Revised Methods for Estimating Abundance of the Eastern North Pacific Stock of Gray Whales

JEFFREY M. BREIWICK[1], ANDRÉ E. PUNT[2], DAVID J. RUGH[1], JEFFREY L. LAAKE[1] AND RODERICK C. HOBBS[1]

Contact e-mail: jeff.breiwick@noaa.gov

1 – National Marine Mammal Laboratory, 7600 Sand Point Way NE, Bldg. 4, Seattle, WA 98115, USA
2 - School of Aquatic and Fishery Sciences, Box 355020, University of Washington, Seattle, WA 98195-5020, USA

ABSTRACT
Counts of migrating whales at Yankee Point and Granite Canyon, California, form the basis for estimation of trends and abundance for the eastern North Pacific stock of gray whales. Data from these surveys have been collected and analysed by methods that have become more sophisticated over time. We outline an approach and work plan which, given the available data, will lead to a set of abundance estimates that are suitable for trend analysis based on a more consistent approach.

KEYWORDS: GRAY WHALES; NORTH PACIFIC; ABUNDANCE ESTIMATION; TRENDS

INTRODUCTION

Abundance estimates for the eastern North Pacific (ENP) gray whale (*Eschrichtius robustus*) have been made for 23 of the 42 years since 1967/68 using data from shore-based counts of the southbound migration going past Yankee Point or Granite Canyon, near Monterey, California. These shore-based counts have usually begun about 10 December of one year and ended in February or early March of the next year (hence the notation 1967/68 for the respective abundance year). The gray whale migration follows a straight path parallel and close to shore in the Granite Canyon area such that shore-based observers have a consistent view and can see across most of the migratory corridor (Shelden and Laake 2002). The research station at this site (owned by NOAA, but operated by the State of California Department of Fish and Game) also makes it a convenient location from which to conduct this study. The routine nature of these shore-based whale counts lend them to inter-annual trend analyses.

The dates of each census, the number of watch periods per day, type of distance data collected and whether or not independent, paired sightings took place are shown in Table 1. Over the course of time, improvements in the quality of data have occurred. For example, prior to 1985/86, observers made non-calibrated estimates of distance to a sighting using predetermined distance intervals (e.g., ¼, ½, ¾ mile). However, since then, observers have used reticle marks in binoculars to obtain fairly precise distance measurements. Paired, independent sightings (which allow for the estimation of the probability of pods being missed by an observer) have taken place since 1985/86. Prior to that date, each observer worked alone without any tests of sighting rates. Throughout the years, observers recorded wind direction and force. Sky condition was noted prior to 1978, but visibility has been recorded instead of sky condition since then. In the analysis, visibility has been found to be the most useful environmental variable. In all analyses, only sightings made at visibility codes less than or equal to 4 have been used (except 1992/93 and 1993/94, when visibility codes less or equal to 3 were used).

Abundance estimates have generally followed the same methodology throughout the 40-year period (1967 to 2007): the observed number of pods (under acceptable visibility conditions), *m*, was multiplied by correction factors for: (1) pods passing outside watch periods - f_t; (2) night travel rate - f_n; (3) pods missed during watch periods - f_m; and (4) bias in pod size estimation, f_s. The product of *m* and the above correction factors was then multiplied by the mean pod size, $\overline{s}$, to yield an abundance estimate:

$$\hat{N} = m\overline{s}f_t\, f_n\, f_m\, f_s \quad (1)$$

Data from 1967/68 – 1979/80 were originally analysed by Reilly *et al.* (1980, 1983) and reanalysed by Buckland and Breiwick (2002). Data from subsequent surveys have not been analysed in a way totally consistent with Reilly *et al*. (1980, 1983). For example, since 1995/96 the correction factor, f_t, for pods passing outside the watch period has been based, not on the number of pods per hour, as previously, but on the estimated number of whales (i.e. sum of pods after correction for pod size bias and missed pods) per hour. This was done to account for differential sightability by pod size and the covariance within the estimated number of whales sighted when corrections are applied to individual sightings of pods (Hobbs *et al.* 2004). Thus, since 1995/96 the Hermite polynomial correction factor, f_t, has been based on the number of animals rather than the number of pods. Three of the correction factors, f_n, f_m, and f_s were not estimated in each survey. Corrections for night travel rates, f_n, were established through thermal imagery in tests conducted from 1994 to 1996 (Perryman *et al*. 1999); a correction factor for missed pods, f_m, has been made since 1987/88 when paired, independent counts were done, providing mark-recapture (sight-resight) data; and pod size bias corrections, f_s, were done when an aircraft was available (Laake *et al*., 1994).

