SC/60/PFI 5

Measurement error in survey and experiment, Norwegian minke whale surveys 2002-2007

GJERMUND BØTHUN, NILS ØIEN, HANS J. SKAUG

Institute of Marine Research, P.O.Box 1870 Nordnes, N-5817 Bergen, Norway
gjermund@imr.no

ABSTRACT

As part of the Norwegian multiyear surveys to estimate minke whale abundance in the Northeast Atlantic, experiments to evaluate naked eye distance estimation and angle board reading are conducted. Several error models for the experimental data collected over the recent six-year period 2002-2007 are explored and these results are also compared to information from survey duplicate data. The duplicate observations from the survey show a smaller variance in radial distance than the same values from the experiments, while the situation is opposite for angle readings. The mean radial distance and angle are larger in the experiments than in the surveys but it is unclear how this may effect variance.

KEYWORDS: VESSEL SURVEY, MEASUREMENT ERROR, LINE TRANSECT

INTRODUCTION

In line transect surveys the estimation of radial distance and angle to the detected object is fundamental. The method used to analyze the line transect data is robust with regard to variance estimation, but could be influenced by bias in estimated perpendicular distances to the whales. As an integral part of the NILS surveys over the period 2002 to 2007, experiments to collect data on measurement error in radial distance and angle have also been conducted. Here we present analyses of these data as well as models for the measurement error. This paper follows up the results from Aldrin & Tvete (2002).

In distance sampling the perpendicular distance from the transect line to the object is of great importance (Buckland et al. 2001). In line transect surveys for marine mammals the perpendicular distance is calculated from radial distance to the object and the angle between the transect line and the object. During the NILS surveys the radial distance is estimated by the naked eye and angle is measured with help of an angle board.

In the experiment the observer estimates the distance to a buoy with a radar reflector or a GPS mounted. At the same time as the observer is told to estimate distance and angle to the buoy, the true values are recorded from the radar or calculated from the GPS mounted on the ship and the buoy. We will first assume that the true values are recorded without error; we will then modify this assumption.

We will also model the measurement error from duplicate observation (from two independent observers) from the experiment and survey data.

DATA

One or two buoys are dropped into the sea at a distance of 1000 to 4000 meters from the vessel. The vessel moves towards the buoys at a speed between 6 and 10 knots at different courses. On a signal from the cruise leader the observer estimates the angle and distance to the buoy. At the same time the true values are recorded by the ships radar operator. All measurements are recorded in the same way as during the survey. This procedure is repeated several times as the vessel approaches the buoys. Since several recording are made within each drop of the buoys, one expects the data to be positively serial correlated. The cruise leader switches between the buoys at random and the vessel also changes the course several times during the approach towards the buoys to reduce the serial correlation. After having tested one observer, the vessel moves away from the buoys and a new observer is tested in the same way.

We will let R and r denote the observed and true radial distance, respectively. Further, Θ and θ denote the corresponding observed and true angle. Sometimes the observer or the radar operator has looked at the wrong buoy which will result in a large discrepancy in distance and/or angle. Some readings are incomplete, missing either distance or angle. Only observations with complete readings from radar and observer have been used in the analyses. We started with n=2640 complete observations. After having truncated observations that differ with more than 20 degrees in angle and observations outside the interval [100, 2000] meters, n=2265 observations are left.

For some of the experiments two observers have observed the same object at the same time from two independent platforms, giving us duplicate observations. After having truncated these observations in the same way as the rest of the experimental data, we were left with a data set of n=832 duplicate observations.

From the ordinary sightings surveys 2002-2007we have n=273 duplicate observations. From these data we have removed incomplete observations (either one or both platforms missing distance and/or angle), observations that differ by more than 60 seconds as well as observations made behind abeam. The remaining data have been truncated in the same way as for the experimental data, leaving us with a data set of n=205 duplicate survey observations.

