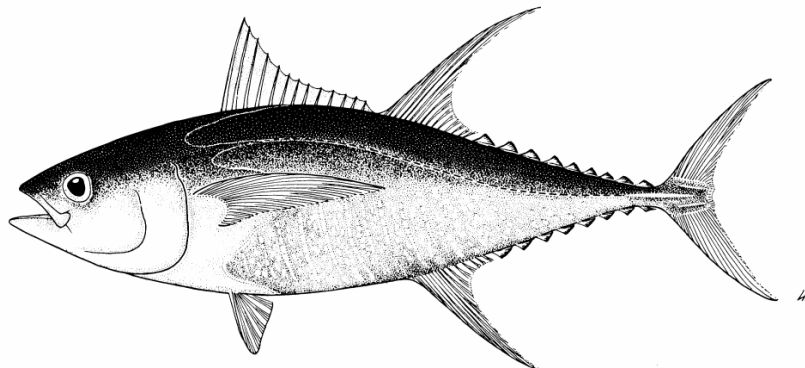




**Longline CPUE indices for bigeye and yellowfin in the  
Pacific Ocean using GLM and statistical habitat  
standardisation methods.**



**Adam Langley<sup>1</sup>, Keith Bigelow<sup>2</sup>, Mark Maunder<sup>3</sup> and Naozumi Miyabe<sup>4</sup>**

<sup>1</sup>Oceanic Fisheries Programme, Secretariat of the Pacific Community, Noumea, New Caledonia.

<sup>2</sup>National Marine Fisheries Service, Honolulu, Hawaii, USA.

<sup>3</sup>Inter-American Tropical Tuna Commission, La Jolla, California, USA.

<sup>4</sup>Japan National Research Institute of Far Seas Fisheries, Shimizu, Japan.

July 2005

# 1 Introduction

Since the early 1950s, the Japanese longline fleet has been the dominant fleet operating in the Pacific Ocean. The catch and effort data from the fleet are a critical input into the stock assessments for both bigeye tuna (*Thunnus obesus*) (Hampton et al. 2005a) and yellowfin tuna (*Thunnus albacares*) (Hampton et al. 2005b). The catch and effort series represent the principal indices of relative abundance for the longline exploitable biomass. However, over the history of the fishery, there have been systematic changes in the operation of the Japanese longline fleet that are likely to have influenced the catchability of the two species, in particular changes in the geographic area fished and the configuration of the longline gear, most notably increases in the number of hooks between floats (HBF).

To account for these temporal changes in species-specific catchability of the longline fishery, the Japanese effort data have been standardized using a variety of approaches; most recently using generalised linear modeling (GLM, Langley 2003) and statistical habitat based modeling (statHBS, Bigelow et al. 2004) techniques. The resulting region-specific standardised effort series are then integrated into the MULTIFAN-CL (MFCL) assessments of yellowfin and bigeye in the WCPO and the Pacific-wide bigeye assessments (Hampton et al. 2003).

This report documents the standardised analyses undertaken to provide indices for the 2005 stock assessments. The approach used to calculate the GLM indices is equivalent to previous years, while some modifications have occurred in the statHBS approach. There has also been a restructuring of the regional stratification of the MFCL model areas and the resulting region-specific indices are consistent with these new boundaries. Further, a significant change in the MFCL assessments from 2004 is the application of the results of the effort standardisation to derive species-specific regional scaling factors for the two models (Hampton et al. 2005a, 2005b). These parameters are influential on the results of the overall stock assessments as they have resulted in a down-weighting of the biomass in the more peripheral regions of the two assessments. The approach used to calculate the regional scaling factors is also documented in this paper.

## 2 Methods

The MFCL models for bigeye and yellowfin are stratified into six regions (1–6) and the Pacific-wide bigeye assessment includes two additional model regions (7 and 8) (Figure 1). This section describes the methodology used to calculate standardised CPUE indices for bigeye and yellowfin in each model region. The methodology is similar to previous years with indices derived using a generalised linear modelling (GLM) approach (Langley 2003) and the statistical habitat based modeling approach (statHBS) (Bigelow et al. 2004).

In addition, the methods used to calculate the relative biomass between MFCL model regions are described. These scaling factors are an important input to the MFCL model as the assessment model scales the individual CPUE indices between regions.

### 2.1 Data summary

The data included in the analysis were longline catch and effort data from the Japanese longline fleet operating in the entire Pacific Ocean from 1952 to 2004. Data were provided by the National Research Institute of Far Seas Fisheries. From 1962 to 2004, these data are aggregated by year, month, and degree of latitude and longitude, while earlier data are aggregated at the five degree spatial resolution. Catches of bigeye and yellowfin were recorded as number of fish caught, effort expressed as the number of hooks set.

From 1975 onwards, the data are further aggregated by the configuration of the longline gear, i.e., the number of hooks deployed between each float (HBF). Only limited HBF data are available for most years prior to 1975, although for the purpose of this analysis it was assumed that all longline sets during the early period were similar in gear configuration to that deployed during the early 1970s, i.e., essentially shallow sets deploying five (5) HBF.

Overall, the data set represents at least 70–80% of the annual Japanese longline effort in the Pacific Ocean, with the exception of lower coverage during the mid 1960s to mid 1970s (about 50%).

For descriptive purposes, annual trends in catch and effort and nominal CPUE for both bigeye and yellowfin were summarised by MFCL region.

## 2.2 Generalised linear models (GLM)

For both species, region-specific GLM indices were calculated by quarter for 1952–2004. The data set included the following variables: year, quarter, latitude (5 degree), longitude (5 degree), HPB, hundreds of hooks, bigeye catch (number), and yellowfin catch (number).

For each region, data were aggregated by year, quarter, 5 degree latitude/longitude cell and HBF category. Latitude/longitude cells with a small number of records (5 or less) were excluded from the data set.

**Table 1. Number of year/quarter and latitude/longitude categories included in the region-specific GLMs for bigeye and yellowfin.**

Variable	MFCL Region							
	1	2	3	4	5	6	7	8
<b>Yr/qtr</b>	212	211	212	212	210	189	182	201
<b>Lat/long</b>	39	32	61	48	28	43	28	158

The dependent variable in the GLMs is the natural logarithm of the catch (in numbers) of the species. Records with a zero catch of the species were excluded. The GLMs all had an equivalent model structure, including the categorical variables year/quarter and latitude/longitude and the HBF, number of hooks, and catch rate (number per hundred hooks) of the other species (bigeye or yellowfin) as third-order polynomial functions. For bigeye, the natural logarithm of the catch (in numbers) in year/quarter ( $u$ ) and five degree latitude/longitude cell ( $v$ ) is predicted as follows.

$$\ln(\text{CATCH}_{(u,v)}) = a\text{YRQTR}_{(u)} + b\text{LATLONG}_{(v)} + c\text{HBF} + d\text{HBF}^2 + e\text{HBF}^3 + f\text{HOOKS} + g\text{HOOKS}^2 + h\text{HOOKS}^3 + i\text{CPUE\_YFT} + j\text{CPUE\_YFT}^2 + k\text{CPUE\_YFT}^3 + \varepsilon_{(u,v)}$$

The equivalent GLM was applied to predict region-specific yellowfin catches, with the substitution of bigeye CPUE as a predictor variable in the model.

The CPUE index was the exponentiated year/quarter coefficients ( $a$ ) from the region-specific GLM. The relationships between predicted catch and the dependent variables included in the GLM were examined for each model.

### 2.3 Area weighting of CPUE indices for MFCL

The region specific CPUE indices represent a key index of exploitable longline biomass in each of the MFCL regions that comprise the total area included within the stock assessment. The models typically assume that the catchability of the Japanese longline fleet is constant over time (allowing for seasonal variation) and is constant between the individual model regions. Nevertheless, there are considerable differences in the respective size of each model region and the relative abundance of bigeye/yellowfin within each region. A methodology was developed to derive region-specific scaling factors that account for these differences, thereby, enabling trends in regional CPUE to be proportional to trends of absolute stock biomass over the entire model area.

The analysis was based on a subset of the Japanese longline data used in the GLM analysis described in the previous section. The data set was restricted to 1960 to 1986 — the period that encompasses the widest geographical distribution of the Japanese fleet within the Pacific Ocean. As previously described, the data set was stratified by year, quarter, HBF, and spatially at a resolution of 5 degrees of latitude and longitude.

The data set was aggregated by spatial cell (5 degree latitude/longitude) and the cumulative catch and the number of effort records was determined for each spatial strata. Strata with a cumulative catch of less than 5000 fish and/or less than 10 unique year/quarter/HBF categories fished during 1960–86 were considered to account for an insignificant proportion of the total stock biomass and, consequently, were excluded from the data set.

A simple GLM was fitted to the final data set to estimate the relative CPUE in each individual lat/long cell included in the entire model area. The dependent variable ( $CPUE$ ) in the model was the natural logarithm of the catch (number of fish) divided by the number of hooks set in the respective year/quarter, latitude/longitude 5-degree cell and HBF category. Both latitude/longitude and HBF were included as categorical variables in the simple model. The predicted  $CPUE$  in lat/long cell  $k$  using longline gear with HBF  $j$  is given by

$$CPUE_{(k,j)} = aLATLONG_{(k)} + bHBP_{(j)} + \varepsilon_{(k,j)}$$

The region specific weighting factor ( $W_R$ ) for the CPUE index was calculated as the sum of the exponentiated coefficients of the lat/long cells ( $a$ ) included within the respective region ( $R$ ), i.e.

$$W_R = \sum_{i=1}^n \exp(a_{(R,i)})$$

Where  $n$  is the number of lat/long cells included in the region. The individual region weighting factors were then normalised relative to the average value for all regions.

To determine the weighted species-region-specific CPUE index, the GLM indices (from Section 2.2) were scaled relative to the mean of the series and the resulting series was scaled by the respective weighting factor.

## 2.4 Statistical habitat based models (stat HBS)

The methodology for the calculation of the statHBS indices for bigeye and yellowfin was similar to the approach described in Bigelow et al. 2004. The main differences from the previous approach was the estimation of a single habitat preference for the entire Pacific Ocean, rather than individual habitat preferences for each MFCL region, and the estimation of area effects within each region.

In the statHBS model, habitat preferences were structured as parameters. Fifteen ambient temperature preferences derived at 2°C intervals from 3.5°C to 33.5°C were used as priors. Fifteen priors were used for oxygen at 0.5 ml l<sup>-1</sup> intervals from 0 to 7.5 ml l<sup>-1</sup>. All priors were non-informative (uniform distribution) with a mean of zero.

The data are spatially stratified by 5 degrees of latitude and longitude. An area effect was included in the statHBS model by estimating an individual parameter for each 5° spatial cell — analogous to including the latitude/longitude categorical variable in the GLM model. The data set is limited to the latitude/longitude cells fished by the Japanese fleet since 1975. For the entire Pacific Ocean, this represents a total of 316 latitude/longitude parameters included in the model.

A likelihood function (a log-transformed least squares) was used as a measure of how well the predicted catch from the various effort series fit the observed catch:

$$L = (\ln(C_{i,y} + \delta) - \ln(\hat{C}_{i,j,y} + \delta))^2$$

where  $C_{i,y}$  is the observed catch for observation  $i$  in region/year/quarter/lat\*long  $y$ ;  $\hat{C}_{i,j,y}$  is the predicted catch for observation  $i$ , effort series  $j$  and region/year/quarter/lat\*long  $y$ . A constant ( $\delta = 0.0001$ ) was added to the observed and predicted catch to avoid computational problems when observed catch was zero.

For individual observations ( $i$ ) from an effort ( $E$ ) series  $j$ , an estimate of catch ( $C$ ) in region/year/quarter/lat\*long  $y$  is obtained as:

$$\hat{C}_{i,j,y} = E_{i,j,y} q_y B_y$$

where  $q$  is overall catchability and  $B$  is abundance. Year effects ( $\theta_y = qB_y$ ) are estimated because both  $q$  and  $B$  are unknown.

The contribution of habitat priors to the objective function are:

$$\Theta_h = \sum \varepsilon^2 / 2\sigma^2$$

where  $\varepsilon$  are residuals of the habitat preferences and  $\sigma$  is the standard deviation of the prior distribution which has mean  $\mu$ . The negative log-likelihood is minimised by simultaneously estimating various parameters with the function minimiser in AD Model Builder ([www.otterrsch.com](http://www.otterrsch.com)). The statHBS models were applied to CPUE from 1975 to 2004 when gear configuration was known. Predicted habitat and latitude/longitude parameters from the statHBS models were then applied to the 1952 to 2004 time-series to derive the region-specific year/quarter indices for the entire period.

In addition, a separate analysis which excluded the latitude/longitude factor was also undertaken for each MFCL region and habitat preferences were estimated separately for each region. This approach is equivalent to the analysis undertaken by Bigelow et al. (2004).

Similar to the GLM analysis, the region-specific year/quarter indices derived from the statHBS model were scaled by a weighting factor that accounts for the relative abundance of longline exploitable biomass between regions. This was undertaken by multiplying the region-specific year/quarter indices by the summation of the latitude/longitude parameters from the specific region.

## **3 Results**

### **3.1 Data summary**

For each model region, annual trends in catch, effort, and nominal CPUE are presented in Figure 2 and Figure 3. The main trends evident in these data are as follow.

- The spatial expansion of the fishery from region 1 in the early 1950s to the entire WCPO (regions 1–6) by the early 1960s and the subsequent development of the fishery in the EPO (regions 7 and 8).
- The high (low) proportion of bigeye (yellowfin) in the catch from regions 1, 2, 7, and 8.
- The low (high) proportion of bigeye (yellowfin) in the catch from regions 3 and 5.
- Comparable catches of bigeye and yellowfin in region 4 and high initial catches of yellowfin in region 6.
- The decline in fishing effort in the EPO (regions 7 and 8) in the last decade. The very low level of effort in region 6 in the last decade.
- The extremely high nominal CPUE for bigeye during the initial years of the fishery in regions 7 and 8 and, to a lesser extent, in region 4 and the subsequent large decline in CPUE. A similar trend is also evident for yellowfin in regions 4–6 and region 8.
- The steady decline in yellowfin nominal CPUE in region 3 and, to a lesser extent, in region 4.
- Declining bigeye nominal CPUE in regions 2, 4, 7, and 8. Stable or increasing nominal CPUE for bigeye in region 1, 3, and 6.

### **3.2 Generalised linear models (GLM)**

#### **3.2.1 Bigeye**

Overall, the individual region-specific GLMs explained a high proportion (60–70%) of the observed variation in the catch of bigeye (Table 2). A high proportion (70–80%) of the explained variation was attributable to the inclusion of the effort term (number of hooks) as a predictor variable, while year/qtr typically accounted for about 5–10% of the explained variation.

**Table 2. Explanatory power (% , adjusted R<sup>2</sup>) of the region-specific GLMs for bigeye and yellowfin tuna.**

<b>Region</b>	<b>Bigeye</b>	<b>Yellowfin</b>
1	58.6	53.8
2	67.2	61.7
3	71.3	75.2
4	69.6	75.3
5	64.3	69.4
6	64.6	70.8
7	71.9	
8	67.6	

The most influential variable included in the bigeye GLMs is the number of hooks set. At low levels of effort catches (and catch rates) are predicted to be very low, while catches steadily increase with increased effort over the main data range (Figure 4). For most regions, catches reach a plateau at relatively high levels of effort (about 500,000 to 800,000 hooks per quarter) and for higher levels of effort predicted catches are depressed (Figure 4).

