Test fishing analyses for Nightingale Island

S.J. Johnston, D.S. Butterworth, A. Brandão and A. Muller

MARAM

Department of Mathematics and Applied Mathematics

University of Cape Town

Rondebosch, 7700

Abstract

Because there was no pre-OLIVA fishing in the month of January 2011, the simpler earlier approach of Johnston (2011) to analyse test fishing data by month to assess the OLIVA impact needs to be refined to be able to include data from January 2012. General Linear Models are used for this purpose; these have the additional advantage that they can take more data into account in adjusting for monthly patterns. These models also reveal a regional pattern in the OLIVA impact, which is least in the north and greatest in the southeast. There is some indication that the impact has decreased slightly over time, but the trend is not statistically significant. The best estimate obtained for the OLIVA impact is a decrease of about 50% (SE=5%) in abundance of lobsters above the size limit.

Introduction

Earlier calculations (Johnston, 2011) of the OLIVA impact compared test fishing results each month to the average CPUE for that month in the two preceding seasons. However there was no pre-OLIVA fishing in January 2011 (the 2010) season, necessitating the development of a refined method of analysis. Note that test fishing data for both long-lines and powerboats are available for July, September, October, November and December 2011 as well now as for January 2012. This paper reports the refined methods applied, which involve General Linear Models (GLMs), and the implications of the results obtained.

The Data

Longlines: Average nominal CPUE (kg/trap) values have been calculated from the existing logbook database from 1997-2010. (Here the "year" refers to calendar year.) CPUE values are calculated for each record (or "line"), and these are averaged over records for each month and year to obtain a mean CPUE value for that month and year, as well as the standard error.

Table 1 reports the mean CPUE values with their standard errors, and Figure 1 plots these values for each year along with error bars corresponding to two standard errors (SEs).

Powerboats: Average nominal CPUE (kg/trap) values have been calculated from the existing logbook database from 1997-2010. (Here the "year" refers to calendar year.) CPUE values are calculated for each record, and these are averaged over records for each month and year to obtain a mean CPUE value for that month and year, as well as the standard error. Table 2 reports the mean CPUE values, and Figure 2 plots these values for each year along with errors as in Figure 1.

Powerboat test fishing data are available only as *total* catch and *total* effort (# traps) for each month; thus no standard errors can be calculated for these test fishing months. Figure 3 plots the CV of past CPUE values against $\frac{1}{\sqrt{E}}$ for Nightingale powerboat data, and reports the linear regression fitted to these data (with the regression forced through the origin). Assuming this relationship holds for the test fishing data, and using the total effort (E) reported for the test fishing each month, a CV for these test fishing data can be inferred. All the SEs for test fishing reported in Table 2 have been calculated from these CVs.

GLM models fitted to the data

A General Linear Model (GLM) was first fitted to both the longline and powerboat CPUE data reported in Tables 1 and 2 excluding the data collected post-OLIVA. The data were grouped into the corresponding "season-year" which runs from September to August, i.e. season 1999 refers to CPUE data collected from September 1999 to August 2000.

The GLM model fitted was:

$$\ln(CPUE_{season,month}) = \mu + \alpha_{season} + \beta_{month} + \varepsilon \qquad (1)$$

where

$CPUE_{season,\ month}$ is the CPUE in a given season and month,

μ is the intercept (constant),

α_{season} is the season factor (i.e. the seasons 1997-2010),

β_{month} is a the month factor with levels associated with the fishing months (9-12, and 1), and

ε is an error term assumed to be normally distributed.

The CPUE was standardised relative to the month of *September* and season *2009*. The standardised CPUE series is then obtained from:

$$CPUE_{season,\ month} = \exp(\mu + \alpha_{season} + \beta_{month}) \qquad (2)$$

Equation (2) was then used to provide a predicted CPUE value for each of the test fishing months in the 2011 season, which is done by assuming an underlying abundance equal to the average of the last two years, i.e. α is set equal to the average of the values for the 2009 and 2010 seasons. These values, together with their standard errors, are shown along with the observed test fishing values in Figures 4a (longline) and 4b (powerboat) and reported in Tables 3 (longline) and 4 (powerboat). The ratio of the observed/predicted CPUE values can also be calculated and these are also shown in the bottom plots of Figures 4a and 4b.

