Parameter estimation for stage-structured projection models using real data

Lynne McArthur[a], **John Boland**[a] **and Fleur Tiver**[b]

[a]Centre for Industrial and Applied Mathematics, University of South Australia, Mawson Lakes, Sth Australia

[b]Sustainable Ecosystems Group, University of South Australia, Mawson Lakes, South Australia.

Abstract: Herbivores, particularly introduced ones such as sheep and rabbits, have reduced the levels of regeneration of a number of species of native trees and shrubs in the arid rangelands of southern Australia since European settlement. These reduced levels of regeneration mean that populations of some species are declining whilst others, less palatable to herbivores, are being maintained, hence gradually altering the composition and/or overall cover of vegetation. In the long term, this could result in loss of palatable species valuable to the pastoral industry for animal production. To investigate the effects of grazing, we describe a model which projects the age-structured populations of some native Australian tree and shrub species under grazing and non-grazing regimes. The technique we are applying is an extension of the Leslie growth model, where the grazing and rainfall events are incorporated by the use of separate transition matrices. The preliminary results indicate that rangeland management practices will need to be adjusted to ensure the long-term sustainability, biodiversity and usability of the arid regions. The purpose of this paper is to frame the context for study in the application of this methodology, as well as to indicate the problems involved in parameter estimation and delineate the extent of our deliberation thus far.

Keywords: *Leslie projection matrix; matrix population models; stochastic environmental factors; arid rangelands; grazing; long-term growth rate.*

1. INTRODUCTION

Ecological processes commonly operate over a wide range of time scales, (Tyre et al. 2000), often over decades or longer. This is particularly so with woody trees and shrubs, most of which have a longer life-span than humans, meaning that studying them by traditional direct monitoring techniques will take too long to generate results that can be used to modify existing management practices.

Several studies have indicated the adverse effects of grazing by sheep, kangaroos and rabbits in the arid and semi-arid regions of Australia (Cunningham and Walker 1973; Lange and Purdie 1976; Crisp 1978; Crisp and Lange 1976; Harrington 1979; Lange and Willcocks 1980; Silander 1983; Lange and Graham 1983; Chesterfield and Parsons 1985; Lange 1985; Eldridge et al. 1990; Auld 1995; Tiver and Andrew 1997; Hunt 2001).

Because of the long life-spans of woody trees and shrubs, there has tended to be comparatively little research to follow the progress of plant cohorts. However, since we know that herbivores are affecting regeneration of trees and shrubs in the arid zone, it is essential that we develop tools that will enable us to predict the sorts of changes to communities which are occurring as a result. It is only when we can predict these changes, and relate these to the types of herbivores and levels of grazing, that we will be able to make recommendations for pastoral and biodiversity management.

Age or stage structured matrix models are widely used in ecology to study population dynamics. These models, which are based upon the Leslie growth matrix (Leslie 1945) are used to determine the long-term growth rate and to project future populations under prescribed conditions. The 'conditions' in the present study include different grazing regimes and environmental factors such as rainfall events.

1.1 Matrix Population Models

Matrix population models (MPM) predict future populations from past states (Caswell 2001). The models can be represented by

$$x(t+1) = Ax(t) \tag{1}$$

where

$$A = \begin{bmatrix} P_1 & F_2 & \cdots & F_{n-1} & F_n \\ S_1 & P_2 & \cdots & 0 & 0 \\ 0 & S_2 & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & S_{n-1} & P_n \end{bmatrix} \tag{2}$$

is the transition matrix between populations at time t and $t+1$. The vector $\boldsymbol{x}(t)$ is the n x 1 population vector at time t, with $n = 1,2,3,\ldots$, consisting of the number of individuals in each of the n classes. The parameters F_i, and S_i, where $F_i > 0$, $0 \leq S_i \leq 1$, $i = 1, 2,\ldots, n$, refer to the fertility and survival rates of the individual stages respectively. The survival rates represent the probability of surviving from class i to class $i + 1$. The parameters P_i, $0 \leq P_i \leq 1$, represent the probability of staying in the same class given survival. These rates and probabilities are collectively known as the *vital rates*.

