Western and Central Pacific Fisheries Commission

SCIENTIFIC COMMITTEE
TWENTY-FIRST REGULAR SESSION

Tonga
13-21 August 2025

Data-moderate assessment approaches for Southwest Pacific Ocean (SWPO) striped marlin

WCPFC-SC21-2025/SA-IP-XX
10 June 2025

N. Ducharme-Barth[1]

[1]NOAA Fisheries, Pacific Islands Fisheries Science Center, 1845 Wasp Boulevard, Bldg. 176, Honolulu HI 96818

Table of contents

1 Introduction **3**

2 Methods **4**

2.1 Model Framework . . . 4

2.1.1 Input data . . . 4

2.1.2 Population Dynamics . . . 5

2.1.3 Observation Model . . . 5

2.1.4 Catch Treatment . . . 5

2.2 Prior Development . . . 5

2.2.1 Biological Simulation Framework . . . 5

2.2.2 Maximum Intrinsic Rate of Increase (R_{max}) Prior . . . 6

2.2.3 Depletion Prior . . . 6

2.3 Model Implementation . . . 7

2.4 Model Diagnostics . . . 7

2.4.1 Convergence Diagnostics . . . 8

2.4.2 Data Fits . . . 8

2.4.3 Posterior Predictive Checks . . . 8

2.4.4 Retrospective Analysis . . . 8

2.4.5 Hindcast Cross-Validation . . . 9

2.5 Forecasts . . . 9

3 Results **9**

4 Discussion **9**

5 Declaration of Generative AI use **9**

6 References **10**

7 Tables **12**

8 Figures **15**

9 Technical Annex **18**

9.1 Equilibrium Age-Structured Population Dynamics . . . 18

9.1.1 Survival . . . 18

9.1.2 Fecundity . . . 18

9.1.3 Stock-Recruitment Relationship . . . 18

9.2 Transitional Mean Weight Model . . . 22

1 Introduction

Assessment of Southwest Pacific Ocean (SWPO) striped marlin (*Kajikia audax*) has been challenging. As documented in the joint NOAA-SPC modeling meeting report (Ducharme-Barth et al., 2025), key issues identified by WCPFC SC20 included poor fits to size composition data and relative abundance indices, conflicts between different data sources, and difficulties in estimating model initial conditions.

Discussions at the 2025 SPC Pre-Assessment Workshop also highlighted challenges related to the estimation of total population scale, where the low biomass scale indicated by previous versions of the assessment appears largely driven by the size composition from multiple fisheries. While estimation of low biomass scale is not inherently problematic, the resulting high ratio of fishing mortality (F) to natural mortality (M) that is maintained over multiple decades in the face of relative stability in catches and catch-per-unit-effort (CPUE) is suspicious and potentially indicative of additional model mis-specification.

The complexity of integrated age-structured models, while providing detailed population dynamics representation, can become problematic when fundamental data conflicts exist or when key biological parameters are uncertain. For SWPO striped marlin, these challenges are compounded by spatiotemporal heterogeneity in fleet coverage, potential non-representativeness in mixed-fleet composition data, and limitations in age data from opportunistic sampling. In the current case, changes to productivity (growth, natural mortality, maturity, and/or steepness) or selectivity assumptions will impact biomass scaling through predictions of expected size composition data.

Scaling back model complexity to a *data-moderate* approach, like Bayesian surplus production models (BSPMs), offers analysts a simplified yet robust alternative for stock assessment when data limitations or conflicts present challenges for more complex models. These models have proven effective in recent pelagic fish stock assessments (Winker et al., 2018), and are routinely applied for WCPFC shark assessments (ISC, 2024; Neubauer et al., 2019, 2024). They also facilitate a more tractable exploration of model assumptions by distilling productivity and fishing assumptions into a restricted parameter subset. The BSPM framework allows for explicit incorporation of parameter uncertainty through informative priors while maintaining computational efficiency and interpretability.

One of the advantages of the Bayesian approach is through the use of prior information to propagate uncertainty in model assumptions, and biological simulation approaches can provide a robust methods for parameterizing key population dynamics parameters (ISC, 2024; Pardo et al., 2016). For species with complex life histories like striped marlin, these simulation-based priors can incorporate uncertainty in growth, maturity, natural mortality, and reproductive parameters to derive realistic distributions for maximum intrinsic rate of increase and carrying capacity. Additionally, prior pushforward checks can be used to further refine parameter distributions by identifying parameter combinations that produce clearly improbable outcomes.

This analysis presents a data-moderate approach for SWPO striped marlin that addresses some of the key technical challenges identified in previous assessments while providing a robust framework for stock status evaluation. The model incorporates biological uncertainty through simulation-based priors and includes approaches to address catch uncertainty and population depletion. It is important to note that this analysis complements the 2025 revised stock assessment report (Castillo-Jordan et al., 2025). While the relative simplicity of data-moderate approaches offer some advantages relative to more data-intensive approaches like fully-integrated age-structured stock assessment models, particularly when uncertainty in data or key

assumptions are high, this comes at the cost of simplifying the model fishery and population dynamics. Over-simplification of key age-structured processes, or lack of spatiotemporal specificity could result in biases in data-moderate approaches. Taking multiple modeling approaches into account can give managers a holistic view on the likely status and trajectory of the stock.

2 Methods

2.1 Model Framework

In order to address some of the issues raised in Section 1, a series of Bayesian state-space surplus production models (BSPMs) spanning the period 1952-2022 were developed following the Fletcher-Schaefer production model framework (Edwards, 2024; Winker et al., 2020). Development of the BSPM followed the approach of (ISC, 2024; Neubauer et al., 2019) and incorporated recent best practices for surplus production models (Kokkalis et al., 2024) and Bayesian workflows for stock assessment (Monnahan, 2024) as guides for model development, analysis, and presentation.

The models were implemented in the Stan probabilistic programming language (Team, 2024b) to take advantage of enhanced convergence diagnostics, greater efficiency in posterior sampling through the no-U-turn (NUTS) Hamiltonian Monte Carlo algorithm (Betancourt & Girolami, 2013), and greater flexibility with model configuration and prior specification. Implementation in R using the *rstan* package (Team, 2024a) provides connection to an ecosystem of additional R packages for model diagnostics and validation: *bayesplot* (Gabry & Mahr, 2024) for posterior visualization and *loo* (Vehtari et al., 2024).

The model begins from an assumed unfished state in 1952 and incorporates process error in population dynamics, observation error in abundance indices, and uncertainty in biological parameters. Two model variants were developed: a baseline model (`0001-2024cpueExPrior`) and a depletion-constrained model (`0002-2024cpueDepPrior`) that incorporates additional information about historical stock status from mean size data. The `0002-2024cpueDepPrior` includes a mean weight based prior on 1988 depletion, and more detail on the development of this prior can be found in Section 2.2.3. Otherwise, both models used the same prior parameterizations (Table 2) based on the biological simulation filtering described in Section 2.2.

