

Simple Stochastic Models and Their Power-Law Type Behaviour

Matt J. Keeling

Zoology Department, Cambridge University, Downing Street, Cambridge CB2 3EJ,
United Kingdom

E-mail: matt@zoo.cam.ac.uk

Received June 20, 1999

A power-law relationship between the mean and variance of ecological time series has been shown to hold for a vast number of species. Here we examine the behaviour of single-species stochastic models and concentrate in particular on the mean–variance relationship as the carrying capacity becomes large. Single-species stochastic models can be written as Markov chains, and the long-term distribution of population sizes and hence power-law scaling can be found analytically. The various power-law scalings that arise have very different biological implications for the effects of stochasticity and the departure from the deterministic paradigm. Finally we extend our analysis to consider the complicating factors of spatial heterogeneity, nontrivial deterministic dynamics, and multispecies models. © 2000 Academic Press

1. INTRODUCTION

Ecologists are often interested in the time series of population levels in different locations and in particular with the fluctuations that are observed (Sutcliffe *et al.*, 1996; Earn *et al.*, 1998; Blasius *et al.*, 1999; Bjørnstad *et al.*, 1999). As such, Taylor's power law (Taylor 1961) gives a robust method of relating the variance to the mean population size. In this paper Taylor's power law is considered for a variety of single-species stochastic models, and I show that for large populations the power-law relationship can often be predicted analytically.

Consider a set of time series, x_t^i , for the number of individuals in population i at time t . For example, we may know the number of weekly cases of a disease in various towns and cities (Keeling and Grenfell, 1999), or we may have annual abundance data for a variety of distinct habitats (e.g., Sutcliffe *et al.*, 1996; Ranta *et al.*, 1997). For such a set of time series, it is interesting to consider if there exists a relationship between the mean M_i and the variance V_i :

$$M_i = \frac{1}{T} \sum_{t=1}^T x_t^i, \quad V_i = \frac{1}{T} \sum_{t=1}^T (x_t^i - M_i)^2.$$

It has been observed (Taylor, 1961; Taylor and Woiod, 1980; Anderson *et al.*, 1982; Hanski and Tiainen, 1989; Boag *et al.*, 1992; Keitt and Stanley, 1998) that for many species there is a power-law relationship between the mean and variance of population size,

$$V_i \propto M_i^\alpha. \quad (1)$$

For the vast majority of species, the power-law scaling parameter, α , is found to lie between one and two, with many species lying close to the extremes (Anderson *et al.*, 1982). In this paper we shall examine the behaviour of stochastic population models to allow us to interpret this ecological parameter α .

Throughout this work statistical artifacts which may arise from the consideration of short, noisy time series will be ignored (Perry and Woiod, 1992) and instead only the results from long-term computer simulations or exact numerical calculations will be considered. It is also assumed that the variation in the mean (M_i) between populations is due solely to differences in the carrying capacity (K_i) of each populations. This variation in K may of course be a reflection of other underlying ecological

processes, such as different levels of predation or the availability of food. Data from ecological time series could also be confounded by large temporal variations in the carrying capacity or other parameters, and hence great care must be used when interpreting the results of wild populations.

2. BASIC FORMULATION

Let us consider a single population of size $n \in \mathbb{N}$. We assume that only two types of event can occur, “births” at a rate B_n which increase the population size by one, or “deaths” at a rate D_n which decrease the population size by one. (Note that B and D may also include changes caused by immigration and emigration, respectively.) This formulation gives us a Markov process, as the rates of increase and decrease depend only on the current population size. It may often be conceptually easier to compare biological data to the per capita rates of increase and decrease, b_n and d_n , (cf. birth and death rates), where

$$b_n = \frac{B_n}{n} \quad \text{and} \quad d_n = \frac{D_n}{n} \quad \forall n > 0.$$

We can write a set of coupled ODEs (the master equations) for the distribution of population sizes, $P_n = \text{Prob}(\text{population size} = n)$,

$$\frac{dP_n}{dt} = B_{n-1}P_{n-1} + D_{n+1}P_{n+1} - (B_n + D_n)P_n.$$

So the fixed point solution must satisfy a *detailed balance* with transitions from P_{n-1} to P_n , balanced by transitions from P_n to P_{n-1} . Hence, the unique solution is given by

$$\begin{aligned} D_n P_n &= B_{n-1} P_{n-1} \\ \Rightarrow P_n &= P_0 \prod_{m=1}^n \frac{B_{m-1}}{D_m}, \end{aligned} \quad (2)$$

where P_0 is defined such that the distribution sums to one.

The results from this stochastic approach can be compared to standard deterministic models where the population size, $x \in \mathbb{R}^+$, is given by

$$\frac{dx}{dt} = f(x) = B(x) - D(x). \quad (3)$$

The functions $B(x)$ and $D(x)$ are taken as the natural continuous population versions of the rates B_n and D_n (see Appendix I). Throughout this work we shall consider the form of $f(x)$ that is implied by the birth and death rates and will concentrate on the situation where there exists a single, globally attracting fixed point, x^* , of Eq. (3).

We find that when the distribution P_n is bounded (i.e., $\sum_{n=1}^{\infty} \prod_{m=1}^n B_{m-1}/D_m < \infty$) it is stable so long as the following conditions are satisfied

$$\begin{aligned} \exists n^* \text{ s.t. } \forall n < n^* \quad B_n > D_n \quad \text{and} \quad \forall n > n^* \quad B_n < D_n \\ B_0 > 0, \quad D_0 = 0. \end{aligned} \quad (4)$$

The condition $B_0 > 0$ means that the zero state is not absorbing, whereas $D_0 = 0$ is used to prevent negative population sizes. It should be noted that in general we will also wish the distribution P_n to possess a finite mean and variance, in which case the conditions become more restrictive.

