Age is not just a number: how incorrect ageing impacts close-kin mark-recapture estimates of population size

Felix T. Petersma[1]*, Len Thomas[1], Danielle Harris[1], Darcy Bradley[2], Yannis Papastamatiou[3]

[1]Centre for Research into Environmental and Ecological Modelling, University of St Andrews, St Andrews, UK

[2]Bren School of Environmental Science & Management, University of California, Santa Barbara, California, USA

[3]Institute of Environment, Department of Biological Sciences, Florida International University, North Miami, Florida, USA

Abstract

Population size is a key parameter for the conservation of animal species. Close-kin mark-recapture (CKMR) relies on the observed frequency and type of kinship among individuals sampled from the population to estimate population size. This approach requires being able to determine the age of sampled individuals. One common approach, particularly in fish studies, is to measure animal length and using an assumed age-length relationship (a 'growth curve'). We used simulation to test the effect of misspecifying the length measurement error and the growth curve on population size estimation. Simulated populations represented two fictional shark species, one with a relatively simple life history and the other with a more complex life history based on the grey reef shark (*Carcharhinus amblyrhynchos*).

We estimated sex-specific adult abundance, which we assumed to be constant in time. We observed small median biases in these estimates ranging from 1.35 to 2.79% when specifying the correct measurement error and growth curve, with true coefficients of variation between 21.56 and 28.50%. Introducing error via misspecified growth curves resulted in changes in the magnitude of the estimated total population, with upwards shifts negatively biasing abundance estimates. Over- and underestimating the length measurement error did not introduce a bias and had negligible effect on the variance in the estimates.

Our findings show that assuming an incorrect length measurement error has little effect on estimation, but having an accurate growth curve is crucial for CKMR whenever ageing is based on

*ftp@st-andrews.ac.uk

length measurements. If ageing could be biased, researchers should be cautious when interpreting CKMR results and consider the potential biases arising from inaccurate age inference.

Keywords: Grey reef shark, *Carcharhinus amblyrhynchos*, abundance estimation, simulation, measurement error, von Bertalanffy growth function

1 Introduction

Close-kin mark-recapture (CKMR) is a method for estimating population size and other key parameters such as fecundity using data on the relatedness of individuals sampled from the population (Skaug, 2001; Bravington et al., 2016b). The key rationale is that small populations will tend to contain a higher proportion of closely-related individuals than large populations.

One of the main advantages of CKMR over capture-recapture (Otis et al., 1978) and its extensions such as spatial capture-recapture (Borchers and Efford, 2008), is that it can be applied in cases when sampling is necessarily lethal, such as fisheries, where alternative metrics are often relative (e.g., catch-per-unit-effort) and potentially unreliable (Bravington et al., 2016a). This is because CKMR does not require the recapturing of individuals, but rather their genetic markers. Offspring share genetic information with their parents (hence 'kin'), thus they 'mark' their parents when born; through modern genetics we can compare sampled individuals with one another to see if these marks are 'recaptured'. So far, CKMR has been developed for parent-offspring pairs (POPs; e.g., Bravington et al., 2016a; Ruzzante et al., 2019; Trenkel et al., 2022), half-sibling pairs (HSPs; e.g., Hillary et al., 2018; Bravington et al., 2019; Patterson et al., 2022), and the combination of both (e.g., Bradford et al., 2018). Regardless of the type of kinship used, CKMR on its own can only ever be used to estimate abundance of the breeding population. Here, we focus only on POPs. For any comparison between two individuals, the probability that a potential offspring truly is the offspring of the parent is inversely related to the number of mature animals alive in the birth year of the offspring. Probabilities of finding a kin pair are expressed as a function of the expected relative reproductive output (ERRO) of the parent in the year that the offspring was conceived. This approach is parent-centric, as it starts from the point that the parent is sampled and then formulates a probability for a PO relationship (an alternative, offspring-centric formulation was proposed by Skaug (2017)). In the simplest scenario, the probability of any adult being the parent of a juvenile reduces to one over the number of potential parents; in reality, this probability is often more complicated, e.g., when reproductive output is related to age, or when there is stock structure or population trend.

To use relatedness to estimate adult population abundance, it is therefore essential to accurately age the studied animals because birth year is derived from their age. However, ageing can be challenging: for example, epigenetic ageing requires calibration using individuals of known age (Polanowski et al.,

2014; De Paoli-Iseppi et al., 2017), which is not always possible. Ageing via otoliths, which are calcium carbonate structures in the inner ear, can be relatively accurate (Campana, 2001) but requires lethal sampling. Alternatively, length can sometimes be used to infer age through growth curves, and length is often recorded during sampling. Accurate estimates for growth curves of the studied species are not always available, and age as a function of length (age-at-length) can vary substantially between populations of the same species (e.g., Bradley et al., 2017). Moreover, length measurements often involve measurement error. Through simulation, we explore the effects of incorrect ageing on the CKMR estimator. Error is introduced in two ways: i) through misspecified growth curves, and ii) through incorrect length measurement, i.e., measurement error. In reality, error could also be (and almost surely is) introduced through natural variation in length-at-age. In this study we assume that all sharks follow the growth curve perfectly; however, one could readily interpret the length measurement error as the joint error of length measurement and length-at-age variation, or even solely as length-at-age variation if that is more appropriate for a particular case study. Simulation is an important tool to assess the robustness of statistical methods to violations of model assumptions (DiRenzo et al., 2023) and their performance more generally (Morris et al., 2019). Previously, Conn et al. (2020) used simulation to study the effects of unmodelled spatial heterogeneity on CKMR estimation and found that this can induce a negative bias in the abundance estimates.

Our simulation study uses two fictional shark species, that are based on a grey reef shark population (*Carcharhinus amblyrhynchos*) at Palmyra Atoll, in the central Pacific Ocean (Bradley et al., 2017; Papastamatiou et al., 2018). This case study was motivated by genetic samples that were collected from this population in 2013 and 2014. One fictional species is a simplification of the real species (hereafter referred to as the 'simple species'), and was included to test the basic performance of the model. The other fictional species has more realistic life history traits (hereafter referred to as the 'complex species'), and was included to more closely match a real empirical study. We also compare the results for both species. It is paramount to first explore the feasibility of CKMR, for example through simulation, before committing the resources and time required for analysis of the samples. Moreover, the findings will be be relevant to other CKMR studies when age is uncertain.

2 Materials and Methods

We first present our setup of the simulations for the two fictional shark species. Simulated time series are 100 years long, with sampling assumed to occur in the final two years (mimicking the two years of sampling at Palmyra). Following that, we present the POP-based CKMR models for our two species, followed by our estimation method and performance diagnostics. We assume that kinship relationships

are known with certainty; in real life situations, one would need to account for uncertainty in this process (Bravington et al., 2016b). All variables and quantities used in this study are summarised in Table 2.1. Code for the simulation and fitting of models was written in `R` and `C++` (R Core Team, 2021; Eddelbuettel, 2013).

Table 2.1: Summary of notation.

Symbol	Description	Type
	General	
n	Number of sampled individuals (individuals can be sampled more than once)	Observed
p	Detection probability of an individual	Function
f	Probability density/mass function	Function
K	Kinship category	Latent/observed
N_s^A	Adult abundance of sex s	Parameter
	Quantities related to a captured individual	
y	Birth year	Latent
c	Capture/sampling year	Observed
ℓ	Length (when captured)	Observed
σ_ℓ	Standard deviation of the length measurement error	Parameter
a	Age (when captured)	Latent
s	Sex	Observed
$\boldsymbol{z}$	Vector of observed covariates at time of capture/sampling	Observed
	Population dynamics and demography	
ϕ	Survival probability from one year to the next	Parameter
r	Growth rate parameter from one year to the next	Parameter
α	Age of maturity	Parameter
	Subscript	
i, j	Individual i and j	
$\female, \male$	Sex, either female or male	
t	Year	

2.1 Simulation

We used stochastic individual-based ('agent-based') simulation. Two different 'species' were simulated separately, one with simple life history characteristics, and one with a more complex life history. For each simulation, sampling in the last two years was random, and mating occurred at random as well, i.e., mothers and fathers were matched at random, where all non-gestating mothers mated and mature males could father multiple litters in the same mating cycle. Females of the simple species always produced two offspring, whereas the litter size for the complex species ranged from 3 to 6, with equal probability. Females of the simple species reproduced every year as gestation was negligible; females of the complex species gestated for a year and therefore reproduced every other year. Newborns had age zero and sex was assigned at random with an expected sex ratio of 50:50. The survival process was Bernoulli where the annual survival probability ϕ was the same for all ages and sexes, but different between the two species

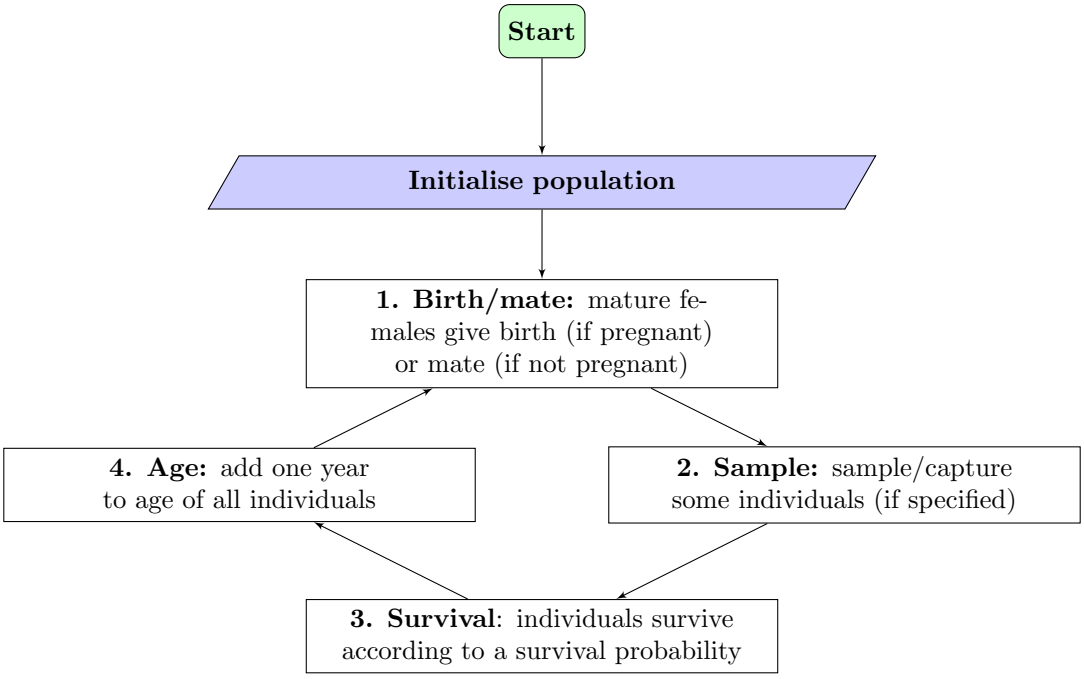

Figure 2.1: Flowchart representing the different stages of the life-cycle for the simulation. A population is initialised at the start of a simulation. Following that, it loops through stage 1–4 every year the simulation runs.

and empirically set at a level that resulted in the yearly population growth rate equaling approximately one, i.e., no growth. Natural mortality was the only source of mortality we considered, and all individuals that reached the maximum age perished at the next survival event, i.e., animals could go through at most $a_{\max} + 1$ yearly cycles. The maximum age for sharks of the simple species was 19 years and 63 years for the complex species, where the latter matches the results from Bradley et al. (2017). For a given species, all individuals of the same sex matured in the same year: males and females in the simple species matured at 10 years old, whereas in the complex species males matured at 17 years and females matured at 19 years of age. The length of an animal was the same for all individuals of a certain age, irrespective of sex and species. After the initialisation of a population in year zero, the simulation looped through four distinct events: a birthing/mating event, a sampling event (only in the final two years of the simulation), a survival event, and an ageing event (Figure 2.1).

