

**Report on the Potential for Use of Mark-Recapture Methods to Estimate the Size of the Pacific Walrus Population**

Produced for  
United States Fish and Wildlife Service  
1011 East Tudor Road, Anchorage, Alaska 99503

By  
Bryan F.J. Manly  
Western Ecosystems Technology Inc.  
2003 Central Avenue, Cheyenne, WY 82001

**Contents**

Summary ..... 2

1. Introduction ..... 4

2. Simulation Experiment 1 ..... 5

3. Simulation Experiment 2 ..... 9

4. Some Further Simulation Results ..... 11

5. Discussion ..... 12

6. Recommendations ..... 14

References ..... 15

Appendix A: The Computer Program CMRSIM ..... 16

    Purpose of the Program ..... 16

    Summary of Options ..... 18

    Example ..... 19

Appendix B: The Computer Program CMRSIM1 ..... 28

    Purpose of the Program ..... 28

    Summary of Options ..... 29

    Example ..... 29

## Summary

- ! Simulation has been used to assess the potential use of mark-recapture methods for estimating the size of the Pacific walrus (*Odobenus rosmarus*) population. The simulation includes the generation of male and female populations at a number of different sampled and non-sampled sites, movement between sites, births and deaths, and the generation of mark-recapture data which is assumed to come from tissue samples and DNA analysis.
- ! With the method of Manly and Parr (1968) as it is usually applied the probability of an animal being captured at time  $t$  is estimated as the proportion of animals captured out of the group consisting of animals known to be alive at time  $t$  because they were captured before and after this time. Once this probability is estimated it is relatively straightforward to estimate the population size at time  $t$ . A survival rate can also be estimated, and all estimation can be done separately for males and females. The Manly and Parr method of estimation makes a minimum of assumptions and is suitable for an initial assessment of the value of mark-recapture methods with the Pacific walrus. Using it, it is found that even with relatively large numbers of animals sampled (5-10% of those at sampling sites) the accuracy of estimates may not be very good.
- ! If individual animals can be assigned to the age classes 0 (under one year old), 1 (from one to less than two years old), 2 (from two to less than three years old), and 3+ (three years old or more) then this information can be used with the Manly and Parr (1968) method of analysis. For example, all animals observed to be of age 1, 2 or 3+ in year  $t + 1$  were certainly alive in year  $t$ . They can therefore be added to the group that is used for estimating the probability of capture in year  $t$ . Using the Manly and Parr method in this way leads to very much improved estimates.
- ! It is concluded that sampling with probabilities of capture of about 0.01 or more for three or four years with age information should lead to estimates with reasonable precision (about a 20% coefficient of variation), where this sampling intensity corresponds to about 1,500 animals sampled each year.
- ! Computer programs are available to simulate estimation with and without using information on the ages of animals. These programs can be used to examine the value of different sampling schemes such as (i) sampling with a constant probability of capture, (ii) sampling with a relatively high probability of capture at the start and end of the sampling period, (iii) sampling with varying effort at different sampling sites, and (iv) sampling only in some years, etc.
- ! Estimates from mark-recapture data may be biased because of the unavailability of some animals for sampling in any year. The extent of this is difficult to assess because of lack of information about the location and movement of the animals. Movement data from a mark-recapture study would help to assess this problem, and it may be possible

to overcome any biases with an appropriate analysis of data taking the location of captures and recaptures into account.

- ! Because of the complex nature of Pacific walrus mark-recapture data, with age and sex related survival, movement and capture probabilities being likely to occur, a study of potential estimation methods is recommended. However, mark-recapture data would need to be collected before modeling issues can be properly addressed.
- ! It is recommended that a pilot study of mark-recapture be carried out for the dual purpose of assessing the extent to which random tissue sampling is possible, and to provide important information on movement patterns.

## 1. Introduction

Mark-recapture sampling has potential as a means for estimating the total size of the Pacific walrus (*Odobenus rosmarus*) population that mainly inhabits the Bering and Chukchi Seas, with 'marking' and 'recapture' data obtained by taking DNA samples of individual animals at land haul-out sites during the summer season (possibly with sampling on ice as well). Providing such samples allow individual animals to be identified with certainty, the animals can be considered to be 'marked' when they are first sampled, and 'recaptured' when they are sampled again, just as if they received a real physical mark.

The main purpose of this report is to describe the results of simulation studies that were designed to assess the value of mark-recapture methods to estimate the population size for Pacific walrus. Two approaches were considered to this end. First, a conventional mark-recapture study was considered, with no assumptions being made about the extent to which walrus can be aged in the field. With this situation the Manly and Parr (1968) method was used for estimating the population size, with separate estimates for males and females. Constant annual survival rates were also estimated separately for males and females. Obviously males and females have to be identified in order to calculate separate estimates for each sex. It is assumed that if necessary the DNA samples can provide this information.

A mark-recapture study without the use of information on the ages of walrus requires a very large amount of sampling. The second approach to mark-recapture therefore examined the extent to which results are improved if some limited information is available on the age of animals. In particular, it was assumed that it is possible to age animal as 0 (in the first year of life), 1 (in the second year of life), 2 (in the third year of life) or 3+ (three years of age or more). Aging to this extent does seem realistic with trained observers (Kelly *et al.*, 2000), and it improves estimation markedly.

A computer program CMRSIM to simulate walrus populations at different sampling sites, with movement between sites, births and deaths, sampling and estimation using no information on ages is described in Appendix A. This Appendix also includes an example set of input and output based on the use of six sampling sites for walrus, with various proportions of the populations at these sites and various probabilities of movement between the sites. There is also a seventh site corresponding to walrus at non-sampled locations, as explained in the Appendix. This Appendix should be consulted for details about the assumptions and methods used in the program because they also generally apply for the simulations described in the next sections of this report.

The computer program CMRSIM1 is very similar to CMRSIM in every respect except that it is assumed that animals can be aged as 0, 1, 2 or 3+, and this information is used to assist the estimation of population sizes and survival rates. The input file has the same format for both programs. CMRSIM1 is described in Appendix B, which also includes some example output. Like Appendix A, this second Appendix contains important

information for a proper understanding of the simulation results in the following sections of this report.

The programs CMRSIM and CMRSIM1 are written in FORTRAN and compiled using the WATCOM 77 compiler (Sybase, 2000).

## 2. Simulation Experiment 1

Some initial simulations were carried out to determine the likely effects on the accuracy of estimation of site fidelity (a tendency for individual walrus to stay at the same site from one year to the next), walrus moving between sites in groups, different numbers of sample years, and different probabilities of capture. For these and the other simulations considered in this report the population always consisted of 150,000 males and the same number of females.

There were 24 different scenarios considered altogether, with the number of sites, male and female survival probabilities, harvesting probabilities and proportions at the sites being as used in the example in Appendices A and B (Table A1). The 24 scenarios had a 3x2x2x2 factorial design, with the factors being:

- A: The number of sample occasions was low (3 years), medium (5 years) or high (10 years);
- B: The probability of capturing an individual walrus at sites 1 to 6 was low (0.05) or high (0.10);
- C: Site fidelity was either present (the individuals at a site returned there with probability 0.5) or absent (the probability of an individual being at site  $j$  was the same for all years); and
- D: Movement was either independent for all individuals, or the individuals at a site joined one of 20 movement groups at random and moved to the same site as all other walrus in that group.

The two levels of factor C require more explanation. When site fidelity was absent the probability of a walrus being at a particular site in one year was the same, irrespective of where it was the year before. The actual probabilities used are listed in Table A1, and are based on the proportions of males and females at different sites calculated from a 1990 survey, as shown in Table A2. When site fidelity was present the individual animals stayed at the same site from one year to the next with probability 0.5. For those that moved the probability of going to a site was the same as when there was no site fidelity.

The 24 scenarios corresponded to the 24 different combinations of the levels for factors A to C. For each scenario 10 populations were generated and were sampled 20 times each. Tables 1 and 2 summarise the results obtained.

The number of populations generated, and the number of times that each population is sampled are chosen by the users of CMRSIM and CMRSIM1. Here  $10 \times 20 = 200$  sets of data were generated, with the 20 samples from each population allowing coefficients of variation to be estimated for each population. The number of simulations is sufficient here to display clearly the effects of different factors.

Table 1 summarises the average percentage bias and the average percentage coefficient of variation for survival and population size estimates, for simulation results from program CMRSIM. With survival rates the presence of site fidelity is the only important source of bias with 5 or 10 years of sampling, with the negative bias indicating that estimates tend to be too low when individuals have a tendency to be at the same site from one year to the next. The bias is higher for males than for females. There is generally a negative bias with three years of sampling, but this is most likely due largely to the fact that survival probability estimates larger than 1.0 were not permitted. As a result, large under-estimates could occur, but not large over-estimates.

With population size site fidelity generally introduces a negative bias in estimates, which is larger for males than for females. However, in the absence of site fidelity there is a positive bias which only becomes small when there are 10 samples and a 0.10 probability of capture.

The bias in population size estimates when there is no site fidelity is a 'small sample' problem. The estimates are mostly quite reasonable but there is an occasional very large over-estimate of size. This can be seen, for example, in the example that is used in Appendix A. The estimates of the number of males at sample time 4 are 122298, 138370, 132344, 168222, 461315 for population 1, and 155783, 131364, 123888, 188233, 119932 for population 2, where the true population size was 150000 in both cases. Therefore, six out of the ten estimates were less than the true size but over all there is a positive bias because of the single large estimate of 461315.

