Evolutionary Ecology Research, 2000, **2**: 769–789

Evolutionary fitness in ecology: Comparing measures of fitness in stochastic, density-dependent environments

T.G. Benton[1]* and A. Grant[2]

[1]*Institute of Biological Sciences, University of Stirling, Stirling FK9 4LA and* [2]*School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK*

ABSTRACT

Several different measures of 'fitness' are commonly used in evolutionary studies. Each measure makes different assumptions, so is appropriate only in some circumstances. Many recent studies have recommended using invasibility arguments to identify the evolutionarily unbeatable strategy (EUS), rather than choosing a measure of fitness to maximize, thereby avoiding the potential pitfalls in choosing. Here we use the 'invasion exponent' to determine the EUS of reproductive allocation in environments that vary in density dependence and environmental stochasticity. We then compare the EUS effort with that predicted by a range of other measures of fitness: r (the deterministic per capita rate of increase), a (the stochastic per capita rate of increase), R_0 (lifetime reproductive success) and population size (arithmetic and geometric means). When the population is at an equilibrium in a constant environment, different measures of fitness predict the same optima. However, when population size is not constant (either due to environmental variation or non-equilibrium dynamics), no single fitness measure universally predicts the EUS. In most circumstances, measures of population size perform best followed by measures of reproductive performance. Measures of population growth perform least well.

Keywords: density dependence, elasticity, evolutionarily unbeatable strategy, fitness, invasion exponent, Lyapunov exponent, life history, lifetime reproductive success, reproductive effort, stochasticity.

INTRODUCTION

Following the publication of *The Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life* in 1859, Darwin encountered some hostility to the use of the phrase 'natural selection' (Burkhardt *et al.*, 1993, p. 371), principally because 'selection' suggested to some that there must be some entity doing the selecting. However, the label 'natural selection' was soon widely established and Darwin realized there was little point in trying to alter it to his, with hindsight, preferred 'Natural Preservation' (Burkhardt *et al.*, 1993, p. 389). In his *Principles of Biology* (1864), Herbert Spencer coined the term 'survival of the fittest' as a popularization of Darwin's term,

---

* Author to whom all correspondence should be addressed. e-mail: t.g.benton@stir.ac.uk

without the teleological implications of ‘selection’ (Desmond and Moore, 1991). It was Alfred Russel Wallace who, in 1866, wrote to Darwin to suggest the adoption of this expression: ‘The term “survival of the fittest” is the plain expression of the fact; “natural selection” is a metaphorical expression of it, and to a certain degree indirect and incorrect, since . . . Nature . . . does not so much select special varieties as exterminate the most unfavourable ones’ (A.R. Wallace in F. Darwin, 1887, p. 45).

Darwin then adopted ‘survival of the fittest’ as a synonym of natural selection in the sixth edition of *The Origin of Species* in 1872. The concept of ‘fitness’ as a measure of ‘success at being naturally selected’ has since been adopted and used in a multitude of ways within evolutionary biology. Richard Dawkins (1982) discusses some of the technical usages of the term in a chapter entitled ‘An agony in five fits’, and concludes that the term has caused so much confusion we would be better off without it. Part of this confusion stems from the fact that there are several commonly used fitness measures, each with their own explicit and implicit assumptions. Stearns (1992, p. 33) reminds us that ‘all fitness definitions are tools invented by scientists to analyse natural selection’. Fitness is used as a performance index in evolutionary arguments: the trait with the greatest fitness is assumed to be the best in optimality studies. To have value, any measure of fitness must be related to the outcome of natural selection – that is, the number of descendants (of a genotype, phenotype or even individual) into the future. As such, fitness must be a property of the life history as a whole, not just a small part thereof.

Measures of fitness are context-dependent, in that there is no single quantity that natural selection can be said to ‘maximize’. For example, lifetime reproductive success is often used as a measure of fitness, but a strategy maximizing reproductive success does not necessarily mean that that strategy will be most successful in leaving descendants into the future. This is for two reasons. First, having a large number of eggs does not need to imply a large number of surviving offspring, let alone grand-offspring. This is because there are likely to be trade-offs between offspring number and their size and thus ability to survive – having many offspring does not ensure that many will survive. The environment is also a stochastic place; sometimes ‘fit’ individuals will fail to breed due to extrinsic, rather than intrinsic, reasons. Second, the ‘value’ of an offspring will depend on whether the population is growing, shrinking or constant. In a growing population, it is better to produce fewer offspring at an earlier time, rather than more later.

Many different measures of fitness have been used within evolutionary ecology. Broadly speaking they fall into four groups.

1. Those based on measures of population growth (such as r).
2. Those based on measures of reproductive success (such as R_0, LRS, inclusive fitness).
3. Those based on population size (such as K).
4. Others. For example, within ecology, the time to extinction is often used as a measure of population ‘fitness’ (see Foley, 1997, for a recent review).

However, here we concentrate on the three most common measures of fitness in behavioural and evolutionary ecology.

Within life-history theory, the most commonly used measure of fitness is r, the instantaneous per capita growth rate (Stearns, 1992). Strictly, use of r assumes that the life history is unaffected by density-dependent effects and that the environment is constant (Charlesworth, 1994). A stochastic analogue of r, termed a by Tuljapurkar and Orzack (1980), has been described, which allows the relaxation of the constant environment assumption.

Within behavioural ecology, the most commonly used measures of fitness are measures of reproductive success, including *lifetime reproductive success* (LRS or R_0) (Clutton-Brock, 1988). For LRS to function as a measure of fitness, the population size must be constant. Then all that matters is how many offspring are produced, not when they are produced. This is a strong assertion for natural populations (Giske *et al.*, 1993). In most populations, there may be *on average* no net change in population size over time, such that $r \approx 0$, but any change in population size will change the value of offspring and so devalue LRS as a fitness measure. Even when the population is constant, it is often assumed that use of r or R_0 is a 'matter of taste' (Kozlowski, 1993). However, it has recently been emphasized that evolutionary predictions depend on the precise form of the density dependence (Mylius and Diekmann, 1995; Metz *et al.*, 1996), and sometimes r, rather than LRS, may be maximized in a stationary population. Therefore, choice of r or LRS should not be a matter of taste, but driven by a biological understanding of how density dependence operates.

The third fitness measure often used is a measure of the carrying capacity (K) – the population size at equilibrium – of the environment for the strategy in question. If two genotypes are competing for limiting resources, the fittest is the one that is most 'efficient' and converts the limiting resource into the most bodies (MacArthur and Wilson, 1967; Charlesworth, 1994).

There are thus a variety of different measures of fitness which make different assumptions and thus do different jobs. The use of any one of them usually requires at least one of the simplifying assumptions: the environment is constant, or there is no density dependence. In many cases, neither of these assumptions will hold.

Increasingly, adaptational analyses are couched in terms of invasion dynamics (Metz *et al.*, 1992; Rand *et al.*, 1994; Ferriere and Gatto, 1995; Mylius and Diekmann, 1995; Gurney and Middleton, 1996; Van Dooren and Metz, 1998; Benton and Grant, 1999a; Diekmann *et al.*, 1999): to predict the outcome of selection one should model the invasion of a mutant into a population of residents and find the evolutionarily unbeatable strategy (EUS; Van Dooren and Metz, 1998). The invasion of a mutant phenotype into an established population is measured by a mathematical quantity called the dominant Lyapunov exponent. This is termed ϑ, the invasion exponent, by Rand *et al.* (1994). In simple terms, if a mutant arises in an established population, ϑ is estimated as the growth rate of the invader population. Invasion can only occur if $\vartheta > 0$, leading to the identification of ϑ as fitness (which by necessity depends on the environment in which the strategy finds itself) (Metz *et al.*, 1992, 1996; Van Dooren and Metz, 1998).

