CCSBT-ESC/0709/12

Aerial Survey: updated index of abundance and preliminary results from calibration experiment

Paige Eveson
Mark Bravington
Jessica Farley

Prepared for the CCSBT 8th Meeting of the Stock Assessment Group and the 12th Meeting of the Extended Scientific Committee
4-8 September, and 10-14 September 2007, Hobart, Australia

Table of Contents

List of Tables

List of Figures

Abstract

The preliminary point estimate from the 2007 scientific aerial survey is lower than the 2006 estimate, and well below the average level in the mid-1990s. Although the trend appears to have continued downwards, the high uncertainty in recent estimates (due primarily to low search effort and numbers of sightings) means that we cannot yet be sure that this trend is not just due to chance fluctuation.

The aerial survey analysis methods were the same as those used last year. A random-effects extension of the models is still being investigated, which should better handle strata where there is low effective sampling effort. Preliminary results suggest that the relative abundance indices obtained using the random effects models will be very similar to those obtained using the current methods, but that we may be able to achieve higher precision (i.e., smaller CVs).

A large-scale calibration experiment was undertaken in 2007, with the primary purpose of comparing SBT sighting rates by one observer versus two observers in a plane (in light of the fact that future surveys may have only one observer in a plane). An initial investigation of the data has brought up a number of potential issues for analysis. Analysis of the data will be completed prior to the 2008 survey, as it is not required for analysis of the 2007 data.

Introduction

The index of juvenile Southern Bluefin Tuna (SBT) abundance based on a scientific aerial survey in the Great Australian Bight (GAB) is one of the very few fishery-independent indices available for assessment and management purposes. The aerial survey was conducted in the GAB between 1991 and 2000, but was suspended in 2000-01 due to logistic problems of finding trained, experienced spotters and spotter-pilots. (Note that the terms 'spotter' and 'observer' are interchangeable). The suspension also allowed for further data analysis and an evaluation of the effectiveness of the survey. Analysis of the data was completed in 2003 and it showed that the scientific aerial survey does provide a suitable indicator of SBT abundance in the GAB (Bravington 2003). In the light of serious concerns about the reliability of historic and current catch and CPUE data, this fishery-independent index is even more important.

In 2005 and 2006, a full scientific line-transect aerial survey in the GAB was re-established. New analysis methods were developed, and all data were re-analysed. Based on these analyses, an index of abundance was constructed for the surveys in 1993-2000 and in 2005-2006.

Recognising the importance of having a time-series of indicators, the full scientific line-transect aerial survey continued in 2007.

In addition, a large-scale calibration experiment was conducted using a second plane containing a third observer, running in parallel with the main survey. The calibration experiment was considered particularly important in the light of the over-catch issues noted above, and the potential requirement to continue the survey into the future. This experiment was necessary to:

(1) compare SBT sighting rates by one observer versus two observers in a plane - in light of the pending retirement of the only experienced spotter-pilot available for the survey (i.e., in case future surveys have only one observer in a plane), and

(2) provide additional data on observer variability to help reduce this source of uncertainty in the aerial survey indices.

This report summarises the field procedures and data collected during the 2007 season for the scientific line-transect aerial survey and the calibration experiment. Preliminary results for the aerial survey index of abundance are presented, but analyses of the calibration experiment data are still underway.

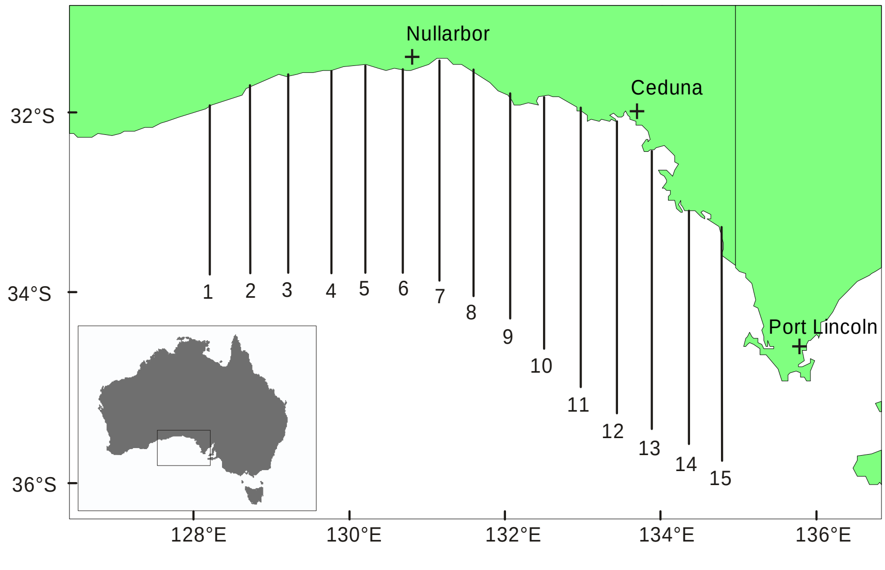
Line-transect aerial survey

Field procedures

The line-transect aerial survey was conducted in the GAB between January and March 2007, although one additional line was surveyed on 2 April because the weather conditions were suitable. As in previous years, the plane used in the 2007 survey was a Rockwell Aero Commander 500S. Two observers were employed for each survey flight, one spotter-pilot and one spotter. The same spotter-pilot employed for the 2005 and 2006 surveys was also used throughout the 2007 survey. This year, two spotters were employed as the second observer in the survey plane, one of whom was used in both the 2005 and 2006 surveys. Each spotter participated in approximately half of the line transect survey flights.

The line-transect survey followed the protocols established for the 2000 aerial survey (and used in all subsequent surveys) regarding the area searched, plane height and speed, environmental conditions, and time of day the survey was conducted (Cowling 2000). The survey area lies between 128°E and 135°E, running from the coast to just off the continental shelf. Fifteen north-south transect lines (Figure 1) were searched by the observers. A complete replicate of the GAB consists of only 12 lines divided into 4 blocks. The remaining 3 lines in a replicate (either {1, 3 and 14} or {2, 13 and 15}) were not searched to save on time, and SBT abundance is historically low in these areas. The blocks were flown from west to east, and the lines within each block were flown in a pre-set order (sequence and direction).

Figure 1. Location of the 15 transect-lines for the scientific aerial surveys in the GAB.



The survey was only conducted on days when the environmental conditions were met. The minimum environmental conditions required were: less than 1/3 cover of cloud at or below1500 ft, visibility at 1500 ft must be greater than 7 nautical miles (nm), and wind speed at the sea surface must be 8 knots or less. However, once the survey had started, it continued as long as the wind speed did not exceed 10 knots.

A Garmin 176 GPS was used to log the position of the plane (15 second intervals) and waypoints during the survey. Transects are flown at 120 knots and at an altitude of 1500 ft. Each observer searched the sea surface from straight ahead through to 90° on their side of the plane (abeam of the plane) for surface patches (schools) of SBT. Occasionally the observer would search both sides of the plane if the other observer was unable to observe. A data recorder sat behind them recording environmental and sighting information in a logbook, and monitored the GPS.

When a sighting of SBT was made, a waypoint position (and time) was recorded in the GPS. The plane continued along the transect line until the observer judged that the sighting was at 90° to the plane. At that point, the plane left the transect line and flew directly to the sighting and circled it. Each sighting can contain one or more schools (or patches) of SBT. The two observers independently estimated a range for the size of individual fish in each school (in kg) and the size of each school (in tonnes). Another waypoint was recorded over the school, and then the plane flew back to the point it left the transect to resume searching. For each sighting of SBT, the behaviour of the fish was recorded as "deep" or "shallow" and "feeding", "rippling" or "fattening". Information was only collected on those sightings for which some part of the grouping was within 7 nm of the transect line. While flying out to a sighting, the observers refrained from looking at the areas that had not yet been searched. This reduced the possibility of additional (secondary) sightings. If secondary sightings were made when flying off the transect, they were only recorded if they were within the 7 nm limit, and were in areas not already searched. If a recorded secondary sighting could be seen from the transect (when the plane returned), then it was changed to be a primary sighting. Only primary sightings were included in the analysis (i.e., original sightings made from the transect and secondary sightings that could be seen from the transect on return).

