Title

Confronting transitions in fishery fleet structure and selectivity: Practical recommendations for integrated age-structured stock assessments based on simulation analysis

Authors

Matthew L.H. Cheng[1*], Daniel R. Goethel[2], Peter-John F. Hulson[2], Curry J. Cunningham[1]

[1] Department of Fisheries at Lena Point, College of Fisheries and Ocean Sciences, University of Alaska Fairbanks, 17101 Point Lena Loop Rd, Juneau, Alaska 99801, USA

[2]National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center, Auke Bay Laboratories, 17109 Point Lena Loop Road Juneau, Alaska 99801, USA

*Corresponding Author

Abstract

Dynamic shifts in fleet structure and gear usage lead to complex implications for representing fishery selectivity in stock assessment models. There is generally a lack of consensus on how assessment models should be configured to confront changes in fishery fleet structure or associated selectivity forms, while balancing complexity-parsimony tradeoffs. We conducted a simulation analysis to evaluate the performance of alternative assessment models when confronted with fleet transitions among gear types, which included differences in: 1) rates of transition (i.e., a fast or slow transition among gears), and 2) selectivity forms for each modeled fleet (i.e., asymptotic or dome-shaped). In general, explicitly modelling fleet structure (i.e., multi-fleet models) performed well, but demonstrated bias in biomass estimates and management reference points when selectivity forms were mis-specified. Single-fleet models were only unbiased when time-varying selectivity (e.g., using time blocks or continuous formulations) was estimated to account for changes among gear types. Our results suggest that single-fleet models with time-varying fishery selectivity are adequate for operational management advice, but research oriented multi-fleet models should be used as validation tools to explore model consistency within single-fleet models.

Keywords: fishery selectivity, fishing fleet structure, fisheries management, simulation, stock assessment

1 Introduction

Changes in harvest methods within commercial fisheries are common, and can be influenced by market forces, technological advancements, interactions with non-target species, and regulatory frameworks (Branch et al. 2006; Watson and Kerstetter 2006; Eigaard et al. 2014). These changes can be gradual or rapid in nature, which can involve gear modifications (e.g., mesh size), developing new fishing technology to improve fishery yield (Beverton and Holt 1957; Sainsbury 1984; Pauly 1998), or altering the spatial distribution of fishing effort in response to regulatory changes (Beare et al. 2013). For instance, in a Hawaiian longline tuna fishery, gradual transitions in hook shape and widths increased the selection of larger and more valuable individuals, while reducing bycatch rates of non-target species (Gilman et al. 2012). Similarly, attempts to reduce juvenile mortality of North Sea plaice (*Pleuronectes platessa*) resulted in the implementation of an area closure (known as the "plaice box") for the fishery, rapidly altering the seasonal and spatial distribution of fishing effort (Pastoors 2000; Aarts and Poos 2009). Understanding harvester and management-driven changes in fishery practices is critical, given the strong influence of fishery processes on the demographics of a population (Brunel and Piet 2013) and the provision of management advice (Beverton and Holt 1957; Scott and Sampson 2011; Sampson 2014).

Stock assessment models, which estimate the impact of harvest on fish populations while accounting for critical biological processes (e.g., recruitment, growth, and natural mortality) that govern population dynamics, commonly form the scientific basis for fisheries management advice (Quinn and Deriso 1999). Most contemporary assessments utilize statistical catch-at-age models (hereafter, stock assessment models) via an integrated analysis framework, where several data sources (e.g., catch, abundance indices, and age or length composition data) are integrated into a single analysis to estimate population status and project population dynamics under alternative harvest strategies (Fournier and Archibald 1982; Maunder and Punt 2013). Under the integrated analysis framework, removals due to harvest from the population are characterized by defining one or more fishery fleets, also referred to as the fleet structure. Each fishery fleet is often associated with catch and compositional data, as well as a selectivity curve to describe age- or length-specific removals (hereafter, fishery selectivity). More generally, fishery selectivity in stock assessment models encompasses both the probability of capturing an individual when encountered (i.e., contact selectivity) and the probability of spatial and temporal overlap with individuals during fishing operations (i.e., availability; Sampson, 2014).

Defining fleet structure and parameterizing fishery selectivity is a primary assumption in stock assessment models, necessitating explicit decisions regarding the number of fleets to represent, the shape of the selectivity curve, how that curve is parameterized, and the potential for variation in selectivity over time (Punt et al. 2014a). The number of fishery fleets to model depends on the availability of fleet-specific data, the degree of contrast in fleet dynamics, and the management structure (e.g., whether quotas are fleet-specific). Although explicitly modelling the full diversity of fleets (e.g., gears) in a fishery may better represent removal processes within the population and allows for the provision of fleet-specific management advice, there is potential for introducing additional uncertainty in model estimates if multi-fleet models are not supported by the available data. Another important modelling consideration involves determining the shape of the selectivity curve (e.g., asymptotic or dome-shaped) for these fleets. Several approaches can be utilized to represent the shape of the selectivity curve, which include parametric and non-parametric approaches (Thorson and Taylor 2014; Privitera-Johnson et al. 2022). The former generally provides a more parsimonious approach due to a reduced number of parameters, but the latter is more flexible and robust to model misspecification, enabling the characterization of a wider range of possible shapes. Regardless of how selectivity curves are implemented, there is risk of bias in management advice if selectivity is specified incorrectly (Maunder and Piner 2015). Lastly, assumptions regarding potential time-variation in fishery selectivity is a critical decision that must be addressed. While changes in selectivity may occur due to fluctuations in harvest methods or market demands; Eigaard et al., 2014, 2011; Sampson and Scott, 2012), time-invariant selectivity is a common assumption in many stock assessment models, given limitations in the available age-or length-composition data. When assessment models are not limited by the available data, time-variation in selectivity can be accounted for by allowing continuous changes using autoregressive models (Linton and Bence, 2011; Xu *et al.*, 2019). Discrete time blocks can also be implemented, where selectivity is estimated for a pre-defined block and assumed to remain constant within time blocks. The choice of pre-defined blocks is subjective, but it is typically based on an observable major change in the fishery (e.g., the introduction of new gear types). Properly addressing time-varying dynamics in fishery selectivity is critical for providing adequate management advice and inappropriate assumptions can potentially manifest as consistent directional biases in stock assessment estimates for biomass (as was demonstrated in the example of the Pacific Halibut (*Hippoglossus stenolepis*) assessment; Stewart and Martell, 2014).

There is a wide range of fishery fleet structure complexity that can be integrated into an assessment model depending on the spatial, temporal, gear, and stock dynamics present. For instance, if multiple gear types (e.g., trawl and hook-and-line) exist within a fishery, each gear could be represented as its own fleet (i.e., a multi-fleet model), with removals resulting from gear-specific selectivity patterns. Similarly, fleet structure can also be defined to represent removals occurring in different sectors (e.g., commercial and recreational; Bohaboy *et al.*, 2022) or areas (Cope and Punt, 2011; Berger *et al.*, 2012; Hurtado-Ferro *et al.*, 2014), with removals represented with a sector or area-specific selectivity pattern. Alternatively, fleets can be aggregated across gears or areas (i.e., a single-fleet model), which can reduce complexity and improve tractability of an assessment, particularly when data available to inform fleet-specific processes are limited (e.g., age or length compositions). Aggregation of fishery fleets is a common assumption in many assessments and generalized platforms (Nielsen et al. 2021), but the implications of ignoring complex fleet structure have yet to be thoroughly evaluated.

To date, there has been limited analysis of how best to account for fleet structure transitions over time within stock assessment models (Cheng et al., 2024) or how to select among different selectivity parameterizations for newly-emerging fleets. In cases where multiple fishery fleets have operated and been explicitly managed as discrete units for extensive periods, multi-fleet models are often already utilized. Given the existing need to provide catch advice specific to each fleet, fleet-specific monitoring provides the data necessary to support the implementation of multi-fleet models (e.g., as is done in the Gulf of Mexico red snapper assessment; SEDAR 2018). In these instances, incorporating transitions in fleet structure is easily achieved, given the explicit representation of fleet structure within the modelling framework. However, addressing transitions in fleet structure is more challenging when distinctions among fleets are uncertain, the implementation of multi-fleet models are unsupported by the available data, or when a new fishery sector emerges over time. To address gradual transitions in fleet structure, Nielsen et al. (2021) showed that estimates from single-fleet models assuming non-parametric time-varying selectivity were consistent with multi-fleet models for North Sea and Western Baltic herring (*Clupea harengus*). In the presence of a rapid (i.e., less than five years) and near complete change in gear type usage (i.e., a transition from longline hooks to longline pots), Cheng et al. (2024) compared disaggregated fleet and aggregated fleet models for the Alaska sablefish (*Anoplopoma fimbria*) assessment. Their results indicated that an aggregated fleet model adequately addressed changes in fishery dynamics by defining a discrete time block that approximately coincided with the change in fleet structure, whereas data limitations impeded the estimation of selectivity parameters in disaggregated fleet models, resulting in management advice that was likely overly optimistic.

Although numerous methods exist for addressing changes in fishery dynamics within assessment models, it remains ambiguous how practitioners should simultaneously address uncertainties in selectivity forms, time-variation, and transitions in fishery fleet structure (or the potential benefits of disaggregating fishery fleets), while balancing complexity-parsimony tradeoffs. To address these uncertainties, we performed a simulation experiment using an age-structured operating model to evaluate the performance of alternative assessment models when confronted with transitions among gear types, which included variability in the: 1) rates of transition (i.e., a slow or fast transition among gears), and 2) selectivity forms for modeled fleets (e.g., asymptotic or dome-shaped). Insights from our study offer pragmatic guidance to stock assessment practitioners seeking to determine assessment model configurations for addressing changes in fishery fleet structure and selectivity.

2 Methods

To explore how fleet structure and selectivity parameterizations may impact assessment performance, we developed operating models (OMs) that emulated the biology and recent fleet transitions that have occurred in the Alaska sablefish fishery (Cheng et al. 2024). Each OM assumed two fishery fleets were operating. To investigate model performance across a range of scenarios, we also developed OMs that differed in their rates of transition among gear types and their assumed selectivity forms. These OMs were the basis of comparison and represented the truth, while also providing the simulated data to which estimation models (EMs) were fit. In total, 10 EMs with differing assumptions regarding fleet structure and selectivity were applied to these simulated datasets following a full-factorial design. To understand the influence of available data on model performance following a change in fleet structure, all EMs were applied to three assessment periods in each OM. These three periods represented different intervals after a fleet structure change began (further described in *Operating Model Configurations*; Fig. 1A; colored lines). Model estimates were compared to the true dynamics generated from respective OMs to identify model robustness and performance. In each OM and EM combination, AIC model selection was also conducted to evaluate this criterion's reliability to select assessment models that were correctly parameterized (i.e., EMs matched the OM structure), and its ability to determine parsimonious EMs (i.e., those demonstrating minimal bias with intermediate model complexity). Analyses were conducted in the R statistical environment and EMs were configured in Template Model Builder (TMB; Kristensen *et al.*, 2016). Code associated with this study can be found at https://github.com/chengmatt/Fleet_Selex_Sim. A description of OM and EM configurations are provided in the following sections, and further details can be found in Supplementary Material 1.

2.1 Operating Model Configurations

OMs were sex- and age-structured and represented a single homogeneous population. Annual recruitment was simulated based on a Beverton-Holt stock-recruit relationship, with steepness set at 0.85 (Francis 1992). Dynamics in the OM were generally based on the life-history characteristics and estimated parameter values from the 2021 Alaska sablefish stock assessment (i.e., the OMs were conditioned on the dynamics from the sablefish stock; Goethel et al., 2021). Alaska sablefish are a fast-growing and long-lived species (individuals can live up to 90 years) that exhibit spasmodic recruitment and sexually dimorphic growth, where females reach a larger asymptotic size compared to males. Simulations were based on Alaska sablefish given interest in developing good practices to account for changes in fishery fleet structure (Goethel et al. 2022; Cheng et al. 2024), as observed in the Alaska sablefish fishery starting in 2017. In particular, the fixed-gear fishery (hook-and-line and pot gear) experienced a rapid transition in fleet structure (within 5 years) during this period. Prior to 2017, removals from pot gear were minimal (~5%), while the majority of removals were predominately from hook-and-line gear. However, following a regulatory change that allowed for pot gear use in the Gulf of Alaska in 2017 and the emergence of a new gear type ("slinky" pots), total removals from pot gear increased to comprise ~80% of total removals from the fixed-gear fishery by 2022 (Goethel et al. 2022, 2023). Although aspects of this simulation study are specific to Alaska sablefish, alternative removal scenarios are introduced to encompass a wider range of potential changes in fishery fleet structure that may be applicable to other fisheries.

Six distinct OMs were developed to explore the combinatory effects of different rates of transition in fleet structure (i.e., fast, or slow; expressed through changes in fleet-specific fishing mortality rates) and selectivity forms (see Table 1). Each OM includes two fishery fleets and a single fishery-independent survey, all of which operated continuously across the time-series. The predominant fishery fleet's (i.e., the fleet exhibiting the highest fishing mortality) selectivity form at the start of each simulation was always logistic, generally resembling the hook-and-line fishery for Alaska sablefish. For clarity, all OM names are non-italicized and will be denoted with the rate of transition among gear types followed by the selectivity form of the predominant fleet after the transition. For example, 'Fast-Logistic' denotes a fast transition in fishing mortality rates from a predominant gear with logistic selectivity at the start of the time series to a predominant gear type also with logistic selectivity at the end of the time series (Table 1).

Two annual trends in fishing mortality were simulated to represent different rates with which a new fishery fleet might develop. Simulating various ways in which fishery fleet structure changes allows the utility of alternative EMs in addressing such changes to be compared. First, we simulated a "fast" transition where the fishing mortality rate from fishery fleet 2 increased starting in year 25, from 5% of the total fishing mortality to 75%, over a span of 5 years (i.e., the fleet transition ended in year 30; Fig. 1A). A total of 50 years was simulated for the fast transition scenario. A fast transition is akin to fishery dynamics for Alaska sablefish as described above, wherein a regulatory change and the emergence of a new gear type precipitated a rapid transition in removals among two gear types. Next, we simulated a "slow" transition where the fishing mortality rate from fishery fleet 2 increased gradually starting in year 25 and reached an apex in year 50 (i.e., the transition occurred across a span of 25 years), comprising 75% of the total fishing mortality and remained at that level for the remainder of the simulation. In the slow case, a total of 70 years were simulated (Fig. 1A, Table 1). The slow scenario is similar to Nielsen et al. (2021) and can be conceived as gradual improvements to fishing gear. A total of 50 years were simulated in the first case and 70 years in the second case to ensure that both fast and slow scenarios had 20 years with their respective fisheries at a new fleet transition equilibrium post change.

