CCSBT-ESC/0809/24

CSIRO

The aerial survey index of abundance: updated analysis methods and results

Paige Eveson
Mark Bravington
Jessica Farley

Prepared for the CCSBT 5th Management Procedure Workshop 2-7 September and the 13th Meeting of the Extended Scientific Committee 8-12 September 2008
Rotorua, New Zealand

TABLE OF CONTENTS

Abstract

The estimate of relative abundance of juvenile southern bluefin tuna (SBT) in the Great Australian Bight from the 2008 scientific aerial survey is higher than the 2005-2007 estimates. However, the 90% confidence interval on the 2008 estimate overlaps with the confidence intervals for the previous 3 years. Therefore, we cannot conclude that the increase is statistically significant. The 2008 estimate remains significantly below the average level in the mid-1990s.

The models used to analyse the data were modified from the previous year to include random strata effects, which can better handle strata (e.g., year/month/area combinations) where there is low effective sampling effort. The relative abundance indices obtained using these models are very similar to those obtained using the previous fixed strata effects models, but achieve higher precision (i.e., smaller CVs) due to the reduced number of parameters estimated (i.e. increased degrees of freedom).

A large-scale calibration experiment was undertaken in 2008, with the primary purpose to compare the number of sightings and estimated total biomass of SBT observed by the calibration plane versus the survey plane (in light of the fact that future surveys may have only one observer in a plane). An initial investigation of the data suggests that the calibration plane made significantly fewer sightings than the survey plane; however, this difference still needs to be rigorously quantified. Note that the mean estimates of biomass per sighting and number of patches per sighting were similar between the calibration plane and survey plane. Full analysis of the calibration experiment will be reported to the 2009 SAG and SE meetings.

Introduction

The index of juvenile Southern bluefin tuna (SBT) abundance from the scientific aerial survey in the Great Australian Bight (GAB) is one of the 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 2001 due to logistic problems of finding trained, experienced spotters and spotter-pilots. (Note that the terms ‘spotter’ and ‘observer’ are used interchangeably). The suspension also allowed for further data analysis and an evaluation of the design and effectiveness of the survey. A decision to continue or end the scientific aerial survey could then be made on the merits of the data, in particular the ability to detect changes in abundance.

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 (Anon, 2007). In 2005, the full scientific line-transect aerial survey was re-established in the GAB, and new analysis methods were developed for estimating an index of abundance over all survey years. The aerial survey has been conducted each year since, and the analysis methods have also been refined to provide a robust estimator and confidence intervals.

In 2007, a large-scale calibration experiment was conducted alongside the scientific aerial survey, 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). While the study provided useful information about differences in sightings between observers (e.g., sightings made by one observer are often missed by another observer), it proved difficult to answer the main question of how much better two observers are than one.

Thus, in 2008, another calibration experiment was designed and run in parallel with the full scientific line-transect aerial survey. This time, the calibration experiment was designed to compare the number of SBT sightings and total estimated biomass of SBT seen by the calibration plane versus the survey plane over similar areas and time periods, as opposed to trying to “match up” sightings as in the 2007 calibration experiment.

This report summarises the field procedures and data collected during the 2008 season for the aerial survey and the calibration experiment. The most recent methods for analysing the aerial survey data are described, and results are presented. Analysis of the calibration data is still underway, but preliminary results are presented.

Line-transect aerial survey

Field procedures

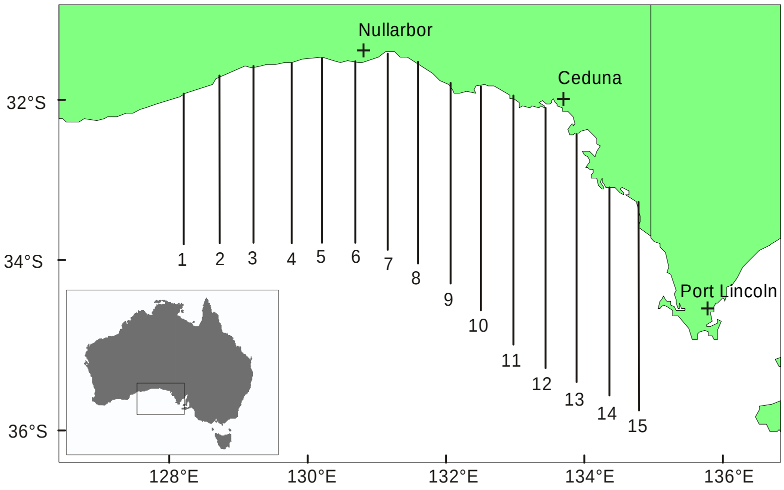
The line-transect aerial survey was conducted in the GAB between January and March 2008. As in previous years, the plane used was a Rockwell Aero Commander 500S. Two observers were employed for each survey flight, one spotter-pilot and one spotter. Unfortunately, the spotter-pilot employed for the 2005-2007 surveys was not available in 2008. However, a spotter-pilot that participated in the 1993-2000 surveys was available, and as he was already calibrated into the survey, he was suitable for the 2008 survey. This year, two spotters were employed as the second observer in the survey plane, one of whom was used in the 2005-2007 surveys and the other in the 2007 survey. Each spotter participated in approximately half of the line transect survey flights during January and February (i.e. during the time that the calibration experiment was conducted – see Calibration experiment section below) and only one spotter was participated in the survey plane in March.

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 (Fig. 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).

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.

Figure 1. Location of the 15 north-south transect lines for the scientific aerial survey in the GAB.



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.