ERROR IN RADIAL DISTANCE

Normal distribution

In Figure 1 all (R-r) vs. r have been plotted. Based on this Figure and histogram plots we first assume that (R-r) is normally distributed. Looking at the data at the original scale in Figure 1, it looks like the observers estimate the true distance unbiased at distances below 1000 meters. Above 1000 meters, however, the observers seem to underestimate the true distance.

We tried to simultaneously estimate the parameters for two regression lines and the break point (r=bp) between the lines by minimizing sum of squares. This gave a very weak fit. We therefore set the intercept of the second line in a way that ensures that the two lines connect in the breakpoint ($a_0=b_0+(b_1-a_1)*bp$). Removing one parameter stabilizes the solution. The best breakpoint lies at a place between 965 and 1093 meters, depending on starting conditions. We therefore fixed the breakpoint at r=1000 meters. We estimated the parameters in the model

$$R = \begin{cases} \beta_0 + \beta_1 * r + \varepsilon, & r < bp \\ \alpha_0 + \alpha_1 * r + \varepsilon, & r \geq bp \\ \varepsilon \sim \mathcal{N}(0, \sigma^2), & bp = 1000 \end{cases} \tag{1}$$

by minimizing sum of squares. The parameters for the two regression lines are given in Table 1

Table 1. Regression parameters for stepwise linear models.

	Intercept	Slope	Sigma (σ)
Above 1000 meters	356.6	0.577	235.6
Below 1000 meters	39.1	0.895	

We also estimated the parameters in a polynomial model of order two;

$$R = \beta_0 + \beta_1 * r + \beta_2 * r^2 + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2) \tag{2}$$

The intercept (β_0) is not significantly different from zero, so we fitted the model with β_0 fixed. The parameters for the model are summarized in Table 2. The fitted regression of models (1) and (2) are shown in Figure 1.

Table 2. Regression parameters for the polynomial model.

Intercept (β_0)	Slope (β_1)	Slope (β_2)	Sigma (σ)
0	1.067	-0.000161	235.7

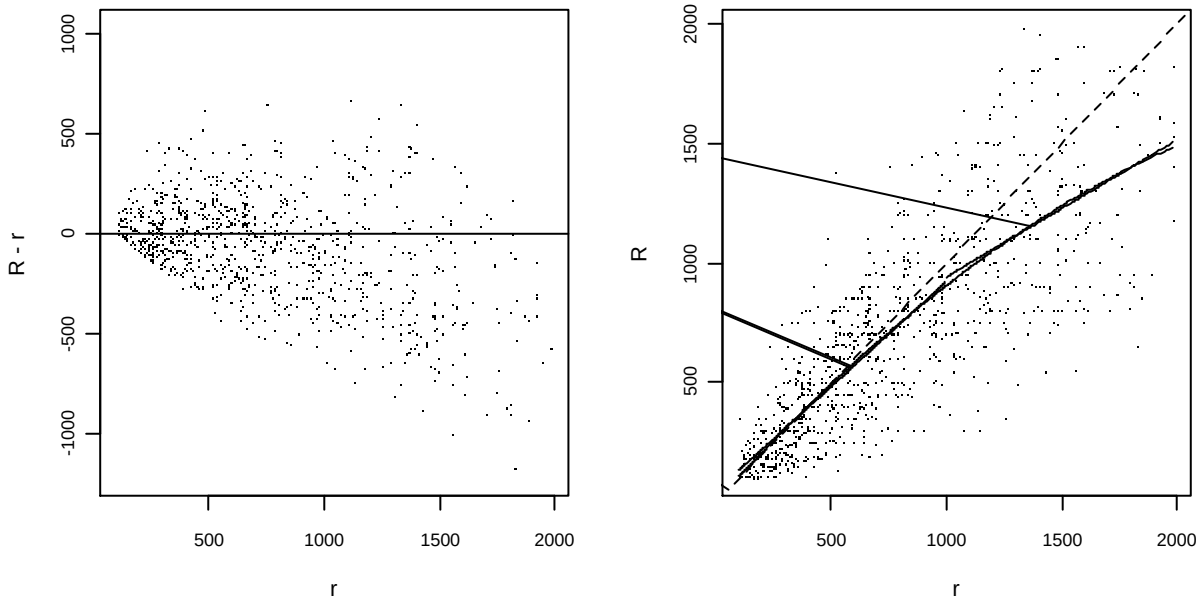

Figure 1. Left panel: (R-r) vs. r. Right panel: R vs. r, linear regression with breakpoint and polynomial regression.