To explore how differences in fleet structure transition rates, compounded with contrast in selectivity among fishery fleets may influence EM performance, three selectivity scenarios were simulated (Fig. 1B). In the first selectivity scenario, removal patterns from both fishery fleet 1 and fishery fleet 2 demonstrated logistic selectivity (Logistic):

Eq. 1

$$sel_{t,a,s,f} = \left[1 + e^{-k_{s,f}\left(a - a_{s,f}^{50}\right)}\right]^{-1}$$

where subscripts *t, a, s,* and *f* index years, ages, sexes, and fleets, $sel_{t,a,s,f}$ represents selectivity, $k_{s,f}$ is the slope of the selectivity curve, and $a_{s,f}^{50}$ is the age-at-50% selectivity. Fishery fleet 1 selected younger individuals from the population, while fishery fleet 2 selected older individuals (Fig. 1B). This selectivity pattern can be envisioned as the introduction of a new gear type that better targets older individuals or reduces the selection of younger individuals (e.g., through changes in mesh sizes or hook types).

The other two selectivity scenarios (Gamma) assumed removal patterns from fishery fleet 1 resulted from logistic selectivity (Eq. 1), while removal patterns from fishery fleet 2 were parameterized as a gamma function, to allow for dome-shaped selectivity (Punt *et al.*, 1996). Here, the oldest individuals were less vulnerable to harvest compared to the Logistic scenario:

Eq. 2

$$sel_{t,a,s,2} = \left(\frac{a}{a_s^{max}}\right)^{\left(\frac{a_s^{max}}{p_s}\right)} e^{\frac{a_s^{max} - a}{p_s}}$$

$$p_s = 0.5 * \left[\sqrt{a_s^{max2} + 4\gamma_s^2} - a_s^{max}\right]$$

where a_s^{max} describes the age-at-maximum selection, γ_s represents the slope of the ascending and descending limbs, and p_s is a quantity derived from a_s^{max} and γ_s. This selectivity can be envisioned as an introduction of a new gear type, with a distinct pattern of harvesting fewer older fish compared to the logistic selectivity assumed in fishery fleet 1. Two versions of the gamma selectivity function were implemented for fishery fleet 2, which were Gamma-Old and Gamma-Young, and differed in their degree of doming in selectivity (Figure 1). In particular, the Gamma-Old scenario had an older age of maximum selection and selected older individuals. Conversely, Gamma-Young selected comparatively younger individuals. The Gamma-Young scenario can be envisioned as the emergence of a novel market (i.e., small fish) or a regulation change to protect larger, mature fish (e.g., a harvest slot; Bohaboy et al. 2022). For all scenarios, selectivity patterns were specified to be time-invariant for a given fleet, while males were selected at an older age compared to females (i.e., given smaller size-at-age for male sablefish). Across all OMs in this study, the survey fleet was represented with time-invariant logistic selectivity (Eq. 1), which is consistent with the current understanding of survey selectivity for Alaska sablefish. While alternative selectivity forms could have been utilized, logistic selectivity was assumed for the survey fleet to reduce the potential for model confounding, particularly when coupled with a fishery that had dome-shaped selectivity.

Several data types were generated from the six OMs, which included catch data, age-composition data, and an abundance index. Data were simulated for both fishery and survey fleets across the entire modeled time-series. Observed catch data for each fishery fleet were simulated with negligible observation error (CV = 0.001) assuming a lognormal distribution. Fishery age-composition data were generated following a multinomial distribution. The associated input sample size (the sample size that reflects the over-dispersion of compositional data, ISS) varied in proportion to the annual instantaneous fishing mortality rates specified for each fleet, which increased samples for fleets with higher fishing effort (i.e., as would be the case for real world observer coverage and monitoring; Fig. 1A):

Eq. 3
$$\mathrm{ISS}_{t,s,f} = [\frac{F_{t,f} - \min(\boldsymbol{F}_f)}{\max(\boldsymbol{F}_f) - \min(\boldsymbol{F}_f)}\left(\mathrm{ISS}_{s,f}^{\max} - \mathrm{ISS}_{s,f}^{\min}\right)] + \mathrm{ISS}_{s,f}^{\min}$$

where $\mathrm{ISS}_{t,s,f}$ is the input sample size and $F_{t,f}$ is the fleet-specific instantaneous fishing mortality rate. $\mathrm{ISS}_{s,f}^{\min}$ and $\mathrm{ISS}_{s,f}^{\max}$ are pre-defined minimum and maximum values of input sample sizes, specified at 50 and 100 and are distributed across sexes based on their sex-ratios (i.e., to reflect sex-specific availability), respectively. Observations from the fishery-independent survey included an abundance index that was simulated with lognormal error (CV = 0.2). Age-composition data for the survey were generated following a multinomial distribution with a constant ISS of 100. A total of 200 replicate datasets were simulated to encapsulate variation in both observation and process error.

Lastly, for each OM, three different assessment periods were used to evaluate how model performance may depend on the length of the available data time series following the change in fishery fleet structure. These included: 1) when the instantaneous fishing mortality for the two fleets intersected (Fast: year 27; Slow: year 40), 2) when the fleet transition concluded (Fast: year 30; Slow: year 50), and 3) the terminal period, which was 20 years after the completed transition (Fast: year 50; Slow: year 70; Fig. 1A; colored lines). Collectively, these OM scenarios aim to provide pragmatic guidance for EM parameterizations (i.e., fleet structure, selectivity forms, and time-variation), while considering the dependence of model parameterizations on available data. For a summary and abbreviation of OM scenarios, see Table 1.

<u>*2.2 Estimation Model Configurations*</u>

A total of 10 EMs were configured to assess model performance, which represented common stock assessment approaches utilized when practitioners are confronted with complex fleet structure and fleet transitions. All EMs were single area sex- and age-structured models, configured as either a multi-fleet or single-fleet model (Supplementary Material 1). In general, EMs mimicked the structure of the OMs, except for assumptions regarding fishery fleet structure, the treatment of time-varying selectivity, and selectivity functional forms. Each EM was applied to all OM scenarios, following a full factorial design. EM names are italicized and are first denoted with the assumed fleet structure (i.e., *2Fleet* or *1Fleet*). This is then followed by the assumption regarding time-variation, which only applies to *1Fleet* models (i.e., *TimeInvar, Block*, *RandWlkPar, SemiPar*, see Single Fleet Models section below for further details). Finally, the name concludes with the assumed selectivity for the predominant fleet following the transition in fleet structure (e.g., *Logistic* or *Gamma*). For instance, *2Fleet-Logistic* represents a EM estimating two fishery fleets and assumes logistic selectivity for fleet 1 and fleet 2 (note that fleet 1 in multi-fleet models is always logistic). Conversely, *1Fleet-Block-Gamma* is a single fleet model that includes a time block to account for the fleet transition, where the selectivity after the fleet transition is parametrized with a gamma function (see Table 2 for all OM and EM scenarios and associated names).

Values for weight-at-age, maturity-at-age, natural mortality, steepness, the recruitment deviation parameter, and observation errors (i.e., index CV and ISS) were set to their true values to focus on the impacts of fleet structure and selectivity. The primary estimated parameters included: virgin recruitment, annual recruitment deviations, annual fishing mortality multipliers, selectivity parameters, and survey catchability. A description of specific EMs used in this study is provided below and in Table 2. In the following sections, references to logistic and gamma selectivity correspond to Eq. 1 and Eq. 2, respectively.

2.2.1 Multi-fleet models (2Fleet)

A total of two multi-fleet models were evaluated in this study, with differing parameterizations of time-invariant selectivity. Variants of multi-fleet models included the case where: 1) both fishery fleet 1 and fishery fleet 2 assumed logistic selectivity (*2Fleet-Logistic*), and 2) fishery fleet 1 assumed logistic selectivity, while fishery fleet 2 assumed gamma selectivity (*2Fleet-Gamma)*. Both models serve as a basis of comparison for when EM and OM structures align (i.e., correct assumptions regarding fleet structure, selectivity functional form, and time-variation) or provide context on the implications of mis-specifying selectivity when correctly accounting for fleet structure.

2.2.2 Single-fleet models (1Fleet)

2.2.2.1 Time-invariant (TimeInvar)

To understand the consequences of ignoring temporal changes in fleet structure and potential misspecification of selectivity forms, single-fleet EMs assuming time-invariant logistic selectivity (*1Fleet-TimeInvar-Logistic*) or time-invariant gamma selectivity (*1Fleet-TimeInvar-Gamma*) were explored.

2.2.2.2 Time block (Block)

Two single-fleet EMs with time blocked selectivity were used to evaluate the utility of time blocks in addressing temporal changes in fleet structure. Here, a total of two time blocks were specified. For both EMs, the first time block assumed logistic selectivity from the first year until the start of the fleet transition, years $t \in \{1,2,...24\}$ (Fig. 1A). Selectivity for the second time block was defined in years $t \in \{25,26,...T\}$, where T denotes the terminal year of the assessment period. Selectivity for the second time block was assumed to be either logistic selectivity (*1Fleet-Block-Logistic*) or gamma selectivity (*1Fleet-Block-Gamma*).

2.2.2.3 Random Walk (RandWlkPar)

In addition to discrete temporal changes in selectivity, EMs that allowed for continuous time-varying dynamics in selectivity parameters were also investigated. These EMs were implemented to evaluate if allowing selectivity to vary continuously as a parametric form performed better than simple time blocks when fishery fleets had distinct selectivity patterns (i.e., a logistic curve shifting towards a gamma curve). Separate EMs assuming either logistic (*1Fleet-RandWlkPar-Logistic*) or gamma selectivity (*1Fleet-RandWlkPar-Gamma*) were explored, and parameters for a given selectivity form varied as a random-walk over time (similar to Ianelli *et al.*, 2016):

Eq. 4

$$\omega_{t,s} = \begin{cases} \omega_{1,s} & t = 1 \\ \omega_{t-1,s} e^{\epsilon^{\omega}_{t,s}} & t > 1 \end{cases}$$

$$\epsilon^{\omega}_{t,s} \sim N(0, \sigma^{RW})$$

where $\omega_{t,s}$ represents a given selectivity parameter (i.e., $a^{50}_{t,s}$,$k_{t,s}$,$a^{max}_{t,s}$,$\gamma_{t,s}$), which were estimated as fixed-effect parameters in the first year. $\epsilon^{\omega}_{t,s}$ denotes annual deviations for a given selectivity parameter, which is governed by a normal distribution with mean 0 and standard deviation σ^{RW}. In this parameterization, all parameters defining a given selectivity form varied (e.g., $a^{50}_{t,s}$ and $k_{t,s}$ both varied in *1Fleet-RandWlkPar-Logistic*). Although σ^{RW} is theoretically estimable by integrating out $\epsilon^{\omega}_{t,s}$ using marginal maximum likelihood via Laplace Approximation

(Nielsen and Berg 2014; Kristensen et al. 2016), these values were subjectively tuned in this study. This was done to minimize the computational demands for this factorial simulation experiment. Briefly, we searched across a coarse range of values for σ^{RW} (i.e., 0.25 – 2.0) and selected a value that allowed for adequate fits to composition data without introducing unnecessary flexibility. We assumed the same σ^{RW} value for all selectivity parameters within a given EM to limit the range of values searched across. This resulted in σ^{RW} values of 1.25 and 2.0 being selected for *1Fleet-RandWlkPar-Logistic* and *1Fleet-RandWlkPar-Gamma,* respectively. Thus, deviations were estimated using penalized maximum likelihood. Preliminary investigations indicated that pre-specified values of σ^{RW} were comparable to those estimated using marginal maximum likelihood.

2.2.2.4 Semi-parametric (SemiPar)

Lastly, EMs assuming semi-parametric logistic (*1Fleet-SemiPar-Logistic*) or gamma (*1Fleet-SemiPar-Gamma*) selectivity were implemented in this study to understand the performance of EMs specified with a high degree of flexibility. While non-parametric time-varying selectivity allows for additional flexibility, initial explorations indicated that these models were not feasible in the current study, considering the number of ages represented within the model (n_{ages} = 30). Thus, semi-parametric EMs were pursued instead. Here, deviations were estimated across both ages and years, and were imposed on an assumed selectivity functional form:

Eq. 5

$$sel_{t+1,a,s,f} = s_{t,a,s,f}e^{\epsilon^{sel}_{t,a,s}}$$

$$\epsilon^{sel}_{t,a,s} \sim N(0, \sigma^{sel})$$

$$\text{nLL}_{\text{CurvaturePenalty}} = \sum_{a=2}^{n-1} \left(\Delta^2 sel_{t,a,s,f}\right)^2$$

where $\epsilon_{t,a,s}^{sel}$ are deviations about the assumed selectivity functional form governed by a normal distribution with mean 0 and standard deviation σ^{sel}. Thus, estimates of selectivity under this approach are able to exceed 1, but are constrained using a curvature penalty of squared second differences to provide regularity along the age axis (as is done in Ianelli *et al.*, 2016). To aid in model estimation and convergence, we assumed deviations were constant within age-blocks (i.e., with binning of every three ages for parsimony) given by the following:

Eq. 6

$$\epsilon_{t,a,s}^{sel} = \begin{cases} \epsilon_{t,1,s}^{sel} & \text{for } 1 \leq a \leq 3 \\ \epsilon_{t,4,s}^{sel} & \text{for } 4 \leq a \leq 6 \\ \vdots & \\ \epsilon_{t,28,s}^{sel} & \text{for } 28 \leq a \leq 30 \end{cases}$$

where $\epsilon_{t,a,s}^{sel}$ was defined in groups of three (10 age groups modelled), similar to the approach in Xu et al. (2019). As is the case for the *RandWlkPar* model variants, selectivity deviations were estimated using penalized maximum likelihood and σ^{sel} was subjectively tuned across a range of values (i.e., 0.25 – 2.0), which was assumed to be identical across ages and years for a given EM. A value of 0.75 for σ^{sel} was selected for both *1Fleet-SemiPar-Logistic* and *1Fleet-SemiPar-Gamma*.

While the EMs investigated in the current study are not an exhaustive list of possible model configurations, they represent a broad range of approaches that can be considered in contemporary stock assessment models. In particular, the choice of logistic and gamma functional forms in this study serves as an initial foundation for practitioners. Although more complex selectivity forms (e.g., double normal or double logistic; Methot and Wetzel, 2013) are viable options, these were not included for the purposes of brevity and to maintain comparability among EMs.

