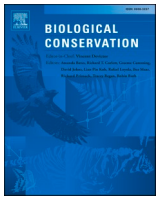
Biological Conservation 289 (2024) 110381

Contents lists available at ScienceDirect

Biological Conservation

journal homepage: www.elsevier.com/locate/biocon



Rate and extent of genetic diversity loss under non-equilibrium scenarios of habitat loss

Qian Tang [a], Tak Fung [a], Drew E. Terasaki Hart [b,c], Frank E. Rheindt [a,*]

[a] National University of Singapore, Department of Biological Sciences, Singapore
[b] The Nature Conservancy, USA
[c] University of California, Berkeley, USA

ARTICLE INFO

Keywords:
Anthropogenic extinction crisis
Biodiversity
Conservation planning
Landscape genetics
Protected area
SLOSS debate

ABSTRACT

Habitat loss is one of the greatest threats to biodiversity, but there is considerable debate over the interplay between the total amount of habitat lost versus the degree of habitat fragmentation. Previous studies on this topic focused on the effects of habitat loss on species richness or genetic diversity over long timescales, while neglecting shorter timescales that are of immediate conservation concern. To address this knowledge gap, we examined the rate and extent of genetic diversity loss under different non-equilibrium scenarios of habitat loss, by performing analytical calculations for a non-spatial setting and individual-based simulations for spatially explicit settings, including a real-world case study of malleefowl (*Leipoa ocellata*, Galliformes) populations in Australia. Our work revealed that the total amount of habitat lost had the biggest negative effect on genetic diversity via reductions in population abundance and associated genetic drift, with the degree of fragmentation having smaller but nonetheless substantial negative effects. The latter result suggested that to optimize the conservation of genetic diversity, it is better to preserve a single large reserve over several small ones. Furthermore, reductions in population abundance led to loss of genetic diversity in the population only after long time-lags, which highlights the potential for genetic rescue shortly after habitat loss. The malleefowl case study revealed how sampling uncertainty due to low sample sizes can blur the effects of habitat loss on genetic diversity, underscoring the limitations of conservation genetic studies based on small sample size and uneven spatial distribution.

1. Introduction

Evolutionary potential determines the long-term viability of populations and constitutes a fundamental aspect of biodiversity (Lerouzic and Carlborg, 2008). Genetic diversity, as estimated from genome-wide markers (Benestan et al., 2016; DeWoody et al., 2021; Harrisson et al., 2014; Lande and Shannon, 1996), is a sensible indicator of evolutionary potential. In a population under equilibrium, the genetic diversity lost via drift, selection, and emigration should equal the gain in genetic diversity via mutation and immigration (Franklin, 1980). However, under recent and dramatic anthropogenic pressures such as harvesting and habitat loss, the equilibrium often breaks down due to a corresponding rapid loss of individuals and hence genetic diversity, which is not fully compensated by mutation and immigration (Pérez-Pereira et al., 2022). Consequently, the reduced genetic diversity may lead to increased risk of population collapse due to environmental fluctuations (Kardos et al., 2021). Conserving genetic diversity has recently been proposed as part of a global arsenal of strategies for biodiversity conservation (Díaz et al., 2020). However, our current knowledge of how human impacts change genetic diversity is in its infancy and there is an urgent need for greater understanding (Exposito-Alonso et al., 2022; Miraldo et al., 2016).

In this study, we focus on habitat loss, which is one of the main drivers of reductions in biodiversity in the Anthropocene (Díaz et al., 2019; Symes et al., 2018). Given the large negative impacts of habitat loss, nature reserves that conserve habitat are still one of the most effective methods for biodiversity conservation (Gaston et al., 2008). But economic development driven by human population growth and rapid urbanization has led to intensified competition for land (Kim et al., 2014). As a result, nature reserves are becoming smaller and more fragmented (Lewis et al., 2019; Volenec and Dobson, 2020), thus highlighting a need for more research on how to design the spatial configuration of reserves to meet conservation needs.

* Corresponding author at: 16 Science Drive 4, Block S3 #04-01, 117558, Singapore.
E-mail address: dbsrfe@nus.edu.sg (F.E. Rheindt).

https://doi.org/10.1016/j.biocon.2023.110381
Received 29 May 2023; Received in revised form 6 November 2023; Accepted 15 November 2023
Available online 17 December 2023

The relative importance of the spatial configuration and extent of habitats has been the subject of a recent debate (Fahrig, 2017; Fletcher et al., 2018) among conservationists, keen to inform the optimal design of reserves to achieve maximum functionality of ecosystems in the face of ongoing habitat loss (Cabeza, 2003). This recent exchange is reminiscent of the classic SLOSS debate on whether to conserve a single large or several small reserves of equal combined size (Diamond, 1975; Simberloff and Abele, 1982; Wilcox and Murphy, 1985). Empirical evidence at different spatial scales often suggests that the total amount of habitat lost is the predominant factor (see review in Fahrig, 2017), but ambiguities regarding the effects of different habitat loss scenarios suggest that a more thorough understanding is still required (Fletcher et al., 2018). Most previous studies have based their conclusions on species richness. A few simulation studies have investigated the effects of different habitat loss scenarios on genetic diversity (Jackson and Fahrig, 2016, 2014), but only examined the long-term effects after 1000 generations, with population size and genetic diversity reaching stable values. This long-term focus neglects the transient changes that occur on

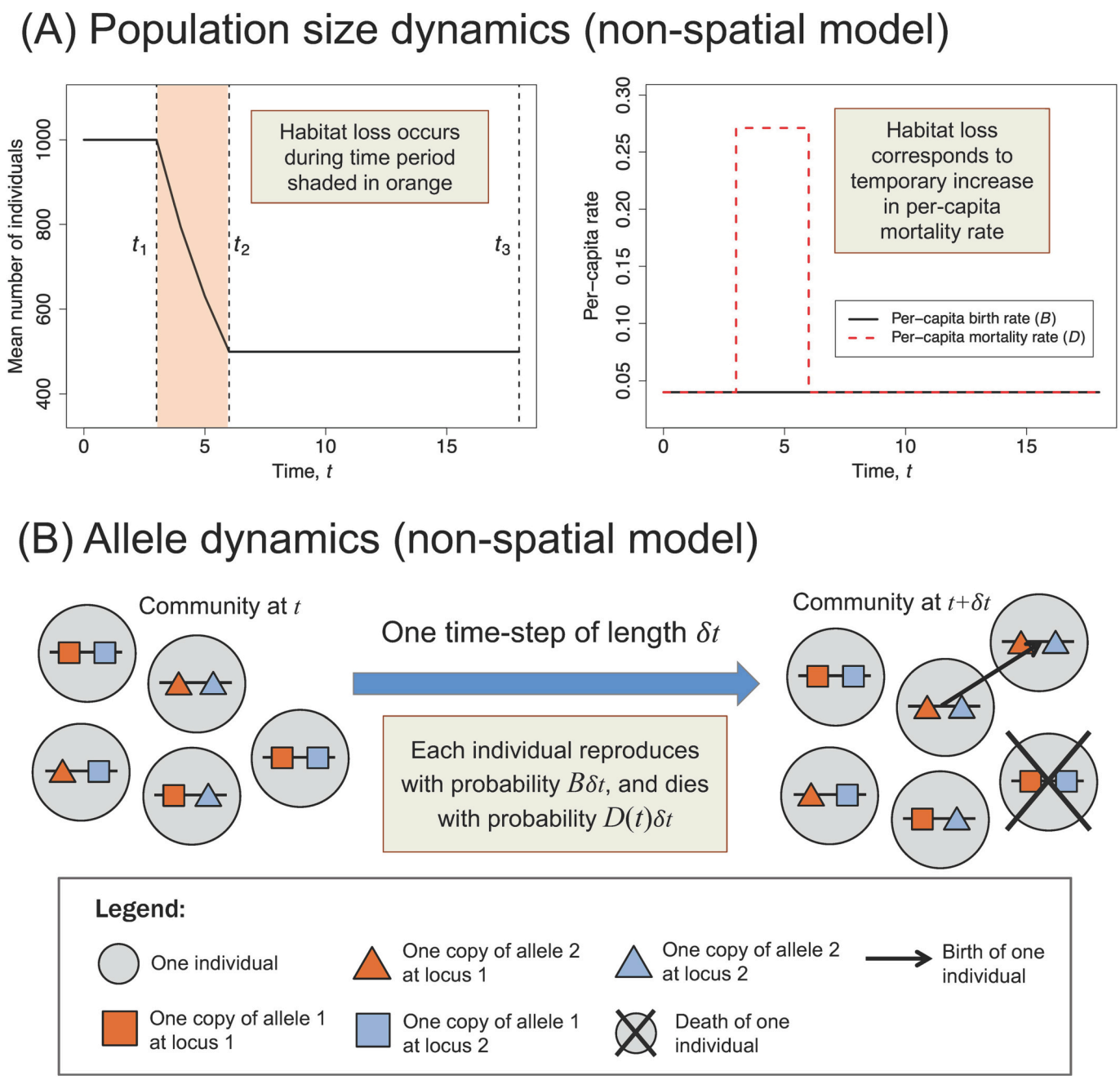


Fig. 1. Schematic diagram of the non-spatial model used, illustrating model structure and dynamics. (A) Illustration of the population size dynamics from the model. The mean population size (number of individuals) is reduced by a period of habitat loss from time t_1 to t_2, after which the mean population size stabilizes at a lower value. The period of habitat loss corresponds to a period of elevated per-capita mortality rate. (B) Illustration of allele dynamics from the model and how they are related to the population dynamics. In the illustration, there is a community of five haploid individuals, each with two biallelic loci. The diagram shows an example of allele dynamics for the community of individuals over one short time-step of length δt, from time t to time $t+\delta t$. In this time-step, each individual reproduces asexually with probability $B\delta t$ and dies with probability $D(t)\delta t$. In the example dynamics shown, one individual reproduces and one individual dies in the time-step considered. The random births and deaths introduce demographic stochasticity (ecological drift) into the model, which in turn causes random changes in allele frequencies (genetic drift). The model is non-spatial and hence the spatial locations of the five individuals do not affect model dynamics.

shorter timescales, which are relevant to more-immediate conservation applications.

In this study, we address this critical knowledge gap by performing mathematical analyses and individual-based simulations to investigate transient changes in population abundance and genetic diversity across species under various scenarios of habitat loss. These scenarios relate to different types of spatial configurations by which habitat is reduced and fragmented, thus disrupting population and allele dynamics from their equilibrium state. In particular, fragmentation can lead to more isolated subpopulations whereas a reduced extent of habitat can lead to a lower population size, both of which increase genetic drift and hence cause genetic diversity to be temporarily non-equilibrium. Thus, we refer to these scenarios as non-equilibrium scenarios of habitat loss. Within the parameters of our analyses, our results provide guidance on the potential negative effects of habitat loss on genetic diversity and hence evolutionary potential, with implications for efforts to optimize the spatial configuration of reserves.

2. Methods

2.1. Non-spatial population model

We first performed mathematical analyses on a simplified non-spatial individual-based population model to provide baseline expectations of how habitat loss affects genetic diversity, which were then compared with results from simulations of a much more complex spatially explicit population model.

The non-spatial model represented the abundance and allele dynamics of a population of haploid individuals that reproduced asexually and underwent mortality. These individuals were assumed to have the same fitness, such that there were no selection effects. The individuals experienced random birth and death events at specified (mean) per-capita rates, which introduced demographic stochasticity (also known as ecological drift) into the population. This demographic stochasticity meant that the time trajectory of population size differed each time the model was run. In addition, the random births and deaths inherent in demographic stochasticity resulted in random changes in the number of copies of each allele, i.e., genetic drift. The per-capita birth rate was assumed to be constant over time, whereas the per-capita mortality (or death) rate was changed over time to represent the effects of habitat loss. This habitat loss resulted in a reduction in population size and hence an increase in genetic drift, thus causing genetic diversity to be non-equilibrium.

For a particular model run (see Fig. 1 for a schematic diagram), let the population size be $N(t)$ at time t. Also, let the constant per-capita birth rate be B and the time-varying per-capita mortality rate be $D(t)$ at time t, which were both assumed to be density-independent. Across model runs, denote the mean population size at time t by $\overline{N(t)}$. To model habitat loss, the mean population size was held constant at $\overline{N(0)}$ over the time period $[0, t_1)$, before being decreased over $[t_1, t_2)$ from $\overline{N(0)}$ to $\overline{N(t_2)} = p\overline{N(0)}$, where $0 \le p < 1$. Thereafter, during $[t_2, t_3]$, the mean population size remained constant at $\overline{N(t_2)}$. These changes in mean population size corresponded to the per-capita birth and mortality rates being equal during $[0, t_1)$, i.e. $B = D(0) = D_0$; the per-capita mortality rate increasing from D_0 to $D_1 > D_0$ during $[t_1, t_2)$; and the per-capita mortality rate decreasing back to D_0 during $[t_2, t_3]$. More specifically, the dynamics of $\overline{N(t)}$ were specified by the equation:

$$\frac{d\overline{N(t)}}{dt} = (B - D(t))\overline{N(t)}. \tag{1}$$

Solving Eq. (1) during $[t_1, t_2)$ gave:

$$\overline{N(t_2)} = \overline{N(0)}e^{(B-D_1)(t_2-t_1)}, \tag{2}$$

which was rearranged to give:

$$D_1 = B - \frac{ln(p)}{t_2 - t_1}. \tag{3}$$

We considered L independent loci each with two alleles, corresponding to single nucleotide polymorphisms (SNPs). The mutation rate was considered to be so small that it can be neglected during the ecological timescales considered. We derived formulae specifying how genetic diversity changed before, during, and after habitat loss using two indicators of genetic diversity: mean nucleotide diversity (Nei and Li, 1979) and mean proportion of fixed loci (Hedrick, 1994). These formulae together with details of their derivation are novel and are hence presented in the Results section.

