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]

December 18, 2024

[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 both abundance and demographic parameters are key parts of the management of wildlife populations. Obtaining these estimates with minimal bias and low uncertainty is critical for conservation. Mark-recapture (CITE) provides both estimates of abundance and demographic parameters simultaneously, at the cost of having to be able to identify individuals. Marks can be natural (e.g., callosity patterns on the heads of right whales; CITE), artificial (e.g., bird ringing; CITE) or genetic (e.g., walrus crossbow shenanigans; CITE), but in all cases the method relies on the recapture of individuals to give information for estimation of survival, fecundity and abundance. Various models have been proposed over the last XXX years (CITE a bunch of Richard Cormack's papers? Or Ruth's textbook?), taking into account many common situations (e.g., open vs. closed populations; CITE Richard Barker, Matt from Otago etc). Model parameters can be decomposed so that they depend on time, space and other environmental covariates (CITE Otis et al).

If we use genetic marks to identify individuals, as in genetic individual mark recapture (IMR) we can make use of additional genetic information in these samples. DNA sequenced from an individual will tell us about its relatives as well as the individual itself. Close-kin mark-recapture (CKMR) was developed to exploit the additional information which we obtain when taking genetic samples. So, as well as confirming that an individual as been recaptured, we can also assess the probability that an individual has a given kin relationship with another (e.g., making statements like "individual a is the mother of individual b with probability x"). CKMR also has the advantage that both non-lethal and lethal (from sampling, hunting, natural mortality etc.) samples can be used. By harnessing this extra information and putting kinship into a probabilistic framework, we can gain more accurate estimates of abundance and demographic parameters. CKMR is still unfamiliar to many, but there are quite a few published applications by now, mostly on fish and sharks but some on mammals. We will not give an exhaustive list here, but see Bravington et al. (2016) for a general (albeit mathematical) framework, Conn et al. (2020) for a CKMR design exercise on bearded seals, and Lloyd-Jones et al. (2023) for a real-life application to flying-foxes (fruit bats).

The kinds of kin relationships that can be inferred from DNA samples has increased with advances in sequencing technology. Parent-offspring pairs provide the most basic level of information, but we might also want to find full- or half-sibling pairs. It's also worth noting that even if we never sample an individual, we still use information about them to estimate parameters of interest: e.g., we might never sample a father, but if we sample his children they will show up as half-siblings and so we will

know that he existed. For parent-offspring pairs, how do we know which individual is the offspring and which is the parent? If we know the age of the individuals with complete accuracy then this problem is solved (the older one has to be the parent). We can't always determine the exact age, but may be able to estimate it to a reasonable range (or at least life stage). With the addition of age information we can include further kin relationships: even though grandparent-grandchild pairs are not genetically distinguishable from thiatic (niece/nephew-uncle/aunt) relationship, with age information (and some information about the life history of the organism) we can include these relationships in our analysis. The advent of epigenetic age determination gives us a probabilistic quantification of age which is already being used in conjunction with CKMR (CITE some fish paper of Mark's).

Calculating kinship probabilities also requires in-depth consideration on the life history and genetics of the organism in question. For example, CKMR will not work on species that reproduce via parthenogenesis (as offspring are genetically identical to parents, such as aphids) or species that give birth to clones (e.g., armadillos). CKMR would also be fairly ineffective for species that observe monogamous relationships, as the information from each additional capture is minimal. Knowing about reproductive strategy is also important for both sampling and probability of kinship: does the animal in question breed in one place and at one time? Does it use a polygynous mating structure? Is it a broadcast spawner?

The monetary cost of doing CKMR is decreasing in terms of sequencing as new technology comes online. The non-lab 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. To design an effective survey, it is important to consider whether desired monitoring goals can be met given available time, money, and effort. This leads us to the question of "how many samples?". In the case of CKMR, this is relatively easy to calculate. Given a target coefficient of variation and 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 of samples. In this paper we show how to perform these calculations (and the information needed to do so) using a case study on walrus in the North Pacific.

