Occam's shadow: levels of analysis in evolutionary ecology—where to next?

EVAN G. COOCH[1], **EMMANUELLE CAM**[2] **& WILLIAM LINK**[2],
[1]*Department of Natural Resources, Cornell University, Ithaca, New York, USA and* [2]*Biological Resources Division, US Geological Survey, Patuxent Wildlife Research Center, Laurel, Maryland, USA*

ABSTRACT   *Evolutionary ecology is the study of evolutionary processes, and the ecological conditions that influence them. A fundamental paradigm underlying the study of evolution is natural selection. Although there are a variety of operational definitions for natural selection in the literature, perhaps the most general one is that which characterizes selection as the process whereby heritable variation in fitness associated with variation in one or more phenotypic traits leads to intergenerational change in the frequency distribution of those traits. The past 20 years have witnessed a marked increase in the precision and reliability of our ability to estimate one or more components of fitness and characterize natural selection in wild populations, owing particularly to significant advances in methods for analysis of data from marked individuals. In this paper, we focus on several issues that we believe are important considerations for the application and development of these methods in the context of addressing questions in evolutionary ecology. First, our traditional approach to estimation often rests upon analysis of aggregates of individuals, which in the wild may reflect increasingly non-random (selected) samples with respect to the trait(s) of interest. In some cases, analysis at the aggregate level, rather than the individual level, may obscure important patterns. While there are a growing number of analytical tools available to estimate parameters at the individual level, and which can cope (to varying degrees) with progressive selection of the sample, the advent of new methods does not reduce the need to consider carefully the appropriate level of analysis in the first place. Estimation should be motivated* a priori *by strong theoretical analysis. Doing so provides clear guidance, in terms of both (i) assisting in the identification of realistic and meaningful models to include in the candidate model set, and (ii) providing the appropriate context under which the results are interpreted. Second, while it is true that selection (as defined)*

Correspondence: E. G. Cooch, Department of Natural Resources, Cornell University, Ithaca, New York, USA.

operates at the level of the individual, the selection gradient is often (if not generally) conditional on the abundance of the population. As such, it may be important to consider estimating transition rates conditional on both the parameter values of the other individuals in the population (or at least their distribution), and population abundance. This will undoubtedly pose a considerable challenge, for both single- and multi-strata applications. It will also require renewed consideration of the estimation of abundance, especially for open populations. Thirdly, selection typically operates on dynamic, individually varying traits. Such estimation may require characterizing fitness in terms of individual plasticity in one or more state variables, constituting analysis of the norms of reaction of individuals to variable environments. This can be quite complex, especially for traits that are under facultative control. Recent work has indicated that the pattern of selection on such traits is conditional on the relative rates of movement among and frequency of spatially heterogeneous habitats, suggesting analyses of evolution of life histories in open populations can be misleading in some cases.

1 Introduction

If a population has variation among individuals in one or more traits affecting fitness, and if this variation is to some degree genetically heritable, then the distribution of these traits may change in predictable ways over generations. This, of course, is the process of natural selection (Endler, 1986). Perhaps the most fundamental task in evolutionary ecology is the estimation of differences in fitness amongst individuals. Over the past 20 years, we have developed growing expertise in the ability to estimate a variety of demographic parameters relevant to characterizing fitness (e.g. survival rate, recruitment rate), particularly using data collected from marked individuals in the wild (Endler, 1986; Manly, 1990). Schwarz & Seber (1999) recently completed a comprehensive summary of the current 'state of the field' of methods used in estimation from marked individuals. Earlier reviews by Lebreton *et al.* (1993), Clobert (1995) and Lebreton (1995) provide an additional perspective. Yet, despite significant advances in our ability to estimate various parameters, we believe there is reason to consider several *conceptual* issues relevant to the *application* of these methods to questions in evolutionary ecology.

At present, many of the estimation procedures currently in wide use rely on analysis of aggregations of individuals. In some cases, this is perhaps quite reasonable. Often, we group individuals based on an *a priori* expectation that they share a common (perhaps heritable) trait of interest, and that the groupings of individuals may successfully account for the major axes of variation in the data (e.g. allowing for differences among age classes). However, one of the classic generalizations in evolutionary ecology is that selection operates at the level of the individual, although the pattern of selection can frequently be modified by higher levels of organization (family units and other forms of social structure, for example, can often influence the fitness of an individual within the group; Reeve & Keller, 1999; Clobert 2001). Evolutionary ecology is implicitly concerned with heterogeneity, since selection operates on differences among individuals. Notably, many of the recent advances in estimation and modelling have been focused on accounting for sources of heterogeneity among individuals, in some instances by increasing the dimensionality (structure) of the models. However, in cases where the groupings based on *a priori* criterion are orthogonal to, or at least covary in unexpected ways with biologically relevant groupings, then aggregate-based techniques may not adequately address heterogeneity.

However, increasing technical capability to accommodate heterogeneity does not

reduce the need for careful consideration of the choice, and the interpretation, of a particular analytical method. In fact, we submit that the ability to estimate robustly a variety of parameters of potential interest to evolutionary ecologists should increase the degree of attention given to the underlying theoretical considerations motivating the study.

In this essay, we consider some of the significant analytical and conceptual challenges that we suspect will require careful consideration in the future. In particular, we focus on the general question of 'levels of analysis', with reference both to the increasingly sophisticated array of analytical techniques, which are available, and to underlying conceptual issues in evolutionary ecology. Our intention throughout is not to provide definitive answers, or exhaustive review, but rather to provoke renewed discussion of what we submit are some key issues, and to make a few general suggestions for future consideration.

2 Levels of analysis—seeing the trees despite the forest

The question of whether modelling and estimation should be focused on aggregates of individuals, or on the individuals themselves is very much at the heart of evolutionary ecology. In fact, this basic dichotomy (for purposes of convenience) also has strong parallels in modelling, and in the statistical theory of data analysis in general. The following example clearly illustrates several of the basic issues (described in Brown, 1995; pp. 16-17).

An odorous gas is released into a room, and individuals sitting in the room are asked to indicate when they first sense the gas. Those individuals sitting nearest to the point of the release will typically indicate detection first, followed by a 'wave' of detection-responses over time with increasing distance from the point of release. Of course, this pattern reflects the random diffusion of the gaseous particles from an area of high concentration (at the point of release) to an area of lower concentration.

Clearly, the pattern of movement of the gas could be assessed at a variety of levels. Each individual gas molecule took a specific path (over space and time), the process of which was governed by a number of factors, including interactions with other gas molecules. It might be possible to characterize the physical 'rules' governing those interactions, from which a 'mechanistic' model could be constructed to describe the dynamics of the 'population' of gas molecules under a specified set of conditions. Alternatively, the 'average' behaviour of individual molecules can be studied, and a 'phenomenological' model based on statistical mechanics can be derived to characterize the pattern, if not the process, of diffusion of the gas through the room.

Which is the 'appropriate level' for analysis—should diffusion of gases be considered (analysed, modelled) from the perspective of individual molecules, or from emergent properties at the population level? While the answer clearly depends on the question, at least to some degree, whatever decision is made has significant consequences in terms of the parameters we need to estimate, and the means by which they are estimated.

2.1 Occam's shadow—model complexity, parsimony, and individual di erences

'Everything should be as simple as possible, but not simpler'

Albert Einstein

'Science may be described as the art of systematic over-simplification'

Karl Popper

Many decisions concerning model complexity reflect this basic contrast between 'phenomenological' models and 'mechanistic' models; the former generally attempt to describe faithfully the dynamics at population or higher organizational scales, while the latter attempt to describe flexibly the dynamics at a variety of scales (Judson, 1994). The relative merits of either approach arguably rest on the degree of coupling of processes operating at different organizational scales. For example, some behavioural choices (e.g. where to forage) may have a strong 'bottom-up' influence on the long-term dynamics of the population, so a robust population model might need to incorporate these individual behavioural decisions. Alternatively, some processes at the individual level may be only weakly coupled to higher-level dynamics, and perhaps can be safely ignored (or abstracted into more phenomenological models; Caswell & John, 1992).

The distinction between phenomenological and mechanistic models can be viewed, at least structurally, as a question of the most appropriate degree of model complexity, for a given purpose. For example, comparatively simple phenomenological models may be more efficient (in some senses) than complex models in prediction of stationary multi-year time series of population dynamics (as in the example of modelling the diffusion of gaseous molecules described earlier). However, non-stationary dynamics generally require models that are more complex (Turchin, 1990; Botsford, 1992; Judson, 1994).

Perhaps the best-known example of this in ecology involves the modelling of density-dependent population dynamics. The ubiquity of density-dependence as a logical necessity in biological systems is beyond debate, although the form by which density operates on the population remains frustratingly elusive, and often vague. To some degree, this difficulty has motivated the adoption of a parsimonious approach to incorporating density dependence into models of population growth. In fact, the commonly used logistic (Verhulst) equation

$$\frac{\mathrm{d}N}{\mathrm{d}t} = rN\left(1 - \frac{N}{K}\right)$$

is merely the smallest number of terms of a Taylor expansion for $\mathrm{d}N/\mathrm{d}t$ that can satisfactorily characterize regulated population growth (Lotka, 1925), and thus is arguably the most parsimonious model describing the phenomenon of regulated growth.

However, despite the simplicity of this expression, the logistic growth model itself is not a particularly realistic one, and naïve application of it can lead to all sorts of spurious conclusions. The biggest weakness of the logistic model is that it is entirely phenomenological; it is entirely motivated by the abundance dynamics of the population, and contains no information whatsoever concerning the interaction of the population (at any organizational level) with the environment. For example, the 'carrying capacity' K contains no real information about the environment, but is simply the equilibrium density of the population under density-dependent feedback (Berryman, 1992). Further, the logistic equation is a first-order model; while such models may be adequate for characterizing the dynamics of populations that are regulated around a stable-point equilibrium (Turchin, 1993), they are inadequate for modelling multi-dimensional and generally non-stationary dynamics (Ginzberg *et al.*, 1992). As noted by Turchin (see also Botsford, 1992; Getz, 1993; Taneyhill, 1993), the addition of any 'mechanistic' feature describing individual behaviour will satisfy the need to accommodate higher dimensionality (Ginzberg *et al.*, 1992).

To some degree, this consideration is the conceptual motivation for the increase in interest in individually based models (Huston *et al.*, 1988; Botsford, 1992; Łomnicki, 1988,1992; Wilson, 1998). The addition of one or more dimensions in ecological models to accommodate differences among individuals has been strongly paralleled by, and indeed has perhaps partially motivated, increasing interest in analytical methods for 'accommodating' individual heterogeneity in estimation of various demographic parameters.

2.1.1 The danger within—ignoring heterogeneity. Perhaps the best-known example of clear differences between patterns described in aggregate and in individuals concerns senescence (Vaupel & Yashin, 1985a,b,c; Service, 2000). There is a very large body of literature devoted to trying to explain the presence of senescence— the progressive decline in one or more components of fitness with age, after some threshold age is reached (e.g. Hamilton, 1966; Partridge & Barton, 1993; Charlesworth, 1994). While senescence of several fitness traits has been demon-strated conclusively in the laboratory, there are few clear demonstrations in the wild (McDonald *et al.*, 1996). For example, consider the potential for senescent declines in survival. Often, there is little if any evidence of such senescent declines with age. The possible reasons for this are many, but the most obvious are that the sample of individuals living to older ages gets progressively smaller, so there is a basic statistical issue of diminishing sample sizes.

