

INCORPORATING LANDSCAPE STOCHASTICITY INTO POPULATION VIABILITY ANALYSIS

RYAN A. CHISHOLM^{1,2,3} AND BRENDAN A. WINTLE¹

¹*School of Botany, The University of Melbourne, Victoria 3010, Australia*

²*Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08540 USA*

Abstract. The importance of incorporating landscape dynamics into population viability analysis (PVA) has previously been acknowledged, but the need to repeat the landscape generation process to encapsulate landscape stochasticity in model outputs has largely been overlooked. Reasons for this are that (1) there is presently no means for quantifying the relative effects of landscape stochasticity and population stochasticity on model outputs, and therefore no means for determining how to allocate simulation time optimally between the two; and (2) the process of generating multiple landscapes to incorporate landscape stochasticity is tedious and user-intensive with current PVA software. Here we demonstrate that landscape stochasticity can be an important source of variance in model outputs. We solve the technical problems with incorporating landscape stochasticity by deriving a formula that gives the optimal ratio of population simulations to landscape simulations for a given model, and by providing a computer program that incorporates the formula and automates multiple landscape generation in a dynamic landscape metapopulation (DLMP) model. Using a case study of a bird population, we produce estimates of DLMP model output parameters that are up to four times more precise than those estimated from a single landscape in the same amount of total simulation time. We use the DLMP modeling software RAMAS Landscape to run the landscape and metapopulation models, though our method is general and could be applied to any PVA platform. The results of this study should motivate DLMP modelers to consider landscape stochasticity in their analyses.

Key words: Brown Creeper; DLMP; landscape; population viability analysis, PVA; RAMAS.

INTRODUCTION

Population viability analysis (PVA) is a decision-making tool used widely in conservation management (e.g., Boyce 1992, Akçakaya and Sjørgen-Gulve 2000, Menges 2000, Beissinger and McCullough 2002). The results of PVA are used to assess the vulnerability of species, rank management options, and guide research and data collection (Possingham et al. 1993, Akçakaya et al. 1995, Lindenmayer and Possingham 1996, Akçakaya and Sjørgen-Gulve 2000). The widespread application of PVA in conservation planning requires that PVA methods be developed and refined to ensure that they have maximum relevance and utility in decision making. Most PVA applications utilize dynamic population models and static models of the environment and habitat, though the need to incorporate dynamic habitats in PVA has long been acknowledged (Burgman et al. 1993, Holt et al. 1995). The importance of incorporating environmental and landscape change in PVA is particularly acute when recognizing the impact that dynamic processes such as climate change and habitat fragmentation are likely to have on biodiversity

(Thomas et al. 2004). Recent developments in PVA technology (Akçakaya et al. 2004) enable the incorporation of landscape and habitat dynamics and these developments show great potential for broader application in conservation planning (Wintle et al. 2005).

Predictions made by population viability models including metapopulation and landscape dynamics models (dynamic landscape metapopulation models, DLMP; Wintle et al. 2005) are affected by two sources of variation: landscape stochasticity and population stochasticity. Variability in parameter estimates such as expected minimum population size (EMP; McCarthy and Thompson 2001) made from model outputs thus depends on the number of different landscapes considered (landscape model repetitions), the number of population repetitions performed on each landscape (population model repetitions), the inherent variability in model outcomes associated with the variable landscape and the inherent variability associated with population factors.

With current DLMP software, landscape model repetitions must be performed manually and are computationally intensive, while population model repetitions can be automated and are generally much faster. Because of this there is a temptation to use a very high ratio of population repetitions to landscape repetitions (e.g., 1000:1), or even to consider only one

Manuscript received 10 October 2005; revised 22 June 2006; accepted 30 June 2006; final version received 1 August 2006.
Corresponding Editor: J. M. Marzluff.

