CSIRO

The aerial survey index of abundance: updated analysis methods and results for the 2010/11 fishing season

Paige Eveson
Jessica Farley
Mark Bravington

Table of Contents

Abstract

The estimate of relative juvenile abundance from the 2011 scientific aerial survey shows a substantial increase from 2010 and is similar to the 1993 estimate, which was the highest estimate of all survey years.

The 2011 survey was the first in which all flights had only one spotter (i.e., no flights had a spotter-pilot). Calibration experiments conducted in 2008 and 2009 showed that a plane with only one spotter makes fewer sightings than a plane with two spotters. A method of estimating this difference (which we call the "calibration factor") and accounting for it in the analysis of the survey data was developed in Eveson et al. (2009, 2010). The presence of a high proportion of large schools this year led us to reconsider the calibration factor estimated previously, since a plane with one spotter is less likely to miss very large schools. Re-analysing the data from the calibration experiments leaving out very small sightings (<2 tonnes) led to a revised calibration factor estimate of 0.7 instead of 0.5 (i.e., a plane with only one spotter makes approximately 70% as many sightings as a plane with two spotters). Methods to incorporate uncertainty in the calibration factor estimate were also developed and applied this year, so the CVs for the relative abundance estimates include this additional uncertainty.

Further complicating the analysis this year was the high proportion of schools comprised of small (estimated to be 1-year-old) fish. Such small fish were far less common in past survey years. In the CCSBT operating model (OM) and management procedures (MP), the aerial survey index is assumed to provide a relative time series of age 2-4 abundance in the Great Australian Bight. Thus, for consistency with the OM and MP as well as general consistency in interpretation of the index across years, schools of small fish (<8 kg, estimated to be 1-year-olds) were omitted from the analysis.

The environmental conditions during the 2011 survey were variable, with above average conditions in January and early February, followed by poor conditions in late February and March.

Introduction

The index of juvenile southern bluefin tuna (SBT) abundance based on a scientific aerial survey in the Great Australian Bight (GAB) is one of the few fishery-independent indices available for monitoring and assessment of the SBT stock. 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 observers (spotters). The suspension also allowed for further data analysis and an evaluation of the 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 and weak year classes in the late 1990s and early 2000s, this fishery-independent index is even more important (Anon 2008).

In 2005, the full scientific line-transect aerial survey was re-established in the GAB, and this survey has been conducted each year since. New analysis methods were developed and have subsequently been refined. Based on these methods, an index of abundance across all survey years has been constructed.

In addition, in 2007 a large-scale calibration experiment was initiated with the primary purpose of comparing SBT sighting rates by one observer versus two observers in a plane. This was done in light of the fact that future surveys would have only one observer in a plane (as was the case for one of the two planes flying in the 2010 survey and both planes flying in the 2011 survey). The data provided useful information about differences in sightings between observers (e.g., sightings made by one observer are often missed by another observer). However, it proved difficult to definitively estimate the effect of the number of observers on the index.

In 2008 and 2009 a new calibration experiment was designed and run in parallel with the full scientific aerial survey. This calibration experiment was designed to compare:

- the number of SBT sightings;

- and total estimated biomass of SBT observed;

by the single observer plane versus the survey plane (with two observers) over the same area and time strata.

This report summarises the field procedures and data collected during the 2011 season, which was the first season that both survey planes had only one observer. A method for accounting for the fact that a plane with one observer makes fewer sightings than a plane with two observers was developed in Eveson et al. (2009, 2010) based on data from the calibration experiments. These methods were refined this year and applied in the analysis. A couple of small additional changes were made to the analysis methods this year, including a change to the way in which the observer effect is included in the sightings model (as discussed in Appendix A), and the omission of schools of 1-year-old fish from the analysis (as discussed in the Data preparation section). The current methods for analysing the data are described, and results are presented from applying these methods to the data from all survey years.

Field procedures

The 2011 scientific aerial survey was conducted in the GAB between 1 January and 31 March 2011. As for previous surveys (e.g. Eveson et al. 2009; 2010), two Rockwell Aero Commander 500S were chartered for the season - one for the full three months (plane 1) and a second for January and February only (plane 2). Each plane contained one observer and a non-spotting pilot. The same observers employed for the 2007 to 2010 surveys were used in the 2011 survey. In addition, a spotter-pilot used in the 1993-1998 and 2008-10 surveys was used as an observer in one of the planes in January.

The survey followed the protocols established for the 2000 survey (Cowling 2000) and used in all subsequent surveys with respect to the area searched, plane flying height and speed, minimum environmental conditions, time of day the survey lines were flown, and data recording protocols. Fifteen north-south transect lines (Figure 1) were surveyed. A complete replicate of the GAB consists of a subset of 12 (of the 15) 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, as SBT abundance is historically low in these areas and surveying a subset increases the number of complete replicate of the GAB in the survey.

When flying along a line, the single observer searched the sea surface for patches of SBT from his side of the plane (the right side) through 180° to the other side of the plane (the left side). When both planes were surveying, they always surveyed neighbouring blocks. The blocks were chosen with the aim of allowing both planes to complete each block at least once per replicate. When conditions allowed for only one plane to survey (e.g. only one block was suitable), then preference was given to plane with the observer that had not surveyed that block.

The 2011 field operation was successful, largely due to the above average weather conditions in January and early February, and the availability of two planes at that time. The weather in late February and March was particularly windy, cloudy and wet. Despite this, almost 7 replicates of the GAB were completed in 2011, which is similar to 2010 but higher than the 3-5 replicates for the preceding 5 years. The total flying time (transit and transect time) for the 2011 survey was 204.8 hours, compared to 213.6 hours in 2010.

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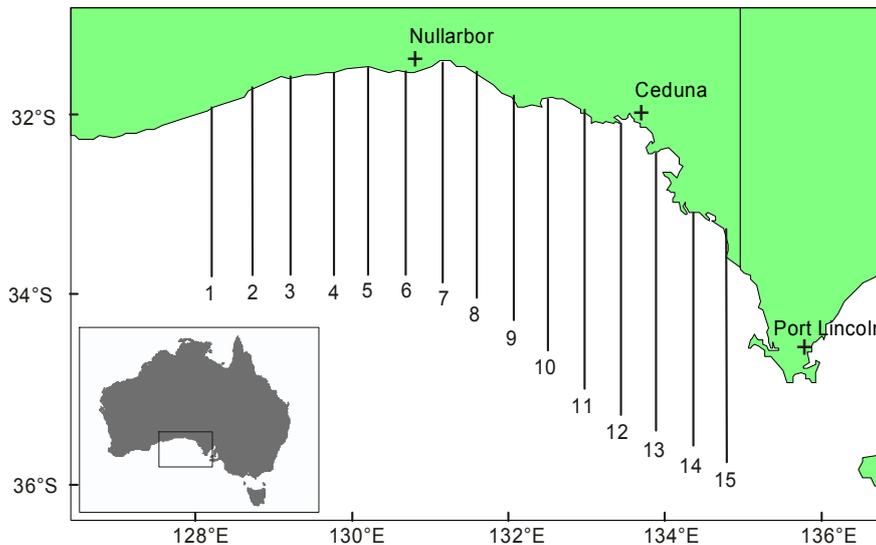
Data preparation

The data collected from the 2011 survey were loaded into the aerial survey database and checked for any obvious errors or inconsistencies and corrections made as necessary. In order for the analyses 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 sightings made

within 6 nm of a transect line were included (see Basson et al. 2005 for details). In cases where a sighting consisted of more than one school, then the sighting was included if at least one of the schools 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 details).

This year's data included a high proportion of schools of small (1-year-old) fish, which was unusual compared to past survey years. The percent of the total tonnage of SBT spotted that was estimated to be fish less than 8 kg was 30.8% in 2011. The next highest was 16.1% in 2010, followed by 13.2% in 2009, and ranging from 0.6 to 8.8% in other years (Table 1). Because the spotters estimate average weight, not age, of fish in schools, we use 8 kg as an average weight cut-off between 1-year-olds and 2-year-olds. An 8 kg cut-off was used because it equates to a length of 73 cm using the weight-length relationship of Robins (1963), and a 73 cm fish is a reasonable estimate of the upper limit of a 1 year old fish (and lower limit of a 2 year old fish) based on both the GAB direct age data (Farley et al., 2011) and the CCSBT growth curve (Eveson, 2011). When two spotters are in a plane, they both make independent estimates of the size of fish in each school, so we averaged their estimates to come up with a single estimate of average fish size. Although early validation experiments found that spotters' estimates of average fish size within a school could be inconsistent (Cowling et al. 2002), we are fairly confident based on conversations with the spotters that they can distinguish schools of very small (1-year-old) fish from schools of larger/older fish.

