Inclusion of skipper effect in the GLM standardisation of the CPUE abundance indices for orange roughy off Namibia

Anabela Brandão and Doug S. Butterworth

Marine Resource Assessment & Management Group (MARAM)
Department of Mathematics and Applied Mathematics
University of Cape Town
Rondebosch, 7701, Cape Town

January 2007

Abstract

Various GLM analyses are used to standardise the CPUE data for Namibian orange roughy to investigate the impact of using information on skippers rather than on vessels. The inclusion in the GLMs of a skipper factor rather than a vessel factor results in lower CPUE indices for the most recent years. Classifying skippers as experienced or inexperienced in orange roughy fishing does not have an appreciable impact on standardised CPUE trends. However, classifying individual initially inexperienced skippers as experienced after their first year in the fishery results in lower CPUE indices, in particular for the last four years. In broad terms, there is not a clear difference between skippers pre-classified as experienced or inexperienced.

Introduction

The delta-lognormal model, as first proposed in Brandão and Butterworth (2002), is used to standardise the commercial orange roughy CPUE data. This model includes a vessel factor so that the CPUE indices can be standardised for differences in the operation of vessels and to act as a surrogate for possible differences in the performance of skippers. This is not a totally satisfactory way to account for differences in skippers in the standardised CPUE as a vessel will have had different skippers at various times.

In this paper, information on which skipper was in operation is included in the GLM analyses to investigate their effect on the CPUE. Skipper information was not available for all the companies that have been in operation since the start of the fishery, so only commercial tow data for which the skipper information is available is used in the present analyses (11 430 records out of 17 971). To simplify the GLM that is normally used to standardise the CPUE data, a lognormal model is fitted to the CPUE data instead of fitting a delta-lognormal model.

In addition, to investigate the effect of both vessel and skipper on CPUE, a General Linear Mixed Model (GLMM) was used to standardise the commercial orange roughy CPUE data, with the factor for skipper considered as a random effect in this case.

The Models

A lognormal GLM is used to standardise the commercial orange roughy CPUE data. Various models have been fitted to take different aspects of skipper information into account and also to investigate the effect on CPUE when using data on vessel rather than skipper. The general model is of the form:

$$\ln(CPUE + \delta) = \mu + \alpha_{aggregation} + \beta_{year} + \gamma_{captain} + \lambda_{vessel} + \phi_{month} + \eta_{year \times aggregation} + \theta_{captain \times aggregation} + \tau_{vessel \times aggregation} + \varepsilon$$

where:

δ	is a constant added to the CPUE to allow for the occurrence of zero catches (taken to be 10% of the average of the nominal CPUE);
μ	is the intercept;
aggregation	is a factor that represents the aggregation in which the tow took place, three methods of defining this factor were investigated: i) *strata* is a factor with 2 levels associated with the “south” and the combined “known” aggregations, ii) *agg* is a factor with 5 levels associated with the “south” and the individual “known” aggregations, and iii) *sub-agg* is a factor with 16 levels associated with the sub-aggregations for the “south” and the “known” aggregations;
year	is a factor with 8 levels associated with the years 1997–2004;
captain	is a factor that represents the skipper; four methods of defining this factor were investigated: i) *skipper* is a factor with 13 levels associated with the individual skipper, ii) *newboy* is a factor with 15 levels associated with the individual skipper, but those that were identified as inexperienced and were in the fishery for more than one year

had a different level associated with their first year in the fishery than for all subsequent years,

iii) *novice* is a factor with 2 levels associated with whether or not the skipper was an experienced orange roughy skipper, and

iv) *novice1* is a factor with 2 levels associated with whether or not the skipper was an experienced orange roughy skipper, except that all inexperienced skippers were only categorised as such in their first year and then were denoted as experienced in all subsequent years;

vessel	is a factor with 6 levels associated with the vessels for which there is information on the skipper;
month	is a factor with 12 levels (January– December);
year×*aggregation*	is the interaction between year and aggregation;
captain×*aggregation*	is the interaction between captain and aggregation;
vessel×*aggregation*	is the interaction between vessel and aggregation; and
ε	is the error term assumed to be normally distributed.

