CCSBT-ESC/2308/32
(ESC Agenda item 7.2)

Mitigating the effects of increasing effort concentration by developing indices based on data from multiple fleets

28th Extended Scientific Committee Meeting, CCSBT,
28 August – 2 September 2023

Simon Hoyle
Hoyle Consulting, simon.hoyle@gmail.com

Mitigating the effects of increasing effort concentration by developing indices based on data from multiple fleets.

Simon Hoyle

Hoyle Consulting, simon.hoyle@gmail.com.

Working Paper prepared for the 28th Extended Scientific Committee Meeting, CCSBT, 28 August – 2 September 2023.

Abstract

This paper reports work to explore the potential to develop CPUE indices for SBT based on data for multiple fleets in addition to Japan. The analyses use generalized additive models (GAMs) with spatiotemporal smoothers, and a delta lognormal approach. The temporal and spatial distributions of both fishing effort and the highest catch rates have changed between 1986 and 2022, while the spatial and temporal extents of fishing effort have declined. Simulated data were generated from the best models fitted to the aggregated dataset and used to explore the effectiveness of different model configurations for dealing with these changing distributions. The principal GAM models produced unbiased estimates with the simulated data, while GLM models and less flexible GAM smoothers provided biased indices. Manipulating the simulated dataset to produce a large rapid change in fish distribution resulted in moderately biased indices. Increasing the effort concentration through time to focus effort on areas with higher CPUE also resulted in estimation bias, particularly at the end of the time series when concentration was greatest. This bias may be due to loss of information from the dataset rather than model failure, and it may be helpful to increase the information via models that include data from other fleets as well as Japan. In general, GAM models provided less biased indices than either a GAM equivalent to the variable squares method (GAM_VS) or a combined model (w0.8) approach.

Introduction

The CPUE standardization methods used for SBT have been updated to address problems with recent CPUE estimates, particularly an anomalously high value in 2018 (CCSBT, 2020). The main reason for these analytical problems was identified as increasing aggregation of fishing effort, together with a method that relied on data availability in all strata. Sparse data caused parameter estimation problems (ESC 25, para 37). Analyses between 2020 and 2022 (Hoyle, 2021; Hoyle, 2022; Hoyle, 2020) developed an alternative approach using generalized additive models (GAMs) implemented with the R package *mgcv* (Wood, 2011). Data were fitted with multi-dimensional smoothers which share information among adjacent values of continuous variables. Further analysis (Hoyle, 2022) indicated that the principal GAM models produced unbiased estimates with the simulated data, while GLM models and less flexible GAM smoothers provided biased indices, particularly at the end of the time series as effort became more concentrated, and data became sparse.

However, manipulating the simulated dataset to produce a large rapid change in fish distribution resulted in moderately biased indices from all models. Increasing the effort concentration through time to focus effort on areas with higher CPUE also resulted in estimation bias, particularly at the end of the time series when concentration was greatest. This bias may be due to loss of information from the dataset rather than model failure, and ESC 27 concluded that it may be helpful to increase the available information via models that include data from other fleets as well as from Japan.

Work for 2023 involved exploring the spatio temporal effort distributions of fleets other than Japan, to help understand whether they might usefully contribute to maintaining through time the coverage of the SBT population distribution, and thereby reducing the risk of parameter estimation difficulties. This involved plotting the distribution of data from these fleets and exploring patterns of effort concentration among all fleets.

Methods

Input data

These analyses were based on two datasets, files 'CPUEInputs_2023_June.txt' (CPUEinputs) and 'CatchEffort_2023_June.txt' (CatchEffort), available from the private area of the CCSBT website. These data are aggregated by year, month, and 5° latitude and longitude. In the CPUEinputs file, catches are reported by age class based on spatially and temporally stratified size sampling, whereas the CatchEffort file reports only numbers and weights.

The following processes were then applied to the CatchEffort_2023_June.txt dataset:

- Filter effort as follows:
    - Year > 1985.
    - Statistical areas 4 to 9.
    - Months 4 to 9.
    - latitudes north of 50° S.
    - gear code = 'LL'.
    - dataset_code is not 'JP_RTMP'.
    - fleet_code is not NZA or NZP.
- Data cleaning
    - filter out one misplaced stratum.
    - CPUE is not missing.
    - CPUE < 120 (outlier).

- Create categorical areaf variable, which merges statistical area 4 with 5 and statistical area 6 with 7.
- Adjust numeric longitude variable (lon) by adding 360 to all longitudes between -180 and -100, to provide continuity across the spatial domain of the fishery. Longitudes are recorded as -180 to 180 and so the range of the adjusted longitude variable was from -95 to 260.
- Create categorical variables yf, latf, lonf, and mf, for year, latitude, longitude, and month.
- Create categorical llf variable, indicating 5° square that combines latitude and longitude.
- Create numeric catch variable, the sum of NUMBER_RETAINED and NUMBER_DISCARDED.
- Create numeric cpue variable = catch per 1000 hooks.

For these analyses, data preparation did not remove strata with fewer than 10000 sets, except where stated otherwise. This was to avoid complications due to different levels of filtering depending on how many fleets were fishing in a stratum.

Aggregated datasets differ from the Japanese operational dataset used for primary analyses. The primary dataset is only available to Japanese scientists. The main differences between these datasets are listed below.

- The primary dataset is available as operational (set by set) data (but may be aggregated for the main analysis) whereas the available dataset is aggregated by month and 5-degree cell.
- The primary dataset uses a set of core vessels that have high SBT catches for at least 3 years, whereas the available dataset includes data from all vessels.
- The primary dataset includes catches of bigeye and yellowfin tuna, but the available dataset does not.

Data exploration

Several approaches were used to explore changes in effort distribution and concentration through time. CPUE estimates are affected by effort concentration both spatially and by month within years, since data are stratified by 5-degree cell and month. We therefore explored changes in both the numbers of strata fished and the number of operations per fished stratum, and their variation through time both within individual fleets and within datasets combined across fleets.

Maps

In order to explore how effort concentration has changed through time by fleet, a map series showing data coverage through time was generated for each individual fleet code, for each year between 1986 and 2022 in which data were available. These were annual plots of the number of month x spatial cell strata fished. An additional map series was generated that compared data coverage between Japanese effort and total effort for all fleets.

The 252 maps are not included in this paper, but available for download from the github repository.

Concentration indices

Indices of fishing effort concentration were also calculated, including the Gini coefficient (Gini, 1912) and Gulland's index of concentration (Gulland, 1956). The Gini coefficient is best known as an indicator of wealth concentration but can be used to measure aggregation of any quantity. We use it to estimate the spatial aggregation of the catch of each species, and effort, in each region. A higher Gini coefficient indicates that more of the catch (or effort) is being taken from fewer spatial cells. We estimated values separately for each year, where the values y_i are catches or effort per 5° x 5° cell, ranked from lowest to highest, and including zeroes for unfished cells. Cell areas are assumed to be uniform.

$$Gini = \frac{2\sum_{i=1}^{n} i y_i}{n\sum_{i=1}^{n} y_i} - \frac{n+1}{n}$$

Gulland's index of concentration measures the extent to which a fleet has concentrated its fishing effort in areas with higher than average catch rates. The weighted version of the index is calculated as follows, where y_i is the catch in the ith stratum, e_i is the effort in the ith stratum, and N is the number of exploited strata.

$$Gulland = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} e_i} \cdot \frac{1}{\sum_{i=1}^{N} \frac{y_i}{e_i N}}$$

This index varies from year to year depending on both the distribution of the effort, and the distribution of the catch rates. If effort is evenly distributed with respect to catch rate then the index will average 1, whereas it will be higher than 1 if effort is preferentially targeted to areas with higher than average catch rate (Hoyle, 2014).

CPUE standardization

Using the variables and interactions previously selected (Hoyle, 2022), models were run using the delta lognormal approach using the R package *mgcv*. This package uses the offered terms and initial basis dimension (*k*) as a starting point for a search. The *k* parameter sets the upper limit on the degrees of freedom associated with a single smooth, while for a tensor product smooth the upper limit is the product of the *k* values for each marginal smooth.

Delta models were fitted using the function gam() and restricted maximum likelihood (REML), and positive models were fitted using the function bam() and generalized cross-validation (GCV).

Wood (2011) recommends that models with multiple levels of interactions should specify main effects using either *s()* or *ti()* and interaction terms with *ti()*. Models were fitted using *ti()* for all terms, so that all terms were fitted using the default cubic splines.

Binomial: cpue != 0 ~ yf + ti(lon2, k=40) + ti(LAT, k=4) + ti(MONTH, k = 6) + ti(lon2, LAT, k = c(40,4)) + ti(MONTH, LAT, k = c(6,4)) + ti(lon2, MONTH, k = c(10, 5)) + ti(YEAR, LAT, k = c(20, 4)) + ti(YEAR, MONTH, k = c(20, 5)) + ti(lon2, YEAR, k =c(10, 9)) + ti(LAT, lon2, MONTH, k = c(4,15, 6)) + ti(LAT, lon2, YEAR, k = c(4,10, 9)) + ti(log(N_HOOKS), k = 10)

Lognormal: log(cpue)= 0 ~ yf + ti(lon2, k=40) + ti(LAT, k=4) + ti(MONTH, k = 6) + ti(lon2, LAT, k = c(40,4)) + ti(MONTH, LAT, k = c(6,4)) + ti(lon2, MONTH, k = c(10, 5)) + ti(YEAR, LAT, k = c(20, 4)) + ti(YEAR, MONTH, k = c(20, 5)) + ti(lon2, YEAR, k =c(10, 9)) + ti(LAT, lon2, MONTH, k = c(4,15, 6)) + ti(LAT, lon2, YEAR, k = c(4,10, 9))

As a final step the model was specified using a 'shrinkage' version of the cubic spline smooth (bs = "cs"), which if warranted will penalise a curve to zero and effectively eliminate it from the model.

The lognormal model used *log(cpue)* as the response, with identity link and gaussian error distribution, while the binomial model used '*cpue* != 0' as the response, with logit link function. Binomial models included effort in the formula to account for the effect of effort on the probability of non-zero catch in a stratum. Effort was included as a spline rather than a straight line or offset, to allow for potential nonlinearity in the relationship.

Table 1: Settings used in mgcv to compare models with different distributions.

Distribution	Family	Dataset	response	Link function	Likelihood
Binomial (DLN)	Binomial	all	cpue > 0	logit	REML
Lognormal (DLN)	Gaussian	nonzero	log(cpue)	identity	GCV

delta <- gam(cpue > 0 ~ formula, data = a, gamma = 2, method = 'REML', family = binomial)

pos <- bam(log(cpue) ~formula, data = apos, gamma = 2, method = 'GCV')

Preparation of indices

The density of SBT in each stratum (year by month by grid cell) was predicted and stratum abundances predicted by multiplying ocean area (Hoyle and Langley, 2020) stratum density.

A time series of predicted abundance was calculated by summing predicted abundances across strata for each year. Abundance indices were obtained by dividing the abundances for each year by the mean of all years.

Spatial cells were included in the locations used for prediction if the original dataset contained at least 15 nonzero effort observations (strata) associated with the cell.

Several alternate methods were used to generate indices.

- CV_limit indices were prepared by including predictions only for strata for which the coefficient of variation (CV) of the predicted density was less than a predefined limit.
- Fished_strata: indices were developed by predicting only for strata (5-degree cell by month) that were fished at least once during the period 1986 to 2022. Given the limit of 10000 hooks for including strata in the dataset, this implies at least quarter and month with more than 3 sets.
- Fished_strata_nyears: indices were developed by predicting only for strata (5-degree cell by month) that were fished at least once during a predefined period, for a range of different periods. Periods used were 2017, 2018, 2019, 2020, 2021, and 2017-2022.

To check the utility of conclusions based on the CatchEffort_2023_June.txt dataset, indices were developed using both the CPUEinputs_2023_June.txt and the CatchEffort_2023_June.txt datasets and compared on the same plot.

To check the influence of omitting strata with fewer than 10000 hooks, indices were developed using the CPUEinputs_2023_June.txt dataset both with and without these strata and compared on the same plot.

To check the influence of sharing catchability between the main Japanese fleet and the Japanese charter fleets fishing in Australian and New Zealand waters (fleet codes AUC and NZC), models were run with both the standard approach and with separate fleet effects estimated.

Indices for combined fleets

Indices for combined fleets were developed by fitting models as follows, which estimates independent catchability by fleet in both model components. For these analyses, Japanese charter fleets were assigned the same fleet code as the Japanese domestic fleet, so as to assume consistent catchability across the entire Japanese fleet.

Binomial: cpue != 0 ~ yf + FLEET_CODE + ti(lon2, k=40) + ti(LAT, k=4) + ti(MONTH, k = 6) + ti(lon2, LAT, k = c(40,4)) + ti(MONTH, LAT, k = c(6,4)) + ti(lon2, MONTH, k = c(10, 5)) + ti(YEAR, LAT, k = c(20, 4)) + ti(YEAR, MONTH, k = c(20, 5)) + ti(lon2, YEAR, k =c(10, 9)) + ti(LAT, lon2, MONTH, k = c(4,15, 6)) + ti(LAT, lon2, YEAR, k = c(4,10, 9)) + ti(log(N_HOOKS), k = 10)

Lognormal: log(cpue)= 0 ~ yf + FLEET_CODE + ti(lon2, k=40) + ti(LAT, k=4) + ti(MONTH, k = 6) + ti(lon2, LAT, k = c(40,4)) + ti(MONTH, LAT, k = c(6,4)) + ti(lon2, MONTH, k = c(10, 5)) + ti(YEAR, LAT, k = c(20, 4)) + ti(YEAR, MONTH, k = c(20, 5)) + ti(lon2, YEAR, k =c(10, 9)) + ti(LAT, lon2, MONTH, k = c(4,15, 6)) + ti(LAT, lon2, YEAR, k = c(4,10, 9))

First, indices were developed by combining Japanese data with one fleet at a time, for each of the Australian, Korean, New Zealand, South African, and Taiwanese fleets. Next, indices were developed after combining data from all fleets. Then, indices were developed by combining Japanese, Australian, Korean, and New Zealand data.

CPUE by year and fleet
To explore the nominal catch rates by year and fleet, and factors affecting recent catch rates, histograms of CPUE were plotted by year and fleet.

Nominal CPUE by fleet (without combining with Japanese data) was estimated by fitting the simple model cpue ~ year_factor.

A simplified version of standardized CPUE was estimated by fitting a delta lognormal model as follows, with all variables included as factors.
mod_delta <- gam(cpue != 0 ~ year + month + cell)

mod_pos <- gam(log(cpue) ~ year + month + cell)

R code
All R code used in and developed for this study is available at the github repository https://github.com/hoyles/R_ccsbt_cpue. Please email simon.hoyle@gmail.com to request permission to access the repository.