However, the concept of 'quality' can also be invoked as a 'reason' underlying the failure to find evidence of senescence. Suppose, for example, the population studied is composed of individuals with different 'baseline' survival rates (e.g. each individual has its own mortality risk; Service, 2000). The term 'baseline' is used to describe the individual's survival rate at the start of the study (at birth for example, or at the start of reproductive life if the analysis is restricted to individuals that recruited into the breeding segment of the population). Here, we assume that 'quality' reflects differences among individuals in baseline survival. Further, assume that survival decreases at older age (i.e. there is a senescent decline) in every individual. This can be visualized as a set of parallel curves describing variation in survival as a function of age, where each curves corresponds to one individual. Last, suppose the sample of individuals ringed as offspring represents a random sample with respect to bird 'quality'. As individuals with lower survival rates die earlier, the composition of the sample progressively changes with time (or age at the level of cohorts). The progressive increase in the proportion of individuals with higher baseline survival rates results in increased survival assessed in the aggregate. Depending on the initial proportions of individuals of different quality, the average survival rates in older individuals (this subsample is mostly composed of higher quality individuals) may be higher than the average survival rate corresponding to younger age-classes (where the sample includes a higher proportion of lower-quality individuals) even though the survival rate of each individual actually decreases at older age. In this situation, senescent declines expressed within individuals are masked when addressed in the heterogeneous aggregate. This phenomenon has motivated considerable efforts to describe the genuine influence of age on survival in human populations using frailty models (e.g. Manton & Stallard, 1981, 1984; Manton *et al.*, 1981; Trussell & Richard, 1985; Hougaard 1986, 1991). Interestingly, Vaupel & Yashin (1985b) noted that heterogeneity among individuals could lead to an apparent decrease in survival at the aggregate level in older age-classes (i.e. apparent senescent decline), while survival actually

increases at the individual level. The possibility that genotypes that do not senesce and others that do coexist in populations has been proposed by Carey *et al.* (1992) and Curtsinger *et al.* (1992).

Other examples are well known to most people working in evolutionary ecology. For example, the increase in fecundity with age is ubiquitous in many species of birds (and indeed is often invoked as a selective pressure favouring delayed breeding in some species—see below). Is the increase in fecundity with age a function of an increase within individuals, or is it an artefact of analysing a progressively 'selected' sample? The increase in fertility with age could reflect an increase in the proportion of good 'quality' birds in the population, which happen to also have higher fertility (Curio 1983; Wooller *et al.*, 1990; Forslund & Pärt, 1995). This hypothesis of differences in individual quality becomes important at several levels, and is an issue we will revisit at various points in this paper.

In addition to the simple example of detection of senescence detailed above, access to individual heterogeneity is important for several reasons. First, the range of variation in fitness components among individuals (after accounting for the influence of variables such as age, time or sex for example) is of particular interest in evolutionary ecology. As emphasized earlier, the question of heterogeneity can be addressed at several levels relevant to evolutionary ecology. For example, changes in variance in individual response (e.g. breeding performance) over the lifespan of the individual (Lin *et al.*, 1997) may provide insight into the influence of experience on reproduction for example. Second, failure to account for heterogeneity among individuals can lead to failure to detect the influence of covariates in statistical analyses (Bryk & Raudenbush, 1992). Similarly, failure to account for heterogeneous variances among groups in the trait addressed can have the same consequence. Referring to the question of detection of senescence, heterogeneous variances among groups (e.g. sex) in baseline survival rates for example may explain why senescence has been detected in some groups and not the others, as the strength of mortality selection may depend on the degree of heterogeneity among individuals.

2.2 The grit beneath the gloss—estimation in the face of heterogeneity

It is clear from the preceding discussion that analysis in aggregate can potentially obscure important patterns. In this section, we consider the problem of heterogeneity from a more technical perspective. We begin with analysis of a simple example involving differences among individuals in age of recruitment. We show that failure to account for such heterogeneity can lead to significantly biased estimates of survival in some cases. We follow this with a discussion of some approaches that have been suggested as possible solutions to individual differences.

2.2.1 Heterogeneity and estimator bias. In his review of various challenges in the analysis of data from marked individuals, Lebreton (1995) remarked that 'heterogeneity in capture probability tends to bias survival estimates to a relatively limited extent', although it was noted that such heterogeneity can profoundly impact estimates of population abundance. However, the impact of heterogeneity in capture rate (and in other parameters) can be important in some cases, particularly where the pattern of heterogeneity covaries with other demographic variables (for example, age).

Consider the following example. A sample of individual birds is marked as

nestlings. Conditional on surviving and returning to the sample area (with probability ϕ), there is some chance that the individual will be encountered on a subsequent occasion (probability p). Typically, the probability of a given encounter history is written in terms of these two probabilities only (Lebreton *et al.*, 1992).

However, in many species of birds, (i) individuals are sampled during breeding (when they are typically easier to catch), and (ii) only breeding individuals return to the sampling area. Ascension to breeding stage is a permanent developmental state transition, from non-breeder (i.e. pre-recruit) to breeder (i.e. recruit). The probability of recruitment is determined by an underlying latent parameter, a. Let a_x be the probability that an individual of age x makes a permanent state transition from a 'pre-recruit' (a bird that has never bred), to a 'recruit' (a bird that has bred at least once; this does not necessarily mean it will breed every year, merely that it has bred at least once). For simplicity, we assume that once a bird has been 'recruited', it breeds every year, and that all breeding birds have the same probability of capture, regardless of the age at which they started breeding (Clobert *et al.*, 1994). Thus, the probability of the initial encounter following marking depends on both surviving and being a recruit. For example, consider the encounter history '101'. The individual was marked on the first occasion, released, and was seen again on the third occasion. Assuming that ϕ_1 is independent of recruitment state, there are two different ways this history could be achieved (Clobert *et al.*, 1994; Pradel & Lebreton, 1999): (1) an individual survived the first interval with probability ϕ_1 (which we know must be true since it was seen on the final occasion), did not recruit (with probability $1 - a_1$) by the second occasion, and thus was not seen on the second occasion, did recruit on the final occasion (with probability a_2) and was seen on the final occasion (with probability p_3; thus, $\phi_1(1 - a_1)\phi_2 a_2 p_3$). Alternatively, (2) the individual could have survived the first interval, recruited on the second occasion (with probability a_1) but was not seen on the second occasion (with probability $1 - p_2$), survived the second interval (with probability ϕ_2) and was seen on the last occasion (with probability p_3; thus, $\phi_1 a_1(1 - p_2)\phi_2 p_3$). The overall probability of the encounter history '101' then is the sum of these two expressions, $\phi_1\phi_2 p_3[a_1(1 - p_2) + a_2(1 - a_1)]$.

The presence of recruits and non-recruits in the population constitutes a source of capture heterogeneity, and is clearly a violation of the assumption that all marked individuals are equally likely to be encountered, conditional on being alive. At minimum, the assumption would only hold for recruited individuals (the assumption that capture rate among recruits is independent on age of recruitment may be reasonable in some situations). To illustrate the potential impact of ignoring heterogeneity due to differences in age of recruitment, we simulated a data set where we assumed that a was constant over all years for all individuals in the population; $a = 0.3$. In other words, in any given year, an individual bird is recruited with probability $a = 0.3$. We also assumed that survival and 'true' recapture rate are constant over time ($\phi = 0.85$, $p = 0.75$), with no 'true' differences in either among 'age classes'. We simulated encounter histories for a single cohort of 10 000 individuals over seven occasions, and fit a series of models using the standard Cormack-Jolly-Seber (CJS) models to the data, following the basic model selection paradigm espoused by Burnham & Anderson (1998). We used the CJS parameterization since this approach is commonly used for birds marked as young. Our intent is to consider the degree to which estimates of survival might be biased using this approach whenever $a < 1$. The model set included model $\{\phi p\}$ (constant over time for both parameters), since it might be assumed (*a priori*) that this will

be the most parsimonious model given that both parameters were held constant over time, and that *a* also did not vary with time.

However, our analysis of the simulated data showed that a model where both survival and recapture rate were allowed to vary with time, $\{\phi_t p_t\}$, was the most parsimonious model in the candidate model set; no other model had any appreciable degree of support. Estimates for both ϕ_i and p_i are shown in Fig. 1(a).

Two things are notable. First, the 'apparent' recapture rates increase monotonically over time. Although the latent recruitment rate was fixed at $a = 0.3$, the increase in apparent recapture rate is expected, since (i) only individuals that have recruited are potentially catchable after marking, and (ii) there is an increasing proportion of recruited individuals within a single cohort over time (i.e. there are very few non-recruits by occasion 5 (i.e. age 5)). The recapture rate estimated for birds marked as young using the CJS approach is confounded by recruitment rate. Comparison of apparent recapture rates among age classes can (under certain assumptions) provide an *ad hoc* estimate of the parameter *a* (Clobert *et al.*, 1992, 1994; Pradel & Lebreton, 1999). Second, the estimated survival rate the year after marking is significantly negatively biased; it is approximately 25% lower than the 'true' value of 0.85. Further, the estimates for the second and following occasions tend to initially be higher than the true value of 0.85, eventually declining to approximately the true value by occasion 5. The magnitude of the negative bias for the first estimate, and the degree of 'overshoot' of the second estimate is inversely proportional to *a* (Fig. 1(b)).

2.2.2 A first step at accommodating heterogeneity—structural stratification.   Clearly, there is significant potential for estimator bias in some cases. The preceding example is particularly striking in that the simulated survival rates were in fact constant over time. For samples of individuals marked as young, we generally expect that survival over the first year after marking (offspring or juvenile survival) will be lower than survival in subsequent years (i.e. adult survival). In such cases, where heterogeneity among age classes is anticipated, it is standard practice to include a model allowing for such 'age' differences in survival. Such 'age' models (Pollock, 1981) are often used to accommodate the effects of heterogeneity among newly marked individuals, and represents what we refer to as a 'stratification approach'—the parameterization of the model is stratified to allow for specific forms of differences among individuals.

For 'true' age effects, the heterogeneity arises because of maturation—survival rates of young birds are generally different (lower) from survival rates of older birds. In such cases, the heterogeneity is temporal or developmental heterogeneity—and occurs within birds. In other cases, the heterogeneity occurs among individuals. In both cases, 'age' models can often be used to minimize the effects of heterogeneity (Burnham & Rexstad, 1993; Pradel, 1993; Pradel *et al.*, 1997; Prévot-Julliard *et al.*, 1998).

However, in our hypothetical example, such an 'age' model would, in fact, have been selected as the most parsimonious model, even though there were no actual age effects on survival in the data! The detected 'age difference' (which would have been consistent with biological expectations) in this case is actually an artefact of heterogeneity in the recruitment 'states' of the individuals in the sample—some individuals are recruits, while others are not. Unless this heterogeneity is explicitly modelled (e.g. Clobert *et al.*, 1994, 1994; Pradel & Lebreton, 1999; Spendelow
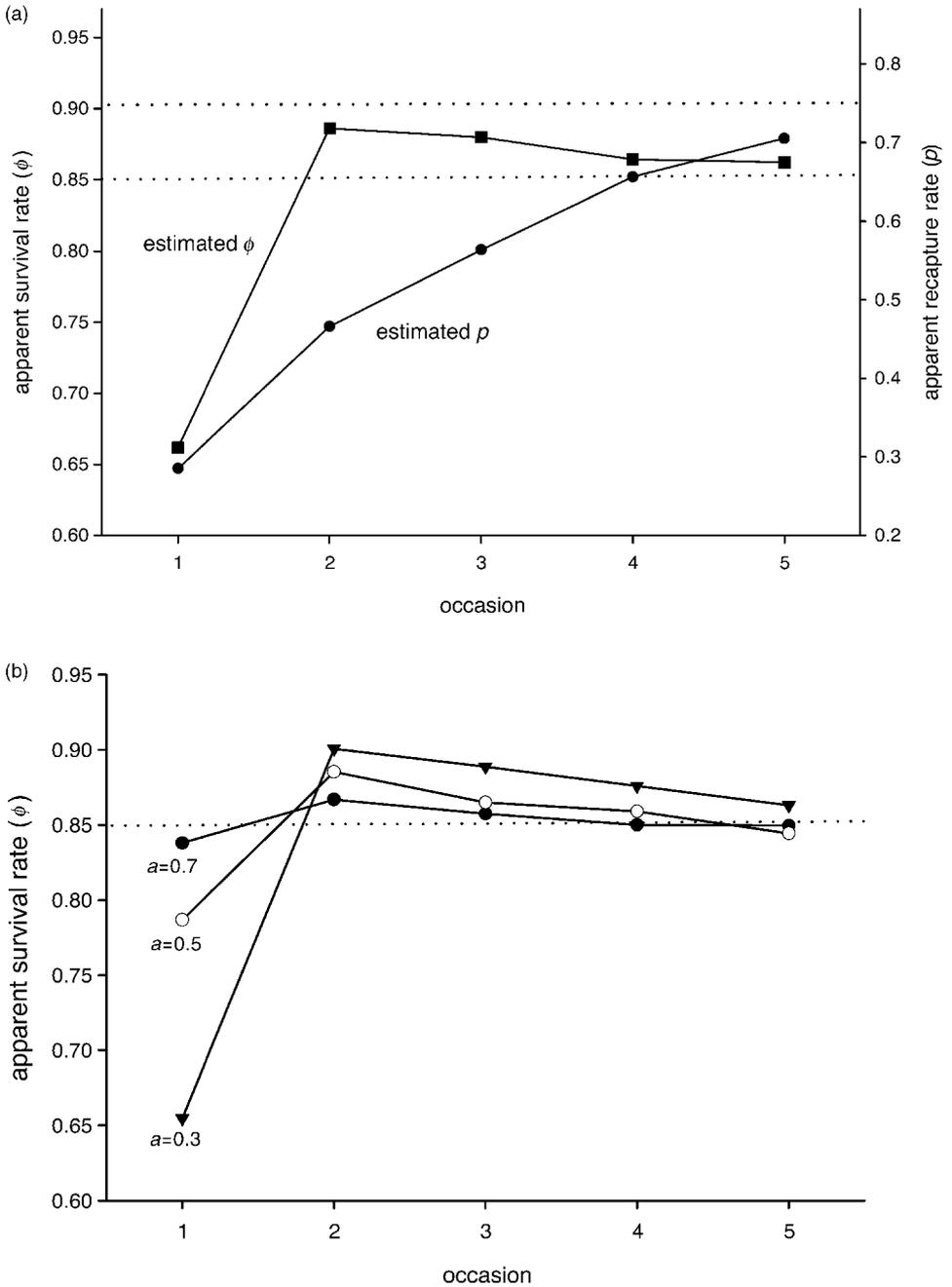
FIG. 1. (a) Estimated survival (ϕ) and recapture rate (p) for encounter data ($R = 10\,000$) simulated in the presence of variable probability of recruitment (a; details in text). Parameter estimates from time dependent Cormack-Jolly-Seber model. (b) Relationship between probability of recruitment (a) and the magnitude of the bias in estimates of apparent survival (ϕ) for the first interval following marking. Estimates derived from fitting time-dependent Cormack-Jolly-Seber model to encounter data dd($R = 10\,000$) simulated in the presence of variable probability of recruitment (a; details in text).