In comparison with previous theoretical analyses of neutral models in population genetics, our analyses are novel because we consider a model with time-varying population size. The classic Wright-Fisher model (Fisher, 1923; Wright, 1931) represents the stochastic neutral allele dynamics of a population with constant size and non-overlapping generations. The Wright-Fisher model has been analyzed to derive quantities such as the mean time for an allele to be lost given a particular initial number of copies of the allele, and the steady-state distribution of allele frequencies at an equilibrium between genetic drift and mutation (Ewens, 2004; Kimura, 1955). Another classic model that has been analyzed in similar ways is the neutral Moran process, which represents the stochastic neutral allele dynamics of a population with constant size and overlapping generations (Ewens, 2004; Moran, 1958). If the population size in the Moran process is large enough, then the assumption of a population with constant size can be relaxed to a population with constant *mean* size, with little effect on steady-state dynamics (a result that was derived in the context of neutral models in ecology, with a community instead of a population and species instead of alleles; Volkov et al., 2003). The Wright-Fisher model and neutral Moran process have been extended in various ways, such as by adding selection, geographic structure and age structure (Ewens, 2004), but to our knowledge has not been extended to include time-varying population size. The neutral non-spatial model with time-varying mean population size that we construct and analyze in our study can be conceptualized as an extension of the neutral Moran process with constant mean population size.

2.2. Spatially explicit population model

We produced individual-based, spatially explicit forward-time simulations for diploid populations subjected to different scenarios of habitat loss using *Geonomics* v1.3.8 (Terasaki Hart et al., 2021). The model operated in discrete time with timesteps of equal length. In each timestep, individuals moved in a random direction and the distance traveled was specified by a lognormal distribution, which is a parsimonious leptokurtic distribution that generally describes animal movement in many empirical studies; see Hawkes (2009) and references therein. Afterwards, mature individuals randomly searched for and chose a mate of the opposite sex within a specified radius (Table 1), with successful mating resulting in production of recombinant offspring. Lastly, individuals experienced density-dependent mortality according to a discrete-time logistic equation. The specific values of the life-history parameters that we used were loosely based on birds in general (Table 1), with one timestep being defined as one year. We examined three species of different mobility, measured as the distance traveled per year. To choose values for this distance, we noted that the natal and breeding distances of most resident bird species fall within the range between 0.1 km to 5 km (Martin and Fahrig, 2018). Thus, mean distance traveled per year was on the order of tens of meters to kilometers, and we chose to examine the values 0.2 km, 0.6 km, and 1 km.

For each of the three species, we produced a raster of 20 cells × 20 cells representing a square habitat of dimension 20 km × 20 km. Initially, all 400 cells were designated as pristine habitat, each with a carrying capacity of 10 individuals km^{-2} that was based on a median number from a meta-analysis (Stephens et al., 2019). We tested 18

Table 1
Biological parameters and values used in simulations of the spatially explicit population model.

Parameter	Sp. A	Sp. B	Sp. C	Malleefowl
Carrying capacity in pristine habitat (km^{-2})	10 (Stephens et al., 2019)			10 (Frith, 1962)
Carrying capacity in cleared habitat (km^{-2})	Variable			2 (Frith, 1962)
Age at sexual maturity (years)	1 (Claramunt, 2021)			3 (Department of Parks and Wildlife, 2016)
Sex ratio (male:female)	1			1 (Department of Parks and Wildlife, 2016)
Intrinsic population growth rate ($year^{-1}$)	0.5			0.5
Probability of mating for a pair	0.1			0.1
Mean no. of offspring per mating pair	7 (Claramunt, 2021)			7 (Department of Parks and Wildlife, 2016)
Mating radius (km)	1			1
Maximum age (years)	25 (Minias and Podlaszczuk, 2017)			28 (Department of Parks and Wildlife, 2016)
Movement distance (km $year^{-1}$)	0.2	0.6	1	1 (Stenhouse and Moseby, 2023)
Neutral mutation rate (per base)	1.00×10^{-9}			1.00×10^{-8} (Bergeron et al., 2023)

habitat loss scenarios, which were all combinations of two different spatial configurations of habitat preservation (random and clumped), three different proportional amounts of habitat cleared (20 %, 50 % and 80 %) and three different carrying capacities in the cleared habitat cells (10 %, 20 %, and 50 % of pristine habitat cells). Since the mortality of individuals in a cell was calculated based on the individual density (number of individuals) and carrying capacity of the cell (according to a discrete-time logistic equation), individuals that enter a cleared habitat cell experienced greater density-dependent mortality than individuals in a pristine habitat cell. Thus, individuals survived and moved about in cleared habitat cells but at a lower density than in pristine habitat cells. With respect to the two different spatial configurations, the random configuration produced habitat maps with a Moran's I of −0.0105 to 0.0143, whereas the clumped configuration produced habitat maps with a larger Moran's I of 0.0756 to 0.308. The random configuration corresponded to a high degree of habitat fragmentation and was achieved by randomly picking pristine habitat cells to preserve. In contrast, the clumped configuration corresponded to a low degree of habitat fragmentation and consisted of randomly choosing which pristine habitat cells to preserve under the constraint that each cell chosen must be adjacent (horizontally or vertically) to a cell that has already been chosen (see Fig. 2 for examples).

Each habitat loss scenario started with 4000 individuals randomly distributed across the entire area, with allele frequencies for 1000 biallelic loci (corresponding to SNPs) being chosen randomly according to a uniform distribution and alleles distributed randomly among individuals. Results from our non-spatial model indicated that trends in genetic diversity under habitat loss were qualitatively robust to use of different initial distributions of allele frequencies (see Fig. 3 in the Results and Figs. S1, S2 and S3 in Supplementary material). Thus, we did not explore other initial distributions of allele frequencies for our spatial model. The 1000 loci were not linked and hence underwent free recombination. After initialization of individuals and alleles in the spatial model, the model dynamics were simulated for 60–70 years (each year corresponding to a timestep) of burn-in until the size and distribution of the population had stabilized. After the burn-in period, the model was simulated for ten further years without habitat loss, which was only applied in year 11 ($t = 11$). After the habitat loss event, dynamics were simulated for a further 199 years (until $t = 210$). For all three species, counts of individuals were collected every year whereas genetic data for the 1000 loci were collected every five years to calculate genetic diversity in the form of nucleotide diversity and the proportion of fixed loci (using *VCFtools* v0.1.16; Danecek et al., 2011). To investigate the level of inbreeding and population genetic spatial structuring ("genetic structuring" hereafter), which are potential mechanisms mediating genetic diversity loss, we calculated the inbreeding coefficient (*FIS*, as calculated with *VCFtools*) and the correlation coefficient between spatial and genetic distances (Mantel *r* from Mantel tests as calculated with *R* package *vegan*; Dixon, 2003). For each species and habitat loss scenario, we performed 100 repeat simulations to account for stochasticity. In total, we examined 57 different scenarios consisting of 18 habitat loss scenarios and a null scenario (with no habitat loss) for three species (Table S1). The loss of habitat resulted in a reduction in the extent of habitat and fragmentation of the habitat, thus altering population and allele dynamics and causing genetic diversity to be non-equilibrium.

We also used *Geonomics* to perform simulations for scenarios of habitat loss based on a real-world example from South Australia, the malleefowl (*Leipoa ocellata*, Galliformes). The malleefowl is generally monogamous, though not strictly so, and tends to maintain a single territory (Weathers et al., 1990). No distinct population structure in malleefowl has been observed across the Eyre Peninsula, South Australia (Stenhouse et al., 2022), even though dispersal after breeding is relatively rare (Stenhouse and Moseby, 2023). The malleefowl is classified as vulnerable under the IUCN Red List due to likely future declines under a drying climate and more frequent fires (BirdLife International, 2023). In our simulations, we used values of biological parameters, such as lifespan, mobility and carrying capacity, based on previous studies of malleefowl (Table 1). The simulations were performed on seven rasters of 20 cells × 20 cells, with each raster corresponding to a 20 × 20 km area in South Australia where at least three malleefowl individuals had been sampled by an empirical study (Stenhouse et al., 2022). Because the sampling was carried out using transects, we refer to each 20 × 20 km area as a "transect". Within each transect, the current habitat map was constructed using vegetation data collected from the National Vegetation Information System (NVIS) Version 6.0. The data was categorized into two classes, namely native vegetation and disturbed vegetation. The disturbed vegetation was subject to anthropogenic grazing and fires, which reduced the carrying capacity of malleefowl by 80 % compared with the carrying capacity inside native vegetation (Frith, 1962). Thus, we considered native and disturbed vegetation to correspond to pristine and cleared habitat in the simulation model. The majority of habitat clearing in South Australia occurred during the 19th century (Bradshaw, 2012). Thus, we simulated two simple habitat loss scenarios: one in which all habitat loss (clearance) occurred instantaneously during the mid-19th century in 1850, and the other in which habitat was lost continuously from 1800 to 1900. For each habitat loss scenario, we performed 100 repeat simulations to account for stochasticity. Also, counts of individuals were collected every year whereas genetic data for 1000 loci were collected every five years to calculate genetic diversity in the form of nucleotide diversity and the proportion of fixed loci. We calculated genetic diversity under two sampling schemes: one in which all simulated individuals were sampled (complete sampling scheme) and the other in which individuals were only sampled from locations corresponding to the sampling locations in the accompanying empirical study (Stenhouse et al., 2022; real sampling scheme). The simulation results were compared with the empirical data to determine the extent to which trends in the data could have been explained by habitat loss or by incomplete sampling.

3. Results

3.1. Non-spatial population model

Firstly, we present details on the derivation of formulae specifying the mean nucleotide diversity at time t under the habitat loss scenarios considered. Let locus l ($1 \leq l \leq L$) have $n_{l1}(t)$ copies of the first allele and

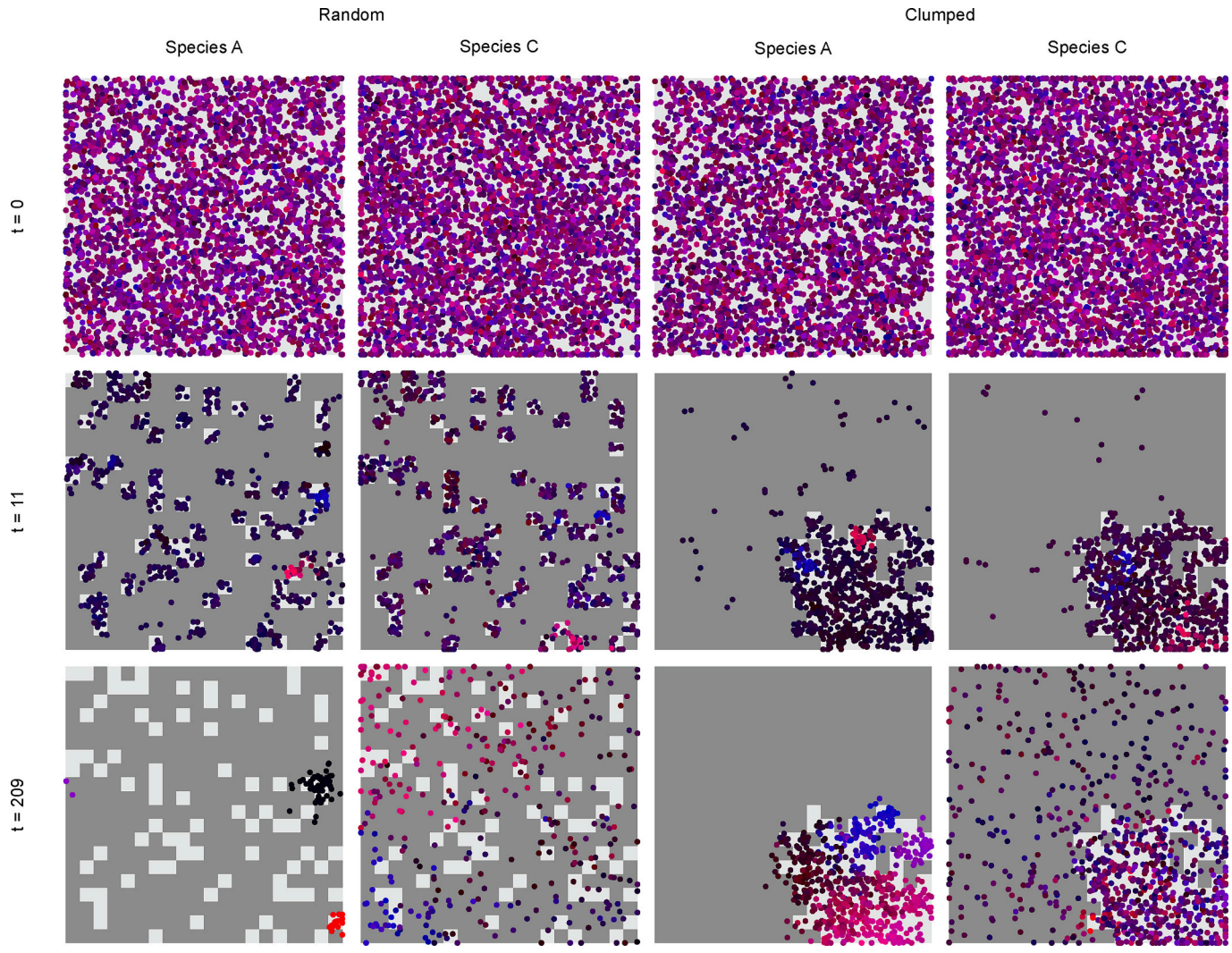


Fig. 2. Temporal changes in spatial genetic structure of two species with different mobility from our spatially explicit simulations, under two habitat loss scenarios with preserved areas being randomly chosen or spatially clumped. Individuals (represented by circles) are colored according to their genetic structure, which was calculated by converting the top two principal components in a principal component analysis into scaled values for red and blue color channels, such that individuals with more similar colors were more similar genetically. Species A had low mobility whereas Species C had high mobility, with mean dispersal distances per year of 0.2 and 1 km respectively. For each species and habitat loss scenario, spatial genetic structure is shown before habitat loss, immediately after habitat loss, and several decades after habitat loss, corresponding to the different rows. Light and dark gray cells correspond to pristine and cleared habitat respectively, and in both habitat loss scenarios, 80 % of pristine habitat was cleared. Cleared habitat had a carrying capacity that was 20 % of that for pristine habitat. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$n_{l2}(t)$ copies of the second allele at time t. Considering just this locus, the nucleotide diversity at time t for an extant population was specified by Nei and Li (1979) as

$$\pi_l(t) = 2\left(\frac{n_{l1}(t)}{n_{l1}(t)+n_{l2}(t)}\right)\left(\frac{n_{l2}(t)}{n_{l1}(t)+n_{l2}(t)}\right) = 2\left(\frac{n_{l1}(t)}{N(t)}\right)\left(\frac{N(t)-n_{l1}(t)}{N(t)}\right) = 2\left(\frac{n_{l1}(t)}{N(t)} - \left(\frac{n_{l1}(t)}{N(t)}\right)^2\right), \tag{4}$$