2.3 Sensitivity Analysis

2.3.1 Time Block Sensitivity

A sensitivity analysis was conducted to determine the implications of implementing alternate periods for when a time block occurs in the *1Fleet-Block* EMs. Sensitivities were performed where the time block was implemented in years other than year 25, which was when the fleet structure change begins. This sensitivity run also sought to evaluate the utility of AIC in identifying the correct time block specification. Here, time block EMs (*1Fleet-Block-Logistic* and *1Fleet-Block-Gamma*) were only applied to OM scenarios with a fast change in fleet structure (Fast-Logistic, Fast-Gamma-Old, and Fast-Gamma-Young). Incremental time blocks were implemented and tested in the EM in one-year increments ranging from five years prior to and five years after the start of the fleet transition (i.e., year 25). EMs were only applied to the conclusion of the fleet transition (year 30) and the terminal period (year 50) because the year that the fleet transition intersected (i.e., year 27) only contained a total of 2 years of data following the start of the transition. Additionally, when applying time block EMs to the shorter data time series (i.e., at the conclusion of the fleet transition in year 30), selectivity parameters associated with time blocks in the terminal year were not identifiable because only 1 year of data existed for the second time block. Therefore, the specified breakpoints were limited to years $t \in \{20,21\ldots29\}$. The sensitivity run involved the following steps:

1. A time block EM was parametrized with a discrete change in selectivity specified to occur in a given year $t \in \{20,21\ldots30\}$.
2. Each time block EM, with discrete changes defined to occur within a specific year *t*, was applied to the three OMs that exhibited a fast change in fleet structure.

3. This process was repeated for each of the 200 simulated datasets and for the two assessment periods.
4. Convergence rates, AIC values, and relative error in SSB were computed for each model run.
5. Comparisons of AIC and SSB across EMs and assessment periods were undertaken to determine whether the specification of a time block year had a large impact on model bias, and whether AIC was a reliable metric for identifying when a selectivity change might have occurred.

2.3.2 Survey Data Time-Series Sensitivity

Because the availability of survey data for the entire modeled time-series is often not realistic, we also performed a sensitivity test to evaluate the implications of only having survey data (i.e., abundance indices and age-composition data) available for the latter half of the time-series. In general, the sensitivity run followed the full-factorial experimental design discussed in previous sections, where each EM was applied to all OM scenarios. EMs were only applied to the terminal assessment period (Fast: year 50, Slow: year 70) and focused on the comparison of SSB estimates.

2.4 Evaluation of Model Performance

To evaluate model performance, only model runs that converged (i.e., positive definite Hessian matrix and a maximum absolute gradient < 0.001) were analyzed. Convergence rates were computed for each model run to assess tradeoffs between model complexity and the ability of EMs to achieve stable solutions. Metrics pertinent to management were calculated, which included the time series of spawning stock biomass (SSB) and the catch advice resulting from fishing at $F_{40\%}$ (Acceptable Biological Catch; ABC). $F_{40\%}$ is the fishing mortality rate that would result in 40% of unfished SSB-per-recruit (see Supplementary Material 1 for further details). Estimates of ABC and $F_{40\%}$ are management reference points that are commonly used in fisheries management and is the current management strategy for Alaska sablefish (Clark 2002; Goethel et al. 2023). Relative error (RE) was computed for ABC and SSB (denoted as θ):

Eq. 7

$$\mathrm{RE}_{\theta} = \frac{\theta_{est} - \theta_{truth}}{\theta_{truth}}$$

where RE_{θ} is the relative error for metric θ, θ_{est} is the estimated value for an EM, and θ_{truth} is the true value defined in the OM. RE_{θ} was then summarized by computing the median and its corresponding 95% simulation intervals.

AIC values were also computed for each EM run to determine the utility of AIC in detecting the correct selectivity form for multi-fleet models, and its ability to identify parsimonious EMs for single-fleet models, especially when a limited post-transition data time-series exists to adequately parameterize multi-fleet EMs. We compared AIC values within each assessment period and fleet transition scenario, as well as EMs with identical fleet structure assumptions, to ensure that comparisons were only made among EMs utilizing the same dataset. Finally, to determine which EM configuration was the most robust to different fishery dynamics (i.e., fleet structure and selectivity forms), we used the minimax method with SSB as the summary statistic (Punt et al. 2014b; McGilliard et al. 2015). Here, the Median Absolute Relative Error (MARE) of SSB across the estimated time-series was computed, and the maximum MARE for each EM within a given assessment period and across all OM scenarios was identified. The EM configuration with the smallest maximum MARE was considered the most robust model as it is likely to be the least biased across the range of uncertainties explored in the current study. Minimax solutions were compared across all EM configurations and OM scenarios within a given assessment period (i.e., to determine if the most robust model depended on the time series of data available).

3 Results

Overall, the EMs in the study demonstrated high convergence rates (mean = 98.7%; Fig. S1). Convergence rates were lower when EMs assumed deviations on logistic selectivity parameters (*1Fleet-RandWlkPar-Logistic*; mean = 90.4%), likely because of a complex likelihood surface and high correlations between selectivity parameters. Further investigations suggested that the *1Fleet-RandWlkPar-Logistic* model had positive definite Hessian matrices but were unable to reach a maximum absolute gradient < 0.001 without providing alternative starting values. In the subsequent sections, biases associated with SSB and ABC are discussed and will refer to the maximum median bias observed, unless stated otherwise. Furthermore, the following descriptors are used to characterize the range of absolute bias: small (< |10%|), moderate (> |10%| and < |20%|), and large (> |20%|).

3.1 Trends in Spawning Stock Biomass

In general, the magnitude and patterns of bias in SSB for OMs with fast or slow transitions in fishery fleet structure were consistent, but with some exceptions. The largest biases in SSB occurred during the terminal assessment period (~|40%|) in EMs that ignored changes in fleet structure by assuming a single-fleet with time-invariant selectivity (*1Fleet-TimeInvar-Logist* and *1Fleet-TimeInvar-Gamma*) (Fig. 2; fast transition, and Fig. 3; slow transition). For these EMs, moderate biases were detected during the conclusion of the fleet transition, while minimal biases were observed during the earlier fleet intersection period. Furthermore, biases in SSB typically peaked following the change in fleet structure and trended less biased towards the end of the time series (Fig. 2 and Fig. 3). Positive biases in SSB developed when single-fleet EMs assuming time-invariant selectivity were applied to OMs where fishery removals resulted from logistic selectivity only (Fast-Logistic and Slow-Logistic) or from an old-selecting gamma curve (Fast-Gamma-Old and Slow-Gamma-Old). Conversely, negative biases in SSB developed in the OM with a young-selecting gamma curve (Fast-Gamma-Young and Slow-Gamma-Young).

When fast changes in fleet structure occurred (Fig. 2), assuming time blocked selectivity (*1Fleet-Block-Logistic* and *1Fleet-Block-Gamma)* reduced biases in SSB compared to time-invariant selectivity EMs (< |10%|). Biases in SSB were relatively low across all assessment periods, although slightly larger biases were observed when time block EMs were applied toward the conclusion of the fleet transition. However, when the true fishery removals were represented by a young-selecting gamma curve, moderate negative biases (~ -20%) persisted in the EM approach that assumed logistic selectivity across both time blocks (*1Fleet-Block-Logistic*), especially when applied to the terminal assessment period (Fast-Gamma-Young; Fig. 2). Conversely, under slow changes in fleet structure, time block EMs exhibited moderate biases (~|25%|), which were detected for assessment periods occurring at the conclusion of the fleet transition and the terminal assessment periods (Fig. 3).

EMs assuming deviations on selectivity parameters (*1Fleet-RandWlkPar-Logistic* and *1Fleet-RandWlkPar-Gamma*) generally demonstrated small biases (< |5%| bias) across OMs. However, under both fast and slow fleet structure changes, moderate negative biases developed

(~ -15%) when these EMs assumed a selectivity form that largely differed (e.g., *1Fleet-RandWlkPar-Logistic*) from the OM (e.g., Fast-Gamma-Young and Slow-Gamma-Young; Fig. 2 and Fig. 3). By contrast, assuming semi-parametric selectivity (i.e., the *1Fleet-SemiPar-Logistic* and *1Fleet-SemiPar-Gamma* EMs) exhibited consistently low bias in SSB for all OMs (<5%; Fig. 2 and Fig. 3).

As expected, EMs with correctly specified fleet structure and selectivity (i.e., the *2Fleet-Logistic* and *2Fleet-Gamma* EMs when applied to OMs with matching selectivity assumptions for fleet 2 demonstrated the least bias across all OMs, with consistent results across assessment periods. While multi-fleet EMs that mis-specified selectivity for the second fleet performed well for some scenarios (i.e., <|5%| bias), consistent negative biases were detected when the assumed EM selectivity was mis-specified for the second fleet (e.g., the *2Fleet-Logistic* EM applied to data simulated from the Fast-Gamma-Young and Slow-Gamma-Young OMs; ~ -15% bias; Fig. 2 and Fig. 3).

<u>*3.2 Management Reference Points*</u>

Overall, the magnitude and pattern of bias in ABC remained consistent across OMs simulated with either a fast or slow change in fleet structure, albeit with a few exceptions. For both scenarios of fleet structure change, biases in ABC were large when single-fleet time-invariant EMs (*1Fleet-TimeInvar-Logistic* and *1Fleet-TimeInvar-Gamma*) were applied during the terminal assessment period (~|20 - 35%| bias) but were smaller in magnitude if applied shortly after the change in fleet structure (i.e., during the fleet transition period, Fig. 4; fast transition, and Fig. 5; slow transition).

Although time block EMs (*1Fleet-Block-Logistic* and *1Fleet-Block-Gamma*) reduced biases relative to single-fleet time-invariant EMs, biases in ABC were larger when time block EMs were confronted with slow changes in fleet structure (Fast: ~|8%|; Slow: ~|15%|; Fig. 4 and Fig. 5, respectively). EMs assuming continuous time-varying selectivity also performed better than time block EMs across most OM scenarios and assessment periods (< |13%| bias), with relatively small differences in median bias for ABC between EMs assuming deviations on selectivity parameters (*1Fleet-RandWlkPar-Logistic* and *1Fleet-RandWlkPar-Gamma*) or semi-parametric selectivity (*1Fleet-SemiPar-Logistic* and *1Fleet-SemiPar-Gamma*; Fig. 4 and 5).

Lastly, biases in ABC were negligible (~0%) when both fleet structure and selectivity were correctly specified for both fast and slow changes in fleet structure (Fig. 4 and Fig. 5). Generally, multi-fleet EMs with mis-specified selectivity resulted in minimal biases in ABC across assessment periods (~|5%|). However, assuming logistic selectivity for both fleets (*2Fleet-Logistic*) for OMs with strong dome-shaped selectivity (Fast-Gamma-Young and Slow-Gamma-Young) often resulted in negative biases (~ -13%; Fig. 4 and Fig. 5).

<u>*3.3 Model selection using AIC*</u>

AIC consistently detected the correct multi-fleet EM, with mean differences in AIC exceeding 100 units between the correct and incorrect EM (Fig. S2). For single-fleet EMs, AIC preferred EMs assuming time-invariant logistic selectivity during the fleet intersection period, and time blocked EMs during the fleet transition or the terminal assessment periods (Fig. S3). However, AIC-based model selection exhibited variable performance in identifying an appropriate selectivity functional form for single-fleet EMs. Importantly, AIC did not consider continuous time-varying EMs to be parsimonious, despite demonstrating minimal bias in derived quantities across OMs. This is not surprising given that continuous time-varying EMs typically estimated 300 – 1500 parameters, whereas time block or time-invariant selectivity EMs estimated 100 – 200 parameters, when utilized during the terminal period.

3.4 Time Block Sensitivity Analysis

Generally, time blocks employed 2 - 3 years (i.e., years 27 - 28) after the initial transition in fleet structure (i.e., year 25) were preferred by AIC (Fig. S4) and resulted in reduced bias in SSB (Fig. S5 and Fig. S6). However, specifying time blocks prior to the change in fleet structure (i.e., before year 25), demonstrated increasing levels of bias in SSB. The use of AIC in selecting time blocks was more variable during the conclusion of the fleet transition but had increased precision in selecting the correct transition timing when used during the terminal period (Fig. S4), suggesting that an extended data time series may facilitate the identification of appropriate breakpoints.

3.5 Survey Data Time-Series Sensitivity Analysis

When survey data were only available for the latter half of the time-series, the magnitude of biases in SSB increased, relative to those observed in the primary analyses (Fig. S7, S8). Although patterns of bias in SSB generally remained consistent with those previously described, *1Fleet-RandWlkPar-Gamma* was an exception, demonstrating comparatively poorer model performance. Specifically, when applied to most OM scenarios, large positive biases were detected (~ +25%) during the beginning of the time-series (Fig. S7, S8), which were not originally observed in the primary analyses. Additionally, when fishery removals in OMs shifted from a logistic selectivity curve into an old-selecting gamma curve (Fast-Gamma-Old and Slow-

Gamma-Old), large positive biases were also detected towards the terminal year of the assessment period (Fig. S7, S8).

3.6 Minimax Solution

The multi-fleet EM assuming gamma selectivity (*2Fleet-Gamma*) proved to be the most robust across the different rates of change in fleet structure, selectivity parametrizations, and assessment periods that were explored. Here, *2Fleet-Gamma* had the lowest value of maximum MARE in SSB across all OM scenarios and assessment periods (< 3; bolded in Table 3).

4 Discussion

Across the scenarios explored in this study, ignoring changes in fleet structure by assuming a single-fleet with time-invariant selectivity led to substantial biases in management quantities. Thus, assuming a single-fleet model with time-invariant selectivity when changes in fleet structure have occurred is inadequate and alternative approaches to account for such changes are warranted. The implementation of selectivity time blocks improved model performance over time-invariant selectivity models but were only adequate to address fast changes in fishery fleet structure, and generally depended on the assumed post-transition selectivity form. Specifically, biases were reduced only when time blocks were specified to occur after the start of the fleet transition.

Models assuming continuous time-varying selectivity generally performed well across both fast and slow changes in fleet structure, although their performance sometimes depended on the selectivity form assumed, as well as the availability of survey data. Despite continuous time-varying selectivity models often demonstrating minimal bias in most scenarios, these models were seldom considered parsimonious when using AIC-based model selection for single-fleet models. This likely occurred because continuous time-varying EMs estimated up to 1000 parameters and marginal AIC was applied under a penalized maximum likelihood framework (Maunder and Harley, 2011; Punt *et al.*, 2014; Privitera-Johnson *et al.*, 2022), leading to an overestimation of the number of effective parameters. Formulating these models under a state-space framework may produce different outcomes (Nielsen and Berg, 2014; Stock and Miller, 2021), but were not attempted given computational demands, and should be a future area of research.

