



How sampling affects estimates of demographic parameters

Mark R. Lesser & Simon Brewer

Keywords

Logistic regression; *Pinus ponderosa*;
Population growth; Sampling error and bias;
Sampling method; Simulation modelling

Received 29 November 2011

Accepted 26 March 2012

Co-ordinating Editor: Francisco Pugnaire

Lesser, M.R. (corresponding author,
mrlesser@syr.edu): Program in Ecology,
Department of Botany, University of Wyoming,
1000 E. University Ave., Laramie, WY, USA,
82071

Brewer, S. (simon.brewer@geog.utah.edu):
Department of Geography, University of Utah,
260 S. Central Campus Dr., Salt Lake City, UT,
USA, 84112

Lesser, M.R.: Current address: Department of
Biology, Syracuse University, Life Sciences
Complex, Syracuse NY, 13244

Abstract

Question: Demographic rates are often modelled using small data sets over short time frames. Here, we use fully sampled populations as a basis for testing how the intensity of two different sampling approaches (individual random-tree and n -tree distance plots) can affect estimates of growth parameters and the timing of population development. How do sampling method and intensity affect estimates of early stages of population growth?

Location: North-central Wyoming, USA.

Methods: We used a data set in which every individual in each of four discrete ponderosa pine populations was mapped and aged. We calculated cumulative population growth and fitted it to a logistic regression model. Based on this model, we estimated population growth rate, first colonization, timing of population growth initiation, maximum growth rate and growth saturation. We conducted simulations for two sampling methods. First, individual trees were chosen at random, with different percentages of the full population being chosen. Second, we simulated n -tree distance plot sampling, where we changed the number of plots that were laid in each population. For each method and at each intensity, 10 000 simulation runs were performed. The simulation results were fitted to a logistic regression model. We then looked at the difference between the full and partially sampled population results to examine how lowering sampling intensity affected the results.

Results: Population growth rate was not significantly affected by sampling intensity except at low levels of sampling. However, first colonization and timing of population initiation were affected by sampling intensity. For both parameters, the individual random-tree method produced more accurate results than the n -distance method as sampling intensity decreased.

Conclusions: Accurate estimation of population growth parameters is critical for both ecological understanding and resource management. Results are encouraging in that they indicate that moderate levels of sampling will reliably estimate population growth parameters. However, our results are specific to ponderosa pine and may not apply to other species with different life-history characteristics. Our results also highlight the fact that population structure can play a major role in sampling accuracy and needs to be considered in choosing the appropriate method and intensity.

Introduction

Understanding of demographic rates and the development of new populations are essential elements of population ecology and conservation biology (Crawley 1990; Morris & Doak 2002; Sibly & Hone 2002). However, the data required to fully understand these rates and processes can

be difficult to obtain, especially for many plant species where large population sizes and long generation times put them beyond the scope of traditional observational studies (Crawley 1990). Despite these difficulties, little attention has been paid to the influence of sampling methods and intensity on estimation of demographic parameters (Doak et al. 2005).

For many tree species, long generation times and large population sizes are more of a problem than with other plants. Traditional observational studies are incapable of capturing population dynamics as they unfold over the course of multiple centuries. Dendroecology makes it possible to examine population growth over long time periods with high temporal and spatial precision (Fritts & Swetnam 1989). However, while dendroecology can overcome temporal issues, scale may still pose difficulties. Large populations can make measuring every tree difficult, if not impossible. Therefore, various sampling techniques are typically employed to estimate the population parameters. However, the effects of sampling on estimates of population growth parameters are unclear (Meyer et al. 1986). We distinguish between a *population* growth curve, based on all individuals of a population, and the *sample* growth curve, based on a sample of the population.

It is important to consider how different sampling techniques will affect modelling and parameter estimation. Ideally, in any sampling strategy to estimate population structure, an unbiased random sample of individuals should be used. However, this approach is usually not feasible because the position of every individual in the population needs to be known before sampling. Sampling techniques employed in dendroecological studies include transects (e.g. Norman & Taylor 2005), plot-sampling methods (e.g. League & Veblen 2006; Kaye et al. 2010) and non-random sampling based on size (e.g. White 1985), fire scars (e.g. Veblen et al. 2000) or evidence of old age (Gray et al. 2004; Biondi et al. 2011). Many dendroecological-based studies have used *n*-tree distance plot sampling methods (e.g. Brown & Wu 2005; Brown 2006; Heyerdahl et al. 2006), whereby the number of trees measured is held constant and plot size is adjusted based on the density of the population (Jonsson et al. 1992; Lessard et al. 2002). This method is preferred over fixed-radius plot sampling when population density is uneven (Lessard et al. 2002).

Another consideration in modelling population growth is that a crucial period in understanding population development is during the early stages of colonization and expansion (Stacey & Taper 1992). During initial stages of colonization, populations may experience extended lags in growth, or conversely grow more rapidly than expected due to stochastic events. This is a particularly difficult stage of population development to study using dendroecology because a population must not only contain the initial colonizing individuals, but those individuals need to be represented in the sample. The proportion of colonizing individuals is likely to be very small for a large, well-established population.

Here, we use data from four ponderosa pine (*Pinus ponderosa* Dougl. Ex. Laws) populations, where the entire

history of population growth, from the initial colonization event to the present, is known (Lesser & Jackson 2012). Because these data represent the entire population, except for a small percentage that was unable to be aged because of heart rot (<15% of combined populations), the population growth curve can be estimated. We used these data to test for error and bias in estimates of population growth parameters. In particular, we examined how sampling affects estimation of timing parameters at different stages of population development. We used simulation models to test how two different sampling strategies, individual random-tree and *n*-tree distance plots, and sampling intensity affected these parameters compared to the known full populations.