Considering the coefficients of variation, it is interesting to see that site fidelity has reduced the variation in estimates, at the same time as introducing some bias. The type of movement (individual or grouped) seems to have little effect. As expected, the percentage coefficient of variation decreases with an increase in the probability of capture and the number of samples.

**Table 1 The average percentage bias and average percentage coefficient of variation in estimates of survival rates and population size from simulating 20 samples from each of 10 populations, for 24 different scenarios, with program CRMSIM.**

Samples	Capture probability	Site Fidelity	Type of movement	Percentage Bias in Estimates				Percentage Coefficient of Variation			
				Survival probability		Population size		Survival probability		Population size	
				Male	Female	Male	Female	Male	Female	Male	Female
3	0.05	Yes	Individual	-8.4	-8.1	-4.3	-1.3	20.8	17.2	48.7	40.0
3	0.05	Yes	Grouped	-7.7	-5.4	3.2	-0.4	19.1	16.6	64.1	35.5
3	0.05	No	Individual	-9.2	-7.4	16.9	28.2	23.5	20.9	75.7	81.1
3	0.05	No	Grouped	-11.8	-11.8	-3.3	28.8	25.1	22.3	69.0	93.7
3	0.10	Yes	Individual	-7.9	-3.6	-18.1	-9.7	10.7	9.6	10.4	10.1
3	0.10	Yes	Grouped	-7.8	-3.7	-20.9	-9.8	10.4	9.6	10.9	11.0
3	0.10	No	Individual	-1.6	0.3	7.4	6.5	13.5	11.6	27.3	20.8
3	0.10	No	Grouped	-5.3	-5.3	10.1	-1.5	13.8	11.7	26.9	18.1
5	0.05	Yes	Individual	-3.7	-1.7	-8.8	-3.4	8.5	6.6	23.7	20.3
5	0.05	Yes	Grouped	-2.7	-2.3	-6.9	-4.4	7.6	7.1	25.8	25.1
5	0.05	No	Individual	-1.7	0.1	15.8	10.8	10.5	8.4	57.4	40.5
5	0.05	No	Grouped	-0.3	0.2	18.8	18.5	9.5	8.5	59.8	53.5
5	0.10	Yes	Individual	-4.7	-2.4	-14.0	-7.7	2.9	2.7	7.4	6.7
5	0.10	Yes	Grouped	-4.0	-1.8	-14.6	-12.1	3.0	2.4	7.0	6.9
5	0.10	No	Individual	0.9	0.5	2.6	1.9	4.9	4.3	15.3	11.9
5	0.10	No	Grouped	0.2	0.4	3.2	3.5	4.8	4.0	13.9	12.5
10	0.05	Yes	Individual	-1.6	-0.8	-7.6	-7.0	1.6	1.4	11.9	12.0
10	0.05	Yes	Grouped	-1.8	-0.5	-6.9	-6.7	1.7	1.6	12.7	12.2
10	0.05	No	Individual	-0.1	0.2	4.3	4.4	2.6	2.2	23.1	22.6
10	0.05	No	Grouped	-0.4	0.0	2.8	2.3	2.5	2.0	24.1	23.1
10	0.10	Yes	Individual	-1.7	-1.0	-9.3	-8.5	0.6	0.5	4.6	4.5
10	0.10	Yes	Grouped	-1.4	-0.9	-8.4	-7.9	0.6	0.5	4.7	4.7
10	0.10	No	Individual	0.1	0.1	0.9	0.8	1.0	0.7	8.1	7.8
10	0.10	No	Grouped	-0.3	-0.1	1.2	-0.3	1.0	0.7	8.2	7.7

**Table 2 The average percentage bias and average percentage coefficient of variation in estimates of survival rates and population size from simulating 20 samples from each of 10 populations, for 24 different scenarios, with program CRMSIM1.**

Samples	Capture probability	Site Fidelity	Type of movement	Percentage Bias in Estimates				Percentage Coefficient of Variation			
				Survival probability		Population size		Survival probability		Population size	
				Male	Female	Male	Female	Male	Female	Male	Female
3	0.05	Yes	Individual	1.9	2.0	-7.2	-3.9	8.2	6.7	5.4	4.6
3	0.05	Yes	Grouped	3.0	0.4	-8.4	-3.5	6.9	7.6	4.9	4.7
3	0.05	No	Individual	-1.6	-0.5	0.4	0.5	11.8	9.3	7.6	6.4
3	0.05	No	Grouped	-4.1	-1.2	-0.3	1.1	11.6	8.9	8.3	6.5
3	0.10	Yes	Individual	1.8	1.0	-8.8	-4.5	4.5	4.1	2.6	2.4
3	0.10	Yes	Grouped	1.8	0.7	-8.7	-4.5	4.1	3.7	2.5	2.4
3	0.10	No	Individual	0.2	-0.4	0.4	-0.2	6.5	5.8	4.0	3.5
3	0.10	No	Grouped	-6.1	-0.5	-2.5	-1.5	4.9	4.6	4.2	3.2
5	0.05	Yes	Individual	0.2	-0.2	-5.6	-3.1	2.8	2.5	4.7	4.3
5	0.05	Yes	Grouped	-0.0	-0.4	-5.8	-3.6	2.8	2.3	4.4	4.3
5	0.05	No	Individual	0.3	0.1	0.9	0.5	3.8	3.0	7.0	5.5
5	0.05	No	Grouped	0.0	0.2	-0.3	0.0	4.0	3.0	6.3	5.6
5	0.10	Yes	Individual	-0.2	-0.4	-6.8	-3.6	1.3	1.4	2.3	2.2
5	0.10	Yes	Grouped	-0.1	-1.4	-5.2	-4.6	1.3	1.2	2.3	2.1
5	0.10	No	Individual	-0.2	0.0	-0.0	0.2	2.0	1.6	3.5	2.9
5	0.10	No	Grouped	-2.9	1.8	-1.0	0.6	1.8	1.5	3.2	3.0
10	0.05	Yes	Individual	-0.5	-0.2	-5.2	-2.6	0.6	0.5	4.0	3.8
10	0.05	Yes	Grouped	-0.4	-0.2	-4.9	-2.2	0.6	0.5	4.2	3.8
10	0.05	No	Individual	-0.1	-0.0	0.1	0.3	0.9	0.7	5.8	4.9
10	0.05	No	Grouped	-0.2	0.3	-0.1	0.3	0.9	0.7	5.8	4.8
10	0.10	Yes	Individual	-0.6	-0.4	-5.8	-3.0	0.3	0.3	2.1	1.9
10	0.10	Yes	Grouped	-1.0	-0.1	-6.4	-2.7	0.3	0.3	2.0	1.9
10	0.10	No	Individual	-0.0	-0.1	0.2	0.1	0.5	0.4	3.0	2.5
10	0.10	No	Grouped	-0.1	-0.1	-0.6	0.1	0.4	0.4	3.0	2.5



Table 2 summarises the results obtained using the program CMRSIM1, and hence on the assumption that animals can be assigned to the age classes 0, 1, 2 and 3+ in the field. The estimates are very much better in all respects, demonstrating the very definite value of using information on ages. The main message from this table in terms of population size is that a capture probability of 0.05 is far more than is required in order to produce a low coefficient of variation, even for a study only lasting for three years. However, site fidelity at the simulated level leads to a negative bias in estimates in the order of 5 to 9% for males, and about half that size for females.

### 3. Simulation Experiment 2

The results in Table 2 show that an acceptable coefficient of variation (say less than 20%) can in principle be obtained with less than 5% of the population captured each year, providing that animals can be assigned to the ages 0, 1, 2 and 3+ in the field. To investigate this further a second simulation experiment was carried out. This was similar to the first simulation experiment except that shorter studies of 2, 3 and 4 years duration were considered, and capture probabilities were reduced to 0.01 and 0.02. Thus the factors were the same as before, but with the following levels:

- A: The number of sample occasions was 2 years, 3 years or 4 years;
- B: The probability of capturing an individual walrus at sites 1 to 6 was 0.01 or 0.02;
- C: Site fidelity was either present (the individuals at a site in one year returned there the next year with probability 0.5) or absent (the probability of an individual being at site  $j$  was the same for all years); and
- D: Movement was either independent for all individuals, or the individuals at a site joined one of 20 movement groups at random and moved to the same site as all other walrus in that group.

This gives 24 different scenarios (factor combinations) as for experiment 1. However, only program CMRSIM1 was used for this simulation because estimation without using age information would clearly give very poor results with these low capture probabilities. Also, as explained in Appendix B, the estimation of population size is still possible with two years of sampling when animals can be aged, but this is not the case without age information. Table 3 summarises the simulation results in the same way that Tables 1 and 2 summarised the results for the first simulation experiment.

