Fisheries New Zealand
Tini a Tangaroa

# Stan-ASD: a new age-structured stock assessment model, with an application to sub-Antarctic hake (*Merluccius australis*) and ling (*Genypterus blacodes*)

New Zealand Fisheries Assessment Report 2021/59

D.N. Webber,
A. Dunn,
S. Mormede

ISSN 1179-5352 (online)
ISBN 978-1-99-101935-6 (online)

**October 2021**

New Zealand Government

Requests for further copies should be directed to:

Publications Logistics Officer
Ministry for Primary Industries
PO Box 2526
WELLINGTON 6140

Email: brand@mpi.govt.nz
Telephone: 0800 00 83 33
Facsimile: 04-894 0300

This publication is also available on the Ministry for Primary Industries websites at:
http://www.mpi.govt.nz/news-and-resources/publications
http://fs.fish.govt.nz go to Document library/Research reports

# TABLE OF CONTENTS

# EXECUTIVE SUMMARY

**Webber, D.N.[1]; Dunn, A.[2]; Mormede, S.[3] (2021). Stan-ASD: a new age-structured stock assessment model, with an application to sub-Antarctic hake (*Merluccius australis*) and ling (*Genypterus blacodes*).**

***New Zealand Fisheries Assessment Report 2021/59*. 43 p.**

General purpose software packages currently available for implementing stock assessments are limited by the long software development cycle and the high cost of investment. While such general-purpose packages have provided a consistent basis for age-structured stock assessments in New Zealand, they are not easily modified to be able to evaluate new techniques, estimation methods, or alternative parameterisations. We use the Stan programming language to describe a new generalised age-structured stock assessment model, the Stan age-structured dynamics model (Stan-ASD), to provide a platform that would allow for rapid development of methods and improvements for stock assessments.

Stan is a state-of-the-art platform for statistical modelling and high-performance statistical computation and was used because it has a robust and stable Markov chain Monte Carlo (MCMC) algorithm that deals well with parameters that are poorly informed by data, and may be highly correlated or confounded. The Stan-ASD model partition includes only two dimensions: age and category. Category can refer to sex, maturation stage, area, species, or some other factor or combination of factors. The Stan-ASD model has an annual time period and events are ordered and proportionally applied in time steps within the annual time period. The model has been set up so that it can replicate and extend equivalent CASAL models in annual cycle and structure.

We have tested and validated the Stan-ASD code using two examples: the most recent Sub-Antarctic hake stock assessment and the most recent Sub-Antarctic ling (LIN 5&6) stock assessment. These two models were similar but had different model structures and observations available for model inference. A good match was achieved between the CASAL and Stan-ASD models for both the hake and ling examples.

We see the Stan-ASD model as being complimentary to other generalised stock assessment software, such as CASAL, as it allows the rapid coding and implementation new ideas/concepts and to test different aspects of age structured stock assessment models. If such alternatives or ideas are then found to be worthwhile, these new concepts can be incorporated into other generalised stock assessment software (e.g., CASAL) as standard features for a wider range of stock assessment scientists to make use of. We recommend further development of the Stan-ASD model so that it can be made more accessible to other stock assessment scientists and tested under a wider range of scenarios.

---

[1] Quantifish Ltd., Tauranga, New Zealand.
[2] Ocean Environmental Ltd., Wellington, New Zealand.
[3] SoFish Consulting Ltd., Wellington New Zealand.

# 1. INTRODUCTION

General purpose software packages currently available for implementing stock assessments are limited by the long software development cycle and the high cost of investment. While such general-purpose packages have provided a consistent basis for age-structured assessments in New Zealand, they are not easily modified to be able to evaluate new techniques, estimation methods, or alternative parameterisations. We use the Stan programming language (Stan Development Team 2021) to describe a new generalised age-structured stock assessment model, the Stan age-structured dynamics model (Stan-ASD), to provide a platform that would allow for rapid development of methods and improvements for stock assessments.

Stan is an open-source package that can be run from the command line using *CmdStan* or from within the statistical programming language *R* (R Core Team 2019) using the *RStan* package. Stan is a state-of-the-art platform for statistical modelling and high-performance statistical computation and was used as the basis for Stan-ASD because it uses a robust and stable no-U-turn sampler (NUTS) Markov chain Monte Carlo (MCMC) algorithm that deals well with parameters that are poorly informed by data and/or are highly correlated.

Stan-ASD borrows some of its ideas and code from the length-structured dynamics (LSD) model that is used for New Zealand rock lobster stock assessments (Webber et al. 2018) but is more general and has been developed for age-based assessments.

We note that the Stan-ASD model can be used as a platform for rapidly coding and evaluating new ideas or concepts in stock assessment models. If proved valuable, these can be included in other generalised stock assessment software (e.g., CASAL) as standard features for use in a wider range of stock assessments. However, Stan-ASD is not only limited to model testing, but also is capable of producing stock assessments for informing managers. Furthermore, Stan's MCMC algorithm is less sensitive to parameter correlation, is faster, and generally mixes better than other commonly used MCMC algorithms (e.g., Metropolis-Hastings).

We tested and validated the Stan-ASD code using two examples: the most recent Sub-Antarctic hake (*Merluccius australis*) stock assessment (Dunn et al. in press) and the most recent Sub-Antarctic ling (*Genypterus blacodes*) (LIN 5&6) stock assessment (Mormede et al. in press). These two models were similar but had slightly different model structures and observations available for model inference.

# 2. METHODS

The Stan-ASD model partition includes two dimensions: age and category. Category can refer to sex (e.g., males and females), maturation stage (e.g., immature and mature), area, species, or some other factor (e.g., tagging) or combination of factors. The model has an annual time period, in which events (e.g., recruitment, natural mortality, fishing mortality) are ordered and proportionally applied in time steps within the annual time period. It has been set up so that it can replicate and extend equivalent CASAL (Bull et al. 2012) models in annual cycle and structure.

The annual cycle in Stan-ASD is similar to that of CASAL and contains the following processes:

1. Ageing
2. Recruitment
3. Spawning
4. Natural and fishing mortality

If two or more processes are specified to occur within the same time step then they will occur in the above order. More processes (e.g., maturation, migration) can be incorporated in the future. We describe each of the process implemented in detail below.

In the following sections, we use $a$ to index age for $a = \{a_{\min}, \dots, A\}$, $c$ to index category, and $y$ to index year. The numbers at age, in each category, during each year are represented as $N_{a,y,c}$. We also refer to different fisheries using the subscript $f$, different indices of abundance using $i$, and different sets of proportions at age using $j$. However, we drop the category subscript ($c$) and fishery subscript ($f$) to simplify the following equations.

## 2.1 Mean weight at age

The mean weight at age ($w_a$) is defined as

$$w_a = \alpha(L_a)^{\beta} \times (1 + c^2)^{\frac{\beta(\beta-1)}{2}}$$

where $L_a$ is the mean size at age, $\alpha$ and $\beta$ are the size-weight parameters, and $c$ is the CV of size at age. The CV term corrects for bias in estimating mean weight for a lognormally distributed size at age model (see Bull et al. 2012 for further details).

## 2.2 Selectivity at age

Selectivity at age can be defined for each fishery, index of abundance, or set of proportions at age observations. Selectivity at age is a function of parameters that define the shape of a selectivity ogive (e.g., two parameters for a logistic ogive or three parameters for a double normal ogive) and can be specified with a parameter that defines the height of the selectivity ogive (i.e., the maximum proportion selected). For example, a model that includes a different selectivity ogive for males and females may specify a logistic ogive (i.e., maximum height of 1) for males but estimate the maximum height parameter for females (i.e., logistic capped). This allows the stock assessment model to define the proportions of ages selected and the proportion of males versus females that are selected.

The logistic ogive is defined as

$$S_a = 1/\left(1 + 19^{(a_{50}-a)/a_{to95}}\right)$$

with parameters $a_{50}$ and $a_{to95}$ which define the age at 50% selectivity and the age to 95% selectivity, respectively. The logistic capped ogive is

$$S_a = a_{\max}/\left(1 + 19^{(a_{50}-a)/a_{to95}}\right)$$

where $a_{\max}$ defines the maximum height of the selectivity curve.

The double normal ogive is defined as

$$S_a = \begin{cases} 2^{-((a-a_1)/a_L)^2}, & a \le a_1 \\ 2^{-((a-a_1)/a_R)^2}, & a > a_1 \end{cases}$$

with parameters $a_1$, $a_L$, and $a_R$ which define the age at maximum selectivity, the steepness of the left-hand limb, and the steepness of the right-hand limb, respectively. The double normal capped ogive is

$$S_a = \begin{cases} a_{\max} \times 2^{-((a-a_1)/a_L)^2}, & a \le a_1 \\ a_{\max} \times 2^{-((a-a_1)/a_R)^2}, & a > a_1 \end{cases}$$

where $a_{\max}$ defines the maximum height of the selectivity curve. Selectivities can be defined for different fisheries ($f$), different indices of abundance ($i$), and different sets of proportions at age ($j$).

## 2.3 Spawning stock biomass

The spawning stock biomass (SSB) is the biomass (tonnes) of mature individuals each year and was defined as

$$SSB_y = \sum_a N_{a,y} m_a w_a$$

where $N_{a,y}$ is the number of individuals at age in the model each year and $m_a$ is the proportion mature at age. The proportion mature is defined using a logistic curve as

$$m_a = 1/\left(1 + 19^{(m_{50}-a)/m_{to95}}\right)$$

where $m_{50}$ and $m_{to95}$ represent the age at 50% maturity and the age to 95% maturity, respectively. The unfished equilibrium SSB is called *B₀* and is defined as

$$B_0 = SSB_y \quad \text{for} \quad y = 1.$$

## 2.4 Vulnerable biomass

The vulnerable biomass (tonnes) for each year has been defined as

$$V_y = \sum_a N_{a,y} w_a S_a$$

where $N_{a,y}$ is the numbers at age each year, $w_a$ is the mean weight at age, and $S_a$ is the selectivity at age. The vulnerable biomass can be calculated for a specific fishery ($f$) or abundance index ($i$).

## 2.5 Ageing

To age individuals, all of the individuals in the partition are moved one cell to the right

$$N_{a,y} = N_{a-1,y-1}$$

and all of the oldest individuals are absorbed into a final age group ($A$) (the plus group, i.e., the accumulative number of all individuals of age $A$ and older)

$$N_{A,y} = N_{A-1,y-1} + N_{A,y-1}.$$

We refer to the ages being modelled as ages $a_{\min}$-$A$+ (e.g., 1-30+).

## 2.6 Recruitment

Recruitment each year has been defined as

$$R_y = R_0 \times YCS_{y-e} \times SR\left(SSB_{y-e}\right)$$

where $R_0$ is the average recruitment (see Section 2.8 for its derivation), $YCS_y$ is the year class strength for each year, and $SR\left(SSB_{y-e}\right)$ is the stock recruitment multiplier which is a function of the spawning stock biomass in some years prior $e$ (e.g., Bull et al. 2012). The Beverton & Holt (1957) stock recruitment relationship, parameterised with the steepness parameter ($h$) is used

$$SR(SSB_y) = \frac{SSB_y}{B_0} \Bigg/ \left(1 - \frac{5h-1}{4h}\left(1 - \frac{SSB_y}{B_0}\right)\right)$$

where $h = SR(0.2B_0)$.

To ensure that $R_0$ has the correct meaning, the YCSs are standardised (i.e., transformed by dividing the YCS vector by the mean of that vector) to have mean 1, and a penalty is applied to reduce the impact of this standardisation. The penalty applied is

$$\left(\frac{1}{n_S}\sum_{y\in S} YCS_y - 1\right)^2 \times \vartheta$$

where, $S$ is the set of years for which YCSs are estimated, $n_S$ is the number of years for which YCSs are estimated, and $\vartheta$ is the YCS penalty multiplier.

## 2.7 Natural and fishing mortality

Both natural and fishing mortality for each fishery ($f$) are removed from the partition within a time step using

$$N'_{a,y} = N_{a,y}e^{-\tau M}\left(1 - \sum_f S_{a,f}U_{y,f}\right)$$

where $M$ is the natural mortality, $\tau$ is the proportion of natural mortality to apply within a time step, $S_{a,f}$ is the selectivity at age for each fishery, and $U_{y,f}$ is the exploitation rate each year for each fishery which was defined as

$$U_{y,f} = \frac{C_{y,f}}{V_{y,f}}$$

where $C_{y,f}$ is the catch (tonnes) each year associated with each fishery.

## 2.8 Unfished numbers at age

The unfished numbers at age are found by iterating over ageing, recruitment (specifying $R_0 = 1$), and natural mortality until the equilibrium distribution of numbers at age is found. During the final iteration, the average of the numbers at age before and after natural mortality are calculated as

$$N_a^0 = N_a^1(1-\lambda) + \lambda N_a^2$$

where $N_a^1$ is the numbers at age before applying natural mortality, $N_a^2$ is the numbers at age after applying natural mortality, and the spawning mortality proportion is $\lambda = 0.5$. This is then used to derive an equilibrium SSB

$$SSB_0 = \sum_a N_a^0 m_a w_a.$$

Finally, $R_0$ is solved using

$$R_0 = \frac{B_o}{SSB_0}.$$

## 2.9 Observations and likelihoods

### 2.9.1 Abundance indices

Each index of abundance ($i$) is assumed to be lognormally distributed and defined as

$$\log(I_{y,i}) \sim N(q_i V_{y,i}, \sigma^2_{y,i})$$

where $q_i$ is the catchability coefficient for each index of abundance, $V_{y,i}$ is the vulnerable biomass each year for each index of abundance, and the variance is defined as

$$\sigma^2_{y,i} = \sqrt{\log(1 + \varsigma^2_{y,i})} \times \frac{1}{\omega_i}$$

and

$$\varsigma_{y,i} = \sqrt{c^2_{y,i} + \pi^2_i}$$

where $c_{y,i}$ is the survey coefficient of variation (CV) each year for each survey, $\pi_i$ is the additional process error CV for each survey, and $\omega_i$ is the likelihood weighting for each survey.

### 2.9.2 Proportions at age

We include ageing error in proportions at age observations to be modelled explicitly. After the expected values are calculated for proportions at age observations, misclassification rates are applied to them, which has the effect of 'smearing' the proportions at age. The resulting 'smeared' proportions at age are used in calculating the objective function. The ageing error model used assumes that individuals of age $a$ are classified as ages which are normally distributed with mean $a$ and constant CV. Note that if the CV is high enough, some fish of the younger age classes may disappear.

Proportions at age observations ($j$) are assumed to be multinomially distributed

$$(P_a)_j \sim M(\eta_j, (p_a)_j)$$

where $\eta_j$ is the effective sample size for each set of proportions at age observations, and $(p_a)_j$ is the predicted set of proportions at age in the model. Proportions at age observations can be fitted either jointly across different categories (i.e., the proportions at age for males and females are concatenated and sum to 1) or separately (i.e., the proportions at age for each sex are fitted independently).

## 2.10 Generated quantities

Generated quantities are values derived after model fitting in Stan and can include quantities of interest such as additional model diagnostics or outputs useful to managers.

The SSB in the final model year is referred to as 'B current' (e.g., $B_{2021}$) and is often related to the unfished SSB (e.g., $B_{2021}/B_0$).

