Report of the Intersessional IA Workshop on estimating abundance of Antarctic minke whales

1. INTRODUCTORY ITEMS

The Workshop was held at the Marine Research Institute, Bergen, Norway, from 18-20 January, 2011 and was continued in Tromsø from 28 May-1 June, 2011. The list of participants is given as Annex A.

1.1 Convenor's opening remarks

Walløe welcomed the participants to the meeting and emphasised the importance of resolving the differences between the two sets of Antarctic minke whale abundance estimates - one from the OK model (Okamura and Kitakado, 2010) and one from the SPLINTR model (Bravington and Hedley, 2010) - so that an estimate could hopefully be agreed at SC63.

1.2 Appointment of rapporteurs

No formal appointments were made but a willingness to record the Workshop's discussions was indicated by Butterworth and Hedley.

1.3 Documents available

The new documents considered by the Workshop were SC/J11/AE1-12, as shown in Annex B.

1.4 Report format

The report is divided into two sections. Sections 2-4 report on work discussed in Bergen; sections 5-7 report on work completed in Tromsø.

2. COMPARISON OF HAZARD PROBABILITY AND TRACKLINE INDEPENDENCE APPROACHES

In order to lay the ground for an agreed abundance estimate during SC/63, the first step was to compare the estimation methods of the OK and SPLINTR estimation methods, in particular to examine the different ways that the two methods estimated detection probabilities and $g(0)$. Papers SC/J11/AE3 and SC/J11/AE5 used simplified forms of these approaches applied to the data to try to better understand the possible biases associated with implementations of the OK and SPLINTR methods.

SC/J11/AE3 showed that for school size 1 (s=1), density estimated assuming Trackline Conditional Independence (TLI) exceeded that from the 'standard method' (with $g(0)$ assumed equal to 1) by 70%, but that estimate density from the TLI model was less than that from the HR approach by 20%. It was noted that the results in this paper were unreliable for TLI when s>1 because of poor prediction of proportion of sightings that were duplicates.

SC/J11/AE5 developed an Empirical Hazard Probability (EHP) model for sightings within 0.3 n.miles of trackline. The model suggested that for s=1, the density for TLI was less than that indicated from the EHP model by 25%. It was thought that the bias arises from a high proportion of simultaneous duplicates. Also for s=1, the density estimates from the EHP model were less than those from comparable OK estimates, but the reasons for this were not well understood. Further work was needed; future investigations would include examining the sensitivity of the results to alternative cue rates and truncation distances other than 0.3 n.miles.

The Workshop noted some concerns regarding the OK approach, as follows:

(1) Lack of fit: the model predicts more duplicates close to the trackline than observed.
This could be investigated through extension of the EHP model to larger perpendicular distance, and in IWC simulation trials (but with parameters closer to actual SH minke situation).

(2) Low estimates of $g(0)$ for s=1.
The reasons underlying this concern are lower densities for s=1 from EHP approach compared to OK (see above), and the higher $g(0)$ estimate of about 0.25 in Burt *et al.* (2009) using Buckland-Turnock data collected on SOWER cruises as an independent check. Possible areas for further examination in OK approach are alternative specifications for confirmed/unconfirmed status, and treating s=1 parameters in the HR formulation as separate estimable factors rather than part of a continuous function in school size.

3. RESULTS FROM COMPARISON TESTS REQUESTED AT SC/62 MEETING

Documents SC/J11/AE2 and SC/J11/AE6 provided results from sensitivity runs of the OK and SPLINTR models respectively. In addition, SC/J11/AE4 provided some results from sensitivity runs for the 'IWC Standard analysis' method, in order to try to understand why previous SPLINTR model results were so similar to results which assumed $g(0)=1$ (e.g. Branch, 2006). The sensitivity runs were as specified in Bravington *et al*. (2011).

SPLINTR-like treatment of confirmed/unconfirmed

- Increases density by 8% for OK (though larger increase for simplified OK version in AE3), but decreases by 7% for standard, compared to SPLINTR.
- Thus increases discrepancy between OK and SPLINTR abundance estimates.

No error in estimates of school size

- Surprisingly large difference in effect for OK and for SPLINTR.
- Comparison is confounded by other differences in OK and SPLINTR implementations which are being addressed in the Work Plan.

Fix school size to 1

- Decreases density for OK by 67% and for standard by 56%; difference consistent with small $g(0)$ for $s=1$ for OK.

