



**SCIENTIFIC COMMITTEE
SIXTEENTH REGULAR SESSION**
Online
11–20 August 2020

**Analysis of Pacific-wide operational longline dataset for bigeye and yellowfin tuna
catch-per-unit-effort (CPUE)**

WCPFC-SC16-2020/SA-IP-07

N. Ducharme-Barth¹ and M. Vincent¹

¹Oceanic Fisheries Programme, The Pacific Community

Contents

1 Executive summary	3
2 Background information	3
3 Data preparation	4
3.1 Data cleaning	4
3.2 Data rescue	5
3.2.1 HBF	5
3.2.2 JP fleet	6
3.3 Data summaries	7
3.4 Species clustering	7
3.5 Clustering results	8
3.6 Oceanography	8
4 Modeling approach	9
4.1 Computational challenges	9
4.1.1 Spatial knots	9
4.1.2 Sub-sampling	9
4.2 Vessel covariates	10
4.2.1 Vessel ID	10
4.3 Species clustering	10
4.4 Oceanographic covariates	10
4.5 Flag-group	11
4.5.1 HBF	11
4.6 Error structure	11
4.7 Model structure	12
4.7.1 delta-GLM	12
4.7.2 Spatiotemporal delta-GLMM	12
5 Results	13
5.1 Indices	13
5.2 Catchability covariates	14
5.3 Estimated spatial patterns	14
5.4 Diagnostics	15
6 Discussion	15
7 Acknowledgments	17
8 Tables	21
9 Figures	22

1 Executive summary

This paper summarizes the characteristics of the operational longline data-set and describes the data preparation, analysis, and spatiotemporal modeling required to create indices of relative abundance for the 2020 assessments of bigeye tuna and yellowfin tuna. A VAST delta-gamma mixed-effects spatiotemporal modeling approach with two catchability covariates *FLAG* and *hooks-between-floats (HBF)* was used. A random forest (RF) machine learning approach was used to predict *HBF* for records where it was not available. The resulting indices for both bigeye tuna and yellowfin tuna showed declines in many of the key stock assessment regions and were consistent with the trends seen in both the previous 2017 analysis and the nominal catch-per-unit-of-effort (CPUE).

A number of recommendations are made for future research to improve the quality of these indices:

- Data reconstruction to improve the number of covariates available across fleets for the duration of the model period.
- Interviews and surveys with vessel operators and captains to identify species specific targeting practices and relevant covariates to standardize on. This will guide model development and can direct changes to the observer program where needed to collect data on relevant covariates.
- Create a recent period index to take advantage of improved covariate coverage using either the observer data or a combination of observer and operational data with an expanded covariate set.
- Simulation study to determine the benefits and limitations associated with splitting the index into an early period and recent period, both in the ability to capture the underlying abundance trend and for assessing the implications of using a split index in a stock assessment.
- Revisit previous analyses of catch-per-effective-effort which account for the vertical distribution of tropical tuna biomass and longline hooks in the water column.
- Allocate funding to produce a spatiotemporal sub-surface oceanographic product which includes dissolved oxygen for the Pacific Ocean for the entire assessment period.
- The analysis of archival tagged, longline-vulnerable bigeye and yellowfin tuna needs to be continued and supplemented by additional releases in multiple locations across the WCPFC in order to complete our understanding of how vertical position in the water column is a function of fish length, time of day, and oceanographic conditions.
- Basic research on the depth distribution of longline hooks across vessels from different distant water fishing nations (DWFNs) and Pacific Island Countries and Territories (PICTs).

2 Background information

This paper describes the processes and analyses used to generate indices of relative abundance from fisheries-dependent longline operational (set-level) catch and effort data. These indices are used as inputs in the 2020 Western and Central Pacific Fisheries Commission (WCPFC) bigeye tuna and yellowfin tuna stock assessments ([Figure 1](#)). Given the spatial and temporal scale of the different longline fleets operating in the region, the resultant indices of abundance represent an important input to the stock assessments of both bigeye tuna and yellowfin tuna. The operational longline data-set is an amalgamation of operational level data from the distant-water fishing nations (DWFNs) and Pacific-Island countries and territories (PICTs) longline fleets operating in the Pacific

basin. This data-set is the most complete spatiotemporal record of longline fishing activity in the Pacific, spanning from 1952 to the present and is the result of a tremendous collaborative, data-sharing effort from the countries involved.

This data-set was first created in 2015 in support of the Pacific-wide bigeye tuna stock assessment (McKechnie et al., 2015a,b), and was subsequently analyzed to generate indices of relative abundance (McKechnie et al., 2017b; Tremblay-Boyer et al., 2017a; Tremblay-Boyer and Pilling, 2017) for the 2017 WCPFC bigeye tuna and yellowfin tuna stock assessments (McKechnie et al., 2017a; Tremblay-Boyer et al., 2017b). The 2015 analysis describes in detail the process used to identify species targeting via catch composition clustering, analyses on the inclusion of gear and vessel covariates that are unavailable for the full time period of interest, and choice of model error structure (McKechnie et al., 2015b). In 2017, work was done to identify methods for generating vessel identifier “proxies” for those records where vessel identifier was unavailable (Tremblay-Boyer and Pilling, 2017). Additionally, 2017 saw the first application of spatiotemporal modeling approaches (previously referred to as “geostats”) in support of WCPFC stock assessments (Tremblay-Boyer et al., 2017a). The current work builds off of these previous efforts and analyses of the operational longline data, as well as the spatiotemporal modeling done in support of the 2019 WCPFC skipjack tuna stock assessment (Ducharme-Barth et al., 2019; Kinoshita et al., 2019). For additional background and description of both the operational longline data-set and previous analyses please consult these earlier reports.

3 Data preparation

3.1 Data cleaning

The operational data-set consists of 10,897,686 longline set-level records from the commercial longline logbooks of 27 different fishing nations from a period of 1952 through the present day (Figure 2 - 3). In the consolidated data-set, each record contains the vessel’s *flag*, vessel’s *fleet*, *date*, *location* (to the nearest $1^\circ \times 1^\circ$ spatial cell), *effort* (number of hooks fished), and species catch (albacore tuna, bigeye tuna, striped marlin², swordfish, and yellowfin tuna) (Figure 4). Nominal catch-per-unit-of-effort (CPUE defined as catch per 100 hooks fished) indicate declines in catch rate over the model period in key regions for each species (Figure 5 - 6). *Hooks-between-floats* (HBF) and unique *vessel ID* are also available though coverage is incomplete, particularly in the early years. Prior to analysis, the data were screened similarly to the method first described in McKechnie et al. (2015b). This data cleaning process occurred in the following steps:

1. Removed records with missing year: -2 records (negligible %)
2. Removed records from 2019 or 2020: -98,064 records (-0.90 %)
3. Removed records from outside of the Pacific basin: -19,376 records (-0.18 %)
4. Removed records with more than 50 *HBF* outliers: -26,736 records (-0.25 %)
5. Removed records with number of hooks fished, per set, greater than 5,000 or less than 150: -7,499 records (-0.07 %)
6. Removed records with more fish caught than number of hooks fished: -18 records (negligible %)

²Striped marlin catch was unavailable for Japanese records prior to 1967

7. Removed records with vessels that did not fish at least 10 quarters³ or made less than 30 sets: -704,247 records (-6.46 %)
8. Removed records flagged under Belize (BZ), Panama (PA), or Spain (ES): -21,192 records (-0.19 %)
9. Removed Japanese records that could not be attributed to the offshore (OS) or distant-water (DW) components of their fishery: -39 records (negligible %)
10. Remaining records: 10,020,513 records (91.95 %)

Those familiar with the data cleaning process used by McKechnie et al. (2015b) will note that unlike previous analyses, this analysis retained records without any catch of the five reported species. In a spatiotemporal modeling framework, presence of zero catches can be informative of the species distributional limits assuming that there are no major changes to either species targeting or reporting. An examination of the location of these sets showed that close to half of them occurred south of 40° S latitude and beyond the distributional limit of bigeye tuna and yellowfin tuna. Though these sets are indicative of longline fishing effort directed towards southern bluefin tuna they can help provide information on the southern extent of the range of these species.

Records that were missing *HBF* were not excluded in order to explore the effects of filling in missing *HBF*. Records missing *vessel ID* were retained since *vessel ID* proxies were not considered for standardization.

3.2 Data rescue

Given the limited number of covariates that were available across fleets for the totality of the assessment period, it was necessary to “fill-in” missing covariates particularly if they were systematically missing from segments of the fishery. Otherwise, including these covariates using a “dummy” level for the records where the covariate was missing risked introducing temporal discontinuities in the standardized index.

These data rescue operations were conducted to salvage missing observations of *HBF* (predominantly from Japanese records prior to 1967, but also records from a variety of flags in assessment region 6; [Figure 7](#)), and also missing fleet identifiers, OS or DW, for Japanese records prior to 1957. In both cases a formulation of the Random Forests (RF) machine learning approach ([Breiman, 2001](#)), Ensemble Random Forests or ERF ([Siders et al., 2020](#)), was applied to predict the missing observations. RF is a computationally efficient machine learning algorithm capable of handling large quantities of “training” data, non-linear interactions between covariates, and can produce predictions either via regression or multiple-category classification.

3.2.1 HBF

From records where *HBF* was recorded we noticed patterns in the distribution of *HBF* ([Figure 8](#)). We developed flag-specific ERF models which predicted the *HBF* bin (bins of 5 *HBF*) using fleet id, year, month, longitude, latitude, number of hooks fished, total catch in numbers, and proportion of species catch for albacore tuna, bigeye tuna, striped marlin, swordfish, and yellowfin tuna. Predictions were made at the 5 hook bin level, rather than the individual *HBF* level, in order to improve the classification accuracy of the algorithm. Bins of 5 were selected for analysis based on the recommendation of the 2020 Pre-Assessment Workshop (PAW) as a compromise between capturing

³Vessels that entered the fishery within the last 10 quarters of the model period were not subject to this exclusion.

fine-scale changes in CPUE as a result of different *HBF* fished and not overly degrading classification accuracy by presenting the model with too many potential categories. A brief description of the methodology and data processing for developing the flag-specific predictions of *HBF* follows.

1. Subset total dataset to data from a single flag.
2. Randomly partition the data where *HBF* was recorded into two segments: 90% for training the RF classification model and 10% for validating the RF classification.
3. For the training subset, select 100 random subsets of the data to be passed to each RF in the ensemble.
4. Fit a RF model with 500 trees to each of the 100 random partitions of data.
5. These 100 RF models (50,000 trees) are aggregated into a single, ensemble model.
6. The ensemble model is then used to predict on the data held aside for validation. Classification accuracy is assessed using the accuracy as measured by the True Positive classification rate and Cohen's κ .
7. The ensemble model is applied to the data where *HBF* was missing to generate predictions of *HBF* fished at 5 hook intervals.
8. This process is repeated for all flags. The flag-specific predictions of *HBF* for the records where *HBF* had been missing ($n = 2,342,441$) are then combined with the observed *HBF* that had been rounded up to the nearest value of 5 to generate a complete record of *HBF* at 5 hook intervals.

One of the challenges faced with predicting missing *HBF* bins is that for each flag, observations of *HBF* are not distributed evenly across the 10 bins (0-5, 6-10,..., 46-50). For example, records from New Caledonian flagged vessels rarely showed observations outside of the 26-30 *HBF* bin. A model trained on this data would be dominated by these 26-30 *HBF* observations and absent sufficient contrast in the data may perform poorly on predicting the rarer observations. To compensate for this, the data can be re-balanced or “over-sampled” (resampling minority classes with replacement) so that the model is trained on data where each class is represented equally. This alternate, *over-sampling* approach was considered in addition to the default approach where the models were trained on data that maintained the ratio of classes in the original data-set (*ratio-maintained* approach).

Performance of the classification algorithm was satisfactory. Given the validation datasets, the median classification accuracy across flags was 84% for the *ratio-maintained* approach, and 77% for the *over-sampling* approach. Cohen's κ indicated moderate performance (McHugh, 2012), the median Cohen's κ across flags was 70% for the *ratio-maintained* approach, and 62% for the *over-sampling* approach. Not all flags had the same number of observations or the same number of missing observations. Accounting for these differences did result in lower accuracy (78% *ratio-maintained*; 70% *over-sampling*) and Cohen's κ (68% *ratio-maintained*; 60% *over-sampling*). This indicates that the fishing signature of flags with fewer observations may have been easier to isolate, possibly as a result of reduced variability in fishing location and/or target species.

3.2.2 JP fleet

Japanese records with fleet (OS or DW) missing ($n = 250$ 147) were predicted with the same classification approach described for the multi-class *HBF* problem with the exception that only

the ratio-maintained approach was considered. Data were processed according to the steps above though the classification model differed slightly. Missing fleet prior to 1957 was predicted as a function of month, longitude, latitude, number of hooks fished, total catch in numbers, and proportion of species catch for albacore tuna, bigeye tuna, swordfish, and yellowfin tuna. The training dataset were the Japanese records with fleet recorded from 1958-1967.

Performance of the classification algorithm was satisfactory. Given the validation dataset, the classification accuracy was 91% and Cohen's κ of 81% indicated almost perfect performance (McHugh, 2012).

3.3 Data summaries

Spatiotemporal summaries of the data along with the response of CPUE (catch in numbers of individuals per 100 hooks fished) to different gear/vessel characteristics is shown in the following figures. The spatial distribution of nominal CPUE for bigeye tuna (Figure 5) and yellowfin tuna (Figure 6) both appear to indicate declines in the intensity of hot-spots of nominal CPUE. For bigeye tuna, these hot-spots are located along the equatorial counter-currents (particularly in the eastern Pacific Ocean), and around the Hawaiian islands. Yellowfin tuna shows a hot-spot in the equatorial western and central Pacific Ocean. The response of nominal CPUE (by species) to *HBF* (Figure 9) and *vessel ID* (Figure 10), when these covariates were recorded, show non-linear relationships for the different flag groups. The response of catch by species to hooks fished per set (Figure 11) appears to be positive, increasing, and more consistent across DWFN fleets.

3.4 Species clustering

Previous CPUE analyses of the operational longline data-set clustered records according to species catch composition by "trip". Clustering was done at the "trip"-level rather than the set-level as random variation between sets targeting the same species could result in very different species composition. Clustering at the "trip"-level was thought to better capture the overall targeting patterns while removing some of the set-level random variation. In the 2015 analyses, the CPUE data were modeled independently within each assessment region and this species clustering variable was used to subset data to records that were likely targeting bigeye tuna or yellowfin tuna (McKechnie et al., 2015b). The clustering was done independently within each assessment region. Clustering was also done by decade to determine whether clusters had remained static over time. In 2017, the species cluster was included in the CPUE standardization model as a catchability covariate. It was thought that the difference in species landed within a "trip" was indicative of subtle changes to the gear configuration that enhanced the catchability for the target (or most frequently caught) species.

Species clustering of bigeye tuna, yellowfin tuna, albacore tuna, and swordfish catch at the "trip"-level was identified using k-means clustering similar to what was done for the previous analyses (see McKechnie et al. 2015). Since the designation of a unique trip was unavailable, when *vessel ID* was available, "trips" were defined as all records from the same vessel within the same month. In the absence of *vessel ID*, the previous analyses defined trips as all records from the same *flag*, month, and $1^\circ \times 1^\circ$ spatial cell. For the current analysis, the spatial cell definition was relaxed to $5^\circ \times 5^\circ$ since this more accurately reflects the spatial footprint of a longline vessel's operations within a one month period.

Alternative clustering approaches to the k-means clustering algorithms could yield different results. However, given the magnitude of the number of observations and the computational demands of the other approaches, k-means was felt to be the best alternative at this stage. K-means clustering

requires an *a priori* assumption on the number of groups to cluster records into, and clustering with 2 - 7 groups were tried. Each algorithm, was “jittered” 30 times to find a globally optimal, stable clustering solution.

3.5 Clustering results

Given the potential number of groups attempted (2 - 7), the optimal number of clusters was determined to be 4 following a visual inspection of the within cluster sum-of-squares ([Figure 12](#)). These four clusters can be broadly categorized as bigeye tuna, “other” (mainly swordfish), yellowfin tuna, and albacore tuna clusters and are largely stable across time and stock assessment region ([Figure 13](#)). There are two notable exceptions to this: 1) the bigeye tuna and yellowfin tuna clusters in assessment regions 5 & 6 contain a large proportion of albacore tuna catch and 2) there is an increasing trend in swordfish catch across time which coincides with a decrease in catch of striped marlin for the “other” cluster. Looking at the spatial distribution of cluster by decade ([Figure 14](#)), the distributional patterns of cluster grouping are largely consistent and appear to correlate strongly with perceived abundance patterns of the 3 major tuna species: albacore tuna, bigeye tuna, and yellowfin tuna.

3.6 Oceanography

Oceanography was first incorporated in CPUE analysis for WCPFC assessments by Tremblay-Boyer et al. (2017a). One of the benefits of using spatiotemporal models under the VAST framework (Thorson et al., 2015; Thorson, 2019) is that oceanographic covariates can either be included as an effect on abundance or as an effect on catchability that is then “standardized” out of the abundance predictions. The previous analysis utilized subsurface temperature profiles from the ECMWF ocean reanalysis system ORAS4 (Balmaseda et al., 2013) to calculate two intermediate variables: depth of the 15° C isotherm and the depth difference (in meters) between the 12° C and 18° C isotherms ($\Delta \text{ depth } (12^\circ - 18^\circ C)$). Both variables were included in the analysis as catchability covariates implemented as splines in the standardization model. The depth of the 15° C isotherm was selected as it related to both bigeye tuna and yellowfin tuna abundance, and $\Delta \text{ depth } (12^\circ - 18^\circ C)$ characterizes the vertical compression of the subsurface environment and is potentially indicative of availability to the gear.

Unfortunately, the current analysis was not able to replicate the previous oceanographic analysis as the ECMWF oceanographic product used is unavailable beyond 2015. The current analysis made use of a similar product which spans the entire model period: the monthly $1^\circ \times 1^\circ$ EN4 quality controlled subsurface ocean temperature and salinity objective analysis (Good et al., 2013; Gouretski and Reseghetti, 2010). In addition to the two variables previously calculated, the following variables were also derived: salinity (5 m), the difference in salinity from the surface to 200m ($\Delta \text{ salinity } (5 - 200\text{m})$), salinity variability in the upper 200m ($\sigma \text{ salinity } (<200\text{m})$), temperature variability in the upper 200m ($\sigma \text{ temperature } (<200\text{m})$), and temperature (5 m).

Salinity itself has not been shown in previous studies to be an important factor in the vertical distributions of either bigeye tuna or yellowfin tuna (Arrizabalaga et al., 2015; Song et al., 2008, 2009). However a derived variable based on salinity, such as variability of salinity in the upper 200m of the water column, could be indicative of mesoscale oceanographic features that affect abundance or catchability.

Lastly, subsurface dissolved oxygen is likely to be a very important oceanographic covariate in its ability to explain vertical compression of available habitat in the water column and thus availability

to the gear. Unfortunately, a suitable oceanographic product that has the appropriate spatial and temporal resolution for the entire model region and time period was unavailable. This represents a critical gap in any modeling that attempts to account for oceanographic variability.

4 Modeling approach

4.1 Computational challenges

As described in previous works ([McKechnie et al., 2015b](#); [Tremblay-Boyer et al., 2017a](#)), the volume of records in the operational longline data-set poses some unique computational challenges. Furthermore, modeling the data in a spatiotemporal framework using the VAST package is slower than when compared to a traditional delta-lognormal generalized linear model (GLM) without interactions between the spatial and temporal terms. The run time for a VAST model is determined, in part, by three “axes”: 1) the number of spatial effects or “knots” in the model, 2) the number of time steps estimated, and 3) the amount of data that the model is being fit to. Given the temporal span (1952-2018) and the quarterly time-step of the dynamics in the assessment model, there is no flexibility on the 268 time steps so computational gains have to come from the other two “axes”.

4.1.1 Spatial knots

Ideally, there would be a spatial effect estimated at least for each $5^\circ \times 5^\circ$ spatial cell in the model region. However, at the scale of the Pacific-basin, this results in ~ 600 spatial effects which quickly balloons to 321 600 spatiotemporal random effects (knots \times time steps \times 2 components in the delta-model). While there may be data available to estimate this many spatiotemporal effects, such a model is not feasible with current computational resources. In order to balance computational feasibility with an appropriate spatial resolution, a model with 150 spatial knots was selected. This estimates a spatial effect for approximately each $10^\circ \times 10^\circ$ spatial cell, and $\sim 90\,000$ spatiotemporal random effects. Sensitivity analyses showed that the estimated indices were robust to models with 100, 150, & 175 knots except for some slight variability in the smallest region, region 9. Based on the recommendations from the 2020 PAW, 4 additional knots were added around regions 8 & 9 in order to increase the spatial resolution of the model and produce more robust indices for the smallest assessment regions ([Figure 15](#)).