The only other generated quantity coded in Stan-ASD (so far) is the posterior predictive distribution for the indices of abundance. The posterior predictive distribution is a useful diagnostic to explore how well the variability estimated within a model agrees with the variability observed within the data. In Stan-ASD we simulated the posterior predictive distribution for each survey during year $y$. The posterior predictive distribution for each survey was simulated as

$$\log\left(I_{y,i}^{PP}\right) \sim N\left(q_i V_{y,i}, \sigma_{y,i}^2\right)$$

where the inputs are described above in Section 2.9.1.

## 2.11 MCMC diagnostics

Standard diagnostic plots of model parameters (e.g., trace and autocorrelation plots) are trivial to produce from Stan-ASD model runs using several purpose-built *R* packages (e.g., *RStan* and *bayesplot*) as such packages are built to work with general Stan output. Additionally, several diagnostic statistics are generated from Stan-ASD automatically during a model run including:

- the bulk effective sample size (bulk-ESS);
- the tail effective sample size (tail-ESS);
- the potential scale reduction statistic ($\hat{R}$ or R-hat); and
- the effective sample size for each parameter (effective N).

When a Stan-ASD MCMC has finished, the user will be warned if either the bulk-ESS or the tail-ESS is low, giving an indication that the MCMC may not be adequate. The bulk-ESS is a useful measure of sampling efficiency in the bulk of the distribution (and is related to the efficiency of the mean and median estimates) and is well defined even if the chains do not have finite mean or variance. The tail-ESS is estimated by computing the minimum of effective sample sizes for 5% and 95% quantiles. The tail-ESS is a useful measure for sampling efficiency in the tails of the distribution (and is related to efficiency of variance and tail quantile estimates). Both the bulk-ESS and the tail-ESS should be at least 100 (approximately) per Markov chain to be reliable and indicate that estimates of respective posterior quantiles are reliable (Vehtari 2019).

An alternative way to monitor whether a chain has converged to the equilibrium distribution is to compare its behaviour with other randomly initialised chains. This is the motivation for the potential scale reduction statistic ($\hat{R}$). The $\hat{R}$ statistic measures the ratio of the average variance of draws within each chain to the variance of the pooled draws across chains; if all chains are at equilibrium, these will be the same and $\hat{R}$ will be one. If the chains have not converged to a common distribution, the $\hat{R}$ statistic will be greater than one (see Gelman et al. 2013, Vehtari 2019, Stan Development Team 2021). The Stan developers recommend running at least four chains by default and only using the samples if $\hat{R}$ is less than 1.05.

Another technical difficulty posed by MCMC methods is that the samples are typically autocorrelated (or anticorrelated) within a chain. This increases the uncertainty of the estimation of posterior quantities of interest, such as means, variances, or quantiles (see Geyer 1992, 2011). To help address this, Stan-ASD estimates an effective sample size for each parameter, which plays the role in the MCMC central limit theorem (CLT) as the number of independent draws plays in the standard CLT. Unlike most packages, the particular calculations used by Stan follow those for $\hat{R}$, which involve both cross-chain (mean) and within-chain calculations (autocorrelation) (see Gelman et al. 2013).

# 3. VALIDATION

To test and validate the code, we replicated two models implemented in CASAL. We tested each model using three different Stan-ASD model runs, by:

1. using the median of the posterior density (MPD) values from base case CASAL assessment in a fixed parameter run in Stan-ASD, and comparing the resulting model fits and likelihoods;

2. finding the maximum *a posteriori* (MAP[4]) of the parameter values for the Stan-ASD model, and comparing this with the CASAL MPD;
3. sampling from the posterior distribution of the Stan-ASD model using the Stan MCMC algorithm and comparing this with the posterior distribution obtained using the CASAL MCMC algorithm.

In addition, we compared several aspects each model run including the likelihood components, prior contributions, and penalty values; and other key output quantities such as initial and final numbers at age, annual recruitments, spawning stock biomass (SSB) trajectories, expected values of survey biomass indices, and expected values of proportions at age observations.

## 3.1 Hake

The base case Sub-Antarctic hake stock assessment model was described by Dunn et al. (in press). The model partition was split by sex and included ages 1–30+, for years from 1975 to 2021. The annual cycle for the model consisted of three time steps:

1. Step 1
    a. Ageing
2. Step 2
    a. Spawning
    b. Recruitment
    c. November survey
    d. Summer fishery
3. Step 3
    a. April survey
    b. Winter fishery

The MCMC consisted of four chains, each run for 4000 iterations. The first 2000 samples were dropped during the warm-up period, and we then retained every second sample, resulting in a total of 4000 samples from the posterior distribution.

All likelihood components and most of the prior densities matched adequately when taking the MPD values from base case CASAL assessment and doing a fixed parameter run within Stan-ASD (Table 1). The only notable differences were the YCS prior density and penalty. These differed because CASAL evaluates the YCS prior density and penalty for those years that are fixed to one, whereas Stan-ASD only evaluates the YCS prior density and penalty for the set of years that YCS is estimated. Evaluating the YCS prior density and penalty for the Stan-ASD model outside the model resulted in identical values.

Comparisons of the initial and final numbers at age show that the Stan-ASD model matched the CASAL model very closely (Figure 1, Figure 2). Recruitment and SSB comparisons also indicated an excellent match, with a few minor differences in the posterior distribution of the early recruitment and SSB during the 1980s and 1990s (Figure 3, Figure 4). The fit to survey biomass also matched well and the posterior predictive distribution aligned well with the observations suggesting that the combination of the survey CVs and the additional process error described the uncertainty adequately (Figure 5). Finally, the age compositions aligned well when comparing the CASAL MPD with the Stan-ASD MAP (Figure 6, Figure 7, Figure 8); however, the posterior distribution from the Stan-ASD model were slightly different for the April survey (Figure 7).

Standard diagnostic plots of key model parameters suggested that the MCMC chains were mixing adequately with the exception of two right-hand limb selectivity parameters (Figure 9). These were improvements on the CASAL model trace diagnostics (see Figure 23 in Appendix I). Autocorrelation

---

[4] Bayesians usually refer to the penalised maximum likelihood as the maximum *a posteriori* (MAP). However, the term median of the posterior density (MPD) is commonly used in fisheries science. We use MAP throughout this document when referring to model fits when using Stan-ASD and MPD when referring to model fits done using CASAL.

was acceptable for most parameters with the exception of $B_0$, natural mortality ($M$), and a two of the selectivity parameters (Figure 10). The bulk-ESS and the tail-ESS were both good suggesting that posterior quantiles were reliable. The $\hat{R}$ statistic was at or above 1.005 for five model parameters ($B_0$, natural mortality, and a few of the selectivity parameters) indicating that the MCMC may not have converged to the equilibrium distribution (Table 2). The effective N was low for the same parameters (Table 2).

The differences between the CASALs MPD and the Stan-ASD MAP were very small with only minor differences in parameter estimates and derived quantities (Table 3). The same was true of the posterior distributions from the CASAL and the Stan-ASD models (Table 3).

**Table 1: The prior, likelihood, and penalty components from the base case CASAL model and the Stan-ASD fixed parameter model run for hake. The largest differences are indicated in red.**

| Component | Component type | CASAL MPD | Stan-ASD fixed parameter | Difference |
|---|---|---|---|---|
| April survey | Likelihood | -4.614 | -4.625 | 0.011 |
| November survey | Likelihood | -17.904 | -17.973 | 0.069 |
| April age | Likelihood | 62.194 | 62.190 | 0.003 |
| November age | Likelihood | 356.717 | 356.723 | -0.006 |
| Summer fishery age | Likelihood | 537.758 | 537.857 | -0.099 |
| $q$ (November survey) | Prior | -1.489 | -1.489 | 0.000 |
| $q$ (April survey) | Prior | -2.248 | -2.248 | 0.000 |
| $B_0$ | Prior | 11.006 | 11.006 | 0.000 |
| $M$ | Prior | 0.021 | 0.021 | 0.000 |
| YCS | Prior | -14.951 | -15.347 | 0.396 |
| YCS penalty | Penalty | 6.944 | 8.296 | -1.352 |
| Total | | 933.434 | 934.412 | -0.978 |

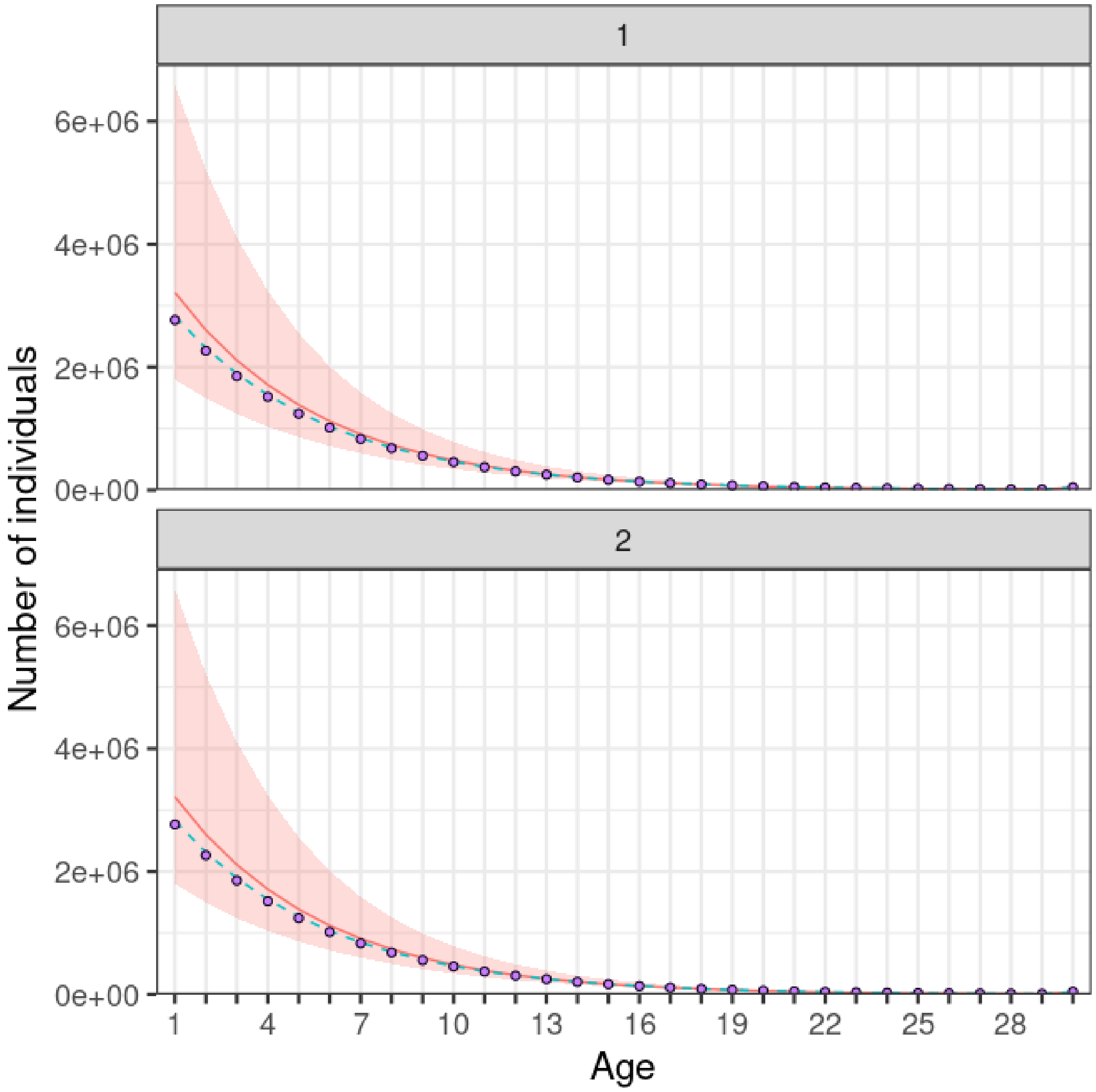

**Figure 1:** Initial numbers at age by sex (1=males, 2=females) for the CASAL models MPD (purple points), the CASAL MPD parameter values pushed through the Stan-ASD model as a fixed parameter run (open black points), the Stan-ASD MAP (dashed blue line), and the Stan-ASD posterior distribution (median is red line and red shading is the 90% credible interval).

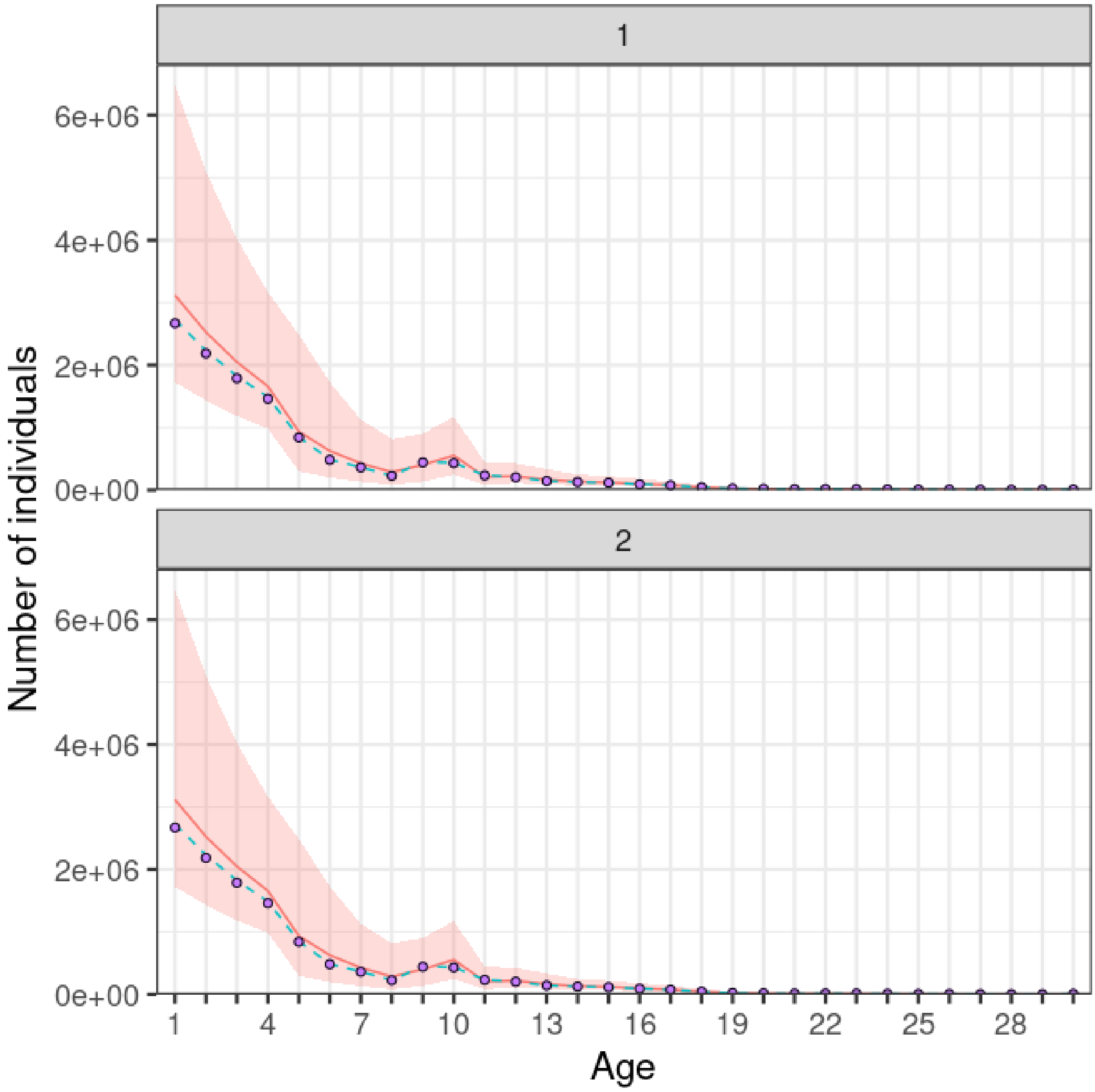

**Figure 2:** **Final numbers at age by sex (1=males, 2=females) for the CASAL models MPD (purple points), the CASAL MPD parameter values pushed through the Stan-ASD model as a fixed parameter run (open black points), the Stan-ASD MAP (dashed blue line), and the Stan-ASD posterior distribution (median is red line and red shading is the 90% credible interval).**

