SCIENTIFIC COMMITTEE
THIRTEENTH REGULAR SESSION
Rarotonga, Cook Islands
9–17 August 2017

Use of operational vessel proxies to account for vessels with missing identifiers in the development of standardised CPUE time series

WCPFC-2017-SC13/SA-WP-04

L. Tremblay-Boyer[1], G.M. Pilling

[1]Oceanic Fisheries Programme, The Pacific Community

1 Executive summary

CPUE standardization involves modelling CPUE over time to account for additional variables that can impact CPUE but which are not reflective of abundance trends. One variable that has consistently been found to be influential in standardization is the vessel identifier, that is, a variable that identifies which fishing sets are performed by a given vessel. For the 2017 assessments, an extensive operational LL dataset has been amalgamated from the SPC-held data together with the data held by all the important distant water fishing nations. This provided a significantly extended temporal span for longline indices and as such, the assessments for bigeye and yellowfin in 2017 have series that extend back in time for more than 25 years in some regions compared to the previous assessments for this species. A challenge is that in many assessment model regions, the bulk of the historical effort (e.g. 1950 to the 1980s) does not have unique vessel identifier information. This is a common issue across tuna RFMOs. The *de facto* solution is to estimate a generic missing vessel ID for that historical period, combined with vessel effects for a core fleet. However, as vessel IDs only become available later in the time series, this causes, by definition, an abrupt change in the value for this variable over time, thereby introducing a temporal bias in the standardization for the variable.

The idea for this paper was initially presented at the 2017 Pre-assessment workshop in Noumea and consists of a proof-of-concept rather than a comprehensive development of the new methodology. As such, the generated indices were not used as the basis for the diagnostic case for this year's bigeye and yellowfin assessment, but used instead as a one-off sensitivity (referred to under *CPUE-Proxy* in McKechnie et al., 2017a and Tremblay-Boyer et al., 2017).

We use the term *vessel proxy* to refer to a new variable that groups individual fishing sets via clustering based on a set of explanatory variables. These variables are present in the dataset on a longer timespan than the vessel IDs are, even in very early years, and would be expected to be roughly similar (i.e. distributed around an average) for a given vessel. Our aim is not to exactly recreate vessel IDs but to create groupings that would be representative of sets of operational characteristics present within a fleet, hence the term 'proxy'. The potential efficacy of the approach is examined primarily through correlation between the true and proxy vesssel effects, using the extensive Japanese fleet data as a basis because (1) this fleet represents the bulk of the early longline effort; and, as such, (2) has the highest proportion of historical missing vessel identifiers.

The approach appears to work well in particular model regions for both bigeye and yellowfin, potentially due to the history of fishing effort by various fleets within a region, and the fact that Japanese effort reduces to very low levels in some WCPO regions after the 1980s, which means that even the vessel standardised index is highly variable. In addition, the approach appears to have performed worst in region 5, where effort is highly seasonal, suggsting some form of seasonality in the classification might improve results. Finally there is currently no clear indication of the optimal number of clusters to provide to the algorithm. An experimental approach has been used here, where the approach was fitted to a training dataset with vessel IDs, and the assumption made that the same rule could apply to earlier years. This is a strong assumption, in particular because vessel operational characteristics such as hooks between floats appeared to be less variable earlier in the data set. However, there is the potential to examine the performance of the approach through simulation, where these factors can be defined and controlled.

The method as is appears to show promise, but there are a number of aspects that need to be explored and could be optimized:

1. Current validation has focused on mapping the estimated proxy effects against true vessel

effects and extrapolate from there to select an optimal vessel proxy number for the prediction dataset. This assumes that the testing dataset is representative of the prediction dataset in the way the operational covariates map out to vessel effects, and also that the optimal cluster number is related to effort. Given the importance of the relative change in time in the value of covariates on standardized indices, this aspect could also be explored as a way of assessing performance.

2. Given the trend in lognormal proxy effects performing better than binomial ones, other error distributions that only require one effect, such as the negative binomial, might be useful to explore.

3. The inclusion of additional covariates should be considered, such as set start time, although some reconstruction work might be needed. National scientists might also be able to suggest other non-standard covariates which are available when vessel IDs are missing, and might give further information on vessel ID clusters.

4. The definition for the testing dataset needs to be expanded, as in the current case, the Japanese core fleet became quite small in some regions such that it was unclear if the poor performance of the method was due to the method itself or the testing dataset being very variable.

5. Performance was sensitive to the number of clusters assumed, in some instances in unpredictable ways. Given its potential influence, this part of the methodology needs to be refined as a priority. Hierarchical clustering might be a more practical approach as the algorithm only needs to be run once to generate a range of cluster numbers. Supervised classification methods should also be explored.

Overall, additional investigation could improve performance, and would greatly benefit from collaboration with other national agencies with better intrinsic understanding of the drivers of fleet performance across the region.

We invite SC13 to:

- discuss the approach used here, the areas of potential development, and its potential for future stock assessments within the WCPO;
- note the importance of vessel and gear information as inputs into CPUE standardisation;
- discuss collaborative work between the scientific services provider and national scientists to enhance the analyses.

2 Introduction

Time-series of relative abundance are one of the key inputs in stock assessments, including those conducted for tuna, billfish and sharks in the WCPO. In this region, these time-series are generated by standardizing catch-per-unit-effort (CPUE) data, typically derived from longline fisheries given their long history of fishing in the area, anticipated closer link between CPUE and underlying stock abundance, and the broad spatial span of their activity.

CPUE standardization involves modelling CPUE over time to account for additional variables that can impact CPUE but which are not reflective of abundance trends. Those variables can include gear type, species targeting, time of set, area fished, etc. One variable that has consistently been found to be influential in standardization is the vessel identifier, that is, a variable that identifies which fishing sets are performed by a given vessel. This variable represents an array of operational features for which the fine-scale data are frequently not available, such as vessel features, gear type/configuration used, the experience of the skipper, the fishing strategy used by the crew, etc.

Recently distant-water fishing nations agreed to provide their operational longline data to SPC for use as the basis for abundance indices in WCPO assessments (OFP, 2015). This provided a significantly extended temporal span for longline indices and as such, the assessments for bigeye and yellowfin in 2017 have series that extend back in time for more than 25 years in some regions compared to the previous assessments for this species (see McKechnie et al., 2017b).

The extended time series of operational data brings additional challenges. In many regions, the bulk of the effort from 1950 to the 1980s does not have unique vessel identifier information. This problem also occurs in other t-RFMOs (e.g. Hoyle et al., 2016) and has proved an ongoing challenge to the calculation of both point-estimates and measures of variation for standardized indices. Although attempts were made by the nations concerned to investigate these information gaps (outlined in OFP, 2015 and further discussed in Pilling and Brouwer, 2017a), further information could not be obtained, although some additional information on fleet ID (e.g. coastal, offshore, distant water) is available. The absence of this information means the standardization model cannot capture some key operational features, such that a potentially influential variable has to be excluded from the analysis.

To make the best use of the available information, a short and a long CPUE time-series were presented in previous analyses of this dataset (McKechnie et al., 2015), whereby the short contained vessel id for a core fleet subset, and the long time-series had a generic missing value for vessel ID. The latter approach tends to be used *de facto* when vessel information is missing for part of the time series (see also Hoyle et al., 2016). If the distribution of fishing effort over a covariate such as vessel ID does not change over time, standardizing for this covariate should have little impact on the indices. Conversely, if the distribution of effort over a covariate whose levels show a signal in CPUE response, for example relative levels of effort on different types of hook, of vessel, a fleet, etc., the impact of standardizing for this covariate will be important. This is why the *de facto* approach of estimating a generic missing vessel ID together with vessel effects for a core fleet is problematic: vessel IDs only start being available later in the time series, which causes, by definition, an abrupt change in the value for this variable over time. Since vessel IDs typically capture, in part, the impact of evolving fishing practices (techniques, technologies, targeting, etc.), which are known to impact catch rates (Harley et al., 2001; Ward and Hindmarsh, 2007), a temporal bias in the standardization for the variable is therefore likely to be introduced. Figure 1 to 9 show the proportion of sets with missing vessels for key fleets by region from the 2017 updated regional structures for the bigeye and yellowfin stock assessments.

In this paper, we propose a new approach that 'reconstructs' historical vessel IDs based on other variables available throughout the time series of the dataset which have the potential to be representative of the vessel effects. This proxy variable can then be used in lieu of a generic 'missing vessel' effect for sets lacking vessel identifiers. The idea for this paper was initially presented at the 2017 Pre-assessment workshop in Noumea (Pilling and Brouwer, 2017a) and consists of a proof-of-concept rather than a comprehensive development of the new methodology. As such, the generated indices were not used as the basis for the diagnostic case for this year's bigeye and yellowfin assessment, but used instead as a one-off sensitivity (referred to under *CPUE-Proxy* in McKechnie et al., 2017a and Tremblay-Boyer et al., 2017).

3 Methods

The analyses presented here focused on the Japanese fleet, because (1) this fleet represents the bulk of the early longline effort; and, as such, (2) has the highest proportion of historical missing vessel identifiers.

We use the term *vessel ID* to refer to the categorical variable which assigns individual fishing sets to a specific vessel. In practice, the definition differs between countries, with some referring to the physical vessel throughout its lifetime in the fishing fleet, while others use the international call sign or the vessel name (which could change between owners even though the physical vessel is the same). We do not address these discrepancies since they would have minimal impact here given that we are dealing with a single fleet whose vessel IDs, when available, have been confirmed to be unique to an individual vessel (Pilling and Brouwer, 2017b).

The term *vessel proxy* is used here to refer to a new variable that groups individual fishing sets via clustering based on a set of explanatory variables. These variables are present in the dataset on a longer timespan than the vessel IDs are, even in very early years, and would be expected to be roughly similar (i.e. distributed around an average) for a given vessel. Our aim is not to exactly recreate vessel IDs but to create groupings that would be representative of sets of operational characteristics present within a fleet, hence the term 'proxy'. Since the goal of standardizing is to capture relative changes in the amount of fishing effort expended in categories with characteristics that may impact catch rates, we do not necessarily need to recreate exact vessel IDs, as long as we can create groupings that are representative of these key features and change over time at comparable rates.

The formula for a typical GLM used to standardize longline catch rates for tuna in the WCPO, (e.g. McKechnie et al., 2015), would be:

response $\sim \beta_0 + YQ_i + cell_i + cluster_i + vessel_i$

where YQ is the year/quarter, cell is the lat/lon, cluster is the assigned cluster, and vessel is the vessel ID.

The value for the vessel coefficient $vessel_i$ is the one we are trying to reproduce here, noting that the actual value should have minimal influence as long as the relative ranking between vessel coefficients is on average preserved and, similarly, the relative change over time in how fishing effort is expended between low and high CPUE vessels (or vessel proxies). The approach taken is described in the following sections.