2008 survey year

The line-transect survey was successfully completed in 2008. Approximately 53 lines were completed, compared to 35 in 2006, 44 in 2005 and 38 in 2006. The increased number of lines surveyed this year was predominantly due to good weather conditions experienced in March. The total flying time (transit and transect time) for the 2008 survey was 146.2 hours.

Data preparation

The data collected from the 2008 survey were loaded into the aerial survey database, which contains the data collected from all surveys. The 2008 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 Bravington 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. 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.

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
2008	7465	122	7929	3.5	24	18.7	12.5	312

---

[1] The biomass statistics differ slightly from those reported in Table 2 of the 2007 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.

The total distance searched increased substantially in 2008 compared to the previous three survey years. As noted above, this was primarily due to the unusually good flying conditions in March. The sightings rate (number of sightings per nm) also increased significantly to almost the highest of all survey years (Fig. 2). However, this did not translate to such a large increase in the amount of biomass per nm, because the average patch size was the smallest of any survey year (Fig. 2; Table 1). Almost all SBT sightings in 2008 were made in the eastern half of the survey area, with the distribution of sightings clustered particularly along transect lines 9 and 10 (Fig. 3).

The average and maximum number of patches per sighting were much higher than they have been in recent survey years, and similar to those seen in the early survey years (Table 1). On the contrary, the average and median biomass per patch were much smaller than in any of the previous years (Table 1; Fig. 4).

Estimates of the average size of SBT within a patch have been found to be inconsistent between different observers (Cowling et al. 2002). In past reports, we were able to make valid comparisons between years by only considering fish size estimates made by a single observer who had operated in all survey years. Unfortunately, this observer did not operate in 2008. Furthermore, the fish size data are not used in the analysis to estimate the abundance index. Thus, we have not updated the figure showing fish size data in this year's report.

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.***

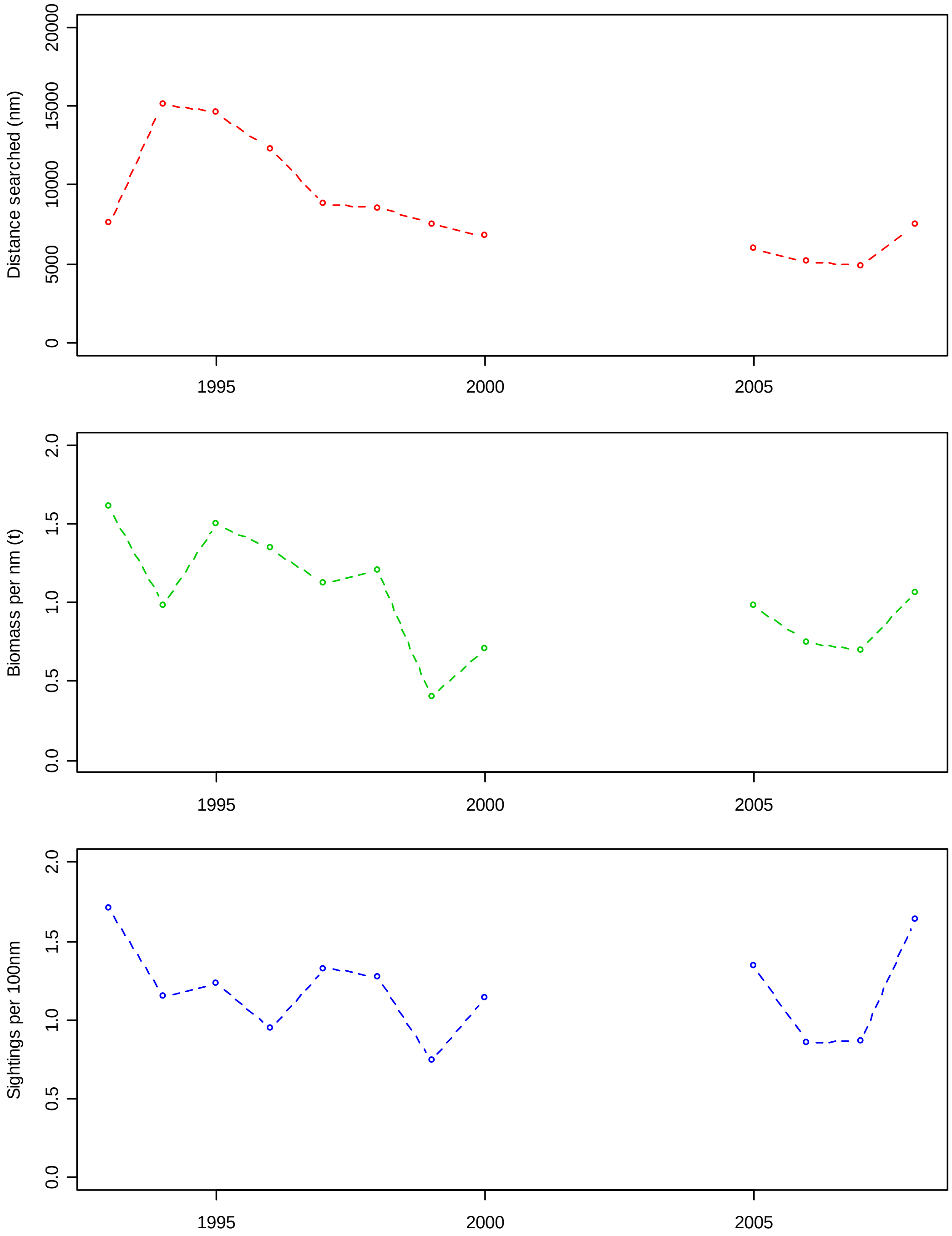


Figure 3. Distribution of SBT sightings made during each aerial survey year. Red circles show the locations of SBT sightings, where the size of the circle is proportional to the size of the sighting, and grey lines show the north/south transect lines that were searched.