For most regions, predicted bigeye catches are generally correlated with yellowfin CPUE (catch per hundred hooks), i.e., larger catches of bigeye generally occur when the catch rate of yellowfin is higher (Figure 4). However, the magnitude of this effect varies between regions being stronger in the EPO regions (7 and 8) and the sub-equatorial regions of the WCPO (1, 2, 5, and 6) and weaker in the equatorial regions of the WCPO (3 and 4). This suggests that there are differences in the level of association of the two species between the various regions.

For the main subequatorial regions of the WCPO (1, 2, and 5) and the EPO (region 8), there is a general increase in the predicted catch (and catch rate) of bigeye with increased HBF (correlated with increased fishing depth) (Figure 4). However, this relationship is not apparent in the data from the other regions with no apparent trend in bigeye catch over the range of HBF in the equatorial regions of the WCPO (3 and 4) and the northern EPO (region 7), while the contrary trend in bigeye catch is evident for region 6 — bigeye catches declining with increased HBF (Figure 4). These regional differences may reflect differences in oceanographic conditions between regions or be partly attributable to interactions with other factors included in the model, for example, the inclusion of the yellowfin CPUE variable.

For each region, the quarterly indices for bigeye are presented in Figure 5. Most regions, with the exception of 5 and 6, are characterised by a sharp decline in CPUE (greater than 50%) during the first 10 years of fishing. Over the subsequent period, most regions reveal a general steady decline in CPUE (regions 1, 2, 4, 5, 7, and 8) or relative stable CPUE (region 3). CPUE indices in region 6 are highly variable throughout the time-series. Region 1 and, to a lesser extent, region 2 are characterized by high seasonal variability in the catch rate of bigeye (Figure 5).

The precision of the indices, as illustrated by the confidence intervals, is lowest at the start of the time-series and for regions 6 and 7 where low fishing effort occurs (Figure 5).

In all regions, the GLM does not fit the low values of observed catch particularly well; consequently, the distribution of the residuals are negatively skewed (Figure 6). The proportion of

records with low (less than -1.5) residuals was highest in region 1 reflecting a larger percentage of low bigeye catches. For all regions, there was no time-series trend in the proportion of low residuals in the GLMs (Figure 7), indicating that the poor fit to the low values was not introducing a significant bias to the resulting standardised indices.

For most regions, the trend in the GLM index is generally comparable to the nominal indices (Figure 8), although the GLM index is generally more pessimistic for the recent period (post 1990).

### **3.2.2 Yellowfin**

Overall, the individual region-specific GLMs explained a moderate proportion (50–75%) of the observed variation in the catch of yellowfin, with the models for the two equatorial regions having the highest explanatory power (Table 2). As with bigeye, the effort term (number of hooks) accounts for a high proportion of the explained variation (50–85%). Year/quarter is most influential in the non-equatorial regions accounting for 15–25% of the explained variation (compared to 5–10%), probably reflecting the more seasonal nature of these fisheries.

For all regions, the predicted catch of yellowfin generally increases with increasing bigeye CPUE, reaching a maximal threshold and then declining when bigeye catch rates are highest declines for (Figure 4). This relationship may reflect changes in the species composition of the catch under different targeting strategies.

The HBF variable does not contribute substantially to the explanatory power of the yellowfin CPUE models. For some regions (1, 2, and 5) there is a slight positive relationship between yellowfin catch and HBF, while only region 6 exhibits the more intuitive decline in predicted yellowfin catch with increased HBF (Figure 4). It is likely that the inclusion of bigeye CPUE as a predictor variable in the model is accounting for the expected decline in yellowfin catch with increased depth of set (correlated with higher bigeye CPUE).

The most influential variable included in the yellowfin GLMs is the number of hooks set. At low levels of effort catches (and catch rates) are predicted to be very low, while catches steadily increase with increased effort over the main data range (Figure 4). At high levels of effort, most region models predict catches to increase at a lower rate, stabilise at a plateau, or be slightly lower than the maximal level.

For each region, the quarterly indices for yellowfin are presented in Figure 9. Regions 5 and 6 reveal the steep initial decline in yellowfin CPUE evident in the bigeye CPUE indices for most regions. The equatorial regions (regions 3 and 4) and, to a lesser extent, region 6 reveal an increase in CPUE during the mid 1970s and a subsequent decline in CPUE from the late 1980s (Figure 9). Catch rates in the equatorial regions have been very low in the last few years. For the other regions, CPUE trends have been relatively variable although the two northern regions also exhibited very low CPUE since 2000.

The precision of the CPUE indices, as indicated by the confidence intervals, is lowest for the northern regions 1 and 2 (Figure 9). There is also high uncertainty associated with the higher CPUE indices from region 6 during the latter half of the time-series.

In all regions, the GLM does not fit the low catch observations particularly well; consequently, the distribution of the residuals are negatively skewed (Figure 10). The proportion of records with low (<-1.5) residuals was highest in region 1 (Figure 11) reflecting a larger percentage of low



yellowfin catches. For all regions, there is no systematic trend in the proportion of low residuals in the GLM.

The trend in the GLM index is comparable to the nominal indices for most regions (Figure 12), although the GLM indices from region 3 and, to a lesser extent, region 5 are less pessimistic than the corresponding nominal index.

### 3.3 Statistical habitat based models (*stat HBS*)

#### 3.4.1 Summary of explanatory power

Overall, the *statHBS* models (including area) explained 25% and 35% of the variation in catch rate (catch per hundred hooks) of bigeye and yellowfin in the Pacific-wide dataset (Table 5). The relative explanatory power of the area effects and the Pacific-wide habitat preferences was investigated by comparing a base model (estimating year/quarter parameters only) with models including the area and habitat preference variables separately.

For bigeye, the habitat preferences accounted for a very small increase in the explanatory power of the model once the area effect (latitude/longitude cell) was included (Table 5). The habitat preferences were more influential in the yellowfin model; the inclusion of the preferences in the base model substantially improved the explanatory power of the model. However, the largest improvement was gained with the addition of the area effect. The inclusion of the area effect also negated much of the explanatory power attributable to the habitat preferences and any additional explanatory power (attributable to the habitat preferences) was minimal (Table 5).

For both models, it is evident that the habitat preferences are not adequate to predict the geographical differences in observed catch rate. Clearly, further refinement of the parameterisation of the habitat preferences is required before the model is sufficiently reliable to standardise CPUE to account for geographical differences in the operation of the fishery. In the interim, these limitations can, to some extent, be addressed by the inclusion of the area effect within the *statHBS* models.

**Table 3. Comparison of *statHBS* models to estimate standardized CPUE for bigeye and yellowfin tuna from 1975 to 2004.**

Bigeye - Model	Parameters	Likelihood	Improvement	Improvement per parameter	Pseudo- $R^2$
Base (yr:quarter)	120	761,200			0.01
Habitat	150	709,420	51,780	1,726.0	0.08
Area	435	581,111	180,089	571.7	0.25
Area + habitat	465	577,254	183,946	533.2	0.25
Yellowfin - Model					
Base (yr:quarter)	120	1,958,460			0.01
Habitat	150	1,469,040	489,420	16,314.0	0.26
Area	435	1,292,480	665,980	2,114.2	0.35
Area + habitat	465	1,285,110	673,350	1,951.7	0.35

### 3.4.2 Bigeye

The area coefficients estimated from the bigeye statHBS model reveal highest catch rates within the north-eastern and eastern equatorial Pacific and a general westward decline in bigeye CPUE (Figure 13). Bigeye CPUE is predicted to be lowest within the north- and south-west Pacific.

A spatial representation of the mean residual in each 5° cell was used to compare the performance of the statHBS with and without the inclusion of area effects (c.f. Figure 14 and Figure 15). The statHBS without area effects (habitat preferences only) had larger negative residuals in the eastern tropical Pacific (region 8) and large positive residuals in the central Pacific (regions 2 and 4) (Figure 14). The inclusion of the area effect greatly reduced the spatial trend in residuals (Figure 15).

The fitted habitat parameters are also presented for statHBS models with and without area effects (Figure 16). For a statHBS without area effects, the fitted temperature parameters are indicative of temperature preferences and represent a bimodal temperature preference with peaks at 15° and 25°C (Figure 16). The fitted temperature preference from the statHBS with the inclusion of an area effect is substantially different and most of the habitat preference is in the lowest temperature category. However, the preferences are not comparable between the two models as the inclusion of area effects is likely to account for the geographic distribution of the species with respect to temperature, while the habitat preferences will explain the additional variation associated with each area effect.

There was little difference in the estimated oxygen habitat preference between the statHBS models including and excluding the area effect (Figure 16).

The year/quarter indices derived from the statHBS model (including area effects) are very similar to the GLM indices presented in Section 3.2 (Figure 17). For comparison, the statHBS indices were also recalculated using the approach presented in Bigelow et al (2004) where separate habitat preferences were calculated for each MFCL region and no area effect was included. Overall, these indices were more optimistic than either the statHBS-area model or the GLM indices, revealing less of a decline in CPUE, particularly in the last decade (Figure 17).

### 3.4.3 Yellowfin

For the statHBS without area effects, the fitted temperature parameters for yellowfin reveal a unimodal temperature preference with a peak at from 27° to 29°C (Figure 18). The same model also indicates a decline in habitat preference at lower oxygen concentrations (from 3.5 to 1.0 ml·l<sup>-1</sup>). There is also an unrealistic decline at higher oxygen concentrations, but this is related to an interaction of temperature which is the dominant preference in most areas and also that some strata in the ocean do not have high oxygen values and are thus under-represented in the model.

As with the bigeye models, the habitat preferences from the statHBS model including the area effect differ significantly from the model excluding area (Figure 18). There is a strong interaction between area and the habitat preferences, particularly the temperature preference, and, therefore, the preferences are likely to be explain some of the deviance associated with the individual area effects. These are likely to include seasonal trends in each area as well as differences in vertical thermal structure and fish distribution.

The year/quarter indices derived from the statHBS model (including area) were generally comparable to the corresponding GLM indices (Figure 19), although there was some deviation in the indices in the southern regions (5 and 6) and region 2. As for bigeye, the yellowfin statHBS

indices were also recalculated with region-specific habitat preferences (as per Bigelow et al. 2004). For most regions, the resulting indices deviated considerably from both the GLM and statHBS-area indices. For the equatorial areas (regions 3 and 4) and region 1, these alternative indices were more optimistic in recent years (post 1990) compared to the other series, while for the southern regions (5 and 6) the indices were more pessimistic (Figure 19).

### 3.4 Area weighting of CPUE indices for MFCL

During 1960–1986, the Japanese longline fleet operated throughout the Pacific Ocean between latitudes 40°S and 40°N (Figure 20). However, fishing effort was generally concentrated in the equatorial region and was relatively low in the south-eastern Pacific. The selection criteria for inclusion of lat/long cells in the area weighting analysis generally resulted in the removal of those cells that had a level of effort below 10% of the global average. This is evident by comparing the distribution of fishing effort (Figure 20) with the cells included in the bigeye (Figure 21) and yellowfin (Figure 22) GLMs. Overall, the qualifying criteria resulted in a significant reduction in the number of cells included in regions 6–8 for bigeye and region 2 and 6 for yellowfin (Table 4).

**Table 4. The number of 5 degree latitude/longitude cells included in the entire data set and the number of cells included in the bigeye and yellowfin GLM area weighting analyses by MFCL region.**

Region	Total	Qualifying	
		BET	YFT
1	38	32	31
2	32	31	23
3	61	53	57
4	48	48	47
5	28	22	25
6	45	16	24
7	28	18	-
8	158	101	-

The coefficients from the bigeye model reveal highest catch rates are within the eastern equatorial Pacific and in the north-eastern Pacific (Figure 21). Catch rates generally decline westwards along the equator and are low at higher latitudes in the western and south Pacific.

In contrast, the highest catch rates of yellowfin were within 5 degree latitude/longitude cells in the western equatorial region, centred on the area of Papua New Guinea archipelagic waters (Figure 22). Catch rates steadily decline with increasing latitude and eastwards along the equator.

For both bigeye and yellowfin, the GLM relative area weightings were determined by calculating the sum of the lat/long coefficients within each of the regions, thus, incorporating both the relative catch rate and the size of the region. For bigeye, the area weightings indicate that longline exploitable biomass is highest in region 8, followed by regions 4, 3, 7, and 2, while relative biomass in regions 1, 5, and 6 is approximately 5% of the level of region 8 (Table 5 and Figure 23).

The area weighting derived from the statHBS model (including an area effect) was calculated in a similar manner as the GLM weighting (see Section 2.4). For bigeye, the statHBS weightings differed marginally from the GLM values (Table 5 and Figure 23). Relative weightings based on

the statHBS model were slightly higher for regions 2 and 7 and lower in region 8 compared to the GLM approach.

For yellowfin in the WCPO, the GLM area weightings indicate the highest levels of biomass are within region 3 followed by regions 4, 5, and 6, with limited biomass in the two northern regions (1 and 2) (Table 5 and Figure 24). Comparable results were obtained from the statHBS approach with the only pronounced difference occurring in region 6; this region is given a higher weighting using the GLM approach (Table 5 and Figure 24).

**Table 5. Relative region scaling factors for bigeye (Pacific wide) and yellowfin (WCPO) from the GLM and statHBS approaches. The scaling factors are relative to the average of each series.**

Region	Bigeeye		Yellowfin	
	GLM	statHBS	GLM	statHBS
1	0.154	0.145	0.232	0.247
2	0.114	0.228	0.603	0.849
3	3.092	2.965	0.966	0.830
4	1.495	1.640	1.576	1.512
5	0.746	0.946	0.143	0.104
6	0.398	0.077	0.173	0.144
7			0.812	1.157
8			3.495	3.158

## 4 Discussion

The GLM indices presented in this paper are not directly comparable to those indices presented in Langley (2003) due to changes in the regional structure of the bigeye and yellowfin MFCL assessments. Nevertheless, the GLM model structure and the global data set are very similar and, therefore, any differences in the indices are likely to be solely due to the changes in the regional boundaries.

The main change to the statHBS approach was the inclusion of area (latitude and longitude) effects. These area effects substantially improved the explanatory power of the model compared to models with just habitat preferences estimated. The resulting model was much closer in structure to the GLM model which also includes latitude and longitude as an interaction term. Consequently, the resulting statHBS indices were similar to those derived from the GLM approach and the trends in the indices differed from the statHBS indices derived from the methodology used in the previous analysis (Bigelow et al. 2004).

In summary, the results of this study reveal the performance of the statHBS model is substantially improved with the inclusion of area effects (latitude and longitude stratification), particularly in the case of bigeye tuna where Pacific-wide habitat preferences contributed only marginally to the improvement of the performance of the model. The case was similar for yellowfin and, while the Pacific-wide habitat preferences explained a higher proportion of the variance than for bigeye, the habitat preferences contributed little to the model once the area effect was included. Therefore, it would seem that the habitat preferences are, at best, explaining the prevailing conditions in a geographic area, particularly for surface orientated species such as yellowfin.