2.1.1 Input data

The BSPM was fit to catch and standardized CPUE data for SWPO striped marlin. Annual catch data (in numbers) spanning 1952-2022 were compiled from the 2024 assessment (Castillo-Jordan et al., 2024) sources and aggregated into total removals by year.

The catch series (Figure 1) shows initial low removals in 1952-1953, followed by a high peak in 1954. Overall, however, catches have been relatively stable though catches in recent decades show a slight decline. A single standardized CPUE index (1988-2022) from the diagnostic case of the 2024 assessment (Castillo-Jordan et al., 2024) was used. The index is very noisy (Figure 2). In general, however, there is a slight declining trend over the index period. The decline was most pronounced after 2000, though the trend stabilized somewhat, prior to showing a slight increase in recent years.

2.1.2 Population Dynamics

The population dynamics follow a Fletcher-Schaefer surplus production model with state-space formulation where population depletion x_t (relative to carrying capacity K) evolves according to:

$$x_t = \begin{cases} (x_{t-1} + R_{max} x_{t-1}(1 - \frac{x_{t-1}}{h}) - C_{t-1}) \times \epsilon_t, & x_{t-1} \leq D_{MSY} \\ (x_{t-1} + x_{t-1}(\gamma \times m)(1 - x_{t-1}^{n-1}) - C_{t-1}) \times \epsilon_t, & x_{t-1} > D_{MSY} \end{cases}$$

where R_{max} is the maximum intrinsic rate of increase, n is the shape parameter, and ϵ_t represents multiplicative process error with $\epsilon_t = \exp(\delta_t - \sigma_P^2/2)$ and $\delta_t \sim N(0, \sigma_P)$. The intermediate parameters are: $D_{MSY} = (1/n)^{1/(n-1)}$, $h = 2D_{MSY}$, $m = R_{max}h/4$, and $\gamma = n^{n/(n-1)}/(n-1)$.

2.1.3 Observation Model

The model fits to a standardized CPUE index using a lognormal likelihood:

$$I_t \sim \text{Lognormal}(\log(q \times x_t) - \frac{\sigma_{O,t}^2}{2}, \sigma_{O,t})$$

where catchability q is analytically derived assuming an uninformative uniform prior, and total observation error $\sigma_{O,t} = \sigma_{O,input} + \sigma_{O,add}$ combines fixed input uncertainty with an estimated additional error component.

2.1.4 Catch Treatment

Annual fishing mortality is estimated directly $F_t \sim N^+(0, \sigma_F)$ with catch fitted using lognormal likelihood with uncertainty $\sigma_C = 0.2$

Catch is predicted as:

$$C_t = (x_t + \text{surplus production}_t) \times (1 - \exp(-F_t)) \times \epsilon_t \times K$$

2.2 Prior Development

2.2.1 Biological Simulation Framework

Informative priors for key population parameters were developed through biological simulation using Monte Carlo sampling from biologically plausible parameter distributions. The simulation framework incorporated uncertainty and correlations in life history parameters:

- **Growth parameters**: Francis parameterization with independent sampling
- **Natural mortality**: Reference mortality M_{ref} negatively correlated with maximum age (r = -0.3)
- **Maturity**: Length-based logistic function parameters
- **Reproduction**: Steepness parameter for stock-recruit relationship and sex ratio
- **Weight-length relationship**: Correlated allometric parameters (r = -0.5) with biological constraints
- **Fishery selectivity**: Length at first capture for depletion prior

Parameter combinations (Table 1) were filtered to ensure biological realism and included correlations between life history parameters reflecting biological trade-offs (e.g., shorter lived individuals having higher natural mortality rates).

2.2.2 Maximum Intrinsic Rate of Increase (R_{max}) Prior

R_{max} was calculated by numerically solving the Euler-Lotka equation using bounded optimization (nlminb function in R) to find the value of R_{max} that satisfies the equation within convergence criteria. Starting values and bounds were set to ensure biologically realistic solutions. Formulation of the Euler-Lotka equation followed the approach implemented in openMSE (Hordyk et al., 2024; Stanley et al., 2009):

$$\alpha \sum_{a=1}^{A_{max}} l_a f_a \exp(-R_{max} \times a) = 1$$

where α is the slope at the origin of the stock-recruitment curve, l_a is the probability of survival to age a, and f_a is the fecundity (reproductive output) at age a. Additional technical detail can be found in Section 9.1 of the Technical Annex.

Repeating this calculation across each of the 500,000 parameter combinations resulted in a prior distribution of R_{max} conditioned on plausible biology and life-history characteristics. The resulting distribution was first filtered to values of $R_{max} < 1.5$ representative of what the literature (Gravel et al., 2024; Hutchings et al., 2012) indicates is most likely for teleosts. A prior pushforward analysis, similar to (ISC, 2024), was used to further refine this distribution and develop a lognormal prior for R_{max}. Briefly, random values of R_{max} along with random values of the other leading parameters of the BSPM (drawn from the prior distributions described in Table 2) were used to drive simulated populations forward subject to direct removal of the observed catch (Figure 1). Based on the results of the prior pushforward (Figure 3), the R_{max} distribution was further filtered based on the following criteria:

- Depletion in the terminal year $> 2\%$ and $R_{max} < 1$
- Depletion in the terminal year $> 2\%$, $R_{max} < 1$, and trend in depletion over the last decade between -5% and 10%

This last filtering step approximates the recent relative trend seen in the index (Figure 2). Fitting a lognormal distribution to the remaining R_{max} values results in the prior distribution specified in Table 2 and shown in Figure 4. Note that though initial filtering restricted $R_{max} < 1$ the final lognormal prior still assigns some prior weight to $R_{max} > 1$.

2.2.3 Depletion Prior

It is not possible to fit to size composition data in a BSPM framework, however information on fishing mortality and depletion may be contained in the long time series of New Zealand recreational fishing data. This fishery routinely catches the largest individuals and shows a decline in average size since the start of the model period (Figure 5), prior to which limited large-scale commercial fishing occurred. Assuming constant availability and selectivity, declining mean weight can be an indication of increased mortality. Assuming natural mortality M has remained constant, this indicates that the population may be in a more depleted state than in the initial period. The `0002-2024cpueDepPrior` variant was developed to capitalize on this potential additional source of information.