et al., 2002), biased estimates of one or more parameters are a possibility, as this example clearly shows.

2.2.3 Multi-strata models—a generalized stratification framework.   Recently, Lebreton *et al.* (1999) (see also Lebreton, 1995) suggested that a multi-strata approach (Arnason, 1973; Brownie *et al.*, 1993) may provide a unified conceptual framework for analysis of unseen events (e.g. recruitment), allowing efficient use of mixtures of information, significantly increasing the degree of stratification in the model. The key conceptual step in applying the multi-strata approach as a general framework is in recognizing that unseen events simply constitute different strata— individuals move in or out of a particular stratum or state, which may (or may not) be observable. In the present case, the multi-strata approach allows separate estimation of survival, recapture and movement between recruitment 'states'. The capacity to accommodate multiple strata in an analysis significantly improves the degree of stratification, since it allows explicit modelling of probabilistic transitions among individuals classified by 'state'.

In the present example, we wish to estimate the probabilities of survival, recapture, as well as the probability of permanently moving from a pre-recruit, non-breeding state to a recruited, breeding state (see Pradel & Lebreton, 1999, p. S76, for details). Fitting a multi-strata model to the simulated data yielded estimates of $\hat{\phi} = 0.849$, $\hat{p} = 0.750$, and $\hat{a} = 0.297$, which are essentially identical to the true parameter values. Our purpose here is not to suggest that multi-strata models will solve all identifiability issues (although they do have the potential to do so in many cases, as this example clearly indicates), merely to point out the potential problems that can result from not explicitly accommodating heterogeneity in the analysis. In this case, the heterogeneity can be accommodated by explicitly including a 'movement' parameter to specify the probability of the permanent state transition from 'pre-recruit' to 'recruit'.

However, despite the significant potential advantages in accommodating certain forms of heterogeneity, and the apparent success in this example, application of the multi-strata approach, and indeed, ultrastructral stratification in general, may often require specific assumptions which, ironically, may obscure potentially impor- tant differences among individuals. In the recruitment example, we were able to estimate survival rate despite heterogeneity in recruitment 'state' among individuals by applying a multi-strata approach, but only if survival rates did not differ between recruitment 'states' (i.e. that survival of a recruit and a pre-recruit over a particular interval were equal). As clearly noted by Pradel & Lebreton (1999), this is a particularly strong assumption, particularly in the context of evolutionary ecology, where the relative survival rate of breeders and non-breeders is central to many theories of the evolution of this trait (as discussed below). In addition, numerical convergence required setting the terminal value for *a* to 1.0 (i.e. assuming that full breeding is attained at some age). To some degree, both assumptions may simply represent a limitation of the current state of development of multi-state models (Lebreton *et al.*, 1999). However, they may also reflect fundamental limitations of this approach in certain applications (J. D. Nichols, personal communication).

2.2.4 Fork in the road—where to next? While in some cases the 'problem of heterogeneity' might be 'corrected' by improved generality (through increasing stratification) of model structure to data (see above), structural solutions may have significant technical limitations in many cases. The question of the influence of

heterogeneity and stratification of data in statistical analysis in general is extensively discussed in Cohen (1986).

Other approaches may be needed. One obvious candidate is the possibility of modelling survival as a function of one or more individual covariates (Hoffman & Skalski, 1995; White & Burnham, 1999). This would seem to be a logical solution (at some levels) to accommodating individual differences. However, as noted by Nichols (2001), use of individual covariates is simply a different type of stratification, since in effect it aggregates individuals as a function of the covariates included in the model.

Alternatively, there have been several notable advances in statistical inference methods to address individual heterogeneity motivated by questions relevant to other areas of research extensively using longitudinal data (mostly 'biomedical, educational, pharmacological, psychological, and sociological studies'; Morrel, 1998). These approaches are generally based on hierarchical random effects and frailty models accounting for unobserved heterogeneity, often executed in a Bayesian framework. Transferring such approaches to the typical situation where recapture probability has also to be taken into account is a challenge, but this has been envisioned by Lebreton (1995) as a promising idea (e.g. Dupuis, 1995). In addition, Pledger & Schwarz (2001) have also suggested the use of finite mixture methods as a method to allow for heterogeneity of survival (where survival probability is a random variable from a finite mixture; heterogeneity is thus modelled as a random effect). The finite mixture approach has the potential to help estimate the numbers of different 'types' of individuals in the sample, which is clearly an important advance.

However, while applications of these approaches hold considerable promise, they cannot eliminate the reliance on aggregates (Nichols, 2001). It may be that such a goal is, perhaps, unattainable, and, from a conceptual perspective, undesirable in the ultimate sense; some level of aggregation may be a prerequisite to conduct any robust scientific inquiry involving retrospective analysis (Nichols, 2001). If correct, then it may be that the most reasonable approach is to carefully evaluate the criterion used in allowing for differences among individuals in our analyses. However, while this might be the most robust approach in retrospective analysis, strong inference may ultimately require a manipulative experimental approach, to better delineate sources of variation among individuals (Nichols, 2001). We address both issues in the rest of this paper.

3 Grist for the evolutionary mill—heterogeneity and evolutionary ecology

In the preceding section, we discussed the issue of heterogeneity from a decidedly statistical perspective, considering both the potential implications of ignoring (or failing to accommodate) individual differences, as well as some possible analytical solutions. In this section, we consider the broader 'conceptual' role of heterogeneity in evolutionary ecology.

For the evolutionary ecologist, the ultimate context for estimation is the degree to which selection and the fitness differences upon which selection operates translate into evolutionary change. In the following, we briefly review the basic concepts underlying this larger purpose, concentrating on some of the fundamental processes that both contribute to heterogeneity among individuals, and condition the *interpretation* of fitness differences associated with this heterogeneity; in particular, the genetical context of the analysis of selection. Variation among individuals

reflects the interaction of the individual genotype with the environment, and consideration of fitness differences among individuals ultimately is conditioned on this relationship. For readers with extensive backgrounds in quantitative genetics, much of this presentation will be decidedly (and quite intentionally) simplified. Recent texts by Roff (1997) and Lynch & Walsh (1997) are comprehensive yet quite accessible treatments.

We follow this section with consideration of two examples of 'typical' problems in evolutionary ecology. In both cases, we focus on the relationship between estimation and interpretation, under the assumption that our ultimate concern is on the degree to which our results are relevant with respect to the evolution of one or more traits in the long-term. In particular, we consider aspects of the role of individual heterogeneity on optimization, selection, and the evolution of life histories.

3.1 Selection and heterogeneity—a short review of the underlying context

Much of our focus in terms of parameter estimation has been directed (at least implicitly) towards characterizing fitness differences associated with a particular phenotype (i.e. analysis of variation of one or more fitness components, see Endler, 1986; Manly, 1990). Such analyses are undoubtedly familiar to most biologists and, increasingly, among the statisticians recruited to assist with the analyses. For example, do large birds have higher survival, do older birds have higher fecundity, are species with greater levels of dispersal better able to track 'patchy' environments (either spatially or temporally)?

In effect, such analyses partition (at one or more levels) covariation between fitness and variation in the phenotype. We will assume for the moment that 'fitness' can be defined and assessed in some unambiguous fashion. The selection differential S is the difference between the mean phenotypes of the population before ($\bar{z}^\star$) and after the selection episode ($\bar{z}$):

$$S = \bar{z}^\star - \bar{z}$$

that, for quantitative traits, can be expressed as (Lande & Arnold, 1983)

$$S = \text{cov}(w, z)$$

where w is the scaled relative fitness ($w/\bar{w}$). Thus, the selection differential (i.e. the short-term result of selection) can be estimated as the slope of the regression of fitness w on the value of the phenotype z. Estimated in this fashion, the selection coefficient S refers to the degree of directional selection (increasing, or decreasing) on a particular trait. (Including the square of the deviations of the individual phenotypes z^2 from the mean permits assessment of change in phenotypic variance following selection—since the sum of the squared deviations from the mean is the variance.)

There are several well-known limitations to univariate selection analysis (Endler, 1986; Lynch & Walsh, 1997), but the most notable involves the problem of selection on correlated characters (Crespi, 1990). How do we know that selection is acting on the measured phenotype? In some (perhaps most) cases, there may be another unmeasured variable correlated with the first, such that the pattern of selection measured against the one trait is significantly biased by the presence of the other trait. The solution to this problem is clearly to account for the covariance among the characters, using a multivariate approach (Lande & Arnold, 1983). In

the multivariate case, the observed selection differential on a particular trait j, S_j, will be the sum of the direct and indirect effects of selection (where the indirect effects are mediated by the covariance among the characters).

$$S_j = \sum_i P_{ij}\beta_j$$

where P_{ij} are the elements of the phenotypic (not genotypic) covariance matrix. It is the phenotypic covariance matrix, **P**, that reflects the expression of the interaction of the genotype and the environment; this, ultimately, is the underlying source of heterogeneity—variation among individuals in their genetic constituency, and differences in the expression of those genes in different environments (Stearns, 1992; Roff, 1997; Lynch & Walsh, 1997).

Estimates of the β_j values follow

$$\mathbf{P}\beta = \mathbf{S}$$

$$\beta = \mathbf{P}^{-1}\mathbf{S}$$

which is simply the matrix version of multiple regression fitness w on trait z.

$$\beta = \mathbf{P}^{-1}\text{cov}(w, z)$$

However, demonstrating a statistically significant relationship between fitness and phenotypic variation is merely sufficient to demonstrate that natural selection is operating. This is a minimum, but not complete requirement for the evolution of a particular set of traits. In order for evolution to occur, fitness differences must be transmitted (i.e. inherited) among generations (Endler, 1986). Surprisingly, many studies make only passing reference to a 'presumed' (or assumed) genetic relationship between the phenotype and fitness. This oversight can severely complicate interpretation of the results from typical analysis of variation in fitness components.

In particular, responses to selection may be constrained by a lack of sufficient genetic variation, developmental constraints, or antagonistic pleiotropy among characters (e.g. Clark, 1987). This awareness has led to an increased appreciation of the importance of genetic architecture of the population upon which selection is operating. Genetic structure is typically described by the additive genetic variance-covariance matrix, **G**, which estimates the degree to which phenotypic (co)variation is a consequence of genotype by statistically partitioning total phenotypic variation and covariation into components attributable to additive genetic and environmental factors (Falconer, 1989; Lande, 1982; Lynch & Walsh, 1997).