The following section summarizes the previous basis for estimating abundance for the ENP gray whales in more detail. Key citations for each abundance year or series of years are given in square brackets.

1967/68 – 1987/88 [Buckland and Breiwick, 2002]
From 1967/68 to 1973/74, gray whale counts were conducted from a site at Yankee Point, approximately 5km north of the Granite Canyon research station, which was used during all subsequent years. There were two watch periods per day of 5h each. One observer would be on the morning watch and another on the afternoon watch. Prior to 1978/79 no visibility codes were recorded; instead data showed sky conditions or sometimes noted miles as an indication of visibility. Starting in 1978/79, visibility was documented according to six subjective categories ranging from 1 for excellent to 6 for useless. In the analysis, watch periods for which visibility code exceeded 4 at any time were discarded (as per IWC protocol; see Buckland *et al*., 1993). Distance to sightings were based on estimated intervals (0-1/4 nm, 1/4-1/2 nm, etc.) prior to 1985. However, non-calibrated estimates of distance are not considered reliable. Therefore, distance estimates have been based on binoculars with reticles since 1985. No adjustments for biased pod size estimates or pods missed during watch periods were made. Hermite polynomial models were fitted to the unadjusted pod frequency counts. The relative abundance estimates for 1967/78–1985/86 were rescaled so that the 1987/88 estimate (see next section) passed through the ‘best’ estimate from Buckland *et al.* (1993), in which sighting heterogeneity was modelled. Thus, the indices of abundance were multiplied by 20869/15954 (1.3081). Independent observer

data have been collected since 1985/86 to estimate the proportion of missed pods. Abundance estimates were further corrected for the night travel rate based on radio-tagged whales which showed that gray whales travel slightly faster at night than during the day (Swartz *et al.* 1987).

1987/88 [Buckland *et al.* 1993]
Paired, independent count data were modelled using iterative logistic regression to allow for the effects of covariates on detection probability. None of the tested factors (observer, watch period, station [North or South]) were significant. The detection probability increased with pod size, with the rate of passage and with migration date, and decreased with an increase in visibility code (lower detection in worsening visibility).

1992/93 – 1993/94 [Laake *et al.*, 1994]
The methodology applied by Laake *et al*. (1994) was much the same as Buckland *et al.* (1993). The night travel rate correction factor (f_n) was assumed to be 1.02 (SE = 0.023), the same as used by Buckland *et al.* (1993). New correction factors for missed pods (f_m) and bias in recorded pod size (f_s) were provided. Missed pod rates were estimated from additional data on paired, independent observers, and pod size bias was estimated by using aircraft circling over whale pods to accurately establish pod size.

1995/96 [Hobbs *et al.*, 2004]
Hobbs *et al.* (2004) changed methodology so that the abundance was estimated based on the number of whales passing during watch periods ($\hat{W}$) multiplied by a Hermite polynomial correction factor for whales missed during watch periods (f_t,) based on the number of whales rather than pods, and the same correction factor for night travel (f_n), was used as by Buckland *et al.* (1993) and Laake *et al.* (1994). Recorded pod sizes were converted to estimated number of whales accounting for differential sightability by pod size, and the covariance was estimated within the estimated number of whales sighted when corrections were applied to individual sightings of pods. The number of pods of each size passing during watch periods was estimated by multiplying the recorded pod size by a correction for bias in estimating pod size and by the missed pod correction factor.

1997/98, 2000/01, 2001/02 [Rugh *et al.*, 2005] and **2006/07** [Rugh *et al.*, 2008]
The methodology of Hobbs *et al.* (2004) was used, including the night travel rate correction factor based on Perryman *et al.* (1999), instead of the correction used previously (from Swartz *et al*. 1987). This night correction factor was also applied to all previous abundance estimates 1967/68 – 1995/96.

