

# Dynamic models in behavioural and evolutionary ecology

Alasdair Houston\*, Colin Clark†, John McNamara‡ & Marc Mangel§

\* Department of Zoology, University of Oxford, Oxford OX1 3PS, UK

† Institute of Applied Mathematics, University of British Columbia, Vancouver V6T 1W5, Canada

‡ School of Mathematics, University of Bristol, Bristol BS8 1TW, UK

§ Departments of Agricultural Economics, Entomology and Mathematics, University of California, Davis, California 95616, USA

*The maximization of fitness is often used to analyse the action of natural selection on the life history of animals but short periods of behaviour receive ad hoc treatment. This article describes a dynamic, stochastic model for analysing behaviour in terms of the maximization of fitness.*

THE Darwinian concept of evolution by natural selection<sup>1</sup> has led to the introduction of optimization methods in biology<sup>2-4</sup>. The basic idea, as illustrated by life history theory<sup>5-7</sup>, is that behaviours have evolved in a way that maximizes some measure of reproduction (usually called 'fitness'). In general, life history theory deals with relatively gross decisions about allocation, such as yearly reproductive effort, and is not well suited for the description of fine-scale decisions, such as when a bird should remove an empty egg shell that might attract predators to the nest<sup>8</sup>. The classic work of ethologists such as Tinbergen and Lorenz showed that such fine-scale behavioural traits are also subject to the principles of natural selection. Here we describe mathematical methods that can be used to treat behaviour from an evolutionary standpoint by describing a framework for dynamic modelling of behavioural decisions in the context of the life history of an animal<sup>9,10</sup>.

Earlier attempts to apply the principle of natural selection to fine-scale decisions typically maximized a simple surrogate currency instead of fitness<sup>11</sup>. For example, most models that come under the heading of optimal foraging theory use the rate of energy gain as a currency, and assume that maximizing this rate is equivalent to maximizing fitness. Although these models sometimes can predict short term behaviours<sup>12</sup>, they have limitations such as not being able to compare predation and starvation risks, not considering sequences of actions, and ignoring the state of the animal and the information it has about the environment. To include these aspects, a dynamic, state-variable approach to the analysis of behavioural decisions is required. For example, consider a fish in a lake which contains a number of different habitats<sup>13</sup>. If the potential food intake and risk of predation vary between habitats, how would it judge which was the better foraging habitat? Ideally, an optimization model should be able to explain both short-term (daily) movements between habitats and long-term habitat shifts over the lifetime of the fish. Thus we are concerned with both entire life histories and shorter periods in the life of an animal that involve detailed choices.

The four components of the framework we propose are a set of variables characterizing the state of the animal, a set of actions that the animal can perform, dynamics which specify the relationship between actions and subsequent states, and a state-dependent reward function that specifies future reproductive success in terms of the state of the animal at the end of a relatively short interval. To illustrate the framework, we begin with a simple paradigm for habitat selection<sup>9</sup>, then provide

examples concerning the behaviour of a number of different organisms.

## A simple model

Imagine an animal that can choose one of  $n$  'habitats' ( $H_i$ ) in which to forage on a particular day  $t$ . Each habitat  $H_i$  is characterized by three parameters:  $\beta_i$ , probability of death due to predation,  $Y_i$ , energetic value of food items,  $\lambda_i$ , probability of discovering a food item.

The most interesting situation arises when the riskier habitats also happen to be the most productive. The question then is, assuming that evolution has led to the selection of 'optimal' foraging behaviour, which habitat would we expect the animal to choose?

To answer this question we introduce a state variable,  $X_t$ , representing the animal's bodily energy reserves at the start of day  $t$ . Thus energy reserves are increased by daily food consumption  $Z_t$ , and decreased by daily metabolic expenditure, denoted by  $\alpha$ , so that reserves at the start of day  $t+1$  are given by the equation:

$$X_{t+1} = X_t - \alpha + Z_t \quad (1)$$

$Z_t$  is a random variable whose distribution is determined by the animal's choice of habitat. If habitat  $H_i$  is chosen, then  $Z_t$  equals  $Y_i$  with probability  $\lambda_i$ , and equals zero with probability  $1 - \lambda_i$ .

Energy reserves cannot be negative, and are bounded by the animal's capacity Cap:

$$0 \leq X_t \leq \text{Cap} \quad (2)$$

If energy reserves decline to a low level, the forager is likely to die of starvation; we simplify this situation by assuming that the animal dies if  $X_t = 0$ .

This framework now allows us to compare, in terms of fitness, the benefits of food consumption with the risks of predation. Death may occur either from starvation or predation. If, for the moment, we limit attention to behaviour during a nonbreeding phase in the animal's life, it may be reasonable to identify fitness with the probability of surviving until the end of the nonbreeding period. We then adopt the hypothesis that an optimal behavioural strategy is one that maximizes the forager's probability of surviving from  $t=0$  to  $t=T$ , where  $T$  denotes a fixed horizon, corresponding to the onset of breeding activity. (This simple paradigm can be extended in many directions, one of which includes breeding activities.)

We assume that the animal 'knows' the parameter values  $\beta_i$ ,  $Y_i$ ,  $\lambda_i$ , characterizing each habitat  $H_i$ , and that it can assess its

¶ To whom correspondence should be addressed.

