Title: SPASAM-MSE: a generalized spatial management strategy evaluation framework for exploring the impacts of spatial structure on stock assessments and management outcomes

Running: SPASAM-MSE: A Generalized Spatial MSE Tool

Authors: Chengxue Li[1*], Jonathan J. Deroba[1], Daniel R. Goethel[2], Aaron M. Berger[3], Amy M. Schueller[4], Brian J. Langseth[5], Emily Liljestrand[1], Dana H. Hanselman[2], Timothy J. Miller[1]

Affiliations:
[1]Northeast Fisheries Science Center, NMFS-NOAA, 166 Water St., Woods Hole MA, 02543, USA

[2]Alaska Fisheries Science Center, Auke Bay Laboratories, NMFS-NOAA, 17109 Point Lena Loop Road, Juneau, AK 99801, USA

[3]Northwest Fisheries Science Center, Fisheries Resource and Monitoring Division, NMFS-NOAA, 2032 S.E. OSU Drive, Newport, OR 97365, USA

[4]Southeast Fisheries Science Center, Beaufort Laboratory, NMFS-NOAA, 101 Pivers Island Road, Beaufort, NC 28516, USA

[5]Northwest Fisheries Science Center, NMFS-NOAA, 2725 Montlake Blvd E, Seattle, WA, 98112, USA

**Corresponding author:*
Chengxue Li
Email: chengxue.li@noaa.gov

Intended Outlet: *CJFAS* (Special Issue on State-Space Methods)

Article Type: Original Research

Abstract (165 words, *CJFAS*)

Management strategy evaluations (MSEs) are helpful simulation tools for exploring the expected benefits and tradeoffs among management objectives produced by various management procedures. Few spatially explicit MSE tools exist, however, to facilitate evaluation of spatial complexity in stock assessment and management. We describe a generalized MSE R package tool (Spatial Processes and Stock Assessment Methods MSE, SPASAM-MSE) that integrates spatial population dynamic options in both the operating and assessment models to aid implementation of spatial MSE applications. The tool facilitates exploration of assessment and management performance across conditions that create spatial structure (e.g., biocomplexity, connectivity, demographics, fisheries, and management) while utilizing contemporary statistical methodologies (e.g., random effects and state-space modeling features). An example application is provided to demonstrate the utility of understanding tradeoffs in spatial management decisions. SPASAM-MSE provides a straightforward interface to consider spatial complexity in biological processes, assessment configurations, and management actions. We envision that the SPASAM-MSE tool will help facilitate increased operational implementation of robust spatial management procedures and aid management decision-making.

Keywords: management strategy evaluation (MSE), spatial stock assessment, state-space models, fisheries management, stock structure

1. Introduction

Spatial dynamics in marine species and associated fisheries are widespread, often necessitating management considerations to conserve biocomplexity and provide sustainable harvest recommendations through the use of stock assessments (Smedbol and Stephenson, 2001; Berger et al., 2017). Stock assessments typically make use of nonlinear statistical models that estimate population trends and status by fitting observed data (Hilborn, 2012). Increasingly, stock assessment modeling platforms have integrated the ability to model spatial structure and dynamics in innovative ways (Goethel et al., 2011; Berger et al., 2024), amid wider recognition that spatial structure needs to be considered to avoid inappropriate management advice or outputs at scales irrelevant to decision-making (Ying et al., 2011; Kerr et al., 2017). Although spatially explicit stock assessments are not always required to provide robust management advice (Benson et al., 2015; Li et al., 2017; Punt et al., 2017), determination of need is often case specific and requires thorough simulation testing (Cadrin et al., 2023; Goethel et al., 2023a). However, in the absence of a comprehensive analytical tool, such as a structured decision analysis framework, resource constraints often impede the development, testing, and broader evaluation of simulation-based spatial applications.

Management strategy evaluation (MSE) is a structured decision-making approach used in fisheries systems to assess the entire management process (i.e., the performance of a management procedure which includes data collection, stock status determination such as through application of a stock assessment, and implementation of harvest rule to determine resultant management action; Cox and Kronlund 2008) within a closed-loop feedback simulation. This enables systematic comparison of trade-offs among alternative management scenarios, supports the identification of management procedures that are robust to system uncertainty or biological complexity, and provides a means to explore the impact of a given management procedure before actually implementing it in the real-world. For instance, an MSE simulation typically consists of the biological population and associated processes that represents current scientific knowledge or hypotheses (i.e., the underlying 'true' dynamics of the system), management procedures, and fishery dynamics (Punt et al., 2016). Tradeoffs are explored within and across management or uncertainty scenarios using performance metrics representing desired conservation and fishery outcomes (Punt et al., 2016).

However, biological populations that display spatial population structure, such as spatial heterogeneity generated from ontogenetic life history traits that are inherent in many managed fish stocks, cannot be considered in most MSE tools. Development and implementation of spatial MSE tools is necessary to move the assessment discipline away from the de facto postulation that spatial structure is not an important driver of population dynamics, which is inherent in the unit population assumption of most stock assessment applications (Berger et al., 2024). The development of high resolution spatial simulation tools for MSEs, such as operating models conditioned on observed dynamics and the integration of realistic levels of uncertainty, can be time consuming and overwhelm resource availability during time-limited stock assessment cycles where management advice must be developed (Goethel et al., 2024).

While the use of spatially explicit operating models to explore management performance has increased (e.g., Babcock and MacCall, 2011; Gruss et al., 2016; Cunningham et al., 2019; Smith et al., 2021), few examples exist that additionally integrate spatially explicit assessments into a

MSE (e.g., Ying et al., 2011; Li et al., 2015; Kapur et al., 2024). On the contrary, most spatial MSE applications examine the implications of ignoring spatial structure by testing a spatially aggregated or spatially implicit assessment model against a spatial operating model (e.g., Benson et al., 2015; Punt et al., 2017; Jacobsen et al., 2022). Development of spatially explicit assessments can be time consuming because of added complexity, which may be one reason that MSE tools sparingly explore the integration of spatial assessments (Goethel et al., 2023a, 2024). Moreover, many MSE studies aim to understand management performance for a specific application requiring that the spatial MSE tool be tailored and conditioned to the species of interest. Tailored MSE tools help ensure the population, fishery, and management objectives are adequately reflected by the modeling framework (e.g., the two region red snapper MSE tool (Zhang et al., 2024; http://gomredsnappermsetool.fiu.edu/), the multi-region Atlantic bluefin tuna MSE (https://iccat.github.io/abft-mse/), and the five region Alaska sablefish MSE R package (https://github.com/ovec8hkin/SpatialSablefishMSE). Less attention has been devoted to the development of generalized spatial MSE packages. Although a variety of general MSE tools are available (e.g., openMSE [https://openmse.com/] and SSMSE [Doering et al., 2023; https://github.com/nmfs-fish-tools/SSMSE]), the spatial capabilities are limited. For example, openMSE has a spatial version, but users are warned that implementing models with more than two regions may be difficult (https://openmse.com/features-multimse/). Similarly, SSMSE is limited by the structure of the SS3 assessment platform, which can only accommodate a population structure with spatial heterogeneity (Berger et al., 2024).

The Spatial Processes and Stock Assessment Methods MSE (SPASAM-MSE) tool was developed to provide practitioners with a flexible MSE simulation tool that can integrate all common spatial population structures into operating models (OMs) and estimation models (EMs), which may differ in their spatial complexity. SPASAM-MSE is a fully generalized, spatially explicit MSE that can be applied to almost any species and fishery, while explicitly integrating associated spatial dynamics across multiple areas. Development of SPASAM-MSE is a result of ongoing work (e.g., Goethel et al., 2019, 2021; Bosley et al., 2019, 2022; Berger et al., 2021, Li et al., *In revision*) aimed at increasing the use of spatial stock assessments to provide robust fisheries management advice. In this article, we first describe the model structure and primary spatial features of SPASAM-MSE. We then demonstrate the spatial capabilities, types of standard performance metrics available, and visualizations that are readily available to summarize MSE results using an example application. Key objectives of the versatile SPASAM-MSE tool include explicitly exploring spatial dynamics, promoting greater recognition and integration of spatial structure in stock assessments, enhancing understanding of the associated trade-offs affecting management performance, and lessening the resource burden that often acts as a barrier to exploring spatial management procedures.

2. Model Description

The SPASAM-MSE package, developed in R (https://lichengxue.github.io/whamMSE/), is a flexible and modular framework for conducting MSE, with a strong emphasis on spatial processes. Simulations can be conditioned on a range of inputs, including observational data, stock assessment outputs, and real-world management scenarios. The package also supports simulations of generic fish populations using life-history traits drawn from a built-in library of generalized life-history types (e.g., short-, medium-, and long-lived species) (Wiedenmann et al., 2017). Spatial structure can be incorporated throughout the MSE framework, including within population dynamics (e.g., demography, movement, and population structure), fleet dynamics (e.g., spatial fleet behavior and fishing pressure), data generation and collection (e.g., spatially explicit data streams), assessment model configuration (e.g., panmictic, spatially implicit, spatially explicit), biological reference point estimation (e.g., accounting for movement, population structure, and regional dynamics), and other aspects of the management procedure [e.g., harvest control rules (HCRs), catch apportionment; Figure 1].

2.1 Model Development and Design Considerations

The SPASAM-MSE framework was developed with the primary goal of evaluating the implications of spatial structure on stock assessments and fisheries management outcomes. The intent was to build an MSE tool that explicitly prioritized spatial realism in both population and fleet dynamics (where fleet is defined as a specific gear type in a specific region; Bosley et al., 2019; Berger et al., 2021). Key spatial processes, such as spatial heterogeneity in demography, fishing pressure, and fleet behavior, as well as more complex population structures and movement dynamics, including metapopulations and natal homing, were central to the design.

Although SPASAM-MSE incorporates many key processes relevant to spatial fisheries management, it is not intended to serve as a universal MSE tool. For example, although the framework has a modular structure that allows users to implement custom HCRs, it does not currently include the full range of HCR types available in other generalized MSE platforms. The tools primary strength lies in the ability to explore spatial complexity rather than to provide comprehensive coverage of all possible management scenarios. We view this specialization as a core strength for users focused on spatial questions, while maintaining the structure extensibility needed to support broader applications (e.g., simulation-estimation experiments).

2.2 MSE framework

The SPASAM-MSE framework consists of an OM that simulates the true population and fishery dynamics along with ‘observed’ pseudo-data; an EM that estimates population parameters and population status; a catch projection module that provides catch advice based on estimated status and the HCR; and a feedback loop that returns catch advice to the OM to simulate population and fishery dynamics in the following year(s). The MSE simulation period is split into two time periods, where the ‘historical’ period simulates the dynamics based on user-defined inputs to initialize the model up until the start of the ‘feedback’ period at which time the full MSE closed loop management simulations begin and the desired management procedures are applied. The package allows users to customize nearly all components of the MSE framework to evaluate alternative management strategies (Table 1).

2.2.1 User Interface and R Package

SPASAM-MSE is implemented as an R package designed to support flexibility, transparency, user customization, and collaboration. Users interact with the framework primarily through a set of wrapper functions that define biological, fishery, and spatial dynamics. Several features are included to streamline model development and reduce technical barriers, including an R Shiny application for configuring movement patterns, built-in libraries for generic life-history types, and utility functions for visualizing outputs and model diagnostics. These tools help users build, run, and interpret complex spatial MSEs without needing to write custom code for every component. In addition, the modular coding design allows users to easily modify or extend the package to meet their specific simulation needs.

2.2.2 Overview of Modeling Framework

SPASAM-MSE uses the WHAM (Woods Hole Assessment Model; https://timjmiller.github.io/wham), an age-structured state-space stock assessment platform, as the foundation for its modeling framework (Stock and Miller, 2021; Miller et al., 2025). The WHAM provides the flexibility, extensibility, advanced state-space modeling features, and spatially explicit modeling capabilities needed to support SPASAM-MSE's objectives. WHAM is currently used in most fish stock assessments in the U.S. Northeast region and offers a well-tested infrastructure for conducting assessments that can be adapted for spatial applications.

The SPASAM-MSE framework builds upon the WHAM modeling framework, which already supports (1) spatial heterogeneity in demography and complex population structures; (2) integration of process error through random effects; (3) the ability to link environmental covariates to both biological and fishery processes; and (4) incorporation of multiple observation types with varying distributional assumptions. Building on this foundation, SPASAM-MSE extends WHAM's capabilities to support a broader range of applications, including complex spatial population structures (e.g., metapopulations) and diverse movement patterns such as ontogenetic movement (see https://github.com/lichengxue/wham).

2.2.3 Model Conditioning

The OM is conditioned using a series of user inputs organized into five main components: (1) fish biology, including recruitment, weight-at-age, maturity-at-age, natural mortality, and movement; (2) fishery and survey characteristics, such as fishing mortality, gear selectivity, and survey catchability; (3) observation model inputs, including the coefficient of variation and effective sample sizes for aggregate catch observations across fleets and surveys, and the likelihood distributions for age composition data; (4) process error associated with biological and fishery processes; and (5) environmental covariates and their effects on underlying processes. Inputs for OM conditioning can be derived from a stock assessment fit to real-world data or specified based on user-defined values and expert judgement (e.g., for generic or exploratory applications).

2.3 Operating Model (OM)

2.3.1 Age and Temporal Structure

SPASAM-MSE is fundamentally an age-structured model, though users may input length-based relationships (e.g., length–weight, length-at-age), which are internally converted into age-based quantities for use in the model. At present, SPASAM-MSE also does not support sex-specific functionality.

SPASAM-MSE defines population dynamics using a flexible temporal structure, in which each year is divided into user-defined seasonal time steps. A single time step per year is also supported for models without seasonal dynamics. Both the number and duration of seasonal steps can be specified by the user, allowing the model to accommodate a range of biological or fishery time scales. Key demographic and fishery processes, such as natural mortality, movement, and fishing mortality, are represented through probability transition matrices applied within each time block. These matrices are multiplied sequentially across all seasonal steps to produce the full annual transition matrix (Miller et al., 2025).

2.3.2 Population and Spatial Structure

Users must define the number of geographic regions to model (hereafter "regions"), along with the number of population components, which are interpreted differently depending on the assumed population structure. SPASAM-MSE supports a range of population configurations, including panmictic, spatial heterogeneity, natal homing, and metapopulation structures (Figure 2). In the panmictic configuration, a single population is assumed to be homogeneously distributed across a single region, with no spatial differentiation. In the spatial heterogeneity configuration, a single biological population spans multiple regions but remains reproductively and demographically well-mixed. However, a spatially heterogeneous population may exhibit regional phenotypic variation in biological traits, environmentally driven differences in recruitment, or regional fishing pressure. Regional recruitment in this case is apportioned from a global stock–recruit relationship. Natal homing population structure allows for multiple population components, each assigned to a unique natal spawning region with a unique stock–recruit relationship. Mature individuals return to their natal region to spawn each year, although limited reproductive straying can be modeled. During non-spawning periods, populations may overlap spatially. The metapopulation structure assumes that each region contains a sub-population with its own local stock–recruit relationship and that extensive reproductive mixing occurs across regions. Individuals adopt the biological and spawning characteristics of the sub-population in the region they currently occupy. In a meta-population, region and sub-population are synonymous, such that all fish in a given region are assigned to the associated spawning component, and only one spawning population is permitted per region. In addition, users may set recruitment in any region to zero, effectively designating it as a sink supported solely by immigration from source regions.

2.3.3 Demographics

The package accommodates a range of life-history traits by allowing users to specify either generic or custom biological parameters. Users can select from pre-specified short-, medium-, or long-lived life-history profiles available in the built-in library, which includes parameter sets for length-at-age (via von Bertalanffy growth functions), length–weight, and maturity-at-age relationships. These are automatically generated based on the selected life-history profile and the user-defined maximum age or plus group (Wiedenmann et al., 2017). Alternatively, users may directly provide annual values for biological parameters such as weight-at-age, maturity-at-age, and natural mortality. Weight-at-age can be specified as time-varying to reflect interannual variability in growth and can differ across fleets, surveys, regions, and populations. Other biological parameters can also vary to represent spatial heterogeneity. For example, maturity-at-age can vary by population, natural mortality by population and region, and the stock–recruit relationship by population.

2.2.4 Movement Dynamics

SPASAM-MSE adopts WHAM's movement modeling approach, which allows movement to be modeled either simultaneously with survival or sequentially after survival within each time interval. In WHAM, movement is represented by mean rates that may differ across populations, regions, age classes, and seasons, with additional variability introduced through random effects (Miller et al., 2025). Building on this foundation, SPASAM-MSE supports ontogenetic movement by allowing mean movement rates to vary across ages, thereby capturing more biologically realistic patterns observed in many marine species (e.g., Liljestrand et al., 2019; Jacobsen et al., 2022). Ontogenetic movement can be modeled using flexible functional forms, such as increasing or decreasing logistic curves, or dome-shaped double logistic functions, to reflect age-specific movement behaviors (e.g., Figure 3b and 3d). For instance, older individuals may exhibit greater movement due to expanded home ranges, while juveniles may disperse in search of nursery habitats. Users can also manually assign movement rates to specific age classes to reflect case-specific or application-specific scenarios. Variability in ontogenetic movement can also be incorporated via random effects (e.g., Figures 3a and 3c). Additionally, users may impose a directional trend by adding a constant annual increase to the mean movement rate, allowing nonstationary movement dynamics (Figure 3c). This feature allows SPASAM-MSE to simulate long-term directional changes in movement behavior, such as climate-driven poleward shifts in species distributions.

Configuring movement dynamics typically requires detailed input, often in the form of high-dimensional arrays. As spatial and temporal resolution increases and movement behavior becomes more complex, input preparation can become tedious and error-prone. To simplify this process, an R Shiny application built into SPASAM-MSE was developed to assist users in setting up movement configurations (Figure 4). The app provides a user-friendly, interactive interface aligned with the current built-in movement options, streamlining setup and reducing input errors. The built-in R function provides several preset movement patterns with user-defined movement rates, which can also be used to generate movement configurations for simulation studies.

2.2.5 Fleet Dynamics

Users can specify the number of fleets operating within the spatial domain and assign each fleet to a specific region. Each fleet can have a unique selectivity-at-age pattern, defined either by functional forms (increasing logistic, decreasing logistic, or double logistic) with associated input parameters (e.g., slope and age at 50% selectivity), or by directly specifying selectivity at each age. Users can also define historical fishing mortality (F) for each fleet. These configurations play a critical role in shaping population dynamics and determining population status at the start of the feedback period. Several built-in options are available for defining historical F (e.g., Table S1). By default, historical F is specified at the age corresponding to the maximum fully selected total F. That is, the age at which the sum of F across all fleets is highest. Alternatively, users may provide a fully customized, fleet-specific F matrix to override the default settings.

2.2.6 Random Effects and Environmental Covariates

In the SPASAM-MSE framework, the structure and magnitude of OM process random effects are inherited directly from the WHAM modeling framework (Stock and Miller, 2021; Stock et al., 2021; Miller et al., 2025). Users can set a random number seed, and the model will automatically generate random effects according to the specified error structure (e.g., independent or autocorrelated). Random effects can be applied to biological and fishery processes such as recruitment, numbers-at-age transitions, natural mortality, selectivity, survey catchability, and movement, allowing direct control over the magnitude and structure of simulated variability across populations, regions, ages, and seasons.

Environmental covariates are also supported in the OM, consistent with WHAM, including their process error structure, observation error, and parameters that govern mechanistic linkages to biological processes (e.g., recruitment, natural mortality, survey catchability, and movement). These linkages can be modeled using linear or nonlinear relationships (e.g., polynomial) and may include temporal lags. Multiple environmental covariates can act simultaneously on different processes within and across populations and regions, providing flexibility to explore complex, biologically realistic scenarios. The environmental process itself can be configured as a state-space model (with both process and observation error). Users may either generate pseudo time series by specifying mean values and error magnitudes (with random draws governed by the seed) or provide observed environmental time series directly. This design provides flexibility in how environmental variability is represented and how it influences simulated population dynamics.