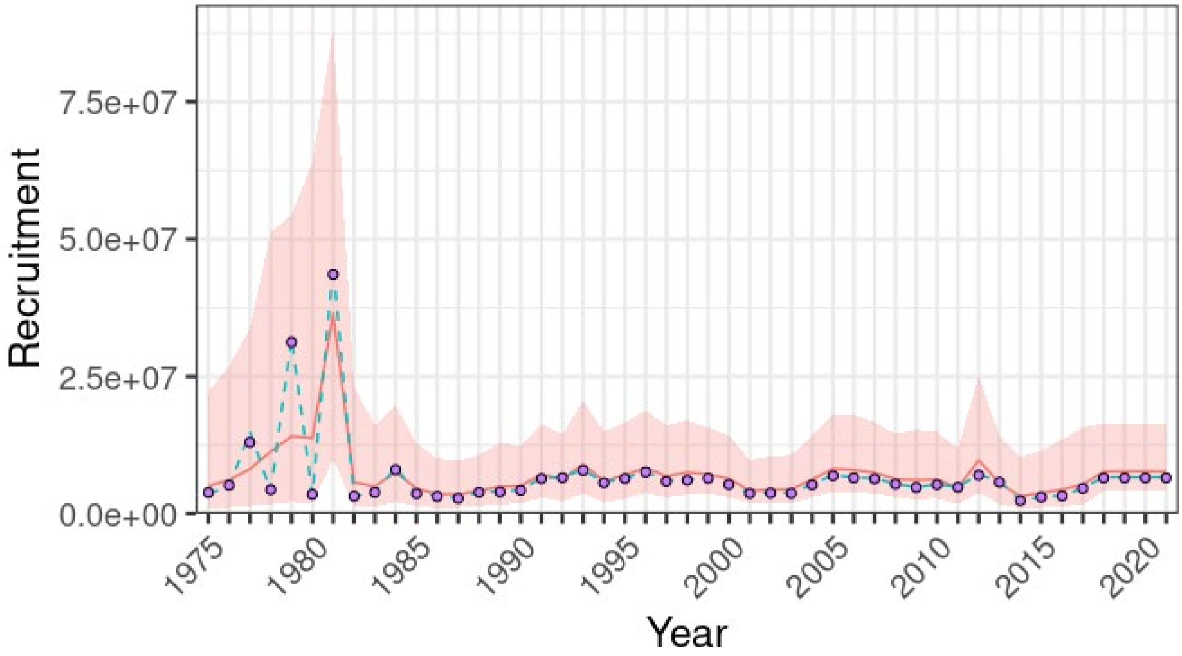

Figure 3: Recruitment by year for the CASAL models MPD (purple points), the CASAL MPD parameter values pushed through the Stan-ASD model as a fixed parameter run (open black points), the Stan-ASD MAP (dashed blue line), and the Stan-ASD posterior distribution (median is red line and red shading is the 90% credible interval).

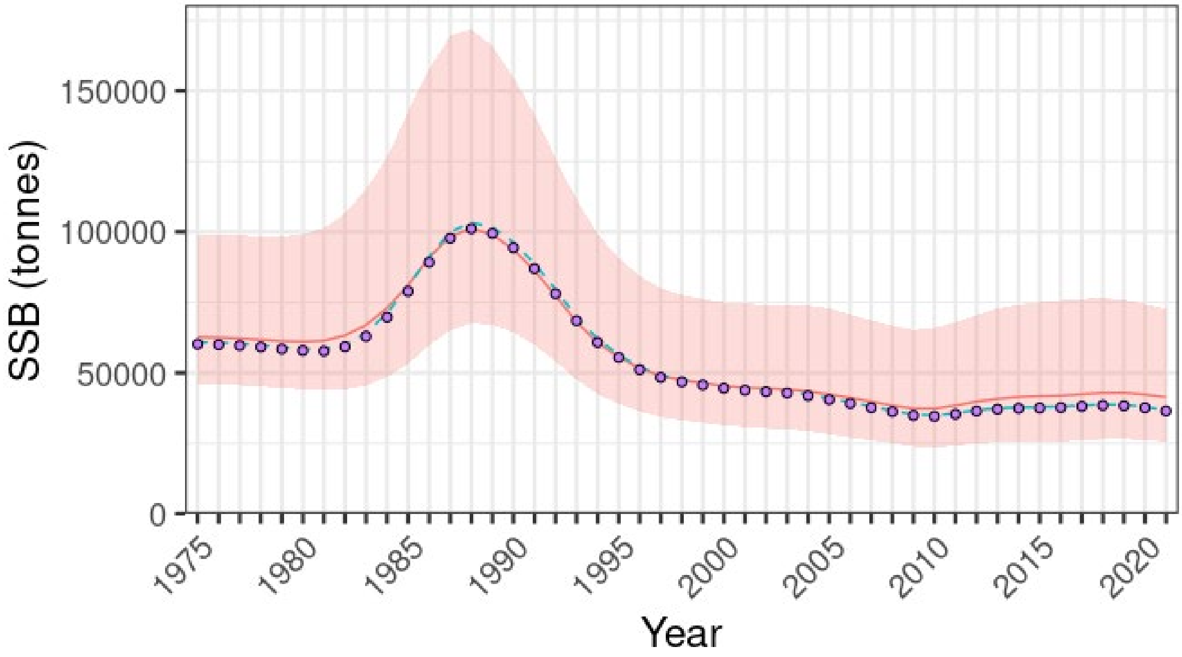

Figure 4: Spawning stock biomass (SSB) by year for the CASAL models MPD (purple points), the CASAL MPD parameter values pushed through the Stan-ASD model as a fixed parameter run (open black points), the Stan-ASD MAP (dashed blue line), and the Stan-ASD posterior distribution (median is red line and red shading is the 90% credible interval).

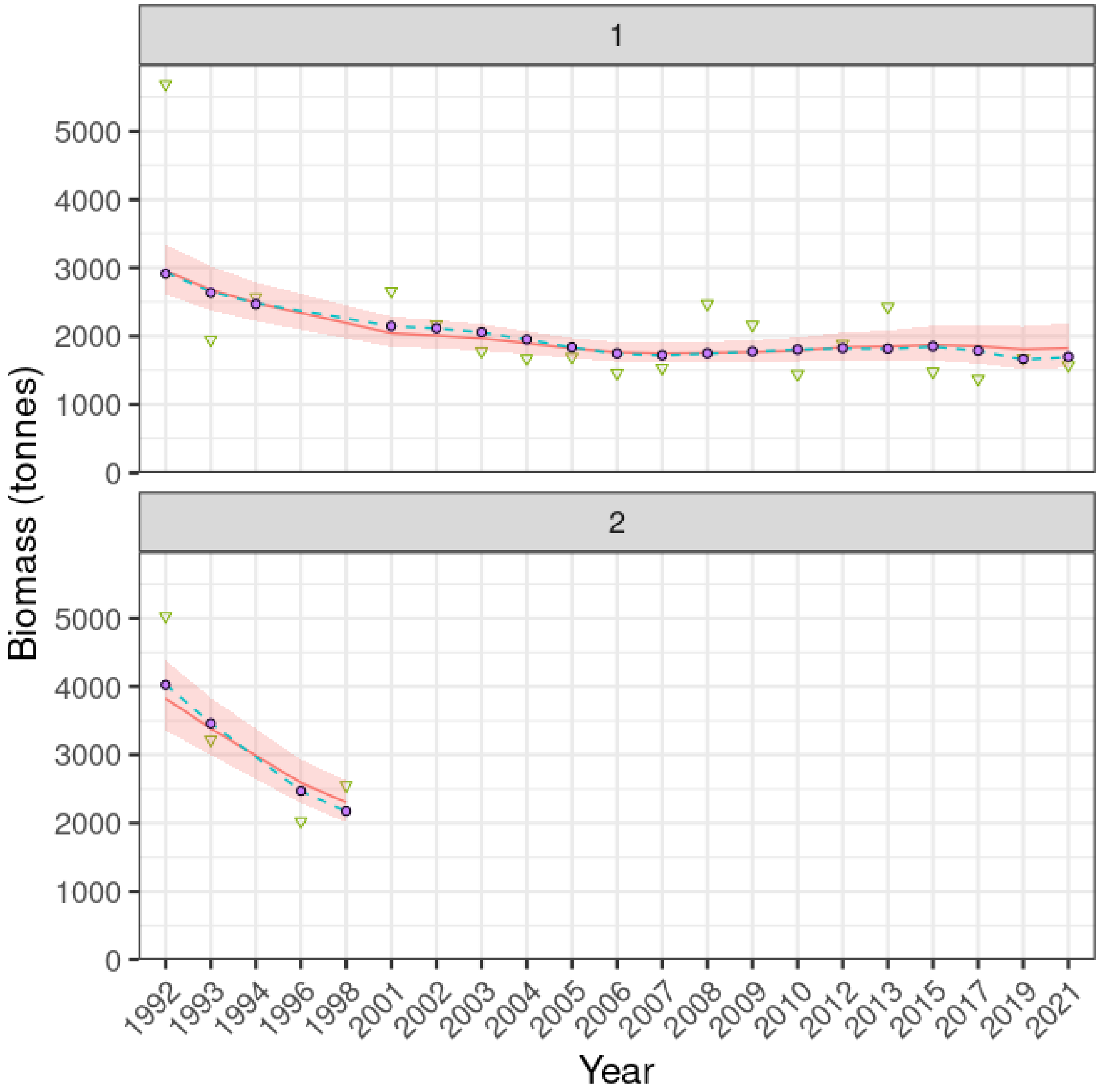

**Figure 5:** Survey biomass by year showing the observations (green triangles), the exploratory CASAL models MPD (purple points), the MPD values pushed through the Stan-ASD model (open black points), the Stan-ASD MAP (dashed blue line), the Stan-ASD posterior median and 90% credible interval (red line and inner red shaded region), and the Stan-ASD posterior predictive distribution (outer shaded region).

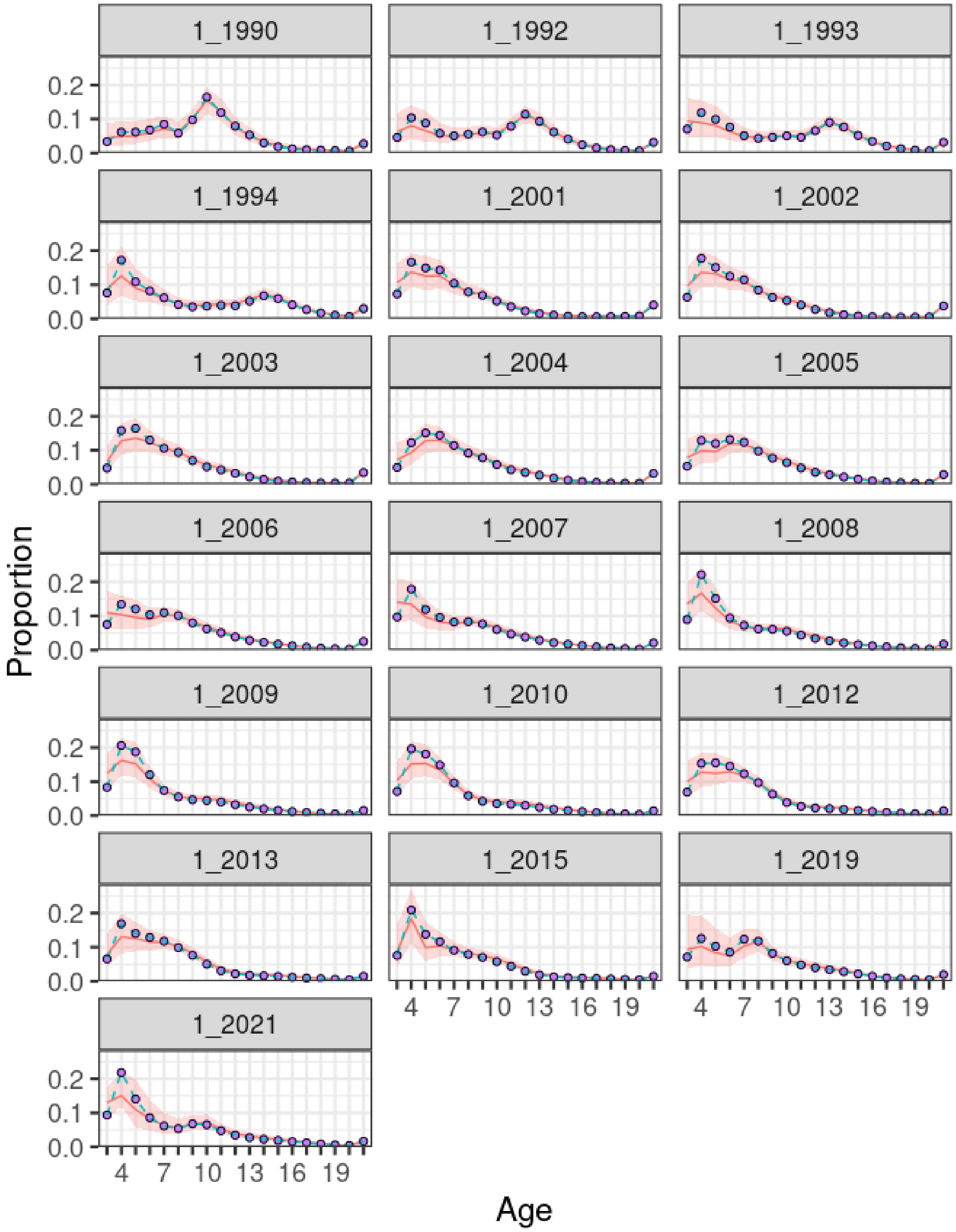

**Figure 6:** November survey age composition year for the CASAL models MPD (purple points), the MPD values pushed through the Stan-ASD model (open black points), the Stan-ASD MAP (dashed blue line), and the Stan-ASD posterior distribution (median is red line and red shading is the 90% credible interval).

