CCSBT-ESC/1009/14

The aerial survey index of abundance: updated analysis methods and results for the 2009/10 fishing season

Paige Eveson
Jessica Farley
Mark Bravington

Prepared for the CCSBT 15th Meeting of the Extended Scientific Committee
4-10 September 2010, Taipei, Taiwan

Table of Contents

Abstract

The estimate of relative juvenile abundance from the 2010 scientific aerial survey is the highest estimate since 1996; however, it remains lower than the average level in the period 1993-1996.

The current year (2010) was the first year that planes with only one observer flew as part of the survey. Calibration experiments conducted in 2008 and 2009 showed that a plane with only one observer makes fewer sightings than a plane with two observers. Thus, in order to include the data collected in the 2010 survey from flights with only one observer, the analysis needs to take into account the fact that the number of sightings is expected to be fewer with only one observer than with two observers. A method for doing so was proposed in last year's CCSBT report (Eveson et al. 2009) and is described in greater detail in this report.

In order to make our results directly comparable to results provided in past, the data were first analysed using only data from flights with two observers. These results were provided to the CCSBT as part of the 2010 data exchange. The analysis was then redone including data from flights with one observer using the methods described for dealing with such data. The results from the analysis using only two-observer data and from the analysis using both one and two-observer data are very similar. If the methods for including one-observer data are accepted by the CCSBT, results based on all data will be provided to the CCSBT in future.

The environmental conditions in 2010 were highly favourable, with lower average wind speed and higher average sea surface temperature (SST) than experienced in any of the past surveys. New data, especially data from extreme conditions, can significantly affect the estimated model coefficients and, consequently, the relative abundance estimates; thus, we explored the environmental covariates being included in the models and their influence on the abundance indices. Results are presented which highlight the complexity of model selection and the importance of regularly exploring the models as new data become available.

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 might have only one observer in a plane (as was the case for one of the two planes flying in the 2010 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 2010 season. This was the first year that a plane with only one observer flew as part of the survey, in addition to a plane with two observers. The current methods for analysing the data are described, and results are presented from applying these methods to data from all survey years and flights with two observers (i.e., omitting the data collected in 2010 from flights with one observer in order to make the results directly comparable to those from previous years). A method for including the data from one-observer flights is proposed and results from applying this method are presented. Finally, explorations regarding the environmental covariates in the models are presented, motivated by the fact that the environmental conditions during the 2010 survey were better than in any past survey year.

Field procedures

The 2010 scientific aerial survey was conducted in the GAB between 1 January and 31 March 2010. As for previous surveys (e.g. Eveson et al. 2009), the plane type used was a Rockwell Aero Commander 500S. Two planes were chartered in 2010, one for the full three months (plane 1) and a second for January and February only (plane 2). Three observers were employed; plane 1 had two observers (an observer-pilot and a dedicated observer), while plane 2 had a single dedicated observer (and a non-observing pilot). The two dedicated observers were swapped between planes in January and February, but the pilots (one also an observer) remained in their respective planes. The same observers employed for the 2007 to 2009 surveys were used throughout the 2010 survey.

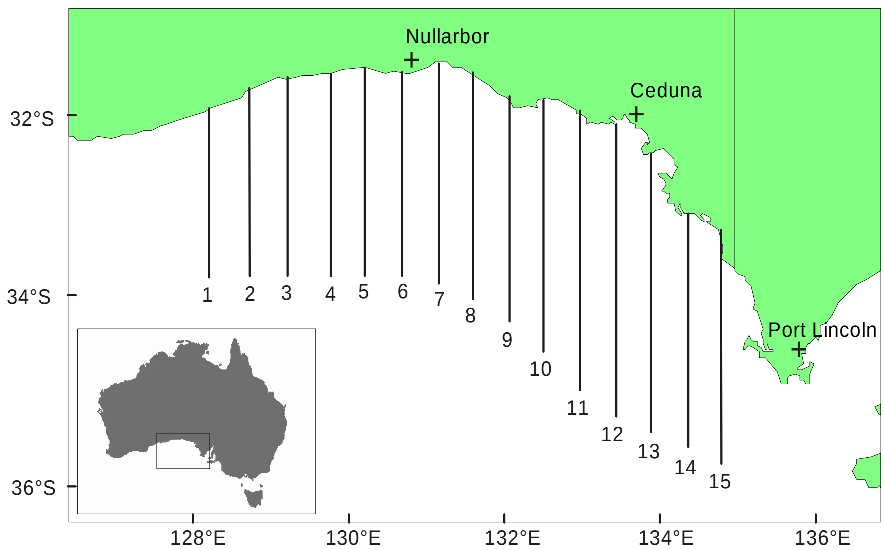
The survey followed the protocols established for the 2000 survey (Cowling 2000) and used in all subsequent surveys, with respect to the area searched, plane height and speed, minimum environmental conditions, time of day the survey lines were flown, and data recording protocols. Fifteen north-south 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 two planes used slightly different methods to search for SBT. For plane 1 (i.e. with two observers), each observer searched the sea surface from straight ahead through to 90° to their side of the plane (abeam of the plane) for surface patches (schools) of SBT. An observer would occasionally search both sides of the plane, if the other observer was temporally unable to observe. For plane 2, the single observer searched for patches of SBT from his side of the plane through 180° to the other side of the plane.

When two 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 1 with two ovservers.

The 2010 field operation was very successful, largely due to the prevailing whether conditions in January and March and the availability of two planes for two months of the survey. Nearly 7 replicates of the GAB were completed, compared to between 3 and 5 for the 5 years prior. The total flying time (transit and transect time) for the 2010 survey was 213.6 hours.

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 2010 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).

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. 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. For 2010, the summary statistics are presented using data from all flights, as well as broken down by flights with two observers and flights with one observer.

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

Survey year	Total distance searched (nm)	Number SBT sightings	Total biomass	Average patches per sighting	Max patches per sighting	Average biomass per patch	Median biomass per patch	Max biomass per patch
1993	7603	130	12235	3.9	76	24.4	18.8	203
1994	15180	174	15055	3.3	23	26.5	21.6	246
1995	14573	179	22046	3.6	38	34.7	27.9	225
1996	12284	116	16638	4.1	46	34.9	27.8	147
1997	8813	117	9965	3.0	18	28.1	22.7	202
1998	8550	109	10392	2.3	21	40.9	20.3	964
1999	7555	56	3033	2.4	21	23.0	16.6	121
2000	6775	77	4838	2.6	17	24.1	20.0	100
2005	5968	80	6096	2.4	17	31.8	25.2	194
2006	5150	44	4037	2.0	8	46.9	31.6	268
2007	4872	42	3510	2.6	11	32.8	24.6	121
2008	7462	122	8054	3.5	24	19.0	13.4	314
2009	8101	154	8651	2.6	22	22.0	14.5	170
2010	10559	208	21921	3.9	41	26.8	18.0	634
2010, 2-obs	6736	176	19073	4.3	41	25.5	17.0	634
2010, 1-obs	3823	32	2848	2.2	7	41.3	32.8	234

[1] The biomass statistics differ slightly from those reported in Table 1 of the 2009 CCSBT report (Eveson et al. 2009) because the patch size estimates used in calculating these statistics have been corrected for differences between observers (see Appendix A). Observer differences are re-estimated each year using all available data and thus the corrected patch size estimates can change slightly.

The total distance searched in 2010, if all flights (with one or two observers) are considered, was the highest since 1996. If only two-observer flights are considered, the distance searched was close to the average since 1996. The sightings rate (number of sightings per 100 nm) for two-observer flights was higher than in previous survey years, as was the amount of biomass per nm (Figure 2). This was due in part to very good environmental conditions for spotting fish (see next section). The sightings rate for flights with only one-observer was much lower; this was expected from the calibration experiments performed in 2008 and 2009 (see last year's report, Eveson et al. 2009). The overall mean and median biomass per patch were slightly smaller in 2010 (based on all data) than the average over all years, but larger than the previous two years (Figure 3; Table 1).

Similar to last year, sightings in 2010 were concentrated in the eastern half of the survey area, with the greatest concentration along the shelf-break (Figure 4). There appears to have been a general eastward shift in the distribution of SBT sightings over the years.