For both species, we ran the simulations for one hundred years, to ensure that all animals of the initial populations would have died off. Every simulation started with 8500 individuals. At each sampling event, 375 individuals were non-lethally sampled, where re-captures were possible between sampling events, which resulted in at most 750 unique sampled individuals across the two years of sampling. For every sampled individual, the age, year of capture, and sex were recorded; true length was derived through a von Bertalanffy growth function (VBGF; von Bertalanffy, 1938; Francis, 1988), that was specified as

$$l(a) = l_\infty \times (1 - e^{-k(a-a_0)}), \tag{1}$$

where $\ell_\infty = 163$ cm is the asymptotic length, $a_0 = -8.27$ the theoretical age at length zero, and $k = 0.0554$ denotes the growth coefficient. These values match the estimates of the best model in Table 2 of Bradley et al. (2017). Gaussian noise was added to reflect length measurement error with variance $\sigma_\ell^2 = 2.89^2$, after which length was rounded to the nearest integer. Based on these parameters, we generated 1000 different realisations of a 100-year-long population history for each species, using functions based on those from the `fishSim`-package (Baylis, 2022).

2.2 POP-based estimator

In this study, we developed estimators for both populations based only on POPs. CKMR models use a pseudo-likelihood, which is constructed from the joint distribution of all pairwise comparisons between the samples, i.e., the product of approximately n^2 Bernoulli trials for a POP, where n is the number of samples. It is not a true likelihood as we only consider pairwise comparisons and treat these as independent, whereas they clearly are not: an offspring can only have one parent of each sex. Working with the pseudo-likelihood does not affect the point estimates but could affect other properties of likelihood-based estimation, such as variance estimation, although this effect is likely minor or even negligible provided that a small proportion of the total population is sampled, i.e., $n \ll N$ (Skaug, 2001; Bravington et al., 2016b). Because length is measured with error and age is inferred from length, age is uncertain and hence we cannot assume directionality in the comparison, i.e., who is the parent and who is the offspring. Therefore, for any comparison for individual i and j, we test both directions (parent-offspring and offspring-parent), denoted PO/OP. The pseudo-log-likelihood is given by

$$\begin{aligned}\log \mathcal{L}_P(\boldsymbol{\theta}|\boldsymbol{x}) = &\ell_P(\boldsymbol{\theta}|\boldsymbol{x}) = \\ &\sum_i \sum_j \log \Big\{ \Pr(K_{ij} = \text{PO/OP}|\boldsymbol{z}_i, \boldsymbol{z}_j)^{\omega_{ij}} (1 - \Pr(K_{ij} = \text{PO/OP}|\boldsymbol{z}_i, \boldsymbol{z}_j))^{1-\omega_{ij}} \Big\},\end{aligned} \tag{2}$$

where $\boldsymbol{\theta}$ is the parameter vector, $\boldsymbol{x}$ denotes the observed data, K_{ij} is the kinship between i and j, Pr() is the probability function, ω_{ij} is an indicator that is 1 if the kinship between i and j is observed to be PO/OP and 0 otherwise, and $\boldsymbol{z}$ denotes the information recorded about a captured individual, such as length. Age is required to calculate the probability of observing kinship, and therefore we sum over all potential ages for i, j and multiply by the probability density of that age given the measured length, $f(a|l^*)$:

$$\begin{aligned}\ell_P(\boldsymbol{\theta}|\boldsymbol{x}) = \sum_i \sum_j \log \Big\{ &\sum_{a_i} \sum_{a_j} \Pr(K_{ij} = \text{PO/OP}|\boldsymbol{z}_i, \boldsymbol{z}_j, a_i, a_j)^{\omega_{ij}} \times \\ &(1 - \Pr(K_{ij} = \text{PO/OP}|\boldsymbol{z}_i, \boldsymbol{z}_j, a_i, a_j))^{1-\omega_{ij}} \times f(a_i|l_i^*) f(a_j|l_j^*) \Big\}.\end{aligned} \tag{3}$$

We will now specify the two main elements of Equation (3), namely the probability of observing the PO/OP kinship, and the probability density of age given length.

2.2.1 Probability of kinship

We modelled the female and male abundance separately, and thus for every PO/OP comparison between two individuals, we had to consider four scenarios, which were all combinations of which individual is older and thus the potential parent, and whether this parent was male or female. We will first present the formulae for the simple species, followed by those for the complex species. The probability of any comparison between i and j being PO/OP is the same as the product of the testing for PO and OP separately (or the sum as we are working with the log-likelihood), thus we only present the PO probabilities. For the simple species this became

$$\Pr(K_{ij} = \text{FO}|\boldsymbol{z}_i, \boldsymbol{z}_j, a_i, a_j) = \\ \mathbb{I}(y_i + \alpha_{\female} \leq y_j) \times \left(N^A_{\female, y_j}\right)^{-1} \times \begin{Bmatrix} 1; & \text{if } c_i \geq y_j \\ \phi_i(c_i, y_j); & \text{if } c_i < y_j \end{Bmatrix} \tag{4}$$

for the males, and

$$\Pr(K_{ij} = \text{MO}|\boldsymbol{z}_i, \boldsymbol{z}_j, a_i, a_j) = \\ \mathbb{I}(y_i + \alpha_{\male} \leq y_j) \times \left(N^A_{\male, y_j}\right)^{-1} \times \begin{Bmatrix} 1; & \text{if } c_i \geq y_j \\ \phi_i(c_i, y_j); & \text{if } c_i < y_j \end{Bmatrix} \tag{5}$$

for the females. Here, $\mathbb{I}()$ is an indicator function that returns 1 if its argument is true and 0 otherwise, FO and MO refer to father-offspring and mother-offspring, respectively, y denotes the birth year, α the age of maturity, $N^A_{s,t}$ the total adult abundance of sex s in year t, c the year of capture, and $\phi_i(t_1, t_2)$ the survival function for individual i from t_1 to t_2; as survival was assumed constant, $\phi_i(t_1, t_2)$ was defined as $\phi^{t_2 - t_1}$. Even though females could only have one litter whereas males could father multiple litters, their ERROs were formulated similarly, i.e., the reciprocal of the total mature abundance of their respective sexes. For the complex species, the probability of an MO pair thus became

$$\Pr(K_{ij} = \text{MO}|\boldsymbol{z}_i, \boldsymbol{z}_j, a_i, a_j) = \\ \mathbb{I}(y_i + \alpha_{\female} \leq y_j - 1) \times \left(N^A_{\female, y_j - 1} \times \phi\right)^{-1} \times \begin{Bmatrix} 1; & \text{if } c_i \geq y_j \\ \phi_i(c_i, y_j); & \text{if } c_i < y_j \end{Bmatrix} \tag{6}$$

and the probability of an FO pair became

$$\Pr(K_{ij} = \text{FO}|\boldsymbol{z}_i, \boldsymbol{z}_j, a_i, a_j) = \\ \mathbb{I}(y_i + \alpha_{\male} \leq y_j - 1) \times \left(N^A_{\male, y_j - 1}\right)^{-1} \times \begin{Bmatrix} 1; & \text{if } c_i \geq y_j - 1 \\ \phi_i(c_i, y_j - 1); & \text{if } c_i < y_j - 1 \end{Bmatrix}. \tag{7}$$

The two key differences between the complex species relative to the simple one were that 1) a potential father only needed to have been alive the year before the birth of the offspring, whereas a potential mother needed to have survived until birthing, and 2) the potential parents needed to have matured at least one year before the birth year. To illustrate this, imagine that we are comparing two individuals from the complex species, where the parent is female, and we know the individuals' ages. The offspring was caught in year 50 at age 3, and thus born in year 47. The potential parent was female, and caught in year 45 and would have needed to survive for at least two years in order to be a potential parent; she was 36 years old at the time of capture, and thus born in year 9. The ERRO for this parent in the year of mating, i.e., the year before the birth year of j, is the reciprocal of the number of females alive in that year who also survived one year of gestation, which is ϕ. Therefore, the probability that i is the mother of j would be:

$$\begin{aligned} \Pr(K_{ij} = \text{MO}|\boldsymbol{z}_i, \boldsymbol{z}_j, a_i, a_j) =& \mathbb{I}(9 + 19 \leq 47 - 1) \times \left(N^A_{\female, y_j - 1} \times \phi\right)^{-1} \times \phi^{47-45} \\ =& 1 \times \left(N^A_{\female, y_j - 1} \times \phi\right)^{-1} \times \phi^2 \\ =& \phi \left(N^A_{\female, y_j - 1}\right)^{-1} \end{aligned} \tag{8}$$

Every comparison, given a_i and a_j, contains a signal about the adult population in a specific year. We assumed a constant population size, and thus $N^A_{s,t} = N^A_s$. We also developed and tested a model that included sex-specific growth parameters. This model was internally inconsistent and therefore not included in the main body of this manuscript for any formal inference. However, we did include the derivations and some results in Appendix C.

