Equivalence relationships between stage-structured population models

Jonathan M. Yearsley [a,*], David Fletcher [b]

[a] *The Macaulay Institute, Craigiebuckler, Aberdeen, AB15 8QH, UK*
[b] *Department of Mathematics and Statistics, University of Otago, Dunedin, New Zealand*

Abstract

Matrix population models are widely applied in conservation ecology to help predict future population trends and guide conservation effort. Researchers must decide upon an appropriate level of model complexity, yet there is little theoretical work to guide such decisions. In this paper we present an analysis of a stage-structured model, and prove that the model's structure can be simplified and parameterised in such a way that the long-term growth rate, the stable-stage distribution and the generation time are all invariant to the simplification. We further show that for certain structures of model the simplified models require less effort in data collection. We also discuss features of the models which are not invariant to the simplification and the implications of our results for the selection of an appropriate model. We illustrate the ideas using a population model for short-tailed shearwaters (*Puffinus tenuirostris*). In this example, model simplification can increase parameter elasticity, indicating that an intermediate level of complexity is likely to be preferred.
© 2002 Elsevier Science Inc. All rights reserved.

1. Introduction

Stage-structured population models are a popular method in conservation biology for analysing the viability of a population. Once the form of a model has been chosen, a wide range of analytical tools are available to investigate the dynamics of a population [1]. For example, we can

---

[*] Corresponding author: Tel.: +44-1224 498 200; fax: +44-1224 318 556.
E-mail address: j.yearsley@macaulay.ac.uk (J.M. Yearsley).

estimate the long-term population growth rate (λ), stable-stage distribution [1], the sensitivity of λ to changes in parameter values [1–3], and population viability can be predicted [4,5]. Such models can also include the effects of stochasticity [6,8] and density dependence [7,9,10]. With the practical limitations on time and resources often faced by conservation projects, these analyses are increasingly being used to help to guide species management plans by highlighting the life-history stages towards which conservation efforts should be focused.

Powerful as many of these analysis tools are, their validity ultimately rests upon the underlying model. Previous studies have found that results can depend upon the structure of the model [11–13] and even the programs used to analyse the data [14]. Overly complex models are commonly discouraged on the grounds that the results of an analysis are less clear to interpret, less robust to changes in the model's details whilst the model itself is difficult, time consuming and expensive to accurately parameterise. Yet an overly simple model may have little relevance for practical conservation issues. This raises the question: how much complexity should we incorporate in a model?

In this paper we show that simple population models can be parameterised in such a way that they are equivalent to more complex models, in terms of three essential properties; the long-term growth rate, the stable-stage distribution and the generation time. We derive this parameterisation for a commonly applied stage-structured model, where simplification involves the aggregation of stages. Other properties of the more complex model will not necessarily be conserved, and we discuss the importance of these variations. It is shown that for a sub-set of models the parameterisation of the simple model requires less effort in data collection. Finally, we present an application of our results to a model of short-tailed shearwaters (*Puffinus tenuirostris*) and discuss the differences between the complex and simple models, as well as the implications for practical use of such models.

2. The baseline model

We consider a 'baseline model' to be any stage-structured population model that can be represented by the life-cycle graph shown in Fig. 1 (Table 1). The model has C discrete stages. This structure is commonly used for modelling animal populations. For example, recent examples include studies on birds [15–18], mammals [19–22], insects [23,24] and fish [25]. The model has a
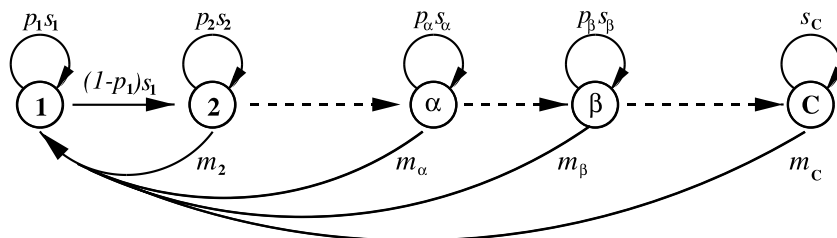

Fig. 1. The life-cycle graph for the baseline model. All individuals must pass through the same juvenile stage class 1. Note that because there are no stages beyond C, that p_C is necessarily equal to one. The parameters are explained in the text and in Table 1.