The Pacific walrus (*Odobenus rosmarus divergens*, hereafter, 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, occupy 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 decline 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). Finally, the walrus was an endangered species candidate until the 2017 decision not to list—a decision which is currently being litigated (citation?—when will new SSA be out?).

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% 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 (USFWS) initiated a genetic individual mark-recapture (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. Biologists attempted to completely sample walruses in the accessible portion of the marginal ice zone in each year a research 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 this individual genetic mark recapture 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 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 individual genetic mark recapture without increasing sampling effort.

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; however 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 extend traditional CKMR to include IMR as an additional kinship type. We then explore different demographic and design scenarios for Pacific walrus using IMR, CKMR, and CKMR + IMR and demonstrate how the latter can be used to substantially reduce the overall amount of survey effort required for adequate monitoring. [1875 words]

2 Methods

2.1 Model Framework

2.1.1 Life-history considerations

Walrus are long-lived post maturity and give birth to one calf at a time. Walrus experience long gestation times (~14 months, CITE), therefore, females cannot give birth to calves in consecutive years, and are unlikely to give birth to calves every second year. Individuals are unlikely to re-mate with the same partner (therefore there will be few, if any, full sibling or full thiatic pairs; FSPs or FTPs). While males form leks and compete for breeding access to females, we do not expect females to vary in terms of individual fecundity. The population of Pacific walrus is well-mixed across the range, with no evidence of site fidelity within or between years, and with no genetic substructure expected.

Adult males are inaccessible to this study given season sex-segregation and the geographical coverage of sampling effort (see Section 1). Adult males are also likely to have persistent individual variability in breeding success. Therefore, we focus on female-only dynamics and do not include paternal father-offspring or half-sibling pairs (FOPs or XpHSPs) in the model. However, we do use sampled males as potential offspring whose mothers might also be captured and potential half-siblings of other individuals.

2.1.2 Stage-structured quasi-equilibrium dynamics

For the purposes of the CKMR model, we assume that female adult walrus have constant survival over age and time, and that the total number of female adults is increasing or decreasing exponentially over the period covered by the population dynamics (2000–2027; the lower limit of 2000 is set because there were fairly drastic changes in the population prior to that). We also assume that the age composition is fairly stable over that period ("quasi-equilibrium" or "quasi-stable"), so that it should follow a geometric distribution. The parameter of that distribution is determined by the simple relations

$$N_{y+1,a+1} = N_{y,a}\phi_{\mathrm{A}}$$

$$N_{y+1,a} = e^{r} N_{y,a}$$

$$\implies N_{y+1,a+1} = N_{y+1,a} \times \phi_{\mathrm{A}} e^{-r}, \tag{1}$$

where $\phi_{\mathrm{A}}$is female adult annual survival, $N_{i,j}$ is the abundance in year i for animals of age j, r is the rate of increase. From this, the adult numbers-at-age (and at year) can be inferred. Juvenile females age 1–5 have a different survival rate (ϕ_{J}), again assumed constant within (juvenile) ages and over time. We assume that the juvenile age composition is stable over time. Our formulation of IMR does assume that sampling is unselective within juveniles, and also within adults (but not between the two), and the age composition enters indirectly because of the way that juvenile abundance is calculated.

2.1.3 The breeding cycle

We use a Markov model to describe the walrus breeding cycle. We assume that female adult walrus exist in one of three breeding states: (1) pregnant, (2) with calf or, (3) neither of the above, and that next year's state depends only on this year's. From state 1 (pregnant), next year's state must be 2 (with calf). From state 2, a female may next year either return to state 1 (become pregnant again), with probability ψ_2, or move to state 3 (neither pregnant nor with calf) with probability $1 - \psi_2$. From state 3, she will next year either move to state 1 (become pregnant) with probability ψ_3, or remain in state 3 with probability $1 - \psi_3$. Note that due to long gestation times (14 months), walrus cannot give birth to calves in two consecutive years. We also allow $\psi_2 \neq \psi_3$ in case females tend to rest for a year after rearing a calf before becoming pregnant again. This is shown in Figure 1. Female survival probability 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 ψ_2 and ψ_3, this leads to a ramping-up in effective fecundity (i.e., probability of being in state 2) over the first few years of adult life. Both ψ_2 and ψ_3 are free parameters, to be 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 mHSPs.