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. For 2010, only the data from flights with two observers are included. ***Note that these plots are based on raw data, which has not been corrected for environmental factors or observer effects.***

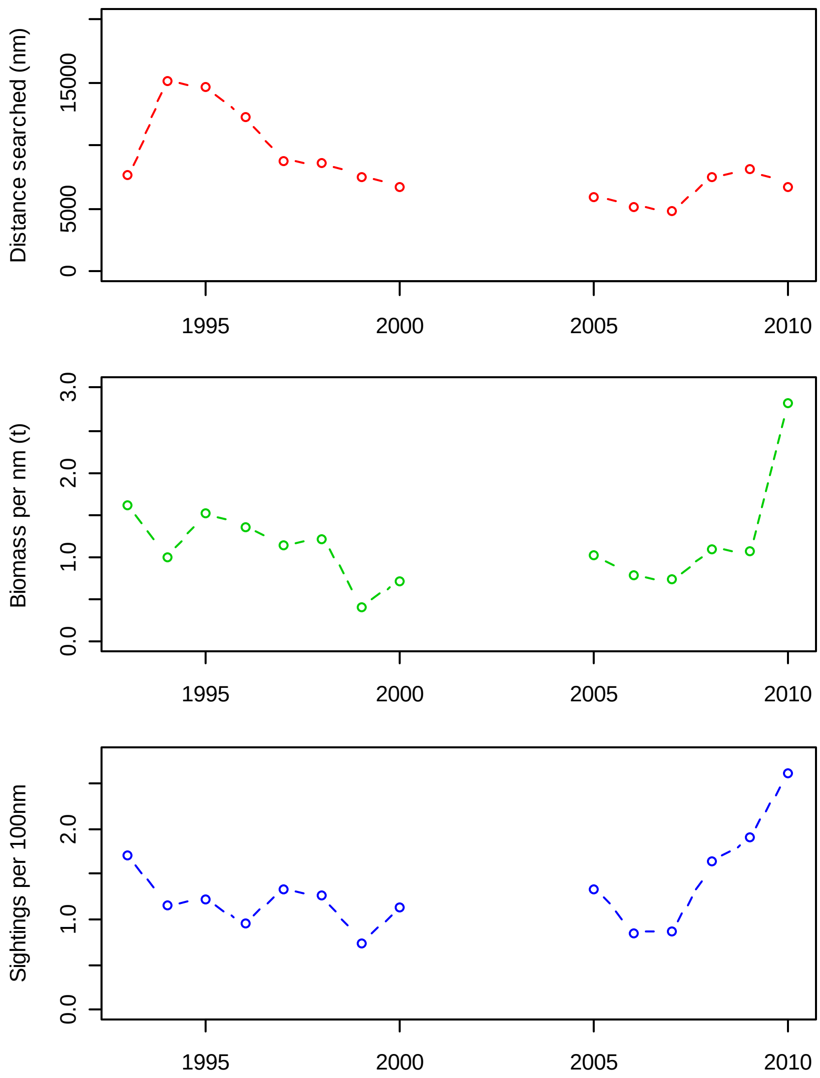


Figure 3. Frequency of SBT patch sizes (in tonnes) by survey year.

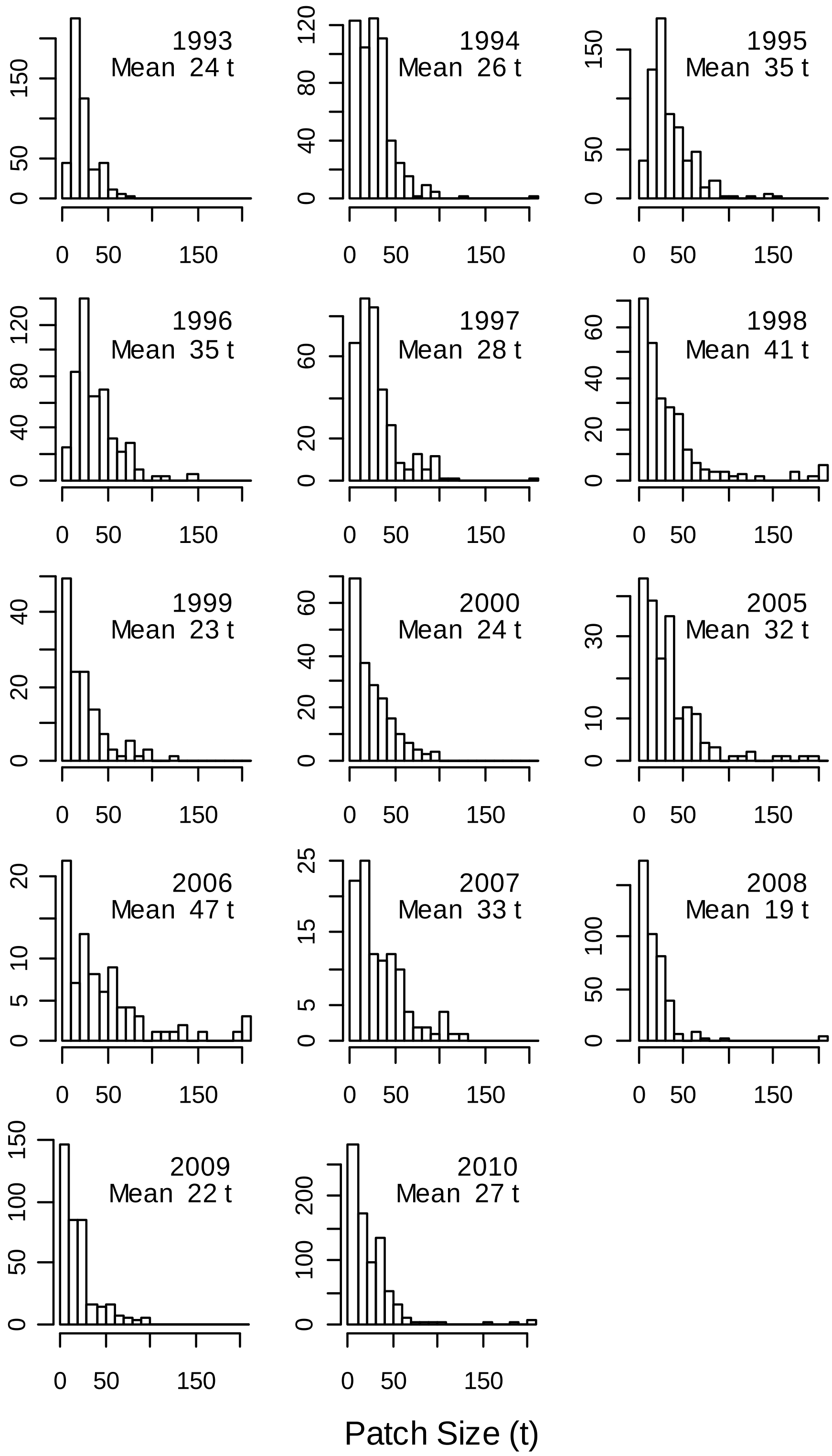
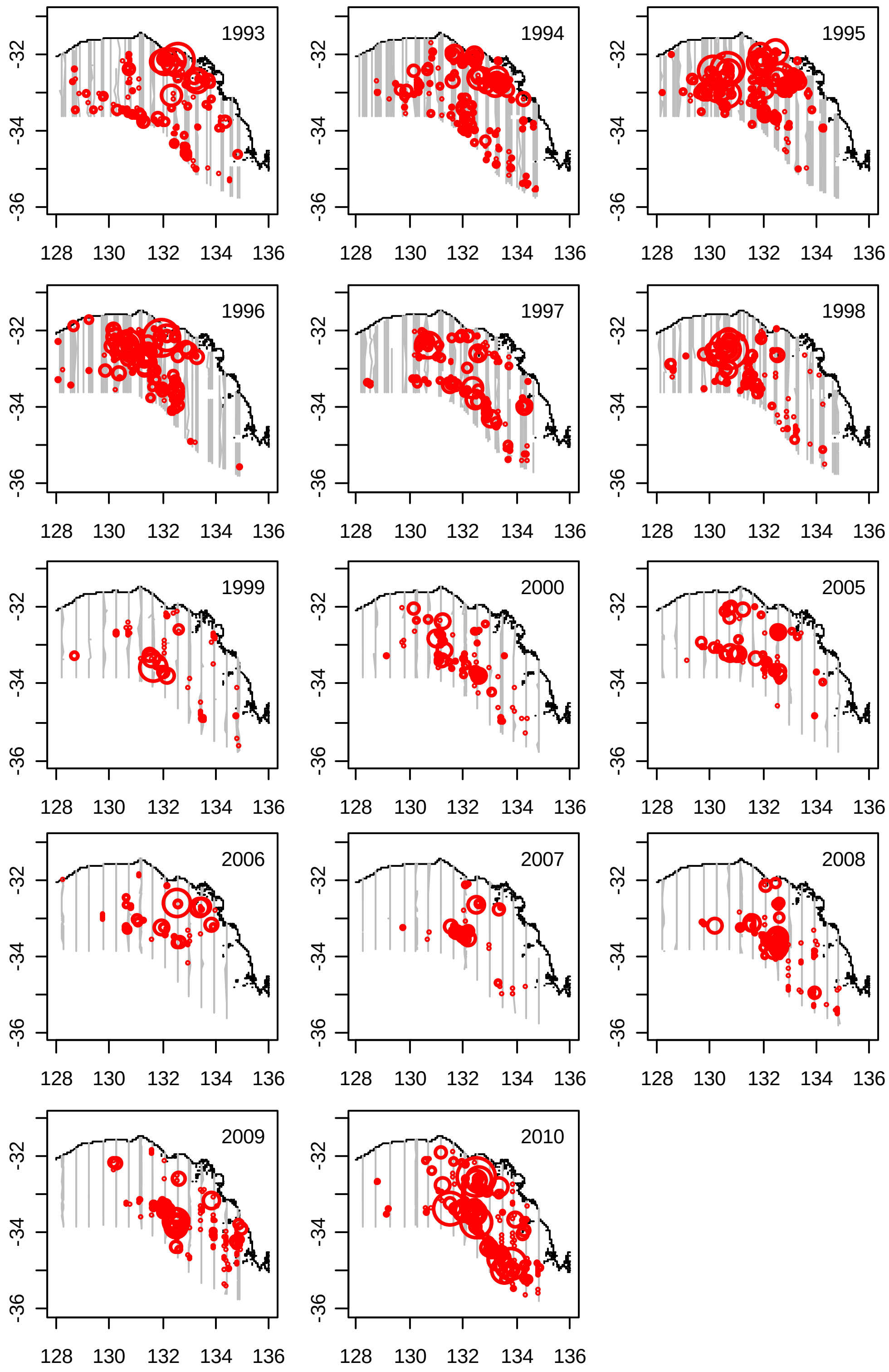