3.1 Description of the testing dataset

For each assessment model region, a testing dataset was identified consisting of the core Japanese fleet, where that core fleet was defined using the approach from previous assessments based upon region-specific minimum threshold of vessel activity (quarters active, number of sets per quarter, etc.). This threshold is necessary to ensure that sets retained in the analysis are representative of overall fishing activity in the region, and that vessel effects are only estimated (and thus standardized against) for vessels with enough fishing activity for the estimated effect to not be fully collinear with a few year-quarters. the testing data sets consisted of vessels of known ID, but vessel proxies were also developed so that the results of GLMs using the proxies could be compared with those using the actual vessel IDs.

The testing was done individually for each region of the 2017 bigeye and yellowfin stock assessments. The exception was that region 6 was not included in the analysis as it is the one region for which there is a good history of vessel ID presence in the dataset (see Figure 6). Region 9 was also not included as it had very little effort before the 1980s, and vessel IDs were present in the dataset once effort levels increased (see Figure 6). The results shown here are for the updated 2017 region structure, but 2014 region indices were also produced.

3.2 Covariate selection and formatting

The covariates available for the Pacific-wide longline operational dataset consisted of: date of longline set, longitude and latitude to the nearest degree, hooks used per set, hooks between floats, set start time, species catch in number for albacore, bigeye, yellowfin, skipjack, swordfish. Set start was not considered in the current analysis because of its partial coverage over time and issues in reporting format, but would be interesting to include in the future pending further developments.

As a first exploratory analysis to confirm that the available operational covariates could be used to predict vessel effects, estimated vessel effects from GLMs applied to the already filtered and clustered core fleet-short datasets (see detailed methods in McKechnie et al. (2015) and this year's settings in McKechnie et al. (2017b)) were mapped against summary statistics (mean, median, quantiles, etc.) of candidate covariates over the sets performed by the vessel.

3.3 Cluster identification

Vessel proxies were developed based upon clustering analyses. We used the k-means algorithm to assign fishing sets to clusters based on the set of covariates described above, which were standardized across sets so that each covariate had a mean of zero and a standard deviation of 1. This removed covariate scale effects (e.g. if left raw latitudes would be between -50 and 50, years between 1950 and 2015) in cluster identification (since k-means is a distance based algorithm). This algorithm assigns observations (i.e. sets) to a fixed number of clusters, as specified by the user.

Based on the exploration in Section 3.2 we elected to retain all covariates and perform the clustering for sets split by fleet type, as there was a distinct signature for this variable mapping to distinct covariate combinations in most regions. Set-level species catch information was also included. However, as vessels can switch targeting between seasons or have sets that differ in species composition due to chance, these four variables, that is, catch in individuals for albacore, bigeye, yellowfin and swordfish, were further processed to lower their weight in comparison to the operational characteristics. This was done by collapsing the 4 variables of species catch into a single axis *via* a principal component analysis, and using the first PCA axis (i.e. the one that explains the most variation) as the 'species composition' covariate in the clustering algorithm.

With k-means, the user must specify the desired number of clusters, i.e. vessel proxies, to be returned by the algorithm. Each observation gets assigned to one of N_c clusters, that is, the sets get split into N_c partitions. In order to explore how the results were impacted by the choice of N_c, 25, 50, 100, 150, 200 and 250 clusters were separately evaluated as a trial, and their performance assessed (see Section 3.4). These numbers were chosen arbitrarily to cover a range from a relatively small number of vessel proxies to around the number of vessels present in a 'typical' core fleet. The clustering was performed separately for each region to match the approach used in CPUE analyses (McKechnie et al., 2015), and for each fleet within that region. The total number of clusters in each region was set to match the N_c for that specific run, the number of clusters for each fleet in that region was a proportion of that fleet's effort (in sets) over the total effort in the region. For instance, for region X under the $N_c = 50$ scenario, 40 clusters were assigned to fleet A, and 10 clusters to fleet B, as fleet A exerted 80% of the effort in sets in the testing dataset. The final output for each fishing set in the core fleet of each region was thus a value of candidate vessel proxies going from 1 to N_c, depending on the defined cluster number scenario.

All analyses were performed in R (R Core Team, 2017).

3.4 Model runs and validation for the testing dataset

Vessel proxies were identified for sets of the Japanese core fleet in the step above. To evaluate how CPUE standardization with true vessel IDs compared to that with vessel proxies, multiple GLMs were performed on the testing dataset for each region (i.e. the Japanese core fleet) which included as the vessel effect either the actual vessel ID variable, or the estimated vessel proxies under each of the N_c scenarios. The GLM models were the same as used for standardization on this dataset in recent years, and used a delta-lognormal approach with a 5×5 cell effect and a cluster in addition to the year effect, as defined in McKechnie et al. (2015), which means a total of 14 models were run for each of the 7 regions tested (2×6 vessel proxy N_c scenarios, and an extra two with the true vessel ID for comparison). The estimated binomial and lognormal vessel effects were then extracted for each vessel or vessel proxy, and an average vessel proxy effect for each true vessel ID was calculated as a weighted mean across each of the effects for the proxy in which that vessel was found, weighted by the number of sets for that vessel which were assigned to that proxy. This weighting step was required as sets from a single vessel were found to usually be assigned to more than one vessel proxy.

To select the number of vessel proxies N_c* that best reproduced the actual vessel ID information for each region, the 'true' vessel effect was plotted against its estimated proxy effects under each cluster number scenario and model component (binomial, lognormal). For each region $\times$ model component combination, N_c* was defined as the N_c that resulted in the highest correlation between the weighted proxy effect and the true vessel effect.

3.5 Cluster selection for extension to long time-series

The first step required to move from the testing of the proxy approach to its actual application to sets where vessel IDs are missing (the 'prediction' dataset) was to decide on a number of proxies (N_c) to use for each region. Based on the analyses above which explored a range of candidate N_c values for the Japanese core fleet of each region, we calculated the N_c by scaling the N_c* identified for each region above as a function of the effort present in the testing dataset *vs.* that in the prediction dataset. This was done by first selecting the highest values of N_c* if it differed between the binomial and the lognormal component, and then multiplying it by the ratio of effort in the prediction dataset to the testing dataset.

Second, fleet type was readily available in the testing dataset but was missing for some of the earliest years in some regions (typically before 1960). In most cases however, all sets in years when fleet ID started being reported belonged to a single fleet (usually distant-water), such that in those instances, fleet ID for sets in earlier years could safely be assumed to also belong to this fleet. For region 1, more than one fleet was active in the years where fleet information become available and there was no clear way to assign a fleet ID to sets. For this region, we thus filtered sets to start the vessel proxy predictions when the fleet ID became available.

The vessel proxies were predicted for each region as described above for the testing dataset, except that the sets on which the prediction was made had no pre-existing information on vessel ID. The final GLMs were performed on a hybrid data set containing: (1) sets with a previously missing vessel IDs, now replaced with a vessel proxy ID across the whole time series; and (2) the usual core fleet, with their true vessel ID. Therefore, the vesselID of the core fleet was used when available, else the vessel was predicted. To examine the influence of these approaches on the resulting indices, a comparison by model region was made where no vessel effect was used within the GLM, vessels have either a true vessel ID or a generic 'Missing' vessel ID, and where all vessels have either their true vessel ID, if available, or, for Japanese sets with missing ID, a vessel proxy ID, generated following the approach above. These last indices were the ones used under the *CPUE-Proxy* one-off sensitivities for this year's bigeye and yellowfin assessments (McKechnie et al., 2017a; Tremblay-Boyer et al., 2017).

4 Results

Based on the exploration in Section 3.2 we elected to retain all covariates and perform the clustering for sets split by fleet type, as there was a distinct signature for this variable mapping to specific covariates combination in most regions. This is shown as an example for the lognormal vessel effects for bigeye tuna in Figures 10 to 16. Depending on the region there is not necessarily a clear relationship between the estimated vessel effect and covariate on their own, but combined, there are clear patterns of pooling by fleet ID, as well as correlations across multiple covariates, highlighting that some form of classification approach across these covariates might be effective at modelling vessel effects. For instance, hooks per set and hooks-between-floats are often related, and there is also often a relationship of hooks-per-set over time. In parallel these variables often map out to a degree to either longitude or latitude. Such relationships were to be expected based on common sense but it is useful to see that it is captured by the vessel coefficients estimated by GLMs, as it underscores that it might not be necessary to have access to the actual vessel ID to capture (at least part of) the effect it has on catch rates in a GLM.

The comparison of effects estimated for true *vs.* proxy vessel IDs in the testing dataset highlighted that the method as applied here performs with variable success across regions (Tables 1 and 2, Figures 17–18 for bigeye tuna, and 19–20 for yellowfin tuna). For regions 1, 2, 3, and 7 for yellowfin tuna, performance was quite good, especially for lognormal indices where the correlations were between 0.6 and 0.9. For other regions however, such as regions 3, 4 and 8 for bigeye tuna, and 4 and 5 for yellowfin tuna, the performance was quite poor, with correlations often close to zero or even negative. In general, the lognormal tended to perform better than the binomial component, and correlations tended to improve with increasing number of assumed clusters, although after a certain level little additional improvement was gained and performance sometimes dropped (e.g. yellowfin binomial indices for region 8).

When comparing the standardised indices resulting from the use of vessel proxies to those with full 'true' vessel IDs and the exclusion of vessel information, differences in performance are seen between regions (Figures 22 and 21). In general, the exclusion of a vessel ID can lead to increased variability in the standardised CPUE (e.g. yellowfin region 1, bigeye region 7), to differences in the CPUE trend (e.g. the later years of bigeye and yellowfin in region 3) or to the relatively flat CPUE series for yellowfin, but with high variability (e.g. region 5). The inclusion of the vessel proxy results in historical trends closer to those found with the full vessel ID, although differences can still be seen. Since we do not know what the true relative trend in abundance for each species is, it is hard to validate the method by comparing the indices standardized under each approach, but it is still useful to note that the indices estimated with the vessel proxies retain features similar accuracy and precision to those typically seen in the region.

5 Discussion

This method does not aim to estimate missing vessels perfectly, but to improve on the current approach of estimating a generic missing vessel, or not explicitly accounting for operational covariates at all. Performance of the vessel proxy, as evaluated by its relationship with true vessel effects for the testing dataset, varies greatly by region. This might be due on the one hand to the history of fishing effort by various fleets, and on the other hand because Japanese effort falls to very low levels in some WCPO regions after the 1980s, such that even the vessel standardized index is highly variable in the testing dataset. In addition, region 5, where effort is highly seasonal, appears to have performed the worst overall, indicating that some form of seasonality in the classification might improve results.

The performance of the approach was assessed mostly by the correlation between the true ID and proxy ID approaches. The performance could not be assessed for the long dataset as true vessel IDs were, by definition, unknown early in the data set, and we have little control over the testing dataset as it is dictated by the historical effort dynamics of fleets across the Pacific. Finally, as there is currently no clear indication of the optimal number of clusters to provide to the algorithm, an experimental approach has been taken here, where the approach was fitted to a training dataset with known vessel IDs, and the assumption made that the same rule could apply to earlier years. This is a strong assumption, in particular because vessel operational characteristics such as hooks between floats appeared to be less varied earlier in the data set. However, there is the potential to examine the performance of the approach through simulation, where these factors can be defined and controlled.

