Abundance and population genetics of Hog Deer (*Axis porcinus*) in Victoria

D.S.L. Ramsey, C. Pacioni, and E. Hill

November 2019

Front cover photo: Male Hog Deer (Arthur Rylah Institute).

Accessibility

If you would like to receive this publication in an alternative format, please telephone the DELWP Customer Service Centre on 136 186, email customer.service@delwp.vic.gov.au or contact us via the National Relay Service on 133 677 or www.relayservice.com.au. This document is also available on the internet at www.delwp.vic.gov.au

Abundance and population genetics of Hog Deer (*Axis porcinus*) in Victoria

D.S.L. Ramsey[1], C. Pacioni[1], and E. Hill[2]

In partnership with

[1]Arthur Rylah Institute for Environmental Research
123 Brown Street, Heidelberg, Victoria 3084

[2]La Trobe University, Plenty Road Bundoora, Victoria 3086

Acknowledgements

Contents

Tables

Figures

Summary

Context:

Hog Deer (*Axis porcinus*) were introduced into Victoria in the 1860s but are currently largely confined to the coastal areas of south and east Gippsland. Hog Deer are a highly valued games species, and because of their relative rarity there is some concern among hunters that the Hog Deer population is in decline due to factors such as illegal hunting and loss of habitat.

Aims:

This study aimed to estimate the abundance and distribution of Hog Deer across their range as well as investigate the genetics of the Hog Deer population to examine genetic diversity, population structure and connectivity between local populations as well as effective population size.

Methods:

The abundance and density of Hog Deer was estimated using data from 100 camera traps set at 50 sites across their range in coastal Gippsland from Lower Tarwin to Point Hicks during November and December 2018. Monitoring was undertaken predominately on publicly accessible land such as State Game Reserves, balloted hunting reserves and National Parks, and hence did not include areas of private land managed for Hog Deer. Four distance markers were placed in each camera's field of view and used to estimate the distance between each deer and the camera location. Abundance (N) and density of deer was estimated using point transect distance sampling methods. A similar analysis was undertaken for Sambar Deer (*Rusa unicolor*) images detected by the camera array.

Tissue samples collected from Hog Deer between 2015 and 2019 were subject to DNA extraction, and 16 genetic markers (microsatellites) were identified and used in subsequent genetic analyses. Population structure was investigated using spatially explicit and non-spatial Bayesian clustering analysis and sex-biased dispersal patterns of Hog Deer were investigated using spatial autocorrelation analysis. The number of breeders contributing to the genetic pool of the population in each generation (effective population size–N_e) was also estimated using the linkage disequilibrium and coalescent based approaches. In addition, descriptive statistics of genetic diversity were also calculated.

Results:

Hog Deer images were recorded from 36 cameras at 22 of the 50 sites. Notably, no images of Hog Deer were obtained from cameras located east of Lakes Entrance. The analysis of Hog Deer abundance was restricted to the 1762 km^2 area comprising native vegetation within the Hog Deer breeding range. The total abundance of Hog Deer within this area was estimated to be 3000 (95% confidence interval: 1858–4845). Most of the population was estimated to be on Wilsons Promontory (2079 deer, 95% confidence interval: 1054–4103), Snake Island had a population of 246 (95% confidence interval: 160–378) and a further 675 (95% confidence interval: 481–947) were estimated for the Gippsland Lakes region.

Predicted abundances do not include contributions from areas of private land managed for Hog Deer, so they are likely to underestimate the total number of Hog Deer in the Gippsland Lakes region. Most Sambar Deer were detected east of Lakes Entrance (188 deer; 95% confidence interval: 108–334). Fallow Deer (*Dama dama*) were detected at two camera locations near Port Welshpool and Lower Tarwin.

Results from the Bayesian clustering analysis identified the presence of three genetic clusters (subpopulations), centred on Wilsons Promontory (western subpopulation), Snake Island, and the Gippsland Lakes (eastern subpopulation). Genetic variability within each cluster was very low, with average heterozygosity around 40% and an average of only two alleles per locus. Genetic differentiation between the western and eastern subpopulations was also very low. Genetic relatedness was greater than expected by chance for Hog Deer at Boole Poole, Sunday Island, Blond Bay and Clydebank. Spatial autocorrelation analysis found that likely dispersal in the eastern Hog Deer population was less than 60 km for females and 140 km for males. No dispersal patterns were discernible for the western subpopulation. Effective population size (N_e) was 110 for the western subpopulation, 62 for the eastern subpopulation and 10 for the

Snake Island subpopulation. The values for the western and Snake Island subpopulations are lower than the recommended value (N_e =100) for ensuring genetic viability within each subpopulation in the long term.

Conclusions and implications:

The Hog Deer population inhabiting three core areas in coastal Gippsland between Lower Tarwin and Point Hicks was predicted to be low, with less than 1000 individuals estimated to inhabit areas outside Wilsons Promontory. However, as our monitoring design did not include areas of private land managed for Hog Deer (e.g. Sunday Island and some mainland private properties), our estimates of abundance are likely to underestimate the total population of Hog Deer in this region. Future monitoring programs should attempt to represent the contributions of these managed areas to the Hog Deer population to improve the estimates of abundance for the region.

The Hog Deer population is characterised by low genetic variability and low effective population size, putting it at a high risk of inbreeding depression, especially for the Snake Island and eastern subpopulations. Hence the long-term viability of the Gippsland Lakes and Snake Island subpopulations of Hog Deer are vulnerable to potential impacts such as overharvesting, disease or habitat loss.

There are also significant barriers to successful dispersal and colonisation of suitable habitat outside the breeding range. For example, the likelihood of significant breeding populations establishing east of Lakes Entrance are hypothesised to be limited because of predation by wild dogs and/or competition with Sambar Deer in this area. If Hog Deer are to be managed as a sustainable hunting resource, management of the current extant core subpopulations is essential to ensure their long-term viability.

Recommendations

The following recommendations are made for the Hog Deer population inhabiting publicly accessible areas where hunting is permitted (e.g. State Game Reserves from Jack Smith Lake to Ewing Morass) as well as areas designated for balloted hunting (e.g. Snake Island, Blond Bay, Boole Poole).

- Investigate the sustainable level of hunting to ensure the long-term viability of Hog Deer populations in these areas.

- Monitor Hog Deer populations periodically to determine population trends, especially in State Game Reserves and areas designated for balloted hunting. Future monitoring should include private land managed for Hog Deer to determine the contributions of these areas to the overall Hog Deer population.

- Manage some State Game Reserves to increase the carrying capacity and productivity of Hog Deer populations where this is feasible and does not impact other biodiversity values.

- Undertake studies of fitness parameters, such as reproductive and survival rates of Hog Deer, and use the results to trigger management actions.

- Develop a management plan for Hog Deer (on publicly accessible areas where hunting is permitted) that critically examines the feasibility and biodiversity impacts of potential management actions to ensure the long-term viability of core subpopulations.

1   Introduction

The Hog Deer (*Axis porcinus*) is a small (30–50 kg) species of deer native to southern Asia. It was introduced into the Gippsland region of Victoria in the 1860s and 1870s, and by the 1940s had spread from the initial release sites near Port Welshpool and Sale to colonise the coastal strip between the Tarwin River and Lakes Entrance (Mayze and Moore 1990; Menkhorst 1995). Recent investigations examining locations of Hog Deer sightings from various sources has concluded that the breeding range of Hog Deer encompasses a 2336 km^2 coastal strip between the Tarwin River and Point Hicks. Although Hog Deer have been seen outside these areas, notably as far east as Mallacoota and as far west as the Otway Ranges, none of these areas currently support breeding populations (Forsyth *et al.* 2016).

Hog Deer commonly inhabit coastal shrublands and swamps, including Manna Gum and *Banksia* woodlands and *Leptospermum*, *Melaleuca* and *Acacia* scrub (Menkhorst 1995). Although their diet in their native range consists primarily of grasses, in Victoria their diet consist predominantly of dicots, including both forbs and shrubs (Davis *et al.* 2008). Adult Hog Deer are mainly solitary, but females may be accompanied by subadult offspring, and aggregations may sometimes occur where food is plentiful (Menkhorst 1995). In Victoria breeding occurs mainly in December and January, and births peak in August and September following a gestation period of approximately 240 days. Females may breed in their first year and produce an average of 1.2–1.4 young per year; males develop their first set of antlers at around 10 months of age (Mayze and Moore 1990; Menkhorst 1995).

Because of its relative rarity, the Hog Deer is a highly valued game species in Victoria. Hunting of Hog Deer is highly regulated: licensed hunters are allowed to take one male and one female deer during April, and the current harvest is around 150 animals per year (Moloney and Turnbull 2018). While Hog Deer are appreciated for their aesthetics and are a valued hunting resource, they can have a negative impact on biodiversity and may pose problems for private landholders (Côté *et al.* 2004; Davis *et al.* 2016). There is also some concern among hunters that the Hog Deer population is declining as a result of various factors such as habitat degradation and illegal hunting (Slee 1985). Extensive areas of suitable Hog Deer habitat (e.g. wetlands/marshes and coastal scrub) that once existed between Westernport Bay and the Gippsland lakes have been progressively lost or degraded through agricultural development. This loss of habitat is thought to have restricted movements of Hog Deer, resulting in increasingly fragmented populations, which have become more vulnerable to other impacts such as hunting (Mayze and Moore 1990).

Among the various population parameters, population size is arguably one of the most important, because knowledge about how abundance or density varies across the landscape can provide important information about prevailing trends (e.g. increasing or declining) as well as whether management actions are effective. However, very little information on densities or abundance are available for any deer species in Australia (Davis *et al.* 2016). In addition to abundance, information on other aspects of population dynamics such as sex-specific dispersal patterns and 'effective' population size obtained from molecular data can provide important insights into the connectivity, sustainability and genetic condition of populations. Knowledge about these population parameters for Hog Deer, and their relationship to various habitat characteristics, would provide valuable information that could be used to guide the management of Hog Deer populations.

Here we report on a study to estimate the current abundance and densities of Hog Deer in south and east Gippsland using data collected from a random sample of locations across the known range. We have also undertaken an analysis of Hog Deer genetics using tissue samples taken from Hog Deer checking stations during the hunting season, to investigate the degree of population structuring and connectivity between local populations as well as the effective population size. Using this information, we make some recommendations for managing Hog Deer in Victoria.

2 Methods

2.1 Hog Deer abundance and density estimates

2.1.1   Monitoring

Monitoring to assess the abundance and density of Hog Deer was undertaken across the range of the species in coastal Gippsland, from Lower Tarwin to Point Hicks (Figure 1). The area from which sampling sites were drawn (the sampling frame) was estimated from a habitat suitability model fitted to the known locations of Hog Deer, extracted from the Victorian Biodiversity Atlas database and other sources. We then intersected this habitat map with the estimate of the Hog Deer breeding range (Forsyth *et al.* 2016) and then subdivided the resulting map into units of 2×2 km cells that could be sampled. A spatially balanced random sample of at least 50 cells was then selected, with the probability of a cell being selected weighted by the habitat suitability value (Figure 1).

A spatially balanced sampling design was preferred over simple random or stratified sampling because it ensures good coverage of the environmental space by selecting samples that minimise the spatial autocorrelation among sampled units. This improves efficiency by ensuring that each sampled unit provides as much unique information about the environmental space as possible (Foster *et al.* 2017). Monitoring was conducted primarily on public land such as State Game Reserves and National Parks and hence, did not include areas of private land managed for Hog Deer.

At each of the 50 cells (sites) selected for monitoring, two heat-in-motion cameras (Reconyx HC600/HC900 Hyperfire H.O. Covert IR) were deployed at the approximate centroid of the cell, separated by at least 50 metres. Cameras were situated at locations where deer could potentially be detected and to ensure good visibility of deer. Four plastic markers were placed in the midline of the field of view of each camera at 2.5, 5, 7.5, and 10 metres from the camera location, which were used to classify images of deer into distance classes from the camera. Reflective tape was applied to each marker to ensure good visibility at night (Figure 2). The distance measurements of Hog Deer from the camera were then used to estimate deer density using point distance sampling methods (Buckland *et al.* 2006; Howe *et al.* 2017). Cameras were set out for approximately one month during November and December 2018, a period chosen so that potential disturbance of the sites by hunters or other people was likely to be minimal.

2.1.2   Density and abundance estimates

Images of deer from each camera were classified into species, and also sex and age class (juvenile/adult) if possible, by volunteers from the Australian Deer Association. For each deer image, the distance class was also recorded using the plastic markers as distance guides. We used the point transect distance sampling model of Howe *et al.* (2017) to estimate the densities of Hog Deer at each camera location and then used these to estimate the population abundance of Hog Deer in the breeding range. Briefly, this method assumes that cameras are deployed independently of animal locations at a location k for a period of time T_k and captures images for as long as an individual is present to trigger the camera. Images are then obtained at a predetermined set of instants, t units of time apart. Temporal effort at each camera is then calculated as T_k / t. Howe *et al.* (2017) suggested that a useful range for t is 0.25 to 3 seconds with values at the lower end being more suitable for fast-moving or rarer species. If the camera covers a horizontal angle of view of θ radians then the fraction of the circle observed by the camera field of view is $\theta / 2\pi$. Hence, the data consist of a series of snapshot instants taken t units of time apart with overall sampling effort at each location k equal to $(\theta T_k)/2\pi t$ (Howe *et al.* 2017). Estimates of density ($\widehat{D}_k$) follows standard point transect methods (Buckland *et al.* 2006),

$$\widehat{D}_k = \frac{n_k}{\pi w^2 \, e_k \hat{p}_k}$$
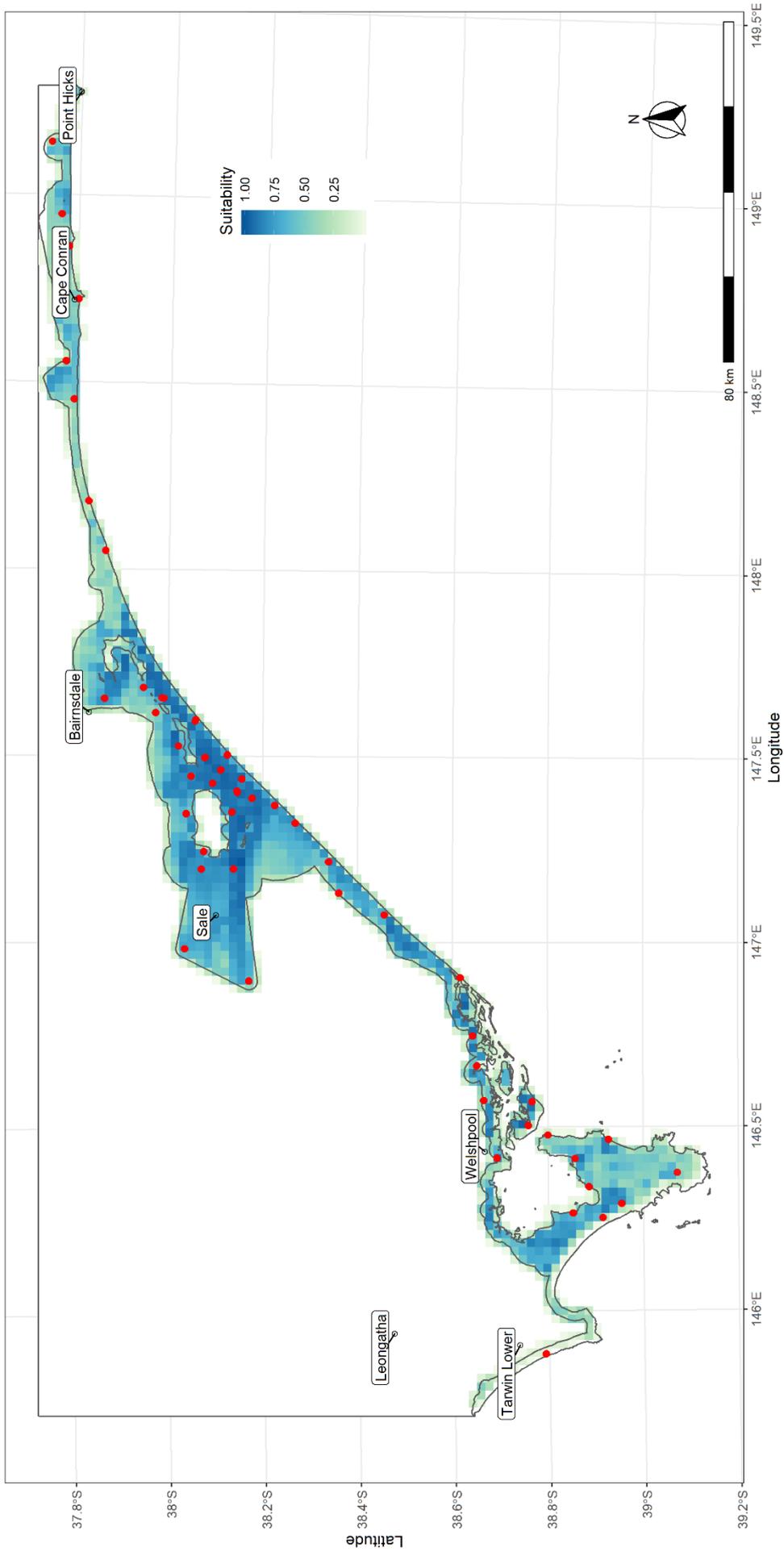
Figure 1. Habitat suitability (0–1) within the Hog Deer breeding range at a resolution of 2 × 2 km. Red circles indicate sampled sites used to estimate Hog Deer abundance and density.

where n_k is the number of observations of deer at location k, w is the maximum observation distance from the camera (truncation distance), $\hat{p}_k$ is the probability of detecting an individual that is within w distance from the camera and e_k is the overall sampling effort $(\theta T_k)/2\pi t$.

