CCSBT-ESC/2508/11
(ESC Agenda item 7.2)

Development of joint CPUE indices for southern bluefin tuna based on data from multiple fleets

30th Extended Scientific Committee Meeting, CCSBT,
25-30 August 2025

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

Development of joint CPUE indices for southern bluefin tuna based on data from multiple fleets

Simon Hoyle

Hoyle Consulting, simon.hoyle@hoyleconsulting.co.nz.

Working Paper prepared for the 30th Extended Scientific Committee Meeting, CCSBT, 25-30 August 2025.

Executive Summary

Background and Objectives

The Commission for the Conservation of Southern Bluefin Tuna (CCSBT) has historically relied on catch per unit effort (CPUE) indices from the Japanese longline fishery as a primary indicator of southern bluefin tuna (SBT) abundance. Previous analytical approaches using generalized linear models (GLMs) encountered difficulties due to increasing spatial and temporal aggregation of fishing effort, leading to sparse data and parameter estimation problems, particularly in recent years. These problems were resolved via the development of GAM-based approaches using spatial-temporal smoothing. Nevertheless, concerns remain that increasing effort concentration may reduce the accuracy and precision of SBT abundance indices.

This analysis has developed joint CPUE indices using operational longline data from Australian, New Zealand, and Korean fleets, combined with aggregated Japanese fleet data, to address limitations of single-fleet approaches and with the aim of improving the reliability of abundance indices for stock assessment.

Methods

The study incorporated longline catch and effort data from four fishing fleets spanning different time periods. Fleet-specific data processing addressed differences in data formats, coordinate systems, and species reporting. Hierarchical cluster analysis was applied to operational fleet data to identify distinct fishing strategies based on species composition, which may reflect different catchability patterns.

Standardization employed generalized additive models (GAMs) implemented through the R package mgcv, using a delta-lognormal approach to handle zero-inflated catch data. The models incorporated spatial, temporal, and operational covariates, with fleet-specific effects to account for differences in catchability between fishing operations.

Key Findings

Data characterization described the substantial differences in fishing patterns among fleets, including variations in temporal coverage, spatial distribution, gear configurations, and target species composition. Cluster analyses identified between 2-5 distinct fishing strategies within each operational fleet, and meaningful differences in SBT catchability that warranted separate treatment in standardization models.

Preliminary model results suggest that joint fleet analyses can provide abundance indices that maintain consistency with historical patterns while potentially offering improved precision and reduced sensitivity to fleet-specific operational changes.

Implications and Limitations

The multi-fleet approach represents a methodological improvement that may provide more robust abundance indices for southern bluefin tuna stock assessment. The improved spatial and temporal coverage afforded by incorporating multiple fleets could help mitigate biases associated with increasing effort concentration and sparse data in single-fleet analyses.

However, several limitations should be acknowledged. The analysis relies on the assumption that standardization models adequately account for differences in catchability between fleets and fishing strategies. The quality and consistency of data reporting across fleets may vary, potentially affecting model performance. Combining aggregated and operational data in a single analysis is a novel but flawed approach, and the resulting indices should not be seen as representative of stock trends. Representative indices should be based on operational data from all fleets.

Recommendations

Further validation of the joint modeling approach through simulation testing and retrospective analysis would strengthen confidence in the resulting indices. Continued testing and refinement of clustering methods and consideration of additional operational variables could improve the identification of distinct fishing strategies.

The joint CPUE approach developed in this analysis provides a foundation for enhanced abundance monitoring of southern bluefin tuna, though continued evaluation and refinement of methods will be important as additional data become available and fishing patterns continue to evolve.

Introduction

The CPUE standardization methods used for SBT have been updated to address problems with recent CPUE estimates. Analytical problems arose from 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).

A new approach has been developed and adopted (Hoyle, 2021; Hoyle, 2022; Hoyle, 2020; Itoh and Takahashi, 2022) that uses generalized additive models (GAMs) implemented with the R package *mgcv* (Wood, 2011). The principal GAM models produced unbiased estimates with simulated data, while GLM models and less flexible GAM smoothers provided biased indices. This bias was particularly pronounced at the end of the time series as effort became more concentrated, and data became sparse (Hoyle, 2022).

However, simulations indicated that biased indices would result from increasing effort concentration through time, as vessels focused effort on areas with higher CPUE. This effect was strongest 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. ESC 27 concluded that it may be helpful to increase available information via models that include data from other fleets in addition to 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 coverage of the SBT population distribution through time, thereby reducing the risk of parameter estimation difficulties. Results showed that data from other fleets can significantly improve coverage throughout the time series, particularly in recent years. Catch rates of most other fleets showed similar trends to indices from the Japanese fleet.

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., 2024; Hoyle et al., 2018; Hoyle et al., 2015; Kitakado et al., 2021). Such analyses require significant work to prepare data and ensure they are compatible for a joint analysis. Different fishing methods are used by different fleets, and by different groups and vessels within fleets, resulting in variation in catchability.

There is likely 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. These analyses require operational data.

Before jointly analysing national datasets, each dataset needs to 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. It is also necessary to remove effort where there may be issues with reporting quality or the representativeness of the sampling frame.

Work for 2025 involved obtaining, preparing, and analysing operational data for the Australian, Korean, and New Zealand fleets, and combining it with Japanese data to develop joint indices. The specific objectives of this analysis were to:

1. Characterize the fishing patterns and data quality of each fleet
2. Identify distinct fishing strategies within fleets through cluster analysis
3. Develop standardized CPUE indices that account for fleet-specific catchability differences
4. Evaluate the performance and robustness of the joint modeling approach

Methods

Data Sources and Processing

This analysis incorporated longline catch and effort data from four fishing fleets targeting Southern Bluefin Tuna (SBT) in the southern hemisphere. Operational data were provided by the data management agencies for the Australian, New Zealand, and Korean fleets. Aggregated data for the Japanese fleet were obtained from CPUE inputs file provided by the CCSBT.

The operational datasets provided set-level information including species-specific catch, effort, spatial and temporal details, and vessel identifiers, while the Japanese dataset consisted of aggregated catch and effort data by stratum.

Fleet-Specific Data Processing

Korean Fleet (KR): Operational data were obtained from CSV files covering the period 1996-2023. Initial processing involved coordinate conversion from degree-minute format with hemisphere indicators (NS: 1=North, 2=South; EW: 1=East, 2=West) to decimal degrees. Longitudes were standardized to ensure proper range (-180° to 180°).

Species catch data included Southern Bluefin Tuna (SBT), Albacore (ALB), Bigeye (BET), Yellowfin (YFT), Swordfish (SWO), and billfish species (Blue Marlin, Black Marlin, Striped Marlin). Processing approaches were based on the approach in Hoyle et al. (2022).

New Zealand Fleet (NZ): Data were processed from CSV files with set-level information including precise datetime stamps in the Pacific/Auckland time zone. Coordinate conversion from truncated positions to decimal degrees was performed, followed by calculation of 1° and 5° grid references.

Species mapping included conversions from CCSBT codes (STN=SBT, BIG=Bigeye, YFN=Yellowfin, BKM=Black Marlin, BEM=Blue Marlin, STM=Striped Marlin, MAK=Shortfin Mako).

Australian Fleet (AU): Data integration involved merging three Excel files: effort, catch, and bait information. Complex temporal processing included time zone lookup using coordinates, sunrise time calculations relative to set times, and lunar illumination calculations for moon phase effects.

Spatial processing involved longitude transformation (adding 360° to longitudes between -180° and -100°) to maintain continuity across the fishing domain. Vessel identification used unique bt_id values, and species data processing included handling of retained and discarded catch separately.

Japanese Fleet (JP): Aggregated data were processed from text files with stratum-level (month × 5° grid cell) information. The dataset included total hooks and SBT catch by stratum, with CPUE calculated as catch per 1000 hooks. Longitude adjustment involved adding 360° to negative longitudes to generate spatial continuity across the 180° boundary.

Data Filtering and Quality Control

Data filtering was applied to ensure data quality and analytical consistency across fleets:

Effort and catch-based filters:

- Minimum hooks: 200 (AU, NZ), 500 (KR) hooks per set (fleet-specific thresholds)
- Maximum hooks: 4,500 (KR, AU, NZ), hooks per set to remove outliers
- Hooks Between Floats (HBF): 3-50 floats where available, with correction of obvious data entry errors
- At least 50 sets per vessel

Spatial filters:

- Valid coordinates: latitude within ±90°, longitude within ±180° or 0°-360°.
- CCSBT statistical areas 4-9 only (core SBT fishing grounds)
- Latitude > -50° (southern boundary of analysis area)
- Area 9 restriction: latitude <= -35° OR not in area 9 (excluding northern portion)

Temporal filters:

- Operational months 4-9 (April-September, core SBT fishing season)
- Years > 1985 (JP), 1990 (NZ, KR), 2000 (AU)
- Valid date information required for all records

CPUE-based filters:

- Maximum CPUE: catch per 1000 hooks < 200 (AU, NZ, KR) or 120 (JP) to remove extreme outliers

Vessel-based filters:

- Valid vessel identification required
- Consistent vessel codes within fleets
- Minimum activity: vessels with >= 50 sets total, >= 2 years fishing.
- CPUE distribution: vessels with p(positive) >= 0.05 and <= 0.975.

Data Standardization and Harmonization

All datasets were transformed to a common structure with standardized variable names and units:

Coordinate Processing and Spatial Transformations: All fleet datasets required coordinate processing to ensure spatial consistency and compatibility with CCSBT area definitions. Longitudes were transformed to maintain spatial continuity across the Pacific domain. For Australian and New Zealand datasets, longitudes between -180° and -100° were converted by adding 360° (e.g., -175° became 185°). This transformation prevents artificial discontinuities across the 180° meridian where fishing grounds extend across the line.

CCSBT Statistical Area Assignment: All datasets were assigned to CCSBT statistical areas using a standardized lookup function. CCSBT areas 1-15 cover the full range of fishing grounds, with areas 4-9 representing the core SBT fishing areas that were retained for analysis.

Temporal standardization:

- Operation date (op_date) as Date class
- Operation year (op_yr), month (op_mon), and quarter (op_qtr) as integers
- Year factor (year_factor) for categorical modeling

Effort standardization:

- Log-transformed hooks (log_hooks) for modeling
- Hooks Between Floats (HBF) where available

Species standardization:

- Species catch in numbers of fish
- CPUE calculated as SBT per 1000 hooks
- Binary catch indicator (cpue_binary: 0/1)
- Log-transformed CPUE (log_cpue) for positive catches only

Fleet identification:

- Fleet codes: AU, NZ, KR, JP
- Vessel identifiers: anonymized and fleet prefix added (e.g., "AU_123")
- Fleet-specific factors created for modeling

The standardization process resulted in a harmonized dataset with consistent variable definitions, units, and coding schemes across all fleets, enabling joint analysis while preserving fleet-specific characteristics important for CPUE standardization.

Data Characterization

Fleet-specific data characterization examined temporal coverage, spatial distribution, effort characteristics, and species composition patterns. Temporal trends in fishing effort (number of sets, vessels, and hooks) were analysed to identify changes in fleet activity over time. Spatial distributions were mapped to understand fishing ground usage and potential range shifts. Operational characteristics including hook numbers, hooks between floats (HBF), and other gear configurations were summarized to identify fleet-specific practices.

Cluster Analysis

Cluster analysis was conducted to identify distinct fishing strategies within each operational fleet based on species composition. This approach recognizes that different targeting strategies result in different catchability coefficients for SBT, which must be accounted for in CPUE standardization.

Individual fishing sets were aggregated to 10-day trips using vessel and temporal identifiers to reduce the influence of short-term operational decisions and focus on strategic targeting choices. Species catch data were converted to proportions relative to total catch within each trip to standardize for different trip durations and catch magnitudes.

Species were selected for clustering based on their prevalence in each fleet's dataset (minimum 5% of trips) and their importance in the fishery. Core species included Southern Bluefin Tuna (SBT), Bigeye Tuna (BET), Albacore (ALB), Yellowfin Tuna (YFT), Swordfish (SWO), and other billfish species, where they were present in sufficient quantities.

Clustering Methodology

Principal Component Analysis (PCA) was first applied to species proportion data using fourth-root transformation to stabilize variance and reduce the influence of rare species. The optimal number of principal components was determined using the Kaiser criterion and scree plot analysis.

Hierarchical clustering was performed using Ward's linkage method (ward.D2) on Euclidean distances calculated from the transformed species proportions. The optimal number of clusters was determined based primarily on silhouette analysis (Rousseeuw, 1987), while also using the Calinski-Harabasz Index (Caliński and Harabasz, 1974) and the elbow method for validation. Final selection also considered biological interpretability and statistical robustness.

Alternative clustering methods (K-means and CLARA) were evaluated for comparison and validation. The hierarchical clustering approach was selected as the primary method based on superior interpretability and consistency with previous analyses in the region. The CLARA method provided higher silhouette scores in all cases and may be explored in future.

Each cluster was characterized by its mean species composition, spatial and temporal distribution, operational characteristics (mean hooks per trip, HBF), and SBT catch rates.

CPUE Standardization

Model Framework

CPUE standardization was conducted using a delta-lognormal approach implemented with Generalized Additive Models (GAMs). This approach consists of two components: a binomial model for the probability of catching SBT (presence/absence) and a lognormal model for positive catch

rates when SBT was caught. The final standardized index was calculated by multiplying predictions from both components.

Model specifications

Three model configurations were evaluated.
Base model (Model 1): cluster effects as factors.
Fixed vessel model (Model 2): Base model + vessel effects as factors.
Random vessel model (Model 3): Base model + vessel effects as random effects.

The GAM models were implemented using the R package mgcv (Wood, 2017) with the following general structure:

Binomial component (presence/absence):

```
cpue_binary ~ year_factor + cluster_factor + ti(lon, k=20) + ti(lat, k=4) +
              ti(op_mon, k=5) + ti(log_hooks) + ti(lon, lat, k=c(10,4), bs="cs") +
              ti(op_mon, lat, k=c(5,4), bs="cs") + ti(lon, op_mon, k=c(10,5), bs="cs") +
              ti(op_yr, lat, k=c(3,4), bs="cs") + ti(op_yr, op_mon, k=c(3,5), bs="cs") +
              ti(lon, op_yr, k=c(10,3), bs="cs")
```

Positive component (lognormal):

```
log(cpue) ~ year_factor + cluster_factor + ti(lon, k=20) + ti(lat, k=4) +
              ti(op_mon, k=6) + ti(log_hooks) + ti(lon, lat, k=c(20,4), bs="cs") +
              ti(op_mon, lat, k=c(6,4), bs="cs") + ti(lon, op_mon, k=c(20,6), bs="cs") +
              ti(op_yr, lat, k=c(20,4), bs="cs") + ti(op_yr, op_mon, k=c(20,6), bs="cs") +
              ti(lon, op_yr, k=c(20,20), bs="cs") + ti(lat, lon, op_mon, k=c(4,4,4), bs="cs") +
              ti(op_yr, lon, op_mon, k=c(4,4,4), bs="cs") + ti(lat, lon, op_yr, k=c(4,4,4), bs="cs") +
              ti(lat, op_mon, op_yr, k=c(4,3,3), bs="cs")
```

The positive model includes four three-way tensor product interactions to capture complex spatio-temporal relationships in SBT catch rates when present. The binomial model uses only two-way interactions to avoid convergence issues while maintaining adequate model complexity for presence/absence patterns.

