SC/69A/ASI/15/Rev1

Sub-committees/working group name:

Addressing recommendations by the IWC Scientific Committee to improve estimates of abundance of the franciscana dolphin

Federico Sucunza, Daniel Danilewicz and Alexandre N. Zerbini

Addressing recommendations by the IWC Scientific Committee to improve estimates of abundance of the franciscana dolphin.

Federico Sucunza, Daniel Danilewicz, and Alexandre N Zerbini

Introduction and background

Estimates of abundance of the franciscana dolphin (*Pontoporia blainvillei*) have been under by the IWC Scientific Committee (SC) as part of the review of the status of the franciscana by the Committee following the process established by the working group on Abundance Estimates, Status of Stocks and International Cruises (ASI) (IWC, 2021a, 2022). As part of the review process, the SC provided several suggestions to improve correction factors (CFs) to estimate density of franciscana dolphins as well as to improve estimates of abundance of this species and agreed the following tasks should be performed before the review of the status of the franciscana is completed by the Committee (IWC, 2021b [SC/68C/REP/02], Item 3.8, p. 12):

1. The CV of the correction factor for visibility and group size bias presented in Sucunza et al. (2020a [SC/68B/ASI/04_rev1]) be revised and the estimates computed with this correction factor will be updated.
2. An estimate of uncertainty for group size bias correction computed by Sucunza et al. (2020a [SC/68B/ASI/04_rev1]) be calculated and the estimate of abundance in Franciscana Management Area (FMA) II (Sucunza et al. 2020b) be updated.
3. The estimate of abundance for the 2014 FMA III survey (Sucunza et al., 2020c [SC/68B/ASI/06]) be revised following the discussion provided in item 3.5 of IWC (2021b [SC/68C/REP/02]).
4. If the situation with the coronavirus pandemic improves, surveys in FMA IV will be completed and new estimates of abundance will be computed.

In this document, we address points (1) to (3) above. Point (4) is addressed separately in Crespo et al. (this meeting).

In addition to the tasks above, a number of suggestions were made by the ASI reviewers to improve the estimates of abundance for franciscanas in the various management areas. Annex 1 of this document provides a description of how authors addressed those suggestions to improve estimates for FMAs Ia, Ib, II and III.

Point 1 – Revision of the CV visibility and group size bias correction factor

A study to compute factors to correct for availability, perception and group size bias on franciscana aerial surveys was conducted in Babitonga Bay, Brazil in 2011. Preliminary results of this study were presented by Zerbini et al. (2011) and these were subsequently updated by Sucunza et al. (2020a). The latter provided estimates of the correction factor (CF) for all sources of bias (visibility and group size) as well as separate estimates of each source of bias (group size, perception and availability).

In Sucunza et al. (2020a), the CF for visibility and group size bias had been estimated at 4.42 (CV=0.02). This estimate was revised following the recommendations of the SC at 4.76 (CV=0.25). This study has since been published in Frontiers of Marine Science (Sucunza et al., 2022). We propose the new estimate be endorsed by the SC.

The SC categorized estimates of abundance for FMA Ia in 2018 (Sucunza et al., 2020d [SC/68B/ASI/05]) and Ib in 2011 and 2017 (Danilewicz et al., 2020 [SC/68B/ASI/07_rev1]) as “Provisional” and indicated that they could be elevated to Category 1 once the CV of the CF for visibility and group size bias was updated (IWC, 2021). Revised estimates for these areas are provided in Table 1 and we propose they are endorsed as Category 1 by the Committee.

Table 1 – Revised estimates of franciscana abundance ($\hat{N}$) in FMA Ia and Ib.

Area	Year	$\hat{N}$	CV
FMA Ia	2018	1,183	0.76
FMA Ib	2011	1,590	0.53
FMA Ib	2017	1,521	0.47

Point 2 – Compute uncertainty of the correction factor for group size bias in aerial surveys for franciscanas and update the estimate of abundance for FMA II.
Sucunza et al. (2020a) demonstrated that estimates of group sizes by observers on an airplane were 36% lower than those in a surface platform (a vessel in this case). Therefore, these authors recommended that estimates of franciscana abundance computed from aerial surveys be corrected by a factor of 1.36 (CV = 0.11). The correction factor was built using a log-linear model incorporating the effect of platform (boat and airplane), and the uncertainty around these parameters was combined using the delta method to obtain the variance of the bias in group size estimates (Sucunza et al. 2022).

In 2021, the SC endorsed an estimate of abundance of franciscanas in FMA II in 2009 (N=6,827, CV=0.26, Sucunza et al. 2020b) as Category 2 but noted that that it could be elevated to Category 1 once corrected for group size bias. The estimate for FMA II corrected for group size bias using the 1.36 (CV=0.11) factor is 9,284 (CV = 0.28). We propose this estimate be elevated to Category 1A.

Point 3- Revised estimate of abundance for FMA III
Suggestions to improve an estimate of abundance of franciscanas in FMA III in 2014 (Sucunza et al. 2020c) were provided in section 3.5 of IWC (2021). These suggestions are addressed (Appendix 1) and the revised estimate for FMA III corrected for visibility and group size bias is 9,437 (CV = 0.34).

References

Danilewicz, D., Sucunza, F., Ott, P.H., Ferreira, E., Perez, M.S., Berchieri, N., Alvares, D., Andriolo, A.,
Secchi, E.R., Flores, P.A.C., Farro, A.P., Martins, A. and Zerbini, A.N. 2020. Abundance and distribution of franciscanas (*Pontoporia blainvillei*) in northern Rio de Janeiro (FMA Ib), Brazil. SC/68B/ASI/07_rev1, 15pp.

IWC. 2021a. Report of the Scientific Committee (SC/68C). 200pp.

IWC. 2021b. Report of the Workshop on the Review of the Status of the Franciscana, 7-9 April 2021. Paper SC/68C/REP/02, 22 pp.