where $n_{l1}(t)+n_{l2}(t) = N(t)$, the non-zero population size at time t. Thus,

$$\overline{\pi_l(t)} = 2\left(\overline{\frac{n_{l1}(t)}{N(t)}} - \overline{\left(\frac{n_{l1}(t)}{N(t)}\right)^2}\right) = 2\left(\overline{\frac{n_{l1}(t)}{N(t)}} - \left(V\left(\frac{n_{l1}(t)}{N(t)}\right) + \left(\overline{\frac{n_{l1}(t)}{N(t)}}\right)^2\right)\right), \tag{5}$$

where V denotes the variance (across model runs) due to genetic drift. It can be shown (see Supplementary material) that

$$\overline{\frac{n_{l1}(t)}{N(t)}} \approx \overline{\frac{n_{l1}(t)}{N(t)}}\left(1+\frac{V(N(t))}{(\overline{N(t)})^2}\right) - \frac{V(n_{l1}(t))}{(\overline{N(t)})^2}, \tag{6}$$

$$V\left(\frac{n_{l1}(t)}{N(t)}\right) \approx \frac{V(n_{l1}(t))}{(\overline{N(t)})^2}\left(1-2\overline{\frac{n_{l1}(t)}{N(t)}}\right) + \left(\overline{\frac{n_{l1}(t)}{N(t)}}\right)^2\frac{V(N(t))}{(\overline{N(t)})^2}, \tag{7}$$

where the approximations arose due to a second-order Taylor-series expansion of $n_{l1}(t)/N(t)$ around the mean (across model runs) and

$$\overline{\frac{n_{l1}(t)}{N(t)}} = \frac{n_{l1}(0)}{N(0)}, \tag{8}$$

$$\frac{V(N(t))}{(\overline{N(t)})^2} = \frac{\int_0^t e^{\rho(y)}(B+D(y))dy}{N(0)}, \tag{9}$$

$$\frac{V(n_{l1}(t))}{(\overline{N(t)})^2} = \frac{n_{l1}(0)}{(N(0))^2}\int_0^t e^{\rho(y)}(B+D(y))dy. \tag{10}$$

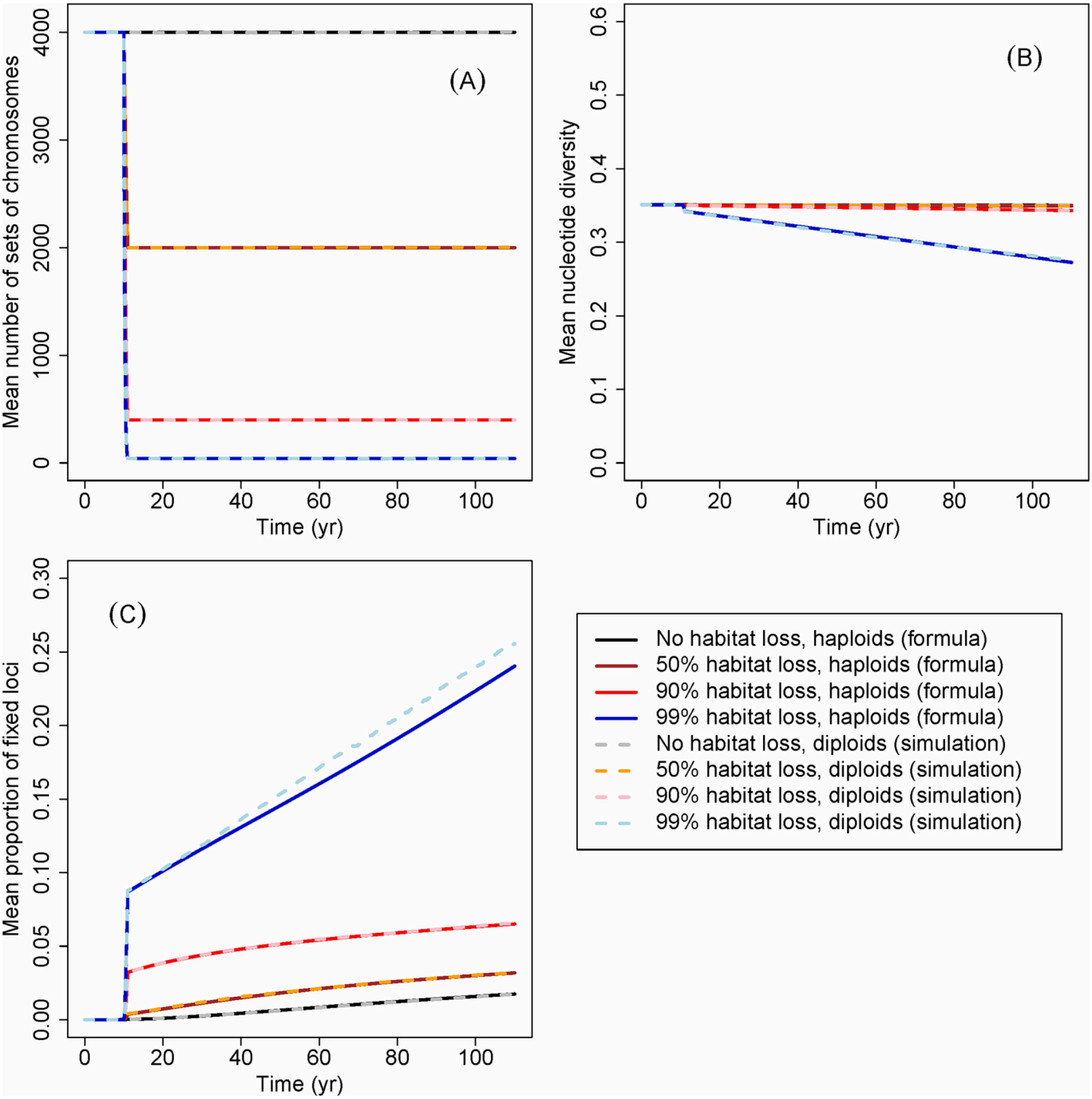


Fig. 3. Mean (A) number of sets of chromosomes, (B) nucleotide diversity and (C) proportion of fixed loci over time from the non-spatial population model, under three different habitat loss scenarios with different proportional amounts of habitat lost, as derived using formulae (haploid case) and simulations (diploid case). In each habitat loss scenario, habitat loss occurred between $t_1 = 10$ years and $t_2 = 11$ years. The per-capita death rate was $D_0 = 0.04$ year^{-1} when there was no habitat loss and was given by D_1 (Eq. (3)) during habitat loss. The value of 0.04 year^{-1} was based on the maximum life span of a typical bird species (Table 1). The per-capita birth rate was set constant at $B = 0.04$ year^{-1}, such that the mean population size was constant when there was no habitat loss. 100 loci, each with two alleles, were considered. For each locus, the initial number of copies of the first allele was drawn from a random uniform distribution bounded by 0 and the initial population size. The initial population size was 4000 and 2000 for the haploid and diploid cases, respectively. For comparison, results from a baseline scenario with no habitat loss are also shown. We obtained qualitatively similar results when the initial number of copies of the first allele was drawn using the steady-state allele frequency distribution of a neutral model with constant population size (Figs. S2 and S3).

In Eqs. (8)–(10), $\rho(y) = 0$ for $[0, t_1)$, $\rho(y) = (D_1 - B)(y - t_1)$ for $[t_1, t_2)$ and $\rho(y) = (D_1 - B)(t_2 - t_1)$ for $[t_2, t_3]$, where y is a dummy variable used in the integrals. Analogously, the mean nucleotide diversity considering all L independent loci was

$$\overline{\pi_L(t)} = \frac{1}{L}\sum_{l=1}^{L}\overline{\pi_l(t)} \approx \frac{2}{L}\sum_{l=1}^{L}\left(\frac{\overline{n_{l1}(t)}}{N(t)} - \left(V\left(\frac{n_{l1}(t)}{N(t)}\right) + \left(\frac{\overline{n_{l1}(t)}}{N(t)}\right)^2\right)\right). \tag{11}$$

We note that $\overline{\pi_L(t)}$ depended on $\rho(y)$ and hence the timing, duration and extent of habitat loss. For example, if the extent of habitat loss increased, then D_1 and hence $\rho(y)$ increased, such that $V(N(t))/\overline{N(t)}^2$ and $V(n_{l1}(t))/\overline{N(t)}^2$ increased (as per Eqs. (9) and (10)), with a corresponding change in $\overline{\pi_L(t)}$ (as per Eqs. (7) and (11)).

It can be shown (see Supplementary material) that if $\overline{N(t)} \gg 2V(n_{l1}(t))/\overline{n_{l1}(t)}$ for all l, then Eq. (11) can be simplified to

$$\overline{\pi_L(t)} \approx \frac{2}{L(N(0))^2}\sum_{l=1}^{L} n_{l1}(0)n_{l2}(0), \tag{12}$$

which was constant and equal to the initial value. The quantity $V(n_{l1}(t))/\overline{n_{l1}(t)}$ increased with t (see Supplementary material) due to genetic drift, such that the inequality $\overline{N(t)} \gg 2V(n_{l1}(t))/\overline{n_{l1}(t)}$ only held when t was sufficiently small (Fig. 3). For large t, the inequality ceased to hold and it can be shown (see Supplementary material) that $\overline{\pi_L(t)}$ was lower than its initial value, reflecting loss of nucleotide diversity due to genetic drift (Fig. 3).

Secondly, we derived formulae specifying the mean proportion of fixed loci at time t under the habitat loss scenarios considered. Again, let locus l have $n_{l1}(t)$ copies of the first allele and $n_{l2}(t)$ copies of the second allele at time t. Then following Kendall (1948) and Fung and Chisholm (2023), the probability of the number of copies of the first allele being zero at time t was

$$P(n_{l1}(t) = 0) = \left(1 - \frac{1}{1 + \int_0^t \sigma e^{\rho(y)} dy}\right)^{n_{l1}(0)}, \tag{13}$$

where $\sigma = D_0$ for $[0,t_1)$ and $[t_2,t_3]$, and $\sigma = D_1$ for $[t_1,t_2)$. We note that $P(n_{l1}(t)=0)$ depended on $\rho(y)$ and hence the timing, duration and extent of habitat loss. For example, if the extent of habitat loss increased, then D_1 and hence $\rho(y)$ increased, such that $P(n_{l1}(t)=0)$ increased (as per Eq. (13)). Analogously, $P(n_{l2}(t)=0)$ was given by Eq. (13) with $n_{l1}(0)$ replaced by $n_{l2}(0)$, and $P(n_{l1}(t)+n_{l2}(t)=0)$ was given by Eq. (13) with $n_{l1}(0)$ replaced by $n_{l1}(0)+ n_{l2}(0)$. Then using Bayes' Theorem, the probability that locus l was fixed at time t given that the population was still extant was $[P(n_{l1}(t)=0)(1-P(n_{l2}(t)=0))+P(n_{l2}(t)=0)(1-P(n_{l1}(t)=0))]/(1-P(n_{l1}(t)+n_{l2}(t)=0)) = f(n_{l10}(0),n_{l2}(0),t)$. Thus, the mean proportion of fixed loci at time t was

$$p_{\text{fixed}}(t) = \frac{1}{L}\sum_{l=1}^{L} f(n_{l10}, n_{l20}, t). \tag{14}$$

We see that $p_{\text{fixed}}(t)$ depended crucially on the extinction probabilities of alleles at time t, as given by Eq. (13). These probabilities were very low if the number of copies of each allele was high, which was often the case if the initial population size was large and t was sufficiently small, such that the mean population size at time t, $\overline{N(t)}$, was reasonably large and genetic drift was weak. Thus, like the mean nucleotide diversity $\overline{\pi_L(t)}$, $p_{\text{fixed}}(t)$ changed by little if $\overline{N(t)}$ was sufficiently large (Fig. 3). However, when habitat loss increased from 90 % to 99 %, $\overline{N(t)}$ decreased from 400 to a low value of 40, with a corresponding nonlinear increase in $p_{\text{fixed}}(t)$ (Fig. 3). This suggested that with respect to preserving genetic diversity, a population size on the order of hundreds of individuals at least is required. Given that genetic diversity is required for ensuring a high probability of survival of a population, this result is related to the idea of a minimum viable population (MVP; Shaffer, 1981). However, because MVPs depend a lot on context (Flather et al., 2011), further work is required for our result here to be used for estimating MVPs.

To confirm that our formulae were correct, we compared values of $\overline{N(t)}$, $\overline{\pi_L(t)}$ and $p_{\text{fixed}}(t)$ from our formulae with those from simulations, and found that there was very good agreement (Figs. 3, S1 and S3). In addition, we expect our formulae to give good approximations in the case of diploids with random mating and no linkage disequilibrium, because in this case allele dynamics at one chromosome in an individual are largely independent of allele dynamics at the other chromosome. We confirmed this expectation for the habitat loss scenarios we examined, using simulations with randomly mating diploids and no linkage disequilibrium (Fig. 3). The initial population size was set to half that in the haploid case, to ensure that the total number of chromosomes were comparable. The sex of each initial diploid individual was randomly assigned as male or female, as was the sex of each newborn individual. Birth and death events occurred randomly as per the haploid case, except that when an individual reproduced, it randomly chose an individual of the opposite sex for sexual reproduction. If all individuals were of the same sex, then reproduction could not occur.

3.2. Spatially explicit population model

Spatially explicit simulations of three stylized bird species under 18 different habitat loss scenarios demonstrated that the effects of habitat loss on genetic diversity were not as pronounced as on population abundance (Fig. 4). The population size of each species decreased almost immediately after habitat loss under any habitat loss scenario and was followed by increasing levels of inbreeding (*FIS*) and genetic structuring (Mantel *r*), whereas noticeable loss of genetic diversity (both nucleotide diversity and proportion of fixed loci) only manifested decades later, if at all. In terms of decreases in population size, species with a higher mobility were generally more sensitive to a higher proportional loss of habitat, lower carrying capacity of cleared habitat and decreasing spatial autocorrelation of the preserved habitat (Figs. 4 and S4). However, species with a higher mobility exhibited lower levels of inbreeding and genetic structuring after habitat loss (Figs. 2 and 4). For species of all mobility, genetic diversity loss was generally small for all the habitat loss scenarios and did not substantially deviate from the null scenario, regardless of whether the mean nucleotide diversity or the mean proportion of fixed loci was used as an indicator. Genetic diversity loss was most noticeable under the most extreme habitat loss scenarios: random habitat preservation with 50 % or 80 % of the area cleared, and with the carrying capacity of a cleared cell being 10 % that of a pristine habitat cell (Fig. 4).

