Comparisons of measured and estimated distances and angles from sightings surveys

RUSSELL LEAPER*, LOUISE BURT+, DOUGLAS GILLESPIE# AND KELLY MACLEOD^

Contact e-mail: russell@ivyt.demon.co.uk

ABSTRACT

Photogrammetric systems using video cameras were used to measure radial distances to sightings during the SCANS-II, CODA and SOWER surveys. These surveys included sightings of a variety of species from harbour porpoise, at distances of a few hundred metres, to large baleen whales at distances greater than 10km. A total of 910 initial sightings with estimated distances from reticles and measured distances from video, using 7 × 50 (636) or 25× 'Big Eye' (274) binoculars, were compared. Bearings to sightings were also measured from still images. The CV_{RMSE} in distances varied between 0.19 and 0.33 for reticle binoculars. Comparisons of measured distances to simultaneous sightings by other observers using naked eye gave a CV_{RMSE} of 0.39 for naked eye estimates. There was a consistent, non-linear pattern in all data sets, of over-estimating close distances to sightings of surfacing cetaceans and under-estimating those further away. However, this pattern was not evident from the distance experiments on SOWER to fixed targets which also had a much lower variance (CV_{RMSE} = 0.13). Bearing data from SCANS-II and CODA showed around 5% of estimates had gross errors greater than 20º that were attributed to mistakes. For the remaining values, RMS errors were in the range 5.7º–7.2º for SCANS-II and CODA and 4.9º for SOWER. Both distance and angle errors will make a substantial contribution to the variance of abundance estimates and simulated data showed that the observed non-linear nature of distance errors may cause considerable bias even when linear regressions might suggest little bias. There still remain technological challenges in operating complex electronic systems at sea to measure distances and bearings, but investment in these methods should be a cost effective way of reducing bias and improving precision of cetacean abundance estimates.

KEY WORDS: SURVEY–VESSEL; PHOTOGRAMMETRY

INTRODUCTION

Distances and angles to sightings during line-transect surveys are critical data items but often rely on estimates from observers that may be subject to considerable error. These errors are a widely acknowledged problem for cetacean abundance estimation (Williams *et al.*, 2007). Photogrammetric methods have been used for some time to measure distances and angles to cetacean sightings and have been incorporated into the data collection system on recent surveys (Gillespie *et al.*, 2010). On the SCANS-II[1] (Small Cetaceans in the European Atlantic and North Sea) and CODA[2] (Cetacean Offshore Distribution and Abundance) surveys in the Northeast Atlantic in 2005 and 2007, photogrammetric systems were part of a fully integrated, computer-based data collection system. On the IWC SOWER (Southern Ocean Whale and Ecosystem Research) surveys in 2006/07 and 2007/08, the use of video cameras to measure distances and digital still cameras to measure angles was limited to experimental periods.

The implications of measurement error for bias and precision in abundance estimates have been examined for theoretical models, showing the potential for severe bias in the case of both large unbiased measurement error and biased errors (Marques, 2007). Distance and angle experiments to artificial visual targets such as buoys are also conducted during many surveys to assess the variance and, sometimes, to try to correct for distance errors. However, the extent to which such experiments are representative of the real situation for cetacean sightings is difficult to assess. Most methods to correct for distance errors have also relied on either additive models (e.g. Chen, 1998; Chen and Cowling, 2001) or linear multipliers (e.g. Marques, 2004). Such models may not always be appropriate for correcting distance errors. For example, Alldredge *et al.* (2007) reported non-linearities in distance errors to calling birds and suggested the need for more complex error correction methods. The aim of this paper is to compare measured and estimated values to the sightings made during surveys and examine the implications of measurement error for abundance estimates.

SURVEY METHODS

The integrated data collection system used on the SCANS-II and CODA surveys, described in Gillespie *et al.* (2010) included photogrammetric measurement of distances and angles to sightings using the methods of Leaper and Gordon (2001). Observers on the surveys consisted of two 'Primary' observers searching with naked eye and two 'Tracker' observers, one searching with 7 × 50 binoculars and one with 25× 'Big Eyes' (Monk Leviathan) to implement Mark Recapture Distance Sampling methods (Buckland and Turnock, 1992). Measurements from digital video sequences

* *International Fund for Animal Welfare, 87-90 Albert Embankment, London, SE1 7UD, UK.*
+ *RUWPA, University of St Andrews, The Observatory, Buchanan Gardens, St Andrews, Fife, KY18 9LZ, UK.*
# *Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews, Fife KY16 8LB, UK.*
^ *SMRU Ltd., Scottish Oceans Institute, New Technology Centre, North Haugh, St Andrews, Fife. KY16 9SR, UK.*
[1] *http://biology.st-andrews.ac.uk/scans2/*
[2] *http://biology.st-andrews.ac.uk/coda/*

of cetacean surfacings were used to calculate distances and digital still images were used to calculate angles for the observers using binoculars. The general principle behind the use of video cameras to measure distances at sea is the same as with using reticle binoculars and involves measuring the angle of dip from the horizon to the whale from a platform of known height. Eye heights on SCANS-II and CODA vessels varied between 10 and 14m. One of the main challenges to the system is capturing an image of the first surfacing reported by the observer of sufficient quality to allow measurements to be made. Photogrammetric measurement of bearings used a downward pointing camera taking a still image of reference marks on the deck of the vessel. These methods can only be used for observers searching with binoculars but some estimates of errors from naked eye observations were available from the SCANS-II survey where photogrammetric measurements from the Tracker could be compared with naked eye estimates from the Primary for the same surfacing event. Estimated angles were obtained using angle boards for the Primary observers and using angle pointers attached to the binocular mounts for the Trackers. The 7 × 50 binoculars were supported on a monopod with the angle pointer at the bottom, the Big Eyes were on a fixed pedestal with an angle scale just below the binoculars.

A subset of the full system described in Gillespie *et al*. (2010) was used during experimental periods of the IWC SOWER surveys in the Southern Ocean. The video system was used for observers in the top observation barrel on SOWER cruises in 2006/07 and 2007/08 from an eye height of 20.5m. Observers on SOWER use 7 × 50 binoculars with a non-linear reticle scale which is marked in nautical miles: angles are measured using angle boards (see Fig. A1 in Appendix). On the SOWER 2007/08 cruise, the video system was also used during a distance estimation experiment where distances were also measured to a buoy in the water using radar. This experiment served as an additional calibration check. Observer distance estimation errors during standard distance experiments were also compared to the errors to whale sightings.