Even though year-quarter and lon-lat are already present in the typical CPUE standardization model, such that including them to predict vessel proxies might appear circular, in practice 'true' vessel IDs are already collinear with these variables as individual vessels typically tend to focus their effort within a restricted area, and are also active within a continuous window of time over one part of the time-series, and then tend to gradually phase out. Using this variable also helps to prevent grouping together sets occurring in completely different periods of time, or in completely different areas within the region.

Finally, the method as is shows promise, but there are a number of aspects that need to be explored and could be optimized:

1. Current validation has focused on mapping the estimated proxy effects against true vessel effects and extrapolate from there to select an optimal vessel proxy number for the prediction dataset. This assumes that the testing dataset is representative of the prediction dataset in the way the operational covariates map out to vessel effects, and also that the optimal cluster number is related to effort. Given the importance of the relative change in time in the value of covariates on standardized indices, this aspect could also be explored as a way of assessing performance.
2. Given the trend in lognormal proxy effects performing better than binomial ones, other error distributions that only require one effect, such as the negative binomial, might be useful to explore.
3. The inclusion of additional covariates should be considered, such as set start time, although some reconstruction work might be needed. National scientists might also be able to suggest other non-standard covariates which are available when vessel IDs are missing, and might give further information on vessel ID clusters.
4. The definition for the testing dataset needs to be expanded, as in the current case, the Japanese core fleet became quite small in some regions such that it was unclear if the poor performance of the method was due to the method itself or the testing dataset being very variable.
5. Performance was sensitive to the number of clusters assumed, in some instances in unpredictable ways. Given its potential influence, this part of the methodology needs to be refined as a priority. Hierarchical clustering might be a more practical approach as the algorithm only needs to be run once to generate a range of cluster numbers. Supervised classification methods should also be explored.

6 Conclusion

Given the prevalence of the issue of the absence of vessel ID information in data sets across RFMOs, the importance of accounting for operational covariates in CPUE standardizations that are important for regional stock assessments, the practical advantages of this method *cf.* the estimation of standard errors for indices, and the potential to be more flexible with the core fleet definition, it was felt useful to present the approach at this stage to SC for review. The work presented here is preliminary and further exploration of the approach is needed. For this reason it was used as a one-step sensitivity scenario in the 2017 bigeye and yellowfin assessments, and not for the diagnostic case. However, early results appear promising for some model regions, and there are a number of paths for improvements which could increase the performance for others. Additional investigation could improve performance, and would greatly benefit from collaboration with other national agencies with better intrinsic understanding of the drivers of fleet performance across the region.

References

Harley, S., Myers, R., and Dunn, A. (2001). Is catch per unit effort proportional to abundance? *Canadian Journal of Fisheries and Aquatic Sciences*, 58:1760–1772.

Hoyle, S. D., Kim, D. N., Lee, S. I., Matsumoto, T., Satoh, K., and Yeh, Y.-M. (2016). Collaborative study of tropical tuna cpue from multiple indian ocean longline fleets in 2016. IOTC–2016–WPM07–11, , 7th Working Party on Methods, Victoria, Seychelles.

McKechnie, S., Pilling, G., and Hampton, J. (2017a). Stock assessment of bigeye tuna in the western and central Pacific Ocean. WCPFC-SC13-2017/SA-WP-05, Rarotonga, Cook Islands, 9–17 August 2017.

McKechnie, S., Tremblay-Boyer, L., and Harley, S. J. (2015). Analysis of Pacific-wide operational longline CPUE data for bigeye tuna. WCPFC-SC11-2015/SA-WP-03, Pohnpei, Federated States of Micronesia, 5–13 August 2015.

McKechnie, S., Tremblay-Boyer, L., and Pilling, P. (2017b). Background analyses for the 2017 stock assessments of bigeye and yellowfin tuna in the western and central Pacific Ocean. WCPFC-SC13-2017/SA-IP-06, Rarotonga, Cook Islands, 9–17 August 2017.

OFP (2015). Report of the workshop on operational longline data. Technical Report WCPFC-SC11-2015/SA-IP-01, Pohnpei, Federated States of Micronesia, 5–13 August 2015.

Pilling, G. and Brouwer, S. (2017a). Report from the spc pre-assessment workshop, noumea, april 2017. Technical Report WCPFC-SC13-2017/SA-IP-02, Rarotonga, Cook Islands, 9–17 August 2017.

Pilling, G. and Brouwer, S. (2017b). Report of the workshop on analysis of cpue for stock assessments, noumea, april 2017. Technical Report WCPFC-SC13-2017/SA-IP-01, Rarotonga, Cook Islands, 9–17 August 2017.

R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Tremblay-Boyer, L., McKechnie, S., Pilling, G., and Hampton, J. (2017). Stock assessment of yellowfin tuna in the Western and Central Pacific Ocean. WCPFC-SC13-2017/SA-WP-06, Rarotonga, Cook Islands, 9–17 August 2017.

Ward, P. and Hindmarsh, S. (2007). An overview of historical changes in the fishing gear and practices of pelagic longliners, with particular reference to japan's pacific fleet. *Reviews in Fish Biology and Fisheries*, 17(4):501–516.

7 Tables

Table 1: Coefficients for Bigeye

coef.table	Species	Model	Region	X25	X50	X100	X150	X200	X250
1	Bigeye	Lognormal	1	0.74	0.77	0.79	0.79	0.81	0.80
2	Bigeye	Lognormal	2	0.70	0.77	0.78	0.78	0.80	0.80
3	Bigeye	Lognormal	3	0.34	0.50	0.61	0.64	0.65	0.66
4	Bigeye	Lognormal	4	0.45	0.46	0.48	0.50	0.51	0.51
5	Bigeye	Lognormal	5	0.78	0.80	0.78	0.78	0.77	0.79
6	Bigeye	Lognormal	7	0.33	0.33	0.36	0.34	0.36	0.31
7	Bigeye	Lognormal	8	0.14	−0.02	−0.15	0.10	0.01	0.09
8	Bigeye	Binomial	1	0.36	0.41	0.49	0.48	0.44	0.42
9	Bigeye	Binomial	2	0.39	0.47	0.51	0.42	0.36	0.42
10	Bigeye	Binomial	3	−0.13	−0.20	−0.10	0.01	−0.07	−0.01
11	Bigeye	Binomial	4	0.12	0.41	0.39	0.37	0.43	0.34
12	Bigeye	Binomial	5	0.51	0.69	0.62	0.65	0.67	0.70
13	Bigeye	Binomial	7	0.20	0.28	0.35	0.31	0.28	0.25
14	Bigeye	Binomial	8	0.10	0.17	0.10	0.01	−0.17	0.05

Table 2: Coefficients for Yellowfin

coef.table	Species	Model	Region	X25	X50	X100	X150	X200	X250
1	Yellowfin	Lognormal	1	0.82	0.87	0.89	0.88	0.88	0.88
2	Yellowfin	Lognormal	2	0.72	0.77	0.77	0.75	0.76	0.76
3	Yellowfin	Lognormal	3	0.61	0.62	0.61	0.60	0.60	0.60
4	Yellowfin	Lognormal	4	0.13	0.14	0.20	0.21	0.22	0.22
5	Yellowfin	Lognormal	5	−0.34	−0.41	−0.06	0.05	−0.02	−0.17
6	Yellowfin	Lognormal	7	0.61	0.59	0.60	0.62	0.62	0.65
7	Yellowfin	Lognormal	8	0.52	0.82	0.80	0.63	0.27	0.53
8	Yellowfin	Binomial	1	0.70	0.72	0.74	0.75	0.74	0.73
9	Yellowfin	Binomial	2	0.44	0.49	0.49	0.46	0.52	0.48
10	Yellowfin	Binomial	3	0.30	0.35	0.43	0.44	−0.18	0.36
11	Yellowfin	Binomial	4	0.07	0.14	0.15	0.05	0.07	0.09
12	Yellowfin	Binomial	5	−0.24	−0.38	−0.05	0.04	−0.08	0.00
13	Yellowfin	Binomial	7	0.48	0.42	0.46	0.48	0.50	0.48
14	Yellowfin	Binomial	8	0.28	0.85	0.86	0.85	0.83	0.66

8 Figures

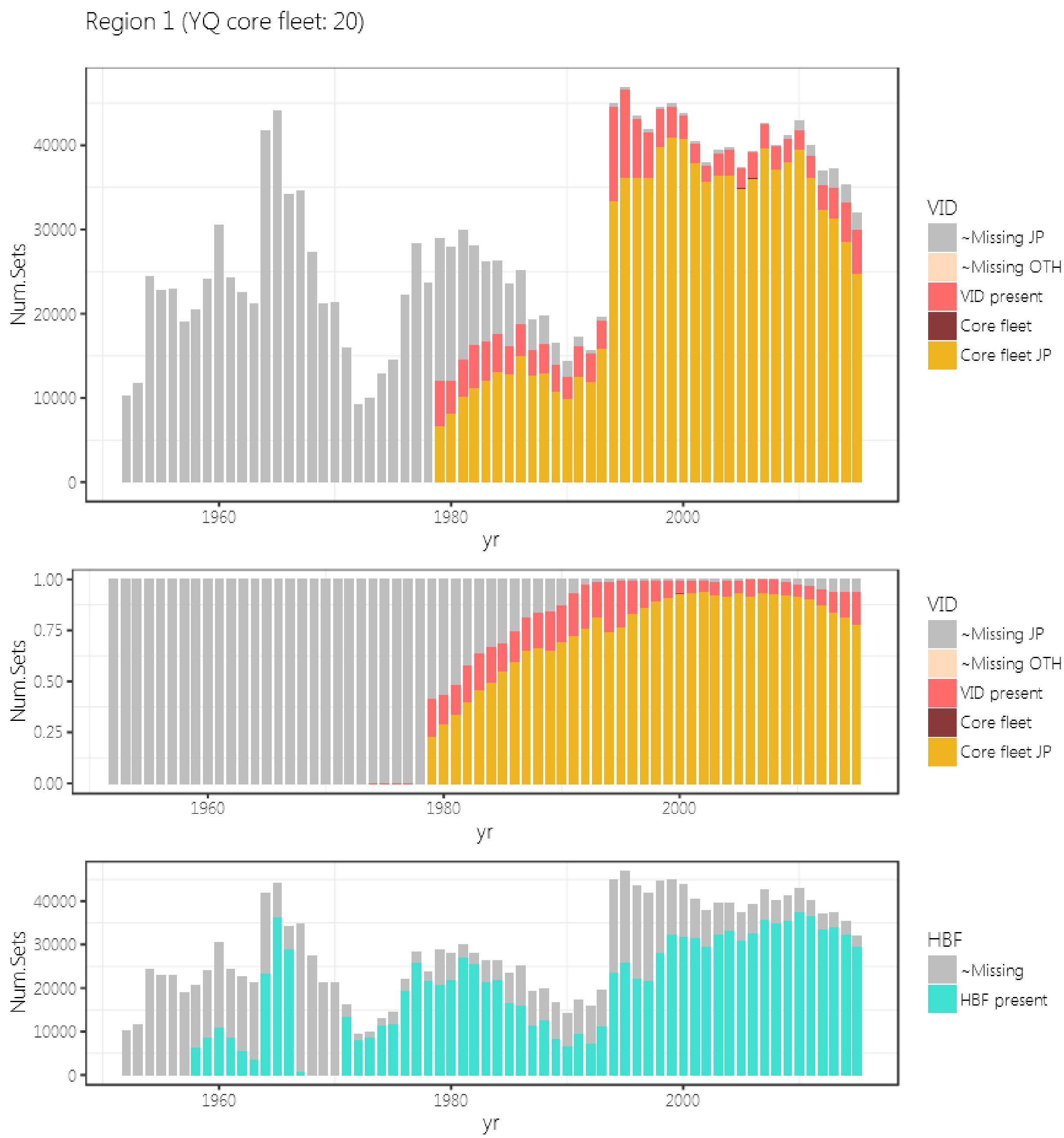

Figure 1: Proportion of sets with missing vessel ID by key fleet category for region 1