We found that the proportion of total habitat cleared had a greater negative effect on population size than the spatial configuration of the preserved habitat (Fig. 4). For example, increasing the proportion of habitat cleared to 80 % typically reduced population size to below 1000, representing a decrease of >75 % of the original population size. In contrast, decreasing the spatial autocorrelation of the preserved habitat (moving from “clumped” to “random”) typically resulted in decreases in population size that were smaller, although the effects can still exceed 1000 individuals (Fig. 4). The proportion of habitat cleared had a nonlinear interaction with the spatial configuration of the preserved habitat in determining genetic diversity loss (Fig. 4). A large proportion of habitat cleared on its own was typically not sufficient for substantial genetic diversity loss. Rather, low spatial autocorrelation of the preserved habitat was often required as well. Compared with loss of genetic diversity when the preserved area was clumped, loss of genetic diversity when the preserved area was randomly chosen was up to 60 % higher, with the largest differences occurring under the most severe habitat loss scenarios of 50 % or 80 % of the area cleared, and the carrying capacity of a cleared cell being 10 % or 20 % that of a pristine habitat cell (Table S2 and Fig. S4).

Simulations for the malleefowl in South Australia were overall consistent with the simulations for the three stylized bird species, in the sense that population abundance and genetic diversity responded more negatively to a greater proportional loss of habitat cover and patterns of habitat preservation that were more random (Fig. 5). Comparing simulations with instant versus gradual habitat loss, population abundances decreased more gradually under gradual habitat loss but by the end of the 19th century were comparable to those under instant loss, whereas genetic diversity was similar under both types of habitat loss (Fig. 5).

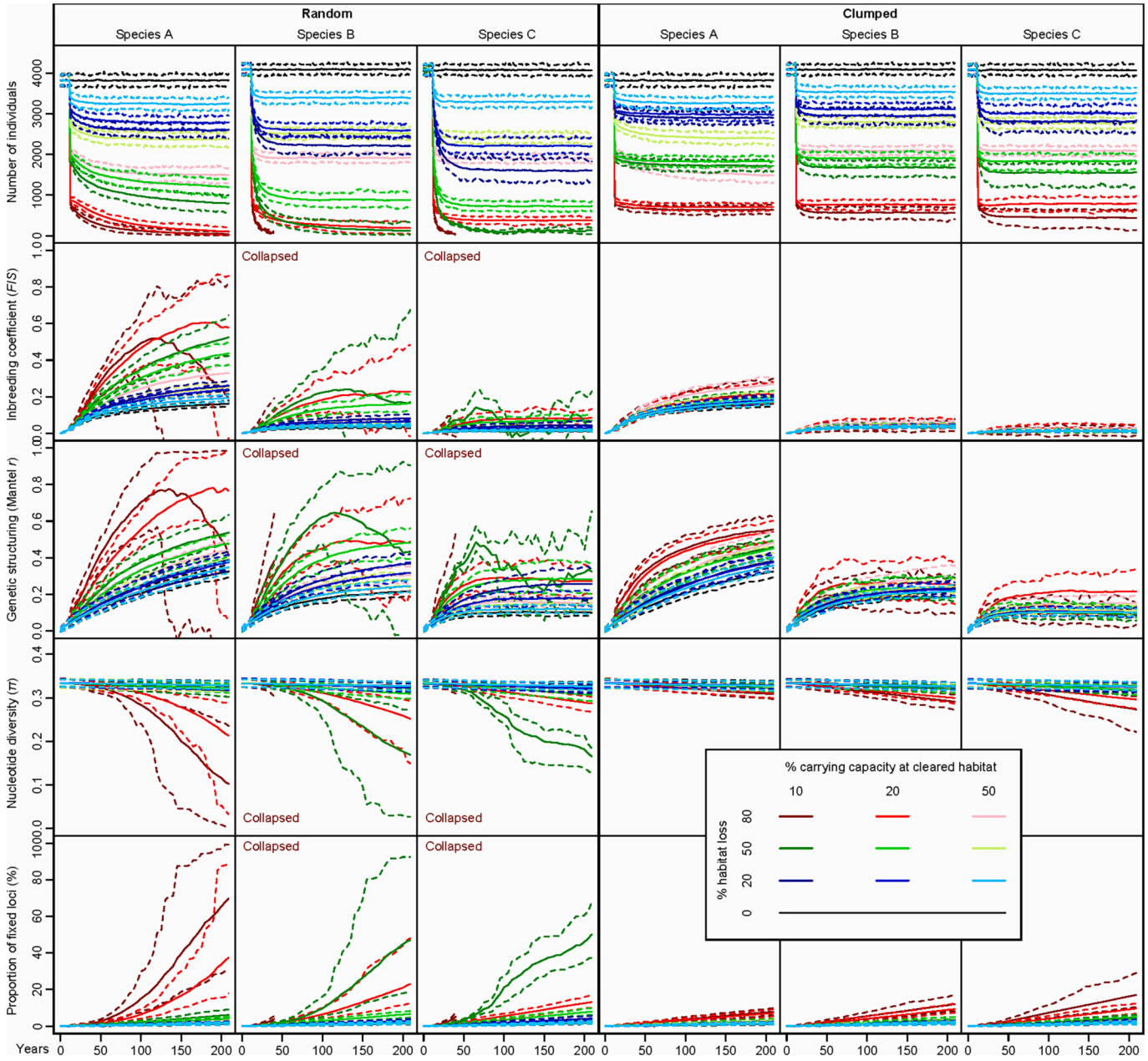


Fig. 4. Temporal changes in population size, inbreeding, genetic structuring and genetic diversity across 18 habitat loss scenarios from our spatially explicit simulations, for three species with different mobility. For comparison, results from a null scenario with no habitat loss are also shown. Genetic diversity was measured by nucleotide diversity and proportion of fixed loci. Species A, B, and C have a mean dispersal distance per year of 0.2 km, 0.6 km, and 1 km respectively. The different habitat loss scenarios were distinguished by the proportion of habitat that was lost, the spatial autocorrelation of the preserved area (either randomly chosen or clumped), and the carrying capacity of a cleared habitat cell expressed as a proportion of the carrying capacity of a pristine habitat cell. These three factors are represented in the figure as different colors, columns, and shades of color, respectively. For each scenario and quantity (population size, inbreeding coefficient, amount of genetic structuring or genetic diversity indicators), we show the mean (solid line) and middle 95 % of values (delineated by two dashed lines surrounding the solid line) from 100 stochastic simulations. For Species B and C, the population collapsed and went extinct (between the 40th to 60th year) in 98 % of simulations in the habitat loss scenario with 80 % of habitat lost, the preserved area being randomly chosen, and the carrying capacity of a cleared habitat cell being 10 % of its pristine value. Thus, the inbreeding coefficient, amount of genetic structuring and genetic diversity was only calculated for the first 30 years after the habitat loss in these cases, and we have marked these cases with "Collapsed" on the corresponding graphs.

Genetic diversity calculated using the real sampling scheme, which sampled 0.17 % to 1.69 % of the estimated total number of individuals across different transects, tended to be substantially lower than genetic diversity calculated using all individuals (complete sampling scheme), especially for the proportion of fixed loci. Empirical values of the proportion of fixed loci were generally consistent with simulated values under the real sampling scheme, whereas empirical values of nucleotide diversity tended to be somewhat lower than simulated values under the real sampling scheme (Fig. 5).

4. Discussion

Among all factors considered, the total amount of habitat loss had the greatest negative effect on the loss of genetic diversity, with the spatial

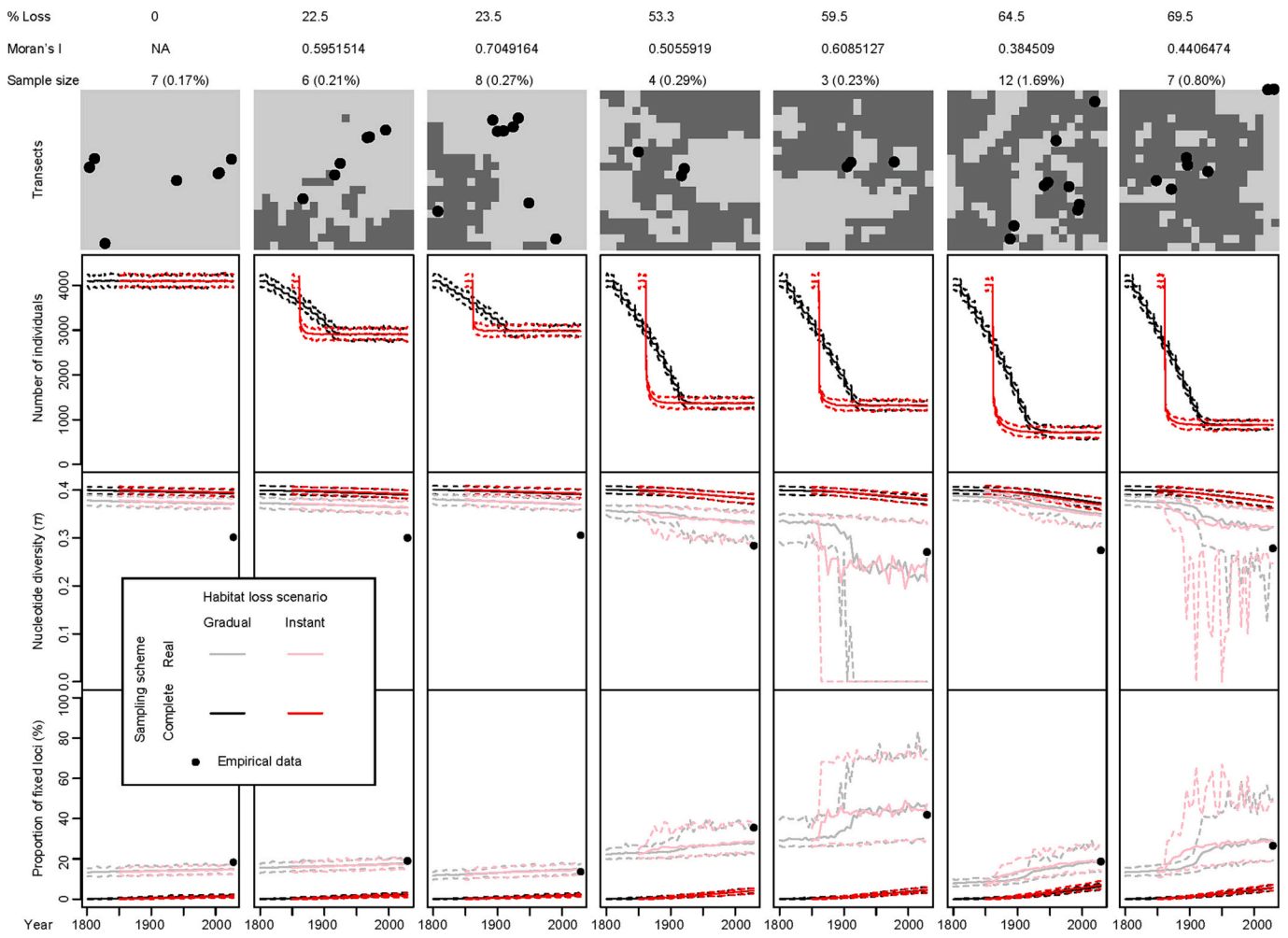


Fig. 5. Temporal changes in population size and genetic diversity of malleefowl across seven transects in South Australia under gradual and instant historical habitat loss. The top row shows current habitat maps for the seven transects, with light and dark gray cells representing pristine habitat and habitat that has been disturbed by anthropogenic activities, respectively. The black dots in each map indicate the locations of sampled individuals in an empirical study (Stenhouse et al., 2022). The proportional amount of habitat lost, the spatial pattern of habitat loss (as measured by Moran's I) and sample size (the percentage in brackets refers to sample size expressed as a percentage of the estimated abundance) are shown above each map. For each habitat loss scenario and quantity (population size or genetic diversity indicator), the mean (solid line) and middle 95 % of values (delineated by two dashed lines surrounding the solid line) from 100 stochastic simulations are shown.

configuration of habitat fragmentation exerting smaller but nonetheless noticeable negative effects. In contrast to population abundance, which generally declined proportionally with the amount of habitat lost, genetic diversity only exhibited considerable decreases in extreme cases when abundance was reduced substantially, corresponding to the most severe habitat loss scenarios associated with highly fragmented habitat configurations and large amounts of habitat cleared. This result complements previous simulation studies that examined long-term scenarios (Jackson and Fahrig, 2016, 2014), in the sense that these previous studies found that the amount of habitat loss and – to a lesser extent – the degree of fragmentation had negative effects on genetic diversity in the long term. However, our result is also in contrast to these previous studies in the sense that we found that over shorter timescales corresponding to non-equilibrium situations, the amount of habitat loss and the degree of fragmentation generally do not have pronounced negative effects on genetic diversity, with the exception of extreme scenarios in which the amount of habitat loss and the degree of fragmentation interacted non-linearly.

Our results and corresponding conclusions were based on a combination of analyses of a simple non-spatial mathematical model and a complex spatially explicit simulation model, in scenarios of habitat loss that caused genetic diversity to be non-equilibrium. The two models complemented each other in their strengths and weaknesses. The non-spatial model was simple enough to allow for the derivation of formulae for genetic diversity, which gave clear insights into how genetic diversity changed with habitat loss. However, the non-spatial model lacked biological realism, notably with no explicit representation of space. Nonetheless, results from the non-spatial model provided a useful baseline set of results, in particular in terms of identifying the critical dependence of loss of genetic diversity on low population sizes with strong genetic drift. The explicit representation of space in the spatially explicit model allowed for an investigation of how habitat loss resulted in clustering of individuals, which led to a spatial restructuring of genetic diversity and inbreeding. By comparing the non-spatial (Fig. 3) and spatial models (Fig. 4), we found that the spatial model exhibited generally higher rates of loss of genetic diversity after habitat loss, which indicated that inbreeding and genetic structuring acted at the subpopulation level to accelerate genetic drift.