Omission of platform C (upper bridge) sightings

- Decreases (surprisingly) both OK and SPLINTR estimates by 7%.
- Previous results had suggested an appreciable and unexpected difference for OK.

Fix *g*(0)=1 for SPLINTR

- Roughly has only a 20% effect on density whether comparison within SPLINTR or on a per-stratum basis compared to the standard method.

The Workshop agreed that the original concern about the near equivalence of standard and SPLINTR abundance estimates had been addressed; the SPLINTR estimate would become notably greater after an appropriate decrease in the standard estimate to allow for spatial modelling. Furthermore there would be an increase in the SPLINTR estimate to allow for bias introduced through trackline independence assumption. In conclusion, there is no need to consider the IWC-Standard method further.

4. BIASING FACTORS AND A WORKPLAN

The Workshop agreed that from the SPLINTR/Non-Spatial SPLINTR comparison, the achieved survey tracks did not fully achieve the design requirements to eliminate bias. Thus the absence of consideration of spatial effects in the OK approach means that the associated abundance estimate is biased upwards. On the other hand, the analyses in SC/J11/AE3 and SC/J11/AE5 showed that the TLI assumption in SPLINTR leads to bias; consequently the associated SPLINTR abundance estimate is biased downward.

A procedural workplan for the period before SC/63 was developed in order to examine the quantitative effects of different factors in the models, and to allow appropriate adjustments to estimates in order to compensate for different applications of these factors (Annex C).

5. DECISIONS ON PREFERRED FACTOR SELECTIONS IN ADJUSTING OK AND SPLINTR ESTIMATES OF MINKE WHALE ABUNDANCE

The results from the procedural workplan described above were presented in SC/J11/AE8-12. A summary table of the effects of the various factors is given in Table 1.

Table 1

Factor adjustments (percentages) calculated by single-factor changes from all OK-preferred options to all SPLINTR-preferred options, based on the median ratio of abundance estimates across strata. See Section 6 for an explanation of the SPHAZ in the middle of the table.

FACTOR	ALL CP	CP2	CP3
OK preferred everything	0	0	0
Reference data instead of OK-pref	-3.4	-3.9	-1.6
No IO confirmation data	-10	-12.4	-7.7
All CL as confirmed	4.4	3.8	5.6
Recent SSX added	-2	-2.7	-1
HP à la SPHAZ not à la OK	12	15.5	8.8
TLI not HP à la SPHAZ	-40.8	-47.3	-38
Last instead of first perp dist	2.2	16.1	-0.2
Sightability instead of Beaufort	-7.8	-8	-7.4
Pre-extension à la SPLINTR	-3.1	-2.7	-3.4
Spatial model à la SPLINTR	-10.9	-15.3	-3.4

Reference dataset

The reference dataset had been developed primarily to facilitate the comparison exercise for which it would serve as a basis, in particular by excluding some data with missing fields that were hard to handle in SPLINTR. While some of the changes made from the OK preferred dataset might be argued to be desirable, this would not be the case for others. Pragmatically therefore it was decided not to implement this factor when adjusting OK estimates of abundance. For the purposes of this exercise, for SPLINTR, the reference dataset is the preferred dataset.

Confirmation status – treat all sightings in IO-mode as unconfirmed

About 15% of school size estimates in IO mode are confirmed; note that the confirmation process is very different to that in CL mode. SPLINTR proposes ignoring the fact of IO-confirmation data on the basis that for behavioural or other reasons, these IO-confirmed sightings may not constitute a random subset of all the IO-mode sightings in terms of school size. Hence using CL-mode data to infer school sizes for the IO-unconfirmed balance of these sightings would introduce a bias. The existence of such a bias was considered confirmed by the fact that this adjustment led to a difference in abundance estimates of some 10%. Accordingly it was decided to apply this adjustment factor.

Confirmation status – treat all sightings in CL-mode as confirmed

About 25% of school size estimates in CL mode are unconfirmed, and the recorded school size estimate is mostly 1 in these cases. The question is what to do about these CL-unconfirmed estimates. The concern about SPLINTR's approach of treating the estimates as accurate is that the school size for the CL-unconfirmed sightings may be underestimated because of absence of close approach, suggesting a downward bias if this procedure is adopted. On the other hand, indications of different behavior by singleton animals (which have proved more difficult to track in dive time experiments) suggest that treating CL-unconfirmed school size estimates as in the OK approach will preferentially exclude smaller schools and hence lead to a positive bias. The relative extents of these biases are difficult to estimate, so that the intermediate approach of applying half the factors estimated was adopted.