2.2.2 Probability density of age given length

We had an assumed true length-at-age curve $l(a)$ (Equation (1)) and we knew that there was measurement error on lengths. We denote the *measured* length as l^*. With this information we derived the probability density $f(a|l^*)$ using Bayes' rule as follows:

$$f(a|l^*) = f(l^*|a) f(a) f(l^*)^{-1}. \tag{9}$$

Measured length given age $l^*|a$ was assumed to follow a discretised Normal distribution, as lengths were rounded to the nearest centimetre. We followed Roy (2003) in defining this distribution as

$$f(l^*|a) = \Phi\left(\frac{l^* - \mu - 0.5}{\sigma_l}\right) - \Phi\left(\frac{l^* - \mu + 0.5}{\sigma_l}\right), \tag{10}$$

where Φ denotes the standard normal cumulative distribution function, the expectation μ is given by Equation (1), and standard error σ_l captures the measurement error. As the sampling probability in the simulation was unrelated to age, the age distribution of sampled individuals was the same as the age distribution in the whole population, and we did not need to distinguish between the two. We assumed that the population had a stable age distribution with no growth, which meant that the distribution of ages, had we not imposed a maximum age, would have been geometric with shape parameter being equal to the mortality rate, which is $1 - \phi$. Acknowledging that there was a maximum age, $a_{\max}$, we needed to condition on the age being at most this age, and thus

$$f(a) = \begin{cases} \frac{(\phi)^a(1-\phi)}{1-(\phi)^{a_{\max}+1}}; & \text{if } 0 \leq a \leq a_{\max} \\ 0; & \text{otherwise} \end{cases}, \tag{11}$$

where the numerator and denominator were the geometric probability mass and cumulative distribution functions, respectively. Note here that we used the definition of a geometrically distributed variable being the number of failures (survival) until a success (death) occurs. Finally, the probability density function on measured length became

$$f(l^*) = \sum_{a=0}^{a_{\max}} f(l^*|a) f(a). \tag{12}$$

2.3 Fitting

The parameters in the CKMR model were estimated by maximising the pseudo-likelihood, which can involve prohibitively long computation time. To resolve this, we restricted the number of pairwise comparisons. Many pairwise comparisons resulted in identical probabilistic statements, and thus in practice only needed to be derived once. As we considered abundance for both sexes separately, we estimated two parameters: $N_{\female}^A$ and $N_{\male}^A$. All other parameters, such as ϕ, were assumed known and fixed. To each of the 2000 population realisations (1000 for each species) we fitted the appropriate POP-model with varying length measurement errors and growth curves, which was achieved by altering some of the fixed parameters. Specifically, we considered five different assumed length measurement errors: the correct one, a 33% and 67% underestimate, and a 33% and 67% overestimate. We also considered five growth curves: the correct one, two that were shifted upwards by 5% and 10%, and two that were shifted downwards by 5% and 10%. This resulted in a total of 25 combinations, or scenarios. We labelled these

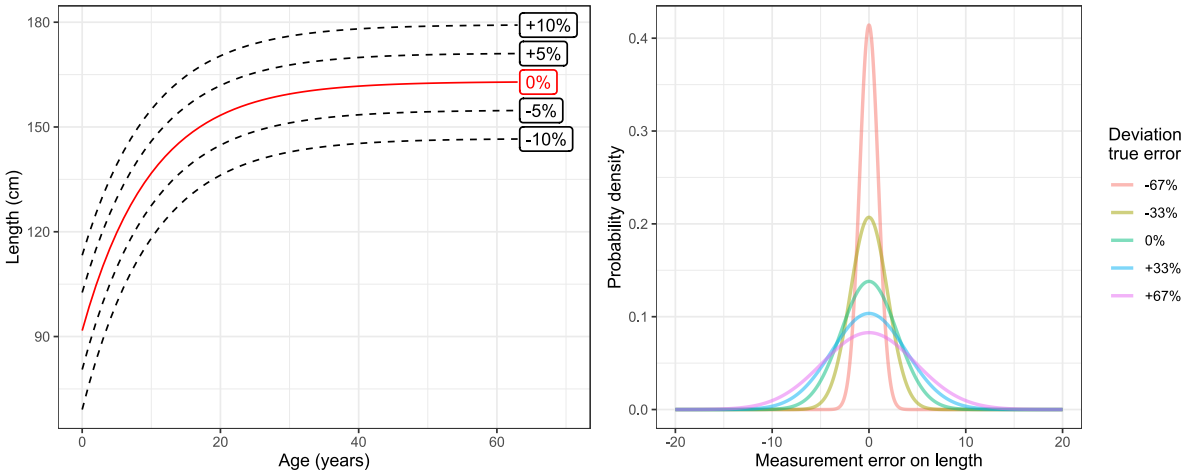

Figure 2.2: The left panel shows the five growth curves that were used in the scenarios tested in this study. The true growth curve is indicated in red; the black dotted-lines show the incorrect ones, which were constructed by shifting the growth curve up and down in steps of 5 percent. The right panel shows five measurement errors used in this study. The true simulated error was 2.89 cm, and the other measurements errors were chosen by deviating from this error in both directions.

scenarios using the format 'ME±XX:GC±YY', where ME refers to the measurement error, XX denotes the percentage over- or underestimate, GC stands for growth curve, and YY denotes the percentage of up- or downwards shifting; for example, the scenario with a 33% overestimated length measurement error and a 5% downshifted growth curve had the label ME+33:GC-5. These errors and growth curves are visualised in more detail in Figure 2.2. Considering 25 scenarios for every simulation resulted in the fitting of 50,000 models in total. To keep computation time to a minimum, we implemented most of the fitting process in `C++`, which was linked with `R` through the `Rcpp`-library (R Core Team, 2021; Eddelbuettel, 2013).

2.4 Variance and performance

To evaluate the performance of the estimator, we present the following metrics: i) mean error and mean relative error to evaluate a potential bias; ii) median error and median relative error to evaluate the median bias, which uses the median instead of the mean, as the median is often more appropriate when distributions are skewed. In addition, the mean absolute error (MAE) and root mean squared error (RMSE) are presented in supplemental tables. Furthermore, we evaluated the performance of the uncertainty estimator by comparing the average estimated (model-based) coefficient of variation (CV) with the true (empirical) CV. We evaluated the CV as opposed to the variance or standard deviation as we believed a relative measure of variation was more valuable in this study. However, we appreciated that a bias in the estimator would have induced a bias of opposite direction in estimated CV; thus, we also included the true and estimated standard deviation in supplemental tables. The average estimated variance was estimated from the Hessian matrix produced by the maximum likelihood estimation, and

averaged over these 1000 estimated variances. We can treat the pseudo-likelihood as a true likelihood as long as sampling was sparse (see Section 2.2). We sampled at most 750 unique individuals out of a population with an expected size of 8,500 individuals, so it is unclear if these meets that criterion. The occurrence of recaptures, i.e., that some individuals were sampled at more than one sampling event, suggests that it has not, as this clearly indicates that the samples are not independent. Consequently the estimated variance will be negatively biased as the pairwise comparisons are not approximately independent. To explore the extent of this bias, we compared the average estimated variance to the true variance, which was the empirical variance in population estimates across the 1000 simulations for each species.

3 Results

The mean number of POPs for all sampling realisations was 48.6 (range: 25–76) for the simple species and 55.6 (range: 31–90) for the complex species. Mean simulated adult abundances in the final year of the simulation were 794 and 793 (range: 630–992 and 600–1019; ♀ and ♂) for the simple species and 514 and 650 (range: 400–683 and 516–824; ♀ and ♂) for the complex species. Small numbers of recaptures occurred in every simulation, ranging from 4 to 25 recaptured individuals for the simple species and 5 to 30 individuals for the complex species. The simulated mean annual growth rate was 0.999 for both sexes of the simple species, and 1.001 and 0.998 for the males and females of the complex species, respectively; the mean annual growth for any simulation was always within 0.3 percent point from the mean across all simulations. The fitting algorithm did not always converge when the measurement error and/or growth curve was (very) negatively biased. Whenever this happened, it happened for the most of the simulations in that scenario. Therefore, we excluded the scenarios where this happened from the analysis, which led to the exclusion of scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10.

For the simple species, median errors for $\widehat{N}^A$ when using correct measurement error and growth curve specification (ME+0:GC+0) were 20.83 and 22.52 (relative: 2.57% and 2.79%; ♀ and ♂) individuals (Figure 3.1; Tables A.1 and A.2). For the complex species, median errors for $\widehat{N}^A$ when using correct measurement error and growth curve specification were 6.48 and 10.28 (relative: 1.35% and 1.59%; ♀ and ♂) individuals (Figure 3.2; Tables A.3 and A.4). For the simple species, median relative errors in abundance estimates were positive but close to zero for all deviations from the true length measurement error provided that the growth curve was correctly specified, although they were slightly larger for the females (Figure 3.1, also Tables A.1–4). For any given measurement error, we observed a trend from a positive median error to a negative median error as we shifted the growth curve upwards (Figures 3.1 and 3.2). When growth curves were shifted down 5%, this resulted in median relative errors of around

30% for the simple species, and between 30 and 60% for the complex species. Shifting growth curves up by 5% resulted in median relative errors of -30% for the simple species, and between -30 and -40% for the complex species.

True CVs for the same estimates ranged from 21.56 to 28.50% for the simple species, and 20.17 to 33.97% for the complex species (Table B.4). Given a measurement error, the true CV increased as the growth curve was shifted away from the truth, in either direction. Given a growth curve, changing the length measurement error had a negligible effect on the CV. Overall, little difference in true CV was noted between the two species. Estimated CVs for adult abundance estimates ranged from 20.57 to 20.67% for the simple species, and 18.67 to 19.64% for the complex species (Table B.2). Overall spread was higher for the complex species compared to the simple species. Moreover, the estimated CV was always lower than the true CV, for all parameters. Averaged over the scenarios, estimated CVs exhibited a negative bias in $\widehat{N}^A_{♂}$ of 11.2% and in $\widehat{N}^A_{♀}$ of 14.3% for the simple species, and a negative bias of $\widehat{N}^A_{♂}$ of 16.3% and in $\widehat{N}^A_{♀}$ of 21.1%.

4 Discussion

In this study, we explored the effects of incorrect age inference from length measurements on CKMR estimates through misspecifying the length measurement error and the growth curve in various ways. Overall, an incorrect measurement error mostly impacted the convergence likelihood of the fitting algorithm, whenever the measurement error was assumed to be smaller in the fitting than was true for the simulation. Whenever the measurement error was high enough to allow for convergence, it made little difference whether it was the true error or if a much higher error was assumed. This would suggest that, if researches are ever unsure about whether their assumed length measurement error is correct, it is safer to overestimate it. A misspecified growth curve, on the other hand, had drastic effects on the estimation of all parameters, both through the point and the uncertainty estimates.

The model performed well under correct specification (scenario ME+0:GC+0), although the positive median relative error in adult abundance estimates suggest a positive median bias. This error was more extreme for the simple than for the complex species. A bias in the estimates is not uncommon for maximum likelihood methods when sample size is small, which could be true in our study as the number of sampled POPs never surpassed 76 for the simple species and 90 for the complex species. However, it could also be that this shows a slight bias in the method, especially as a previous CKMR simulation study by Conn et al. (2020) found positive biases in the abundance estimates too.

