SEDAR

Southeast Data, Assessment, and Review

SEDAR 87

Gulf Brown Shrimp

SECTION III: Assessment Process Report

Updated: 5 June 2025
Original Release: 28 May 2025

SEDAR
4055 Faber Place Drive, Suite 201
North Charleston, SC 29405

Document History

May 28, 2025 Original release

June 5, 2025 First letter of all species, regions, and size classes capitalized for consistency.

SEDAR87 Gulf Brown Shrimp Benchmark Assessment

Gulf Branch
Sustainable Fisheries Division
NOAA Fisheries - Southeast Fisheries Science Center

June 2025

Table of contents

1. Assessment Process Proceedings

On January 20, 2025, President Trump issued Executive Order 14172 to rename the Gulf of Mexico as the Gulf of America. Any reference to Gulf of America Brown Shrimp in SEDAR reports and other documents refers to the same species and fishery listed in *50 CFR part 622, Subpart C* (Shrimp Fishery of the Gulf of Mexico). As of the publication of this report, all efforts were made to use "Gulf of America" per Executive Order 14172. However, previous NOAA reports (cited herein) may have referred to this water body as the "Gulf of Mexico".

1.1 Introduction

1.1.1 Workshop Time and Place

The SEDAR 87 Assessment Process (AP) for Gulf Brown Shrimp was conducted via a series of webinars held between October 2024 and February 2025.

1.1.2 Terms of Reference

1. Review any changes in data or analyses following the Data Workshop. Summarize data as used in each assessment model. Provide justification for any deviations from Data Workshop recommendations.
2. Develop a management advice framework. Consider data availability (e.g., landings and catch-per-unit-effort [CPUE]) and management needs (e.g., harvest controls, stock status), and particular needs of the fishery and the biology of the resource.
3. Examine the impacts of social science factors on biological reference points as informed by stakeholders through industry input.
4. Recommend biological reference points for use in management.
    - Consider how reference points could be affected by management, ecosystem, climate, species interactions, habitat considerations, social or economic drivers, and/or episodic events.
5. Provide estimates of stock population parameters, including: Fishing mortality, biomass, selectivity, and/or other parameters as necessary to describe the population.
6. Characterize uncertainty in the assessment and estimated values.

- Consider uncertainty in input data, modeling approach, and model configuration.
- Provide appropriate measures of model performance, reliability, and 'goodness of fit'.
- Provide measures of uncertainty for estimated parameters and derived quantities such as biological reference points and stock status if feasible.

7. Provide recommendations for future research and data collection. Emphasize items that will improve future assessment capabilities and reliability. Consider data, monitoring, and assessment needs.

8. Complete an Assessment Workshop Report in accordance with project schedule deadlines.

1.1.3 List of Participants

Assessment Process Participants

Lisa Ailloud, Co-Lead Analyst ..... NMFS Miami
Molly Stevens, Co-Lead Analyst ..... NMFS Miami

Don Behringer ..... NMFS SEFSC
Jie Cao ..... NCSU/GMFMC SSC
Steve Munch ..... NOAA NMFS
Jim Nance ..... GMFMC SSC
Jason Saucier ..... MS DMR
Katie Siegfried ..... NMFS SEFSC
Brendan Turley ..... NMFS

Appointed Observers

Leann Bosarge ..... Industry Rep
Glenn Delany .....

Staff

Julie Neer ..... SEDAR
Emily Ott ..... SEDAR
Matt Freeman ..... GMFMC Staff
Dominique Lazarre ..... SERO
Michelle Masi ..... SERO
Ryan Rindone ..... GMFMC Staff
Carrie Simmons ..... GMFMC Staff

Assessment Process Webinar Observers

Sarina Atkinson ..... NMFS Miami
Peyton Cagle ..... LWFD
Judd Curtis ..... SAFMC Staff
Kyle Detloff ..... NMFS SEFSC
Traci Floyd ..... MS DMR
Carissa Gervasi ..... NOAA
Bob Gill ..... GMFMC

David Hanisko ..... NMFS SEFSC
Kimberly Johnson ..... NMFS SEFSC
Tricia Kimball .....
Christopher Liese ..... NMFS SEFSC
Alan Lowther ..... NMFS SEFSC
Richard Malinowski ..... NOAA NMFS
Jessica Marchant ..... AL DCNR
Fernando Martinez-Andrade ..... TPWD
Akbar Marvasti ..... NOAA
Cassidy Peterson ..... NMFS SEFSC
Cheston Peterson ..... NOAA
Adam Pollack ..... NMFS SEFSC
David Records ..... NOAA
Sarah Roberts ..... NOAA
Skyler Sagarese ..... NMFS SEFSC
Andrew Scalisi ..... LDWF
Chris Schieble ..... LDWF
Rebecca Smith ..... NOAA
Jim Tolan .....
Michael Travis ..... NOAA
Jo Williams ..... NMFS SEFSC

1.1.4 List of Assessment Process Working Papers and Reference Documents

Document #	Title	Authors	Date Submitted
Documents Prepared for the Assessment Process			
SEDAR87-AP-01	Development of estuarine environmental indices for SEDAR 87 Gulf of Mexico White, Pink, and brown shrimp stock assessment	Brendan Turley, Lisa Ailloud, and Molly Stevens	25 July 2024
SEDAR87-AP-02	Price Indices for Shrimp Imports and Gulf of Mexico Shrimp Landings by Size and Season	Christopher Liese	18 December 2024
SEDAR87-AP-03	Developing a fishery-independent index of relative abundance for Gulf of Mexico Brown Shrimp using VAST	Lisa Ailloud, Molly Stevens, Brendan Turley, Adam Pollack, and David Hanisko	31 January 2025

SEDAR87-AP-04	Developing a fishery-independent index of relative abundance for Gulf of Mexico Pink Shrimp using VAST	Lisa Ailloud, Molly Stevens, Brendan Turley, Adam Pollack, and David Hanisko	31 January 2025
SEDAR87-AP-05	Developing a fishery-independent index of relative abundance for Gulf of Mexico White Shrimp using VAST	Lisa Ailloud, Molly Stevens, Brendan Turley, Adam Pollack, and David Hanisko	31 January 2025
Reference Documents			
SEDAR87-RD12	JABBA: Just Another Bayesian Biomass Assessment	Henning Winker, Felipe Carvalho, Maia Kapur	
SEDAR87-RD13	Empirical dynamic modeling for sustainable benchmarks of short-lived species	Cheng-Han Tsai, Stephan B. Munch, Michelle D. Masi, and Molly H. Stevens	
SEDAR87-RD14	Recent developments in empirical dynamic modelling	Stephan B. Munch, Tanya L. Rogers, George Sugihara	
SEDAR87-RD15	Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat	James T. Thorson and Lewis A. K. Barnett	

2. Data Review and Update

The following list summarizes the data inputs (and units) used in the assessment modeling process along with their corresponding available temporal scale based upon recommendations from the Data Workshop process. Two assessment modeling platforms were considered: a Bayesian surplus production model, JABBA (Just Another Bayesian Biomass Assessment), and an Empirical Dynamic Modeling (EDM) platform (see *Section 3*). Data for JABBA were on an annual time scale and included commercial landings (in million pounds of tails) and an index of abundance built with SEAMAP and Texas Park and Wildlife (TPWD) survey data using Vector Auto-Regressive Spatio-Temporal (VAST) modeling (Ailloud et al. 2025). EDM explored all the datasets listed below using various levels of stratification. JABBA allowed for different start years of data inputs, while EDM was limited by the start year of the survey data. For EDM, data were stratified by fishing area [A (*Figure 1*) : 1-10, 11-17, 18-21], size [S: >67 (Small), 67-31

(Medium), <=30 (Large) tails per pound], and quadrimester of the year [Q: January-April (Winter), May-August (Summer), September-December (Fall)] where possible, and are indicated as such in the data list. Stratifications were defined based on existing definitions of the ecological distribution of shrimp and the shrimping industry.

1. Commercial landings (million pounds of tails): 1960-2022 [A, S, Q]
2. SEAMAP survey data (number of shrimp per trawl hour): 1987-2022 [A, S, Q]
3. Ex-Vessel price indices (2022 dollars): 1960-2023 [S, Q]
4. Imports (product volume in 100 million pounds): 1972-2022 [Q]
5. Salinity (practical salinity unit): 1987-2022 [A]
6. Bottom temperature (degrees Celsius): 1987-2022 [A]

Brown Shrimp are distributed primarily in the western Gulf (*Figure 1* : 11-21). Possible data stratifications for Brown Shrimp EDM were defined as:

A) Aggregated: ANNUAL ; SIZE BINS AGG ; AREA AGG (11:21)
B) Area: ANNUAL ; SIZE BINS AGG ; AREA (11:17, 18:21)
C) Size: ANNUAL ; SIZE BINS (>67, 67-31, <=30) ; AREA AGG (11:21)
D) Size_Area: ANNUAL ; SIZE BINS (>67, 67-31, <=30) ; AREA (11:17, 18:21)
E) Season: SEASONAL (SUMMER, FALL+WINTER) ; SIZE BINS AGG ; AREA AGG (11:21)
F) Area_Season: SEASONAL (SUMMER, FALL+WINTER) ; SIZE BINS AGG ; AREA (11:17, 18:21)
G) Size_Season: SEASONAL (SUMMER, FALL+WINTER) ; SIZE BINS (>67, 67-31, <=30) ; AREA AGG (11:21)
H) Size_Area_Season: SEASONAL (SUMMER, FALL+WINTER) ; SIZE BINS (>67, 67-31, <=30) ; AREA (11:17, 18:21)

For Brown Shrimp, additional strata were included that aggregated the Small and Medium size classes into a Smedium size class containing all shrimp smaller than >31 tails per pound.

Csm) Size: ANNUAL ; SIZE BINS (>30, <=30) ; AREA AGG (11:21)
Gsm) Size_Season: SEASONAL (SUMMER, FALL+WINTER) ; SIZE BINS (>30, <=30) ; AREA AGG (11:21)
Hsm) Size_Area_Season: SEASONAL (SUMMER, FALL+WINTER) ; SIZE BINS (>30, <=30) ; AREA (11:17, 18:21)

2.1 Stock Structure and Management Unit

The SEDAR 87 Gulf Brown Shrimp Benchmark Assessment stock boundary extends from the United States–Mexico border in the west through the northern Gulf of America waters (hereafter referred to as the Gulf) to the Dry Tortugas and Florida Keys. This includes all waters within the Gulf of Mexico Fishery Management Council (hereafter referred to as the Gulf Council) boundaries and extends to include fishing areas split by the eastern boundary off the Florida Keys (*Figure 1*: Areas 002, 001) in their entirety due to complications with reporting over time

(Atkinson et al. 2024). This stock boundary distinction is most important for Pink Shrimp due to its distribution being centered in the eastern Gulf, but it was applied to all Gulf shrimp species.

2.2 Fishery-Independent Survey Data

2.2.1 Southeast Area Monitoring and Assessment Program (SEAMAP)

The Southeast Area Monitoring and Assessment Program (SEAMAP) survey is a collaborative effort between federal, state, and university programs, designed to collect, manage, and distribute fishery-independent data throughout the region. The SEAMAP survey design was improved and expanded in 2008 and was deemed representative for Brown Shrimp since 1987 for all provided size classes (Pollack and Hanisko 2024). The SEAMAP survey operates in the Fall and Summer, and seasonality was explored categorically (*Figure 2*) and continuously (*Figure 3*). The Summer survey tracks Brown Shrimp populations after they have recruited to the population, and the Fall generally surveys what is left after most of the fishery has taken place. Raw annual indices for Brown Shrimp are shown by size class and season in *Table 1*.

Due to the global COVID-19 pandemic, the 2020 Summer SEAMAP survey did not operate. There are variable time step methods available for dealing with missing data in EDM, but these approaches considerably increase model complexity and are typically only used when variable-sized gaps in data are present throughout a long time series. Using a variable time step method was determined to be inappropriate for addressing a single missing time step of data; however, it was also not desired to stop the model in 2019 when data were available through 2022. Therefore, the 2020 missing data point was replaced by an average of 2019 and 2021 on the finest resolution possible (i.e., annual aggregations averaged 2019/21, but seasonal models only averaged Summer 2019/21 and directly included Fall 2020).

2.2.2 Vector Autoregressive Spatio-Temporal (VAST) Index

VAST is a spatio-temporal modeling platform that can be used for standardizing indices of relative abundance. Data from one or more surveys are combined to predict population density based on both habitat covariates (that impact abundance) and spatial and spatio-temporal random effects, while controlling for catchability covariates (that impact sampling efficacy). A VAST index was developed for Brown Shrimp based on data from SEAMAP and the TPWD surveys for input into JABBA. Details of the VAST index are documented in Ailloud et al. (2025).

2.3 Fishery-Dependent Data

2.3.1 Commercial Landings

Commercial landings of Brown Shrimp were constructed using data from the Gulf Shrimp System (GSS) and state trip ticket programs. Species-specific Gulf shrimp landings have been collected since the late 1950s, and their complex history within the federal and state databases, including justifications for the relative coefficients of variance (CVs) through time, is documented in great detail in Atkinson et al. (2024). Landings were converted to tail weight for input to the assessment model.

Landings from the Winter months (JFMA) were minimal. For EDM, they were aggregated with the previous years' Fall (SOND) landings to account for removals during this time period and

match the seasonal timestep of the SEAMAP survey (*Figure 4*). Shrimp landings have been sold and recorded by eight market categories, which were aggregated into three general size classes. These are shown broken out seasonally in *Figure 5* and with the real-time seasonal fluctuations by size class in *Figure 6*. The Small size category displays the largest fluctuations, where new recruits enter the population in the Summer and are either caught or have grown out to larger size classes by the Fall (*Table 2*).

Brown Shrimp landings peaked in 1990 with landings totaling 105.91 million pounds of tails. Landings have been declining since the mid-2000s due to economic conditions described in the following section, Griffith et al. (2023), and Atkinson et al. (2024).

2.4 Economics and Social Sciences

2.4.1 Imports and Ex-vessel Price Indices

Imported shrimp have exceeded the volume of domestically caught shrimp since the 1980's (Lowther 2023; Atkinson et al. 2024). In the mid-2000s, the volume of imported shrimp increased dramatically, particularly for Large shrimp which has a higher market value, causing domestic ex-vessel prices to plummet (*Figure 7*). Time series of imports and ex-vessel prices were both considered during EDM development (Liese 2024).

2.4.2 Industry Impacts

The globalization of the shrimp market with a focus on cheap aquaculture has resulted in dire economic operating conditions for the domestic fleet (Griffith et al. 2023). Increasing fuel costs and plummeting ex-vessel prices have created a situation in which most vessels struggle to remain profitable. Further, many vessels have exited the fleet, and those that remain may oscillate between narrowing profit margins and losses (*SEDAR87 data workshop report* 2023 pp. 84–94). With fewer vessels operating, the shrimping effort and associated landings have decreased overall, and the shrimp population size has increased.

Industry impacts were documented during a stakeholder listening session at the Data Workshop, with the intention of holding additional listening sessions throughout coastal Gulf shrimping communities. During this session, resource users stated that the troubles of the Shrimp Fishery cannot be improved by domestic fishery management solutions. The bulk of the problems are globally influenced, and this fishery was recommended to the National Seafood Strategy to address these problems if possible, informed by additional information gathered through the newly formed *Shrimp Futures Project*.

2.5 Environmental Indices

Annual shrimp recruitment has been tied to environmental drivers (Browder et al. 2002; Zink et al. 2018; Schlenker et al. 2023). Within an assessment modeling framework, it is important to include drivers of abundance at the most meaningful spatio-temporal scale. At the SEDAR 87 Data Workshop, the Environment and Industry Working Group recommended that salinity and temperature in the nursery grounds during the months that the shrimp were inhabiting the area were likely the primary environmental drivers for shrimp abundance. These two variables were hypothesized to best explain the magnitude of recruits into the population each year. The methodology used to derive Brown Shrimp temperature and salinity indices was outlined in

Turley et al. (2023). These indices were considered in the construction of the VAST index and development of EDM.

Brown Shrimp is in its inshore nursery grounds February through May each year throughout its coastal range. It was hypothesized that the environment would affect the overall population abundance more directly through its impact on the young of the year in this volatile habitat. While there may be some impacts of seasonal differences in rainfall and temperature fluctuations affecting local abundance, the trends of data from both Texas (TX; north of Laguna Madre) and Louisiana (LA) appeared to follow strikingly similar trends, indicating consistency throughout the range (*Figure 8*).

2.5.1 Temperature

Temperature in the western Gulf follows nearly identical trends of state temperature averages from TX (north of Laguna Madre) to LA, with TX experiencing more extreme lows in some years. The standardized temperature index represents the nursery conditions well, which were remarkably similar on average, throughout these variable estuarine habitats.

2.5.2 Salinity

While salinity in TX was much higher compared to salinity in LA, they generally experienced co-occurring peaks and troughs, resulting in a standardized index that tracks changes in salinity well throughout the region.

3. Stock Assessment Model Configurations and Methods

Two modeling frameworks were evaluated for the Gulf Brown Shrimp SEDAR 87 Benchmark Assessment: Just Another Bayesian Biomass Assessment (JABBA) Model and Empirical Dynamic Modelling (EDM). These are described below.

3.1 Just Another Bayesian Biomass Assessment (JABBA) Model

JABBA is a Bayesian state-space surplus production model (SPM) framework that is documented in Winker et al. (2018) and is available as an R package on *GitHub*. SPMs pool the overall effects of recruitment, somatic growth, natural mortality, and associated density-dependent processes into a single production function dealing with undifferentiated biomass (Haddon 2021). The state-space formulation allows for the estimation of observation and process error, and the Bayesian formulation allows the user to define prior distributions for each parameter in the model to represent the initial beliefs about the parameter before observing any data. Primary data inputs into JABBA are indices of abundance proportional to the exploitable part of the stock biomass and a time series of fishery removals. The time series of removals can begin prior to the indices of abundance, and contrast in the data is required to appropriately map the stock dynamics.

The generalized surplus production function (Pella and Tomlinson 1969) used by JABBA is defined as

$$SPM_t = \frac{r}{m-1}\left(1 - {\frac{B_t}{K}}^{m-1}\right)$$

where r is the intrinsic rate of population increase at time t, K is the carrying capacity, B is the stock biomass at time t, and m is the shape parameter that determines at which B/K ratio maximum surplus production is attained. The Pella-Tomlinson function above is a generalized production function with Schaefer ($m = 2$) and Fox ($m = 1$) as special cases. The Schaefer may be the most well-known, with a symmetrical production curve and Maximum Sustainable Yield (MSY) attained at half the carrying capacity, $B = K/2$.

JABBA has several features including the ability to a) fit multiple CPUE time series and associated standard errors, b) estimate or fix the process variance, c) estimate additional observation variance on individual or grouped CPUE series, and d) specify either a Fox, Schaefer or Pella-Tomlinson production function. A full JABBA model description, including formulation and state-space implementation, prior specification options, and diagnostic tools is available in Winker et al. (2018).

3.1.1 Estimated Parameters

JABBA model parameters are defined in greater detail below.

K: Carrying capacity (million lb tail weight)

m: Shape parameter of the Pella-Tomlinson that determines at which B/K ratio maximum surplus production is attained. If $m = 2$, the model reduces to the Schaefer form, with the surplus production (SP) attaining MSY at exactly $K/2$. If $0 < m < 2$, SP attains MSY at biomass levels smaller than $K/2$; the converse applies for values of m greater than 2.

ψ: Ratio of the spawning biomass in the first year to K.

q: Catchability coefficient.

r: Intrinsic rate of population increase.

σ^2: Process variance.

τ^2 : Additional observation variance for the survey index.

3.1.2 Model Configurations and Prior Assumptions

The final VAST index built on SEAMAP and Texas Park and Wildlife (TPWD) survey data presented in Ailloud et al. (2025) was used as input to JABBA alongside an annual time series of commercial catches spanning 1960-2022 (*Section 2.3.1*). The following CVs were recommended by the WG and input into JABBA to reflect uncertainty in landings based on changes in the sampling programs through time. 1960-1983: CV = 0.2, 1984-2015: CV = 0.1, 2016-2022: CV = 0.05. The time series and associated confidence intervals are shown in *Figure 9* and *10*. Model configurations and prior distributions were defined as follows:

Carrying capacity (K): uninformative prior. Lognormal distribution specified using the "range" option in JABBA with lower and upper values ranging from maximum catch to 10x maximum catch (*Figure 11*)

Production function: Pella-Tomlinson (MSY at $B_{MSY}/K = 0.4$; $CV = 0.3$) where B_{MSY} is the biomass at MSY (*Figure 12*)

Process error variance (σ^2): Default $1/\gamma(4,0.01)$ (*Figure 13*). This matches the level of process error where state-space SPMs are most likely to adequately perform.

Observation error variance (if estimated) (τ^2): Default $\sim 1/\gamma(0.001,0.001)$ (*Figure 14*)

r prior: informative priors were developed based on the Medium (0.2-0.8) and High (0.6-1.5) resilience categories in FishBase (Froese et al. 2019). Given that FishBase does not include any crustaceans and that shrimp are likely on the higher range of r compared to most fishes, an additional Very High (1.2-3) prior was tested (*Figure 15*)

Initial biomass depletion ratio (ψ): two alternative priors were tested to reflect Low initial depletion $lognormal(0.9,0.25)$ and and High initial depletion $lognormal(0.25,0.5)$ at the beginning of the catch time series ($\psi = B_{1960} \,/\, K$) (*Figure 16*)

A factorial design was used to test a suite of models with alternative prior assumptions about r, ψ and τ^2. The naming convention for candidate model is as follows:

SpeciesCode_ModelRun_ProductionCurve_rPrior (H:High,M:Medium,V:Very High)_PsiPrior(High:0.2,Low:0.9)_ObservationError(T=TRUE,F=False)_StartYearCatches

For example,

BSH_1_P_rH_psil0.9_sigT_60 : Brown Shrimp (BSH_) run number 1 (_1) using a Pella-Tomlinson surplus production curve (_P), high r prior (_rH), low initial depletion (_psil0.9) with additional observation error being estimated (_sigT) and a catch time series starting in 1960 (_60)

3.1.3 Model Diagnostics

Candidate models were assessed based on the following four criteria (Carvalho et al. 2021):

3.1.3.1 Model Convergence

The Geweke convergence diagnostic (CONV_gw) compares the mean of the first and last part of Markov chain to see if they are significantly different. Z scores near 0 (between -1.96 and 1.96) are considered acceptable (Geweke 1992).

Heidelberger and Welch stationarity diagnostic (CONV_hs) shows the iteration number from which the chain is considered to have converged and an associated p value, where the null hypothesis is that the sampled values come from a stationary distribution (Heidelberger and Welch 1983). 'Failure' of the stationarity test indicates that a longer MCMC run is needed. The Heidelberger and Welch half-width test (CONV_hw) checks whether the Markov chain sample size is adequate to estimate the mean values accurately (Heidelberger and Welch 1983).

3.1.3.2 Model Fit

Catch-per-unit-effort (CPUE) residuals runs test: CPUE indices pass the runs test (CPUE_rt_rand) if there is no evidence of a non-random residual pattern ($p > 0.05$). Any year where the residuals are larger than the threshold limit [3 standard deviations (sd) away from the mean (Anhøj and Olesen 2014)] fail the outlier test (CPUE_rt_outl).

3.1.3.3 Model Consistency

Retrospective analysis: This test checks for systematic bias in the stock status estimates. The procedure involves sequentially removing all data from the most recent period (i.e. peeling), refitting the model, and then comparing terminal year estimates of stock status [e.g. spawning stock biomass (SSB), fishing mortality (F)] to the full model. A guiding practice proposed by Hurtado-Ferro et al. (2015), suggests values of Mohn's rho (RETRO_) that fall outside a set range (-0.22 to 0.30) for shorter-lived species indicates an undesirable retrospective pattern. In addition, the direction of the retrospective bias has implications for characterizing risk associated with management advice.

Process error: The annual process error deviations should exhibit a stochastic pattern with a constant average centered around the zero (ProcB_mu) and 95% credibility intervals covering the zero value (ProcB_CI).

3.1.3.4 Prediction Skill

Hindcast cross-validation (Kell et al. 2016, 2021): this test is to check that the model has prediction skill of future states under alternative management scenarios. The procedure involves sequentially removing CPUE data from the most recent period, refitting the model with the remaining data, and then comparing known CPUE values (observations) to model estimates.

Mean Absolute Scaled Error (HX_MASE): The MASE score scales the mean absolute error of the prediction residuals to the mean absolute error of a naive in-sample prediction (i.e. equal to the last observed value). A score of 0.5 indicates that the model forecasts of CPUE values are twice as accurate as a naive in-sample prediction, indicating that the model has prediction skill. A score higher than 1 indicated that the model forecasts are no better than a random walk. If MASE < 1, the model has some level of prediction skill and passes the test.

3.1.4 Goodness of Fit

Deviance Information Criteria (DIC) was used for model selection purposes, where a lower value generally indicates a better model fit. Root-Mean-Squared-Error (RMSE) was used to quantitatively evaluate the randomness of model residuals. These criteria were used to determine the best model of those that passed the model diagnostic tests described in the previous section.

3.2 Empirical Dynamic Modelling (EDM)

Empirical Dynamic Modelling (EDM) uses lags of time series data to reconstruct the state-space of a system (Sugihara 1994; Sugihara et al. 2012; Munch et al. 2017, 2022). This form of modeling is particularly useful for short-lived species with chaotic population dynamics where drivers are often not observed directly, yet the information is embedded within the time series of abundance. Lags of abundance indices are used to reconstruct the full dynamics of the system without needing data on variables impacting abundance or specifying model form. Gaussian-Process EDM (GP-EDM) version 0.0.0.9010 on *GitHub* was used to fit the SEAMAP survey data aggregated at levels defined in *Section 2*. We also tested the inclusion of economic and environmental variables as covariates since they are hypothesized drivers of shrimp abundance where measurements do exist.

3.2.1 Model Configurations

3.2.1.1 Formulation with Fishery Removals

Gaussian Process regression was used to approximate the Brown Shrimp population delay-embedding map f

$$P\left[y_t \,\middle|\, f, (X_{t-m} - q * C_{t-m}), z, V_e\right] \sim Normal(f(X_{t-m} - q * C_{t-m}, z), V_e)$$

where the probability of observing abundance y at time t is dependent on the function approximation f, vector of abundance indices X with m lags ($X_{t-m} = x_{t-1}, \ldots, x_{t-m}$), optional covariates z, and process variance V_e. The delay embedding map defined above was expanded to include removals (C, catch or landings) scaled by a catchability parameter q which can be fit within or among populations. Here, catchability is a scalar used to translate units of landings into survey units. Covariates (z) can be included as direct drivers of abundance where measurements exist. Fitting to 'escapement', the composite variable $X_{t-m} - q * C_{t-m}$ is the number of individuals remaining after harvesting. GP-EDM with a single lag $m = 1$ can be thought of as a nonparametric production model (Thorson et al. 2014). f is dependent on the inverse length scales $\Phi = \phi_1, \ldots, \phi_{i=m+z}$ and pointwise prior variance τ and follows a Gaussian Process prior with mean zero and covariance function Σ, which assumes no relationship on the shape function.

$$P\left[f \,\middle|\, \Phi, \tau\right] \sim GP(0, \Sigma)$$

The covariance function Σ is defined for abundance y

$$\Sigma(y_t, y_s) = \tau * exp[-\Sigma_{i=1}^{m+z} \phi_i (X_{it} - X_{is})^2]$$

at times t and $s \in T$ where T is the time series length (Munch et al. 2022). The inverse length scale parameters ϕ and abundance observations X are provided for each $i = m + z$ where m is the lags of abundance and z is the covariates. This function is scaled by τ, and a prior is applied here that constrains the total variance of the predicted population size (y_{T+1}) to be less than twice the observed variance in $y_1, \ldots, y_T$. This prior specification for process and observed variances and length scale parameters are represented by

$$P[V_e, \tau, \Phi]$$

The covariance function and inverse length scales jointly control the degree of nonlinearity of the shape function f, where $\phi = 0$ indicates a flat relationship and a large estimate for ϕ indicates a higher degree of nonlinearity. The covariance function Σ can either tighten the relationship around the observed data, favoring a smaller length scale (i.e. a larger inverse length scale parameter) or relax the relationship, facilitating a smoother function with a larger length scale (smaller ϕ). Detailed GP prior specification for EDM variance and length scale parameters can be found in Munch et al. (2017).

An optional feature of GP-EDM is to assign a linear prior on f which can aid in grounding the population to 0 as the harvest rate, U, approaches 1 (i.e. the entire population is harvested). The linear prior option assumes that the mean function for the GP is linear with respect to the first input and fits the model on the residuals of

$$y_t = \beta_0 + \beta_1[x_{t-1} - q * c_{t-1}] + f(X_{t-m} - q * C_{t-m}, z)$$

where $[x_{t-1} - q * c_{t-1}]$ is first lag of escapement and f is the GP function approximation. If $y_t = log(x_{t+1}/x_t)$ and is backtransformed, this is equivalent to a Ricker model excluding f (Ricker 1954). In this case, we're working on deviations from growth under the assumed Ricker model. The model fits similarly to the previous configuration, but the primary difference can be observed outside of the range of observed data. This configuration helps linearly ground the fishery model abundance to zero as simulated removals approach the total population size. Without this prior, it's possible that outside of the observed range of the data, the abundance levels out to the flat prior where the population may never reach zero (and can result in extraordinarily high landings under simulated high harvest rates).

3.2.1.2 Embedding Dimension

EDM embedding dimension E is limited by the length T of the time series. An approximate maximum embedding dimension is $E \leq \sqrt{T}$. In the case of continuous seasonal data, the maximum embedding dimension is larger since the time series T is longer. Models were configured using Summer and Fall seasons as continuous time steps throughout a year and as a population-specific level within a hierarchical EDM, which will be explained in further detail below. The embedding dimension is defined as the number of population lags m (and covariates z if included) plus one, $E = m + z + 1$. For Brown Shrimp, the first year of the SEAMAP survey was 1987, resulting in 36 years of data, and a maximum embedding dimension of approximately 6 on an annual scale.

3.2.1.3 Hierarchical Model Scaling

Prior to fitting EDM models, all input data are standardized to a mean of 0 and standard deviation of 1. In the context of EDM, the term 'populations' is used to define data aggregations where information is expected to be informative. For Brown Shrimp, data aggregations and resulting populations that could be used to delineate levels of EDM are defined at the start of *Section 2*. For systems with multiple populations, these could be fit within a hierarchical EDM or independently.

In hierarchical models, the data must be scaled globally or locally across populations. For global scaling, the data across populations are expected to have the same mean. For Brown Shrimp, global scaling is likely inappropriate for most strata defined here. For example, we never expect the abundance of Large shrimp to equal the abundance of Small shrimp as would be implied by global scaling. Local scaling allows us to scale the data within the defined population time series of available data for each respective lag of population abundance or covariate. Both global and local scaling are applied within each predictor, not across all data. For example, each predictor is scaled to a mean of 0 and standard deviation of 1 for each lag and covariate. For global scaling, all data from all populations are used to scale the data; for local scaling, this is done within populations.

In independent models, definition of global or local scaling is obsolete because all data are scaled to a mean of 0 and standard deviation of 1. Independent models were tested for all data aggregations to ensure information was gained through the increased complexity and shared information from hierarchical models and with dynamic correlation.

