SCIENTIFIC COMMITTEE
THIRD REGULAR SESSION

13-24 August 2007
Honolulu, United States of America

REGIONAL WEIGHTING FACTORS FOR YELLOWFIN TUNA IN WCP-CA STOCK ASSESSMENTS

WCPFC-SC3-ME SWG/WP-1

Simon D. Hoyle[1] and Adam D. Langley[1]

[1] Oceanic Fisheries Program, Secretariat of the Pacific Community, Noumea, New Caledonia.

Regional weighting factors for yellowfin tuna in WCP-CA stock assessments

Simon D. Hoyle and Adam D. Langley

Oceanic Fisheries Program

Secretariat of the Pacific Community

Nouméa, New Caledonia

Introduction

The Japanese longline fishery abundance indices, highly influential in the yellowfin and bigeye Multifan-CL stock assessment models, are calculated from catch and effort data using generalized linear modelling (GLM). An index is estimated for each region. Regional weighting is applied to adjust these independently-estimated abundance indices for the relative abundances in each region (Langley *et al.* 2005). Its calculation is based on the assumption that catchability is the same in each region. In principle, abundance indices could be weighted automatically by fitting the GLM to all regions simultaneously and assuming uniform catchability. However, computer memory constraints do not permit this.

Although it is assumed that catchability estimates can be equalized across the regions, catchability varies with hooks between floats (HBF) and other explanatory variables, and abundance varies with latitude and longitude (also referred to henceforth as 'latlong'). In analyses of this type, catchability and abundance are confounded; the temporal index and spatial effects are assumed to reflect abundance, and other effects to reflect catchability. HBF can also be thought of as reflecting abundance by depth, given the distribution of 'habitat', but it is simpler to consider it in terms of catchability and view the population in two dimensions.

The following described the methods used in 2006 (Hampton *et al.* 2006) for the reweighting GLM and the index of abundance GLMs.

1. Data selection.

The 2006 reweighting GLM aggregated Japanese longline data from 1960 to 1986 into strata by latlong (5 degree square), quarter, and HBF. The period was selected because it represented maximum spatial operation of the longline fleet in the WCPO; i.e. followed the period of initial fishery expansion during the 1950s and preceded the contraction of fishing effort that occurred following the declaration of EEZs.

Strata after 1975 without HBF were omitted, and strata before 1975 without HBF were assigned HBF of 5. Strata at the latlong.qtr.HBF level with zero catch were omitted. Latlongs were omitted if they had cumulative catch less than 5000 fish, or data from 10 or fewer quarter / HBF strata.

For the 2006 indices of abundance, data from 1952 to 2004 were used. The stratification approach was the same as for the reweighting GLM, but data selection differed in that latlongs with cumulative catch less than 5000 fish were included, and the threshold number of quarter / HBF strata per latlong was 5 rather than 10. A separate analysis was carried out for each region.

2. Analysis methods

In the 2006 reweighting GLM, the regional weighting factors W_R were calculated by exponentiating and then summing the estimated coefficients of the latlong cells ($a_{R,i}$) included in each region (R). $W_R = \sum_{i=1}^{n_R} \exp\left(a_{R,i}\right)$, where n_R is the number of latlong cells included in the region.

The region-specific indices from the abundance GLM were each normalized so that the average of the series was 1. Then, the regional adjustment factors adj_R were calculated by dividing the regional weighting factors by the average index during the period used in the reweighting analysis: 1960-1986 in this case: $adj_R = \frac{W_R}{\sum_{i=1}^{n} I_{rt_i,R} / n}$,

where $I_{t,R}$ is the unweighted normalized index, W_R is the weight for region R, and rt_i is the ith year in the regional weighting series. Finally the normalized indices were scaled by applying the adjustment factors; i.e., $I'_{t,R} = I_{t,R} \cdot adj_R$, where $I'_{t,R}$ is the reweighted index,

Possible improvements

With the approach described, weights will be biased if both abundance and the number of strata in the model change through time. This applies both to CPUE trends and to quarterly variation. For example, if abundance in a region is high early in a time series and more data (strata) are available from the region during this period, the relative weight of the region will be higher than if data were available from later in the series. Such bias can be eliminated by fitting a time effect, but if trends vary between regions then a region.time effect must be fitted.

In addition, the 2006 reweighting GLM shared the HBF parameter across all regions, but HBF has a different relationship with catchability in each region. This suggests that the HBF parameter should be estimated separately by region, so that the weight of a region will be not biased by the catch rate in the HBF strata with more data. However, estimating separate HBF effects by region makes the catchability estimates incompatible, and an assumption must be made about catchability’s relationship with HBF across regions. Catchability may be assumed to be the same across regions at a particular HBF. Alternatively, HBF may be assumed to be the same at an appropriately weighted average of the HBF levels.

Methods

The effects of including region.time and region.HBF in the analysis were examined, by applying the following approach to the catch of yellowfin tuna.

1. Strata in the reweighting analysis were counted by year and region

2. Relative weights were estimated for alternative time periods (1960 to 1986, 1960 to 1974, 1975 to 1986, and 1966 + 1975-1986), and using alternative data selection criteria (number of yellowfin > 1000, 2500, 5000; number of strata > 5, 10). Given the data included, the ability of alternative selection methods to estimate weights for all regions was examined.