Results

Patterns in the Japanese effort
Plots of patterns in the Japanese effort by latitude (Figure 1) show that both the number of fished strata and total effort have in general declined through time in all latitudes, but with considerable temporal variation. Effort in the southernmost latitude centred on 47.5 S has declined almost to zero, with occasionally a small amount of fishing in a few strata. Effort in the northernmost stratum centred on 32.5 S has been comparatively stable. Most of the effort in (to a lesser extent) the majority of fished strata has consistently occurred in the central latitudes centred on 37.5 S and 42.5 S. Relative effort in these strata has oscillated. The period 2017 - 2020 saw very high effort at 42.5 S, but by 2022 some effort had shifted north so that 37.5 S and 42.5 S had similar levels of effort.

In the 1990s, apportionment of effort by month (Figure 2, lower right) showed highest effort in May and June, followed by July. From 2010 - 2020, May had the most effort followed by April and June. The most recent year (2022) saw an unusually high proportion of effort occurring in August, and a very low proportion in April. Apportionment of fished strata by month was more stable through time than effort (Figure 2, upper right). The most obvious changes were a steady reduction in the proportion of strata fished in August, and increases in the proportions of strata fished in May, June, and particularly April.

When viewed in terms of statistical areas (Figure 3), area 9 has always seen the largest number and the highest proportion of strata fished. This has not been the case for effort however, with similar amounts of effort allocated to area 8 since about 2009. The year 2022 saw approximately twice as much effort in area 8 as in area 9. Significant Japanese effort also remains in area 7. However, effort in areas 4, 5, and 6 has been minimal since 2015, although reasonable numbers of strata are still fished except in area 6.

Comparisons by fleet

Operations per stratum by flag, statistical area and year were plotted to compare changes in effort concentration through time among flags (Figure 4). Australia shows relatively stable effort through time, mostly in areas 4 and 5. There was decline in Japanese operations per stratum until 2010, with more stability since then, and considerable variability among statistical areas. Korean effort has shown a high degree of increasing concentration into area 9, with increasing numbers of operations per stratum. New Zealand data show relatively stable operations per stratum since about 2005. Taiwanese effort has somewhat increased the numbers of operations per stratum in both areas 8 and 9. South African operations per stratum have declined since 2010, and no data are available since 2019.

The number of fished strata per year has declined steadily from a very high level for the Japanese fleet (Figure 5). Fished strata have also declined for the Korean fleet. The Australian and New Zealand fleets have been relatively stable in recent years, and the Taiwanese fleet increased its number of fished strata from 2003 until around 2015. Combining all fleets results in much more stable coverage through time than coverage for just the Japanese fleet, with considerably higher numbers of strata (Figure 6). When data are cleaned so that there are at least 10000 hooks per stratum, the number of strata in the Japanese dataset declines from a high of 176 in 1998 to a low of 56 strata in 2022.

The comparison by fleet is made more obvious when directly comparing the Japanese coverage with the improvement in coverage when combined with effort from another flag (Figure 7). To aid comparison with results in Japanese and Korean documents, Figures 8 and 9 replicate the approach in those documents to show both the number of fished cells per statistical area and year and the number of operations per cell per year on the same plot.

Indices of concentration

Gini coefficients by month for Japanese effort (Figure 10) showed a general pattern of increasing concentration through time in all months. Gini coefficients for all fleets' effort (Figure 11) showed similar trends in months 8 and 9 but were more stable in for months 4 to 7. In addition, the average level of the coefficients was lower, indicating that effort was in general less concentrated. The equivalent figures for SBT catch showed similar relationships between Japanese CPUE and CPUE of all fleets (Figures 12 and 13), although the differences between Japan and all fleets for CPUE were smaller than they were for effort.

Similarly comparing Gulland's indices of concentration for Japanese catch and effort (Figure 14) with those for all fleets' catch and effort (Figure 15) showed that Japanese effort was somewhat more concentrated into areas of higher CPUE than was the effort of all fleets combined. This is not surprising, since the effort of all fleets includes effort from the Taiwanese fleet which targets species other than SBT and reports significantly lower SBT catch rates.

CPUE indices

CPUE indices were developed for the combination of Japanese data and data from each other fleet (Figure 16). In general, the indices were similar to the estimated Japanese indices, for most of the time series. Divergence increased through time as effort became more concentrated and other fleets comprised a higher proportion of effort. The largest differences occurred for indices that included Taiwanese data and South African data.

Additional CPUE indices were developed for datasets that a) combined all fleets, and b) combined all except the Taiwanese and South African fleets. Both diverged from the Japanese index primarily in the period since 2010. The greatest divergence was for the all-fleets dataset, which reached a very high level for the 2022 index.

The array of strata used for predicting indices had relatively little impact on the CPUE index time series (Figure 18). The greatest differences between sets of predicted indices were for the 2015 and 2022 predictions. In 2022 the all-strata prediction was lower than any of the fished strata indices, whereas for 2022 the all-strata prediction was higher than any of the fished strata indices.

Frequency histograms of CPUE by year and fleet (Figure 19) did not suggest any inconsistencies between the high index estimate in 2022 and the observed catch rates. Catch rates in 2022 were notably high in data from Japan and New Zealand, but not for any of the other fleets. The uniformly low catch rates in the Taiwanese data may have obscured any temporal patterns in that dataset.

Nominal CPUE indices by fleet (Figure 20) and standardized indices using a simplified GLM model (Figure 21) were similarly consistent with the GAM standardized indices. Strong increases were observed since 2010 in all indices, with the 2022 index particularly high for Japan and New Zealand.

Indices based on the CV_limit method (Figure 22) included predictions only for strata where the CV of the predicted density was less than a predefined limit. Application of this limit tended to change the trend, because uncertainty was greater in later parts of the time series. Since strata that were more uncertain were not summed to form the index, indices were more negatively biased in periods with greater uncertainty.

Inclusion of strata with fewer than 10000 hooks had a significant effect on the index in several years, particularly 2014 and 2015 (Figure 23).

Indices based on to Japanese data were slightly different depending on whether they were based on the CPUEInputs.txt file or the CatchEffort.txt file. These files differ in that fish less than 4 years old were omitted in the analyses of the CPUEInputs.txt file.

Maps

For maps of the data coverage see the Github repository (https://github.com/hoyles/R_ccsbt_cpue). Please email simon.hoyle@gmail.com to request access.

Discussion

The distributions of both fish and effort change through the time series, in addition to varying seasonally. The increasing sparsity of the Japanese fishing distribution is marked, with far fewer fished strata in 2022 than in 1986. The factors motivating the effort contraction are less well understood. Initially the contraction would been driven by reduced catch quotas. More recently, catch rates that increased faster than quotas are likely to have allowed vessels to catch their quotas with less effort. The contributions of improvements in fishing and communication technology, remote sensing, and understanding of tuna behaviour are unclear but may have increased vessels' ability to identify and target areas of higher abundance, thus increasing catch rates and reducing the effort required to take the quota.

GAM models perform better than GLM models at estimating indices from sparse data. They suffer less from the problem of inestimable parameters since they share information across strata. Nevertheless, there are limits to their ability to derive information from sparse data, and this is reflected in increasing uncertainty about recent catch rates. Lower spatiotemporal coverage is inevitably associated with higher uncertainty and increases the risk that fished strata are not representative of the overall biomass trend.

Analyses in this study have shown that data from other fleets can significantly improve coverage throughout the time series, and particularly in recent years. Catch rates of most other fleets are likely to contain useful information about abundance trends. They show similar trends to indices from the Japanese fleet, as shown by indices based on nominal and standardized CPUE.

Joint analysis using data from multiple fleets fishing on the same stock is increasingly applied as a way to increase the coverage and representativeness of CPUE indices (Hoyle et al., 2018; Hoyle et al., 2015b; Kitakado et al., 2021). Such analyses require significant work to prepare data, to ensure they are compatible for a joint analysis. Different fishing methods are used by different fleets, and by different groups and even different vessels within fleets, resulting in variation in catchability. The current study has to some extent accounted for fleet-level catchability variation by using the fleet as a covariate – the only option available with aggregated data. However, aggregated data are likely to be unsuitable for joint analyses - operational data will be required. There is likely to be considerable catchability variation within fleets other than Japan, given the diversity of vessel size, experience, equipment, bait use, and targeting practices within domestic fleets compared to distant water fishing fleets. These sources of variability can be addressed using a combination of techniques, such as the inclusion of vessel ids, identification of targeting practices, and auxiliary analyses using additional covariates.

Targeting behaviour is an important issue that can significantly affect catch rates. A common approach is the use of clustering based on species composition (He et al., 1997; Hoyle et al., 2022; Hoyle et al., 2015b).

Before jointly analysing national datasets, each dataset should be thoroughly explored and characterised to identify factors that may need to be accounted for during the standardization, and to eliminate sources of data conflict (e.g., Hoyle et al., 2015a; Hoyle and Okamoto, 2015; Hoyle et al., 2015c). It will also be necessary to remove effort where there may be issues with reporting quality or the representativeness of the sampling frame (e.g., Hoyle et al., 2015b).

Previous arrangements for joint standardization of operational DWFN tuna catch and effort data (Hoyle et al., 2018; Hoyle et al., 2015b) have involved joint in-person meetings of one week to 10 days at which all data are shared amongst all participants. The first exercise of this type also involved several week-long in-country meetings to characterise data and develop the code used to explore, clean, and prepare data (https://github.com/hoyles/cpue.rfmo). At the joint workshops, analysts have run models for their own datasets, and the lead analyst has provided assistance and run both individual fleet and joint models. Joint analysis requires that all datasets have the same fields, so data preparation code needs to be consistent and shared. Joint analyses involve very large datasets, so standardization models run slowly. Past analyses have used GLM methods, which run much faster than GAMs and use less RAM. Code development would be helpful to permit GAM runs to be parallelized for much greater speed. Pilot studies will also be needed to identify computer hardware requirements. The process would be greatly facilitated if data sharing or cloud computing arrangements could be identified that allowed analyses to be done outside a relatively short in-person meeting.

Bibliography

CCSBT. Report of the Twenty Fifth Meeting of the Scientific Committee. Online: Commission for the Conservation of Southern Bluefin Tuna; 2020

Gini, C. Variabilità e mutabilità. Reprinted in Memorie di metodologica statistica (Ed Pizetti E, Salvemini, T) Rome: Libreria Eredi Virgilio Veschi. 1; 1912.

Gulland, J.A. On the fishing effort in English demersal fisheries. Fisheries Investigations Series II. 20:41 pp; 1956.

He, X.; Bigelow, K.A.; Boggs, C.H. Cluster analysis of longline sets and fishing strategies within the Hawaii-based fishery. Fisheries Research. 31:147-158; 1997.

Hoyle, S. Potential CPUE model improvements for the primary index of Southern Bluefin Tuna abundance. CCSBT Extended Scientific Committee for the 26th Meeting of the Scientific Committee. Online; 2021

Hoyle, S. Validating CPUE model improvements for the primary index of Southern Bluefin Tuna abundance. Working Paper prepared for the 12th Operating Model and Management Procedure Technical Meeting, CCSBT, 20-24 June 2022, Hobart, Australia: CCSBT; 2022

Hoyle, S.D. Spatial considerations in bigeye and yellowfin CPUE from Japanese and Taiwan, China longline fisheries in the Indian Ocean, IOTC–2014–WPTT16–27 Rev_1. Working Party on Tropical Tunas 16, Bali, Indonesia: Indian Ocean Tuna Commission; 2014

Hoyle, S.D. Investigation of potential CPUE model improvements for the primary index of Southern Bluefin Tuna abundance, CCSBT-ESC/2008/29. CCSBT Extended Scientific Committee 25. online; 2020

Hoyle, S.D.; Huang, J.H.-w.; Kim, D.N.; Lee, M.K.; Matsumoto, T.; Walter III, J. Collaborative study of bigeye tuna CPUE from multiple Atlantic Ocean longline fleets in 2018. 2018

Hoyle, S.D.; Langley, A.D. Scaling factors for multi-region stock assessments, with an application to Indian Ocean tropical tunas. Fisheries Research. 228:105586; 2020. 10.1016/j.fishres.2020.105586

Hoyle, S.D.; Lee, S.I.; Kim, D.N. CPUE standardization for southern bluefin tuna (Thunnus maccoyii) in the Korean tuna longline fishery, accounting for spatiotemporal variation in targeting through data exploration and clustering. PeerJ. 10:e13951; 2022.

Hoyle, S.D.; Lee, S.I.; Kim, Z.G. Descriptive analyses of the Korean Indian Ocean longline fishery, focusing on tropical areas. Indian Ocean Tuna Commission Working Party on Tropical Tunas; 2015a

Hoyle, S.D.; Okamoto, H. Descriptive analyses of the Japanese Indian Ocean longline fishery, focusing on tropical areas. Indian Ocean Tuna Commission Working Party on Tropical Tunas; 2015

Hoyle, S.D.; Okamoto, H.; Yeh, Y.-m.; Kim, Z.G.; Lee, S.I.; Sharma, R. IOTC–CPUEWS02 2015: Report of the 2nd CPUE Workshop on Longline Fisheries, 30 April – 2 May 2015. Indian Ocean Tuna Commission; 2015b

Hoyle, S.D.; Yeh, Y.-M.; Chang, S.-T.; Wu, R.-F. Descriptive analyses of the Taiwanese Indian Ocean longline fishery, focusing on tropical areas. Indian Ocean Tuna Commission Working Party on Tropical Tunas; 2015c

Kitakado, T.; Satoh, K.; Lee, S.I.; Su, N.; Matsumoto, T.; Yokoi, H.; Okamoto, K.; Lee, M.K.; Lim, J.; Kwon, Y. Update of Trilateral Collaborative Study Among Japan, Korea and Chinese Taipei for Producing Joint Abundance Indices for the Atlantic Bigeye Tunas Using Longline Fisheries Data up to 2019. Collective Volume of Scientific Papers, ICCAT. 78:169-196; 2021.

Wood, S.N. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 73:3-36; 2011.

