**Title: Panmictic Panacea? Implementing Good Practices for Developing Spatial Stock Assessments through Application with Alaskan Sablefish (*Anoplopoma fimbria*)**

**Running header:** Spatial Stock Assessment for Alaskan Sablefish

**Authors:** Craig A. Marsh1, Daniel R. Goethel1\*, Aaron M. Berger2, Katy Echave1, Pete Hulson1, and Ben Williams1

**Affiliations:** 1Auke Bay Lab, Alaska Fisheries Science Center, National Oceanic and Atmospheric Administration, Juneau, AK, USA.

2Fishery Resource Analysis and Monitoring Division, Northwest Fisheries Science Center, NMFS-NOAA, 2032 S.E. OSU Drive, Newport, OR, USA.

\*Corresponding author: [daniel.goethel@noaa.gov](mailto:daniel.goethel@noaa.gov)

**Journal:** Canadian Journal of Fisheries and Aquatic Sciences or Fish and Fisheries

**ORCID:** Craig Marsh,

Daniel Goethel, <https://orcid.org/0000-0003-0066-431X>

Aaron Berger, <https://orcid.org/0000-0002-1408-7122>

Katy Echave,

Pete Hulson,

Ben Williams,

# **Abstract (word limit from journal?)**

Marine species and associated fisheries demonstrate complex spatial dynamics driven by a myriad of biological, ecosystem, and socioeconomic factors. Recent improvements in data resolution and data quality have led to increased awareness that spatiotemporal processes should be integrated into the stock assessment models that inform fishery management advice. We utilize a case study with Alaskan sablefish (*Anoplopoma fimbria*) to demonstrate how good practices for developing spatial stock assessments can be implemented in a real-world application. Our framework emphasizes implementation of flexible and reproducible workflows to enable group model building and testing of multiple hypotheses, utilization of high resolution data analysis to inform key model structure decisions, co-development of single region and spatially explicit assessments to improve insight, and performing thorough simulation self-testing to ensure model tractability. Results indicate that the current single region model for Alaskan sablefish is likely adequate for management decision-making, given the truly panmictic nature of the population. However, spatial models identified regional differences in availability of sablefish by age class and differential harvesting by management unit. Moreover, by utilizing a fully integrated modeling approach, new insights were garnered regarding movement potential of this highly mobile species, indicating that movement across regions is likely lower than indicated when analyzing tagging data in isolation. We recommend that the single region model be used for quota setting, but semi-regular updates of the spatial model could provide further insight into the distribution of incoming year classes and the potential for regional depletion of older sablefish. The sablefish application provides one of the first end-to-end demonstrations of implementing spatial stock assessment good practices. Thus, it will help future practitioners better recognize important decision points and associated analyses to inform them when attempting to build a spatial stock assessment.

# **Introduction**

Spatial delineation of the world’s oceans has progressed rapidly in recent decades as marine spatial planning initiatives have attempted to address the competing needs of the myriad sectors increasingly being involved in the blue economy (Bax et al., 2022). Concurrently, improvements in biological understanding of the spatial structure of marine populations and the role of spatial dynamics in governing the productivity and sustainable harvest of living marine resources has led to increasing implementation of spatial fisheries management actions (Cianneli et al., 2013; Ciu and Kritzer, 2016). Thus, provision of scientifically-informed management and ocean policy requires development and utilization of spatially explicit modeling frameworks that can better match the scale of decision-making (Goethel et al., 2016). Across the array of disciplines that constitute fisheries science, spatial models have proliferated to address emerging needs and understanding, including early life history dynamics (e.g., larval individual-based models; ), spatial utilization across the entire life cycle of marine species (e.g., ), fishery processes (), whole of ecosystem approaches, and to inform marine spatial planning (). Although the full array of modeling frameworks help guide spatial fisheries management decisions, advances in spatial stock assessment models could provide the most direct benefits to management advice, given that stock assessments are utilized for estimating population status and projecting sustainable catch quotas based on observed data (Hilborn, 2012; Goethel et al., 2023).

Spatial stock assessments are able to integrate the complex spatiotemporal dynamics of marine species, which can help better elucidate drivers of productivity and refine estimates of sustainable regional harvests, thereby, providing advice at the scales relevant to decision-makers (Goethel et al., 2011; Berger et al., 2017). Moreover, by integrating data at finer resolutions, spatial assessments can increase the information content of most data inputs (e.g., fishery catch, indices of abundance, and age or length compositions; Punt et al., 2019b). Traditionally, stock assessment models have ignored spatial processes, instead assuming a unit population exists with homogeneous dynamics and negligible immigration or emigration (Goethel and Cadrin, 2021). Non-spatial assessments can provide adequate management advice when the boundaries of the modeled unit correspond with primary biological structure, even when movement occurs outside of the population (Cadrin et al., 2019; Cadrin, 2020). However, historic management unit delineations often do not match those of biological populations (Kerr et al., 2017). Moreover, fine-scale biological processes are increasingly being recognized, including both inter- and intra-population structure due to spatially varying demographics and connectivity (Smedbol and Stephenson, 2001; Cianelli et al., 2013). Therefore, finer resolution assessment models are required to avoid violation of assumptions and associated bias in estimates of population trajectories, parameters, and harvest advice (Punt, 2019a; Goethel et al., 2021; Bosley et al., 2022).

Simulation analysis has been widely implemented to test and compare the performance of spatially explicit and single region (i.e., non-spatial) assessment models when confronted with spatial dynamics (Punt, 2019a,b; Goethel and Cadrin, 2021). Oftentimes, spatial models are able to adequately account for complex spatial structure when present, while also providing management quantities at the resolution needed to maintain inter- and intra-population integrity (e.g., Carruthers et al., 2015; McGilliard et al., 2015; Goethel et al., 2021; Bosley et al., 2022). However, when poorly specified, spatial models may perform worse than spatially implicit (i.e., areas-as-fleets, AAF) or single region models for estimation of total population biomass (Lee et al., 2017; Guan et al., 2019). Quota apportionment methods are often then required to assign population-level quotas based on non-spatial assessments to management regions to match the scale of management advice, and adequate catch apportionment may be difficult to achieve except when high resolution fishery-independent survey data is available (e.g., Bosley et al., 2019). In an attempt to provide further insight on the relative performance of spatial and single region models, Goethel et al. (2024) implemented a first of its kind international, blinded, spatially explicit simulation experiment using a high spatiotemporal resolution operating model. The experiment emulated the uncertainty associated with a real-world assessment application (i.e., by providing limited a priori information on population dynamics aside from what could be gleaned from the simulated data), and demonstrated that no single spatial structure (i.e., spatially explicit compared to single region assessment models) performed best. Thus, further evidence was provided to support the prevailing sentiment in the stock assessment community that the utility of a given assessment spatial structure depends on the application, and that multiple structures should be developed in tandem to improve ecological and management insight (Goethel et al., 2024).

For operational management advice, applied spatial assessments remain scarce (Berger et al., 2017; Punt, 2019b). A number of broadly distributed Pacific tunas (e.g., Indian Ocean yellowfin tuna, *Thunnus albacares*; Fu et al., 2019) are assessed using spatial models. Conversely, only a handful of demersal species are managed based on the outputs of spatial models, and these assessments often relying on simplified spatial processes (e.g., multiple regions with no post-larval dispersal) when implemented (e.g., Gulf of Mexico red snapper, *Lutjanus campechanus*, and Northern shrimp, *Pandalus borealis*, in the North Sea; SEDAR, 2018; Cardinale et al., 2023). Although applied spatial assessments are rapidly increasing in research contexts (e.g., Carruthers et al., 2015; Goethel et al., 2015; de Moor et al., 2017), adoption for the purpose of management decision-making is often impeded by institutional inertia to changes in assessment and management approaches, perceived increases in model and output complexity, and potential mismatches in assessment and management boundaries (Kerr et al., 2017; Berger et al., 2017; Cadrin, 2020). Additionally, dissemination of spatial assessment applications within a management context is typically limited to the grey literature, which results in advances or lessons learned from these applications being insular within regional fisheries management organizations (Goethel et al., 2023, 2024).

However, there is clear momentum towards wider application of spatial assessment frameworks (Goethel and Cadrin, 2021), which is evidenced by myriad options for integrating spatial dynamics within the primary generalized assessment platforms utilized globally (Berger et al., 2024). Furthermore, attempts to document initial good practices for spatial models and associated identification of spatial population structure have been undertaken, but these are mostly pragmatic suggestions given the dearth of applications upon which to base recommendations (e.g., Punt, 2019b; Cadrin et al., 2023; Goethel et al., 2023). Moreover, many spatial assessments adopt the default spatial structure of the existing management framework, which can detrimentally impact catch advice when misaligned with the underlying biological structure (Berger et al., 2021). Therefore, there has been broad acknowledgement that increased documentation of applied spatial assessments is needed, including evidence for the implemented spatial structure, which would help improve knowledge sharing while better highlighting the benefits and pitfalls of spatial stock assessments (Goethel et al., 2023, 2024). Exemplar spatial model applications should include the full model development process, including identification of appropriate spatial structure (i.e., based on population or stock identification and high resolution data analysis), development of conceptual models representing the full spatiotemporal dynamics, stepwise and iterative spatial model building approaches, diagnostic and self-testing results, comparisons across spatial model archetypes, and determination of the most appropriate spatial model configurations (i.e., based on management needs, data limitations, and model tractability; Goethel et al., 2024).

Utilizing the good practices outlined by Punt (2019b) and Goethel et al. (2023) for implementing spatial stock assessments, we develop the first case study documenting the entire model development approach for a spatial stock assessment with application to sablefish (*Anoplopoma fimbria*) in Alaska. Our approach emphasizes the high resolution data analysis utilized to identify appropriate spatial structure and inform adequate model assumptions regarding spatial dynamics. Given the limited documentation of spatial assessment applications, fully describing the model development process (i.e., including exploratory data analysis, model comparison techniques, and goodness of fit methods) will help identify and disseminate lessons learned to aid future spatial applications. In particular, we highlight model decision points and key considerations for transitioning from a single region (i.e., panmictic) assessment towards a spatially explicit model, which should be of broad interest given that such situations are the most common for fisheries management globally (Kerr et al., 2017). The Alaska sablefish application provides a interesting exemplar species in this context, because it is a relatively mobile demersal species with limited genetic structure (i.e., it is truly a panmictic stock), but which requires regional management (i.e., catch apportionment to management units). As such, the methods and results of this work will provide a useful template for future spatial model practitioners that can be used as a guide for developing models with other species. Moreover, by comparing the results of spatial and non-spatial assessments, it provides direct input for Alaska sablefish management by providing insight into the adequacy of current regional catch apportionment strategies from single region models.