3. The interaction of region and time was included in the analysis. Relative abundances by year and region were estimated using the sum of the spatial effects plus the region.year and region.qtr effects for that time period and region.

$$\ln\left(Catch / Hooks\right)_{i,y,qtr,HBF} \sim a_i + b_{R,y} + c_{R,qtr}$$

$$W_R = \sum_{i=1}^{n_R} \sum_{y_{min}}^{y_{max}} \sum_{qtr=1}^{4} \exp\left(a_{R,i} + b_{R,y} + c_{R,qtr}\right)$$

4. Regional weights were estimated with region-specific HBF for the years 1966 and 1975-1986, when HBF was available. Catchability was assumed to be the same for shallow sets with HBF of 5.

$$\ln\left(Catch / Hooks\right)_{i,y,qtr,HBF} \sim a_i + b_{R,y} + poly\left(d_{R,HBF}, 3\right)$$

$$W_R = \sum_{i=1}^{n_R} \sum_{y=y_{min}}^{y_{max}} \exp\left(a_{R,i} + b_{R,y} + poly\left(d_{R,HBF=5}, 3\right)\right)$$

5. Steps 3 and 4 were combined to give regional weights based on the interactions of year, quarter, and HBF with region.

$$\ln\left(Catch / Hooks\right)_{i,y,qtr,HBF} \sim a_i + b_{R,y} + c_{R,qtr} + poly\left(d_{R,HBF}, 3\right)$$

$$W_R = \sum_{i=1}^{n_R} \sum_{y=y_1}^{y_{\max}} \sum_{qtr=1}^{4} \exp\left(a_{R,i} + b_{R,y} + c_{R,qtr} + poly\left(d_{R,HBF=5}, 3\right)\right)$$

The CPUE time series were estimated by region. Each time series was normalized, and the mean of the normalized values was calculated for the period used in the weighting factor analysis. Regional index adjustment factors were calculated by dividing the regional weighting factors by the means calculated above.

$$adj_R = \frac{W_R}{\sum_{i=1}^{n} I_{rt_i,R} / n}$$

6. Diagnostics were examined to check for violation of the assumptions of the analysis.

Results

The stratification approach used for the 2006 model, based on the period 1960 to 1986, resulted in more strata for 1966 and the period from 1975 to 1986, because data from those years were stratified by HBF whereas the rest of the data were not (Table 1). Because each stratum has the same weight in the model, the HBF-stratified data from 1966 and 1975-86 had more influence on the relative weighting of the regions. Accordingly, we examined relative weights based on a period with data largely lacking HBF information (1960-1974), a continuous period with HBF information (1975-1986), and all the data up to 1986 with HBF information (1966 + 1975-1986).

Spatial coverage of data from these periods varied (Figure 1 to Figure 4). Given the small amount of effort in region 6 between 1975 and 1986, the weight for this region was very low when data before 1975 were excluded (Table 2).

Selection criteria interacted with the length of the time series to influence the relative weights of the regions. When the time series was shorter all latlongs had less catch, taking some below the 5000 fish threshold, or below the 5 time/HBF strata threshold. This occurred more significantly in region 6, giving the region less weight. Relative weights were rebalanced somewhat by reducing the selection thresholds to 1000 fish and 5 time/HBF strata.

Time trends and seasonal effects were apparent in all regions (Figure 5), suggesting the need to include time effects in the reweighting standardization. Including the full time effect region*year*qtr was not possible due to memory constraints, so the model was fitted with region*yr + region*qtr (Table 3).

HBF appeared to affect CPUE differently by region (Figure 6), suggesting that region*HBF should be included in the analysis. Regional weights were estimated with region-specific HBF, assuming that catchability among regions was the same for shallow sets with HBF 5.

Finally, the region.HBF analysis was combined with the region.yr+qtr analysis to give a region.yr + region.qtr + region.HBF analysis. This model had the best AIC (Table 4). If HBF is included then years without true HBF should be omitted. The period 1966 + 1975-1986 was selected.

Means of normalized abundance indices were calculated for the period in the reweighting GLM (Table 5), and used to calculate the regional index adjustment factors for all options (Table 6). Results did not differ substantially from those used in 2006. Weight for region 3 was 3% lower, regions 2 and 5 weights were each 2% higher, and region 4 weight was 1% higher (Figure 7).

Diagnostics showed slight skewness in the residuals (Figure 8, Figure 9), but no serious breaches the assumptions of normality and homoscedasticity (Figure 10).

Discussion

Regional weighting factors are influential components of the stock assessments for yellowfin and bigeye tuna in the WCPFC. Catch rate by region has changed through time, and in relation to season and HBF. We have therefore investigated the effect of including these factors in the regional weighting factor standardization for yellowfin tuna. Results of the analysis suggest that including these factors improves the model. The model using data from 1966 and 1975-1986, and fitting to region.yr, region.qtr, and region.HBF was selected as the best model based on having substantially the best AIC. However, altering the model has not substantially changed the estimated weighting factors.

References

Langley, A. Bigelow, K. Maunder, M.N. and Miyabe, N. (2005) Longline CPUE indices for bigeye and yellowfin in the Pacific Ocean using GLM and statistical habitat standardisation methods. WCPFC–SC1 SA WP–8, 40p.

Hampton, J., Langley, A., Kleiber, P.(2006). Stock assessment of yellowfin tuna in the western and central Pacific Ocean, including an analysis of management options. WCPFC–SC2 SA WP–1, 103p.

Tables

Table 1: Number of strata in the 2006 regional reweighting analysis, by year and region, using the criteria of at least 5000 fish and 10 quarter/HBF strata. The only year with HBF data before 1975 was 1966.

	1	2	3	4	5	6
1960	97	72	167	133	54	64
1961	93	46	175	141	45	62
1962	96	66	205	159	78	66
1963	93	66	179	176	81	74
1964	98	61	190	179	88	70
1965	120	63	191	174	87	65
1966	428	135	623	428	242	204
1967	106	69	208	173	88	66
1968	106	73	213	152	86	55
1969	107	62	198	156	74	34
1970	110	75	182	161	73	35
1971	104	69	195	154	82	42
1972	77	57	175	147	72	43
1973	94	64	185	145	62	22
1974	92	55	202	146	77	18
1975	320	128	825	502	76	15
1976	387	277	887	746	82	16
1977	411	265	951	613	51	16
1978	366	296	901	575	80	11
1979	403	365	854	670	126	7
1980	448	239	951	715	153	33
1981	493	279	979	561	243	29
1982	454	254	768	632	227	35
1983	442	233	612	503	203	25
1984	425	202	751	557	170	28
1985	362	240	758	572	182	19
1986	462	246	629	480	165	38