Figures

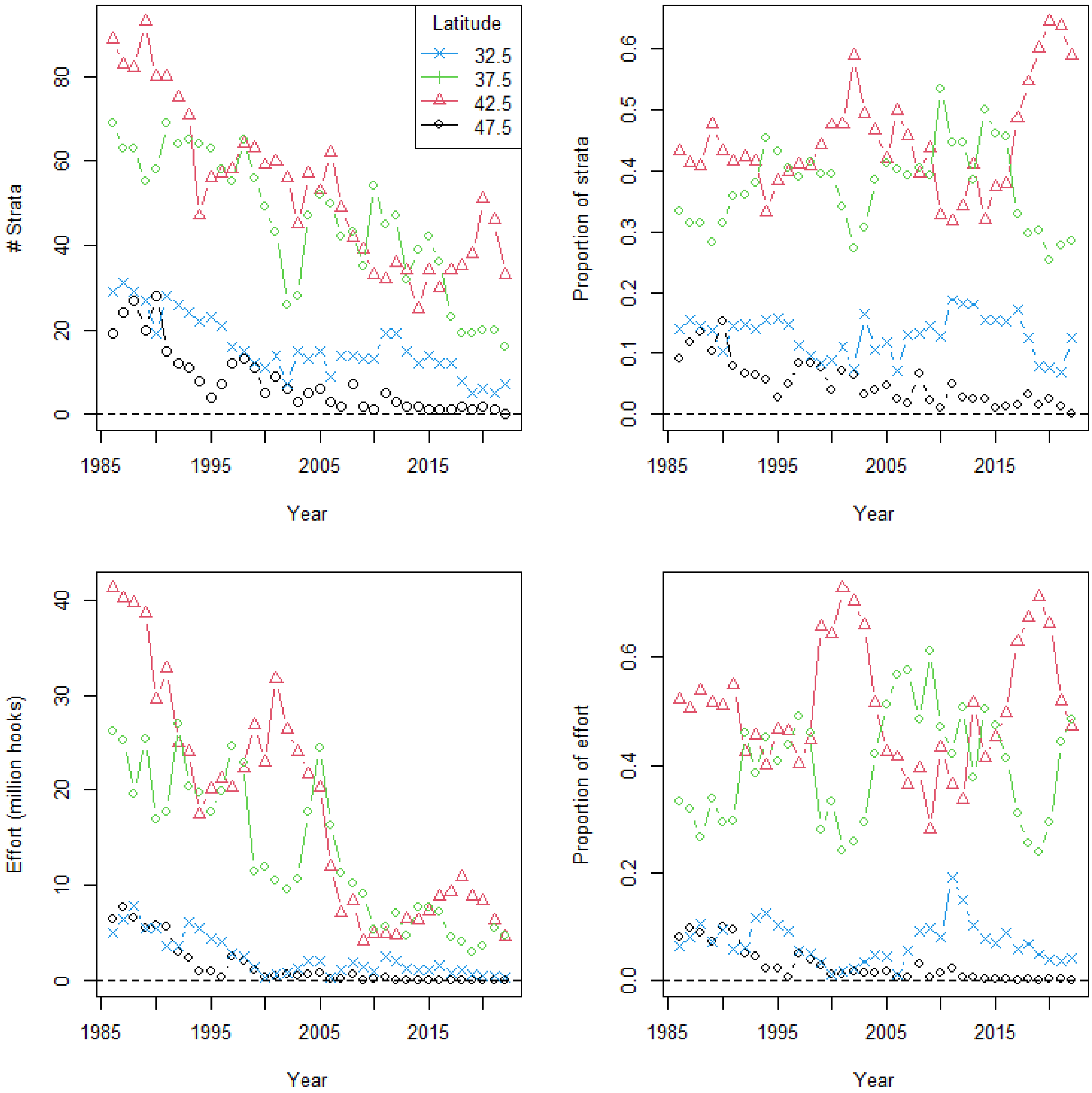

Figure 1: For the Japanese fleet, the number of strata fished (above) and total effort (below) by latitude and year, with totals on the left and proportions per latitude on the right.

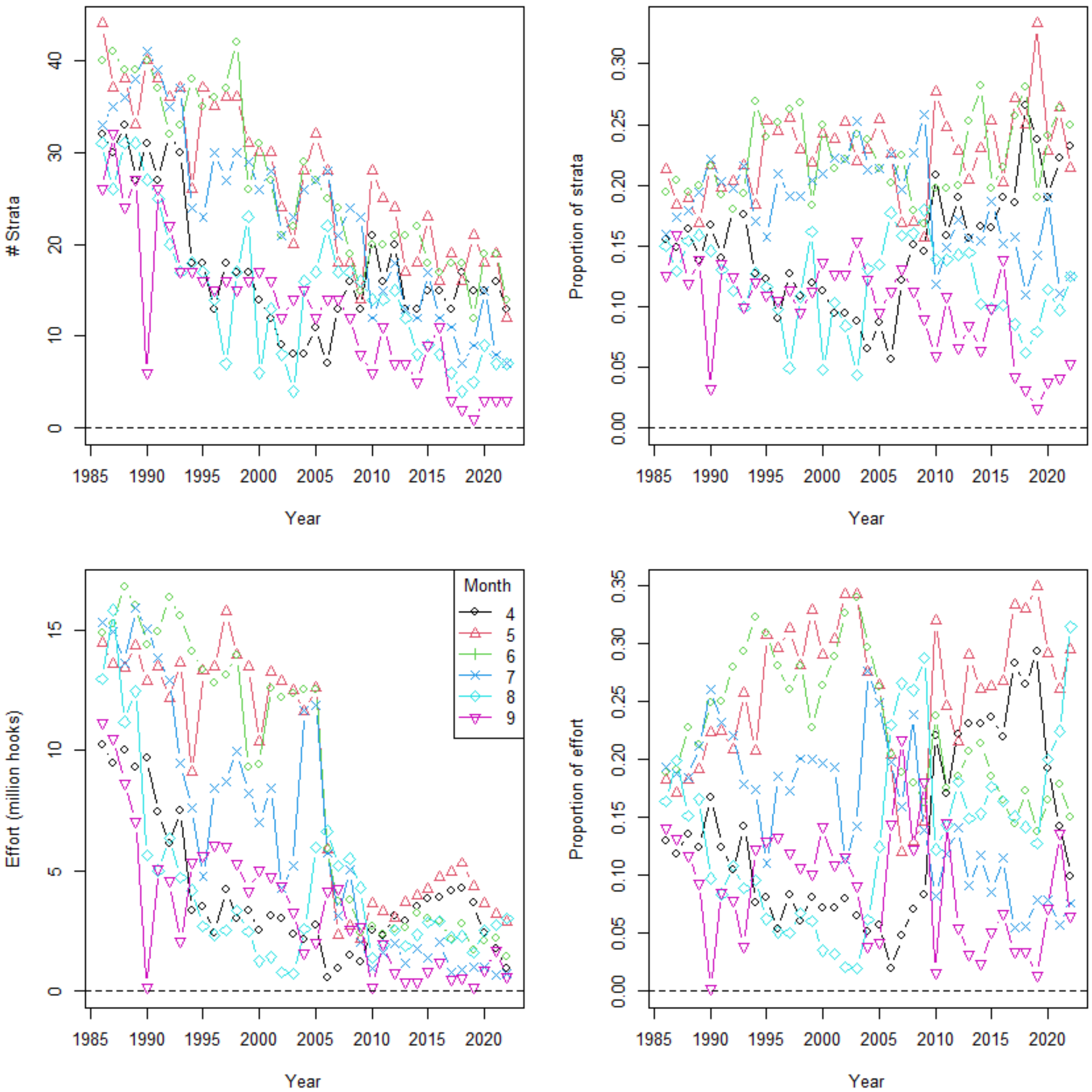

Figure 2: The number of fished strata (above) and total effort (below) by month and year, with totals on the left and proportions per month on the right.

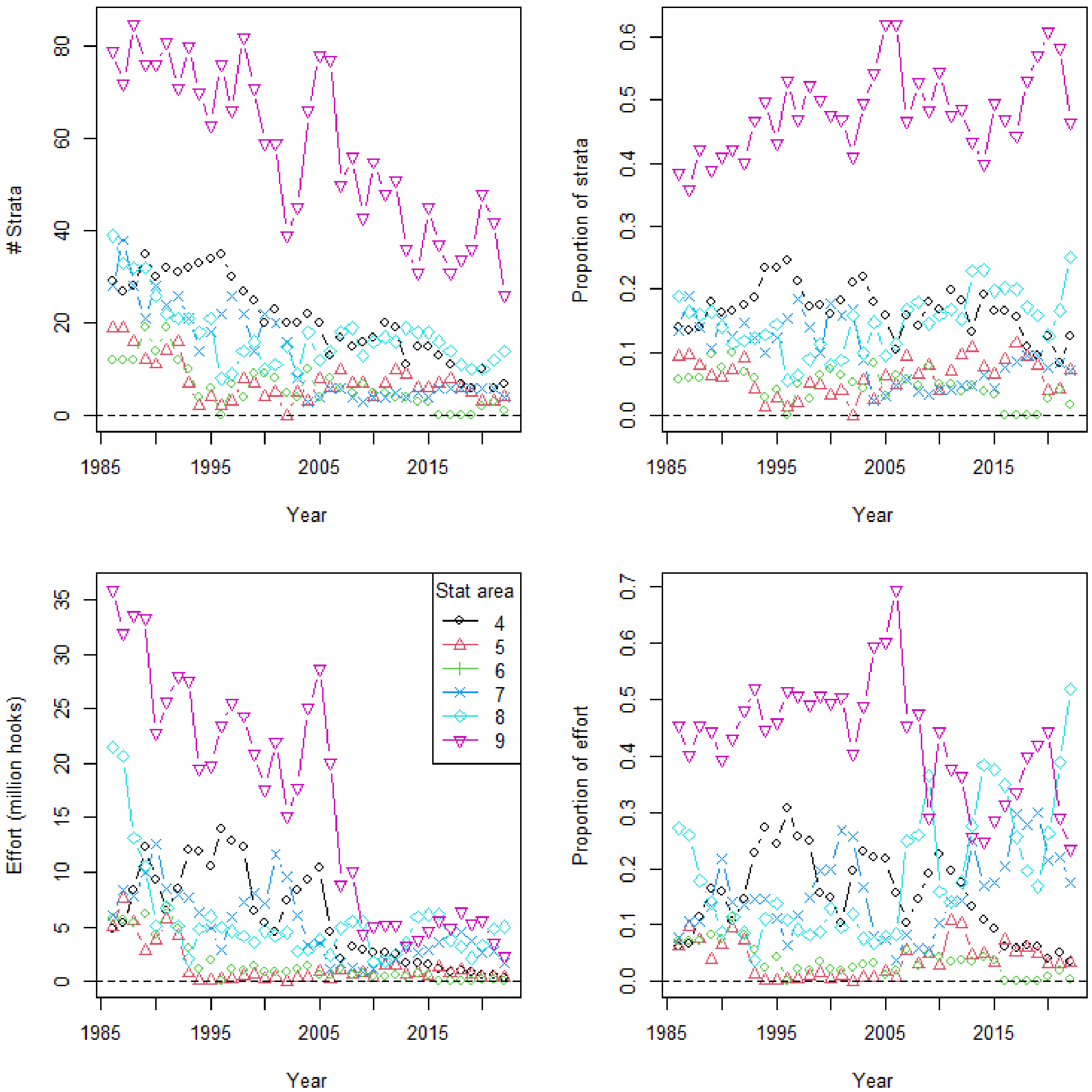

Figure 3: For the Japanese fleet, the number of fished strata (above) and total effort (below) by statistical area and year, with totals on the left and proportions per statistical area on the right.

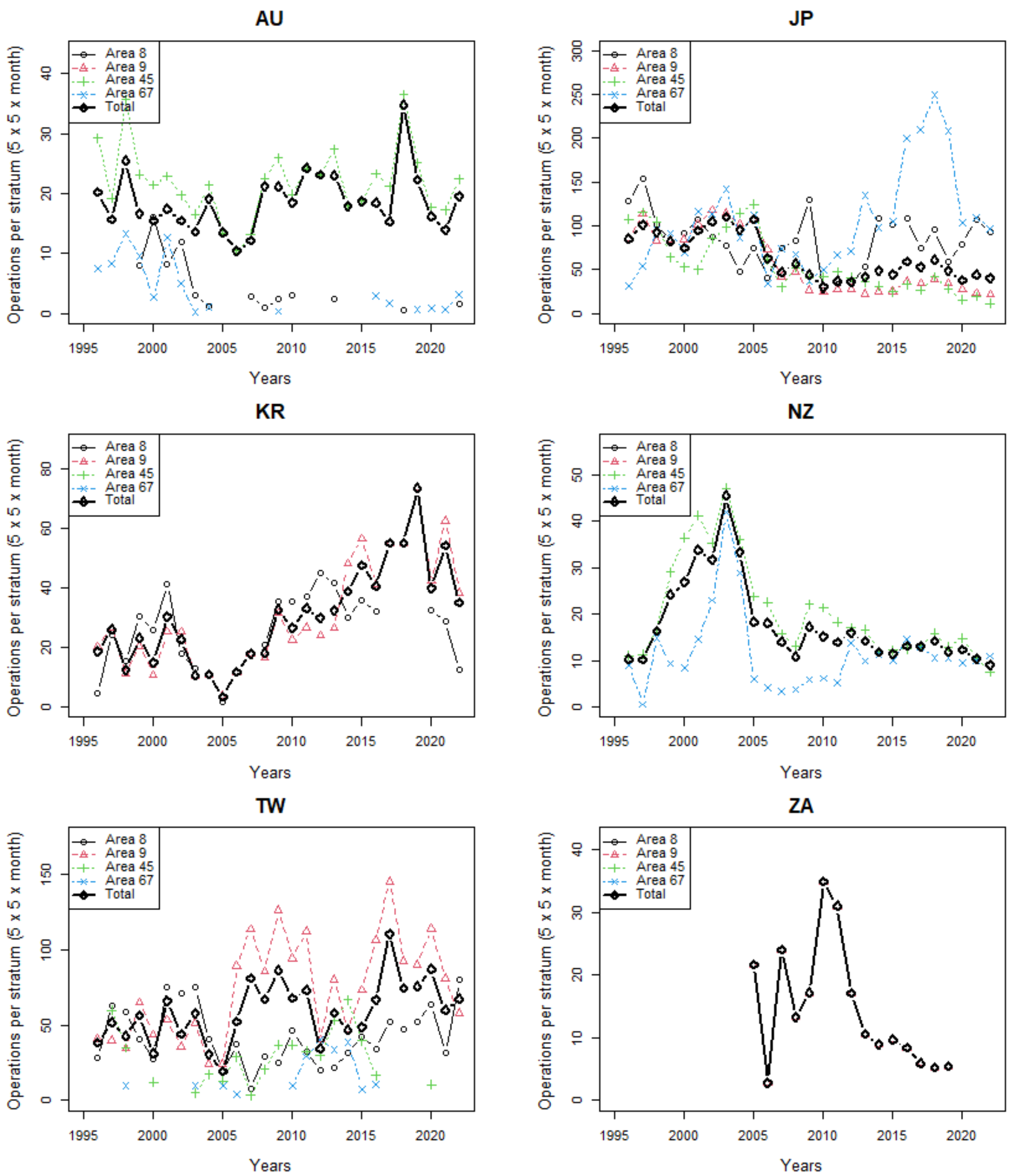

Figure 4: For each flag, the average number of operations per stratum (5 x 5 x month) by statistical area, year, and flag. The number of operations is estimated by assuming 3000 hooks per operation.