Recent developments to the VAST package have added the capability for bi-linear interpolation of abundance between spatial knots. Though this produces “prettier” (i.e. smoother) abundance distributions, applying this capability does not meaningfully change the estimated index given the spatial scale of the assessment regions though it does increase the computation time. Based on this, the feature was not used in the current analysis

4.1.2 Sub-sampling

Another approach to managing the computational challenge presented by this data-set is to randomly sub-sample the number of records to reduce the computational overhead of the model. This is the approach that was applied in [McKechnie et al. \(2017b\)](#); [Tremblay-Boyer et al. \(2017a\)](#); [Tremblay-Boyer and Pilling \(2017\)](#). In the current analysis, we randomly sub-sampled 5 observations within predefined strata of time step \times spatial cell/knot \times Flag-group. This *post hoc* stratification resulted in a more tractable, and more spatio-temporally balanced final dataset of 572,395 records (5.71% of total records). Even with the resampling, the dataset is still more heavily weighted towards the tropical regions due to some knots having a higher number of Flag-groups operating

within their boundary (Figure 16; Shown for yellowfin tuna. The same data-set was used for bigeye tuna.).

Sensitivity analyses showed that indices were robust across a range of samples per strata (3, 5, and 6), and across random sub-samples of data with the same number of samples per strata.

4.2 Vessel covariates

4.2.1 Vessel ID

Vessel ID represents an important variable that could explain changes in catchability over time as seen in the 2019 CPUE analysis of Japanese pole-and-line skipjack tuna (Kinoshita et al., 2019). Similar to *HBF*, *vessel ID* is missing for a large proportion of the observations, particularly early in the model period. Previous work (Tremblay-Boyer and Pilling, 2017), explored the use of *vessel ID* proxies in the CPUE standardization model. These proxies were constructed by clustering observations according to the date of the record, location, hooks fished, *HBF*, and species catch composition. Feedback from members of the 2020 PAW placed this as a lower priority item and it was not given further consideration. It remains an area of future research.

4.3 Species clustering

Though species clustering could contain information about gear configuration related to increased catchability of target species, the spatial patterns in species cluster are highly correlated with patterns in species abundance. Given that the proposed spatiotemporal modeling approach explicitly accounts for the spatial structure of the data, and that all assessment regions were modeled in a single unified analysis; we did not feel like it would be appropriate to include the species clustering variable given the potential for spatial confounding with abundance.

4.4 Oceanographic covariates

As mentioned before, one of the advantages to the VAST modeling framework is the inclusion of oceanographic covariates either as abundance or catchability covariates. Analyses on the inclusion of oceanographic covariates were limited to 3 variables that showed the least correlation: temperature (5 m), depth of the 15° C isotherm , and salinity variability (<200m). Absent specific feedback from the 2020 PAW, oceanographic covariates were assumed to impact abundance in model runs. Oceanographic covariates were included as a polynomial spline with 3 degrees of freedom using the B-spline basis in the *splines* package (R Core Team, 2019).

Sensitivity analyses to the inclusion of abundance covariates using the bigeye tuna model showed that they had no impact on the trend of the estimated indices. However, the estimated regional weights (driven by the magnitude of the biomass estimates in each region) showed some variability to the inclusion of oceanographic covariates and their assumed functional forms. Given this uncertainty, and the fact that the perceived most influential covariate, sub-surface dissolved oxygen, was unavailable; oceanographic covariates were not included in the standardization model. Further research is recommended to ensure that the relationship between oceanographic covariates, bigeye tuna & yellowfin tuna abundance and longline CPUE is modeled appropriately, and to see how differing estimates of regional weights influence the assessment outcomes.

4.5 Flag-group

The Flag groupings used in the standardization models are defined in [Table 1](#). The distant water fishing nations records for Japan, Chinese Taipei, and the United States were split by fleet group based on differences in the area of operation and/or gear configuration. The Fiji charter vessels were also split apart from the domestic Fiji records. Records from PICTs and Indonesia were grouped into 3 categories, broadly corresponding to pelagic eco-regions in the WCPO ([Longhurst, 1999](#)) and patterns in species catch composition: countries with exclusive economic zones (EEZ) overlapping with the northern extent of the western Pacific warm pool (NP), countries with EEZs mainly within the equatorial waters of 10°N and 10°S latitude (EQ), and countries with EEZs primarily south of 10°S latitude (SP).

4.5.1 HBF

Longline fishers are able to manipulate the characteristics of their gear in order to target depths associated with particular species, often on a set-by-set basis. This is done through a combination of adjustments to the line setting speed, float line length, branch line length and *HBF*. Even still, variability in current, surface winds, and water density can result in longlines with the same “configuration” fishing at effectively different depths. Additionally, the material of the mainline will also affect the position of hooks in the water column though this is unlikely to change from set-to-set. This is information commonly recorded by onboard observers, however it is largely unavailable (except for *HBF*) in the operational longline data-set consolidated across all flags. As a result, *HBF* is the only available covariate which could be used to model the effects of gear configuration on bigeye tuna and yellowfin tuna catch rates.

Based on the recommendation of the 2020 PAW, *HBF* “bins” were modeled using a polynomial spline with 3 degrees of freedom using the B-spline basis in the *splines* package ([R Core Team, 2019](#)). This enforced a correlation structure on the data such that adjacent *HBF* “bins” were estimated to have a similar effect.

Though the relationship between *HBF* and species CPUE appeared to vary between the different flag-groups ([Figure 9](#)), a sensitivity analysis did not show a meaningful difference between indices estimated from models where the effects of *HBF* and Flag-group were additive or interactive. Additionally, due to some Flag-groups having too few observed *HBF* bins, some of the Flag-group \times *HBF* effects were poorly resolved which resulted in the interactive model failing to produce standard errors. As a result, the more parsimonious, additive model was used.

4.6 Error structure

Typically, continuous error structures (Lognormal or Gamma) have been used to model the positive component of the delta CPUE standardization model since the response variable is usually defined as

$$\frac{\text{Catch numbers}}{\text{Hooks fished}} \quad (1)$$

However, a more appropriate representation may be to use a discrete error structure (Poisson or Negative Binomial) with hooks fished used as an offset since this will more faithfully replicate the data generation process. Sensitivity analyses showed that the the Negative Binomial model was not computationally feasible and that the Poisson model failed to appropriately capture the dispersion in the data. These were not considered for further analysis.

There was also very little difference between estimated indices which modeled CPUE using either of the trialed continuous error distributions. The gamma distribution was selected for the spatiotemporal delta-GLMM “geostats” model based on an examination of the residual diagnostics.

4.7 Model structure

4.7.1 delta-GLM

The previous WCPFC stock assessments for bigeye tuna and yellowfin tuna (McKechnie et al., 2017a,b; Tremblay-Boyer et al., 2017b) utilized a standardized CPUE index calculated from a GLM fit to independent partitions of data from the operational longline data within each assessment region. This model structure was not altered as these indices are used in initial phases of the stock assessment in order to maintain continuity with the stepwise model progression from the 2017 diagnostic cases to the proposed 2020 diagnostic cases. Delta models consist of two sub-models: a logistic model (binomial component) predicting the probability of a positive catch occurring and a second model (positive component) predicting the magnitude of the positive catch rate (Lo et al., 1992; Stefansson, 1996). The two models are defined as follows:

Binomial component

$$y_i \sim \text{Bernoulli}(p_i) \quad (2)$$

$$\log \frac{p_i}{1 - p_i} = \beta_0 + \text{YrQtr}_i + 5^\circ \text{Cell}_i + \text{SpeciesCluster}_i + \beta_1(\text{HooksFished})_i \quad (3)$$

Note that the linear effect of HooksFished_i is defined in terms of 100s of hooks fished per set.

Positive component

$$\log c_i \sim \text{Normal}(\log \mu_i, \sigma^2) \quad (4)$$

$$\log \mu_i = \beta_0 + \text{YrQtr}_i + 5^\circ \text{Cell}_i + \text{SpeciesCluster}_i \quad (5)$$

where $\log c_i$ is the log CPUE of either bigeye tuna or yellowfin tuna defined as numbers caught per 100 hooks fished per set.

4.7.2 Spatiotemporal delta-GLMM

Spatiotemporal models have been shown to be more accurate and less biased than equivalently structured delta-GLMs when fit to fisheries dependent data (Grüss et al., 2019; Zhou et al., 2019). Additionally, explicitly modeling the spatiotemporal structure of the data allows these models to cope with non-stationary effort distributions like the ones exhibited in the operational longline dataset (Ducharme-Barth et al. 2019). The VAST spatiotemporal modeling approach (Thorson et al., 2015; Thorson, 2019) was used to generate relative abundance indices from the operational longline catch and effort data for the 2020 stock assessments of bigeye tuna and yellowfin tuna. In the VAST spatiotemporal modeling approach, the relative abundance index is the spatial average of predicted abundance once catchability effects have been “standardized” out. These indices are assigned to “index fisheries” in the stock assessments similar to what was done for the 2018 south Pacific albacore tuna stock assessment (Tremblay-Boyer et al., 2018). Additionally, the spatial abundance distributions predicted from the VAST model are used to calculate the regional weighting factors for the assessment regions.

The model implemented by the VAST package (version 3.3.0), a spatiotemporal delta generalized linear mixed model (GLMM), is an extension of the delta-GLM described in the previous section,

the key difference being how space is modeled. An interactive relationship between space and time, as opposed to an additive one, is specified using Gaussian random fields to define the spatial and spatiotemporal components of the model (Thorson et al., 2015). These Gaussian random fields are defined with a Matern covariance function. Using the estimated correlation structure of the data, spatiotemporal delta-GLMMs can simultaneously interpolate abundance of unobserved strata. The delta-GLMM structure implemented in R using the VAST package (Thorson et al., 2015; Thorson, 2019) is defined below:

Binomial component⁴

$$y_i \sim \text{Bernoulli}(p_i) \quad (6)$$

$$\log \frac{p_i}{1 - p_i} = \text{YrQtr}_i + \omega_1(s_i) + \phi_1(s_i, t_i) + s(\text{HBF}_i) + \text{Flag}_i + \epsilon_1 \quad (7)$$

Positive component

$$c_i \sim \text{Gamma}(\log \mu_i, \sigma^{-2}, \lambda \sigma^2) \quad (8)$$

$$\log \mu_i = \text{YrQtr}_i + \omega_2(s_i) + \phi_2(s_i, t_i) + s(\text{HBF}_i) + \text{Flag}_i + \epsilon_2 \quad (9)$$

where ω is the spatial random effect, ϕ is the spatiotemporal random effect, $s(\text{HBF})$ is the spline on *HBF* and *Flag* is the additive effect of Flag-group.

5 Results

5.1 Indices

Nominal CPUE of bigeye tuna and yellowfin tuna showed similar levels of decline across the assessment regions (Figure 17 & Figure 18). The exception being the non-equatorial regions for yellowfin tuna. The northern regions (1 & 2) showed moderately flat trends while the southern regions (5 & 6) both showed a precipitous decline in the early years of the model period. Trends in nominal CPUE was largely consistent between the full data set (10 020 513 observations) and the sub-sampled data set used for the spatiotemporal model (572 395 observations).

The standardized indices, both those estimated using the delta-GLM model described in Section 4.7.1 and the spatiotemporal model are consistent with both the nominal CPUE and the indices used in the diagnostic cases of the 2017 WCPFC stock assessments for bigeye tuna and yellowfin tuna. For bigeye tuna, the spatiotemporal approach estimates slightly higher relative abundance than the delta-GLM in the early period (pre-1960) for both of the northern regions (1 & 2), and slightly lower abundance in the early period of region 6. The spatiotemporal indices for yellowfin tuna also estimate slightly higher relative abundance than the delta-GLM in the initial years for regions 1, 2, & 4. The delta-GLM indices for yellowfin tuna estimate slightly higher initial abundance than the spatiotemporal indices. Both the spatiotemporal indices and the delta-GLM indices smooth out what appears to be an anomalous spike in nominal CPUE in region 1 during the late 1970s, and also the rapid initial decline in CPUE in regions 5 & 6. For both species the spatiotemporal model is able to interpolate the abundance trend for region 9, even when it was sparsely sampled, by taking advantage of the estimated correlation structure of the data and borrowing information heavily from the adjacent regions.

⁴The version of VAST used in this analysis (3.3.0) does not allow for covariates to be defined separately for each component of the delta model. Given that a continuous error distribution is used for the positive component and HooksFished is already included in the response variable, we were unable to use it as an offset in the binomial component of the model.

5.2 Catchability covariates

Inclusion of catchability covariates resulted in minor changes to the estimated index for bigeye tuna ([Figure 19](#)). Including in Flag-group resulted in a slight downwards revision of the standardized index relative to the nominal in the early years, and a slight upwards revision relative to the nominal in the later years. This is likely due to the change over time in fleet composition ([Figure 20](#)) from predominantly Japanese vessels, which were estimated to have the highest encounter probabilities, to vessels with lower encounter probabilities ([Figure 21](#)). Adding in the spline on *HBF* resulted in the opposite effect with the index being revised upwards in the early years, most notably in regions 1 & 2. This is because the predominant method of fishing in the early years used low numbers of *HBF*. These were estimated to have a low encounter probability and catch rate ([Figure 21](#)). This is an expected result given our understanding of *HBF* as a proxy for depth fished and the deeper swimming depth of bigeye tuna. However, as gear configurations changed over time to fishing with a greater number of *HBF* ([Figure 22](#)), which more efficiently caught bigeye tuna, the influence of this covariate also increased ([Figure 23](#)). When *HBF* is included in the standardization model, the removal of its effect results in an increase in the standardized abundance index relative to the nominal CPUE early in the model period.

The estimated index for yellowfin tuna did not meaningfully change with the inclusion of the catchability covariate for Flag-group and spline on *HBF* ([Figure 24](#)) despite the model showing trends in the influence of both covariates ([Figure 25](#)). The model predicted high probabilities of encounter for vessels from Australia, Japan, and Korea, and from sets with 20 - 30 *HBF* ([Figure 26](#)). Vessels from countries that operate predominantly in cooler waters such as New Zealand, or sets made with either very low *HBF* or very high *HBF* were predicted to have low encounter probability for yellowfin tuna. That being said, when vessels from New Zealand did encounter yellowfin tuna they tended to have the highest catch rate along with Australian vessels and those participating in the US Hawaiian shallow-set fishery. Catch-rates also tend to be highest for sets made with 25-30 *HBF*.

5.3 Estimated spatial patterns

Encounter rate patterns for bigeye tuna were stationary across time, though the intensity did diminish throughout the model period ([Figure 27](#)). Bigeye tuna were predicted to be most frequently encountered in a band just north of the equator extending northeast around the Hawaiian islands and eastwards into the eastern Pacific Ocean (EPO). Model predicted outputs for positive catch-rate were highest in the equatorial EPO early in the model period ([Figure 28](#)). In the WCPO, the area around Hawaii in assessment region 2 was predicted to have the highest positive catch rate.

Yellowfin tuna were predicted to have the highest encounter rate in the equatorial western Pacific, and in a broad, tapering band towards the EPO ([Figure 29](#)). Encounter rate diminished outside of the tropics and towards the poles, though perhaps not as rapidly as initially expected. These encounter rates were predicted to be highest at the beginning of the model period. Once encountered, yellowfin tuna were predicted to have the highest catch-rates in the equatorial western Pacific ([Figure 30](#)). These also were predicted to decline from the start of the model period.

The estimated regional weights (given as percentages) and defined as the average predicted abundance in each of the 9 assessment regions are listed as follows for bigeye tuna and yellowfin tuna:

- Bigeye tuna: 10.18, 31.13, 9.38, 23.67, 3.06, 8.05, 12.79, 1.58, 0.15
- Yellowfin tuna: 8.22, 5.38, 15.43, 19.72, 12.34, 10.99, 19.37, 7.55, 1.00

The spatiotemporal standardization models predicts a higher concentration of abundance in the eastern WCPO (model regions 2 and 4) for bigeye tuna and the equatorial WCPO (model regions 3, 4, and 7) for yellowfin tuna relative to the null hypothesis that regional abundance is proportional to area size ([Figure 31](#)).

5.4 Diagnostics

Based on feedback received from the 2020 PAW, DHARMA-style residuals were used to assess model fit to the data ([Hartig, 2020](#)). The VAST modeling framework is built upon TMB ([Kristensen et al., 2016](#)) which has the capability to simulate samples of the observations given the assumed model structure and the estimated fixed and random effects. DHARMA-style residuals take advantage of this approach by defining the residuals based on where the observations fall in the distribution of simulated samples ([Dunn and Smyth, 1996](#)). In addition to accounting for the random effects, another advantage of this approach is that it allows the entire model structure, both components of the delta model, to be evaluated using a quantile-quantile (QQ) plot of the uniformly distributed model residuals. Previously, QQ plots were generated separately for each model component.

For the bigeye tuna model, the model appears to fit the data well in the aggregate sense ([Figure 32](#)) without indication of non-uniformity. Visualizing the residuals spatially and temporally by model component, we see some evidence for model mis-specification. Though there do not appear to be persistent spatiotemporal patterns in the residuals for encounter rate ([Figure 33](#)), there do appear to be patterns in the residuals for positive catch-rate ([Figure 34](#)). Within the assessment region, the model appears to be underestimating catch rates in the north Pacific through the 1950s. This pattern persists in the EPO through the 1960s.

Similar to bigeye tuna, yellowfin tuna showed a good fit to the data in the aggregate sense ([Figure 35](#)). Examining the spatiotemporal distribution of model component residuals, the model persistently overestimated the encounter rate in assessment region 7 during recent years, as well as in assessment region 2 early in the model period ([Figure 36](#)). The model also appeared to underestimate the positive catch-rate in the equatorial western pacific and EPO through the early 1960s ([Figure 37](#)).

For both species, residual patterns were largely absent as the model approached the present day, indicating a better fit to the data in recent years. This lack of fit earlier in the model period is captured in the estimated model uncertainty which is highest when the model begins, and gradually diminishes over time ([Figure 17 & Figure 18](#)).

6 Discussion

The approach taken for the 2020 stock assessments predicts relative abundance indices for bigeye tuna and yellowfin tuna that show consistent trends with those from previous analyses, and also similar trends to the nominal CPUE. Given that average fishing efficiency is believed to increase between 2-4% per year ([Palomares and Pauly, 2019](#)), a greater difference between the nominal and standardized indices would be expected over the 67 year model period. This likely indicates that the standardization model is missing a key element. This analysis highlights a number of areas where improvements can be made.

Longline fishing in the Pacific has changed dramatically from its pre-World War II industrial beginnings. Vessels have grown larger and more powerful, with onboard electronic systems for measuring sea surface temperature, current speed, and thermocline depth. Targeting practices have also

changed from a focus on albacore for cannery-processing to targeting the tropical tunas, bigeye tuna and yellowfin tuna for fresh or flash frozen sashimi. Perhaps most important are the gear changes such as the switch from rope to monofilament mainlines. One of the challenges with attempting to model CPUE across such a long period of time, and using a data set from 27 different fishing nations is that it greatly reduces the minimum set of covariates usable for standardization. The residual patterns, particularly those early in the model period, indicate that the model is too simple and is unable to account for changes in catch rate. Furthermore, the utility of the *HBF* covariate as a proxy for fishing depth is reduced without having information on the other characteristics of the longline set such as mainline material, branch line construction, or line setting speed as all effect the depth of the fished hooks. If there is a continued desire for a unified analysis of the operational longline dataset for the entire model time period, it is recommended that further data rescue exercises be undertaken to augment the number of usable catchability covariates.

Given the described covariate deficiencies in the operational longline data-set, another option would be to estimate an index from data collected as a part of the Regional Observer Program. This would allow additional covariates shown to influence catch rates such as time of day, bait type, light stick use, hook type, and hook spacing to be included in the standardization model (Campbell, 2018; Monnahan and Stewart, 2018). It is recommended that interviews and surveys with boat captains and vessel operators be conducted to identify species-specific targeting practices and an appropriate covariate set to capture these targeting behaviors. Using the observer data presents a trade-off between reduced spatial, temporal coverage and a potential increase in explanatory power as more covariates can be included. A compromise could be reached by splitting the index into an early period fit to the operational data, and overlapped with a second index fit to recent period observer data or a combination of observer and operational data with an expanded covariate set. If a split analysis is to be considered this must be done carefully to ensure a proper transition between the two indices. Any hard temporal discontinuities in the indices can cause problems for the assessment model. It is recommended that future simulation work be undertaken to determine the benefits and limitations associated with splitting the index, both in the ability to capture the underlying abundance trend and for assessing the implications of using split indices in the stock assessment model.

