# Combining individual and close-kin mark-recapture to design an effective survey for Pacific walrus

Eiren K. Jacobson[1*], Mark V. Bravington[2], Rebecca L. Taylor[3], Irina S. Trukhanova[4,5], David L. Miller[6], William S. Beatty[3,7]

18th January 2025

[1] Centre for Research into Ecological & Environmental Modelling and School of Mathematics & Statistics, University of St Andrews, St Andrews, Scotland

[2] Estimark Research, Hobart, Australia

[3] US Geological Survey, Alaska Science Center, Anchorage, Alaska

[4] North Pacific Wildlife Consulting, LLC, Seattle, Washington

[5] US Fish and Wildlife Service, Marine Mammals Management, Anchorage, Alaska

[6] 165 Perth Road, Dundee, Scotland

[7] US Geological Survey, Upper Midwest Environmental Sciences Center, La Crosse, Wisconsin

[*]Corresponding author email: ej45@st-andrews.ac.uk

## Acknowledgments

## Data availability

## Conflict of Interest statement

## Author contribution statements

## Statement on inclusion (optional)

**Abstract**

The Pacific walrus *(Odobenus rosmarus divergens)* is an ice-associated marine mammal found in the Bering and Chukchi Seas, where they have been hunted for subsistence for time immemorial. In the late 20th century, the population declined, likely because it had reached carrying capacity and was subject to high harvests. Currently, Pacific walrus is species of conservation concern due to the potential impacts of climate change, particularly related to loss of sea ice. To reduce uncertainty in estimates of population size and trend, researchers undertook an individual genetic mark-recapture (IMR) sampling campaign from 2013-2017 and collected tissue samples from over 8,000 individuals. Another campaign of a similar scale is ongoing (2023-2027). While sample collection was designed for IMR, advances in close-kin mark-recapture (CKMR) methodology and associated molecular techniques mean these samples could also be suitable for CKMR. The advantages of CKMR over IMR include increased effective sample size (since each individual tags not only itself, but also its parents, siblings, and offspring) and additional insights into demographic quantities of interest. Here, we combine individual and close-kin mark-recapture in a single modelling framework (ICKMR) and investigate whether different sampling strategies can increase precision in estimates of abundance and trend. Our modelling approach includes special considerations for walrus life-history, including a multi-year inter-birth interval. We implement our model in R and use an individual-based simulation to test performance of the ICKMR model. Something here about survey design. We find that the expected precision of the ICKMR estimates of abundance are higher than those expected from IMR alone. This result suggests that ICKMR is a promising approach for assessing population size and trend of species which have been difficult to survey using more traditional methods. [285/350]

Keywords: Close-kin mark-recapture, individual genetic mark-recapture, survey design, walrus

# 1 Introduction

Estimation of abundance and of other demographic parameters such as survival is a key part of wildlife management and conservation. Traditional mark-recapture analysis (Williams et al., 2002) can deliver estimates with low bias and uncertainty, provided that enough individual animals i) are naturally, artificially, or genetically "marked" and identifiable and ii) can be recaptured over time. If genotypes are used as the marks, as in genetic individual mark-recapture (IMR; Palsbøll et al., 1997), then kinship patterns amongst the samples (parents, siblings, etc) contains additional information relevant to demographics (Skaug, 2001). Close-kin mark-recapture (CKMR; see Bravington et al., 2016 for more detail on the points made below) is a framework for using these kinships, as inferred from genotypes, to estimate abundance and demographic parameters. CKMR provides additional flexibility compared with IMR since lethal samples (from sampling, hunting, natural mortality etc.) and/or non-lethal samples can be used. As of 2025, most CKMR projects have been for commercial fish (e.g., Davies et al., 2020) or sharks (e.g., Hillary et al., 2018), but there are also some for mammals, including Conn et al. (2020)'s modeling study of bearded seals and its implementation by Taras et al. (2024), and Lloyd-Jones et al. (2023) for flying foxes.

The principle behind CKMR is that every individual has (or had) one mother and one father; thus, for a given sample size, in a large population there will be few "recaptures" of parents or their other descendants, while in a small population there will be many. In practice, the data for CKMR comprise the outcome of pairwise kinship checks amongst samples (e.g., whether two particular samples are genetically parent-and-offspring), plus covariates associated with each sample such as its date of capture, age, size, sex etc. The CKMR model has two components: a population-dynamics part driven by the demographic parameters; and formulae for the expected frequencies of different kinship types in pairwise comparisons, conditional on sample covariates and population dynamics. By combining the kinship data with the model, parameters can be estimated using maximum-likelihood or Bayesian methods.

Most are potential candidates for CKMR, except semelparous (breed-once-and-die) species and a few pathological cases mentioned Bravington et al. (2016). The data requirements are otherwise quite flexible; the kinships used in any particular study can vary depending on logistic and modeling considerations. However, CKMR does require fairly unambiguous pairwise kinship determination which, at least with current genetic methods, limits the usable types of "close" kin to three or four: Parent-Offspring Pairs (POPs), Full-Sibling Pairs (FSP), and "second-order kin", which are primarily

Half-Sibling Pairs (HSP; one shared parent). Any other kinship constitutes an "Unrelated Pair" (UP) for modelling purposes, even if there is genetic evidence of weak relatedness. Some information about sample age is required. Often this comes from "hard parts" such as teeth or otoliths or, much less accurately, from covariates such as body size; however, the advent of epigenetic ageing (Polanowski**; Mayne**lungfish) makes it possible to use just the biopsy sample itself. Uncertainty in age and other covariate estimates can be accommodated within the model, though the precision of parameter estimates is of course linked both to sample sizes, and to the accuracy of covariate measurements.

Devising the formulae for the expected frequencies of kinships requires consideration of the life history and reproductive biology of the organism in question, as well as aspects of the sampling. Knowing about reproductive strategy is also important for both sampling and probability of kinship: does the population breed in one place and at one time? Does the species have a polygynous mating structure? Are breeding animals as likely to be sampled as "resting" ones?

Need a transition here

To design an effective survey, it is important to consider whether desired monitoring goals can be met given available time, money, and effort. Because of the genetic component, CKMR and genetic IMR do incur extra costs relative to, say, photographic IMR, but that aspect of cost continues to decrease as new sequencing technology comes online. The non-laboratory costs remain as they ever were for mark-recapture (or any other observational study): person-hours in the field. As is the case with all fieldwork, researchers aim to maximize information gathering while minimizing cost and also keeping people safe. This leads us to the question of "how many samples?". In the case of CKMR, this is relatively easy to calculate. Given some information on demographics and abundance (from existing genetic surveys, pilot studies, or the literature), we can calculate how many kin pairs one is likely to obtain from a given number and composition (e.g., by age) of samples, and what CV is likely to be obtained for various demographic parameters of interest. In this paper we show how to do these calculations using a case study on the Pacific walrus (*Odobenus rosmarus divergens*; hereafter, walrus) in the North Pacific.

The walrus is a gregarious, ice-associated pinniped inhabiting continental shelf waters of the Bering and Chukchi seas. During winter (when sea ice forms south of the Bering Strait) virtually all walruses occupy the Bering Sea (Fay, 1982). In summer (when sea ice is absent from the Bering Sea) almost all juvenile and adult female walruses, and some adult male walruses, migrate north to the Chukchi Sea. When walruses rest offshore on sea ice floes, their distribution is dynamic, because it generally follows

the marginal ice zone (a moving, changing habitat which contains a mix of ice floes and water) but also concentrates in regions of high benthic productivity. This allows walruses to forage for benthic invertebrates while simultaneously having access to a nearby substrate for hauling out.

Sea ice has declined for decades (Perovich and Richter-Menge, 2009; Stroeve et al., 2012; Stroeve and Notz, 2018), and coupled global atmospheric-ocean general circulation models predict its continued decline (Årthun et al., 2021). When sea ice recedes from the continental shelf, walruses come on shore to rest in large herds at sites termed haulouts, from which they make long trips to foraging hotspots (Jay et al., 2012). This change in their activity budgets (Jay et al., 2017) may ultimately lead to a decline in body condition and an increase in mortality or a decrease in reproduction (Udevitz et al., 2017). Furthermore, disturbance at haulouts can cause stampedes, resulting in mass calf and juvenile mortality. Continued sea-ice loss and a concomitant increase in the intensity and expansion of industrial and shipping activities in Pacific Arctic waters (Silber and Adams, 2019) are expected to drive a substantial population decline (Garlich-Miller et al., 2011; MacCracken et al., 2017; Johnson et al., 2023; Johnson et al., 2024).

Range-wide abundance and demographic rate estimates are crucial for understanding population status, as well as for developing and implementing harvest management plans. In particular, subsistence walrus harvests in Alaska and Chukotka exceed 4,000 animals annually (USFWS, 2023), and indigenous peoples need information on the status of the walrus population in order to manage these harvests sustainably. Furthermore, in the United States, the Marine Mammal Protection Act (MMPA) requires a determination of potential biological removal for walrus, which in turn, requires a precise abundance estimate (Gilbert, 1999; Wade and DeMaster, 1999).

Scientists have attempted to ascertain walrus population size since at least 1880 (Fay et al., 1989), and until very recently, unsuccessfully. The most concerted effort was the 1975-2006 range-wide airplane-based surveys conducted collaboratively with the Soviet Union and then Russian Federation. However, resulting estimates were biased and imprecise, and count-based methods were abandoned after the 2006 survey which, despite a rigorous design, innovative field methods, and sophisticated analyses, yielded a 95% confidence interval (CI) on the population size estimate of 55,000–507,000 animals (CV = 0.93). The extensive imprecision in the estimate resulted from the walrus population being widely dispersed with unpredictable local clumping (Speckman et al., 2011; Jay et al., 2012), which is, in turn, due to the large area of arctic and subarctic continental shelf over which they forage, their gregarious nature, and the dynamic nature of the marginal ice zone.

The first rigorous walrus survival rate estimates were obtained within the past decade via Bayesian integrated population models (IPMs), which combined multiple data sources to estimate demographic rates and population trend over multiple decades (Taylor and Udevitz, 2015; Taylor et al., 2018). However, the original problems with the aerial survey data continued to preclude conclusions about population abundance in the IPMs (Taylor and Udevitz, 2015).

In 2013, the U.S. Fish and Wildlife Service (FWS) initiated a genetic IMR project to estimate walrus abundance and demographic rates. Under this approach, genetic "marking" via skin biopsy samples (Palsbøll et al., 1997) provided a major advantage over traditional marking techniques because walruses are extremely difficult to handle physically. Over five years of research cruises, biologists attempted to collect a representative sample of walruses in the accessible portion of the marginal ice zone in each year a cruise was conducted, although Russian waters were not accessible in all years. Sampling focused on groups of adult females and juveniles, as these classes are the demographically important population segments of this polygynous species (Fay, 1982). Further methods for the IMR study are detailed by Beatty et al. (2020; 2022).

Data analysis from the first generation of walrus research cruises (2013–2017) used a Cormack-Jolly-Seber multievent model to estimate survival rates, and a Horvitz-Thompson-like estimator to obtain population size. The total abundance of 257,000 had a 95% credible interval (CrI) of 171,000–366,000 (CV=0.19; Beatty et al. 2022). Although the precision of the abundance estimate from the IMR study was much improved over the final aerial survey, the IMR study required extensive investment of human and financial resources (i.e, USD $5,000,000). A more cost-effective approach is needed to assess the walrus population on a regular interval. As mentioned above, biopsy samples also contain information about kin relationships, which, through CKMR, can substantially augment the information content of genetic IMR without increasing sampling effort.

***NEED something about walrus moving about all over the goddamn place, from IMR data and (more likely) sat tags :) Some of that *could* go to the Discussion, but I think aat least a pre-mention here, coz it will otherwise be in the alert reader's mind as they look at the model structure

CKMR has mostly been used in situations where self-recaptures are unlikely or impossible (e.g., because the population is very large, or because sampling is lethal) although Lloyd-Jones et al. (2023) did include IMR in a CKMR study but not integrated into a single model. Here, we focus on a population where IMR was the original project goal and therefore we extend traditional CKMR to include IMR in the same model as an additional kinship type. We then explore different demographic

and design scenarios for walrus using IMR alone versus CKMR + IMR = ICKMR and demonstrate how the latter can be used to substantially reduce the overall amount of survey effort required for adequate monitoring. [1665 words]

***NEED a description of how the rest of the paper works— unless the Methods section kicks off with that, I guess; haven't looked yet.