ANALYSIS METHODS

Calibration tests of the photogrammetric systems are described in Leaper and Gordon (2001) and indicated sufficiently small errors (root mean square error in distance of 3.5% and in angle of 1.5°) that measured values were treated as 'true' values when compared with visual estimates for the analyses in this paper. We only used initial sightings, except for analysis of simultaneous sightings by naked eye observers during SCANS-II, to avoid autocorrelation, because distance estimation errors to re-sightings may be strongly influenced by the initial estimates.

Unlike the computer controlled system for capturing angle images on SCANS-II and CODA, which used webcams, the still camera system on the SOWER surveys needed to be completely self-contained. This system is described in Appendix I and followed similar experiments on a previous cruise in the same series of surveys in 1983/84 (Thompson and Hiby, 1985). In addition to using the still camera to measure angles to sightings with images captured when the observer pressed a button, images were also captured at intervals to examine observer scanning patterns. On SCANS-II and CODA, images were captured at random intervals with a mean interval of 30s. On SOWER, the interval was fixed at 30s because the camera did not support random intervals but the variation in the observer's scanning patterns would effectively generate a random sample.

Following the surveys, pairs of simultaneous distances and angles were compared. These are referred to as 'measured' for values derived from the photogrammetric system and 'estimated' for naked eye estimates, reticle and angleboard readings by the observers. For distances, the errors are likely to scale with the distance and a convenient measure is the CV of the root mean squared error (CV_{RMSE}) defined as the root mean squared error divided by the mean of the observed values. For angles, the root mean squared (RMS) error is more appropriate.

For analysis of naked eye estimates, distances and angles to reported surfacings from the Primary and Tracker from SCANS-II that occurred close together in time and location but were not necessarily classed as duplicate sightings, were compared. Sightings had to occur within 10 seconds (the Tracker sighting did not necessarily need to occur first as with usual duplicate sightings) and on a similar bearing (±10°).

It was anticipated that patterns of errors in distance estimates would be complex and non-linear (Williams *et al.*, 2007). In addition to simple linear regression of estimated distance against measured, non-linear effects were investigated by plotting log(estimated)–log(measured) against log(measured). Investigative analyses were also carried out using Generalized Additive Models (GAMs).

Simulation study of the effects of measurement error on estimated strip widths

For non-linear errors, simple simulations of the detection process were used to investigate some of the effect of measurement error on estimated strip width. Errors will affect both the accuracy and precision of estimates with the effect on precision being strongly influenced by the number of data points. To investigate bias, 10,000 simulated sightings were generated with and without distance error. Software DISTANCE (Thomas *et al.*, 2010; Thomas *et al.*, 2006) was used to fit detection functions to these two data sets so that estimated strip widths could be compared.

To simulate sightings, a fixed vessel speed of $5ms^{-1}$ was assumed with whales distributed randomly within a box ahead of the vessel. The probability P, that a whale surfacing at a particular location was detected, was modelled by the hazard probability function:

$$P(r,\theta) = \frac{e^z}{1 + e^z} \quad (1)$$

Where r is the radial distance, θ is the angle from the trackline and

$$z = a_0 + a_1 r + a_2 r^3 \quad (2)$$

Where a_0, a_1 and a_2 are parameters of the detection function

The choice of functional form for the hazard probability function was based on sightings from surveys of minke whales (Cooke and Leaper, 1998). The aim of the simulations was to investigate general implications of distance error rather than specific results for any particular survey. Thus parameters were not species specific and a pattern of a dive time of 120s followed by 3 surfacings was assumed in all cases with whales travelling in a straight line with a speed of $0.5ms^{-1}$. Whales were introduced into the box according to the method of Hiby (1982) in order to ensure the correct distribution of whale headings. Values of a_1 and a_2 were adjusted to create detection functions with different effective strip widths. The inclusion of whale movement and multiple surfacings was designed to ensure that simulated data without error did not fit perfectly to a simple parametric detection function but were a more realistic representation of real data, even though the parameters themselves were not conditioned to any actual data. These simulations did not include a term in θ or the implications of angle error, but the data on search patterns using binoculars gathered by the photogrammetric systems on these surveys did allow the detection probability by angle to be estimated.

For each simulated sighting that occurred with position (r,θ), the position with distance error (r_e, θ) was generated by calculating r_e from r using the regressions derived from the data for the survey and observation method being investigated.

RESULTS

Performance of video systems

Gillespie *et al.* (2010) describe the performance of the video systems on the SCANS-II and CODA surveys. On SCANS-II the majority of sightings were of harbour porpoise (*Phocoena phocoena*) and distances were successfully measured on video from 448 (37%) of 1,211 sightings. The CODA survey had a greater variety of species, including large whales, and 405 (48%) of 843 sightings were measured. The combined success rate for two of the three CODA vessels analysed was 67% whereas the third vessel suffered a total failure of the video recording system. The higher success rate on CODA was likely due to a combination of larger, more visible species and the use of high definition video.

On the 2006/07 SOWER cruise, seven minke whale (*Balaenoptera bonaerensis*) surfacings were measured on video out of a total of 21 sequences that were recorded (33%). The main reason for sightings not being detectable on video appeared to be related to image quality and the characteristics of minke whale blows. The maximum distance that a minke body was detected on video was 3.6km and the maximum distance that a blow was detected was 1.9km. Even at this distance, this sighting was only detected for certain due to being a combined blow/body cue. On the 2007/08 cruise, experiments were mainly conducted in the presence of large baleen whales using a high definition video camera. In this case, 34 video measurements were obtained out of a possible 64 sequences (53%). Large baleen whale blows were detected out to measured ranges of 10km.

The experiment on the SOWER cruise that compared distances to a buoy between radar and the video gave the linear regression *Video = 1.03 × Radar* with $CV_{RMSE} = 0.05$ assuming radar measurements had no error. However, it is not known which of radar or video is more accurate and the small bias of 3% apparent in the video could be explained by refraction effects (which would affect reticle binoculars in the same way). These results were consistent with the calibration tests reported in Leaper and Gordon (2001).

Angle measurements were obtained for 94% of sightings on SCANS-II and 85% on CODA. The lower success rate on CODA was due to conflicts between USB devices connected to the computer which caused the webcams to stop working periodically.

