eq. 12).

Discussion

Diversity as an ffRn. The diversity model was written in a non-standard way (eq. 1), because in this form, its basic behavior is transparent. When $h > 0$, the fractional term is the highest-order derivative; at high frequencies it therefore dominates the zeroth-order (D) term, such that at short lags, $\Delta t < \tau$, diversity fluctuations $\Delta D \propto \Delta T$ such that D follows the temperature. However, at low frequencies ($\Delta t > \tau$), the zeroth-order term dominates, and we have instead $\Delta D \propto \Delta E_+$. By inspection, the model therefore reproduces the crossover at lag τ, and the crossover will be scaling due to the scaling of T, E_+ (eq. 2). The mathematical structure of the model is clearer if we substitute the drivers in terms of their own Gaussian forcings γ_T, γ_E (eq. 2), rewriting equation 1 as:

$$\tau^h \frac{d^h D}{dt^h} + D = \tau^{-\alpha} \frac{d^{-\alpha}}{dt^{-\alpha}}(s_T\gamma_T + s_E\gamma_E) \tag{7}$$

The term on the right-hand side represents the total forcing of the diversity. (Note: $d^{-\alpha}/dt^{-\alpha}$ is a fractional integral order α: for Weyl derivative and integrals it is the inverse of the α-order derivative d^α/dt^α).

The linear combination of white noises $s_T\gamma_T + s_E\gamma_E$ is also a white noise. The D equation is thus an order h-ordered fractional relaxation equation forced by an order α fractionally integrated white noise, that is, it is a "fractionally integrated fractional relaxation" process (ffRn; Lovejoy 2022a). The basic "unit" ffRn process $U_{h,\alpha}(t)$ satisfies:

$$\left(\frac{d^{h+\alpha}}{dt^{h+\alpha}} + \frac{d^\alpha}{dt^\alpha}\right) U_{\alpha,h} = \gamma \tag{8}$$

where γ is the unit white noise defined earlier (eq. 4), and we have used the fact that for Weyl fractional derivatives, fractional differentiation and integration commute. If time is rescaled ($t \to t/\tau$), we see (from eq. 7) that D is proportional to $U_{\alpha,h}$. We note that if $h = 1$, the D equation (eq. 1) would be a classical relaxation equation, and if forced by a white noise (i.e., if $\alpha = 0$), D would be a classical Ornstein-Uhlenbeck (OU) process. OU processes are thus the classical special cases of crossover phenomena involving high-frequency processes with exponential decorrelations (e.g., Markov processes) that lead to white noise behavior at low frequencies. They are currently the conventional approaches to the modeling and analysis of microevolutionary as well as macroevolutionary dynamics (Khabbazian et al. 2016; Bartoszek et al. 2017; Liow et al. 2022).

Deterministic Behavior: Impulse Response Functions. The D process—the solution to equation (7)—is the response of the operator $\left(\frac{d^{h+\alpha}}{dt^{h+\alpha}} + \frac{d^\alpha}{dt^\alpha}\right)$ to a white noise forcing. The general behavior of responses to linear operators is determined by their impulse response (Green's) functions $G_{\alpha,h}$ that satisfy:

$$\left(\frac{d^{h+\alpha}}{dt^{h+\alpha}} + \frac{d^\alpha}{dt^\alpha}\right) G_{\alpha,h} = \delta(t) \tag{9}$$

(Lovejoy 2022a), where $\delta(t)$ is the Dirac ("delta") function. $G_{\alpha,h}$ can be expressed in terms of "generalized exponentials" or Mittag-Leffler functions $e_{h,h+\alpha}$ as:

$$\begin{aligned} G_{\alpha,h}(t) &= t^{h-1+\alpha} e_{h,h+\alpha}(-t^h) = t^{\alpha-1} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{t^{nh}}{\Gamma(\alpha + nh)}; \quad t \geq 0 \\ & \qquad\qquad\qquad\qquad\qquad\qquad 0 < \alpha < 1/2; \quad 0 < h < 2 \\ G_{\alpha,h}(t) &= 0; \qquad\qquad\qquad\qquad\qquad\qquad t < 0 \end{aligned} \tag{10}$$

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

$$e_{a,b}(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(an+b)}$$

At small t, the leading order term is therefore $G_{\alpha,h}(t) \approx \frac{t^{\alpha-1+h}}{\Gamma(\alpha+h)}$. The large t (asymptotic) expansion is:

$$G_{\alpha,h}(t) = t^{\alpha-1} \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(\alpha - nh)} t^{-nh}; \quad t >> 1 \qquad (11)$$

(Podlubny 1999). Whereas the small t expansion is $t^{\alpha-1}$ times terms of positive powers of h, the large t expansion is in terms of $t^{\alpha-1}$ times terms in negative powers of h, with leading term $G_{\alpha,h}(t) = \frac{t^{\alpha-1}}{\Gamma(\alpha)}$. Unless $h = 0$, $G_{\alpha,h}(t)$ therefore transitions between two different power laws. The special case $h = 0$ that applies to the temperature and turnover forcings (eq. 2), corresponds to the pure power law $G_{\alpha,0}(t) = \frac{t^{\alpha-1}}{\Gamma(\alpha)}$. $G_{\alpha,h}$ has the property that if it is (fractionally) integrated ζ times, the result is just $G_{\alpha+\zeta,h}$. As explained in Appendix 1, $G_{\alpha,h}$ is useful for numerical simulations.

At a typical highest resolution of global datasets with timescales ≈ of 1 Myr, $G_{\alpha,h}$ gives the deterministic response to forcings that are effectively impulses at this scale, for example, a bolide strike (Alvarez et al. 1980; During et al. 2022), supernova or gamma ray burst (Fields et al. 2020), or even much slower hyperthermal event such as the Paleocene–Eocene thermal maximum (Gingerich 2006; McInerney and Wing 2011) or the Cenomanian–Turonian event (Eaton et al. 1997; Meyers et al. 2012; Venckutė-Aleksienė et al. 2018), extensive volcanic eruption episodes, or other short-timescale (below ≈ 1 Myr) stressors. Figure 2 shows a plot of impulse responses for temperature and turnover demonstrating their singular nature for the empirical parameters estimated in SL ($\alpha \approx 0.25$, $h \approx 0.5$). These singular responses combine apparently contradictory features: on the one hand, the falloff at short times following the impulse is very sharp, whereas on the other hand, the decay is very slow at long times, so its effects take a long time to disappear. The sharpness feature is desirable; because mass extinctions, and potentially other episodes of dramatic biotic change, effectively represent periods of almost infinite turnover rate, it is near instantaneous or singular-like (Foote 1994, 2005). Indeed, the global stratigraphic stages and substages are based on the episodes of turnover. Figure 3 shows the impulse response for the diversity that

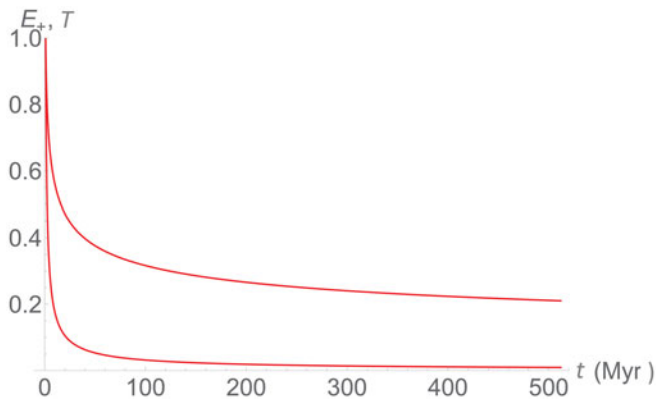


Figure 2. The impulse (delta function) response $G_{\alpha,0}(t) = t^{\alpha-1}/\Gamma(\alpha)$ for fractional integrals of order α normalized for the same response after 1 Myr. The bottom corresponds to the turnover (E_+) response $\alpha = 1/4$, and the top corresponds to the temperature (T) response with $\alpha = 3/4$. Notice the long-term effects. Due to causality, the impulse response is 0 for $t < 0$.

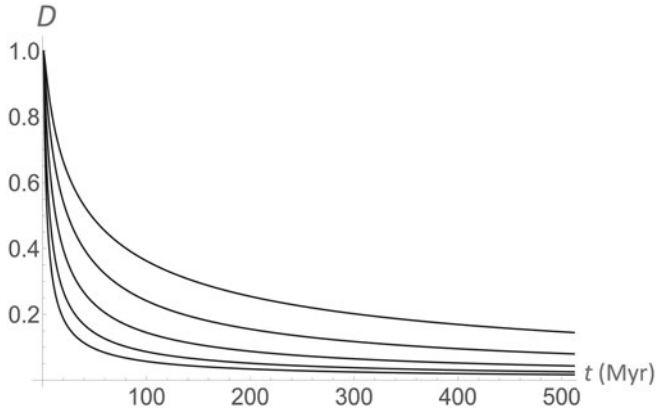


Figure 3. The impulse response $G_{\alpha,h}(t/\tau)$, with $\alpha = 1/4$, $h = 1/2$, corresponding to the diversity (D) response, for critical transition times $\tau = 1$, 4, 16, 64, and 256 Myr (bottom to top). The empirical value is $\tau \approx 40$ Myr (SL). Due to causality, the impulse response is 0 for $t < 0$.

has two different power law regimes: a high-frequency $t^{\alpha-1+h}$ regime and a low-frequency $t^{\alpha-1}$ regime (the leading terms in eqs. 10, 11) with a transition at various scales τ indicated. As expected, the former (temperature dominance) corresponds to the singularity $t^{-0.25}$ and the latter to the singularity $t^{-0.75}$ (turnover dominance; both are shown in Fig. 2). Note that because the equations are linear, the impulse responses from the deterministic forcings shown in Figures 2 and 3 will be superposed onto the stochastic white noise–driven responses that we calculate in the section entitled "Numerical Simulations."

The other aspect of the singular impulse responses is their long time decays that are much slower than conventional exponential decays and can be thought of as a kind of memory that characterizes the responses to both deterministic and stochastic forcing. Our model thus predicts that there will be long-term consequences of bolide impacts or other catastrophic events. This is in accord, for example, with the findings of Krug et al. (2009) and Krug and Jablonski (2012) that the Cretaceous–Paleogene (K-Pg) mass extinction caused by the effects of the Chicxulub asteroid impact changed long-term origination rates and their spatial distribution, a situation that persists today, 66 Myr after the event, in accord with this long-memory feature of the FMEM model. The slow decay in the response is also a desired property in modeling macroevolutionary dynamics, as it can replicate effects of phylogenetic inertia or "phylogenetic constraint" (Gould 2002) and also inertia of persistence of geological structures that can affect dynamics of biodiversity for periods, eras, or even eons. The genetic composition is basically the material trans-generational memory of biological individuals of all levels of the Linnaean hierarchy (Eldredge 1996). Therefore, extinction (or origination) of genera or taxa of higher ranks can change the functioning and the properties of the biosphere for hundreds of millions of years. The same goes for the formation of such structures as oceanic basins or continents with their myriads of possible configurations—their origins and disappearances impose new boundary conditions for geophysical and macroevolutionary dynamics for hundreds of millions of years (Nance 2022; Spiridonov et al. 2022).

We could further note that the long memory can be exploited to make future predictions. This is because for Gaussian forcing (eq. 2), E_+ and the increments of T are long-memory fGn processes and—as discussed earlier—D is an ffRn.

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

The predictability of the former has been mathematically studied in Gripenberg and Norros (1996) and has been exploited for monthly and seasonal forecasts in Del Rio Amador (2019, 2021) and Lovejoy and Del Rio Amador (2023), and the predictability properties of ffRn processes have been studied in Lovejoy (2022).

Solving the Model. Readers only interested in results can skip ahead until the basic model equations (38, 39) that are given at the end of the section "Full model statistics".

Fractional derivatives are generally convolutions (with power laws, eq. 3), hence different ranges of integration in the convolution yield different fractional derivatives. Most often (e.g., the Riemann-Liouville and Caputo fractional derivatives), the latter are taken from time = 0 to t, in which case the initial conditions are important and dealing with them is technically somewhat complex. In these cases, the main tool is the Laplace transform. Here, however, we consider statistically stationary white noise forcing that starts at time = $-\infty$. In this case, we can use the "Weyl" fractional derivative (a convolution from $-\infty$ to t, eq. 3) whose Fourier transform ("F.T.") is particularly simple:

$$\frac{d^h}{dt^h} \overset{F.T.}{\leftrightarrow} (i\omega)^h \tag{12}$$

where ω is the Fourier conjugate variable, that is, the frequency (when h is an integer, the formula may be familiar to the reader). If we Fourier transform equations (1) and (2), we can write the model in matrix form as:

$$\underline{\widetilde{S}}(\omega) = (i\omega\tau)^{-\alpha}\underline{\underline{\widetilde{F}}}(\omega)\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ s_T & s_E \end{pmatrix}\begin{pmatrix} \widetilde{\gamma}_T \\ \widetilde{\gamma}_E \end{pmatrix}$$

$$\underline{\widetilde{S}}(\omega) = \begin{pmatrix} \widetilde{T} \\ \widetilde{E_+} \\ \widetilde{D} \end{pmatrix}$$

$$\underline{\underline{F}}(\omega) = \begin{pmatrix} (i\omega\tau)^{-h} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \dfrac{1}{1+(i\omega\tau)^h} \end{pmatrix} \tag{13}$$

(the single underlining indicates a vector, the double underlining, a matrix, the Fourier transform is denoted with a tilde).

As noted earlier, the D forcing is a linear combination of white noises (eq. 7), such that the sum on the right-hand side of equations (7) and (13) is a correlated white noise. However, from the data (see fig. 5), we see that E_+ and T are themselves correlated. We therefore rewrite the model in terms of two statistically independent ($\langle\gamma_1\gamma_2\rangle = 0$) unit ($\langle\gamma_1^2\rangle = \langle\gamma_2^2\rangle = 1$) white noise drivers γ_1, γ_2:

$$\begin{pmatrix} \gamma_T \\ \gamma_E \end{pmatrix} = \begin{pmatrix} \sigma_T & 0 \\ 0 & \sigma_E \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \rho_E & \sqrt{1-\rho_E^2} \end{pmatrix}\begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \tag{14}$$

So that:

$$\sigma_T^2 = \langle\gamma_T^2\rangle; \quad \sigma_E^2 = \langle\gamma_E^2\rangle; \quad \rho_E = \frac{\langle\gamma_T\gamma_E\rangle}{\sigma_T\sigma_E}; \quad \langle\gamma_T\rangle = \langle\gamma_E\rangle = 0 \tag{15}$$

where σ_T is the standard deviation of γ_T, σ_E of γ_E, and ρ_E is the T, E_+ correlation. Equation (14) is the standard Cholesky decomposition for two correlated random variables, noises.

Fourier transforming equation (14) and using equation (13), we can write the model as:

$$\underline{\widetilde{S}}(\omega) = (i\omega\tau)^{-\alpha}\underline{\underline{F}}(\omega)\underline{\underline{\sigma}}\underline{\underline{\rho}}\underline{\widetilde{\gamma}} \tag{16}$$

$$\underline{\underline{\sigma}} = \begin{pmatrix} \sigma_T & 0 & 0 \\ 0 & \sigma_E & 0 \\ 0 & 0 & \sigma_D \end{pmatrix}; \quad \underline{\underline{\rho}} = \begin{pmatrix} 1 & 0 \\ \rho_E & \sqrt{1-\rho_E^2} \\ \rho_D & \mathrm{sgn}(r)\sqrt{1-\rho_D^2} \end{pmatrix}; \quad \underline{\widetilde{\gamma}} = \begin{pmatrix} \widetilde{\gamma}_1 \\ \widetilde{\gamma}_2 \end{pmatrix}$$

Where the parameters:

$$\sigma_D = s_T\sigma_T\sqrt{1+2\rho_E r+r^2}; \quad r = \frac{s_E\sigma_E}{s_T\sigma_T}$$
$$\rho_D = \frac{1+r\rho_E}{\sqrt{1+2r\rho_E+r^2}} \tag{17}$$

depend on both the driver statistics (σ_T, σ_E, and ρ_E) and the model parameters s_T, s_E. While σ_D does parametrize the amplitude of the diversity fluctuations, unlike σ_T, σ_E (which must be ≥ 0), it is not a true standard deviation: if $s_T < 0$, it will be negative. Similarly, we will see that ρ_D determines the D, E_+, and T correlations but is not itself a correlation coefficient and depends on the sign of the ratio r.

Stochastic Response to White Noise Forcing

Scaling Processes: fGn and fBm. We are interested in the statistical properties of the solutions $\underline{\widetilde{S}}(\omega)$. These can be expressed in terms of fGn, fBm, and ffRn processes. Before discussing the full statistics that include the cross-correlations, let us therefore discuss the statistics of each of the main variables individually.