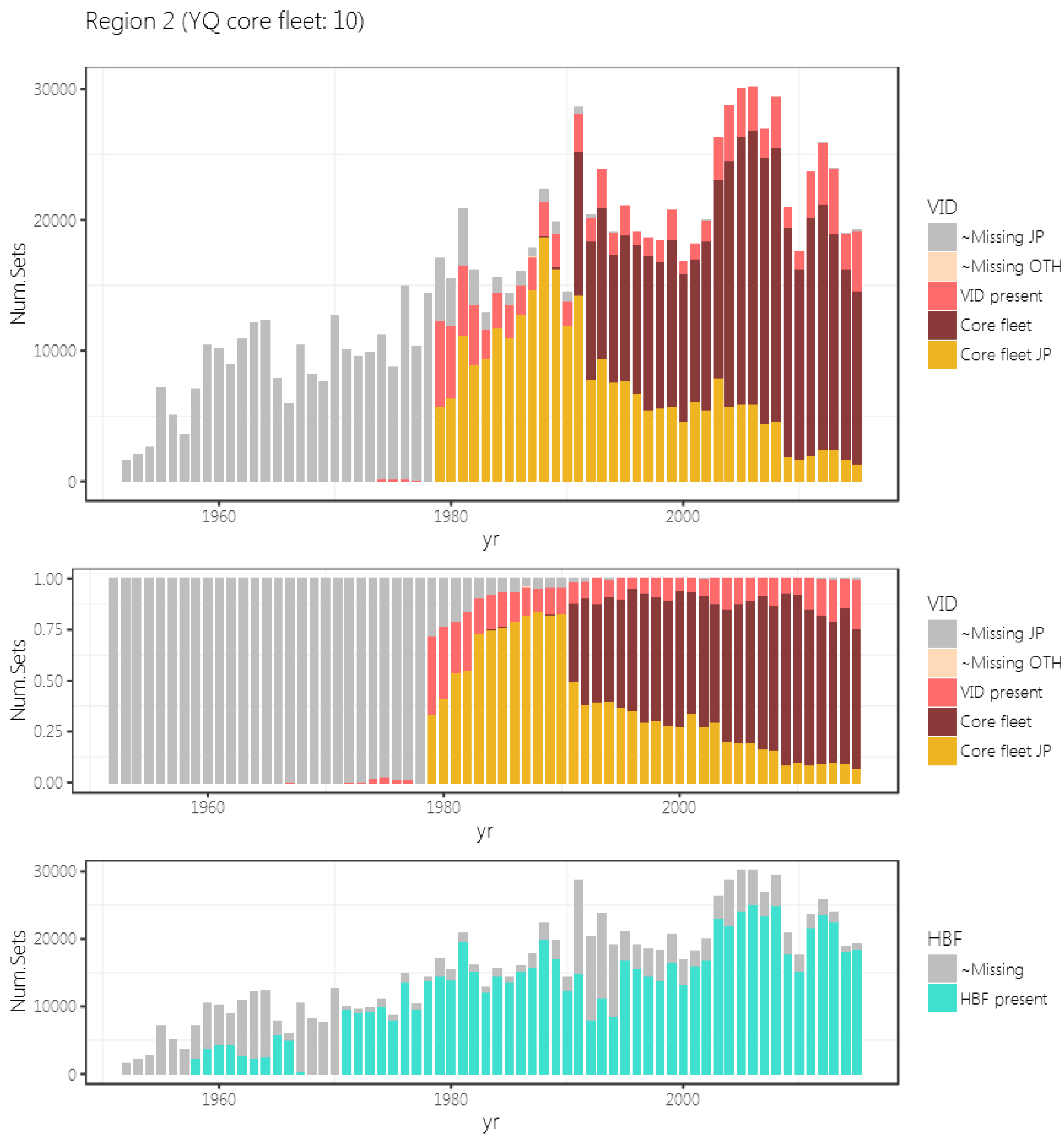

Figure 2: Proportion of sets with missing vessel ID by key fleet category for region 2

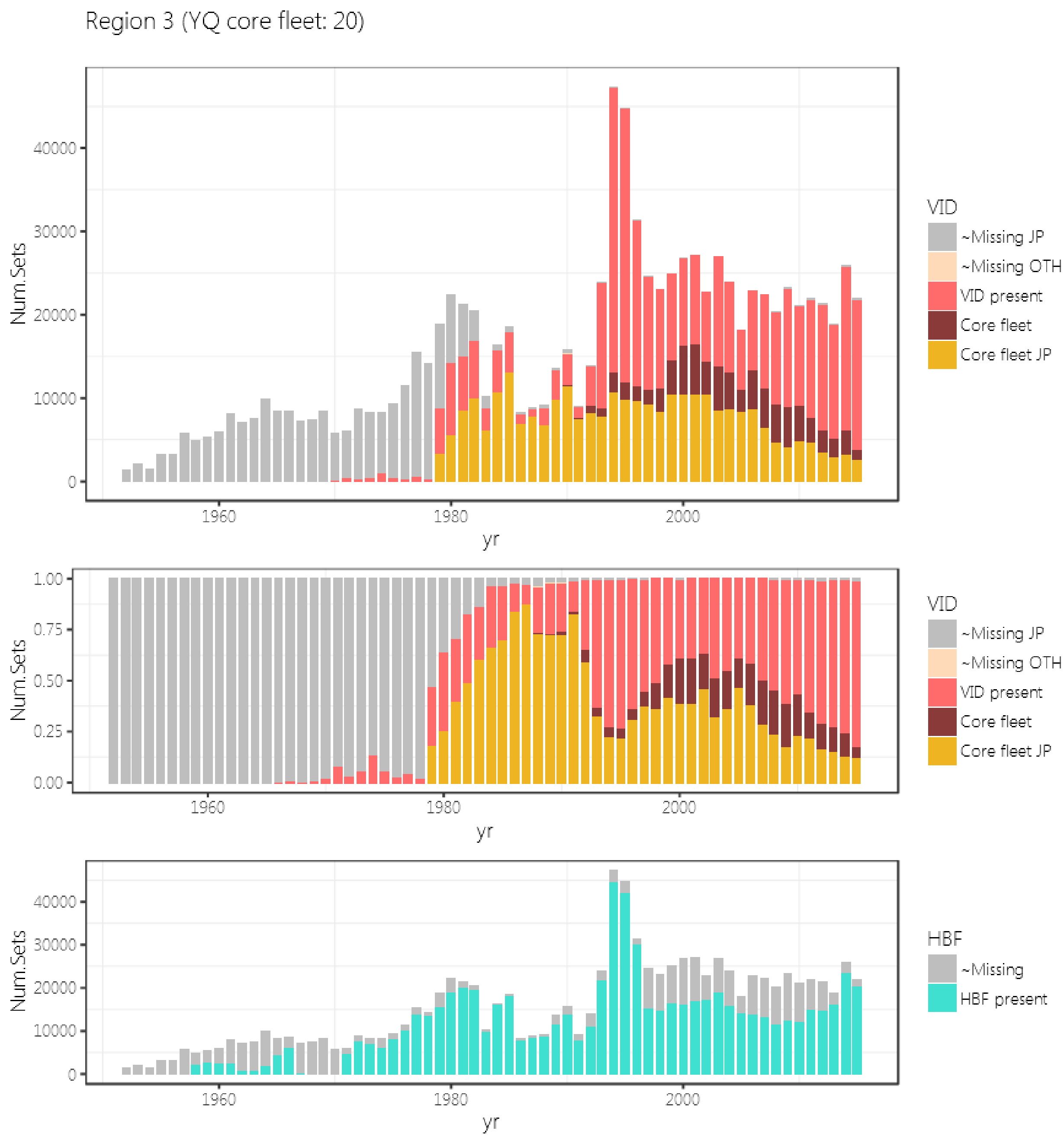

Figure 3: Proportion of sets with missing vessel ID by key fleet category for region 3