Figure 4. Distribution of SBT sightings made during each aerial survey year. Red circles show the locations of SBT sightings, where the size of the circle is proportional to the size of the sighting, and grey lines show the north/south transect lines that were searched.



Environmental variables

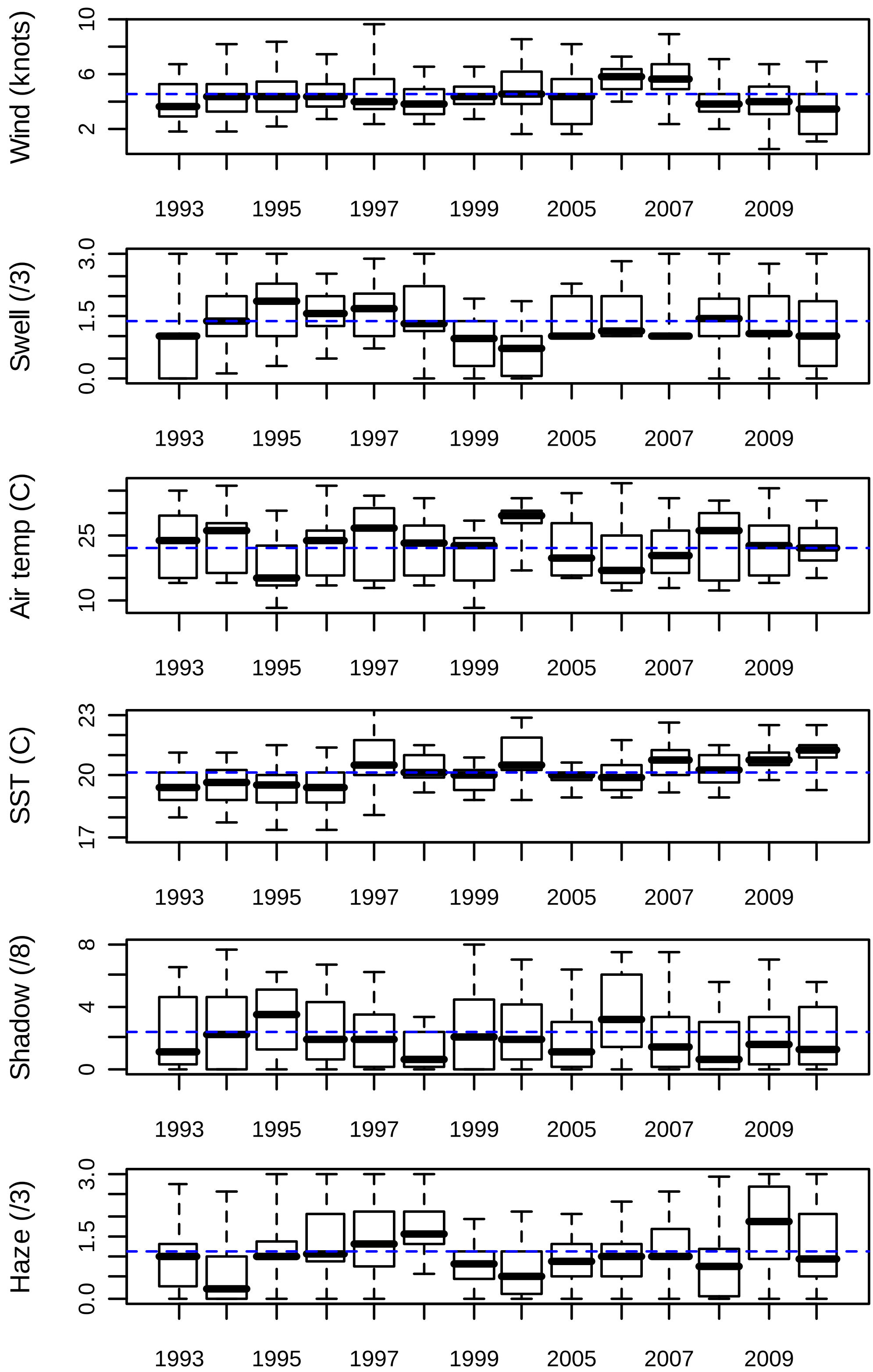
Table 2 and Figure 5 summarize the environmental conditions that were present during valid search effort in each survey year. All the environmental variables presented were recorded by the survey plane, with the exception of sea surface temperature (SST), which was extracted from the 3-day composite SST dataset produced by CSIRO Marine and Atmospheric Research's Remote Sensing Project (see Eveson et al. 2006 for more details).

The wind speed during the 2010 survey was lowest on average than all other survey years, and SST was also higher and less variable compared to other survey years (Table 2; Figure 5). Sighting rates tend to be higher with low wind speed and high SST, so overall the conditions were very favourable in 2010 for making sightings.

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

Survey year	Wind speed (knots)	Swell height (0-3)	Air temp (°C)	SST (°C)	Sea shadow (0-8)	Haze (0-3)
1993	3.9	0.8	24.4	19.6	1.8	0.9
1994	4.1	1.5	20.6	19.7	2.7	0.5
1995	4.4	1.7	18.7	19.6	2.7	1.1
1996	4.5	1.6	22.9	19.6	2.1	1.2
1997	4.1	1.7	25.3	21.1	1.6	1.3
1998	3.7	1.7	22.3	20.4	0.9	1.7
1999	4.1	0.9	22.0	19.9	2.9	0.7
2000	4.3	0.6	27.5	20.7	2.6	0.7
2005	4.7	1.5	21.7	20.1	1.6	0.8
2006	5.6	1.5	20.0	20.1	3.5	1.0
2007	5.8	1.3	21.6	20.8	2.0	1.3
2008	3.8	1.4	24.2	20.4	1.4	0.9
2009	3.8	1.4	22.2	21.0	2.1	1.7
2010	3.5	1.1	23.6	21.2	1.8	1.2

Figure 5. Boxplots summarizing the environmental conditions present during valid search effort for each aerial survey year. The thick horizontal band through a box indicates the median, the length of a box represents the inter-quartile range, and the vertical lines extend to the minimum and maximum values. The dashed blue line running across each plot shows the overall average across all survey years.