In the CCSBT operating model (OM) and management procedures (MP), the aerial survey index is assumed to provide a relative time series of age 2-4 abundance in the Great Australian Bight. Thus, for consistency with the OM and MP as well as general consistency in interpretation of the index across years, schools estimated to be comprised of 1-year-old fish (i.e., that had an average fish size estimate of less than 8 kg) were omitted from the analysis.

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. All of the values are based on raw data, which have *not* been corrected for environmental factors or observer effects. This table, and all summary information and results presented in this report, include only the data outlined in the previous section as being appropriate for analysis. Recall the change this year, in that we are omitting schools comprised of fish less than <8 kg on average. This has the biggest effect on the current year, but affects all years to some degree. Also note that the summary statistics for 2010 include data from all flights, some of which had only one observer; for 2011, all data comes from flights with only one observer.

The total distance searched in 2011 was similar to last year, which was the greatest since 1996. This was due to having two planes available to fly during the first half of the survey when weather conditions were favourable. The raw sightings rate (number of sightings per 100 nm) was about average, but the biomass per nm was highest of all survey years due to a lot of large patches (Table 1, Figure 2). However, it is important to remember that the statistics for 2010 and 2011 include data from flights with only one

observer, so caution must be used in comparing them to previous years (which is why we have chosen not to include the figure with the raw sightings rate and biomass per nm plotted against year).

Similar to last year, sightings in 2011 were concentrated in the eastern half of the survey area, with the greatest concentration along the shelf-break (Figure 3). The data this year continues to support that there has been a general eastward shift in the distribution of SBT sightings over the years.

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. The statistics differ from those reported in previous reports because schools of small fish (<8 kg) are being omitted. **All values in the table are based on raw data, which have not been corrected for environmental factors or observer effects.**

Survey year	Total distance searched (nm)	Number SBT sightings	Sightings per 100nm	Total biomass	Biomass per nm	Average patches per sighting	Max patches per sighting	Average biomass per patch	Max biomass per patch	% biomass < 8 kg
1993	7603	129	1.70	12212	1.61	4.0	76	24.5	203	0.2
1994	15180	160	1.05	13948	0.92	3.3	23	26.4	246	7.3
1995	14573	165	1.13	20101	1.38	3.5	38	34.6	225	8.8
1996	12284	110	0.90	16020	1.30	4.0	46	36.4	147	3.7
1997	8813	101	1.15	9145	1.04	3.2	18	28.5	202	8.2
1998	8550	104	1.22	9750	1.14	2.2	21	42.0	964	6.2
1999	7555	50	0.66	2992	0.40	2.5	21	24.1	121	1.4
2000	6775	76	1.12	4797	0.71	2.6	17	24.7	100	0.8
2005	5968	79	1.32	5968	1.00	2.4	17	31.8	194	2.1
2006	5150	43	0.83	4011	0.78	2.0	8	47.2	268	0.6
2007	4872	41	0.84	3510	0.72	2.6	11	33.1	121	0
2008	7462	121	1.62	7979	1.07	3.5	24	19.0	314	0.7
2009	8101	145	1.79	7875	0.97	2.5	22	22.1	170	13.2
2010[1]	10559	184	1.74	18398	1.74	4.0	41	24.8	531	16.1
2011[2]	10148	135	1.33	18457	1.82	2.7	37	49.9	400	30.8

[1] Data comes from flights with one observer as well as flights with two observers.
[2] All data comes from flights with one observer.

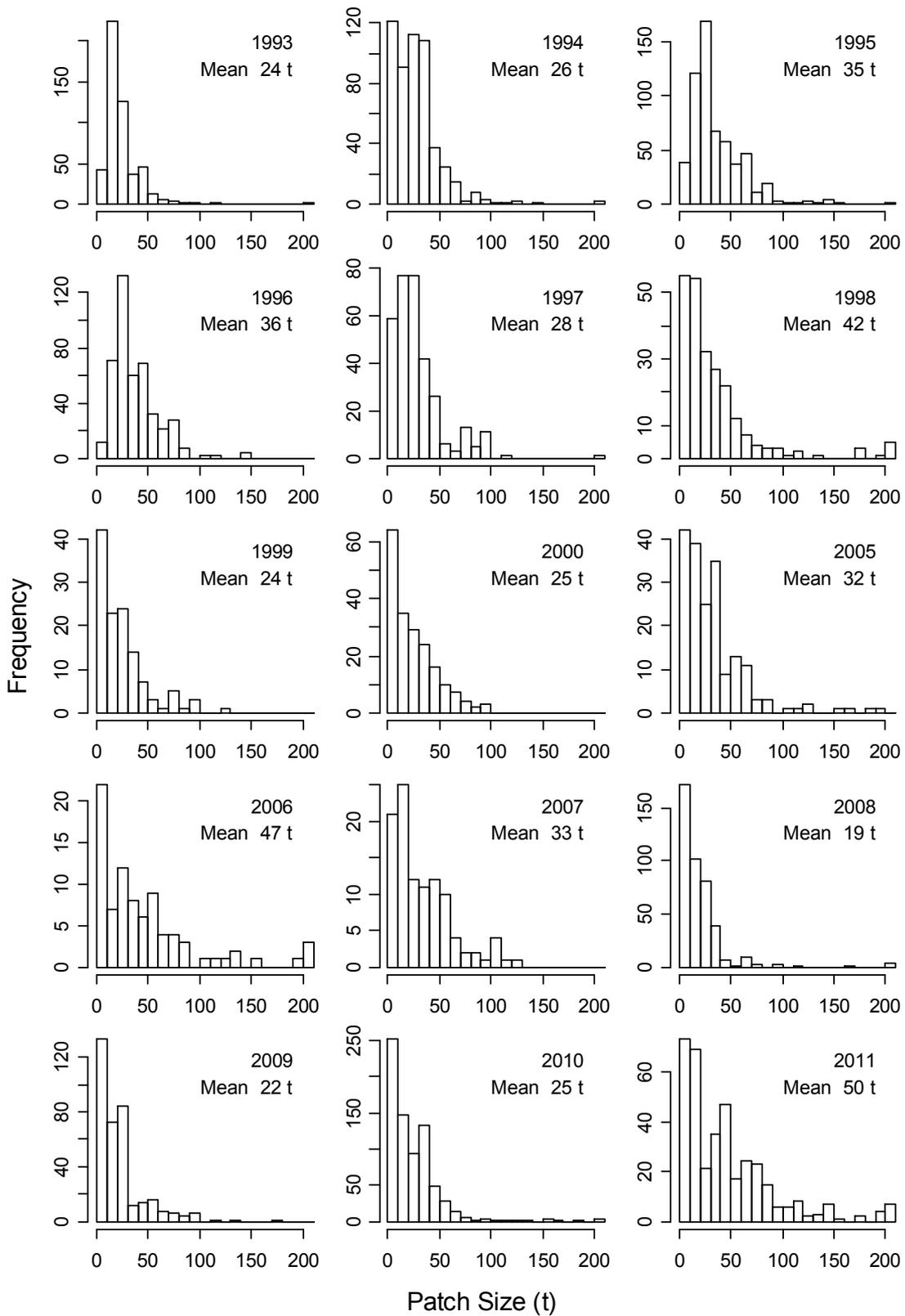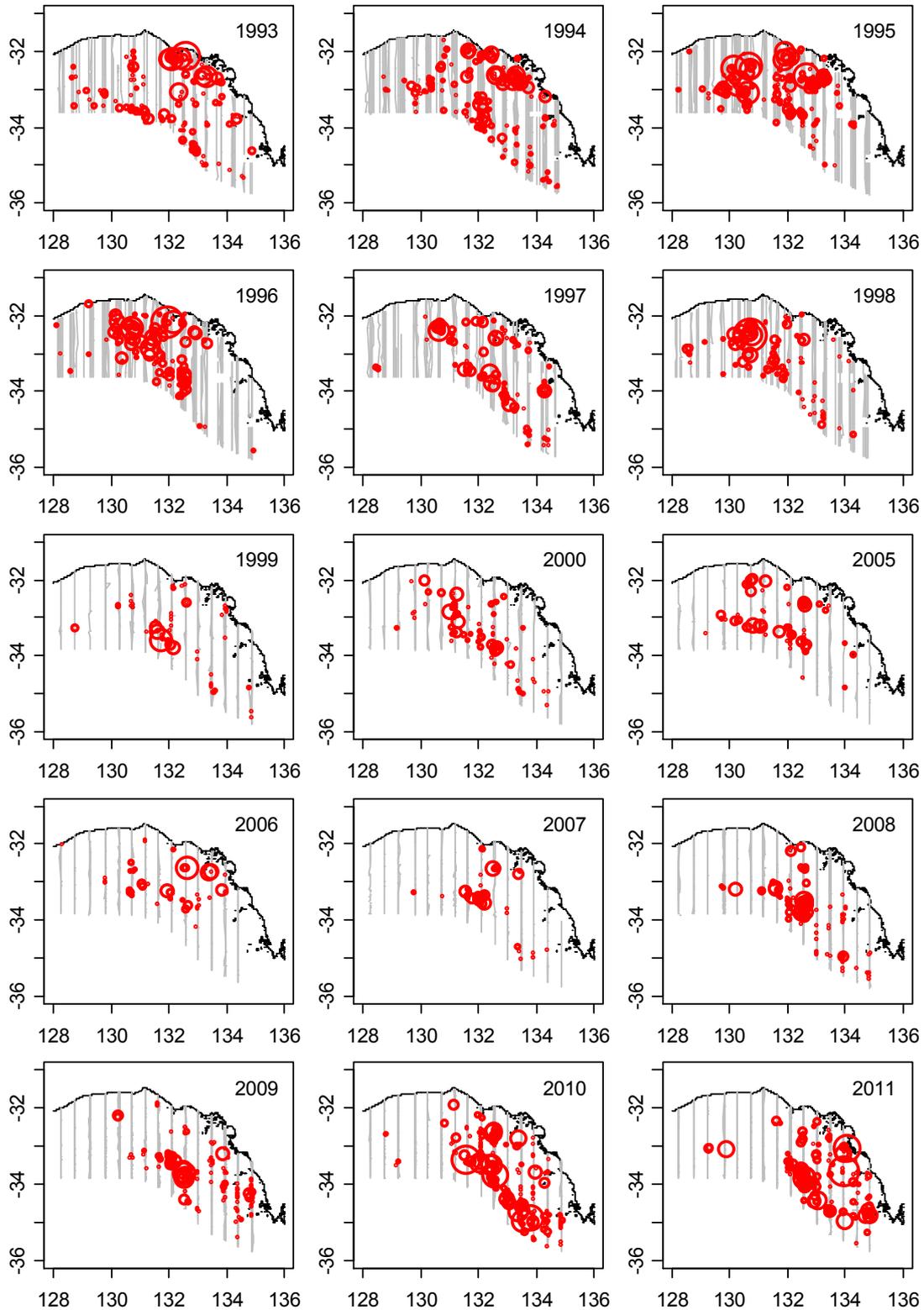Figure 2. Frequency of SBT patch sizes (in tonnes) by survey year.