2.2.7 Observation Model

The observation model simulates how data are generated from the "true" population and fishery dynamics produced by the OM by incorporating observation error. Two main data types are supported: region-specific indices (e.g., total catch or survey indices, either abundance- or biomass-based) and age composition data. Indices are generated as aggregated summaries at the OM regional level and may represent either fishery-dependent (e.g., fleet catch) or fishery-independent (e.g., survey catch) sources (Tables S5–S6). Users can specify whether catch and survey data are provided only as indices or are also accompanied by age composition data. They can also define whether surveys are abundance- or biomass-based, set survey catchability coefficients, and specify the timing of survey operations. Observation error for catch and index data is modeled using a log-normal distribution, with variability controlled by user-specified coefficients of variation.

For age composition data, users can select from various likelihood types and specify either effective sample sizes or standard deviations to control the precision of proportions-at-age. Available likelihoods include count-based options (e.g., multinomial, Dirichlet–multinomial) and proportion-based options (e.g., Dirichlet, logistic-normal) (Table S7). When using self-weighting likelihoods (e.g., Dirichlet–multinomial or logistic-normal), users can additionally specify parameters that govern overdispersion or covariance structure.

Currently, tagging data are not simulated in the SPASAM-MSE framework, which is a known limitation. Likewise, fishery catch per unit effort (CPUE) cannot be explicitly simulated, but a survey index with similar characteristics to a given fleet could be included to approximate that type of data, including time-varying catchability that is common in fishery CPUE indices.

2.3 Data Collection and Processing

The data processing module serves as the critical bridge between the OM and the EM in the spatial-MSE feedback loop by enabling selection, transformation, aggregation, and integration of spatially structured data to match the spatial resolution of the EM. When aggregation across regions is required, for example to translate spatially explicit OM outputs into region-aggregated or panmictic EM inputs, the module supports multiple schemes. The simplest is sum-based aggregation, in which total values, such as catch or abundance indices, are summed directly across regions. For age-structured data where maintaining demographic composition is important, catch-weighted aggregation is used: catch-at-age from each fleet or region is multiplied by its regional total catch, summed by age across regions, and then converted to proportions by dividing each age-specific catch by the total catch. An additional equal-weighted aggregation option treats each region equally, which can help reduce bias when catch is not a reliable proxy for population distribution or spatial sampling effort is uneven.

The OM retains both the "true" values (e.g., regional catch or survey indices) and corresponding "observed" values, which incorporate simulated observation error. This framework enables users to either: (1) specify the observation error for aggregated data directly, or (2) derive that error structure by comparing true and observed values at the aggregate level. For example, observation error in total catch or indices can be quantified by evaluating the variability between the sum of true values across regions and the sum of observed (i.e., error-affected) values. This method provides a flexible means to define or revise observation-error assumptions in the EM, depending on how users choose to simulate the realism and uncertainty of available data.

Users can specify the number of years of data passed from the OM to the EM as either a fixed constant (e.g., always the most recent 20 years) or as an increasing number of years as the feedback loop progresses. They can also specify observation error settings within the EM, including coefficients of variation for aggregate catch or indices, effective sample sizes or standard deviations for age composition, and choice of likelihood type and catchability coefficients for surveys. The EM specifications do not have to match those of the OM.

The module also supports the aggregation of biological data (e.g., weight-at-age, maturity-at-age) to ensure consistency with the spatial resolution of the EM by combining values across regions. Both maturity-at-age and weight-at-age are combined among regions for each fleet or survey using the catch weighted average.

2.4 Estimation Model (EM)

All components of the OM, including temporal structure, spatial and population configuration, demographics, movement, fleet dynamics, and environmental variability, can be mirrored in the EM, enabling the EM to match the OM's complexity when fitting data from a spatially explicit OM. The SPASAM-MSE framework also allows for simpler EM structures to explore the effects of spatial mismatch. A detailed comparison of EM configurations [Panmictic (PAN), Fleets-as-Areas (FAA), Separate Panmictic (SEP), Spatially Disaggregated (SpD), Spatially Explicit (SpE)] is provided in Table 2.

In the current version of SPASAM-MSE, the EM can match most movement structures specified in the OM. Exceptions include structural forms of ontogenetic movement (e.g., age-based logistic movement) and systematic temporal trends, which cannot be exactly replicated when movement is estimated within the EM. Instead, the EM approximates these dynamics through movement random effects, which can optionally be correlated among ages or years. These random effects capture variability consistent with the OM but do not reproduce its exact structural forms.

The structural differences between the OM and EM influence how well spatial dynamics are represented. If the OM includes movement that the EM ignores or simplifies, then bias is likely to occur in estimated recruitment or mortality. In other words, the EM may mistakenly attribute movement-driven changes to other biological processes, caused by demographic leakage where demographic signals that "leak" across space are not correctly modeled with movement. This type of leakage has been shown to introduce bias in regional estimates when spatial boundaries are incoherent between models and reality (Kerr et al., 2017; Berger et al., 2021). Such bias can undermine the effectiveness of management procedures that rely on spatially aggregated indicators.

In addition, SPASAM-MSE offers flexible spatial data aggregation so users can simplify their EM relative to the OM. For example, if the OM includes multiple regions, then users can combine data from any combination of regions or omit regions altogether. This flexibility allows for evaluation of how merging or excluding spatial units affects assessment and management. For example, this feature would be useful to evaluate the effects of a marine protected area or offshore wind farm that can create spatial heterogeneity in fishing pressure, survey availability, or fish density.

2.5 Management Module

The management module includes determining how frequently the EM is applied in the feedback period, calculating biological reference points, evaluating population status to inform the HCR, and applying the HCR to project catch. The module can also incorporate management implementation error, which may result in realized catch differing from the projected catch.

2.5.1 Biological Reference Points (BRPs)

The method used to calculate global biological reference points (BRPs) for spatially explicit models follows the approach described in detail in Miller et al. (2025). The global $F_{X\%}$ is defined as the fishing mortality rate that reduces the global spawning potential ratio (SPR) to X%, where SPR is the ratio of spawning biomass per recruit (SSBPR) at a given fishing mortality rate to that in the unfished state. When a Beverton-Holt or Ricker stock-recruit relationship is assumed, the global F_{MSY} is computed as the fishing mortality that maximizes the equilibrium yield, defined as the product of equilibrium recruitment and yield per recruit (YPR). Inputs to the SSBPR and YPR calculations are averaged over a user-specified set of years to reflect prevailing biological and fishery conditions.

In a single-region panmictic model (e.g., PAN or FAA), only one SPR is calculated, which by definition serves as the global BRP. Similarly, in SEP when BRPs are estimated separately by region, each region-specific SPR functions as a distinct local BRP, consistent with the assumption that regions represent biologically independent populations.

For BRP calculations in spatially structured models, a total F-at-age is computed by summing F-at-age values across fleets and regions. The age with the highest total F-at-age defines the fully selected age, and the corresponding F is considered the global fully selected F. The equilibrium global SPR is then calculated as a weighted sum of SPRs across populations and regions, conditioned on prevailing conditions (average weight-at-age, maturity-at-age, natural mortality, selectivity, and movement over the last n years). Weights may be based on (1) average regional recruitment over the past n years, (2) average regional fleet catch or survey indices over the past n years, or (3) user-specified values.

Fishing mortality in a given region may affect multiple populations that reside in or migrate through an area due to movement among regions. Thus, even though each population has a distinct SPR curve, the total F experienced by individuals of a population can be spatially distributed. This means that calculating an $F_{X\%}$ for a given population must account for the cumulative F it experiences across regions. Since different populations may have different biological parameters and movement patterns, the same total F may not produce an SPR of exactly X% for all populations simultaneously. A global $F_{X\%}$ therefore represents a weighted SPR across all populations, with weights defined as above.

Under natal homing dynamics, each population returns to its natal region to spawn, meaning that regions and populations are effectively synonymous in BRP estimation. Because reproductive mixing is limited, biological characteristics such as weight-at-age, maturity-at-age, natural mortality, and the stock–recruit relationship, can be defined at the regional level and are assumed to remain consistent within each region during the spawning season.

Under metapopulation dynamics, individuals from multiple subpopulations may occupy the same region during spawning. For reference point calculations, it is assumed that incoming individuals adopt the biological characteristics of the population associated with the region they enter, rather than retaining traits from their natal origin. Recruitment in each region is assumed to depend solely on the local SSB, without tracking the original source population of the spawners.

2.5.2 Harvest Control Rules (HCRs)

Users have the option to set a variety of HCRs based on $F_{X\%}$ or F_{MSY} (when a Beverton-Holt or Ricker stock–recruit relationship is assumed). For instance, the F reference level can be set at a constant rate ($F_{X\%}$ or a user input value), a biomass-based HCR with target F dependent on stock status can be defined (e.g., a threshold or sloped HCR), or a constant catch can be input (Deroba and Bence, 2008). Biomass-based HCR implementations adjust the target F (i.e., fraction of the F reference point) linearly based on the ratio of current SSB to the reference SSB (e.g., $SSB_{X\%}$ or SSB_{MSY}). Users can configure the upper and lower SSB thresholds (SSB_{Thresh_up} and SSB_{Thresh_low}) between which F changes, as well as the maximum and minimum percentages of $F_{X\%}$ or F_{MSY} to apply. When $SSB \geq SSB_{Thresh\_up}$, the maximum percentage of $F_{X\%}$ or F_{MSY} is used; when $SSB \leq SSB_{Thresh\_low}$, the minimum percentage of $F_{X\%}$ or F_{MSY} is used. Between the thresholds, a linear interpolation is used to scale the applied rate. Currently, the code for HCRs is designed to be modular in SPASAM-MSE, which allows users to easily incorporate other candidate HCRs to meet specific management needs.

2.5.3 Catch Projection

During interim years between assessments in the feedback period, the EM projects the population forward under prevailing conditions, calculated as the average selectivity-at-age, maturity-at-age, weight-at-age, natural mortality, and (for spatially explicit EMs) movement parameters over a user-specified number of years preceding the most recent assessment. SPASMA-MSE supports a flexible range of catch projection options (see available options in Table 1). For EMs that incorporate process error, users can specify whether the process errors continue into the projection periods. The same is also true of any environmental covariate processes. Catch in each projection year is determined by the selected HCR. The way catch is projected varies by EM type (Table 2). In the PAN model, catch projections are fleet-specific, and by definition only has one region. In contrast, the FAA, SEP, SpD, and SpE models all perform fleet- and region-specific catch projections. This structure allows for spatially explicit projections that account for regional dynamics and heterogeneity in fleet operations. Environmental covariates, if present, can also be carried forward into projection years to inform catch advice.

2.5.4 Catch Apportionment Strategy

When the EM uses a coarser spatial resolution than the OM, catch advice from the EM must be apportioned to the finer resolution of the OM. Catch apportionment is a critical step, as the method used to allocate catch back to individual fleets and regions can significantly influence management outcomes. Using poorly aligned catch apportionment methods can misrepresent local dynamics, leading to unintended depletion in some areas and compromising management effectiveness (Bosley et al., 2019; Berger et al., 2020), whereas a well-designed strategy can help mitigate negative consequences due to a lack of spatial resolution and model misspecification in the EM. The package provides flexible options for apportioning catch to the appropriate spatial units in the OM, accommodating complex combinations of fleets, gear types, and regions. Users can choose from four major groups of weighting schemes (details in Table S2): (1) equal weights, which splits catch evenly across fleets or regions; (2) catch-based weights, which uses historical fleet-, gear-, or region-specific catch data; (3) survey-based weights, which allocates catch using one or multiple survey indices; and (4) user-specified weights. Users may also specify the number of years to use when calculating the catch apportionment weights.

2.5.5 Implementation error

Implementation error produces a realized catch that is applied in the OM that differs from the target, projected catch from the EM. Users have four options to define implementation error: 1) realized catch is the product of target catch and a lognormal random variable with user specified variance; 2) realized catch is the product of target catch and a normal random variable with user specified variance; 3) realized catch is the product of target catch and a multiplier that is drawn from a uniform distribution with a user specified range; and 4) realized catch is the product of target catch and a user specified constant multiplier (e.g., to produce something like systematic misreporting).

2.6 Feedback Loop and Iteration

During each iteration of the feedback period, annual catch quotas by fleet (produced from the EM based on the HCR and catch apportionment strategy) are passed to the OM, where the corresponding fleet-specific F are derived using a Newton algorithm. If management implementation error is included, the projected catch is first adjusted using a specified error model to simulate deviations between recommended and realized catch. The adjusted catch is then used to compute the fleet- and region-specific F values. Population dynamics are updated sequentially for each year in the interim period between assessments by running the OM forward with the derived F values, and new data are simulated accordingly. The EM is then refitted in the next iteration of the assessment using the extended time series of simulated data. This feedback process is run iteratively for the user specified number of years in the feedback period.

The user must specify the number of replicates to run for each MSE scenario, with each replicate using a unique random number seed to generate a distinct realization of stochastic processes. These seeds are stored to ensure consistency and reproducibility across scenarios, enabling direct comparisons among management strategies when needed. In each realization, a time series of unobserved process error (i.e., random effects) is applied to key biological and fishery processes (e.g., recruitment, survival, movement) based on the specified error distribution for each component. A corresponding set of observation errors is then applied on top of the resulting "true" population dynamics. Together, these sources of variability produce a unique trajectory of population dynamics and a corresponding pseudo dataset for each replicate. For each replicate, different management strategies can be applied to the same underlying realization, allowing direct comparison of performance across strategies. Because the process error and observation error remain fixed within a replicate, differences in outcomes can be attributed solely to the management strategy rather than stochastic variability.

The runtime of the MSE is heavily influenced by factors such as the number of regions, the length of the historical and feedback periods, the frequency of assessments during the feedback period, and the number of age classes. It is also affected by available computing power and memory. While in theory the number of regions that can be modeled is limited only by computational resources, in practice runtime constraints can be substantial.

Additional drivers of computational burden include the number and complexity of process errors in the EM. For example, modeling autocorrelation in process errors (e.g., AR(1) structures) increases the number of parameters to be estimated. This becomes particularly challenging for multidimensional process errors, such as random effects on numbers-at-age transitions, which can scale with both the number of populations and regions. Moreover, the use of self-weighting likelihoods for age composition data introduces extra distributional parameters that expand quickly as more fleets and surveys are included. These modeling choices can greatly increase runtime and may render some scenarios impractical to run on standard computing hardware.

2.7 Performance Metrics and Visualization

MSE results, including outputs from the OM and EM, are collected at the end of the feedback period. Users have the option to save each EM output generated during the feedback period. In addition, the package supports diagnostics for each EM, such as convergence, retrospective analysis, and residual analysis (e.g., one-step-ahead residuals). This functionality allows users to leverage the package as a platform for simulation-estimation experiments, allowing evaluation of model performance for EMs with different structural assumptions.

MSE results are summarized internally, and a range of performance metrics are automatically calculated based on flexible, user-specified options (Table S3). Performance can also be evaluated in relative terms by comparing metrics to a user-defined baseline management strategy, allowing users to explore trade-offs between the baseline and alternative management strategies. The user has the option to summarize results as a "global" median among years and realizations (i.e., median of n years × n realizations), a median among realizations of the means among years, or a median among realizations of the medians among years. These options were intended to allow users to evaluate whether the distribution of the results and subsequent method for calculating summary statistics influence relative performance of management procedures and conclusions.

The function supports output in multiple formats, including PDF, HTML, and PNG. The example in the following section provides a snapshot of the types of performance metrics and figures produced by the SPASAM-MSE package. For a complete list of available outputs and customization options, please see https://lichengxue.github.io/whamMSE/.

3. Example Application, Model Setup, and Results

The results presented in this example are intended solely for illustrative purposes. They are not meant to inform real-world management decisions but rather to demonstrate the capabilities of the SPASAM-MSE framework and to guide users in implementing MSE within the platform.

3.1 Background

In this section, we present an example application of the SPASAM-MSE framework to illustrate how MSE can address a central question in fisheries stock assessment: How does misalignment in spatial structure between the OM and EM affect management performance? Spatial misalignment can lead to biased parameter estimates and, in turn, suboptimal management advice. Using SPASAM-MSE, we evaluate trade-offs among estimation models with varying spatial structures to examine how these differences influence management outcomes.

3.2. OM Conditioning

The OM consisted of two distinct natal populations of a medium-lived fish species, each occupying a separate spatial region. Each region was associated with its own natal population, fleet, and survey. The species was modeled with 10 age classes, with life-history parameters (growth, weight-at-age, maturity-at-age) drawn from Wiedenmann et al. (2017), which provides generic values for a medium-lived groundfish species. A summary of the OM inputs and assumptions is provided in Table 3.

The OM included spatially explicit fleet dynamics and survey operations, with each fleet and survey acting independently within its designated region. Recruitment for each natal population was modeled as a stochastic process with annual deviations around a mean, with deviations following an independent and identically distributed (IID) log-normal distribution with a standard deviation of 0.8. Similarly, numbers-at-age transitions (i.e., survival) were modeled as log-normal random processes with a standard deviation of 0.3 (the same for each population and region). Natural mortality was fixed at 0.2 and was constant across all ages, regions, and populations. Seasonal movement followed a natal homing pattern: during spring and winter, fish could move between regions (North → South: 0.3; South → North: 0.1); in summer, individuals returned to their natal regions by the end of the season; and in autumn, the spawning season, no movement occurred.

The OM included a 30-year historical period to simulate baseline population dynamics prior to the implementation of management strategies. Fishing pressure was specified externally during this phase and applied separately to each fleet. For the first 15 years, a fully selected F of 0.2 was applied to each fleet in each region, reflecting a period of higher exploitation. In the following 15 years, F was reduced to 0.1, representing a management response or reduction in fishing pressure. Fleets in both regions operated year-round, and fishing effort was applied uniformly across seasons.

3.3. Data Observations and Stochasticity

Given that identical initial numbers-at-age and fishing pressure were assumed across populations in the OM, variability in population dynamics across realizations arises solely from stochastic processes. These are introduced through annual random effects, drawn independently for each realization (or replicate) using a unique random seed, based on the specified error distributions for recruitment and numbers-at-age transitions. For each realization, observational time series (i.e., fleet catch and survey indices, and age composition) are generated by applying annual observation errors drawn using the same random seed. In this example, we included 25 realizations, each consisting of a distinct set of true population dynamics and associated observational data used in the assessment during the feedback loop.

3.3 EM Structure

To evaluate how different assumptions about spatial structure influence management outcomes, we developed five EMs representing varying degrees of spatial complexity. These EMs serve as alternative management strategies within the MSE framework, each making different assumptions about population structure, fleet and survey treatment, and movement dynamics. The EMs range from a simple panmictic model to spatially explicit models with fixed or estimated movement. A summary of EM configurations and assumptions is provided in Table 4.

3.4 MSE Settings

In this example, we specified a 15-year feedback period, with assessments conducted every 3 years, resulting in two assessments over its duration. The HCR was defined as a constant F at 75% of $F_{40\%}$. Projections for each EM were conducted over 3-year periods (matching the assessment interval) and were based on prevailing biological conditions (5-year averages of life history traits). Recruitment in the projection period was defined as the average of the estimated recruitment time series from the EM, with stochastic recruitment deviations continuing throughout the projection period. In each assessment, the entire time series up to the most recent year of available data was used. A summary of MSE configurations is provided in Table 5.

3.5 Data Visualization and MSE Results

Runtime for each replicate (i.e., seed) under a given EM scenario is approximately 30 minutes, and up to 60 minutes for spatially explicit EMs. The MSE simulation supports parallel computing across replicates, allowing multiple seeds to be run simultaneously without loss of efficiency. The number of replicates that can be executed in parallel depends on the number of cores available on the user's machine or computing cluster. After the simulations are completed, the package provides a comprehensive suite of tools for evaluating and comparing management strategies, with outputs summarized across realizations. In this example, we present a selection of MSE outputs using a variety of data visualization approaches (Figures 6–8 and Figures S2–S10).