**Table 3 The average percentage bias and average percentage coefficient of variation in estimates of survival rates and population size from simulating 20 samples from each of 10 populations, for 24 different scenarios, using program CRMSIM1.**

Samples	Capture probability	Site Fidelity	Type of movement	Percentage Bias in Estimates				Percentage Coefficient of Variation			
				Survival probability*		Population size		Survival probability		Population size	
				Male	Female	Male	Female	Male	Female	Male	Female
2	0.01	Yes	Individual			1.9	5.7			28.9	30.2
2	0.01	Yes	Grouped			1.8	3.6			25.4	25.3
2	0.01	No	Individual			15.3	4.4			56.7	28.6
2	0.01	No	Grouped			5.7	9.0			34.6	34.6
2	0.02	Yes	Individual			-5.0	-2.0			14.1	12.9
2	0.02	Yes	Grouped			-4.9	-0.2			15.6	14.8
2	0.02	No	Individual			5.0	2.6			21.1	17.4
2	0.02	No	Grouped			2.1	2.6			21.5	16.1
3	0.01	Yes	Individual	-5.2	-8.8	1.7	1.1	22.6	23.1	24.5	22.7
3	0.01	Yes	Grouped	-3.7	-8.9	1.9	2.1	19.3	23.4	23.1	20.8
3	0.01	No	Individual	-17.1	-17.0	7.9	7.0	31.4	29.8	34.8	28.2
3	0.01	No	Grouped	-16.4	-15.2	5.9	9.2	28.4	27.6	28.6	30.2
3	0.02	Yes	Individual	-0.4	-3.0	-4.6	-2.4	15.0	14.3	12.7	11.5
3	0.02	Yes	Grouped	-1.8	0.8	-6.4	-1.8	14.8	13.5	11.3	11.4
3	0.02	No	Individual	-8.4	-3.9	6.3	2.6	22.3	17.0	20.0	15.0
3	0.02	No	Grouped	-7.4	-6.4	5.1	0.7	19.0	17.8	18.2	14.3
4	0.01	Yes	Individual	-3.8	-2.9	0.2	1.2	14.8	13.9	21.5	17.1
4	0.01	Yes	Grouped	-0.8	-0.6	-0.3	2.7	13.5	13.9	19.9	19.5
4	0.01	No	Individual	-7.3	-4.3	6.9	6.0	19.0	15.6	31.5	23.1
4	0.01	No	Grouped	-6.9	-4.4	6.2	3.4	17.7	16.9	28.1	22.4
4	0.02	Yes	Individual	1.4	1.3	-2.9	-2.0	8.3	7.8	11.0	9.8
4	0.02	Yes	Grouped	0.5	-1.0	-4.0	-1.4	9.2	8.8	10.9	10.3
4	0.02	No	Individual	-1.3	-1.4	1.7	1.0	11.0	10.9	14.2	13.2
4	0.02	No	Grouped	-0.7	-0.9	1.3	2.7	10.7	10.5	15.3	13.8

\*Survival probabilities cannot be estimated with only two years of sampling.

Table 3 shows some substantial negative biases in survival probability estimates. To some extent this is the result of not allowing estimates to exceed 1.0. The result is then a mean for estimates that tends to be lower than the true survival probability. The bias is substantial (up to 7% or more) unless the study lasts four years with a probability of capture of 0.02.

For population size estimation an average coefficient of variation of about 20% or less is just about possible with a two year study and a 0.02 probability of capture. However, some bias in estimation of the order of  $\pm 5\%$  seems possible, particularly for the estimation of the number of males. As was the case with survival estimation, a four year study with a 0.02 probability of capture seems necessary to keep the bias less than 5%.

#### **4. Some Further Simulation Results**

The results presented above are for situations where the amount of sampling effort is constant during a study. In practice this may not be desirable. For example it might be reasonable to have limited sampling effort in the first year of a study, increase the effort in the second year when any problems with sampling methods have been overcome, and then reduce the effort again, possibly maintaining a low level of sampling indefinitely.

Some further simulations were carried out to examine the results for this type of situation. It was assumed that a study would continue for five years. In year 1 the sampling effort is relatively low, with a probability of capturing animals being 0.01 unless they are in the part of the population that is not available for capture. In year 2 the probability of capture increases to 0.02, and then reduces to 0.01 again in year 3. Finally, in years 4 and 5 the probability of capture is only 0.005. Other aspects of the study were the same as for the simulations described above. First, 10 populations were simulated with site fidelity, and each was sampled 20 times. Then 10 populations were simulated without site fidelity and again sampled 20 times each.

The results are summarised in Table 4. If anything, the results are better with site fidelity. In terms of survival rates the results are not really very satisfactory because the coefficients of variation (10-14%) mean that the standard error of estimation is about 0.1, so that 95% confidence intervals are something like the estimated values  $\pm 0.2$ . However, population size estimates are quite good for years 1 and 2, reasonable for year 3, but very poor for year 4.

**Table 4 Summary of the results from simulations with probabilities of capture of 0.01, 0.02, 0.01, 0.005 and 0.005 for years 1 to 5 in terms of the average percentage bias and the average percentage coefficient of variation (CV).**

Parameter	Results with Site Fidelity		Results without Site Fidelity	
	Mean Bias (%)	Mean CV (%)	Mean Bias (%)	Mean CV (%)
Male survival rate	0.8	10.7	-2.6	13.8
Female survival rate	0.1	10.0	-1.5	11.9
Male Population size year 1	-1.3	15.1	4.4	21.2
Male Population size year 2	-0.7	10.8	0.5	14.9
Male Population size year 3	1.5	20.6	10.1	28.6
Male Population size year 4	19.5	84.6	32.7	97.4
Female Population size year 1	1.9	14.2	3.3	14.4
Female Population size year 2	-2.2	10.0	1.1	11.7
Female Population size year 3	1.2	18.0	6.7	22.8
Female Population size year 4	19.1	72.4	23.7	66.4

If nothing else, this example illustrates that a sampling plan with variable effort in different years may be worthy of further consideration.

## 5. Discussion

The simulation results presented above suggest that a reasonable level of accuracy for the estimation of the size of the Pacific walrus population (a coefficient of variation of 20% or less) can be obtained with sampling probabilities of the order of 0.01 or more for three or four years, or higher probabilities for two or three years. The assumption made in obtaining this result is that there are about 150,000 male and 150,000 female walrus, of which about half of them are unavailable for capture because of their location (Table A2). This means that a 0.01 sampling probability corresponds to about 0.5% of the population being sampled, which is about 750 males and 750 females. Presumably this level of effort is quite practical. Therefore, mark-recapture sampling does seem to be feasible for estimating the size of the Pacific walrus population providing that the individual animals can be accurately aged until they are three years old.

Because this report is an initial investigation into the feasibility of using mark-recapture with Pacific walrus, some simplifying assumptions have been made. In particular, it has been assumed that males of all ages have the same yearly survival rate. Similarly, females of all ages are assumed to have the same survival rate. Also, all the reported simulation results are based on a single survival rate of 0.90 per year, although the simulation model allows the male and female survival rates to be different. In reality, pups and juveniles will have low survival rates compared to adults and this would be allowed for in the analysis of the real data.

It is also assumed in the simulation model that males of all ages have the same movement patterns. Females are allowed to have different movement patterns from males, but again this is independent of age. In reality, movement patterns will presumably be related to age. Hence this is likely to have to be allowed for in the analysis of the real data.

There are some uncertainties that may hinder obtaining good estimates from a mark-recapture study. The first of these relates to the fraction of the population that is unavailable for sampling in any year. This does not include individual animals that are temporarily unavailable for sampling at a site because they are at sea. It is assumed that sampling will take place over several days or weeks so that most animals at a site will be available for sampling at least part of the time. Rather, the unavailable animals are assumed to be at locations that are not sampled at all, either because they are unknown, or because they are not accessible.

The fact that some animals are unavailable for sampling in any year does not in itself upset mark-recapture estimates. If different animals are unavailable in different years then it causes no biases at all. However, if there is a tendency for animals to remain unavailable for several years then this will tend to result in the estimated population size being too low. It is obvious, for example, that if there is a group of animals that are never present at sampling sites then they cannot be included in the population size that is estimated by mark-recapture sampling.

In the simulations considered in this report it was assumed that there are initially 51% of females and 41% of males "at sea" and unavailable for capture. The actual percentages are unknown for the real population. If higher proportions at sea were assumed for the simulations then this would reduce sample sizes (because each walrus that is available has a certain probability of capture) and make recaptures more difficult. This implies that accuracy would improve if there were smaller percentages at sea, but would get worse if there were higher percentages at sea. However, the effect will also depend on whether there is site fidelity or not. Large percentages of males and females at sea with a tendency to stay there would presumably introduce a high bias in estimation. The percentages are fixed in the input to the simulation programs and are easily changed to investigate the effect of doing this.

With the knowledge about walrus movement that exists at the present time it is not really possible to be sure about the extent of the bias, if any, that might result from animals being unavailable for sampling either temporarily or permanently. However, mark-recapture sampling would in itself provide more information on movement, and it is likely that an allowance for movement would be included in the model used to analyse the mark-recapture data obtained from a real study.

Most models for mark-recapture data assume that when a sample is taken there is a certain probability  $p$  (say) of an animal being sampled. This probability may vary with the time and location of sampling, but should be constant at one time in one place. With

Pacific walrus it seems that there may be difficulties in sampling all animals with the same probability because of the different behavior of males and females and animals of different ages. However, this may not be a problem as the final method for data analysis is likely to produce separate estimates of population size for males and females of different ages. Thus what will be needed will be that effectively random samples will be taken of males in each of the four age classes 0, 1, 2 and 3+ years and females in each of the same age classes. The sampling probability need then not be the same for each of the eight sex-age classes. Presumably this type of stratified random sampling of different ages and sexes will be more practical than attempting to take a completely random sample of all individuals.

Another point to note here is that it may not be desirable to sample the recently born 0 age individuals. In that case the mark-recapture approach can still be used to estimate the numbers of individuals and survival rates for individuals aged 1 year old or more. Extending this idea, it would also be possible to only sample mature adults and estimate the sizes and survival rates just for this part of the population.

With real data some age stratification of sampling is likely to be needed with harvested animals as well, because of the apparent selection for mature animals and calves of the year.