# 2 Methods

To evaluate our proposed survey designs, we must first construct our ICKMR model for walrus. We encode our knowledge about walrus biology and life history to (*i*) build a model of walrus population dynamics, including the breeding cycle, and (*ii*) formulate kinship probabilities between pairs of individuals. The population dynamics model incorporates demographic parameters that will need to be estimated: survival rates, adult abundance in some reference year, trend, and so on. The kinship probabilities depend on the population dynamics. Given a real dataset, we would estimate the parameters by maximizing the log-likelihood that combines the kinship probabilities with the actual outcomes of all pairwise comparisons. For design purposes, though, we instead use a computational shortcut to predict the precision of the estimates that would be expected under different sampling designs. Although it is not strictly necessary to simulate any data in this process, we did use simulations to check that our CKMR model was appropriately formulated. This section describes our population dynamics model, kinship probability formula, design calculations, and simulation setup.Some of that para might repeat stuff in the Intro, or might better belong there. dunno

## 2.1 Biological considerations

Adult males are inaccessible to this study given seasonal sex-segregation and the geographical coverage of sampling effort (see Section 1). They also form leks and compete for breeding access to females, so it is plausible that adult males might also exhibit persistent individual variability in breeding success, which would considerably complicate the interpretation of paternal half-sibling kinship data (see Discussion). Therefore, we restrict attention to female-only dynamics, and consider only three types of kinship: Mother-Offspring Pair (MOP), Cross-cohort maternal Half-Sibling Pair (XmHSP), and Self Pair (SP), i.e. recapture of an individual. Our samples comprise juvenile and adult females, plus juvenile males; the problems are with modeling males as parents, but we can safely use sampled

juvenile malesas potential offspring of females and as potential maternal half-siblings of other (female or male) samples, as below. We do not expect females to vary much in terms of individual fecundity. The population is well-mixed across its range, with no evidence of site fidelity within or between years (CITE)I reckon this should be covered in the Intro, and xrefed here back to that. Also, NB that "genetic substructure" (ie classical pop gen differences) is not the point with CKMR; if you have non-heritable site fidelity, you're still in trouble...coming next...

We assume that age estimates will be available for all samples, based on epigenetic data (ref**). Visual classification is only accurate over the first couple of years of life (ref**), which could be problematic for CKMR. Our model is structured to allow for errors in estimated age (with standard deviation assumed known, i.e. after calibration of epigenetic against known-age samples), though the results here assume that there are no errors; see Discussion.

### 2.1.1 Stage-structured quasi-equilibrium dynamics

For our female-only population dynamics model, we could choose between a fully-age-structured or a stage-structured (juvenile+adult) approach. We opted for the the latter because (i) most female adults are expected to have similar reproductive capacity and chance of survival regardless of age (unlike with teleost fish, say); and (ii) stage-structured models are simpler to code for CKMR and require fewer parameters— although with IMR data as well, the simplicity claim is weaker, For real data analysis, we might move to an age-structured formulation (see Discussion) instead, but the stage-structured results should be quite adequate for design purposes; the fundamental role of total (not age-specific) adult abundance and survival is very similar in both.

We used two stages: juveniles aged 1–5, and adults aged 6+ (the first age at which an accompanying calf is common). We did not include calves (age 0) in the model or in the samples considered, to avoid extra parameters and complications around mother-calf sampling. "Age" here means "at sampling", i.e. in the summertime; births are in spring. We assume that survival over age and time within each stage is constant ($\phi_{\mathrm{A}}$ and $\phi_{\mathrm{J}}$), and that an offspring's survival from age 1 onwards is independent of its mother's survival, whether or not the offspring has weaned yetWalrus reprod biol should describe how this stuff works. It's uttlerly different from most seals.. Further, we assume that adult female abundance is increasing or decreasing exponentially over the period covered by the population dynamics (2000–2027; the lower limit of $y_0$ =2000 is set because there were fairly drastic changes in the population prior to that). we haveElsewhere we have $N(y, A)$ I think. I don't mind the bracketed version. If we

do go with subscripts instead, then I'd prefer to drop commas wherever possible, coz theyre bloody hard to read in subscript font. EG $N_{y_0 A}$ is fine coz there's no ambiguity. Also I wonder if we have a bit of notational trouble here: $N_{ya}$ usually means "numbers at year and age" (with fish...) and we'll need that below; but here $N_{y\mathrm{A}}$ is meaning "all Adults"...

$$N_{y,\mathrm{A}} = N_{y_0,\mathrm{A}} e^{r(y-y_0)} \tag{1}$$

where $N_{y,\mathrm{A}}$ is the abundance of adult females in year $y$. $e^r$ is the erate of population increase.

Age composition within stage is not required for calculating MOP and XmHSP probabilities as below, but it is relevant for SPs. For that purpose, we assume that age composition over the period is adequately described by the stable-age or "quasi-equilibrium" distribution consistent with survial $\phi_{\mathrm{A}}$ and rate-of-increase $e^r$. As shown in e.g. Keyfitz and Caswell (2005) section 5, this iscan someone please check this; I've just typed from intuition!

$$N_{ya} \propto N_{y\mathrm{A}} \phi_{\mathrm{A}}^a e^{-ra} \tag{2}$$

### 2.1.2 The breeding cycle

It's confusing to have digits used for the three states, but also for ages! eg "females enter state 3 at age 4"... Could be alleviated by always calling the states B1,B2,B3 or S1... I also suspect this is a bit wordy, esp given the nice figure.We use a Markov model to describe the walrus breeding cycle. We assume three breeding states: (1) pregnant, (2) with young-of-the-year (YOTY) calf or, (3) non-breeding; i.e., neither of the above. The Markov property assumes that next year's state depends only on this year's. From state 1 (pregnant), next year's state must be 2 (with YOTY calf). From state 2, a female may next year either return to state 1 (become pregnant again), with probability $\psi_2$, or move to state 3 (neither pregnant nor with calf) with probability $1 - \psi_2$. From state 3, she will either move to state 1 (become pregnant) with probability $\psi_3$, or remain in state 3 with probability $1 - \psi_3$. Due to long gestation times (~14 months), walrus cannot give birth to calves in two consecutive years (CITE). We also allow $\psi_2 \neq \psi_3$ as they are unlikely to give birth to calves every second year (CITE). This is shown in Figure 1. Survival is assumed to be independent of breeding state. Females enter state 3 (i.e., reach sexual maturity) on reaching age 4, and therefore can become pregnant at age 5 and give birth at age 6. Depending on the values of $\psi_2$ and $\psi_3$, this leads to a ramping-up in effective fecundity (i.e.,

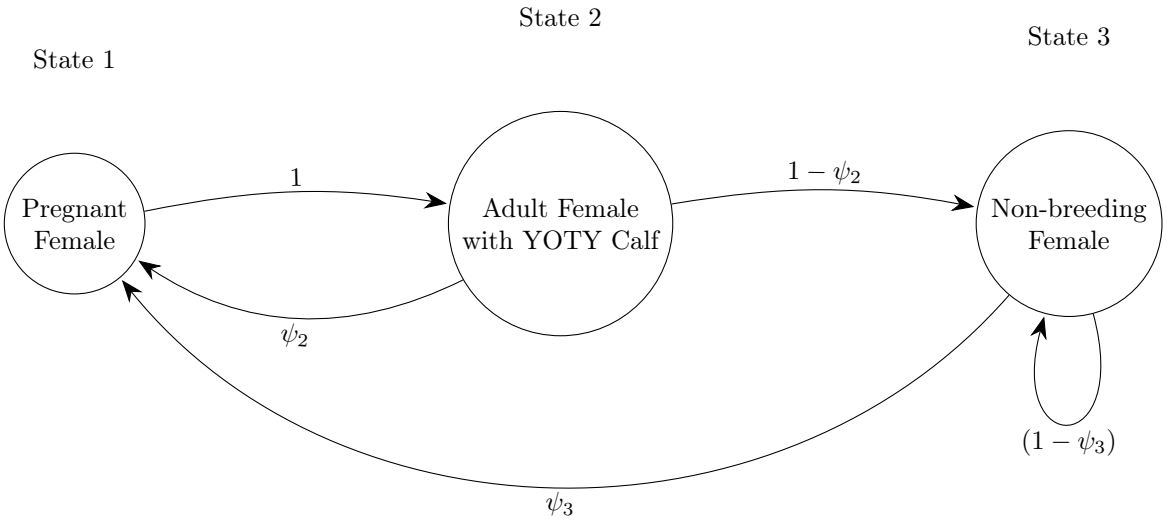


Figure 1: Directed cyclic graph showing the breeding cycle for walrus as represented in our Markov model. Nodes in the graph show the states (pregnant, with young-of-the-year (YOTY) calf, or non-breeding) and edges give the probabilities of transition between those states. Walrus reach sexual maturity at age 4, so enter the graph at node non-breeding.

probability of being in state 2) over the first few years of adult life. Both $\psi_2$ and $\psi_3$ are estimated from the data. We do not use any data on whether females were with/without calf when sampled, so the information on those parameters comes indirectly, through the distribution of birth-gaps between maternal half-sibling pairs.

### 2.1.3 Formulating kinship probabilities

There's scope for confusion between "sample" and "individual" here. In a purist and not-helpful sense, comparisons are just between samples. In fact, we take trouble to only compare a single sample from each individual (in those cases where there's >1) with other reprsenetative samples. But when we get onto SPs, then it really is samples... We now need to formulate the demographic probabilities that two ¿samples/individuals? have a given kinship. We write the kinship for individuals $i$ and $j$ as $K_{ij}$, which in our case may be MOP, XmHSP, SP, or UP (i.e. none of the others).I'd drop all that; it's substantially not right as it is presently written (eg AFAIK expected number of juveniles never matters! and the number of possible $i$'s doesn't either) and it's actually really hard to get right in a general sense. And then there's SelfPs.... I get the reason for wanting to say something, but it's too hard! The core concept is ERRO, as per BSA2016, but that doesn't cover SPs. Let's just plunge into specific cases...

Throughout we use the following notation: for individual $i$, sampled at age $a_i$ in year $y_i$ with birth year $b_i \triangleq y_i - a_i$. As noted above, we only consider female abundance, so throughout, $N$ refers to females only. When female abundance is considered for a given year and development stage (respectively), it is written with two arguments, $N_{y,\mathrm{D}}$ (note that we use upright letters for development stages A and J)$N_{y\mathrm{A}}$ has already been used in the pop dyn bit, so sequence has gone awry. We define the binary variable $L$no commas coz there is more than one binary variable in existence to indicate lethality of sampling ($L_i = 1$ indicating lethal sampling for individual $i$). We use $\mathbb{I}()$ as an indicator function, giving 1 when the condition inside the brackets is true, else 0. Kinship probabilities are functions of demographic parameters such as $\phi_{\mathrm{A}}$and $N_{y_0\mathrm{A}}$; throughout we use $\boldsymbol{\theta}$ as shorthand for the full set of parameters, which become explicit in later iterations of the formulae.

### 2.1.4 Mother-offspring pairs (MOPs)

Suppose we are about to compare a potential mother individual $i$, sampled at age $a_i$ in year $y_i$, to a potential offspring $j$, sampled at age $a_j$ in year $y_j$, and therefore born in year $b_j \triangleq y_j - a_j$. We restrict attention to comparisons that satisfy the following:

- $i$ is female (though $j$ need not be) (only female adult dynamics are being modelled);
- $a_j \geqslant 1$ (no calf samples);
- $b_j \geqslant 2000$ (population dynamics runs from year 2000).

We can now distinguish two cases: $y_i < b_j$ and $y_i \geqslant b_j$.

