A simple model for skewed species-lifetime distributions

Yohsuke Murase¹, Takashi Shimada¹ and Nobuyasu Ito

Department of Applied Physics, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
E-mail: murase@serow.t.u-tokyo.ac.jp, shimada@ap.t.u-tokyo.ac.jp and ito@ap.t.u-tokyo.ac.jp

New Journal of Physics 12 (2010) 063021 (9pp)
Received 8 March 2010
Published 11 June 2010
Online at http://www.njp.org/
doi:10.1088/1367-2630/12/6/063021

Abstract. A simple model of a biological community assembly is studied. Communities are assembled by successive migrations and extinctions of species. In the model, species are interacting with each other. The intensity of the interaction between each pair of species is denoted by an interaction coefficient. At each time step, a new species is introduced to the system with randomly assigned interaction coefficients. If the sum of the coefficients, which we call the fitness of a species, is negative, the species goes extinct. The species-lifetime distribution is found to be well characterized by a stretched exponential function with an exponent close to 1/2. This profile agrees not only with more realistic population dynamics models but also with fossil records. We also find that an age-independent and inversely diversity-dependent mortality, which is confirmed in the simulation, is a key mechanism accounting for the distribution.

Contents

1. Introduction
2. Model
3. Result
4. Discussion
Acknowledgments
References

1 Authors to whom any correspondence should be addressed.
1. Introduction

The dynamics of ecosystems in evolutionary timescales has attracted much attention since it is a well-known and good example of understanding complex open systems, where various interactions among species may cause collective behavior [1, 2]. The development of theoretical models will help us to better interpret fossil records and further understand the mechanisms of evolutionary processes.

In this paper, we focus on the distribution of the lifetime of species. The lifetime of a species is the duration from its emergence to its extinction. The distribution of families (or genera) estimated from fossil records exhibits an interesting profile: it is neither a simple exponential nor a clear power law in the sense that it is convex on a log–log plot and concave on a semi-log plot [3–5] (also see figure 1). This implies that the distribution has a heavier tail than that of an exponential function but decays faster than a power law.

There are several theoretical explanations for lifetime distributions. The red-queen hypothesis states that environmental changes are so rapid and extreme that older species cannot obtain evolutionary advantage over a younger species; hence the mortality should be independent of time. This leads to an exponential function for the species-lifetime distribution: \( \exp(-t/\tau)/\tau \) [6]. Another group of simple theories supports power-law distributions. Assume that the fitness of each species (which can be the number of individuals, or any other measure of the distance from extinction) follows a neutral random walk in one-dimensional space. Then the lifetime is the first return time of a random walk, which obeys the \( t^{-3/2} \) power-law distribution for large \( t \). If the dynamics of a species is a critical Galton–Watson branching process [7] instead of a random walk, the return time distribution shows a power law with an exponent \(-2\). A more general stochastic process predicts power laws with exponential cutoffs [8]. In addition to these stochastic processes, the so-called ‘self-organized criticality’ (‘SOC’) models predict power laws [1, 9].

In addition to these simple models, population dynamics models are also proposed with the aim of bridging the differences in ecological and evolutionary descriptions. These include the tangled-nature models [10, 11] and their variants [12–16], the web-world model [17, 18], the scale-invariant model [19, 20] and others [21, 22]. All those models have population dynamics and rules for emergence and extinctions of species.

Recently, models with random migration rules were found to present the skewed species-lifetime distributions for various functional forms of population dynamics [16, 19]. In these models, migrations are modeled as additions of new species whose interaction coefficients are randomly assigned from certain distributions irrespective of resident species. The resulting species-lifetime distribution shows a curved line that lies somewhere between a power law and an exponential distribution: concave on a log–log plot and convex on a semi-log plot. It is worth noting that this skewed profile seems to be observed universally with respect to the functional forms of population dynamics and model parameters. Although these models have different numbers of species and various community structures, they all share similar profiles of species-lifetime distributions.