Table 1
The parameters and variables used in the analysis of the baseline stage-structured population model and its simplification (see Fig. 1)

Parameter name	Description of the parameter
C	Total number of stages of the baseline model
N	Total number of stage in the simplified model
i, j, k	Indices for a stage of the model
α	Index of the first stage in the set to be aggregated
β	Index of the last stage in the set to be aggregated
s_i	Survival probability of an individual in stage i
m_i	Fecundity of an individual in stage i
p_i	Probability that an individual stays in stage i for the following time step
n_i	Number of individuals in stage i
f_i	Proportion of the population in stage i
$\bar{s}$	Average survival probability from stages α through to β
$\bar{m}$	Average fecundity from stages α through to β
$\tilde{s}$	Survival probability of the aggregated stage
$\tilde{m}$	Fecundity of the aggregated stage
$\tilde{p}$	Probability of staying in the aggregated stage for the following time step
λ	Long-term population growth rate (variable)
$\hat{\lambda}$	Calculated long-term population growth rate for a parametrised model
T_g	Generation time of the baseline model

time step of one breeding cycle, and a projection matrix is used to evolve the present population structure forwards in time by one breeding cycle. The ith stage has a breeding population of n_i individuals (where a breeding individual is usually taken to mean a breeding female). All individuals within a stage are assumed to be identical and each stage has three parameters; the fecundity of a breeding individual m_i, the individual probability of survival until the next breeding cycle s_i and the probability that an individual stays in the same stage the following breeding cycle p_i. Since p_i is a constant, this model implies a geometric distribution for stage duration [1], with the expected duration in a stage, conditional on survival, being $(1 - p_i)^{-1}$. Our analysis could, in principle, be extended to more general life-cycles, with the only requirement being that all individuals must have at least one stage in common during their lifetime. By appropriately parametrising the model, this structure can describe more complicated scenarios. For example, emigration, immigration and proportions of the breeding population that skip breeding.

At equilibrium, ignoring stochastic effects from parameter fluctuations and finite population size, the population of each stage will be growing at the same rate, λ [1]. This implies that the proportion of the population in each stage is constant, and can be described by the characteristic equation. We can calculate λ by solving the characteristic equation, which for this model has the form [1]

$$1 = \sum_{i=1}^{C} m_i \mathscr{L}_i, \tag{1}$$

where the derived quantity $\mathscr{L}_i$ can be written in terms of the basic model parameters as

$$\mathscr{L}_i = \mathscr{L}_{i-1}s_{i-1}(1 - p_{i-1})\frac{1}{\lambda - s_i p_i} \tag{2}$$

and $\mathscr{L}_1 = 1/(\lambda - s_1 p_1)$. In a population with a constant size ($\hat{\lambda} = 1$, where $\hat{\lambda}$ is the numerical value rather than the variable), the derived quantity $\mathscr{L}_i$ can be interpreted as the probability that an individual has survived from birth up to stage i (usually denoted by l_i [26]). In a growing or shrinking population $\mathscr{L}_i$ no longer has a simple interpretation unless $p_i = 0$, in which case $\mathscr{L}_i \lambda^i$ is now the probability that an individual survives up to stage i.

We can also define the stable-stage distribution, and the generation time of the model as follows.

Definition 1. The stable-stage distribution for the baseline model is the fraction, f_i, of the population present in stage i at equilibrium [1], and can be written as

$$f_i = \frac{n_i}{\sum_{j=1}^{C} n_j} = \frac{\mathscr{L}_i}{\sum_{j=1}^{C} \mathscr{L}_j}. \tag{3}$$

Definition 2. The generation time of the baseline model, T_g, is defined as the expected mean age of mothers from a set of new-born individuals when the population has a stable-stage distribution. It can be calculated by taking the derivative of the characteristic equation (Eq. (1)) with respect to $\log(\lambda)$ [26]. The generation time can be written as

$$T_g \equiv \sum_{i=1}^{C} T(i), \tag{4a}$$

where

$$T(i) = -\lambda m_i \frac{\partial \mathscr{L}_i}{\partial \lambda} = m_i \mathscr{L}_i S_{1,i} \tag{4b}$$

and

$$S_{i,j} = \begin{cases} \sum_{k=i}^{j} \frac{\lambda}{\lambda - s_k p_k} & \text{if } i \leqslant j, \\ 0 & \text{if } i > j. \end{cases} \tag{4c}$$

There are many ways of defining a generation time [1,26], but the above definition is a natural choice which emerges from the mathematical analysis.

3. Model simplification

The baseline model can be simplified by aggregating a set of stages, say those indexed from α through to β (see Fig. 1), into a single stage. The most extreme simplification would reduce all C stages of the baseline model to just one stage. The aggregate stage will have three parameters, $\tilde{m}$, $\tilde{s}$ and $\tilde{p}$, which must be defined in some way. The simplified model carries the advantage of having fewer parameters to estimate, resulting in either reduced fieldwork effort or increased precision in estimating the parameters. However, the behaviour of the simplified model will not necessarily be

identical to the baseline model, leading to the possibility of decreased confidence in the output from a simplified model.