Our results indicated a substantial time-lag between the reduction of population abundance and subsequent genetic diversity loss, typically on the order of decades (Fig. 4). This time-lag – sometimes referred to as genetic extinction debt (Plue et al., 2017; Vranckx et al., 2012) – is alarming because it implies a locked-in, long-term impact on a population's viability. We found that the time-lag was due to the slow rate at which genetic drift acted to reduce genetic diversity, and was mediated by inbreeding and genetic structuring at a subpopulation level (Figs. 3

and 4). Thus, management actions that rebuild genetic connectivity and prevent inbreeding, such as assisted migration (Fitzpatrick et al., 2023; Grummer et al., 2022) and managed breeding (Frankham, 2010), can mitigate the loss of genetic diversity within a time window of decades after habitat loss. It is possible to draw parallels here with research documenting how species richness is reduced by habitat loss via reductions in population abundances of species, which increase the probability of species extinctions via ecological drift (Thompson et al., 2019). Like genetic drift, ecological drift is a slow random process and hence only acts strongly at very small population abundances (Lande, 1993). Thus, loss of species richness is expected to occur slowly and the extinction debt due to habitat loss would take a long time to be cleared, on the order of hundreds of generations for tree communities (Thompson et al., 2019). These parallels between the underlying causes of extinction debt and genetic extinction debt emphasize the importance of conservation measures aimed at maintaining sufficiently high levels of population abundances.

We also found that the effects of habitat loss on genetic diversity in non-equilibrium scenarios depended on the mobility of species, with more-mobile species suffering greater reductions in genetic diversity. This outcome corroborates results presented in another recent simulation study, albeit at much longer timescales (Jackson and Fahrig, 2014). The underlying mechanism seems to be the same regardless of timescale: more-mobile species have a greater chance of dispersing from suitable to non-suitable habitat, thus reducing their population abundance and genetic diversity. In our simulations, we found that this mechanism operated more strongly in landscapes with randomly placed versus clumped suitable habitat, because individuals were more likely to disperse to unsuitable habitat that surrounded patches of randomly placed suitable habitat (Fig. 2). However, we note that this result depends on species being unable to choose where they disperse to, which may be unrealistic in many contexts. Species that are capable of detecting the quality of neighboring habitat would partially be able to mitigate the negative effects of dispersing into unsuitable habitat (Jackson and Fahrig, 2016). In addition, although less-mobile species had lower reductions in genetic diversity overall, they became spatially clumped into genetically homogeneous clusters within islands of remaining suitable habitat (Fig. 2). Isolated subpopulations, receiving little influx of migrants because of the species' low mobility, could be especially vulnerable to local extinction because of environmental perturbation or ecological drift – indeed, the latter appears to have caused local extinctions within some of the habitat islands in our simulations. Therefore, it will be important for future work, especially in real systems of conservation concern, to examine temporal trends in genetic diversity at multiple spatial scales.

In terms of the SLOSS debate (Diamond, 1975; Simberloff and Abele, 1982; Wilcox and Murphy, 1985), our results imply that a single large reserve is more beneficial than several smaller reserves of the same total area for the conservation of genetic diversity. All species in our spatial simulations exhibited larger population size in a single large reserve of contiguous pristine habitat, within a matrix of cleared habitat that had lower carrying capacity (Figs. 4 and S4). Moreover, within the time frame of our simulations of 210 years, all species exhibited a smaller loss of genetic diversity when there was a single large reserve, especially for species of high mobility under the more severe habitat loss scenarios (Table S2 and Fig. S4). Although species of low mobility generally exhibited smaller gains in genetic diversity in a single large reserve, they often exhibited substantially lower levels of inbreeding and genetic structuring (Fig. 4), which may help prevent considerable loss of genetic diversity in the long run (beyond the timescale of our simulations). We caution that our study is at a spatial scale with thousands of individuals, and thus our suggestion regarding the SLOSS debate may not be applicable at larger scales (e.g., country-wide scale), with hundreds of thousands of individuals that may have adapted to differences in environmental conditions over a large geographical range. We also caution that our results were based on models parameterized for bird species, and hence our results should not be simply generalized to other taxa, especially those that are unable to fly and hence have much lower capacity for dispersal (e.g., herptiles). We encourage future studies that examine loss of genetic diversity under habitat loss for other taxa, which together with our results could then be used to inform conservation of species richness in whole communities of multiple taxa.

Our simulation results for the malleefowl case study in South Australia followed the general trends described for the stylized bird species. Thus, simulations under a complete sampling scheme suggested that loss of genetic diversity was greatest for the transect with the greatest amount of habitat loss and degree of fragmentation, and genetic diversity loss occurred slowly after relatively rapid declines in population abundance (Fig. 5). However, by comparing results using simulated data under a real sampling scheme, which reflected how empirical data (Stenhouse et al., 2022) was collected, with results using simulated data under a complete sampling scheme, we found that sampling a small fraction of the population created substantial uncertainty in estimates of genetic diversity. This sampling uncertainty was strong enough to blur the effects of habitat loss. In particular, the sixth and seventh transects had the greatest amount of habitat loss and degree of fragmentation, which led to the greatest loss of genetic diversity in the populations as reflected in the simulated data under a complete sampling scheme (Fig. 5), yet sample genetic diversity was lowest in the fourth and fifth transects as reflected in the simulated data under a real sampling scheme, because of sampling uncertainty from small sample sizes. Given that empirical studies of genetic diversity often use small sample sizes with relatively uneven spatial distributions, mathematical corrections of sampling bias can be helpful (Bashalkhanov et al., 2009). However, such corrections may not incorporate important factors for accurate estimation, such as species-specific relationships between the number of genetic variants and area (Exposito-Alonso et al., 2022). In our case study, there was generally good agreement between simulated (under a real sampling scheme) and empirical values of proportion of fixed loci, but simulated values of nucleotide diversity were substantially higher than empirical values in four transects. A possible reason was our use of random initial allele frequencies for simulated loci, which does not fully reflect the actual allele frequencies of malleefowl a couple of centuries ago. Another possible reason was our use of simplified habitat loss scenarios in simulations (either instantaneous or gradual loss at a constant rate), which does not fully reflect the real way in which habitat was lost in South Australia. Overall, we recommend interpreting conservation genetic data obtained under partial sampling schemes with caution. Moreover, our simulations in the malleefowl case study suggest that the use of spatially explicit forward-time simulations, as implemented in software packages such as *Geonomics* (Terasaki Hart et al., 2021), is a promising avenue for the correction of sampling bias while accounting for spatial unevenness of populations.

Our study has examined genetic diversity indicators that were calculated using neutral genetic markers, which follow neutral evolutionary processes. This practice is consistent with previous studies (Jackson and Fahrig, 2016, 2014) and hence facilitates comparison with these studies. On the other hand, selection – especially linked background selection – can have a pronounced footprint across large portions of the genome (Pouyet et al., 2018), which may help mitigate loss of genetic diversity (Exposito-Alonso et al., 2022). Therefore, simulations purely based on neutral genetic markers may overestimate the total rate of genetic diversity loss. The use of neutral genetic markers as a good indicator of overall evolutionary potential has become uncertain under the recent debate over the neutral gene theory (Jensen et al., 2019; Kardos et al., 2021; Teixeira and Huber, 2021). However, accounting for adaptive genetic markers in simulations is challenging because of their specificity to certain organisms and environments (e. g. Tournebize et al., 2022). The incorporation of adaptive genetic markers and corresponding genomic linkage information could form part of the next great frontier in future simulation studies on diversity loss.

5. Conclusion

Using non-spatial modelling and spatial simulation, our study captures the transient dynamics of genetic diversity at non-equilibrium stages after habitat loss. Our results are consistent with the following sequence of events leading to a loss of genetic diversity: loss of habitat ➔ loss of individuals ➔ elevated levels of inbreeding and population spatial structuring ➔ loss of genetic diversity. This sequence of events is associated with a time lag between habitat loss and genetic diversity loss, which could sometimes afford a window of opportunity for conservation measures, such as assisted migration and managed breeding, to counter the effects of inbreeding. Our spatial simulation explores combinations of different amounts and configurations of habitat loss to demonstrate that "single large" reserves are overall more suitable for conserving genetic diversity than "several small" ones in terms of the SLOSS debates. By comparing empirical data with results of simulations, our study reveals that spatial modelling is a promising avenue in correcting estimates of genetic diversity that may be subject to sampling artifacts.

Declaration of competing interest

The authors declare no competing interests.

Data availability

A guide to our simulation analyses is deposited on GitHub: https://github.com/qt37t247/HabitatLossGenetics.

Acknowledgements

Q. T. and F. E. R. were supported by the Ministry of Education - Singapore under their MOE Tier 2 grant (grant number WBS A-0004777-00-00), whereas T. F. was supported by the Ministry of Education - Singapore under another MOE Tier 2 grant (grant number WBS A-8001046-00-00). A portion of the time D. E. T. H. spent on this work was funded by a Bezos Earth Fund grant to The Nature Conservancy. All authors are grateful to the editor and three anonymous reviewers for their insightful and constructive suggestions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biocon.2023.110381.

References

Bashalkhanov, S., Pandey, M., Rajora, O.P., 2009. A simple method for estimating genetic diversity in large populations from finite sample sizes. BMC Genet. 10, 84. https://doi.org/10.1186/1471-2156-10-84.

Benestan, L.M., Ferchaud, A.-L., Hohenlohe, P.A., Garner, B.A., Naylor, G.J.P., Baums, I. B., Schwartz, M.K., Kelley, J.L., Luikart, G., 2016. Conservation genomics of natural and managed populations: building a conceptual and practical framework. Mol. Ecol. 25, 2967–2977. https://doi.org/10.1111/mec.13647.

Bergeron, L.A., Besenbacher, S., Zheng, J., Li, P., Bertelsen, M.F., Quintard, B., Hoffman, J.I., Li, Z., St. Leger, J., Shao, C., 2023. Evolution of the germline mutation rate across vertebrates. Nature 615, 285–291.

BirdLife International, 2023. Species factsheet: *Leipoa ocellata*. Downloaded from. http://datazone.birdlife.org/species/factsheet/malleefowl-leipoa-ocellata. (Accessed 11 March 2023).

Bradshaw, C.J., 2012. Little left to lose: deforestation and forest degradation in Australia since European colonization. J. Plant Ecol. 5, 109–120.

Cabeza, M., 2003. Habitat loss and connectivity of reserve networks in probability approaches to reserve design. Ecol. Lett. 6, 665–672.

Claramunt, S., 2021. Flight efficiency explains differences in natal dispersal distances in birds. Ecology 102 (9), e03442.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., 2011. The variant call format and VCFtools. Bioinformatics 27, 2156–2158.

Department of Parks and Wildlife, 2016. Fauna profiles: Malleefowl Leipoa ocellata. Retrieved from: http://www.dpaw.wa.gov.au/.

DeWoody, J.A., Harder, A.M., Mathur, S., Willoughby, J.R., 2021. The long-standing significance of genetic diversity in conservation. Mol. Ecol. 30, 4147–4154. https://doi.org/10.1111/mec.16051.

Diamond, J.M., 1975. The island dilemma: lessons of modern biogeographic studies for the design of natural reserves. Biol. Conserv. 7, 129–146.

Díaz, S., Settele, J., Brondízio, E.S., Ngo, H.T., Agard, J., Arneth, A., Balvanera, P., Brauman, K.A., Butchart, S.H., Chan, K.M., 2019. Pervasive human-driven decline of life on Earth points to the need for transformative change. Science 366, eaax3100.

Díaz, S., Zafra-Calvo, N., Purvis, A., Verburg, P.H., Obura, D., Leadley, P., Chaplin-Kramer, R., De Meester, L., Dulloo, E., Martín-López, B., Shaw, M.R., Visconti, P., Broadgate, W., Bruford, M.W., Burgess, N.D., Cavender-Bares, J., DeClerck, F., Fernández-Palacios, J.M., Garibaldi, L.A., Hill, S.L.L., Isbell, F., Khoury, C.K., Krug, C.B., Liu, J., Maron, M., McGowan, P.J.K., Pereira, H.M., Reyes-García, V., Rocha, J., Rondinini, C., Shannon, L., Shin, Y.-J., Snelgrove, P.V.R., Spehn, E.M., Strassburg, B., Subramanian, S.M., Tewksbury, J.J., Watson, J.E.M., Zanne, A.E., 2020. Set ambitious goals for biodiversity and sustainability. Science 370, 411–413. https://doi.org/10.1126/science.abe1530.

Dixon, P., 2003. VEGAN, a package of R functions for community ecology. J. Veg. Sci. 14, 927–930. https://doi.org/10.1111/j.1654-1103.2003.tb02228.x.

Ewens, W.J., 2004. Mathematical Population Genetics, Interdisciplinary Applied Mathematics. Springer New York, New York, NY. https://doi.org/10.1007/978-0-387-21822-9.

Exposito-Alonso, M., Booker, T.R., Czech, L., Gillespie, L., Hateley, S., Kyriazis, C.C., Lang, P.L.M., Leventhal, L., Nogues-Bravo, D., Pagowski, V., Ruffley, M., Spence, J. P., Toro Arana, S.E., Weiß, C.L., Zess, E., 2022. Genetic diversity loss in the Anthropocene. Science 377, 1431–1435. https://doi.org/10.1126/science.abn5642.

Fahrig, L., 2017. Ecological responses to habitat fragmentation per se. Annu. Rev. Ecol. Evol. Syst. 48, 1–23. https://doi.org/10.1146/annurev-ecolsys-110316-022612.

Fisher, R.A., 1923. XXI.—on the dominance ratio. Proc. R. Soc. Edinb. 42, 321–341.

Fitzpatrick, S.W., Mittan-Moreau, C., Miller, M., Judson, J.M., 2023. Genetic rescue remains underused for aiding recovery of federally listed vertebrates in the United States. J. Hered. 114 (4), 354–366 esad002.

Flather, C.H., Hayward, G.D., Beissinger, S.R., Stephens, P.A., 2011. Minimum viable populations: is there a 'magic number' for conservation practitioners? Trends Ecol. Evol. 26, 307–316.

Fletcher, R.J., Didham, R.K., Banks-Leite, C., Barlow, J., Ewers, R.M., Rosindell, J., Holt, R.D., Gonzalez, A., Pardini, R., Damschen, E.I., Melo, F.P.L., Ries, L., Prevedello, J.A., Tscharntke, T., Laurance, W.F., Lovejoy, T., Haddad, N.M., 2018. Is habitat fragmentation good for biodiversity? Biol. Conserv. 226, 9–15. https://doi.org/10.1016/j.biocon.2018.07.022.

