Use of TVH Logbook Data to construct an Annual Abundance Index for Torres Strait Rock Lobster – 2016 Update

Robert Campbell, Darren Dennis, Eva Plaganyi, Roy Deng

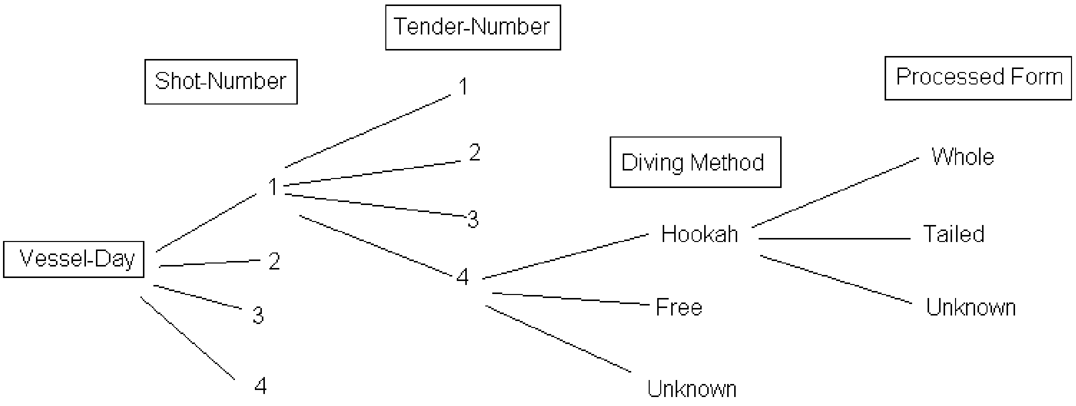
CSIRO Oceans and Atmosphere

December 2016

1. TVH Data

Logbook data obtained from AFMA consists of 96,215 individual catch records for the TVH rock-lobster fishery for the 23 years from 1994 to 2016. The structure of the data is shown in Figure 1. For each vessel-day there can be multiple shots (up to 4) with each shot consisting of up to 8 tenders. Each tender has a catch recorded by diving method (hookah, free or unknown) and the catch is recorded by processed form (whole, tailed or unknown). The data was aggregated so that each record refers to the catch for a unique vessel-day, shot, tender and diving method. This gave 66,366 records.

Figure 1. Structure of the TVH data



The distribution of these 66,366 catch records by year and month, diving method, processed state of catch and MSE-area are given in Tables 1-3. It is apparent that there has been little if any effort during October and November before 2006 and since 2006 there has been zero effort in the months October-to-January. As such the analysis was limited to the 8 months between February and September. Similarly the analysis was also limited to those records with a known MSE-area (i.e. areas designated 0 and -99 were excluded) though areas 201 and 202 were combined (to provide a better data coverage, and designated as area 110) and area 401 (GBR) was also excluded.

In the past CPUE has been recorded as the catch-per-tender-set. However, as there can be multiple shots-per-day the duration of a tender-set can obviously vary and each tender-set cannot be assumed to be equivalent to a tender-day. The catch data also contains a field “Hours-Fished” which records the duration of the fishing trip for each tender-set and this was deemed to be a better measure of tender effort than assuming

Table 1. Number of TVH catch records by year and month.

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	TOTAL
1994	84	105	236	448	347	364	227	310	270			54	2445
1995	23	116	123	147	185	220	121	239	238	3		220	1635
1996	366	237	447	247	378	264	356	517	411			324	3547
1997	383	232	307	239	598	333	438	538	327	18		598	4011
1998	445	739	551	484	486	587	553	603	493		9	231	5181
1999	117	98	262	242	208	214	161	132	146			235	1815
2000	196	240	349	215	328	370	342	232	99		66	274	2711
2001	375	97	223	65	259	270	206	174	119	9	1	87	1885
2002	26	285	365	295	401	400	360	492	398			89	3111
2003	100	461	488	393	490	518	527	596	413			176	4162
2004	24	607	712	571	662	761	729	633	395			106	5200
2005	13	662	615	543	519	538	552	533	323			4	4302
2006		409	436	361	286	206	349	289	92				2428
2007		288	427	446	542	489	402	184	91				2869
2008		133	222	113	161	96	159	175	152				1211
2009		148	227	174	201	200	125	163	70				1308
2010		255	333	302	324	292	309	294	253		6		2368
2011		286	384	371	322	380	356	310	261				2670
2012		166	344	371	311	336	318	264	201				2311
2013		461	383	414	424	324	374	385	243				3008
2014		357	395	297	433	408	445	274	291		1		2901
2015		419	408	441	355	313	253	357	137				2683
2016	12	500	444	315	379	334	313	183	124				2604
Total	2,164	7,301	8,681	7,494	8,599	8,217	7,975	7,877	5,547	30	83	2,398	66,366

Table 2. Annual number of TVH catch records by diving method and TVH catch by processed state.

Diving Method			Total	Catch by Processed State (kg)			Total		
Hookah	Free	Unknown	Records	Tails	Whole	Unknown	Catch	%Tails	%Whole
1,505	136	804	2,445	123,006	0	0	123,006	100.0%	0.0%
947	59	629	1,635	100,407	635	0	101,042	99.4%	0.6%
1,609	87	1,851	3,547	219,045	7,810	0	226,855	96.6%	3.4%
1,890	112	2,009	4,011	273,151	1,880	8	275,039	99.3%	0.7%
2,681	169	2,331	5,181	310,635	18,922	0	329,557	94.3%	5.7%
1,412	38	365	1,815	88,416	6,681	0	95,097	93.0%	7.0%
2,330	114	267	2,711	118,824	10,038	0	128,862	92.2%	7.8%
812	26	1,047	1,885	66,347	2,729	0	69,076	96.0%	4.0%
1,721	10	1,380	3,111	108,216	39,471	0	147,687	73.3%	26.7%
3,958	104	100	4,162	255,447	105,964	0	361,411	70.7%	29.3%
5,045	154	1	5,200	317,467	163,651	0	481,118	66.0%	34.0%
4,101	199	2	4,302	484,497	60,480	0	544,977	88.9%	11.1%
2,307	119	2	2,428	108,909	26,539	0	135,448	80.4%	19.6%
2,829	39	1	2,869	207,463	61,133	0	268,596	77.2%	22.8%
1,205	6	0	1,211	63,378	37,060	0	100,438	63.1%	36.9%
1,281	27	0	1,308	51,322	39,729	10	91,061	56.4%	43.6%
2,356	12	0	2,368	67,817	214,797	0	282,614	24.0%	76.0%
2,668	1	1	2,670	171,469	332,064	0	503,533	34.1%	65.9%
2,311	0	0	2,311	65,282	305,198	2	370,482	17.6%	82.4%
3,006	2	0	3,008	61,631	300,030	0	361,661	17.0%	83.0%
2,901	0	0	2,901	42,054	230,511	120	272,685	15.4%	84.5%
2,678	1	4	2,683	22,479	130,231	0	152,710	14.7%	85.3%
2,592	12	0	2,604	41,702	195,911	0	237,613	17.6%	82.4%
54,145	1,427	10,794	66,366	3,368,964	2,291,464	140	5,660,568	59.5%	40.5%

Table 3. Number of TVH catch records by MSE-area.

	Northern	Mabuiag	Badu	Thurs Is.	Central	Warrior	Warraber	Kirkaldie	Adolphus	East TS	East TS	GBR	East Coast		
YEAR	A101	A102	A103	A104	A105	A106	A107	A108	A109	A201	A202	A401	A0	A-99	TOTAL
1994	51	257		11	119	252	926	64	89	106	177	1		392	2445
1995	106	289	2	41	83	187	487	111	26	36	32	4	8	223	1635
1996	620	1152	2	11	51	269	719	41	37	1	32		4	608	3547
1997	425	1324	21	21	73	524	881	4	21	52	33	2		630	4011
1998	463	1681	51	130	107	661	1042	160	16	31	45			794	5181
1999	158	457	34	33	66	254	348	177	17	14	30	15		212	1815
2000	137	252	66	48	51	825	605	229	59	7	22	35	5	370	2711
2001	42	70	5	44	26	712	366	83	40	3	41	44	4	405	1885
2002	107	278	18	176	44	692	592	718	48		17	16	4	401	3111
2003	808	719	115	317	344	404	432	832	96	7	49	3	3	33	4162
2004	921	765	209	163	551	344	980	970	205	11	58	4	10	9	5200
2005	682	588	164	196	164	203	511	1680	90	3	18	1	2		4302
2006	301	332	21	130	187	300	440	355	276	34	48	4			2428
2007	362	417	42	146	134	323	367	980	62	10	24	2			2869
2008	227	63	6	91	53	238	240	206	48	2	31	3	1	2	1211
2009	272	42	5	80	145	371	231	47	26	23	59	7			1308
2010	493	138	101	102	31	197	206	997	43	12	32	14	2		2368
2011	389	111	34	83	17	159	430	1406	25		14		2		2670
2012	417	217		14	46	155	1166	267	18	5	5	1			2311
2013	719	239	34	16	63	168	469	1267	6	6	21				3008
2014	777	263	15	27	165	268	780	445	47	14	93		7		2901
2015	176	173	45	5	116	876	660	486	25		121				2683
2016	57	12	62	7	202	681	454	914	18	131	46		20		2604
Total	8,710	9,839	1,052	1,892	2,838	9,063	13,332	12,439	1,338	508	1,048	156	72	4,079	66,366

Figure 2. The total number of TVH catch records each year and the number of records for which the corresponding effort data is available. The percentage of records for which no effort is recorded is also shown (right hand axis).

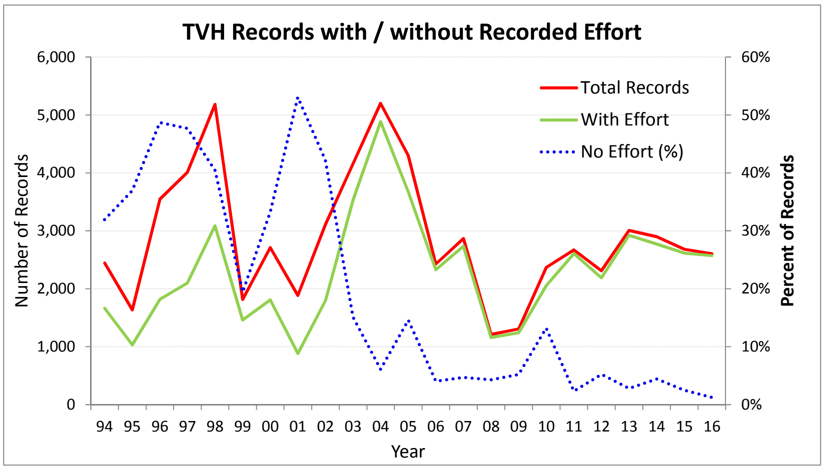


Figure 3. The percent of total TVH catch each year (a) caught by each fishing method, and (b) landed as Tails or Whole weight.