In order to predict the expected change in the distribution of the phenotype over the long run (i.e. the pattern of evolution of the trait or traits), we need to have some estimate of the genetic variance-covariance matrix, **G** (where the diagonal elements are the additive genetic variances of the different characters, and the off-diagonal elements are the additive genetic covariances among different traits). Given an estimate of **G**, then the expected response to selection acting on the vector **z** ($\Delta\bar{\mathbf{z}}$, where the elements of **z** are the phenotypic values of the various phenotypic traits) is (see Lande, 1988; Lande & Arnold, 1983; Price & Grant, 1985; Lofsvold, 1986; references in Lynch & Walsh 1997).

$$\Delta\bar{\mathbf{z}} = \mathbf{G}\beta$$

$$= \mathbf{G}\mathbf{P}^{-1}s$$

More difficult, however, is the estimation of the structure of the genetic variance-covariance matrix **G**. Commonly, **G** is estimated as the genetic variances and pairwise genetic covariances of traits measured on adult organisms under defined environmental conditions (the complete structure of **G** would minimally include dimensions for ontogeny).

Obviously, it is simpler to estimate the phenotypic variance-covariance matrix, **P**. If **P** is, in fact, a good predictor of **G** (which it appears to be in some cases, especially for morphological traits; Cheverud, 1988; Stearns, 1992; Roff, 1994, 1997) then $\Delta \bar{z}$ can be estimated, at least in the short-run. Nonetheless, long-run predictions still require assumptions of constancy of **G**, which is certainly untenable at some scales. Unless at least the segregational component of **G** is variable over time, no phenotypic evolution is possible Turelli (1988).

Thus, estimates of **G**, or at least the degree to which **P** and **G** covary, are needed to assess the evolution of traits *over the long-run*. Within generation, we are focusing on the process of selection, and require only estimates of the phenotypic variance-covariance matrix, **P**. However, we must always be aware that estimating fitness differences is not the same as assessing the evolution of a trait. We generally accept that organisms can adapt to environmental factors in a number of ways. However, the degree to which any population can change and adapt to environmental variation is limited (ultimately) by the genetic variation in the population (which is influenced to varying degrees by selection, gene flow, mutation, and genetic drift), and the plasticity of interaction of individual genotypes with the environment. In order to characterize the degree to which phenotypic variation in a population is partitioned between genetic and environmental factors, which is fundamental if we are to be able to distinguish between proximate mechanisms (natural selection) and ultimate consequences (evolution), a large sample of marked individuals of known lineages must be followed, over multiple environments. Systematic variation in the expression of individual genotypes in different years can be shown analytically to reflect differences in the environmental component of the phenotype. Such changes can attenuate the response of the mean of a trait to directional selection (e.g. Cooke *et al.*, 1990; Frank & Slatkin, 1992). As such, to study the demography of a population fully, and to be able to interpret meaningfully the consequences of selection (and not simply the pattern and process), will require assessment of the degree to which various demographic traits reflect differences due to genetic or environmental factors (Endler, 1986; Stearns, 1992; Roff, 1997; Lynch & Walsh, 1997).

Of course, we are not suggesting that, in the absence of any information concerning the additive genetic covariance matrix **G**, there is no point in proceeding. Studies of the process of selection, robust demonstrations of differences in one or more transitions (vital rates), are crucial and, in some cases, estimates of the phenotypic correlation matrix alone may be sufficient, if applied cautiously (Cheverud, 1988). We are merely drawing attention to the fact that demonstrating selection, especially for a single population (on a local scale) is in and of itself only one part of a larger study of evolutionary ecology, a fact that should condition, *and often moderate*, the interpretation of the results. This may be especially relevant in studies attempting to demonstrate the presence or absence of 'trade-offs', a subject that we address later.

3.2 Informed use—estimation in the broader context

> 'I propose some simple mathematical models of.... Such models have an
> obvious air of unreality when compared to ... qualitative and verbal
> models ... They have the corresponding advantage of forcing one to make
> one's assumptions clearer. The purpose of mathematical formulation in
> this case is almost entirely to clarify the assumptions made.'
>
> *John Maynard Smith*

The role of theory and, in particular, the use of mathematical models as a way of
codifying 'what we know' or 'what we believe', has a long, and at times controversial
history in evolutionary ecology. Theory should *motivate* collection and analysis of
data, the results of which should then be used to revise theory, and so forth. This
process is succinctly described by Nichols (2000). The purpose of a model is an
important consideration. Often, with the purpose clearly in mind, theory can
provide a useful guide to specifying the necessity for estimating certain parameters,
and can often inform the selection of models for inclusion in the candidate
model set.

Given this, theoretical models have an important role in science in demonstrating
the possibilities of certain outcomes of a class of interactions. Such outcomes are
occasionally testable, allowing for refutation of the model (although testability of a
model is not its sole criterion for usefulness). Models are often used to predict the
magnitude and pattern of response of a system to the various processes included
in the model. They may tell us nothing about any individual system; however, they
can help us improve our understanding of the range of possible interactions between
the elements of the system.

In the following, we discuss in some detail two relatively simple examples of
situations where careful theoretical analysis can help guide our approach to
addressing problems in modelling and estimation. Each example represents a
typical analysis that relates, at some level at least, to issues of general interest in
evolutionary ecology. In particular, these examples are intended to demonstrate
the necessity for careful consideration of what parameters are being estimated, and
how the results are interpreted, especially in the larger context of individual
differences, interactions of individuals with their environment, and levels of analysis.

3.2.1 Example 1—age of first breeding. For our first example, we consider estima-
tion of what we will refer to as 'age of first breeding'. There has been considerable
interest in the technical challenges of methods of estimation of this parameter;
recent papers by Schwarz & Arnason (2000) and Pradel & Lebreton (1999) provide
excellent reviews of several important technical considerations.

General interest in age of first breeding can be traced to several fundamental
principles in life history theory. It is well known that the relative growth rate of an
age-structured population very much reflects the timing (schedule) of reproduction
across various age classes, except in stationary populations where the net reproduc-
tive rate adequately characterizes growth (we will assume that this is a rare event).
Early seminal work on the question of the optimal timing of the start of reproduction
can be attributed to Cole (1954). Since their early work, there has been a large
volume of literature devoted to theoretical explorations of age of first breeding (see
recent texts by Stearns, 1992 and Roff, 1992, for an introduction to much of
the relevant literature). Many of these investigations have considered species with
indeterminate growth, where fitness differences are often directly related to body

size. In such cases, timing of first breeding is often considered to reflect the optimiza-
tion of a *trade-o* of increased size (and potential fecundity) by delaying and growing
larger and the increased risks of mortality accrued by delaying (Sibly & Calow,
1983, 1984; Charnov, 1990; Stearns, 1992; Roff, 1992). Mertz (1971) showed that
in a growing population, any delay in maturation rate reduces λ, and will be selected
against in most cases, especially when pre-recruitment survival probability depends
on maturation rate (Caswell, 1989). In fact, increasing population growth strongly
favours individuals with reduced age of first breeding since a reduction in matura-
tion rate will tend to increase λ. In contrast, if $\lambda < 1$ (i.e. in a declining population),
the opposite is true; delaying maturation will act to slow the overall rate of decline
of population growth. Thus, consideration of how age of first breeding might change
due to selection must clearly consider the growth rate of the population (Mertz,
1971; Wiley, 1974). Estimates of realized growth rate of a population are readily
derived using standard mark-recapture approaches (Pradel, 1996; Nichols & Hines,
this volume). However, since it is doubtful that a population will be in decline for
a sufficiently long time such that age of first breeding will be selectively increased,
it is reasonable to conclude that under most conditions, *any* delay in maturation
will tend to reduce growth, and be selected against. This result is perhaps not
surprising: if an organism waits to breed, it will be more likely to suffer mortality
without ever having bred. Further, early breeding individuals contribute their genes
to future generations more rapidly than late breeding individuals; this compounding
will infer greater fitness if in fact age of first breeding is a heritable trait.

However, this is clearly at odds with the observation that in many taxa, age of
first breeding is significantly delayed beyond the age of physical maturation
(Williams, 1992; Charlesworth, 1994). This seems especially true among species
with high adult survival rates and low fecundity (Charnov, 1990, 2000). Why do
some individuals delay breeding beyond the age of physical maturity? For delayed
breeding to evolve, something else must occur; indeed, what is necessary for
delayed breeding to evolve is some 'trade-off', such that a delay in breeding is
compensated by a change in some other vital rate such that population growth is
actually increased by delaying reproduction. Mertz (1971; following Lewontin,
1965) and Wiley (1974) derived general conditions under which delayed breeding
can evolve in species where physical maturity is reached before breeding (see
Charlesworth, 1994, for a general review; see also Stearns, 1992 for a review of a
similar analysis for taxa with indeterminate growth, and Emlen, 1984, for the
simpler situation of semelparous birth).

We refer to the connection (trade-off or otherwise) between delayed breeding
and changes in other vital rates parenthetically, since the way in which these traits
are connected is important for understanding the general issue of the evolution of
a particular life history. In the absence of a 'trade-off' between delayed reproduction
and increases in other vital rates, delayed breeding is unlikely to evolve. Since
delayed breeding is evident in several taxa, it is perhaps *logical* to conclude that
some levels of trade-offs are operating.

In fact, this is the essence of the 'trade-off' approach to understanding life
histories—the assumption that the life history represents an optimal allocation of
resources among various activities, as reflected in the mortality and fertility
schedule. The principle of allocation (e.g. Levins, 1968; Calow, 1979) provides a
simple motivation for the presence of trade-offs when the traded traits compete for
a common and limited resource pool. When nutrient resources are limited,
resources used for one trait will not be available for another and vice versa. This

logic provides a theoretical justification for the costs of reproduction (Williams, 1966) as well as a possible context for pleiotropic genes to act on reproduction and growth (Stearns *et al.*, 1991).

Much of the support for trade-offs is based on an apparent logical necessity for such trade-offs to exist, as an explanation for the observed variation in life histories. In some cases, perhaps in general, there is a risk of tautology (logical consistency of trade-offs with life history variation is not proof of their existence). To a large degree, many studies on variation of age of first breeding (for example) are largely motivated by an implicit assumption of a trade-off between one or more 'vital rates', and age of first breeding. For example, a progressive increase in adult survival might select for an increase in age of first breeding, assuming a negative trade-off between the two traits (Emlen, 1984). Thus, given evidence of increasing adult survival, due to some extrinsic factors, it might be reasonable to predict a systematic change in age of first breeding. Such an expectation implicitly assumes some trade-off between adult survival and selection for age of first breeding.

However, there are several important considerations. Trade-offs *within* an individual must be true (Tuomi *et al.*, 1983; Emlen, 1984); if an individual is forced to expend greater energy on one activity, then this necessarily reduces the amount of energy available for another activity. However, this does not necessarily mean that trade-offs occur *among* individuals. This is important, since natural selection operates on the additive genetic covariance *among* individuals, not from correlations *within* individuals (i.e. the **G** matrix, discussed previously; Endler, 1986; Roff, 1997; Lynch & Walsh, 1997). There have been significant difficulties in demonstrating genetically-based trade-offs, especially in wild populations where estimates of the genetic structure of the population are often unknown (van Noordwijk & de Jong, 1986; Reznick, 1985, 1992; Stearns, 1992; but see Partridge, 1992; Viallefont *et al.*, 1995).

Further, can we differentiate between restraint (as an evolved 'strategy') and constraint (Maynard Smith *et al.*, 1985; Sih & Gleeson, 1995; Prévot-Julliard *et al.*, 1999)? Is a particular life history as expressed by a given individual the 'optimal fitness solution' for that individual (Berrigan & Koella, 1994), or evidence of an individual 'doing the best it can'? Is the variation observed in age of first breeding among individuals evidence of individual optimization, or individual constraint?