The detection probability within the area covered by the camera $\hat{p}_k$ is estimated from the detection function, which models how detection probability declines with distance from the camera using the observed distance class information recorded for each individual deer (Buckland *et al.* 1993). We investigated two forms for the detection function (half-normal and hazard-rate); the main difference between these two forms is that the hazard-rate function has a more pronounced shoulder than the half-normal. The fit of each of these functions to the distance sampling data can be improved by including adjustment terms (Miller *et al.* 2019b). We examined the addition of up to two cosine adjustment terms to each detection function and compared their relative fit with the Akaike information criterion (AIC) (Burnham and Anderson 1998). Once a suitable detection function was found, the function with respect to distance was integrated to give the marginal or 'average' probability of detection $\hat{p}_k$.

Since the survey design consisted of spatially referenced samples from camera traps, we used density surface modelling (DSM) (Miller *et al.* 2013) to estimate Hog Deer densities and abundance within the area defined by the Hog Deer range (Figure 1). DSM is a spatial model that seeks to construct a relationship between spatially varying abundance and corresponding environmental variables, in order to predict abundance or density over the entire study region, not just the areas sampled. In the context of the present study, this required us to estimate a predictive relationship between the counts of Hog Deer at each camera location and environmental variables. Because the counts were collected using point transect distance sampling, a two-stage modelling approach was adopted. A detection function was first fitted to the distance data to estimate the average detection probability $\hat{p}_k$. The detection probability was then incorporated into a generalised additive model with the observed counts of individuals as the response and environmental covariates as the potential explanatory variables. The model for the observed counts at each camera location was therefore

$$E(n_k) = \hat{p}_k e_k \exp\left[\beta_0 + \sum_j f_j(z_{jk})\right]$$

Where β_0 is an intercept term and f_j are (possibly smooth) functions of the covariates z_{jk}. The overall sampling effort e_k and the detection probability $\hat{p}_k$ were fitted in the model as an offset, which effectively corrects the expected counts for both imperfect detection and sampling effort (Miller *et al.* 2013). The Poisson distribution would usually be specified for the observed counts. However, the presence of significant overdispersion (such as excess zeros) might result in a poor fit for the Poisson distribution. Alternative distributions that can be considered for over-dispersed count data include the negative binomial, quasi-Poisson and Tweedie distribution (Candy 2004).

The covariates examined in the DSM included both biotic (e.g. vegetation type) and abiotic (temperature/rainfall) variables as well as models including smooth functions of spatial location (Table 1). Variables relating to temperature and rainfall were derived from BioClim (https://www.worldclim.org/bioclim), which consists of a set of global climate layers for the period 1970–2000 at a spatial resolution of approximately 1 km (30 seconds). We included BioClim layers relating to mean annual temperature (°C), temperature seasonality (standard deviation), minimum and maximum annual temperatures, annual rainfall (mm), rainfall in the wettest quarter and rainfall in the driest quarter (Table 1). Vegetation types were derived from Ecological Vegetation Classes (EVC) extracted from Victorian Bioregions spatial layers available at https://services.land.vic.gov.au (NV2005_EVCBCS layer) (Table 1). We combined similar EVCs to derive five vegetation types as some EVCs were either not sampled or had inadequate sample sizes for analysis. These comprised (1) coastal scrub/shrublands, (2) heathlands or heathy woodlands, (3) woodlands or forest, (4) herb-rich woodlands and (5) wetlands (see Appendix 1 for further details). We constructed a suite of plausible models using combinations of these variables and compared their relative fit using AIC (Burnham and Anderson 1998). The model with the lowest AIC had higher support than the alternative models examined and hence, was then used to predict deer abundance for the Hog Deer breeding range. All models were fitted using the dsm package (Miller *et al.* 2019a) in R (ver. 3.5.3) (R Development Core Team 2018). Potential non-linear relationships were explored by including smooth functions of continuous

variables in models using thin-plate regression splines available in the mgcv package (Wood 2017), which is utilised by dsm. Smooth functions of variables were indicated by s( ) enclosing the respective variable. The variance of overall Hog Deer abundance was estimated using the parametric moving-block bootstrap method in dsm (Miller *et al.* 2013).

Figure 2. Camera trap images showing the plastic markers used to indicate distance from the camera. Distance intervals of 2.5 m were used for this study.

Table 1. Variables used in the density surface model (DSM) to investigate relationships with deer abundance.

Variable	Description	Source
MeanTemp	Mean annual temperature (℃)	Bioclim https://www.worldclim.org/bioclim
TempSeas	Temperature seasonality (℃)	Bioclim https://www.worldclim.org/bioclim
MaxTempW	Maximum temperature of warmest month (℃)	Bioclim https://www.worldclim.org/bioclim
MinTempC	Minimum temperature of coldest month (℃)	Bioclim https://www.worldclim.org/bioclim
AnnRain	Mean annual rainfall (mm)	Bioclim https://www.worldclim.org/bioclim
RainW	Rainfall in the wettest quarter (mm)	Bioclim https://www.worldclim.org/bioclim
RainD	Rainfall in the driest quarter (mm)	Bioclim https://www.worldclim.org/bioclim
EVC	Vegetation type - five classes (see Appendix 1 for details)	Derived from EVC layer NV2005_EVCBCS https://services.land.vic.gov.au
s(X,Y)	Spatial smooth of location	Site locations (easting/northing)

2.2 Population genetics

Genetic data were used to provide complementary information on Hog Deer ecology and population dynamics. Genetic data from tissue samples (tongue, liver or skin) obtained from wild, free-ranging Hog Deer in Victoria were available from a previous study (Hill, unpublished data). These samples were collected between 2008 and 2017 by hunters during the hunting season (April), and during annual culls on Wilsons Promontory in August 2015 and 2016. These data were integrated with data obtained from skin samples collected during the 2019 hunting season to maximise the coverage across the Hog Deer distribution in Victoria. Sample locations were identified by the closest geographical feature (town, reserve or check station), so they are only indicative of the general area. As a result, several samples had the same coordinates even though they were possibly sampled at some distance from each other.

We used primers for 16 microsatellite loci that were known to be polymorphic in Hog Deer (see Appendix 2 for details) and removed from the dataset samples with genotypes missing for more than five loci, leaving a total of 403 samples across 54 sites for analysis.

2.2.1 Population structure

We evaluated whether we could use the genetic data to identify demographically discrete populations, and if so, whether we could delineate the boundaries of these populations using population structure analyses. These analyses can also be used to identify 'migrants' (animals that move from one population to another) and admixed individuals (animals whose genome originates partly from a different population). Population structure was investigated with the spatially explicit Bayesian clustering analysis implemented in Geneland (Guillot *et al.* 2005) and the (non-spatial) Bayesian assignment tests implemented in Structure (Pritchard *et al.* 2000; Falush *et al.* 2003). Geneland and Structure implement models with population admixture as well as models that assume uncorrelated allele frequencies between different populations or a relaxed model that allowed for correlated allele frequencies between clusters (Falush *et al.* 2003; Guillot 2008). (Further details of these analyses are provided in Appendix 2.)

Using Structure output for the admixture model with correlated allele frequencies, we also calculated the number of direct migrants between populations by calculating the number of individuals that were assigned to a genetic cluster different to that in which they were sampled. For this analysis only individuals that had more than 75% of their genome assigned to a different cluster were considered migrants.

Descriptive statistics of genetic diversity parameters were calculated in Genalex (Peakall and Smouse 2006), including the expected heterozygosity H_e (Hartl and Clark 1997), average of observed and expected number

of alleles (Brown and Weir 1983) and average number of private alleles P_A. Cluster differentiation (F_{ST}) was calculated in Genalex under the AMOVA framework.

2.2.2 Dispersal

Genetic data allows an evaluation of the distances at which related individuals within each sex disperse to assess the degree of sex-biased dispersal (if any is present). This can be achieved by investigating the spatial organization of genotypes across the landscape. For this purpose we used spatial autocorrelation analysis, which compares pairwise genetic distances with geographic distances (Smouse and Peakall 1999). Distance classes were evenly spaced at 20 km intervals. Confidence intervals were calculated using 1000 bootstraps of the autocorrelation coefficients (r) and to test statistical significance, we ran 1000 random permutations and conducted a test for heterogeneity (Peakall *et al.* 2003; Smouse *et al.* 2008; Banks and Peakall 2012).

To further explore the relationship between individuals we also computed the (within each cluster) mean pairwise relatedness coefficient (Queller and Goodnight 1989) using Genalex and tested whether the animals sampled within the same site (where more than 10 individuals were sampled from any given site) were more related than randomly expected using 1000 bootstrap permutations.

2.2.3 Effective population size

The effective population size (N_e) is, in simple terms, an estimate of the number of breeders that contribute to the genetic pool of the population in each generation. Hence N_e is also related to population abundance (N). The ratio N_e / N has been investigated in several taxa and is on average between 0.1 and 0.23 (Palstra and Fraser 2012; Frankham *et al.* 2014) although it is not a fixed value and depends on the interplay between mortality and reproductive rates (Waples *et al.* 2011). It is a very important parameter when evaluating the genetic health of wild populations because the rate of genetic diversity loss and inbreeding is related to the inverse of N_e (i.e. a small N_e leads to increased genetic loss and inbreeding), ultimately leading to reduction of fitness and higher extinction probability (Frankham 1995).

We estimated the current effective population size (N_e) using the linkage disequilibrium approach (Waples 2006) using the program NeEstimator (Do *et al.* 2014). We report the parametric 95% confidence intervals around the estimate as well as those computed with the JackKnife approach (Jones *et al.* 2016). Despite the fact that this method does not explicitly take into account gene flow between populations (and the expected admixture between populations), it has been demonstrated that it can perform an accurate estimation of the effective population size if the migration rate is less than 10% and the population is not excessively large (i.e. N_e < 500) (Waples and England 2011). Because allele frequencies that are close to 0 or 1 can strongly bias N_e estimation (Waples 2006), we evaluated the effect of three allele frequencies thresholds: 0.01, 0.02, and 0.05. That is, alleles whose frequencies is below the threshold are dropped out in this analysis. It is important to note that especially in presence of migration, the use of higher thresholds is recommended (Gilbert and Whitlock 2015).

To further explore the demographic dynamics of Hog Deer in Victoria, we used coalescent based approaches to examine changes in the effective population size over time. Hence, this analysis examines whether the effective population size N_e is stable or undergoing change (increasing or decreasing). The parametric exponential growth (Griffiths and Simon 1994) and the non-parametric Extended Bayesian Skyline Plot (EBSP) (Heled and Drummond 2008) models were fitted to two datasets, one with the samples from Snake Island alone and one with a subset of samples (n = 81) stratified by sampling locations across all Victoria. Further details of these analyses are provided in Appendix 2.

3 Results

3.1 Hog Deer abundance and density

We obtained 877 observations of distances between cameras and Hog Deer. Hog Deer were recorded on 36 cameras at 22 of the 50 sites. Notably, no images of Hog Deer were obtained from cameras located east of Lakes Entrance (Figure 3). One camera in the south of Wilsons Promontory recorded over 280 images of Hog Deer (11 images/day), more than twice as many as the camera with the next highest encounter rate (Figure 3). This camera was situated in an area where a high proportion of deer were recorded feeding, spending extended periods in front of the camera. There was also a large number of images of a resting deer, a situation not recorded on any other camera. Because the inclusion of data from this camera had a correspondingly large influence on density and abundance estimates, we excluded this data from further analysis. However, data from the second camera at this site was retained.
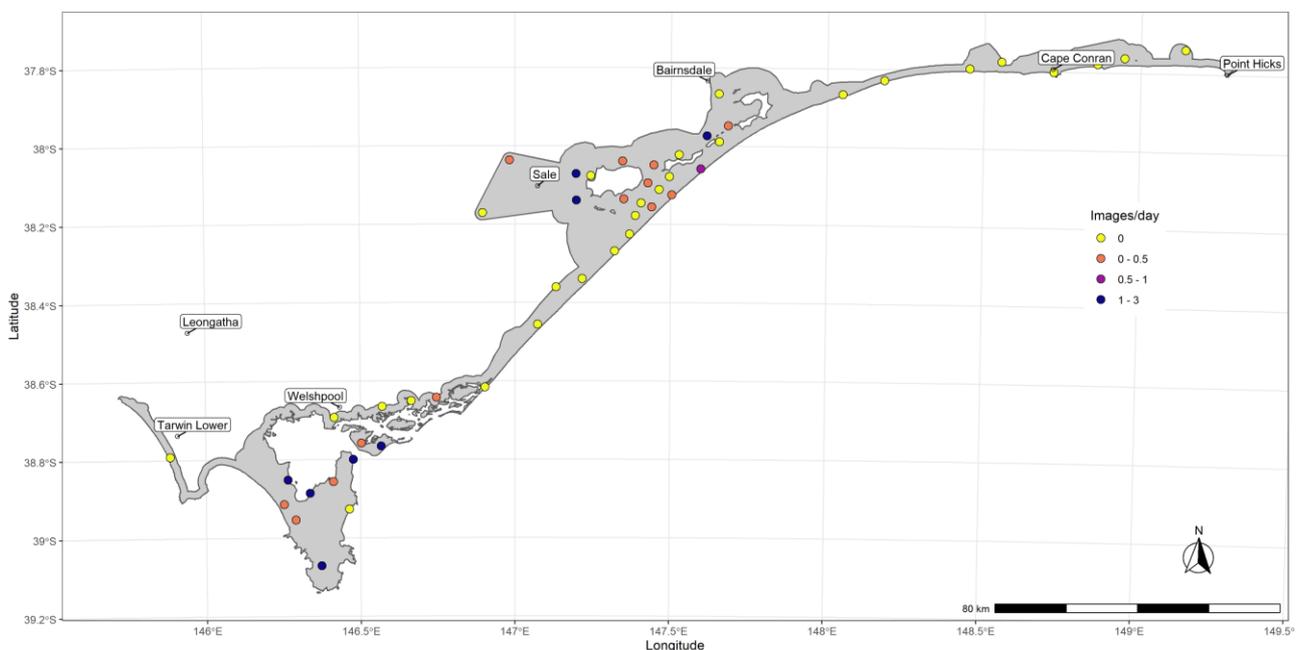
Figure 3. Encounter rate of Hog Deer (deer images/day) recorded at each site.