Where:

- year_factor: categorical year effect (the target abundance signal)
- cluster_factor: fishing strategy clusters from species composition analysis
- ti(): tensor product interactions using cubic splines with shrinkage (bs="cs")
- k: basis dimension controlling smoothness
- op_yr, op_mon: continuous operation year and month
- log_hooks: log-transformed hook numbers

Model Implementation

Models were fitted using the mgcv package in R (Wood, 2017), with restricted maximum likelihood (REML) estimation for both components. Following Wood's recommendations for multi-level interactions, all terms were specified using tensor product interactions (ti()) to ensure proper separation of main effects and interactions.

Binomial models included effort (log_hooks) to account for the effect of effort on the probability of encountering SBT. Both models used cubic spline smooths with shrinkage (bs="cs"), allowing automatic variable selection by penalizing terms to zero when not supported by the data.

Both model components were fitted with a gamma=2 smoothing parameter multiplier to increase the smoothing penalty and reduce overfitting.

```
binomial_model <- mgcv::gam(formula, family = binomial(),
                    data = data, method = "REML", gamma = 2)
positive_model <- mgcv::gam(formula, family = gaussian(),
                    data = positive_data, method = "REML", gamma = 2)
```

Model Selection and Evaluation

Multiple model configurations were tested, varying in the inclusion of:

- Cluster assignments (included vs. excluded)
- Vessel effects (fixed vs. random)
- Interaction complexity

Model selection was based on multiple criteria including:

- Akaike Information Criterion (AIC)
- Deviance explained
- Residual diagnostics
- Biological plausibility of trends

Index Calculation

Standardized annual indices were calculated by predicting abundance in each fished stratum (year × month × 5° spatial cell) and aggregating across strata weighted by ocean area. Reference conditions for prediction included median values for continuous variables and the most common observation for factors.

Ocean areas for each 5° × 5° grid cell were calculated following Hoyle and Langley (2020), with predictions limited to cells that were fished during the study period. Annual indices were normalized to have a geometric mean of 1.0 across the time series.

Uncertainty Estimation

Standard errors for annual indices were calculated using the delta method to propagate uncertainty from both model components. Confidence intervals were constructed assuming log-normal distributions for the annual index values.

Diagnostic Assessment

Model diagnostics included examination of:

- Residual patterns (temporal, spatial, by fitted values)
- Quantile-quantile plots for distributional assumptions
- Influence measures for outlier detection
- Partial effects plots for smooth terms
- Basis dimension adequacy using gam.check()

Sensitivity Analysis

Sensitivity of the standardized index was evaluated by varying reference conditions for key variables including:

- Reference month (seasonal baseline)
- Reference hook numbers (effort baseline)
- Reference spatial location (geographic baseline)
- Cluster assignment (targeting strategy baseline)

Software and Packages

All analyses were conducted in R version 4.3.3 using the following primary packages:

- mgcv (Wood, 2017): GAM model fitting and diagnostics
- tidyverse (Wickham et al., 2019): data manipulation and visualization
- cluster: clustering algorithms and validation
- factoextra (Kassambara and Mundt, 2017): cluster visualization and assessment
- fastcluster (Müllner, 2013): rapid implementation of hclust in particular.
- corrplot: correlation analysis and visualization
- viridis: color palettes for visualization
- lubridate (Grolemund and Wickham, 2011): date and time manipulation
- gridExtra and ggpubr: plot arrangement and publication formatting

Additional packages for specialized analyses included DHARMa (Hartig, 2020) for model diagnostics, gratia and visreg (Breheny and Burchett, 2017) for GAM visualization, and future for parallel processing of computationally intensive model fitting procedures.

Models were run using the OpenBLAS library to improve the efficiency of matrix calculations (Xianyi et al., 2012).

Results

Characterization

Temporal Coverage and Fleet Characteristics

Data from Australia, New Zealand and Korea were available since 2000, 1990, and 1994 respectively (Figure 1). After cleaning the final dataset included 17742, 68507, and 9588 sets from Australia, New Zealand, and Korea respectively (Table 1), and 3642 strata in the aggregated Japanese dataset. The number of fish caught by species and fleet are provided in Table 2.

Operational patterns and targeting behaviour

Much of the effort in the Australian and New Zealand datasets had low SBT catch rates (Figure 2) and was likely targeted towards species other than SBT. Hooks per set tended to be lower in the relatively small-scale vessels operating in the Australian and New Zealand fleets, compared to the Korean fleet (Figure 3).

Fishing seasonality varied among fleets, as expected given differences in targeting approaches and fishing locations (Figure 4). The number of active vessels in the Australian fleet declined considerably from 2000 to 2010, while the Korean fleet was relatively stable through time, and the New Zealand fleet showed more variability (Figure 5). The Australian and New Zealand fleets showed significant variability in the proportions of positive sets, while the relatively targeted Korean and Japanese fleets were much more stable (Figure 6).

Species composition patterns

Species compositions differed between fleets (Figure 7). The Australian and New Zealand fleets caught high proportions of albacore, which increased through time in the Australian fleet but declined from 2005 in the New Zealand fleet. The Korean fleet showed a large increase in the proportion of albacore catch between approximately 2006 and 2015, almost inversely related to SBT, with the exception of a pulse of yellowfin tuna in 2004-2005.

Yellowfin tuna was significant for the Australian fleet but declined as a proportion from about 2011 as SBT catch increased. SBT catch also increased substantially in the NZ fleet, from a low point in 2006, to reach its highest proportion in the most recent year.

Fleet participation and spatial distribution

Fleet participation by vessels revealed distinct patterns (Figure 8). There was an early loss of vessels in the Australian fleet with limited replacement, relatively steady turnover in the NZ fleet, and numerous KR vessels with very short periods of participation. These differences likely reflect the differing natures of the fleets and the data provision process.

The domestic Australian and New Zealand fleets tend to continue fishing in the same jurisdictions and areas (Figure 9), and the longline effort provided was not restricted to vessels that caught SBT. Vessels in the distant water Korean longline fleet may be less constrained to fishing in the same parts of the ocean or to targeting SBT. Importantly, the KR longline effort provided included only vessels that caught some SBT in a given year.

Cluster analysis

Silhouette scores and Calinski-Harabasz indices suggested 2-3 clusters for the Australian fleet and 2-4 clusters for the New Zealand fleet (Figure 10). A minimum selection of three clusters was applied, because overestimating the number of clusters is preferable to underestimating when accounting for targeting heterogeneity. Five clusters were supported for the Korean fleet.

SBT catch rates differed substantially between clusters, particularly in the Australian and New Zealand datasets (Figure 11). These differences demonstrate meaningful variation in SBT targeting intensity among operational strategies within each fleet.

In the Australian dataset, cluster 3 with high SBT catch rates was dominated by SBT and albacore (Figure 12). Cluster 2 was dominated by albacore, while cluster 1 was dominated by yellowfin tuna.

In the New Zealand dataset, the cluster 2 had high SBT catch rates and was similar to the Australian cluster 3 in being dominated by SBT and albacore. The other two clusters were dominated y albacore (cluster 1) and swordfish (cluster 3).

In the Korean dataset, clusters 1, 2, and 5 were all dominated by SBT, while clusters 3 and 4 were dominated by yellowfin and albacore respectively.

The temporal patterns of cluster allocation (Figure 13) reflect the patterns of species composition seen in Figure 7. The Australian SBT-focused cluster 3 expanded starting in 2013. The NZ SBT-focused cluster 3 contracted through 1995 and then expanded from 2011 to 2022. The KR SBT-focused cluster 1 became the only cluster represented in that dataset from 2016 onward.

Modeling

Model fitting and selection

All three model configurations fitted successfully. Model fitting times varied considerably. In Model 1 base, the binomial took 3 minutes to converge while the positive model took 25 minutes. For Model 2 vessel fixed, the binomial model fitted in 21 minutes, and the positive model took 58 minutes. For Model 3 vessel random, the binomial model took 14 minutes to converge, and the positive model took 60 minutes.

The model with fixed vessel effects was selected as the preferred model because it had the best AIC value in both the binomial and positive components (Table 3). All model components were highly significant (Tables 4 to 7).

Model performance and diagnostics

Model diagnostics were expected to show some lack of fit, given the dataset included both operational and aggregated data. These data types have different error distributions, and the error distribution of aggregated data is always problematic because strata with higher CPUE tend to have more sets, and therefore smaller variance.

The binomial model fit was relatively good with only minor deviations from expected patterns (Figure 14). The positive model showed greater signs of variation, with particular differences by fleet, which was anticipated given the heterogeneity in fishing operations spatial distribution, and data aggregation across fleets.

Plots showing expected values and partial effects of individual covariates, while holding other parameters constant (Figure 15), showed very strong spatial patterns, particularly in the binomial component. Clusters also had a large explanatory effect. There were surprising and somewhat inconsistent patterns in the vessel effects, with higher rates of positive catches for the Australian fleet. This may reflect the relative confounding between vessels and clusters at the fleet level. These patterns were not observed in the same plots for model 3 which used random rather than fixed effects for the vessel ids. However, these issues are not expected to affect the indices.

Residual boxplots for latitude, longitude, month, and year (Figure 16) showed substantial patterns associated with longitude, and minor patterns associated with latitude and month in the positive component. The longitude patterns may indicate that more flexibility is needed in the basis dimensions (k values) used in the model. There was insufficient time to explore a range of k values for the main effects and the interaction terms.

DHARMa residual plots by variable (Figure 17) were consistent with the box plot diagnostics, showing moderate departure from model assumptions for the positive component.

Plots of the smooth model components are provided in figures 18 (binomial) and 19 (positive). They include similar information to Figure 15, but without the partial effects and with the addition of the interaction terms.

Standardized indices from the 3 model configurations are presented in Figure 20. Results are very similar across all three models, indicating relatively small impact associated with the treatment of vessel effects. However, substantial uncertainty is evident at the end of the time series, reflecting reduced data coverage and increased fleet concentration in recent years.

The estimated trend appears to differ from the Japanese index for the period 1990-2022 (Itoh and Takahashi, 2025). These differences likely reflect the influence of the Australian and New Zealand datasets. improved spatial and temporal coverage provided by the multi-fleet approach, but require careful interpretation given the methodological differences between the analyses.

Discussion

This analysis, which combines data from multiple fleets in a single standardization analysis, was motivated by the progressive spatial and temporal concentration of the Japanese fleet through time. This increasing concentration reduces the amount of information available to the model by reducing the number of strata with information about catch rates. Loss of information can also cause instability in model estimates (e.g., Hoyle, 2021) and tends to increase uncertainty in the indices of abundance, as estimates increasingly rely on information sharing among adjacent strata.

A further concern is the theoretical risk that vessels are partly achieving this greater effort concentration by increasingly using information obtained from one another, from remote sensing,

and from oceanographic models If so, this technological enhancement may improve their ability to focus on areas with higher catch rates and avoid areas likely to have low catch rates. Such a process would tend to positively bias CPUE estimates over time.

Combining information from multiple fleets, rather than relying on data from Japan alone, has the potential to the improve the analysis by increasing the number of informative strata available to the analysis. The multi-fleet approach also provides additional spatial and temporal coverage that can help maintain representativeness of the fishing grounds throughout the time series.

Several unique difficulties were associated with this analysis. The first challenge was the highly distinct spatial coverage of each dataset. The New Zealand dataset is based entirely in the New Zealand region in the far east of the spatial domain. The Australian dataset covers an even smaller areas based in eastern Australia and largely north of 40°S. The Korean dataset has broad coverage in the west but has no spatial overlap with either the Australian or the New Zealand datasets. The Japanese aggregated dataset has by far the broadest overage, covering almost all spatial cells covered by the other datasets, plus additional areas.

Lack of overlap between the operational datasets was problematic for the analysis, because overlap throughout the dataset is needed to avoid singularities and ensure that parameters are estimable. Spatial overlap makes it possible to estimate the relative catch rates among areas without making unreasonable assumptions. This problem was resolved by including the aggregated Japanese data in the analysis, which provided the necessary spatial connectivity.

A further difficulty was the need to jointly analyse operational and aggregated data in the same model, which raised issues around parameter estimability and data weighting. The aggregated dataset lacked vessel ids and cluster allocations, which were assigned as dummy values. All Japanese effort was allocated the same vessel id and cluster factor, which initially led to lack of identifiability due to perfect collinearity in the vessel-as-factor model. This was resolved by using the Japanese vessel and cluster as the reference levels for each factor in the GAM, which set their values to zero.

Calibrating data weights between datasets was also problematic. Equal weights were initially given to all sets and to all strata, although each row of aggregated data is more informative about catch rates than individual operational sets, because aggregated data combines catch rates from multiple sets. However, aggregated data rows are not necessarily more informative in binomial regression, because each stratum provides only one piece of information: whether the aggregated stratum is zero or non-zero. Effort-dependent data weights were therefore omitted from the binomial model, though they should be considered in future in the positive model.

Estimation difficulties occurred when fitting the binomial model, associated with vessels with extreme catch success rates (100%, very high proportions, very low proportions, or 0% of non-zero catches). Vessels that either always or never catch something are problematic in fixed-effect binomial models because they have infinite parameter values and provide no useful information to the model. Vessels that almost always or almost never catch something are also problematic, because they can cause numerical instability, and they provide little useful information to the model. This separation problem was resolved by restricting the data set to vessels making at least 50 sets, fishing in at least 2 years, and with between 5% and 97.5% positive catches. However, vessels with more than 97.5% positive catches may provide useful information to the positive model, and their inclusion should be considered in future analyses.

The basis dimensions (k values) were initially set to the same values used in the Japanese GAM (Itoh and Takahashi, 2025), with reductions for the op_yr effects to allow for the shorter time series.