It seems clear that in the absence of an estimate of the additive genetic covariance between age of first breeding and other life history characters, it is difficult to make predictions about long-term evolution of the trait, much less to draw robust conclusions about the causes, and consequences of variation in this trait over time (Clark, 1987; Dorn & Mitchell-Olds, 1991). As argued by Reznick (1992), simple phenotypic correlations, as measured on a series of individuals or populations, or life history manipulation experiments (Nur, 1990), do not provide adequate estimates of the cost of reproduction (or other trade-offs), in the context of evolutionary theory (but see Partridge, 1992, for an alternative view). The former method considers both environmental and genetic effects (the phenotype is effectively the expressed interaction of the two), whereas the latter represents analysis of a fixed environmental effect, and is generally independent of the genetic background. Variation in resource abundance as well as in the *quality of individuals* (i.e. heterogeneity) may mask underlying genetic trade-offs (Bell & Koufopanou, 1986; van Noordwijk & de Jong, 1986). This clearly points towards the need to consider differences among individuals.

Criticisms regarding interpretation of purported trade-offs have been countered

in many instances by studies involving experimental manipulations. However, there are also well-known limits to the general utility of phenotypic manipulations, most notably, the difficulty of interpreting the results of a simple manipulation experiment relative to what might be optimal for a given individual. For example, it is possible to mechanically delay individual breeding, such that survival of artificially 'delayed' individuals can be compared to survival of a control group that, on average, bred earlier. However, at the individual level, we do not know if individuals that delayed breeding 'naturally' would have had lower survival than if they had bred earlier. (The related question of individual optimization is addressed in the next example.)

This issue is reviewed in Viallefont *et al.* (1995). In order to reveal both the environmental and genetic components of trade-offs, one should compare genetically identical individuals along environmental gradients as well as genetically different individuals under constant environmental conditions. As Reznick comments, 'the mechanism that underlies the response to a phenotypic manipulation may be different from the organism's capacity for evolutionary change. It is this difference. . . that is crucial'.

In short, estimates of variation in age of first breeding (or indeed, any life history parameter) need to be evaluated within a larger context; the ability to estimate a parameter does not mitigate the necessity for careful consideration of how the results are to be interpreted. A delay in the age of first breeding, for example, appears to be constrained to *evolve* only in the presence of trade-offs with measures of fecundity and survival. In the absence of unambiguous evidence for such trade-offs, how is adaptive variation in age of first breeding to be interpreted? It is critical to consider whether a study provides any evidence concerning restraint (as indicated by the evolution of the trait) or simple environmental (perhaps mediated by phylogenetic limits) constraint (which tells us something about the mechanism, but not the ultimate consequence) of variation in the parameter. In fact, this point is implicitly made by Nichols *et al.* (1994), who note that multi-state models may be well suited to estimation of the *immediate* costs of reproduction (see also Viallefont *et al.*, 1995). In fact, this particular paper is often cited as a good example of an approach to analysing trade-offs, in particular costs of reproduction. However, Nichols *et al.* point out '. . . the multi-state estimation models permit . . . estimation and testing for studies directed at *phenotypic* correlations'. They correctly note (citing Reznick, 1992) that such studies *may* be inadequate for evolutionary analysis, unless we have relevant *a priori* criteria for classification of individuals into 'quality' categories, or when combined with manipulative experiments. However, this may not solve the problem of interpretation (i.e. constraint versus optimal response), or the relevance of studies focusing on the phenotypic level only.

3.2.2 Example 2—state-dependent movements.   In the preceding example, the issue concerned inference and the relationship between trade-offs and estimates of age of first breeding in the presence of heterogeneity among individuals and environmental effects. In essence, the difficulty relates to the more general problem of individual optimization.

We consider this issue in the present example, which considers aspects of animal movement. Animal movement has been of particular interest to ecologists for many years, perhaps since the nascent stages of the development of ecology as a science. The challenges of characterizing patterns of movement, in terms of causes, and consequences, are often complicated in systems where individuals are not easily

observed, for one or more reasons. Clearly, such situations are familiar to many (if not most) practising ornithologists, who work with taxa characterized in many cases by the ability to move. Approaches to the analysis of animal movement using data from marked individuals have been recently reviewed by Nichols & Kaiser (1999), and references cited therein. We first outline some of the theoretical considerations that may be relevant to understanding why individual animals move. Specifically, we consider movement among foraging patches, in response to differences in individual physiological state. We then discuss some of these considerations within the context of estimation of movement rates.

Suppose an individual bird must choose between one of several different 'states'. These states could be either physical locations (e.g. nesting areas, or foraging patches), or reproductive states (e.g. breeding or non-breeding), or any other situation with two or more discrete states among which the organism must choose. For this example, we will assume that the states represent discrete physical locations (patches), among which an individual must choose where to reside at a particular time. Consider a small bird attempting to survive the winter period, which for convenience we will divide into a number of equal-length time units. At the beginning of each time unit, the individual bird needs to make a decision as to which habitat patch to reside in. Each patch is characterized by differences in relative food abundance, and the risk of mortality, perhaps due to predation. In many cases, a patch that has abundant food might also be riskier (for example, if predators focus on patches known to be attractive to birds because of abundant food; however, see below). A key application of evolutionary ecology to behavioural ecology is the expectation that natural selection will (in simplest form) optimize the temporal sequence of patch choices, such that on average most individuals will select patches that will maximize some fitness measure. The probability of survival over any given time period may be determined by the probability of starvation, and the risk of mortality due to predation, both of which are functions of which foraging patch is selected. In making a particular patch choice, the individual balances the competing risks of starvation due to lack of sufficient energy reserves, versus the chance that it will suffer mortality due to predation.

Body size (or, more accurately, physiological condition) is a dynamic variable—and (in this example) is a 'state' in which the animal finds itself at the start of each period. Thus, we are considering two different states simultaneously—the fixed 'patch' state among which the individual has to choose to reside, and the variable individual physiological 'state' (body condition) which motivates the patch selection made by each bird at each time step over winter. The optimal patch i at time t is that which maximizes $F(x, t, T)$, which for the present example is the probability that the bird survives from the start of period t to period T (where period T is the final time interval), given that at period t the forager is alive, and the value of the state variable (energy reserves) is x. The optimal sequence of patch selections can be solved analytically by application of a recursive dynamic programming algorithm, in a straightforward manner (see Mangel & Clark, 1988; Clark, 1993).

In general, if individuals are acting optimally (which we discuss later) then, for some individuals at least, the optimal strategy at a given time t may be to move among the various patches, with the pattern of movements conditioned by energy state, and time left in the season. Frequently, with increasing time to go until the time horizon is reached, the decision strategy becomes stationary for given energy states.

How can this 'optimality analysis' be useful in helping us understand the

movement and consequences of movement of individuals? Moreover, does such an approach inform our thinking about what parameters we should estimate, and how? Even consideration of simple patch selection models usefully highlights several key points. Specifically, we expect heterogeneity among individual birds in the pattern of movement at a given point in time, conditional on (i) individual energy state, (ii) the assumption that individuals are acting optimally for their given energy state, and (iii) that the decisions of any particular individual are made independently of actions of other individuals in the population. We discuss the first two conditions first.

Suppose we are interested in modelling the movements of birds, using data from marked individuals, where location is recorded for all living birds remaining in the sample area. The preceding analysis clearly suggests that there should be marked differences in the pattern of movement over the course of the season. Nevertheless, how can the implicit individual heterogeneity expected under state-dependent patch selection be accommodated analytically? As noted by Mangel & Clark (1988), since the process that is being modelled is stochastic, it is not possible for the researcher to predict the future state, or the actual sequence of decisions of any *individual*.

Are there any options available if, in fact, the individual state cannot be routinely assessed? In most situations, state cannot be assigned for individuals that are not encountered at a particular sampling occasion. Perhaps ironically, at this point, the simplest option is to test predictions at the population level, by using the model to derive predictions about the expected proportion of individuals in the population choosing a particular patch at a particular occasion.

Mangel & Clark (1988) describe a simple technique for doing this. They define a transition density $w(x, t|z)$, the probability that an organism at time t has energy state x, given (i) energy state z at $t - 1$, and (ii) given that the optimal patch is chosen between $t - 1$ and t. Given a set of optimal decisions $c'(x, t)$, where c' is the optimal patch to select in period t if the energy state at the start of the period is x, then

$$P(x, t) = \sum_z w(x, t|z) P(z, t - 1)$$

which can be iterated forwards in time given an initial density $P(x, 1)$. As noted by Mangel & Clark, this expression describes a non-stationary Markov chain (the chain is stationary if $w(x, t|z)$ is independent of t). The calculation of the transition matrix $w(x, t|z)$ proceeds in a straightforward manner (see Mangel & Clark, 1988, p. 78). Given $P(x, t)$, the fraction of organisms expected to choose patch i in period t, $f_i(t)$ can be calculated as the sum of the probabilities for values of the energy state variable for which patch i is optimal

$$f_i(t) = \sum_x P(x, t) \delta_{i, c'(x, t)}$$

where $\delta_{i, c'(x, t)}$ is a Kroenecker delta function where $\delta_{i, c'(x, t)} = 1$ if $c'(x, t) = i$, 0 otherwise. This approach is similar to the use of Markov chains and models for individual histories described by Lebreton (1995).

Clearly, the observed proportions of individuals choosing particular patches at each occasion will potentially differ from expectations for several possible reasons. Most obviously, the model itself may be an inadequate representation of reality. However, lack of fit may also indicate that the fitness of suboptimal strategies is not significantly lower than the fitness expected for optimal strategies, such that the strength of selection favouring the optimal patch selection strategy is low.

It therefore appears that one option is to estimate abundance at each patch, at each occasion, ignoring individual transition probabilities. As we discuss in the next section, the question of abundance estimation may be increasingly important. However, perhaps more importantly, this again calls into question the larger issue of the intent of the estimation of movement rates. Clearly, *individuals* should make specific movements over specific intervals, conditional on being in a particular physiological 'state' at the beginning of the interval. Suppose we are able to assess physiological 'state' for each individual at each occasion; following Nichols *et al.* (1994), a multi-strata approach might be used to determine whether individuals in poor condition (lower physiological 'state') have a greater propensity to move among patches, or lower survival, than do individuals in good condition. Nichols & Kendall (1995) suggested that the application of a multi-strata approach might be well suited to assessing the costs associated with particular behavioural decisions.

We feel there is some risk that this perspective is too limited in the larger context of selection on particular behaviours. If a particular movement is sub-optimal for a given bird, given its condition, then we expect there to be some 'cost' (as suggested by Nichols & Kendall, 1995). However, if the strength of selection against sub-optimal strategies is significant, then the likelihood of an *individual* bird making a sub-optimal movement will be low (which will be reflected in the frequency table of specific movements) and, by extension, there will be little evidence for trade-offs if indeed individuals are acting optimally. Viallefont *et al.* (1995) made essentially the same point, noting that 'actual trade-offs are certainly rarer than potential ones, because animals regulate their own reproductive success (*or, specifically, various actions which potentially a ect their individual fitness*) to avoid such costs, and because only one strategy is likely to be favored by natural selection.'

This raises the interesting question of selection and optimality, and the role of individual heterogeneity. The utility (if not relevancy) of optimality models is an issue that has been debated at length (Gould & Lewontin, 1979; Parker & Maynard Smith, 1990; Seger & Stubblefield, 1996). A trait is optimal if the strategy characterized by that trait (in this case, a particular pattern of movement) cannot be invaded by an alternative strategy (Orzack & Sober, 1994). A trait is locally optimal if predictions derived from the optimality model match the observations collected in the field (statistically), and all or almost all of the individuals exhibit the same trait. This condition is in fact the critical one; adaptation concerns variation at the individual level, and consideration of selection and optimality must consider individual heterogeneity (Orzack & Sober, 1994). Suppose that the abundance of individual animals in each patch matched closely the relative frequencies predicted by the optimization model (as described earlier; Mangel & Clark, 1988). Yet, without data concerning individual decisions, individual patterns of movement, it is not possible to conclude locally optimal behaviour—optimality is defined with respect to what an individual does (the essence of state-dependent modelling), not with respect to the average trait of a group of individuals (Orzack & Sober, 1994). To test the hypothesis of local optimality, multiple measurements from individuals are required. Heterogeneity among individuals would be consistent with the hypothesis that selection on the trait has not been particularly strong. As in our previous comments, our purpose here is simply to urge caution in the application of increasingly accessible methods for estimation of various parameters. Theory clearly advances by careful assessment of data collected in the field. However, theory must also inform data collection and analysis. While a failure to detect a cost of reproduction, for example, may be attributed to lack of robust

estimation techniques (Viallefont *et al.*, 1995), it may also reflect the fact that such negative trade-offs are unlikely to occur at all in some cases, especially under 'average conditions' (Stearns, 1992), except at the intraindividual level. Further, estimation is necessarily conditioned on the realized sequence of decisions; what may be more relevant in many cases are estimates of transition rates given a different sequence of decisions for that individual.