Fitting the half-normal and hazard-rate detection functions to the distance data revealed that models with at least a 2nd or 3rd order cosine adjustment term resulted in minimal AIC (Table 2). The half-normal detection function with 2nd and 3rd order cosine terms had the lowest AIC and resulted in a good fit to the data (Figure 4). Detection probability of Hog Deer was close to 1.0 within two metres of the camera dropping to around 0.5 at 5 metres from the camera and almost zero beyond 8 metres (Figure 4). The average probability of detection $\hat{p}_k$ out to a truncation distance of 12.5 metres was 0.13.

Table 2. Results of model selection among six models for the detection function fitted to the Hog Deer distances recorded from each camera. AIC–Akaike's Information Criterion, ΔAIC–difference between AIC and the model with the lowest AIC (in bold).

Detection model	Adjustments	AIC	ΔAIC
Half-normal	none	2080.1	6.5
Half-normal	2nd order cosine	2081.4	7.9
Half-normal	2nd and 3rd order cosine	**2073.5**	**0.0**
Hazard-rate	None	2081.3	7.8
Hazard-rate	2nd order cosine	2073.7	0.2
Hazard-rate	2nd and 3rd order cosine	2075.5	1.9
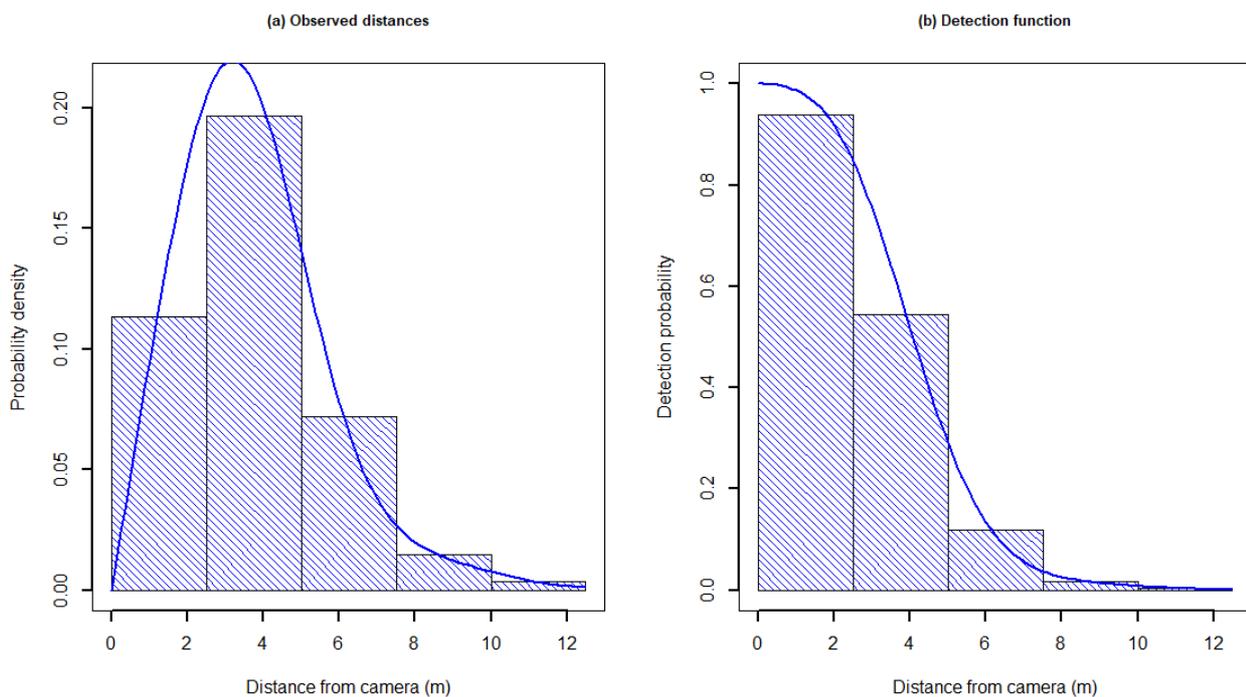
Figure 4. The probability density of observed distances (left) and the fitted half-normal detection function (right) estimated from distances recorded between the camera and images of Hog Deer.

Because of the large number of sites with zero counts of Hog Deer, we modelled the counts of Hog Deer images using the Tweedie distribution rather than a Poisson distribution, as the former resulted in a much better fit to the data. Based on the potential covariates available to use in the DSM (Table 1), initial exploratory analysis indicated that smooth functions of MeanTempW, MinTempW and RainD explained significant variation in DSM models so were used in further analyses along with vegetation class (EVC) and spatial location, s(X,Y). Six plausible models were constructed using various combinations of these variables to model the distribution of Hog Deer (Table 3).

Table 3. Details of models fitted to the Hog Deer encounter rate in cameras to estimate the spatial distribution and abundance of Hog Deer. Non-linear smooth functions of variables are indicated by `s()` enclosing the respective variable. AIC–Akaike's Information Criterion, ΔAIC–difference between AIC and the model with the lowest AIC (in bold). All models were fitted in the dsm package.

Model	Variables	df	AIC	ΔAIC	%Deviance
1	s(MeanTemp) + s(MinTempW)+ s(RainD)	17	411	18	54
2	EVC	7	424	31	24
3	s(X,Y)	23	398	5	71
4	s(MeanTemp) + s(MinTempC) + s(RainD) + EVC	19	407	14	60
5	s(MeanTemp) + s(MinTempC) + s(RainD) + s(X,Y)	23	402	9	69
6	EVC + s(X,Y)	27	**393**	**0**	77

Results of model selection for the six dsm models indicated that model 6, including the effects of ecological vegetation class (EVC) and spatial location s(X,Y), was the most supported by the data, being the model with lowest AIC. This model explained 77% of the deviance in the data (Table 3). Model 3, which included only spatial location was the next most supported model, having a difference in AIC of 5 with the most supported model and explained 71% of the deviance in the data. A check on the adequacy of Model 6 by plotting the deviance residuals against the theoretical quantiles (QQ-Plot) indicated some underestimation of Hog Deer counts by the model at higher abundance (Figure 5). A plot of the randomised quantile residuals vs fitted values indicated that the model was an adequate fit to the data with no obvious residual trends.
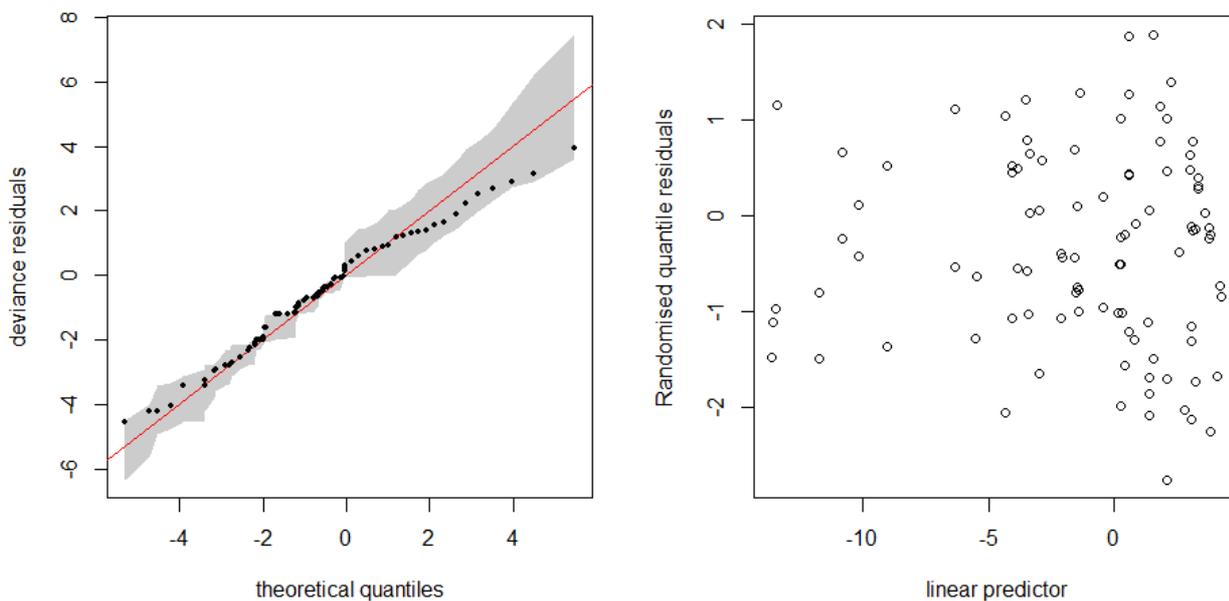
Figure 5. Quantile–quantile (QQ) plot (left) and randomised residual plot (right) for the most supported dsm model, model 6.

Prediction of hog deer abundance was restricted to the 1762 km² area comprising native vegetation occurring within the Hog Deer breeding range (Figure 6). Since there was a relatively small difference in the level of support for model 6 and model 3 by the data (ΔAIC of 5) as well as the percentage deviance explained by each model (77% vs 71%), we predicted the total abundance from both models because they gave quite different overall predicted abundances, ranging from 2199 to 3802 deer within the breeding range (Table A2, Appendix 1). We also averaged the predicted abundance over both models (model averaged prediction) to incorporate the model selection uncertainty into predicted abundances. The model averaged

prediction assumed equal weighting for both models, which is a reasonable approach when averaging over a small set of plausible models (Dormann *et al.* 2018). We calculated the variance of the model averaged prediction by combining bootstrap replicates from both models with equal weighting. The resulting model-averaged prediction indicated that the total abundance of Hog Deer within the species' breeding range was 3000 deer, with a 95% confidence interval of 1858–4845 (Table 4).

The average density ($\hat{D}$) of Hog Deer across the Hog Deer breeding range was estimated to be 1.7 deer/km^2 (Table 4). However, spatial variation in density across the range was high, having a standard deviation of 4.4 deer/km^2 (Table 4). This was exemplified by the estimate of the median density, which was only 0.2 deer/km^2, almost an order of magnitude lower than the mean density indicating a high degree of skew in the distribution of density estimates across the range (Table 4). The predicted distribution of Hog Deer density exhibited three core subpopulations. One subpopulation was centred on Wilsons Promontory where the highest densities were apparent on the north-west, south and north-east coast (Figure 6). The second subpopulation was centred on Snake Island, and with the third on the Gippsland Lakes, especially the region around Lake Wellington and Blond Bay (Figure 6). Given the existence of these three subpopulations, we produced separate abundance and density estimates for three regions within the Hog Deer breeding range (Table 4). These regions comprised the areas to the west of Port Welshpool, including the Lower Tarwin and Wilsons Promontory (Wilsons Promontory region), Snake Island (Snake Island region), and the area to the east of Port Welshpool (Gippsland Lakes region) (Figure 7). Abundance was much higher in the Wilsons Promontory region (2079 deer) compared with the Snake Island and Gippsland Lakes regions (246 & 675 deer, respectively) (Table 4), with the majority of deer predicted to occur in the southern half of Wilsons Promontory. However, the estimate for Wilsons Promontory was also the most imprecise, having a coefficient of variation (relative error) of 36% (Table 4). Mean hog deer density was highest on Snake Island (4.6 deer/km^2) and lowest in the Gippsland Lakes region (0.6 deer/km^2) (Table 4).

Table 4. Estimates of the total abundance ($\hat{N}$) and density ($\hat{D}$, deer/km^2) of Hog Deer in native vegetation in the Hog Deer breeding range (1762 km^2) as well as for three regions within the breeding range (Figure 7). se–standard error; cv–coefficient of variation; LCL, UCL–lower and upper 95% confidence intervals; sd–standard deviation.