Methods of analysis

Analysis of data from flights with two observers

The methods of analysis used when only data from flights with two observers were included remained the same as last year. A brief description of the methods is repeated here, but readers are referred to Appendix A for full details.

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. Specifically, the models can be expressed as

BpS model: log**E**(Biomass) ~ Year*Month*Area + SST + WindSpeed

SpM model: log**E**(N_sightings) ~ offset(log(Distance)) + Year*Month*Area + log(ObsEffect) + SST + WindSpeed + Swell + Haze + 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).

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, the details of which are found in Appendix B. Briefly, we first obtained standard errors (SEs) for the predicted values of 'standardized SpM' and 'standardized average BpS' in each year/area/month stratum. These were used to calculate SEs for the stratum-specific abundance estimates, which were in turn used to calculate SEs for the annual abundance estimates. Lastly, we applied the delta method to determine SEs for the relative abundance indices. Note that CVs are given simply by dividing the SE of each index estimate by the estimate. We calculated confidence intervals for the indices based on the assumption that the logarithm of the indices follows a normal distribution, with standard errors approximated by the CVs of the untransformed indices.

Analysis of data from flights with one or two observers

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. Thus, in order to include the data collected in the 2010 survey from flights with only one observer, our analysis needs to take into account the fact that the number of sightings is expected to be fewer with only one observer than with two observers. A method for doing so was proposed in last year's report (Eveson et al. 2009), and was applied this year to the data from the 2010 survey.

Based on our analysis of the calibration experiment data conducted last year (Eveson et al. 2009), we expect that, on average, a plane with one observer will make about half (0.496) as many sightings as a plane with two observers. We refer to this factor as the "calibration factor". Although we found evidence that the calibration factor differs between observers, the data are too limited given its highly variable nature to draw any definitive conclusions. For instance, further investigation showed that this result was driven mainly by one flight where a large number of sightings were made by the survey plane and not the calibration plane. Furthermore, if an observer other than those that participated in the 2008 and 2009 calibration experiments were to fly as a solo observer in future surveys, then we would not have a calibration factor for him/her; thus, it is preferable not to use observer-specific calibration factors in the analysis unless it is considered essential.

Recall that prior to fitting the SpM model, we run a pair-wise observer analysis to estimate the relative sighting abilities of all observer pairs that have been involved in past and present surveys (see "Sightings per mile (SpM) model" section in 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 (in past and current surveys), and multiplied it by the estimated calibration factor. For example, one of the observers who flew as a solo observer in the 2010 survey has flown as part of two different observer pairs in past and present 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 0.496, this gives a relative sighting ability estimate for this observer when flying solo of 0.45. Thus, we now have relative sighting ability estimates (also referred to as "observer effect" estimates) for all observer combinations, so we can proceed with fitting the SpM model in the usual way.

The BpS model can be fitted the same way as in the previous section, but now including data from the flights with one observer. The only difference 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).

Once the BpS and SpM models have been fitted, the calculation of the relative abundance indices can proceed in the usual way. We calculated the CVs and confidence intervals using the same methods as in the previous section as well; however, we acknowledge that there is extra uncertainty in the observer effect estimates for solo observers that are not being accounted for in the SpM model. This is an issue for future work, but will only affect the CV of the 2010 estimate in the current analysis (i.e., the CV for the 2010 estimate should be slightly larger than reported here).

Results

Figure 6 shows the estimated time series of relative abundance indices with 90% confidence intervals when data from flights with two observers only are included in 2010. These results were provided to the CCSBT for use in the management procedures operating model, since they are directly comparable to results provided in past (i.e., they were obtained using the same models and do not require new methods for calibrating data from one-observer planes). The point estimates and CVs corresponding to Figure 6 are reproduced in Table 3. Although the 2010 point estimate remains lower than the average level in the period 1993-1996, it is the highest estimate since 1996. The confidence interval on the 2010 point estimate is notably smaller when data from one-observer flights are included. This is expected since the amount of distance searched and number of sightings increases significantly (see Table 1); recall, however, from the Methods section that this confidence interval is an underestimate because it does not account for the extra uncertainty in the observer effect estimates for solo observers.

Figure 7 compares the results obtained when data from flights with two observers only are included in 2010 (as in Figure 6) versus when data from all flights (one or two observers) are included. Even though 2010 is the only year with data from flights with one observer, the estimates can change for all years since the models are refit using all data, meaning the model coefficients, and hence the predicted values, change. However, as we would expect, estimates in past years remain very similar, and 2010 is the year with the largest change. The point estimate for 2010 is lower when data from flights with one observer are included, but the difference is not significant when confidence intervals are taken into account.

Figure 6. Time series of relative abundance estimates with 90% confidence intervals. Results were obtained using data from flights with two observers only.

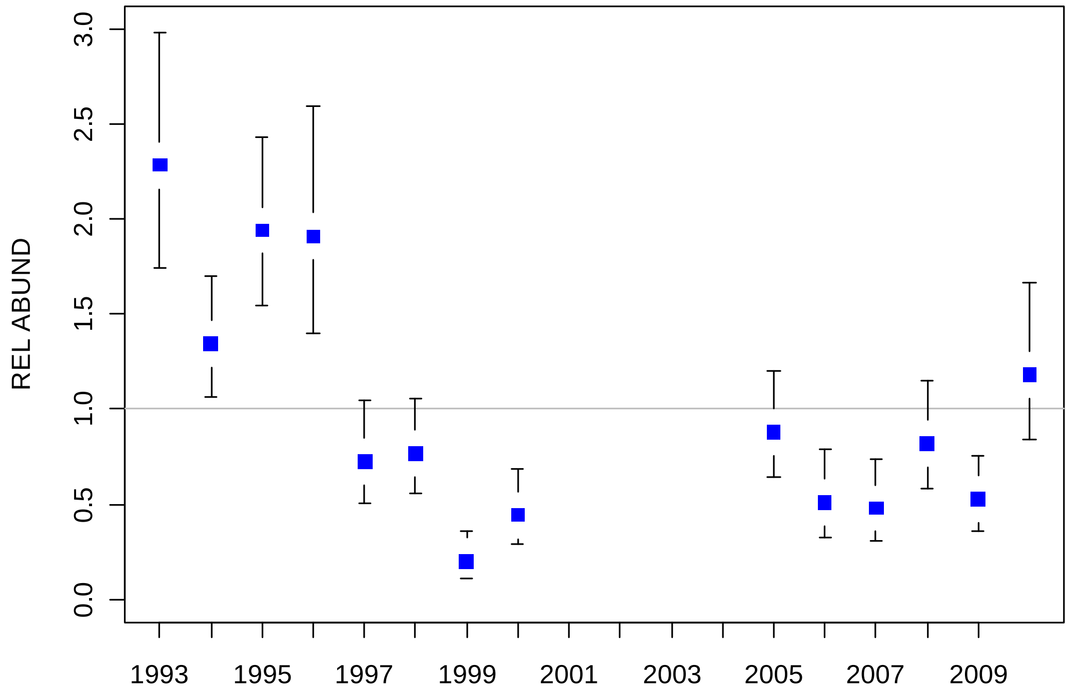
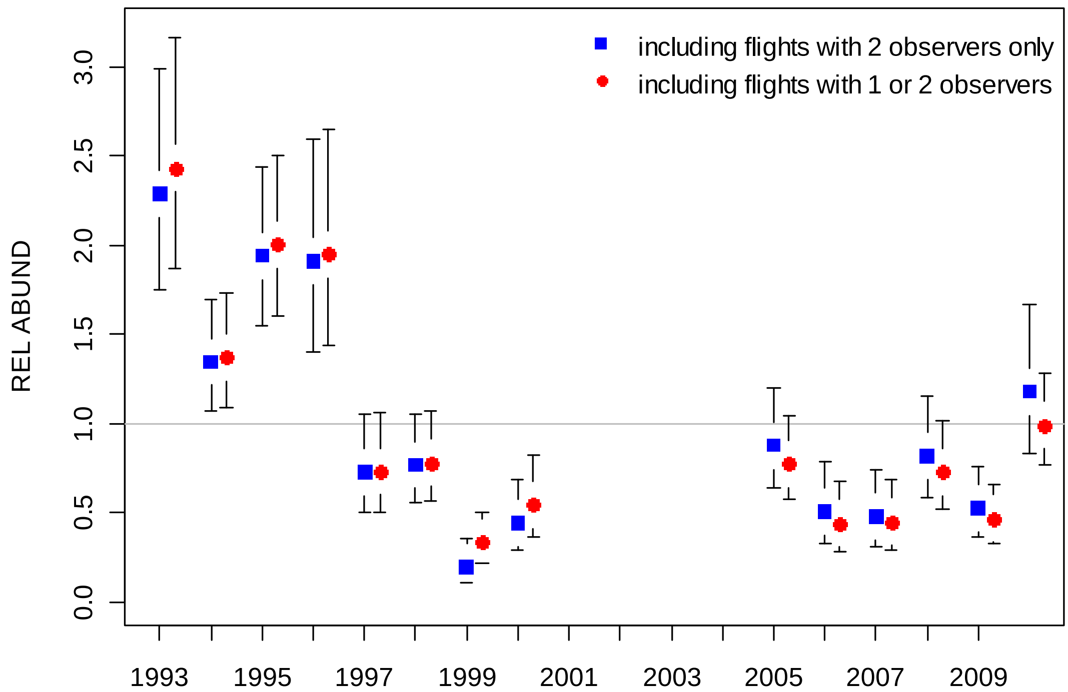