It is sometimes suggested that theory is based on empirical statements at a high level of generality. Often, this is regarded as an underlying reason for the palpable gulf between 'theoreticians' and 'empiricists', who in the latter instance have occasionally rejected theory since it fails to account adequately for the individual heterogeneity observed in the field. Nichols (2001) has suggested that some degree of generalization (and, by extension, aggregation) is necessary for science to progress. However, the two preceding examples point clearly to extensions of 'generalized theory' that are based specifically on individual optimization. In fact, it might be argued that the generality of theory is in fact increased by implicitly modelling individual decisions. Empirical tests of individually-based theory will depend to a large extent on our ability to estimate individual-specific rates, which may not be possible in all cases. Further, the covariance between environmental factors (cues) and fitness consequences of unrealized responses to those cues is rarely estimable (Williams & Nichols, 1984; Viallefont *et al.*, 1995). The critical factor may be in deciding the appropriate level of both theory and empirical test. As noted previously, in situations where individual decisions may be only weakly coupled to higher-level dynamics, such decisions might be safely ignored (or abstracted into more phenomenological models; Caswell & John, 1992). However, if this is not the case, then it may be important to consider methods for estimation of individual-specific rates, which may require considerable extension of current methodologies based on analysis of aggregates of individuals. Further, the estimation and interpretation of individual vital rates may be strongly conditioned by interactions with other individuals, and with the larger environment. We discuss some examples of this in the next section.

4 The complexities of life—some challenges for the future

> 'The universe is not only queerer than we imagine, but it is queerer than we can imagine'
>
> *J. B. S. Haldane*

4.1 *Variable states and reaction norms—the challenge of measuring selection in a heterogeneous world*

Life history theory and indeed evolutionary ecology in general, is largely the study of the pattern (and underlying causes) of the particular pattern of variation, at one of many different levels. The range of responses of individual organisms to various environmental factors will ultimately reflect the interaction of the genotype with a particular environment, and thus characterizing the pattern and consequences of selection must be conditioned on the particular form of this interaction. Variation in the phenotype at the population level reflects differences in the underlying genetics among individuals, as well as the degree to which the phenotype of each genotype varies in response to environmental change. This general variability in the phenotype of a given genotype in different environments is generally referred to as phenotypic plasticity. The range of variation in the expression of the genotype

in different environments is generally known as the norm of reaction (Stearns, 1992, and references therein). While for some traits the phenotype commonly varies among but not within individuals (i.e. the expression of the trait for an individual is fixed after some point in development), for others there is variation both among and within individuals; phenotypic expression for some traits may vary significantly (over some temporal scale) within individuals.

The dynamic state variable modelling paradigm discussed in the previous example rests very much on the idea that the expression of particular life-history traits depends upon the state of the external environment (e.g. Abrams & Rowe, 1996). In this sense, Clark & Mangel (2000) argue that dynamic state variable models are a robust generalization of traditional life-history theory (see also Clark, 1993).

If the phenotype is indeed the vehicle upon which selection operates (Reeve & Keller, 1999), and if phenotypic variation reflects differences in both the underlying genetics, as well as environmental differences, then analysis of selection without controlled experiments, and where the underlying genetic structure is unknown (i.e. some estimate of the genetic variance-covariance matrix $\mathbf{G}$) will be complicated by heterogeneous environments. At the extreme, it may suggest that inference beyond the local environment (i.e. perhaps a single population) is tenuous.

Originally, it was assumed that plasticity limited the degree of selection on the genotype by providing a 'buffer' in different environments, slowing the rate of evolutionary change (Levins, 1968). However, if plasticity is itself under selection, and if the norm of reaction is genetically based, then plasticity may evolve. In fact, phenotypic plasticity has been shown to be highly heritable (e.g. Mazer & Schick, 1991; Via *et al.*, 1995, and references cited therein). There is also increasing evidence that, under some conditions, the plasticity of a trait may evolve separately from the mean of the trait itself (for example, the variation in clutch size as a norm of reaction to different environmental conditions may evolve separately from the mean clutch size itself). Reaction norms are adaptations to heterogeneous habitat conditions experienced by offspring (Stearns & Koella, 1986; Thompson, 1991; Via *et al.*, 1995); thus, ancestors must experience a range of environments (i.e. offspring must disperse into new habitats with different environmental conditions) for plasticity to evolve.

Consideration of the optimal reaction norm among different habitats can markedly affect the reaction norms predicted by traditional life history theory. For example, delayed age of first breeding is generally expected in situations where there is high juvenile mortality (Emlen, 1984). However, Kawecki & Stearns (1993) show that this is only true if growth rate calculated over all habitats (Lebreton, 1996) is decreasing (consistent with Mertz, 1971). Unless this is true, then we should not necessarily expect delayed maturation in habitats with high juvenile mortality; such an expectation would depend upon the growth rate of the population as a whole. Kawecki & Stearns (1993) showed that under 'source-sink' conditions, where individuals may be expected to move from poor to better quality habitats, the optimal phenotype in a given habitat depends not only on conditions there, but is linked to the performance of other individuals in other habitats. Further, the performance of other individuals may be conditioned upon the abundance of individuals in the particular habitat (discussed below), and the frequency of habitats of differing qualities, such that the predicted optimal norm of reaction calculated over all habitats differs from the optimal life history for each habitat considered alone. Houston & McNamara (1992) derived essentially the same result, using a state-dynamics approach. Thus, in a spatially variable environment, fitness cannot be considered separately for each habitat, but must be calculated over all habitats

in which the organism might potentially find itself. Moreover, since most of the successful offspring will originate from the best habitats, and from habitats that are at high frequency, then natural selection in these habitats can be expected to be the strongest. Thus, adaptations should be more precise (i.e. there should be less genetic variation around the optimal norm of reaction) in better and more frequent habitats.

These results clearly point to the need to consider implicitly the population structure in any analysis of fitness and selection, beyond the obvious impacts of gene flow (Nichols & Kendall, 1995). It also suggests that the fitness of a particular adaptation in a given set of environmental conditions cannot be assessed precisely without knowledge of the conditions encountered by other members of the population. For example, in a given habitat, whether it is better to lay a clutch of two or three eggs may depend on the pattern and extent of dispersal of offspring among habitats, again suggesting strongly the need for estimates of juvenile dispersal rate. As noted by Kawecki & Stearns, 'the fitness of a phenotype expressed in a given habitat cannot be defined except in the context of the environmental variation encountered by the whole population. *This result emphasizes the need for thorough studies of natural source-sink populations and their reaction norms*' (italics ours).

4.2 Frequency- and density-dependence (or, why abundance matters)

While state-variable optimization models can significantly improve our understanding of sources of individual heterogeneity in many respects, several key assumptions are often made in general application. First, it is often assumed that the risk (and the probability of reward in terms of energy gain) due to selecting a particular strategy (say, choosing a particular patch) is fixed over time. Second, the optimal sequence of choices is derived for an individual, conditional on that individual's 'state' at a particular point in time. Finally, it is implicitly assumed that individuals act optimally (at least, on average). Even if an individual state can be assessed at each occasion, there are still several potential challenges in analysing the pattern of movement, related to these assumptions.

For example, while it seems reasonable that estimates of risk could be obtained using standard protocols, in many cases, the estimated probability of survival implicitly assumes that the risk of mortality for a given individual is independent of the risk of mortality for any other individual. Even if we simplify the situation and consider the survival of only males or females, minimizing (or eliminating) non-independence of mortality among paired individuals, there is still the possibility that risk of mortality is frequency-dependent. This is especially likely for risk associated with the probability of predation. In this case, calculating the risk faced by an individual is significantly complicated by frequency-dependence—estimates would necessarily have to condition on the number of conspecifics in the population over a given interval over which the individual was exposed to risk. In some cases, risk of mortality may also be density-dependent, reflecting increased intraspecific competition for other resources, which may increase risk of mortality. Further, it is also conceivable that other transitions are influenced by the number of individuals occupying each patch. In most metapopulation models, dispersal rate (or, the probability of moving between patches) is assumed to be density-independent. However, this is inconsistent with the evolutionary expectation that, in many cases, migration (dispersal) should in fact be density-dependent. If increasing density within a patch decreases the fitness of individuals remaining in the patch, then

fitness-dependent dispersal can be shown to have a strong stabilizing effect on population dynamics (Ruxton & Rohani, 1998). Again, it is perhaps worth considering estimation of movement rates among patches conditional on the size of the population in each patch, either explicitly within the estimation itself, or using abundance estimates as a covariate in the analysis (as a measure of habitat quality; see Nichols & Kendall, 1995).

While the stationary solution to this problem is related to concepts of the ideal free distribution (Fretwell & Lucas, 1970), at least in some cases, this does not adequately account for changes in the optimal strategy as a function of proximity to the terminal time-horizon. Clark & Mangel (2000) discuss a variety of ways in which frequency- and density-dependence can be incorporated into a dynamic programming framework (see also McNamara & Houston, 1996), although the solution is potentially difficult, since it involves calculating a dynamic ESS (evolutionarily stable strategy). However, to further complicate matters, the risk of predation does not have a simple ESS solution, since the predation risk also varies as a function of the distribution of predators. It is reasonable to assume that predators are as free to move among patches as their prey. Thus, the calculation of the optimal strategy given variation in both predator and prey movements must consider significant variation in the temporal dynamics of both species, as the predators follow the prey, and the prey attempt to evade the predators (Hugie & Dill, 1994; Weber, 1998; Lima & Bednekoff, 1999). One approach to accommodating the ephemeral nature of the temporally variable risk (or food supply) is to use a Bayesian updating approach (Clark & Mangel, 2000; Mangel, 1990), where the individual bird uses the previous experience to update its estimate of a particular parameter for a given patch. Not only does this consideration influence the optimal strategy for a given individual, but it will also clearly pose significant challenges for parameter estimation.

In addition, these general considerations of frequency- and density-dependence become even more complicated when considered in a higher-order context (e.g. metapopulation, community), and analysis of selection—both conceptually and technically—may increasingly require consideration of factors at a much broader scale than has been traditional (Reeve & Keller, 1999; Dieckman *et al.*, 2001; Clobert, 2002).

5 Summary

We are at an exciting, and perhaps critical, juncture in our ability to consider the pattern and process of natural selection in the wild, as part of broader studies in evolutionary ecology. The considerable advances in our ability to estimate a variety of important parameters, with recent consideration of methods for estimating individual latent parameters (in the context of frailty and random-effects models), have placed us on an unprecedented threshold. With this opportunity comes the need to consider carefully the appropriate level of analysis, particularly the general question of individual or aggregate, and the context in which the analysis is conducted in the first place. Not so long ago, much of the debate on several important issues in evolutionary ecology was cast in qualified statements reflecting our 'inability to estimate parameters of interest'. Such qualifications are increasingly moot, at least at some levels. This presents us with an important and critical decision—now that we increasingly have the analytical tools, are we finally at the point where we can adequately differentiate theory from tautology? The complexities of analysis in

wild populations—variable and state-dependent phenotypes, spatial heterogeneity, density- and frequency-dependence—pose considerable challenges for the future.

Acknowledgements

REFERENCES

ABRAMS, P. A. & ROWE, L. (1996) The effects of predation on the age and size of maturity of prey, *Evolution*, 50, pp. 1052-1061

ARNASON, A. N. (1973) The estimation of population size, migration rates and survival in a stratified population, *Research in Population Ecology*, 15, pp. 1-8.

BELL, G. & KOUFOPANOU, V. (1986) The cost of reproduction, *Oxford Surveys in Evolutionary Biology*, 3, pp. 83-131.

BERRIGAN, D. & KOELLA, J. C. (1994) The evolution of reaction norms: simple models for age and size at maturity, *Journal of Evolutionary Biology*, 7, pp. 549-566.

BERRYMAN, A. A. (1992) The logistic equation, *Trends in Ecology and Evolution*, 7, p. 316.