While skewed profiles are observed not only for several population dynamics models but also for fossil records, a mechanism that gives rise to such a skewed profile is still unclear. The universality of the profile suggests that the underlying mechanism should be captured by a simple model. For this reason, we propose a further simplified model that reproduces the skewed profile, and we then analyze it in the following.
Figure 1. (a, b) Lifetime distributions of families estimated from fossil records after [4]. The original data contain the list of presence of families (present, absent or unknown) in each geological era. Circles (a) and triangles (b) represent the estimates where unknown cases are assumed as present and absent, respectively. (c–e) Species-lifetime distributions for (c) the scale-invariant model [19], (d) the tangled-nature model A and (e) the tangled-nature model B [16] with random migration rules. The data for mathematical models are shifted in the horizontal and vertical directions so that they collapse onto the fossil data. (f, g) Fittings (f) by a $q$-exponential function (equation (1)) with $\tau = 33.0$ and $q = 1.23$ and (g) by a stretched-exponential function (equation (2)) with $\tau = 10.0$ and $\beta = 0.5$.

2. Model

A community in our model is organized as successive migrations and extinctions of species. Each species has a state variable called ‘fitness’, which depends on interactions of that species with other species. The interaction of species $i$ with species $j$ is described by the coefficient $a_{ij}$, and the fitness of the $i$th species is defined as the sum of all interactions the species $i$ has, given by $f_i = \sum_j a_{ij}$. The species $i$ can survive if $f_i$ is larger than or equal to zero. If the fitness is negative, that species becomes extinct and is eliminated from the system. Although the assumptions here are similar to those suggested in the model used in [23], our model accommodates changes in the number of species and suggests much simpler rules for driving evolution.

At each time step, a new species migrates into the system, which may cause successive extinctions. We take the following procedure for each time step. Firstly, a new species is added to the system. For each new interaction coefficient, we assign a nonzero random value from the Gaussian distribution with mean 0 and variance 1 with probability $c$ and 0 with probability $1 - c$. Coefficients for intra-species interactions $a_{ii}$ are zero. Secondly, after the migration, fitness of each species $f_i$ is calculated and a species with the minimum fitness is identified. If the minimum fitness is non-negative, no extinction happens at that time step. Otherwise, the species with the minimum fitness is eliminated from the system (extinction), and $f_i$ for all species are recalculated and another species with the minimum fitness is identified again. We repeat this
Figure 2. An example of the model dynamics. Nodes and arrows denote species and interactions, respectively. Numbers shown next to arrows denote interaction coefficients $a_{ij}$. (i) Before the migration of species $D$ (shown in red), species $A$, $B$ and $C$ coexist. (ii) When species $D$ immigrates into the community, species $B$ goes extinct. (iii) Then, another resident species $A$ goes extinct due to the extinction of species $B$. After extinctions of species $A$ and $B$, all species ($C$ and $D$) have positive fitness values. All of the above happens within one time step. In the next time step, another new species appears.

Fitness of a species changes only by migrations and extinctions of other species. In a migration event, a change in $f_i$ is caused by the $a_{ij}$ assigned between the species $i$ and the new entrant. Such changes follow the Gaussian distribution with mean 0. In an extinction event, a change in $f_i$ is caused by a loss of $a_{ij}$ between species $i$ and the species going extinct. The mean of lost $a_{ij}$ is expected to be positive since the sum of all interaction coefficients held by resident species must always be positive. In our simulation, the average of all interaction coefficients $a_{ij}$ of the emerging communities is confirmed to be slightly positive. An extinction of a certain species reduces the fitness of other species on average, and the dynamics of species’ fitness can no longer be described by a neutral random walk.