Table 3. Results from the aerial survey analysis, obtained using data from flights with two observers only.

Year	Index	SE	CV	CI.05	CI.95
1993	2.29	0.37	0.16	1.75	2.99
1994	1.35	0.19	0.14	1.07	1.70
1995	1.94	0.27	0.14	1.55	2.44
1996	1.91	0.36	0.19	1.40	2.60
1997	0.72	0.16	0.22	0.50	1.05
1998	0.77	0.15	0.19	0.56	1.06
1999	0.20	0.07	0.37	0.11	0.36
2000	0.44	0.12	0.27	0.29	0.69
2005	0.88	0.17	0.19	0.64	1.20
2006	0.51	0.14	0.27	0.33	0.79
2007	0.48	0.13	0.26	0.31	0.74
2008	0.82	0.17	0.21	0.58	1.15
2009	0.52	0.12	0.23	0.36	0.76
2010	1.18	0.25	0.21	0.84	1.67

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.

Figure 7. Comparison of relative abundance estimates and 90% confidence intervals obtained using data from flights with two observers only versus data from flights with one or two observers.



Model explorations

As already noted, the environmental conditions in 2010 were highly favourable, with lower average wind speed and higher average sea surface temperature (SST) than experienced in any of the past surveys. New data, especially data from such extreme conditions, can significantly affect the estimated model coefficients and, consequently, the relative abundance estimates; thus, we explored the environmental covariates being included in the models and their influence on the abundance indices. The last time model selection procedures were carried out was in 2006.

Exploratory plots of all the available environmental covariates versus the response variables—log(BpS) and log(SpM)—are given in Figure 8 and Figure 9 respectively. SST and air temperature are reasonably correlated (Table 4), so we only consider using SST in the models. The correlation between other variables is quite low.

Table 4. Correlation matrix for environmental variables.

	WindSpeed	Haze	Swell	SST	LowCloud	SeaShadow	AirTemp
WindSpeed	1.00	0.11	0.11	-0.08	0.14	0.07	-0.13
Haze	0.11	1.00	0.12	0.10	0.11	0.02	-0.21
Swell	0.11	0.12	1.00	-0.14	0.00	0.15	-0.27
SST	-0.08	0.10	-0.14	1.00	-0.13	-0.21	0.42
LowCloud	0.14	0.11	0.00	-0.13	1.00	0.07	-0.17
SeaShadow	0.07	0.02	0.15	-0.21	0.07	1.00	-0.29
AirTemp	-0.13	-0.21	-0.27	0.42	-0.17	-0.29	1.00

The current BpS model has just 2 environmental covariates: wind speed and SST. After looking at the exploratory plots of the raw data (Figure 8) and carrying out model selection procedures using a fixed effects version of the model (since it much quicker to fit than the mixed effects version[1]), we added haze and moon phase to the BpS model. The results are as follows:

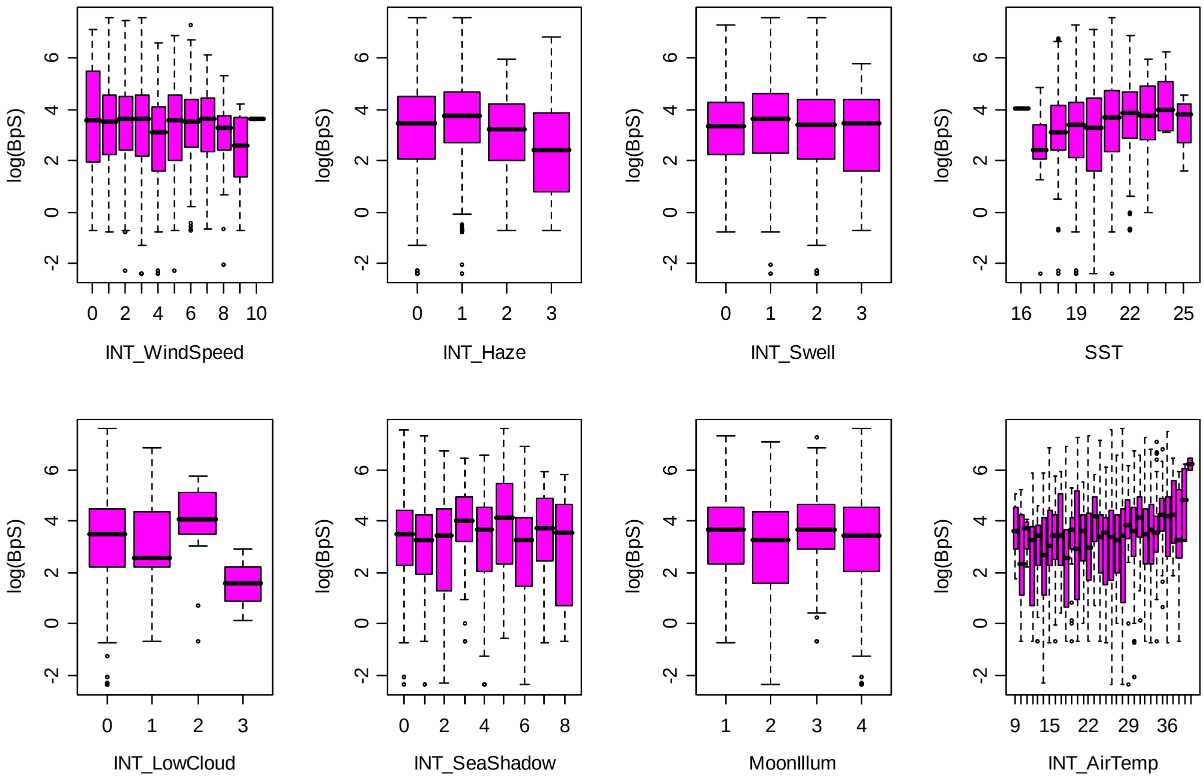
Current covariates	Est	StdErr	t-value	P(>| t |)	
SST	0.335	0.057	5.854	0.000	***
INT_WindSpeed	-0.024	0.028	-0.869	0.385	
Deviance explained = 40.7%					

New covariates	Est	StdErr	t-value	P(>| t |)	
SST	0.335	0.058	5.739	0.000	***
INT_WindSpeed	-0.016	0.029	-0.545	0.586	
INT_Haze	-0.081	0.063	-1.285	0.199	
factor(MoonPhase)2	-0.127	0.150	-0.846	0.398	
factor(MoonPhase)3	0.501	0.202	2.485	0.013	*
factor(MoonPhase)4	-0.051	0.123	-0.414	0.679	
Deviance explained = 40.8%					

[1] Recall that the terms being treated as random effects are the Year, Month and Area 2-way and 3-way interaction terms.