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.

Environmental variables

Table 2 and Figure 4 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).

The environmental conditions during the 2011 survey were fairly average when taken over all months of the survey (Table 2; Figure 4). However, as noted in the Field procedures section, the conditions were highly variable, with above average conditions in January and early February, followed by poor conditions in late February and March.

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	1.9	0.9
1994	4.1	1.5	22.6	19.7	2.8	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	1.6	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	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	20.1	1.6	0.8
2006	5.6	1.5	20	20.1	3.5	0.9
2007	5.8	1.3	21.6	20.8	2	1.3
2008	3.8	1.4	24.2	20.4	1.4	0.9
2009	3.8	1.4	22.3	21	2.2	1.7
2010	3.5	1.1	23.6	21.2	1.8	1.2
2011	3.9	1.3	20.2	20.2	2.8	1.4

Figure 4. 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 average across all survey years.
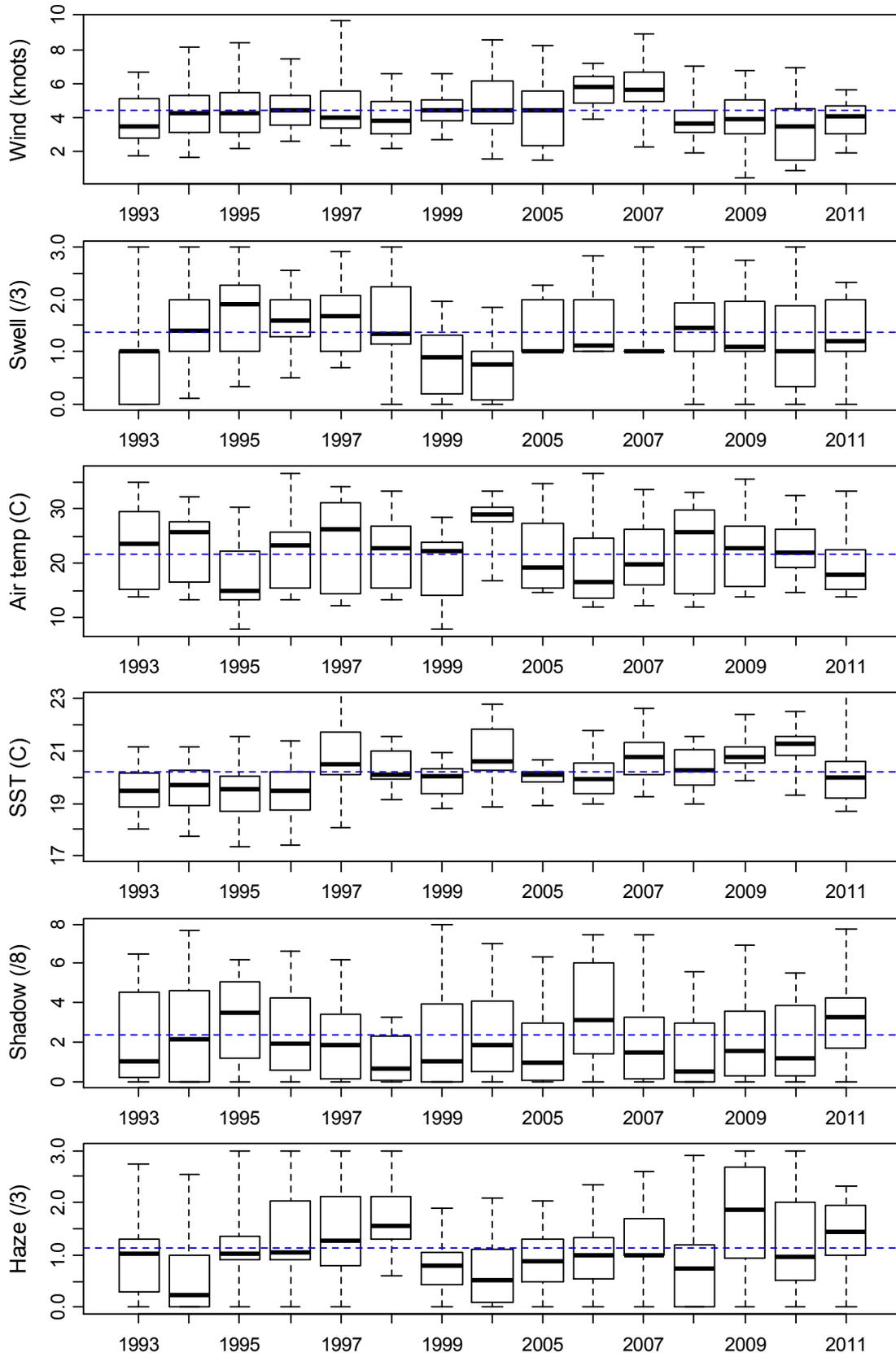
Methods of analysis

The methods of analysis this year follow the methods described last year for including data from one-observer flights, with a few small changes. Details of the methods can be found in Appendix A, but here we give a brief description highlighting where the analysis has been changed.

We fit generalized linear models to two different components of observed biomass—biomass per sighting (BpS) and sightings per nautical mile of transect line (SpM). We included the same environmental and observer variables in the models as in the past several years. The only change is that in the SpM model we include "observer effect" as an offset (i.e., as known) rather than as a linear covariate. The reason for this is discussed in Appendix A. Because of this, we need to account for uncertainty in the observer effect estimates through other methods. Such methods have been developed but we were not successful in implementing them in time for this report. Thus, the standard errors, CVs and confidence intervals for the relative abundance indices reported in Table 3 do not include uncertainty in the observer effects for the SpM model (meaning they are slightly too small).

Specifically, the models can be expressed as:

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

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

Year, Month, Area and MoonPhase were fit as factors; all other explanatory variables were fit as linear covariates. Note that the term Year*Month*Area encompasses all 1-way, 2-way and 3-way interactions between Year, Month and Area (i.e., it is equivalent to writing Year + Month + Area + Year:Month + Year:Area + Month:Area + Year:Month:Area).

In both models, the 2-way and 3-way interaction terms between Year, Month and Area were fit as random effects, whereas the 1-way effects were fit as fixed effects. Many of the 2-way and 3-way strata have very few (sometimes no) observations, which causes instabilities in the model fits when treated as fixed effects. One main advantage of using random effects is that when little or no data exist for a given level of a term (say for a particular area and month combination of the Area:Month term), we still have information about it because we are assuming it comes from a normal distribution with a certain mean and variance (estimated within the model).

In order to account for including data from one-observer flights, the only change to the BpS model is that there is only one biomass estimate per school, so it is not necessary to take an average over the estimates made by two observers (refer to "Biomass per sighting (BpS) model" section in Appendix A).