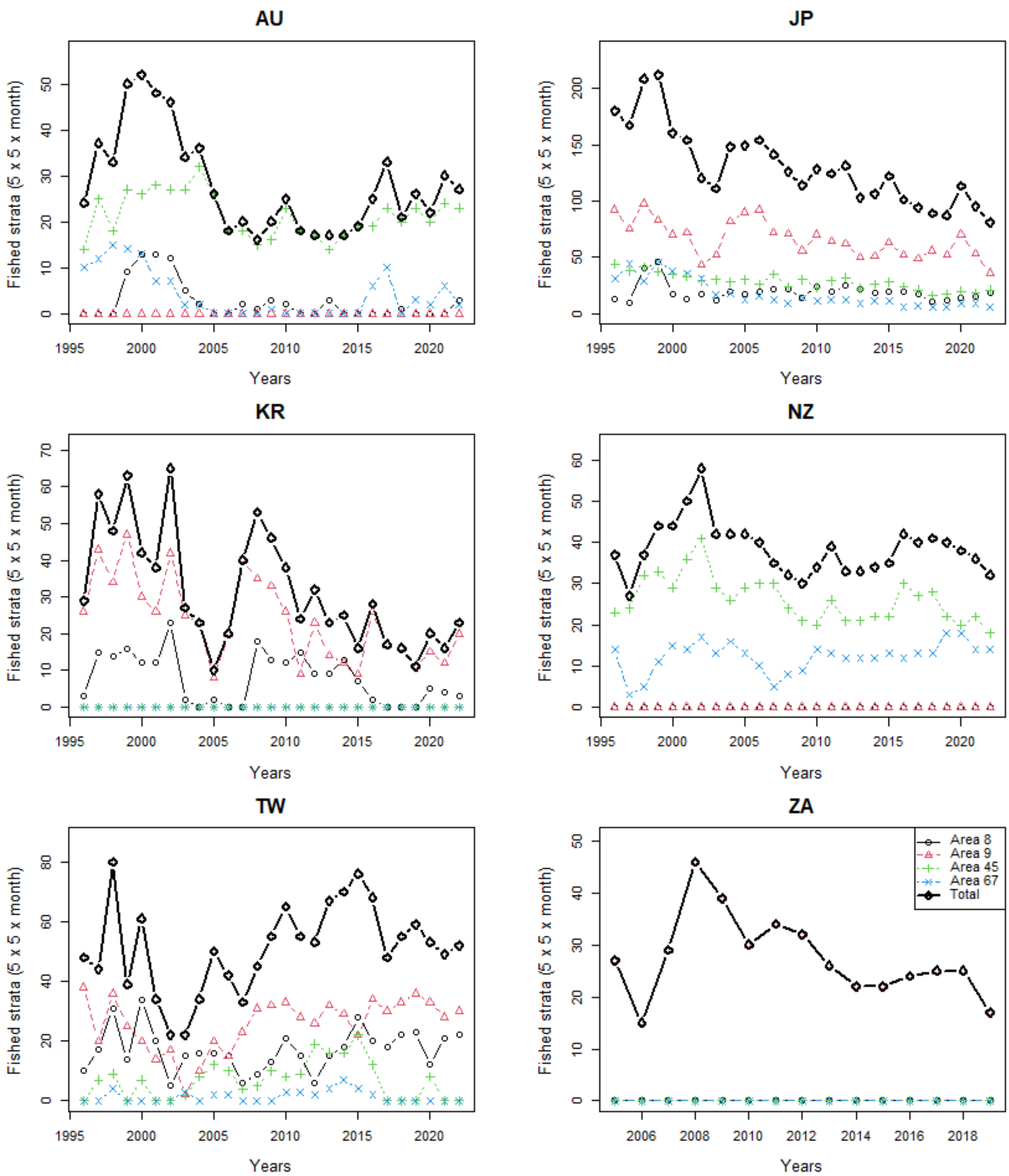

Figure 5: For each fleet, the number of fished strata (5 x 5 x month) per statistical area, year, and flag.

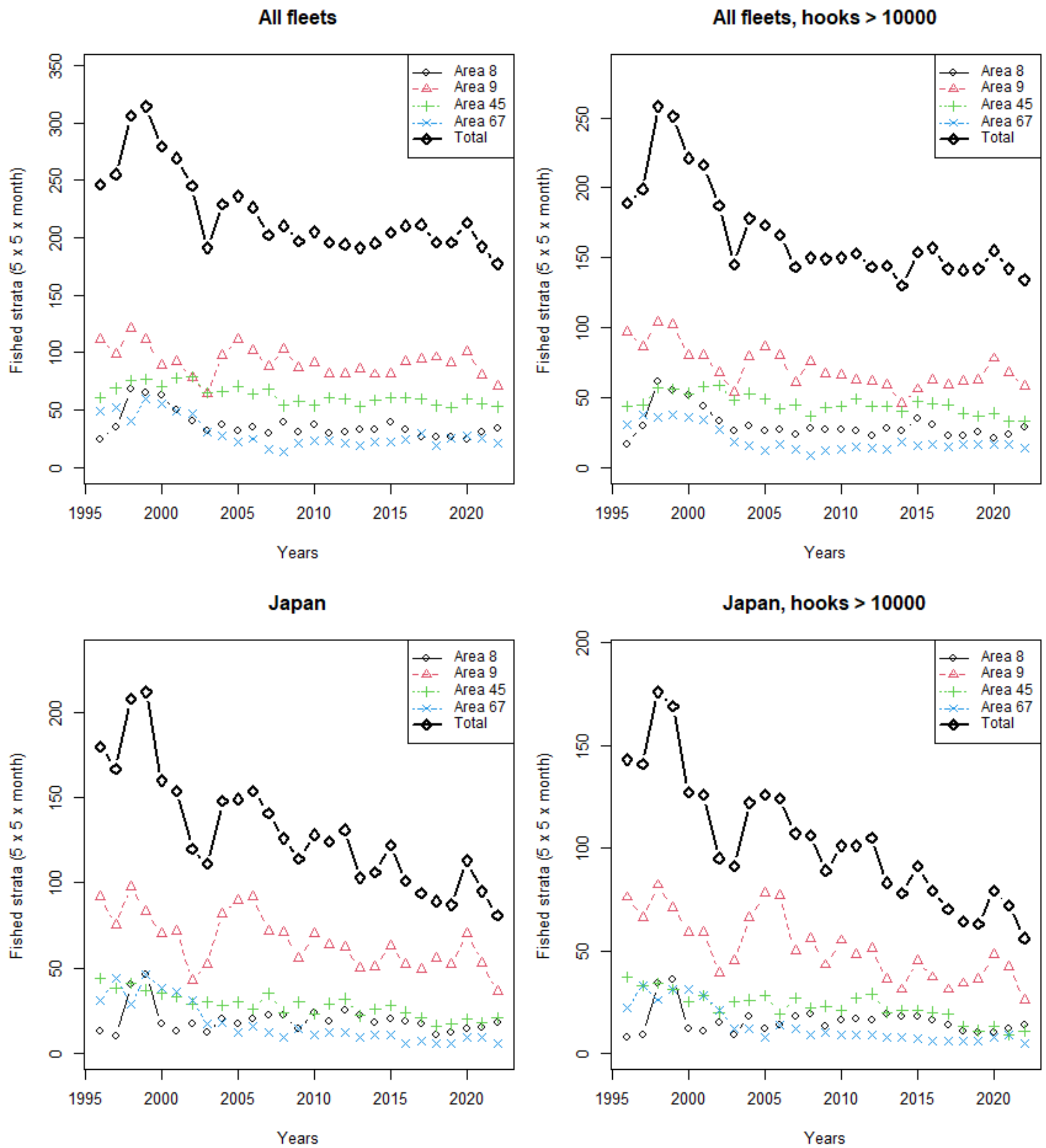

Figure 6: The number of fished strata per statistical area and year for all fleets combined (above) and for Japan alone (below).

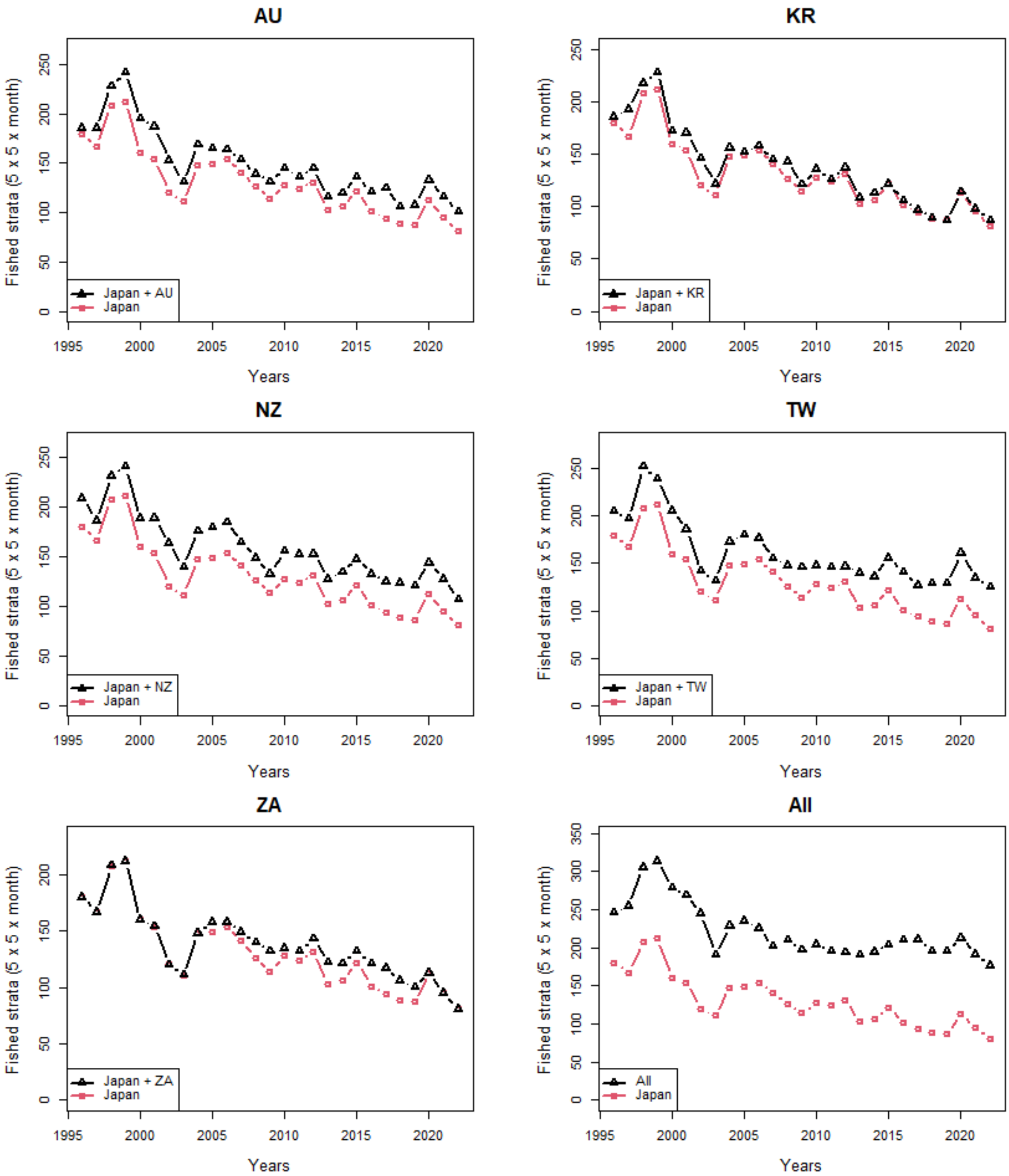

Figure 7: The annual numbers of fished strata for Japan (red) and for the combination of Japan and another flag (black). The 'All' subplot compares Japan (red) with the strata fished by all fleets combined (black).

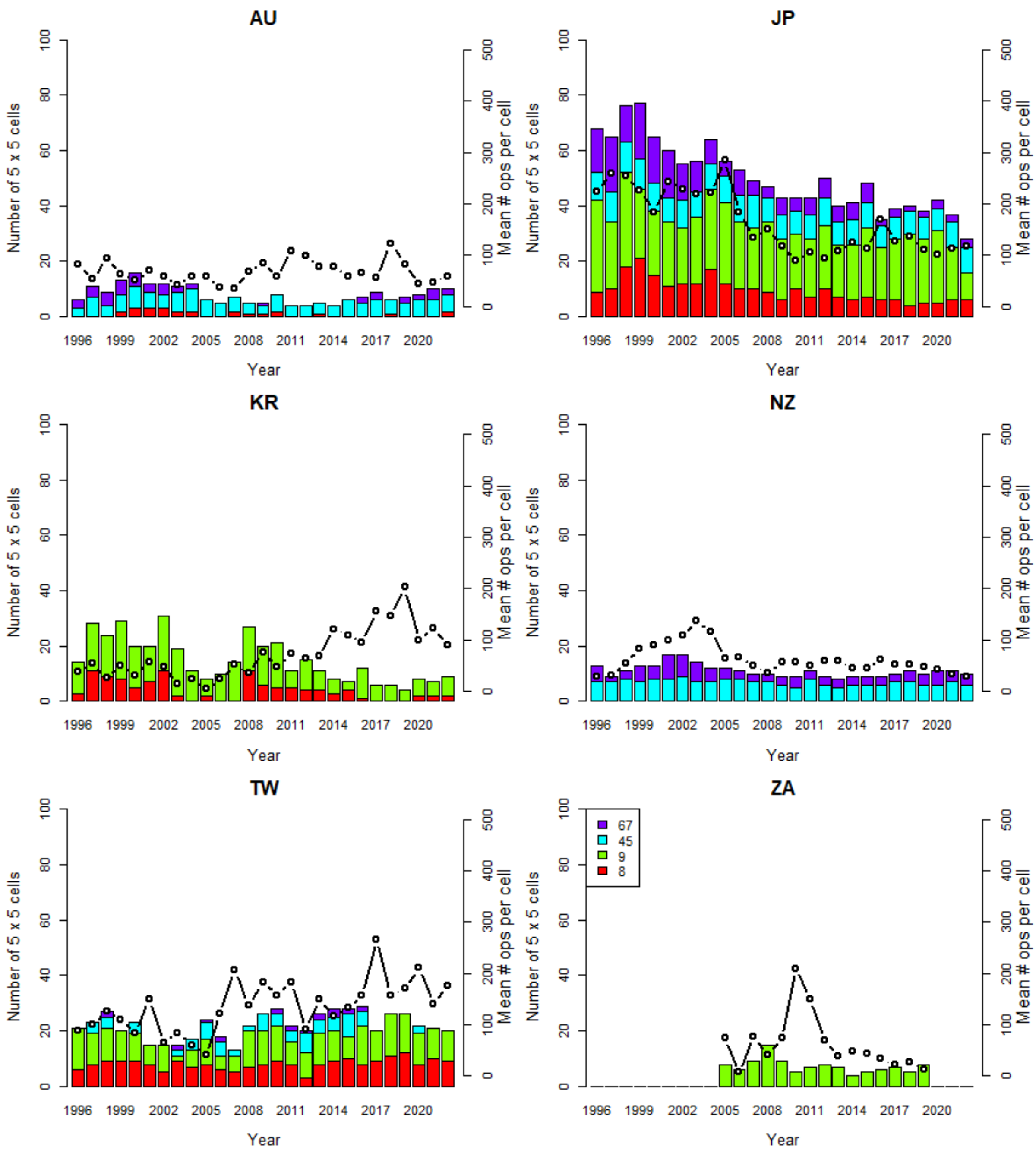

Figure 8: Combined plot showing both the number of fished cells (5 x 5) per statistical area, year, and flag (bar plot) and the number of operations per cell per year and flag (line plot).