However, the statHBS model, as currently parameterised, appears to perform poorly with respect to defining the vertical preference of the species — the key rationale for the approach in the first instance. In the case of bigeye, differences in relative catch rate between areas are poorly explained by the species estimated preferences for both temperature and oxygen concentration. While these variables may be important in delineating the extremes of the range for the species, within that range there are likely to be other factors more important in defining the species habitat and, thereby, the relative abundance of the species. The most notable is the availability of prey items.

Archival tagging studies of bigeye tuna (Musyl et al 2003, Gunn et al 2005) have revealed strong diurnal trends in the depth distribution of the species. It has been hypothesised that changes in the vertical distribution of bigeye are attributable to the foraging strategy of the species, whereby bigeye follow the vertical migration of the prey in the deep scattering layer (DSL). On this basis, it would be more appropriate to formulate a habitat preference, at least on the vertical scale, that was based on the distribution of bigeye relative to the depth of the DSL or some other proxy such as light intensity.

In the case of yellowfin, a comparison of time-at-temperature profiles from different areas of the Pacific has indicated that ambient temperature is not a reliable indicator of habitat preference for the species over a wide spatial scale (Brill et al. 1999). Brill et al. proposed an alternative vertical preference for yellowfin based on the differential from the temperature at the surface. This formulation of the yellowfin habitat preference should be investigated within the context of the statHBS model.

There is a range of other important variables that are likely to influence the distribution of a pelagic species. Most critically will be the underlying productivity of an area. This may be easiest to parameterise simply by including an area effect as per the statHBS model presented in this report. However, this approach does not capture fine scale variation in productivity, such as associated with oceanic fronts, or temporal (principally seasonal) variation in the productivity of an area. This level of refinement of the statHBS model is likely to be beyond the current temporal and spatial resolution of the catch and effort and oceanographic data available, particularly for the historical time-series. Nevertheless, it may be an appropriate development of the statHBS approach for fisheries where high resolution data are available.

In conclusion, there is considerable scope to improve the current statHBS approach, particularly with respect to how the habitat preferences of individual species are formulated. This is likely to require a more thorough analysis of the results from archival tagging projects to define the variables (or correlated variables) that explain the vertical and spatial distribution of the respective species.

In the meantime, it is prudent to persist with the standardisation of longline CPUE data using the more established GLM approach. The approach is not sensitive to assumptions regarding the habitat preference of the species and, with the inclusion of the appropriate predictor variables, should be adequate to capture broad-scale changes in longline catchability.

A direct statistical comparison of the statHBS and GLM models was not possible due to the differences in spatial structure of the two types of models. Nevertheless, a comparison of the respective fits for the WCPO bigeye and yellowfin MULTIFAN-CL assessment models, integrating all other sources of data from the fishery, reveals that the GLM standardised effort series provides a better overall fit to the data (Hampton et al. 2005a, 2005b). For both species and particularly for bigeye, the fits to the data are substantially better for the model runs including the

GLM standardised effort series than for the equivalent models with the statHBS standardised effort, as indicated by the total likelihood functions from the models (Table 6). Essentially, the GLM effort standardisation results in an overall fit to the data that is more consistent with the trends in the catch and effort data from the other fisheries included in the model. This would be indicated by the lower penalties (included in the total likelihood function) associated with the effort deviations for the principal longline fisheries. On this basis, it was considered that the model incorporating the GLM standardised effort series should represent primary assessment for each species.

A significant change to the 2005 assessments for bigeye and yellowfin in the WCPO was the scaling of the standardised effort series by the respective species-specific regional scaling factor. The relative weighting factors effectively scale the biomass level in each of the model regions. In previous assessments, these weightings were determined by a qualitative examination of the species-specific CPUE data, essentially determining the area within each region where the species was caught. However, the statistical approach used to determine these weighting factors presented in this paper is more defensible as it accounts for both the size of the area fished and the relative CPUE level. It is encouraging that for both species the GLM and statHBS approaches yielded comparable regional scaling factors. The resulting weighting factors are also broadly consistent with the overall distribution of species catch between regions.

## 5 References

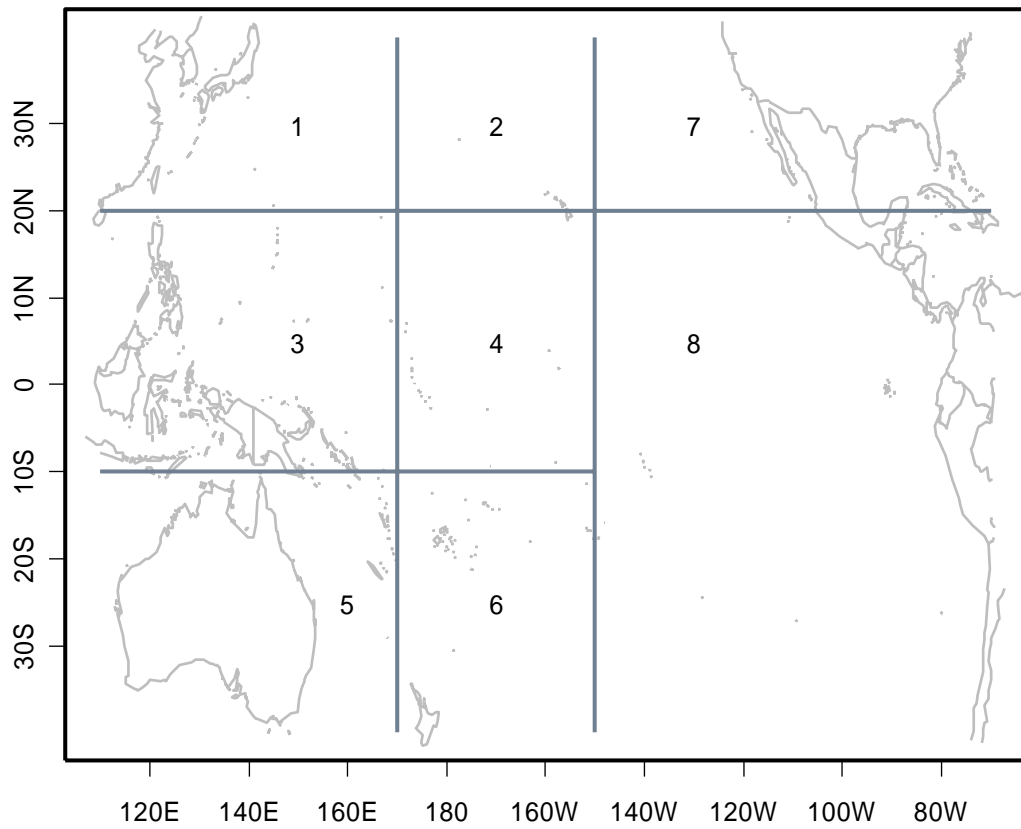
- Bigelow, K., A. Langley, T. Patterson (2004) Relative abundance indices of the Japanese longline fishery for bigeye and yellowfin tuna in the western and central Pacific Ocean. Working Paper SA-7. Seventeenth Meeting of the Standing Committee on Tuna and Billfish. Majuro, Marshall Islands. 9<sup>th</sup>–18<sup>th</sup> August 2004.
- Brill, R.W., Block, B.A., Boggs, C.H., Bigelow, K.A., Freud, E.V., Marcinek, D.J. (1999). Horizontal movements and depth distribution of large adult yellowfin tuna (*Thunnus albacares*) near the Hawaiian Islands, recorded using ultrasonic telemetry: implications for the physiological ecology of pelagic fishes. *Marine Biology* 133: 395–408.
- Gunn, G., Hampton, J., Evans, K., Clear, N., Patterson, T., Langley, A., Bigelow, K., Leroy, B., Williams, P., Miyabe, N., Bestley, S., Hartmann, K., Sibert, J. (2005). Migration and habitat preferences of bigeye tuna, *Thunnus obesus*, on the east coast of Australia – a project using archival and conventional tags to determine key uncertainties in the species stock structure, movement dynamics and CPUE trends. Final Report FRDC Project no. 1999/109. Published by CSIRO Marine Research.
- Hampton, J., P. Kleiber, A. Langley, Y. Takeuchi & M. Ichinokawa (2005a). Stock assessment of yellowfin tuna in the western and central Pacific Ocean. WCPFC–SC1 Working Paper SA-WP1. Oceanic Fisheries Programme, Secretariat of the Pacific Community, Noumea, New Caledonia.
- Hampton, J., P. Kleiber, A. Langley, Y. Takeuchi, M. Ichinokawa & M. Maunder (2005b). Stock assessment of bigeye tuna in the western and central Pacific Ocean, with comparisons to a Pacific-wide assessment. WCPFC–SC1 Working Paper SA-WP2. Oceanic Fisheries Programme, Secretariat of the Pacific Community, Noumea, New Caledonia.
- Hampton, J., Kleiber, P., Takeuchi, Y., Kurota, H., and Maunder, M. 2003. Stock assessment of bigeye tuna in the western and central Pacific Ocean, with comparisons to the entire Pacific Ocean. WP BET-1, SCTB 16, Mooloolaba, Australia, 9–16 July 2003.
- Langley, A. (2003) Standardised analysis of yellowfin and bigeye cpue data from the Japanese longline fleet, 1952 to 2001. Working Paper RG-2. Sixteenth Meeting of the Standing Committee on Tuna and Billfish. Mooloolaba, Australia. 9<sup>th</sup>–16<sup>th</sup> July 2003.

Musyl, M.K., Brill, R.W., Boggs, C.H., Curran, D.S., Kazama, T.K., Seki, M.P. (2003) Vertical movements of bigeye tuna (*Thunnus obesus*) associated with islands, buoys and seamounts near the main Hawaiian Islands from archival tagging data. *Fish Oceanogr* 12:1–18

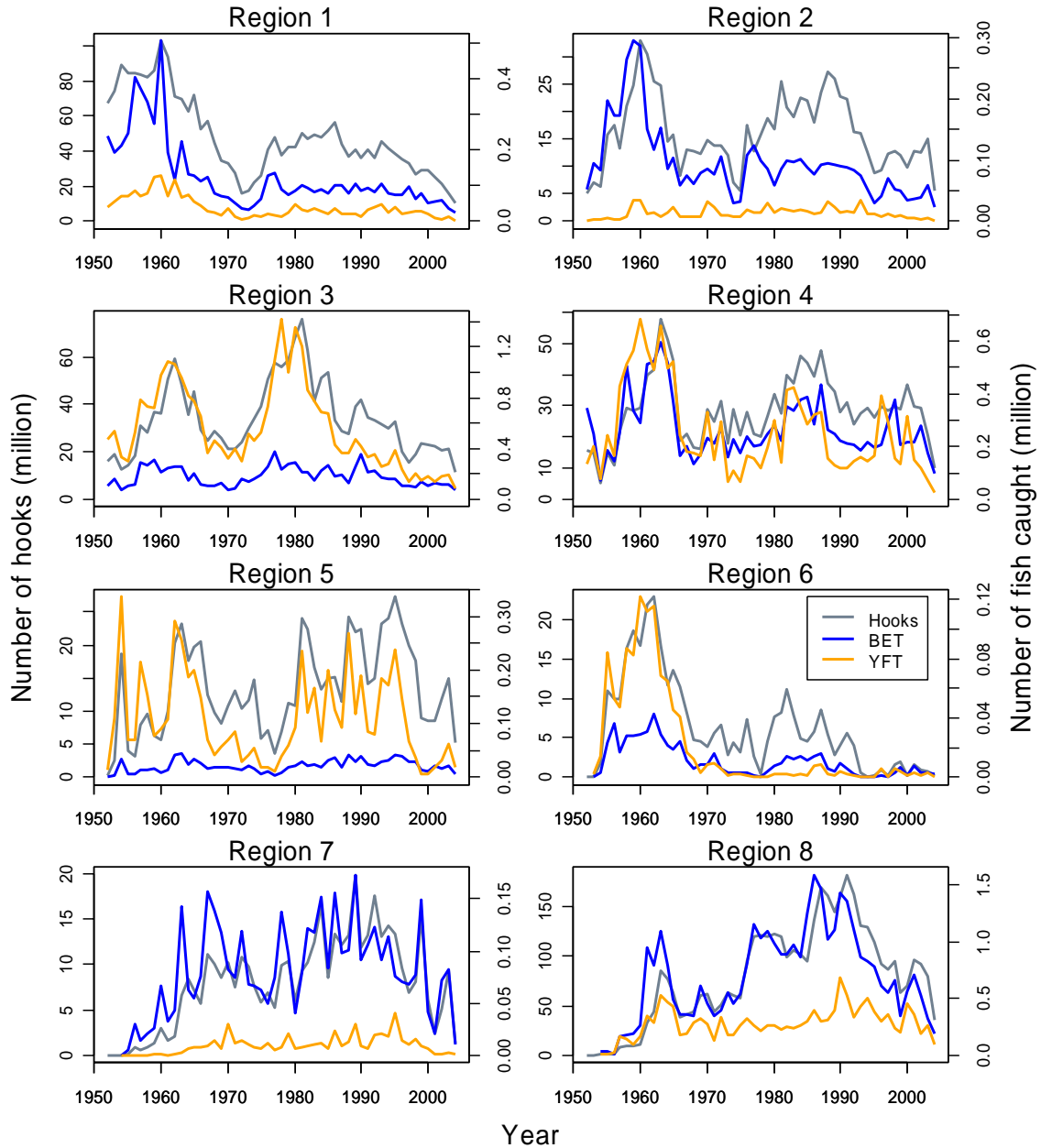
**Table 6. Total likelihood values for the WCPO bigeye and yellowfin MFCL assessments for alternative models using either GLM or statHBS standardised Japanese longline effort series. Values are for model runs using a fixed natural mortality at age (see Hampton et al. 2005a, 2005b for details). Lower values represent an improved fit to the model data set.**

<b>Effort standardisation</b>	<b>Bigeye</b>	<b>Yellowfin</b>
<b>GLM</b>	-788,905	-841,036
<b>statHBS</b>	-786,613	-839,309