With regard to the SpM model, we know from the calibration experiments conducted in 2008 and 2009 that a plane with only one observer makes fewer sightings than a plane with two observers. Based on an analysis of the calibration experiment data conducted in 2009, we estimated that, on average, a plane with one observer will make about half as many sightings as a plane with two observers (Eveson et al. 2009). We referred to this factor as the

"calibration factor". The presence of many large schools of SBT this year (Figure 2) led us to reconsider our calibration experiment results, since a plane with one spotter seems less likely to miss very large schools (in fact, it was the spotters involved in this year's survey who raised this issue). Re-analysing the data from the calibration experiments leaving out sightings of small schools confirmed that the calibration factor (i.e., the proportion of sightings made by one observer compared to two observers) did indeed increase when small sightings were excluded (Figure 5). We can see that the estimates increase rapidly when really small sightings are left out of the analysis (<2 t), then start to level off. We fit a power curve through the data, and used the fitted value corresponding to a sightings cut-off of 2 t, namely 0.7, as our revised calibration factor estimate (i.e., a plane with only one observer makes approximately 70% as many sightings as a plane with two observers).

Figure 5. Estimated proportion of sightings made by one observer compared to two observers (i.e., calibration factor) when small sightings (less than x tonnes for various values of x) are omitted from the analysis. The red line is a power curve fit through the point estimates.
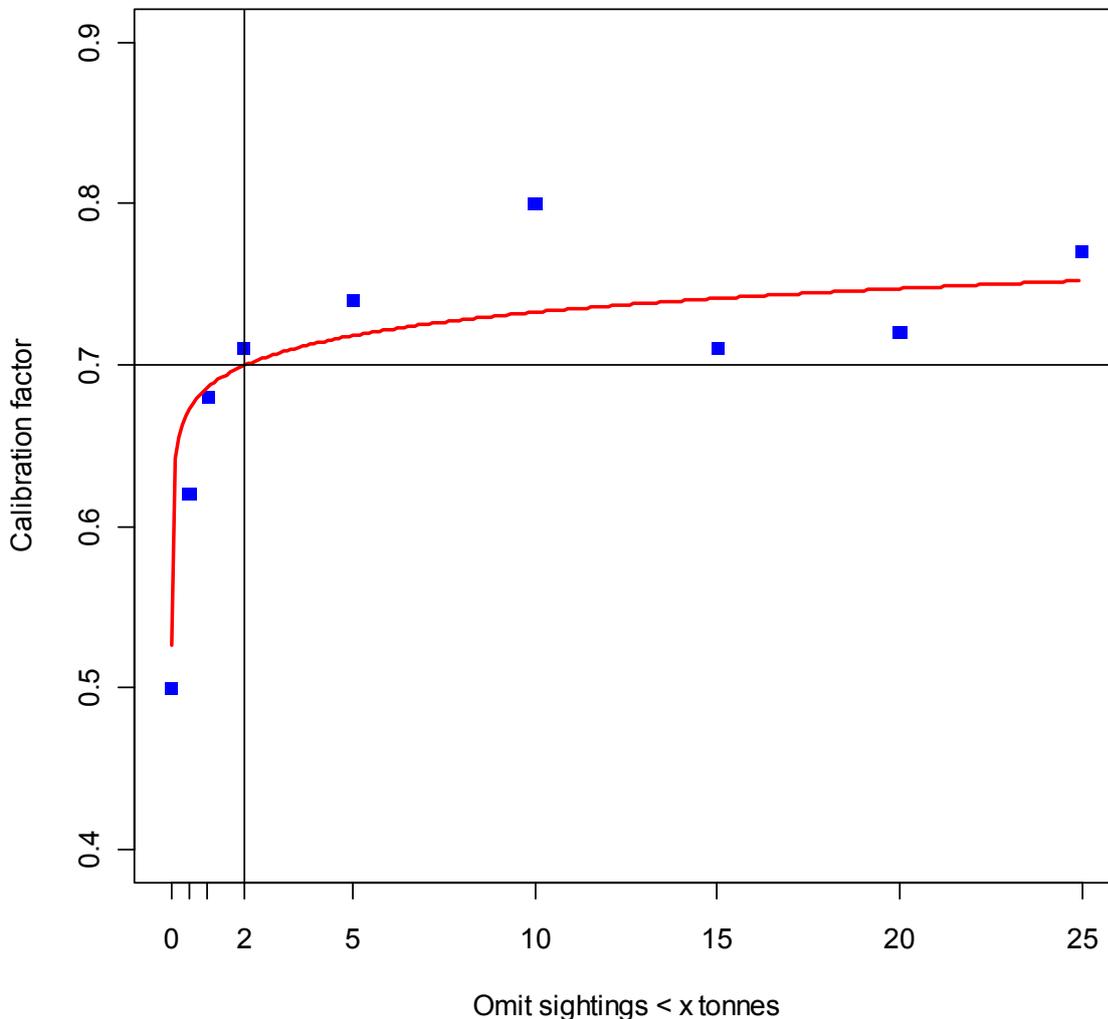
Note that using 0.7 instead of 0.5 as the calibration factor is a more robust approach because even if one observer only sees 50% as many really small schools as two observers, but we assume he sees 70%, this equates to very little difference in overall "missed" biomass. However, on the contrary, if we assume one observer sees only 50% as many big schools as

two observers, when in fact he sees 70% as many, we will be overestimating the amount of missed biomass by a lot.

Once the calibration factor has been estimated, it is used to calculate a relative sighting ability (i.e., an "observer effect") estimate for solo observers. Recall that the "observer effect" estimates for the SpM model are calculated based on a pair-wise observer analysis to estimate the relative sighting abilities of all observer *pairs* that have been involved in past surveys (see Appendix A). In order to estimate a relative sighting ability for a solo observer, we took the average of the relative sighting ability estimates from when he flew as part of a pair, and multiplied it by the estimated calibration factor. For example, one of the observers who flew as a solo observer in the 2010 and 2011 surveys has flown as part of two different observer pairs in past surveys, with relative sighting ability estimates of 0.90 and 0.92. If we take the average of these two relative sighting ability estimates and multiply it by the calibration factor of 0.7, this gives a relative sighting ability estimate for this observer when flying solo of 0.64 Thus, we now have "observer effect" estimates for all observer combinations, so we can proceed with fitting the SpM model in the usual way.

Once the models were fitted, the 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 emphasise 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, except we now include uncertainty in the calibration factor estimate, details of which can be found in Appendix B. Recall from above that there is still uncertainty in the observer effect estimates (now being input to the SpM model as an offset) that are not currently being accounted for.

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

(Model results and diagnostics for the BpS and SpM models are provided in Appendix C.)

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 given in Table 3.

The 2011 point shows a substantial increase from 2010 and is similar to the 1993 estimate, which was the highest estimate of all survey years. The confidence interval on the 2011 estimate is quite wide, but taking this into account, it is still significantly higher than other estimates in the 2000s. (We should recall from the Methods section that all of the confidence

intervals are being slightly underestimated because they do not account for uncertainty in the observer effect estimates.)

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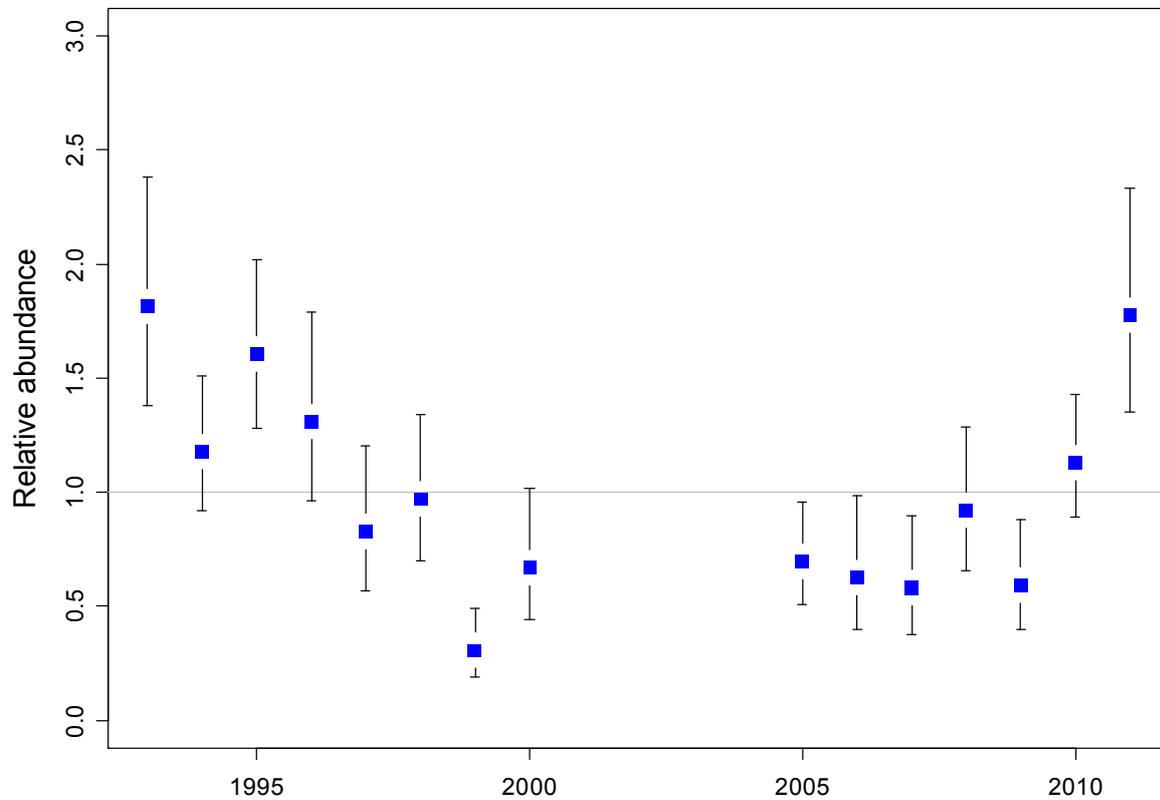

Table 3. Results from the aerial survey analysis.