### 1. Derivation of the stochastic dynamic programming equation

Recall that  $J(x, t, T)$  is the maximal probability of survival from day  $t$  to day  $T$  given that  $X_t = x$ . Hence, when  $t = T$ ,  $J(x, T, T) = 1$  as long as  $x > 0$ . For values of  $t < T$ , we proceed as follows. To survive from day  $t$  to  $T$ , the organism must first survive day  $t$  and then survive from day  $t+1$  to  $T$ , taking into account changes in the state variable that occur on day  $t$ .

Suppose that the organism chooses habitat  $H_i$  on day  $t$ . If the animal avoids predation on this day, its energy reserves at the beginning of day  $t+1$  can have one of two values,  $x - \alpha$  and  $x'_i$ , where  $x'_i$  is the minimum of  $x - \alpha + Y_i$  and  $\text{Cap}$ . These values correspond to finding no food and finding a food item of energetic value  $Y_i$  and occur with probabilities  $1 - \lambda_i$  and  $\lambda_i$ , respectively. The corresponding maximal probabilities that the animal survives from day  $t+1$  to  $T$  are  $J(x - \alpha, t+1, T)$  and  $J(x'_i, t+1, T)$  respectively. Thus, conditional on the animal avoiding predation on day  $t$ , its probability of surviving until  $T$  is

$$(1 - \lambda_i)J(x - \alpha, t+1, T) + \lambda_iJ(x'_i, t+1, T).$$

Taking into account predation risk on day  $t$  the survival probability is

$$\begin{aligned} P_i(x, t, T) = & (1 - \beta_i)[(1 - \lambda_i)J(x - \alpha, t+1, T) \\ & + \lambda_iJ(x'_i, t+1, T)]. \end{aligned}$$

Because  $J(x, t, T)$  is the maximum probability of surviving from day  $t$  to day  $T$ , we simply maximize  $P_i$  over habitats to find  $J(x, t, T)$ , which gives equation (4) of the text.

energetic state  $X_t$  on each day  $t$ . The computation of the optimal strategy for selection of habitat can be carried out using the method of stochastic dynamic programming. Details are given in Ross<sup>14</sup>, see also Box 1. Let  $J(x, t, T)$  denote the probability of survival from day  $t$  until day  $T$ , given that  $X_t = x$ , and assuming that the animal uses the best strategy. We then have

$$J(x, T, T) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \end{cases} \quad (3)$$

(the animal is alive on day  $T$  if and only if  $x > 0$ ).

It is shown in Box 1 that in general  $J(x, t, T)$  satisfies the equation

$$J(x, t, T) = \max P_i(x, t, T) \quad x > 0 \quad (4)$$

where

$$\begin{aligned} P_i(x, t, T) = & (1 - \beta_i)[\lambda_i J(x'_i, t+1, T) \\ & + (1 - \lambda_i)J(x - \alpha, t+1, T)], \end{aligned} \quad (5)$$

and where  $x'_i = \min(x - \alpha + Y_i, \text{Cap})$  (6)

Equation (4) is known as the stochastic dynamic programming equation.

If the values of the model parameters are known, then  $J$  can be found by an iterative process called backward induction. We first take  $t = T - 1$ . Because of equation (3), all expressions on the right side of equation (5) are known. Thus  $P_i(x, T-1, T)$  is known for each choice of habitat  $H_i$  and  $J(x, T-1, T)$  can be found from equation (4). The value  $i^*$  of  $i$  that yields the maximum in equation (4) specifies the optimal habitat  $H_{i^*}$ . In general  $i^*$  depends on  $x$ . Having computed  $J(x, T-1, T)$  as above, one can then use the same process to compute  $J(x, T-2, T)$ , and subsequently  $J(x, T-3, T), \dots, J(x, 0, T)$ . The computations are easily coded for automatic computation, and can be performed very rapidly on a desktop microcomputer. The procedure is illustrated in Box 2.

The qualitative predictions are that (i) foragers will accept greater risks of predation in habitats with higher food availability, and (ii) a forager will accept greater predation risks when hungry than when it is well-fed. This provides an explanation for the decrease in feeding rate as total intake increases<sup>15</sup>. Such

### 2. Illustration of backward induction

We take the upper limit on reserves,  $\text{Cap}$ , to be 20 and consider three habitats, with  $\beta_1 = 0$ ,  $\beta_2 = 0.01$ ,  $\beta_3 = 0.05$ ,  $\lambda_1 = 0.4$ ,  $\lambda_2 = 0.6$ ,  $\lambda_3 = 0.8$  and  $Y_1 = Y_2 = Y_3 = 2$ ,  $\alpha = 1$ . Thus as the predation risk in a habitat increases, so does the mean net gain. The table gives  $P_i(x, t, T)$  for a range of  $x$  and  $t$ . (Top entry,  $P_1$ ; middle entry,  $P_2$ ; bottom entry,  $P_3$ ). To compute these values, start at the right-hand column, where  $J(x, T, T)$  is given by equation (3). At time  $T-1$ ,  $P_i(x, T-1, T)$  is given by equation (5), and  $J(x, T-1, T)$  is equal to the maximum value of  $P_i$ . This value is underlined. It can be seen that it is optimal to choose habitat 3 only if reserves are one. Calculating  $J(x, T-1, T)$  for all  $x \geq 2$  shows that habitat 1 is chosen for all  $x \geq 2$ . The values of  $J(x, T-1, T)$  can be used to find  $P_i(x, T-2, T)$  and hence  $J(x, T-2, T)$  etc. Using this procedure it can be shown that, for  $T-t$  sufficiently large, the optimal policy has the form 'choose habitat 3 if  $1 \leq x \leq 2$ , choose habitat 2 if  $3 \leq x \leq 11$  and choose habitat 1 if  $12 \leq x \leq 20$ '. This is the long-term policy for maximizing survival in this environment.