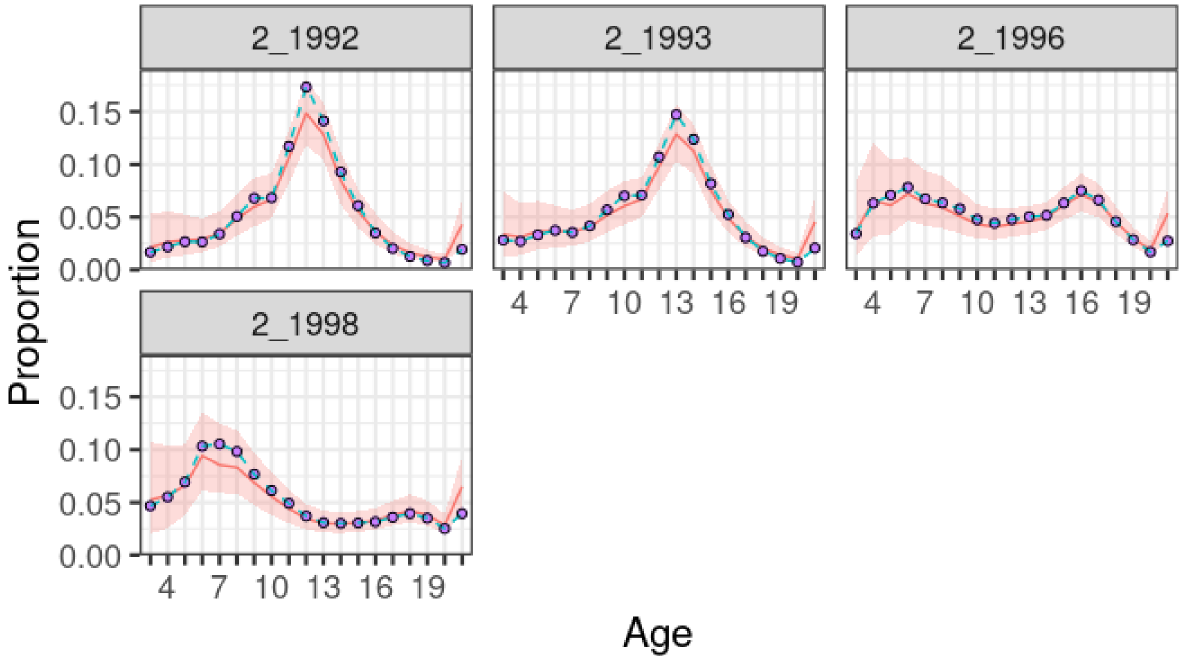

**Figure 7:** April survey age composition by year for the CASAL models MAP (purple points), the MPD values pushed through the Stan-ASD model (open black points), and the Stan-ASD posterior distribution (median is red line and red shading is the 90% credible interval).

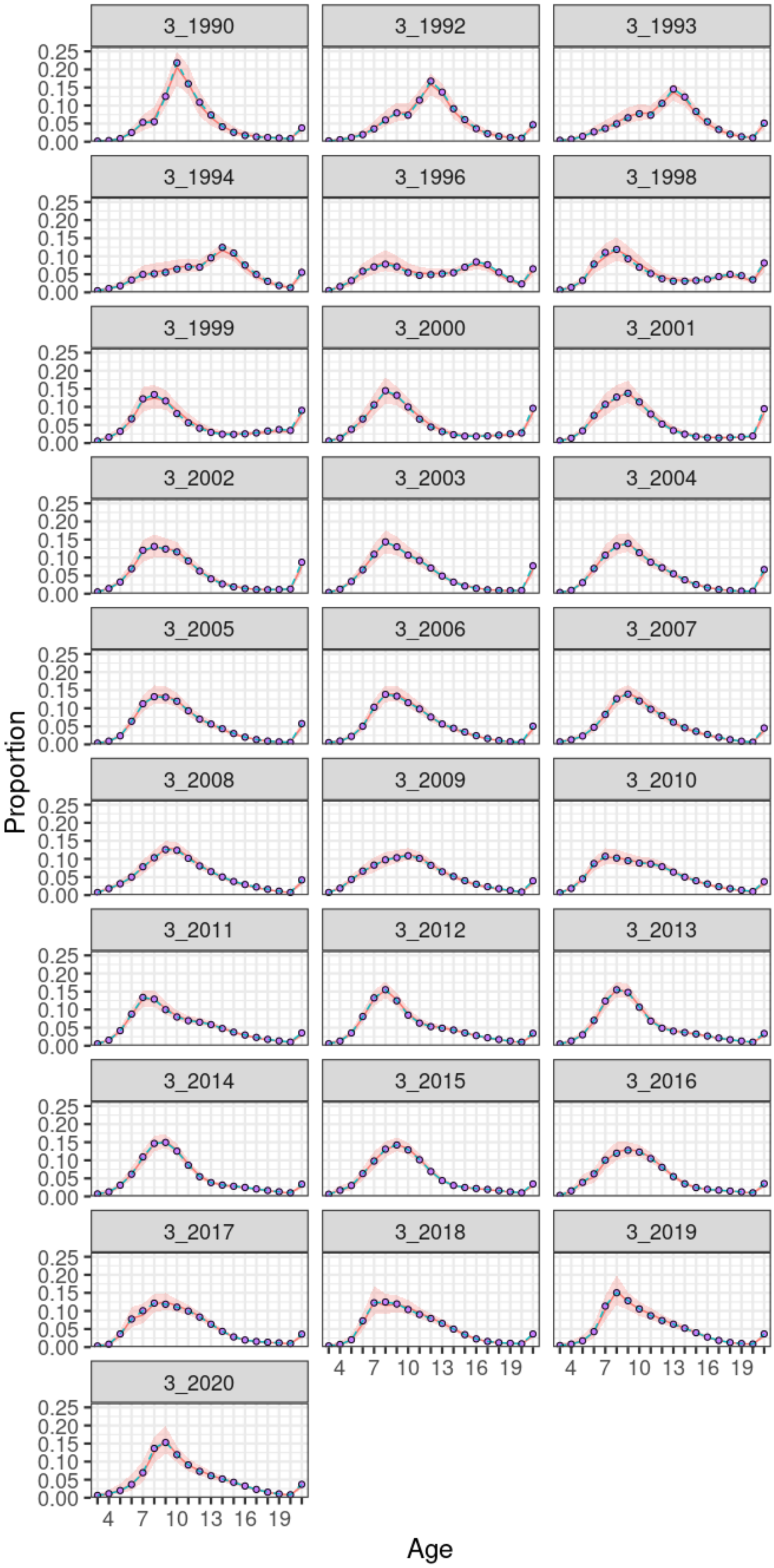

**Figure 8:** **Longline age composition by year for the CASAL models MAP (coloured points), the MPD values pushed through the Stan-ASD model (open black points), and the Stan-ASD posterior distribution (median is red line and red shading is the 90% credible interval).**

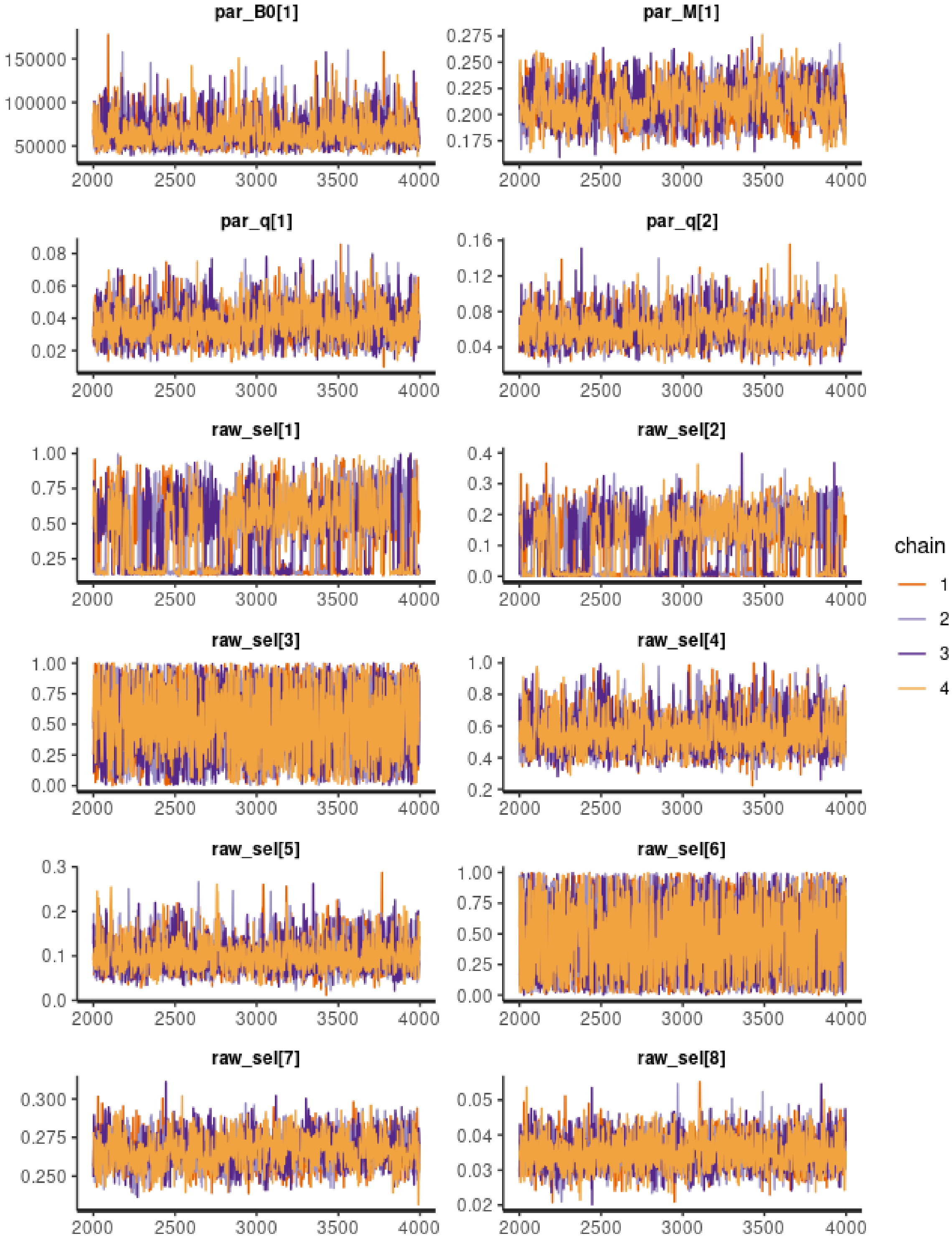

**Figure 9:** **Markov chain Monte Carlo (MCMC) trace plots for key model parameters by chain in the Stan-ASD hake model.**

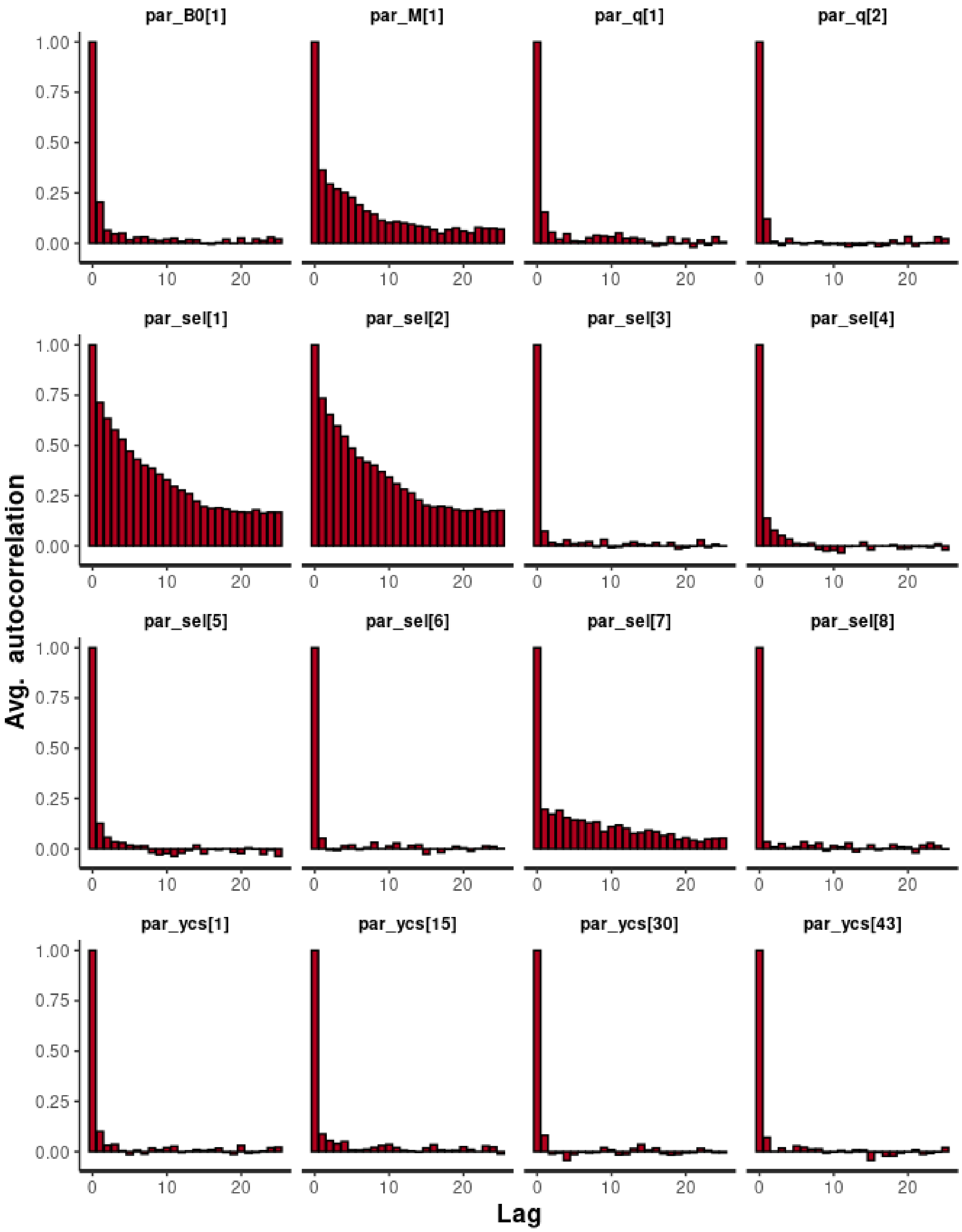

**Figure 10:** Markov chain Monte Carlo (MCMC) autocorrelation for key model parameters for lags of 0–25 in the Stan-ASD hake model.