The current paper outlines a work plan which will be used to develop a time series of abundance estimates for the ENP gray whales by applying a more uniform approach for all years of sightings data from 1967/68 to 2006/07. This approach will be better suited to trend analysis.

WORKPLAN FOR UPDATING THE ESTIMATES OF ABUNDANCE

1. The database for 23 years of shore-based counts will be reviewed and reformatted. The data are currently in four different formats due to available programs and variations in data collection procedures over time. A common set of data fields will be developed and the flat files will be reworked to fit this format. Careful documentation and development of metadata will result in a standalone dataset that will be provided to the IWC with a uniform data

analysis program (Gray Whale Abundance Estimation Program, ERAbund, Vers. 2.43, 22 April 2003).

2. Pod size bias correction factors will be revised to reconcile the available tests of observers' pod size estimates: via aircraft (Reilly, 1981; Laake *et al*., 1994), video imagery (Perryman *et al*. 1999), tracking teams (Rugh *et al*., 2008) and paired pod size estimates from the standard watch observer data. An approach for compiling these calibration results is outlined in Appendix A.
3. Corrections for pods missed during watch periods, f_m, have been established using paired, independent observational effort conducted for several weeks during each of the studies since 1985/86. To date, most analyses (e.g. Hobbs et al. 2004) have applied the data from the respective field season. An analysis including all of the double count data from all years will be conducted to determine a generalized correction method, and this method will be applied to all of the years with data collected at Granite Canyon including the years when there was no paired observational effort. Options for extending this analysis to include the data collected at Yankee Point will be explored. It is thought that f_m will decline as sighting rates increase, so density will be included as a covariate in the analysis.
4. Corrections for pods missed during non-effort periods, f_t, using Hermite polynomials, were based on pods per hour before 1987/88, but since then the corrections have been applied to whales per hour. Accordingly, Hermite polynomials will be applied to whales (pods corrected by pod size bias) per hour for all available years of sighting data.
5. An empirical approach will be used to scale the earlier estimates to the most recent estimates (sensu Buckland *et al*. 1993) if it is not possible to reconcile the differences in data collection methods. One way to achieve this would be to construct separate time-series for 1967/68-1985/86 and 1987/88-2006/07, and use the assessment method to estimate a calibration factor.
6. The increased data requirements placed on a single observer as in recent years will be examined. Where possible data from both observers will be used to maximize the data value. These data will also be analysed as a single observer data set to determine the risk of bias compared to a single observer. The two sets of results will be compared to determine which minimizes the risk of bias and increased variance.

ACKNOWLEDGEMENTS
The National Marine Mammal Laboratory (NMML) of the Alaska Fisheries Science Center, NMFS, NOAA, provided support for this study. We are grateful to the scores of observers who have collected sighting observations over the years, making this project possible. Paul Wade, Devin Johnson and Nancy Friday (NMML/NOAA) are thanked for their comments on a draft of this paper.

REFERENCES

Buckland, S.T. and Breiwick, J.M. 2002. Estimated trends in abundance of eastern Pacific gray whales from shore counts, 1967/68 to 1995/96. *J. Cetacean Res. Manage*. 4(1):41-8.

Buckland, S.T., Breiwick, J.M., Cattanach, K.L. and Laake, J.L. 1993. Estimated population size of the California gray whale. *Mar. Mamm. Sci*. 9(3):235-49.

Hobbs, R.C., Rugh, D.J., Waite, J.M., Breiwick, J.M. and DeMaster, D.P. 2004. Abundance of North Pacific gray whales on the 1995/96 southbound migration. *J. Cetacean Res. Manage.* 6(2):115-20.

Laake, J.L., Rugh, D.J., Lerczak, J.A. and Buckland, S.T. 1994. Preliminary estimates of population size of gray whales from the 1992/93 and 1993/94 shore-based surveys. Paper SC/46/AS7, presented to the IWC Scientific Committee, May 1994 (unpublished). 16 p.