Based on our results, incorporating spatial structure into stock assessment models provided clear benefits for maintaining SSB (Figure 6 and Figure S3) and reducing the risk of overfishing at both regional (Figures S5&S6) and global scales (Figure 7). However, these benefits often came at the risk of reduced catch relative to the PAN (Figure 6 and Figure S3). Management outcomes improved even when spatial structure was implicitly accounted for, as in the FAA EM, which outperformed the PAN EM by better maintaining SSB while sustaining relatively high catch levels (Figure 6 and Figure S3). The SpD model appeared to offer an intermediate management strategy, balancing catch and SSB (Figure 6 and Figure S3) while avoiding overfishing at both regional (Figures S5&S6) and global scales (Figure 7), performing comparably to the SpE EMs.

4. Discussion

SPASAM-MSE offers a spatially explicit platform for conducting MSE, supporting improved understanding of spatial processes in stock assessments by leveraging advanced features of state-space models. It is the first fully generalized and modular framework for spatial MSEs that integrates biocomplexity across both OMs and EMs. Although previous MSE tools have provided limited support for spatial structure, SPASAM-MSE introduces an unprecedented level of spatial realism and flexibility.

The SPASAM-MSE is designed with user-friendly flexibility. It supports comprehensive MSE design and analysis, and is applicable across a broad spectrum of scenarios, from realistic, data-rich fisheries to theoretical simulations involving generic fish species with varying life history traits. This versatility makes SPASAM-MSE a powerful tool both for testing management strategies under real-world conditions and for investigating broader hypotheses related to spatial dynamics, model structure, and data limitations. Additionally, it could serve as a collaborative platform by fostering engagement among stakeholders when defining HCRs and evaluating tradeoffs. In short, SPASAM-MSE fills a critical gap in the global fisheries management toolkit by fully integrating spatial complexity into MSE workflows, thereby advancing robust, science-based, and sustainable fisheries management.

4.1. Potential Applications

The SPASAM-MSE package enables users to address a wide range of research and management questions relevant to spatial fisheries assessment and management. For example, it can be used to evaluate catch apportionment strategies and to assess whether a single panmictic assessment is sufficient for achieving management objectives within spatially structured populations or fisheries.

Users can also test the consequences of misalignment between the true spatial structure in the OM and the structure assumed in the EM or compare the performance of global versus spatially explicit biological reference points under various movement scenarios. The framework also allows exploration of whether random effects on key processes (e.g., numbers-at-age transitions; Li et al., *In revision*) can serve as proxies for unmodeled spatial dynamics, such as movement. Users can also leverage the framework as a simulation-estimation platform for evaluating the impacts of model misspecification.

SPASAM-MSE also enables the evaluation of sampling design and data availability. In particular, the data processing module supports scenarios where a portion of the spatial domain, such as a MPA, wind farm, or habitat refuge, is closed to survey or fishery access. Users can simulate these unfished areas and test how their presence affects assessment quality and management performance. Furthermore, the data module allows for the aggregation of new data streams from closed areas into adjacent regions of the EM, allowing exploration of whether incorporating unobserved areas improves model accuracy and decision-making.

One of the most valuable applications of SPASAM-MSE lies in its ability to guide decisions regarding the appropriate level of model, data, or harvest rule complexity. This challenge represents a classic bias–variance trade-off, where overly simplistic models may fail to capture critical system dynamics, while overly complex models may suffer from reduced stability, robustness, or interpretability. SPASAM-MSE offers a structured framework for comparing alternative model configurations within a spatial context, enabling users to evaluate trade-offs between realism and reliability. By systematically assessing the consequences of different modeling choices, the framework facilitates the identification of a practical "minimally complex, maximally robust" management procedure (Goethel et al., 2023b), representing a level of complexity that adequately captures key biological and management processes without compromising performance, ultimately supporting more effective and resilient spatial fisheries management.

Additional applications include identifying trade-offs among management strategies under varying types of spatial movement dynamics, including ontogenetic or environmentally-induced range shifts. SPASAM-MSE supports the use of random effects to model stochastic movement, capturing interannual variability or uncertainty in migration rates and directionality. This enables exploration of how movement uncertainty affects population connectivity, data availability, and the spatial allocation of harvest.

SPASAM-MSE supports the inclusion of environmental covariates in the OM, the EM, and in EM catch projections (for generating catch advice). In the OM, environmental covariates can drive spatial and temporal variability in key biological processes (e.g., recruitment, natural mortality, or movement), linking environmental conditions to “true” population dynamics. In the EM, covariates may either be excluded, which allows users to evaluate management strategies and outcomes when assessment models ignore environmental drivers, or incorporated as explanatory variables to test whether the assessment model can detect and account for environmentally driven variability. Environmental covariates can also be carried forward into catch projections, allowing projected catch advice to reflect expected environmental variability. This functionality enables users to evaluate how environmental heterogeneity influences estimation accuracy, model performance, and management outcomes, while also providing a framework to explore strategies under changing environmental conditions.

Other applications include testing the consequences of biased or imprecise data. For example, SPASAM-MSE could be used to evaluate the consequences of misreporting or other types of management implementation error (Perretti et al., 2020). Furthermore, evaluations related to data quality and quantity could be conducted to inform discussions around resource allocation or regional data collection priorities.

4.2 Limitations

Although SPASAM-MSE offers substantial flexibility and analytical capability, several important limitations and considerations remain. The number of built-in HCRs in SPASAM-MSE is currently limited. However, the framework is modular and user-friendly, allowing users to easily add custom HCRs within the HCR module to support region-specific policies or risk-based management strategies. Additionally, while management implementation error can be represented as a latent state within the OM (Perretti et al., 2020), it is not explicitly modeled within the estimation model, potentially limiting the realism of feedback simulations.

A notable limitation of both SPASAM-MSE and WHAM, is the current lack of support for tagging data, which are critical not only for estimating movement parameters but also for informing natural mortality rates (e.g., Goethel et al., 2019). Although SPASAM-MSE allows movement rates to be estimated using priors, integrating tagging data would provide a stronger empirical foundation for modeling spatial dynamics and validating movement assumptions. This limitation is particularly relevant given that, for many species, empirical knowledge of movement is sparse or entirely lacking. As a result, defining a "true" OM often involves hypothetical assumptions rather than evidence-based structures. This reflects a broader challenge in spatial stock assessment and management: the need for foundational data collection to support spatial model development, including empirical data on movement patterns, spatially resolved abundance indices, habitat-associated ecological processes, and stock of origin of harvested fish.

Another limitation is that SPASAM-MSE currently uses age as the fundamental partition for tracking population dynamics, which may restrict its applicability to fisheries or regions where length-based assessments are the standard. While a length-based extension has been developed within WHAM (Correa et al., 2023), it has not yet been implemented in SPASAM-MSE.

Finally, from a computational perspective, model run times can be intensive, particularly for scenarios with high spatial resolution, long feedback periods, highly migratory or long-lived species, or complex random effects. Although the package leverages Template Model Builder (TMB) for efficient optimization and supports parallel processing to reduce overall runtime, complex OM–EM configurations may still require substantial computational resources as spatial and temporal resolution increases. In practice, users may need to make pragmatic simplifications, for example, limiting the number of regions (e.g., ≤ 4), reducing age structure for long-lived species, or shortening feedback periods to ensure tractable runtimes. Even with these simplifications, useful inference can still be drawn from MSE scenarios that capture the essential dynamics while balancing realism with computational feasibility.

4.3 Future Work

Future developments of SPASAM-MSE will focus on expanding capabilities, improving accessibility, and broadening applicability across fisheries and user communities. Planned enhancements include support for tagging data, both in simulation and estimation, to improve validation and parameterization of spatial dynamics; incorporation of additional reference point structures (e.g., region-specific or multi-stock BRPs) to support more flexible and spatially nuanced HCRs; and continued development of the built-in species library. The library currently consists of generic life-history types, but will be expanded to include species-specific templates, enabling direct application of SPASAM-MSE to management-relevant case studies.

Recognizing that SPASAM-MSE currently lacks a graphical user interface (GUI), future updates will focus on improving usability, particularly for non-R users, through the development of R-Shiny applications and interactive dashboards. These tools will enhance visualization, facilitate interpretation of complex results, and support stakeholder engagement. In the longer term, SPASAM-MSE is intended to serve as a platform for collaborative management strategy development, with strong potential for use in stakeholder workshops and advisory processes to co-develop, test, and refine sustainable management strategies. The modular, open-source design of the code base also encourages users to contribute their own extensions and new features, consistent with the principles of open and transparent science. As the user community grows, we anticipate parallel growth in MSE capabilities, particularly through the addition of novel HCRs, BRPs, and movement dynamics.

Recognizing that SPASAM-MSE currently lacks a graphical user interface (GUI), future updates aim to improve usability, particularly for non-R users, through the development of R-Shiny applications and interactive dashboards. These tools will enhance visualization, facilitate interpretation of complex results, and support stakeholder engagement. In the long term, SPASAM-MSE is envisioned as a platform for collaborative management strategy development, with strong potential for use in stakeholder workshops and advisory processes to co-develop, test, and refine sustainable management strategies. Moreover, by adhering to the tenets of open and transparent science, the modular and open-source nature of the code base is meant to encourage users to develop their own code and add new features. Thus, we envision that as the user base grows, there will be a simultaneous increase in MSE capabilities, particularly in the form of new and unique HCRs, BRPs, and movement dynamics.

5. Acknowledgements

We thank the NOAA internal reviewers, the Associate Editor, and anonymous reviewers for their constructive feedback. Funding was provided by NOAA Fisheries through the Northeast, Northwest, and Alaska Fisheries Science Centers. We are also grateful to Grumpy's Pub in Falmouth, MA, for providing a convivial atmosphere that fostered stimulating discussions and ultimately improved this research. The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the authors and do not necessarily reflect those of NOAA, the U.S. Department of Commerce, or the authors' affiliated organizations.

6. Competing interests statement

Timothy J. Miller serves as a Guest Editor for the *Canadian Journal of Fisheries and Aquatic Sciences* special issue on State-Space Methods. Daniel R. Goethel serves as an Associate Editor for *Canadian Journal of Fisheries and Aquatic Sciences*. All other authors declare that they have no competing financial interests or personal relationships that could have influenced the work reported in this paper.

7. CRediT authorship contribution statement

- **Chengxue Li**: Conceptualization, Methodology, Software, Formal analysis, Visualization, Writing – original draft, Writing – review & editing.
- **Jonathan J. Deroba**: Supervision, Funding acquisition, Conceptualization, Methodology, Writing – review & editing.
- **Daniel R. Goethel**: Funding acquisition, Conceptualization, Methodology, Writing – review & editing.
- **Aaron M. Berger**: Funding acquisition, Conceptualization, Methodology, Writing – review & editing.
- **Amy M. Schueller**: Funding acquisition, Methodology, Writing – review & editing.
- **Brian J. Langseth**: Funding acquisition, Methodology, Writing – review & editing.
- **Emily Liljestrand**: Software, Formal analysis, Writing – review & editing.
- **Dana H. Hanselman**: Conceptualization, Writing – review & editing.
- **Timothy J. Miller**: Software, Writing – review & editing.

8. Funding statement

This work was funded by NOAA Fisheries through the Northeast Fisheries Science Center (NEFSC), Alaska Fisheries Science Center (AFSC), and Northwest Fisheries Science Center (NWFSC).

9. Data availability statement

The data underlying this article are available on GitHub at: https://github.com/lichengxue/SPASAM-MSE-Method-Paper

10. References

Babcock, E.A., and MacCall, A.D. 2011. How useful is the ratio of fish density outside versus inside no-take marine reserves as a metric for fishery management control rules? Can. J. Fish. Aquat. Sci. 68: 343–359. https://doi.org/10.1139/F10-141

Benson, A.J., Cox, S.P., and Cleary, J.S. 2015. Evaluating the conservation risks of aggregate harvest management in a spatially structured herring fishery. Fish. Res. 167: 101–113. https://doi.org/10.1016/j.fishres.2015.02.003

Berger, A.M., Barceló, C., Goethel, D.R., Hoyle, S.D., Lynch, P.D., McKenzie, J., Dunn, A., Punt, A.E., Methot, R.D., Hampton, J., and Porch, C.E. 2024. Synthesizing the spatial functionality of contemporary stock assessment software to identify future needs for next generation assessment platforms. Fish. Res. 275: 107008. https://doi.org/10.1016/j.fishres.2024.107008

Berger, A.M., Goethel, D.R., Lynch, P.D., Quinn, T., McKenzie, J., and Punt, A.E. 2021. Incoherent dimensionality: a framework for aligning population structure and stock assessment. ICES J. Mar. Sci. 77: 2195–2207. https://doi.org/10.1093/icesjms/fsaa100

Berger, A.M., Goethel, D.R., Lynch, P.D., Quinn, T., Mormede, S., McKenzie, J., and Dunn, A. 2017. Space oddity: the mission for spatial integration. Can. J. Fish. Aquat. Sci. 74: 1698–1716. https://doi.org/10.1139/cjfas-2016-0297

Bosley, K.M., Goethel, D.R., Berger, A.M., Deroba, J.J., Fenske, K.H., Hanselman, D.H., Langseth, B.J., and Schueller, A.M. 2019. Overcoming challenges of harvest quota allocation in spatially structured populations. Fish. Res. 220: 105344. https://doi.org/10.1016/j.fishres.2019.105344

Cadrin, S.X., Goethel, D.R., Berger, A., and Jardim, E. 2023. Best practices for defining spatial boundaries and spatial structure in stock assessment. Fish. Res. 262: 106650. https://doi.org/10.1016/j.fishres.2023.106650

Carruthers, T.R., Hordyk, A.R., and Price, S. 2018. The Data-Limited Methods Toolkit (DLMtool): an R package for informing management of data-limited populations. Methods Ecol. Evol. 9: 2383–2387. https://doi.org/10.1111/2041-210X.13081

Correa, G.M., Monnahan, C.C., Sullivan, J.Y., Thorson, J.T., and Punt, A.E. 2023. Modelling time-varying growth in state-space stock assessments. ICES J. Mar. Sci. 80: 2036–2049. https://doi.org/10.1093/icesjms/fsad133

Cox, S.P., and Kronlund, A.R. 2008. Practical stakeholder-driven harvest policies for groundfish fisheries in British Columbia, Canada. Fish. Res. 94: 224–237. https://doi.org/10.1016/j.fishres.2008.08.004

Cunningham, C.J., Anderson, C.M., Wang, J.Y.L., Link, M., and Hilborn, R. 2019. A management strategy evaluation of the commercial sockeye salmon fishery in Bristol Bay, Alaska. Can. J. Fish. Aquat. Sci. 76: 1669–1683. https://doi.org/10.1139/cjfas-2018-0395

Deroba, J.J., and Bence, J.R. 2008. A review of harvest policies: understanding relative performance of control rules. Fish. Res. 94: 210–223. https://doi.org/10.1016/j.fishres.2008.07.003

Doering, K.L., Vaughan, N.R., Walter, J.F., Methot, R.D., Sagarese, S.R., Smith, M., Farmer, N.A., Calay, S., Cummings, N.J., Johnson, K.F., and Marshall, K. 2023. SSMSE: an R package for Management Strategy Evaluation with Stock Synthesis operating models. J. Open Source Softw. 8: 4937. https://doi.org/10.21105/joss.04937

Goethel, D.R., Quinn, T.J., and Cadrin, S.X. 2011. Incorporating spatial structure in stock assessment: movement modeling in marine fish population dynamics. Rev. Fish. Sci. 19: 119–136. https://doi.org/10.1080/10641262.2011.557451

Goethel, D.R., Bosley, K.M., Hanselman, D.H., Berger, A.M., Deroba, J.J., Langseth, B.J., and Schueller, A.M. 2019. Exploring the utility of different tag-recovery experimental designs for use in spatially explicit, tag-integrated stock assessment models. Fish. Res. 219: 105320. https://doi.org/10.1016/j.fishres.2019.105320

Goethel, D.R., Bosley, K.M., Langseth, B.J., Deroba, J.J., Berger, A.M., Hanselman, D.H., and Schueller, A.M. 2021. Where do you think you're going? Accounting for ontogenetic and climate-induced movement in spatially stratified integrated population assessment models. Fish Fish. 22: 141–160. https://doi.org/10.1111/faf.12510

Goethel, D.R., Berger, A.M., and Cadrin, S.X. 2023a. Spatial awareness: good practices and pragmatic recommendations for developing spatially structured stock assessments. Fish. Res. 264: 106703. https://doi.org/10.1016/j.fishres.2023.106703

Goethel, D.R., Omori, K.L., Punt, A.E., Lynch, P.D., Berger, A.M., de Moor, C.L., Plagányi, É.E., Cope, J.M., Dowling, N.A., McGarvey, R., and Preece, A.L. 2023b. Oceans of plenty? Challenges, advancements, and future directions for the provision of evidence-based fisheries management advice. Rev. Fish Biol. Fish. 33: 375–410. https://doi.org/10.1007/s11160-022-09726-7

Goethel, D.R., Berger, A.M., Hoyle, S.D., et al. 2024. Drivin' with your eyes closed: results from an international, blinded simulation experiment to evaluate spatial stock assessments. Fish Fish. 25: faf12819. https://doi.org/10.1111/faf.12819

Grüss, A., Harford, W.J., Schirripa, M.J., Velez, L., Sagarese, S.R., Shin, Y.J., and Verley, P. 2016. Management strategy evaluation using the individual-based, multispecies modeling approach OSMOSE. Ecol. Model. 340: 86–105. https://doi.org/10.1016/j.ecolmodel.2016.09.019

Jacobsen, N.S., Marshall, K.N., Berger, A.M., Grandin, C.J., and Taylor, I.G. 2021. Management strategy evaluation of Pacific hake: exploring the robustness of the current harvest policy to spatial stock structure, shifts in fishery selectivity, and climate-driven distribution shifts. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-NWFSC-168. Available from https://repository.library.noaa.gov/view/noaa/30919

Jacobsen, N.S., Marshall, K.N., Berger, A.M., Grandin, C., and Taylor, I.G. 2022. Climate-mediated stock redistribution causes increased risk and challenges for fisheries management. ICES J. Mar. Sci. 79: 1120–1132. https://doi.org/10.1093/icesjms/fsac029

Kapur, M.S., Haltuch, M.A., Connors, B., Berger, A., Holt, K., Marshall, K., and Punt, A.E. 2024. Range-wide contrast in management outcomes for transboundary Northeast Pacific sablefish. Can. J. Fish. Aquat. Sci. 81: 810–827. https://doi.org/10.1139/cjfas-2024-0008

Kerr, L.A., Hintzen, N.T., Cadrin, S.X., Clausen, L.W., Dickey-Collas, M., Goethel, D.R., Hatfield, E.M.C., et al. 2017. Lessons learned from practical approaches to reconcile mismatches between biological population structure and stock units of marine fish. ICES J. Mar. Sci. 74: 1708–1722. https://doi.org/10.1093/icesjms/fsx059

Kvamsdal, S.F., Eide, A., Ekerhovd, N.A., Enberg, K., Gudmundsdottir, A., Hoel, A.H., Mills, K.E., Mueter, F.J., Ravn-Jonsen, L., Sandal, L.K., and Stiansen, J.E. 2016. Harvest control rules in modern fisheries management. Elem. Sci. Anth. 4: 000114. https://doi.org/10.12952/journal.elementa.000114

Hilborn, R. 2012. The evolution of quantitative marine fisheries management 1985–2010. Nat. Resour. Model. 25: 122–144. https://doi.org/10.1111/j.1939-7445.2011.00105.x

Li, Y., Bence, J.R., and Brenden, T.O. 2018. Can spawning origin information of catch or a recruitment penalty improve assessment and fishery management performance for a spatially structured stock assessment model? Can. J. Fish. Aquat. Sci. 75: 2136–2148. https://doi.org/10.1139/cjfas-2017-0523

Li, C., et al. In review. Random effects on numbers-at-age transitions implicitly account for movement dynamics and improve stock assessment and management. Can. J. Fish. Aquat. Sci. [In review].