# **Methods**

The process undertaken to develop a spatially explicit assessment model for Alaskan sablefish generally follows the guidance provided by Goethel et al. (2023). By utilizing the step-by-step approach outlined by Goethel et al. (2023), the current work provides an applied archetype for how good practices for spatial model development can be used in practice. Our approach utilized four main steps, each with multiple sub-modules, including (Table 1): assessing the need for a spatial model and performing a data inventory, identifying appropriate and tractable model structures, implementing and refining spatial model alternatives, and choosing a final spatial model and comparing results with non-spatial alternatives. Although the results of the final spatial model provide important insight for sablefish management, the purpose of the study was not necessarily to replace the existing single region sablefish assessment. Instead, the goal was to demonstrate how the recommendations of Goethel et al. (2023) could be implemented with a real-world example, while also illustrating how single and spatial assessments can be implemented synergistically and leverage strengths of each structure to improve scientifically-informed management advice. Given these study expectations, emphasis is placed on model development and associated decision point analyses. To a lesser extent, brief comparisons among different model spatial structures are utilized to demonstrate management implications.

Paragraph…summarizes the 4 main components below

## *Sablefish Background*

* Why Sablefish – mobile species, great data, a lot of tagging data, previous studies to lean on but none have integrated tagging data with the assessment data

We documented the development of a spatially explicit tag integrated age structured assessment model for Alaskan Sablefish (*Anoplopoma fimbria*), covering the Gulf of Alaska, Bering Sea, and Aleutian Islands regions (Goethel et al. 2023). This stock provides a unique case study to illustrate the spatial modeling process, due to its vast mobility, economic importance, and a long-term dedicated longline survey and over 40 years of tag releases and recaptures. Currently a single, panmictic population is modeled with quotas apportioned to management areas based on area-specific survey biomass (SAFE reference). However, significant spatial heterogeneity exists in the population distribution resulting from multiple hypothesized factors such as age-based habitat preferences and migration patterns.

Short paragraph on summary of tagging models, then how Kari used these in a spatial assessment (not tag-integrated)…highlight limitations of non-integrated approach

Many spatial tagging models have been developed and or included the Alaskan Sablefish stock (Hanselman et al 2015, Heifetz and Fujioka 1991, Bracken 1982, Kimura & Shavy 1998) but none have attempted to integrate the tagging data with the assessment data and the assessment population structural assumptions in a full integrated assessment model. Computational power limitations previously hindered fully integrated analyses in past studies, but recent advancements in computers and efficient statistical software like TMB and Stan (references) have alleviated these barriers.

## *Identifying Spatial Model Need and Data Availability*

Highlight that use questions of Table 2 to guide the approach here

### Literature review

Expanding panmictic stock assessments to include spatial regions will have existing assessment models and supporting material present in the grey literature. We found this to be the best initial resource for understanding what has been explored in addition to providing insight into the most important process dynamics for the stock.

The primary literature was the next step for providing *a priori* evidence for spatial delineations and spatial structure.

Examples: genetics, tagging, morphometric, fishery analysis, stock structure, etc.

However, we recognize these resources are not available for many stocks globally. In these cases, environmental or ecosystem meta studies could be used as proxies for spatial distribution factors and delineation (Planque et al. 2011, Landa et al. 2011).

### Manager and Stakeholder Engagement

Outline approach for discussing needs with managers (NPFMC/SSC), engaging stakeholders through informal discussions and NPFMC meetings (PT), presentations, etc.

### Data Inventory and Exploratory Analyses

Reporting all data sets and the finest spatial resolution available is critical when considering how granular the assessment model can be. The dataset with the coarsest spatial resolution can set an upper limit for the number of regions considered during model development (Cadrin 2020). The limiting data set for sablefish was early catch history prior to 1979 (Table 4), which was only available at the stock level (Fenske 2023).

## *Defining and Reducing Model Structure*

Highlight that using Table 3 here (decision points) to determine major spatial processes that need to address

### Conceptual Spatial Model Development

Summarize approach for developing a conceptual model (i.e., based on lit rev, data availability, known dynamcis….developed the ‘ideal’ model structure…used this as most complex starting point to then trim down to a tractable starting model and permutations models using following data analyses and model sensitivity runs). This approach allowed freedom in exploring all possible avenues, before model triage was conducted to identify a ‘starting’ model for each spatial resolution.

### Exploratory methods used for spatial boundary delineation

One approach utilized in this study to inform spatial boundaries was that described by Lennert-Cody et al. (2010, 2013). This analyzed length frequencies from both fishery and survey data sets at the one-degree latitude and longitude spatial resolution. This method uses a regression tree to identify clusters based on seasonal, latitudinal and longitudinal breaks as covariates.

Age-length pairs were also visually plotted by FMP regions for each sex to inspect growth rate variation among regions. A more objective approach at identifying growth delineation is described by (Kapur et al 2020).

Fishing effort distribution plots were used to provide insight into the possibility of changing fishing selectivities over space and time and help inform fleet assumptions in the base model. The fishing effort variables available for our case study were depth, duration, latitude and longitude (See appendix Figure A 1 – Figure A 4). These show a shift from hook and line (HAL) to pot fishing (POT), but no systematic shifts that would highlight the need to incorporate spatially varying selectivities.

In addition to the above explorations, we explored tag-recovery matrices to understand the mixing of fish between FMP regions (Figure 3). This analysis aggregates all recoveries by release region and size group and ignores spatial-temporal recovery rates as well as time-at-liberty. The aim was to get a general understanding of source and sink dynamics.

Additional exploratory analysis available but not used in this study for informing spatial boundaries include; hierarchical clustering analyses from spatiotemporal models (Gruss et al. 2023), otolith chemistry (Tanner et al. 2016), and morphometric analysis (Kapur et al 2020).

## *Initial Spatial Model Implementation*

### Model Structure

A generalized spatial model was developed in TMB for this project, a full description of the model is given in Appendix (**Currently in a standalone word document to reduce clutter**). The model allowed for a general number of spatial regions, input observations and movement/tagging assumptions, and is hereby referred to throughout this paper as “the model”. The model was a initially based on the current stock assessment (Goethel et al. (2021), which assumed a single annual timestep which applied the following population dynamics;

1. Recruitment and tag releases

2. Total mortality and ageing

3. Markovian movement

4. Tag-shedding

The key population dynamics had spatial flexibility, recruitment allowed for both global and regional estimable annual deviations, and selectivities and catchabilities also could be estimated by region.

An initial model structure was used for each spatial model configuration (1-area, 3-area & 5-area). From the initial model, we then explored a range of model assumptions and used model selection criteria to identify a suite of candidate models.

An overview of the core process dynamics are given in Table 5. Observations available to each spatial model are given in Appendix B, the initial model assumed composition data were multinomially distributed with bootstrapped standard errors converted to an equivalent input sample size and abundance and catch data were lognormally distributed with design based standard errors.

### Identifying Key Uncertainties and Implementing Sensitivity Runs

Again, based on table 3 (decision points)

A range of alternative model assumptions were explored from the starting model structure for each spatial resolution. These included,

* Model start year (equilibrium vs non-equilibrium starting age-structure)
* Age-varying movement
* Spatially-varying recruitment deviations
* Time-varying selectivities
* Selectivity parameterizations
* Spatially-varying catchabilities
* Time-varying reporting rate
* Inclusion of the tag-recovery data
* Tag-likelihood (Poisson vs Negative Binomial)
* Compositional likelihood (Multinomial vs Dirichlet-Multinomial)

### Initial Diagnostics and Simulation Self-Tests

Model convergence was used to omit assumptions that resulted in unsatisfactory inference. A model was deemed not-converged if it exhibited any of the following conditions

1. Parameters estimated at or near boundaries
2. Non positive definite hessian
3. Maximum absolute gradient of any estimated parameter > 0.001

All models that converged were subjected to a “self-test” (Deroba et al, 2015). The self-test assumes the operating model (OM) and estimation model (EM) have identical model assumptions and are employed as an integrity check, to ensure model robustness and help identify coding errors. Any model that could not produce unbiased estimates of SSB was classified as non-converged.

For the models that passed the above convergence criteria had residuals assessed for all data sets. Given the large number of observations we frequently found that alternative model assumptions, would result in a similar total objective score, but different fits amongst the observed data sets, complicated model comparison. For this reason, we prioritized survey abundance index fits assuming the fits to composition and tagging data were satisfactory. This was firstly for convenience in model selection, but also follows the view of Francis 2011 & Francis 2014, where fits to indices of abundance should be prioritized over other data fits when providing catch-based management advice.

## *Final Models and Comparisons*

Summarize how chose final model structures based on diagnostics/sensitivity runs in previous section

### Model Performance and Diagnostics

General description of how plan to compare the final models (e.g., some quick diagnostics like residual plots, self-test performance…then comparing estimates of total and regional SSB, exploitation, etc.)

# **Results**

## *Why is a Spatial Model Warranted for Alaska Sablefish?*

Summarize Lit review, management discussions, and input from stakeholders

Can format this as a kind of Q+A based on Table 2

### Data Availability and Collation

Northwest pacific sablefish is fortunate to be the focus of genetics studies (Jasonowicz et al. 2017, **One more?**), morphometric studies (M. Kapur et al. 2020) and tagging studies (Hanselman et al 2015, Tripp-Valdez et al. 2012, Kimura and Shavy 1998, Heifetz and Fujioka 1991, Bracken 1982).

Sablefish is also fortunate to have a dedicated fishery independent survey that provides a long-time series of highly spatial length, age and abundance data (Kimura and Zenger Jr 1997). Since 1978, the Alaskan sablefish survey has been conducted during the summer using a systematic design, offering increased flexibility for generating model inputs at different spatial resolutions (e.g., three-region vs. five-region models). This flexibility arises from the uniform sampling intensity along the coast, distinguishing it from commonly used stratified survey designs. Spatially stratified surveys typically optimize variance reduction across space and can serve as initial spatial boundaries for a spatial assessment model. However, most surveys have multiple objectives and consider ecosystem factors rather than being solely optimized for a single species.

We found flexible estimators, such as model-based estimators (VAST Thorson 2015 and sdmTMB Anderson et al. 2022)) were advantageous over design-based estimators due to their ability to generate assessment inputs such as length frequencies and abundance indices for various spatial resolution assumptions. This flexibility enables modelers to easily configure and test alternative spatial structures, which is a large source of uncertainty in spatially explicit age-structured models. However, applying model-based estimators to Alaskan sablefish was challenging, due to their preference to occupy narrow area on the slope but have long north-south spatial extent along the slope from Canada to the Bering Sea. This spatial extent complicated the use of spatial model-based estimators, as they required fine spatial resolution to capture abundance changes across the slope resulting in sluggish and impractical use.