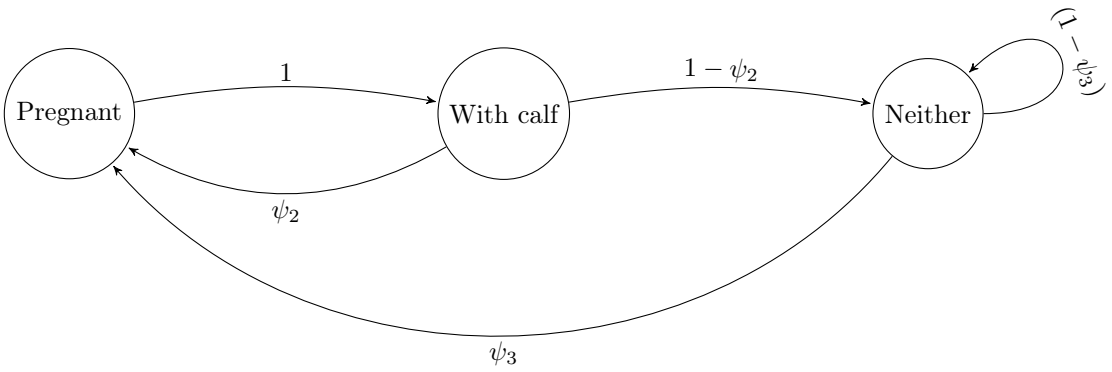

Figure 1: Directed cyclic graph showing the breeding cycle for walrus as represented in our Markov model. Nodes in the graph show the states (pregnant, with calf, or neither) and edges give the probabilities of transition between those states. Walrus reach sexual maturity at age 4, so enter the graph at node "neither".

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 specifically restrict attention to comparisons that satisfy the following criteria:

- Individual i is female (though individual j need not be); only female adult dynamics are being modelled.
- $y_j = y_i$ because there is some possibility that j and her mother are still physically associated, so within a single year their chances of being sampled are positively correlated;
- $a_j \geqslant 1$ because calves are not being used in the current model.
- $b_j \geqslant 2000$ because we are only modelling population dynamics from the year 2000 onwards.

We set the number of comparisons and the number of MOPs for excluded cases to zero. With those restrictions in mind, we can distinguish two cases: $y_i < b_j$ and $y_i \geqslant b_j$. In the former, 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). The MOP probability is

$$\mathbb{P}\left[K_{ij} = \text{MOP} | a_i, y_i, b_j, \text{nonlethal}_i\right] = \frac{i\text{'s Expected Reproductive Output in } b_j | i \text{ alive and age } a_i \text{ in } y_i}{\text{Total Reproductive Output (of females) in } b_j}. \tag{2}$$

Expected reproductive output (ERO) and total reproductive output (TRO) are in units of "number of calves", so the TRO is $\bar{\beta}_2 N(b_j)$ where $\bar{\beta}_2$ is the proportion of females with calves at any one time.

There are two components to i's ERO: first, she has to survive; second, she has to be calving in b_j. This can be written as

$$\mathrm{ERO}_i(b_j|y_i, a_i) = \Phi(y_i - b_j, a_i)\,\mathbb{P}\left[\mathrm{Br}(\mathrm{age} = a_i + (b_j - y_i)) = 2\right], \tag{3}$$

where $\Phi(\Delta t, a)$ gives the probability of survival for Δt years, starting from age a (obtained by the product of the corresponding annual juvenile and adult survival probabilities). $\mathrm{Br}(\cdot)$ is an individual's breeding state and involves individual i's age at b_j (assuming she survives).