A supplementary analysis was used to estimate relative population depletion in 1988 using New Zealand recreational fishing weight data. A modified Gedamke-Hoenig approach (Gedamke & Hoenig, 2006) for using size composition data to estimate mortality in non-equilibrium situations was applied to estimate change in total mortality from Z_1 (initial period) to Z_2 (final period) at a known change point (1952) when industrial fishing commenced at the start of the model period (more detail in Technical Annex Section 9.2). These two total mortality rates were calculated for each of the biological parameter combinations described in Section 2.2.1 and used to calculate relative spawning stock biomass depletion as the ratio of spawning potential under the two mortality regimes:

$$\text{Relative Depletion} = \frac{SPR_{Z_2}}{SPR_{Z_1}} = \frac{\sum_{a=1}^{A_{max}} l_{a,Z_2} f_a}{\sum_{a=1}^{A_{max}} l_{a,Z_1} f_a}$$

where $l_{a,Z}$ is survival to age a under total mortality Z, and f_a is reproductive output at age a (incorporating maturity, weight, sex ratio, and reproductive cycle). Additional technical detail on the population dynamics components of this equation can be found in Section 9.1 of the Technical Annex.

The resulting depletion estimates across successful parameter combinations were used to construct an informative lognormal prior for stock depletion that could be applied in the BSPM in 1988 (Figure 6).

2.3 Model Implementation

Models were implemented using the No-U-Turn Sampler (NUTS) in Stan. Each model was run using 5 independent Markov chains with randomly generated starting values to ensure robust exploration of the posterior parameter space. Each chain was sampled for 3,000 iterations, with the first 1,000 iterations discarded as warmup to allow for algorithm adaptation. To reduce posterior sample autocorrelation and storage requirements, every 10 samples was retained from the post-warmup iterations, yielding a final sample size of 1,000 draws from the joint posterior distribution across all chains. The NUTS algorithm was configured with strict adaptation parameters to ensure reliable convergence. The adaptation delta was set to 0.99 to reduce the step size and minimize divergent transitions, while the maximum tree depth was increased to 12 to allow for longer trajectories during the dynamic sampling process.

Model convergence and performance were assessed using multiple diagnostic criteria recommended for Bayesian stock assessment workflows (Monnahan, 2024). The potential scale reduction factor ($\hat{R}$) was required to be less than 1.01 for all parameters, indicating convergence across chains. Effective sample sizes were required to exceed 500 to ensure adequate precision in posterior estimates. Additionally, posterior predictive checks were conducted to evaluate model adequacy through comparison of observed versus predicted index and catch values, while parameter identifiability was assessed through visual examination of posterior distributions and correlation structures among estimated parameters (Gabry et al., 2019).

2.4 Model Diagnostics

Model performance was comprehensively evaluated using multiple diagnostic criteria recommended for Bayesian stock assessment workflows (Monnahan, 2024). The diagnostic framework included Stan model convergence criteria, data fits, posterior predictive checks, retrospective analysis, and hindcast cross-validation.

2.4.1 Convergence Diagnostics

Conventional Stan model convergence diagnostics and thresholds were employed to identify whether posterior distributions were representative and unbiased (Monnahan, 2024). Models were considered to have 'converged' to a stable, unbiased posterior distribution if they satisfied the following criteria:

- **Potential scale reduction factor** ($\hat{R}$) less than 1.01 for all leading model parameters, indicating convergence across chains
- **Bulk effective sample size** (N_{eff}) greater than 500 for all leading model parameters, ensuring adequate precision in posterior estimates

- **No divergent transitions** during the NUTS sampling process, which can indicate problematic posterior geometry
- **Maximum tree depth** not exceeded during sampling, ensuring adequate exploration of parameter space

Additionally, trace plots and rank plots were examined to visually assess mixing and convergence behavior across chains (Gabry et al., 2019).

2.4.2 Data Fits

Model fit to different data sources (abundance indices and catch observations) was quantified using normalized root-mean-squared error (NRMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}$$

$$NRMSE = \frac{RMSE}{\sum_{i=1}^{n} y_i/n}$$

where y_i are the observed values and $\hat{y}_i$ are the model predictions. The average NRMSE across all posterior samples was calculated for each data component.

2.4.3 Posterior Predictive Checks

Posterior predictive checks were conducted to evaluate model adequacy by comparing observed data with data simulated from the fitted model. For each posterior sample, synthetic datasets were generated using the estimated parameters, and agreement between observed and predicted datasets was assessed visually.

2.4.4 Retrospective Analysis

Retrospective analysis was conducted for each model by sequentially peeling off a year from the terminal end of the fitted index and re-running the model. Data were removed for each year up to five years from 2022 to 2018. Estimates of x in the terminal year of each retrospective peel were compared to the corresponding estimate of x from the full model run to better understand any potential biases or uncertainty in terminal year estimates. The Mohn's ρ statistic (Mohn, 1999) was calculated and presented. This statistic measures the average relative difference between an estimated quantity from an assessment (e.g., depletion in final

year) with a reduced time-series of information and the same quantity estimated from an assessment using the full time-series.

Mohn's ρ was calculated as: $\rho = \frac{1}{n}\sum_{i=1}^{n}\frac{\hat{\theta}T-i,T-i-\hat{\theta}T,T-i}{\hat{\theta}_{T,T-i}}$ where $\hat{\theta}T-i,T-i$ is the estimate of parameter θ in year $T-i$ from the model with data through year $T-i$, and $\hat{\theta}T,T-i$ is the estimate for year $T-i$ from the full model.

Additionally, following Kokkalis et al. (2024) the proportion of retrospective peels where the relative exploitation rate (U/U_{MSY}) and relative depletion (D/D_{MSY}) inside the credible intervals of the full model run were calculated. This metric provides an additional perspective on retrospective consistency by evaluating whether the uncertainty estimates from the full model appropriately capture the range of estimates obtained when using reduced datasets.

2.4.5 Hindcast Cross-Validation

Hindcast cross-validation (Kell et al., 2021) was conducted for each index to determine the performance of the model to predict the observed CPUE I one-step-ahead into the future relative to a naïve predictor. Briefly, the 'model-free' approach to hindcast cross-validation was used, and made use of the same set of five retrospective peels described in Section 2.4.4.

The 'model-free' hindcast calculation is described using the model from the last peel $BSPM_{2017}$ as an example. This model fit to index data through 2017 but included catch through 2022. The model estimates of predicted CPUE in 2018 based on $BSPM_{2017}$ is the 'model-free' hindcast for 2018, $\hat{I}_{2018}$. The naïve prediction of CPUE in 2018 is simply the observed CPUE from 2017, $\ddot{I}_{2017} = I_{2018}$. The absolute scaled error (ASE) of the prediction is:

$$ASE_{2018} = \frac{|I_{2018}-\hat{I}_{2018}|}{|I_{2018}-\ddot{I}_{2018}|}$$

Repeating this calculation across all retrospective peels for years 2019-2022 and taking the average across ASE values gives the mean ASE or MASE for the model. An MASE value less than one indicates that the model has greater predictive skill than the naïve predictor.

2.5 Forecasts

Simple, status quo based catch projections over a 10 year window are provided for each model. [More detail to be filled in here...]