In an extension to allow for the possibility that the OLIVA impact is dependent on region, the model of equation (1) was extended to allow for a regional effect, and also to estimate the OLIVA impact by region and month:

$$\ln(CPUE_{season,month,region}) = \mu + \alpha_{season} + \beta_{month} + \gamma_{region} + \delta_{2011,month} + \omega_{2011,region} + \varepsilon \quad (3)$$

where

γ_{region} is the region factor (north, northeast, south, southeast , west),

$\delta_{2011,month}$ is a the monthly pattern of the OLIVA impact in the 2011 season, and

$\omega_{2011,region}$ is a the regional pattern of the OLIVA impact in the 2011 season.

In fitting this model, the longline data fitted now include also the test fishing data for the 2011 season.

Inclusion of region dependence of the OLIVA effect in 2011 is AIC justified ($\Delta AIC = -29$ for inclusion of the $\omega_{2011,region}$ terms), but not any month dependence ($\Delta AIC = +1$ if the $\delta_{2011,month}$ terms are included in addition).

In estimating the impact on overall abundance (B_{season}) from the OLIVA incident, an index of this abundance is provided by the area (A_{region}) weighted sum of the CPUEs predicted in each region by the GLM model of equation 3, i.e.:

$$B_{season} = \sum_{regions} A_{region} . CPUE_{season,region} \quad (4)$$

Thus the ratio (r) of abundance in 2011 relative to the geometric mean abundance over 2009 and 2010 seasons is given by:

$$r = \sum_{region} A_{region} . \frac{CPUE_{2011,region}}{\sqrt{\{\sum_{region} A_{region} . CPUE_{2009,region}\} . \{\sum_{region} A_{region} . CPUE_{2010,region}\}}} \quad (5)$$

Discussion

The GLM approach adopted here has the further advantage (compared to Johnston, 2011) of being able to use data from earlier years in estimating the monthly pattern in CPUE (the β factors).

The plots shown in Figures 1 and 2 depict the September 2011 – January 2012 test fishing CPUE estimates in relation to those for previous years. These plots readily show that both the longline and powerboat test fishing CPUE values for September-December 2011 are significantly less than in immediately preceding years and clearly the lowest values on record (except for the November series for powerboats). January 2012 test fishing CPUEs however seem on par with January 2010 values, so that it becomes important to use the GLM approach to determine whether or not the OLIVA effect has diminished.

Figure 4 shows that with the exception of January for the more variable powerboat data, the test fishing results are all well below the values predicted from the GLM on the basis of past monthly patterns. The less variable longline data are the more reliable, and there is a hint of an upward trend in the observed (from test fishing): predicted ratio from October to January. What is important though is to establish whether that trend is statistically meaningful, given the variances of the estimates which can be computed for the longline data.

Two models have been fitted to the longline ratio estimates (together with their standard errors) in the lower plot in Figure 4a.

Model I: ratio = r $\quad$ r = 0.581 (SE = 0.098) $\quad$ (6a)

Model II: ratio = $r + mt$ (t = month with September = 1) $\quad$ (6b)

r = 0.390 (SE = 0.257) $\quad$ m = 0.058 (SE = 0.072)

The estimates given are maximum likelihood estimates (i.e. omitting the degrees of freedom correction for bias when providing the SE estimate). The model with an increase over time has a slope (m) which is clearly not statistically significant at the 5% level in these terms because of the large SEs associated with each individual ratio estimate. The relative plausibility of these two models can be assessed using AIC_c, which corrects for the small sample size involved. The difference in AIC_c (ΔAIC_c) for the two models is 19.38, so that is in AICc weigthing terms, the weight to be accorded Model II relative to Model I is $e^{(-\frac{\Delta AIC_C}{2})}$=0.00006 which is negligible.

This suggests that the most appropriate inference from these data is that the ratio has not changed over the five-month period of the test fishing. However, there may be covariances between the monthly ratio estimates because of some test fishing places which are common from one month to the next. This could increase the relative weight due to Model II, and is investigated below using the GLM model of equation (3) which takes account of this by the inclusion of region factors.