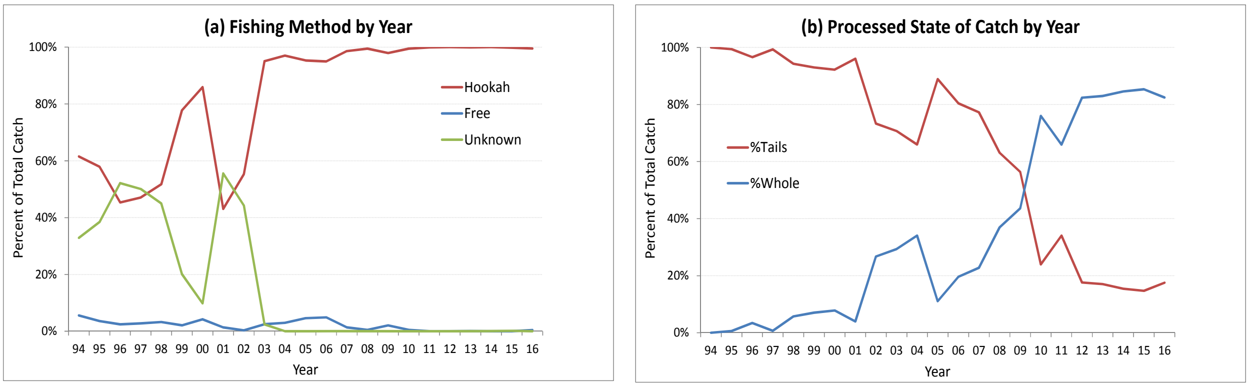


Figure 4. Distribution of (a) effort, (b) catch and (c) CPUE for the 44,958 records for which effort was recorded on TVH logbooks.

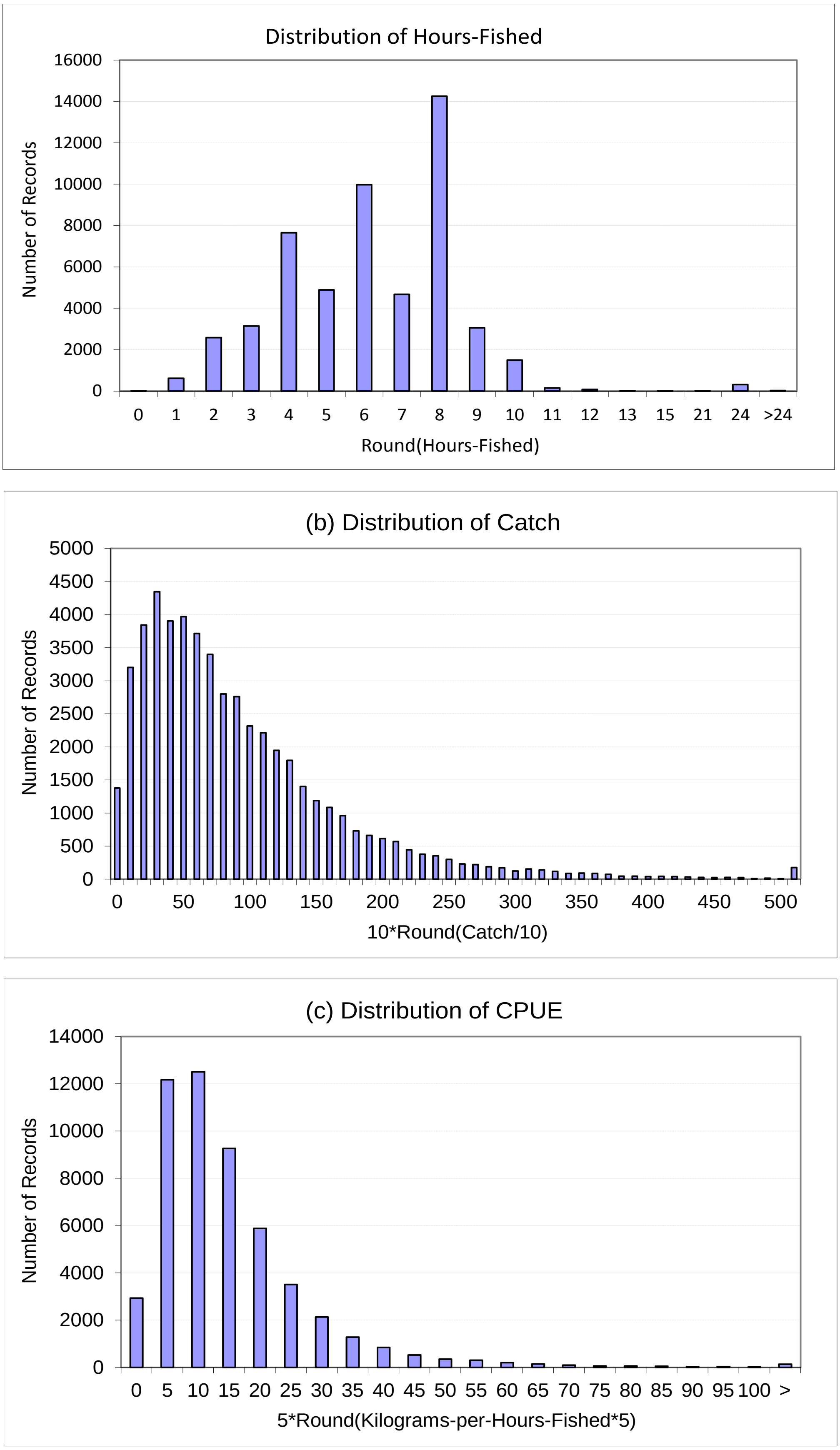
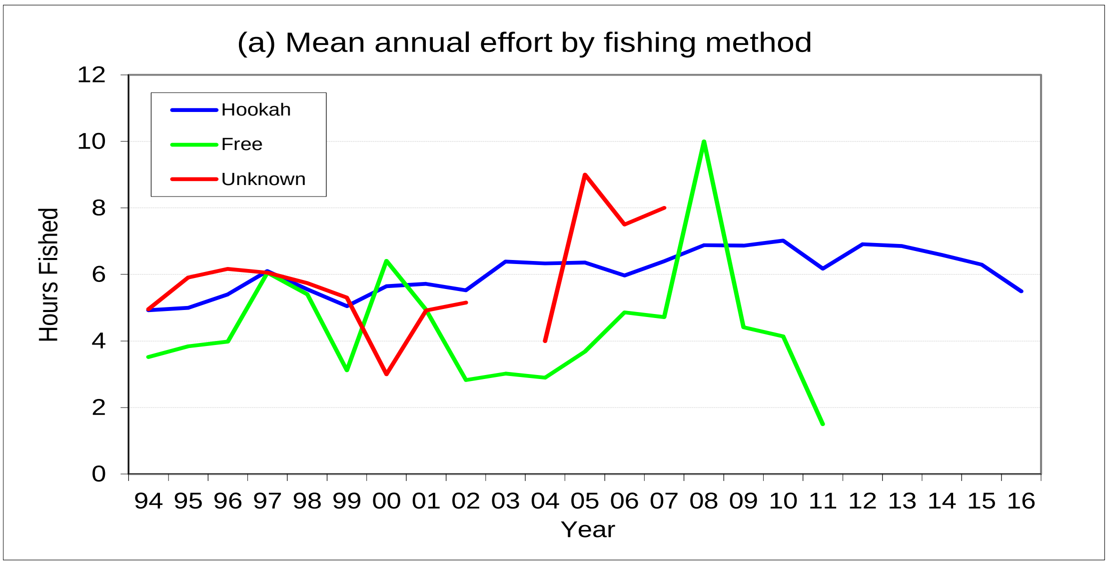
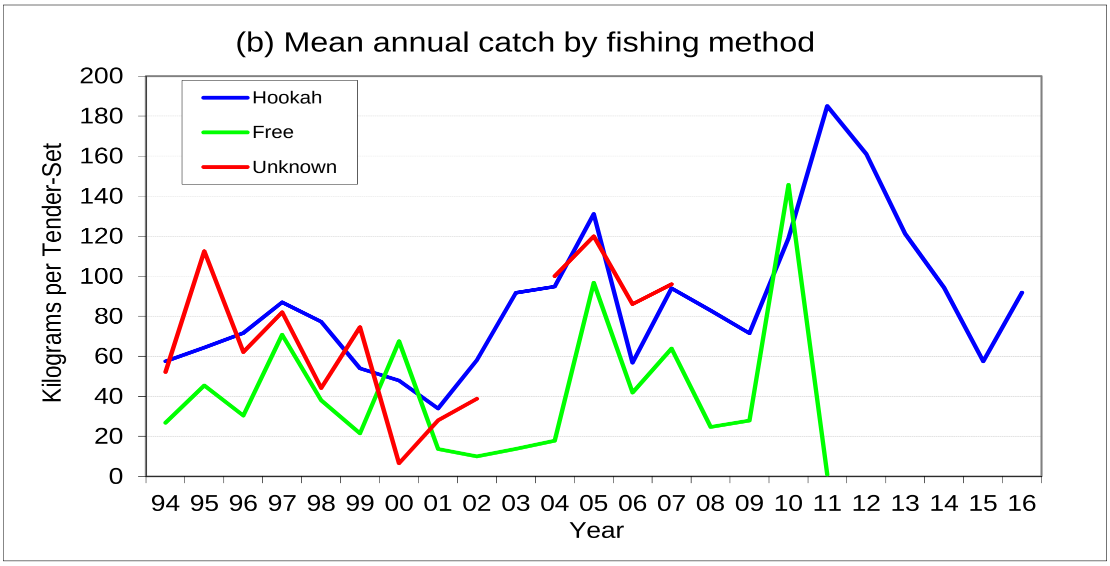
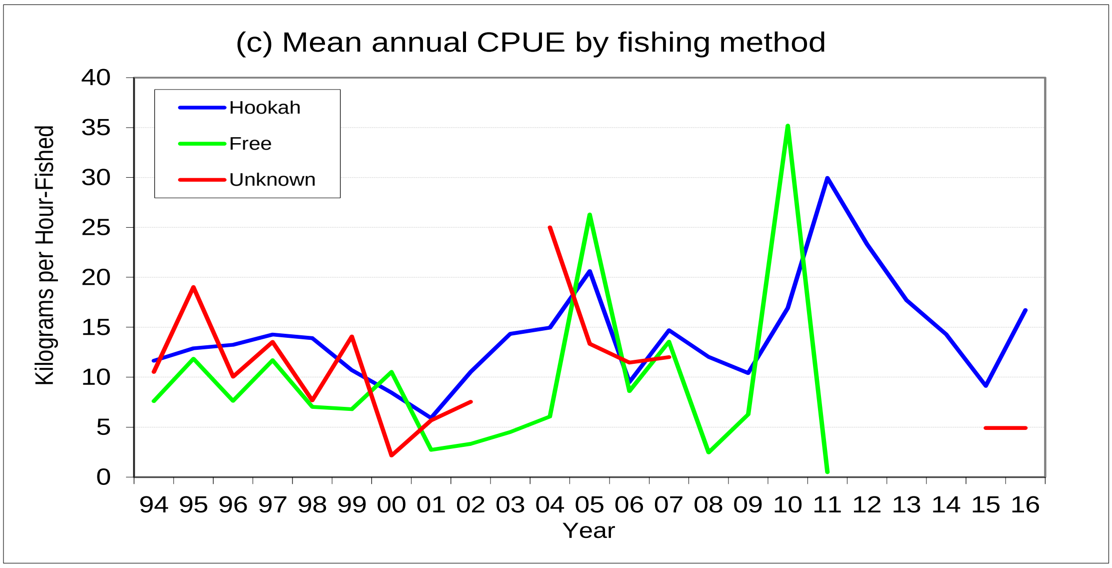


Figure 5. Mean (a) effort, (b) catch and (c) CPUE by fishing method and year for the 48,030 unique vessel-day, shot, tender and diving method records for which this effort was between 0 and 12 hours and areas and months restricted as described in the text.







each tender-set is equivalent to a day's effort. Unfortunately this field has not been completed for all tender-sets (c.f. Figure 2b), with the number of hours fished recorded for only 52,944 (79.8%) of the 66,366 records. The distribution of hours fished for these records is shown in Figure 4. The number of recorded hours fished was between 0.15 hours and 96 hours, though was less than12 hours for 99.3% of all records. All records where the recorded hours-fished was greater than 12 hours were considered suspect due to possible recording errors and as such only those records where the hours-fished was 12 hours or less were included in the analysis. The five records where effort was less than 0.5 hours were also excluded. Note, the number of hours fished was recorded as 24 hours for 315 records and was assumed to represent a "day's" fishing.