Multi-fleet models also proved effective in addressing changes in fleet structure. Moreover, the use of AIC in selecting among alternative selectivity forms appeared reliable for multi-fleet models, wherein the correct selectivity form was always selected as the most parsimonious (Fig. S2). In general, multi-fleet structures performed reasonably, even with misspecification of selectivity forms and may serve as a promising approach for practitioners to explore if sufficient fleet-specific compositional data are available. Our results indicate that, given parsimony-complexity tradeoffs and data limitations as new fishery fleets develop, single-fleet models with time-varying effects are adequate for operational management advice when confronted with fleet transitions. However, research oriented multi-fleet models should be used as a validation tool to explore consistency in population trends across alternative model structures.

4.1 Interpreting Bias Trends

Across various single-fleet EMs, a consistent pattern emerged where biases in SSB were generally small during the beginning of the modeled time-series, peaked prior to the fleet transition, and became less pronounced towards the terminal assessment period. To illustrate these biases, we consider the application of single-fleet EMs assuming time-invariant selectivity (i.e., EMs *1Fleet-TimeInvar-Logist* and *1Fleet-TimeInvar-Gamma*) under the Fast-Logistic OM scenario (Fig. 6).

Towards the beginning of the time-series and prior to the fleet transition, estimated selectivities in single-fleet time-invariant EMs favored the capture of older individuals over younger individuals, deviating from the true simulated selectivity form (Fig. 6). This divergence likely stemmed from estimated selectivities being a compromise to represent data from the two distinct fishery fleets, manifesting as a weighted average between them. Consequently, the assumed reduced capture of younger individuals led to their accumulation within the estimated population, resulting in a positive bias in SSB estimates. Given divergences in estimated selectivities, fishing mortality multipliers were concomitantly overestimated to adequately fit to the observed catch data (Fig. 6). Biases in SSB were presumably minimal during the initial period, due to the relatively low weight-at-age of young individuals and their consequently minor contribution to SSB. Following the fleet transition, estimated selectivities incorrectly exhibited an increased preference towards removal of younger individuals, depleting the accumulation of individuals from the previous period, and precipitating a decreasing trend in SSB bias over time in the absence of those individuals contributing to the spawning population. To reconcile an increased selection of younger individuals with observed catch data, fishing mortality multipliers were underestimated as a result (Fig. 6). Although the underestimation of fishing mortality led to an increasing bias for older individuals, their contribution to SSB was minimal, given their low abundance within the population. Similar trends in SSB biases were observed in single-fleet EMs applied to OMs characterized by extreme dome-shaped selectivity (e.g., Gamma-Young), albeit with biases that were in the opposite direction (i.e., initial negative bias, followed by decreasing bias; Fig. 2 and 3). The mechanisms underlying these patterns resemble those in the example described above, except that selectivities initially favored younger individuals, followed by a preference towards older individuals during the post-transition period (Fig. S9, S10, S11, and S12). The biases described in Figure 6 are generally specific to the selectivity and fishing mortality scenarios evaluated. In particular, because composition data were a catch-weighted average of the two fishery fleets, and catches were generally higher before the fleet transition, the estimated time-invariant fishery selectivities better resembled the population selectivity curves from the pre-transition period (Fig. S9, S10, S11, and S12).

In most scenarios and assessment periods, the bias trends described were consistent across EMs, but were greatly reduced as flexibility in selectivity parametrization increased (e.g., by introducing time blocks, continuous parametrizations, or allowing for multiple fleets). However, there were some exceptions to these trends. In particular, biases for most EMs applied during the fleet-intersection period were negligible, presumably due to the incomplete transition of the true simulated selectivity towards the second fleet, and the available data predominately reflecting fishery dynamics prior to the fleet transition. Moreover, multi-fleet models with both fleets assuming logistic selectivity, consistently exhibited negative biases in SSB when applied to OMs characterized by strong dome-shaped selectivity. This was likely attributed to increased removals of intermediate to older-aged individuals, despite not being removed in the OM. Exceptions to the trends observed in the primary analyses also arose when survey data were only available for the latter half of the modeled time-series, where we detected large biases in single-fleet models that assumed deviations on gamma selectivity parameters (EM: *1Fleet-RandWlkPar-Gamma*; Fig. S7, S8). These biases likely manifested from reductions in survey data, which provided information on population age-structure. Consequently, fishery age-composition data were likely overfitted, resulting in a poorly estimated descending limb of the selectivity curve, which could have otherwise been better informed in the presence of informative survey data.

4.2 Pragmatic Recommendations for Addressing Fleet Structure Transitions

Complications can arise in single fleet models when the fraction of catch from different fleets changes over time and if the selectivity of these fleets differs (i.e., changes in fleet structure occur), which can manifest in complex time-varying selectivity patterns (e.g., Lee et al. 2017). Herein, we provide considerations for parametrizing stock assessment models when confronted with changes in fleet structure or removal patterns (Fig. 7). We preface these recommendations with the caveat that they are generally specific to data-rich fisheries with long data time-series. However, we also provide some guidance for fisheries that may be more data-moderate. Firstly, we recommend practitioners begin by defining the maximum number of fishery fleets as proposed by Punt et al. (2014) (Fig. 7). This process can involve defining fleets as different gears, areas, or seasons to the finest resolution feasible, and will likely depend on the characteristics of the fishery. In a spatial context, this can be done using multivariate regression trees (Lennert-Cody et al. 2010, 2013). Hypotheses of plausible fleet-specific selectivity forms and the timing of changes in fleet structure should then be developed using *a priori* knowledge of fishery dynamics and communicating with stakeholders. Justification for these hypotheses should explicitly consider processes governing contact selectivity and availability.
Concomitantly, a thorough analysis of composition data should be conducted to explore differences among candidate fleets and to identify locations (i.e., geospatial and depth strata) in

which samples were collected. Given that access to compositional data can sometimes be limited, sample sizes and data quality should be evaluated to identify whether there is sufficient information to support the development of a multi-fleet model. In the case where data are insufficient to support multi-fleet EMs, single-fleet EMs should be pursued and hypotheses regarding fleet-specific selectivities previously formulated can be used to infer appropriate selectivity parameterizations. If sufficient data exists, multi-fleet models should be implemented to represent removal processes from the fishery. Considering that the use of a multi-fleet model was the most robust in this study (Table 3) and AIC-based model selection consistently detected the correct functional form for the selectivity curve, these models can serve as a valuable starting point in ensuring removal processes are adequately represented. Moreover, multi-fleet models can help validate subsequent single-fleet assessment models (Nielsen et al. 2021; Cheng et al. 2024). We further recommend analysts employ traditional model diagnostics (e.g., residual analysis and likelihood profiles; Carvalho et al. 2017, 2021; Trijoulet et al. 2023) in tandem with previously developed hypotheses on selectivity forms to determine biologically plausible models (Hulson and Hanselman 2014; Punt et al. 2020; Carvalho et al. 2021; Privitera-Johnson et al. 2022). Residual diagnostics can be particularly useful in this context, given that the presence of systematic patterns across ages could indicate a mis-specified selectivity form, while patterns across years or cohorts could suggest the need to consider time-varying selectivity.

Under slow shifts in fishery fleet structure, our simulation study indicated that both multi-fleet models and single-fleet models with time-varying selectivity performed reasonably, consistent with findings from Nielsen et al. (2021). Using flexible time-varying approaches (e.g., non-parametric or semi-parametric) will likely achieve adequate model performance in most scenarios, although multi-fleet models without time variation in selectivity can potentially be more parsimonious in some cases (i.e., if process deviation parameters are treated as independent). Results from our study also indicated that time-varying selectivity assuming deviations on parameters was only appropriate when fleet-selectivities were similar (e.g., fleets have the same functional form) and should be implemented with caution. Furthermore, when employing continuous selectivity approaches, additional care is warranted to ensure the biological plausibility of estimated selectivities, especially in data-moderate situations. This was evident in our sensitivity analyses, which demonstrated that, when survey data were only available for part of the modeled time-series, continuous time-varying selectivity approaches constrained to dome-shaped forms could overfit age-composition data and degrade fits to other data sources (Martell and Stewart 2014; Punt 2023). Therefore, in data-moderate contexts (such as when limited survey data are available) where gradual changes in fishery fleet structure are expected and a single-fleet model is pursued, it may be practical to assume asymptotic rather than dome-shaped time-varying selectivity to avoid overfitting data. However, it is important to explicitly recognize that model results are likely to be biased towards low biomass estimates (Privitera-Johnson et al. 2022). Similarly, given that modelling time-variation on an incorrect process or the estimation of implausible time-varying selectivity forms can lead to the provision of suboptimal management advice (Szuwalski et al. 2018; Fisch et al. 2023; Cheng et al. 2024), we emphasize the need to further consider *a priori* knowledge of fishery dynamics when implementing flexible time-varying selectivity approaches. We suggest that these models be validated against estimates from multi-fleet models when possible, assuming that selectivity from multi-fleet models is adequately characterized.

For fast fleet transitions, we similarly found that multi-fleet models and single-fleet models assuming time-varying semi-parametric selectivity demonstrated minimal bias. We also found that time block approaches were appropriate in addressing fast changes in fleet structure, similar to findings from Cheng et al. (2024). However, time block models did not perform well when selectivity assumptions largely diverged from the simulated truth (i.e., assuming logistic, but selectivity was strongly dome-shaped), underscoring the sensitivity to assumed selectivities for this approach. Therefore, when fast shifts in fishery fleet structure are present (e.g., regulatory change or adoption of a new gear), we recommend that practitioners implement time blocked selectivity, following the selectivity forms identified for multi-fleet models. However, in data-moderate scenarios, the development of multi-fleet models may not be supported, and it may be necessary to proceed directly with a time blocked single-fleet model. Here, previously developed hypotheses about fleet-specific characteristics can similarly be useful for guiding appropriate parametrizations of time blocked selectivity within the context of a single-fleet model. The breakpoints defined for time blocks should then be evaluated across a range of plausible periods using model selection tools to determine optimal breakpoints (typically several years after a change in fleet structure is suspected) (Fig. 7). While multi-fleet models and single-fleet models coupled with flexible time-varying selectivity parameterizations are also plausible under such circumstances, time block approaches are likely more practical in data-moderate scenarios. Additionally, time blocked selectivity approaches can potentially be more parsimonious in some applications, enabling practitioners to explore other unmodelled dimensions that are influential to population dynamics (e.g., sex, time, and age-varying natural mortality, time-varying growth; Deroba and Schueller, 2013; Johnson *et al.*, 2015; Correa *et al.*, 2021). However, it should also be noted that discrete time blocked parametrizations will require frequent and repeated re-evaluation of blocking assumptions if fleet structure continues to change over time.

4.3 Caveats and Future Work

Like many simulation studies, aspects of this study were limited and could be expanded upon in future studies. First, several parameters were set at their true values, which may lead to overly optimistic model performance (e.g., natural mortality, steepness). Given that natural mortality and dome-shaped selectivity are confounded, it would be of interest to assess model performance when natural mortality is simultaneously estimated with dome-shaped selectivity (Thompson, 1994; Clark, 1999). Additionally, the current study only evaluated the life-history characteristics of Alaska sablefish and future studies could extend this work by incorporating additional life-histories. We acknowledge the simplicity of selectivity forms used in this study, which were also specified to be time-invariant. The use of simple selectivity forms in this study likely led to optimistic performance of EMs assuming dome-shaped selectivity detected in our primary results. These EMs often demonstrated minimal bias, even when the true removal patterns were represented by asymptotic selectivity. While there is a general expectation of overestimating biomass (i.e., through the development of cryptic biomass) when selectivity is mis-specified to be dome-shaped (Cadrin et al. 2016), these biases were not detected in our primary analyses. Presumably, this is attributed to the relatively simple selectivity forms utilized in the data-generating process, the presence of informative survey data, and how OM selectivities were conditioned (i.e., both fishery fleet 1 and the survey fleet exhibited logistic selectivity). Indeed, through limited sensitivity analyses, we found that informative data on population age-structure from the survey fleet was necessary to mitigate the effects of selectivity misspecification, allowing EMs that incorrectly assumed dome-shaped selectivity to perform well. As such, we caution against overinterpreting the optimistic performance of EMs assuming dome-shaped selectivity in this study. Furthermore, our investigations were also limited in the number of fishery fleets evaluated. The incorporation of additional fishery fleets, particularly in the context of combining gears with limited catches, or using fleets to represent spatial dynamics (e.g., closed areas, seasonal and/or ontogenetic migrations) could be a fruitful avenue for future research (e.g., expanding on the work of Lee et al., 2017). Lastly, we recognize the use of a multinomial distribution to simulate and fit composition data may not fully capture the complexities of real-world sampling variability. While we did not examine the influence of alternative compositional likelihoods in this study, prior research suggests that the Dirichlet-Multinomial distribution may be more suitable within the context of estimating time-varying selectivity (Xu et al. 2020). Given that composition data often exhibit positive correlations and overdispersion, which are not adequately captured by a multinomial distribution, the findings of this study likely represent a best-case scenario (Francis 2014).

4.4 Conclusions

Ignoring changes in fleet structure or emerging fleets may result in inadequate management advice, while data limitations can hinder implementation of multi-fleet models. Models of intermediate complexity (e.g., time blocked or continuous time-varying selectivity models), complemented by research-oriented multi-fleet models are likely suitable for most applications. However, other considerations may necessitate the use of multi-fleet models. For instance, certain management frameworks may require advice on fleet-specific catch or it may be important to monitor spatial and fleet-specific discard patterns and harvester behaviors (Marchal 2002; Branch et al. 2006; Eigaard et al. 2011). Within the context of developing closed-loop feedback control systems (i.e., management strategy evaluations), multi-fleet models enable the exploration of fleet-specific behavioral responses (Van Putten et al. 2012) and allow harvesters to consider performance measures that are tailored to their needs, which may contribute to the development of more robust management procedures (Bastardie et al. 2010b, 2010a; Fernández et al. 2010; Pascoe et al. 2010; Nielsen et al. 2021). Importantly, multi-fleet models may better represent removal patterns as observed by harvesters, which can foster stakeholder trust and engagement in the fishery management process. Ultimately, the exploration of multi-fleet models, whether in operational or research-oriented contexts, is likely valuable in guiding informed decision-making within the fisheries management process.

Acknowledgements

We thank Kimberly Fitzpatrick, Genoa Sullaway, Maia Kapur, Franz Mueter, the Associate Editor for CJFAS, and two anonymous reviewers for their thoughtful comments, which helped improve the manuscript. MC was supported by the National Science Foundation Graduate Research Fellowship during the time of this study. Figures 6 and 7 were created with BioRender.com. Findings and conclusions of this study are those of the authors, and do not necessarily represent the views of the National Marine Fisheries Service, NOAA.