Using matrix algebra, the population after $t = k$ time steps is therefore given by

$$x(k) = A^k \, x(0) \tag{3}$$

where vector $\boldsymbol{x}(0)$ is the initial n x 1 population vector.

Matrix population models can be easily adapted to 'stage'-structured models, rather than 'age'-structured models. This is particularly useful when there is no direct linear relationship between the age and the size of the individual, creating difficulties in age determination (Caswell 2001). The probability of remaining in the same stage-class can be estimated from information on the duration of the stages.

The challenge of these models, as with all models, is to estimate the parameters, which in this case are the vital rates. With short-lived species, estimation of the vital rates is possible through observation and controlled experiments. However, many of the native tree and shrub species in South Australia live for more than 250 years. This makes it very difficult to measure or even estimate survival rates or indeed longevity accurately. In addition, the fertility parameters refer to the number of offspring produced, on average, by each stage class. This can often be estimated from animal populations reasonably accurately, but with plant populations, the formation, distribution, germination and survival of seeds is a process that is dependent on many varied and random events. In addition, it is often not possible to determine from which age-class the seeds originated.

1.2 Model Description

The models will be constructed for individual tree and shrub species. The life cycle is divided into stage-classes, as defined by Tiver and Andrew (1997) and for this study five classes are used as outlined in Table 1 for *Myoporum platycarpum.* The third column in Table 1 contains the relative fertility, which is discussed in Section 2.2.

Determination of the long-term growth rate, λ, which is the largest real eigenvalue of the projection matrix, indicates the likelihood of sustainability. Obviously, this parameter is a function of the vital rates of the matrix. In many ways our focus is to fully delineate this dependence and thus determine the most significant contributing factors to the long-term sustainability of the population.

Many theoretical studies have been undertaken in the last few decades on the dynamics of populations evolving in random environments (Sykes 1969; Gerrodette *et al.* 1985; Mode and Root 1998; Tuljapurka 1990; Runge and Moen 1998; Caswell 2001; Todd and Ng 2001; Hunt 2001) and citations therein. There are essentially three methods of incorporating stochasticity into population projections (Sykes 1969). The first method concerns the addition of random errors into a deterministic model. The second method describes the vital rates as random variables with specified means, variances and covariances between the vital rates. The last method uses several transition matrices that are randomly selected for each realisation of the projection. The method that we feel is most suitable for this study is the last. We have decomposed the projection matrix $\boldsymbol{A}$ in a manner that reflects differing conditions comprising grazing or restriction from grazing and random significant rainfall events. This is implemented by constructing four separate projection matrices specifically reflecting these events resulting in a model which comprises four projection matrices in Table 2, for each species.

Table 1. The Population Stages with associated relative fertilities for *Myoporum platycarpum.* The juvenile stage is non-reproductive.

Stage	Description (Duration years)	Relative Fertility, *f*
I, II, III	juvenile, (10)	0.0
IV	young mature (30)	0.75
V	mature (45)	1.00
VI	old mature (150)	0.5
VII	senescent (300)	0.2

Table 2. The projection matrices corresponding to the four different regimes.

Regime	Projection Matrix
Ungrazed/ no rain	$\boldsymbol{U}_n$
Ungrazed/ rain	$\boldsymbol{U}_r$
Grazed/ no rain	$\boldsymbol{G}_n$
Grazed/ rain	$\boldsymbol{G}_r$

To illustrate this method, consider the following sequence of events: On a particular paddock grazing is allowed for two years with a significant rainfall event in the second year, followed by two years of no grazing and one more year of grazing with no significant rainfall. This would then be represented by

$$x(5) = G_n U_n^2 G_r G_n x(0)\,. \tag{4}$$

There are two aspects of the problem well illustrated by this example. One is that in order to simulate the system under random rainfall events we will have to perform statistical and time series analysis on rainfall to determine return rates. Secondly, the sequencing of grazing and non-grazing regimes will significantly alter the long-term sustainability of the populations. Simulation using various permutations of the four projection matrices will therefore give an indication of the optimal grazing regime. In addition, the combination can be analysed by calculating the growth rate for each sequence. This presents a valuable opportunity for managers and graziers to perform 'what if' scenarios.