Year	Index	SE	CV	CI.05	CI.95
1993	1.81	0.30	0.17	1.38	2.38
1994	1.18	0.18	0.15	0.92	1.51
1995	1.61	0.22	0.14	1.28	2.02
1996	1.31	0.25	0.19	0.96	1.79
1997	0.83	0.19	0.23	0.57	1.20
1998	0.97	0.19	0.20	0.70	1.34
1999	0.31	0.09	0.29	0.19	0.49
2000	0.67	0.17	0.26	0.44	1.02
2005	0.70	0.13	0.19	0.51	0.95
2006	0.62	0.17	0.28	0.40	0.98
2007	0.58	0.15	0.26	0.38	0.90
2008	0.92	0.19	0.20	0.66	1.29
2009	0.59	0.14	0.24	0.40	0.88
2010	1.13	0.16	0.14	0.89	1.43
2011	1.78	0.29	0.17	1.35	2.33

Index = relative abundance point estimates; SE= standard error; CV = coefficient of variation; CI.05 and CI.95 = lower and upper range of 90% confidence interval.

Summary

The estimate of relative juvenile abundance from the 2011 scientific aerial survey shows a substantial increase from 2010 and is similar to the 1993 estimate, which was the highest estimate of all survey years. The confidence interval on the 2011 estimate is quite wide, but taking this into account, it is still significantly higher than other estimates in the 2000s.

The 2011 survey was the first in which all flights had only one spotter (i.e., no flights had a spotter-pilot). Calibration experiments conducted in 2008 and 2009 showed that a plane with only one spotter makes fewer sightings than a plane with two spotters. We call the proportion of sightings made by one spotter compared to two spotters the "calibration factor". A method of estimating the calibration factor and accounting for it in our analysis of the survey data was developed in Eveson et al. (2009, 2010). A very high proportion of large schools this year led us to reconsider the calibration factor reported in Eveson et al. (2010), because a plane with one spotter is less likely to miss very large schools. Re-analysing the data from the calibration experiments leaving out very small sightings (<2 tonnes) led to a revised calibration factor estimate of 0.7 instead of 0.5. Methods to incorporate uncertainty in the calibration factor estimate were also developed and applied this year, so the CVs for the relative abundance estimates include this additional uncertainty.

This year's data included a high proportion of schools comprised of small (<8 kg, estimated to be 1-year-old) fish, which was unusual compared to past survey years. In the CCSBT OM and MP, the aerial survey index is assumed to provide a relative time series of age 2-4 abundance in the Great Australian Bight. Thus, for consistency with the OM and MP as well as general consistency in interpretation of the index across years, schools of 1-year-old fish (in all years) were omitted from the analysis.

The only other change to the analysis compared to last year is that the observer effect estimates for the sightings (SpM) model are now being included as an offset (i.e., as known) rather than as a linear covariate. The reason for this is discussed in Appendix A. As a result, we need to account for uncertainty in the observer effect estimates through other methods. Such methods have been developed but we were not successful in implementing them in time for this report. Thus, the CVs for the relative abundance indices reported here do not yet include uncertainty in the observer effects for the SpM model (i.e., they are slightly too small).

References

Anonymous. 2008. Report of the Thirteenth Meeting of the Scientific Committee, Commission for the Conservation of Southern Bluefin Tuna, 5-12 September 2008, Rotorua, New Zealand.

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

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

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

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

Eveson, P. 2011. Updated growth estimates for the 1990s and 2000s, and new age-length cut-points for the operating model and management procedures. CCSBT-ESC/1107/09.

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., Farley, J., and Bravington, M. 2009. The aerial survey index of abundance: updated analysis methods and results. CCSBT-ESC/0909/12.

Eveson, P., Farley, J., and Bravington, M. 2010. The aerial survey index of abundance: updated analysis methods and results for the 2009/10 fishing season. CCSBT-ESC/1009/14.

Farley, J., Eveson, P. and Clear, N. 2011. An update on Australian otolith collection activities, direct ageing and length at age in the Australian surface fishery. CCSBT-ESC/1107/17.

Robins, J.P. 1963. Synopsis of biological data on bluefin tuna, *Thunnus thynnus maccoyii* (Castlenau) 1872. FAO Fisheries Report 6: 562:587.

Acknowledgements

Appendix A – Methods of analysis

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

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

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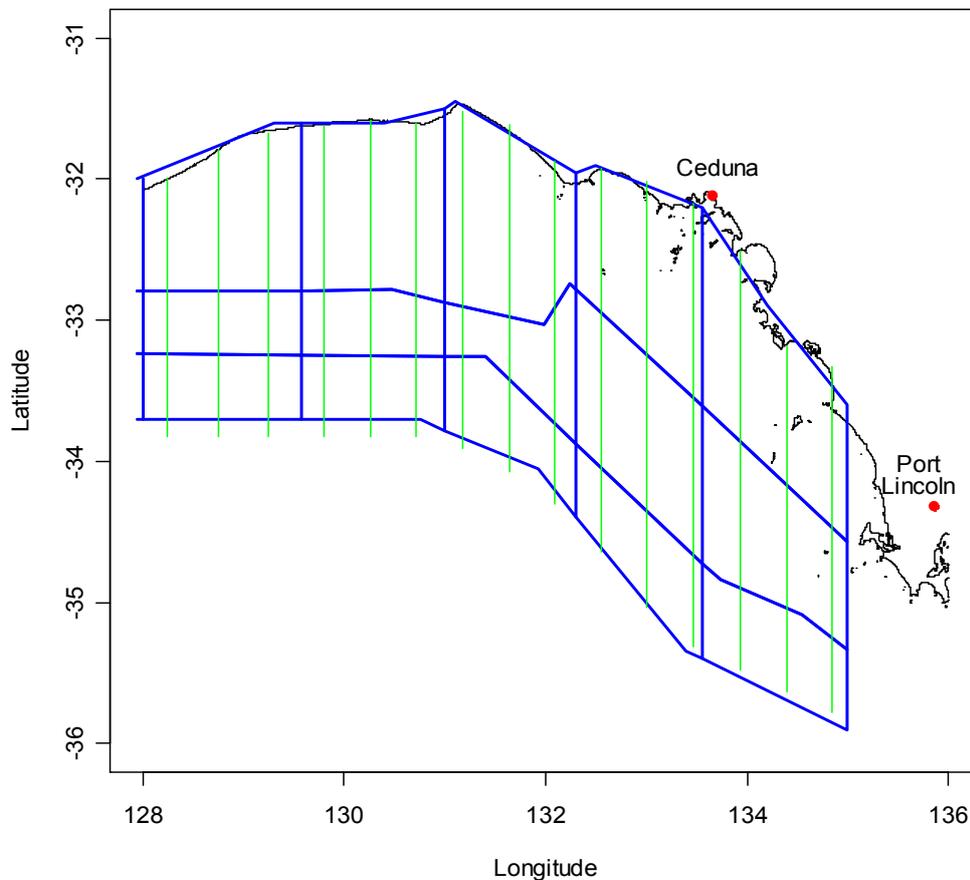
Biomass per sighting (BpS) model

For the BpS model, we first estimated relative differences between observers in their estimates of patch size (using the same methods as described in Bravington 2003). As in Bravington (2003), we found good consistency between observers. In particular, patch size estimates made by different observers tended to be within about 5% of each other, except for one observer, say X, who tended to underestimate patch sizes relative to other observers by about 20%. The patch size estimates were corrected using the estimated observer differences (e.g. patch size estimates made by observer X were scaled up by 20%). Because the observer differences 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 and area).