Comparison of estimated and measured distances

After a first comparison of the estimated and measured data, the 90th percentile of largest distance errors were re-examined for errors due to data recording, transcribing or measurement mistakes.

Plots of estimated against measured radial distances are shown by survey and binocular type in Figs 1a–e together with linear regressions (regression coefficients are given in Table 1). Rounding to certain reticle values, indicated by points in a horizontal line, is particularly apparent for the

Table 1

Comparison of estimated and measured radial distances.

Survey	Searching method	n	Linear regression slope m (with intercept forced to 0)	CV_{RMSE}	Regression on log of distance: a, standard error in ()	Regression on log of distance: b, standard error in ()
SCANS-II	7 × 50	245	0.93	0.31	–0.13 (0.03)***	0.96 (0.18)***
SCANS-II, 5% of furthest estimated distances (>3.4km) truncated	7 × 50	233	0.97	0.36	–0.15 (0.03)***	1.05 (0.21)***
CODA	7 × 50	321	0.83	0.32	–0.08 (0.02)**	0.44 (0.14)**
CODA, 5% of furthest estimated distances (>6.5km) truncated	7 × 50	305	0.78	0.35	–0.12 (0.02)***	0.70 (0.15)***
SCANS-II	Big Eye	136	1.07	0.33	–0.20 (0.04)***	1.61 (0.31)***
SCANS-II, 5% of furthest estimated distances (>4km) truncated	Big Eye	129	1.06	0.36	–0.25 (0.05)***	1.94 (0.35)***
CODA	Big Eye	138	0.97	0.19	–0.18 (0.03)***	1.45 (0.21)***
CODA, 5% of furthest estimated distances (>8km) truncated	Big Eye	131	0.94	0.21	–0.20 (0.03)***	1.62 (0.22)***
Combined SCANS-II and CODA	Big Eye	274	0.98	0.23	–0.18 (0.02)***	1.42 (0.16)***
SOWER	7 × 50	41	0.86	0.26	–0.22 (0.07)**	1.74 (0.59)**
SOWER truncated at estimated distances >6km (10 values removed)	7 × 50	31	0.92	0.23	–0.21 (0.10)*	1.65 (0.82)
SCANS-II	Naked eye	28	0.81	0.39	–0.36 (0.17)*	2.00 (1.0)
SCANS-II truncated at estimated distances >600m (3 values removed)	Naked eye	25	0.74	0.44	–0.53 (0.18)**	2.95 (1.1)*

***Regression significant at p<0.001. **Regression significant at p<0.01 *Regression significant at p<0.05. n is the number of observations.
m is the slope term *estimated* = m × *measured*. a and b are the slope and intercept in ln (*estimated*)-ln(*measured*) = a × ln(*measured*) + b.

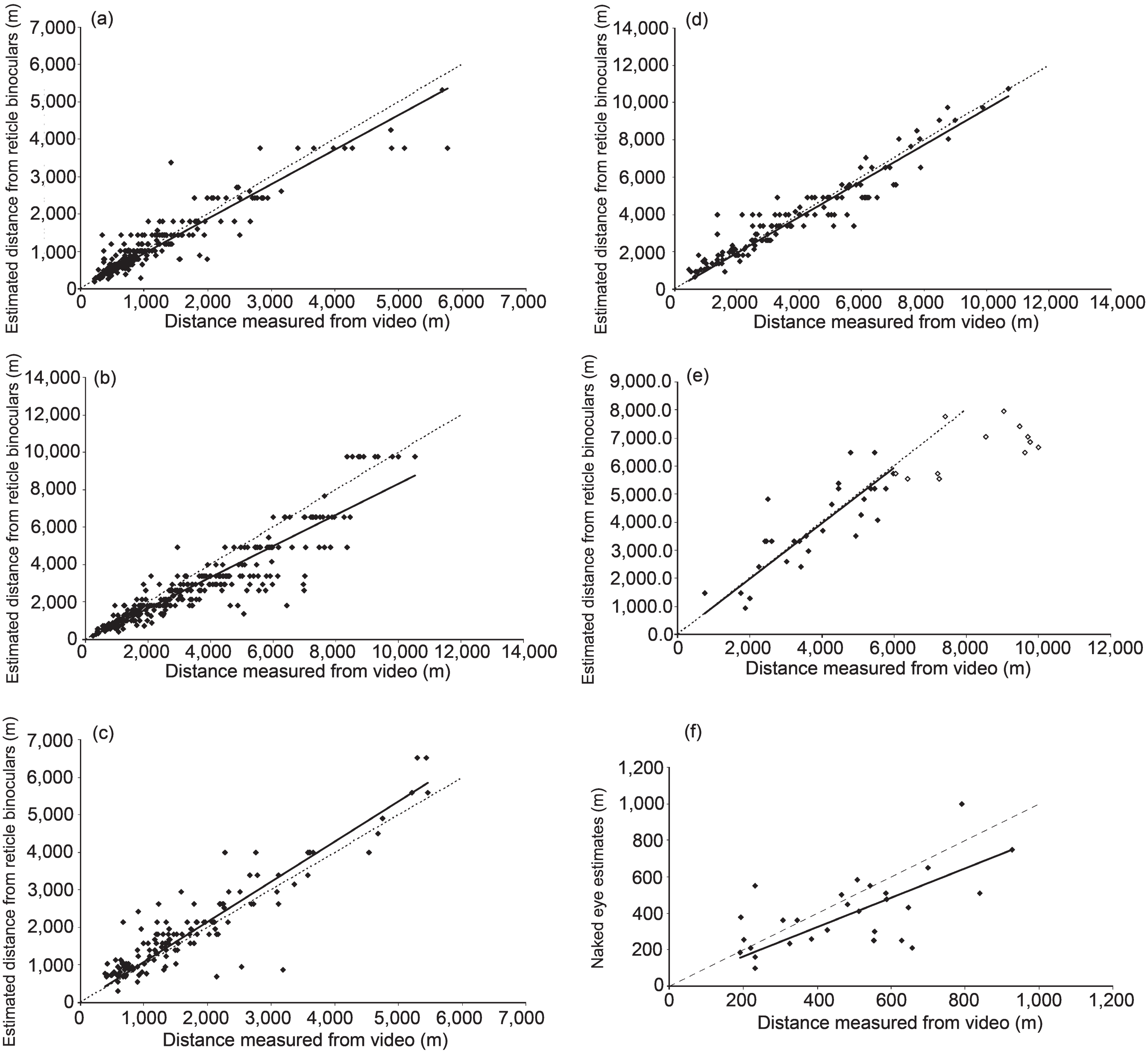


Fig. 1a. Estimated against measured distances from SCANS-II for 7 × 50 binoculars. Dotted line indicates no error, solid line indicates fitted linear regression.