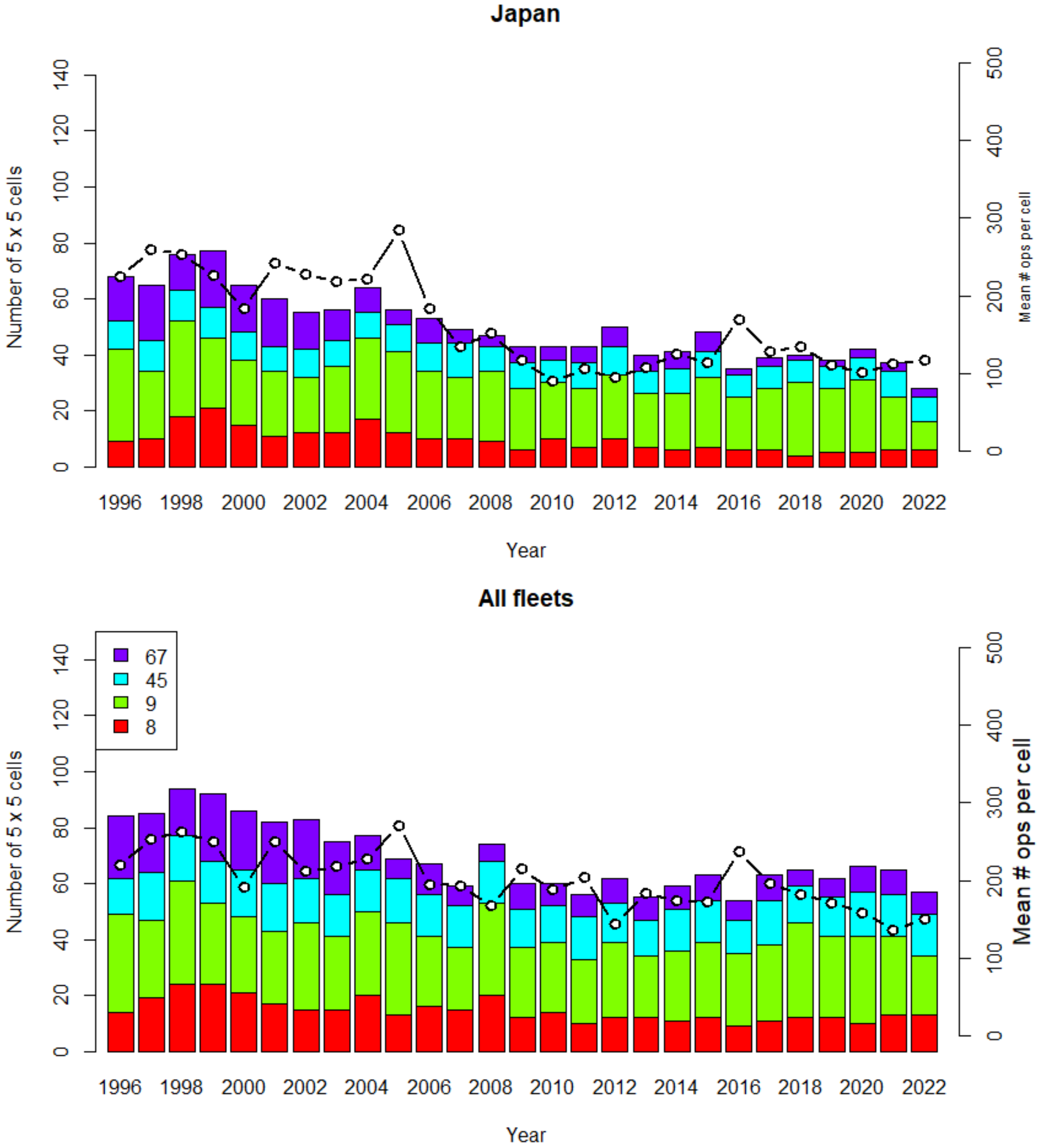

Figure 9: Combined plot showing both the number of fished cells (5 x 5) per statistical area and year (bar plot) and the number of operations per cell per year (line plot) for Japan (above) and all flags combined (below).

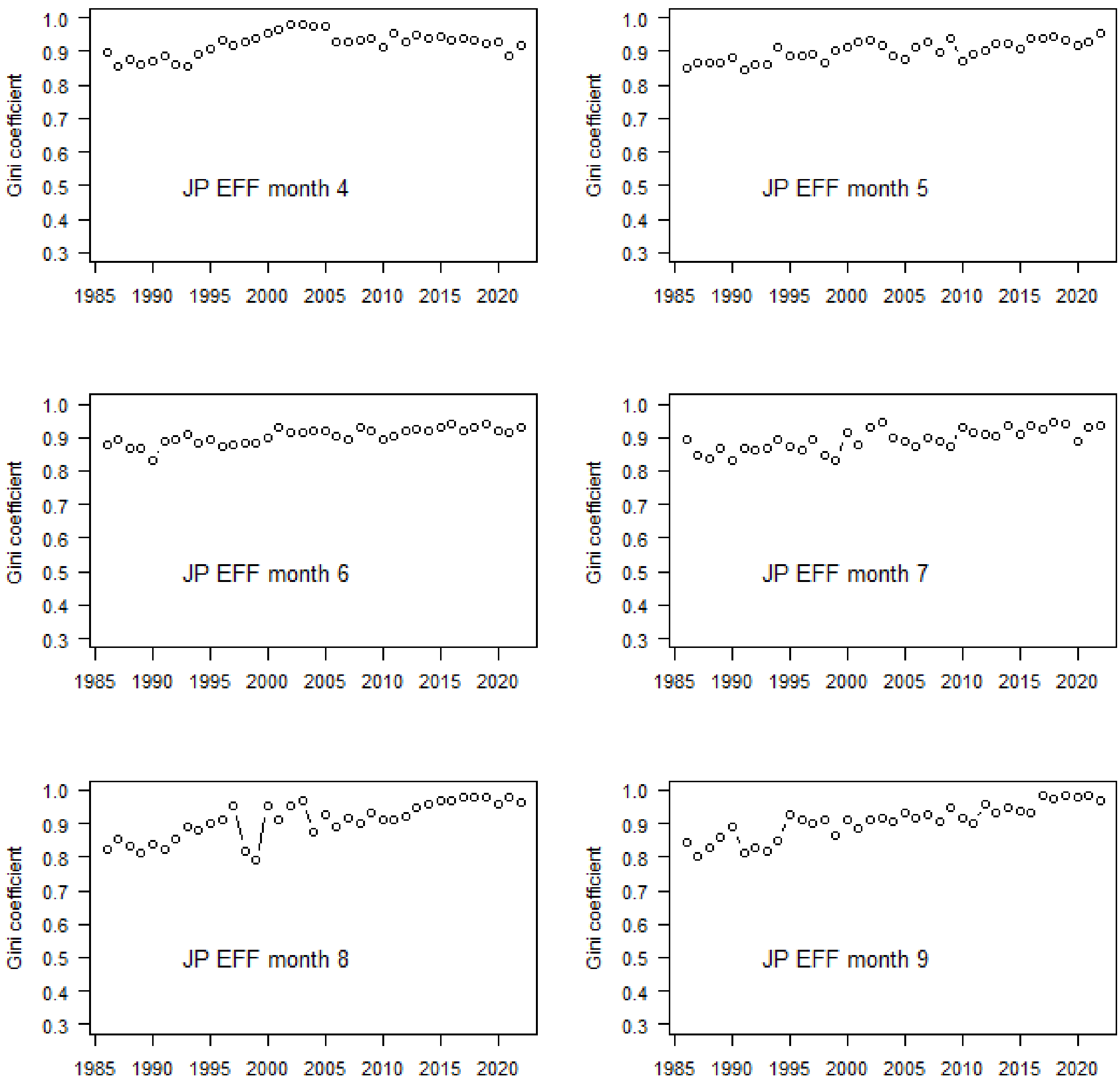

Figure 10: Gini coefficients for effort distribution by the Japanese fleet, by year and month.

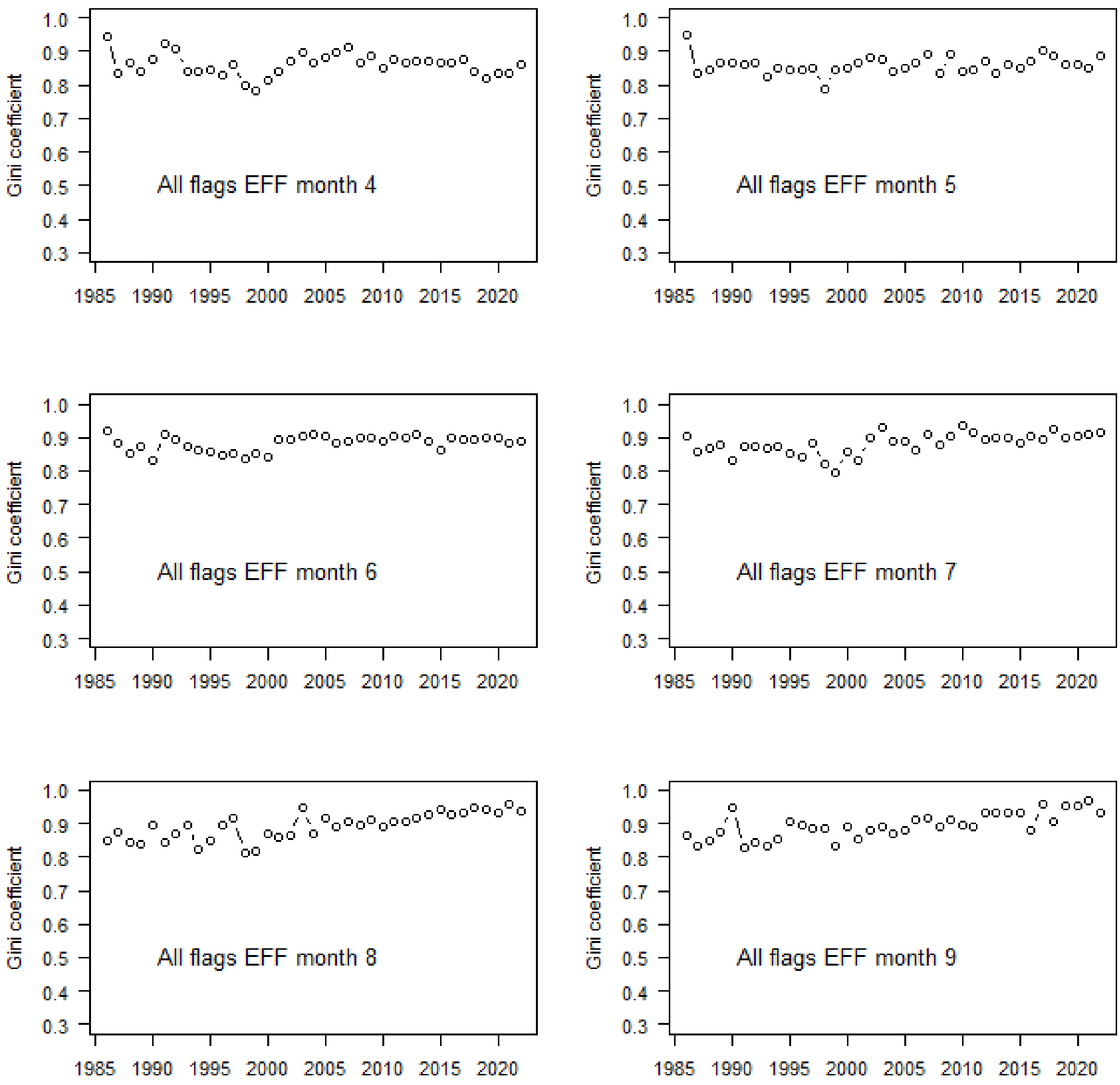

Figure 11: Gini coefficients for effort distribution across all fleets, by year and month.

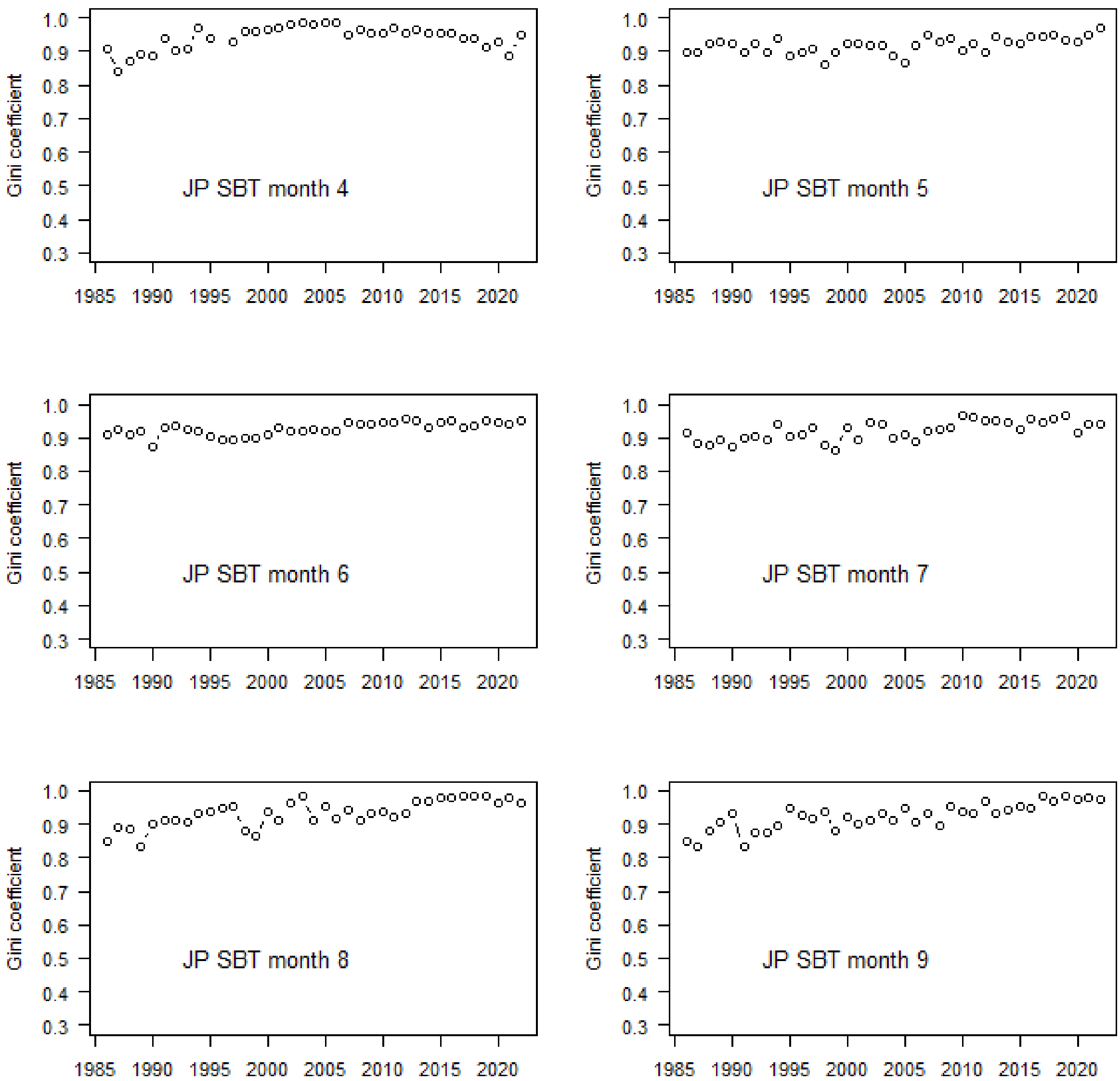

Figure 12: Gini coefficients for SBT catch distribution for the Japanese fleet, by year and month.