Few studies have compared the different measures of fitness to illustrate their strengths and weaknesses. The aim of this study is to identify the optimal values of a trait in a range of environments, characterized by differences in density dependence and variability. Comparison of the optimal life histories predicted using a range of fitness measures with the evolutionarily unbeatable strategy will be informative in guiding the choice of measure for individual studies.

METHODS

We model the population dynamics using a two class matrix population model (see Caswell, 1989, for a review of matrix models). The standard two class matrix model is:

$$\begin{pmatrix} F_1 & F_2 \\ S_1 & S_2 \end{pmatrix} \cdot \begin{pmatrix} n_{1(t)} \\ n_{2(t)} \end{pmatrix} = \begin{pmatrix} n_{1(t+1)} \\ n_{2(t+1)} \end{pmatrix}$$

where the F_x represent the number of offspring born to an individual in stage class x and surviving until the next census. S_1 represents survival of an individual from stage 1 to 2 and S_2 represents the persistence of individuals in stage 2. Multiplication of the life-history matrix by the vector containing the numbers in each stage (the $n_{x(t)}$) gives the numbers in each age class at the next time step.

This basic model is modified in three ways. First, we simulate the addition of stochastic variation in the vital rates (ε_x). Second, we assume that there is a trade-off between current fecundity and survival, measured by the proportional reproductive effort, E. We model this trade-off as a non-linear function, by multiplying the survivorship terms by $(1 - E^z)$, following Cooch and Ricklefs (1994). Ricklefs (1977) reviewed the literature and suggested that $z = 6$ was realistic (which is the default value taken below). Third, we assume there is density-dependent feedback on the fecundity terms (such as if there were density-dependent survival of juveniles between birth/hatching and the first census), the survival terms, or both. Our model then becomes:

$$\begin{pmatrix} (F_1 \pm \varepsilon_1) \cdot E \cdot f(\Delta), & (F_2 \pm \varepsilon_2) \cdot E \cdot f(\Delta) \\ (S_1 \pm \varepsilon_3) \cdot (1 - E^z) \cdot f(\Delta), & (S_2 \pm \varepsilon_4) \cdot (1 - E^z) \cdot f(\Delta) \end{pmatrix} \cdot \begin{pmatrix} n_{1(t)} \\ n_{2(t)} \end{pmatrix} = \begin{pmatrix} n_{1(t+1)} \\ n_{2(t+1)} \end{pmatrix}$$

The impact of stochasticity was investigated by varying:

1. The amount of variation (as measured by the coefficient of variation, CV, the ratio of the standard deviation to the mean).
2. The correlations between the variations in the vital rates. Sampling was conducted so that the rank correlations took the values 1, 0 or −1. This resulted in four patterns of variations, each of which can be described by three correlations: between fecundities, between survivorships and between fecundities and survivorships. The four patterns are: (a) variations across all vital rates were positvely correlated – a good year for one vital rate was also a good year for all others (correlations = 1, 1, 1); (b) variations across all vital rates were uncorrelated (correlations = 0, 0, 0); (c) variations between the fecundity terms and the survivorship terms were positively correlated, but between fecundities and survivorships there was a negative relationship – a good year for fecundities was a bad year for survivorships (correlations = 1, 1, −1); (d) variations between the fecundities and survivorships were positively correlated, but there was no correlation between fecundities and survivorships (correlations = 1, 1, 0).
3. The sampling distribution from which the variations were drawn. Three basic patterns of variation were used. First, sampling for variations in fecundity was taken from a log normal distribution and variations in survival were from a beta distribution. Second, we used a two-state discrete distribution, whereby for each time period the value of the vital rate was taken as the mean ± (CV*mean). Whether the (CV*mean) term was added or subtracted was decided independently for each time step. Such a ‘good year’ or ‘bad year’ scenario was used by Schaffer (1974). Recently, Sibly *et al.* (1991) showed that the predictions Schaffer made with such a distribution were robust to the choice of distribution. Third, we modelled a situation where organisms live in a constant environment but which is liable to periodic ‘catastrophic years’. In randomly chosen

'years', but at a given average frequency, some (or all) of the vital rates were reduced to 1% or 10% of their average value.

In all simulations, random deviates for each year were independent; we made no attempt to investigate the impact of serial autocorrelation

The impact of the density-dependent function was investigated by density dependence acting on the fecundities, the survivorships, or both. In addition, four models of density dependence were used. The density dependence was such that the vital rates were a function of total population size, rather than the population density of a subset (e.g. juveniles) in the population.

1. *Ricker-type*. Each term is reduced by an amount given by $\exp(1 - *\beta*(N_{\text{tot}}))$, where β is a scaling constant and $N_{\text{tot}} = N_1 + N_2$ is the total population size.
2. *Beverton and Holt-type*. Each term is multiplied by $1/(1 + 2*\beta*(N_{\text{tot}}))$.
3. *Usher-type*, *with gradual onset of density dependence*. The density multiplier was given by $1/(1 + \exp(1.25*\beta*(N_{\text{tot}}) - 50{,}000*\beta)$.
4. *Usher-type*, *with sudden onset of density dependence*. The density multiplier was given by $1/(1 + \exp(12.5*\beta*(N_{\text{tot}}) - 500{,}000*\beta)$.

The parameters chosen were such that the equilibrium population size under each type of density dependence was within 10%. For all models, $\beta = 0.00002$. For most life histories and most values of reproductive effort, the Usher-sudden-onset density dependence made the underlying deterministic population dynamics chaotic. The vast majority of other combinations of density dependence, life history and E resulted in equilibrium dynamics, although a few examples of bifurcations to limit cycles were encountered with large values of E (i.e. nearly semelparous life histories).

We investigated the impact of stochasticity and density dependence on 15 life histories (Table 1) to represent a range from semelparous to iteroparous. As the basic life histories shown in Table 1 are subject to the reproductive effort trade-off (so an 'iteroparous' life history could become semelparous with a high reproductive effort), the values shown should be viewed as constraints, or boundaries on the actual vital rates expressed.

The optimal value of the proportional reproductive effort was identified using a method first suggested by Grant (1997) to estimate the elasticity of vital rates in stochastic, density-dependent environments and is based on the mathematics of invasion developed in recent years (Metz *et al.*, 1992; Rand *et al.*, 1994; Ferriere and Gatto, 1995). Rand *et al.* (1994) showed that ϑ, the invasion exponent, can be used to find the evolutionarily stable attractor (ESA) in a dynamical system, whatever the underlying dynamics. ϑ is equivalent to the dominant Lyapunov exponent of the system, as discussed by Metz *et al.* (1992), Ferriere and Gatto (1995) and Tuljapurkar (1989). The EUS value of reproductive effort is found by estimating the elasticity of ϑ to changes in E, using numerical differentiation. In short, a population of residents is modelled until the population dynamics have stabilized on the attractor (whatever its underlying dynamics). An invader population with the same life history as the resident, but with E very slightly different (here we use $0.995*E$), is then simulated, starting with a small population and ignoring demographic stochasticity. The invader population does not contribute to the density dependence, thus linearizing the system about the point where the invading population is zero. The sign of ϑ then measures

