CCSBT-ESC/0609/16

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

Paige Eveson
Mark Bravington
Jessica Farley

Prepared for the CCSBT 7th Meeting of the Stock Assessment Group (SAG7) and the 11th Meeting of the Extended Scientific Committee (ESC11)
4-11 September, and 12-15 September 2006, Tokyo, Japan

TABLE OF CONTENTS

Abstract

The estimate from the 2006 scientific aerial survey shows a small drop from the 2005 estimate, and both are below average levels in the mid-1990s. The mean of the 2005 and 2006 estimates is about 66% of the mean of the 1994-1998 estimates. The coefficient of variation (CV) on the 2006 estimate is higher than in most years, including 2005, due to the low amount of search effort and number of sightings in 2006, as well as to the anomalous environmental conditions.

Since last year's analysis (Bravington et al. 2005), we have made some changes to the analysis methods and to the historical database. The main impact has been the use of SST data to improve the adjustments for environmental variability; this has a fairly strong effect on the index in some years, including 2006 for which the SST conditions were unusual. Other changes to the methods and database have streamlined the analysis considerably, but have had comparatively minor impacts on the abundance index itself.

Variation between observers remains the chief source of uncertainty in the aerial survey indices, and really needs to be resolved if the survey is to continue usefully. If the aerial survey goes ahead in 2007, it should certainly include from the outset a calibration experiment using an extra plane containing a third observer, running in parallel with the main survey. This is also crucial in light of the possible retirement of the current spotter/pilot in the near future.

Background

One of the key aspects of the Recruitment Monitoring Program is the development of a fishery-independent index of juvenile SBT recruitment based on a scientific aerial survey in the Great Australian Bight (GAB). 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. 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).

Given the logistic problems of finding spotters and spotter/pilots to conduct the scientific line-transect survey, the feasibility of using experienced commercial tuna spotters to collect data on SBT sightings in the GAB during and between commercial operations, was explored after the suspension of the line-transect survey. This approach consisted of two parts:

- (1) a voluntary reduced line-transect component based on the 2000 scientific aerial survey design (Cowling 2000), and
- (2) a 'commercial' spotting component based on SBT sighted per unit of searching effort (a SAPUE index).

The voluntary nature of the reduced line-transect component led to it being both substantially reduced and highly ad-hoc in terms of the timing, location and number of transects flown. This, together with the high variability in estimates resulting from reduced effort (CV ~45-154%), suggests that an ad-hoc reduced survey would not provide a reliable indicator of juvenile SBT abundance. The commercial spotting data provided preliminary fishery-dependent indices of SBT abundance (SAPUE index) for 2002-2004. However, the indices are difficult to interpret (e.g. different ways of defining type of effort), and suffer from many of the problems which make longline catch per unit effort (CPUE) difficult to interpret (e.g. substantial changes in coverage over time; non-random coverage and areas with no coverage in some years). Although the SAPUE index may provide a qualitative indicator of juvenile SBT abundance in the GAB, it has always been recognised that a line-transect survey with consistent design and protocols from year to year is highly preferable.

In 2005, a full scientific line-transect aerial survey in the GAB was re-established. The data from all survey years (1993-2000 and 2005) were analysed using an adaptation of previously developed methods, and an index of abundance was constructed (Bravington et al. 2003). In addition, SBT sightings data from commercial tuna spotters continued to be collected over the 2005 fishing season.

Recognising the importance of having a time-series of indicators, the full scientific line-transect aerial survey continued in 2006, along with collection of commercial spotting data. During this year, improvements to the analysis methods for the aerial survey data also continued to be explored. This report pertains only to the scientific aerial survey; it summarises the field procedures and data collected during the 2006 season, describes changes made to the analysis methods, and presents results from applying these methods to the data from all survey years.

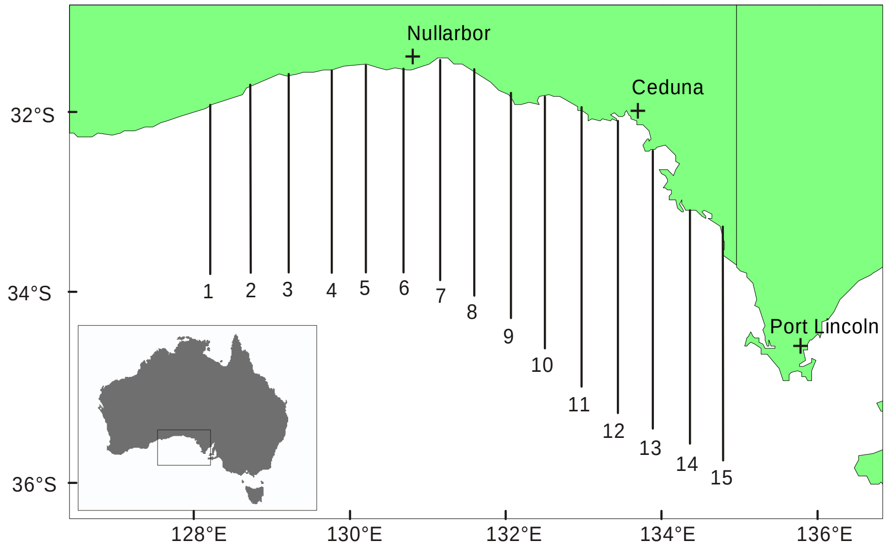
Field procedures

A line-transect aerial survey was conducted in the GAB between 7 January - 18 March 2006. The same spotter, spotter-pilot and data recorder employed for the 2005 survey were used in 2006. The survey followed the protocols used in the 2000 and 2006 aerial 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 (i.e. spotter and spotter-pilot). 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 transect-lines for the scientific aerial surveys 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 the secondary sightings could be seen from the transect (when the plane returned), that was recorded. Only secondary sightings that could be seen from the transect were included in the analysis.

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.

2006 survey year

The line-transect survey was successfully completed in 2006. A total of 3.5 replicates of the GAB were completed, compared to 3.75 (1 extra line) in 2005. The total flying time (transit and transect time) for the 2006 survey was 151.3 hours.