3 Results

4 Discussion

5 Declaration of Generative AI use

A generative artificial intelligence (AI) assistant, Anthropic Claude Sonnet 4.0, was used to parse model code in the preparation of an initial draft of this report.

6 References

Betancourt, M. J., & Girolami, M. (2013). *Hamiltonian Monte Carlo for Hierarchical Models*. https://doi.org/10.48550/ARXIV.1312.0906

Castillo-Jordan, C., Day, J., Teears, T., Davies, N., Hampton, J., McKechnie, S., Magnusson, A., Peatman, T., Vidal, T., Williams, P., & Hamer, P. (2024). *Stock Assessment of Striped Marlin in the Southwest Pacific Ocean: 2024* (WCPFC-SC20-2024/SA-WP-03 (Ver. 2.01)). https://meetings.wcpfc.int/node/23120

Castillo-Jordan, C., Day, J., Teears, T., Davies, N., Hampton, J., McKechnie, S., Magnusson, A., Peatman, T., Vidal, T., Williams, P., & Hamer, P. (2025). *Revised Stock Assessment of Striped Marlin in the Southwest Pacific Ocean: 2025* (WCPFC-SC21-2025/SA-WP-XX).

Ducharme-Barth, N., Castillo-Jordan, C., Sculley, M., Ahrens, R., & Carvalho, J. (2025). *Summary report from the NOAA-SPC assessment model meeting on SWPO striped marlin* (WCPFC-SC21-2025/SA-IP-XX).

Edwards, C. (2024). *Bdm: Bayesian biomass dynamics model* (R package version 0.0.0.9036).

Farley, J., Eveson, P., Krusic-Golub, K., & Kopf, K. (2021). *Southwest Pacific Striped Marlin Population Biology (Project 99)* (WCPFC-SC-17/SA-IP-11). https://meetings.wcpfc.int/node/12569

Gabry, J., & Mahr, T. (2024). *Bayesplot: Plotting for Bayesian Models* (R package version 1.11.1). https://mc-stan.org/bayesplot/

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). Visualization in Bayesian Workflow. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *182*(2), 389–402. https://doi.org/10.1111/rssa.12378

Gedamke, T., & Hoenig, J. M. (2006). Estimating Mortality from Mean Length Data in Nonequilibrium Situations, with Application to the Assessment of Goosefish. *Transactions of the American Fisheries Society*, *135*(2), 476–487. https://doi.org/10.1577/T05-153.1

Gravel, S., Bigman, J. S., Pardo, S. A., Wong, S., & Dulvy, N. K. (2024). Metabolism, population growth, and the fast-slow life history continuum of marine fishes. *Fish and Fisheries*, *25*(2), 349–361. https://doi.org/10.1111/faf.12811

Hamel, O. S., & Cope, J. M. (2022). Development and considerations for application of a longevity-based prior for the natural mortality rate. *Fisheries Research*, *256*, 106477. https://doi.org/10.1016/j.fishres.2022.106477

Hordyk, A., Huynh, Q., & Carruthers, T. (2024). *openMSE: Easily Install and Load the 'openMSE' Packages* (R package version 1.0.1). https://openmse.com/

Hutchings, J. A., Myers, R. A., García, V. B., Lucifora, L. O., & Kuparinen, A. (2012). Life-history correlates of extinction risk and recovery potential. *Ecological Applications*, *22*(4), 1061–1067. https://doi.org/10.1890/11-1313.1

ISC. (2024). *Stock Assessment of Shortfin Mako Shark in the North Pacific Ocean through 2022* (WCPFC-SC20-2024/SA-WP-14). https://meetings.wcpfc.int/node/22828

Kell, L. T., Sharma, R., Kitakado, T., Winker, H., Mosqueira, I., Cardinale, M., & Fu, D. (2021). Validation of stock assessment methods: Is it me or my model talking? *ICES Journal of Marine Science*, *78*(6), 2244–2255. https://doi.org/10.1093/icesjms/fsab104

Kokkalis, A., Berg, C. W., Kapur, M. S., Winker, H., Jacobsen, N. S., Taylor, M. H., Ichinokawa, M., Miyagawa, M., Medeiros-Leal, W., Nielsen, J. R., & Mildenberger, T. K. (2024). Good practices for surplus production models. *Fisheries Research*, *275*, 107010. https://doi.org/10.1016/j.fishres.2024.107010

Mohn, R. (1999). The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. *ICES Journal of Marine Science*, *56*(4), 473–488. https://doi.org/10.1006/

jmsc.1999.0481

Monnahan, C. C. (2024). Toward good practices for Bayesian data-rich fisheries stock assessments using a modern statistical workflow. *Fisheries Research*, *275*, 107024. https://doi.org/10.1016/j.fishres.2024.107024

Neubauer, P., Kim, K., Large, K., & Brouwer, S. (2024). *Stock Assessment of Silky Shark in the Western and Central Pacific Ocean 2024* (WCPFC-SC20-2024/SA-WP-04-Rev2). https://meetings.wcpfc.int/node/23088

Neubauer, P., Richard, Y., & Tremblay-Boyer, L. (2019). *Alternative assessment methods for oceanic whitetip shark* (WCPFC-SC15-2019/SA-WP-13).

Pardo, S. A., Kindsvater, H. K., Reynolds, J. D., & Dulvy, N. K. (2016). Maximum intrinsic rate of population increase in sharks, rays, and chimaeras: The importance of survival to maturity. *Canadian Journal of Fisheries and Aquatic Sciences*, *73*(8), 1159–1163. https://doi.org/10.1139/cjfas-2016-0069

Stanley, R. D., McAllister, M., Starr, P., & Olsen, N. (2009). *Stock assessment for bocaccio (*Sebastes paucispinis*) in British Columbia waters* (Canadian {Science} {Advisory} {Secretariat} Research Document 2009/055). Fisheries; Oceans Canada. https://www.dfo-mpo.gc.ca/csas-sccs/publications/resdocs-docrech/2009/2009_055-eng.htm

Team, S. D. (2024a). *RStan: The R interface to Stan* (R package version 2.32.6). https://mc-stan.org/

Team, S. D. (2024b). *Stan Modeling Language Users Guide and Reference Manual, v 2.32.6.*

Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Burkner, P., Paananen, T., & Gelman, A. (2024). *Loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models* (R package version 2.7.0). https://mc-stan.org/loo/

Winker, H., Carvalho, F., & Kapur, M. (2018). JABBA: Just Another Bayesian Biomass Assessment. *Fisheries Research*, *204*, 275–288. https://doi.org/10.1016/j.fishres.2018.03.010

Winker, H., Mourato, B., & Chang, Y. (2020). Unifying parameterizations between age-structured and surplus production models: An application to atlantic white marlin (*Kajiki albidia*) with simulation testing. *Collect. Vol. Sci. Pap. ICCAT*, *76*(4), 219–234.