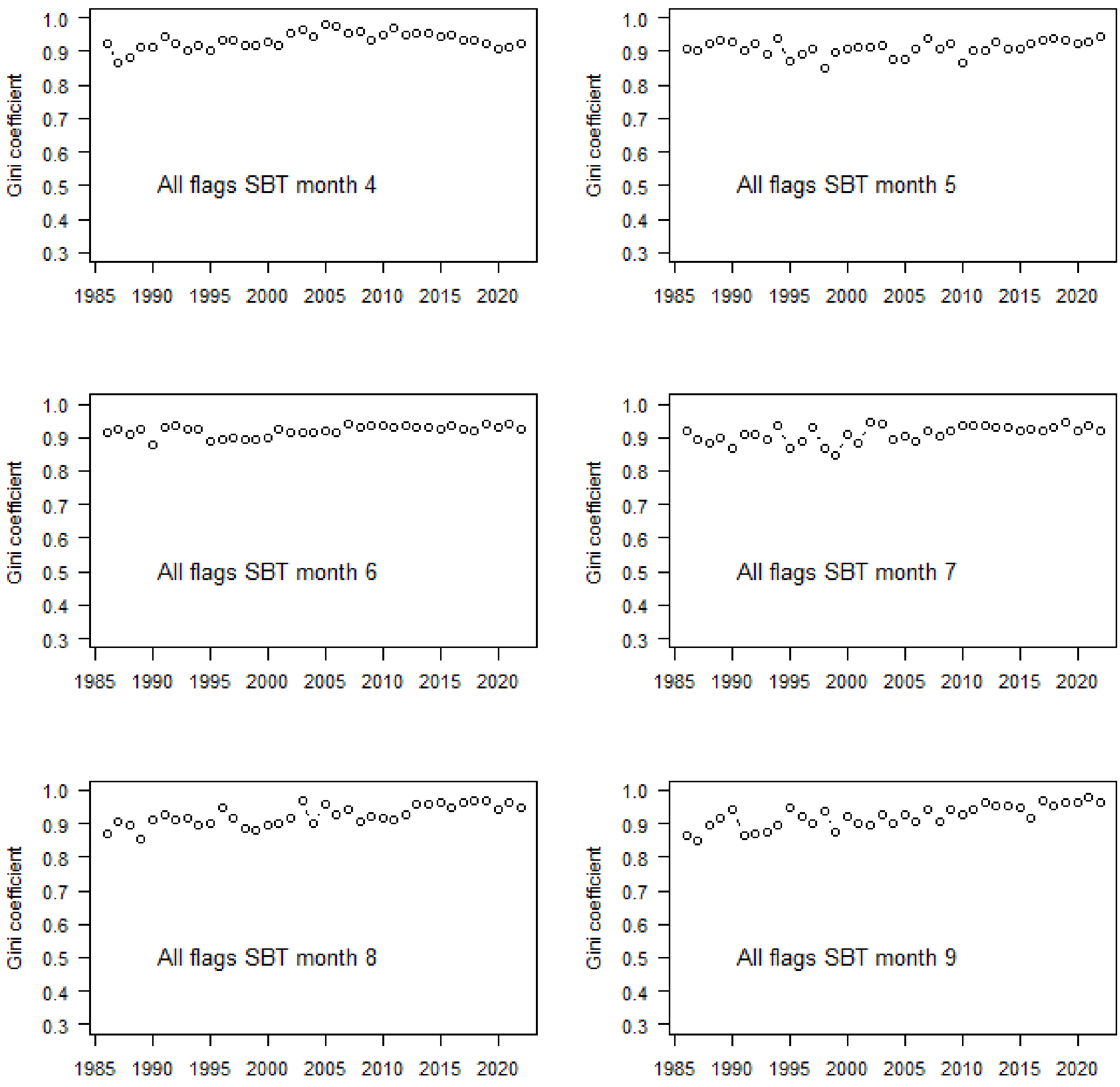

Figure 13: Gini coefficients for SBT catch distribution across all fleets, by year and month.

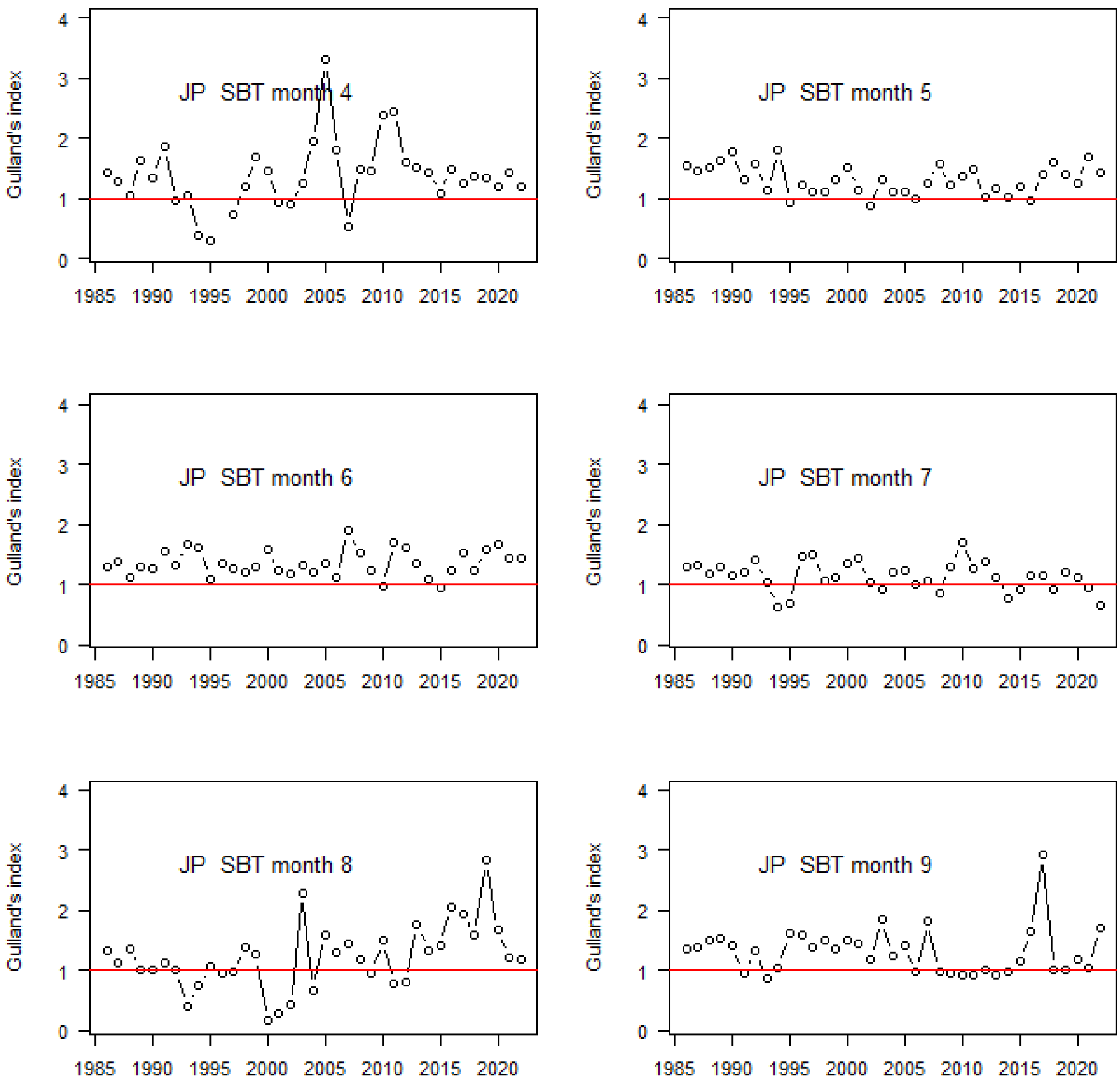

Figure 14: Gulland's coefficient of effort concentration applied to SBT catch and effort by Japan, by year and month.

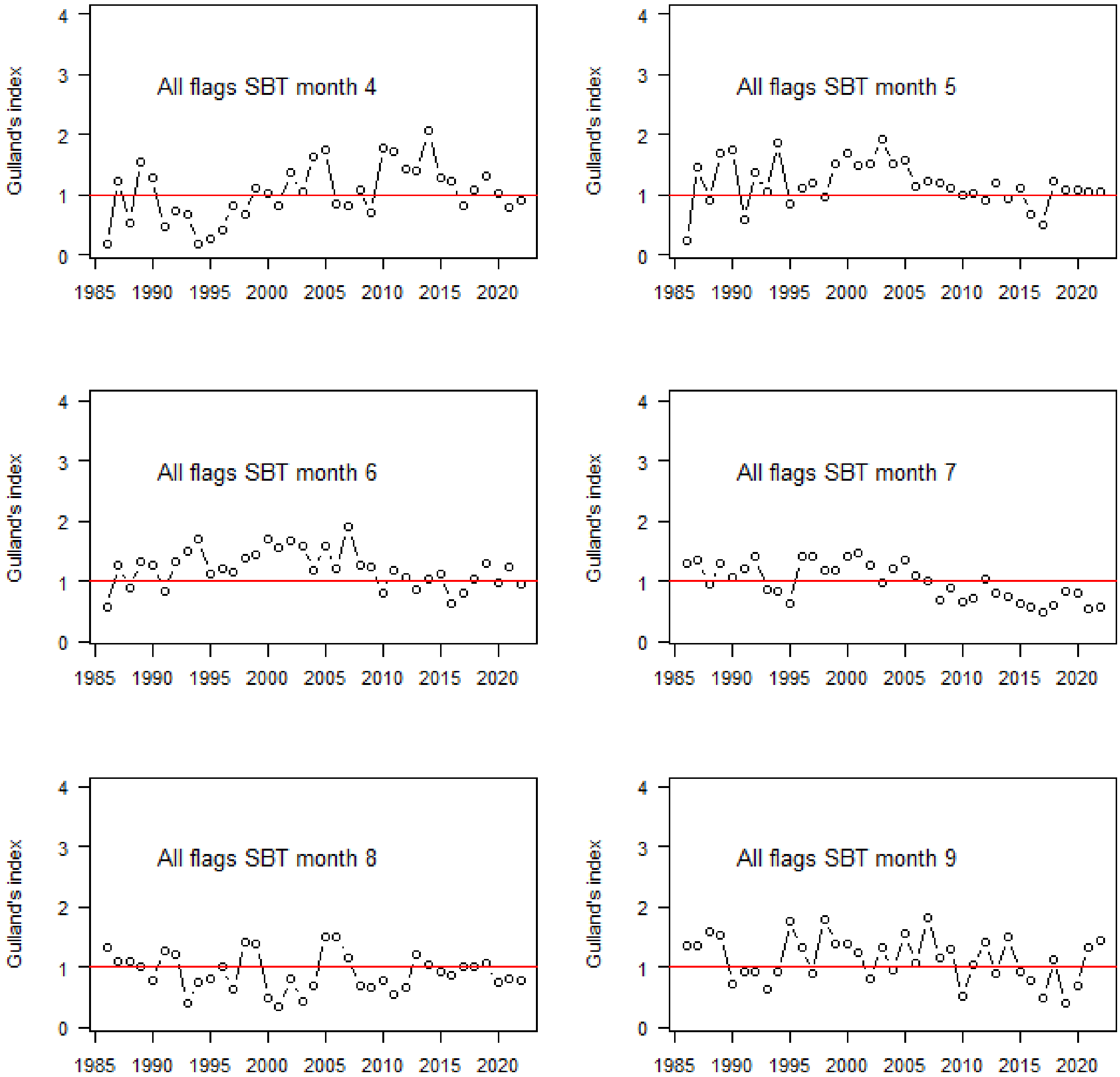

Figure 15: Gulland's coefficient of effort concentration applied to SBT catch and effort from all fleets combined, by year and month.

Figure 16: Standardized CPUE indices based on data from Japanese vessels only (black circles), and Japanese vessels plus one additional fleet at a time.