Miller, T.J., Curti, K.L., and Hansell, A.C. 2025. Space for WHAM: a multi-region, multi-stock generalization of the Woods Hole Assessment Model with an application to black sea bass. Can. J. Fish. Aquat. Sci. In press. https://doi.org/10.1139/cjfas-2025-0097

Perretti, C.T., Deroba, J.J., and Legault, C.M. 2020. Simulation testing methods for estimating misreported catch in a state-space stock assessment model. ICES J. Mar. Sci. 77: 911–920. https://doi.org/10.1093/icesjms/fsz249

Smedbol, R.K., and Stephenson, R. 2001. The importance of managing within-species diversity in cod and herring fisheries of the north-western Atlantic. J. Fish Biol. 59: 109–128. https://doi.org/10.1111/j.1095-8649.2001.tb01381.x

Smith, J.A., Tommasi, D., Welch, H., Hazen, E.L., Sweeney, J., Brodie, S., Muhling, B., Stohs, S.M., and Jacox, M.G. 2021. Comparing dynamic and static time-area closures for bycatch mitigation: a management strategy evaluation of a swordfish fishery. Front. Mar. Sci. 8: 630607. https://doi.org/10.3389/fmars.2021.630607

Stock, B.C., and Miller, T.J. 2021. The Woods Hole Assessment Model (WHAM): a general state-space assessment framework that incorporates time- and age-varying processes via random effects and links to environmental covariates. Fish. Res. 240: 105967. https://doi.org/10.1016/j.fishres.2021.105967

Stock, B.C., Xu, H., Miller, T.J., Thorson, J.T., and Nye, J.A. 2021. Implementing two-dimensional autocorrelation in either survival or natural mortality improves a state-space assessment model for Southern New England–Mid Atlantic yellowtail flounder. Fish. Res. 237: 105873. https://doi.org/10.1016/j.fishres.2021.105873

Punt, A.E., Dunn, A., Elvarsson, B.P., Hampton, J., Hoyle, S.D., Maunder, M.N., Methot, R.D., and Nielsen, A. 2020. Essential features of the next-generation stock assessment package: a perspective. Fish. Res. 229: 105617. https://doi.org/10.1016/j.fishres.2020.105617

Punt, A.E. 2017. Strategic management decision-making in a complex world: quantifying, understanding, and using trade-offs. ICES J. Mar. Sci. 74: 499–510. https://doi.org/10.1093/icesjms/fsv193

Punt, A.E., and Smith, A.D.M. 1999. Harvest strategy evaluation for the eastern stock of gemfish (Rexea solandri). ICES J. Mar. Sci. 56: 860–875. https://doi.org/10.1006/jmsc.1999.0538

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A., and Haddon, M. 2016. Management strategy evaluation: best practices. Fish Fish. 17: 303–334. https://doi.org/10.1111/faf.12104

Wiedenmann, J., Wilberg, M., Sylvia, A., and Miller, T. 2017. An evaluation of acceptable biological catch (ABC) harvest control rules designed to limit overfishing. Can. J. Fish. Aquat. Sci. 74: 1028–1040. https://doi.org/10.1139/cjfas-2016-0381

Ying, Y., Chen, Y., Lin, L., and Gao, T. 2011. Risks of ignoring fish population spatial structure in fisheries management. Can. J. Fish. Aquat. Sci. 68: 2101–2120. https://doi.org/10.1139/f2011-118

Zhang, Y., Goethel, D.R., Smith, M.W., Picariello, L., Chen, Y., Li, Y., and Agar, J. 2024. A web-based decision support tool for performing management strategy evaluation on the Gulf of Mexico red snapper resource. Ecol. Inform. 82: 102756. https://doi.org/10.1016/j.ecoinf.2024.102756

11. Tables

Table 1. Summary of the major functionalities of the SPSAM-MSE framework.

Functionality	Description
Operating Model	Model that governs “true” population, fleets, and surveys dynamics.
Temporal Structure	Supports annual or multi-season time blocks
Spatial Structure	Supports multiple spatial regions within the model domain.
Population Structure	Supports panmictic population, panmictic population with spatial heterogeneity, natal-homing, and metapopulation configurations.
Demographics	Supports flexible life-history traits (growth, maturity, weight, natural mortality, lifespan)
Movement Dynamics	Supports flexible movement rates by population, region, age, and season, with optional directional trends.
Fleet Dynamics	Supports spatially structured fishing pressure across regions and fleets.
Random Effects	Supports random effects for recruitment, survival, selectivity, survey catchability, and natural mortality, and movement.
Environmental Covariates	Supports linking environmental drivers to processes.
Observation Model	Module responsible for generating pseudo-observational data.
Data Type	Supports annual total catch or survey indices (biomass or abundance) and age compositional data.
Observation Error	Supports user-specified coefficient of variation for aggregate data and effective sample size for age composition data.
Age Compositional Likelihoods	Supports likelihoods including count-based options (e.g., multinomial, Dirichlet–multinomial) and proportion-based options (e.g., Dirichlet, logistic-normal).
Data Collection and Processing	Module responsible for aggregating, filtering, and processing observational data for input to the assessment model.
Data Time Series	Defines the length of the observation time series used in the assessment.
Data Aggregation	Supports aggregation (across regions) of catch and survey data (including composition data) for the assessment.
Survey Inclusion/Exclusion	Supports inclusion or exclusion of specific survey data from the assessment model.
Region Reduction	Supports inclusion or exclusion of regions from the assessment model.
Estimation Model	Model that estimates population status based on fitting observational data.
Assessment Model Structure	Supports the full range of spatial and population structures (PAN, FAA, SEP, SpD, SpE), including corresponding movement configurations.
Observation Error	Supports user-specified coefficient of variation for aggregate data and effective sample size for age composition data.
Age Compositional Likelihoods	Supports likelihoods including count-based options (e.g., multinomial, Dirichlet–multinomial) and proportion-based options (e.g., Dirichlet, logistic-normal).
Random Effects	Supports random effects for recruitment, survival, selectivity, survey catchability, natural mortality, and movement.
Environmental Covariates	Supports linking environmental drivers to processes.
MSE Feedback Loop	Module that governs the closed-loop simulation process, including assessment frequency, quota setting, population projections, and evaluation of management strategy performance.
Feedback Period	Defines the number of years for the feedback period during which the full management procedure is implemented and simulated.
Assessment Frequency	Defines how often assessments occur in the feedback period.

Functionality	Description
Biological Reference Points	Supports flexible weighting schemes for calculating biological reference points.
Harvest Control Rules	Supports different harvest control rules, including constant F or catch, threshold or sloped rules, F at X% SPR or X% F_{MSY}
Catch Projections	Supports flexible projection options: terminal or average F over n years; F at X% SPR or X% F_{MSY}; user-defined F or total/fleet-specific catch; continuation/specification of environmental covariates and random effects; and user-specified biological parameters (e.g., weight-at-age, maturity-at-age, recruitment).
Catch Apportionment	Distributes global catch advice (e.g., from a panmictic EM) to spatial units using a given allocation strategy.
Management Implementation Error	Supports modeling management implementation error, where actual catch advice may deviate from projections, and offers flexible selection of error distributions (e.g., lognormal, normal, uniform).
Result Collection & Visualization	Supports summary and export of results into figures and tables.
Model Diagnostics	Supports convergence checks, residual analysis, retrospective analysis, and one-step-ahead residuals for each assessment model.
Performance Evaluation	Built-in tools to compute, compare, and summarize performance metrics across management strategies.
Visualization & Reporting	Generates visualizations and exports comprehensive MSE results.

Table 2. Summary of the estimation model (EM) structures available in the SPASAM-MSE framework.

Feature	**Panmictic (PAN)**	**Fleets-as-Areas (FAA)**	**Separate Panmictic (SEP)**	**Spatially Disaggregated (SpD)**	**Spatially Explicit (SpE)**
Spatial Structure	Single region	Single region	Multiple regions	Multiple regions	Multiple interacting regions
Population Structure	Single panmictic population	Single panmictic population	Region-specific panmictic populations	Region-specific panmictic populations	Spatial heterogeneity, metapopulation, or natal-homing structures
Demographic Variation	Global life-history traits	Global life-history traits	Regional life-history traits	Regional life-history traits	Regional or population-level life-history traits
Stock-recruit Relationship (SRR)	Global SRR	Global SRR	Local SRR per region	Local SRR per region	Global or local SRR per region
Movement/Connectivity	None	None	None	None	Movement fixed or estimated
Fleet Structure	Single fleet and survey across all regions (Or multiple fleets/surveys across all regions due to different types)	Single fleet and survey per region (Or multiple fleets/surveys per regions due to different types)	Single fleet and survey per region (Or multiple fleets/surveys per regions due to different types)	Single fleet and survey per region (Or multiple fleets/surveys per regions due to different types)	Single fleet and survey per region (Or multiple fleets/surveys per regions due to different types)
Observation Data Aggregation	Aggregated data across regions	Disaggregated data per region.	Disaggregated data per region	Disaggregated data per region	Disaggregated data per region
Biological Data Aggregation	Aggregated data across regions	Aggregated data across regions	Disaggregated data per region	Disaggregated data per region	Disaggregated data per region
Likelihood	Single likelihood	Single likelihood	Independent likelihood per region	Single joint likelihood aggregated across regions	Single joint likelihood aggregated across regions
Model Parameterization	Single global parameter set	Shared biology; fleet/survey-specific selectivity	Separate parameter sets per region, estimated independently	Separate parameter sets per region, estimated jointly	Region- & population-level parameters plus movement parameters
Biological Reference Point (BRP)	Global BRP; movement not included	Global BRP; movement not included	Regional BRP; movement not included	Global BRP (aggregated from regional BRPs via user-specified weighting); movement not included	Global BRP (aggregated from regional BRPs via user-specified weighting); movement included
Catch Projection	Global catch projections by fleet (including fleet-level breakdowns if multiple fleets are present).	Global catch projections (including fleet-level breakdowns if multiple fleets are present in a region)	Region-specific catch projections (including fleet-level breakdowns if multiple fleets are present in a region)	Region-specific catch projections (including fleet-level breakdowns if multiple fleets are present in a region)	Region-specific catch projections (including fleet-level breakdowns if multiple fleets are present in a region)

Table 3. Summary of the operating model (OM) configuration used for the SPASAM-MSE example application.

Component	Description
Operating Model (OM)	
Regions & Populations	Two regions (North and South), each with its own natal population, fleet, and survey.
Seasonal Structure	The year is divided into 4 seasons of equal length (0.25 each): spring, summer, autumn, and winter.
Number of Age Classes	10
Life-history Parameters	Life-history parameters were derived for a medium-lived species from Wiedenmann et al. (2017).
Spawning Timing	Spawning occurs at year fraction 0.625.
Historical Period	30-year historical period
Fishing Mortality	Fleet-specific $F = 0.2$ for first 15 years, then reduced to $F = 0.1$ for next 15 years
Fleet Operations	Fleets in both regions operate year-round.
Survey Timing	Surveys in both regions at 0.625 of the year.
Coefficient of Variation (CV) for Fleet and Survey	Fleet catch CV = 0.2; Survey index CV = 0.1
Age Composition Likelihood for Fleets and Surveys	Multinomial
Effective Sample Size (ESS) for Fleet and Survey Age Composition	ESS = 100
Survey Catchability	Catchability = 0.2 for both surveys
Movement Dynamics	Seasonal movement: Spring/Winter—movement allowed (North→South: 0.3, South→North: 0.1); Summer—natal homing; Autumn (spawning season)—no movement.
Movement Assumptions	Movement rates during spring and winter are region-specific and assumed to be constant across ages and years.
Selectivity	Logistic selectivity curves are used. Parameters are the same for both fleets and the same for both surveys: fleets use ($a_{50} = 3$, slope = 1); surveys use ($a_{50} = 2$, slope = 1). Selectivity is constant over time.
Natural Mortality	Set to 0.2; assumed constant across time, age, regions, and populations.
Recruitment Model	No stock–recruit relationship assumed. Recruitment for each population is modeled as random deviations around a mean of e^{10}, with independently and identically distributed (IID) with a standard deviation (SD) of 0.8.
Numbers-at-age (NAA) Transitions	Modeled as independently and identically distributed (IID) random effects with a standard deviation (SD) of 0.3
Initial Numbers-at-age	Equilibrium-based; with year 1 recruitment = e^{10}; fully-selected $F = 0.1$.

Table 4. Summary of estimation model (EM) configurations used for the SPASAM-MSE example application.

Component	**Description**
Estimation Model (EM)	
Panmictic (PAN)	Single-region, single-population model. Recruitment and NAA random effects were estimated for the entire population. Fleet and survey data were aggregated across regions. Catch CV = 0.2 and ESS = 100 for fleet age composition; survey CV = 0.1 and ESS = 100 for survey age composition.
Fleets-as-Areas (FAA)	Single-region, single-population model with two fleets and two surveys retained separately (not aggregated), resulting in two sets of fleet and survey parameters. Recruitment and NAA random effects were estimated for the entire population. While the model assumes no explicit spatial structure in population dynamics, the separate treatment of fleets and surveys implicitly accounts for spatial data structure. Fleet data: CV = 0.2, ESS = 100; Survey data: CV = 0.1, ESS = 100.
Separate Panmictic (SEP)	Multi-region model with independent panmictic populations in each region. Recruitment and NAA random effects were estimated separately for each region, with no movement among regions. Fleets and surveys were retained without aggregation, resulting in separate parameterizations and independent likelihoods for each region. Fleet data: CV = 0.2, ESS = 100; Survey data: CV = 0.1, ESS = 100.
Spatially Disaggregated (SpD)	Two-region, two-population model. Recruitment and NAA random effects were modeled separately for each population. Fleets and surveys were retained without aggregation, yielding two sets of fleet and survey parameters. No movement was estimated, but the structure explicitly represented spatial separation within a joint likelihood framework. BRPs were calculated at the global level, aggregated across regions. Fleet data: CV = 0.2, ESS = 100; Survey data: CV = 0.1, ESS = 100.
Spatially Explicit (SpE)	Two-region, two-population model. Recruitment and NAA random effects were estimated separately for each population. Fleets and surveys were retained without aggregation, resulting in two sets of fleet and survey parameters. Movement patterns and rates were assumed to be known and fixed (i.e., not estimated) and matched the true OM configuration. Fleet data: CV = 0.2, ESS = 100; Survey data: CV = 0.1, ESS = 100.

Table 5. Summary of management strategy evaluation (MSE) configurations used for the SPASAM-MSE example application.

Component	Description
Closed-loop Feedback	
Number of Realizations	25
Feedback Period	15 years
Assessment Interval	Every 3 years
Number of Assessments in Total	5 (within the 15-year feedback period)
A Fraction of $F_{40\%}$ Used as the Harvest Control Rule (HCR)	75% of $F_{40\%}$
Projection Settings (EM)	Recruitment random effects are assumed to continue. For the PAN EM, projected total catch is apportioned equally across regions. For all other EMs, projected catch is already region-specific.

12. Figures

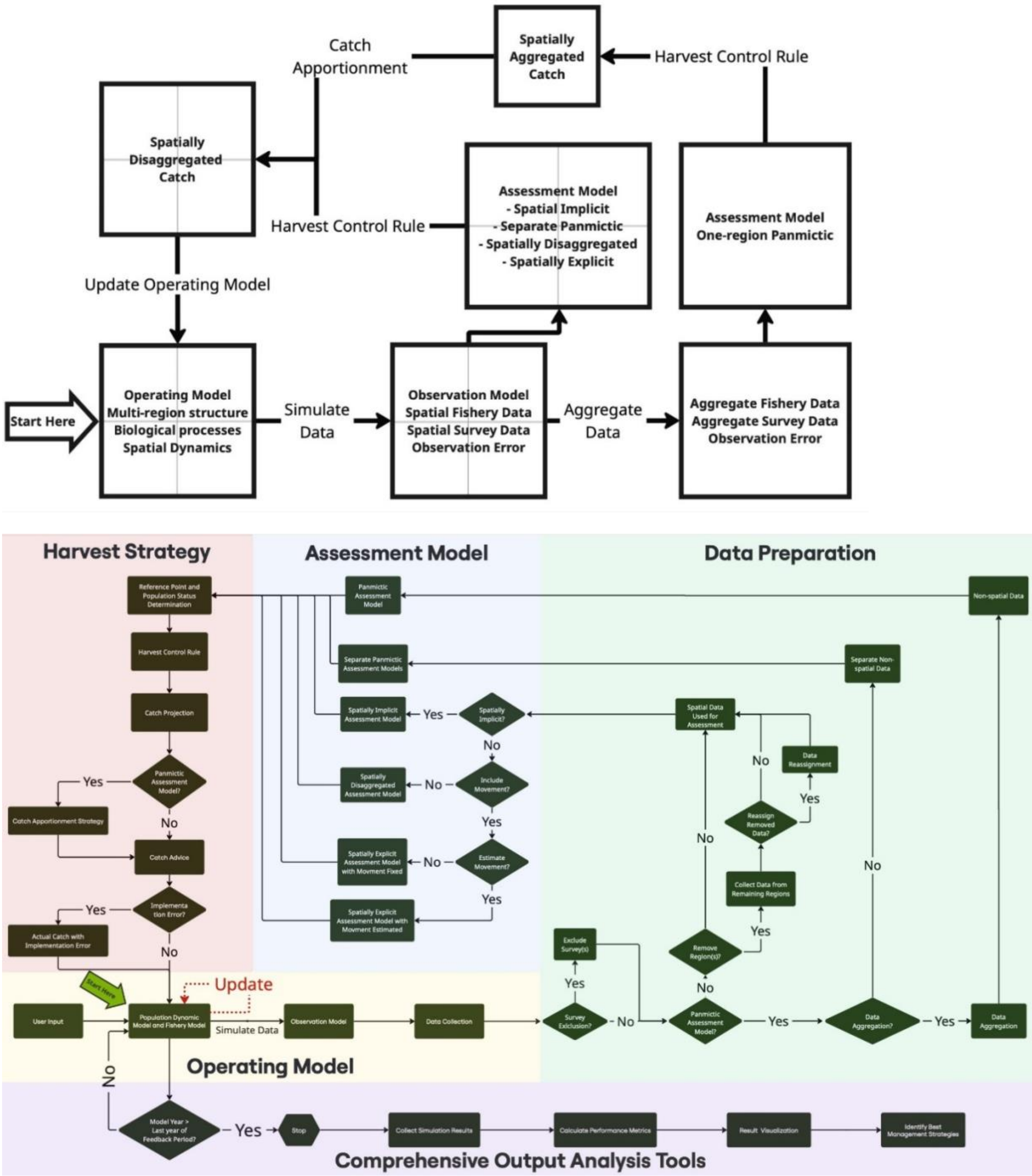


Figure 1. (A) High-level overview of the SPASAM-MSE framework, highlighting its spatial capabilities; (B) Detailed flowchart of the SPASAM-MSE framework, illustrating the structure and functionality of each component of the MSE process, including operating model conditioning, data generation and preparation, assessment model development, and harvest strategy implementation within the feedback loop.