## *Spatial Model Building Process*

Summarize Table 3 results

### High Resolution Data Analysis

This approach identified three longitudinal breaks, two of which fell in the middle of an FMP boundary (Figure 2). The split in the Aleutian Island FMP was ignored due to the lack of catch and data available to the west of that split point. However, the longitudinal splits in the central Gulf region, did correspond with an ecosystem-based break identified by Kapur et al. 2020. Due to the spatial resolution of catch, an apportionment assumption would be required to allocate the Central Gulf catch to either side of this spatial boundary. A sensitivity simulation was thought necessary to identify if splitting the Central Gulf region at this break point would impact management advice, as of yet this sensitivity has not been explored.

Incorporating spatially specific growth in an age-based model requires expanding the numbers age to track length (or length group), to ensure the correct length composition among regions. Adding this extra dimension to the numbers at age matrix would have a significant computational cost in the context of a spatially explicit tag-integrated model. Due to this consideration, we decided that very stark spatial growth patterns would be required to justify the inclusion of spatially varying growth which is why we chose to use visual plots instead of a more objective statistical approach (Figure 3). **Probably need to re-word this**

Given the recorded spatial resolution of historical catch was at the FMP region and the length frequency analysis only had one split in the middle of an FMP region, we decided to develop a five-area model with the Eastern Gulf region consisting of both East/West Yakutat (Figure 1), due to low data availability. In the Western regions We also developed a three-area model which aggregated Bering Sea, Aleutian Islands and Western Gulf region similar to that developed by Fenske (2023) as well as a single panmictic model.

Tagged fish are released by the summer survey, which records the length of each fish released. There are two large benefits of having the survey release tags, the first being the presence of an age-length transition matrix. This is used to convert unsexed lengths frequencies of tagged fish at release to sex and age disaggregated releases which are compatible with the sex and age structured estimation model. The second, is the spatial coverage of releases, which results in a balanced spatial design.

Tagged fish are in the partition are assumed to have the same ageing, mortality and growth assumptions as the untagged members of the partition. When tagged fish are released into the partition, they are indexed by the release event index denoted by which is region and year specific (). Tag induced mortality and initial tag-loss is applied as an initial mortality rate denoted by ,

Ongoing annual tag-shedding denoted by is also applied as a mortality process whereby tagged fish are deleted from the partition as

Values for and were taken from the tagging study by Beamish and McFarlane (1988).

The final model consideration for tag-releases was for how many years to track the tagged cohorts within the model partition, also known as time-at-liberty denoted by . Figure 5 shows the skewed distribution of time-at-liberty for all tag-recoveries, which highlights sablefish’s ability to remain at liberty for long periods of time. We chose to use for the three-area model and for the five-area model, which was a compromise between computation time and potential movement information. These values were larger than the 9 years at liberty considered by Beamish and McFarlane (1988). The spatial model used a tag-release conditioned likelihood (Vincent, McGarvey) this means the model tracks numbers at age for each region and tag-release cohort. For a five-area model () without tagging and max age of 30 the partition has 150 elements (5 30). When tagging is included and , the partition has 7650 elements (150 + 30 5 ).

Two assumptions regarding tag-recoveries were investigated which included the mixing periods and tag-recovery method.

* mixing period - Figure 6 displays the relative distribution of releases and recoveries by release year and years-at-liberty (normalized to have a max value of 1 for each year at liberty). This shows recovered fish that have been at liberty between 0 to 1 years have quite a different distribution to fish that were at liberty for 2+ years. We interpreted this as mixing signal, where by fish at liberty for one year were more likely to be captured than fish from 2+ years. For this reason, we chose to have a mixing period of two years before considering tag-recovery observations. This approach differed from those discussed in Kolody and Hoyle (2015), firstly it was qualitative in nature and secondly, we focused on the temporal distribution of tag recoveries by release events. The fishery dependent catch per unit effort data was not a the spatially granularity to test for gradients in recoveries from release events as proposed by Kolody and Hoyle (2015).
* Tag recovery was fishery dependent, choice of gear type – Most of the returns are from the fixed gear fishery (Longline & Trap/Pot). One trend in recent years was a dramatic drop off in tag-recoveries. Due to this trend, we did not consider recoveries after 2016 (Figure 7).

### Conceptual Spatial Model

Quick summary of ‘optimal’ complexity model and why can’t achieve this

### Key model Uncertainties and Sensitivity Runs

Use table 3 to highlight the key decision point uncertainties

An additional sensitivity analysis focusing on the observation sub model of each converged model was also conducted. This iterated over each data set and dropped it during estimation whilst keeping all other datasets (except catch which was needed to estimate annual fishing mortalities). Spatial models that estimated movement would not converge when tagging data was omitted. This highlighted the necessity of tagging data to account for diffusion among regions (**Are there papers which show you don’t need tagging data? That either disagree or agree with this finding**).

Perhaps the most sensitive model assumption was the start year of the model. Across all three spatial resolution models we would get similar absolute SSB estimates over the data period but varying estimates of B0 and thus percent B0 depending on if we started the model in 1960, 1977 or 1990. This highlighted a key model sensitivity that was further explored using simulation. The issue with starting the model far away from B0 was estimability of the non-equilibrium parameters. The self-test frequently showed theses parameters were biased. There is a parallel project which is investigating optimal parameterization for spatial models that start in non-equilibrium states.

### Diagnostic and Self-Test Results

Model assumptions that had convergence problems included age-varying movement, certain time-varying selectivity models, and specific selectivity parameterizations. In the case of age-based movement the self-test simulation showed that all parameters were identifiable for a conditioned OM. However, understanding why the model did not converge when estimated with actual data was not ascertained during this project, due to the number of possible reasons and time constraints.

There was an *a priori* preference in the observation sub-model for distributions that had estimable dispersion parameters such as the negative binomial vs the Poisson and Dirichlet-Multinomial vs Multinomial (Thorson 2018). These distributions would remove the need to do any iterative data-weighting commonly conducted in stock assessments.

## *Final Models*

3 models structures: single region, 3 region, 5 region, including 3 different starting years

### Spatial Model Structure

### Diagnostic and Self-Test Results

The following shows the results from the self-test Figures ?? show SSB trajectories from the 1A, 3A and 5A models (move these to appendix, not really that important). The self-test was found to be the criteria which rejected the most model formulations. In particular there were issues with early and later recruitment deviations which lacked data were fixed due to the self-test in the spatial models. Even though the more traditional convergence criteria wanted more sophisticated model the self-test would frequently expose that they could not produce unbiased management estimates.

Relationship between and , interestingly with the estimated movement from the 3-Area model, it estimated a large for the Aleutian Islan/Bering Sea/Western gulf complex (AI/BS/WGOA), but a small , see Table 3. This highlights a source sink dynamic from AI/BS/WGOA down to the Eastern Gulf of Alaska (EGOA).

### Model Comparisons

SSB, depletion, regional exploitation

# **Discussion**

* General summary - This project applied and documented methods and key decision points for developing a spatially explicit age-structured stock assessment model to the Gulf of Alaska Sablefish stock. This resulted in a suite of spatial candidate models that resulted in “similar” estimates of global SSB during the data period (Figure 5). When examining global depletion (Figure 6), there was a much more variability among candidate models due to the uncertainty in estimates of .
* How can you use these results in management - One approach often used when stock assessments produce a suite of models is to use a model ensemble approach to inform management decisions (Ducharme-Barth and Vincent 2022). Model ensembles utilize all models in a weighted framework, which can account for model-structural uncertainty, which was shown in the sablefish case with spatial assumptions resulting in variable estimates of depletion.
  + How to use this tool?
    - Run it in parallel with the assessment, use it as a “conditioned OM” to test assessment model
    - Use it as a sanity check on regional abundance and catch advise.
  + How to handle spatial BRPs…do we need spatial BRPs or can more empirical approaches be used or in conjunction with single region BRPs (e.g., population wide biomass target and empirical BRPs to get regional apportionment, etc.)
* Discuss how this work builds on Kari’s work, movement assumptions.
  + Are results consistent with her findings, is there anything that differs.
    - Spatial EM chapter 2, movement is fixed.
    - We estimate movement that is consistent with observed data (not just from tagging data which has inherent bias)
* Caveats
  + Data
    - Mainly tagging data limitations
      * Only point of release/recapture
      * Long time at liberty
      * Nuisance parameters (tag reporting, tag mixing)
  + Time-varying parameters (movement, apportionment, reporting rate)
    - Room for random effects in the future (barring time/memory constraints in current implementations of state-space models?)
  + Demographic leakage across boundaries (ADFG/DFO management areas)
  + Discussion topics - Data-weighting
  + Things I would have changed and suggest others do (other than not doing a postdoc of course) - Change parameterization of (or ) to estimate a total value and apportionment parameters, that way you can profile the total to get an understanding of which data sets were influencing the absolute scale of the model which was sensitive to starting condition assumptions. The model developed, estimated a separate for each region, which made using log-likelihood profiles difficult to diagnose the source of this uncertainty.
  + Exploratory analysis found possible boundary splits in the middle of Central Gulf that were not practical to implement in the model due to data reporting restrictions, which are often pinned to management regions (circularity).
    - How to handle imperfect boundaries (Berger et al., 2021)
    - How to rectify imperfect boundaries given management inertia (Cadrin, 2021; Cadrin et al., 2023)
    - Usefulness of the regression tree approach (Lennerty-Cody 2013)
    - Implications of changes in biological parameters across space and how to deal with them
* Future Work
  + A model assumption that should be explored in future research is the uncertainty from converting unsexed lengths of tag-releases to sex and age disaggregated tag releases which are inputs in the model and assumed known without error. One approach that we wanted to explore consisted bootstrapping the age-length transition matrix and re-running the model with different tag-releases at age and sex from each bootstrapped age-length transition matrix. This is quite computer intensive and should only be conducted on final models.
  + Integrating multiple tag types (Thorson et al., 2021) and high resolution tagging submodel to better inform interannual variability in movement, deal with tagging nuisance parameters, and to integrate path dependency of tagging data (i.e., move away from point estimates only of tag release and recapture)
  + MSE next steps - The next step of this research would be to use a full management strategy evaluation (MSE) to identify the “best” performance spatial resolution model.
    - We used simulations to identify model pathologies and biases using the “self-test”. However, to truly compare management performance a full MSE is required which is the next step.
    - Need to determine if spatial models even warranted…what is the minimally complex, maximally robust management strategy?
      * Do simpler EMs perform ok
      * Eg Punt et al. (2017)
      * Potential for hybrid management strategies (e.g., using CKMR) and how would spatial dynamics potentially complicate…
    - Lessons learned from PSTAT sablefish MSE work (Kapur et al., IN review)
    - Potentially touch on next steps for sablefish MSE work (not necessarily spatial, but spatial dynamics in OM might be incorporated as part of testing alternate HCRs)
  + MSE may also identify alternative stock status reference points other than B0 which we found to be sensitivity to model assumptions such as the start year and conditions i.e., use where is the first year with reliable data (Punt 2003).
* Conclusion

# **Acknowledgements**

# **Data Availability Statement**

# **References**

# **Tables**

## **Table 1.** Recommended good practices from Goethel et al. (2023) for developing spatial stock assessments and how this advice was utilized during the current study, the benefits of the advice, and difficulties encountered in implementing it.