The 2-way and 3-way interaction terms between Year, Month and Area were fit as random effects, whereas the 1-way effects were fit as fixed effects. Many of the 2-way and 3-way strata have very few (sometimes no) observations, which causes instabilities in the model fits when treated as fixed effects. One main advantage of using random effects is that when little or no data exist for a given level of a term (say for a particular area and month combination of the Area:Month term), we still have information about it because we are assuming it comes from a normal distribution with a certain mean and variance (estimated within the model). Having decided on the overall structure, we then decided what environmental variables to include in the model. Based on exploratory plots and model fits, we determined the two environmental covariates that had a significant effect on the biomass per sighting were wind speed and, especially, SST.[1] 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 covariates (note that $\mathbf{E}$ 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.

---

[1] Note that the selection of environmental covariates for the BpS and SpM models was undertaken as part of the 2006 analysis. We re-investigated last year (see Eveson et al. 2010) and found some suggestion that wind speed was not necessary in the BpS model and that sea shadow should be included in the SpM model, but we have chosen to keep the covariates the same for now for consistency purposes, and because it made very little difference to the relative abundance estimates.

This analysis gives estimates of the relative sighting efficiencies for the 18 different observer pairs that have flown at some point in the surveys. The observer pairs ranged in their estimated sighting efficiencies from 72% to 97% compared to the pair with the best rate.

This year, we include the (logged) estimates of relative observer pair efficiencies as an offset when fitting the SpM model (i.e., as a predictor variable with a known, rather than estimated, coefficient). In past, we have included the logged estimates as a covariate in the SpM model rather than as an offset, with the coefficient (i.e., "slope") to be estimated, with the notion that if the relative efficiencies from the pairwise analysis are correct, the slope estimate should be close to one. By doing so, we were attempting to capture some of the uncertainty in the estimates. As discussed in last year's report (Eveson et al. 2010), the coefficient of the log observer effect term was actually coming out close to negative one, and although it was coming out as insignificant, this was still a worry because it suggests that the sightings rate actually declines as our estimates of relative sighting ability increases. Complex interactions between the space-time strata and the observer effects seemed to be the cause of this. As such, we have chosen to include the log relative efficiencies as an offset, and account for their uncertainty through other methods. Such methods have been developed but we were not successful in implementing them in time for this report. Thus, the standard errors and CVs for the relative abundance indices reported in Table 3 do not include uncertainty in the observer effects for the SpM model (which means they are slightly too small). We aim to get the methods of including observer uncertainty implemented correctly in the coming year.

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 terms for year, month and area, as well as all possible interactions between them, in the SpM model, and we fitted the 2-way and 3-way interaction terms as random effects (see BpS model section). We determined what environmental variables to include in the model based on exploratory plots and model fits (see footnote 1). The final model fitted was:

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

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. Note, as discussed above, that the log observer effect estimates are now being included as an offset rather than a linear covariate.

---

[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 average conditions calculated from all the data.

Appendix B – CV calculations

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

Let $\hat{B}_{ijk}$ be the predicted value of BpS in year i, month j and area k under standardized environmental/observer conditions (see footnote 3), and $\hat{\sigma}\left(\hat{B}_{ijk}\right)$ be its estimated standard error. Similarly, let $\hat{S}_{ijk}$ be the predicted value of SpM in year i, month j and area k under the same environmental/observer conditions, and $\hat{\sigma}\left(\hat{S}_{ijk}\right)$ be its estimated standard error. Then,

$$\hat{A}_{ijk} = \hat{S}_{ijk}\hat{B}_{ijk}$$

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

Since $\hat{B}_{ijk}$ and $\hat{S}_{ijk}$ are independent, the variance of $\hat{A}_{ijk}$ is given by

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

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

---

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

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:

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

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 $\left(\dfrac{\partial \hat{A}_i}{\partial \theta}\right)$ 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].

Now, to account for the additional variance due to uncertainty in the calibration factor, we use a similar approach as above to account for additional variance due to uncertainty in the environmental coefficients. Namely, from the GLM used to estimate the calibration factor, which we will call α, we get an estimate of its variance, which we will call V_α. Then, the variance in the abundance estimates due to uncertainty in α can be estimated by

$$V_\alpha\left(\hat{A}_i\right) = \left(\frac{\partial \hat{A}_i}{\partial \alpha}\right)' V_\alpha \left(\frac{\partial \hat{A}_i}{\partial \alpha}\right)$$

---

[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.

where $\left(\dfrac{\partial \hat{A}_i}{\partial \alpha} \right)$ is the derivative of $\hat{A}_i$ with respect to α (in essence, it is the amount that the abundance estimate $\hat{A}_i$ changes when the calibration factor is tweaked slightly). Thus, we revise our estimate of $V\left(\hat{A}_i \right)$ by adding on to it $V_\alpha \left(\widehat{A}_i \right)$.

So we have variance estimates for the abundance estimates, but we also want to calculate the variance for the mean-standardized estimates (referred to as the relative abundance indices), calculated as:

$$\hat{I}_i = \frac{\hat{A}_i}{\dfrac{1}{n} \sum_{i=1}^{n} \hat{A}_i}$$

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

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

Appendix C: Results and diagnostics

Biomass per sighting (BpS) model

Extract from the output produced by the software used to fit the model (the gam function in the R statistical package mgcv):

```
Family: Gamma
Link function: log

Formula:
Biomass ~ factor(Year) + factor(Month) + factor(Area) + SST + WindSpeed +
Y.M + Y.A + M.A + Y.M.A - 1

Parametric Terms:
     Covariate  Estimate       SE   t-value   p-value
           SST     0.128    0.041     3.092     0.002
     WindSpeed    -0.043    0.026    -1.699     0.090

R-sq.(adj) =  0.0828   Deviance explained = 38.6%
GCV score = 2.0143  Scale est. = 1.7649    n = 1637
```

The results suggest that the size of a sighting tends to increase as SST increases and decrease as wind speed increases, but that relationship with SST has much greater statistical significance. These results are roughly supported by plots of observed biomass per sighting (on a log scale) versus the environmental covariates. Nevertheless, we prefer to include wind speed in the model for consistency with previous years, and because the relative abundance estimates were very similar when wind speed was omitted.

Figure C2 shows some standard diagnostic plots for generalized linear models, and Figure C3 shows the residuals plotted against a number of factors. These plots do not suggest major problems with the model fit. Ideally there should be no trend in the plots of the square root of the absolute residuals against the fitted values (i.e., lower half of Fig. C2, with left-hand side being on the link scale and the right-hand side being on the response scale); although there is a small kink revealed by a smooth through the data (red line), there is not a consistent increasing or decreasing trend.