The correspondence between the behaviour of the simplified model and the baseline model depends upon the choice of the three parameters $\tilde{m}$, $\tilde{s}$ and $\tilde{p}$. It is of interest to consider how the choice of values for these three parameters affects the outputs from the simplified model. Important outputs are the growth rate λ, the generation time T_g, the stable-stage distribution f_i and the parameter sensitivities. Can the values of $\tilde{m}$, $\tilde{s}$ and $\tilde{p}$ be chosen so as to preserve some of these properties of the baseline model? We focus here on three of these properties: λ, f_i and T_g. The long-term growth rate is of prime importance because it is commonly used as a measure of population viability, and is often a key quantity in many analyses [1]. The stable-stage distribution enters into the model sensitivity [1] and approximations of the stochastic growth rate [6,8] by providing the relative weightings of the individual stages. Finally, the generation time scales the model's sensitivity to all parameters and emerges from a number of analyses. It should be emphasised that other properties of the baseline model will not necessarily be conserved in the simplified model. In calculating $\tilde{m}$, $\tilde{s}$ and $\tilde{p}$ such that the three properties above are preserved, we are providing an 'ideal parameterisation' upon which applications to real scenarios can be judged.

Before deriving the 'ideal parameterisation' we define the following quantities for an aggregated stage

Definition 3. For stages α through to β, an average fecundity $\bar{m}$, average survival $\bar{s}$, aggregated-stage survivorship $\overline{\mathscr{L}}$, and aggregated generation time $\overline{T}_g$ can be defined as

$$\bar{m} \equiv \sum_{i=\alpha}^{\beta} m_i \frac{\mathscr{L}_i}{\overline{\mathscr{L}}}, \tag{5a}$$

$$\bar{s} \equiv \sum_{i=\alpha}^{\beta} s_i \frac{\mathscr{L}_i}{\overline{\mathscr{L}}}, \tag{5b}$$

$$\overline{\mathscr{L}} \equiv \sum_{i=\alpha}^{\beta} \mathscr{L}_i, \tag{5c}$$

$$\overline{T}_g = \sum_{i=\alpha}^{\beta} T(i) = T_g - \sum_{i=1}^{\alpha-1} T(i) - \sum_{i=\beta+1}^{C} T(i). \tag{5d}$$

The values $\bar{m}$ and $\bar{s}$ can be interpreted as the average fecundity and average survival probability from an unbiased random sample of individuals in a population that has reached the stable-stage distribution (see Eq. (3)).

Theorem 1. *Given a baseline model with a long-term population growth rate, λ, a stable-stage distribution f_i and a generation time T_g, there exists an aggregated model whose parameterisation of the aggregated-stage $(\tilde{m}, \tilde{s}, \tilde{p})$ gives the same long-term population growth rate, same stable-stage distribution and same generation time as the baseline model. Specifically, this parameterisation is*

$$\tilde{m} = \bar{m}\frac{\overline{\mathscr{L}}}{\widetilde{\mathscr{L}}}, \tag{6a}$$

$$\tilde{s} = \tilde{s}\tilde{p} + \Psi(\lambda - \tilde{s}\tilde{p}), \tag{6b}$$

$$\tilde{p} = \frac{\tilde{s}\tilde{p}}{\tilde{s}}, \tag{6c}$$

where

$$\widetilde{\mathscr{L}} = \mathscr{L}_{\alpha-1}\frac{s_{\alpha-1}(1 - p_{\alpha-1})}{\lambda - \tilde{s}\tilde{p}}, \tag{7}$$

$$\tilde{s}\tilde{p} = \lambda\left(1 - \frac{B}{A}\right), \tag{8a}$$

$$A = \overline{T}_g - S_{1,\alpha-1}\bar{m}\overline{\mathscr{L}} + S_{\alpha,\beta}\sum_{i=\beta+1}^{C} m_i\mathscr{L}_i, \tag{8b}$$

$$B = \bar{m}\overline{\mathscr{L}} + \sum_{i=\beta+1}^{C} m_i\mathscr{L}_i \tag{8c}$$

and

$$\Psi \equiv \frac{\tilde{s}(1 - \tilde{p})}{\lambda - \tilde{s}\tilde{p}} = \frac{\mathscr{L}_{\beta+1}}{\mathscr{L}_{\alpha-1}}\frac{\lambda - s_{\beta+1}p_{\beta+1}}{s_{\alpha-1}(1 - p_{\alpha-1})}. \tag{9}$$

Proof. The proof is in three parts. First, the characteristic equation is used to define $\tilde{m}$. Second, the stable-stage distribution is used to relate $\tilde{s}$ to $\tilde{s}\tilde{p}$. Finally, the generation time for the baseline model is compared with the generation time for the simplified model, from which $\tilde{s}\tilde{p}$ is calculated, leading to expressions for $\tilde{s}$ and $\tilde{p}$.