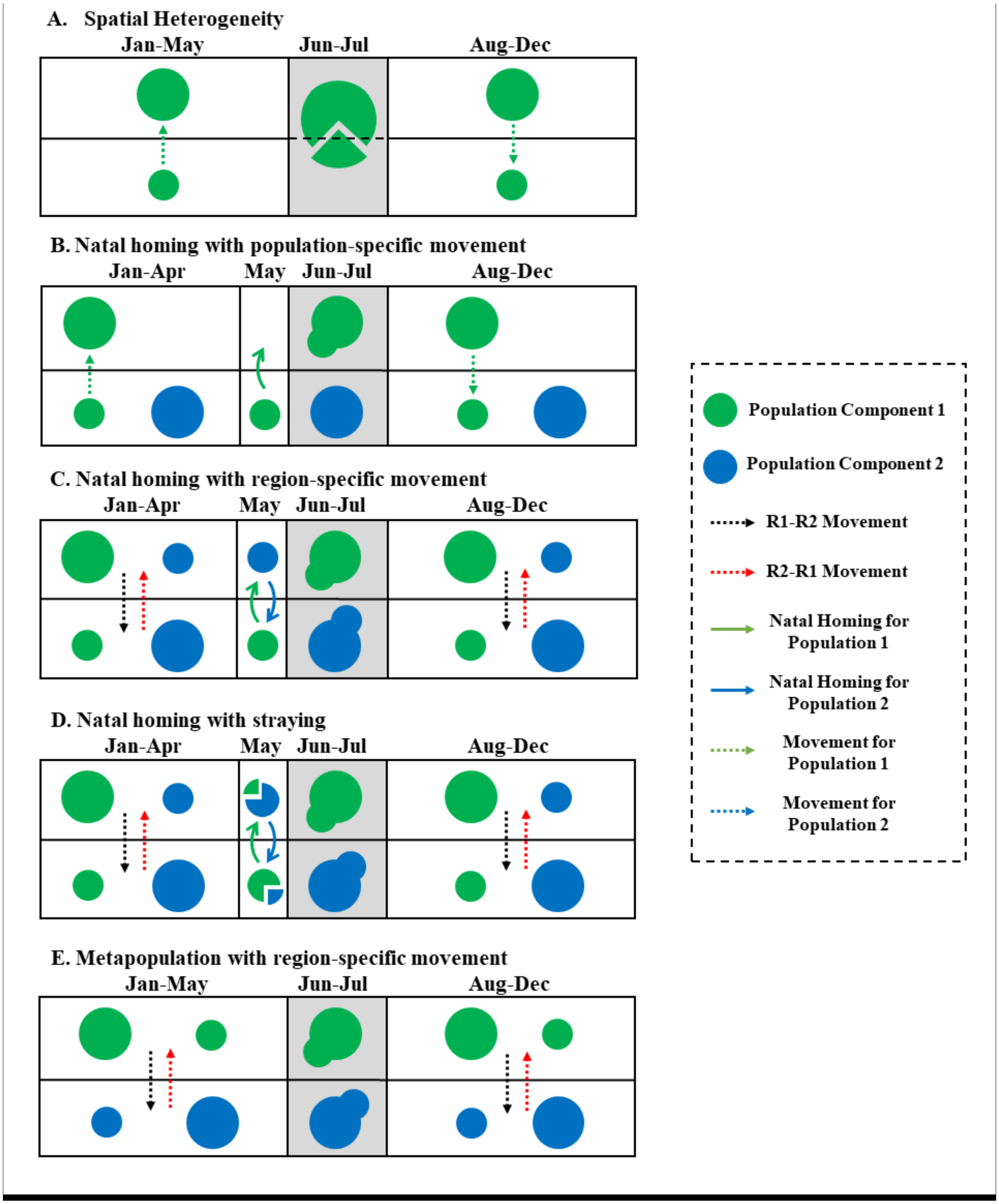


Figure 2. Examples of population structure and movement scenarios available in SPASAM-MSE: (A) spatial heterogeneity with seasonal movement; (B) natal homing with population-specific movement; (C) natal homing with region-specific movement; (D) natal homing with straying; and (E) metapopulation with region-specific movement. Spawning is assumed to occur in June–July (grey background). These examples represent common configurations but do not capture the full spectrum of spatial and movement scenarios supported by SPASAM-MSE.

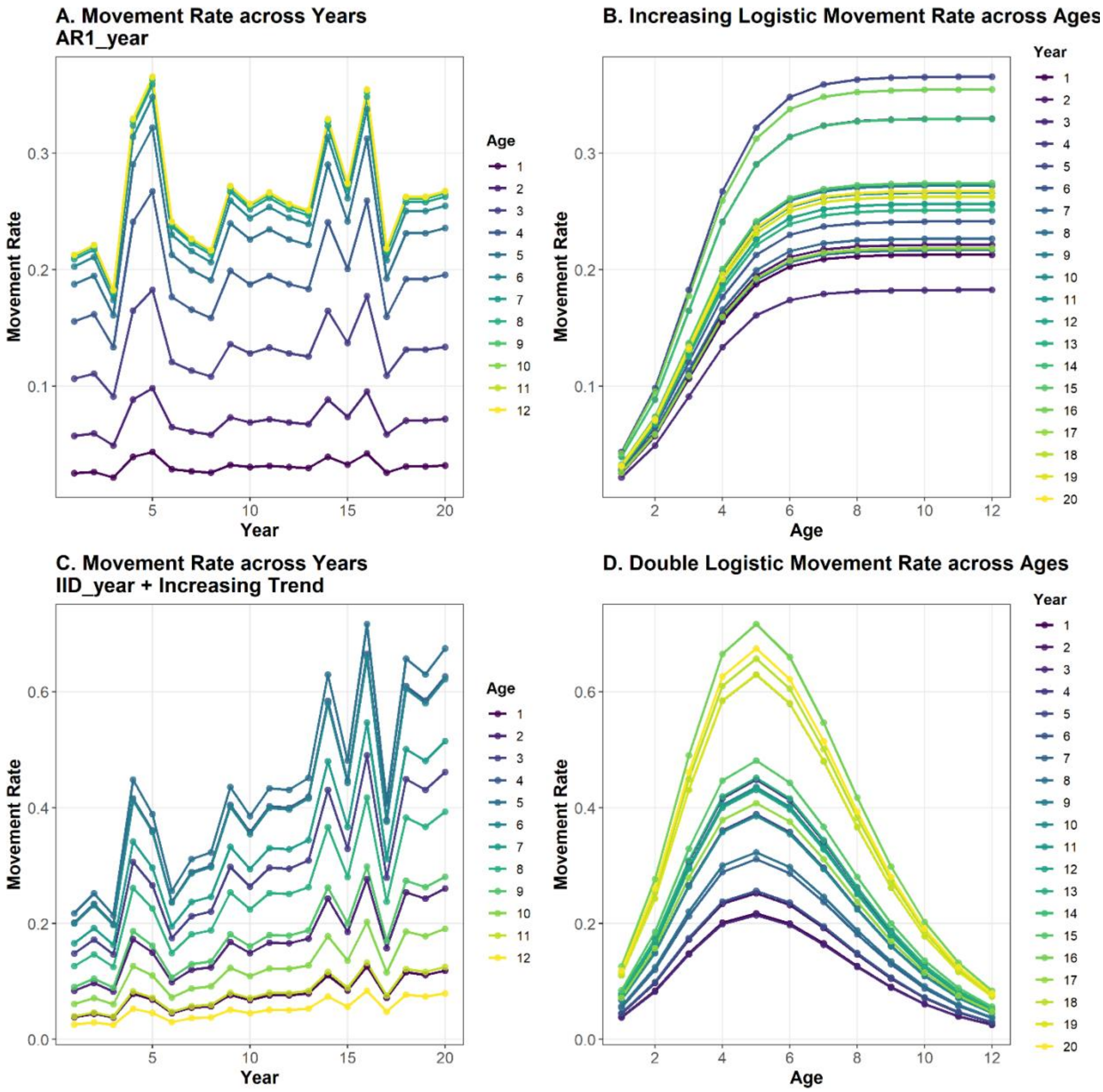


Figure 3. (A-B) Ontogenetic movement modeled as an increasing logistic curve with random effects following a stationary AR(1) process across years, representing stochastic variability; (C-D) Ontogenetic movement modeled as dome-shaped double-logistic curve with random effects following an independent and identically distributed (IID by years) process and a yearly increasing trend in the mean movement rate, representing nonstationary stochastic movement variability.

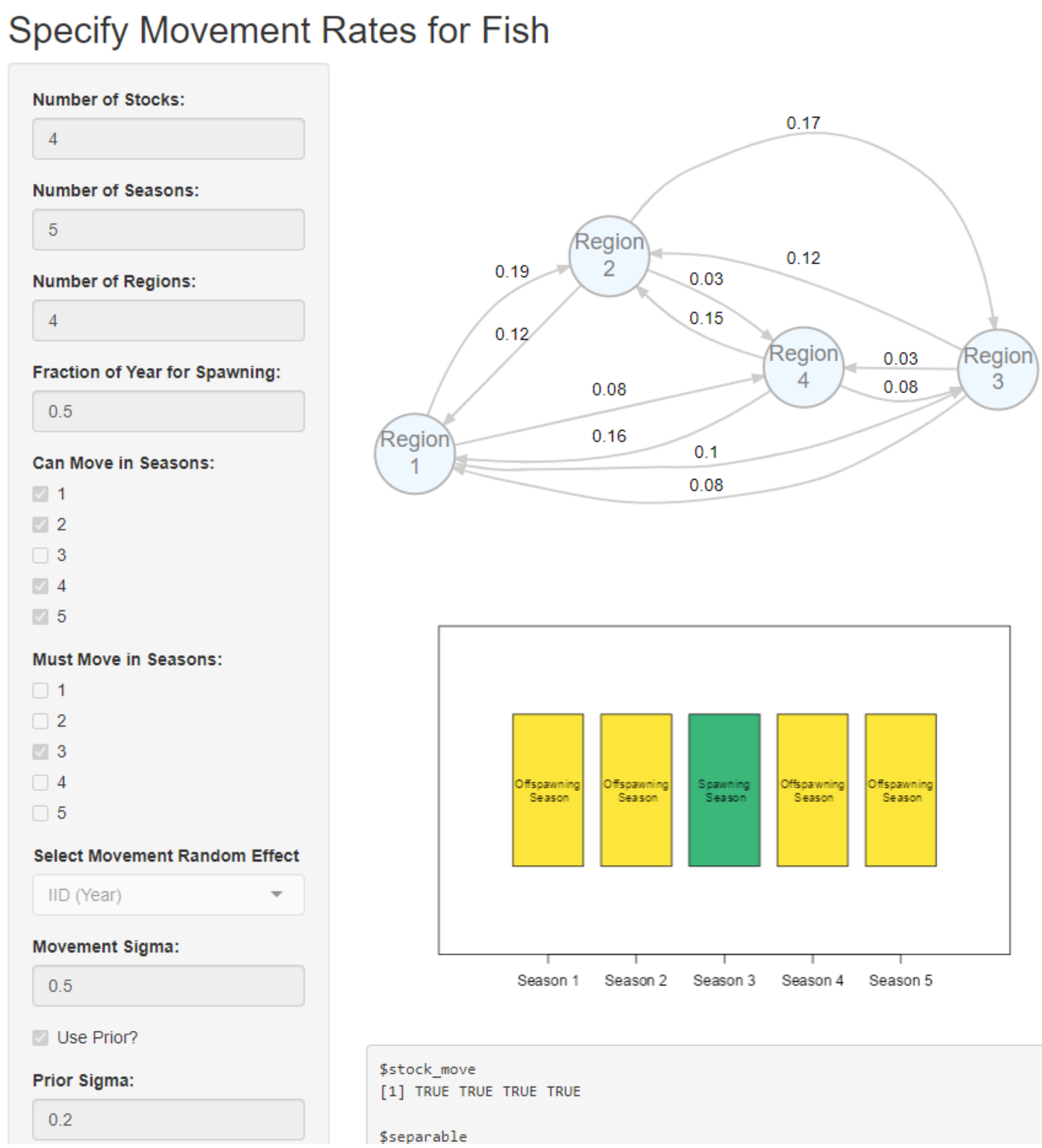


Figure 4. The R Shiny application built into SPASAM-MSE that enables user-friendly configuration of complex movement dynamics. Users can specify inputs such as the number of populations and regions, seasonal structure, movement patterns, mean movement rates, and random effect options, with outputs updated interactively.

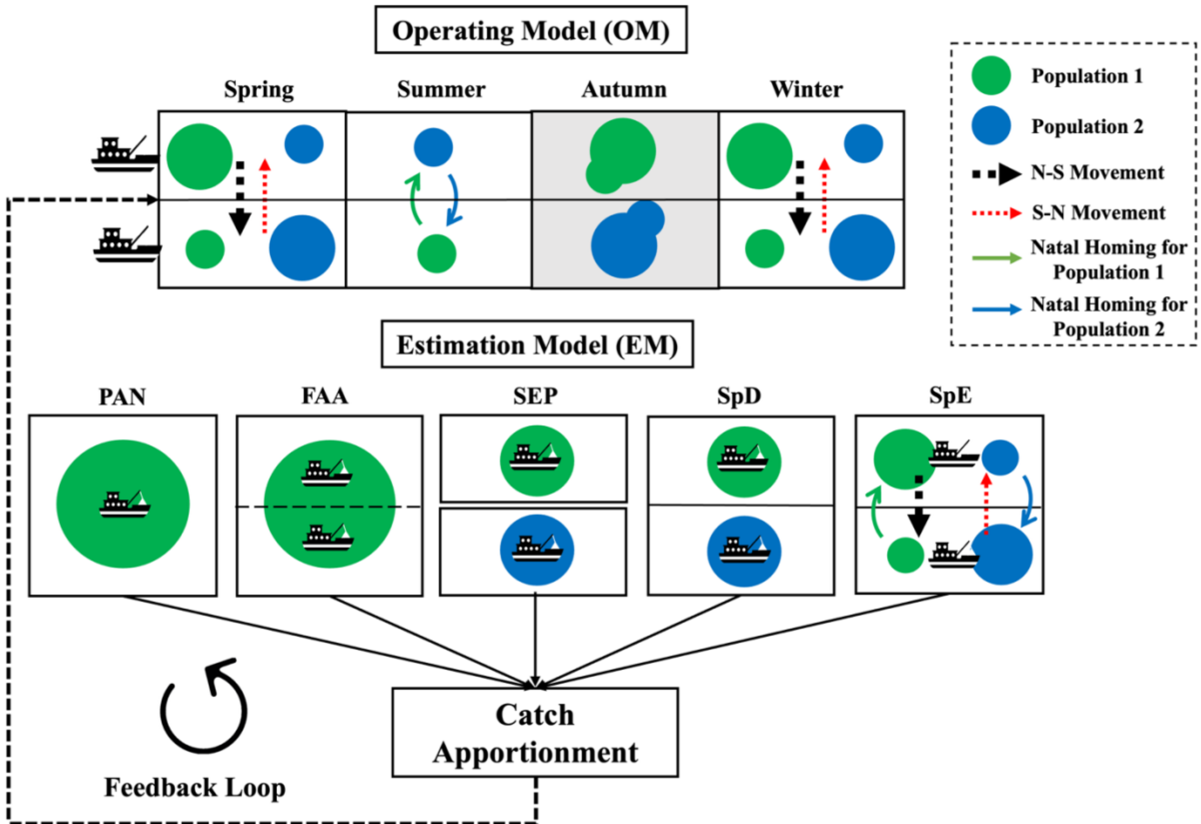


Figure 5. Schematic of the SPASAM-MSE framework as used in the example application. The operating model (OM) is spatially explicit with natal homing movement dynamics and includes two fleets and two surveys in two regions. Spawning is assumed to occur in Autumn (grey background). Five candidate estimation models (EMs) are evaluated: Panmictic (PAN), Fleets-as-Areas (FAA), Spatially Disaggregated (SpD), Spatially Explicit with Fixed Movement (SpE-F), and Spatially Explicit with Movement Estimated (SpE-E), each differing in spatial structure and movement assumptions. Projected catch from each candidate EM is passed to the catch apportionment module, which generates fleet-specific catch advice that feeds back to the OM to update population dynamics.

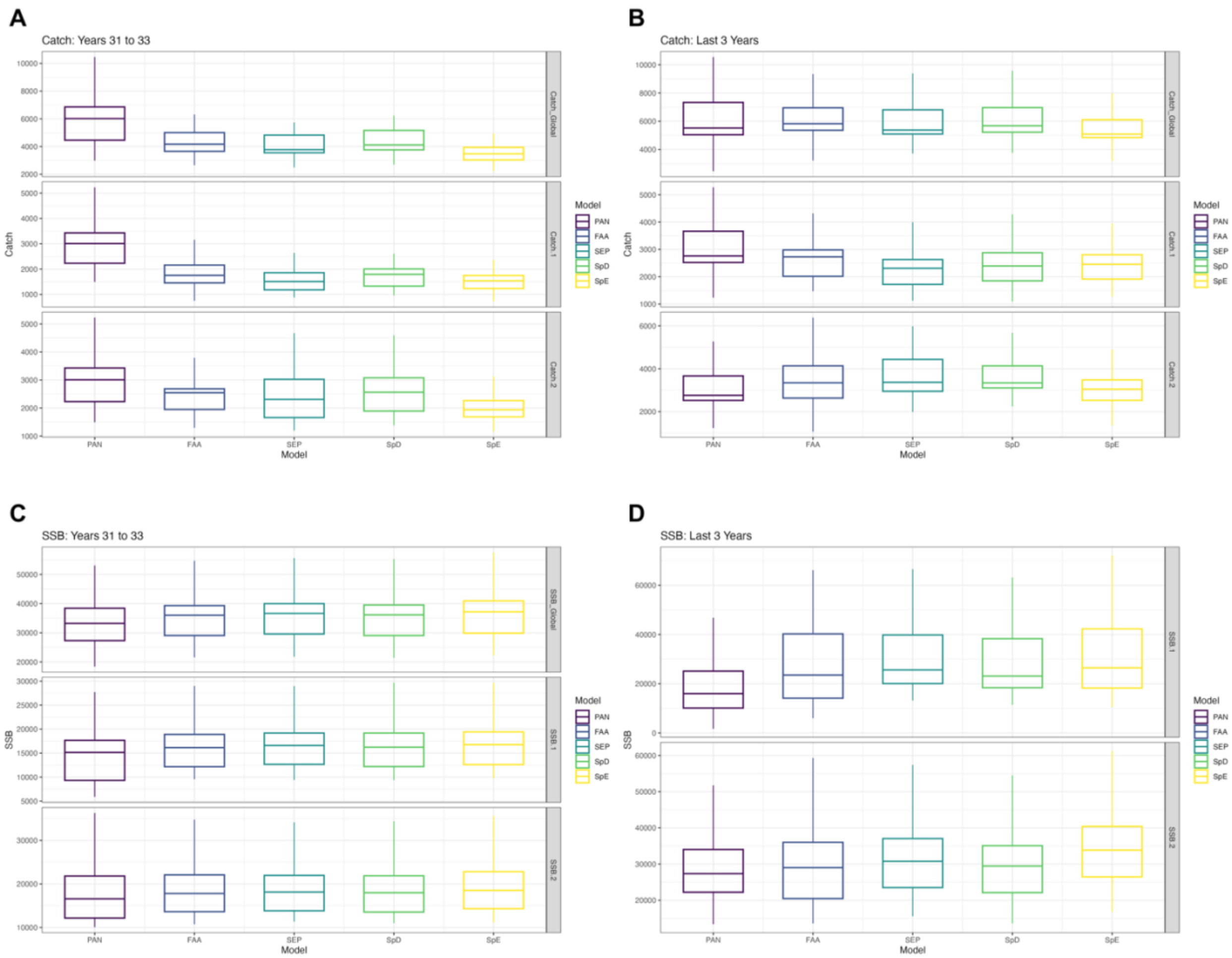


Figure 6. Performance of estimation models (EMs) in terms of catch and spawning stock biomass (SSB), summarized across 25 realizations. For each realization, catch and SSB were calculated as the mean over the evaluation window (short term = first 3 years of the feedback loop; long term = last 3 years). Panels show short-term average catch (A), short-term average SSB (B), long-term average catch (C), and long-term average SSB (D). Boxplots summarize the distribution of realization-level means across replicates: horizontal lines indicate medians, boxes represent the interquartile range (25th–75th percentiles), and whiskers extend to 1.5× the interquartile range.