Methods

Study populations

We used four isolated, colonizing populations of ponderosa pine in the Bighorn Basin in north-central Wyoming (Table 1). These populations were isolated from other ponderosa pine populations by >15 km and had easily distinguishable boundaries beyond which no ponderosa pine could be found. Nearly every tree in each of these populations was mapped and aged (Lesser & Jackson 2012), creating a unique data set where the population growth curves can be estimated.

Population density at all sites was low, with average densities >10 trees·ha⁻¹ (Fig. 1), but highly uneven. The number of aged trees in each population ranged from 71 at Grass Creek to 835 at Cottonwood Creek (Table 1). Initial colonization of sites occurred between the 1530s and 1650s (Appendix S1). Lesser & Jackson (2012) provide additional site information.

Sampling designs and simulation modelling

To investigate the effect of sampling, we simulated two sampling approaches. First, we simulated random sampling of individual trees. The number of trees sampled was chosen to represent 90% down to 10% of the full population in 10% step increments. For each percentage incre-

Table 1. Study site locations, elevations, area of population and number of trees included in the analysis.

Population	Latitude	Longitude	Elevation (m)	Area (ha)	No. of trees
Castle garden	43.96N	107.52W	1450–1550	94	144
Grass creek	43.88N	108.63W	1700–1900	136	71
cottonwood Creek	43.80N	108.70W	1700–1850	173	835
Anchor dam	43.67N	108.83W	1950–2050	10	128

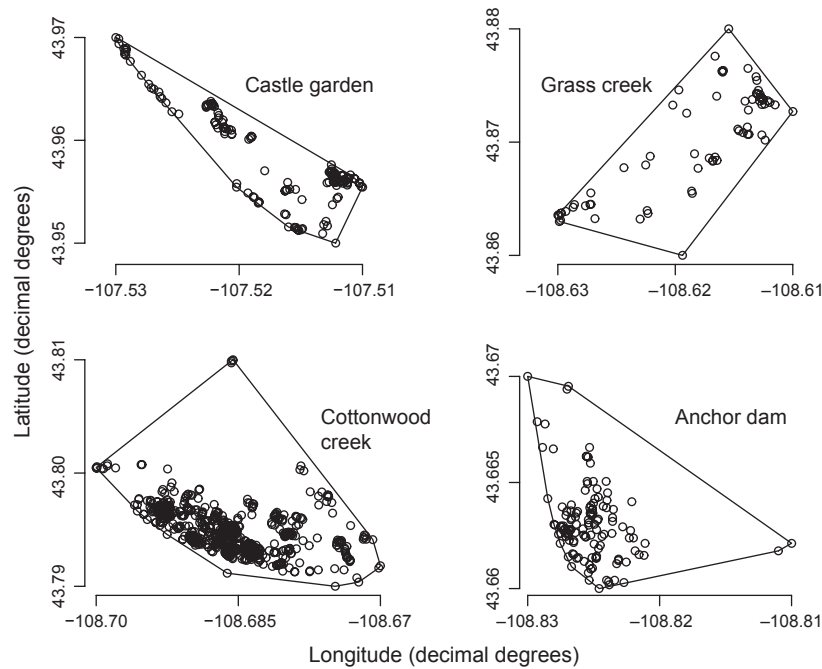


Fig. 1. Maps of the four study populations showing tree locations (open circles) and the bounding box used to contain the *n*-distance sampling plot placement.

ment of the full population we ran the simulation 10 000 times.

The second sampling technique simulated was an *n*-tree distance plot sampling method, which we chose as a representative of plot-based sampling methods. In the *n*-tree distance approach a plot ‘centre’ is randomly placed within the population. We defined the limits of the population with a convex-hull bounding box placed around the outermost trees (Fig. 1). From the plot centre, every tree occurring within a pre-set distance was sampled. We varied the number of plots from one to ten in each population. For each number of plots, 10 000 simulations were run with plots randomly placed within each population. Due to the different densities and overall sizes of the populations, specific parameters on minimum and maximum plot size and the minimum number of trees to be included were adjusted on a site-by-site basis (Table 2).

Our goal was to examine how sampling intensity affected estimates of growth rate and timing of population development for each method. To examine growth rate we

calculated cumulative population size through time from the output of each run. Cumulative population size was calculated as the ratio of trees establishing in a decade to the number of reproductively mature trees present in the population at the beginning of that decade (Lesser & Jackson 2012; Morris & Doak 2002). We then fitted a logistic regression model to cumulative growth. This model was used based on Lesser & Jackson (2012), which showed that this model was a good fit for the populations. We calculated the median slope and intercept and the lower 5% quantile of the slope from the 10 000 runs performed for each sampling intensity under both tested methods. The slope of the regression model was interpreted as the population growth rate and is comparable to the value *r*, or the rate of natural increase (Sibly & Hone 2002). We compared the simulated model median results and the full population model results using a z-score to determine if there was a significant difference in the shape of the model. Z-scores were calculated as: Eq. (1)

$$z = \frac{\beta_1 - \beta_2}{\sqrt{SE\beta_1^2 + SE\beta_2^2}} \tag{1}$$

where β_1 is the population slope, β_2 is the sample population slope, and SE is the standard error (Clogg et al. 1995). Z-scores were also used to compare the 5% quantile results from the simulations to the full population model, to assess the worst-case scenario for that sampling intensity.

Table 2. Site specific parameters for *n*-distance sampling design.

Population	Max. no. trees in plot	Min. no. trees in plot	Max. plot radius (m)
Castle garden	20	5	120
Grass creek	10	5	160
Cottonwood creek	30	5	160
Anchor dam	12	5	40

The decades of initial and second establishment were calculated from the simulations. Based on our calculation of cumulative growth, the second establishment event represents the initiation of population growth. We also calculated the time of growth saturation and the time of maximum growth rate. Saturation was calculated as the time when 95% of the current population was present and is comparable to k , or the carrying capacity of the population. The timing of maximum growth was calculated as half the time to reach the saturation point, where the slope of the logistic regression model is maximized.