Environmental observations were recorded at the start and end of each transect and at 30 minute intervals during the transect flight, or when the conditions changed significantly. The observations include wind speed and direction, air temperature, amount of high and low cloud, glare, haze and swell.

The line-transect survey was successfully completed in 2007. Approximately 35 lines were completed, compared to 44 in 2005 and 38 in 2006. The reduced number of lines surveyed this year was due to the poor weather conditions experienced in February and March. The total flying time (transit and transect time) for the 2007 survey was 151.3 hours.

Data preparation

The data collected from the 2007 survey were loaded into the aerial survey database, which contains the data collected from all surveys. The 2007 data were checked for any obvious errors or inconsistencies and corrections were made where necessary.

In order for the data and results to be comparable between all survey years, only data collected in a similar manner from a common area were included in the data summaries and analyses presented in this report. In particular, only search effort and sightings made along north/south transect lines in the unextended (pre-1999) survey area, and only sightings made within 6 nm of a transect line, were included (refer to the 2005 Final Report to DAFF for details; Basson et al., 2005). Note that if a sighting consisted of more than one cluster, then the sighting was included if at least one of the clusters was within 6 nm of the line. In accordance with last year's analysis, we excluded secondary sightings and any search distance and sightings made during the aborted section of a transect line (see Eveson et al., 2006, for more details).

Search effort and SBT sightings

A summary of the total search effort and SBT sightings made in each survey year is given in Table 1. These numbers, as well as all summary information and results presented in this report, include only the data outlined in the previous section as being appropriate for analysis.

The total distance searched has been continually declining since 1994, and was below 5000 nm in 2007. The sightings rate (number of sightings per nm) was very similar in 2007 as in 2006, but the amount of biomass per nm has declined slightly due to average patch size declining (Figure 2; Table 1). The distribution of SBT sightings in 2007 was fairly clustered in the eastern-central part of the survey area (transect lines 8-12), but the number of sightings is too sparse to make meaningful comparisons with previous years (Figure 3).

Table 1. Summary of aerial survey data by survey year. Only data considered suitable for analysis (as outlined in text) are included. All biomass statistics are in tonnes.[1]

Survey year	Total distance searched (nm)	Number SBT sightings	Total biomass	Average patches per sighting	Max patches per sighting	Average biomass per patch	Median biomass per patch	Max biomass per patch
1993	7603	130	12225	3.9	76	24.4	18.8	203
1994	15169	174	15010	3.3	23	26.4	21.5	245
1995	14573	179	21971	3.6	38	34.6	27.9	224
1996	12284	116	16487	4.1	46	34.6	27.3	147
1997	8813	117	9790	3.0	18	27.6	22.2	197
1998	8550	109	10226	2.3	21	40.3	20.3	943
1999	7555	56	3024	2.4	21	22.9	16.5	120
2000	6775	77	4817	2.6	17	24.0	20.0	100
2005	5968	80	6162	2.4	17	32.1	25.1	198
2006	5152	44	4095	2.0	8	47.6	32.2	272
2007	4870	42	3491	2.6	11	32.6	24.2	124

The average and maximum number of patches per sighting were up slightly from 2006, but still lower than in early survey years (Table 1). The size of a patch tended to be typical of previous years, with the average and median being very close to the average of these statistics from previous years (Table 1; Figure 4).

Estimates of the average size of SBT within a patch have been found to be inconsistent between different observers (Cowling et al. 2002), however we assume that fish size estimates should be comparable for the same observer. Therefore, as in the 2005 and 2005 Final Reports, we consider the fish size estimates made by a single observer who has operated in all survey years, including 2007. According to this observer, the SBT present in the survey area in 2007 were larger on average than in 2006, but close to the average size over all previous survey years (Figure 5).

[1] The biomass statistics differ very slightly from those reported in Table 2 of the 2006 Final Report to DAFF because the patch size estimates used in calculating these statistics have been corrected for differences between observers (see “Methods of analysis” section). Observer differences are re-estimated each year using all available data and thus the corrected patch size estimates can change slightly.

Figure 2. Plots of a) total distance searched (i.e. effort) by year; b) biomass per mile by year; c) number of sightings per 100 miles by year. *Note that these plots are based on raw data, which has not been corrected for environmental factors or observer effects.*

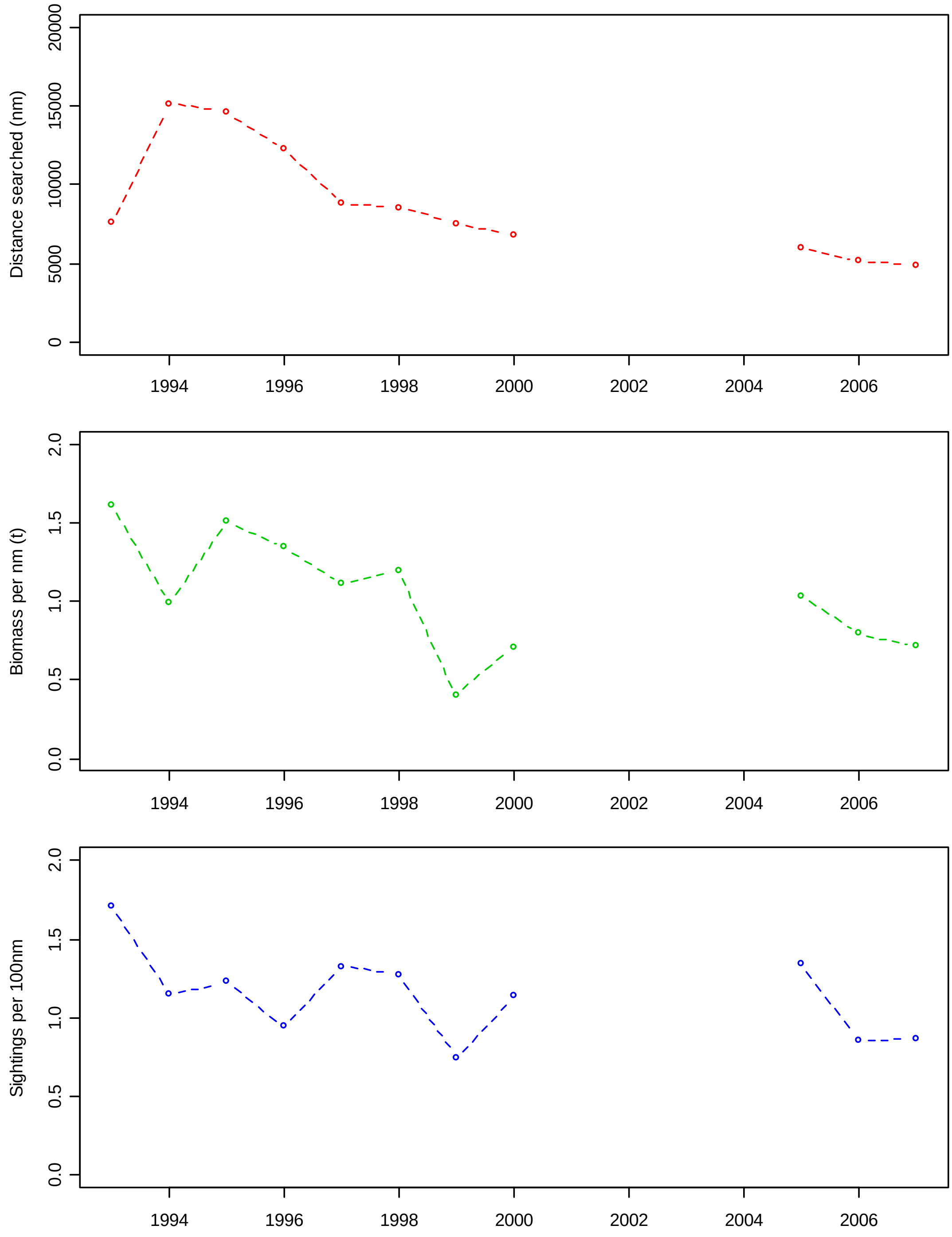


Figure 3. Distribution of SBT sightings made during each aerial survey year. Red circles show the locations of SBT sightings, where the size of the circle is proportional to the size of the sighting, and grey lines show the north/south transect lines that were searched.