Let us start with the scaling processes T, E_+ that are of the form:

$$\frac{d^{\alpha_X}X}{dt^{\alpha_X}} = \gamma \overset{F.T.}{\leftrightarrow} (i\omega)^{\alpha_X}\widetilde{X} = \widetilde{\gamma} \tag{18}$$

For the statistics, we can determine the power spectrum:

$$E_X(\omega) = \left\langle|\widetilde{X}|^2\right\rangle = \frac{1}{2\pi}|\omega|^{-\beta_X}; \quad \beta_X = 2\alpha_X \tag{19}$$

where β_X is the spectral exponent, and we have used the fact that the spectrum of a Gaussian white noise is flat:

$$\langle|\widetilde{\gamma}(\omega)|^2\rangle = \frac{1}{2\pi}\langle\gamma^2\rangle = \frac{1}{2\pi} \tag{20}$$

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

$E_X(\omega)$ is thus the basic form of the T, E_+ spectra (not to be confused with the extinction rate E). From the Wiener-Khintchin theorem, the (real-space) autocorrelation function $R_X(\Delta t)$ is the inverse transform:

$$\begin{aligned} R_X(\Delta t) &= \langle X(t)X(t-\Delta t)\rangle \propto \Delta t^{H_X} \overset{F.T.}{\leftrightarrow} \widetilde{R}_X(\omega) \\ &= E_X(\omega) = \left\langle \left|\widetilde{X(\omega)}\right|^2 \right\rangle \propto |\omega|^{-\beta_X}; \\ H_X &= \frac{\beta_X - 1}{2} = \frac{\alpha_X}{2} \end{aligned} \tag{21}$$

The technical difficulty is that due to a low-frequency divergence, the inverse transform of pure power spectra (eq. 19) only converges for $\beta_X < 1$ (i.e., $\alpha_X < ½$, $H_X < 0$); this is the fGn regime appropriate for E_+. Even here, $R_X(\Delta t)$ is infinite for $\Delta t = 0$. Because $R_X(0)$ is the variance, fGn processes are (like the white noise special case $\alpha_X = 0$) generalized functions that must be averaged (integrated) over finite intervals in order to represent physical processes. Averaging to yield a finite resolution process is adequate for $\beta_X > -1$ ($\alpha_X > -½$, $H_X > -1$) such that the fGn process is defined for $-1 < \beta_X < 1$ (i.e., $-½ < \alpha_X < ½$, $-1 < H_X < 0$). After X is averaged over a finite resolution τ_r, the result is $X\tau_r$ with root-mean-square (RMS) statistics $\left\langle X_{\tau_r}^2 \right\rangle^{1/2} \propto \tau_r^{H_X}$. Because $H_X < 0$, the empirical statistics will be highly sensitive to the resolution τ_r.

When $\alpha_X \geq ½$, the low-frequency divergences imply that the $X(t)$ process is nonstationary (the process generally "wanders off" to plus or minus infinity). However, for $1 < \beta_X < 3$ (i.e., $½ < \alpha_X < 3/2$, $0 < H_X < 1$; this is the range appropriate for T: $H_T \approx 0.25$, $\beta_T \approx 3/2$), its increments are (stationary) fGn processes; this regime defines the fBm process. Finally, because all physical scaling processes exist over finite ranges of scale, there will be finite outer (longest) timescale (smallest frequency) such that truncating the spectrum at low frequencies (as for the ffRn processes, see the section "Two Scaling Regimes: *fRn* and *ffRn*") leads to an overall stationary process.

When analyzing paleo-series, it is convenient to analyze the statistics in real space, the main reason being that these are easier to interpret (the difficulty in interpretation is the cause of the recently discovered quadrillion error in climate temperature spectra [Lovejoy 2015]). An additional reason is that uniformly sampled paleo-series are typically not available: indeed, the geochronologies themselves are typically scaling (see Appendix 3 and Fig. A3.1 for more discussion). For the moment, the problem of spectral estimation from data with scaling measurement densities (i.e., scaling number of measurements per time interval) is an unsolved problem.

We have already noted that the autocorrelation functions are only adequate for $H_X < 0$ ($\alpha_X < ½$, $\beta_X < 1$), this is why, when $0 < H_X < 1$, it is conventional to define fluctuations using differences $\Delta X(\Delta t) = X(t - \Delta t) - X(t)$, which are stationary over this range. Differences avoid low-frequency divergences but will still have high-frequency divergences when $H_X < 0$. To avoid problems at both the small scale (resolution dependencies) and large scales (nonstationarity), it is convenient to use Haar fluctuations. Over the interval Δt, the Haar fluctuation $\Delta X(\Delta t)$ is defined as the difference between the average of the first and second halves of the interval.

$$\left\langle \Delta X(\Delta t)^2 \right\rangle^{1/2} \propto \Delta t^{H_X} \leftrightarrow E_X(\omega) \propto \omega^{-\beta_X}; \quad \begin{matrix} -1 < H_X < 1 \\ -1 < \beta_X < 3 \\ \beta_X = 2H_X + 1 \end{matrix} \tag{22}$$

(valid for Haar fluctuations). Over the indicated range of parameters, the Haar fluctuations are stationary and are independent of the resolution.

Comparing equations (7) and (2), we find:

$$\begin{aligned} \left\langle \Delta E_+(\Delta t)^2 \right\rangle^{1/2} &\propto \Delta t^{H_E}; \quad H_E = \alpha - \frac{1}{2} \\ \left\langle \Delta T(\Delta t)^2 \right\rangle^{1/2} &\propto \Delta t^{H_T}; \quad H_T = h + \alpha - \frac{1}{2} \end{aligned} \tag{23}$$

Two Scaling Regimes: fRn and ffRn. From equations (8) and (9), the basic Fourier transforms of ffRn processes and their impulse responses are:

$$\begin{aligned} \widetilde{U}_{\alpha,h}(\omega) &= \frac{\widetilde{\gamma}}{(i\omega)^\alpha (1 + (i\omega)^h)}; \\ \widetilde{G}_{\alpha,h}(\omega) &= \frac{1}{(i\omega)^\alpha (1 + (i\omega)^h)}; \end{aligned} \quad 0 < \alpha < 1/2; \quad 0 < h < 2 \tag{24}$$

The fractional relaxation noise (fRn) process is the special case where $\alpha = 0$. The ffRn power spectrum is therefore:

$$E_{\alpha,h}(\omega) = \left\langle \left|\widetilde{U}_{\alpha,h}\right|^2 \right\rangle = \frac{1}{2\pi|\omega|^{2\alpha}|1 + (i\omega)^h|^2} \tag{25}$$

$E_{\alpha,h}(\omega)$ is thus the basic form of the D spectrum.

The full statistical properties of ffRn processes (including series expansions) are discussed in Lovejoy (2022); however, for our purposes, the low- and high-frequency scaling exponents are sufficient. For these, equation (25) yields:

$$E_{\alpha,h}(\omega) \propto |\omega|^{-\beta}; \quad \begin{matrix} \beta_l = 2\alpha; & \omega << 1 \\ \beta_h = 2(\alpha + h); & \omega >> 1 \end{matrix} \tag{26}$$

("h" for high frequency, "l" for low frequency). To obtain the basic fluctuation statistics, it is sufficient to apply equation (22) to each regime separately. Indeed, more generally, "Tauberian theorems" (e.g., Feller 1971) imply that if the spectrum is a power law over a wide enough range, then the corresponding (second-order) real-space statistics will also be scaling. Therefore:

$$\left\langle \Delta U_{\alpha,h}(\Delta t)^2 \right\rangle^{1/2} \begin{matrix} \propto \Delta t^{H_l}; & H_l = \alpha - \frac{1}{2}; & \Delta t >> 1 \\ \propto \Delta t^{H_h}; & H_h = \alpha + h - \frac{1}{2}; & \Delta t << 1 \end{matrix} \tag{27}$$

Using the empirical values $\alpha \approx 0.25$, $h \approx 0.5$, we see that E_+ is a fractional Gaussian noise and that T is an fBm process. Also, we find (cf. eqs. 7, 27) that $H_D \approx H_T$ ($\Delta t \ll \tau$) and $H_D \approx H_E$ ($\Delta t \gg \tau$).

The Full-Model Statistics: Spectra, Correlations

The Basic Model. The model is linear and has stationary Gaussian (white noise) forcing (T, E_+), therefore D, E_-, E, O are also Gaussian, such that their statistics are determined by spectra and cross-spectra, or equivalently in real space (via the Wiener-Khintchin theorem), by the autocorrelations and cross-correlations:

$$R_{ij}(\Delta t) = \left\langle S_i(t) S_j(t - \Delta t) \right\rangle \overset{F.T.}{\leftrightarrow} \widetilde{R}_{ij}(\omega) = \left\langle \widetilde{S}_i(\omega) \widetilde{S}_j^*(\omega) \right\rangle \tag{28}$$

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

(the diagonal terms are the spectra of the components: $\widetilde{R}_{ii}(\omega) = E_i(\omega)$, the asterisk indicates the complex conjugate). In matrix notation:

$$\begin{aligned}\underline{\underline{\widetilde{R}}}(\omega) &= \left\langle \underline{\widetilde{S}}\underline{\widetilde{S}}^{T*} \right\rangle = |\omega\tau|^{-2\alpha} \underline{\underline{F}}(\omega) \underline{\underline{\sigma\rho}} \left\langle \underline{\widetilde{\gamma}}\underline{\widetilde{\gamma}}^{T*} \right\rangle \underline{\underline{\rho}}^{T*} \underline{\underline{\sigma}}^{T*} \underline{\underline{F}}(\omega)^{T*} \\ &= \frac{|\omega\tau|^{-2\alpha}}{2\pi} \underline{\underline{F}}(\omega) \underline{\underline{\sigma\rho\rho}}^{T*} \underline{\underline{\sigma}}^{T*} \underline{\underline{F}}(\omega)^{T*}\end{aligned} \tag{29}$$

where the superscript T indicates the transpose, the asterisk indicates the complex conjugate, and we have used:

$$\left\langle \underline{\widetilde{\gamma}}.\underline{\widetilde{\gamma}}^{*T} \right\rangle = \left\langle \begin{pmatrix} \widetilde{\gamma}_1 \\ \widetilde{\gamma}_2 \end{pmatrix} (\widetilde{\gamma}_1 \quad \widetilde{\gamma}_2) \right\rangle = \frac{1}{2\pi} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2\pi} \underline{\underline{1}} \tag{30}$$

The key correlation matrix (from eq. 16) is:

$$\underline{\underline{\rho\rho}}^{T*} = \begin{pmatrix} 1 & \rho_{TE} & \rho_{TD} \\ \rho_{TE} & 1 & \rho_{ED} \\ \rho_{TD} & \rho_{ED} & 1 \end{pmatrix} \tag{31}$$

where

$$\rho_{TE} = \rho_E; \quad \rho_{TD} = \rho_D; \quad \rho_{ED} = \rho_E\rho_D + \mathrm{sgn}(r)\sqrt{1-\rho_E^2}\sqrt{1-\rho_D^2} \tag{32}$$

and

$$\underline{\underline{\sigma\rho\rho}}^{T*} \underline{\underline{\sigma}}^{T*} = \begin{pmatrix} \sigma_T^2 & \rho_{TE}\sigma_T\sigma_E & \rho_{TD}\sigma_D\sigma_T \\ \rho_{TE}\sigma_T\sigma_E & \sigma_E^2 & \rho_{DE}\sigma_E\sigma_D \\ \rho_{TD}\sigma_D\sigma_T & \rho_{DE}\sigma_E\sigma_D & \sigma_D^2 \end{pmatrix} \tag{33}$$

Completing the Model: The Diagnostic Equation for E_. Before writing down the final spectra, let us complete the system with the help of the diagnostic equation that allows us to determine E_- from D (and hence E, O, eq. 6).

The Fourier transform of the diagnostic equation (eq. 5) is:

$$\widetilde{E}_- = \left(\frac{\tau_D}{\tau}\right)(i\omega\tau)\widetilde{D} \tag{34}$$

Therefore the full system is:

$$\begin{aligned}\begin{pmatrix} \widetilde{T} \\ \widetilde{E}_+ \\ \widetilde{D} \\ \widetilde{E}_- \end{pmatrix} &= (i\omega\tau)^{-\alpha} \begin{pmatrix} (i\omega\tau)^{-h} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \dfrac{1}{1+(i\omega\tau)^h} & 0 \\ 0 & 0 & 0 & \dfrac{i\omega\tau}{1+(i\omega\tau)^h} \end{pmatrix} \\ &\times \begin{pmatrix} \sigma_T & 0 & 0 & 0 \\ 0 & \sigma_{E_+} & 0 & 0 \\ 0 & 0 & \sigma_D & 0 \\ 0 & 0 & 0 & \dfrac{\tau_D}{\tau} \end{pmatrix} \\ &\times \begin{pmatrix} 1 & 0 \\ \rho_E & \sqrt{1-\rho_E^2} \\ \rho_D & \mathrm{sgn}(r)\sqrt{1-\rho_D^2} \\ \rho_D & \mathrm{sgn}(r)\sqrt{1-\rho_D^2} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}\end{aligned} \tag{35}$$

From this, we can find E, O:

$$\begin{aligned}\widetilde{E} &= \frac{1}{2}(\widetilde{E}_+ - \widetilde{E}_-) \\ \widetilde{O} &= \frac{1}{2}(\widetilde{E}_+ + \widetilde{E}_-)\end{aligned} \tag{36}$$

The explicit formulae for $E_\pm$ are:

$$\begin{aligned}\widetilde{E}_+ &= (i\omega\tau)^{-\alpha}\left[\sigma_{E_+}\left(\rho_E\widetilde{\gamma}_1 + \sqrt{1-\rho_E^2}\widetilde{\gamma}_2\right)\right] \\ \widetilde{E}_- &= (i\omega\tau)^{-\alpha}\frac{i\omega\tau_D}{1+(i\omega\tau)^h}\left(\rho_D\widetilde{\gamma}_1 + \mathrm{sgn}(r)\sqrt{1-\rho_D^2}\widetilde{\gamma}_2\right)\end{aligned} \tag{37}$$

The overall final statistics are:

$$\begin{aligned}\underline{\underline{\widetilde{R}}}(\omega) &= \begin{pmatrix} \left\langle |\widetilde{T}|^2 \right\rangle & \left\langle \widetilde{T}\widetilde{E}_+^* \right\rangle & \left\langle \widetilde{T}\widetilde{D}^* \right\rangle & \left\langle \widetilde{T}\widetilde{E}_-^* \right\rangle \\ \left\langle \widetilde{E}_+\widetilde{T}^* \right\rangle & \left\langle |\widetilde{E}_+|^2 \right\rangle & \left\langle \widetilde{E}_+\widetilde{D}^* \right\rangle & \left\langle \widetilde{E}_+\widetilde{E}_-^* \right\rangle \\ \left\langle \widetilde{D}\widetilde{T}^* \right\rangle & \left\langle \widetilde{D}\widetilde{E}_+^* \right\rangle & \left\langle |\widetilde{D}|^2 \right\rangle & \left\langle \widetilde{D}\widetilde{E}_-^* \right\rangle \\ \left\langle \widetilde{E}_-\widetilde{T}^* \right\rangle & \left\langle \widetilde{E}_-\widetilde{E}_-^* \right\rangle & \left\langle \widetilde{E}_-\widetilde{D}^* \right\rangle & \left\langle |\widetilde{E}_-|^2 \right\rangle \end{pmatrix} \\ &= |\omega\tau|^{-2\alpha} \begin{pmatrix} |\omega\tau|^{-2h}\sigma_T^2 & (i\omega\tau)^{-h}\rho_{TE}\sigma_T\sigma_E & \dfrac{\rho_{TD}\sigma_D\sigma_T}{(i\omega\tau)^h(1+(-i\omega\tau)^h)} & \dfrac{\rho_{TD}\tau_D\sigma_T(-i\omega\tau)^{1-h}}{\tau(1+(i\omega\tau)^h)} \\ (-i\omega\tau)^{-h}\rho_{TE}\sigma_T\sigma_E & \sigma_E^2 & \dfrac{\rho_{ED}\sigma_E\sigma_D}{(1+(-i\omega\tau)^h)} & \dfrac{\rho_{ED}\tau_D\sigma_E(-i\omega\tau)}{\tau(1+(i\omega\tau)^h)} \\ \dfrac{\rho_{TD}\sigma_D\sigma_T}{(-i\omega\tau)^h(1+(i\omega\tau)^h)} & \dfrac{\rho_{ED}\sigma_E\sigma_D}{(1+(i\omega\tau)^h)} & \dfrac{\sigma_D^2}{|1+(i\omega\tau)^h|^2} & \dfrac{\sigma_D\tau_D(-i\omega\tau)}{\tau|1+(i\omega\tau)^h|^2} \\ \dfrac{\rho_{TD}\tau_D\sigma_T(i\omega\tau)^{1-h}}{\tau(1+(i\omega\tau)^h)} & \dfrac{\rho_{ED}\tau_D\sigma_E(i\omega\tau)}{\tau(1+(i\omega\tau)^h)} & \dfrac{\sigma_D\tau_D(i\omega\tau)}{\tau|1+(i\omega\tau)^h|^2} & \dfrac{\tau_D^2|\omega\tau|^2}{\tau^2|1+(i\omega\tau)^h|^2} \end{pmatrix}\end{aligned} \tag{38}$$

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

Using equations (36) and (37), the spectra of E, O can be determined:

$$\begin{aligned}\left\langle|\widetilde{E}|^2\right\rangle &= \frac{1}{4}\left(\left\langle|\widetilde{E_+}|^2\right\rangle + \left\langle|\widetilde{E_-}|^2\right\rangle - 2\left\langle\widetilde{E_+}\widetilde{E_-^*}\right\rangle\right) \approx \frac{1}{4}\left\langle|\widetilde{E_+}|^2\right\rangle \\ \left\langle|\widetilde{O}|^2\right\rangle &= \frac{1}{4}\left(\left\langle|\widetilde{E_+}|^2\right\rangle + \left\langle|\widetilde{E_-}|^2\right\rangle + 2\left\langle\widetilde{E_+}\widetilde{E_-^*}\right\rangle\right) \approx \frac{1}{4}\left\langle|\widetilde{E_+}|^2\right\rangle\end{aligned} \quad (39)$$

The far-right approximation can be seen from equation (37) using the fact that τ_D is the resolution of the series, such that for the full range of empirically accessible frequencies, we have $\omega\tau_D < 1$. In addition, because $\tau > \tau_D$, the factor $|\omega\tau_D/(1 + (i\omega\tau)^h)| < <1$.