In this study, we assumed that all individual sharks followed the specified growth curve perfectly, and any variation in lengths for a given age came from measurement error. This is a simplification of reality, and future research could focus on ways to accommodate natural variation in length at a given age,

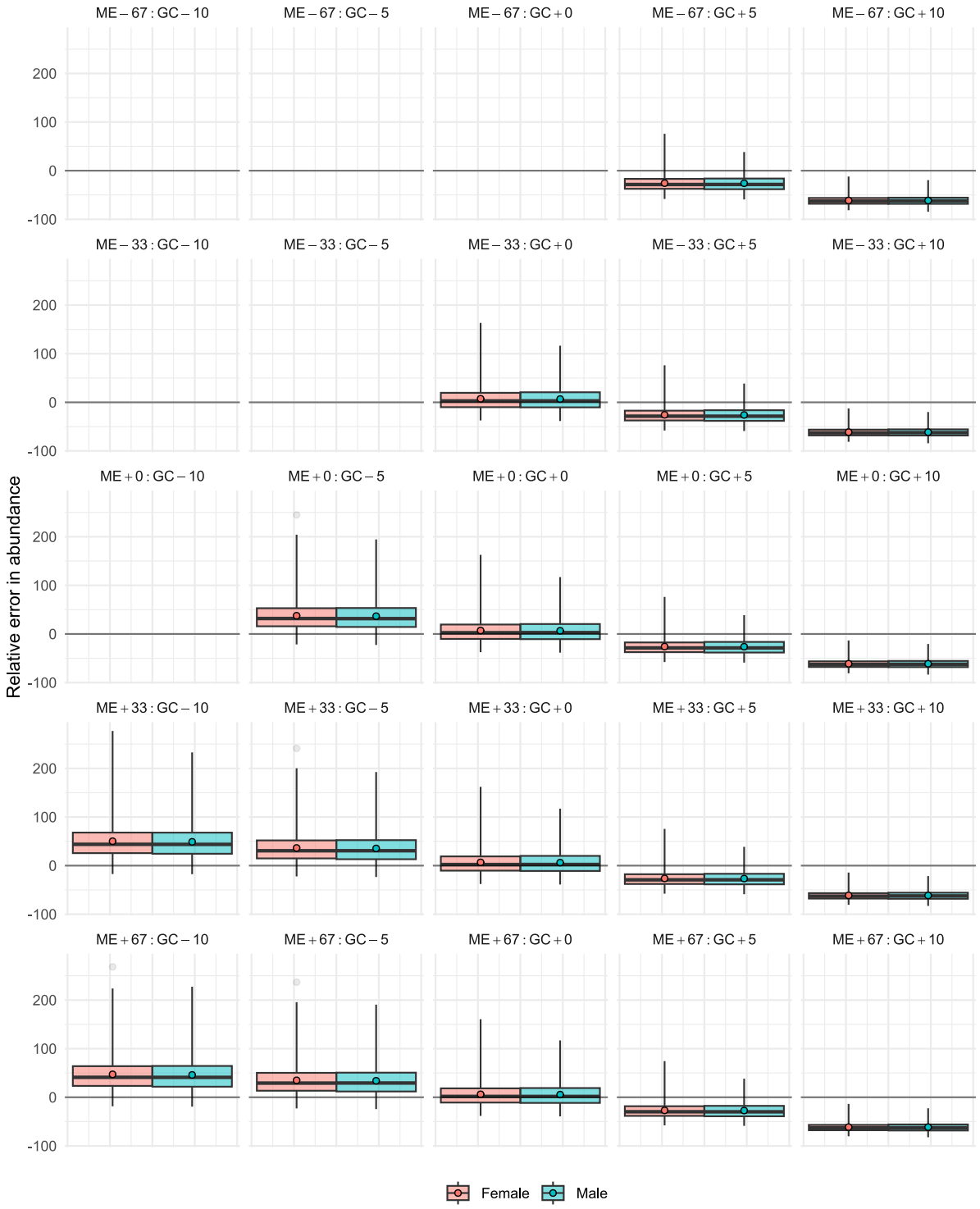

Figure 3.1: Box plots for estimated sex-specific adult abundance for the simple species. We only present the results for scenarios in which the optimiser consistently converged; this meant that some scenarios were left blank. Box plots show the interquartile range (IQR) and the median; the mean is indicated by the darker filled circle; the vertical lines cover five times the IQR; and all values outside of that are indicated as outliers. The scenarios were labelled using the format 'ME±XX:GC±YY', where ME refers to the measurement error, XX denotes the percentage over- or underestimate, GC stands for growth curve, and YY denotes the percentage of up- or downwards shifting; for example, the scenario with a 33% overestimated length measurement error and a 5% downshifted growth curve had label ME+33:GC-5.

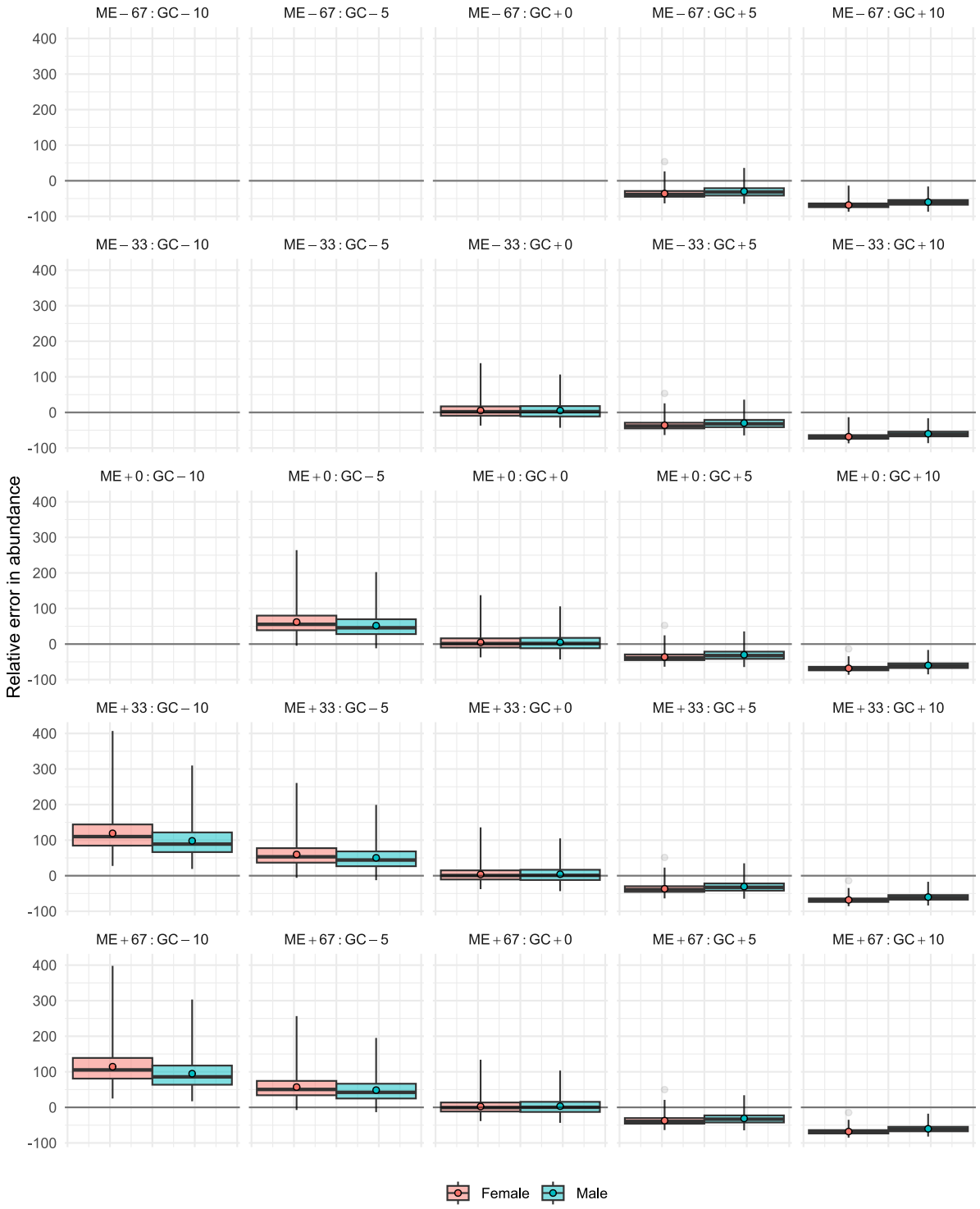

Figure 3.2: Box plots for estimated sex-specific adult abundance for the complex species. We only present the results for scenarios in which the optimiser consistently converged; this meant that some scenarios were left blank. Box plots show the interquartile range (IQR) and the median; the mean is indicated by the darker filled circle; the vertical lines cover five times the IQR; and all values outside of that are indicated as outliers. The scenarios were labelled using the format 'ME±XX:GC±YY', where ME refers to the measurement error, XX denotes the percentage over- or underestimate, GC stands for growth curve, and YY denotes the percentage of up- or downwards shifting; for example, the scenario with a 33% overestimated length measurement error and a 5% downshifted growth curve had label ME+33:GC-5.

which could be a function of age in itself. As an incorrect length measurement error seemed to have little effect on point estimates, we believe that, when in doubt, it is preferable to assume a higher length measurement error as this improves how likely it is that the fitting algorithm converges.

The effects of deviating from the true growth curve on the abundance estimates were substantial. When growth curves were shifted by only 5%—an amount that we considered small—we often observed median relative errors of over 30%. This strongly highlights the sensitivity of the method to correct age estimation. The estimated CVs were consistent for any given parameter throughout the various scenarios. As the true CVs increased when the growth curve was misspecified, irrespective of direction, this meant that shifting the growth curve away from the truth increased the underestimation of the CV.

The true CVs we observed in this study were generally acceptable, especially when the correct scenario was selected. However, occasionally true CVs were over 25%, which we believe is mostly due to the relatively small number of POPs detected in our sample, as these positive detections contain most of the information for parameter estimation. The complex species had slightly higher CVs in the scenarios that were closer to the truth. We believe that this could be due to the randomness in reproduction of the complex species, which resulted in more variation in the population dynamics, which likely inflated the CVs.

Even though CKMR models normally assume that recaptures do not occur, they did occur in our simulations. This violation of the independence assumption resulted in underestimated average CVs and standard deviations on the parameter estimates, for all scenarios. Averaged over the scenarios, estimated CVs exhibited a negative bias in $\widehat{N}^A_{\male}$ of 11.2% and in $\widehat{N}^A_{\female}$ of 14.3% for the simple species, and a negative bias in $\widehat{N}^A_{\male}$ of 16.3% and in $\widehat{N}^A_{\female}$ of 21.1%. When the correct scenario (ME+0:GC+0) was used, the degree of underestimation was below 5% for the simple species, which is a difference that could be acceptable; for the complex species, however, the CV for $\widehat{N}^A_{\male}$ was underestimated by 13.7%, which is substantially higher. We believe that this difference in underestimation of spread is mainly due to the longer-lived nature of the complex species. The violation of independence could therefore be more severe for that species, resulting in a increased difference between model and empirical CVs. Nonetheless, the presence of recaptures does create the potential for extending CKMR by incorporating some form of capture-recapture into the method (Otis et al., 1978). Additionally, it could also allow us to fit the growth curve jointly with the CKMR model, instead of assuming it to be known by extrapolating from other studies. This could create a situation where one collects new samples every year, thereby continuously improving the estimates not only of the abundance and trend, but also of the growth curve: in a Bayesian framework, one could use the initial growth curve as prior information, and then update the posterior every year as more information is collected.