The two-plane experiments to calibrate between observers were not conducted in 2006. As in 2005, we planned to conduct these experiments in March when observers were likely to become available, and after considerable effort had been put to the survey itself. Unfortunately, when observers and planes were available in March, the weather was not suitable. We were able to charter the survey plane for an additional week at the beginning of April, but again the weather was unsuitable for the experiments. Given that we have been unable to conduct the calibration experiments for two seasons now, it seems unlikely that this approach will succeed. Better results may come from running a calibration experiment during the main survey, using an extra "follower" plane containing a third observer. Preliminary analysis suggests that at least 50-70 sightings of SBT are needed during the calibrations to achieve reasonable levels of precision.

Data preparation

In the months prior to the 2006 survey, a thorough investigation of the aerial survey database maintained by CSIRO was conducted. A number of data inconsistencies and data entry errors were identified, investigated and corrected where possible. The whole process was time-consuming, but the historical database now appears to be in good shape and this task should not need revisited in future years (at least not to the same extent). Also, the incorporation of future data should be simplified by new automated error-checking code, though human scrutiny will always remain important. We also plan to increase the number of automated checks in future.

Small changes to the design of the database were proposed to make it more convenient for subsequent analyses, and these changes were implemented by the database manager after consultation. The most important of these changes was adding a field for each recorded interval of a flight that specifies whether the plane was in search mode along a transect line, or whether it was off the line, either investigating a sighting or else in transit between lines. The time during which the plane is off the line should not be included as valid search effort. Whether the plane is on or off the line is not directly recorded during flight, so it must be inferred from other recorded information. While there was already a field in the database intended to convey this information, the automated process used to produce it often gave the wrong answer in complicated situations (e.g. when a second sighting was made and investigated prior to the initial sighting being investigated) or in situations with missing data (e.g. waypoints that were not recorded). The new field is also automatically produced but it is more robust to such errors.

The data collected from the 2006 survey were loaded into the aerial survey database, which already contained the data collected from all previous surveys. The 2006 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 are 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 un-extended (pre-1999) survey area are included – refer to last year's CCSBT report (Bravington et al. 2005) for details. In keeping with previous analyses, only sightings made within 6 nm of a transect line are included (again, refer to the Bravington et al. 2005 for details). Note that if a sighting consists of more than one cluster, then the sighting is included if at least one of the clusters is within 6 nm of the line.

Every patch in the database is recorded as being either primary or secondary, where a secondary patch is one that was not part of the original sighting. For example, while the plane is off the transect line investigating a potential SBT sighting, additional patches not seen from the transect line may be spotted. These additional patches could be part of the sighting being investigated (that were not seen from the line), or else part of a new sighting. In principle, secondary patches should not be included in the analysis to estimate abundance. Unfortunately, the protocols (or interpretation of the protocols) for determining what is a secondary patch appears to have changed over the years. This is suggested by the fact that there were a large number of patches recorded as being secondary in early survey years, and almost none in recent years (Table 1). This issue is discussed in the Aerial Survey Design Workshop Report 1999 (20-21 September 1999, Queensland) and possible explanations are suggested, but no conclusions are reached. In the 2005 analysis, we included secondary patches based on the hypothesis that the protocols have been carried out differently over time, so that patches that were considered secondary in early years would have been considered primary in later years. However, on further investigation and discussion, we have decided to exclude secondary patches from the current analysis. The survey protocol changed in 1998 so that sightings beyond 7 nm from the line were no longer investigated or recorded. The fact that over 50% of the secondary patches in the early years were located more than 7 nm from the transect line supports the alternate hypothesis that these patches were part of distant sightings no longer investigated in later years. Furthermore, biomass estimates were recorded for only 10% of secondary patches. Thus, even if we were to include secondary patches in our analysis, they would contribute very little information (over half would be excluded because they are more than 6 nm from the line and those remaining would have very little biomass information). The decision to exclude secondary sightings was based on our best judgment in light of available information; although unlikely, this decision could change in future should additional information come available on which to better base it.

Table 1. Number of primary and secondary patches recorded in the database by survey year.

Patch type	1993	1994	1995	1996	1997	1998	1999	2000	2005	2006
Primary	624	708	761	658	448	262	135	213	206	108
Secondary	114	69	46	1	0	0	15	0	1	0

Another issue that was revisited in the current analysis was what to do with transect lines that have been aborted due to poor weather and, thus, poor sighting conditions. In the 2005 analysis, data from aborted lines were included in the analysis. On further deliberation, we concluded that the search effort made prior to aborting a line is not comparable to search effort made in good conditions (in case tuna systematically move away from, say, windy areas where they are difficult to sight to calm areas where they are easier to sight). Thus, we decided for the current analysis to exclude any search distance and sightings made during the aborted section of a transect line. More specifically, each transect line is broken into 3 sections for analysis purposes (see upcoming 'Methods of analysis' section for details), and only data from the section in which the line was aborted is omitted.

Search effort and SBT sightings

A summary of the total search effort and SBT sightings made in each survey year is given in Table 2. 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. Because of a few changes in the criteria used to determine which data are appropriate for analysis (see previous section), the numbers in Table 2 are slightly different than those presented in Table 1 of last year's CCSBT report (Bravington et al. 2005). In terms of total distance searched, the differences mainly result from the improvements to the database that allow for better determination of when the plane was in valid search mode on a transect line, and also due to search effort made during aborted sections of lines being excluded. In terms of the information about SBT sightings, the differences mainly result from patches recorded as secondary and sightings made during aborted sections of lines no longer being included.

The total distance searched was lowest in 2006 so, not surprisingly, the number of sightings was also lowest. However, even after standardizing for distance searched, the number of sightings and total biomass were still below average in 2006 (Figure 2). While too sparse to say much, the distribution of SBT sightings appears to be similar to last year (Figure 3).

Table 2. Summary of aerial survey data by survey year. Only data considered suitable for analysis (as outlined in text) are included. All biomass figures are in tonnes.