Funding Statement

This research was supported by the National Science Foundation Graduate Research Fellowship Program (Grant # 2235201).

Competing Interests Statement

The authors declare there are no competing interests.

Data Availability Statement

Code required to generate simulated data utilized in this study can be found at https://github.com/chengmatt/Fleet_Selex_Sim.

References

Aarts, G., and Poos, J.J. 2009. Comprehensive discard reconstruction and abundance estimation using flexible selectivity functions. ICES Journal of Marine Science **66**(4): 763–771. doi:10.1093/icesjms/fsp033.

Bastardie, F., Nielsen, J.R., and Kraus, G. 2010a. The eastern Baltic cod fishery: a fleet-based management strategy evaluation framework to assess the cod recovery plan of 2008. ICES Journal of Marine Science **67**(1): 71–86. doi:10.1093/icesjms/fsp228.

Bastardie, F., Vinther, M., Nielsen, J.R., Ulrich, C., and Paulsen, M.S. 2010b. Stock-based vs. fleet-based evaluation of the multi-annual management plan for the cod stocks in the Baltic Sea. Fisheries Research **101**(3): 188–202. doi:10.1016/j.fishres.2009.10.009.

Beare, D., Rijnsdorp, A.D., Blaesberg, M., Damm, U., Egekvist, J., Fock, H., Kloppmann, M., Röckmann, C., Schroeder, A., Schulze, T., Tulp, I., Ulrich, C., Van Hal, R., Van Kooten, T., and Verweij, M. 2013. Evaluating the effect of fishery closures: Lessons learnt from the Plaice Box. Journal of Sea Research **84**: 49–60. doi:10.1016/j.seares.2013.04.002.

Berger, A.M., Jones, M.L., Zhao, Y., and Bence, J.R. 2012. Accounting for spatial population structure at scales relevant to life history improves stock assessment: The case for Lake Erie walleye Sander vitreus. Fisheries Research **115–116**: 44–59. doi:10.1016/j.fishres.2011.11.006.

Beverton, R., and Holt, S. 1957. On the Dynamics of Exploited Fish Populations. Min. Agric. Fish and Food U.K. Fish Invest. Ser., London.

Bohaboy, E.C., Goethel, D.R., Cass-Calay, S.L., and Patterson, W.F. 2022. A simulation framework to assess management trade-offs associated with recreational harvest slots, discard mortality reduction, and bycatch accountability in a multi-sector fishery. Fisheries Research **250**: 106268. doi:10.1016/j.fishres.2022.106268.

Bosley, K.M., Schueller, A.M., Goethel, D.R., Hanselman, D.H., Fenske, K.H., Berger, A.M., Deroba, J.J., and Langseth, B.J. 2022. Finding the perfect mismatch: Evaluating misspecification of population structure within spatially explicit integrated population models. Fish and Fisheries **23**(2): 294–315. doi:10.1111/faf.12616.

Branch, T.A., Hilborn, R., Haynie, A.C., Fay, G., Flynn, L., Griffiths, J., Marshall, K.N., Randall, J.K., Scheuerell, J.M., Ward, E.J., and Young, M. 2006. Fleet dynamics and fishermen behavior: lessons for fisheries managers. Can. J. Fish. Aquat. Sci. **63**(7): 1647–1668. doi:10.1139/f06-072.

Brunel, T., and Piet, G.J. 2013. Is age structure a relevant criterion for the health of fish stocks? ICES Journal of Marine Science **70**(2): 270–283. doi:10.1093/icesjms/fss184.

Cadrin, S.X., DeCelles, G.R., and Reid, D. 2016. Informing fishery assessment and management with field observations of selectivity and efficiency. Fisheries Research **184**: 9–17. doi:10.1016/j.fishres.2015.08.027.

Carvalho, F., Punt, A.E., Chang, Y.-J., Maunder, M.N., and Piner, K.R. 2017. Can diagnostic tests help identify model misspecification in integrated stock assessments? Fisheries Research **192**: 28–40. doi:10.1016/j.fishres.2016.09.018.

Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripa, M., Kitakado, T., Yemane, D., Piner, K.R., Maunder, M.N., Taylor, I., Wetzel, C.R., Doering, K., Johnson, K.F., and Methot, R.D. 2021. A cookbook for using model diagnostics in integrated stock assessments. Fisheries Research **240**: 105959. doi:10.1016/j.fishres.2021.105959.

Cheng, M.L.H., Goethel, D.R., and Cunningham, C.J. 2024. Addressing complex fleet structure in fishery stock assessment models: Accounting for a rapidly developing pot fishery for Alaska sablefish (Anoplopoma fimbria). Fisheries Research **271**: 106917. doi:10.1016/j.fishres.2023.106917.

Clark, W.G. 1999. Effects of an erroneous natural mortality rate on a simple age-structured stock assessment. Canadian Journal of Fisheries and Aquatic Sciences **56**(10): 1721–1731. doi:doi:10.1139/f99-085.

Clark, W.G. 2002. $F_{35\%}$ Revisited Ten Years Later. North American Journal of Fisheries Management **22**(1): 251–257. doi:10.1577/1548-8675(2002)022<0251:FRTYL>2.0.CO;2.

Cope, J.M., and Punt, A.E. 2011. Reconciling stock assessment and management scales under conditions of spatially varying catch histories. Fisheries Research **107**(1–3): 22–38. doi:10.1016/j.fishres.2010.10.002.

Correa, G.M., McGilliard, C.R., Ciannelli, L., and Fuentes, C. 2021. Spatial and temporal variability in somatic growth in fisheries stock assessment models: evaluating the consequences of misspecification. ICES Journal of Marine Science **78**(5): 1900–1908. doi:10.1093/icesjms/fsab096.

Deroba, J.J., and Schueller, A.M. 2013. Performance of stock assessments with misspecified age- and time-varying natural mortality. Fisheries Research **146**: 27–40. doi:10.1016/j.fishres.2013.03.015.

Eigaard, O.R., Marchal, P., Gislason, H., and Rijnsdorp, A.D. 2014. Technological Development and Fisheries Management. Reviews in Fisheries Science & Aquaculture **22**(2): 156–174. doi:10.1080/23308249.2014.899557.

Eigaard, O.R., Thomsen, B., Hovgaard, H., Nielsen, A., and Rijnsdorp, A.D. 2011. Fishing power increases from technological development in the Faroe Islands longline fishery. Can. J. Fish. Aquat. Sci. **68**(11): 1970–1982. doi:10.1139/f2011-103.

Fernández, C., Cerviño, S., Pérez, N., and Jardim, E. 2010. Stock assessment and projections incorporating discard estimates in some years: an application to the hake stock in ICES Divisions VIIIc and IXa. ICES Journal of Marine Science **67**(6): 1185–1197. doi:10.1093/icesjms/fsq029.

Fisch, N., Shertzer, K., Camp, E., Maunder, M., and Ahrens, R. 2023. Process and sampling variance within fisheries stock assessment models: estimability, likelihood choice, and the consequences of incorrect specification. ICES Journal of Marine Science **80**(8): 2125–2149. doi:10.1093/icesjms/fsad138.

Fournier, D., and Archibald, C.P. 1982. A General Theory for Analyzing Catch at Age Data. Can. J. Fish. Aquat. Sci. **39**(8): 1195–1207. doi:10.1139/f82-157.

Francis, R.I.C.C. 1992. Use of Risk Analysis to Assess Fishery Management Strategies: A Case Study using Orange Roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. Can. J. Fish. Aquat. Sci. **49**(5): 922–930. doi:10.1139/f92-102.

Francis, R.I.C.C. 2014. Replacing the multinomial in stock assessment models: A first step. Fisheries Research **151**: 70–84. doi:10.1016/j.fishres.2013.12.015.

Gilman, E., Chaloupka, M., Read, A., Dalzell, P., Holetschek, J., and Curtice, C. 2012. Hawaii longline tuna fishery temporal trends in standardized catch rates and length distributions and effects on pelagic and seamount ecosystems. Aquatic Conservation **22**(4): 446–488. doi:10.1002/aqc.2237.

Goethel, D., Hanselman, D., Rodgveller, C., Echave, K.B., Williams, B., Shotwell, S.K., Sullivan, J., Hulson, P., Malecha, P., and Siwicke, K. 2021. 3. Assessment of the Sablefish Stock in Alaska. : 347.

Goethel, D.R., Cheng, M.L.H., Echave, K.B., Marsh, C., Rodgveller, C.J., and Siwicke, K. 2023. 3. Assessment of the Sablefish Stock in Alaska. North Pacific Fishery Management Council, Anchorage, AK.

Goethel, D.R., Rodgveller, C.J., Echave, K.B., Shotwell, S.K., Siwicke, K.A., Malecha, P.W., Cheng, M., Williams, M., Omori, K., and Lunsford, C.R. 2022. Assessment of the Sablefish Stock in Alaska. : 182.

Hulson, P.-J.F., and Hanselman, D.H. 2014. Tradeoffs between bias, robustness, and common sense when choosing selectivity forms. Fisheries Research **158**: 63–73. doi:10.1016/j.fishres.2013.12.016.

Hurtado-Ferro, F., Punt, A.E., and Hill, K.T. 2014. Use of multiple selectivity patterns as a proxy for spatial structure. Fisheries Research **158**: 102–115. doi:10.1016/j.fishres.2013.10.001.

Ianelli, J., Honkalehto, T., Barbeaux, S., Fissel, B., and Kotwicki, S. 2016. Assessment of the walleye pollock stock in the Eastern Bering Sea. Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska. North Pacific Fishery Mngt. Council: 126.

Johnson, K.F., Monnahan, C.C., McGilliard, C.R., Vert-pre, K.A., Anderson, S.C., Cunningham, C.J., Hurtado-Ferro, F., Licandeo, R.R., Muradian, M.L., Ono, K., Szuwalski, C.S., Valero, J.L., Whitten, A.R., and Punt, A.E. 2015. Time-varying natural mortality in fisheries stock assessment models: identifying a default approach. ICES Journal of Marine Science **72**(1): 137–150. doi:10.1093/icesjms/fsu055.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., and Bell, B. 2016. TMB: Automatic Differentiation and Laplace Approximation. J. Stat. Soft. **70**(5). doi:10.18637/jss.v070.i05.

Lee, H.H., Thomas, L.R., Piner, K.R., and Maunder, M.N. 2017. Effects of age-based movement on the estimation of growth assuming random-at-age or random-at-length data: age-based movement and estimation of growth. J Fish Biol **90**(1): 222–235. doi:10.1111/jfb.13177.

Lennert-Cody, C.E., Maunder, M.N., Aires-da-Silva, A., and Minami, M. 2013. Defining population spatial units: Simultaneous analysis of frequency distributions and time series. Fisheries Research **139**: 85–92. doi:10.1016/j.fishres.2012.10.001.

Lennert-Cody, C.E., Minami, M., Tomlinson, P.K., and Maunder, M.N. 2010. Exploratory analysis of spatial–temporal patterns in length–frequency data: An example of distributional regression trees. Fisheries Research **102**(3): 323–326. doi:10.1016/j.fishres.2009.11.014.

Linton, B.C., and Bence, J.R. 2011. Catch-at-age assessment in the face of time-varying selectivity. ICES Journal of Marine Science **68**(3): 611–625. doi:10.1093/icesjms/fsq173.

Marchal, P. 2002. Area-based management and fishing efficiency. Aquatic Living Resources **15**(2): 73–85. doi:10.1016/S0990-7440(02)01157-9.

Martell, S., and Stewart, I. 2014. Towards defining good practices for modeling time-varying selectivity. Fisheries Research **158**: 84–95. doi:10.1016/j.fishres.2013.11.001.

Maunder, M.N., and Harley, S.J. 2011. Using cross validation model selection to determine the shape of nonparametric selectivity curves in fisheries stock assessment models. Fisheries Research **110**(2): 283–288. doi:10.1016/j.fishres.2011.04.017.

Maunder, M.N., and Piner, K.R. 2015. Contemporary fisheries stock assessment: many issues still remain. ICES Journal of Marine Science **72**(1): 7–18. doi:10.1093/icesjms/fsu015.

Maunder, M.N., and Punt, A.E. 2013. A review of integrated analysis in fisheries stock assessment. Fisheries Research **142**: 61–74. doi:10.1016/j.fishres.2012.07.025.

McGilliard, C.R., Punt, A.E., Methot, R.D., and Hilborn, R. 2015. Accounting for marine reserves using spatial stock assessments. Can. J. Fish. Aquat. Sci. **72**(2): 262–280. doi:10.1139/cjfas-2013-0364.

Methot, R.D., and Wetzel, C.R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research **142**: 86–99. doi:10.1016/j.fishres.2012.10.012.

Nielsen, A., and Berg, C.W. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research **158**: 96–101. doi:10.1016/j.fishres.2014.01.014.

Nielsen, A., Hintzen, N.T., Mosegaard, H., Trijoulet, V., and Berg, C.W. 2021. Multi-fleet state-space assessment model strengthens confidence in single-fleet SAM and provides fleet-specific forecast options. ICES Journal of Marine Science **78**(6): 2043–2052. doi:10.1093/icesjms/fsab078.

Pascoe, S., Punt, A.E., and Dichmont, C.M. 2010. Targeting ability and output controls in Australia's multi-species Northern Prawn Fishery. European Review of Agricultural Economics **37**(3): 313–334. doi:10.1093/erae/jbq022.

Pastoors, M. 2000. Effects of a partially closed area in the North Sea ("plaice box") on stock development of plaice. ICES Journal of Marine Science **57**(4): 1014–1022. doi:10.1006/jmsc.2000.0586.

Pauly, D. 1998. Beyond Our Original Horizons: the Tropicalization of Beverton and Holt. : 28.

Privitera-Johnson, K.M., Methot, R.D., and Punt, A.E. 2022. Towards best practice for specifying selectivity in age-structured integrated stock assessments. Fisheries Research **249**: 106247. doi:10.1016/j.fishres.2022.106247.

Punt, A.E. 2023. Those who fail to learn from history are condemned to repeat it: A perspective on current stock assessment good practices and the consequences of not following them. Fisheries Research.

Punt, A.E., Hurtado-Ferro, F., and Whitten, A.R. 2014a. Model selection for selectivity in fisheries stock assessments. Fisheries Research **158**: 124–134. doi:10.1016/j.fishres.2013.06.003.

Punt, A.E., Pulfrich, A., Butterworth, D.S., and Penney, A.J. 1996. The effect of hook size on the size-specific selectivity of hottentot *Pachymetopon blochii* (Val.) and on yield per recruit. South African Journal of Marine Science **17**(1): 155–172. doi:10.2989/025776196784158473.

