Annex E

Is SS2 an adequate replacement for SS3 as a way of implementing closing-when abeam protocols?

Mark Bravington (CSIRO CMIS, Hobart, TAS)

In the SOWER/IDCR surveys, the minimum school size estimates made during IO mode tend to be underestimates of true school size. To obtain accurate school size estimates, Closing mode is used instead. By comparing the reported school size distributions in IO mode with those in Closing mode (while allowing for sighting distance, environmental conditions, and for local spatial variations in school size distribution– in principle, anyway), it is possible to infer the extent of underestimation in IO mode. However, the comparison is only indirect, because we can only compare frequency distributions of reported school sizes in the two modes. This is a result of the particular protocols of Closing mode, whereby both platforms' normal observation processes are interrupted whenever either platform makes a sighting; hence we do not know what the reported school size of each sighting made in Closing mode would have been if the sighting had been made in IO mode instead. Underestimation of school size is an important "parameter" of all the new methods being developed for SOWER/IDCR analyses, so it is important to get a better handle on it than can be obtained from standard Closing/IO comparisons. There are three uses: direct inclusion in, or diagnostic checks for, the new methods; guiding the development of realistic simulation trials; and variations of Closing mode protocol in future surveys, so that the utility of the school size data can be improved.

In 20067, therefore, the SOWER cruise included some experiments to try to assess directly the extent of underestimation in IO mode. Statistically, the ideal way to infer underestimation in IO mode would be to run some of the survey using exactly the same platform setup and school size estimation protocols as in standard IO mode until the sighting comes abeam of the vessel, at which point closing is attempted. That way, each school size estimate can be made using IO protocols and then compared directly with the true school size obtained after closing. This direct sighting by sighting comparison is much more statistically powerful (and simpler) than the indirect comparison between frequency distributions required with standard Closing mode data. This approach is called "SS 3", because it is a successor to the "SS 2" experiments conducted in the mid1980s. The difference between SS 2 and SS 3 is that the IO barrel operates in SS 3 but not in SS 2.

The SS 3 experiment in 20067was conducted successfully, and there were not many cases of inability to close after waiting for the sighting to come abeam (note that there are always some sightings which cannot be closed on anyway, even when the protocol is to close immediately on sighting). Since the data from SS 3 is definitely more useful for abundance estimation than the data from standard Closing mode, this suggests using SS 3 as a replacement for Closing mode in future. However, the problem with SS 3 is that it requires the same number of on duty crew as IO mode, so in the context of a regular survey it does not provide adequate rest breaks. Given the current logistics of SOWER, it is not feasible to operate SS 3 as a replacement for Closing mode.

For this reason, the 20067 survey also ran an experiment in SS 2 mode, i.e. without IO, and this was also operationally successful. Since SS 2 is no more difficult than Closing mode in terms of crew schedules, it could potentially be used as a replacement for Closing mode in future. However, the lack of an IO means that the sample of sightings, and the process of estimating school size itself, is not exactly the same as in IO mode. Potentially, therefore, SS 2 data might give a biased picture of under-estimation compared to the "gold standard" of SS 3. This seems unlikely a priori, because relatively few sightings are made by the IO alone, and most (all?) of the school size estimation process does not depend on the IO. However, to fully justify a switch away from historical Closing mode to SS 2 (a.k.a. Closing abeam) mode, it is necessary to check whether the results of the recent SS 2 and SS 3 experiments are consistent. This working paper reports on some analyses.

1 DATA & METHODS

There were 90 sightings in SS 3 mode and 28 in SS 2 mode. True school size is an important covariate for the probability of underestimation (clearly, if the true school size is 1, the probability of under-estimation is 0). To prevent the sample size per "stratum" becoming too low, I grouped school size into four categories (1, 2, 3/4, 5+) and only consideration underestimations that crossed a category boundary. Other potentially important covariates include Sightability (cut into 2/3, with 43 sightings, and 4, with 75 sightings), and perpendicular distance at first sighting.