|  |  |  |  |
| --- | --- | --- | --- |
| **Good Practice Advice** | **Implementation Approach** | **Benefits** | **Difficulties Encountered** |
| Identify need for spatial model and perform literature review |  |  |  |
| Engage stakeholders |  |  |  |
| Develop Flexible Data and Assessment Workflows |  |  |  |
| Develop Conceptual Model of Spatial Dynamics |  |  |  |
| Narrow Model Options through High Resolution Data Analysis |  |  |  |
| Implement Single Region and Spatial Models |  |  |  |
| Use Diagnostics to Identify Final Suite of Models |  |  |  |
| Implement Simulations and Self-Tests |  |  |  |
| Identify Spatial Reference Points |  |  |  |
| Document |  |  |  |
| Iterate |  |  |  |

## **Table 2.** Primary questions used to identify the need for a spatial assessment (from Goethel et al., 2023) and associated answers for the sablefish case study used to narrow modeling options based on the literature review and initial data explorations.

|  |  |  |
| --- | --- | --- |
| **Question** | **Answer** | **Citation** |
| Management Needs? | Currently modeled as single panmictic model but annual catch allocation done spatially outside of assessment model. |  |
| Data Availability and Resolution? |  |  |
| Population Structure? |  |  |
| Degree and Drivers of Movement? | Ontogenetic, seasonal |  |
| Does Biology Vary Spatially? | Recruitment is assumed to vary, given the movement dynamics and observed spatial abundance |  |
| Impacts of Climate Change? |  |  |
| What is minimal complexity needed to inform management decisions? |  |  |
|  |  |  |

## **Table 3.** Major decision points encountered when developing spatial models as outlined by Goethel et al. (2023) with the type of analyses used to inform each decision, the final model parametrization chosen, and difficulties or uncertainties encountered during the model building process for the Alaskan sablefish case study.

|  |  |  |  |
| --- | --- | --- | --- |
| **Spatial Decision Point** | **Analyses Used** | **Final Parametrization** | **Difficulties Encountered** |
| Population Structure | Spatial clustering algorithm using survey lengths |  |  |
| Temporal Structure |  |  |  |
| Spatial Resolution | Tested multiple |  |  |
| Fleet Structure |  |  |  |
| Recruitment Dynamics | Global recruitment vs spatial recruitment devs, BH vs no SR |  |  |
| Regional Scaling |  |  |  |
| Dispersal |  |  |  |
| Movement | Markovian after the recruitment age |  |  |
| Demographic Variation | Constant growth among regions |  |  |
|  |  |  |  |
|  |  |  |  |

## **Table 4.** Input data sets and recorded spatial resolution.

|  |  |
| --- | --- |
| Data Set | Spatial Resolution |
| Catch pre-1979 | Available at stock resolution by gear |
| Catch post-1978 | Available at Fishing Management Plan (FMP) boundaries by gear (Figure 1) |
| Observer data (age, length and catch) | Latitude and Longitude positions |
| Survey data (age, length and catch) | Latitude and Longitude positions |
| Tagging data | Latitude and Longitude for releases, Latitude and longitude for approximately (get %) recoveries |

## **Table 5.** Initial model process dynamic assumptions.

|  |  |
| --- | --- |
| Process dynamic | Assumptions |
| Recruitment | No stock recruitment relationship,  Regional specific and annual recruitment deviations |
| Fishing Mortality | Fishery gear specific selectivity constant across regions. Annual regional fishing mortalities solved using a Newton Raphson algorithm |
| Movement | Markovian movement, which is age & time-invariant |
| Natural mortality | Not estimated, age and time invariant |
| Tag release | See next section |

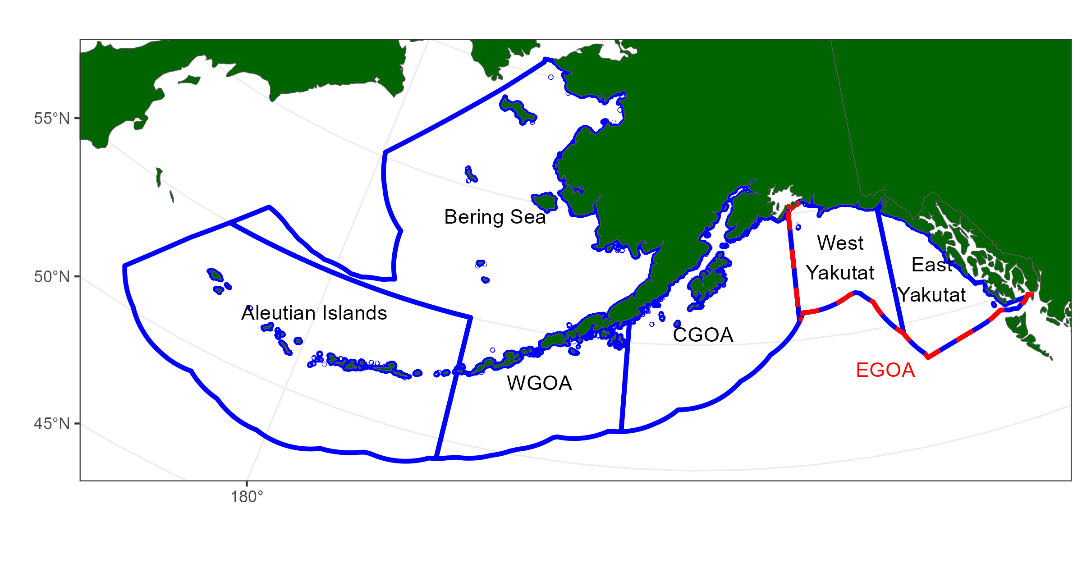
## **Table 6.** Key model quantities from the 1977 three area model.

|  |  |  |  |
| --- | --- | --- | --- |
|  | (millions) | (kilo tonnes) | Total biomass |
| AI/BS/WGOA | 12.11 | 45.97 |  |
| CGOA | 3.60 | 107.45 |  |
| EGOA | 1.52 | 79.0 |  |

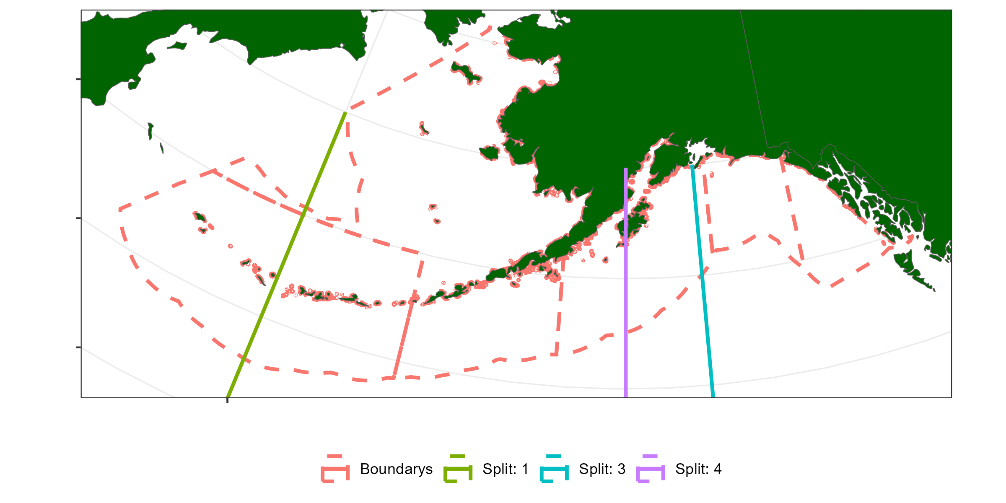
## **Table 7.** Movement matrix from 3-Region model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 3 Region Model |  | AI/BS/WGOA | CGOA | EGOA |
| AI/BS/WGOA | 0.826 | 0.117 | 0.0567 |
| CGOA | 0.0262 | 0.885 | 0.0885 |
| EGOA | 0.0415 | 0.0952 |  |
| 3 Region Model (Fenske, 2023) | AI/BS/WGOA | 0.675 | 0.223 | 0.102 |
| CGOA | 0.239 | 0.371 | 0.39 |
| EGOA | 0.079 | 0.282 | 0.639 |
| AI/BS/WGOA | 0.675 | 0.223 | 0.102 |

# **Figures**



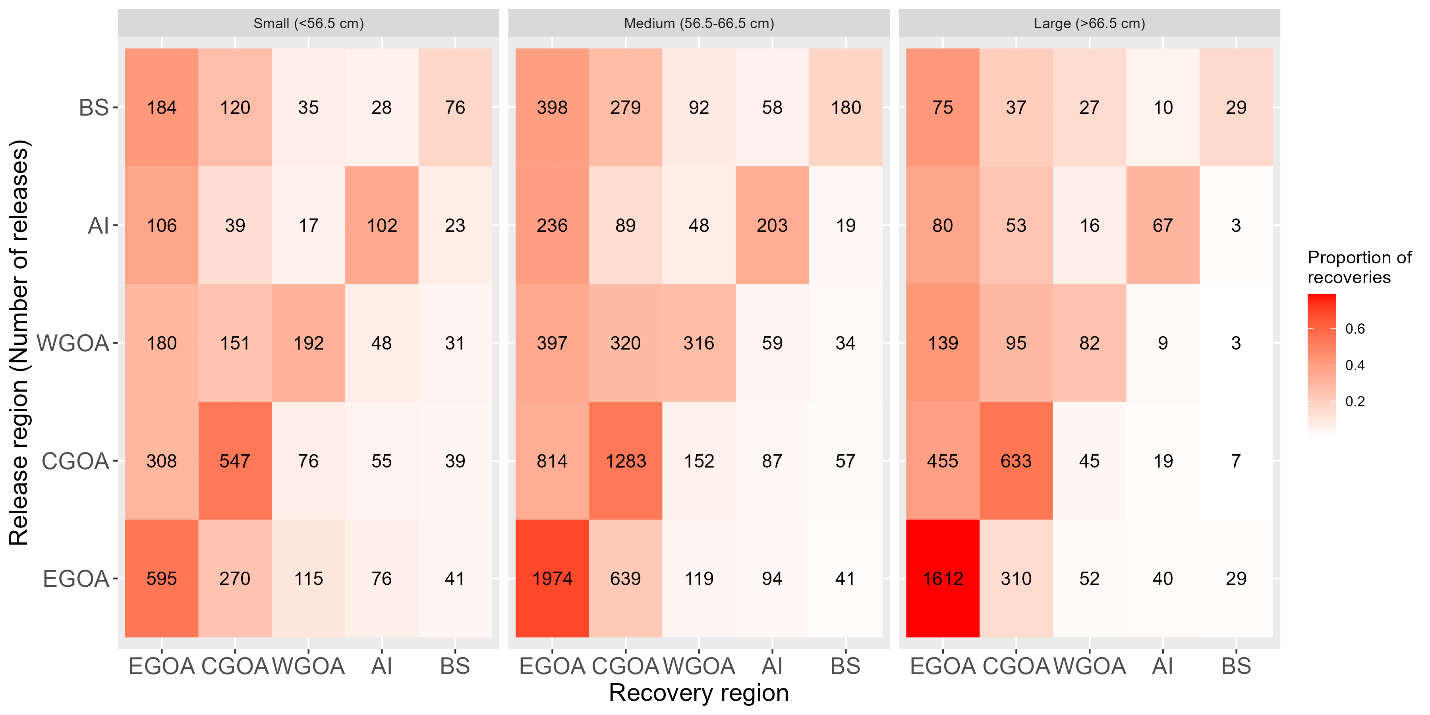
## **Figure 1.** Fishery management plan (FMP) boundaries. The Eastern Gulf is often reported at subareas split up be East/West Yakutat and sometimes Southeast.