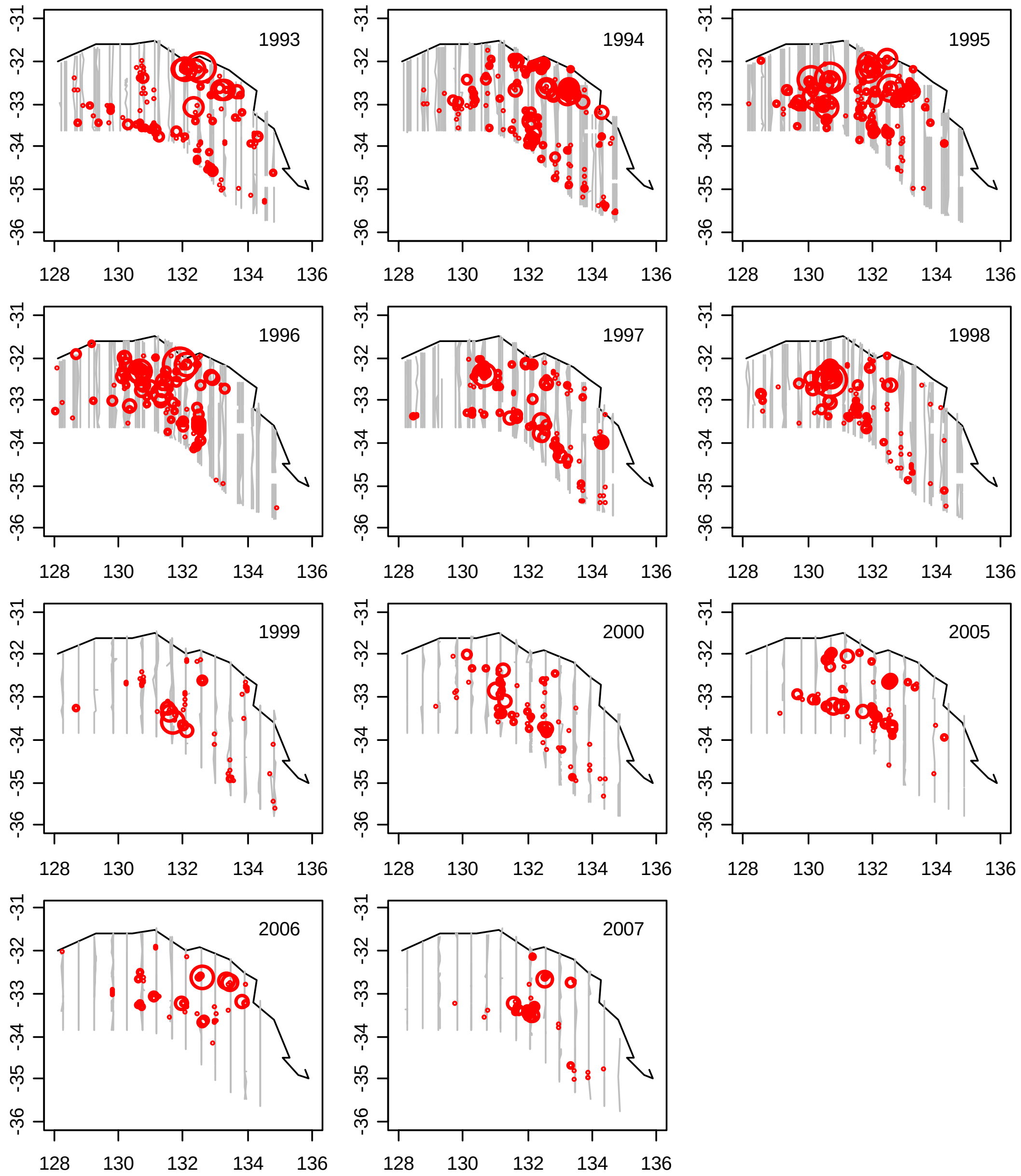


Figure 4. Frequency of SBT patch sizes (in tonnes) by survey year.

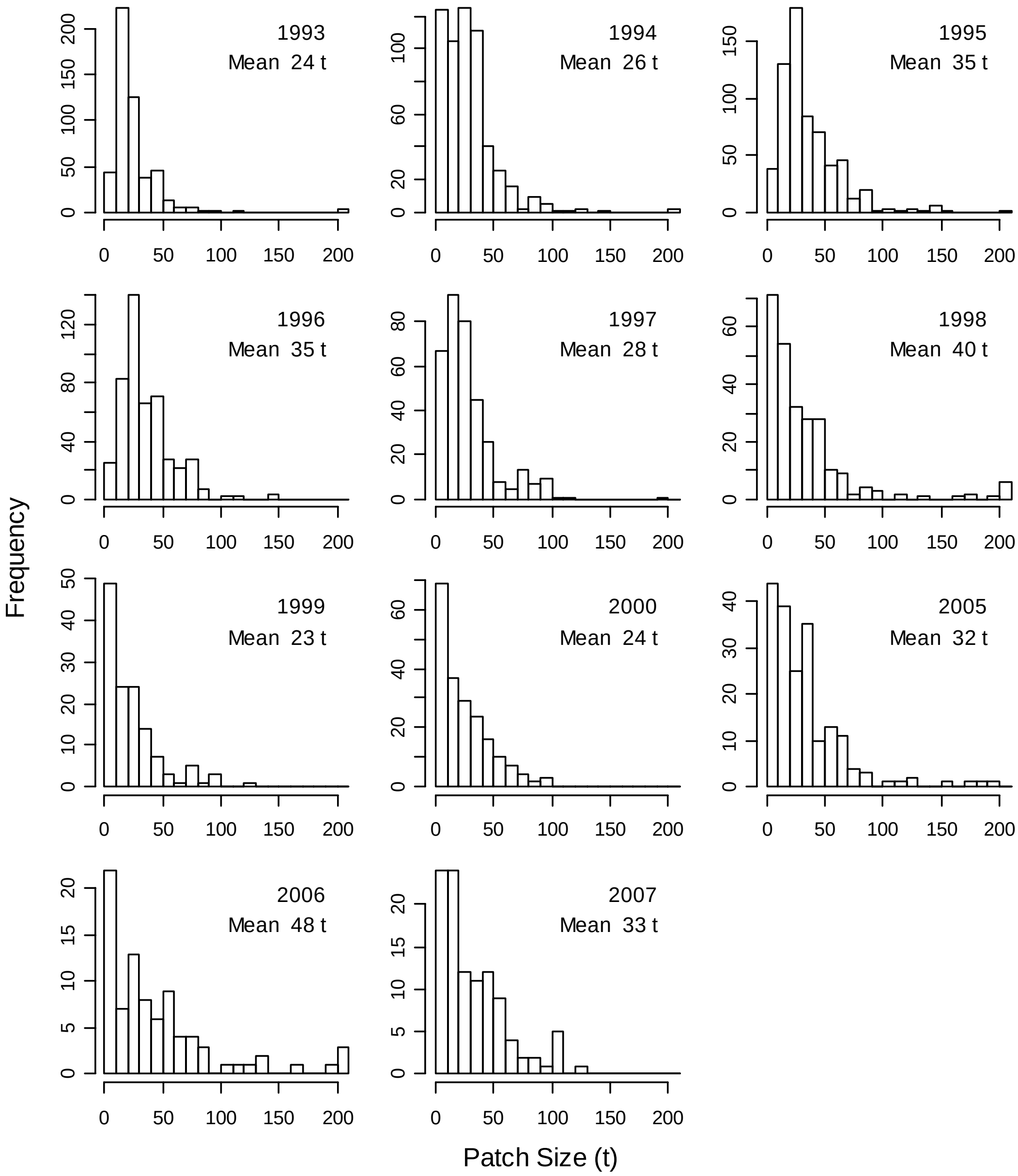
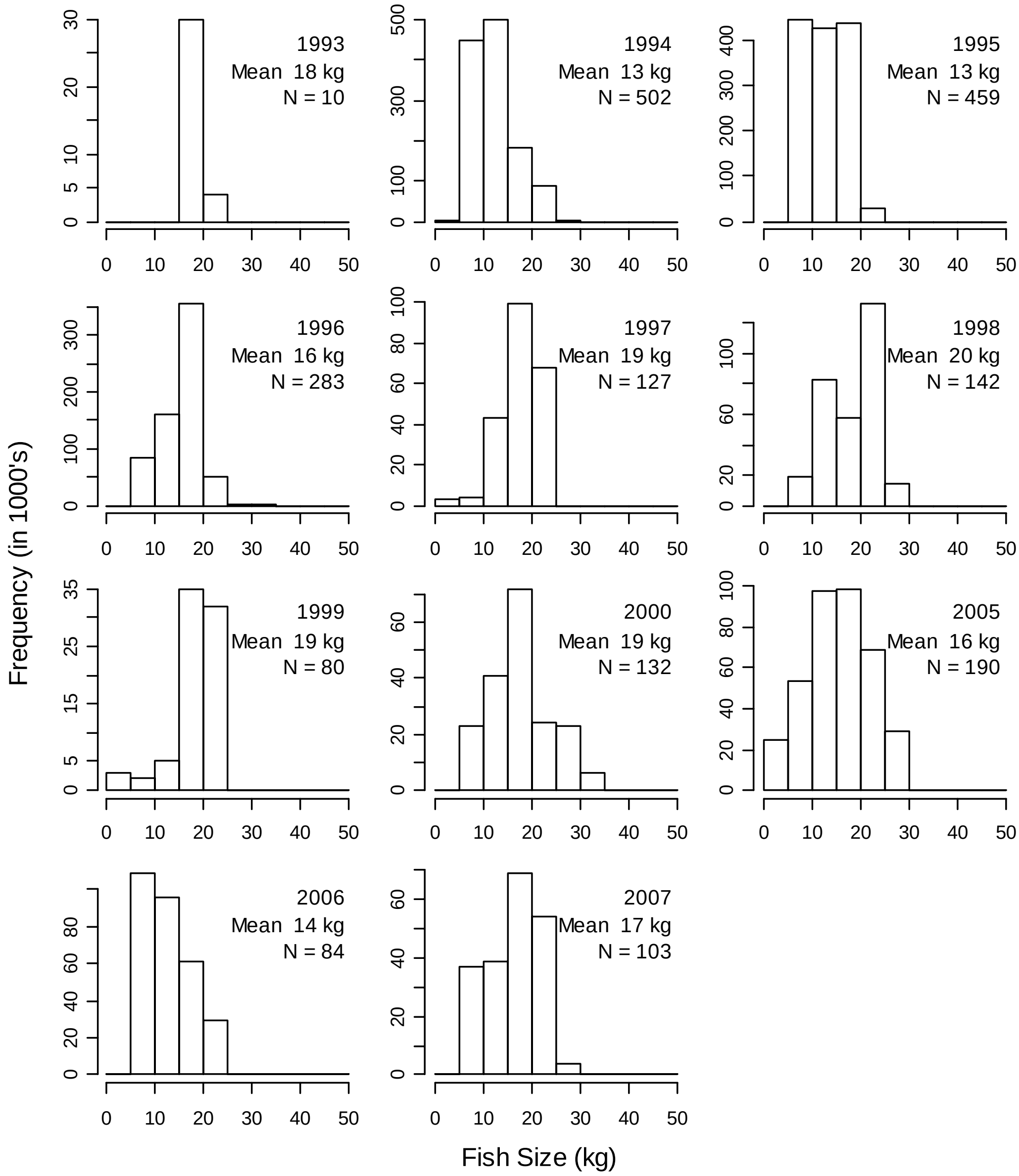