However, the positive model failed to converge (still running after 5+ hours) until the initial degrees of freedom (k values) in the three-way interaction terms were reduced considerably. This problem may have been caused by variation through time in the spatial distribution of data among fleets. For example, non-zero data from the Australian fleet were particularly sparse until about 2010. Restricting the flexibility of the three-way interactions allowed the model to fit successfully, but further model exploration is recommended.

Model diagnostics were expected to show some lack of fit given that the dataset included both operational and aggregated data. These data types have different error distributions, and the error distribution of aggregated data is inherently problematic because strata with higher CPUE tend to have more sets and therefore smaller variance. The fit of the binomial model was relatively good with only minor deviations from expected patterns. The positive model showed greater signs of variation, with particular differences by fleet, which was anticipated given the heterogeneity in fishing operations across fleets.

Cluster assignments showed large explanatory effects, confirming the importance of accounting for different targeting strategies in multi-fleet analyses. Spatial patterns were very strong, particularly in the binomial component. Vessel effects showed some unexpected patterns, with higher rates of positive catches for the Australian fleet. This may reflect confounding between vessels and clusters at the fleet level. Notably, these patterns were not observed in Model 3, which used random rather than fixed effects for vessel identifiers. Residual patterns indicated that additional flexibility may be needed in the longitude terms, but time constraints prevented full exploration of alternative basis dimensions for the main effects and interaction terms.

The resulting indices are to some extent different from the Japanese indices (Itoh and Takahashi, 2025) during the same period, with variation at different times and a larger increase at the end of the time series. Some differences are expected given the different data inputs, but more validation is needed to make sure that the results are credible. The combination of aggregated and operational data in a single analysis, while necessary for spatial connectivity, represents a methodological compromise. Representative indices would ideally be based on operational data from all fleets, though this is not currently feasible given data availability constraints. A likely contributor to the differences is the relatively low statistical weighting given to the Japanese data, since each Japanese stratum receives the same weight as a single set from the other 3 fleets. It would be useful to explore giving higher weights to the aggregated data, since each of these strata represents between 4 and 1600 sets, with an average of 88 sets.

These analyses were made feasible by using OpenBLAS libraries, which speed up matrix operations considerably compared to the standard Basic Linear Algebra Subprograms (BLAS) libraries. Otherwise, GAM analyses of such large datasets take far longer and can be impractical. OpenBLAS, Intel MKL, and the BLAS implementation used in Microsoft R Open are all optimized implementations. Microsoft R Open is no longer being developed, and the most recent version is based on R 4.0.2 from September 2020.

This project has successfully developed a new R codebase for joint analysis of longline CPUE data using GAMS; and used it to generate joint indices of abundance by combining data from multiple fleets. Further testing and development will be required if the CCSBT decides to continue down this path. Ideally, the aggregated Japanese data should be replaced with operational data. The index may have a long-term role as a comparator to the Japanese index, or it may be developed as a replacement if its performance is found to be superior.

Recommendations for Future Work

1. Enhanced data analyses: Work towards developing collaborations that can use operational data from all fleets, to eliminate the need for mixed data types.
2. Model refinement: Further explore clustering methods, basis dimensions, model complexity, and data weighting approaches to optimize fit while maintaining stability.
3. Simulation testing: Validate the multi-fleet approach through simulation studies to better understand potential biases.
4. Compare GAM approach with alternative spatiotemporal models such as tinyVAST or sdmTMB.

Declaration of generative AI use

A generative artificial intelligence (AI) assistant, Anthropic Claude Sonnet 4.0, was used to parse R code during model development. It was also used to provide a first draft of the methods section.

Bibliography

Breheny, P.; Burchett, W. Visualization of regression models using visreg. The R Journal. 9:56-91; 2017.

Caliński, T.; Harabasz, J. A dendrite method for cluster analysis. Communications in Statistics-theory and Methods. 3:1-27; 1974.

Grolemund, G.; Wickham, H. Dates and times made easy with lubridate. Journal of statistical software. 40:1-25; 2011.

Hartig, F. DHARMa: residual diagnostics for hierarchical (Multi-Level/Mixed) regression models. 2020

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. 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.; Campbell, R.A.; Ducharme-Barth, N.D.; Grüss, A.; Moore, B.R.; Thorson, J.T.; Tremblay-Boyer, L.; Winker, H.; Zhou, S.; Maunder, M.N. Catch per unit effort modelling for stock assessment: A summary of good practices. Fisheries Research. 269:106860; 2024.

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.; 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; 2015

Itoh, T.; Takahashi, N. Development of the new CPUE abundance index using GAM for southern bluefin tuna in CCSBT. CCSBT-OMMP/2206/08. CCSBT Operating Model and Managament Procedure Meeting. Hobart, Tasmania; 2022

Itoh, T.; Takahashi, N. Update of CPUE abundance index using GAM for southern bluefin tuna in CCSBT (GAM22) up to the 2024 data. CCSBT-OMMP/2507/06. 15th Operating Model and Management Procedure Technical Meeting. Online: Center for the Conservation of Southern Bluefin Tuna; 2025

Kassambara, A.; Mundt, F. Package 'factoextra'. Extract and visualize the results of multivariate data analyses. 76:10.18637; 2017.

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.

Müllner, D. fastcluster: Fast hierarchical, agglomerative clustering routines for R and Python. Journal of Statistical Software. 53:1-18; 2013.

Rousseeuw, P.J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics. 20:53-65; 1987.

Wickham, H.; Averick, M.; Bryan, J.; Chang, W.; McGowan, L.D.A.; François, R.; Grolemund, G.; Hayes, A.; Henry, L.; Hester, J. Welcome to the Tidyverse. Journal of open source software. 4:1686; 2019.

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.

Wood, S.N. Generalized additive models: an introduction with R: Chapman and Hall/CRC; 2017

Xianyi, Z.; Qian, W.; Yunquan, Z. Model-driven level 3 BLAS performance optimization on Loongson 3A processor. 2012 IEEE 18th international conference on parallel and distributed systems: IEEE; 2012

Tables

Table 1: Number of records available by fleet before and after filtering, quality control, and removal of records outside the key period of April to September and SBT areas 4 to 9.

	AU	NZ	KR
Initial records	45758	114802	33518
After filtering		112880	33507
After QC	45507	112098	33504
After area/month	21975	73770	23806
After 5% - 97.5% positive vessels	17742	68507	9588

Table 2: Number of fish recorded by fleet and by species in the final dataset

	AU	NZ	KR
sbt	112,643	282,299	147,083
alb	314,834	1,457,321	78,759
bet	45,371	57,157	9,884
yft	197,829	7,100	33,399
swo	39,623	142,171	1,414
mls	9,499	2,057	36
bsh	972		
sma	7,349	65,671	
ceo	21,612		
dol	11,013		
poa	45,894		
blm			10
bum			19

Table 3: Comparison of AIC values across the three alternative models.

Model	**AIC Positive**	**AIC Binomial**	**AIC Total**	**dAIC**
model1_base	125944.04	56331.27	182275.31	1370.3
model2_vessel_fixed	125054.97	55850.04	180905.01	0
model3_vessel_random	125206.16	55912.17	181118.33	213.32

Table 4: Summary of parametric terms and associated statistics from the binomial sub-model of model 2.

Term	df	Chi-square	p-value
year_factor	32	1581	<0.001
cluster_factor	11	2129	<0.001
vessel_factor	290	1034	<0.001

Table 5: Summary of smooth terms and associated statistics from the binomial sub-model of model 2 fixed.

Term	edf	Ref.df	Chi-square	p-value
ti(lon)	16.017	17.345	542.32	<0.001
ti(lat)	2.603	2.804	933.98	<0.001
ti(op_mon)	3.826	3.953	1197.81	<0.001
ti(log_hooks)	1.003	1.007	178.15	<0.001
ti(lon,lat)	16.675	27	633.00	<0.001
ti(op_mon,lat)	7.865	12	376.29	<0.001
ti(lon,op_mon)	23.861	36	428.97	<0.001
ti(op_yr,lat)	4.895	6	78.95	<0.001
ti(op_yr,op_mon)	6.469	8	85.43	<0.001
ti(lon,op_yr)	14.006	18	623.48	<0.001

Table 6: Summary of parametric terms and associated statistics from the positive sub-model of model 2 fixed.

Term	df	F	p-value
year_factor	32	63.875	<0.001
cluster_factor	10	290.620	<0.001
vessel_factor	291	5.373	<0.001

Table 7: Summary of smooth terms and associated statistics from the positive sub-model of model 2 fixed.

Term	edf	Ref.df	F	p-value
ti(lon)	14.945	16.206	51.894	<0.001
ti(lat)	2.796	2.919	153.023	<0.001
ti(op_mon)	4.711	4.901	131.063	<0.001
ti(lon,lat)	36.405	57	13.987	<0.001
ti(op_mon,lat)	7.939	15	15.405	<0.001
ti(lon,op_mon)	21.837	36	9.712	<0.001
ti(op_yr,lat)	27.861	33	13.403	<0.001
ti(op_yr,op_mon)	25.626	44	7.000	<0.001
ti(lon,op_yr)	50.752	72	8.367	<0.001
ti(lat,lon,op_mon)	4.005	27	1.852	<0.001
ti(op_yr,lon,op_mon)	18.455	27	12.586	<0.001
ti(lat,lon,op_yr)	12.815	27	7.539	<0.001
ti(lat,op_mon,op_yr)	4.713	12	5.511	<0.001

Table 8: •Annual indices and standard errors (SE) for each year, presented for Model 1 base, Model 2 fixed, and Model 3 random.

	Model 1		Model 2		Model 3	
Year	Index	SE	Index	SE	Index	SE
1990	0.4294	0.0106	0.4001	0.0085	0.4352	0.0102
1991	0.3218	0.0062	0.3411	0.0058	0.3323	0.0061
1992	0.3682	0.0057	0.4086	0.0057	0.3763	0.0055
1993	0.4071	0.0063	0.4454	0.0063	0.4151	0.0062
1994	0.7369	0.0120	0.6799	0.0107	0.7398	0.0118
1995	0.6574	0.0106	0.6597	0.0104	0.6783	0.0108
1996	0.5248	0.0096	0.4966	0.0089	0.5225	0.0095
1997	0.4973	0.0088	0.5203	0.0087	0.5140	0.0087
1998	0.5018	0.0097	0.5232	0.0091	0.5166	0.0094
1999	0.5337	0.0103	0.5330	0.0090	0.5408	0.0097
2000	0.5127	0.0097	0.5178	0.0084	0.5217	0.0090
2001	0.5381	0.0098	0.5220	0.0081	0.5382	0.0090
2002	0.4941	0.0085	0.4900	0.0072	0.5026	0.0080
2003	0.3999	0.0074	0.3839	0.0060	0.4049	0.0070
2004	0.3573	0.0071	0.3440	0.0057	0.3637	0.0067
2005	0.3973	0.0084	0.3758	0.0065	0.3979	0.0076
2006	0.3860	0.0086	0.3519	0.0061	0.3768	0.0074
2007	0.4339	0.0096	0.4089	0.0068	0.4243	0.0080
2008	0.7404	0.0150	0.6883	0.0105	0.7264	0.0126
2009	0.6960	0.0134	0.6533	0.0099	0.6928	0.0118
2010	0.7438	0.0138	0.7168	0.0114	0.7520	0.0130
2011	0.7959	0.0141	0.7668	0.0122	0.8044	0.0136
2012	0.9608	0.0170	0.9497	0.0158	0.9839	0.0167
2013	1.0987	0.0223	1.1011	0.0216	1.1239	0.0219
2014	1.3907	0.0348	1.3800	0.0335	1.4048	0.0334
2015	1.6993	0.0485	1.7192	0.0467	1.7332	0.0460
2016	2.0210	0.0698	2.0080	0.0644	2.0159	0.0636
2017	1.8352	0.0730	1.8048	0.0647	1.8152	0.0648
2018	2.0700	0.0950	2.0752	0.0828	2.0714	0.0835
2019	1.5098	0.0973	1.4869	0.0825	1.4724	0.0834
2020	1.7516	0.1601	1.7410	0.1378	1.7131	0.1373
2021	2.7194	0.4100	2.7961	0.3715	2.6826	0.3578
2022	4.4697	1.1229	4.7103	1.0731	4.4071	0.9981

Figures

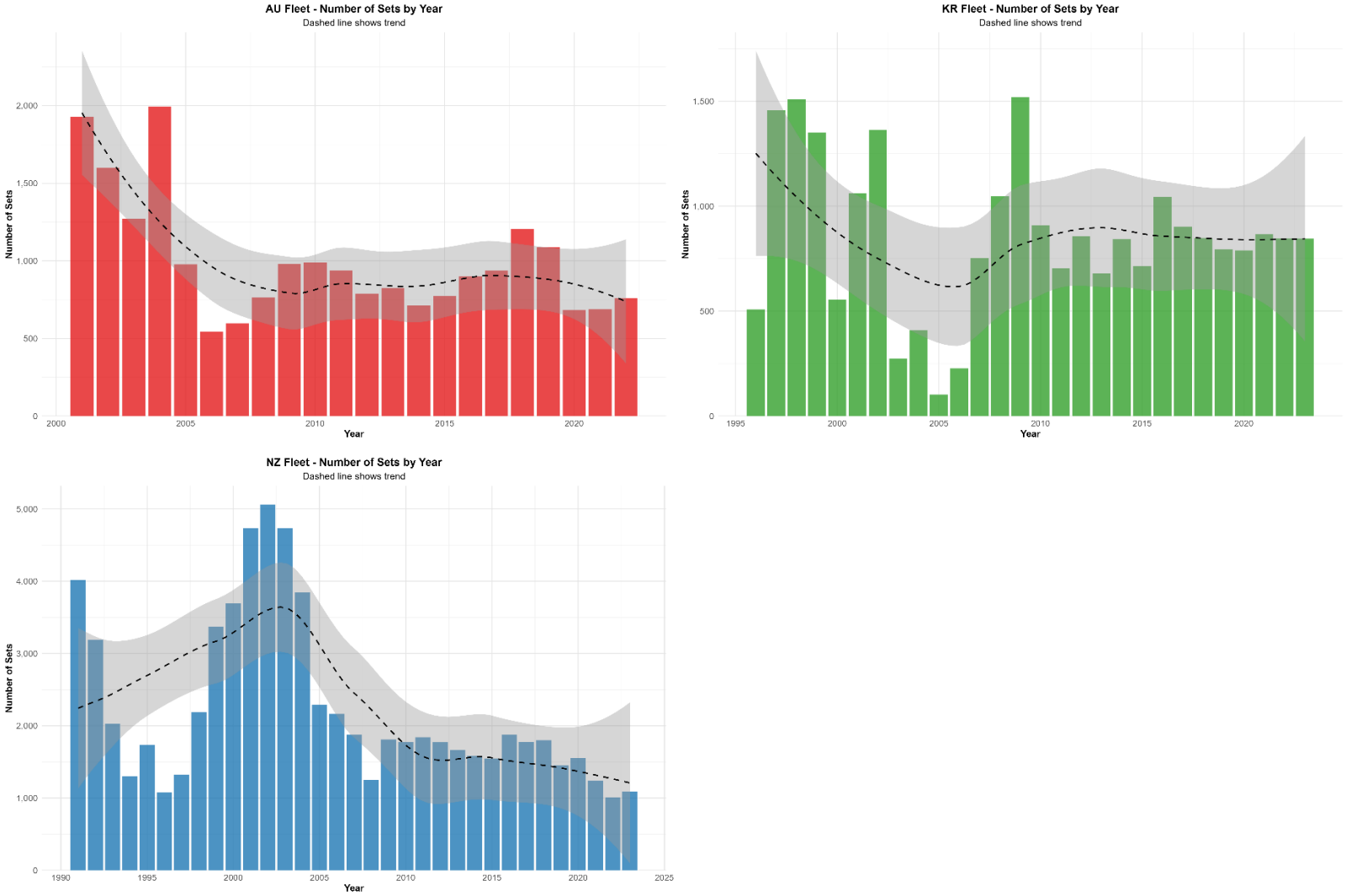

Figure 1: Time series plots showing number of sets per year for operational fleets (AU, KR, NZ).