³E-mail: chisholm@princeton.edu



PLATE 1. The Brown Creeper (*Certhia americana*), subject of the population viability analysis in this study, is a small North American forest bird that forages for insects on tree trunks. Photo credit: Nathan Gregory.

realization of the dynamic landscape. In such cases, modelers may be underestimating variance in their parameter estimates because they are excluding landscape stochasticity. Both Akçakaya et al. (2004), in a PVA of the Sharp-tailed Grouse (*Tympanuchus phasianellus*), and Larson et al. (2004), in a PVA of the Ovenbird (*Seiurus aurocapillus*), conducted demographic simulations on dynamic landscapes. Both studies considered only one landscape per scenario, although Larson et al. (2004) conducted a preliminary analysis that suggested the impact of landscape stochasticity in their model was sufficiently small to be ignored. Wintle et al. (2005), on the other hand, found that the variation in model outputs arising from landscape model stochasticity in a PVA for the Brown Creeper (*Certhia americana*; see Plate 1) was at least as large as variation due to alternative management scenarios. They attempted to encapsulate landscape and population stochasticity by considering 50 landscapes for each scenario with three population simulations on each landscape; however, they were unable to determine whether this combination of landscape and population simulations minimized the variance of their EMP estimates.

Given that landscape stochasticity is an important contributor to the variance of model outputs in at least some PVAs on stochastic dynamic landscapes, modelers are confronted with the question of how to apportion simulation time between generating different dynamic landscapes and running multiple demographic simulations on these landscapes. The goal is to find an optimal solution to this problem that will minimize the overall variance of parameter estimates such as EMP. Minimizing the variance of parameter estimates will improve predictions about the effects of management actions (or competing management options) on sensitive species.

METHODS

Throughout this manuscript and the appendices, we use the following formula for sample variance:

$$s^2 = \frac{1}{n} \sum (y_i - \bar{y})^2$$

where n is the number of observations, y_i are the observations and \bar{y} is the mean observation; an unbiased estimator of the population variance is thus $[n/(n-1)]s^2$.

Optimal population repetitions

Expected minimum population size (EMP) is a metric for assessing threat in populations of organisms (McCarthy and Thompson 2001). EMP is estimated as the average of minimum population sizes across multiple runs of a population model. In DLMP models, variation in minimum population size (MP) arises from two sources: landscape stochasticity and population stochasticity. Here, we make the assumption that the random component of MP associated with landscape stochasticity (X_L) is a normal random variable:

$$X_L \sim \mathcal{N}(\mu_L, \sigma_L^2)$$

where μ_L is the mean and σ_L^2 is the variance of the distribution. We also assume that the random component of MP associated with population stochasticity (X_P) is a normal random variable with zero mean and constant variance σ_P^2 independent of the particular landscape realization on which the simulation takes place. The independence assumption will also be referred to as the assumption of homogeneity of variance (i.e., homogeneity of variance across landscapes). So

$$X_P \sim \mathcal{N}(0, \sigma_P^2).$$

From the assumptions that X_L and X_P are independent and normally distributed, it follows that the random variable representing MP is also a normal random variable:

$$X = X_L + X_P \sim \mathcal{N}(\mu_L, \sigma_L^2 + \sigma_P^2) = \mathcal{N}(\mu_L, \sigma_T^2)$$

where the total variance of the random variable X can be expressed as $\sigma_T^2 = \sigma_L^2 + \sigma_P^2$.

Suppose that to estimate $EMP = E(X)$ we perform n_L landscape repetitions and n_P population repetitions on each landscape. Because we assume equal population variance σ_P^2 on each landscape, and we perform the same number of population repetitions n_P on each landscape, an unbiased estimator of $E(X)$ weights each observed MP equally, such that

$$\begin{aligned} \bar{x} &= \frac{1}{n_L n_P} \sum_{i=1}^{n_L} \sum_{j=1}^{n_P} x_{ij} = \frac{1}{n_L n_P} \sum_{i=1}^{n_L} \sum_{j=1}^{n_P} (x_{L,i} + x_{P,ij}) \\ &= \frac{1}{n_L} \sum_{i=1}^{n_L} x_{L,i} + \frac{1}{n_L n_P} \sum_{i=1}^{n_L} \sum_{j=1}^{n_P} x_{P,ij} \end{aligned}$$

where x_{ij} is the MP from the j th population simulation on the i th landscape, $x_{L,i}$ is the underlying mean MP on the i th landscape (a realization of X_L), and $x_{P,ij} = x_{ij} - x_{L,i}$ (a realization of X_P).