The Properties of the Model

Scaling Properties

High- and Low-Frequency Exponents. Given that both the model equations and the corresponding model statistics are both readily expressible in the spectral domain, it is tempting to empirically evaluate the model directly using spectra. Unfortunately, although spectra are commonly used, they suffer from various technical issues such as spectral leakage (the "smearing out" of spectral variance across a range of frequencies), effects that are important whenever strong spectral peaks are present. Problematic peaks occur not only in quasi-periodic signals (where most of the effort has been made [e.g., Springford et al. 2020]) but also scaling signals, especially when the spectral exponent β (eq. 19) is significant (β = 0 for white noise [Hébert et al. 2021]). As we discuss in Appendix 3, the challenges for estimating spectra are much greater when the data are sampled at irregular intervals (as they typically are in paleobiology), and in particular, when the chronology itself (i.e., the temporal measurement density) is scaling (Appendix 3, Fig. A3.1). Indeed, the problem of how best to estimate the spectra of data with scaling chronologies (i.e., with holes over a wide range of scales) is still unsolved, and many nonstandard approaches give large biases.

Even if one can carefully handle the technical aspects, spectra still have the difficulty that their interpretation remains problematic. This was dramatically illustrated when Lovejoy (2015) discovered that the iconic (and still frequently cited) spectrum of Mitchell (1976) was in error by a factor of 10^{15}, an error that had not been noticed for four decades and was probably a consequence of spectra that are often not plotted with amplitude units at all or with undetected errors in the units. If Mitchell's spectrum had been accurate, two consecutive million year average Earth temperatures would only have differed by about 10 microKelvins (μK)—yet this patently false implication was not noticed because the units of the spectrum (K^2yr) were not intuitive, whereas an RMS Haar fluctuation of 10 μK would have been obviously problematic. Even spectral updates as recent as 2020 are in error by a factor 10^{11} (see the review by Lovejoy [2023]). The comparison of the fluctuation analyses in this paper with those of the spectra (Appendix 3) highlights the power of fluctuation analysis when applied to irregularly sampled data.

Fortunately, over scaling ranges, the fluctuation and spectral statistics are both scaling with exponents as indicated in the previous section. Therefore, to interpret the statistics (eqs. 38, 39) in real space, it suffices to use the fact that Fourier scaling implies real-space scaling and to use the above relations between real-space and Fourier scaling exponents (eq. 22). In matrix form, the spectral exponents are therefore:

$$\beta_h = \begin{pmatrix} 2(\alpha+h) & 2\alpha+h & 2(\alpha+h) & 2(\alpha+h)-1 \\ 2\alpha+h & 2\alpha & 2\alpha+h & 2\alpha+h-1 \\ 2(\alpha+h) & 2\alpha+h & 2(\alpha+h) & 2(\alpha+h)-1 \\ 2(\alpha+h)-1 & 2\alpha+h-1 & 2(\alpha+h)-1 & 2(\alpha+h-1) \end{pmatrix}$$

$$\beta_l = \begin{pmatrix} 2(\alpha+h) & 2\alpha+h & 2\alpha+h & 2\alpha+h-1 \\ 2\alpha+h & 2\alpha & 2\alpha & 2\alpha-1 \\ 2\alpha+h & 2\alpha & 2\alpha & 2\alpha-1 \\ 2\alpha+h-1 & 2\alpha-1 & 2\alpha-1 & 2(\alpha-1) \end{pmatrix} \quad (40)$$

(The elements correspond to T, E_+, D, and E_-, left to right, top to bottom). Using the relationship between H and β (eq. 22), the high and low frequency (here small and large times, t) have exponents:

$$H_h = \begin{pmatrix} \alpha+h-\frac{1}{2} & \alpha+\frac{h-1}{2} & \alpha+h-\frac{1}{2} & \alpha+h-1 \\ \alpha+\frac{h-1}{2} & \alpha-\frac{1}{2} & \alpha+\frac{h-1}{2} & \alpha+\frac{h}{2}-1 \\ \alpha+h-\frac{1}{2} & \alpha+\frac{h-1}{2} & \alpha+h-\frac{1}{2} & \alpha+h-1 \\ \alpha+h-1 & \alpha+\frac{h}{2}-1 & \alpha+h-1 & \alpha+h-\frac{3}{2} \end{pmatrix} \quad (41)$$

While at low frequencies, large Δt (i.e., large lags) we have:

$$H_l = \begin{pmatrix} \alpha+h-\frac{1}{2} & \alpha+\frac{h-1}{2} & \alpha+\frac{h-1}{2} & \alpha+\frac{h}{2}-1 \\ \alpha+\frac{h-1}{2} & \alpha-\frac{1}{2} & \alpha-\frac{1}{2} & \alpha-1 \\ \alpha+\frac{h-1}{2} & \alpha-\frac{1}{2} & \alpha-\frac{1}{2} & \alpha-1 \\ \alpha+\frac{h}{2}-1 & \alpha-1 & \alpha-1 & \alpha-\frac{3}{2} \end{pmatrix} \quad (42)$$

We should add here that because E, O are linear combinations of E_+, E_-, their exponents will the maximum of those of E_+, E_-, so that:

$$\begin{aligned} H_{h,E_\pm} &= \max\left(\alpha-\frac{1}{2}, \alpha+h-\frac{3}{2}\right) = \alpha-\frac{1}{2}; \quad h<1 \\ H_{l,E_\pm} &= \max\left(\alpha-\frac{1}{2}, \alpha-\frac{3}{2}\right) = \alpha-\frac{1}{2} \end{aligned} \quad (43)$$

We see that for the physically relevant parameters, $H = \alpha - ½ = -0.25$ for both E, O, over the whole range (close to the data, see SL and Fig. 4).

To get a concrete idea of the implications of model, let us use the rough empirical estimates from SL of $\alpha = 0.25$, $h = 0.5$. Plugging these values into equations (41) and (42), we obtain:

$$H_h = \begin{pmatrix} 0.25 & 0 & 0.25 & -0.25 \\ 0 & -0.25 & 0 & -0.5 \\ 0.25 & 0 & 0.25 & -0.25 \\ -0.25 & -0.5 & -0.25 & -0.75 \end{pmatrix}$$

$$H_l = \begin{pmatrix} 0.25 & 0 & 0 & -0.5 \\ 0 & -0.25 & -0.25 & -0.75 \\ 0 & -0.25 & -0.25 & -0.75 \\ -0.5 & -0.75 & -0.75 & -1.25 \end{pmatrix} \quad (44)$$

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

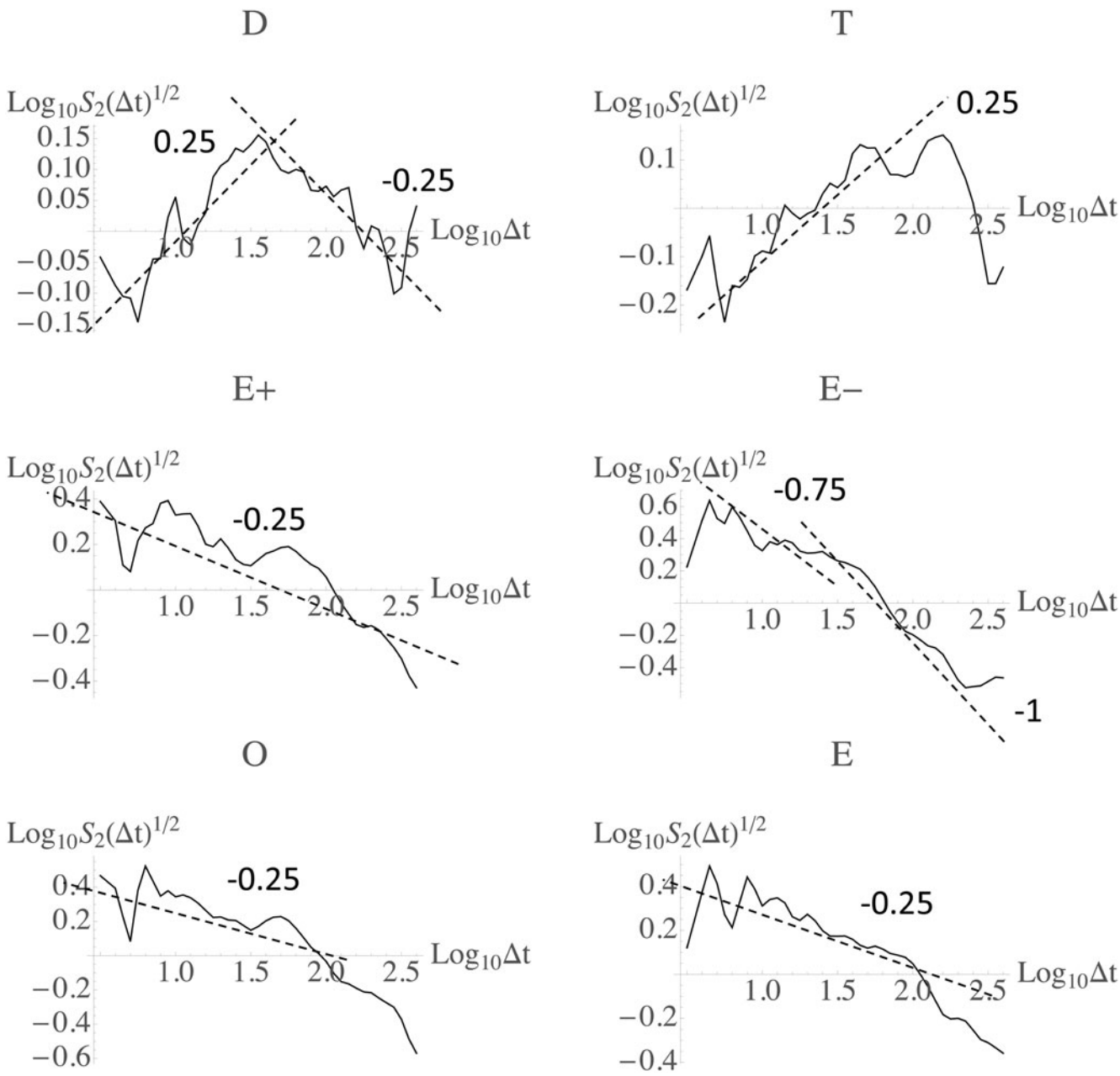


Figure 4. This shows the Phanerozoic marine animal macroevolutionary analysis of the six series discussed in this paper; *D*, *T*, *O*, *E* are replotted from SL. The dashed lines show the theory slopes (eq. 44) with transition at $\Delta t \approx 40$ Myr, i.e., $\log_{10}\Delta t \approx 1.6$.

(again, for *T*, E_+, *D*, and E_-, left to right, top to bottom). We can see that the Haar fluctuations will be useful for all the series over the whole range of frequencies/scales, the only exception being $\Delta D(\Delta t)$ at long lags ($H_l < -1$, lower right corner of the H_l matrix with $H_l < -1$). In this case, the Haar fluctuations "saturate," and the spurious (limiting) value $H_l = -1$ is obtained.

With these results, we can comment on the issue of the stationarity/nonstationarity of the statistics. In the real world, scaling regimes only exist over finite ranges of scale, they have finite inner and outer limits. However, the outer scaling limits in mathematical series may be infinite. In this case, scaling series with low frequency $H_l < 0$ are stationary, whereas if $H_l > 0$, they will be nonstationary (although even here, if $H_l < 1$, the increments of the series will be stationary). From equation (44), we see that the only correlation with $H_l > 0$ is for the temperature; hence all the evolutionary responses (including cross-correlations) are stationary. However, even the temperature will be stationary at timescales beyond the outer limit of the scaling regime (which we argue is at least of the order of the length of the Phanerozoic). That is why we prefer the term "wandering" when $H_l > 0$: for such scaling processes, the nonstationarity may be spurious, the series only *appears* to be nonstationary over the observed (finite) range of scales. On a more fundamental note, no empirical series is stationary or nonstationary, the latter are properties of models, not data.

Normalized Correlations. The cross-spectra and cross-covariances (eq. 38) can be used to determine the normalized

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

correlations that were estimated in SL:

$$\rho_{jk}(\Delta t) = \frac{\langle \Delta S_j(\Delta t) \Delta S_k(\Delta t) \rangle}{\langle \Delta S_j(\Delta t)^2 \rangle^{1/2} \langle \Delta S_k(\Delta t)^2 \rangle^{1/2}} \quad (45)$$

(using Haar fluctuations). However, from equations (41) and (42), we find that their exponents (whether at high or low frequencies) are $2H_{jk} - (H_{jj} + H_{kk}) = 0$, that is, they are *not* power laws and only vary at sub–power law rates, and they are therefore nontrivial (i.e., they are significant) over the whole range of Δt. Because there are six series (T, E, D, O, E_+, E_-) there are 15 pairs whose fluctuation correlations may be determined over the observed range of $3 \approx < \Delta t \approx < 400$ Myr; see Figure 5. The key correlations are those that correspond to the model parameters: $\rho_E = \rho_{TE}$, $\rho_D = \rho_{TD}$, see equation (32). We can already see that the correlations are quite noisy, a consequence of the low resolution and variable sampling of the series. To make a proper model–data comparison, we therefore turn to numerical simulations.

Numerical Simulations

The Statistics of the Simulated Series

The model has two fundamental exponents (α, h), two basic correlations ($\rho_E = \rho_{TE+}$, $\rho_D = \rho_{TD}$), and a crossover timescale τ.

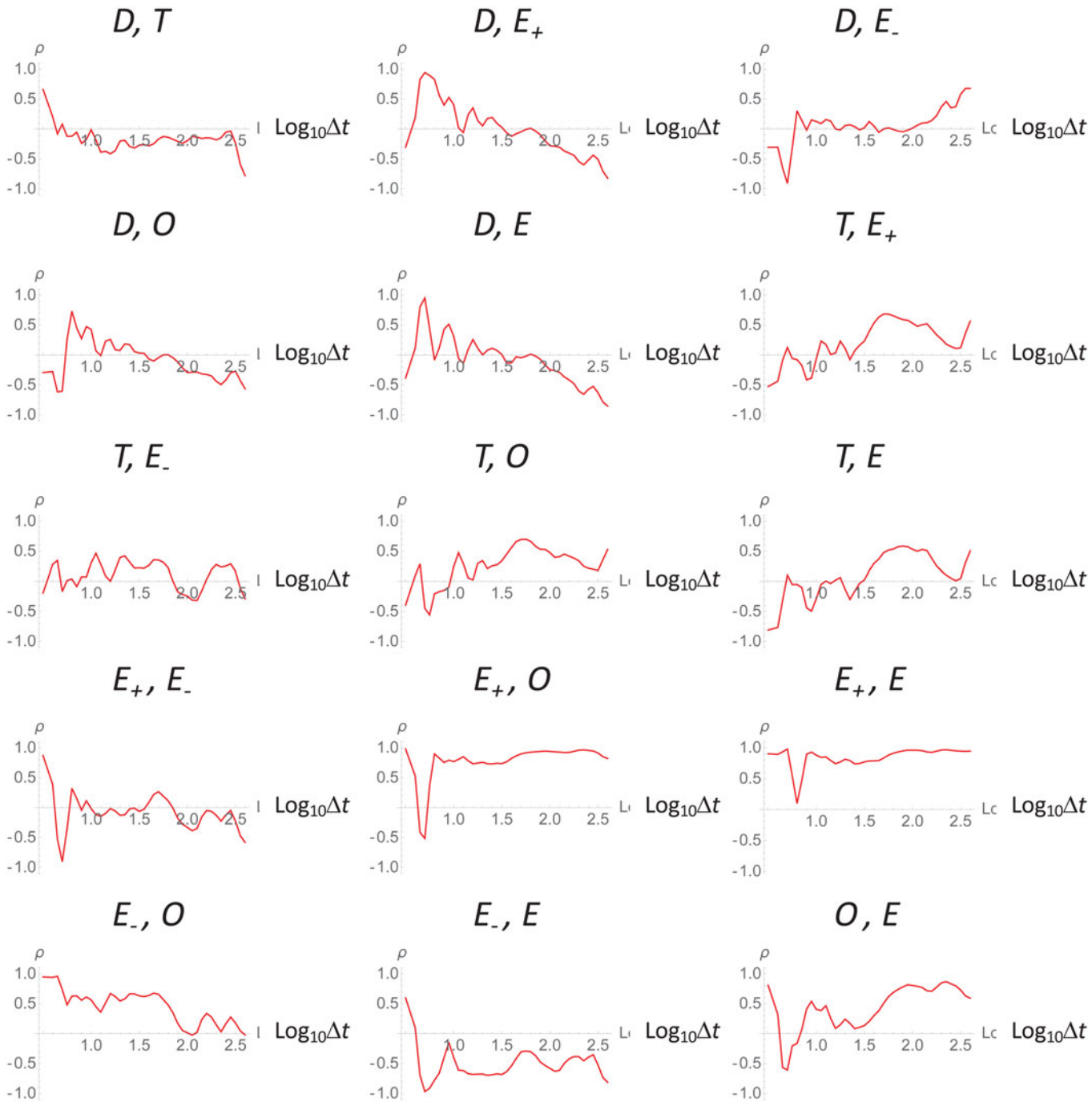


Figure 5. The (normalized) pairwise correlations of the 15 pairs of the six series as functions of lag. Several of these are reproduced from SL.