The characteristic equation for the baseline model can be written as

$$1 = \sum_{i=1}^{C} m_i\mathscr{L}_i = \sum_{i=1}^{\alpha-1} m_i\mathscr{L}_i + \bar{m}\overline{\mathscr{L}} + \sum_{i=\beta+1}^{C} m_i\mathscr{L}_i. \tag{10}$$

Similarly the characteristic equation for the simplified model can be written as

$$1 = \sum_{i=1}^{\alpha-1} m_i\mathscr{L}_i + \tilde{m}\widetilde{\mathscr{L}} + \sum_{i=\beta+1}^{C} m_i\mathscr{L}_i. \tag{11}$$

Equating Eqs. (10) and (11) and simplifying gives Eq. (6a) for $\tilde{m}$.

For the stable-stage distribution to remain invariant to stage aggregation it is necessary that $\mathscr{L}_i$ is preserved outside the region of aggregation. Since the $\mathscr{L}_i$s are linked by the recursion relationship of Eq. (2), the $\mathscr{L}_i$ for $i < \alpha$ are unaffected by the model simplification, whilst the $\mathscr{L}_i$ for

$i > \beta$ are invariant to the model simplification if $\mathscr{L}_{\beta+1}$ remains invariant. Using Eq. (2), the requirement that $\mathscr{L}_{\beta+1}$ remains invariant can be written as

$$\mathscr{L}_{\beta+1} = \widetilde{\mathscr{L}} \frac{\tilde{s}(1 - \tilde{p})}{\lambda - s_{\beta+1} p_{\beta+1}}. \tag{12}$$

Substituting in Eq. (7) for $\widetilde{\mathscr{L}}$ and rearranging gives

$$\mathscr{L}_{\beta+1}(\lambda - s_{\beta+1} p_{\beta+1}) = \mathscr{L}_{\alpha-1} s_{\alpha-1} (1 - p_{\alpha-1}) \frac{\tilde{s}(1 - \tilde{p})}{\lambda - \tilde{s}\tilde{p}} \tag{13}$$

which can be rewritten to give Eq. (6b) for $\tilde{s}$ in terms of $\tilde{s}\tilde{p}$.

Finally, the parameter $\tilde{s}\tilde{p}$ can be calculated using the requirement that the generation time is preserved. Using Eq. (4a)–(4c), the generation time for the simplified model can be written as

$$\widetilde{T}_g = \sum_{i=1}^{\alpha-1} T(i) + \tilde{m}\widetilde{\mathscr{L}} \left(S_{1,\alpha-1} + \frac{\lambda}{\lambda - \tilde{s}\tilde{p}} \right) + \sum_{i=\beta+1}^{C} T(i) + m_i \mathscr{L}_i \left(\frac{\lambda}{\lambda - \tilde{s}\tilde{p}} - S_{\alpha,\beta} \right), \tag{14}$$

where $S_{i,j}$ is defined by Eq. (4c). Looking at the difference between Eqs. (4a)–(4c) and (14) gives

$$\sum_{i=\alpha}^{\beta} T(i) = \sum_{i=\alpha}^{\beta} m_i \mathscr{L}_i \left(S_{1,\alpha-1} + \frac{\lambda}{\lambda - \tilde{s}\tilde{p}} \right) + \sum_{i=\beta+1}^{C} m_i \mathscr{L}_i \left(\frac{\lambda}{\lambda - \tilde{s}\tilde{p}} - S_{\alpha,\beta} \right). \tag{15}$$

Rearranging Eq. (15) gives the expression for $\tilde{s}\tilde{p}$ of Eq. (8a), which can then be used to calculate $\tilde{s}$ (Eq. (6b)) and $\tilde{p}$ (Eq. (6c)). $\quad\square$

ffTwo issues arise when selecting an appropriate level of model structure for application to a real example; the equivalence between a simplified model and the baseline model, and the efficient parameterisation of the simplified model from field data. This paper concentrates upon the issue of equivalence, and does not deal in detail with the parameterisation issue. Nevertheless, it can be seen that Eqs. (5c) and (8b) require a knowledge of the parameters m_i, s_i and p_i for stages of the baseline model which are inside the region of simplification (i.e., stages α through to β). In general, therefore, the parameterisation of the simplified model depends upon the details of the baseline model, which is undesirable if the motivation for using the simplified model is to avoid excessive parameter estimation. This problem may be overcome if new sampling procedures could be developed to estimate Eqs. (5c) and (8b) more directly. At present though, it is more likely that average quantities, such as $\bar{m}$ and $\bar{s}$, can be estimated directly from field data (assuming no sampling bias between stages), without a detailed knowledge of the underlying structure in the baseline model.