## 6. Recommendations

- ! Relatively simple methods of estimation were used for the simulations described in this report. However, as noted in Appendix A, there are a number of methods that are available for analysing mark-recapture data from several geographical locations (Schwarz and Seber, 1999, Section 4.2.3), and it is possible that one of these methods would be more appropriate for analysing the real data. If mark-recapture is to be considered further as an option for estimating the size of the Pacific walrus population then it is recommended that a study of these methods of analysis should be made to decide which of them are of potential use. It seems that what is required is a method that allows for movement between sites, different sampling probabilities at different sites and for different ages and sexes, takes into account the known ages of animals, and allows information on harvested animals to be used. Because of the complex nature of the data that would be obtained, it might be necessary to modify an existing method of analysis. Although a preliminary study of methods of analysis could be made at any time, the patterns of movement in the real population will be an important consideration for data analysis and the final decision about how to model the data for analysis will depend on the nature of the data actually obtained. Consequently, data would have to be collected for two years at least before modeling issues can be properly assessed.
  
- ! In view of the uncertainties about the extent to which random sampling can be carried out (with or without stratification by sex and age), and uncertainties about how mark-

recapture data will need to be modeled, it would be very advantageous to carry out a pilot study as soon as possible to collect some data. This would provide information about the ease with which DNA samples can be taken, and the extent to which these can be taken at random. It would also provide information on movement patterns so that method of estimation can be investigated better. Unless there are serious problems with sampling, it should be possible to use the data obtained for the start of a long-term study.

## References

- Chivers, S.J. (1999). Biological indices for monitoring population status of walrus evaluated with an individual-based model. In *Marine Mammal Survey and Assessment Methods* (eds. G.W. Garner, S.C. Amstrup, J.L. Laake, B.F.J. Manly, L.L. McDonald and D.G. Robertson), pp. 239-47. Balkema, Rotterdam.
- Fedoseyev, G.A. and Goltsev, V.N. (1969). Age-sex structure and reproductive capacity of the Pacific walrus population. *Zoological Journal* 48: 407-13. Canadian translation of fisheries and aquatic sciences 5503, 1990.
- Gilbert, J.R., Fedoseyev, G.A., Seagars, D., Razlivalov, E. and LaChugin, A. (1992). *Aerial Census of Pacific Walrus, 1990*. USFWS R7/MMM Technical report 92-1.
- Kelly, B.P., Taras, B. and Quakenbush, L. (2000). *Monitoring Age and Sex Composition of Pacific Walrus: 12-30 July 1999 Cruise Report*. Report to the North Pacific Marine Research Program and the Eskimo Walrus Commission.
- Manly, B.F.J. and Parr, M.J. (1968). A new method of estimating population size, survivorship, and birth rate from capture-recapture data. *Transactions of the Society for British Entomology* 18: 81-9.
- Schwarz, C.J. and Seber, G.A.F. (1999). Estimating animal abundance: review III. *Statistical Science* 14: 427-56.
- Sybase (2000). Information on WATCOM FORTRAN 77. Available at the web site <http://www.sybase.com/products/archivedproducts/watcomfortran77/>.
- U.S. Fish and Wildlife Service (1998). *Pacific Walrus (Odobenus rosmarus divergens): Alaska Stock*. Stock Assessment, Marine Mammals Management, 1011 East Tudor Road, Anchorage, AK 99503.

## Appendix A: The Computer Program CMRSIM

### Purpose of the Program

The Windows computer program CMRSIM was written to simulate the movement between land haul-out sites, survival, mark-recapture sampling, and the estimation of the population size and the survival rate per year for the Pacific walrus (*Odobenus rosmarus divergens*). Different distributions over the haul-out sites are assumed for males and females, and also different demographic parameters, and possibly different sampling probabilities. In brief, the program assumes that the walrus population is distributed at a number of land haul-out sites where they can be sampled in the summer months. Some part of the population is also 'at sea' and not available for sampling. The proportions at different sites and at sea are different for males and females.

This program differs from CMRSIM1 (described in Appendix B) only in terms of how data are analysed. With CMRSIM it is not assumed that data will be available on the age of animals. However with CMRSIM1 it is assumed that walrus can be accurately assigned to the ages 0 (the first year of life), 1 (the second year of life), 2 (the third year of life), or 3+ (older than 3 years).

Sampling is assumed to be repeated for a number of years. What is necessary for the mark-recapture method to work is that there is movement of the walrus between sites from one year to the next. In particular, it is necessary that the walrus that are at sea in one year become available for sampling in other years. Therefore an important part of the input to the program is a site transition probability matrix. If there are  $S$  haul-out sites then this matrix has  $S + 1$  rows and columns, with the last row and column representing the walrus that are at sea. The element in the  $i$ th row and  $j$ th column of the matrix is the probability that a walrus at site  $i$  in year  $t$  moves to site  $j$  in year  $t + 1$ . There are two transition probability matrices, one for males and one for females. This is necessary because the distribution over sites is different for males and females.

An important consideration is whether the walrus have site fidelity or not. Site fidelity means that a walrus at site  $i$  in one year has a tendency to remain at that site. This will cause problems if the walrus that are at sea in one year tend to stay at sea in the following years because they will then represent a part of the population that is not sampled properly. On the other hand, if all walrus have the same probability of being at site  $i$  in a year, irrespective of where they were the previous year then estimation from the mark-recapture data will be straightforward.

Another important consideration is whether individuals move between sites independently or in groups. Either of these possibilities is allowed for in CMRSIM. If there is group behavior then each site in year  $t$  has  $G$  movement groups. All of the individuals in a group move to the same site in year  $t + 1$ , where this is chosen using the transition probability matrix. Each year, each individual joins movement group  $i$  with probability  $1/G$ . There are separate movement groups for males and females.



It is assumed that there is a certain probability of sampling a walrus at a haul-out site, which may vary from site to site, from year to year. It may also vary for males and females at the same site.

The subsistence harvest of walrus is another source of sample data. It is assumed that there is a certain probability of a male walrus being harvested and identified from a DNA sample. Females also have a probability of being harvested, which may be different from the male probability. The natural mortality rate per year must be set for the simulated population. This may be different for males and females. It includes the mortality from the subsistence harvest. To maintain a constant population size, all deaths are replaced by births of new individuals.

There are a number of ways to analyse mark-recapture data from a population distributed over several geographical locations, as reviewed by Schwarz and Seber (1999, Section 4.2.3). CMRSIM uses a very simple approach based on the method of Manly and Parr (1968). The method for estimating the size of the population in year  $i$  consists of identifying those individuals that were captured both before and after this year, so that they were certainly in the population. Let  $p_i$  be the proportion of these individuals that are captured in year  $i$ , which is assumed to be an unbiased estimate of  $p_i$ , the proportion of all individuals captured in year  $i$ . Also, let  $N_i$  be the total population size in year  $i$ , and  $n_i$  be the total number of animals sampled in that year. Then it can be expected that  $n_i \approx p_i N_i$ , so that

$$\hat{N}_i = n_i / p_i \quad (A1)$$

is a moment estimator of  $N_i$ , which can be applied for  $i = 2, 3, \dots, T - 1$ , where  $T$  is the total number of sampling years. Having estimated the total population size, this can be allocated out to the sites where the population is according to the distribution for the individuals known to be in the population. These calculations are done separately for males and females.

The estimation of survival is also based on a very simple approach. Let  $s$  be the yearly survival rate, let  $M_i$  be the number of individuals first captured in year  $i$ , and let  $R_i$  be the number of individuals recaptured in year  $i$ . Then the expected total number of recapture in years 2 to  $T - 1$  is

$$E\left(\sum_{i=2}^{T-1} R_i\right) = M_1(s^1!_2 + s^2!_3 + \dots + s^{T-1}!_T) + M_2(s^0!_3 + s^1!_4 + \dots + s^{T-2}!_T) + \dots + M_{T-2} s^0!_{T-1}. \quad (A2)$$

If the left-hand side of this equation is replaced by the total observed number of recaptures and  $p_i$  is replaced by  $p_i$ , then the result is an equation with  $s$  as the single unknown value. Therefore  $s$  is estimated by solving this equation. This is done separately for males and females. Recaptures in the last year of the study cannot be included in this calculation because the probability of capture cannot be estimated for this year.

## Summary of Options

CMRSIM is designed to simulate the collection of mark-recapture data to a reasonable level of realism, and summarise the properties of the estimates that are produced from the data. The key features are:

- ! The male and female components of the population are treated separately, with different distributions over sites, different probabilities of moving between pairs of sites, and possibly different capture probabilities.
- ! Males of all ages have the same yearly probability of survival. Females of all ages also have the same yearly probability of survival, which may or may not be equal to the male probability.
- ! There can be up to 50 sampling sites and up to 20 sample times (assumed to be different years). In practice the sites might be land haul-out locations or locations on ice.
- ! Sampling does not have to take place every year. Setting all sampling probabilities to zero in a year simulates the situation where there is no sampling in that year. Sampling can also be missed for just some of the sites in some of the years.
- ! Input parameters for the simulation (which are separate for males and females) are a yearly survival probability, a yearly probability of an animal being harvested and a DNA sample taken, the total number of animals, the proportion of animals at each of the sampling sites (with the balance at sea and not available for sampling), the probability of an animal moving from site  $i$  to site  $j$  between two sample times for all  $i$  and  $j$  (with an extra site for animals at sea), and capture probabilities for each site at each sample time.
- ! The movement from site to site between sample times can either be done independently for each animal, or in groups. In the case of groups, there are  $G$  movement groups at each site and an animal joins one of these chosen at random and moves to the same site as the rest of the group. Males and females have separate movement groups.
- ! Given a set of input parameters, the user chooses how many populations to simulate, and how many sets of mark-recapture data to generate from each population.
- ! Estimates of population size are calculated using the Manly and Parr (1968) method because this is simple and makes relatively few assumptions compared to some other methods.