It is conceivable that both the vessel and the skipper in operation in a fishery would have an effect on CPUE. This is readily taken into account by including both factors in a GLM. However, most skippers operated in one year only, so that including both vessel and skipper as factors in a GLM causes aliasing, i.e. some of the covariates are identical to combinations of other covariates so that the parameters for those covariates cannot be distinguished. To overcome this problem, a GLMM was investigated in which the factor *vessel* is considered as a fixed effect but the factor *skipper* is taken to be a random effect, i.e. instead of estimating the effect of each skipper on CPUE, the variance of the distribution assumed for the effect of the skippers is estimated.

Model Implementation

The GLMs to investigate the effect of skipper on the CPUE of orange roughy are fitted to the commercial tow data inside the known aggregations of orange roughy in Namibia for the fishing years (July–June) 1997 to 2004 for which information on skipper is available. A total of 11 430 tows was available for the analyses (including tows that were taken in the “south” aggregation).

GLM Results and Discussion

Table 1 gives various models fitted and the adjusted R^2 (the percentage of the variance accounted for by the model). The inclusion of skipper or vessel alone makes large difference in terms of the percentage of the variance accounted for, with skipper having a larger R^2_{adj}. However, once other factors are included in the GLM (especially including the factor for aggregation *strata*), although (in terms of R^2_{adj}) the factor skipper still performs better than the factor vessel, the difference is not that much greater. By differentiating the first year of inexperienced orange roughy skippers improves the fit slightly. By considering the factor *novice* (or *novice1*), i.e. whether the skipper is experienced or not, does not improve the model fits.

Table 2 shows the skipper effect on CPUE (in the logarithmic scale) when the first year of an inexperienced orange roughy skipper is differentiated from subsequent years. One would expect this effect to be lower in the first year that a novice skipper enters the fishery. However, this result is not observed consistently for various models and for the two skippers that were inexperienced in the orange roughy fishery and who remained in the fishery for more than one year. Fitting only skipper in the GLM has the effect of the first year in the fishery acting in the opposite direction to that expected for both skippers. GLMs including the factor *novice* result in effects for some experienced skippers which are worse than those for novices.

Figure 1 gives the skipper effect for two GLMs: one also includes the factor aggregation *strata*, and the second includes the factors *strata* and *year*. The skippers in the plot are ordered from experienced (denoted with a star) to novices, and within these categories in the order in which they entered the fishery. If there was more than one new skipper in one year the order is random. The skipper effect does not show the expected upward trend linking experience to CPUE.

Figure 2 shows CPUE trends for various models fitted. If only the year factor is fitted, the trend in CPUE for 2004 is very different when skipper is included in the model, while the nominal trend is very similar to that when vessel is included in the model. The big difference in the CPUE index for 2004 disappears when *strata* is taken into account in the GLM (Fig. 3). Again the trend when vessel is included is similar to that when only the factors *year* and *strata* are taken into account. Replacing *vessel* by *skipper* shows a bigger decline in CPUE since 2001. The trend in CPUE for the GLMM with *vessel* as a fixed effect and *skipper* as a random effect is basically the same as the trend of the model with *skipper* as a fixed effect. Including the *novice1* factor does not change the CPUE trend when only *year* and *strata* are fitted in the GLM, but distinguishing new skippers to the orange roughy fishery as experienced after their first year results in lower CPUE indices for 1997 and from 2001 onwards (Figure 4).

Fitting a GLM or a GLMM with skipper as a random effect does not alter the CPUE trend nor does it change the uncertainty in the CPUE indices substantially (Figure 5).