In our model, we did not allow for any growth or decline in the population size over time. Our simulated populations exhibited no systematic growth, but the stochastic nature of the process did lead

to some random growth/decline. One could consider estimating a growth rate, or assume a small range of growth rates (see Hillary et al. (2018) for an example with white sharks (*Carcharodon carcharias*)). The main challenge would be to understand how including a growth rate parameter affects the assumed age distribution. We can imagine three general population growth scenarios. If a population is stable but growing or in decline, the assumed age distribution will be geometric and depend on a combination of survival and growth rate (Caswell, 2000, Section 4.5.2.1). The second scenario is when a population exhibits changing growth or decline, in which case there is no stable age structure. We believe that this scenario is intractable, and it would make a good subject for a robustness study to see how much it affects estimation. The third scenario would be where there is no expected population growth or decline but there is demographic stochasticity, which in practice could result in deviations from the stable age structure. For this scenario, an option could be to use the method described by Hillary et al. (2018), where the measured lengths were binned and a multinomial distribution was fit to these binned data to estimate the distribution of sampled ages. Still, this could be a topic for future research to see what other methods exist to find the distribution of (sampled) ages.

CKMR involves many pairwise comparisons, which often involves many identical probabilistic statements. To limit computation time, we only unique probabilistic statements once. If further computational improvements are required, it is possible to reduce the number of pairwise comparisons that are evaluated by excluding a subset of comparisons from the analysis. For example, the length-age relationships are often much clearer for younger animals, and therefore one could choose to only consider animals up to a certain size as potential offspring (Trenkel et al., 2022).

In our simulation and model, we assumed some life history traits to be fixed and known, but this is not always required for CKMR. We estimated only abundance in our model and assumed quantities such as survival to be known and fixed. In order to relax the assumption of a fixed and known survival parameter ϕ, one could estimate it by including half-sibling pairs alongside parent-offspring pairs (Bravington et al., 2016b). Parent-offspring pairs can be used to model fecundity as long as the parameter appears explicitly in the model, which could be the case when fecundity varies with the size or age of animals (Bravington et al., 2016b, Section 3.1.4) We are unaware of any attempts to estimate time-varying fecundity or survival, and we believe this to be a potential direction for future research. Moreover, we assumed maturity to be knife-edge as it slightly reduces complexity of the model. However, if maturity occurs more gradually, then this can be accommodated by adding a fecundity curve to the model (e.g., a logistic curve; Conn et al., 2020). We also imposed a fixed and known maximum age in the simulation, mostly to reduce computation time. In reality, animals do not always have a maximum age; in such cases, one could set the maximum age equal to an age that the animal has practically zero probability of reaching.

When a promising method like CKMR is first presented, one can see the appeal to start studying populations as quickly as possible. Benchmark comparisons could be useful (e.g., Ruzzante et al., 2019)

to compare a new method to some 'truth'. However, these comparisons can be ambiguous when it is unclear how accurate the benchmark truly is. Simulation studies, such as this one (and see Conn et al. (2020) for the effects of unmodelled spatial heterogeneity on CKMR), are a key part of understanding when the CKMR method works well and when it does not. We believe the CKMR method has great potential and, in some cases, is an improvement over other methods, but our study confirms that care that needs to be taken when ageing is biased. In such cases epigenetic ageing could be preferable, even though epigenetic ageing can still involve substantial uncertainty (e.g., Larison et al., 2021; Prado et al., 2021) and relies strongly on the quality of the training data (Mayne et al., 2023).

Author contributions

Felix T. Petersma: conceptualisation (lead); simulation (lead); methodology (lead); formal analysis (lead); visualisation (lead); writing - original draft(lead); writing - review and editing (equal). Len Thomas: conceptualisation (supporting); simulation (supporting); formal analysis (supporting); visualisation (supporting); writing - original draft(supporting); writing - review and editing (equal). Danielle V. Harris: conceptualisation (supporting); simulation (supporting); formal analysis (); visualisation (); writing - original draft(supporting); writing - review and editing (equal). Darcy Bradley: conceptualisation (supporting); simulation (supporting); formal analysis (supporting); writing - review and editing (equal). Yannis Papastamatiou: conceptualisation (supporting); simulation (supporting); formal analysis (supporting); writing - review and editing (equal).

Acknowledgements

We would like to thank Mark Bravington for his time and enthusiasm whenever we discussed the methodology and worked on deriving the correct probabilities over email and in-person.

Data availability statement

The `R` scripts used for the simulation and analysis in this study are publicly available at `https://github.com/fpetersma/CKMR_grey_reef_sharks`.

Competing interests statement

The authors declare no conflicts of interest.

ORCID

Felix T. Petersma: https://orcid.org/0000-0001-8915-0627

Len Thomas: https://orcid.org/0000-0002-7436-067X

Danielle V. Harris: https://orcid.org/0000-0003-1447-1420

Darcy Bradley: https://orcid.org/0000-0003-2581-8768

Yannis Papastamatiou: https://orcid.org/0000-0002-6091-6841

References

Baylis, S. M. (2022). fishSim: Simulate Populations With Multiple Stocks, Movement, Mating, Mortality, and Parentage.

Borchers, D. L. and Efford, M. G. (2008). Spatially explicit maximum likelihood methods for capture-recapture studies. *Biometrics*, 64(2):377–385.

Bradford, R. W., Thomson, R., Bravington, M. V., Foote, D., Gunasekera, R., Bruce, B. D., Harasti, D., Otway, N., and Feutry, P. (2018). A close-kin mark-recapture estimate of the population size and trend of east coast grey nurse. Technical Report July, CSIRO, Hobart.

Bradley, D., Conklin, E., Papastamatiou, Y. P., McCauley, D. J., Pollock, K., Kendall, B. E., Gaines, S. D., and Caselle, J. E. (2017). Growth and life history variability of the grey reef shark (*Carcharhinus amblyrhynchos*) across its range. *PLoS ONE*, 12(2):1–20.

Bravington, M. V., Feutry, P., Pillans, R. D., Johnson, G., Saunders, T., Gunasekera, R., Bax, N. J., and Kyne, P. M. (2019). Close-Kin Mark-Recapture population size estimate of *Glyphis garricki* in the Northern Territory. Technical Report November, CSIRO, Hobart.

Bravington, M. V., Grewe, P. M., and Davies, C. R. (2016a). Absolute abundance of southern bluefin tuna estimated by close-kin mark-recapture. *Nature Communications*, 7(1):1–8.

Bravington, M. V., Skaug, H. J., and Anderson, E. C. (2016b). Close-kin mark-recapture. *Statistical Science*, 31(2):259–274.

Campana, S. E. (2001). Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. *Journal of Fish Biology*, 59(2):197–242.

Caswell, H. (2000). *Matrix Population Models*. Sinauer Sunderland, MA.

Conn, P. B., Bravington, M. V., Baylis, S., and Ver Hoef, J. M. (2020). Robustness of close-kin mark–recapture estimators to dispersal limitation and spatially varying sampling probabilities. *Ecology and Evolution*, 10(12):5558–5569.

De Paoli-Iseppi, R., Deagle, B. E., McMahon, C. R., Hindell, M. A., Dickinson, J. L., and Jarman, S. N. (2017). Measuring animal age with DNA methylation: From humans to wild animals. *Frontiers in Genetics*, 8:2010–2017.

DiRenzo, G. V., Hanks, E., and Miller, D. A. (2023). A practical guide to understanding and validating complex models using data simulations. *Methods in Ecology and Evolution*, 14(1):203–217.

Eddelbuettel, D. (2013). *Seamless R and C++ Integration with Rcpp*. Springer New York, New York, NY.

Francis, R. (1988). Maximum likelihood estimation of growth and growth variability from tagging data. *New Zealand Journal of Marine and Freshwater Research*, 22(1):43–51.

Hillary, R. M., Bravington, M. V., Patterson, T. A., Grewe, P., Bradford, R., Feutry, P., Gunasekera, R., Peddemors, V., Werry, J., Francis, M. P., Duffy, C. A. J., and Bruce, B. D. (2018). Genetic relatedness reveals total population size of white sharks in eastern Australia and New Zealand. *Scientific Reports*, 8(1):2661.

Larison, B., Pinho, G. M., Haghani, A., Zoller, J. A., Li, C. Z., Finno, C. J., Farrell, C., Kaelin, C. B., Barsh, G. S., Wooding, B., Robeck, T. R., Maddox, D., Pellegrini, M., and Horvath, S. (2021). Epigenetic models developed for plains zebras predict age in domestic horses and endangered equids. *Communications Biology*, 4(1):1412.

Mayne, B., Berry, O., and Jarman, S. (2023). Calibrating epigenetic clocks with training data error. *Evolutionary Applications*, 16(8):1496–1502.

Morris, T. P., White, I. R., and Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, 38(11):2074–2102.

Otis, D. L., Burnham, K. P., White, G. C., and Anderson, D. R. (1978). Statistical inference from capture data on closed animal populations. *Wildlife Monographs*, (62):3–135.

Papastamatiou, Y. P., Bodey, T. W., Friedlander, A. M., Lowe, C. G., Bradley, D., Weng, K., Priestley, V., and Caselle, J. E. (2018). Spatial separation without territoriality in shark communities. *Oikos*, 127(6):767–779.

Patterson, T. A., Hillary, R., Feutry, P., Gunasekera, R., Marthick, J., and Pillans, R. D. (2022). Rapid estimation of cryptic adult abundance and breeding dynamics in a critically endangered elasmobranch from close-kin mark recapture. *bioRxiv Ecology*, pages 1–34.

Polanowski, A. M., Robbins, J., Chandler, D., and Jarman, S. N. (2014). Epigenetic estimation of age in humpback whales. *Molecular Ecology Resources*, 14(5):976–987.

Prado, N. A., Brown, J. L., Zoller, J. A., Haghani, A., Yao, M., Bagryanova, L. R., Campana, M. G., E. Maldonado, J., Raj, K., Schmitt, D., Robeck, T. R., and Horvath, S. (2021). Epigenetic clock and methylation studies in elephants. *Aging Cell*, 20(7):1–11.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Vienna, Austria.

Roy, D. (2003). The discrete normal distribution. *Communications in Statistics - Theory and Methods*, 32(10):1871–1883.

Ruzzante, D. E., McCracken, G. R., Førland, B., MacMillan, J., Notte, D., Buhariwalla, C., Mills Flemming, J., and Skaug, H. (2019). Validation of close-kin mark–recapture (CKMR) methods for estimating population abundance. *Methods in Ecology and Evolution*, 10(9):1445–1453.

Skaug, H. J. (2001). Allele-sharing methods for estimation of population size. *Biometrics*, 57(3):750–756.