After applying each of the following filters to the data:

- Exclude MSE-areas 0, 401 and -99
- Exclude Month<2 and Month>9
- Exclude Hours-Fished less than 0.5 hour and greater than 12 hours

the number records included in the data for further analysis was reduced to 48,030. The mean (a) effort, (b) catch and (c) CPUE by fishing method and year for these records are shown in Figure 5.

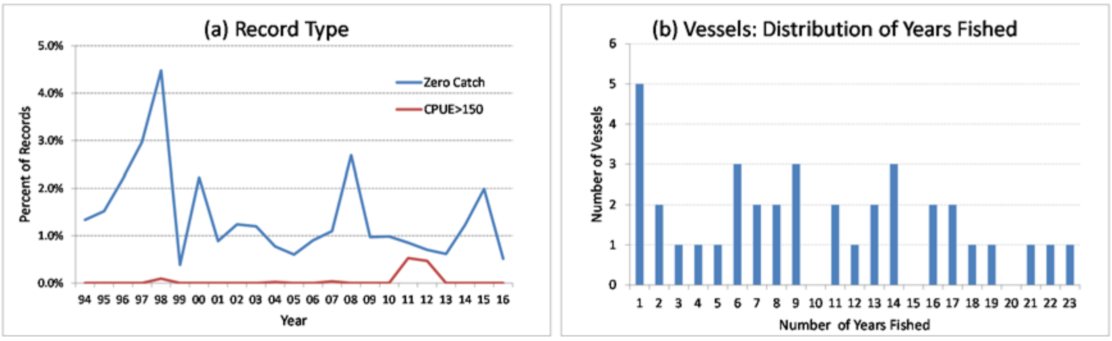
2. GLM Analysis

i) Fitted Data

Of the 48,030 records selected above for analysis it was noted that there were a small percentage of records (617 or 1.3%) where the catch was zero. The inclusion of such records in the GLM analyses can cause problems. The percentage of such records each year is shown in Figure 5a and varies from 4.48% in 1998 to 0.39% in 1999. Nevertheless, apart from the relatively high values in 1997 and 1998 there does not appear to be a trend in the percentage of zero catches in the data over time. As such, and as recommended for the analyses undertaken previously, these zero catch records were excluded from the analyses. Note, to retain the zero-catch records in the analysis a two-stage analysis of the data can be undertaken where one first models the probability of obtaining a positive catch following by a separate analysis where one models the size of the positive catch. The results of each analysis can then be combined to obtain the required standardised CPUE index. Such an approach was not considered appropriate for this data due to the small percentage of zero-catch records in the data.

Further inspection of the data also indicated a number of records having a very high CPUE (kilograms of catch per hour fished) value and which could be considered outliers in the data due to errors in either the recording of the catch or effort. To exclude this possibility the 27 records having a CPUE>150 kgs/hour were deleted from the data (cf. Figure 6a). Finally, due to the observation that Vessel-Names and Vessel-Symbols are not always matched (likely due to the switching of licences between vessels) a combination of Vessel-Name and Vessel-Symbol was adopted to identify vessels in the data. Of the 88 vessels identified in this manner in the selected data, only the data pertaining to the 46 vessels which had fished for 3 or more years and for which there were more than 50 data records were included in the analysed data (c.f. Figure 6b). Combined with the other two filters the total number of records remaining in the data for analysis was 41,882.

Figure 6. (a) Percentage of records in the data, by year, where either the catch is zero, or the CPUE>150 kg/hour, and (b) histogram of the number of vessels (distinguished by vessel symbol) by the number of years they have fished in the fishery.



The number of *Area-Month* strata fished each year and the number of vessels fishing each year in the data selected for inclusion in the GLM analyses is shown in Figure 7 while a bubble plot displaying the number of observations for each vessel each year in this data is shown in Figure 8. A summary of the number of observations and nominal CPUE (kilograms per hour) within each *Year*Area*, *Year*Month* and *Area*Month* strata is provided in the Appendix.

Figure 7. (a) Number of *Area-Month* strata fished each year and (b) the number of vessels fishing each year in the data selected for inclusion in the GLM analyses.

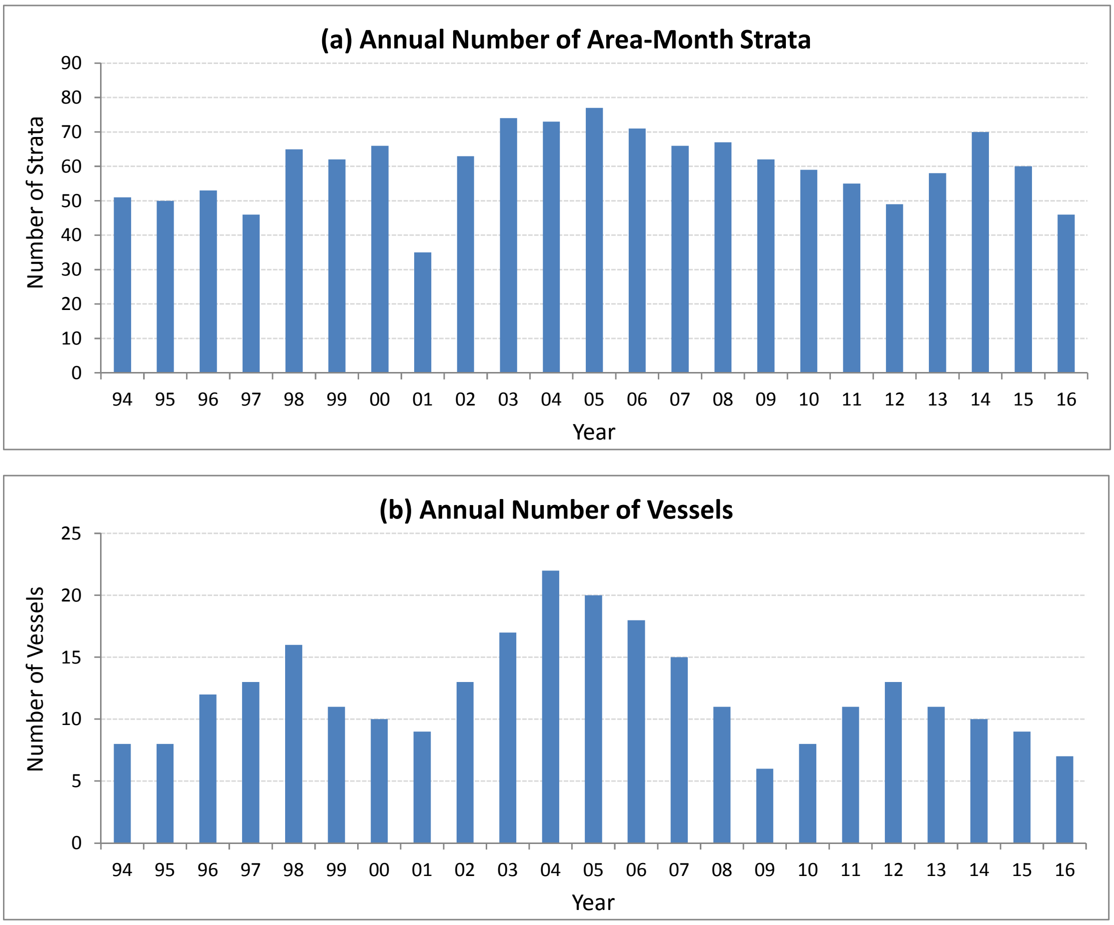
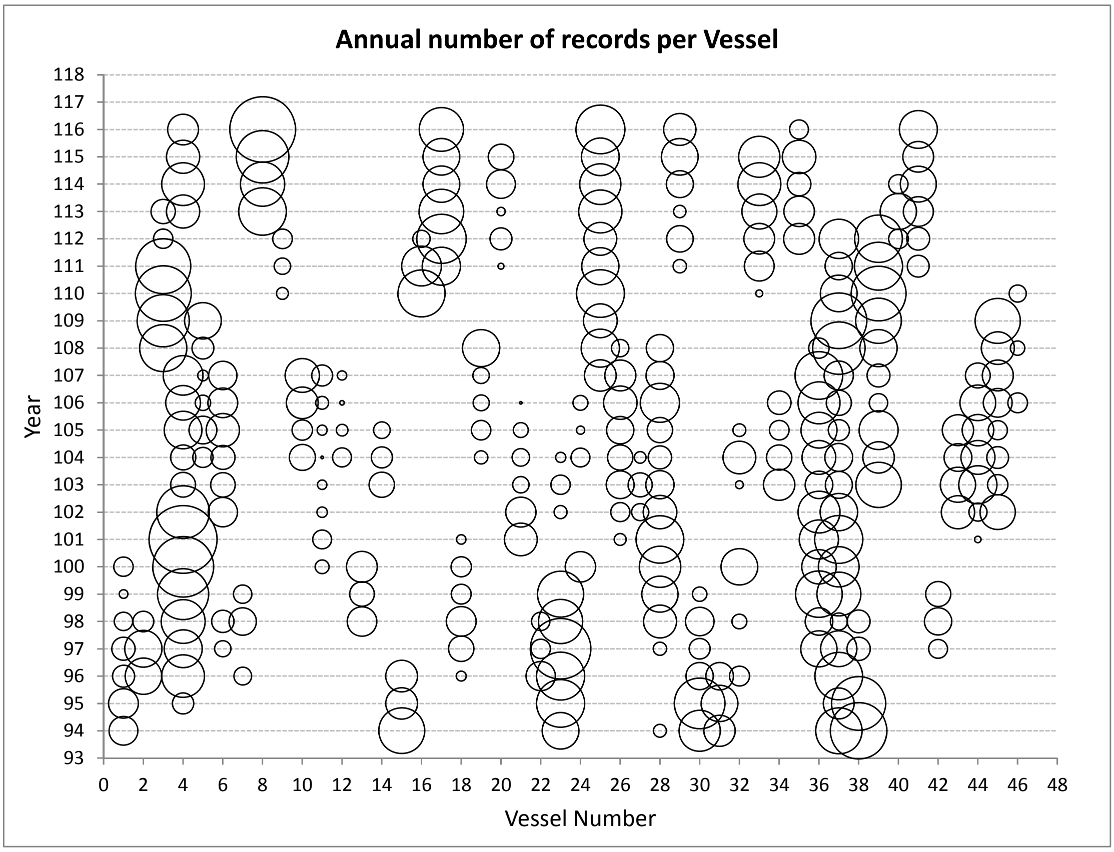


Figure 8. Bubble plot displaying the number of observations for each vessel each year in the data selected for inclusion in the GLM analyses.



ii) GLM Models

Several different General Linear Models (GLMs) were adopted for analysing the data in order to obtain a standardised index of stock abundance in each year.