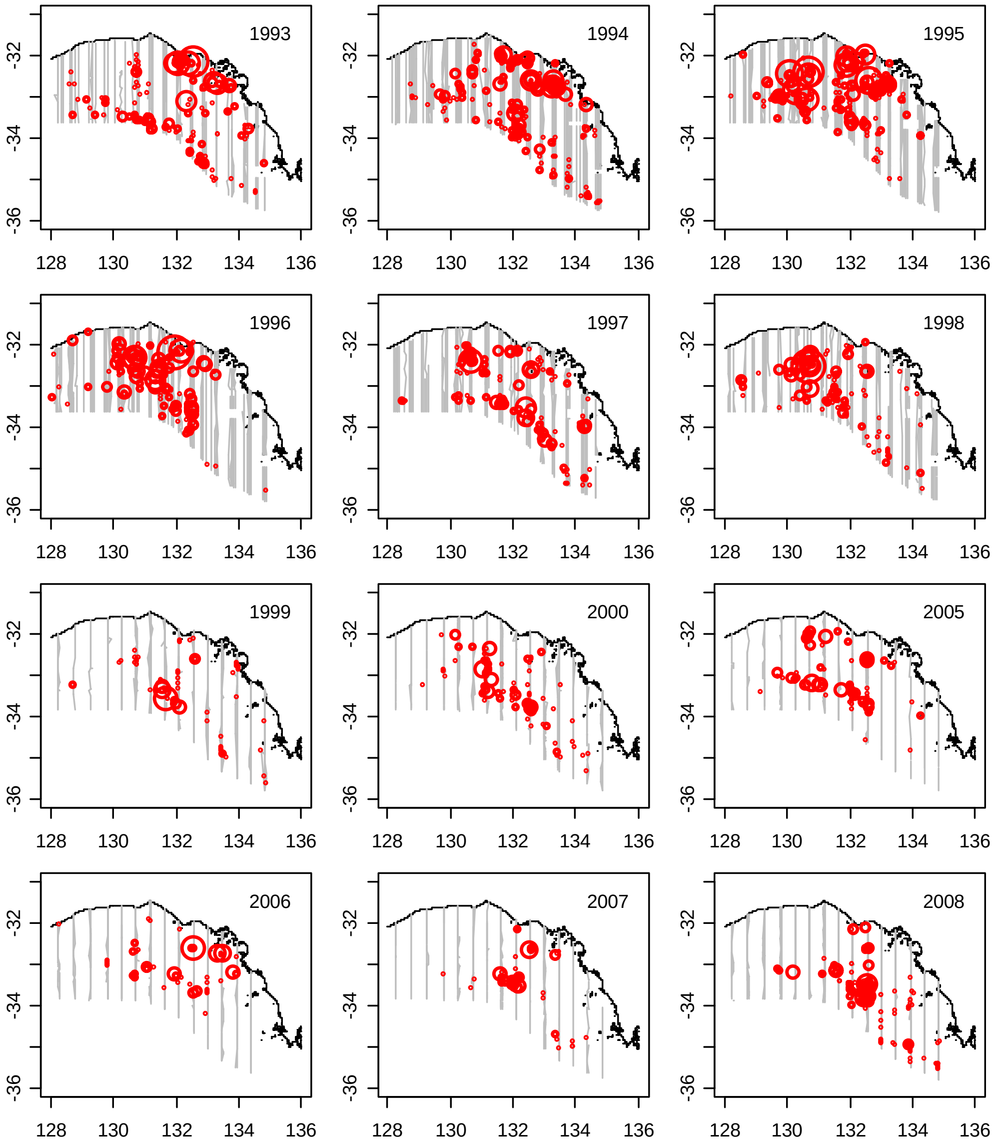
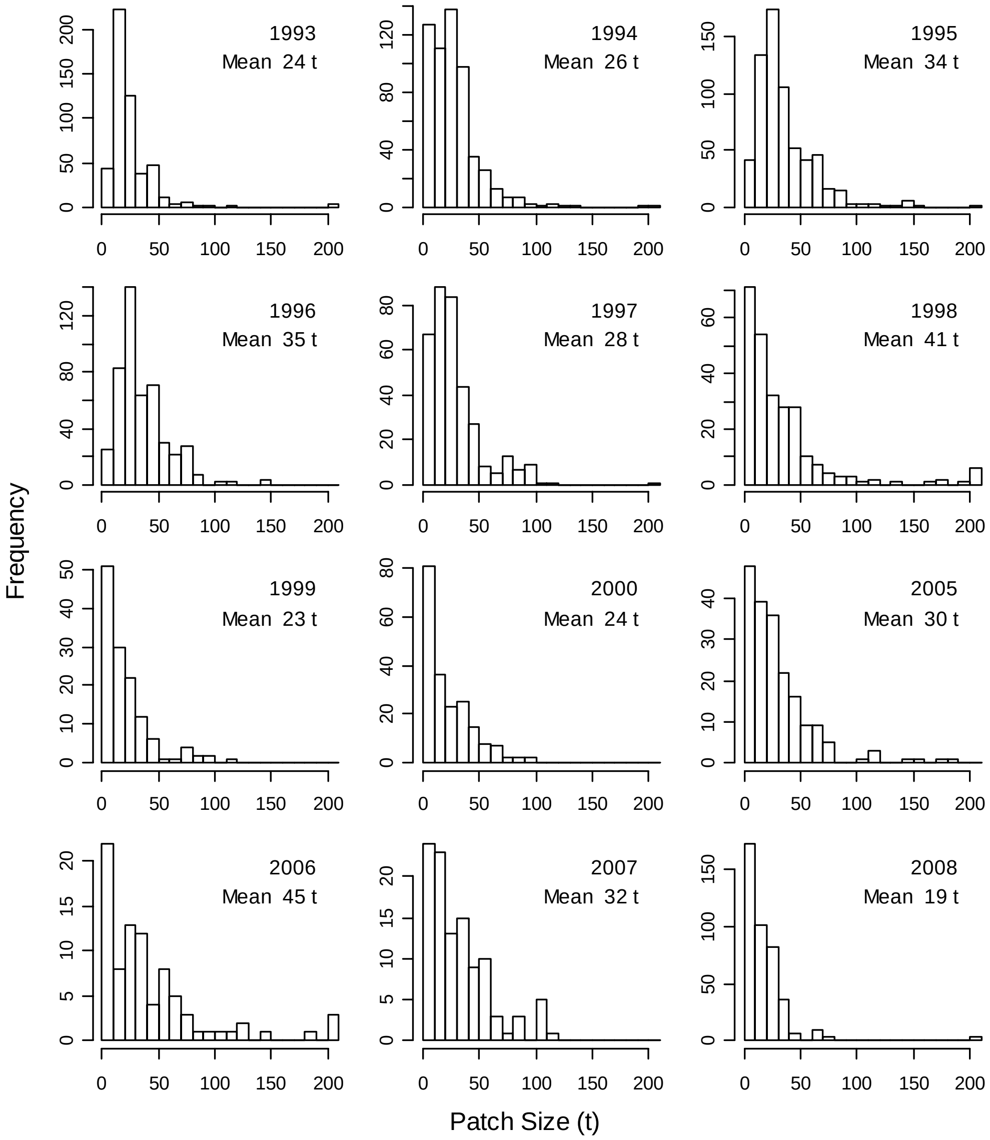