Figure 5. Size frequency of SBT (in kg) recorded by one observer who operated in all survey years. Data are weighted by patch size. N = number of patches for which fish size was estimated.



Environmental variables

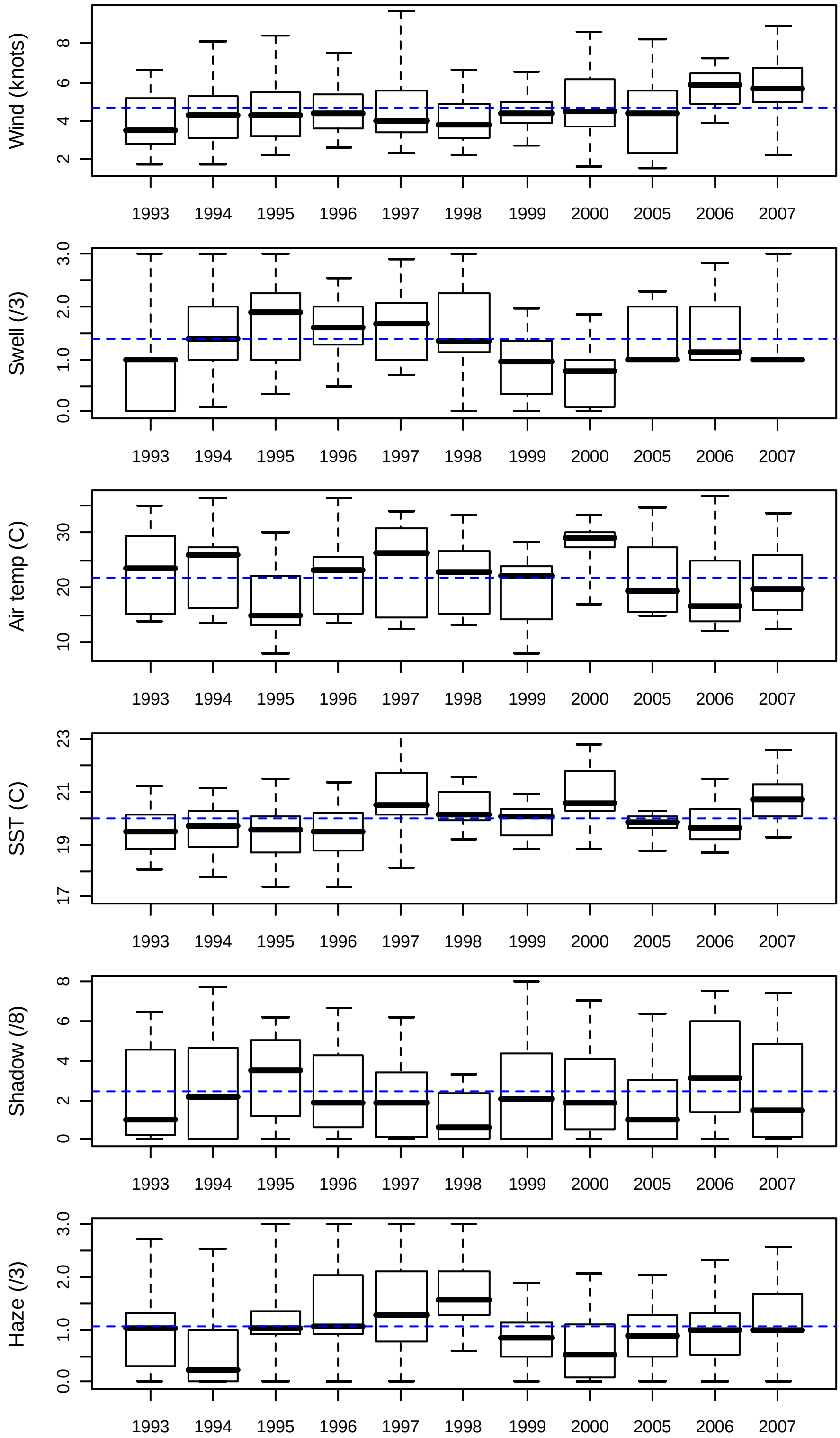
Table 2 and Figure 6 summarize the environmental conditions that were present during valid search effort in each survey year. All the environmental variables presented were recorded by the survey plane, with the exception of sea surface temperature (SST), which was extracted from the 3-day composite SST dataset produced by CSIRO Marine and Atmospheric Research's Remote Sensing Project (see Eveson et al. 2006 for more details).

Notably, the wind speed during the 2007 survey was very high on average compared to other survey years, and was the reason for the low amount of search effort that could be completed. SST also tended to be higher than average. In terms of the expected number of sightings, these variables have opposing effects: high winds tend to mean fewer sightings (either due to reduced sightability or reduced surfacing behaviour), whereas high SST tends to mean more sightings (presumably due to increased surfaced behaviour). Increases in SST also tend to be associated with larger sightings (i.e., greater biomass per sighting).

Table 2. Average environmental conditions during search effort for each aerial survey year.