In addition to finding the optimal policy, it is possible to calculate the probability that an animal following it will be in any given state. From this, the probability that a given habitat is selected can be found. In this example, the probabilities of selecting habitats 1 to 3, respectively, are 0.526, 0.471 and 0.003. Thus in a large population of animals, the percentage of animals in these habitats would be 52.6, 47.1 and 0.3.

x	$J(x, t, T)$ is underlined				$J(x, T, T)$	
	$t = T-4$	$t = T-3$	$t = T-2$	$t = T-1$		
4	<u><math>P_1(x, t, T)</math></u>	0.9713	1.0000	1.0000	1.0000	1
	$P_2(x, t, T)$	0.9711	0.9900	0.9900	0.9900	
	$P_3(x, t, T)$	0.9409	0.9500	0.9500	0.9500	
3	<u><math>P_1(x, t, T)</math></u>	0.9426	0.9426	1.0000	1.0000	1
	$P_2(x, t, T)$	0.9521	0.9521	0.9900	0.9900	
	$P_3(x, t, T)$	0.9318	0.9318	0.9500	0.9500	
2	<u><math>P_1(x, t, T)</math></u>	0.7933	0.8560	0.8560	1.0000	
	$P_2(x, t, T)$	0.8378	0.8950	0.8950	0.9900	
	$P_3(x, t, T)$	0.8542	0.9440	0.9440	0.9500	
1	<u><math>P_1(x, t, T)</math></u>	0.3618	0.3618	0.4000	0.4000	1
	$P_2(x, t, T)$	0.5372	0.5372	0.5940	0.5940	
	$P_3(x, t, T)$	0.6873	0.6873	0.7600	0.7600	
0	$J(0, t, T)$	0	0	0	0	0

predictions are hardly surprising, but one cannot expect to derive deep insights from such a simple model. The intuitively appealing prediction that foraging behaviour should be affected by the forager's current state of hunger is one that does not arise from classical foraging models. Box 2 also shows that quantitative predictions can be made about the long-term distribution of animals among habitats.

The dynamic modelling approach offers several advantages over previous methods. First, the use of dynamic state variables means that models can be made more realistic and biologically meaningful. Constraints on variable values and rates are easily included—indeed they can hardly be omitted. The models can be made time-dependent and fully stochastic, realistically reflecting environmental conditions. Alternative behavioural choices (for example, foraging, resting or reproductive activities) can be treated simultaneously in a unified way<sup>9,10,16</sup>. Evolutionary fitness can be modelled in a direct manner, by including reproduction as well as survival. In fact, this approach is a generalization of classical life history theory<sup>5-7,16</sup>, but, by working with a representation of the animal's state, a better insight into the nature of life-history tradeoffs is obtained<sup>16</sup>. Moreover, the approach enables us to calculate the loss in expected future reproductive success that results from adopting a suboptimal action<sup>10</sup>, referred to as the canonical cost of an action<sup>10</sup>. These costs can be used to assess the robustness of the conclusions of

a model and the selection pressure for a given behavioural strategy<sup>10,17</sup>.

In short, the stochastic programming approach to behaviour modelling is extremely flexible, and can be used to help organize and understand a wide variety of both field and experimental data; some recent applications are described below.

### Foraging in lions

The African lion is the only social member of the cat family<sup>18</sup>. Lions live in prides typically consisting of up to 18 related adult females plus offspring and unrelated adult males. Female lions from a given pride hunt prey in cooperative groups ranging from one to eight, depending on prey size and other circumstances.

Caraco and Wolf<sup>19</sup> analysed data collected in the Serengeti by Schaller<sup>20</sup> and calculated average daily food intake per lion as a function of hunting group size. In all cases, groups of size  $n = 2$  maximized the average individual food intake. For large prey such as zebra and wildebeest, however, the average number of lions observed feeding at kills ranged from four to eight. Caraco and Wolf calculate that these large groups would experience up to 50% reduction in food intake, compared with groups of two. When hunting small prey such as Thomson's gazelle, on the other hand, lions either hunt individually or in small groups of two or three.

Packer<sup>18</sup> has re-examined the results of Caraco and Wolf and concludes that "there are no good data showing that cooperative hunting is in fact beneficial to individual lions" in terms of average rate of food intake. Instead, Packer suggests that lion sociability is primarily an adaptation to the opportunity for scavenging from conspecifics in a region where both lions and their prey occur at high densities. When food items are large and infrequently obtained, variance in the amount of food acquired may have as great an influence on fitness as does the average feeding rate. By increasing the frequency of kills and decreasing the food per hunter per kill, group foraging generally decreases the variance in food intake.