3.2.1.4 Dynamic Correlation

Dynamic correlation ρ is defined as the degree to which the EDM population dynamics are correlated. This quantifies the similarity of population responses across predictor space and ranges from 0 to 1. Populations in hierarchical models will share the same embedding parameters and inverse length scale parameters (this includes models with $\rho = 0$, or independent dynamics). A dynamic correlation $\rho = 1$ means the dynamics of each population are identical. In other words, we assume that all delay vectors come from the same attractor. If fitting a single population or independent model, ρ reverts back to the mode of the prior, 0.5.

In hierarchical models, the dynamic correlation can be fixed or estimated. In cases where dynamic correlation is set to 0 within a hierarchical model, this will still yield different results when compared to independently fit models. This is because the hierarchical model shares information among the estimated length scale parameters ϕ for each embedding parameter.

3.2.1.5 Length Scale

Length scale parameters ϕ and the number of model inputs ($i = m + z$) define the complexity of the function represented by the GP. Each model input i incorporates an additional dimension of space, and their associated length scale parameter ϕ_i defines the wiggliness in that dimension. Low values of ϕ indicate stiff and mostly linear relationships, and large values of ϕ indicate more nonlinear relationships. A model with a single input and large ϕ_1 would have many degrees of freedom, while a model with many inputs but all ϕ_i close to 0 would have relatively few degrees of freedom (Tsai et al. 2024).

3.2.1.6 Data Transformations

Possible data transformations on the population are defined below. This is referred to as 'ytrans' in the GP-EDM R Package, but was defined as X_t above. This is the transformation that is applied before fitting the model.

- none: no transformation
- log: log transformation ($log(X_t)$)
- gr1: log difference transformation ($log(X_t/X_{t-1})$)
- gr2: log difference transformation on escapement ($log\big(X_t/(X_{t-1} - q * C_{t-1})\big)$)

3.2.1.7 Covariates

The underlying theory of EDM is that lags of the population have information on population drivers embedded within them (Munch et al. 2020). It is possible to include some covariates directly in EDM that are believed to influence population abundance. In the case of Gulf penaeid shrimp, economic conditions have had a massive impact on the domestic fishery, which in turn directly influences the amount of shrimp left in the water. Additionally, it has been hypothesized that environmental drivers such as salinity and temperature in the shrimp nursery grounds may have a direct impact on recruitment to the population the following year (Turley et al. 2023).

While covariates have the potential to improve model fits and short-term predictive accuracy, relying on lags of the population alone for estimating the biological MSY is simpler from an operational standpoint. Including covariates in the model requires making some assumption

about the future states of that covariate in projections, which cannot be done with high confidence in this context. In addition, some of the variables considered may contain some level of covariation which the model is not set up to account for in its present form.

3.2.1.8 Cross Validation

Two different cross validation approaches were explored to evaluate prediction accuracy: "leave time out" and "sequential". Prediction method "leave time out" leaves out all data points (i.e., survey data, catch, covariates) taken at the same time across all populations where population is specified within hierarchical models. The "sequential" prediction method leaves out all future time points across all populations where population is specified. In both of these methods, training data are iteratively omitted for the predictions, but the inverse length scales and variances used are those obtained using all of the training data under the originally fit model. We anticipate that "sequential" would perform worse when compared to "leave time out". Both cross validation approaches were applied to all model configurations, but ultimately the "sequential" method was preferred for model selection because our ultimate objective is to project landings and harvest rates into the future in order to accurately estimate the system's maximum sustainable yield for fishery management.

3.2.2 Goodness of Fit

Goodness of fit was measured through the estimation of R^2.

In sample fit statistics for each prediction method:

- R^2 - proportion of variance explained by model (independent or hierarchical)
- R^2_{pop} - proportion of variance explained for each population within a hierarchical model
- R^2_{scaled} - proportion of variance explained by a hierarchical model, centered and scaled by population means
- $rmse$ - root mean square error
- df - degrees of freedom, trace of the smoother matrix

Out-of-sample fit statistics for each prediction method:

- R^2_{out} - out-of-sample R^2
- R^2_{outpop} - out-of-sample R^2_{pop}
- $R^2_{outscaled}$ - out-of-sample R^2_{scaled}
- $rmse_{out}$ - out-of-sample $rmse$

These fit statistics measure the models' overall performance and ability to perform outside of the training data. Within hierarchical models, population-specific R^2_{pop} metrics measure the model's ability to track the individual populations. For example, a model may be able to track one population well, but may fit another poorly. These population-specific R^2_{pop} metrics were centered and scaled around their respective model means in the R^2_{scaled} fit statistics to more appropriately measure the overall model performance. Population-specific R^2_{pop} and R^2_{scaled}

statistics were compared to R^2 statistics obtained from independent model fits of each population to ensure that the complexity of the hierarchical model was warranted (i.e. improved overall prediction skill).

3.2.3 Estimated Parameters

Parameters estimated and priors specified in GP-EDM are defined below.

- $\phi_1 : \phi_i$ - length scale parameters for $1 : i$ where i is the total m lags and z covariates ($i = m + z$); priors are set such that the expected number of local extrema for each ϕ_i is 1 (Munch et al. 2017)
- V_e - process variance
- τ - pointwise prior variance in f
- ρ - dynamic correlation between populations where values range from 0 to 1, with 0- independent no correlation and 1- identical dynamics
- q- catchability scalar that translates the units of landings into units of survey CPUE

The relative magnitude of the pointwise prior variance τ and process noise V_e gives information on how important the function is relative to the noise. Process variance is represented as a percentage of the total variance, whereas the pointwise prior variance cannot be directly translated to variance percentage because it interacts with the length scale parameters. If the model is purely deterministic, $V_e = 0$ and $\tau \approx 1$. If the model is not fitting the data well, τ is small and the process variance is close to 1.

Catchability could be estimated jointly (q=shared) or separately for each population in each model configuration. In some instances, the model obtained very good fits, but estimated catchability $q = 0$ and ignored the landings altogether. For the purposes of our work here, the link to landings is critical. To select a representative model for estimating MSY, the models were filtered to exclude any model where catchability < 0.001 (where the observed catchability in the data were typically above 0.01).

3.2.4 Estimating Maximum Sustainable Yield (MSY) with EDM

Maximum Sustainble Yield (MSY) estimates were generated following the methodology outlined in Tsai et al. (2024). Harvest rates ranging from 0:1 were projected into the future and an average of the long-term dynamics were taken for each population, then added up to obtain estimates of long-term landings. These averages were used to identify the harvest rate that maximizes landings. Models that were configured seasonally required landings and associated harvest rates to be translated to annual scales. Translating catch from a seasonal to an annual time scale was fairly simple

$$C_t = 2 * C_{t/2}$$

where t is defined as one year here, and $t/2$ represents 2 seasonal steps per year. Annual harvest rate U_t was estimated from a seasonal harvest rate $U_{t/2}$ as

$$U_t = 1 - \left(1 - U_{t/2}\right)^2$$

where the new estimated harvest rate U_t captures the portion of the population (0:1) removed via landings over the course of a year. Here, the estimated long-term biomass associated with the rate of removals does not need to be changed. The annual harvest rate U_t was further translated to an annual fishing mortality rate $F_t = -ln(1 - U_t)$. This allows for the calculation of the more familiar benchmark F_t/F_{MSY}, which is a measure of overfishing (estimated to be occurring if $F_t/F_{MSY} > 1$).

3.2.5 Model Diagnostics

Models were diagnosed and deemed reliable based on a set of criteria defined below. This methodology worked well for all Gulf shrimp species assessed within SEDAR 87. These decisions were applied to 'no covariate' models, since assumptions on the cyclical nature of environmental variables and the relationship between harvest rate and economic variables would be required for projections. It was determined that these assumptions should be avoided for the purposes of defining biological maxima if possible. The projection period was initially set to 50 timesteps then extended to 80 to ensure the reference points had stabilized before taking an average. The duration over which to average was determined by the length of a cycle, which was typically driven by the seasonal time steps in the model if present. The estimate of MSY is sensitive to setting an appropriate projection period that ensures the population has stabilized and an appropriate save interval that ensures only complete cycles are clipped, the latter ensures the estimate is not biased high or low (as would be observed if the time step just outside of a completed cycle is increasing or decreasing, respectively).

3.2.5.1 Model Fitting Performance

Model performance was determined by considering the suite of Goodness of Fit parameters defined above. The top 30 models from the hierarchical overall R^2_{out} and top 30 models from the $R^2_{outscaled}$ were pulled, and any overlapping models were considered. The top 5 from each of these criteria and the top 5 aggregated Gulf-wide models were considered to evaluate what was gained from added complexity.

3.2.5.2 Model Projection Performance

Projection performance was evaluated to ensure models extrapolate to MSY in a reasonable way. Model selection was already performed with this goal in mind when relying on predictmethod=sequential to obtain fit statistics. Additional diagnostics were developed to cull out unreasonable models. This included removing models that maximized catch at $U = 1$, which generally happened when models would predict that the population returns to the flat prior outside of the observed range of the data. These models were often paired with unrealistically high catch estimates due to the coupling of extreme harvest rates with populations that did not always ground to zero. It is intuitively not sustainable to remove the entire population, so these models were removed. Unrealistically high estimates of MSY were defined as greater than ten times the highest historic landings.

3.2.5.3 Model Robustness

From the remaining set of models that (1) had good fits, (2) did not solve on a bound ($U = 1$), and (3) did not estimate MSY at greater than 10x historical landings records, a retrospective pattern analysis was carried out where 1 to 5 time steps were peeled back and MSY was re-

estimated. The Model Projection Performance selection criteria defined above were applied to each of these iterations. If any iteration failed, it was dropped from further consideration. This resulted in a final selection that balances model complexity and relative stability.

4. Stock Assessment Model Results

4.1 JABBA Results

4.1.1 Model Fit and Diagnostics

Diagnostic results for the top performing JABBA model runs are presented in *Figure 17*. Run 1 (BSH_1_P_rH_psil0.9_sigT_60) was the best performing model; it had the lowest DIC among the runs that passed the most diagnostic tests (i.e. all runs with 'high' or 'very high' priors on r). All candidate runs failed the CPUE outlier test, but that was due to a single data point (1988) falling outside the 'three-sigma limit' (*Figure 18*). Run 1 assumed a high r prior, low initial depletion and allowed for additional observation error to be estimated. Prior and posterior distributions resulting from that run are shown in *Figure 19*. All models had a MASE at or below 1, indicating that the average model forecasts are better than a naïve baseline prediction (*Figure 17*). The hindcasting cross-validation results for Run 1 show predictions within limits of the 95% credible intervals suggesting a good prediction skill (*Figure 20*). The model had difficulty fitting to recent years, as the catch size composition has likely shifted to target larger shrimp in recent years due to economic demand. This was introduced in *Section 2.4* and will be elaborated further in the Discussion. A retrospective analysis of derived benchmarks also showed notable consistency as years of the model were peeled back (*Figure 21*).

4.1.2 Estimated Parameters and Derived Quantities

Parameter estimates and associated uncertainty for all top performing JABBA models are shown in *Figure 22*, *Figure 19* and *Table 3*. The candidate models appeared relatively unaffected by the assumption placed on initial depletion (ψ). The highest levels of uncertainty were observed for runs that did not allow for additional observation error to be estimated (runs 16, 79, 82). These runs also showed a consistent pattern of retrospective bias in the absolute metrics of biomass and fishing mortality (though no retrospective bias in the relative management quantities B/B_{MSY} and F/F_{MSY}). These runs also failed the process error test with large variances being estimated. Most model posteriors did not deviate significantly from the priors as there was not much contrast in the data to inform the underlying surplus production model. Posterior distributions for the top model run (run 1) are shown in *Figure 19*. For that run, K and r appear well informed.

Overall, biomass trajectories were fairly similar across runs and uncertainty was highest at the start of the time series before the index enters the model and when CVs on landings are relatively high. Uncertainty was lowest in the first year when the index enters the model (*Figure 23*). Although estimates of K (*Table 4*) and the resulting surplus production function (*Figure 24*) were highly variable, estimates of MSY were remarkably consistent, around 90 million pounds of tails across all runs (*Table 4*). The estimated time series of B/B_{MSY} are also relatively similar across runs (*Figure 25*) with most models dipping below $B/B_{MSY} = 1$ in the late 1980's / early 1990's. Two clusters of models are visible. Runs 16, 79, and 82, which do not allow for additional observation error, estimate a higher but more uncertain B/B_{MSY} in the terminal year than the remainder of the models that allow for additional observation error to be estimated.

4.2 EDM Results

Over 7,500 model configurations were evaluated for Brown Shrimp to explore assumptions and ensure that results from the various iterations made sense. Up to the maximum embedding dimension was considered, with preference given to simpler models where possible. Estimated parameters, model fits, and projection capabilities are discussed below, resulting in the recommendation of a single model by the end of this section.

4.2.1 Model Configurations

Model configurations were examined to test assumptions and ensure results were as expected. The impact of using the 'sequential' method when defining the training dataset for prediction accuracy is shown in *Figure 26*, where 'leave-time-out' almost always yielded a higher out-of-sample R^2_{out} fit. In hierarchical models, large differences in population means could artificially inflate the R^2_{out} metric. Therefore, metrics were calculated to estimate goodness-of-fit that were centered and scaled around the population mean, $R^2_{scaledout}$. Hierarchical out-of-sample R^2_{out} generally yielded a higher value than the out-of-sample scaled by population-specific fits, $R^2_{scaledout}$. With these models, the goal is to fit and project each population within the model well, and $R^2_{scaledout}$ was the primary metric used to gauge model fits going forward.

Assumptions of global and local scaling were shown across all model configurations (embedding dimension, population, y transformation) using out-of-sample scaled $R^2_{scaledout}$ fit statistic. Local scaling generally yielded better fits compared to global scaling, which agrees with what we understand about these populations and their relative biomass (*Figure 27*). At this stage, only models with local scaling were considered for all configurations (except for models with one "population" where global scaling is inherent).

Some of the reported R^2 metrics were associated with models that ignored landings (i.e. $q \approx 0$). *Figure 28* shows the distribution of R^2 after these models were removed. From the set of models that account for landings, additional models were excluded due to the fact that they included covariates. In *Figure 29*, information can be inferred about the relative scales of population sizes and landings, where model configurations fit better to shared catchabilities (e.g. similar scales between population and landings) compared with models that assumed distinct catchabilities (e.g., different scales between populations and landings). This figure shows the model configurations that were analyzed for fit and eventual MSY estimation.

4.2.2 Model Fit and Residual Analysis

From the set of models described in the previous section, the procedures outlined in *Section 3.2.5.1* were applied, i.e. ranking the models by out-of-sample prediction accuracy for the model as a whole (R^2_{out}), scaled by population ($R^2_{outscaled}$), or both (*Section 3.2.2*). This resulted in 54 models going through MSY estimation and further model diagnostics (*Table 5*). These models had out-of-sample prediction accuracies ranging from 0.287 up to 0.828. Scaled population R^2_{scaled} metrics ranged from 0 up to 0.511, where a zero here would indicate that one population prediction was no better than a random forecast. These were overall very good fits to the data, and in-sample fit statistics were even greater.

4.2.3 Model Diagnostics

The subset of 54 models with the best fits was further reduced down to 5 models after testing for projection ability and model robustness as outlined in *Section 3.2.5* (*Table 6*). Some of these models were characterized by instability, with a few retrospective peels spiking up over 10x historical record landings. One model run that exceeded 5x historical landings was characterized with retrospective bias and was removed from consideration, where the results were trending upwards as years of data were peeled back (*Table 6*, Run BSH_G21023). The remaining 2 top performing models were size class models with a seasonal time step (strata=G, time=YEAR2). They were locally scaled, had a shared catchability parameter, and had the gradient transformation on escapement (ytrans=gr2) for the predictor variable (y_t). One of the models aggregated the Small and Medium size classes (embedding dimension of 4), and the model with Small, Medium, and Large separated had an embedding dimension of 5. These observations line up with how we understand EDM, in that an additional driver (distinction between Small and Medium shrimp), may require an additional lag to explain the variation in these data and to capture the dynamics of the more complex system with an additional dimension. Top performing model parameterizations are summarized below.

Run	Catchability	Population	Time Step	Lags	Scaling	Transformation
G10323	Shared	Size (Sm,L)	Seasonal	4	Local	gr2
G20023	Shared	Size (S,M,L)	Seasonal	5	Local	gr2

Variable harvest rate projections of CPUE by size class for the recommended model are shown in *Figure 30* - *32* and associated landings are shown in *Figure 33* - *35*. These data series were used to generate average biomass and landings under all harvest rates 0:1 in *Figure 36* and *Figure 37*. An additional consideration when interpreting these results, is there is no feedback loop here for fishing out a size class. The model with finer resolution of size classes (model BSH_G20023) and a more consistent harvest rate that achieves peak landings by size class (*Figure 36*) was recommended for providing estimates of MSY. This can be contrasted with *Figure 38* and *Figure 39*, where the harvest rate that maximizes 'Smedium' shrimp landings is realized at a harvest rate with no 'Large' shrimp landings remaining. The recommended model (BSH_G20023) has stable landings and a similar harvest rate for all size classes, which is expected to be more robust than a model that may have unaccounted negative feedback loops between size classes and future recruitment or growth across seasons. The total metrics for B_{MSY} and MSY are shown in *Figure 40* and *Figure 41*. B_{MSY} estimates were left in the seasonal (summer vs. fall+winter) scale, while MSY was translated to annual values and rates (vertical dashed lines in these two figures represent equivalent harvest rates on the seasonal and annual scales). The horizontal dotted line on the MSY figure shows the maximum landings ever caught by the fishery, 105.91 million pounds of tails in 1990, which is less than half the projected MSY here, 215.07 million pounds of tails.

4.2.4 Estimated Parameters and Derived Quantities

Estimated parameters from the top-performing model are shown in *Table 7*. The function-space complexity is defined jointly by the length scales, which define the degree of nonlinearity, and the covariance matrix, which can open up the ability of the model to vary within a smoother

space. The estimated length scale parameters ϕ_i are all less than 1, and are linear and smooth for all populations in the model (*Figure 42* - *46*). The pointwise prior variance was estimated to be 0.689 and exceeded the function process variance of 0.103, indicating an informative model. The dynamic correlation of the model was 0.957, indicating a high correlation between these data, which we can observe visually most easily in *Figure 2* and *Figure 5*. These populations had a shared catchability, $q = 0.402$, that translates fishery removals to the units of the SEAMAP survey (number of shrimp per trawl hour divided by million pounds of tails).

The R^2 statistics were similar for the overall and population-scaled metrics, both approximately 0.7 (*Table 8*). Out-of-sample $R^2_{out} = 0.485$ performed slightly better than the $R^2_{outscaled} = 0.364$ scaled by population means, which was expected since the magnitude of the differences in means can inflate the estimated model fit (this effect is removed in the $R^2_{outscaled}$ statistic). Derived population benchmarks and associated rates are shown in *Table 8* alongside model fit statistics. Annual MSY was estimated as 215.07 million pounds of tails, occurring at $F_{MSY} = 0.617$ where the population biomass at this rate is $B_{MSY} = 405.39$ million pounds of tails.

4.2.5 Fishing Mortality

Estimated fishing mortality rates through time are shown in *Table 9*. In 2022, the stock experienced <2% F_{MSY} and the stock size was >4x B_{MSY}. The highest rates of fishing mortality were observed in the 1980's prior to the economic collapse of the fleet due to aquaculture imports. Even at that time, the Brown Shrimp stock was not undergoing overfishing.

4.2.6 Biomass and Abundance Trajectories

Table 9 provides the time series of estimated biomass over time and associated reference points. The stock is not estimated to have undergone overfishing over the duration of the assessment timeframe (1987-2022) and is only estimated to have been overfished in 1988. This could be explained by the oscillating nature of this stock, where landings in 1988 were high, particularly for Large shrimp, but the population was in a trough, resulting in $B/B_{MSY} = 0.90$. During this year, under multiple model parameterizations, the stock was estimated as overfished. Given the oscillating nature of EDM, it is possible that when using averaged MSY projections, the true sustainable fishing levels in any given year could be above or below the average MSY, but it is not expected to be an issue unless the system begins chronically dipping below sustainable levels. Fits of the preferred model are shown in *Figure 47* and *Figure 48*.

5. Discussion

EDM is particularly suitable for studying populations that exhibit non-equilibrium dynamics and nonlinear state-dependent behavior (i.e. where interactions change over time and as a function of the system state). JABBA relies on very rigid SPM assumptions about stock and fishery dynamics that likely do not hold true for shrimp. EDM models performed very well and had high levels of prediction accuracy, therefore we recommend that EDM be used for providing management advice.

The JABBA models were generally well behaved but the results are limited by the general constraints of surplus production models which aggregate dynamics, not accounting for size- or spatial- differences. For Gulf Brown Shrimp, it is well documented that the fishery size

composition has changed through time (*Figure 49*). Global market conditions and increasing demand and prices for Large shrimp have shifted the composition of landings away from Small shrimp towards larger, more valuable shrimp. This apparent change in selectivity may bias results if unaccounted for. Additionally, as the domestic fleet consolidated, larger and more efficient vessels remained and could be trawling in a different habitat compared to the historic distribution of the fleet. These factors are justifications against using surplus production models that assume catch levels reflect only changes in stock abundance and that patterns of exploitation are primarily driven by shrimp availability rather than environmental or economic considerations.

EDM models showed very good diagnostics and prediction accuracy. The biomass of the Brown Shrimp population and removals were modeled and predicted well. Size-structure was included through the use of populations within a hierarchical GP-EDM, which further improved the model fits. One caveat of the current EDM configurations explored here is there is no feedback loop from smaller size classes to larger size classes. For example, there is no penalty on Medium and Large shrimp for removing too many Small too early under a high harvest rate. In reality, a harvest rate that maximizes Small shrimp may cause a negative feedback loop on Large shrimp that is not accounted for here. In some simulations, the peak landings for Large shrimp was at a much lower harvest rate than the Small, and when aggregating these size classes to approximate a total MSY, it is feasible that the realized Large shrimp landings would be lower due to the lack of Small shrimp growing out to Medium and Large size classes. Accounting for this negative feedback loop through mixed-age configurations is possible (Dolan et al. 2023), but it is complicated by mixing landings across calendar years to fit to population escapement, which would markedly increase management complexity, perhaps unnecessarily. The recommended EDM configuration here maximized landings of all size classes at approximately the same harvest rate, removing the need to explore this caveat further.

EDM was able to capture the cyclical nature of shrimp population abundance, resulting in a more accurate population model. Lags of the population retain information on sometimes immeasurable drivers, including abundance of predators and some environmental influences. Direct inclusion of environmental and economic covariates improved model fits further, but they were not used in the final model because additional assumptions would be required on the future state of the industry and environment. Furthermore, relationships between the simulated harvest rates and these covariates would need to be addressed, and may respond in unexpected ways. Given the goal of providing a biological MSY estimate for this fishery, biological models only were used for this purpose. The models with covariates may serve other purposes and could be used to predict year-ahead abundance and landings more accurately than the model with lags alone, particularly for those tied to economic drivers.

Finally, providing management advice for this fishery using static estimates of MSY may not be appropriate due to the highly cyclical nature of this stock which is not fully captured in a long-term average. The model itself captures the dynamics, but the methods to obtain MSY through a long-run average do not. In high productivity years, the fishery may be able to harvest more than the average MSY, in low productivity years, it may push the stock into an ‘overfished’ status (see: 1988). To provide management advice for a population with such large estimates of sustainable landings, the long run average should be used to ensure that the stock does not undergo overfishing. In the event of improved economic conditions where the fleet expands and prevailing environmental conditions change, this assumption could be revisited and management

advice could be provided on a finer scale. Updating the model with seasonal inputs as they become available could account for the peaks and dips in productivity, allowing the fleet to take advantage of high productivity years or potentially sit out low productivity years. Because the fleet is mainly limited by the economics of the fishery, these kinds of model explorations are recommended as a future research recommendation.

6. Research Recommendations

The models provided in this report are sufficient to provide management advice for the stock. However, should future research funding become available, we have provided suggestions below.

Potential improvements to the modeling framework include accounting for removal of shrimp as it pertains to harvest rates that are optimized at varying size classes. Creating a feedback loop that appropriately represents the removal of larger shrimp that may not contribute to future generations as well as the removal of smaller shrimp that may not grow into Large shrimp should be accounted for. Sensitivities of these potential feedback loops and their impact on estimating optimal harvest rates should be investigated in both directions (i.e. Large to Small and Small to Large impacts).

Additional research into covariates may also be investigated. Direct inclusion of covariates generally resulted in improved model fits and could likely improve forecasting efficiency for trends of abundance. To forecast MSY, covariates would need to be projected into the future. For environmental covariates, the cyclical nature of these trends would need to be captured. For economic covariates, the relationship with projected harvest rates would need to be explicitly defined.

As funding for scientific surveys is becoming increasingly sparse, implications of using an average of 2019/21 for missing summer 2020 SEAMAP data and resulting effects on model diagnostics should be investigated. EDM performs best on continuous, long time series of data, and quantifying implications of future gaps in survey data would be valuable.

7. Acknowledgements

The SEDAR 87 Gulf Brown Shrimp Benchmark Assessment would not have been possible without the efforts of the numerous NMFS, SEFSC, SERO, and GMFMC staff along with the many academic, research partners, and stakeholders involved throughout the Gulf listed in *Section 1.1.3*.

8. References

Ailloud, L., M. Stevens, B. Turley, A. Pollack, and D. Hanisko. 2025. Developing a fishery-independent index of relative abundance for Gulf of Mexico Brown Shrimp using VAST. Page 26pp. SEDAR, SEDAR87-AP-03, North Charleston, SC.

Anhøj, J., and A. V. Olesen. 2014. *Run charts revisited: A simulation study of run chart rules for detection of non-random variation in health care processes*. PLoS One 9(11):Article e113825.

Atkinson, S., A. Lowther, K. Dettloff, and S. Smith. 2024. Gulf of Mexico commercial brown, pink and white shrimp landings. Page 38pp. SEDAR, SEDAR87-DW-06, North Charleston, SC.

Browder, J. A., Z. Zein-Eldin, M. M. Criales, M. B. Robblee, S. Wong, T. L. Jackson, and D. Johnson. 2002. *Dynamics of pink shrimp (Farfantepenaeus duorarum) recruitment potential in relation to salinity and temperature in Florida Bay*. Estuaries 25(6):1355–1371.

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

Dolan, T. E., E. P. Palkovacs, T. L. Rogers, and S. B. Munch. 2023. *Age structure augments the predictive power of time series for fisheries and conservation*. Canadian Journal of Fisheries and Aquatic Sciences:cjfas-2022-0219.

Froese, R., H. Winker, G. Coro, N. Demirel, A. C. Tsikliras, D. Dimarchopoulou, G. Scarcella, M. L. D. Palomares, M. Dureuil, and D. Pauly. 2019. *Estimating stock status from relative abundance and resilience*. ICES Journal of Marine Science 77(2):527–538.

Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. Pages 169–193 *in* A. P. D. J. M. Bernardo J. O. Berger and A. F. M. Smith, editors. Bayesian statistics. Clarendon Press, Oxford.

Griffith, D., C. Liese, M. Travis, M. Freeman, and D. Records. 2023. Social dimensions of Gulf of Mexico shrimping. Page 12pp. SEDAR, SEDAR87-DW-15, North Charleston, SC.

Haddon, M. 2021. Surplus production models. Pages Chapter 7 Using r for modelling and quantitative methods in fisheries. Chapman; Hall, London, UK.

Heidelberger, P., and P. D. Welch. 1983. Simulation run length control in the presence of an initial transient. Operations Research 31:1109–1144.

Hurtado-Ferro, F., C. S. Szuwalski, J. L. Valero, S. C. Anderson, C. J. Cunningham, K. F. Johnson, R. Licandeo, C. R. McGilliard, C. C. Monnahan, M. L. Muradian, K. Ono, K. A. Vert-Pre, A. R. Whitten, and A. E. Punt. 2015. *Looking in the rear-view mirror: Bias and retrospective patterns in integrated, age-structured stock assessment models*. ICES Journal of Marine Science 72(1):99–110.

Kell, L. T., A. Kimoto, and T. Kitakado. 2016. *Evaluation of the prediction skill of stock assessment using hindcasting*. Fisheries Research 183:119–127.

Kell, L. T., R. Sharma, T. Kitakado, H. Winker, I. Mosqueira, M. Cardinale, and D. Fu. 2021. *Validation of stock assessment methods: Is it me or my model talking?* ICES Journal of Marine Science 78(6):2244–2255.

Liese, C. 2024. Price indices for shrimp imports and Gulf of Mexico shrimp landings by size and season. Page 10pp. SEDAR, SEDAR87-AP-2, North Charleston, SC.

Lowther, A. 2023. Shrimp import data. Page 2pp. SEDAR, SEDAR87-DW-10, North Charleston, SC.

Munch, S. B., A. Brias, G. Sugihara, and T. L. Rogers. 2020. *Frequently asked questions about nonlinear dynamics and empirical dynamic modelling*. ICES Journal of Marine Science 77(4):1463–1479.

Munch, S. B., V. Poynor, and J. L. Arriaza. 2017. *Circumventing structural uncertainty: A Bayesian perspective on nonlinear forecasting for ecology*. Ecological Complexity 32:134–143.

Munch, S. B., T. L. Rogers, and G. Sugihara. 2022. *Recent developments in empirical dynamic modelling*. Methods in Ecology and Evolution:1–14.

Pella, J. J., and P. K. Tomlinson. 1969. *A generalized stock production model*. Inter-American Tropical Tuna Commission Bulletin 13:419–496.

Pollack, A. G., and D. S. Hanisko. 2024. Brown, white and pink shrimp abundance indices from SEAMAP groundfish surveys in the northern Gulf of Mexico. Page 113pp. SEDAR, SEDAR87-DW-13, North Charleston, SC.

Ricker, W. 1954. Stock and recruitment. J. Fish. Res. Bd. Canada 11(5):559--623.

Schlenker, L. S., C. Stewart, J. Rock, N. Heck, and J. W. Morley. 2023. *Environmental and climate variability drive population size of annual penaeid shrimp in a large lagoonal estuary*. PLOS ONE 18(5):e0285498.

SEDAR87 Data Workshop Report. 2023. Page 94pp. SEDAR, SEDAR87, North Charleston, SC.

Sugihara, G. 1994. Nonlinear forecasting for the classification of natural time series.

Sugihara, G., R. May, H. Ye, C. Hsieh, E. Deyle, M. Fogarty, and S. Munch. 2012. *Detecting Causality in Complex Ecosystems*. Science 338(6106):496–500.

Thorson, J. T., K. Ono, and S. B. Munch. 2014. *A Bayesian approach to identifying and compensating for model misspecification in population models*. Ecology 95(2):329–341.

Tsai, C.-H., S. B. Munch, M. D. Masi, and M. H. Stevens. 2024. *Empirical dynamic modeling for sustainable benchmarks of short-lived species*. ICES Journal of Marine Science:1–12.

Turley, B., L. Ailloud, and M. Stevens. 2023. Development of estuarine environmental indices for SEDAR 87 Gulf of Mexico white, pink, and brown shrimp stock assessment. Page 15pp. SEDAR, SEDAR87-AP-01, North Charleston, SC.

Winker, H., F. Carvalho, and M. Kapur. 2018. *JABBA: Just Another Bayesian Biomass Assessment*. Fisheries Research 204:275–288.