Table 1. The life histories used in this study

	Matrix elements								
Life history	F_1	F_2	S_1	S_2	Deterministic ESS effort[a]		T_0	R_0	λ
1	0.1	2	0.8	0.95	F	0.463	3.72324	12.3745	1.45963
2	1.0	10.0	0.6	0.85	F	0.556	2.16102	19.0493	2.49475
					S	0.99	1.23194	1.3558	1.27395
3	0.5	5.0	0.5	0.75	F	0.556	2.46952	5.23446	1.68594
					F&S	0.69	2.29758	4.99493	1.75807
					S	0.747	2.2188	4.43075	1.74486
4	1.0	5.0	0.75	0.85	F	0.558	2.20526	12.1108	5.12193
					F&S	0.724	2.00235	9.25488	2.25025
					S	0.99	1.1695	1.21863	1.1819
5	4.0	10.0	0.6	0.85	F	0.565	1.461	20.7216	3.48962
					F&S	0.880	1.19507	8.71134	4.26273
6	1.0	4.0	0.5	0.75	F	0.614	2.09291	4.61903	1.74103
7	1.0	10.0	0.3	0.6	F	0.658	1.92385	4.70033	1.95248
					S	0.99	1.14074	1.17013	1.1464
8	0.85	15.0	0.1	0.5	F	0.695	1.85831	2.2534	1.48179
					F&S	0.748	1.7932	2.21089	1.49259
					S	0.847	1.6388	1.89052	1.43544
9	1.0	5.0	0.3	0.5	F	0.701	1.76368	2.35785	1.54226
					F&S	0.763	1.68721	2.29757	1.55756
					S	0.99	1.07973	1.07952	1.07324
10	4.0	10.0	0.3	0.6	F	0.703	1.21649	6.7369	3.45377
					F&S	0.981	1.02061	4.26629	4.005321
11	1.0	10.0	0.3	0.3	F	0.720	1.71385	3.2262	1.872
					S	0.99	1.13691	1.16691	1.14426
12	1.0	10.0	0.1	0.3	F	0.765	1.51715	1.56969	1.32743
					S	0.99	1.05474	1.04897	1.04632
13	4.0	10.0	0.3	0.3	F	0.785	1.15286	5.48215	3.66511
					F&S	0.984	1.01722	4.21604	4.00447
14	4.0	10.0	0.1	0.3	F	0.910	1.03002	4.09181	3.74866
					F&S	1.000	1.000	4.000	4.000
15	5.0	2.5	0.5	0.3	F	0.912	1.02349	5.1147	4.66664

[a] The deterministic ESS effort is for $z = 6$. F refers to environments where density affects fecundity terms only, S survivorship and F&S both together.

Note: The parameters T_0, R_0 and λ are calculated for the mean matrix before the application of density dependence.

whether the invader manages to invade, and the magnitude of the partial differential $(\delta\vartheta/(\partial\vartheta/\partial\log E)$ is a measure of the proportional sensitivity of ϑ to change in E, and thus a measure of the selection pressure on E (Benton and Grant, 1999b). The evolutionarily stable attractor (equivalent to the EUS) for E is given by the condition $\partial\vartheta/\partial\log E = 0$. An example analysis is given in Benton and Grant (1999a).

Procedures were written in Fortran 77 to estimate ϑ. Computationally, ϑ is measured as the growth rate of the invader population in an environment dominated by the residents, where the invaders do not contribute to the density dependence. The invader population's growth rate is estimated by the average slope of the least-squares regression of the logarithm of invader population size on time (Rand *et al.*, 1994; Grant, 1997). Invasion attempts were of length 50 time steps for most simulations, but 100 for the ones simulating catastrophic events. Changing the length of the invasion attempts did not change anything but the precision of the results. Each value of ϑ was calculated from 1000 independent replicate invasion attempts, or 500 in the case of catastrophes. The EUS, where $\partial\vartheta/\partial \log E = 0$, was estimated in a number of ways, depending on the scatter in the relationship between E and $\partial\vartheta/\partial \log E$. When scatter is low, the value can be calculated by interpolation. In most situations, $\partial\vartheta/\partial \log E = 0$ was read directly off a graph. In cases where scatter is high, polynomial regression was used. The precision of these methods was estimated using the same life history, at medium variation ($CV = 0.5$) and high variation ($CV = 0.9$), with density dependence specified by the four different functions outlined above. The EUS value of E was estimated 10 times independently for each combination of CV and density dependence, and the repeatability (intraclass correlation coefficient; Lessells and Boag, 1987) was calculated. The repeatabilites were 0.9988 for $CV = 0.5$, 0.9536 for $CV = 0.9$ and 0.9925 combining the analyses. So, the precision of the estimate of the EUS declines with increasing variation, but can nonetheless be estimated with some confidence.

For each value of E, the following were calculated:

1. Two matrices based on the average vital rates: one based on the vital rates after the impact of stochasticity (the density-independent, DI, average) and one based on the vital rates after the impact of both stochasticity and density (the density-dependent, DD, average). Each matrix was an average of the 50,000 realizations used per simulation. Using each type of matrix, r, a and R_0 were calculated. a, the stochastic analogue of r, was estimated using the approximation developed by Tuljapurkar (1982). DI-r, DI-a and DI-R_0 can be thought of as measuring the 'intrinsic' vital rates modified by the impact of environmental stochasticity. DD-r, DD-a and DD-R_0 measure the vital rates after both stochasticity and density dependence have been added.
2. The average population size (K). For each simulation, there were 50,000 time steps. The population size at each time step was used to calculate the arithmetic mean population size (K_{am}) and the geometric mean population size (K_{gm}).

For measures of fitness, the optimal effort was taken as that value of E which maximized fitness. Simulations were conducted on the University of Stirling's HP9000/887 computer. Random deviates were selected using routines from the NAG library. Copies of the programs are available on request.

Statistical analysis was used to investigate the strength of the relationships between the EUS and the optima predicted using the various measures of fitness. As the optima are related manifestations of a single model, there must be exact mathematical relationships which relate them. However, such relationships would be difficult to define given the model's complexity. Clearly, the statistical significance of the relationships will depend on the life histories and environments used, but the use of statistics is a way of measuring the strength of association of the performance of the different fitness measures, within the confines of the data used here.

RESULTS

Constant density-dependent environments

The correspondence between the various fitness measure estimates of optimum effort and the EUS estimated by ϑ is shown in Fig. 1 for the different life histories in a constant environment. Even in constant environments, none of these relationships is a simple y = x line. For the arithmetic and geometric mean sizes, 70–80% of the points fall exactly on a y = x line, showing correspondence between the different ways of estimating the optimum effort. However, the other points show no relationship to the y = x line (Fig. 1D). These outliers represent cases where the underlying population dynamics are at non-equilibrium (cycles or chaos), mainly when the density dependence is modelled using the Usher model, with abrupt onset of density dependence. For DI-R_0, some 60% of points fall on the y = x line and the outliers represent non-equilibrium density dependence or cases where density dependence acts on survivorships or both fecundities and survivorships, rather than fecundities alone (Fig. 1A). As expected, DI-r and DI-a optima are identical in deterministic environments. There is a weak relationship between the DI-r and DI-a optima and the ESS effort predicted by ϑ ($R^2 = 17\%$, $P < 0.001$), but it is too weak to have much predictive power (Fig. 1B). Because, in a constant environment, there is no population growth, all values of effort have the same fitness when fitness is measured using the density-

