CCSBT-ESC/0809/25

Commercial spotting in the Australian surface fishery, updated to include the 2007/8 fishing season

Jessica Farley
Marinelle Basson

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

Table of Contents

Abstract

Data on the sightings of SBT schools in the GAB were collected by experienced tuna spotters during commercial spotting operations between December 2007 and March 2008. Spotting data has now been collected over seven fishing seasons (2001-02 to 2007-08). In all seasons, the majority of search effort occurred in December to March, and the areas of highest SBT abundance per nautical mile searched were within a “core fishing area” close to the shelf-break, and around the inshore lumps/reefs. The commercial spotting data was used to produce nominal and standardised fishery-dependent indices of SBT abundance (surface abundance per unit effort – a SAPUE index).

Introduction

In the summer of 2001-02 (called the 2002 season), a pilot study was conducted to investigate the feasibility of using experienced industry-based tuna spotters to collect data on the sightings of SBT during commercial spotting operations in the Great Australian Bight. The data provided a preliminary fishery-dependent index of SBT abundance (surface abundance per unit effort – a SAPUE index) for that fishing season. Recognising the importance of time-series of indicators, we continued to collect and analyse SBT sightings data from commercial tuna spotters over the following 6 fishing seasons (2003-2008). Interpretation of the results are difficult as the data suffers from many of the same problems that affect catch per unit effort (e.g. changes in coverage over time, lack of coverage in areas where commercial fishing is not taking place, and changes in operations over time), but it may provide a qualitative indicator of juvenile SBT abundance in the GAB. It has always been recognised, however, that a line-transect survey with consistent design and protocols from year to year is highly preferable. In 2008, we continued to collect SBT sightings data from commercial spotters. This report summarises the field procedures and data collected, and provides results of analyses for all seven seasons (2002-2008).

Field procedures

Data were collected on SBT schools sighted by four spotters engaged between December 2007 and March 2008 (called the 2008 fishing season). In previous seasons, data has been collected from up to 6 spotters, but this year only three spotters were required by Industry. In the 2002 to 2005 fishing seasons, <1% of search effort occurred before December or after March; thus data were only collected from December to March since the 2006 season.

The spotting data were collected following the protocols used in the previous six fishing seasons. Within each plane there was a spotter and pilot. For most flights, the spotter searched the sea surface on both sides of the plane for surface schools of SBT. During some flights, the pilot also searched for schools. A GPS was used to log the position of the plane and record waypoints. Sighting information and environmental conditions were recorded by the spotter and/or pilot in a logbook (not by a separate data recorder). The start and end of “search” periods were recorded, so that transit time to and from the fishing area, or periods of time when the spotter was not searching for fish, could be removed from the analysis. There were no restrictions on the environmental conditions for commercial spotting operations, although they rarely occurred when wind speeds were above 10-15 knots.

When a “sighting” of SBT was made, a waypoint (position and time) was recorded over the school (or schools). The spotter estimated a range for the size of fish in the schools (in kg) and the biomass of each school (in tonnes). It is important to note that many SBT schools are recorded as single schools (34-62% each season). Some schools, however, are recorded in groups of 2-10 or even 50+ schools. Environmental observations were recorded at the start and end of each flight and when the conditions changed significantly during the day. The environmental observations included wind speed and direction, air temperature, cloud, visibility, spotting conditions and swell. The spotter also recorded the type of search effort (restricted or broad scale) undertaken during the flight. The target species of each flight (SBT, skipjack tuna, mackerel, or a combination of these) was also recorded.

Results

Search effort and SBT sightings

Data were collected for 93 commercial spotting flights in the 2008 fishing season. Due to a problem with one spotter's GPS, flight path data for 8 flights in March could not be obtained and thus the proportion of search time and biomass sighted in the 'core' fishing are unknown for those flights. However, the total search effort, biomass recorded and environmental conditions for the flights are known and are included in the analyses.

The relative contribution to the total search effort by spotter is given in Table 1, and details of search effort and SBT sightings are given in Table 2. SBT were recorded on 75 of the 93 commercial flights in 2008 (80.6%). Note, however, that the total biomass shown in Table 2 does not represent the total biomass of SBT present in the survey area, as many schools were potentially recorded several times (either by different spotters on the same day or over several days).

The location of SBT sightings varied slightly between seasons (Figure 1) but the areas of highest SBT sighted per nautical mile searched remains within the same 'core fishing area' (130.2-132.9°E and 32.7-34.0°S) and around the inshore lumps/reefs each season. Figure 2 and Figure 3 show the size of SBT schools and fish recorded by spotter 1 between 2002 and 2008. Using data from one spotter removes the problem of differences between spotters in their estimates of school and fish size. Spotter 1 was selected because he had collected data on the greatest number of SBT schools each season. On average, it appears that the mean size of schools has increased, and the mean size of fish has decreased, since 2004.

Table 1. Relative contribution (%) by spotters to the total search effort (time) by fishing season

Season	Spotter 1	Spotter 2	Spotter 3	Spotter 4	Spotter 5	Spotter 6
2002	61.3	7.6	11.7	-	5.6	13.9
2003	20.2	11.5	33.2	1.2	4.4	29.5
2004	42.2	15.2	19.4	-	-	23.2
2005	39.7	9.3	19.5	-	5.0	26.5
2006	44.2	11.6	-	-	14.8	29.5
2007	38.0	11.1	-	-	22.1	28.8
2008	37.3	23.7	-	-	-	39.0

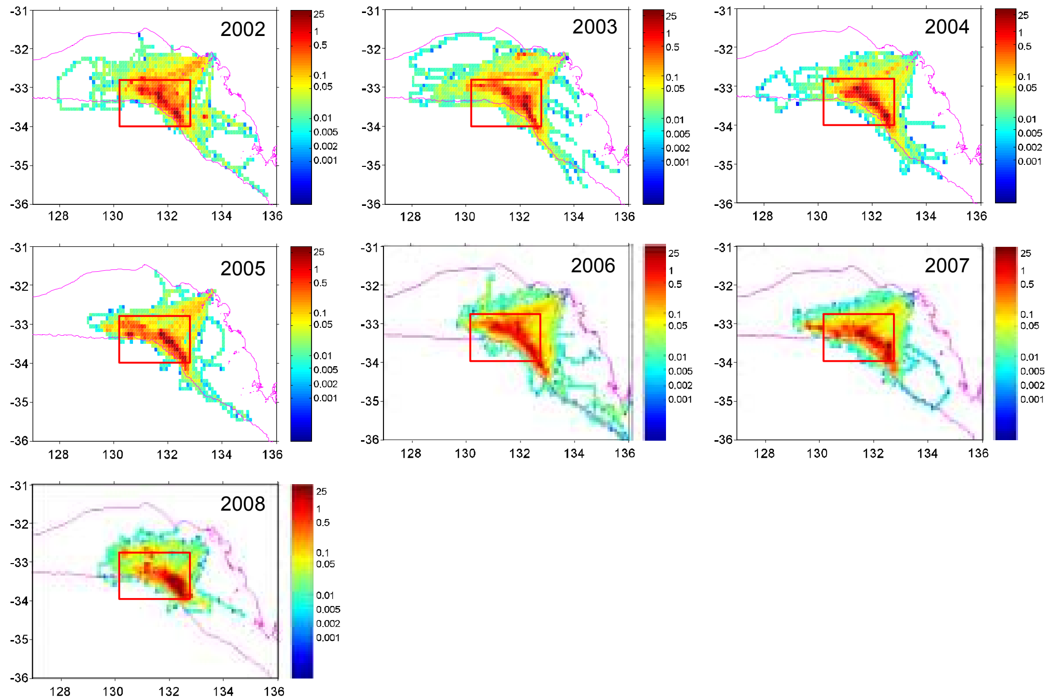
Table 2. Search effort and SBT sighted by spotters in the 2002-2008 fishing seasons.