Figure 4. Frequency of SBT patch sizes (in tonnes) by survey year.



Environmental variables

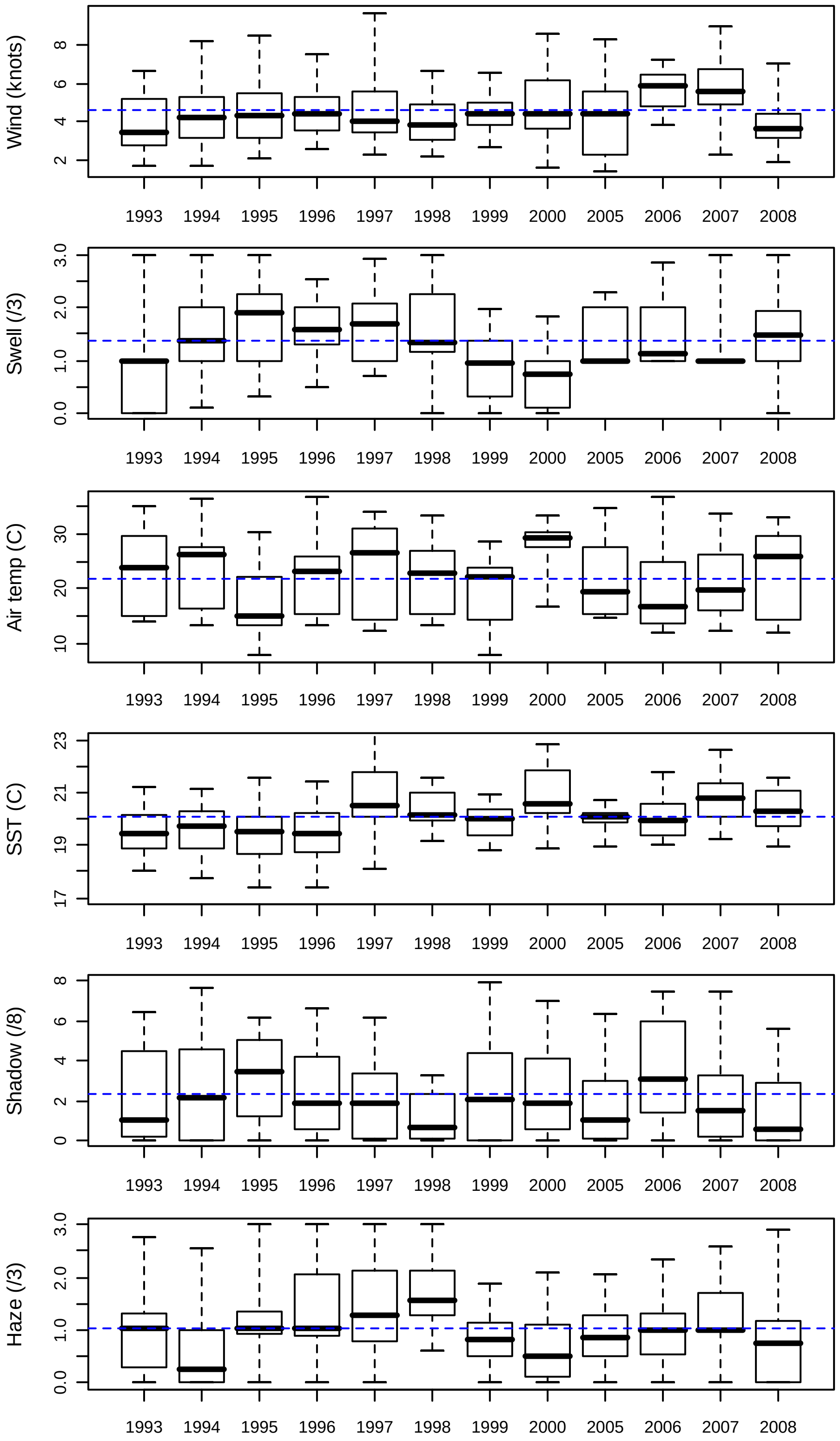
Table 2 and Figure 5 summarize the environmental conditions that were present during valid search effort in each survey year. All the environmental variables presented were recorded by the survey plane, with the exception of sea surface temperature (SST), which was extracted from the 3-day composite 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 2008 survey was much lower on average compared to other survey years (Table 2; Fig. 5), and was the main reason for the higher amount of search effort that could be completed. SST was fairly average, whereas the amount of sea shadow and haze tended to be lower than in other years (Table 2; Fig. 5). Overall, the conditions were above average for making sightings.