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

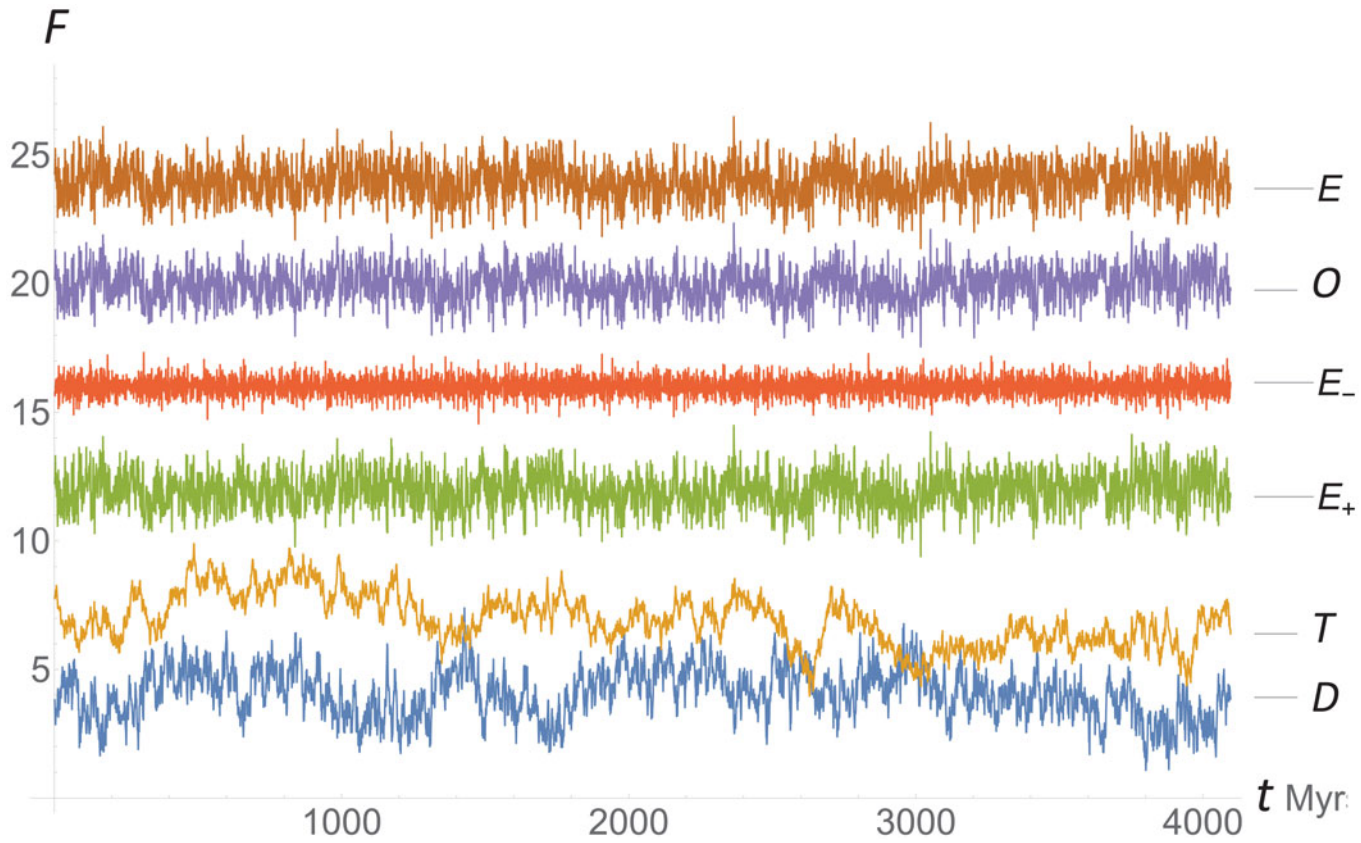


Figure 6. The previous 2^{14} simulation degraded from ¼ Myr resolution to 1 Myr. Curves normalized by their standard deviations and then offset by 5 units in the vertical for clarity.

The third correlation (ρ_{DE}) is a derived parameter (eq. 32). In addition, there are two amplitude factors, σ_T, σ_E, but these will depend on the nondimensionalization/normalization of the series; on log-log plots, they correspond to an up–down shift, and on (normalized) correlation plots, the normalization eliminates them; they will not be considered further.

We used the results of SL to fix the values $\alpha = 0.25$, $h = 0.5$, $\tau =$ 32 Myr. (This is the nearest power of 2 to the slightly larger—but only roughly estimated—value $\tau = 40$ Myr in SL. Also, due to the scaling of the model, the more significant parameter is log τ and $\log_2 40 = 5.3$, not far from $\log_2 32 = 5$.) This leaves the only unknown parameters as the TE and TD correlations ($\rho_E = \rho_{TE+}$, $\rho_D = \rho_{TD}$; Fig. 5).

Before comparing the model directly to the (noisy) data, we checked that we were able to numerically reproduce the theoretically expected behavior. The basic modeling technique is to use convolutions with various (impulse response) Green's functions; this is detailed in Appendix 1, but follows the methods described in Lovejoy (2022). The main numerical problems are the small scales that have singular power law filters that are not trivial to discretize, and there are some (easier to handle) long-time (low-frequency) issues.

Rather than attempting to rigorously determine optimum parameters, as indicated earlier, we fixed the exponents $\alpha =$ 0.25, $h = 0.5$, and the crossover scale $\tau = 32$ Myr. With guidance of the Figure 5 correlations for ρ_{TE+}, ρ_{TD} and some numerical experimentation, we took $\rho_E = 0.5$, $\rho_D = -0.1$ (hence $\rho_{TE+} = 0.5$, $\rho_{TD} = -0.1$, $\rho_{DE+} = -0.9$; i.e., the sign of r was taken as negative, eq. 32). We then performed simulations at a resolution of 250 kyr resolution, with simulation length of ≈4 Gyr ($2^{14} = 16{,}384$ points), shown in Figure 6. This single very large realization has statistics that are close to those of an infinite ensemble. We postpone a discussion of the significance of the correlations to the section "The Statistics of the Simulated Series Resampled at the Data Sampling Times."

According to the model (see the diagonal elements in eq. 44), the only series with positive low-frequency scaling exponent ($H_l > 0$) is the temperature ($H_l = 0.25$), which indeed shows "wandering" behavior (second from the bottom in Fig. 6); from the figure, one can see its long-range correlations as low-frequency undulations. This is also true for D, but only up to the crossover scale (≈32 Myr), after which consecutive 32 Myr intervals tend to cancel ($H_l < 0$, eq. 44). The other series are, on the contrary, "canceling" ($H_l < 0$, $H_h < 0$) especially E_- (eq. 44). We can also visually make out some of the correlations, but this is clearer at lower resolution, as discussed later.

On these simulations, we can check that the theoretical scaling is obeyed; this was done using Haar fluctuations; see Figure 7 where the theory slopes (from eqs. 43, 44) are shown as reference lines. Note that because the Haar analysis "saturates" at $H = -1$, the low-frequency $H_l = -1.25$ value for E_- (eq. 44, lower right-hand diagonal element) yields a slope of −1 (not −1.25); however, the other slopes are accurately estimated. Note that the theory–simulation agreement is not perfect, mostly because the theory is for the average statistics over an infinite ensemble, whereas Figure 7 is from a single—albeit large—simulation.

We can also work out the 15 correlations as functions of lag (Fig. 8). The figure shows the model parameters $\rho_{TE+} = 0.5$

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

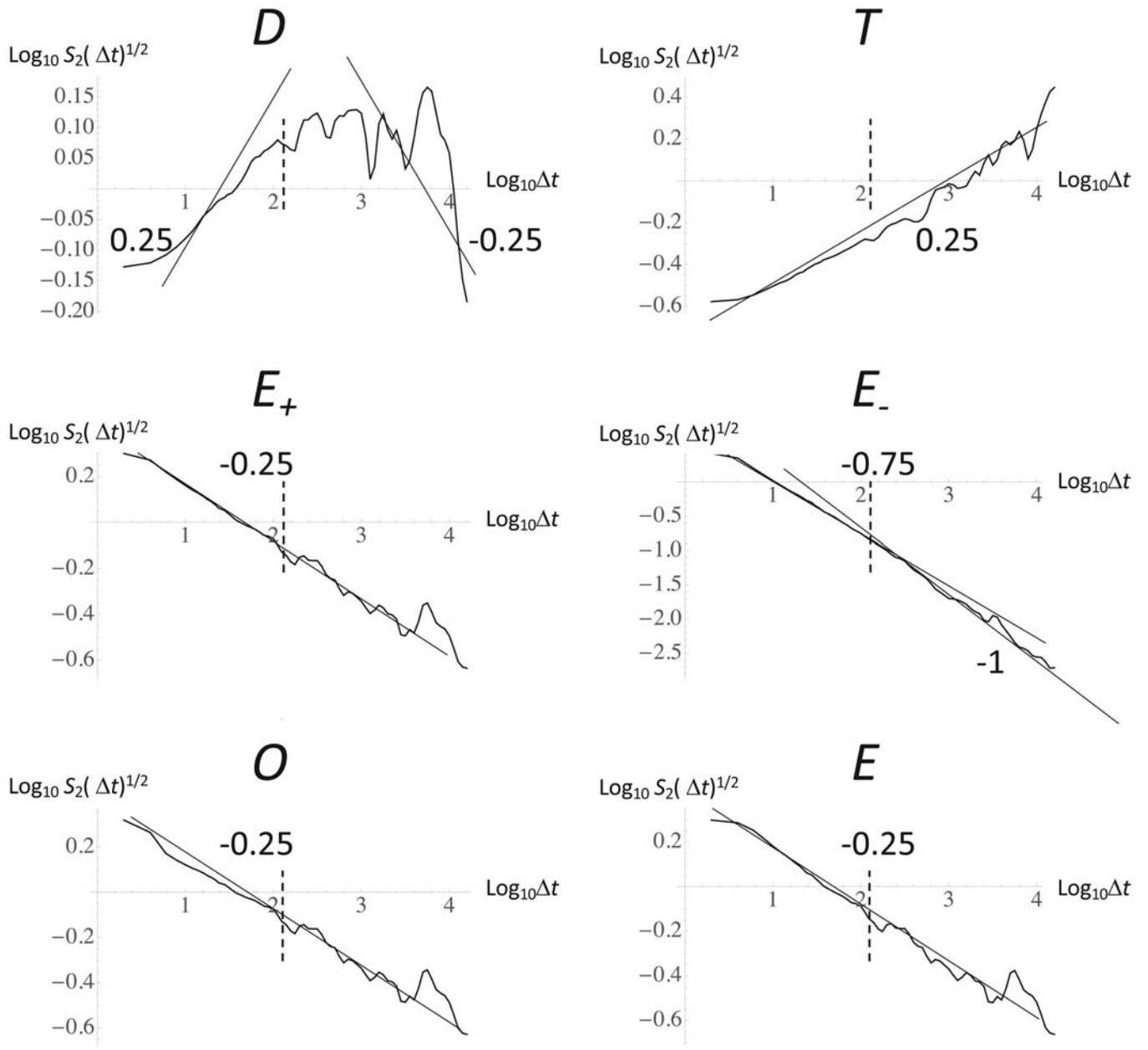


Figure 7. Simulation 2^{14} =16,384 points with theoretical slopes indicated. The transition scale τ is 2^7 = 128 units, indicated by dashed vertical lines. The could represent a series modeled at 250 kyr resolution with a total simulation length of 4 Gyr and with crossover at 32 Myr. Due to its length, this simulation has statistics that are close to the ensemble averaged statistics. The parameters are: $\alpha = 0.25$, $h = 0.5$, $\rho_E = \rho_{TE} = 0.5$, $\rho_D = \rho_{TD} = -0.1$ (with derived *DE* correlation $\rho_{DE} = -0.9$).

($= \rho_E = 0.5$), $\rho_{TD} = -0.1$ ($= \rho_D = -0.1$) as solid black reference lines, and the derived correlation $\rho_{DE+} = -0.9$ (eq. 32) as a dashed reference lines. Also shown are dashed theory lines for the *TE*, *TO* correlations (predicted to be equal to equal to TE_+ at long lags, eq. 39) and the *DE*, *DO* correlations (predicted to be equal to DE_+ at long lags, see eq. 39). We can see that the correlations approach the theoretical correlations at large lags, although the results are somewhat noisy.

The Statistics of the Simulated Series Resampled at the Data Sampling Times

Before making more effort at parameter fitting and comparing the model to data, it is important to take into account the small number of empirical data points and their irregular sampling (the corresponding geochronology itself turns out to be scaling; see Appendix 3, Fig. A3.1). Figure 9 shows the result for a simulation with the same parameters, but with a 1 Myr temporal resolution (right-hand side), resampled at the same times as the data (left-hand side). Because the model and data are only expected to have similar statistics, the detailed "bumps" and "wiggles" are unimportant, but one can nevertheless make out realistic-looking variability, including correlations between the series. Note that the model respects causality, so when there is a large extinction event, the curve is asymmetric with a rapid upturn being followed by a slower downturn (however, we have followed convention such that the present is at the left and the past at the right).

We can now consider the fluctuation scaling and correlation statistics on the resampled series and compare them with both the data and the results from the same simulations but at a regular

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

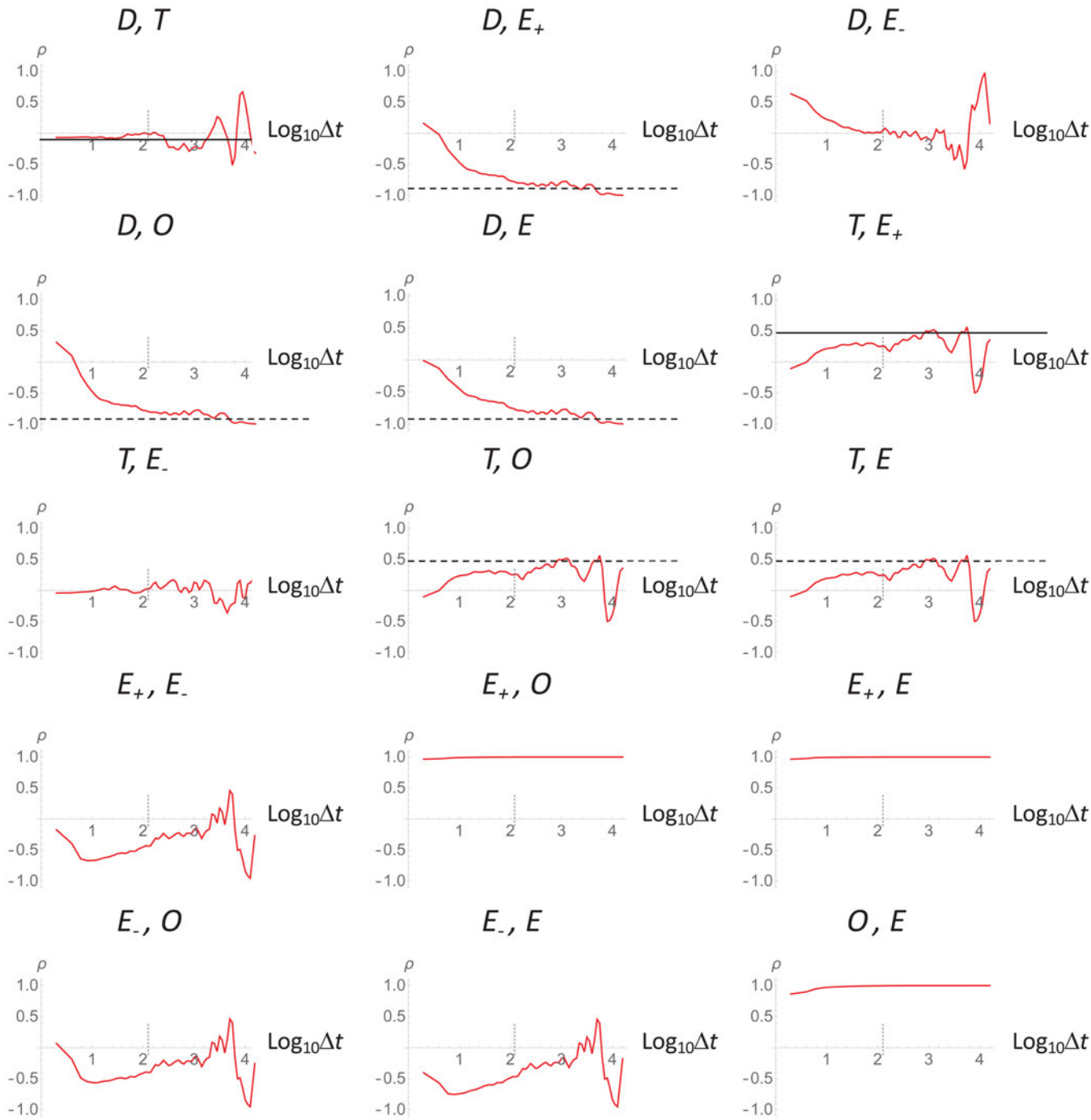


Figure 8. The 15 pairwise correlations from the 2^{14} realization in Fig. 7. Only two of the correlations were prescribed, and this only at a single resolution; the rest are consequences of the model, the two exponents α, h, and the crossover time $\tau = 2^7$ (shown as short dashed vertical lines). The two prescribed correlations (*DT*, TE_+) are shown as solid horizontal lines, and the derived correlations are shown as dashed lines (DE_+ from *DT*, TE_+, eq. 32) and then *TE*, *TO* (predicted to be equal to equal to TE_+ at long lags, eq. 39) and *DE*, *DO* (predicted to be equal to DE_+, at long lags, eq. 39). Note that these are from a single realization of the process, not the ensemble average. In addition, the statistics of some are fairly sensitive to irregularly sampled (and small size) of the empirical data; compare with Fig. 11.