Define fecundity as a function of age

$$F(a) \triangleq \frac{\mathbb{P}\left[\mathrm{Br}(a) = 2\right]}{\bar{\beta}_2}, \tag{4}$$

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

$$\begin{gathered}\mathbb{P}\left[K_{cj} = \mathrm{MOP}|a_i, y_i, b_j, \mathrm{nonlethal}_i, y_i < b_j\right] = \\ \frac{\Phi(y_i - b_j, a_i)\,F(a_i + (b_j - y_i))}{N(b_j)}.\end{gathered} \tag{5}$$

The second case is if i is sampled after j's birth. In that case, we 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 $F(a \leqslant 0) = 0$ (walruses cannot breed before birth), we can write

$$\begin{gathered}\mathbb{P}\left[K_{ij} = \mathrm{MOP}|a_i, y_i, b_j, \mathrm{nonlethal}_i, b_j < y_i\right] = \\ \frac{F(a_i - (y_i - b_j))}{N(b_j)}.\end{gathered} \tag{6}$$

2.1.5 Maternal half-sibling pairs (XmHSPs)

Suppose we are now about to compare individual k, born in year b_k (which can be calculated based on her age at sampling) to individual l, born in b_l, to check whether they have the same mother. This time, we impose the following criteria:

- $b_l > b_k$ to avoid double-counting. Also, $b_k \neq b_l$ since walrus give birth to a single offspring.

- $b_k \geqslant 2000$ since the population dynamics model does not go any further back.

Imagine that individual m is the mother of k. Individual l must have had *some* mother; what is the probability she was m, given what we know about m? The latter amounts to two things:

- m was alive, mature, and in breeding state 2 at k's birth;
- 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 l's mother, three things have to happen:

1. she survives until b_l;
2. she is in the right breeding state in b_l;
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 probability that a female adult survives at least Δt more years from "now"; and let Ψ be the breeding-cycle transition matrix. The probability 3-vector of an animal being in each state (1,2,3) at time t is $p^{[t]}$. Then her probability vector at time $t+1$ is $p^{[t+1]} = \Psi p^{[t]}$. Now define $p^{[0]} = (0,1,0)^\top$ which is the probability vector of m's breeding state at k's birth (i.e., certain to be in state 2, with calf), and recall that $\bar{\beta}_2$ is the proportion of adult females in breeding State 2. Then:

$$\begin{aligned}
&\mathbb{P}\left[K_{kl} = \text{XmHSP} | b_k, b_l, \theta\right] \\
&= \mathbb{P}\left[k\text{'s mum was also } l\text{'s mum} | b_k, b_l, \theta\right] \\
&= \frac{\Phi\left(b_\text{B} - b_\text{A} - 1\right)\left[\Psi^{b_\text{B} - b_\text{A}} p^{[0]}\right]_2}{N_{\text{ad}\female}\left(b_\text{B}\right)\bar{\beta}_2},
\end{aligned} \tag{7}$$

where the subscript "2" means that we need the second element of the vector inside the square-brackets, i.e. the probability that m (given she was alive) was in breeding State 2 at l's birth.

Walrus life-history (or our assumptions about it) has actually made this 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 **??**.

2.1.6 Self-recaptures (SelfPs)

As noted previously, we only consider female samples for self-recapture, since juvenile males are prone to "permanent emigration" as well as true mortality, so do not yield readily-interpretable inferences.

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

$$\mathbb{P}\left[K_{12} = \mathrm{SP} | y_1, a_1, y_2, a_2\right] = \frac{\mathbf{I}\left(y_2 - a_2 = y_1 - a_1\right)}{N_{\female}\left(y_1, a_1\right)}. \tag{8}$$

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

The big problem with applying (8) 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.

In this project we have gone for a different option, which only uses "stage-structured" dynamics (adults and juveniles). This keeps the population dynamics simple, but to handle the IMR data, we do have to make extra assumptions about sampling selectivity. We assume that 1) all (female) adults are equally likely to be sampled in a given year (i.e., that probability does not depend on age, provided the animal is adult), and 2) that all juveniles age 1–6 are also equally likely— although juveniles overall may have different sampling probabilities to adults.

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 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).