Figure 6 gives CPUE trends for GLMs that incorporate a factor for aggregation that has as its levels the “south” and the individual “known” aggregations (*agg*). Other factors in the model consist of *year*, *month*, and an interaction between *year* and *agg*. The distinguishing term in the two GLMs fitted is whether a factor for skipper or for vessel is included with the other variables in the model. Replacing vessel with skipper results in increased CPUE indices in the first few years but lower indices in the remainder of the series for all aggregations except the “South”. For this aggregation, there is no clear reversal in indices in about 2000, but the last two years nevertheless have lower CPUE indices when skipper rather than vessel is included in the GLM.

There is no appreciable difference in CPUE indices whether the factor *agg* or *strata* is included in a GLM (Figure 7). The additional inclusion of skipper lowers the CPUE indices.

Figure 8 shows similar results to Figure 6 except that in this case the factor for aggregation has been categorised into the different sub-aggregations of the “south” and the “known” aggregations (*sub-agg*). Differences between CPUE trends for each aggregation whether a vessel or skipper factor is used in the GLM generally decrease when the aggregations are split into their various sub-aggregations. Differences in CPUE indices (especially in *Frankies*) depending on whether the factor *agg* or *sub-agg* is fitted in the GLM are most probably because of the large amount of aliasing that occurs when the aggregations are split into the various sub-aggregations.

Acknowledgements

Data were provided by NatMIRC for this study. Herman Krauze kindly provided skipper information. Funding from the Namibian Deepwater Fishing Industry is gratefully acknowledged.

References

Brandão, A. and Butterworth, D.S. 2002. Standardised CPUE abundance indices for orange roughy off Namibia based on lognormal and delta-lognormal models. Namibian Ministry of Fisheries and Marine Resources document: DWFWG/WkShop/Feb02/Doc 1.

Table 1. Adjusted R^2 for various GLMs fitted.

Model	R^2_{adj}
skipper	18.4
newboy	19.4
vessel	7.5
novice	0.8
novice1	1.2
year	16.7
strata	33.5
month	1.5
year+skipper	21.4
year+vessel	18.2
year+novice	16.9
year+novice1	16.7
year+newboy	21.7
year+strata+skipper	38.7
year+strata+newboy	38.6
year+strata+novice1	37.6
year+strata+vessel	38.3
year+strata+skipper+strata.skipper	38.9
year+strata+vessel+strata.vessel	38.4
year+strata+novice1+novice1.year	37.8
year+strata+novice1+novice1. strata	37.6
year+strata+novice1+vessel	38.3
year+strata+newboy+ strata.newboy	38.8
year+strata+skipper+strata.skipper+year.strata	38.9
year+strata+skipper+vessel	38.7
year+strata+newboy+vessel	38.8

Table 2. Skipper effect (standard error) when the first year of a novice skipper is differentiated for various models.

	Model			
Skipper level	**Skipper only**	**Skipper + year**	**Skipper + year+strata**	**Skipper + year+strata+ skipper.strata**
Franz (first year)	1.947 (0.183)	0.361 (0.212)	0.530 (0.188)	0.563 (0.188)
Franz (subsequent years)	-0.056 (0.110)	0.724 (0.127)	0.757 (0.112)	0.743 (0.113)
Jenner (first year)	-0.624 (0.159)	1.322 (0.268)	0.849 (0.238)	0.890 (0.246)
Jenner (subsequent years)	-1.466 (0.117)	-0.979 (0.380)	1.013 (0.339)	1.003 (0.357)
Novice – experienced level	0.6693 (0.070)	-0.365 (0.075)		
Novice1 – experienced level	1.0592 (0.089)	-0.049 (0.103)	-0.036 (0.089)	