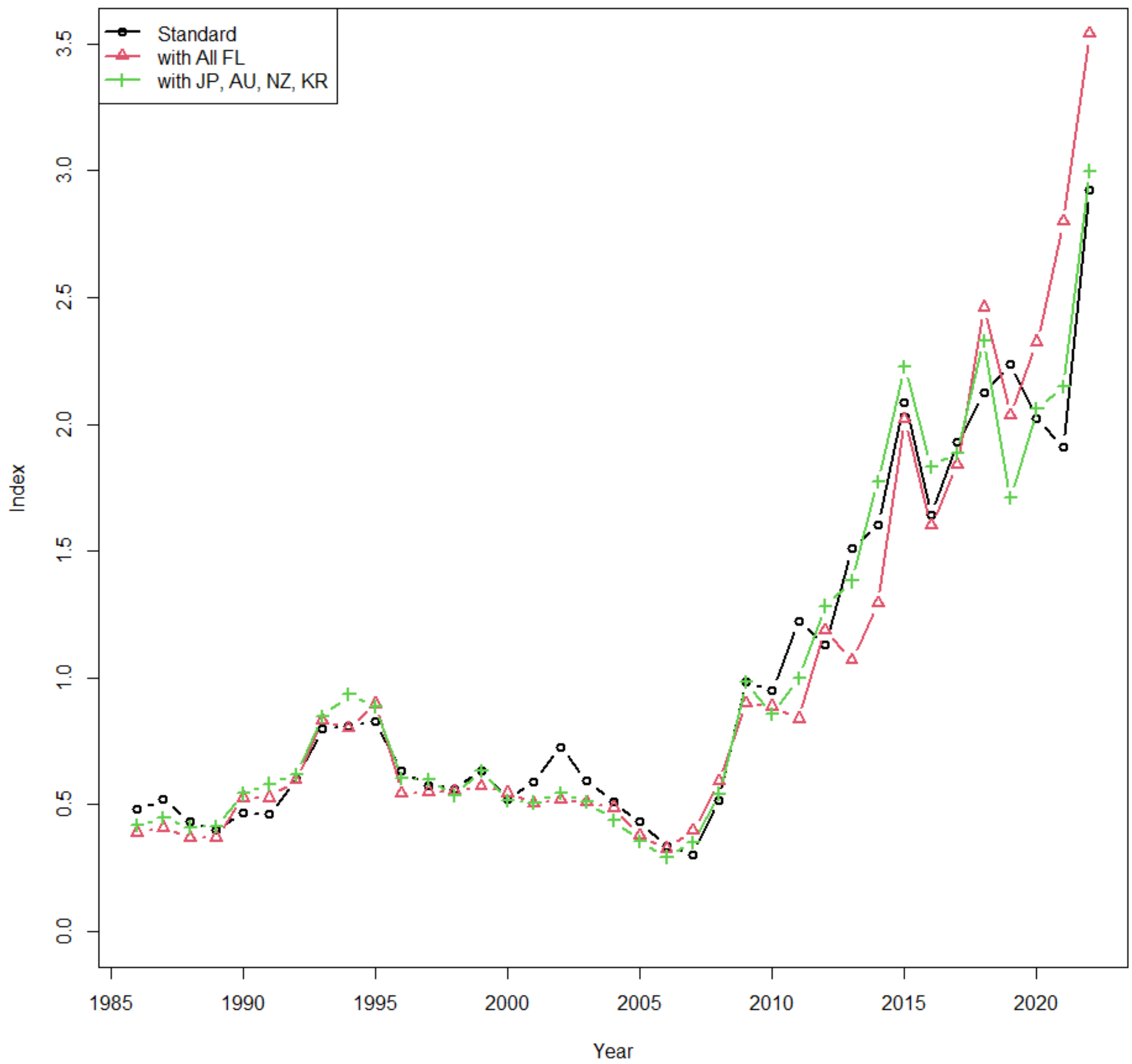

Figure 17: Standardized CPUE indices based on data from Japanese vessels (black circles), all fleets (red triangles), and Japanese, Australian, New Zealand and Korean vessels (green crosses).

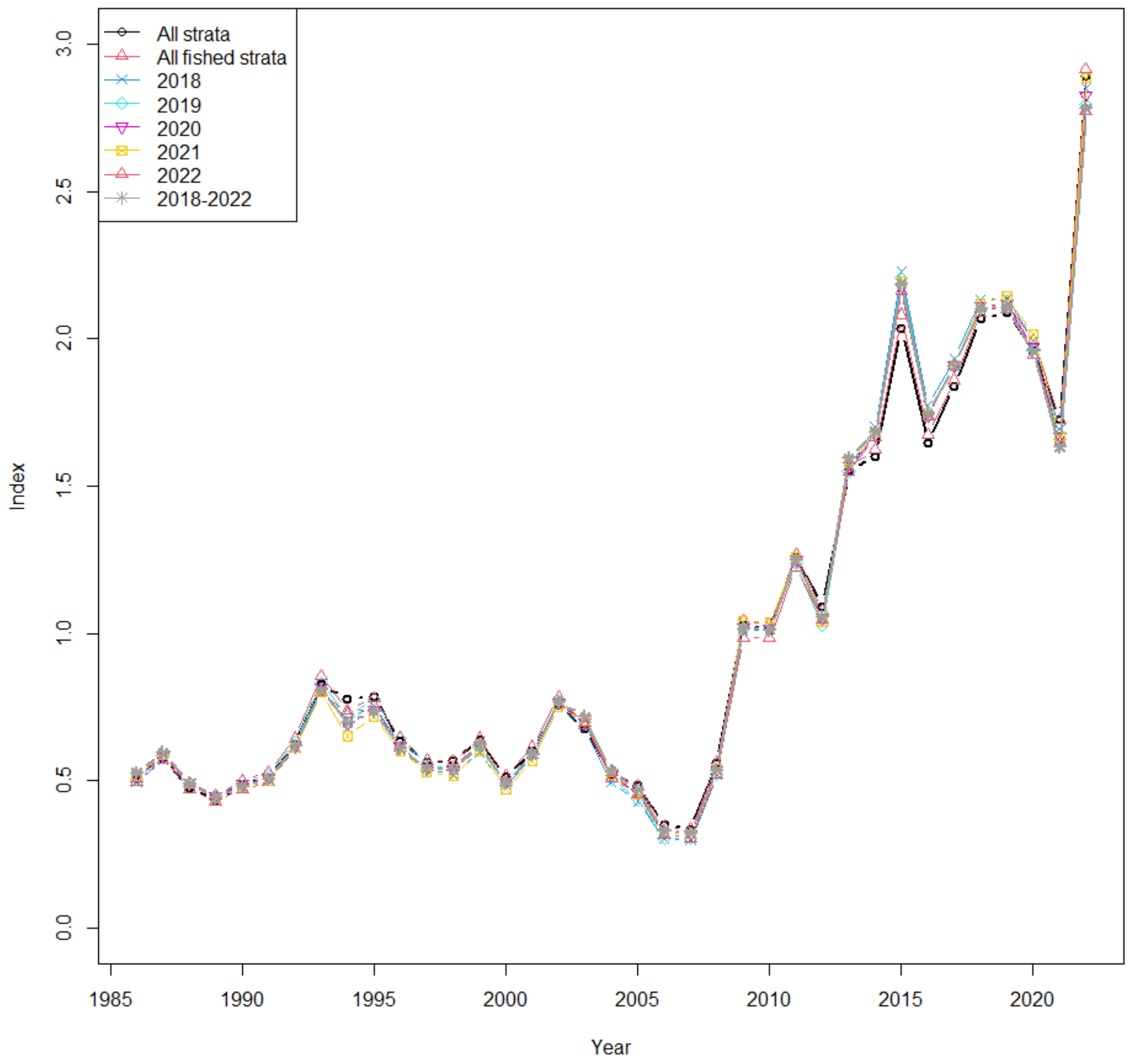

Figure 18: Comparison of indices based on predictions from different groups of cell-month strata, fitted using the CPUEinputs.txt dataset. Indices are based on all strata (black circles), strata fished between 1985 and 2022 (red triangles), and strata fished either during a single year, or during the period 2018-2022.

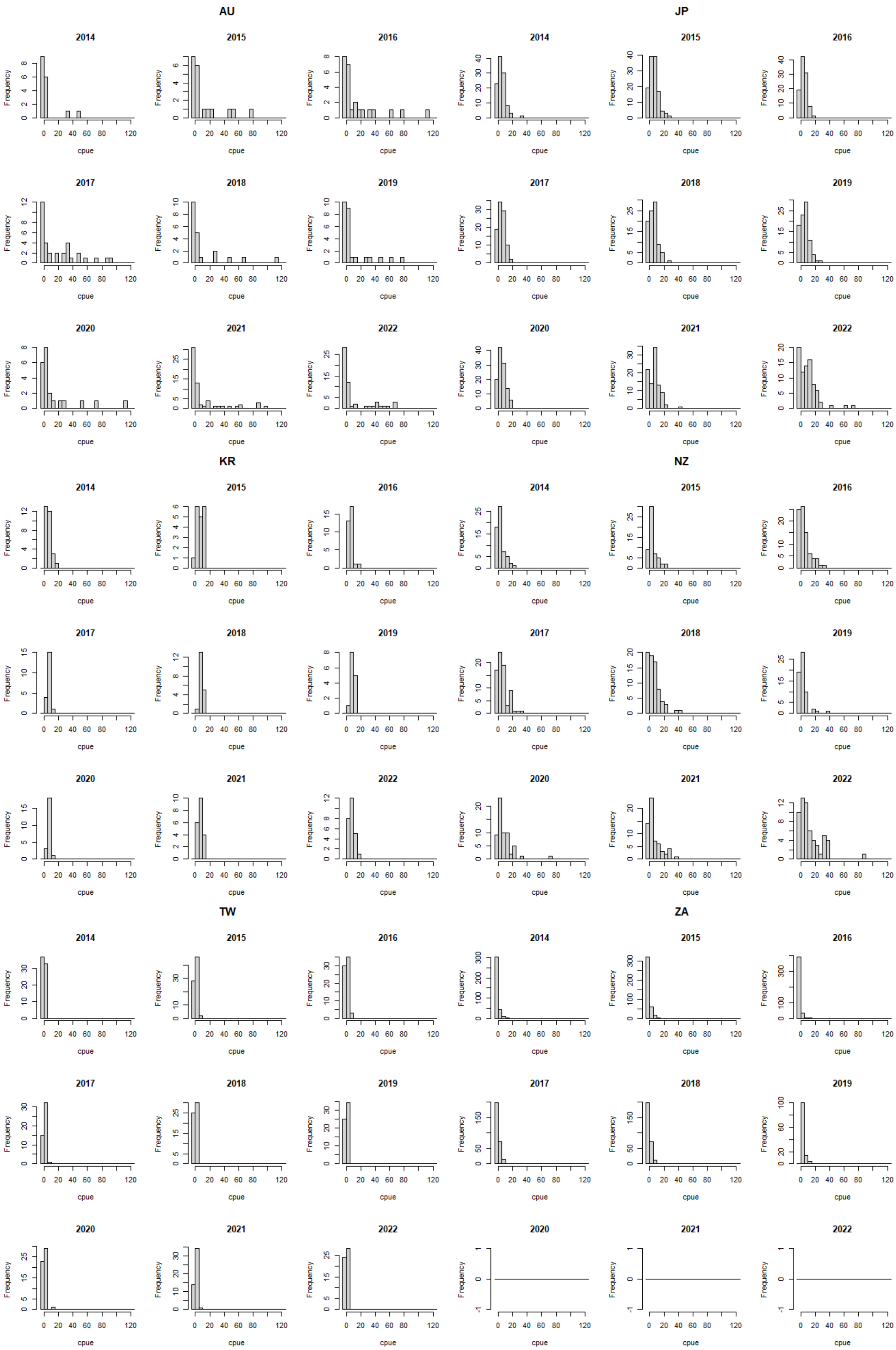

Figure 19: Frequency histograms of CPUE per cell-month stratum by year and fleet, for the period 2014 to 2021, and the Australian, Japanese, Korean, New Zealand, Taiwanese and South African fleets.

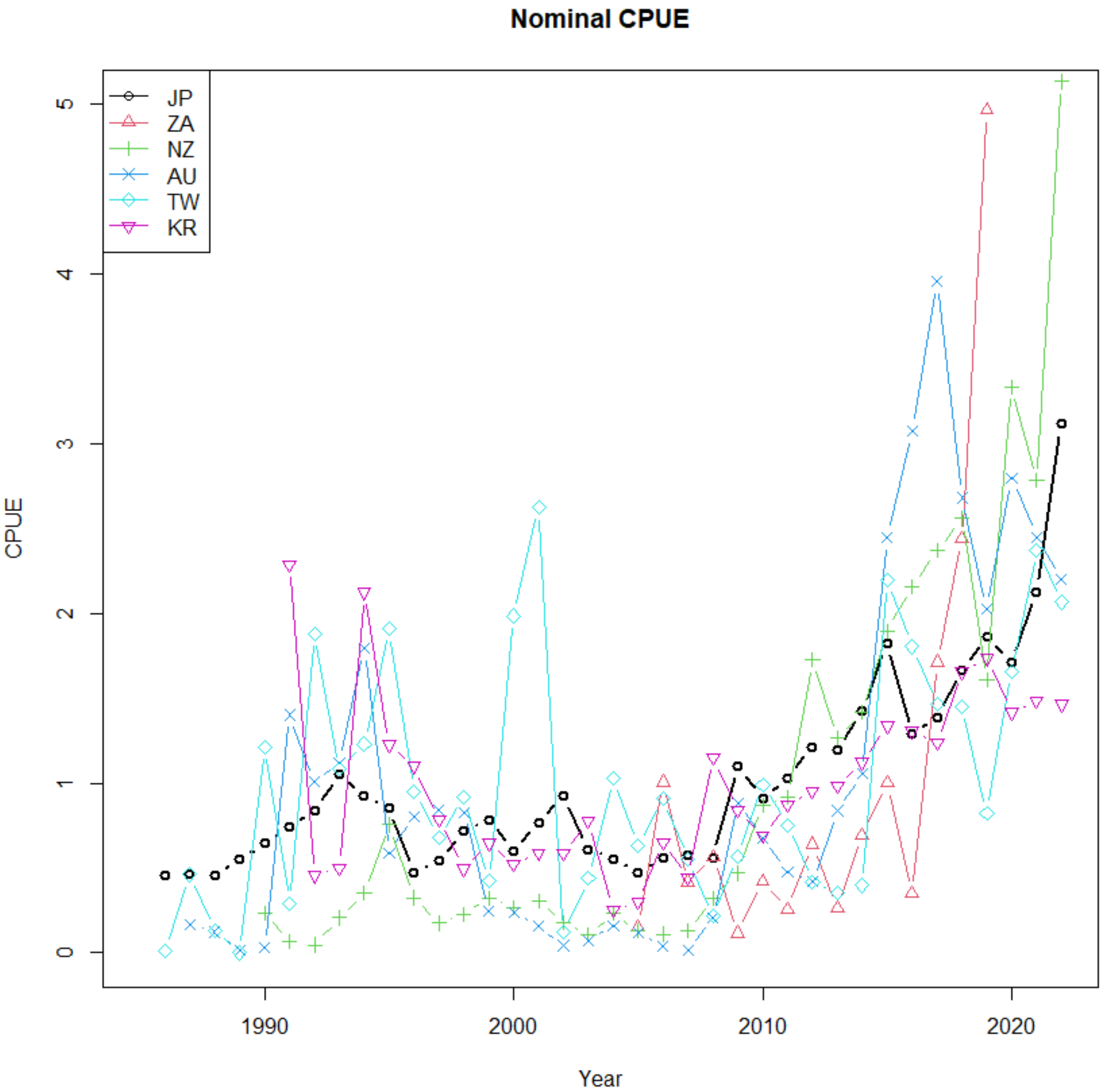

Figure 20:Nominal CPUE based on catch and effort data for individual fleets (not combined with data for the Japanese fleet).