Table 2: Relative weights estimated using alternative data combinations

	Min cum catch	Min qtrs	1	2	3	4	5	6
1960-1986	5000	10	0.026	0.019	0.515	0.249	0.124	0.066
1960-1974	1000	5	0.024	0.016	0.526	0.252	0.109	0.074
		10	0.025	0.016	0.518	0.256	0.110	0.075
	2500	5	0.022	0.014	0.523	0.258	0.111	0.070
		10	0.022	0.014	0.523	0.258	0.111	0.070
	5000	5	0.021	0.014	0.526	0.263	0.115	0.060
		10	0.021	0.014	0.526	0.263	0.115	0.060
1975-1986	1000	5	0.027	0.024	0.549	0.266	0.132	0.001
		10	0.027	0.025	0.557	0.270	0.120	0.001
	2500	5	0.028	0.022	0.565	0.265	0.119	0.001
		10	0.028	0.022	0.565	0.265	0.119	0.001
	5000	5	0.024	0.020	0.590	0.260	0.106	0.001
		10	0.024	0.020	0.590	0.260	0.106	0.001
1966 + 1975-1986	1000	5	0.026	0.022	0.517	0.251	0.130	0.053
		10	0.027	0.023	0.538	0.261	0.116	0.034
	2500	5	0.025	0.020	0.534	0.257	0.127	0.037
		10	0.026	0.021	0.548	0.263	0.113	0.029
	5000	5	0.026	0.018	0.560	0.257	0.128	0.010
		10	0.027	0.018	0.570	0.261	0.113	0.010

Table 3: Include parameters interacting with region in the analysis, minimum cumulative catch of 1000 yellowfin, and 5 strata.

Years	Interactions with region	1	2	3	4	5	6
1960-1986	none	0.027	0.020	0.512	0.239	0.120	0.082
1960-1974	none	0.024	0.016	0.526	0.252	0.109	0.074
1975-1986	none	0.027	0.024	0.549	0.266	0.132	0.001
1966+1975-1986	none	0.026	0.022	0.517	0.251	0.130	0.053
1960-1986	yr, qtr	0.027	0.018	0.512	0.246	0.122	0.076
1960-1974	yr, qtr	0.025	0.016	0.517	0.25	0.117	0.074
1975-1986	yr, qtr,	0.027	0.022	0.551	0.264	0.133	0.003
1966+1975-1986	yr, qtr	0.026	0.02	0.52	0.245	0.13	0.059
1960-1986	HBF	0.021	0.017	0.529	0.245	0.11	0.077
1960-1974	HBF	0.024	0.016	0.526	0.25	0.111	0.074
1975-1986	HBF	0.017	0.022	0.606	0.245	0.108	0.001
1966+1975-1986	HBF	0.018	0.02	0.565	0.239	0.114	0.045
1960-1986	yr,qtr, HBF	0.023	0.017	0.529	0.247	0.114	0.07
1960-1974	yr,qtr, HBF	0.025	0.017	0.518	0.249	0.118	0.074
1975-1986	yr,qtr, HBF	0.017	0.02	0.581	0.267	0.112	0.002
1966+1975-1986	yr,qtr, HBF	0.018	0.019	0.547	0.246	0.119	0.053

Table 4: Model selection using Akaike information criterion (AIC).

Model (1966 + 1975-1986)	AIC
Basic	75886
+ yr, qtr	73272
+ HBF	75044
+ yr + qtr + HBF	72652

Table 5: Mean of normalized yellowfin CPUE time series by region for each of the periods used in the regional weighting analysis.

Period	1	2	3	4	5	6
1960-1986	1.069	1.084	1.029	0.998	0.836	0.788
1960-1974	1.364	1.159	1.006	1.136	0.995	0.899
1975-1986	0.665	0.995	1.083	0.856	0.631	0.604
1966+1975-1986	0.697	0.982	1.092	0.906	0.661	0.631

Table 6: Regional index adjustment factors for all options. The factors used in 2006 and the optimal set of factors selected in this analysis are in bold.

Period	1	2	3	4	5	6
2006	**0.024**	**0.017**	**0.490**	**0.242**	**0.146**	**0.081**
1960-1986	0.025	0.018	0.484	0.233	0.140	0.101
1960-1974	0.018	0.014	0.540	0.229	0.113	0.085
1975-1986	0.037	0.022	0.464	0.284	0.191	0.002
1966+1975-1986	0.034	0.021	0.434	0.254	0.180	0.077
1960-1986	0.025	0.016	0.484	0.240	0.142	0.094
1960-1974	0.019	0.014	0.532	0.228	0.122	0.085
1975-1986	0.037	0.020	0.464	0.282	0.192	0.005
1966+1975-1986	0.034	0.019	0.435	0.247	0.180	0.085
1960-1986	0.019	0.015	0.502	0.240	0.129	0.095
1960-1974	0.018	0.014	0.540	0.227	0.115	0.085
1975-1986	0.024	0.021	0.525	0.268	0.160	0.002
1966+1975-1986	0.024	0.019	0.483	0.246	0.161	0.067
1960-1986	0.021	0.016	0.499	0.241	0.133	0.089
1960-1974	0.019	0.015	0.533	0.227	0.122	0.084
1975-1986	0.026	0.021	0.497	0.285	0.167	0.004
1966+1975-1986	**0.024**	**0.019**	**0.459**	**0.250**	**0.164**	**0.083**