Though the VAST framework presents a unique modeling approach to the inclusion of spatiotemporally varying environmental covariates, it is limited in its ability to account for variability in catch rates due to the vertical distribution of tuna within the water column. Particularly for bigeye tuna, diving behavior as it relates to environmental preferences is believed to have a major impact on availability to longline gear. Early work by Bigelow et al. (2002) accounting for the vertical distribution of both tuna and fishing effort (hooks) to estimate catch-per-unit-of-effective-effort (CPEE), where effective effort is the numbers of longline hooks fished at different depth bins weighted by the perceived vertical distribution of tuna showed larger deviations from the nominal index than those seen in the current analysis. It is recommended that this approach be revisited. However, in order to improve on this existing work, 3 data gaps must be addressed: 1) funding is needed to produce a spatiotemporal sub-surface dissolved oxygen oceanographic product for the Pacific Ocean spanning the assessment model period; 2) the analysis of archival tagged, longline-vulnerable bigeye and yellowfin tuna needs to be continued and supplemented by additional releases in multiple locations across the WCPO. Existing studies point to the potential for daytime and night-time vertical distributions that vary by size and potentially regional differences in oceanography or even FAD-density within the WCPO (Scutt Phillips et al., 2017; Abascal et al., 2018; Scutt Phillips et al., 2019); 3) basic research on the fishing depth of longline hooks such as that described in Boggs (1992) or Bigelow et al. (2006) needs to be conducted across vessels from different DWFNs and PICTs in

order to understand spatial patterns in the depth distribution of hooks fished.

Finally, persistent temporal patterns exist for both the Flag-group and *HBF* covariates so it is possible that the estimated categorical effects and covariate influence are confounded with the trend in abundance despite the moderate level of overlap between categories. For both bigeye tuna and yellowfin tuna there is a matching decline in the influence of the Flag-group covariate on the index over time. This matches with a transition of records coming predominantly from Japanese vessels, which are estimated to have higher catch rates, to Korean, US, and PICT vessels. This has the effect of standardizing the index “down” in the early period and standardizing the abundance index “up” in later years as shown for bigeye tuna. Japanese vessels, however, may be estimated to have a higher effect given that the majority of their records come the earlier periods when the stock was less depleted. Future research is needed to understand the modeling limitations when both the underlying abundance and characteristics of the fishery show persistent temporal patterns.

7 Acknowledgments

We are very appreciative of those members of the ePAW who took the time to review and comment on a proposal of the CPUE modeling approach prior to the 2020 meeting. This early feedback helped structure many of the decisions made in this analysis. Thank you also to members of the SPC-OFP staff Tiffany Vidal, Joe Scutt Phillips, and Paul Hamer for providing feedback at various stages of the analysis.

References

- Abascal, F., Peatman, T., Leroy, B., Nicol, S., Schaefer, K., Fuller, D., and Hampton, J. (2018). Spatiotemporal variability in bigeye vertical distribution in the Pacific Ocean. *Fisheries Research*, 204:371–379.
- Arrizabalaga, H., Dufour, F., Kell, L., Merino, G., Ibaibarriaga, L., Chust, G., Irigoien, X., Santiago, J., Murua, H., Fraile, I., Chifflet, M., Goikoetxea, N., Sagarminaga, Y., Aumont, O., Bopp, L., Herrera, M., Marc Fromentin, J., and Bonhommeau, S. (2015). Global habitat preferences of commercially valuable tuna. *Deep Sea Research Part II: Topical Studies in Oceanography*, 113:102–112.
- Balmaseda, M. A., Mogensen, K., and Weaver, A. T. (2013). Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quarterly Journal of the Royal Meteorological Society*, 139(674):1132–1161.
- Bigelow, K., Musyl, M. K., Poisson, F., and Kleiber, P. (2006). Pelagic longline gear depth and shoaling. *Fisheries Research*, 77(2):173–183.
- Bigelow, K. A., Hampton, J., and Miyabe, N. (2002). Application of a habitat-based model to estimate effective longline fishing effort and relative abundance of Pacific bigeye tuna (*Thunnus obesus*). *Fisheries Oceanography*, 11(3):143–155.
- Boggs, C. (1992). Depth, capture time, and hooked longevity of longline-caught pelagic fish: Timing bites of fish with chips. *Fishery Bulletin*, 90:642–658.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

- Campbell, R. (2018). Aggregate and size-based standardised CPUE indices for longline target species caught within the ETBF – 2018 Update. Working paper presented to the 21st meeting of the Tropical Tuna RAG, Mooloolaba, 16-17 July 2018.
- Ducharme-Barth, N., Vincent, M., Pilling, G., and Hampton, J. (2019). Simulation analysis of pole and line CPUE standardization approaches for skipjack tuna in the WCPO. Technical Report WCPFC-SC15-2019/SA-WP-04, Pohnpei, Federated States of Micronesia 12-20 August 2019.
- Dunn, P. K. and Smyth, G. K. (1996). Randomized Quantile Residuals. *Journal of Computational and Graphical Statistics*, 5(3):236–244.
- Good, S. A., Martin, M. J., and Rayner, N. A. (2013). EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates: THE EN4 DATA SET. *Journal of Geophysical Research: Oceans*, 118(12):6704–6716.
- Gouretski, V. and Reseghetti, F. (2010). On depth and temperature biases in bathythermograph data: Development of a new correction scheme based on analysis of a global ocean database. *Deep Sea Research Part I: Oceanographic Research Papers*, 57(6):812–833.
- Grüss, A., Walter, J. F., Babcock, E. A., Forrestal, F. C., Thorson, J. T., Lauretta, M. V., and Schirripa, M. J. (2019). Evaluation of the impacts of different treatments of spatio-temporal variation in catch-per-unit-effort standardization models. *Fisheries Research*, 213:75–93.
- Hartig, F. (2020). DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models.
- Kinoshita, J., Aoki, Y., Ducharme-Barth, N., and Kiyofuji, H. (2019). Standardized catch per unit effort (CPUE) of skipjack tuna of the Japanese pole-and-line fisheries in the WCPO from 1972 to 2018. Technical Report WCPFC-SC15-2019/SA-WP-14, Pohnpei, Federated States of Micronesia.
- Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., and Bell, B. M. (2016). **TMB** : Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software*, 70(5).
- Lo, N. C. H., Jacobson, L. D., and Squire, J. L. (1992). Indexes of relative abundance from fish spotter data based on delta-lognormal models. *Canadian Journal of Fisheries and Aquatic Sciences*, 49(12):2515–2526.
- Longhurst, A. (1999). *Ecological geography of the sea*. Academic Press, San Diego. OCLC: 1085815782.
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3):276–282.
- McKechnie, S., Hampton, J., Abascal, F., Davies, N., and Harley, S. J. (2015a). Sensitivity of WCPO stock assessment results to the inclusion of EPO dynamics within a Pacific-wide analysis. Technical Report WCPFC-SC11-2015/SA-WP-04, Pohnpei, Federated States of Micronesia, 5–13 August 2015.
- McKechnie, S., Pilling, G., and Hampton, J. (2017a). Stock assessment of bigeye tuna in the western and central Pacific Ocean. Technical Report WCPFC-SC13-2017/SA-WP-05, Rarotonga, Cook Islands, 9–17 August 2017.

- McKechnie, S., Tremblay-Boyer, L., and Harley, S. J. (2015b). Analysis of Pacific-wide operational longline CPUE data for bigeye tuna. Technical Report WCPFC-SC11-2015/SA-WP-03, Pohnpei, Federated States of Micronesia, 5–13 August 2015.
- McKechnie, S., Tremblay-Boyer, L., and Pilling, G. (2017b). Background analyses for the 2017 stock assessments of bigeye and yellowfin tuna in the western and central Pacific Ocean. Technical Report WCPFC-SC13-2017/SA-IP-06, Rarotonga, Cook Islands.
- Monnahan, C. C. and Stewart, I. J. (2018). The effect of hook spacing on longline catch rates: Implications for catch rate standardization. *Fisheries Research*, 198:150–158.
- Palomares, M. L. D. and Pauly, D. (2019). On the creeping increase of vessels’ fishing power. *Ecology and Society*, 24(3):art31.
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Scutt Phillips, J., Escalle, L., Pilling, G., Sen Gupta, A., and van Sebille, E. (2019). Regional connectivity and spatial densities of drifting fish aggregating devices, simulated from fishing events in the Western and Central Pacific Ocean. *Environmental Research Communications*, 1(5):055001.
- Scutt Phillips, J., Pilling, G. M., Leroy, B., Evans, K., Usu, T., Lam, C. H., Schaefer, K. M., and Nicol, S. (2017). Revisiting the vulnerability of juvenile bigeye (*Thunnus obesus*) and yellowfin (*T. albacares*) tuna caught by purse-seine fisheries while associating with surface waters and floating objects. *PLOS ONE*, 12(6):e0179045.
- Siders, Z., Ducharme-Barth, N., Carvalho, F., Kobayashi, D., Martin, S., Raynor, J., Jones, T., and Ahrens, R. (2020). Ensemble Random Forests as a tool for modeling rare occurrences. *Endangered Species Research*, In Press.
- Song, L., Zhou, J., Zhou, Y., Nishida, T., Jiang, W., and Wang, J. (2009). Environmental preferences of bigeye tuna, *Thunnus obesus*, in the Indian Ocean: an application to a longline fishery. *Environmental Biology of Fishes*, 85(2):153–171.
- Song, L. M., Zhang, Y., Xu, L. X., Jiang, W. X., and Wang, J. Q. (2008). Environmental preferences of longlining for yellowfin tuna (*Thunnus albacares*) in the tropical high seas of the Indian Ocean. *Fisheries Oceanography*, 17(4):239–253.
- Stefansson, G. (1996). Analysis of groundfish survey abundance data: Combining the GLM and delta approaches. *Ices Journal of Marine Science*, 53(3):577–588.
- Thorson, J. T. (2019). Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. *Fisheries Research*, 210:143–161.
- Thorson, J. T., Shelton, A. O., Ward, E. J., and Skaug, H. J. (2015). Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. *Ices Journal of Marine Science*, 72(5):1297–1310.
- Tremblay-Boyer, L., Hampton, J., McKechnie, S., and Pilling, G. (2018). Stock assessment of South Pacific albacore tuna. Technical Report WCPFC-SC-14-2018/SA-WP-05, Busan, South Korea, 8–16 August 2018.

- Tremblay-Boyer, L., McKechnie, S., Pilling, G., and Hampton, J. (2017a). Exploratory geostatistical analyses of Pacific-wide operational longline CPUE data for WCPO tuna assessments. Technical Report WCPFC-SC13-2017/SA-WP-03, Rarotonga, Cook Islands, 9-17 August 2017.
- Tremblay-Boyer, L., McKechnie, S., Pilling, G., and Hampton, J. (2017b). Stock assessment of yellowfin tuna in the Western and Central Pacific Ocean. Technical Report WCPFC-SC13-2017/SA-WP-06, Rarotonga, Cook Islands, 9-17 August 2017.
- Tremblay-Boyer, L. and Pilling, G. (2017). Use of operational vessel proxies to account for vessels with missing identifiers in the development of standardised CPUE time series. Technical Report WCPFC-2017-SC13/SA-WP-04, Rarotonga, Cook Islands.
- Wood, S. N. (2006). *Generalized Additive Models: An Introduction with R.* Chapman and Hall/CRC.
- Zhou, S., Campbell, R. A., and Hoyle, S. D. (2019). Catch per unit effort standardization using spatio-temporal models for Australia's Eastern Tuna and Billfish Fishery. *ICES Journal of Marine Science*.

8 Tables

Table 1: Representation of flag-groups in the final data-set used for the CPUE standardization model.

Flag.group ⁵	Flag.fleet	N	Percent
JP.DW	JP.DW	215,150	37.59
JP.OS	JP.OS	94,845	16.57
KR	KR	71,651	12.52
TW.DW	TW.DW	54,475	9.52
PICT.SP	CK, FJ.FJ, NC, NU, PF.PF, TO, TV, US.AS, VU, WS	43,524	7.60
TW.OS	TW.OS	21,084	3.68
CN	CN, CN.CN, CN.DW	20,581	3.60
US.HW.D	US.HW (Hawaii deep set, >15 HBF)	12,342	2.16
PICT.EQ	ID, KI, PG, PG.DW, SB, SB.DW	8,783	1.53
US.HW.S	US.HW (Hawaii shallow set, <15 HBF)	7,368	1.29
AU	AU.AU, AU.CV	7,172	1.25
PICT.NP	FM, GU, MH, PW	5,786	1.01
FJ.charter	FJ.AU, FJ.CK, FJ.CN, FJ.KR, FJ.NZ, FJ.TW, FJ.US	5,756	1.01
NZ	NZ.JP, NZ.NZ	3,878	0.68

⁵ISO 3166 alpha-2 country codes are used to refer to the flags of individual nations.

9 Figures

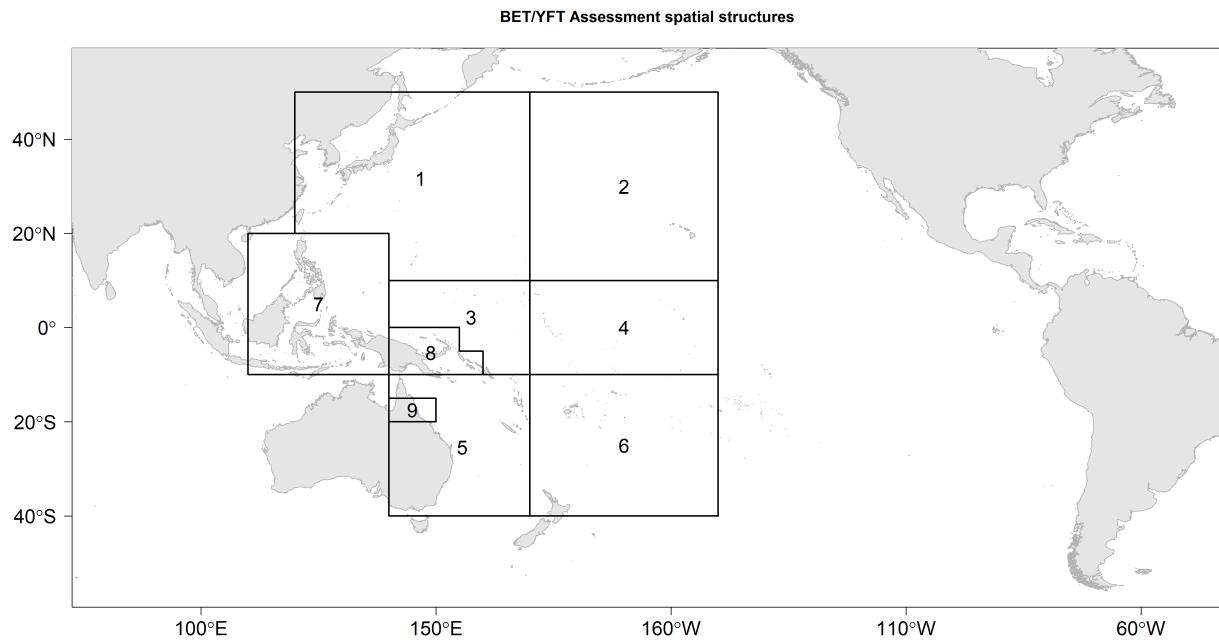


Figure 1: Spatial structure for the bigeye tuna and yellowfin tuna stock assessments.

Operational longline sets per region

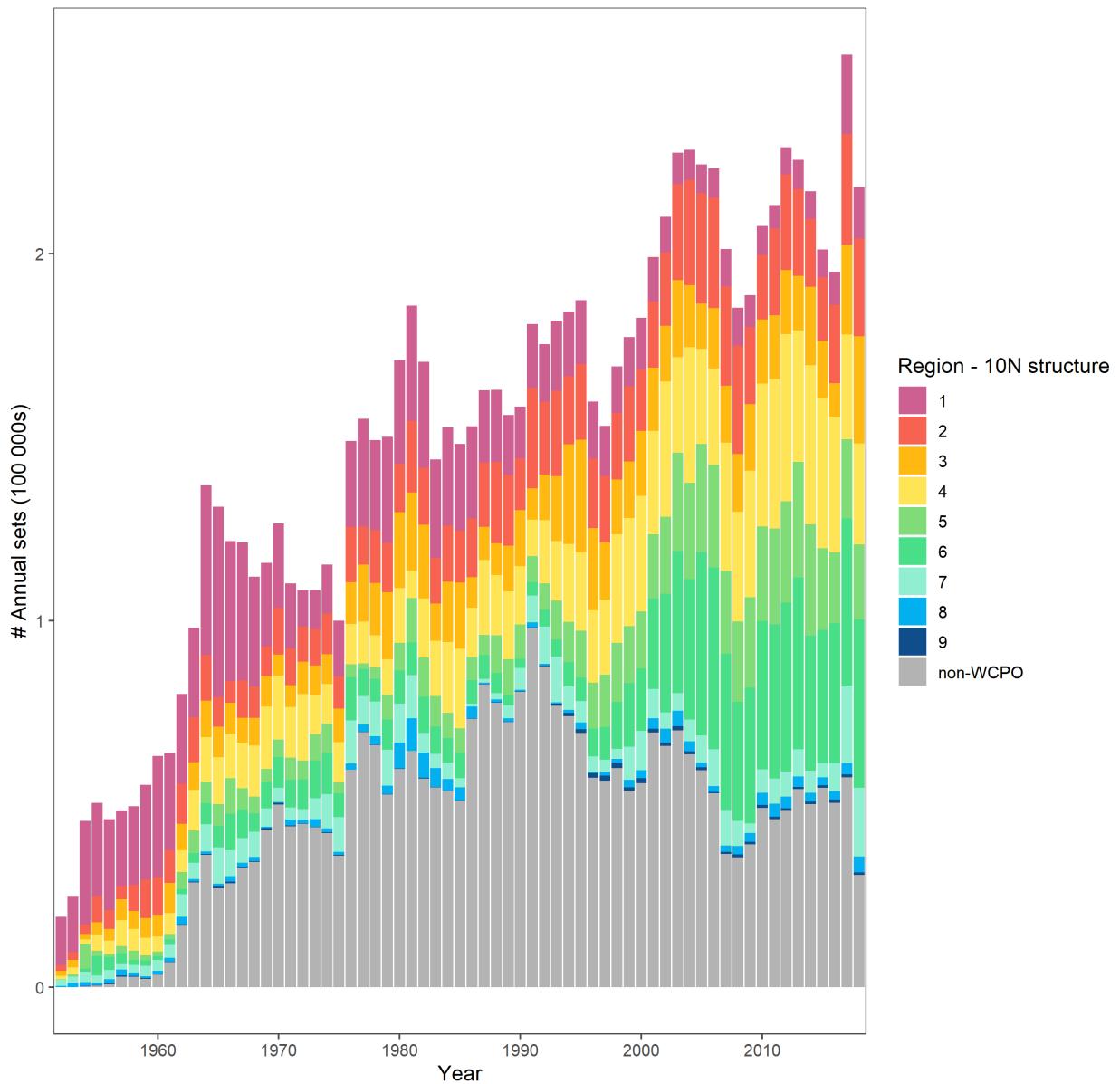


Figure 2: The number of operational longline records within each assessment region over time.

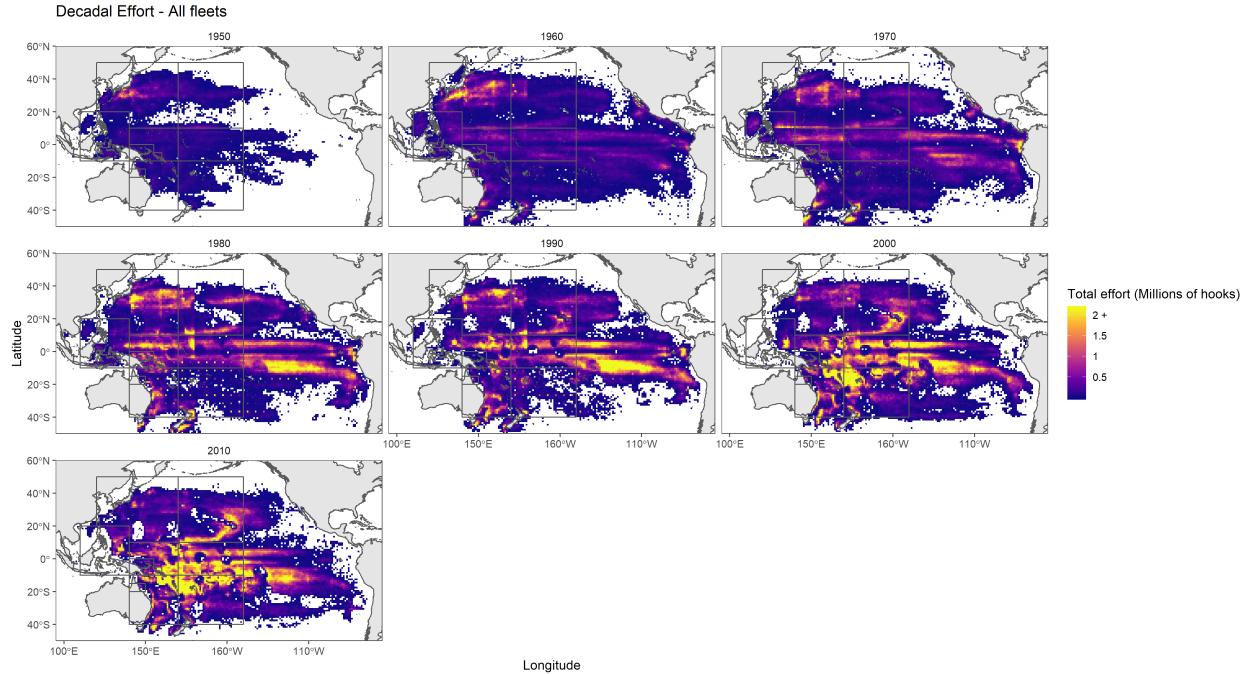


Figure 3: The decadal distribution of longline fishing effort (defined as hooks fished) across all fishing fleets in the operational longline data set.