Region	$\hat{N}$	se($\hat{N}$)	cv($\hat{N}$) %	95% LCL	95% UCL	$\hat{D}$ (mean)	$\hat{D}$ (median)	sd($\hat{D}$)
Wilsons Promontory	2079	743	36	1054	4103	3.9	1.7	7.1
Snake Island	246	55	22	160	378	4.6	4.1	4.5
Gippsland Lakes	675	117	17	481	947	0.6	0.07	1.3
Breeding range	3000	745	25	1858	4845	1.7	0.2	4.4
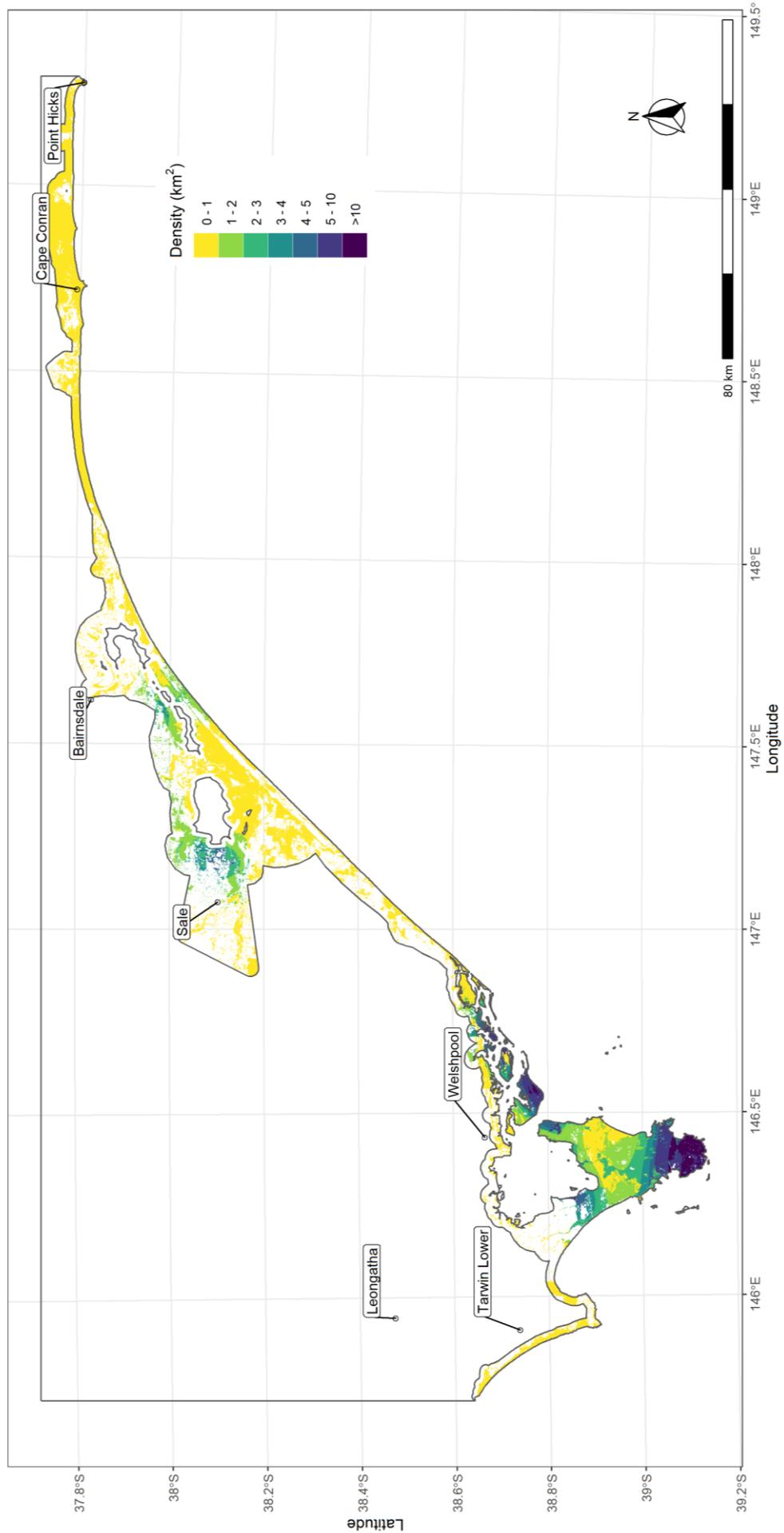
Figure 6. Distribution of Hog Deer density (deer/km2) within native vegetation within the Hog Deer breeding range (1762 km2). Predictions are from the model-averaged estimates from model 6 and model 3, assuming equal weights.
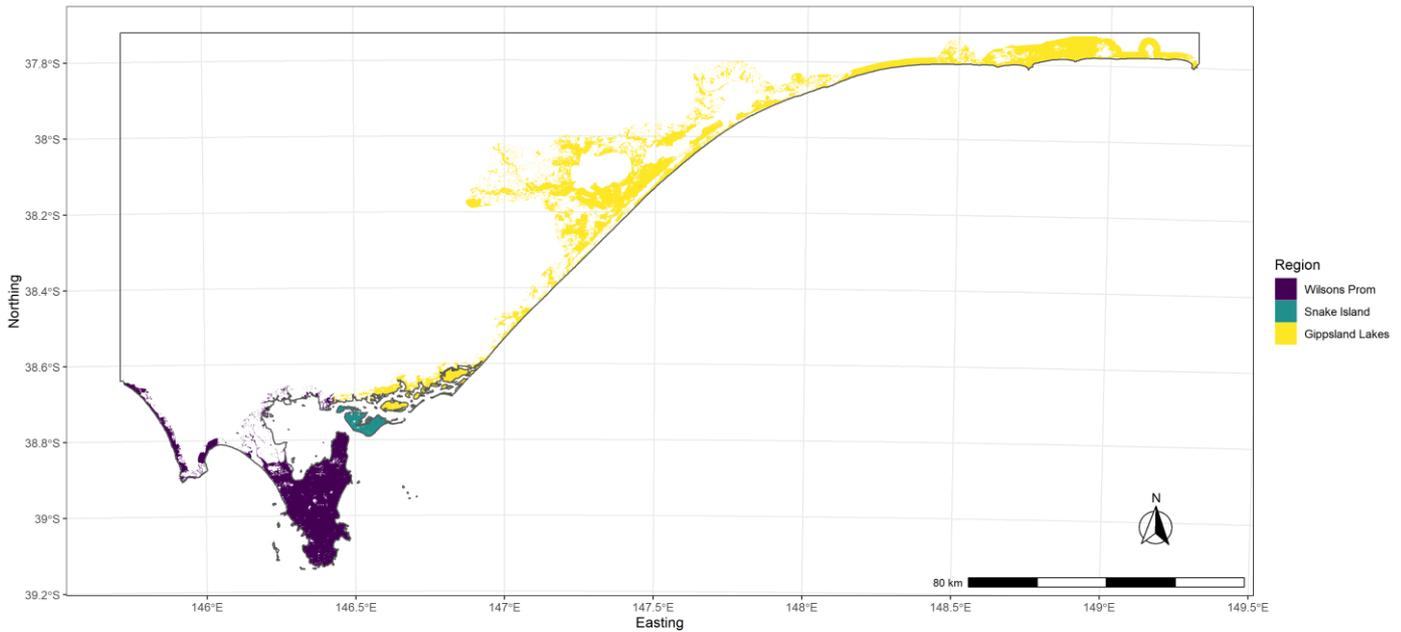
Figure 7. Regions within the Hog Deer breeding range. Each region contains a core Hog Deer subpopulation with abundance and density estimates provided in Table 4.

3.2 Other deer species

Camera sampling also detected the presence of two other deer species: Sambar Deer (*Rusa unicolor*) and Fallow Deer (*Dama dama*). Sambar Deer were detected at 16 camera locations at 10 sites (Figure 8) while Fallow Deer were only detected at four camera locations at two sites (Port Welshpool and Lower Tarwin) (Figure 9). We used the data on Sambar Deer detections to estimate the abundance and density of Sambar Deer within the Hog Deer breeding range. However, there were insufficient detections of Fallow Deer for abundance estimation, so these data were not analysed further.
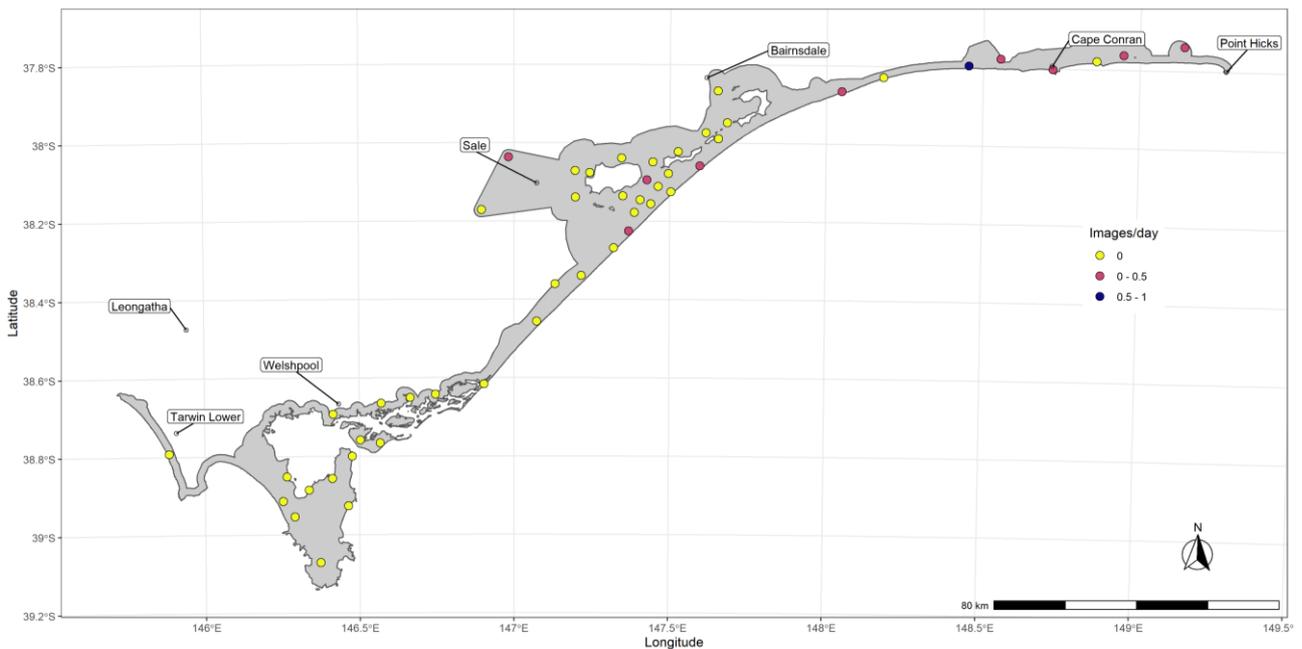
Figure 8. Encounter rate of Sambar Deer (deer images/day) recorded at each site.

Fitting the half-normal and hazard-rate detection functions to the Sambar Deer distance data revealed that the hazard-rate model with a 2nd order cosine adjustment term resulted in minimal AIC (Table A3, Appendix 1) and resulted in a good fit to the data (Figure 10). Despite the different form of the detection function for Sambar Deer detections compared with Hog Deer (hazard-rate vs half-normal), the detection probabilities were similar. The average probability of detection $\hat{p}_k$ of Sambar Deer out to a truncation distance of 12.5 metres was 0.18, slightly higher than that for Hog Deer (0.13).
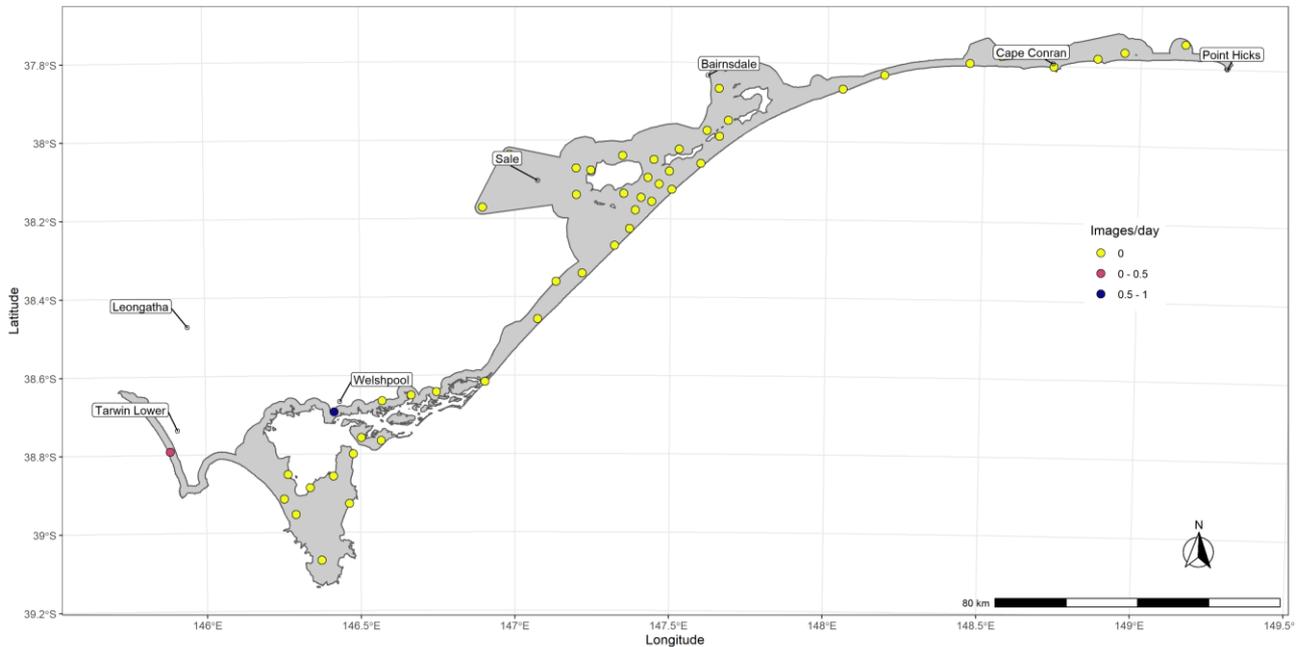
Figure 9. Encounter rate of Fallow Deer (deer images/day) recorded at each site.
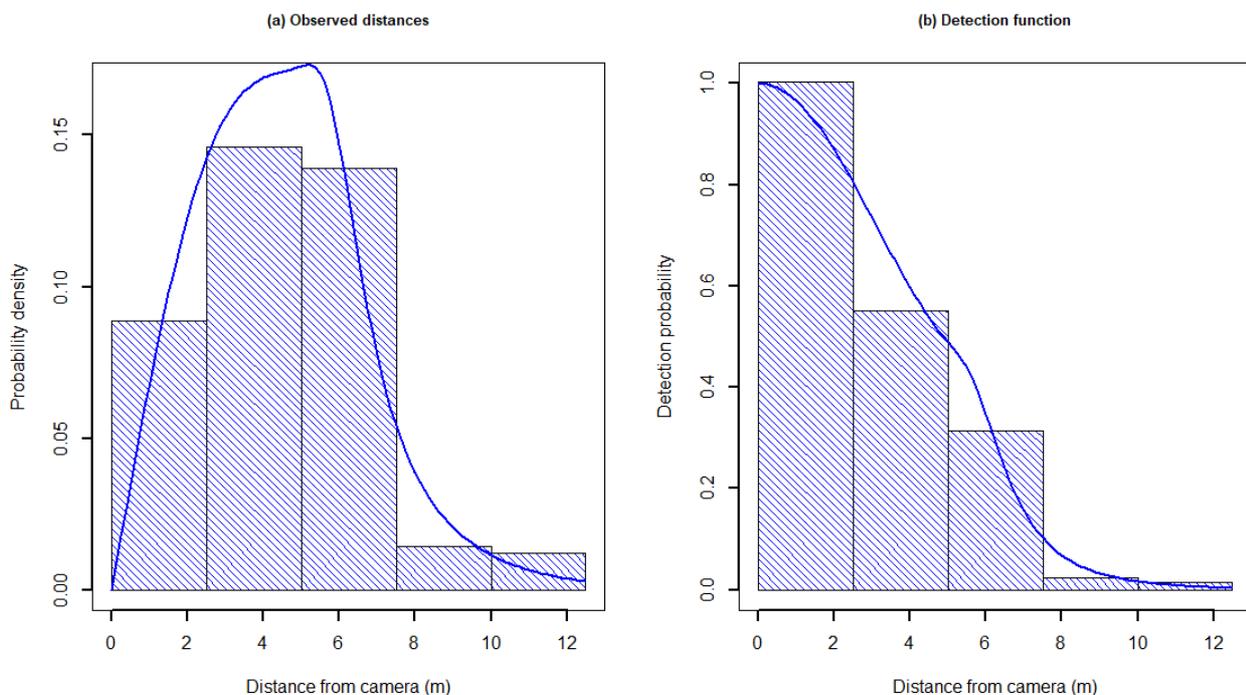
Figure 10. Probability density of observed distances (left) and the fitted hazard-rate detection function (right) estimated from distances recorded between the camera and images of Sambar Deer.

Because of the high number of sites with zero counts of Sambar Deer, counts of Sambar Deer images were modelled using the Tweedie distribution, rather than the Poisson as the former resulted in a much better fit to the data. Based on the potential covariates available to use in the DSM (Table 1), initial exploratory analysis indicated that smooth functions of TempRange, and a linear function of MeanTemp explained significant variation in DSM models so were used in further analyses. The occurrence of zero counts of Sambar Deer within some of the vegetation classes (EVC) resulted in numerical errors during model fitting. Hence, EVC was modelled as a smooth random effect rather than a fixed effect as for Hog Deer. The model selection results for the six models fitted to the Sambar Deer counts suggested that Model 1, consisting of the linear effect of MeanTemp and a smooth function of TempRange was the most supported by the data, having the lowest AIC (Table 5). Model 4 also had near equal support but as this model gave similar estimates of abundance as Model 1, we preferred Model 1 for inference about Sambar Deer density due to its simpler structure. This model explained 67% of the deviance in the data and was an adequate fit to the data, based on inspection of QQ plots and randomised quantile residuals, with no strong residual trends evident (Figure A1, Appendix 1).

Table 5. Details of models fitted to the Sambar Deer encounter rate at cameras to estimate the spatial distribution and abundance of Sambar Deer within the Hog Deer breeding range. Non-linear smooth functions of variables are indicated by an s() enclosing the respective variable. AIC–Akaike's Information Criterion, ΔAIC–difference between AIC and the model with the lowest AIC (in bold). All models were fitted in the dsm package.