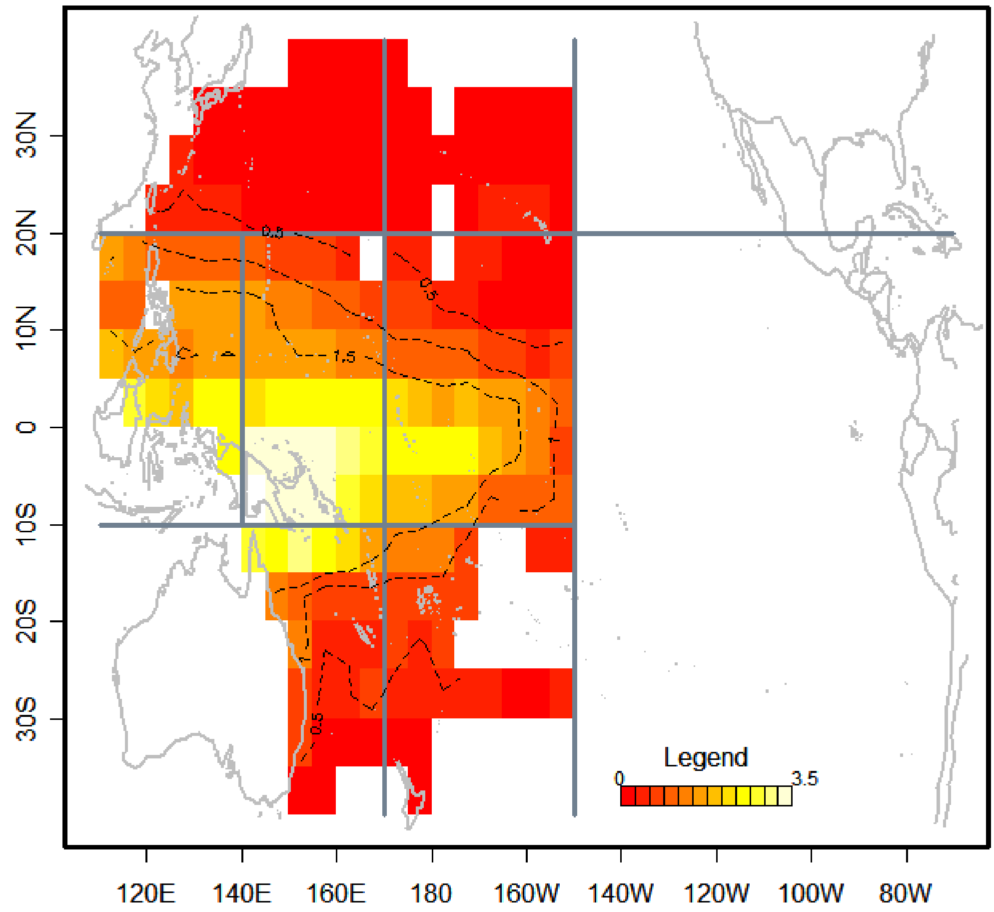

Figure 1: Heat map of relative CPUE by 5 degree square, estimated using the method used in 2006, with data from 1960 to 1986, cumulative catch of at least 5000 yellowfin, and at least 10 quarters.

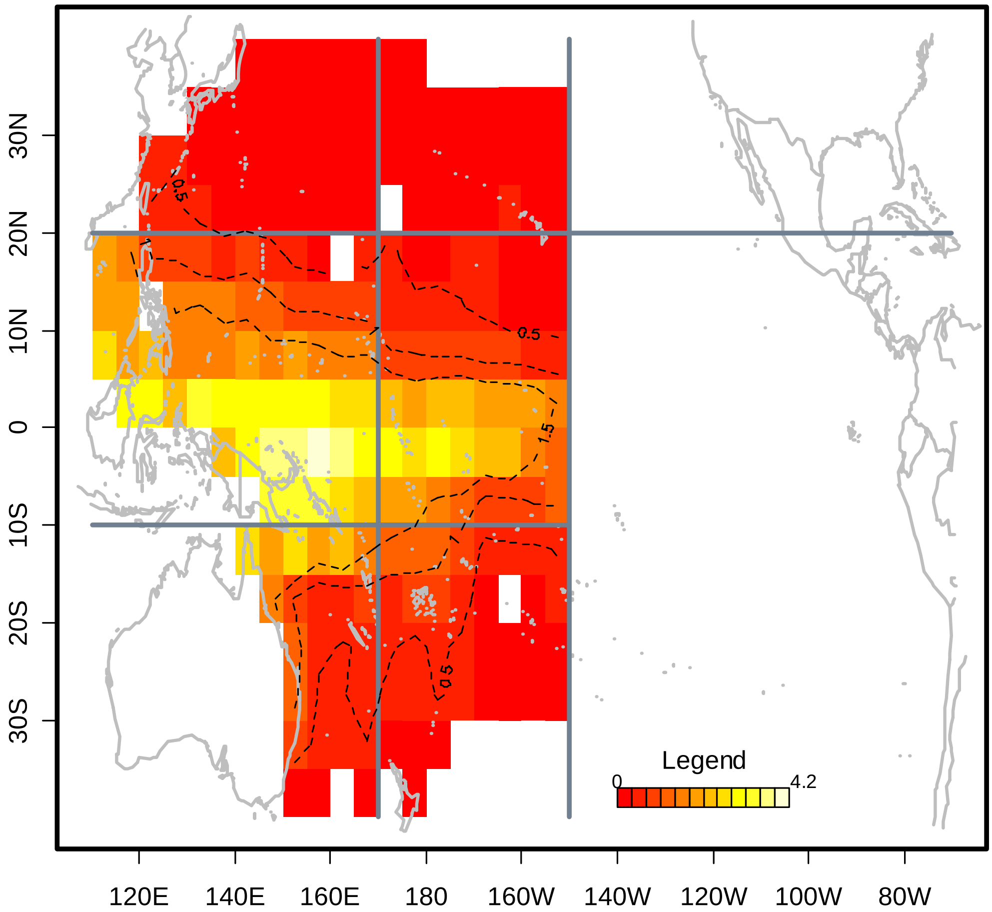

Figure 2: Estimates using data from 1960 to 1974, cumulative catch of at least 1000 yellowfin, and at least 5 quarters of data