Fishing season	No. flights	Search effort (hrs)	% flights with SBT recorded	Total number of schools	Total biomass[1] recorded	% of effort in the core[2]	% of biomass in the core[2]
2002	86	325	84	1182	44626	75.4	91.8
2003	102	425	82	1301	38559	80.2	78.4
2004	118	521	77	1133	33982	89.0	75.7
2005	116	551	94	2395	87447	89.5	83.2
2006	102	452	82	1554	50524	83.2	72.4
2007	120	600	92	2600	94018	86.5	80.0
2008	93	451	81	2529	100341	85.1	78.6

[1] The total biomass recorded does not represent the total biomass of SBT present in the survey area, as many schools were potentially recorded several times (either by different spotters on the same day or over several days).

[2] Does not include data for flights where flight path data was not obtained. i.e. 20 flights in 2005 (See Basson and Farley, 2005; CCSBT-ESC/0509/23) and 8 flights in 2008 (see above).

Search effort



SAPUE

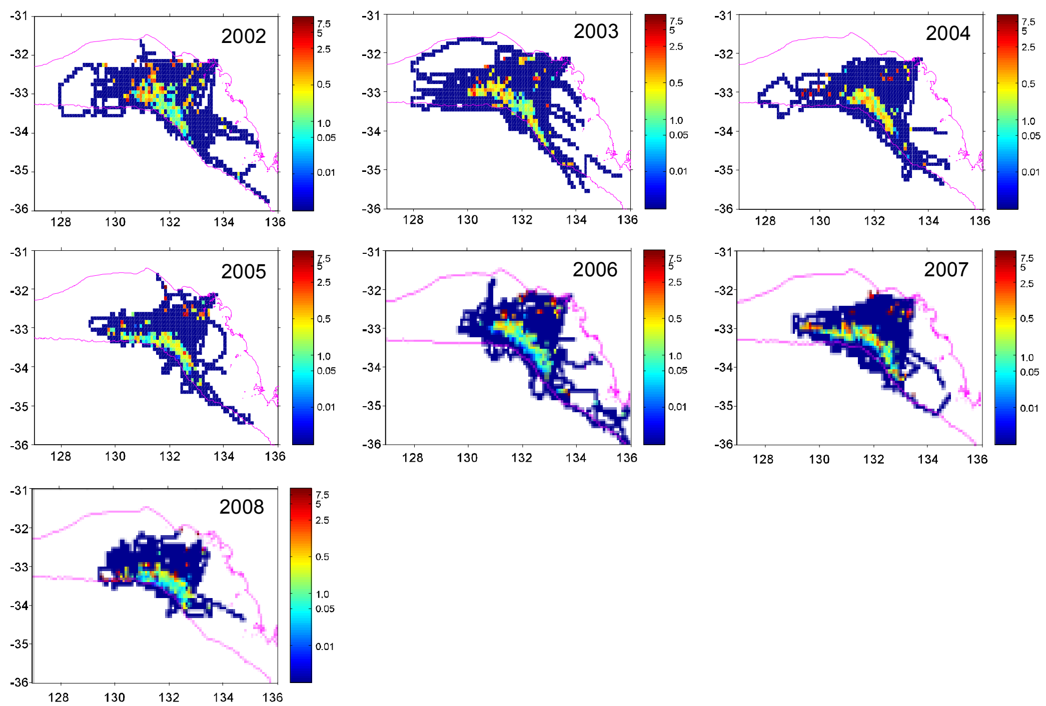


Figure 1. Search effort (nm flown/0.1° square), locations of SBT sightings, and SAPUE (tonnes/nm/0.1° square) in the GAB by fishing season. SAPUE data are displayed as the % of total effort for the season. Areas of darkest blue in the SAPUE plot indicate zero SAPUE. Note the log scale for effort and SAPUE. The core fishing area is shown by a red square.

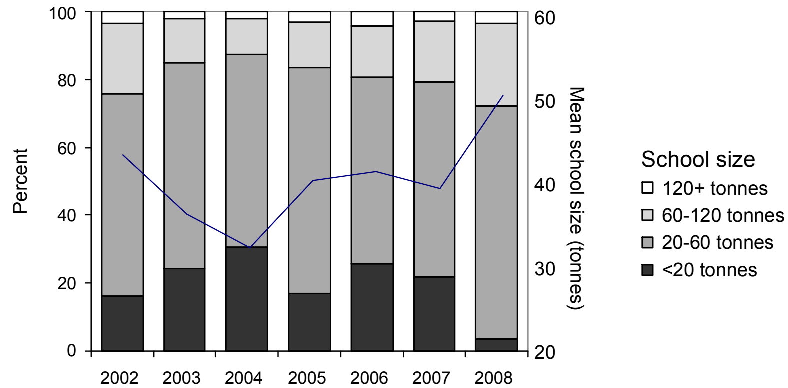


Figure 2. Proportion of SBT schools by size class (bars) and mean school size (line) recorded by one spotter in the 2002-2008 fishing seasons. Total number of school size estimates = 5,439.

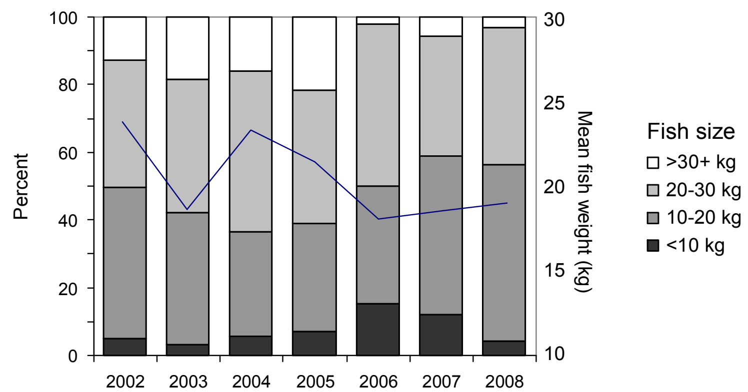


Figure 3. Proportion of SBT by fish size class (bars) and mean fish size (line) recorded by one commercial spotter in the 2002-2008 fishing seasons. Data are weighted by school size. Fish size data collected for 5,398 schools.