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	12222	3.9	76	24.4	18.8	203
1994	15169	174	14996	3.3	23	26.4	21.5	245
1995	14573	179	21948	3.6	38	34.5	27.9	224
1996	12284	116	16487	4.1	46	34.6	27.3	147
1997	8813	117	9804	3.0	18	27.6	22.3	198
1998	8550	109	10236	2.3	21	40.3	20.3	944
1999	7555	56	3021	2.4	21	22.9	16.5	120
2000	6775	77	4811	2.6	17	23.9	20.0	100
2005	5968	80	6121	2.4	17	31.9	25.0	196
2006	5152	44	4064	2.0	8	47.3	31.9	270

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.***

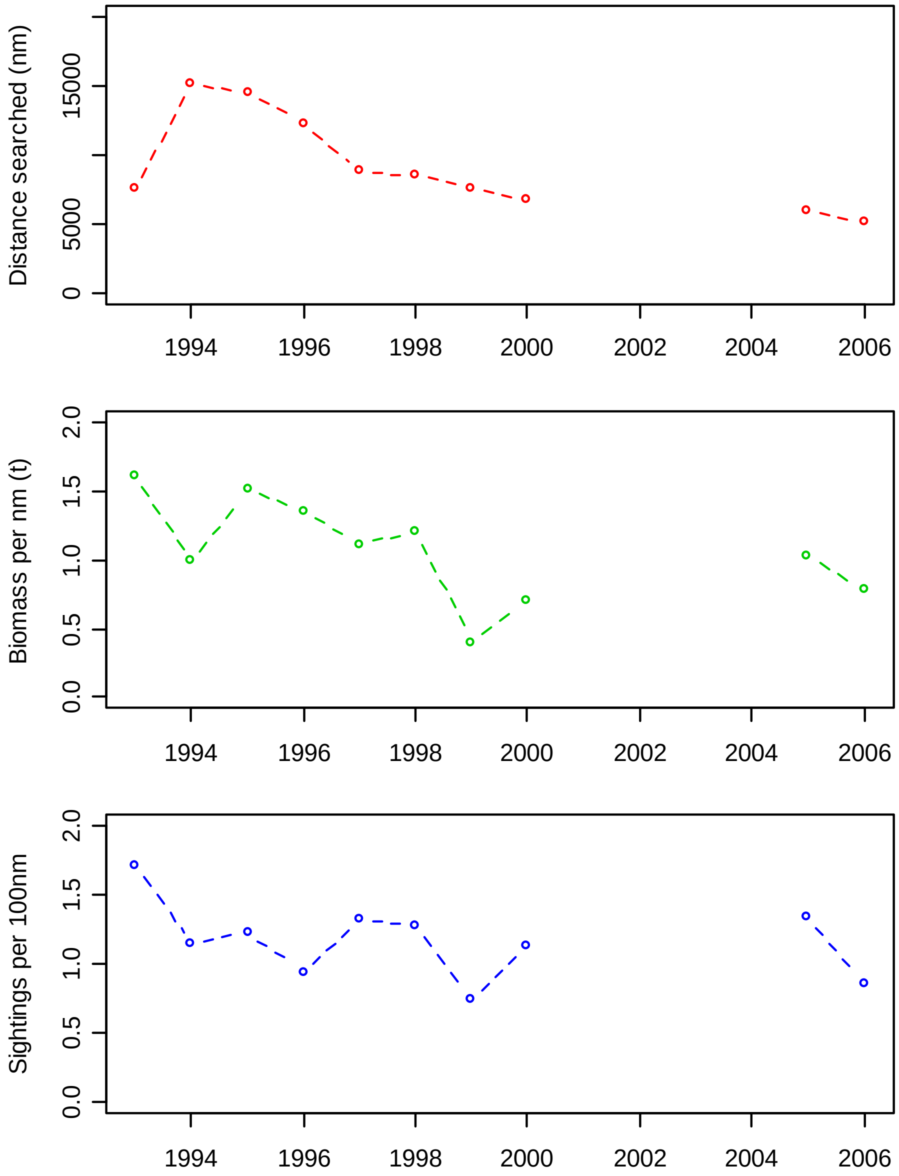
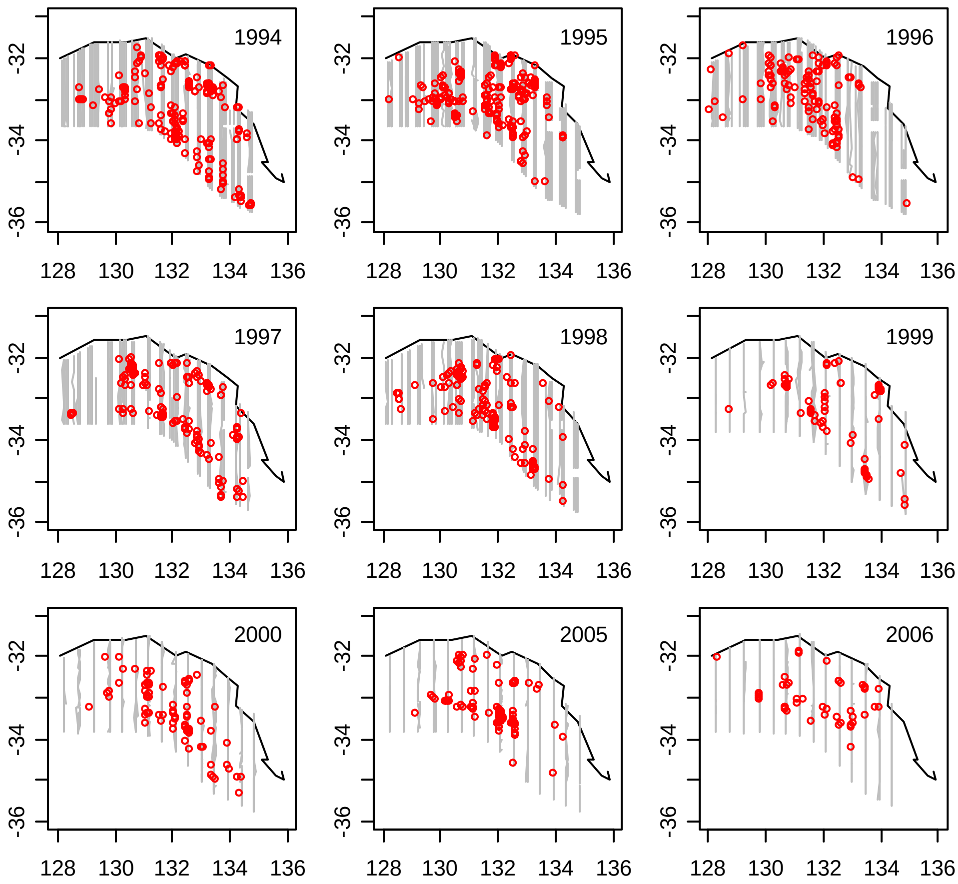