IWC. 2022. Report of the Scientific Committee (SC/68D). 218pp.

Sucunza, F., Danilewicz, D., Secchi, E.R., Andriolo, A., Cremer, M., Flores, P.A.C., Ferreira, E., Alves, L.C.P., De Castro, F., Pretto, D., Sartori, C.M., Schulze, B., Denuncio, P., Perez, M.S. and Zerbini, A.N. 2020a. Assessing bias in aerial surveys for threatened cetaceans: results from experiments conducted with the Franciscana (*Pontoporia blainvillei*). SC/68B/ASI/04_rev1, 33pp.

Sucunza, F., Danilewicz, D., Andriolo, A., Azevedo, A.F., Secchi, E.R. and Zerbini, A.N. 2020b. Distribution, habitat use, and abundance of the endangered franciscana in southeastern and southern Brazil. *Marine Mammal Science* 36(2): 421-35. https://doi.org/10.1111/mms.12650

Sucunza, F., Danilewicz, D., Cremer, M., Ferreira, E. and Denuncio, P. 2020c Abundance of the endangered franciscana in southern Brazil. SC/68B/ASI/06, 16pp.

Sucunza, F., Danilewicz, D., Ott, P.H., Neves, M., Berchieri, N., Farro, A.P., Martins, A. and Zerbini, A.N. 2020d. Population size and IUCN Red Listing of the isolated population of the franciscana (*Pontoporia blainvillei*). SC/68B/ASI/05, 25pp.

Sucunza, F., Danilewicz, D., Andriolo, A., de Castro, F.R., Cremer, M., Denuncio, P., Ferreira, E., Flores, P.A.C., Ott, P.H., Perez, M.S., Pretto, D., Sartori, C.M., Secchi, E.R. and Zerbini, A.N. 2022. Assessing bias in aerial surveys for cetaceans: Results from experiments conducted with the franciscana dolphin. Front. Mar. Sci. 9:1016444. doi: 10.3389/fmars.2022.1016444

Zerbini, A.N., Danilewicz, D., Secchi, E.R., Andriolo, A., Cremer, M., Flores, P.A.C., Ferreira, E., Alves, L.C.P.d.S., Sucunza, F., de Castro, F.R., Pretto, D., Sartori, C.M., Schulze, B., Denuncio, P. and Laake, J.L. 2011. Assessing bias in abundance estimates from aerial surveys to improve conservation of threatened franciscana dolphins: preliminary results from a survey conducted off southern Brazil. SC/63/SM9, 13pp.

Appendix 1

Reviews of estimates of abundance and of correction factors for perception, availability andgroup size for the franciscana dolphin (*Pontoporia blainvillei*): response to reviewers

Reviews: Correction Factors

Relevant document: SC/68B/ASI/04_Rev1

Reviewer 1

Methods
Experiment 1 uses two methods. One is an estimation of group size, density and abundance based on standard CDS/MCDS techniques to compute a correction factor for aerial surveys. This analysis is straightforward, usesthe *mrds* functions and was already approved by the IWC in Zerbini et al. (2011). However, the results presentedin ASI/4_rev1 differ slightly because the number of aerial sightings used here was lower (by 11) than in Zerbiniet al. (2011). I was not able to understand where the difference comes from. Similarly, there are some differencesin the distance calibration results between the two versions of the analysis and it is not clear why (possibly because the radial distance data were log-transformed in ASI/4_rev1 but this transformation is not mentioned in Zerbini et al. 2011?).

Authors: Preliminary results provided by Zerbini et al. (2011) differ slightly because that dataset was refined and re-analyzed, and the results presented in Sucunza et al. 2022 supersedes those from this preliminary report

The second method (also from Zerbini et al. 2011, with the same unexplained difference in sample size) uses a GLM to assess differences in group sizes across platforms, with model selection and averaging among four candidate models. This approach is valid, the best models are significantly better than the null model, and modelaveraging was warranted by the small difference in AIC between the two best models. Two remarks: there is noinformation on model fit and no CV is given for the 33% greater group size in boat surveys, even though it should be possible to compute one from the GLM result. Before this value can be used as a correction factor forother surveys, it will need a measure of uncertainty.

Authors: Model assumptions were verified by plotting residuals versus fitted values and versus each covariate in the model (Zuur and Ieno, 2016). The variance of the bias in group size estimates from the airplane was approximated by the delta method (Seber, 1982).

Predicted group sizes for each platform computed from the model parameters estimates were revised and indicated that groups seen from the boat are, on average, 36% (CV = 0.11) larger than those seen from the airplane (Sucunza et al. 2022). This suggests that estimates of abundance for franciscanas using aerial surveys is biased low if average group size estimates are used to compute density and should be corrected by a factor of 1.36 (CV = 0.11). Estimation of uncertainty for this correction factor was recommended by the IWC SC and the CV provided her addresses this recommendation.

The resulting correction factor (CF = 4.42, CV = 0.04) for visibility and group size bias for aerial surveys was developed as the ratio of the density estimated by boats (D = 2.99 ind/km2, CV = 0.23) and by the aircraft (D = 0.68 ind./km2, CV = 0.28). The text mentions that the CV for the ratio was computed using the delta-method, but the resulting CV of 0.04 seems very small and I was not able to replicate it. (Unless there is a covariance component that is not mentioned here?)

Authors: The correction factor was revised (Sucunza et al., 2022) and recomputed as 4.76 (CV = 0.25) based on density estimates from boat (D = 3.20 groups/km^2, CV = 0.22) and airplane (D = 0.67 groups/km^2, CV = 0.28) surveys. Revision of these quantities were also recommended by the IWC SC and the new values presented in Sucunza et al. (2022) address the recommendation.