Lognormal distribution

We now assume that ln(R/r) is *normally distributed*. We model ln(*R)* as a linear function of ln(*r)*:

(3) $$\ln R = \beta_0 + \beta_1 * \ln r + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

The model is maximized by maximum likelihood. From Figure 2 we expect that the intercept and slope will be close to 0 and 1 respectively. The parameter estimates are summarized in Table 3.

Table 3. Parameter estimates from model (3).

β_0	β_1	σ
0.602	0.888	0.380

Model (3) can now be formulated as:

$$R = 1.826 * r^{0.888} + \varepsilon, \;\; \varepsilon \sim \mathcal{N}(0, 0.38^2)$$

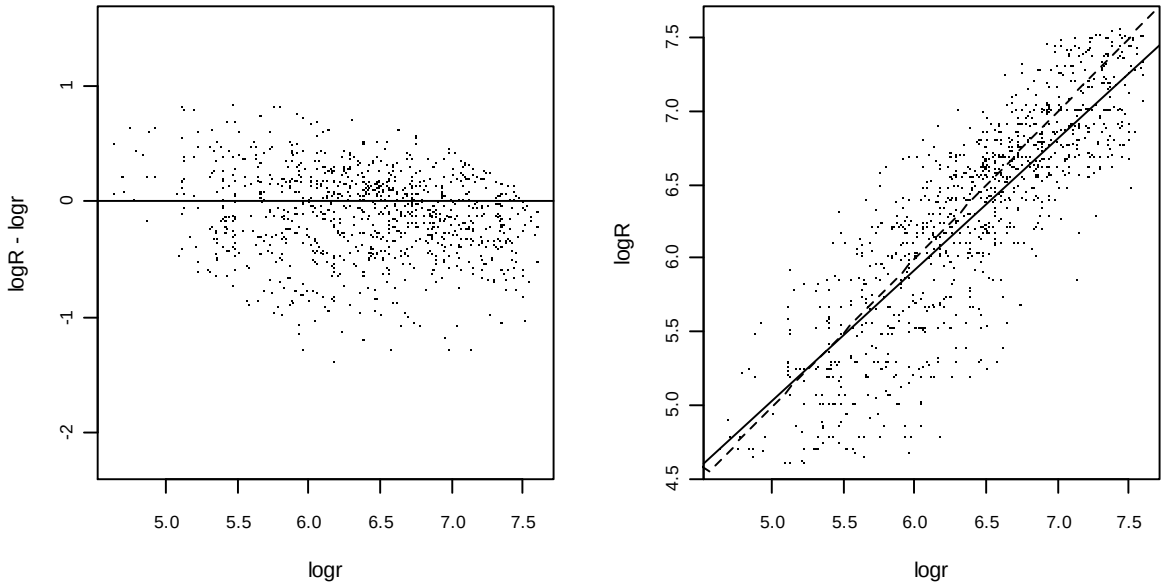

Figure 2. Left panel: ln(R/r) vs ln(r). Right panel: ln(R) vs.ln(r), fitted and straight line.

We will now study the residuals

$$\hat{\varepsilon} = \ln R - (\hat{\beta}_0 + \hat{\beta}_1 \ln r)$$

from model (3). We assume that $\hat{\varepsilon}^2$ is gamma distributed with expectation $\sigma^2(r)$. We investigate the relationship between σ and r with a generalized additive model (Hastie & Tibshirani 1990). We fitted the following model:

$$\mathrm{E}(\hat{\varepsilon}^2) = \ln\big(\sigma^2(r)\big) = 2\ln\big(\sigma(r)\big) = s(\ln(r))$$

where s is a spline function with approximately 3 degrees of freedom. The fitted spline function with 95% confidence interval is shown in Figure 3. Because of very few data points at the ends, the confidence interval gets wide there. The plot shows decreasing values of σ for increasing values of r in the mid region. The value of R-squared, deviance and GCV for the fitted spline all indicate only a weak relationship.