Survey year	Wind speed (knots)	Swell height (0-3)	Air temp (°C)	SST (°C)	Sea shadow (0-8)	Haze (0-3)
1993	3.9	0.8	24.4	19.6	1.8	0.9
1994	4.1	1.5	20.6	19.7	2.7	0.5
1995	4.4	1.7	18.7	19.6	2.7	1.1
1996	4.5	1.6	22.9	19.6	2.1	1.2
1997	4.1	1.7	25.3	21.1	1.6	1.3
1998	3.7	1.7	22.3	20.4	0.9	1.7
1999	4.1	0.9	22.0	19.9	2.9	0.7
2000	4.3	0.6	27.5	20.7	2.6	0.7
2005	4.7	1.5	21.7	19.8	1.6	0.8
2006	5.6	1.5	20.0	19.9	3.5	1.0
2007	5.8	1.3	21.6	20.8	2.0	1.3

Figure 6. Boxplots summarizing the environmental conditions present during valid search effort for each aerial survey year. The thick horizontal band through a box indicates the median, the length of a box represents the inter-quartile range, and the vertical lines extend to the minimum and maximum values. The dashed blue line running across each plot shows the overall average across all survey years.



Methods of analysis

We applied the same analysis methods as last year to the updated data. Details of the methods can be found in Appendix A. In summary, separate models were constructed to describe two different components of observed biomass: i) biomass per patch sighting (BpS), and ii) sightings per nautical mile of transect line (SpM). Each model component was fitted using a generalized linear model. Since environmental conditions affect what proportion of tuna are available at the surface to be seen, as well as how visible those tuna are, and since different observers can vary both in their estimation of school size and in their ability to see tuna patches, the models include 'corrections' for environmental and observer effects in order to produce standardized indices that can be meaningfully compared across years.

For the purpose of analysis, we defined 45 area/month strata: 15 areas (5 longitude blocks and 3 latitude blocks; see Figure A1 in Appendix A) and 3 months (Jan, Feb, Mar[2]). The SpM and BpS model results were used to predict what the number of sightings per mile and the average biomass per sighting in each of the 45 area/month strata in each survey year would have been under standardized environmental/observer conditions. Using these predicted values, we calculated an abundance estimate for each stratum as 'standardized SpM' multiplied by 'standardized average BpS'. We then took the weighted sum of the stratum-specific abundance estimates over all strata within a year, where each estimate was weighted by the geographical size of the stratum in nm^2, to get an overall abundance estimate for that year. Lastly, the annual estimates were divided by their mean to get a time series of relative abundance indices.

It is important to have not only an estimate of the relative abundance index in each year, but also of the uncertainty in the estimates. To calculate CVs for the indices was a fairly involved process, the details of which can be found in Appendix B. Briefly, we first obtained standard errors (SEs) for the predicted values of 'standardized SpM' and 'standardized average BpS' in each year/area/month stratum. These were used to calculate SEs for the stratum-specific abundance estimates, which were in turn used to calculate SEs for the annual abundance estimates. Lastly, we applied the delta method to determine SEs for the relative abundance indices. Note that CVs are given simply by dividing the SE of each index estimate by the estimate. We calculated confidence intervals for the indices based on the assumption that the logarithm of the indices follow a normal distribution, with standard errors approximated by the CVs of the untransformed indices (as calculated above).

As recommended last year (Eveson et al., 2006), we have been investigating using a random-effects extension of the BpS and SpM models to better handle strata where there is low effective sampling effort. Preliminary results suggest that the relative abundance indices obtained using the random effects models will be very similar to those estimated using the methods described above, but that we may be able to achieve higher precision (i.e., smaller CVs). For next year, we hope to present results that not only incorporate random effects, but also formally merge the estimation of observer effects into the models to correctly propagate all the uncertainty into the final CVs. Currently, variation between observers remains the most troublesome source of uncertainty in the aerial survey indices.

Results

Figure 7 shows the estimated time series of relative abundance indices with 90% confidence intervals. The point estimates and CVs corresponding to Figure 7 are reproduced in Table 3. The point estimate for 2007 is lower than the 2006 estimate, and well below the average level

[2] Note that the one flight on 2 April 2007 was included as part of March.

in the mid-1990s. The trend appears to have continued downwards, however the high uncertainty in recent estimates (due primarily to low search effort and numbers of sightings) means that we cannot yet be sure that this trend is not just due to chance fluctuation.

Figure 7. Time series of relative abundance estimates with 90% confidence intervals.

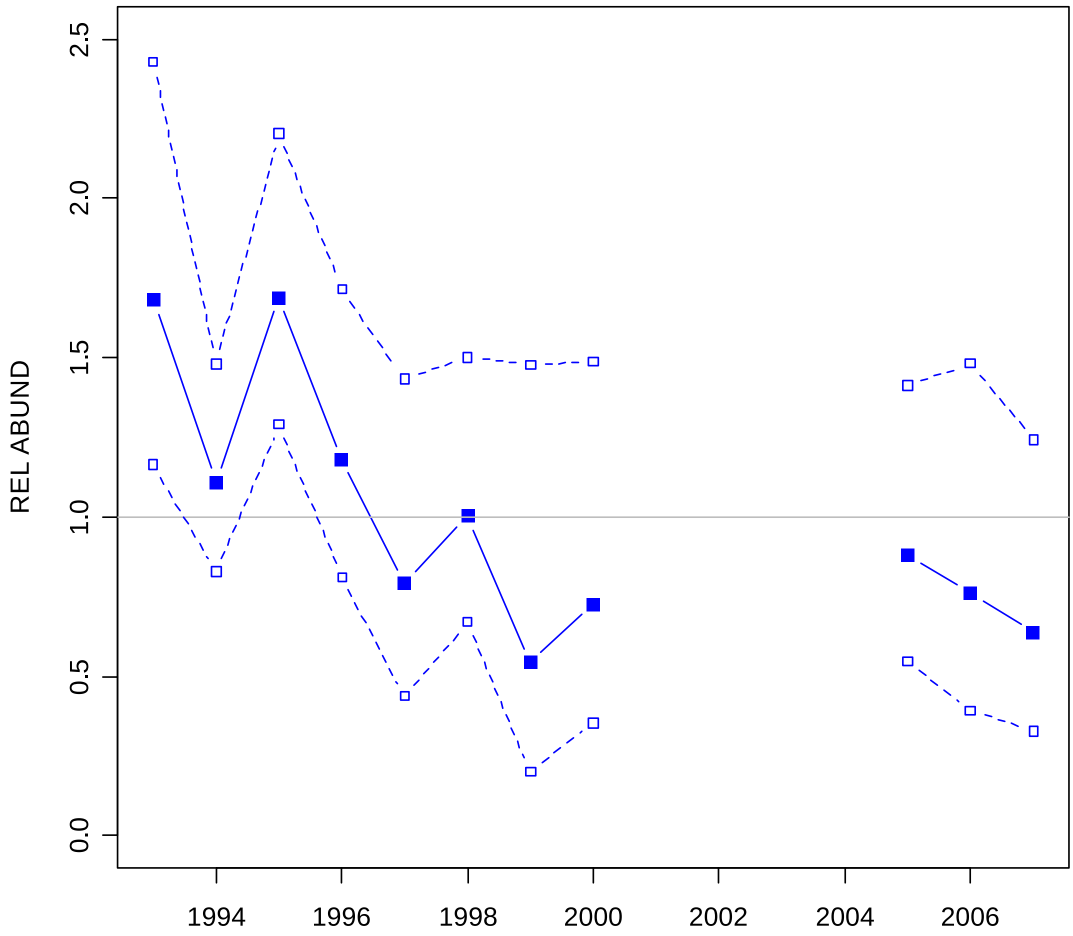


Table 3. Aerial survey relative abundance point estimates, standard errors and CVs (as per Figure 7).

	1993	1994	1995	1996	1997	1998	1999	2000	2005	2006	2007
Estimate	1.681	1.108	1.688	1.180	0.793	1.003	0.543	0.726	0.880	0.761	0.637
SE	0.375	0.195	0.274	0.269	0.285	0.246	0.330	0.316	0.254	0.308	0.259
CV (%)	22.3	17.6	16.2	22.8	36.0	24.5	60.9	43.6	28.8	40.5	40.6

Calibration experiment

Field procedures

The second component of the 2007 program was to undertake a large-scale calibration experiment to compare the sighting rates of one observer (spotter) versus two observers in a plane (in case future surveys have only one observer in a plane), and to collect additional data on observer variability to help reduce this source of uncertainty in the aerial survey indices.