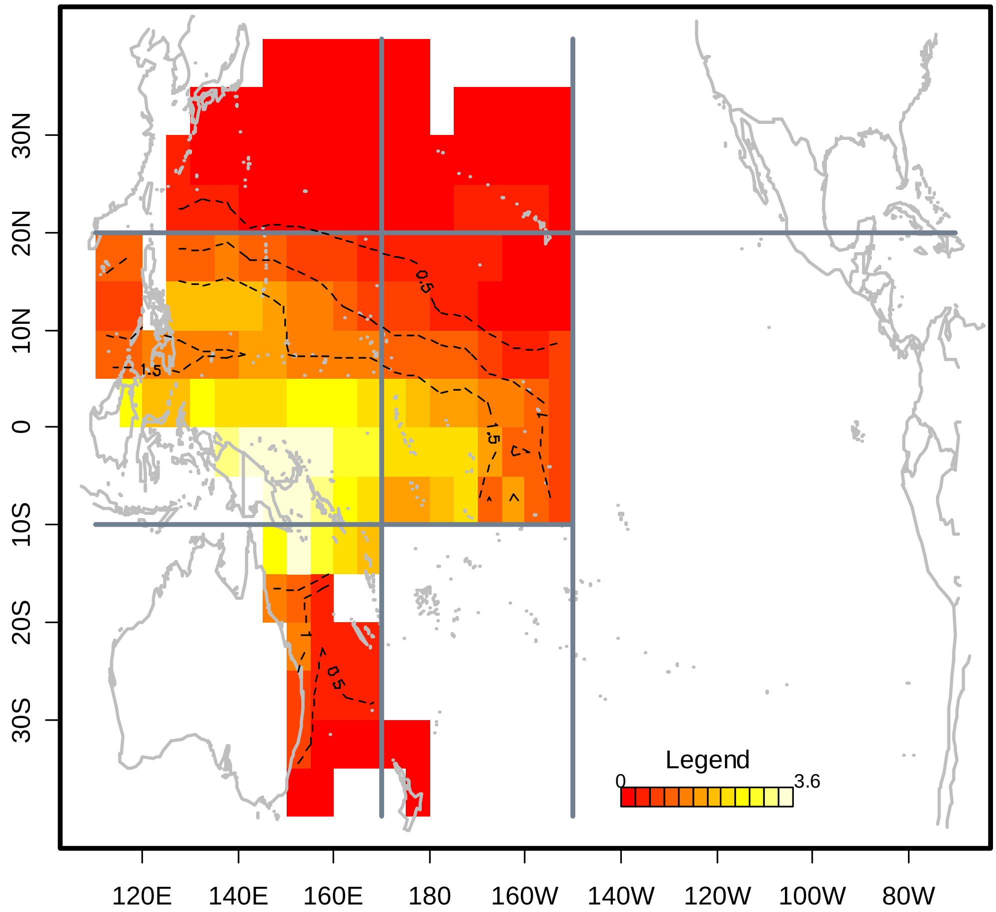

Figure 3: Estimates using data from 1975 to 1986, cumulative catch of 1000 yellowfin and at least 5 quarters

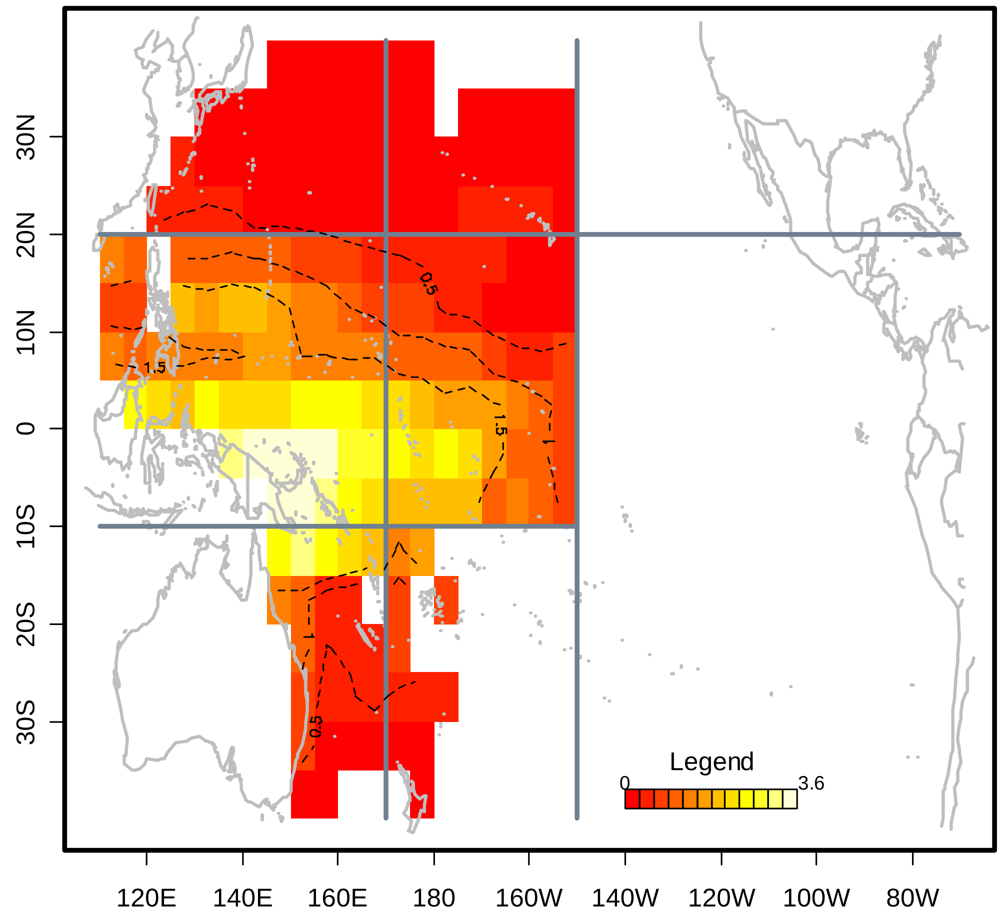

Figure 4: Estimates using data from 1966 + 1975-86, cumulative catch of 1000 yellowfin, and at least 5 quarters