**Table 2: Maximum *a posteriori* (MAP) and Markov chain Monte Carlo (MCMC) statistics for the Stan-ASD model runs for hake. Failed $\hat{R}$ statistics are indicated in red.**

| | MAP | | | | | | | | | | MCMC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Parameter | | Mean | SE mean | SD | 2.5% | 25% | 50% | 75% | 97.5% | Effective N | $\hat{R}$ |
| $B_{2021}$ | 36 718 | 44 054 | 408 | 14 702 | 23 800 | 33 742 | 41 366 | 51 335 | 80 449 | 1 296 | 1.001 |
| $B_0$ | 60 919 | 66 128 | 384 | 16 812 | 44 366 | 54 187 | 62 692 | 73 944 | 108 871 | 1 917 | 1.000 |
| $B_{2021} / B_0$ | 0.603 | 0.656 | 0.002 | 0.077 | 0.513 | 0.603 | 0.654 | 0.707 | 0.815 | 962 | 1.003 |
| $M$ | 0.201 | 0.211 | 0.001 | 0.018 | 0.178 | 0.197 | 0.210 | 0.223 | 0.248 | 422 | 1.006 |
| $q$ (November survey) | 0.031 | 0.036 | 0.000 | 0.010 | 0.019 | 0.029 | 0.035 | 0.042 | 0.060 | 1 948 | 1.005 |
| $q$ (April survey) | 0.058 | 0.061 | 0.000 | 0.017 | 0.033 | 0.048 | 0.060 | 0.071 | 0.100 | 3 221 | 1.001 |
| sel[1] (November survey) | 4.375 | 11.002 | 0.455 | 5.828 | 4.323 | 4.697 | 11.937 | 15.678 | 21.951 | 164 | 1.027 |
| sel[2] (November survey) | 1.000 | 6.140 | 0.362 | 4.413 | 1.023 | 1.282 | 6.997 | 9.777 | 13.941 | 149 | 1.028 |
| sel[3] (November survey) | 199.994 | 102.402 | 1.074 | 56.323 | 8.173 | 54.246 | 101.714 | 149.915 | 195.456 | 2 752 | 1.002 |
| sel[4] (April survey) | 12.277 | 14.229 | 0.058 | 2.843 | 9.685 | 12.318 | 13.819 | 15.677 | 21.159 | 2 401 | 1.000 |
| sel[5] (April survey) | 4.704 | 5.876 | 0.035 | 1.764 | 3.216 | 4.643 | 5.568 | 6.792 | 10.182 | 2 576 | 1.000 |
| sel[6] (April survey) | 10.115 | 93.020 | 0.986 | 59.260 | 6.348 | 38.611 | 90.985 | 144.332 | 194.647 | 3 615 | 1.000 |
| sel[7] (fishery) | 7.206 | 7.390 | 0.011 | 0.250 | 6.933 | 7.220 | 7.378 | 7.555 | 7.905 | 540 | 1.008 |
| sel[8] (fishery) | 2.635 | 2.724 | 0.004 | 0.226 | 2.314 | 2.567 | 2.718 | 2.871 | 3.174 | 3 059 | 1.001 |
| Objective function | -934 | -998 | 0.114 | 5.506 | -1 009 | -1 001 | -997 | -994 | -988 | 2 318 | 1.001 |

**Table 3: Maximum *a posteriori* (MAP) and Markov chain Monte Carlo (MCMC) statistics for CASAL and Stan-ASD model runs for hake.**

| | MAP | | CASAL MCMC | | | | Stan-ASD MCMC | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Parameter | CASAL | Stan-ASD | Mean | 2.5% | 50% | 97.5% | Mean | 2.5% | 50% | 97.5% |
| $B_{2021}$ | 36 373 | 36 718 | 38 445 | 22 246 | 36 490 | 65 509 | 44 054 | 23 800 | 41 366 | 80 449 |
| $B_0$ | 60 224 | 60 919 | 61 690 | 43 281 | 59 212 | 95 602 | 66 128 | 44 366 | 62 692 | 108 871 |
| $B_{2021} / B_0$ | 0.604 | 0.603 | 0.638 | 0.332 | 0.616 | 1.130 | 0.656 | 0.513 | 0.654 | 0.815 |
| $M$ | 0.200 | 0.201 | 0.199 | 0.173 | 0.198 | 0.225 | 0.211 | 0.178 | 0.21 | 0.248 |
| $q$ (November survey) | 0.032 | 0.031 | 0.033 | 0.018 | 0.033 | 0.051 | 0.036 | 0.019 | 0.035 | 0.06 |
| $q$ (April survey) | 0.060 | 0.058 | 0.061 | 0.033 | 0.059 | 0.098 | 0.061 | 0.033 | 0.06 | 0.1 |
| sel[1] (November survey) | 4.374 | 4.375 | 4.704 | 4.265 | 4.611 | 5.724 | 11.002 | 4.323 | 11.937 | 21.951 |
| sel[2] (November survey) | 1.000 | 1.000 | 1.297 | 1.007 | 1.201 | 2.133 | 6.14 | 1.023 | 6.997 | 13.941 |
| sel[3] (November survey) | 199.996 | 199.994 | 112.946 | 34.188 | 111.268 | 194.332 | 102.402 | 8.173 | 101.714 | 195.456 |
| sel[4] (April survey) | 12.300 | 12.277 | 13.854 | 9.463 | 13.415 | 20.524 | 14.229 | 9.685 | 13.819 | 21.159 |
| sel[5] (April survey) | 4.713 | 4.704 | 5.865 | 3.225 | 5.610 | 10.116 | 5.876 | 3.216 | 5.568 | 10.182 |
| sel[6] (April survey) | 10.102 | 10.115 | 93.299 | 6.667 | 90.689 | 195.222 | 93.02 | 6.348 | 90.985 | 194.647 |
| sel[7] (fishery) | 7.186 | 7.206 | 7.221 | 6.840 | 7.211 | 7.642 | 7.39 | 6.933 | 7.378 | 7.905 |
| sel[8] (fishery) | 2.638 | 2.635 | 2.677 | 2.260 | 2.667 | 3.152 | 2.724 | 2.314 | 2.718 | 3.174 |
| Objective function | -915.014 | -934.189 | -947.062 | -958.787 | -946.702 | -936.690 | -998 | -1 009 | -997 | -988 |

## 3.2 Ling

The base case Sub-Antarctic (LIN 5&6) ling stock assessment model was described by Mormede et al. (in press). The model partition was split by sex and included ages 3–25+, for years from 1972 to 2021. The annual cycle for the ling model consisted of two time steps:

1. Step 1
    a. Ageing
    b. Spawning
    c. Fishing (mid-step)
    d. Summer survey
2. Step 2
    a. Recruitment
    b. Autumn survey

In the base case ling model, the catchability coefficients (*q*) were treated as nuisance parameters. This was done because the longline CPUE abundance indices were small numbers (approximately 5 on average), requiring a very small *q* that is close to the lower bound of zero. CASAL applies an unseen penalty to push model parameters away from their bounds and in this case the penalty did not allow a *q* value small enough to fit to the CPUE series properly (i.e., *q* was forced to be higher than it should have been) when *q* was estimated as a free parameter. This resulted in the CASAL model with free *q* not fitting to the CPUE observations (Figure 11).

We did not code nuisance *q* as an option in Stan-ASD because it does not apply a penalty to push parameter values away from bounds. Furthermore, the nuisance *q* concept is undesirable in a Bayesian context, and the Stan MCMC algorithm is capable of sampling from the posterior distribution when *q* is estimated. Instead, we altered the ling base case model by multiplying the longline CPUE series by 1000 so that its magnitude was similar to that of the trawl survey series (Figure 12) and estimated *q* as a parameter.

Additionally, the right-hand limbs of the male and female trawl fishery selectivity ogives were fixed at 100 in CASAL but were not fixed in Stan-ASD because it is not yet possible to fix some of the parameters of a selectivity and estimate the others. Instead, these parameters were bounded to be between 99.99 and 100.01 which should result in minimal differences within the model. The ability to fix some selectivity parameters could be coded in future implementations.

Two different CASAL runs were done:

1. a model run that estimated all catchability coefficients; and
2. a model run that estimated all catchability coefficients but with the rescaled CPUE longline abundance index.

A total of five different model runs were done in Stan-ASD including:

1. an exploratory fixed parameter model run;
2. an exploratory MAP;
3. the final fixed parameter run;
4. the final MAP; and
5. the final MCMC.

The MAP model runs estimated the CV process error parameter, whereas this value was fixed (to the CASAL MPD) during the MCMC. The MCMC consisted of four chains, each run for 4000 iterations. The first 2000 samples were dropped during the warm-up period, and we then retained every second sample, resulting in a total of 4000 samples from the posterior distribution.

All likelihood components and most of the prior densities matched adequately when taking the MPD values from base case CASAL assessment and doing a fixed parameter run within Stan-ASD (Table 4). The only notable differences were in the YCS prior density and penalty. These differed because CASAL evaluates the YCS prior density and penalty for those years that are fixed to one, whereas Stan-ASD only evaluates the YCS prior density and penalty for the set of years that YCSs were estimated. Evaluating the YCS prior density and penalty for the Stan-ASD model outside the model resulted in identical values.

Comparisons of the initial and final numbers at age show that the Stan-ASD model matched the CASAL model well (Figure 13, Figure 14). Recruitment and SSB comparisons also indicated an excellent match (Figure 15, Figure 16). The fit to CPUE and survey biomass also matched well, and the posterior predictive distribution aligned well with the observations suggesting that the combination of the CPUE/survey CVs and the additional process error described the uncertainty adequately (Figure 12). Finally, the age compositions aligned well when comparing the CASAL MPD with the Stan-ASD MAP (Figure 17, Figure 18, Figure 19, Figure 20); however, the posterior distribution from the Stan-ASD model was slightly different for the autumn survey (Figure 18).

Standard diagnostic plots of key model parameters suggested that the MCMC chains were mixing adequately (Figure 21) and that autocorrelation was acceptable (Figure 22). These were improvements on the equivalent CASAL model trace diagnostics (see Figure 24 in Appendix I). The bulk-ESS and the tail-ESS were both good suggesting that posterior quantiles were reliable. The $\hat{R}$ statistic was less than 1.005 for all model parameters, except one left hand-limb selectivity, indicating that the MCMC was likely to have converged to the equilibrium distribution (Table 5), and the effective N was acceptable for all but one parameter (Table 5).

The differences between the CASAL MPD and the Stan-ASD MAP were very small with only minor differences in parameter estimates and derived quantities (Table 6). The same was true of the posterior distributions from the CASAL and the Stan-ASD models (Table 6).

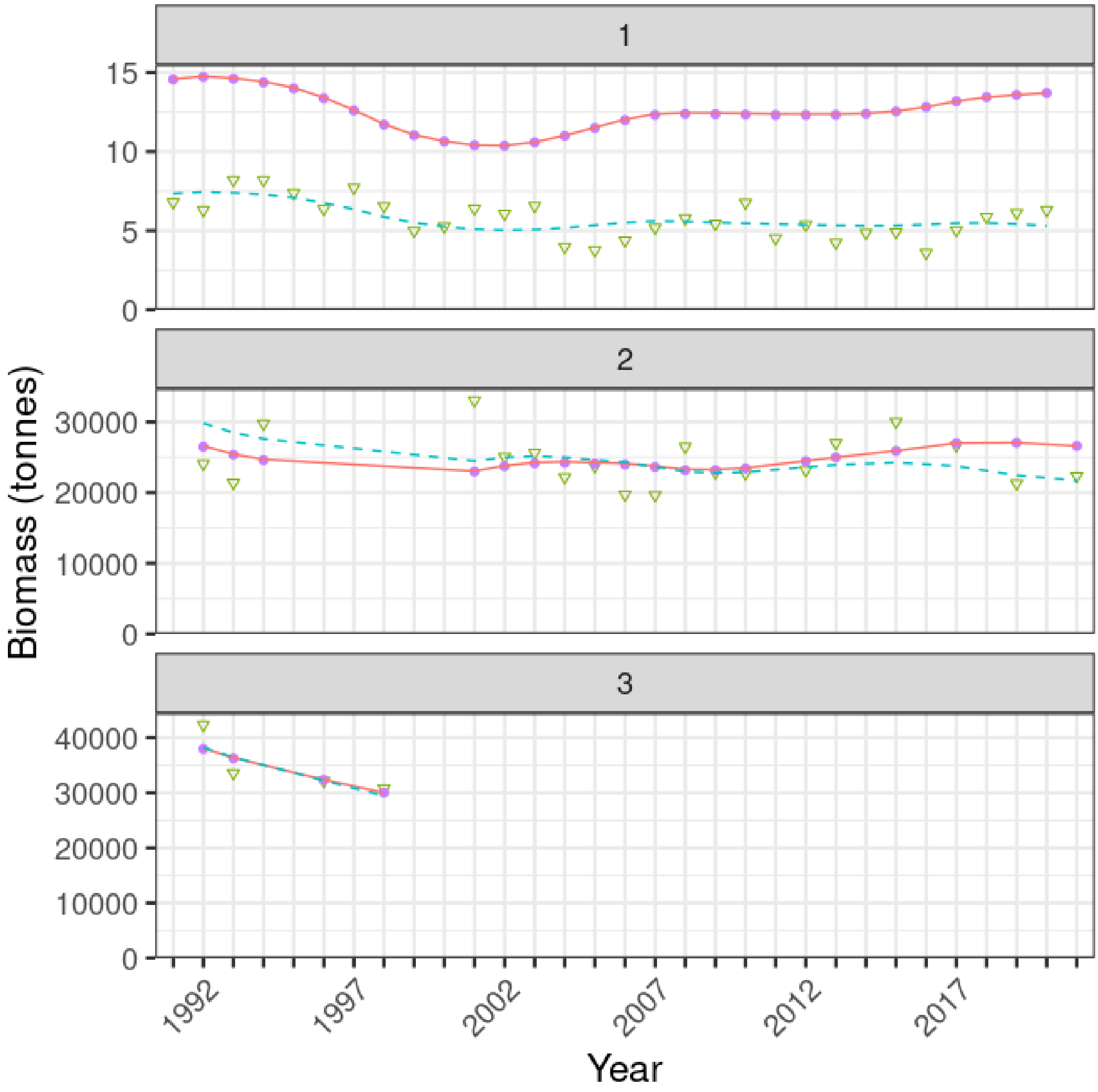

Figure 11: Survey biomass by year for each survey (1=longline CPUE, 2=summer survey, 3=autumn survey) showing the observations (green triangles), the exploratory CASAL model MPD (purple points), the MPD values pushed through the Stan-ASD model (red line), and the Stan-ASD MAP (dashed blue line). The CASAL model has free *q* and no scaling of the longline CPUE, leading to an over-estimation of *q*.