The variance of this estimate is

$$\text{var}(\bar{x}) = \text{var}\left(\frac{1}{n_L} \sum_{i=1}^{n_L} x_{L,i} + \frac{1}{n_L n_P} \sum_{i=1}^{n_L} \sum_{j=1}^{n_P} x_{P,ij}\right)$$

$$\begin{aligned} &= \frac{1}{n_L^2} \text{var}\left(\sum_{i=1}^{n_L} x_{L,i}\right) + \frac{1}{n_L^2 n_P^2} \text{var}\left(\sum_{i=1}^{n_L} \sum_{j=1}^{n_P} x_{P,ij}\right) \\ &= \frac{1}{n_L} \left(\sigma_L^2 + \frac{\sigma_P^2}{n_P}\right). \end{aligned} \tag{1}$$

It is clear from Eq. 1 that ignoring landscape variance (i.e., assuming $\sigma_L^2 = 0$), as many studies have done, produces an underestimate of the variance in EMP. Our goal is to produce the best estimate of EMP in a given period of time, t . Suppose that to make an observation on X_L takes t_L seconds, and to make an observation on X_P takes t_P seconds. Then, if we are performing n_P population repetitions per landscape, we must set n_L such that

$$t_L n_L + t_P n_L n_P = t.$$

On rearranging and substituting into Eq. 1 for n_L , we obtain

$$\text{var}(\bar{x}) = \frac{t_L + t_P n_P}{t} \left(\sigma_L^2 + \frac{\sigma_P^2}{n_P}\right). \tag{2}$$

We seek to minimize this quantity, which is the variance of our EMP estimate, by choosing an optimal value of n_P , the number of population repetitions per landscape. To do this, we take the derivative of Eq. 2 with respect to n_P and equate to zero

$$\frac{d\text{var}(\bar{x})}{dn_P} = \frac{t_P}{t} \left(\sigma_L^2 + \frac{\sigma_P^2}{n_P}\right) + \frac{t_L + t_P n_P}{t} \left(-\frac{\sigma_P^2}{n_P^2}\right) = 0.$$

Simplifying and rearranging gives

$$n_P = \frac{\sigma_P}{\sigma_L} \sqrt{\frac{t_L}{t_P}}. \tag{3}$$

This is the number of population repetitions per landscape that minimizes the expected variance of our estimate of EMP. Note that the absence of t from this equation means that the optimal value of n_P is independent of the number of landscape simulations (n_L) and therefore the total duration of the simulation set.

The REPEATER program

We developed a C++ computer program that puts the optimization principles described in the previous section into practice. We chose to work with the RAMAS Landscape software (Akçakaya et al. 2004), but could just have easily applied the same principles to VORTEX (Lacy 1993), ALEX (Possingham and Davies 1995), or any other DLMP software. The RAMAS landscape software integrates the population modeling software RAMAS GIS (Akçakaya 2002) and the forest dynamics modeling software LANDIS (Mladenoff et al. 1996, Mladenoff and He 1999).

The REPEATER program automates the execution of multiple runs of the landscape and population

