SC/61/RMP 4

Further analysis of measurement error in radial distance

GJERMUND BØTHUN

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

ABSTRACT

Bias and variability in observed radial distances is estimated from buoy sighting experiments. Earlier analysis of the same dataset have shown that observers have a tendency of underestimating radial distances (Bøthun et. al., 2007). In the present paper we account for serial correlation and observer specific effects and conclude that estimated radial distances are unbiased.

Introduction

One underlying assumption in distance sampling it that the sighting distance (the sampling unit) is recorded without measurement error (Buckland et al., 2004). This is probably not true in line transect survey for marine mammals. Experiments designed to test individual observers ability to estimate distances and bearing at sea; with respect to bias and precision was conducted during the North East Atlantic independent line transect survey (NAILS) in the period 2002 – 2007. Bøthun et al., (2007) analyzed this dataset and developed models describing bias and precision. Similar data from the survey period 1996-2001 is analyzed in Schweder, (1997) and Aldrin and Tvete, (2002). In this paper we focus on measurement error in radial distances and extend the models presented in Bøthun et al., (2007) by introducing observer specific effects in bias and precision. Further we investigate the data for serial correlation.

Material and methods

The distance and bearing experiments are carried out as an integrated part of the survey, and are carried out under conditions as similar to sighting conditions as possible. The experiment protocol is described in details in Bøthun et al., (2007) and Schweder, (1997). During the survey period a total of 138 experiments was carried out onboard five different vessels. 47 different observers was tested. This results in 2266 distance estimates. For two experiment there are not given any explicit information about beaufort, but based on qualitative information in the survey report we have coded the missing values as beaufort 0.

Bøthun et al., (2007) shows that a log-log transformation of the data improve the homogeneity of the residuals in the model. Figure 1 indicates that there still are heterogeneity between observers. Motivated by figure 1 we allow for individual variation. Figure 1 also indicates that it may be necessary with individual models for the expectation.

The way the experiment is conducted may introduce serial correlation. We model this correlation as a general auto regressive moving average (ARMA) time series process.

Observer specific biases is accounted for as a random effect. This give us a underlying "human model" accounting for the individual effect rather than a model that estimate the individual effect for the tested observers.

All models was optimized with the nlme software (Linear and nonlinear mixed effects models; Pinheiro, 2008).

Results

Fixed effects

In equation 1 we present a model with common intercept and slope for all observes. Further the model include beaufort and platform as covariate. Beaufort is treated as a numerical variable.

$$(1) \qquad \log(R_{ijk}) = \alpha + \beta * \log(r_{ijk}) + \gamma * plt_2 + \delta * bft + e_{ijk}$$

Where R_{ijk} is the k'th observation (radial distance) in the j experiment for observer i. r_{ijk} is the corresponding distance from radar or gps, threaded as an exact measurement. Platform (*plt_2*) is a variable with values 0 and 1, where the upper platform (platfrom1/barrel) is coded 0 and the lower platform (platfrom2/wheelhouse roof) is coded 1. Beaufort (*bft*) is the variable describing the meteorological observation of sea state on the beaufort scale. e_{ijk} is the residuals in the model; assumed normally distributed (o,σ) and independent for all values of i,j and k. The model is fitted with restricted maximum likelihood (REML). The parameter estimate is given in table 1.

Individual residuals

In model 2 we relax the the assumption about the residuals in model 1 by assuming that e_{ijk} are normally distributed (0, $\sigma\gamma_i$). The platform effect and the beaufort effect is not significantly different from zero in this model. In terms of Akaike's information criterion (AIC; Akaike, 1973) model 2 is an improvement over model 1 with Δ_{AIC}=249, but we also noted that in terms of the Bayesian information criterion (BIC; Schwarz, 1978) model 2 perform almost equal to model 1.

Individual bias (random effect)

Model 3 introduce individual (observer specific effects) biases in intercept and slope as random effects.

$$(2) \qquad \log(R_{ijk}) = (\alpha + a_i) + (\beta + b_i) * \log(r_{ijk}) + \gamma * plt2 + \delta * bft + e_{ijk}$$

Model 3 is expressed in equation 2. a_i is distributed N(0,σ_a) and b_i is distributed N(0,σ_b). Estimated parameters for model 3 are given in table 1 and table 3. In terms of AIC and BIC, model 3 is an strong improvement over model 1 (Δ_{AIC}=917).