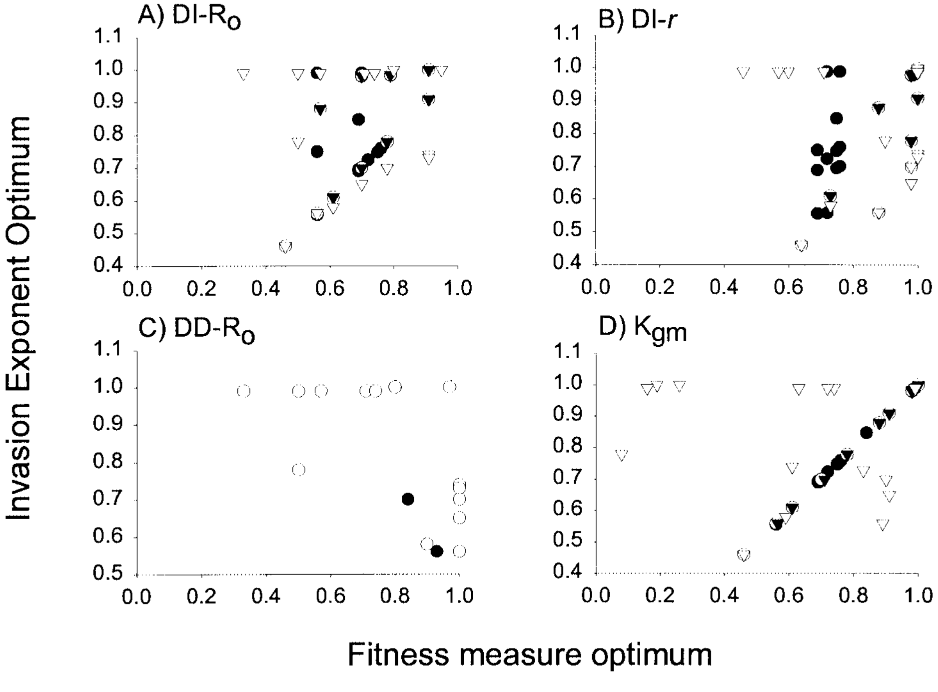

Fig. 1. Relationship between the optimum effort predicted by maximizing different measures of fitness and the EUS effort predicted using the invasion exponent in a constant environment. ● = Ricker DD, ○ = B&H DD, ▼ = Usher-gradual-onset DD, ▽ = Usher-sudden-onset DD. (A) DI-R_0 ($n = 64$, 38 of which lie on y = x); (B) DI-r ($n = 64$); (c) DD-R_0 ($n = 76$, 60 of which have no optimum effort as all values of E gave same R_0); (d) K_{gm} ($n = 76$, 64 of which lie along y = x).

dependent measures of r, a and R_0. However, when the deterministic population dynamics are a limit-cycle or chaos, these quantities are not necessarily 0 (DD-r and DD-a) or 1 (DD-R_0) and may vary with E. In these cases, the optimum predicted by using these measures bears no relation to the EUS effort (Fig. 1C).

Stochastic density-dependent environments

Optimum effort was estimated using each of the different measures of fitness in 525 different 'environments' characterized by differences in variability, the correlations between variations in the vital rates, the sampling distribution from which the stochasticity is drawn and different types of density dependence. The relationships between these optima and the EUS effort predicted by ϑ are shown in Figs 2–6. Summary statistics for the regressions are given in Table 2. From Table 2, it is immediately apparent that the performance of a fitness measure is strongly influenced by the vital rates affected by density. When both fecundities and survivorships are modified by density dependence, the optima predicted by optimizing the fitness measures correspond most strongly with the EUS predicted using ϑ (except for DI-R_0). When survivorships alone are modified by density, there is generally the lowest correspondence between the predicted optima.

r as a measure of fitness

DI-r is generally a better predictor of the EUS effort than DD-r (Table 2). When densities modify survivorships alone, there is no relationship between optima and the EUS. When both fecundities and survivorships are modified by density, the relationship between the DI-r optima and EUS effort accounts for 79% of the variation in EUS effort. The equivalent relationship accounts for only 28.7% of the variation with DD-r as the predictor of EUS effort. When density modifies fecundities only, the figures are 43.5% and 4.6% respectively.

a as a measure of fitness

Irrespective of the vital rates affected by density, none of the relationships between the optima predicted by DD-a and ϑ were significant. DI-a performed better when density affects the fecundities and survivorships or the fecundities alone. In these cases, the optima are closer to the EUS efforts than those predicted using DI-r.

R_0 as a measure of fitness

Neither DI-R_0 nor DD-R_0 predict the EUS effort when density acts on the survivorships. Even when density acts on both fecundities and survivorships, DD-R_0 explains little variation in the EUSs ($R^2 = 23\%$). DI-R_0 is a better predictor, accounting for 62% of the variation. When density affects fecundities alone, DI-R_0 performs moderately well ($R^2 = 74\%$), whereas DD-R_0 explains less than 2% of the variation in ESSs.

Population size (K) as a measure of fitness

By far the best predictors of the optimum effort calculated using the invasion exponent are measures of the population size. With equilibrium population dynamics, and pooling all data irrespective of the vital rates affected by density, geometric mean population size

Table 2. Regression statistics for relationships between optimal efforts predicted by different measures of fitness and the EUS effort

Fitness measure	Data in analysis[a]	R^2	P^b	F	d.f.	Regression equation	Notes[c]
DI-r	All	15.5	$<0.0005^\dagger$	97.1	1,524	$0.503 + 0.329*r$	
	DD on S	0.0	0.476	0.5	1,133		
	DD on F&S	85.4	$<0.0005^\dagger$	684.8	1,117	$1/(\text{EUS}^2) = 3.90 - 2.78*r$	
	DD on F	43.5	$<0.0005^\dagger$	209.2	1,269	$0.240 + 0.563*r$	
DD-r	All	0.8	0.024	5.1	1,518	$0.747 + 0.0446*r$	
	DD on S	1.5	0.09	2.9	1,129		
	DD on F&S	28.7	$<0.0005^\dagger$	48.8	1,118	$0.687 + 0.178*r$	Heteroscedastic
	DD on F	4.6	$<0.0005^\dagger$	14.0	1,267	$0.614 + 0.138*r$	
DI-a	All	13.4	$<0.0005^\dagger$	82.1	1,524	$0.528 + 0.309*a$	
	DD on S	4.3	0.009	7.0	1,133	$0.945 - 0.149*a$	
	DD on F&S	90.5	$<0.0005^\dagger$	1125.0	1,118	$1/(\text{EUS}^2) = 4.15 - 3.13*a$	
	DD on F	52.3	$<0.0005^\dagger$	297.1	1,269	$0.175 + 0.652*a$	
DD-a	All	0.1	0.237	1.4	1,501		
	DD on S	1.5	0.09	3.0	1,133		
	DD on F&S	0.7	0.175	1.9	1,118		Heteroscedastic
	DD on F	0.0	0.34	0.9	1,246		
DI-R_0	All	25.9	$<0.0005^\dagger$	179.1	1,511	$\ln(\text{EUS}) = -0.690 + 0.608*R_0$	
	All	34.6	$<0.0005^\dagger$	233.1	1,434	$\ln(\text{EUS}) = -0.765 + 0.713*R_0$	–All US
	DD on S	0.0	0.764	0.1	1,120		
	DD on F&S	62.0	$<0.0005^\dagger$	195.1	1,118	$0.349 + 0.609*R_0$	
	DD on F	74.3	$<0.0005^\dagger$	780.6	1,269	$0.147 + 0.810*R_0$	

DD-R_0	All	0.6	0.046	4.0	1,505		
	DD on S	0.1	0.295	1.1		ln transformed	
	DD on F&S	23.2	$<0.0005^{\dagger}$	37.0		$\ln(\text{EUS}) = -0.322 + 0.178{*}R_0$	
	DD on F	1.9	0.015	6.0	1,263	$0.684 + 0.0738{*}R_0$	
K_{am}	All	47.2	$<0.0005^{\dagger}$	469.3	1,522	$0.345 + 0.553{*}K_{am}$	Outliers
	All	81.3	$<0.0005^{\dagger}$	2251.5	1,516	$0.137 + 0.805{*}K_{am}$	–US outliers
	All	90.4	$<0.0005^{\dagger}$	4240.0	1,448	$0.118 + 0.840{*}K_{am}$	–All US
	DD on S	72.0	$<0.0005^{\dagger}$	339.9	1,131	$\ln(\text{EUS}) = -0.820 + 0.721{*}\,K_{am}$	
	DD on F&S	98.1	$<0.0005^{\dagger}$	5690.6	1,112	$0.0261 + 0.972{*}K_{am}$	–US outliers
	DD on F	75.1	$<0.0005^{\dagger}$	817.5	1,269	$0.126 + 0.815{*}K_{am}$	Heteroscedastic
K_{gm}	All	44.4	$<0.0005^{\dagger}$	407.9	1,508	$0.346 + 0.571{*}K_{gm}$	Outliers
	All	91.0	$<0.0005^{\dagger}$	5036.1	1,495	$0.0588 + 0.931{*}K_{gm}$	–US outliers
	All	93.7	$<0.0005^{\dagger}$	6412.0	1,430	$0.0526 + 0.942{*}K_{gm}$	–All US
	DD on S	86.7	$<0.0005^{\dagger}$	849.3	1,129	$0.225 + 0.736{*}K_{gm}$	
	DD on F&S	91.6	$<0.0005^{\dagger}$	1189.8	1,108	$0.0712 + 0.937{*}K_{gm}$	
	DD on F	90.6	$<0.0005^{\dagger}$	2459.2	1,255	$0.0565 + 0.923{*}K_{gm}$	

[a] DD on S = cases where density acts solely on survival terms; DD on F = density acts solely on fecundity terms; DD on F&S = density acts on both survival and fecundity terms; All = all data included in analysis.