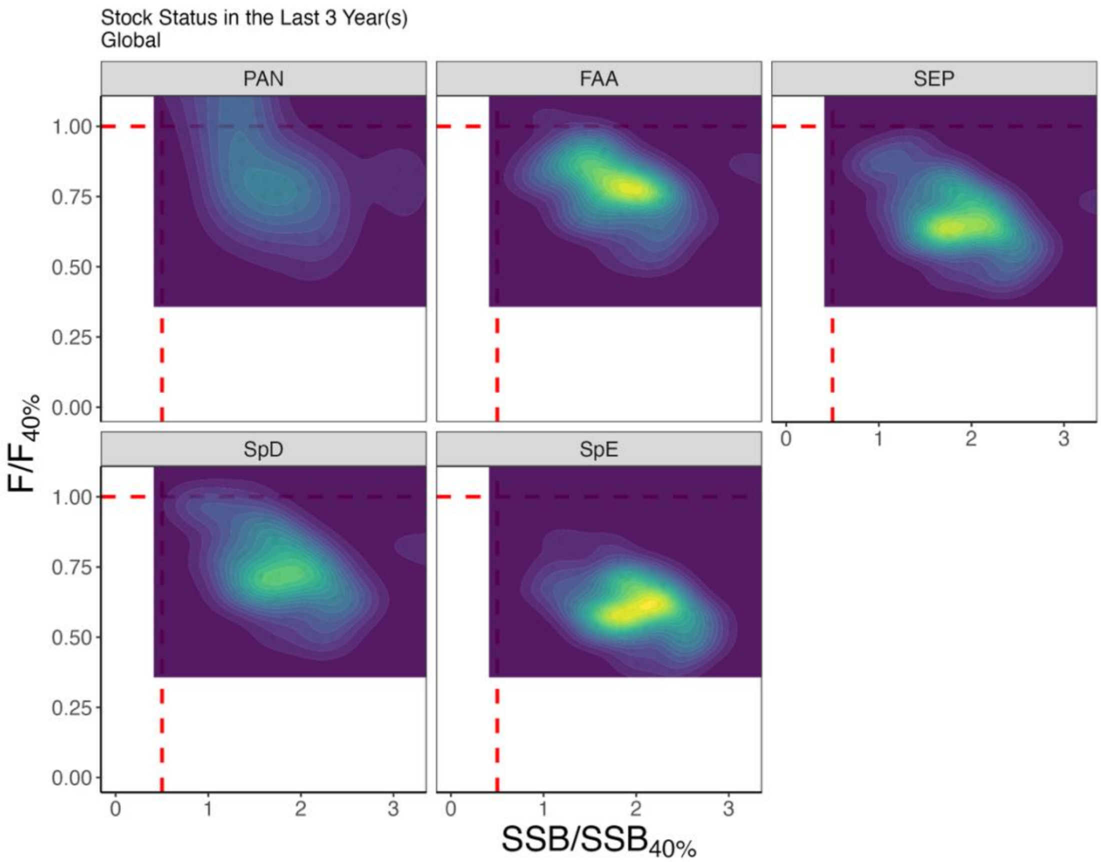


Figure 7. Kobe plot illustrating the global population status over the last 3 years of the feedback period for each estimation model (EM), summarized across 25 realizations.

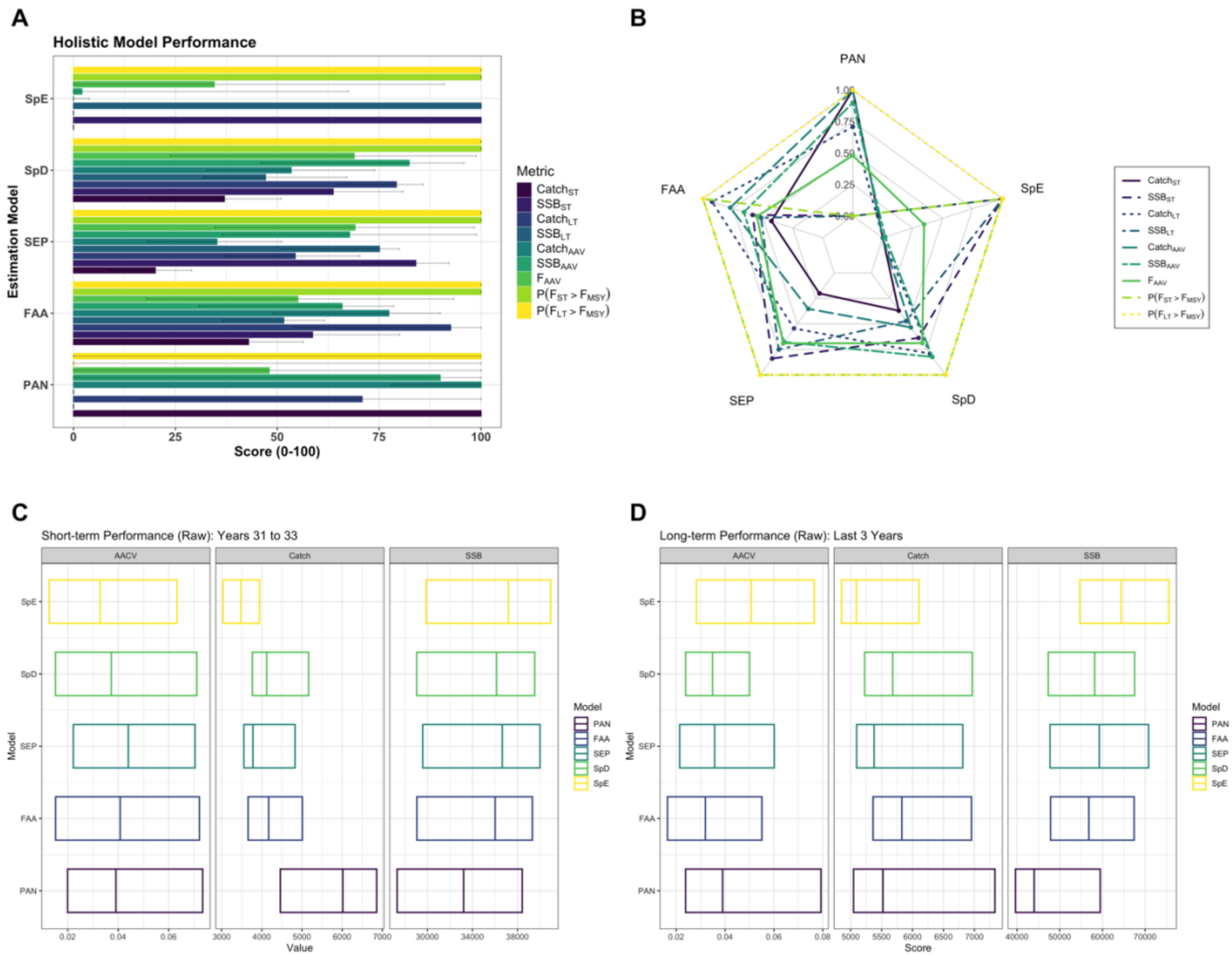


Figure 8. Global-scale performance of estimation models (EMs) summarized across 25 realizations. (A) Standardized performance scores (0–1) for short- and long-term catch, fishing mortality (*F*), and spawning stock biomass (SSB); higher scores indicate better performance. Bars show medians, with horizontal lines indicating interquartile ranges (25th–75th percentiles). (B) Radar chart of median scores for short- and long-term catch and SSB, annual average variation in catch (AACV), SSB (AASV), and *F* (AAFV), and the probability of overfishing in the short and long term. (C–D) Global AACV, catch, and SSB in the short term (C) and long term (D). Boxplots display medians (horizontal line), interquartile ranges (box), and whiskers extending to 1.5× the interquartile range.

13. **Supplementary Files**

Table S1. Description of fishing mortality (*F*) configuration options available in the SPASAM-MSE framework. These options allow users to define initial *F*, select among historical fishing patterns, or provide user-specified *F* values.

Option	Description
F.year1	Fishing mortality in the first year
Fhist	Pattern of historical fishing mortality. Available options:
	“constant”: Constant across years
	“updown”: Increase to “Fmax” until change point, then decrease to “Fmin”
	“downup”: Decrease to “Fmin” until change point, then increase to “Fmax”
	“F-H-L”: Constant “F.year1 × Fmax” until change point, followed by constant “F.year1 × Fmin”
	"F-L-H": Constant “F.year1 × Fmin” until change point, followed by constant “F.year1×Fmin”
Fmax	Maximum F (or multiplier when Fhist = "F-H-L" or Fhist = "F-L-H")
Fmin	Minimum F (or multiplier when Fhist = "F-H-L" or Fhist = "F-L-H")
change_time	Proportion of the time series (0–1) indicating when the change in fishing mortality occurs.
user_F	Optional matrix (`n_years × n_fleets`) of user-specified fishing mortality values.

Table S2. Catch apportionment strategies available in the SPASAM-MSE.

Strategy Name	Catch Allocation Methods	Equation
Equal allocation	Total catch is equally allocated across all fleets.	$C_{t,f} = \frac{C_t}{N_F}$
Equal by gear	Gear-specific catch is equally divided among fleets using the same gear.	$C_{t,f} = \frac{C_{t,g}}{N_F(g)}, \text{with} \sum_{g=1}^{N_G} C_{t,g} = C_t$
Region-weighted	Catch allocated by historical regional catch; equally split among fleets in region.	$w_r = \frac{\sum_{y=t-Y}^{y=t-1} C_{y,r}}{\sum_{y=t-Y}^{y=t-1} C_y}$ $C_{t,r} = w_r C_t$ $C_{t,f} = \frac{C_{t,r}}{N_G(r)}$
Gear-weighted	Catch allocated by historical gear-specific catch; split among regions using that gear.	$w_g = \frac{\sum_{y=t-Y}^{y=t-1} C_{y,g}}{\sum_{y=t-Y}^{t-1} C_y}$ $C_{t,g} = w_g C_t$ $C_{t,f} = \frac{C_{t,g}}{N_R(g)}$
Region × Gear weighted	Allocated based on both regional and gear-specific history.	$w_r = \frac{\sum_{y=t-Y}^{y=t-1} C_{y,r}}{\sum_{y=t-Y}^{y=t-1} C_y}$ $w_g = \frac{\sum_{y=t-Y}^{y=t-1} C_{y,g}}{\sum_{y=t-Y}^{y=t-1} C_y}$ $C_{t,f} = w_r w_g C_t$
Fleet-specific weighted	Catch allocated by historical catch of each fleet.	$w_f = \frac{\sum_{y=t-Y}^{y=t-1} C_{y,f}}{\sum_{y=t-Y}^{y=t-1} C_y}$ $C_{t,f} = w_f C_t$
Survey-based regional split	Survey-based catch allocation across regions; split equally among fleets in region.	$w_r = \frac{\sum_{y=t-Y}^{y=t-1} I_{y,r}}{\sum_{y=t-Y}^{y=t-1} I_y}$ $C_{t,r} = w_r C_t$ $C_{t,f} = \frac{C_{t,r}}{N_G(r)}$
Survey × Gear weighted	Regional survey-based weights; split among fleets by gear history.	$w_r = \frac{\sum_{y=t-Y}^{y=t-1} I_{y,r}}{\sum_{y=t-Y}^{y=t-1} I_y}$ $w_g = \frac{\sum_{y=t-Y}^{y=t-1} C_{y,g}}{\sum_{y=t-Y}^{y=t-1} C_y}$ $C_{t,f} = w_r w_g C_t$
Multi-survey index equal	Multi-survey-based allocation to regions; equally split among fleets. Catch is equally distributed among fleets within the region.	$w_r = \frac{\sum_{y=t-Y}^{y=t-1} I_{y,r}^{(i)}}{\sum_{y=t-Y}^{y=t-1} I_y^{(i)}}$ $C_{t,r} = w_r C_t$ $C_{t,f} = \frac{C_{t,r}}{N_G(r)}$

Strategy Name	Catch Allocation Methods	Equation
Multi-survey × Gear weighted	Allocated based on multi-survey regional index and gear-specific historical catch.	$w_r = \frac{\sum_{y=t-Y}^{y=t-1} I_{y,r}^{(i)}}{\sum_{y=t-Y}^{y=t-1} I_{y}^{(i)}}$ $w_g = \frac{\sum_{y=t-Y}^{y=t-1} C_{y,g}}{\sum_{y=t-Y}^{y=t-1} C_{y}}$ $C_{t,f} = w_r w_g C_t$
User-defined fleet weights	User provides weights for each fleet.	$C_{t,f} = w_f C_t$
User-defined region weights	User provides weights for regions; catch is equally split among fleets in the region.	$C_{t,r} = w_r C_t$ $C_{t,f} = \frac{C_{t,r}}{N_G(r)}$

Note:

- C_t: Total catch in year t.
- $C_{t,f}$: Catch allocated to fleet f in year t.
- $C_{t,r}$: Catch allocated to region r in year t.
- $C_{t,g}$: Total catch for gear type g in year t.
- $I_{y,r}$: Survey catch for region r in year y.
- I_y: Survey total catch in year y.
- $I_{y,r}^{(i)}$: Survey i catch for region r in year y.
- N_G: Total number of gear types.
- N_F: Total number of fleets.
- N_R: Total number of regions.
- $N_F(g)$: Total number of fleets using gear type g.
- $N_G(r)$: Total Number of gear types in region r.
- w_f: Weight assigned to fleet f.
- w_r: Weight assigned to region r.
- w_g: Weight assigned to gear type g.
- Y: Number of years used for calculating the average.

Table S3. Summary of performance metrics available in SPASAM-MSE.

Performance Metric	**Equation**	**Description**
Short-term Average Catch	$\frac{1}{k}\sum_{y=1}^{k}\text{Catch}_y{}^{(t)}\vert t=1,\ldots,N$	Average catch over the first k years of the feedback period. Summarized across N realizations using the mean or median.
Short-term Average Fully Selected *F*	$\frac{1}{k}\sum_{y=1}^{k}F_y{}^{(t)}\vert t=1,\ldots,N$	Average fully selected F over the first k years of the feedback period. Summarized across N realizations using the mean or median.
Short-term Average *SSB*	$\frac{1}{k}\sum_{y=1}^{k}SSB_y{}^{(t)}\vert t=1,\ldots,N$	Average *SSB* over the first k feedback period. Summarized across N realizations using the mean or median.
Long-term Average Catch	$\frac{1}{m}\sum_{y=n-m+1}^{n}\text{Catch}_y{}^{(t)}\vert t=1,\ldots,N$	Average catch over the last m years of the feedback period. Summarized across N realizations using the mean or median.
Long-term Average Fully Selected *F*	$\frac{1}{m}\sum_{y=n-m+1}^{n}F_y{}^{(t)}\vert t=1,\ldots,N$	Average fully selected fishing mortality over the last m years of the feedback period. Summarized across N realizations using the mean or median.
Long-term Average *SSB*	$\frac{1}{m}\sum_{y=n-m+1}^{n}SSB_y{}^{(t)}\vert t=1,\ldots,N$	Average *SSB* over the last m years of the feedback period. Summarized across N realizations using the mean or median.
Relative Difference in Short-term Average Catch	$\frac{1}{k}\sum_{y=1}^{k}(\frac{\text{Catch}_{y,\text{EM}}{}^{(t)}-\text{Catch}_{y,\text{baseline}}{}^{(t)}}{\text{Catch}_{y,EM}{}^{(t)}})\vert t=1,\ldots,N$	Average relative differences in catch between a candidate EM and a baseline EM over the first k years of the feedback period. Summarized across N realizations using the mean or median.
Relative Difference in Short-term Average Fully Selected *F*	$\frac{1}{k}\sum_{y=1}^{k}(\frac{F_{y,\text{EM}}{}^{(t)}-F_{y,\text{baseline}}{}^{(t)}}{F_{y,EM}{}^{(t)}})\vert t=1,\ldots,N$	Average relative differences in F between a candidate EM and a baseline EM over the first k years of the feedback period. Summarized across N realizations using the mean or median.
Relative Difference in Short-term Average *SSB*	$\frac{1}{k}\sum_{y=1}^{k}(\frac{SSB_{y,\text{EM}}{}^{(t)}-SSB_{y,\text{baseline}}{}^{(t)}}{SSB_{y,EM}{}^{(t)}})\vert t=1,\ldots,N$	Average relative differences in *SSB* between a candidate EM and a baseline EM over the first k years of the feedback period. Summarized across N realizations using the mean or median.
Relative Difference in Long-term Average Catch	$\frac{1}{m}\sum_{y=n-m+1}^{n}(\frac{\text{Catch}_{y,\text{EM}}{}^{(t)}-\text{Catch}_{y,\text{baseline}}{}^{(t)}}{\text{Catch}_{y,EM}{}^{(t)}})\vert t=1,\ldots,N$	Average relative differences in Catch between a candidate EM and a baseline EM over the last m years of the feedback period. Summarized across N realizations using the mean or median.
Relative Difference in Long-term Average Fully Selected *F*	$\frac{1}{m}\sum_{y=n-m+1}^{n}(\frac{F_{y,\text{EM}}{}^{(t)}-F_{y,\text{baseline}}{}^{(t)}}{F_{y,EM}{}^{(t)}})\vert t=1,\ldots,N$	Average relative differences in F between a candidate EM and a baseline EM over the last m years of the feedback period. Summarized across N realizations using the mean or median.
Relative Difference in Long-term Average Fully Selected *SSB*	$\frac{1}{m}\sum_{y=n-m+1}^{n}(\frac{SSB_{y,\text{EM}}{}^{(t)}-SSB_{y,\text{baseline}}{}^{(t)}}{SSB_{y,EM}{}^{(t)}})\vert t=1,\ldots,N$	Average relative differences in *SSB* between a candidate EM and a baseline EM over the last m years of the feedback period. Summarized across N realizations using the mean or median.

Performance Metric	Equation	Description
Probability of Overfishing	$P_{(t)} = \frac{1}{n}\sum_{y=1}^{n} I\left(\frac{F_y{}^t}{F_{MSY_y}{}^t} > 1\right) \mid t = 1, \dots, N$	Proportion of years where *F* is over F_{MSY} or the proxy. Summarized across *N* realizations using the mean or median.
Probability of Overfished	$P_{(t)} = \frac{1}{n}\sum_{y=1}^{n} I\left(\frac{SSB_y{}^t}{SSB_{MSY_y}{}^t} < 0.5\right) \mid t = 1, \dots, N$	Proportion of years where *SSB* is below SSB_{MSY} or the proxy. Summarized across *N* realizations using the mean or median.
Overfishing Status	$\frac{1}{n}\sum_{y=1}^{n} \frac{F_y{}^{(t)}}{F_{MSY_y}{}^{(t)}} \mid t = 1, \dots, N$	Overfishing status over *n* years summarized over *N* realizations. Summarized across *N* realizations using the mean or median.
Overfished Status	$\frac{1}{n}\sum_{y=1}^{n} \frac{SSB_y{}^{(t)}}{SSB_{MSY_y}{}^{(t)}} \mid t = 1, \dots, N$	Overfished status over *n* years. Summarized across *N* realizations using the mean or median.
Average Annual Catch Variation	$\frac{\sum_{y=2}^{n} \lvert \mathrm{Catch}_y{}^{(t)} - \mathrm{Catch}_{y-1}{}^{(t)} \rvert}{\sum_{y=2}^{n} \mathrm{Catch}_{y-1}{}^{(t)}} \mid t = 1, \dots, N$	Interannual variability in catch. Summarized across *N* realizations using the mean or median.
Average Annual *F* Variation	$\frac{\sum_{y=2}^{n} \lvert F_y{}^{(t)} - F_{y-1}{}^{(t)} \rvert}{\sum_{y=2}^{n} F_{y-1}{}^{(t)}} \mid t = 1, \dots, N$	Interannual variability in *F*. Summarized across *N* realizations using the mean or median.
Average Annual *SSB* Variation	$\frac{\sum_{y=2}^{n} \lvert SSB_y{}^{(t)} - SSB_{y-1}{}^{(t)} \rvert}{\sum_{y=2}^{n} SSB_{y-1}{}^{(t)}} \mid t = 1, \dots, N$	Interannual variability in *SSB*. Summarized across *N* realizations using the mean or median.
Relative Bias in Model Parameter	$\frac{\hat{\beta}^{(t)} - \beta}{\beta} \mid t = 1, \dots, N$	Relative bias of the estimated model parameter $\hat{\beta}$ to the true model parameter β. Summarized across *N* realizations using the mean or median.
Mean Relative Bias in Management Quantity	$\frac{1}{n}\sum_{y=1}^{n} \frac{\widehat{\theta_y}^{(t)} - \theta_y{}^{(t)}}{\theta_y{}^{(t)}} \mid t = 1, \dots, N$	Mean relative bias of the estimated management quantity $\hat{\theta}$ compared to the true management quantity θ over years. Summarized across *N* realizations using the mean or median.
Score for Catch and *SSB* Related Performance Metrics	$\frac{\mathrm{Metric}_{y,\mathrm{EM}}{}^{(t)} - \min(\mathrm{Metric}_{y,\mathrm{EM}}{}^{(t)})}{\max(\mathrm{Metric}_{y,\mathrm{EM}}{}^{(t)}) - \min(\mathrm{Metric}_{y,\mathrm{EM}}{}^{(t)})} \mid t = 1, \dots, N$	Normalized performance scores for each realization, scaled from 0 (worst) to 1 (best) across EMs. Summarized across *N* realizations using mean or median.
Score for *F* and Annual Variation Related Performance Metrics	$1 - \frac{\mathrm{Metric}_{y,\mathrm{EM}}{}^{(t)} - \min(\mathrm{Metric}_{y,\mathrm{EM}}{}^{(t)})}{\max(\mathrm{Metric}_{y,\mathrm{EM}}{}^{(t)}) - \min(\mathrm{Metric}_{y,\mathrm{EM}}{}^{(t)})} \mid t = 1, \dots, N$	Inverse normalized performance scores for each realization, scaled from 0 (worst) to 1 (best) across EMs. Summarized across *N* realizations using mean or median.