Zink, I. C., J. A. Browder, D. Lirman, and J. E. Serafy. 2018. *Pink shrimp Farfantepenaeus duorarum spatiotemporal abundance trends along an urban, subtropical shoreline slated for restoration*. PLOS ONE 13(11):e0198539.

9. Tables

Table 1: SEAMAP CPUE in number of shrimp per trawl hour for Brown Shrimp by size and season.

Year	Season	Large	Medium	Small
1987	Summer	7.15	61.59	135.94
1987	Fall	8.58	23.54	17.23
1988	Summer	3.45	40.02	65.94
1988	Fall	10.53	15.68	10.28
1989	Summer	19.72	187.27	104.88
1989	Fall	16.65	38.18	15.59
1990	Summer	12.72	133.25	133.50
1990	Fall	13.02	35.72	26.83
1991	Summer	28.17	163.16	188.50
1991	Fall	18.28	45.74	29.15
1992	Summer	9.45	54.91	60.34
1992	Fall	14.91	37.30	16.65
1993	Summer	6.64	38.55	59.20
1993	Fall	9.67	33.92	31.06
1994	Summer	11.21	88.46	133.71
1994	Fall	14.92	35.79	20.05
1995	Summer	14.16	136.05	152.24
1995	Fall	17.31	55.82	25.61
1996	Summer	6.80	37.48	94.90
1996	Fall	10.82	32.86	20.78
1997	Summer	6.18	30.94	92.48
1997	Fall	12.02	38.70	26.24
1998	Summer	7.57	82.02	176.89
1998	Fall	14.31	33.21	17.84
1999	Summer	12.62	110.40	161.81
1999	Fall	13.28	35.08	43.64
2000	Summer	18.35	137.73	199.81
2000	Fall	17.96	45.84	30.49
2001	Summer	14.85	122.23	81.89
2001	Fall	13.40	43.12	25.44
2002	Summer	13.35	112.17	130.22
2002	Fall	14.00	38.32	34.50
2003	Summer	15.78	114.21	292.44

Table 1 Continued: SEAMAP CPUE in number of shrimp per trawl hour for Brown Shrimp by size and season.

Year	Season	Large	Medium	Small
2003	Fall	22.54	34.33	15.83
2004	Summer	35.23	140.01	112.72
2004	Fall	24.32	36.10	23.29
2005	Summer	14.62	87.57	89.84
2005	Fall	25.88	41.47	48.23
2006	Summer	48.20	310.37	243.29
2006	Fall	48.07	47.53	45.61
2007	Summer	27.79	130.60	151.33
2007	Fall	33.02	42.89	20.79
2008	Summer	23.24	126.62	203.78
2008	Fall	53.63	73.34	58.25
2009	Summer	43.73	200.94	166.50
2009	Fall	79.16	72.96	24.84
2010	Summer	88.75	264.00	129.60
2010	Fall	103.36	78.78	10.23
2011	Summer	85.28	309.20	152.37
2011	Fall	84.88	49.64	15.36
2012	Summer	62.16	206.58	166.56
2012	Fall	63.44	88.52	13.61
2013	Summer	41.63	92.55	157.58
2013	Fall	47.18	53.40	20.97
2014	Summer	18.19	59.18	116.18
2014	Fall	60.34	79.91	31.62
2015	Summer	34.83	179.27	200.21
2015	Fall	69.01	82.80	32.28
2016	Summer	47.72	123.54	100.24
2016	Fall	45.61	51.88	18.89
2017	Summer	34.20	142.58	122.11
2017	Fall	44.38	81.76	38.13
2018	Summer	34.43	152.43	86.72
2018	Fall	51.09	60.85	32.05
2019	Summer	24.48	86.84	187.10
2019	Fall	36.54	55.57	30.76
2020	Summer	36.70	195.85	276.67
2020	Fall	58.16	98.59	39.78

Table 1 Continued: SEAMAP CPUE in number of shrimp per trawl hour for Brown Shrimp by size and season.

Year	Season	Large	Medium	Small
2021	Summer	48.92	304.86	366.23
2021	Fall	43.31	107.40	30.71
2022	Summer	67.08	314.12	137.17
2022	Fall	37.78	100.90	32.62

Table 2: Landings of Brown Shrimp in millions of pounds of tails by size and season, where Fall landings include the following year's Winter landings due to the associated timing of the SEAMAP survey.

Year	Season	Large	Medium	Small
1987	Summer	6.72	32.72	33.17
1987	Fall	10.62	7.65	1.56
1988	Summer	8.57	32.82	23.84
1988	Fall	10.81	5.52	0.72
1989	Summer	8.32	36.57	28.78
1989	Fall	13.01	9.13	1.16
1990	Summer	10.47	35.65	40.41
1990	Fall	12.32	8.41	0.72
1991	Summer	10.61	27.34	29.70
1991	Fall	12.49	6.60	1.45
1992	Summer	4.37	22.28	22.20
1992	Fall	12.14	8.05	1.91
1993	Summer	3.71	16.92	31.25
1993	Fall	9.66	6.56	0.65
1994	Summer	5.51	22.21	21.99
1994	Fall	11.81	6.25	1.36
1995	Summer	6.21	20.03	34.10
1995	Fall	11.19	6.05	0.74
1996	Summer	4.69	20.84	33.18
1996	Fall	10.95	6.17	0.72
1997	Summer	3.44	16.33	29.48
1997	Fall	9.05	7.95	1.69
1998	Summer	4.24	24.01	34.41
1998	Fall	12.03	7.22	2.05
1999	Summer	6.27	20.11	38.47
1999	Fall	9.95	6.43	2.24
2000	Summer	9.60	28.18	37.08
2000	Fall	12.33	8.71	1.13
2001	Summer	7.91	28.32	38.96
2001	Fall	9.54	5.48	0.95
2002	Summer	5.90	20.82	32.84
2002	Fall	8.21	6.90	1.87

Table 2 Continued: Landings of Brown Shrimp in millions of pounds of tails by size and season, where Fall landings include the following year's Winter landings due to the associated timing of the SEAMAP survey.

Year	Season	Large	Medium	Small
2003	Summer	5.89	22.77	39.91
2003	Fall	9.75	6.96	1.08
2004	Summer	6.97	16.49	37.75
2004	Fall	10.03	5.41	0.54
2005	Summer	5.58	12.51	25.56
2005	Fall	10.08	5.38	1.92
2006	Summer	9.33	21.84	29.38
2006	Fall	19.37	8.64	0.30
2007	Summer	8.93	17.86	27.62
2007	Fall	11.97	5.52	0.47
2008	Summer	6.86	11.36	17.77
2008	Fall	12.99	5.08	0.82
2009	Summer	11.40	17.97	18.89
2009	Fall	20.18	7.58	0.41
2010	Summer	4.94	8.05	10.39
2010	Fall	11.89	8.57	1.58
2011	Summer	7.43	17.83	24.25
2011	Fall	15.21	8.37	1.02
2012	Summer	8.21	18.65	14.96
2012	Fall	12.75	8.81	1.06
2013	Summer	8.13	14.52	23.12
2013	Fall	12.13	7.82	1.12
2014	Summer	4.84	14.59	25.83
2014	Fall	13.20	9.24	1.43
2015	Summer	11.92	13.21	17.72
2015	Fall	12.50	7.51	0.89
2016	Summer	9.64	10.18	13.78
2016	Fall	9.76	5.39	2.08
2017	Summer	9.03	12.13	16.11
2017	Fall	9.09	7.54	2.89
2018	Summer	10.23	13.78	24.66
2018	Fall	12.02	8.67	0.77

Table 2 Continued: Landings of Brown Shrimp in millions of pounds of tails by size and season, where Fall landings include the following year's Winter landings due to the associated timing of the SEAMAP survey.

Year	Season	Large	Medium	Small
2019	Summer	6.50	7.95	12.20
2019	Fall	8.13	5.45	0.49
2020	Summer	6.99	11.81	8.45
2020	Fall	8.33	5.05	0.37
2021	Summer	4.16	12.14	10.97
2021	Fall	5.87	8.25	0.63
2022	Summer	4.34	7.92	7.03
2022	Fall	5.27	5.89	0.43

Table 3: Brown Shrimp parameter estimates from JABBA where Runs are described using unique identifiers (1:90), P indicates a Pella-Tomlinson surplus production curve with estimated shape parameter m, r is the relative level of the intrinsic rate of growth prior (M-Medium, H- High, V- Very High), psi is the initial depletion prior (0.9 low, 0.2 high), sig indicates whether additional observation error tau2 is estimated (T/F), and the last two numbers are the start year of the landings (1960). Median parameter estimates are provided with lower and upper credible intervals. K is reported in million lb tail weight.

Run	Parameter	Estimate	LCI.95	UCI.95
BSH_1_P_rH_psil0.9_sigT_60	K	266.48	207.41	335.41
BSH_1_P_rH_psil0.9_sigT_60	r	0.88	0.67	1.10
BSH_1_P_rH_psil0.9_sigT_60	q	0.01	0.00	0.01
BSH_1_P_rH_psil0.9_sigT_60	psi	0.89	0.54	1.28
BSH_1_P_rH_psil0.9_sigT_60	sigma2	0.00	0.00	0.02
BSH_1_P_rH_psil0.9_sigT_60	tau2	0.10	0.05	0.20
BSH_1_P_rH_psil0.9_sigT_60	m	0.88	0.60	1.33
BSH_4_P_rM_psil0.9_sigT_60	K	372.24	237.03	656.62
BSH_4_P_rM_psil0.9_sigT_60	r	0.55	0.35	0.87
BSH_4_P_rM_psil0.9_sigT_60	q	0.01	0.00	0.01
BSH_4_P_rM_psil0.9_sigT_60	psi	0.89	0.55	1.31
BSH_4_P_rM_psil0.9_sigT_60	sigma2	0.00	0.00	0.02
BSH_4_P_rM_psil0.9_sigT_60	tau2	0.11	0.05	0.23
BSH_4_P_rM_psil0.9_sigT_60	m	0.73	0.45	1.16
BSH_16_P_rM_psil0.9_sigF_60	K	303.87	170.86	1,120.18
BSH_16_P_rM_psil0.9_sigF_60	r	0.56	0.26	0.96
BSH_16_P_rM_psil0.9_sigF_60	q	0.01	0.00	0.01
BSH_16_P_rM_psil0.9_sigF_60	psi	0.88	0.55	1.31
BSH_16_P_rM_psil0.9_sigF_60	sigma2	0.04	0.01	0.05
BSH_16_P_rM_psil0.9_sigF_60	tau2	0.00	0.00	0.01
BSH_16_P_rM_psil0.9_sigF_60	m	0.45	0.26	0.98
BSH_49_P_rV_psil0.9_sigT_60	K	172.43	125.88	272.69
BSH_49_P_rV_psil0.9_sigT_60	r	1.36	1.02	1.88
BSH_49_P_rV_psil0.9_sigT_60	q	0.01	0.01	0.02
BSH_49_P_rV_psil0.9_sigT_60	psi	0.89	0.55	1.29
BSH_49_P_rV_psil0.9_sigT_60	sigma2	0.00	0.00	0.01
BSH_49_P_rV_psil0.9_sigT_60	tau2	0.09	0.04	0.21
BSH_49_P_rV_psil0.9_sigT_60	m	0.89	0.59	1.62

Table 3 Continued: Brown Shrimp parameter estimates from JABBA where Runs are described using unique identifiers (1:90), P indicates a Pella-Tomlinson surplus production curve with estimated shape parameter m, r is the relative level of the intrinsic rate of growth prior (M- Medium, H- High, V- Very High), psi is the initial depletion prior (0.9 low, 0.2 high), sig indicates whether additional observation error tau2 is estimated (T/F), and the last two numbers are the start year of the landings (1960). Median parameter estimates are provided with lower and upper credible intervals. K is reported in million lb tail weight.

Run	Parameter	Estimate	LCI.95	UCI.95
BSH_73_P_rH_psil0.2_sigT_60	K	229.60	172.69	425.35
BSH_73_P_rH_psil0.2_sigT_60	r	0.92	0.66	1.18
BSH_73_P_rH_psil0.2_sigT_60	q	0.01	0.00	0.01
BSH_73_P_rH_psil0.2_sigT_60	psi	0.27	0.08	0.60
BSH_73_P_rH_psil0.2_sigT_60	sigma2	0.00	0.00	0.02
BSH_73_P_rH_psil0.2_sigT_60	tau2	0.10	0.04	0.21
BSH_73_P_rH_psil0.2_sigT_60	m	0.78	0.47	1.35
BSH_79_P_rH_psil0.2_sigF_60	K	229.04	150.85	639.26
BSH_79_P_rH_psil0.2_sigF_60	r	0.79	0.49	1.16
BSH_79_P_rH_psil0.2_sigF_60	q	0.01	0.00	0.02
BSH_79_P_rH_psil0.2_sigF_60	psi	0.24	0.10	0.60
BSH_79_P_rH_psil0.2_sigF_60	sigma2	0.04	0.01	0.05
BSH_79_P_rH_psil0.2_sigF_60	tau2	0.00	0.00	0.01
BSH_79_P_rH_psil0.2_sigF_60	m	0.48	0.29	1.20
BSH_82_P_rM_psil0.2_sigF_60	K	386.91	210.91	800.41
BSH_82_P_rM_psil0.2_sigF_60	r	0.46	0.28	0.80
BSH_82_P_rM_psil0.2_sigF_60	q	0.01	0.00	0.01
BSH_82_P_rM_psil0.2_sigF_60	psi	0.24	0.10	0.57
BSH_82_P_rM_psil0.2_sigF_60	sigma2	0.04	0.03	0.05
BSH_82_P_rM_psil0.2_sigF_60	tau2	0.00	0.00	0.01
BSH_82_P_rM_psil0.2_sigF_60	m	0.49	0.31	0.86
BSH_85_P_rV_psil0.2_sigT_60	K	166.55	129.90	235.01
BSH_85_P_rV_psil0.2_sigT_60	r	1.47	0.91	1.97
BSH_85_P_rV_psil0.2_sigT_60	q	0.01	0.01	0.02
BSH_85_P_rV_psil0.2_sigT_60	psi	0.23	0.09	0.56
BSH_85_P_rV_psil0.2_sigT_60	sigma2	0.00	0.00	0.01
BSH_85_P_rV_psil0.2_sigT_60	tau2	0.10	0.04	0.20
BSH_85_P_rV_psil0.2_sigT_60	m	0.97	0.62	1.39

Table 4: Brown Shrimp reference points from selected JABBA models in Table 3. K, Bmsy and MSY are reported in million lb tail weight.

Run	K	Bmsy	Fmsy	MSY
BSH_1_P_rH_psil0.9_sigT_60	266.48	91.46	0.98	89.96
BSH_4_P_rM_psil0.9_sigT_60	372.24	115.54	0.77	88.66
BSH_16_P_rM_psil0.9_sigF_60	303.87	69.12	1.33	90.95
BSH_49_P_rV_psil0.9_sigT_60	172.43	61.20	1.48	90.60
BSH_73_P_rH_psil0.2_sigT_60	229.60	75.75	1.18	89.29
BSH_79_P_rH_psil0.2_sigF_60	229.04	55.38	1.65	91.17
BSH_82_P_rM_psil0.2_sigF_60	386.91	92.94	1.00	93.16
BSH_85_P_rV_psil0.2_sigT_60	166.55	59.94	1.51	90.69

Table 5: Brown Shrimp fit statistics for top performing models where run names are described by strata A:H, species, start year, landings units, shared catchability b (T/F), population, time step (YEAR2 is seasonal), embedding dimension E, scaling (global vs. local), and y transformations (log, gr1, gr2, none).

Run	R2_out	R2_outscaled
A211_BSH1987_CPUEtailmp_bshareF_GULFYEAR_E3_global_ytransnone	0.340	
A218_BSH1987_CPUEtailmp_bshareF_GULFYEAR_E4_global_ytransnone	0.323	
A20016_BSH1987_CPUEtailmp_bshareF_GULFYEAR_E5_global_ytransnone	0.322	
A127_BSH1987_CPUEtailmp_bshareF_GULFYEAR_E3_global_ytranslog	0.307	
A134_BSH1987_CPUEtailmp_bshareF_GULFYEAR_E4_global_ytranslog	0.287	
C21032_BSH1987_CPUEtailmp_bshareF_SIZEYEAR_E5_local_ytransnone	0.828	0.388
C21016_BSH1987_CPUEtailmp_bshareT_SIZEYEAR_E5_local_ytransnone	0.826	0.384
C10435_BSH1987_CPUEtailmp_bshareF_SIZEYEAR_E3_local_ytransnone	0.820	0.371
C10218_BSH1987_CPUEtailmp_bshareT_SIZEYEAR_E4_local_ytransnone	0.814	0.339
G10890_BSH1987_CPUEtailmp_bshareF_SEAS_SIZEYEAR_E4_local_ytransnone	0.808	0.389
C10211_BSH1987_CPUEtailmp_bshareT_SIZEYEAR_E3_local_ytransnone	0.803	0.340
G10442_BSH1987_CPUEtailmp_bshareT_SEAS_SIZEYEAR_E4_local_ytransnone	0.800	0.359
G21040_BSH1987_CPUEtailmp_bshareF_SEAS_SIZEYEAR_E5_local_ytranslog	0.786	0.347
C10267_BSH1987_CPUEtailmp_bshareF_SIZEYEAR_E3_local_ytranslog	0.785	0.333
G10876_BSH1987_CPUEtailmp_bshareF_SEAS_SIZEYEAR_E3_local_ytransnone	0.783	0.342
G21016_BSH1987_CPUEtailmp_bshareT_SEAS_SIZEYEAR_E5_local_ytransgr1	0.729	0.336
G10554_BSH1987_CPUEtailmp_bshareF_SEAS_SIZEYEAR_E4_local_ytranslog	0.791	0.332
G10106_BSH1987_CPUEtailmp_bshareT_SEAS_SIZEYEAR_E4_local_ytranslog	0.785	0.300
G21008_BSH1987_CPUEtailmp_bshareT_SEAS_SIZEYEAR_E5_local_ytranslog	0.782	0.318
C10274_BSH1987_CPUEtailmp_bshareF_SIZEYEAR_E4_local_ytranslog	0.781	0.238
C21004_BSH1987_CPUEtailmp_bshareT_SIZEYEAR_E5_local_ytranslog	0.776	0.245
G21032_BSH1987_CPUEtailmp_bshareT_SEAS_SIZEYEAR_E5_local_ytransnone	0.775	0.194
G21064_BSH1987_CPUEtailmp_bshareF_SEAS_SIZEYEAR_E5_local_ytransnone	0.771	0.290
G10428_BSH1987_CPUEtailmp_bshareT_SEAS_SIZEYEAR_E3_local_ytransnone	0.768	0.302
D10442_BSH1987_CPUEtailmp_bshareF_SIZE_AREAYEAR_E4_local_ytransnone	0.758	0.286
C10043_BSH1987_CPUEtailmp_bshareT_SIZEYEAR_E3_local_ytranslog	0.758	0.204
G10540_BSH1987_CPUEtailmp_bshareF_SEAS_SIZEYEAR_E3_local_ytranslog	0.752	0.278
D10435_BSH1987_CPUEtailmp_bshareF_SIZE_AREAYEAR_E3_local_ytransnone	0.748	0.297
D10211_BSH1987_CPUEtailmp_bshareT_SIZE_AREAYEAR_E3_local_ytransnone	0.744	0.282
C21020_BSH1987_CPUEtailmp_bshareF_SIZEYEAR_E5_local_ytranslog	0.743	0.301
D10218_BSH1987_CPUEtailmp_bshareT_SIZE_AREAYEAR_E4_local_ytransnone	0.740	0.282
G21024_BSH1987_CPUEtailmp_bshareT_SEAS_SIZEYEAR_E5_local_ytransgr2	0.740	0.270
C10050_BSH1987_CPUEtailmp_bshareT_SIZEYEAR_E4_local_ytranslog	0.738	0.270
C21012_BSH1987_CPUEtailmp_bshareT_SIZEYEAR_E5_local_ytransgr2	0.736	0.000

Table 5 Continued: Brown Shrimp fit statistics for top performing models where run names are described by strata A:H, species, start year, landings units, shared catchability b (T/F), population, time step (YEAR2 is seasonal), embedding dimension E, scaling (global vs. local), and y transformations (log, gr1, gr2, none).

Run	R2_out	R2_outscaled
G10092_BSH1987_CPUEtailmp_bshareT_SEAS_SIZEYEAR_E3_local_ytranslog	0.732	0.123
G21023_BSH1987_CPUEtailmp_bshareT_SIZEYEAR2_E5_local_ytransgr2	0.628	0.511
G10435_BSH1987_CPUEtailmp_bshareT_SIZEYEAR2_E4_local_ytransnone	0.714	0.503
G21031_BSH1987_CPUEtailmp_bshareT_SIZEYEAR2_E5_local_ytransnone	0.706	0.497
F421_BSH1987_CPUEtailmp_bshareT_AREA_ASSESSYEAR2_E3_local_ytransnone	0.599	0.481
E20064_BSH1987_CPUEtailmp_bshareF_SEASONYEAR_E5_local_ytransnone	0.657	0.479
E890_BSH1987_CPUEtailmp_bshareF_SEASONYEAR_E4_local_ytransnone	0.660	0.477
E20032_BSH1987_CPUEtailmp_bshareT_SEASONYEAR_E5_local_ytransnone	0.651	0.464
E442_BSH1987_CPUEtailmp_bshareT_SEASONYEAR_E4_local_ytransnone	0.654	0.462
E20040_BSH1987_CPUEtailmp_bshareF_SEASONYEAR_E5_local_ytranslog	0.645	0.455
E554_BSH1987_CPUEtailmp_bshareF_SEASONYEAR_E4_local_ytranslog	0.647	0.452
E20008_BSH1987_CPUEtailmp_bshareT_SEASONYEAR_E5_local_ytranslog	0.638	0.440
E106_BSH1987_CPUEtailmp_bshareT_SEASONYEAR_E4_local_ytranslog	0.641	0.437
G20047_BSH1987_CPUEtailmp_bshareF_SIZEYEAR2_E5_local_ytransgr1	0.427	0.416
G10323_BSH1987_CPUEtailmp_bshareT_SIZEYEAR2_E4_local_ytransgr2	0.550	0.409
F20015_BSH1987_CPUEtailmp_bshareT_AREA_ASSESSYEAR2_E5_local_ytransgr1	0.443	0.370
E428_BSH1987_CPUEtailmp_bshareT_SEASONYEAR_E3_local_ytransnone	0.606	0.367
E876_BSH1987_CPUEtailmp_bshareF_SEASONYEAR_E3_local_ytransnone	0.606	0.367
G20023_BSH1987_CPUEtailmp_bshareT_SIZEYEAR2_E5_local_ytransgr2	0.485	0.364
C21024_BSH1987_CPUEtailmp_bshareF_SIZEYEAR_E5_local_ytransgr1	0.729	0.356

Table 6: Retrospective analysis of the Brown Shrimp MSY estimates for top tier performing models with increasing peels. _0 indicates the base model, and _1:5 indicates 1 through 5 time steps of data peeled back. The maximum landings throughout the history of the fishery is 105.91 million pound of tails, and MSY_factor is the amount of times MSY is over this value. MSY_drop indicates whether the average MSY estimate was greater than 5 or 10 times the historical high, and F_drop indicates that the model solved at harvest rate U=1 and was excluded from further consideration. MSY, and Bmsy are in millions of pounds of tails. Run details are included in the previous table.

Run	MSY	BMSY_mp	MSY_factor	MSY_drop5	MSY_drop10	F_drop
BSH_G21023_0	560.61	403.97	5.29	1	0	0
BSH_G21023_1	567.79	409.14	5.36	1	0	0
BSH_G21023_2	636.95	445.87	6.01	1	0	0
BSH_G21023_3	682.58	464.53	6.44	1	0	0
BSH_G21023_4	718.63	475.85	6.79	1	0	0
BSH_G21023_5	727.73	481.88	6.87	1	0	0
BSH_G10435_0	797.10	813.70	7.53	1	0	0
BSH_G10435_1	919.45	866.40	8.68	1	0	0
BSH_G10435_2	1,379.09	1,299.53	13.02	1	1	0
BSH_G10435_3	597.27	584.64	5.64	1	0	0
BSH_G10435_4	985.47	754.50	9.30	1	0	0
BSH_G10435_5	804.44	679.57	7.60	1	0	0
BSH_G21031_0	726.11	741.24	6.86	1	0	0
BSH_G21031_1	755.01	711.45	7.13	1	0	0
BSH_G21031_2	1,219.94	1,149.56	11.52	1	1	0
BSH_G21031_3	607.46	620.53	5.74	1	0	0
BSH_G21031_4	971.04	743.45	9.17	1	0	0
BSH_G21031_5	813.86	687.57	7.68	1	0	0
BSH_G10323_0	371.83	337.40	3.51	0	0	0
BSH_G10323_1	398.06	325.08	3.76	0	0	0
BSH_G10323_2	410.61	324.51	3.88	0	0	0
BSH_G10323_3	440.11	336.96	4.16	0	0	0
BSH_G10323_4	439.59	336.56	4.15	0	0	0
BSH_G10323_5	436.77	334.40	4.12	0	0	0
BSH_G20023_0	215.07	405.39	2.03	0	0	0
BSH_G20023_1	217.66	410.20	2.06	0	0	0
BSH_G20023_2	219.29	383.76	2.07	0	0	0
BSH_G20023_3	233.54	357.61	2.21	0	0	0
BSH_G20023_4	228.08	349.24	2.15	0	0	0
BSH_G20023_5	232.40	379.59	2.19	0	0	0

Table 7: Brown Shrimp parameter estimates for the top performing model.

Parameter	BSH_G20023
Catchability	0.402
DynamicCorrelation	0.957
LengthScale1	0.204
LengthScale2	0.251
LengthScale3	0.017
LengthScale4	0.026
LengthScale5	0.001
PointwisePriorVariance	0.689
ProcessVariance	0.103

Table 8: Brown Shrimp MSY estimates for the top performing model.

Statistic	BSH_G20023
MSY_mptails	215.069
Fmsy	0.617
Umsy_annual	0.460
Umsy_seasonal	0.265
Bmsy_mp	405.394
df	28.529
R2	0.729
R2Scaled	0.700
R2_outsample	0.485
R2Scaled_outsample	0.364

Table 9: Brown Shrimp status through time based on benchmarks from the recommended model. Landmp- landings in millions of pound of tails, Frate- fishing mortality rate, Best_mp- estimate of population size in millions of pound of tails, FFmsy- Frate relative to Fmsy, BBmsy- Best relative to Bmsy.

Year	landmp	Frate	Best_mp	FFmsy	BBmsy
1987	92.43	0.16	632.28	0.26	1.56
1988	82.28	0.26	363.13	0.42	0.90
1989	96.96	0.11	951.47	0.17	2.35
1990	107.97	0.13	883.70	0.21	2.18
1991	88.20	0.08	1,177.30	0.13	2.90
1992	70.95	0.16	481.78	0.26	1.19
1993	68.76	0.17	445.59	0.27	1.10
1994	69.15	0.10	757.00	0.16	1.87
1995	78.33	0.08	998.54	0.13	2.46
1996	76.54	0.16	506.87	0.27	1.25
1997	67.93	0.14	514.11	0.23	1.27
1998	83.96	0.11	825.92	0.17	2.04
1999	83.46	0.09	937.90	0.15	2.31
2000	97.04	0.09	1,120.45	0.15	2.76
2001	91.15	0.13	749.01	0.21	1.85
2002	76.56	0.09	852.61	0.15	2.10
2003	86.36	0.07	1,232.36	0.12	3.04
2004	77.20	0.09	925.07	0.14	2.28
2005	61.03	0.08	765.65	0.13	1.89
2006	88.86	0.05	1,849.48	0.08	4.56
2007	72.37	0.07	1,011.53	0.12	2.50
2008	54.87	0.04	1,341.21	0.07	3.31
2009	76.43	0.05	1,463.83	0.09	3.61
2010	45.43	0.03	1,679.33	0.04	4.14
2011	74.12	0.04	1,734.12	0.07	4.28
2012	64.44	0.04	1,495.54	0.07	3.69
2013	66.83	0.07	1,028.70	0.11	2.54
2014	69.13	0.08	909.52	0.13	2.24
2015	63.75	0.04	1,489.38	0.07	3.67
2016	50.84	0.05	965.41	0.09	2.38
2017	56.79	0.05	1,152.78	0.08	2.84
2018	70.12	0.07	1,039.32	0.11	2.56
2019	40.71	0.04	1,048.57	0.06	2.59
2020	41.00	0.02	1,756.58	0.04	4.33
2021	42.02	0.02	2,243.64	0.03	5.53
2022	30.87	0.02	1,716.53	0.03	4.23

10. Figures

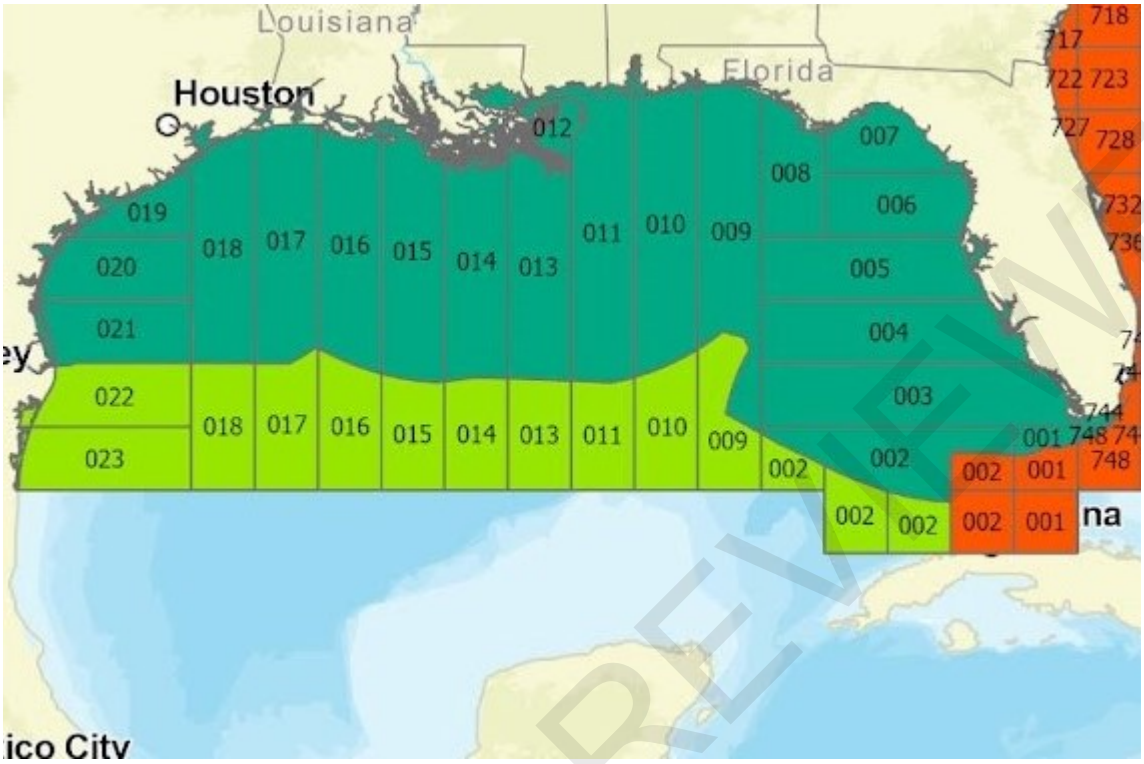

Figure 1: Map of the Gulf of America, where dark green is the Gulf defined by Gulf of Mexico Fishery Management Council boundaries, light green is Gulf international waters, and red is typically managed by the South Atlantic Fishery Management Council. Fishing areas 001 and 002 in their entirety were included in the analyses here per the recommendation of WP-06.