Using Eq. (2) we can compute numerically the population distribution and hence the mean and variance as the underlying carrying capacity is varied. Log-log plots of mean against variance can be used to obtain local values for the power-law gradient α . More precisely the local power law is given by

$$\alpha(K) = \frac{d \log(V)}{dK} \left(\frac{d \log(M)}{dK} \right)^{-1} = \frac{M}{V} \frac{dV}{dM}.$$

3. SIMPLE FORMS

We shall now consider what forms of B_n and D_n satisfy the imposed conditions (4). The simplest model is linear behaviour,

$$B_n = c + Rn, \quad D_n = Rn + \frac{cn}{K}$$

$$\text{or} \quad b_n = R + \frac{c}{n} \quad d_n = R + \frac{c}{K} \quad (5)$$

$$\Rightarrow \quad f(x) = c \left(1 - \frac{x}{K} \right).$$

This corresponds to a sink-like population, where the natural death rate of individuals ($R + \frac{c}{K}$) is greater than the per capita birth rate (R), but the population is maintained by imports (c) from some external source. We notice that the term R occurs in both the birth and the death rates and hence cancels out in the calculation of the

deterministic $f(x)$. R can therefore be used to increase the magnitude of the stochasticity relative to the deterministic dynamics. This illustrates one of the main conceptual problems with moving between stochastic and deterministic representations; two very different stochastic processes can possess the same underlying deterministic equations.

For this system, Eq. (2) leads to a negative binomial distribution of population sizes,

$$P_n = \frac{\left(\frac{c}{R} + n - 1\right)!}{n! \left(\frac{c}{R} - 1\right)!} \left(\frac{RK}{c}\right)^n \left(1 + \frac{RK}{c}\right)^{-(c/R) - n},$$

with mean $M = K$ and variance $V = \frac{R}{c}K^2 + K$ (Fig. 1a). Hence, a single power law does not hold for all K ; however, this system does show many of the features we observe for many models. First, the local power-law gradient, $\alpha(K)$, lies between 1 and 2. Second, as R and hence the relative magnitude of the stochasticity increases, the variance behaves more like the square of the mean and $\alpha(K)$ approaches 2. Finally, as the carrying capacity, K , becomes large α tends to an asymptotic value, which is 2 for this simple model.

A more complex form which displays richer dynamics is the logistic equation with imports:

$$\begin{aligned} B_n &= c + (R + 1)n & D_n &= Rn + \frac{n^2}{K} \\ \text{or } b_n &= R + 1 + \frac{c}{n} & d_n &= R + \frac{n}{K} \end{aligned} \quad (6)$$

$$\Rightarrow f(x) = c + x \left(1 - \frac{x}{K}\right).$$

This model again assumes a small amount of immigration (c), which has been included to prevent permanent extinctions, and a constant birth rate per individual. However, the per capita death rate is now considered to increase linearly with population size. The deterministic equations predict a globally attracting equilibrium just above the carrying capacity.

For a wide choice of the parameters c and R this logistic-based system demonstrates localised power-law scaling; in particular there are two regions of power-law behaviour separated by a marked drop in the variance (Fig. 1b). This drop corresponds to the change from a system where localised extinctions are common to one where extinctions are rare. The point where this phase

transition occurs can therefore be compared to the critical community size associated with many diseases (Bartlett, 1957; Keeling, 1997; Swinton *et al.*, 1998). Above the phase transition point, as the carrying capacity K becomes large, α asymptotes to 1.

The behaviour at this transition can be attributed to the anomalous nature of extinct populations, which are only affected by an occasional import. Figure 2 shows the distribution of population sizes, P_n , as K is increased and we pass through the phase transition point. When K is small, extinctions dominate (Fig. 2a), whereas when K is large the population is distributed close to the mean (Fig. 2c). It is this change from bimodality (where P_n has local maxima at $n = 0$ and close to $n = K$) to unimodality (where P_n is maximal close to $n = K$ and there are very few extinctions) that causes the observed drop in the variance. No such behaviour is observed for the simplest model (5), which is always unimodal and therefore does not display a rapid change in variance. Below the phase transition point, the local power law is noninteger (lying between one and two) and is found to be an increasing function of the stochasticity R . We speculate that this noninteger power-law behaviour is attributable to the disparity in time-scales between the rate of events for extinct and extant populations.

What causes the difference between these two models ((5) and (6)) and in particular the difference in the asymptotic behaviour as the carrying capacity (and hence the population size) becomes large? In the linear model (5) both the deterministic and the stochastic components scale like n , whereas for the logistic model (6) the deterministic component scales like n^2 but the stochastic part only scales like n . Therefore, it is not surprising that the linear model has a larger variance to mean ratio than the logistic model. We now extend this general argument to predict the power-law behaviour of a general class of models when the carrying capacity is large.

3.1. Analytical Calculation

We restrict our attention to the situation where the underlying deterministic dynamics have a single globally attracting fixed point, in particular

$$f(x) = \left(1 - \frac{x}{K}\right) g(x) \quad (7)$$

for some positive definite function g . However, this still leaves some ambiguity in the values of the birth and death rates; hence we shall set