Experiment 2 uses a GLMM to model the effect of covariates on the proportion at surface, following Sucunza et al. (2018). The use of a mixed-effect model is appropriate given that measures of complete dive cycles are repeated for several of the groups observed (similar to what is done in analyses of dive data from tags). Then the authors use a standard approach, Laake et al. (1997)'s model, to calculate the probability of a dolphin group being available within the field of view given the speed of the aircraft. Note that the formula on p. 12 is incorrect (but calculations are correct): the first term's denominator should be E(s)+E(d) rather than E(s)E(d) and E(m) (which is not defined) should be replaced with E(d).

Authors: addressed.

The additional data show that there are differences in mean surface and dive times between the two areas but that the resulting proportion spent at the surface is similar. These are useful results, suggesting that instantaneouscorrection factors for availability bias (e.g., those used for photo surveys) will not vary much across regions andwater conditions, but that the adjusted correction when taking into account the time-in-view might differ (for instance in the deeper waters Ubatuba where dives are longer, the adjusted correction would be 0.38). However, the additional data do not change the result significantly (the adjusted correction for the combined two areas still rounds up to 0.39).
Four remarks on this part: first, the SE of 0.009 (i.e., a CV of 0.02) calculated for the adjusted correction of 0.39seems unrealistically small, given that the CV around the proportion at surface is much larger (0.54 based on Table 9). This seems to be the result of the nonparametric bootstrap procedures. I think we need more details onhow this is implemented? In each replication, did the E(s) and E(d) correspond to observed pairs of surface anddive cycles (presumably, longer dives were followed by longer surface times, which may have kept the resultingvariation low). It just seems weird that someone using these results would end up with 0.35 (CV 0.54) for instantaneous correction but 0.39 (CV 0.02) for the same correction adjusted for time-in-view. (In other studiesI know, the CV of the adjusted correction is assumed to be the same as that of the instantaneous correction.)

Authors: Barlow et al. (1988) model was used to calculate the probability of a dolphin group being available (*Pr*) and the variance was estimated by the delta method (Seber 1982), given by the equation presented by Crespo et al. (2010). New values of *Pr* and CV were computed at 0.39 (CV = 0.39) (Sucunza et al. 2022). These new values result in a CV for the instantaneous correction that is much more realistic than the one reported previously.

Second, rather than use an average value of 6 seconds for the time-in-view (which may not be valid for all aerial surveys in the franciscana range), it would be possible to calculate the actual time-in-view for the perpendicular distance of each sighting (i.e., use the empirical distribution of the time-in-view for specific survey datasets).

Authors: although we agree with the idea, this point was not included in our revised manuscript once it is not clear if a group just below the surface will at $x > 0$ will by visual available as a group at $x = 0$.

Third, the conclusion rely heavily on the assumption that $g(0)$ for boat surveys was equal to 1. Rather than assume unity, the Laake et al. formula could be used to calculate the exact bias based on boat speed, time window for the corresponding field of view, and the recorded dive/surface times. It willlikely yeld a value very close or equal to 1 but it would remove the need to assume.

Authors: We computed g(0) as suggested by the reviewer using the Barlow et al. (1988) method and the resulted value was 1.

Fourth, we are told that the depth at which a dolphin can be seen from an aerial platform varies from 1.40 m in murky waters to 2.25 m clearer waters. This means that for the same diving behaviour, the apparent surface times (as defined here in terms of visibility from a helicopter) would increase and the dive times would decrease in deeper waters. Since the calculated proportion of time was almost the same in Ubatuba than in BB, does this means that the calculations of the proportion at the surface include both effect of water murkiness and that of potentially different diving behaviour? Does that create a limitation to the applicability of the factor (e.g., if applied to an area with clear water but shorter dive times)?

Authors: yes, this is a potential problem. Different diving behavior with respect to the proportion of time at surface will result in different estimates of availability bias. In this sense, Sucunza et al. 2022 indicate that experiments to address availability of franciscana groups to aerial platforms are recommended in other regions to compute improved and/or area-specific correction factors.

Reviewer 2

The manuscript did not state whether the data for experiment 2 were categorized by behavior, whether there were differences in the typical behaviors of franciscana in the two study areas during the survey periods, or how the typical behaviors of franciscana in the two study areas during the survey periods compared to that of other locations and times.

Authors: although it is possible to identify some behaviour categories (e.g. feeding) from the helicopter, it is extremely hard to do that from the airplane during the aerial surveys. In this sense, it was not intended to produce behaviour-specific correction factors for availability bias during experiment 2. Group behaviour was not systematically recorded during experiment 2, however, because multiple groups of different sizes and compositions were recorded in various habitats within two very different geographic regions, it is expected that most, if not all, behavior types displayed by the species have been sampled throughout our study.

I think there is an error in the GLM equation in section 2.1.3.1. Here, the authors state that the Poisson GLMtakes the form

Authors: The GLM equation has been corrected as suggested by the reviewer.

Additionally, there was a typo in the equation for availability bias on the bottom of pg. 12 (i.e., Laake et al.'seqn 4). The equation at the bottom of page 12 uses both E(m) and E(d) for average (expected) dive duration.

Authors: This has also been corrected.

Are the conclusions appropriate given the results?

For the most part, yes, the conclusions are appropriate given the results. There are two notable exceptions.

First, as mentioned above, the results from experiment 2 likely depend on behavior state because Bordino et al. (1999) noted important differences in surface/dive synchrony and duration depending on behavior state.

Authors: see response above.

Second, the authors state, "The new estimates of availability of franciscana groups reported in this study as well as the independent estimate of groups size bias can be used independently of the assumptions described to the correction factor." It is important to note that the estimates of E(s) and E(d) from experiment 2 can be used independently (with the caveat above regarding potential behavioral differences). However, the value for w(x), the amount of time that the ocean at distance x is available to be seen by the observers, will vary depending on the field of view of each individual aircraft and on survey speed. I have found that the radius of the observer's view area from McLaren's model is quite difficult to estimate in practice
Authors: agree.