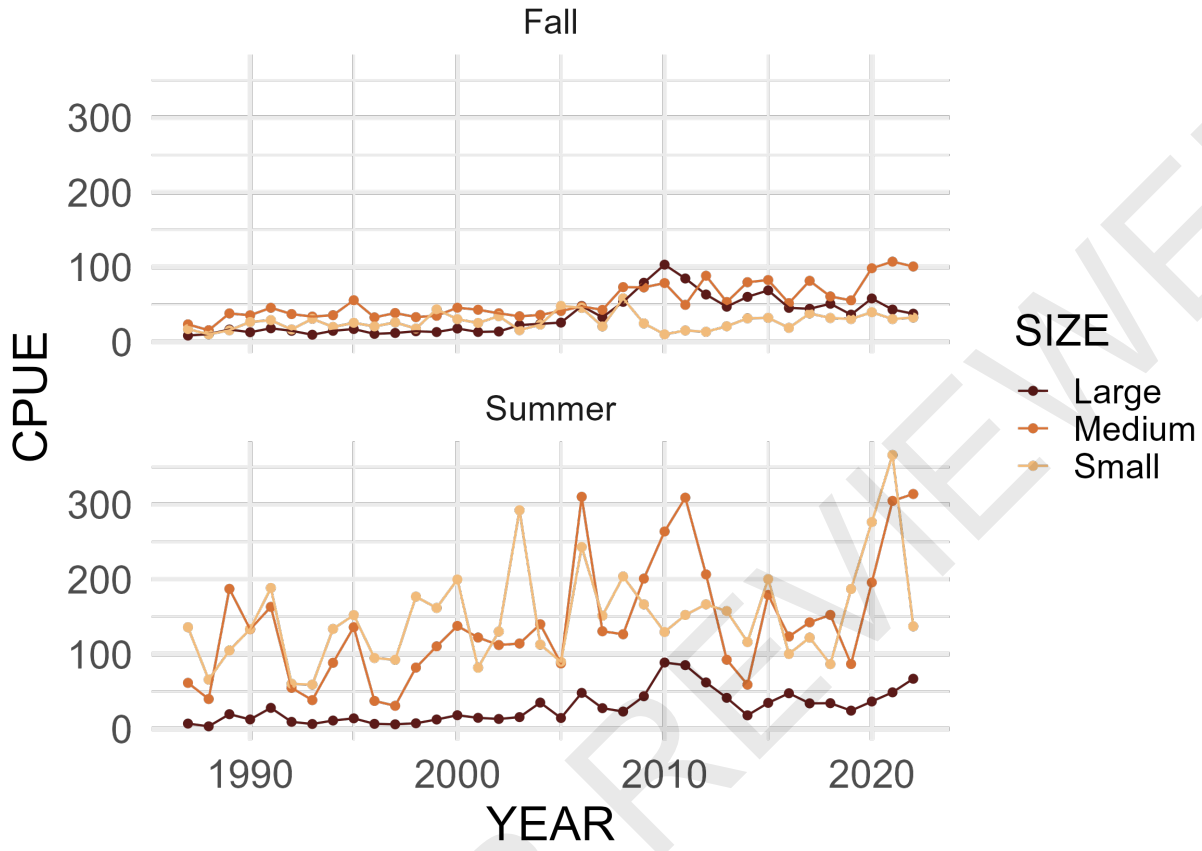

Figure 2: Brown Shrimp CPUE separated by size and season.

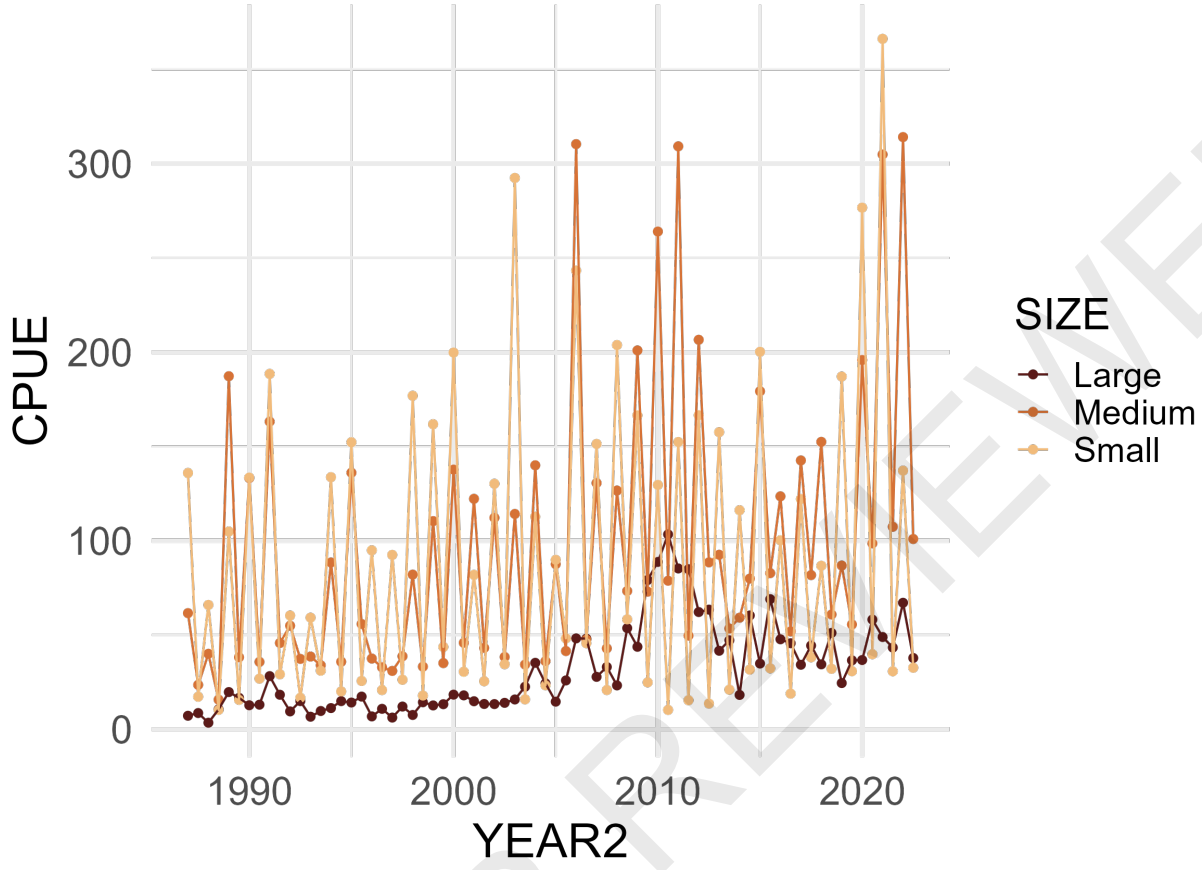

Figure 3: Size-stratified Brown Shrimp CPUE with continuous seasonal oscillations.

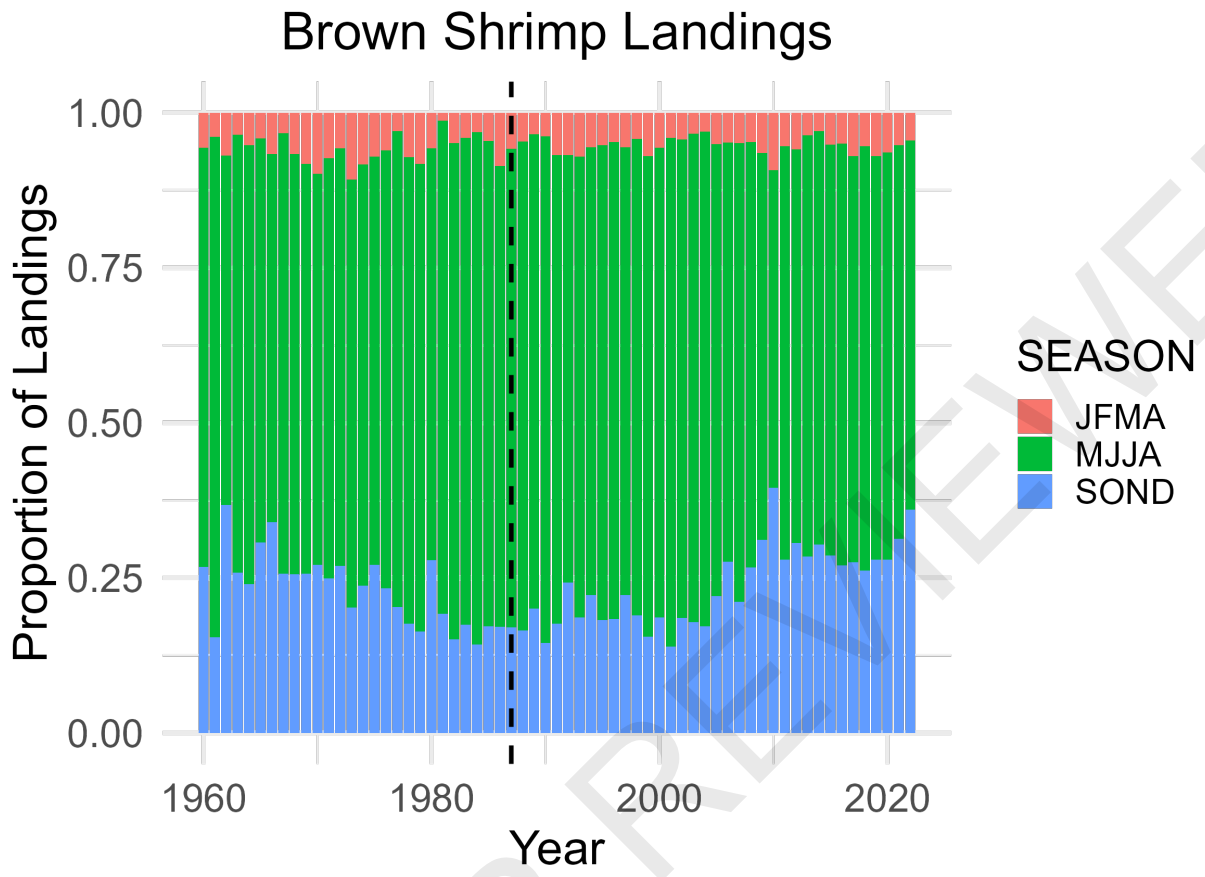

Figure 4: Seasonal distribution of landings, where the minimal landings from Winter (JFMA) of the previous year were aggregated with Fall (SOND) to approximately match the timing of removals lined up with the SEAMAP survey. Brown Shrimp receive a fresh pulse of recruits annually in the Summer.

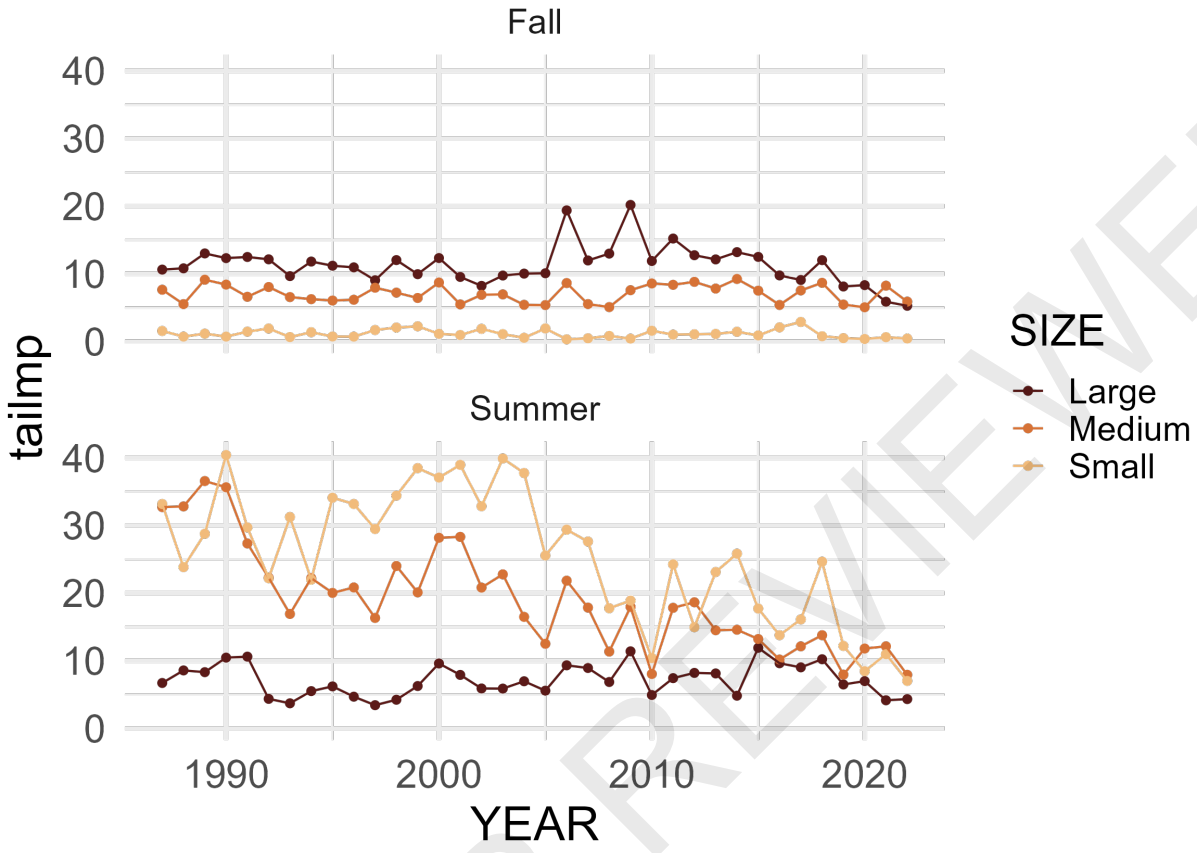

Figure 5: Brown Shrimp landings in millions of pounds of tails separated by size and season.

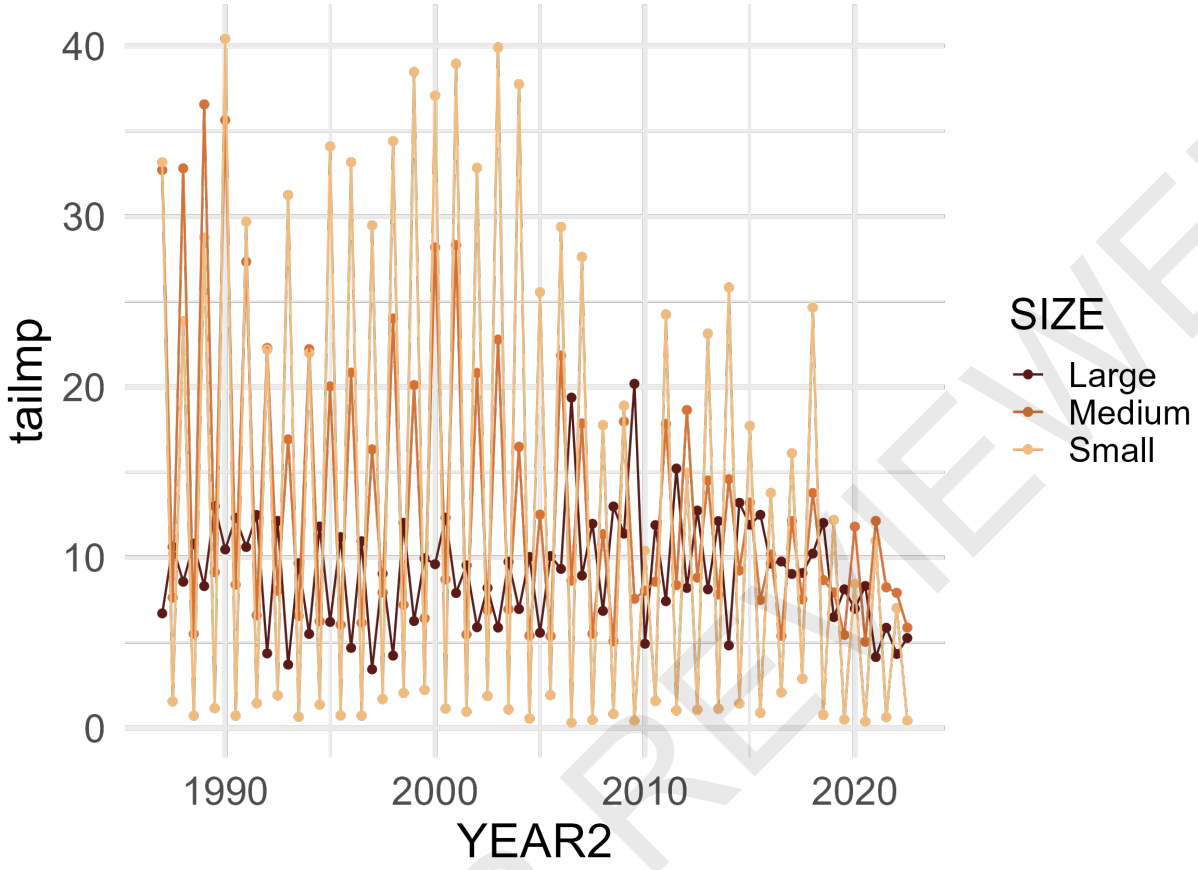

Figure 6: Size-stratified Brown Shrimp landings with continuous seasonal oscillations.

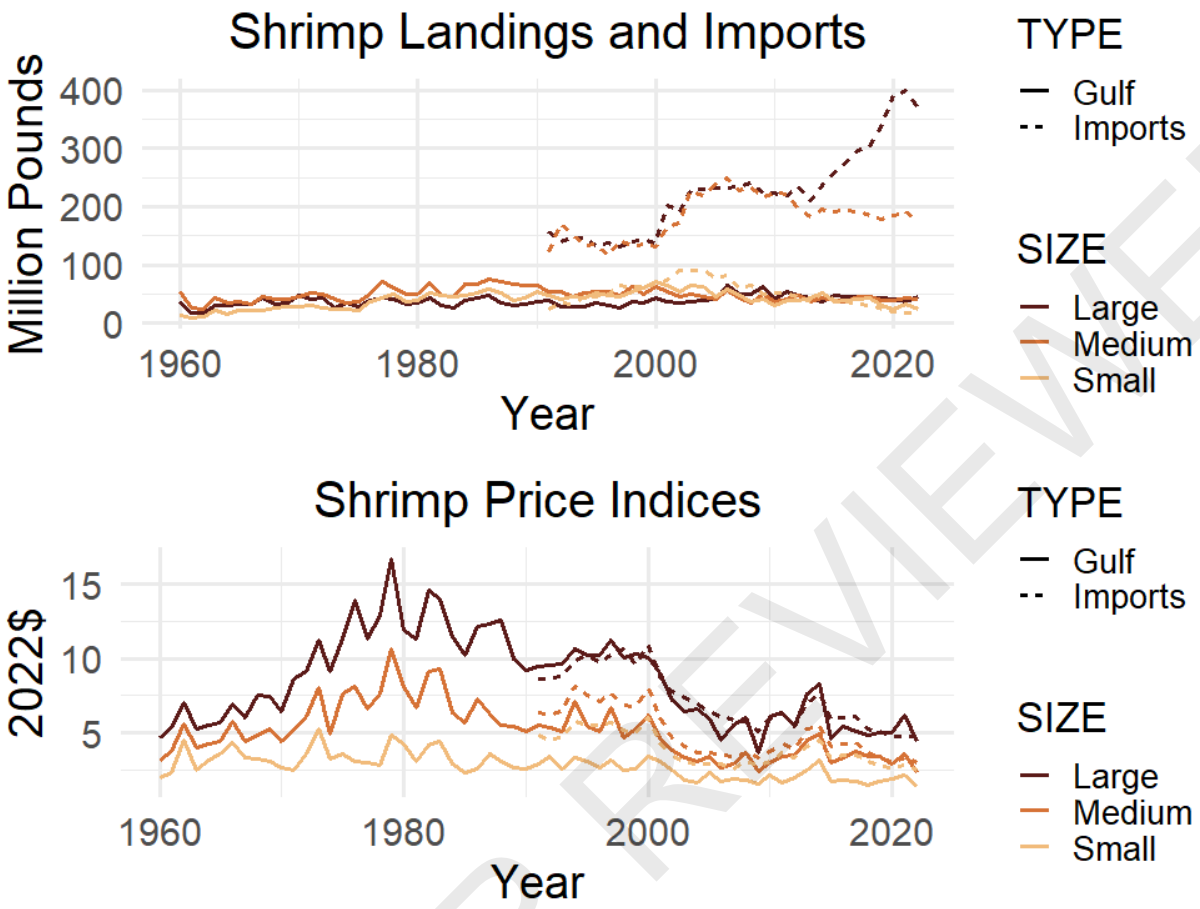

Figure 7: Domestic Gulf shrimp landings compared to global imports into the US by size category (top panel). This increase in supply has resulted in a crash of the ex-vessel price and domestic price index by size category, with all sizes decreasing, but Large yielding the highest amount (bottom panel).

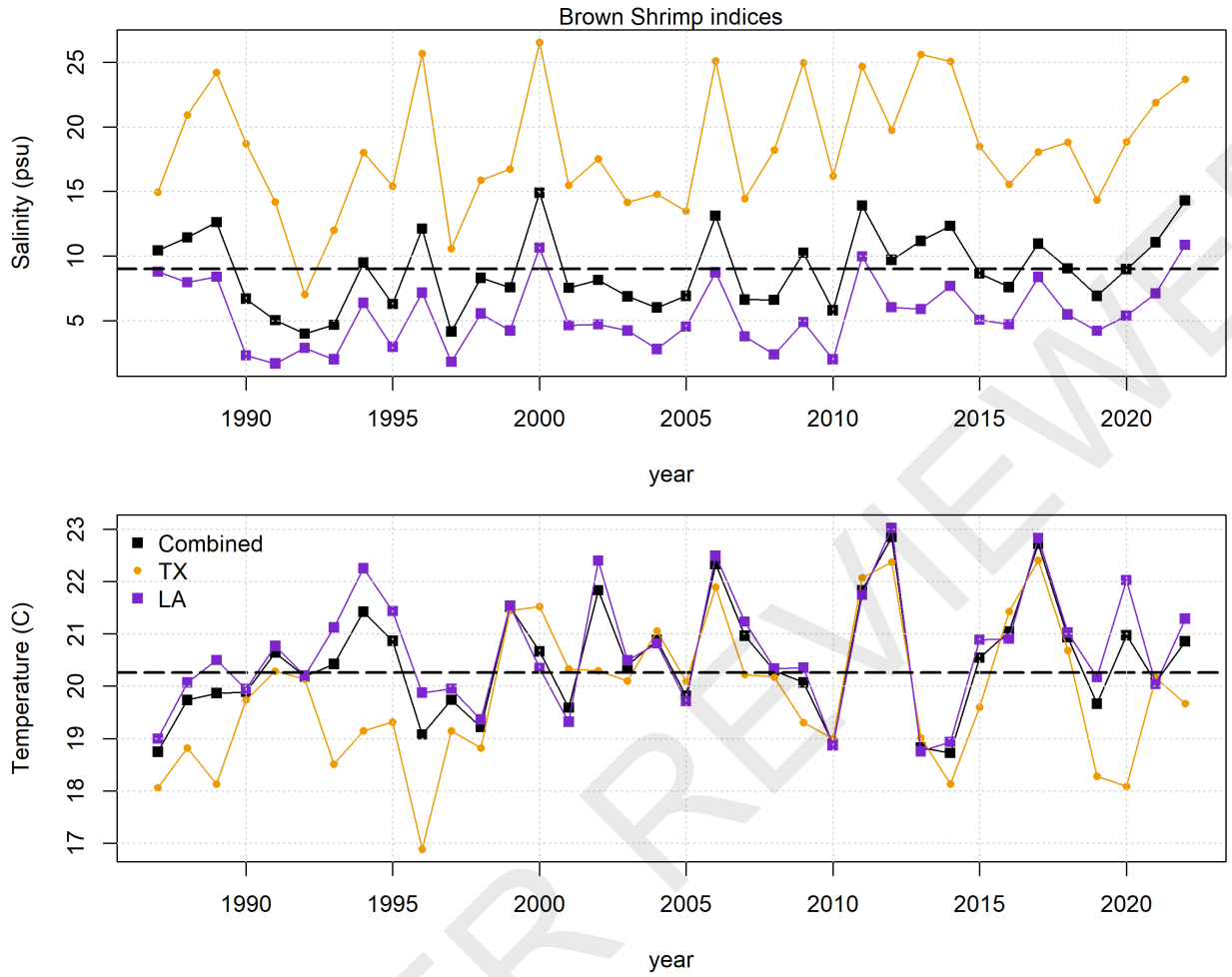

Figure 8: Brown Shrimp combined TX and LA environmental indices, truncated when TX data become available in 1987.

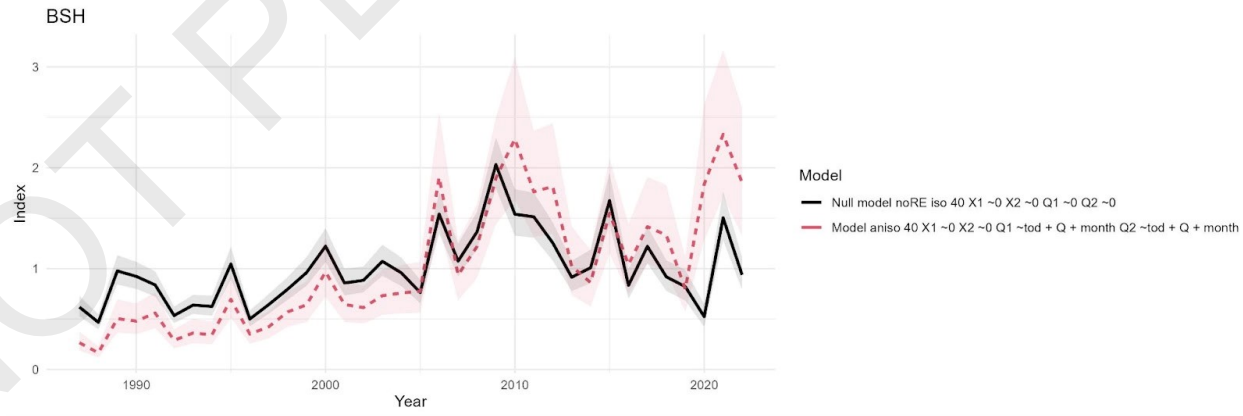

Figure 9: Final VAST index (red dashed line) and associated 95% confidence interval (red shading) incorporated into the JABBA model.

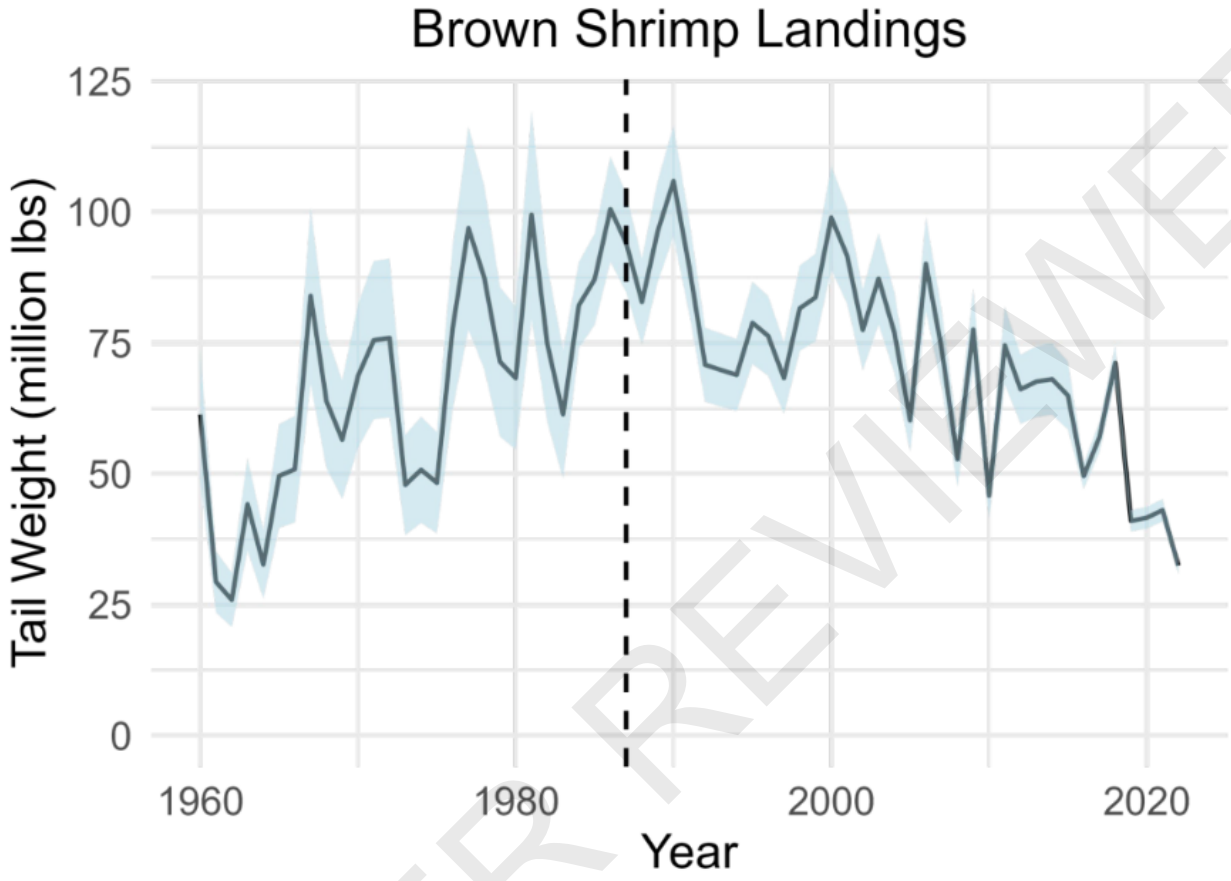

Figure 10: Final landings (blue line) and associated error (blue shading) input into JABBA. The dashed line indicates the start year of the index of relative abundance.

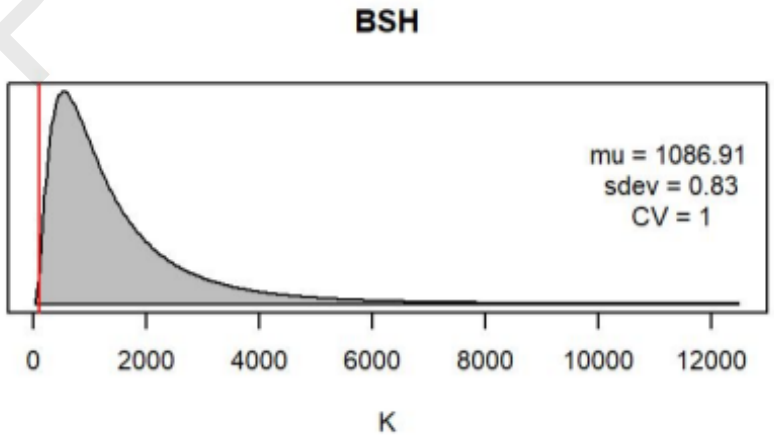

Figure 11: JABBA prior for carrying capacity, K, for all model configurations.