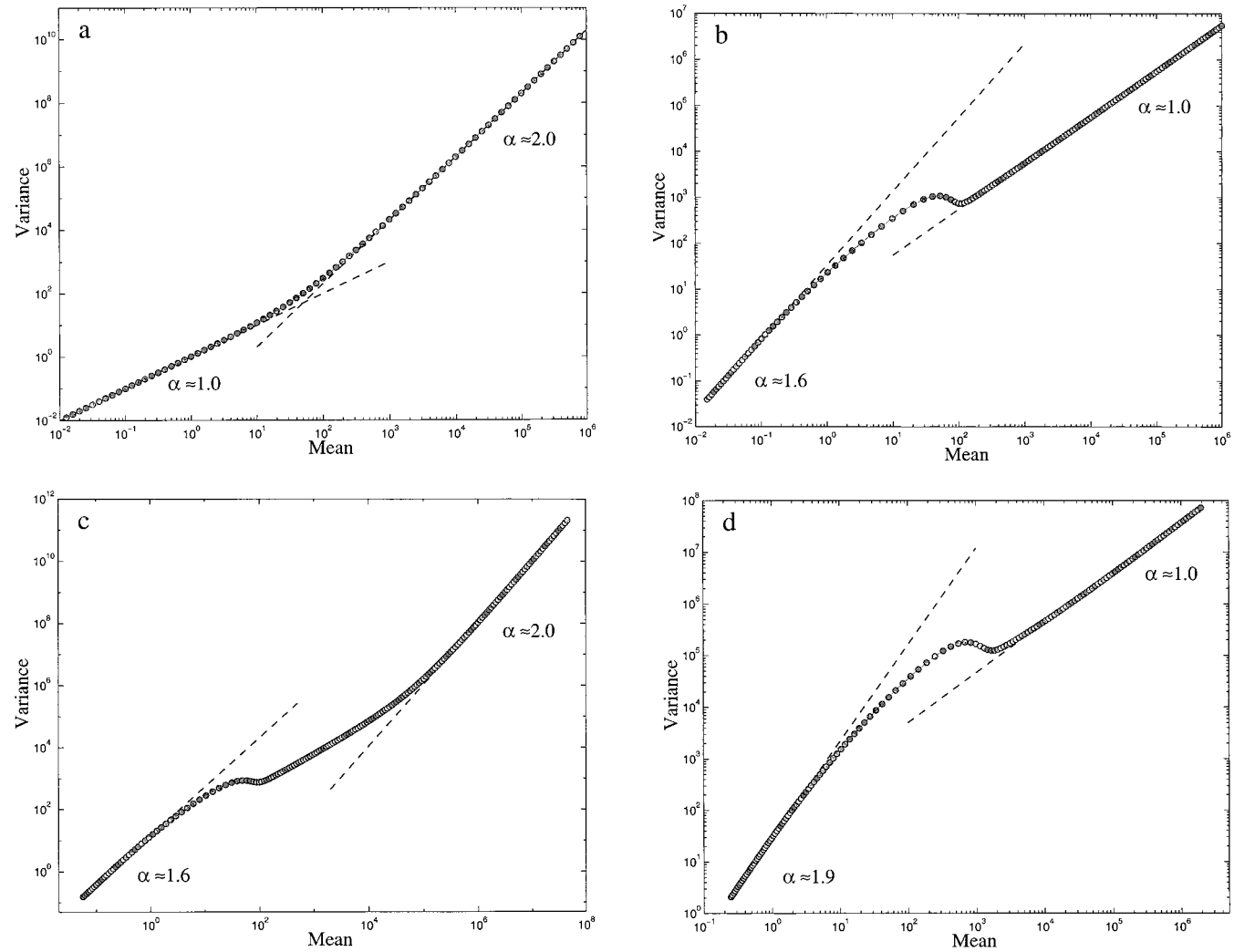


FIG. 1. (a) Results from the negative binomial distribution from the linear model: $B_n = c + Rn$, $D_n = Rn + cn/K$ ($c = 50$, $R = 1$). For large carrying capacity, K , the power law $\alpha \rightarrow 2$. (b) Results from the logistic equation; $B_n = c + (R + 1)n$, $D_n = Rn + n^2/K$ ($c = 0.025$, $R = 5$). For small carrying capacity, K , the gradient $\alpha \approx 1.6$; after the phase transition $\alpha \rightarrow 1$. (c) Results from the logistic equation with quadratic stochasticity; $B_n = c + (R + 1)n + rm^2$, $D_n = Rn + rm^2 + n^2/K$ ($c = 0.1$, $R = 5$, $r = 10^{-4}$). For small carrying capacity, $\alpha \approx 1.6$. After the phase transition $\alpha \rightarrow 2$; although there is initially a phase where $\alpha \approx 1$, this is due to the small value of r . (d) Results from the Ricker equation: $B_n = c + bn(b/d)^{-n/K}$, $D_n = dn$ ($c = 0.1$, $d = 5$, $b = 5.05$). For small carrying capacity $\alpha \approx 1.9$; after the phase transition $\alpha \rightarrow 1$.

$$\begin{aligned}
 B_n &= g(K; n) + h(K; n) \\
 D_n &= \frac{n}{K} g(K; n) + h(K; n)
 \end{aligned}
 \quad \forall n \in \mathbb{N}. \quad (8)$$

Here, the function $h \geq 0$ plays the same role as R in Eqs. (3.1) and (3.2), controlling the level of stochasticity. In what follows, we implicitly assume that g and h are positive-definite rational polynomials.

The behaviour of the mean and variance can be seen to depend on the asymptotic relationship between h and g . In all of what follows, we shall assume that for sufficiently large K there exists an $N \in \mathbb{N}$ ($1 \ll N \ll K$) such that.

- above N all functions (and in particular $\frac{h}{g}$) behave like their asymptotic forms,
- $P_N \ll P_K$ so that the behaviour below N can be ignored,

If we wish to consider the per capita rates of increase and decrease (b_n and d_n), then we can write

$$\begin{aligned}
 b_n &= \frac{1}{n} g(K; n) + \frac{1}{n} h(K; n) \\
 d_n &= \frac{1}{K} g(K; n) + \frac{1}{n} h(K; n),
 \end{aligned} \quad (9)$$