The experiment involved a second plane with only one spotter (and a non-spotting pilot) flying simultaneously with the survey plane, and recording potential SBT sightings (to be compared with the SBT sightings made by the survey plane). Given that preliminary analysis suggested that at least 50-70 sightings of SBT (not individual schools) were required for meaningful comparison, the calibration experiment was conducted parallel to the dedicated line-transect aerial survey during most of survey. The only times the planes did not fly in parallel were when not enough observers were available, and also one day in March when a large number of SBT sightings were made and the two planes could not stay close enough together for valid comparisons (see Issues for Analysis below).

A second plane (Rockwell Aero Commander) was chartered for the experiment, and a non-spotting pilot was employed. As mentioned in section 3, a spotter-pilot was employed to fly in the survey plane for the entire survey, and two dedicated spotters were swapped between the calibration and survey planes throughout the season.

The calibration experiment followed the protocols used in the line-transect survey regarding the area searched, plane height and speed, environmental conditions, and time of day the survey was conducted. The calibration plane (with only one observer) recorded sightings as it flew along the transect lines approximately 5-7 nautical miles ahead of the dedicated survey plane. However, since the aim of the experiment was to compare sighting rates (i.e. the ability of one observer to detect schools of SBT compared to two observers), the calibration plane did not leave the line to investigate the sightings (as this would alert the trailing survey plane to the sighting). As such, the calibration plane could not verify SBT sightings, but instead recorded the observer's best guess as to what the sighting comprised (e.g., SBT, skipjack, bait fish, wind, et cetera) and whether or not he would have left the line to check the sighting had he been flying in the survey plane. For those sightings that would have been checked, the side of the plane that the sighting was made on, an estimate of the perpendicular distance of the sighting from the plane (i.e., the transect line), and an estimate of the number of schools in the sighting were all recorded.

The survey plane conducted the line-transect survey as normal behind the calibration plane – leaving the transect line to investigate sightings as required. When the survey plane left the transect line to investigate a sighting, the spotter-pilot radioed to the calibration plane to stop searching and circle on the spot until given the signal to start searching again. Radio communication was maintained throughout the survey to monitor the distance between planes and to communicate start/stop. However, communication about SBT sightings did not occur at any time. As for the line-transect survey, a Garmin 176 GPS was used to log the position of the calibration plane and waypoints during the survey by a data-recorder sitting in the back of the plane.

Issues for analysis

While no formal analysis of the calibration experiment data has yet been undertaken, an initial investigation of the data has brought up a number of potential issues for analysis. The primary aim of analysis will be to determine, for each sighting made by the survey plane that was verified to be SBT, whether there was a "match" in the data from the calibration plane (i.e., whether the calibration plane made the same sighting and deemed that the sighting should be checked). Initially, we had hoped that matches could be determined using an automated procedure based on a set of criteria, but this appears to be unfeasible – the data are too complex to set rules that will work in all situations. Instead, each confirmed SBT sighting will need to be evaluated "by hand" and a judgement made as to whether or not a match exists in the calibration data. This evaluation will be based on comparing the locations, times, estimated numbers of schools, and estimated distances of sightings recorded by the calibration plane with the data recorded by the survey plane for the SBT sighting in question. Unfortunately, even with careful consideration given to each SBT sighting, there are a number of reasons, as outlined below, which will make it challenging in many cases to definitively determine whether a match exists.

The distance of 5-7 nautical miles between planes was maintained for safety reasons, whilst still reducing the likelihood of schools having moved (either horizontally or vertically) during the time interval between the two planes crossing the same point on the transect line. Problems occurred, however, when the planes passed through areas of very high school density. Although the calibration plane was able to move quickly through these areas, the survey plane was required to stop and examine each school. In some cases, this took a significant amount of time and resulted in periods of >30 minutes between the two planes crossing the same point. In these situations it will be difficult to judge whether the sightings recorded by the two planes were in fact the same sightings, because schools could have moved, surfaced or gone under during this time.

The calibration plane did not always record an estimate of the distance and number of schools for sightings deemed to potentially be SBT (i.e., for sightings recorded as "would be checked"). This was especially true near the start of the experiment when logistics were still being worked out. In such cases, the amount of data available for determining matches is reduced.

In a couple of situations the survey plane recorded two SBT sightings very close together, whereas the calibration plane recorded only one sighting that could be considered a match. It is quite possible that the observer in the calibration plane judged from the transect line (since this plane cannot leave the line) that all schools belonged to the same sighting; thus, it is not clear whether to consider both SBT sightings to have matches or only one of them. Along the same lines, in situations where the calibration plane missed seeing two SBT sightings that were close together, it is unclear whether to consider this one miss or two.

Initial investigation of the data suggests that the calibration plane missed a number of SBT sightings made by the survey plane; however, what also needs to be considered is that the survey plane missed sightings recorded as most likely being SBT by the calibration plane. Such reverse comparisons are difficult because sightings made by the calibration plane cannot be confirmed as SBT; nevertheless, investigation of this issue is highly important and necessary before drawing any conclusions regarding sighting rates of one observer versus two.

Another matter for consideration is whether there is any potential to include sightings of other species (skipjack, in particular) in the comparisons. Ultimately, we are not interested in whether the sightings rate for skipjack differs between the two planes, but these additional data could increase our sample sizes to more adequate numbers. The difficulty with using skipjack sightings is that the observer in the calibration plane may not have treated these sightings in the same way as SBT (i.e., may not have been as diligent about recording every skipjack sighting) and, also, he was not as faithful about recording additional information for these sightings (i.e., estimated distance from the line and estimated number of schools).

We are also hopeful that the data from the calibration experiment can provide some insight into observer variability, which could help to reduce this source of uncertainty in the aerial survey indices. We have not yet had time to look into this issue.

Analysis of the calibration data will be undertaken during the spring of 2007 and should be completed prior to the next survey.

Summary

The preliminary point estimate from the 2007 scientific aerial survey is lower than the 2006 estimate, and well below the average level in the mid-1990s. The trend appears to have continued downwards, however the high uncertainty in recent estimates (due primarily to low search effort and numbers of sightings) means that we cannot yet be sure that this trend is not just due to chance fluctuation.

An initial investigation of the calibration experiment data has brought up a number of potential issues for analysis. Analysis of the data will be completed prior to the 2008 survey, as it is not required for analysis of the 2007 data. The primary purpose of the calibration data was to compare SBT sighting rates by one observer versus two observers in a plane; however, it is anticipated that the data can also provide some insight into observer variability, which could help to reduce this source of uncertainty in the aerial survey indices.

Acknowledgements

There are many people we would like to recognise for their help and support during this project. We would especially like to thank this years spotters, pilots and data recorders; John-Jay Dent, Chris Hickman, Andrew Jacob, Lyell Jaensch, Darren Tressider and Brett Warren. We also thank the tuna fishing companies in Port Lincoln for their support of the project. This study was funded by AFMA, DAFF, Australian Industry, and CSIRO Marine and Atmospheric Research as part of the Japanese/Australia SBT Recruitment monitoring Program.

References

Basson, M., Bravington, M., Eveson, P. and Farley, J. 2005. Southern bluefin tuna recruitment monitoring program 2004-05: Preliminary results of Aerial Survey and Commercial Spotting data. Final Report to DAFF, June 2005.

Bravington, M. 2003. Further considerations on the analysis and design of aerial surveys for juvenile SBT in the Great Australian Bight. RMWS/03/03.

Cowling, A. 2000. Data analysis of the aerial surveys (1993-2000) for juvenile southern bluefin tuna in the Great Australian Bight. RMWS/00/03.

Cowling A., Hobday, A. and Gunn, J. 2002. Development of a fishery independent index of abundance for juvenile southern bluefin tuna and improvement of the index through integration of environmental, archival tag and aerial survey data. FRDC Final Report 96/111 and 99/105.