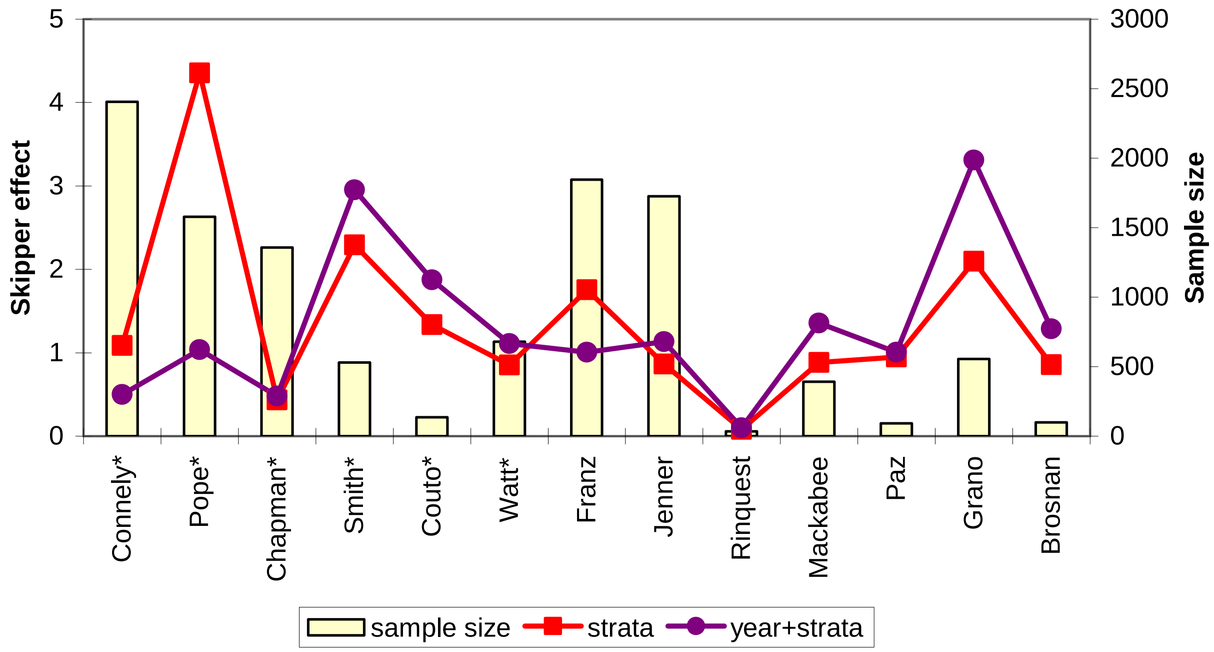

Figure 1. The skipper effect (i.e. exp($\gamma_{captain}$)) for two GLMS fitted: *skipper+strata* and *skipper+strata+year*. Each series is normalised to its geometric mean. Skippers are listed in order of appearance in the fishery; those with asterisks had previous experience in other orange roughy fisheries. The number of tows carried out by each skipper is also shown.

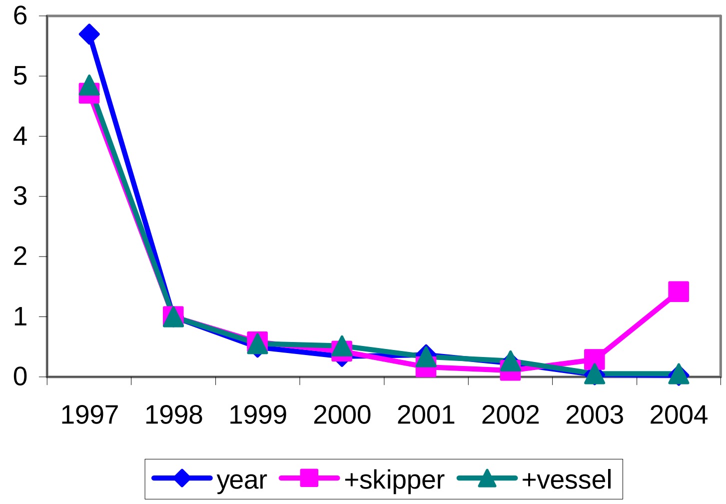

Figure 2. Standardised CPUE indices for various models fitted. Models shown are: *year* only, *year+skipper* and *year+vessel*.