Results for the AIC-preferred version of equation (3), which includes 2011 interaction terms with region but not with month, are shown in Table 5 and Figure 5. This shows that the OLIVA impact clearly varies regionally, and perhaps surprisingly is least in the north area where the OLIVA grounded.

If instead of a separate factor for each month in 2011 (the $\delta_{2011,month}$ factor), a linear trends is assumed, the estimate of an increase in CPUE from September to January of 15% with an associated SE of 19% is not significant, so again statistical evaluation does not lend support for a decrease over time in the impact from the OLIVA incident during the 2011 season.

The ratio of abundance in 2011 relative to the geometric mean for 2009 and 2010 is estimated from equation (5) to be:

$$r = 0.476,\ \mathrm{SE} = 0.046 \tag{7}$$

where this SE was estimated using the delta approximation. As anticipated, this estimate of the OLIVA impact is more precisely determined (SE = 0.046) than for the model based on equation (1) which ignores regional effects (SE = 0.098 - see equation (6a)).

Conclusion

Since there is a clear regional dependence of the OLIVA impact, results from the approach of equations (3) to (5) seem the most reliable to use. The resultant estimate of the longline CPUE in 2011, expressed relative to the 2009-2010 geometric mean, of r = 0.476 suggests an OLIVA impact of about 50% on abundance. Any reduction in this impact over the period (September-January) of test fishing is slight – the estimated increase over this period is 15% of the September abundance (i.e. an increase in 2011 CPUE relative to the 2009-2010 mean from about 0.45 in September to 0.51 in January), but that estimated increase is not significant a the 5% level.

Reference

Johnston, S.J. 2011. Further test fishing analyses for Nightingale island. MARAM/Tristan/2011/Nov/17.

Table 1: Long-line mean CPUE (kg/trap) values. Standard errors are provided in parentheses. Highlighted values are those from the 2011 season test fishing.

Month	1	2	3	7	9	10	11	12	Total
1997	0.108 (0.016)	0.093 (0.010)	0.053 (0.003)		4.894 (0.264)	4.137 (0.435)			1.517 (0.127)
1998	3.721 (0.232)	2.928 (0.172)	2.546 (0.098)		2.460 (0.095)	2.010 (0.196)			2.860 (0.078)
2000							4.028 (0.302)	4.023 (0.249)	4.025 (0.193)
2001	5.210 (0.299)				2.750 (0.108)	4.214 (0.299)	3.761 (0.183)		3.516 (0.098)
2002					3.239 (0.123)	3.563 (0.136)	3.773 (0.727)		3.359 (0.091)
2003					3.691 (0.178)	3.048 (0.199)	7.857 (0.346)	10.753 (0.890)	5.320 (0.226)
2004	9.943 (0.489)				5.989 (0.213)	5.473 (0.132)	5.652 (0.208)	7.609 (0.952)	5.868 (0.105)
2005							7.363 (0.219)	7.113 (0.215)	7.228 (0.154)
2006									
2007					4.976 (0.311)	4.596 (0.316)		5.746 (0.256)	5.052 (0.178)
2008		7.489 (0.350)							7.489 (0.350)
2009					4.890 (0.184)	4.328 (0.265)	3.987 (0.352)	4.428 (0.271)	4.544 (0.127)
2010	3.699 (0.169)				3.967 (0.186)		5.710 (0.244)	4.739 (0.232)	4.548 (0.113)
2011				**0.923 (1.667)**	**2.048 (0.493)**	**1.422 (0.292)**	**2.746 (0.402)**	**3.113 (0.473)**	
2012	**3.420 (0.604)**								

Table 2: Powerboat mean CPUE (kg/trap) values. Standard errors are provided in parentheses. Highlighted values are those from the 2011 test fishing.