Nominal SAPUE

As for previous years, the duration of "search" sectors during flights were calculated using the GPS logged position and time. The logbook data on SBT sightings were summarised to give the total number of sightings, schools, and total biomass per plane per day. The data were extracted to ensure consistency between seasons (e.g. flights in November and April, outside the main fishing season and with relatively low coverage, were excluded; flights with less than 30 minutes of search effort were excluded because these were considered too short to have a meaningful SAPUE estimate). As these data were removed for all seasons, it should not affect the relative index of abundance. Nominal (unstandardised) indices of juvenile SBT abundance (surface abundance per unit effort – SAPUE) were calculated, based on the mean of biomass sighted (tonnes) per unit of search effort (minutes). The SAPUE indices were calculated by geographic area (whole GAB and core fishing area), by search type (broad and restricted), and for flights where SBT was/was not targeted.

Three nominal SAPUE indices of juvenile abundance are shown in Figure 4a. These indices show declines prior to 2004 then all four fluctuated similarly between 2005 and 2008, and showed increases in the last three seasons. Not surprisingly, mean SAPUE was higher for flights that SBT were recoded on, but as there were relatively few flights where SBT were not recorded (13.9%), it makes little difference to the overall SAPUE index obtained by month. The lower mean SAPUE for the core fishing area in many seasons is also not surprising given that search effort is highest in this area. Figure 4b shows the comparison of mean SAPUE by search type. Since the type of search effort (broad/restricted) was not recorded in 2002, these indices can only be calculated for six seasons. The nominal SAPUE estimate for 2008 is higher than the 2006-2007 seasons for broad search effort, but similar to the 2007 estimate for restricted search effort.

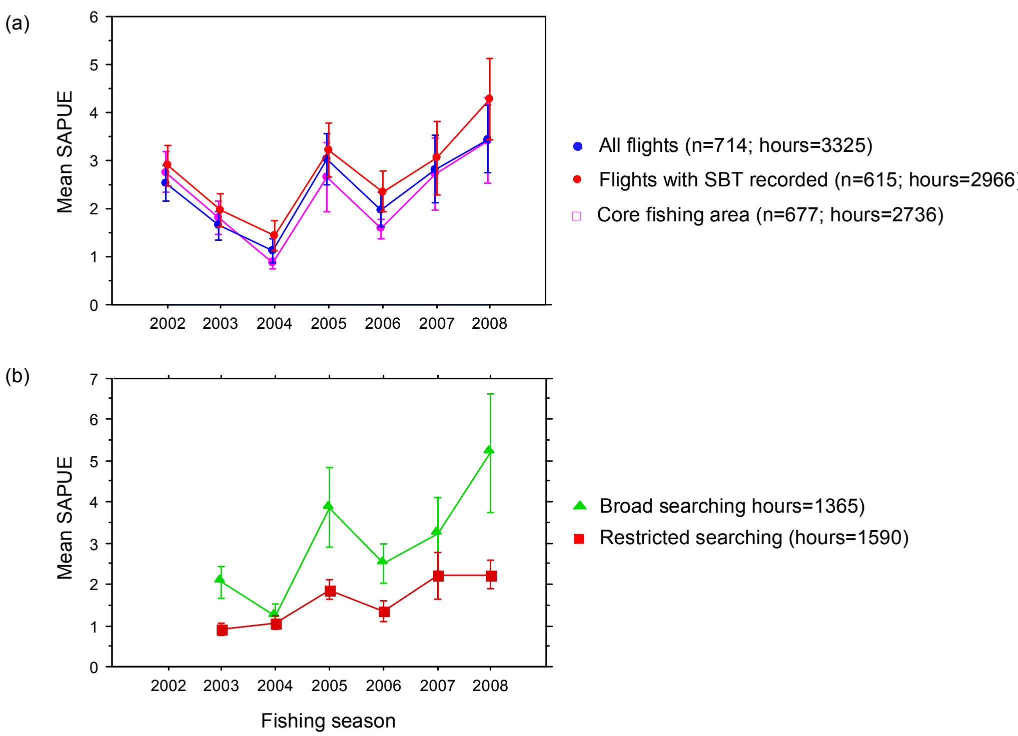


Figure 4. Nominal SAPUE indices (+/-se) (tonnes of SBT sighted per minute searching) for the 2002-2008 fishing seasons (a) for all flights, flights in the core area, or flights that SBT were recorded, and (b) by search effort type. Classifying search effort as either broad or restricted started in 2003 (i.e. the 2002/03 fishing season). Note that only flights in December to March were included, and when search effort was >30 minutes.

Standardised SAPUE

There are now seven years worth of commercial spotting data which can potentially be standardised to obtain an index of juvenile abundance (ages 2-4 primarily) in the GAB between December and March. Although data from 5 companies are available, summaries of the number of days flown in each month and season show that two of the companies flew a limited number of days and only in some months/seasons (Table 3). This is understandable because these companies took a relatively small proportion of the surface fishery catch, and it should be remembered that the commercial spotting is directly and strongly linked to the commercial fishing operations. The number of spotters required by Industry in 2008 decreased to only three, as there has been a tendency over time for fewer fishing companies to catch tuna for the other companies in the fishery. This is important from the point of view of interpretation of the data. The commercial spotting data can suffer from many of the same hard-to-quantify biases that affect catch per unit effort, for example, changes in coverage over time, lack of coverage in areas where commercial fishing is not taking place –for whatever reasons – and changes in operations over time. From a statistical perspective, the scientific aerial survey, which uses a line transect design and consistent protocols, is far preferable as an approach to an index compared to the commercial spotting. However, these additional (commercial spotting) data can potentially provide further insights given the relatively large amount of effort (hours flown).

Table 3. Number of days flown by spotter, year and month within year. Note that the 'season' is the same as the 'year' for all months except December; for example December 2001 will fall in the 2002 Season.

Year	Month	spotter1	spotter2	spotter3	spotter5	spotter6
2001	Dec	14		8		4
2002	Jan	7	5	5		7
2002	Feb	7	3	3	4	4
2002	Mar	11				
2002	Dec			10		10
2003	Jan	10	6	9	5	10
2003	Feb	2	3	6	1	4
2003	Mar	5		6		4
2003	Dec			11		10
2004	Jan	9	7	5		11
2004	Feb	15	10	9		6
2004	Mar	16		2		4
2004	Dec			4		3
2005	Jan	11	7	9	1	7
2005	Feb	9	2	10	6	16
2005	Mar	19		2		8
2005	Dec	9			3	4
2006	Jan	8	4		3	8
2006	Feb	9	8		9	9
2006	Mar	12			4	10
2006	Dec	6			2	7
2007	Jan	15	7		10	14
2007	Feb	9	6		7	7
2007	Mar	12			11	6
2007	Dec	5				11
2008	Jan	11	11			9
2008	Feb	11	6			12
2008	Mar	8	5			4

Given the changes in spotting effort by the 6 spotters over time (Table 3), exploratory analysis was again done using different combinations of spotters. Results proved not to be particularly sensitive to the choice, so final analyses were done using data for spotters 1,2,5 and 6, as well as for the two spotters who have consistently spotted in all four months and contributed most of the effort, namely spotters 1 and 6. Data from all months (Dec, Jan, Feb and March) were included in the analyses. As noted in the past, the change in the 2006 season (Table 3), when the effort for spotter 3 (also referred to as company 3 in previous working papers[1]) dropped to zero, but that for spotter 5 increased, causes several difficulties for the analysis. It has become more difficult to fit models with an interaction term between spotter and season due to the unbalanced data. In 2007, spotter 3 again did no spotting, though spotter 5 continued and contributed a relatively large number of days' data to the database. In 2008, however, spotter 5 did no spotting, but spotter 2 contributed more effort and spotted in 3 months.