To assess the implications of group foraging in greater detail, Clark<sup>21</sup> developed a dynamic model of lion hunting behaviour. In simplified form, the model is equations (1) and (2) with  $Z_t$  a random variable given by

$$\text{probability}(Z_t = E/n) = \lambda_n, \quad \text{probability}(Z_t = 0) = 1 - \lambda_n$$

where  $X_t$  denotes stomach contents at the beginning of day  $t$ ,  $Z_t$  is food intake,  $n$  denotes hunting group size,  $\alpha$  is the daily minimum food requirement per lion ( $= 6$  kg), Cap is the stomach capacity ( $= 30$  kg) and  $E$  is the average food content per prey item ( $= 164$  kg for a zebra and 12 kg for a gazelle; zebra carcasses last up to three days; gazelle are consumed immediately). The kill probabilities  $\lambda_n$  depend on group size  $n$ , and also on the type of prey and habitat. We wish to find the behaviour that maximizes survival over a time period of  $T$  days. Thus equation (3) holds, and the resulting dynamic programming equation is similar to that given by equations (4)–(6) above but with  $\beta_i = 0$ , because lions are not preyed upon.

Typical predictions of this model are shown in Table 1a, which shows survival probabilities  $P$  and optimal hunting groups sizes  $n^*$  as functions of initial stomach contents  $x$  ( $T = 30$  days), for three sets of prey/habitat combinations.

Notice that the optimal group size  $n^*$  predicted by the dynamic model depends on the lion's current state  $x$ , as well as on prey type and habitat. For example, well-fed lions should hunt zebra in groups of size six, rather than two as predicted by Caraco and Wolf's model. In fact, hunting in groups of size two is severely suboptimal, yielding only a 78% probability of survival over a 30-day period. The main reasons for this result are (i) small groups hunting zebra have reduced kill probabilities ( $\lambda_2 = 0.30$ ,  $\lambda_6 = 0.43$ ), and (ii) two lions cannot consume a whole zebra before the carcass rots so that much meat is wasted if only two lions share a zebra kill. The importance of this constraint

**Table 1** Probability of survival and hunting group size

$x$ (kg)	<i>a</i>						<i>b</i>		
	(1)		(2)		(3)		$N$	$P$	$n^*$
	$P$	$n^*$	$P^a$	$n^*$	$P$	$n^*$			
5	0.86	3	0.26	2	0.06	1	2	0.18	2
10	0.90	4	0.43	2	0.13	1	4	0.75	2
20	0.97	4	0.64	2	0.26	2	6	0.95	2
30	0.99	6	0.74	2	0.35	2	8	0.993	2
							10	0.999	2

*a*, Maximum probability of survival ( $P$ ) and optimal hunting group size ( $n^*$ ), as functions of current stomach contents ( $x$ ) with a 30-day horizon, for lions hunting (1) zebra in wet-season Serengeti habitat, (2) Thomson's gazelle in wet-season habitat, and (3) gazelle in dry-season habitat. From Clark<sup>21</sup>. *b*, Maximum probability of survival ( $P$ ) and optimal hunting group size ( $n^*$ ) for prides that communally share kills, as a function of pride size (number of adult females),  $N$ . Prey is zebra, dry-season habitat.

is made obvious by the dynamic approach, but was overlooked in previous analyses of lion hunting behaviour.

A second prediction of the dynamic model is that hungry lions should hunt in smaller groups than well-fed lions. In times of prey scarcity, for example, lions may be hungry most of the time, and our prediction is that smaller hunting groups will then be observed. The breakup of lion prides during the dry season in the Kalahari has been described by Owens and Owens<sup>22</sup>.

This model can be modified to allow for scavenging of kills by pride mates, which Packer suggests as the main mechanism underlying lion social behaviour. Consider a pride containing  $N$  adult females, and suppose that the pride hunts in  $k$  separate groups each containing  $n$  females so that  $N = kn$ . Kills are assumed to be shared equally among pride members. These assumptions lead to a binomial distribution for  $Z_t$ .

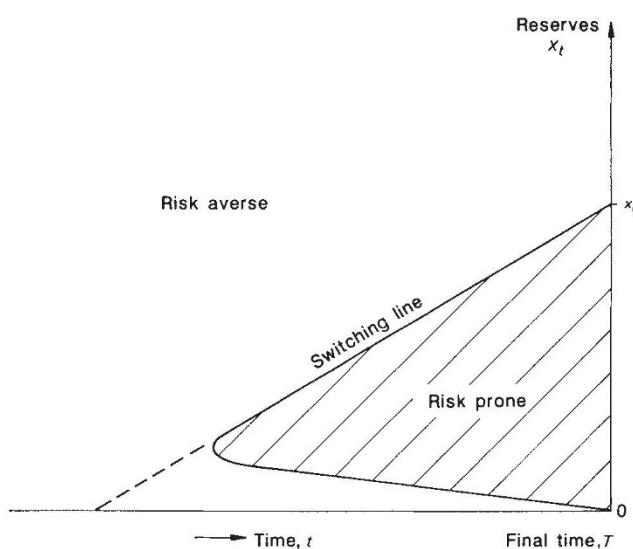
Table 1b shows the optimal hunting group sizes  $n^*$  and 30-day survival probabilities  $P$  for this model of communal scavenging, as a function of pride size  $N$ . The corresponding results for a pride of size  $N = 6$  that does not share kills are  $n^* = 6$  and  $P = 0.42$ , indicating that communal scavenging of kills may greatly increase survival rates for individual lions under certain circumstances—namely where prey is large, relatively scarce, and perishable, and where ecological conditions permit separate hunting groups to communicate to pride mates the fact that a kill has been made. These circumstances are characteristic of the African savannah.