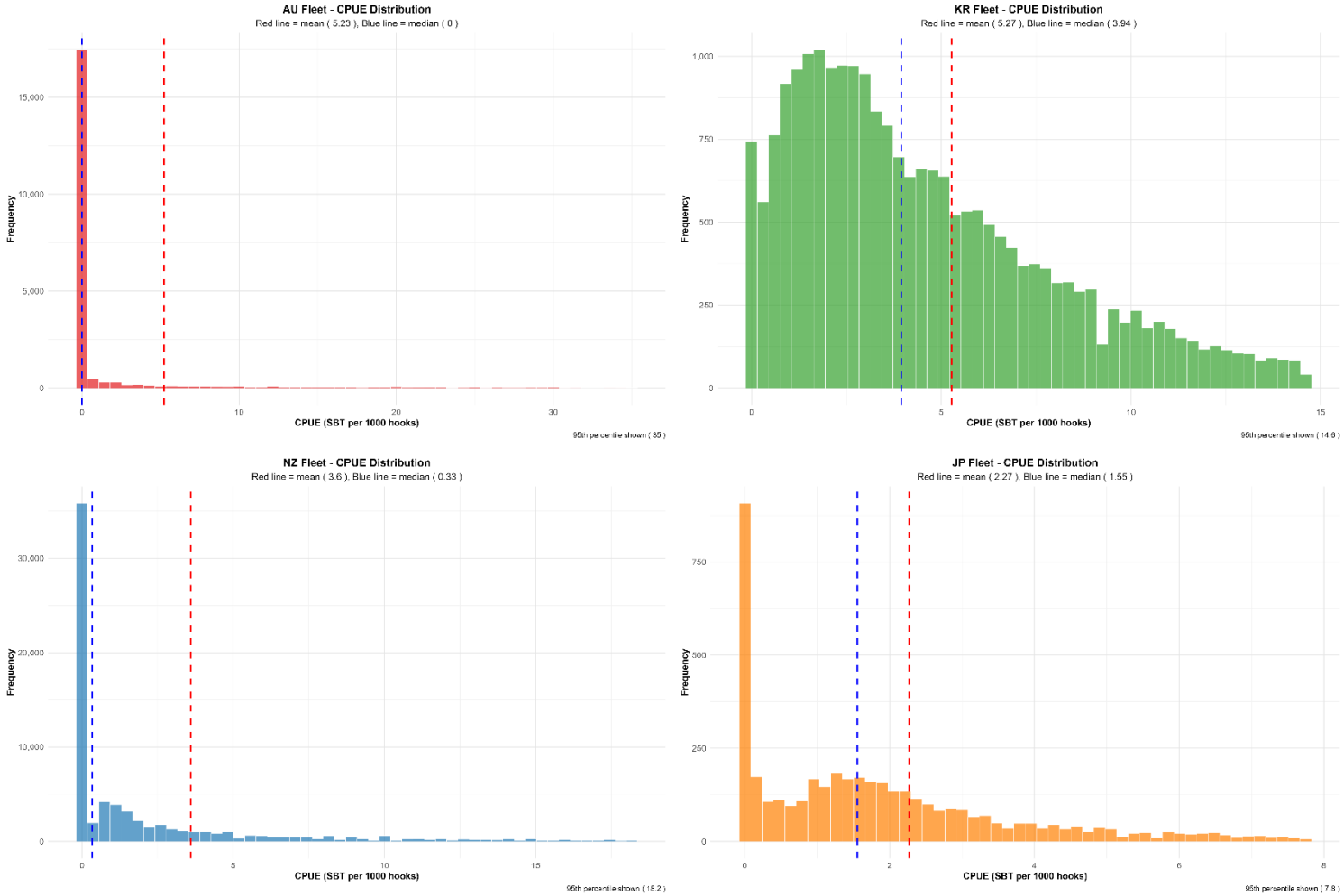

Figure 2: Frequency histograms of CPUE (SBT per 1000 hooks) for operational fleets (AU, KR, NZ) by individual set and aggregated Japanese (JP) fleet by stratum.

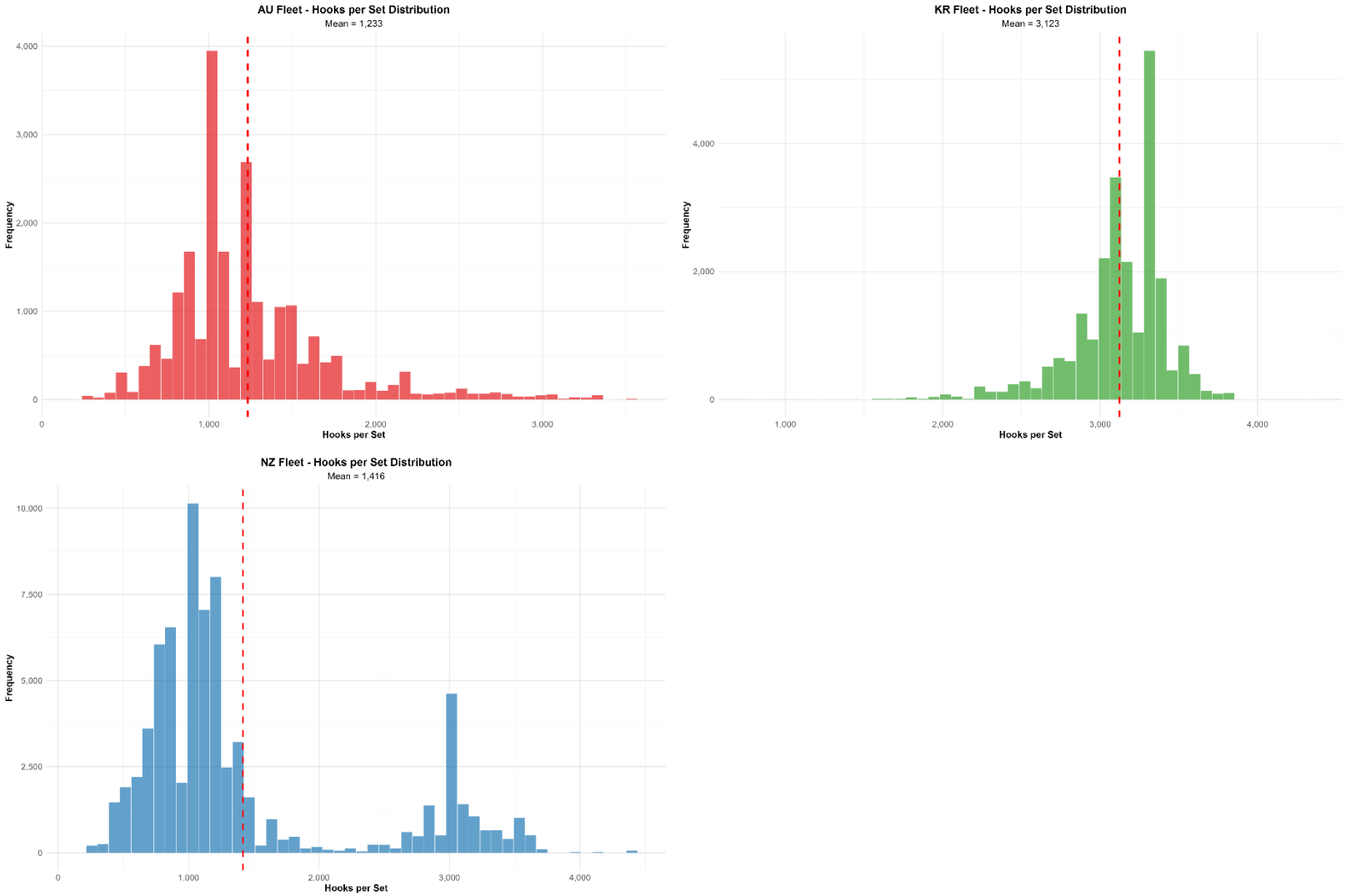

Figure 3: Frequency histograms showing distribution of hooks per set for operational fleets (AU, KR, NZ).

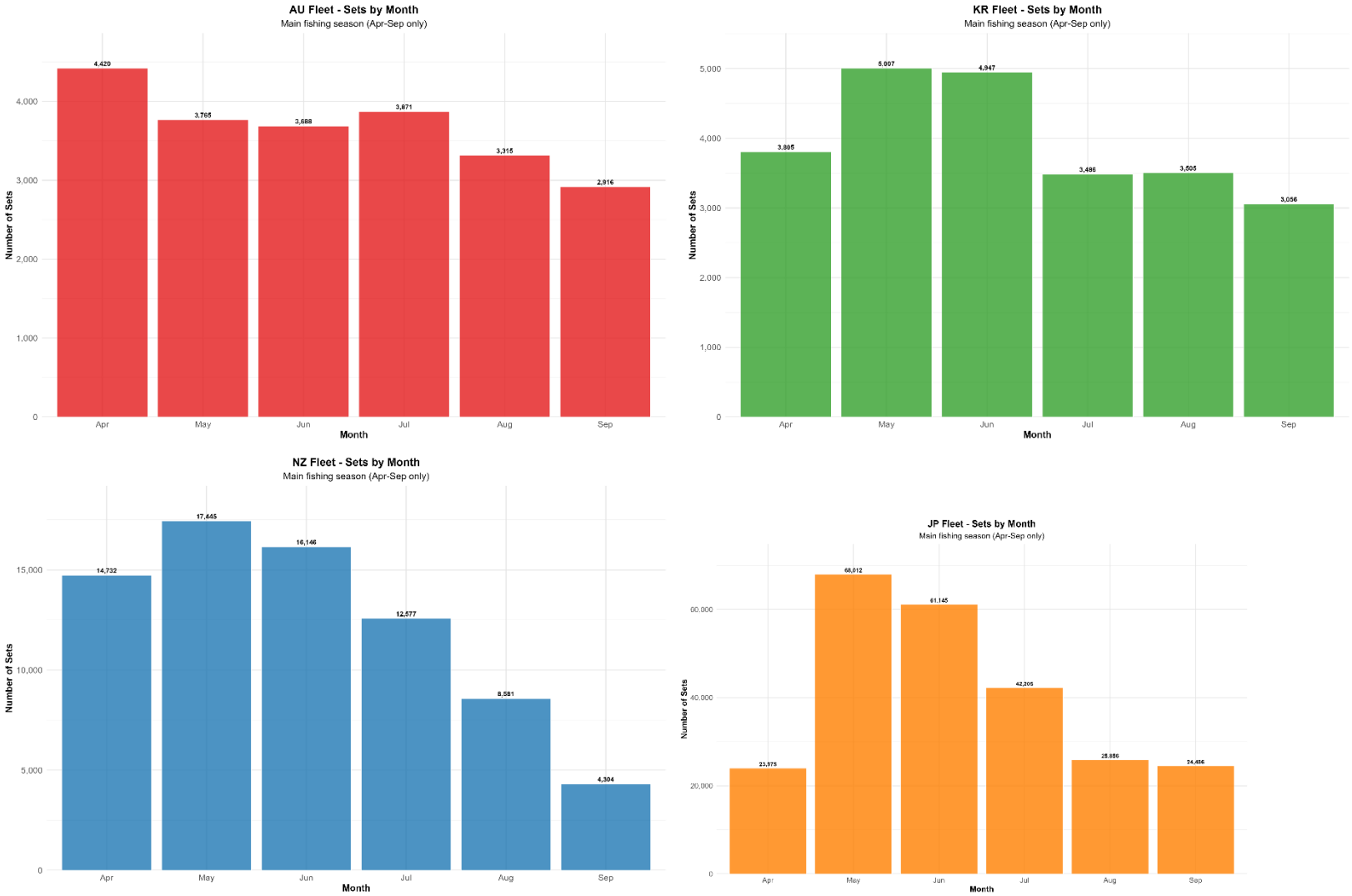

Figure 4: Bar plots showing monthly distribution of sets for operational fleets (AU, KR, NZ) and for the aggregated Japanese (JP) fleet, aggregated across all years. Japanese set counts are calculated based on 3000 hooks per set.