1 Myr resolution (Fig. 10; see Fig. A2.1 for the *E*, *O* plots for the alternative Sepkoski stages database; Appendix 2, the *E*, *O* scaling and correlations are nearly identical). The figure shows a log-log plot of the RMS fluctuations as a function of the lag. To make the comparison, they were normalized by their standard deviations, but this is somewhat arbitrary, such that the up–down displacement (corresponding to a different nondimensionalization/ normalization) is unimportant. To judge the realism of the model, the appropriate comparison is between the shapes of the resampled model output (red) and the data (black). We can see that the two are fairly close, although both model and data are noisy due to the small number of points and the irregular sampling. The agreement must be assessed not only by allowing for (relative) vertical shifts, but also by noting that the scales on

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

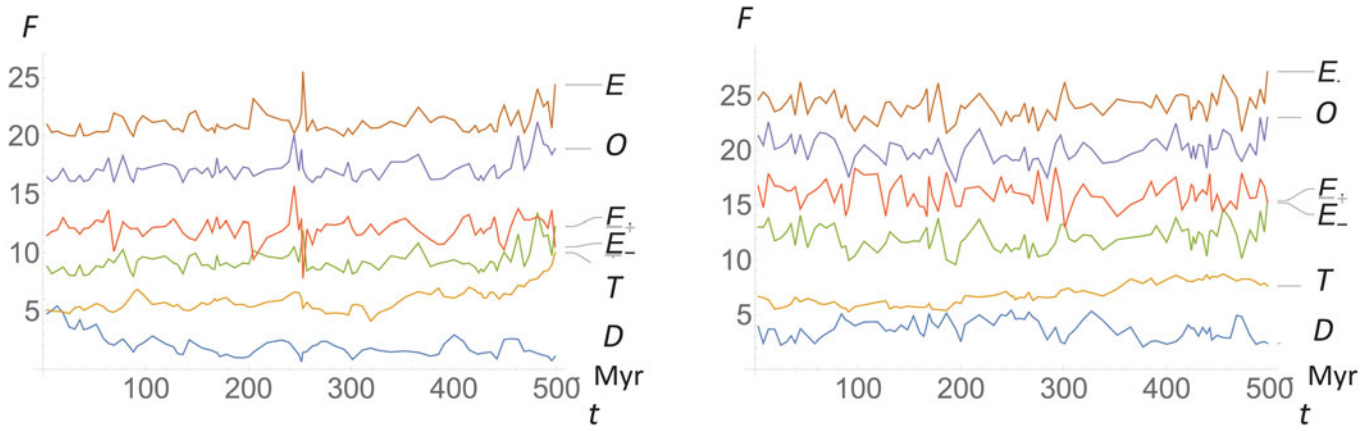


Figure 9. Model–simulation comparison with all series normalized by their standard deviations. The simulation was at 1 Myr resolution, and it was sampled at the same (irregular) times as the data (84 points over the last 500 Myr). Each curve was displaced by 5 units in the vertical for clarity. Due to causality, the series are asymmetric, with time running from right to left. The simulation is on the right.

the top D, T comparisons are such that the fluctuations vary only over a small range (for the data, factors of ≈1.7 for D and ≈2 for T) for lags varying over range of about a factor 100. In comparison, the E_+, E_-, E, O ranges are closer to factors of 10. Aside from this, these basic fluctuation statistics are fairly close to the data.

Figure 10 also gives important information about the effect of the sampling: compare the resampled (red) and uniformly sampled analyses (brown). The resampling is particularly important for E_+, E_-, E, O, although the effects are mostly at small lags for E_+, E, O but at large lags for E_-. This information should prove useful in interpreting a variety of real-world extinction and origination data.

Finally, we can compare the 15 pairwise correlations (Fig. 11). Again, to judge the realism of the model, compare the red and black correlations. Although—as expected—these are fairly noisy, we see that the agreement is quite good, significantly, it is generally much better than the agreement between the uniformly sampled correlations (brown curves) and data (black). By comparing the red (resampled) and brown (uniformly sampled) correlations, we see that the resampling is especially important for the DE_+, DO, DE, E_+E_-, E_-O, OE, correlations and to a lesser extent, the OE, TE_+ comparisons; for the others, it is about the same. We could note the successful prediction that the E_+E, E_+O, OE correlations should be ≈1 and the E_-E correlations should be ≈ −1. Interestingly, the prediction that the E_-O correlations should be ≈ −1 (eq. 39) is verified with the uniform sampling (i.e., it is indeed a property of the model), yet the resampling (red in the lower left graph in Fig. 11) makes it > 0 and aligns it closely with the observations. In other words, when the pure model predictions are poor (brown vs. black), there are many instances where the effects of nonuniform sampling are particularly strong, such that overall the model explains the data fairly well: overall 6 fluctuation plots (Fig. 10) and 15 correlations (Fig. 11) with 5 adjustable parameters (α, h, τ, ρ_E, ρ_T).

Discussion of the Model and Physical Significance of the Correlations

The model was motivated by an attempt to model the diversity process as a scaling crossover phenomenon with wandering climate (paleotemperature) and stabilizing life (turnover) scaling drivers. In the course of the model development, it became clear both theoretically (due to the definition of the diversity, eq. 5) and empirically that rather than E, O being fundamental, it was rather the turnover E_+ that is fundamental (indeed, the E_+ and E_- statistics are quite different (Figs. 4, 10), and the E_+E_- correlations are nearly zero (Figs. 5, 11). In any event, the model predicted that E, O would follow the E_+ statistics (eq. 39; Figs. 4, 10, and the E_+E and E_+O correlations in Figs. 4, 11).

A more counterintuitive finding concerns the correlations. To start with, the model specifies that the diversity is primarily driven by the temperature up until the crossover scale, yet the temperature and diversity are negatively correlated over the entire range! Although at any given time lag, the DT correlation is small (−0.1), it means that there is a (weak) tendency for the diversity fluctuations to decrease when temperature fluctuations increase and vice versa, but this is not enough to offset the overall temperature control of the diversity that implies that consecutive temperature fluctuations tend to add up ($H_T = 0.25 > 0$), and this is a stronger overall effect.

There is an additional more subtle effect. Consider that at each scale, the imposed TE_+ correlation is moderate and positive ($\rho_{TE+} = 0.5$), and together, ρ_{TD} (the temperature diversity correlation) and ρ_{TE+} (the temperature turnover correlation with $r < 0$, eq. 32) imply that at each lag, DE_+ is negatively correlated (reaching the theory value $\rho_{DE+} \approx -0.9$, at long lags; see the DE_+ correlation, the brown curve in Fig. 11). As the turnover E_+ also drives the diversity (eq. 1), at each scale, we thus have a tendency for T and E_+ fluctuations to increase (or decrease) together but for D and E_+ (and hence T and D) to have opposite tendencies. The overall result is that the weak anticorrelation of D with T and D with E_+ at any fixed scale is still dominated by the stronger effect of T fluctuations growing with scale and dominating the E_+ driver at lags < τ.

We could remark that $\rho_{TE+} = 0.5 > 0$ indicates a tendency for temperature changes to "stimulate" the turnover: periods of increasing temperatures tending to be associated with increasing turnovers, and periods of decreasing temperatures with decreasing turnovers. Also there is a strong anticorrelation between D and E_+ ($\rho_{DE+} \approx -0.9$) that indicates that turnover decreases with diversity (note that this anticorrelation seems to nearly disappear after the nonuniform sampling; see Fig. 11, second in the top row). However over the range of scales that E_+ dominates dynamics of D (i.e., $\Delta t > \tau$, as $H_{E+} \approx -0.25 < 0$), successive E_+ fluctuations

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

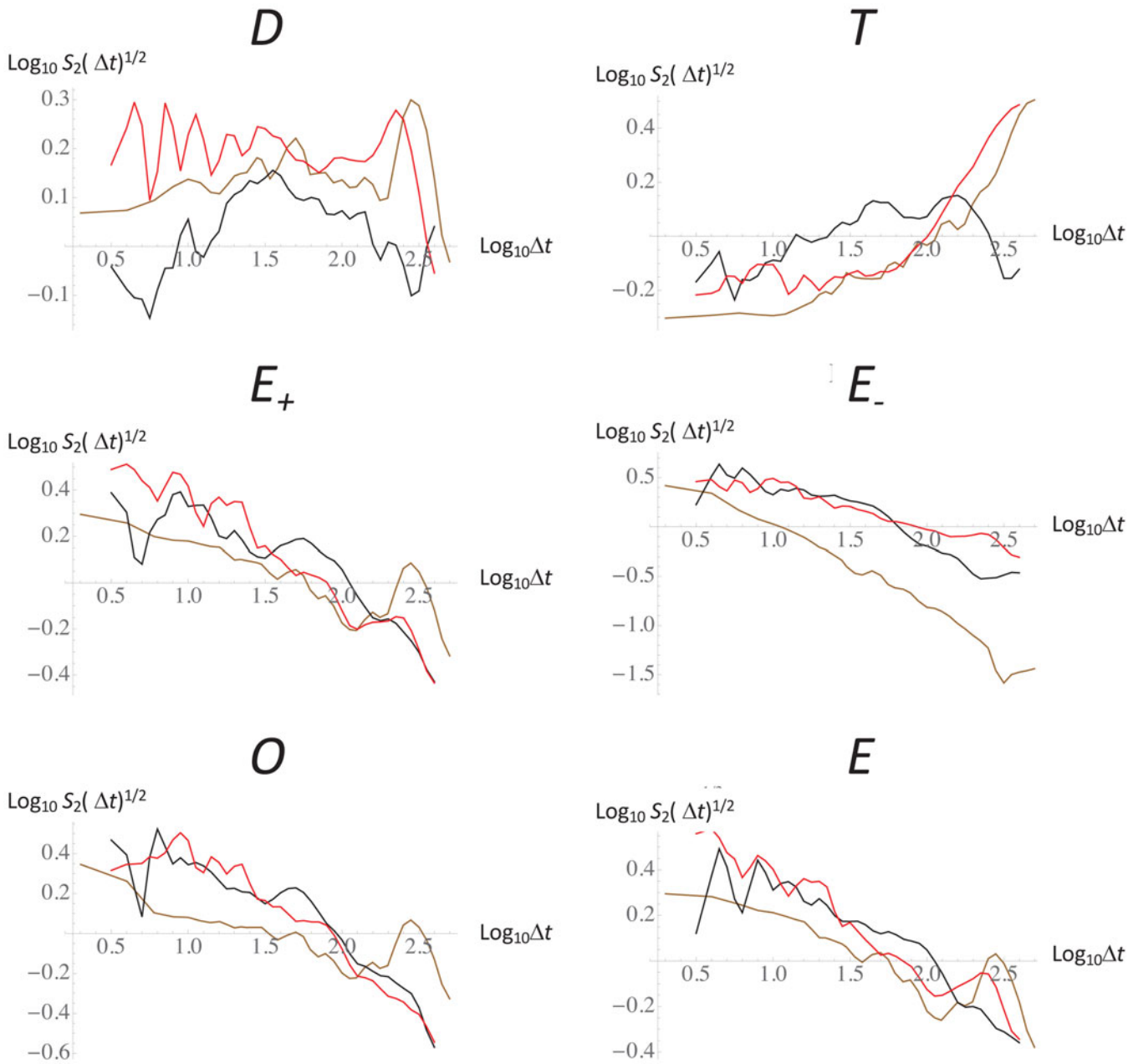


Figure 10. A comparison of the RMS Haar fluctuations for the 1 Myr resolution simulations discussed earlier (brown), from the simulation resampled at the data times (red), and from the data (black), these two irregularly sampled series are shown in Fig. 9. The relative vertical offsets of the curves are not significant; they correspond to specific normalizations/nondimensionalizations. We see that in general, the resampling at the data times (red) yields a closer fit to the data (black) than the analysis of the simulation at uniform (1 Myr) intervals; this is especially true for E_-, O, E, E_+. FMEM, Fractional MacroEvolution Model.

tend to cancel, and on long timescales, the latter effect is dominant, such that $H_D = H_{E+} \approx -0.25$—this is a scaling region of biotic self-regulation.

Discussion and Conclusions

The driver of macroevolutionary biodiversity has famously been reduced to a dichotomy between life and the environment: the metaphor of Red Queen versus Court Jester (Van Valen 1973; Barnosky 2001). Using genus-level time series from the PBDB, Spiridonov and Lovejoy (2022; SL) systematically analyzed fluctuations in extinction (E) and origination (O) rates, biodiversity (D), and paleotemperatures (T) over the Phanerozoic. They did this as a function of timescale from the shortest (≈3 Myr) to longest lags available (≈400 Myr), and their analysis included the correlations of the fluctuations at each scale. They concluded that T, E, O—the basic climate and life parameters—showed evidence of wide range scaling, supporting the hypothesis that over this range, there is a single biogeological "megaclimate" (Lovejoy 2015) regime with no fundamental timescale. However, they found that D followed the T fluctuations up until a critical time $\tau \approx 40$ Myr, whereas at longer timescales, it followed life (E, O): D was a scaling crossover phenomenon. At the shorter timescales $\Delta t < \tau$, following the temperature, the D scaling exponent $H_D \approx +0.25$ (i.e., > 0), indicating

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

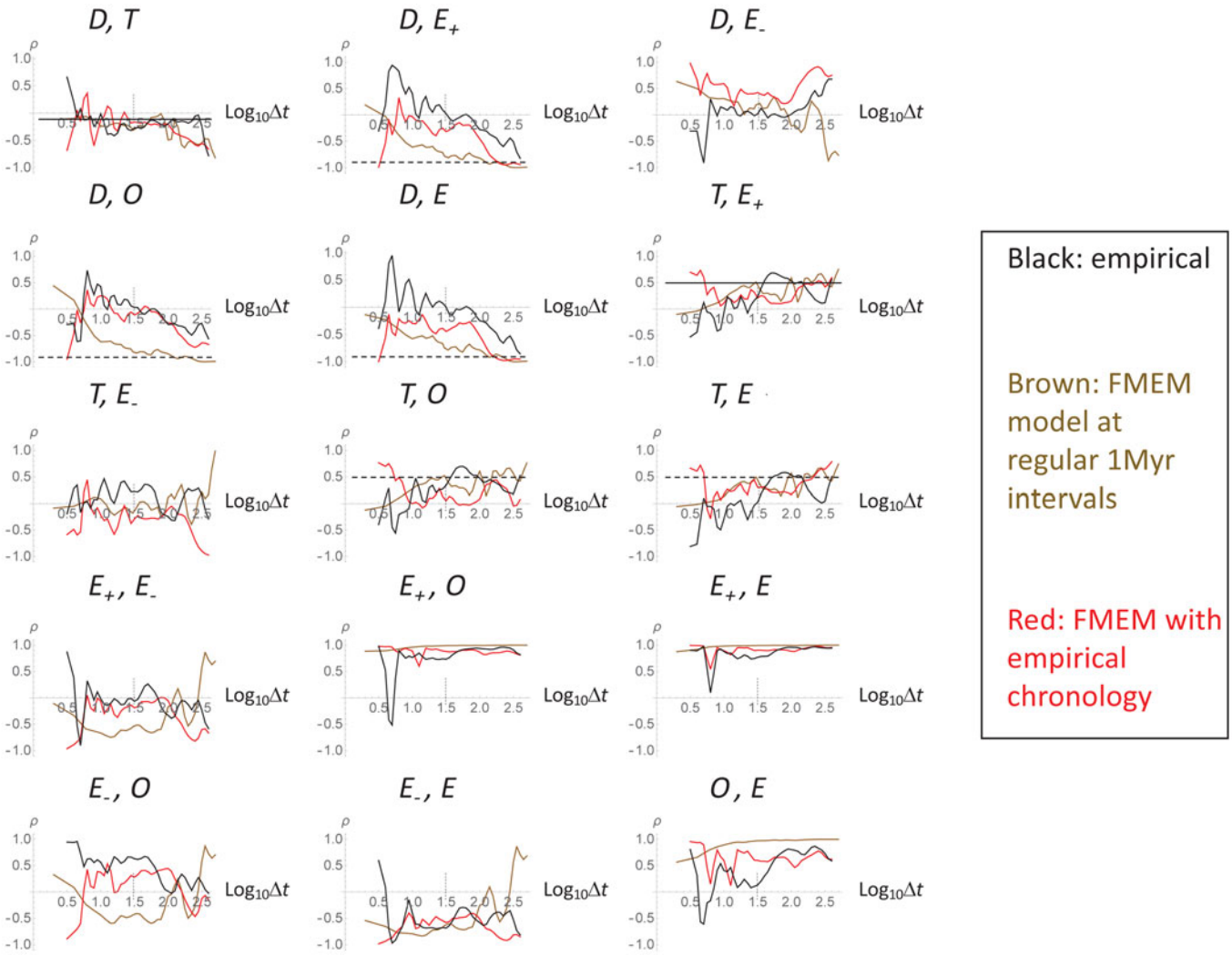


Figure 11. The pairwise correlations from the same three series as in fig. 10 with the same color codings: i.e., data (black), brown the simulation at a uniform 1 Myr resolution, and (red), the simulation resampled at the data times. The resampling notably improves the correlations for DE_+, DO, DE, E_+E_-, E_-O, OE, and to a lesser extent the OE, TE_+ comparisons; for the others, it is about the same. FMEM, Fractional MacroEvolution Model. The solid and dashed horizontal lines are the same as those in figure 8.

that fluctuations tended to grow with scale, leading to "wandering" behavior. In contrast, for time lags $\Delta t > \tau$, following E, O, its scaling exponent was $H_D \approx -0.25$ (i.e., < 0), hence successive fluctuations tended to cancel, resulting in long-time stabilization of diversity by life.