Model 4 is expressed in equation 3. This is a nested random structure with observer as the primary effect and experiment within observer as an nested effect. In this model the common slope (β) is not significantly different from 1. We noted that the observer effect and the experiment effect is of the same magnitude.

$$(3)\quad \log(R_{ijk}) = (\alpha + a_i + c_{ij}) + (\beta + b_i + d_{ij}) * \log(r_{ijk}) + \gamma * plt2 + \delta * bft + e_{ijk}$$

Individual bias and residuals

Model 5 is expressed in equation 2, but with e_{ijk}~N(0,σ_i). Model 1 to 3 are all sub models of model 5. Estimated parameters for model 5 are given in table 1 and table 2. This model perform better than model 1 and 2 in terms of both AIC and BIC. But compared to model 3 there is some improvement in terms of AIC (Δ_{AIC}=150), but in terms of BIC this model perform worse (Δ_{BIC}=-114) because of the high number (46) of new parameters introduced.

Serial correlation

The experiment is designed with two buoys to reduce the effect of serial correlation within experiment. Figure 4 shows residuals for model 1 for some randomly selected observers and experiment. The residuals for the two buoys are shown with different symbols and there is a strong indication of a experiment effect rather than a buoy effect. Including buoy in the model results in a poorer fit. Figure 3 and figure 4 give indication of a strong serial correlation. In model 6 we extend equation 1 by modelling the structure of the residuals as an ARMA process. Both ARMA(2,0) and ARMA(1,1) models fits the data well. It is more important to model the correlation structure in a meaningful way rather than the best possible way (Schabenberger, 2002), therefore we chose to use a auto regressive process of order 2 (AR(2)). Model 1 is a sub model of model 6. Model 6 is a clear improvement in terms of AIC and BIC (Δ_{AIC}=1800, Δ_{BIC}=1799). Accounting for the serial correlation did not affect the parameter estimates of the intercept and slope much, but the estimated platform and beaufort effect are no longer significant. Further the estimated residual part of the model (sigma) increased after accounting for serial correlation.

Individual residuals and serial correlation

Model 7 extend model 6 by also allowing for individual residuals. According to AIC this is a improvement(Δ_{AIC}=123), but according to BIC model 6 perform best (Δ_{BIC}=-151).

Individual bias and serial correlation

Model 8 is a generalization of both model 3 and 6, and perform better than both according to AIC and BIC. The common slope and intercept are not significantly different from 1 and 0 respectively. Further there are not any significantly effect of platform or beaufort.

Individual bias and residuals and serial correlation

Model 9 is an extension of model 7 and 8. Model 9 perform better than model 7 according to both AIC and BIC (Δ_{AIC}=292, Δ_{BIC}=286). Model 9 perform better than model 8 according to AIC, but not according to BIC (Δ_{AIC}=98, Δ_{BIC}=-177). In model 9 the common slope in not significantly different from 1.

Discussion

The nested mixed-effect model, with observer and experiment effects, is the single factor that improve the fit of the model most. The second most important factor is to account for the serial correlation in the residuals. The nested mixed-effect model introduce an induced correlation that probably account for some of the serial correlation, but we believe that the underlying structure is a serial correlation structure because of the experimental design. The nested mixed-effect models

shows that for a given observer the variation among experiments is of the same magnitude as between observers.

The abundance estimation working group agreed on using a common measurement error model for all vessels and observers (SC/48/Rep1, 1997) after accounting for random effects.

Cooke and Leaper, (1996) concluded that accounting for observer effects would not produce a net benefit, because of a very high CV in the parameter estimates.

In all of the mixed-effects models, with the expectation of model 3, the fixed intercept is not different from one. We noted that the correlation between the random intercept and the random slope is highly negative (around -0.98) for all models.

According to the BIC model 8 (mixed-effect & serial correlation) is the best model, according to AIC model 9 (mixed-efffect, serial correlation & individual variance) is the best model. Model 9 got a penalty by the BIC because of the high number of parameters. One solution could be to model the gamma's as a random effect rather than estimating then as fixed parameters, but this is not possible within the NLME framework.

For the vessel "Nysleppen" (NYS) only the upper platform was used from 2002 to 2004 because the lower platform was too close to the radar. There is a co linearity between observer and vessel; one observer tend to stay on the same vessel from year to year. Further most of the observers only participate in a sub period of the survey period. Therefore the observed platform effect may be flavoured by year, vessel and observer effects or interaction between those.

For some of the cruises there have been allocated to little time to conduct experiment under proper conditions. As a compromise experiment have been conducted close to the shore line or within fjords, in bad weather conditions. This may explain way beaufort level 4 have a significantly higher value than beaufort 1-3.