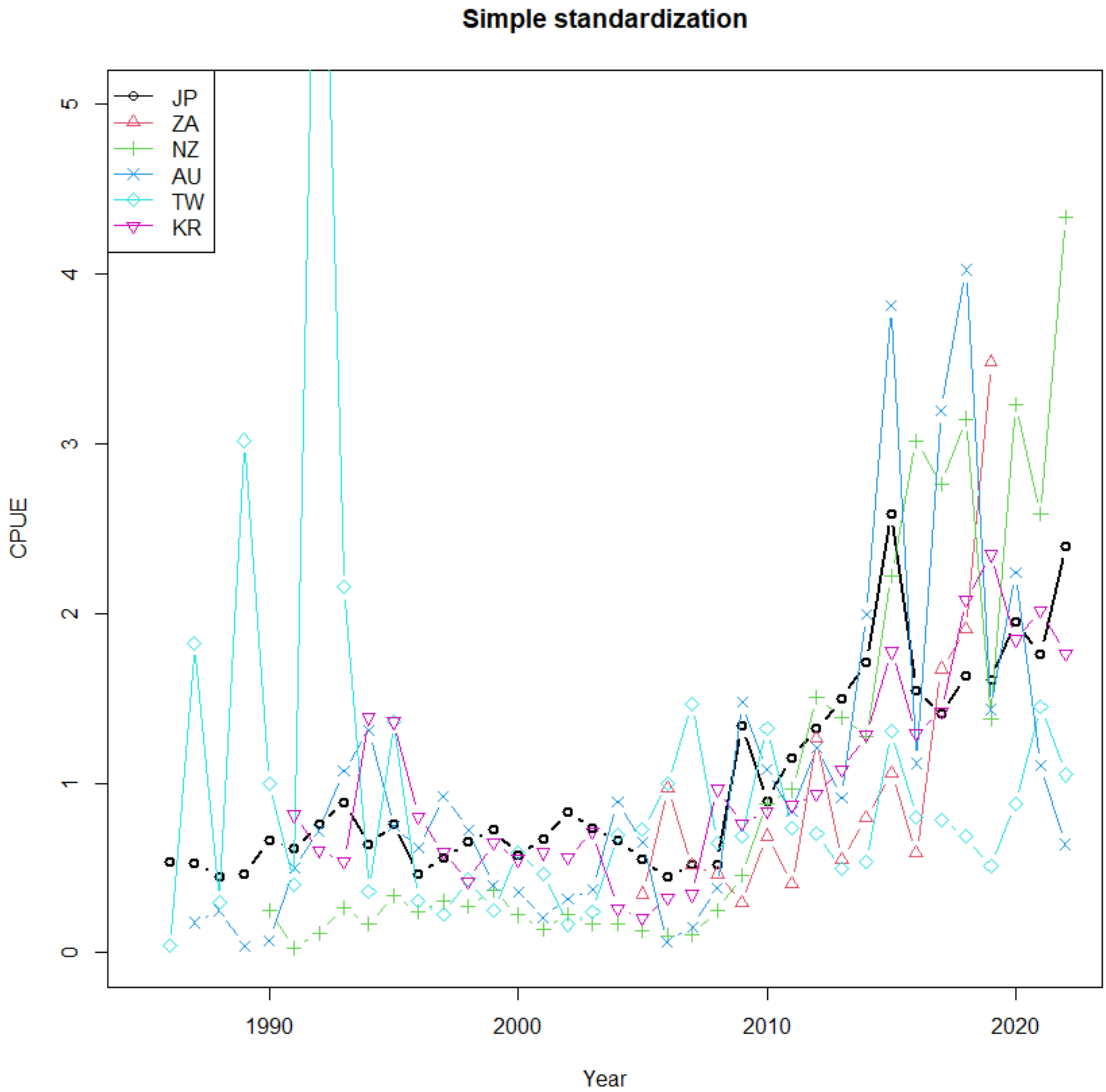

Figure 21: Indices of abundance based on standardized CPUE for individual fleets (not combined with data for the Japanese fleet), using the simple delta lognormal standardization model 1) CPUE !=0 ~ year + cell + month; and 2) log(CPUE) ~ year + cell + month.

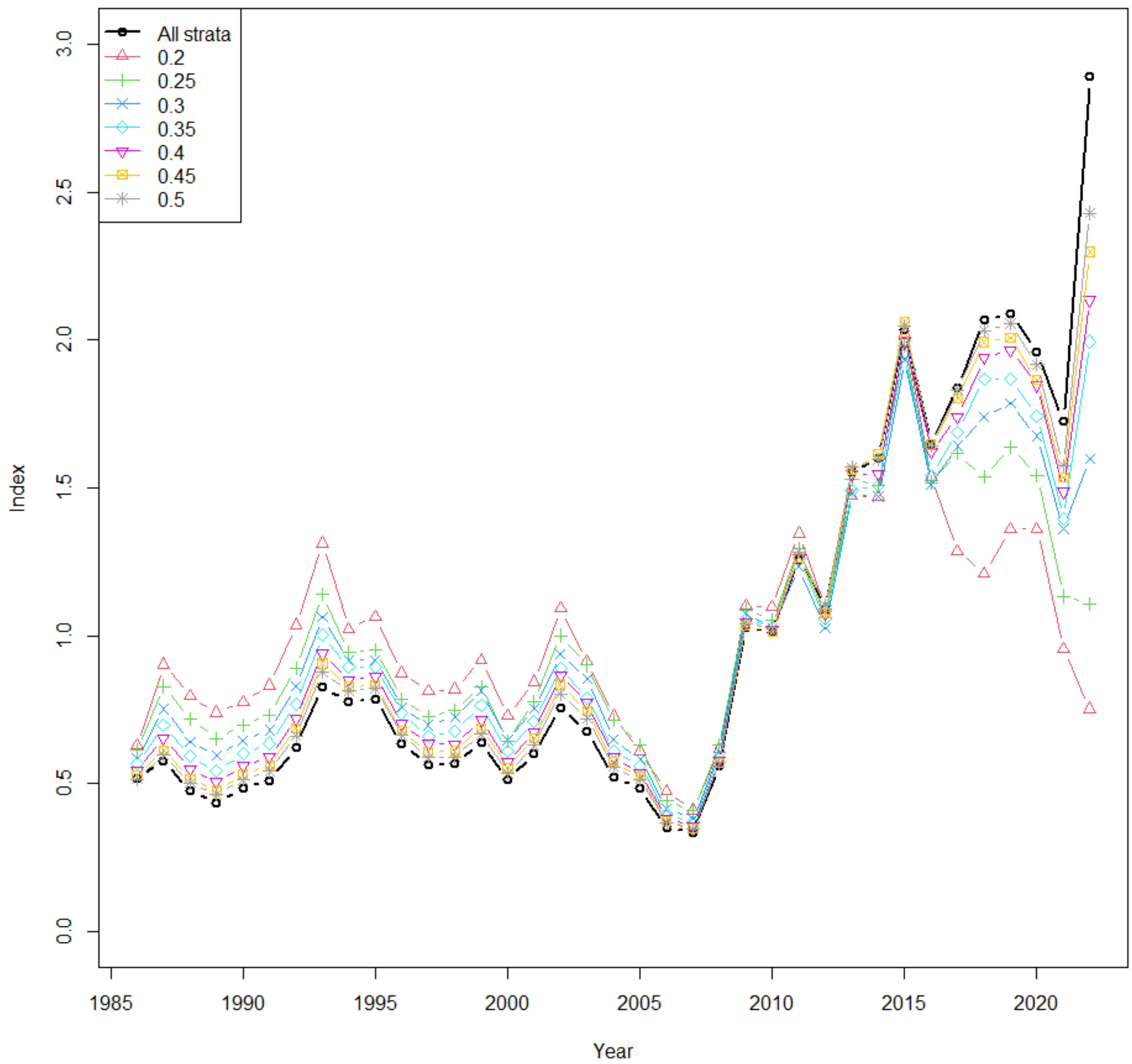

Figure 22: Indices of abundance based on the standard CPUEInputs file, based on the sums of predictions for strata with CVs less than the thresholds listed in the legend.

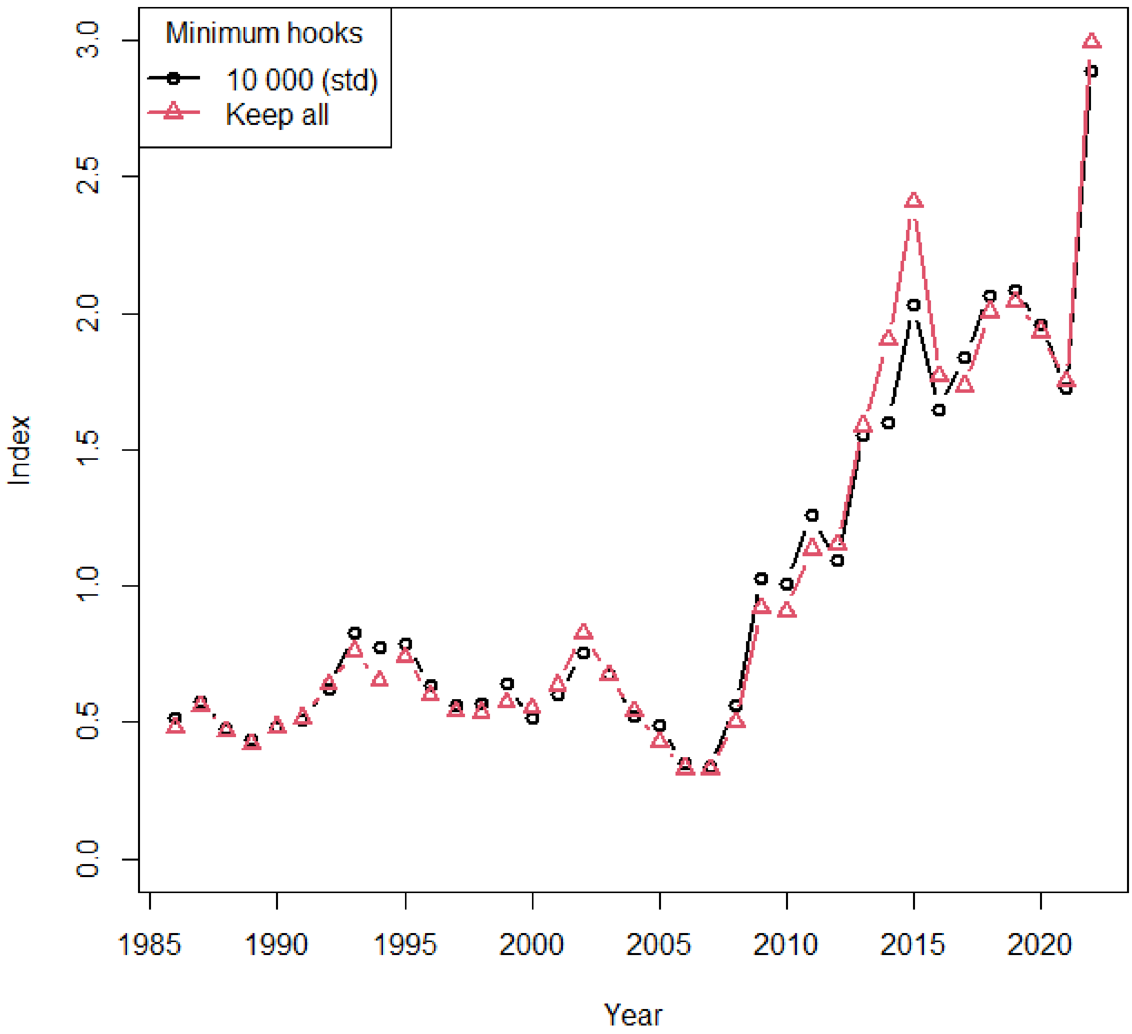

Figure 23: Indices of abundance based on the CPUEinputs file, where the model either retains all strata or omits strata with fewer than 10000 hooks.

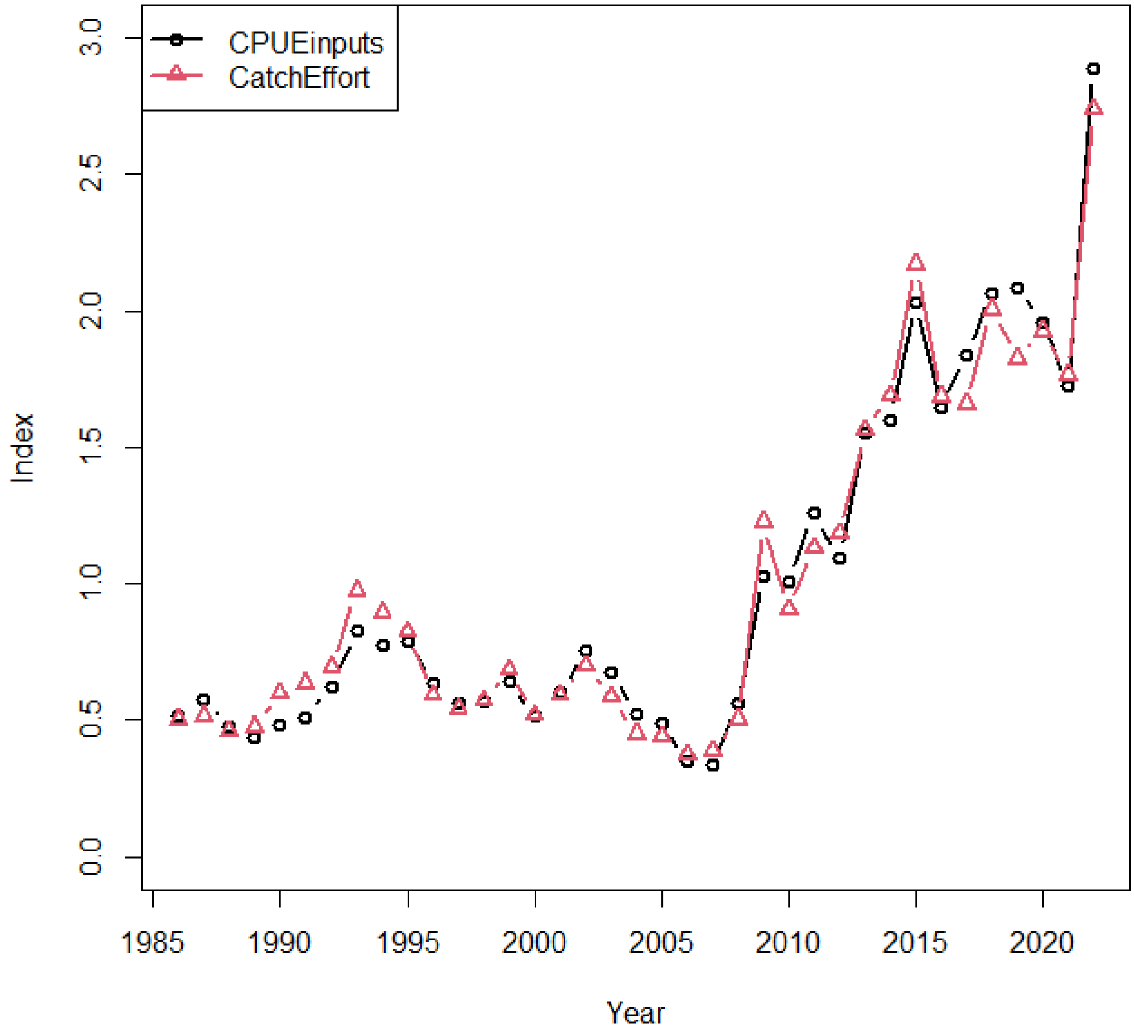

Figure 24: Comparison of indices estimated using the CPUEinputs and CatchEffort files. In both cases strata with fewer than 10000 hooks are omitted, and data are included from the Japanese fleet and Japanese charter vessels fishing under Australian and New Zealand flags.