Figure C1. Plots of observed biomass per sighting, on a log scale, versus the covariates included in the model; shown is the mean +/- 2 standard deviations.
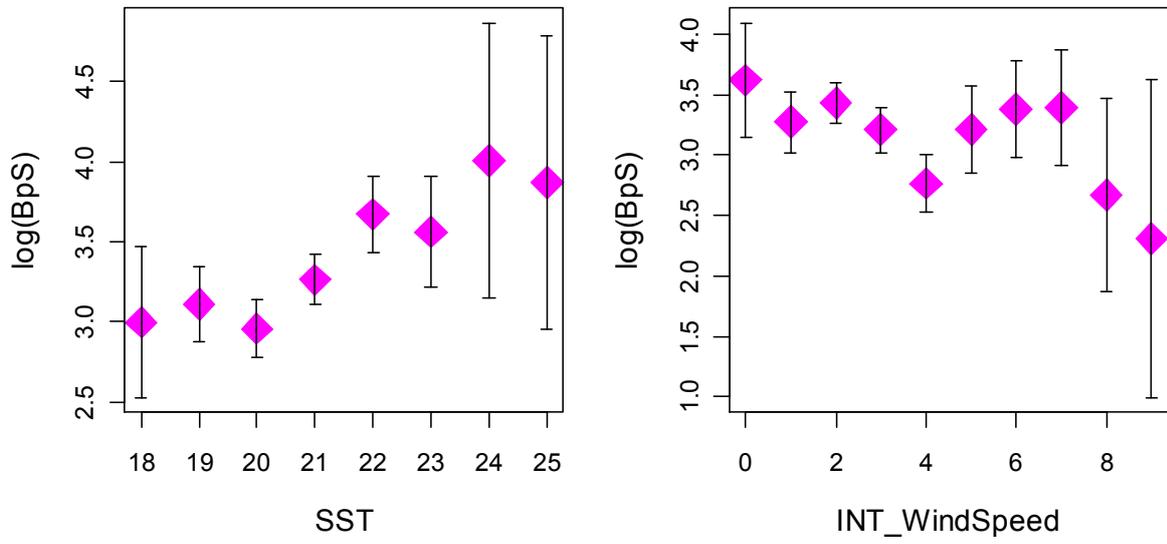
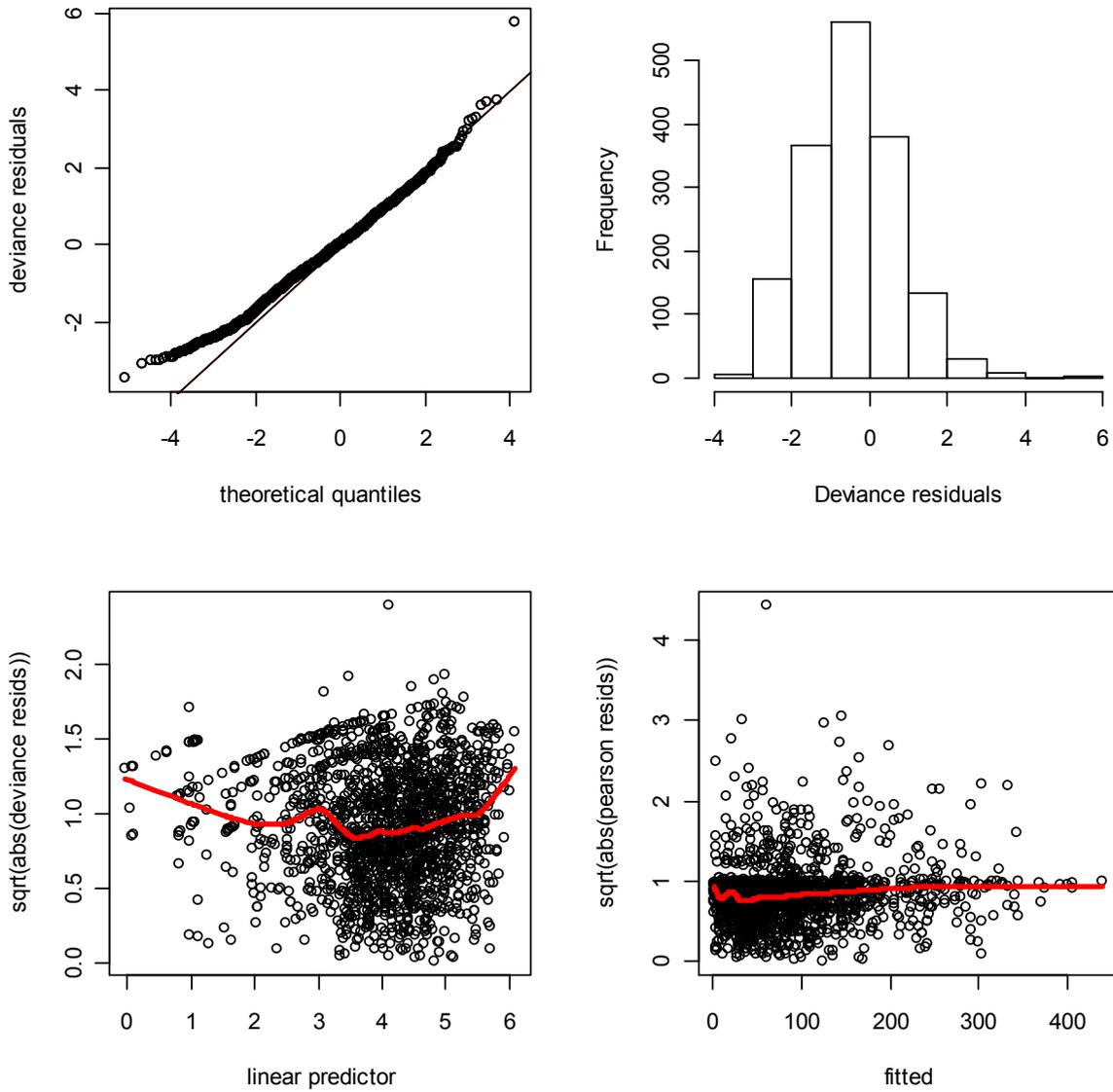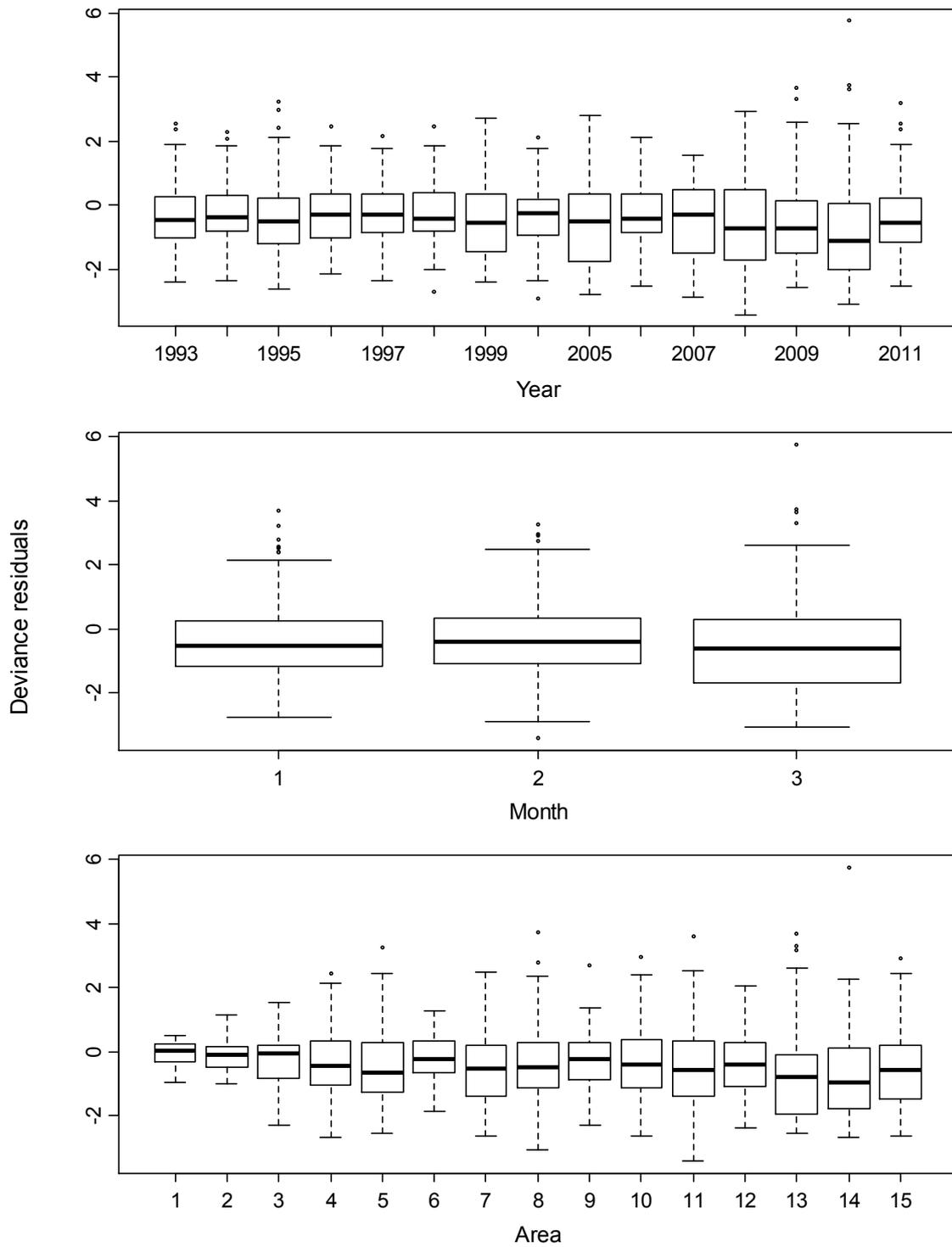Figure C2. Standard diagnostic plots for biomass per sighting (BpS) model.

Figure C3. Boxplots of deviance residuals by year, month and area for biomass per sighting (BpS) model.

Sightings per mile (SpM) model

Extract from the output produced by the software used to fit the model (the gam function in the R statistical package mgcv):

```
Family: quasipoisson
Link function: log

Formula:
N_sightings ~ offset(log(as.numeric(Distance))) + factor(Year) +
factor(Month) + factor(Area) + Y.M + Y.A + M.A + Y.M.A +
log(ObserverEffect) + AvgWindSpeed + AvgSST + AvgSwell + AvgHaze +
factor(MoonPhase) - 1
```

Parametric Terms:

Covariate	Estimate	SE	t-value	p-value
AvgWindSpeed	-0.262	0.023	-11.209	0.000
AvgSST	0.211	0.036	5.913	0.000
AvgSwell	-0.189	0.055	-3.452	0.001
AvgHaze	-0.122	0.049	-2.471	0.014
factor(MoonPhase)2	-0.127	0.102	-1.247	0.213
factor(MoonPhase)3	-0.008	0.126	-0.064	0.949
factor(MoonPhase)4	0.186	0.086	2.148	0.032

```
R-sq.(adj) =  0.475   Deviance explained = 65.7%
GCV score = 1.5291  Scale est. = 1.2395    n = 1620
```

The results suggest that there is a tendency for the rate of sightings to increase as SST increases, and to decline as wind speed, haze and swell increase (all highly significant). The relationship with moon phase is more complex, with the sightings rate being greater when the moon phase is 1 (fraction of moon illuminated is 0-25%) or 4 (fraction of moon illumination is 75-100%), but this relationship is not as significant.