- ! Estimates of survival are based on assuming that the yearly survival rate is constant. The estimates are those values that make the observed total number of recaptures equal to the expected number, except that estimates over 1.0 are not allowed.
- ! The simulated population size is kept constant, with each death giving rise to the birth of a new animal.

## Example

As an example, consider the set of input data for the program that is shown in Table A1. Here there are 6 sites corresponding to Bristol Bay and the eastern Bering Sea, south Kamchatka, north Kamchatka, coastal southern Chukotka, northern Chukotka, and Wrangel island (Figure A1). These areas were chosen based on site groupings from the 1990 walrus survey (Gilbert *et al.*, 1992, Table 6). Sampling is for 5 years.

**Table A1 Input data for a simulation. Each line has a free format.**

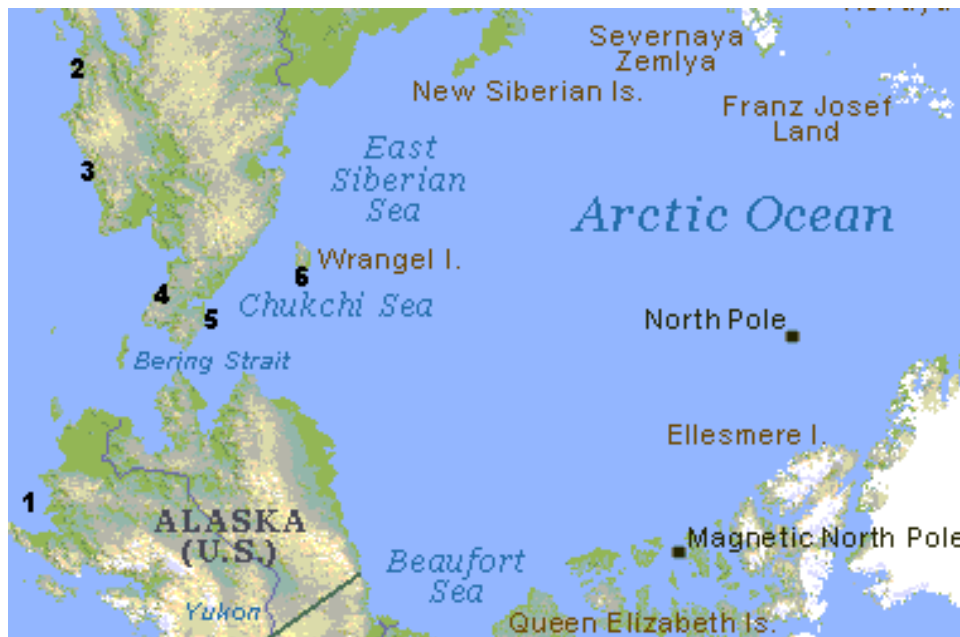
<p>SIMULATION DATA FOR PACIFIC WALRUS  6 5  0.90 0.90  0.010 0.010  150000 150000  0.0451 0.0421 0.0301 0.0481 0.0561 0.2642  0.0050 0.0047 0.0033 0.0053 0.0838 0.4887  0.0451 0.0421 0.0301 0.0481 0.0561 0.2642 0.5144  0.0451 0.0421 0.0301 0.0481 0.0561 0.2642 0.5144  0.0451 0.0421 0.0301 0.0481 0.0561 0.2642 0.5144  0.0451 0.0421 0.0301 0.0481 0.0561 0.2642 0.5144  0.0451 0.0421 0.0301 0.0481 0.0561 0.2642 0.5144  0.0451 0.0421 0.0301 0.0481 0.0561 0.2642 0.5144  0.0050 0.0047 0.0033 0.0053 0.0838 0.4887 0.4092  0.0050 0.0047 0.0033 0.0053 0.0838 0.4887 0.4092  0.0050 0.0047 0.0033 0.0053 0.0838 0.4887 0.4092  0.0050 0.0047 0.0033 0.0053 0.0838 0.4887 0.4092  0.0050 0.0047 0.0033 0.0053 0.0838 0.4887 0.4092  0.0050 0.0047 0.0033 0.0053 0.0838 0.4887 0.4092  0.05 0.05 0.05 0.05 0.05  0.05 0.05 0.05 0.05 0.05  0.05 0.05 0.05 0.05 0.05  0.05 0.05 0.05 0.05 0.05  0.05 0.05 0.05 0.05 0.05  0.05 0.05 0.05 0.05 0.05  0.05 0.05 0.05 0.05 0.05  0.05 0.05 0.05 0.05 0.05  0.05 0.05 0.05 0.05 0.05  0.05 0.05 0.05 0.05 0.05  0.05 0.05 0.05 0.05 0.05</p>	<p><i>Title for the data*</i>  There are 6 sites and 5 years of sampling  The yearly survival probability is 0.90 for both males and females  The probability of being included in the subsistence harvest is 0.010 for males and females  There are 150,000 males and 150,000 females in the population  These are the initial proportions of males at the 6 sampling sites  These are the initial proportions of females at the 6 sampling sites  This and the following 6 rows give the probabilities  of moving between sites for males. The first of the 7  rows gives the probability of a male at site 1 staying  at site 1 (0.0451), moving to site 2 (0.0421), moving to  site 3 (0.0301). etc. Row <i>i</i> gives the probabilities for  a male at site <i>i</i>. Site 7 is at sea (not available for sampling)  This and the next 6 rows gives the transition probabilities  for females  This and the next 5 rows gives probabilities of capture for males, with  a row for each of the 6 sites and a column for each of the 5 sample times  The last part of the input is the capture probabilities for females</p>
---	--

\*Text in italics are comments.

The yearly survival probability is assumed to be 0.9, which seems realistic for all ages combined (Fedoseyev and Goltsev, 1969, p. 9; Chivers, 1999). The yearly probability of being included in the subsistence harvest and also sampled is assumed to be 0.010, based on an estimated total reported harvest (excluding 'struck and lost') of about 3,000 animals (U.S. Fish and Wildlife Service, 1998). There are assumed to be 150,000 males and 150,000 females in the population, where these totals include those that are not available for sampling at any point in time.

The proportions of males and females initially present at each sampling site are chosen to match the total numbers from the 1990 survey of Gilbert *et al.* (1992), as shown in Table A2. For this example these proportions are also used for the probabilities of movement between sites, such that the probability of an animal moving to site *i* between two sample times is equal to the proportion initially at site *i*, irrespective of where the movement is from. In general, it might be thought that it is more reasonable to assume some site fidelity, so that an animal at site *i* in one year has a tendency to still be at that site in the following year. The proportion of the population assumed to be unavailable for sampling is 0.51 for males and 0.41 for females.

**Figure A1** The six sites used for the example simulation, from site 1 (Bristol Bay and the east Bering sea) to site 6 (Wrangel Island). Each 'site' is actually a set of haulout locations in the same general area.



Finally, the probabilities of capture are set at 0.05 for each of sites 1 to 6 for all five sample times, individual animals moved independently, 2 populations were generated, and each population was sampled 5 times.

The output obtained from the data listed in Table A1 is shown in Table A3 together with various explanatory comments. Additional options were that individuals moved independently between site (there are no movement groups), two populations were generated, and each population was sampled five times. These additional options are chosen after the program has been started.

**Table A2 The assumed initial distribution of walrus at the six haul-out sites.**

Site		1990 Survey	Used in simulation	Numbers by sex		Proportions	
				Male	Female	Male	Female
1	Bristol Bay & E. Bering Sea	7522	7500	6750	750	0.0451	0.0050
2	S. Kamchatka	6700	7000	6300	700	0.0421	0.0047
3	N. Kamchatka	5295	5000	4500	500	0.0301	0.0033
4	S. Chukotka	8380	8000	7200	800	0.0481	0.0053
5	N. Chukotka	21002	21000	8400	12600	0.0561	0.0838
6	Wrangel Island	112848	113000	39550	73450	0.2642	0.4887
7	Unavailable for capture	Unknown	138500	77000	61500	0.5144	0.4092
			300000	149700	150300	1.0000	1.0000

It is important to note that the input to the program CMRSIM has the following structure, rather than being exactly as in Table A1:

- Line 1: NSD, which is an integer (1 or more) indicating how many different simulation scenarios are being considered.
- Line 2 on: The input parameters for the first simulation scenario, in the format of Table A1.
- Next: If NSD = 1 then there is no more input. Otherwise the parameters for a second scenario are input in the format of Table A1. Further sets of parameters for more scenarios follow if NSD > 2.

For example, Table A4 shows a full set of data for running the program for two scenarios (NSD = 2). A set of data like this will generally just be set up in an ASCII text editor. Note that when CMRSIM is started it describes the format needed for the data file.