Reviewer 3

Experiment 1: simultaneous ship and plane *Data*

The track lines are only 600 m apart. Is this too close together where the animals could react and move to the next track line or split up the groups? Could the ship have split up the groups and then the plane flew over and saw smaller group sizes? Maybe not?

Authors: based on field observations during the experiment, groups of franciscana dolphins don't seem to have split up. If they had been split up, group size estimates for boat surveys would have been different than those computed using other methods (e.g., helicopter surveys), but that was not documented in experiments reported in Sucunza et al. (2018, 2022).

Method

To test difference between group sizes used a GLM with Poisson error where distance and platform were predictor variables. Since the effect of distance is different for each platform, it seems it is more appropriate to standardize the distance from the two platforms or else only test group sizes for the distances only within the flat part of the respective detection function.

Authors: because the interaction between the two variables was not significant, we preferred to use all the data to improve models fit.

Distance calibration correction factors should be by distance for each observer, unless there is not a difference by distance, which was not discussed.

Authors: addressed. "Perpendicular distance estimated from the boats were corrected for each observer considering the calibration experiments described above prior to estimation of detection probability for that platform." (Sucunza et al. 2022)

Availability correction factor equation from Laake et al. 1997 appears to be incorrect, where E(m) should be E(d). Not sure if they used just an average w(0)=6 sec for each observation or if they used the w(x) for each sighting's individual x. Reported correction factor was 0.39 (SE=0.009). This SE seems way too low.

Authors: addressed (see description above about recalculation of the SE of the correction factor of 0.39).

Reviews FMA Ia

Relevant document: SC/68B/ASI/05.

Reviewer 1

Have the data used in the analysis, including any pre-selection process, been clearly specified?
I am not clear how many observations were used finally. In page 7, it says n=64. In page 8 it says 17 observations (14 on effort, 3 off effort). In page 11, Table 2, it appears 8 observations. Maybe I missed some clear explanation but I didn´t see it. Why there are three different n provided? Which is the final one and where do the other n come from?

Authors: a total of 27 on effort recorded franciscana groups (n = 8 transect lines, n = 19 transit lines) were used to fit detection function models, and juts the 8 groups recorded during the transect lines were used to estimate density.

Was the method implemented as intended?
Yes, although in page 7 it says that with the small number of observations (n=64….see point raised above), they only used a null model (no covariates) and half-normal key function. If n is actually 64, then it is sufficiently large as to also try the hazard-rate (which by seeing the histogram under the detection function would probably fit better) and also some covariates. The way it was done, there is no possible model selection as only one option was attempted.

Authors: A set of detection function models were fitted to the binned data, including the uniform, the half normal the hazard-rate key functions with cosine, Hermite polynomial and simple polynomial adjustment terms and the covariate group size, following standard combinations proposed in Thomas et al. (2010). Models with acceptable fit based on visual assessment and on goodness-of-fit statistics where ordered based on the Akaike Information Criterion (AIC) values. A set of detection function models were fitted (Table 1) and bootstrap analysis was performed to estimate density and abundance (Table 2).

Table 1. Proposed detection function models (Model), estimated average detection probability ($\hat{P}$), coefficient of variation (CV) of $\hat{P}$, Akaike's information criterion differences between the model in question and the most parsimonious model (ΔAIC), proportion of the bootstrap resamples for which the model in question was selected (*w*boot). Uniform (Unif), half-normal (Hn) and hazard-rate (Hr) key functions, cosine (cos), Hermite polynomial (Herm) and polynomial (poly) adjustment terms of order(x), and group size covariate (size).

Model	$\hat{P}$	CV($\hat{P}$)	ΔAIC	*w*boot
Hn	0.616	0.176	0.000	38%
Unif+cos(1)	0.607	0.145	0.005	38%
Hn + size	0.616	0.177	1.998	10%
Hn + cos(2)	0.618	0.327	1.999	2%
Hn + Herm(4)	0.616	0.274	1.999	<1%
Hr	0.652	0.240	2.010	6%
Hr + poly(4)	0.636	0.337	3.997	<1%
Hr + size	0.651	0.242	4.009	6%

Table 2. Density and abundance estimates of franciscana dolphins in Espírito Santo State, southeastern Brazil, through the study period. Franciscana Management Area Ia (FMA Ia). FMA Ia north and FMA Ia south correspond to geographic regions (i.e. strata) used for density estimation. Coefficient of variation(CV). *n* = number of sightings used for density estimation (averaged over all bootstrap resamples), ER = number of franciscana groups detected/km *on effort* of planned effort (averaged over all bootstrap resamples), Es = estimated average group size (averaged over all bootstrap resamples), $\widehat{Du}$ = estimated uncorrected density of individuals/km^2 (averaged over all bootstrap resamples), $\widehat{Dc}$ = estimated density of individuals/km^2 corrected for visibility bias and group size bias (averaged over all bootstrap resamples), $\widehat{Nc}$ = abundance corrected for visibility bias and group size bias (averaged across all bootstrap replicates), CI = confidence intervals.

Strata	*n*	ER (CV)	Es (CV)	$\widehat{Du}$ (CV)	$\widehat{Dc}$ (CV)	$\widehat{Nc}$ (CV)	95% IC
north	4	0.007 (0.572)	1.388 (0.191)	0.049 (0.708)	0.232 (0.764)	336 (0.764)	31 – 960
south	4	0.010 (0.507)	4.083 (0.201)	0.160 (0.572)	0.760 (0.640)	847 (0.640)	132 – 2,190
FMA Ia (Total)	8	0.008 (0.540)	2.562 (0.226)	0.097 (0.676)	0.462 (0.756)	1,183 (0.756)	163 – 3,150

In the case of 'standard' methods:
• *have the version and options used been specified?*
Which release version of R was used?, and which engines/libraries were used to model the detection function? I would like to see a bit more

Authors: bootstrap analysis was performed using a set of customized functions in "Distance" version 1.0.5 (Miller et al., 2019) and "mrds" version 2.2.6 (Laake et al. 2022) packages in R version 4.1.1 (R Core Team 2021).