Punt, A.E., Smith, A.D.M., Smith, D.C., Tuck, G.N., and Klaer, N.L. 2014b. Selecting relative abundance proxies for BMSY and BMEY. ICES Journal of Marine Science **71**(3): 469–483. doi:10.1093/icesjms/fst162.

Punt, A.E., Tuck, G.N., Day, J., Canales, C.M., Cope, J.M., de Moor, C.L., De Oliveira, J.A.A., Dickey-Collas, M., Elvarsson, B.Þ., Haltuch, M.A., Hamel, O.S., Hicks, A.C., Legault, C.M., Lynch, P.D., and Wilberg, M.J. 2020. When are model-based stock assessments rejected for use in management and what happens then? Fisheries Research **224**: 105465. doi:10.1016/j.fishres.2019.105465.

Quinn, T.J.I., and Deriso, R.B. 1999. Quantitative fish dynamics. Oxford University Press, Oxford.

Sainsbury, K.J. 1984. Optimal mesh size for tropical multispecies trawl fisheries. ICES Journal of Marine Science **41**(2): 129–139. doi:10.1093/icesjms/41.2.129.

Sampson, D.B. 2014. Fishery selection and its relevance to stock assessment and fishery management. Fisheries Research **158**: 5–14. doi:10.1016/j.fishres.2013.10.004.

Sampson, D.B., and Scott, R.D. 2012. An exploration of the shapes and stability of population-selection curves: Shapes and stability of population-selection curves. Fish and Fisheries **13**(1): 89–104. doi:10.1111/j.1467-2979.2011.00417.x.

Scott, R.D., and Sampson, D.B. 2011. The sensitivity of long-term yield targets to changes in fishery age-selectivity. Marine Policy **35**(1): 79–84. doi:10.1016/j.marpol.2010.08.005.

SEDAR. 2018. SEDAR 52 - Gulf of Mexico Red Snapper stock assessment report. SEDAR, North Charleston SC. : 434.

Stewart, I.J., and Martell, S.J.D. 2014. A historical review of selectivity approaches and retrospective patterns in the Pacific halibut stock assessment. Fisheries Research **158**: 40–49. doi:10.1016/j.fishres.2013.09.012.

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. Fisheries Research **240**: 105967. doi:10.1016/j.fishres.2021.105967.

Szuwalski, C.S., Ianelli, J.N., and Punt, A.E. 2018. Reducing retrospective patterns in stock assessment and impacts on management performance. ICES Journal of Marine Science **75**(2): 596–609. doi:10.1093/icesjms/fsx159.

Thompson, G.G. 1994. Confounding of Gear Selectivity and the Natural Mortality Rate in Cases where the Former is a Nonmonotone Function of Age. Can. J. Fish. Aquat. Sci. **51**(12): 2654–2664. doi:10.1139/f94-265.

Thorson, J.T., and Taylor, I.G. 2014. A comparison of parametric, semi-parametric, and non-parametric approaches to selectivity in age-structured assessment models. Fisheries Research **158**: 74–83. doi:10.1016/j.fishres.2013.10.002.

Trijoulet, V., Albertsen, C.M., Kristensen, K., Legault, C.M., Miller, T.J., and Nielsen, A. 2023. Model validation for compositional data in stock assessment models: Calculating residuals with correct properties. Fisheries Research **257**: 106487. doi:10.1016/j.fishres.2022.106487.

Van Putten, I.E., Kulmala, S., Thébaud, O., Dowling, N., Hamon, K.G., Hutton, T., and Pascoe, S. 2012. Theories and behavioural drivers underlying fleet dynamics models. Fish and Fisheries **13**(2): 216–235. doi:10.1111/j.1467-2979.2011.00430.x.

Waterhouse, L., Sampson, D.B., Maunder, M., and Semmens, B.X. 2014. Using areas-as-fleets selectivity to model spatial fishing: Asymptotic curves are unlikely under equilibrium conditions. Fisheries Research **158**: 15–25. doi:10.1016/j.fishres.2014.01.009.

Watson, J.W., and Kerstetter, D.W. 2006. Pelagic Longline Fishing Gear: A Brief History and Review of Research Efforts to Improve Selectivity. mar technol soc j **40**(3): 6–11. doi:10.4031/002533206787353259.

Xu, H., Thorson, J.T., and Methot, R.D. 2020. Comparing the performance of three data-weighting methods when allowing for time-varying selectivity. Can. J. Fish. Aquat. Sci. **77**(2): 247–263. doi:10.1139/cjfas-2019-0107.

Xu, H., Thorson, J.T., Methot, R.D., and Taylor, I.G. 2019. A new semi-parametric method for autocorrelated age- and time-varying selectivity in age-structured assessment models. Can. J. Fish. Aquat. Sci. **76**(2): 268–285. doi:10.1139/cjfas-2017-0446.

Tables

Table 1. Descriptions of the operating model (OM) scenarios. 'Assessment Periods' represent the various points in the time series when a given EM was applied (i.e., representing different data quantity scenarios), while the term 'Intersection' in this column indicates the year in which the fishing mortality multiplier first intersects between fleets (Fig. 1).

OM Abbreviations	Fleet Structure Change	Assessment Periods	Selectivity Functional Form	Description of OM
Fast-Logistic	Fast	Fleet Intersection: Years 1 - 27 Fleet Transition End: Years 1 - 30 Terminal: Years: 1 - 50	Fleet 1: Logistic Fleet 2: Logistic	Gear change occurs rapidly with little difference in selectivity functional forms among fleets.
Fast-Gamma-Old	Fast	Fleet Intersection: Years 1 - 27 Fleet Transition End: Years 1 - 30 Terminal: Years: 1 - 50	Fleet 1: Logistic Fleet 2: Gamma with moderate dome	Gear change occurs rapidly, but new fishery exhibits moderately reduced selectivity of older individuals.
Fast-Gamma-Young	Fast	Fleet Intersection: Years 1 - 27 Fleet Transition End: Years 1 - 30 Terminal: Years: 1 - 50	Fleet 1: Logistic Fleet 2: Gamma with strong dome	Gear change occurs rapidly, but new fishery exhibits increased selectivity of young individuals and strongly decreased selectivity of older individuals.
Slow-Logistic	Slow	Fleet Intersection: Years 1 - 40 Fleet Transition End: Years 1 - 50 Terminal: Years: 1 - 70	Fleet 1: Logistic Fleet 2: Logistic	Gear change occurs slowly with little difference in selectivity functional forms among fleets.
Slow-Gamma-Old	Slow	Fleet Intersection: Years 1 - 40 Fleet Transition End: Years 1 - 50 Terminal: Years: 1 - 70	Fleet 1: Logistic Fleet 2: Gamma with moderate dome	Gear change occurs slowly, but new fishery exhibits moderately reduced selectivity of older individuals.
Slow-Gamma-Young	Slow	Fleet Intersection: Years 1 - 40 Fleet Transition End: Years 1 - 50 Terminal: Years: 1 - 70	Fleet 1: Logistic Fleet 2: Gamma with strong dome	Gear change occurs slowly, but new fishery exhibits increased selectivity of young individuals and strongly decreased selectivity of older individuals.

1079 **Table 2.** Description of estimation models (EMs) evaluated.

EM Abbreviations	Fleet Structure	Selectivity Functional Forms	Time-variation Parameterization	Description of EM
2Fleet-Logistic	Two fleets	Fleet 1: Logistic Fleet 2: Logistic	Time-invariant	Both fleets assume time-invariant logistic selectivity
2Fleet-Gamma	Two fleets	Fleet 1: Logistic Fleet 2: Gamma	Time-invariant	Fleet 1 assumes time-invariant logistic selectivity, while fleet 2 assumes time-invariant gamma selectivity.
1Fleet-TimeInvar-Logistic	One fleet	Logistic	Time-invariant	Single-fleet model assuming time-invariant logistic selectivity.
1Fleet-TimeInvar-Gamma	One fleet	Gamma	Time-invariant	Single-fleet model assuming time-invariant gamma selectivity.
1Fleet-Block-Logistic	One fleet	Time Block 1 (Year 1 – 24): Logistic; Time Block 2 (Year 25 – Terminal): Logistic	Time block	Single-fleet model assuming time-varying selectivity as a time block. Both time blocks assume logistic selectivity.
1Fleet-Block-Gamma	One fleet	Time Block 1 (Year 1 – 24): Logistic; Time Block 2 (Year 25 – Terminal): Gamma	Time block	Single-fleet model assuming time-varying selectivity as a time block. Time block 1 assumes logistic selectivity and time block 2 assumes gamma selectivity.
1Fleet-RandWlkPar-Logistic	One fleet	Logistic	Random walk deviations on parameters	Single-fleet model assuming continuous time-varying logistic selectivity with deviations on selectivity parameters (i.e., $a^{50}$and k).
1Fleet-RandWlkPar-Gamma	One fleet	Gamma	Random walk deviations on parameters	Single-fleet model assuming continuous time-varying gamma selectivity with deviations on selectivity parameters (i.e., $a^{max}$and γ).
1Fleet-SemiPar-Logistic	One fleet	Logistic	Semi-parametric	Single-fleet model assuming continuous time-varying logistic selectivity with deviations on selectivity values by age and year.
1Fleet-SemiPar-Gamma	One fleet	Gamma	Semi-parametric	Single-fleet model assuming continuous time-varying gamma selectivity with deviations on selectivity values by age and year.

1080

Table 3. Minimax solutions for each estimation model (EM; rows) across operating model (OM) scenarios (columns) and within assessment periods (i.e., when the stock assessment was carried out; nested rows). Values are Median Absolute Relative Errors (MAREs) in SSB summarized across all years and simulation replicates for a given EM. Values in bold identify the minimax solution for a given assessment period, which is the EM that has the smallest value of maximum MAREs across all OM scenarios.

	Fast-Logistic	**Fast-Gamma-Old**	**Fast-Gamma-Young**	**Slow-Logistic**	**Slow-Gamma-Old**	**Slow-Gamma-Young**
Assessment Period: Fleet Intersection						
2Fleet-Logistic	0.0293	0.0361	0.1348	0.0209	0.0292	0.1289
2Fleet-Gamma	0.0308	**0.0312**	0.0268	0.0243	0.0201	0.0240
1Fleet-TimeInvar-Logistic	0.0295	0.0320	0.0289	0.0347	0.0266	0.0263
1Fleet-TimeInvar-Gamma	0.0322	0.0344	0.0421	0.0235	0.0233	0.0371
1Fleet-Block-Logistic	0.0287	0.0330	0.0280	0.0328	0.0250	0.0240
1Fleet-Block-Gamma	0.0301	0.0326	0.0282	0.0261	0.0232	0.0252
1Fleet-RandWlkPar-Logistic	0.0280	0.0321	0.0274	0.0263	0.0232	0.0240
1Fleet-RandWlkPar-Gamma	0.0324	0.0324	0.0359	0.0257	0.0244	0.0371
1Fleet-SemiPar-Logistic	0.0338	0.0377	0.0353	0.0249	0.0272	0.0303
1Fleet-SemiPar -Gamma	0.0347	0.0404	0.0371	0.0261	0.0293	0.0304
Assessment Period: Fleet Transition End						
2Fleet-Logistic	0.0239	0.0340	0.1427	0.0186	0.0286	0.1312
2Fleet-Gamma	**0.0308**	0.0279	0.0259	0.0256	0.0203	0.0225
1Fleet-TimeInvar-Logistic	0.0861	0.0627	0.0648	0.0831	0.0604	0.0428
1Fleet-TimeInvar-Gamma	0.0499	0.0399	0.0737	0.0489	0.0387	0.0624
1Fleet-Block-Logistic	0.0514	0.0438	0.0390	0.0618	0.0486	0.0307
1Fleet-Block-Gamma	0.0332	0.0306	0.0324	0.0470	0.0369	0.0304
1Fleet-RandWlkPar-Logistic	0.0275	0.0300	0.0343	0.0269	0.0265	0.0303
1Fleet-RandWlkPar-Gamma	0.0316	0.0319	0.0332	0.0294	0.0304	0.0346
1Fleet-SemiPar-Logistic	0.0317	0.0323	0.0350	0.0260	0.0259	0.0294
1Fleet-SemiPar -Gamma	0.0312	0.0323	0.0363	0.0258	0.0261	0.0303

Assessment Period: Terminal						
2Fleet-Logistic	0.0171	0.0310	0.1456	0.0154	0.0330	0.1329
2Fleet-Gamma	**0.0273**	0.0207	0.0235	0.0196	0.0180	0.0203
1Fleet-TimeInvar-Logistic	0.1840	0.1526	0.0931	0.1310	0.0998	0.1061
1Fleet-TimeInvar-Gamma	0.1435	0.1183	0.1115	0.1011	0.0728	0.1552
1Fleet-Block-Logistic	0.0490	0.0318	0.1077	0.0506	0.0427	0.0534
1Fleet-Block-Gamma	0.0440	0.0371	0.0301	0.0489	0.0425	0.0667
1Fleet-RandWlkPar-Logistic	0.0316	0.0276	0.1113	0.0390	0.0245	0.0753
1Fleet-RandWlkPar-Gamma	0.0381	0.0462	0.0296	0.0421	0.0357	0.0309
1Fleet-SemiPar-Logistic	0.0320	0.0302	0.0359	0.0322	0.0263	0.0316
1Fleet-SemiPar -Gamma	0.0318	0.0299	0.0372	0.0327	0.0264	0.0327