7 Tables

Table 1: Biological parameter distributions and sources used in simulation framework

Parameter	Symbol	Distribution	Mean/Mode	CV/Range	Correlation	Source/Notes
Length at age 1	L_1	Lognormal	60 cm	0.2	Independent	(Ducharme-Barth et al., 2025)
Length at age 10	L_2	Lognormal	210 cm	0.2	Independent	(Ducharme-Barth et al., 2025)
Growth coefficient	k	Beta(6.5, 3.5)	~0.65	-	Independent	(Ducharme-Barth et al., 2025)
Length CV	cv_{len}	Uniform	-	[0.05, 0.25]	-	Growth variability
Maximum age	A_{max}	Lognormal	15 years	0.2	r = -0.3 with M_{ref}	(Farley et al., 2021)
Reference mortality	M_{ref}	Lognormal	0.36 yr^{-1}	0.44	r = -0.3 with A_{max}	(Hamel & Cope, 2022) ($5.4/A_{max}$)
Maturity slope	a_{mat}	Normal	-20	0.2	-	Logistic steepness
Length at 50% maturity	L_{50}	Lognormal	184 cm	0.2	-	(Farley et al., 2021)
Weight coefficient	a_w	Lognormal	5.4×10^{-7}	0.05	r = -0.5 with b_w	(Castillo-Jordan et al., 2024)
Weight exponent	b_w	Normal	3.58	0.05	r = -0.5 with a_w	(Castillo-Jordan et al., 2024) Constrained: 150-350 kg at 300 cm
Steepness	h	Beta(3, 1.5)	~0.73	-	-	Censored to [0.2, 1.0]
Sex ratio (prop. female)	s	Normal	0.5	0.05	-	Constrained [0.01, 0.99]
Reproductive cycle	-	Fixed	1 year	-	-	Annual spawning
Length at first capture	L_c	Lognormal	140 cm	0.125	-	Fishery selectivity

Parameter	Symbol	Distribution	Mean/Mode	CV/Range	Correlation	Source/Notes
Reference ages	age_1, age_2	Fixed	0, 10 years	-	-	Francis parameterization

Table 2: Prior distributions for Bayesian surplus production model parameters

Parameter	Distribution	Mean	SD	Description
Core Population Parameters				
K	Lognormal	$\log(1,049,036)$	0.46	Carrying capacity
R_{max}	Lognormal	$\log(0.5099)$	0.46	Maximum intrinsic rate of increase (from biological simulation)
n	Lognormal	$\log(2)$	0.1	Shape parameter (Pella-Tomlinson production function); $n = 2$ is a Schaefer model
Process and Observation Error				
σ_P	Lognormal	$\log(0.0533)$	0.27	Process error standard deviation (Winker et al., 2018)
$\sigma_{O,add}$	Half-Normal	0	0.2	Additional observation error
σ_F	Half-Normal	0	0.025	Fishing mortality variability
Depletion Constraint (Constrained Model Only)				
x_{1988}	Lognormal	$\log(0.6374)$	0.2	Initial depletion at $t = 37$ (this analysis)

8 Figures

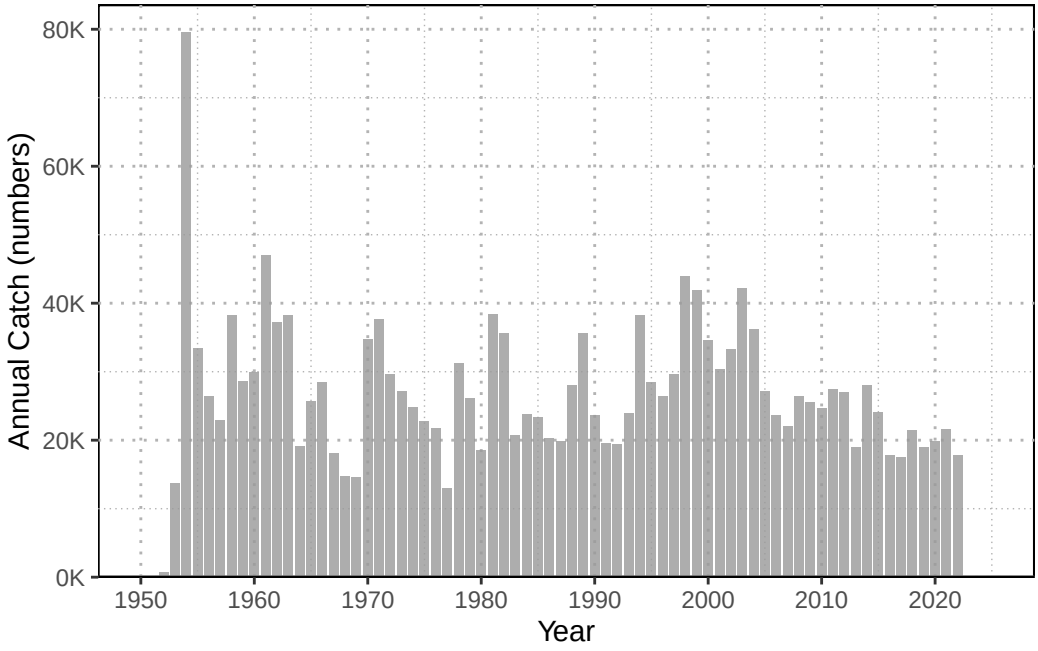

Figure 1: Annual catch of Southwest Pacific Ocean striped marlin (1952-2022). Catch data shows initial low removals in early years, a peak in 1954, followed by relatively stable catches with a slight decline in recent decades.