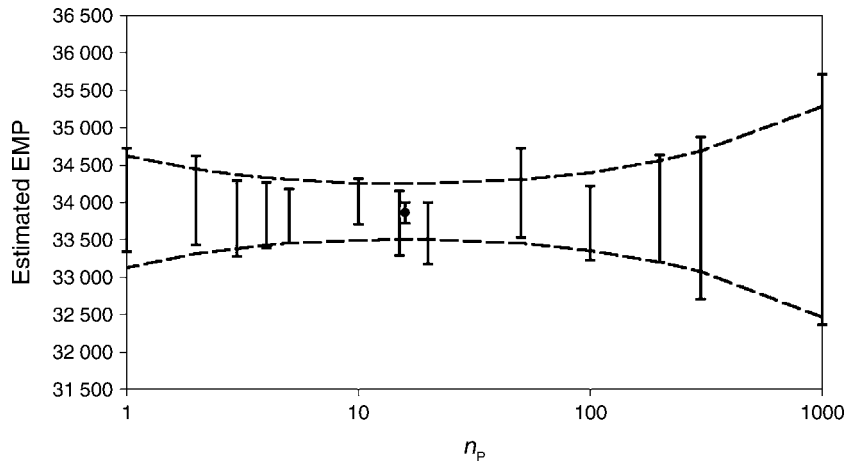


FIG. 1. Estimated values of expected minimum population (EMP) in the Brown Creeper model resulting from two-hour simulations. The x -axis represents number of population repetitions used per landscape (n_p); the y -axis represents EMP. Each set of error bars represents a 68% confidence interval (mean \pm SD) for EMP values generated from two-hour simulations ($t = 7200$ s) using the relevant value of n_p , based on the results of 15 independent two-hour simulations. Note that the right-most error bars ($n_p = 1000$) correspond to a situation in which only one landscape is considered per two-hour simulation; in this case, landscape stochasticity is effectively ignored. The dashed upper and lower curves show the theoretical standard deviations of EMP estimates calculated from Eq. 2. A separate 16-hour simulation with the optimal-iterative- n_p option enabled was used to produce estimates of the parameters in Eq. 2 ($\sigma_L = 1454.8$, $\sigma_P = 4006.5$, $t_L = 215.8$ s, $t_P = 6.5$ s). The solid circle with error bars shows estimated optimal n_p and estimated EMP (\pm SE) from the 16-hour simulation.

models, a process that previously required continual user interaction. The total simulation time (t) and the number of population repetitions per landscape (n_p) are configurable within the REPEATER program. The user can also specify that n_p should be estimated using Eq. 3. It is intended that use of this option will lead to the minimum-variance estimate of EMP for a given simulation time. Data on carrying capacity and population trajectory are collated to produce estimates of the final population size and graphs of carrying capacity, population trajectory and interval extinction risk. Details of software implementation and the mathematics of simulation time and variance estimation are included in Appendices A and B.

The Brown Creeper case study

To the extent that a particular model violates the assumptions of normality and homogeneity of variance mentioned in the previous section, Eq. 3 and our implementation of it in REPEATER may not produce an optimal solution to the problem of estimating EMP. Furthermore, the mathematical properties of expectation and unbiasedness do not guarantee that the estimate of n_p given by Eq. 3 will be unbiased, even though the estimates of σ_L^2 and σ_P^2 are unbiased.

To test whether, in practice, REPEATER produces close to optimal output, we used a DLMP model of the Brown Creeper (*Certhia americana*) in the boreal forests of northern Ontario (Wintle et al. 2005). First, we ran REPEATER for 16 hours with the optimal-iterative n_p option enabled. This produced an estimate of the optimal number of population repetitions per landscape, as well as estimates of the variances σ_L^2 and σ_P^2 . We then

conducted 13 sets of 15 two-hour simulations, each set having a different fixed value of n_p . The n_p values tested were 1, 2, 3, 4, 5, 10, 15, 20, 50, 100, 200, 300, and 1000. The resulting distribution of estimated EMPs was used to determine whether the optimal value of n_p previously determined was indeed leading to a minimum-variance estimate of EMP.

RESULTS

For the Brown Creeper model, the assumption of normality of the distribution of MP values was upheld for most landscapes (Kolmogorov-Smirnov Tests on 10 landscapes: $P = 0.10, 0.63, 0.69, 0.58, 0.10, 0.56, 0.95, 0.00, 0.29, 0.55$). The homogeneity of MP variance test for the same 10 landscapes failed (Levene tests: $P < 0.0001$), although this was largely attributable to the outlying landscape that failed the normality of variance test.

The results from the optimal-iterative- n_p simulation were consistent with the results from the fixed- n_p simulations (Fig. 1). The optimal-iterative approach suggested an optimal n_p value of about 15, and the fixed- n_p simulations suggested that any value of n_p in the range 3–100 produces close to optimal results. The observed variance of estimates was consistent with Eq. 1 across the spectrum of n_p values.

In the Brown Creeper model, most of the variance in EMP (85%) is associated with population stochasticity, but there are substantial gains to be made in the precision of EMP estimates by optimizing the value of n_p and thereby accounting also for the 15% of variance that is associated with landscape stochasticity (Fig. 1).

The $n_p = 1000$ simulations (right-most error bars in Fig. 1), for which the standard error of the EMP estimate was 1767.7, represent the extreme case of not incorporating landscape stochasticity at all, because in this case a single landscape run took the full two hours of allocated time for each simulation; this is to be compared with the optimal value of $n_p = 15$ for which $n_L = 21$ landscapes were generated within the two hours and the standard error of EMP estimates was only 427.8 (Fig. 1). Thus, incorporating landscape stochasticity and using the optimal procedure for estimating the number of population repetitions reduced the standard error of our estimate by a factor of four.