Month	1	2	3	7	9	10	11	12	Total
1997	1.878 (0.260)	1.177 (0.058)	0.835 (0.053)		2.606 (0.129)	3.226 (0.139)	2.550 (0.168)		1.911 (0.081)
1998	2.183 (0.144)	1.310 (0.090)	1.157 (0.101)						1.613 (0.078)
1999		1.829 (0.128)	0.464 (0.075)		2.736 (0.128)	2.176 (0.114)	2.718 (0.121)		2.149 (0.071)
2000					3.062 (0.401)	2.586 (0.192)	5.217 (0.396)	4.113 (0.250)	3.250 (0.153)
2001	7.121 (0.624)				2.274 (0.158)	2.066 (0.215)	1.337 (0.093)		2.332 (0.138)
2002					2.430 (0.197)	2.076 (0.117)	1.582 (0.228)		2.245 (0.125)
2003					2.844 (0.268)	2.782 (0.154)			2.819 (0.170)
2004									
2005									
2006									
2007					9.833 (0.197)	7.843 (0.474)		4.915 (0.348)	7.567 (0.487)
2008		1.804 (0.123)							1.804 (0.123)
2009					2.456 (0.167)	4.306 (0.355)	4.338 (0.315)	3.448 (0.335)	3.229 (0.179)
2010	1.957 (0.162)				2.384 (0.141)		3.674 (0.174)	2.841 (0.190)	2.706 (0.107)
2011				**0.704** (0.127)	**1.604** (0.245)	**0.917** (0.161)	**2.708** (0.477)	**1.200** (0.259)	
2011	**2.575** (0.556)								

Table 3: Observed (test fishing) and GLM predicted **longline** CPUE for the 2011 season. Standard errors are also reported in parentheses.

Month	Observed	Predicted	Ratio of observed/predicted
September 2011	2.048 (0.493)	3.945 (0.382)	0.519 (0.260)
October 2011	1.422 (0.292)	3.791 (0.586)	0.375 (0.257)
November 2011	2.746 (0.402)	4.648 (0.636)	0.591 (0.200)
December 2011	3.113 (0.473)	5.096 (0.568)	0.611 (0.189)
January 2012	3.420 (0.604)	4.926 (0.433)	0.694 (0.197)

Table 4: Observed (test fishing) and GLM predicted **powerboat** CPUE for the 2011 season.

Month	Observed	Predicted	Ratio of observed/predicted
September 2011	1.604	1.132	0.517
October 2011	0.917	1.145	0.292
November 2011	2.708	1.168	0.842
December 2011	1.200	1.025	0.431
January 2012	2.575	0.602	1.410

Table 5: Region factors estimated for the GLM of equation (3) which includes such effects (note that 2010 and north are included in the intercept).

	Estimate (SE)	Value relative to 2009-2010 geometric mean ($e^{-0.5\alpha_{2009}+\alpha_{2011}+\omega_{2011,region}}$	Size of region (km^2) A_{region}
α_{2009}	-0.119 (0.099)		
α_{2011}	-0.438 (0.161)		
γ_{north}	-		12.13
$\gamma_{northeast}$	-0.154 (0.071)		3.29
$\gamma_{southeast}$	-0.126 (0.070)		3.02
γ_{south}	-0.094 (0.070)		9.00
γ_{west}	-0.156 (0.074)		5.87
$\omega_{2011,north}$	-	0.684	
$\omega_{2011,northeast}$	-0.630 (0.201)	0.365	
$\omega_{2011,southeast}$	-1.060 (0.201)	0.238	
$\omega_{2011,south}$	-0.840 (0.201)	0.296	
$\omega_{2011,west}$	-0.296 (0.202)	0.509	

Figure 1: Nightingale long-line CPUE values for September, October, November and December. Means +/- 2 SEs are shown.