The model assumes, first, that in the Phanerozoic, dynamical processes occur over a wide range of timescales, and second, that the basic driving processes (here T, E_+) are also scaling, such that over a wide range, they define no characteristic timescale. This is the simplest assumption, yet it is compatible with rich behavior. The critical scale for the diversity (estimated at ≈40 Myr) is simply the scale at which one process (life/turnover) starts to dominate the other (climate/temperature).

The Phanerozoic evolution of life on our planet is full of contrasting patterns, quantitative and qualitative dynamical transitions, yet it may be compatible with an underlying scale symmetry. For example, transitions related to era boundaries include changes in complexity of marine animal communities in post-Paleozoic times (Wagner et al. 2006), the increase in longevities of animal genera in the same post-Paleozoic period (Miller and Foote 2003), or changes in the structure of bivalve communities after the K-Pg mass extinction (Aberhan and Kiessling 2015), yet each of these processes occur over wide ranges of scale.

In addition, it could be argued that the dynamics of biodiversity should be subdivided into these smaller (substage) time intervals, potentially extending the scaling range to shorter timescales (≈1 Myr or less). Important boundaries that modified macroevolutionary dynamics at global scales can also be found at period and subperiod boundaries: apparently the appearance of calcifying plankton at the beginning of the Jurassic (or so called "post-Conodontozoic"; Ferretti et al. 2020) should have had a significant effect in stabilizing hyperthermal events and thus decreasing the volatility of biospheric evolution (Wignall 2015; Eichenseer et al. 2019). In addition, presumably the appearance, global spread, and deep impact on the climate of terrestrial plants in the Silurian is highly significant (Lenton et al. 2016). There are also shorter timescale transitions—such as transitions toward more volatile dynamics of zooplankton between the Early and Middle Ordovician toward the Late Ordovician (Crampton et al. 2016).

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

Such changes are thus numerous and add to the plausibility that such diverse and sometimes sharply varying processes may be scaling. Therefore, the subdivision of the Phanerozoic into shorter time intervals in search of local scaling laws—in our view at least at this stage of the understanding and data availability—is rather premature and arbitrary and could preclude the search of overarching commonalities that can hide in this apparent heterogeneity. Moreover, scaling multifractal processes are more general than the Gaussian ones considered here; they are intermittent, typically involving occasional large changes in intensity (e.g., apparent transitions), occurrences of extremes. Physically, this could correspond to local dominance of controlling mechanisms (which can be hierarchically modulated by feedbacks from biota or external forcing) at all timescales where scaling regimes apply.

In this study, as well as in the study which first described scaling properties of Phanerozoic marine genus-level evolution, SL analyzed the processes and time series as a homogenous stationary system. Although this is an assumption that the *statistical* properties are constant over the eon, it is nevertheless consistent with high variability and sudden transitions (i.e., volatile and intermittent behavior). Although the exact FMEM model presented here was not intermittent, intermittency and its associated sharp transitions could easily be introduced simply by replacing the (Gaussian) white noise forcings by multifractal ones.

To clarify our ideas, to better understand the geobiodynamics, and to better understand and quantify the limitations, biases, and other data issues, we proposed the simple FMEM to reproduce the observations. It is a model of macroevolutionary biodiversity driven by paleotemperature (the climate proxy) and the turnover rate ($E_+ = O + E$), the life proxy. To fit with basic empirical scaling statistics and theoretical ideas about the macroclimate regime (form timescales of roughly 1 Myr to at least 400 Myr), these drivers were taken to be scaling with climate dominating at short timescales and life at long timescales. Therefore, FMEM suggests a possible way to combine into a single stochastic framework both: (1) the destabilizing geophysical (and possibly astrophysical) processes (Raup 1991, 1992a,b; Melott and Bambach 2014; Fields et al. 2020) with (2) the stabilizing, density-dependent, and self-regulating biotic processes. The model is specified by a simple parametrization based on two scaling exponents and two pairwise correlations (between T and E_+ and between T and D).

Because the FMEM model is linear, it can be used to model a superposition of stochastic and deterministic processes such as bolide impacts: the responses to the deterministic (external) and stochastic (internal) forcings simply add. As the impulse response function is a (singular) power law, a short impulse on the system will generate a sharp initial response followed by a long power law decay (i.e., much slower than the classical exponentials), decaying with somewhat different exponents depending on whether the impulse is on the temperature or turnover or both (see Fig. 2).

The model has two unusual characteristics: first, it is stochastic, such that the crossover from climate to life dominance is thus a scaling (power law) not standard exponential (i.e., Markov process–type) transition. Stochastic models involve infinite dimensional probability spaces, they are therefore natural model types in systems with huge numbers of degrees of freedom. We believe that they are intrinsically more realistic than strongly nonlinear but deterministic chaos-type models (including those that are deterministic but are perturbed by noises). When the intermittency is strong, scaling stochastic models must be nonlinear (e.g., multifractal cascade processes), and this can easily be included in further model improvements—the Gaussian forcing (γ_1, γ_2, eq. 14) need only be replaced by a multifractal one. Here, intermittency is neglected, and linear stochastic equations with Gaussian white noise forcings are used (linear stochastic models can often be used even when the underlying dynamics are strongly nonlinear).

The other unusual FMEM characteristic is that it is a system of fractional differential equations. Unlike the familiar integer-ordered differential equations that typically have exponential impulse response functions (Green's functions), fractional equations typically have power law response functions and are natural ways to model scaling processes. These impulse response functions are physical models of bolide impacts and similar nearly instantaneous processes, and we discussed some implications.

The model is also highly parsimonious with two scaling exponents and a crossover time τ determined by the PBDB data as analyzed SL. These determined the basic scaling characteristics of the six series: T, E_+, D, E_- ($= O - E$), O, E. In addition, the model had two correlations that were specified: those between T and E_+ and between T and D. From these, the other 13 pairwise correlations (out the 15 possible pairs of the six series) were implicitly determined and were compared with the data.

The fractional derivatives were of the Weyl type such that their Fourier transforms are simply power laws. Because the system is ultimately forced by two Gaussian white noises, only the second-order statistics (i.e., the spectra and correlation functions) are needed, and these are easily obtained: the basic solutions are ffRns that were recently introduced (Lovejoy 2022a). In future, more realistic intermittent (multifractal) forcings could be used instead of the Gaussian white noise. Beyond exhibiting the full solution to the equations with a full statistical characterization, we then implemented the model numerically, first verifying the model against the theoretically predicted behavior. By producing simulations at 1 Myr resolution, we are able to resample the output at the same irregular sampling times as the PBDB. The statistical characteristics of the results (the six scaling curves showing the fluctuations as functions of timescale) plus the 15 pairwise correlations as functions of timescale are all quite close to the data, and in several cases, the agreement could be clearly attributed to the limitations, biases, and so on of the data. In particular, this was the case of the DE_+, DO, DE, E_+E_-, E_-O, OE correlations that are much closer to the data following the irregular sampling than with the original model outputs uniformly sampled at 1 Myr resolution.

Given the model's simplicity, it thus is remarkably realistic. This is fortunate, as until higher-resolution (global-scale) time series become available (e.g., Fan et al. 2020), more complex models may not be warranted. In any case, the model is able to help explain some subtle points about the interaction of different correlated series that are also strongly self-correlated over wide ranges of timescales, and this with quantitatively and qualitatively different scaling behaviors ("wandering" vs. "canceling"/self-stabilizing).

Acknowledgments. We would like to thank the editors L. H. Liow and M. J. Hopkins and the reviewers, C. Marshall and three anonymous reviewers, for many helpful suggestions that significantly strengthened and clarified the article. The research project was supported by grant S-MIP-21-9 "The Role of Spatial Structuring in Major Transitions in Macroevolution" and National Science and Engineering Research Council grant 222858 "Scaling and Multifractals in Geosystems." This paper is Paleobiology Database official publication no. 470.

Competing Interests. The authors declare no competing interests.

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

Literature Cited

Aberhan, M., and W. Kiessling. 2015. Persistent ecological shifts in marine molluscan assemblages across the end-Cretaceous mass extinction. *Proceedings of the National Academy of Sciences USA* **112**:7207–7212.

Alroy, J. 2010a. Geographical, environmental and intrinsic biotic controls on Phanerozoic marine diversification. *Palaeontology* **53**:1211–1235.

Alroy, J. 2010b. The shifting balance of diversity among major marine animal groups. *Science* **329**:1191–1194.

Alroy, J. 2015. A more precise speciation and extinction rate estimator. *Paleobiology* **41**:633–639.

Alroy, J., C. Marshall, R. Bambach, K. Bezusko, M. Foote, F. Fürsich, T. A. Hansen, *et al.* 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. *Proceedings of the National Academy of Sciences USA* **98**:6261–6266.

Alroy, J., M. Aberhan, D. J. Bottjer, M. Foote, F. T. Fürsich, P. J. Harries, A. J. Hendy, S. M. Holland, L. C. Ivany, and W. Kiessling. 2008. Phanerozoic trends in the global diversity of marine invertebrates. *Science* **321**:97–100.

Alvarez, L. W., W. Alvarez, F. Asaro, and H. V. Michel. 1980. Extraterrestrial cause for the Cretaceous–Tertiary extinction. *Science* **208**:1095–1108.

Baleanu, D., K. Diethelm, E. Scalas, and J. J. Trujillo. 2012. *Fractional calculus: models and numerical methods*. World Scientific, Singapore.

Barnosky, A. D. 2001. Distinguishing the effects of the Red Queen and Court Jester on Miocene mammal evolution in the northern Rocky Mountains. *Journal of Vertebrate Paleontology* **21**:172–185.

Bartoszek, K., S. Glémin, I. Kaj, and M. Lascoux. 2017. Using the Ornstein–Uhlenbeck process to model the evolution of interacting populations. *Journal of Theoretical Biology* **429**:35–45.

Beaufort, L., C. T. Bolton, A.-C. Sarr, B. Suchéras-Marx, Y. Rosenthal, Y. Donnadieu, N. Barbarin, S. Bova, P. Cornuault, and Y. Gally. 2022. Cyclic evolution of phytoplankton forced by changes in tropical seasonality. *Nature* **601**:79–84.

Benton, M. J. 1995. Diversification and extinction in the history of life. *Science* **268**:52–58.

Benton, M. J. 2009. The Red Queen and the Court Jester: species diversity and the role of biotic and abiotic factors through time. *Science* **323**:728–732.

Boyle, J. T., H. D. Sheets, S.-Y. Wu, D. Goldman, M. J. Melchin, R. A. Cooper, P. M. Sadler, and C. E. Mitchell. 2013. A re-examination of the contributions of biofacies and geographic range to extinction risk in Ordovician graptolites. *GFF* **136**:38–41.

Brayard, A., G. Escarguel, H. Bucher, C. Monnet, T. Brühwiler, N. Goudemand, T. Galfetti, and J. Guex. 2009. Good genes and good luck: ammonoid diversity and the end-Permian mass extinction. *Science* **325**:1118–1121.

Caraballoa, T., R. Colucci, and X. Han. 2016. Non-autonomous dynamics of a semi-Kolmogorov population model with periodic forcing. *Nonlinear Analysis: Real World Applications* **31**:661–680.

Carrillo, J. D., S. Faurby, D. Silvestro, A. Zizka, C. Jaramillo, C. D. Bacon, and A. Antonelli. 2020. Disproportionate extinction of South American mammals drove the asymmetry of the Great American Biotic Interchange. *Proceedings of the National Academy of Sciences USA* **117**:26281–26287.

Casey, M. M., E. E. Saupe, and B. S. Lieberman. 2021. The effects of geographic range size and abundance on extinction during a time of "sluggish" evolution. *Paleobiology* **47**:54–67.

Cornette, J. L., and B. S. Lieberman. 2004. Random walks in the history of life. *Proceedings of the National Academy of Sciences USA* **101**:187–191.

Crampton, J. S., R. A. Cooper, P. M. Sadler, and M. Foote. 2016. Greenhouse–icehouse transition in the Late Ordovician marks a step change in extinction regime in the marine plankton. *Proceedings of the National Academy of Sciences USA* **113**:1498–1503.

Crampton, J. S., S. R. Meyers, R. A. Cooper, P. M. Sadler, M. Foote, and D. Harte. 2018. Pacing of Paleozoic macroevolutionary rates by Milankovitch grand cycles. *Proceedings of the National Academy of Sciences USA* **115**:5686–5691.

Cuthill, J. F. H., N. Guttenberg, and G. E. Budd. 2020. Impacts of speciation and extinction measured by an evolutionary decay clock. *Nature* **588**:636–641.

Del Rio Amador, L., and Lovejoy, S. 2019. Predicting the global temperature with the Stochastic Seasonal to Interannual Prediction System (StocSIPS) *Climate Dynamics* **53**:4373–4411.

Del Rio Amador, L., and S. Lovejoy. 2021. Using regional scaling for temperature forecasts with the Stochastic Seasonal to Interannual Prediction System (StocSIPS). *Climate Dynamics* **57**:727–756.

During, M. A., J. Smit, D. F. A. E. Voeten, C. Berruyer, P. Tafforeau, S. Sanchez, K. H. Stein, S. J. Verdegaal-Warmerdam, and J. H. Van Der Lubbe. 2022. The Mesozoic terminated in boreal spring. *Nature* **603**:91–94.

Eaton, J. G., J. I. Kirkland, J. H. Hutchison, R. Denton, R. C. O'Neill, and J. M. Parrish. 1997. Nonmarine extinction across the Cenomanian–Turonian boundary, southwestern Utah, with a comparison to the Cretaceous–Tertiary extinction event. *Geological Society of America Bulletin* **109**:560–567.

Eichenseer, K., U. Balthasar, C. W. Smart, J. Stander, K. A. Haaga, and W. Kiessling. 2019. Jurassic shift from abiotic to biotic control on marine ecological success. *Nature Geoscience* **12**:638–642.

Eldredge, N. 1985. *Unfinished synthesis: biological hierarchies and modern evolutionary thought*. Oxford University Press, New York.

Eldredge, N. 1989. *Macroevolutionary dynamics: species, niches and adaptive peaks*. McGraw-Hill, New York.

Eldredge, N. 1996. Hierarchies in macroevolution. Pp. 42–61 *in* D. Jablonski, D. H. Erwin, and J. H. Lipps, eds. *Evolutionary paleobiology*. University of Chicago Press, Chicago.

Eldredge, N. 2003. The sloshing bucket: how the physical realm controls evolution. Pp. 3–32 *in* J. P. Crutchfield and P. Schuster, eds. *Evolutionary dynamics: exploring the interplay of selection, accident, neutrality, and function*. SFI Studies in the Sciences of Complexity Series. Oxford University Press, New York.

Erwin, D. H. 2011. Evolutionary uniformitarianism. *Developmental Biology* **357**:27–34.

Erwin, D. H. 2012. Novelties that change carrying capacity. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **318**:460–465.

Erwin, D. H. 2016. *Wonderful Life* revisited: chance and contingency in the Ediacaran–Cambrian radiation. *Chance in Evolution* 279–298.

Fan, J.-X., S.-Z. Shen, D. H. Erwin, P. M. Sadler, N. MacLeod, Q.-M. Cheng, X.-D. Hou, *et al.* 2020. A high-resolution summary of Cambrian to Early Triassic marine invertebrate biodiversity. *Science* **367**:272–277.

Feller, W. 1971. *An introduction to probability theory and its applications*, Vol. 2. Wiley, New York.

Ferretti, A., A. M. Bancroft, and J. E. Repetski. 2020. GECkO: Global Events Impacting Conodont Evolution. *Palaeogeography, Palaeoclimatology, Palaeoecology*, Special Issue **549**:109677.

Fields, B. D., A. L. Melott, J. Ellis, A. F. Ertel, B. J. Fry, B. S. Lieberman, Z. Liu, J. A. Miller, and B. C. Thomas. 2020. Supernova triggers for end-Devonian extinctions. *Proceedings of the National Academy of Sciences USA* **117**:21008–21010.

Finnegan, S., J. L. Payne, and S. C. Wang. 2008. The Red Queen revisited: reevaluating the age selectivity of Phanerozoic marine genus extinctions. *Paleobiology* **34**:318–341.

Foote, M. 1994. Temporal variation in extinction risk and temporal scaling of extinction metrics. *Paleobiology* **20**:424–444.

Foote, M. 2000. Origination and extinction components of taxonomic diversity: general problems. *Paleobiology* **26**:74–102.

Foote, M. 2005. Pulsed origination and extinction in the marine realm. *Paleobiology* **40**:6–20.

Gingerich, P. D. 1993. Quantification and comparison of evolutionary rates. *American Journal of Science* **293**:453–478.

Gingerich, P. D. 2001. Rates of evolution on the time scale of the evolutionary process. *Genetica* **112–113**:127–144.