Reviewer 2

Could appreciably more precise or reliable estimates have been produced using a different analysis? (In the case of data collected in major international programmes, the SC has a particular responsibility to ensure that these are effectively used.)
A model-based abundance estimation using environmental variables could be tested in addition, as a possibilty to include the additional sightings that were collected 'offeffort'. A model-based abundance estimate would allow for analysis of non-design-based data, therefore inclusion of the additional transects that were excluded for abundance estimation. But this would just be an additional analysis – the current one is totally fine, since the criteria for a design-based survey and analysis have been met.

Authors: agree with the suggestion, and it will be used in future works.

Reviews FMA Ib

Relevant documents: SC/68B/ASI/07_Rev1.

Reviewer 1

Was the method implemented as intended?
Yes, although I would consider more diagnostics before deciding on the final detection function. For example, for 2017, the Hn key function was chosen, and looking at the plot in Figure 6, I would think that a Hr key function would be more appropriate. The AIC is only 0.44 points smaller in Hn than in Hr because, I assume, it is just more parsimonic with only one parameter instead of two. In these cases, I like to look into other diagnostics, particularly the Cramer von Misses value, and most probably it will be much smaller for Hr than Hn meaning a much better fit (and the p-value much higher, as happens with the chi-square GOF). I also like to look at the CV of the detection function, which in this case is much smaller for Hr than for Hn meaning that it is capturing more of the variability. Looking at all that (and without seeing the CvM value and p-value and the shape of the Hr detection function, which I would like to see), I would have chosen the Hr.

Particularly when the sample size is small, the AIC tends to penalize a lot the number of parameters, so using other criteria in parallel with the AIC may be very useful. For 2011, the detection function shown in Figure 6 seems to be the one for model 2 or 3, not 1 (null model). There are two lines of dots, meaning there are two shapes in there, probably two different sea states or two different levels of water transparency. If it was the Hr without covariates, it would only shown one line of dots, like for 2017. Also the shape looks more like a Hn than a Hr. Please check that the right plots are shown.

Authors: the dataset was refined and re-analyzed and new results are presented here.

"Detection probability (*P*) was estimated using Conventional (CDS) and Multiple Covariate Distance Sampling (MCDS) methods (Buckland et al. 2001, Marques & Buckland 2003). Only the half-normal and the hazard-rate detection functions were proposed to fit distance data. Sighting data from other franciscana surveys (Sucunza et al. 2020, Sucunza et al. 2022) carried out by the same observers were combined with 2011 and 2017 surveys data to increase sample size and better estimate detection probability in this study. Exploratory analyses indicated that adequate fits were obtained by modelling binned perpendicular distance data (grouping intervals: 0m, 29m, 59m, 129m, 198m and 275m) and by left truncating the perpendicular distance data at 20m. Beaufort sea state (factor covariate with two levels: "calm", Beaufort sea state between 0 and 2, and "high" between 3 and 4), glare (factor covariate with two levels "presence" and "absence"), water transparency (factor covariate with two levels "clear" and "murky"), group size (numerical covariate) were considered as covariates to model distance data. Models with acceptable fit based on visual assessment and on goodness-of-fit statistics were ordered based on the Akaike Information Criterion (AIC) values. Analyses were performed using a set of customized functions (mrds v.2.2.0, Laake et al., 2018) in R, version 4.1.1 (R Core Team, 2021).

Uncorrected (for visibility bias and group size bias) density of franciscanas (D_u) was estimated using the Horvitz-Thompson estimator (Marques & Buckland, 2003). Expected mean group size was estimated by dividing density of individuals (D_u) by density of groups (Innes *et al*., 2002). Variance was estimated using the analytical estimator of Innes et al. (2002) and Log-normal 95% confidence intervals were computed as suggested by Buckland et al. (2001). A correction factor for visibility bias (Marsh and Sinclair 1989) and group size bias computed to correct abundance estimates of franciscana from aerial survey data (CF = 4.76, CV = 0.25; Sucunza et al., 2022) was applied to correct the density estimates. The corrected density (D_c) was computed by multiplying the uncorrected estimate (Du) by the CF. Abundance was then estimated as the product of corrected density and the total survey area in each year. Variance for Dc was approximated by the delta method (Seber, 1982)."

Table 1. Summary of models proposed to fit perpendicular distance data for franciscana dolphins in southern Brazil. Hn: half-normal key function, Hr: hazard-rate key function, Δ AIC: Akaike's information criterion differences between the model in question and the most parsimonious model, *wi*: Akaike weight, *P*: overall probability of detection, CV *P*: coefficient of variation of *P*.

Models	Δ AIC	*wi*	*P*	CV *P*
Hr + f(glare)	0.000	0.335	0.723	0.053
Hr	0.845	0.220	0.710	0.058
Hn + f(sea state)	1.498	0.159	0.724	0.056
Hn + f(water transparency)	1.802	0.136	0.712	0.056
Hr + group size	2.696	0.087	0.707	0.058
Hn	3.331	0.063	0.598	0.077