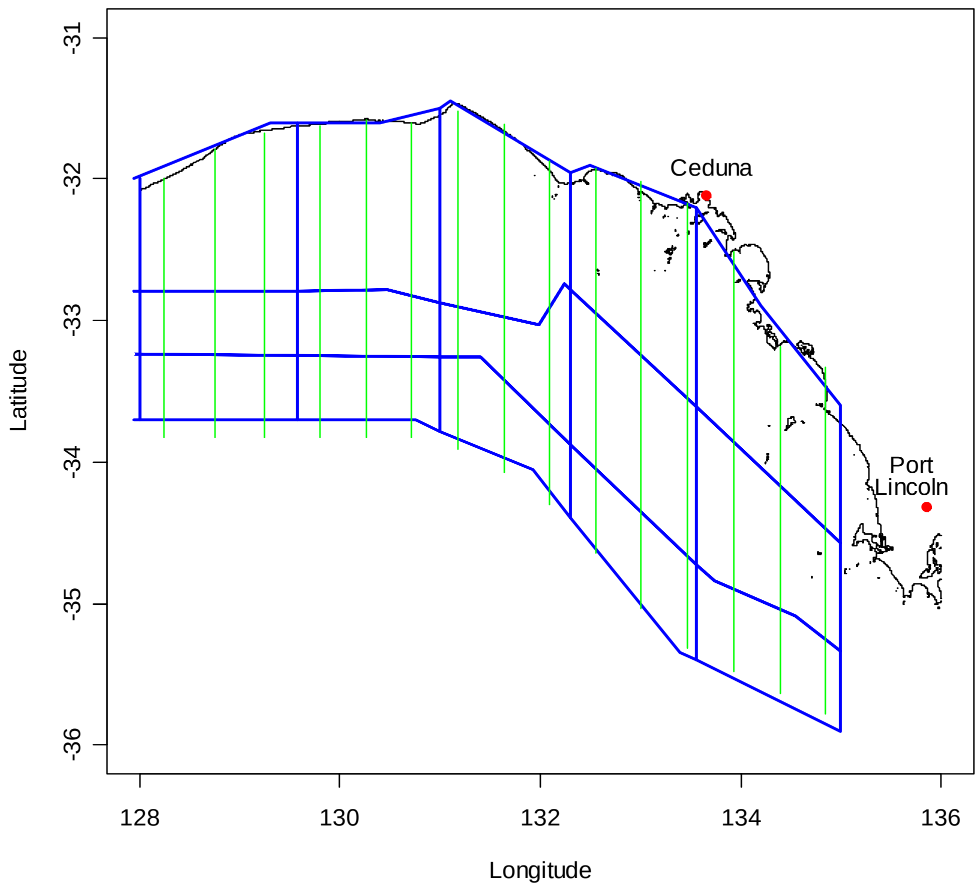
Eveson, P., Bravington, M. and Farley, J. 2006. The aerial survey index of abundance: updated analysis methods and results. CCSBT-ESC/0609/16.

Appendix A – Methods of analysis

Separate models were constructed to describe two different components of observed biomass: i) biomass per patch sighting (BpS), and ii) sightings per nautical mile of transect line (SpM). Each component was fitted using a generalized linear model (GLM), as described below. Since environmental conditions affect what proportion of tuna are available at the surface to be seen, as well as how visible those tuna are, and since different observers can vary both in their estimation of school size and in their ability to see tuna patches, the models include 'corrections' for environmental and observer effects in order to produce standardized indices that can be meaningfully compared across years.

For the purposes of analysis, we defined 45 area/month strata: 15 areas (5 longitude blocks and 3 latitude blocks, as shown in Figure A1) and 3 months (Jan, Feb, Mar). The latitudinal divisions were chosen to correspond roughly to depth strata (inshore, mid-shore and shelf-break).

Figure A1. Plot showing the 15 areas (5 longitudinal bands and 3 latitudinal bands) into which the aerial survey is divided for analysis purposes. The green vertical lines show the official transect lines for the surveys conducted in 1999 and onwards; the lines for previous survey years are similar but are slightly more variable in their longitudinal positions and also do not extend quite as far south (which is why the areas defined for analysis, which are common to all survey years, do not extend further south).



Biomass per sighting (BpS) model

For the BpS model, we first estimated relative differences between observers in their estimates of patch size (using the same methods as described in Bravington 2003). As in Bravington (2003), we found good consistency between observers. In particular, patch size estimates made by different observers tended to be within about 5% of each other, except for one observer, say X, who tended to underestimate patch sizes relative to other observers by about 20%. The patch size estimates were corrected using the estimated observer effects (e.g. patch size estimates made by observer X were scaled up by 20%). Because the observer effects were estimated with high precision, we treated the corrected patch size estimates as exact in our subsequent analyses. The final biomass estimate for each patch was calculated as the average of the two corrected estimates (recall that the size of a patch is estimated by both observers in the plane). The final patch size estimates were then aggregated within sightings to give an estimate of the total biomass of each sighting. It is the total biomass per sighting data that are used in the BpS model.

The BpS model was fitted using a GLM with a log link and a Gamma error structure. We chose to fit a rather rich model with 3-way interaction terms between year, month and area. This is true not only for the BpS model but also for the SpM model described below. In essence, the 3-way interaction model simply corrects the observation (the total biomass of a sighting in the case of the BpS model; the number of sightings in the case of the SpM model) for environmental effects, which are estimated from within-stratum comparisons (i.e. within each combination of year/month/area). Using this model has the advantage of minimizing the risk of bias in the estimates. A more parsimonious model, such as one based on 2-way interactions only, would give predictions with lower variance, but the predictions would be contingent on the assumed model structure being correct, which might not be the case. For example, a model with year*month, year*area and area*month interactions would not be able to reflect different within-season patterns of spatial distribution in different years. Low bias seems more important than low variance here.

Apart from variance, the other main reason for using a less-rich model would be to allow extrapolation to unsurveyed strata. Since the aerial survey has a fairly systematic design, with some coverage in almost all year/month/area combinations, it is possible to get away with a rich model in the interests of minimizing bias. Out of the 495 year/month/area combinations (11 years and 45 area/month strata), there are in fact 19 with no search effort, and for these it is necessary to make predictions using a less-rich model. Since the proportion of unsurveyed strata is so low, the overall abundance index is not much affected.

Having decided on the overall structure, we then decided what environmental variables to include in the model. Based on exploratory plots and model fits, we determined the two environmental covariates that had a significant effect on the biomass per sighting were wind speed and, especially, SST (note that these environmental covariates were determined as part of the 2006 analysis and were not re-investigated in 2007). Thus, the final model fitted was

$$\log\mathbf{E}(\text{Biomass}) \sim \text{Year*Month*Area} + \text{SST} + \text{WindSpeed}$$

where Year, Month and Area are factors, and SST and WindSpeed are linear covariate (note that **E** is standard statistical notation for expected value).

Sightings per mile (SpM) model

For the SpM model, we first updated the pairwise observer analysis described in Bravington (2003), based on within-flight comparisons of sighting rates between the various observers. This analysis gives estimates of the relative sighting abilities for the 18 different observer pairs that have flown at some point in the surveys. The observer pairs ranged in their estimated sighting rates from 54% to 98% compared to the pair with the best rate.

Although this analysis gives reasonable certainty about the relative ranking of different observer pairs, the data provide much less information about the relative efficiency; for example, even if it is clear from the data that A & B together would see more schools than C & D together under the same conditions, it is less clear whether A & B would see 100% more or only 10% more. If there was good certainty about the relative efficiencies, we could just include the estimates from the pairwise model as a known offset (i.e., as a predictor variable with known, rather than estimated, coefficients) when fitting the SpM model. However, because of the uncertainty in the relative efficiencies, we chose instead to include log-relative-efficiency as a covariate in the SpM model rather than as an offset, with the effect size (i.e., “slope”) to be estimated. If the relative efficiencies from the pairwise analysis are correct, the slope estimate should be close to one. This approximation is not perfect, because there is still uncertainty about the relative rankings which we have ignored; in future, we hope to formally merge the pairwise model with the SpM model to correctly propagate all the uncertainty into the final CVs.