For a sub-class of models, where $p_i = 0$ for all stages inside the region of simplification (as is commonly the case for Leslie matrix models), the simplified model can be derived with only a knowledge of $\bar{m}$, $\bar{s}$ and the invariant quantities λ and T_g.

Corollary 1. *If the parameters p_α through to p_β of the baseline model are zero then the simplified model can be parameterised without a detailed knowledge of the parameters m_i and s_i for stage α through to β. Specifically, Eq. (5c) for $\overline{\mathscr{L}}$ and Eq. (8b) for A can now be written in terms of the average quantities $\bar{m}$, $\bar{s}$ and parameters from outside the region of aggregation.*

$$\overline{\mathscr{L}} = \frac{s_{\alpha-1}\mathscr{L}_{\alpha-1} - \lambda\mathscr{L}_{\beta+1}\left(\lambda - s_{\beta+1}p_{\beta+1}\right)}{\lambda - \bar{s}}, \tag{16}$$

$$A = \overline{T}_g - \bar{m}\overline{\mathscr{L}}S_{1,\alpha-1} + (\beta - \alpha + 1)\sum_{i=\beta+1}^{C} m_i\mathscr{L}_i. \tag{17}$$

Eq. (16) can be substituted into the characteristic equation to solve for the long-term growth rate, which can then be used to calculate the other properties of the simplified model.

Corollary 2. *If the region of aggregation extends up to stage C (i.e., $\beta = C$) then the parameters for the aggregated region simplify still further to give*

$$\tilde{m} = \frac{\bar{m}}{\lambda - \bar{s}(1 - p_{\alpha-1})}\left(\frac{\overline{T}_g}{\bar{m}\overline{\mathscr{L}}} - S_{1,\alpha-1}\right)^{-1}, \tag{18}$$

$$\tilde{s} = \lambda\left[1 - \left(\frac{\overline{T}_g}{\bar{m}\overline{\mathscr{L}}} - S_{1,\alpha-1}\right)^{-1}\right], \tag{19}$$

$$\tilde{p} = 1, \tag{20}$$

$$\widetilde{L} = L_{\alpha-1}s_{\alpha-1}(1 - p_{\alpha-1})\left(\frac{\overline{T}_g}{\bar{m}\overline{\mathscr{L}}} - S_{1,\alpha-1}\right), \tag{21}$$

$$\overline{L} = \frac{L_{\alpha-1}s_{\alpha-1}}{\lambda - \bar{s}}. \tag{22}$$

Corollary 3. *If the parameters of the baseline model are stage independent, such that $m_i = m$, $s_i = s$, $p_i = p$ (with the exception that $p_C = 1$, and providing that $m \neq 0$ and $s \neq 0$) then the results simplify down to the well-known equations*

$$\lambda = m + s \tag{23}$$

and

$$T_g = \frac{m\lambda}{(\lambda - s)^2} = 1 + \frac{m}{s} \tag{24}$$

and the parameterisation of the aggregated stage (Eqs. (6a)–(6c)) become

$$\tilde{m} = \bar{m} = m, \tag{25a}$$

$$\tilde{s} = \bar{s} = s, \tag{25b}$$

$$\tilde{p} = \begin{cases} p & \text{if } \beta < C \\ 1 & \text{if } \beta = C. \end{cases} \tag{25c}$$

In this case no advantage is gained by analysing a complex model.

Table 2
The parameters of the age-dependent model for short-tailed shearwaters

Age class	Stage distribution	Age-dependent parameter value		
	f_i	m_i	s_i	p_i
Pre-breeders				
1	0.4652	0	0.938	0.8934
Breeders				
2	0.0384	0.1375	0.9222	0
3	0.0360	0.0945	0.9471	0
4	0.0346	0.1076	0.9590	0
5	0.0336	0.1189	0.9612	0
6	0.0328	0.1283	0.9566	0
7	0.0318	0.1358	0.9477	0
8	0.0306	0.1415	0.9367	0
9	0.0291	0.1453	0.9254	0
10	0.0273	0.1474	0.9151	0
11	0.0254	0.1481	0.9070	0
12	0.0233	0.1476	0.9017	0
13	0.0214	0.1460	0.8995	0
14	0.0195	0.1437	0.9004	0
15	0.0178	0.1407	0.9041	0
16	0.0163	0.1374	0.9096	0
17	0.0151	0.1339	0.9160	0
18	0.0140	0.1305	0.9217	0
19	0.0131	0.1273	0.9248	0
20	0.0123	0.1244	0.9231	0
21	0.0115	0.1221	0.9140	0
22	0.0107	0.1204	0.8946	0
23	0.0097	0.1195	0.8615	0
24	0.0085	0.1195	0.8111	0
25	0.0070	0.1195	0.8111	0
26	0.0151	0.1230	0.6415	1.0000

m is the fecundity, s is the probability of surviving one time step, and p is the probability of staying in the same age class until the following year. Pre-breeding females are represented by the first age class, and breeding females by the following 25 age classes. This model gives a long-term growth rate of $\hat{\lambda} = 0.986$ year^{-1} and a generation time of $T_g = 15.6$ years.