Perryman, W.L., Donahue, M.A., Laake, J.L. and Martin, T.E. 1999. Diel variation in migration rates of eastern Pacific fray whales measured with thermal imaging sensors. *Mar. Mammal Sci.* 15(2):426-45.

Reilly, S.B. 1981. Population assessment and dynamics of the California gray whale (*Eschrichtius robustus*). Unpubl. Ph.D. dissertation, Univ. Wash., Seattle.

Reilly, S.B., Rice, D.W. and Wolman, A.A. 1980. Preliminary population estimate for the California gray whale based upon Monterey shore censuses, 1967-68 to 1978-79. *Rep. int. Whal. Commn* 30:359-68.

Reilly, S.B., Rice, D.W. and Wolman, A.A. 1983. Population assessment of the gray whales, *Eschrichtius robustus*, from California shore censuses, 1967-80. *Fish. Bull.* 8(2):267-81.

Rugh, D.J., Hobbs, R.C., Lerczak, J.A. and Breiwick, J.M. 2005. Estimates of abundance of the eastern North Pacific stock of gray whales (*Eschrichtius robustus*) 1997-2002. *J. Cetacean Res. & Manage.* 7(1):1-12.

Rugh, D., Breiwick, J., Muto, M., Hobbs, R., Shelden, K., D'Vincent, C., Laursen, I.M., Reif, S., Maher, S. and Nilson, S. 2008. Report of the 2006-2007 census of the eastern North Pacific stock of gray whales. AFSC Processed Rep. 2008-03, 157 p., Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.

Rugh, D.J., Muto, M.M., Hobbs, R.C. and Lerczak, J.A. 2008. An assessment of shore-based counts of gray whales. *Marine Mammal Science* 24(4):864-80.

Shelden, K.E.W. and Laake, J.L. 2002. Comparison of the offshore distribution of southbound migrating gray whales from aerial survey data collected off Granite Canyon, California, 1979-96. *J. Cetacean Res. Manage.* 4(1):53-6.

Swartz, S.L., Jones, M.L., Goodyear, J., Withrow, D.E. and Miller, R.V. 1987. Radio-telemetric studies of gray whale migration along the California coast: a preliminary comparison of day and night migration rates. *Rep. int. Whal. Commn* 37:295-9.

Table 1. Gray whale shore-based count locations, dates and field methods.

Year	Location[1]	Start date	End date	Watch periods per day[2]	Paired obs.	Distance data[3]	Visibility
1967/68	YP	18-Dec-1967	03-Feb-1968	two – 5h ea.	-	Intervals	Sky/dist
1968/69	YP	10-Dec-1968	06-Feb-1969	two – 5h ea.	-	Intervals	Sky/dist
1969/70	YP	08-Dec-1969	08-Feb-1970	two – 5h ea.	-	Intervals	Sky/dist
1970/71	YP	09-Dec-1970	12-Feb-1971	two – 5h ea.	-	Intervals	Sky/dist
1971/72	YP	18-Dec-1971	07-Feb-1972	two – 5h ea.	-	Intervals	Sky/dist
1972/73	YP	16-Dec-1972	16-Feb-1973	two – 5h ea.	-	Intervals	Sky/dist
1973/74	YP	14-Dec-1973	08-Feb-1974	two – 5h ea.	-	Intervals	Sky/dist
1974/75	GC	10-Dec-1974	07-Feb-1975	two – 5h ea.	-	Intervals	Sky/dist
1975/76	GC	10-Dec-1975	03-Feb-1976	two – 5h ea.	-	Intervals	Sky/dist
1976/77	GC	10-Dec-1976	06-Feb-1977	two – 5h ea.	-	Intervals	Sky/dist
1977/78	GC	10-Dec-1977	05-Feb-1978	two – 5h ea.	-	Intervals	Sky/dist
1978/79	GC	10-Dec-1978	08-Feb-1979	two – 5h ea.	-	Intervals	Vis codes
1979/80	GC	10-Dec-1979	06-Feb-1980	two – 5h ea.	-	Intervals	Vis codes
1984/85	GC	27-Dec-1984	31-Jan-1985	two – 5h ea.	-	Intervals	Vis codes
1985/86	GC	10-Dec-1985	07-Feb-1986	three – 3 or 3.5h ea	✓	Intervals/reticles	Vis codes
1987/88	GC	10-Dec-1987	07-Feb-1988	three – 3 or 3.5h ea	✓	Reticles	Vis codes
1992/93	GC	10-Dec-1992	07-Feb-1993	three – 3 or 3.5h ea	✓	Reticles	Vis codes
1993/94	GC	10-Dec-1993	18-Feb-1994	three – 3 or 3.5h ea	✓	Reticles	Vis codes
1995/96	GC	13-Dec-1995	23-Feb-1996	three – 3h ea.	✓	Reticles	Vis codes
1997/98	GC	13-Dec-1997	24-Feb-1998	three – 3h ea.	✓	Reticles	Vis codes
2000/01	GC	13-Dec-2000	05-Mar-2001	three – 3h ea.	✓	Reticles	Vis codes
2001/02	GC	12-Dec-2001	05-Mar-2002	three – 3h ea.	✓	Reticles	Vis codes
2006/07	GC	12-Dec-2006	22-Feb-2007	three – 3h ea.	✓	Reticles	Vis codes