3. Result

Figure 3(a) shows a temporal dynamics of the number of species, $N$ for $c = 0.2$. It fluctuates in a finite range with its average 61.7. As shown in figure 3(b), the probability density function of $N$ peaks at around 50 and decays exponentially for large $N$. The power spectrum for this time series is calculated and found to be approximated by $1/f^2$ with saturation at low frequency, which indicates that the evolutionary process reproduced by our model is an Ornstein–Uhlenbeck process. We next examined the distribution of extinction sizes, which is the number of species that go extinct at a single time step. The distribution exhibits a simple exponential form, which is different from the power laws typically obtained from the SOC models and similar to the one obtained from the population dynamics models with random migrations [16, 20]. It is confirmed that the statistics discussed below does not change qualitatively against various values of $c$, although the characteristic number of species and the timescale change. Thus, we show the results only for $c = 0.2$ in the following.
Figure 3. (a) A typical temporal fluctuation of the number of species for $c = 0.2$. (b) A probability density function of the number of species. The semi-log plot is shown in the inset.

The species-lifetime distribution is shown in figure 4. One of the fitting functions proposed to account for this skewed profile is a $q$-exponential function [4],

$$p(t) \propto [1 - (1 - q)(t/\tau)]^{1/(1-q)},$$

where $q$ is an exponent to characterize the skewness and $\tau$ is a characteristic timescale at which $p(t)$ crosses over from the exponential function to the power law $t^{1/(1-q)}$. This function is naturally derived from a simple age-dependent mortality function. If the mortality function $m(t)$ is an inverse of a linear function of age, a $q$-exponential species-lifetime distribution is naturally derived. Another candidate of fitting is a stretched exponential function,

$$p(t) \propto \exp\left[-(t/\tau)^\beta\right],$$

where $\beta$ and $\tau$ are fitting parameters. This function is often used phenomenologically to describe relaxation dynamics in disordered materials such as glasses [24].

Among the three candidates shown in the figure, the stretched exponential function fits best, while the $q$-exponential function also fits reasonably well. The exponent $\beta$ for the stretched exponential function is approximately $1/2$. The fitting parameter $q$ for the $q$-exponential distribution is 1.16, which is smaller than the value estimated from fossil records, 1.23.

4. Discussion

While both stretched exponential and $q$-exponential functions exhibit a ‘skewed’ profile, the pictures these distributions stem from are different. A $q$-exponential function originates from a simple age-dependent mortality function [4]. In this picture, long-living species are expected to have advantageous features to survive. If a species survives for a longer duration, that species has a higher survival probability than other species surviving for a shorter duration.

On the other hand, a stretched exponential function is typically obtained by a linear combination of exponential functions with different decay times. If we assume that the dynamics of $N$ is characterized by an asymmetric random walk, a stretched exponential function with exponent $1/2$ is naturally obtained. Let us denote the probability of $N$ becoming $N + 1$ by $p$
and becoming \( N - 1 \) by \( 1 - p \). Here we assume \( p < 1/2 \) and \( N \) is bound to be positive. In addition, we assume that \( p \) is independent of \( N \) and that every species goes extinct with an equal probability irrespective of their age. Such an asymmetric random walk model yields an exponential decay for probability density functions of \( N \). Independence of \( p \) from \( N \) means that the extinction probability of each species living in a community of \( N \) species is proportional to \( 1/N \); hence the characteristic time to extinction \( \tau \) is proportional to \( N \). A species-lifetime distribution is then expressed as

\[
p(t) = \int_0^\infty \frac{\exp(-t/\tau)}{\tau} b \exp(-b \tau) \, d\tau
\]

\[
= 2b K_0(2\sqrt{bt})
\]

\[
\approx \sqrt{\pi} (bt)^{-1/4} \exp(-2\sqrt{bt}) \quad (t \gg 1),
\]

where \( K_\nu(x) \) is the modified Bessel function of the second kind and \( b \) is a coefficient that depends on \( p \). We call this set of assumptions in the above discussion the ‘modified red-queen hypothesis’, since it assumes an equal mortality for all the species in each moment.