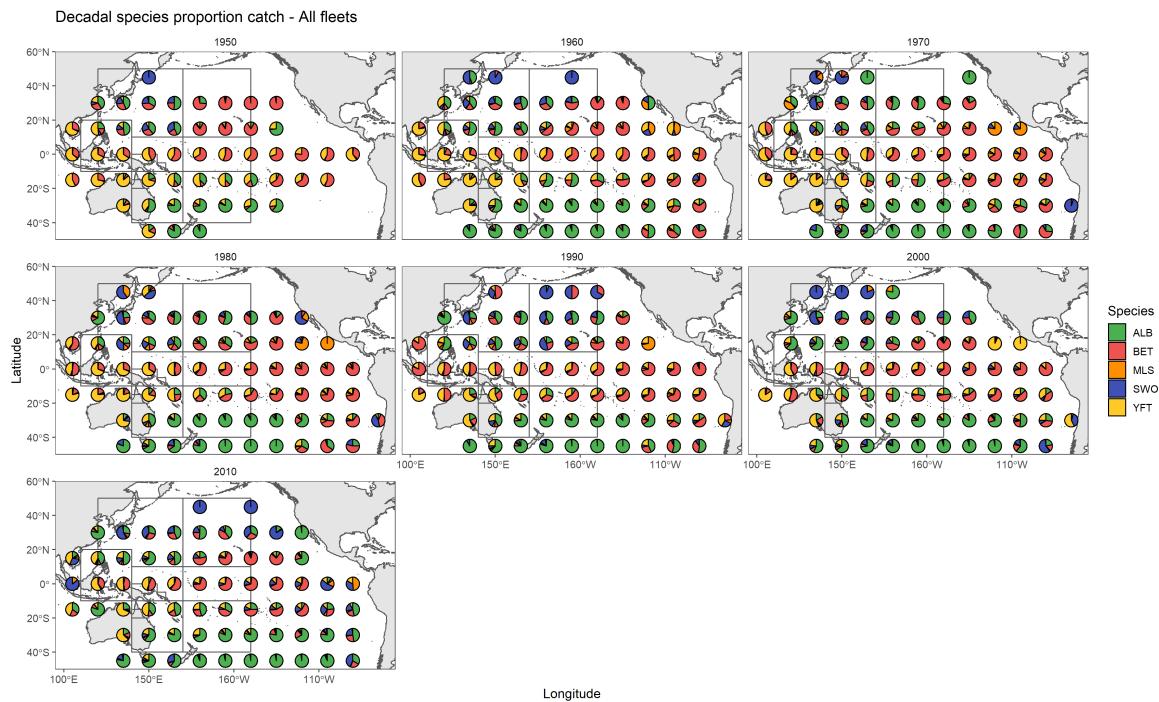


Figure 4: The decadal distribution of proportion of species caught across all fishing fleets by $15^\circ \times 15^\circ$ spatial cell in the operational longline data set.

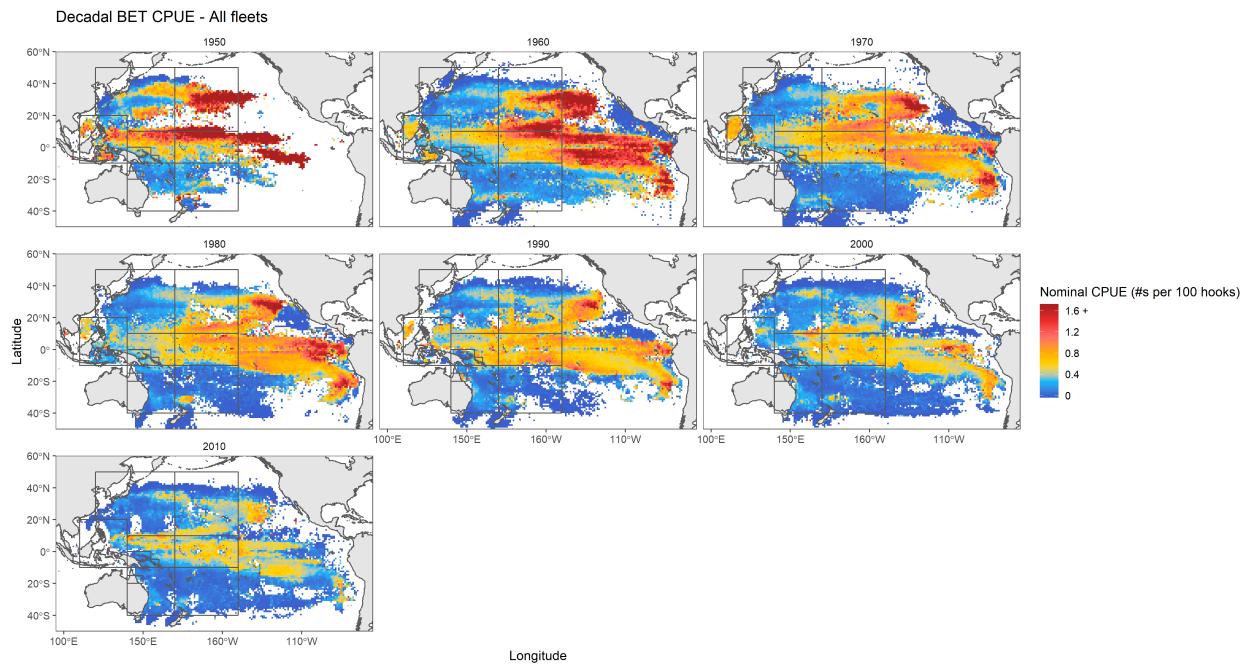


Figure 5: The decadal distribution of bigeye tuna nominal CPUE (numbers per 100 hooks fished) across all fishing fleets in the operational longline data set.

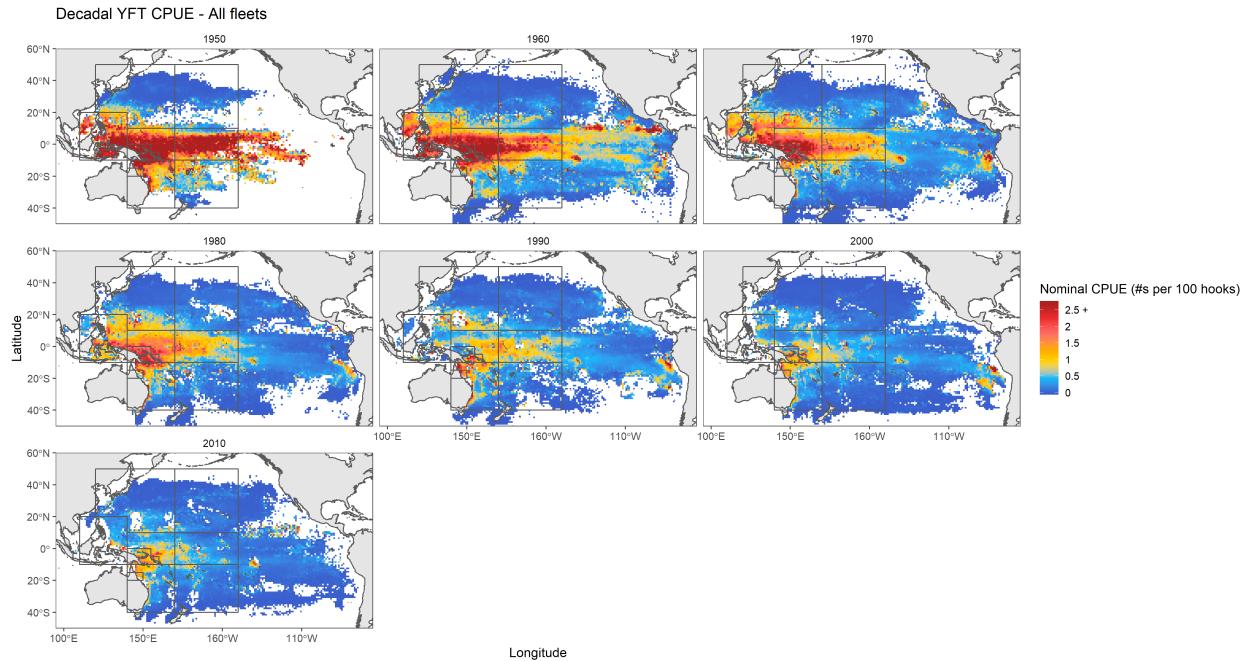


Figure 6: The decadal distribution of yellowfin tuna nominal CPUE (numbers per 100 hooks fished) across all fishing fleets in the operational longline data set.

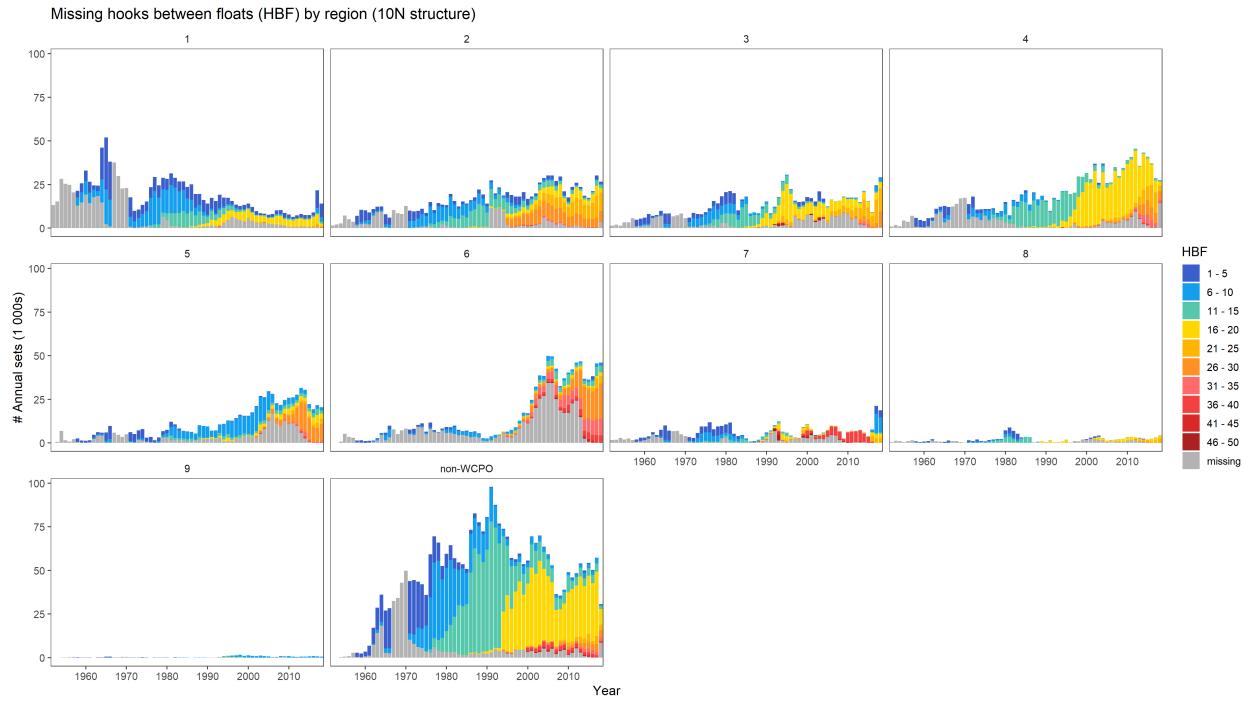


Figure 7: The number of records with and without HBF by assessment region. The color of the bar indicates the “bin” of HBF that each record was assigned to.

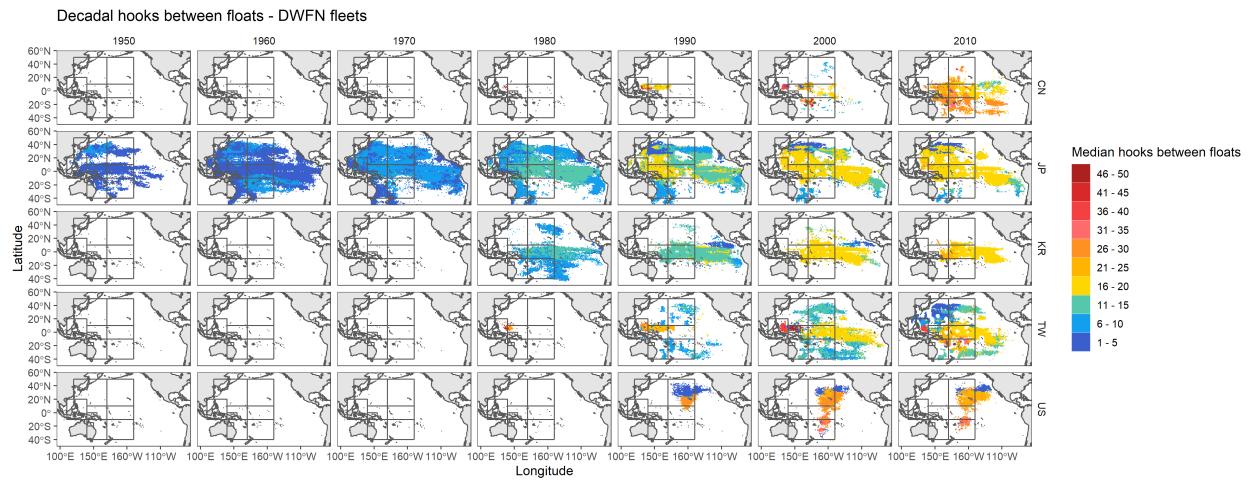


Figure 8: The decadal (columns) distribution of HBF by major DWFN fishing fleet (rows). The color of the bar indicates the median value “bin” of HBF

CPUE by hooks between floats (HBF) - Flag groups

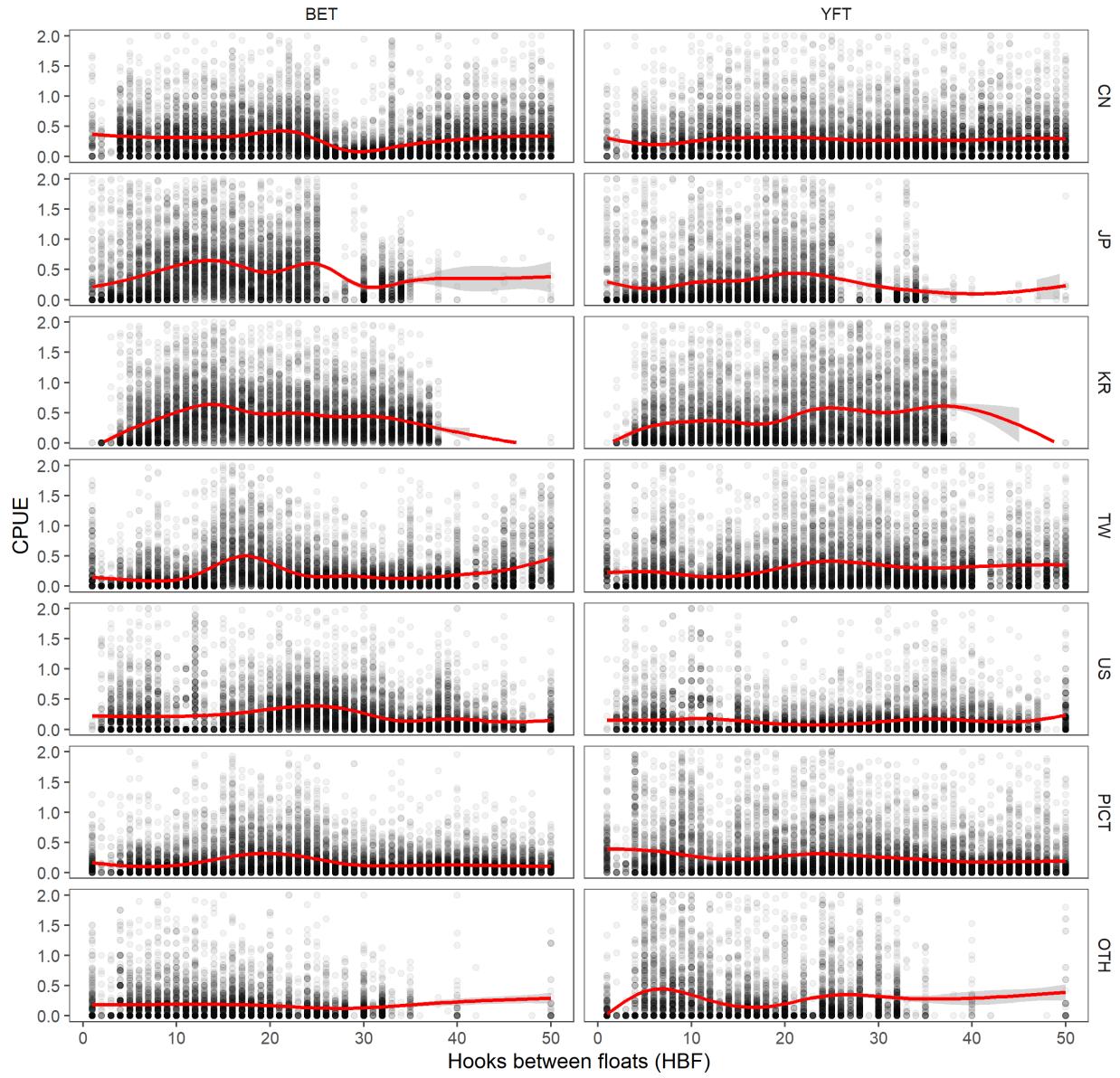


Figure 9: The response of nominal CPUE (numbers per 100 hooks fished) by species to hooks-between-floats (HBF) by flag group. The red line is a generalized additive model (GAM; Wood 2006) smooth fit through the data.

Mean CPUE by vessel - Flag groups

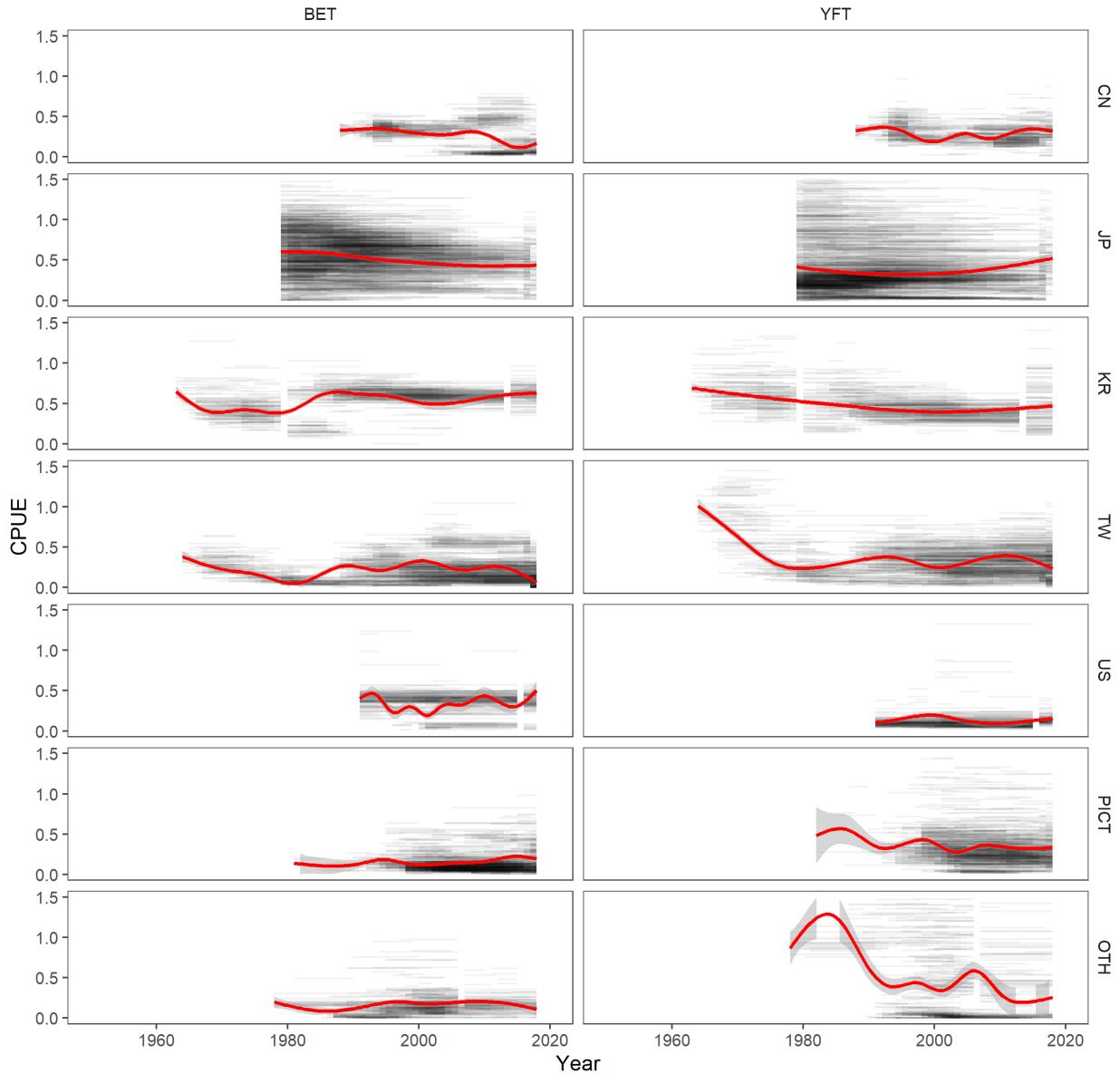


Figure 10: The average nominal CPUE (numbers per 100 hooks fished) by vessel ID for each species by flag group. The horizontal lines indicate the duration of activity for each individual vessel. The red line is a generalized additive model (GAM; Wood 2006) smooth fit through the data.