### Small birds in winter

During the winter small birds must forage most of the daylight period if they are to obtain enough food to satisfy their daily requirement<sup>23</sup>. Under these conditions adverse fluctuations in the amount of food obtained can easily result in starvation. Foraging activities also expose a bird to predators, so that there is often a trade off between the need to obtain food and the need to avoid predators.

To model the foraging decisions of a bird over a single day, let time  $t = 0$  be dawn and let time  $t = T$  be dusk. At each of the times  $t = 0, 1, \dots, T-1$  the bird must choose a foraging option from some set of available options. The option chosen at time  $t$  determines the distribution of the amount of food obtained between time  $t$  and  $t+1$  and the predation risk. The option also determines metabolic consumption between  $t$  and  $t+1$ . Overnight the bird uses  $x_c$  units of energy reserves.

To model foraging behaviour in winter, we assume that each day in winter can be described in this way. Given a terminal reward (expected future reproductive success) for the various energy states at the end of winter, the optimal policy on each day can be found by working backwards from the terminal reward. When there are sufficiently many days left in winter,



**Fig. 1** The region of risk-prone behaviour as a function of energy reserves and time of day. An animal has two possible courses of action,  $a_H$  and  $a_L$ , of high and low variance, respectively, in the energetic gain resulting from the course of action. The dashed line has slope  $\mu$  equal to the mean rate of gain under both foraging options. In the shaded region of the figure it is optimal to choose the 'risky' option  $a_H$ .

the optimal policy on all days is the same. This policy is the one that maximizes the bird's long-term survival probability.

This sort of model can be used to analyse the value of a given increase in energy reserves, the trade off between starvation and predation risks, or the importance of variability in foraging returns. We discuss some results concerning the last of these topics.

The calculations that yield the optimal policy also give the long-term survival probability as a function of energy reserves. In particular, the survival probability of a bird at dusk is a non-linear function of its energy reserves. This function is convex for low levels of reserves and concave for high levels of reserves. One can regard this function as a terminal reward function at dusk. The consequences of non-linear reward functions are well-known (see ref. 12 for review). Given a choice between two options that result in the same mean gain but have different variances, it is optimal to maximize variability (risk-prone behaviour) when the derivative is increasing and to minimize variability (risk-averse behaviour) when the derivative is decreasing. In the small-bird-in-winter paradigm the situation is complicated by the need to avoid daytime starvation, and can be summarized as follows.

Suppose that two actions  $a_H$  and  $a_L$  are available at each of the times of day  $t = 0, 1, \dots, T-1$ . These actions have no predation risk and result in the same mean energetic gain  $\mu$  per time unit, but the variance in gain is greater under  $a_H$  than under  $a_L$ . We refer to  $a_H$  as the high-variance option and  $a_L$  as the low-variance option. The general form of the optimal policy is shown in Fig. 1. The switching line in this figure has slope  $\mu$ . An animal whose reserves are below this line will not survive the night if it obtains energy at the mean rate. It is thus optimal to take risks in the shaded region of the graph by choosing the high variance option  $a_H$ . When reserves are very low, it is always optimal to be risk-averse and choose the low variance option, because the high variance option has a greater risk of immediate starvation. An animal with reserves above the switching line expects to survive the night and therefore plays safe by being risk-averse. The result of these effects is to give a wedge-shaped region in which it is optimal to be risk-prone.

Given a choice between several actions we define the canonical cost of choosing a particular action as the loss in expected reproductive success which is incurred as a result of this choice<sup>10</sup>. In the small-bird-in-winter model the canonical cost of action  $a$ ,  $c(a, x, t)$ , depends on reserves and time of day, and is essentially the reduction in overwinter survival probability which results from choosing  $a$ .

In this example the canonical cost of choosing  $a_L$  when  $a_H$  is optimal and of choosing  $a_H$  when  $a_L$  is optimal can be quite high. But the probability of being in the shaded region in which  $a_H$  is optimal is small. Houston and McNamara<sup>17</sup> use this to argue that the selection pressure to be risk-prone under the appropriate circumstances is much smaller than the corresponding selection pressure to be risk-averse. This conclusion is supported by analysing the total mortality under various strategies. If a bird with a single behavioural option is given a further option yielding the same mean amount of food but having a higher variance, the bird can slightly reduce its probability of starvation by choosing the high variance option in appropriate circumstances. If instead the second option has a lower variance it is optimal for the bird to use this option most of the time, and the resultant drop in starvation probability is usually massive. In line with this, the data suggest that risk-prone behaviour is a much less robust phenomenon than risk-averse behaviour (see refs 17, 24 for further discussion and references).

### The dawn chorus

In spring, many songbirds produce an intense burst of song at dawn, the dawn chorus. McNamara *et al.*<sup>25</sup> use a dynamic model to analyse the factors that can produce this chorus. The model considers a male bird that sings to attract a mate. A period of five days is modelled. The daylight period of each day is divided into 96 equal time intervals. During each interval the bird can either forage or sing. If it forages it gains no mate but finds food of energy value  $Z_t$ . The energy expended during the interval is a linearly increasing function of the bird's energy reserves (and hence body mass). If reserves at time  $t$  are  $x$  then reserves at time  $t+1$  satisfy.