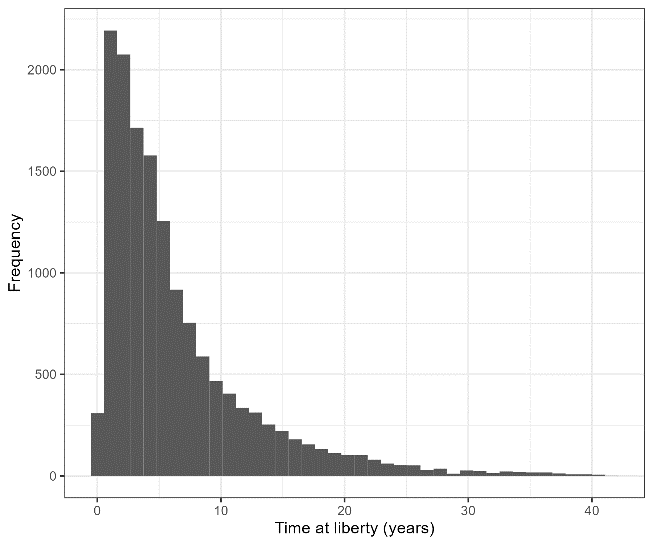
## **Figure 2.** Longitude splits based on regression tree analysis.

|  |  |
| --- | --- |
|  |  |

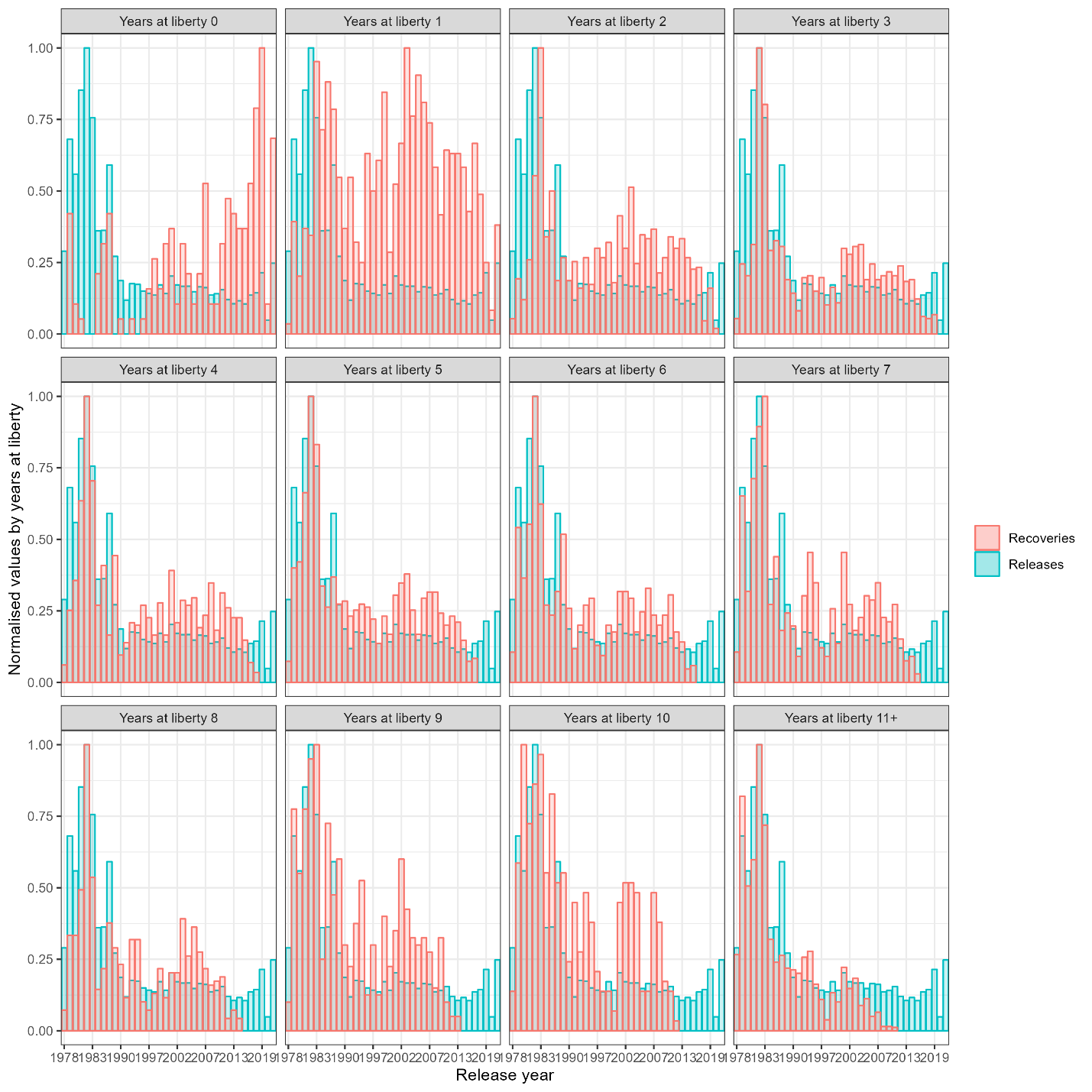
## **Figure 3.** Age-length pairs from the survey for males (left panel) and females (right panel) amongst each FMP region.



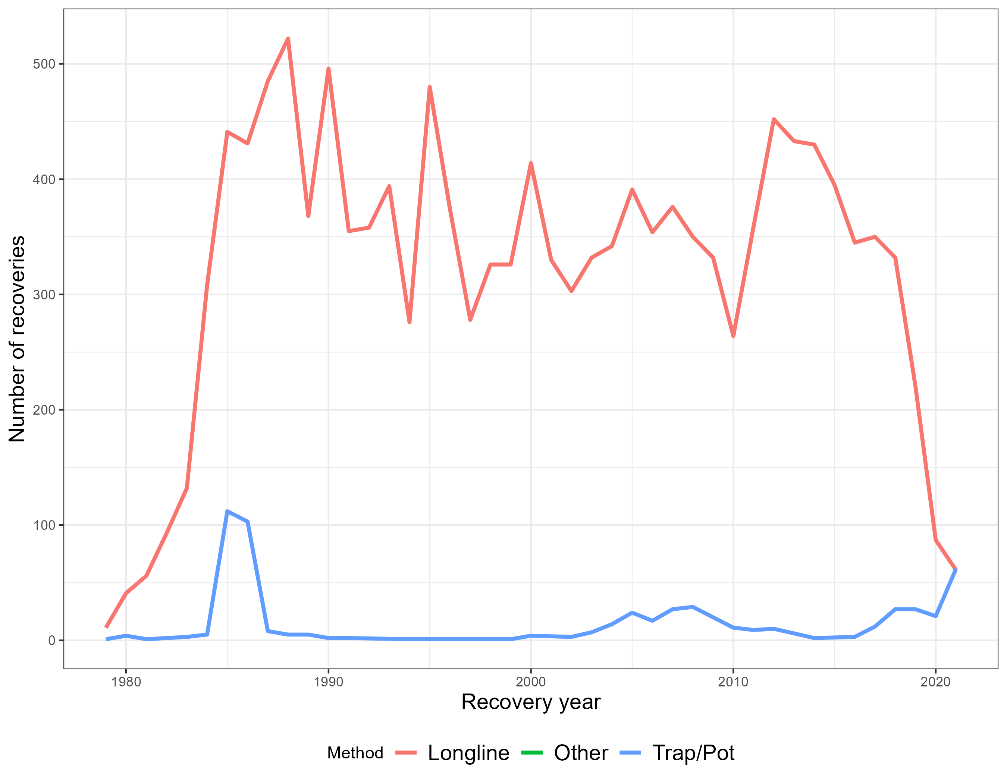
## **Figure 4.** Tag-recoveries by FMP region and size group for all tags recovered.



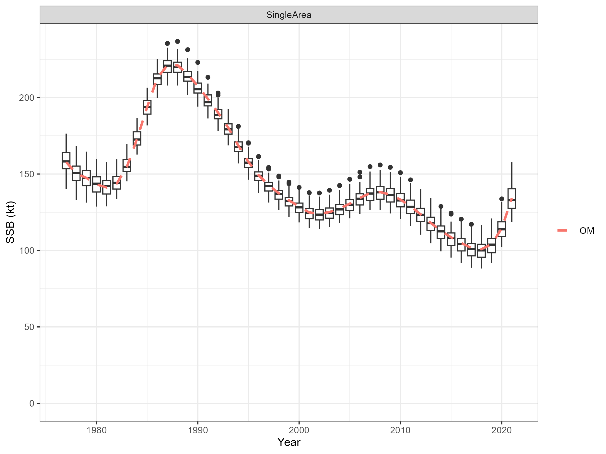
## **Figure 5.** Time at liberty from all tag-recoveries.



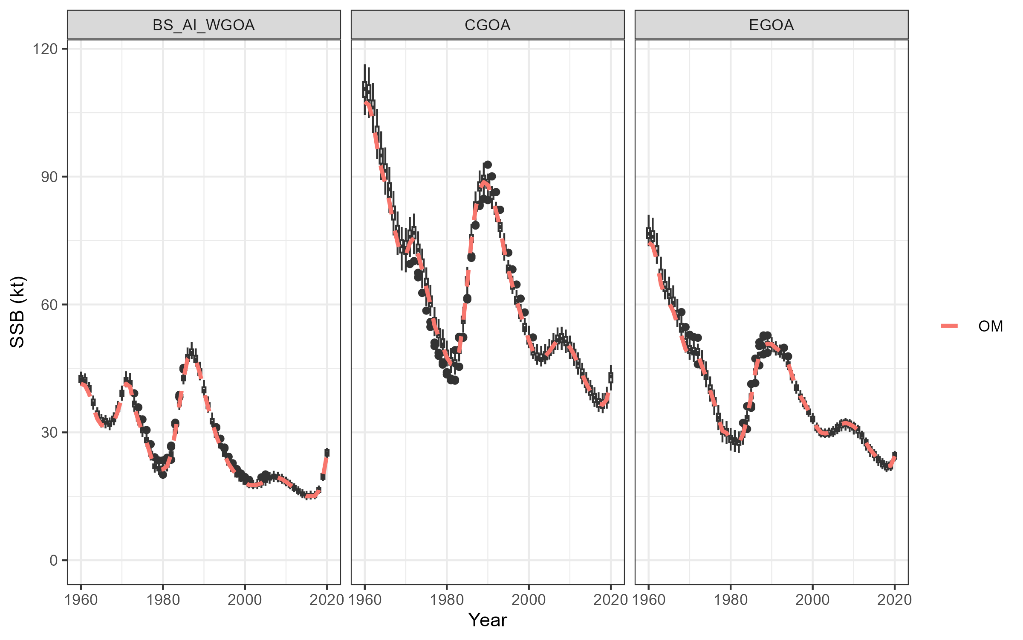
## **Figure 6.** Relative distribution of tag recoveries and tag releases, by release year (x-axis) and time-at liberty (panels).