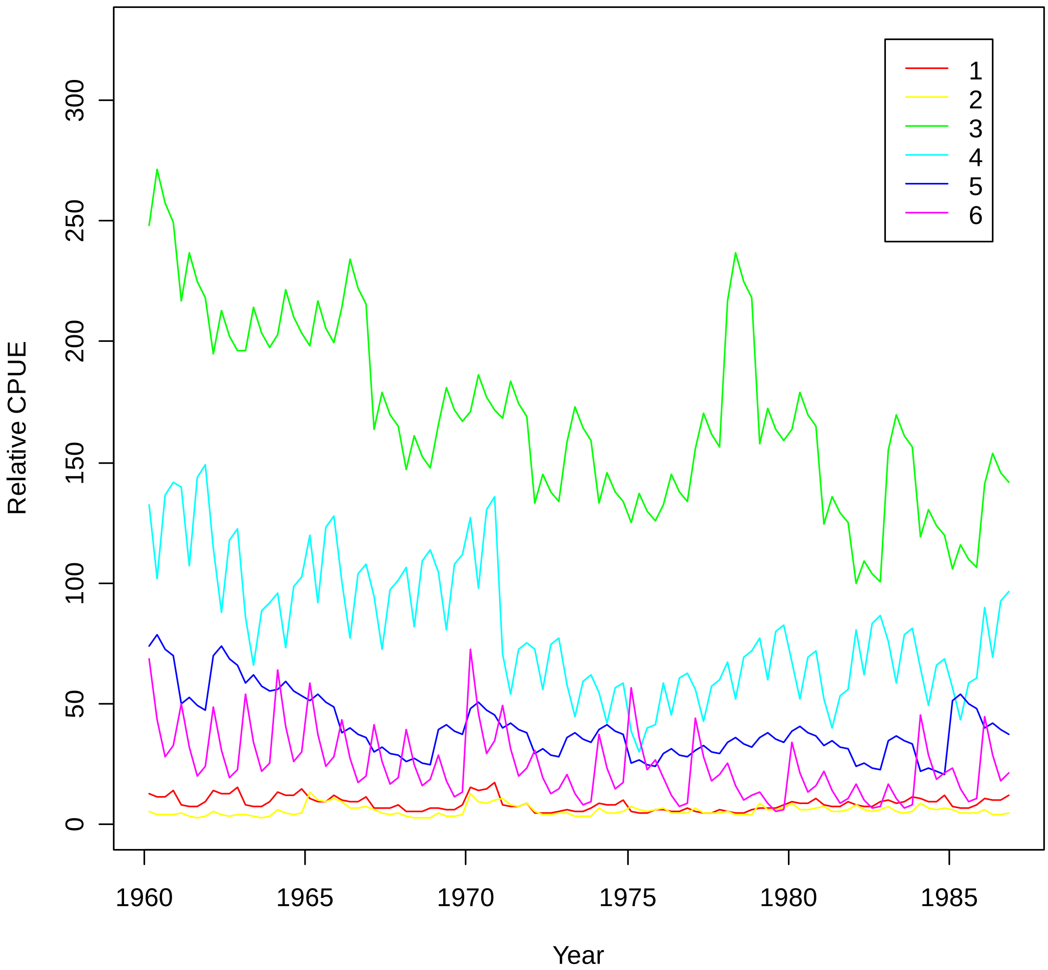

Figure 5: Catch rate of yellowfin by region for 1960-1986

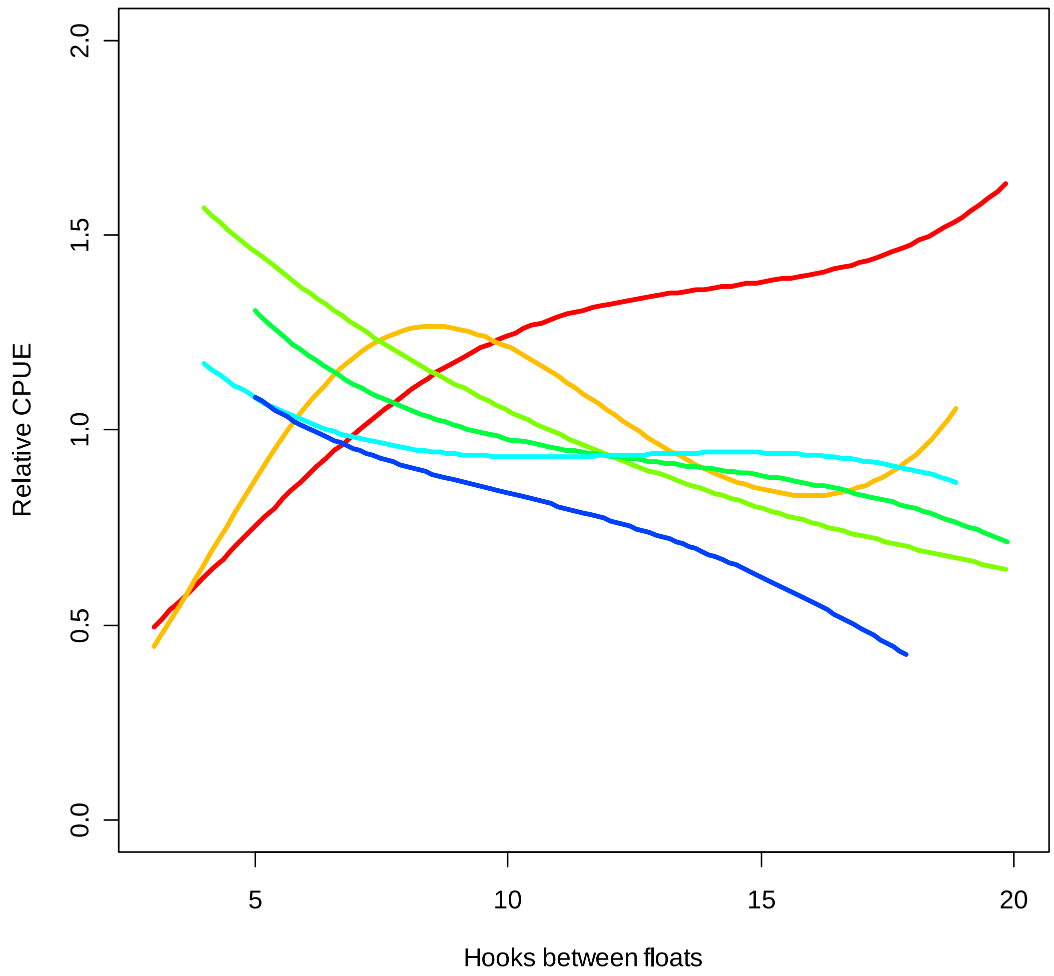

Figure 6: Relationship between HBF and