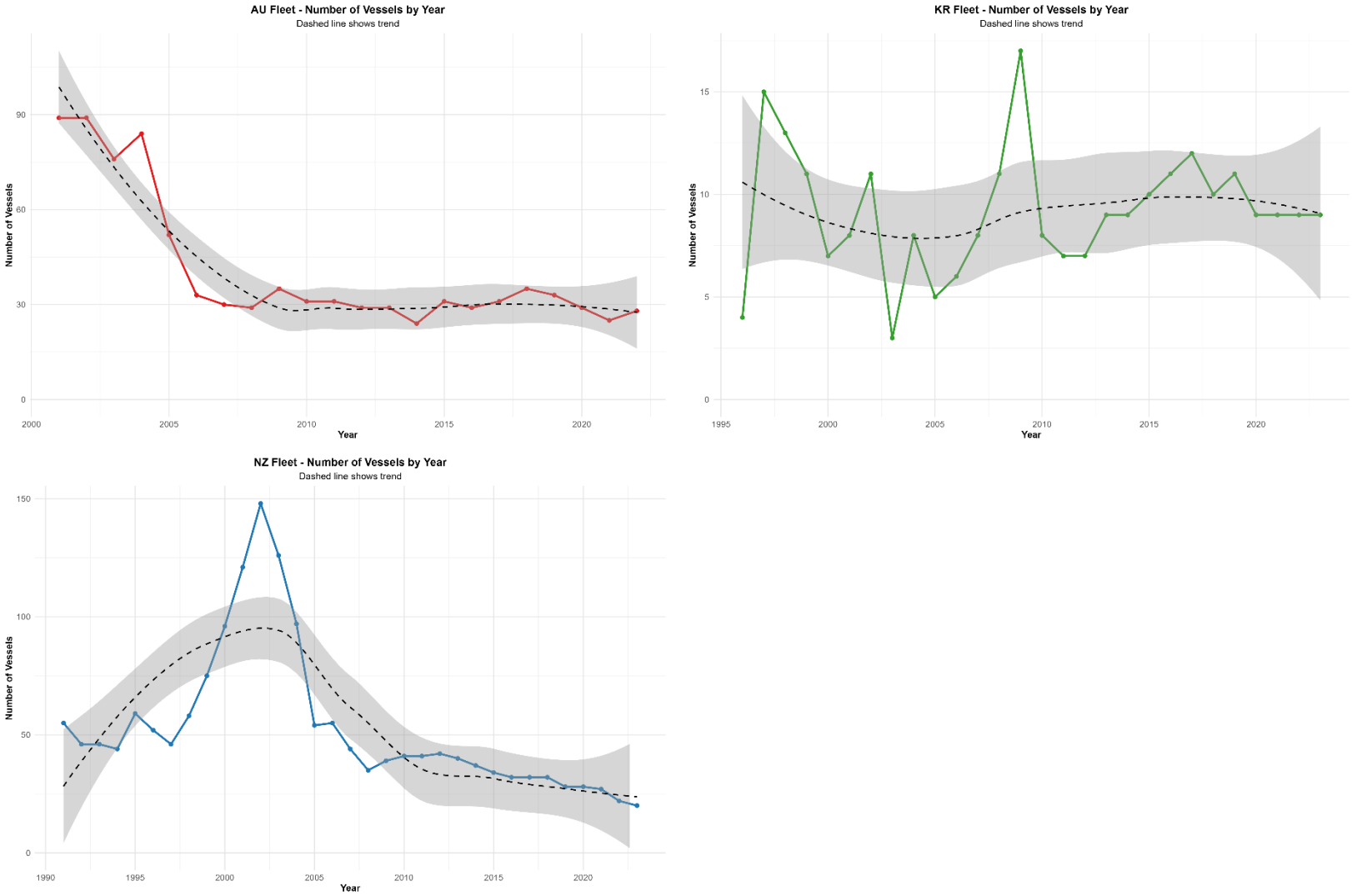

Figure 5: Time series plots showing number of active vessels per year for operational fleets (AU, KR, NZ).

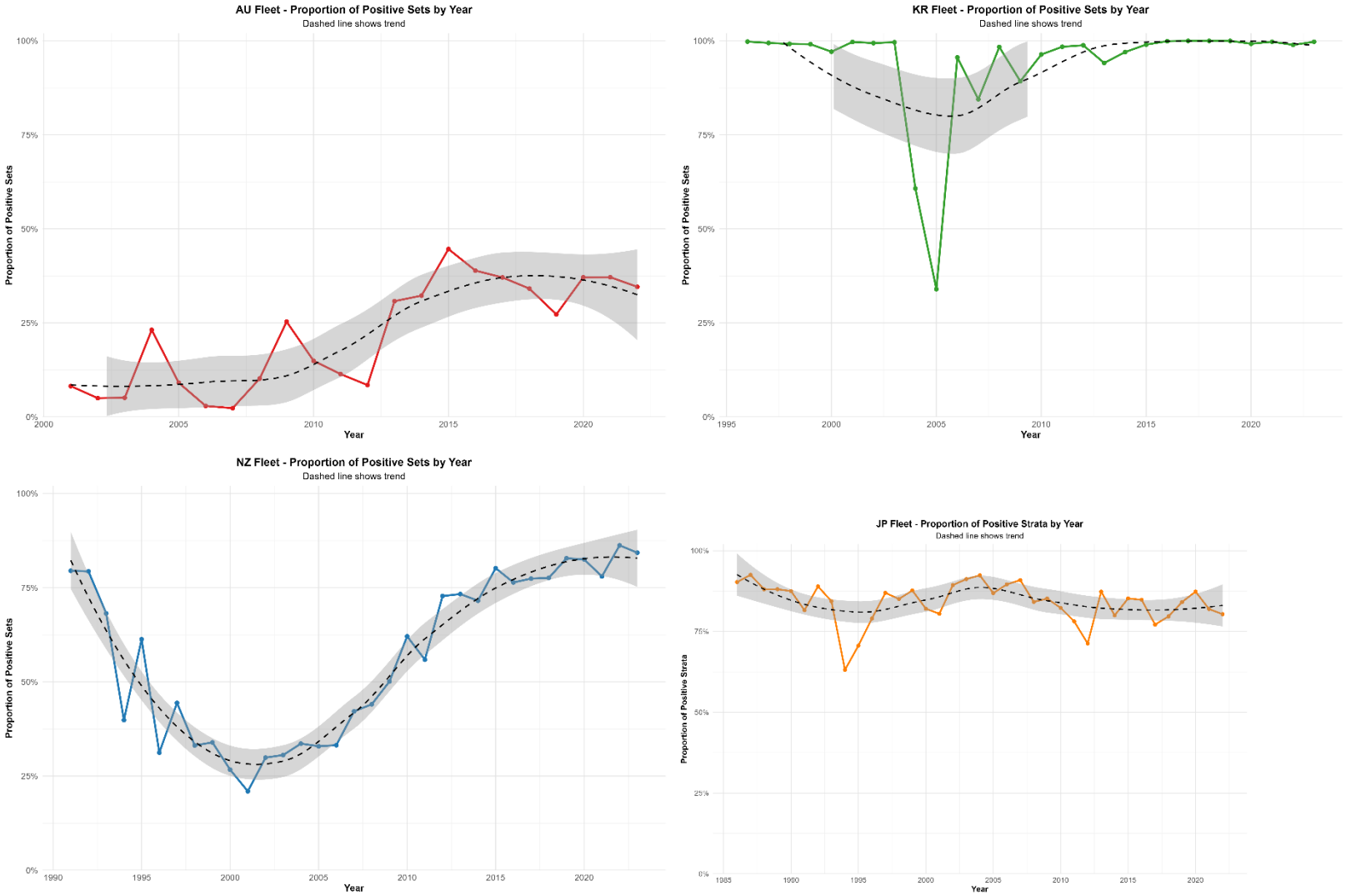

Figure 6: Time series plots showing proportion of sets with SBT catch for operational fleets (AU, KR, NZ) and proportion of strata with SBT catch for aggregated Japanese (JP) fleet.

Figure 7: Stacked area plots showing relative catch proportions by species for operational fleets (AU, KR, NZ).

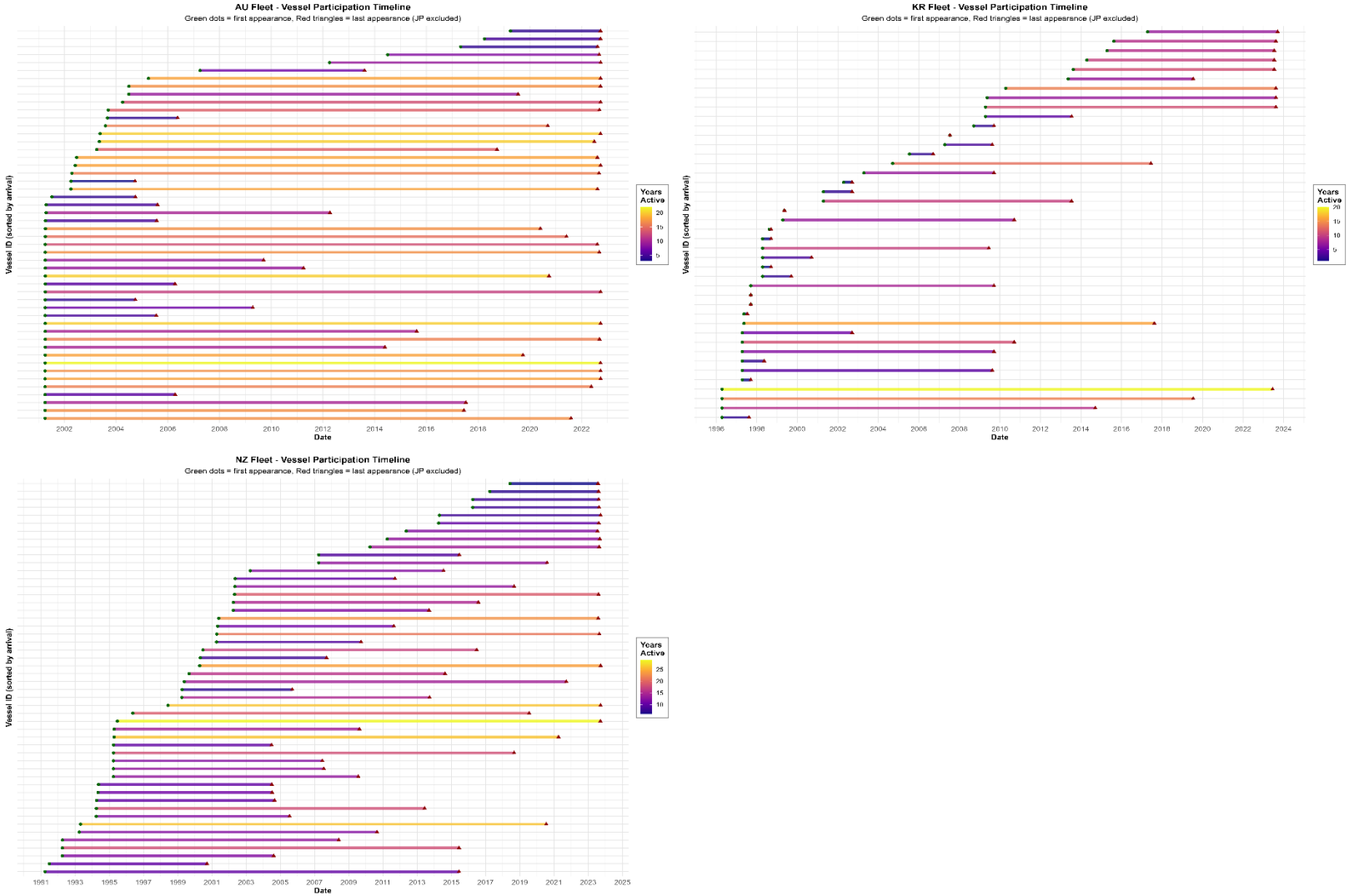

Figure 8: Line plots showing data coverage periods for individual vessels within operational fleets (AU, KR, NZ). Vessels providing fewer than 2 years of data were subsequently excluded, as were those with less than 5% or more than 97.5% positive observations.

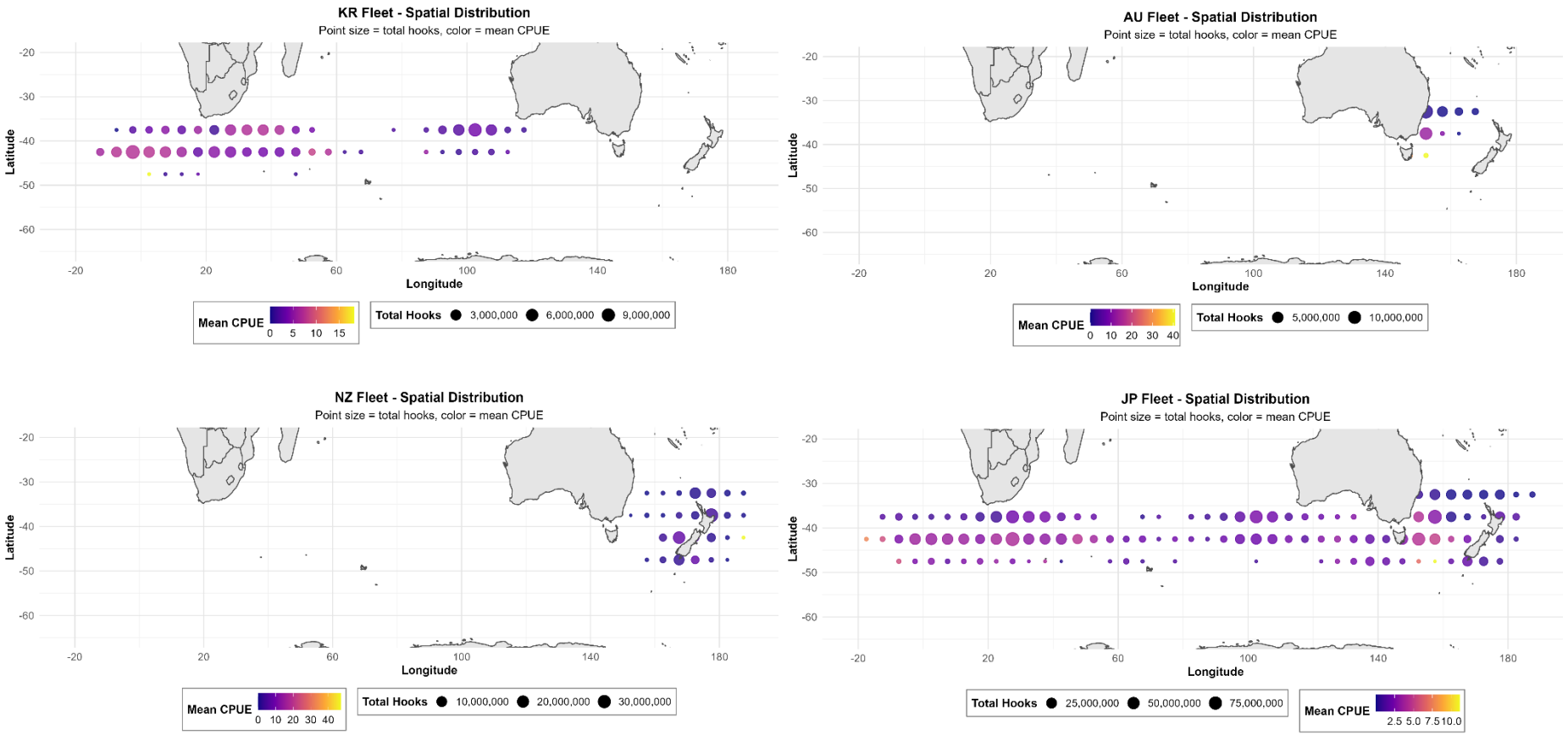

Figure 9: Maps showing geographic distribution of fishing effort for operational fleets (AU, KR, NZ) and aggregated Japanese (JP) fleet, with point size indicating effort and colour indicating CPUE.