1089

Figures

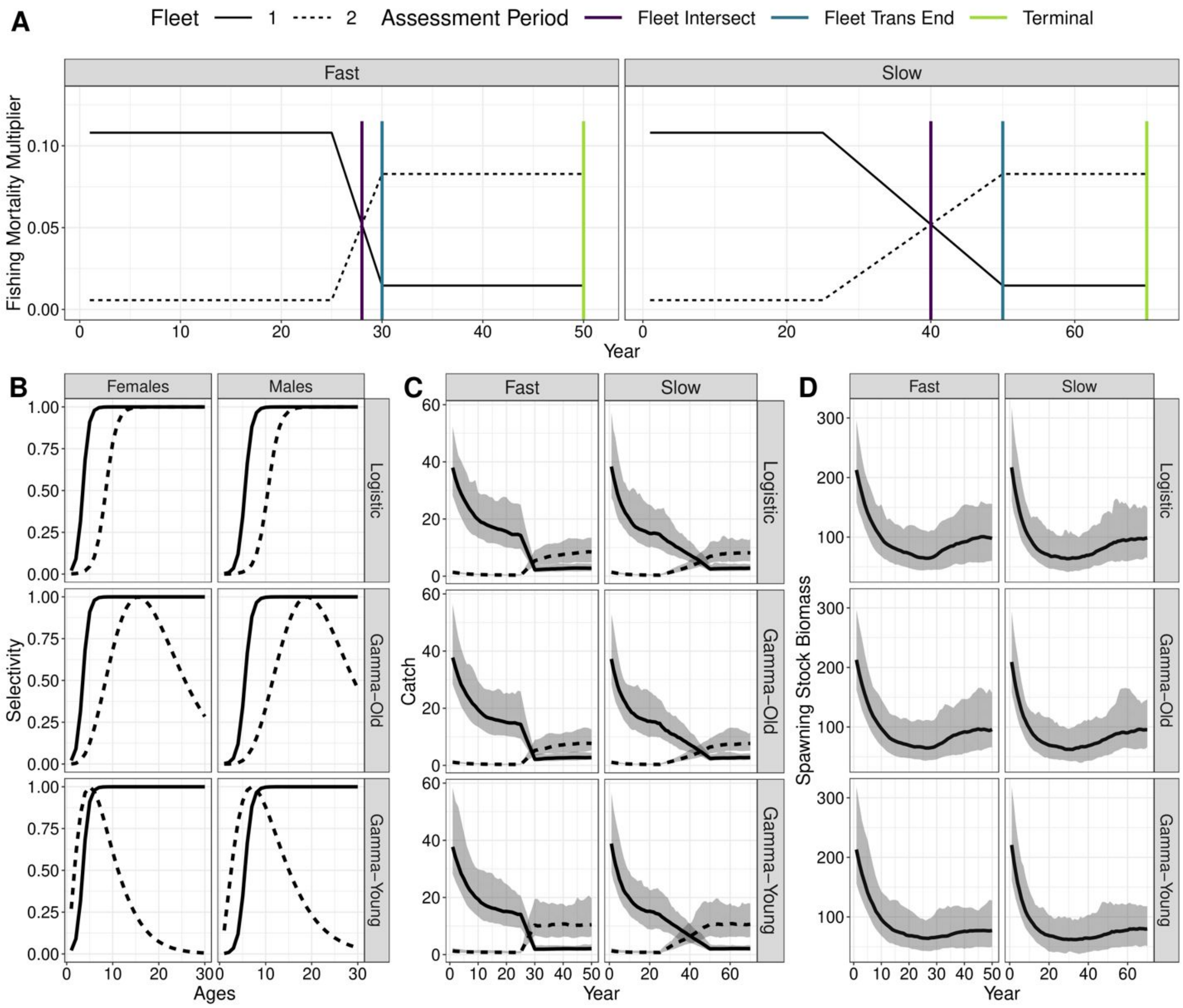


Figure 1: Overview of the operating model settings. Solid lines refer to fishery fleet 1 and dashed lines refer to fishery fleet 2. Panel A depicts the two different fleet structure transition scenarios (Fast and Slow) and the three assessment periods (vertical colored lines), where the y-axis represents fleet-specific instantaneous fishing mortality rates. Panel B depicts the three different selectivity scenarios evaluated. Selectivity for fleet 1 was always modeled with a logistic curve. For fishery fleet 2, selectivity was represented with a logistic curve, a gamma distribution selecting older fish (age-at-maximum selection: females age 15.5, males age 19), or a gamma

distribution that selected younger fish (age-at-maximum selection: females age 5, males age 7). Panel C displays the simulated catch resulting from each fishery fleet, while panel D demonstrates the resulting spawning stock biomass trajectories, where units for both panels are on the same scale. Shading in panels C and D represent 95% simulation intervals.

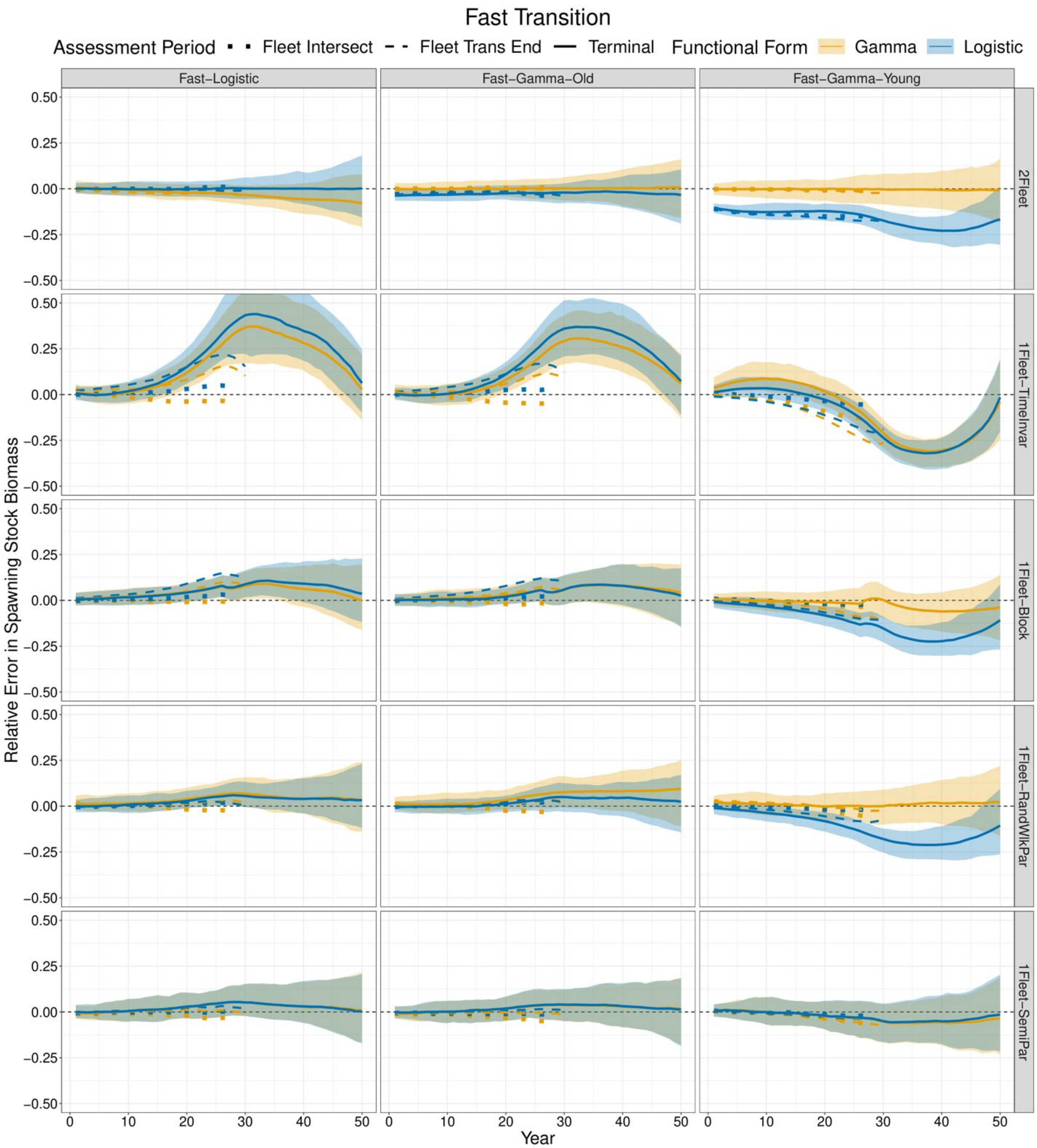
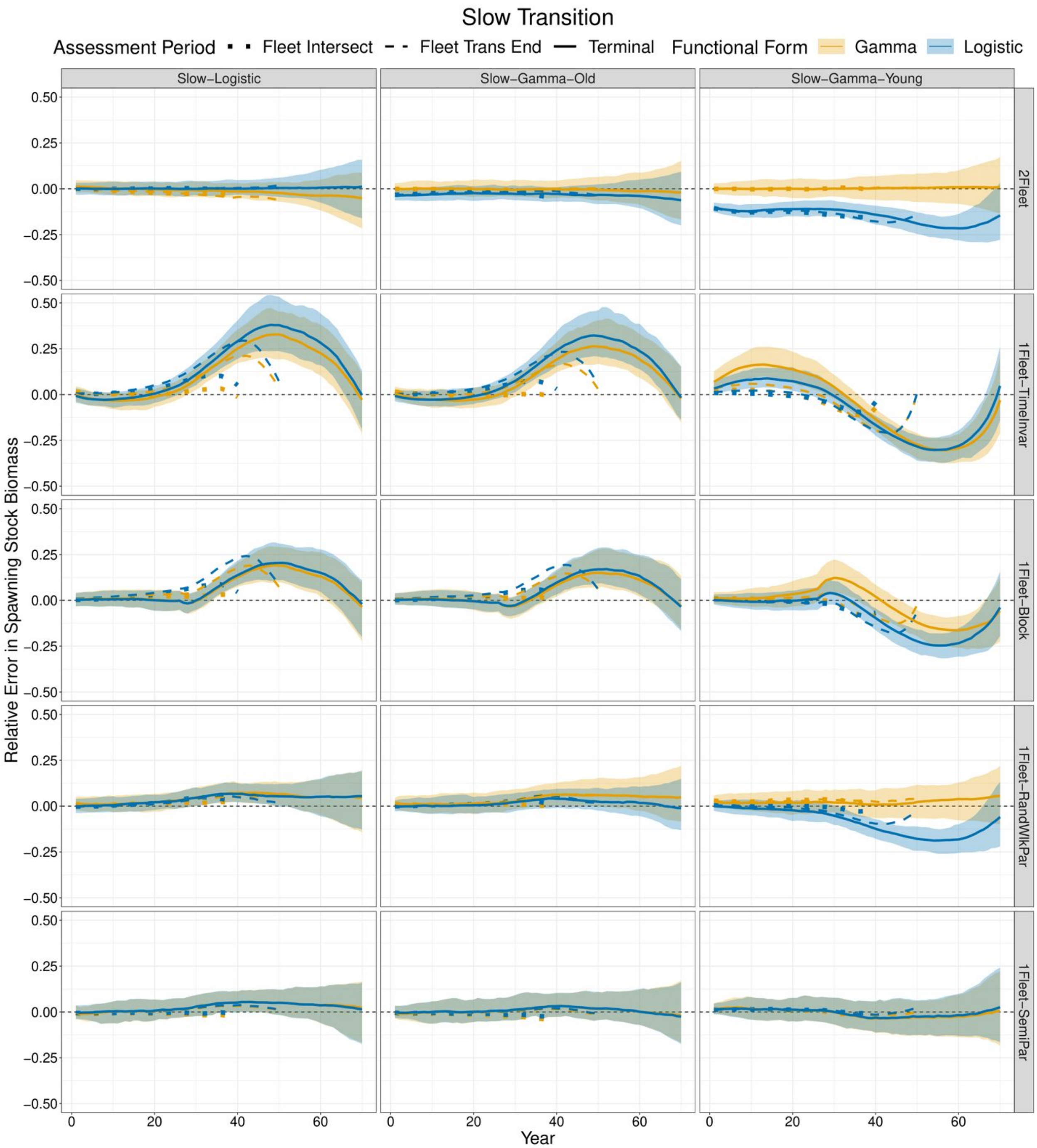


Figure 2: Relative error in estimated annual spawning stock biomass across operating model (OM) scenarios where a fast change in fleet structure was simulated. Only results from converged models are presented here. Column panels are OMs, while row panels (describing fleet structure and selectivity time-variation assumptions) in combination with colored lines

(orange: *Gamma*; blue: *Logistic*) denote estimation models (EMs). Line types describe the different assessment periods during which EMs were applied to. Lines represent the median relative error. The shading represents the 95% simulation intervals for each EM type applied during the terminal assessment period (to aid in clarity of visualizations, simulation intervals are only shown for EMs applied to the terminal period). The black horizontal line represents 0% relative error.



1150

1151 Figure 3: Relative error in estimated annual spawning stock biomass across operating model

1152 (OM) scenarios where a slow change in fleet structure was simulated. Only results from

1153 converged models are presented here. Column panels are OMs, while row panels (describing

1154 fleet structure and selectivity time-variation assumptions) in combination with colored lines (orange: *Gamma*; blue: *Logistic*) denote estimation models (EMs). Line types describe the different assessment periods during which EMs were applied to. Lines represent the median relative error. The shading represents the 95% simulation intervals for each EM type applied during the terminal assessment period (to aid in clarity of visualizations, simulation intervals are only shown for EMs applied to the terminal period). The black horizontal line represents 0% relative error.

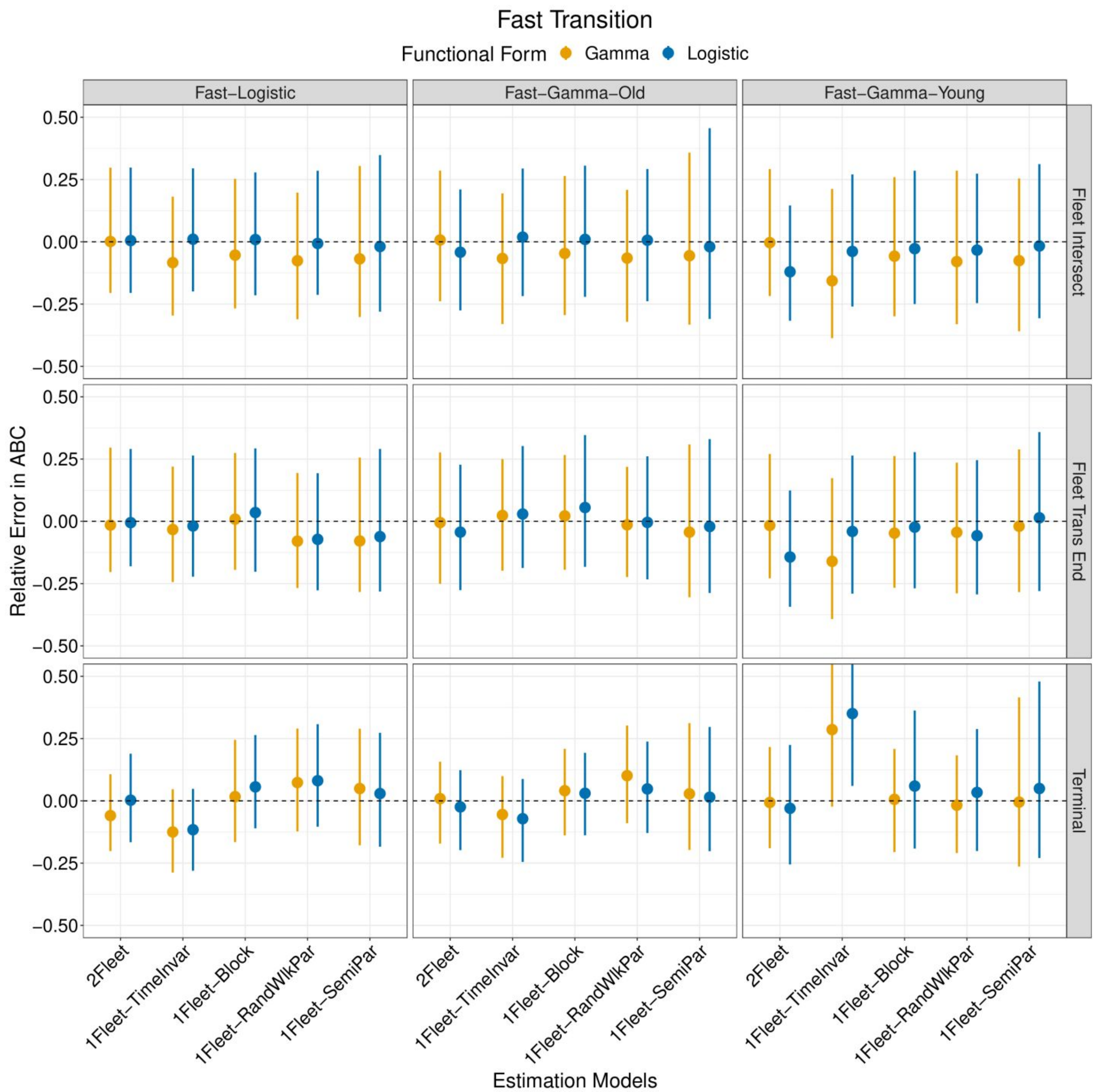


Figure 4: Relative error in Acceptable Biological Catch (ABC) across operating model (OM) scenarios where a fast change in fleet structure was simulated. Only results from converged models are presented here. Column panels represent the different OM scenarios. The x-axis (describing fleet structure and selectivity time-variation assumptions) in combination with colored points (orange: *Gamma*; blue: *Logistic*) denote estimation models (EMs). Row panels describe the different assessment periods during which EMs were applied to. Points represent the

median relative error and line ranges are the 95% simulation intervals. The black horizontal line represents 0% relative error.