catch rate of yellowfin in each of the 6 regions

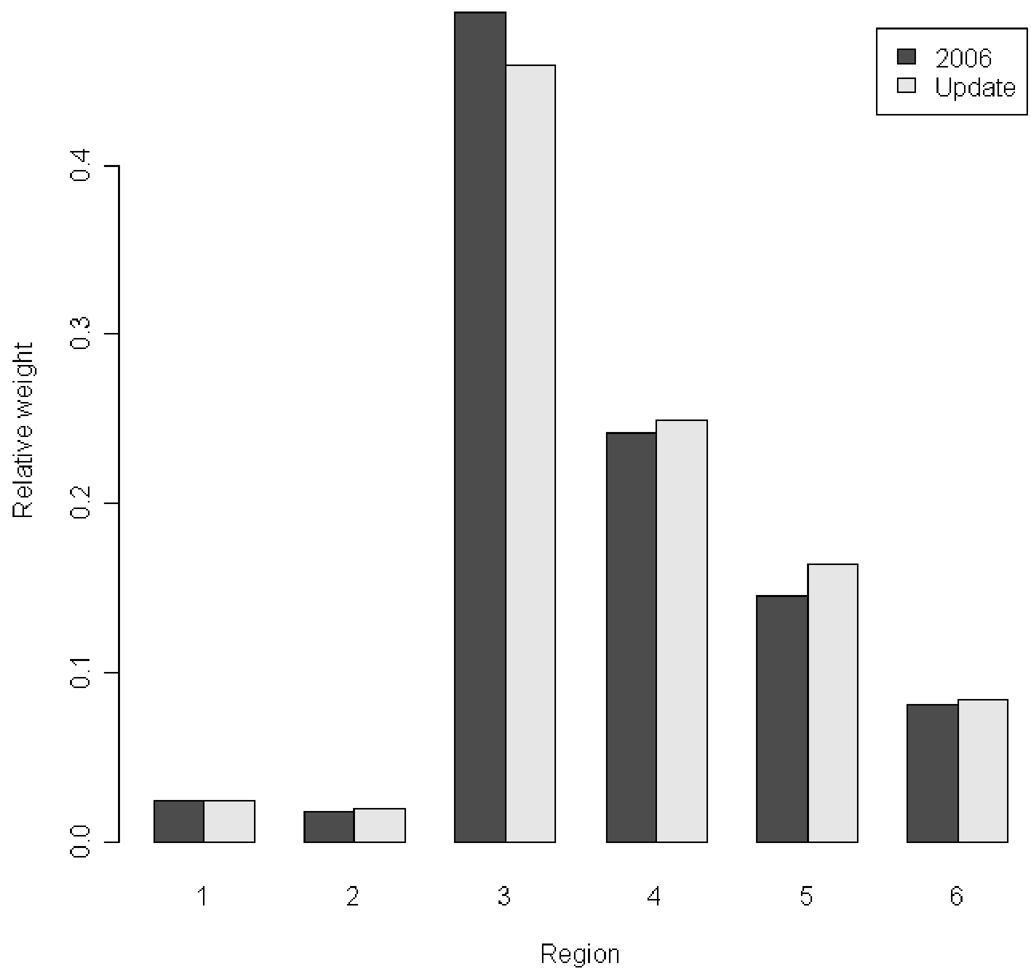

Figure 7: Comparison of 2006 and updated normalized CPUE index adjustment factors by region.