Catch by hooks fished - Flag groups

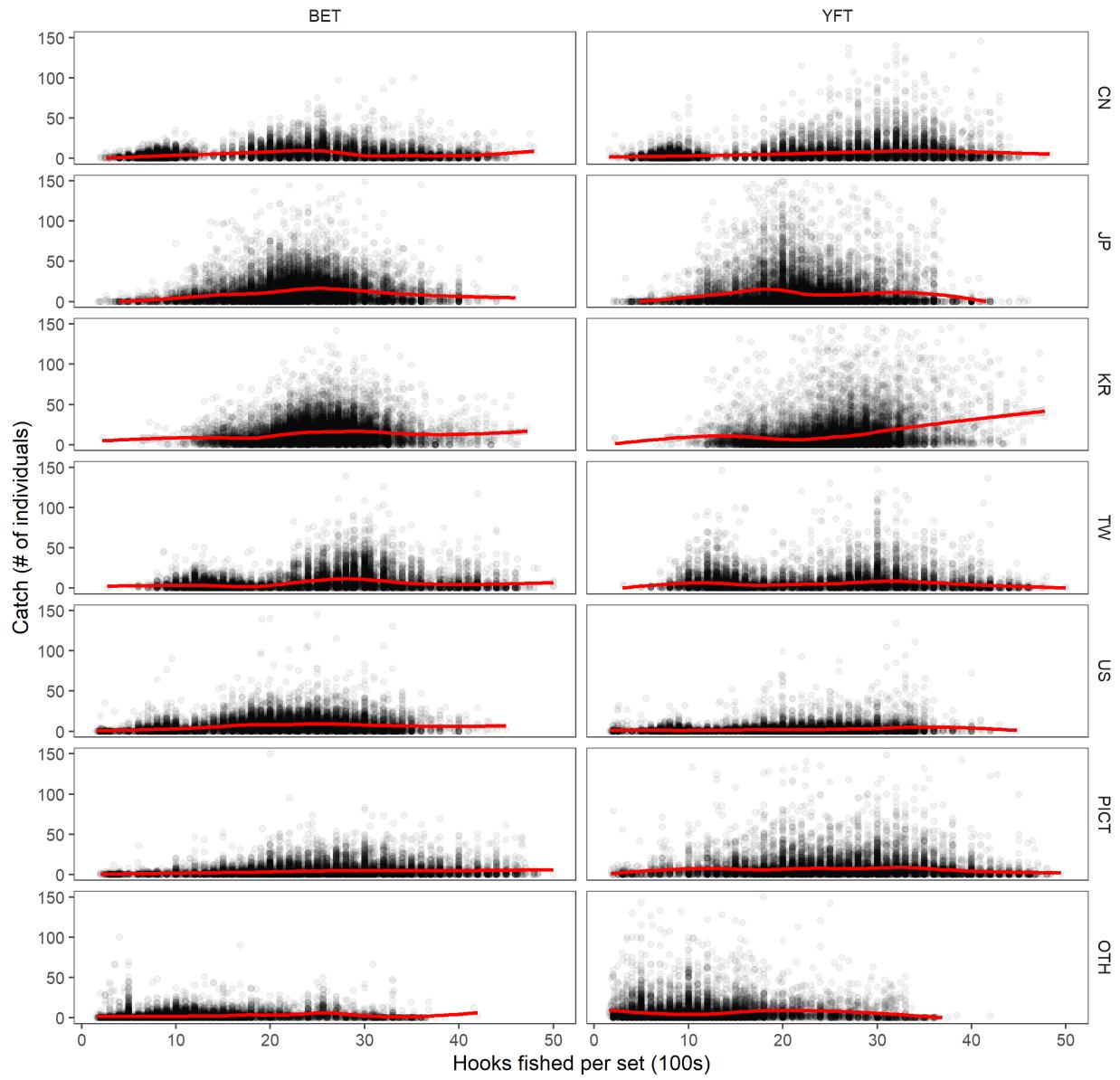


Figure 11: The response of catch (numbers) by species to 100 hooks fished per set by flag group. The red line is a generalized additive model (GAM; [Wood 2006](#)) smooth fit through the data.

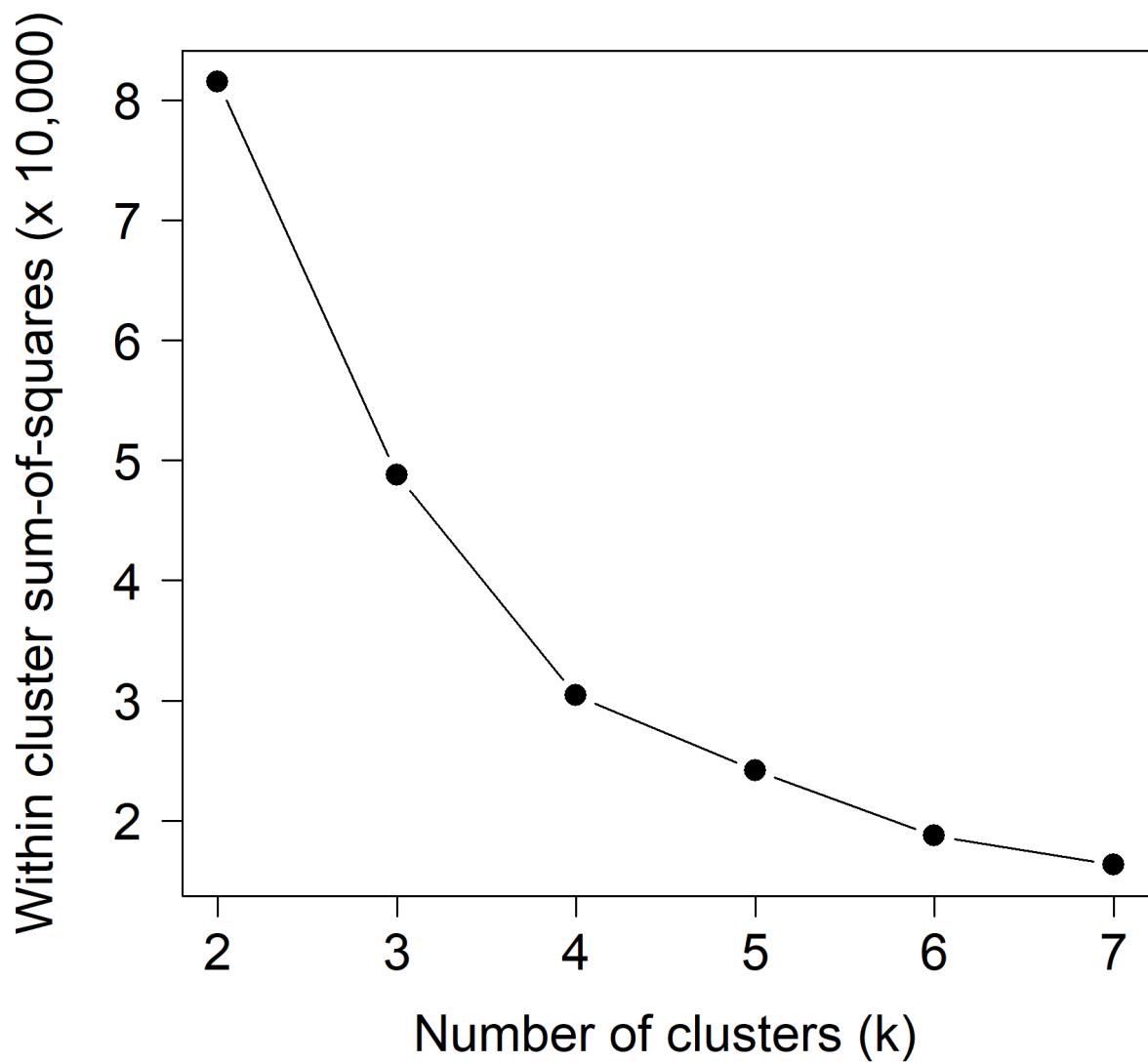


Figure 12: “Elbow”-plot from k-means clustering of “trips” based on the proportion of each species caught.

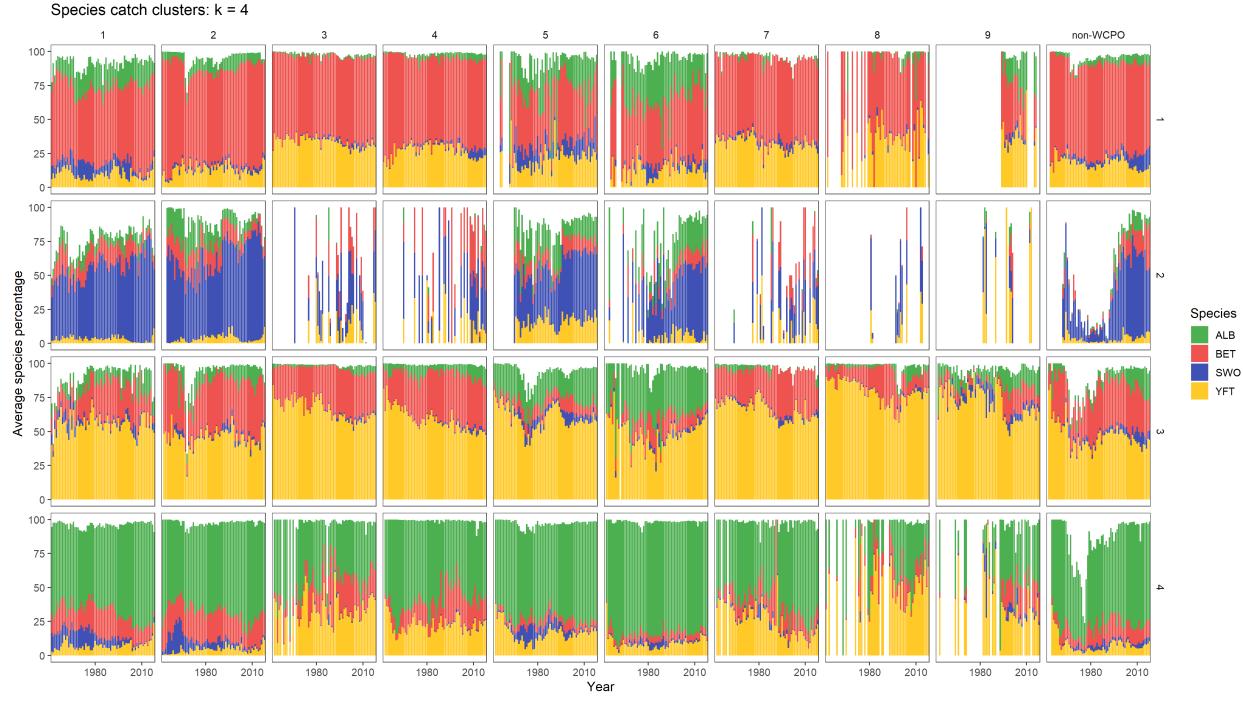


Figure 13: The average species composition by cluster and assessment region over time from the k -means clustering analysis.

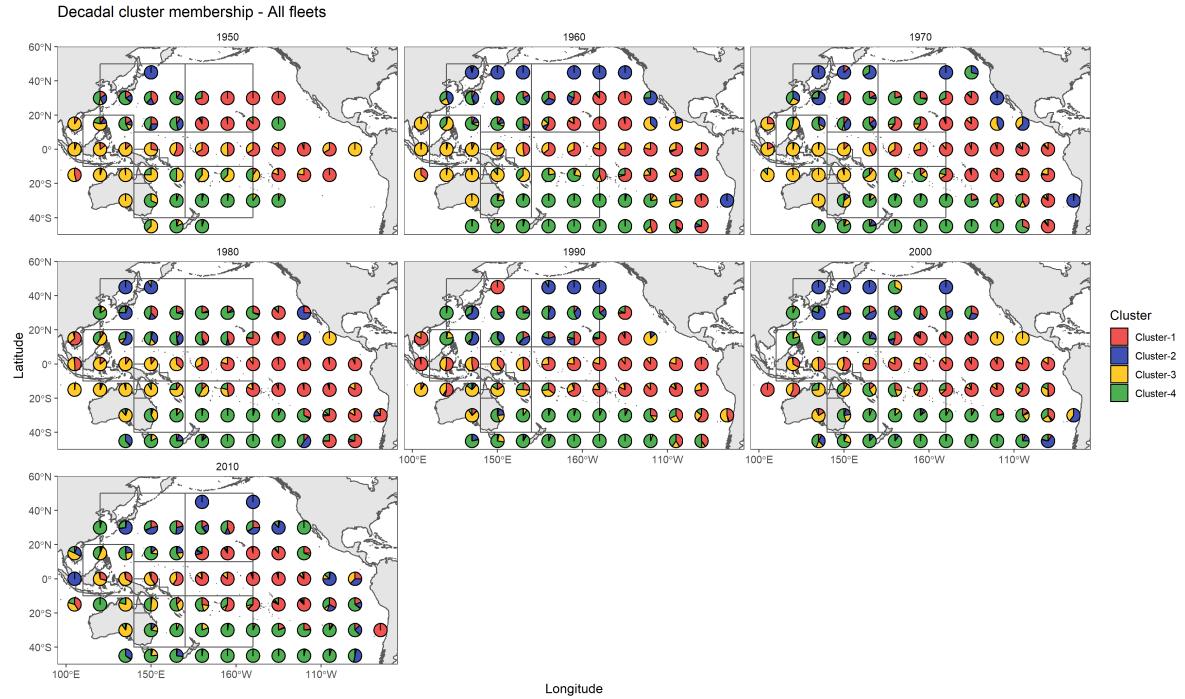


Figure 14: The decadal distribution of proportion of sets belonging to each species cluster across all fishing fleets by $15^\circ \times 15^\circ$ spatial cell in the operational longline data set.

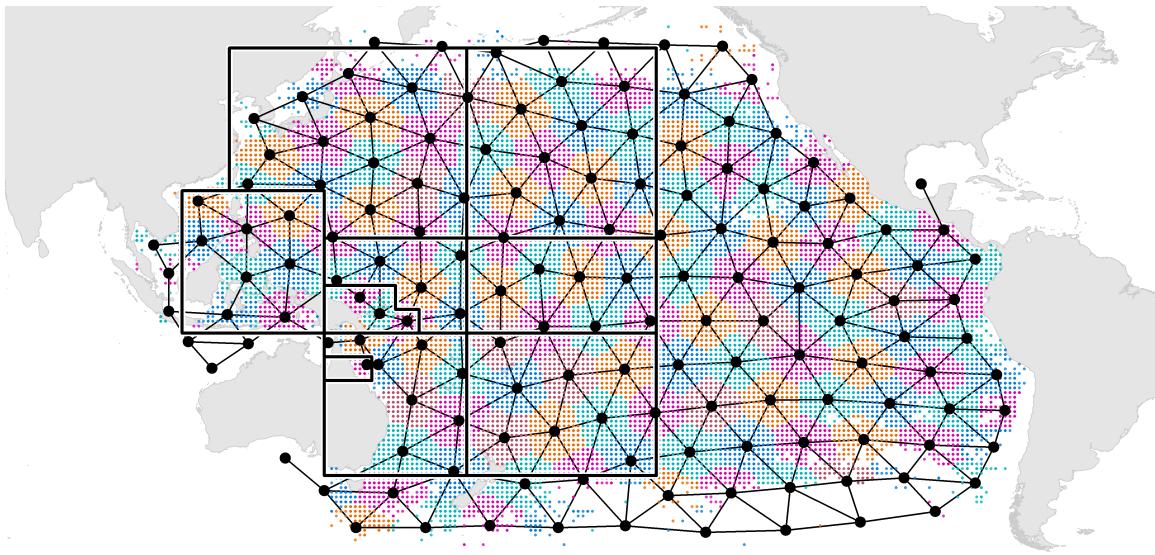


Figure 15: The distribution of spatial knots used to define the mesh for the spatiotemporal standardization model. Extrapolation grid cells are color coded to show the knot that they are associated with.

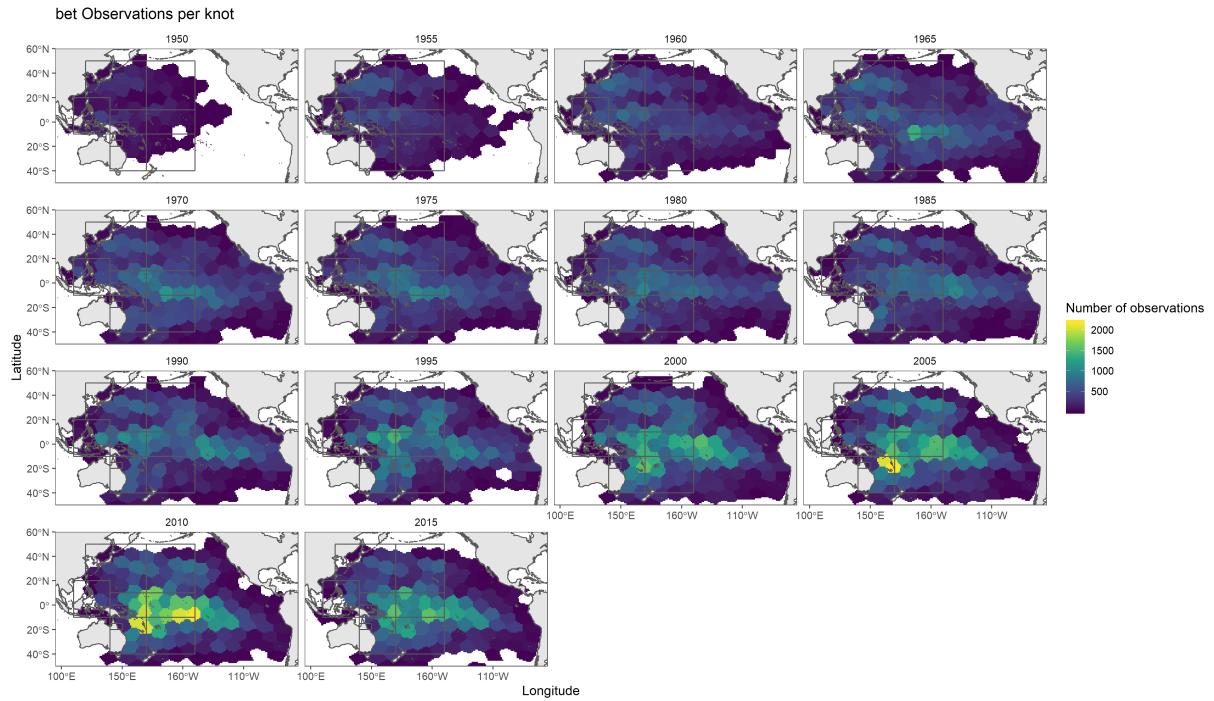


Figure 16: The temporal distribution of observations by spatial knot from the final sub-sampled data set used to estimate the indices.