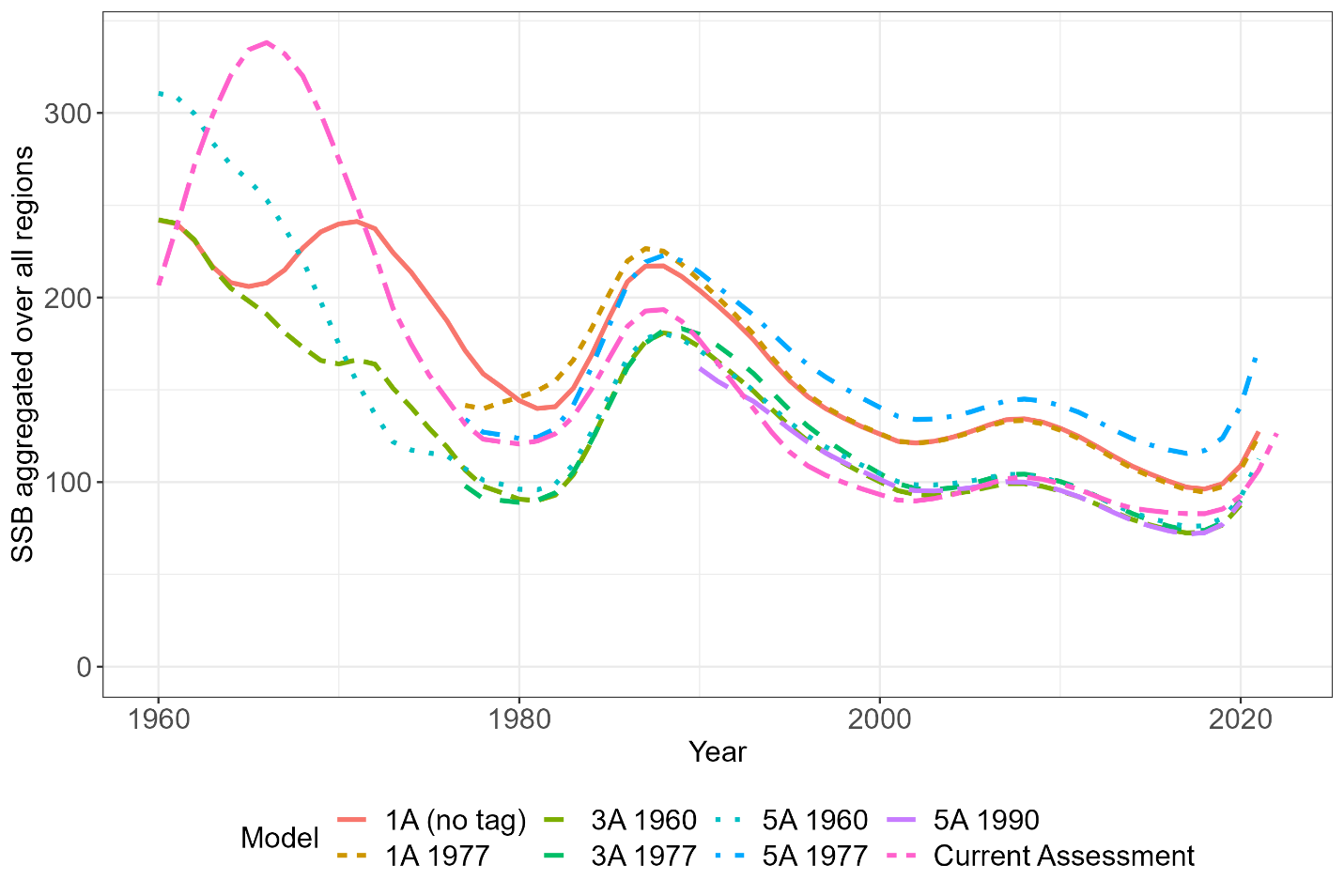
## **Figure 7.** Tag recoveries by gear type. The fixed gear type includes both Longline and Trap/Pot.



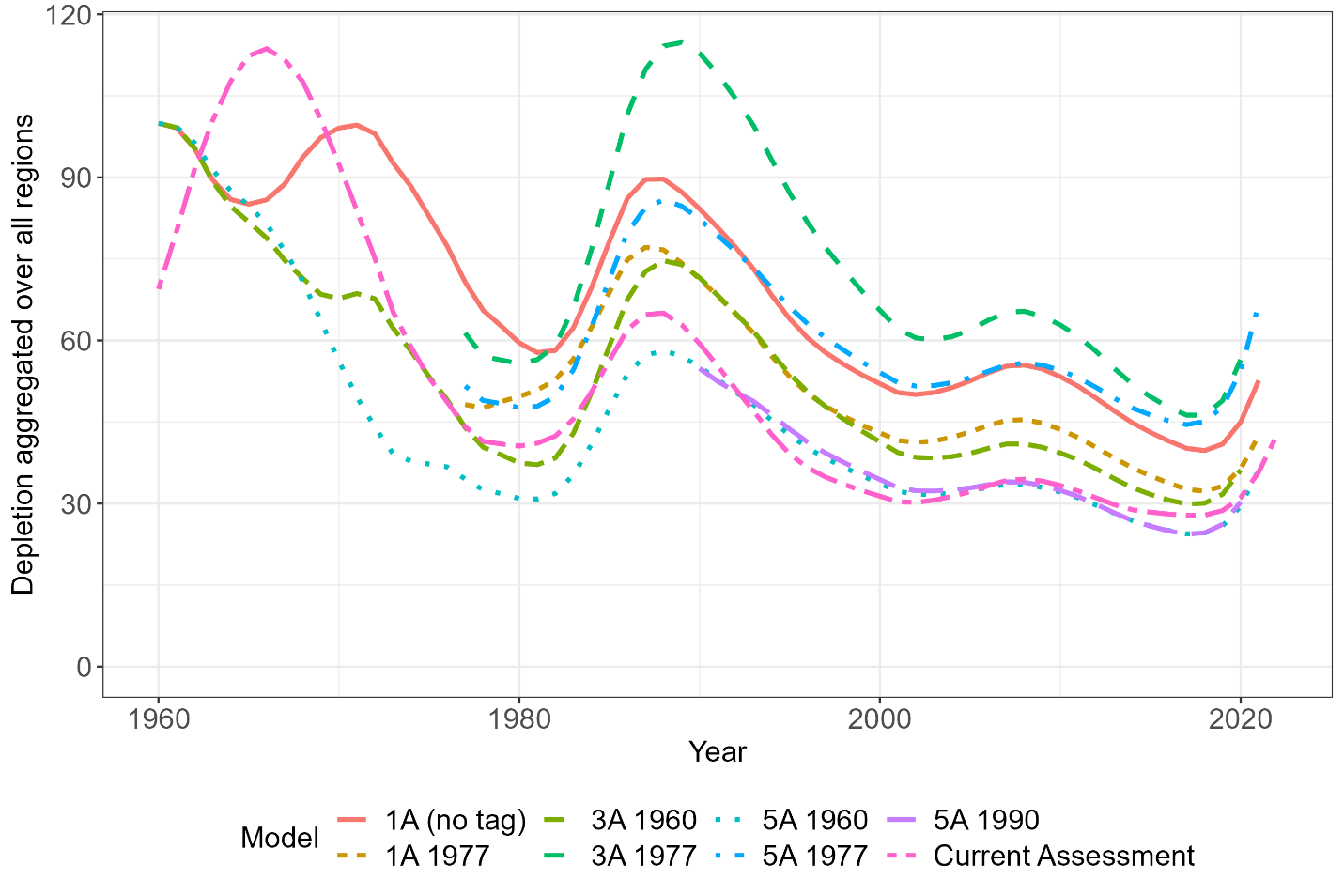
## **Figure 8.** Estimated SSBs from the self-test for the 1-Area model.