1.3 Sensitivity Analysis

Lambda is the growth rate, and a value greater than one indicates that the population is increasing. A value less than one will indicate a declining population. The magnitude of the change in lambda due to a change in one of the vital rates is known as the *sensitivity* of lambda to that particular rate. Perhaps more useful quantities are the elasticities of lambda, which are interpreted as the relative contributions of each of the vital rates to lambda. For example we might say that the probability of surviving stage II and progressing to stage III contributes about 40% to the value of lambda. This then might indicate the most important vital rates on which future experiments should be based. If for example, a model indicates that a population is very sensitive to some factor about which no data exist, then it is vital to collect such data. Sensitivity analysis will thus indicate which of the vital rates are the most important in determining long-term sustainability of the species. The method is applied to the problem of describing the effects of sheep grazing on *Myoporum platycarpum* as follows.

2. MATRIX PROJECTION MODEL OF *MYOPORUM PLATYCARPUM*

2.1 Introduction

Preliminary investigation was carried out on *Myoporum platycarpum*, a tree commonly known as false sandlewood. Experiments conducted by Tiver (2002) on this species include monitoring the survival and growth of *Myoporum* under three different fencing regimes: those enclosed in rabbit proof fencing, those enclosed in sheep proof fencing and those without fencing. The plants are all in stage two, being less than four years old. In addition, counts of the population in each stage class have been conducted at three different sites at Middleback station, about 400 km north-west of Adelaide.

2.2 Fertility and Rainfall Factors

The fertility of the different stages is composed of the probability of producing seeds, multiplied by the probability of those seeds germinating, multiplied by the probability of surviving to the seedling stage. Although it is not possible to obtain accurate estimates of all of these probabilities, it is possible to estimate the fertilities themselves, using knowledge of the average relative flowering capacity and frequency of the five stages of the species. This produces the relative values given in Table 1. The absolute values of fertility can then be determined by optimisation methods, under the assumption of population stability, given no grazing. This assumption we feel is valid, given the longevity of the species, and in the absence of information to the contrary. Thus, if the population is stable then lambda must be close to one, and using this information a *fertility factor*, ϕ, can be determined which will multiply the fertility ratios, f, to produce F_i. In addition, a *rainfall factor*, ρ_n, was introduced under the assumption that some rain must fall for the species to survive. This is in contrast to the 'significant rainfall events' which occur randomly every 20 years or so. Assume that an *ideal* rainfall regime has a rainfall factor equal to one. The years without significant rainfall will reduce rates of survival, germination and growth and so ρ_n will be less than one. Assume also that in the years where significant, or greater than average, rainfall is observed the rainfall factor is doubled, so $\rho_r = 2\rho_n$.

3. PARAMETER ESTIMATION

3.1 Initial Populations

Since we are dealing with a stage-structured model rather than an age-structured model, the duration, d_i, of each stage must be taken into account when determining the vital rates in each class. Since these species are so long-lived, the estimates of the duration of each stage has to be estimated from experimental data and historical records (Tiver and Andrew 1997), primarily calibrated to major rainfall events. The estimates for the duration of each stage for *Myoporum platycarpum* are given in Table 1. The initial population vectors for the grazed and ungrazed sites are, respectively,

$$x_u(0) = \begin{bmatrix} 556 \\ 205 \\ 28 \\ 58 \\ 40 \end{bmatrix} \quad x_g(0) = \begin{bmatrix} 217 \\ 236 \\ 44 \\ 218 \\ 134 \end{bmatrix}. \tag{5a,b}$$

Determining whether the ungrazed distribution is a *stable* stage-distribution is not a simple task, since one needs to know the survival probabilities and the growth rate to calculate the ratios. (Caswell 2001). However, since we are assuming for the present that the ungrazed population is in equilibrium, we will assume a stable stage-distribution. This assumption allows calculation of the transition probabilities for the ungrazed projection matrix as described below. There are further considerations, not discussed here, regarding the age distributions within each stage.