Figure 3. Distribution of SBT sightings made during each aerial survey year. Red circles show the locations of SBT sightings and grey lines show the north/south transect lines that were searched. (1993 excluded for display purposes but has similar distribution to 1994 – see last year's report, i.e. Bravington et al. 2005)



The average and maximum number of patches per sighting were lower in 2006 than in all previous survey years, and there appears to have been a declining trend in number of patches per sighting over the years of the survey (Table 2). On the other hand, the size of a patch tended to be larger than average in 2006, with the average and median being larger than in any previous year (Table 2; 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 between years for the same observer. Therefore, as in last year's report, we consider the fish size estimates made by a single observer who has operated in all survey years, including 2006. According to this observer, the SBT present in the survey area in 2006 were, on average, the smallest seen since the earliest survey years (Figure 5).

Figure 4. Frequency of SBT patch sizes (in tonnes) by survey year (excluding 1993 for display purposes).

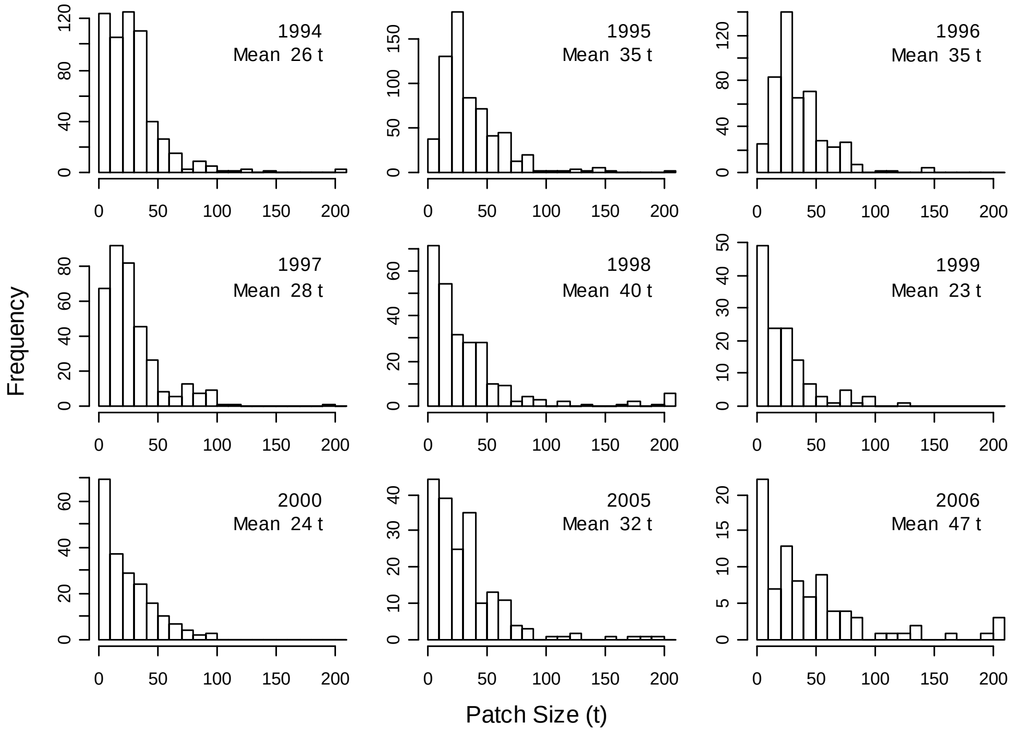
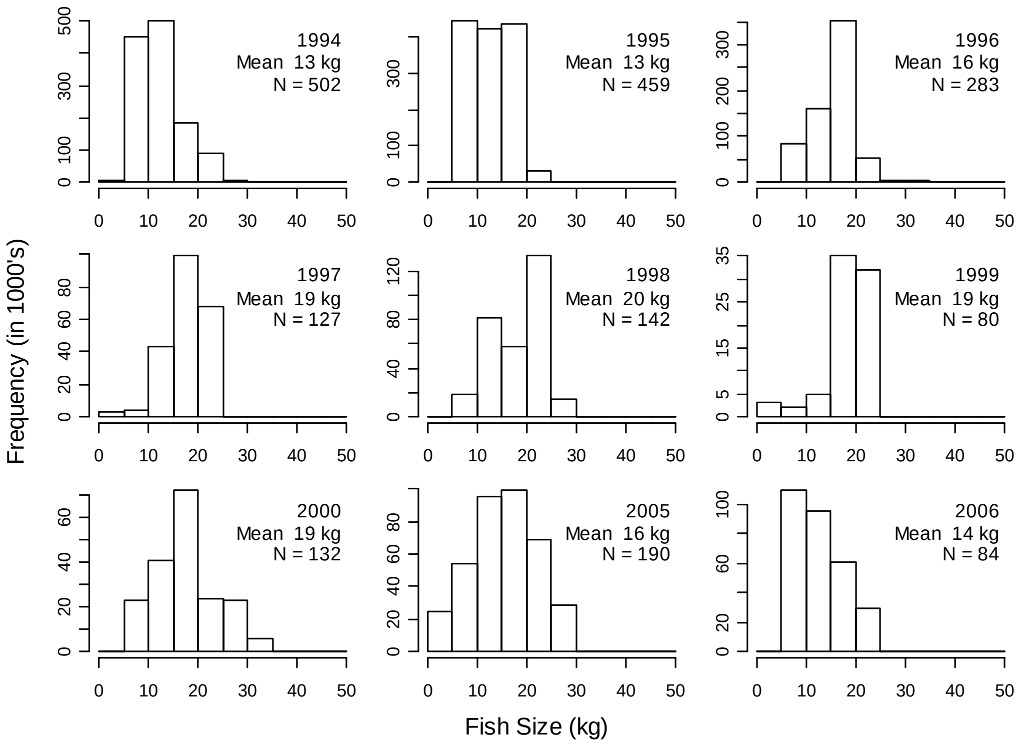


Figure 5. Size frequency of SBT (in kg) recorded by one observer who operated in all survey years. Data are weighted by patch size. N = number of patches for which fish size was estimated (1993 again excluded but has very little data, N=9).