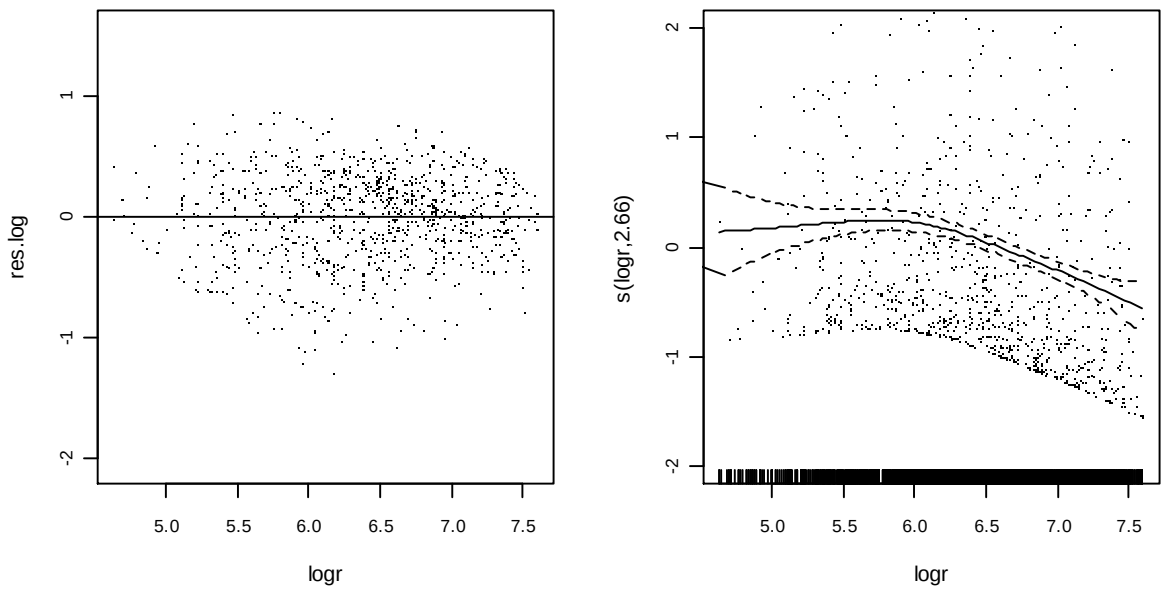

Figure 3. Left panel: Residuals vs log(r). Right panel: spline function estimate of log(σ^2(r)) vs log(r).

Even if it is not a clear linear relationship between σ and ln(r), we fitted a linear model (with a threshold at log(r) = log(403) = 6) as outlined in (Aldrin & Tvete 2002). We fitted the models

$$\ln R = \beta_0 + \beta_1 * \ln r + \varepsilon(r), \quad \varepsilon(r) \sim \mathcal{N}(0, \sigma^2(r))$$

and $$\ln \sigma(r) = \alpha_0 + \alpha_1 max(\ln r, 6)$$

simultaneously by maximum likelihood. The four parameters are shown in table 4.

Table 4. log r and log sigma.

β_0	β_1	α_0	α_1
0.6238	0.8844	0.6155	-0.2445

ERROR IN ANGLE

We started with a linear model

(4) $\Theta = \beta_1 \theta + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma)$

The fitted model is shown in Figure 4. The observer will normally be situated off the center of the angle board, and this could lead to a bias that is dependent on θ. There will be some delay between the times the observer are asked (the time the angle and distance is measured in the wheelhouse) to estimate distance and angle until he or she is finished with the experiment. The vessel will move a distance dl during this time (we used dl=20 meters). We have calculated the angle velocity Ω for each observation as:

$$\Omega = \tan^{-1}\left(\frac{\sin\theta}{\cos\theta - dl/r}\right) - \theta$$

We also found that

$$\Omega' = \frac{\theta}{\log(r)}$$

is a good proxy for Ω. In addition to model (4) we also estimated the parameters in the following models:

(5) $\Theta = \beta_1 \theta + \beta_2 \Omega + \varepsilon,$

(6) $\Theta = \beta_1 \theta + \beta_2 \Omega' + \varepsilon,$

(7) $\Theta = \beta_1 \theta + \beta_2 \sqrt[-1]{r} + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma)$

The parameter estimates for model (4) to (7) is summarized in Table 5. In model (5) and (6), β_1 and β_2 are highly correlated. Therefore we fixed $\beta_1 = 1$ in these models. Including angle velocity gave a lower residual standard error. In model (5) and (6) there are no θ dependent bias.

Table 5. Error in angle, estimated parameters.

Model	β_1	β_2	σ	AIC
4	1.072		4.28	13013
5	1	1.089	4.31	13045
6	1	0.442	4.26	12993
7	1.073	-20.3	4.18	12909

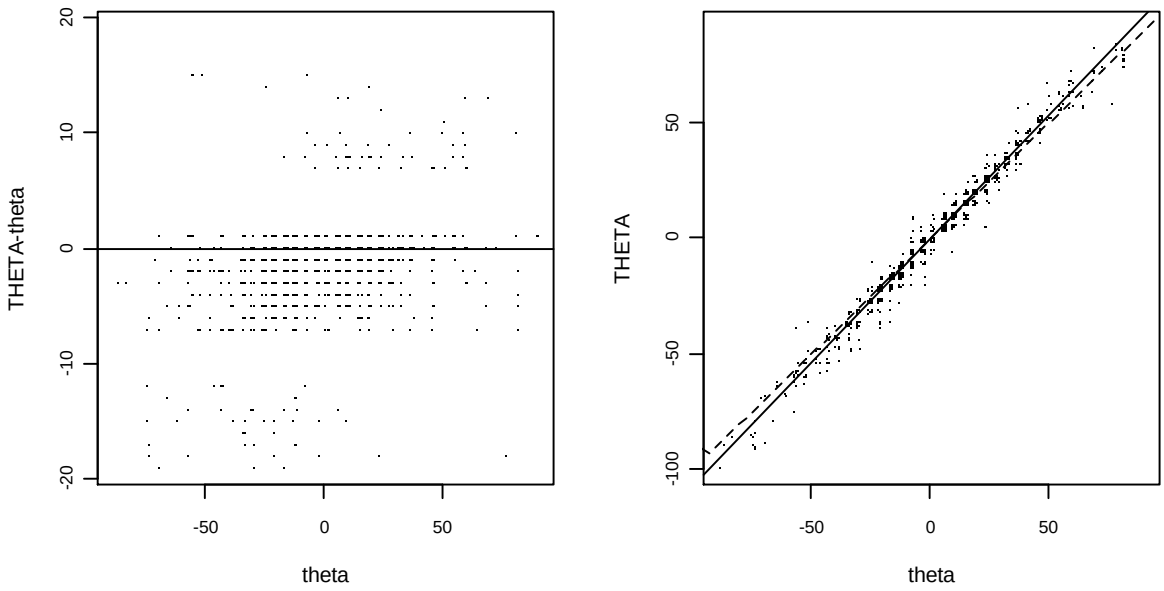

Figure 4. THETA and THETA-theta vs theta, and fitted model.

We will now study the residuals

$$\hat{\varepsilon} = \Theta - (\hat{\beta}_1 \theta)$$

from model (4). We assume that $\hat{\varepsilon}^2$ is gamma distributed with expectation $\sigma^2(\theta, r)$. We investigate the relationship between σ, θ and r with a generalized additive model (Hastie and Tibshirani 1990). σ increases with increasing |θ| and decreases with increasing r. Therefore we tried to describe variation in σ as a function of angle velocity. We fitted the following model:

$$\ln\big(\sigma^2(\theta, r)\big) = 2\ln\big(\sigma(\theta, r)\big) = s(\Omega')$$

where s is a spline function with approximately 4 degrees of freedom. The fitted spline function with 95% confidence interval is shown in Figure 5. Because of very few data point at the ends the confidence interval gets very wide there.