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. However, we do condition on whether the second sample was juvenile (1–6) or adult (7+) at sampling, i.e., its developmental stage. If the age of the first, plus the intervening time between samples, would place the first sample in a different developmental stage to the second sample, then the two cannot be the same animal. If the first sample would have reached the right developmental stage, 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. The other consideration, of course, is that the first sample may not have survived the

intervening years. The corresponding equation is

$$\begin{aligned}&\mathbb{P}\left[\text{Sample 2 turns out to be the same animal as sample 1}|a_1 y_1 d_2 y_2\right]\\&= \frac{\mathbf{I}\left[d\left(a_1 + (y_2 - y_1)\right) = d_2\right] \times \Phi\left(y_2 - y_1; a_1\right)}{N_{\female}\left(y_2, d_2\right)}\end{aligned} \tag{9}$$

where: $a/y/d_{\mathrm{F/S}}$ is the age, sampling year, or developmental stage of Freya or Sandy as appropriate; $d\left(a\right)$ is the function that maps age to developmental stage, with $d\left(a < 7\right) =$ "Juvenile" and $d\left(a \geqslant 7\right) =$ "Adult"; and $N_{\female}\left(y, d\right) = \sum_a \mathbf{I}\left[d\left(a\right) = d\right] N_{\female}\left(y, a\right)$. For the developmental stage "Adult , the abundance is just $N_{\mathrm{ad}\female}\left(y\right)$ which is explicitly a variable in the population dynamics but, since the latter is not formulated with explicit age structure, some more work is required to deduce $N_{\female}\left(y, \text{Juvenile}\right)$.

The key is the quasi-equilibrium assumption, i.e. that the age composition is (except perhaps for the oldest adults) stable over time, even if abundance as a whole is increasing or decreasing. The basic idea is that the number of Juveniles leaving that developmental stage should, after mortality, match the number of adults entering the Adult developmental stage in the next year; and the basic tool is the Geometric Progression. Let α be the first age of adulthood, and assume that survival continues "forever" so that we can use infinite sums (and so we can avoid several other inconveniences in the CKMR equations that would require an explicit representation of age structure). Dropping the female

subscript for brevity, and writing $\phi_a = \Phi(1, a)$ for one-year survival, we have

$$N(y+1, a+1) = \phi_a N(y, a) \ \forall\, y, a \text{ [survival]}$$

$$N(y+1, a) = \rho N(y, a) \ \forall\, y, a \text{ [increase/decrease at rate } \rho\text{]}$$

$$\implies N(y, a+1) = \phi_a N(y-1, a) = \frac{\phi_a}{\rho} N(y, a)$$

$$\implies N(y, a+t) = \left(\frac{\phi_a}{\rho}\right)^t N(y, a) \text{ if } \phi\text{is constant between } a \text{ and } a+t-1$$

$$\implies N_{\text{ad}}(y) = \sum_{a \geqslant \alpha} N(y, \alpha) \left(\frac{\phi_{\text{ad}}}{\rho}\right)^{a-\alpha} = \frac{N(y, \alpha)}{1 - \phi_{\text{ad}}/\rho} \text{ [total adult abundance relative to new incoming adults]}$$

$$\implies N(y, \alpha) = (1 - \phi_{\text{ad}}/\rho) \times N_{\text{ad}}(y)$$

$$\implies N(y-1, \alpha-1)\, \phi_{\text{ju}} = (1 - \phi_{\text{ad}}/\rho) \times N_{\text{ad}}(y) \text{ [same cohort one year earlier]}$$

$$\implies N(y, \alpha-1) = \phi_{\text{ju}}^{-1} (1 - \phi_{\text{ad}}/\rho)\, N_{\text{ad}}(y+1)$$

$$= (\rho\phi_{\text{ju}})^{-1} (1 - \phi_{\text{ad}}/\rho)\, N_{\text{ad}}(y)$$

$$\implies N(y, 1) = \left(\frac{\phi_{\text{ju}}}{\rho}\right)^{-((\alpha-1)-1)} (\rho\phi_{\text{ju}})^{-1} (1 - \phi_{\text{ad}}/\rho)\, N_{\text{ad}}(y) \text{ in terms of new post-calves}$$