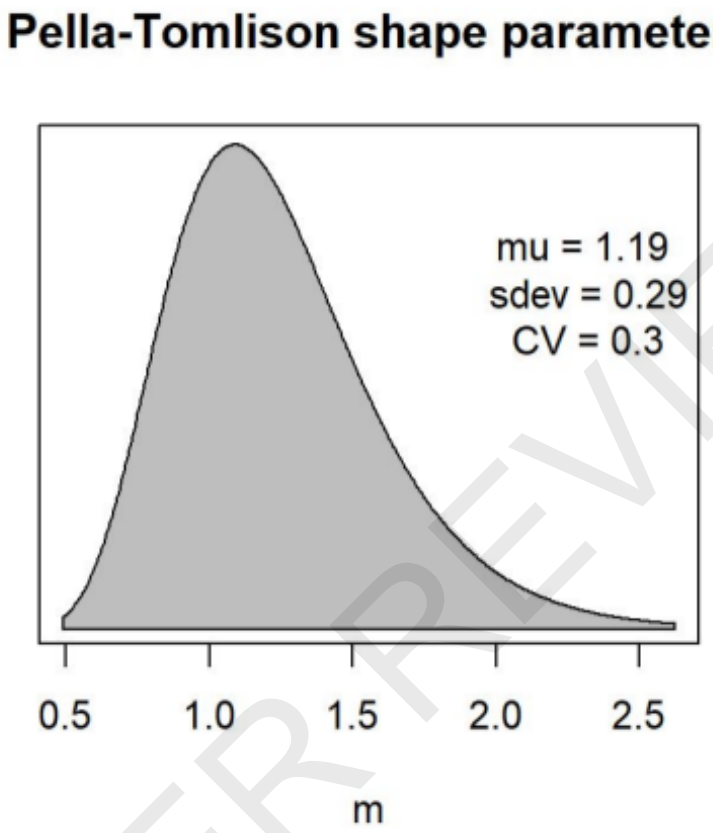

Figure 12: JABBA prior for Pella Tomlinson production function shape parameter, m, for all model configurations.

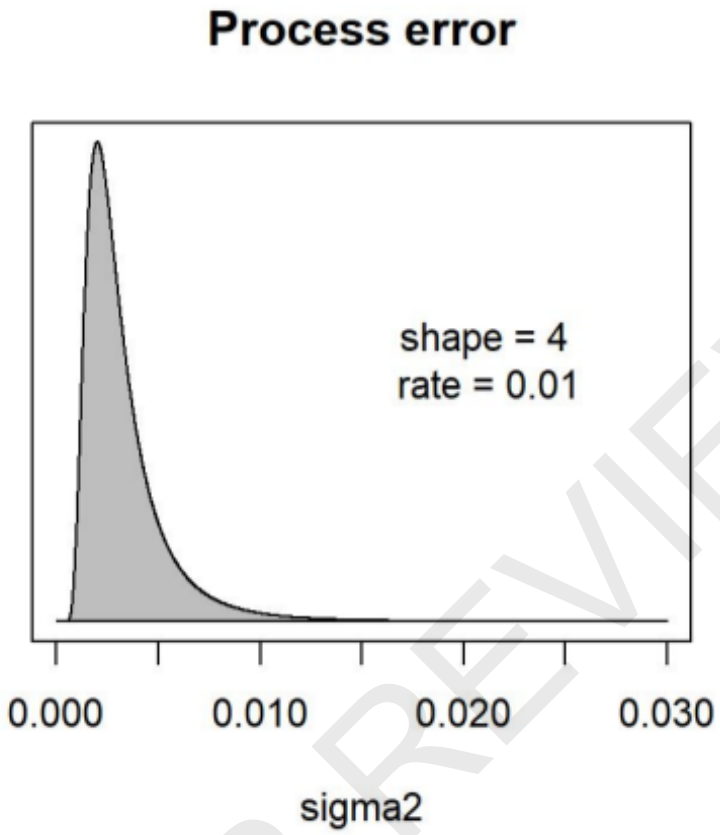

Figure 13: JABBA prior for process error variance for all model configurations.

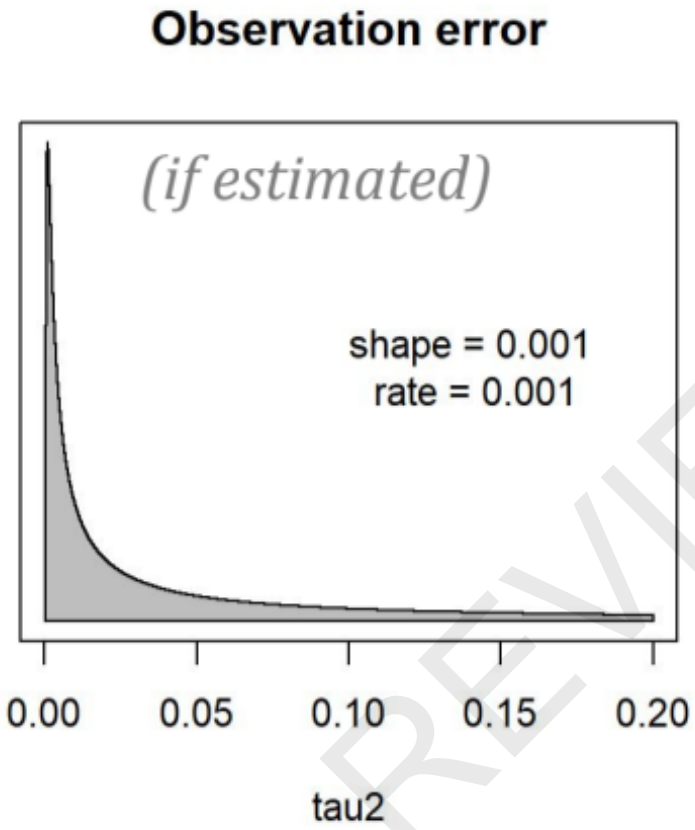

Figure 14: JABBA prior for observation error variance for all model configurations where estimated.

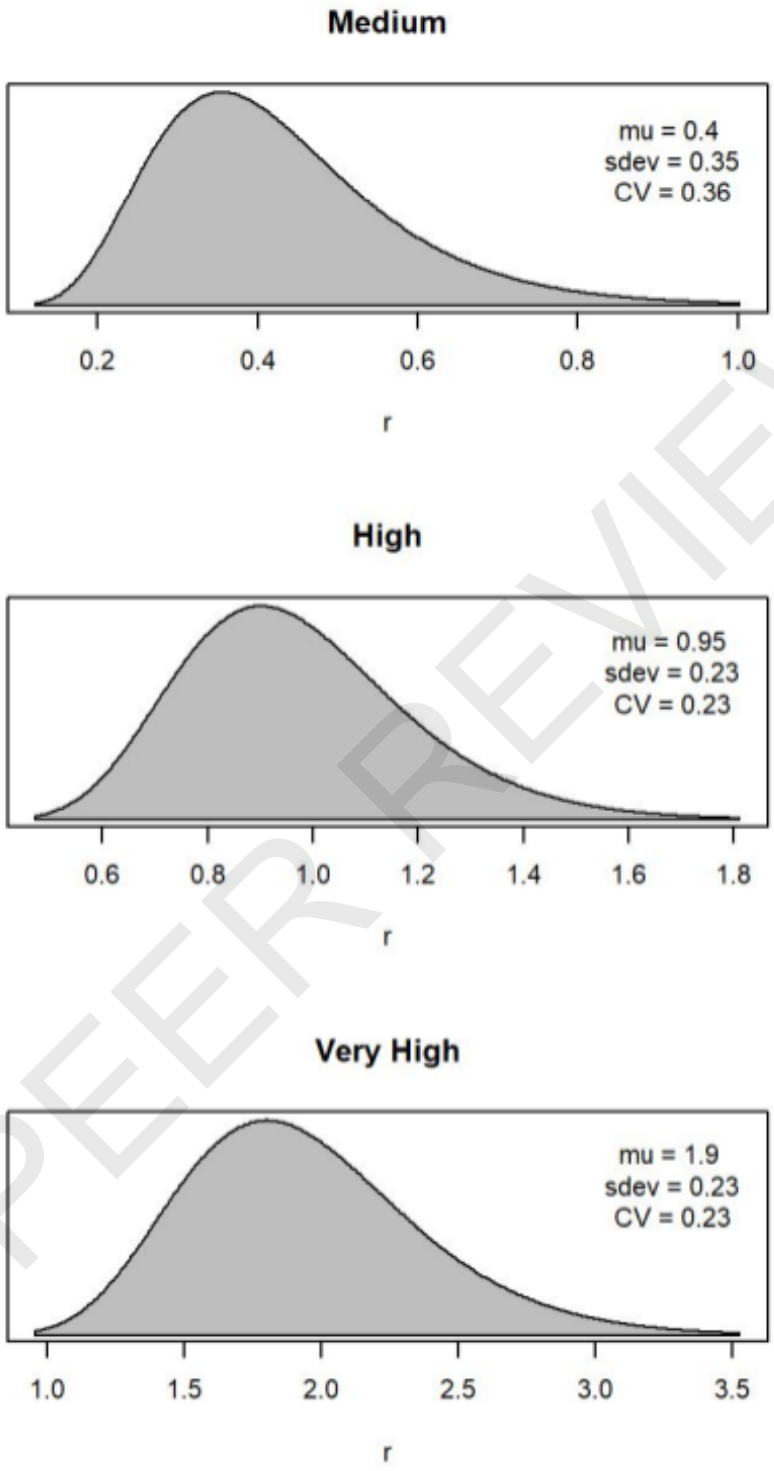

Figure 15: JABBA alternative prior assumptions for the intrinsic growth rate r.

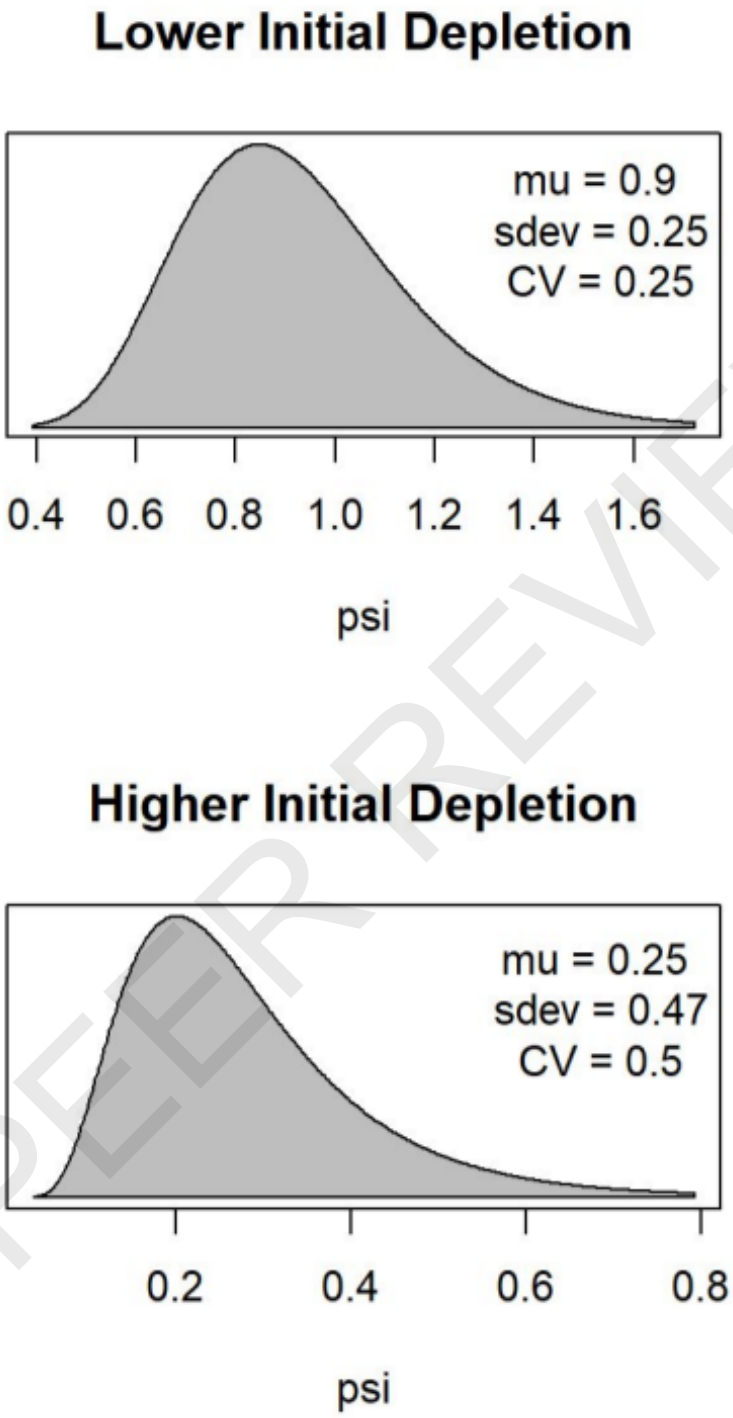

Figure 16: JABBA alternative prior assumptions for the initial biomass depletion ratio psi.

	Model Convergence			Model Fit		Model Consistency				Process Error	Prediction Skill	
run	CONV_gw	CONV_hw	CONV_hs	CPUE_rt_rand	CPUE_rt_outl	RETRO_B	RETRO_F	RETRO_B.Bmsy	RETRO_F.Fmsy	ProcB_CI	HX_MASE	DIC
BSH_1_P_rH_psil0.9_sigT_60	PASS	PASS	PASS	PASS	FAIL	0.02	-0.02	-0.02	0.02	PASS	0.85	-469.80
BSH_4_P_rM_psil0.9_sigT_60	FAIL	PASS	PASS	PASS	FAIL	0.21	-0.15	-0.04	0.03	PASS	1.00	-461.50
BSH_16_P_rM_psil0.9_sigF_60	PASS	PASS	PASS	PASS	FAIL	-0.29	0.45	0.03	0.00	FAIL	0.82	-525.60
BSH_49_P_rV_psil0.9_sigT_60	PASS	PASS	PASS	PASS	FAIL	0.03	-0.03	-0.05	0.05	PASS	0.91	-451.60
BSH_73_P_rH_psil0.2_sigT_60	PASS	PASS	PASS	PASS	FAIL	0.15	-0.12	-0.05	0.04	PASS	0.84	-461.30
BSH_79_P_rH_psil0.2_sigF_60	PASS	PASS	PASS	PASS	FAIL	-0.34	0.52	0.03	0.01	FAIL	0.73	-524.30
BSH_82_P_rM_psil0.2_sigF_60	FAIL	PASS	PASS	PASS	FAIL	-0.43	0.85	0.03	0.02	FAIL	0.82	-522.80
BSH_85_P_rV_psil0.2_sigT_60	PASS	PASS	PASS	PASS	FAIL	0.10	-0.08	-0.01	0.01	PASS	0.85	-466.90

Figure 17: Diagnostic tests for top performing JABBA models, where Run 1 was the best model that passed most of the diagnostic tests.

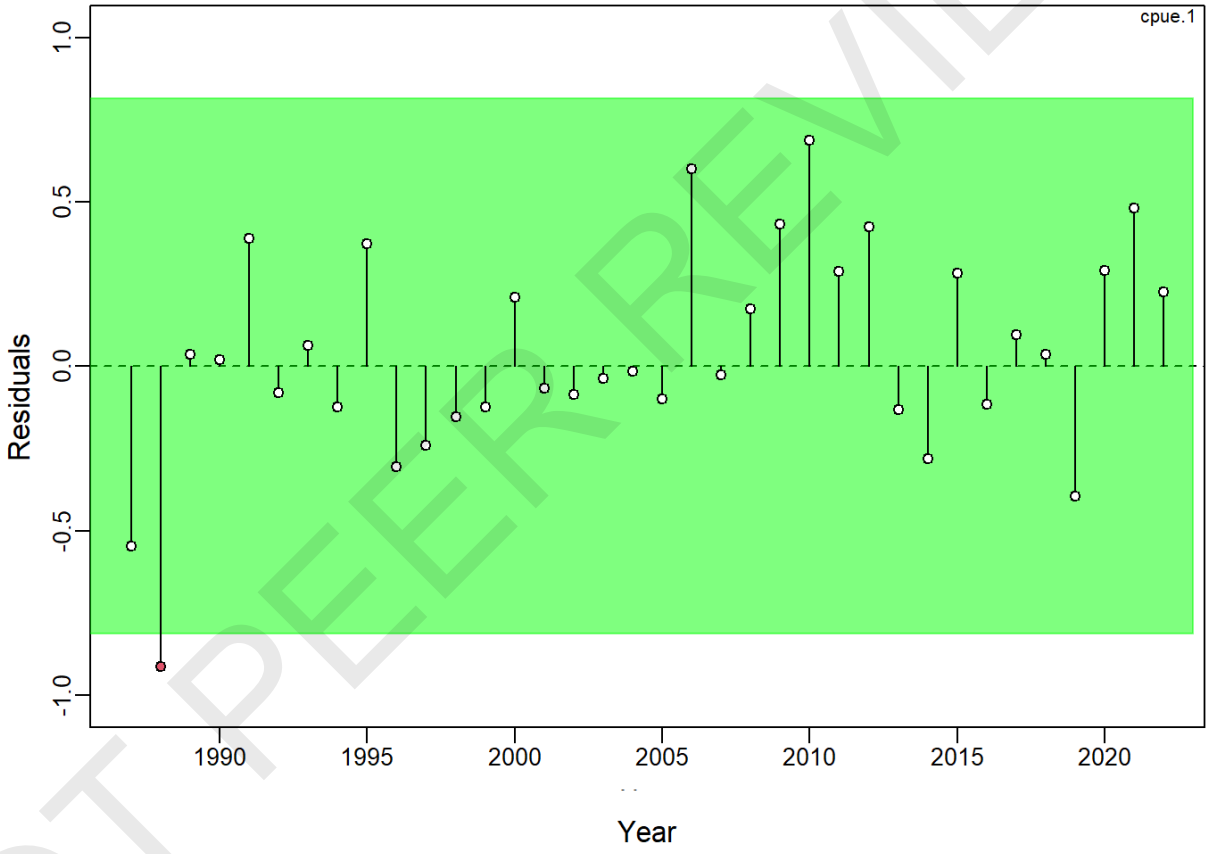

Figure 18: Residual runs test for top performing JABBA model. Green shading indicates no evidence (p = 0.05) and red shading evidence (p < 0.05) to reject the hypothesis of a randomly distributed time-series of residuals, respectively. The shaded (green/red) area spans three residual standard deviations to either side from zero, and the red points outside of the shading violate the 'three-sigma limit' for that series.

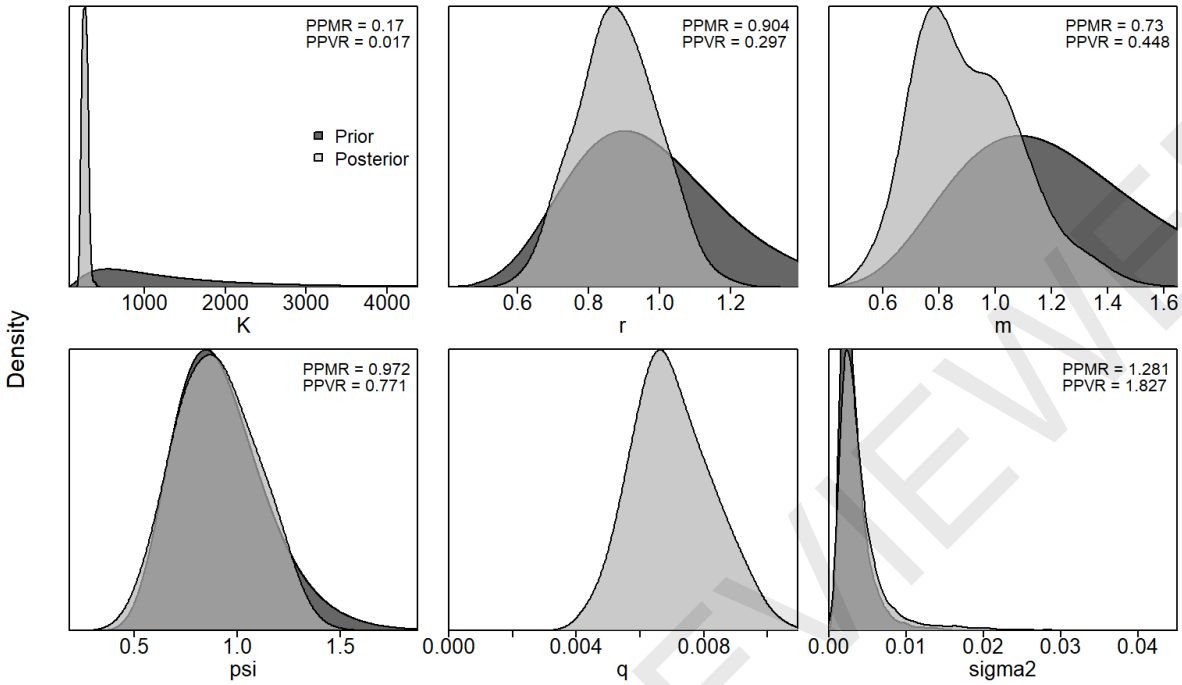

Figure 19: Prior and posterior distributions of key model parameters for top JABBA model run.

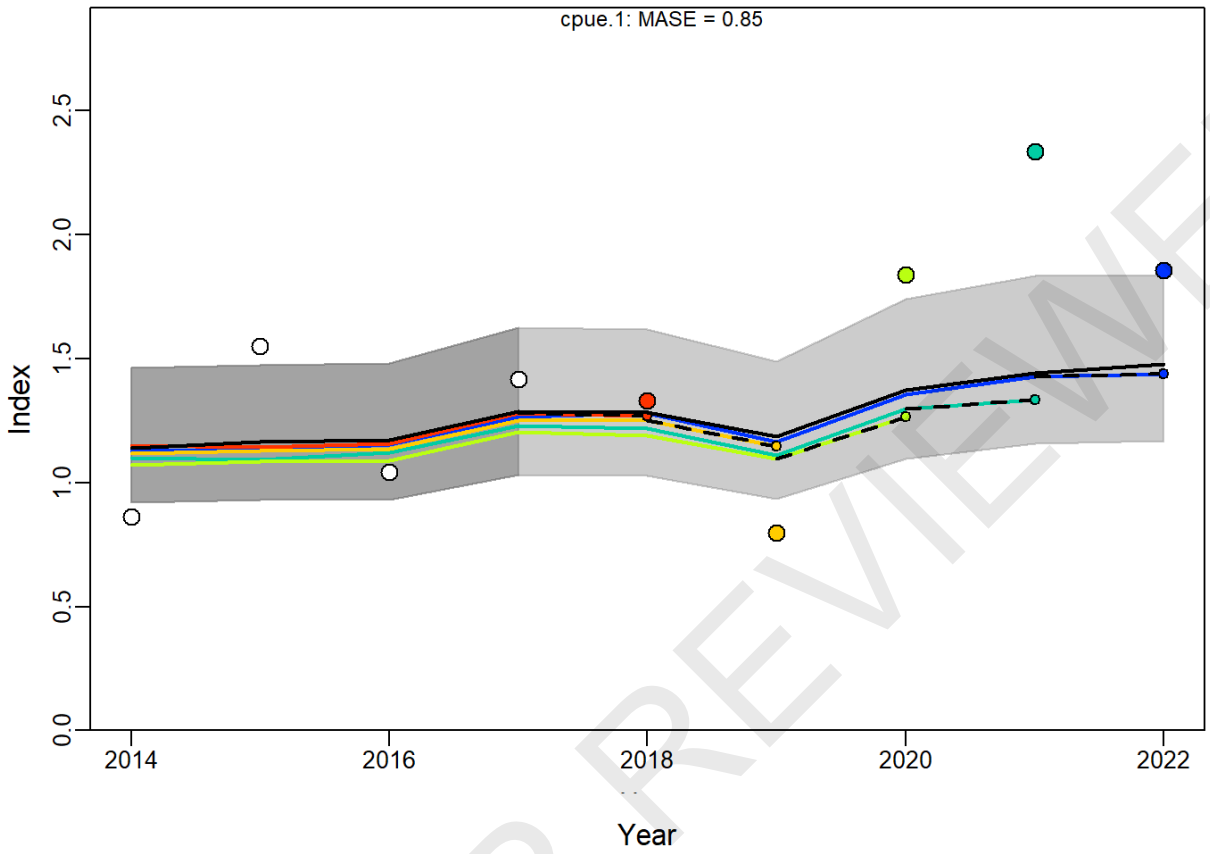

Figure 20: Hindcasting cross-validation (HCxval) results from CPUE fits, showing observed (large points), fitted (solid lines) and one-year ahead forecast values (small terminal points). HCxval was performed using one reference model (black line) and five hindcast model runs (colored lines with terminal years 2018 to 2022) relative to the expected CPUE. The mean absolute scaled error (MASE) score scales the mean absolute error (MAE) of forecasts (i.e., prediction residuals) to MAE of a naïve in-sample prediction (CPUE value this year = CPUE value from last year).

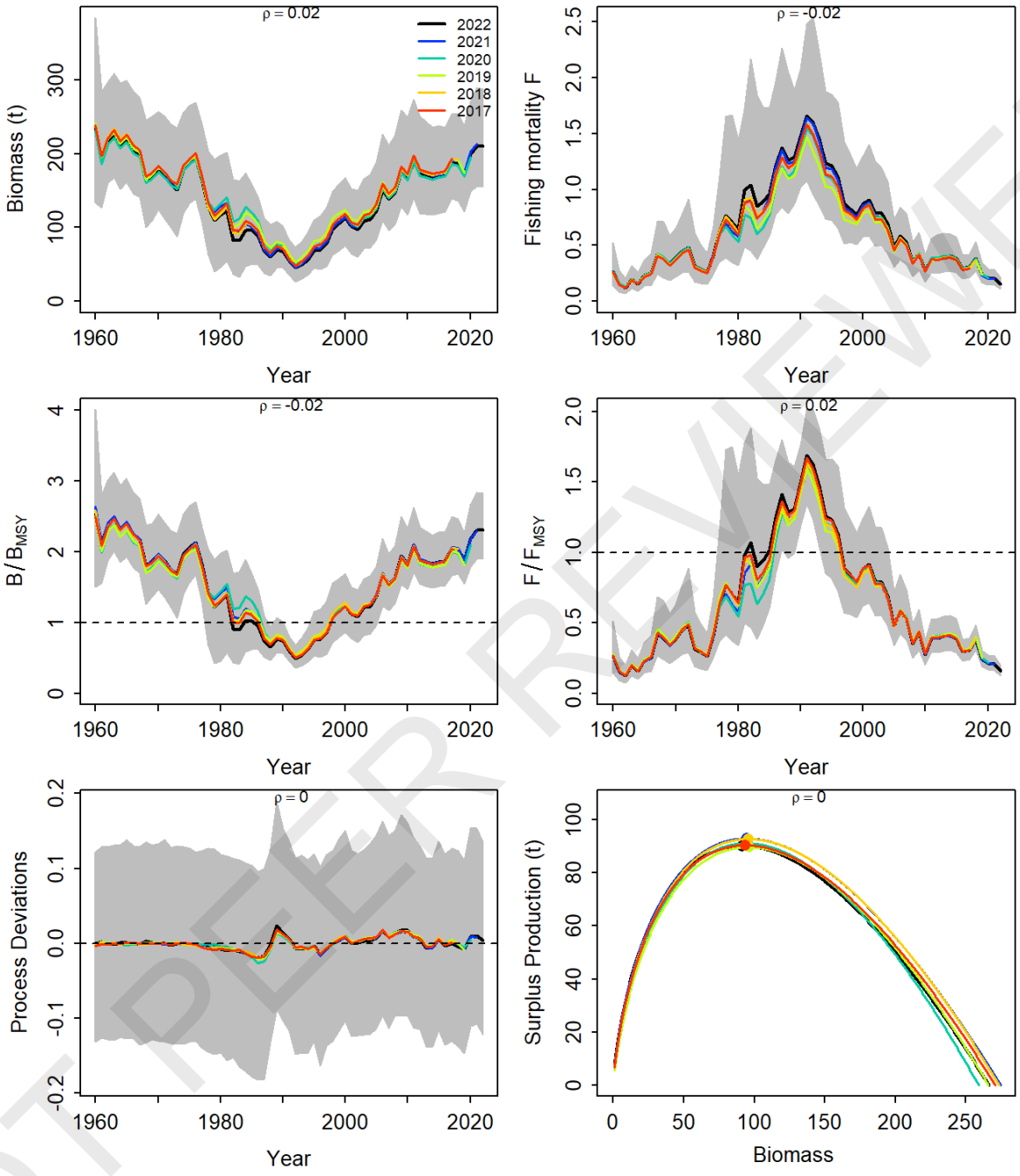

Figure 21: Retrospective analysis of key parameters and management quantities for top performing JABBA model run, with the line color corresponding to terminal years of data ranging from 2017:2022. Mohn's rho statistic (ρ) are denoted on top of the panels. Grey shaded areas are the 95% credible intervals from the reference model. Biomass and surplus production are reported in million lb tail weight.

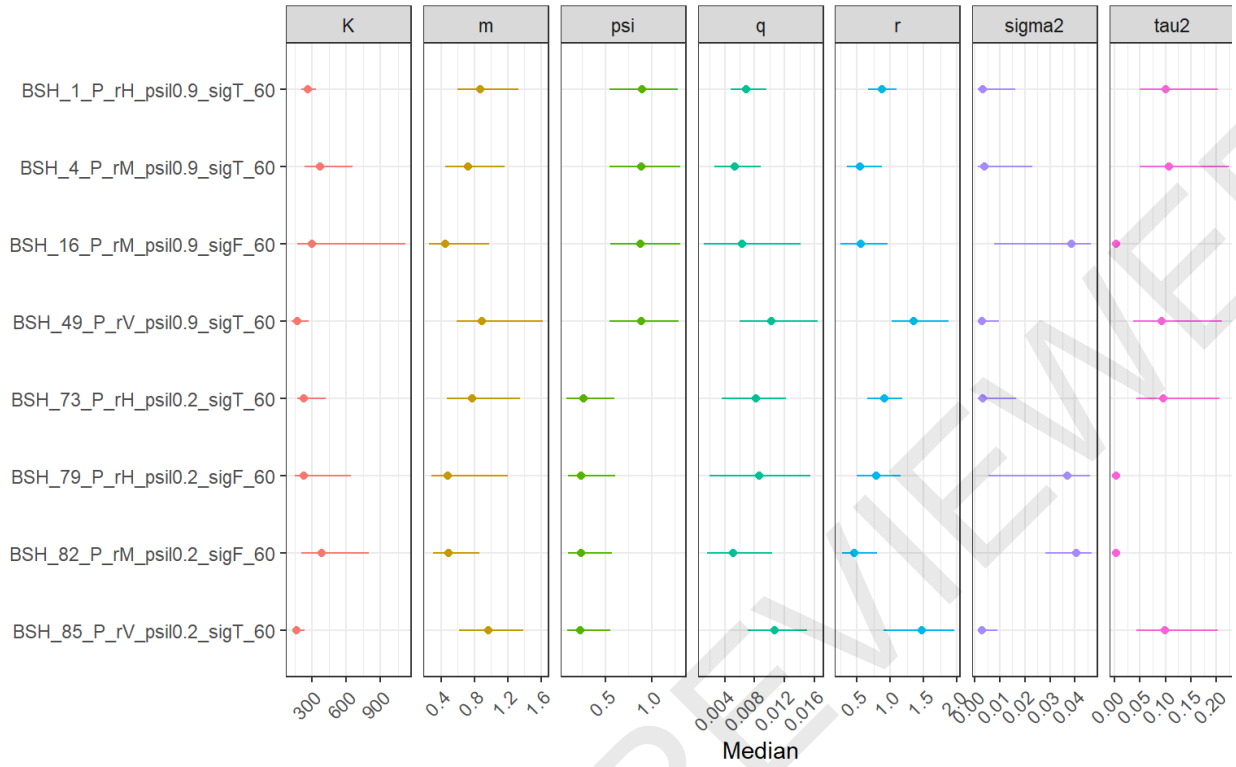

Figure 22: Parameter estimates and error for top performing JABBA models.