Environmental variables

Even though sea surface temperature (SST) is known to affect tuna surfacing behaviour, we have since 2002 omitted it from analysis because of problems in deriving a consistent SST index. In 2006, though, we obtained for the first time a reliable source of SST data, which we have now incorporated into our models. The SST data for 1994 onwards were extracted from the 3-day composite SST dataset produced by CSIRO Marine and Atmospheric Research's Remote Sensing Project[1]. This dataset, which is derived from NOAA environmental satellite AVHRR thermal imagery, has a spatial resolution of 0.036^{o} latitude by 0.042^{o} longitude and a nominal temporal resolution of one day, achieved by averaging raw data from 3 consecutive days. The CSIRO dataset does not cover 1993, so for 1993 we used the AVHRR Oceans Pathfinder 8-day composite SST dataset available from the Physical Oceanography Distributed Active Archive Center[2]. The spatial resolution of this dataset is similar to the CSIRO data, but the temporal resolution is much coarser at 8 days; thus, the 1993 SST data are less reliable than for other years.

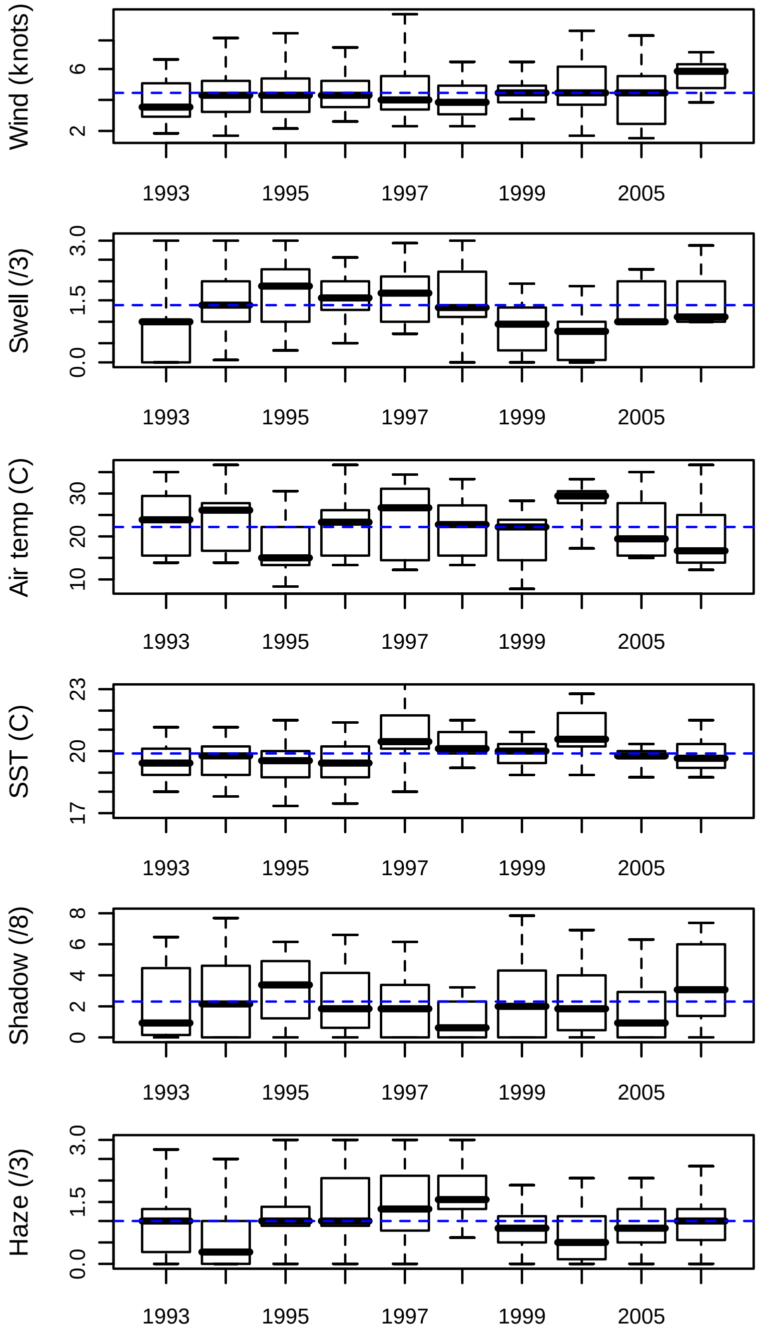
Table 3 and Figure 6 summarize the environmental conditions that were present during valid search effort in each survey year. All variables were recorded onboard the plane with the exception of SST, as described above. The conditions during the 2006 survey were generally poor, with the wind speed being much higher on average than in any other year, as well as the amount of sea shadow (also referred to as high cloud). The other environmental variables also tended to be slightly worse than in other years. The poor conditions in 2006 mean that standardization for environmental variables is very important before valid comparisons can be made between years. Unfortunately, the poor conditions are also reflected in the amount of search effort that was able to be completed in 2006, which translate to greater uncertainty in the estimate for this year. In fact, the table and figure somewhat disguise the unusual nature of 2006. January was exceptionally warm (the highest average SST of any January in the survey) while February was exceptionally cold (the lowest average SST of any February in the survey), and flying time in February was very limited. This has a big impact on the results, as discussed further below.