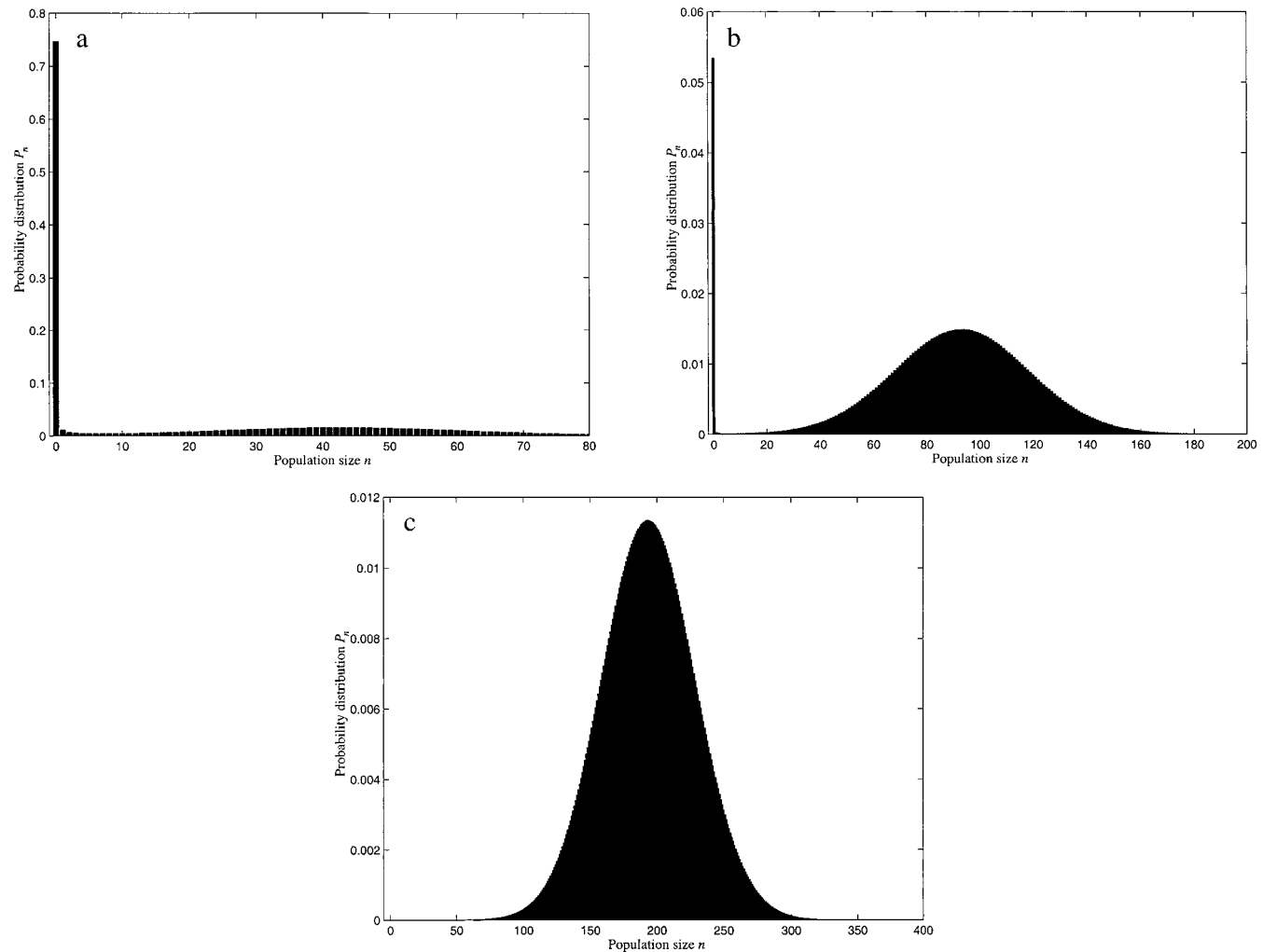


FIG. 2. The fixed point population distribution P_n for the logistic equation of Fig. 1b: $B_n = c + (R + 1)n$, $D_n = Rn + n^2/K$ ($c = 0.025$, $R = 5$). In graph (a), when $K = 50$, the population is extinct for over 70% of the time. In graph (b), when $K = 100$, much more of the time is spent near the carrying capacity. Finally by the time $K = 200$, in (c) extinctions are negligible.

in which case the main parameter of the system, $\frac{h(n)}{g(n)}$, can be related to the per capita changes by

$$\frac{h(n)}{g(n)} = \frac{d_n - b_n \frac{n}{K}}{b_n - d_n},$$

where $b_n - d_n$ is the expected per capita rate of increase of a population of size n .

We note that the imposed deterministic equation (7), together with the assumption that g and h are both positive, places additional restrictions on the birth and death rates. For population sizes below the carrying capacity ($n < K$),

$$d_n K > b_n n > d_n n,$$

whereas above the carrying capacity,

$$b_n n > d_n K > b_n K.$$

Other values of the birth and death rates can be accommodated by changing the deterministic form assumed in Eq. (7).

We now consider three scenarios in turn, with all mathematical results given in Appendix II. By considering the leading order terms for the distribution P_n , we are able to predict the power-law relationship between V and M as the carrying capacity becomes large; however to calculate the coefficient of proportionality requires far more detail about g and h , and inclusion of lower order terms.

When h is of the same order as g ($\frac{h(n)}{g(n)} \rightarrow H$, $H = O(1)$), then the amount of stochasticity is comparable with the

deterministic component of the dynamics. This is the behaviour seen in the logistic equation and the Ricker equation (Figs. 1b and 1d). For this scenario, α will asymptote to 1 as the carrying capacity becomes large, and the variance becomes proportional to the mean. Because the standard error ($\sqrt{V/M}$) decays to zero as K becomes large, the distribution of population sizes becomes relatively more tightly clustered and the system will behave more like its deterministic counterpart. This gives mathematical weight to the biological intuition that for many species stochasticity is only important for small populations. In the case where $H = 0$, the stochasticity is reduced such that the expected distributions are close to Poisson, and hence $V = M$.

When $h(n)$ is of the same order as $ng(n)$ ($\frac{h(n)}{g(n)} \rightarrow Hn + H_0$, $H = O(1)$, $H_0 = O(1)$), we find that α will asymptote to 2. This is observed in Figs. 1a and 1c, when the stochasticity is of the same order as the deterministic component of the dynamics. For this type of system, when the carrying capacity is large, the standard deviation scales linearly with the mean and the standard error is constant. This means that the shape of the population distribution is conserved under rescaling and therefore the individual nature of the population has no effect. With these models, the variance in population numbers always has an influence on the expected dynamics, but (once K is large) one can safely scale out the population size.

Finally, for any system where $h(n)$ increases like $n^2g(n)$ or faster, the distribution (2) will not be bounded. If we were to perform multiple simulations of such a system then some populations may go to zero, while others tend to infinity. Such a system would not be biologically realistic or viable; for large population sizes ($n \gg K$) we always expect the death rate to be larger than the birth rate so that population sizes cannot grow unbounded.