Skaug, H. J. (2017). The parent–offspring probability when sampling age-structured populations. *Theoretical Population Biology*, 118:20–26.

Trenkel, V. M., Charrier, G., Lorance, P., and Bravington, M. V. (2022). Close-kin mark–recapture abundance estimation: practical insights and lessons learned. *ICES Journal of Marine Science*, 79(2):413–422.

von Bertalanffy, L. (1938). A quantitative theory of organic growth. *Human Biology*, 10(2):181–213.

Appendix A Performance metrics

Table A.1: Performance metrics for the estimation of parameter $N_{♀}^{A}$, extracted from 1000 simulations of the simple shark population. The columns are, from left to right: scenario, median relative error, mean relative error, median error, mean error, mean absolute error, and root mean square error. Scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10 were not included as (most of) the simulations did not fit successfully. Scenario 3-3 uses the correct measurement error (2.89) and growth curve specification.

Scenario	Mdn. Rel. Err	Mean Rel. Err.	Mdn. Err.	Mean Err.	MAE	RMSE
ME-67:GC+5	-28.48	-25.66	-223.83	-203.94	220.61	245.14
ME-67:GC+10	-62.69	-61.36	-495.14	-487.72	487.72	495.50
ME-33:GC+0	2.88	7.32	22.72	58.23	142.91	201.11
ME-33:GC+5	-28.54	-25.71	-224.94	-204.40	220.93	245.38
ME-33:GC+10	-62.78	-61.34	-495.26	-487.58	487.58	495.28
ME+0:GC-5	31.80	37.56	252.60	298.59	304.39	390.33
ME+0:GC+0	2.57	7.14	20.83	56.78	142.47	200.48
ME+0:GC+5	-28.72	-25.85	-227.03	-205.50	221.78	246.05
ME+0:GC+10	-62.70	-61.23	-494.06	-486.74	486.74	494.33
ME+33:GC-10	43.85	50.21	343.32	399.15	400.98	488.70
ME+33:GC-5	30.59	36.34	243.51	288.88	295.53	381.69
ME+33:GC+0	2.23	6.76	18.13	53.74	141.65	199.10
ME+33:GC+5	-29.21	-26.29	-230.69	-209.01	224.53	248.47
ME+33:GC+10	-62.81	-61.25	-494.00	-486.85	486.85	494.26
ME+67:GC-10	41.04	47.21	320.66	375.24	377.73	466.13
ME+67:GC-5	29.16	34.90	232.41	277.39	285.22	371.65
ME+67:GC+0	1.79	6.12	13.53	48.64	140.45	196.86
ME+67:GC+5	-29.80	-26.97	-235.30	-214.35	228.79	252.27
ME+67:GC+10	-62.85	-61.37	-495.30	-487.84	487.84	495.04

Table A.2: Performance metrics for the estimation of parameter $N^A_{\male}$, extracted from 1000 simulations for the simple shark population. The columns are, from left to right: scenario, median relative error, mean relative error, median error, mean error, mean absolute error, and root mean square error. Scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10 were not included as (most of) the simulations did not fit successfully. Scenario 3-3 used the correct measurement error (2.89) and growth curve specification.

Scenario	Mdn. Rel. Err	Mean Rel. Err.	Mdn. Err.	Mean Err.	MAE	RMSE
ME-67:GC+5	-28.40	-25.93	-223.56	-206.14	217.13	243.78
ME-67:GC+10	-62.28	-61.38	-488.98	-487.42	487.42	494.97
ME-33:GC+0	2.96	6.74	24.43	52.92	144.84	189.98
ME-33:GC+5	-28.53	-25.99	-224.78	-206.64	217.45	244.05
ME-33:GC+10	-62.41	-61.37	-489.77	-487.34	487.34	494.79
ME+0:GC-5	31.59	36.63	247.66	289.87	297.53	374.35
ME+0:GC+0	2.79	6.55	22.52	51.44	144.49	189.38
ME+0:GC+5	-28.72	-26.14	-226.29	-207.83	218.33	244.83
ME+0:GC+10	-62.30	-61.28	-489.39	-486.58	486.58	493.89
ME+33:GC-10	43.72	49.00	341.71	387.75	390.42	470.34
ME+33:GC-5	30.62	35.41	238.47	280.12	289.00	365.82
ME+33:GC+0	2.34	6.16	19.50	48.30	143.76	188.06
ME+33:GC+5	-29.16	-26.59	-228.60	-211.41	221.07	247.38
ME+33:GC+10	-62.32	-61.30	-489.56	-486.78	486.78	493.90
ME+67:GC-10	40.98	46.03	320.23	364.13	367.87	448.20
ME+67:GC-5	29.36	33.96	229.31	268.62	279.08	355.90
ME+67:GC+0	1.72	5.51	13.73	43.13	142.59	185.95
ME+67:GC+5	-29.89	-27.27	-233.68	-216.81	225.38	251.36
ME+67:GC+10	-62.58	-61.44	-490.18	-487.86	487.86	494.77

Table A.3: Performance metrics for the estimation of parameter $N^A_{\female}$, extracted from 1000 simulations of the complex shark population. The columns are, from left to right: scenario, median relative error, mean relative error, median error, mean error, mean absolute error, and root mean square error. Scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10 were not included as (most of) the simulations did not fit successfully. Scenario 3-3 used the correct measurement error (2.89) and growth curve specification.

Scenario	Mdn. Rel. Err	Mean Rel. Err.	Mdn. Err.	Mean Err.	MAE	RMSE
ME-67:GC+5	-38.21	-35.96	-194.00	-185.43	187.72	199.33
ME-67:GC+10	-69.41	-68.34	-355.48	-352.09	352.09	356.40
ME-33:GC+0	1.89	5.30	9.12	26.84	82.06	112.47
ME-33:GC+5	-38.37	-36.07	-194.96	-185.99	188.23	199.71
ME-33:GC+10	-69.34	-68.26	-354.98	-351.70	351.70	355.94
ME+0:GC-5	55.54	61.97	282.61	318.31	318.38	365.22
ME+0:GC+0	1.35	4.64	6.48	23.44	81.55	111.32
ME+0:GC+5	-38.63	-36.31	-195.61	-187.25	189.38	200.68
ME+0:GC+10	-69.19	-68.16	-353.91	-351.15	351.15	355.29
ME+33:GC-10	109.97	119.20	556.22	612.37	612.37	666.21
ME+33:GC-5	53.20	59.79	270.34	307.08	307.24	354.45
ME+33:GC+0	0.57	3.70	2.63	18.60	80.97	109.82
ME+33:GC+5	-39.05	-36.72	-197.27	-189.35	191.32	202.37
ME+33:GC+10	-69.23	-68.08	-353.09	-350.72	350.72	354.71
ME+67:GC-10	105.31	114.40	534.45	587.65	587.65	641.11
ME+67:GC-5	50.39	57.13	258.11	293.33	293.59	341.34
ME+67:GC+0	-0.63	2.49	-3.03	12.36	80.40	108.17
ME+67:GC+5	-39.51	-37.25	-200.26	-192.09	193.91	204.65
ME+67:GC+10	-69.15	-68.02	-352.44	-350.42	350.42	354.27

Table A.4: Performance metrics for the estimation of parameter $N_{\male}^{A}$, extracted from 1000 simulations for the complex shark population. The columns are, from left to right: scenario, median relative error, mean relative error, median error, mean error, mean absolute error, and root mean square error. Scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10 were not included as (most of) the simulations did not fit successfully. Scenario 3-3 used the correct measurement error (2.89) and growth curve specification.

Scenario	Mdn. Rel. Err	Mean Rel. Err.	Mdn. Err.	Mean Err.	MAE	RMSE
ME-67:GC+5	-31.74	-29.79	-203.73	-194.20	201.00	220.13
ME-67:GC+10	-61.34	-60.03	-393.89	-390.66	390.66	397.78
ME-33:GC+0	2.13	5.35	12.74	34.22	113.74	152.24
ME-33:GC+5	-31.95	-29.89	-204.63	-194.80	201.48	220.51
ME-33:GC+10	-61.30	-59.99	-393.09	-390.44	390.44	397.48
ME+0:GC-5	45.75	51.88	297.86	336.61	337.30	401.99
ME+0:GC+0	1.59	4.91	10.28	31.35	113.06	151.03
ME+0:GC+5	-32.33	-30.13	-206.50	-196.36	202.81	221.63
ME+0:GC+10	-61.23	-59.98	-392.55	-390.36	390.36	397.27
ME+33:GC-10	88.97	98.09	579.49	636.75	636.75	703.85
ME+33:GC-5	44.11	50.38	288.00	326.83	327.73	392.72
ME+33:GC+0	1.05	4.24	6.59	27.01	112.18	149.32
ME+33:GC+5	-32.61	-30.56	-208.75	-199.14	205.21	223.70
ME+33:GC+10	-61.30	-60.04	-392.71	-390.76	390.76	397.51
ME+67:GC-10	85.75	94.83	559.04	615.54	615.54	682.57
ME+67:GC-5	42.24	48.54	276.01	314.85	316.10	381.44
ME+67:GC+0	0.14	3.35	0.95	21.20	111.13	147.24
ME+67:GC+5	-33.16	-31.13	-211.30	-202.85	208.45	226.55
ME+67:GC+10	-61.38	-60.19	-394.36	-391.68	391.68	398.23

Appendix B Variance estimates

Table B.1: Likelihood-based estimates for the standard deviation of the estimated parameters, averaged over 1000 simulations. The first number in a scenario label refers to the assumed measurement error on length; the second number refers to the used growth function. Scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10 were not included as (most of) the simulations did not successfully fit. Scenario 3-3 used the correct measurement error (2.89) and growth curve specification.

	Simple species		Complex species	
Scenario	$N^A_{\male}$	$N^A_{\female}$	$N^A_{\male}$	$N^A_{\female}$
ME-67:GC+5	125.31	123.96	62.82	92.58
ME-67:GC+10	64.96	64.47	30.90	52.52
ME-33:GC+0	181.21	178.88	103.71	139.35
ME-33:GC+5	125.23	123.87	62.73	92.47
ME-33:GC+10	65.01	64.50	30.99	52.58
ME+0:GC-5	232.63	229.27	160.18	201.50
ME+0:GC+0	180.93	178.60	103.08	138.79
ME+0:GC+5	125.01	123.64	62.51	92.17
ME+0:GC+10	65.21	64.68	31.11	52.61
ME+33:GC-10	254.39	250.33	217.42	263.40
ME+33:GC-5	230.61	227.25	158.05	199.53
ME+33:GC+0	180.33	177.97	102.19	137.94
ME+33:GC+5	124.30	122.90	62.13	91.63
ME+33:GC+10	65.21	64.66	31.20	52.54
ME+67:GC-10	249.34	245.36	212.67	259.08
ME+67:GC-5	228.22	224.87	155.44	197.12
ME+67:GC+0	179.28	176.91	101.03	136.79
ME+67:GC+5	123.19	121.79	61.63	90.90
ME+67:GC+10	65.02	64.45	31.28	52.38

Table B.2: Likelihood-based estimates for the coefficient of variation (CV; %) of the estimated parameters, averaged over 1000 simulations. Here, CV is defined as the Hessian-based estimated of the standard error divided by the parameter estimate. The first number in a scenario label refers to the assumed measurement error on length; the second number refers to the used growth function. Scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10 were not included as (most of) the simulations did not successfully fit. Scenario 3-3 used the correct measurement error (2.89) and growth curve specification.