Main Effects Model

In order to explore the impact of each fitted effect, the first set of analyses were based on the following model where no interactions between main effects were included:

$$\begin{aligned} CPUE &= Intercept + Year + Month + Area + Vessel + Fishing\text{-}Method \\ &\quad + Proportion\ of\ Catch\ Landed\ as\ Tails + Southern\ Oscillation\ Index \\ &\quad \text{/ distribution = gamma, link = log} \end{aligned}$$

$$= I + Y + M + A + V + F + P + SOI \text{ / dist= gamma, link=log}$$

The SAS GENMOD procedure was used to fit the model. All effects *Year, Month, Area*, *Vessel* and *Method* (Hookah, Free and Unknown) were fitted as class variables except for the SOI index which was fitted as a continuous variable. The *Proportion-Tails* was also fitted as a class variable with each record classified as one of the following five levels: (<20%, 20% to <40%, 40% to <60%, 60% to <80%, >=80%). A log-gamma distribution was assumed for the distribution of CPUE values. The annual index and abundance was determined using the method described in the section below.

For each of the main effects, a measure of the impact of each level on the modelled CPUE was obtained by taking the exponent of the estimated parameter for each level. The impact of each level was then compared to the impact of a reference level. For each main effect these reference levels were:

Month	September
Area	Eastern Torres Strait
Method	Hookah diving
Vessel	Vessel with the largest number of records
Proportion-tails	>80%

Finally, the annual influence of each of the main effects on the resulting index of abundance was calculated using the method described in Bentley et al (2012).

As shown in Campbell (2004) a bias in the annual abundance index can result when there is an unequal number of observations within each spatial-temporal strata used for calculating the abundance index. In order to overcome this problem a weighting of the observations needs to be incorporated when fitting the data to the GLM. Each observation was therefore weighted such that the sum of the weights for all observations in each of the *Year-Month-Area* strata was the same for all strata. Furthermore, in order to account for the weighting given each observation in determination of the annual influence of each main effect the sum of the weights for all observation within a given level was used instead of just the number of observations.

Interactions Models

The second set of analyses was undertaken in order to explore whether the inclusion of 2-way interactions between the main spatial-temporal effects improved the model fit to the data. Specifically, the following five models were examined:

Int-1:
*CPUE = Intercept + Year +Month + Month*Area*
+ Vessel +Fishing-Method + Proportion-Tails + SOI
/ distribution = gamma, link = log

Int-2A:
*CPUE = Intercept + Year*Month + Month*Area*
+ Vessel +Fishing-Method + Proportion-Tails + SOI
/ distribution = gamma, link = log

Int-2B:
*CPUE = Intercept + Year*Area + Month*Area*
+ Vessel +Fishing-Method + Proportion-Tails + SOI
/ distribution = gamma, link = log

Int-2C:
*CPUE = Intercept + Year*Month +Year*Area*
+ Vessel +Fishing-Method + Proportion-Tails + SOI
/ distribution = gamma, link = log

Int-3:
*CPUE = Intercept + Year*Month +Year*Area + Month*Area*
+ Vessel +Fishing-Method + Proportion-Tails + SOI
/ distribution = gamma, link = log

where * indicates an interaction between the related effects. The inclusion in these 2-way interactions allows for the relative distribution of the resource between the different areas and months to be different between years.

ii) Derivation of Annual Index

Using the results from each GLM an annual abundance index was constructed based on the standardised CPUE.

For the model which included the three 2-way interactions the standardised CPUE within each Year-Month-Area strata was calculated as follows:

$$stdCPUE(year = y, month = m, area = a) =$$
$$\exp(I + Y.M_{ym} + Y.A_{ya} + M.A_{ma} + F_h + V_{ref} + P_{ref})$$

where $Y.M_{ym}$, $Y.A_{ya}$, $M.A_{ma}$, F_h, V_{ref} and P_{ref} are the parameters estimates relating to each of the terms included in the model. Note, due to the over-parameterization inherent in the GLM both F_h=0, V_{ref} =0 and P_{ref}=0 as these respectfully to relate the last levels in each of the *Fishing-Method*, *Vessel* and *Proportion-Tails* factors included in the model. In total there are 1840 (=23 years x 8 months x 10 areas) *Year-Month-Area* strata. As the standardised-CPUE is taken as an index of the density of fish within each strata, an index of the abundance of lobsters across the fishery in each year and month is given by:

$$Index(year = y, month = m) = \frac{1}{\sum_{a=1}^{NA} Area_a} \sum_{a=1}^{NA} Area_a.stdCPUE(y, m, a)$$

where $Area_a$ is the spatial size of each of the *NA Area* effects included in the GLM. Finally, an index of abundance for each year can be obtained by taking the average of the *NM* monthly indices in each year.

$$Index(year = y) = \frac{1}{NM} \sum_{m=1}^{NM} \left[\frac{1}{\sum_{a=1}^{NA} Area_a} \sum_{a=1}^{NA} Area_a.stdCPUE(y, m, a) \right]$$

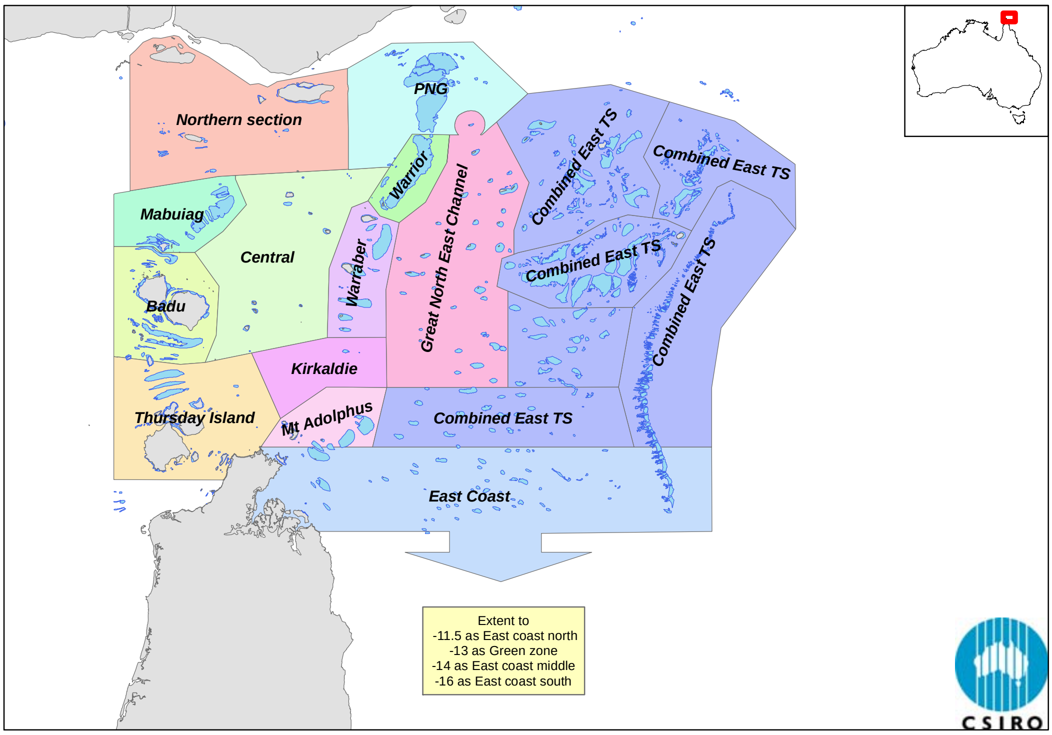
Finally, a relative annual abundance index, B_y, was calculated such that the mean index over all years equals 1, i..e:

$$B_y = \frac{Index(year = y)}{\frac{1}{NY} \sum_{i=1}^{NY} Index(year = i)}$$

Two different sets of spatial sizes, $Area_a$, were used in calculating the above abundance index. These were:

1. The total spatial size of the each MSE area shown in Figure 9.
2. The spatial extent of each MSE area which had been fished between 1994 and 2013. This was based on the number of 0.1x0.1-degree squares in which a fishing operation had been reported in each area during this period. For those squares which included more than one MSE area, the square was apportioned between the different areas based in the total number of records in each area.

Figure 9. Map of the MSE regions used as the area effects in the GLM.



In order to ascertain the spatial size of each MSE area used in the GLM-analysis, the number of 0.1x0.1-degree squares fished (based on the location of the mother ship recorded in the TVH logbook) within each region was determined for each year (c.f. Table 4). Across the entire Torres-Strait region the number of squares fished each year has varied between 31 (in 1995) and 101 (in 2004). Across all years, the maximum and mean number of squares fished within each area was determined together with the number of unique squares fished and the spatial size of each area in 10,000 hectares. Each size metric for each area was then expressed as a percentage of the combined total across all areas. These calculations are shown in Table 4 and displayed in Figure 10. For each area the relative sizes based on the maximum and mean number of squares fished are similar and for the GLM analysis the size of each area was taken to be the mean of these two metrics (see GLM area in Table 4).

The derivation of the abundance index based on the GLMs which included less than three 2-way interaction terms is similar to that shown above. However, it can be noted that for those models which do not included an interaction with the Year effect (i.e. the main effects and Int-1 models), the relative abundance index, B_y, reduces to the simpler form:

$$B_y = \frac{\exp(Y_Y)}{\frac{1}{NY}\sum_{i=1}^{NY}\exp(Y_i)}$$

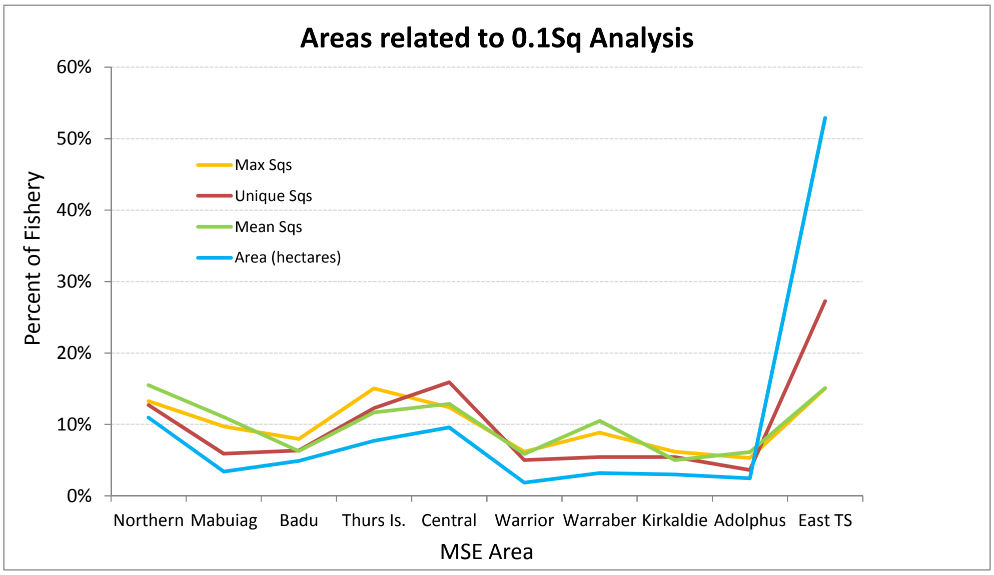
where Y_i , $i=1, NY$ are the parameters estimates relating to *NY Year* effects included in the model. In these situations the abundance is independent of the relative size of each *Area* effect included in the GLM.