One's instinctive reaction is that generally only the first situation ($\frac{h}{g} \rightarrow H$) will be applicable to ecological systems. However, the following set of (not-unbelievable) birth and death rates gives rise to the entire spectrum of behaviour:

$$\begin{aligned} b_n &\rightarrow R + \frac{1}{n^p} & d_n &\rightarrow R + \frac{1}{Kn^{p-1}} \\ \Rightarrow & \frac{h(n)}{g(n)} & \rightarrow Rn^p. \end{aligned}$$

This shows that, in principle, all behaviours are possible. However, if we want the per capita birth rate (b_n) to decrease to zero with increasing n , and the death rate (d_n) to increase, then we must be in the first situation if conditions (7) and (8) hold. For the second situation to occur,

then from Eq. (9) we must have $b_n - d_n \rightarrow 0$ as K and n become large. Which of these two situations is most likely to occur depends on the organism in question and can only be assessed by a detailed study of individuals.

From the above analysis we can conclude that understanding the population changes at a deterministic level is often insufficient; we need to consider behaviour at the event level. The two models in Figs. 1b and 1c, have the same overall deterministic formulation; however, the stochastic versions possess strikingly different behaviour. Finally, a large population size is no guarantee that stochastic effects can be neglected; for the second class of models (e.g., Figs. 1a and 1c) the variance always affects the mean.

4. COMPLICATIONS

It should be realised that not all problems can be classed as one of the three situations given above. For some functional forms it is not possible to find an N which satisfies the necessary conditions; e.g., if

$$\frac{h}{g} = n + K,$$

this form does not show its asymptotic behaviour until $n \gg K$.

Another difficulty that arises is if the leading order terms in h or g are a function of K . Our analysis fails for functions such as

$$\frac{h}{g} = Kn \quad \text{or} \quad \frac{h}{g} = \frac{n}{K}$$

which may lead to asymptotic power laws other than one or two. For example, if $g \rightarrow 1$ and $h \rightarrow Kn$, we obtain the situation where the variance increases dramatically with the mean and $\alpha \rightarrow 3$. We believe that such scenarios may explain the presence of large local power laws in some ecological time series. Notice that with $\alpha = 3$, because the standard error increases with the mean, the likelihood of extinctions will presumably increase with population size which goes against normal intuition.

A further complication arises if the underlying deterministic dynamics have multiple stable states. This is contrary to condition (4) and essentially means that the distribution will be multimodal for any value of K . This in turn means that α will asymptote to two. However, we may find that transitions from the vicinity of one fixed point to the vicinity of another occur less and less

frequently as K increases. Hence, a stochastic simulation may spend all of its time close to just one of the fixed points, so that its dynamics will be highly dependent on the initial conditions.

For all these complications, however we can still use Eq. (2) to predict the mean and variance numerically. We shall now expand our consideration to three common situations which cannot be addressed using this basic equation. For each of these scenarios we offer some preliminary explorations, suggesting how the power law α will be affected.

4.1. Spatial Heterogeneity

So far we have assumed homogeneity, so that even very large populations are fully mixed. However, this is rarely the case. When spatial models are considered, we often find that at large distances, parts of the population can become decorrelated; this would lead to a power-law scaling of $\alpha \rightarrow 1$ as the habitat area and hence the population becomes large (Keeling *et al.* 1997).

Here we consider how the power law α varies as spatial heterogeneities are introduced. We let the correlation between two spatial locations decrease exponentially with the distance between them:

$$\text{Corr}(x(\underline{y}), x(\underline{z})) = \exp(-A \|\underline{y} - \underline{z}\|).$$

This exponential decrease with distance is a common property of many spatial models (Keeling *et al.*, 1997) and we would suppose that such behaviour often occurs in nature. By assuming that our habitat occupies a disk (D_r) of radius r , we can relate how the mean and variance change as r is increased. Obviously, we expect

$$M_r = \pi r^2 \mu, \quad (10)$$

where μ is the average density of individuals. The variance is now calculated as

$$V_r = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left[\int_{D_r} x(\underline{y}) d\underline{y} - M_r \right]^2 dt \quad (11)$$

$$\begin{aligned} &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \int_{D_r} \int_{D_r} x(\underline{y}) x(\underline{z}) d\underline{z} d\underline{y} dt - M_r^2 \\ &= V_0 \int_{D_r} \int_{D_r} \exp(-A \|\underline{y} - \underline{z}\|) d\underline{z} d\underline{y} \\ &= \frac{V_0}{A^4} \int_{D_R} \int_{D_R} \exp(-\|\underline{y} - \underline{z}\|) d\underline{z} d\underline{y}, \end{aligned}$$

$$\text{where } R = Ar. \quad (12)$$

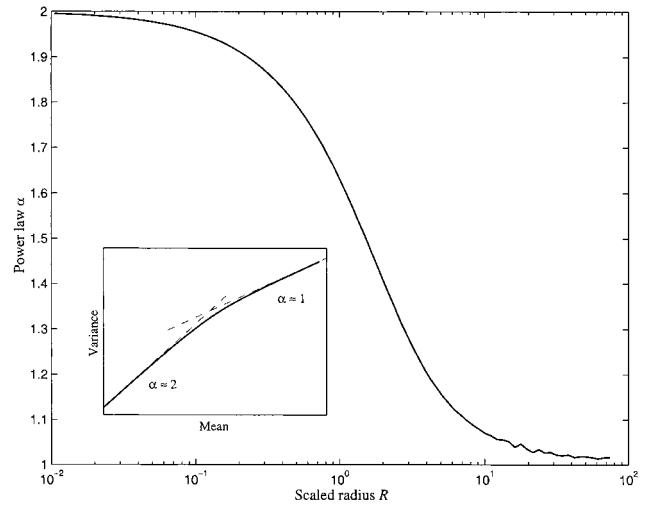


FIG. 3. The local power-law scaling α for heterogeneous systems, where the correlation between sites decays exponentially with distance. When the scaled radius R is small (and hence the mean population size is small) α asymptotes to 2, whereas for large R and large populations α asymptotes to 1. The inset graph shows the mean–variance plot on a log–log scale, clearly showing the change in power law.