School size experiment (SSX) data

The SSX experiment is informative as it provides pairwise comparisons. However these data were collected during CPIII only, and school size distributions differed during CPII. The adjustment factor would be applied for CPIII but not for CPII.

Perpendicular sighting distance of duplicates

It was decided to adopt the OK approach of using the last (rather than, say, the first) sighting. Distance measurement errors are a substantial issue in SOWER and they are reduced at shorter distances from the vessel, so basing perpendicular distance on the final sighting (which will usually be closer to the vessel) should give the least distance error. A potential problem is that whale movement from its initial position relative to the vessel may introduce bias, but it was noted that the approaches had performed satisfactorily in the simulation trials (see Palka, 2010) which had included whale movement.

Weather covariates in detection function
While sightability in principle provides more information than Beaufort alone, a difficulty is that it has a subjective component and this may have changed over time. Efforts are being made to link sightability to objective measurements, so that such predictors of sightability could be used in place of the current partially subjective recordings. Further work along these lines was encouraged, but until such time as it may have been brought to a successful conclusion, adjustment by Beaufort alone was considered the preferred approach.

Pre-extend transect legs
While the in principle reasons for such a modification are recognised, there is as yet no clear basis for quantification of such an adjustment. Accordingly it was agreed not to apply this adjustment factor.

Spatial or stratified estimator
Achieved cruise track design differed more from that desirable under a random stratified design for CPII than was the case for CPIII. The adjustment factor for moving to a spatial estimator is larger for CPII, consistent with expectations for this reason. Differential adjustment of estimates for reasons of failure to achieve random stratified design objectives is considered desirable, and the method in SPLINTR has been simulation tested and is considered to be the best available approach.

6. RESULTS FROM APPLYING FACTOR ADJUSTMENTS

Having decided which factors to apply, the resulting adjustments (given in Table 2) were applied on a CP-specific basis to each method's preferred abundance estimates to give adjusted estimates (Tables 3-5). The CVs in Table 3 are taken directly from those calculated using the OK model in SC/J11/AE12; CV estimates from OK and SPLINTR have been similar in the past and it is expected that similar CVs would be obtained from any reasonable adaptations or combinations of the methods.

Table 2
Factor adjustments applied to each method (by CP series); see text in section 5 for details of factors. These are the same numbers as shown in Table 1 but here we show only the factors which were actually applied to each method's estimates to obtain adjusted estimates.

OK	CP2	CP3
No IO confirmation data	-12.4	-7.7
All CL as confirmed (upweight)	1.9	2.8
Recent SSX added (CP3)	0	-1
Spatial model	-15.3	-3.4
Cumulative	**-24**	**-9**

SPLINTR	CP2	CP3
Recent SSX removed (CP2)	2.7	0
All CL as confirmed (downweight)	-1.9	-2.8
OK-preferred data instead of reference	3.9	1.6
Last instead of first perp dist	-16.1	0.2
Beaufort not Sightability	8.0	7.4
No pre-extension	2.7	3.4
Cumulative	**+1**	**+11**

Table 3

Abundance estimates for CP2 and CP3 for original and adjusted methods. Note that there is one major adjustment that has not been made yet: from Trackline Independence to some form of Hazard Probability. There is as yet no quantitative estimate of the adjustment, but the preliminary results were closer to 'adjusted OK' than to 'adjusted SPLINTR'.

CP2 estimates:

MA	1	2	3	4	5	6	TOTAL
OK	166050	245183	172823	106420	606516	108272	1405264
adjusted-OK	125545	185375	130666	80461	458568	81861	1062477
adjusted-SP	82118	118385	67833	47059	254154	42690	612240
SP	81317	117231	67171	46600	251675	42274	606268
CV (OK)	0.20	0.16	0.21	0.18	0.17	0.27	0.12

CP3 estimates:

MA	1	2	3	4	5	6	TOTAL
OK	51878	76689	122217	79773	237367	106205	674129
adjusted-OK	47075	69589	110902	72387	215391	96372	611716
adjusted-SP	41842	55844	69518	36008	151844	66121	421177
SP	37785	50429	62777	32516	137121	59710	380339
CV (OK)	0.12	0.19	0.15	0.32	0.12	0.15	0.09

Table 4

Ratios between estimates for the two methods, before and after adjustment:

	MA	1	2	3	4	5	6	TOTAL
CP2	SP/OK	0.49	0.48	0.39	0.44	0.41	0.39	0.43
	adjusted-SP/OK	0.65	0.64	0.52	0.58	0.55	0.52	0.58
CP3	SP/OK	0.73	0.66	0.51	0.41	0.58	0.56	0.56
	adjusted-SP/OK	0.89	0.80	0.63	0.50	0.70	0.69	0.69

Table 5

Ratios between CP3 and CP2 estimates, by method before and after adjustment:

	MA	1	2	3	4	5	6	TOTAL
OK	CP3/2	0.31	0.31	0.71	0.75	0.39	0.98	0.48
	adjusted-CP3/2	0.37	0.38	0.85	0.90	0.47	1.18	0.58
SP	CP3/2	0.46	0.43	0.93	0.70	0.54	1.41	0.63
	adjusted-CP3/2	0.51	0.47	1.02	0.77	0.60	1.55	0.69

Despite these adjustments, these estimates still differ by 30-40%. The primary reason is methodological: OK estimates are based on the hazard rate modeling approach, whereas SPLINTR assumes trackline independence.

Neither of these sets of estimates can be accepted as they stand. The trackline independence assumption has been checked for the Antarctic minke data, and found to be invalid: the data show a greater proportion of the duplicate sightings to be simultaneous duplicates than is predicted under this assumption. This indicates that the SPLINTR abundance estimates are negatively biased, but by an unknown amount. The OK estimates are problematic because they substantially underestimate the number of duplicate sightings by the barrel and independent observer (by some 40% for CPII), which in turn suggests that they are positively biased as a result of underestimation of $g(0)$.

Thus the sets of estimates above are considered to bracket the true minke whale abundances in the areas surveyed, but the extent of bias of either is currently unquantified. Because a key assumption of the SPLINTR method is violated, it would be difficult to modify this approach in a simple manner to correct for this. However further work [HAZ-SPL; SC/J11/AE11 or SC/63/IA15] has been conducted which replaces the trackline-independence assumption within the SPLINTR approach by a hazard-rate formulation similar in its specifications to that used by OK.

Present results from this 'Hazardised SPLINTR' approach must be considered as preliminary only, as time limitations have precluded their careful checking as yet, and the approach has yet to be tested against the simulation trials under which the existing OK and SPLINTR approaches performed satisfactorily. Initial fits to the data using this approach nevertheless DO show an appreciably better match to the number of duplicate sightings seen by the barrel and independent observer than is the case for OK, and the abundances estimates produced exceed those from OK by about 10% (when all other factors are equal between the models). Although both models use the haz prob framework, there are differences in the formulation that could explain the differences in model fit and abundance. However, there remain likely problems, in that the estimates for $g(0)$ for individual platforms for schools of size 1 seem rather low, and the predicted numbers of resightings after an initial sighting by another platform is less than is observed. An advantage though of this last observation is that it allows a very coarse estimate of the likely extent of the positive bias in estimates of abundance from the hazardized-SPLINTR approach through underestimation of $g(0)$, suggesting that this is in the 10-20% range.

Reliable final estimates for the Antarctic minke populations cannot be provided immediately. They first require rigorous evaluation of the extent of this positive bias in existing hazard-rate approach based estimates. There is though a clear path with promise to resolve this problem. This is adjustment of the functional form used in existing hazard-probability formulations to better fit data on the number of sightings by the various platforms, together with the proportions of resightings. If lack-of-fit features of the existing hazard rate approaches can be resolved in this manner, this will allow reliable final estimates of minke whale abundance from the ICDR/SOWER cruises to be put forward at the next SC meeting.

7. CONCLUSIONS

To summarize, conclusions are as follows:

(1) The large differences in estimates from the two approaches recorded last year no longer appear to be irreconcilable.

(2) The hazard-probability formulation underlying the OK approach is to be preferred to the trackline independence assumption on which the original SPLINTR method was based.

(3) Other agreed adjustment factors (see Annex D) substantially reduce the difference between the current OK and SPLINTR abundance estimates, by about 30%.

(4) Adjusting for these other factors reduces the decline in abundance estimates from CPII to CPIII: from about 50% to 40% for OK, and from 40% to 30% for SPLINTR (see Table 5 for Area-specific estimates).

(5) Adjustment of details of current hazard-rate formulations to better fit data should allow final figures to be agreed at next year's SC meeting.

(6) Preliminary calculations indicate that those final figures will be closer to the adjusted OK than the adjusted SPLINTR estimates.