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
2008	3.8	1.4	24.2	20.4	1.4	0.9

Figure 5. 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

This year we were successful in implementing a random-effects extension of the models, as recommended in last year's report (Eveson et al. 2007). Details of the revised methods (which, for the most part, are the same as last year) can be found in Appendix A. Here, we give a brief description, highlighting where the analysis has been changed.

As before, generalized linear models were fit to two different components of observed biomass—biomass per patch sighting (BpS) and sightings per nautical mile of transect line (SpM). Terms were included in both models for year, month and area, including all 2-way and 3-way interactions (recall that for analysis purposes, we split the survey region into 15 areas; see Figure A1 in Appendix A). We included the same environmental and observer variables in the models this year as we did last year. Thus, the models can still be expressed as:

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

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

where **E** is standard statistical notation for expected value.

The difference in the models this year is that the 3-way year/month/area interaction term is now fit as a random effect instead of a fixed effect. Essentially, this means that we assume the year/month/area effects belong to a normal distribution with a given mean and variance, which are to be estimated within the model (i.e., only two parameters being estimated). Previously, in the fixed effect version of the models, a separate parameter was estimated for each year/month/area combination. Besides the large reduction in the number of parameters needing estimated, the other main advantage of the random effects model is that when no data exist for a given year/month/area stratum, we still have information about it because we are assuming it comes from a normal distribution for which we have parameter estimates. For example, with no other information, we would estimate the coefficient for a missing year/month/area effect to be the mean of the normal distribution. In the fixed effect version of the model, we had to resort to fitting a less rich model with only 2-way interaction terms in order to make inferences for any year/month/area strata with no data (see Appendix A of last year's report; Eveson et al. 2007).

Once the new models were fitted, the analysis proceeded as before. Specifically, 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 area/month 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.

We stress that 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. We used the same process as last year to calculate CVs for the indices, the details of which are repeated 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 follows a normal distribution, with standard errors approximated by the CVs of the untransformed indices.

Results

Figure 6 shows the estimated time series of relative abundance indices with 90% confidence intervals. The point estimates and CVs corresponding to Figure 6 are reproduced in Table 3. The point estimate for 2008 is higher than the 2005-2007 estimates, but because the 90% confidence interval on the 2008 estimate overlaps with those for the previous 3 years, we cannot conclude that the increase is statistically significant. It should also be noted that the 2008 estimate remains significantly below the average level in the mid-1990s.

Note that using the new random strata effects models leads to very similar point estimates as the previous fixed strata effects models, but that the precision on the estimates tends to be higher (i.e., smaller CVs) (Fig. 7).

Figure 6. Time series of relative abundance estimates with 90% confidence intervals.