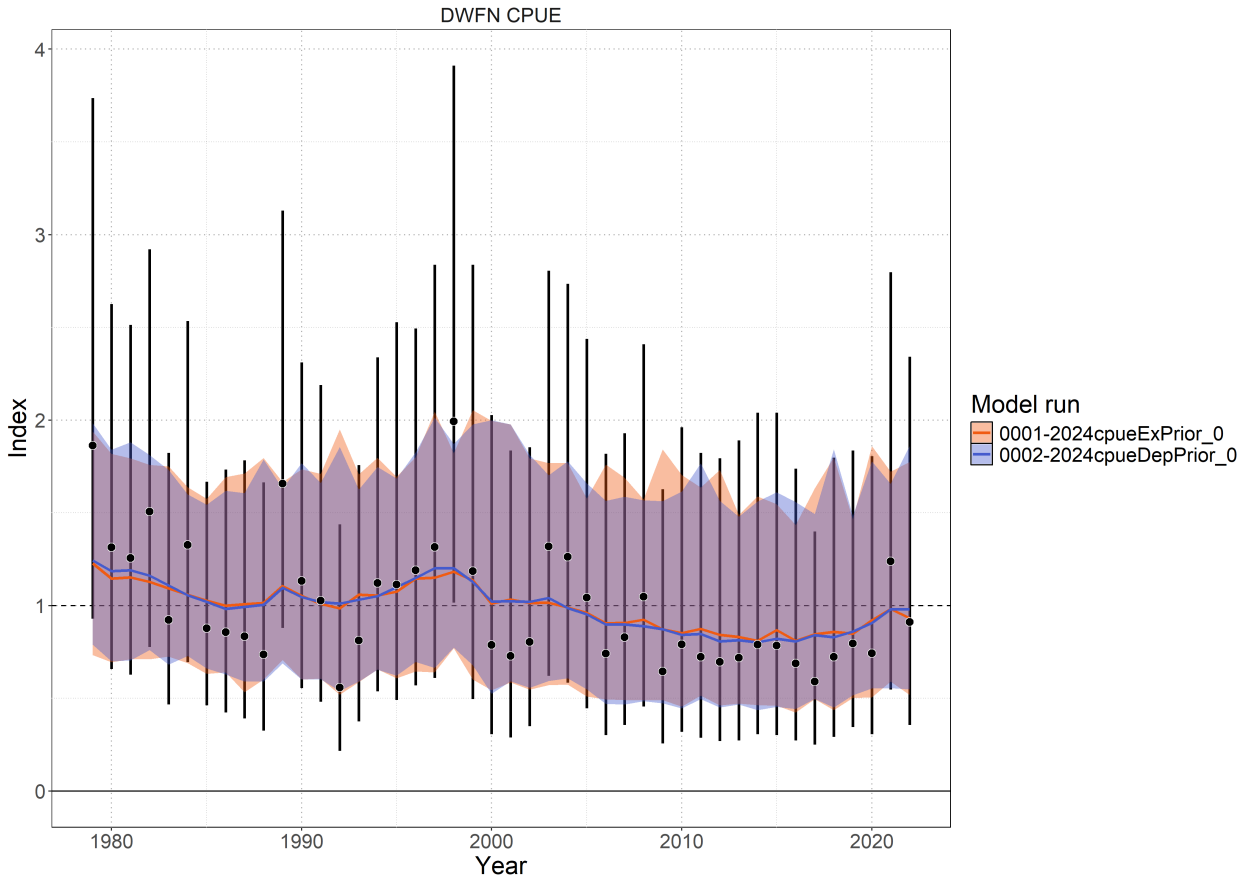

Figure 2: Standardized index (black points), observation error (black bars), and posterior predicted model fits (colored lines) with associated 95% credible interval (colored shading).

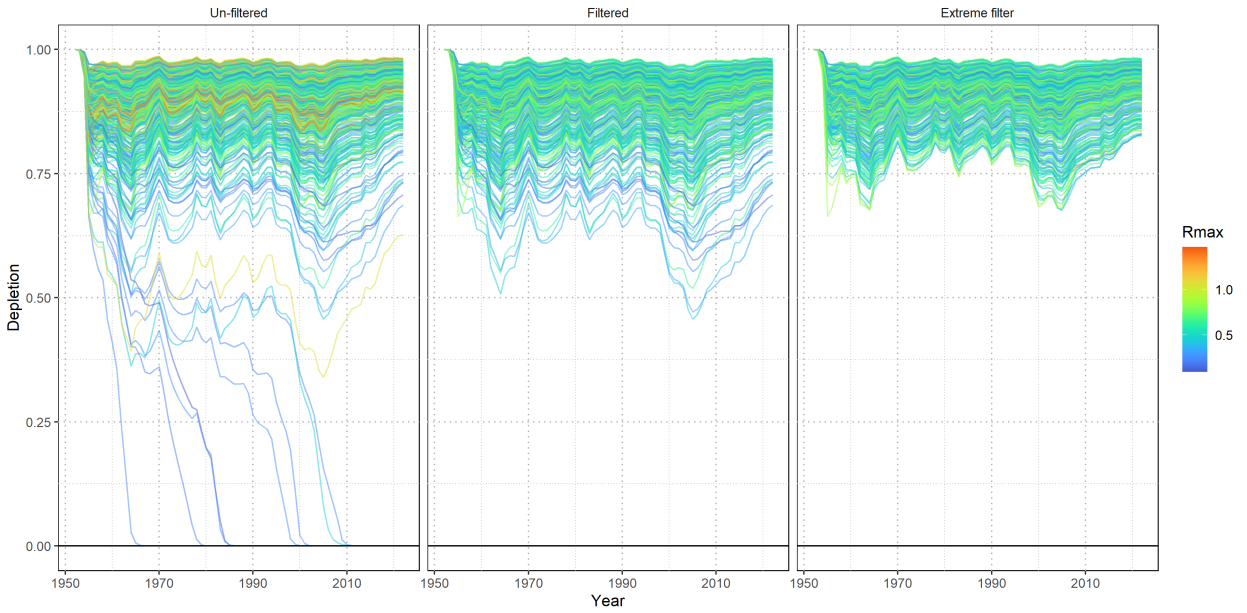

Figure 3: Prior pushforward check for population depletion trajectories under different biological parameter filtering scenarios. Each line represents a simulated population trajectory colored by maximum intrinsic rate of increase (R_{max}). Left panel shows unfiltered R_{max}, middle panel shows filtered R_{max}, and right panel shows extreme filtering.

9 Technical Annex

Additional technical detail of population dynamics components used to solve the Euler-Lotka equation for R_{max} and implement the Gedamke-Hoenig approach (Gedamke & Hoenig, 2006) for using size composition data to estimate mortality in non-equilibrium situations.

9.1 Equilibrium Age-Structured Population Dynamics

9.1.1 Survival

Survival schedule was calculated assuming constant natural mortality:

$$l_a = \begin{cases} 1 & \text{if } a = 1 \\ l_{a-1} \times \exp(-M_{ref}) & \text{if } 1 < a < A_{max} \\ \frac{l_{A_{max}-1}}{1-\exp(-M_{ref})} & \text{if } a = A_{max} \text{ (plus group)} \end{cases}$$

9.1.2 Fecundity

Reproductive output incorporated biological constraints:

$$f_a = \frac{\psi_a \times W_a \times s}{\rho}$$

where: - ψ_a = proportion mature at age a - W_a = weight at age a (proxy for reproductive capacity) - s = sex ratio (proportion female) - ρ = reproductive cycle (years between spawning events)

Maturity-at-age was derived from length-based logistic maturity:

$$\psi_L = \frac{\exp(a_{mat} + b_{mat} \times L)}{1 + \exp(a_{mat} + b_{mat} \times L)}$$

where $b_{mat} = -a_{mat}/L_{50}$, then integrated over the length-at-age distribution to obtain ψ_a.

Length-at-age followed the Francis parameterization:

$$L_a = L_1 + (L_2 - L_1) \times \frac{1 - \exp(-k(a - age_1))}{1 - \exp(-k(age_2 - age_1))}$$

with length variability modeled using probability density functions linking length bins to ages.