6 of the 118 sightings had a post closure estimate lower than the pre abeam estimate (4 of 9 in SS 3, 2 of 28 in SS 2). In 3 of the cases, the overestimation was proportionally small (5 vs 4, and 40 vs 35); in the other 3, the pre-abeam estimate was 2 and the post-closure size was 1. (It may be that the schools genuinely changed in size between the two measurements.) A 3% rate of substantive overestimation seems low enough to ignore, so I did; all overestimations were treated as correct.

For each sighting in either mode, we can determine whether the pre-abeam school size (category) estimate was lower than the post-closure size. A natural way to model the probability of this happening, is to use a logit-link binomial GLM, with covariates formed by school size, Sightability, and perpendicular distance, in various combinations and interactions. To check whether the results are affected by Mode (SS2 or SS3), we can choose one model and refit it with Mode as a covariate. We can investigate the Mode effect by comparing goodness of fit using AIC, or the significant via formal analysis of deviance, or informally by looking at the standard error and point estimate of the Mode coefficient ("Wald's test").

I tried a number of different models with and without interactions between perpendicular distance, school size, and Sightability, plus different category boundaries for school size and Sightability. In all cases, the Mode effect was insignificant; its standard error was always greater than the absolute value of its point estimate. While this is reassuring and perhaps not unexpected a priori, it should be noted that none of the other more obvious covariates emerged as significant, either; the sample sizes are quite limited. AIC always favoured the simpler models. On commonsense grounds, it seems unreasonable to exclude school size, perpendicular distance, or Sightability as covariates, whatever AIC says; the results from the simplest (i.e. no interaction) model that includes these factors is given in Table 1. A priori, we might expect the probability of underestimation to increase with perpendicular distance (which it does somewhat, in the table), and to decrease in good sighting conditions (which it doesn't). The most likely explanation for counterintuitive results is limited sample size and the resulting very high parameter uncertainty. However, one can easily imagine that there might be a genuine interaction between Sightability and perpendicular distance, in that sightings with large perpendicular distances (where underestimation is more likely) may only occur under good Sightabilities. The sample size is too small to reliably fit interaction models even when they are intuitively reasonable, so it is not surprising that some strange point estimates occur with over-simplified models.

Table 1:

Coefficients etc from a binomial GLM: base case for factors is Sch Size=2, Sightability=23, Mode=SS2

	Estimate	SE	Pr(>|z|)
(Intercept)	0.72	0.63	0.25
SchS=34	0.07	0.53	0.90
SchS 5	0.35	0.55	0.47
Perpdist	0.31	0.53	0.55
Sightab=4	0.52	0.44	0.24
SS3	0.34	0.50	0.50

2 DISCUSSION

There is no evidence that SS 2 will give different results about underestimation than SS 3 will. Admittedly, the sample size is very limited (more so for SS 2) and so the power to detect a small difference would not be high. Nevertheless, even if the two protocols would really give slightly different results with a large enough sample, it is important to note that SS 2 is still a better basis for comparisons with IO mode than standard Closing mode is; both SS 2 and Closing lack an IO and are therefore not strictly comparable with IO mode anyway, but direct estimation of school size under-estimation is much more powerful than indirect estimation, so any bias-variance trade off is likely to be in favour of SS 2.

With respect to the potential impact of protocol changes on comparability between past and future SOWER surveys: it is important to note that, based on simulation results in IWC SC IA, the "standard method" is very unlikely to continue to be used for agreed estimates of minke whale abundance from either past or future surveys. The new methods that are being developed all attempt to estimate absolute abundance, and so offer comparability “by design”; they are all (or should be) capable of coping appropriately with these different protocols. The advantages of switching to Closing abeam are that the precision of estimates should be improved appreciably, that analysis of (future) data will be simplified, and that straightforward diagnostics of each method’s model for underestimation of school size become available.

Overall, given the lack of evidence against and the strong in-principle logic in favour, it seems appropriate to support the Scientific Committee’s conditional recommendation of 2007, that Closing mode protocols be changed to the Closing ¬abeam protocol of SS 2.