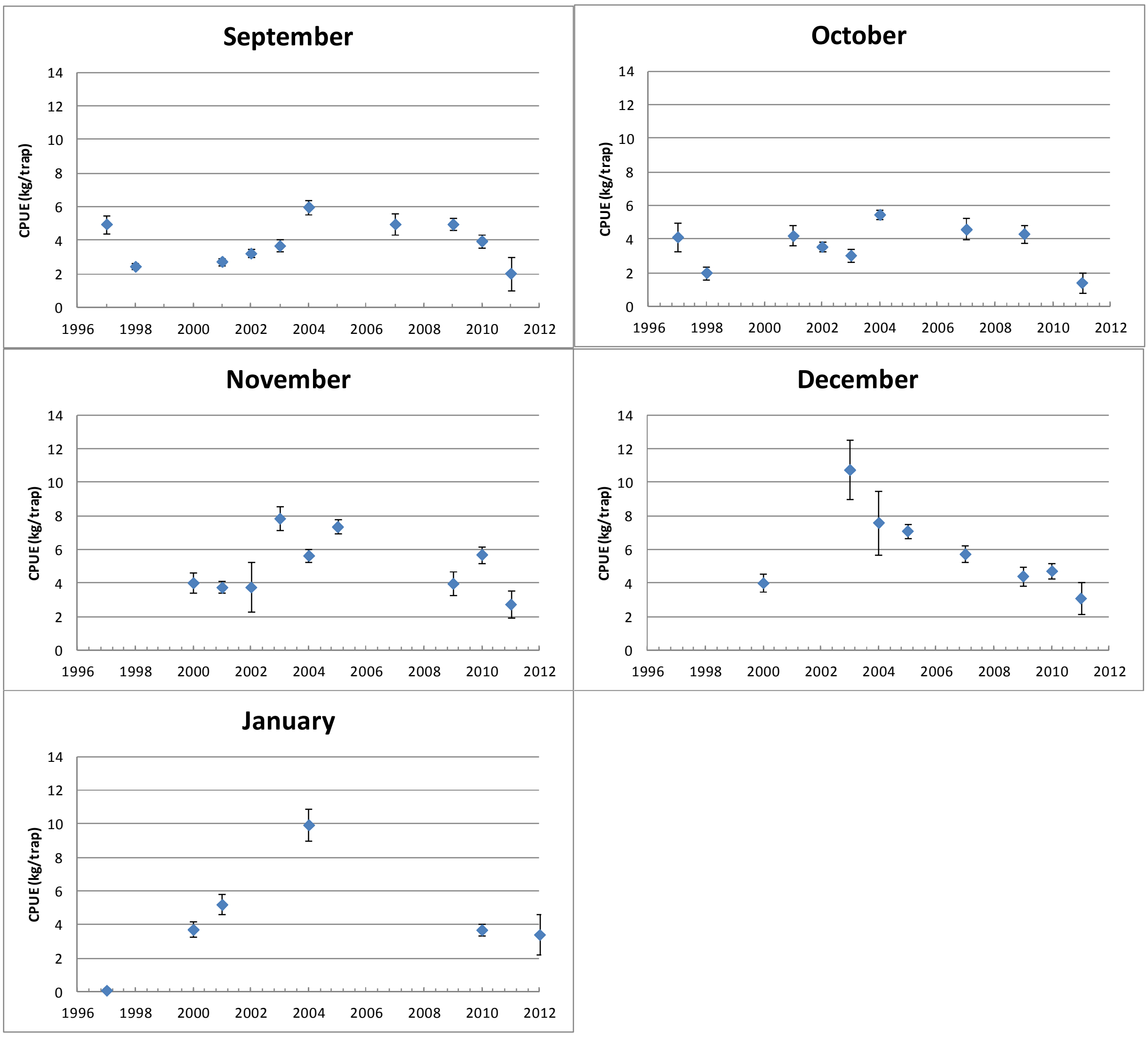


Figure 2: Nightingale powerboat CPUE values for September, October, November and December 2011, and January 2012. Means +/- 2 SEs are shown.