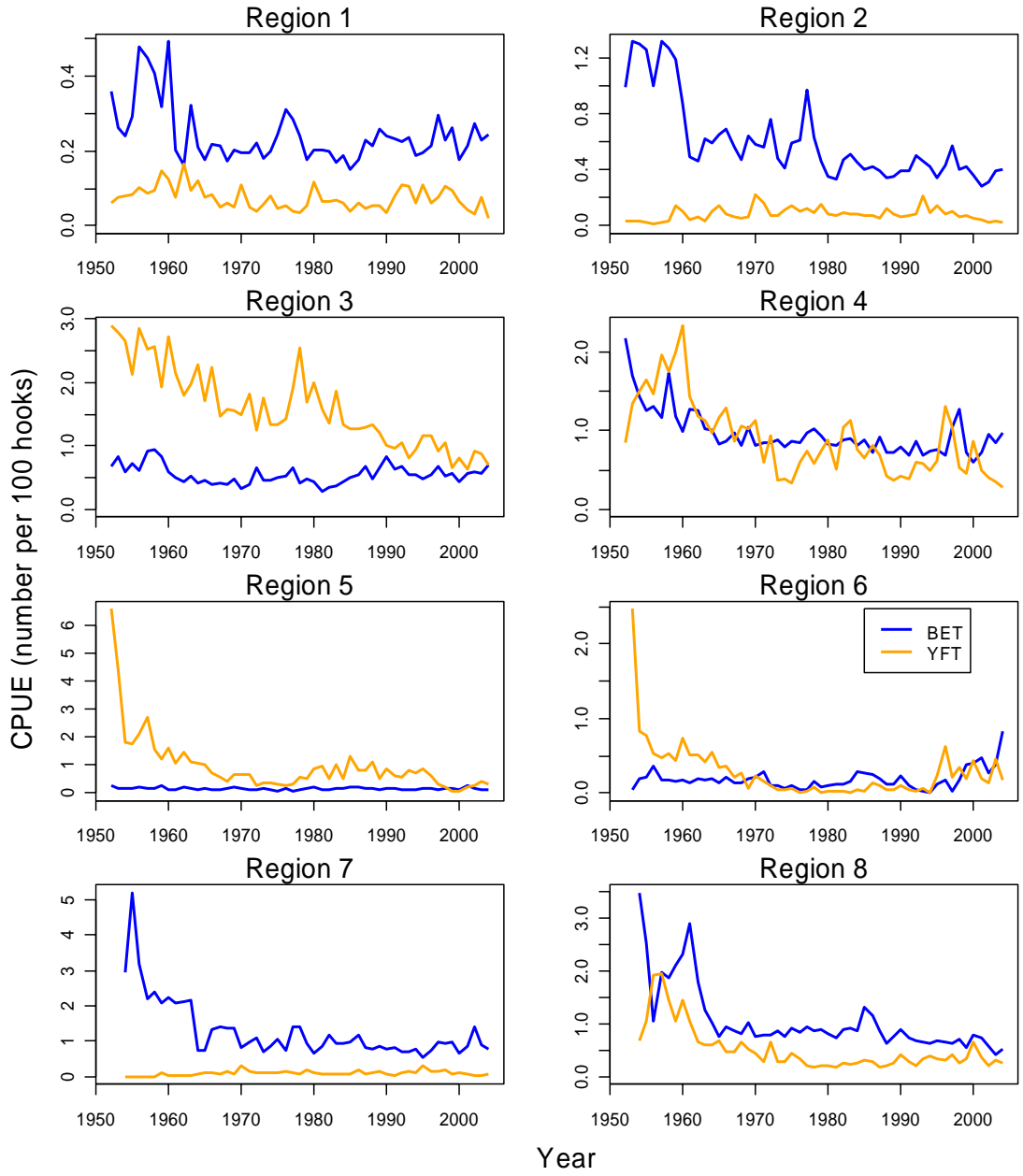




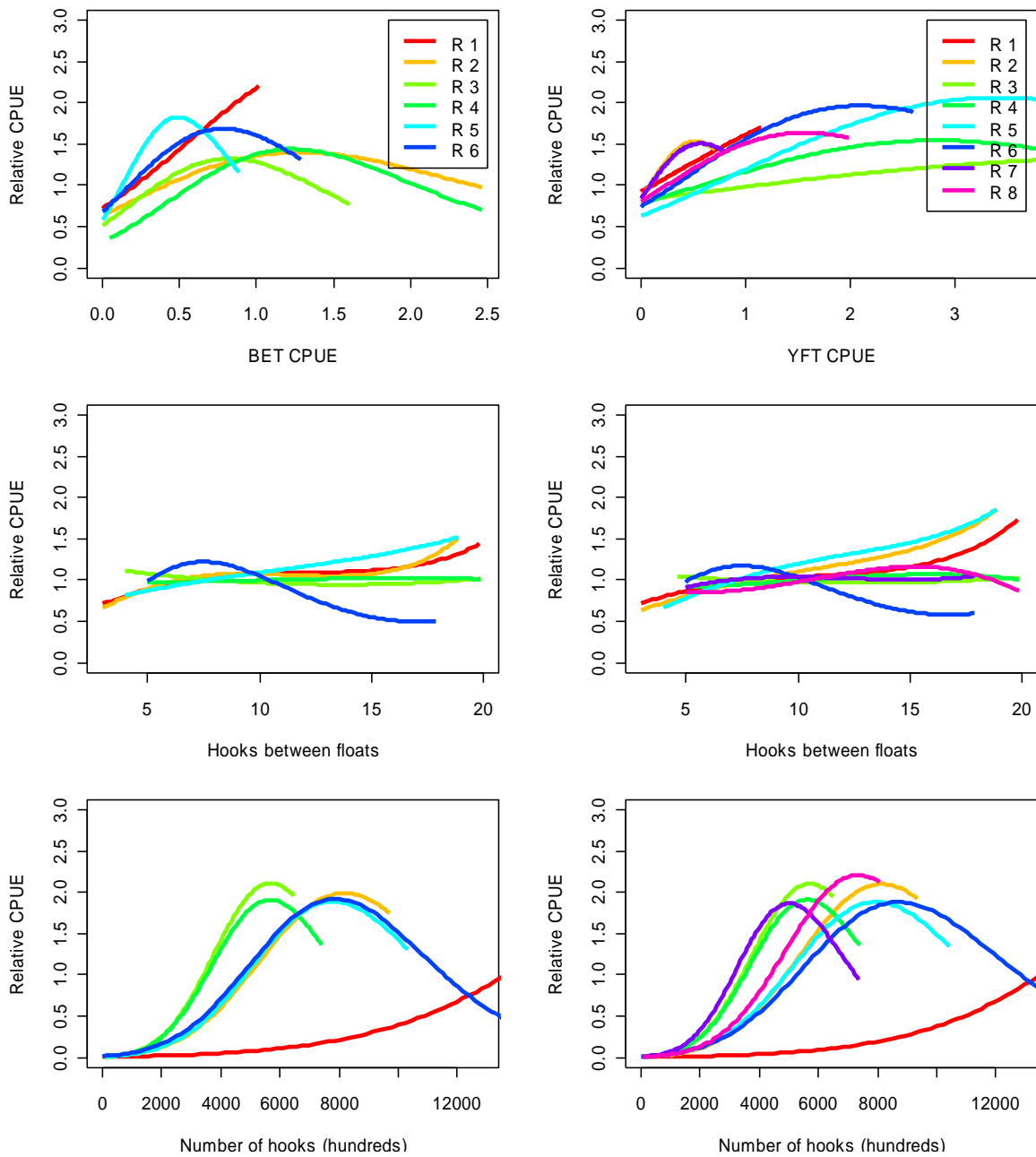
**Figure 1. Definitions of the regions included in the WCPO yellowfin and bigeye and Pacific-wide bigeye MFCL assessments.**



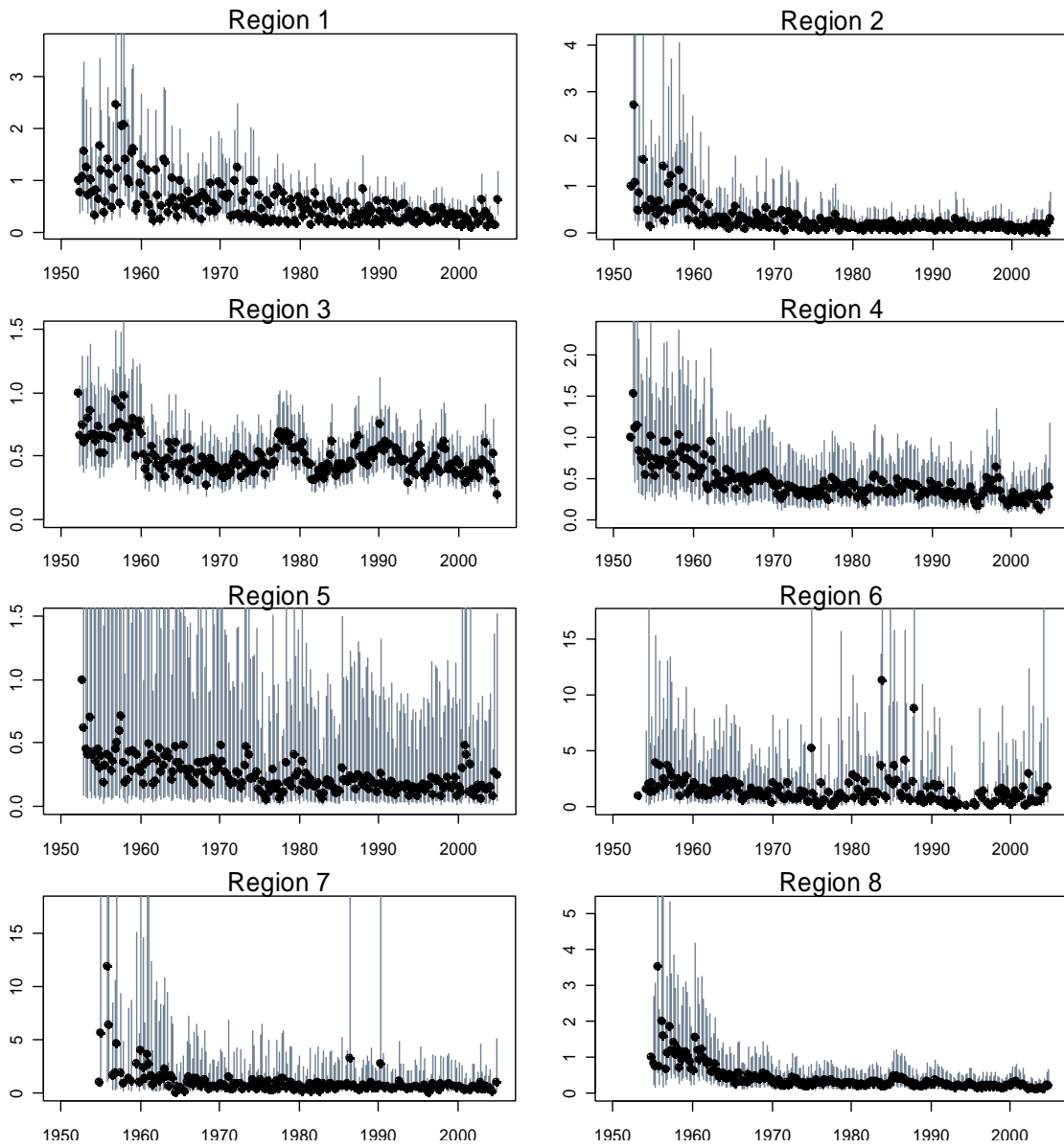
**Figure 2. Annual effort (millions of hooks) and catch (millions of fish) of bigeye and yellowfin by MFCL region included in the Japanese longline data set.**



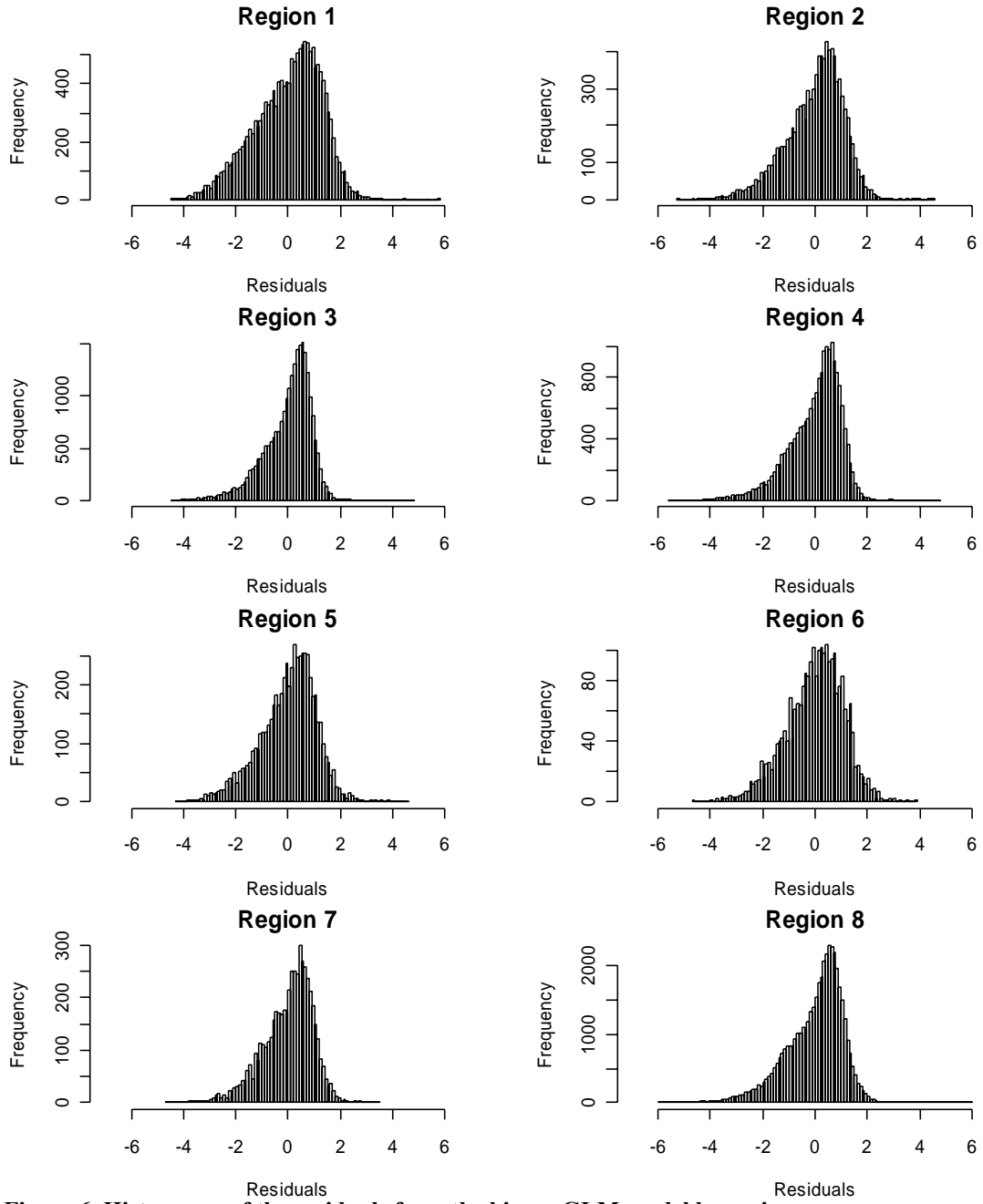
**Figure 3. Annual trends in nominal CPUE (number of fish per 100 hooks) of bigeye and yellowfin by MFCL region.**