Model	Variables	df	AIC	ΔAIC	%Deviance
1	s(MeanTemp) + s(TempRange)	9	**201**	**0**	67
2	s(EVC)	7	220	19	22
3	s(X,Y)	11	206	5	64
4	s(MeanTemp) + s(TempRange) + s(EVC)	9	201	0	67
5	s(MeanTemp) + s(TempRange) + s(X,Y)	13	208	7	67
6	S(EVC) + s(X,Y)	13	209	8	66

Since no Sambar Deer were detected west of the Gippsland Lakes, we restricted predictions of abundance and density to the Gippsland Lakes region (Figure 7). The abundance of Sambar Deer in this region was predicted to be 188 (95% Confidence Interval; 108–334). The average density of Sambar Deer in this region was predicted to be 0.16 deer/km^2 (standard deviation; 2.7). The distribution of Sambar Deer indicated densities were highest east of Lakes Entrance (Figure 11).
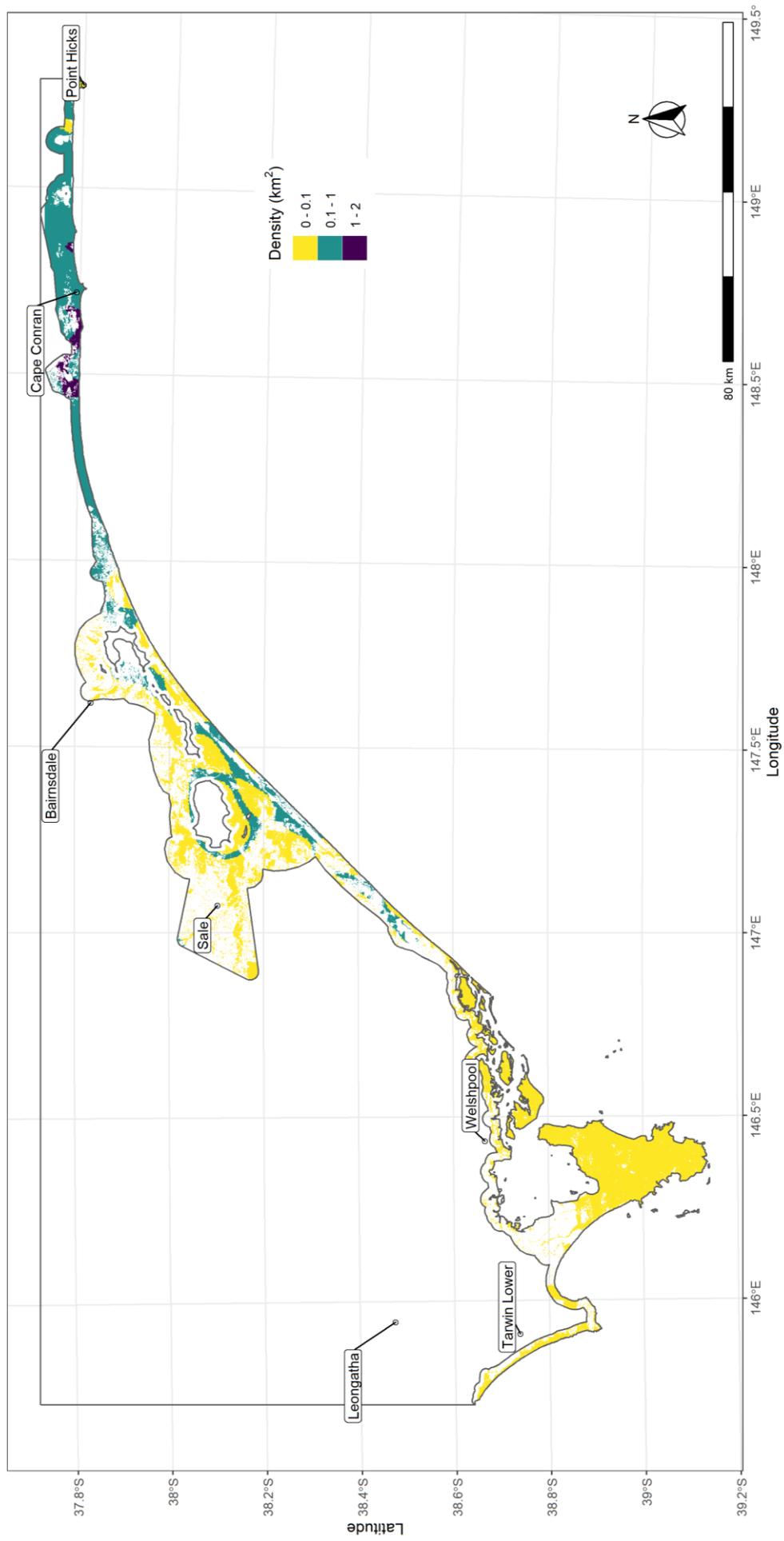
Figure 11. Distribution of Sambar Deer density (deer/km^2) within native vegetation within the Hog Deer breeding range (1762 km^2).

3.3 Population genetics of Hog Deer

3.3.1 Population structure

Bayesian clustering analyses supported models with three genetic clusters as the most likely with only a few differences in individual assignments. We report the results from the strictest model (spatially explicit without admixture and non-correlated frequencies in Geneland), which we used to delineate the boundaries of the three clusters, and the most flexible model (with admixture and correlate allele frequencies in Structure), which we used to evaluate migration/admixture rates. In the former, Snake Island formed a distinct cluster, a second cluster included the eastern sites, including Sunday Island and the Gippsland Lakes while the third cluster encompassed the western sites as far east as Gelliondale (Figure 12). This model did not detect any migrants and the probability of assignment of individuals were relatively high (i.e. > 65%) except the sites at the interface between the western and eastern cluster (Port Welshpool, Gelliondale, Port Albert and Sunday Island) where it was approximately 50%.
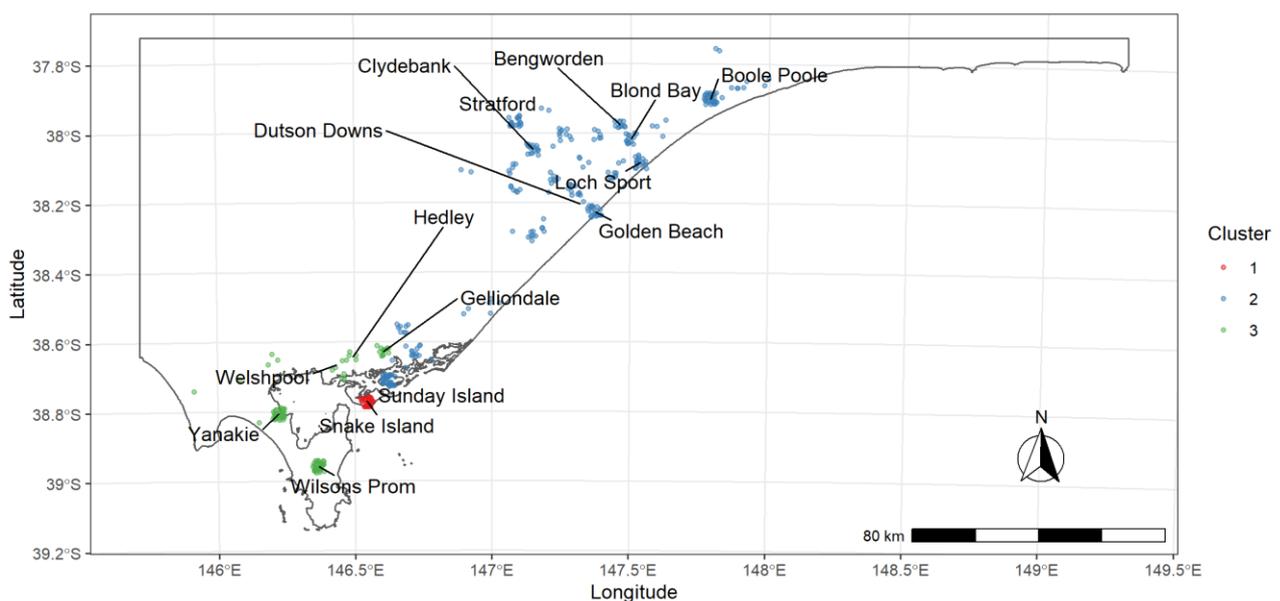
Figure 12. Map of individual assignment (each individual assigned to the most likely genetic cluster) with the non-admixture model with uncorrelated allele frequencies from Geneland.

In the model including admixture with correlated allele frequencies in Structure, the western sites (as far east as Hedley) formed a relatively homogeneous cluster. Another cluster consisted almost exclusively of Hog Deer on Snake Island, except for a few individuals from Loch Sport, Stratford, and Golden Beach assigned to this population. The third cluster consisted of the eastern sites, but sites in the Gippsland Lakes, Blond Bay, Bengworden and Stratford areas showed a high degree of admixture with individuals from the western cluster (Figure 13).

Based on the results presented above, we separated the sampling sites according to the three clusters identified by Geneland under the stricter model (i.e. the model without admixture and uncorrelated allele frequencies), from here onwards identified as Snake Island, western cluster and eastern cluster. Using this subdivision, we then identified migrants and admixed individuals in each cluster using Structure results from the model with admixture and correlated allele frequencies as explained above (Table 6).

Genetic variability within each cluster was very low, with most markers with 2 alleles on average and heterozygosity around 40% (Table 6). Snake Island had slightly lower variability, although it is important to point out that the sample size was also smaller than for the other clusters. Differentiation (based on F_{ST} statistic) between eastern and western clusters was significant but very low (Table 7).
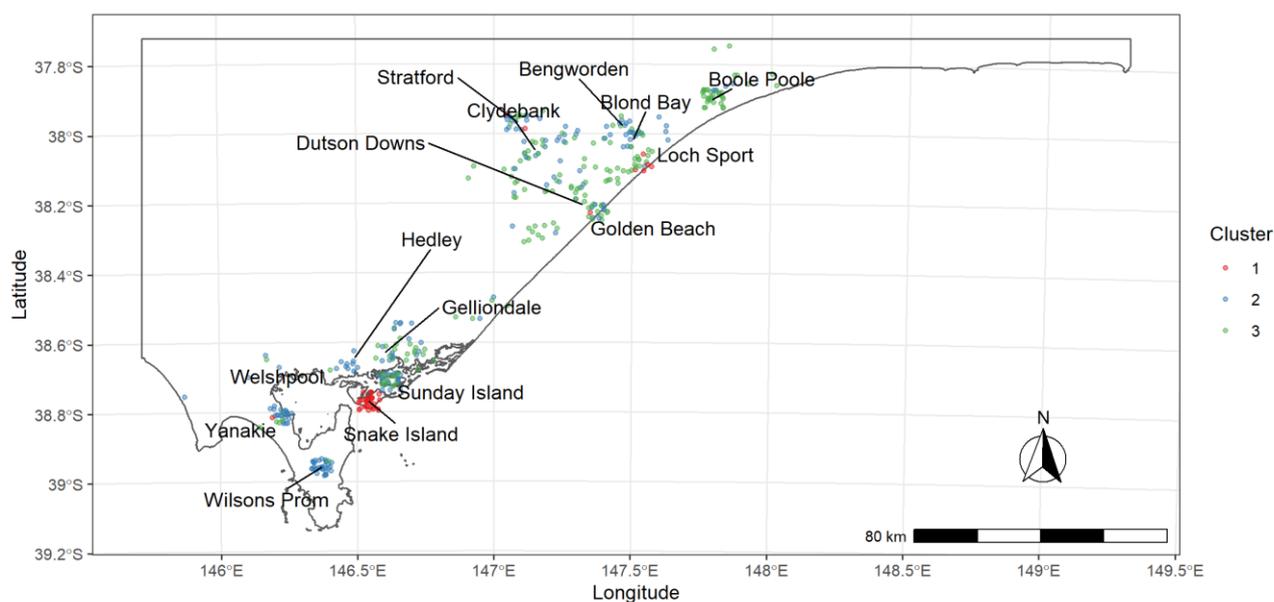
Figure 13. Map of Structure individual assignment (each individual assigned to the most likely genetic cluster) with the admixture model with correlated allele frequencies.

Table 6. Descriptive statistics (mean and standard error) for the three clusters as identified by Geneland for the model without admixture and uncorrelated allele frequencies.

Parameter*	Snake Island	Eastern	Western
n	30	281	92
N_A	2.06 (0.11)	2.75 (0.19)	2.19 (0.14)
N_A Freq. >= 5%	2 (0.13)	2.06 (0.11)	2 (0.13)
N_{EA}	1.62 (0.12)	1.78 (0.13)	1.78 (0.13)
P_A	0 (0)	0.38 (0.13)	0.06 (0.06)
H_E	0.34 (0.04)	0.39 (0.04)	0.39 (0.05)
Resident	1	0.60	0.84
Migrant	0	0.23	0.05
Admixed	0	0.17	0.11

* n – samples size

N_A – mean number of alleles

N_{EA} – mean number of effective alleles

P_A – private alleles

H_E – expected heterozygosity

resident, migrant, admixed – proportion within of each cluster of resident animals (individuals assigned to the cluster where they have been sampled), migrant (individuals with more than 75% ancestry from a different cluster from where they were sampled or admixed if less than 75%

Table 7. Pairwise F_{ST} statistic between genetic clusters. P values are above the diagonal.

	Snake Island	Eastern	Western
Snake Island	—	0.001	0.001
Eastern	0.162	—	0.001
Western	0.147	0.044	—

3.3.2 Dispersal

The spatial autocorrelation analysis was conducted independently for the eastern and western clusters. In the eastern cluster, both sexes displayed significant autocorrelation (Figure 14; $p = 0.001$). Females in the first and third distance class (< 20 and between 40 and 60 km) had a positive and significant spatial correlation coefficient ($p = 0.001$ and 0.023 respectively). That is, female individuals sampled at up to 60 km distance were more likely to be related than expected by chance. The distance classes 80–120 km had a negative and significant spatial correlation coefficient ($p = 0.017$, 0.001 and 0.022 respectively) indicating that individuals in these classes were not likely to be related. Males also had a significant spatial correlation coefficient in the first distance class ($p = 0.001$), although about half that of females (Figure 14). Males had a negative spatial correlation in distance classes > 140 km ($p = 0.001$), indicating that individuals separated by this distance were unlikely to be related. The two autocorrelograms are statistically different ($p = 0.001$), indicating different dispersal patterns between males and females.
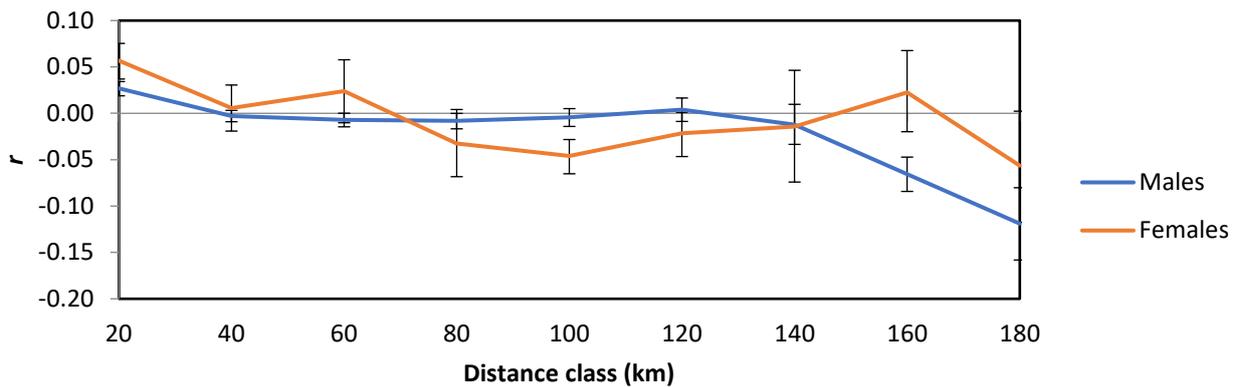
Figure 14. Spatial autocorrelation analysis for the eastern cluster comparing the genetic distance between individuals of each sex at increasing sampling distance. The vertical axis reports the spatial correlation coefficient, r (positive values: related individuals, negative values: unrelated individuals. Error bars including $r = 0$ indicate no significant correlation for that distance class).