[b] Sequential Bonferroni adjustment to α: α_{crit} becomes 0.00269 (at 5% level). † Significant results.

[c] Heteroscedastic = transformation to equalize variances not found; therefore, regression statistics approximate. Outliers = regression relationship strongly affected by small number of points with very large residuals. –US outliers = outliers removed from model if equilibrium population size not constant (e.g. chaos); this occurs when density dependence is Usher-type with sudden-onset. –All US = analysis conducted without data generated when density dependence modelled with Usher-sudden-onset.

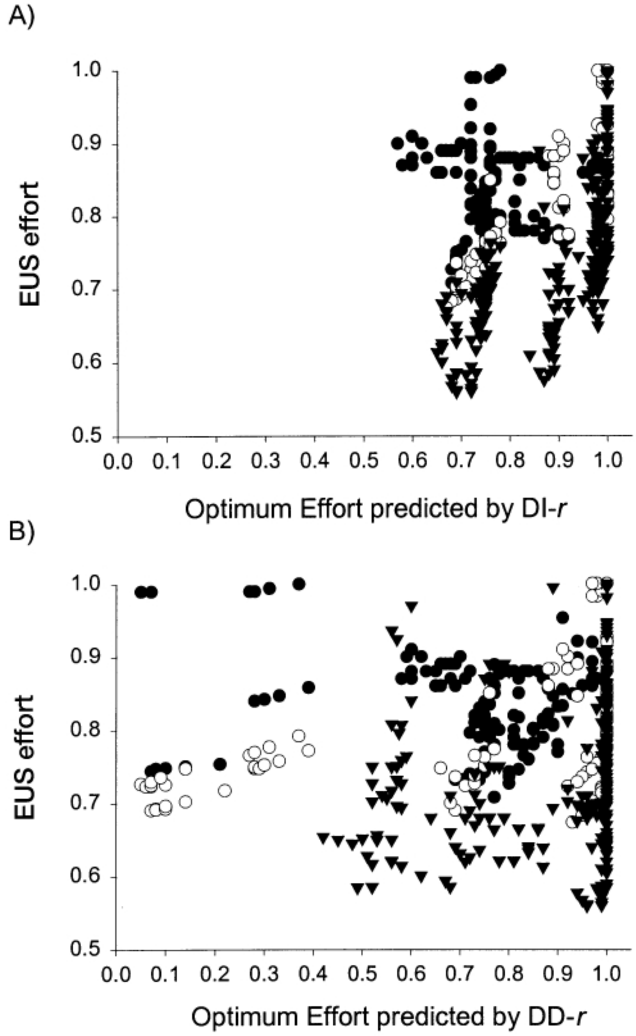

Fig. 2. Relationship between the optimum effort predicted by maximizing per capita growth rate, *r*, and the EUS effort predicted using the invasion exponent. Each point represents a different combination of life history and environment, characterized by the amount of variation, the distribution from which variation is drawn, the correlations between vital rates, the vital rates modified by density, the model of density-dependent feedback, and so on. (A) *r* calculated from the average matrices before modification by density dependence ($n = 536$). (B) *r* calculated from average matrices after modification by density dependence ($n = 530$). ● = survivorships modified by density dependence; ○ = fecundities and survivorships modified by density dependence; ▼ = fecundities alone modified by density dependence.

explains 93.7% of the variation in the EUSs predicted by ϑ. Similarly, arithmetic mean population size explains 90.4%. Conversely, the next best predictor, DI-R_0, explains only 41.4% of the overall variation. However, when the underlying population dynamics are non-equilibrium, the predicted optima may diverge markedly (Figs 1, 5, 6).

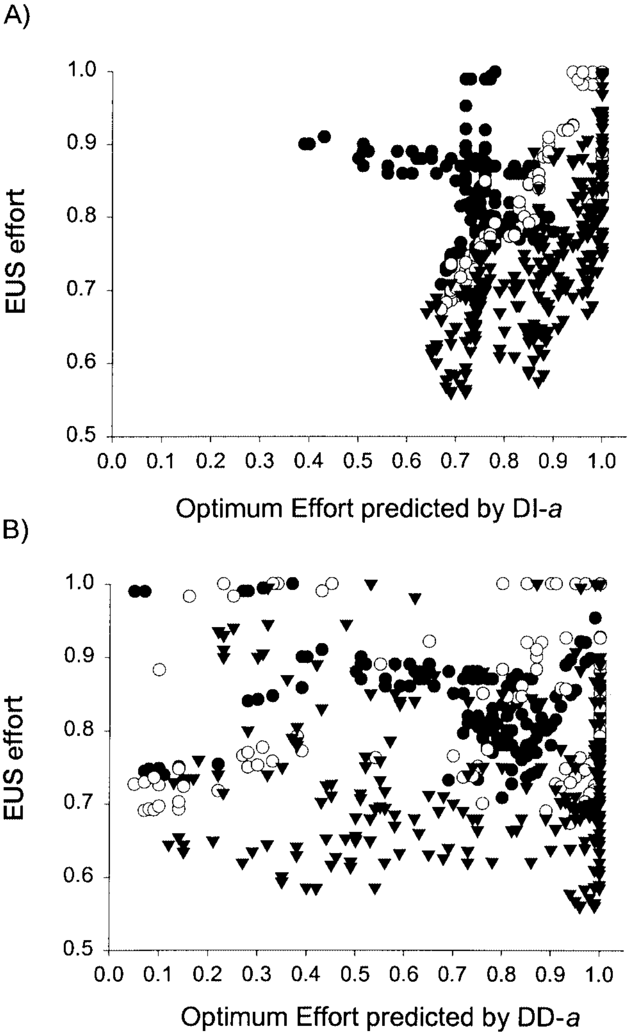

Fig. 3. Relationship between the optimum effort predicted by maximizing stochastic per capita growth rate, *a*, and the EUS effort predicted using the invasion exponent. (A) *a* calculated from the average matrices before modification by density dependence ($n = 536$). (B) *a* calculated from average matrices after modification by density dependence ($n = 530$). Other details as Fig. 2 legend.