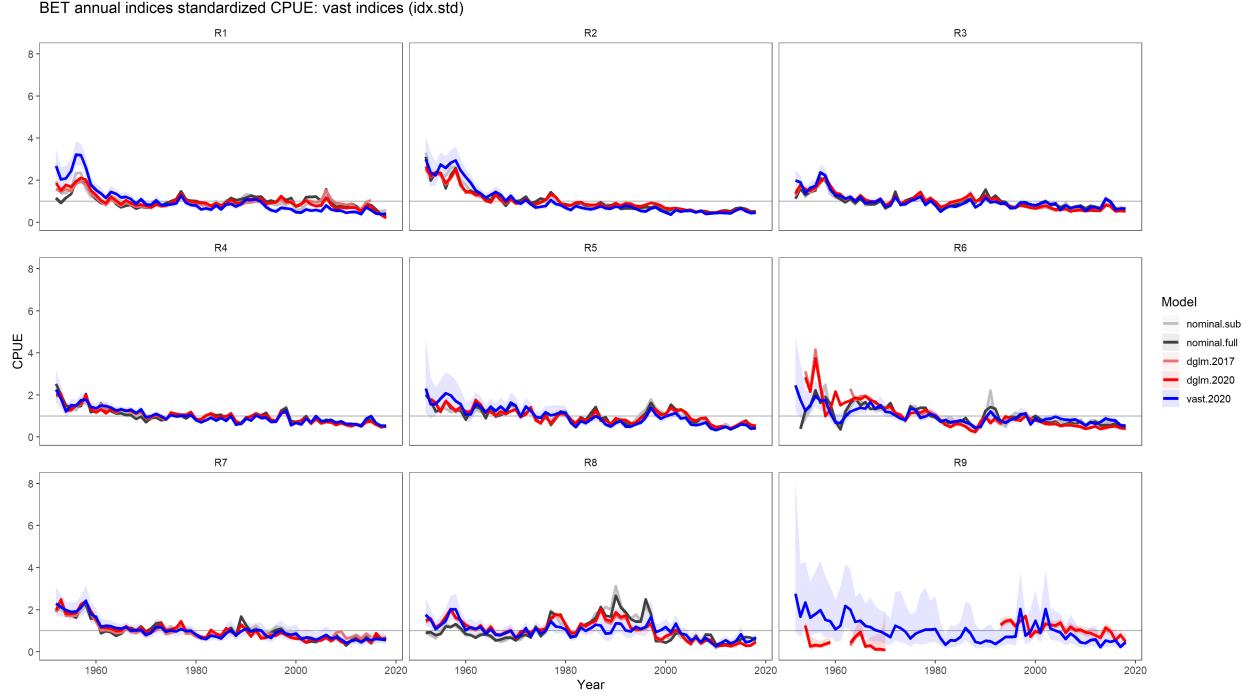


Figure 17: Nominal and standardized indices for bigeye tuna. The nominal index corresponding to the subset of data that the standardization model was fit to is shown in light gray (*nominal.sub*). The nominal index from the full data set is shown in dark gray (*nominal.full*). The delta-GLM index used in the 2017 stock assessment is shown in light red (*dglm.2017*). The delta-GLM index used in the data update step of the 2020 stock assessment is shown in red (*dglm.2020*). The VAST spatiotemporal index used in the diagnostic case of the 2020 stock assessment is shown in blue. (*vast.2020*). The asymptotic 95% confidence intervals are shown via the corresponding shaded polygon.

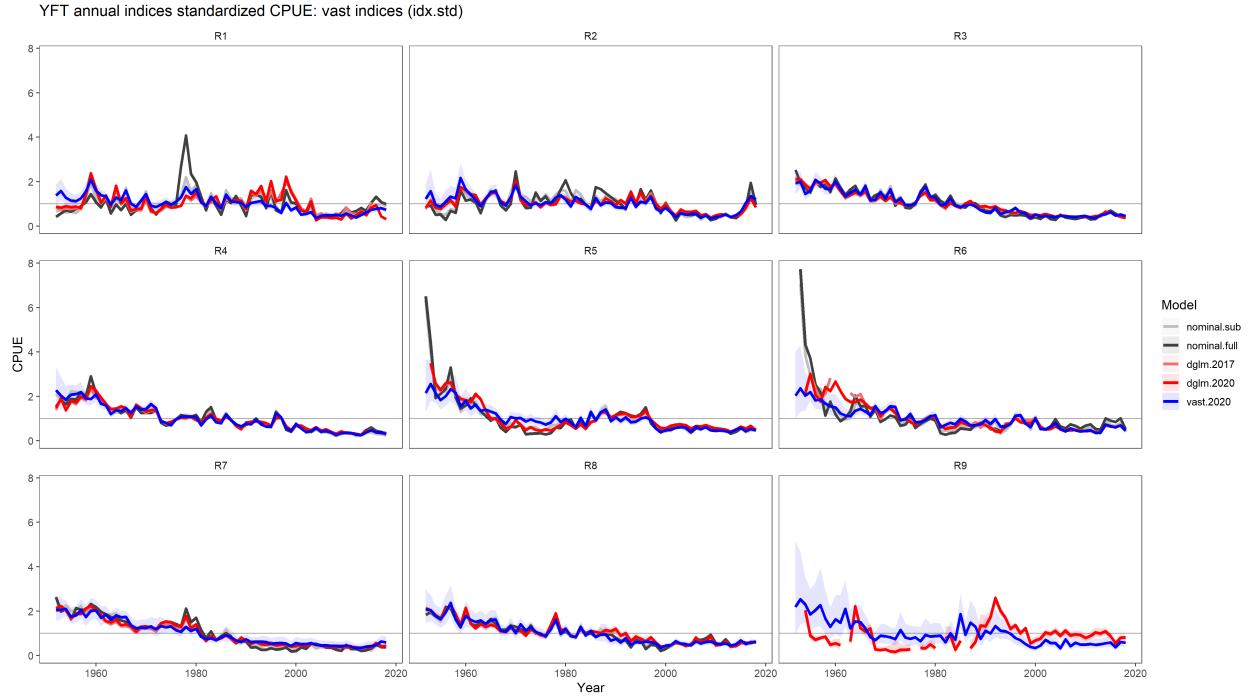


Figure 18: Nominal and standardized indices for yellowfin tuna. The nominal index corresponding to the subset of data that the standardization model was fit to is shown in light gray (*nominal.sub*). The nominal index from the full data set is shown in dark gray (*nominal.full*). The delta-GLM index used in the 2017 stock assessment is shown in light red (*dglm.2017*). The delta-GLM index used in the data update step of the 2020 stock assessment is shown in red (*dglm.2020*). The VAST spatiotemporal index used in the diagnostic case of the 2020 stock assessment is shown in blue. (*vast.2020*). The asymptotic 95% confidence intervals are shown via the corresponding shaded polygon.

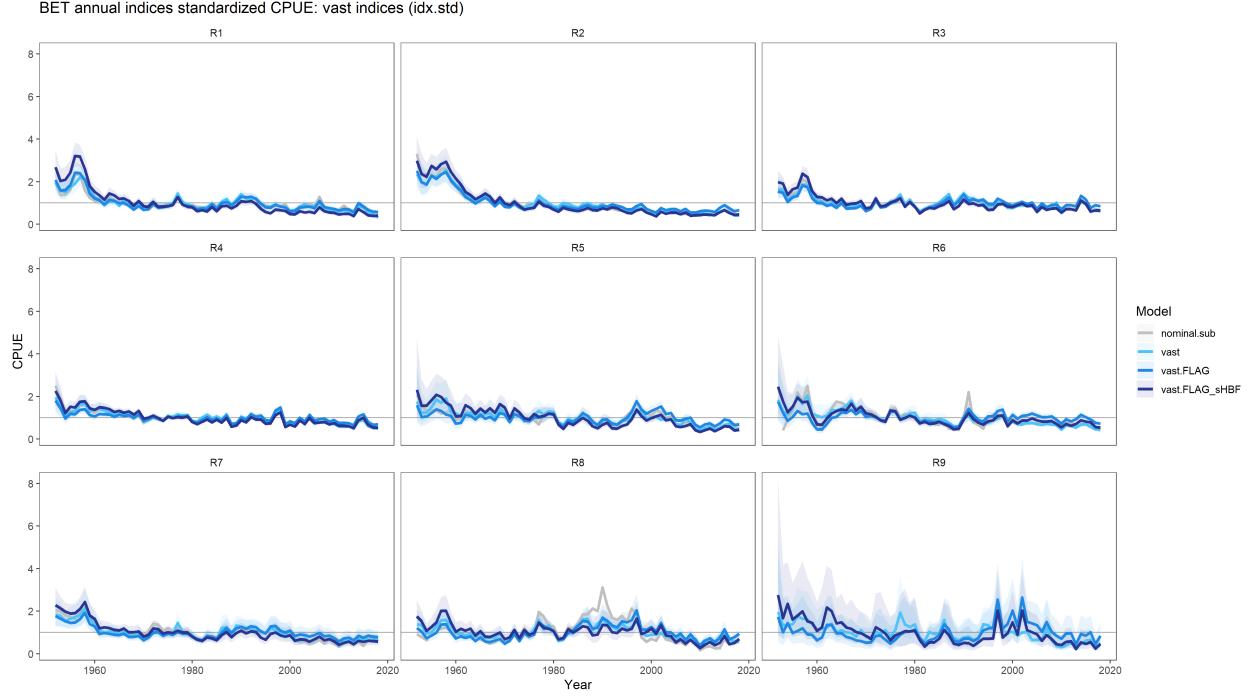


Figure 19: Stepwise plot showing the effect of including each additional covariate on the estimated index for bigeye tuna. The nominal index corresponding to the subset of data that the standardization model was fit to is shown in light gray (*nominal.sub*). The vast index with only spatial and spatiotemporal random effects is shown in light blue (*vast*). The vast index with inclusion of a fixed effect for Flag-group is shown in blue (*vast.FLAG*). The final vast index with inclusion of a fixed effect polynomial spline for *HBF* is shown in dark blue (*vast.FLAG_sHBF*). The asymptotic 95% confidence intervals are shown via the corresponding shaded polygon.



Figure 20: The number of records over time by Flag-group per model region.

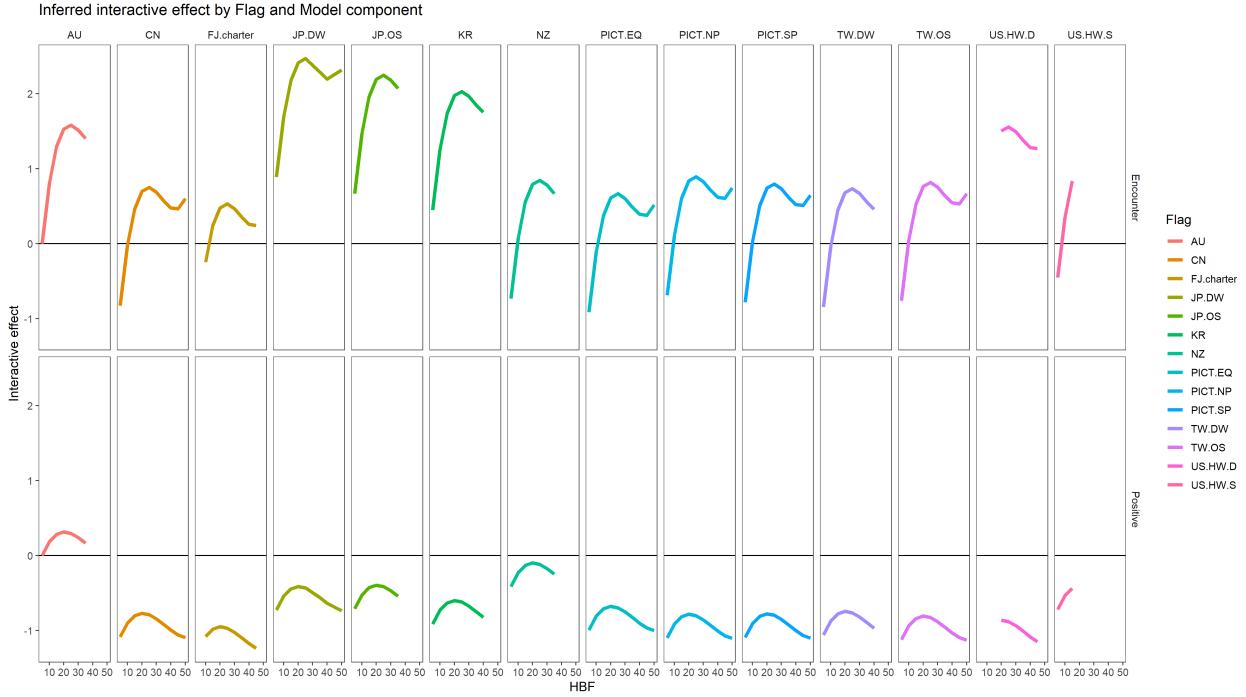


Figure 21: Though interactions were not explicitly modelled, the estimated inferred interactive effects of Flag-group and *HBF* for the two components of the delta model (encounter probability and positive catch) are shown for bigeye tuna. The color denotes the flag group.

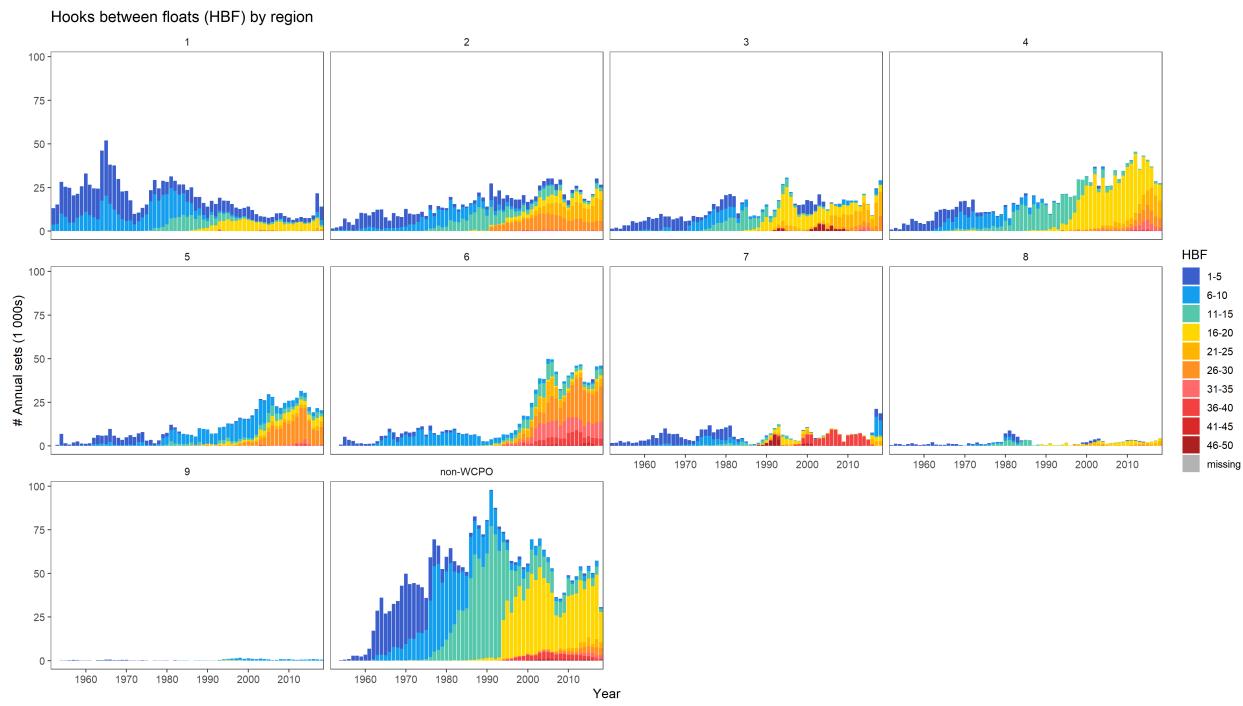


Figure 22: The number of records over time by 5-hook *HBF* bin per model region.

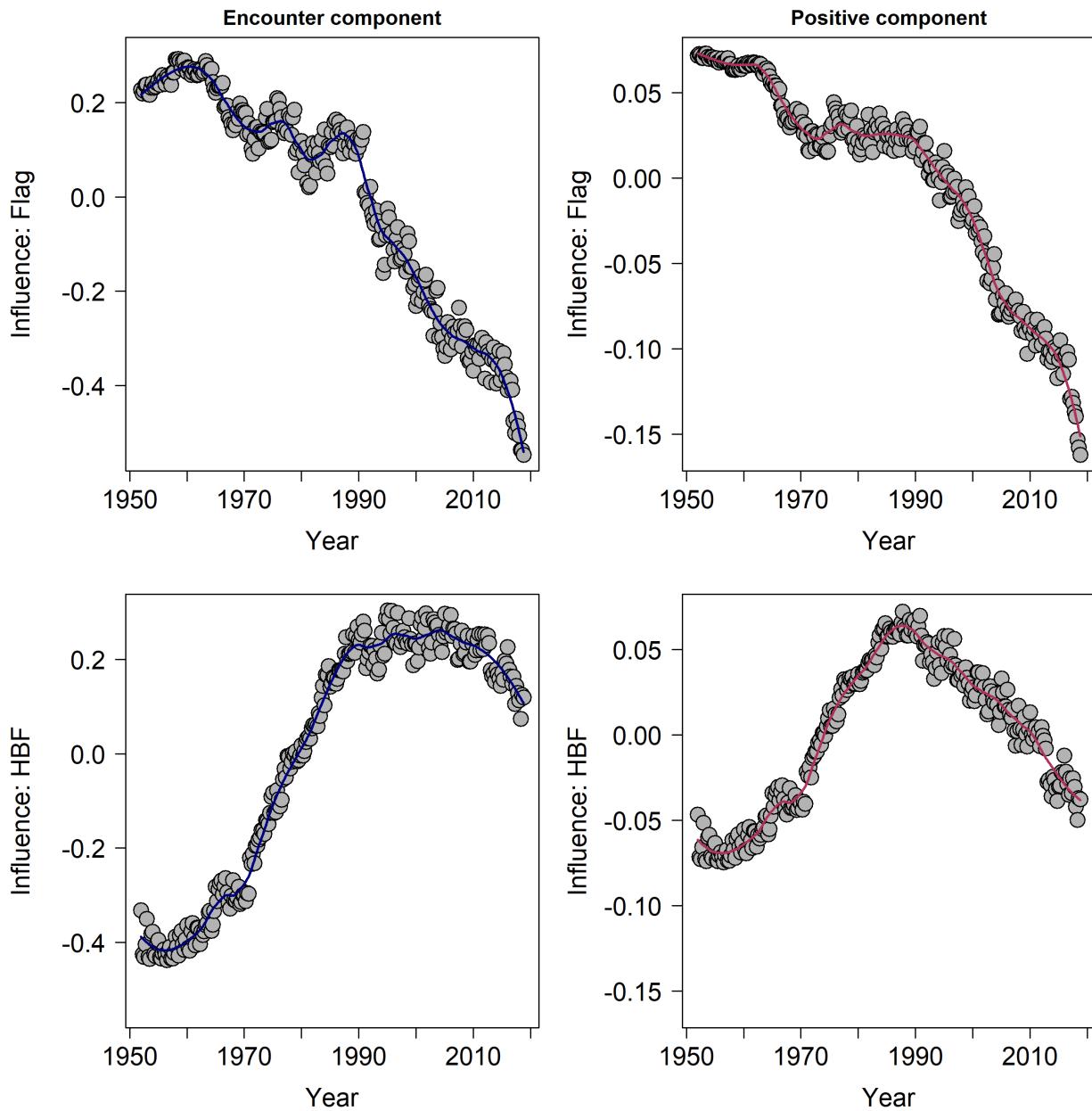


Figure 23: Influence plot (bigeye tuna) for each covariate and each model component, showing the effective correction in the index as a result of a change over time of the associated factor.