Frankham, R., 2010. Challenges and opportunities of genetic approaches to biological conservation. Biol. Conserv. 143, 1919–1927. https://doi.org/10.1016/j.biocon.2010.05.011.

Franklin, I.R., 1980. Evolutionary change in small populations. Sinauer Associates, pp. 135–149.

Frith, H.J., 1962. Conservation of the mallee fowl, Leipoa ocellata Gould (Megapodiidae). CSIRO Wildlife Res. 7, 33–49.

Fung, T., Chisholm, R.A., 2023. Improving the realism of neutral ecological models by incorporating transient dynamics with temporal changes in community size. Theor. Popul. Biol. 149, 12–26.

Gaston, K.J., Jackson, S.F., Cantú-Salazar, L., Cruz-Piñón, G., 2008. The ecological performance of protected areas. Annu. Rev. Ecol. Evol. Syst. 39, 93–113.

Grummer, J.A., Booker, T.R., Matthey-Doret, R., Nietlisbach, P., Thomaz, A.T., Whitlock, M.C., 2022. The immediate costs and long-term benefits of assisted gene flow in large populations. Conserv. Biol. 36, e13911 https://doi.org/10.1111/cobi.13911.

Harrisson, K.A., Pavlova, A., Telonis-Scott, M., Sunnucks, P., 2014. Using genomics to characterize evolutionary potential for conservation of wild populations. Evol. Appl. 7, 1008–1025. https://doi.org/10.1111/eva.12149.

Hawkes, C., 2009. Linking movement behaviour, dispersal and population processes: is individual variation a key? J. Anim. Ecol. 78, 894–906.

Hedrick, P.W., 1994. Purging inbreeding depression and the probability of extinction: full-sib mating. Heredity 73, 363–372.

Jackson, N.D., Fahrig, L., 2014. Landscape context affects genetic diversity at a much larger spatial extent than population abundance. Ecology 95, 871–881.

Jackson, N.D., Fahrig, L., 2016. Habitat amount, not habitat configuration, best predicts population genetic structure in fragmented landscapes. Landsc. Ecol. 31, 951–968. https://doi.org/10.1007/s10980-015-0313-2.

Jensen, J.D., Payseur, B.A., Stephan, W., Aquadro, C.F., Lynch, M., Charlesworth, D., Charlesworth, B., 2019. The importance of the neutral theory in 1968 and 50 years on: a response to Kern and Hahn 2018. Evolution 73, 111–114.

Kardos, M., Armstrong, E.E., Fitzpatrick, S.W., Hauser, S., Hedrick, P.W., Miller, J.M., Tallmon, D.A., Funk, W.C., 2021. The crucial role of genome-wide genetic variation in conservation. Proc. Natl. Acad. Sci. 118, e2104642118 https://doi.org/10.1073/pnas.2104642118.

Kendall, D.G., 1948. On the generalized" birth-and-death" process. Ann. Math. Stat. 19, 1–15.

Kim, T., Cho, S.-H., Larson, E.R., Armsworth, P.R., 2014. Protected area acquisition costs show economies of scale with area. Ecol. Econ. 107, 122–132.

Kimura, M., 1955. Solution of a process of random genetic drift with a continuous model. Proc. Natl. Acad. Sci. U. S. A. 41, 144–150. https://doi.org/10.1073/pnas.41.3.144.

Lande, R., 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. Am. Nat. 142, 911–927.

Lande, R., Shannon, S., 1996. The role of genetic variation in adaptation and population persistence in a changing environment. Evolution 434–437.

Lerouzic, A., Carlborg, O., 2008. Evolutionary potential of hidden genetic variation. Trends Ecol. Evol. 23, 33–37. https://doi.org/10.1016/j.tree.2007.09.014.
Lewis, E., MacSharry, B., Juffe-Bignoli, D., Harris, N., Burrows, G., Kingston, N., Burgess, N.D., 2019. Dynamics in the global protected-area estate since 2004. Conserv. Biol. 33, 570–579.
Martin, A.E., Fahrig, L., 2018. Habitat specialist birds disperse farther and are more migratory than habitat generalist birds. Ecology 99, 2058–2066. https://doi.org/10.1002/ecy.2428.
Minias, P., Podlaszczuk, P., 2017. Longevity is associated with relative brain size in birds. Ecol. Evol. 7, 3558–3566.
Miraldo, A., Li, S., Borregaard, M.K., Flórez-Rodríguez, A., Gopalakrishnan, S., Rizvanovic, M., Wang, Z., Rahbek, C., Marske, K.A., Nogués-Bravo, D., 2016. An Anthropocene map of genetic diversity. Science 353, 1532–1535. https://doi.org/10.1126/science.aaf4381.
Moran, P.A.P., 1958. Random processes in genetics. In: Mathematical Proceedings of the Cambridge Philosophical Society. Cambridge University Press, pp. 60–71.
Nei, M., Li, W.-H., 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. 76, 5269–5273.
Pérez-Pereira, N., Wang, J., Quesada, H., Caballero, A., 2022. Prediction of the minimum effective size of a population viable in the long term. Biodivers. Conserv. 31, 2763–2780. https://doi.org/10.1007/s10531-022-02456-z.
Plue, J., Vandepitte, K., Honnay, O., Cousins, S.A., 2017. Does the seed bank contribute to the build-up of a genetic extinction debt in the grassland perennial Campanula rotundifolia? Ann. Bot. 120, 373–385.
Pouyet, F., Aeschbacher, S., Thiéry, A., Excoffier, L., 2018. Background selection and biased gene conversion affect more than 95% of the human genome and bias demographic inferences. Elife 7, e36317.
Shaffer, M.L., 1981. Minimum population sizes for species conservation. BioScience 31, 131–134.
Simberloff, D., Abele, L.G., 1982. Refuge design and island biogeographic theory: effects of fragmentation. Am. Nat. 120, 41–50.
Stenhouse, P., Moseby, K.E., 2023. Patch size and breeding status influence movement patterns in the threatened Malleefowl (Leipoa ocellata). Austral Ecol. 48, 904–927.
Stenhouse, P., Onley, I.R., Mitchell, K.J., Moseby, K.E., Austin, J.J., 2022. Spatial genetic structure and limited gene flow in fragmented populations of the threatened Malleefowl (Leipoa ocellata). Ecol. Genet. Genom. 24, 100127.
Stephens, P.A., Vieira, M.V., Willis, S.G., Carbone, C., 2019. The limits to population density in birds and mammals. Ecol. Lett. 22, 654–663. https://doi.org/10.1111/ele.13227.
Symes, W.S., Edwards, D.P., Miettinen, J., Rheindt, F.E., Carrasco, L.R., 2018. Combined impacts of deforestation and wildlife trade on tropical biodiversity are severely underestimated. Nat. Commun. 9, 4052.
Teixeira, J.C., Huber, C.D., 2021. The inflated significance of neutral genetic diversity in conservation genetics. Proc. Natl. Acad. Sci. 118, e2015096118.
Terasaki Hart, D.E., Bishop, A.P., Wang, I.J., 2021. Geonomics: forward-time, spatially explicit, and arbitrarily complex landscape genomic simulations. Mol. Biol. Evol. 38, 4634–4646.
Thompson, S.E., Chisholm, R.A., Rosindell, J., 2019. Characterising extinction debt following habitat fragmentation using neutral theory. Ecol. Lett. 22, 2087–2096.
Tournebize, R., Borner, L., Manel, S., Meynard, C.N., Vigouroux, Y., Crouzillat, D., Fournier, C., Kassam, M., Descombes, P., Tranchant-Dubreuil, C., 2022. Ecological and genomic vulnerability to climate change across native populations of Robusta coffee (Coffea canephora). Glob. Chang. Biol. 28, 4124–4142.
Volenec, Z.M., Dobson, A.P., 2020. Conservation value of small reserves. Conserv. Biol. 34, 66–79. https://doi.org/10.1111/cobi.13308.
Volkov, I., Banavar, J.R., Hubbell, S.P., Maritan, A., 2003. Neutral theory and relative species abundance in ecology. Nature 424, 1035–1037.
Vranckx, G.U.Y., Jacquemyn, H., Muys, B., Honnay, O., 2012. Meta-analysis of susceptibility of woody plants to loss of genetic diversity through habitat fragmentation. Conserv. Biol. 26, 228–237.
Weathers, W.W., Weathers, D.L., Seymour, R.S., 1990. Polygyny and reproductive effort in the Malleefowl Leipoa ocellata. Emu 90, 1–6.
Wilcox, B.A., Murphy, D.D., 1985. Conservation strategy: the effects of fragmentation on extinction. Am. Nat. 125, 879–887.
Wright, S., 1931. Evolution in Mendelian populations. Genetics 16, 97.

Supplementary Material for “Rate and extent of genetic diversity loss under non-equilibrium scenarios of habitat loss”

Qian Tang[1], Tak Fung[1], Drew E. Terasaki Hart[2,3], Frank E. Rheindt[1]

[1] National University of Singapore, Singapore

[2] The Nature Conservancy, USA

[3] University of California, Berkeley, USA

Further results for non-spatial model

Deriving formulae specifying mean nucleotide diversity

We first present details of how we derived formulae specifying mean nucleotide diversity for the non-spatial model that we used. As shown in the main text, considering just one biallelic locus l, the mean nucleotide diversity for the model is given by

$$\overline{\pi_l(t)}=2\left(\overline{\frac{n_{l1}(t)}{N(t)}}-\overline{\left(\frac{n_{l1}(t)}{N(t)}\right)^2}\right)=2\left(\overline{\frac{n_{l1}(t)}{N(t)}}-\left(V\left(\frac{n_{l1}(t)}{N(t)}\right)+\left(\overline{\frac{n_{l1}(t)}{N(t)}}\right)^2\right)\right), \quad \text{(S1)}$$

where $n_{l1}(t)$ is the number of copies of the first allele at time t; $N(t)$ is the number of copies of both alleles at time t, which is also equal to the population size at time t (because the population consists of haploid individuals); and V refers to the variance arising from genetic drift. By performing a second-order Taylor series expansion of $n_{l1}(t)/N(t)$ about the mean values $\overline{n_{l1}(t)}$ and $\overline{N(t)}$, and then taking the mean and variance, we get

$$\overline{\frac{n_{l1}(t)}{N(t)}}\approx\frac{\overline{n_{l1}(t)}}{\overline{N(t)}}-\frac{Cov\left(n_{l1}(t),N(t)\right)}{\left(\overline{N(t)}\right)^2}+\frac{\overline{n_{l1}(t)}V\left(N(t)\right)}{\left(\overline{N(t)}\right)^3}, \quad \text{(S2)}$$

$$V\left(\overline{\frac{n_{l1}(t)}{N(t)}}\right)\approx\frac{V\left(n_{l1}(t)\right)}{\left(\overline{N(t)}\right)^2}-2\frac{\overline{n_{l1}(t)}Cov\left(n_{l1}(t),N(t)\right)}{\left(\overline{N(t)}\right)^3}+\frac{\left(\overline{n_{l1}(t)}\right)^2V\left(N(t)\right)}{\left(\overline{N(t)}\right)^4}, \quad \text{(S3)}$$

where Cov refers to the covariance. Noting that $N(t)=n_{l1}(t)+n_{l2}(t)$, where $n_{l2}(t)$ is the number of copies of the second allele at time t, and that $n_1(t)$ and $n_2(t)$ are uncorrelated, we have

$$Cov\left(n_{l1}(t),N(t)\right)=Cov\left(n_{l1}(t),n_{l1}(t)+n_{l2}(t)\right)=Cov\left(n_{l1}(t),n_{l1}(t)\right)+Cov\left(n_{l1}(t),n_{l2}(t)\right)=V\left(n_{l1}(t)\right). \quad \text{(S4)}$$

Using (S4) in (S2) and (S3) gives

$$\frac{\overline{n_{1l}(t)}}{N(t)} \approx \frac{\overline{n_{l1}(t)}}{\overline{N(t)}} - \frac{V\left(n_{l1}(t)\right)}{\left(\overline{N(t)}\right)^2} + \frac{\overline{n_{l1}(t)}}{\overline{N(t)}} \frac{V\left(N(t)\right)}{\left(\overline{N(t)}\right)^2} = \frac{\overline{n_{l1}(t)}}{\overline{N(t)}} \left(1 + \frac{V\left(N(t)\right)}{\left(\overline{N(t)}\right)^2}\right) - \frac{V\left(n_{l1}(t)\right)}{\left(\overline{N(t)}\right)^2}, \quad \text{(S5)}$$

$$V\left(\frac{\overline{n_{l1}(t)}}{N(t)}\right) \approx \frac{V\left(n_{l1}(t)\right)}{\left(\overline{N(t)}\right)^2} - 2\frac{\overline{n_{l1}(t)}}{\overline{N(t)}} \frac{V\left(n_{l1}(t)\right)}{\left(\overline{N(t)}\right)^2} + \left(\frac{\overline{n_{l1}(t)}}{\overline{N(t)}}\right)^2 \frac{V\left(N(t)\right)}{\left(\overline{N(t)}\right)^2} = \frac{V\left(n_{l1}(t)\right)}{\left(\overline{N(t)}\right)^2} \left(1 - 2\frac{\overline{n_{l1}(t)}}{\overline{N(t)}}\right) + \left(\frac{\overline{n_{l1}(t)}}{\overline{N(t)}}\right)^2 \frac{V\left(N(t)\right)}{\left(\overline{N(t)}\right)^2}, \quad \text{(S6)}$$

which are (6) and (7) in the main text. The last expressions in (S5) and (S6) depend on three factors that are given explicitly by the following formulae, following Kendall (1948) and Fung and Chisholm (2023):

$$\frac{\overline{n_{l1}(t)}}{\overline{N(t)}} = \frac{n_{l1}(0)e^{-\rho(t)}}{N(0)e^{-\rho(t)}} = \frac{n_{l1}(0)}{N(0)}, \quad \text{(S7)}$$

$$\frac{V\left(N(t)\right)}{\left(\overline{N(t)}\right)^2} = \frac{N(0)e^{-2\rho(t)} \int_0^t e^{\rho(y)}\left(B + D(y)\right)dy}{\left(N(0)e^{-\rho(t)}\right)^2} = \frac{\int_0^t e^{\rho(y)}\left(B + D(y)\right)dy}{N(0)}, \quad \text{(S8)}$$

$$\frac{V\left(n_{l1}(t)\right)}{\left(\overline{N(t)}\right)^2} = \frac{n_{l1}(0)e^{-2\rho(t)} \int_0^t e^{\rho(y)}\left(B + D(y)\right)dy}{\left(N(0)e^{-\rho(t)}\right)^2} = \frac{n_{l1}(0)}{\left(N(0)\right)^2} \int_0^t e^{\rho(y)}\left(B + D(y)\right)dy, \quad \text{(S9)}$$

where B, $D(y)$ and $\rho(y)$ are as defined in the main text. (S7)-(S9) are equivalent to (8)-(10) in the main text.