1038

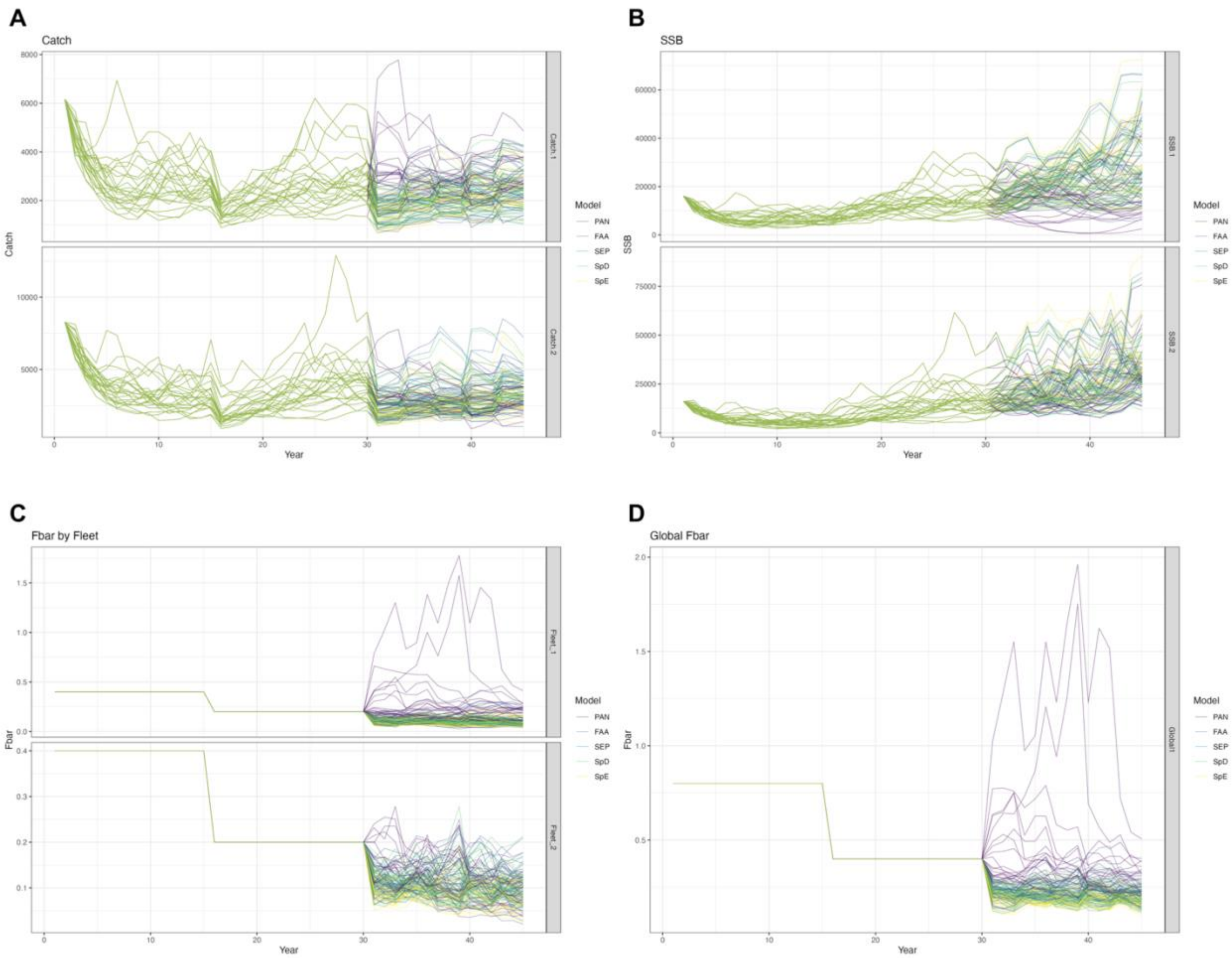


Figure S1. Trajectories of (A) catch, (B) spawning stock biomass (SSB), and (C-D) fishing mortality (*F*) from all estimation models (EMs), across 25 realizations.

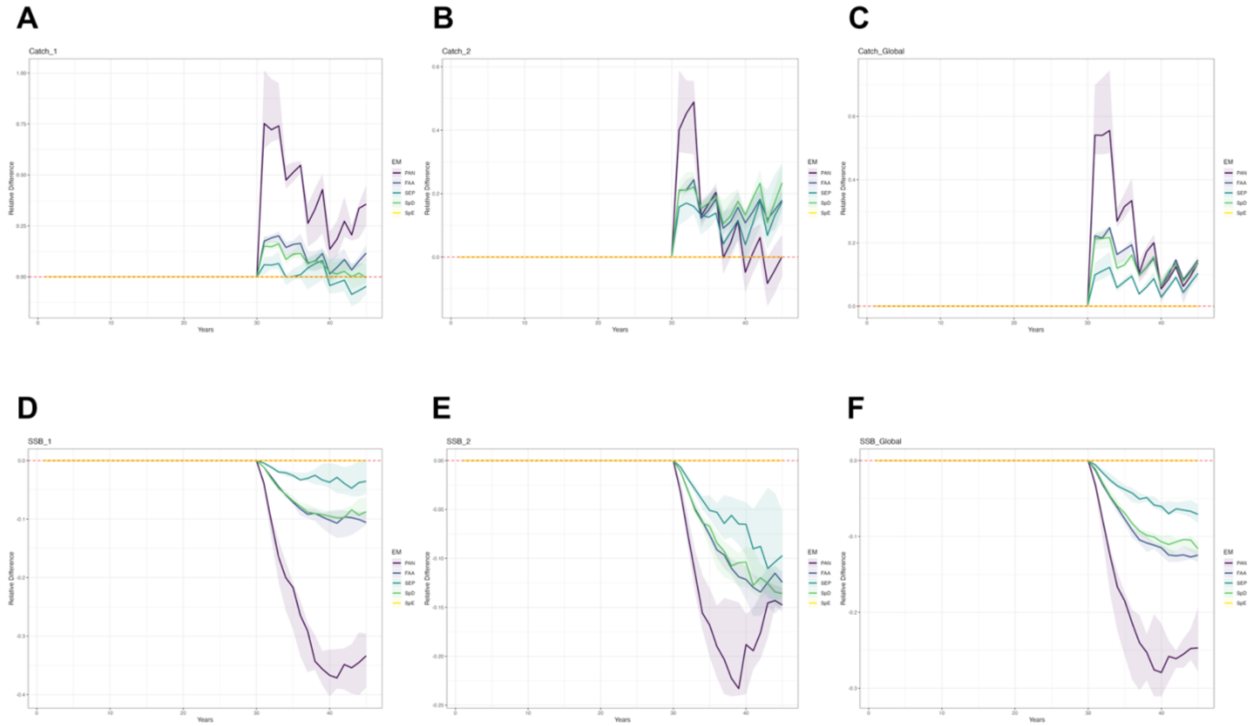


Figure S2. Annual differences in catch and spawning stock biomass (SSB) relative to the spatially explicit EM (SpE), at both regional and global scales. Results are summarized across 25 realizations. Lines indicate the median values, and shaded areas represent the interquantile range from the 40th to 60th percentiles.

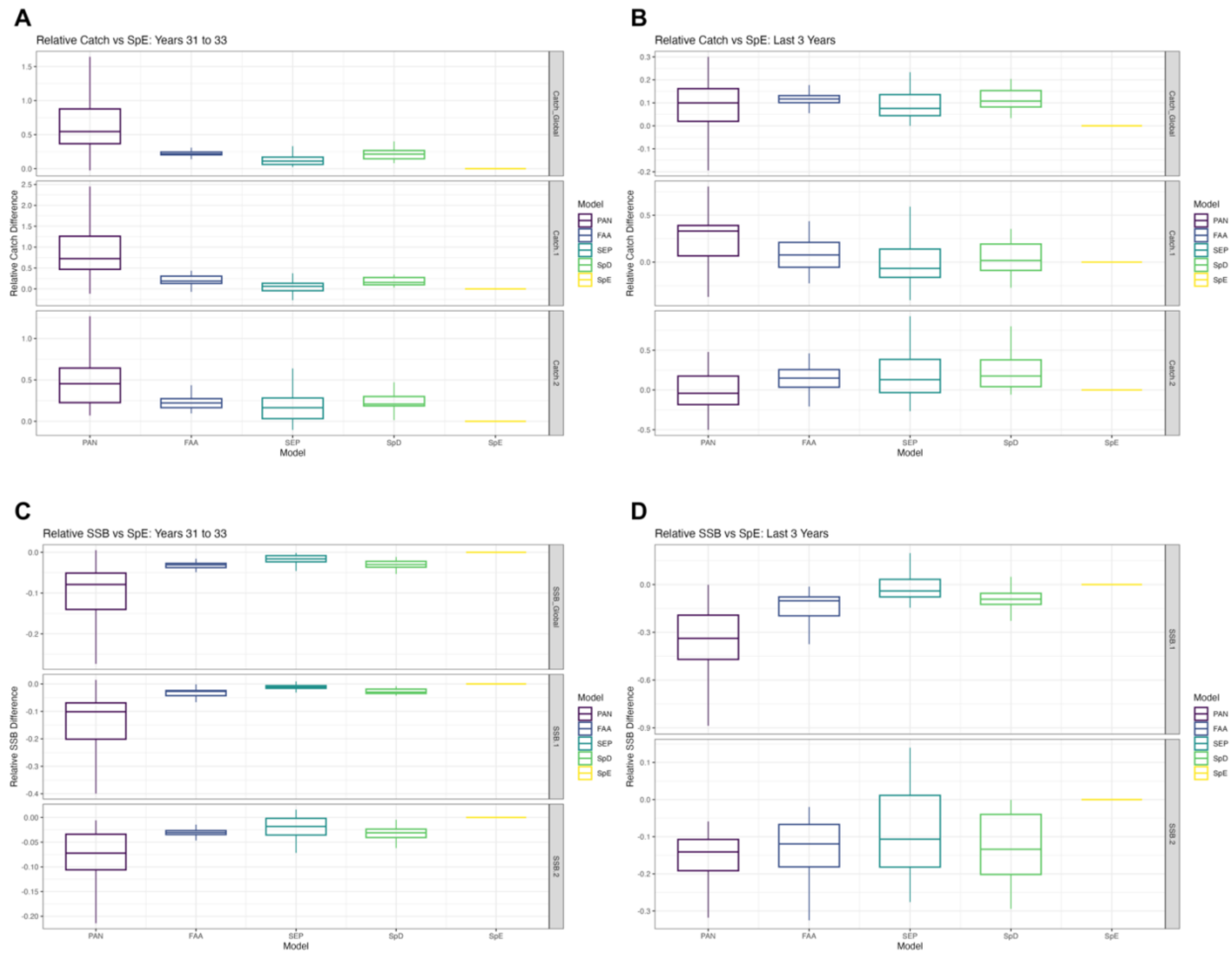


Figure S3. Relative performance of estimation models (EMs) for short-term and long-term catch and spawning stock biomass (SSB), compared to the spatially explicit EM (SpE). For each realization, mean catch or mean SSB was calculated over the evaluation window (short term = first 3 years; long term = last 3 years of the feedback loop). Values for each EM were then expressed relative to the corresponding PAN value within the same realization, such that values >1 indicate better performance than PAN and values <1 indicate worse performance. Panels show relative performance for short-term average catch (A), short-term average SSB (B), long-term average catch (C), and long-term average SSB (D). Boxplots summarize the distribution of these relative performance values across 25 realizations: horizontal lines indicate medians, boxes represent the interquartile range (25th–75th percentiles), and whiskers extend to 1.5× the interquartile range.

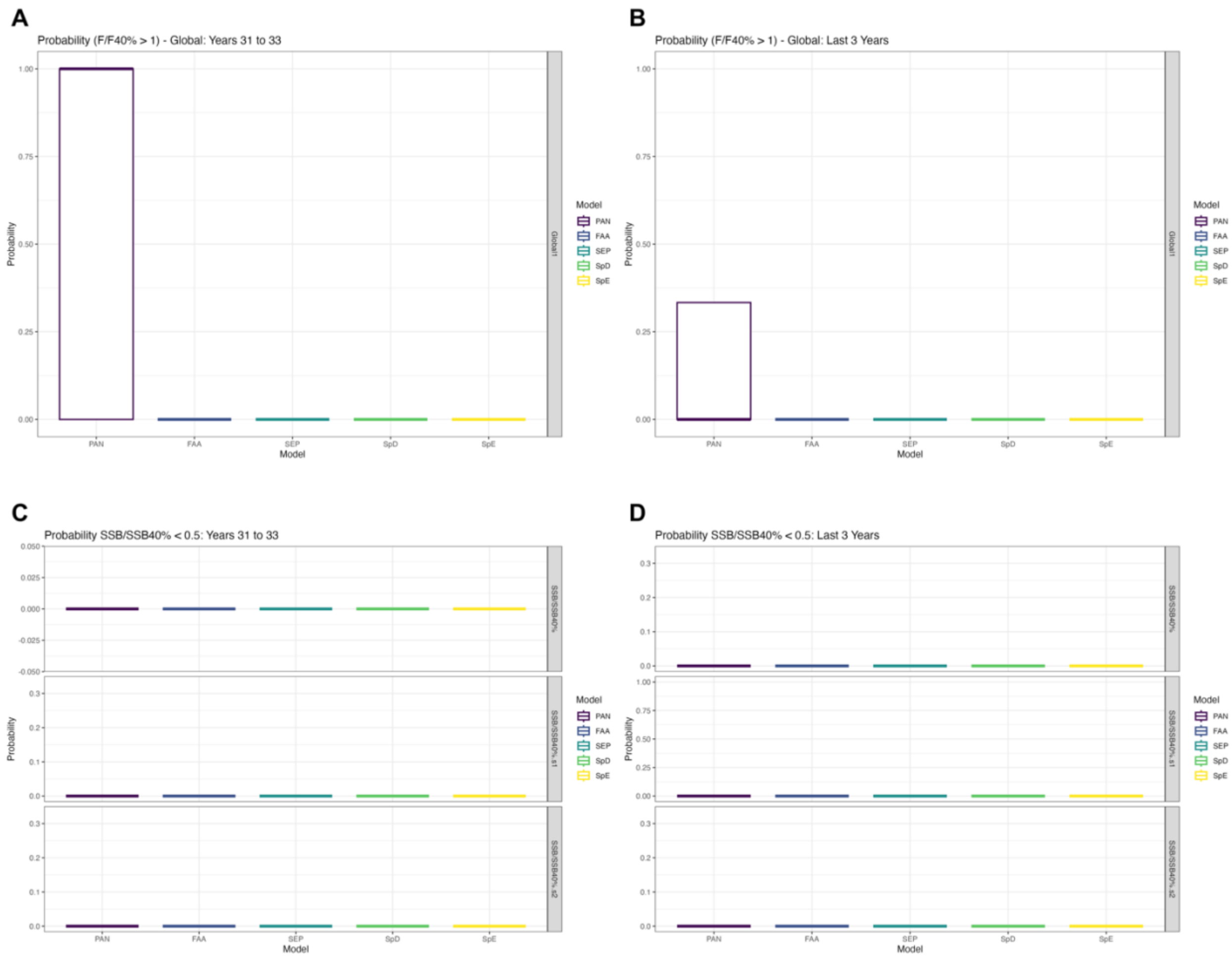
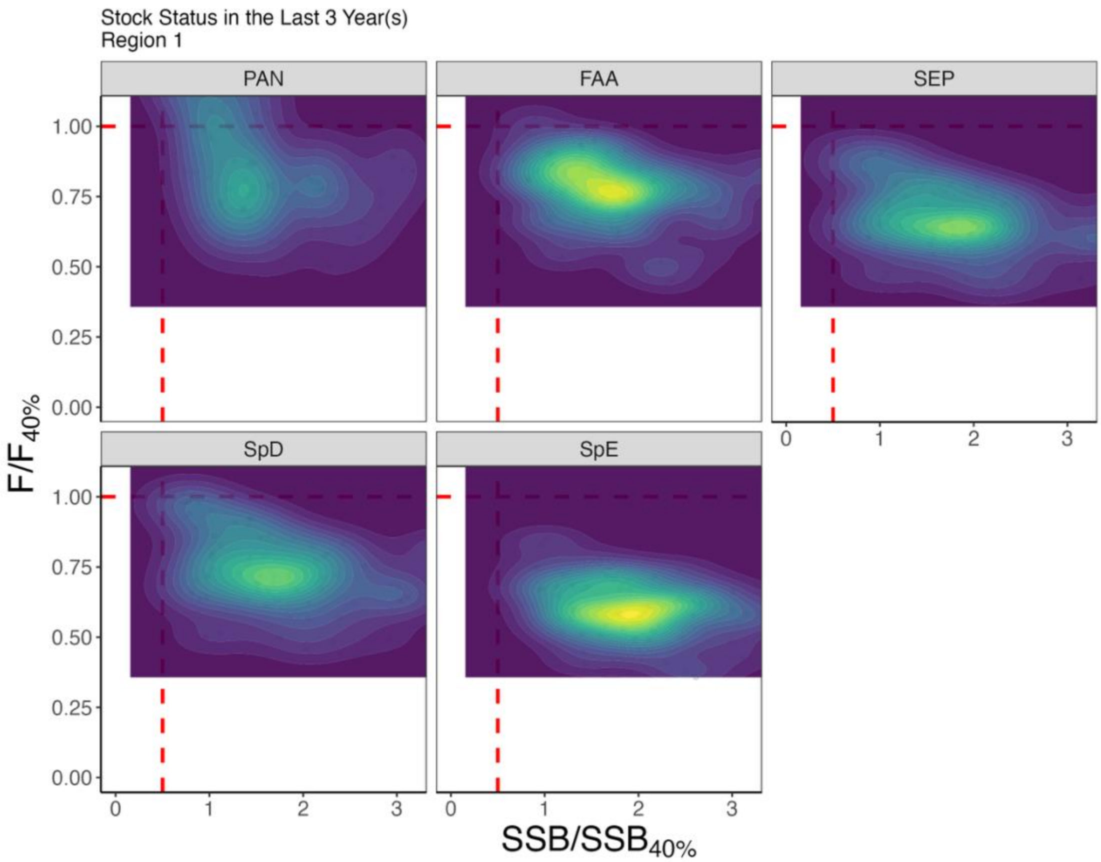


Figure S4. Probability of (A) overfishing in the short-term and (B) long-term, and probability of being overfished in the short-term (C) and long-term (D), across all estimation models (EMs). Results are summarized over 25 realizations. Boxplots show the median and interquartile range (25th–75th percentiles), with whiskers extending to 1.5 times the interquartile range.