Other quantities of interest, such as the various measures of model sensitivity [1] and approximations to the stochastic growth rate, can be calculated once the simplified model has been fully parameterised. In general, these quantities will differ from the baseline model, and will be the deciding factor in choosing the degree of model complexity. This choice is illustrated by the following example.

3.1. Example: a model of short-tailed shearwaters

We will apply our results to a model of the short-tailed shearwater (*Puffinus tenuirostris*), developed by Hunter et al. [18]. Simplifications of this model have been investigated numerically

[27], with the conclusion that the 25 breeding age classes in the model add little to the model's prediction of long-term growth rate. To simplify our analysis we aggregate the original 15 pre-breeding age classes into one pre-breeding class. Our baseline model therefore has 26 age classes (1 pre-breeder class and 25 breeder classes, Table 2).

This model was progressively simplified. Each simplification involved aggregating from stage N through to stage 26, giving a simplified model with N stages ($N = 2, \ldots, 26$). Table 3 shows the survival and fecundity parameters of the aggregated stage (Eqs. (6a)–(6c)) for the series of simplified models, together with the mean elasticities, and the mean ratio of elasticities. All models, irrespective of the number of stages, have the same long-term growth rate, generation time and

Table 3
Simplifications of the short-tailed shearwater model obtained by successively aggregating stages containing the oldest birds

Number of stages in the model, N	Parameters of the aggregated stage		Mean elasticities		Mean ratio of elasticities
			Fecundity	Survival	
	$\tilde{m}$	$\tilde{s}$	$\bar{e}_m$	$\bar{e}_s$	
26	0.12	0.64	0.0018	0.17	5.3
25	0.13	0.65	0.0018	0.17	5.3
24	0.14	0.67	0.0018	0.17	5.4
23	0.15	0.7	0.0019	0.17	5.4
22	0.16	0.72	0.0020	0.17	5.4
21	0.16	0.74	0.0021	0.17	5.4
20	0.17	0.76	0.0022	0.17	5.4
19	0.18	0.77	0.0024	0.18	5.4
18	0.18	0.79	0.0026	0.18	5.4
17	0.18	0.8	0.0029	0.18	5.4
16	0.18	0.81	0.0032	0.18	5.4
15	0.19	0.82	0.0037	0.19	5.4
14	0.19	0.82	0.0043	0.19	5.4
13	0.19	0.83	0.0051	0.20	5.4
12	0.19	0.84	0.0061	0.20	5.4
11	0.19	0.84	0.0073	0.21	5.3
10	0.19	0.85	0.0088	0.23	5.3
9	0.19	0.85	0.011	0.24	5.3
8	0.19	0.86	0.013	0.26	5.2
7	0.19	0.86	0.015	0.28	5.2
6	0.20	0.87	0.017	0.30	5.1
5	0.20	0.87	0.020	0.34	4.9
4	0.19	0.88	0.023	0.37	4.8
3	0.18	0.88	0.027	0.42	4.5
2	0.18	0.89	0.032	0.47	4.5

The parametrisation of the aggregated stage, and the mean elasticity of the long-term growth rate to changes in fecundity and survival are shown (to two significant figures) for successive model simplifications. The parameter $\tilde{p}$ is unity for all models, and is not shown. The long-term growth rate for all models is $\hat{\lambda} = 0.986$ year^{-1} and the generation time is $T_g = 15.6$ years. The mean elasticities are defined as $\bar{e}_m = \sum_{i=1}^{N} f_i e_m(i)$ and $\bar{e}_s = \sum_{i=1}^{N} f_i e_s(i)$ and the mean ratio of the elasticities is $\sum_{i=1}^{N} f_i e_s(i)/e_m(i)$, where $e_s(i) = d\log(\lambda)/d\log(s_i)$, $e_m(i) = d\log(\lambda)/d\log(m_i)$ and f_i is the proportion of the population at equilibrium in stage i (Eq. (3)).

stable-stage distribution. However, there are differences between the models in other respects. For example, as the number of stages in the model is reduced, the mean elasticity of both fecundity and survival increases, implying that the model's predicted growth rate is becoming more sensitive to uncertainties in the parametrisation. Furthermore, the mean ratio of the elasticities tends to decrease, showing that the model is becoming relatively more sensitive to uncertainties in the fecundities as it is simplifed (to calculate the average elasticities it was assumed that there was no covariance between stages). The main point is that not all properties of a complex model can be conserved when creating a simplified model, and these differences can be important. The undesirable effects of covariation between a model's structure and quantities of interest have to be weighed up against the benefits of model simplification, which include increased precision in parameter estimation and reduced data collection effort.