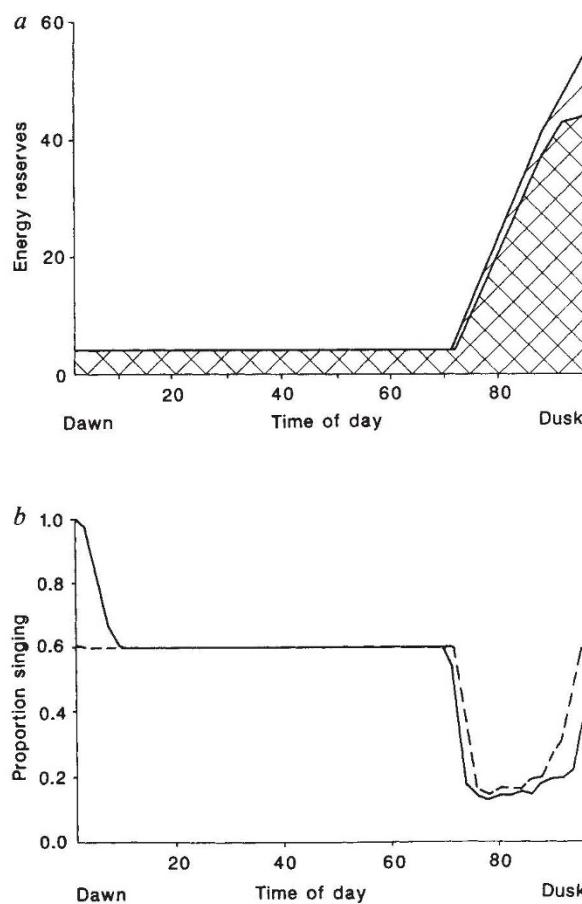
$$X_{t+1} = x + Z_t - (a_F + k_F x)$$

If the bird sings it finds no food and its reserves at time  $t+1$  are given by

$$X_{t+1} = x - (a_S + k_S x)$$

A singing bird attracts a mate with probability  $m_t$ . A bird which manages to attract a mate can spend the remaining time foraging. Birds rest between dusk and the following dawn. Overnight energy consumption is assumed to follow a normal distribution with mean  $\mu_N$  and variance  $\sigma_N^2$ .

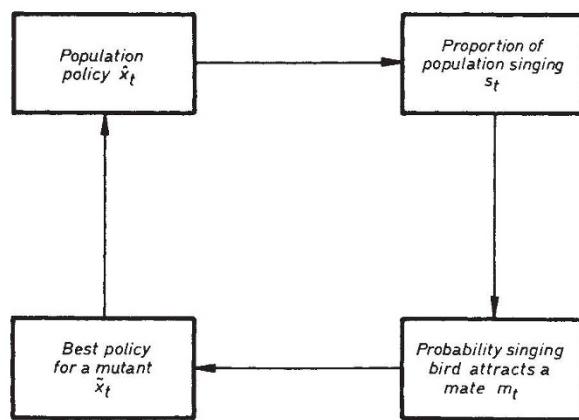
During the five-day period a bird starves if its reserves fall to zero. If the bird survives till dawn on the sixth day it receives a reward of one unit. If it has also obtained a mate by this time it receives a further reward. The optimal policy maximizes



**Fig. 2** *a*, The optimal policy in the model of singing and foraging. The bird can either sing or forage. When there is a constant loss of energy each night ( $\mu_N > 0$ ,  $\sigma_N^2 = 0$ ) it is optimal to forage if and only if reserves are in the double hatched region. When  $\sigma_N^2 = 9$  it is optimal to forage if and only if reserves are in the single or double hatched region. *b*, The proportion of birds singing under the two policies shown in *a*. When  $\sigma_N^2 = 0$ , the proportion is given by the broken line and there is no dawn chorus. When  $\sigma_N^2 = 9$ , the proportion is given by the solid line. There is a marked dawn chorus and a slight reduction in singing at dusk.

expected total reward. In general the optimal action depends on reserves and time of day, but is a very similar function of these on all days except the last.

In describing the results of the model, it is important to distinguish between the optimal policy and the expected behaviour that is generated by following the optimal policy. The optimal policy is a specification of a critical level of reserves  $\hat{x}_t$ , such that if  $X_t > \hat{x}_t$  then it is optimal to sing and if  $X_t < \hat{x}_t$  then it is optimal to forage. The expected behaviour depends on the proportion of time spent above or below  $\hat{x}_t$  and gives the proportion of birds singing in a large population of identical birds. In the absence of an energy loss overnight,  $\hat{x}_t$  is constant and there is no circadian rhythm. Adding a constant loss of energy per night (that is,  $\mu_N > 0$ ,  $\sigma_N^2 = 0$ ) results in a rise in  $\hat{x}_t$  in the afternoon as a bird must build up its reserves to survive the following night (Fig. 2*a*). This results in a dip in the proportion of birds that are singing in the afternoon (Fig. 2*b*). As might be expected, a burst of song at dawn can be produced by increasing the probability of getting a mate  $m_t$  at this time or decreasing the returns from foraging  $Z_t$ . McNamara *et al.*<sup>25</sup> found, however, that making the energy lost overnight a random variable ( $\sigma_N^2 > 0$ ) could also produce a dawn chorus. The reason is that a bird must get its reserves at dusk above the amount