The cause of divergence between estimates of optimum efforts

The divergence between the predictions of the EUS using ϑ and optima from other fitness measures depends on the fitness measure in question (Table 3). For example, the difference between the optimum effort predicted using the geometric mean population size and the EUS depends on the CV, the correlation and the vital rates affected by density dependence. Sixty percent of the variance is unaccounted for, most likely because the estimates of the optima are sufficiently close that the precision of the measurement techniques are being approached. Conversely, the R^2 for the other models, using the less accurate fitness

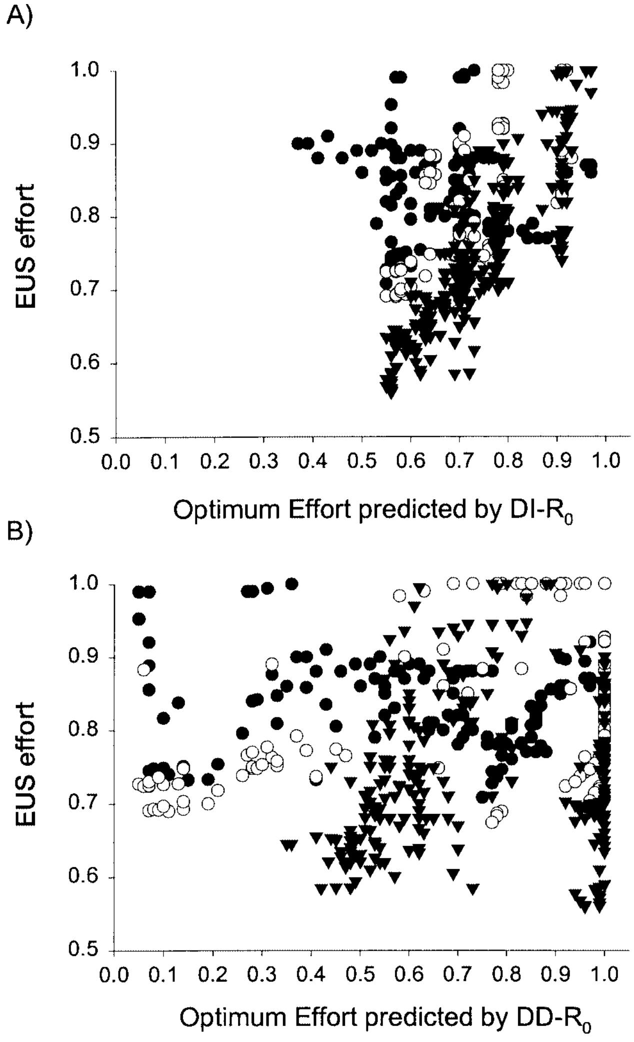

Fig. 4. Relationship between the optimum effort predicted by maximizing reproductive success (R_0) and the EUS effort predicted using the invasion exponent. (A) R_0 calculated from the average matrices before modification by density dependence ($n = 523$). (B) R_0 calculated from average matrices after modification by density dependence ($n = 511$). Other details as Fig. 2 legend.

measures, are considerably higher (> 80%). With these fitness measures, many factors affect the difference between the predicted optima: the CV, the correlations between vital rates, the life history, the deterministic optima, the vital rates affected by density and the interactions of all these factors. With all fitness measures, however, the CV and correlation structure are very important determinants of how good a predictor the fitness measure is. The greater the CV the larger the discrepancy between predictions. Likewise, some combinations of correlations (such as positive correlations between the variations in the vital rates) caused a bigger discrepancy than others (such as zero correlations).

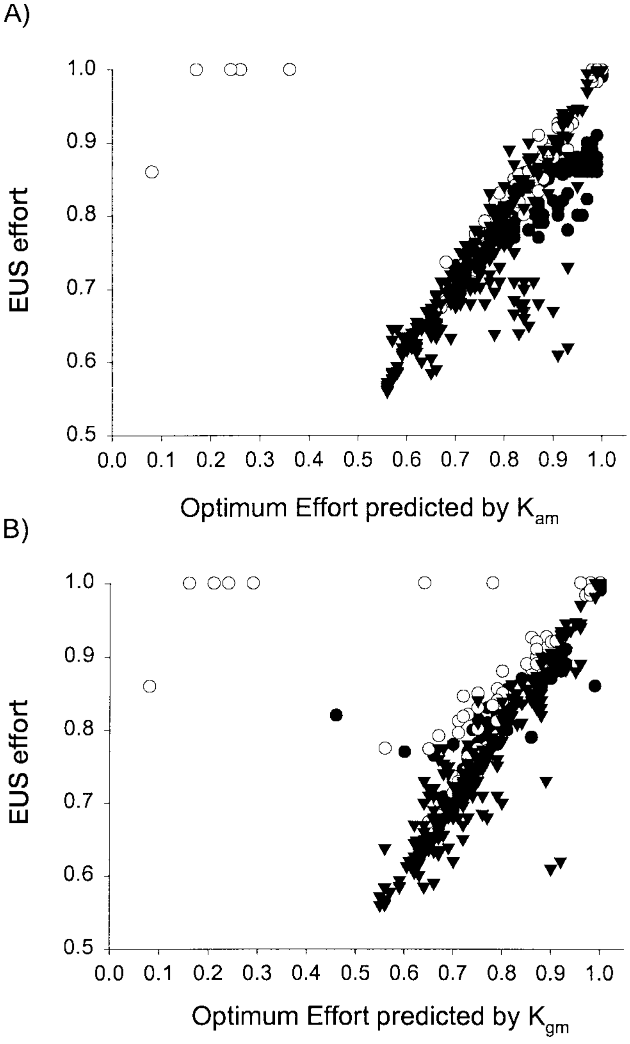

Fig. 5. Relationship between the optimum effort predicted by maximizing average population size (carrying capacity) and the EUS effort predicted using the invasion exponent. (A) Arithmetic mean population size calculated from 50,000 time steps ($n = 534$). (B) Geometric mean population size calculated from 50,000 time steps ($n = 514$). Other details as Fig. 2 legend.

DISCUSSION

We have shown that there is no one universal measure of fitness that natural selection can be said to optimize in all cases; instead, there are a range of measures whose results may generally approximate those of a more exact analysis based on invasibility criteria (the 'master fitness concept' of Metz *et al.*, 1996; Van Dooren and Metz, 1998). Similar conclusions, reached by different routes, have been made by Gurney and Middleton (1996) and Mylius and Diekmann (1995) and build on a growing body of literature specifying the need for invasibility analysis (e.g. Metz *et al.*, 1992; Rand *et al.*, 1994; Ferriere and Gatto, 1995).

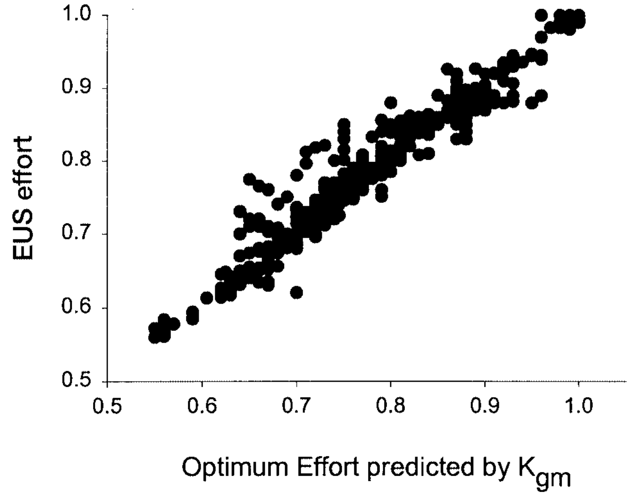

Fig. 6. Relationship between the optimum effort predicted by maximizing geometric mean population size (carrying capacity) and the EUS effort predicted using the invasion exponent. $n = 429$ (as Fig. 5b, but without data where density dependence is modelled using the Usher equation with sudden onset of density dependence). Usher-sudden-onset DD tends to cause equilibrium population size to vary (limit cycles or chaos). A comparison of Fig. 6 and Fig. 5b indicates that most outliers in Fig. 5b occur when underlying equilibrium population size is not constant.