For $y_i < b_j$, individual $i$ still has to survive several years in order to be individual $j$'s mother (note that $i$ may be immature when sampled, but mature by the time of $j$'s birth). In this case $i$'s sampling must be non-lethal ($L_i = 0$). The MOP probability is

$$\mathbb{P}\left[K_{ij} = \mathrm{MOP} | a_i, y_i, b_j, L_i = 0, \boldsymbol{\theta}\right] = \frac{R_i(b_j|y_i, a_i)}{R^+(b_j)}. \tag{3}$$

Where $R_i(b_j|y_i, a_i)$ is the expected reproductive output (ERO) of individual $i$ in year $b_j$ given $i$ is age $a_i$ in year $y_i$. $R^+(b_j)$ is the total reproductive output (TRO) of the whole population in year $b_j$. The units of ERO and TRO can be arbitrary (as long as they are the same), but here it makes sense to pick "number of calves". TRO is the total number of adult females in the population when $j$ is born, $N_{b_j,\mathrm{A}}$, multiplied by the proportion of females with calves (breeding state 2), $\bar{\beta}_2$: $R^+(b_j) = \bar{\beta}_2 N_{b_j,\mathrm{A}}$.

There are two components to $i$'s ERO: first, she has to survive; second, she has to be calving (breeding state 2) in $b_j$:Where is $\bar{\beta}_2$ calculated? Needs to be here... unless it's done earlier, in the breeding cycle bit.

$$R_i\left(b_j|y_i,a_i\right)=\Phi\left(y_i-b_j,a_i\right)\mathbb{P}\left[B\left(a_i+b_j-y_i\right)=2\right], \tag{4}$$

where $\Phi\left(\Delta t,a\right)$ gives the probability of survival for $\Delta t$ years, starting from age $a$ (product of annual juvenile and adult survival probabilities). $B\left(a\right)$ is an individual's breeding state in year $a$, which here is individual $i$'s age at $b_j$ ($a_i+b_j-y_i$, assuming she survives).

Define fecundity as a function of age

$$F\left(a\right)\triangleq\frac{\mathbb{P}\left[B\left(a\right)=2\right]}{\bar{\beta}_2}, \tag{5}$$

so that immature animals have fecundity 0, and an average adult has fecundity 1. Thus

$$\mathbb{P}\left[K_{ij}=\text{MOP}|a_i,y_i,b_j,L_i=0,y_i<b_j,\boldsymbol{\theta}\right]=\frac{\Phi\left(y_i-b_j,a_i\right)F\left(a_i+b_j-y_i\right)}{N_{b_j,\text{A}}}. \tag{6}$$

If $i$ is sampled after $j$'s birth ($b_j<y_i$). We then know $i$ was alive (or not born yet), so there is no survival term to worry about, but she may not have been mature. Letting $F\left(a\leqslant 0\right)=0$ (walruses cannot breed before birth),

$$\mathbb{P}\left[K_{ij}=\text{MOP}|a_i,y_i,b_j,L_i=0,b_j<y_i,\boldsymbol{\theta}\right]=\frac{F\left(a_i-y_i-b_j\right)}{N_{b_j,\text{A}}}. \tag{7}$$

### 2.1.5 Maternal half-sibling pairs (XmHSPs)

We now find the probabilities of cross-cohort (X), maternal (m), half-sibling pairs (HSP): XmHSPs. We want to compare individual $k$, born in year $b_k$ (calculated from age at sampling) to individual $l$, born in $b_l$, to check whether they have the same mother. We impose the following criteria:

- $b_l>b_k$ (avoiding double-counting);
- $b_k\neq b_l$ (walrus give birth to a single offspring at a time);
- $b_k\geqslant 2000$ (population dynamics model starts at 2000).

If $m$ is the mother of $k$, what is the probability that $\ell$'s motherIt's hard to read script-$l$ vs 1 or script-

$i$— so I recommend $\ell$. I have changed a couple here, but not done all. was $m$, given what we know about $m$? The latter amounts to two things: $(i)$ $m$ was alive, mature, and in breeding state 2 at $k$'s birth and, $(ii)$ $m$ survived at least one more year after $k$'s birth, otherwise $k$ would not have lived long enough to be sampled. In order for $m$ to be $\ell$'s mother, three things have to happen:

1. she survives until $b_\ell$;
2. she is in breeding state 2 in $b_\ell$;
3. amongst all the females that are alive and in the right breeding state in year $b_l$, she is the mother.

Let $\Phi(\Delta t)$ be the adult probability of survival for $\Delta t$ years from "now" and let $\Psi$ be the (3×3) breeding cycle transition matrix. The probability 3-vector of an animal being in each state (1,2,3) at time $t$ is $p^{[t]}$. The probability vector at time $t+1$ is then $p^{[t+1]} = \Psi p^{[t]}$. Now define $p^{[0]} = (0,1,0)^\top$which is the probability vector of $m$'s breeding state at $k$'s birth (certain state 2), and recall $\bar{\beta}_2$ is the proportion of adult females in state 2. Then:

$$\begin{aligned}\mathbb{P}\left[K_{kl} = \text{XmHSP} | b_k, b_l, \boldsymbol{\theta}\right] &= \mathbb{P}\left[K_{km} = \text{MOP} \wedge K_{lm} = \text{MOP} | b_k, b_l, \boldsymbol{\theta}\right] \\ &= \frac{\Phi\left(b_l - b_k - 1\right)\left[\Psi^{b_l - b_k} p^{[0]}\right]_2}{N_{b_l,\text{A}} \bar{\beta}_2} \end{aligned} \tag{8}$$

where $[]_2$ gives the second element of the vector, i.e. the probability that $m$ (given she was alive) was in breeding state 2 at $l$'s birth.

There is a practical complication with HSPs: we do not identify them directly, but

**GGPs— exclude big birth-gapsI will write something here...

** False-neg stuff (sigh..!)... and here

** mtDNA

### 2.1.6 Self-recaptures (SPs)

Our stage-structured model keeps the population dynamics simple, but we do have to make extra assumptions about sampling selectivity to include the IMR data. Here, we assume that selectivity only varies by stage (adults and juveniles), not by age within stage. We only consider female samples for self-recapture, since juvenile males are prone to "permanent emigration" (CITE) as well as true mortality, so do not yield readily-interpretable inferences.

To compute stage-structured self-recapture probabilities, we condition on the age of the first sample but *not* explicitly on the age of the second sample; instead we condition on the second sample's developmental stage at sampling. If the first sample would have reached the right developmental stage (otherwise, the two cannot be the same animal), then we assume it is equally likely to be *any* of the females in that developmental stage at that year (i.e., sampling is unselective within developmental stage) and thus the chance it is the same as the second sample is the reciprocal of the developmental stage abundance. We must also include survival for the intervening years. The self-recapture kinship probability between samples 1 and 2 is (where sample 1 was taken before sample 2):

$$\mathbb{P}\left[K_{12} = \mathrm{SP} | a_1, y_1, d_2, y_2, L_1 = 0, \boldsymbol{\theta}\right] = \frac{\mathbb{I}\left[d\left(a_1 + (y_2 - y_1)\right) = d_2\right] \Phi\left(y_2 - y_1, a_1\right)}{N_{y_2, d_2}}, \tag{9}$$

where $d(a)$ is the function that maps age to developmental stage, with $d(a < 6)$ = "juvenile" and $d(a \geqslant 6)$ = "adult". We also condition on the first sample being non-lethal, for obvious reasons. To obtain $N_{y_2,d_2}$, adult abundance is part of the population dynamics model, but some more work is required to deduce juvenile abundance. Assuming stable age composition, we show in Appendix G that for walrus:

$$N_{y,\mathrm{J}}, = N_{y,\mathrm{A}} \frac{\rho - \phi_{\mathrm{A}}}{\rho - \phi_{\mathrm{J}}} \left( \left( \frac{\rho}{\phi_{\mathrm{J}}} \right)^5 - 1 \right),$$

where $\rho = e^r$ is the relative annual population increase/decrease.

## 2.2 Simulations

To test the CKMR model, we developed an individual-based simulation with the life history and population dynamics of Pacific walrus. The simulation was modified from the package `fishSim` by Shane Baylis (`https://github.com/SMBaylis/fishSim`). The individual-based simulation is stochastic and operates on an annual basis. Individuals are tracked through the use of unique identifiers so that kinship pairs can be identified in simulated samples. We initialized the simulation in 1950 with a population of 250,000 animals. These individuals are considered "founders" and do not have mothers or fathers. The age and sex structure of the initial population is determined by the survival rates used in the simulation (Table 1), which were based on rates reported in Taylor et al. XXXX. Individuals that are at or beyond the age of first reproduction mate randomly and males can potentially father

more than one calf. Females reproduction follows Section 2.1.2. Females that are in state 2 of the breeding cycle give birth to a single offspring with 1:1 sex ratio. There is no systematic age-effect on female reproductive dynamics, except that they are guaranteed not-pregnant in the year immediately prior to maturity (Section 2.1.2), which slightly lowers effective fecundity for the first few years of adulthood until the Markov chain reaches equilibrium. We did not include senescence in our CKMR model, but we do include it in our simulations. Parameters in Table 1 have been adjusted to maintain a stable population ($r \approx 0$).

In sampling years, captures are simulated according to either historical or planned future sample sizes. Females are available to be sampled at any age, while only calf and juvenile males are available for sampling. For simulated captures between 2014 and 2017, we used the realized sample sizes by age or age class as the basis for simulation. For simulated captures between 2023 and 2027, we used the target number of samples per age class as the basis for simulation. After sampling, some individuals die (according to age and/or sex specific mortality rates, Table 1). If a female with a young-of-the-year calf dies, her calf also dies. Individuals automatically die if they reach the maximum age. Living individuals then have their age incremented.

The female breeding cycle is as described in Section 2.1.2. Although we assume in the simulation that all pregnancies result in live births, this rate is aliased with the nominal calf-survival probability, since only samples from age 1 onwards are considered; only the product (nominal pregnancy success rate $\times$ nominal calf survival) affects the simulated samples, not the two constituent parameters. Males and females $< 4$ (or $>$ the age of last reproduction; ALR) are exempt from this cycle. The simulation then proceeds to the following year.

All simulations had a starting population size of 250,000 and were run from 1950 to 2030. To evaluate agreement between the simulation and CKMR model, we generated 10 replicate simulated datasets with demographic parameters under a null scenario as in Table 1column D0 and simulated historical and future sampling according to realized or target sample sizes by age class, with effort per year from 2023 as in S0 in Table 2). We checked each of the simulated datasets against the CKMR model for observed versus expected numbers of kin pairs in different categories (MOPs, XmHSPs, and SPs), observed versus expected gaps between half-sibling pairs, and the log-likelihood derivatives at the true parameter values.

Table 1: Demographic parameters for simulation under four scenarios (D0, D1, D2, and D3)

| | Demographic Scenario | | | |
|---|---|---|---|---|
| Parameter | D0 NULL | D1 Stable | D2 Decreasing | D3 Increasing |
| Maximum age (AMAX) | 37 | 37 | 37 | 37 |
| Age at first reproduction for females (AFR) | 6 | 6 | 6 | 6 |
| Age of last reproduction for females (ALR) | 37 | 29 | 29 | 29 |
| Age of first reproduction for males | 15 | 15 | 15 | 15 |
| Young-of-the-year (Age 0 calf) survival | 0.7 | 0.7 | 0.66 | 0.7 |
| Juvenile survival (Ages 1 to 5) | 0.9 | 0.9 | 0.85 | 0.9 |
| Reproductive adult female survival (Ages 6 to ALR) | 0.9622 | 0.99 | 0.985 | 0.99 |
| Non-reproductive adult female survival (Ages ALR to AMAX) | NA | 0.55 | 0.5 | 0.55 |
| Probability of breeding at 2-yr interval ($\psi_2$) | 0.1 | 0.1 | 0.1 | 0.1 |
| Probability of breeding at 3-yr+ interval ($\psi_3$) | 0.5 | 0.5 | 0.5 | 0.5 |
| Resulting rate of increase ($r$) | 0 | 0 | -0.02 | +0.01 |

Table 2: Sampling scenarios

| | | Effort per Year | | | | | |
|---|---|---|---|---|---|---|---|
| Sampling Scenario | Description | 2023 | 2024 | 2025 | 2026 | 2027 | 2028 |
| S0 | NULL: 100% effort 2023-2028 | 1 | 1 | 1 | 1 | 1 | 1 |
| S1 | Reality + 100% effort 2025-2026 | 1 | 0 | 1 | 1 | 0 | 0 |
| S2 | Reality + 100% effort 2025-2027 | 1 | 0 | 1 | 1 | 1 | 0 |
| S3 | Reality + 100% effort 2025-2028 | 1 | 0 | 1 | 1 | 1 | 1 |
| S4 | Reality + 75% effort 2025-2026 | 1 | 0 | 0.75 | 0.75 | 0 | 0 |
| S5 | Reality + 75% effort through 2027 | 1 | 0 | 0.75 | 0.75 | 0.75 | 0 |
| S6 | Reality + 75% effort through 2028 | 1 | 0 | 0.75 | 0.75 | 0.75 | 0.75 |
| S7 | 100% effort 2023-2025 | 1 | 1 | 1 | 0 | 0 | 0 |

## 2.3 Survey design

We were interested in evaluating the performance of CKMR under different demographic and sampling scenarios. The demographic scenarios were a stable population (D1), a slightly decreasing population (D2) and a slightly increasing population (D3). Demographic parameters for these simulated scenarios are shown in Table 1. For these simulations, we simulated historical sampling according to realized sample sizes by age and sex and future sampling by target sample sizes by age class. We simulated scenarios with (L2) and without (L1) the collection of 100 lethal samples per year in sampling years. We also simulated various reductions in sampling effort, either by reducing the number of sampling years or by reducing the amount of sampling effort within years (S1-S7; Table 2). With three demographic scenarios, two lethality scenarios, and seven sampling scenarios, this resulted in a total of 42 simulated datasets from which to evaluate survey design.