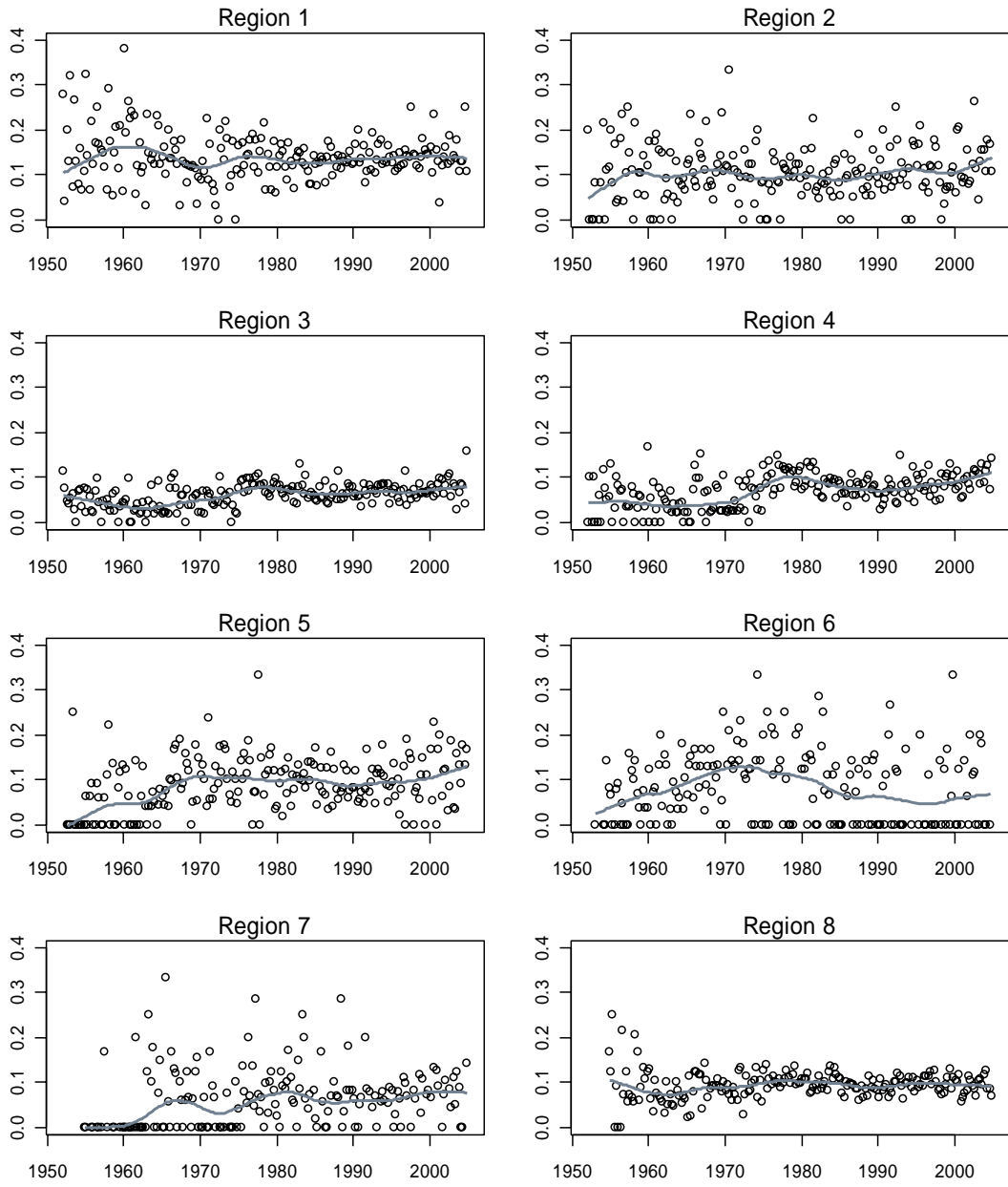
**Figure 4. The predicted relationship between yellowfin (left) and bigeye (right) catch and the continuous variables included in the GLM for each region. CPUE is expressed as catch per observation.**



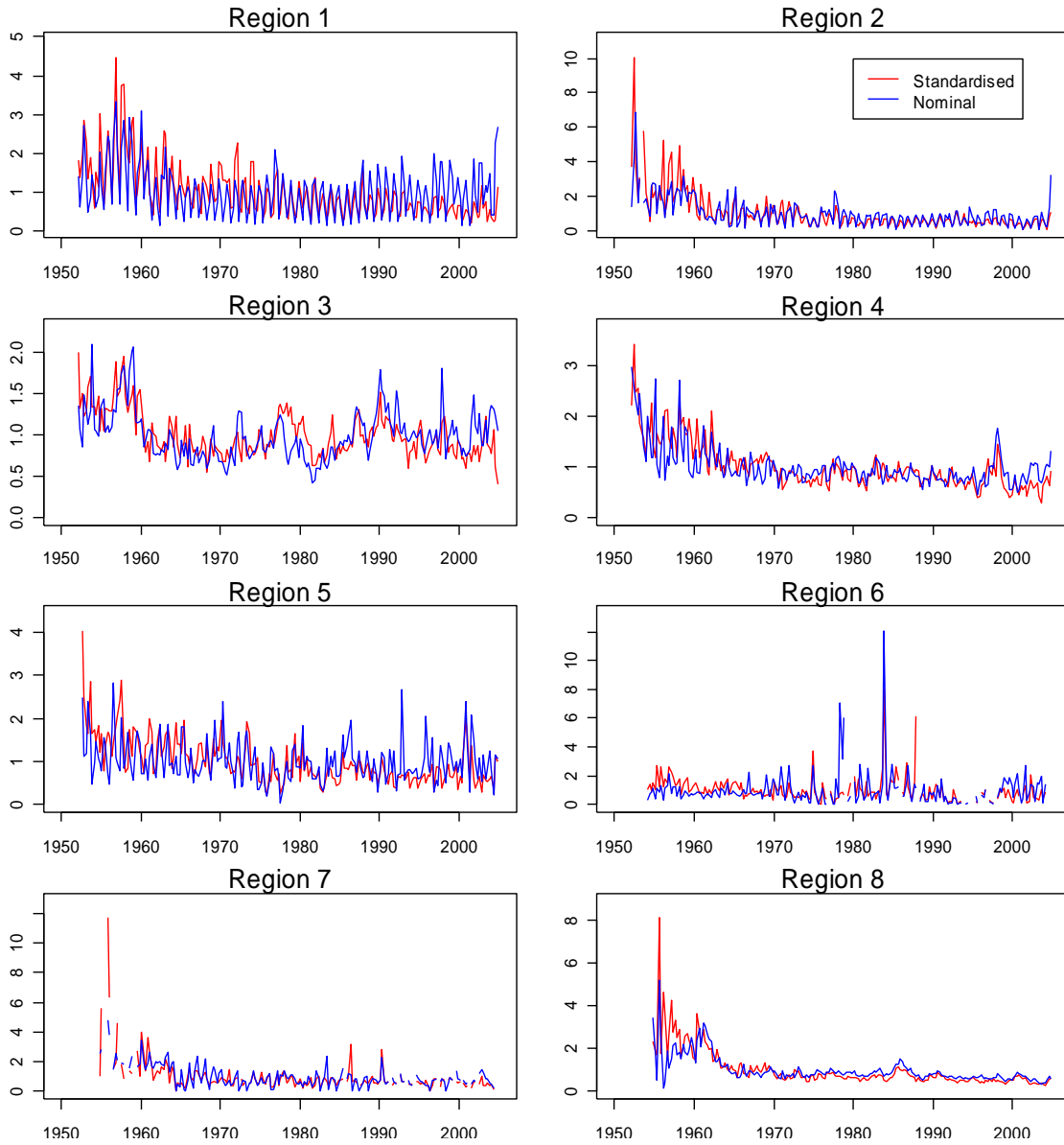
**Figure 5. Quarterly indices for bigeye derived from the GLM for each MFCL region. The indices are standardised relative to the first quarter of the series. The vertical grey lines represent the confidence interval for the index ( $\pm 2$ \*standard error).**



**Figure 6. Histograms of the residuals from the bigeye GLM model by region.**

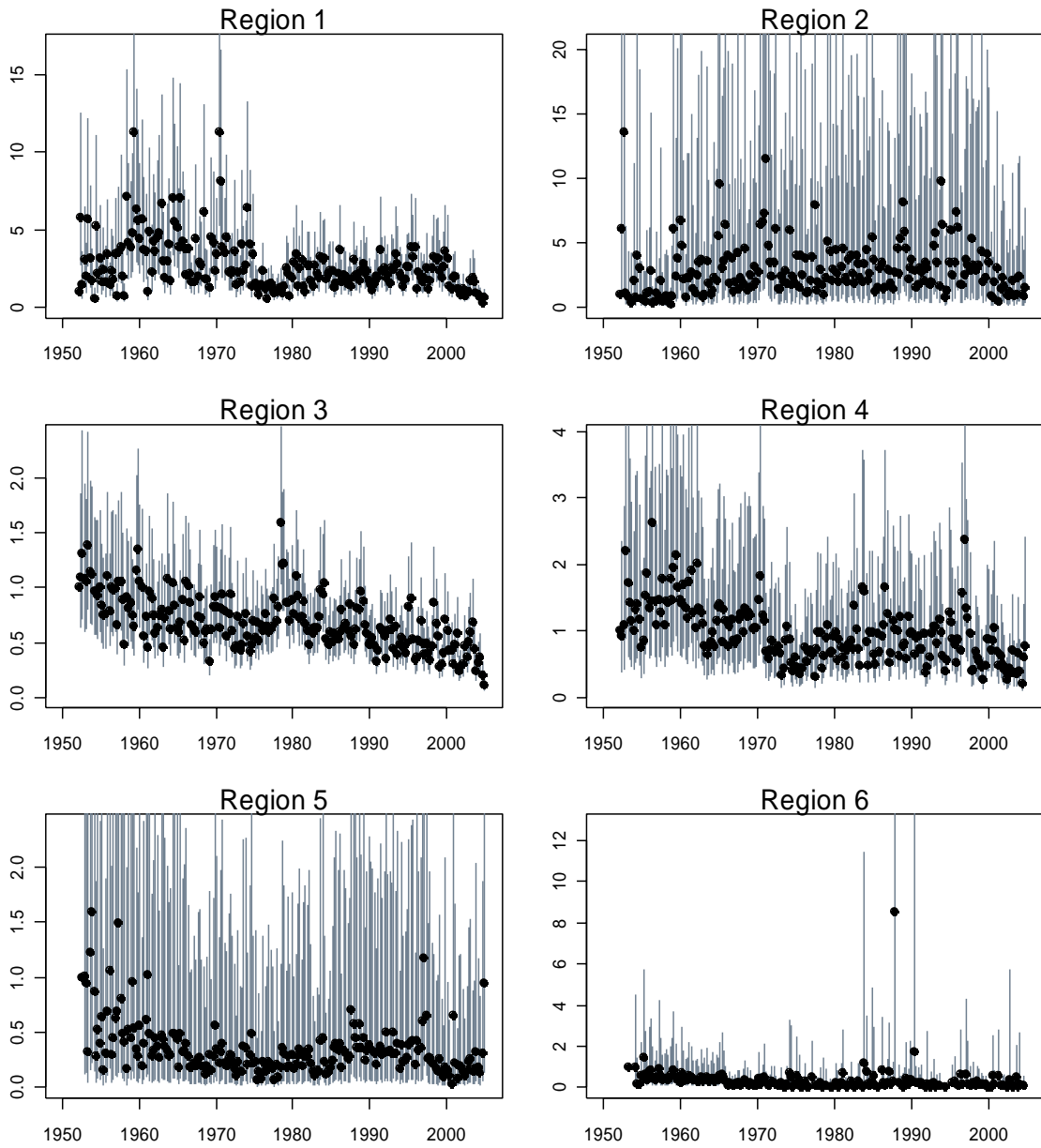


**Figure 7. Proportion of records with low residuals (less than -1.5) in the bigeye GLM models by year/quarter for each region. The line represents the lowest fit to the data points.**

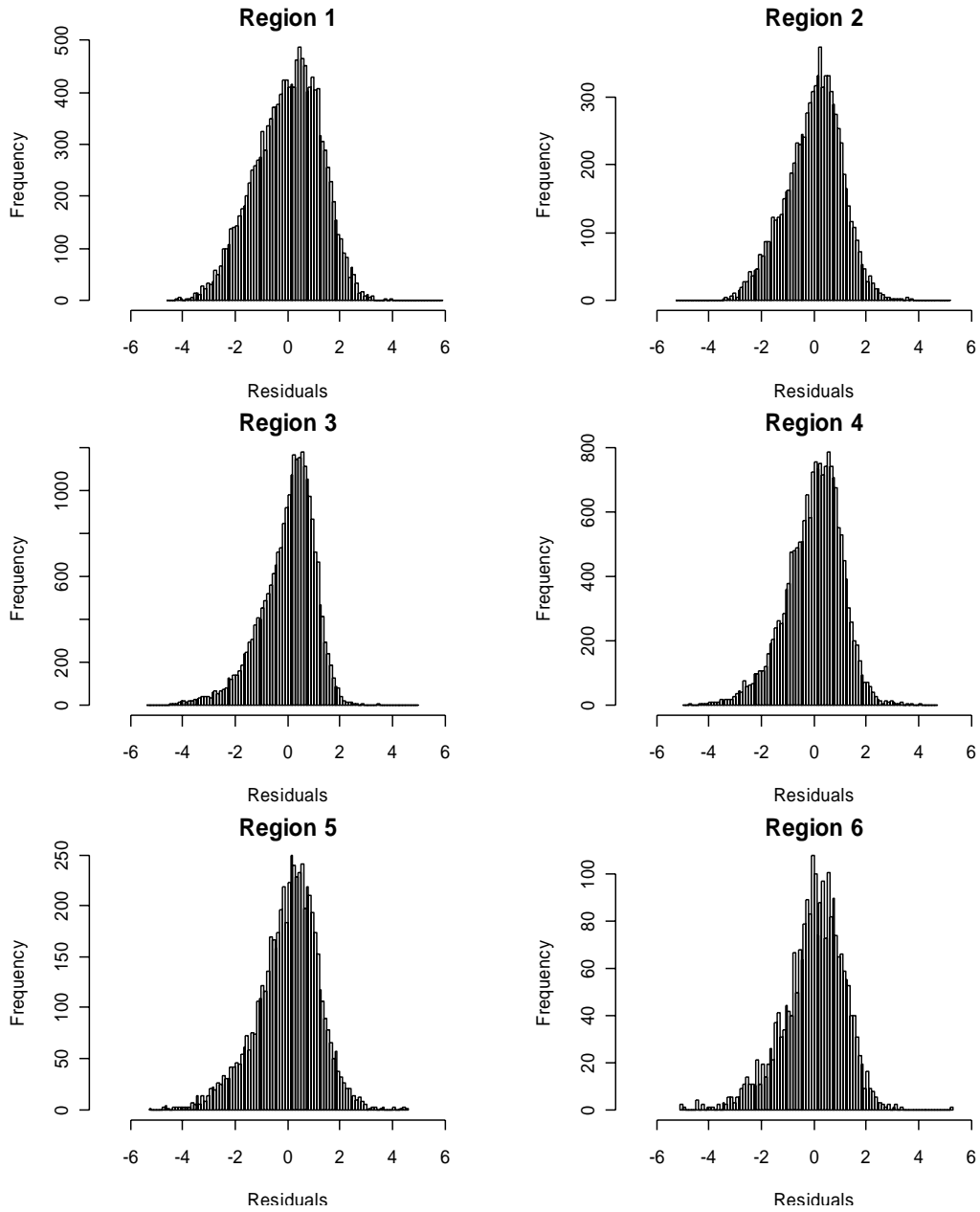


**Figure 8. A comparison of nominal CPUE (catch per hook) for bigeye and the quarterly indices derived from the GLM (normalised to the mean of the series).**