Table 4. Number of 0.1x0.1-degree squares fished (based on location of mother ship) within each MSE areas used in the GLMs fitted to the TVH data.

	N_01	N_02	N_03	N_04	N_05	N_06	N_07	N_08	N_09	N_10	
Year	Northern	Mabuiag	Badu	Thurs Is.	Central	Warrior	Warraber	Kirkaldie	Adolphus	East TS	Total
1994	2	6	0	1	5	1	3	1	1	14	34
1995	4	5	1	2	3	1	4	1	1	9	31
1996	5	6	1	3	3	2	3	1	1	7	32
1997	4	6	5	8	2	2	4	1	1	17	50
1998	5	6	5	6	5	2	4	1	3	13	50
1999	4	6	5	4	3	2	4	1	2	14	45
2000	6	6	4	9	3	2	4	1	2	6	43
2001	4	4	2	5	3	2	5	1	3	4	33
2002	4	5	4	8	3	2	4	3	2	3	38
2003	12	8	7	17	14	7	8	7	6	4	90
2004	14	11	9	12	14	7	10	4	6	14	101
2005	13	10	7	14	13	5	10	6	5	6	89
2006	15	10	5	10	14	5	6	4	5	15	89
2007	13	10	4	9	12	5	5	4	4	8	74
2008	12	6	3	5	9	4	7	2	4	9	61
2009	15	4	2	6	8	6	6	3	4	10	64
2010	11	4	5	9	6	2	4	5	6	6	58
2011	8	3	3	4	2	2	6	3	3	5	39
2012	13	6	0	5	6	2	10	1	4	3	50
2013	9	7	1	1	6	2	6	2	1	4	39
2014	12	4	1	4	9	1	7	3	5	7	53
2015	6	3	3	3	11	6	5	4	4	7	52
2016	4	3	2	2	8	4	7	4	4	5	43
Total	195	139	79	147	162	74	132	63	77	190	1258

mean	8.48	6.04	3.43	6.39	7.04	3.22	5.74	2.74	3.35	8.26	54.70
max	15	11	9	17	14	7	10	7	6	17	113
unique	28	13	14	27	35	11	12	12	8	60	220
mean	15.5%	11.0%	6.3%	11.7%	12.9%	5.9%	10.5%	5.0%	6.1%	15.1%	100.0%
max	13.3%	9.7%	8.0%	15.0%	12.4%	6.2%	8.8%	6.2%	5.3%	15.0%	100.0%
unique	12.7%	5.9%	6.4%	12.3%	15.9%	5.0%	5.5%	5.5%	3.6%	27.3%	100.0%
GLM area	14.4%	10.4%	7.1%	13.4%	12.6%	6.0%	9.7%	5.6%	5.7%	15.1%	100.0%

Figure 10. Relative size of each MSE Area (expressed as a percent of the combined areas) based on the size of each Area as specified in Table 4.

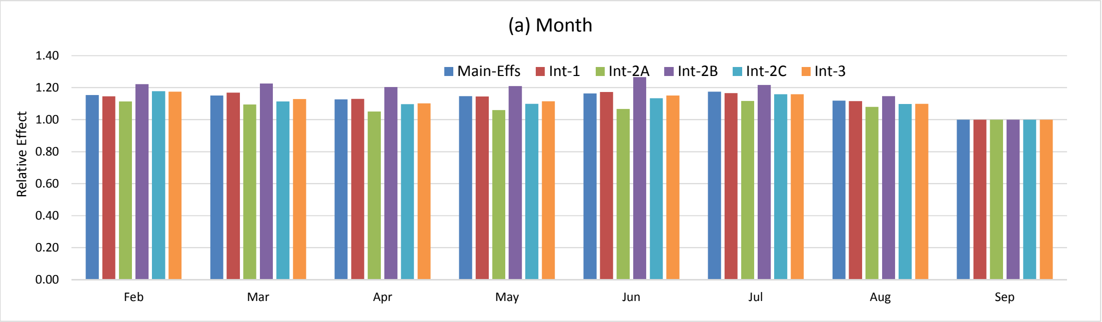


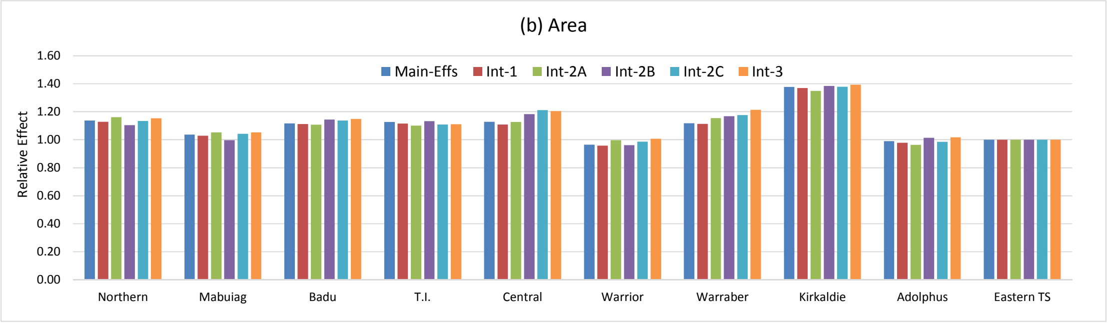
3. Results

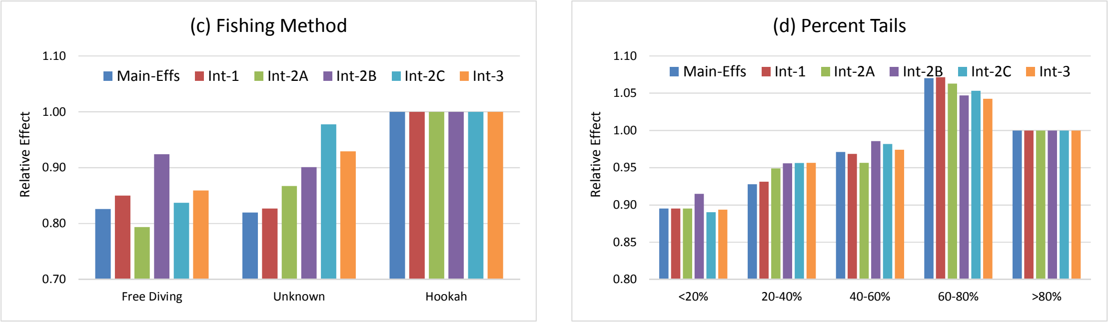
(a) Standardising Effects

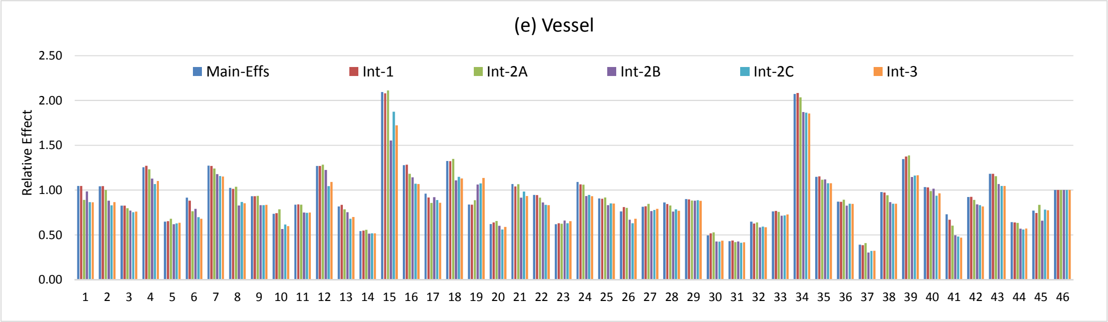
Statistics for the Type 3 contrasts computed for each fitted effect indicated that each effect was highly significant. The relative impact of each level for all effects fitted to each GLM model is shown in Figure 11. For each effect the values have been scaled so that the influence of each level is relative to that of the last level (i.e, *Month*=Sep, *Area*=Eastern TS, *Method*= Hookah and *Proportion-Tails* >80%). For those models which included interactions the *Quarter* and *Area* effects were determined by calculating the mean effect across all *Year*, *Month* and *Area* strata respectively.

Figure 11. The relative impact of each level for each main effect fitted to the each GLM model.









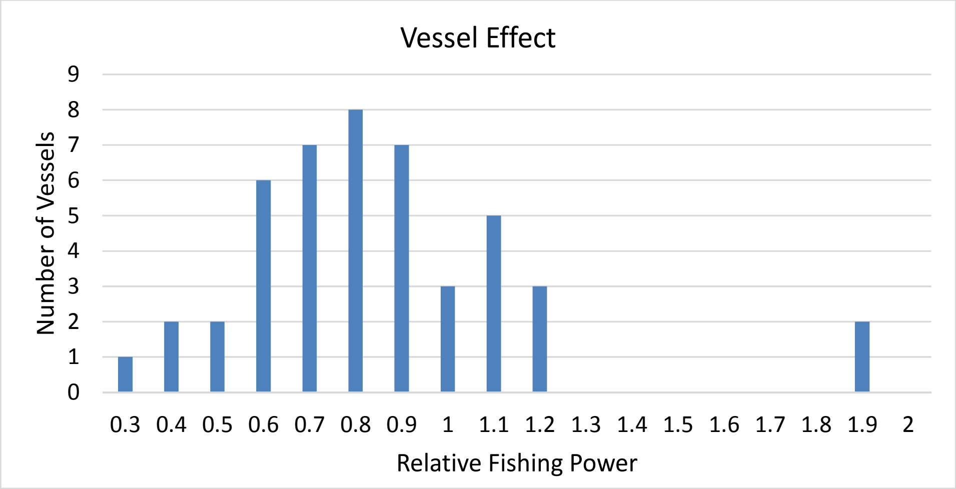
Relative CPUE is relatively constant across the eight months of the year and displays only small variation across the six GLM models, though the CPUE in September is the lowest across all models (c.f. Figure 11a). The greatest variation seen in the relative CPUE within each month is between the results for the 2Ints-A and 2Ints-B models. Taking the average of the relative effect across the results for the six models for each month indicates that the CPUE during February to August is between 11-17% higher than the CPUE in September.

The relative CPUE across the various areas included in the GLM also displays little variation across the six GLM models, though there is some degree of variation across the ten areas (c.f. Figure 11b). Taking the average of the relative effect across the results for the six models for each area indicates that the relative CPUE is lowest in Warrior (98%), Mt Adolphus (99%) and Eastern TS (100%) and highest in Kirkaldie (138%), Warraber (116%) and Central (116%).