Weight-at-age used the allometric relationship:

$$W_a = a_w \times L_a^{b_w}$$

9.1.3 Stock-Recruitment Relationship

The slope at the origin of the stock-recruitment curve was:

$$\alpha = \frac{4h}{(1-h)\phi_0}$$

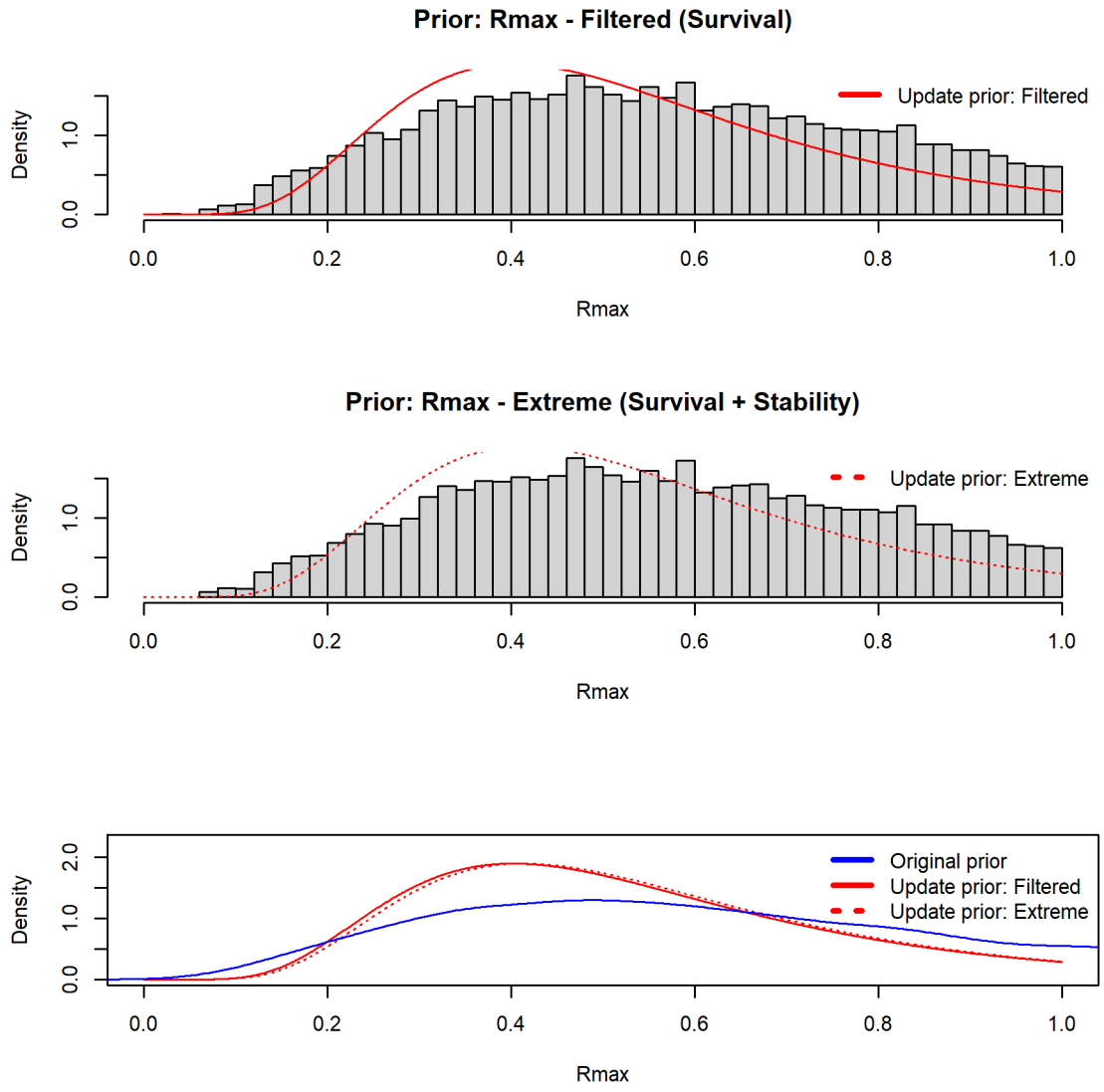

Figure 4: Prior distributions for maximum intrinsic rate of population increase R_{max}. Upper panel: Gray histogram is the R_{max} values from the numerical simulation which meet baseline filtering levels. Red line is fitted lognormal distribution. Middle panel: Gray histogram is the R_{max} values from the numerical simulation which meet extreme filtering levels. Dotted red line is fitted lognormal distribution. Bottom panel: Original distribution of R_{max} values from numerical simulation (gray), those from viable populations (blue), and the two lognormal priors (red).