We used the simulated samples from each simulated dataset to...not sure whether this should be explained here or after design calcs.

## 2.4 Sensitivity analyses

## 2.5 Model checking

## 2.6 Checking model correctness by simulation

Close-kin pairwise probability formulae are usually quite simple, at least with hindsight, but they still can be awkard to get right in the first place. One way to reduce the risk of mistakes is to generate simulated datasets, and check that the CKMR code is giving the expected results when known parameter values are inserted. CKMR simulation code looks utterly different from kinship-probability code, and the chance of "making the same mistake twice" is therefore much less than with many statistical simulations. Robustness is improved even further if two different people are involved, one to simulate and one to write kinship-probability code. Even though simulation is not strictly necessary for most CKMR design exercises, simulation may be worth the additional effort in order to help the whole process, and that is the approach we took for walrus. We did find and fix several mistakes this way, both in the CKMR code and in the simulation code, so the exercise was certainly worthwhile.

The obvious question is how to approach CKMR model-checking when simulated datasets are available. There are various options and no . One thing to avoid, if possible, is the naive and laborious

approach of actually *fitting* a CKMR to each simulated dataset, which can be painfully slow. (Note, perhaps for discussion: We started this project before RTMB became available, expecting that the actual model-fitting code for real data would eventually have to be written in TMB itself, but keen to avoid the complexity of TMB at the design stage. In contrast, design calculations are quick because it is only necessary to calculate probability arrays once, and R alone is adequately fast, without TMB or RTMB. However, it would not be practical to fit even our simple model to multiple datasets without RTMB; and even with RTMB, repeated fitting of a more complicted model, e.g. with copious random effects, might be a challenge.) We used several checks. All are aimed at detecting gross errors (and we did find some); power to detect subtle mistakes is lower, but in our experience subtle mistakes are actually less likely than big ones. The first two checks are based on single realizations of simulated data, and so are also suitable as diagnostics when fitting to real data; the last two require multiple simulated datasets.

- Observed and expected totals of sampled kin-pairs of each type. Clearly, unless these match reasonably well, there must be a major inconsistency between model and simulationg. The definition of "reasonably well" can be guided by the inherent Poisson variability. If an expected total is 227, say, then we would not expect to see observed total much outside, say, the 95% confidence limits for a Poisson distribution with mean (and therefore variance) 227. This can be roughly approximated by $227 \pm 2\sqrt{227}$ or about [195,255]. Clearly, the expected total needs to be fairly large for this to have much power, so it might be useful to increase the simulated sample size for checking purposes.

  **OPTION**list the totals here (for first test dataset, chosen so that sim matches CK code as closely as possible)

- Breakdown of observed and expected kin-pair totals across some covariate of interest. If the totals from the previous step are not matching well, then the breakdown may shed light on where to look for problems. For example: the distribution of birth-gaps between XmHSPs is driven in the longer term by the adult rate mortality rate, so if observed and expected do not correspond, then the treatment of mortality is likely inconsistent. Also, the number of mothers by age-at-birth should fluctuate over the first few years of adulthood because of the typically-three-year breeding cycle (most 6yo have just given birth; most 7yo are still nursing last year's offspring, etc), until it settles down because of the averaging effects of irregularities. If the observed and expected patterns do not match, then the breeding cycle treatment is inconsistent.

**OPTION** show the 2 graphs here.

- P-values of observed kin-totals by type, based on the Poisson distribution as above. Given a reasonable number of simulated datasets (say 20 or more), these should be roughly uniform across the interval [0,1]. Clearly, it would require a large number of simulations to get a precise check here, but precision is not necessary: the goal is to pick up fairly coarse errors.
  **OPTION** show 4 histos here (instead of box'n'whiska)

- Looking at the mean and variance of the derivative of the pseudo-log-likelihood at the true parameter values $\theta_0$ (something which can be calculated fairly quickly by numerical differentiation). The mean should be close to 0 and the variance determines what "close" might mean, given the number of simulations available. This checks the crucial "unbiased estimating equation" (UEE) assumption required by most statistical estimation frameworks, including maximum-likelihood. If UEE does not hold, then by definition there is a mismatch between simulation and model.
  **OPTION** there's some numbers printeed at the end of compare2sims.R, I thnk.

The description so far implicitly assumes that the CKMR model (if working right) corresponds exactly to the data-generation mechanism in the simulations. However, it might be desirable to make the CKMR model simpler, especially for design purposes where the goal is just to make sure that sampling plans are sensible; developing a more complicated and realistic model can often be left until the real data appears. For example, we wanted to avoid reproductive senescence in the CKMR equations, so that all adults could be treated as a single block without requiring age-structured dynamics inside the model. Nevertheless, senescence is likely a reality of the walrus world, and there is such a thing as "too simple to be useful", so it is worth checking whether the simpler formulation is going to run into serious trouble. Simulated datasets can be used to estimate approximate bias in a slightly-mis-specified CKMR model, again without needing to do any estimation. The idea is to approximate the MLE for each dataset, based only on calculations using the true parameter value for the simulations. The MLE $\hat{\theta}$ will by definition satisfy the equation $d\Lambda/d\theta|_{\hat{\theta}} = 0$, and we can take a first-order Taylor expansion around the true value $\theta_0$ to give

$$0 = \left.\frac{d\Lambda}{d\theta}\right|_{\hat{\theta}} \approx \left.\frac{d\Lambda}{d\theta}\right|_{\theta_0} + \left(\hat{\theta} - \theta_0\right) \left.\frac{d^2\Lambda}{d\theta^2}\right|_{\theta_0}$$

$$\implies \hat{\theta} - \theta_0 \approx -\left[\left.\frac{d\Lambda^2}{d\theta^2}\right|_{\theta_0}\right]^{-1} \left.\frac{d\Lambda}{d\theta}\right|_{\theta_0} \tag{10}$$

The square-bracketed term can be replaced (to the same order of accuracy as the rest of the approxmation) by the *expected* Hessian which is the crux of our design calculations anyway, and which of course does not vary from one simulation to the next. Thus, the only quantity that has to be calculated per simulated dataset is $d\Lambda/d\theta|_{\theta_0}$, already required for the unbiased-estimating-equation check above. The estimated bias is the average across simulations of (10). This is quite similar to the UEE check above, but with a change in focus: this time, we may be prepared to tolerate some small violation of UEE, provided that it does not imply substantial bias on the parameter scale. In particular, if the estimated bias for the $r^{\text{th}}$ parameter (i.e. $r^{\text{th}}$ component of $\theta$) is below its sampling variability— say, if bias is less than 1 standard deviation, computed from the square-root of the diagonal of the inverse Hessian or $\sqrt{H^{-1}(r,r)}$— then there is little reason to worry about bias for that particular parameter.

**OPTION** stuff from the end of compare2sims.R

**DISCUSSION?**

In the end, based on the checks above, our estimation and simulation codes did indeed appear consistent, and any bias induced by (among other minor things) ignoring senescence did not seem problematic. Of course, we only reached that position *after* going thru the checking process several times, to find and fix inconsistencies.

## 2.7 Design Calculations for CKMR + IMR

Our goal is to evaluate potential survey designs for a CKMR survey of walrus, i.e. to predict the variance of some quantity-of-interest that would be obtained if the data was collected according to that design. In many other contexts we would need to create a series of simulations for each design, estimate parameters (and the quantities of real interest, which may be functions of the raw parameters) for each simulation, then directly calculate the variance of those estimates. Happily, this cumbersome process can be circumvented almost entirely with the formulation of CKMR+IMR that we use, because of three key properties:

- it is based on a pseudo-log-likelihood that sums over pairwise comparisons between samples;
- the outcome of each comparison is discrete yes-or-no: do these two samples the target kinship, or not?
- individual samples have multiple covariates (e.g. year of sampling, age), but the range of possible values for each covariate is limited.

Compared to general mark-recapture frameworks, this leads to some remarkably simple formulae for computations. In particular, because the pairwise comparisons are almost mutually independent (see below), the parameter covariance matrix can be predicted, without any simulated data, just from two types of information:

- the number of pairwise comparisons with each particular combination of covariates, basically the product of the numbers of relevant samples (which constitute "the design");
- a fairly-easy-to-compute function of the kinship probabilities, which determines how much statistical information about the parameters will be obtained, on average, from a single pairwise comparison of samples with those covariates.

We present a slightly simplified explanation here, dealing first with the latter. Let $y_{ijk}$ be the kinship outcome for samples $i$ and $j$ and target kinship $k$: $y_{ijk} = 1$ if their actual kinship $K_{ij} = k$, or 0 if $K_{ij} \neq k$; and let $y$ be the set of all data (all kinship outcomes). Also define $p_{ijk}(\theta) = \mathbb{P}[K_{ij} = k | z_i, z_j, \theta]$ to be the kinship probability for parameter values $\theta$, computed from a formula such as (6); we just write $p_{ijk}$ where there is no ambiguity about $\theta$. Each comparison has a very low probability of "success" ($y_{ijk} = 1$), on the order of the reciprocal of adult abundance, and is well approximated by a Poisson distribution with mean $p_{ijk}$. The pseudo-log-likelihood $\Lambda$ is thus[1]:

$$\Lambda(\theta; y) = \text{const} + \sum_{i<j;k\in\mathcal{K}} \{-p_{ijk} + y_{ijk} \log p_{ijk}\} \tag{11}$$

With a real dataset $y$, we would estimate $\hat{\theta}(y)$ by maximizing (11), and then infer the variance of $\hat{\theta}(y)$ from the inverse of the Hessian (second derivative matrix) of (11)[2]. For design purposes, we

[1] In practice, certain types of comparison are usually excluded *a priori* based purely on their covariate values and the target kinship (e.g. second-order kin born a long time apart), but this does not alter the basic argument.

[2] Assuming sparse sampling...

work instead with $H(\theta_0)$, the expected value over datasets $Y$ of $d^2\Lambda(\theta_0; Y)/d\theta^2$ at the true parameter value $\theta_0$. Since $\Lambda$ itself consists of a sum of terms over individual comparisons, the same is true of the second derivative and its expectation, which we can write say as $H(\theta_0) = \sum_{i<j;k\in\mathcal{K}} h_{ijk}(\theta_0)$. Some algebra (ref**) shows that the single-comparison expected Hessian is given by

$$h_{ijk}(\theta_0) = 4 d_{ijk}(\theta_0)\, d_{ijk}(\theta_0)^\top \qquad (12)$$
$$\text{where } d_{ijk}(\theta) = \frac{d\sqrt{p_{ijk}}}{d\theta}$$

The term $d_{ijk}(\theta_0)$ is the derivative vector of the square-root of the kinship-$k$ probability for covariate values $z_i$ and $z_j$ with respect to the elements of the parameter vector $\theta$, at the value $\theta = \theta_0$. This can be obtained efficiently for all $(i, j, k)$ by numerical differentation of the probabilities calculated by the CKMR model, using some reasonable guess about $\theta_0$; the whole process takes just a minute or two for our walrus example.