$$\implies N(\text{ju}, y) = \sum_{a=1}^{\alpha-1} N(y, 1) \left(\frac{\phi_{\text{ju}}}{\rho}\right)^{a-1}$$

$$= \frac{N(y, 1)}{1 - \phi_{\text{ju}}/\rho} \left(1 - \left(\frac{\phi_{\text{ju}}}{\rho}\right)^{\alpha-1}\right) \text{ [finite geometric progression]}$$

$$\implies N(\text{ju}, y) = N_{\text{ad}}(y) \times$$

$$\left(1 - \frac{\phi_{\text{ju}}}{\rho}\right)^{-1} \left(1 - \left(\frac{\phi_{\text{ju}}}{\rho}\right)^{\alpha-1}\right) \left(\frac{\phi_{\text{ju}}}{\rho}\right)^{2-\alpha} (\rho\phi_{\text{ju}})^{-1} \left(1 - \frac{\phi_{\text{ad}}}{\rho}\right) \tag{10}$$

2.2 Simulations and model checking

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 are not exhausted; i.e., a male could potentially father more than one calf. Females reproduction follows a Markov process with three States, as described in Section 2.1.3. 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 (again as per Section 2.1.3), which slightly lowers effective fecundity for the first few years of adulthood until the Markov process has settled down. Although, there is an option to induce senescence (end of reproduction) beyond some chosen age, this was not used for the test dataset. The parameters in Table 1 have been adjusted to maintain a stable population, i.e., with RoI close to 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 ($<$ 6yr old) 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. The simulation then ages all individuals that are still alive (i.e., increases their age by 1).

The female breeding cycle is as described in 2.1.3. Females enter State 3 of the cycle (not pregnant) when they turn 4 (since age of first reproduction is 6, they may get pregnant at 5). From State 3 3, they can either move to Stage 1 (pregnant) or stay in Stage 3 according to the probability of breeding at a 3-yr+ interval. From State 1, females must move to State 2: with dependent calf. Although we nominally assume in the simulation that 100% of pregnancies result in live births, this rate is aliased with the nominal calf-survival probability, since (at least for now) only samples from 1yo onwards are considered; only the product (nominal pregnancy success rate $\times$ nominal calf survival) affects the simulated samples, not the two constituent parameters. From State 2, females can either move back to State 1: pregnant according to the probability of breeding at 2-yr interval, or move to State 3. Males and non-breeding females 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 as in Table 1column D0 and simulated historical and future

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	0.1	0.1	0.1	0.1
Probability of breeding at 3-yr+ interval	0.5	0.5	0.5	0.5
Resulting rate of increase (RoI)	0	0	-0.02	+0.01

Table 2: Sampling scenarios

		Effort per Year					
Sampling Scenario	Description	2023	2024	2025	2026	2027	2028
S0	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

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 these 10 replicate simulated datasets against the CKMR model for observed versus expected numbers of kin pairs in different categories (MOPs, XmHSPs, and SelfPs), observed versus expected gaps between half-sibling pairs, and the log-likelihood derivatives at the true parameter values.

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...

2.4 Sensitivity analyses

2.5 Design Calculations for CKMR + IMR

great wodge of theory— NB 2 MB what about HS..?

"flowchart" for what you *actually* need

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 avilable online, eg obs & exp kin-totals? mebbe report the deriv checks for bias, dunno)