**Table A3 Output from running CMRSIM with the data in Table A1, with comments in italics on the right-hand side.**

```
#####
#                               CMRSIM                               #
#                               #                                   #
# A program to simulate mark-recapture sampling and                 #
# estimation for a population that occurs at up to 50              #
# sites, up to 20 sample times, and up to 300,000 males and      #
# 300,000 females sampled altogether. There is movement           #
# between sites between sample times. Males and females         #
# are treated separately for the simulation and may have          #
# different parameters. Harvesting data is used for              #
# improving estimation.                                           #
#                                                                   #
#                               Version 1.1 (October, 2000)        #
#                               Written by Bryan F.J. Manly        #
#                               Western EcoSystem Technology Inc.   #
#                               Cheyenne, Wyoming 82001           #
#                               bmanly@west-inc.com                #
#####
```

Date: 4/11/2000 Time: 13:35

*The first part of the output summarises the input parameters for the simulation.*

Data: SIMULATION DATA FOR PACIFIC WALRUS  
 Number of Sites = 6 Number of Sample Times = 5  
 Male Survival Probability = 0.900 Female Survival Probability = 0.900  
 Male Harvesting Probability = 0.010 Female Harvesting Probability = 0.010  
 (Survival includes survival from harvesting)  
 Total Males = 150000 Total Females = 150000

Male Proportions Initially at Different Sites  
 (Last site is all unknown locations)  
 0.05 0.04 0.03 0.05 0.06 0.26 0.51

Female Proportions Initially at Different Sites  
 (Last site is all unknown locations)  
 0.00 0.00 0.00 0.01 0.08 0.49 0.41

Movement Probability Matrix For Males  
 (Element in row i and column j is the probability of moving from site i to site j in one unit of time)

	1	2	3	4	5	6	7
1	0.05	0.04	0.03	0.05	0.06	0.26	0.51
2	0.05	0.04	0.03	0.05	0.06	0.26	0.51
3	0.05	0.04	0.03	0.05	0.06	0.26	0.51
4	0.05	0.04	0.03	0.05	0.06	0.26	0.51
5	0.05	0.04	0.03	0.05	0.06	0.26	0.51
6	0.05	0.04	0.03	0.05	0.06	0.26	0.51
7	0.05	0.04	0.03	0.05	0.06	0.26	0.51

Movement Probability Matrix For Females  
 (Element in row i and column j is the probability of moving from site i to site j in one unit of time)

	1	2	3	4	5	6	7
1	0.00	0.00	0.00	0.01	0.08	0.49	0.41
2	0.00	0.00	0.00	0.01	0.08	0.49	0.41
3	0.00	0.00	0.00	0.01	0.08	0.49	0.41
4	0.00	0.00	0.00	0.01	0.08	0.49	0.41
5	0.00	0.00	0.00	0.01	0.08	0.49	0.41
6	0.00	0.00	0.00	0.01	0.08	0.49	0.41
7	0.00	0.00	0.00	0.01	0.08	0.49	0.41

Capture Probability Matrix for Males  
 Sample Time ==>

Site	1	2	3	4	5
1	0.05	0.05	0.05	0.05	0.05
2	0.05	0.05	0.05	0.05	0.05
3	0.05	0.05	0.05	0.05	0.05
4	0.05	0.05	0.05	0.05	0.05

5 0.05 0.05 0.05 0.05 0.05  
6 0.05 0.05 0.05 0.05 0.05

Capture Probability Matrix for Females

Sample Time ==>  
Site 1 2 3 4 5  
1 0.05 0.05 0.05 0.05 0.05  
2 0.05 0.05 0.05 0.05 0.05  
3 0.05 0.05 0.05 0.05 0.05  
4 0.05 0.05 0.05 0.05 0.05  
5 0.05 0.05 0.05 0.05 0.05  
6 0.05 0.05 0.05 0.05 0.05

Burn in Movements = 1  
Number of Movement Groups (IGP) = 0  
Number of Simulated Populations = 2  
Number of samples per population = 5  
Random number seed = 100

Summary of Generated Population  
Numbers at Sites

*These are the numbers for simulated population 1.*

Sample Time ==>  
Site 1 2 3 4 5  
1 M 6765 6658 6824 6683 6754  
F 749 755 724 801 772  
2 M 6315 6366 6263 6233 6258  
F 705 640 661 748 683  
3 M 4515 4465 4480 4564 4450  
F 494 455 466 487 477  
4 M 7214 7251 7196 7201 7198  
F 794 759 845 790 824  
5 M 8415 8482 8549 8499 8546  
F 12570 12202 12546 12544 12537  
6 M 39630 39631 39755 39631 39766  
F 73305 73593 73251 73152 73302  
7 M 77146 77147 76933 77189 77028  
F 61383 61596 61507 61478 61405

Sites Where Individual Male Animals Are  
(0 = dead or elsewhere, -1 = harvested)

*First samples from the population.*

Sample Time ==>  
Anim 1 2 3 4 5  
1 1 7 0 0 0  
2 1 7 6 6 6  
3 1 4 2 7 6  
4 1 7 7 7 0  
5 1 7 7 7 7

*Output abbreviated here.*

.  
.  
.  
210165 0 0 0 0 1  
210166 0 0 0 0 7  
210167 0 0 0 0 3  
210168 0 0 0 0 7  
210169 0 0 0 0 7

Sites Where Individual Female Animals Are  
(0 = dead or elsewhere, -1 = harvested)

Sample Time ==>  
Anim 1 2 3 4 5  
1 1 0 0 0 0  
2 1 6 7 7 5  
3 1 6 6 6 6  
4 1 0 0 0 0  
5 1 5 7 6 7

*Abbreviated here as well.*

.  
.  
.  
210132 0 0 0 0 6

210133 0 0 0 0 7  
 210134 0 0 0 0 6  
 210135 0 0 0 0 6  
 210136 0 0 0 0 6

Sites Where Individual Male Animals Are Captured  
 Sample Time ==>

Anim 1 2 3 4 5  
 1 1 0 0 0 0  
 2 0 0 0 0 6  
 3 0 6 0 0 0  
 4 0 0 0 0 4  
 5 0 0 0 0 6  
 .  
 .  
 .  
 23267 0 0 0 0 6  
 23268 0 0 0 0 6  
 23269 0 0 0 0 3  
 23270 0 0 0 0 4  
 23271 0 0 0 0 1

*Abbreviated here as well.*

Sites Where Individual Female Animals Are Captured  
 Sample Time ==>

Anim 1 2 3 4 5  
 1 1 0 0 0 0  
 2 0 0 0 -1 0  
 3 0 0 -1 0 0  
 4 0 -1 0 0 0  
 5 0 0 0 0 6  
 .  
 .  
 .  
 26843 0 0 0 0 6  
 26844 0 0 0 0 6  
 26845 0 0 0 0 6  
 26846 0 0 0 0 6  
 26847 0 0 0 0 6

*Abbreviated here as well.*

Estimated Numbers at Sites *For the first mark-recapture sample from the first population.*

Sample Time ==>  
 Site 1 2 3 4 5  
 1 M 0 49782 0 30575 0  
 F 0 0 0 0 0  
 2 M 0 0 0 0 0  
 F 0 0 0 0 0  
 3 M 0 24891 0 0 0  
 F 0 0 0 0 0  
 4 M 0 0 0 15287 0  
 F 0 0 0 0 0  
 5 M 0 24891 25545 15287 0  
 F 0 8192 12883 31928 0  
 6 M 0 49782 153272 61149 0  
 F 0 106498 103064 127714 0  
 All M 0 149345 178817 122298 0  
 SE 0 60186 66827 42551 0  
 F 0 114690 115947 159642 0  
 SE 0 40735 39536 50210 0

Estimated Male Survival Rate = 0.831  
 Estimated Female Survival Rate = 0.814

*The following table summarises the results  
 from taking 5 samples from each of the 2  
 populations.*

Estimated Parameters From Simulations  
 Survival Estimated Number of Males at Times  
 Pop Spl Male Female 2,3, ... Followed by Female Estimates  
 1 1 0.831 0.814 149345 178817 122298 114690 115947 159642  
 1 2 0.881 0.975 107603 208255 138370 165396 152915 217624  
 1 3 0.845 0.853 196829 161820 132344 116414 174419 157474



1	4	0.855	0.891	110537	98474	168222	111766	114758	212219
1	5	0.973	1.000	93896	111475	461315	282501	145343	184814
% Bias		-2.6	0.7	-12.2	1.2	36.3	5.4	-6.2	24.2
% CV		6.3	8.8	27.9	30.7	96.4	48.6	17.0	18.8
2	1	0.876	0.925	107053	183675	155783	136808	169460	161288
2	2	0.933	0.931	143602	267574	131364	135511	203766	155375
2	3	0.840	1.000	250409	117083	123888	442504	125587	219456
2	4	0.920	0.873	161783	117915	188233	221903	129419	127924
2	5	0.768	0.879	202150	92125	119932	103477	241566	127994
% Bias		-3.6	2.4	15.3	3.8	-4.1	38.7	16.0	5.6
% CV		7.4	5.7	36.8	47.4	19.0	92.2	33.0	24.9

*The % Bias is the average bias from the simulated estimates, as a percentage of the true parameter value. The % CV is the coefficient of variation (the standard deviation of estimates divided by the true parameter value) expressed as a percentage.*