4. Discussion

In this paper we considered a commonly applied stage-structured population model and so by ignoring effects such as stochasticity there is already an implicit simplification in all our models. We have focused upon the structure of the model and the effect of simplifying its structure by aggregating stages. In general, the conclusions from a simplified model depend upon the parametrisation of the aggregated stage. It is shown how to parameterise an aggregated stage so that the simplified model and the original 'baseline' model are identical with respect to three essential properties; the long-term growth rate, stable-stage distribution and generation time. For a sub-set of models, this parametrisation is possible with a knowledge of only the average fecundity and survival for the aggregated stages. In this case the simplified model will require less fieldwork effort to parameterise. For all other cases, either new sampling techniques must be developed to estimate the relevant average quantities, or some knowledge of the original baseline model is required, although this last option, in part, limits the usefullness of a simplified model.

In general, the other properties of the simplified model will not be the same as for the 'baseline' model. Quantities such as the expected lifespan, parameter sensitivity and stochastic growth rate will be a function of a model's structure. The application of our results to a model for short-tailed shearwaters suggests that although precision in parameter estimation may be increased by using a simple model, parameter elasticity may also increase. This might reduce the benefits of using a simple model. In using these models, we need to strike a balance between the benefits of a simplified model and the costs of poor prediction.

Previous studies have also addressed the question of model complexity and structure. Vandermeer [28] and later Moloney [29] proposed algorithms to choose the size categories in a stage-structured model. In practice these algorithms are rarely used, and can give misleading results [1], whilst alternative methods have since been developed which avoid the use discrete stages altogether [30]. More recently the question of model complexity has been addressed by either numerically comparing various models [27], or by defining an index of utility, with which models can be compared [31]. These studies are motivated by the growing application of models to conservation and management decisions, where the choice of model structure may have a critical effect upon the accuracy of a prediction. The general conclusion from these studies is that complexity is rarely justifiable on its own, and that intermediate levels of complexity are often more advisable.

Other studies have shown that simplified models can capture the essential details of a full matrix model [32,33]. Heppel et al. [34] found that use of simple age-classified models compared favourably with the use of full Leslie matrix models in terms of comparing patterns of elasticity, suggesting that complete life-history information is not important when looking at the elasticity patterns of a model.

The results presented here allow the accuracy of the predictions from a simplified model to be calculated, given that the model is 'ideally parameterised', so as to conserve the essential properties of a 'baseline' model. In this way the problem of filling the theoretical gap between a model's structure and the accuracy of its predictions is divided into two parts. Firstly, the behaviour of a simplified model given an ideal parametrisation can be studied. This can be viewed as the effect of model structure. Secondly, the effect of deviations, due to data limitations, from the ideal parametrisation can be considered. Such deviations from the ideal parametrisation have not been considered in this paper and are a topic of current research. An improved theoretical understanding of both of these issues is needed to guide the appropriate level of model complexity in applied situations. Appropriate model selection may be an important consideration for the accurate guidance of conservation and management projects, especially when time and funding is limited.

Acknowledgements

References

[1] H. Caswell, Matrix Population Models: Construction, Analysis, and Interpretation, 2nd Ed., Sinauer Associates, MA, 2000.

[2] T.G. Benton, A. Grant, Elasticity analysis as an important tool in evolutionary and population ecology, Trends Evol. Ecol. 14 (1999) 467.

[3] H. de Kroon, J. van Groenendael, J. Ehrlen, Elasticities: a review of methods and model limitations, Ecology 81 (2000) 607.

[4] M.S. Boyce, Population viability analysis, Ann. Rev. Ecol. Syst. 23 (1992) 481.

[5] J.M. Reed, D.D. Murphy, P.F. Brussard, Efficacy of population viability analysis, Wildlife Soc. Bull. 26 (1998) 244.

[6] S.D. Tuljapurkar, Population dynamics in variable environments. III Evolutionary dynamics of r-selection, Theor. Populat. Biol. 21 (1982) 141.

[7] T. Takada, H. Nakajima, An analysis of life history evolution in terms of the density-dependent Lefkovitch matrix model, Math. Biosci. 112 (1992) 155.