REFERENCES

Branch, T.A. 2006. Abundance estimates for Antarctic minke whales from three completed sets of circumpolar surveys. Paper SC/58/IA18 presented to the IWC Scientific Committee, May 2006, St. Kitts and Nevis, West Indies (unpublished). 28pp. [Paper available from the Office of this Journal].

Bravington, M., Hedley, S., Kitakado, T., Okamura, H. and Skaug, H. 2011. Report of the Scientific Committee. Annex G. Report of the Sub-Committee on In-Depth Assessments. Appendix 3. Intersessional process for resolving differences in minke whale abundance. *J. Cetacean Res. Manage.* 12(Suppl.): In press.

Bravington, M.V. and Hedley, S. 2010. Antarctic minke whale abundance from the SPLINTR model: some 'reference' dataset results and 'preferred' estimates from the second and third circumpolar IDCR/SOWER surveys. Paper SC/62/IA12 presented to the IWC Scientific Committee, June 2010, Agadir, Morocco (unpublished). 15pp plus revised. [Paper available from the Office of this Journal].

Burt, M.L., Ensor, P. and Borchers, D.L. 2009. Detection probability of Antarctic minke whales: analyses of the BT mode experiments conducted on the IWC-SOWER cruises 2005/06-2007/08. Paper SC/61/IA18 presented to the IWC Scientific Committee, June 2009, Madeira, Portugal (unpublished). 14pp. [Paper available from the Office of this Journal].

Okamura, H. and Kitakado, T. 2010. Abundance estimates of Antarctic minke whales from the historical IDCR/SOWER survey data using the OK method. Paper SC/62/IA3 presented to the IWC Scientific Committee, June 2010, Agadir, Morocco (unpublished). 35pp. [Paper available from the Office of this Journal].

Palka, D. 2010. Comparison of results from the OK, SPLINTR, Integrated and standard analytical methods when applied to simulated data, 2004-2008. Paper SC/62/IA14 presented to the IWC Scientific Committee, June 2010, Agadir, Morocco (unpublished). 21pp. [Paper available from the Office of this Journal].

Annex A

List of Participants

Norway
Hans Julius Skaug
Lars Walløe

Japan
Toshihide Kitakado
Hiroshi Okamura

Invited Participants
Mark Bravington
Doug Butterworth
Sharon Hedley

Annex B

List of documents

SC/J11/AE

1. OKAMURA, H. and KITAKADO, T. Revised abundance estimates of Antarctic minke whales from the OK method.

2. OKAMURA, H. and KITAKADO, T. Sensitivity analyses of Antarctic minke whales abundance estimation by the OK method.

3. OKAMURA, H. and KITAKADO, T. Comparison between the point independence model and the hazard probability model using the IDCR/SOWER Antarctic minke whales data.

4. HEDLEY, S. Conventional line transect analyses of IWC IDRC/SOWER data from CPII and CPIII Antarctic minke whale surveys: results for comparison with stratified SPLINTR.

5. BRAVINGTON, M. Empirical investigation of HP and TLI for SOWER.

6. BRAVINGTON, M. SPLINTR homework.

7. OKAMURA, H. Revised Table 2 in SC/J11/AE3.

8. BRAVINGTON, M. Effects of model details on SPLINTR.

9. OKAMURA, H. and KITAKADO, T. Fitting of the OK method for detection pattern of sightings.

10. OKAMURA, H. and KITAKADO, T. Rationales for the OK preferred dataset.

11. BRAVINGTON, M. A hybrid spatial-and-hazard-prob model for SOWER minke data.

12. OKAMURA, H. and KITAKADO, T. Abundance estimates for Antarctic minke whales using the OK method.

Annex C

Procedure and Work Plan to Develop Consensus Estimates

1. Both OK and SPLINTR developers will run their methods for their preferred choices of the factors in Table A1 (their 'Preferred Implementation'). Note that all runs hereafter will be for actual northern boundaries achieved in each survey - see also paragraph (3) of the addendum below.

2. Both OK and SPLINTR developers will also run their methods for the 'Reference Implementation' detailed in Table A1. Note that this Reference Implementation is NOT the group's agreed best choice, but was selected rather on the basis of ease of implementation by the developer needing to make the change, together with reasonable confidence that that change would not cause, for example, convergence problems. Note that the primary purpose of these Reference Implementations is to isolate the impact of the OK method Hazard Rate vs the SPLINTR method trackline independence assumptions.