The remaining requirement for design, is to group similar comparisons, i.e. across all pairs with identical covariate values. Let $z_i$ denote all the covariate values for sample $i$ that are needed to compute $p_{ijk}$ (note that this may vary for different $k$, given different roles that the sample may be playing: for brevity we omit the $k$ here). Now let $m(z)$ denote the number of samples with covariate values $z$. The number of comparisons between samples that have covariates $z_1$ and samples that have covariates $z_2$ is $m(z_1)\, m(z_2)$. The grouped version of the expected Hessian can be written as

$$H(m_{\mathcal{Z}}; \theta_0) = \sum_{z_i < z_2 \in \mathcal{Z}; k\in\mathcal{K}} m(z_1)\, m(z_2)\, h(z_1, z_2, k) \qquad (13)$$

where $h(z_1, z_2, k)$ is the single-comparison expected Hessian for two samples with covariates $z_1$ and $z_2$ respectively[3]. The set $\mathcal{Z}$ comprises all possible combinations of covariates, and $m_{\mathcal{Z}}$ is the corresponding breakdown of total sample size by covariate combinations (e.g. year and age and sex). Once $H$ has been computed, it can be inverted to give the average predicted variance $V(m_{\mathcal{Z}}; \theta_0)$ of a parameter estimate. CVs or standard errors of any quantity-of-interest $g(\theta)$ that can be obtained

[3]The ordering "$z_1 < z_2$" is arbitrary, included just to avoid double-counting. Sometimes it makes sense to also do comparisons with $z_1 = z_2$, in which case an extra factor of $1/2$ is required.

from $\theta$, can then be approximated by the Delta method:

$$\mathbb{V}\left[g\left(\theta\right); m_{\mathcal{Z}}, \theta_0\right] \approx \left[\left.\frac{dg}{d\theta}\right|_{\theta_0}\right] V\left(m_{\mathcal{Z}}, \theta_0\right) \left[\left.\frac{dg}{d\theta}\right|_{\theta_0}\right]^{\top} \tag{14}$$

While a "design" must, by definition, include some specification of sample sizes, it may not specify the full breakdown of samples into specific $z$-categories. For example, the plan might be to sample 1000 adult walruses per year, but the age composition cannot be controlled directly. However, we still need to know that detailed breakdown $m_{\mathcal{Z}}$ in order to apply the above steps, so some extra extra assumptions and calculations might be required. For example, our population-dynamics model does not explicitly represent the adult age composition within the population, let alone within the samples; probability formulae such as (7) are *conditioned* on sample age, but make no prediction about how many samples of each age there will be. It would be possible to calculate expected sample sizes based on quasi-stable age compositions and unselective-sampling assumptions (assumptions that are in fact implicit for the self-recapture probability (9)), but somewhat laborious. Instead, since we are simulating sampled datasets in any case, the simulated sample composition can be used directly for $m_{\mathcal{Z}}$.

The use of the pseudo-log-likelihood Hessian to approximate the inverse variance is not strictly justified in a mathematical sense, because the pairwise comparisons are not fully mutually independent. The proposed walrus sample size (about 15,000 in total) is so large relative to adult abundance (about 70,000 females, although in effect somewhat more because of turnover during the years modeled) that roughly 10% of samples are recaptured multiple times, as self and/or as kin. This means that a comparable proportion of pairwise comparisons have predictable outcomes based on the results of other comparisons, which breaks independence. Thus the "sparse sampling" assumption of Bravington et al. (2016), which underlies the use of the Hessian, is not strictly justified; this does not lead to bias in point estimates, but the Hessian-based approximation is likely to underestimate the true variance somewhat. Accordingly, we have made some simple adjustments to "effective sample size" based on summaries of the simulated datasets, as explained in the Appendix. This should be quite adequate for design purposes— since, in any case, all our variance estimates have to be based on uncertain assumptions about true parameter values— but a more detailed treatment may be worthwhile when it comes to analysing the real data.

## 2.8 Something for the Appendix, perhaps: Adjustments for non-sparse sampling

From experience, any attempt at a comprehensive treatment of non-independence in CKMR is complicated, to say the least. In this paper, we restrict attention to some obvious aspects for walrus that are easy to address. We consider the comparisons in stages: first SelfPs, then MOPs, then XmHSPs. We adjust set the effective sample size for each stage based on recaptures from the preceding stages in one simulated dataset, as follows:

- Sample sizes are initially taken from the simulated dataset (thus allowing detailed breakdown of sample size by age, year, etc). All available samples are used for SelfP comparisons.
- If an individual is self-recaptured, only its final capture will be used in MOP and XmHSP comparisons (i.e. duly adjusting the sample sizes sample sizes for MOPs and XmHSPs, as well as the number of MOPs etc found if that individual is involved).
- Any Offspring $o$ identified in a MOP, will be excluded from XmHSP comparisons (since $o$'s sibship with any other sample $i$ can be deduced from the MOP results, based on whether $i$ is also an offspring of $o$'s Mother).

This deals with the implications of one type of kinship for the others, but does not deal with multiple recaptures within a kinship class (e.g. an individual who is sampled 3 times; given that sample A matches sample B, and B matches C, it is redundant to compare A with C). There are simple ways to handle that with real datasets, as long as age is known fairly accurately, . Taking self-recapture as an example, we can compare the first capture of an individual to other samples from successively later dates, stopping immediately after a second capture (i.e. a self-recapture) if there is one. Thereafter, the first sample is not used in self-comparisons against later samples, but the second sample (the recapture) should still be used until and unless it too is recaptured, and so on. For XmHSPs, where we are really recapturing the shared Mother rather than the samples themselves, we can proceed in an analogous fashion, dropping the firstborn halfsib from subsequent comparisons once its first halfsib has been found, and so on.

It is difficult to follow that approach in a design context, because the number of *possible* triads is extremely large, even though in practice only a limited number will be seen in any real dataset. Instead, we make simple overall adjustments, as follows:

- For SelfPs, we start by tabulating the number of triple captures, etc. If we were using the above scheme for an individual caught 3 times, we would want to end up counting only 2 recaptures, rather than the 3 that arise from all pairwise matches. Generally, for an individual caught $N$ times, we would only count $N-1$ recaptures rather than $N\left(N-1\right)/2$ from all pairings of its samples. Let $\alpha$ be the proportional reduction in the number of SelfP pairs that would result. The number of comparisons should be reduced accordingly, and we apply the same reduction $\alpha$ across-the-board to *all* SelfP comparison categories. (About 2% of walrus self-recaptures in different years are expected to be 3rd or more captures.)

- An exactly analogous approach works for XmHSPs (ignoring yet more complications from false-negatives). (About 6% of walrus XmHSPs are expected to be part of triads.)

This should give a reasonably unbiased adjustment to ensure sample sizes in the design are comparable to the adjusted sample sizes we would get with real data, although it is not quite right because some types of comparison (i.e. pairs of covariates) are liable to be more/less susceptible to multiple recaptures than others. That nuance really does not seem important for design purposes. A more practical problem is that the adjustments are dependent on results from a single realization of sampling, and in particular on the number of triple-captures, which is fairly small and thus subject to some variability. Thus, as usual in statistics, bias correction entails some increase in uncertainty.

It is not necessary to adjust MOP counts in the same way, because the outcomes of comparing one potential-offspring to all adults for MOPship are almost independent[4] of the results for any other potential-offspring; mothers are allowed to have multiple offspring.

All the phenomena above would, if uncorrected, tend to lead to underestimates of variance. However, there is another related phenomenon in non-sparse-sampling which tends to have the *opposite* effect, basically of finite-population-correction. Consider a simplified setting where all adults have the same covariates and only one juvenile, $O$, is sampled. In the MOP comparisons between $O$ and all $m_{\mathrm{ad}}$ adult samples, at most 1 of those adult samples can be $O$'s Mother, whereas the Poisson approximation in theory allows for 2 or more, albeit with low probability. It can be shown that the Fisher information (for a model where the only parameter is $N_{\mathrm{ad}}$, the number of adult females) based on the Poisson approximation is lower than the true Fisher information based on the yes-no outcome to all $m_{\mathrm{ad}}$ com-

[4]Or almost completely independent: if the two potential-offspring are born in the same or successive years, then walrus biology precludes them from both being offspring of the same mother. Another complication arises with non-lethal adult sampling; some MOP pairs will have offspring born after non-lethal sampling of the mother, in which case the mother's known lifespan can be allowed for in comparisons with other samples. We ignore all this.

parisons at once, by a factor $(N_{\text{ad}} - m_{\text{ad}})/N_{\text{ad}}$ in this simple case; that is, the pseudo-log-likelihood Hessian here leads to an *overestimate* of true variance. While the same qualitative effect presumably extends to more complex settings with different covariates, it is by no means obvious how to extend the calculations, and we propose generally ignoring this overestimation of variance. On the whole, it is usually a worse mistake to be over-confident than to be under-confident in an estimate; model-based variance estimates tend to be biased low anyway, through ignoring structural oversimplifications in the model; and with CKMR variance can usually be reduced anyway by modest increases in sample size. For walrus in particular, the total proposed adult sample size is under 10% of the population, so the effect on standard deviations would presumably not exceed 5%.

### 2.8.1 Some code I don't want to delete yet...

```
I ran these snippets inside add_data(), using my debugger

table( table( c( MOPs, XmHSPs, selfPs))
# 1 2 3
# 1524 120 4
length( unique( samples$Me))
# 14385
length( unique( samples$Me)) - length( unique( c( MOPs, XmHSPs, selfPs)))
# 13187 this should go with "0" in the first line

# So, quite a few: 10% of samples are in pairs, 10% of those are in triplets!
# Where are most of these triplets coming from?
table( table( c( XmHSPs)))
# 1   2
# 444 27
# ... so it's mostly not from the XmHSPs

table( metab) # SelfPs before trimming the xtuples
#    1    2    3
# 9286  212    3
```

```
# ... not them either

table( table( MOPs[,1]))
#   1   2
# 419  13
# ... so it's mostly not from multiple offspring of one mother...

table( table( MOPs[,2]))
# 1
# 445
# ... good; no-one has >1 mother!

# Thus, the bulk of the 124 must be "interference" across kinships
# which is probably worth fixing, using the first lot of simple steps above.
```

# 3 Results

obvs the checks passed (eventually...). Not much more to say! (I suspect it's not worth reporting the checks— where do you stop with that? but perhaps they should be available online, eg obs & exp kin-totals? maybe report the deriv checks for bias, dunno)

Present some scenarios and results together (do NOT put "scenarios" into separate earlier section, it just confuses the hell out of readers).

