

Review of “Evidence for high natality rates among Steller sea lions in the Gulf of Alaska”

Eli Holmes, National Marine Fisheries Service, eli.holmes@noaa.gov

April 23, 2009

1 Summary

The ASLC-0901 report analyzes mark-resight data on females visiting the Chiswell rookery in 2003-2008. The analysis purports to show evidence of high natality rates at an eastern CGOA rookery based on the mark-resight data. However, the analysis suffers from two technical problems:

- For an estimate of the average pups per female, the unit of replication is ‘female’. In the report, the unit of replication used is ‘year’ thus the analysis suffers from pseudo-replication and females with high fecundity are overly represented.
- The trailing blanks for each individual are ignored rather than being treated as ‘no pup; status unknown’. Thus all 0s (no pup years) by females not sighted after their last visit to the rookery are ignored.

These two technical problems lead to zeros (no pup data points) being systematically eliminated and high fecundity females being systematically over-represented. This leads to overestimation of the average natality rates across mature females. Below I show how the data might be reanalyzed so that the estimates do not suffer from the trailing blank problem and so that they are unbiased in the face of variation in female fecundity. Using this method and assuming an adult survival for a stable population (ca. 0.86), the point estimate of the probability of pupping per female is 0.52. This is similar to the estimate obtained by the modeling work in Holmes et al. (2007).

2 The analysis

The data are coming from a rookery and thus females that give birth are much more likely to appear in the data set. In addition, when females stop appearing in the dataset, it is unknown whether they died or simply did not give birth and thus were unlikely to revisit the rookery. This is an advantage to some degree since we have a high confidence that we are not missing births. In the dataset, 97% of the births occurred at Chiswell.

Consider a sample of n females that arrive the first time at the rookery in year 2003 and then are recorded from 2004 to 2008 to as a 0, 1, 1*, 0*, or blank (meaning no pup either because it died or didn’t give birth again). *We cannot count the pup status in year 2003 because females show up at the rookery in order to pup, so the first year is skewed towards 1s (pup).* Also the first year might be the female’s first year of maturity and she can’t give birth that year. For this 2003 cohort, the expected number of pups 2004 to 2008

is:

$$\Pi = \sum_{t=2003}^{2008} \bar{p} \times (t - 2003) \times P(died(t))$$

\bar{p} is the mean pupping probability per female. $(t - 2003)$ is the number of breeding seasons the females survived past 2003 if she died in year t . $P(died(t))$ is the probability that the female died in year t after the breeding season, i.e. between breeding season t and $t + 1$. Basically, we take the expected number of pups if a female lived only 0 breeding seasons past 2003 \times the probability of living 0 breeding seasons past 2003, plus the expected number of pups if the female lived only 1 breeding season past 2003 \times the probability of living 1 breeding season past 2003, etc, up to the number of breeding seasons observed past 2003 (in this case 5). It is ok if there is variability across females in p because the number of pups in a sample of n depends only on the average p (\bar{p}) across females. (technically the expected value of Π only depends on \bar{p}).

To get the estimate of \bar{p} for females first sighted in 2003, we find the \bar{p} that minimizes the difference between the observed number of pups (both on Chiswell and off) in the sample versus the predicted number (Π) for a given \bar{p} .

To get the probability that the female died in year t , we need an estimate of survival for reproductive females. The apparent survival in the data set is ca. 0.8 which would imply a rapidly declining population. However because the data were taken from a rookery, females that do not reproduce stop appearing thus it is not surprising that the apparent survival is low. We might think of this as the females ‘emigrating’ to the non-reproductive portion of the population. The estimate of adult survival in the 1970s (stable population) was ca. 0.86 and the estimate for the early 2000s time period from Holmes et al. (2007) is 0.92. Once we have an estimate of survival, $P(died(t))$ can be calculated with the geometric probability distribution:

$$P(died(t)) = (1 - sa)^{(t-2003)} \times sa$$

where sa is the adult survival and $(1 - sa)$ is then the probability of success (death) for the geometric distribution.

For each cohort year, 2003 to 2007, we repeat the analysis to get the estimate of \bar{p} for each year and then take the mean (weighted by cohort sample size) to get the overall mean \bar{p} for the entire data set.

3 Results

Table 1: The estimate of mean pupping probability across cohorts. Note that the year 2003 estimate is inflated upwards because it contains some females who visited Chiswell before year 2003. The adult survival used in the calculation is sa . $sa = 0.8$ is added for completeness but this would imply a declining population. $sa = 0.86$ is the 1970s estimate for a stable population and $sa = 0.92$ is the Holmes et al. (2007) estimate for this time period. The observed number of pups includes pups off the rookery (1* in the dataset).

	sa=0.8	sa=0.86	sa=0.92	sample size	obs. num. pups
2003	0.74	0.63	0.59	58	115
2004	0.31	0.27	0.26	26	19
2005	0.60	0.54	0.53	24	28
2006	0.54	0.49	0.50	26	20
2007	0.59	0.56	0.58	30	14
overall mean	0.59	0.52	0.51	164	196

Table 1 shows the estimates of \bar{p} for each cohort of females that first arrived at Chiswell during the breeding season at year t . The estimates for year t in the table are not the estimate of pupping probability for year

t , but the average pupping probability across the cohort. Only the data after the first year of sighting can be used (otherwise the estimates would be biased because the cohort is defined by showing up at Chiswell and they probably showed up to give birth). Thus the cohort 2003 estimate uses the data on this cohort from 2004 to 2008. The 2008 cohort is missing because there is no data after 2008.