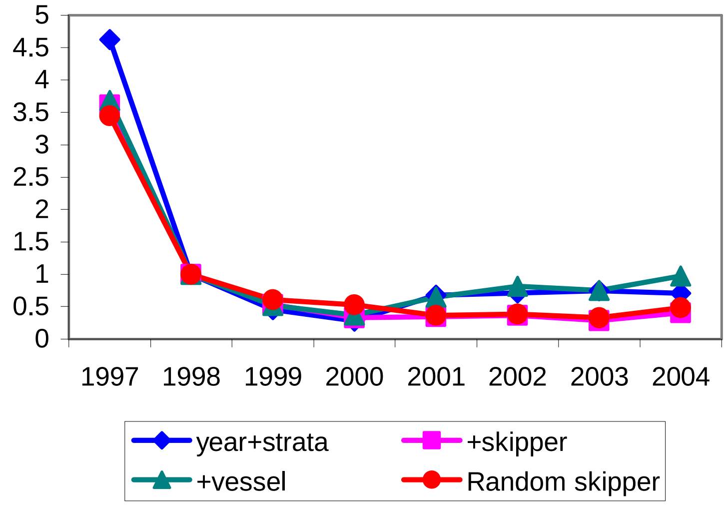

Figure 3. Standardised CPUE indices for various models fitted. The GLMs shown are: *year+strata*, *year+strata+skipper*, *year+strata+vessel* and a GLMM with year+*strata+vessel* as fixed effects and *skipper* as a random effect.

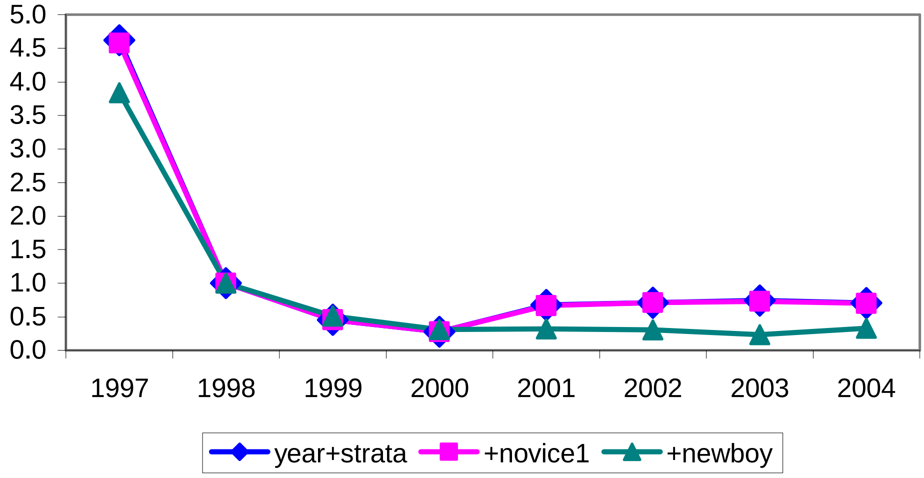

Figure 4. Standardised CPUE indices for various models fitted. Models shown are: *year+strata*, *year+strata+novice1* and *year+strata+newboy*.