TO ADD: Figure of # samples versus # kin pairs observed for each type (MOPS, XmHSPs, Self)

## 3.1 Sample sizes and duration

## 3.2 Sensitivity analyses

- pop-dyn stable/inc/dec (presumably, minimal diff from a Design PoV, ITO how much CKIMR adds rel to IMR— which is the key Design Q)
- turning off CKMR (or IMR)
- value of Lethal

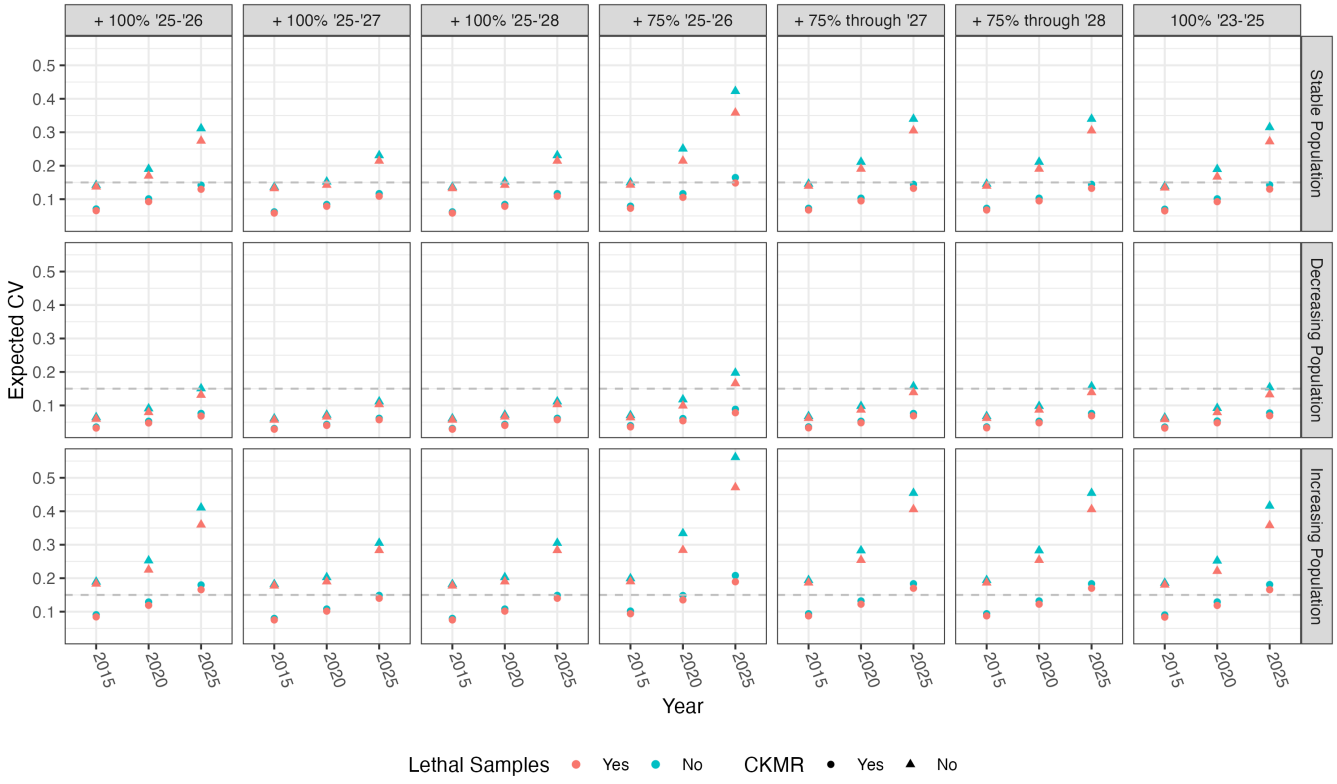


Figure 2: Expected CV of adult female abundance (vertical axis) in different years (horizontal axis) under different demographic (panel rows) and sampling (panel column) scenarios. Triangular points represent expected CVs from IMR alone, while circular points show expected CVs with ICKMR. The inclusion of lethal samples is indicated by red (lethal samples included) or blue (no lethal samples included in sampling years) points. The horizontal dashed line at CV = 0.15 represents an arbitrary threshold for decision making.

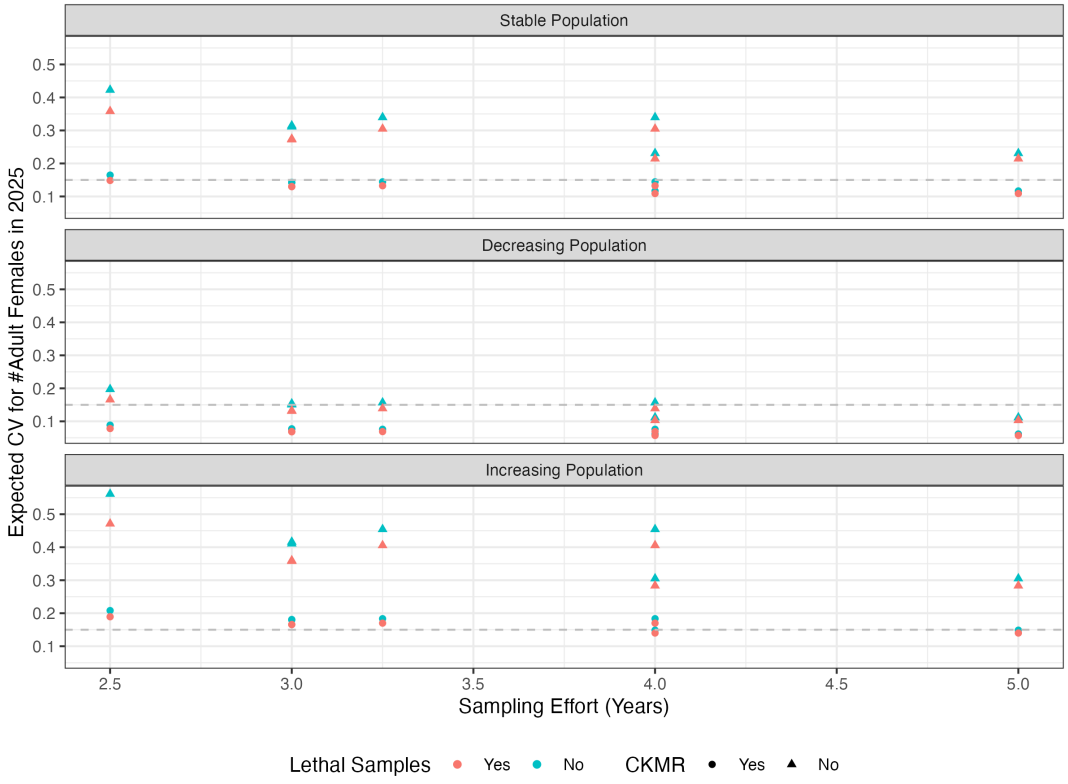


Figure 3: Sampling effort (in number of years, horizontal axis) versus expected CV for adult female abundance in 2025 with IMR alone (triangular points) or with ICKMR (round points) and with (blue points) and without (red points) the inclusion of lethal samples in sampling years. The three panels represent demographic scenarios of a stable population, decreasing population, and increasing population, respectively. The horizontal dashed line at CV = 0.15 represents an arbitrary threshold for decision making.

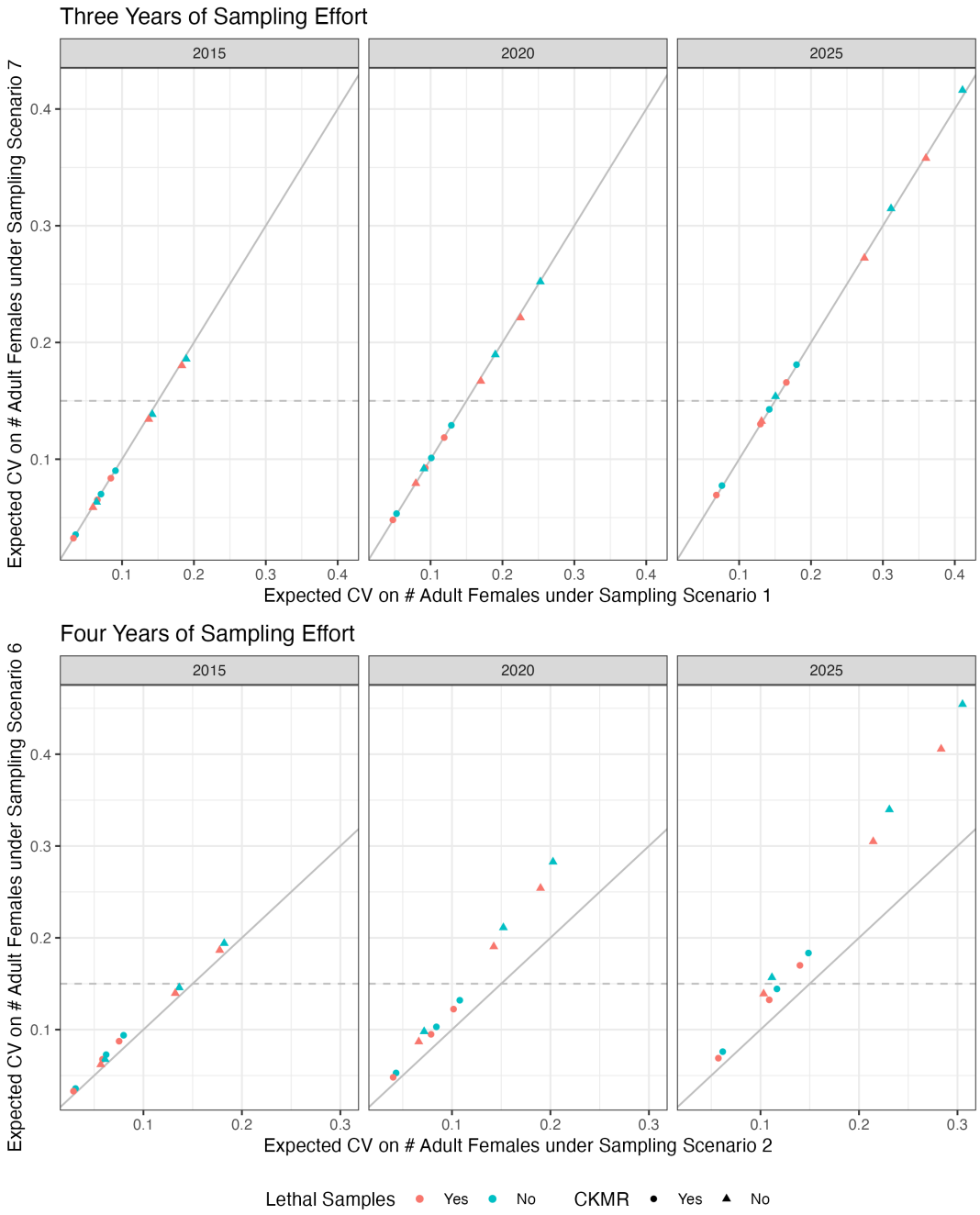


Figure 4: Expected CV on the number of adult females with three years of sampling (top panels) and four years of sampling (bottom panel). In the top panel, the horizontal axis shows expected CVs under sampling scenario 1 and the vertical axis shows expected CVs under sampling scenario 7. In the bottom panels, the horizontal axis shows expected CVs under sampling scenario 2 and the vertical axis shows expected CVs under sampling scenario 7. From left to right, panels indicate expected CVs in 2015, 2020, and 2025. Individual points represent expected CVs under different possible demographic scenarios, with (red) and without (blue) the inclusion of lethal samples, and with (round points) and without (triangular points) the use of CKMR (versus IMR alone). The solid grey line is 1:1. The horizontal dashed grey line represents an arbitrary threshold CV of 0.15.

# 4 Discussion

- We show how sample collection plans could be modified to achieve desired monitoring goals with less sampling effort.
- We didn't bother doing X coz IJAD[5]. For real data analysis, we might do Y instead.
- Ways to extend the model... impact of DNAge
- Future utility of lethal samples (although my guess is: there won't be enough. Glass-half-full, or glass-half-empty, if you're a walrus?)
- The full ramifications of opting for a stage-structured quasi-equilibrium model, which avoids having to model age composition but does entail an *assumption* about selectivity, are not at all obvious, but the model seems to us fairly reasonable; it might be worth revisiting when large numbers of DNAge samples become available. At that point it would be possible to compare the actual age compositions with the predicted compositions assuming partly-unselective sampling and quasi-equilibrium.
- As should be evident from the preceding text and number of authors on this paper, building a close-kin model involves a high level of collaboration between statisticians, biologists and geneticists. CKMR is very much a multidisciplinary methodology and each discipline has a great deal to input into the process of model building.
- Would be great to mention that CKMR was motivated by fisheries and is an example of a shared tool between fisheries scientists and ecologists, maybe cite Schaub et al 2024
- on stage-structured dynamics: That assumption may turn out to be unreasonable for juveniles especially; but it will only be possible to check once enough sample-age-composition data become available. However, if it does turn out to be the case that (say) 2yo are disproportionately likely to be sampled (given their estimated abundance from the fitted model), then it would not be hard to adjust the stage-structured IMR equations to incorporate sample-composition-data and (estimated) selectivity. Sample sizes in this project are large enough that selectivity (i.e., the ratio of age-specific sample compositions to model-estimated population age compositions) should be estimated with respectable precision and without "propagating" a lot of uncertainty into other

[5] It's Just A Design

parameter estimates. We therefore think that our current somewhat crude IMR sub-model should give a reasonable guide to ultimate precision, even if it gets adjusted somewhat in the cold light of real data. Note that similar assumptions appear to be made in Beatty et al. 202 (to be confirmed).