BOTSFORD, L. W. (1992) Individual state structure in population models. *In*: D. L. DEANGELIS & L. J. GROSS (Eds) *Individual-based Models and Approaches in Ecology*, pp. 213-236 (New York, Chapman & Hall).

BROWN, J. H. (1995) *Macroecology* (Chicago, University of Chicago Press).

BROWNIE, C., HINES, J. E., NICHOLS, J. D., POLLOCK, K. H. & HESTBECK, J. B. (1993) Capture-recapture studies for multiple strata including non-Markovian transition probabilities, *Biometrics*, 49, pp. 1173-1187.

BRYK, A. S. & RAUDENBUSH, S. W. (1992) *Hierarchical Linear Models: Applications and Data Analysis Methods* (Newbury Park, Sage Publications).

BURNHAM, K. P. & ANDERSON, D. R. (1998) *Model Selection and Inference—A Practical Information-Theoretic Approach* (New York, Springer-Verlag)

BURNHAM, K. P. & REXSTAD, E. A. (1993) Modeling heterogeneity in survival rates of banded waterfowl, *Biometrics*, 49, pp. 1194-1208.

CALOW, P. (1979) The cost of reproduction—a physiological approach, *Biological Review* 54, pp. 23-40.

CAREY, J. R., LIEDO, P., OROZCO, D. & VAUPEL, J. W. (1992) Slowing of mortality rates at older ages in large medfly cohorts, *Science*, 258, pp. 457-461.

CASWELL, H. (1989) *Matrix Population Models* (Sunderland, Sinauer Press).

CASWELL, H. & JOHN, A. M. (1992) From the individual to the population in demographic models. *In*: DEANGELIS, D. L. & GROSS, L. J. (Eds) *Individual-based Models and Approaches in Ecology*, pp. 36-61 (New York, Chapman & Hall)

CHARLESWORTH, B. (1994) *Evolution in age-structured populations*, 2nd edn (Cambridge, Cambridge University Press).

CHARNOV, E. L. (1990) On evolution of age at maturity and the adult lifespan, *Evolutionary Ecology*, 3, pp. 139-144.

CHARNOV, E. L. (2000) Evolution of life-history variation among species of altricial birds, *Evolutionary Ecology Research*, 2, pp. 365-373.

CHEVERUD, J. M. (1988) A comparison of genetic and phenotypic correlations, *Evolution*, 42, pp. 958-968.

CLARK, A. G. (1987) Genetic correlations: the quantitative genetics of evolutionary constraints. *In*: V. LOESCHKE (Ed.) *Genetic constraints on adaptive evolution*, pp. 25-45 (Berlin, Springer-Verlag)

CLARK, C. W. (1993) Dynamic models of behavior: an extension of life history theory, *Trends in Ecology and Evolution*, 8, pp. 205-209.

CLARK, C. W. & MANGEL, M. (2000) *Dynamic State Variable Models in Ecology—Methods and Applications* (Oxford, Oxford University Press).

CLOBERT, J. (1995) Capture-recapture and evolutionary ecology: a difficult wedding, *Journal of Applied Statistics*, 22, pp. 989-1008.

CLOBERT, J. (2002) Capture-recapture and evolutionary ecology: further comments. *In*: B. J. T. MORGAN & D. L. THOMSON (Eds) *Statistical Analysis of Data From Marked Bird Populations. Journal of Applied Statistics*, this issue.

CLOBERT, J., JULLIARD, R. & MCCLEERY, R. H. (1993) The components of local recruitment. *In*: J.-D. LEBRETON & P. M. NORTH (Eds), pp. 281-293 (Switzerland, Birkhäuser-Verlag).

CLOBERT, J., LEBRETON, J.-D., ALLAINE, D. & GAILLARD, J. M. (1994) The estimation of age-specific breeding probabilities from recaptures or resightings of marked animals. II. Longitudinal models, *Biometrics*, 50, pp. 375-387.

COHEN, J. E. (1986) An uncertainty principle in demography and the unisex issue. *American Statistician*, 40, pp. 32-39

COLE, L. C. (1954) The population consequences of life history phenomena, *Quarterly Review of Biology*, 29, pp. 103-137.

COOKE, F., TAYLOR, P. D., FRANCIS, C. M. & ROCKWELL, R. F. (1990) Directional selection and clutch size in birds, *American Naturalist*, 136, pp. 261-267.

CRESPI, B. J. (1990) The measurement of selection on phenotypic interaction systems, *American Naturalist*, 135, pp. 32-47.

CURIO, E. (1983) Why do young birds reproduce less well?, *Ibis*, 125, pp. 400-404.

CURTSINGER, J. W., FUKUI, H. H., TOWNSEND, D. R. & VAUPEL, J. W. (1992) Demography of genotypes: failure of the limited life-span paradigm in *Drosophila melanogaster*, *Science*, 258, pp. 585-599.

DIECKMAN, U., LAW, R. & METZ, J. A. J. (2000) *The Geometry of Ecological Interactions* (Cambridge, Cambridge University Press).

DIECKMAN U. & METZ, J. A. J. (2001) *Elements of Adaptive Dynamics (Cambridge Studies in Adaptive Dynamics)* (Cambridge, Cambridge University Press).

DORN, L. A. & MITCHELL-OLDS, T. (1991) Genetics of *Brassica campestris* 1. Genetic constraints on evolution of life-history characters, *Evolution*, 45, pp. 371-379.

DUPUIS, J. (1995) Bayesian estimation of movement and survival probabilities from capture-recapture data, *Biometrika*, 82, pp. 761-772.

EMLEN, J. M. (1984) *Population Biology—The Coevolution of Population Dynamics and Behavior* (New York, Macmillan).

ENDLER, J. A. (1986) *Natural Selection in the Wild* (Princeton, Princeton University Press).

FALCONER, D. S. (1989) *Introduction to Quantitative Genetics* (New York, Wiley).

FORSLUND, P. & PÄRT, T. (1995) Age and reproduction in birds—hypotheses and tests, *Trends in Ecology and Evolution*, 10, 374-378.

FRANK, S. A. & SLATKIN, M. (1992) Fisher's fundamental theorem of natural selection, *Tree*, 7(3), 92-95.

FRETWELL, S. D. & LUCAS, H. L., JR. (1970) On territorial behaviour and other factors influencing habitat distribution in birds. I. Theoretical development, *Acta Biotheoretica*, 19, pp. 16-36.

GETZ, W. M. (1993) The logistic equation revisited: final instalment, *Trends in Ecology and Evolution*, 8, pp. 69-70.

GINZBURG, L., WATKINSON, A., MACKENZIE, A., YOUNG, T. P., BERRYMAN, A. A. & OLSON, M. H. (1992) Intuition and the logistic equation (letters and reply), *Trends in Ecology and Evolution*, 7, pp. 314-317.

GOULD, S. J. & LEWONTIN, R. C. (1979) The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme, *Proceedings of the Royal Society* (London, B), 205, 581-598.

HAMILTON, W. D. (1966) The moulding of senescence by natural selection, *Journal of Theoretical Biology*, 12, pp. 12-45.

HOFFMANN, A. & SKALSKI, J. R. (1995) Inferential properties of an individual-based survival model using release-recapture data: sample size, validity and power, *Journal of Applied Statistics*, 22, pp. 579-595.

HOUGAARD, P. (1986) Survival models for heterogeneous populations derived from stable distributions, *Biometrics*, 73, pp. 387-396.

HOUGAARD, P. (1991) Modelling heterogeneity in survival data, *Journal of Applied Probability*, 28, pp. 695-701.

HOUSTON, A. I. & MCNAMARA, J. M. (1992) Phenotypic plasticity as a state-dependent life history decision, *Evolutionary Ecology*, 6, pp. 243-253.

HUGIE, D. M. & DILL, L. M. (1994) Fish and game: a game theoretic approach to habitat selection by predators and prey, *Journal of Fish Biology*, 45, pp. 151-169.

HUSTON, M. A., DEANGELIS, D. L. & POST, W. M. (1988) New computer models unify ecological theory, *BioScience*, 38, pp. 682-691.

JUDSON, O. P. (1994) The rise of the individual-based model in ecology, *Trends in Ecology and Evolution*, 9, pp. 9-14.

KAWECKI, T. J. & STEARNS, S. C. (1993) The evolution of life histories in spatially heterogeneous environments: optimal reaction norms revisited, *Evolutionary Ecology*, 7, pp. 155-174.

LANDE, R. (1982) A quantitative genetic theory of life-history evolution, *Ecology*, 63, pp. 607-615.

LANDE, R. (1988) Genetics and demography in biological conservation, *Science*, 241, pp. 1455-1460.

LANDE, R. & ARNOLD, S. J. (1983) The measurement of selection on correlated characters, *Evolution*, 37, pp. 1210-1226.

LEBRETON, J.-D. (1995) The future of population dynamics studies using marked individuals: a statistician's perspective, *Journal of Applied Statistics*, 22, pp. 1009-1030.

LEBRETON, J.-D. (1996) Demographic models for subdivided populations: the renewal equation approach, *Theoretical Population Biology*, 49, pp. 291-313.

LEBRETON, J.-D., BURNHAM, K. P., CLOBERT, J. & ANDERSON, D. R. (1992) Modeling survival and testing biological hypotheses using marked animals: a unified approach, *Ecological Monographs*, 62, pp. 67-118.

LEBRETON, J.-D., PRADEL, R. & CLOBERT, J. (1993) The statistical analysis of survival in animal populations, *Trends in Ecology and Evolution*, 8, pp. 91-95.

LEBRETON, J.-D., ALMERAS, T. & PRADEL, R. (1999) Competing events, mixtures of information and multistratum recapture models, *Bird Study*, 46, 39-46.

LEVINS, R. (1968) *Evolution in Changing Environments* (Princeton, Princeton University Press).

LEWONTIN, R. C. (1965) Selection for colonizing ability. *In*: H. G. BAKER & G. L. STEBBINS (Eds), *The Genetics of Colonizing Species*, pp. 79-94 (New York, Academic Press).

LIMA, S. L. & BEDNEKOFF, P. A. (1999) Temporal variation in danger drives antipredator behavior: the predation risk allocation hypothesis, *American Naturalist*, 153, pp. 649-659.

LIN, X., RAZ, J. & HARLOW, S. D. (1997) Linear mixed models with heterogeneous within-cluster variances, *Biometrics*, 53, pp. 910-923.

LINK, W. A. (1999) Modeling pattern in collections of parameters, *Journal of Wildlife Management*, 63, pp. 1017-1027.

LOFSVOLD, D. (1986) Quantitative genetics of morphological variation in Peromyscus. 1. Tests of the homogeneity of the genetic covariance structure among species and sub-species, *Evolution*, 40, pp. 559-573.

ŁOMNICKI, A. (1988) *Population Ecology of Individuals* (Princeton, Princeton University Press).

ŁOMNICKI, A. (1992) Population ecology from the individual perspective. *In*: D. L. DEANGELIS & L. J. GROSS (Eds), *Individual-based Models and Approaches in Ecology*, pp. 3-17 (New York, Chapman & Hall).

LOTKA, A. J. (1925) *Elements of Physical Biology* (Baltimore, Williams and Wilkins).

LYNCH, M. & WALSH, B. (1997) *Genetics and Analysis of Quantitative Traits* (Sinauer Associates).

MANGEL, M. (1990) Dynamic information in uncertain and changing worlds, *Journal of Theoretical Biology*, 146, pp. 317-332.

MANGEL, M. & CLARK, C. W. (1988) *Dynamic Modeling in Behavioral Ecology* (Princeton, Princeton University Press).

MANLY, B. F. J. (1990) *The Statistics of Natural Selection* (New York, Chapman & Hall).

MANTON, K. G. & STALLARD, E. (1981) Methods for evaluating the heterogeneity of aging processes in human populations using vital statistics data: explaining the black/white mortality crossover by a model of mortality selection, *Human Biology*, 53, pp. 47-67.

MANTON, K. G. & STALLARD, E. (1984) Heterogeneity and its effects on mortality measurement. *In*: J. VALLIN, J. H. POLLARD & L. HELINGMAN (Eds), *Methodologies for the Collection and Analysis of Mortality Data. International Union for the Scientific Study of Populations* (Liege, Belgium, Ordins Editions).

MANTON, K. G., STALLARD, E. & VAUPEL, J. W. (1981) Methods for comparing the mortality experience of heterogeneous populations, *Demography*, 18, pp. 389-410.