	Simple species		Complex species	
Scenario	$N^A_{\male}$	$N^A_{\female}$	$N^A_{\male}$	$N^A_{\female}$
ME-67:GC+5	20.66	20.59	18.70	19.83
ME-67:GC+10	20.64	20.57	18.67	19.81
ME-33:GC+0	20.66	20.60	18.71	19.84
ME-33:GC+5	20.66	20.59	18.70	19.83
ME-33:GC+10	20.64	20.58	18.67	19.82
ME+0:GC-5	20.66	20.60	18.72	19.84
ME+0:GC+0	20.66	20.60	18.71	19.84
ME+0:GC+5	20.66	20.59	18.70	19.83
ME+0:GC+10	20.65	20.58	18.68	19.82
ME+33:GC-10	20.67	20.60	18.72	19.84
ME+33:GC-5	20.66	20.60	18.72	19.84
ME+33:GC+0	20.66	20.60	18.72	19.84
ME+33:GC+5	20.66	20.59	18.71	19.83
ME+33:GC+10	20.65	20.58	18.68	19.82
ME+67:GC-10	20.67	20.60	18.72	19.84
ME+67:GC-5	20.66	20.60	18.72	19.84
ME+67:GC+0	20.66	20.60	18.72	19.84
ME+67:GC+5	20.66	20.60	18.71	19.83
ME+67:GC+10	20.65	20.59	18.69	19.82

Table B.3: Empirical estimates for the standard deviation of the estimated parameters, derived over 1000 simulations. The first number in a scenario label refers to the assumed measurement error on length; the second number refers to the used growth function. Scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10 were not included as (most of) the simulations did not successfully fit. Scenario 3-3 used the correct measurement error (2.89) and growth curve specification.

	Simple species		Complex species	
Scenario	$N^A_{\male}$	$N^A_{\female}$	$N^A_{\male}$	$N^A_{\female}$
ME-67:GC+5	130.20	136.09	103.71	73.15
ME-67:GC+10	86.14	87.54	74.95	55.27
ME-33:GC+0	182.56	192.59	148.42	109.27
ME-33:GC+5	129.92	135.83	103.38	72.79
ME-33:GC+10	85.56	87.02	74.51	54.80
ME+0:GC-5	237.00	251.53	219.86	179.16
ME+0:GC+0	182.35	192.36	147.82	108.88
ME+0:GC+5	129.46	135.38	102.81	72.22
ME+0:GC+10	84.72	86.31	73.85	54.06
ME+33:GC-10	266.36	282.11	300.08	262.50
ME+33:GC-5	235.40	249.59	217.85	177.10
ME+33:GC+0	181.85	191.81	146.93	108.29
ME+33:GC+5	128.52	134.44	101.96	71.47
ME+33:GC+10	83.65	85.30	72.97	53.13
ME+67:GC-10	261.46	276.67	295.13	256.41
ME+67:GC-5	233.59	247.46	215.43	174.64
ME+67:GC+0	180.97	190.85	145.77	107.51
ME+67:GC+5	127.23	133.08	100.93	70.62
ME+67:GC+10	82.44	84.13	71.95	52.10

Table B.4: Empirical estimates for the coefficient of variation (CV; %) of the estimated parameters, derived from 1000 simulations. Here, CV is defined as the standard deviation of the error parameter estimates divided by the mean estimate for that scenario. The first number in a scenario label refers to the assumed measurement error on length; the second number refers to the used growth function. Scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10 were not included as (most of) the simulations did not successfully fit. Scenario 3-3 used the correct measurement error (2.89) and growth curve specification.

	Simple species		Complex species	
Scenario	$N^A_{\male}$	$N^A_{\female}$	$N^A_{\male}$	$N^A_{\female}$
ME-67:GC+5	22.16	23.03	22.74	22.21
ME-67:GC+10	28.12	28.50	28.88	33.97
ME-33:GC+0	21.56	22.58	21.68	20.17
ME-33:GC+5	22.13	23.01	22.70	22.14
ME-33:GC+10	27.92	28.32	28.68	33.60
ME+0:GC-5	21.87	23.00	22.28	21.50
ME+0:GC+0	21.57	22.59	21.69	20.23
ME+0:GC+5	22.10	22.97	22.65	22.05
ME+0:GC+10	27.58	28.01	28.42	33.03
ME+33:GC-10	22.54	23.63	23.32	23.29
ME+33:GC-5	21.92	23.03	22.30	21.55
ME+33:GC+0	21.60	22.60	21.70	20.30
ME+33:GC+5	22.07	22.95	22.60	21.96
ME+33:GC+10	27.25	27.70	28.13	32.38
ME+67:GC-10	22.58	23.65	23.32	23.26
ME+67:GC-5	21.99	23.08	22.32	21.61
ME+67:GC+0	21.62	22.63	21.71	20.39
ME+67:GC+5	22.05	22.92	22.56	21.88
ME+67:GC+10	26.95	27.40	27.83	31.69

Appendix C Modelling population growth

In an alternative version of this model we included a growth parameter r, which allowed for the estimation of exponential growth or decline in the population size. Including growth in the model meant that the distribution of age given measured length $f(a|\ell^*)$ is no longer defined exclusively by survival parameter ϕ but rather by a combination of ϕ and r (Caswell, 2000, Section 4.5.2.1), given that the population reached a new stable age distribution. Hillary et al. (2018) estimated $f(a|\ell^*)$ by grouping the data and fitting a multinomial distribution, however this was beyond the scope of our research. We did run our simulation study with yearly growth parameters $r_{\female}$ and $r_{\male}$ assuming that the $f(a|\ell^*)$ as presented in Chapter 4 was *approximately* correct. We decided not to include this part of the study in the main body of this thesis, as we did not believe that the results could be used to accurately assess the effect of incorrect ageing on parameter estimation. Nonetheless, we included these results here for completeness as they could contain some valuable insight and form the basis for future research. As we considered abundance for both sexes separately, we estimated the following four parameters: $N^A_{\female,y_0}$, $N^A_{\male,y_0}$, $r_{\female}$ and $r_{\male}$, where y_0 is some reference year.. The kinship probabilities remained the same as presented in the Equations (4.4)–(4.7). As population size was no longer assumed equal for all years, abundances in different years are linked through a geometric population dynamics model:

$$N_y = N_{y_0} r^{y-y_0}, \tag{13}$$

where $r \in (0, \infty)$ denotes the yearly growth rate. We set $y_0 = 2014$ to match Bradley et al. (2017) as closely as possible.

C.1 Estimated abundance through time

We fit our 25 scenarios, consisting of all combinations of 5 different measurement errors and 5 different growth curves, to both populations. We modelled the male and female side of the population separately resulting in four figures, each containing 19 population history plots — six plots are blank since the models in these scenarios did not (all) fit correctly. In Figures C.1– C.4 we notice a similar pattern of over- and underestimation related to shifting the growth curves. However, as we also model exponential growth or decline, we also notice effects of shifting the growth curves on the direction and magnitude of this trend. Albeit potentially informative, due to the inconsistency between modelling growth and the assumed age distribution we believe that these results can not be directly used for inference.

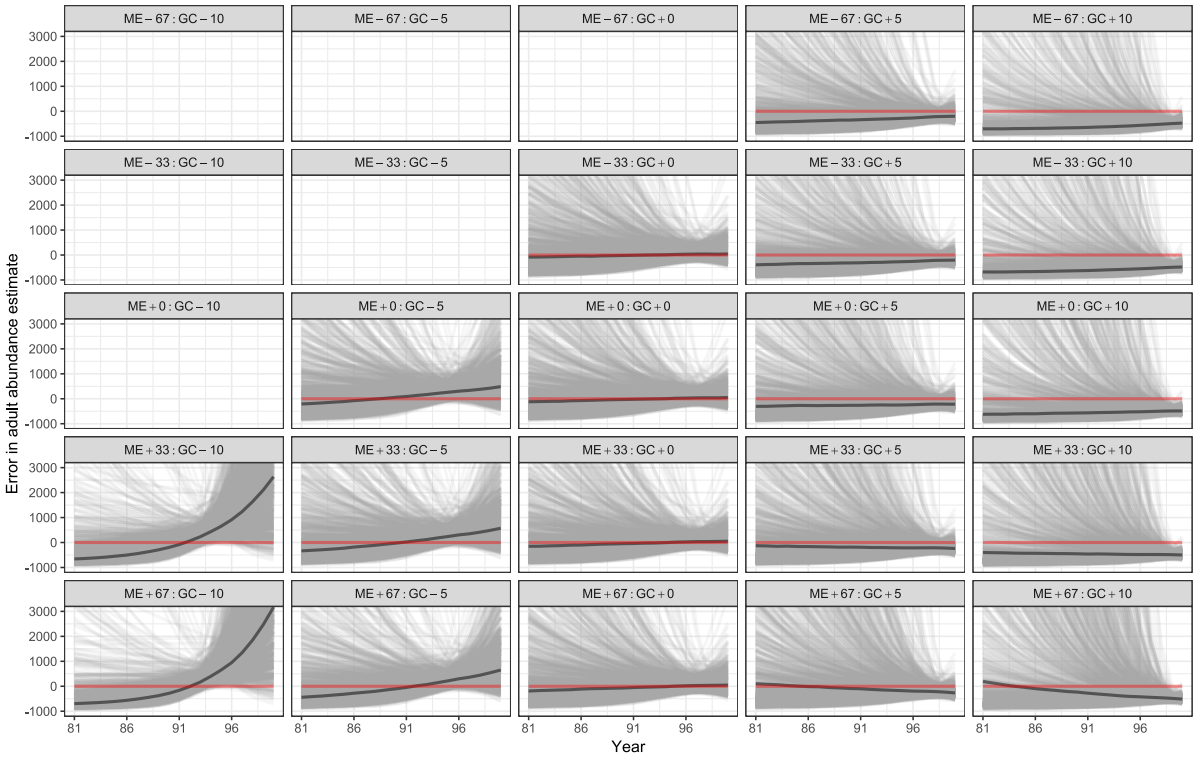

Figure C.1: Plots of the 1000 simple female adult population trends for the simple species for the last twenty years of the simulation, for the 19 scenarios that resulted in successful fits. The median of these 1000 trends is indicated in dark grey, and the truth adult abundance is indicated in red. The scenarios were labelled using the format 'ME±XX:GC±YY', where ME refers to the measurement error, XX denotes the percentage over- or underestimate, GC stands for growth curve, and YY denotes the percentage of up- or downwards shifting; for example, the scenario with a 33% overestimated length measurement error and a 5% downshifted growth curve had label ME+33:GC-5.