From (S1) and (S7), if $V\left(n_{l1}(t)/N(t)\right) \ll \overline{n_{l1}(t)/N(t)}$, then

$$\overline{\pi_l(t)} = 2\left(\frac{\overline{n_{l1}(t)}}{N(t)} - \left(V\left(\frac{n_{l1}(t)}{N(t)}\right) + \left(\frac{\overline{n_{l1}(t)}}{N(t)}\right)^2\right)\right) \approx 2\left(\frac{\overline{n_{l1}(t)}}{N(t)} - \left(\frac{\overline{n_{l1}(t)}}{N(t)}\right)^2\right) = 2\left(\frac{n_{l1}(0)}{N(0)} - \left(\frac{n_{l1}(0)}{N(0)}\right)^2\right) = 2\frac{n_{l1}(0)}{N(0)}\frac{N(0) - n_{l1}(0)}{N(0)} = 2\frac{n}$$

Thus, if there are L loci and $V\left(n_{l1}(t)/N(t)\right) \ll \overline{n_{l1}(t)/N(t)}$ for all $1 \le l \le L$, then the mean nucleotide diversity over L loci is

$$\overline{\pi_L(t)}=\frac{1}{L}\sum_{l=1}^{L}\overline{\pi_l(t)}\approx\frac{2}{L\left(N(0)\right)^2}\sum_{l=1}^{L}n_{l1}(0)n_{l2}(0),\text{(S11)}$$

as per (12) in the main text. From (S5) and (S6), the inequality

$V\left(n_{l1}(t)/N(t)\right)\ll\overline{n_{l1}(t)/N(t)}$ is approximately

$$\frac{V\left(n_{l1}(t)\right)}{\left(\overline{N(t)}\right)^2}\left(1-2\frac{\overline{n_{l1}(t)}}{\overline{N(t)}}\right)+\left(\frac{\overline{n_{l1}(t)}}{\overline{N(t)}}\right)^2\frac{V\left(N(t)\right)}{\left(\overline{N(t)}\right)^2}$$

$$\ll\frac{\overline{n_{l1}(t)}}{\overline{N(t)}}\left(1+\frac{V\left(N(t)\right)}{\left(\overline{N(t)}\right)^2}\right)-\frac{V\left(n_{l1}(t)\right)}{\left(\overline{N(t)}\right)^2}.\text{(S12)}$$

Rearranging gives

$$2\frac{V\left(n_{l1}(t)\right)}{\left(\overline{N(t)}\right)^2}\left(1-\frac{\overline{n_{l1}(t)}}{\overline{N(t)}}\right)$$

$$\ll\frac{\overline{n_{l1}(t)}}{\overline{N(t)}}\left(1+\frac{V\left(N(t)\right)}{\left(\overline{N(t)}\right)^2}\right)-\left(\frac{\overline{n_{l1}(t)}}{\overline{N(t)}}\right)^2\frac{V\left(N(t)\right)}{\left(\overline{N(t)}\right)^2}.\text{(S13)}$$

A stricter inequality is

$$2\frac{V\left(n_{l1}(t)\right)}{\left(\overline{N(t)}\right)^2}\ll\frac{\overline{n_{l1}(t)}}{\overline{N(t)}}.\text{(S14)}$$

If (S14) holds, then so does (S13). (S14) is equivalent to

$$2\frac{V\left(n_{l1}(t)\right)}{\overline{n_{l1}(t)}}\ll\overline{N(t)}.\text{(S15)}$$

Explicitly, following Kendall (1948) and Fung and Chisholm (2023):

$$\frac{V\left(n_{l1}(t)\right)}{\overline{n_{l1}(t)}}=\frac{n_{l1}(0)e^{-\rho(t)}\int_0^t e^{\rho(y)}\left(B+D(y)\right)dy}{n_{l1}(0)e^{-\rho(t)}}=\int_0^t e^{\rho(y)}\left(B+D(y)\right)dy,\text{(S16)}$$

$$\overline{N(t)}=N(0)e^{-\rho(t)}.\text{(S17)}$$

From (S16), $V\left(n_{l1}(t)\right)/\overline{n_{l1}(t)}$ increases with t. Thus, (S15) only holds for sufficiently small t. When the inequality fails to hold, then the variance term in (S1) becomes non-negligible and reduces $\overline{\pi_L(t)}$.

Verifying accuracy of formulae

For the (stochastic) non-spatial model, we derived formulae specifying $\overline{N(t)}$ (as per (S17)); $\overline{\pi_L(t)}$ (as per (11) in the main text); and the mean proportion of fixed loci at time t, $p_{¿}(t)$ (as per (14) in the main text). To verify the accuracy of these formulae, we compared values from the formulae with those from simulations, for the three habitat loss scenarios that we examined and the baseline scenario with no habitat loss (see main text). For each scenario, the model was simulated 1,000 times and values of $\overline{N(t)}$, $\overline{\pi_L(t)}$ and $p_{¿}(t)$ calculated for every value of t that is a multiple of 0.01 yr. The model was simulated according to the schematic diagram shown in Fig. 1, with a timestep of $\delta t = 0.01$ yr during the time periods when no habitat loss occurred (i.e., $t < t_1 = 10\,yr$ and $t \geq t_2 = 11\,yr$) and a smaller timestep of $\delta t = 0.001$yr during the time period when habitat loss occurred in the habitat loss scenarios (i.e., $10\,yr = t_1 \leq t < t_2 = 11\,yr$), to account for the higher mortality rates during this time period. There was very good agreement between values from formulae and simulations (Fig. S1). Let the percentage absolute error between a value from a formula and a corresponding value from simulations be the absolute difference between the values divided by the value from the simulations, expressed as a percentage. Then the mean percentage absolute error for $\overline{N(t)}$ across all values of t was only 0.052–0.809% for the four scenarios. Similarly, the mean absolute percentage error for $\overline{\pi_L(t)}$ was only 0.001–0.114%. For $p_{¿}(t)$, the simulated values were often zero, such that the percentage absolute error was often undefined. Thus, we considered the absolute error instead. The mean absolute error for $p_{¿}(t)$ was only 0.005–0.059.

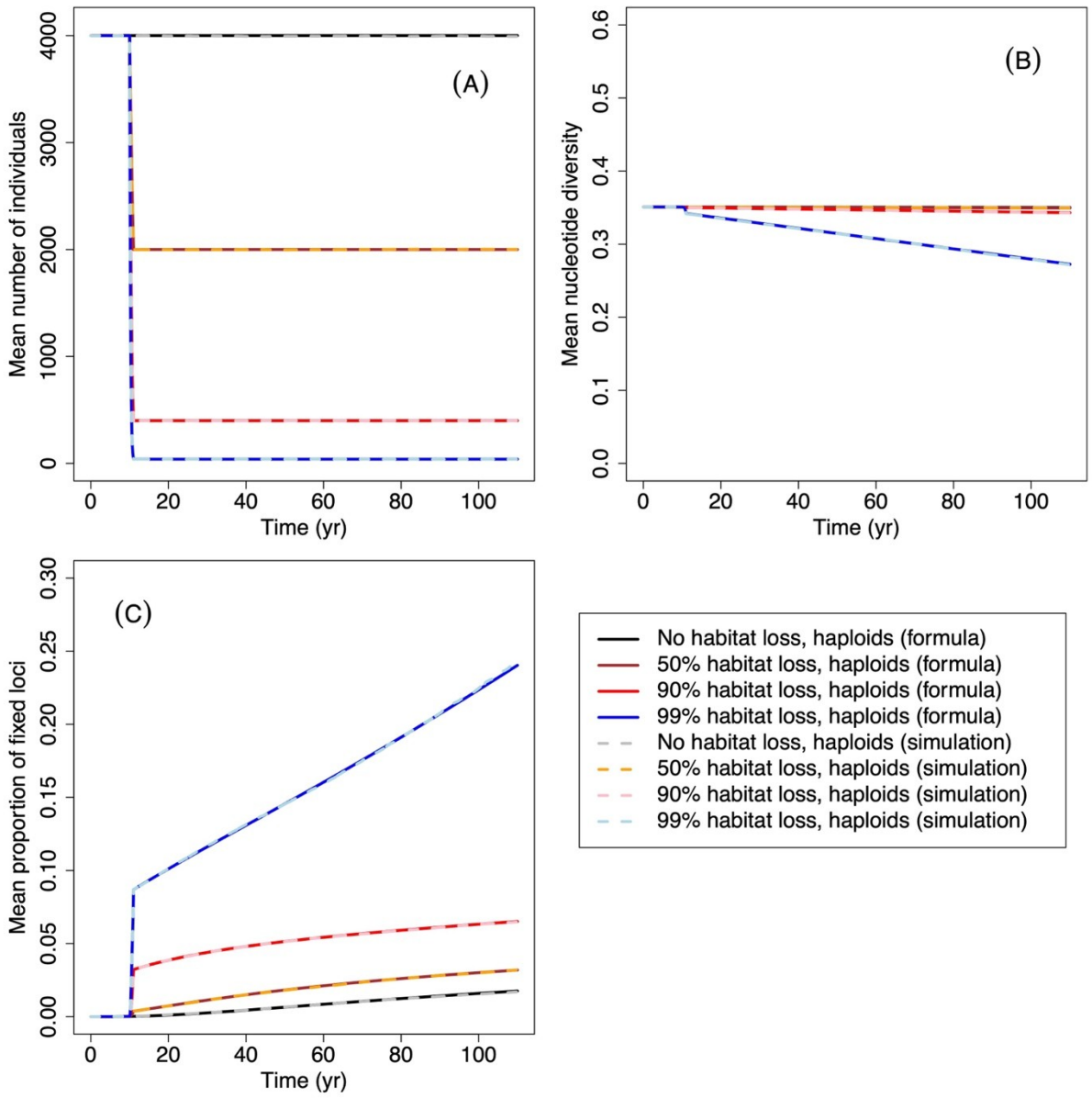


Fig. S1. Mean (A) number of individuals, (B) nucleotide diversity and (C) proportion of fixed loci over time from the non-spatial population model, under three different habitat loss scenarios with different proportional amounts of habitat lost, as derived using formulae and simulations. In each habitat loss scenario, habitat loss occurred between t_1=10 yr and t_2=11 yr. The per-capita death rate was D_0=0.04 yr^{-1} when there was no habitat loss and was given by D_1 (eq. (3) in main text) during habitat loss. The

value of 0.04 yr^{-1} was based on the maximum life span of a typical bird species (Table 1 in main text). The per-capita birth rate was set constant at $B=0.04$ yr^{-1}, such that the mean population size was constant when there was no habitat loss. 100 loci, each with two alleles, were considered. For each locus on a chromosome, the initial number of copies of the first allele was drawn from a random uniform distribution bounded by 0 and the initial population size. The initial population size was 4,000 (haploid individuals). For comparison, results from a baseline scenario with no habitat loss are also shown.

Examining a different distribution of initial allele frequencies

When deriving results for the non-spatial model presented in the main text (graphically illustrated in Fig. 3), we chose the initial allele frequencies by randomly sampling from a uniform distribution. Specifically, for each of the L loci on a chromosome, the initial number of copies of the first allele was drawn from a random uniform distribution bounded by 0 and the initial population size. To test the robustness of these results, we also derived results when the initial number of copies of the first allele was drawn using the steady-state allele frequency distribution of a neutral model with constant population size M.

The neutral model with constant population size is the neutral Moran process that describes the dynamics of two alleles at a locus, with mutation from one allele to the other. In each discrete timestep of the model, one allele is chosen to replicate (representing a haploid individual reproducing asexually) whereas another allele is chosen to disappear (representing a haploid individual dying). With small probably ν, the new allele in a timestep mutates and changes type.

The model eventually reaches a steady-state reflecting a balance between genetic drift and mutation, with the steady-state allele frequency distribution given by equation (10) in Moran (1958), which is

$$P(i/M)=P(0)\frac{M!\Gamma\left(i+\frac{Mv}{1-2v}\right)\Gamma\left(\frac{M(1-v)}{1-2v}-i\right)}{i!(M-i)!\Gamma\left(\frac{Mv}{1-2v}\right)\Gamma\left(\frac{M(1-v)}{1-2v}\right)}, \quad \text{(S18)}$$

where $P(i/M)$ is the probability of the allele frequency being i/M, and $P(0)$ is determined by the probabilities having to sum to 1, i.e. $\sum_{i=0}^{M} P(i/M)=1$. We use (S18) with $M=10^6$ and $v=10^{-9}$. The value of $M=10^6$ corresponds to an arbitrarily large population covering approximately 10^5 km^2 of pristine habitat (as per carrying capacity used in Table 1), whereas the value of $v=10^{-9}$ corresponds to the mutation rate used for the spatially explicit model (see Table 1 in main text). To account for SNPs generally being considered to have allele frequencies greater than 0.01 (Wright, 2005; Keats and Sherman, 2013), we truncate the allele frequency distribution given by (S18) at 0.01 (lower bound) and 0.99 (upper bound) and renormalize. The resulting allele frequency distribution is given in Fig. S2, which exhibits a U-shape.