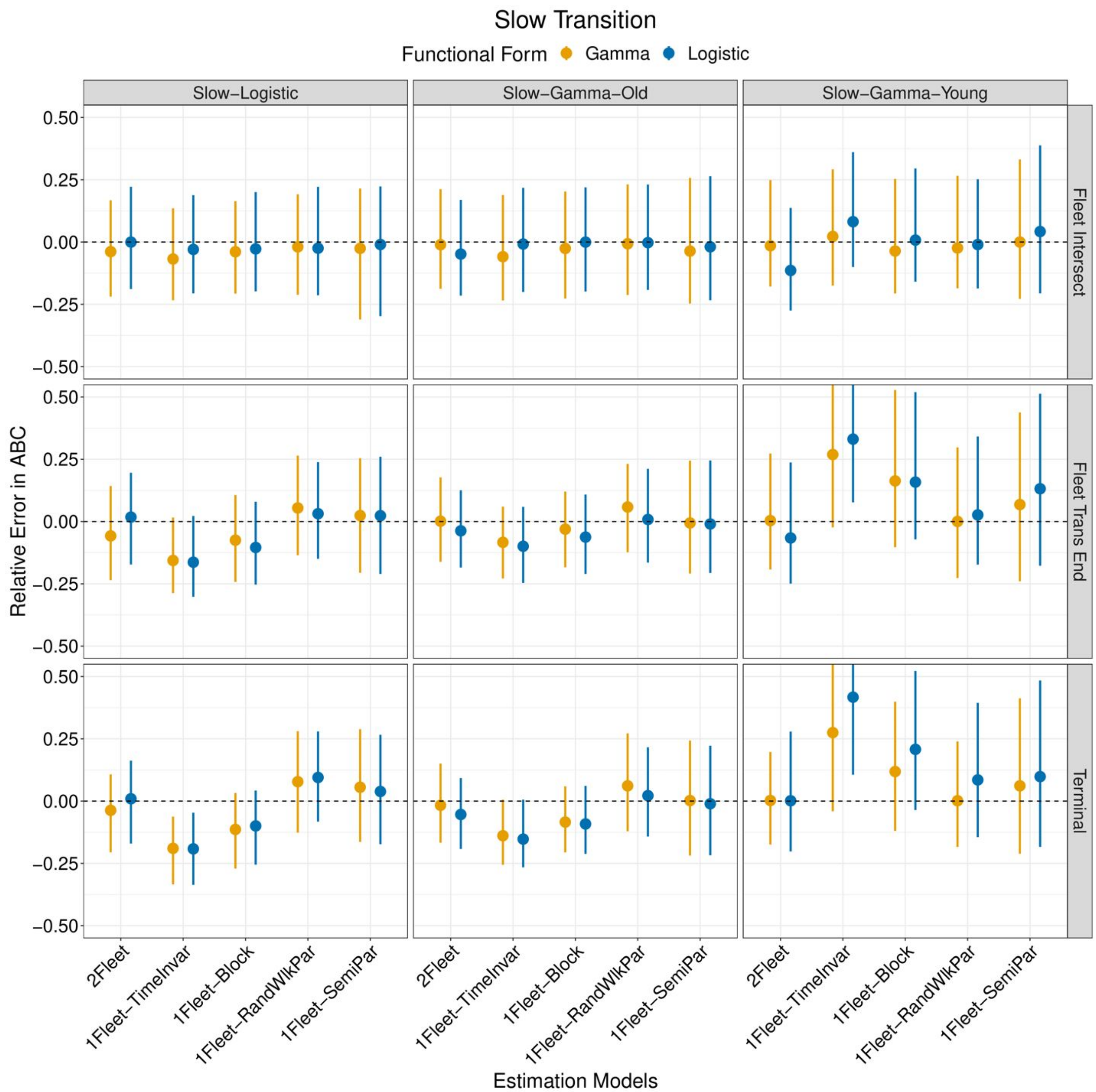


Figure 5: Relative error in Acceptable Biological Catch (ABC) across operating model (OM) scenarios where a slow change in fleet structure was simulated. Only results from converged models are presented here. Column panels represent the different OM scenarios. The x-axis (describing fleet structure and selectivity time-variation assumptions) in combination with colored points (orange: *Gamma*; blue: *Logistic*) denote estimation models (EMs). Row panels describe the different assessment periods during which EMs were applied to. Points represent the

median relative error and line ranges are the 95% simulation intervals. The black horizontal line represents 0% relative error.

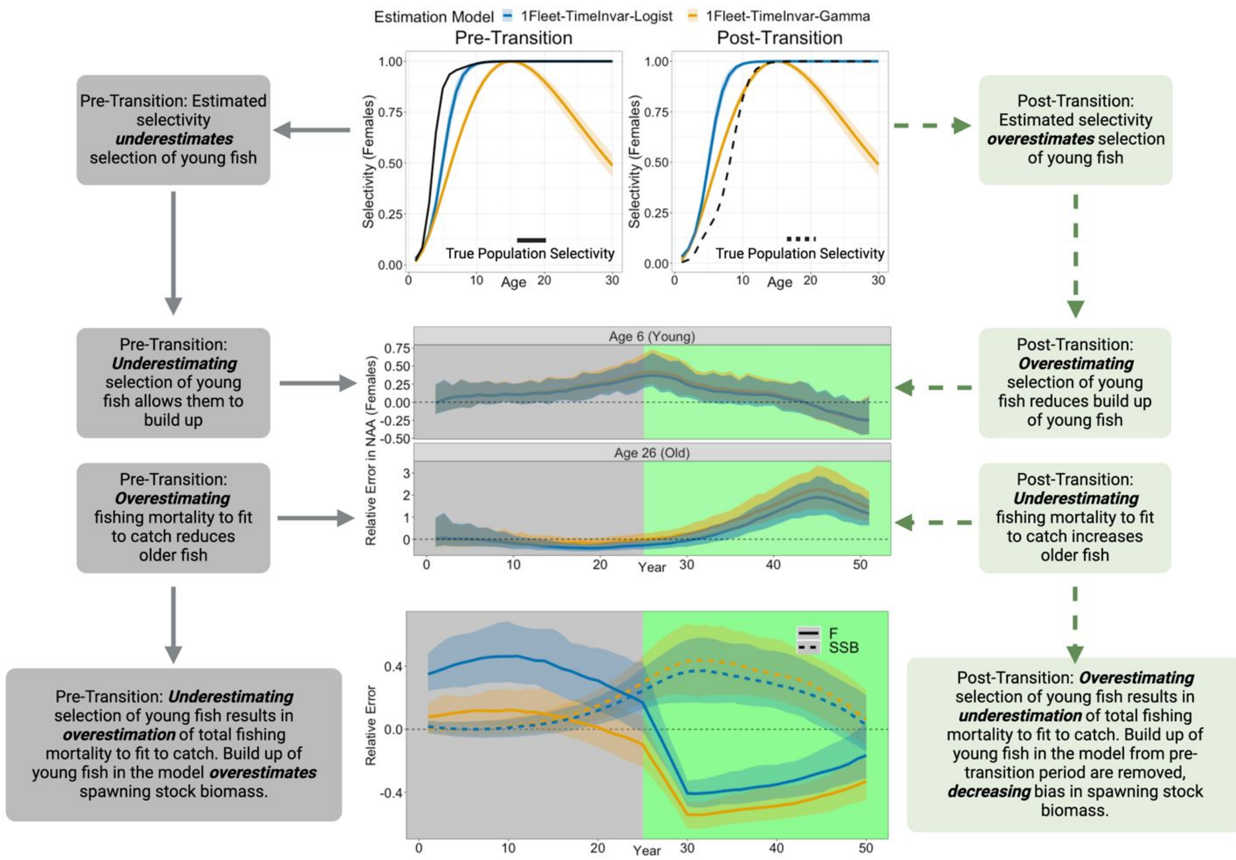


Figure 6. Schematic depicting how biases in spawning stock biomass (SSB) arise when changes in fleet structure are ignored by assuming a single-fleet model with time-invariant selectivity (blue: *1Fleet-TimeInvar-Logist*; orange: *1Fleet-TimeInvar-Gamma*), during the terminal assessment period. The first column describes how biases arise prior to the fleet transition (shown in grey), while the last column describes how biases arise following the fleet transition (shown in green). Lines accompanied with shaded intervals in the middle column are the median error and 95% simulation intervals, respectively. The upper row panel compares estimated selectivities against the true population selectivity for females. Relative error in numbers-at-age (NAA) for females are shown in the middle row panel. Only 2 ages are shown for clarity of visualization, but patterns in relative error of NAA are qualitatively similar between young (ages

1 – 15) and old (ages 16 – 30) individuals. The bottom row panel depicts relative error in SSB (dotted lines) and total fishing mortality (F; solid lines).

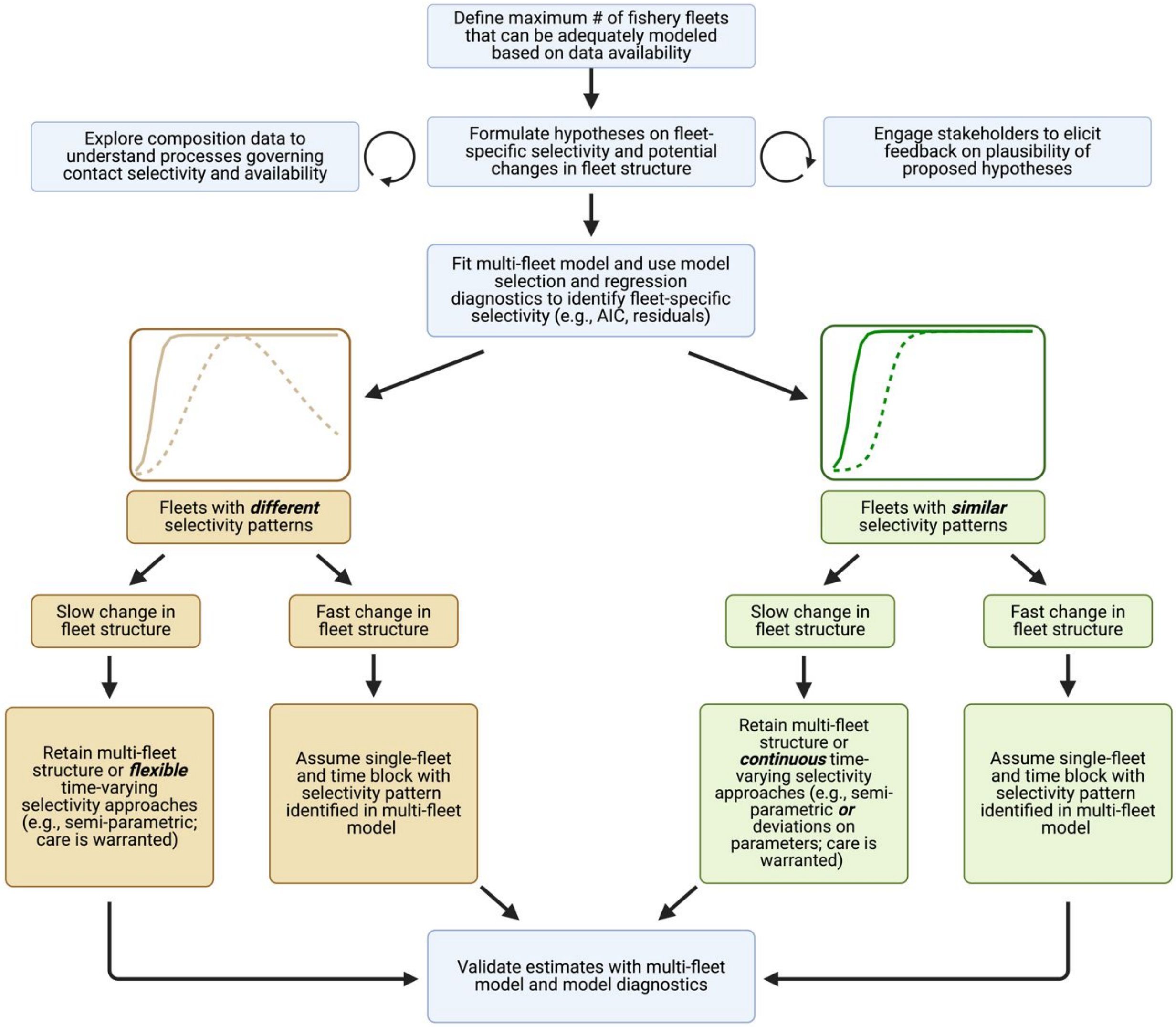


Figure 7: A decision tree portraying decision points for determining parameterizations of fleet structure and selectivity with pragmatic recommendations for each. Recommendations are intended to provide general guidance on model structure and assumes that fleet-specific catch and composition data are available.