**Table A4 Example of a full input file for running the program CMRSIM. Here there are two scenarios simulated, with and without site fidelity. For scenario 1 the sampling probabilities in years 1, 3 and 5 are 0.05, 0.01 and 0.01, respectively, with no samples in years 2 and 4. For scenario 2 there are no samples in years 2 and 3, and sampling probabilities of 0.1, 0.1 and 0.005 in years 1, 4 and 5, respectively. Comments are in italics on the right**

```

2
SIMULATION DATA FOR PACIFIC WALRUS: SITE FIDELITY
  6      5
0.90 0.90
0.010 0.010
150000 150000
0.0451 0.0421 0.0301 0.0481 0.0561 0.2642
0.0050 0.0047 0.0033 0.0053 0.0838 0.4887
0.5225 0.0210 0.0150 0.0240 0.0281 0.1321 0.2572
0.0225 0.5210 0.0150 0.0240 0.0281 0.1321 0.2572
0.0225 0.0210 0.5150 0.0240 0.0281 0.1321 0.2572
0.0225 0.0210 0.0150 0.5240 0.0281 0.1321 0.2572
0.0225 0.0210 0.0150 0.0240 0.5281 0.1321 0.2572
0.0225 0.0210 0.0150 0.0240 0.0281 0.6321 0.2572
0.0451 0.0421 0.0301 0.0481 0.0561 0.2642 0.5144
0.5025 0.0023 0.0017 0.0027 0.0419 0.2443 0.2046
0.0025 0.5023 0.0017 0.0027 0.0419 0.2443 0.2046
0.0025 0.0023 0.5017 0.0027 0.0419 0.2443 0.2046
0.0025 0.0023 0.0017 0.5027 0.0419 0.2443 0.2046
0.0025 0.0023 0.0017 0.0027 0.5419 0.2443 0.2046
0.0025 0.0023 0.0017 0.0027 0.0419 0.7443 0.2046
0.0050 0.0047 0.0033 0.0053 0.0838 0.4887 0.4092
  0.05  0.0  0.01  0.0  0.01
  0.05  0.0  0.01  0.0  0.01
  0.05  0.0  0.01  0.0  0.01
  0.05  0.0  0.01  0.0  0.01
  0.05  0.0  0.01  0.0  0.01
  0.05  0.0  0.01  0.0  0.01
  0.05  0.0  0.01  0.0  0.01
  0.05  0.0  0.01  0.0  0.01
  0.05  0.0  0.01  0.0  0.01
  0.05  0.0  0.01  0.0  0.01
  0.05  0.0  0.01  0.0  0.01
  0.05  0.0  0.01  0.0  0.01
SIMULATION DATA FOR PACIFIC WALRUS: NO SITE FIDELITY
  6      5
0.90 0.90
0.010 0.010
150000 150000
0.0451 0.0421 0.0301 0.0481 0.0561 0.2642
0.0050 0.0047 0.0033 0.0053 0.0838 0.4887
0.0451 0.0421 0.0301 0.0481 0.0561 0.2642 0.5144
0.0451 0.0421 0.0301 0.0481 0.0561 0.2642 0.5144
0.0451 0.0421 0.0301 0.0481 0.0561 0.2642 0.5144
0.0451 0.0421 0.0301 0.0481 0.0561 0.2642 0.5144
0.0451 0.0421 0.0301 0.0481 0.0561 0.2642 0.5144
0.0451 0.0421 0.0301 0.0481 0.0561 0.2642 0.5144
0.0451 0.0421 0.0301 0.0481 0.0561 0.2642 0.5144
0.0050 0.0047 0.0033 0.0053 0.0838 0.4887 0.4092
0.0050 0.0047 0.0033 0.0053 0.0838 0.4887 0.4092
0.0050 0.0047 0.0033 0.0053 0.0838 0.4887 0.4092
0.0050 0.0047 0.0033 0.0053 0.0838 0.4887 0.4092
0.0050 0.0047 0.0033 0.0053 0.0838 0.4887 0.4092
0.0050 0.0047 0.0033 0.0053 0.0838 0.4887 0.4092
0.0050 0.0047 0.0033 0.0053 0.0838 0.4887 0.4092
  0.1  0.0  0.0  0.1  0.005
  0.1  0.0  0.0  0.1  0.005
  0.1  0.0  0.0  0.1  0.005
  0.1  0.0  0.0  0.1  0.005
  0.1  0.0  0.0  0.1  0.005

```

*Two scenarios are considered.  
Start of data for scenario 1  
The format is as in Table A1.*

*Start of data for scenario 2.*

0.1	0.0	0.0	0.1	0.005
0.1	0.0	0.0	0.1	0.005
0.1	0.0	0.0	0.1	0.005
0.1	0.0	0.0	0.1	0.005
0.1	0.0	0.0	0.1	0.005
0.1	0.0	0.0	0.1	0.005
0.1	0.0	0.0	0.1	0.005

## Appendix B: The Computer Program CMRSIM1

### Purpose of the Program

The Windows computer program CMRSIM1 operates in exactly the same way as the program CMRSIM (Appendix A) in terms of the way that movement, survival and sampling of the Pacific walrus population is carried out. The only difference is that for CMRSIM1 it is assumed that age information is available to assist with estimation. Specifically, it is assumed that when a walrus is sampled in the field it is possible to age it as 0 (under one year old), 1 (between one and two years old), 2 (between two and three years old), or 3+ (older than three years). As explained below, this information on age makes it possible to get much improved estimates of probabilities of capture based on the Manly and Parr (1968) method for analysing mark-recapture data. This then leads to improved estimates of population sizes and yearly survival rates.

The data format for program CMRSIM1 is exactly the same as it is for program CMRSIM, as shown in Table A1. The output from the program is also almost the same for the two programs, although the actual estimates will differ even for the same samples taken from the same populations.

In Manly and Parr's (1968) original paper it was suggested that, in order to estimate the probability of capturing an animal in a certain sample, a group of animals should be identified such that they were certainly in the population when the sample in question was taken. For example, one such group is those animals that were seen both before and after the sample time. The probability of capture is then estimated as the proportion of the animals that were certainly in the population that were captured at the sample time.

When animals cannot be aged and births and deaths are occurring, the only group of animals that can be identified as certainly in the population at the time of sample  $i$  are those seen in samples both before and after this time. However, if animals can be aged as 0, 1, 2 and 3+, as defined above, then the group of animals that were certainly alive at the time of sample  $i$  can include all of the following (assuming yearly samples):

- (a) those seen before and after year  $i$  (as is usually used);
- (b) any animal seen to be aged 1 in year  $i + 1$ ;
- (c) any animal seen to be aged 2 in year  $i + 1$  not already included under (a);
- (d) any animal seen to be aged 3+ in year  $i + 1$  not already included under (a);
- (e) any animal seen to be aged 2 in year  $i + 2$  not already included under (b);
- (f) any animal seen to be aged 3+ in year  $i + 2$  not already included under (a), (c) or (d);  
or

(g) any animal seen to be aged 3+ in year  $i + 3$  not already included under (a) to (f).

All of these individual animals must have been in the population in year  $i$ , and they are identified as such irrespective of whether or not they were captured in year  $i$ . The probability of capture in year  $i$  can therefore be estimated by the proportion of these animals that were captured at that time.

Having estimated probabilities of capture, the estimation of population sizes and survival rates can be carried out as explained in Appendix A using equations (A1) and (A2). There is just one important extra point to note. The individuals in categories (b) to (f) above can be identified for the first year of a mark-recapture study. Consequently, the probability of capture can be estimated in year 1, and hence the population size can be estimated. Indeed, the population size in year 1 can be estimated even when sampling only occurs for two years.

## Summary of Options

These are mostly the same as for the program CMRSIM. See Appendix A for a listing of these. However, with CMRSIM1 the age of animals is used in estimation. The initial age distribution is determined as follows:

! As with CMRSIM, the total population size is kept constant, with each death giving rise to the birth of a new animal. The expected age distribution is calculated for a population with this property, and the proportions of animals aged 0, 1, 2 and 3+ are calculated. The individuals in the initial population of animals are randomly assigned ages using these proportions as probabilities. Hence the initial age distribution of animals is the stable age distribution.

## Example

The example used to illustrate the input and output of program CMRSIM can also be used with CMRSIM1. Most of the input data are shown in Table A1. The full data for both programs include slightly more than this, with the following format:

Line 1: NSD, which is an integer (1 or more) indicating how many different simulation scenarios are being considered.  
Line 2 on: The input parameters for the first simulation scenario, in the format of Table A1.  
Next: If NSD = 1 then there is no more input. Otherwise the parameters for a second scenario are input in the format of Table A1. Further sets of parameters for more scenarios follow if NSD > 2.

When the program is started it describes the format needed for the data file in case this has been forgotten.

In brief, this example is for a situation where there are six sampling sites (as shown in Figure A1) plus another area where walrus can be without being available for sampling. Yearly survival probabilities are 0.9 for males and females, the probability of an animal being harvested and sampled is 0.01, the true population size is 150,000 males and 150,000 females, the proportions initially at different sites are based on the results of a 1990 survey (Table A2), animals move freely between sites with no tendency to return to the same site from year to year, and sampling continues for five years with a 0.05 probability of capture for all animals in all years (unless they are in the unsampled area).

Two populations were generated and sampled five times each. The choice of the number of populations and number of samples is made at run time. At this time the number of movement groups is also chosen (as explained in Appendix A). The choice made for this example was no movement groups. The output from the program is shown in Table B1.