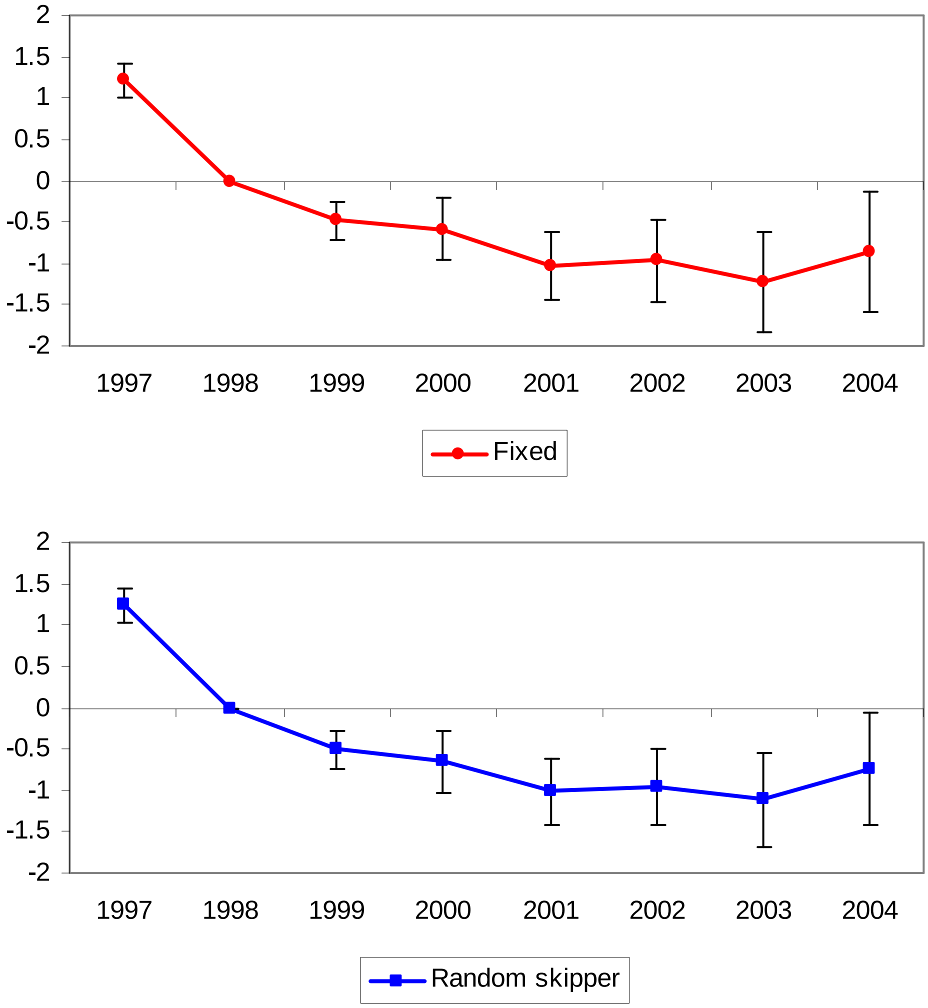

Figure 5. Standardised CPUE indices with 95% confidence intervals for the model: *year+strata+skipper+vessel*, when the model was fitted as a GLM (i.e. all factors fixed effects (top)) and when the model was fitted as a GLMM (i.e. *year+strata+vessel* taken as fixed effects and *skipper* as a random effect (bottom)). Ln (*CPUE*+δ) is plotted for a clearer indication of the confidence intervals.