DISCUSSION

Incorporating both landscape and population stochasticity into models of population viability, such as facilitated by our REPEATER program, can enable modelers to provide conservation managers with more precise estimates of population metrics. In our example, we reduced the variance associated with an estimate of expected minimum population size (EMP) by a factor of four with minimal increases in simulation effort. While we focused on EMP, the importance of accounting for landscape stochasticity is equally applicable to end population size, mean population size or any other metric, provided that the requisite assumptions are satisfied. More precise estimates aid management decisions by reducing type II errors, and consequently, decrease instances in which managers fail to act to prevent population declines because of a lack of statistical evidence.

Our estimators of EMP (or other metric) are always unbiased; the extent to which the variance in these estimates is minimized depends on whether the assumptions of normality and homogeneity of variance are upheld. For our case study of the Brown Creeper, the empirical results were consistent with the theoretical (Fig. 1), even though the homogeneity of variance assumption was violated. This suggests that our approach is robust to violation of the underlying assumptions, but further research in this area is required. In some cases, violation of the normality principle may be redressed by an appropriate transformation on the model output parameter. For instance, if the model output parameter is log-normally distributed one can instead use the logarithm of the parameter as the model output and the assumption of normality will be satisfied. The normality assumption is particularly likely to be violated for species with a high likelihood of extinction, i.e., species whose population metrics are zero inflated. Violation of the assumption of homogeneity of variance corresponds to cases where population variance is different on different landscapes. It seems intuitive to us that, under conditions of heterogeneity of variance, Eq. 3 would underestimate the optimal number population repetitions per landscapes and Eq.

2 would underestimate the variance in EMP estimates, but this requires verification.

Our results demonstrate that, in our case study at least, one only need estimate n_p to about the nearest order of magnitude in order to produce close-to-optimal conditions (Fig. 1). This is further evidence that our method is robust to violation of the underlying assumptions. Moreover, we reiterate that, even if the assumptions of our approach are violated and our method becomes suboptimal, the estimates of population metrics remain unbiased, and that incorporating landscape stochasticity in a suboptimal manner is generally better than ignoring it completely (again, see Fig. 1), as most studies to date have done.

It should be emphasized that the optimal value of n_p will differ from model to model, and will vary with changes to certain parameters within the same model (e.g., in a sensitivity analysis). For instance, changes in the survival or fecundity rate of a species could be expected to affect the variance in EMP associated with population stochasticity (σ_p^2), which would in turn affect the optimal value of n_p . For this reason, one should not assume that a particular optimal value of n_p will be applicable beyond the direct context in which it was estimated.

Modelers may be concerned with the problem of “How long a simulation is long enough?” or “How many landscape replications are required?” In other words, are there substantial gains in precision (reductions in variance of estimates) to be made by running a model for two weeks as opposed to two hours? In accordance with Eq. 2, the variances of parameter estimates are inversely proportional to simulation duration, and therefore the standard errors are inversely proportional to the square root of simulation duration. In general, one can expect a four-fold increase in simulation time to be associated with a two-fold decrease in standard errors. The standard errors can be made arbitrarily low (the parameter estimates can be arbitrarily precise) by running the simulation for long enough. The REPEATER program provides flexibility to the user in the form of multiple stopping rules for simulations. These are (1) number of landscape repetitions to perform; (2) maximum total duration of simulation; and (3) required coefficient of variation of EMP estimate. So the user can specify, for instance, that the simulation should cease when the coefficient of variation of the EMP estimate falls below 10%, or when the simulation duration exceeds 24 hours, whichever happens first. The REPEATER program also provides estimates of the remaining simulation time required until the coefficient of variation would fall below the specified threshold.

ACKNOWLEDGMENTS

We are grateful to Lisa Venier and Jennie Pearce, (Canadian Forest Service), Ascelin Gordon and Sarah Bekessy (RMIT University), and Mark Burgman (Australian Centre of Excellence for Risk Analysis: ACERA) for discussion and ideas. We

thank three anonymous reviewers for helpful comments. Wintle was supported by Australian Research Council grant LP0347473 and ACERA. R. Chisholm is supported by the May Fellowship at Princeton University.