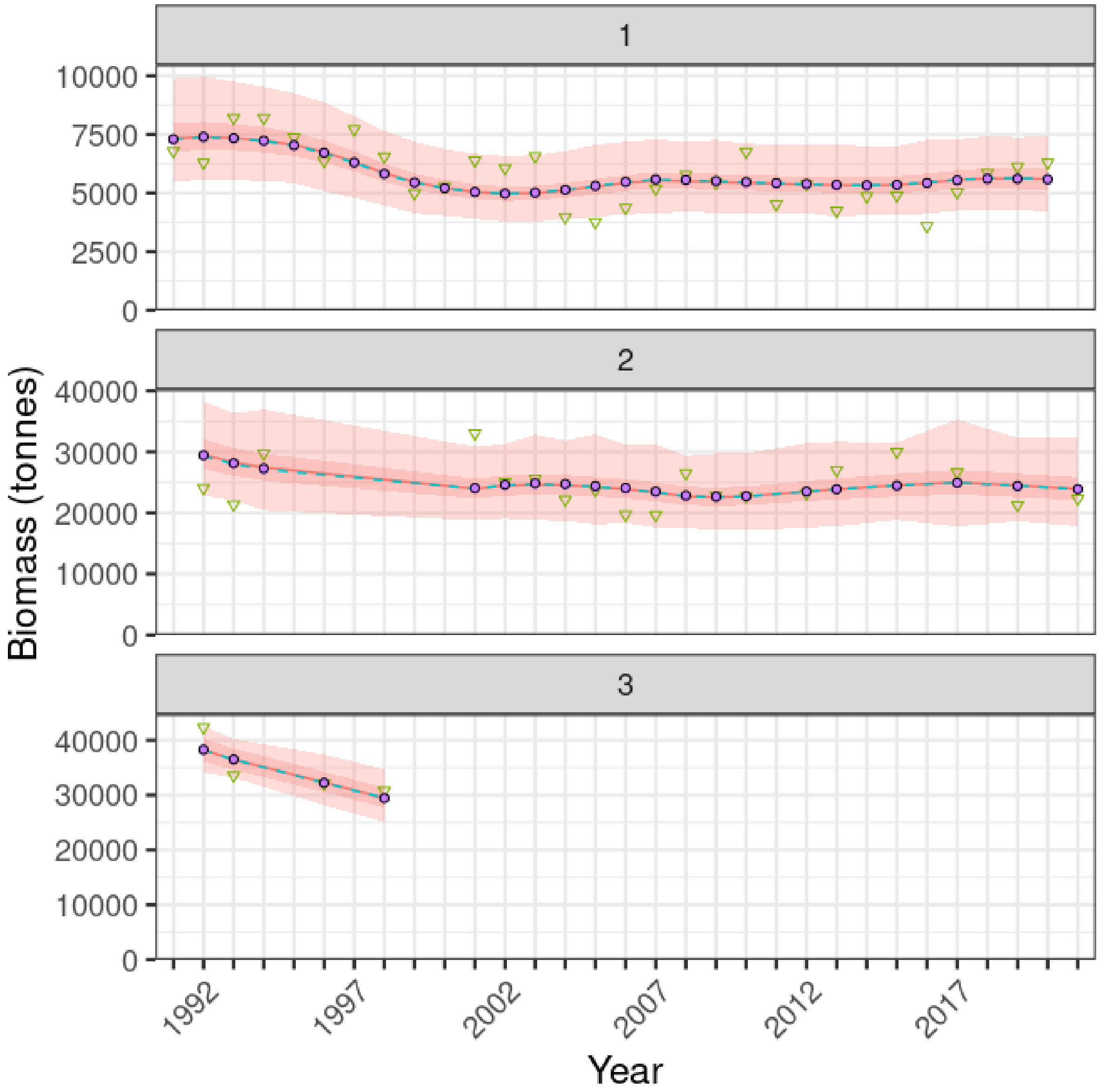

**Figure 12:** Survey biomass by year for each survey (1=longline CPUE, 2=summer survey, 3=autumn survey) showing the observations (green triangles), the exploratory CASAL model MPD (purple points), the MPD values pushed through the Stan-ASD model (open black points), and the Stan-ASD MAP (dashed blue line), the posterior median and 90% credible interval (red line and inner red shaded region), and posterior predictive distribution (outer shaded region).

**Table 4: The prior, likelihood, and penalty components from the base case CASAL model and the Stan-ASD fixed parameter model run for ling. The largest differences are indicated in red.**

| Component | Component type | CASAL MPD | Stan fixed parameter | Difference |
|---|---|---|---|---|
| Longline CPUE | Likelihood | -38.477 | -38.217 | -0.260 |
| Autumn biomass | Likelihood | -7.756 | -7.749 | -0.007 |
| Summer biomass | Likelihood | -25.575 | -25.535 | -0.039 |
| Autumn proportion at age | Likelihood | 166.757 | 166.694 | 0.063 |
| Summer proportion at age | Likelihood | 891.049 | 890.969 | 0.080 |
| Longline fishery proportion at age | Likelihood | 657.560 | 657.574 | -0.014 |
| Trawl fishery proportion at age | Likelihood | 985.767 | 985.709 | 0.058 |
| $q$ (CPUE longline) | Prior | -1.409 | -1.409 | 0.000 |
| $q$ (autumn) | Prior | -1.054 | -1.054 | 0.000 |
| $q$ (summer) | Prior | -2.496 | -2.496 | 0.000 |
| Process error CV (CPUE longline) | Prior | -1.855 | -1.855 | 0.000 |
| Process error CV (autumn) | Prior | -2.044 | -2.044 | 0.000 |
| Process error CV (summer) | Prior | -6.908 | -6.908 | 0.000 |
| $B_0$ | Prior | 12.153 | 12.153 | 0.000 |
| YCS | Prior | -11.939 | -12.537 | 0.598 |
| YCS average | Penalty | 0.626 | 1.083 | -0.457 |
| Total |  | 2 614.400 | 2 614.379 | 0.021 |

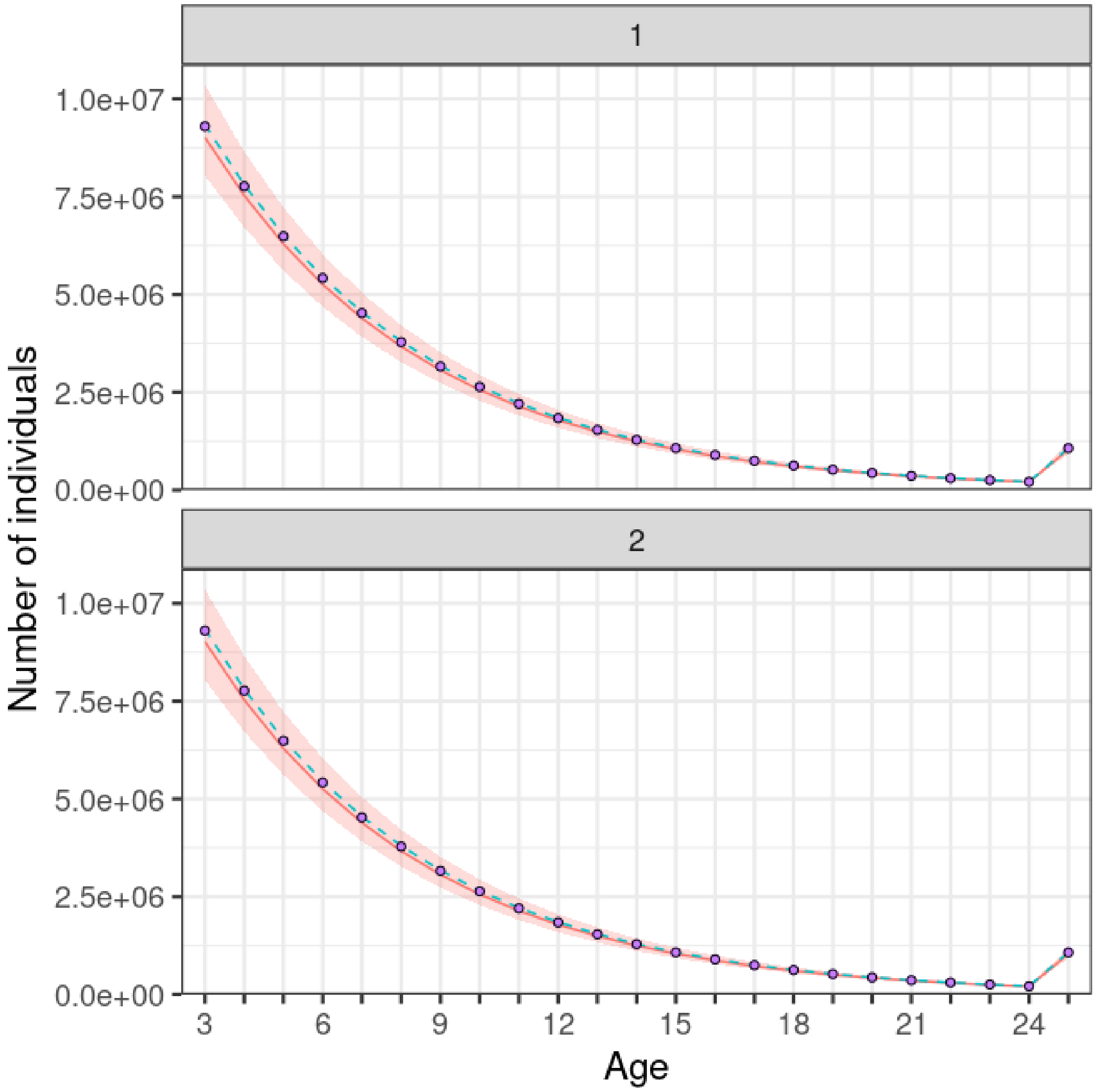

**Figure 13:** Initial numbers at age by sex (1=males, 2=females) for the CASAL models MPD (purple points), the CASAL MPD parameter values pushed through the Stan-ASD model as a fixed parameter run (open black points), the Stan-ASD MAP (dashed blue line), and the Stan-ASD posterior distribution (median is red line and red shading is the 90% credible interval).

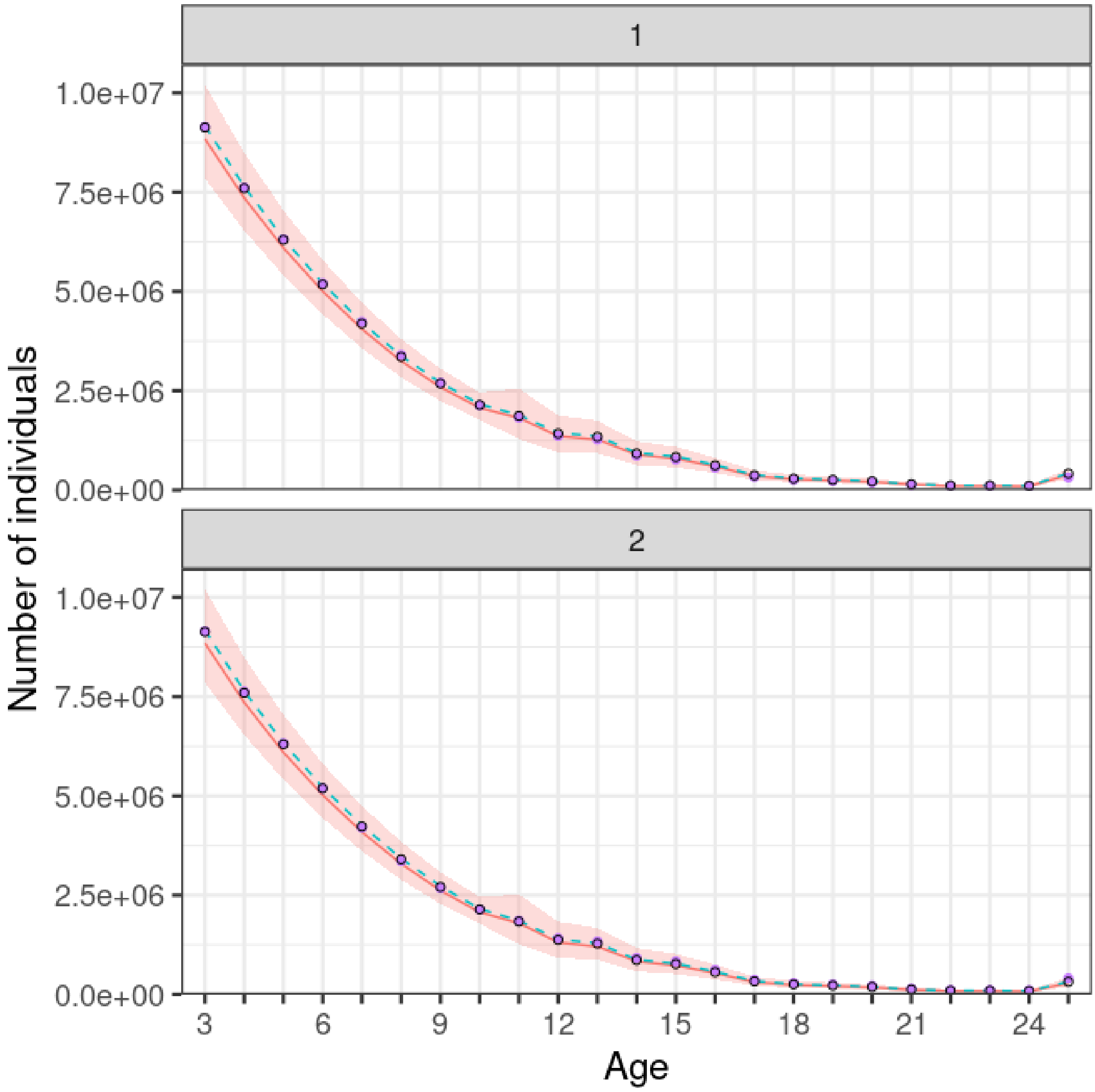

**Figure 14:** Final numbers at age by sex (1=males, 2=females) for the CASAL models MPD (purple points), the CASAL MPD parameter values pushed through the Stan-ASD model as a fixed parameter run (open black points), the Stan-ASD MAP (dashed blue line), and the Stan-ASD posterior distribution (median is red line and red shading is the 90% credible interval).

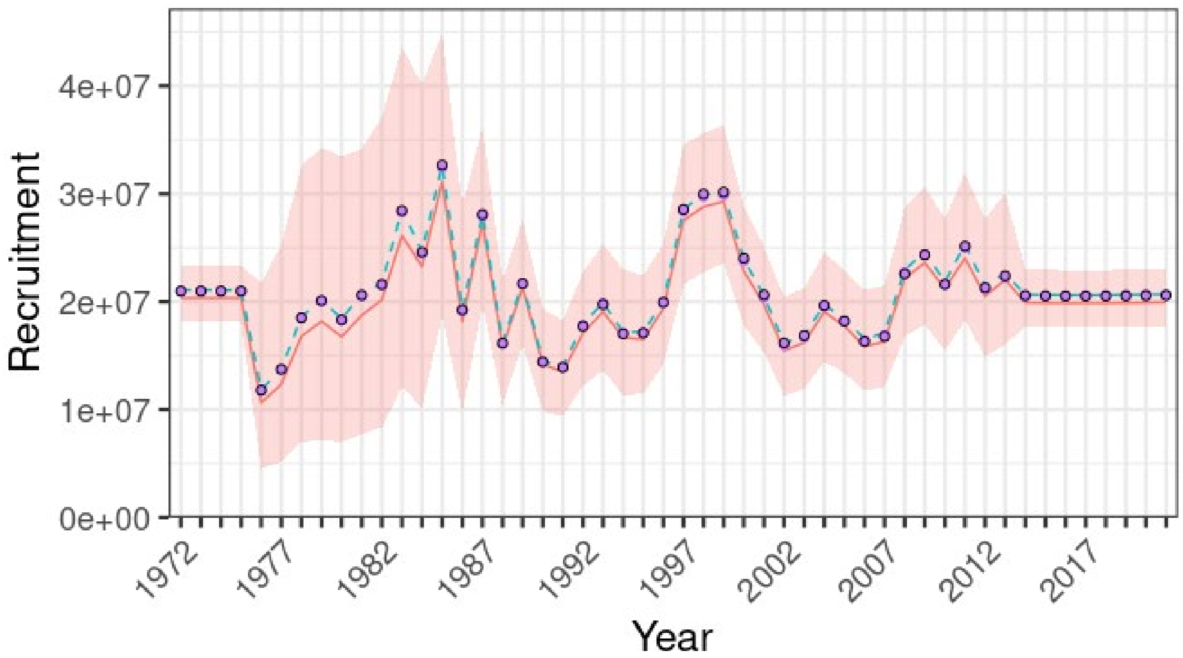

**Figure 15:** Recruitment by year for the CASAL models MPD (purple points), the CASAL MPD parameter values pushed through the Stan-ASD model as a fixed parameter run (open black points), the Stan-ASD MAP (dashed blue line), and the Stan-ASD posterior distribution (median is red line and red shading is the 90% credible interval).