Unlike the previous results, the relative CPUE across the three fishing methods displays some variation across the six GLM models (c.f. Figure 11c). For example, the relative effect of the free-diving method relative to hookah diving varies between 79% and 92% while that for the unknown method varies between 82% and 98%. Across all models, the CPUE for hookah fishing is found to be around 15% higher than for free diving and 11% higher than for unknown method. This latter result is to be expected if this fishing method is a combination of the two other fishing methods

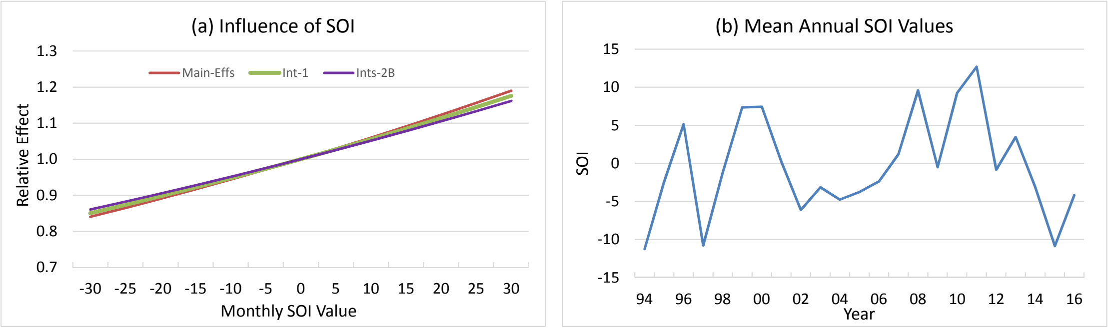
The relative CPUE across all models is similar for each category of the proportion of the catch which is tails with the relative CPUE generally increasing as the *Proportion-Tails* increases in the catch (c.f. Figure 11d). However, the highest CPUE is found for those catches which include 60-80% tails. Across all models, the relative CPUE within each *Proportion-Tails* category is 90%, 95%, 97%, 106% and 100% respectively. Finally, there is substantial variation in the relative CPUE across the 46 vessels included in the GLM models, though the relative effect of each vessel is relatively insensitive to the GLM model used (c.f. Figure 11e). Across all models, the relative fishing power across the fleet varies more than five-fold from 35% to 196% of the standard vessel and the distribution of these effects is shown in Figure 12.

Figure 12. Histogram of the distribution of the relative fishing power of the 46 vessels included in the GLM models.



The monthly value of the SOI was fitted as a simple continuous linear term and the estimated influence of this effect on CPUE based on the results from three of the fitted GLM models is shown in Figure 13. Note, the influence of SOI on CPUE cannot be estimated for several models as the related parameter is aliased when the GLM model includes a *Year.Month* interaction term. The influence of the SOI is seen to be similar for the three models shown in Figure 13, with negative values of the SOI (El Nino conditions) decreasing CPUE while positive values of the SOI (La Nina conditions) increasing CPUE. This indicates that oceanographic conditions may have influenced the high CPUEs experienced in the fishery in 2011 (when the mean SOI value was 12.7) and the low CPUE experienced in the fishery in 2015 (when the mean SOI value was -10.8). However, based on the results shown in Figure 13 the influence on CPUE of the conditions prevailing in these years should have been only 6-7%. Further exploration of the influence of this and other environmental variables is warranted.

Figure 13 (a) Relative influence of the values of the SOI on CPUE and (b) mean annual values of the SOI since 1994.



(b) Annual Abundance Indices

The relative abundance indices based on each of the six GLM models listed in the previous section are listed and displayed in Table 6 and Figure 14 respectively. Relative to the nominal index, each of the standardised indices is similar but is higher at the start of the time-series and lower after 2012. The reasons for these differences can be investigated using the annual influence of each main effect which is shown in Figure 15 for the Main-Effects and Int-1 models. The influence on the annual index is seen to be greatest for the *Vessel* effect followed by the *Proportion-Tails* effect, with the influence of each effect showing an opposing trend over time. The change in the influence of the *Proportion-Tails* effect correlates with the shift from the catch being all tails to now being predominantly whole (c.f. Figure 3b), which decreases CPUE (c.f. Figure 11d) while the change in the influence of the *Vessel* effect is most likely due to an (expected) increase in the relative fishing power of vessels over time. The relative influence of the *Vessel* effect is seen to be greatest towards the start and end of the time-series and explains the divergence seen between the nominal and standardised indices at these times.

The influence of the other effects is seen to be relatively small. For the *Area* and *Month* effects this is likely to be due to the equal weighting given to each *Year-Month-Area* strata in the GLM model analysis. The small but negative trend in the influence of the *Method* effect over the time-series also relates to the fact that there may have been a slight increase in the proportion of catches using hookah diving over time (c.f. Figure 3a) which has the highest CPUE (c.f. Figure 11d)

Table 5. Annual abundance indices for Torres Strait rock lobsters based on the standardised CPUE from the weighted GLM models. The nominal CPUE is also shown for comparison.

Year	Nominal	Main-Effs	Int-1	Int-2A	Int-2B	Int-2C	Int-3
94	0.88	1.39	1.39	1.30	1.44	1.37	1.34
95	0.95	1.30	1.27	1.24	1.34	1.30	1.29
96	0.93	0.92	0.93	0.95	0.94	0.93	0.95
97	1.03	1.17	1.16	1.08	1.20	1.12	1.09
98	0.97	1.04	1.04	1.04	1.08	1.08	1.09
99	0.76	0.64	0.63	0.66	0.63	0.66	0.65
00	0.61	0.63	0.62	0.70	0.58	0.66	0.66
01	0.44	0.48	0.47	0.47	0.50	0.49	0.50
02	0.76	0.70	0.69	0.66	0.62	0.57	0.58
03	1.02	1.08	1.07	1.04	1.06	1.03	1.02
04	1.08	1.17	1.16	1.14	1.06	1.06	1.05
05	1.47	1.48	1.48	1.42	1.46	1.38	1.40
06	0.67	0.68	0.69	0.67	0.67	0.64	0.65
07	1.07	0.97	0.96	0.95	0.96	0.96	0.95
08	0.86	0.83	0.83	0.85	0.89	0.90	0.90
09	0.61	0.63	0.63	0.63	0.66	0.67	0.67
10	1.22	1.12	1.14	1.23	1.18	1.24	1.28
11	2.08	1.75	1.75	1.87	1.96	2.07	2.06
12	1.62	1.40	1.41	1.38	1.30	1.29	1.27
13	1.26	1.16	1.17	1.21	1.15	1.23	1.23
14	1.02	0.90	0.91	0.90	0.88	0.87	0.88
15	0.61	0.58	0.58	0.54	0.51	0.49	0.49
16	1.10	1.00	1.01	1.08	0.95	0.99	1.03
Mean	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Figure 14. Annual abundance indices for Torres Strait rock lobsters based on the standardised CPUE from the Main-Effects and several interaction models. The nominal CPUE is also shown for comparison.

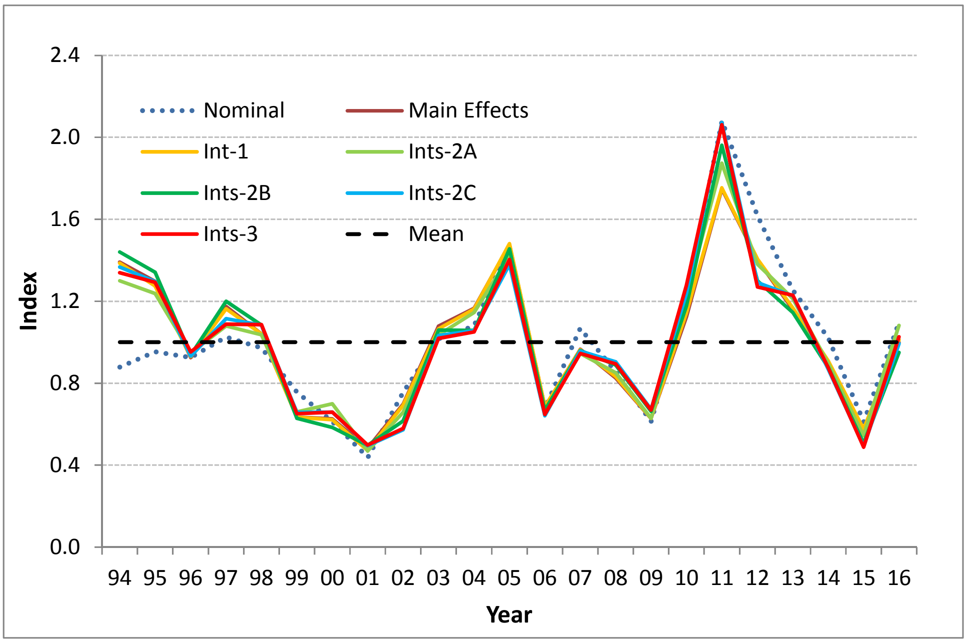


Figure 15. Annual influence of the fixed effects fitted to (a) the Main-Effects model and (b) the Int-1 model.

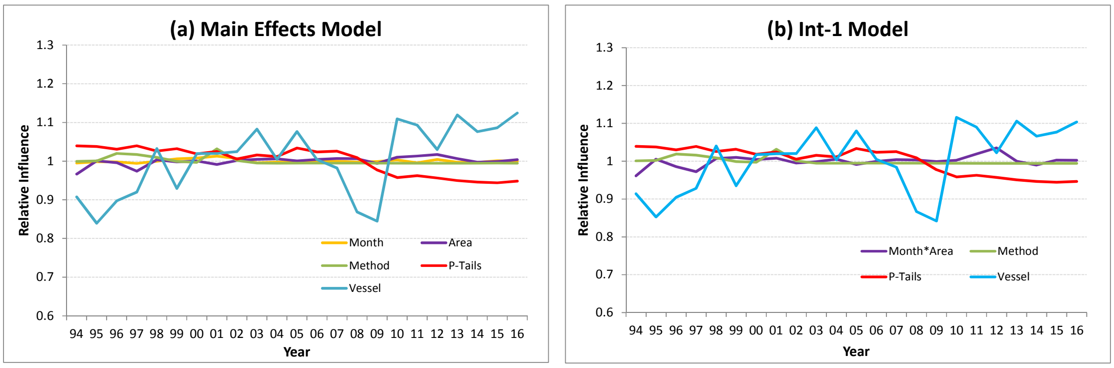


Table 6. Criteria for assessing the goodness-of-fit of each GLM.

GLM	Main	Int-1	Int-2A	Int-2B	Int-2C	Int-3
N-records	42,123	42,123	42,123	42,123	42,123	42,123
df	92	155	307	345	434	497
Deviance	19,485	19,093	17,742	16,947	16,177	15,809
Chi-sq	20,829	20,016	17,969	16,836	15,913	15,502
likelihood	-162,118	-161,630	-159,874	-158,782	-157,680	-157,135
AIC	324,421	323,569	320,362	318,254	316,227	315,263
BIC	325,216	324,910	323,017	321,237	319,980	319,561
N-Strata	1,840	1,840	1,840	1,840	1,840	1,840
Imputed	0	0	10	64	74	74

Several criteria for assessing the goodness-of-fit for each of the GLM models are shown in Table 6. For each criteria shown (where smaller is better) there is an improvement in the fit between each successive model implying that the model which includes all three 2-way interactions provides the best fit to the data. The Int-3 model has considerably greatly flexibility in accounting for inter-annual changes in the distribution of the resource across the different months and areas in comparison to the Main-Effects model which assumes that these distributions are the same for all years. However, the number of parameters (497) estimated in the full interaction model Int-3 is considerably greater than the number of parameters (92) estimated in the Main-Effects model. A consequence of the increase in the number of parameters is that the number of observations on which some of the parameters rely to be estimated can be small (or in some instances zero). A small number of observations increases the likelihood that the corresponding parameter is poorly estimated (or more importantly biased).