Environmental variables

As noted in the past (e.g. CCSBT-ESC/0409/19) sighting conditions and surfacing behaviour are influenced by weather and environmental variables. The environmental variables recorded by season are summarised in Table 4 and Figure 5. Note that the aerial survey transects are only flown during certain conditions, so that summaries of environmental conditions recorded during the aerial survey and during commercial spotting operations would tend to differ. The data show that during the 2008 commercial spotting flights, the average wind speed, air temperature, spotting conditions and visibility were higher than in 2007, while swell height and cloud cover were lower. Overall, however, the environmental data were not particularly unusual compared to previous seasons, apart from visibility which showed greater variability in 2008 compared to previous seasons. It appears this may be due to one spotter/pilot incorrectly recording total visibility on the day, rather than estimating the distance from the aircraft that the spotter is able to search for SBT schools.

We have noted previously (e.g. CCSBT/ESC/0609/17) that although the mean temperature can be quite similar between seasons, the monthly temperatures can be very different. Figure 6 shows the monthly mean temperatures from the data over the past 7 seasons. In 2008, the difference between the average January/February temperature and that for March was quite large. The March average temperature was the highest recorded (the highest overall and the highest March temperature), and the December temperature was the highest of the December temperatures in the dataset.

Analyses of the aerial survey data found that moon illumination was a significant term and it is plausible that this could affect surfacing behaviour. Moon illumination was therefore again included in the standardisation analysis.

[1] Although we use the terms 'company' and 'spotter' interchangeably, the data pertains to a particular spotter.

Table 4. Average environmental conditions during search effort on commercial flights by season (Dec-Mar only). Note visibility was not recorded in 2002.

Fishing season	Wind speed (knots)	Swell height (0-3)	Air temp (°C)	Cloud cover (/8)	Spotting condition (/5)	Visibility (nm)
2002	7.05	1.46	17.91	4.48	2.64	
2003	6.94	1.21	23.35	3.66	2.79	5.54
2004	7.91	1.65	19.73	3.94	2.64	7.77
2005	6.99	1.59	21.14	4.23	2.55	8.95
2006	7.59	1.95	22.11	4.01	2.75	7.64
2007	6.69	1.79	21.10	3.60	2.78	7.92
2008	7.95	1.48	22.87	2.02	2.90	10.80

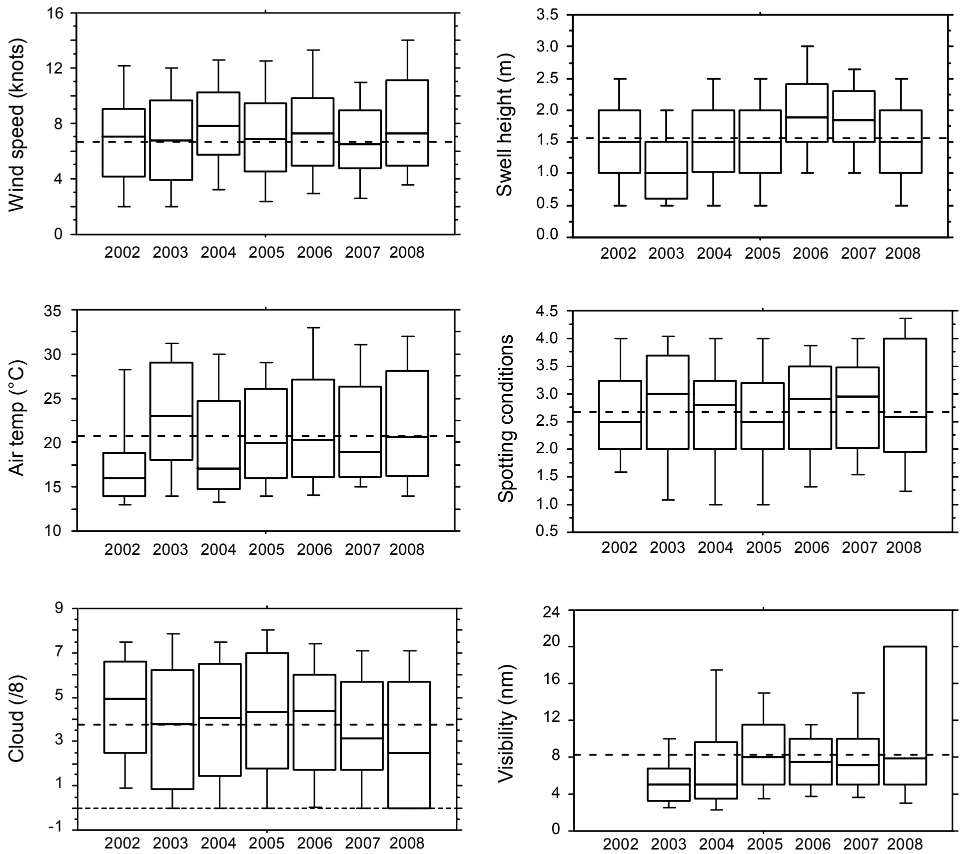


Figure 5. Boxplots summarizing the environmental conditions present during search effort on commercial flights by season (Dec-Mar). The horizontal band through a box indicates the median, the length of a box represents the inter-quartile range, and the vertical lines extend to the minimum and maximum values. The dashed line running across each plot shows the overall average across all survey years. Note visibility was not recorded in 2002.