[1] Location: 1967/68 – 1973/74 censuses were conducted at a private residence in Yankee Point (YP), about 5km north of Granite Canyon. From 1974/75 onwards, censuses have been conducted at Granite Canyon (GC). Other notable dates: 1984/85 census was conducted from a house trailer on the north side of the GC waste pool. 1985/86 first reticled binoculars were used on standard watch, and paired, independent watches were started. 1987/88 was the first year in which wood sheds and Big Eye (25-power) binoculars were used.

[2] 1967/68 – 1984/85: two watch periods per day of 5h each, from 0700–1700 hrs
1985/86 - 1987/88: three watch periods per day, 0700-1030 hrs, 1030–1330 hrs, 1330–1700 hrs
1992/93: three watch periods, 0800-1100 hrs, 1100-1300 hrs, 1300-1600 hrs
1993/94: three watch periods, 0730-1030 hrs, 1030-1330 hrs, 1330-1700 hrs
1995/96 – 2006/07: three watch periods, 0730-1030 hrs, 1030-1330 hrs, 1330-1630 hrs

[3] Intervals were 0 - ¼ nmi, ¼ - ½ nmi, ¾ - 1 nm, 1-1.5 nmi, 1.5-2 nmi, etc. Distances have been based on binocular reticles since 1985/86.

No visibility codes were recorded prior to 1997/98. Instead observers recorded sky conditions and sometimes miles as an indication of visibility. In 1997/98 observers began recording visibility in six subjective categories (1 = excellent; 6 = useless), a system used ever since.

Table 2. Correction factors applied to counts of gray whales.

Year	*Night travel rate* f_n^*	*Missed pods* f_m	*Pod size bias* f_s
1967/68	A	B	F
1968/69	A	B	F
1969/70	A	B	F
1970/71	A	B	F
1971/72	A	B	F
1972/73	A	B	F
1973/74	A	B	F
1974/75	A	B	F
1975/76	A	B	F
1976/77	A	B	F
1977/78	A	B	F
1978/79	A	B	F
1979/80	A	B	F
1984/85	A	B	F
1985/86	A	B	F
1987/88	A	B	F
1992/93	A	C	G
1993/94	A	D	H
1995/96	A	E	I
1997/98	A	E	I
2000/01	A	E	I
2001/02	A	E	I
2006/07	A	E	I

Since 1995/96, the frequency of the estimated number of animals (based on correcting pods by missed pods and pod size bias), rather than pods, was modelled by a normal distribution with Hermite polynomials added to adjust for skewness, kurtosis and higher moments (Buckland *et al.*, 1993).

A = All abundance estimates currently apply the same night travel rate correction factor, based on Perryman *et al.* (1999) using data from January 1994, 1995 and 1996 (total sample size was 116h by day; 146h by night). $f_n^* = 1.0875\ (se = 0.0363)$ Only the latter half of the migration is corrected, i.e., the correction factor is applied to sighting rates only after the median migration date.
B = 1.063 is a constant multiplier to estimate pods missed by observers during their watch; this was based on 1987/88 data (see Buckland *et al.* 1993).
C = 1.193 was the correction factor for missed pods applied in 1992/93.