Histograms of the number of observations per 2-way strata (for which a separate parameter was estimated) are shown in the Appendix. For 9 of the 230 *Year*Area* strata the number of observations was less than 5 (with 2 of these strata having zero observations) while only one of the 180 *Year*Month* strata had less than 7 observations (being zero). On the other hand, the number of observations was greater than 55 for all of the 80 *Area*Month* strata. For those strata for which the number of observations is zero, the related standardised CPUE for these strata needs to be imputed. (Note, the number of strata for which the standardised CPUE needs to be imputed for each model is shown in Table 6.) For this purpose, the corresponding value using the Int-1 model was used as this model allows the standardised CPUE to be calculated within all strata.

For the Int-3 and Int-2C models, the number of *Year-Month-Area* strata where no observations were available for estimating the related model parameters (which then needed to be imputed) was 74 (or 4.0% of the 1840 in total). The number of strata where the parameter estimation was based on less than 5 or 10 observations was 7.4% and 13.1% respectively. Similar, but slightly smaller results, also apply to the Int-2B model. On the other hand, the number of imputed strata for the Int-2A model was only 10 (or 0.5% of all strata) while the number of strata where the parameter estimation was based on less than 10 observations was only 1%. While it is can be considered best practice to select an abundance index where no parameters have had to be estimated (i.e. the Main-Effects or Int-1 models), the small number of estimated parameters in the Int-2A model reduces the likely bias in the corresponding index.

4. Concluding Remarks

The above analyses, and the resulting indices of annual abundance, are based on the number of assumptions about the data and how these data describe fishing behaviour in the fishery. In particular, if there are features of the fishery which are not adequately captured by the data used in these analyses then the GLMs will not be able to standardise the CPUE for these particular features.

For example, even though the inclusion of interactions allows the model the freedom to the resolve differences in the distribution of the resource across the different areas within different years, the model has no ability to resolve changes in the fishery with may take place within any given area (or month). In particular, the GLM assumes that within each year the distribution of fishing effort within any area is random. However, it is possible that with the introduction of new technologies (such as GPS) that over time fishers have been able to more precisely target their fishing effort to sub-regions of preferred habitat (and higher abundance) within a given area. (Note, the location of fishing effort currently recorded in the logbook is the location of the primary vessel and not the associated tenders which can disperse themselves quite widely). Such ‘effort creep’ would result in higher catches and higher CPUE compared to the situation where no new technologies were available. While the fitted GLM models used in the analyses described in this report appear to capture increases in the fishing power of the fleet due to changes in the vessels leaving and entering the fishery, continual increases in the fishing power over time for individual vessels that remain in the fishery will not captured by the available data and fitted models and as such could result in continual biases in the calculated indices of abundance.

To help overcome this problem it would be useful to further investigate whether or not there have been increases in fishing power over time which are not currently captured by the data. With such information in hand one could then decide whether the data currently available adequately captures the strategies used in the fishery. If not, there needs to be a further discussion as to what additional data may need to be collected so that these aspects of the fishery can be taken into account in the statistical analyses used to standardise the data. Of course, this is a discussion that is pertinent to all fisheries.

Finally, the catches and catch-rates achieved in a fishery are also likely to be influenced by changes in oceanographic and environmental conditions which are likely to change on both a seasonal and inter-annual basis. While the current analyses attempt to model the influence of the monthly value of the Southern Oscillation Index (used to distinguish

El Nino and La Nina conditions) on catch rates, the influence of such environmental changes is likely to require a broader understanding of oceanographic processes that impact on the fishery (including delayed effects such as those which influence recruitment and which sub-sequentially propagate through the fishery over time) and again it would be useful to discuss how such processes can be incorporated into these models.

The use of standardised CPUE as an index of resource abundance is an important input to the stock assessments for many fisheries. This is particularly the situation for those fisheries where fishery independent surveys of the resource are not available or feasible (such in fisheries for highly migratory species such as tunas and billfish). However, as noted above the accuracy of these indices is premised on a number of assumptions, particularly the ability of the logbook data used in the analyses to readily capture the important aspects of the fishery which influence catch rates. In these instances, and where possible, it is useful to incorporate fisheries independent data into the stock assessments. In particular, annual indices of resource status based on fishery independent surveys are usually seen as an important adjunct to the fishery dependent data, and where possible their inclusion in the stock assessment is highly recommended. Where such surveys are not available then attention needs to be paid to ensuring that the logbook data from the fishery captures the information necessary to adequately standardise the catch rates in the fishery as discussed above.

For the Torres Strait rock lobster fishery there are currently two sources of catch and effort data, those for the TVH and TIB sectors. The logbook data from the TVH sector is believed to provide a relatively complete and good source of catch and effort data for this sector, though improvements in compliance to ensure that all fields in the logbook are completed (e.g. area fished and hours fished) would improve the utility of these data. Also, a better recording of the locations of the fishing effort (i.e. at the tender level) would also improve the accuracy of the data for standardising catch rates. On the other hand, the data for the TIB sector is considered to be less complete and the measure of effort (days fished) is less accurate and incomplete in many instances. While the utility of these data to provide a useful index of resource abundance has been investigated elsewhere (Campbell, 2016), again greater effort needs to be placed on ensuring the completeness and accuracy of these data for such purposes.

References

Bentley, N., Kendrick, T.H., Starr, P.J., Breen, P.A. 2012. Influence plots and metrics: tools for better understanding fisheries 1 catch per unit effort standardisations. ICES Journal of Marine Science: 69, 84-88.

Campbell, R.A., 2004. CPUE standardization and the construction of indices of stock abundance in a spatially varying fishery using general linear models. Fish. Res. 70, 209–227.

Campbell, R.A., 2016. Use of TIB Docket-Book Data to construct an Annual Abundance Index for Torres Strait Rock Lobster – 2016 update. Information paper presented to the 19th meeting of the Torres Strait Rock Lobster Resource Assessment Group, held 13 December 2016, Cairns.

Punsley, R.G., 1987. Estimation of the relative abundance of yellowfin tuna, *Thunnus albacares*, in the Eastern Pacific Ocean during 1970-1985. Inter-Amer. Trop. Tuna Comm. Bull. 19, 98-131.

Appendix: Summary of Data fitted to GLM

The following three spatial-temporal effects were included in the GLM used to standardise the CPUE for lobsters caught in the Torres Strait:

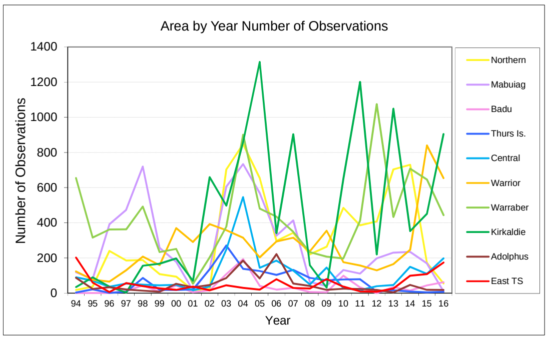
1) Year (all 23 years between 1994 and 2016)
2) Month (all 8 months between February and September)
3) MSE-Area (10 areas)

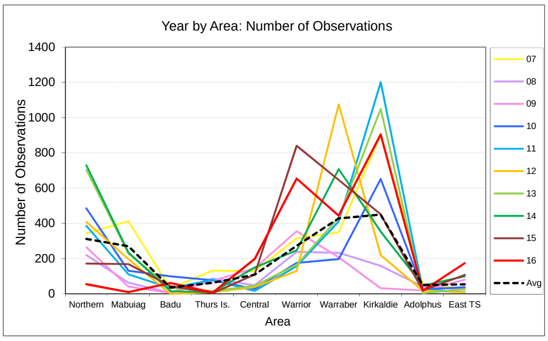
For each 2-way combination of these effects, the following figures provide:

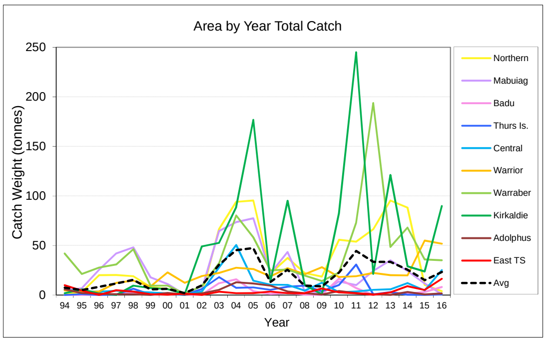
1) Number of data observations
2) Total catch (kilograms of lobsters)
3) Nominal CPUE (kilograms per hour fished)

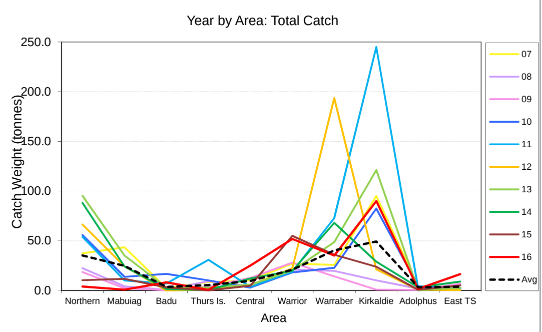
A histogram of the number of observations within each stratum is also shown for each of the above 2-way combination of these effects.