Pairwise relatedness was greater than expected by chance at Boole Poole ($p = 0.001$), Sunday Island ($p = 0.001$), Blond Bay ($p = 0.015$) and Clydebank ($p = 0.02$), although at Blond Bay and Clydebank the confidence intervals encompassed the 95% confidence interval of the null hypothesis of no difference, suggesting that only a few pairs were strongly related. No sites had a significant relatedness coefficient within the western clusters (Figure 15).
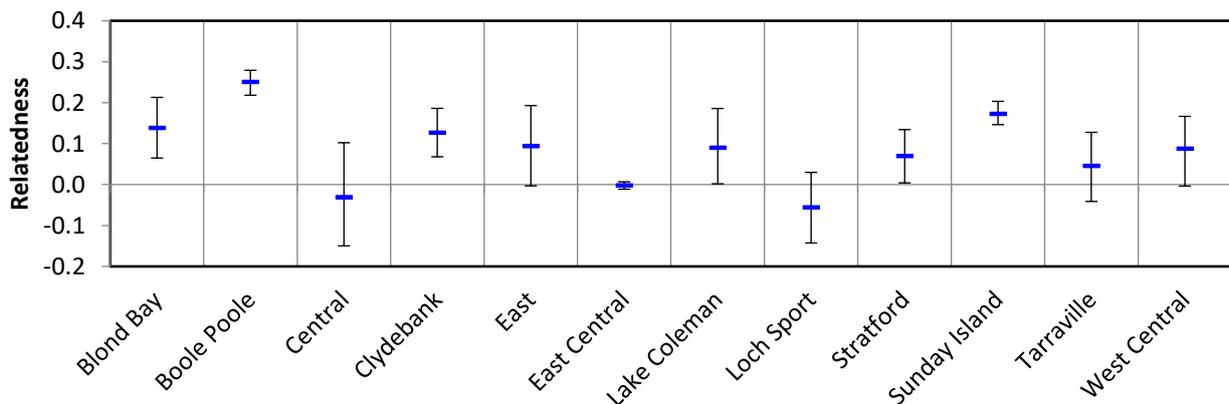
Figure 15. Mean relatedness coefficient (vertical axis) within the eastern cluster for sites with more than 10 samples (some sites grouped by geographical proximity within the cluster i.e. Central, East, East-Central, Lake Coleman, Tarraville, West-Central). Blue bars indicate the point mean relatedness estimate and error bars report the 95% confidence intervals around the mean.

3.3.3 Effective population size

Effective population size estimates N_e for each cluster with the linkage disequilibrium method could be computed for the eastern and western clusters, although for the latter confidence intervals were very wide (Table 8). For Snake Island the computations resulted in a negative estimate (and thus infinite N_e). This numerical problem arose from the adjustment applied to correct N_e bias that is usually present in this estimator (Waples 2006). Recomputing N_e without the bias correction (which is known to slightly underestimate N_e) resulted in a finite value of 10.6. The point estimates of N_e (with a frequency threshold of 0.05) indicated that the western cluster is by far the largest, with the eastern cluster about half the size and the Snake Island population about 90% smaller.

Table 8. Effective population size estimates N_e based on linkage disequilibrium for each cluster, considering three allele frequency thresholds (0.05, 0.02 and 0.01). Last two columns report 95% confidence intervals computed with the parametric and jackknife approach; ∞ means that no valid estimate was computable.

Cluster	Frequency threshold	N_e	Parametric	Jackknife
Snake Island	0.05	10.6	∞	∞
	0.02	10.6	∞	∞
	0.01	10.6	∞	∞
Eastern	0.05	61.6	46.8, 81.7	38.2, 101.3
	0.02	61.6	46.8, 81.7	38.2, 101.3
	0.01	35.6	28.2, 44.7	3.5, 185.5
Western	0.05	110.1	45.6, 552.1	39.6, ∞
	0.02	113.3	48.4, 1628.3	43.6, ∞
	0.01	128.7	51.8, ∞	52.2, ∞

After fitting a coalescent based model to estimate effective population size changes in the Hog Deer population, analysis from the combined dataset did not detect any demographic change with the exponential growth rate estimated to be zero (95% credible interval: 0, 0.00003), indicating a stable population size. Similarly, the EBSP analysis did not support any population change.

However, the analysis of the Snake Island data with the EBSP did support a population change. The reconstruction of the demographic history of this population depicted an ancestral large and stable population size, until sometime close to the present, when a sharp 80 to 90-fold decline in the population was detected (Figure 16).
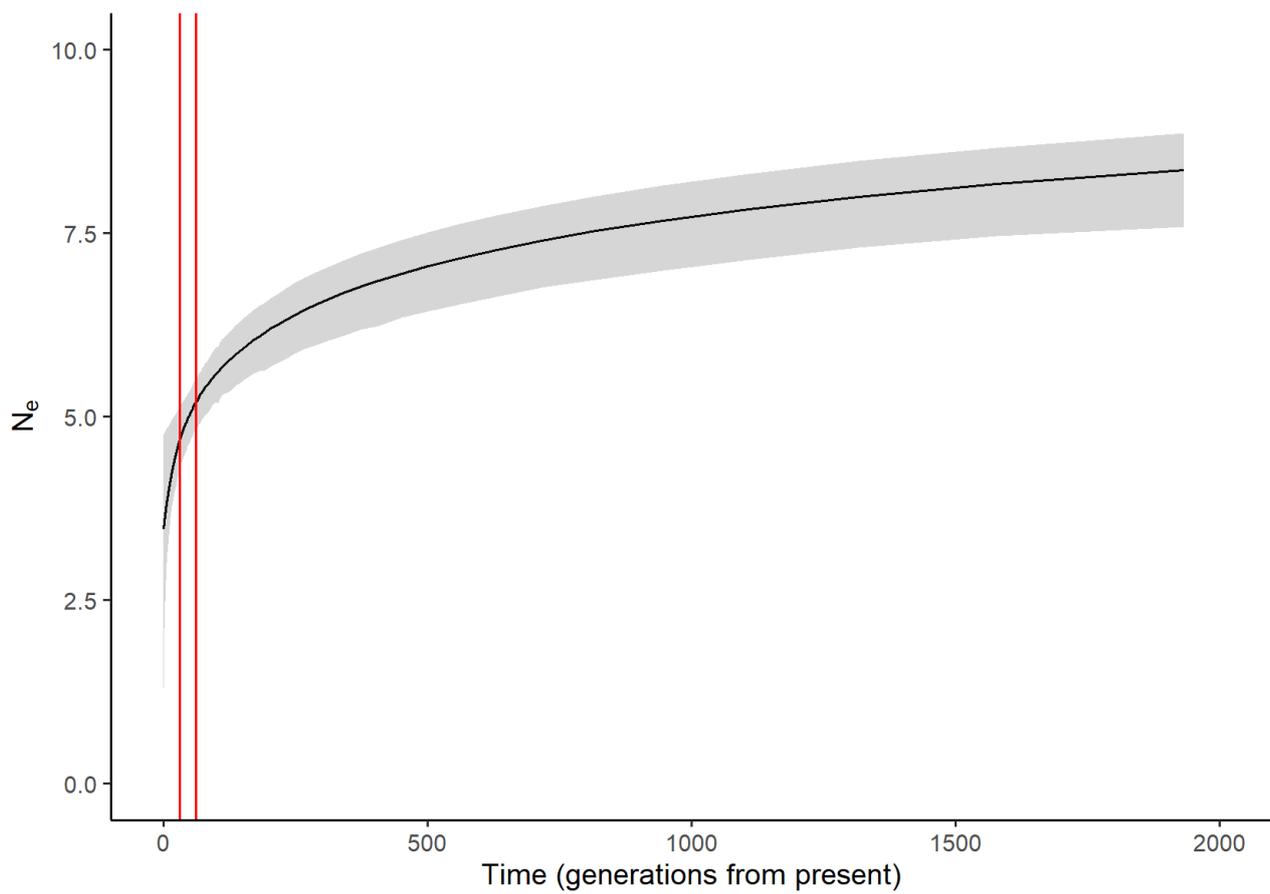
Figure 16. Reconstruction of the demographic history of the Snake Island population, with time expressed in generations (the mean age at reproduction) in the past. The vertical red lines indicate the approximate time window of Hog Deer establishment in Australia (assuming a generation time of 2.5–5 years). Hence, the left-hand side represents the present and right-hand side represents the ancestral population size (i.e. pre-introduction into Victoria). Population size on the vertical axis is expressed as log N_e.

4   Discussion

Estimates of Hog Deer abundance within the breeding range indicated that the population was low, at around 3000 individuals, and most of the population was on Wilsons Promontory. Within Wilsons Promontory, densities were highest in the southern part of the peninsula, around Oberon Bay. However, density estimates in this section of Wilsons Promontory were primarily driven by a single camera site. This site was situated in an ecotone between a forest and open grassland/herbland habitat and was likely to be a feeding area heavily used by deer, and thus may not be representative of the wider habitat in the southern part of the promontory. The predictions of Hog Deer density for the southern part of Wilsons Promontory therefore should be treated with caution, and further sampling is recommended to confirm predictions in this area.

No strong relationships were found between Hog Deer densities and the environmental variables examined. Although the most supported model included habitat type (EVC) as a predictor, the relationship was weak, with variation in Hog Deer densities being driven primarily by spatial location. This is perhaps not surprising because predictions of Hog Deer abundance were confined to the breeding range (Forsyth *et al.* 2016), which is expected to be primarily composed of preferred habitat. Hence, it appears that variation in Hog Deer densities across the breeding range are being driven by other spatially varying factors (e.g. hunting, predation, micro-habitat preferences). Of these, it is perhaps no coincidence that highest densities are to be found in the remote parts of Wilsons Promontory, where there is no legal hunting and illegal hunting would be difficult. However, as our monitoring design excluded areas of private land managed for Hog Deer, our estimates of abundance are likely to underestimate the total population of Hog Deer in areas outside Wilsons Promontory. Future monitoring programs should attempt to represent the contributions of managed areas to the Hog Deer population to enable improved estimates of abundance for the region.

Hog Deer were not detected by cameras located east of Lakes Entrance. While there are certainly records of Hog Deer between Lakes Entrance and Point Hicks (Forsyth *et al.* 2016), the evidence suggests densities are likely to be very low. Hence, despite being introduced into Gippsland Lakes region over 150 years ago (Mayze and Moore 1990), Hog Deer have not successfully colonised East Gippsland. One reason for the low densities of Hog Deer east of Lakes Entrance may be due to high predation rates by wild dogs. Wild dogs were detected on cameras between Lake Tyers and Cann River while no wild dogs were detected west of Lakes Entrance. Given the size of Hog Deer is comparable 'to a large sheep' (Mayze and Moore 1990), it is likely that they would be highly vulnerable to wild dog predation. Since wild dogs are known to be abundant in the forests of East Gippsland, we hypothesise that high predation rates in this area will preclude Hog Deer from ever reaching high densities in suitable habitat in the area from Lakes Entrance to the NSW border. In contrast, an adult Sambar Deer is much larger and less vulnerable to wild dog predation, and estimated Sambar Deer densities were comparatively high in areas east of Lakes Entrance. An alternative hypothesis is that competition with Sambar Deer is largely responsible for the lack of Hog Deer east of Lakes Entrance.

The Hog Deer genetic data showed very little variability and weak spatial organisation, with three clusters (i.e. subpopulations) being identified. Of these, Snake Island was the only subpopulation that was strongly differentiated, while the other two subpopulations showed only a mild differentiation and a high level of admixture. It is very likely that the differentiation of individuals on Snake Island from individuals on the mainland is the result of genetic drift, because the Snake Island population is small and isolated. However, in consideration of the fact that the original founding stock used to establish Australian Hog Deer populations had two origins (Sri Lanka and mainland India), the hypothesis that, by chance, individuals from one of these two sources mostly colonised Snake Island and those form the other source colonised the mainland (with the two mainland clusters being a contemporaneous separation in progress) cannot be completely excluded.

Several individuals sampled on the mainland had a similar genetic profile to individuals on Snake Island. These detected migrants are likely to be the result of Hog Deer translocations that occurred in the 1980s, when approximately 90 Hog Deer from Snake Island were released at Dutson Downs (Mayze and Moore 1990). Based on tags and collar retrieval in successive years, it is known that these animals established in the area and moved within the Gippsland Lakes region. Similarly, animals from Sunday Island and Serendip

Wildlife Park were released at Blond Bay between 1985 and 1987. Released animals were sighted in the area and in Bengworden in successive months, providing a likely explanation of the similarity of the genetic profiles of individuals on Sunday Island with those from Blond Bay and surrounding areas.

The degree of admixture between the two mainland clusters would normally indicate regular, long-distance migration events. However, the history of translocations is a confounder, making it difficult to identify natural dispersal distances based on the Bayesian assignment tests. Nevertheless, a spatial autocorrelation analysis of the eastern cluster supported a sex-biased dispersal, since females sampled within 60 km of each other were more related than expected by chance. Based on this analysis, dispersal distance in females could be as large as 70 km. The negative spatial correlation coefficient for distances between 80 and 120 km further support the inference that related females do not disperse to these distances. On the other hand, males are closely related at distances < 40 km, suggesting that at least a proportion of related males only disperse short distances or that juveniles (i.e. pre-dispersal age) were included in the data. The genetic signal quickly reduced over larger distances, supporting the hypothesis that male dispersal encompassed distances up to 140 km. These results are in line with dispersal distances observed during the Dutson Downs translocation, where more than 80% of the females dispersed less than 30 km, while males displayed a much more variable but generally longer dispersal distances (Mayze and Moore 1990).

The estimation of effective population size N_e in the three Hog Deer clusters were all very low. In conservation settings, an effective population size of about 100 is considered viable only in the short term, with the general recommendation being at least 1000 for long term conservation. Furthermore, any consideration of Snake Island Hog Deer being an 'insurance population' is flawed, as this population is genetically isolated, has a very small N_e and has very limited genetic diversity.

The demographic analysis of the Snake Island data was very informative. While it is well known that the Australian Hog Deer population went through a severe bottleneck because there were only 12 founding individuals, this analysis allowed us to quantify the genetic consequences of this bottleneck. This analysis was also able to estimate the ancestral effective population size (i.e. pre-Australia), which had a 90-fold reduction in N_e, coinciding with the translocation of Hog Deer to Australia.

To maintain the Hog Deer population for sustainable hunting, actions to prevent the onset of fitness reduction (e.g. lower breeding rate) should be considered. Common management actions that are usually considered to counteract genetic drift include increasing gene flow, maximising a breeder's contribution to the genetic pool, and increasing the population size. However, for Hog Deer in Gippsland the first two are probably of limited value because the genetic distance between clusters is limited, so the benefit of supplementation programs would be limited. Breeding manipulation is also not realistic for free-ranging populations. This leaves demographic and habitat management (to increase carrying capacity) as the only feasible options. Monitoring of fitness parameters (reproductive rates, longevity, etc.) could be used to trigger (or intensify) these management actions.

4.1   Conclusion

The Hog Deer population inhabiting three core areas in coastal Gippsland between Lower Tarwin and Point Hicks was predicted to be low, with less than 1000 individuals estimated to inhabit areas outside Wilsons Promontory. However, as our monitoring design did not include areas of private land managed for Hog Deer (e.g. Sunday Island and some mainland private properties), our estimates of abundance are likely to underestimate the total population of Hog Deer in this region. Future monitoring programs should attempt to represent the contributions of these managed areas to the Hog Deer population to improve the estimates of abundance for the region.