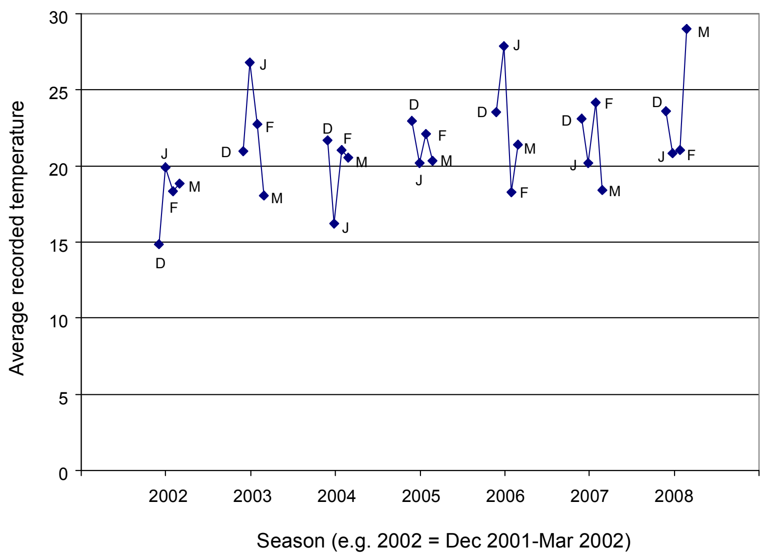


Figure 6. Average monthly temperatures (Dec to Mar) from the spotting data for the past 5 seasons.

The sightings data

As indicated in the past, there are many different ways in which the sightings data could be compiled for analysis. The best way would be to compile the data at as fine a time and spatial scale, to give some chance of partly adjusting for the lack of spread of spatial coverage and the autocorrelation in the observations. This task would, however, be seriously complex and given that an aerial survey was again conducted this season, it is not warranted. Instead, we have followed the approach used in the past. The data are compiled as the biomass sighted and effort in hours flown on each day by each spotter. The associated environmental variables are taken as the means for that day and spotter. The data were compiled as a set for the entire area and all the analyses were done on the ‘whole area’ dataset. Table 5 shows a summary of the number of days flown with no biomass sighted. This information can be treated as a simple ‘presence’/’absence’ index. The percentage days with no sightings were below average in 2005 and 2007.

Table 5. Number of days flown with no biomass sighted and days with some biomass sighted, for all companies combined and all months. Since different levels of effort are associated with each day, the % effort in hours associated with days when no biomass was sighted is also shown.

Season	Zero biomass days	Positive biomass days	Total days	% days with Zero biomass	% effort (hours) associated with zero biomass
2002	10	72	82	12.2	10.0
2003	15	76	91	16.5	11.9
2004	25	90	115	21.7	15.7
2005	6	108	114	5.3	4.1
2006	16	84	100	16.0	11.5
2007	9	110	119	7.6	4.8
2008	18	75	93	19.4	16.3

Modelling approach

We used the same modelling approach as in the past and updated those analyses with data from the 2008 season. The main intention of modelling of these data is to standardise the raw index (e.g. average biomass per unit effort sighted) for differences between spotters and different environmental, weather and spotting conditions from year to year. Some of the variables (e.g. moon illumination) most likely only affect surfacing behaviour of tuna, whereas others (e.g. wind, swell) may affect both spotting ability and surfacing behaviour. The "regression model" used must be able to cope with the zero observations, and with the strong dependency of the variance on the mean. A convenient way to do this is to fit GLMs using the Tweedie family of distributions (Jørgensen, 1997; see also Candy 2004) with a log-link, so that different factors combine multiplicatively. The mean-variance relationship in Tweedie distributions follows a power-law with adjustable exponent Φ, and for $\Phi<2$ there is no problem with zero observations. When fitting the models, the exponent Φ was entered ($1< \Phi <2$). Note that the value of $\Phi=1$ coincides with the Poisson distribution, and a value of $\Phi=2$ with the Gamma distribution. In the past different values of Φ were tried and the deviance residuals were checked to ensure that they were relatively similar over the range of predicted values. This showed that a value of $\Phi=1.5$ is acceptable, and this value was used for all analyses in this working paper.

All analyses were done in R using library(Tweedie) to enable use of "family=tweedie()" in the standard GLM routine. The Akaike information criterion (AIC) statistic was primarily used to compare model fits and bootstrapping was used to explore the estimated variance of parameter estimates.

In the past, data and model exploration, suggested that all the environmental covariates in the dataset were important, though swell was only marginally relevant – including or excluding it had little effect on results or on the AIC statistic. In general, records with missing values for any of the environmental variables in the model are excluded. However, spotter 5 often did not record swell, and if those records are excluded, the dataset (particularly in the most recent 2 seasons) is much smaller. We again explored models which include or exclude 'swell' as a covariate. All results are based on the following model with swell either included or excluded:

Full model without interaction:
biomass ~ as.factor(season) + as.factor(company) + as.factor(month) + wind + spotcon + swell + cloud + temperature + moonillum + offset(log(effort))

As noted in previous working papers on the standardisation of the commercial spotting data, an interaction term between spotter and season appeared to be important, though it does make interpretation of results rather difficult because this model implies a different index time-series for each spotter. However, the changes in effort for all spotters, particularly since the 2006 season, have lead to an unbalanced dataset. We consider that it is still not meaningful to use the model with interaction term to obtain a standardised index of abundance for the whole period[2]. All results are therefore for the no-interaction model. We did, however, look at the sensitivity of the index to using data for different groups of spotters.

[2] The index is constructed by predicting the biomass per unit effort at average values for covariates and a given reference level (for factor variables) using the model. In this case, however, the predictions are not reliable because the model matrix is rank deficient.

Results

Results for six datasets are shown in Figure 7: (i) all spotters (1,2,3,5,6), (ii) spotters 1,3,5 and 6, and (iii) only spotters 1 and 6 with swell included; (iv) spotters 1,3,5,6 with swell excluded, (v) spotters 1,2,5,6 with swell excluded and (vi) spotters 1 and 6 only with swell excluded. The estimated index series are all very similar. The main difference between point-estimates of the index occurs when spotter 3 is not included; the relative values for 2003 and 2004 change, but other values are almost unaffected by the choice of spotters.