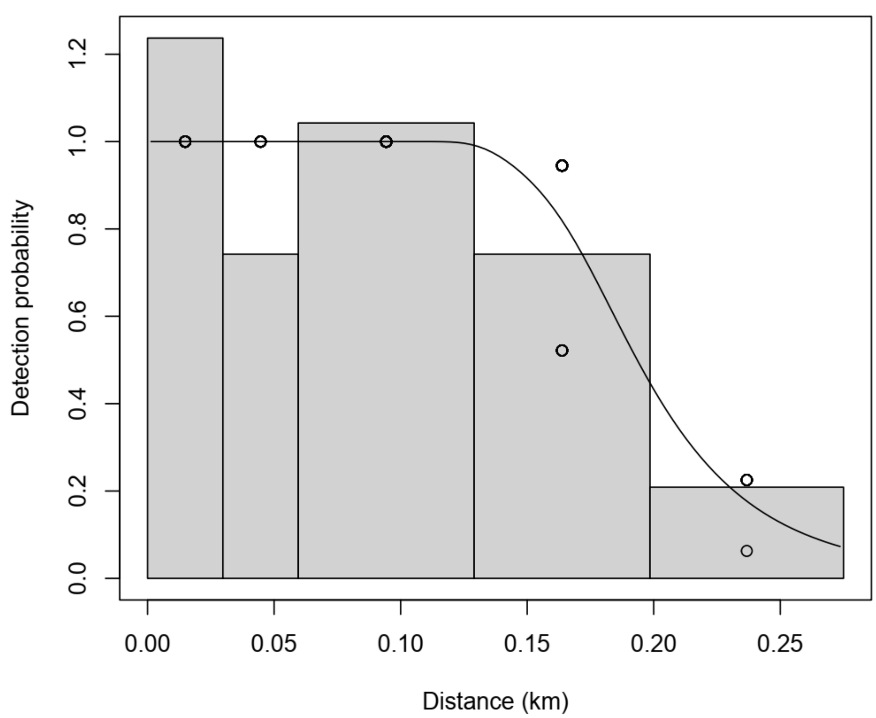

Figure 1. Detection functions for the hazard-rate model with Glare covariate (model #1 Table 1). Km = kilometer.

Table 2. Quantities used for estimation of density of franciscana dolphins in north Rio de Janeiro State (FMA1b), southeast Brazil, in 2011 and 2017. Coefficients of variation and 95% confidence intervals (CI) are shown in parenthesis when applicable.

	2011	2017
Total area (km^2)	4,821	5,731
Survey effort (km)	1,009	2,301
On effort sightings	11	20
Encounter rate (n groups/km)	0.011 (0.46)	0.009 (0.33)
Number of sightings used in fitting the detection function	162	162
Average detection probability (*P*)	0.71 (0.06)	0.71 (0.06)
Average expected group size	2.53 (0.15)	2.51 (0.14)
Average density (*n*_individuals/km^2) corrected for visibility and group size bias	0.33 (0.53)	0.26 (0.47)
Abundance	1,590 (0.53, 95%CI 598 – 4,225)	1,521 (0.47, 95%CI 637 – 3,631)

Reviewer 2

Do the estimates of precision adequately reflect all the major sources of uncertainty?
I think so. Estimation of the detection function was based on data from multiple years and locations. It wouldbe important to confirm that uncertainty associated with variation in factors that vary between locations/yearsis adequately assessed. Region and water transparency did not appear to be significant factors in the model selection exercise, hence I conclude that the authors' approach was suitable.

Authors: see response above.

Reviews FMA II

Relevant documents: Sucunza et al (2020)

Reviewer 1
As ASI/8 Annex 4 is essentially a summary of Sucunza et al (2020) I review these together and refer to Zerbiniet al (2010) where necessary. Where no reference is made, information is taken from ASI/8. Page references (to journal, rather than PDF pages) are to Sucunza et al (2020).

General comments
In Sucunza et al (2020) authors try to address all sources of detection/perception/availability bias using corrections from their own survey and other data/studies. The issue of group size estimation is not addresseddirectly but Zerbini et al (2010) is referenced.

Authors: The revised estimate of abundance for FMA II incorporating the group size bias correct factor was computed as 9,284 individuals (CV = 0.28).

Data

Describes an aerial survey taking place in December 2008 and January 2009 in passing mode in “generally good” weather (Beaufort ≤ 3). Potential issue with survey extent: depths beyond 30m not surveyed (p.425§2.3.1).

Right truncation was performed at the largest distance and left truncation was at 80m in accordance with the field of view. n = 36 + 14 = 41. Number of transects seems confused between the two documents: ASI/8 reports36 tracklines + 13 “transit” lines but Table 1 of Sucunza et al (2020) seems to indicate at least 23+45 (FMA IIa + FMA IIb). Strip transects used for 0 − 80m observations (are these included in the n = 36 here? Based on p.426§2.3.2 I think the answer is no, and that number is 11, but based on p.426§2.3.3 this number is 8?).

Methods

Detection function used binned distances (“Exploratory analyses indicated that the distance data should be grouped (grouping intervals: 0–50m, 51–100m, 101–150m, 151–200 m)”, p.426§2.3.2). I assume this was due to heaping, in which case it’s curious that the cutpoints are at such round numbers, as one might expect the observers to round to the nearest 50m (even +80m, which I assume was already subtracted), so non-multiples of10, with more bins closer to 0m would be preferred. Plots show histograms after this transformation and Figure3 of Zerbini et al (2010) uses the same binning for plotting at least, so I can’t see what’s going on here. Withoutfurther explanation as to why this happened it is hard to infer if this would make any difference to the results.

Authors: Preliminary results provided by Zerbini et al. (2010) and ASI/8 differ slightly because that dataset was refined and re-analyzed, and the results presented in Sucunza et al. 2020 supersedes those from this preliminary report.

Reviewer 2

Methods

Overall, this is a straightforward re-analysis of existing data with similar methods applied to new boundaries. The post-stratification is justified and leaves a sufficient number of transect lines and sightings in each sub-area for robust inference. As in Zerbini et al. (2010), a 80 m left-truncation was used to make the two observer positions (front with bubble windows and rear with flat windows) more comparable. One difference from the previous analysis is that here, the distance data were grouped into 4 intervals “to achieve better model fits”. It is unclear why this was necessary when the fits in the (almost identical) dataset in Zerbini et al. (2010) seemed satisfactory without this binning (binning was used in the plots but not mentioned in the text). The actual goodness-of-fit are not available to check this statement, but the resulting CVs around detection probability P are actually higher here than they were in the previous analysis, and represent a large proportion of the total variance (which is often dominated by the encounter rate component).