Table 3. 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

[1] http://www.marine.csiro.au/remotesensing/oceancurrents/ten_years_of_SST.doc.
[2] http://podaac.jpl.nasa.gov/products/product102.html

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

During the past year, improvements to the aerial survey analysis methods continued to be explored and some changes were made. The main differences in this year's analysis compared to last year's (Bravington et al. 2005) are as follows:

(i) The statistical models are simplified from three components down to two.

(ii) SST is included as a covariate in the models (see previous section), and all environmental covariates are re-investigated.

(iii) Data from all months (January-March) are included, as opposed to last year where only data from January and February were included.

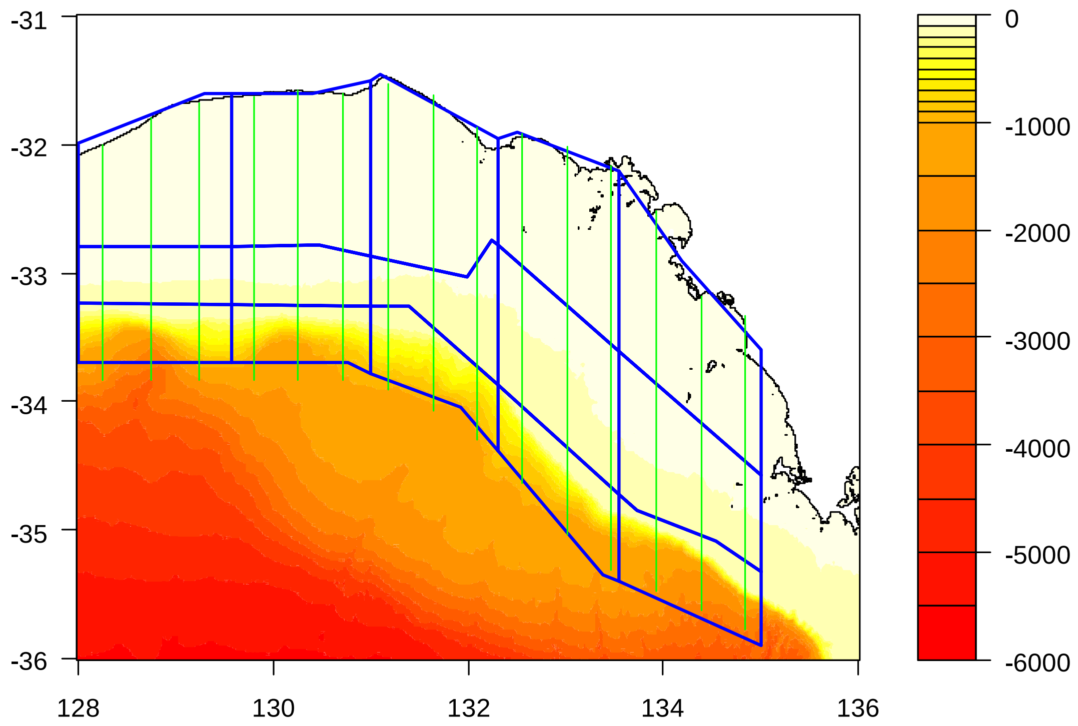
Each of these changes is discussed in more detail below.

Following on from the extensive body of work by Cowling and others in the 1990s, Bravington (2003) developed methods for analysing the aerial survey data to produce a time series of annual relative abundance indices and corresponding estimates of precision. Separate models were constructed to describe three different components of observed biomass: biomass per patch (BpP), patches per sighting (PpS), and sightings per nautical mile of transect line (SpM). Since environmental conditions affect what proportion of tuna are available at the surface to be seen, as well as how sightable 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 included 'corrections' for environmental and observer effects in order to produce standardized indices that could be meaningfully compared across years.

Last year we re-analysed all the data, including 2005, using similar models to Bravington (2003). However, in the interests of speed and simplicity, we avoided the most complicated aspects of the earlier analysis and did not re-investigate all possible environmental covariates. In 2006, we revisited the models and determined that using just two components of observed biomass—biomass per sighting (BpS) and sightings per mile (SpM)—had advantages over the three-component model in terms of parsimony and goodness-of-fit. A generalized linear model (GLM) is still used for each component, with the BpS model being very similar in structure to the previous BpP model, and the new SpM model being very similar in structure to the previous SpM model. Furthermore, we re-investigated the environmental covariates being included in the models. Most importantly, we were able to investigate the influence of SST data in the models, which had been unavailable for consideration in previous analyses.

Following the analysis in 2005, we defined 45 area/month strata as the basic units of analysis: 15 areas (5 longitude blocks and 3 latitude blocks) and 3 months (Jan, Feb, Mar). Figure 7 shows the 15 areas, and also shows how the latitudinal divisions were chosen to correspond roughly to depth strata (inshore, mid-shore and shelf-break).

Figure 7. Plot showing the 15 areas (5 longitudinal bands and 3 latitudinal bands) into which the aerial survey is divided for analysis purposes. It can be seen from the depth profile that the latitudinal bands correspond roughly to depth strata (inshore, mid-shore and shelf-break).The green 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, 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 450 year/month/area combinations (10 years and 45 area/month strata), there are in fact 13 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 investigated 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. 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 55% 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 effect size ("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 plan to formally merge the pairwise model with the SpM model to correctly propagate all the uncertainty into the final CVs.

Variation between observers is probably the biggest source of uncertainty in the aerial survey indices, and really needs to be resolved if the survey is to continue usefully. If the aerial survey goes ahead in 2007, it should certainly include a calibration experiment using an extra plane containing a third observer, running in parallel with the main survey.

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. The only differences between these covariates and those used in the previous SpM model are that SST in now included and haze is substituted for the amount of low cloud, as haze was found to be a better predictor.

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[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.

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, so we leave the details for Appendix A. However, a very brief outline is as follows. We first obtained standard errors (SEs) for the predicted values of 'standardized

[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.

[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).

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 are still investigating using a random-effects extension of the BpS and SpM models to better handle strata where there is low effective sampling effort. This should not change the point estimates much for most years, but may have a larger effect on the CVs.