Figure 4. Species-lifetime distribution on log–log scales and linear-log scales (inset). Fitting curves are also shown with a stretched exponential function, a \( q \)-exponential function and a simple exponential function. Fitting parameters and the reduced chi-square \( \sigma^2 \) (the degrees of freedom is 2988) for the stretched exponential function are \( \beta = 0.576 \), \( \tau = 36.8328 \) and \( \sigma^2 = 17.67 \). For the \( q \)-exponential function, we obtain \( q = 1.16 \) and \( \tau = 93.2 \) with larger deviation from the data: \( \sigma^2 = 123.0 \). For the simple exponential function, which has a smaller number of parameters and hence its degrees of freedom is 2989, we have the worst fit \( \sigma^2 = 430.3 \) with the fitting parameter \( \tau = 134.622 \). In the inset, a fitting with a simple exponential function with \( \tau = 350 \) is also shown as a guide to the eyes.
In order to investigate which picture is more reasonable for the proposed model, the mortality of each species is estimated from the simulations. Using the communities obtained from the migration-and-extinction process, the probability that each resident species goes extinct in the next time step (mortality) is calculated by adding a test immigrant many times (100 times in our simulation). The relationship between mortality and age of a species is shown in figure 5 for different numbers of species. For all N, mortalities show a sharp decrease at \( t \sim 0 \) and then converge to a constant value for \( t \gg 1 \). Mortality in the large \( t \) regime shows significant dependence on \( N \), but not so much on \( t \). And the mortality is approximately proportional to \( 1/N \) as shown in figure 5(b). These are consistent with the modified red-queen hypothesis: extinction risk of a species in the model is almost independent of its age and inversely related to the number of species of a resident community. Note also that this modified red-queen hypothesis is a good approximation, but not an exact reduction of the model. We can find a deviation in the fitting of simulation results with a huge number of samples, which appears as the large values of reduced chi-square.

The analytical treatment for the current model has not been obtained. The difficulty comes from the fact that obtaining a skewed distribution for species lifetime requires not only dynamics of each species’ fitness but also the fluctuations of \( N \). In a model with a constant \( N \) in which only one species with the minimal fitness is removed at each time step, the species lifetime shows an exponential distribution. The graph model we study in this paper is so simple that most of the ecological details are omitted, as in Hubbel’s neutral theory [25]. In spite of the absence of population dynamics, the model shares a similar species-lifetime distribution with several population dynamics models [16]. This fact makes us speculate that a further fundamental mechanism might be shared by the graph model and population dynamics models. We believe that models with a more complex population dynamics model will show a similar skewed profile.

In our model, new species are assumed to have completely random phenotype. This is also an idealized aspect of the model. Since the empirical data shown in figure 1 are collected at the level of families, we speculate that phenotypic correlation between resident and new species...
may be irrelevant. Effects of phenotypic correlation may alter the profile of species-lifetime distributions. The understanding of these effects on species-lifetime distributions is still open.

In summary, we proposed a simple model to explain a skewed species-lifetime distribution observed in the fossil record and population dynamics models. The skewed profile is fitted well by a stretched exponential function with exponent $1/2$. The mortality of a species is found to be independent of age and inversely dependent on the number of species. Therefore, the skewness found in the lifetime distributions is compatible with the modified red-queen hypothesis. We believe that such a skewed profile will also appear in other fields of study such as sociology or economics because features of this model are not necessarily biology specific. Similar lifetime distributions are already observed in other empirical data [26], while data showing simple exponential functions are also common [27].