Gingerich, P. D. 2006. Environment and evolution through the Paleocene–Eocene thermal maximum. *Trends in Ecology and Evolution* **21**:246–253.

Gingerich, P. D. 2009. Rates of evolution. *Annual Review of Ecology, Evolution, and Systematics* **40**:657–675.

Gould, S. J. 1990. *Wonderful life: the Burgess Shale and the nature of history*. Norton, New York.

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

Gould, S. J. 2001. Contingency. Pp. 195–198 *in* D. E. G. Briggs and P. R. Crowther, eds. *Palaeobiology II*. Blackwell Science, Malden, Mass.

Gould, S. J. 2002. *The structure of evolutionary theory*. Harvard University Press, Cambridge, Mass.

Gould, S. J., D. M. Raup, J. J. Sepkoski Jr, T. J. Schopf, and D. S. Simberloff. 1977. The shape of evolution: a comparison of real and random clades. *Paleobiology* **3**:23–40.

Gripenberg, G., and I. Norros. 1996. On the prediction of fractional Brownian motion. *Journal of Applied Probability* **33**:400–410.

Grossman, E.L., and M. M. Joachimski. 2022. Ocean temperatures through the Phanerozoic reassessed. *Scientific Reports* **12**:1–13.

Halliday, T. J. D., P. A. Holroyd, E. Gheerbrant, G. V. R. Prasad, A. Scanferla, R. M. D. Beck, D. W. Krause, and A. Goswami. 2020. Leaving Gondwana: the changing position of the Indian subcontinent in the Global Faunal Network. Pp. 227–249 *in* G. V. R. Prasad and R. Patnaik, eds. *Biological Consequences of Plate Tectonics: New Perspectives on Post-Gondwana Break-up—A Tribute to Ashok Sahni*. Springer, Cham, Switzerland.

Hannisdal, B., and S. E. Peters. 2011. Phanerozoic Earth system evolution and marine biodiversity. *Science* **334**:1121–1124.

Harmon, L. J., M. W. Pennell, L. F. Henao-Diaz, J. Rolland, B. N. Sipley, and J. C. Uyeda. 2021. Causes and consequences of apparent timescaling across all estimated evolutionary rates. *Annual Review of Ecology, Evolution, and Systematics* **52**:587–609.

Hébert, R., K. Rehfeld, and T. Laepple. 2021. Comparing estimation techniques for temporal scaling in palaeoclimate time series. *Nonlinear Processes in Geophysics* **28**:311–328.

Hébert, R., U. Herzschuh, and Laepple, T. 2022. Millennial-scale climate variability over land overprinted by ocean temperature fluctuations. *Nature Geoscience* **15**:899–905.

Hilfer, R., ed. 2000. *Applications of fractional calculus in physics*. World Scientific, Singapore.

Hoffman, A. 1987. Neutral model of taxonomic diversification in the Phanerozoic: a methodological discussion. Pp. 133–146 *in* M. H. Nitecki and A. Hoffman, eds. *Neutral models in biology*. Oxford University Press, New York.

Huisman, J., and F. J. Weissing. 1999. Biodiversity of plankton by species oscillations and chaos. *Nature* **402**:407–410.

Jablonski, D. 1986. Background and mass extinctions: the alternation of macroevolutionary regimes. *Science* **231**:129–133.

Jablonski, D. 2008. Biotic interactions and macroevolution: extensions and mismatches across scales and levels. *Evolution* **62**:715–739.

Jablonski, D., K. Roy, and J. W. Valentine. 2006. Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. *Science* **314**:102–106.

Jernvall, J., and M. Fortelius. 2002. Common mammals drive the evolutionary increase of hypsodonty in the Neogene. *Nature* **417**:538–540.

Khabbazian, M., R. Kriebel, K. Rohe, and C. Ane. 2016. Fast and accurate detection of evolutionary shifts in Ornstein–Uhlenbeck models. *Methods in Ecology and Evolution* **7**:811–824.

Kiessling, W., C. Simpson, and M. Foote. 2010. Reefs as cradles of evolution and sources of biodiversity in the Phanerozoic. *Science* **327**:196–198.

Klafter, J., S. Lim, and R. Metzler, eds. 2012. *Fractional dynamics: recent advances*. World Scientific, Singapore.

Kocsis, A. T., C. J. Reddin, J. Alroy, and W. Kiessling. 2019. The R package divDyn for quantifying diversity dynamics using fossil sampling data. *Methods in Ecology and Evolution* **10**:735–743.

Krug, A. Z., and D. Jablonski. 2012. Long-term origination rates are reset only at mass extinctions. *Geology* **40**:731–734.

Krug, A. Z., D. Jablonski, and J. W. Valentine. 2009. Signature of the end-Cretaceous mass extinction in the modern biota. *Science* **323**:767–771.

Lenton, T. M., T. W. Dahl, S. J. Daines, B. J. W. Mills, K. Ozaki, M. R. Saltzman, and P. Porada. 2016. Earliest land plants created modern levels of atmospheric oxygen. *Proceedings of the National Academy of Sciences USA* **113**:9704–9709.

Lidgard, S., E. Di Martino, K. Zágoršek, and L. H. Liow. 2021. When fossil clades "compete": local dominance, global diversification dynamics and causation. *Proceedings of the Royal Society of London B* **288**:20211632.

Lieberman, B. S. 2003. Unifying theory and methodology in biogeography. Pp. 1–25 *in* R. J. Macintyre and M. T. Clegg, eds. *Unifying theory and methodology in biogeography*. Evolutionary Biology, Vol. 33. Springer, Boston, Mass.

Lieberman, B. S., and N. Eldredge. 1996. Trilobite biogeography in the Middle Devonian: geological processes and analytical methods. *Paleobiology* **22**:66–79.

Lieberman, B. S., W. Miller III, and N. Eldredge. 2007. Paleontological patterns, macroecological dynamics and the evolutionary process. *Evolutionary Biology* **34**:28–48.

Liow, L. H., T. Reitan, and P. G. Harnik. 2015. Ecological interactions on macroevolutionary time scales: clams and brachiopods are more than ships that pass in the night. *Ecology Letters* **18**:1030–1039.

Liow, L. H., J. Uyeda, and G. Hunt. 2022. Cross-disciplinary information for understanding macroevolution. *Trends in Ecology and Evolution* **38**:250–260.

Lomb, N. R. 1976. Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science* **39**:447–462.

Lovejoy, S., and D. Schertzer. 2013. *The weather and climate: emergent laws and multifractal cascades*. Cambridge University Press, New York.

Lovejoy, S. 2015. A voyage through scales, a missing quadrillion and why the climate is not what you expect. *Climate Dynamics* **44**:3187–3210.

Lovejoy, S. 2019. *Weather, macroweather and climate: our random yet predictable atmosphere*. Oxford University Press, New York.

Lovejoy, S. 2022a. Fractional relaxation noises, motions and the fractional energy balance equation. *Nonlinear Processes in Geophysics* **29**:93–121.

Lovejoy, S. 2022b. The future of climate modelling: weather details, macroweather stochastics—or both? *Meteorology* **1**:414–449.

Lovejoy, S. 2023. Scaling, dynamical regimes and stratification: How long does weather last? How big is a cloud? *Nonlinear Processes in Geophysics* **30**:311–374.

Lovejoy, S. and L. Del Rio Amador. 2023. CanStoc: a hybrid stochastic—GCM system for monthly, seasonal and interannual predictions. *Meteorology* **2**:509–529.

Lovejoy, S. 2013. What is climate? *EOS* **94**:1–2.

Lovejoy, S., D. Schertzer, and P. Ladoy. 1986. Fractal characterisation of inhomogeneous measuring networks, *Nature*, **319**, 43–44.

Lovejoy, S., W. J. C. Currie, Y. Tessier, M. Claeredeboudt, J. Roff, E. Bourget, and D. Schertzer. 2001. Universal multfractals and ocean patchiness phytoplankton, physical fields and coastal heterogeneity. *Journal of Plankton Research* **23**:117–141.

Lovejoy, S., L. Del Rio Amador, R. Hébert. 2015. The ScaLIng Macroweather Model (SLIMM): using scaling to forecast global scale macroweather from months to decades. *Earth System Dynamics* **6**:637–658.

Mandelbrot, B. B., and J. W. Van Ness. 1968. Fractional Brownian motions, fractional noises and applications. *SIAM Review* **10**:422–450.

Markov, A. V., and A. V. Korotayev. 2007. Phanerozoic marine biodiversity follows a hyperbolic trend. *Palaeoworld* **16**:311–318.

Marshall, L. G., S. D. Webb, J. J. Sepkoski, and D. M. Raup. 1982. Mammalian evolution and the great American interchange. *Science* **215**:1351–1357.

Mathes, G. H., J. Van Dijk, W. Kiessling, and M. J. Steinbauer. 2021 Extinction risk controlled by interaction of long-term and short-term climate change. *Nature Ecology and Evolution* **5**:304–310.

Mayhew, P. J., M. A. Bell, T. G. Benton, and A. J. Mcgowan. 2012. Biodiversity tracks temperature over time. *Proceedings of the National Academy of Sciences USA* **109**:15141–15145.

McInerney, F. A., and S. L. Wing. 2011. The Paleocene–Eocene thermal maximum: a perturbation of carbon cycle, climate, and biosphere with implications for the future. *Annual Review of Earth and Planetary Sciences* **39**:489–516.

Melott, A., and R. Bambach. 2014. Analysis of periodicity of extinction using the 2012 geological timescale. *Paleobiology* **40**:176–195.

Meyers, S. R., S. E. Siewert, B. S. Singer, B. B. Sageman, D. J. Condon, J. D. Obradovich, B. R. Jicha, and D. A. Sawyer. 2012. Intercalibration of radioisotopic and astrochronologic time scales for the Cenomanian–Turonian boundary interval, Western Interior Basin, USA. *Geology* **40**:7–10.

Miller, A. I., and M. Foote. 2003. Increased longevities of post-Paleozoic marine genera after mass extinctions. *Science* **302**:1030–1032.

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

Miller, A. I., and M. Foote. 2009. Epicontinental seas versus open-ocean settings: the kinetics of mass extinction and origination. *Science* **326**:1106–1109.

Miller, K. S., and B. Ross. 1993. *An introduction to the fractional calculus and fractional differential equations*. Wiley, New York.

Mitchell, E. G., S. Harris, C. G. Kenchington, P. Vixseboxse, L. Roberts, C. Clark, A. Dennis, A. G. Liu, and P. R. Wilby. 2019. The importance of neutral over niche processes in structuring Ediacaran early animal communities. *Ecology Letters* **22**:2028–2038.

Mitchell, J. M., 1976. An overview of climatic variability and its causal mechanisms. *Quaternary Research* **6**:481–493.

Nance, R. D. 2022. The supercontinent cycle and Earth's long-term climate. *Annals of the New York Academy of Sciences* **1515**(1):33–49.

Navascués, M. A., C. Pacurar, and V. Drakopoulos. 2022. Scale-free fractal interpolation, *Fractal and Fractional* **6**(10):602.

Nee, S. 2006. Birth-death models in macroevolution. *Annual Review of Ecology, Evolution, and Systematics* **37**:1–17.

Newman, M., and R. Palmer. 2003. *Modeling extinction*. Oxford University Press, Oxford.

Newman, M. E. 1997. A model of mass extinction. *Journal of Theoretical Biology* **189**:235–252.

Oldham, K. B., and J. Spanier. 1974. *The fractional calculus*. Reprint, 2006, Dover, Mineola, N.Y.

Owolabi, K. M., and A. Atangana. 2019. *Numerical methods for fractional differentiation*. Springer, Singapore.

Payne, J. L., and S. Finnegan. 2007. The effect of geographic range on extinction risk during background and mass extinction. *Proceedings of the National Academy of Sciences USA* **104**:10506–10511.

Payne, J. L., and N. A. Heim. 2020. Body size, sampling completeness, and extinction risk in the marine fossil record. *Paleobiology* **46**:23–40.

Peters, S. 2013. Sepkoski's Online Genus Database. University of Wisconsin–Madison. https://strata.geology.wisc.edu/jack.

Peters, S. E., and M. Foote. 2002. Determinants of extinction in the fossil record. *Nature* **416**:420–424.

Petras, I. 2011. *Fractional-order nonlinear systems: modeling, analysis and simulation*. Higher Education Press, Beijing; Springer-Verlag, Berlin.

Podlubny, I. 1999. *Fractional differential equations*. Academic Press, San Diego, Calif.

Procyk, R., S. Lovejoy, and R. Hébert. 2022. The fractional energy balance equation for climate projections through 2100. *Earth System Dynamics* **13**:81–107.

Raja, N. B., E. M. Dunne, A. Matiwane, T. Ming Khan, P. S. Nätscher, A. M. Ghilardi, and D. Chattopadhyay. 2022. Colonial history and global economics distort our understanding of deep-time biodiversity. *Nature Ecology and Evolution* **6**:145–154.

Raup, D. M. 1985. Mathematical models of cladogenesis. *Paleobiology* **11**:42–52.

Raup, D. M. 1991. A kill curve for Phanerozoic marine species. *Paleobiology* **17**:37–48.

Raup, D. M. 1992a. *Extinction: bad genes or bad luck?* Norton, New York.

Raup, D. M. 1992b. Large-body impact and extinction in the Phanerozoic. *Paleobiology* **18**:80–88.

Raup, D. M. 1994. The role of extinction in evolution. *Proceedings of the National Academy of Sciences USA* **91**:6758–6763.

Raup, D. M., and J. W. Valentine. 1983. Multiple origins of life. *Proceedings of the National Academy of Sciences USA* **80**:2981–2984.

Reitan, T., and L. H. Liow. 2017. An unknown phanerozoic driver of brachiopod extinction rates unveiled by multivariate linear stochastic differential equations. *Paleobiology* **43**:537–549.

Roopnarine, P. D. 2003. Analysis of rates of morphologic evolution. *Annual Review of Ecology, Evolution, and Systematics* **34**:605–632.

Sadler, P. M. 1981. Sediment accumulation rates and the completeness of stratigraphic sections. *Journal of Geology* **89**:569–584.

Saupe, E., H. Qiao, Y. Donnadieu, A. Farnsworth, A. Kennedy-Asser, J. Ladant, D. Lunt, A. Pohl, P. Valdes, and P. Finnegan. 2019. Extinction intensity during Ordovician and Cenozoic glaciations explained by cooling and palaeogeography. *Nature Geoscience* **13**:65–70.

Schopf, T. J. M. 1979. Evolving paleontological views on deterministic and stochastic approaches. *Paleobiology* **5**:337–352.

Sepkoski, J. J. 1984. A kinetic model of Phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinctions. *Paleobiology* **10**:246–267.

Sepkoski, J. J., Jr. 1996. Competition in macroevolution: the double wedge revisited. Pp. 211–255 *in* D. Jablonski, D. H. Erwin, and J. H. Lipps, eds. *Evolutionary paleobiology*. University of Chicago Press, Chicago.

Sepkoski, J. J., Jr. 2002. A compendium of fossil marine animal genera. *Bulletins of American Paleontology* **363**:1–560.

Song, H., P. B. Wignall, H. Song, X. Dai, and D. Chu. 2019. Seawater temperature and dissolved oxygen over the past 500 million years. *Journal of Earth Science* **30**:236–243.

Spiridonov, A., and S. Lovejoy. 2022. Life rather than climate influences diversity at scales greater than 40 million years. *Nature* **607**:307–312.

Spiridonov, A., A. Brazauskas, and S. Radzevičius. 2015. The role of temporal abundance structure and habitat preferences in the survival of conodonts during the mid-early Silurian Ireviken mass extinction event. *PLoS ONE* **10**: e0124146.

Spiridonov, A., A. Brazauskas, and S. Radzevičius. 2016. Dynamics of abundance of the mid- to late Pridoli conodonts from the eastern part of the Silurian Baltic Basin: multifractals, state shifts, and oscillations. *American Journal of Science* **316**:363–400.

Spiridonov, A., D. Kaminskas, A. Brazauskas, and S. Radzevičius. 2017a. Time hierarchical analysis of the conodont paleocommunities and environmental change before and during the onset of the lower Silurian Mulde bioevent—a preliminary report. *Global and Planetary Change* **157**:153–164.

Spiridonov, A., R. Stankevič, T. Gečas, T. Šilinskas, A. Brazauskas, T. Meidla, L. Ainsaar, P. Musteikis, and S. Radzevičius. 2017b. Integrated record of Ludlow (Upper Silurian) oceanic geobioevents—coordination of changes in conodont, and brachiopod faunas, and stable isotopes. *Gondwana Research* **51**:272–288.

Spiridonov, A., J. Samsonė, A. Brazauskas, R. Stankevič, T. Meidla, L. Ainsaar, and S. Radzevičius. 2020a. Quantifying the community turnover of the uppermost Wenlock and Ludlow (Silurian) conodonts in the Baltic Basin. *Palaeogeography, Palaeoclimatology, Palaeoecology* **549**:109128.

Spiridonov, A., R. Stankevič, T. Gečas, A. Brazauskas, D. Kaminskas, P. Musteikis, T. Kaveckas, *et al.* 2020b. Ultra-high resolution multivariate record and multiscale causal analysis of Pridoli (late Silurian): implications for global stratigraphy, turnover events, and climate-biota interactions. *Gondwana Research* **86**:222–249.

Spiridonov, A., L. Balakauskas, and S. Lovejoy. 2022. Longitudinal expansion fitness of brachiopod genera controlled by the Wilson cycle. *Global and Planetary Change* **216**:103926.