Fig. 1b. Estimated against measured distances from CODA for 7 × 50 binoculars. Dotted line indicates no error, solid line indicates fitted linear regression.

Fig. 1c. Estimated against measured distances from SCANS-II for Big Eye binoculars. Dotted line indicates no error, solid line indicates fitted linear regression.

Fig. 1d. Estimated against measured distances from CODA for Big Eye binoculars. Dotted line indicates no error, solid line indicates fitted linear regression.

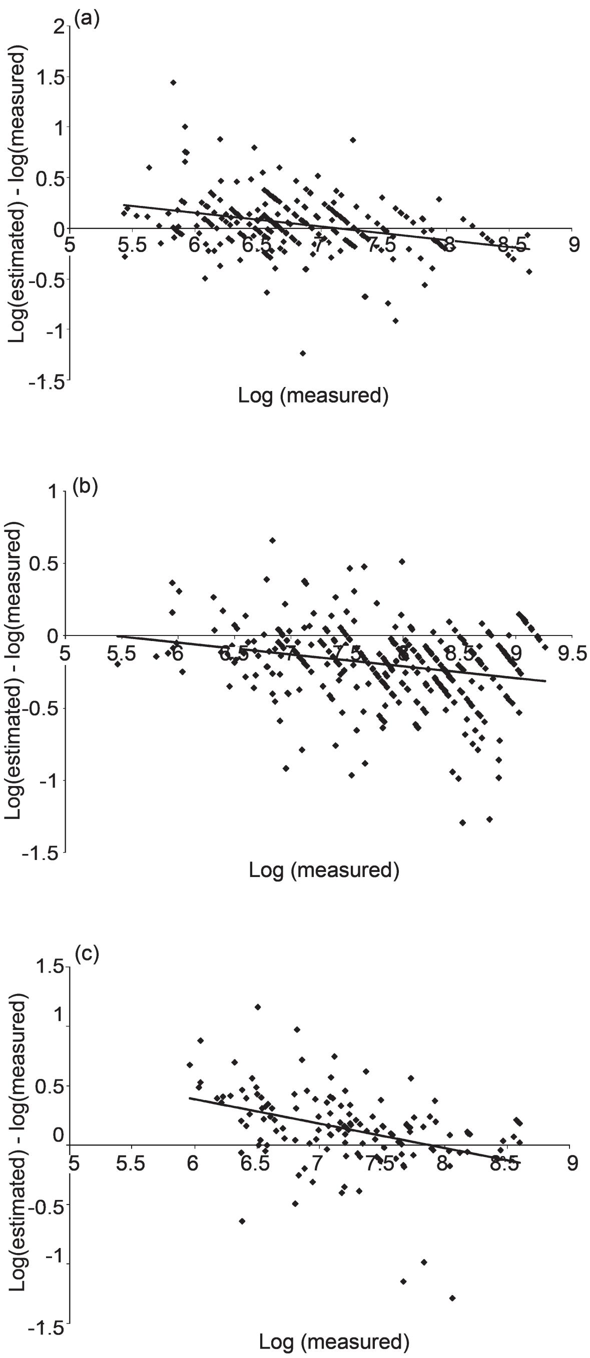
Fig. 1e. Estimated against measured distances from SOWER cruises in 2006/07 and 2007/08. Solid circles represent measured distances <6km, open circles>6km. Dotted line indicates no error, solid line indicates fitted linear regression up to a truncation of measured distances of 6km.

Fig. 1f. Estimated distances from naked eye against measured distances from Tracker from SCANS-II. Dotted line indicates no error, solid line indicates fitted linear regression.

SCANS-II and CODA 7 × 50 binoculars, particularly at larger distances (Figs 1a and b). Less rounding is apparent with the Big Eyes which have a finer reticle scale. The 7 × 50 estimates on CODA and the full data set from SOWER (including distances out to 10km) were the only ones that showed overall bias of greater than 10%. For SOWER, this can be explained by the difficulties of using reticles to estimate very small angles of dip for whales close to the horizon. When distances were truncated at measured values >6km, the bias was negligible (Fig. 1e). However, truncating on the basis of estimated distances >6km did less to reduce the bias (Table 1). The reason for bias in the CODA 7 × 50 binoculars is unclear but these were a different model to those used on SCANS-II whereas the same Big Eyes were used in both surveys.

The magnitude of the errors is indicated by the CV_{RMSE} given in Table 1. These varied between 0.19 for the CODA Big Eyes to 0.33 for the SCANS-II Big Eyes. The CV_{RMSE} of the CODA Big Eye sightings was strongly influenced by a single observer who accounted for 36% of all the sightings and had an individual CV_{RMSE} of 0.09. For the SOWER data, the CV_{RMSE} for the data truncated at 6km (approximately the maximum distance used in the buoy experiments) was 0.23, considerably greater than the CV_{RMSE} of 0.13 from all observers in the buoy experiments on the 2007/08 cruise.