Acknowledgments

We are grateful to P A Rikvold for helpful comments on the manuscript. This work was partially supported by award no. KUK-I1-005-04 from King Abdullah University of Science and Technology (KAUST) and Grant-in-Aid for Young Scientists (B) no. 21740284 to TS from the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

References

[1] Newman M E J and Palmer R G 2003 Modeling Extinction (Oxford: Oxford University Press)
[2] Drossel B 2001 Biological evolution and statistical physics Adv. Phys. 50 209–95
[3] Solé R V and Bascompte J 1996 Are critical phenomena relevant to large-scale evolution? Proc. R. Soc. B 263 161–8
[4] Shimada T, Yukawa S and Ito N 2003 Life-span of families in fossil data forms $q$-exponential distribution Int. J. Mod. Phys. C 14 1267–71
[5] Finnegan S, Payne J L and Wang S C 2008 The red queen revisited: reevaluating the age selectivity of phanerozoic marine genus extinctions Paleobiology 34 318–41
[6] van Valen L 1973 A new evolutionary law Evol. Theor. 1 1–30
[7] van Kampen N G 2007 Stochastic Processes in Physics and Chemistry (Amsterdam: North-Holland)
[8] Pigolotti S, Flammini A, Marsili M and Maritan A 2005 Species lifetime distribution for simple models of ecologies Proc. Natl Acad. Sci. USA 102 15747–51
[9] Bak P and Sneppen K 1993 Punctuated equilibrium and criticality in a simple model of evolution Phys. Rev. Lett. 71 4083–6
[10] Hall M, Christensen K, di Collobiano S A and Jensen H J 2002 Time-dependent extinction rate and species abundance in a tangled-nature model of biological evolution Phys. Rev. E 66 011904
[11] Christensen K, di Collobiano S A, Hall M and Jensen H J 2002 Tangled nature: a model of evolutionary ecology J. Theor. Biol. 216 73–84
[12] Rikvold P A and Zia R K P 2003 Punctuated equilibria and $1/f$ noise in a biological coevolution model with individual-based dynamics Phys. Rev. E 68 031913
[13] Zia R K P and Rikvold P A 2004 Fluctuations and correlations in an individual-based model of biological coevolution J. Phys. A: Math. Gen. 37 5135–55
[14] Sevim V and Rikvold P A 2005 Effects of correlated interactions in a biological coevolution model with individual-based dynamics J. Phys. A: Math. Gen. 38 9475–89
[15] Rikvold P A 2007 Self-optimization, community stability, and fluctuations in two individual-based models of biological coevolution J. Math. Biol. 55 653
[16] Murase Y, Shimada T, Ito N and Rikvold P A 2010 Random walk in genome space: a key ingredient of intermittent dynamics of community assembly on evolutionary time scales J. Theor. Biol. 264 663–72
[17] Caldarelli G, Higgs P G and McKane A J 1998 Modelling coevolution in multispecies communities J. Theor. Biol. 193 345–58
[18] Drossel B, Higgs P G and McKane A J 2001 The influence of predator–prey population dynamics on the long-term evolution of food web structure J. Theor. Biol. 208 91–107
[19] Shimada T, Yukawa S and Ito N 2002 Self-organization in an ecosystem Artif. Life Robot. 6 78
[20] Murase Y, Shimada T and Ito N 2009 Phase diagram and stability of ecosystems Artif. Life Robot. 13 460–3
[21] Chowdhury D, Stauffer D and Kunwar A 2003 Unification of small and large time scales for biological evolution: deviations from power law Phys. Rev. Lett. 90 068101
[22] Tokita K and Yasutomi A 2003 Emergence of complex and stable network in a model ecosystem with extinction and mutation Theor. Pop. Biol. 63 131–46
[23] Solé R V and Manrubia S C 1996 Extinction and self-organized criticality in a model of large-scale evolution Phys. Rev. E 54 R42–5
[24] Palmer R G, Stein D L, Abrahams E and Anderson P W 1984 Models of hierarchically constrained dynamics for glassy relaxation Phys. Rev. Lett. 53 958–61
[25] Hubbell S P 2001 The Unified Neutral Theory of Biodiversity and Biogeography (Princeton, NJ: Princeton University Press)
[26] Mizuno T and Takayasu M 2009 The statistical relationship between product life cycle and repeat purchase behavior in convenience stores Prog. Theor. Phys. Suppl. 179 71–9
[27] Fujiwara Y 2004 Zipf law in firms bankruptcy Physica A 337 219–30