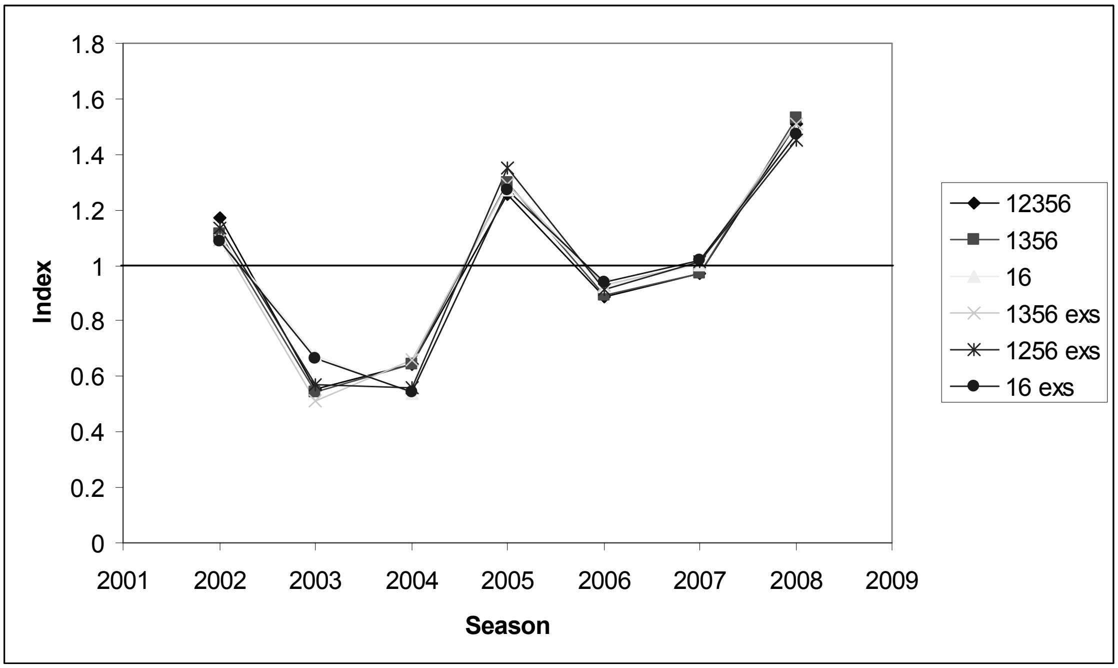


Figure 7. Time-trends of the standardised SAPUE indices (surface abundance per unit effort) scaled to the mean for 3 models with swell included and 3 models with swell excluded. The legend shows which spotters are included (e.g. spotters (1,2,3,5,6)='12356') and 'exs' indicates that swell is excluded. Season refers to the 2nd year e.g. 2006 indicates the 2005/06 season.

The coefficients of variation (CVs) of the season-coefficients are also hardly affected by the inclusion or exclusion of swell as a covariate, irrespective of the fact that there are more records when swell is excluded. Recall that it is not just the model that is different when swell is excluded, but also the dataset (when swell is excluded, records which have swell recorded as missing are included in the analysis). In particular, the flight in 2007 that appears to have been done more like a 'stock-take' flight, and discussed in the section on 'Nominal SAPUE' is included in the model/data which excludes swell. The residual for this data point is in fact obvious in the diagnostics (Appendix, Figure A 2, third panel); the point with highest observed biomass. The comparisons of estimated index values from different model/data combinations, however, suggest that this single point is not having a disproportionate effect on the standardised results.

Given the similarities in results and diagnostics for the 6 datasets shown above, we present results in full for only two of these. For consistency with results presented last year and provided in the data exchange in 2007, we present the model with <u>spotters 1,2,5,6 and with swell excluded</u>. Although spotter 2 has only contributed records in January and February in most years, spotter 2 contributed data for 3 months (January, February and March) in 2008, and there are data for all 6 seasons. Given the fact that spotters 1 and 6 has spotted most consistently (all months in all years), and that spotter 5 did no commercial spotting in 2008, we also consider a model with just these two spotters (1 and 6) and with swell excluded.

Diagnostics for the two models show that residuals are reasonably well-behaved, though the qq-plots are rather poor, and not linear as expected (see Appendix, Figure A 1 and Figure A 2). This is unlikely to badly affect the point-estimates of coefficients, but does indicate a 'fat' tail in the data. In a relative analysis such as this, where the focus is on year-to-year comparisons, poor qq-plots do not generally imply bias in the point-estimates, but do point to the need to validate standard errors. This is done by bootstrap analyses, discussed below.

In the past we explored the effect of using a different assumption about the mean-variance relationship through different values of the Tweedie parameter, Φ. Those results suggested a value of 1.5 is appropriate. Here we checked whether this is still the case. Figure A 3 (Appendix) shows the deviance residuals (square root of the absolute values) plotted against the fitted values for $\Phi=1.5$. The smoother through the data shows that this value is still appropriate, since the smoother is relatively 'flat' (slope close to 0) for this assumption. A slightly lower value, $\Phi=1.4$, was also tried with the model based on data from spotters 1 and 6 only. Although this leads to a slightly flatter relationship, the estimated index is essentially unaffected.

Comparisons between estimated standard errors from the GLM model and estimates from bootstrap analysis, as described in Basson and Farley (2005; CCSBT-ESC/0509/23), were made in the past. These comparisons showed that the model estimates of standard deviations were no smaller than the bootstrap estimates (from 500 replicates). We have not redone this analysis here under the continued assumption that the standard errors from the model can be used to indicate the uncertainty in the index. This assumption should be rechecked in future. As in the past, we note that the standard errors describe only the uncertainty about the season level given the available data; there is an extra layer of uncertainty, about how many SBT were in the GAB outside the area covered by the SAPUE, that the model cannot reveal.

Figure 8 shows results of the standardised index for the two models, both excluding swell, but one for spotters 1,2,5,6 and the other for spotters 1 and 6. The ranges were obtained by taking the predicted values + or – 2 standard deviations on the log scale and then converting to the normal scale. Note though, that the standard deviations themselves take into account the fact that the index has been scaled to the mean.

Results of the estimated index value and standard error are shown in tabular form in Table 6. Note that since the index is scaled to the series mean, values for earlier years will change as new seasons' data are added to the analysis. The index values for the seasons 2002 – 2007, model 1, are therefore not identical to the values that were exchanged in the data exchange of last year (2007). We note that there is no basis for preferring one model over another at this stage, and results are so similar that the choice should not matter. For consistency with last year, we suggest that the model using data for spotters 1,2,5 and 6 (Model 1) be used in the next data exchange for this index.

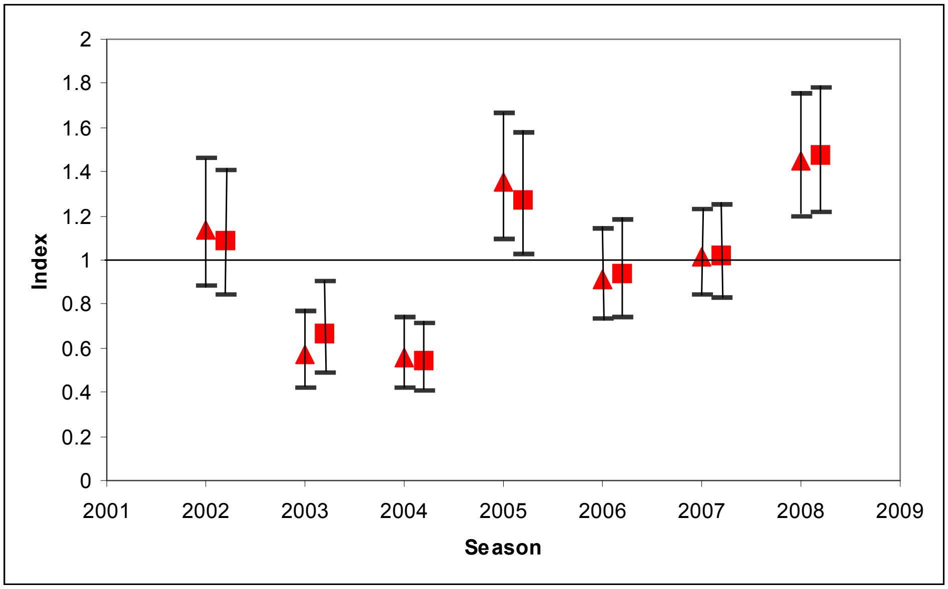


Figure 8. Estimates of standardised relative surface abundance, scaled to the mean over the period, for models with companies 1,2, 5 and 6 for (i) swell included as a covariate (triangles) and (ii) swell excluded as a covariate (squares). All months were included (December – March). The median and exp(predicted value + or – 2 standard errors) are shown. Values are scaled to the mean over the period, so the horizontal line at 1 indicates the mean. 'Season' is indicated by the second year in a split year so that, e.g. 2002 implies the 2001/2002 season.