**Table B1 Output from running CMRSIM1 with the data in Table A1, with comments in italics on the right-hand side.**

```
#####
#                               CMRSIM1                               #
#                               #                                     #
# A program to simulate mark-recapture sampling and                 #
# estimation for a population that occurs at up to 50               #
# sites, up to 20 sample times, and up to 300,000 males and       #
# 300,000 females sampled altogether. There is movement            #
# between sites between sample times. Males and females          #
# are treated separately for the simulation and may have           #
# different parameters. Harvesting data is used for               #
# improving estimation. It is assumed that samples are           #
# taken every year, and that it is possible to recognize          #
# animals aged 0, 1 or 2 years old, with others being seen       #
# as 3+ years old.                                               #
#                                                                     #
#                               Version 1.0 (December, 2000)       #
#                               Written by Bryan F.J. Manly        #
#                               Western EcoSystem Technology Inc.   #
#                               Cheyenne, Wyoming 82001           #
#                               bmanly@west-inc.com                #
#####

Date: 20/12/2000 Time: 11:16

Data: SIMULATION DATA FOR PACIFIC WALRUS
Number of Sites = 6 Number of Sample Times = 5
Male Survival Probability = 0.900 Female Survival Probability = 0.900
Male Harvesting Probability = 0.010 Female Harvesting Probability = 0.010
(Survival includes survival from harvesting)
Total Males = 150000 Total Females = 150000

Male Proportions Initially at Different Sites
(Last site is all unknown locations)
0.05 0.04 0.03 0.05 0.06 0.26 0.51

Female Proportions Initially at Different Sites
(Last site is all unknown locations)
```

*The first part of the output summarises the input parameters for the simulation*

0.00 0.00 0.00 0.01 0.08 0.49 0.41

Movement Probability Matrix For Males  
(Element in row i and column j is the probability of moving from site i to site j in one unit of time)

	1	2	3	4	5	6	7
1	0.05	0.04	0.03	0.05	0.06	0.26	0.51
2	0.05	0.04	0.03	0.05	0.06	0.26	0.51
3	0.05	0.04	0.03	0.05	0.06	0.26	0.51
4	0.05	0.04	0.03	0.05	0.06	0.26	0.51
5	0.05	0.04	0.03	0.05	0.06	0.26	0.51
6	0.05	0.04	0.03	0.05	0.06	0.26	0.51
7	0.05	0.04	0.03	0.05	0.06	0.26	0.51

Movement Probability Matrix For Females  
(Element in row i and column j is the probability of moving from site i to site j in one unit of time)

	1	2	3	4	5	6	7
1	0.00	0.00	0.00	0.01	0.08	0.49	0.41
2	0.00	0.00	0.00	0.01	0.08	0.49	0.41
3	0.00	0.00	0.00	0.01	0.08	0.49	0.41
4	0.00	0.00	0.00	0.01	0.08	0.49	0.41
5	0.00	0.00	0.00	0.01	0.08	0.49	0.41
6	0.00	0.00	0.00	0.01	0.08	0.49	0.41
7	0.00	0.00	0.00	0.01	0.08	0.49	0.41

Capture Probability Matrix for Males

Sample Time ==>

Site	1	2	3	4	5
1	0.05	0.05	0.05	0.05	0.05
2	0.05	0.05	0.05	0.05	0.05
3	0.05	0.05	0.05	0.05	0.05
4	0.05	0.05	0.05	0.05	0.05
5	0.05	0.05	0.05	0.05	0.05
6	0.05	0.05	0.05	0.05	0.05

Capture Probability Matrix for Females

Sample Time ==>

Site	1	2	3	4	5
1	0.05	0.05	0.05	0.05	0.05
2	0.05	0.05	0.05	0.05	0.05
3	0.05	0.05	0.05	0.05	0.05
4	0.05	0.05	0.05	0.05	0.05
5	0.05	0.05	0.05	0.05	0.05
6	0.05	0.05	0.05	0.05	0.05

Burn in Movements = 1

Number of Movement Groups (IGP) = 0

Number of Simulated Populations = 2

Number of samples per population = 5

Random number seed = 100

Summary of Generated Population

Numbers at Sites

Sample Time ==>

Site		1	2	3	4	5
1	M	6765	6658	6824	6683	6754
	F	749	755	724	801	772
2	M	6315	6366	6263	6233	6258
	F	705	640	661	748	683
3	M	4515	4465	4480	4564	4450
	F	494	455	466	487	477
4	M	7214	7251	7196	7201	7198
	F	794	759	845	790	824
5	M	8415	8482	8549	8499	8546
	F	12570	12202	12546	12544	12537
6	M	39630	39631	39755	39631	39766
	F	73305	73593	73251	73152	73302
7	M	77146	77147	76933	77189	77028

*These are the numbers at different sites for the first simulated population, with site 7 being unavailable for capture (at sea).*

F 61383 61596 61507 61478 61405

Sites Where Individual Male Animals Are and Ages of the Animals at Time 1  
(0 = dead or elsewhere, -1 = harvested)

```
      Sample Time ==>
Anim Age 1 2 3 4 5
  1  0  1  7  7  0  0
  2  0  1  0  0  0  0
  3  2  1  4  2  7  6
  4 21  1  7  7  7  7
  5  0  1  7  7  7  7
```

*Abbreviated output.*

.  
.

```
209935 -4 0 0 0 0 7
209936 -4 0 0 0 0 7
209937 -4 0 0 0 0 6
209938 -4 0 0 0 0 4
209939 -4 0 0 0 0 6
```

*Age -4 means not born for 4 years.*

Sites Where Individual Female Animals Are and Ages of the Animals at Time 1  
(0 = dead or elsewhere, -1 = harvested)

```
      Sample Time ==>
Anim Age 1 2 3 4 5
  1 17  1  6  7  7  6
  2  2  1  0  0  0  0
  3 18  1  6  6  6  6
  4 11  1  6  6  7  7
  5 21  1  5 -1  0  0
```

*Abbreviated output.*

.  
.

```
210177 -4 0 0 0 0 7
210178 -4 0 0 0 0 5
210179 -4 0 0 0 0 7
210180 -4 0 0 0 0 6
210181 -4 0 0 0 0 6
```

Sites Where Individual Male Animals Are Captured and Their Ages at Time 1

```
      Sample Time ==>
Anim Age 1 2 3 4 5
  1  0  0  0  0 -1  0
  2 14  1  0  0  0  0
  3 14  0  0  2  0  0
  4  1  0  0  3  0  0
  5 15  1  0  0  0  0
```

*Abbreviated output.*

.  
.

```
23348 -4 0 0 0 0 5
23349 -4 0 0 0 0 2
23350 -4 0 0 0 0 6
23351 -4 0 0 0 0 6
23352 -4 0 0 0 0 6
```

Sites Where Individual Female Animals Are Captured and Their Ages at Time 1

```
      Sample Time ==>
Anim Age 1 2 3 4 5
  1 21  1  0 -1  0  0
  2  1  0  0  0  6  0
  3  2  0  6  0  0  0
  4 19  1  0  0  0  0
  5  4  0  6  0  0  0
```

*Abbreviated output.*

.  
.

```
26729 -4 0 0 0 0 6
26730 -4 0 0 0 0 6
26731 -4 0 0 0 0 6
26732 -4 0 0 0 0 6
26733 -4 0 0 0 0 6
```

Estimated Numbers at Sites

*For the first set of mark-recapture samples at population 1.*



Site	Sample Time ==>				
	1	2	3	4	5
1 M	10229	14889	11789	15937	0
F	1975	1850	1003	2826	0
2 M	12786	11809	9015	21733	0
F	1317	1850	501	942	0
3 M	10229	9241	11789	7244	0
F	988	1480	501	0	0
4 M	21480	16943	13176	10142	0
F	1317	740	501	942	0
5 M	14320	16943	10402	18835	0
F	19425	20354	20056	15073	0
6 M	85920	84713	95702	86931	0
F	119844	126565	129364	131886	0
All M	154962	154538	151875	160822	0
SE	8426	8432	9830	14861	0
F	144867	152840	151928	151668	0
SE	7164	7633	8850	12917	0

Estimated Male Survival Rate = 0.911  
 Estimated Female Survival Rate = 0.905

*Summary of the results for sampling two populations five times each.*

Estimated Parameters From Simulations

Pop	Spl	Survival		Estimated Number of Males at Times									
		Male	Female	1,2, ...	Followed by	Female	Estimates						
1	1	0.911	0.905	154962	154538	151875	160822	144867	152840	151928	151668		
1	2	0.909	0.885	142605	156300	141931	159974	163255	156490	142080	161593		
1	3	0.848	0.905	154099	146731	157378	138462	152729	146720	156223	141783		
1	4	0.880	0.882	152591	145790	149938	145647	156686	143585	144792	138800		
1	5	0.900	0.927	147506	156789	147574	149011	150541	152954	164518	167396		
% Bias		-1.2	0.1	0.2	1.4	-0.2	0.5	2.4	0.3	1.3	1.5		
% CV		2.9	2.0	3.5	3.6	3.8	6.4	4.6	3.5	6.0	8.2		
2	1	0.895	0.899	138394	149140	145624	144545	164943	158453	151555	157118		
2	2	0.831	0.932	150378	162579	145077	129325	144942	151578	158399	155521		
2	3	0.906	0.900	147371	155972	139958	147384	150991	158314	146028	152202		
2	4	0.846	0.898	160304	151641	139132	139870	139932	152601	151849	133488		
2	5	0.933	0.888	151089	152194	164370	173977	165887	156853	158390	165031		
% Bias		-2.0	0.4	-0.3	2.9	-2.1	-2.0	2.2	3.7	2.2	1.8		
% CV		4.7	1.9	5.2	3.5	6.8	11.0	7.8	2.2	3.5	7.8		

*The % Bias is the average bias from the simulated estimates, as a percentage of the true parameter value. The % CV is the coefficient of variation (the standard deviation of estimates divided by the true parameter value) expressed as a percentage.*