Sigma increases linearly for theta above approximately 30 degrees. Furthermore sigma decreases linearly for radial distances from 0 to around 900 meters. This shows that sigma is highest at small distances and large angles. In other words sigma follows the angle speed.

The value of R-squared, deviance and GCV for the fitted spline function all indicate only a weak relationship.

We fitted the models

(8) $$\Theta = \beta_1 * \theta + \varepsilon(\Omega'), \quad \varepsilon(\Omega') \sim \mathcal{N}(0, \sigma^2(\Omega'))$$

and $$\ln \sigma(r) = \alpha_0 + \alpha_1 \Omega'$$

simultaneously by maximum likelihood. The three parameters are shown in table 6.

Table 6. Parameters estimated for model (8)

β_1	α_0	α_1
1.073	1.453	0.003964

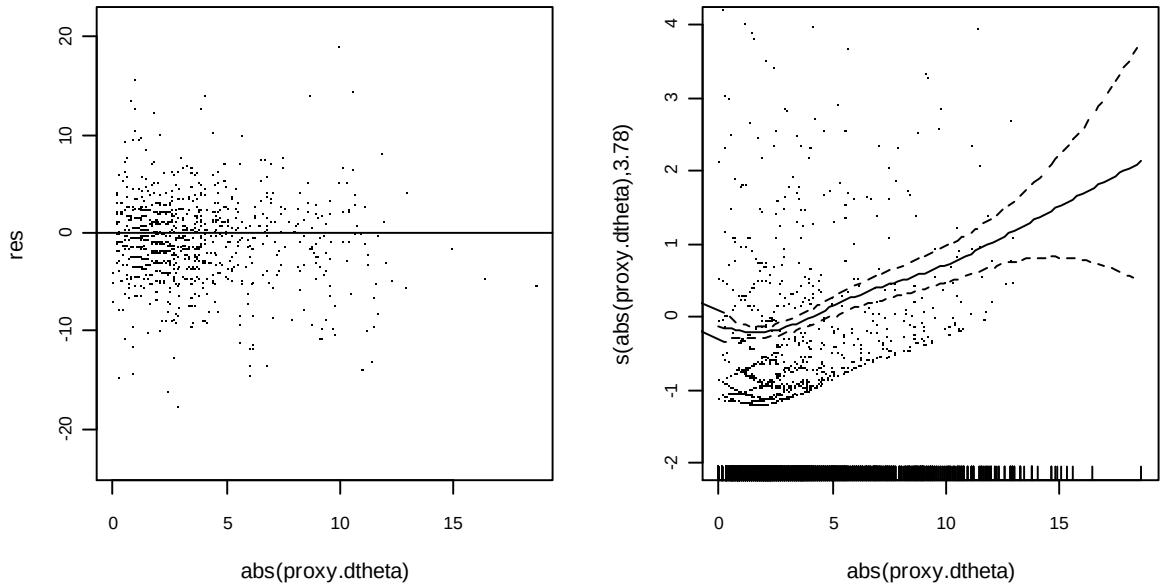

Figure 5. GAM log(sigma) and fitted spline function.

DUPLICATES

Duplicates from experiments

By excluding the reference information from the radar, one can use information in duplicate readings by two observers to estimate the variability in the different models.

$$\varepsilon_A = \nu_A + \omega, \qquad \varepsilon_B = \nu_B + \omega$$
$$var(\nu_A) = var(\nu_B) = \gamma^2 \quad , \quad var(\omega) = \varphi^2$$
$$\sigma^2 = var(\varepsilon) = \gamma^2 + \varphi^2$$

By assuming that the radial distances R_A and R_B observed from platforms A and B are *normally distributed*, $Var(R_A\text{-}R_B) = 2\gamma^2$, we find $\tilde{\gamma}$ = 214 meters.