The Hog Deer population is characterised by low genetic variability and low effective population size, putting it at a high risk of inbreeding depression, especially for the Snake Island and eastern subpopulations. Hence the long-term viability of the Gippsland Lakes and Snake Island subpopulations of Hog Deer are vulnerable to potential impacts such as overharvesting, disease or habitat loss.

There are also significant barriers to successful dispersal and colonisation of suitable habitat outside the breeding range. For example, the likelihood of significant breeding populations establishing east of Lakes Entrance are hypothesised to be limited because of predation by wild dogs and/or competition with Sambar

Deer in this area. If Hog Deer are to be managed as a sustainable hunting resource, management of the current extant core subpopulations is essential to ensure their long-term viability.

4.2   Recommendations

The following recommendations are made for the Hog Deer population inhabiting publicly accessible areas where hunting is permitted (e.g. State Game Reserves from Jack Smith Lake to Ewing Morass) as well as areas designated for balloted hunting (e.g. Snake Island, Blond Bay, Boole Poole).

- Investigate the sustainable level of hunting to ensure the long-term viability of Hog Deer populations in these areas.

- Monitor Hog Deer populations periodically to determine population trends, especially in State Game Reserves and areas designated for balloted hunting. Future monitoring should include private land managed for Hog Deer to determine the contributions of these areas to the overall Hog Deer population.

- Manage some State Game Reserves to increase the carrying capacity and productivity of Hog Deer populations where this is feasible and does not impact other biodiversity values.

- Undertake studies of fitness parameters, such as reproductive and survival rates of Hog Deer, and use the results to trigger management actions.

- Develop a management plan for Hog Deer (on publicly accessible areas where hunting is permitted) that critically examines the feasibility and biodiversity impacts of potential management actions to ensure the long-term viability of core subpopulations.

References

Banks, S. C., and Peakall, R. (2012). Genetic spatial autocorrelation can readily detect sex-biased dispersal. *Molecular Ecology* **21**, 2092–2105.

Brown, A. H. D., and Weir, B. S. (1983). Measuring genetic variability in plant populations. **In** Isozymes in plant genetics and breeding. Part A (eds S. D. Tanksley and T. J. Orton.) pp. 219–239. (Elsevier: Amsterdam.)

Buckland, S. T., Anderson, D. R., Burnham, K. P., and Laake, J. (1993). 'Distance sampling: estimating abundance of biological populations'. (Chapman & Hall: London.)

Buckland, S. T., Summers, R. W., Borchers, D. L., and Thomas, L. (2006). Point transect sampling with traps or lures. *Journal of Applied Ecology* **43**, 377–384.

Burnham, K. P., and Anderson, D. R. (1998). 'Model selection and inference: a practical information-theoretic approach'. (Springer-Verlag: New York, USA.)

Candy, S. G. (2004). Modelling catch and effort data using generalised linear models, the Tweedie distribution, random vessel effects and random stratum-by-year effects. *CCAMLR Science* **11**, 59–80.

Côté, S. D., Rooney, T. P., Tremblay, J.-P., Dussault, C., and Waller, D. M. (2004). Ecological impacts of deer overabundance. *Annual Review of Ecology, Evolution, and Systematics* **35**, 113–147.

Davis, N. E., Bennett, A., Forsyth, D. M., Bowman, D. M. J. S., Lefroy, E. C., Wood, S. W., Woolnough, A. P., West, P., Hampton, J. O., and Johnson, C. N. (2016). A systematic review of the impacts and management of introduced deer (family Cervidae) in Australia. *Wildlife Research* **43**, 515.

Davis, N. E., Coulson, G., and Forsyth, D. M. (2008). Diets of native and introduced mammalian herbivores in shrub-encroached grassy woodland, south-eastern Australia. *Wildlife Research* **35**, 684–694.

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., and Ovenden, J. R. (2014). NeEstimator v2.0: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources* **14**, 209–214.

Dormann, C. F., Calabrese, J. M., Guillera-Arroita, G., Matechou, E., Bahn, V., Bartoń, K., Beale, C. M., Ciuti, S., Elith, J., Gerstner, K., Guelat, J., Keil, P., Lahoz-Monfort, J. J., Pollock, L. J., Reineking, B., Roberts, D. R., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Wood, S. N., Wüest, R. O., and Hartig, F. (2018). Model averaging in ecology: a review of Bayesian, information-theoretic, and tactical approaches for predictive inference. *Ecological Monographs* **88**, 485–504.

Drummond, A., and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**, 214.

Falush, D., Stephens, M., and Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587.

Forsyth, D. M., Stamation, K., and Woodford, L. (2016). Distributions of Fallow Deer, Red Deer, Hog Deer and Chital Deer in Victoria. Unpublished Client Report for the Biosecurity Branch, Department of Economic Development, Jobs, Transport and Resources. Arthur Rylah Institute for Environmental Research, Department of Environment, Land, Water and Planning, Heidelberg, Victoria.

Foster, S. D., Hosack, G. R., Lawrence, E., Przeslawski, R., Hedge, P., Caley, M. J., Barrett, N. S., Williams, A., Li, J., Lynch, T., Dambacher, J. M., Sweatman, H. P. A., and Hayes, K. R. (2017). Spatially balanced designs that incorporate legacy sites. *Methods in Ecology and Evolution* **8**, 1433–1442.

Frankham, R. (1995). Conservation genetics. *Annual Review of Genetics* **29**, 305–327.

Frankham, R., Bradshaw, C. J. A., and Brook, B. W. (2014). Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation* **170**, 56–63.

Gilbert, K. J., and Whitlock, M. C. (2015). Evaluating methods for estimating local effective population size with and without migration. *Evolution* **69**, 2154–2166.

Griffiths, R. C., and Simon, T. (1994). Sampling theory for neutral alleles in a varying environment. *Philosophical Transactions: Biological Sciences* **344**, 403–410.

Guillot, G. (2008). Inference of structure in subdivided populations at low levels of genetic differentiation—the

correlated allele frequencies model revisited. *Bioinformatics* **24**, 2222–2228.

Guillot, G., Mortier, F., and Estop, A. (2005). Geneland: a computer package for landscape genetics. *Molecular Ecology Notes* **5**, 712–715.

Hartl, D. L., and Clark, A. G. (1997). 'Principles of population genetics' 3rd ed. (Sinauer Associates: Sunderland.)

Heled, J., and Drummond, A. (2008). Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology* **8**, 289.

Howe, E. J., Buckland, S. T., Després-Einspenner, M.-L., and Kühl, H. S. (2017). Distance sampling with camera traps. *Methods in Ecology and Evolution* **8**, 1558–1565.

Jones, A. T., Ovenden, J. R., and Wang, Y. G. (2016). Improved confidence intervals for the linkage disequilibrium method for estimating effective population size. *Heredity* **117**, 217.

Mayze, R. J., and Moore, G. I. (1990). 'The Hog Deer'. (Australian Deer Research Foundation, Melbourne, Victoria.)

Menkhorst, P. W. (1995). 'Mammals of Victoria: Distribution, Ecology and Conservation'. (Oxford University Press: Melbourne, Victoria.)

Miller, D. L., Burt, M. L., Rexstad, E. A., and Thomas, L. (2013). Spatial models for distance sampling data: Recent developments and future directions. *Methods in Ecology and Evolution* **4**, 1001–1010.

Miller, D. L., Rexstad, E. A., Burt, M. L., and Bravington, M. V. (2019a). dsm: Density Surface Modelling of Distance Sampling Data. R package version 2.2.18. Available at: http://github.com/DistanceDevelopment/dsm

Miller, D. L., Rexstad, E., Thomas, L., Marshall, L., and Laake, J. L. (2019b). Distance Sampling in R. *Journal of Statistical Software* **89**, 1–28.

Moloney, P. D., and Turnbull, J. D. (2018). Estimates of the 2017 deer harvest in Victoria. Results from surveys of Victorian Game Licence holders in 2017. Unpublished Client Report for the Game Management Authority. Arthur Rylah Institute for Environmental Research, Department of Environment, Land, Water and Planning, Heidelberg, Victoria.

Palstra, F. P., and Fraser, D. J. (2012). Effective/census population size ratio estimation: a compendium and appraisal. *Ecology and Evolution* **2**, 2357–2365.

Peakall, R. O. D., and Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288–295.

Peakall, R., Ruibal, M., and Lindenmayer, D. B. (2003). Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, Rattus fuscipes. *Evolution* **57**, 1182–1195.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Queller, D. C., and Goodnight, K. F. (1989). Estimating relatedness using genetic markers. *Evolution* **43**, 258–275.

R Development Core Team (2018). R: A language and environment for statistical computing. Available at: http://www.r-project.org

Slee, K. J. (1985). 1985 Hog Deer habitat survey. A report prepared for the Australian Deer Association, Victoria.

Smouse, P. E., and Peakall, R. (1999). Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* **82**, 561–573.

Smouse, P. E., Peakall, R., and Gonzales, E. (2008). A heterogeneity test for fine-scale genetic structure. *Molecular Ecology* **17**, 3389–3400.

Waples, R. S. (2006). A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics* **7**, 167–184.

Waples, R. S., Do, C., and Chopelet, J. (2011). Calculating Ne and Ne/N in age-structured populations: a hybrid Felsenstein-Hill approach. *Ecology* **92**, 1513–1522.

Waples, R. S., and England, P. R. (2011). Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. *Genetics* **189**, 633.

Wood, S. N. (2017). Generalized additive models: an introduction with R. 2nd edn. (CRC Press, Taylor and Francis Group: Boca Raton, Florida.)

Wu, C. H., and Drummond, A. J. (2011). Joint inference of microsatellite mutation models, population history and genealogies using transdimensional Markov Chain Monte Carlo. *Genetics* **188**, 151–164.

Appendix 1: Population abundance

Table A1: Ecological Vegetation Classes (EVC) groupings, description and the derived group variable used to model spatial variation in Hog Deer abundance across the Hog Deer breeding range.

EVC	Description	Derived Group
1	Coastal Scrubs Grasslands and Woodlands	Coastal scrub/shrublands
2	Heathy Woodlands	Heathlands/Heathy woodlands
3	Lowland Forests	Woodlands/Forest
4	Box Ironbark Forests or dry/lower fertility Woodlands	Woodlands/Forest
5	Lower Slopes or Hills Woodlands	Woodlands/Forest
6	Dry Forests	Woodlands/Forest
7	Wet or Damp Forests	Woodlands/Forest
8	Riparian Scrubs or Swampy Scrubs and Woodlands	Coastal scrub/shrublands
9	Rainforests	Woodlands/Forest
10	Montane Grasslands, Shrublands or Woodlands	Woodlands/Forest
11	Sub-alpine Grasslands, Shrublands or Woodlands	Woodlands/Forest
12	Plains Grasslands and Chenopod Shrublands	Woodlands/Forest
13	Plains Woodlands or Forests	Woodlands/Forest
14	Riverine Grassy Woodlands or Forests	Woodlands/Forest
15	Herb-rich Woodlands	Herb-rich woodlands
16	Heathlands	Heathlands/Heathy woodlands
18	Wetlands	Wetlands
19	Salt-tolerant and/or succulent Shrublands	Coastal scrub/shrublands
20	Rocky Outcrop or Escarpment Scrubs	Coastal scrub/shrublands
99	No native vegetation recorded	NA

Table A2. Estimates of the total abundance ($\hat{N}$) and density ($\hat{D}$) of Hog Deer in native vegetation in the Hog Deer breeding range (1762 km^2) for the two most supported models as well as their average. se – standard error; cv – coefficient of variation; LCL, UCL – lower and upper 95% confidence intervals; sd – standard deviation.

Model	$\hat{N}$	se($\hat{N}$)	cv($\hat{N}$) %	95% LCL	95% UCL	$\hat{D}$ (mean)	$\hat{D}$ (median)	sd($\hat{D}$)
Model 6	3802	796	20.9	2533	5706	2.2	0.1	5.8
Model 3	2199	705	32.0	1191	4059	1.2	0.3	2.3
Model average	3000	745	24.8	1858	4845	1.7	0.2	4.4

Table A3. Results of model selection among six models for the detection function fitted to the Sambar Deer distances recorded from each camera. AIC–Akaike's Information Criterion, ΔAIC–difference between AIC and the model with the lowest AIC.

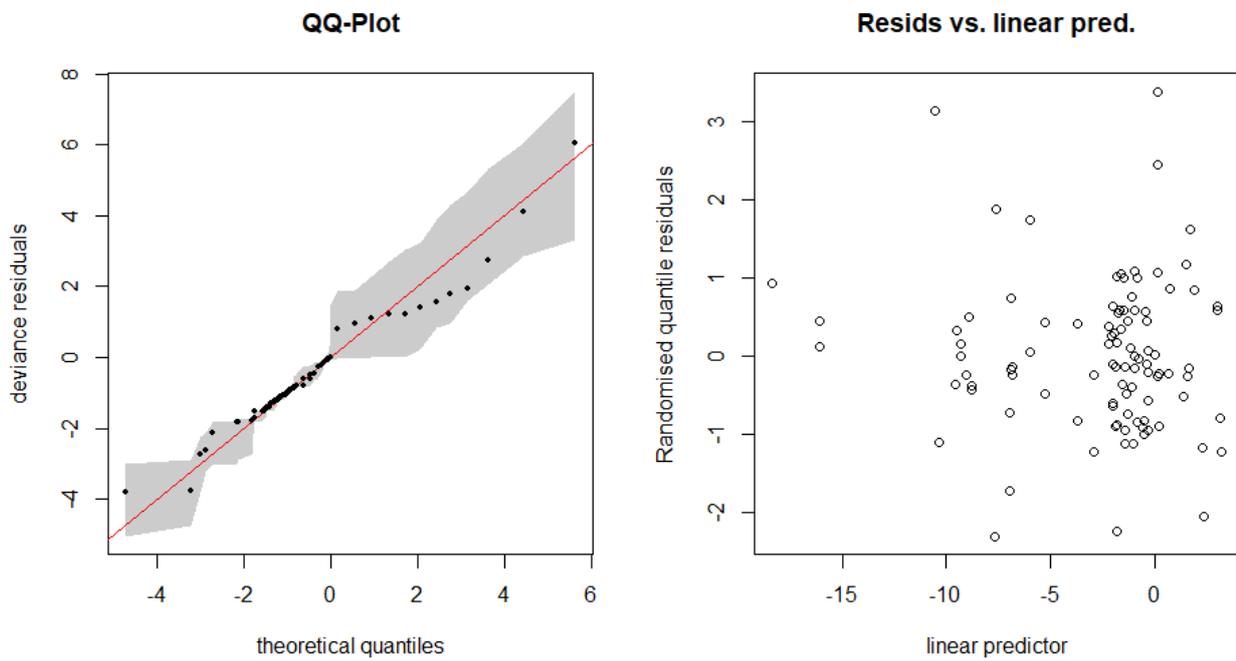| Detection model	Adjustments	AIC	ΔAIC
Half-normal	none	447.0	2.2
Half-normal	2nd order cosine	448.5	3.7
Half-normal	2nd and 3rd order cosine	450.4	5.6
Hazard-rate	None	448.2	3.4
Hazard-rate	2nd order cosine	444.8	0.0
Hazard-rate	2nd and 3rd order cosine	446.8	2.0

Figure A1. Quantile–quantile (QQ) plot (left) and randomised residual plot (right) for the most supported dsm model of Sambar Deer abundance.