In all cases for sightings of surfacing cetaceans, there was



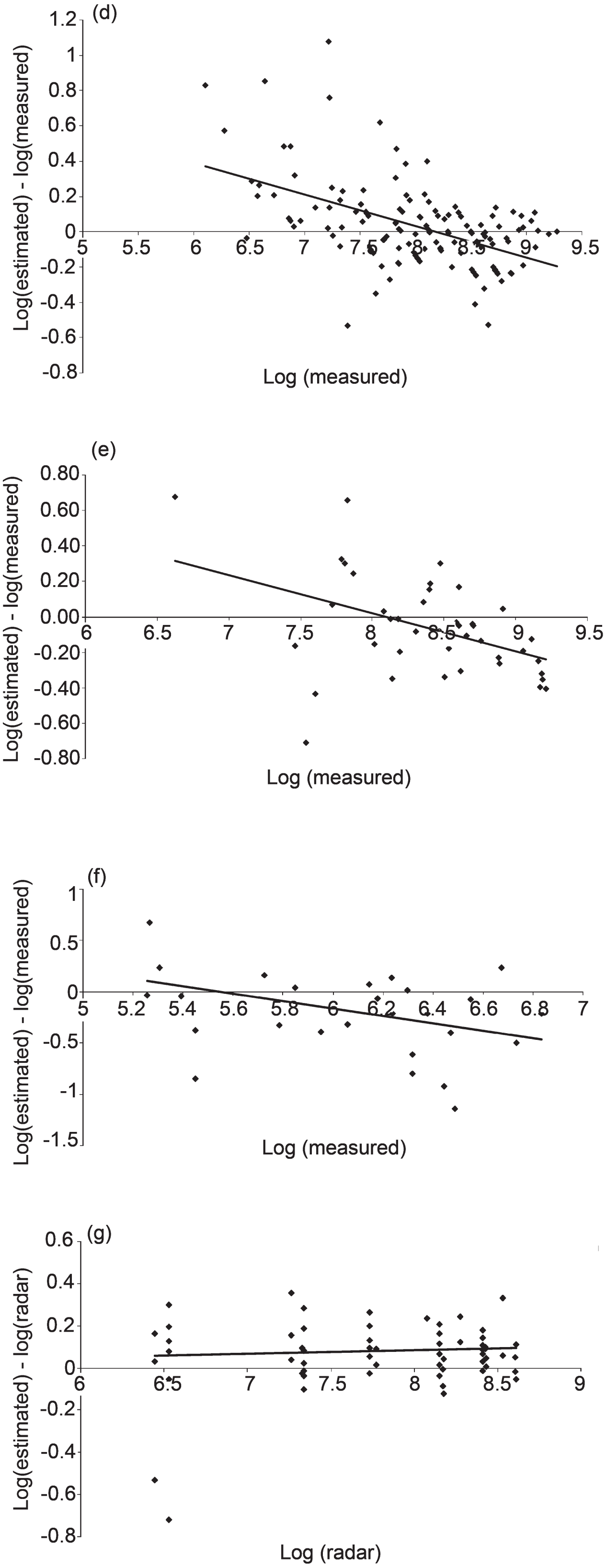


Fig. 2a. Log(estimated)–log(measured) against log(measured) distances from SCANS-II. Measured values are distances from video in metres, estimates are from 7 × 50 reticle binoculars.

Fig. 2b. Log(estimated)–log(measured) against log(measured) distances from CODA. Measured values are distances from video in metres, estimates are from 7 × 50 reticle binoculars.

Fig. 2c. Log(estimated)–log(measured) against log(measured) distances from SCANS-II. Measured values are distances from video in metres, estimates are from Big Eye reticle binoculars.

Fig. 2d. Log(estimated)–log(measured) against log(measured) distances from CODA. Measured values are distances from video in metres, estimates are from Big Eye reticle binoculars.

Fig. 2e. Log(estimated)–log(measured) against log(measured) distances from SOWER cruises in 2006/07 and 2007/08. Measured values are distances from video in metres, estimates are from reticle.

Fig. 2f. Log(naked eye estimate)–log(measured) against log(measured) from simultaneous sightings during SCANS-II. Measured distances are from Tracker platform.

Fig. 2g. Log(estimated)–log(measured) against log(measured) distances from the SOWER distance experiments in 2007/08. Measured values are from radar and estimates from reticle binoculars. Solid line shows linear regression which was not significant.

evidence of a non-linear relationship between error in distance and distance, with over-estimation of close distances and under-estimation of far distances. Figs 2(a–f) shows plots of log(estimated)–log(measured) against log(measured) with regression coefficients in Table 1. The slope of these regressions was significantly different from 0 at $p<0.05$ in all cases, indicating a change in distance bias with distance. By contrast, there was no evidence of a similar pattern in the errors to the fixed buoy in the distance experiments (Fig. 2g). Visual examination of the residuals from each of the regressions in Fig. 2 indicated a uniform spread, suggesting an adequate model. Exploratory investigations with GAMs suggested complex models with 5–8 degrees of freedom. The log based models were chosen for consistency between surveys and simplicity.

For 7 × 50 binoculars, the angle of dip from the horizon to the whale at which distances changed from over to under-estimation was approximately 0.26° for SOWER and 0.37° for SCANS-II (when binocular magnification is taken into account these would result in angles of 1.82 and 2.59° subtended at the eye). For the 25× Big Eyes, these angles were 0.08° and 0.13° (2.00 and 3.25° subtended at the eye). For naked eye on SCANS-II, this angle of dip was 2.6°. These indicate a fairly consistent angle of dip between the horizon and the whale, perceived at the eye, at which distance bias changes from positive to negative.

The effect of truncation at larger radial distances was also investigated for the 7 × 50 and Big Eye data from SCANS-II and CODA. Following the 'rule of thumb' suggested by Buckland *et al*. (1993), the largest 5% of estimated distances were truncated. Unlike for SOWER where there was clear evidence of increasing bias for large distances, truncation of the SCANS-II and CODA data did not generally reduce bias. In addition, the slope of the regression of log(estimated)–log(measured) against log(measured) and the CV_{RMSE} increased in all cases (Table 1). There is also some selectivity in the dataset resulting in both the closest and furthest distances being less likely to be measured from video. For SCANS-II and CODA the closest distance measured on video was 230m and 390m for 7 × 50 and Big Eyes respectively. The effects of truncation were most apparent in the naked eye data from SCANS-II. If 30% of the furthest estimated radial distances (>500m) were truncated then there was no longer a significant correlation between estimated and measured distances ($r = 0.42$, df = 18, $p>0.05$).

Comparison of estimated and measured angles

Where large discrepancies between estimated and measured angles were observed, these were resolved wherever possible by listening to the commentaries and re-analysing the bearing images. Bearing images were taken in sequences, one second apart, and so it was possible to measure whether the observer was looking steadily at a target, or still scanning when the sighting button was pressed. For the 7 × 50 binoculars this resulted in 651 initial sightings where both estimated and measured bearings were available from SCANS-II. Of these, 5% (34 sightings) showed gross errors of more than 20° which could not be resolved and were assumed to be either observer error or related to angle pointers becoming mis-aligned. For the remaining sightings, the RMS error was 7.1° for SCANS-II and 7.2° for CODA. For the Big Eyes there were 355 sightings with both estimated and measured bearings of which 6% of sightings showing errors of more than 20°. Excluding these sightings with large errors gave a RMS error of 6.0° for SCANS-II and 5.7° for CODA. For the simultaneous sightings from naked eye observers during SCANS-II where there was also a measured angle from the Tracker platform, the RMS error was 5.9°. However, this value may be influenced by the selection criteria used for simultaneous sightings; angles needed to be within +/– 10° and hence, sightings with larger angle errors were eliminated.