Table 6. Standardised SAPUE index of juvenile SBT in the GAB for two models including different combinations of spotters (see main text). Season refers to the second year in a split year, i.e. 2002 = the 2001/2002 season. The estimated values are also illustrated in Figure 9 above.

Season	Model 1: spotters 1,2,5,6 (swell excluded)		Model 2: spotters 1 and 6 (swell excluded)	
	Estimate	SE	Estimate	SE
2002	1.14	0.144	1.09	0.140
2003	0.57	0.085	0.66	0.102
2004	0.56	0.079	0.54	0.076
2005	1.35	0.141	1.27	0.136
2006	0.91	0.101	0.94	0.109
2007	1.02	0.096	1.02	0.104
2008	1.45	0.137	1.47	0.140

Summary

We present results of a standardised 'surface abundance per unit effort' (SAPUE) index, based on fitting a general linear model to the data. The model does not have any interaction terms, although past analyses suggested that an interaction between spotter and season is important. Due to the changes in spotter effort in the 2006 and subsequent seasons, the dataset has become unbalanced, making it difficult to obtain a reliable index of abundance for the model with interaction between spotter and season. We considered the sensitivity of results to different combinations of spotters in the analysis, and again to the inclusion or exclusion of 'swell' as a covariate. The reason for this is that one of the spotters has a large number of missing values for this covariate, reducing the dataset if these records are excluded.

The index values for 2003 and 2004 are somewhat sensitive to the choice of spotters included in the model, though the general temporal patterns, particularly in recent years, are not sensitive to the choice of spotters. The estimated index is lowest in 2003 and 2004 (Figure 7). The estimates for 2006 and 2007 are both close to, or slightly above, the average over the past 7 seasons, and the 2008 is above average. The 2008 estimate is the highest of the commercial spotting series (2002-2008), but it is important to consider this in the context of the much longer aerial survey index of abundance (CCSBT-ESC/0809/24). The aerial survey index, which shows a similar pattern for the period of overlap (2005-2008) with the SAPUE index, still estimates the 2008 index as being below the long term average (1993-2008).

Although we tabulate the percentage days on which no biomass was spotted (a ‘presence-absence’ indicator) we consider this to be inappropriate and unreliable as an index of juvenile abundance in the GAB. Note for example that in 2008 no SBT schools were observed on 19% of the days on which spotting occurred. This is the second highest percentage in the time series, and yet the standardised index of SAPUE in 2008 is the highest of the time series. The fact that the pattern of the SAPUE index over the period 2005-2008 is similar to that of the line transect Aerial survey (also see below), supports the argument that a presence-absence type of indicator is inadequate.

The index reflects the abundance of 2, 3 and 4 year olds combined. The two low years would therefore represent the 1999, 2000 and 2001 year-classes (as 4,3,2-year olds in 2003) and the 2000, 2001 and 2002 year classes (as 4,3,2-year olds in 2004). In 2005, there appeared to be many 1-year olds in the bight. This was noticed by industry and mentioned to us, but it was also apparent through the relatively large number of below 10kg fish that were sampled for length from the farming operations. It is unclear and unknown whether the index in 2005 reflects a substantial proportion of age 1 fish or not, compared to other years. It is also well known that not all juveniles spend their summers in the GAB. Unfortunately, there is no direct information about the proportion of the total juvenile population in the GAB each year. This is not a major problem if the proportion has remained approximately constant over time. If, however, there have been substantial changes in the proportion (e.g. through changes in movement dynamics) then it becomes more difficult to know how to interpret this index.

There are now four years of overlap between the SAPUE index and the line-transect aerial survey index (see the update this year in CCSBT/ESC/0809/24). It is encouraging that the overall patterns of the two indices are similar for the four years (2005-2008). Direct comparison is still, however, difficult and should be done with caution. Most importantly, the commercial spotting data are obtained in a substantially different way directly associated with the fishing operation, and covers a much smaller spatial area than the line-transect survey. The changes in the number of spotters, their relative amount and timing of their effort is making standardisation increasingly difficult. We still consider the line-transect aerial survey to be preferable as an approach to an index of juvenile abundance, compared to the commercial spotting.

Acknowledgements

We again especially thank the commercial spotters and pilots for their continued willingness to collect and record sightings data each fishing season. We also thank the tuna fishing companies in Port Lincoln for their continued support of the project. This study was funded by AFMA, DAFF, Australian Industry, and CSIRO’s Wealth from Ocean Flagship.

Appendix

Table A 1. Estimates of coefficients, standard errors and related ‘significance’ quantities for two models (excluding swell as a covariate).

Spotters 1,2,5,6

```
Call:
glm(formula = biomass ~ as.factor(season) + as.factor(spotter) +
    as.factor(month) + wind + spotcon + cloud + temperature +
    moonillum + offset(log(SearchEffort)), family = mvb.tweedie(1.5,
    0), data = tdat2, subset = (spotter != 3))

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-10.536   -4.282   -1.340    1.296   18.526

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)             0.541626   0.393408   1.377 0.169111
as.factor(season)2003 -0.690779   0.215353  -3.208 0.001411 **
as.factor(season)2004 -0.707639   0.201292  -3.515 0.000473 ***
as.factor(season)2005  0.173569   0.185048   0.938 0.348646
as.factor(season)2006 -0.218562   0.184881  -1.182 0.237613
as.factor(season)2007 -0.112391   0.171951  -0.654 0.513610
as.factor(season)2008  0.243939   0.178162   1.369 0.171460
as.factor(spotter)2   -1.627209   0.163151  -9.974  < 2e-16 ***
as.factor(spotter)5    0.115570   0.157149   0.735 0.462380
as.factor(spotter)6   -0.680821   0.109470  -6.219 9.47e-10 ***
as.factor(month)2     -0.303399   0.116817  -2.597 0.009633 **
as.factor(month)3     -0.897781   0.129179  -6.950 9.72e-12 ***
as.factor(month)12     0.258533   0.130443   1.982 0.047948 *
wind                  -0.124943   0.019906  -6.277 6.72e-10 ***
spotcon                0.340168   0.077636   4.382 1.40e-05 ***
cloud                 -0.035968   0.019843  -1.813 0.070407 .
temperature            0.032239   0.007374   4.372 1.46e-05 ***
moonillum             -0.161563   0.126324  -1.279 0.201416
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Tweedie family taken to be 24.48062)

    Null deviance: 29215  on 606  degrees of freedom
Residual deviance: 11856  on 589  degrees of freedom
AIC: 7813.5

Number of Fisher Scoring iterations: 5
```