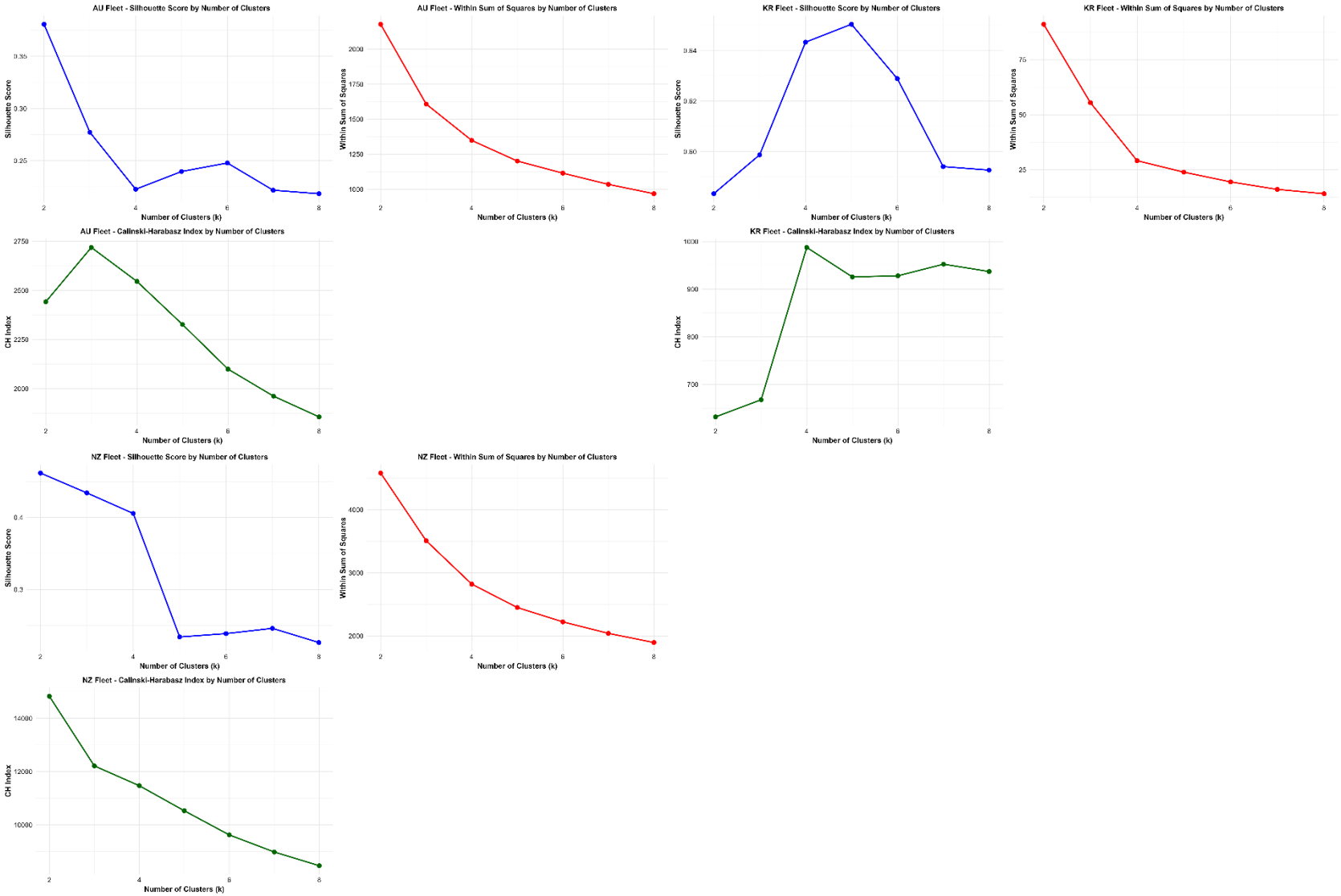

Figure 10: Cluster comparison plots by fleet for the operational data, using silhouette scores, Calinski-Harabasz indices, and elbow plots to select the preferred number of clusters when using hierarchical Ward2 clustering. For the first two approaches a higher score is better, while elbow plots show patterns in the trend of within-sample sums of squares.

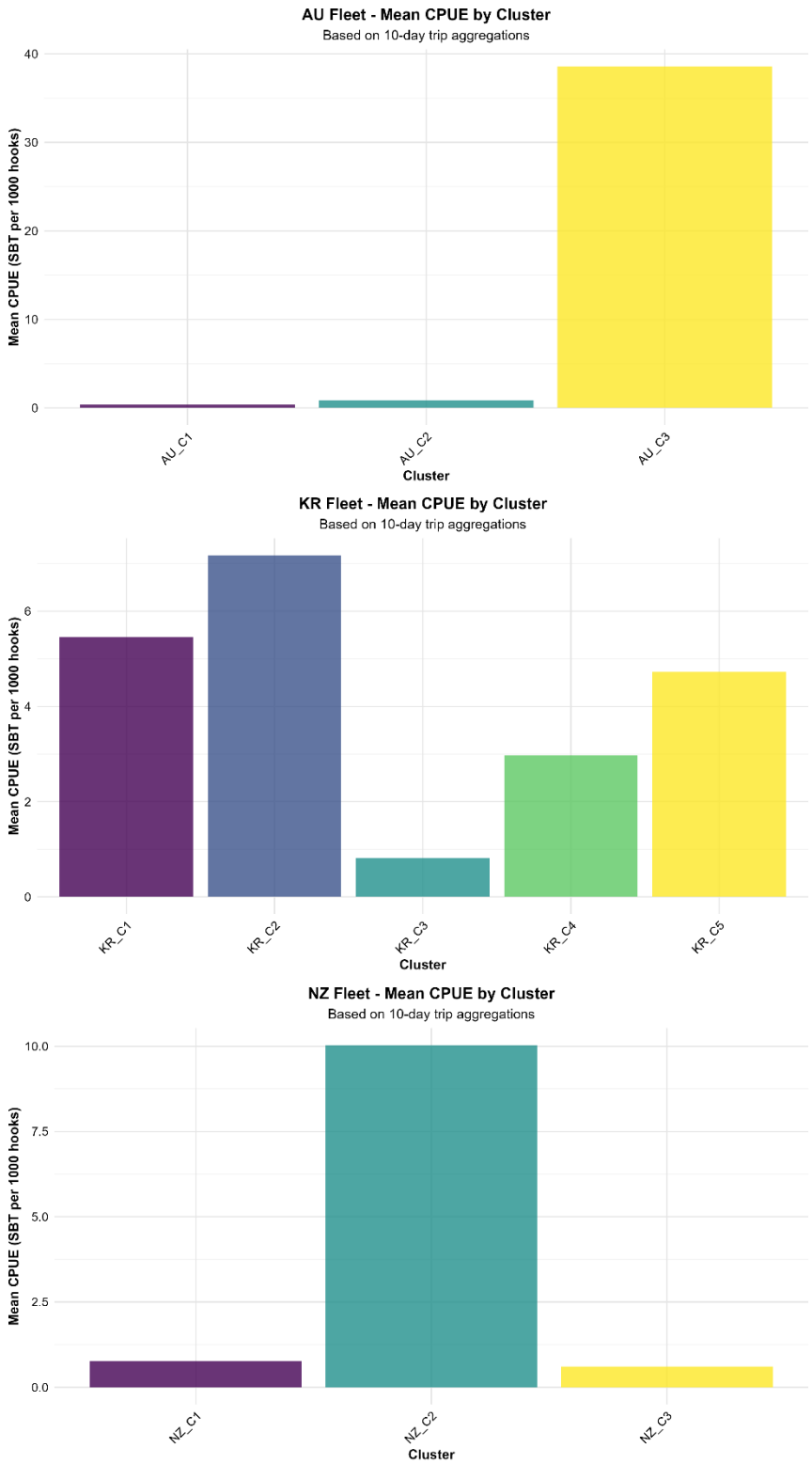

Figure 11: Bar plots showing SBT CPUE distributions by cluster for operational fleets (AU, KR, NZ).

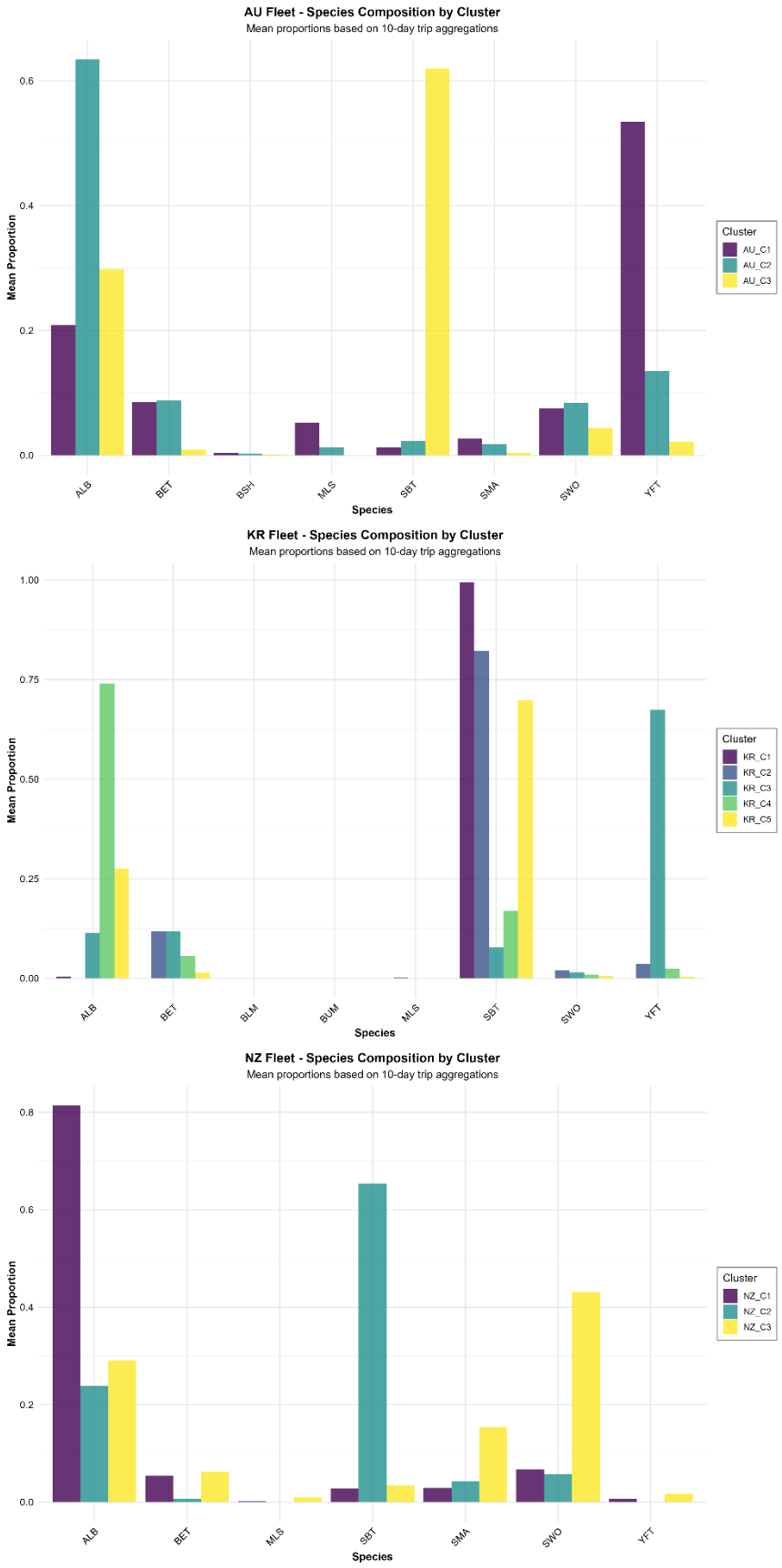

Figure 12: Bar plots (x-axis: species) showing mean catch proportions by cluster for operational fleets (AU, KR, NZ).

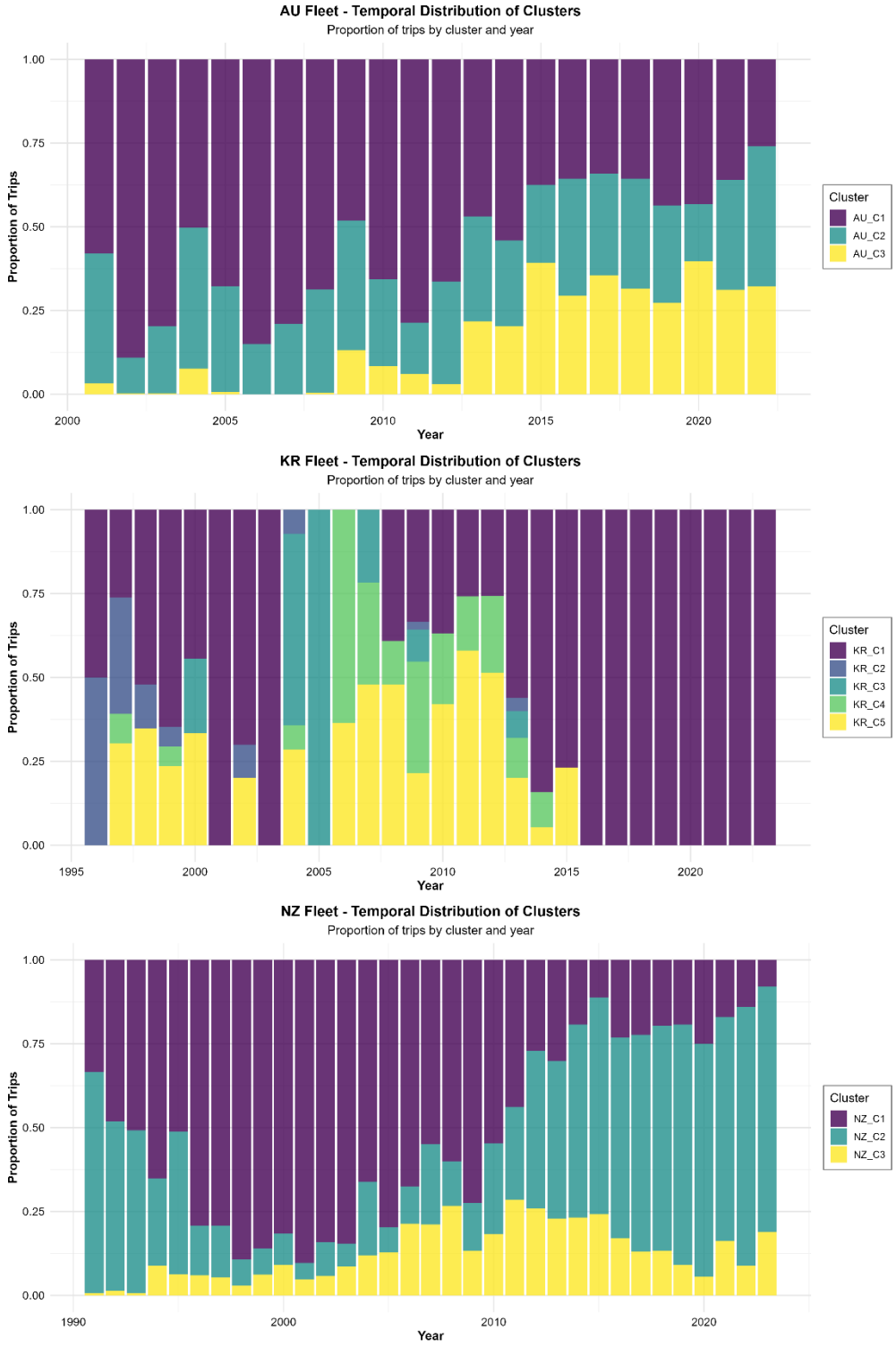

Figure 13: Stacked bar plots showing proportion of effort by cluster for operational fleets (AU, KR, NZ).