3.2 Fertility

The stages were estimated to have fertilities relative to stage V, the most fertile. These relative fertilities are given in Table 1. The fertility factor was found to be 49.8, after optimizing the projection matrix U_n, (ungrazed/no rain) for a lambda equal to one. In addition, during the same optimization process, the rainfall factor, for periods outside the significant rainfall events, was determined to be 0.19. Fertility can now be estimated as the product of the ratios, the fertility factor and the rainfall factor, for each stage, i, as

$$F_i = f_i \phi \rho, \qquad i = 2,3,4,5. \tag{6}$$

3.3 Survival

Data from Tiver (Tiver 2002) indicates that the probability of being grazed given that there is no protective fencing is 0.95, and the probability of dying, given that the individual has been grazed is 0.75. Thus, under a grazing regime, using Bayes' Law, there would be a survival rate for the juvenile stage in the grazed regime of the order of $\sigma_{g1} = 0.237$ or 23.7%. The probability, on the other hand, of surviving without grazing was estimated to be $\sigma_{u1} = 0.62$.

These rates will also be affected by rainfall events and given that the seedlings in the fencing experiment were watered, we define the probability of surviving and progressing from the juvenile stage, I-III, to stage IV as

$$S_1 = \sigma_{u1}\rho. \tag{7}$$

Survival from stage IV to V, V to VI and VI to VII, given by $\boldsymbol{S}_2, \boldsymbol{S}_3, \boldsymbol{S}_4, \boldsymbol{S}_5$ respectively, are simply determined by the ratios of the effective populations in each stage, taking the stage duration d_i into account. Thus,

$$S_2 = \frac{x_{u3}/d_3}{x_{u2}/d_2}, \qquad S_3 = \frac{x_{u4}/d_4}{x_{u3}/d_3},$$

$$S_4 = \frac{x_{u5}/d_5}{x_{u4}/d_4} \tag{8}$$

where the x_{ui} are the entries in the population vector given in equation (5a).

The probability of surviving and staying in the same class, P_i, is determined by the ratio of the duration of stage i to the estimated average total lifespan, for all stages except the juvenile stage. We believe that this stage would be heavily influenced by rainfall events, and although the other stages would also be affected, the degree would be much less. Thus

$$P_1 = \frac{d_1}{\sum_{i=1}^{n} d_i}\rho \text{ and } P_i = \frac{d_i}{\sum_{i=1}^{n} d_i}, \text{ for } i = 2,3,4,5. \tag{9}$$

4. RESULTS

4.1 The Four Projection Matrices

The four projection matrices and their corresponding entries are given by

$$U_n = \begin{bmatrix} 0.0035 & 6.9637 & 9.2850 & 4.6424 & 1.8570 \\ 0.1160 & 0.0561 & 0 & 0 & 0 \\ 0 & 0.0911 & 0.0841 & 0 & 0 \\ 0 & 0 & 0.6214 & 0.2804 & 0 \\ 0 & 0 & 0 & 0.3448 & 0.5607 \end{bmatrix} \tag{10a}$$

$$U_r = \begin{bmatrix} 0.0070 & 13.9900 & 18.6533 & 9.3267 & 3.7307 \\ 0.2320 & 0.0561 & 0 & 0 & 0 \\ 0 & 0.0911 & 0.0841 & 0 & 0 \\ 0 & 0 & 0.6214 & 0.2804 & 0 \\ 0 & 0 & 0 & 0.3448 & 0.5607 \end{bmatrix} \tag{10b}$$

$$G_n = \begin{bmatrix} 0.0029 & 6.9637 & 9.2850 & 4.6425 & 1.8570 \\ 0.0444 & 0.0472 & 0 & 0 & 0 \\ 0 & 0.1243 & 0.0709 & 0 & 0 \\ 0 & 0 & 0.8918 & 0.3937 & \vdots \\ 0 & 0 & 0 & 0.5122 & 0.4724 \end{bmatrix} \tag{10c}$$

$$G_r = \begin{bmatrix} 0.0059 & 13.9275 & 18.5700 & 9.2850 & 3.7140 \\ 0.0889 & 0.0472 & 0 & 0 & 0 \\ 0 & 0.1243 & 0.0709 & 0 & 0 \\ 0 & 0 & 0.8918 & 0.3937 & 0 \\ 0 & 0 & 0.5122 & 0.5122 & 0.4724 \end{bmatrix}. \tag{10d}$$

It can be seen from the above four equations that only the rates in the first two rows are affected by the extreme rainfall events. The fertilities, although not integers, represent the average expected number of offspring that are likely to survive. These values, along with the probability of surviving and staying in stage I-III, and the probability of surviving and progressing to stage IV are both doubled in the event of a significant rainfall event.