1065
1066 Figure S5. Kobe plot illustrating the population status in Region 1 over the last 3 years of the
1067 feedback period for each estimation model (EM), summarized across 25 realizations.

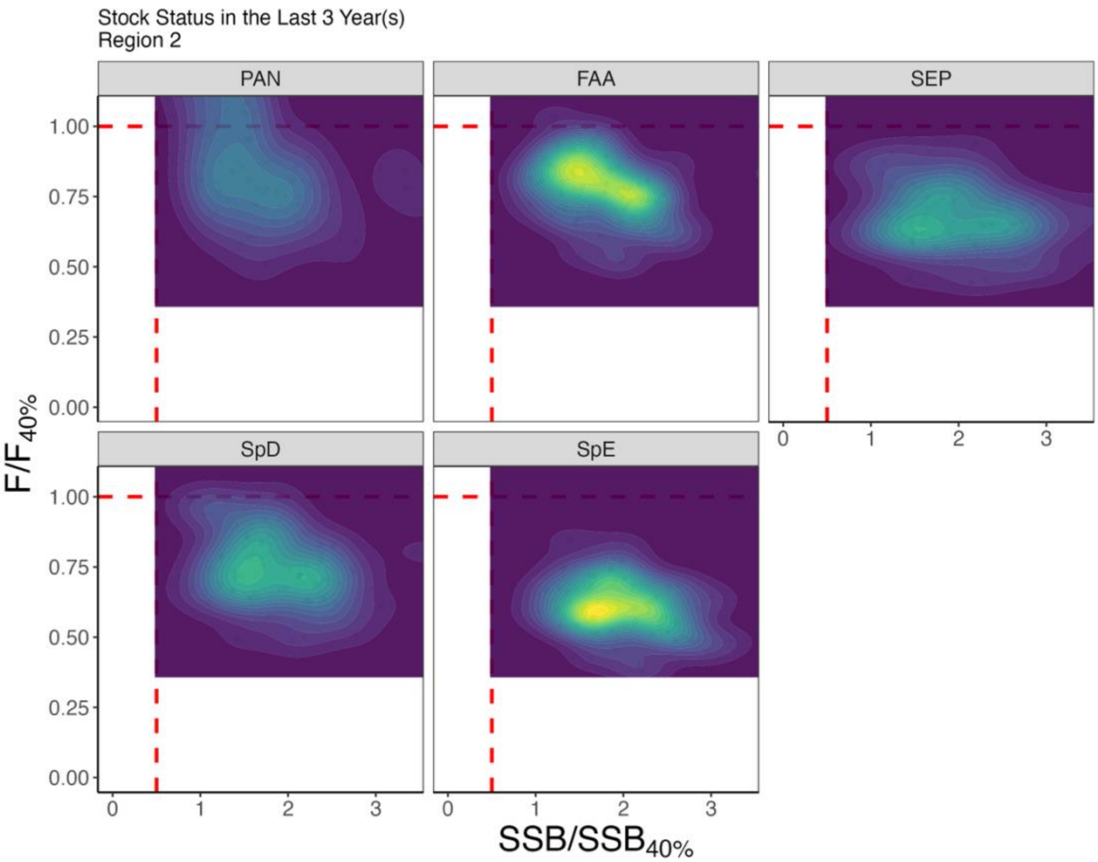


Figure S6. Kobe plot illustrating the population status in Region 2 over the last 3 years of the feedback period for each estimation model (EM), summarized across 25 realizations.

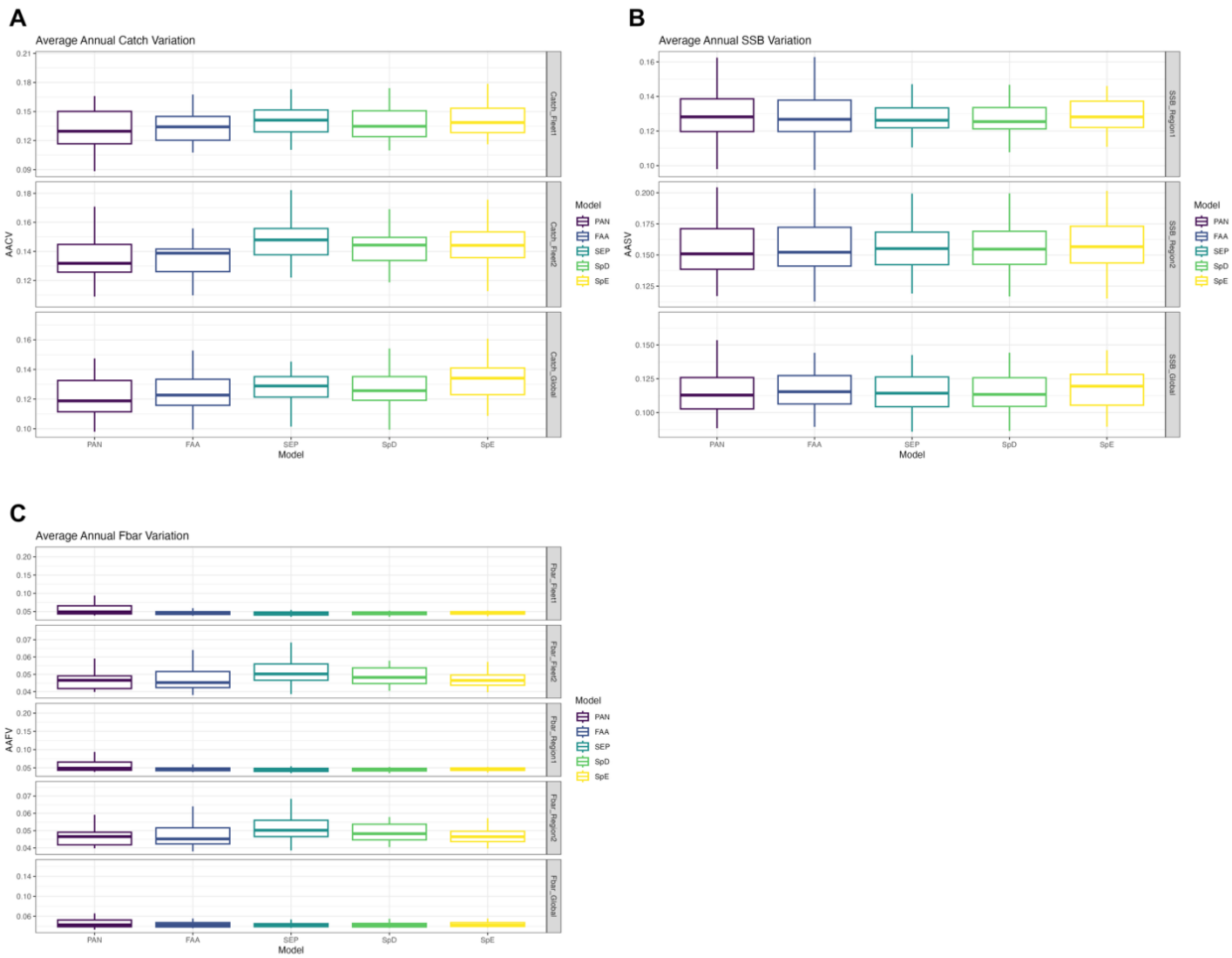


Figure S7. Annual average variation in catch, fishing mortality (*F*), and spawning stock biomass (SSB) at both regional and global scales, calculated over the entire feedback period. Results are summarized across all estimation models (EMs) and 25 realizations. Boxplots show the median and interquartile range (25th–75th percentiles), with whiskers extending to 1.5 times the interquartile range.

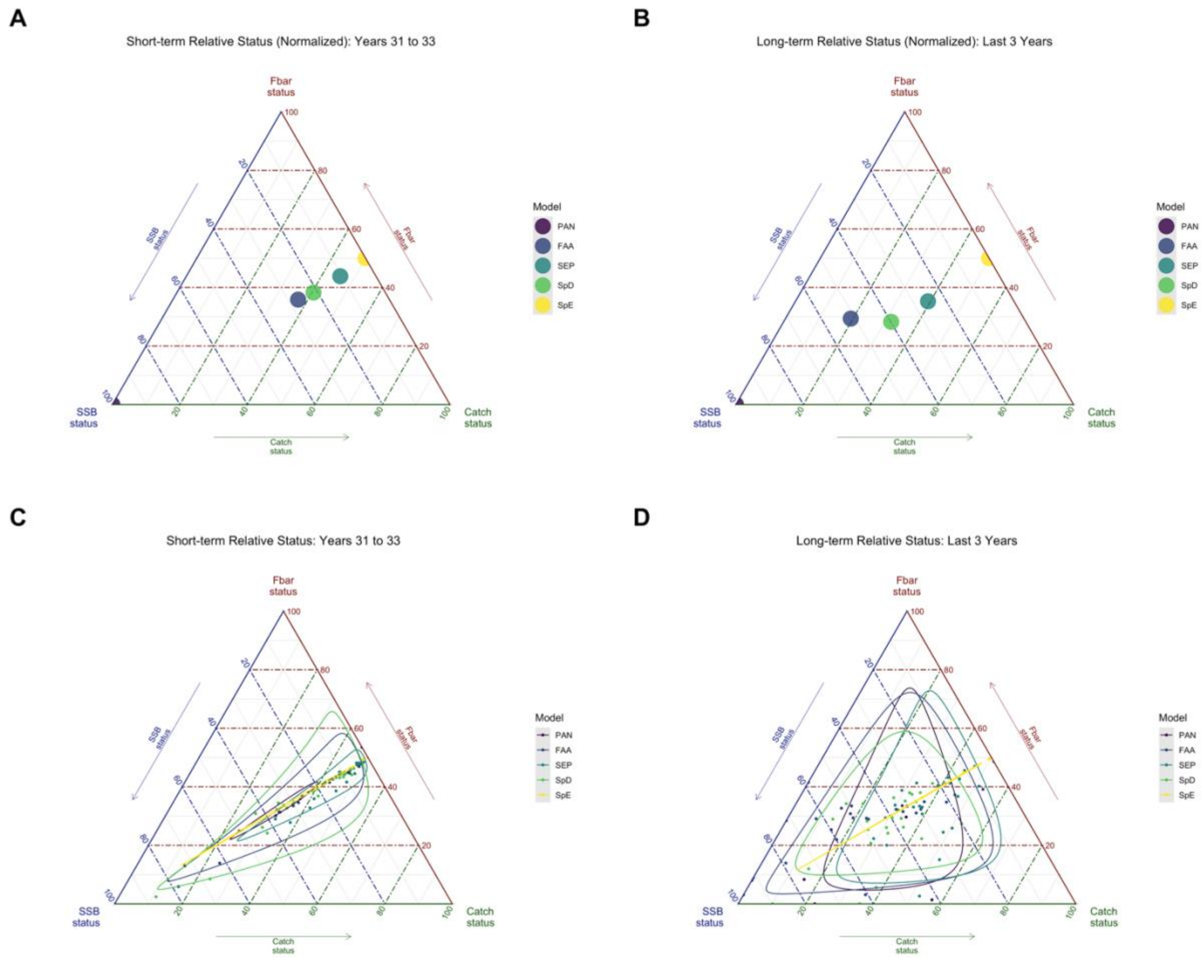


Figure S8. Ternary plots illustrating the trade-offs among three global performance metrics: annual average catch variation (AACV), total catch, and spawning stock biomass (SSB), for each estimation model (EM). Scores were normalized within each realization and inverted for AACV so that higher values consistently represent better performance. Panels A and C show short-term results, and panels B and D show long-term results. In panels A and B, each point represents the median normalized score across 25 realizations per EM. Panels C and D display all raw points from the 25 realizations per EM, along with 95% confidence regions. These plots highlight both central tendencies and variability in model performance across trade-off dimensions.

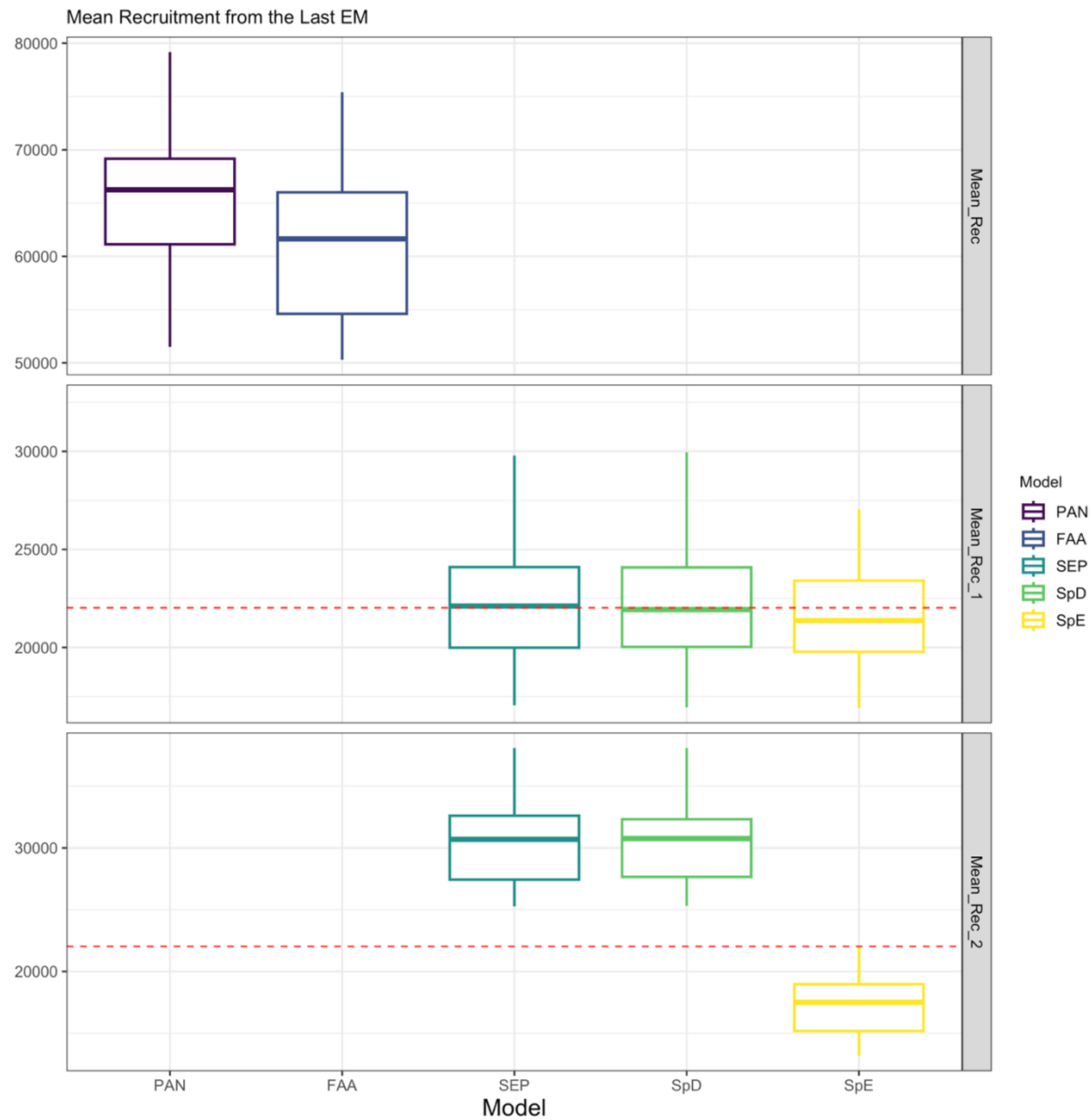


Figure S9. Mean recruitment estimates from each estimation model (EM) based on the final assessment within the feedback period. For panmictic models (PAN and FAA), only global mean recruitment is estimated, while all other EMs estimate region-specific mean recruitment. The red line indicates the true mean recruitment from the operating model (OM). Boxplots show the median and interquartile range (25th–75th percentiles), with whiskers extending to 1.5 times the interquartile range.

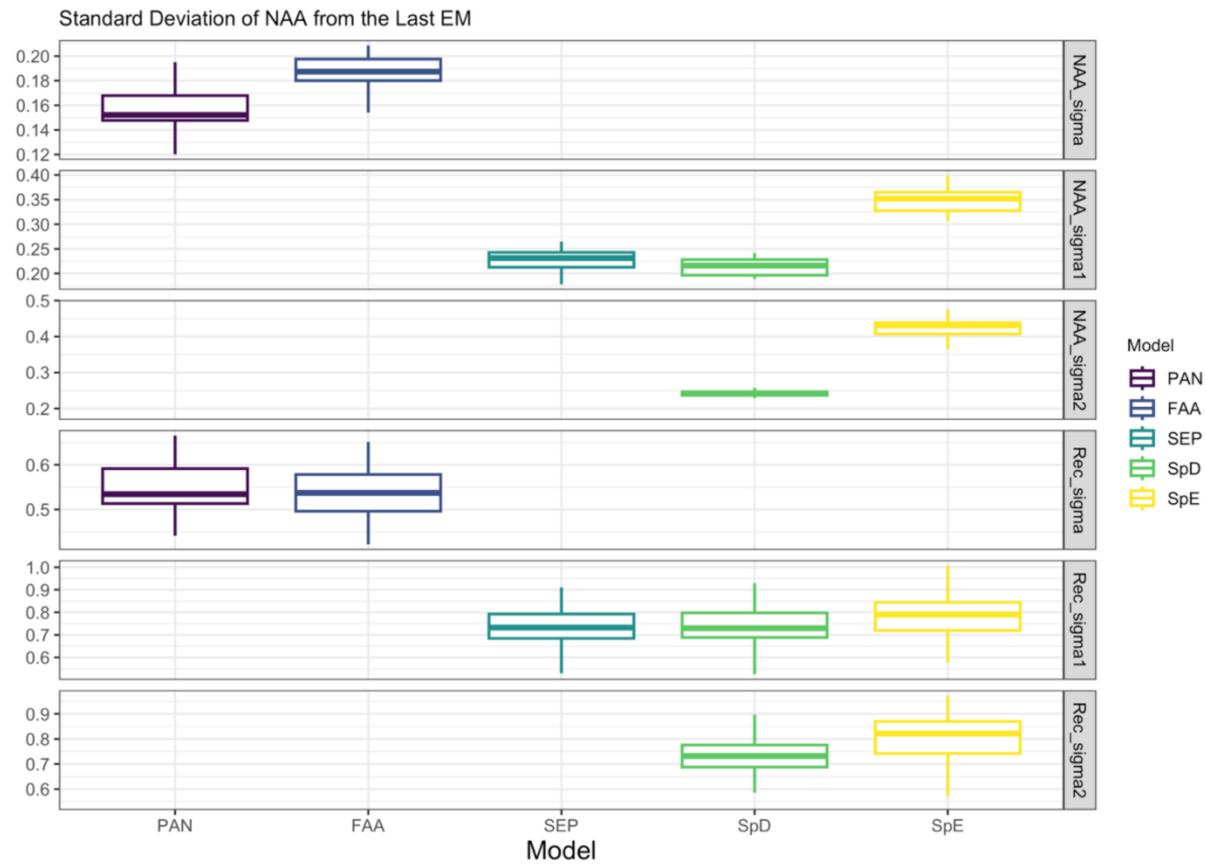


Figure S10. Standard deviation of numbers-at-age (NAA) transitions estimated from each estimation model (EM) based on the final assessment within the feedback period. For panmictic models (PAN and FAA), only global NAA standard deviations were estimated. For all other EMs, region-specific NAA standard deviations were estimated. Boxplots show the median and interquartile range (25th–75th percentiles), with whiskers extending to 1.5 times the interquartile range.