Results

Figure 8 shows the estimated time series of relative abundance indices with 90% confidence intervals. The point estimates and CVs corresponding to Figure 8 are reproduced in Table 4. The estimate for 2006 is slightly lower than the 2005 estimate and both are below the average level in the mid-1990s. The CV for the 2006 estimate is higher than in most years, including 2005, due to the low amount of search effort and number of sightings in 2006, as well as to the more unusual environmental conditions. However, it is still lower than the 1999 or 2000 CVs, which are dominated by uncertainty about observer effects. We should also note that all of the CVs are likely to be on the low side because they do not fully reflect uncertainty in observer effects.

The CVs in Table 4 are lower than those presented in last year's report (Table 2 of Bravington et al. 2005). This will be due in part to the inclusion of SST as a predictor, since its explanatory power was estimated to be quite high in both the BpS and SpM models. It may also be due to the change from a 3-component to 2-component model, given that modelling biomass per sighting seems to provide a better fit than splitting into biomass per patch and patches per sighting. In particular, there were difficulties in finding a good fit for the patches per sighting (PpS) model – see Bravington 2003.

Figure 8. Time series of relative abundance estimates with 90% confidence intervals.

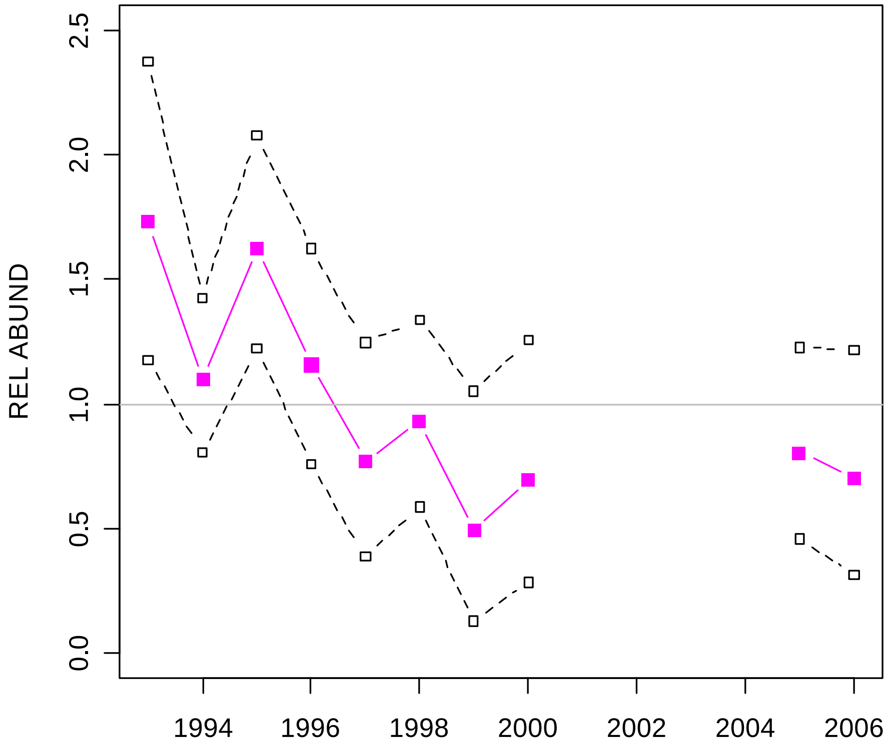
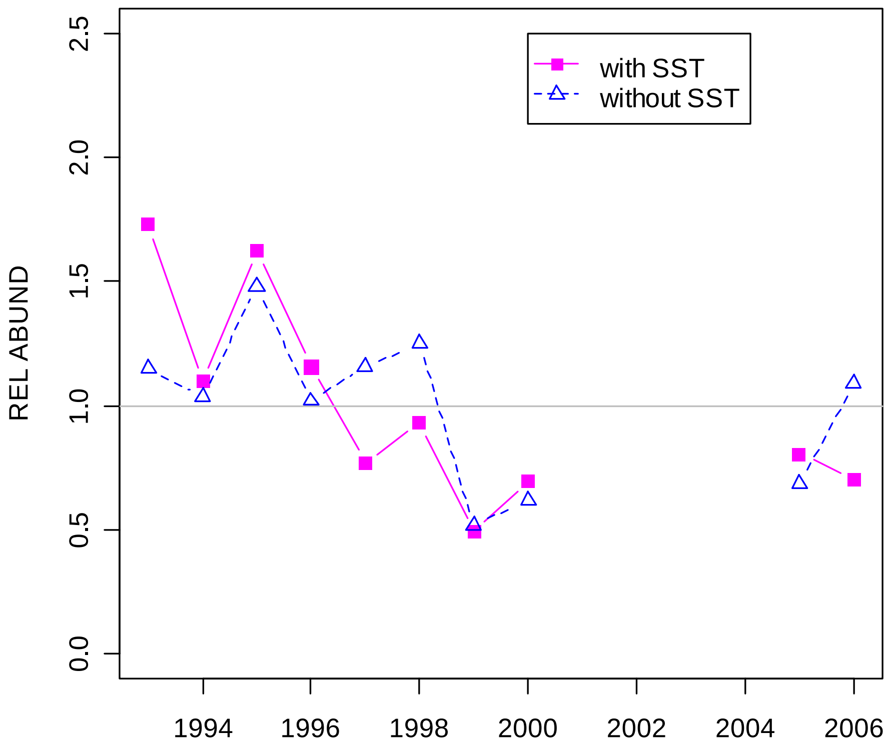


Table 4. Aerial survey index point estimates, standard errors and CVs (as per Figure 8).

	1993	1994	1995	1996	1997	1998	1999	2000	2005	2006
Estimate	1.732	1.097	1.626	1.156	0.766	0.930	0.492	0.695	0.803	0.703
SE	0.367	0.190	0.261	0.265	0.265	0.231	0.293	0.303	0.236	0.280
CV (%)	21.2	17.3	16.0	22.9	34.6	24.8	59.6	43.6	29.4	39.8

For comparison, we redid the analysis using the same environmental covariates as in last years' models, with the most important difference being that last years' models did not include SST. Figure 9 compares the without-SST series (blue triangles) and the with-SST series (pink squares). The without-SST series is very similar to that presented in Figure 6 of last year's report (Bravington et al. 2005), suggesting that the database cleanup and the switch from 3-stage to 2-stage models have had only a small impact. The general trends of the two lines are similar, but the inclusion of SST does make a big difference to the estimates in some years, including 2006. The 1993 estimate should be treated as less reliable than the others because protocols were still being established that year, and also because the SST data for 1993 come from a different source and may not be strictly comparable with subsequent years.