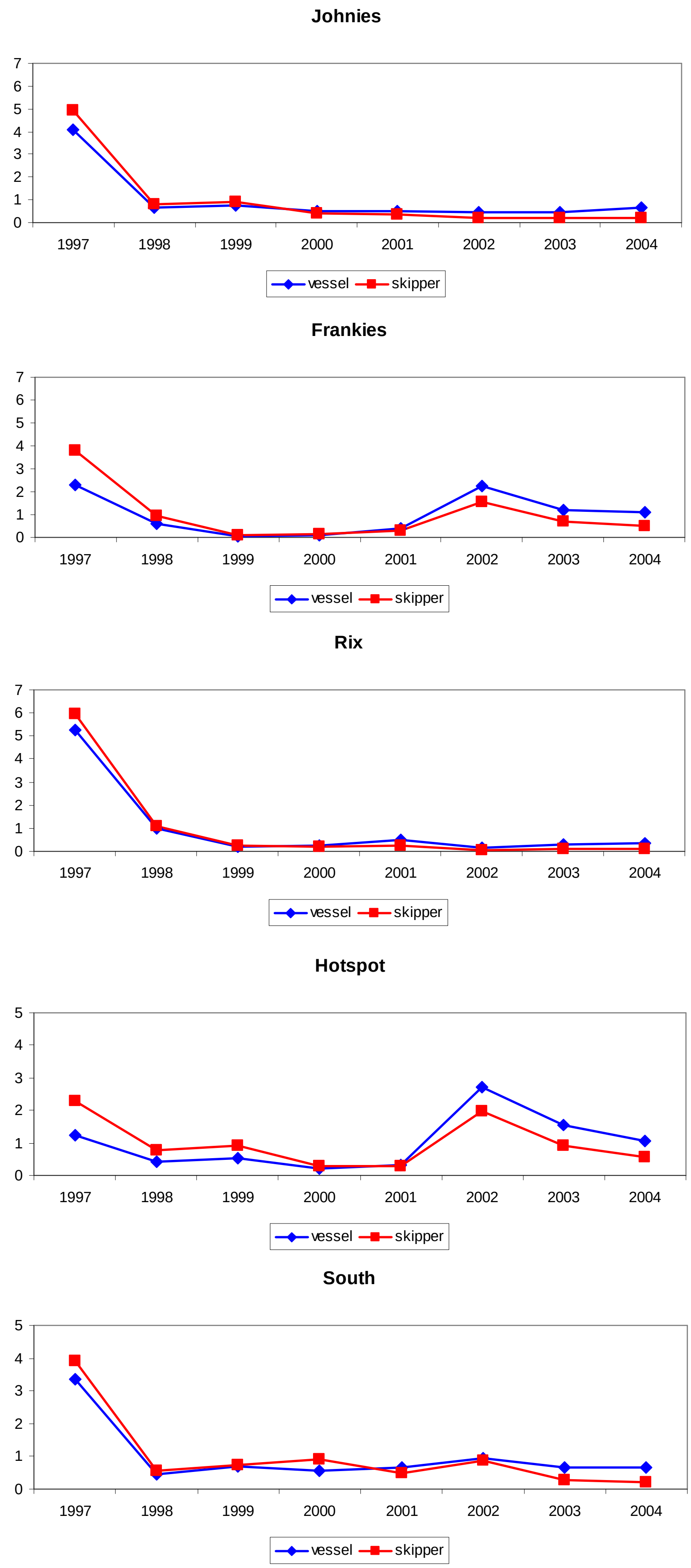

Figure 6. Standardised CPUE indices normalised to their mean for each aggregation. The GLM fitted is: *year+agg+month+year.agg* together with either a main effect for skipper or for vessel.

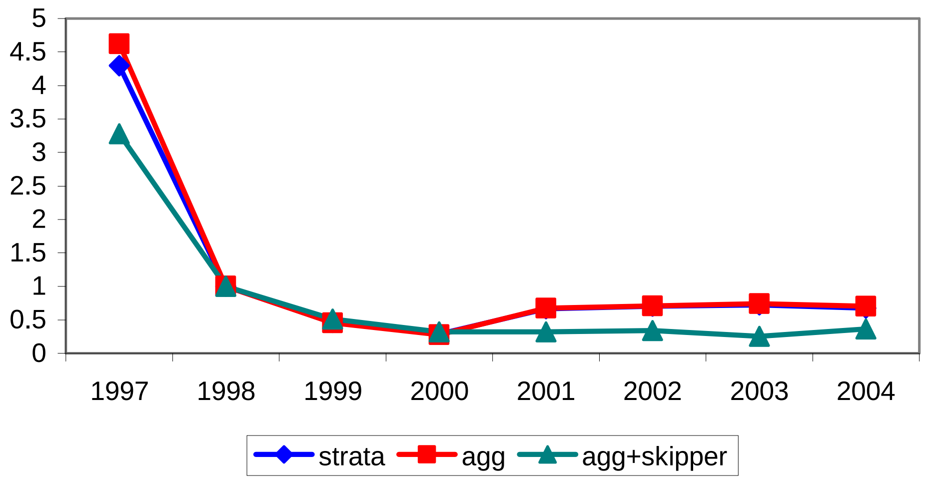

Figure 7. Standardised CPUE indices when the GLMs *year+strata*, *year+agg* and *year+agg+skipper* are fitted to the commercial orange roughy data.