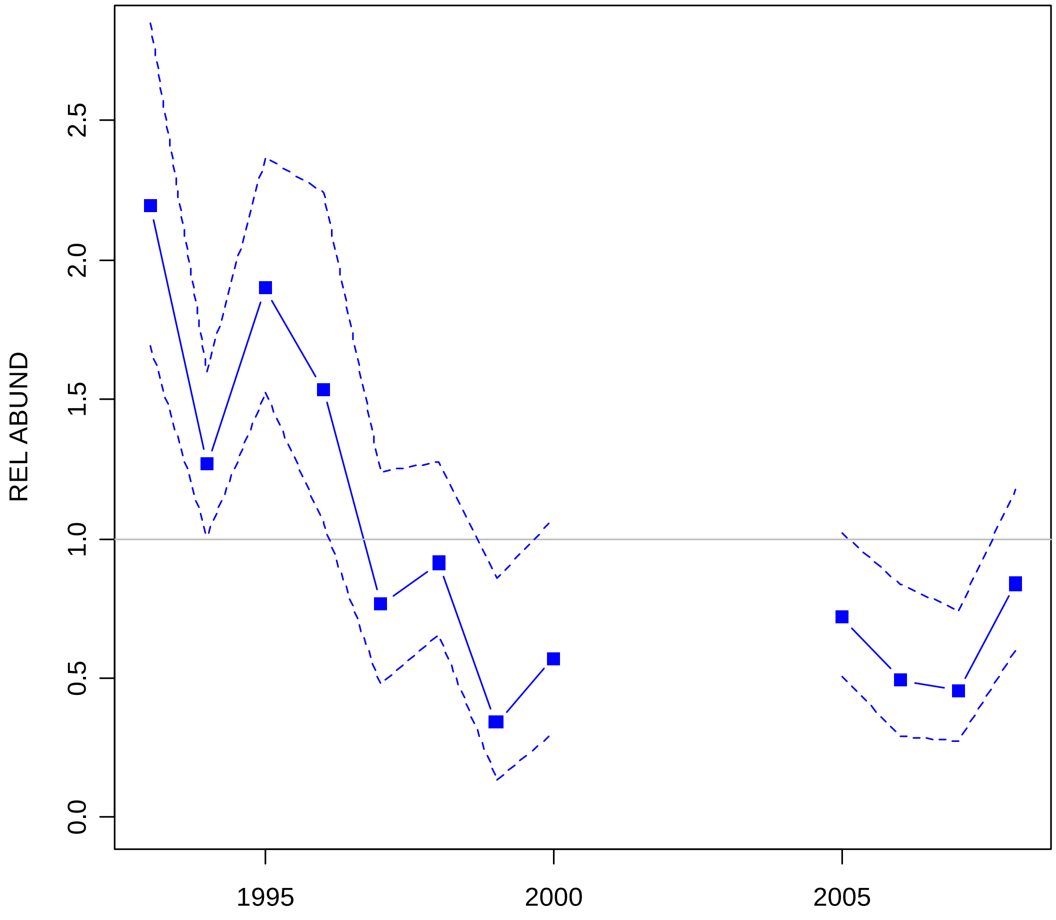
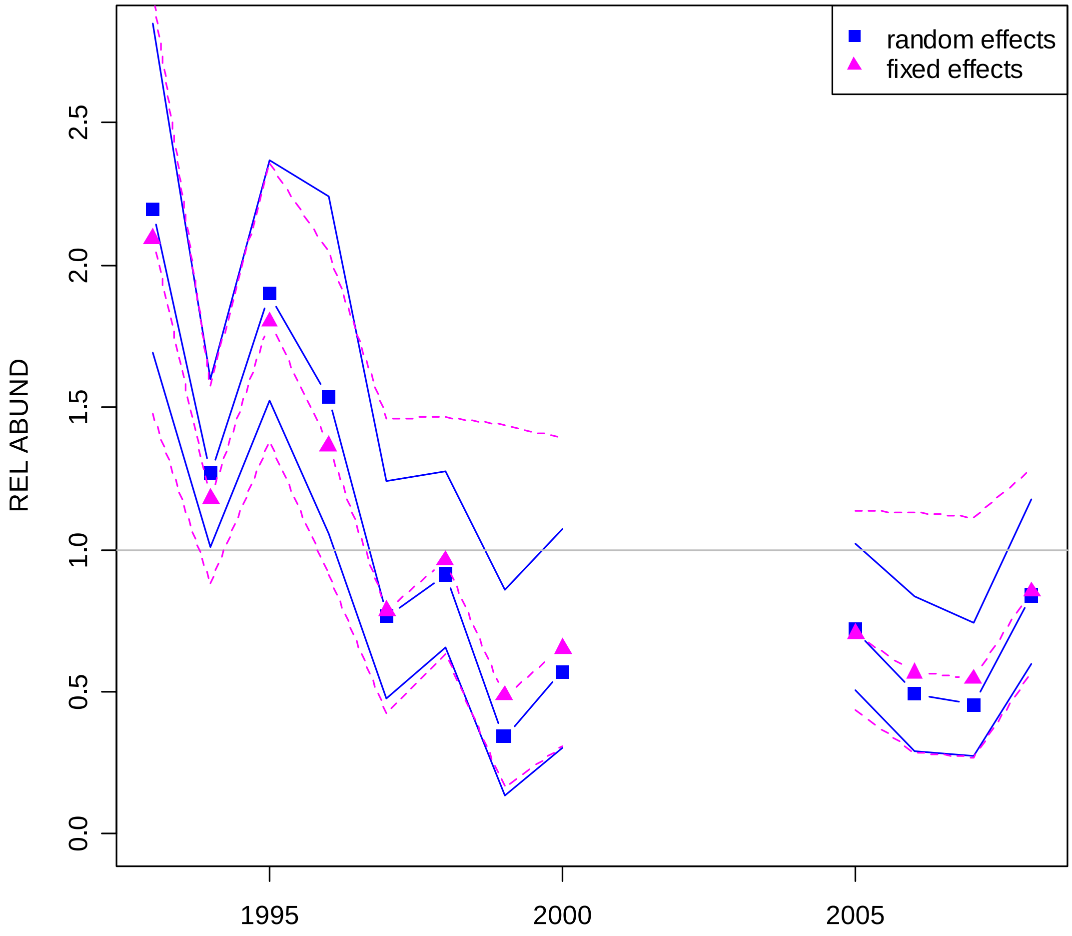


Table 3. Aerial survey relative abundance point estimates, standard errors and CVs (as per Fig. 6).

Year	Estimate	SE	CV
1993	2.20	0.35	15.8%
1994	1.27	0.18	14.1%
1995	1.90	0.25	13.4%
1996	1.54	0.35	22.9%
1997	0.77	0.22	29.2%
1998	0.91	0.19	20.3%
1999	0.34	0.19	55.8%
2000	0.57	0.22	38.3%
2005	0.72	0.15	21.3%
2006	0.49	0.16	31.9%
2007	0.45	0.14	30.3%
2008	0.84	0.17	20.5%

Figure 7. Comparison of relative abundance estimates and 90% confidence intervals obtained using the new random effects models versus the fixed effects models used in last year's analysis.