## **Figure 9.** Estimated SSBs from the self-test for the 3-Area model.



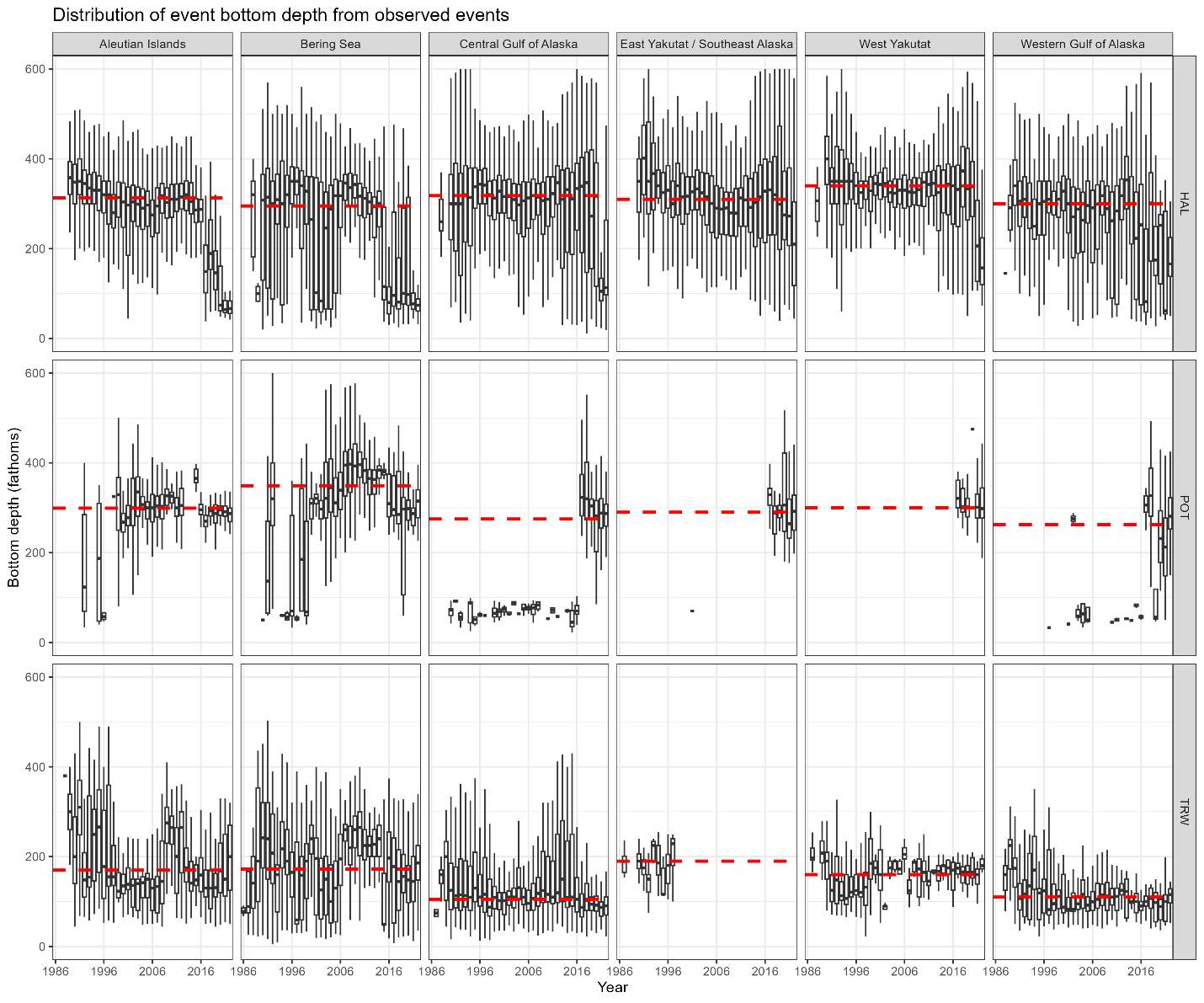
## **Figure 10.** Spatially aggregated SSB comparison of all final models considered.



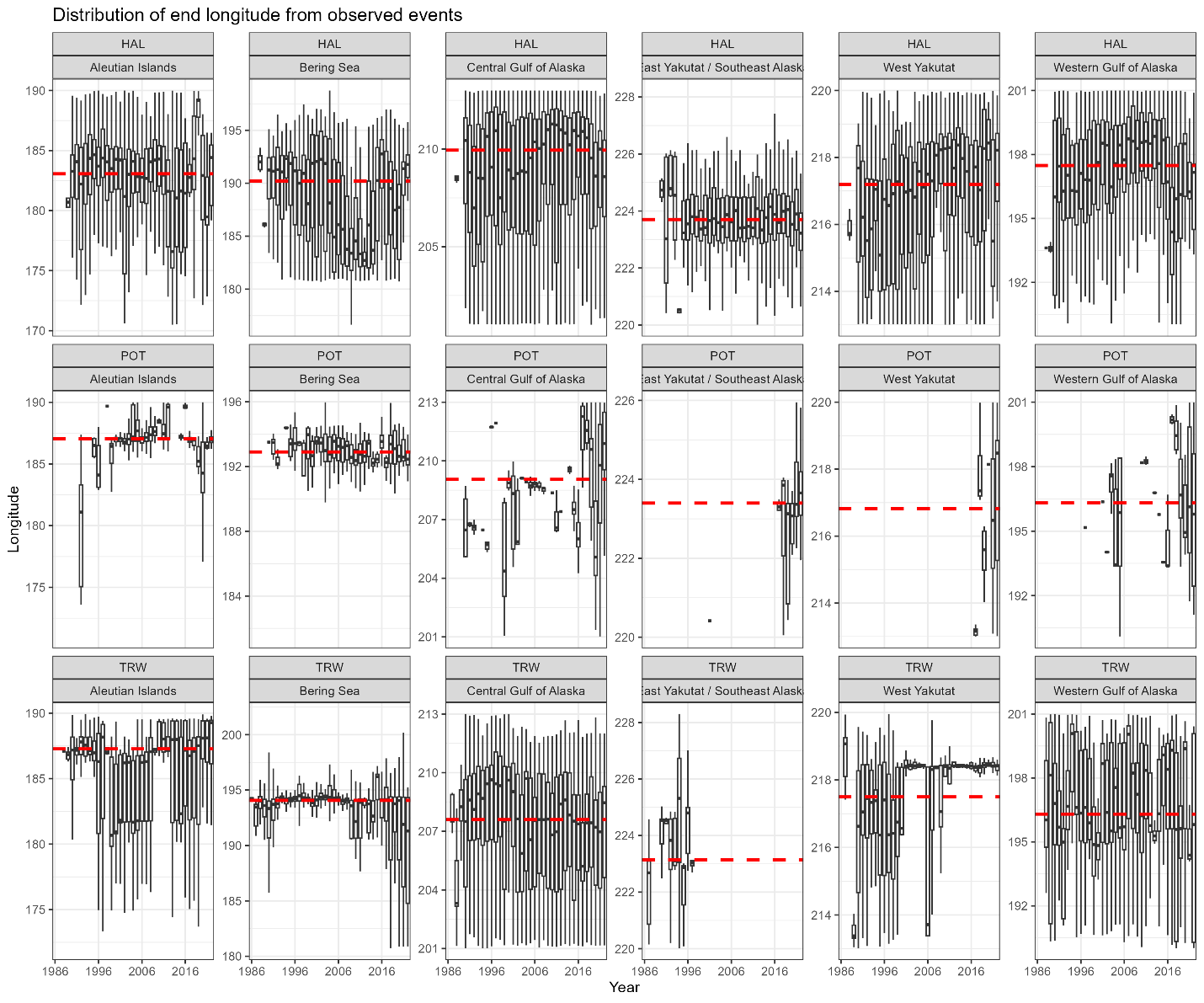
## **Figure 11.** Depletion aggregated across all spatial regions.

## **Figure 12.** Regional exploitation from the spatial models compared to the single region model assuming a domain-wide harvest control rule and regional quota apportionment using the survey.