Authors: it was done to heaping.

Another difference is that the 8 sightings made between 0 and 80 m of the trackline were analyzed as a strip transect, using the estimated probability of detection at 80 m in the distance analysis (i.e., assuming a broad, flat shoulder in the part of the detection curve that was not modelled). It is not clear to me why this approach was necessary (rather than simply applying the density estimated in the 80 – 280 m range to the entire area), since the assumption of equal detection probability of sightings within the left-truncation distance does not always hold true in aerial surveys, even with bubble windows, because of the large apparent speed at which sightings pass by underneath the aircraft (something that is actually observed in the similar survey in Babitonga Bay, figure 2 in Sucunza et al. 2018). In any case, if this approach is kept, it seems to me that the value of P(0) used to correct for perception bias should not be the value given by the MR analysis for combined observers (i.e., something around 0.86?) but rather that of the "observer 1" only (usually output by the *mrds* function but not reported here), since these 8 sightings could only be made by the front observers by definition. The potential biases from these two issues would lead to underestimating abundance.

Authors: the approach of include the sightings made between 0 and 80 m was proposed to increase sample size and better estimate uncertainty. However, we agree that we could have been used the independent estimates of P(0) to each one of the front observers instead of the average platform P(0).

General

Clarification on the issues mentioned above would be useful, but the conclusions are generally supported by theresults. In this study (Sucunza et al. 2020), the correction factor for perception bias is estimated from the data and the correction factor for availability is taken from ASI/4_rev1 (reviewed elsewhere), which is appropriate. However, the 33% group size correction was not applied to the estimate in ASI/8 and there is no clearrecommendation as to whether it should be used.

Authors: The revised estimate of abundance for FMA II incorporating the group size bias correct factor was computed as 9,284 individuals (CV = 0.28).

Review FMA III

Relevant document: SC/68B/ASI/06

Reviewer 1

Methods

- They binned data into what appears to be appropriate bins, but they did not explain why they had to bin thedata in the first place.

Authors: it was done to heaping.

- They only explored the half-normal function, which looks like it fits ok, but it is curious why they did not try the other standard function: the hazard rate function.

They modeled averaged the three models they appeared to try, which is appropriate since the AIC's were soclose. Did they try water transparency as a potential covariate? What was the distribution of group sizes? Maybe pooling the couple large groups would have been better than using group size as a continuous covariate?

Authors: the dataset was refined and re-analyzed and new results are presented here.

"*Line transect analysis methods*

Detection probability was estimated using Conventional (CDS) and Multiple Covariate Distance Sampling (MCDS) methods (Buckland et al. 2001, Marques & Buckland 2003). Due to a spike in sightings data adequate fits were obtained by modeling grouped distance data (grouping intervals: 0 – 21m, >21 – 59m, >59 – 129m, >129 – 198m, >198m – 275m). Only the half-normal and the hazard-rate detection functions were proposed to fit distance data. Beaufort sea state (factor covariate with two levels: "calm", Beafourt sea state between 1 and 2, and "high" between 3 and 4) and group size (factor covariate with two levels: "small", between 1 and 2 individuals, and "large" between 3 and 4 individuals) were considered as covariates to model distance data. Models were ranked according the Akaike Information Criterion (AIC), and model averaging were performed to incorporate unconditional model selection variance in the estimates and confidence intervals (Burnham & Anderson 2002). Analyses were performed using a set of customized functions (mrds v.2.2.0, Laake et al., 2018) in R, version 4.1.1 (R Core Team, 2021). Only data recorded by the front observers in the airplane (bubble windows) were considered in the analysis presented in this study because of the field of view between front and rear observers only partially overlapped.

Abundance Estimation

Density and abundance of franciscanas (Du - density uncorrected for visibility bias and group size bias) were estimated using the Horvitz-Thompson-like estimator (Borchers et al. 1998, Borchers & Burnham 2004). Expected mean group size was obtained as suggested by Innes et al. (2002) and Marques and Buckland (2003). Variance was estimated using the analytical estimator of Innes et al. (2002) and log-normal 95% confidence intervals (Buckland et al. 2001) were computed after unconditional variance was derived (Zerbini et al. 2006).
A correction factor for visibility bias (Marsh & Sinclair 1989) and groups size bias computed to correct abundance estimates of franciscana from aerial survey data (CF = 4.76, CV = 0.25; Sucunza et al., 2022) was applied, and the corrected density estimate (Dc) was computed by multiplying the uncorrected estimate (Du) by CF. Abundance was then estimated as the product of the corrected density and the total area. Variance of Dc was computed by the Delta method (Seber 1982).

Results were presented in Table 3 and 4.

Table 3. Summary of models proposed to fit perpendicular distance data for franciscana dolphins in southern Brazil. Hn: half-normal key function, Hr: hazard-rate key function, Δ AIC: Akaike's information criterion differences between the model in question and the most parsimonious model, *wi*: Akaike weight, *P*: overall probability of detection, CV *P*: coefficient of variation of *P*.

Models	Δ AIC	*wi*	*P*	CV *P*
Hn	0.000	0.324	0.686	0.149
Hr	1.291	0.170	0.681	0.157
Hn + f(sea state)	1.353	0.165	0.771	0.125
Hn + group size	1.458	0.156	0.794	0.103
Hr + group size	1.911	0.125	0.685	0.152
Hr + f(sea state)	3.344	0.061	0.767	0.128

Table 4. Quantities used for estimation of density and abundance of franciscana dolphins in southern Brazil, in March 2014. Coefficients of variation are shown in parenthesis when applicable.