- appendix (F) discusses skip-breeding
- While a stable-age-composition between 2000–2027 is probably not valid for the entire range of adult ages—since older adults would have experienced long periods of increased mortality from hunting—it is perhaps a reasonable assumption for younger adults, and it is only younger adults that matter here because they indirectly determine the number of juveniles. A stable age composition for juveniles seems fairly reasonable, since "recruitment variability" cannot be high for an animal with a litter size of 1, and it only requires a few years for the juvenile distribution to settle down.

# References


Årthun, Marius et al. (2021). "The Seasonal and Regional Transition to an Ice-Free Arctic". en. In: *Geophysical Research Letters* 48.1. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL090825, e2020GL090825. ISSN: 1944-8007. DOI: 10.1029/2020GL090825. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2020GL090825 (visited on 14/12/2024).

Beatty, William S et al. (July 2020). "Panmixia in a sea ice-associated marine mammal: evaluating genetic structure of the Pacific walrus (Odobenus rosmarus divergens) at multiple spatial scales". In: *Journal of Mammalogy* 101.3, pp. 755–765. ISSN: 0022-2372. DOI: 10.1093/jmammal/gyaa050. URL: https://doi.org/10.1093/jmammal/gyaa050 (visited on 14/12/2024).

Beatty, William S et al. (Sept. 2022). "Estimating Pacific walrus abundance and survival with multievent mark-recapture models". en. In: *Marine Ecology Progress Series* 697, pp. 167–182. ISSN: 0171-8630, 1616-1599. DOI: 10.3354/meps14131. URL: https://www.int-res.com/abstracts/meps/v697/p167-182/ (visited on 20/02/2023).

Bravington, Mark V, Hans J Skaug and Eric C Anderson (May 2016). "Close-Kin Mark-Recapture". In: *Statistical Science* 31.2, pp. 259–274. DOI: 10.1214/16-STS552. URL: http://projecteuclid.org/euclid.ss/1464105042.

Conn, Paul B. et al. (2020). "Robustness of close-kin mark–recapture estimators to dispersal limitation and spatially varying sampling probabilities". en. In: *Ecology and Evolution* 10.12. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ece3.6296, pp. 5558–5569. ISSN: 2045-7758. DOI: 10.1002/ece3.6296. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/ece3.6296 (visited on 03/02/2023).

Davies, C.R. et al. (2020). *Next-generation Close-kin Mark Recapture: Using SNPs to identify half-sibling pairs in Southern Bluefin Tuna and estimate abundance, mortality and selectivity*. FRDC report 2016-044. CSIRO.

Fay, Francis H. (1982). "Ecology and biology of the Pacific walrus, Odobenus rosmarus divergens Illiger". In.

Fay, Francis H., Brendan P. Kelly and John L. Sease (1989). "Managing the Exploitation of Pacific Walruses: A Tragedy of Delayed Response and Poor Communication". en. In: *Marine Mammal Science* 5.1. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1748-7692.1989.tb00210.x, pp. 1–16. ISSN: 1748-7692. DOI: 10.1111/j.1748-7692.1989.tb00210.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1748-7692.1989.tb00210.x (visited on 14/12/2024).

Garlich-Miller, Joel et al. (2011). "Status review of the Pacific walrus \textit{(Odobenus rosmarus divergens)}". en. In.

Gilbert, J. R. (1999). "Review of previous Pacific walrus surveys to develop improved survey designs". In: *Marine Mammal Survey and Assessment Methods*. Num Pages: 10. CRC Press. ISBN: 978-1-00-321116-7.

Hillary, R.M. et al. (2018). "Genetic relatedness reveals total population size of white sharks in eastern Australia and New Zealand". In: *Nature Scientific Reports* 8 (1), p. 2661. DOI: 10.1038/s41598-018-20593-w.

Jay, Chadwick V., Anthony S. Fischbach and Anatoly A. Kochnev (2012). "Walrus areas of use in the Chukchi Sea during sparse sea ice cover". en. In: *Marine Ecology Progress Series* 468, pp. 1–13. DOI: 10.3354/meps10057. URL: https://pubs.usgs.gov/publication/70040743 (visited on 14/12/2024).

Jay, Chadwick V. et al. (Mar. 2017). "Walrus haul-out and in water activity levels relative to sea ice availability in the Chukchi Sea". In: *Journal of Mammalogy* 98.2, pp. 386–396. ISSN: 0022-2372.

DOI: 10.1093/jmammal/gyw195. URL: https://doi.org/10.1093/jmammal/gyw195 (visited on 14/12/2024).

Johnson, Devin L. et al. (Oct. 2023). *Assessing the Population Consequences of Disturbance and Climate Change for the Pacific Walrus*. en. Pages: 2023.10.12.562073 Section: New Results. DOI: 10.1101/2023.10.12.562073. URL: https://www.biorxiv.org/content/10.1101/2023.10.12.562073v1 (visited on 14/12/2024).

Johnson, Devin L. et al. (2024). "Assessing the sustainability of Pacific walrus harvest in a changing environment". en. In: *The Journal of Wildlife Management* n/a.n/a. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1 e22686. ISSN: 1937-2817. DOI: 10.1002/jwmg.22686. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jwmg.22686 (visited on 14/12/2024).

Keyfitz, Nathan and Hal Caswell (2005). *Applied mathematical demography*. Springer.

Lloyd-Jones, Luke R. et al. (Aug. 2023). "Close-kin mark-recapture informs critically endangered terrestrial mammal status". en. In: *Scientific Reports* 13.1. Publisher: Nature Publishing Group, p. 12512. ISSN: 2045-2322. DOI: 10.1038/s41598-023-38639-z. URL: https://www.nature.com/articles/s41598-023-38639-z (visited on 20/04/2024).

MacCracken, James G et al. (2017). "Final Species Status Assessment for the Pacific Walrus". en. In.

Palsbøll, Per J. et al. (Aug. 1997). "Genetic tagging of humpback whales". en. In: *Nature* 388.6644. Publisher: Nature Publishing Group, pp. 767–769. ISSN: 1476-4687. DOI: 10.1038/42005. URL: https://www.nature.com/articles/42005 (visited on 14/12/2024).

Perovich, Donald K. and Jacqueline A. Richter-Menge (2009). "Loss of sea ice in the Arctic". eng. In: *Annual Review of Marine Science* 1, pp. 417–441. ISSN: 1941-1405. DOI: 10.1146/annurev.marine.010908.163805.

Silber, Gregory K. and Jeffrey D. Adams (Sept. 2019). "Vessel Operations in the Arctic, 2015–2017". English. In: *Frontiers in Marine Science* 6. Publisher: Frontiers. ISSN: 2296-7745. DOI: 10.3389/fmars.2019.00573. URL: https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2019.00573/full (visited on 14/12/2024).

Skaug, Hans Julius (Sept. 2001). "Allele-Sharing Methods for Estimation of Population Size". In: *Biometrics* 57.3, pp. 750–756. ISSN: 0006-341X. DOI: 10.1111/j.0006-341X.2001.00750.x. URL: https://doi.org/10.1111/j.0006-341X.2001.00750.x (visited on 22/12/2024).

Speckman, Suzann G. et al. (2011). "Results and evaluation of a survey to estimate Pacific walrus population size, 2006". en. In: *Marine Mammal Science* 27.3. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1748

7692.2010.00419.x, pp. 514–553. ISSN: 1748-7692. DOI: 10.1111/j.1748-7692.2010.00419.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1748-7692.2010.00419.x (visited on 14/12/2024).

Stroeve, Julienne C. and Dirk Notz (Sept. 2018). "Changing state of Arctic sea ice across all seasons". en. In: *Environmental Research Letters* 13.10. Publisher: IOP Publishing, p. 103001. ISSN: 1748-9326. DOI: 10.1088/1748-9326/aade56. URL: https://dx.doi.org/10.1088/1748-9326/aade56 (visited on 14/12/2024).

Stroeve, Julienne C. et al. (2012). "Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations". en. In: *Geophysical Research Letters* 39.16. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2012GL052 ISSN: 1944-8007. DOI: 10.1029/2012GL052676. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2012GL052676 (visited on 14/12/2024).

Taras, Brian D. et al. (2024). "Estimating Demographic Parameters for Bearded Seals, Erignathus barbatus, in Alaska Using Close-Kin Mark-Recapture Methods". en. In: *Evolutionary Applications* 17.11. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/eva.70035, e70035. ISSN: 1752-4571. DOI: 10.1111/eva.70035. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/eva.70035 (visited on 19/12/2024).

Taylor, Rebecca L. and Mark S. Udevitz (Jan. 2015). "Demography of the Pacific walrus ( *Odobenus rosmarus divergens* ): 1974-2006". en. In: *Marine Mammal Science* 31.1, pp. 231–254. ISSN: 08240469. DOI: 10.1111/mms.12156. URL: https://onlinelibrary.wiley.com/doi/10.1111/mms.12156 (visited on 20/02/2023).

Taylor, Rebecca L. et al. (Jan. 2018). "Demography of the Pacific walrus ( *Odobenus rosmarus divergens* ) in a changing Arctic". en. In: *Marine Mammal Science* 34.1, pp. 54–86. ISSN: 08240469. DOI: 10.1111/mms.12434. URL: https://onlinelibrary.wiley.com/doi/10.1111/mms.12434 (visited on 20/02/2023).

Udevitz, Mark S. et al. (2017). "Forecasting consequences of changing sea ice availability for Pacific walruses". en. In: *Ecosphere* 8.11. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ecs2.2014, e02014. ISSN: 2150-8925. DOI: 10.1002/ecs2.2014. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/ecs2.2014 (visited on 14/12/2024).

Wade, Paul R. and Douglas P. DeMaster (1999). "Determining the optimum interval for abundance surveys". In: *Marine Mammal Survey and Assessment Methods*. Num Pages: 14. CRC Press. ISBN: 978-1-00-321116-7.

Williams, Byron K., James D. Nichols and Michael J. Conroy (May 2002). *Analysis and Management of Animal Populations*. en. Google-Books-ID: 8emGAwAAQBAJ. Academic Press. ISBN: 978-0-08-057472-1.

# Appendix

## E Derivation of self-recapture "the other way round"

As discussed in Section 2.1.6, (9) can also be formulated "the other way round", i.e., considering whether the second sample is the same as the first. The answer turns out the same, but the derivation is slightly different and *appears* to involve an explicit survival term. Again, suppose two female samples ($y_1, a_1$ and $y_2, a_2$ , where $y_1 < y_2$), then

$$\begin{aligned} &\mathbb{P}\left[K_{21} = \mathrm{SP} | y_1, a_1, y_2, a_2\right] \\ &= \frac{\mathbb{P}\left[\text{Sample 1 survived until Sample 2 was taken}\right] \mathbb{I}\left(y_2 - a_2 = y_1 - a_1\right)}{N\left(y_2, a_2\right)} \\ &= \frac{\Phi\left(y_2 - y_1, a_1\right) \mathbb{I}\left(y_2 - a_2 = y_1 - a_1\right)}{N\left(y_2, a_2\right)}. \end{aligned} \tag{E.1}$$

However, the results are readily seen to be identical because, by definition of "survival", we have

$$N\left(y + t, a + t\right) \equiv N\left(y, a\right) \Phi\left(t, a\right). \tag{E.2}$$

# F Skip breeding

Walrus life-history (or our assumptions about it) has actually made the XmHSP probability rather simpler than it can be in other cases. In particular, all adult female walruses are assumed to have equal survival probabilities, regardless of age and breeding phase; and the breeding cycle is Markovian, so that once a walrus has reached calving-phase (like $m$ at $k$'s birth) her age does not affect her subsequent breeding phases nor her survival. In a "fish-like" situation, $m$'s age (and size) at $k$'s birth would also be important, because it affects both her subsequent future survival probabilities, and her fecundity if she reaches $l$'s birthdate. Thus for "fish" it is necessary to sum across all the possible ages that $m$ might have had, weighted by the probability that she was that age given that she was $k$'s mother.

The importance or otherwise of allowing for skip-breeding in HSP probabilities is not immediately obvious. Clearly, if $m$ breeds only every 3rd year, then her lifetime-average reproductive output is only $1/3$ of a hypothetical breed-every-year female. However, if everybody breeds only every 3rd year, then $m$ is only competing against $1/3$ of the population every time, so those factors broadly cancel out; is lifetime-average reproductive output largely unaffected? To a very coarse approximation the answer is yes; but a more detailed calculation using geometric progressions shows that there is in fact an appreciable bias which depends on the skip-interval and the survival rate (see subsection **??** below). For walrus, the "naive" no-skip-breeding HSP bias might be of the order of 10%, based on $\sim 95\%$ annual adult survival and an average 4-year gap between successful breedings, which is enough to justify avoiding it by the more elaborate calculation in Listing **??**.