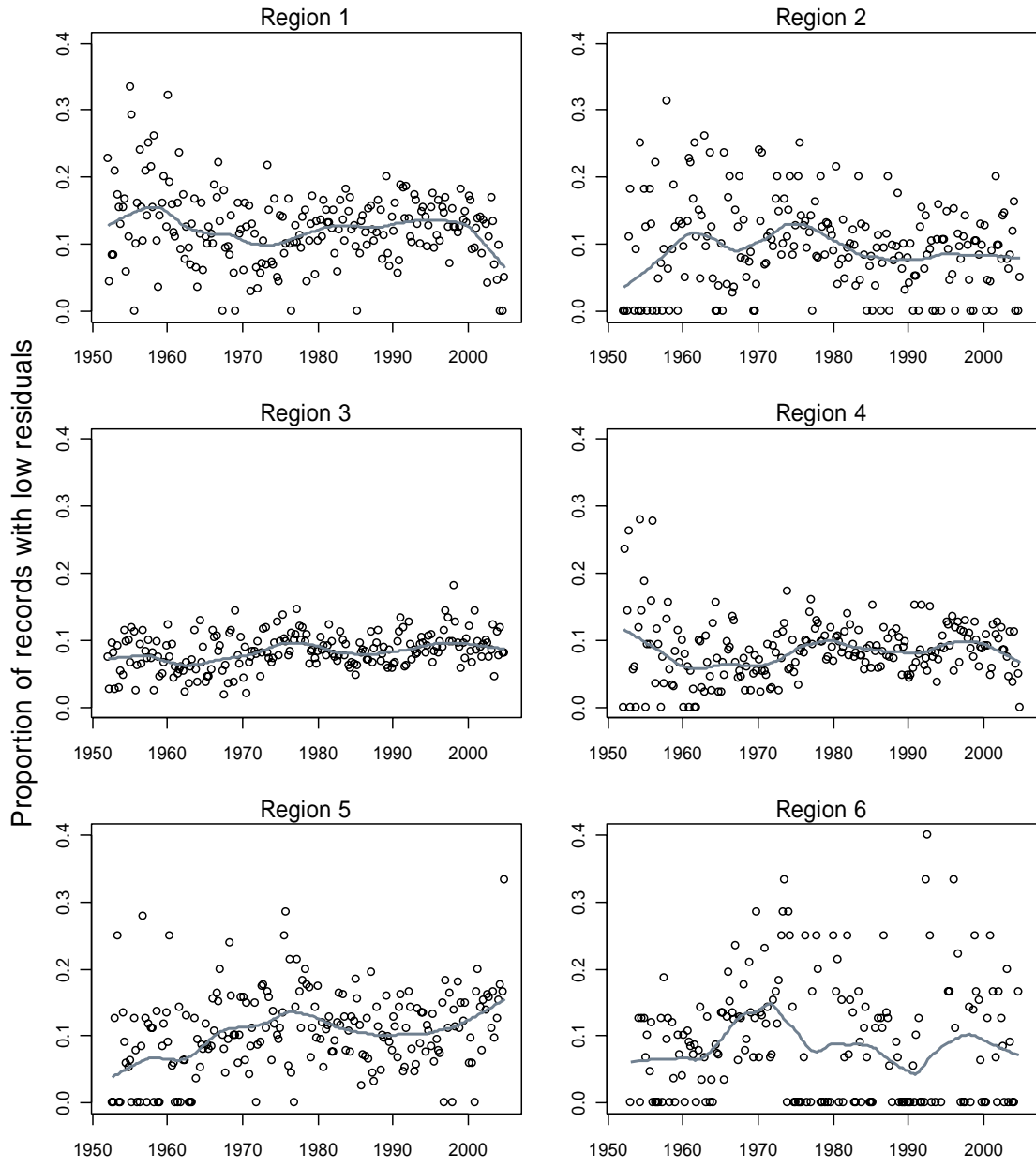




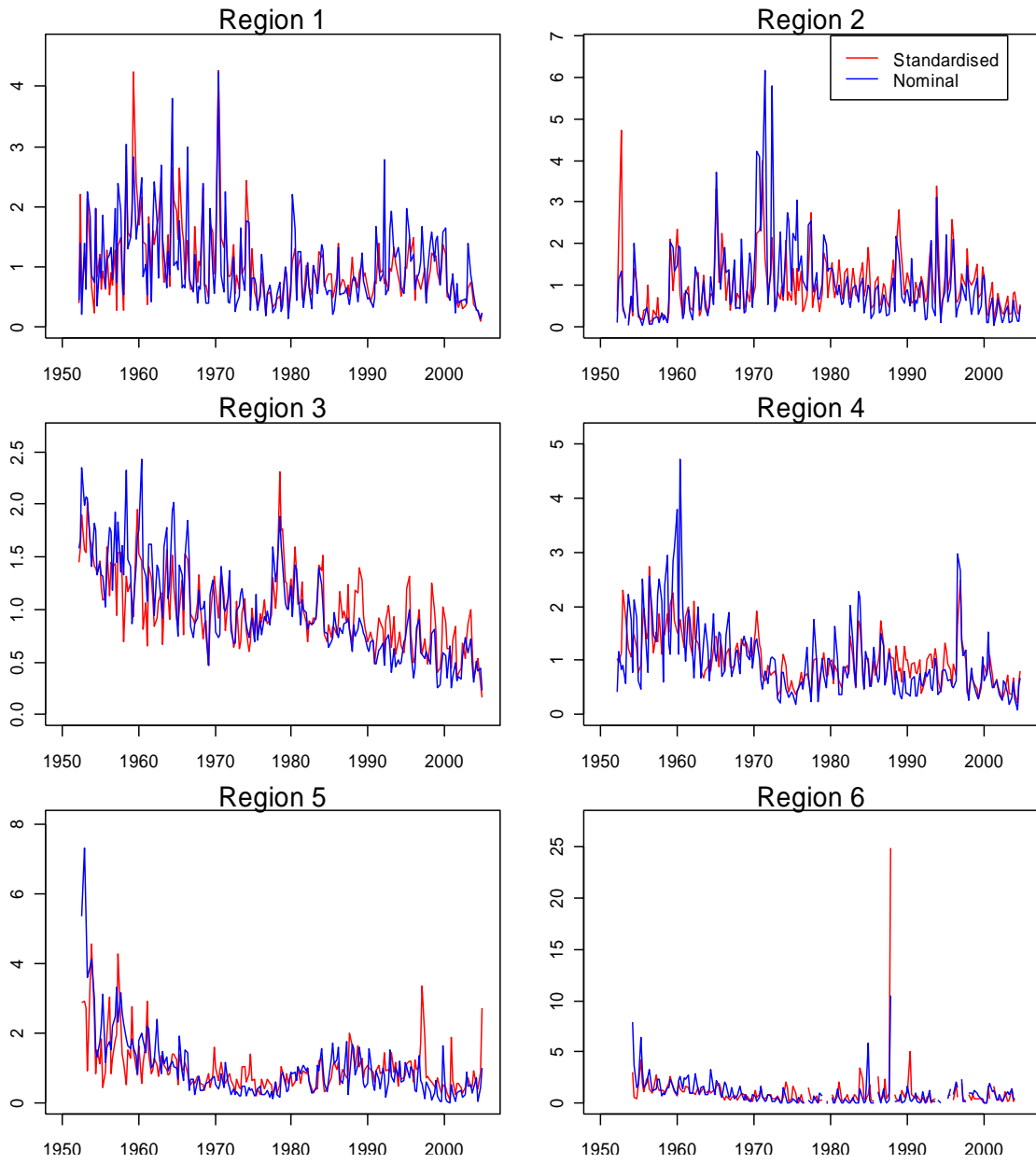
**Figure 9. Quarterly indices for yellowfin derived from the GLM for each MFCL region. The indices are standardised relative to the first quarter of the series. The vertical grey lines represent the confidence interval for the index ( $\pm 2$ \*standard error).**



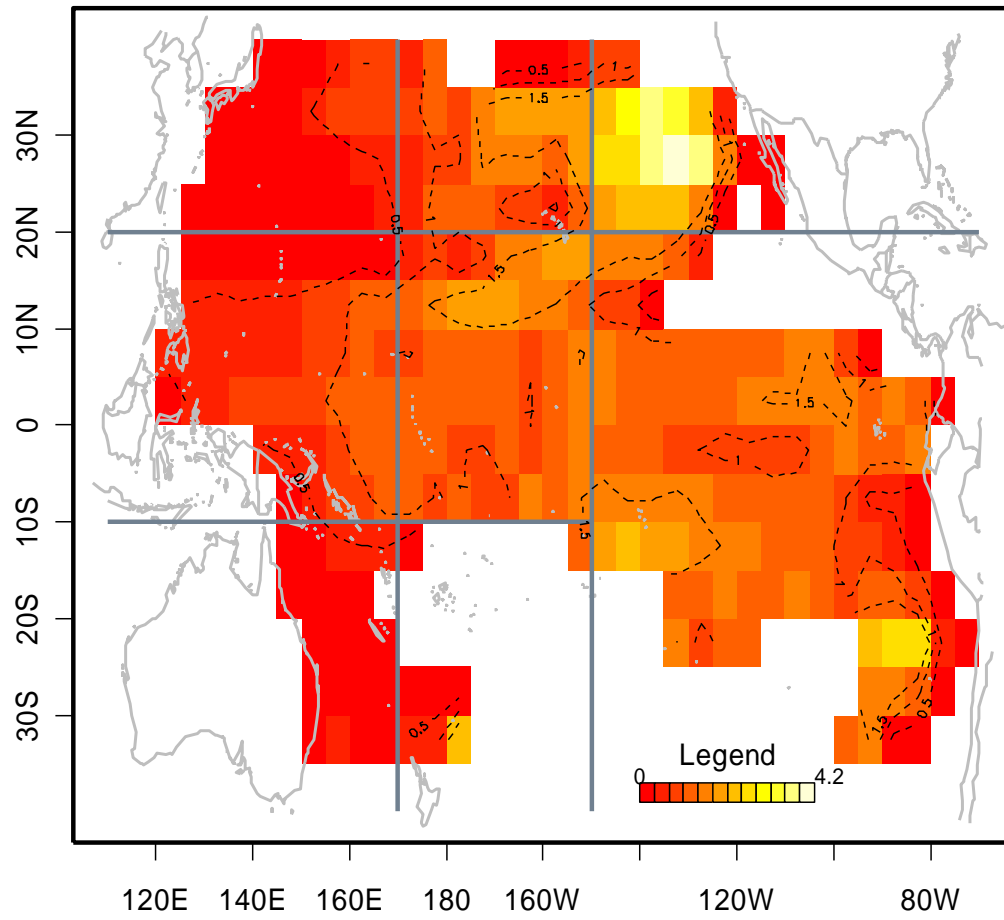
**Figure 10. Histograms of the residuals from the yellowfin GLM models by region.**



**Figure 11. Proportion of records with low residuals (less than -1.5) in the yellowfin GLM models by year/quarter for each region. The line represents the lowess fit to the data points.**



**Figure 12. A comparison of nominal CPUE (catch per hook) for yellowfin and the quarterly indices derived from the GLM (normalised to the mean of the series).**



**Figure 13. Pacific-wide bigeye statHBS latitude and longitude coefficients. Coefficients are scaled to the mean.**

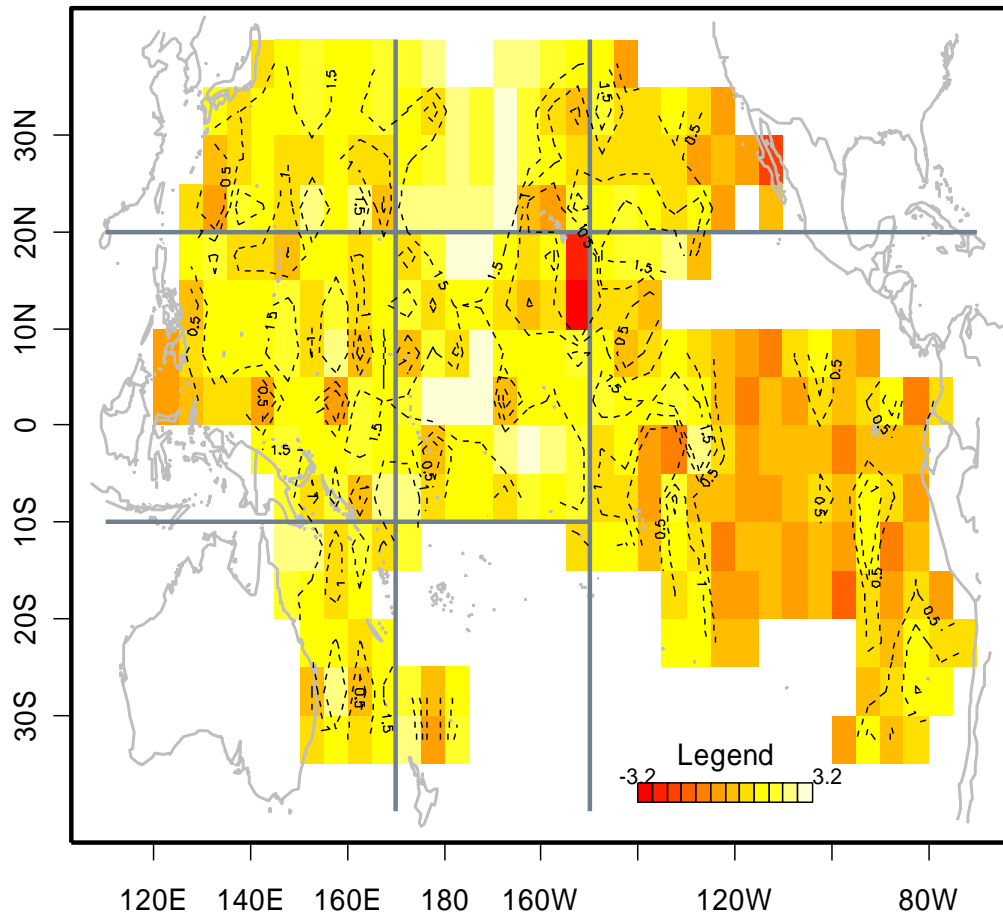


Figure 14. Spatial trend in the mean residuals of the Pacific-wide BET stathBS model fitted without the inclusion of a long/lat effect including data for 1975–2004 only. The mean residual for each cell is normalised to the mean value for all cells. Consistent positive (negative) residuals correspond to values above (below) 1.