Springford, A., G. M. Eadie, and D. J. Thomson. 2020. Improving the Lomb–Scargle Periodogram with the Thomson Multitaper. *Astronomical Journal* **159**(5). https://doi.org/10.3847/1538-3881/ab7fa1.

Stanley, S. M. 1979. *Macroevolution: pattern and process*. Freeman, San Francisco.

Thomson, A. M. 1982. Spectrum estimation and harmonic analysis. *Proceedings of the IEEE* **70**:1055–1096.

Tomašových, A., D. Jablonski, S. K. Berke, A. Z. Krug, and J. W. Valentine. 2015. Nonlinear thermal gradients shape broad-scale patterns in geographic range size and can reverse Rapoport's rule. *Global Ecology and Biogeography* **24**:157–167.

Vakulenko, S. A., I. Sudakov, and L. Mander. 2018 The influence of environmental forcing on biodiversity and extinction in a resource competition model. *Chaos* **28**:031101

Van Dam, J. A., H. A. Aziz, M. Á. Á. Sierra, F. J. Hilgen, L. W. Van Den Hoek Ostende, L. J. Lourens, P. Mein, A. J. Van Der Meulen, and P. Pelaez-Campomanes. 2006. Long-period astronomical forcing of mammal turnover. *Nature* **443**:687–691.

Van Valen, L. 1973. A new evolutionary law. *Evolutionary Theory* **1**:1–30.

Veizer, J., D. Ala, K. Azmy, P. Bruckschen, D. Buhl, F. Bruhn, G. A. Carden, A. Diener, S. Ebneth, and Y. Godderis. 1999. $^{87}Sr/^{86}Sr$, $\delta^{13}C$ and $\delta^{18}O$ evolution of Phanerozoic seawater. *Chemical Geology* **161**:59–88.

Veizer, J., Y. Godderis, and L. M. Francois. 2000. Evidence for decoupling of atmospheric CO_2 and global climate during the Phanerozoic eon. *Nature* **408**:698–701.

Venckutė-Aleksienė, A., A. Spiridonov, A. Garbaras, and S. Radzevičius. 2018 Integrated foraminifera and $\delta^{13}C$ stratigraphy across the

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

Cenomanian–Turonian event interval in the eastern Baltic (Lithuania). *Swiss Journal of Geosciences* **111**:341–352.
Vermeij, G. J. 1977. The Mesozoic marine revolution: evidence from snails, predators and grazers. *Paleobiology* **3**:245–258.
Vermeij, G. J. 2019. Power, competition, and the nature of history. *Paleobiology* **45**:517–530.
Vrba, E. S. 1985. Environment and evolution: alternative causes of the temporal distribution of evolutionary events. *South African Journal of Science* **81**:229–236.
Vrba, E. S. 1993. Turnover-pulses, the Red Queen, and related topics. *American Journal of Science* **293**:418–452.
Wagner, P. J., M. A. Kosnik, and S. Lidgard. 2006. Abundance distributions imply elevated complexity of post-Paleozoic marine ecosystems. *Science* **314**:1289–1292.
West, B. J., Bologna, M., and Grigolini, P. 2003. *Physics of fractal operators*. Springer, New York.
Wignall, P. B. 2015. *The worst of times: how life on Earth survived eighty million years of extinctions*. Princeton University Press, Princeton, N.J.
Ye, S., and S. E. Peters. 2023. Bedrock geological map predictions for Phanerozoic fossil occurrences. *Paleobiology* **49**:394–413.
Zachos, J., M. Pagani, L. Sloan, E. Thomas, and K. Billups. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292**:686–693.
Žliobaitė, I. 2022. Recommender systems for fossil community distribution modelling. *Methods in Ecology and Evolution* **13**:1690–1706.
Žliobaitė, I., M. Fortelius, and N. C. Stenseth. 2017. Reconciling taxon senescence with the Red Queen's hypothesis. *Nature* **552**:92–95.

Appendix 1. Numerical Simulations

Because the model is linear, the obvious simulation method is to use Fourier techniques. The main problem is that the small scales have singular power law filters that are not trivial to discretize, there may also be some long time (low-frequency) issues. A convenient way is to use techniques developed for simulating ffRn processes discussed in Lovejoy (2022). ffRn processes can be simulated by convolving Gaussian white noises with the ffRn Green's function $G_{\alpha,h}$ (eqs. 9, 10). A somewhat better numerical technique is to use the step-response Green's function (= $G_{\alpha+1,h}$; it is the smoother—and hence easier to handle—integral of $G_{\alpha,h}$), followed by a numerical differentiation.

Appendix 2. Sepkoski's Genus Compendium Data and Macroevolutionary Metrics

2.1. The Rates

To explore the robustness of the results based on the PBDB data, we used the Sepkoski database of stratigraphic ranges for marine animal genera (Sepkoski 2002, as implemented in an online tool; Peters 2013). We used the option of including all available fossilized Metazoan phyla. This yielded 19,713 resolved to substage-resolution stratigraphic-range records of the genera. Based on this record, the standing diversities and Foote's per-lineage per-substage extinction (q) and origination (p) rates (Foote 2000; Peters and Foote 2002) were calculated in the web implementation:

$$q = -\ln\left(\frac{N_{bt}}{N_{bl} + N_{bt}}\right) \quad (A2.1)$$

$$p = -\ln\left(\frac{N_{bt}}{N_{ft} + N_{bt}}\right) \quad (A2.2)$$

Here, N_{bt} is the number of taxa crossing bottom and the top of the interval, N_{bl} is the number of taxa crossing bottom of the interval and disappearing in the interval, and N_{ft} is the number of taxa first appearing in a given interval and crossing the upper stratigraphic boundary. Here and in the case of PBDB data, rates were not normalized for the durations of intervals, because most of the evolutionary turnover (as was shown in previous studies) is concentrated at or near the stratigraphic boundaries, which most often are defined by such events (Foote 2000, 2005; Payne and Heim 2020).

Averages of substage boundary ages were used for assigning the diversity, extinction, and origination rate ages. The analyzed substage record consists of 154 points, which started at 535 and ended at 0.9 Myr. The average resolution of the record is 3.5 Myr, being higher than the standard PBDB binning. According to the stratigraphic-range nature of Sepkoski's compendium data (vs. occurrence-based records in the PBDB), no sampling standardization was performed.

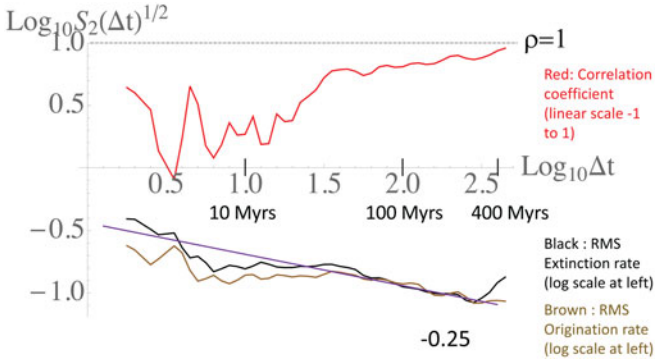


Figure A2.1. Scaling of extinction (black) and origination (brown) rates estimated from Sepkoski's genus-level compendium data. Red line shows time-scaling of correlations between extinctions and originations (for the correlations, the values are not logarithmic: 0 represents no correlation, 1 represents maximum correlation). This analysis is similar to that used in SL but on the Sepkoski stages database (for a detailed comaprison, compare with the bottom left and right of Fig. 9). The bottom curves are the root-mean-square (RMS) origination and extinction rate fluctuations with reference line $H = -0.25$ (as for the Paleobiology Database [PBDB]). The correlations (red, use the same numerical scale as the fluctuations but are linear, varying up to nearly a maximum of 1 indicated by the horizontal dashed line) behave similarly to the PBDB correlations: they are low until about 40 Myr, after which they are high. Approximately at the crossover scale of 40 Myr, macroevolutionary rates functionally track each other, which results in negative scaling of diversity at the longest timescales—therefore showing the same pattern as sample standardized PBDB data for Phanerozoic marine animals and using second-for-third macroevolutionary rates (SL).

2.2. Scaling of Rates in Sepkoski's Genus Compendium Data

The analysis of Foote's rates calculated from the Sepkoski's marine animal compendium data shows an essentially identical pattern of scaling to the second-for-third rates calculated from PBDB data on marine animal genera (Fig. A2.1). Despite the different kinds of data—taxic ranges versus subsamples of occurrences—and differences in the methodologies of estimations in extinction and origination rates, the scaling exponent in both cases was identical $H = -0.25$, which shows that stabilizing behavior and the timescales of relaxation are essentially identical in both cases. Moreover, the transition to synchronization of extinction and origination rates (high correlations) is also achieved at long timescales after the crossover timescale of 40 Myr. Therefore, despite different extinction metrics sometimes showing quite distinct patterns (Foote 1994, 2000; Alroy 2015), the patterns of scaling and the statistical association between extinction and origination rates is robust either to the kind of data chosen and the metrics used. Additionally, the traditional taxic range based Sepkoski's data shows the same qualitative timescale-dependent pattern of transition between incoherent dynamics at shorter timescales (low correlations of E and O) toward stable dynamics (high correlations between E and O) at timescales > 40 Myr.

Appendix 3. Comparing Spectral and Fluctuation Analyses

The FMEM model as well as its corresponding statistical properties are most easily expressed in the Fourier domain; it is therefore natural and tempting

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press

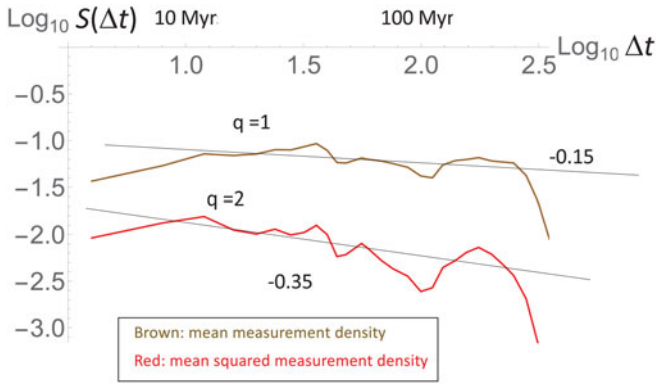


Figure A3.1. The Haar fluctuations of the measurement density for the Paleobiology Database (PBDB) used here. The first two moments (mean, mean square) are shown with reference lines indicating the corresonding scaling exponents (logarithmic slopes). Over the range of ≈10–300 Myr, the chronologies themselves (i.e., ρ(*t*)) do not have well-defined resolutions; they are scaling.

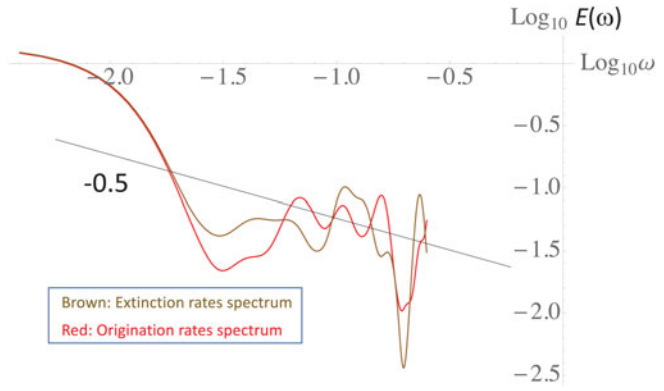


Figure A3.2. Extinction (brown) and origination rates (red) obtained by the Lomb-Scargle method using a Hanning window. The frequencies higher than $(4\ \mathrm{Myr})^{-1}$ were not shown, as they are at higher frequencies than the mean resolution.

to empirically evaluate it via spectral rather than fluctuation analysis. Indeed, had the PBDB been uniformly sampled, this would have been straightforward enough; standard discrete Fourier transforms could be used with standard windowing functions to reduce spectral leakage. This would have been quite adequate for determining scaling regimes and exponents. In certain cases, for example, where accurate determination of frequencies of spectral peaks are required, such as astrophysical applications—multitaper methods (MTM; Thomson 1982) could be used for further precision (as advocated, e.g., in Springford et al. [2020]).

But analyzing paleo-series poses unique problems for data analysis. Not only are the chronologies irregular, but there is increasing evidence that they are in fact themselves typically scaling! This means that the number of data points taken per time interval—the measurement density ρ(*t*)—is scaling, that is, the density fluctuations obey $\langle \Delta\rho(\Delta t)^q \rangle \propto \Delta t^{\xi(q)}$ (see Fig. A3.1 for the PBDB fluctuations). This is the chronological (temporal) counterpart of the spatial scaling of measuring stations reported in Lovejoy et al. (1986); see also Lovejoy and Schertzer (2013: chapter 3). Figure A3.1 also shows that the scaling of the measurement densities is non-Gaussian; this is because $\xi(1) \approx -0.15$ but $\xi(2) \approx -0.35$, such that $\xi(2) \neq \xi(1)$. A full characterization of scaling measurement densities will be discussed in another publication. For now, the key point is that its scaling implies potential biases in spectral exponents, and this is an area that has largely been neglected in recent work on spectral analysis.

To illustrate the difficulties of using spectral instead of fluctuation analysis, we used the Lomb-Scargle method (Lomb 1976) combined with a standard Hann window to reduce spectral leakage. Figure A3.2 shows the resulting spectra for the extinction and origination rates, and Figure A3.3 shows the diversity

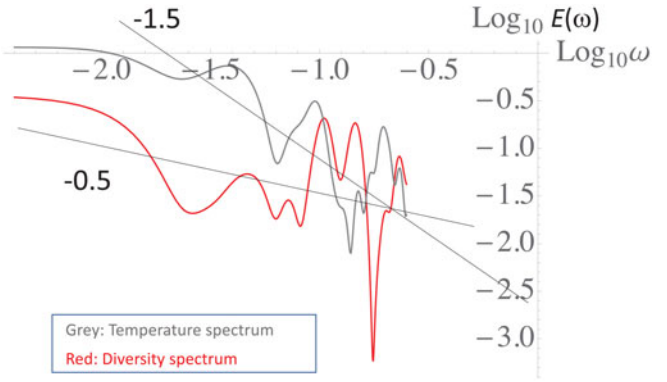


Figure A3.3. Diversity (red) and temperature (gray) spectra. The reference slopes are the theoretical slopes corresonding to $H = -0.25$ (spectral exponent $\beta = 0.5$), and $H = 0.25$ (spectral exponent $\beta = 1.5$).

and temperature spectra. These can be compared with the RMS Haar fluctuations (black curves in Fig. 9 or the bottom curves in Fig. A2.1). The spectrum is sufficiently noisy that if the straight reference lines—whose slopes were inferred from the real-space fluctuation analyses (using $\beta = 1 + 2H$)—had not been added, the scaling would not have been obvious. Indeed, it frequently happens that in order to properly interpret paleo-spectra, performing an initial Haar fluctuation analysis is useful—even necessary—for their interpretation.

Numerical experiments show that in spite of the windowing, the low frequencies are still very sensitive to spectral leakage and so are strongly biased. For example, Hébert et al. (2021) examined this issue numerically by determining the Lomb-Scargle spectrum of scaling processes and found that there were large spectral biases and that these increased as the spectral exponent β increased. This is unsurprising, because as β increases, the low frequencies that are most affected by leakage contain a larger and larger fraction of the total spectral variance. Yet Hébert et al. (2021) underestimated the problem: the numerically simulated data gaps that they used were confined to a fairly narrow range of timescales: they did not consider the effect of *scaling* chronologies—that is, sampling irregulariites over a wide range of scales.

Although spectral leakage is a known—and serious—limitation of Lomb-Scargle spectra, Springford et al. (2020) argue that the problem can be alleviated using MTM (Thomson 1982). The problem with this proposal is that there is no theoretical basis for using the MTM on irregular data: the "Slepian" tapers (window functions) are no longer orthogonal, such that averaging the spectrum of many windowed tapers can make spectral biases worse, not better, a finding that we will elaborate in another publication. An alternative suggested by Hébert et al. (2022) is to fill the gaps with linearly interpolated values and use a regular FFT combined with the MTM method on the resulting uniformly spaced series. Unfortunately, this interpolation method also leads to quite serious biases (especially at high frequencies). This is because the scaling exponent H is the maximum possible order of differentiation, and linear interpolations (i.e., of order 1) have $H = 1$, whereas the series in our case have $H \approx 0.25$ (temperature) and $H \approx -0.25$ (origination and extinction), such that the linear segments filling the gaps are far too smooth. The result is an uncontrolled mixing of data with different H values. A possible alternative might be to first estimate H using another method—for example, Haar fluctuations—and then to use fractal interpolation (e.g., Navascués et al. 2022) to do the interpolation using the correct exponent to fill in the gaps. Unfortunately, even this method may not be used if the measurement density is multifractal (i.e., the H exponent is only a partial characterization of a scaling process).

It should be stressed that real-space (fluctuation) analysis is advantageous, as the interpetation of the fluctuations is straightforward (recall the discussion of the missing quadrillion in the "The Drivers" section), and in addition, it is easy to handle missing data (irregular, including scaling chronologies) and also to determine correlations between series each having uneven chronologies—an advantage that we fully exploited here. In principle, these correlations could be studied by using cross-spectral analysis, but for single realizations (i.e., single series), this requires first breaking the series into segments, and this is more difficult when the chronologies are irregular.

https://doi.org/10.1017/pab.2023.38 Published online by Cambridge University Press