It is clear from Eqs. (10) and (12) that when R is small, such that the correlation is approximately one, $V_r \propto M_r^2$; however it can also be seen from the equations that when $R \rightarrow \infty$, $V_r \propto M_r$. This spatial model should be contrasted with the assumptions of the previous section. Here we suppose that increases in average population size are due to increases in habitat area, with the density of individuals and hence the effective environment each individual observes remaining constant.

Figure 3 shows how the local power-law scaling α varies with the scaled habitat radius R . For large A , such that the correlations decay quickly, the $\alpha = 1$ limit is reached for smaller r and therefore smaller population sizes. The general power-law behaviour is clearly reminiscent of the scaling observed in Section 2. It may therefore be difficult to predict whether power-law scaling originates from stochastic processes at the population level, spatial heterogeneity, or a mixture of the two; although it would appear that spatial heterogeneity alone is unable to produce the jump in the variance that is observed in the homogeneous models.

4.2. Multi-species Models

The previous sections have concentrated on single species models; these have the distinct advantage of possessing an iterative method of calculation (Eq. (2)). For multi-species models such a simple process does not exist; instead we must rely on more complex techniques. In what follows we shall restrict our attention to two

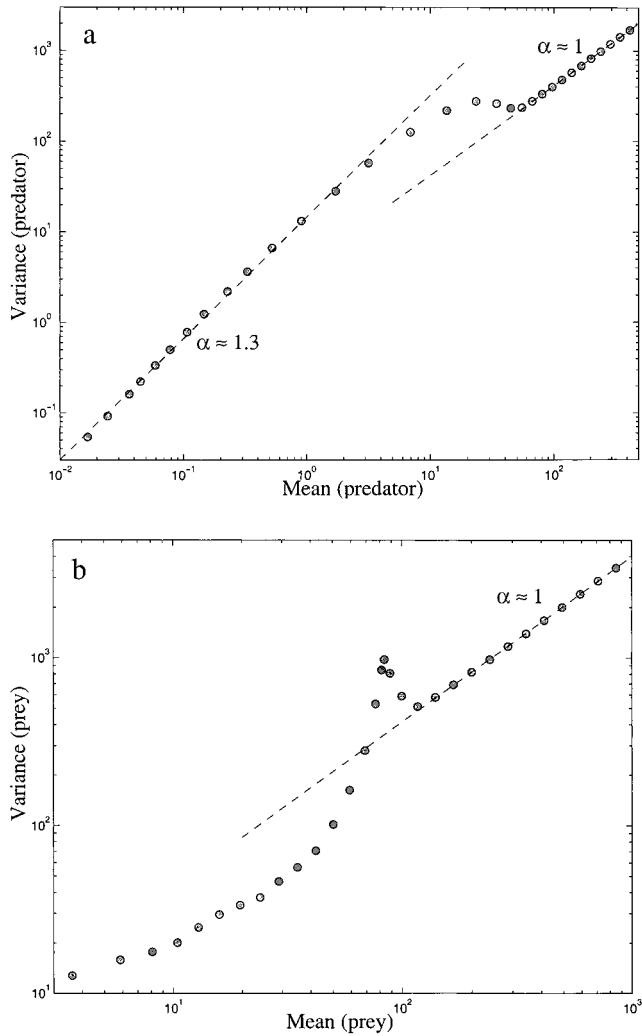


FIG. 4. Mean-variance results from the stochastic Lotka–Volterra predator–prey system (13), ($C^1=0.01$ and $C^2=0.005$). Each point represents the results from stochastic (Monte Carlo) simulations, where the mean and variance were measured over one million time steps. The predator (a) clearly shows two regimes of power-law scaling; when the mean is small $\alpha \approx 1.34$, whereas for larger population sizes $\alpha \rightarrow 1$. The behaviour of the prey (b) is more complex with a distinct increase in the variance just before the large carrying capacity limit is reached.

species models, although the methodology can be adapted for higher dimensional systems.

Suppose that $P_{m,n}$ is the probability that the two populations are of sizes m and n , respectively ($m, n \in \mathbb{N}$). We can again construct an infinite set of coupled ODEs,

$$\begin{aligned} \frac{dP_{m,n}}{dt} = & B_{m-1,n}^1 P_{m-1,n} + B_{m,n-1}^2 P_{m,n-1} \\ & + D_{m+1,n}^1 P_{m+1,n} + D_{m,n+1}^2 P_{m,n+1} \\ & - (B_{m,n}^1 + B_{m,n}^2 + D_{m,n}^1 + D_{m,n}^2) P_{m,n}, \end{aligned}$$

where B^1, B^2, D^1 , and D^2 are the birth and death rates for species one and two respectively. Finding the fixed point solution can be solved by truncating the distribution and solving a matrix equation; however, we would expect the size of the matrix necessary to grow like K^4 . Therefore it is often computationally less intensive to simulate the stochastic process than to solve the underlying matrix equations; this effect becomes far more apparent as the number of species increases.

Figure 4 shows the results from a stochastic Lotka–Volterra type predator–prey system; the rates of increase and decrease are

$$\begin{aligned} B_{v,p}^1 &= C^1 + v & B_{v,p}^2 &= C^2 + \frac{vp}{K} \\ D_{v,p}^1 &= \frac{v^2 + 2vp}{2K} & D_{v,p}^2 &= p, \end{aligned} \quad (13)$$

where v and p are the number of victims (species 1) and predators (species 2), respectively. When the carrying capacity K becomes large the deterministic dynamics predicts that $v \rightarrow K$ and $p \rightarrow \frac{K}{2}$.

From Fig. 4 we see that the same type of local power-law behaviour is observed in this multispecies model, with $\alpha \rightarrow 1$ as the carrying capacity becomes large. This agrees with our intuitions from Section 3, as a power law of one is associated with systems that have a small stochastic component compared to the deterministic dynamics, which is the case here. A rapid change in the variance of both species is seen just before the large K power law begins, and once again this is attributable to the change from bimodal to unimodal distribution.