On the SOWER 2008/09 cruise there were a total of 62 sightings where bearings were both estimated from angle boards and measured photographically. There was evidence of a small systematic bias of around 2° and an overall RMS error of 4.9°. Of the 62 sightings, 45 (73%) were humpback whales (*Megaptera novaeangliae*), nine (15%) were sperm whales (*Physeter macrocephalus*) and five (8%) were southern bottlenose whales (*Hyperoodon planifrons*). There were no significant differences in mean squared error between these species (Anova, df = 2, $p = 0.88$). There were only four sightings where the cue was not recorded as a blow or blow/body and so it was not possible to investigate the accuracy of bearings with respect to cue type.

Perpendicular distance is proportional to the sine of the angle, so this was used to investigate potential bias in perpendicular distance due to angle error. Fig. 3 shows sin(estimated) against sin(measured) for the SCANS-II Tracker angles. The linear regression is given by $y = 1.01x$ showing no evidence of overall bias.

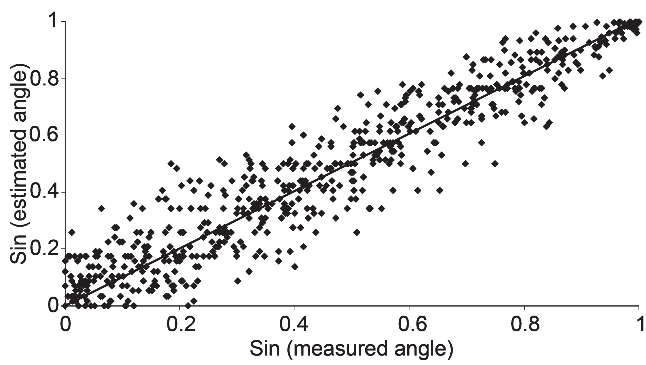


Fig. 3.Angle component of perpendicular distance, sin(measured angle) against sin(estimated angle). Data from SCANS-II survey.

Effect of measurement error on estimated strip widths

The estimated strip widths from simulated data with and without error are given in Table 2. The intercept and slope from the regressions of log(estimated)–log(measured) against log(measured) in Table 1 were used to generate a distance with error (r_e) from the distance a simulated whale was detected (r). It can be seen that there is scope for substantial bias, although the extent of the bias depends on the distribution of observed radial distances in relation to the distance at which distance errors tend from over-estimation to under-estimation. These results should be treated as illustrative of the level of bias that may occur based on the distance error relationships estimated for each survey rather than actual estimates of potential bias for these surveys. The parameters of the detection function were adjusted to

Table 2

Estimated strip widths from simulations with and without measurement error. Parameters for the simulations were adjusted to generate two different estimated strip widths (ESW) for each model to investigate the effects of measurement error under different sighting conditions.

	No measurement error		With measurement error		
Source of measurement error model	ESW (m)	CV	ESW (m)	CV	$\frac{ESW_{error}}{ESW_{noerror}}$
SOWER (all data)	850	0.016	1,229	0.020	1.45
SOWER (all data)	1,386	0.017	1,795	0.021	1.30
SCANS-II 7x50	585	0.016	680	0.018	1.16
SCANS-II 7x50	391	0.013	493	0.015	1.26
Combined SCANS-II/CODA; Big Eye	857	0.016	1,174	0.020	1.37
Combined SCANS-II/CODA; Big Eye	1,956	0.019	2,276	0.026	1.16
SCANS-II naked eye	377	0.016	388	0.021	1.03
SCANS-II naked eye	628	0.016	506	0.024	0.81

Strip widths calculated from simulated data (10,000 sightings) in Distance selecting half-normal key with cosine adjustments based on AIC.

generate different effective strip widths rather than fitted to the data themselves.

The simulations assumed an equal probability of detection for all angles between 0° and 90° (and zero for greater angles). No data were available for angular search effort from naked eye observers, but the angular search effort using binoculars is shown in Fig. 4 for combined 7 × 50 and Big Eyes (CODA survey). The function to describe searching effort by angle fitted to these data by least squares is given in equation 3. If detection probabilities in the simulations were multiplied by the fitted effort function in equation 3, this would reduce estimated strip widths to approximately 50% of what they would be assuming uniform search effort

$$y = \frac{\pi}{2\int^{\pi/2} \cos(x)^{6.9}\, dx} \cos(\theta)^{6.9} \quad (3)$$

DISCUSSION

All the datasets of distances to sightings of surfacing cetaceans showed a consistent pattern of over-estimation of small radial distances and under-estimation of larger ones. This could be a result of rounding effects at small reticle readings if observers tend to round up the reticle reading, and difficulties in counting reticles at larger reticle readings. The same pattern was also apparent in the naked eye estimates but for naked eye this could also be explained by the high variance of the estimates. Williams *et al.* (2007) reported a similar error pattern from an observer using 7 × 50 binoculars from a platform height of 18.3m. In that case, the angle of dip at which errors changed from over-estimates to under-estimates was 0.25º, or 1.8º subtended at the eye. The consistency of the angle subtended at the eye (1.8º–3.25º) may provide some insight into the visual processes involved in distance estimation using reticles and corresponds roughly to the angle of foveal (high acuity) vision. This could be investigated further by specific experiments involving different magnification binoculars and different observation heights. However, such experiments would need to involve real sightings targets because the results indicate that distance experiments to fixed targets do not show the same patterns of distance errors.