Calibration experiment

Field procedures

The second component of the 2008 program was to undertake a calibration experiment in parallel with the scientific line-transect aerial survey. The design and primary goal of the 2008 survey differed from the 2007 study. In particular, the calibration study conducted in 2007 was aimed at determining whether each SBT sighting made by the survey plane was also made by the calibration plane. Unfortunately, even with careful consideration given to each sighting, it proved impossible in many cases to definitively determine whether a match existed (for details refer to Eveson et al. 2007). Thus, the 2008 calibration study was designed to compare the total number of SBT sightings and total estimated biomass of SBT seen by the calibration plane versus the survey plane over similar areas and time periods, as opposed to trying to “match up” individual sightings. As in the 2007 calibration study, the experiment also provided additional data on observer variability, which may be able to help reduce this source of uncertainty in the aerial survey indices.

The calibration study was conducted in January and February 2008. A second plane (also a Rockwell Aero Commander) was chartered and a non-spotting pilot was employed. The experiment involved the second plane with one observer (and a non-spotting pilot) flying in tandem with the survey plane. Radio communication was maintained throughout the study to monitor the distance between planes and to communicate start/stop (see below). However, communication about SBT sightings did not occur at any time. As with the survey plane, a Garmin 176 GPS was used by a data-recorder sitting in the back of the plane to log the position of the calibration plane and waypoints.

As mentioned in the Field procedures section for the line-transect aerial survey, two dedicated spotters were swapped between the calibration plane and survey plane in January and February. The calibration plane followed the same protocols used by the survey plane, except the sole observer searched for patches of SBT from his side of the plane through 180° to the other side of the plane. This allowed for a direct comparison of the number of sightings and biomass recorded. The calibration plane recorded sightings as it flew along the transect lines either ahead or behind the dedicated survey plane. Both planes left the line as required to investigate each sighting and to estimate the number of schools and biomass in the sighting. The distance between the two planes varied depending on the situation. For example, if it was expected that the day was going to be ‘long’ (if the far western block was being surveyed, or it was expected that a larger number of SBT sightings could be encountered – given the weather conditions and location), then the two planes would take off at the same time but start on different lines. Once these lines were completed, the planes would swap lines. The lines were surveyed in the same direction approximately 1 hour apart. If a third line could be flown, it was flown in tandem with ~1 hr between planes. On ‘short’ days, the planes would take off approximately 30 minutes apart with the second plane following the first plane all day, remaining approximately the same time/distance apart. If the survey and calibration planes were too close to each other, the leading plane would instruct the other plane to stop searching and circle on the spot until given the signal to start searching again. This occurred very rarely.

The only times the calibration and survey planes did not fly in tandem were when a large number of SBT sightings were made and the calibration plane would have delayed the survey plane from completing the survey that day (this only occurred on one day in January). On another three days, due to time constraints, the calibration plane only completed two lines while the survey plane completed the full three lines.

Preliminary analysis and results

For each day of the calibration study, we compared the number of SBT sightings and total biomass estimates made by the two planes (Table 4). In doing so, we only included data from sections of lines that were flown by both planes. The summary statistics suggest that the calibration plane consistently made fewer SBT sightings (and estimated less total biomass) than the survey plane. This was true regardless of which spotter was in the calibration plane, although the data are quite limited for observer 50 (Table 4). Note that because the planes flew the same lines within a short time of one another, we do not need to account for different environmental conditions when comparing the number of sightings and biomass estimates between the two planes.

The number of patches per sighting and biomass per patch were, on average, estimated to be quite similar by the two planes (Table 5). This fits our expectation because there is no reason for an observer's ability to estimate the number and size of patches to be affected by whether he is flying in the calibration plane or the survey plane.

Table 4. Number of sightings and total biomass estimates (in tonnes) made by the survey plane versus the calibration plane during each day of the calibration experiment. Data are only included for line segments searched by both planes on the same day.

Year	Month	Day	Calib. spotter	Survey # sightings	Calib. # sightings	**Difference in # sightings**	Survey biomass	Calib. biomass	**Difference in biomass**
2008	1	10	50	27	17	**10**	3705	1758	**1947**
2008	1	12	3	0	0	**0**	0	0	**0**
2008	1	26	3	3	2	**1**	406	76	**329**
2008	1	27	3	0	0	**0**	0	0	**0**
2008	1	28	3	0	0	**0**	0	0	**0**
2008	1	4	50	0	0	**0**	0	0	**0**
2008	1	9	3	1	1	**0**	90	63	**27**
2008	2	16	3	1	0	**1**	0	0	**0**
2008	2	17	3	30	8	**22**	1226	857	**369**
2008	2	18	3	5	4	**1**	1	23	**-23**
2008	2	24	50	1	0	**1**	0	0	**0**
2008	2	26	50	0	0	**0**	0	0	**0**
	Total:			68	32	**36**	5428	2777	**2649**

Table 5. Comparison of summary statistics for calibration plane versus survey plane. Data are only included for line segments searched by both planes on the same day.

Plane	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
Survey	3815	68	5428	3.5	24	23.1	21.9	257
Calibration	3813	32	2777	3.3	15	26.2	17.9	169

Although it is clear that the calibration plane made fewer sightings than the survey plane, we need a statistical method for quantifying this difference (e.g., that takes into account all sources of variability in the results). Recall that for the SpM model, we first perform a pair-wise observer analysis to estimate the relative sighting abilities of all observer pairs that have been involved in past and present surveys (see Appendix A). Essentially, we need to revise this analysis to estimate the relative sighting ability not only of all observer pairs, but also of all solo observers. Appropriate statistical models are currently being explored.

Summary

The preliminary point estimate from the 2008 scientific aerial survey is higher than the 2005-2007 estimates. However, as the 90% confidence interval on the 2008 estimate overlaps with those for the previous 3 years, we cannot conclude that the increase is statistically significant. It should also be noted that the 2008 estimate remains significantly below the average level in the mid-1990s.