3. Full diagnostics must be produced for subsequent reporting of results.

4. Each developer must carry out runs with changes of single factors only between their Preferred and Reference Implementations. It is not necessary to run every possible sequence of changes between these two end points. However sufficient should be run to check whether the impact of each factor on the overall abundance estimate is (reasonably close to being) linear (i.e. if impact of two factors together is well approximated by multiplying their separate proportional effects). Diagnostics need NOT be produced for these 'intermediate' runs.

5. The 'factor impact' will be calculated as the median over strata of the ratios of abundance estimates with or without the change in the direction from Preferred to Reference Implementation for each stratum. Each developer will select an order of application of factor changes in this sequence, and report for each stage both a plot of density and mean school size comparisons by strata and a table of the factor impacts disaggregated at the Circumpolar (CPII or CPIII) and Management Area (I, II, ...VI) level, together with their proposal for aggregated factor impacts where considered appropriate.

6. Full results, including estimates of variance for the Preferred Implementations, will be provided together with full documentation of the approach and reasons for choices amongst the factors in Table A1.

7. In the pre-meeting period before the SC meeting in Tromsø, the group will first review the diagnostics for the Reference and Preferred Implementations for both approaches.

8. The group will then decide its 'Best Implementation' set of choices for each of the factors in Table A1. Note that these need NOT be the same as the Reference Implementation choices, as their purpose is to reflect the group's view of the preferred approach in each instance.

9. Each developer's 'Preferred estimates' of abundance will be converted to 'Method-specific consensus estimates' by multiplication by the proportional 'factor impact' values developed in step 5) above. (This will be where necessary - if runs with a combination of some such changes are already available, their results will be used instead.) Effectively these OK-consensus and SPLINTR-consensus estimates will differ only through the effects of the fundamental differences between these two methodologies.

10. Available information associated with estimates of bias and robustness for both the OK/Hazard Rate and the SPLINTR/Trackline Independence approaches will be debated. Either consensus will be reached on agreed adjustments which then provide unique 'Consensus estimates'. If that does not prove possible, a minimum upward bias for Trackline Independence and minimum doward bias factor for the OK Hazard Rate implementation will be agreed, resulting in two 'Adjusted Method-specific consensus estimates'.

11. The CVs associated with each method-specific consensus estimate will be the CV of the corresponding Preferred estimate (i.e. the factor impacts will be treated as exact).

12. The group will consider simple approaches to combining the two Adjusted Method-specific abundance estimates into unique 'Pragmatic estimates' for proposal should unique consensus estimates not have been achieved.

13. An approach consistent with the development of unique Consensus estimates or Pragmatic estimates will be used to provide CVs for those estimates based on the CVs of the Method-specific consensus estimates. Upper and Lower confidence Limits for all estimates will be developed similarly.

14. For estimates of abundance for boundaries differing from that specified in (1) above are required, factor impacts provided by SPLINTR will be used because this can be implemented more easily than OK.

15. The resultant estimates will be put forward to the IA sub-committee for endorsement. The group will make every effort to either provide unique consensus estimates, or failing that Adjusted Method-specific consensus

estimates that differ by a little as consistent with available scientific analyses and their defensible interpretations. Note that this approach does NOT preclude further development of either approach subsequent to the Tromsø SC meeting, with such results being considered at subsequent meetings to revise estimates hopefully agreed by the SC meeting in Tromsø.

Annex D

Reference dataset and stepwise changes in data to assess linearity of effects

For comparisons between the model approaches, the following factors were examined in a stepwise fashion, starting from each developers' Preferred Implementations.

Factors	**Treatment**	**Reason**	**Which model is changing**
Reference dataset	Use	SPLINTR cannot run very easily otherwise	OK
Confirmation treatment	(1) Treat all sightings in CL-mode as confirmed & all sightings in IO mode as unconfirmed	Hard-wired in SPLINTR	OK
	(2) Use CL-mode classification; treat all IO-mode sightings as unconfirmed	A 'second step' before reaching the OK treatment of using recorded confirmation status in both modes	OK
School size experiment (SSX) data	Use SSX data	SPLINTR needs these data for stable estimation	OK
Sighting distance of duplicates	Use latest (not first) distance	OK model back-calculates earlier sighting distances from the closest one, using recorded time and vessel speed.	SPLINTR
Weather covariates in the detection function	Use Sea State for CPII and CPIII. Good=0-3; Bad=4-5	Sea state probably more consistently recorded across CP series	SPLINTR
Pre-extend transect legs	No	Pre-extension is not usually applied in other studies.	SPLINTR
Spatial or stratified estimator	Stratified	OK model is a stratified model	SPLINTR