Note that year 2003 includes some females that first arrived before year 2003. That is, 2003 was not their first year at the rookery. Females that first visited Chiswell before 2003 should be excluded to provide an unbiased estimate. Because I did not know which females had not been observed before, I included all females. This skews the 2003 estimate upward. The reason is that if you just look at the females present year t , those females are skewed toward the fecund females because fecund females spend more breeding seasons at the rookery. This is a type of sampling bias called the ‘waiting-time bias’. This is not a problem for 2004 to 2008 because, they have been observing Chiswell since 1999 so we can be fairly sure if a female had never been seen 1999 to 2003, then the first year it showed up between 2004 and 2008 was its first year at Chiswell rookery during the breeding season.

4 Comments

- The first year of sighting at Chiswell must be excluded because females show up to give birth. The females sighted year t becomes the sample and the pupping after year t is the data.
- Eliminating all data before the first birth (as done in the report) severely biases the estimates because it preferentially excludes low fecundity females. Eliminating the first year of sighting (as done here) deals with the issue of females not giving birth their first year of maturity.
- Females who are not sighted again are counted as 0 pup. The sampling seems rigorous enough to exclude the possibility that they had a pup at Chiswell but were somehow missed.
- The analysis assumes a closed population. If they are emigrating to another rookery, then there is not much to be done with the data unless that other rookery is also observed.
- Females that are infertile will be largely excluded from the data set. Females that happen to die before giving birth (after their first year of maturity) will also be largely excluded. Thus the estimates are skewed upward to some unknown degree.

5 The ASLC correction for bias toward females that pup

In equation 1, the authors give a correction term for the bias towards females that pup. The authors do not give the logic behind this equation, and I was unable to determine how it was derived. In the text, the correction reduces to

(all observations on Chiswell with pup) divided by (all observations on or off Chiswell minus all observations off Chiswell with pup)

However, equation 1 as described in the text is not what produced the numbers in Table 1a. The numbers in Table 1a appear to be the following (based on back-engineering the numbers):

(all observations with a pup on or off Chiswell) divided by (all observations on Chiswell + all observations off Chiswell + all observations off Chiswell without pup)

I could not determine how this correction would be derived either nor how it would deal with the trailing blanks or pseudo-replication problems.

6 Other criticisms of Holmes et al 2007 by the report authors

The authors make claims that the estimates in Holmes et al. are biased because a too low neonatal mortality was used. This is s_0 in the Holmes et al. paper. The s_0 value drops out when comparing P(pup at 1-month) values across years because s_0 appears in both the nominator and denominator: $P(\text{pup at 1-month}) = P(\text{pup at birth}) \times s_0$. If neonatal mortality has increased substantially then, yes our natality (probability of giving birth) declines are pessimistic (and we mentioned that in the paper). However, we (Holmes et al. authors) have no evidence or reason to believe this has occurred and the authors of ASLC-0901 provide no evidence either. In fact, the tone of this report implies that the authors feel that female condition is superb suggesting that if anything s_0 is higher not lower than before.

The authors state that the sightability estimates in Holmes et al. for nursing females are severely underestimated because the estimates do not adjust for nocturnal feeding. Most of the sightability estimates used in Holmes et al. were 24 hour estimates (because that was all that was available), however one estimate for nursing females was for daylight hours only (in CA) and those estimates were not particularly different than the 24 hour estimates for nursing females. The authors present new sightability estimates. They state these are based on presence/absence data, which suggest it means 'observed at least once between 10am and 6pm'. There is no indication in the report that these estimates are from rookery hourly scan data (such as in Maniscalco 2006). If it is just presence/absence (and not hourly or half-hourly scan data) then the sightability numbers are not (number of hours on land)/(number daylight hours) but rather (number of days observed at least once)/(number of days). We need the fraction of hours on land not fraction of days animal is observed at least once on land. The sightability ranges they suggest, 80-90%, are also incompatible with the numerous ca 60% 24-hour sightability estimates when considered with the information in Millete (1999) [citation in Holmes et al. 2007] which shows data on departure times from Sugarloaf and Lowrie (SE AK). Millete's study found 43-28% of departures in summer happened during daylight hours on Sugarloaf and 54-43% happened during daylight hours on Lowrie. So yes, Steller sea lions have a nocturnal tendency, but an estimate of 80-90% of daylight hours on land would imply that virtually all departures happen at night – which is not supported by field data.