The implications of the compression of the range of true

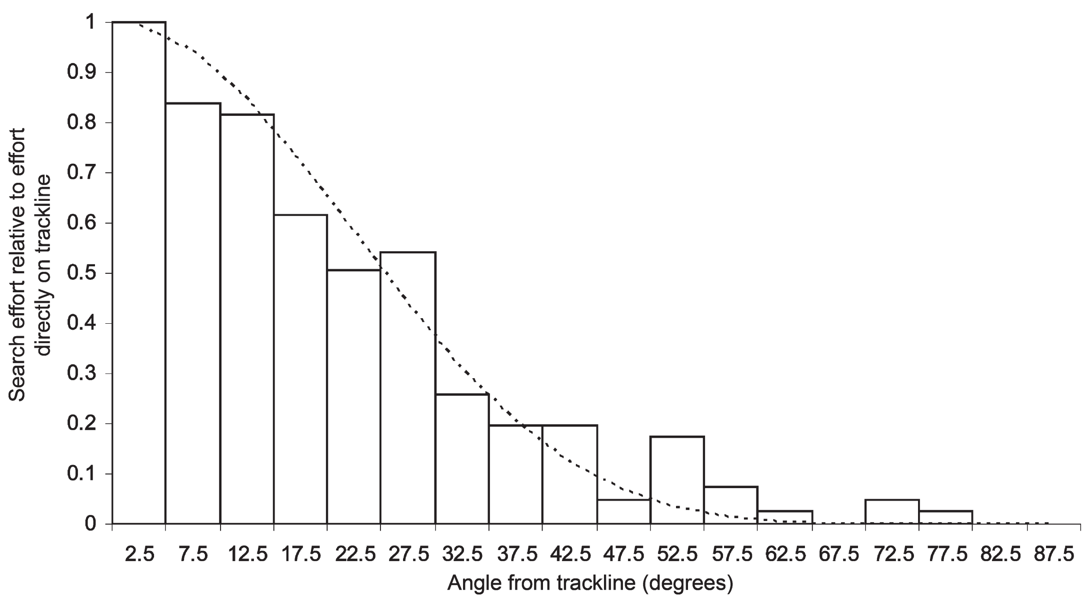


Fig. 4. Search effort by angle from randomised angle measurements from SCANS-II. Fitted function represents best fit by least squares of the form where *a* = 6.9.

distances for abundance estimation are not easy to predict. The overall distribution of radial distances to sightings will affect the direction and extent of any overall bias. We have only investigated the effects on methods using perpendicular distances but cue counting methods may be especially sensitive to non-linearity in errors in distance estimation (e.g. Borchers *et al.*, 2003) because these are based on area (i.e. square of distance).

One result apparent from the SOWER 2007/08 data was the comparison between the distance estimation errors during buoy experiments (CV_{RMSE} = 0.13 for observations from the barrel) and to whales during survey conditions (CV_{RMSE} = 0.24). It would be expected that estimated distances to a stationary object that remains at the surface are more accurate than those to whales and this is apparent from these results. There was also no evidence of the non-linear pattern in distance errors to buoy experiments that is common to the other datasets to actual sightings (Fig. 2g). These results suggest that distance experiments using fixed buoys may not yield much information about the errors that occur under real conditions. Williams *et al*. (2007) reached a similar conclusion, finding that errors in distances to transient cues were larger than those to cues that were visible for a longer period of time. There are also dangers in correcting for estimation error based on simple linear regressions. For example, in the case of the SCANS-II 7 × 50 estimates, a simple linear regression would suggest that distances were underestimated by around 7% (Table 1). Nevertheless, the simulation results in Table 2 would suggest that in this case strip width is likely to be overestimated (by 26% for a strip width without error of 391m). Thus a simple linear multiplier applied to distances would actually exacerbate the error.

The results presented here all involved data that have been through a careful validation process, both at sea and also prior to analysis. Recording distances and bearings by two separate methods allowed an initial screening for gross errors which could then be checked against the complete verbal commentary for each sighting. This validation process involved double checking around 10% of sightings which showed the greatest discrepancies. Although the majority of these cases involved errors with the estimated values, there were also errors in measured values. Errors in measured values could be corrected because all the raw images were stored. Overall, the rate of large discrepancies was higher than might have been expected, but was only apparent because of having two independent sets of data and there was no reason to assume that this was not typical of most surveys.

The patterns of non-linear measurement error observed in this study would appear difficult to correct without at least a substantial number of measurements to real sightings for comparison during a survey. The photogrammetric methods used provide such measurements and as techniques improve, measuring distances should be successful for an increasing proportion of sightings. The use of high definition video has resulted in a marked improvement in image quality on the most recent surveys (CODA and SOWER 2007/08). Detecting minke whale blows in the Southern Ocean on the standard resolution video images was identified as a problem in the 2006/07 SOWER data (Leaper, 2007). However, there were insufficient sightings of minke whales during the video experiments on SOWER 2007/08 to establish whether the high definition video was capable of detecting minke whale blows across the range of distances that blows are detected by visual observers.

Errors in angle measurements appear less likely to cause bias than errors in distances, but will affect the variance of estimates. There was no evidence of changes in angle errors with angle and thus an additive model should be appropriate for angle error. Additive errors for bearings will cause a small bias in perpendicular distances because for a true angle θ and angle error α

$$\frac{\sin(\theta + \alpha) - \sin(\theta)}{\sin(\theta) - \sin(\theta - \alpha)} < 1 \qquad (4)$$

i.e. the increase in perpendicular distance due to a positive angle error will be less than the decrease due to a negative angle error. For a RMSE of α of 7° or less, this bias will be less than 1% for any θ and so is not a major concern. The effect on the variance of the perpendicular distances may need more consideration. Although there was little evidence of angle error causing overall bias, the contribution to the variance will be dependent on the distribution of angles to sightings (Fig. 3). For sightings at 10°, 20° and 30° from the trackline, an RMSE in angles of 7° would contribute to a CV of perpendicular distances of 0.69, 0.34 and 0.21 respectively. Measurements of the proportion of time spent searching by angle sector do show differences between surveys, with 80% of search effort within 26°, 37°, 34° for SCANS-II Big Eye, SCANS-II 7 × 50 and SOWER, respectively. Thompson and Hiby (1985) found that over 80% of sighting effort was within 22.5° of the trackline on the 1983/84 IDCR cruise.

In conclusion, the contribution to the CV of the final abundance estimate from distance and angle estimation errors may be considerably greater than typical CVs for cetacean surveys that do not take these factors into account. In addition, estimation errors may also cause biases of similar or greater magnitude. Although simple linear regressions indicated that none of the surveys showed substantial overall bias, bias can nevertheless occur due to the non-linear relationships between errors and distance. In the case of the simultaneous sightings from SCANS-II, the bias would have been 29% if the survey had been reliant on naked eye estimates. The lack of a significant correlation between the truncated naked eye estimates (over the distance range of 200–500m) and measured distances, highlights the difficulties of estimating distances by naked eye. Distance errors are difficult to predict or correct from typical distance experiments using fixed targets and ultimately there appears no substitute for measuring these at sea. Video systems are still not at the stage where close to 100% success in obtaining images to sightings can be expected, but high definition cameras have allowed considerable improvements. Operating and maintaining complex electronics in harsh marine environments also remains a challenge. For example, one vessel on the CODA survey had major technical problems resulting in no measured distance data. Nevertheless, compared to increased ship time, investment in measurement technologies would appear likely to be a more cost effective way of reducing the CV of the resultant estimates, in addition to reducing the possibility of bias.