Baird and Burkhart, (2000) and Cooke and Leaper, (1996) found a tendency of underestimation of distances at sea by the aid of the naked eye. However Williams et al., (2007) found a tendency of over estimating distances at sea by naked eye, but we noted that these conclusions are based on a very limited numbers of observers and distance ranges not relevant for this study.

This work highlight the importance of a proper model selection among a broad range of models accounting for assumed effects before drawing any conclusions. All models not accounting for serial correlation and/or random effects indicate a tendency of underestimation of radial distance, further this effect is more pronounced at the lower platform. This analysis also indicates a stronger tendency of underestimation for lower values of beaufort, but we note that the experimental design is not randomized according to beaufort.

References

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle., Second International Symposium on Information Theory. Budapest: Akademiai Kiado., pp. 267-281.

Aldrin, M., Tvete, F.I., 2002. Measurment Error Models for the Norwegian minke Whale Surveys 1996-2001. Paper IWC/SC/54/RMP3 presented to the IWC Scientific committee.

Baird, R.W., Burkhart, S.M., 2000. Bias and variability in distance estimation on the wather: implication for the managment of whale watching. Paper IWC/SC/52/WW1 presented to the IWC Scientific committee.

Buckland, S.T., Anderson, D.R., Larsen, F., 2004. Advanced distance sampling. Oxford University Press, Oxford.

Bøthun, G., Øien, N., Skaug, H.J., 2007. Measurment error in survey and experiment 2002-2007. Paper IWC/SC/60/PFI5 presented to the IWC Scientific committee, June 2008.

Cooke, J.G., Leaper, R., 1996. Analysis of angle and distance estimation from NILS95. Paper IWC/A96/AE34 presented to the IWC Scientific committee AE working group.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., the R Core team, 2008. nlme: Linear and Nonlinear Mixed Effects Models.

International Whaling Commission. 1997. Report of the Abundance estimation Working Group. Forty-Seventh Report of the International Whaling Commission. IWC/SC/48/Rep1, 1997.

Schabenberger, O., Pierce, F. J., 2002. Contemporary statistical models for the plant and soil sciences. CRC Press, Boca Raton.

Schwarz, G., 1978. ESTIMATING DIMENSION OF A MODEL. Annals of Statistics 6, 461-464.

Schweder, T., 1997. Measurment Error Models for the Norwegian Minke Whale Survey in 1995. Forty-Seventh Report of the Internatinoal Whaling Commission, Cambrigde, pp. 485-488.

Williams, R., Leaper, R., Zerbini, A.N., Hammond, P.S., 2007. Methods for investigating measurement error in cetacean line-transect surveys. Journal of the Marine Biological Association of the United Kingdom 87, 313-320.

Table 1 parameter estimates, AIC and BIC criterion. F fixed values.

Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\delta}$	$\hat{\sigma}$	$\widehat{AR(2)}$	Mixed	AIC	BIC
1	0.41	0.90	-0.14	0.059	0.37			1958	1986
2	0.41	0.90	-0.10	0.066	$0.29*\gamma_i$			1709	2001
3	0_f	0.96	-0.10	0.067	0.29		Yes	1041	1081
4	-0.25	1_f	-0.11	0.070	0.20		Yes	85	143
5	-0.26	1_f	-0.09	0.073	$0.11*\gamma_i$		Yes	891	1195
6	0.28	0.93			0.39	0.60, 0.23		158	187
7	0.21	0.93	-0.15	0.066	$0.20*\gamma_i$	0.62, 0.22		35	338
8	0_f	1_f			0.32	0.55, 0.25	Yes	-159	-125
9	-0.29	1_f		0.068	$0.17*\gamma_i$	0.56, 0.23	Yes	-257	52

Table 2 summary of models with individual residuals

Model	Min $\widehat{\gamma_i}$	Mean $\widehat{\gamma_i}$	Max $\widehat{\gamma_i}$
2	0.46	1.02	2.20
5	0.70	2.15	3.97
7	1.00	1.82	4.02
9	0.91	1.72	3.51

Table 3 Parameter estimates for random effects

Model	$\widehat{\sigma_a}$	$\widehat{\sigma_c}$	$\widehat{\sigma_b}$	$\widehat{\sigma_d}$	$\widehat{Cor(a,b)}$	$\widehat{Cor(c,d)}$
3	1.20		0.17		-0.98	
4	1.19	1.10	0.18	0.16	-0.98	-0.98
5	1.28		0.18		-0.98	
8	1.07		0.16		-0.98	
9	1.03		0.15		-0.97	