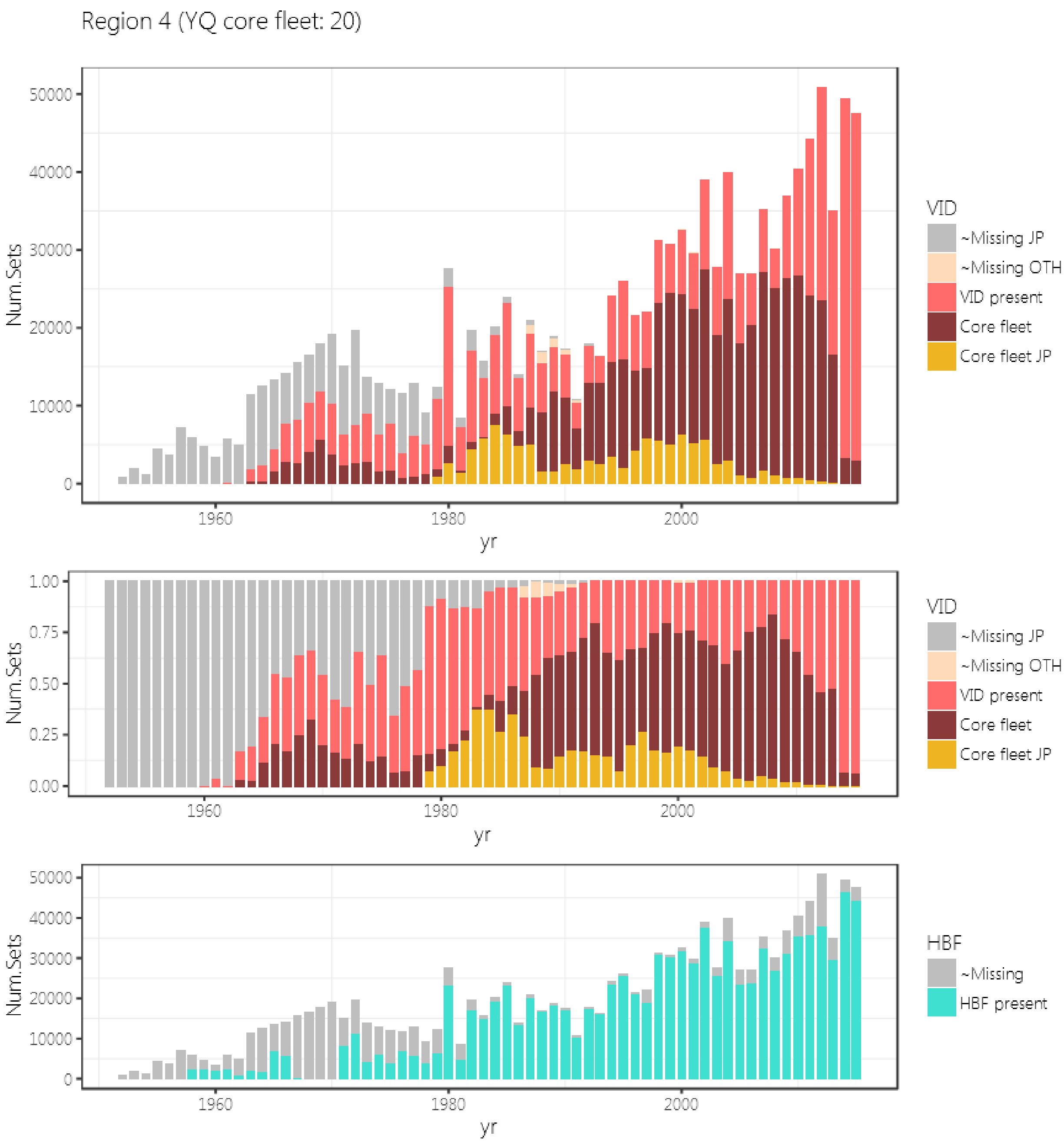

Figure 4: Proportion of sets with missing vessel ID by key fleet category for region 4

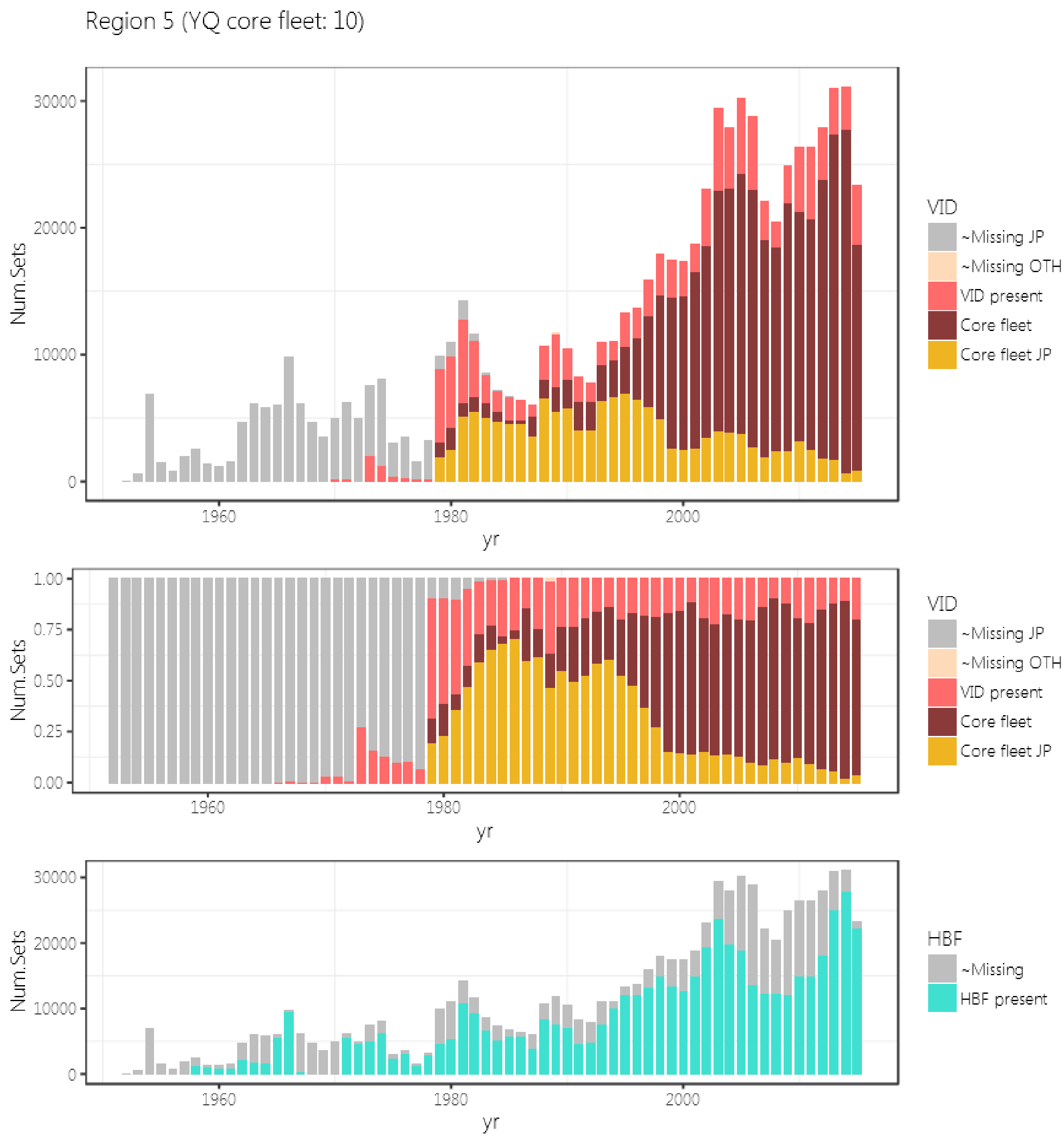

Figure 5: Proportion of sets with missing vessel ID by key fleet category for region 5

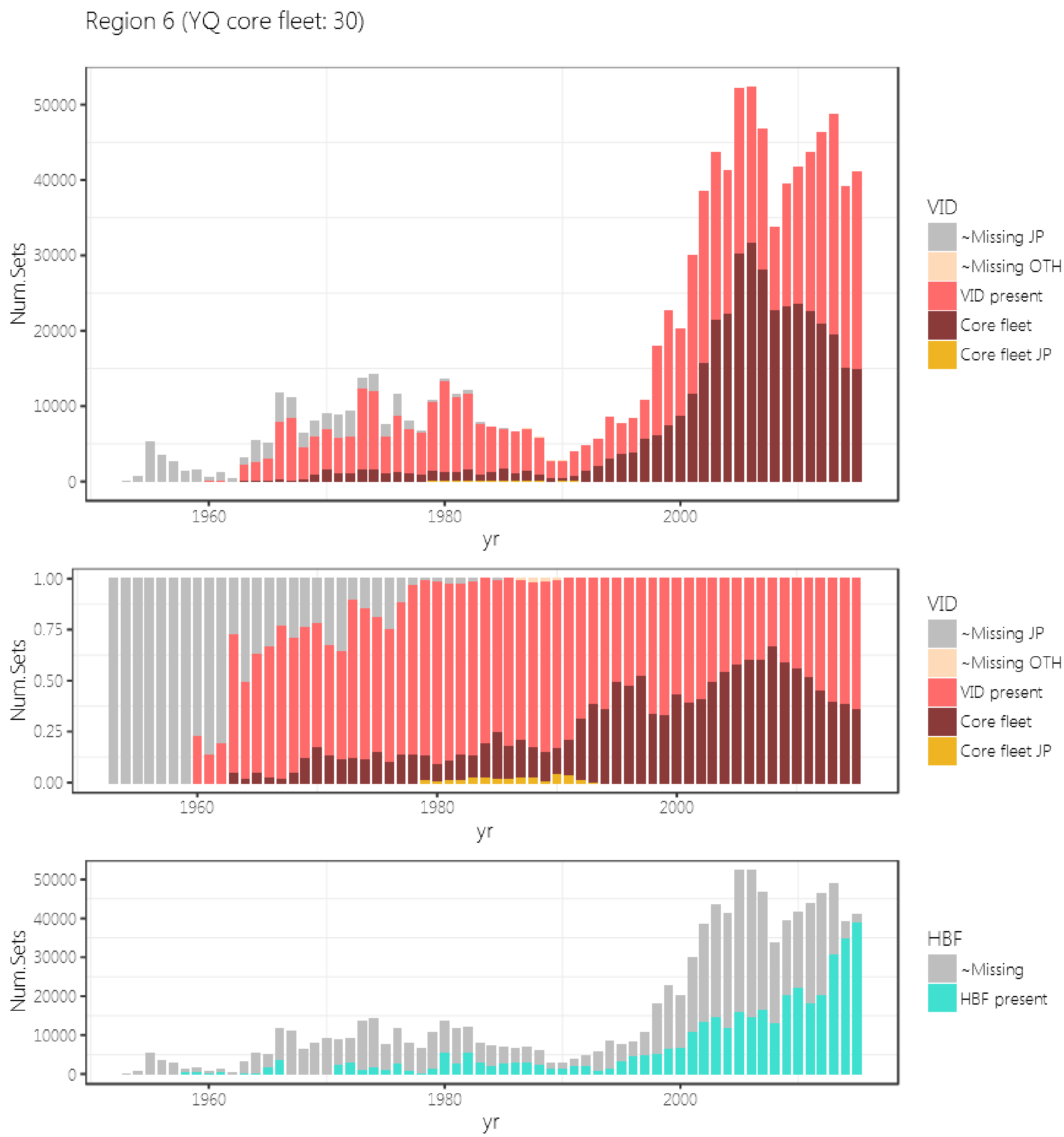

Figure 6: Proportion of sets with missing vessel ID by key fleet category for region 6