Parameters	Results
Survey effort in kilometer - km	3,853
On effort sightings	46
Encounter rate (*n* groups/km)	0.012 (0.178)
Average detection probability	0.721 (0.138)
Average expected group size	2.124 (0.066)

Average uncorrected density (n individuals/km^2)	0.064 (0.231)
Density corrected for visibility and group size bias	0.306 (0.342)
Abundance corrected for visibility and group size bias	9,437 (0.342)

"

Reviewer 2

Have the data used in the analysis, including any pre-selection process, been clearly specified?
The document does not provide much information on data pre-selection (e.g., whether there were off effort/transit sightings and whether/how they were included in the analysis).

Authors: see response above.

Have the methods used been adequately documented?

Model selection was performed using AIC and model averaging was used to incorporate model uncertainty in the estimation of variance. That's good practice, especially if models have similar AIC values. No information on model fit is provided but examination of the plot in Fig. 3 suggests a somewhat reasonable fit. A different binning strategy could potentially minimize discrepancies in the number of sightings per bin and perhaps improve the fit.
Density and abundance corrected for perception, availability and group size bias were computed using the correction factor computed by Sucunza et al., 2020 (CF = 4.42) and variances were computed using the delta method. This seems appropriate, but a revision of the correction factor will be performed by the ASI working group so further discussion may be needed here.

Authors: see response above.

References

<u>*SC/68B Documents*</u>

ASI/04rev1. SUCUNZA, F., DANILEWICZ, D., SECCHI, E.R., ANDRIOLO, A., CREMER, M., FLORES, P.A.C., FERREIRA, E., DE S. ALVES, L.C.P., DE CASTRO, F., PRETTO, D., SARTORI, C.M., SCHULZE, B., DENUNCIO, P., PEREZ, M.S. AND ZERBINI, A.N.
Assessing bias in aerial surveys for threatened cetaceans: results from experiments conducted with the Franciscana (*Pontoporia blainvillei*). 33pp.

ASI/05. SUCUNZA, F., DANILEWICZ, D., OTT, P.H., NEVES, M., BERCHIERI, N., FARRO, A., P., MARTINS, A. AND ZERBINI, A.N.
Population size and IUCN Red Listing of the isolated population of the franciscana (*Pontoporia blainvillei*). 25pp.

ASI/06. SUCUNZA, F., DANILEWICZ, D., CREMER, M., FERREIRA, E. AND DENUNCIO, P. Abundance of the endangered franciscana in southern Brazil. 16pp.

ASI/07rev1. DANILEWICZ, D., SUCUNZA, F., OTT, P.H., FERREIRA, E., PEREZ, M.S., BERCHIERI, N., ALVARES, D., ANDRIOLO, A.,
SECCHI, E.R., FLORES, P.A.C., FARRO, A.P., MARTINS, A. AND ZERBINI, A.N. Abundance and distribution of franciscanas (*Pontoporia blainvillei*) in northern Rio de Janeiro (FMA Ib), Brazil. 15pp.

ASI/08. ZERBINI, A.N., ANDRIOLO, A., CREMER, M., CRESPO, E.A., DANILEWICZ, D., DOMIT, C. AND SUCUNZA, F. Abundance estimate of franciscana dolphins (*Pontoporia blainvillei*): a review and future recommendations. 36pp.

Former SC Documents

Zerbini, A.N., Secchi, E.R., Danilewicz, D., Andriolo, A., Laake, J.L. and Azevedo, A. 2010. Abundance and distribution of the franciscana (*Pontoporia blainvillei*) in the Franciscana Management Area II (southeastern and southern Brazil). Paper SC/62/SM7 presented to the IWC Scientific Committee, June 2010, Agadir, Morocco (unpublished). 14pp. [Paper available from the Office of this Journal].

Zerbini, A.N., Danilewicz, D., Secchi, E.R., Andriolo, A., Cremer, M., Flores, P.A.C., Ferreira, E., Alves, L.C.P.d.S., Sucunza, F., de Castro, F.R., Pretto, D., Sartori, C.M., Schulze, B., Denuncio, P. and Laake, J.L. 2011. Assessing bias in abundance estimates from aerial surveys to improve conservation of threatened franciscana dolphins: preliminary results from a survey conducted off southern Brazil. Paper SC/63/SM9 presented to the IWC Scientific Committee, June 2011, Tromsø, Norway (unpublished). 13pp. [Paper available from the Office of this Journal].

Published References

Bordino, P., Thompson, G. and Iñíguez, M. 1999. Ecology and behaviour of the franciscana (*Pontoporia blainvillei*) in Bahía Anegada, Argentina. *Journal of Cetacean Research and Management* 1(2): 213-22.

Sucunza, F., Danilewicz, D., Andriolo, A., Azevedo, A.F., Secchi, E.R. and Zerbini, A.N. 2020. Distribution, habitat use, and abundance of the endangered franciscana in southeastern and southern Brazil. *Marine Mammal Science* 36(2): 421-35.
Sucunza, F., Danilewicz, D., Cremer, M., Andriolo, A. and Zerbini, A.N. 2018. Refining estimates of availability bias to improve assessments of the conservation status of an endangered dolphin. *PLOS ONE* 13(3): e0194213.

Sucunza, F., Danilewicz, D., Andriolo, A., de Castro, F.R., Cremer, M., Denuncio, P., Ferreira, E., Flores, P.A.C., Ott, P.H., Perez, M.S., Pretto, D., Sartori, C.M., Secchi, E.R., Zerbini, A.N. 2022. Assessing bias in aerial surveys for cetaceans: Results from experiments conducted with the franciscana dolphin. Front. Mar. Sci. 9:1016444. doi: 10.3389/fmars.2022.1016444