**Fig. 3** The iterative procedure for finding a dynamic ESS (after Houston and McNamara<sup>28</sup>). The procedure starts with a population following the policy 'sing if reserves at time  $t$  are greater than  $\hat{x}_t$ '. From this policy, the proportion of the population singing at time  $t$ ,  $s_t$ , is calculated. Using the interaction between singing males, the probability  $m_t$ , that a singing bird attracts a mate at time  $t$  is found. From  $m_t$ , the best policy  $\hat{x}_t$  for a single mutant can be found. The population policy is then replaced by  $\hat{x}_t$  and the process is repeated.

required to survive an unusually bad night. To do so, it sacrifices some singing at dusk (Fig. 2). On most nights it does not use up all the reserves that it has accumulated, and so starts the day with its reserves above  $\hat{x}_t$ . It sings until its reserves fall to  $\hat{x}_t$ . In this case the bird sings at dawn not because dawn is a good time for singing but because the problems posed by variable energy loss overnight require it to trade some singing at dusk for some singing at dawn. Thus the intensity of the dawn chorus may depend on the fluctuations in overnight temperature. Reid found a relationship between overnight temperature and song in the Ipswich sparrow<sup>26</sup>.

This example can be used to introduce the idea of dynamic evolutionarily stable strategies. An evolutionarily stable strategy (ESS) renders a population safe from invasion by a mutant that adopts a different strategy<sup>27</sup>. Most previous ESS models have been concerned with a single choice of action that does not depend on the animal's state. If the probability of attracting a mate by singing depends on whether or not other birds are singing, then a game-theory analysis is required. A dynamic ESS is a rule for determining a sequence of actions as a function of an animal's state, the behaviour of other birds, and time of day. Houston and McNamara<sup>28</sup> consider a large population limit in which the probability of attracting a mate depends on the proportion of other birds that are singing. The procedure for finding the ESS, as shown in Fig. 3, is based on finding the best response for a mutant when a population follows a given strategy. For an initial population strategy, the best response for a mutant is found, and the population strategy is then replaced by the best mutant strategy. The procedure is repeated until the best response for a mutant is the current population strategy. This strategy is then the ESS for the population. The procedure can be applied to a wide range of games; Houston and McNamara<sup>29</sup> use it to analyse a dynamic version of the hawk-dove game.

### Clutch size in parasitic insects

Parasitic insects, such as the wasp *Nasonia vitripennis* or various tephritid fruit flies, provide an ideal setting for the study of reproductive behaviour using dynamic models. The adult of these species is free living, laying its eggs in the host. We begin by identifying one or more state variables. Natural choices are energy reserves, mature eggs remaining, or a combination of

mature eggs and oocytes (potential eggs). For simplicity, we will use only the number of mature eggs remaining as a state variable. The 'reward function' for parasitic insects is related to the fecundity of mature offspring from a particular clutch<sup>30</sup>. This function involves a combination of survivorship, size, and egg production of offspring. Also of particular importance for parasitic insects is adult survivorship from one potential host to another (or from one day to another).

We consider the behaviour of an insect that starts its search for hosts with a complement of  $X_0$  mature eggs, and let  $X_t$  denote the number of eggs remaining at the start of period  $t$ . We assume that foraging for hosts stops at some period  $T$ . For example, if eggs are resorbed at the end of a day, then  $T$  would denote the number of search periods in a day. Alternatively, if we are interested in a lifetime problem,  $T$  might be the time at which the insect dies or the time at which its eggs no longer hatch. We then define  $J(x, t, T)$  as the maximum expected fitness from laying eggs between  $t$  and  $T$ , given that  $X_t = x$ .

We are interested in computing  $J(X_0, 1, T)$ . Assume that there are  $n$  different kinds of hosts (indexed by volume, fruit or insect type) and that  $\lambda_i$  is the probability of encountering a host type  $i$  in a period of unit length. If a clutch of size  $C$  is laid in a host of type  $i$ , the fitness accrued to the mother is  $f_i(C)$  and the handling time is  $\tau_i(C)$ . We specify survival through a function  $P_i(x, t)$  representing the probability of surviving from period  $t$  to period  $t+1$ , given that the insect is alive at period  $t$ , that  $X_t = x$ , and that a host of type  $i$  is encountered. The index  $i=0$  is used to indicate that no host is encountered in a given period,  $\lambda_0 = 1 - \sum \lambda_i$  is the probability of no encounter, and  $P_0(x, t)$  is the survival probability. With these assumptions, the fundamental equation is

$$\begin{aligned} J(x, t, T) &= \sum_{i=0}^n \lambda_i \max_{C \leq x} (f_i(C) + P_i(x, t) \\ &\quad \times J(x - C, t + \tau_i(C), T)) \end{aligned} \quad (7)$$

with  $f_0(C) = 0$  for all values of  $C$ , and  $t + \tau_i(C)$  is replaced by  $T$  if it exceeds  $T$ . This equation shows that the interplay of time and state variable is crucial for the analysis of insect oviposition decisions.

Mangel<sup>16</sup> analyses a simple version of equation (7) in which all  $\tau_i(C) = 1$  and  $P_i(x, t) = P_0 = \text{constant}$ . This model can explain observed frequency distributions of clutch sizes<sup>30</sup>, some aspects of superparasitism<sup>31</sup>, and the response of the apple maggot *Rhagoletis pomonella* (Walsh) to its oviposition-marking pheromone<sup>31</sup>. In the first case, small clutches predominate in

the observed frequency distribution of clutches of the parasitic wasp *N. vitripennis*. This is exactly what is predicted by equation (7). The frequency of small clutches should increase with the time left for searching and the probability of encountering a host. The model can also be used to predict when an insect should oviposit in a previously parasitized host. If the remaining search time is great or the insect has few eggs, then a previously parasitized host should be rejected. As  $t$  approaches the time horizon  $T$ , the theory shows that hosts which would be considered unacceptable for small values of  $t$  become acceptable. In all three cases, the interaction of the time horizon and state variable is needed for understanding the oviposition problem. We see this as a general feature of dynamic models.