This subsection is a mathematical treatment that does not apply directly to walrus, because we have already allowed for the female walrus' breeding cycle in our CKMR probabilities. Its purpose is to explain why failure to allow for semi-regular skip-breeding would otherwise lead to some bias in HSP-derived estimates of $N$ (adult abundance, loosely defined). This is not obvious, so we provide a derivation that provides an indication of the size of bias.

The idea of skip-breeding is that there is a regular *minimum* breeding interval. For simplicity, we just consider the case where breeding can only occur every 2nd (or, later on, every $k^{\text{th}}$) year. To "zeroth order", it might be expected that this makes no difference to the expected number of HSPs overall; although the mother of sample #1 only gets to breed every second year subsequently, whenever she does she only has to compete against half the females in the population, so the two effects cancel out. This is reasonable, but neglects the possibility of the mother dying in-between. Before equationizing all this, it is worth noting that skip-breeding has minimal effect on the estimation of *survival* from

HSPs; the probability of finding an HSP-pair diminishes by a factor of $s^k$ every $k$ years whether there are skips or not.

Consider the expected number of (maternal, say) later-born HSPs of any given sample Sally, based on one comparison per subsequent cohort. Each such comparison has probability $1/N$ of having the same mother as Sally, *iff* that mother is still alive. For a species without skip-spawning (i.e. with a typical breeding interval of just 1 year), this expected number $H_1$ is therefore given by

$$H_1 \triangleq \frac{s}{N} + \frac{s^2}{N} + \frac{s^3}{N} + \cdots = \frac{N^{-1}s}{1-s} \tag{F.1}$$

because there is only probability $s$ that the mother will survive from one year to the next. (Note that $\phi$ is sometimes used instead of $s$ elsewhere in this document.)

For a similar situation in a species with a 2-year gap, the equation is

$$H_2 \triangleq 0 + \frac{s^2}{N/2} + 0 + \frac{s^4}{N/2} + ... = \frac{2N^{-1}s^2}{1-s^2} \tag{F.2}$$

because there cannot be any maternal HSPs from odd-numbered later cohorts, but in the even-numbered later cohorts the mother (if alive) only has to compete against half the overall adults, since the remainder will be skipping that year. Thus we have

$$\frac{H_2}{H_1} = \frac{2s}{1+s} \tag{F.3}$$

and since $s < 1$, the ratio is below 1— although not much below for a long-lived species where $s \approx 1$. Misapplying a "no-skip" $H_1$-style model when there are "only" the number of HSPs from an $H_2$-situation, would lead to a positive bias in $\hat{N}$ (there are generally fewer HSPs from an $H_2$-situation than an $H_1$-situation for the same $N$).

We can generalize this to a $k$-year skip-breeding cycle:

$$\begin{aligned} H_k \triangleq 0 + 0 + \cdots + \frac{s^k}{N/k} + \cdots + \frac{s^{2k}}{N/k} + \cdots &= \frac{kN^{-1}s^k}{1-s^k} \\ \implies \frac{H_k}{H_1} &= \frac{ks^{k-1}(1-s)}{1-s^k} \end{aligned} \tag{F.4}$$

For further algebraic insight, we can define the death rate $d \triangleq 1 - s$ so that $d \ll 1$ for a long-lived species, and expand in powers of $d$:

$$\frac{H_k}{H_1} = \frac{k\left(1-d\right)^{k-1} d}{1-\left(1-d\right)^k}$$
$$= \frac{k\left(1-\left(k-1\right)d + O\left(d^2\right)\right)d}{1-\left(1-kd+\frac{1}{2}k\left(k-1\right)d^2+O\left(d^3\right)\right)}$$
$$= \frac{kd\left(1-\left(k-1\right)d+O\left(d^2\right)\right)}{kd\left(1-\frac{1}{2}(k-1)d+O\left(d^2\right)\right)}$$
$$= \left(1-\left(k-1\right)d+O\left(d^2\right)\right)\left(1+\tfrac{1}{2}\left(k-1\right)d+O\left(d^2\right)\right)$$
$$= 1-\tfrac{1}{2}\left(k-1\right)d+O\left(d^2\right) \tag{F.5}$$

Thus, for given $d$, the bias gets worse as $k$ gets larger; but note that, across taxa, we would generally expect $k$ and $d$ to be negatively correlated since there is no point in having a long skip-breeding interval if you are unlikely to survive until the other end of it.

In some situations, it might reasonably be argued that bias of this magnitude is either unimportant, or reasonably accommodated by just using a "vanilla" (skip-free) HSP model with a pre- or post-calculated adjustment based on (anticipated, or estimated) survival rate and skip-breeding interval. In other words, the "vanilla" probability could be multiplied by a fixed factor $b$ before use in the log-likelihood, in exactly the same way as the false-negative rate is accommodated. For walrus, with $d \approx 0.05$ and $k$ perhaps 3 or 4, perhaps $b = 0.9$ would do a reasonable job. While not strictly valid statistically (because uncertainty in $b$ is ignored, and because $b$ might in practice be "calculated" after-the-fact based on survival estimates, and then re-inserted in a second iteration of estimation), the offence would often be small. That said:

1. the actual bias will also depend on the temporal distribution of samples (the calculations above assume infinite future sampling, and a constant number of samples per year), so the "ideal" correction is not the same as the simplest one; and

2. it is not usually *all* that difficult to devise a bespoke probability formula that properly accommodates biological nuances and avoids "all" bias, just as we have done here for walrus.

Note that *lethal-sample* POPs are hardly affected (unless there is correlation between breeding state and sampling probability, in which case a "vanilla" POP model might not work anyway). By the time the adult is lethally sampled, the offspring must already be born, so there is no "discounting" due to potential mortality. This would usually cover most POPs in the non-lethal case, too. With non-

lethal samples, the effect on POP probabilities would depend whether sampling probability is linked to breeding-cycle status. For this report, we have generally assumed that it does not.

## Self-recapture when exact age is known

Beatty et al. (2022) used a fairly complex IMR formulation to cope with historically-very-imprecise estimates of age (or, more realistically, of "stage") estimates. However, when accurate age data are available, the pairwise comparison probabilities for self-recapture are remarkably simple. Suppose two female samples ($y_1, a_1$ and $y_2, a_2$ , where $y_1 < y_2$). Then the probability that the first one is the same as the second is just

$$\mathbb{P}\left[K_{12} = \mathrm{SP} | y_1, a_1, y_2, a_2\right] = \frac{\mathbb{I}\left(y_2 - a_2 = y_1 - a_1\right)}{N\left(y_1, a_1\right)}. \tag{F.6}$$

The indicator $\mathbb{I}\left(\cdot\right)$ is 1 if the two samples were born in the same year, or 0 if not, The samples can only be from the same animal if they were both born in the same year and if they were, we then need to know how many females of age $a_1$ were alive at $y_1$, $N(y_1, a_1)$. This implicitly assumes that all females of the same age have the same survival and sampling probabilities. (See appendix for the equivalent derivation of $\mathbb{P}\left[K_{21} = \mathrm{SP} | y_1, a_1, y_2, a_2\right]$).

In principle, given unlimited data, we could separately apply (F.6) to each combination of $(y, a)$-consistent pairs, to empirically estimate from all numbers-at-age-and-year from the reciprocal of the observed rates. Then we could apply (E.2) to estimate year-and-age-specific survivals. In practice, that would be ridiculous, since it would require an enormous number of recaptures and would lead to noisy abundance estimates, estimated survivals greater than one, and so on. However, the principle does illustrate the great power of *known-age* mark-recapture data. Note also that there are no assumptions about equiprobable sampling across ages, etc; all probabilities are simply conditioned on observed ages, and it does not particularly matter *why* there are more samples of one age than another.

The big problem with applying (F.6) in an ICKMR setting, i.e., with conditioning on age explicitly, is that it requires explicit calculation of all $N\left(y, a\right)$ within the model. This is normally unnecessary with CKMR for mammal-like species, where the main information is *only* connected with aggregate adult abundance (via TRO). It is extremely convenient to work just with a "homogenous block" of adults, and there is in any case no direct information on population age composition unless extra data are used. One option is "just" to work with a fully-age-structured population dynamics framework—but that is a lot of work to develop (from experience in fisheries work) and requires modelling extra data.

# G Derivation of juvenile abundance

The key point here is that we don't need to decompose the adult stage into separate age classes.

Following notation from the rest of the paper, let the number of adults in year $y$ be $N_{\mathrm{A},t}$ where adulthood means being aged $\alpha$ or older. The number next year will be $\rho N_{\mathrm{A},y+1}$ where $\rho = e^r$ and $r$ is the rate of increase as in (1). That will be made up of survivors from adults at $t$, plus survivors from the incoming cohort of oldest juveniles, aged $\alpha - 1$. Thus

$$N_{y+1,\mathrm{A}} = \rho N_{y,\mathrm{A}} = \phi_{\mathrm{A}} N_{y,\mathrm{A}} + \phi_{\mathrm{J}} N_{y,\alpha-1}. \tag{G.1}$$

Rearranging, we have

$$N_{y,\alpha-1} = \frac{\rho - \phi_{\mathrm{A}}}{\phi_{\mathrm{J}}} N_{y,\mathrm{A}}. \tag{G.2}$$

We now need to infer the numbers in the other juvenile age-classes (not just $\alpha - 1$). Starting with the penultimate juvenile age-class, we have:

$$\begin{aligned} N_{y,\alpha-1} &= \phi_{\mathrm{J}} N_{y-1,\alpha-2} && \text{(survival)} \\ N_{y,\alpha-1} &= \rho N_{y-1,\alpha-1} && \text{(population growth)} \\ \implies N_{y,\alpha-2} &= \frac{\rho}{\phi_{\mathrm{J}}} N_{y,\alpha-1}. \end{aligned}$$

Similar relationships apply to each preceding juvenile age class, down to age 1. The total number of juveniles in year $y$, $N_{y,\mathrm{J}}$, is given by a sum from age $x = \alpha - 1$ down to age 1:

$$\begin{aligned} N_{y,\mathrm{J}} = \sum_{x=1}^{\alpha-1} N_{y,\alpha-x} &= \sum_{x=1}^{\alpha-1} N_{y,\alpha-1} \left(\frac{\rho}{\phi_{\mathrm{J}}}\right)^{x-1} \\ &= N_{y,\alpha-1} \sum_{x'=0}^{\alpha-2} \left(\frac{\rho}{\phi_{\mathrm{J}}}\right)^{x'} \\ &= N_{y,\alpha-1} \frac{1 - (\rho/\phi_{\mathrm{J}})^{\alpha-1}}{1 - \rho/\phi_{\mathrm{J}}}, \end{aligned} \tag{G.3}$$

using the standard result for a geometric series: $\sum_{i=1}^{n} ar^i = a\frac{1-r^n}{1-r}$. Substituting for $N_{t,\alpha-1}$ from

(G.2), we have

$$\begin{aligned} N_{y,\mathrm{J}} &= N_{y,\mathrm{A}} \frac{\rho - \phi_{\mathrm{A}}}{\phi_{\mathrm{J}}} \frac{1 - \left(\frac{\rho}{\phi_{\mathrm{J}}}\right)^{\alpha-1}}{1 - \frac{\rho}{\phi_{\mathrm{J}}}} \\ &= N_{y,\mathrm{A}} \frac{\rho - \phi_{\mathrm{A}}}{\rho - \phi_{\mathrm{J}}} \left( \left(\frac{\rho}{\phi_{\mathrm{J}}}\right)^{\alpha-1} - 1 \right). \end{aligned}$$

Now, for the case of walrus, we know that $\alpha = 6$, so:

$$N_{y,\mathrm{J}} = N_{y,\mathrm{A}} \frac{\rho - \phi_{\mathrm{A}}}{\rho - \phi_{\mathrm{J}}} \left( \left(\frac{\rho}{\phi_{\mathrm{J}}}\right)^{5} - 1 \right).$$