Appendix 2: Population genetic methods

DNA extraction and genotyping

DNA extractions were carried out using a Qiagen DNeasy Blood and Tissue Kit, and DNA was quantified using an Invitrogen Qubit 2.0 Fluorometer. Each sample was diluted to 1–2 ng/µL and genotyped in three multiplexes as described by Hill *et al.* (unpublished data) (Table A4). PCRs were performed in 25 µL reactions, comprised of 12.5 µL of Qiagen Multiplex PCR master mix, 3.5 µL of primer mix, 3.5 µL of Q-solution (Qiagen), 3.5 uL of RNase-free water, and 2 µL of 1–2 ng/µL template DNA. PCR cycling conditions were 95°C for 15 minutes, 35 cycles of 94°C for 30 seconds, annealing step for 90 seconds, 72°C for 2 minutes, followed by a final extension of 60°C for 30 minutes. Annealing temperatures were 58°C for multiplexes 1 and 2, and 61°C for multiplex 3. Positive and non-template controls were run throughout each PCR. Samples were sent for genotyping to the Australian Genome Research Facility (AGRF). Genotypes were scored using the Microsatellite plugin in *Geneious 9.1.8*.

Table A4. Details of genetic markers used for this study.

Locus	Primer sequence (5'-3')	Repeat Motif	Size (bp)	Dye	Multiplex	Reference
ApoV19	F: GCCTTCCTATCCAAGAACAGG R: CGGCACAGTGGAGATACTTG	$(TTTG)_6$	492-500	6-FAM	1	Hill et al. unpublished
ApoV104	F: CACCCAGGCTCCCTTACTTC R: GAGCATGGATTCAAGGGCAC	$(TCC)_7$	352-355	6-FAM	1	Hill et al. unpublished
ApoV103	F: TCGCACAGAGTCAGGCAC R: CATGTGGCCCTTGTGACATC	$(AGC)_8$	202-238	6-FAM	1	Hill et al. unpublished
ApoV54	F: TTATCCATTTCGCCATGCCC R: GGTGGGAGGGTCTCTAAAGC	$(TTTA)_5$	480-484	VIC	1	Hill et al. unpublished
Apo4	F: GGCAGGCAGATTCTTAT R: AGCAGCCAAATGGACTA	$(AC)_{18}$	216-226	VIC	1	Lian et al. 2008
ApoV135	F: CACCCTTGTTCCATGAGAGC R: ACACTCCAGGGAATCAGCTG	$(CTTT)_7$	118-122	VIC	1	Hill et al. unpublished
ApoV53	F: TCCAGAAGCATTCCTGACCC R: TGTGATCCAGGTGAGAGCAG	$(AAAC)_5$	471-475	PET	1	Hill et al. unpublished
Apo5	F: AGAAAGGTTACCGACTC R: CTGCTGCATGAAGAATG	$(GT)_{20}$	190-196	6-FAM	2	Lian et al. 2008
Apo7	F: CTTGATTTGTCCTACCCTCT R: ACTATGCCATGCCCTAT	$(TC)_8(AC)_{21}$	205-213	VIC	2	Lian et al. 2008
ApoV133	F: CTGGAATCTGACCCGAGGAG R: AGCCAGTTACAGTCACCAGG	$(CATT)_7$	168-172	VIC	2	Hill et al. unpublished
ApoV94	F: CTGCTATGTCACTGCATGCC R: CGCCAGACTGAAGAACTCAC	$(AATG)_5$	468-472	PET	2	Hill et al. unpublished
ApoV109	F: TGTCACTGCACACTGTTTACC R: CCTTCCTATCAGCTTTGCAGG	$(AGC)_8$	237-246	PET	2	Hill et al. unpublished
ApoV145	F: TGAGAATGTCCTGGGATGCC R: CACCCACTCATTCATCTACGC	$(GGAT)_7$	206-210	PET	2	Hill et al. unpublished
ApoV118	F: TTGCCTGGTTTCACTTTGCC R: CACCGGTGACATCCTGGG	$(TCCC)_7$	126-130	PET	2	Hill et al. unpublished
ApoV17	F: GTCCCACTCAACTCTACGGG R: AACCCATGTCCTAACCCTGG	$(ATGG)_7$	518-522	6-FAM	3	Hill et al. unpublished
ApoV127	F: ACACACAGCAATGAAGACCAG R: CAGTGTTTCTTCGTCCTTGAGG	$(AAAC)_7$	132-140	6-FAM	3	Hill et al. unpublished

Population structure

In order to estimate and potentially account for the presence of null alleles in the population structure analyses, null allele frequencies were estimated for each locus concurrently with the estimation of the number of clusters in both Structure and Geneland.

A sample could have missing data as a result of specific alleles not amplifying (true null alleles) or for other reasons such as poor quality DNA or other technical issues. As a result, estimating the null allele frequencies in both software requires a judgement about how to handle double missing data for any given sample-locus combination. We considered that when a sample had missing data for three or more loci, these were likely due to poor DNA quality or other issues not related to null alleles. However, we flagged as null alleles situations where the genotypes could be obtained for all loci but one or two. The only exception was for locus ApoV17, because it was known that in a batch of samples this locus was not genotyped because of a technical error with the fragment analyser. To further investigate whether the potential presence of null alleles would bias the analyses, we repeated Structure analyses without the estimation of null frequencies.

In Structure, because the allele frequencies were skewed across many loci because of the presence of rare alleles, rather than using the default (fixed) value for the vector of the parameter λ ($\lambda = 1$, which assumes a balance distributions of allele frequencies), we initially used the model with no admixture and non-correlated allele frequencies to estimate λ with K = 1, and then applied the estimated value to all other analyses in this software to estimate the value of K. This is because concurrently estimating K and λ often leads to identifiability issues. We also evaluated whether results would change if the admixture parameter (α) was allowed to vary across genetic clusters as opposed to the default setting where it is set to be identical in all populations. Five runs with a burn in of 250 000 iterations followed by 100 000 iterations was conducted for each K values (from 1 to 8). Adequate settings the Markov Chain Monte Carlo computation (MCMC) was confirmed by inspecting the traces of the estimated parameters. Results were summarised using structure harvester and the Evanno method was used to select the most likely K value after having ensured by inspecting the plot of the mean log likelihood that K = 1 could be discounted (because the Evanno method calculates a second-order rate of change statistic, K = 1 cannot be evaluated with this method).

In Geneland we ran four analyses for each model for 5×10^6 iterations with up to 20 km uncertainty in sample locations. Outputs files were read and collated with the R package data.table and trace plots, Effective Sample Size (ESS) and the Gelman and Rubin's convergence diagnostic were computed with the Coda package. It is also important to highlight that the admixture model in Geneland is implemented with a different strategy from Structure. Based on a simulation study, the no admixture model was found to correctly identify the number of K even in presence of admixture. Hence, the admixture model in Geneland fixes K to the estimated value from the analysis in the model with no admixture and estimates only the ancestry proportion, the admixture parameter and the spatial scale parameter. The scale parameter indicates the degree of spatial organisation of the cluster with large values suggesting little spatial cluster. The width of the spatial cline is about four times the spatial parameter.

The parameter λ, estimated to be 0.4715 in Structure, was applied as a fixed parameter to all successive analyses in this software. Models with no admixture (correlated and uncorrelated allele frequencies), as well as the admixture model with correlated frequencies, suggested K = 3 as the most likely number of clusters (Figure A2). Individual assignments (Figure A3) and geographical distribution of clusters as well as proportional membership of sampling sites to the three inferred cluster (Figure A2) were very similar across all these models. In these results, the western sites (up to Hedley) grouped together forming a relatively homogeneous cluster. The second cluster was almost exclusively formed by Snake Island, with the exception of a few individuals assigned to this population from other sites (Figure A3), and the third cluster included the eastern sites. In the latter, an admixture gradient with the western cluster was evident, with sites at the eastern extreme of the distribution and central area almost exclusively made up of eastern individuals while Sunday Island and mainland sites in its proximity as well as Gippsland Lakes sites and Stratford showed a high degree of admixture with the 'western' cluster. In line with these results, in the model where α was allowed to vary, it was estimated to be about 0.1 for the eastern and western clusters, and about 0.03 for Snake Island. However, this and the admixture with uncorrelated allele frequencies model suggested K = 2 as the most likely number of clusters. In this case, individuals sampled on the main land and Sunday Island were grouped together and separated from those sampled on Snake Island with the exception of the same individuals that were identified as migrants from Snake Island in the previous models in Loch Sport and Stratford.

Demographic analysis

Changes in the effective population size N_e were estimated using the parametric exponential growth (Griffiths and Simon 1994) and the non-parametric "Extended Bayesian Skyline" (EBSP) (Heled and Drummond 2008) models. Models were fitted to the data using the unbiased, one phase mutational model with a constant rate implemented in a fully Bayesian framework as described by Wu and Drummond (2011) in `BEAST 1.8` (Drummond and Rambaut 2007). We allowed each locus to vary its mutation rates by using a gamma prior with a shape=1 and scale=0.001, but we constrained the mean mutation rates across all loci to be 5 X 10^{-4} mutations/locus/generation, which is the mean mutation rate estimated for mammals. We ran two independent analyses for each model for 4 X 10^8 iterations and sampled the parameters every 50,000. We confirmed convergence, established adequate length for the burn in phase, and ensure that all parameter estimates had adequate effective sample size with the software `Tracer` (Drummond and Rambaut 2007). These analyses were conducted on two datasets, one with the samples from Snake Island alone and one with a subset of samples (n=81) stratified by sampling locations across all Victoria. To this end, the latter analysis should be viewed as a species-wide assessment of demographic changes. By doing this, it is assumed that any separation in subpopulations occurred recently and was negligible.
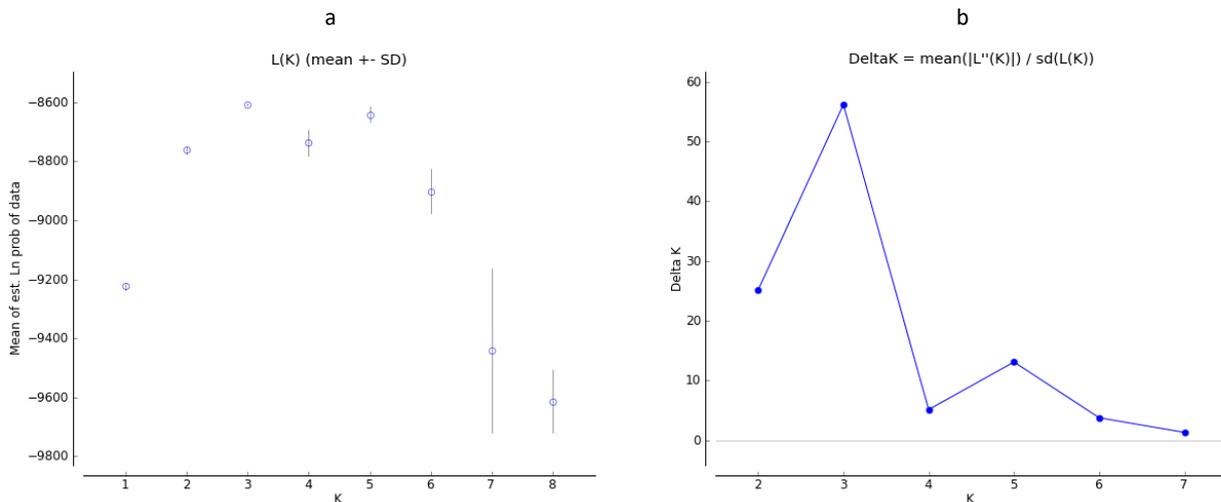

Figure A2. Summary of results for the admixture model with correlated allele frequencies in Structure: (a) plot of the mean log likelihood of the data given K; (b) plot of the deltaK statistics (Evanno method).
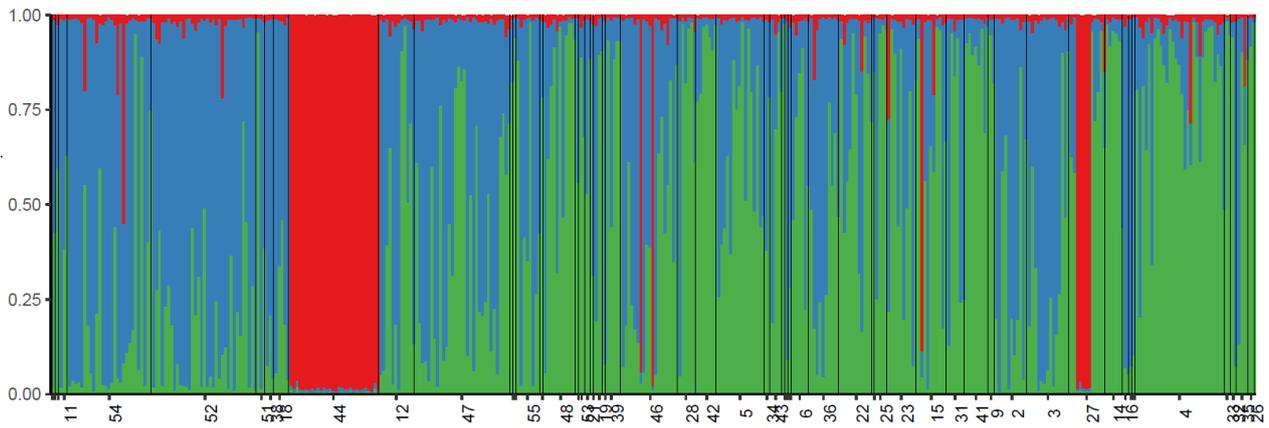
Figure. A3. Bar plot of inferred ancestry (vertical axis) of individuals (each bar on the plot) with the admixture model with correlated allele frequencies in Structure grouped by sampling site (only site with more than 6 samples are labelled, see Table A4 for a list of sites) sorted from west to east. Black vertical lines mark each sampling site. Genetic clusters are colour-coded. Note how all animals sampled on Snake Island (sampling site 44) belong to the same cluster and are well differentiated from all other sampling sites, except five individuals at Loch Sport (27), two at Stratford (46), one at Golden Beach (15), and an F1 (c. 50% of the genome at Snake Island and c. 50% from the western (blue) cluster) at Yanakie (54).

Table A4. List of sampled sites and number of samples used in the genetic analysis.

Site no.	Site	n
1	Alberton	1
2	Bengworden	11
3	Blond Bay	14
4	Boole Poole	30
5	Clydebank	16
6	Dowds Morass	6
7	Dutson	1
8	Dutson Downs	1
9	Eel Farm	2
10	Fish Creek	1
11	Foster	3
12	Gelliondale	12
13	Gellions Run	1
14	Gippsland Lakes Coastal Park	6
15	Golden Beach	10
16	Goon Nure	2
17	Heart Morass	1
18	Hedley	5
19	Jack Smith Lake	3
20	Jack Smith SGR	1
21	Kilmany	2
22	Lake Coleman	11
23	Lake Reeve	10
24	Lake Victoria	1
25	Lake Wellington	4
26	Lakes Entrance	2
27	Loch Sport	12
28	Longford	6
29	Manns Beach	1
30	Marlay Point	1
31	Meerlieu	6
32	Metung	2
33	Mossiface	2
34	Munro	2
35	Nungurner	4
36	Perry Bridge	10
37	Port Albert	1
38	Port Welshpool	3
39	Sale	5
40	Sandy Point	1
41	Seacombe	8
42	Seaspray	7
43	Sentosa	4
44	Snake Island	30
45	Stradbroke	1
46	Stratford	19
47	Sunday Island	32
48	Tarraville	11
49	Tarwin Lower	1
50	Wattle Point	1
51	Welshpool	3
52	Wilsons Prom	35
53	Woodside	2
54	Yanakie	28
55	Yarram	8