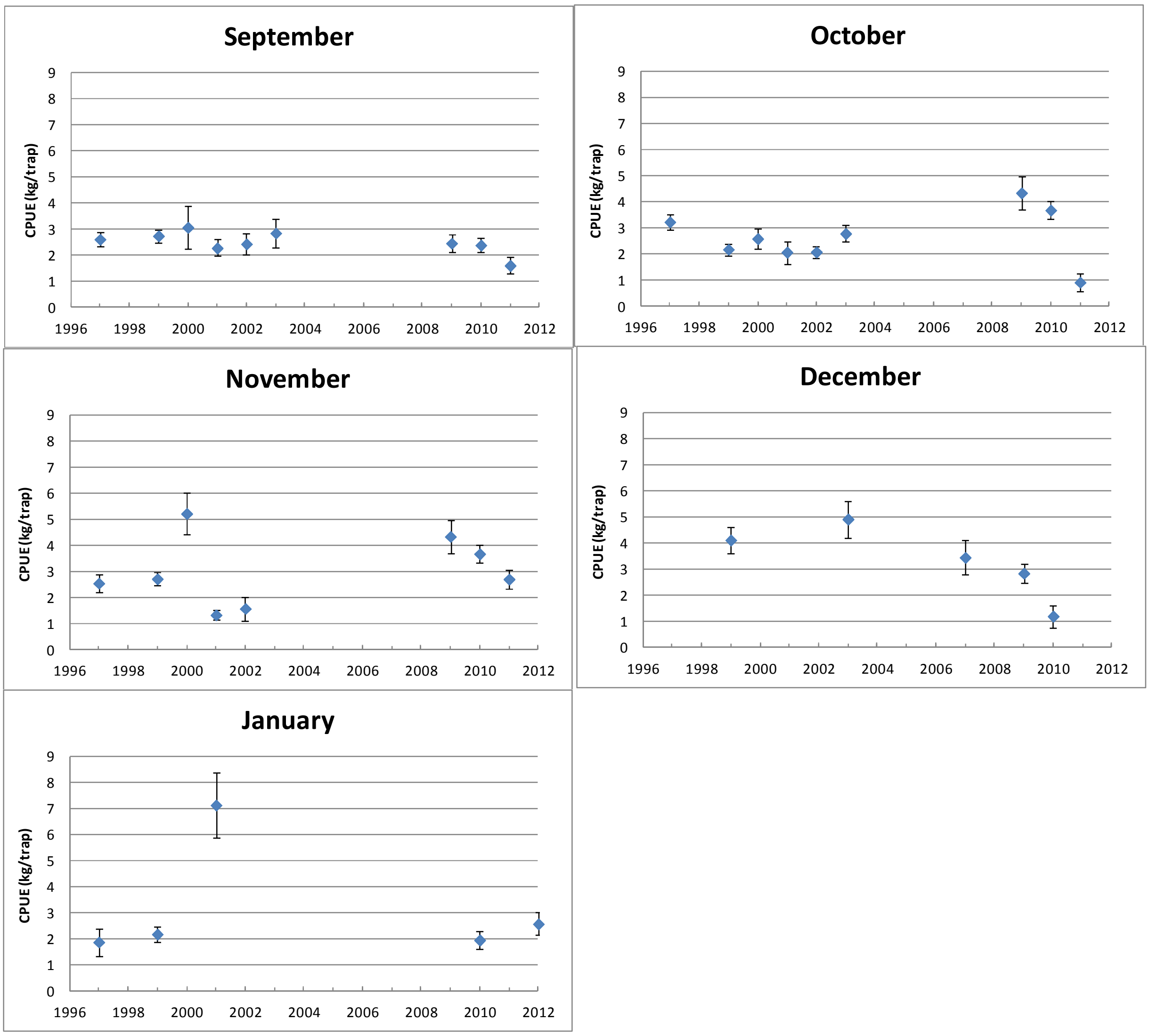


Figure 3: Plot of CV of CPUE values against $\frac{1}{\sqrt{E}}$ for Nightingale powerboat data. The straight line is a linear regression that is fitted to these data, which has been forced to pass through the origin.