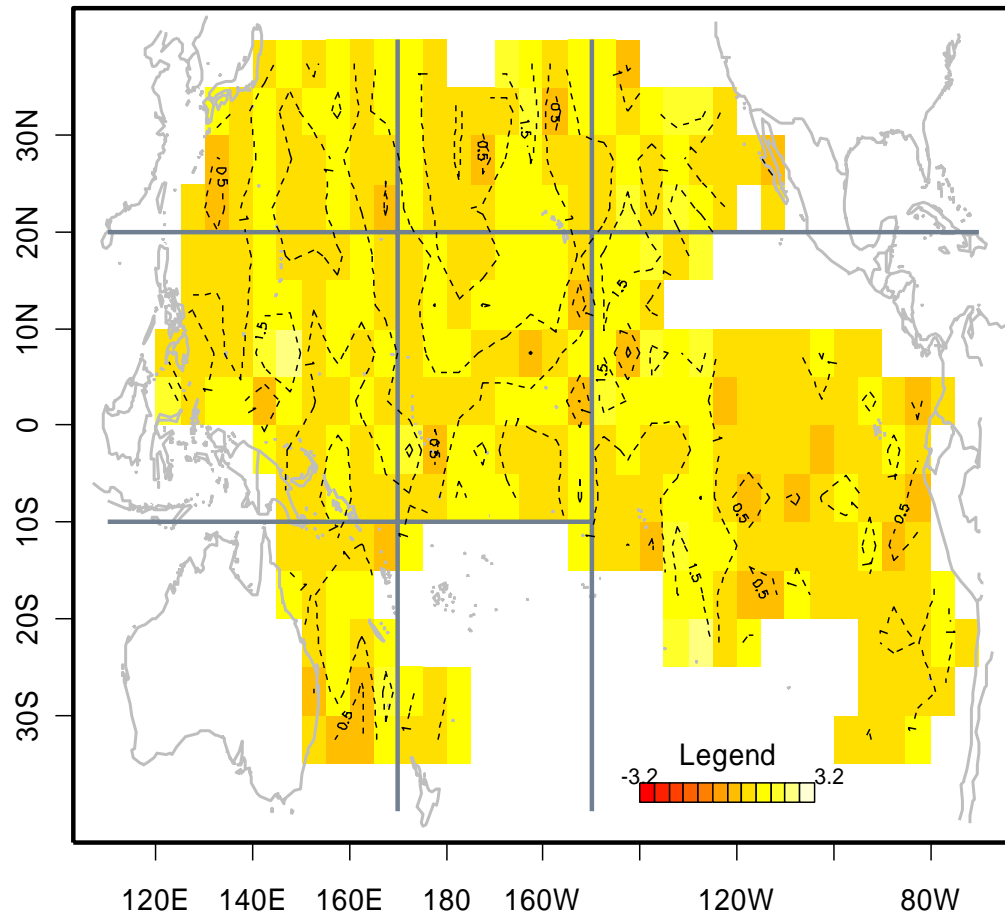
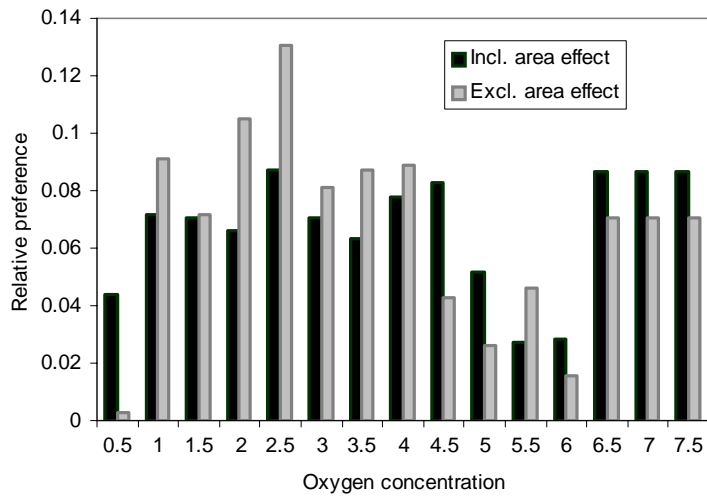
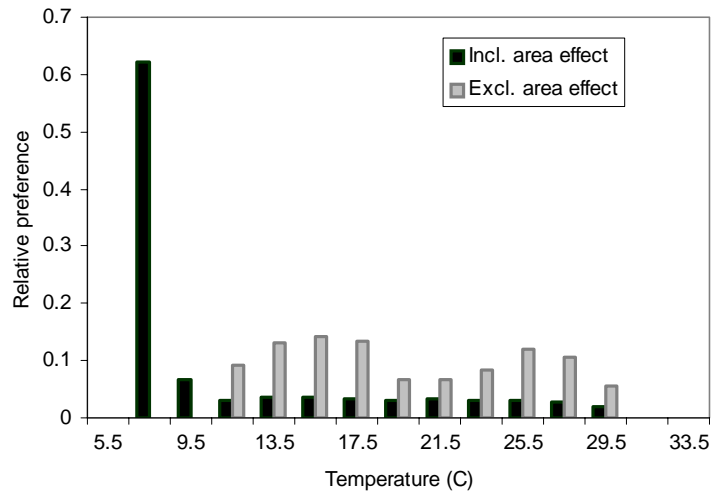
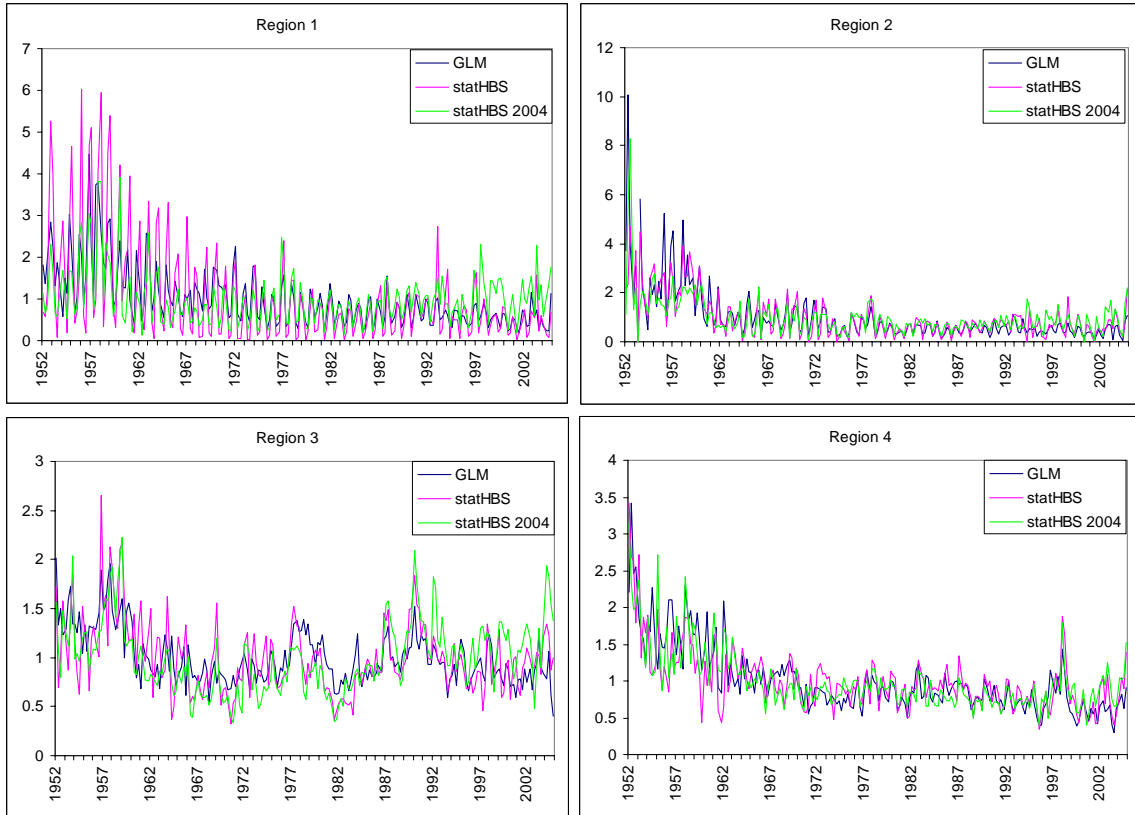


Figure 15. Spatial trend in the mean residuals of the Pacific-wide BET stathBS model fitted with the inclusion of a long/lat effect including data for 1975–2004 only. The mean residual for each cell is normalised to the mean value for all cells. Consistent positive (negative) residuals correspond to values above (below) 1. For comparison, the same scale is used as for the previous figure.

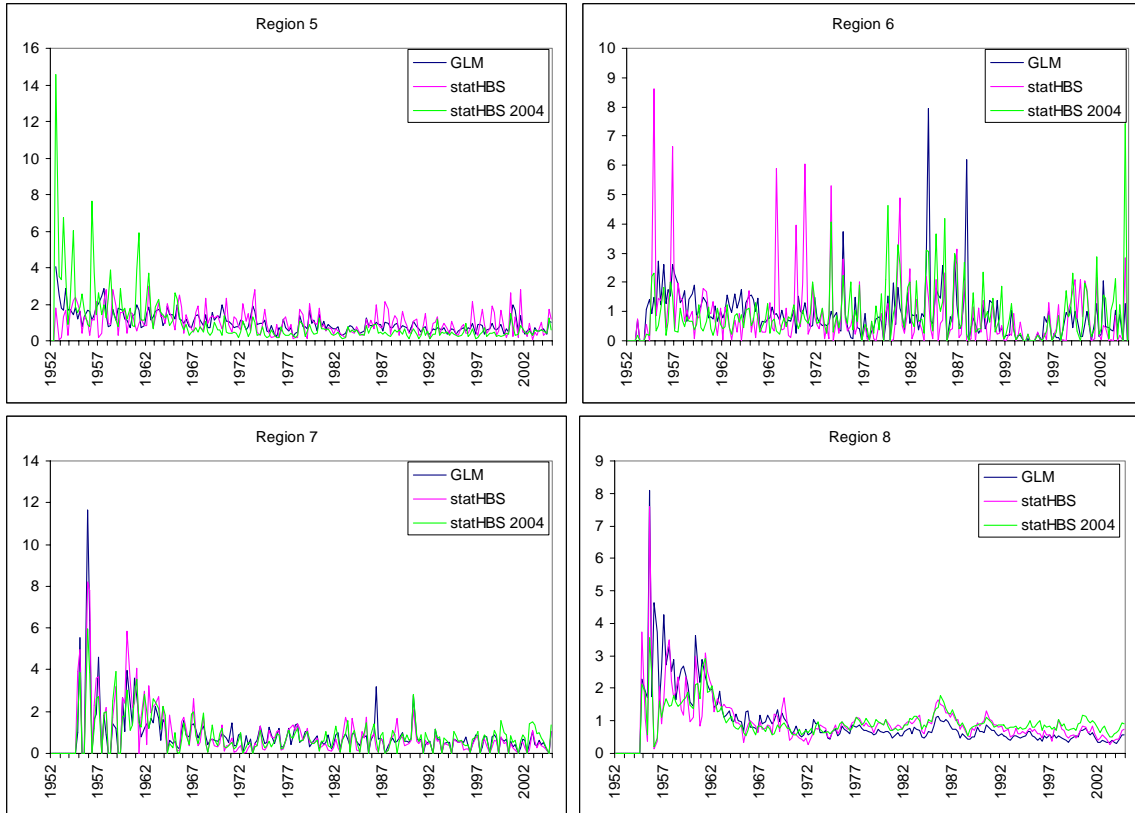


**Figure 16. Estimated bigeye habitat preferences with respect to water temperature (top) and oxygen concentration (ml l<sup>-1</sup>) (bottom) derived from the statHBS models including and excluding the area effect (lat/long interaction term). The value on the x-axis represents the maximum value for the individual category (2 C intervals for temperature, 0.5 ml l<sup>-1</sup> intervals for oxygen concentration).**

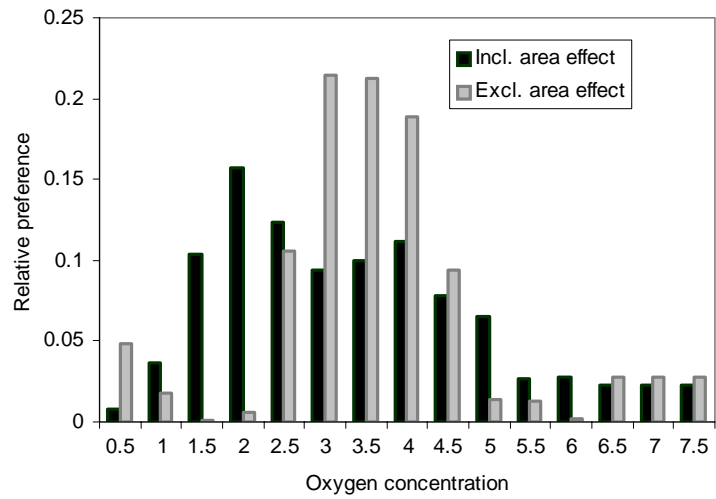
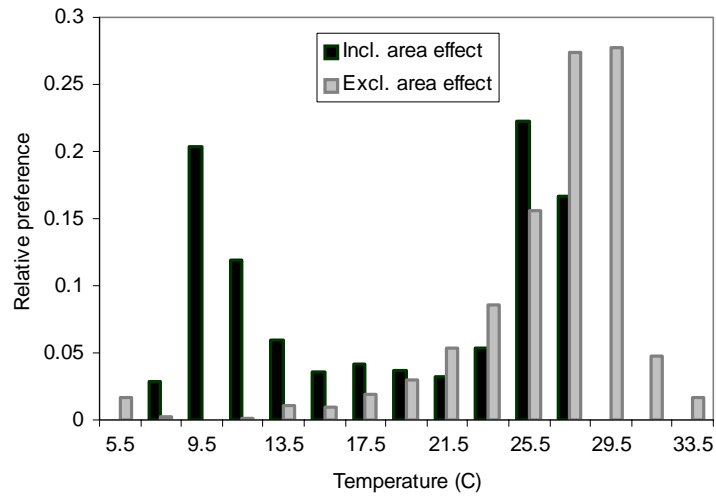