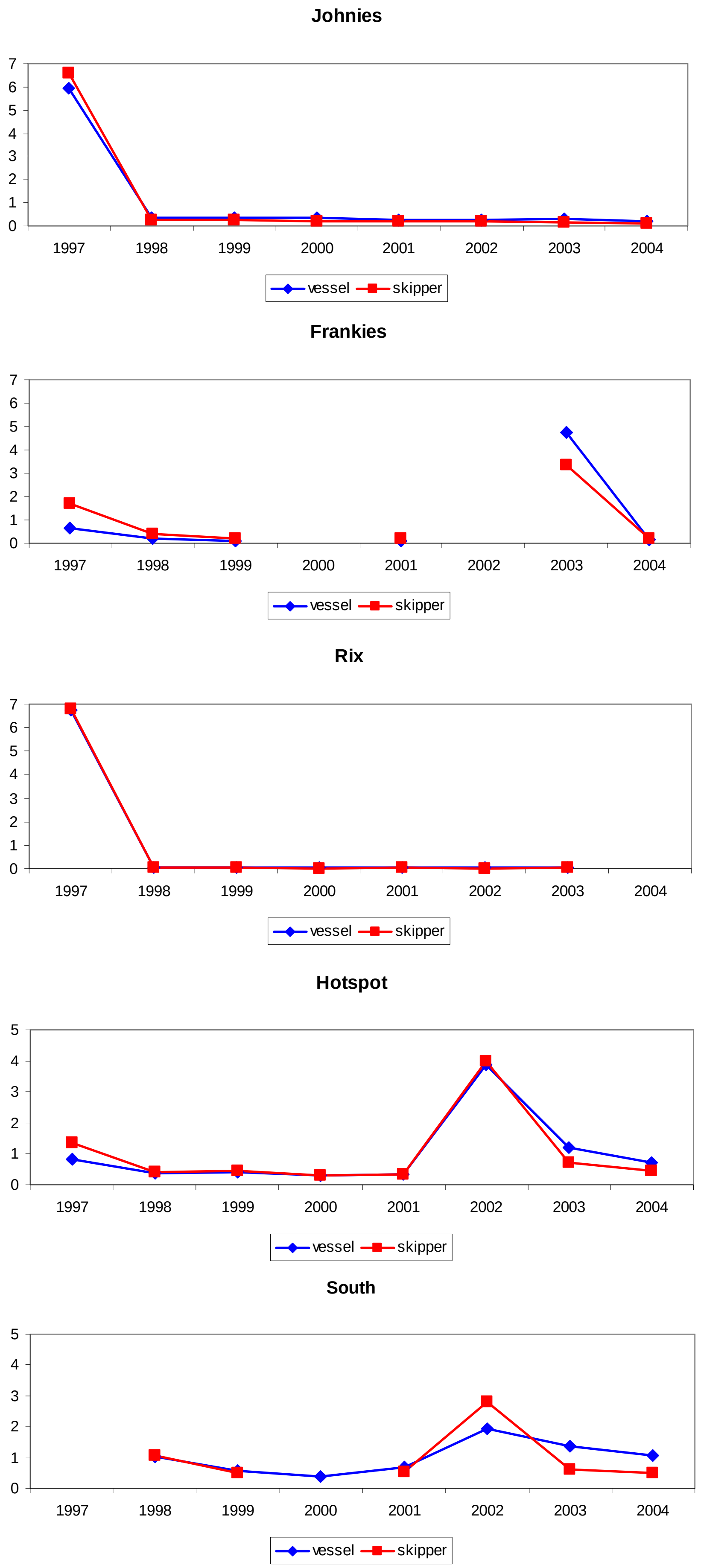

Figure 8. Standardised CPUE indices normalised to their mean for each aggregation. The GLM fitted is: *year+sub-agg+month+year.sub-agg* together with either a main effect for skipper or for vessel.