4.2 Growth Rates of the Projection Matrices

The dominant eigenvalues, λ, for the projection matrices were found using MATLAB©. This value for $\boldsymbol{U}_\mathrm{n}$ was assumed to be equal to one, in order to determine the fertility and rainfall factors. An iterative process of updating λ each time estimates are found possibly improves the estimates, but trials revealed that successive changes in lambda are of the order of 10^{-3}. The remaining growth rates are given in Table 3. As can be seen, the long-term growth rate for the grazed/no rain regime is much less than one. This implies a declining population. The growth rates for the 'rain' regimes both indicate that the population will increase.

The results of the elasticity analysis consistently show across all projection matrices that the most important vital statistic is the survival and progression of the juveniles to young mature adults, which confirms the findings of Tiver and Andrew (Tiver and Andrew 1997) and others. The elasticities of λ to S_i is given in Table 3.

Further analysis can produce the elasticities of lambda to the 'lower-level' parameters, such as the growth rates and probabilities which contribute to the matrix entries (Caswell 2001). These lower level elasticities can then distinguish the effects of the various factors.

Table 3. The growth rates for the four projection matrices, corresponding to the four different regimes. The third column contains the elasticities of lambda to S_1 and F_2.

Projection Matrix	Growth Rate, λ	Dominant elasticities	
		S_1	F_2
$\boldsymbol{U}_\mathrm{n}$	1.0175	0.4240	0.3531
$\boldsymbol{U}_\mathrm{r}$	1.9047	0.4696	0.4346
$\boldsymbol{G}_\mathrm{n}$	0.7598	0.2894	0.1661
$\boldsymbol{G}_\mathrm{r}$	1.2605	0.4215	0.3428

The calculation of the growth rate for the sequence of matrices requires a different approach (Caswell 2001). If the model is to be constructed as described above and illustrated by equation (4), then the *stochastic growth rate*, log λ_s, adopting the notation of Caswell (2001) must be calculated.

5. DISCUSSION

This paper is presented as a preliminary discussion on the possibilities and problems associated with modelling when using real data. In this study, all of the parameters are estimates based on the data or from constrained optimisation. However, they are all biologically realistic, and as mentioned previously, these models can be used as a tool to illustrate consequences given a certain set of assumptions and premises.

The analysis of the growth rate and the elasticities has indicated that the grazing regime in the absence of significant rainfall could cause the population to diminish. Secondly, the elasticity analysis has indicated that it is important to establish more accurate estimates of juvenile survival. An extension of the sensitivity analysis will include the sensitivity of lambda to the lower level parameters, such as the rainfall and fertility factors. As well, the analysis presented here will aid in the design of controlled experiments to be performed in the future.

Further studies could also include delineating the effects of different levels of grazing, and in a similar fashion to Hunt (2001), investigate the variation in the effects of grazing near to watering points.

The estimates presented here are point estimates, which will need confidence limits. The authors intend to expand the study to include vital rates and stage durations that are represented by probability distributions, with given means, variances and covariances. For example, stage durations can be approximated by negative binomial distributions (Caswell 2001). This is logical, since negative binomial distributions are used to determine the probability of an event happening after a certain period of time. This expansion would also include investigation of the total population distributions from simulations.

The age-structure within the stages has also largely been ignored, but there is very little information regarding the life-span of this species, as mentioned above.

Another area of investigation will focus on rainfall. In order to provide a realistic sequence of events as described in equation (4), time series analysis of rainfall for locations in the rangelands needs to be performed. Thus, we can determine the probable return rates and the distribution of intensities. In the subsequent analysis we will also investigate the possibility of using different rainfall factors for fertility and survival. As pointed out by Lange (Lange 1992) more accurate rates cannot be estimated when studies encompass only a small fraction of the life-span of the individuals.

6. REFERENCES

Auld, T. D., The impact of herbivores on regeneration in four trees from arid Australia, *Rangeland Journal,* **17** (2) 213-227, 1995.