The models used to analyse the data were modified from last year to include random year/month/area strata effects, which can better handle strata where there is low (or no) sampling effort. The relative abundance indices obtained using the random strata effects models are very similar to those obtained using the previous fixed strata effects models, but the precision achieved for most estimates is higher.

An initial investigation of the calibration data suggests that the calibration plane consistently made fewer SBT sightings (and estimated less total biomass) than the survey plane. This was true regardless of which spotter was in the calibration plane. In order to rigorously quantify this difference, we need to develop a statistical model that takes into account all sources of variability in the results. Appropriate models are currently being explored.

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; Andrew Jacob, Darren Tressider, Derek Hayman, John Veerhius, Thor Carter and Adam Vasey. 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’s Wealth from Oceans Flagship.

References

Anon, 2007.

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

Bravington, M., Eveson, P., and Farley, J. 2005. The aerial survey index of abundance: updated to include the 2005 survey. CCSBT-ESC/0509/22.

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.

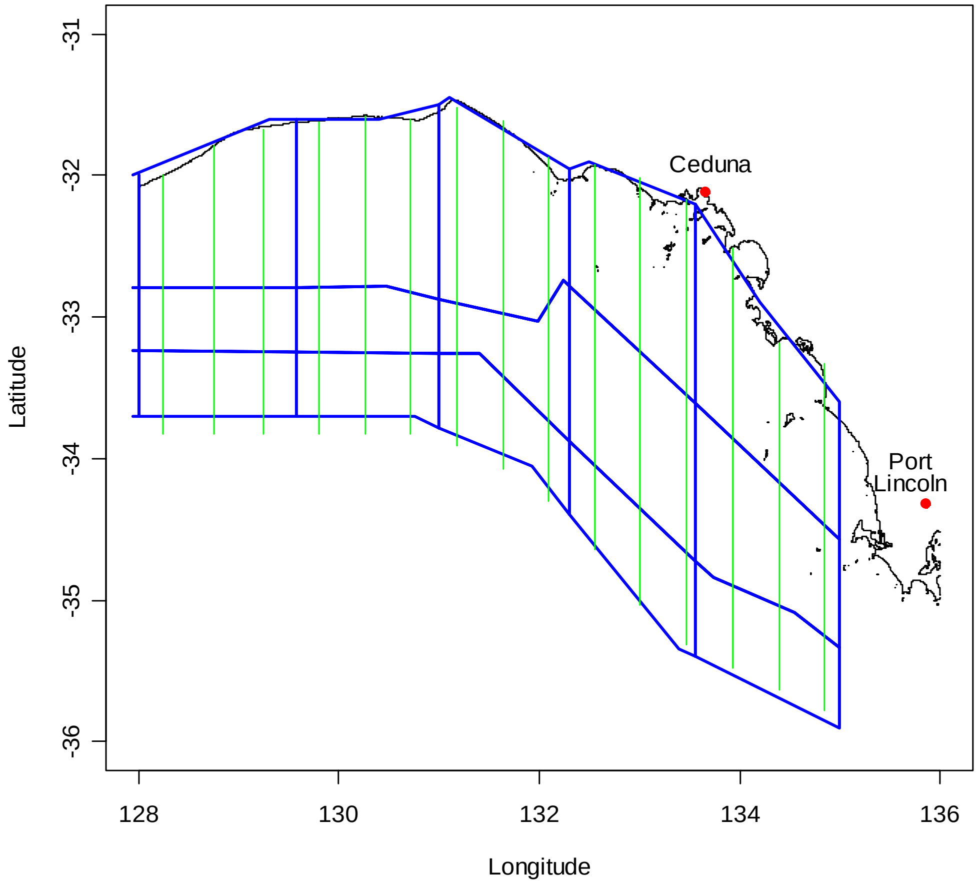
Eveson, P., Bravington, M. and Farley, J. 2007. Aerial survey: updated index of abundance and preliminary results from calibration experiment. CCSBT-ESC/0709/12.

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 observer 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 GLMM (generalized linear mixed model) 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).

As opposed to last year's analysis, where we fitted the 3-way year/month/area interaction terms as fixed effects, we now fit these terms as random effects. This means that we assume the year/month/area interactions belong to a normal distribution with a given mean and variance, which are to be estimated within the model (i.e., only two parameters being estimated). Previously, in the fixed effects version of the model, a separate parameter was estimated for each year/month/area combination. Besides the large reduction in the number of parameters needing estimated, the other main advantage of the random effects model is that when no data exist for a given year/month/area stratum, we still have information about it because we are assuming it comes from a normal distribution for which we have parameter estimates. For example, with no other information, we would estimate the coefficient for a missing year/month/area effect to be the mean of the normal distribution. In the fixed effect version of the model, we had to resort to fitting a less rich model with only 2-way interaction terms in order to make inferences for any year/month/area strata with no data (see Appendix A of Eveson et al. 2007).

We included two environmental covariates, namely wind speed and SST, in the model. These two variables were determined to have a significant effect on the biomass per sighting (note that a rigorous investigation of environmental covariates was conducted in 2006 and has not been repeated since then). 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 62% 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; we still hope in future 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 GLMM 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[2] 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 included 3-way interaction terms to describe year/month/area effects, and we fitted these interactions as random effects (see BpS model section). 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 (note that a rigorous investigation of environmental covariates was conducted in 2006 and has not been repeated since then). 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.

---

[2] 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.

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 would have been under standardized environmental/observer conditions[3]. 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 area/month 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.

---

[3] 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 3 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[4]) 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[5].

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

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

---

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

[5] 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

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