D = 1.16 was the correction factor for missed pods applied in 1993/94.
E = separate missed pod size correction factors by pod size were developed for each year, based on the sighting data.
F = a correction factor of 1.131 (based on data from 1987/88) was used as a scale factor for pod size bias.
G = a correction factor of 1.43 was used for pod size bias in 1992/93.
H = a correction factor of 1.42 was used for pod size bias in 1993/94.
I = the same values from Laake *et al.* (1994) were used (based on data from aerial counts and shore counts during 7-18 Jan 1993 and 5-18 Jan. 1994); variances and covariances were estimated via bootstrap methodology (see Hobbs *et al.* 2004).

Note that in the proposed analysis, all years will be treated with uniform correction factors, varying only as a function of sighting density.

Appendix A

Estimating pod size bias

Let $P^j_{a,b}$ be the probability that observer *i* estimates a pod of true size *a* to be of size *b*. Assuming that this probability is based on a Poisson function, i.e.:

$$P^i_{a,b} = {\mu^i_a}^b e^{-\mu^i_a} / b! \tag{App.1}$$

where μ^i_a is the expected estimated pod size for a pod of actual size *a* by observer *i* (of *I* observers). The Poisson function was tentatively chosen, but alternative forms will be explored and compared. Separate parameter values will be estimated for μ^i_1 and μ^i_2, while μ^i_{3+} will be estimated assuming a linear relationship between μ^i_a and true pod size. This particular parameterization was selected after considering a variety of alternative formulations.

Parameters for equation App.1 will be estimated using the estimated pod size for a set of *I* observers for a set of *J* pods (not all observers provided estimates of pod size for each of the *J* pods). "True" pod sizes were established through calibration efforts by having an aircraft circle a pod while shore-based observers continued their standard counting protocol (Laake *et al*., 1994) or by continual tracking of a set of pods from land by separate observers (Rugh *et al*., 2008). The aircraft provides an excellent view of a whale pod, although there could have been confusion between how to define a pod from the aerial view relative to what shore-based observers saw.

The data can be summarized as $D^i_{j,b}$ (observer *i* recorded the j^{th} pod (of true size a_j) to be *b*) and hence the likelihood maximized is:

$$L \propto \prod_j \prod_i (P^i_{a_j,b})^{D^i_{j,b}} \tag{App.2}$$

The aim of the estimation scheme is to obtain a population-level measure of the probability that a (random) observer estimates a pod of true size *a* to be of size *b*. This could be estimated by imposing a hierarchical structure on the values for μ^i_a, e.g. that μ^i_a is normally distributed, i.e.: $\mu^i_a \sim N(\mu_a, \tau_a)$. However, preliminary analyses showed that the sizes of the random effects were small. In future analyses, there could be an accounting for random effects due to each observed pod.

The observed pod frequencies for a given year, N_b, need to be corrected for pod size estimation errors. In principle, N_b is an outcome of a mixture distribution, i.e. $E(N_b) = \sum_a P_{a,b} Q_a$ where Q_a is the true pod size distribution. In principle, Q_a can be calculated by convolution. However, that can lead to negative estimates for Q_a. Therefore, rather than inverting the *P* matrix, the observed pod frequencies will be treated as a multinomial sample from an underlying distribution and the following likelihood will be maximized to find the estimates for the Q_a :

$$L \propto \prod_b (\sum_a P_{a,b} Q_a)^{N_b} \quad \text{(App.3)}$$

This approach treats the outcome from the first estimation step as known so the standard errors for the mean pod size only account for the error associated with fitting Equation App.3, although sensitivity analyses will explore whether it is possible to estimate pod size bias estimates by combining the estimation of the pod size error matrix and that of the true underling pod size distribution. The true pod size distribution will be assumed to be a two-parameter gamma distribution to avoid an over-parameterized model. This selection is justified given that preliminary analyses suggest that fits are only marginally better when more complicated models (including the saturated model) are selected. The pod size bias model (Equation App.1) will be applied to data from the various experiments separately to assess whether there are significant differences.