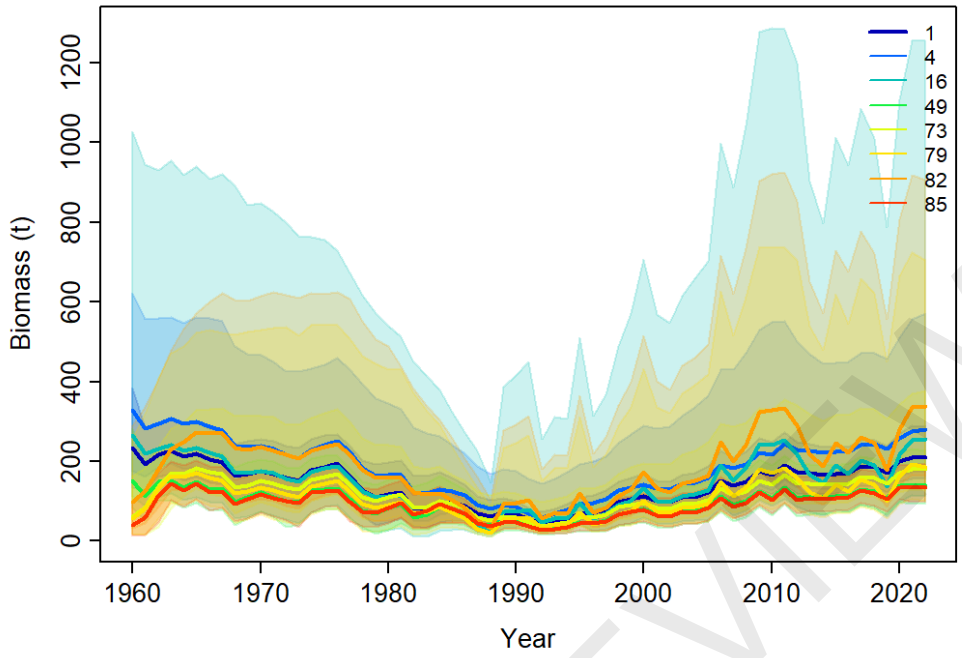

Figure 23: Biomass trajectories for top performing JABBA models (in million lb tail weight), where Run 1 was the best model that passed most of the diagnostic tests.

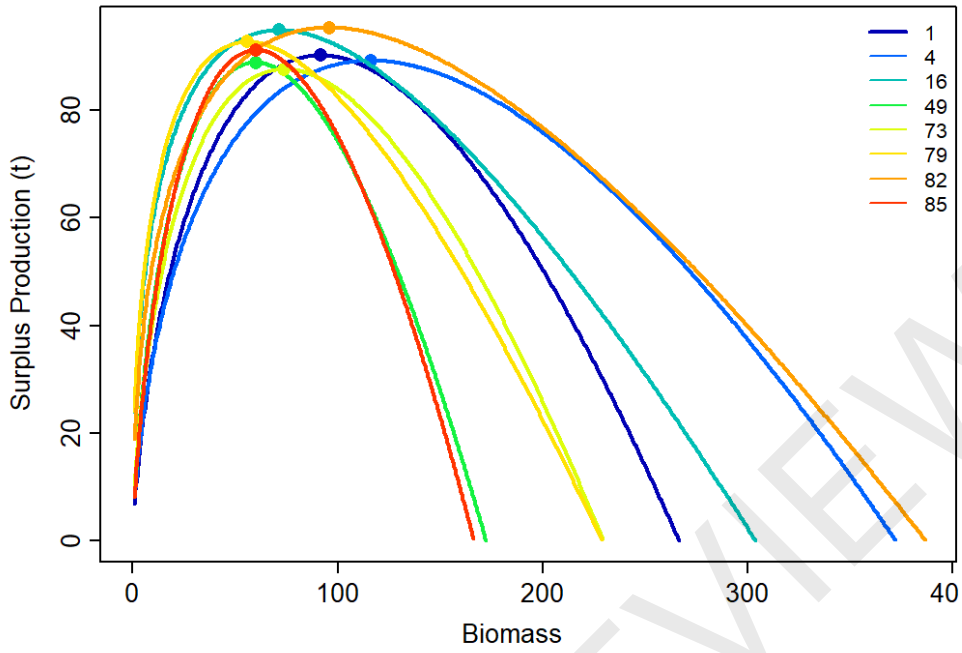

Figure 24: Surplus production and associated biomass estimated for all top performing models (in million lb tail weight). Despite the range of carrying capacities estimated from JABBA, the MSY (peak of the surplus production curve) hovered around 90 million pounds of tails.

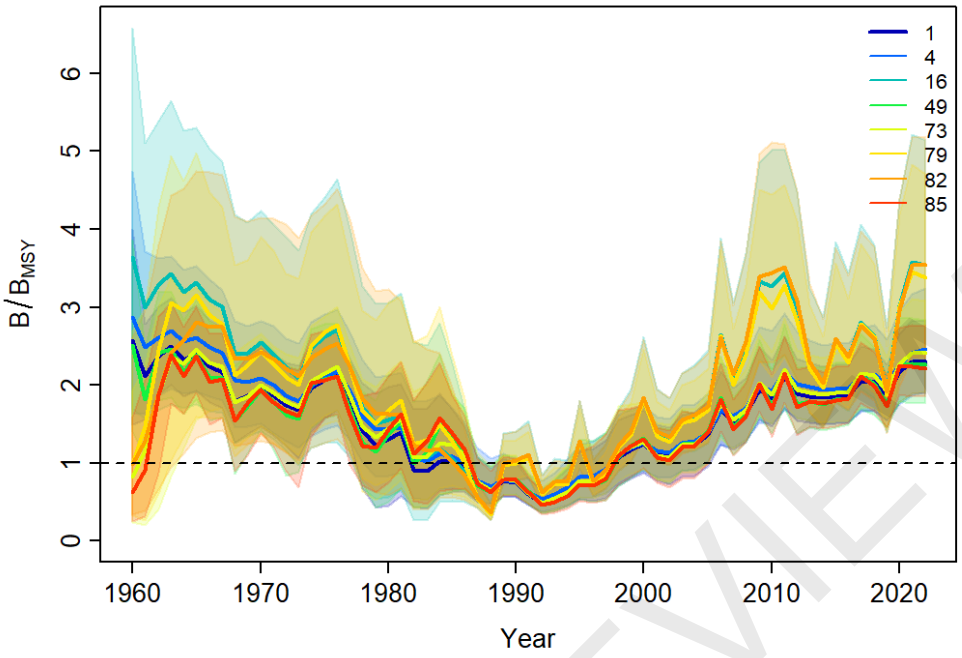

Figure 25: B/Bmsy trajectories for top performing JABBA models, where Run 1 was the best model that passed most of the diagnostic tests.

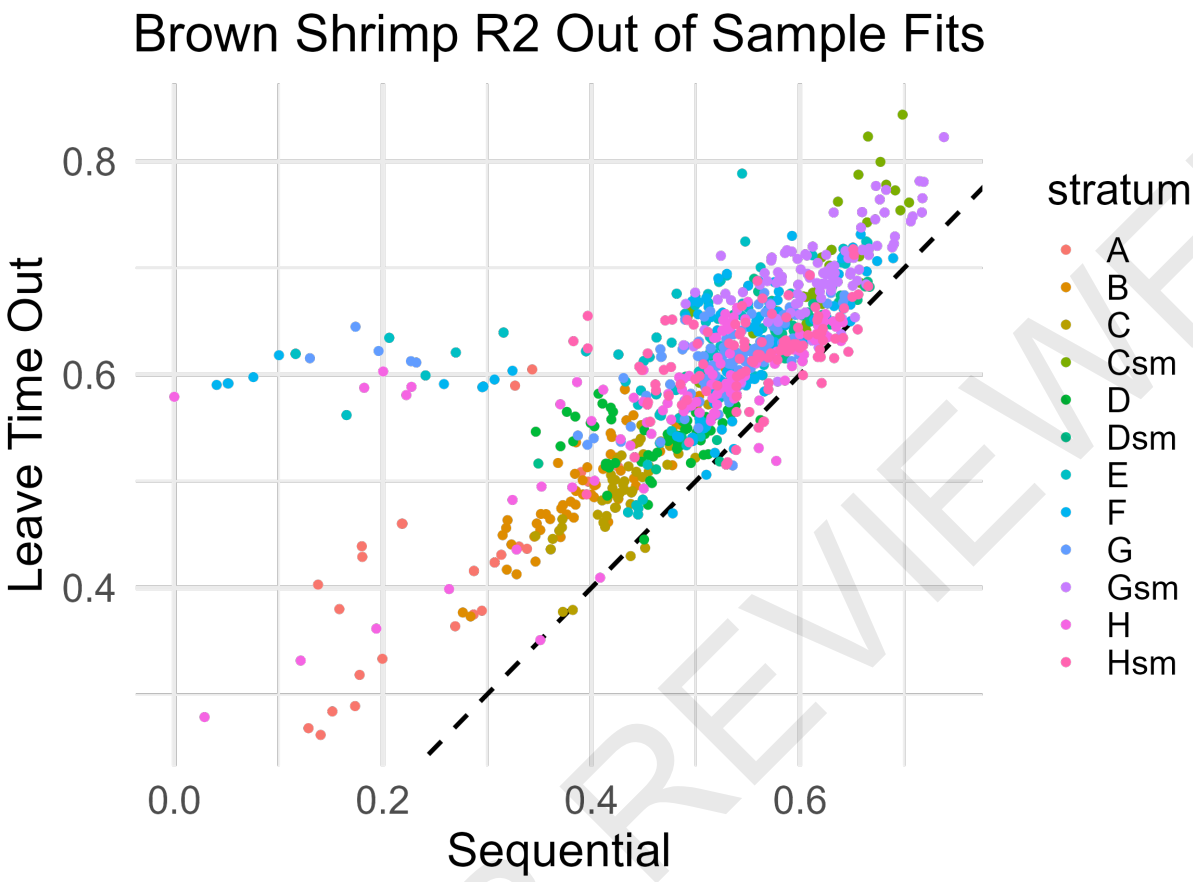

Figure 26: Out-of-sample R2 statistics for each model configuration using the 'leave time out' vs. the 'sequential' cross validation approach. While 'leave-time-out' obtains better model fits, the purpose here is to be able to project well into the future, which is better captured by the 'sequential' approach. Models are filtered based on the R2 statistics from the 'sequential' prediction method going forward.

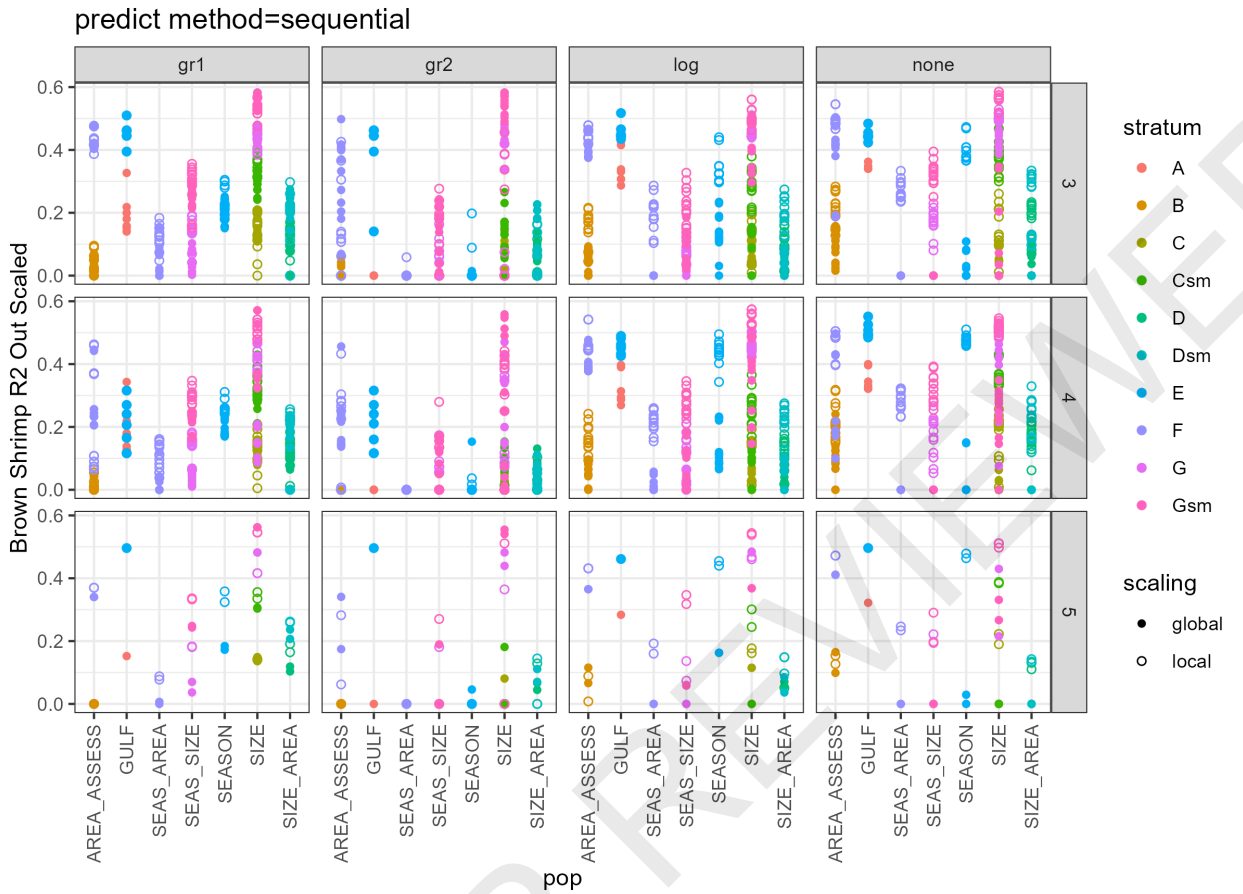

Figure 27: "Sequential" out-of-sample R2 fit statistic resulting from each model run. Facet columns show results based on different data transformations. Facet rows show results based on the embedding dimension. Within each facet, the x axis groups the models by the type of aggregation (spatial, size, season, or a combination). Models that locally scaled the data performed better than models that scaled globally for most transformations, which aligns with how we understand the data (e.g. populations are not directly comparable as defined here).

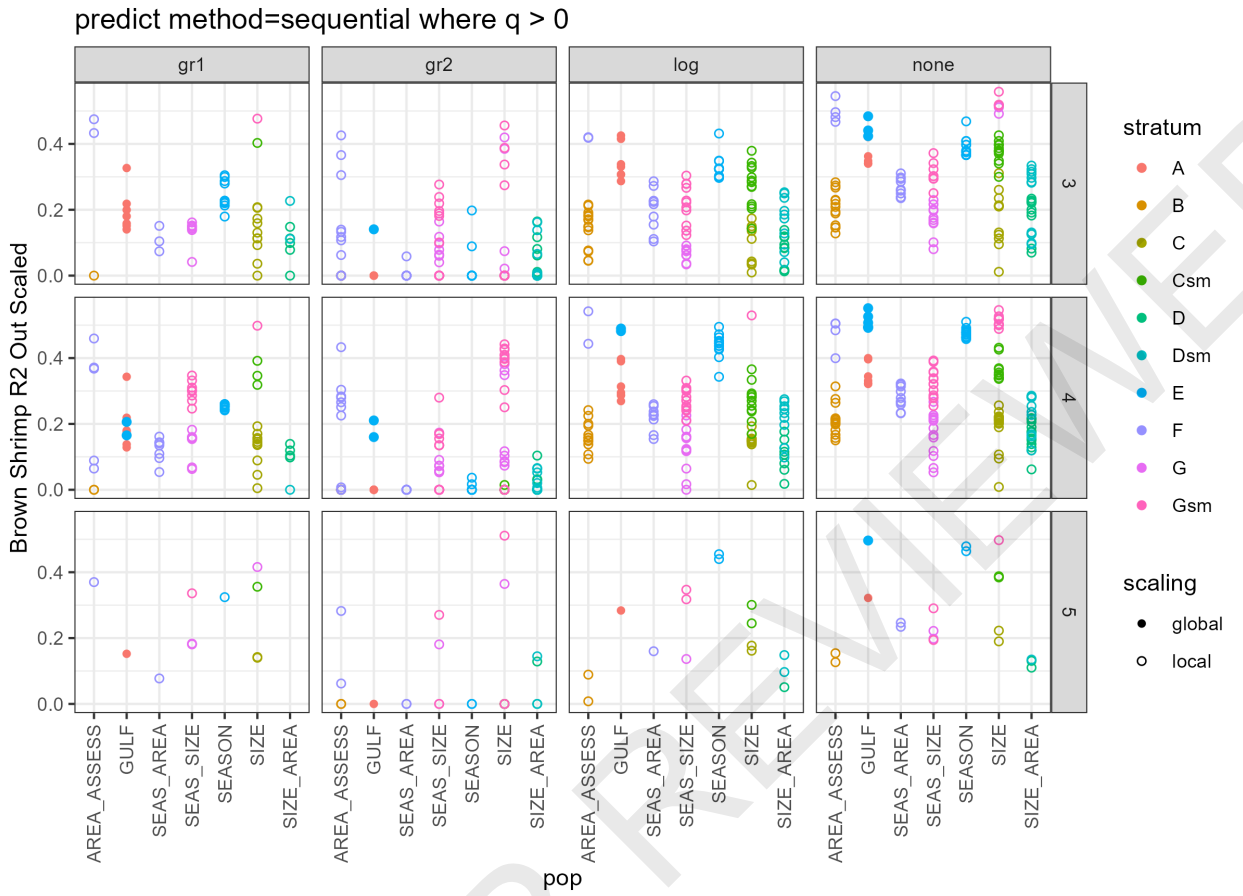

Figure 28: "Sequential" out-of-sample R2 fit statistic resulting from each model run with "local" scaling. Facet columns show results based on different data transformations. Facet rows show results based on the embedding dimension. Within each facet, the x axis groups the models by the type of aggregation (spatial, size, season, or a combination). Models that fit to the survey data and ignored landings (e.g. q=0) were removed from further consideration.

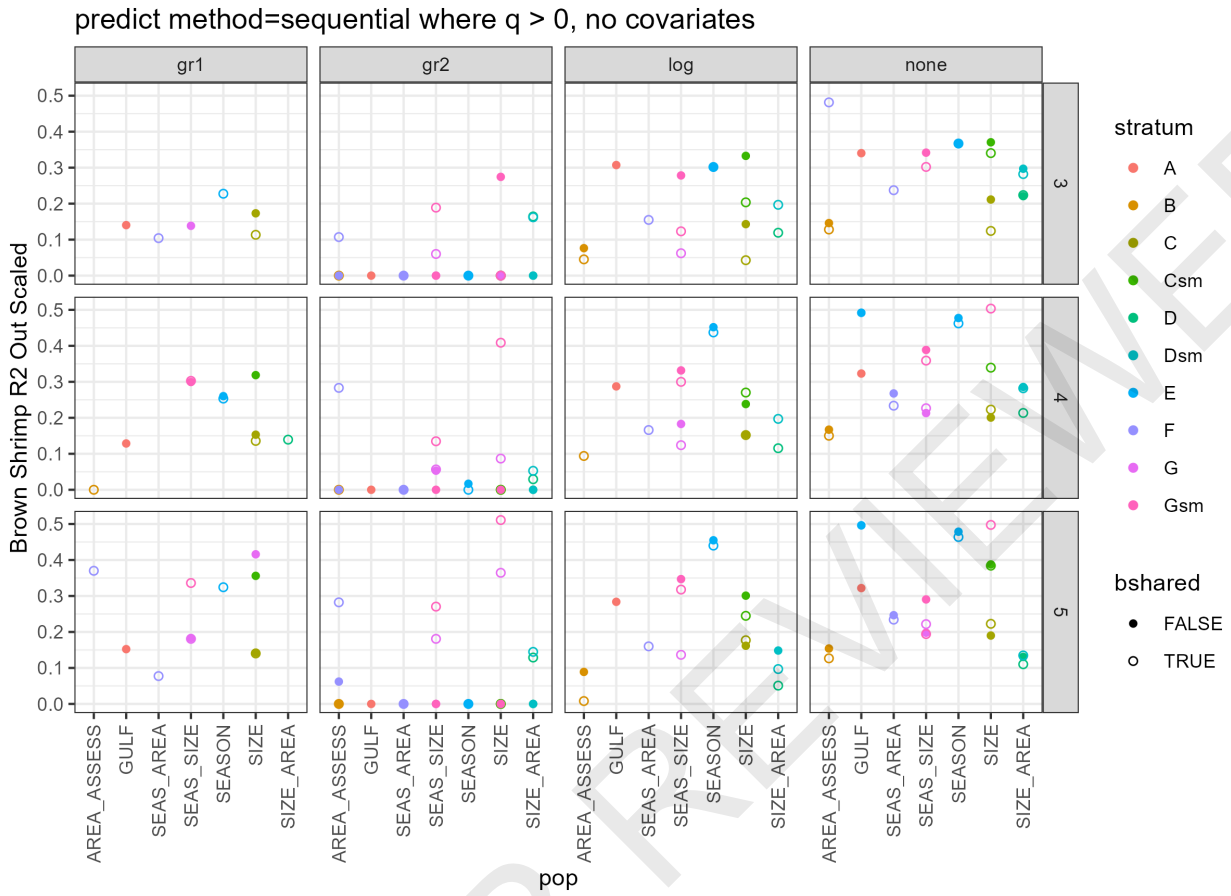

Figure 29: "Sequential" out-of-sample R2 fit statistic resulting from each model run with "local" scaling, q>0.001, and no covariates. Facet columns show results based on different data transformations. Facet rows show results based on the embedding dimension. Within each facet, the x axis groups the models by the type of aggregation (spatial, size, season, or a combination). In this figure, the shape fill was determined by whether or not the catchability parameter was shared among populations in the model (bshared = True / False, respectively).

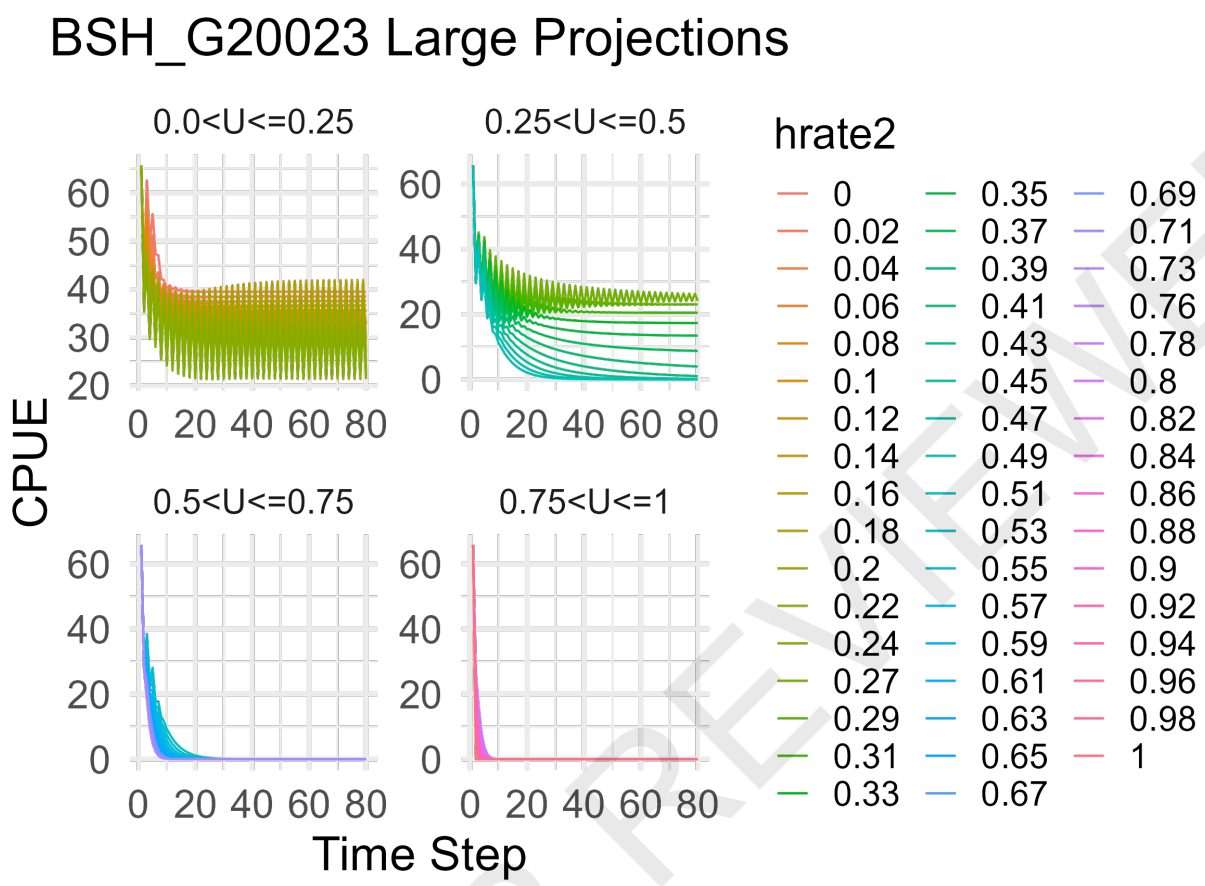

Figure 30: Variable harvest rate projections of CPUE from the best performing run for the Large shrimp population.

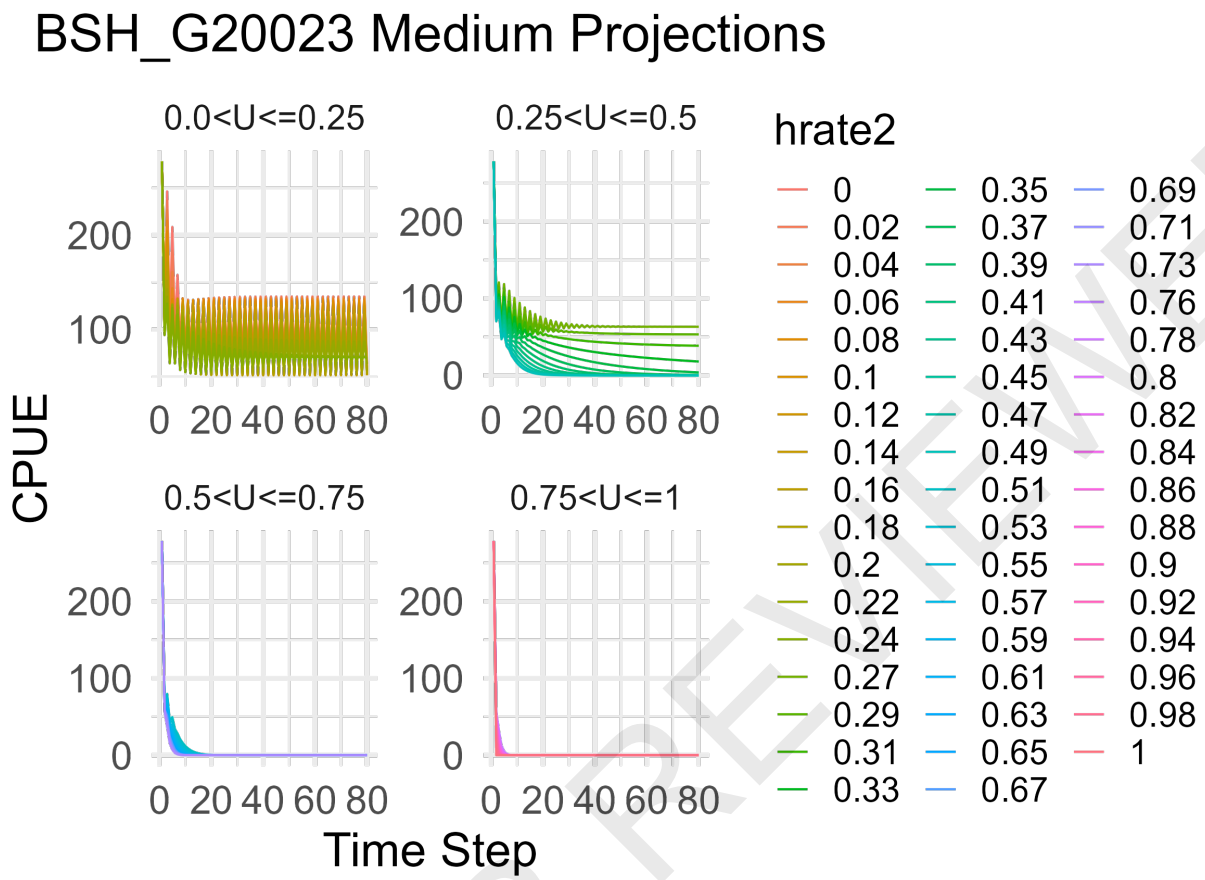

Figure 31: Variable harvest rate projections of CPUE from the best performing run for the Medium shrimp population.