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

3.1 Sample sizes and duration

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

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

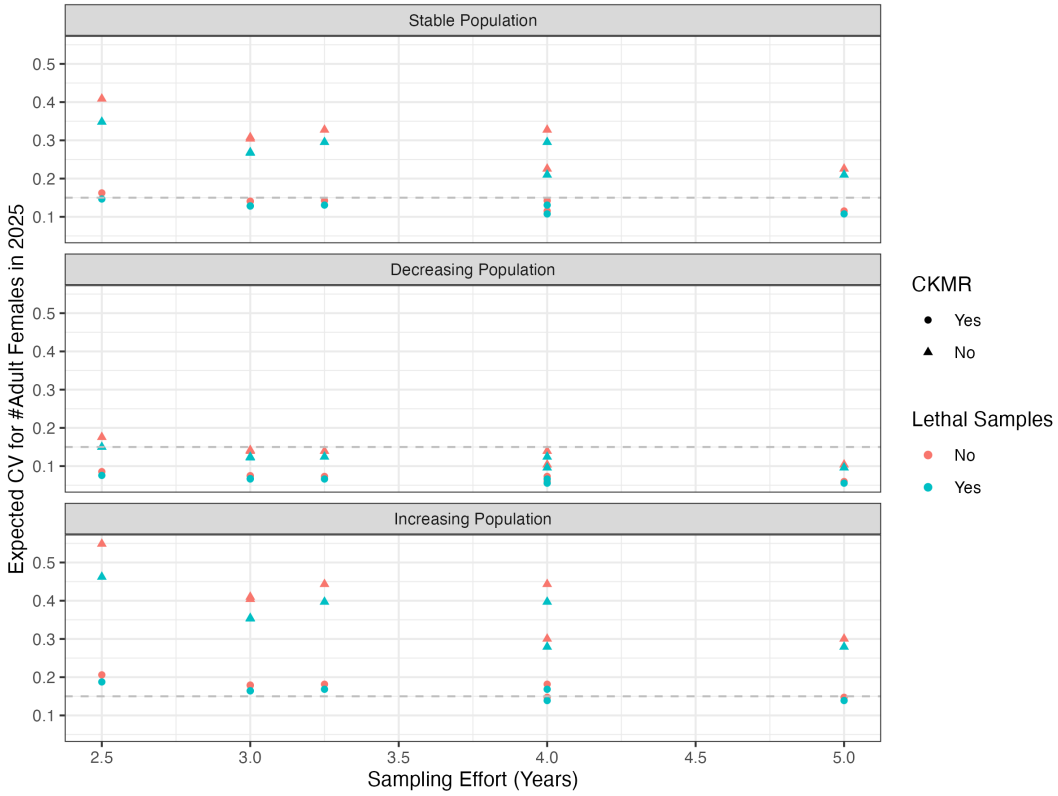

Figure 2: Sampling effort (in number of years, horizontal axis) versus expected CV for adult female abundance in 2025 with IMR alone (triangular points) or with ICKMR (round points) and with (blue points) and without (red points) the inclusion of lethal samples in sampling years. The three panels represent 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.