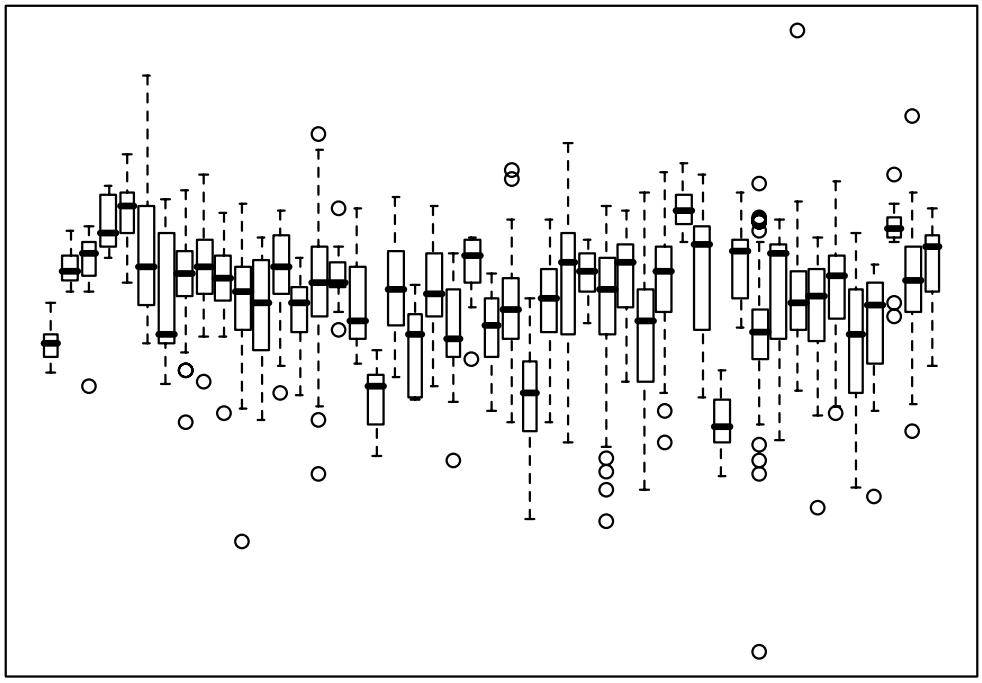

Figure 1 Boxplot of normalized residuals by observers, model 1.

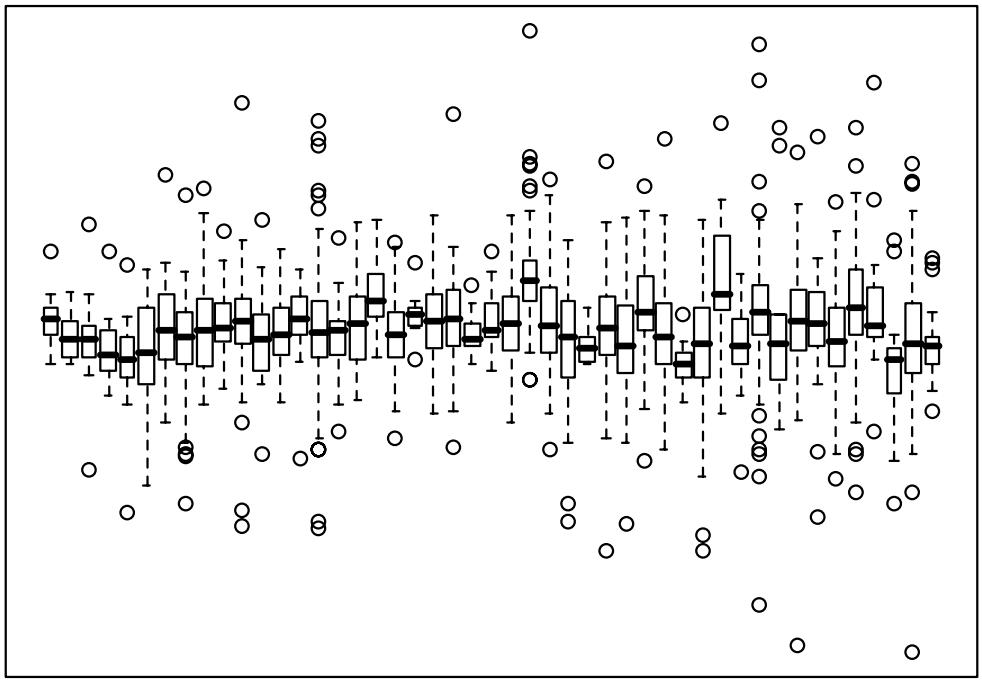

Figure 2 Boxplot of normalized residuals by observers, model 8.