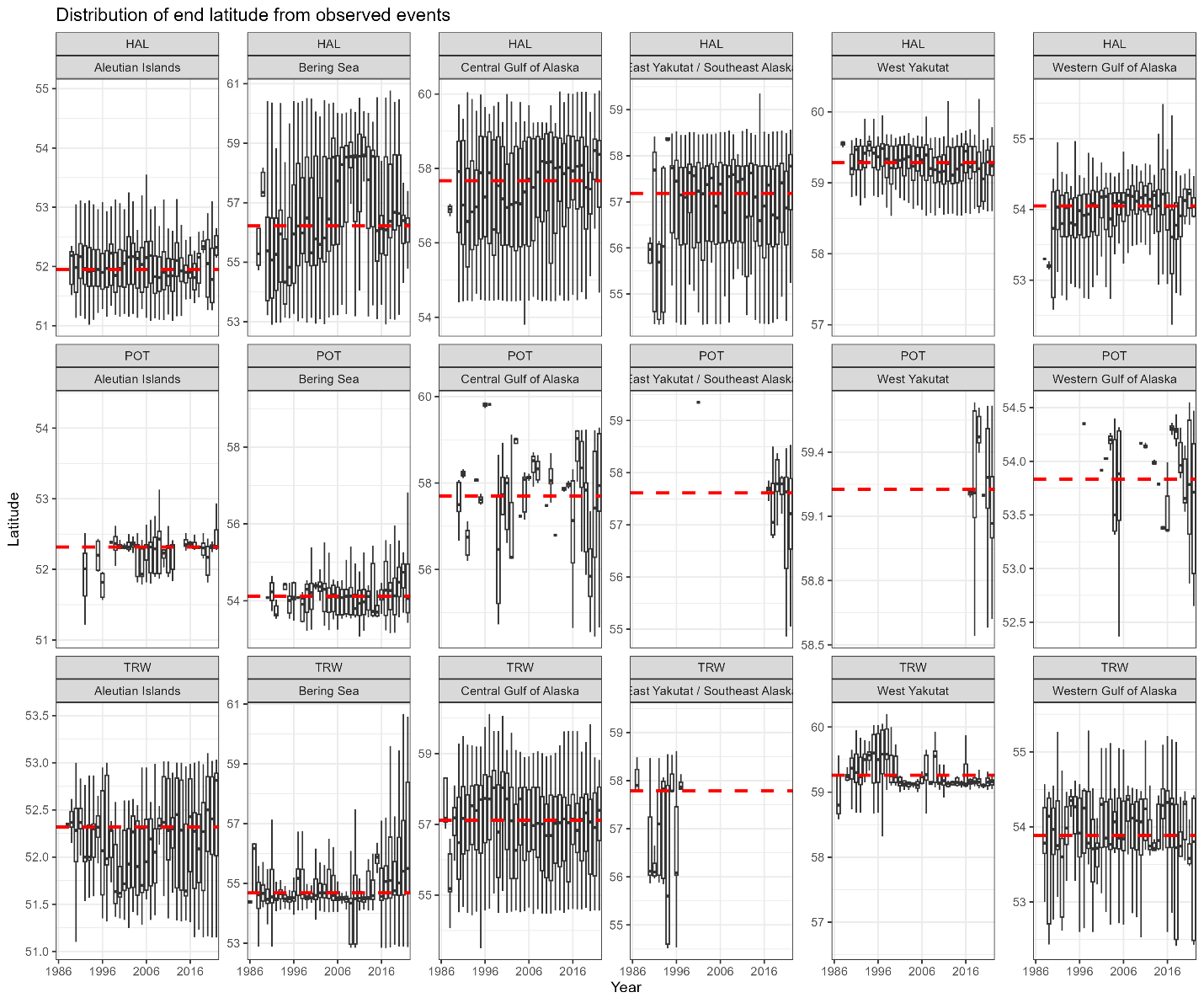
# **Appendix A**



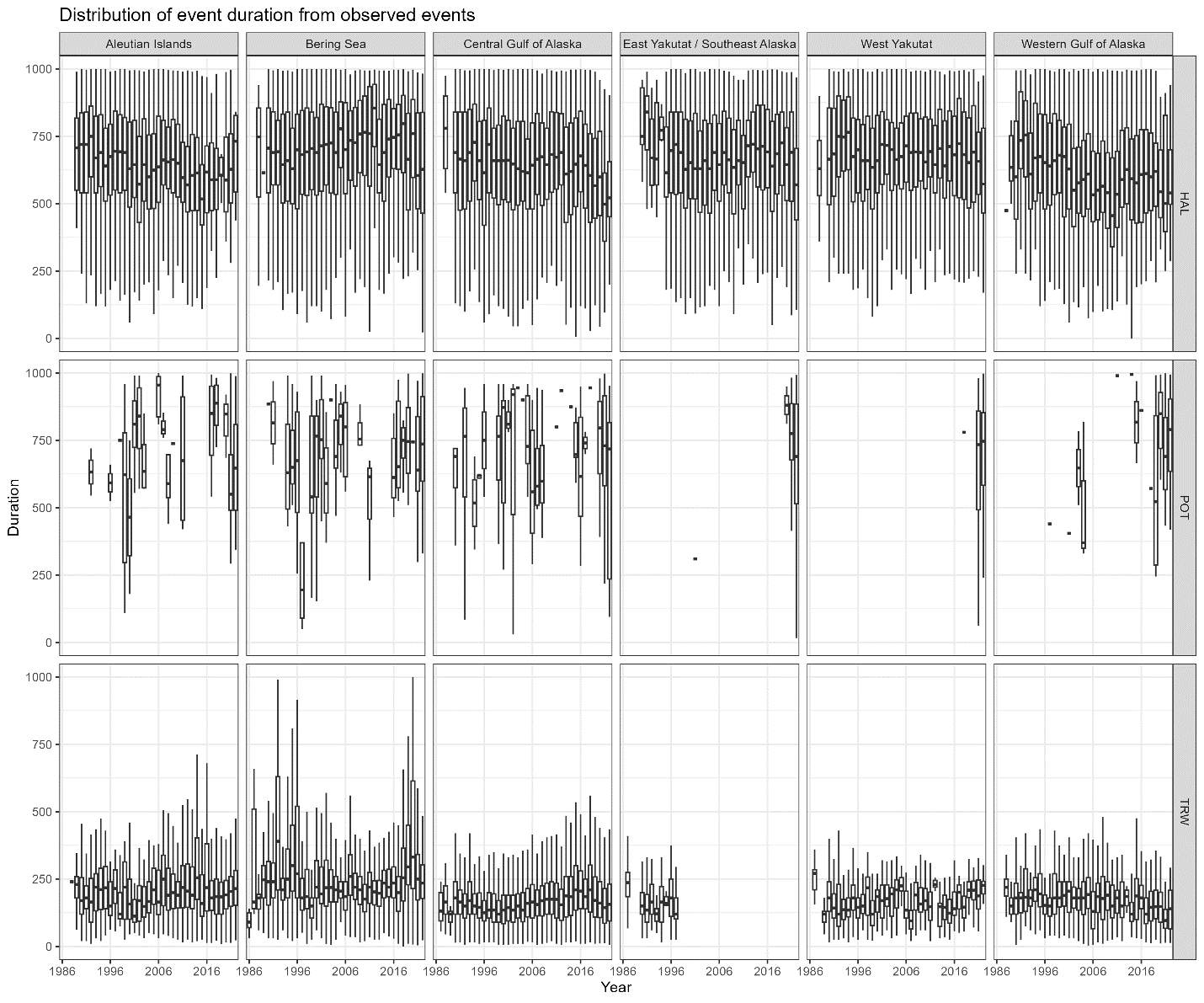
*Figure A 1: bottom depth distributed by year, gear type, and FMP region.*



*Figure A 2: Longitude distributed by year, gear type, and FMP region.*

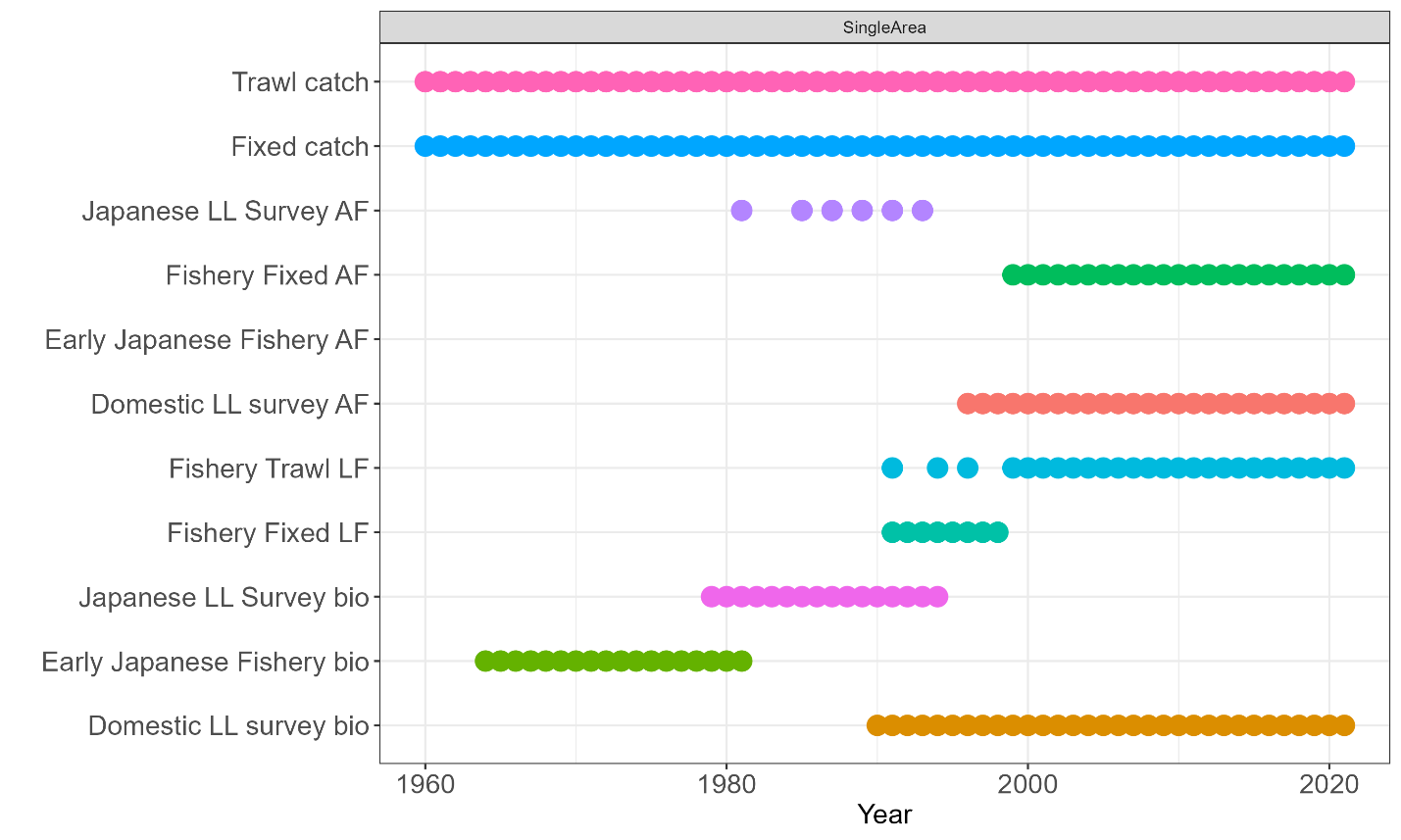


*Figure A 3: Latitude distributed by year, gear type, and FMP region.*

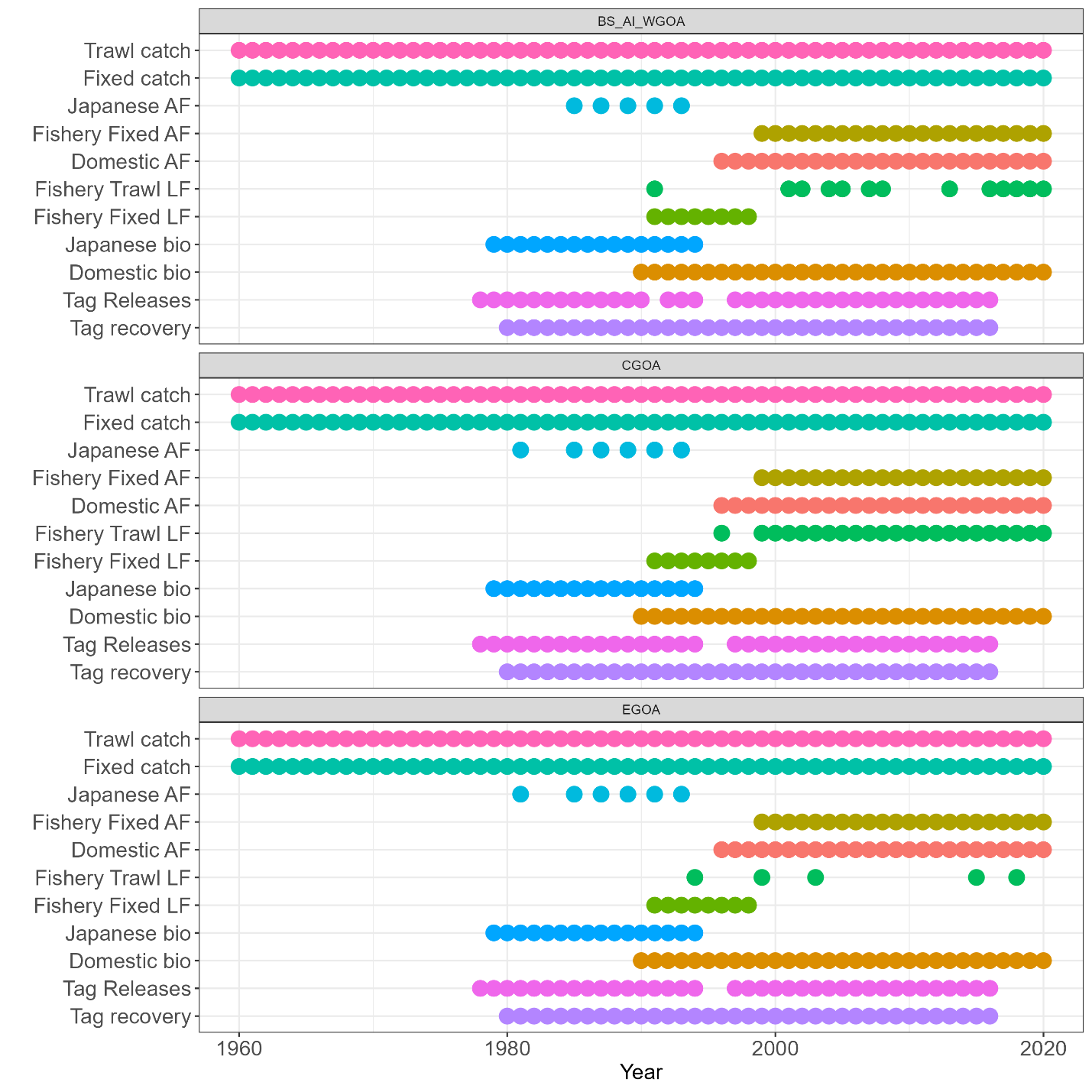


*Figure A 4: Fishing duration distributed by year, gear type, and FMP region.*

# **Appendix B**



*Figure 13:Observation frequency for 1-Area model.*



*Figure 14: Observation frequency of 3 area model.*

# **Appendix C**

## Fits to the 1960 Single area model

Fits to the 1960 Three area model