Figure C5 shows some standard diagnostic plots for generalized linear models, and Figure C6 shows the residuals plotted against a number of factors. The Q-Q plot has a funny kink, but otherwise there are no indications of serious problems with the model fit. The plots of the square root of the absolute residuals against the fitted values (i.e., lower half of Fig. C5, with left-hand side being on the link scale and the right-hand side being on the response scale) look a bit odd, but this is expected because we are modelling count data. A smooth line through these data is reasonably flat, as desired, except for where it follows the residuals for the zero response values (i.e., where the observed number of sightings was zero).

Figure C4. Plots of observed sightings per mile, on a log scale, versus the covariates included in the model; shown is the mean +/- 2 standard deviations.
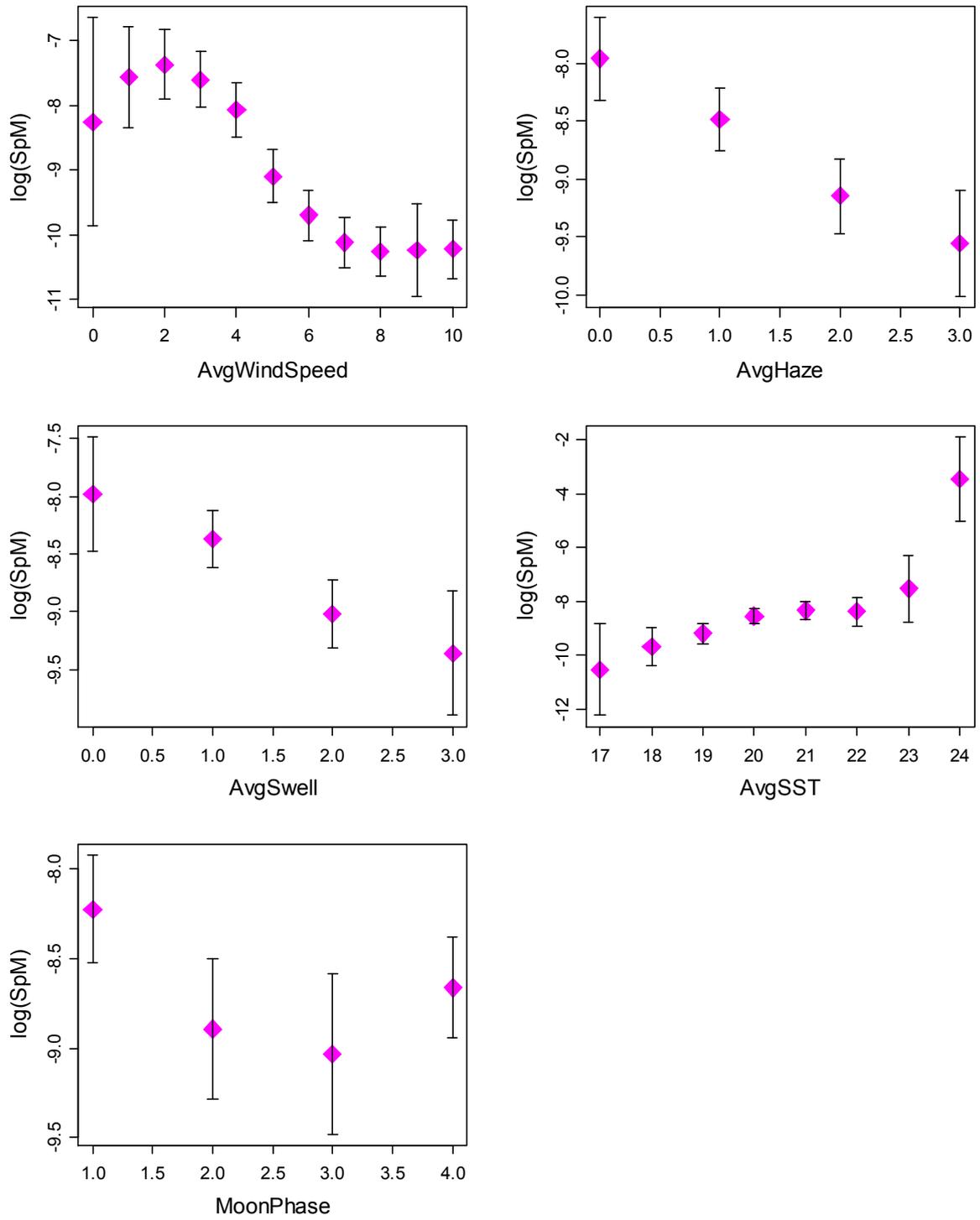
Figure C5. Standard diagnostics plots for sightings per mile (SpM) model.
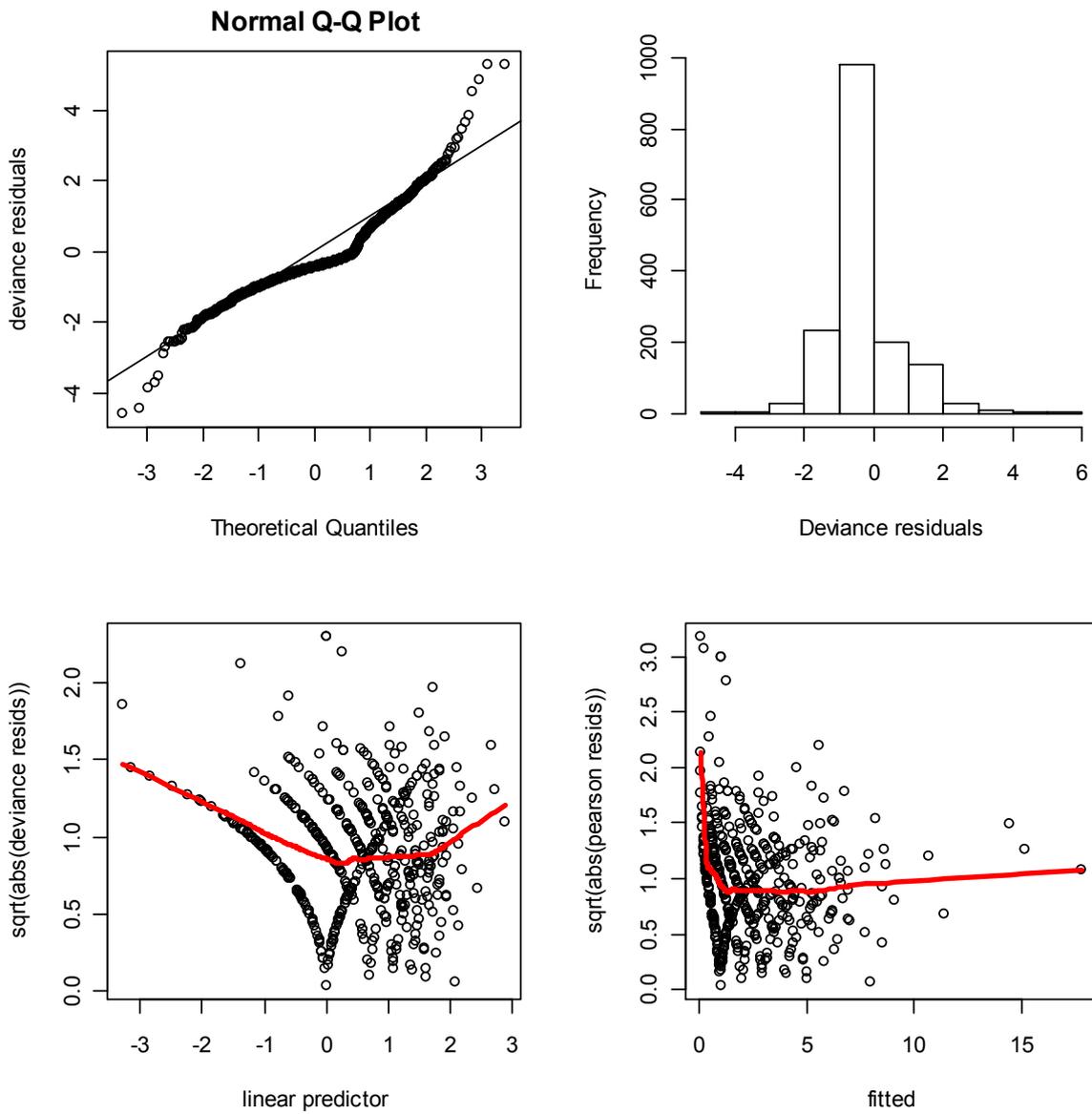
Figure C6. Boxplots of deviance residuals by year, month and area for sightings per mile (SpM) model.