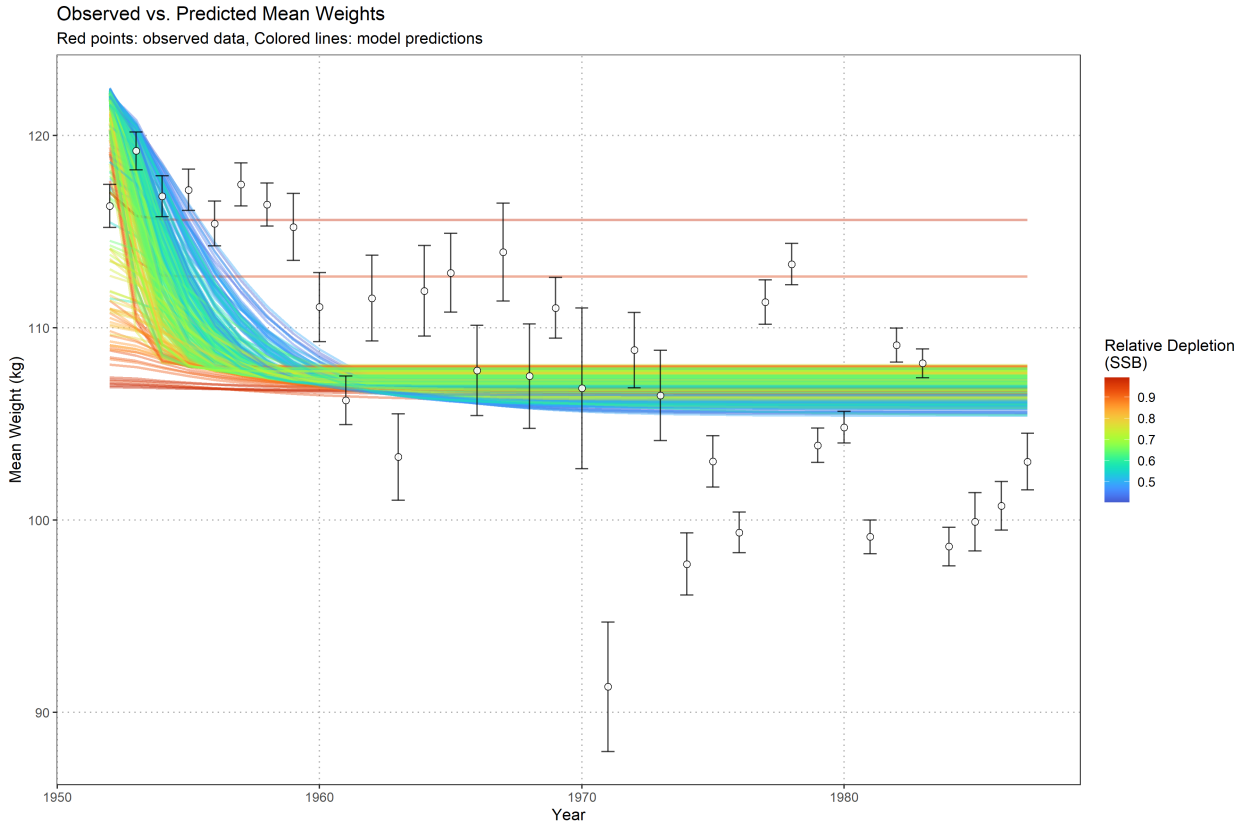

Figure 5: Observed versus predicted mean weights from New Zealand recreational fishery data. Points show observed mean weights with error bars, colored lines show model predictions from the transitional mean weight model colored by relative spawning stock biomass (SSB) depletion.

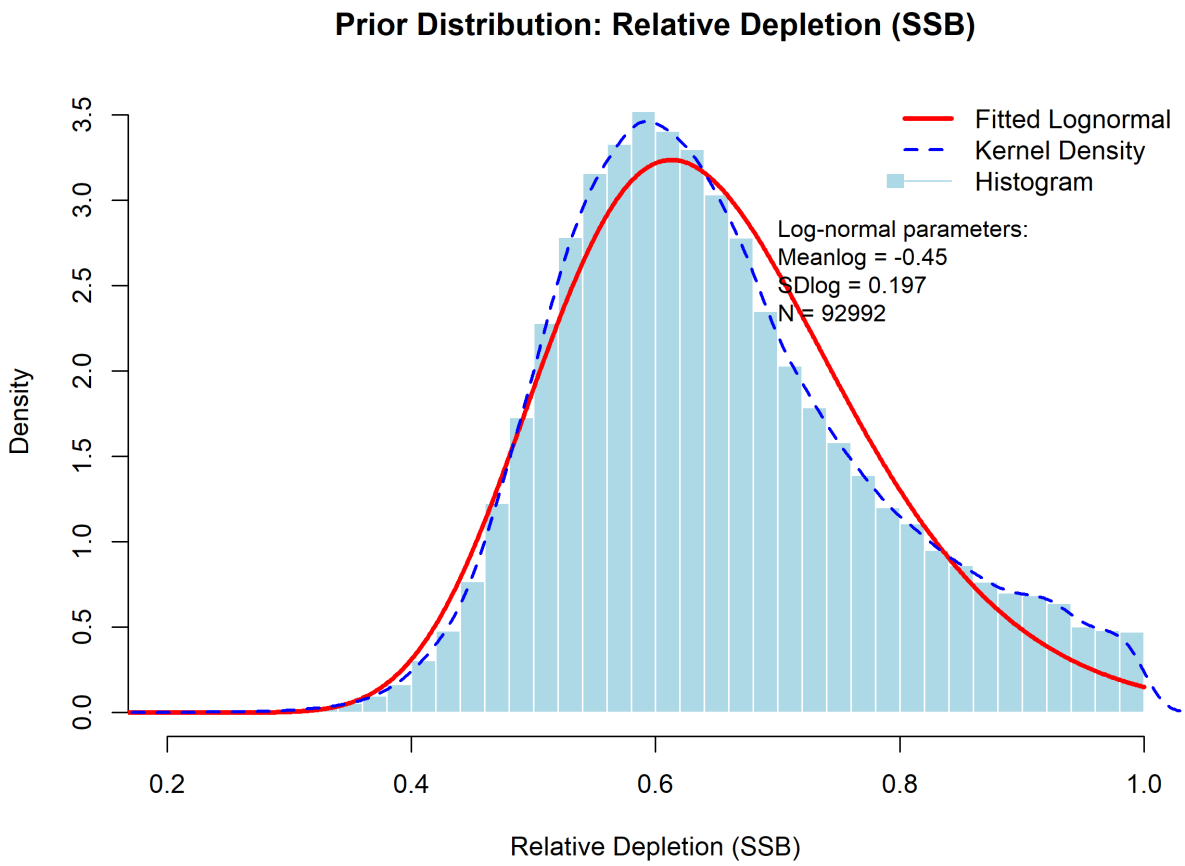

Figure 6: Prior distribution for relative depletion in 1988 derived from New Zealand recreational weight data. Light blue histogram shows depletion estimates from biological parameter combinations that successfully fit the transitional mean weight model. Blue dashed line shows kernel density estimate, red line shows fitted lognormal distribution used as prior in the depletion-constrained BSPM.

The calculation incorporated steepness (h) through the Beverton-Holt stock-recruitment relationship. Spawners-per-recruit in the unfished state was:

$$\phi_0 = \sum_{a=1}^{A_{max}} l_a f_a$$

This α parameter links the intrinsic rate of increase to recruitment compensation, ensuring that R_{max} estimates reflect realistic population productivity under the assumed stock-recruitment dynamics.

Successful simulations (those yielding $R_{max} > 0$ and $R_{max} < 1.5$) were filtered to create a plausible prior distribution (Figure 3). The resulting distribution was fitted to a lognormal prior for use in the BSPM (Figure 4).

9.2 Transitional Mean Weight Model

For each biological parameter combination, total mortality rates Z_1 and Z_2 were estimated by fitting a transitional mean weight model to observed temporal changes in recreational catch weights. The expected mean weight during the transition follows:

$$\bar{W}(d) = W_\infty - \frac{Z_1 Z_2 (W_\infty - W_c)[Z_1 + k + (Z_2 - Z_1)e^{-(Z_2+k)d}]}{(Z_1 + k)(Z_2 + k)[Z_1 + (Z_2 - Z_1)e^{-Z_2 d}]}$$

where: - d = time since mortality change (years after 1952) - W_∞ = asymptotic weight from growth parameters - W_c = weight at first capture (corresponding to L_c) - k = von Bertalanffy growth coefficient

The mortality parameters were estimated in a likelihood framework by maximizing the likelihood of the observed mean weights given a time series of predicted mean weights $\bar{W}(d)$. Optimization was constrained such that $Z_1, Z_2 \geq M_{ref}$ to ensure biologically realistic mortality estimates.