LITERATURE CITED

- Akçakaya, H. R. 2002. RAMAS GIS: linking landscape data with population viability analysis. Version 4.0t. Applied Biomathematics, Setauket, New York, USA.
- Akçakaya, H. R., M. A. McCarthy, and J. L. Pearce. 1995. Linking landscape data with population viability analysis—management options for the Helmeted Honeyeater *Lichenostomus melanops cassidix*. *Biological Conservation* 73:169–176.
- Akçakaya, H. R., V. C. Radeloff, D. J. Mladenoff, and H. S. He. 2004. Integrating landscape and metapopulation modeling approaches: viability of the Sharp-tailed Grouse in a dynamic landscape. *Conservation Biology* 18:526–537.
- Akçakaya, H. R., and P. Sjörgen-Gulve. 2000. Population viability analyses in conservation planning: an overview. *Ecological Bulletins* 48:9–21.
- Beissinger, S. R., and D. R. McCullough. 2002. Population viability analysis. University of Chicago Press, Chicago, Illinois, USA.
- Boyce, M. S. 1992. Population viability analysis. *Annual Review of Ecology and Systematics* 23:481–506.
- Burgman, M. A., S. Ferson, and H. R. Akçakaya. 1993. Risk assessment in conservation biology. Chapman and Hall, London, UK.
- Holt, R. D., S. W. Pacala, T. W. Smith, and J. Liu. 1995. Linking contemporary vegetation models with spatially explicit animal population models. *Ecological Applications* 5:20–27.
- Lacy, R. C. 1993. VORTEX: A computer simulation model for population viability analysis. *Wildlife Research* 20:45–65.
- Larson, M. A., F. R. Thompson III, J. J. Millsbaugh, W. D. Dijak, and S. R. Shifley. 2004. Linking population viability, habitat suitability, and landscape simulation models for conservation planning. *Ecological Modelling* 180:103–118.
- Lindenmayer, D. B., and H. P. Possingham. 1996. Ranking conservation and timber management options for Leadbeaters Possum in Southeastern Australia using population viability analysis. *Conservation Biology* 10:235–251.
- McCarthy, M. A., and C. Thompson. 2001. Expected minimum population size as a measure of threat. *Animal Conservation* 4:341–355.
- Menges, E. S. 2000. Population viability analyses in plants: challenges and opportunities. *Trends in Ecology and Evolution* 15:51–56.
- Mladenoff, D. J., and H. S. He. 1999. Design and behaviour of LANDIS, an object-oriented model of forest landscape disturbance and succession. Pages 125–162 in W. L. Baker, editor. *Spatial modeling of forest landscape change: approaches and applications*. Cambridge University Press, Cambridge, UK.
- Mladenoff, D. J., G. E. Host, J. Boeder, and T. R. Crow. 1996. LANDIS: a spatial simulation model of forest landscape disturbance, management and succession. Pages 175–179 in Second international conference on integrating GIS and environmental modeling. National Center for Geographic Information and Analysis, Santa Barbara, California, USA.
- Possingham, H. P., and I. Davies. 1995. ALEX—A model for the viability analysis of spatially structured populations. *Biological Conservation* 73:143–150.
- Possingham, H. P., D. B. Lindenmayer, and T. W. Norton. 1993. A framework for the improved management of threatened species based on population viability analysis. *Pacific Conservation Biology* 1:39–45.
- Thomas, C. D., A. Cameron, and R. A. Green. 2004. Extinction risk from climate change. *Nature* 427:145–148.
- Wintle, B. A., S. A. Bekessy, L. A. Venier, J. L. Pearce, and R. A. Chisholm. 2005. Utility of dynamic landscape metapopulation models for sustainable forest management. *Conservation Biology* 19:1930–1943.

APPENDIX A

A description of the REPEATER software (*Ecological Archives* A017-013-A1).

APPENDIX B

Unbiased estimates of variance from data sets with unequal underlying group means and unequal group sizes (*Ecological Archives* A017-013-A2).

SUPPLEMENT

The REPEATER software: manual, executable and source code (*Ecological Archives* A017-013-S1).