ACKNOWLEDGEMENTS

The data presented here were collected by a large number of observers on different vessels. We would especially like to thank the data collection methods group for SCANS-II, and the cruise leaders for SCANS-II, CODA and SOWER for their patience and persistence in keeping the data collection systems working at sea. Funding for the analysis by RL was provided by the International Fund for Animal Welfare.

REFERENCES

Alldredge, M.W., Simons, T.R. and Pollock, K.H. 2007. An experimental evaluation of distance measurement error in avian point count surveys. *J. Wildl. Manage.* 71(8): 2759–66.

Borchers, D.L., Pike, D., Gunnlaugsson, T. and Víkingsson, G.A. 2003. Analyses of the NASS 1987 and 2001 minke whale cue counting surveys taking account of distance estimation errors. Paper SC/55/NAM3 presented to the IWC Scientific Committee, May 2003, Berlin (unpublished). 33pp. [Paper available from the Office of this Journal].

Buckland, S.T., Anderson, D.R., Burnham, K.P. and Laake, J.L. 1993. *Distance Sampling: Estimating Abundance of Biological Populations*. Chapman and Hall, New York and London. xii+446pp.

Buckland, S.T. and Turnock, B.J. 1992. A robust line transect method. *Biometrics* 48: 901–09.

Chen, S.X. 1998. Measurement errors in line transect surveys. *Biometrics* 54: 899–908.

Chen, S.X. and Cowling, A. 2001. Measurement errors in line transect surveys where detectability varies with distance and size. *Biometrics* 57: 732–42.

Cooke, J.G. and Leaper, R. 1998. Analysis of covariates for NILS95 minke whale sighting data using the hazard probability method. Paper SC/50/RMP22 presented to the IWC Scientific Committee, May 1988 (unpublished). 12pp. [Paper available from the Office of this Journal].

Gillespie, D., Leaper, R., Gordon, J. and MacLeod, K. 2010. An integrated data collection system for line transect surveys. *J. Cetacean Res. Manage.* 11(3): 217–227.

Hiby, A.R. 1982. The effect of random whale movement on density estimates obtained from whale sighting surveys. *Rep. int. Whal. Commn* 32: 791–93.

Leaper, R. 2007. Results from use of video camera to measure distances to sightings on SOWER cruise 2006/07. Paper SC/59/IA25 presented to the IWC Scientific Committee, May 2007, Anchorage, USA (unpublished). 1pp. [Paper available from the Office of this Journal].

Leaper, R. and Gordon, J. 2001. Application of photogrammetric methods for locating and tracking cetacean movements at sea. *J. Cetacean Res. Manage.* 3(2): 131–41.

Marques, T.A. 2004. Predicting and correcting bias caused by error measurement in line transect sampling using multiplicative error models. *Biometrics* 60: 757–63.

Marques, T.A. 2007. Incorporating measurement error and density gradients in distance sampling surveys. PhD thesis, University of St. Andrews, St. Andrews, UK.

Thomas, L., Buckland, S.T., Rexstad, E.A., Laake, J.L., Strindberg, S., Hedley, S.L., Bishop, J.R.B., Marques, T.A. and Burnham, K.P. 2010. Distance software design and analysis of distance sampling surveys for estimating population size. *J. Appl. Ecol.* 47: 5–14.

Thomas, L., Laake, J.L., Rexstad, E., Strindberg, S., Marques, F.F.C., Buckland, S.T., Borchers, D.L., Anderson, D.R., Burnham, K.P., Burt, M.L., Hedley, S.L., Pollard, J.H., Bishop, J.R.B. and Marques, T.A. 2006. Distance 6.0 Release 4. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. [Available at: http://www.ruwpa.st-and.ac.uk/distance/]

Thompson, D. and Hiby, A.R. 1985. The use of scale binoculars for distance estimation and a time lapse camera for angle estimation during the 1983/84 IDCR minke whale assessment cruise. *Rep. int. Whal. Commn* 35: 309–14.

Williams, R., Leaper, R., Zerbini, A. and Hammond, P.S. 2007. Methods for investigating measurement error in cetacean line-transect surveys. *J. Mar. Biol. Assoc. U.K.* 87: 313–20.

Date received: January 2010
Date accepted: May 2010

Appendix

Leaper and Gordon (2001) described a system for photogrammetric measurement of bearings based on a digital camera attached to the binoculars used by the observer. Mounting the camera on the binoculars has the advantage of moving with the observer and ensuring alignment in a vertical plane because the observer will be holding the binoculars horizontal. The disadvantage is the additional weight for the observer. Observers on the SCANS-II and CODA surveys used a monopod with the 7 × 50 binoculars which took the full weight of the system. On SOWER, observers use a shorter binocular support and are sensitive to additional weight. Thus the system used for SOWER involved downward pointing cameras mounted above the observer. Two cameras were used, one with a remote shutter release (infra-red) which was pressed to obtain a bearing to a sighting and a time-lapse camera taking images every 30s to investigate scanning patterns. This system was very similar to that used on the 1983/84 IDCR cruise (Thompson and Hiby, 1985) except that the cameras were only used to monitor the starboard observer rather than the whole barrel. Two digital cameras, Pentax Optio S10 (for bearings to sightings) and GEC A835 (for time lapse) were mounted in a small, waterproof Lexan case as close to vertically above the observer as possible (Fig. A1). A white stripe was attached along the line of the binoculars to allow measurements. The infra-red remote control for the Pentax Optio was also mounted in a small waterproof box with a large waterproof push button.

It was not possible to position the camera box directly above the observers and so there was some error in bearing measurement due to parallax. This was measured using images of the angle board and found to be less than 1° for all angles within the search area of the starboard observer (the error to the binoculars will be slightly less than this because these were closer to directly beneath the cameras).

Fig. A1. Mounting of digital cameras above observers in the top barrel, SOWER 2008/09 cruise.