By assuming that the observed radial distance is *lognormally distributed*, we find that $\tilde{\gamma}$ = 0.340.

By assuming that the observed angles Θ_A and Θ_B are *normally distributed*, we find that $\tilde{\gamma}$ = 2.80.

Duplicates in survey

By analyzing the data from the survey in the same way as for the experimental data, we find $\gamma_{r,s}$ = 0.291 and $\gamma_{\theta,s}$ = 6.40.

We see that the mean radial distance in the survey data are shorter than the same distance in the experimental data. If the data from the experiment and the survey are comparable, this contradicts the findings in table 4, which says that the variations should be smaller for larger values of r.

The mean observed |θ| in the survey are larger than the same value in the experiment. The survey and experiment also have different mean absolute angle velocity. But the parameterization of the model summarized in table 7 could not explain the differences between survey and experiment alone. In the survey data we allow for

differences of 60 seconds in time for a duplicate observation, this could explain some of the higher variation in the survey data.

Table 7. Parameters estimated from duplicates in survey and experiment.

	Regression	γ_b	γ_s
σ_r	236	214	150
$\sigma_{\log r}$	0.38	0.340	0.291
σ_θ	4.2	2.80	6.40

In Bøthun & Øien (2006) the average standard deviation of radial distance measured with GPS and radar is reported to be sd(R) = 9.53 meters, assuming that the error in GPS and radar to be equal. (Bøthun & Øien 2006) also reports root mean square error (RMS) to be 5.4 meters for one GPS. We use two GPS's to calculate the radial distance between the buoy and the ship, resulting in an RMS equal 7.6 for the error in the calculated radial distance. Using this value we find that sd(R_{radar}) = 12.52 meters and sd(R_{2gps}) = 5.00 meters.
Further, (Bøthun & Øien 2006) report sd(theta) = 1.35 degrees, and simulations show that $\varepsilon_{\theta\ 2gps}$~N(0,0.46^2) for the observed angle and radial distance in the experimental dataset.

So far, the radial distance from radar/gps (r) in model (1) to (3) have been treated as a true non-stochastic value. As shown in Bøthun & Øien (2006), r should be treated as a stochastic value and the models became variance component models. This could explain the difference in variation between the regression model (3) and the variation found by analyzing duplicates.

CONCLUSIONS

We have shown that the regression models (1) to (8) in fact are variance component models. The additional variance from the radar or the gps's can partly describe the difference in estimated variance between the regression models and the results from the duplicate observations in the experiment. The duplicate observations from the survey show a smaller variance than the same values from the experiment. We note that the distribution and mean radial distance are different in the survey and the experiment. It is unclear how this effects the variance.

The duplicate observations of angles in the survey show a higher variance than the same value from the experiment. We note that the distribution and mean value of angle are higher in the survey than in the experiment. We also note that there are large time lags between duplicates in the survey. This may explain the difference in variation between the two datasets.

Aldrin & Tvete (2002) recommends using the regression models to describe the measurement error; because it gives the ability to describe both bias and variation. Aldrin & Tvete (2002) also recommends adjusting the variation according to the variation observed in the survey.

We recommend to only using the regression models to describe the measurement error, even if it probably overestimates the variations slightly.

REFERENCES

Aldrin, M. & Tvete, I. F. 2002. Measurment error models for the Norwegian minke whale surveys 1996-2001. *Paper SC/54/RMP3, IWC Scientific Committee, April 2002.*

Bøthun, G. & Øien, N. 2006, Estimating measurment errors in distance experiments using GPS. *Paper SC/58/RMP5, IWC Scientific Committee, May 2006.*

Buckland, S. T., Anderson, D. R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L. 2001, *Introduction to distance sampling: estimating abundance of biological populations*. Oxford University Press, Oxford, UK.

Hastie, T. J. & Tibshirani, R. J. 1990. *Generalized additive models.* Chapman & Hall, London.