So in the new covariates mixed effects model, only SST comes out strongly significant, with moon phase marginally significant. Based on an analysis of deviance, the new BpS model is not a significant improvement over the current model. These results suggest that only SST needs to be included in the BpS model – we fit this model and found the deviance explained to be 40.7%, which is almost identical to the value obtained from the current and new covariate models, and strongly supports that only SST is important.

Figure 8. Boxplots of observed biomass per sighting, on a log scale, versus environmental conditions present at time of sightings.



The current SpM model has the following environmental covariates: SST, wind speed, haze, swell, and moon phase. Again based on exploratory plots of the raw data (Figure 9) and model selection procedures using a fixed effects version of the model, we added sea shadow to the SpM model. Although the sightings rate appears to decrease with increased low cloud (Figure 9), it turns out that low cloud is recorded as 0 (on a scale of 0 to 3) 97% of the time, so there is not enough contrast in the data to estimate a sensible relationship. Summary results comparing the current SpM model and the new (i.e., with sea shadow) model are as follows:

Current covariates	Est	StdErr	t-value	P(>| t |)	
AvgWindSpeed	-0.302	0.024	-12.499	0.000	***
AvgSST	0.300	0.045	6.624	0.000	***
AvgSwell	-0.176	0.055	-3.18	0.002	**
AvgHaze	-0.145	0.050	-2.891	0.004	**
factor(MoonPhase)2	-0.080	0.105	-0.766	0.444	

factor(MoonPhase)3	-0.139	0.137	-1.012	0.312	
factor(MoonPhase)4	0.179	0.085	2.091	0.037	*
log(ObsEffect)	-1.180	1.132	-1.042	0.297	

Deviance explained = 65.4%

New covariates	Est	StdErr	t-value	P(>| t |)	
AvgWindSpeed	-0.305	0.024	-12.670	0.000	***
AvgSST	0.275	0.046	5.997	0.000	***
AvgSwell	-0.145	0.056	-2.611	0.009	**
AvgHaze	-0.139	0.050	-2.781	0.006	**
AvgSeaShadow	-0.056	0.017	-3.238	0.001	**
factor(MoonPhase)2	-0.076	0.104	-0.730	0.466	
factor(MoonPhase)3	-0.135	0.136	-0.995	0.320	
factor(MoonPhase)4	0.176	0.085	2.063	0.039	*
log(ObsEffect)	-1.285	1.128	-1.139	0.255	

Deviance explained = 66.1%

So in the new covariates mixed effect model, all variables are coming out highly significant except moon phase, which is just marginally significant. Based on an analysis of deviance, the new SpM model (i.e., including sea shadow) does provide a significantly better fit than the current model.

Figure 9. Plots of observed sightings per mile (mean +/- 2 standard deviations), on a log scale, versus average environmental conditions present while searching.

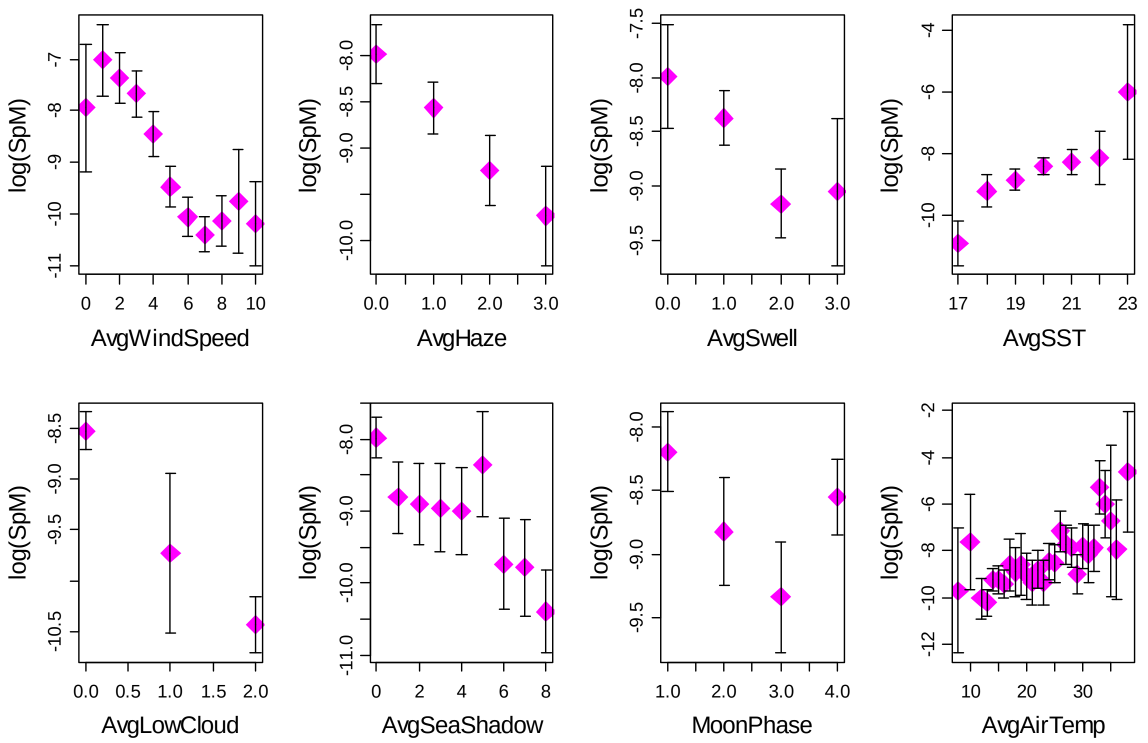
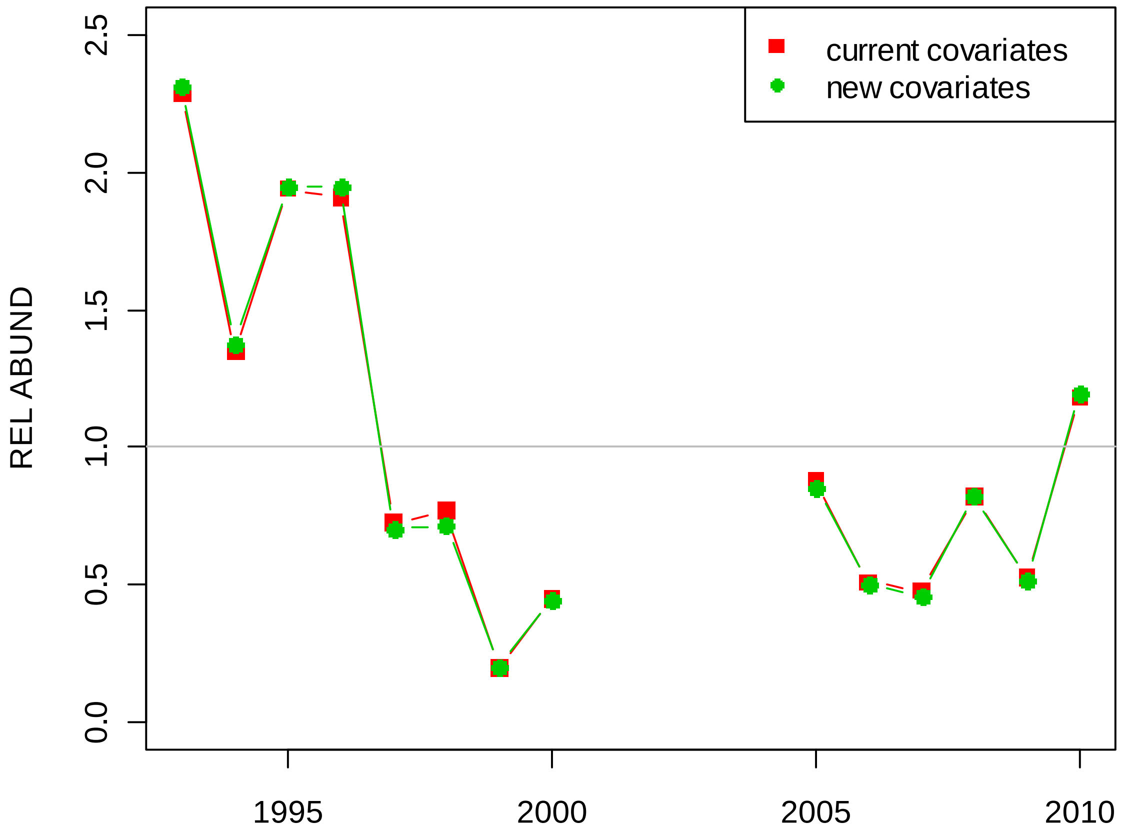


Figure 10 compares the relative abundance time series that are obtained using the current versus new covariate models (i.e., BpS with wind speed dropped and SpM with sea shadow added). There is almost no difference between the results, so the decision of whether or not to make these changes to the models is more a theoretical one than a practical one.