[8] S.D. Tuljapurkar, Stochastic population forecasts and their uses, J. Forecast. 8 (1992) 385.

[9] A. Grant, Selection pressures on vital rates in density-dependent population, Proc. Roy. Soc. London. Ser. B 264 (1997) 303.

[10] A. Grant, T.G. Benton, Elasticity analysis for density-dependent populations in stochastic environments, Ecology 81 (2000) 680.

[11] N.J. Enright, M. Franco, J. Silvertown, Comparing plant life histories using elasticity analysis, Oecologia 104 (1995) 79.

[12] M.B. de Matos, D.M.S. Matos, Mathematical constraints on transition matrix elasticity analysis, J. Ecol. 86 (1998) 706.

[13] L.S. Mills, D.F. Doak, M.J. Wisdom, Reliability of conservation actions based on elasticity analysis of matrix models, Conservat. Biol. 13 (1999) 815.

[14] L.S. Mills, S.G. Hayes, C. Baldwin, C.J. Wisdom, D.J. Mattson, K. Murphy, Factors leading to different viability predictions for a grizzly bear data set, Conservat. Biol. 10 (1996) 863.

[15] S. Wanless, M.P. Harris, J. Calladine, P. Rothery, Modelling responses to herring gull and lesser black-backed gull populations to reduction of reproductive output: implications for control measures, J. Appl. Ecol. 33 (1996) 1420.

[16] J. Real, S. Mañosa, Demography and conservation of western european Bonelli's eagle *Hieraaetus fasciatus* populations, Biologic. Conservat. 79 (1997) 59.

[17] R.N. Holdaway, C. Jacomb, Rapid extinction of the moas (Aves: Dinorinthiformes): Model, test, and implications, Science 287 (2000) 2250.

[18] C.M. Hunter, H. Moller, D. Fletcher, Parameter uncertainty and elasticity analyses of a population model: setting research priorities for shearwaters, Ecologic. Modell. 134 (2000) 299.

[19] S. Focardi, S. Toso, E. Pecchioli, The population modelling of fallow deer and wild boar in a Mediterranean ecosystem, Forest Ecol. Mgmt. 88 (1996) 187.

[20] B.A. Mayle, Progress in predictive management of deer populations in British woodlands, Forest Ecol. Mgmt. 88 (1996) 187.

[21] Z.H. Zhou, W.S. Pan, Analysis of the viability of a giant panda population, J. Appl. Ecol. 34 (1997) 363.

[22] J. Carter, A.S. Ackleh, B.P. Leonard, H. Wang, Giant panda (*Ailuropoda melanoleuca*) population dynamics and bamboo (subfamily Bambusoideae) life history: a structured population approach to examining carrying capacity when the prey are semelparous, Ecologic. Modell. 123 (1999) 207.

[23] R. Wall, K.E. Smith, E. Berriatua, N.P. French, Simulation analysis of the population dynamics of the mite, *Psoroptes ovis*, infesting sheep, Veterinary Parasitol. 83 (1999) 253.

[24] V. Křivan, J. Havelka, Leslie model for predatory gall-midge population, Ecologic. Modell. 126 (2000) 73.

[25] S. Charles, R.B. de la Parra, J.-P. Mallet, H. Persat, P. Auger, Annual spawning migrations in modelling brown trout population dynamics inside an arborescent river network, Ecologic. Modell. 133 (2000) 15.

[26] B. Charlesworth, Evolution in Age-Structured Population, Cambridge University, Cambridge, UK, 1980.

[27] C.M. Hunter, D. Fletcher, Implications of reducing stage-structured models for setting research priorities for shearwaters, Ecologic. Modell., submitted for publication.

[28] J. Vandermeer, Choosing catergory size in a stage projection matrix, Oecologia 32 (1978) 79.

[29] K.A. Moloney, A generalised algorithm for determining category size, Oecologia 69 (1986) 176.

[30] M.R. Easterling, S.P. Ellner, P.M. Dixon, Size-specific sensitivity: applying a new structured population model, Ecology 81 (2000) 694.

[31] S.D. Snowling, J.R. Kramer, Evaluating modelling uncertainty for model selection, Ecologic. Modell. 138 (2001) 17.

[32] S. Brault, H. Caswell, Pod-specific demography of killer whales (*Orca orca*), Ecology 74 (1993) 1444.

[33] L.A. Levin, H. Caswell, T. Bridges, C. DiBacco, D. Cabrera, G. Plaia, Demographic responses of estuarine polychaetes to pollutants: life table response experiments, Ecologic. Appl. 6 (1996) 1295.

[34] S.S. Heppell, H. Caswell, L.B. Crowder, Life histories and elasticity patterns: perturbation analysis for species with minimal demographic data, Ecology 81 (2000) 654.