For each of these four variables the difference between the full population and the sample population parameter was calculated. This was done for each of the 10 000 simulation runs for each sampling method and intensity. To assess both error and bias in the sampling we used box-plots to look at the distributions of each parameter. The deviation of the median simulated value from the full population value was used to assess error in the sampling. The variance in the distribution, or breadth of the box-plot, was used to assess bias, with a wider distribution indicating more bias in the parameter estimation.

Results

Growth rate

Observed population growth followed a logistic pattern in all four populations, with slow initial growth followed by an exponential increase phase, and finally a slowing and almost complete cessation of growth (Lesser & Jackson 2012; Fig. 2). Here we show the logistic regression models that were fitted to the data (Fig. 2). The slope of the fitted logistic regression for the full populations ranged from 0.015 to 0.039 (Appendix S1). For both sampling methods the slope estimates from the fitted logistic regression model become more variable as sampling intensity was decreased (Figs 2, 3).

For the random-tree sampling method z -scores between the slope of the sampled population and the slope of the full population exceeded one standard deviation (SD) only at very low sampling intensities (Appendix S1). As expected, the 5% quantile of the slope, when tested against the full population model, showed a higher degree of difference than the median slope. However, even these values did not become significant until $\leq 40\%$ of the population had been sampled at Castle Garden and Anchor Dam, and $\leq 30\%$ at Cottonwood Creek. Grass Creek showed even less of a difference, where with only 10% of the population sampled the z -score between the simulated population and the full data set was still < 1 SD (Appendix S1).

Results are similar for the n -tree distance-based method. Z -scores increased as fewer trees were sampled (i.e. less plots included). However, the values did not exceed 1 SD until less than two plots were included (Appendix S2).

Z -scores of the 5% quantile against the full population, however, exceeded 1 SD with as many as nine plots at Cottonwood Creek, and seven and five plots, respectively, at Anchor Dam and Castle Garden. Grass Creek showed the least amount of difference, where even at only one plot the z -score remained within 1 SD (Appendix S2).

Estimation of timing parameters

Initial colonization

With the random-tree sampling method, the putative initial establishment event was consistently included in the sample with $>60\%$ sampling (Fig. 4, Appendix S1). At Castle Garden, Cottonwood Creek and Anchor Dam, once the initial tree was missed the estimated decade of initial establishment moved ahead in time in a relatively slow, but systematic way. This is in accordance with the long period of low, but consistent, establishment in the actual data sets (Fig. 2). Accordingly, as sampling intensity increased, the median estimated value deviated less from the putative value, and the variance in the estimates also decreased (Fig. 4).

At Grass Creek, however, where there is a long gap in establishment between the first and second trees (Lesser & Jackson 2012), initial establishment shifted by 140 yrs if the first tree was missed (Fig. 4, Appendix S1). Furthermore, variance in the estimate of initial establishment was much greater at Grass Creek, compared to the other populations, especially at intermediate sampling intensities. At 30% to 70% sampling intensity, both the putative colonizing event and what is actually the second tree to establish are consistently sampled. However, the median estimate switches from the second tree to the putative first tree between 40% and 50% sampling (Fig. 4).

Identifying the assumed initial establishment event was much more variable with the n -tree distance plot method than with the random-tree sampling method (Fig. 5, Appendix S2). The initial establishment event at Cottonwood Creek was not consistently included even at the highest number of plots, and variance around the median was greater (Fig. 5). Differences in population growth initiation were less pronounced at the Castle Garden and Anchor Dam sites, but still showed offsets of over a century when sampling intensity was low (Fig. 5, Appendix S2). As with Cottonwood Creek, variance in estimates was also higher. Grass Creek showed very little difference in timing of initiation. The putative initial establishment was consistently included with three or more plots, and variance around that estimate fell to zero with four or more plots (Fig. 5).

Initiation of population growth

The establishment of the second tree, which represents the inferred initiation of population growth, was consistently