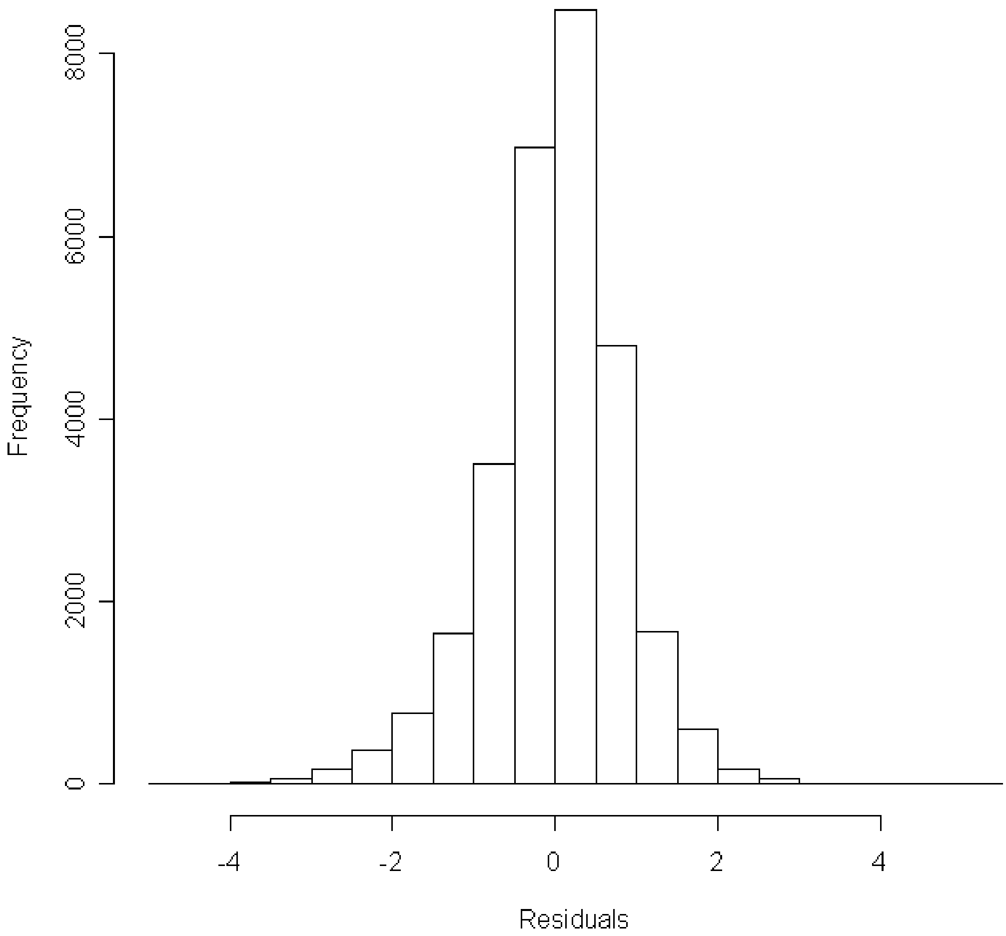

Figure 8: Frequency histogram for residuals from the selected model.

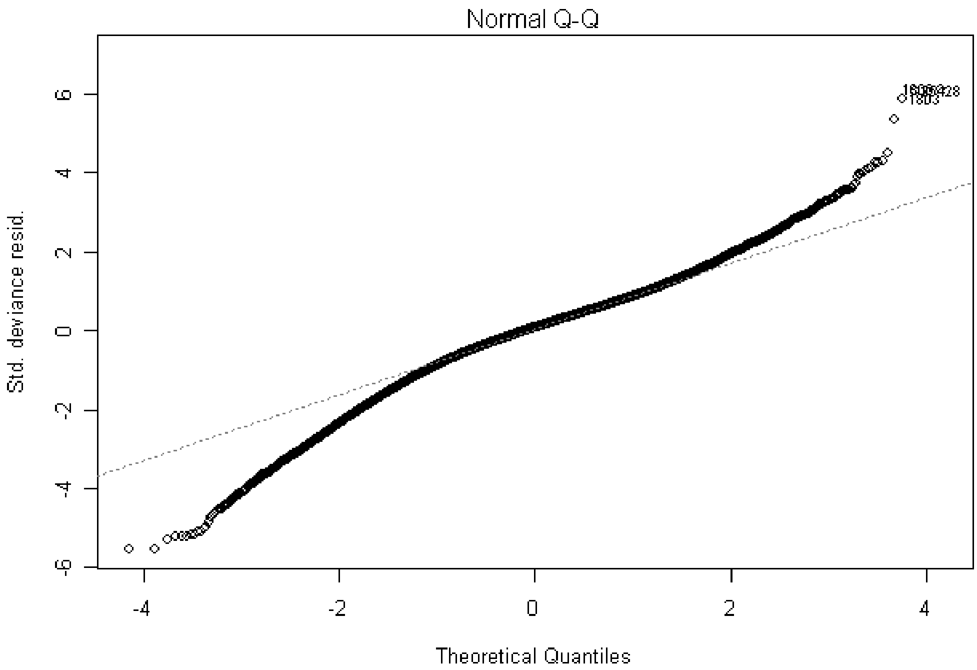

Figure 9: Q-Q plot for the selected model.

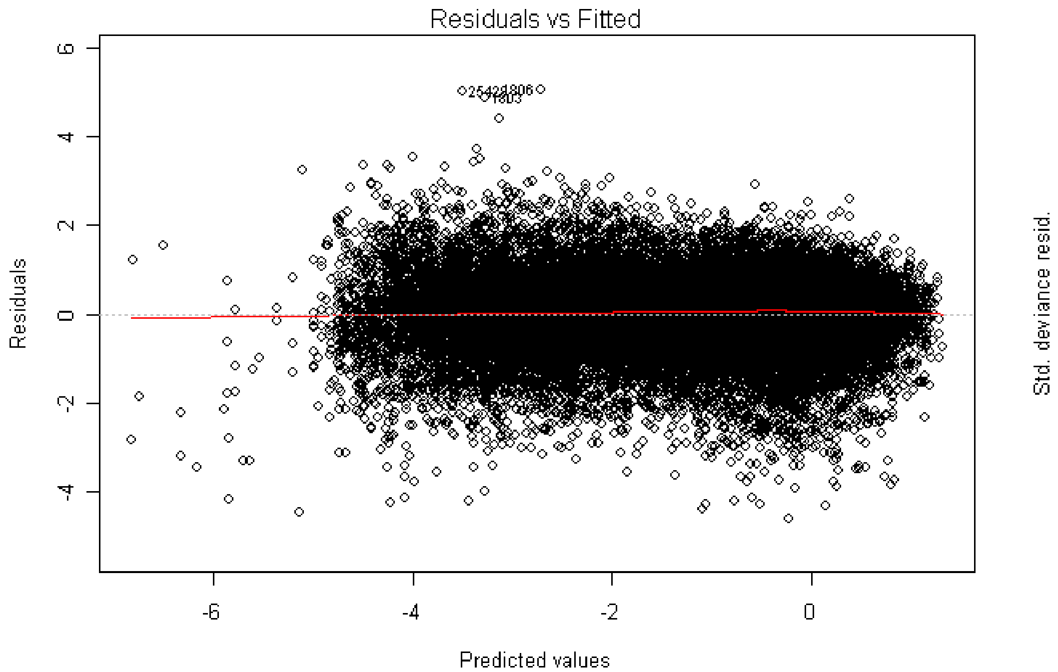

Figure 10: Plot of fitted values versus residuals for the selected model.