4.3. Deterministic Dynamics

If the underlying deterministic system possesses non-trivial dynamics (e.g., periodic orbits or more complex attractors) then the power-law scaling can again be effected. For spatially heterogeneous systems, we would expect distant points to be out of phase with each other. This means that the deterministic oscillations over the entire population only grow like \sqrt{K} ; hence the original conclusion still holds and $\alpha \rightarrow 1$ as K becomes large (Keeling *et al.*, 1997).

On the other hand, for homogeneously mixed (non-spatial) systems the amplitude of the deterministic orbit should scale linearly with population size. Therefore, for large K , the power law α will always approach 2 (Keeling and Grenfell, 1999). For chaotic systems, the fractal nature of the attractor could potentially lead to situations of noninteger power-law scaling (Perry, 1994).

5. DISCUSSION

For a wide variety of models and parameters, a range of noninteger power-law behaviour can be seen when the carrying capacity (and hence the mean) is small. However, for many large populations only two sorts of scaling behaviour exist: $\alpha \rightarrow 1$ or $\alpha \rightarrow 2$. This may help to explain the dominance of these two forms of scaling in a variety of ecological data. For many systems the change to large population behaviour is associated with a rapid decline in the number of extinctions and often with a drop in the variance (Figs. 1b–1d). This drop is due to a change from bimodal distributions (with many extinctions) to unimodal distributions (with few extinctions).

The noninteger power law for small carrying capacities can be attributed to both the increasing importance of the individual nature of the system as population sizes become small and the nongeneric behaviour when the population hits zero. The existence of extinct populations introduces two disparate time scales: a fast time scale for the dynamics within an extant population and a slow time scale for the recovery of extinct populations. In many situations, the presence of fast and slow time scales is a prerequisite for power-law behaviour (Tang and Bak, 1988; Rhodes *et al.*, 1997).

For any given set of birth and death rates (B_n and D_n) we can find a unique function ($f(x)$) which describes the deterministic rates of change. However, the fact that B and D are not uniquely determined by f can cause many difficulties when modellers move from a deterministic to a stochastic framework. Two stochastic models can possess the same underlying deterministic equation, but have radically different behaviour (cf. Figs. 1b and 1c). It is therefore vitally important to formulate individual based models from the behaviour of individuals and not to extrapolate from global observations of the dynamics.

It is interesting to note that when $\alpha \rightarrow 2$, the individual nature of the population is unimportant as the standard deviation scales linearly with the mean. Therefore, once this large population regime is reached, although the variance may play a vital role in the dynamics, its relative effect will not change with carrying capacity; this means we can scale out the population size. However, when there is less stochasticity in the underlying dynamics and $\alpha \rightarrow 1$, then the system behaves more like the underlying deterministic equations as the carrying capacity increases. Therefore, although the variance plays a decreasing role in the dynamics, population size is important and cannot be scaled out, as it determines the convergence to determinism. It is important to realise that the two common modelling steps of scaling out population size and assuming determinism are in fact mutually exclusive.

Many other factors need to be considered before we have a complete understanding of mean–variance relationships. Similar power-law behaviour can also be observed in higher dimensional (multi–species) models, such as the Lotka–Volterra equation for predator–prey systems or the SIR model for epidemics (Keeling and Grenfell, 1999). Again we believe that the scaling at large population sizes is determined by the ratio of stochastic to deterministic behaviour. When the underlying equations possess non-trivial dynamics this adds an extra component to the variance which scales like the mean squared; therefore $\alpha \rightarrow 2$ is not necessarily a sign of high order stochasticity.

The presence of spatial heterogeneities will reduce the observed α (cf. Keeling *et al.*, 1997) as fluctuations about the mean may be out of phase, reducing the overall variance observed. This power-law scaling with the spatial scale of the system should be seen as complementary to the earlier scaling with carrying capacity. For sufficiently large populations spatial decorrelation over large distances is a strong possibility; however, for many populations of ecological importance the assumption of homogeneous mixing is justified.

In conclusion, I believe that careful examination of the mean–variance relation for many ecological systems can shed light on the importance of the stochastic processes present and may even allow us to predict the amount of spatial heterogeneity or the accuracy of deterministic approximations.

APPENDIX I: COMPARISON TO CONTINUOUS POPULATION MODELS

Let us assume that we have a single population, whose underlying deterministic dynamics can be expressed as

$$\frac{dx}{dt} = f(x).$$

We require that there is only one equilibrium point, x^* , which is globally attracting.

$$f(x) > 0 \text{ if } x < x^* \quad f(x) < 0 \text{ if } x > x^*. \quad (14)$$

Note that in particular we need the origin to be unstable. We assert that the average rate of increase of the population is given by the birth rate minus the death rate;

$$f(n) = B_n - D_n.$$

Hence we have a simple link between the discrete population rates and a continuous population differential equation.

However, the reverse situation is not so simple. For any given $f(n)$, we can choose B_n and D_n to produce any required population distribution P_n .

$$\begin{aligned} D_0 &= 0 & B_0 &= f(0) \\ D_n &= \frac{P_{n-1}}{P_n} B_{n-1} & B_n &= f(n) - D_n \quad \forall n \neq 0. \end{aligned}$$

Therefore, while there is a natural means of going from the stochastic rates to a differential equation, the reverse is not true.

Note that Eq. (2) is the discrete population version of the Fokker–Plank equation, whose solution

$$P(x) = P_0 \exp\left(\int_0^x \log\left(\frac{B(y)}{D(y)}\right) dy\right)$$

is often a good approximation to P_n when the carrying capacity is large and the distribution contains few small population sizes.