Table 3. Analysis investigating the causes of divergence between optima predicted using invasion exponents and from optimizing different measures of fitness. Data generated using life histories 3, 4, 8 and 9 (Table 1)

Source	d.f.	Adj R^2	F	P
% Difference between optima predicted by K_{gm} and ϑ				
CV	1	14.0	32.54	0.000
Corr.	3	4.3	3.30	0.022
dp	2	11.8	13.79	0.000
Corr. *CV	3	10.2	7.95	0.000
Error	139	59.7		
Total	148			
% Difference between optima predicted by K_{am} and ϑ				
CV	1	4.7	17.95	0.000
Corr.	3	11.5	14.71	0.000
det EUS	1	23.7	90.69	0.000
Corr.*CV	3	11.7	14.88	0.000
Corr.*det EUS	3	11.5	14.66	0.000
Error	141	36.9		
Total	152			

% Difference between optima predicted by DI-R_0 and ϑ				
CV	1	3.0	21.27	0.000
Corr.	3	16.9	40.37	0.000
dp	2	2.8	10.22	0.000
det EUS	1	9.9	70.95	0.000
LH	3	23.6	56.54	0.000
Corr.*CV	3	1.2	2.81	0.042
CV*det EUS	1	3.6	26.17	0.000
Corr.*det EUS	3	17.7	42.36	0.000
dp*det EUS	2	2.4	8.47	0.000
Error	136	18.9		
Total	155			
% Difference between optima predicted by DI-*r* and ϑ				
CV	1	14.4	114.84	0.000
Corr.	3	5.9	15.84	0.000
LH	3	5.6	14.83	0.000
dp	2	4.8	19.22	0.000
det EUS	1	20.8	166.25	0.000
Corr.*CV	3	3.7	9.81	0.000
CV*det EUS	1	14.1	112.88	0.000
Corr.*det EUS	3	9.5	25.37	0.000
dp*det EUS	2	4.2	16.69	0.000
Error	136	17.0		
Total	155			
% Difference between optima predicted by DI-*a* and ϑ				
CV	1	14.2	109.81	0.000
Corr.	3	6.0	15.57	0.000
LH	3	4.8	12.28	0.000
dp	2	5.5	21.23	0.000
det EUS	1	20.2	156.26	0.000
Corr.*CV	3	3.1	7.89	0.000
CV*det EUS	1	14.0	107.88	0.000
Corr.*det EUS	3	9.9	25.57	0.000
dp*det EUS	2	4.6	17.94	0.000
Error	136	17.6		
Total	155			

Abbreviations: CV = coefficient of variation (covariate in models); Corr. = correlations between vital rates (four combinations: from all variations in vital rates covary with correlation coefficient = 1 to all variations in vital rates independent); dp = vital rates affected by density (Fs only, Ss only, or both); det EUS = deterministic optimum (as covariate); LH = life history.

Although there is recognition of the context-dependence of different measures of fitness (e.g. Stearns, 1992; Mylius and Diekmann, 1995; Metz *et al.*, 1996), many empirical biologists still use reproductive success as the default measure of fitness without proper consideration of its drawbacks. The argument that, in a growing population, offspring produced early in life are 'worth more' than those produced later is a very familiar one. However, in a density-dependent, fluctuating population, an analogous phenomenon occurs. Offspring born at a time when density-dependent influences on juveniles are weak are 'worth more' than offspring born at another time, when, for example, density-dependent mortality is higher. Similarly, offspring are 'worth more' if they themselves reproduce at a time when density dependence is weak (Grant and Benton, 2000). Where the value of an offspring varies markedly over time, LRS is likely to be a poor measure of fitness.

We do not claim that our analyses give an exhaustive representation of all life histories (for example, we study in detail only a two-class model, although similar results are found when the model is extended to more classes). However, within our analysis no single fitness measure universally agrees with the invasion analysis. The proportion of cases of agreement between optimization and invasibility analyses may be smaller or larger with a different choice of models but will be non-zero. In fact, because all our density-dependent functions here depend on the total number of individuals in the population (rather than the number in a subset of the population), the analysis overstates the usefulness of population size as a fitness measure (Grant and Benton, 2000).

Two issues arise from this analysis. First, is 'fitness' – a quantity maximized by natural selection – a meaningful concept? Second, is there a simple measure that 'works' and can be used by empiricists?

Is fitness a useful concept?

Although all existing measures of fitness are context-dependent, they 'should be judged by how well they do [the] job' (Stearns, 1992, p. 33). So although there is not a 'biological' measure of fitness which is 'right', fitness as a concept can still be useful because, in many cases, maximization of particular fitness measures will agree with the appropriate invasibility analysis for the same system. However, to maximize a fitness measure one has to ensure that it is the correct measure to use for that environment, and not ignore its assumptions. For example, the optimal strategy that maximizes LRS will coincide with the EUS, but only in a constant, density-dependent environment, and even then only under some circumstances (Mylius and Diekmann, 1995; Metz *et al.*, 1996). We find, under conditions when the equilibrium population size is constant, measures of population size (especially the geometric mean) are a good predictor of the EUS life history (see also Metz *et al.*, 1996). However, this will not necessarily be the case when density dependence is not a simple function of total population size (Grant and Benton, 2000). When vital rates depend on total population size, as here, and with moderate levels of population variability, the results obtained using population size as a fitness measure match almost exactly those obtained using ϑ. This is especially the case if density dependence affects both survivorships and fecundities. One has a choice of methods of measuring the population size, and the predictions will depend, to an extent, on which method is used. Here we used the arithmetic and geometric means. The geometric mean performs better, unless the population variability (caused by chaos or very high levels of stochasticity) is very large, in which case the arithmetic mean may be better. However, the other measures of fitness often used (r, a, R_0) are, in general, poor predictors of the EUS life history in density-dependent, stochastic environments (although in the right circumstances they may well be accurate predictors). In general, these measures are most accurate when the life histories are compared before the impact of density dependence and also when density dependence acts on both fecundities and survivorships.

This analysis highlights that, for a large range of organisms, predictions made using conventional measures of fitness *may* be inaccurate, sometimes massively so. Hence, a biologist finding apparently non-optimal behaviour in the study organism should perhaps first consider whether the fitness measure being used is appropriate.

Is there a fitness measure that is both simple and works?

The second question that arises is whether there is a 'simple' measure of fitness that can be used to predict the optimal life history, especially by field biologists. Estimation of the carrying capacity from field data is difficult (Dennis *et al.,* 1991) and impractical for measuring the fitness of different strategies.

Estimating reproductive success is much easier and is the most common fitness measure used in the field. Does it work? The answer to this will depend on the environment that the organisms in question experience. If the environment is more-or-less constant, if the equilibrium population size is constant, and if density-dependence affects fecundities alone, or fecundities and survivorships, then some measures of lifetime reproductive success may function tolerably well as predictors of the EUS life history. However, our results suggest that this will only really be the case if the life history can be estimated in the absence of density dependence. With many organisms, density dependence is likely to be strongest on survival of very young offspring (Sinclair, 1989), in which case the number of offspring born/laid should be used in an analysis, rather than the number of surviving offspring. In an environment where vital rates vary (whether the variability is due to environmental stochasticity or non-equilibrium population dynamics), survival of offspring may be so determined by the environment, rather than the intrinsic capabilities of the phenotype, that the number of surviving offspring becomes a poor estimator of the phenotype's ability. In many environments, stochasticity changes the EUS by ≤15% (Benton and Grant, 1999a) and as the optimum predicted by LRS is relatively insensitive to variation in population size, then LRS will approximate the EUS if density dependence acts on fecundities rather than survival. This may be sufficient accuracy. With data collected over a few years only and with small sample sizes, sampling error may make the estimates of optimum effort much more imprecise. Work is currently being undertaken to identify methods that will allow an assessment to be made of the performance of LRS relative to invasibility criteria for a given species, and so allow more exact prescriptions of when LRS can safely be used.

r- and *K*-selection

The measures of fitness that assume no density-dependent changes in the life histories (r and its stochastic analogue, a) are relatively inaccurate predictors of the optimal strategy in a density-dependent world; too often, as one might expect, they predict that an organism should be more semelparous. However, all organisms are likely to experience density dependence, at least occasionally, yet r 'has remained the fitness measure most frequently used in life history theory' (Stearns, 1992, p. 31). These results indicate that, if vital rates vary generally as a function of total population size, then organisms tend to experience K-selection (i.e. they are selected to maximize the number of individuals the environment can support) rather than r-selection (maximization of growth rate) (*sensu* MacArthur and Wilson, 1967). This appears to be the case even when the environment is highly variable. However, if the environment is more or less constant, but punctuated by occasional catastrophic years, the population might be reduced far enough below the carrying capacity for r-selection to have an effect and for r and a to become better approximators of fitness. Such a scenario was envisaged by Caswell (1982). Although most organisms will experience density dependence, even if only occasionally, it is possible that selection on genotypes is density independent even in populations that are strongly regulated (Prout, 1980). Thus, in an explicitly genetic analysis, r and a may perform somewhat more accurately than indicated here.