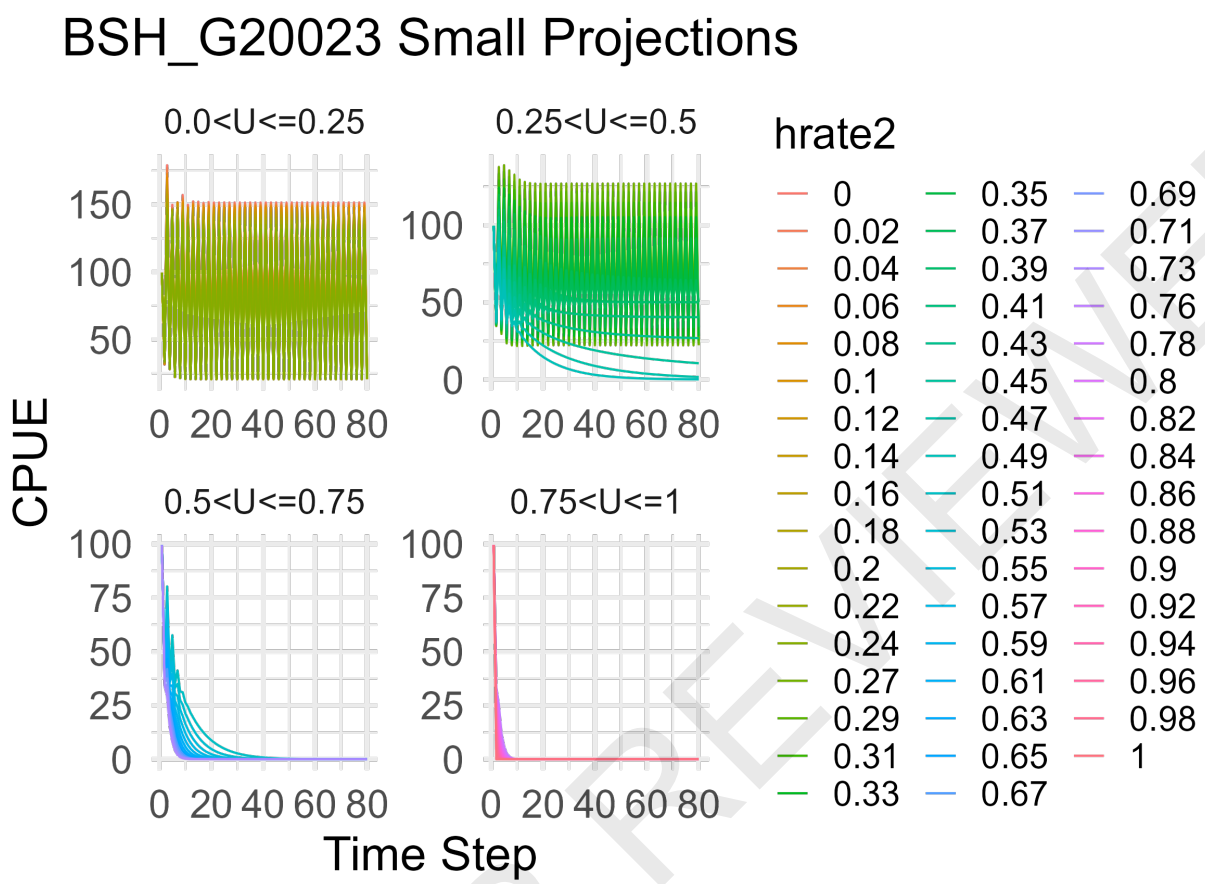

Figure 32: Variable harvest rate projections of CPUE from the best performing run for the Small shrimp population.

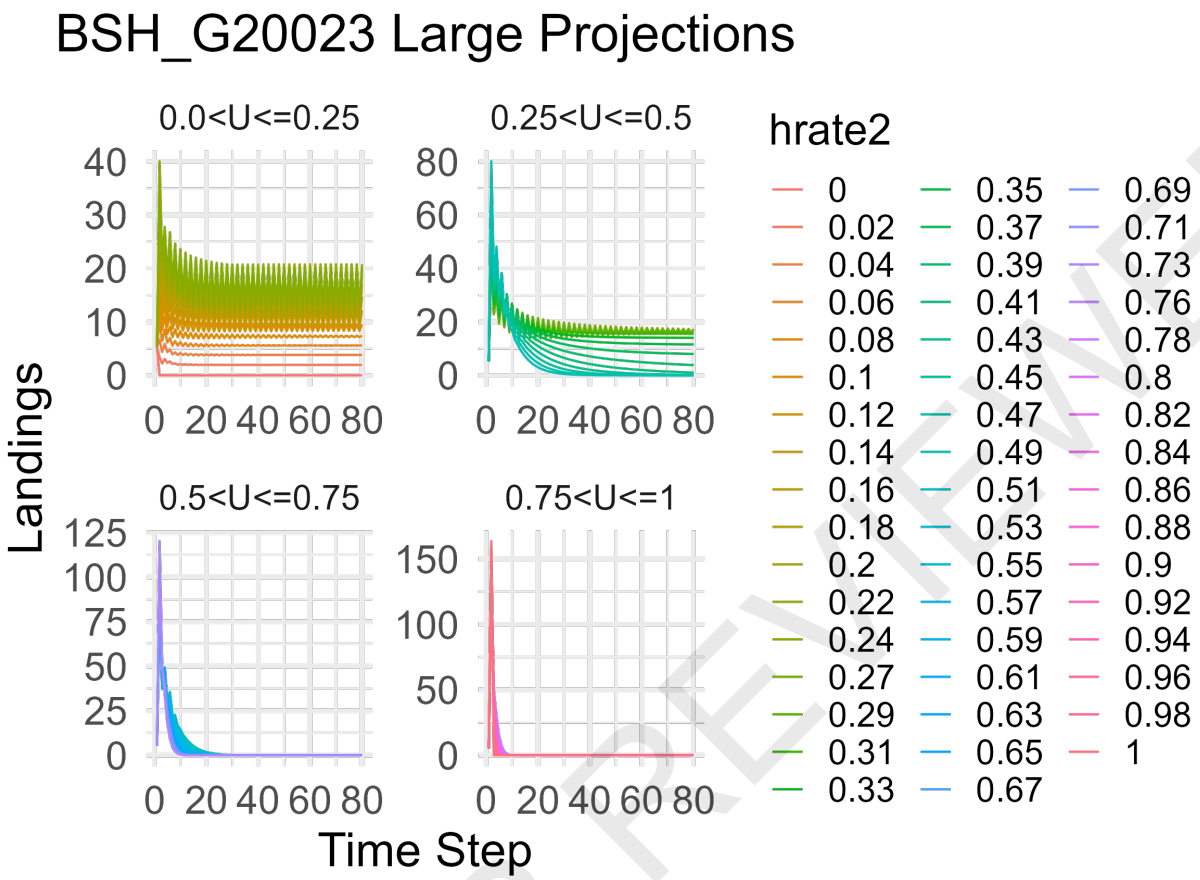

Figure 33: Variable harvest rate projections of landings from the best performing run for the Large shrimp population.

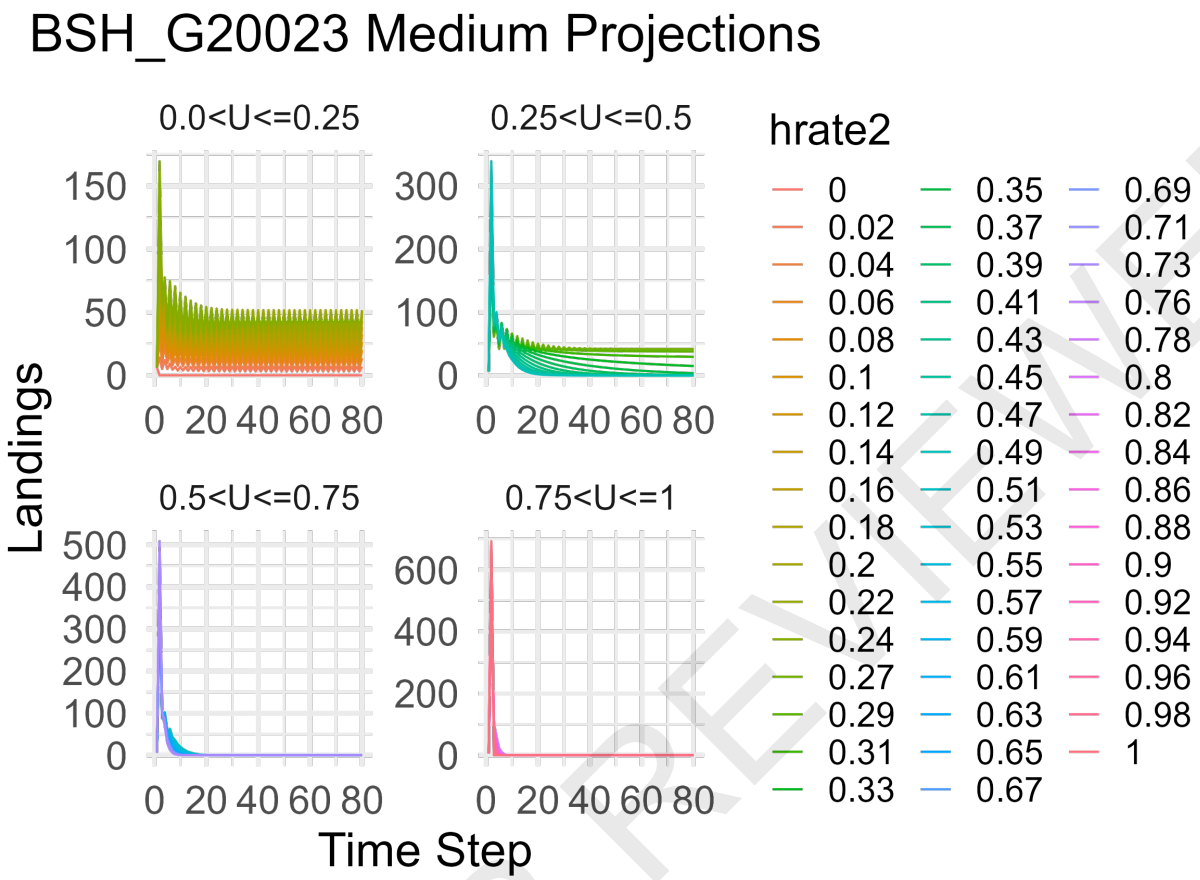

Figure 34: Variable harvest rate projections of landings from the best performing run for the Medium shrimp population.

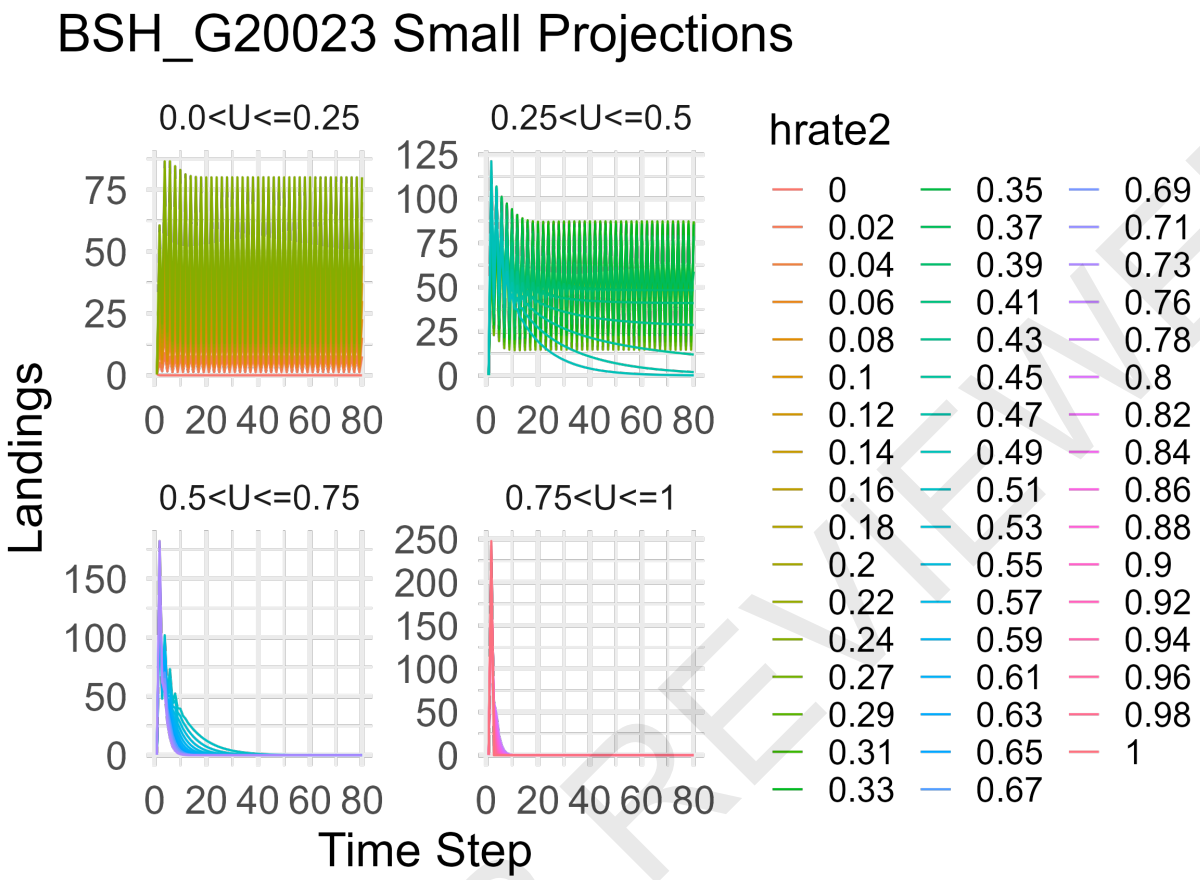

Figure 35: Variable harvest rate projections of landings from the best performing run for the Small shrimp population.

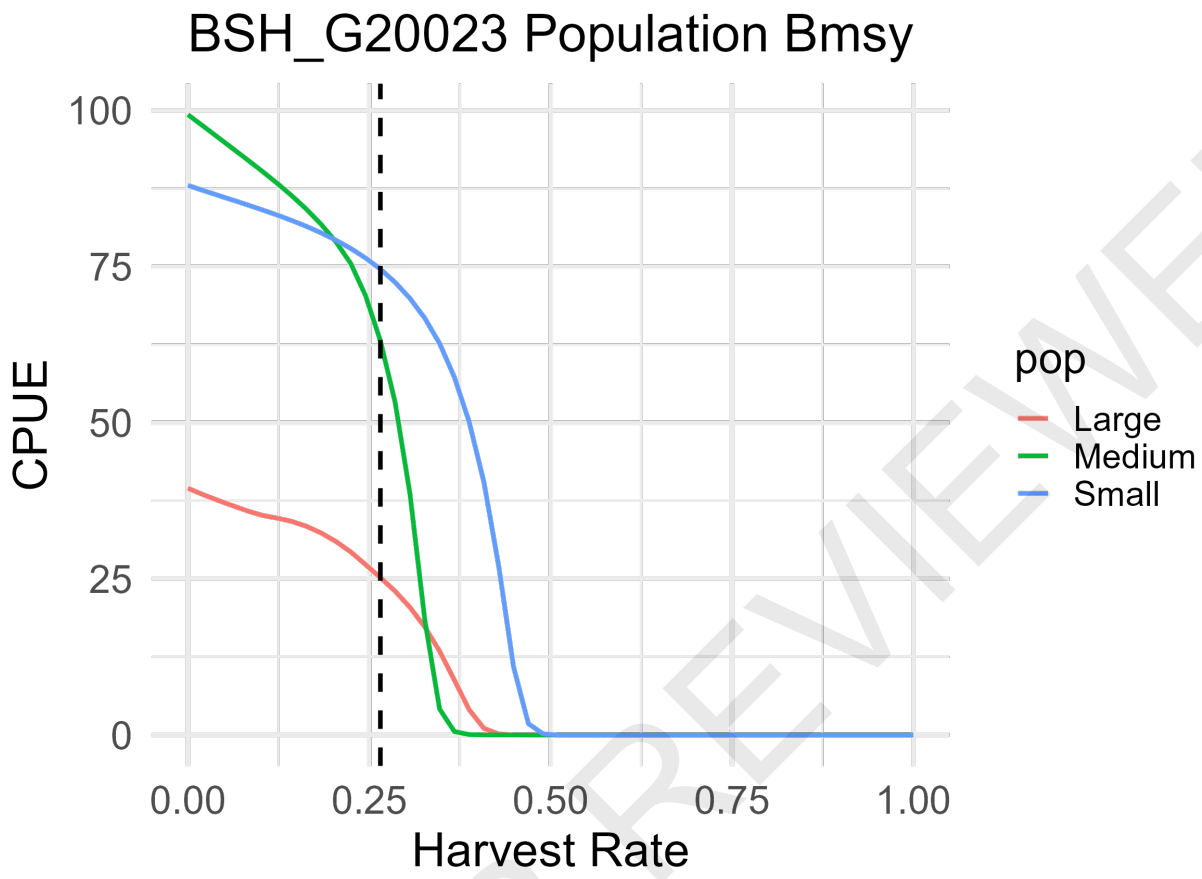

Figure 36: Average CPUE by harvest rate for individual populations for the best performing run. The dashed line indicates the seasonal harvest rate where MSY occurs, indicating population-wide Bmsy in units of CPUE for each population.

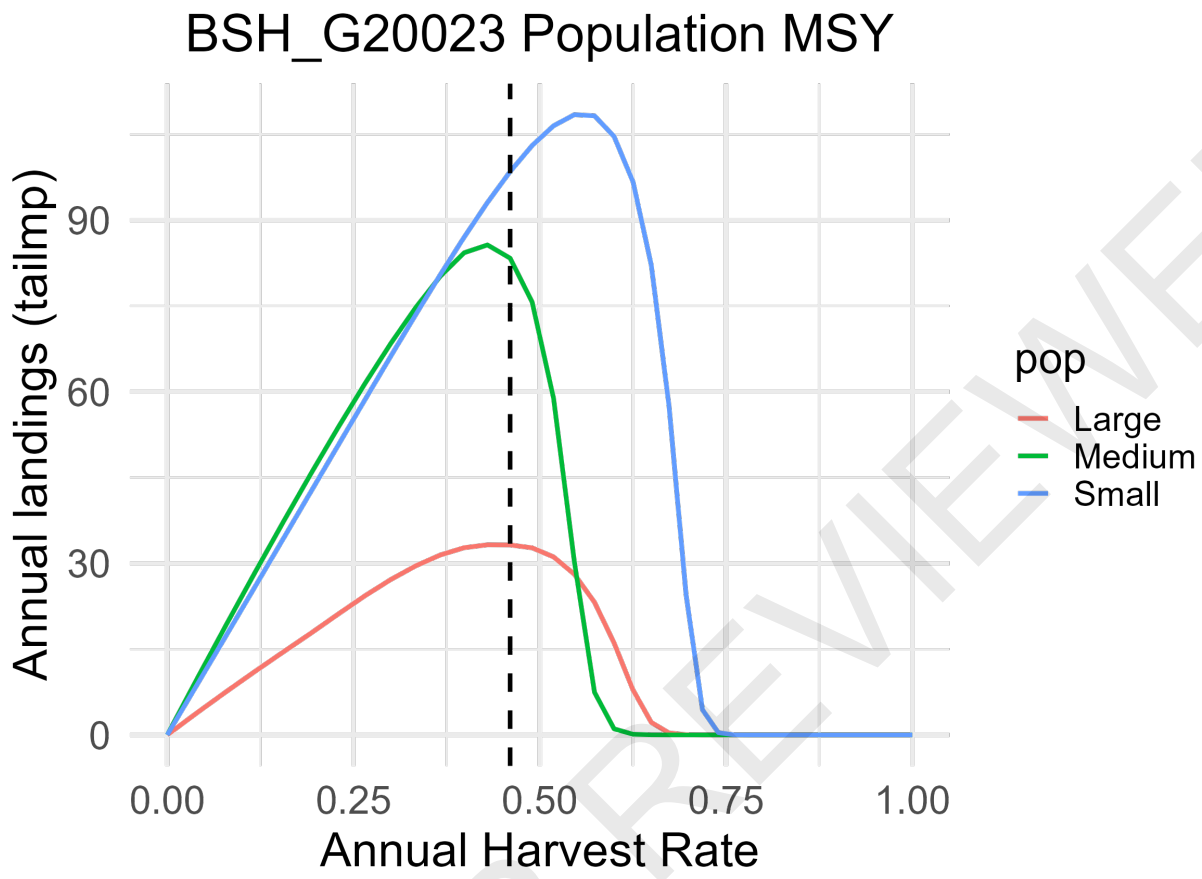

Figure 37: Average landings estimated under a range of annual harvest rates for the best performing run. The optimal annual harvest rate for all populations combined is shown in the dashed line. Individual populations see their landings maximized at slightly different harvest rates.

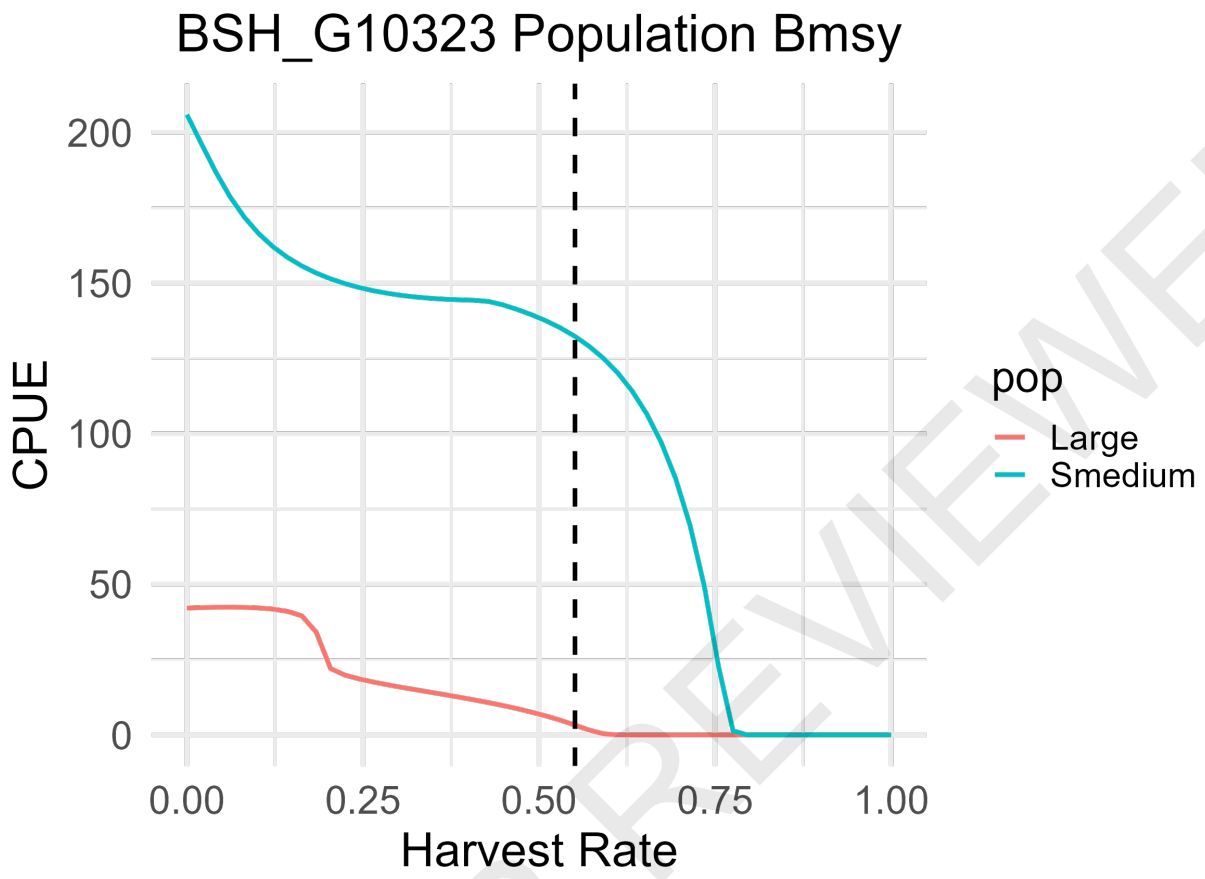

Figure 38: Average CPUE by harvest rate for individual populations for a top performing run. The dashed line indicates the seasonal harvest rate where MSY occurs, indicating population-wide Bmsy in units of CPUE for each population. For this model, the Large shrimp population is approaching zero under the harvest rate that maximizes total landings.

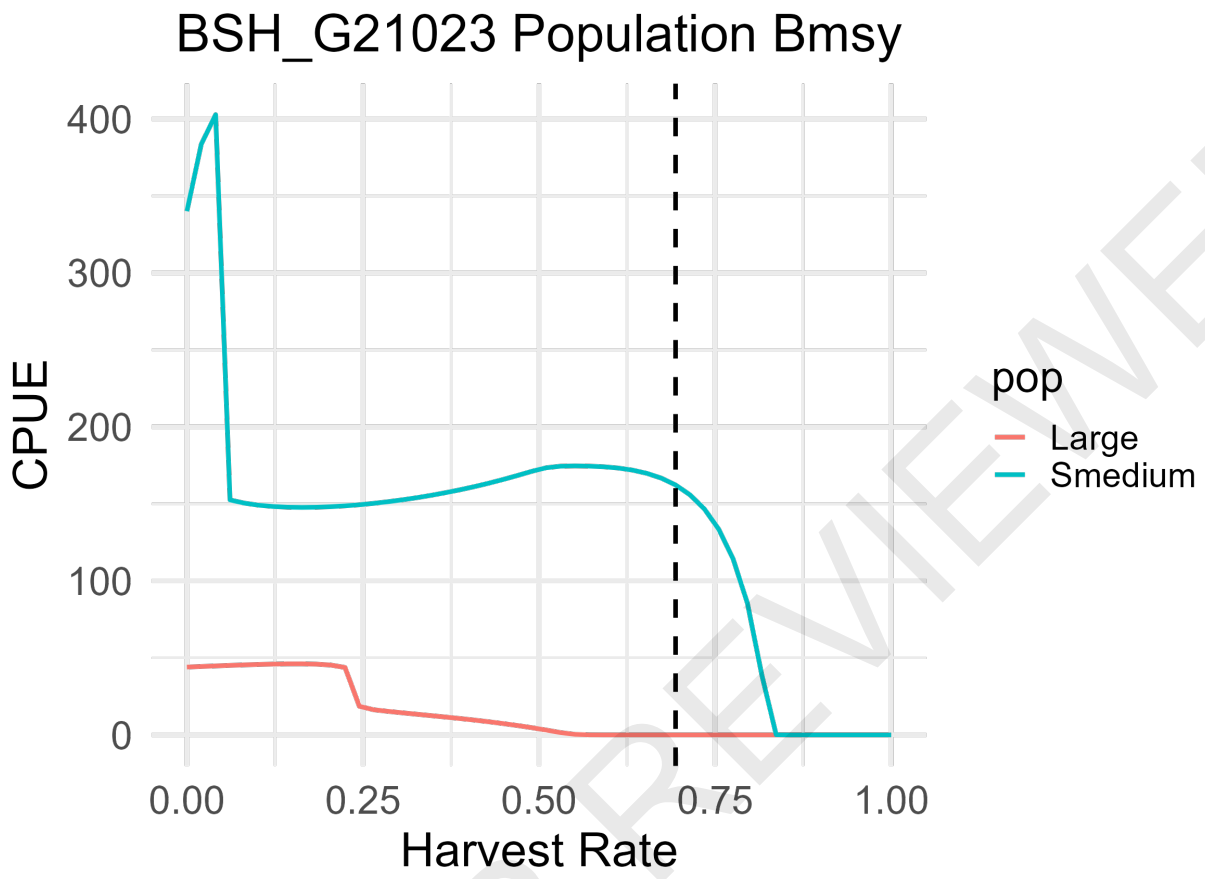

Figure 39: Average CPUE by harvest rate for individual populations for a top performing run. The dashed line indicates the seasonal harvest rate where MSY occurs, indicating population-wide Bmsy in units of CPUE for each population. For this model, the Large shrimp population is nearly zero under the harvest rate that maximizes total landings.

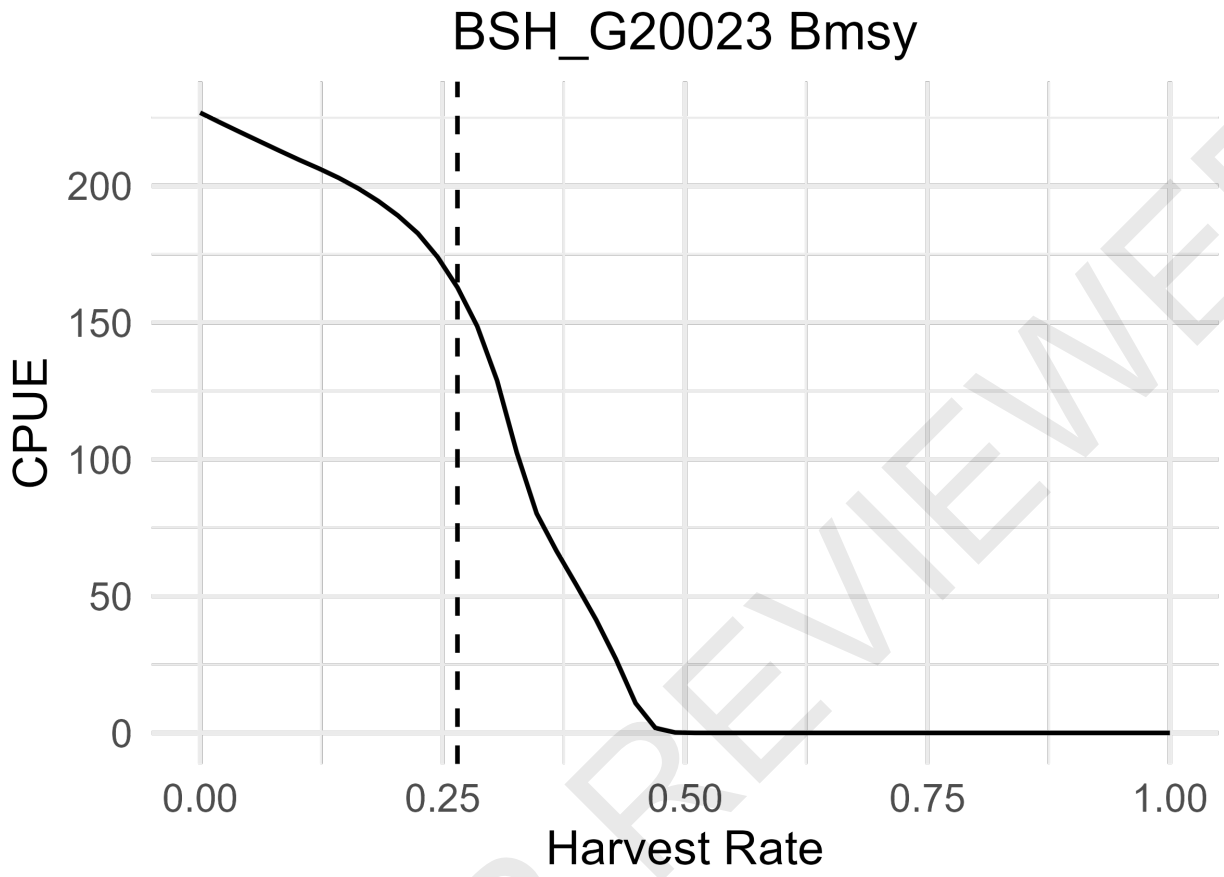

Figure 40: Average CPUE for all populations combined for the best performing run. The dashed line indicates the seasonal harvest rate where MSY occurs, indicating population-wide Bmsy in units of CPUE.