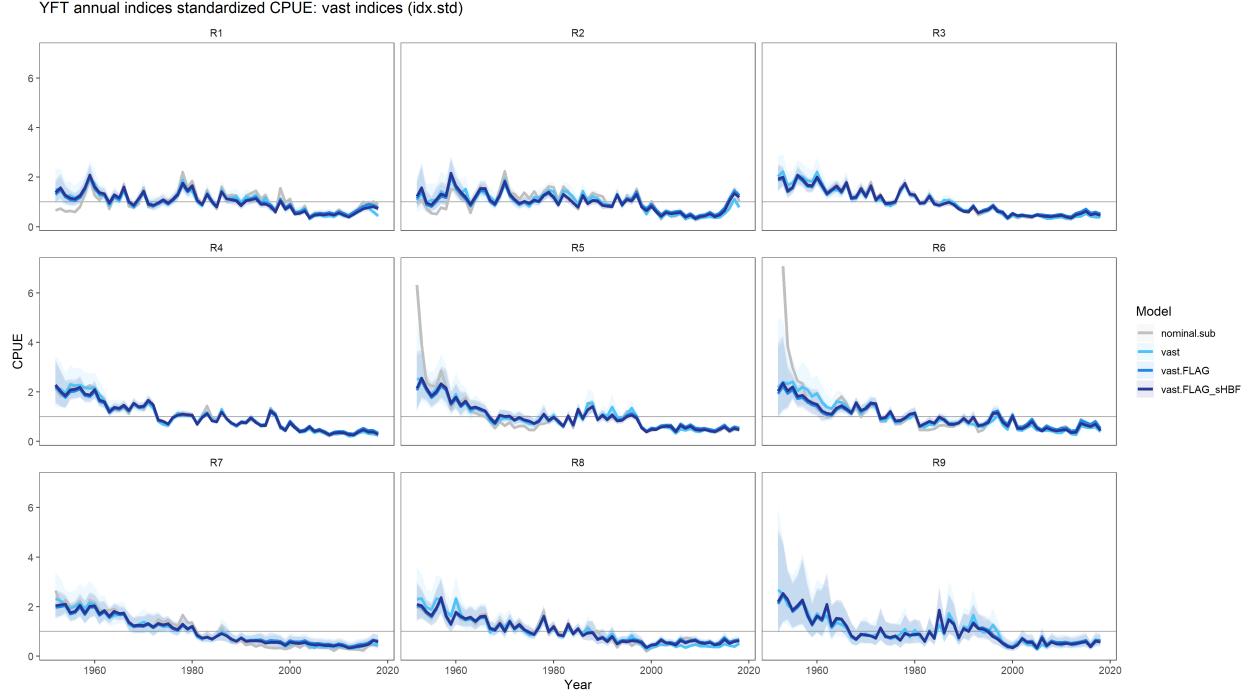


Figure 24: Stepwise plot showing the effect of including each additional covariate on the estimated index for yellowfin tuna. The nominal index corresponding to the subset of data that the standardization model was fit to is shown in light gray (*nominal.sub*). The vast index with only spatial and spatiotemporal random effects is shown in light blue (*vast*). The vast index with inclusion of a fixed effect for Flag-group is shown in blue (*vast.FLAG*). The final vast index with inclusion of a fixed effect polynomial spline for *HBF* is shown in dark blue (*vast.FLAG_sHBF*). The asymptotic 95% confidence intervals are shown via the corresponding shaded polygon.

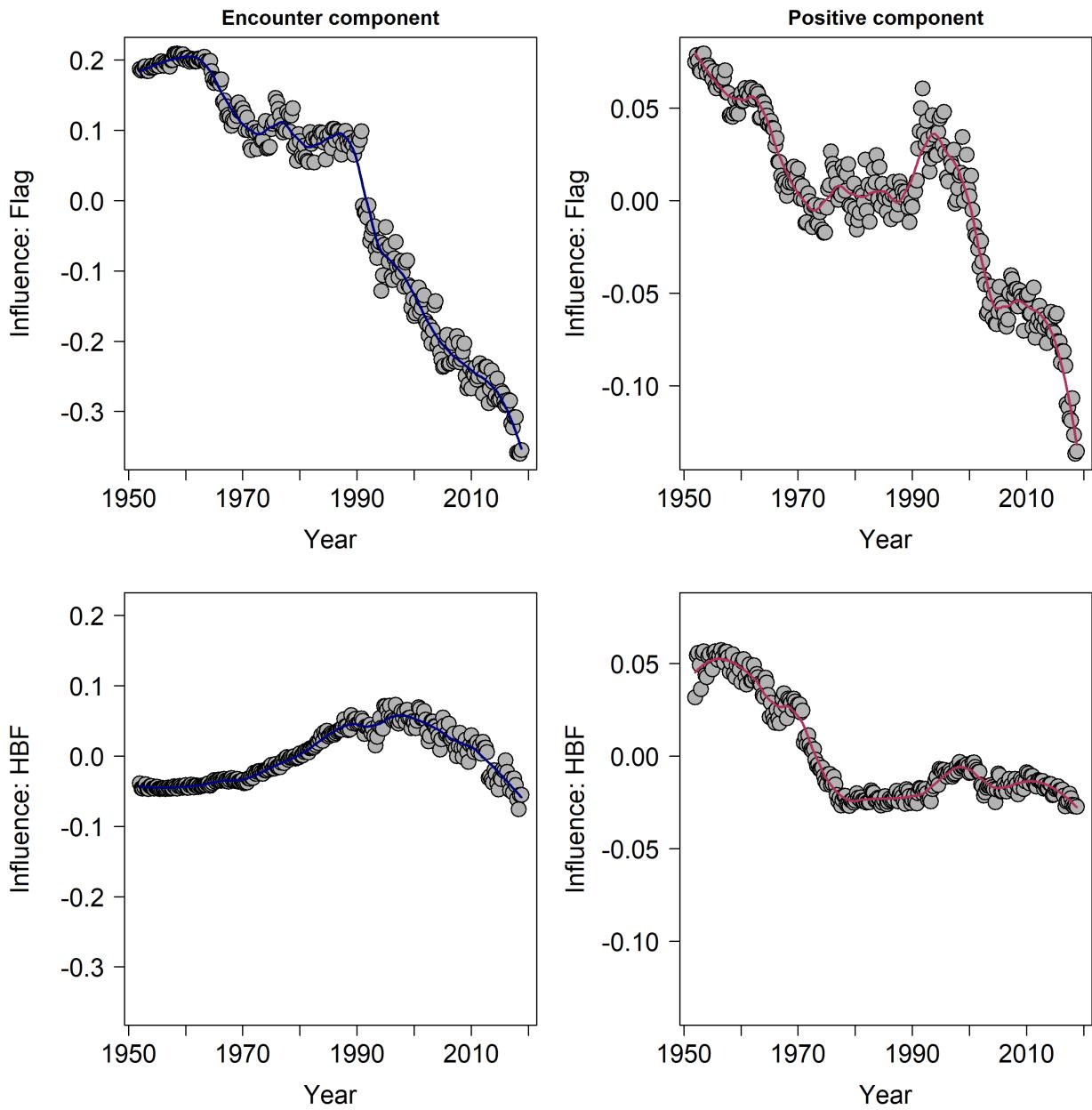


Figure 25: Influence plot (yellowfin tuna) for each covariate and each model component, showing the effective correction in the index as a result of a change over time of the associated factor.

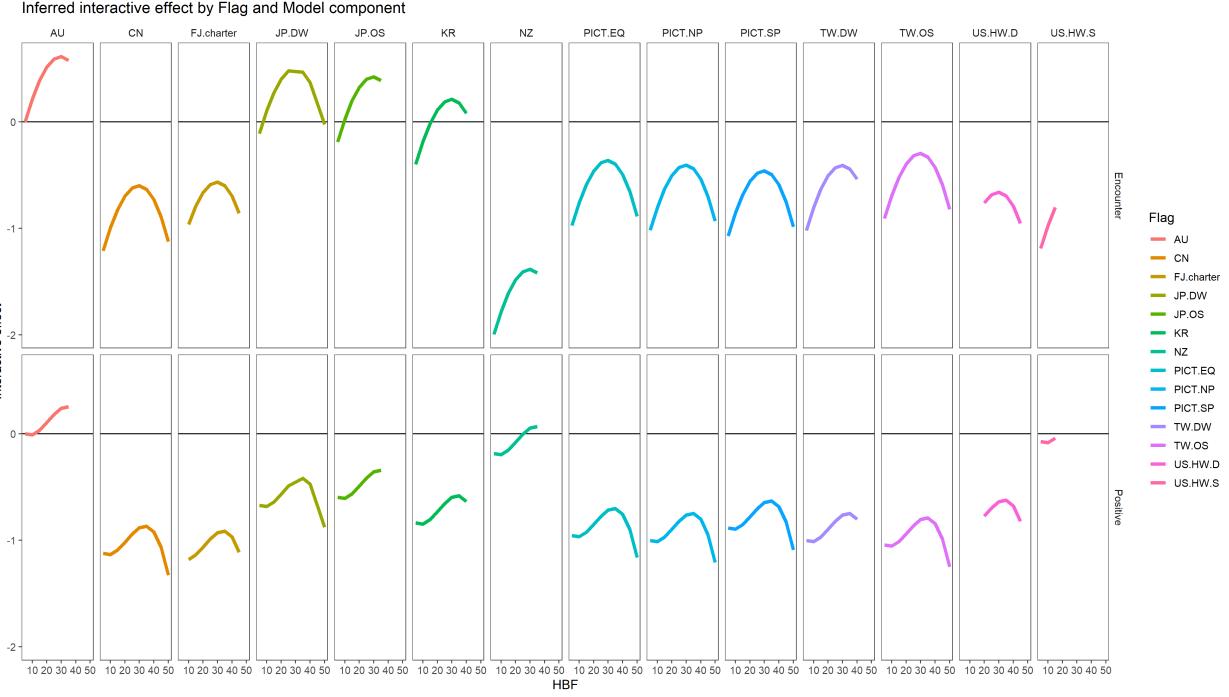


Figure 26: Though interactions were not explicitly modelled, the estimated inferred interactive effects of Flag-group and *HBF* for the two components of the delta model (encounter probability and positive catch) are shown for yellowfin tuna. The color denotes the flag group.

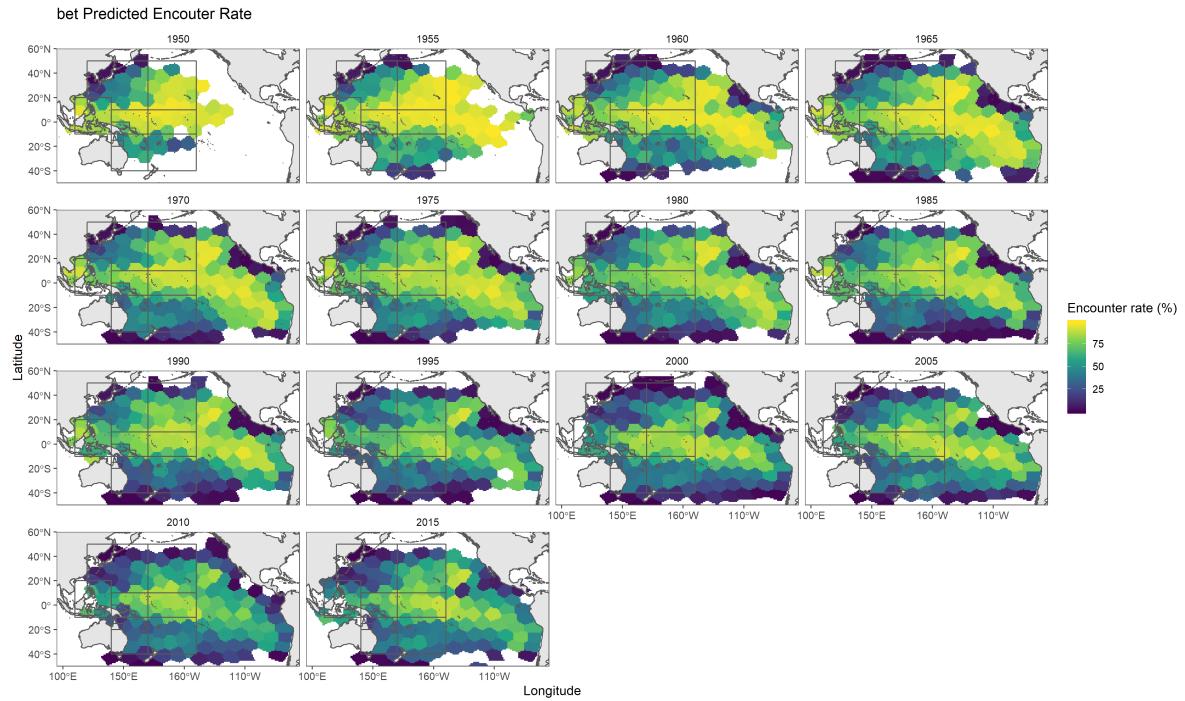


Figure 27: Model predicted estimates of spatiotemporal encounter rate for bigeye tuna

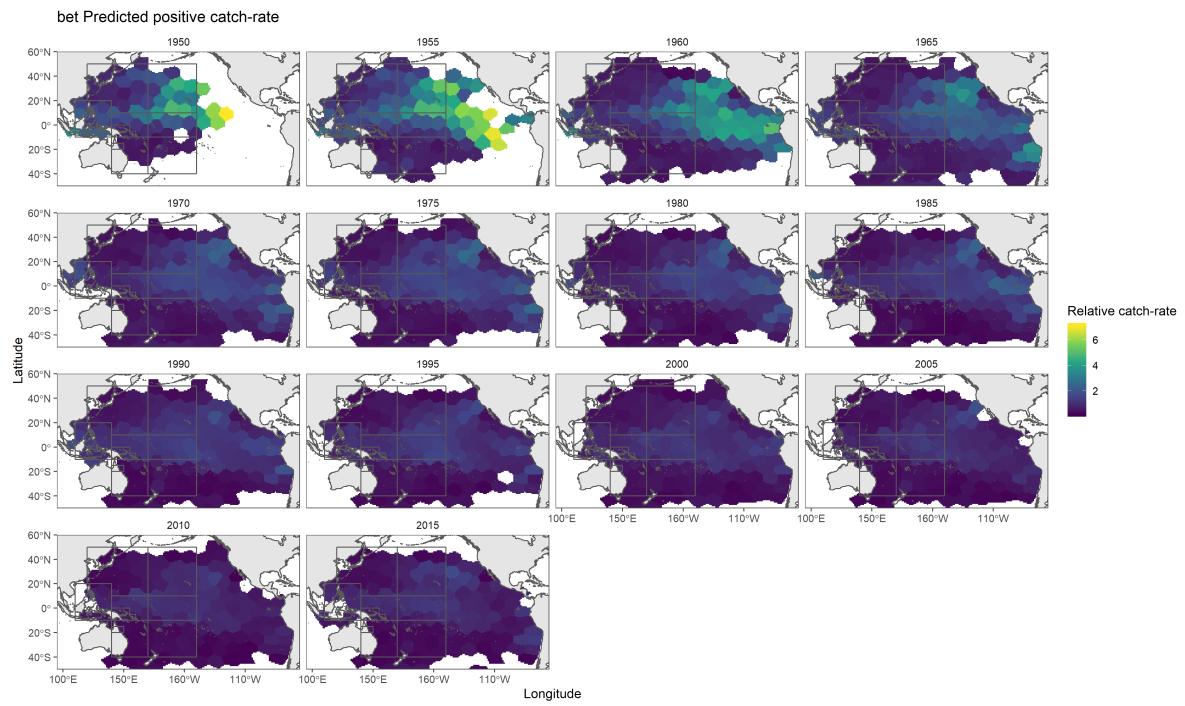


Figure 28: Model predicted estimates of spatiotemporal relative positive catch-rate for bigeye tuna

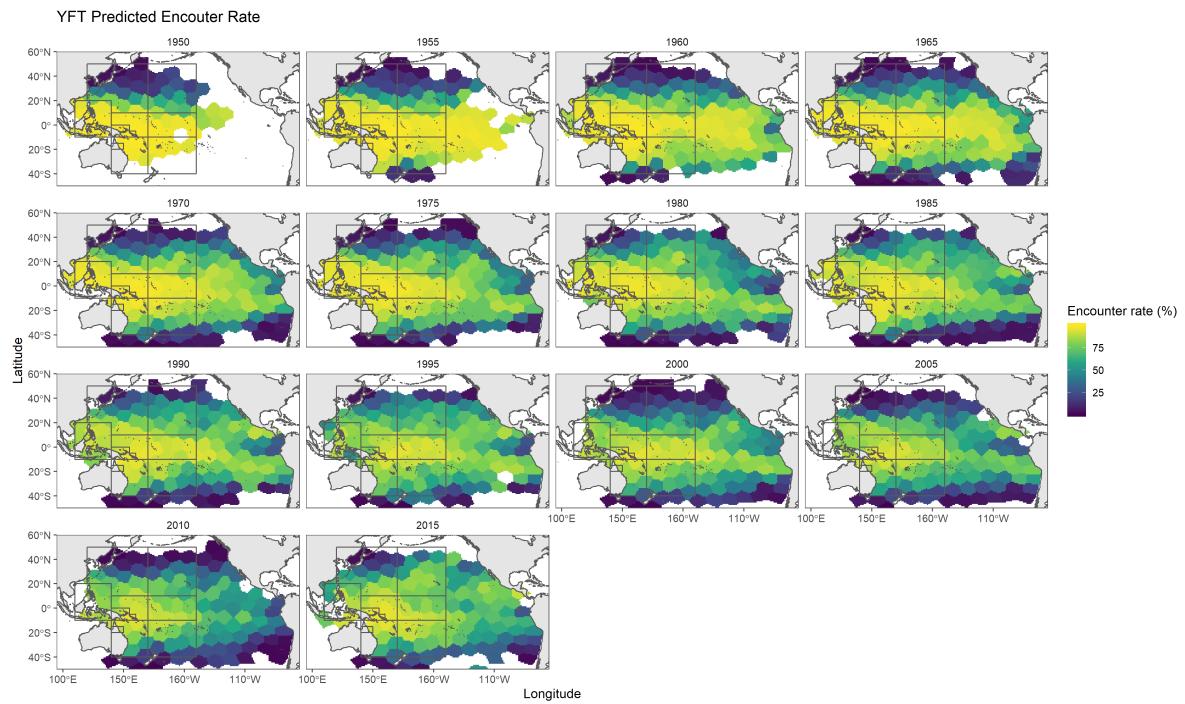


Figure 29: Model predicted estimates of spatiotemporal encounter rate for yellowfin tuna

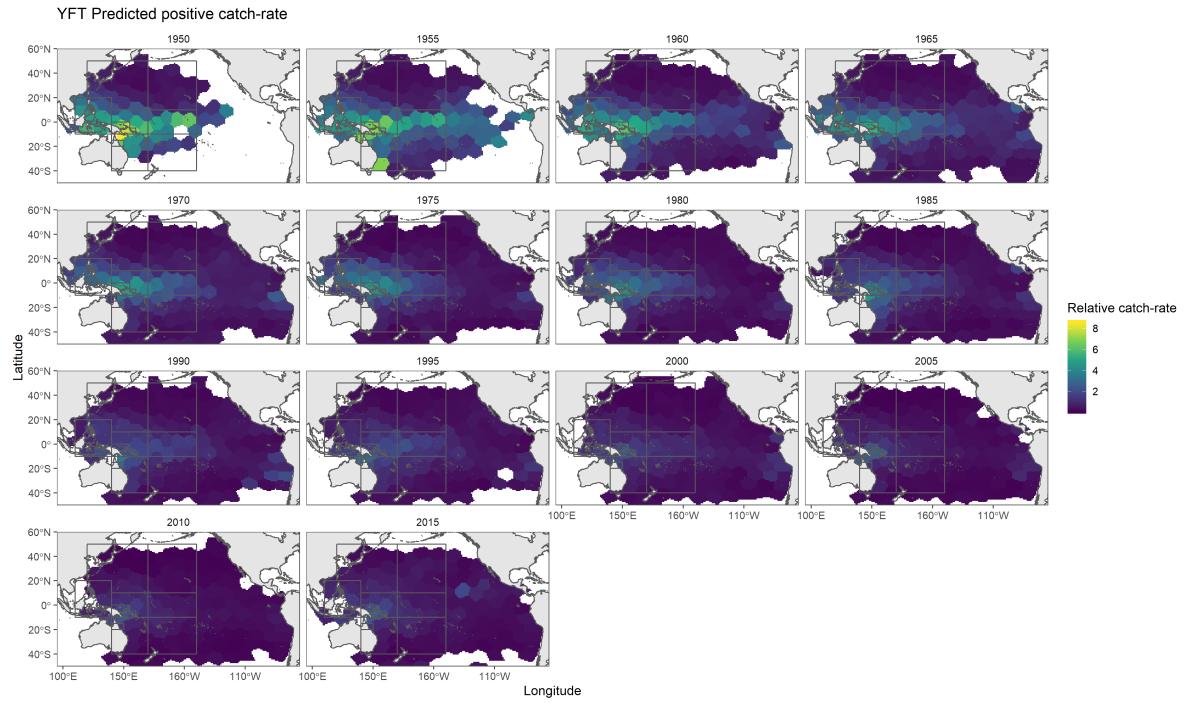


Figure 30: Model predicted estimates of spatiotemporal relative positive catch-rate for yellowfin tuna