For a subset of life histories, in a subset of environments, we analysed the divergence between the optima predicted by the invasion exponents and the fitness measures. For the population-size fitness measures, the divergence between estimates of the optimum life history increases with the CV and is also strongly affected by the correlations between the variations in the vital rates. The accuracy of the other measures of fitness is affected by many other factors, including the life history, the deterministic optimum and the interactions. In fact, fuller analyses suggest that any factor that causes the optimum effort to change from the deterministic optimum affects the accuracy of these measures of fitness. Such additional factors include the model for density dependence, the value of the trade-off exponent, z, and the distribution from which the variation is drawn (Benton and Grant, 1999a).

In conclusion, different measures of fitness may produce widely differing predictions of the optimal life history, even in a deterministic environment. In stochastic, density-dependent environments, the context-dependent nature of fitness measures means some care should be taken in the choice of measure being used in an analysis. Explicit thought should be given to the nature of density dependence and the amount of environmental variability. If the study is of sufficient importance, various scenarios could be modelled with the approach taken here to ensure that a simple choice of fitness does not mislead.

ACKNOWLEDGEMENTS

We thank two anonymous referees and Larry Slobodkin for their constructive comments.

REFERENCES

Benton, T.G. and Grant, A. 1999a. Optimal reproductive effort in stochastic, density-dependent environments. *Evolution*, **53**: 677–688.

Benton, T.G. and Grant, A. 1999b. Elasticity analysis as an important tool in evolutionary and population ecology. *TREE*, **14**: 467–471.

Burkhardt, F., Porter, D.M., Browne, J. and Richmond, M. 1993. *The Correspondence of Charles Darwin. Vol. 8: 1860*. Cambridge: Cambridge University Press.

Caswell, H. 1982. Life history theory and the equilibrium status of populations. *Am. Nat.*, **120**: 317–339.

Caswell, H. 1989. *Matrix Population Models*. Sunderland, MA: Sinauer Associates.

Charlesworth, B. 1994. *Evolution in Age-structured Populations*, 2nd edn. Cambridge: Cambridge University Press.

Clutton-Brock, T.H., ed. 1988. *Reproductive Success*. Chicago, IL: University of Chicago Press.

Cooch, E.G. and Ricklefs, R.E. 1994. Do variable environments significantly influence optimal reproductive effort in birds? *Oikos*, **69**: 447–459.

Darwin, C.R. 1872. *On the Origin of Species*, 6th edn. London: John Murray.

Darwin, F., ed. 1887. *The Life and Letters of Charles Darwin*, Vol. 3. London: John Murray.

Dawkins, R. 1982. *The Extended Phenotype.* London: W.H. Freeman.

Dennis, B., Mulholland, P.L. and Scott, J.M. 1991. Estimation of growth and extinction parameters for endangered species. *Ecol. Monogr.*, **6**: 115–143.

Desmond, A. and Moore, J. 1991. *Darwin*. London: Michael Joseph.

Diekmann, O., Mylius, S.D. and ten Donkelaar, J.R. 1999. Saumon à la Kaitala et Getz, sauce hollandaise. *Evol. Ecol. Res.*, **1**: 261–275.

Ferriere, R. and Gatto, M. 1995. Lyapunov exponents and the mathematics of invasion in oscillatory or chaotic populations. *Theor. Pop. Biol.*, **48**: 126–171.

Foley, P. 1997. Extinction models for local populations. In *Metapopulation Biology: Ecology, Genetics and Evolution* (I.A. Hanski and M.E. Gilpin, eds), pp. 215–246. London: Academic Press.

Giske, J., Aksnes, D.L. and Førland, B. 1993. Variable generation times and Darwinian fitness measures. *Evol. Ecol.*, **7**: 223–239.

Grant, A. 1997. Selection pressures on vital rates in density dependent populations. *Proc. R. Soc. Lond. B*, **264**: 303–306.

Grant, A. and Benton, T.G. 2000. Elasticity analysis for density dependent populations in stochastic environments. *Ecology*, **81**: 680–693.

Gurney, W.S.C. and Middleton, D.A.J. 1996. Optimal resource allocation in a randomly varying environment. *Funct. Ecol.*, **10**: 602–212.

Kozlowski, J. 1993. Measuring fitness in life-history studies. *Trends Ecol. Evol.*, **8**: 84–85.

Lessells, C.M. and Boag, P.T. 1987. Unrepeatable repeatabilities – a common mistake. *Awk*, **104**: 116–121.

MacArthur, R.H. and Wilson, E.O. 1967. *The Theory of Island Biogeography.* Princeton, NJ: Princeton University Press.

Metz, J.A.J., Nisbet, R.M. and Geritz, S.A.H. 1992. How should we define fitness for general ecological scenarios? *Trends Ecol. Evol.*, **7**: 198–202.

Metz, J.A.J., Mylius, S.D. and Diekmann, O. 1996. When does evolution optimise? On the relation between types of density dependence and evolutionary stable life history parameters. IIASA Working Paper WP-96–04, International Institute for Applied Systems Analysis, Austria.

Mylius, S.D. and Diekmann, O. 1995. On evolutionarily stable life histories, optimization and the need to be specific about density dependence. *Oikos*, **74**: 218–224.

Prout, T. 1980. Some relationships between density-independent selection and density-dependent population growth. *Evol. Biol.*, **13**: 1–68.

Rand, D.A., Wilson, H.B. and McGlade, J.M. 1994. Dynamics and evolution: Evolutionarily stable attractors, invasion exponents and phenotype dynamics. *Phil. Trans. R. Soc. B*, **343**: 261–283.

Ricklefs, R.E. 1977. On the evolution of reproductive strategies in birds: Reproductive effort. *Am. Nat.*, **111**: 453–478.

Schaffer, W.M. 1974. Optimal reproductive effort in fluctuating environments. *Am. Nat.*, **108**: 783–790.

Sibly, R.M., Linton, L. and Calow, P. 1991. Testing life-cycle theory by computer simulation. II. Bet-hedging revisited. *Comp. Biol. Med.*, **21**: 357–367.

Sinclair, A.R.E. 1989. Population regulation in animals. In *Ecological Concepts* (J.M. Cherrett, ed.), pp. 197–241. Oxford: Blackwell.

Spencer, H. 1864. *The Principles of Biology*, Vol. 1. London: Williams and Norgate.

Stearns, S.C. 1992. *The Evolution of Life Histories*. Oxford: Oxford University Press.

Tuljapurkar, S.D. 1982. Population dynamics in variable environments. III. Evolutionary dynamics of *r* selection. *Theor. Pop. Biol.*, **21**: 141–165.

Tuljapurkar, S.D. 1989. An uncertain life: Demography in random environments I. *Theor. Pop. Biol.*, **35**: 227–294.

Tuljapurkar, S.D. and Orzack, S.H. 1980. Population dynamics in variable environments. I. Long-run growth rates and extinction. *Theor. Pop. Biol.*, **18**: 314–342.

Van Dooren, T.J.M. and Metz, J.A.J. 1998. Delayed maturation in temporally structured populations with non-equilibrium dynamics. *J. Evol. Biol.*, **11**: 41–62.