The data used for the SpM model were accumulated by flight and area, so that the data set used in the analysis contains a row for every flight/area combination in which search effort was made (even if no sightings were made). Within each flight/area combination, the number of sightings and the distance flown were summed, whereas the environmental conditions were averaged. The SpM model was fitted using a GLM with the number of sightings as the response variable, as opposed to the sightings rate. The model could then be fitted assuming an overdispersed Poisson error structure[3] with a log link and including the distance flown as an offset term to the model (i.e. as a linear predictor with a known coefficient of one). As we did for the BpS model, we used a 3-way interaction to describe year/month/area effects, and we determined what environmental variables to include in the model based on exploratory plots and model fits. A number of environmental covariates correlate highly with the number of sightings made (but not with each other) and came up as significant in the model fits. Again, SST was one of the most influential variables. The final model fitted was:

$$\log \mathbf{E}(\text{N\_sightings}) \sim \text{offset}(\log(\text{Distance})) + \text{Year*Month*Area} + \log(\text{ObsEffect}) \\ + \text{SST} + \text{WindSpeed} + \text{Swell} + \text{Haze} + \text{MoonPhase}$$

where Year, Month and Area are factors, MoonPhase is a factor (taking on one of four levels from new moon to full moon), and all other terms are linear covariates.

Combined analysis

The BpS and SpM model results were used to predict what the number of sightings per mile and the average biomass per sighting in each of the 45 area/month strata in each survey year

[3] Note that the standard Poisson distribution has a very strict variance structure in which the variance is equal to the mean, and it would almost certainly underestimate the amount of variance in the sightings data, hence the use of an overdispersed Poisson distribution to describe the error structure.

would have been under standardized environmental/observer conditions[4]. Using these predicted values, we calculated an abundance estimate for each stratum as 'standardized SpM' multiplied by 'standardized average BpS'. We then took the weighted sum of the stratum-specific abundance estimates over all strata within a year, where each estimate was weighted by the geographical size of the stratum in nm^2, to get an overall abundance estimate for that year. Lastly, the annual estimates were divided by their mean to get a time series of relative abundance indices.

[4] In our predictions, we used above average conditions, namely SST=21, wind speed =3, swell=1, haze=0, low cloud=0, moon phase=4 (full moon), and observer effect=1 (i.e. the 'best' observer pair).

Appendix B – CV calculations

This appendix provides details of how CVs for the aerial survey abundance indices were calculated.

Let $\widetilde{BpS}_{ijk}$ be the predicted value of BpS in year *i*, month *j* and area *k* under standardized environmental/observer conditions (see footnote 4 of main body), and $\hat{\sigma}\left(\widetilde{BpS}_{ijk}\right)$ be its estimated standard error. Similarly, let $\widetilde{SpM}_{ijk}$ be the predicted value of SpM in year *i*, month *j* and area *k* under the same environmental/observer conditions, and $\hat{\sigma}\left(\widetilde{SpM}_{ijk}\right)$ be its estimated standard error. Then,

$$\hat{A}_{ijk} = \widetilde{SpM}_{ijk}\widetilde{BpS}_{ijk}$$

is the stratum-specific abundance estimate for year *i*, month *j* and area *k*.

Since $\widetilde{BpS}_{ijk}$ and $\widetilde{SpM}_{ijk}$ are independent, the variance of $\hat{A}_{ijk}$ is given by

$$\begin{aligned} V\left(\hat{A}_{ijk}\right) &= V\left(\widetilde{SpM}_{ijk}\widetilde{BpS}_{ijk}\right) \\ &= V\left(\widetilde{SpM}_{ijk}\right)E\left(\widetilde{BpS}_{ijk}\right)^2 + V\left(\widetilde{BpS}_{ijk}\right)E\left(\widetilde{SpM}_{ijk}\right)^2 + V\left(\widetilde{SpM}_{ijk}\right)V\left(\widetilde{BpS}_{ijk}\right) \\ &\approx \hat{\sigma}^2\left(\widetilde{SpM}_{ijk}\right)\widetilde{BpS}_{ijk}^2 + \hat{\sigma}^2\left(\widetilde{BpS}_{ijk}\right)\widetilde{SpM}_{ijk}^2 + \hat{\sigma}^2\left(\widetilde{SpM}_{ijk}\right)\hat{\sigma}^2\left(\widetilde{BpS}_{ijk}\right) \end{aligned}$$

The annual abundance estimate for year *i* is given by the weighted sum of all stratum-specific abundance estimates within the year, namely

$$\hat{A}_i = \sum_j \sum_k w_k \hat{A}_{ijk}$$

where w_k is the proportional size of area *k* relative to the entire survey area ($\sum_k w_k = 1$).

If the $\hat{A}_{ijk}$'s are independent, then the variance of $\hat{A}_i$ is given by

$$V\left(\hat{A}_i\right) = \sum_j \sum_k {w_k}^2 V\left(\hat{A}_{ijk}\right)$$

Unfortunately, the $\hat{A}_{ijk}$'s are NOT independent because the estimates of BpS (and likewise, the estimates of SpM) are not independent between different strata. This is because all strata estimates depend on the estimated coefficients of the environmental/observer conditions, so any error in these estimated coefficients will affect all strata. Thus, we refit the BpS and SpM models with the coefficients of the environmental/observer covariates (denote the vector of coefficients by θ[5]) fixed at their estimated values ($\hat{\theta}$). The predictions of BpS and SpM made using the 'fixed environment' models should now be independent between strata, so the stratum-specific abundance estimates calculated using these predictions – which we will denote by $\hat{A}_{ijk}\left(\hat{\theta}\right)$ – should also be independent between strata. Thus, we can calculate the variance of $\hat{A}_i$ conditional on the estimated values of the environmental/observer coefficients as

$$V\left(\hat{A}_i \mid \hat{\theta}\right) = \sum_j \sum_k w_k{}^2\, V\left(\hat{A}_{ijk}\left(\hat{\theta}\right)\right)$$

where $V\left(\hat{A}_{ijk}\left(\hat{\theta}\right)\right)$ is calculated using the formula given above for $V\left(\hat{A}_{ijk}\right)$ but using the BpS and SpM predictions and standard errors obtained from the 'fixed environment' models.

To calculate the unconditional variance of $\hat{A}_i$, we make use of the following equation:

$$\begin{aligned} V\left(\hat{A}_i\right) &= E_\theta\left(V\left(\hat{A}_i \mid \theta\right)\right) + V_\theta\left(E\left(\hat{A}_i \mid \theta\right)\right) \\ &\approx V\left(\hat{A}_i \mid \hat{\theta}\right) + V_\theta\left(\hat{A}_i\right) \end{aligned}$$

where the first term is the conditional variance just discussed and the second term is the additional variance due to uncertainty in the environmental coefficients. The second term can be estimated as follows

$$V_\theta\left(\hat{A}_i\right) \approx \left(\frac{\partial \hat{A}_i}{\partial \theta}\right)' \mathbf{V}_\theta \left(\frac{\partial \hat{A}_i}{\partial \theta}\right)$$

where $\left(\frac{\partial \hat{A}_i}{\partial \theta}\right)$ is the vector of partial derivatives of $\hat{A}_i$ with respect to θ (which we calculated using numerical differentiation), and $\mathbf{V}_\theta$ is the variance-covariance matrix of the environmental coefficients[6].

Finally, the relative abundance index for year *i* is calculated as

$$\hat{I}_i = \frac{\hat{A}_i}{\sum_i \hat{A}_i}$$

[5] θ contains the environmental/observer coefficients from both the BpS and SpM models; i.e. $\theta = (\theta_{\text{BpS}}, \theta_{\text{SpM}})$

[6] Recall that θ contains the environmental/observer coefficients from both the BpS and SpM models, so $\mathbf{V}_\theta = \begin{bmatrix} \mathbf{V}_{\theta_{\text{BpS}}} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{\theta_{\text{SpM}}} \end{bmatrix}$. The variance-covariance matrices for the individual models are returned from the model-fitting software.

Using the delta method, we can approximate the variance of $\hat{I}_i$ by

$$V\left(\hat{I}_i\right) \approx \left(\frac{\partial \hat{I}_i}{\partial \hat{A}_i}\right)^2 V\left(\hat{A}_i\right)$$

Then, the standard error of $\hat{I}_i$ is given by

$$\sigma\left(\hat{I}_i\right) = \sqrt{V\left(\hat{I}_i\right)}$$

and the coefficient of variation (CV) of $\hat{I}_i$ is given by

$$\mathrm{CV}\left(\hat{I}_i\right) = \frac{\sigma\left(\hat{I}_i\right)}{\hat{I}_i}$$