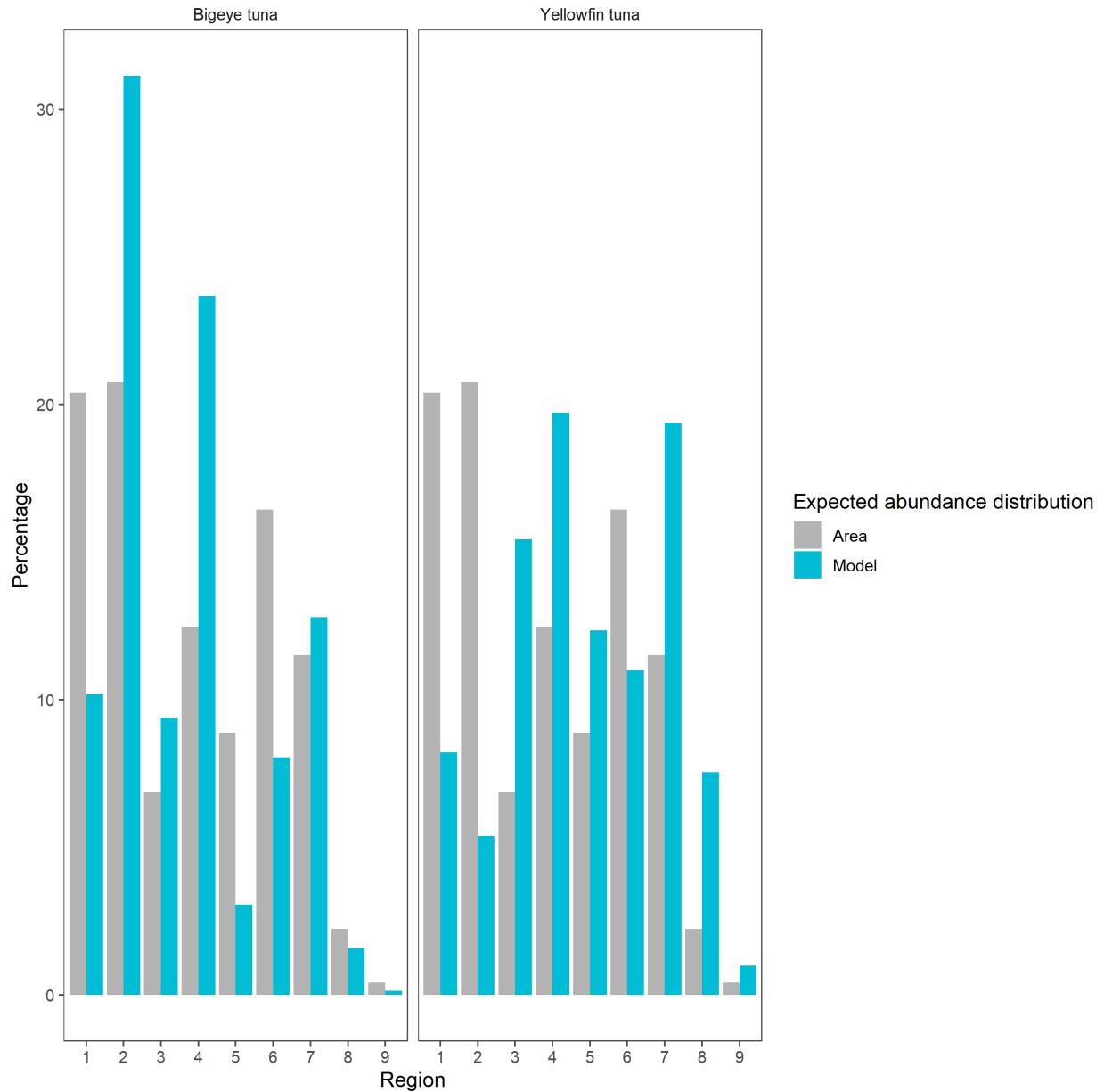


Figure 31: Expected distribution of regional abundance assuming abundance is proportional to the size of the spatial region (*Area*) versus the estimates from the spatiotemporal CPUE standardization model (*Model*).

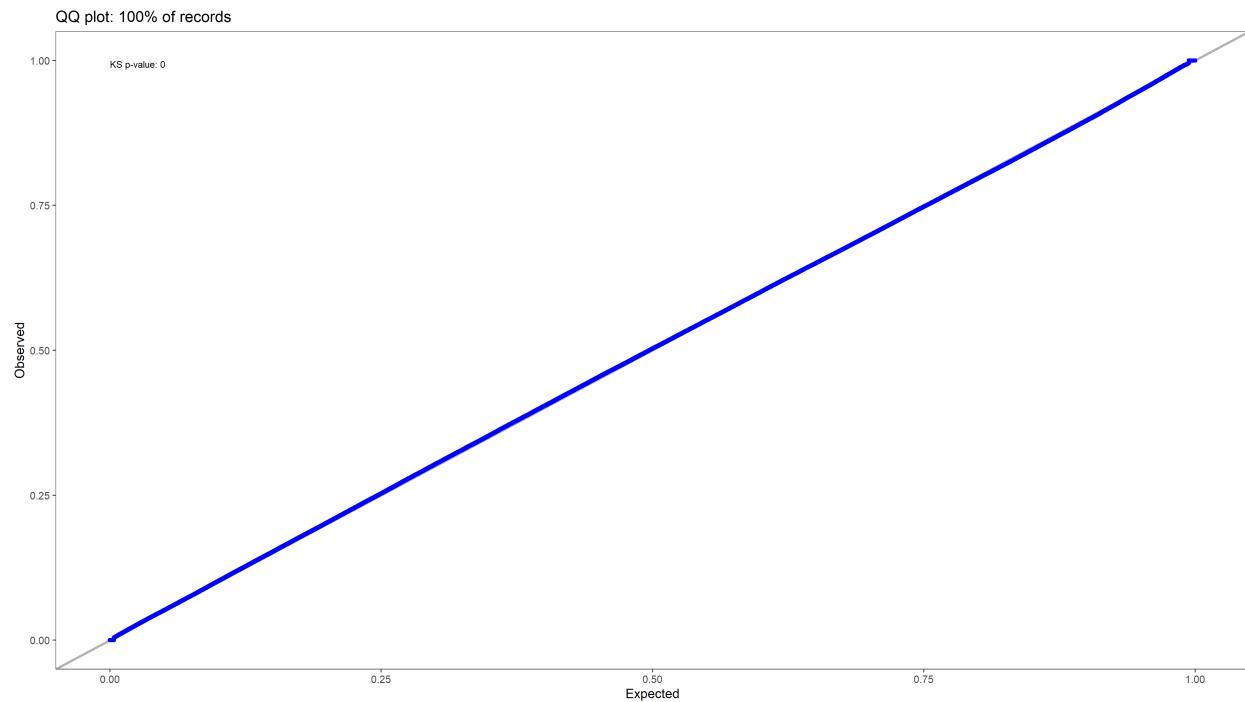


Figure 32: Uniform Quantile-Quantile (QQ) plot of the DHARMA calculated residuals for bigeye tuna.

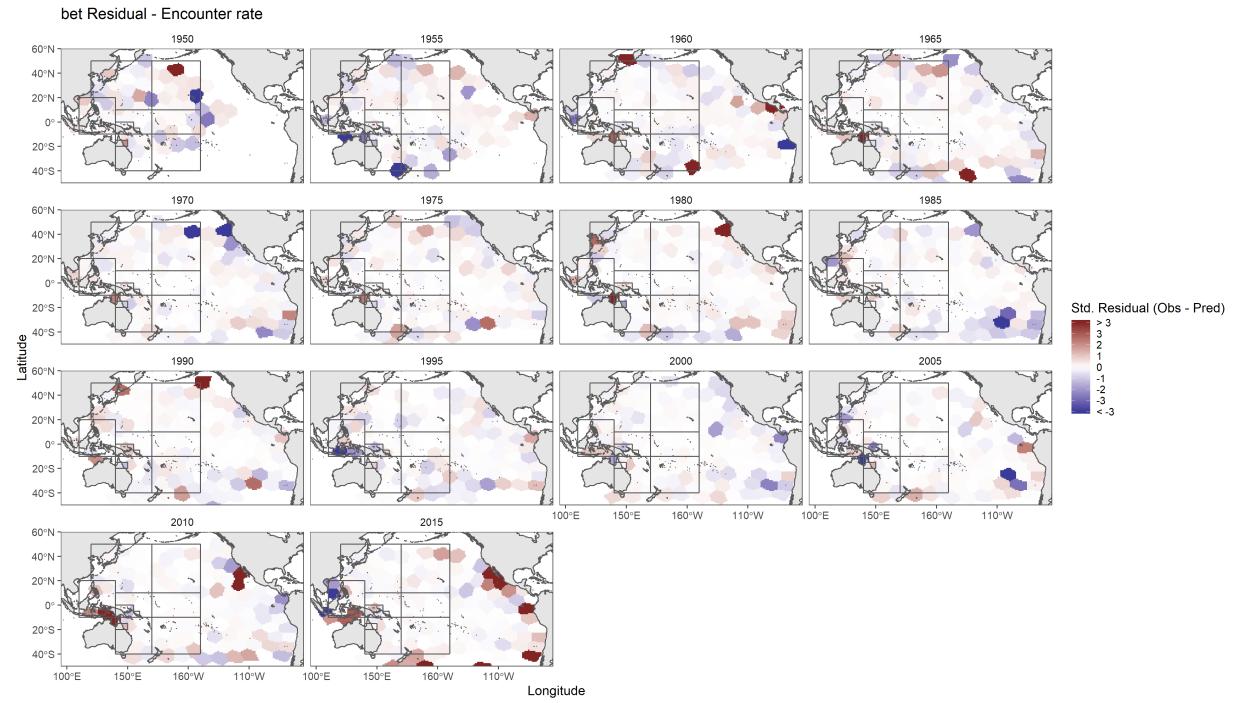


Figure 33: Model residuals of spatiotemporal encounter probability for bigeye tuna

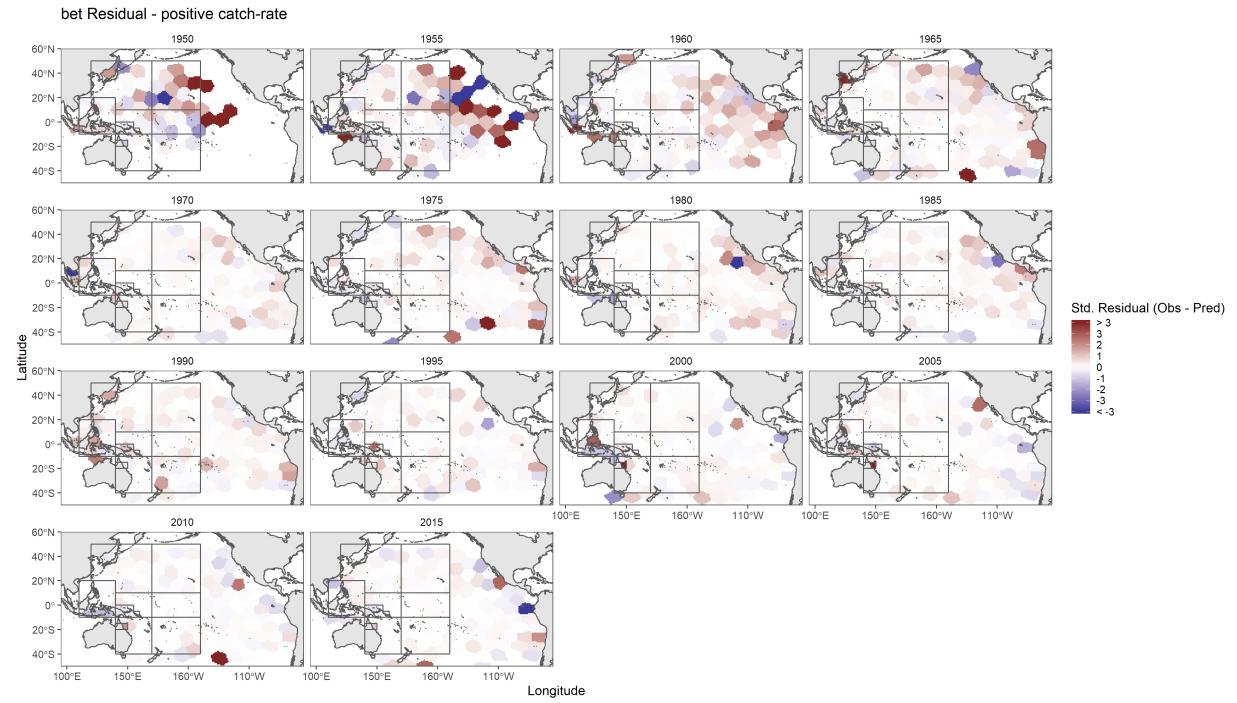


Figure 34: Model residuals of spatiotemporal positive catch-rate for bigeye tuna

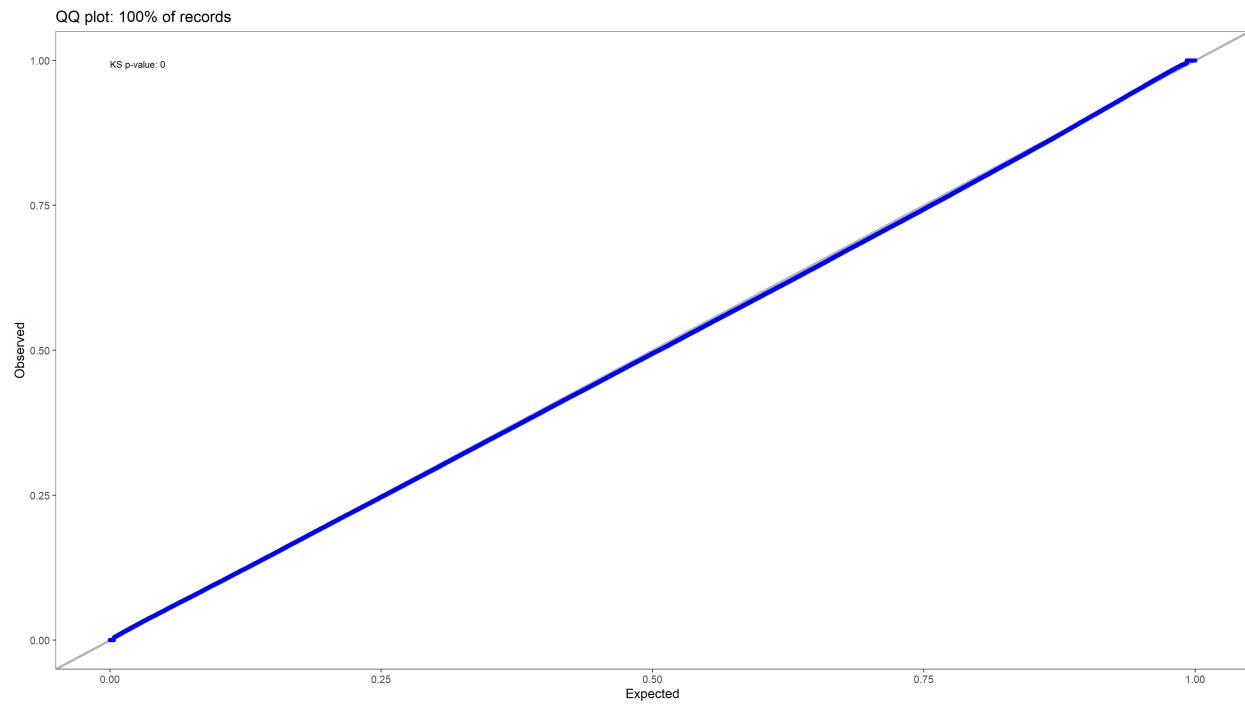


Figure 35: Uniform Quantile-Quantile (QQ) plot of the DHARMa calculated residuals for yellowfin tuna.

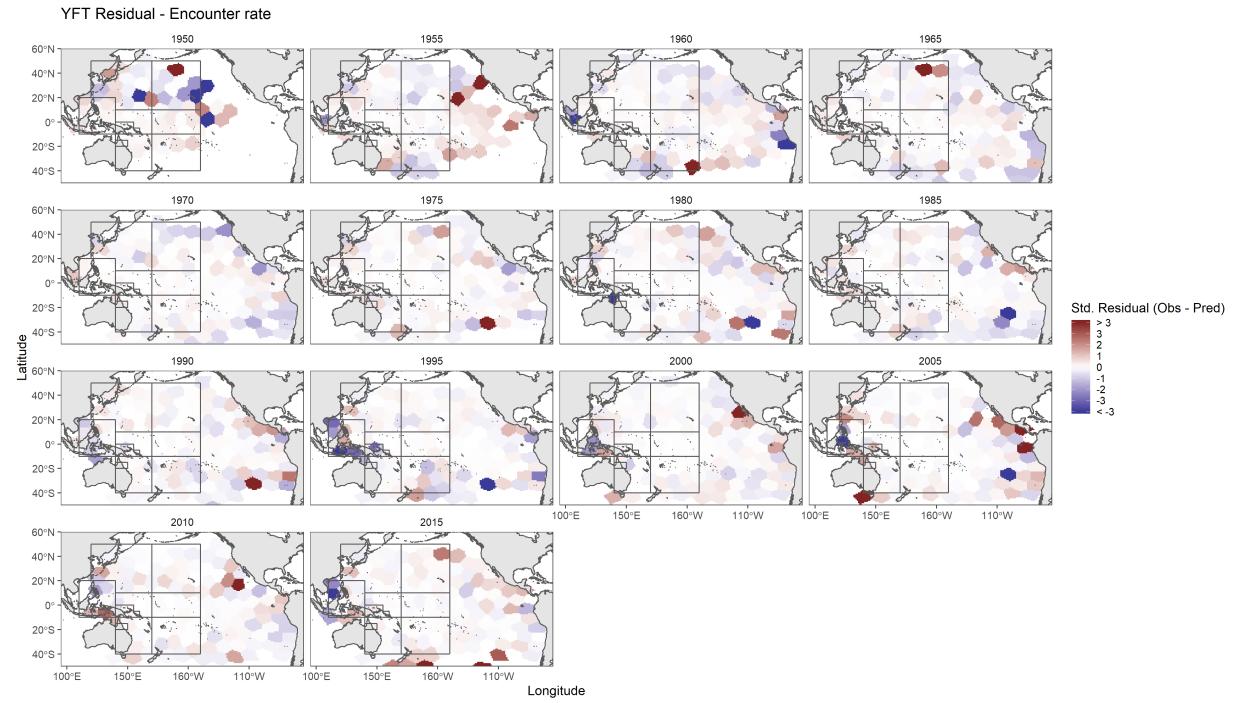


Figure 36: Model residuals of spatiotemporal encounter probability for yellowfin tuna

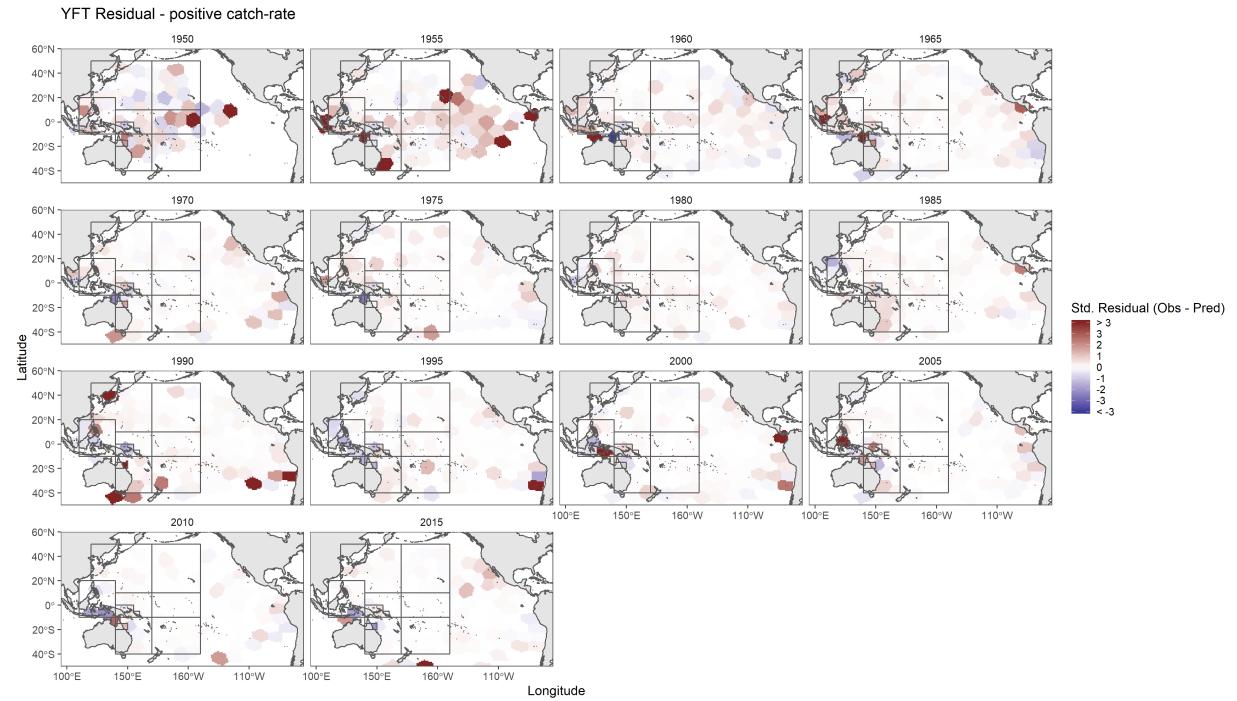


Figure 37: Model residuals of spatiotemporal positive catch-rate for yellowfin tuna