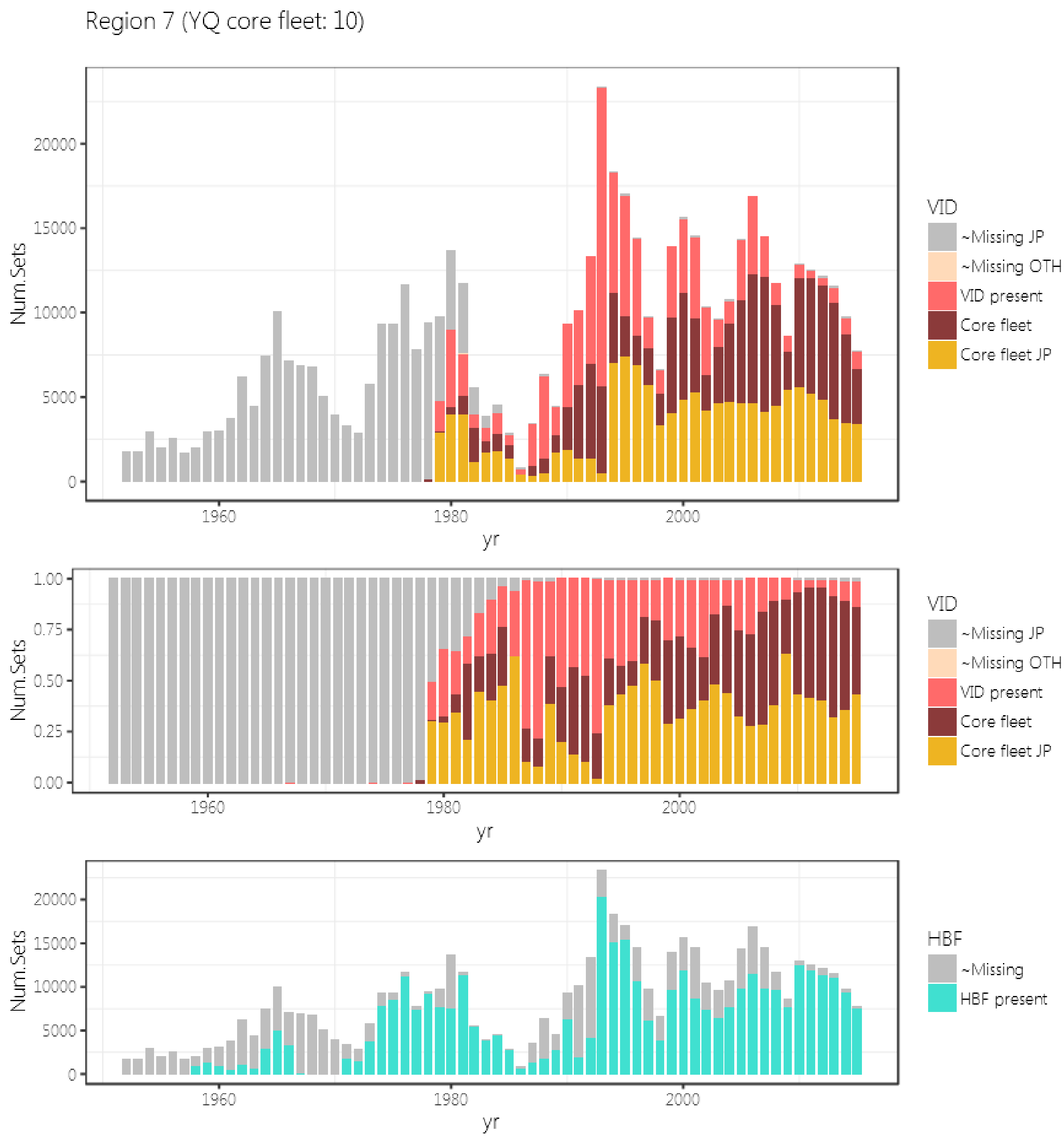

Figure 7: Proportion of sets with missing vessel ID by key fleet category for region 7

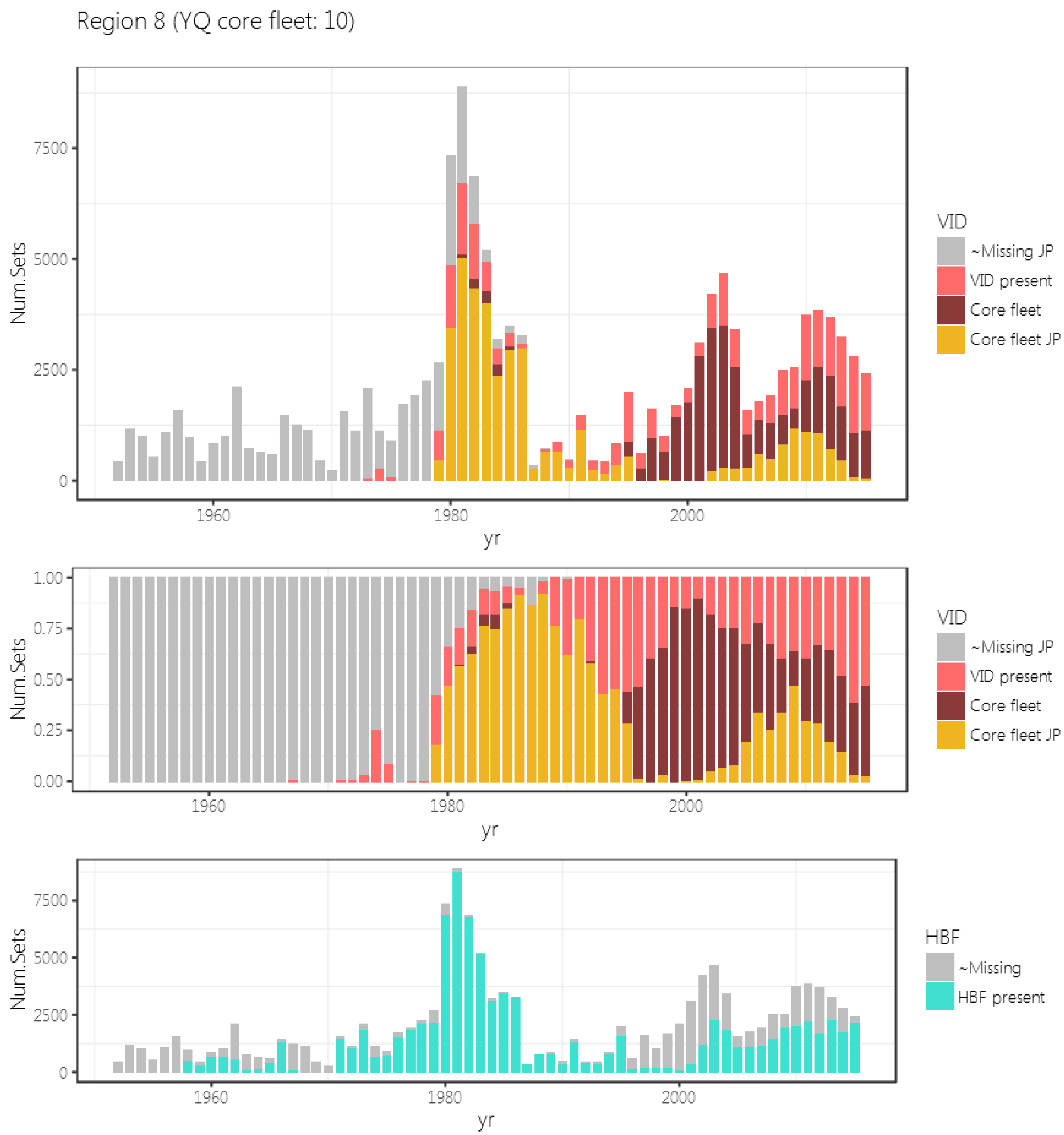

Figure 8: Proportion of sets with missing vessel ID by key fleet category for region 8

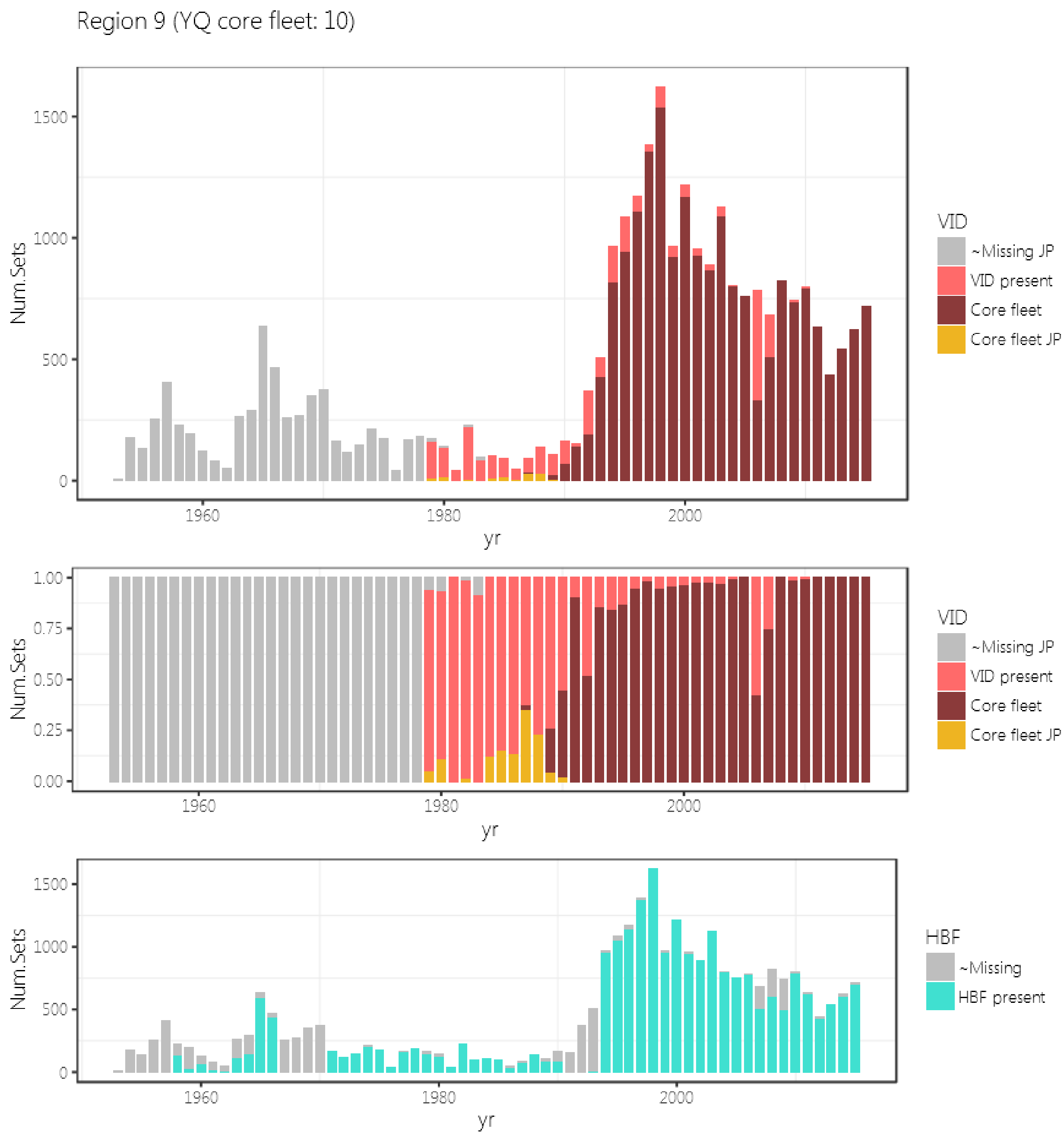

Figure 9: Proportion of sets with missing vessel ID by key fleet category for region 9

Figure 10: The mean value of candidate operational variables (year, hooks by set, hooks-between-floats, latitude, longitude, fleet ID) for the core Japanese fleet against fitted vessel ID lognormal coefficients for bigeye tuna in region 1.