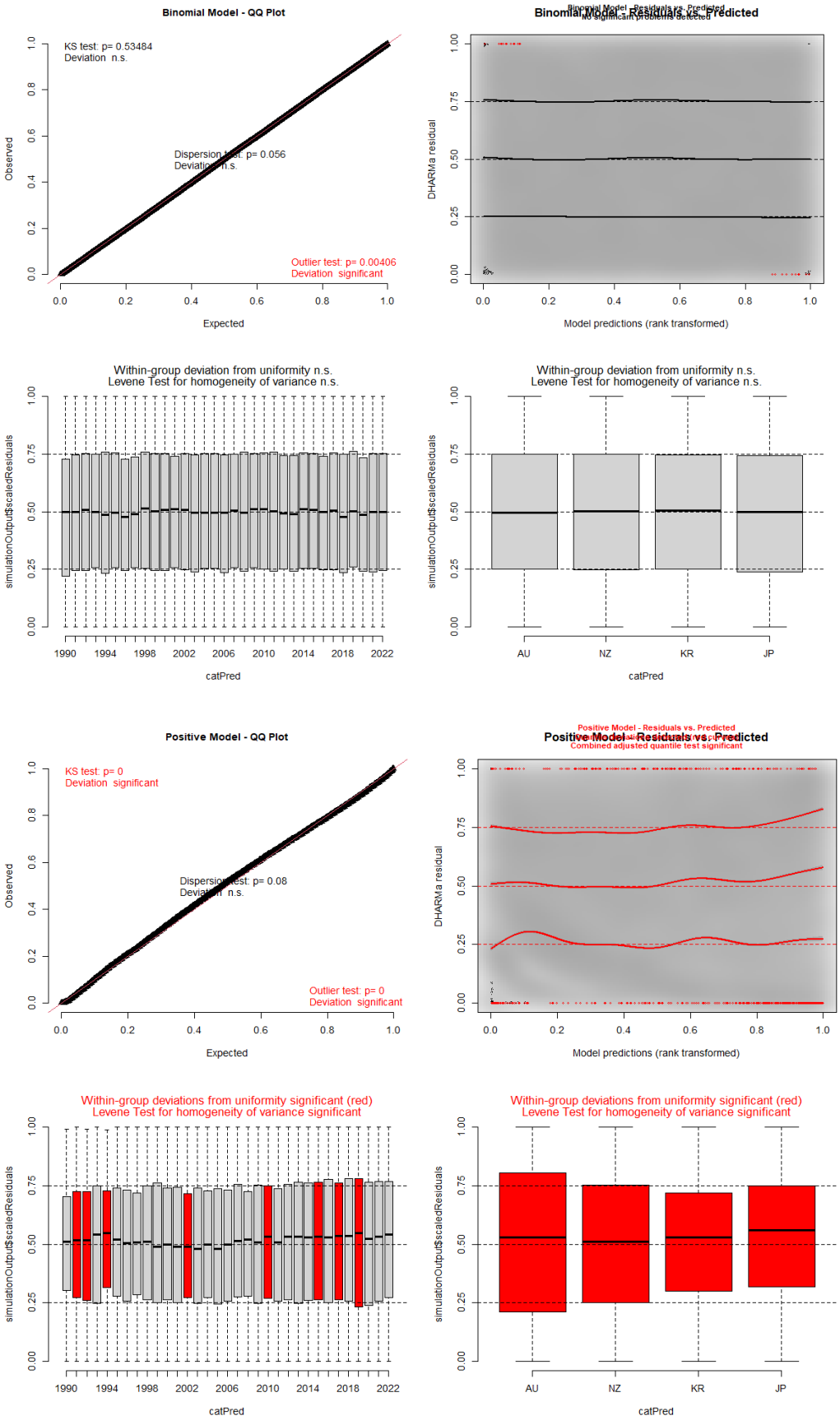

Figure 14: Four-panel plots for the binomial (above) and positive (below) sub-models, showing QQ uniformity, residuals vs. predicted, residuals vs. year, and residuals vs. fleet with DHARMa test statistics.

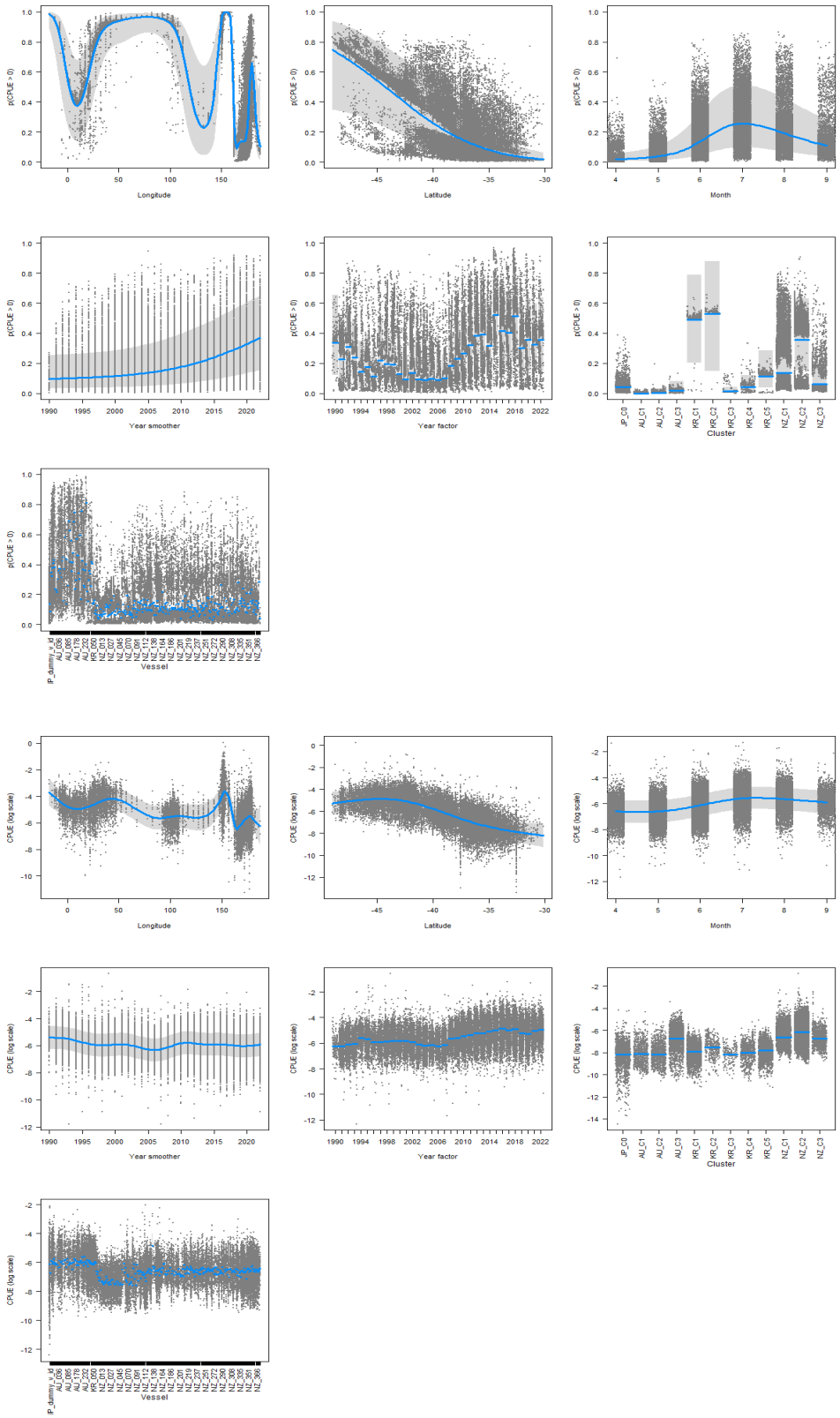

Figure 15: Partial effects of individual covariates on binomial (above) and positive (below) responses with confidence bands, plotted using the visreg package.

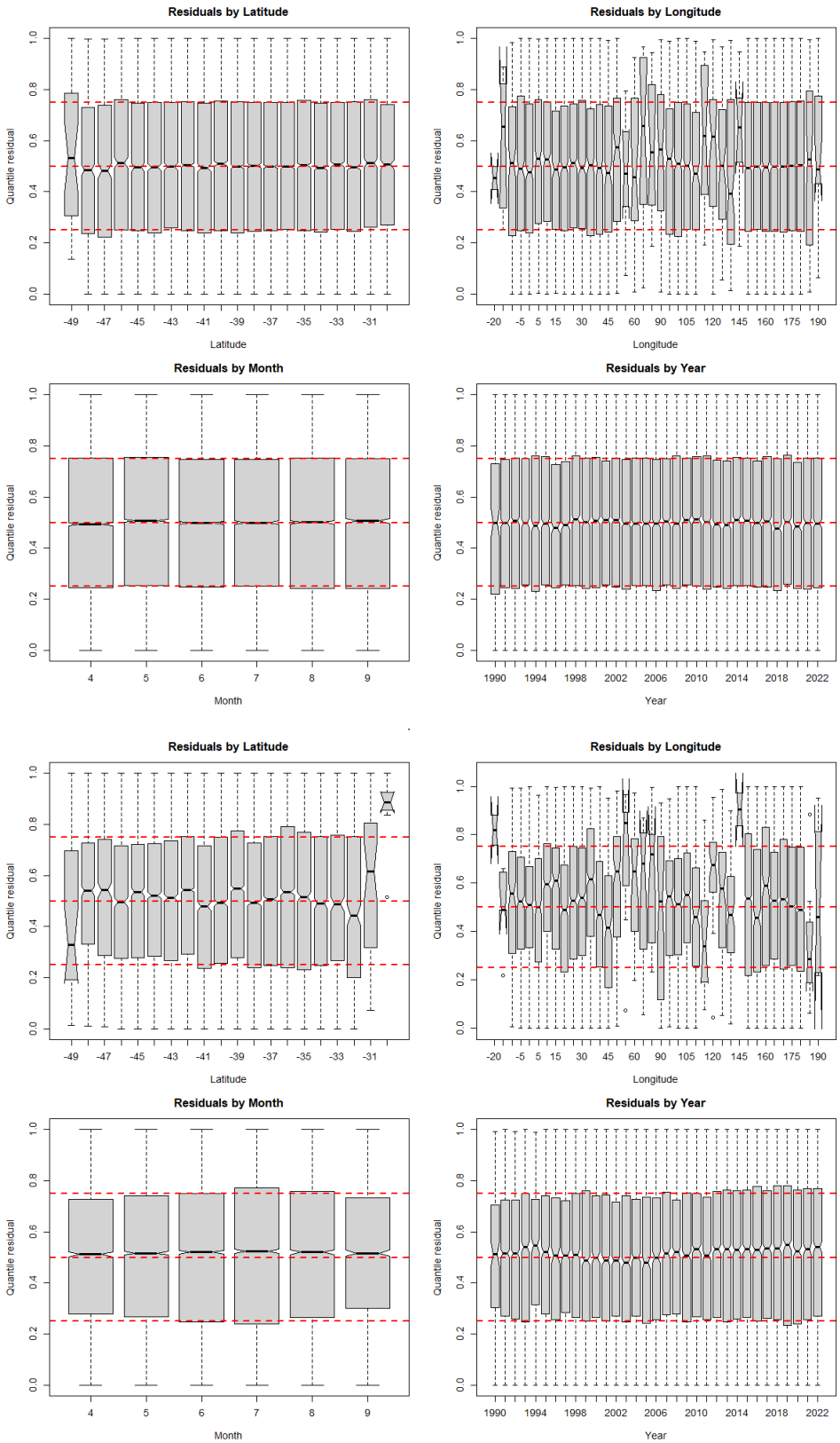

Figure 16: Boxplots for the binomial (above) and positive (below) sub-models, showing quantile residuals (y-axis) against latitude, longitude, month, and year (x-axes) with reference lines at 0.25, 0.5, and 0.75 quantiles.