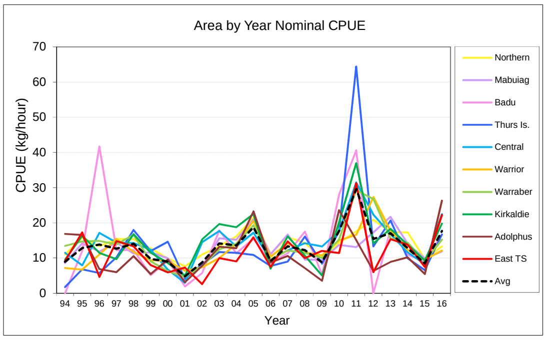
(a) Year*Area

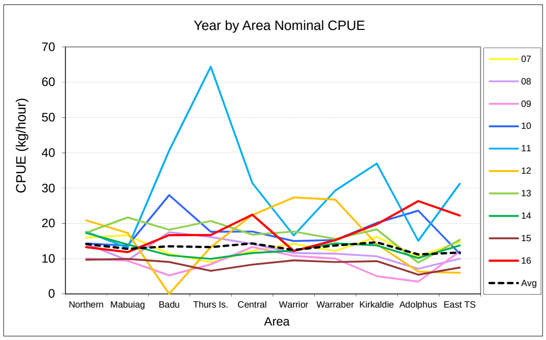








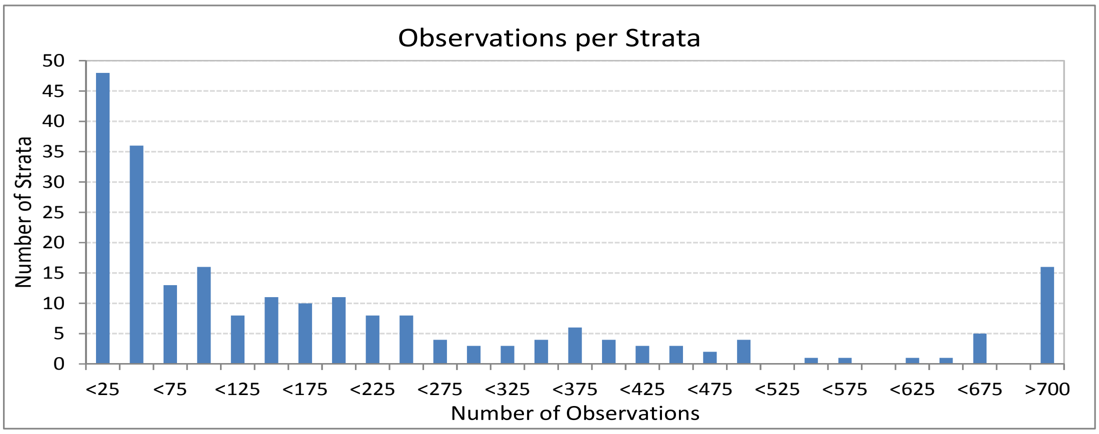




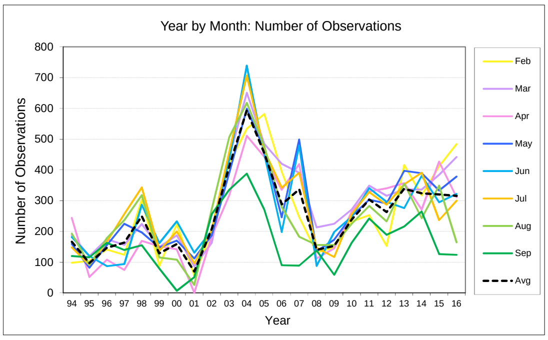
Note: Of the 230 strata (23 years x 10 areas) the number of observations is zero for the following two strata:

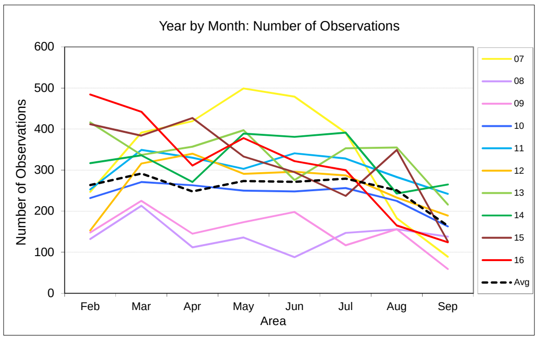
(Year, Area) = (1994, Badu), (2012, Badu)

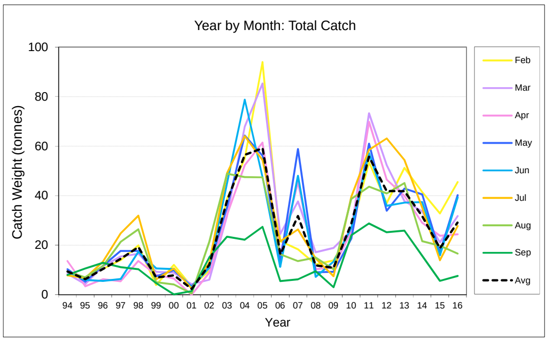
There are a further seven strata where the number of observations was between 1 and 4 and eleven strata where the number of observations was between 5 and 9. The number of observations for all other strata was between 10 and 1315.

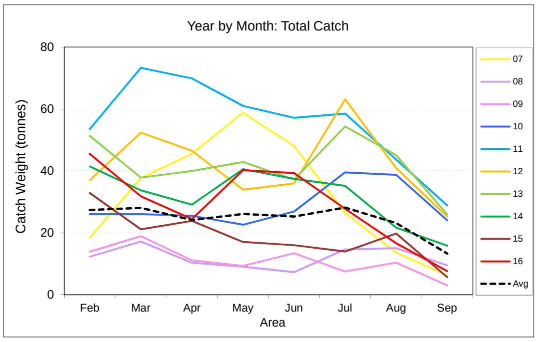


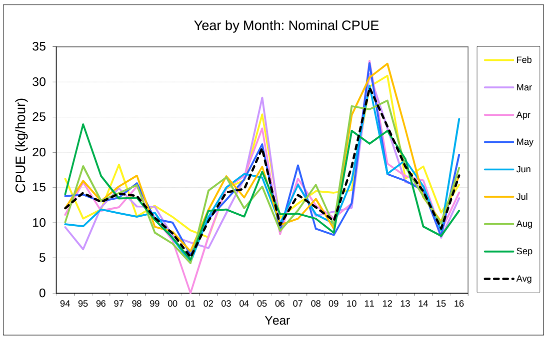
(b) Year*Month

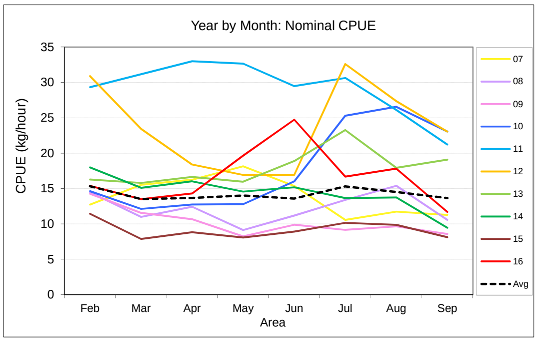




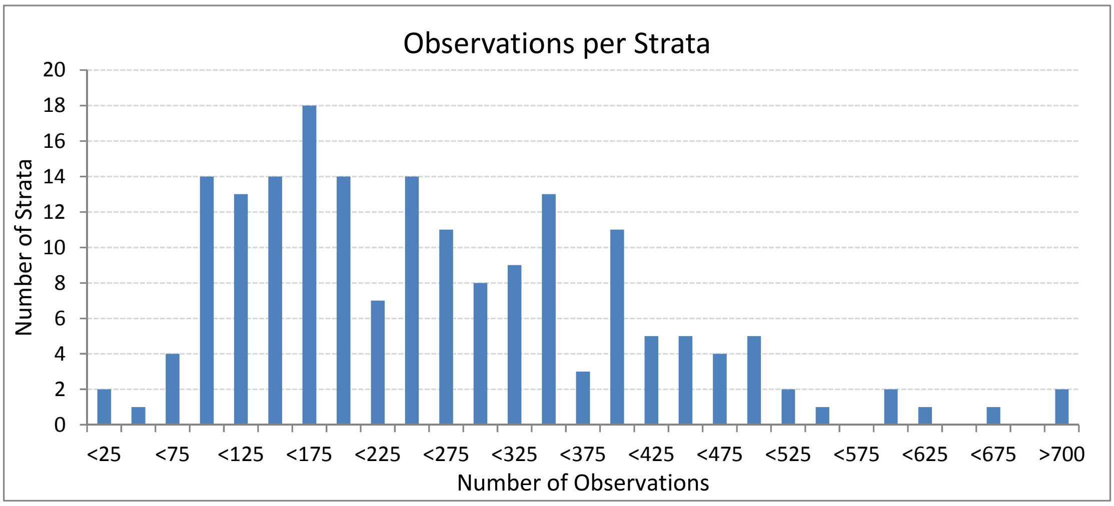




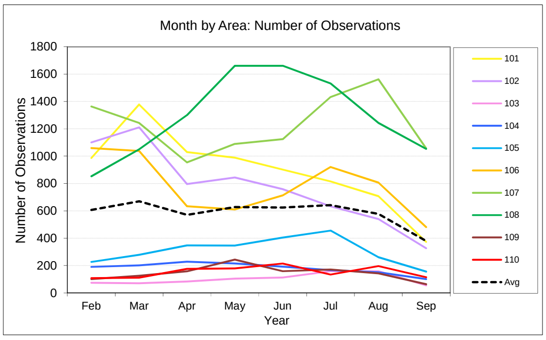


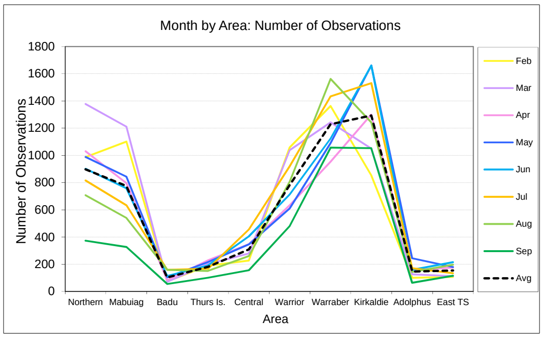


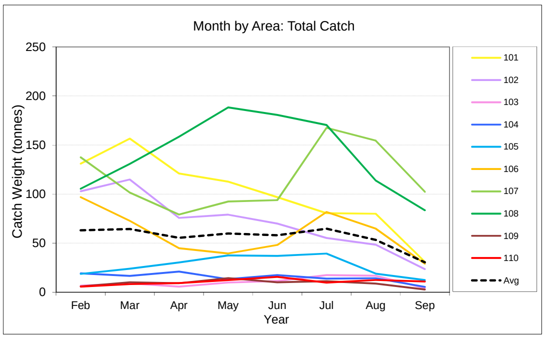
Note: Of the following 184 strata (23 years x 8 months) there were no observations in one strata (2001, April). For the remaining 183 strata the number of observations was between 7 and 739.

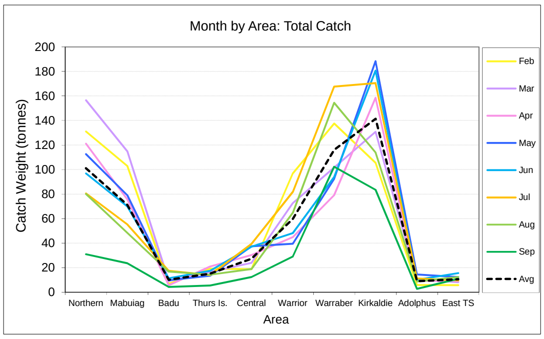


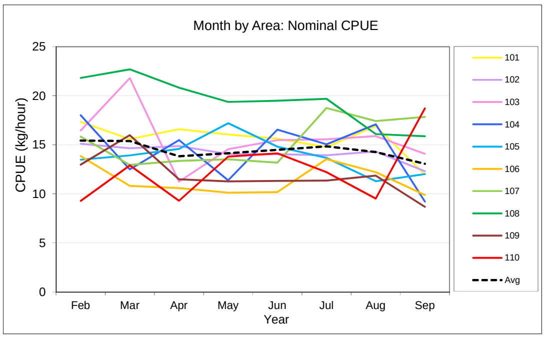
(c) Month*Area

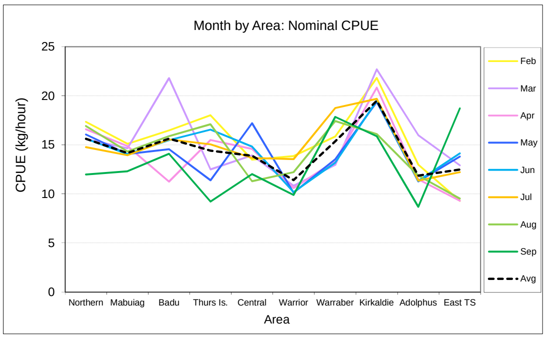












Note: Of the following 80 strata (10 areas x 8 months) the number of observations for all strata was between 55 and 1661.