## Conclusion

The technique for analysing behaviour we present here has the following advantages: (i) It takes account of the state of the animal and how that state changes according to the animal's actions and the environment; (ii) it provides a common currency for assessing behavioural choices in terms of overall fitness, which can be used to analyse trade-offs between different actions; (iii) it includes constraints on state variables or behaviour. Models based on this technique often include a higher degree of biological realism and lead to predictions and insights not provided by simpler models. Stochastic dynamic programming provides a method for finding the optimal behaviour (or behaviours) within the dynamic framework. As a computational technique, stochastic dynamic programming has certain limitations, perhaps the most serious of which is the 'curse of dimensionality' in which computational needs grow vastly as the number of state variables increases. Hence, there is an upper limit to the degree of biological complexity that the method can realistically encompass.

We have applied this dynamic approach to many other situations, such as diel vertical migration of aquatic organisms<sup>32</sup>, optimal choice of prey items<sup>33</sup>, diving behaviour of water birds, growth and migration of salmon, sex change in slugs, web locations of spiders, and food hoarding by small birds.

In our main examples, there are no interactions between animals. We have mentioned that under some circumstances evolutionarily stable dynamic strategies can be modelled in a relatively simple way. In general, the development of such models is likely to encounter conceptual and computational difficulties. Nevertheless, we believe that this is a biologically important issue that deserves further theoretical and empirical research, and that the technique of stochastic dynamic programming will be a powerful tool in this enterprise.

1. Darwin, C. *On the origin of species by natural selection* (Murray, London, 1859).
2. Oster, G. F. & Wilson, E. O. *Caste and Ecology in the Social Insects* (Princeton University Press, 1978).
3. Maynard Smith, J. A. *Rev. Ecol. Syst.* **9**, 31–56 (1978).
4. McFarland, D. J. *Nature* **269**, 15–21 (1977).
5. Stearns, S. C. Q. *Rev. Biol.* **51**, 3–47 (1976).
6. Schaffer, W. M. *Am. Nat.* **121**, 418–431 (1983).
7. Sibly, R. & Calow, P. *J. theor. Biol.* **102**, 527–547 (1983).
8. Tinbergen, N. *et al. Behaviour* **19**, 74–117 (1962).
9. Mangel, M. & Clark, C. W. *Ecology* **67**, 1127–1138 (1986).
10. McNamara, J. M. & Houston, A. I. *Am. Nat.* **127**, 358–378 (1986).
11. Schoener, T. A. *Rev. Syst. Ecol.* **2**, 369–404 (1971).
12. Stephens, D. W. & Krebs, J. R. *Foraging Theory* (Princeton University Press, 1986).
13. Werner, E. E. & Gilliam, J. F. A. *Rev. Ecol. Syst.* **15**, 393–425 (1984).
14. Ross, S. *Introduction to Stochastic Dynamic Programming* (Academic, New York, 1983).
15. McCleery, R. M. *Anim. Behav.* **25**, 1005–1015 (1977).
16. Mangel, M. *J. Math. Biol.* **25**, 1–22 (1987).
17. Houston, A. I. & McNamara, J. M. in *Relevance of models and theories in ethology* (eds Campan, R. & Zayan, R.) 61–75 (Privat, I.E.C., Toulouse).
18. Packer, C. in *Ecological Aspects of Social Evolution* (eds Rubenstein, D. I. & Wrangham, R. W.) 429–451 (Princeton University Press, Princeton, 1986).
19. Caraco, T. & Wolf, L. L. *Am. Nat.* **109**, 343–352 (1975).
20. Schaller, G. B. *The Serengeti Lion* (University of Chicago Press, Chicago, 1972).
21. Clark, C. W. *Anim. Behav.* **35**, 361–368 (1987).
22. Owens, M. J. & Owens, D. D. *Int. Wildl.* **14**, 4–13.
23. Gibb, J. A. *Ibis* **96**, 513–543 (1954).
24. McNamara, J. M. & Houston, A. I. *Acta Biotheoretica* **36**, 3–22 (1987).
25. McNamara, J. M., Mace, R. H. & Houston, A. I. *Behav. Ecol. Sociobiol.* **20**, 399–405 (1987).
26. Reid, M. I. *Anim. Behav.* **35**, 1735–1743 (1987).
27. Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, 1982).
28. Houston, A. I. & McNamara, J. M. *J. theor. Biol.* **129**, 57–68 (1987).
29. Houston, A. I. & McNamara, J. M. *Evol. Ecol.* (in the press).
30. Charnov, E. L. & Skinner, S. W. *Flor. Ent.* **67**, 5–21 (1984).
31. Roitberg, B. D. & Prokopy, R. J. *Ecol. Ent.* **7**, 437–442 (1982).
32. Clark, C. W. & Levy, D. A. *Am. Nat.* (in the press).
33. Houston, A. I. & McNamara, J. M. *Behav. Ecol. Sociobiol.* **17**, 135–141 (1985).