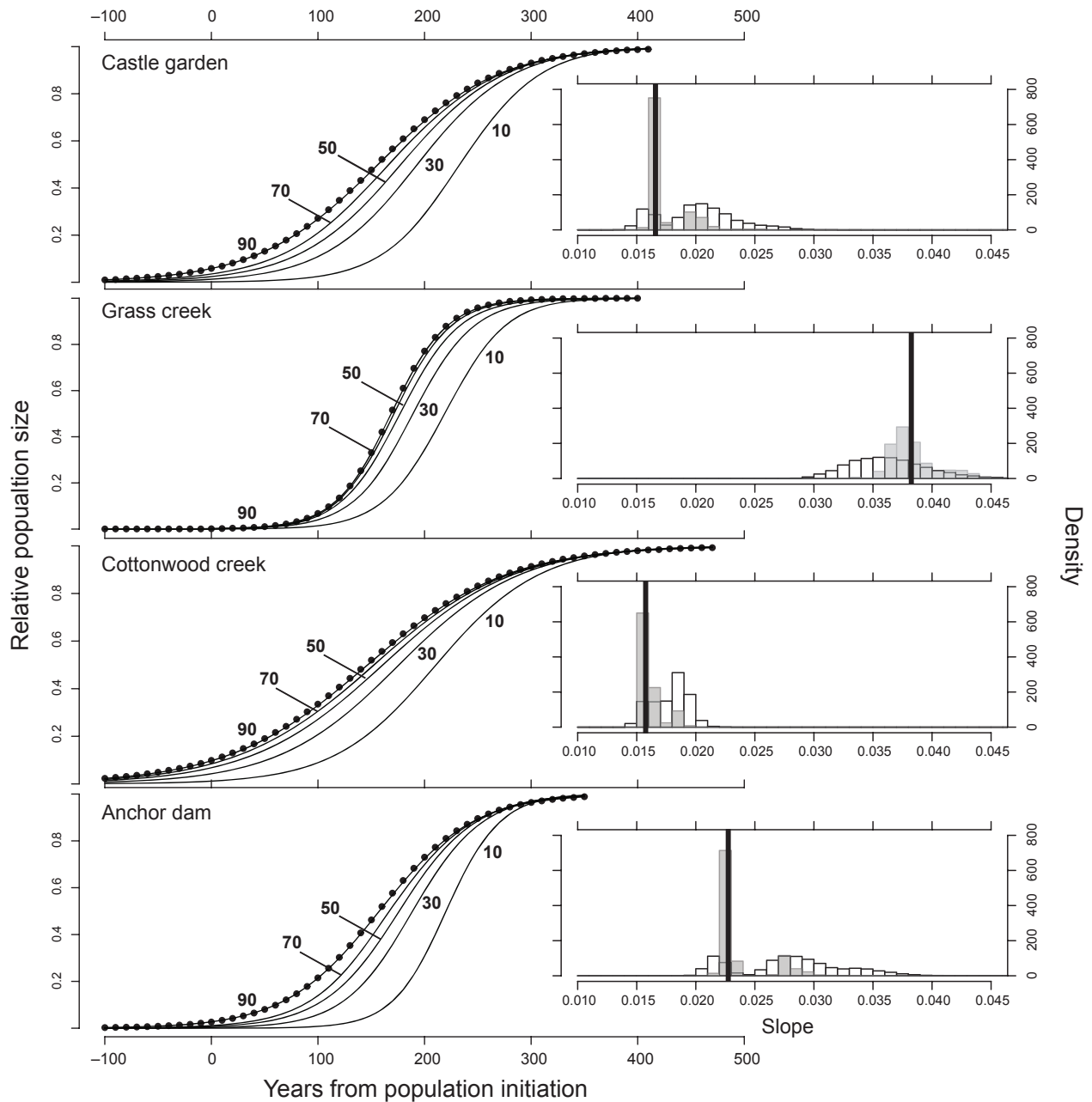


Fig. 2. Modelled population growth using the random individual tree sampling method. Fitted logistic regression lines for relative population size (calculated as a ratio of no. of establishment s/no. of reproductively mature individuals present in the population) against time for four ponderosa pine populations. Black circles indicate the fitted model values for the full population data set (100% sampling). Lines indicate the median sample growth curves for 90–10% of the population randomly sampled. The inset plots show the distribution of the slope of the regression model at 90% sampling (grey bars) and 50% sampling (open bars). The black line indicates the actual slope of the full population data set.

included in the random-tree sampling method when >70% of the population was sampled. At sampling intensities <70%, estimates of the second tree establishment began to shift (Appendix S1), causing shifts in estimated population growth initiation (Fig. 4). The difference between actual population initiation and the sampled

population initiation increased as sampling intensity decreased, and at low levels of sampling the difference is >100 yrs at Castle Garden, Cottonwood Creek and Anchor Dam (Fig. 4). Once again Grass Creek stands out, in this case having the smallest shift and the least amount of variance in the onset of estimated population growth. The