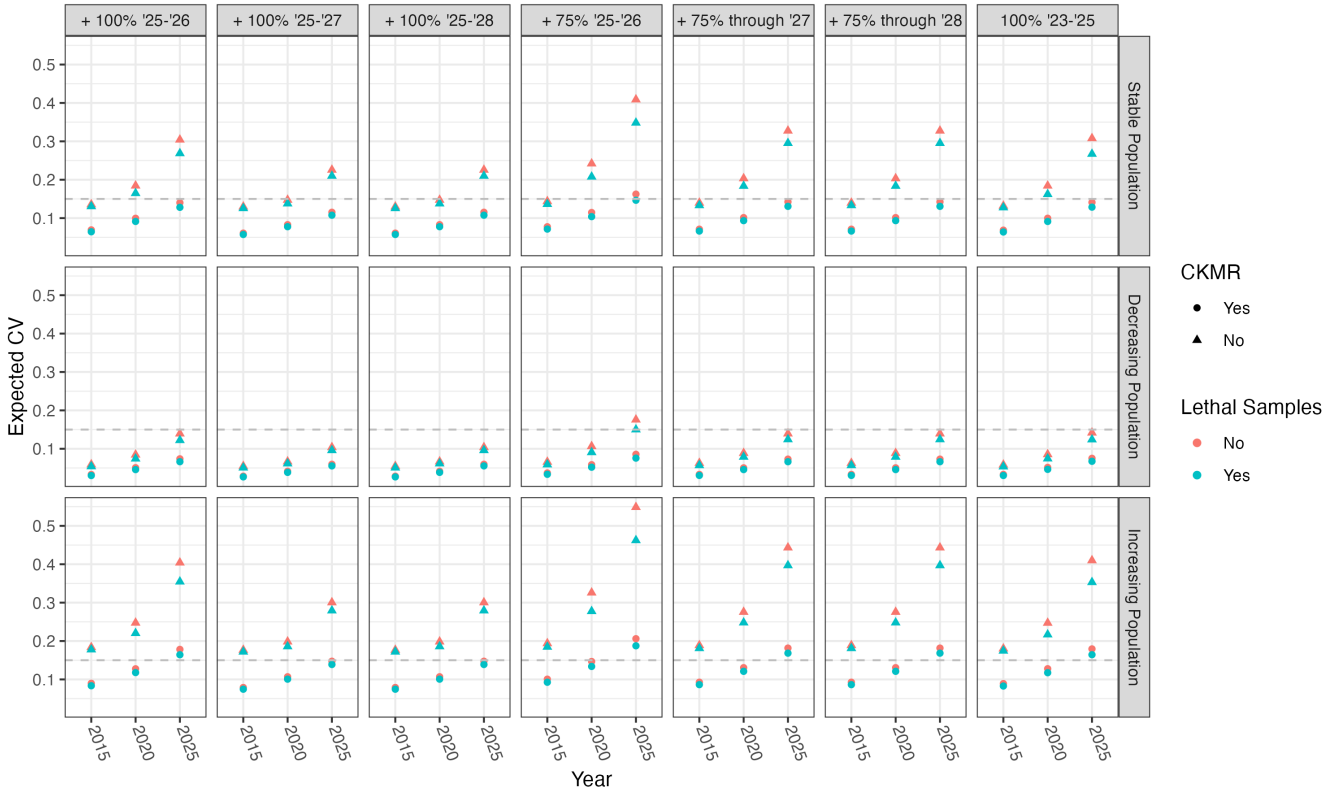

Figure 3: Expected CV of adult female abundance (vertical axis) in different years (horizontal axis) under different demographic (panel rows) and sampling (panel column) scenarios. Triangular points represent expected CVs from IMR alone, while circular points show expected CVs with ICKMR. The inclusion of lethal samples is indicated by red (no lethal samples included) or blue (lethal samples included in sampling years) points. The horizontal dashed line at CV = 0.15 represents an arbitrary threshold for decision making.

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[1]. 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

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 12/14/2024).

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 12/14/2024).

[1] It's Just A Design

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 02/20/2023).

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 02/03/2023).

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 12/14/2024).

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.

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 12/14/2024).

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 12/14/2024).

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 12/14/2024).

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 12/14/2024).

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 04/20/2024).

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 12/14/2024).

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 12/14/2024).

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 12/14/2024).

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 12/14/2024).

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 12/14/2024).

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 02/20/2023).

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 02/20/2023).

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 12/14/2024).

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.

Appendix

Appendix to methods section

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

535 $(y_1, a_1$ and y_2, a_2 , where $y_1 < y_2)$, then

$$\begin{aligned}
&\mathbb{P}\left[K_{21} = \mathrm{SP} | y_1, a_1, y_2, a_2\right] \\
&= \mathbb{P}\left[\text{Sample 2 turns out to be the same animal as sample 1} | y_1, a_1, y_2, a_2\right] \\
&= \frac{\mathbb{P}\left[\text{Sample 1 survived until Sample 2 was taken}\right] \mathbf{I}\left(y_2 - a_2 = y_1 - a_1\right)}{N_{\female}\left(y_2, a_2\right)} \\
&= \frac{\Phi\left(y_2 - y_1, a_1\right) \mathbf{I}\left(y_2 - a_2 = y_1 - a_1\right)}{N_{\female}\left(y_2, a_2\right)}. \qquad (11)
\end{aligned}$$

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

$$N\left(y + t, a + t\right) \equiv N\left(y, a\right) \Phi\left(t, a\right). \qquad (12)$$