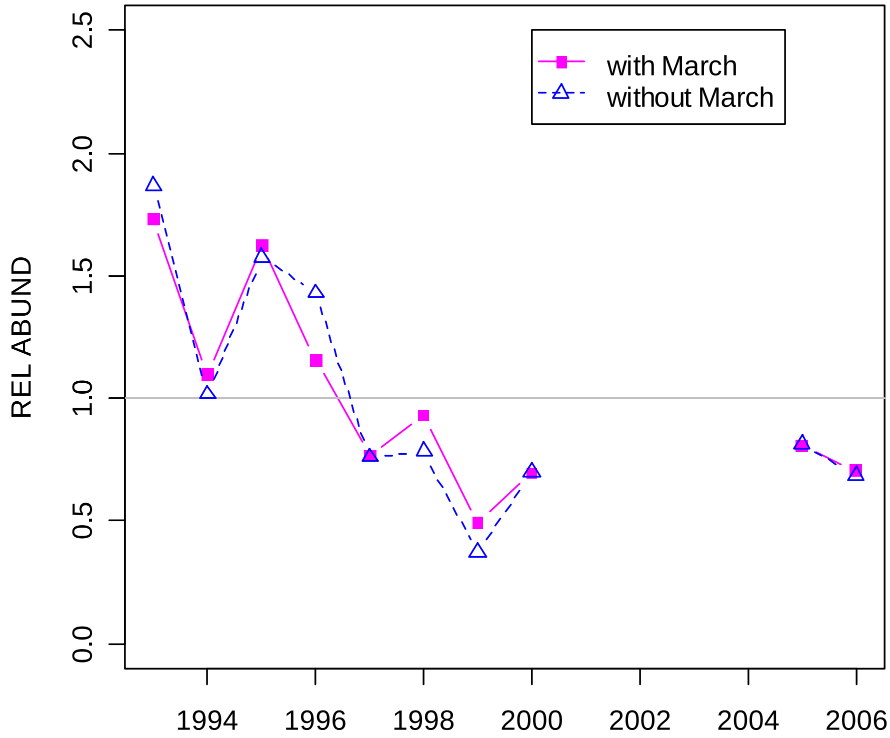
Figure 9. Time series of relative abundance estimates using: a) the newly chosen environmental covariates, which include SST; b) the environmental covariates from the analysis conducted in 2005, which do not include SST.



SST is strongly significant as an explanatory variable, so the model with SST is definitely to be preferred. However, the effect of SST on the 2006 estimate is at first surprising, since average temperatures are similar to 2005. The explanation lies in the temperature variations within the year. The high January temperatures coincided with high effort and high sighting rates, which are scaled down when the high temperatures are accounted for. In February, though, temperatures were generally very low, there were no sightings in 12 of the 15 spatial strata, and the three flights where sightings were made had better-than-average conditions for the month. Accounting for low temperatures should scale up the number of sightings, but if zero sightings are made, then it does not matter how big the multiplier is; zero remains zero. This situation is unusual in that it will only happen when conditions are poor enough to produce large numbers of strata with zero sightings. Note that the 2006 estimate should not be biased by the unusual SST conditions; however, its uncertainty, and its sensitivity to which environmental corrections are applied, is enhanced.

A further difference from last year's results lies in the months used. In the 2005 analysis, we used only historical January and February data in producing the relative abundance index. This was because poor weather in March 2005 led to very little search effort, and also because March has historically been the least informative month, with generally lower search effort and higher variability in search effort than other months. Unfortunately, due to poor weather in February 2006, much of the 2006 search effort came in March, so that March could not be dropped. This means March data from all years had to be included in our analysis in order to get an index that is comparable across years. In order to see how see how much the inclusion of March affected the results, we calculated the index using data from only January and February and compared it with the index calculated using data from all months (Figure 10)—fortunately, the inclusion of March does not alter the index much.

Figure 10. Relative abundance indices calculated using data from all months (January, February and March) compared to those calculated using data from January and February only. Note that in both cases, the new environmental covariates that include SST were used.



Summary

The estimate from the 2006 scientific aerial survey shows a small drop from the 2005 estimate, and both are below the average level in the mid-1990s. The mean of the 2005 and 2006 estimates is about 66% of the mean of the 1994-1998 estimates. The coefficient of variation (CV) of the 2006 estimate is higher than in most years, including 2005, due to the low amount of search effort and number of sightings in 2006, as well as to the anomalous environmental conditions. However, it is still lower than the 1999 or 2000 CVs, which are dominated by uncertainty about observer effects. Variation between observers remains the chief source of uncertainty in the aerial survey indices, and really needs to be resolved if the survey is to continue usefully.

Since last year's analysis (Bravington et al. 2005), we have made some changes to the analysis methods and to the historical database. The main impact has been the use of SST data to improve the adjustments for environmental variability; this has a fairly strong effect on the index in some years, including 2006 for which the SST conditions were unusual. Other changes to the methods and database have streamlined the analysis considerably, but have had comparatively minor impacts on the abundance index itself.

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.

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.

Appendix A – CV calculations

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

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

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

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

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

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

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 because the environmental/observer conditions used in the BpS (likewise SpM) model are correlated between 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. $\theta = (\theta_{\mathrm{BpS}}, \theta_{\mathrm{SpM}})$

[6] Recall that θ contains the environmental/observer coefficients from both the BpS and SpM models, so $\mathbf{V}_\theta = \begin{bmatrix} \mathbf{V}_{\theta_{\mathrm{BpS}}} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{\theta_{\mathrm{SpM}}} \end{bmatrix}$. 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

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