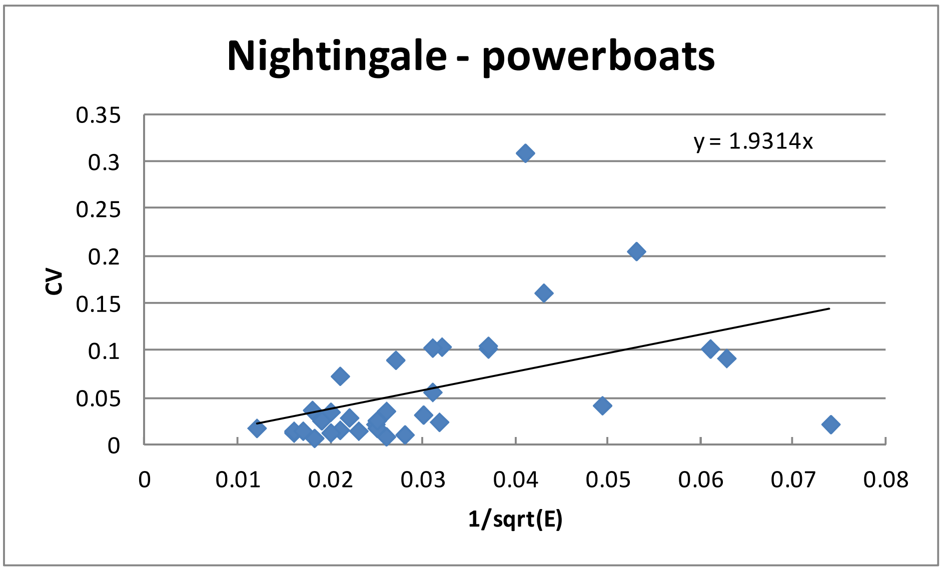
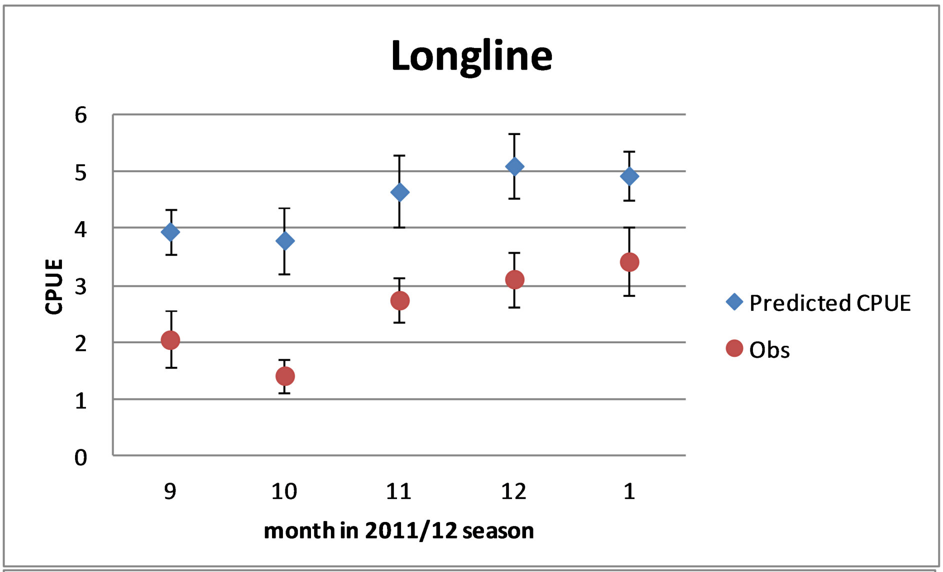


Figure 4a: Longline observed CPUE (test fishing) and GLM predicted values for the 2011 season. The error bars reflect plus and minus one standard error. The bottom plot shows the ratio of the observed to the predicted CPUE values for the 2011 season.