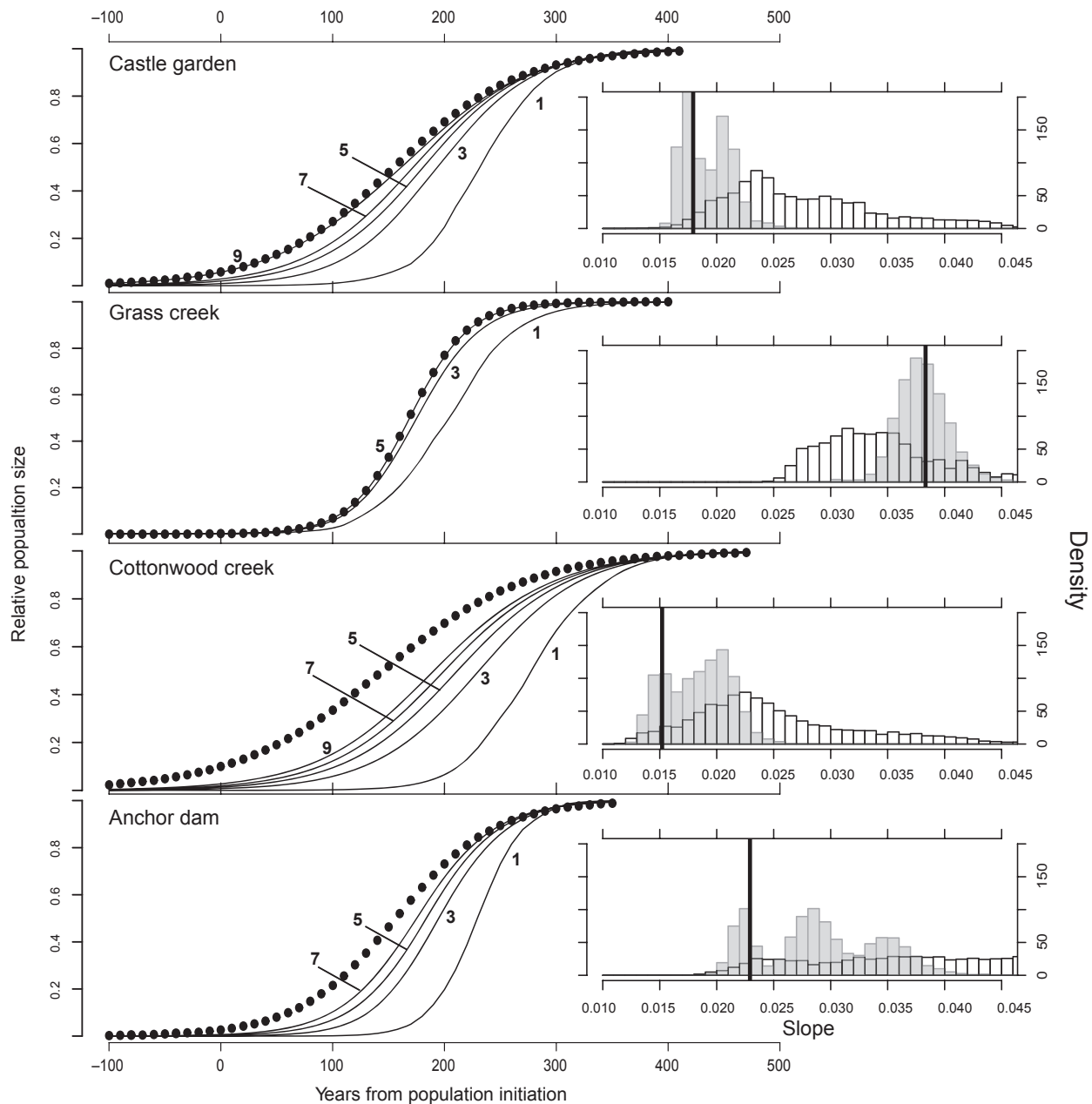


Fig. 3. Modelled population growth using the n -distance plot sampling method. Fitted logistic regression lines for relative population size (calculated as a ratio of no. of establishment s /no. of reproductively mature individuals present in the population) against time for four ponderosa pine populations. Black circles indicate the fitted model values for the full population data set (100% sampling). Lines indicate the median sample growth curves for 9–1 n -distance plots sampled. The inset plots show the distribution of the slope of the regression model at the max. number of plots sampled in each population (grey bars) and two plots sampled (open bars). The black line indicates the actual slope of the full population data set.

demographic pattern at Grass Creek is such that once the putative initial tree is missing, the next two trees are never temporally far apart and therefore there is little underestimation of population growth initiation (Fig. 4, Appendix S1). The other three populations, however, are much more strongly influenced by the separation between the first

and second trees, because more trees are missing from the sample (Fig. 2, Appendix S1).

Estimating the initiation of population growth was much more variable with the n -tree distance plot method (Fig. 5). Even high numbers of plots failed to consistently sample the putative second establishment event at Castle

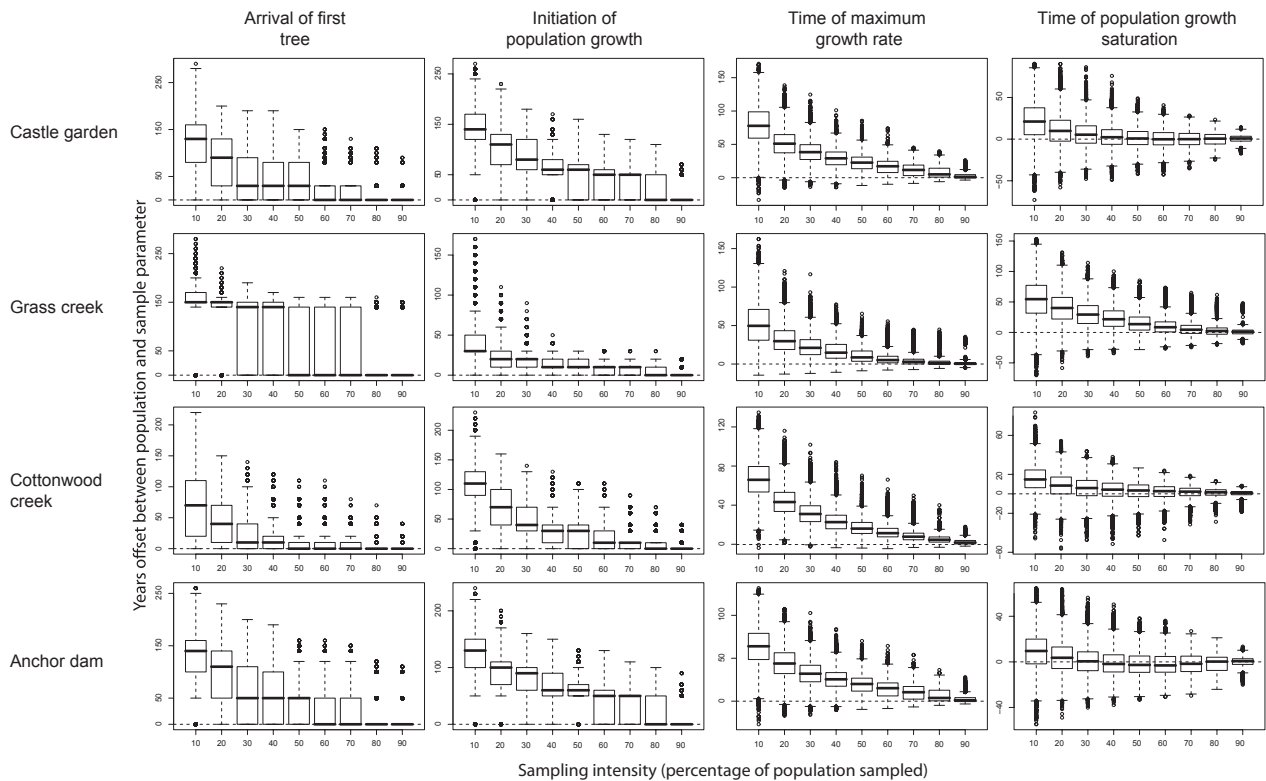


Fig. 4. Box-plots of offset of population growth parameters estimated from 10 000 simulations of random individual tree sampling. Results are shown for the four study populations. For each parameter, estimates are shown at sampling intensities from 10% to 90% of the full population. Values are given in years offset from the full population parameter estimate (population parameter – sample parameter). The box represents the first and third quartiles. The whiskers extend to the largest/smallest data point that falls within 1.5 times the box length from the nearest edge. The thick black bar is the median point of the data. Outliers are represented by black circles.