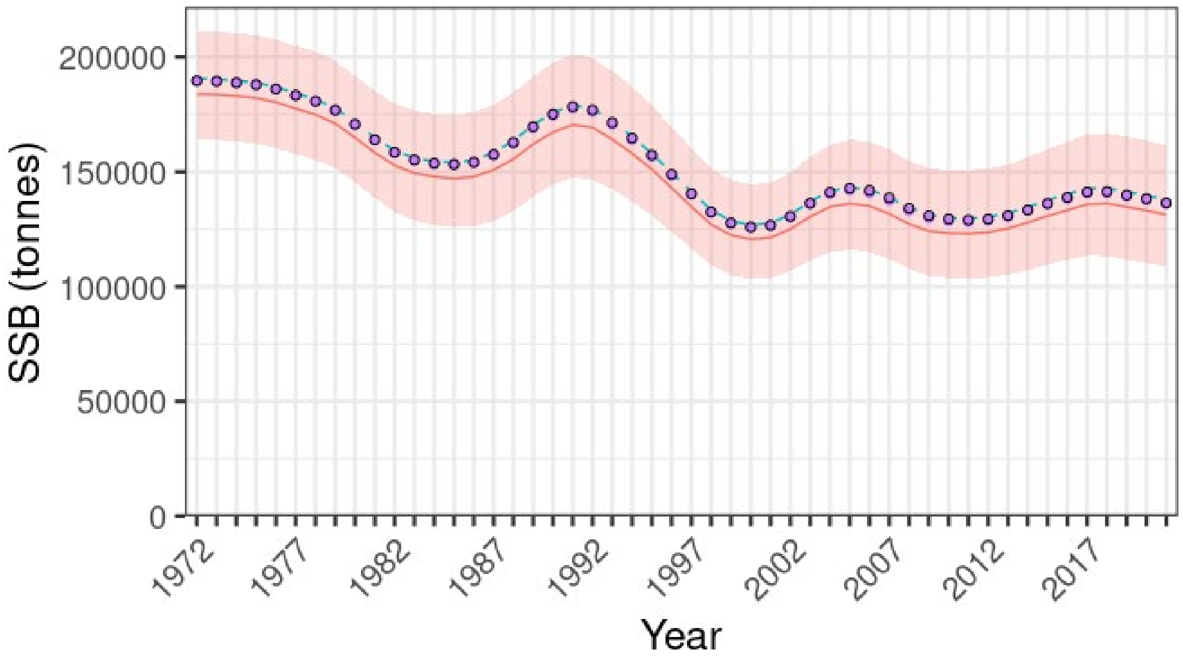

**Figure 16:** Spawning stock biomass (SSB) by year for the CASAL models MPD (purple points), the CASAL MPD parameter values pushed through the Stan-ASD model as a fixed parameter run (open black points), the Stan-ASD MAP (dashed blue line), and the Stan-ASD posterior distribution (median is red line and red shading is the 90% credible interval).

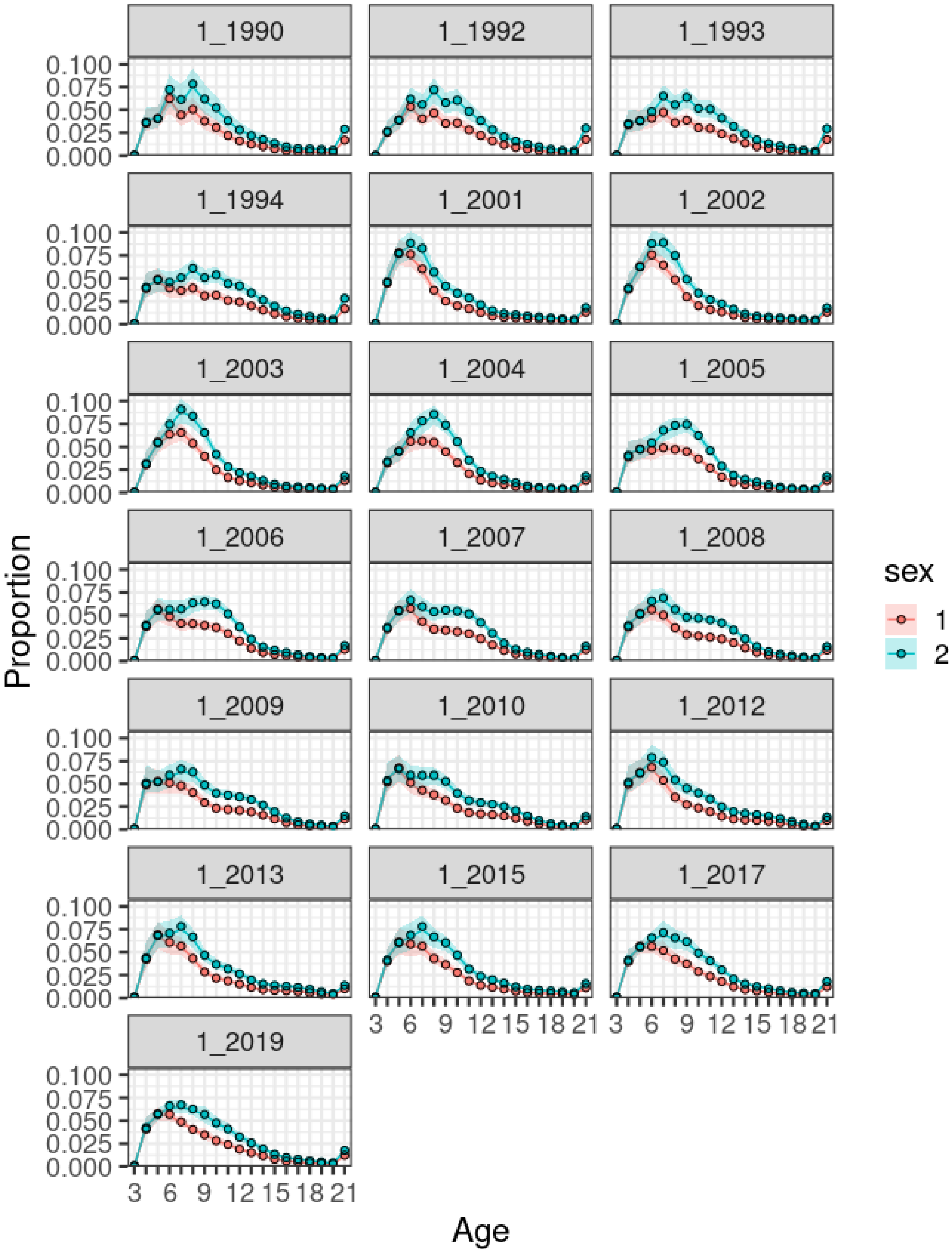

**Figure 17: Summer survey age composition by sex (1=males, 2=females) and year for the CASAL models MPD (coloured points), the MAP values pushed through the Stan-ASD model (open black points), and the Stan-ASD posterior median and 90% credible interval (solid line and shaded region).**

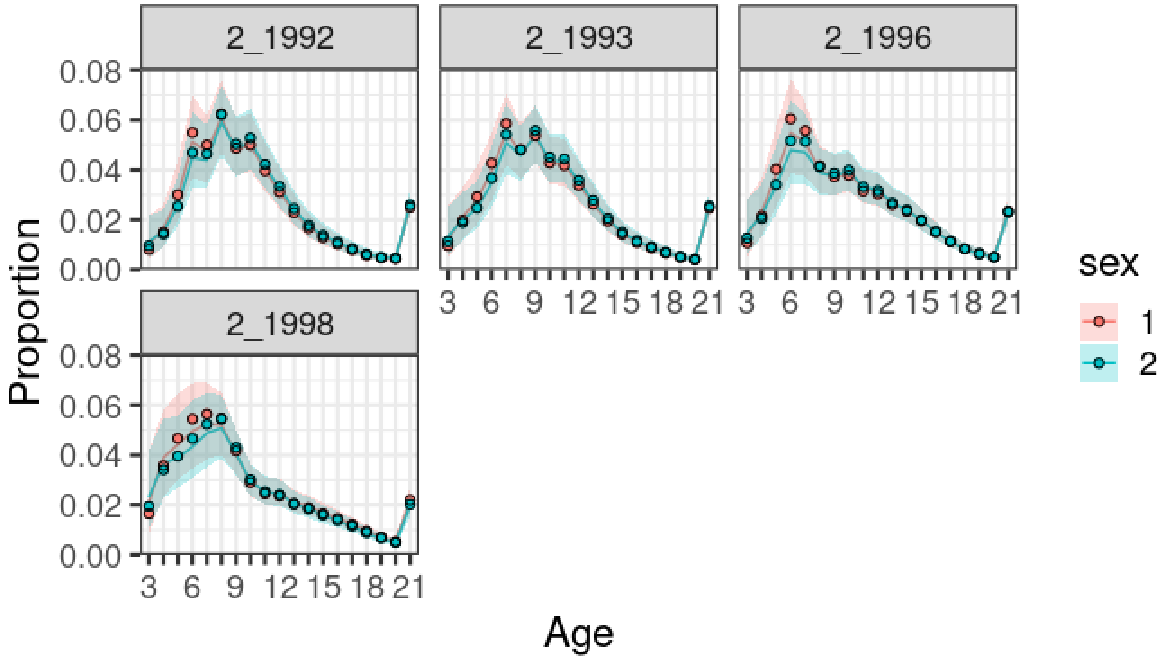

**Figure 18: Autumn survey age composition by sex (1=males, 2=females) and year for the CASAL models MAP (coloured points), the MPD values pushed through the Stan-ASD model (open black points), and the Stan-ASD posterior median and 90% credible interval (solid line and shaded region).**

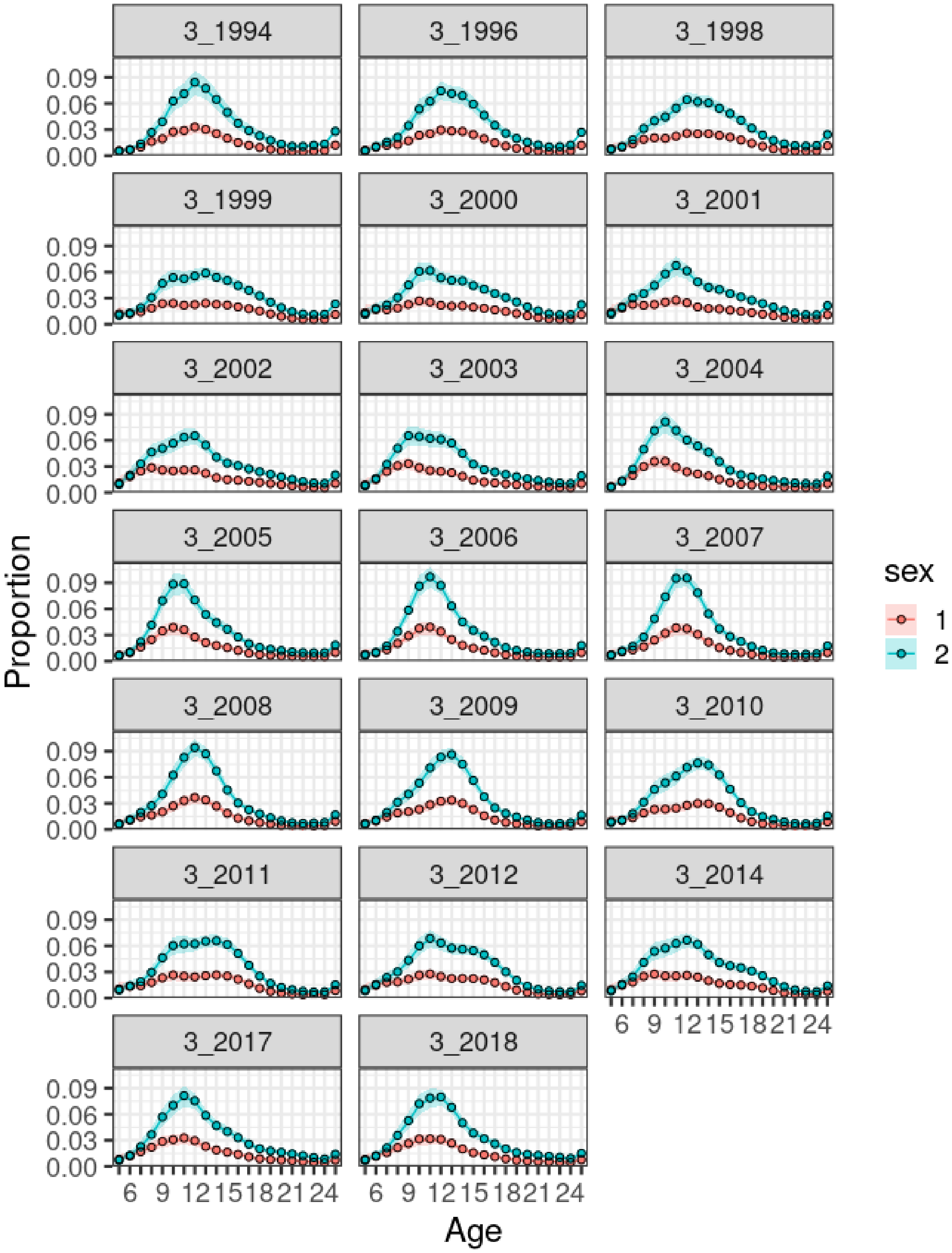

**Figure 19:** Longline age composition by sex (1=males, 2=females) and year for the CASAL models MAP (coloured points), the MPD values pushed through the Stan-ASD model (open black points), and the Stan-ASD posterior median and 90% credible interval (solid line and shaded region).