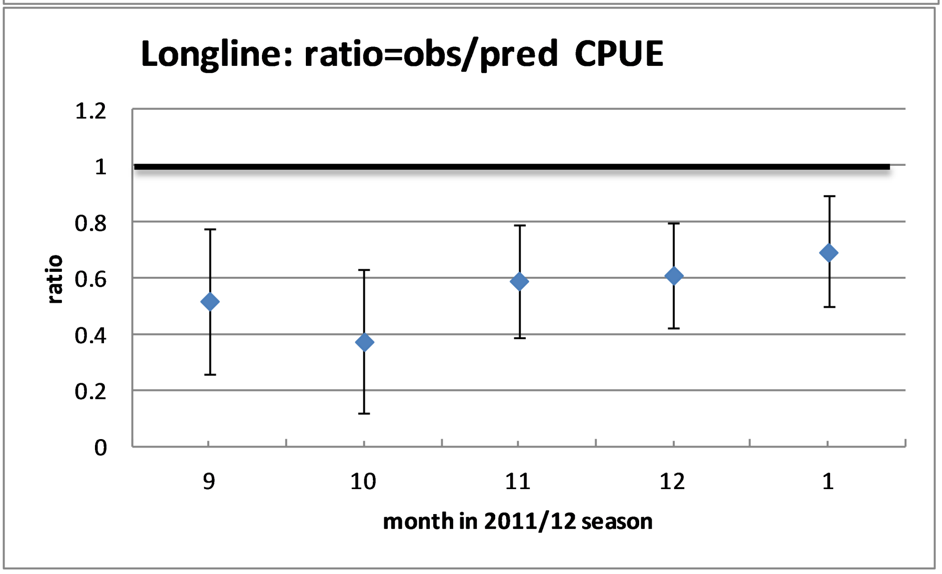
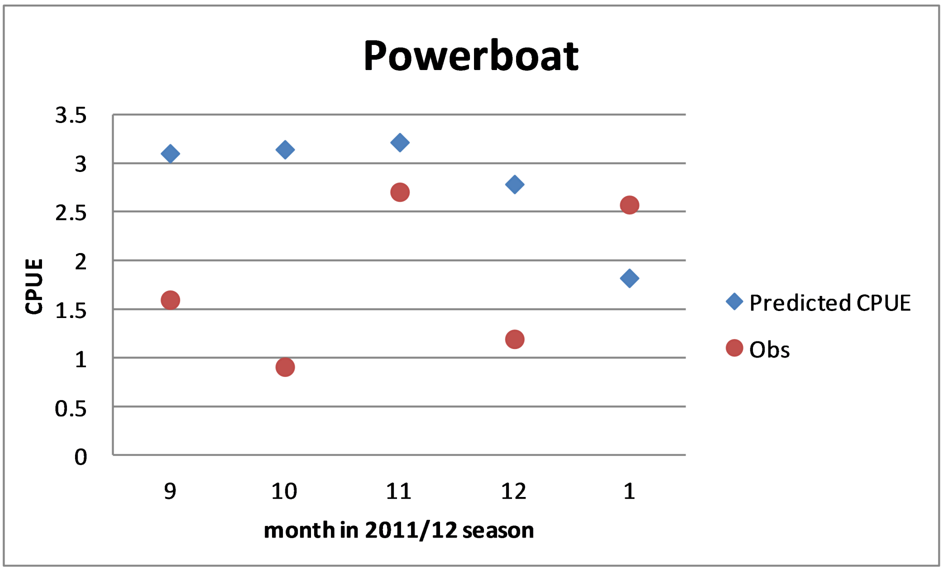


Figure 4b: Powerboat observed CPUE (test fishing) and GLM predicted values for the 2011 season. The bottom plot shows the ratio of the observed to the predicted CPUE values for the 2011 season. Error bars reflecting SEs have not been shown as they can only be inferred for the observed values.



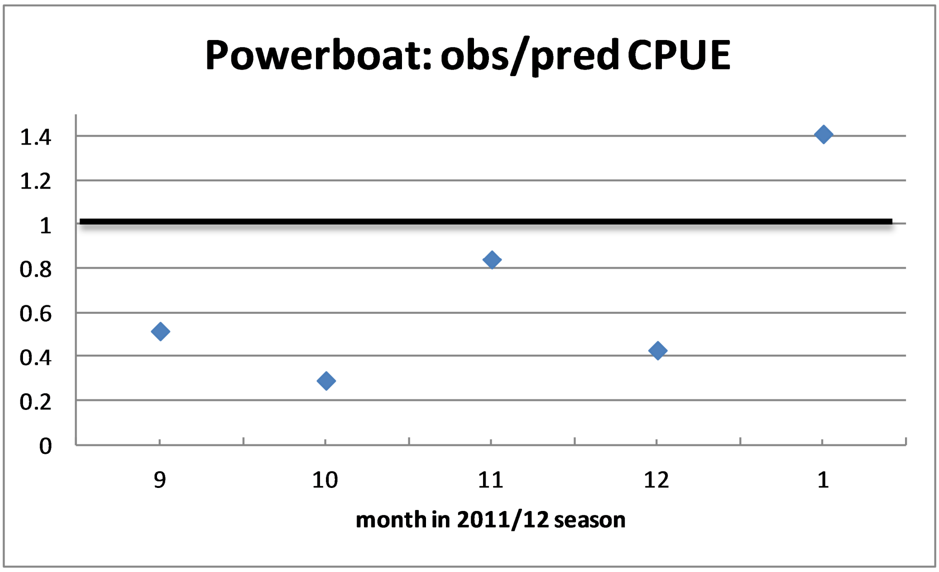


Figure 5: GLM predicted longline CPUE by region in 2011 relative to the geometric means of that in the 2009 and 2010 seasons using a GLM model based on equation (3).