Spotters 1 and 6 only

```
Call:
glm(formula = biomass ~ as.factor(season) + as.factor(spotter) +
    as.factor(month) + wind + spotcon + cloud + temperature +
    moonillum + offset(log(SearchEffort)), family = mvb.tweedie(1.5,
    0), data = workdat08, subset = (spotter != 2 & spotter !=
    3 & spotter != 5))

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-10.455   -4.289   -1.192    1.389   16.556

Coefficients:
```

```
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)              0.603314   0.408934   1.475 0.140833
as.factor(season)2003 -0.495677   0.220742  -2.246 0.025228 *
as.factor(season)2004 -0.699861   0.200911  -3.483 0.000544 ***
as.factor(season)2005  0.155838   0.185217   0.841 0.400587
as.factor(season)2006 -0.149127   0.188864  -0.790 0.430184
as.factor(season)2007 -0.066931   0.176880  -0.378 0.705315
as.factor(season)2008  0.301834   0.179125   1.685 0.092684 .
as.factor(spotter)6   -0.701880   0.104374  -6.725 5.44e-11 ***
as.factor(month)2     -0.200980   0.125897  -1.596 0.111116
as.factor(month)3     -0.787038   0.132589  -5.936 5.90e-09 ***
as.factor(month)12     0.302362   0.129442   2.336 0.019942 *
wind                  -0.132843   0.021174  -6.274 8.39e-10 ***
spotcon                0.294472   0.084520   3.484 0.000543 ***
cloud                 -0.036205   0.020077  -1.803 0.072012 .
temperature            0.033580   0.007906   4.247 2.64e-05 ***
moonillum             -0.144832   0.132001  -1.097 0.273148
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Tweedie family taken to be 21.38009)

    Null deviance: 20692.9  on 458  degrees of freedom
Residual deviance:  8602.5  on 443  degrees of freedom
AIC: 6009.4

Number of Fisher Scoring iterations: 6
```

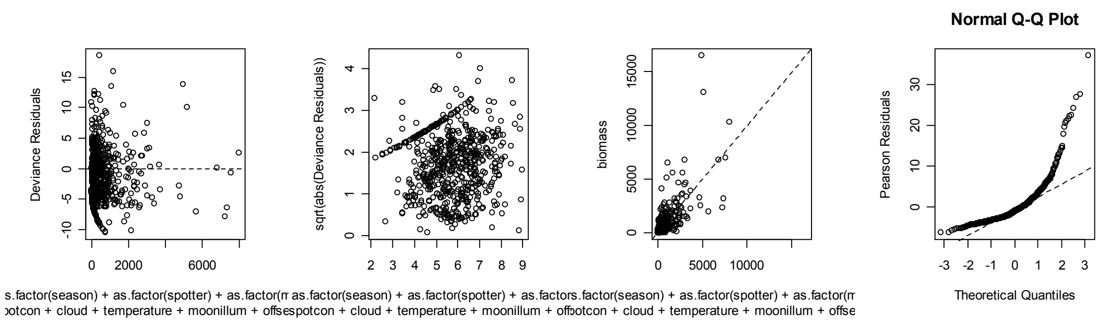


Figure A 1. Diagnostics for the model with spotters 1,2,5,6 and swell excluded. The x-axis text on the first and third panels is the call to the model indicating that predicted values are being plotted.

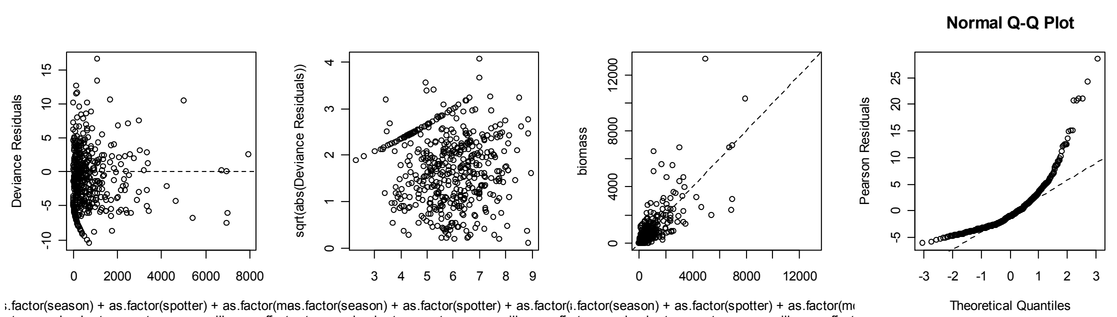


Figure A 2. Diagnostics for the model with spotters 1 and 6 only, and swell excluded. The x-axis text on the first and third panels is the call to the model indicating that predicted values are being plotted.

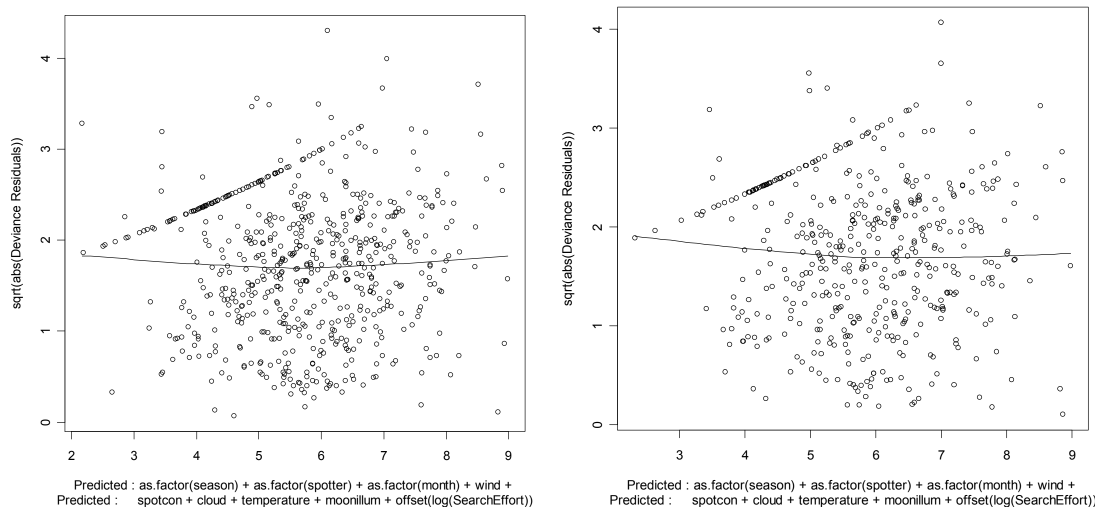


Figure A 3. Deviance residuals (square root of the absolute values) plotted against fitted values for the model with (left panel) spotters 1,2,5,6 and (right panel) spotters 1 and 6 with Tweedie parameter Φ=1.5. In both cases swell is excluded. The solid line is a loess smooth fitted through the data. See main text for details.