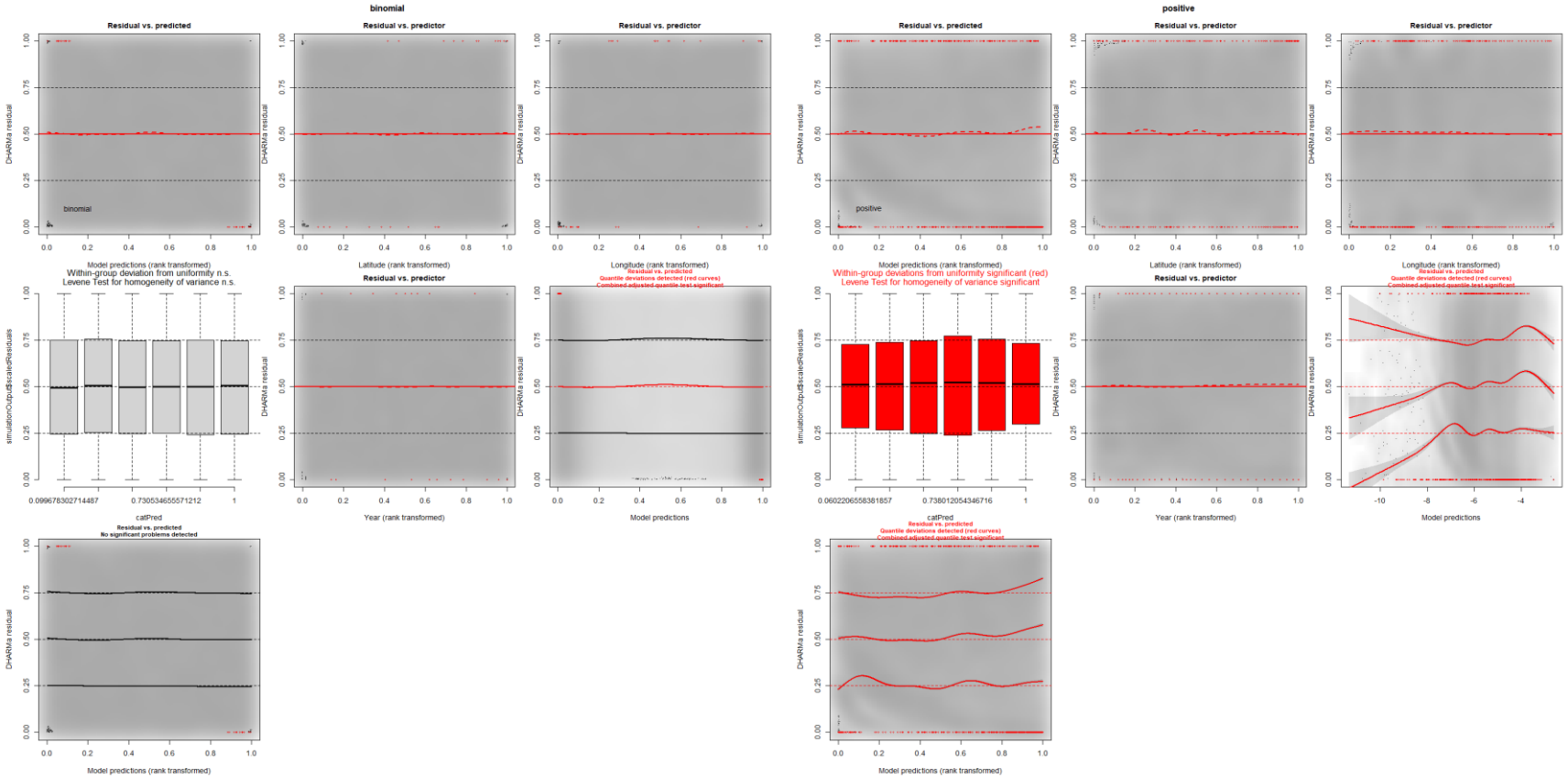

Figure 17: DHARMa diagnostic plots for binomial (left) and positive (right) sub-models, showing quantile residuals vs. predicted values, residuals vs. latitude, longitude, month, and year, with quantile regression lines and uniformity tests.

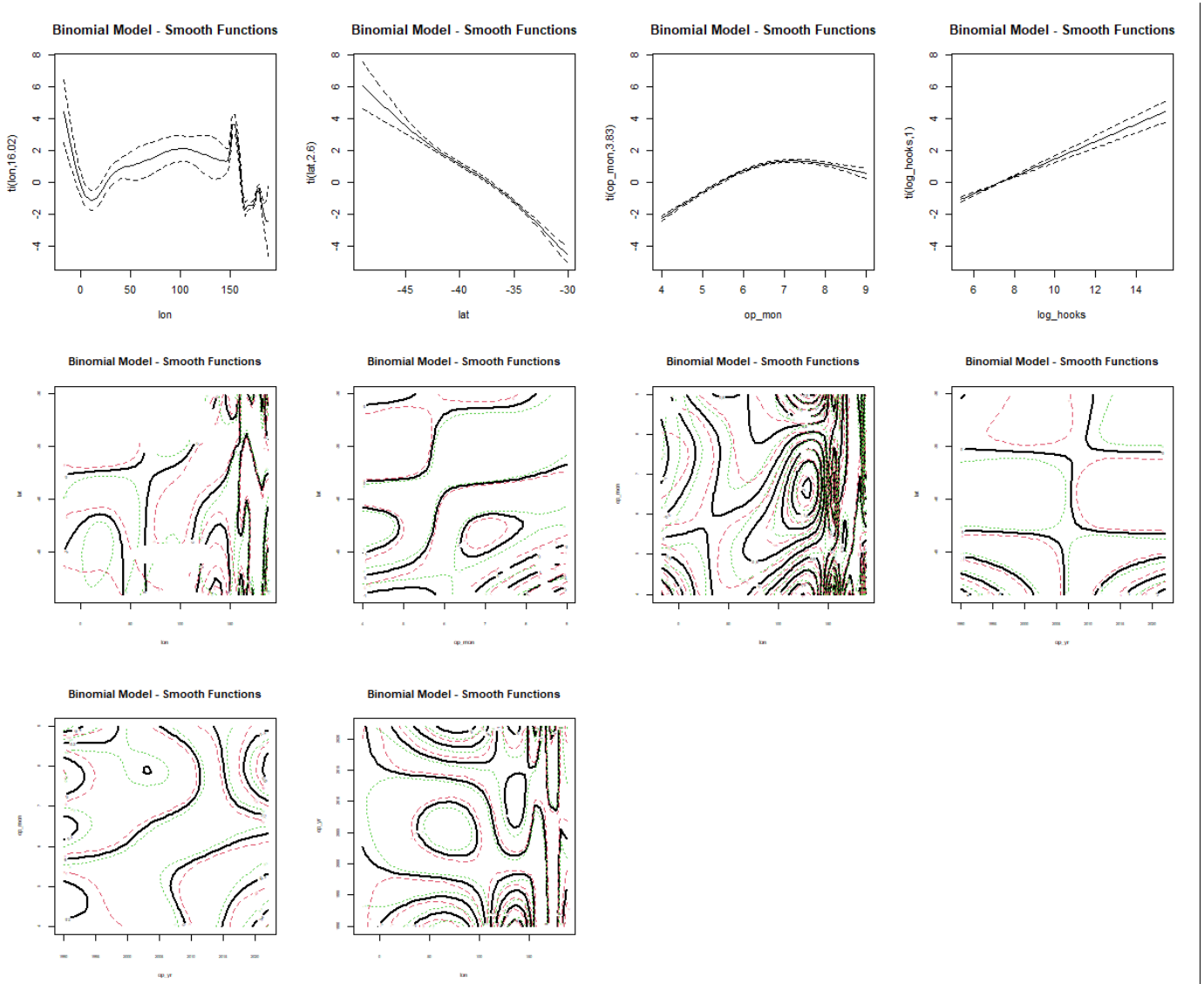

Figure 18: Fitted smooth functions for spatial, temporal, and other covariates in the binomial sub-model with confidence bands.

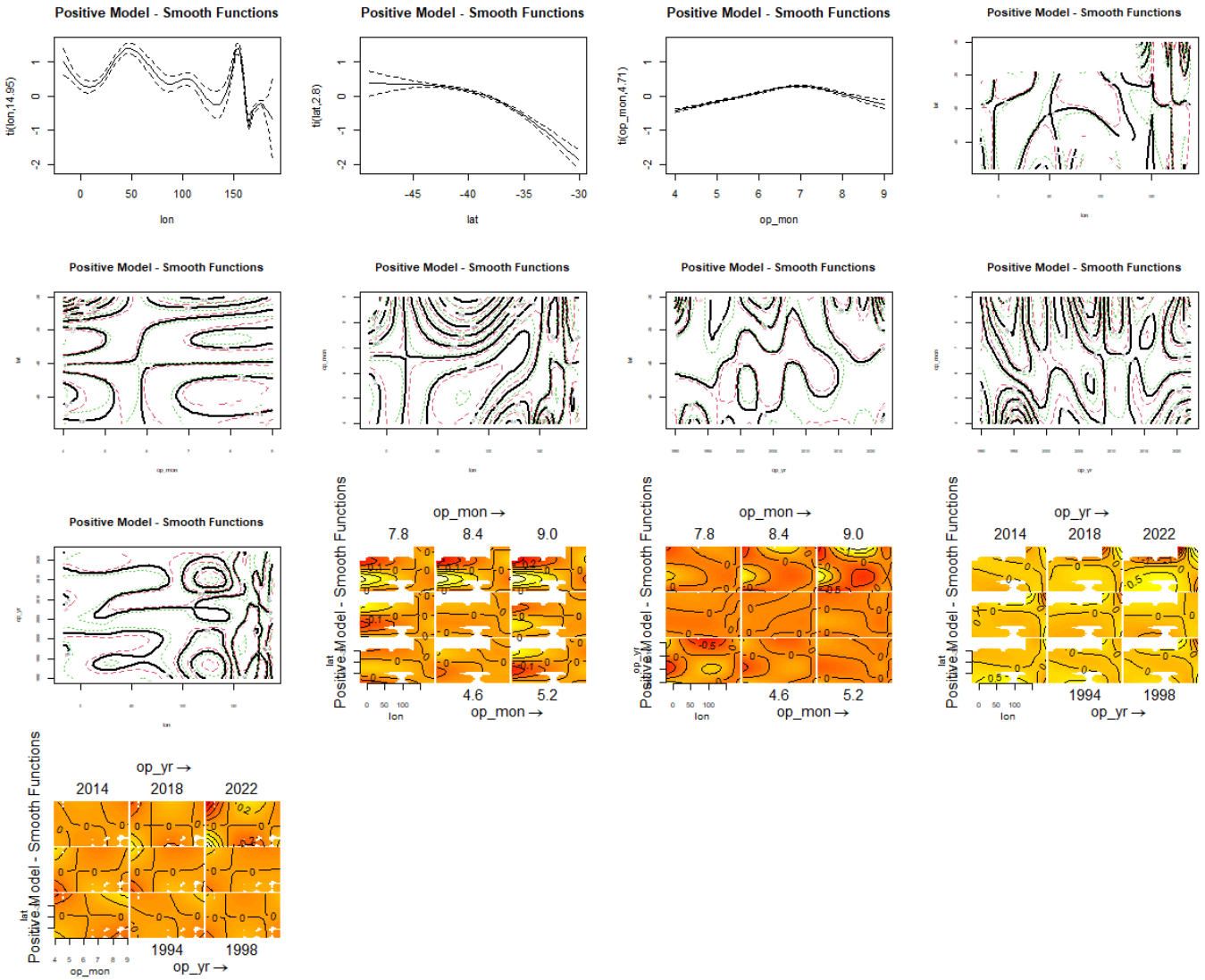

Figure 19: Fitted smooth functions for spatial, temporal, and other covariates in the positive sub-model with confidence bands.

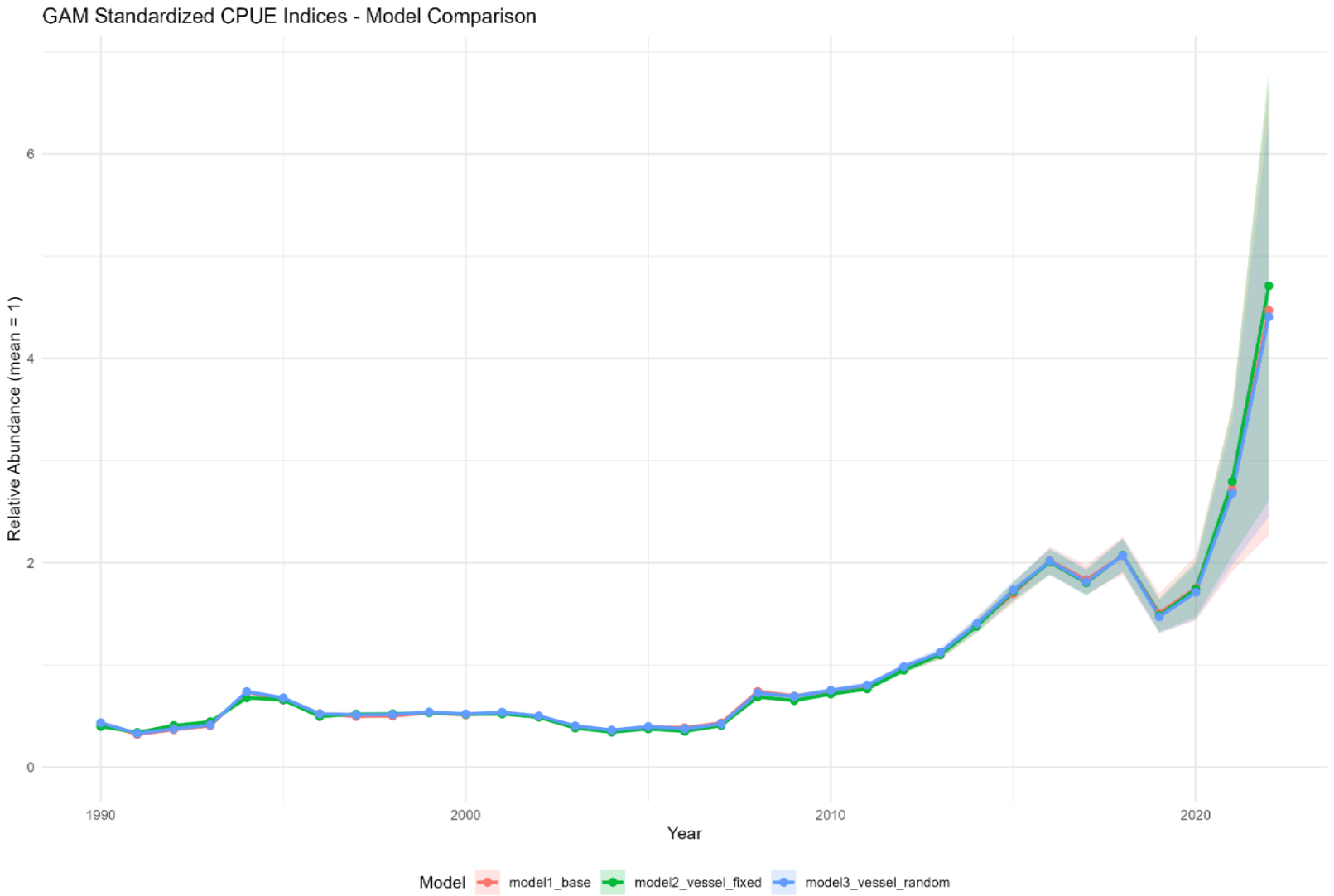

Figure 20: Comparison of indices derived from 3 models: the base model 1 without vessel effects; model 2 with vessel effects implemented as fixed effects, and model 3 with vessel effects implemented as random effects.