Garden, Cottonwood Creek and Anchor Dam (Fig. 5, Appendix S2). At Grass Creek the offset was less, and the assumed event was captured consistently with over three plots. With less than three plots, the offset was only 20 yrs, however variance increased substantially compared to the estimates based on higher numbers of plots (Fig. 5).

Maximum growth rate and growth saturation

All four populations showed a decline in establishment during the later stages of development (Fig. 2). As with the initial stages of population growth, less than complete sampling may easily miss individuals critical to accurately documenting this stage of population growth. Contrary to the results for the initiation of population growth, the greatest impact of sampling intensity on growth saturation for both methods was at Grass Creek (Figs 4, 5). However, even at the lowest sampling intensities, the offset in timing was ≤ 50 yrs. Cottonwood Creek showed an offset of 40 yrs with only one plot in the *n*-tree distance method, but otherwise offsets at Cottonwood Creek, Anchor Dam and Castle Garden were all ≤ 30 yrs (Appendices S1, S2).

Furthermore, the variance in estimates of growth saturation was minimal compared to other sampled variables (Figs 4, 5). Growth saturation was the only measured variable where median estimates underestimated the timing of the event in some populations at some intensities.

The timing of maximum growth rate was calculated as the point where the slope of the logistic regression model was maximized. Offsets in this variable showed consistent patterns between methods, however, the degree of offset and the variance in the estimates were generally higher for the *n*-distance method. At the lowest sampling intensities the median offset of maximum growth rate was between 140 and 30 yrs. In all four populations, for both methods, median estimates decreased systematically as sampling intensity increased. Variance also decreased very systematically with sampling intensity for this variable (Figs 4, 5).

Discussion

The purpose of this study was to investigate the extent to which the estimated shape and timing of population growth is sensitive to sampling method and intensity. We

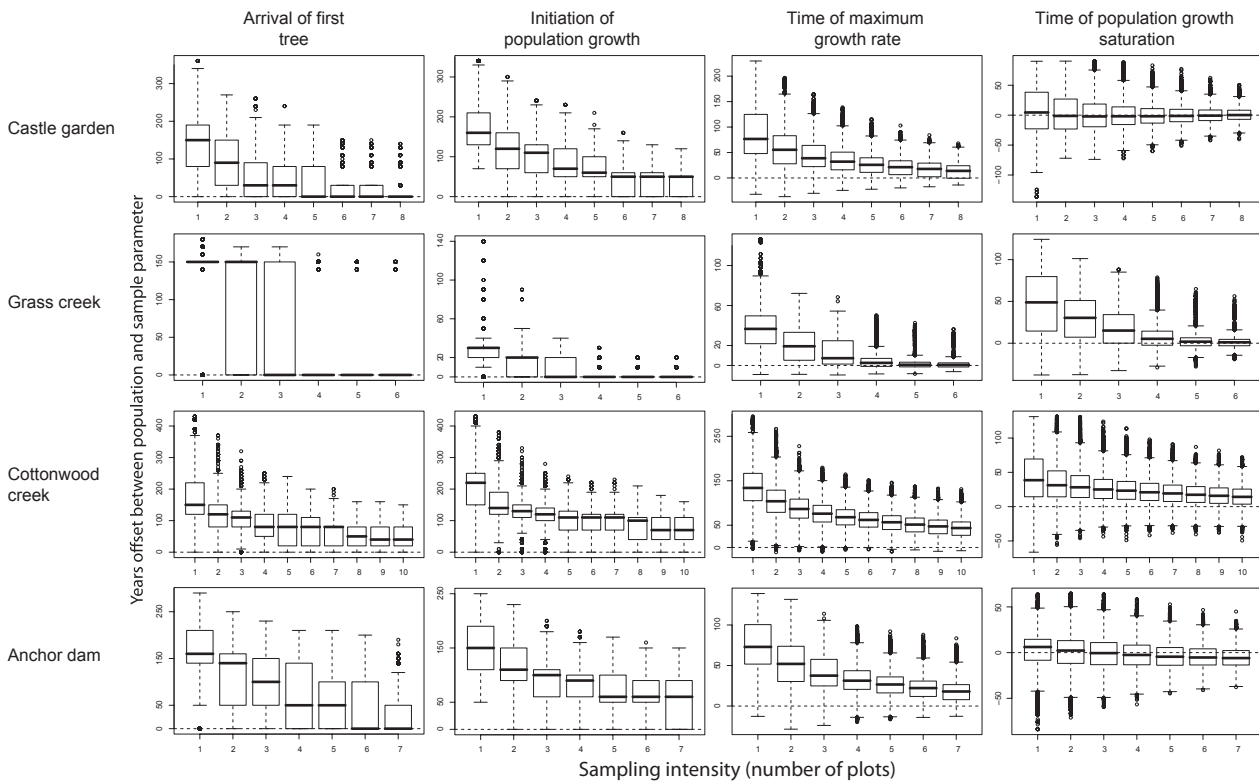


Fig. 5. Box-plots of offset of population growth parameters estimated from 10 000 simulations of n -distance plot sampling. Results are shown for the four study populations. For each parameter, estimates are shown at sampling intensities from one to the maximum number of plots for that population. Values are given in years offset from the full population parameter estimate (population parameter – sample parameter). The box represents the first and third quartiles. The whiskers extend to the largest/smallest data point that falls within 1.5 times the box length from the nearest edge. The thick black bar is the median point of the data. Outliers are represented by black circles.