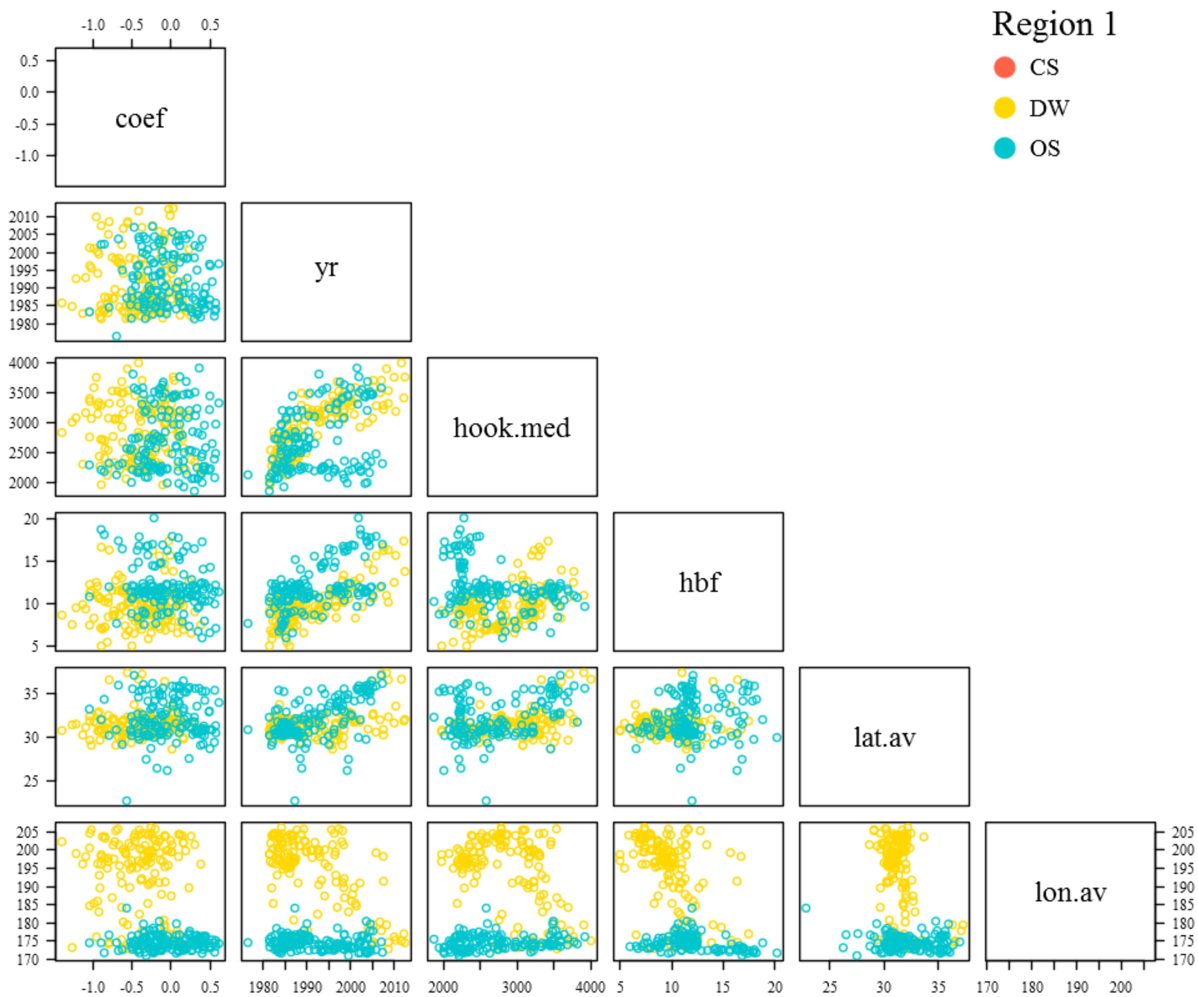

Figure 11: The mean value of candidate operational variables (year, hooks by set, hooks-between-floats, latitude, longitude, fleet ID) for the core Japanese fleet against fitted vessel ID lognormal coefficients for bigeye tuna in region 2.

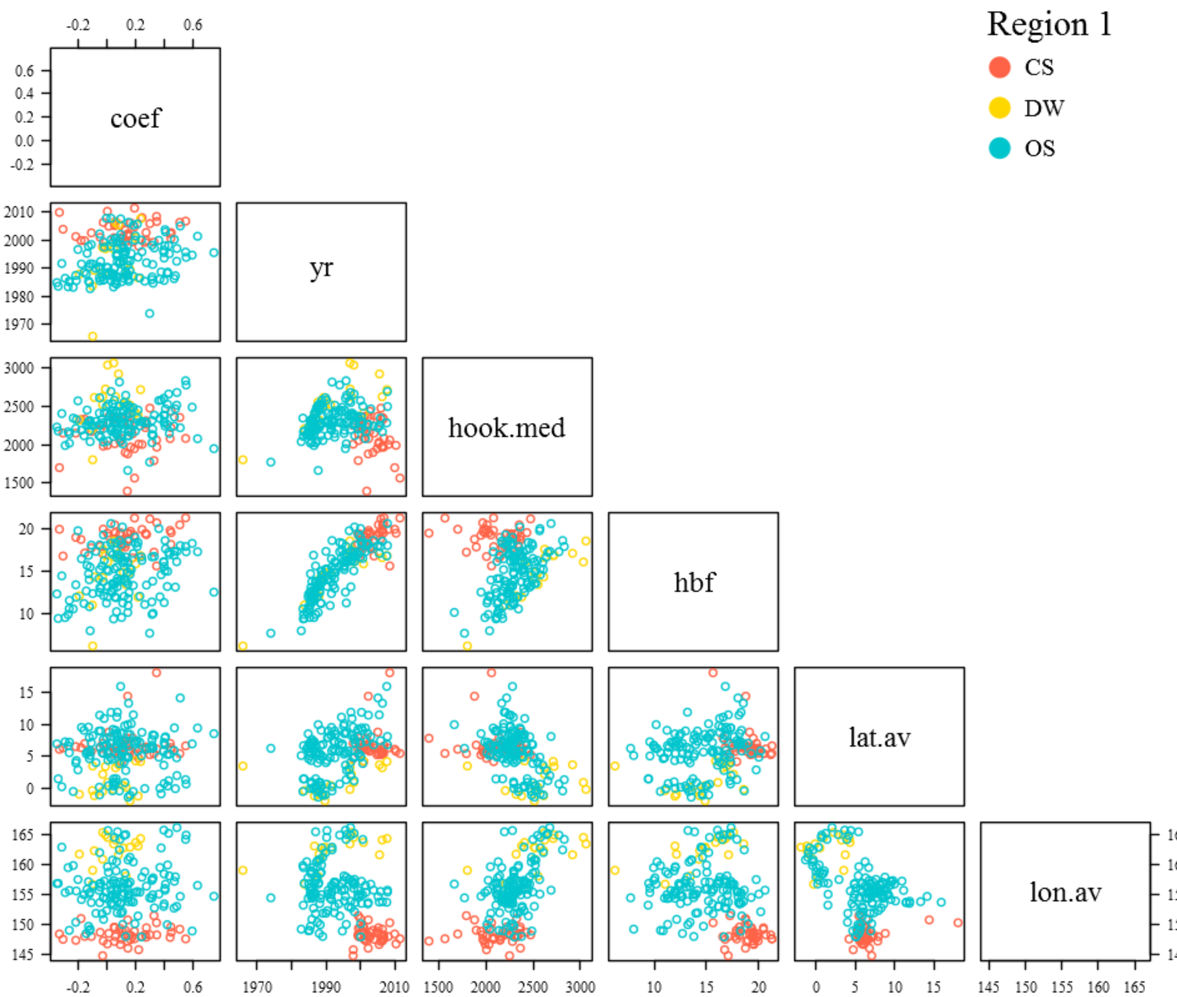

Figure 12: The mean value of candidate operational variables (year, hooks by set, hooks-between-floats, latitude, longitude, fleet ID) for the core Japanese fleet against fitted vessel ID lognormal coefficients for bigeye tuna in region 3.

Figure 13: The mean value of candidate operational variables (year, hooks by set, hooks-between-floats, latitude, longitude, fleet ID) for the core Japanese fleet against fitted vessel ID lognormal coefficients for bigeye tuna in region 4.

Figure 14: The mean value of candidate operational variables (year, hooks by set, hooks-between-floats, latitude, longitude, fleet ID) for the core Japanese fleet against fitted vessel ID lognormal coefficients for bigeye tuna in region 5.

Figure 15: The mean value of candidate operational variables (year, hooks by set, hooks-between-floats, latitude, longitude, fleet ID) for the core Japanese fleet against fitted vessel ID lognormal coefficients for bigeye tuna in region 7.