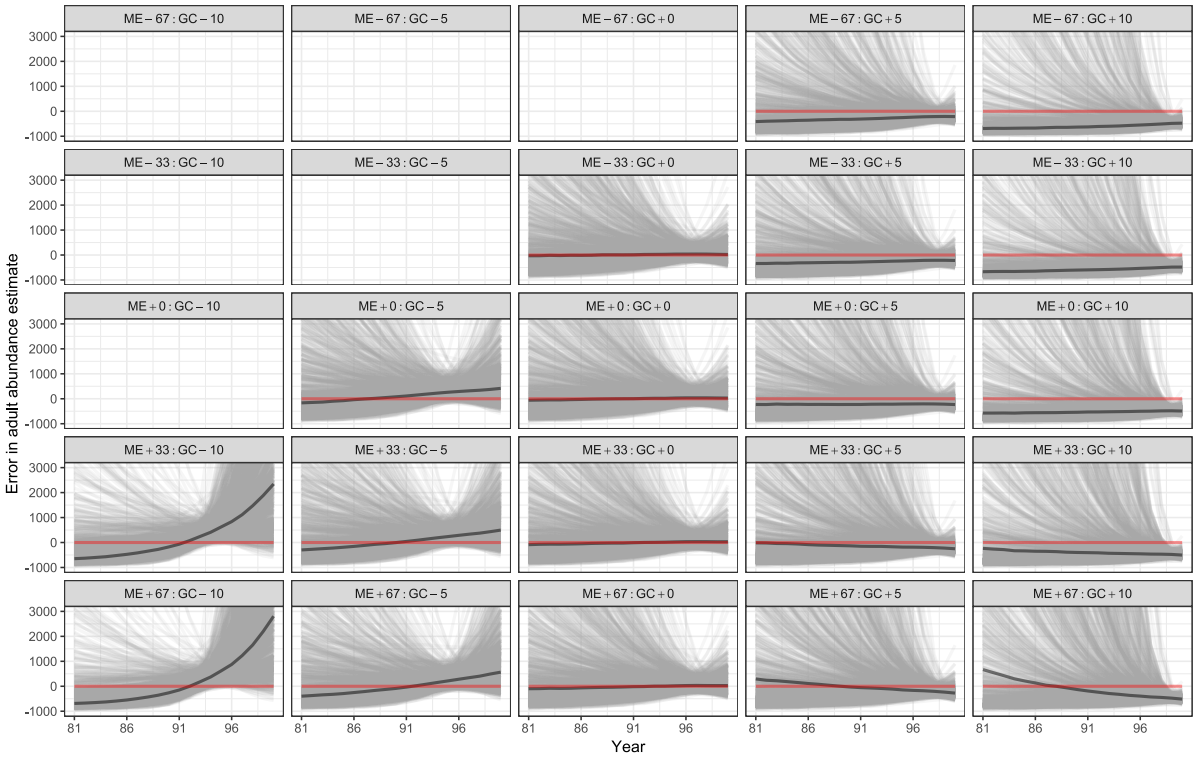

Figure C.2: Plots of the 1000 estimated male adult population trends for the simple species for the last twenty years of the simulation, for the 19 scenarios that resulted in successful fits. The median of these 1000 trends is indicated in dark grey, and the truth adult abundance is indicated in red. The scenarios were labelled using the format 'ME±XX:GC±YY', where ME refers to the measurement error, XX denotes the percentage over- or underestimate, GC stands for growth curve, and YY denotes the percentage of up- or downwards shifting; for example, the scenario with a 33% overestimated length measurement error and a 5% downshifted growth curve had label ME+33:GC-5.

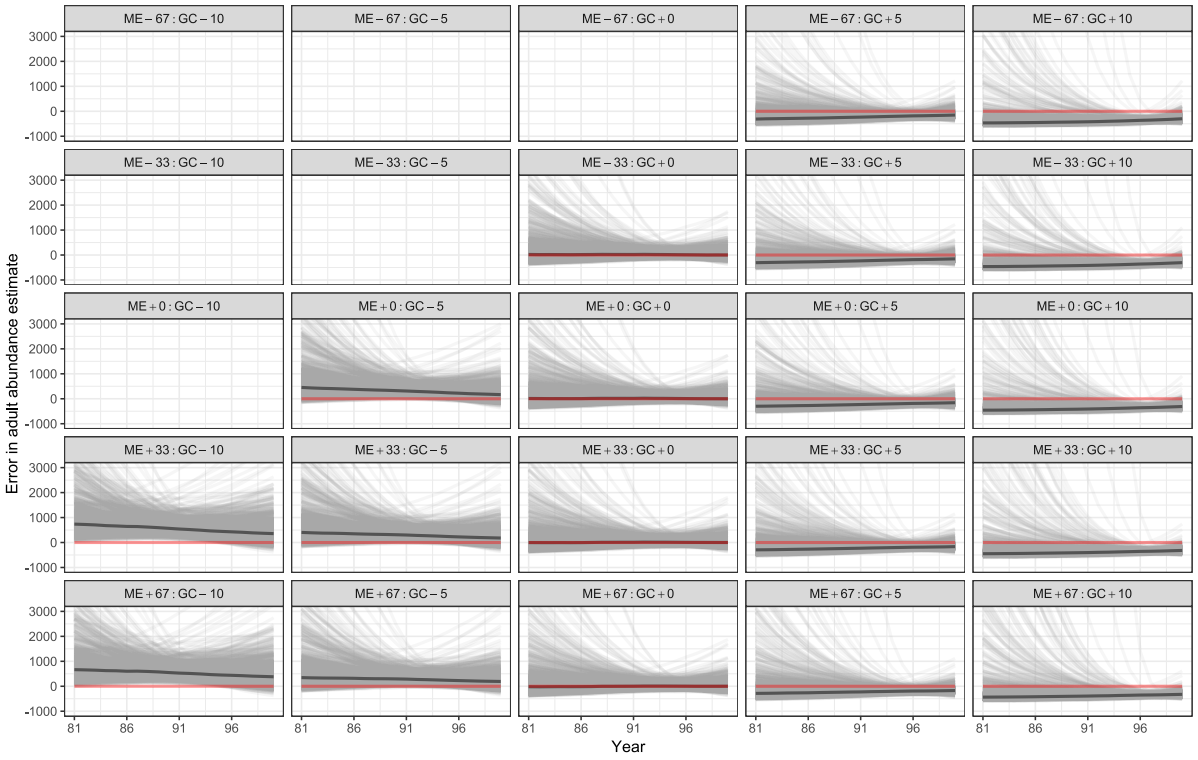

Figure C.3: Plots of the 1000 estimated female adult population trends for the complex species for the last twenty years of the simulation, for the 19 scenarios that resulted in successful fits. The median of these 1000 trends is indicated in dark grey, and the truth adult abundance is indicated in red. The scenarios were labelled using the format 'ME±XX:GC±YY', where ME refers to the measurement error, XX denotes the percentage over- or underestimate, GC stands for growth curve, and YY denotes the percentage of up- or downwards shifting; for example, the scenario with a 33% overestimated length measurement error and a 5% downshifted growth curve had label ME+33:GC-5.

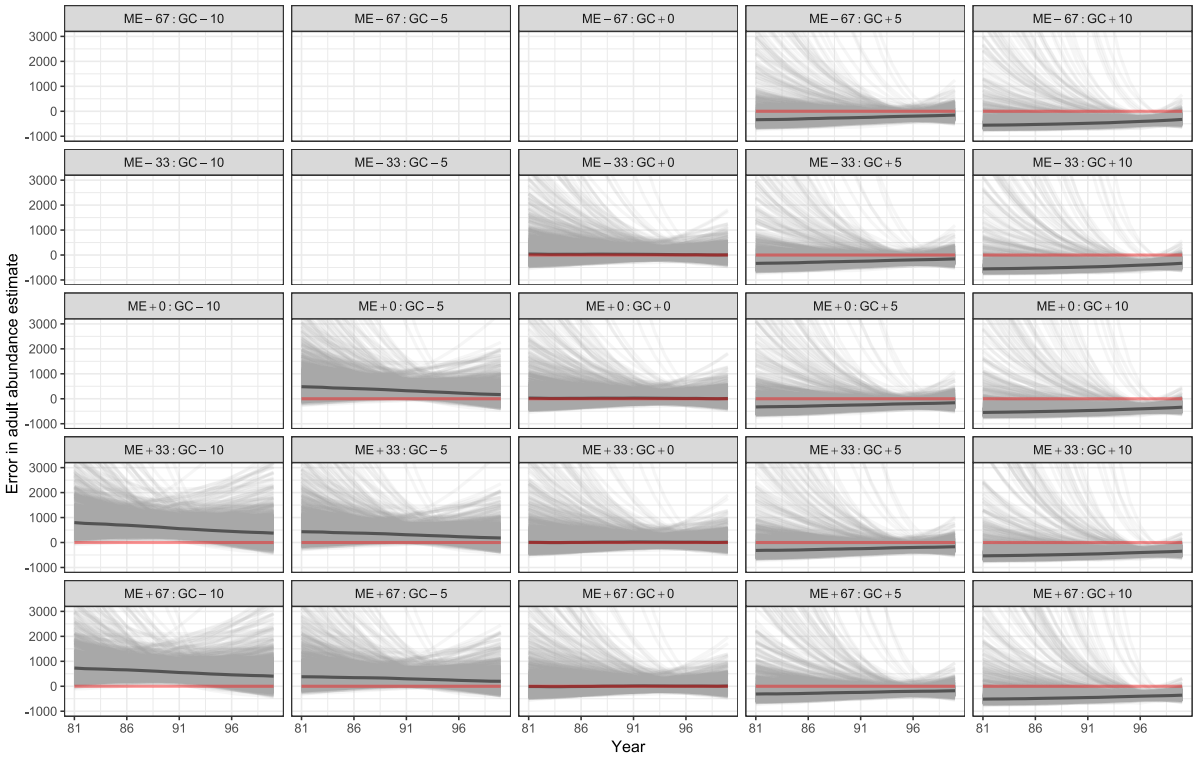

Figure C.4: Plots of the 1000 estimated male adult population trends for the complex species for the last twenty years of the simulation, for the 19 scenarios that resulted in successful fits. The median of these 1000 trends is indicated in dark grey, and the truth adult abundance is indicated in red. The scenarios were labelled using the format 'ME±XX:GC±YY', where ME refers to the measurement error, XX denotes the percentage over- or underestimate, GC stands for growth curve, and YY denotes the percentage of up- or downwards shifting; for example, the scenario with a 33% overestimated length measurement error and a 5% downshifted growth curve had label ME+33:GC-5.