used four natural populations of ponderosa pine where nearly every individual had been previously inventoried and aged. This allowed us to study the effects of sampling in complete populations, rather than simulating age structure and spatial patterns. Other studies using theoretical populations have performed simulations to test the effects of sampling (Doak et al. 2005; Grear & Elder 2008; Abadi et al. 2010), and some studies have used real population data (Fiske et al. 2008; Devenish Nelson et al. 2010; Buckley et al. 2010). However, generalization between species, let alone populations, for growth parameters is difficult (Buckley et al. 2010), and no other study has examined these parameters with a long-lived tree species.

The simulation results show that, for ponderosa pine, a large number of individuals need to be 'missing' (>60%) before significant effects are seen on the shape of the sample growth curve. These results suggest that most sampling strategies will better estimate the population growth rate without exhaustively sampling. However, as 'missing' or unsampled trees increased, the probability of missing early establishment events increased substantially (Appendices S1, S2). This means that while the shape of the sample

growth curve may match that of the population growth curve, the actual time involved in reaching that growth rate could be severely underestimated. The timing of initial colonization and the second establishment event showed the largest offsets and the highest amount of variance of all the measured variables. Variables related to later stages of population development, such as growth saturation and maximum growth rate, showed much less variance and deviation from the estimates based on the full population values, even at low sampling.

For the populations we studied, the individual random-tree method gave more accurate results than the n -tree distance method. However, there are several caveats to this finding. First, the study populations, while being real populations, are fairly unique in that they are all small, disjunct range-margin populations that have developed very slowly. In other population structures (e.g. large continuous forests) or higher-density populations, our findings may not apply and other sampling methods may be better suited. For example, Paulo et al. (2005) found that fixed plot sampling was superior to the n -distance method, and that both of these plot-based methods were superior to

transects in high density (>95 trees·ha⁻¹) cork oak (*Quercus suber*) stands.

Even within the four populations we studied, the effects of different demographic histories are apparent. Lowering the sampling intensity at the Grass Creek population shows much less effect on the timing and shape of population growth than at the other populations. This lowered effect is due to the unique demographic structure at Grass Creek, where there is an extended period (>100 yrs) of no establishment between the first and second individuals (Lesser & Jackson 2012). Furthermore, the bulk of establishment occurs in a relatively short time span in the mid to late 18th century. Combined, these factors create a situation where even at very low sampling intensities the parameters that we measured are relatively unaffected compared to the other populations, which all have a different demographic history. The difference in sampling effect between populations with different structures points to the issue that the reliability of sampling technique and intensity can change dramatically between populations.

Doak et al. (2005) showed that estimates of growth rate could be severely biased even at reasonably high sampling levels. Devenish Nelson et al. (2010) also showed that small sample sizes could produce highly uncertain results when estimating vital rates in red fox (*Vulpes vulpes* L.) populations. However, Fiske et al. (2008) found that while sampling bias was high at low levels of sampling in heliconia (*Heliconia acuminata* Rich.) populations, bias decreased rapidly with increasing sample size. For species such as pines, our results suggest that even modest sampling efforts can produce good estimates of population growth rate and the overall shape of the growth curve. However, our results also demonstrate that the time frames involved in population colonization and growth initiation may be underestimated even at high levels of sampling. Many other estimates such as extinction risk, population viability, elasticity and variability in vital rates are also of interest to demographers, and may be differentially sensitive to sampling effects than the variables we measured (Doak et al. 2005; Fiske et al. 2008). Furthermore, species with different life-history characteristics, such as short-lived herbaceous plants, may exhibit very different population structure and growth. Different dispersal strategies such as zoochory or serotiny may also alter population structure and growth. Hence, different sampling methods and intensities may better estimate population growth parameters for other species.

Understanding how growth parameters are estimated and the effects of sampling methods and intensity is needed to properly design and implement experiments. Our results indicate that, in low-density range-margin populations, for species with life history and recruitment patterns similar to ponderosa pine, growth rate is reliably

estimated with moderate levels of sampling, but that caution needs to be exercised in estimating the timing of early stages population development. Given the anticipated distributional shifts that species might make in response to on-going climate change (Rosenzweig et al. 2007; Williams et al. 2007), underestimation of the time required to successfully colonize new territory could have major ecological and management implications.

Acknowledgements

Support for Lesser was provided by a USDA McIntire-Stennis Competitive Grant and a NSF Doctoral Dissertation Improvement Grant (DEB-0910173). Brewer was supported by NSF Grant EAR-1003848. We thank J. Andersen, C. Benkman, R. Jones, S. T. Jackson, D. Doak and two anonymous reviewers for critical comments and statistical advice.