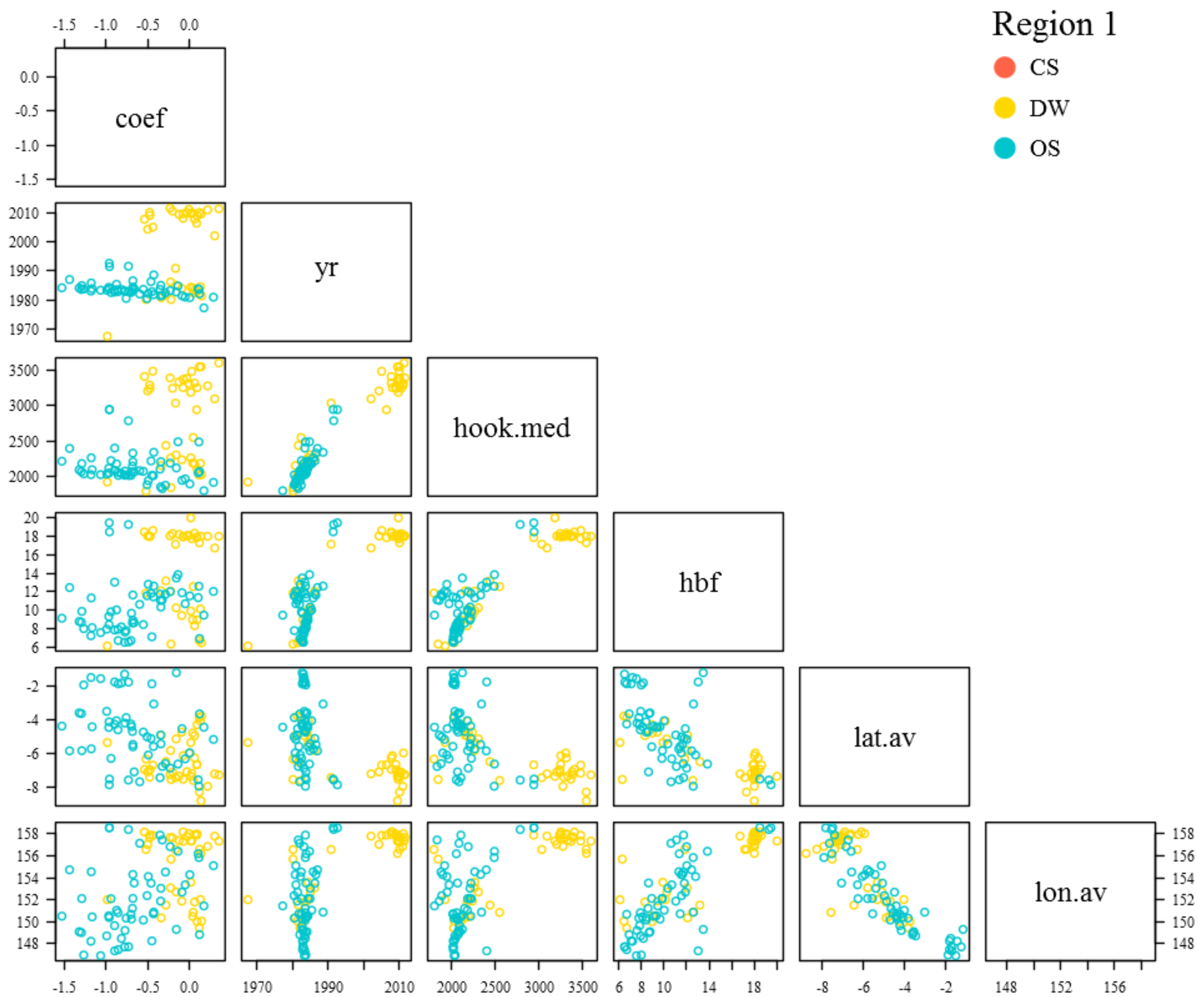

Figure 16: The mean value of candidate operational variables (year, hooks by set, hooks-between-floats, latitude, longitude, fleet ID) for the core Japanese fleet against fitted vessel ID lognormal coefficients for bigeye tuna in region 8.

Figure 17: 'True' vessel effect against proxy vessel effect (weighted average, see 3.4) for the lognormal component of the bigeye GLM standardization model.

Figure 18: 'True' vessel effect against proxy vessel effect (weighted average, see 3.4) for the binomial component of the bigeye GLM standardization model.

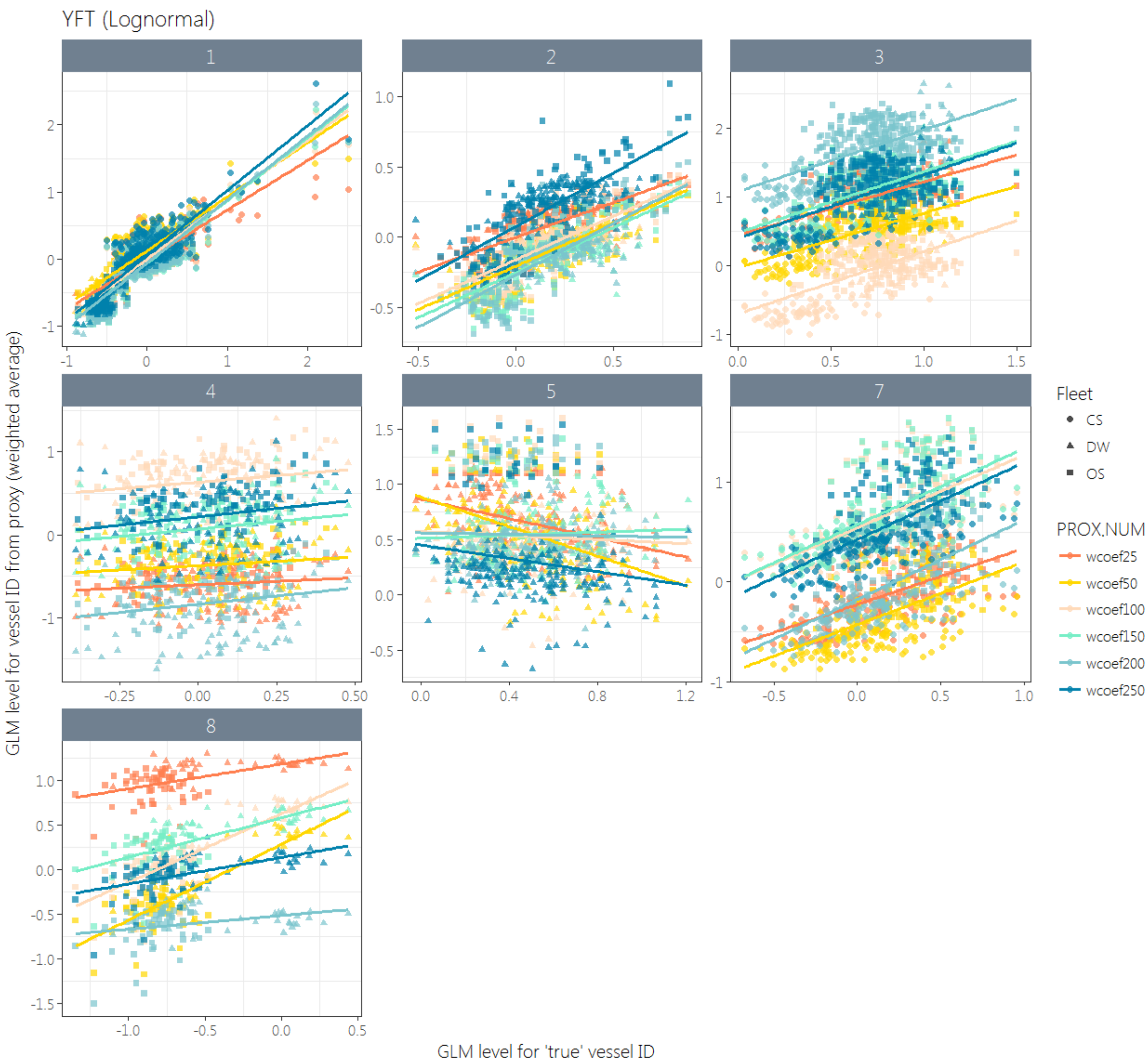

Figure 19: 'True' vessel effect against proxy vessel effect (weighted average, see 3.4) for the lognormal component of the yellowfin GLM standardization model.

Figure 20: 'True' vessel effect against proxy vessel effect (weighted average, see 3.4) for the binomial component of the yellowfin GLM standardization model.

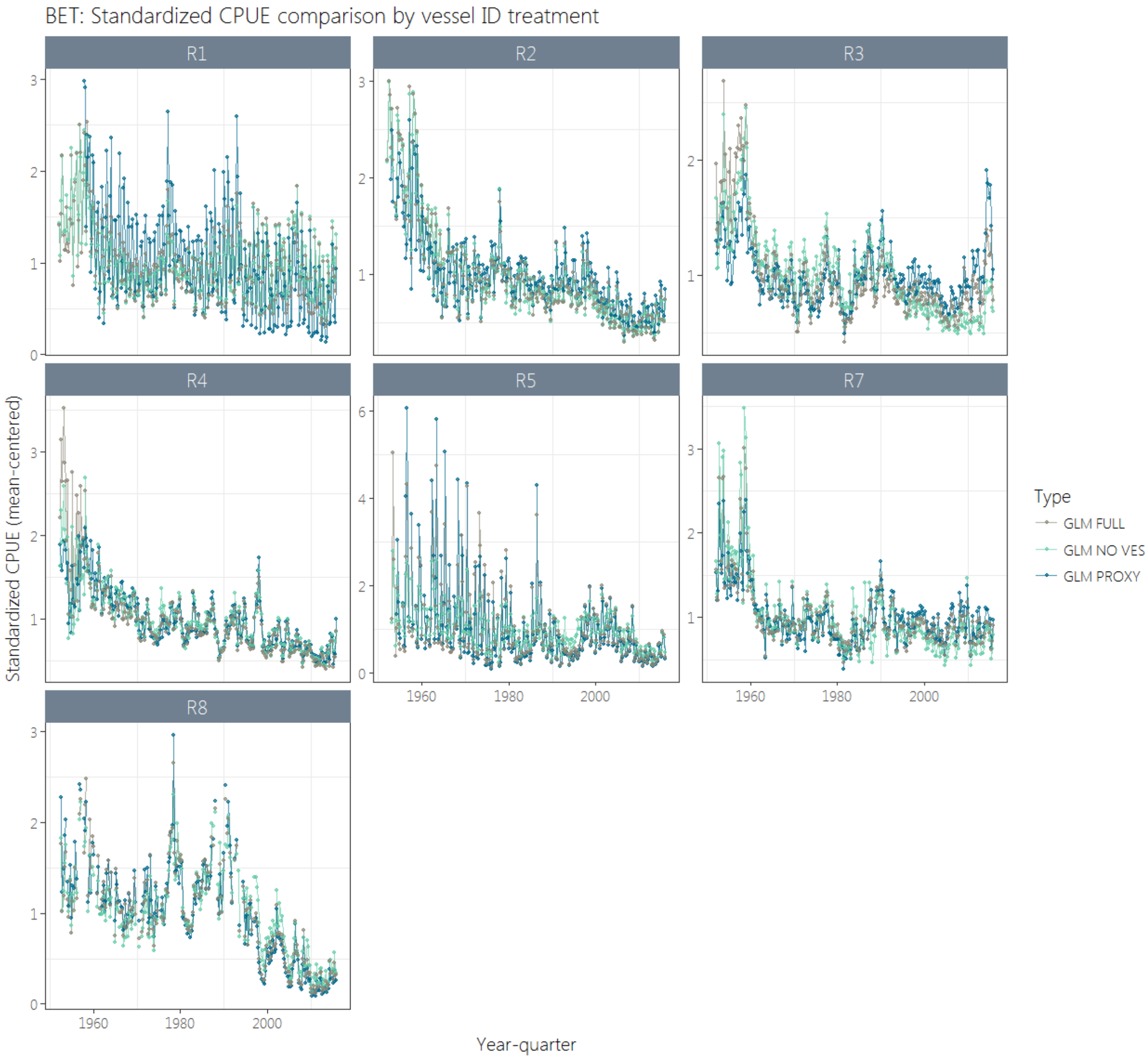

Figure 21: Mean-centered delta-lognomral standardized indices for bigeye tuna under different vessel ID treatment: no vessel effects, vessel effect with generic missing vessel level, hybrid vessel ID/proxy approach with the core fleet with vessel ID and missing sets assigned a vessel proxy.

Figure 22: Mean-centered delta-lognomral standardized indices for yellowfin tuna under different vessel ID treatment: no vessel effects, vessel effect with generic missing vessel level, hybrid vessel ID/proxy approach with the core fleet with vessel ID and missing sets assigned a vessel proxy.