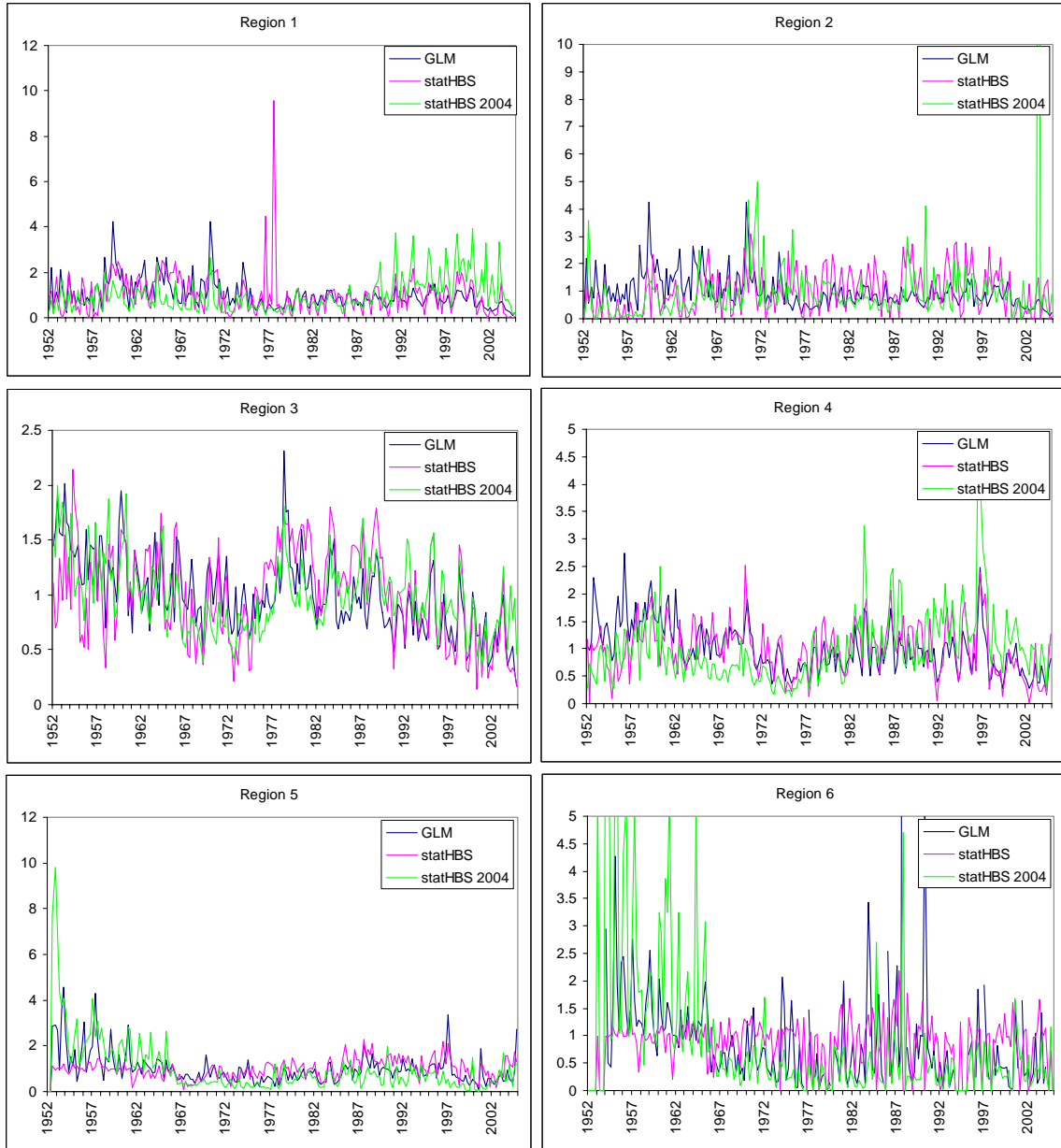
**Figure 17a. A comparison of the standardised CPUE indices for bigeye for the MFCL regions comprising the entire Pacific Ocean; GLM, statHBS, and statHBS following the approach used in Bigelow et al. 2004 (statHBS 2004) whereby habitat preferences were estimated separately for each region and no lat/long effect was included in the model. Each set of indices is scaled to the mean of the series.**



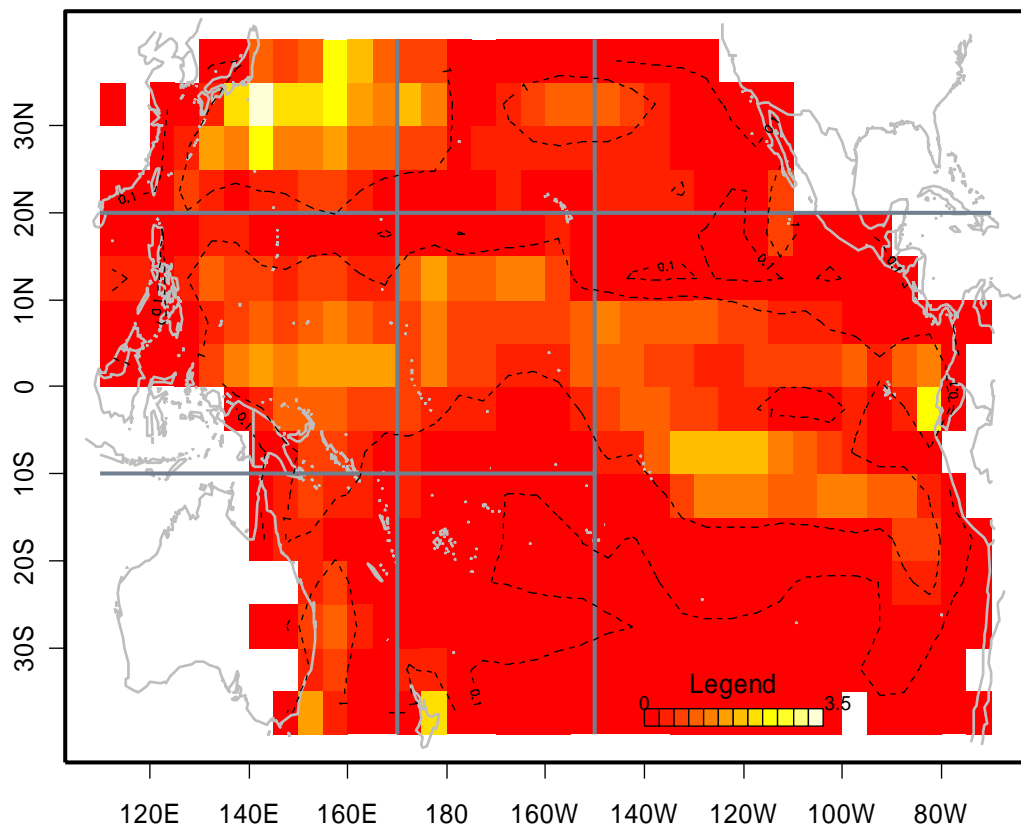
**Figure 17** Error! Reference source not found. **b. continued.**



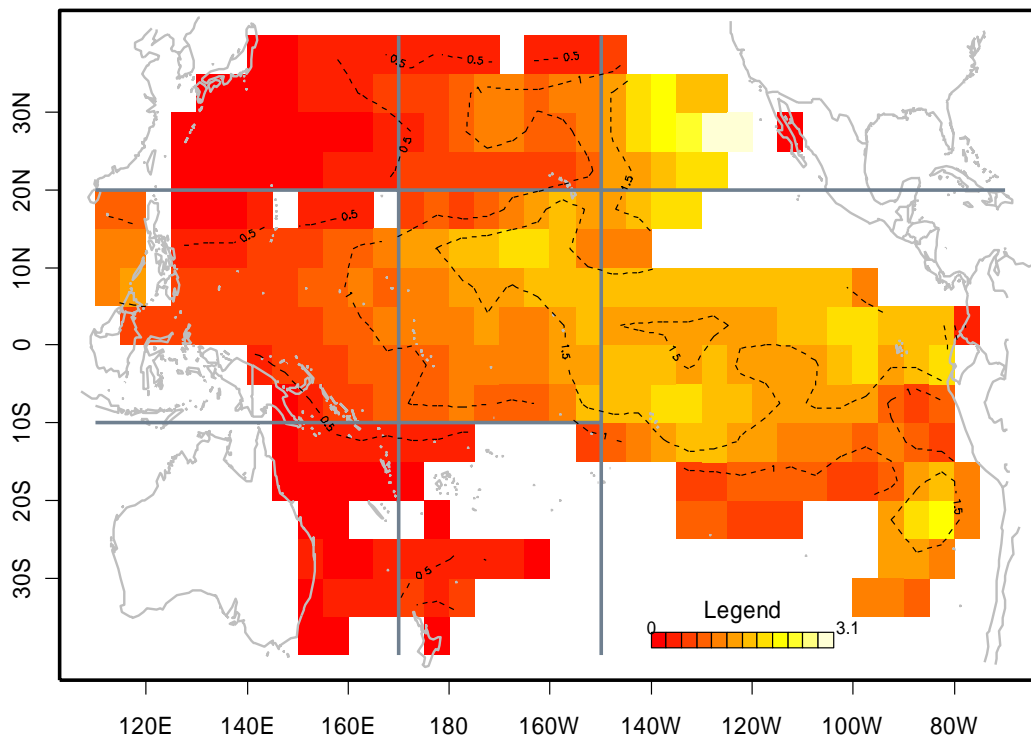
**Figure 18.** Estimated yellowfin habitat preferences with respect to water temperature (top) and oxygen concentration ( $\text{ml l}^{-1}$ ) (bottom) derived from the statHBS models including and excluding the area effect (lat/long interaction term). The value on the x-axis represents the maximum value for the individual category (2 C intervals for temperature, 0.5  $\text{ml l}^{-1}$  intervals for oxygen concentration).



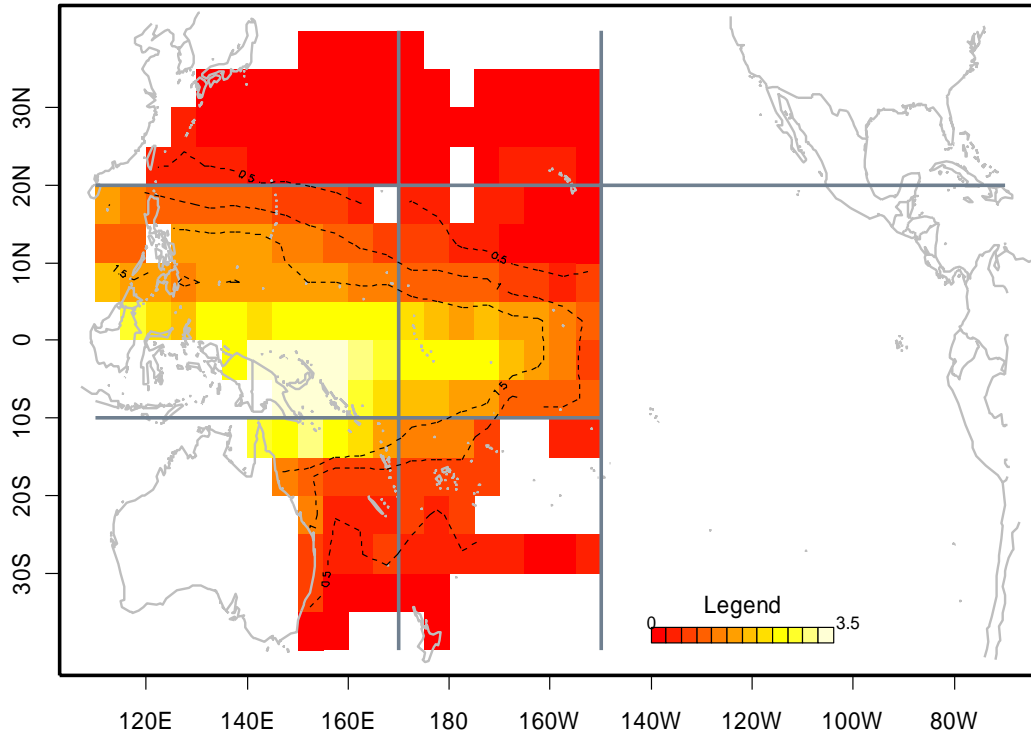
**Figure 19. A comparison of the standardised CPUE indices for yellowfin for the six MFCL regions comprising the WCPO; GLM, statHBS, and statHBS following the approach used in Bigelow et al. 2004 (statHBS 2004) whereby habitat preferences were estimated separately for each region and no lat/long effect was included in the model. Each set of indices is scaled to the mean of the series.**



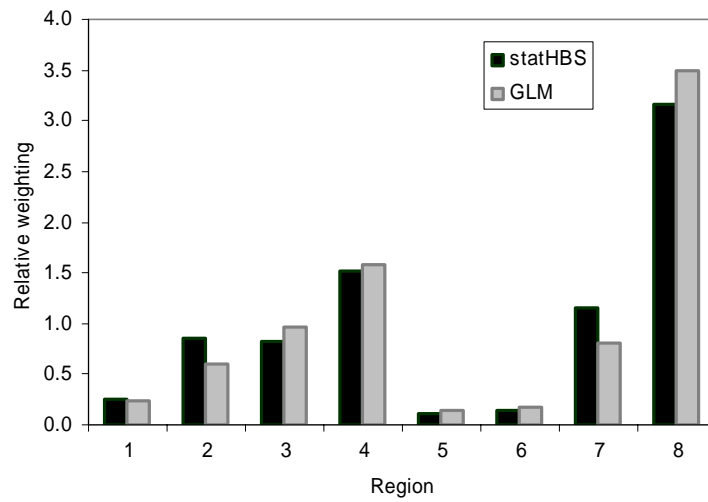
**Figure 20. Relative distribution of Japanese longline effort (number of hooks) in the Pacific Ocean during 1960–1986. The level of effort in each 5 degree square is expressed relative to the mean level of effort for the entire area. The contour lines represent the average level of effort and 10% of the average.**



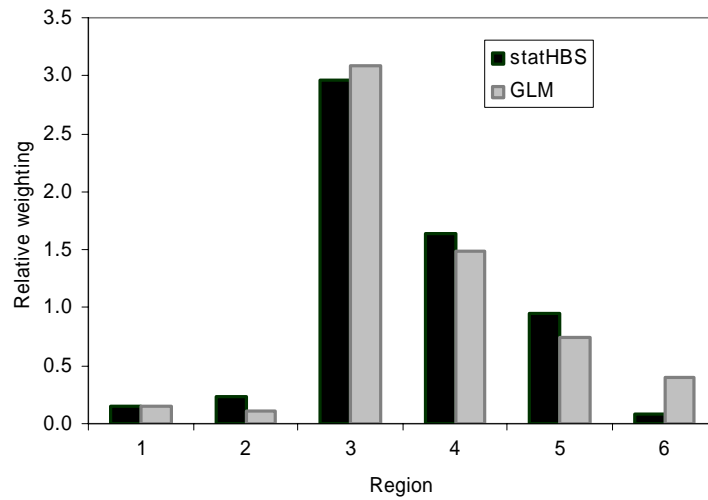
**Figure 21. Relative CPUE of bigeye for each 5 degree latitude and longitude cell as predicted from the GLM used to determine the relative region-specific scaling factors. The model lat/long coefficients are scaled relative to the mean of all coefficients. The contour lines represent 0.5, 1.0, and 1.5 times the mean value.**



**Figure 22. Relative CPUE of yellowfin for each 5 degree latitude and longitude cell as predicted from the GLM used to determine the relative region-specific scaling factors. The model lat/long coefficients are scaled relative to the mean of all coefficients. The contour lines represent 0.5, 1.0, and 1.5 times the mean value.**



**Figure 23. A comparison of the relative biomass weightings for the bigeye longline exploitable biomass from the statHBS and GLM standardisation approaches. The regional weightings are presented scaled to the average of the eight regions encompassing the entire Pacific Ocean.**



**Figure 24. A comparison of the relative biomass weightings for the yellowfin longline exploitable biomass from the statHBS and GLM standardisation approaches. The regional weightings are presented scaled to the average of the six WCPO regions.**