MAYNARD SMITH, J., BURIAN, R., KAUFFMAN, S., ALBERCH, P., CAMPBELL, J., GOODWIN, B., LANDE, R. & RAUP, R. (1985) Developmental constraints and evolution, *The Quarterly Review of Biology*, 60, pp. 265-287.

MAZER, S. J. & SCHICK, C. T. (1991) Constancy of population parameters for life-history and floral traits in Raphanus sativus II. effects of planting density on phenotype and heritability estimates, *Evolution*, 45, pp. 1888-1907.

MCDONALD, D. B., FITZPATRICK, J. W. & WOOLFENDEN, G. E. (1996) Actuarial senescence and demographic heterogeneity in the Florida scrub jay, *Ecology*, 77, pp. 2373-2381.

MCNAMARA, J. M. & HOUSTON, A. I. (1996) State-dependent lift histories, *Nature*, 380, pp. 215-221.

MERTZ, D. B. (1971) Life history phenomena in increasing and decreasing populations. *In*: G. P. PATIL, E. C. PIELOU & W. E. WATERS (Eds), *Statistical Ecology* (Volume 2), pp. 361-399 (University Park, Pennsylvania State University Press).

MORREL, C. H. (1998) Likelihood ratio testing of variance components in the linear mixed-effects model using restricted maximum likelihood, *Biometrics*, 54, pp. 1560-1568.

NICHOLS, J. D. (2000) On the use of models in the conduct of science and management of natural resources. *In*: T. SHENK and A. FRANKLIN (Eds) *Modeling in Natural Resource Management: Development, Interpretation and Application* (Covelo, CA, Island Press) (in press).

NICHOLS, J. D. (2001) Discussion comments on: 'Occam's shadow: levels of analysis in evolutionary ecology—where to next?'. *In*: B. J. T. MORGAN & D. L. THOMSON (eds) *Statistical Analysis of Data From Marked Bird Populations. Journal of Applied Statistics*, in press.

NICHOLS, J. D. & KENDALL, W. L. (1995) The use of multi-state capture-recapture models to address questions in evolutionary ecology, *Journal of Applied Statistics*, 22, pp. 835-846.

NICHOLS, J. D. & KAISER, A. (1999) Quantitative studies of bird movement: a methodological review, *Bird Study*, 46, pp. 279-288.

NICHOLS, J. D., HINES, J. E., POLLOCK, K. H., HINZ, R. L. & LINK, W. A. (1994) Estimating breeding proportions and testing hypotheses about costs of reproduction with capture-recapture data, *Ecology*, 73, pp. 2052-2065.

NUR, N. (1990) The cost of reproduction in birds: evaluating the evidence from manipulative and non-manipulative studies. *In*: J. BLONDEL, A. GOSLER, J.-D. LEBRETON & R. MCCLEERY (Eds), *Population Biology of Passerine Birds: An Integrated Approach*, pp. 281-296, NATO ASI Series, Series G: Ecological Sciences, Vol. 24 (Berlin, Springer-Verlag).

ORZACK, S. H. & SOBER, E. (1994) How (not) to test an optimality model, *Trends in Ecology and Evolution*, 9, pp. 265-267.

PARKER, G. A. & MAYNARD SMITH, J. (1990) Optimality theory in evolutionary biology, *Nature*, 348, pp. 27-33.

PARTRIDGE, L. K. (1992) Measuring reproductive costs, *Trends in Ecology and Evolution*, 7, pp. 99-100.

PARTRIDGE, L. K. & BARTON, N. H. (1993) Optimality, mutation, and the evolution of aging, *Nature*, 362, pp. 305-311.

PLEDGER, S. & SCHWARZ, C. J. (2001) Modelling heterogeneity of survival in band-recovery data using mixtures, *Journal of Applied Statistics*, this issue.

POLLOCK, K. H. (1981) Capture-recapture models allowing for age-dependent survival and capture rates, *Biometrics*, 37, pp. 521-529.

PRADEL, R. (1993) Flexibility in survival analysis from recapture data: handling trap-dependence. *In*: J.-D. LEBRETON & P. H. NORTH (Eds), *Marked Individuals in the Study of Bird Populations*, pp. 29-37 (Basel, Birkhäuser-Verlag).

PRADEL, R. (1996) Utilization of capture-mark-recapture for the study of recruitment and population growth rate, *Biometrics*, 52, pp. 703-709.

PRADEL, R. & LEBRETON, J.-D. (1999) Comparison of different approaches to the study of local recruitment of breeders, *Bird Study*, 46, pp. 74-81.

PRADEL, R., HINES, J. E., LEBRETON, J.-D. & NICHOLS, J. D. (1997) Capture-recapture survival models taking account of transients, *Biometrics*, 53, pp. 60-72.

PRÉVOT-JULLIARD, A.-C., LEBRETON, J.-D. & PRADEL, R. (1998) Re-evaluation of the adult survival of black-headed gull (*Larus ridibundus*) in the presence of recapture heterogeneity, *Auk*, 115, pp. 85-95.

PRÉVOT-JULLIARD, A-C., HENTTONEN, H., YOCCOZ, N. G. & STENSETH, N. C. (1999) Delayed maturation in female bank voles: optimal decision or social constraint? *Journal of Animal Ecology*, 68, pp. 684-697.

PRICE, T. D. & GRANT, P. R. (1985) The evolution of ontogeny in Darwin's finches: a quantitative genetic approach, *American Naturalist*, 125, pp. 169-188.

PRICE, T. & LANGEN, T. (1992) Evolution of correlated characters, *Trends in Ecology and Evolution*, 7, pp. 307-310.

REEVE, H. K. & KELLER, L. (1999) Levels of selection: burying the units-of-selection debate and unearthing the crucial new issues. *In*: L. KELLER (Ed.), *Levels of Selection in Evolution*, pp. 3-14 (Princeton, Princeton University Press).

REZNICK, D. (1985) Cost of reproduction: an evaluation of the empirical evidence, *Oikos*, 44, pp. 257-267.

REZNICK, D. (1992) Measuring the costs of reproduction, *Trends in Ecology and Evolution*, 7, pp. 42-46.

ROFF, D. A. (1992) *Evolution of Life Histories—Theory and Analysis* (New York, Chapman & Hall).

ROFF, D. A. (1994) The estimation of genetic correlations from phenotypic correlations: a test of Cheverud's conjecture, *Heredity*, 74, pp. 481-490.

ROFF, D. A. (1997) *Evolutionary Quantitative Genetics* (New York, Chapman & Hall).

RUXTON, G. D. & ROHANI, P. (1998) Fitness-dependent dispersal in metapopulations and its consequences for persistence and synchrony, *Journal of Applied Ecology*, 67, pp. 530-539.

SCHWARZ, C. J. & ARNASON, A. N. (2000) The estimation of age-specific breeding probabilities from capture-recapture studies, *Biometrics*, 56, pp. 299-304.

SCHWARZ, C. J. & SEBER, G. A. F. (1999) Estimating animal abundance: review III—*Statistical Science*, 14, pp. 427-456.

SEGER, J. & STUBBLEFIELD J. W. (1996) Optimisation and adaptation. *In*: M. R. ROSE & G. V. LAUDER (Eds), *Adaptation* (San Diego, Academic Press).

SERVICE, P. M. (2000) Heterogeneity individual mortality risks and its importance for evolutionary studies of senescence, *American Naturalist*, 156, pp. 1-13.

SIBLY, R. & CALOW, P. (1983) An integrated approach to life-cycle evolution using selective landscapes, *Journal of Theoretical Biology*, 102, pp. 527-547.

SIBLY, R. & CALOW, P. (1984) Direct and absorption costing in the evolution of life cycles, *Journal of Theoretical Biology*, 111, pp. 463-473.

SIH, A. & GLEESON, S. K. (1995) A limits-oriented approach to evolutionary ecology, *Trends in Ecology and Evolution*, 10, pp. 378-382.

SPENDELOW *et al.* (2002) Modelling postfledging survival and age-specific breeding probabilities in species with delayed maturity: a case study of Roseate Terns at Falkner Island, Connecticut, *Journal of Applied Statistics*, this issue, pp. 385-406.

STEARNS, S. C. (1992) *The Evolution of Life Histories* (Oxford, Oxford University Press).

STEARNS, S. C., DE JONG, G. & NEWMAN, R. (1991) The effect of phenotypic plasticity on genetic correlations, *Trends in Ecology and Evolution*, 6, pp. 122-126.

TANEYHILL, D. E. (1993) The logistic equation revisited: final instalment, *Trends in Ecology and Evolution*, 8, p. 70.

THOMPSON, J. D. (1991) Phenotypic plasticity as a component of evolutionary change, *Trends in Ecology and Evolution*, 6(8), pp. 246-249.

TUOMI, J., HAKALA, T. & HAUKIOJA, E. (1983) Alternative concept of reproductive effort, costs of reproduction and selection in life-history evolution, *American Zoologist*, 23, 25-34.

TURCHIN, P. (1990) Rarity of density-dependence or population regulation with lags?, *Nature*, 344, pp. 660-663.

TURCHIN, P. (1993) The logistic equation revisited: final instalment, *Trends in Ecology and Evolution*, 8, pp. 68-69.

TURELLI, M. (1988) Phenotypic evolution, constant covariances, and the maintenance of additive variance, *Evolution*, 42, pp. 1342-1347.

TRUSSELL, J. & RICHARD, T. (1985) Correcting for unmeasured heterogeneity in hazard models using the Heckman-Singer procedure. *In*: N. BRANDON TUMA (Ed.) *Sociological Methodology*, pp. 242-277 (San Francisco, Jossey-Bass).

VAN NOORDWIJK, A. J. & DE JONG, G. (1986) Acquisition and allocation of resources: their influence on variation in life history tactics, *American Naturalist*, 128, 137-142.

VAUPEL, J. W. & YASHIN, A. I. (1985a) Heterogeneity's ruses: some surprising effects of selection on population dynamics, *The American Statistician*, 39, 176-185.

VAUPEL, J. W. & YASHIN, A. I. (1985b) The deviant dynamics of death in heterogeneous populations. *In*: N. BRANDON TUMA (Ed.) *Sociological Methodology*, pp. 179-212 (San Francisco, Jossey-Bass).

VAUPEL, J. W. & YASHIN, A. I. (1985c) Mortality and aging in a heterogeneous population: a stochastic process model with observed and unobserved variables, *Theoretical Population Biology*, 27, pp. 154-175.

VIA, S., GOMULKIEWICZ, R., DE JONG, G., SCHEINER, S. M., SCHLICHTING, C. D. & VAN TIENDEREN, P. H. (1995) Adaptive phenotypic plasticity: consensus and controversy, *Trends in Ecology and Evolution*, 10, pp. 212-217.

VIALLEFONT, A., COOCH, E. G. & COOKE, F. (1995) Estimation of trade-offs with capture-recapture models: a case study on the lesser snow goose, *Journal of Applied Statistics*, 22, pp. 847-861.

WEBER, T. P. (1998) News from the realm of the ideal free distribution, *Trends in Ecology and Evolution*, 13(3), pp. 89-90.

WHITE, G. C. & BURNHAM, K. P. (1999) Program MARK: survival estimation from populations of marked individuals, *Bird Study*, 46, pp. 120-139.

WILEY, R. H. (1974) Effects of delayed reproduction on survival, fecundity, and the rate of population increase, *American Naturalist*, 108, pp. 705-709.

WILLIAMS, B. K. & NICHOLS, J. D. (1984) Optimal timing in biological processes, *The American Naturalist*, 123, pp. 1-19.

WILLIAMS, G. C. (1966) Natural selection, the cost of reproduction, and a refinement of Lack's principle, *American Naturalist*, 100, pp. 687-690.

WILLIAMS, G. C. (1992) *Natural Selection: Domains, Levels, and Challenges* (New York, Oxford University Press).

WILSON, W. G. (1998) Resolving discrepancies between deterministic population models and individual-based simulations, *The American Naturalist*, 151, pp. 116-134.

WOOLLER, R. D., BRADLEY, J. S., SKIRA, I. J. & SERVENTY, D. L. (1990) Reproductive success of short-tailed shearwater *Puffinus tenuirostris* in relation to their age and breeding experience, *Journal of Animal Ecology*, 59, pp. 161-170.