Caswell, H., *Matrix Population Models: Construction, Analysis, and Interpretation*, Sinauer Associates, Sunderland, MA, 2001.

Chesterfield, C. J. and R. F. Parsons, Regeneration of three tree species in arid south-eastern Australia, *Australian Journal of Botany,* **33** 715-732, 1985.

Crisp, M. D., Demography and survival under grazing of three Australian semi-desert shrubs, *Oikos,* **30** 520-528, 1978.

Crisp, M. D. and R. T. Lange, Age structure, distrubtion and survival under grazing of the arid-zone shrub *Acacia burkittii*, *Oikos,* **27** 86-92, 1976.

Cunningham, G. M. and P. J. Walker, Growth and survival of mulga (*Acacia aneura F. Muell. Ex Benth*) in western New South Wales, *Tropical Grasslands,* **7** (1) 69-77, 1973.

Eldridge, D. J., M. Westoby and R. J. Stanley, Population dynamics of the perennial rangeland shrubs *Atriplex vesicaria*, *Maireana astrotricha* and *M. pyramidata* under grazing, 1980-1987, *Journal of Applied Ecology,* **27** 502-512, 1990.

Gerrodette, T., D. Goodman and J. Barlow, Confidence limits for population projections when vital rates vary randomly, *Fishery Bulletin,* **83** (3) 207-217, 1985.

Harrington, G., The effects of feral goats and sheep on the shrub populations in a semi-arid woodland, *Australian Rangeland Journal,* **1** (4) 334-345, 1979.

Hunt, L. P., Heterogeneous grazing causes local extinction of edible perennial shrubs: a matrix analysis, *Journal of Applied Ecology,* **38** 238-252, 2001.

Lange, R. T., Spatial distributions of stocking intensity produced by sheepflocks grazing Australian chenopod shrublands, *Transactions of the Royal Society of South Australia,* **109** (4) 167-174, 1985.

Lange, R. T., Growth rates of western myall (*Acacia papyrocarpa* Benth.) during its main phase of canopy spreading, *Australian Journal of Ecology,* **17** 315-320, 1992.

Lange, R. T. and C. R. Graham, Rabbits and the failure of regeneration in Australia arid zone *Acacia*, *Australian Journal of Ecology,* **8** 377-381, 1983.

Lange, R. T. and R. Purdie, Western myall (*Acacia sowdenii*), its survival prospects and management needs, *Australian Rangeland Journal,* **1** 64-69, 1976.

Lange, R. T. and M. C. Willcocks, Experiments on the capacity of present sheep flocks to extinguish some tree populations of the South Australian arid zone, *Journal of Arid Environments,* **3** 223-229, 1980.

Leslie, P. H., On the use of matrices in certain population mathematics, *Biometrika,* **33** 183-212, 1945.

Mode, C. J. and T. Root, Projecting Age-Structured Populations in a Random Environment, *Mathematical Biosciences,* **88** (2) 223-245, 1998.

Runge, M. and A. Moen, A modified model for projecting age-structured populations in random environments, *Mathematical Biosciences,* **150** 21-41, 1998.

Silander, J. A., Demographic variation in the Australian desert cassia under grazing pressure, *Oecologia,* **60** 227-233, 1983.

Sykes, Z., Some Stochastic versions of the Matrix Model for Population dynamics, *American Statistical Association Journal,* **64** 111-130, 1969.

Tiver, F., *Unpublished data*, Pers.Comm. 2002.

Tiver, F. and M. H. Andrew, Relative effects of herbivory by sheep, rabbits, goats and kangaroos on recruitment and regeneration of shrubs and trees in eastern South Australia, *Journal of Applied Ecology,* **34** 903-914, 1997.

Todd, C. R. and M. P. Ng, Generating unbiased correlated random survival rates for stochastic population models, *Ecological Modelling,* **144** 1-11, 2001.

Tuljapurka, S., *Population Dynamics in Variable Environments*, Springer-Verlag, Berlin, 1990.

Tyre, A. J., B. Tenhumberg, M. A. McCarthy and H. P. Possingham, Swapping space for time and unfair tests of ecological models, *Austral Ecology,* **25** 327-331, 2000.