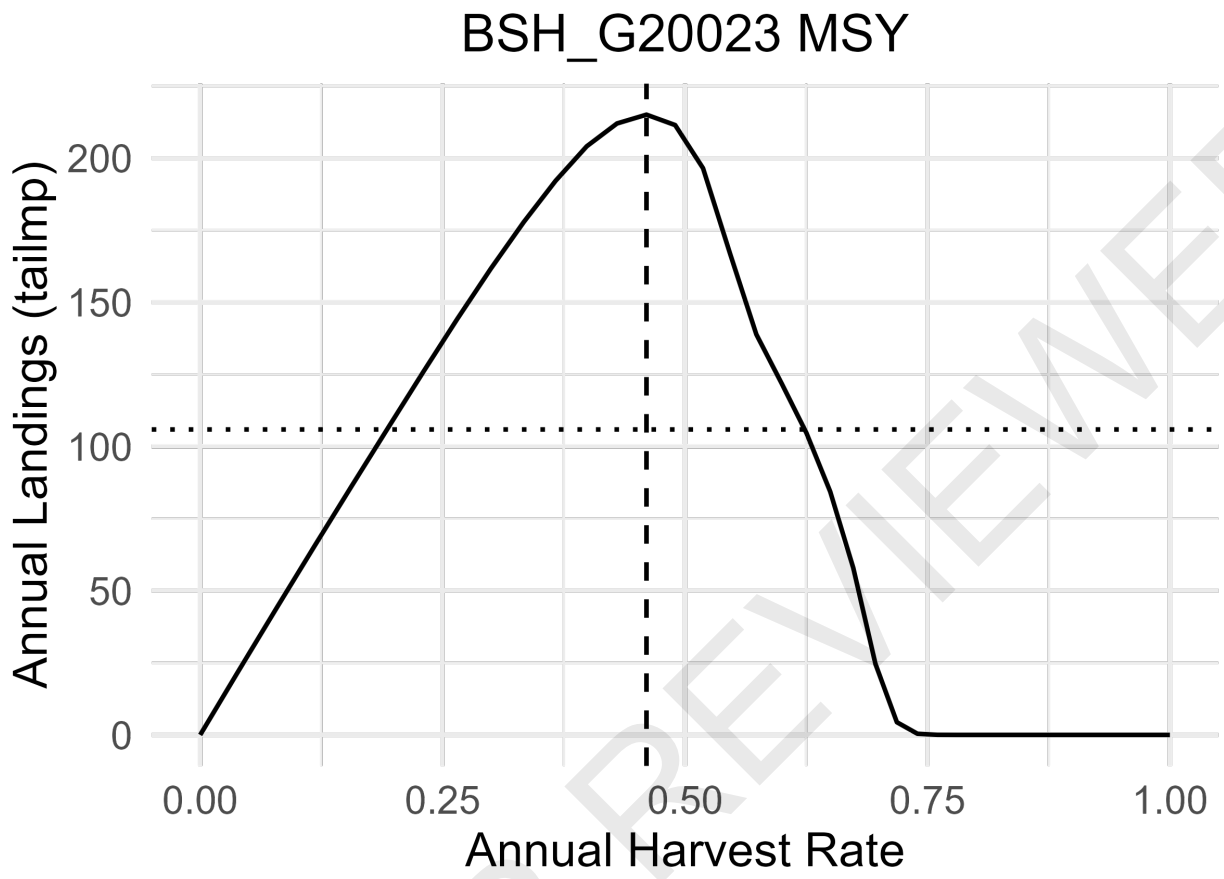

Figure 41: Average landings estimated under a range of annual harvest rates for the best performing run. The optimal annual harvest rate for all populations combined (MSY) is marked with a vertical dashed line. The maximum historical landings are marked with a horizontal dotted line, which is less than half the estimated MSY.

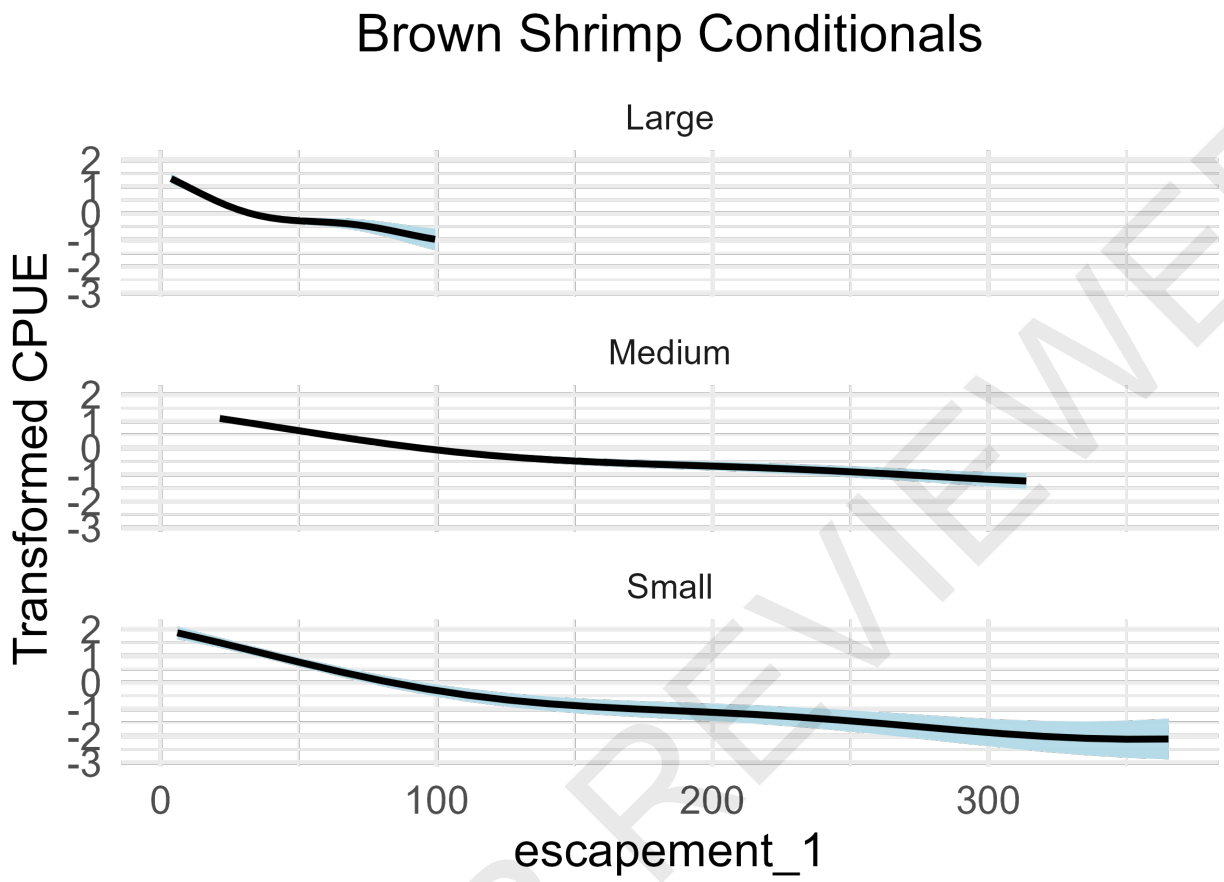

Figure 42: Length scale parameters for the 1st lag of abundance from the best performing model.

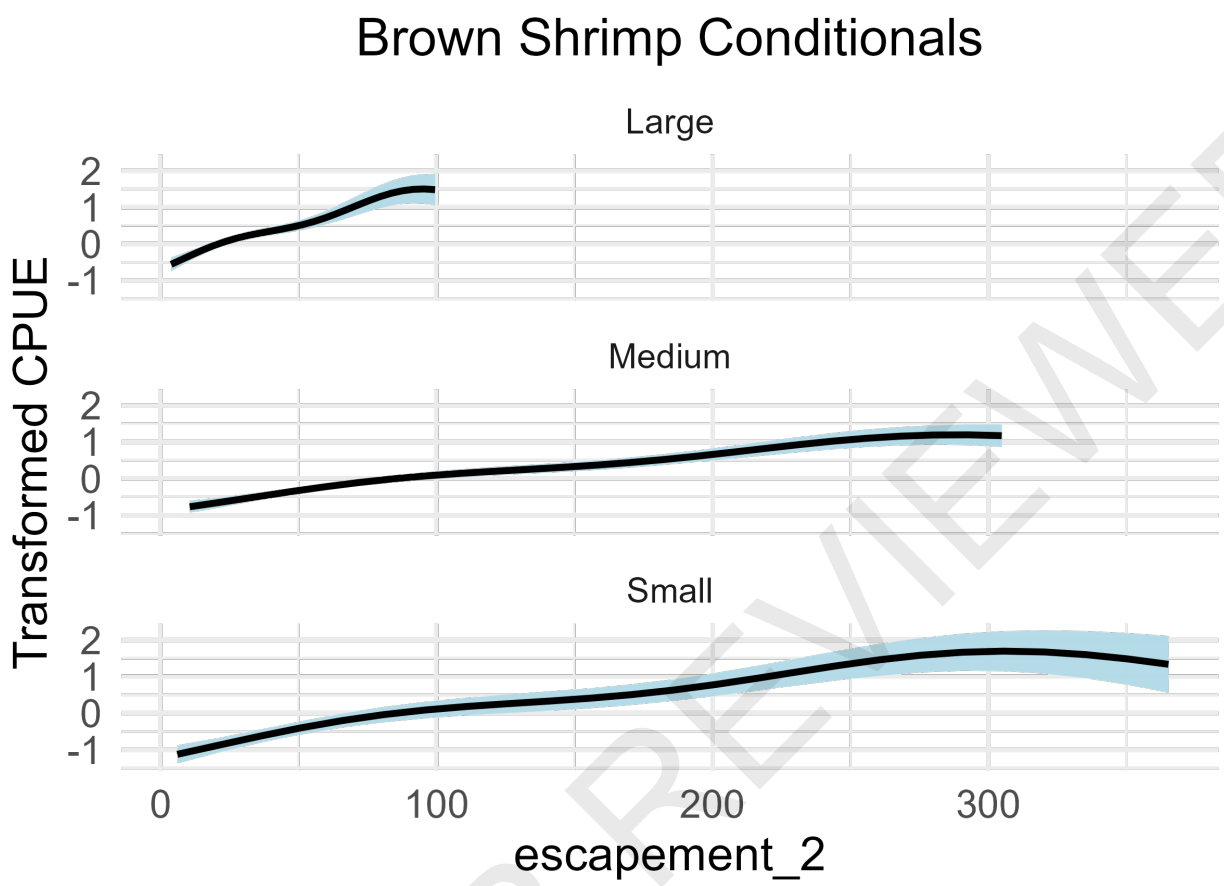

Figure 43: Length scale parameters for the 2nd lag of abundance from the best performing model.

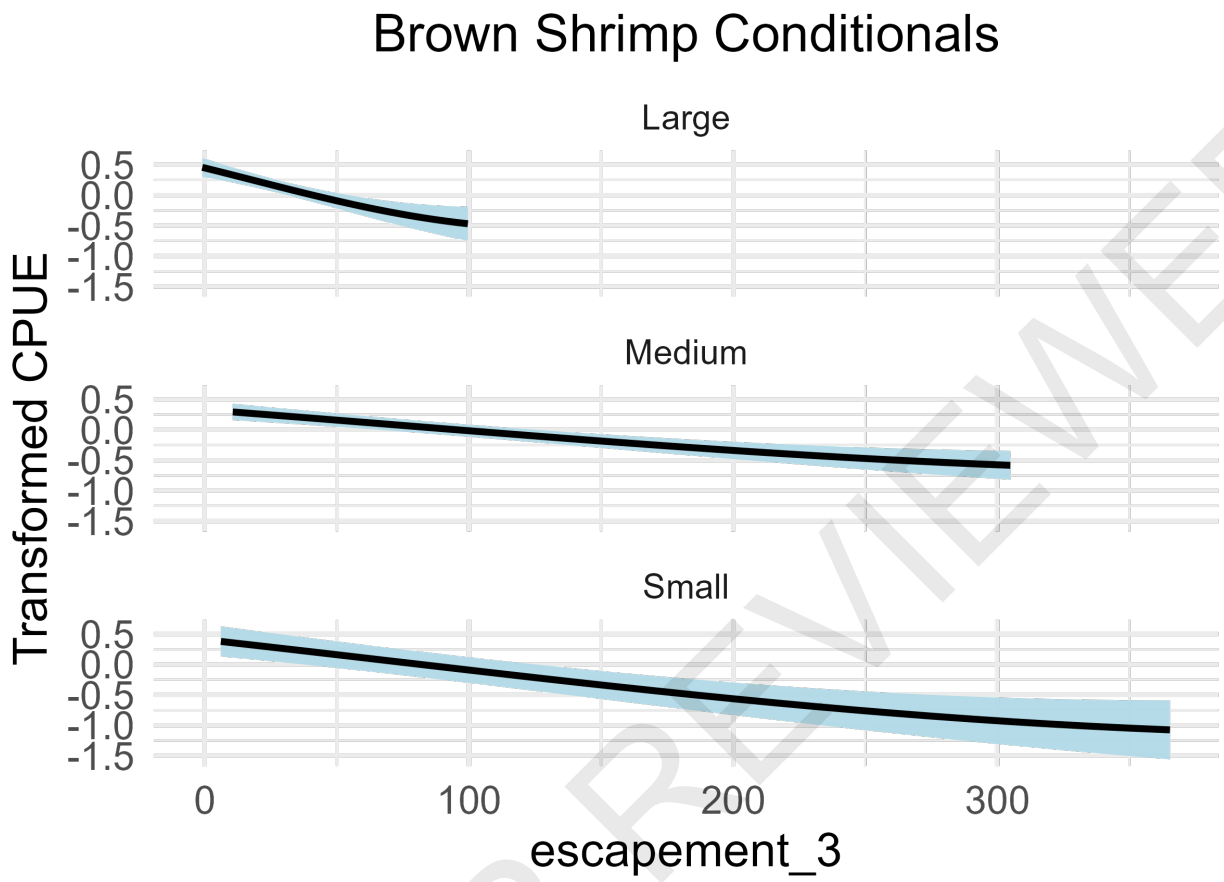

Figure 44: Length scale parameters for the 3rd lag of abundance from the best performing model.

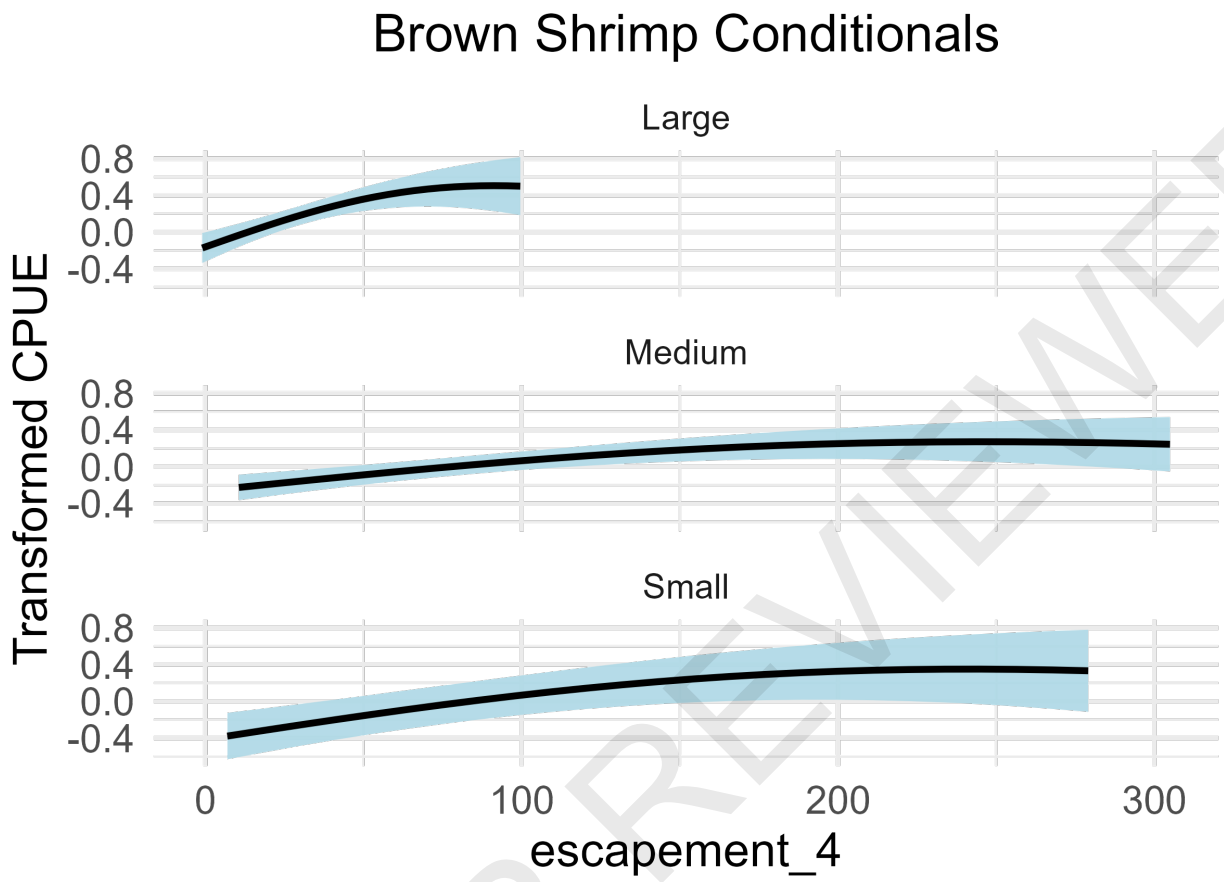

Figure 45: Length scale parameters for the 4th lag of abundance from the best performing model.

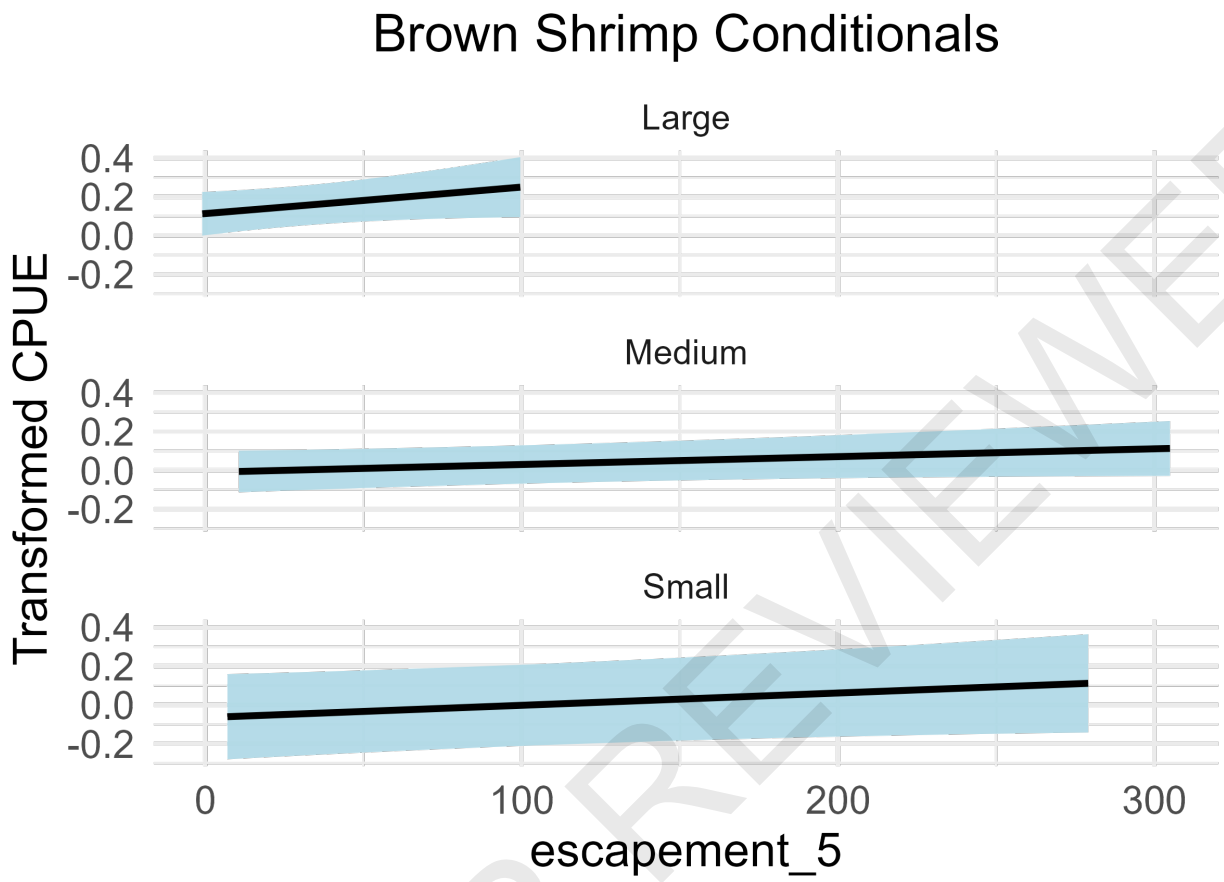

Figure 46: Length scale parameters for the 5th lag of abundance from the best performing model.

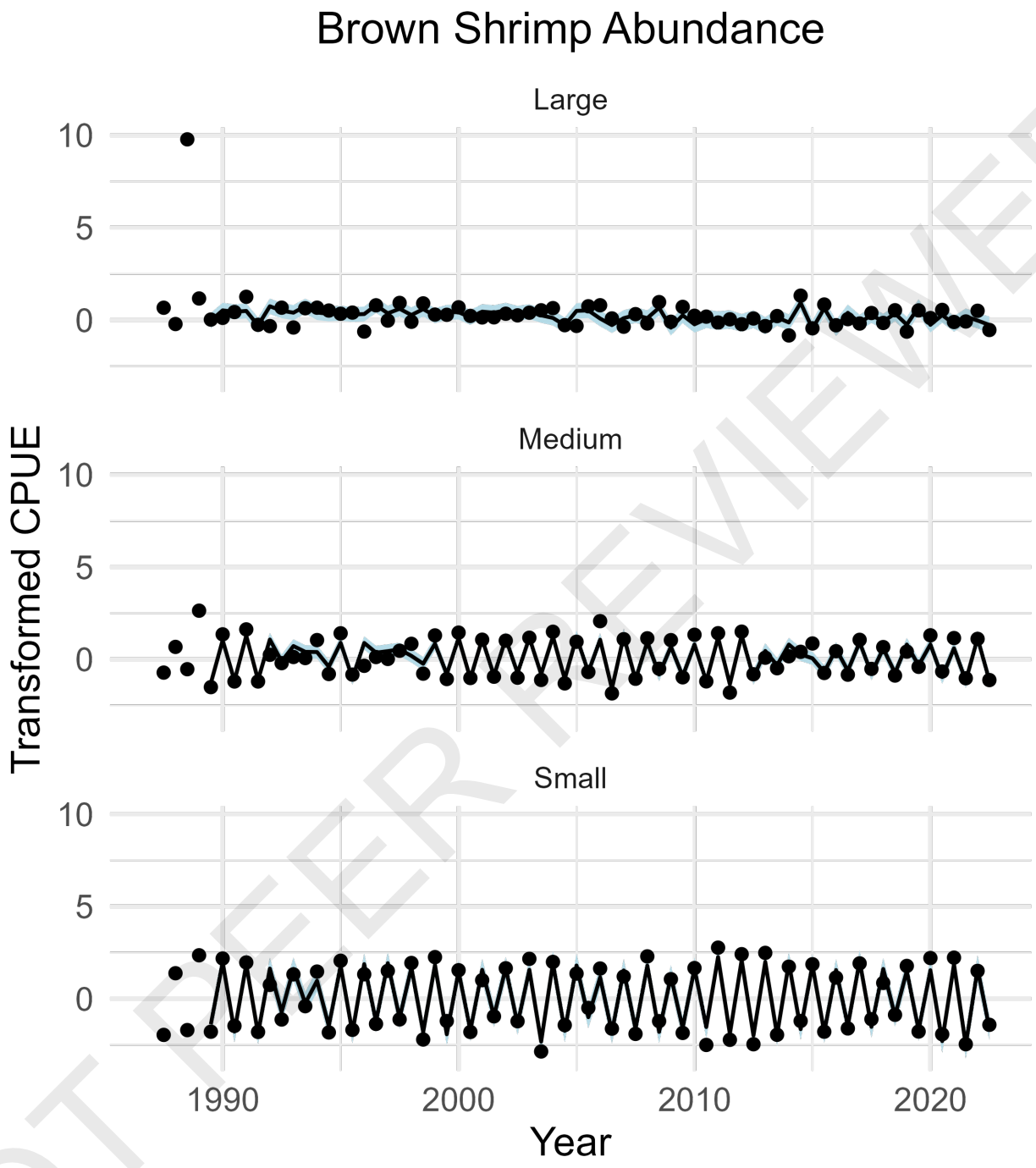

Figure 47: EDM model fits for the best performing run, transformed with error bars.

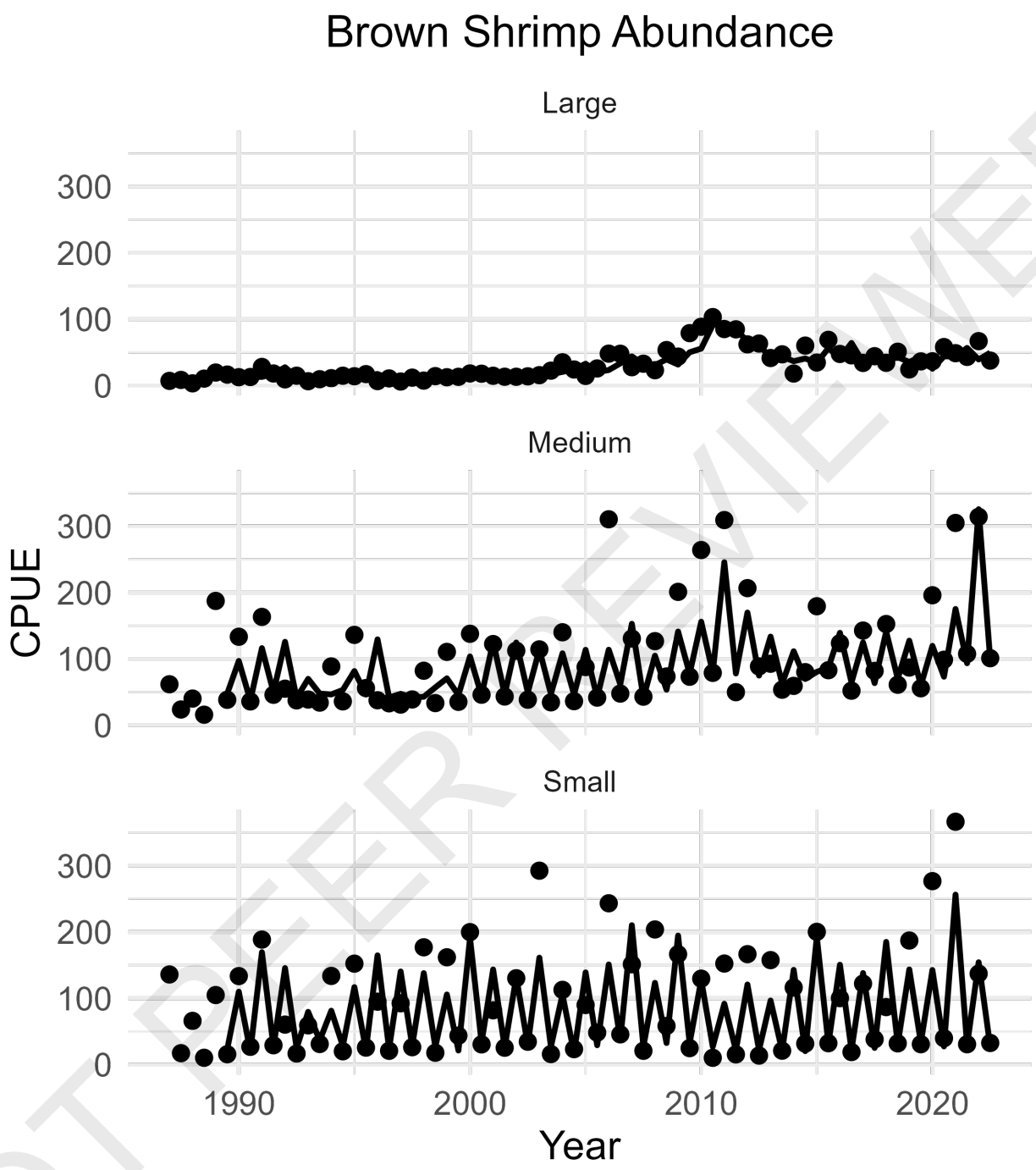

Figure 48: EDM model fits for the best performing run in raw units of SEAMAP CPUE.

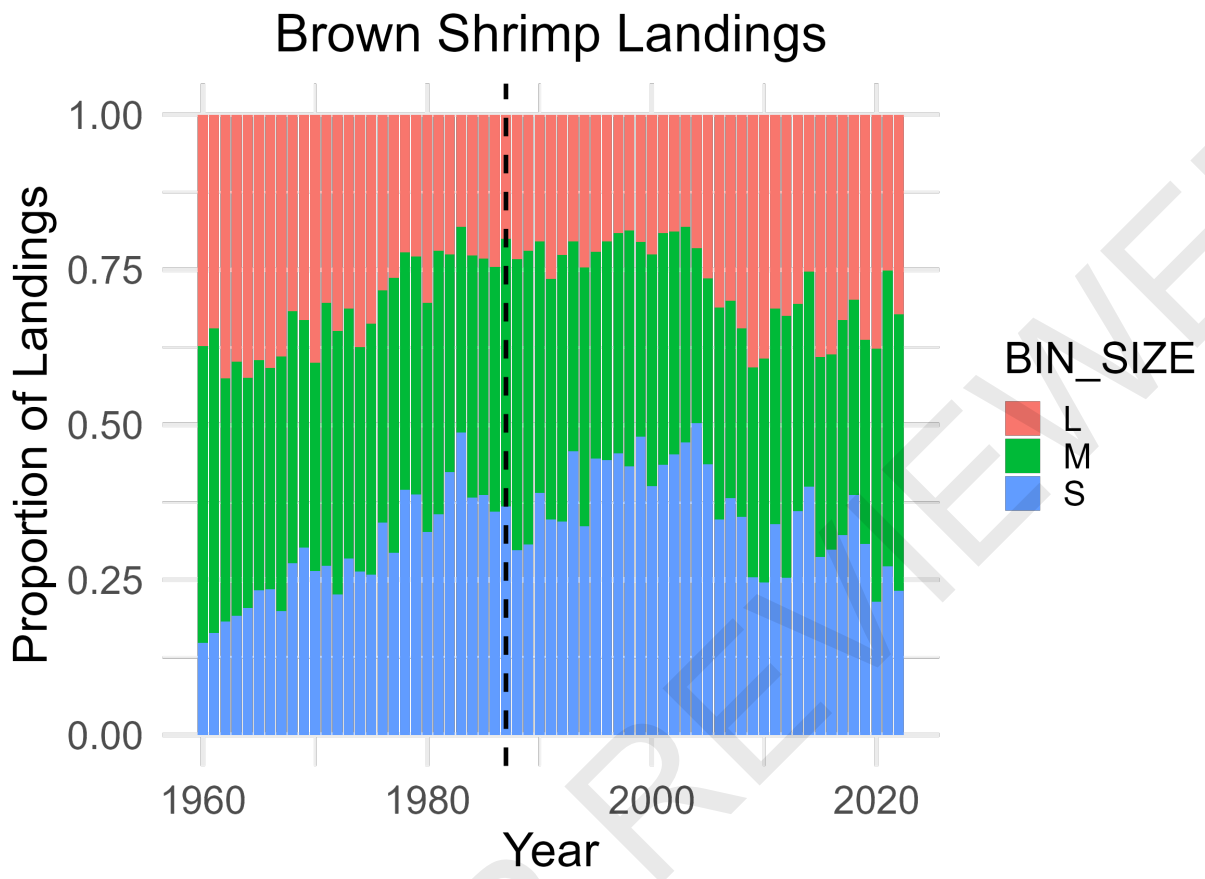

Figure 49: Proportion of landings by size class. The dashed line indicates the first year of the VAST index.