References

- Abadi, F., Gimenez, O., Arlettaz, R. & Schaub, M. 2010. An assessment of integrated population models: bias, accuracy, and violation of the assumption of independence. *Ecology* 91: 7–14.
- Biondi, F., Jamieson, L.P., Strachan, S. & Sibold, J. 2011. Dendroecological testing of the pyroclimatic hypothesis in the central Great Basin, Nevada, USA. *Ecosphere* 2: 1–20.
- Brown, P.M. 2006. Climate effects on fire regimes and tree recruitment in Black Hills ponderosa pine forests. *Ecology* 87: 2500–2510.
- Brown, P.M. & Wu, R. 2005. Climate and disturbance forcing of episodic tree recruitment in a southwestern ponderosa pine landscape. *Ecology* 86: 3030–3038.
- Buckley, Y.M., Ramula, S., Blomburg, S.P., Burns, J.H., Crone, E.E., Ehrlen, J., Knight, T.M., Pichancourt, J., Quested, H. & Wardle, G.M. 2010. Causes and consequences of variation in plant population growth rate: a synthesis of matrix population models in a phylogenetic context. *Ecology Letters* 13: 1182–1197.
- Clogg, C.C., Petkova, E. & Haritou, A. 1995. Statistical methods for comparing regression coefficients between models. *American Journal of Sociology* 100: 1261–1293.
- Crawley, M.J. 1990. The population dynamics of plants. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 330: 125–140.
- Devenish Nelson, E.S., Harris, S., Soulsbury, C.D., Richards, S.A. & Stephens, P.A. 2010. Uncertainty in population growth rates: determining confidence intervals from point estimates of parameters. *PLoS ONE* 5: e13628.
- Doak, D.F., Gross, K. & Morris, W.F. 2005. Understanding and predicting the effects of sparse data on demographic analysis. *Ecology* 86: 1154–1163.

- Fiske, I.J., Bruna, E.M. & Bolker, B.M. 2008. Effects of sample size in estimates of population growth rates calculated with matrix models. *PLoS ONE* 3: e3080.
- Fritts, H.C. & Swetnam, T.W. 1989. Dendroecology: a tool for evaluating variations in past and present forest environments. *Advances in Ecological Research* 19: 111–175. (Begon, M., MacFadyen, A., Fitter, A.H. & Ford, E.D., eds.) Academic Press, San Diego, CA, USA.
- Gray, S.T., Fastie, C.L., Jackson, S.T. & Betancourt, J.L. 2004. Tree-ring based reconstructions of precipitation in the Big-horn Basin, Wyoming since 1260 A.D. *Journal of Climate* 17: 3855–3865.
- Grear, J.S. & Elder, B.D. 2008. Bias in population growth rate estimation: sparse data, partial life cycle analysis and Jensen's inequality. *Oikos* 117: 1587–1593.
- Heyerdahl, E.K., Miller, R.F. & Parsons, R.A. 2006. History of fire and Douglas-fir establishment in a savanna and a sagebrush-grassland mosaic, southwestern Montana, USA. *Forest Ecology and Management* 230: 107–118.
- Jonsson, B., Holm, S. & Kallur, H. 1992. A forest inventory method based on density-adapted circular plot size. *Scandinavian Journal of Forest Research* 7: 405–421.
- Kaye, M.W., Woodhouse, C.A. & Jackson, S.T. 2010. Persistence and expansion of ponderosa pine woodlands in the west-central Great Plains during the past two centuries. *Journal of Biogeography* 37: 1668–1683.
- League, K. & Veblen, T. 2006. Climatic variability and episodic *Pinus ponderosa* establishment along the forest-grassland ecotones of Colorado. *Forest Ecology and Management* 228: 98–107.
- Lessard, V.C., Drummer, T.D. & Reed, D.D. 2002. Precision of density estimates from fixed-radius plots compared to N-tree distance sampling. *Forest Science* 48: 1–6.
- Lesser, M.R. & Jackson, S.T. 2012. Making a stand: five centuries of population growth in colonizing stands of *Pinus ponderosa*. *Ecology* 93: 1071–1081.
- Meyer, S., Ingersoll, C.G., McDonald, L.L. & Boyce, M.S. 1986. Estimating uncertainty in population growth rates: jackknife vs. bootstrap techniques. *Ecology* 67: 1156–1166.
- Morris, W.F. & Doak, D.F. 2002. *Quantitative conservation biology: theory and practice of population viability analysis*. Sinauer, Sunderland, MA, US. 480.
- Norman, S.P. & Taylor, A.H. 2005. Pine forest expansion along a forest-meadow ecotone in northeastern California, USA. *Forest Ecology and Management* 215: 51–68.
- Paulo, M.J., Tomé, M., Otten, A. & Stein, A. 2005. Comparison of three sampling methods in characterization of cork oak stands for management purposes. *Canadian Journal of Forest Research* 35: 2295–2303.
- Rosenzweig, C., Casassa, G., Karoly, D.J., Imeson, A., Liu, C., Menzel, A., Rawlins, S., Root, T.L., Seguin, B. & Tryjanowski, P. 2007. Assessment of observed changes and responses in natural and managed systems. In: Parry, M.L., Canziani, O.F., Palutikof, J.P., van der Linden, P.J. & Hanson, C.E. (eds.) *Climate change 2007: Impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, pp.79–131. Cambridge University Press, Cambridge, UK.
- Sibly, R.M. & Hone, J. 2002. Population growth rate and its determinants: an overview. *Philosophical Transactions of the Royal Society of London, series B: Biological Science* 357: 1153–1170.
- Stacey, P.B. & Taper, M. 1992. Environmental variation and the persistence of small populations. *Ecological Applications* 2: 18–29.
- Veblen, T.T., Kitzberger, T. & Donnegan, J. 2000. Climatic and human influences on fire regimes in ponderosa pine forests in the Colorado Front Range. *Ecological Applications* 10: 1178–1195.
- White, A.S. 1985. Presettlement regeneration patterns in a southwestern ponderosa pine stand. *Ecology* 66: 589–594.
- Williams, J.W., Jackson, S.T. & Kutzbach, J.E. 2007. Projected distributions of novel and disappearing climates by 2100 AD. *Proceedings of the National Academy of Sciences, USA* 104: 5738–5742.

Supporting Information

Additional supporting information may be found in the online version of this article:

Appendix S1. Random-tree sampling method simulations.

Appendix S2. *n*-Distance sampling method simulations.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.