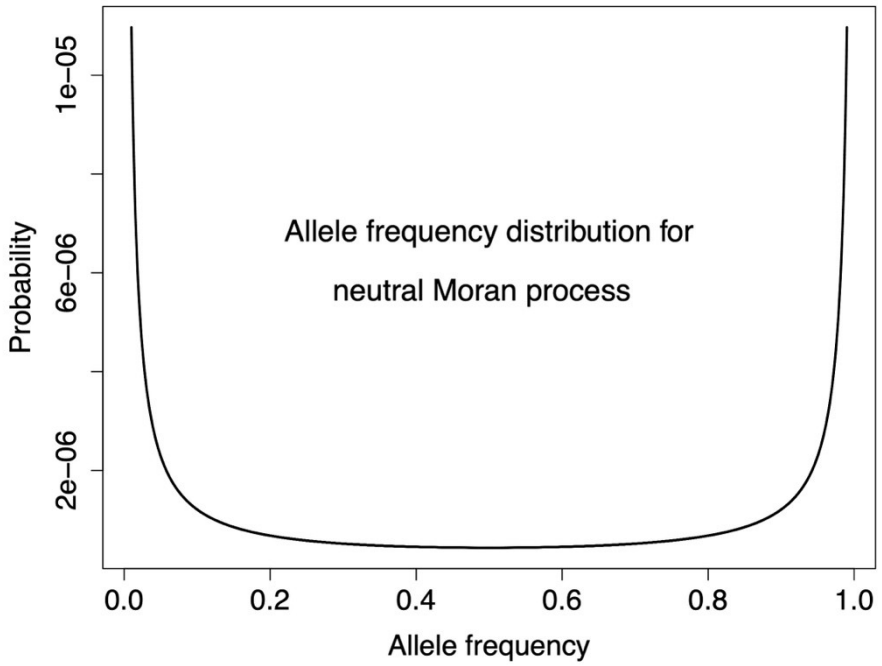


Fig. S2. Probability distribution of allele frequencies from the steady-state of a neutral Moran process with two alleles and mutation from one allele to another. The distribution was calculated using equation (10) in Moran (1958), with population size $M = 10^6$ and mutation rate $\nu = 10^{-9}$, with truncation of the distribution to reflect SNPs generally being considered to have allele frequencies greater than 0.01.

For each of the L loci on a chromosome, we randomly choose an allele frequency according to the distribution shown in Fig. S2 – this specifies how many of the M individuals have a copy of the first allele, and hence how many of them have a copy of the second allele. Afterwards, we randomly sample $N = 4{,}000$ of these M individuals to use in our non-spatial model of habitat loss. Results for our non-spatial model with this method of generating initial allele frequencies are shown in Fig. S3. Compared with results with the initial allele frequencies generated using a uniform distribution (Fig. 3 in main text), the mean nucleotide diversity is lower and the mean proportion of fixed loci is higher (Fig. S3), reflecting more rare alleles arising from the U-shaped initial allele frequency distribution (Fig.

S2). However, the trends in these two indicators of genetic diversity are qualitatively similar (cf. Figs. 3 and S3).

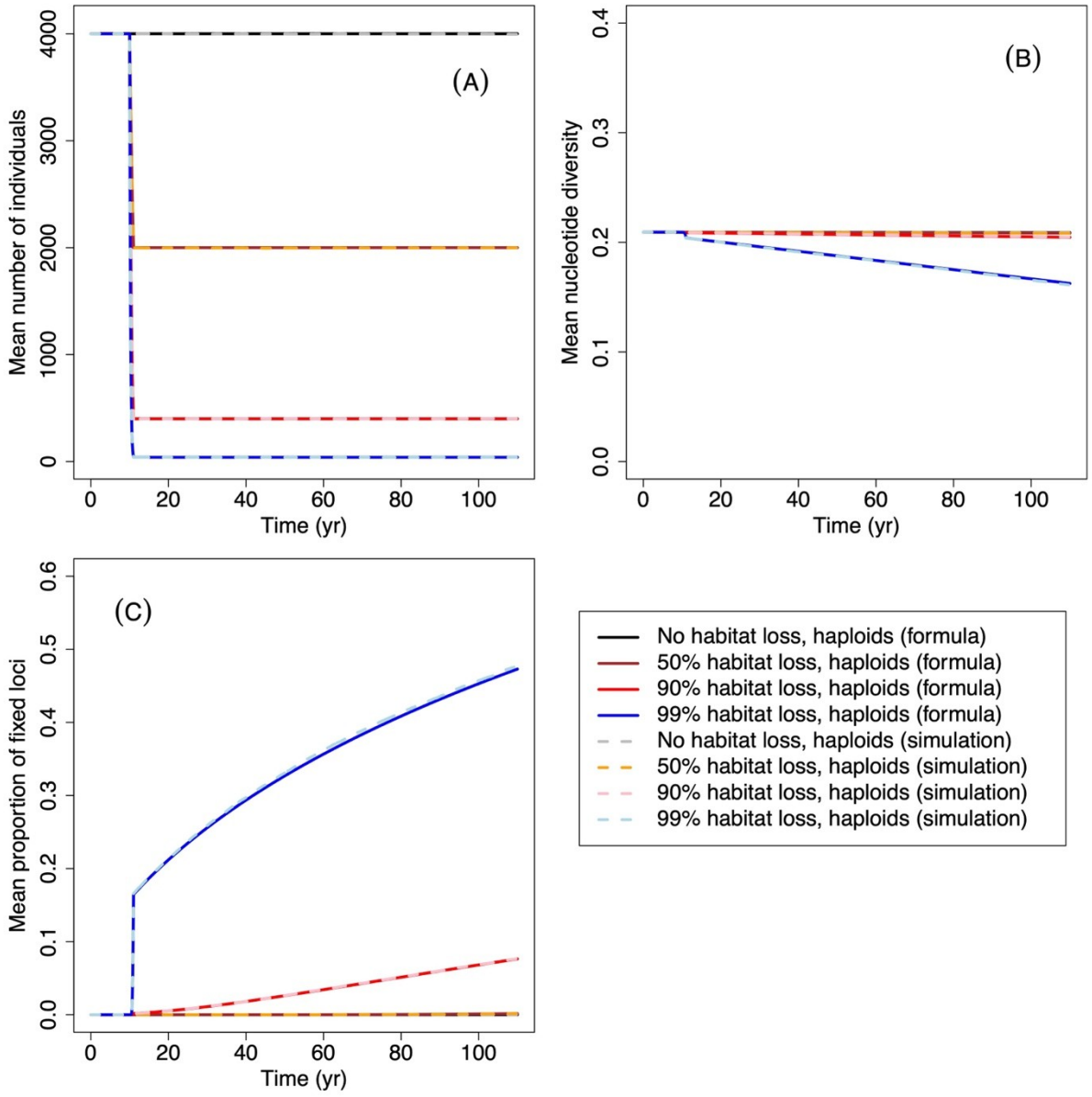


Fig. S3. Same as Fig. S1, except that the initial allele frequencies were drawn using the steady-state allele frequency distribution of a neutral Moran process with two alleles and mutation from one allele to another, as shown in Fig. S2.

Further results for spatial model

Table S1. All scenarios for simulations of the spatial model. Names of scenarios correspond to the names of folders in the online repository (https://github.com/qt37t247/HabitatLossGenetics), which contain the scripts to run the simulations.

						Species		
						A	**B**	**C**
Spatial configuration of preserved habitat	Random	**Cleared habitat (%)**	20	**Carrying capacity of cleared habitat (individuals)**	1	A0C2 D1	B0C2 D1	C0C2 D1
					2	A0C2 D2	B0C2 D2	C0C2 D2
					5	A0C2 D5	B0C2 D5	C0C2 D5
			50		1	A0C5 D1	B0C5 D1	C0C5 D1
					2	A0C5 D2	B0C5 D2	C0C5 D2
					5	A0C5 D5	B0C5 D5	C0C5 D5
			80		1	A0C8 D1	B0C8 D1	C0C8 D1
					2	A0C8 D2	B0C8 D2	C0C8 D2
					5	A0C8 D5	B0C8 D5	C0C8 D5
	Clumped		20		1	A1C2 D1	B1C2 D1	C1C2 D1
					2	A1C2 D2	B1C2 D2	C1C2 D2
					5	A1C2 D5	B1C2 D5	C1C2 D5
			50		1	A1C5 D1	B1C5 D1	C1C5 D1
					2	A1C5 D2	B1C5 D2	C1C5 D2
					5	A1C5 D5	B1C5 D5	C1C5 D5
			80		1	A1C8 D1	B1C8 D1	C1C8 D1
					2	A1C8 D2	B1C8 D2	C1C8 D2
					5	A1C8 D5	B1C8 D5	C1C8 D5
	N/A		N/A		N/A	Anull	Bnull	Cnull

Table S2. Average differences in genetic diversity at the end of simulations of the spatial model (each lasting 210 yr) for three stylized bird species, comparing random with clumped configuration of preserved habitat (positive value means average difference is greater in random versus clumped configuration). Significant values are highlighted with asterisks.

				Nucleotide diversity			Proportion of fixed loci (%)		
Species				A	B	C	A	B	C
Cleared habitat (%)	20	**Carrying capacity of cleared habitat (individuals)**	1	-0.00073	0.003534	0.005475	0.203613	0.725205	1.808306
			2	-1.3E-05	0.000834	0.003741	0.095876	0.440031	0.731571
			5	-0.00128	-0.00028	0.000234	-0.10959	-0.00589	0.165306
	50		1	0.008057	0.150626*	0.151088*	3.186791	43.1322*	45.48566*
			2	0.003162	0.012164	0.012266	1.350343	3.126857	3.913392*
			5	-0.00069	-0.00052	0.000563	-0.0833	0.09985	0.195568
	80		1	0.205753*	-	-	61.97188*	-	-
			2	0.097235*	0.047749	0.008966	30.47473*	13.72931	3.495596
			5	-0.00017	3.01E-05	-0.00095	0.146208	0.161301	0.156566

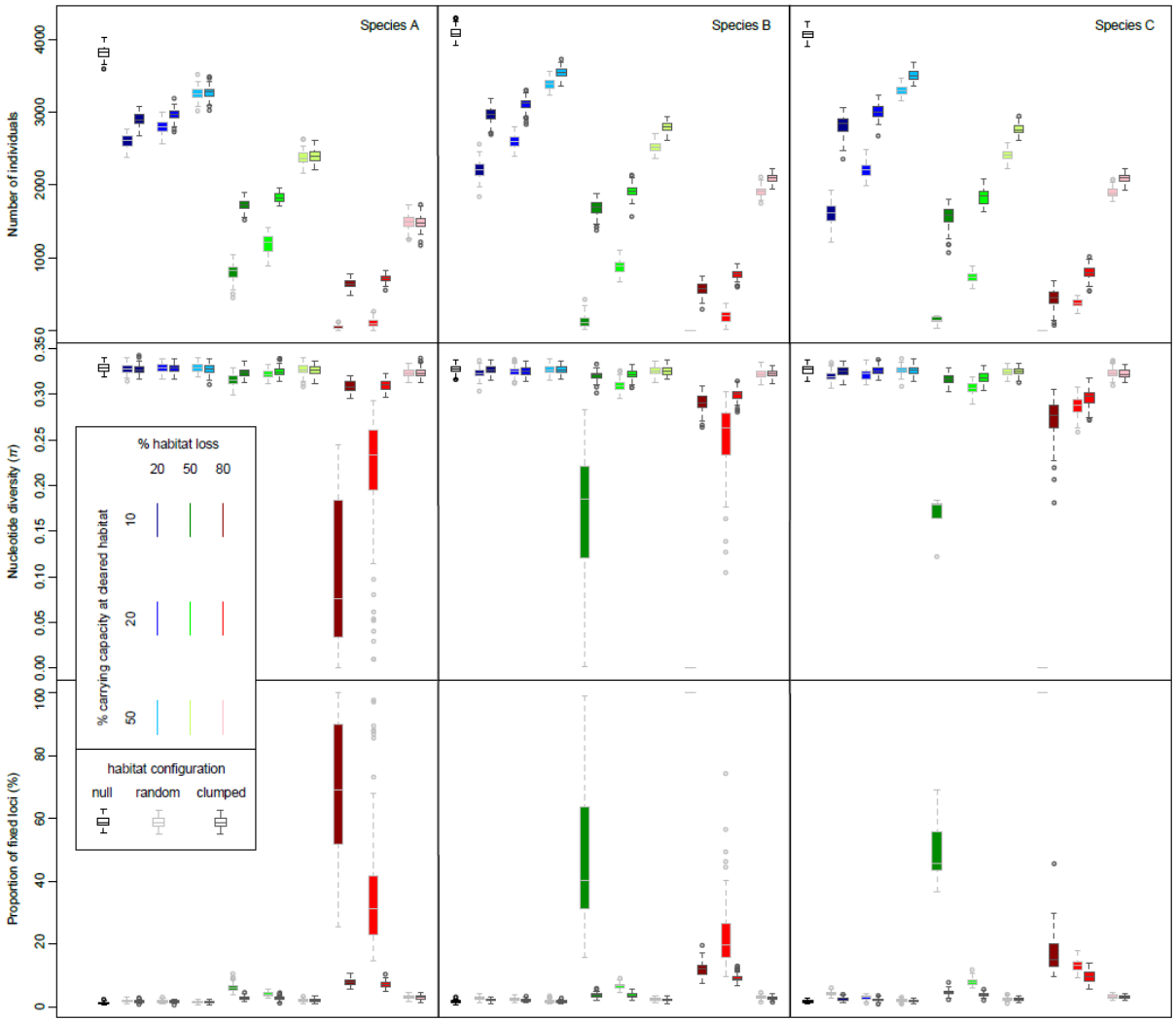


Fig. S4. Genetic diversity at the end of simulations of the spatial model (each lasting 210 yr), for three stylized bird species under different habitat loss scenarios.

References for Supplementary Material

Fung, T., Chisholm, R.A., 2023. Improving the realism of neutral ecological models by incorporating

transient dynamics with temporal changes in community size. Theor. Pop. Biol. 149, 12–26.

Keats, B.J.B., Sherman, S.L., 2013. Population genetics. In: Rimoin, D., Pyeritz, R., Korf, B. (Eds.),

Emery and Rimoin's Principles and Practice of Medical Genetics, Sixth Edition. Elsevier,

Amsterdam, Netherlands, Chapter 13 pp. 1–12. https://doi.org/ 10.1016/b978-0-12-383834-

6.00015-x.

Kendall, D.G., 1948. On the generalized "birth-and-death" process. The annals of mathematical

statistics 19, 1–15.

Moran, P.A.P., 1958. Random processes in genetics. Mathematical Proceedings of the Cambridge

Philosophical Society 54, 60–71.

Wright, A.F., 2005. Genetic variation: polymorphisms and mutations. In: Encyclopedia of Life

Sciences, John Wiley & Sons, Ltd., Hoboken, New Jersey, USA.

https://doi.org/10.1038/npg.els.0005005.