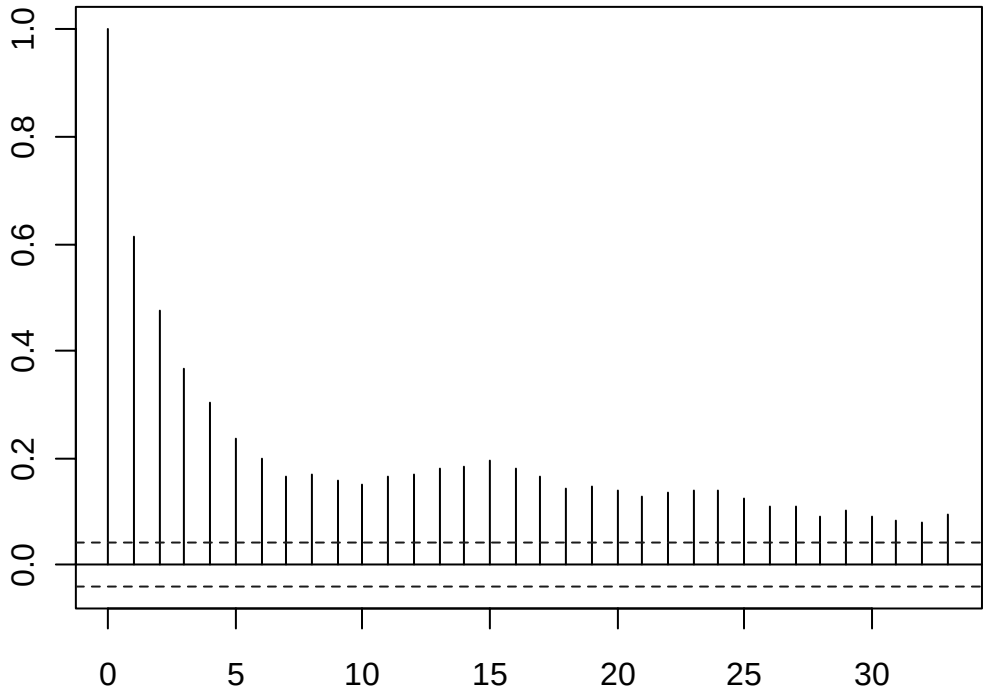

Figure 3 Autocorrelation versus lag, model 1.

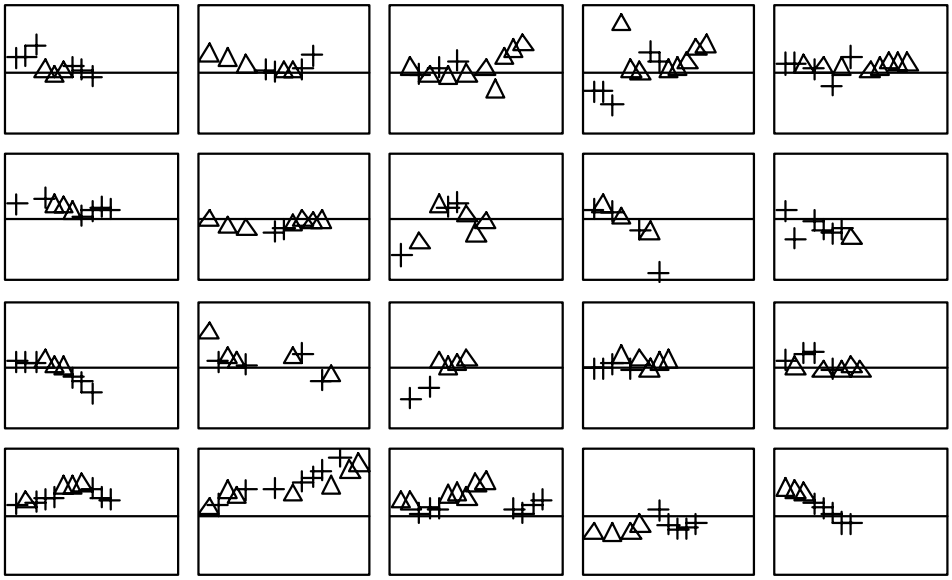

Figure 4 Residuals for some randomly selected observers and experiment, model 1. Plus-sign: port buoy, pyramids starboard buoy.