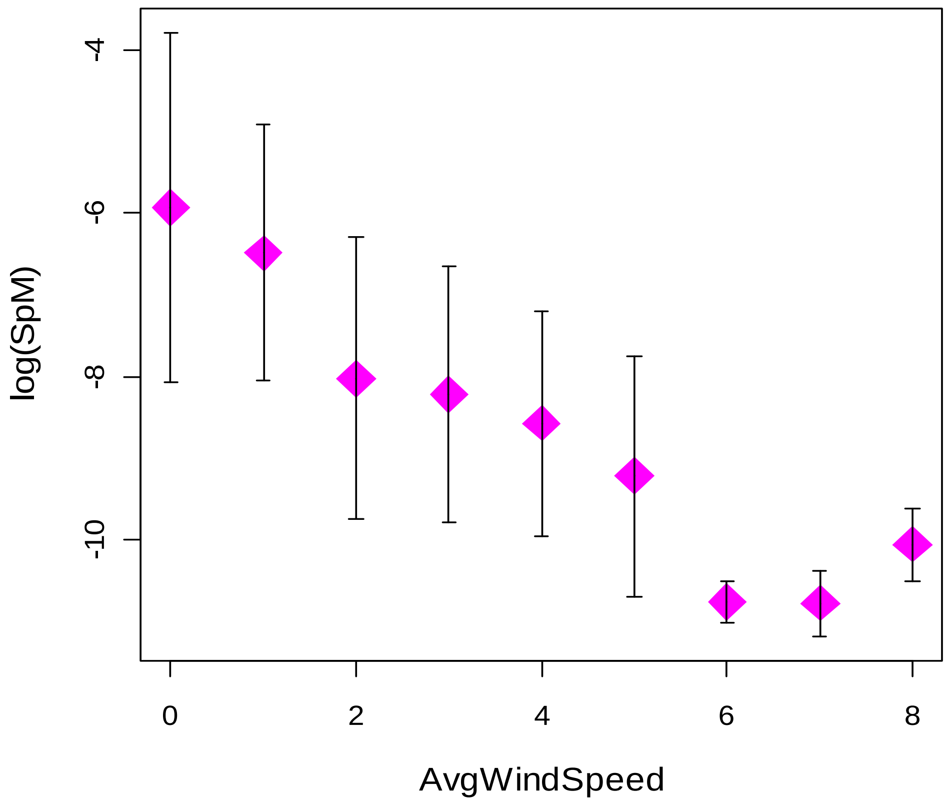
Figure 10. Comparison of relative abundance estimates obtained using the current versus new covariate models (i.e., BpS with wind speed dropped and SpM with sea shadow added). Results were obtained using only data from flights with two observers.



The plot of log(SpM) versus wind speed in Figure 9 suggests that the relationship may be more complex than linear, with the sightings rate actually lower when there is no wind (0 knots) than when the wind speed is ≥1 knot, and levelling off at wind speeds ≥6 knots. This non-linearity was noted in the past but was not considered a problem since almost all observations were within the linear portion of the curve – e.g., of the 1278 sightings made prior to 2009, only 7 were made at a wind speed of 0 and 21 at wind speeds above 6 knots. However, in 2009 and 2010, the weather conditions were above average (2010 in particular) and the number of sightings made at 0 knots was 16 and 56 respectively. This means that the relative abundance estimates for these years (2010 in particular) can be affected significantly by the relationship assumed for wind speed. More specifically, if we use a relationship that has a lower sightings rate at 0 knots compared to a linear relationship, then the 2010 estimate will be larger because it will not get "corrected" as much when we use the model to predict the sightings rate under standardized conditions (but see below, and Figure 11).

Because the relationship between wind speed and sightings rate can have a significant affect on the results, it is worth further consideration. However, for now, there are a couple of reasons why we think it is best to continue using a linear relationship: 1) we have not found a theoretical reason to explain why sightings would decline when there is no wind (glare/reflection on the surface is one potential, but we did not find a relationship between glare and wind speed in the data; plus, observers claim such "millpond" conditions are ideal), and 2) if we redo the plot in Figure 9 using only data from 2010, then the relationship does in fact look almost linear below 7 knots (Figure 11). Thus, the decline at 0 knots seen in Figure 9 is being driven by years with much less data at this wind speed.

Figure 11. Plot of observed sightings per mile (mean +/- 2 standard deviations), on a log scale, versus average wind speed during the 2010 survey only.

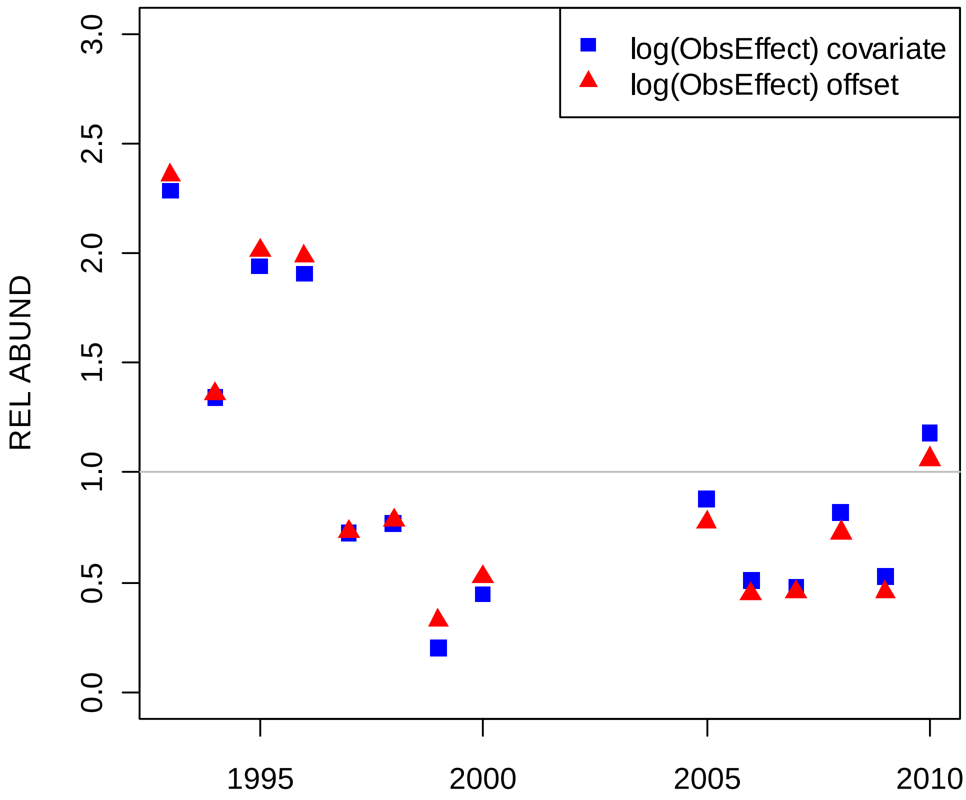


Another issue that came to light with our model explorations was that the observer effect is not being estimated sensibly in the SpM model. As described in the "Sightings per mile (SpM) model" section of Appendix A, prior to fitting the SpM model we estimate relative sighting abilities for all observer pairs, with estimates ranging from 0.72 up to 1.0 for the best pair. If there was good certainty about these estimates, we could just include them as a known offset in the SpM model. However, because there is large uncertainty in the relative efficiencies, we chose instead to include log-relative-efficiency as a covariate in the SpM model, with the effect size (i.e., "slope") to be estimated. If the relative efficiencies estimates are correct, the slope estimate should be close to one. However, as seen in the SpM results presented above, the coefficient of log(ObsEffect) is actually coming out close to negative one. Even though it is also coming out non-significant, it is still a worry because this is highly counter-intuitive (suggesting that the sightings rate actually declines as our estimates of relative sighting ability increases). We investigated further and found that if the SpM model is fitted with only the environmental and observer covariates (leaving out the year, month and area terms), then the coefficient of log(ObsEffect) does in fact come out positive (and significant) as we would expect. This suggests complex interactions between the space-time strata and the observer effects which we have not yet had time to investigate further.