APPENDIX II: CALCULATION OF THE POWER-LAW FOR LARGE POPULATIONS

$$(1) \quad \underline{h(n)/g(n) \rightarrow H, H = O(1)}.$$

$$\begin{aligned} \forall n > N \quad P_n &= P_N \prod_{m=N+1}^n \frac{B_{m-1}}{D_m} \\ &= P_N \prod_{m=N+1}^n \frac{K \frac{g(m-1)}{g(m)} + K \frac{h(m-1)}{g(m)}}{m + K \frac{h(m)}{g(m)}} \\ \Rightarrow P_n &= P_N \prod_{m=N+1}^n \frac{K(H+1)}{m + KH} + O\left(\frac{1}{m^2}\right) \\ &= P_N k^{n-N} (H+1)^{n-N} \frac{\Gamma(M + KH)}{\Gamma(n + KH)} \\ &\quad + O\left(\frac{1}{m^2}\right) \end{aligned}$$

$$\begin{aligned} \Rightarrow M &= \sum_n n P_n \\ &= \sum_n (n + KH) P_n - \sum_n K H P_n \\ &= K(H+1) + O(1) - KH = K + O(1) \end{aligned} \quad (15)$$

$$\begin{aligned} \Rightarrow V &= \sum_n n^2 P_n - M^2 \\ &= \sum_n (n + KH)(n + KH - 1) P_n \\ &\quad - \sum_n (2KH - 1) n P_n \\ &\quad - \sum_n (KH - 1) K H P_n - M^2 \\ &= K^2(H+1)^2 + O(K) - (2KH - 1) M \\ &\quad - (KH - 1) KH - M^2 \\ &= O(K) \end{aligned} \quad (16)$$

Therefore, as the carrying capacity becomes large, $\alpha \rightarrow 1$.

$$(2) \quad \underline{h(n)/g(n) \rightarrow Hn + H_0, H = O(1), H_0 = O(1)}.$$

$$\begin{aligned} \forall m > N \quad \frac{B_{m-1}}{D_m} &= \frac{K + KH_0 + KHm}{m + KH_0 + KHm} + O\left(\frac{1}{m^2}\right) \\ &= \frac{(KH-1)m + K}{KHm} + O\left(\frac{1}{m^2}\right) \end{aligned}$$

This leads to a negative

$$\text{binomial distribution} \Rightarrow M = (H+1)K + O(1) \quad (17)$$

$$\Rightarrow V = (H+1)HK^2 + O(K) \quad (18)$$

Therefore, as the carrying capacity becomes large, $\alpha \rightarrow 2$.

$$(3) \quad \underline{h(n)/g(n) \rightarrow Hn^2 + O(n), H = O(1)}.$$

$$\begin{aligned} \forall m > N \quad \frac{B_{m-1}}{D_m} &= \frac{K + KHm^2 + O(m)}{m + KHm} \\ &= \frac{KH}{KH+1} m + O(1). \end{aligned} \quad (19)$$

Therefore, as $\frac{h}{g}$ is an increasing function of m the distribution P does not converge.

ACKNOWLEDGMENTS

This research was supported by the Royal Society. I wish to thank Bryan Grenfell, Pej Rohani, Jonathan Swinton, the two anonymous referees and the associate editor for their helpful comments.

REFERENCES

- Anderson, R. M., Gordon, D. M., Crawley, M. J., and Hassell, M. P. 1982. Variability in the abundance of animal and plant species, *Nature* **296**, 245–248.
- Bartlett, M. S. 1957. Measles periodicity and community size, *J. R. Stat. Soc. A* **120**, 48–70.
- Bjørnstad, O. N., Stenseth, N. C., and Saitoh, T. 1999. Synchrony and scaling in dynamics of voles and mice in northern Japan, *Ecology* **80**, 622–637.
- Blasius, B., Huppert, A., and Stone, L. 1999. Complex dynamics and phase synchronization in spatially extended ecological systems, *Nature* **399**, 354–359.
- Boag, B., Hackett, C. A., and Topham, P. B. 1992. The use of Taylor power law to describe the aggregated distribution of gastrointestinal nematodes of sheep, *Int. J. Parasitol.* **22**, 267–270.
- Earn, D. J. D., Rohani, P., and Grenfell, B. T. 1998. Persistence, chaos and synchrony in ecology and epidemiology, *Proc. Roy. Soc. Lond. B* **265**, 7–10.
- Hanski, I., and Tiainen, J. 1989. Bird ecology and Taylor variance-mean regression, *Ann. Zool. Fennici* **26**, 213–217.
- Keeling, M. J. 1997. Modelling the persistence of measles, *Trends Microbiol.* **5**, 513–518.
- Keeling, M. J., and Grenfell, B. T. 1999. Stochastic dynamics and a power law for measles variability, *Phil. Trans. R. Soc. Lond. B* **354**, 769–776.
- Keeling, M. J., Mezic, I., Hendry, R. J., McGlade, J., and Rand, D. A. 1997. Characteristic length scales of spatial models in ecology via fluctuation analysis, *Phil. Trans. R. Soc. Lond. B* **352**, 1589–1601.
- Keitt, T. H., and Stanley, H. E. 1998. Dynamics of North American breeding bird populations, *Nature* **393**, 257–260.
- Perry, J. N. 1994. Chaotic dynamics can generate Taylor's power-law, *Proc. R. Soc. Lond. B* **257**, 221–226.
- Perry, J. N., and Woiwod, I. P. 1992. Fitting Taylor's power law, *Oikos* **65**, 538–542.
- Ranta, E., Kaitala, V., and Lindstrom, J. 1997. Dynamics of Canadian lynx populations in space and time, *Ecography* **20**, 454–460.
- Rhodes, C. J., Jensen, H. J., and Anderson, R. M. 1997. On the critical behaviour of simple epidemics, *Proc. Roy. Soc. Lond. B* **264**, 1639–1646.
- Sutcliffe, O. L., Thomas, C. D., and Moss, D. 1996. Spatial synchrony and asynchrony in butterfly population-dynamics, *J. Animal Ecol.* **65**, 85–95.
- Swinton, J., Harwood, J., Grenfell, B. T., and Gilligan, C. A. 1998. Persistence thresholds for phocine distemper virus infection in harbour seal *Phoca vitulina* metapopulations, *J. Animal Ecol.* **67**, 54–68.
- Tang, C., and Bak, P. 1988. Critical exponents and scaling relations for self-organized critical phenomena, *Phys. Rev. Lett.* **60**, 2347–2350.
- Taylor, L. R. 1961. Aggregation, variance and the mean, *Nature* **189**, 732–735.
- Taylor, L. R., and Woiwod, I. P. 1980. Temporal stability as a density-dependent species characteristic, *J. Animal Ecol.* **49**, 209–224.