Moreover, we found that when we fit the SpM model including data from flights with only one observer, the problem corrected itself – i.e., the coefficient for log(ObsEffect) came out very close to one and was significant (estimate = 0.96, p-value = 0.003). Presumably this is due to having more contrast in the data: the relative sighting ability estimates for the two observers who flew solo are 0.45 and 0.49, and since the sightings rate was lower on flights with one observer, the slope of the relationship is more obviously positive. When only data from two-observer flights are included, the relative sighting ability estimates are all quite high (between 0.72 and 1), and they do not always correlate well with the observed sightings rate (e.g., an observer pair with a relative sighting ability estimate of say 0.8 can, on average, have a higher sightings rate than an observer pair with a relative sighting ability estimate of 0.9).

Clearly we do not want to continue making predictions using an SpM model with an illogical coefficient estimate for the observer effect. Although this will not be the case if we use the results which include data from flights with one observer, it is disconcerting that the approach is not robust to changes in the data. One possible solution is to include the relative sighting ability estimates as offsets to the model and develop alternative ways to propagate the uncertainty in these estimates through to the model predictions; we are currently investigating ways to do so. In the meantime, we can see how much the relative abundance point estimates (without CVs) differ when the SpM model is fit to data from two-observer flights with log(ObsEffect) included as an offset opposed to as a covariate (since the estimates provided to CCSBT for the data exchange and for inclusion in MP work were based on the latter). The results are shown in Figure 12; fortunately, the estimates are not too different.

Figure 12. Comparison of relative abundance estimates obtained when log(ObsEffect) is included as an offset in the SpM model versus when it is included as a linear covariate. Results shown were obtained using only data from flights with two observers.



Summary

This was the first year that planes with only one observer flew as part of the scientific survey. In order to make our results directly comparable to results provided in past, we first analysed the data using only data from flights with two observers. We also redid the analysis including data from flights with one observer; this required using new methods for calibrating the data from the one-observer flights. The results from the analysis using only two-observer data are very similar to the results using both one and two-observer data. In both cases, the relative abundance point estimate for 2010 is the highest estimate since 1996, but remains lower than the average level in the period 1993-1996. We note that the CV for the 2010 estimate obtained using one and two-observer data is an underestimate because there is extra uncertainty in the observer effect estimates for solo observers that are not being accounted for in the SpM model – we plan to address this issue in the coming year.

The environmental conditions in 2010 were highly favourable, with lower average wind speed and higher average sea surface temperature (SST) than experienced in any of the past surveys. New data, especially data from such extreme conditions, can significantly affect the estimated model coefficients and, consequently, the relative abundance estimates. Thus, we explored the environmental covariates being included in the models and their influence on the abundance indices. These explorations brought up some issues for further consideration, the most important being: 1) whether a linear relationship between wind speed and sightings rate is adequate; and 2) the current method of including observer effects in the SpM model is not robust and can lead to counter-intuitive results. In terms of 1), we suggest that the evidence against a linear relationship is not strong enough to change the model. In terms of 2), we are currently investigating an alternative method for including observer effects in the SpM model that should correct the problem.

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.

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.

Acknowledgements

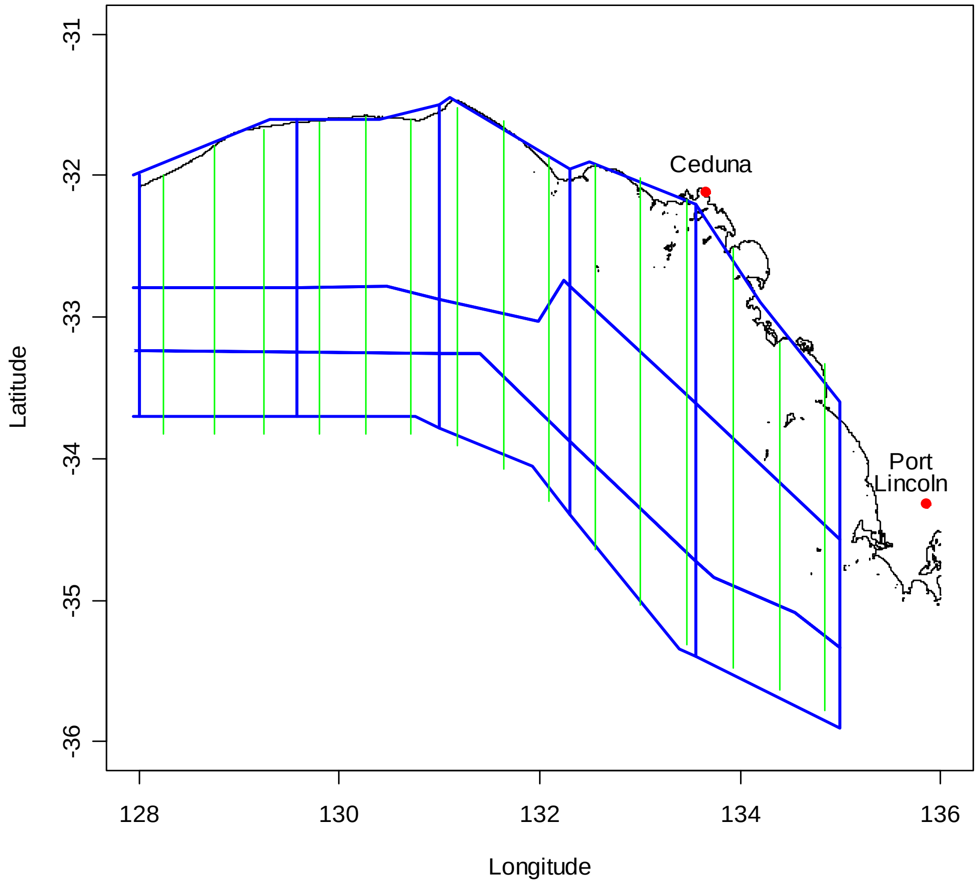
There are many people we would like to recognise for their help and support during this project. We would especially like to thank this year's observers (spotters), pilots and data recorders; Andrew Jacob, Darren Tressider, Derek Hayman, John Veerhius, Thor Carter and Jarred Wait. This study was funded by AFMA, DAFF, the Australian SBT Industry, and CSIRO's Wealth from Oceans Flagship.

Appendix A – Methods of analysis

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

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

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



Biomass per sighting (BpS) model

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

The BpS model was fitted using a 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.[2] 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 **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

[2] Note that the selection of environmental covariates was undertaken as part of the 2006 analysis, and has not been repeated since then. However, now that 4 more years of environmental and survey data have been collected, we plan to re-visit the selection of model covariates in the coming year.

pairs that have flown at some point in the surveys. The observer pairs ranged in their estimated sighting rates from 72% to 97% compared to the pair with the best rate.

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

The data used for the SpM model were accumulated by flight and area, so that the data set used in the analysis contains a row for every flight/area combination in which search effort was made (even if no sightings were made). Within each flight/area combination, the number of sightings and the distance flown were summed, whereas the environmental conditions were averaged. The SpM model was fitted using a 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[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 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. 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 (see footnote 2). The final model fitted was:

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

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

Combined analysis

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

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

would have been under standardized environmental/observer conditions[4]. Using these predicted values, we calculated an abundance estimate for each stratum as 'standardized SpM' multiplied by 'standardized average BpS'. We then took the weighted sum of the stratum-specific abundance estimates over all 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.

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

Appendix B – CV calculations

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

Let $\hat{B}_{ijk}$ be the predicted value of BpS in year *i*, month *j* and area *k* under standardized environmental/observer conditions (see footnote 4 above), 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[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

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

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

---

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

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

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

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

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