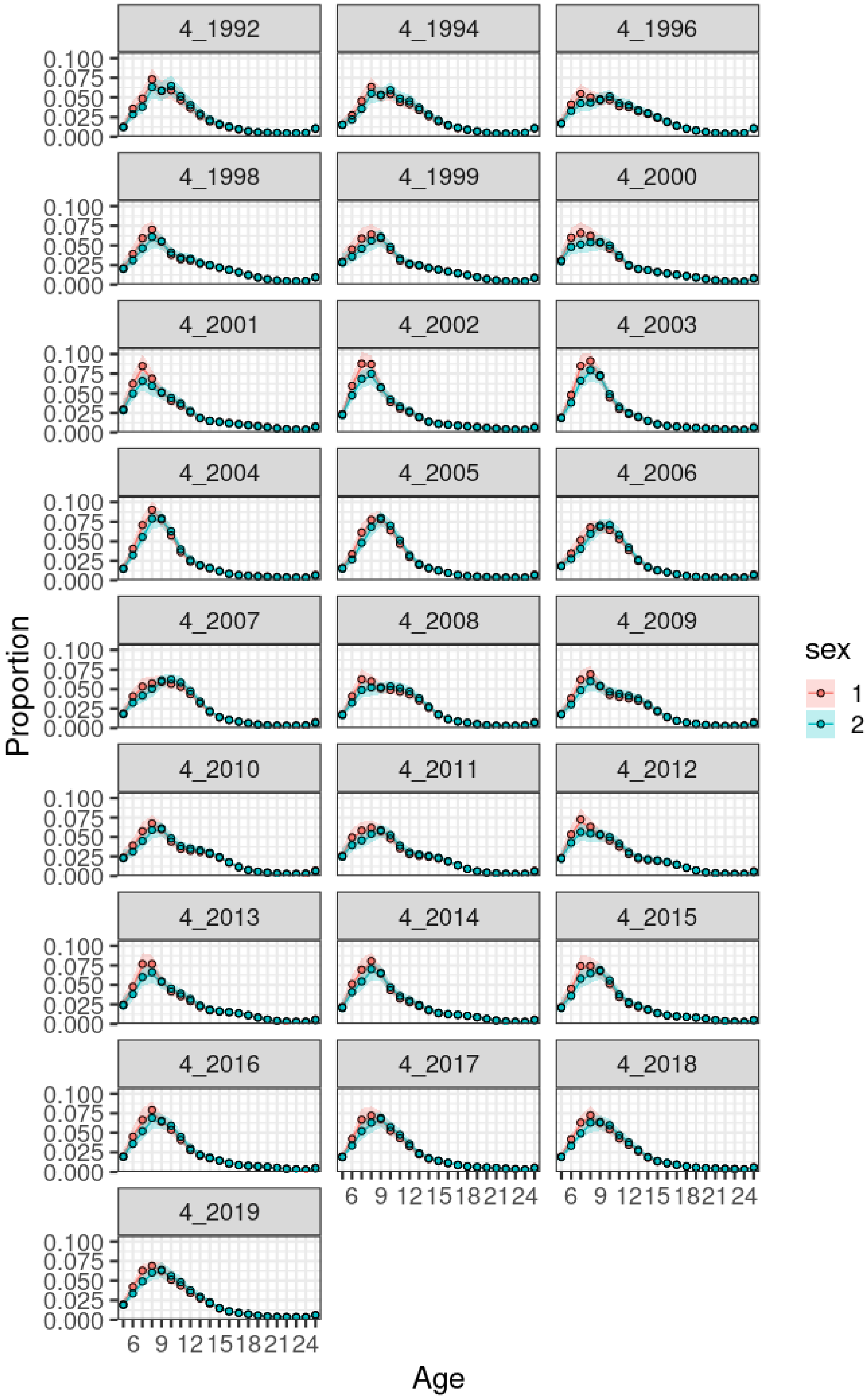

**Figure 20:** Trawl age composition by sex (1=males, 2=females) and year for the CASAL models MAP (coloured points), the MPD values pushed through the Stan-ASD model (open black points), and the Stan-ASD posterior median and 90% credible interval (solid line and shaded region).

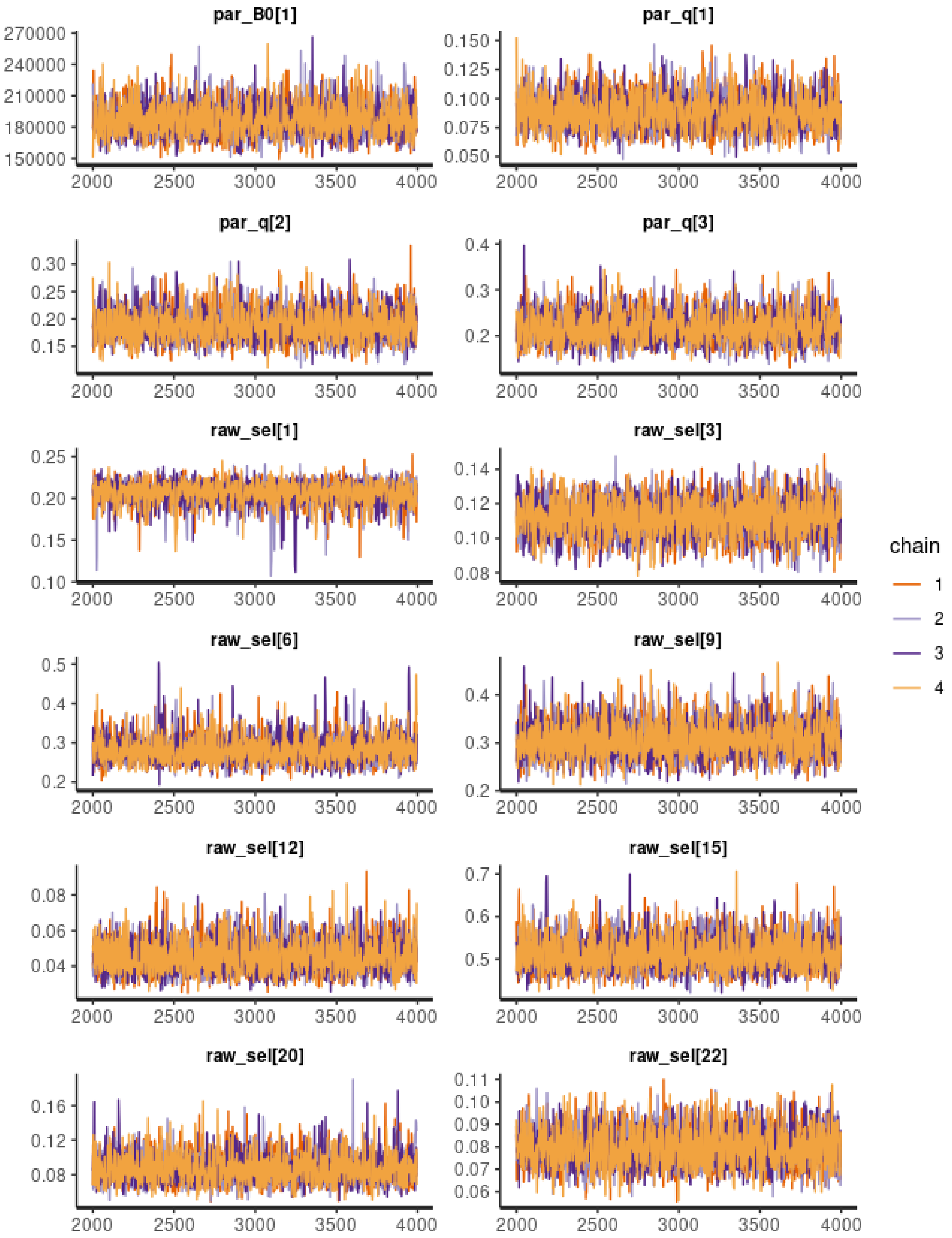

**Figure 21: MCMC trace plots for key model parameters by chain in the Stan-ASD ling model.**

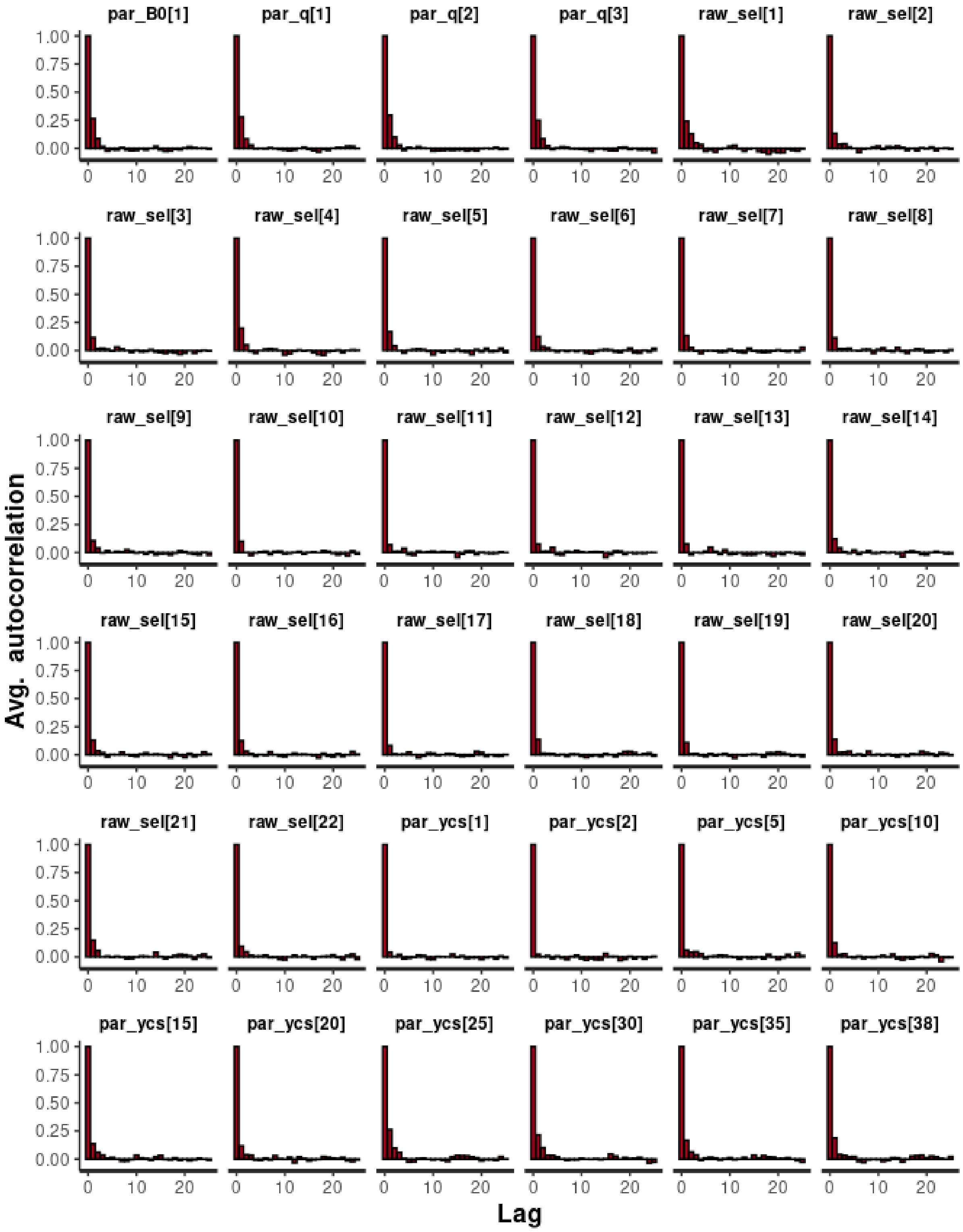

**Figure 22: MCMC autocorrelation for key model parameters for lags of 0–25 in the Stan-ASD ling model.**

**Table 5: Maximum *a posteriori* (MAP) and Markov chain Monte Carlo (MCMC) statistics for the Stan-ASD model runs for ling. Fixed parameters are indicated by grey shading. Although the right-hand limbs of the trawl selectivity ogives were fixed at 100 in CASAL, they were not fixed in Stan-ASD but were instead bounded to be between 99.99 and 100.01. Failed $\hat{R}$ statistics are indicated in red.**

| | MAP | MCMC | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Parameter | | Mean | SE mean | SD | 2.5% | 25% | 50% | 75% | 97.5% | Effective N | $\hat{R}$ |
| $B_{2021}$ | 137 841 | 132 479 | 320 | 16 176 | 104 909 | 121 331 | 131 291 | 142 129 | 168 999 | 2 548 | 1.001 |
| $B_0$ | 190 563 | 184 952 | 304 | 14 442 | 160 653 | 174 849 | 183 856 | 193 440 | 217 151 | 2 260 | 1.001 |
| $B_{2021} / B_0$ | 0.723 | 0.714 | 0.001 | 0.036 | 0.646 | 0.690 | 0.713 | 0.738 | 0.785 | 2 832 | 1.000 |
| $q$ (CPUE longline) | 0.082 | 0.089 | 0.000 | 0.014 | 0.064 | 0.079 | 0.088 | 0.098 | 0.120 | 2 020 | 1.001 |
| $q$ (autumn) | 0.175 | 0.192 | 0.001 | 0.028 | 0.143 | 0.173 | 0.190 | 0.209 | 0.254 | 2 025 | 1.001 |
| $q$ (summer) | 0.203 | 0.220 | 0.001 | 0.033 | 0.164 | 0.197 | 0.217 | 0.240 | 0.293 | 2 129 | 1.001 |
| Process error CV (CPUE longline) | 0.158 | 0.156 | - | - | 0.156 | 0.156 | 0.156 | 0.156 | 0.156 | - | - |
| Process error CV (autumn) | 0.129 | 0.130 | - | - | 0.130 | 0.130 | 0.130 | 0.130 | 0.130 | - | - |
| Process error CV (summer) | 0.001 | 0.001 | - | - | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | - | - |
| sel[1] (summer survey male) | 4.177 | 4.102 | 0.007 | 0.285 | 3.420 | 3.974 | 4.133 | 4.270 | 4.570 | 1 913 | 1.006 |
| sel[2] (summer survey male) | 2.007 | 2.661 | 0.025 | 1.132 | 1.223 | 1.972 | 2.450 | 3.090 | 5.380 | 2 033 | 1.003 |
| sel[3] (summer survey male) | 0.570 | 0.557 | 0.001 | 0.053 | 0.453 | 0.521 | 0.556 | 0.592 | 0.663 | 2 534 | 1.000 |
| sel[4] (summer survey female) | 5.405 | 5.640 | 0.010 | 0.443 | 4.961 | 5.322 | 5.578 | 5.890 | 6.668 | 2 094 | 0.999 |
| sel[5] (summer survey female) | 3.929 | 4.721 | 0.023 | 1.185 | 2.864 | 3.874 | 4.571 | 5.390 | 7.486 | 2 543 | 0.999 |
| sel[6] (autumn survey male) | 5.350 | 5.649 | 0.012 | 0.652 | 4.612 | 5.210 | 5.574 | 5.990 | 7.119 | 2 804 | 0.999 |
| sel[7] (autumn survey male) | 2.584 | 3.306 | 0.024 | 1.203 | 1.659 | 2.514 | 3.060 | 3.834 | 6.402 | 2 578 | 0.999 |
| sel[8] (autumn survey male) | 0.926 | 0.958 | 0.003 | 0.179 | 0.646 | 0.836 | 0.944 | 1.059 | 1.364 | 2 783 | 1.000 |
| sel[9] (autumn survey female) | 5.869 | 6.132 | 0.013 | 0.704 | 4.930 | 5.638 | 6.068 | 6.553 | 7.676 | 3 084 | 1.000 |
| sel[10] (autumn survey female) | 3.280 | 3.920 | 0.021 | 1.168 | 2.100 | 3.085 | 3.774 | 4.578 | 6.644 | 3 088 | 1.000 |
| sel[11] (trawl male) | 8.472 | 8.755 | 0.010 | 0.610 | 7.741 | 8.327 | 8.678 | 9.108 | 10.164 | 3 624 | 1.000 |
| sel[12] (trawl male) | 2.043 | 2.243 | 0.007 | 0.430 | 1.550 | 1.944 | 2.186 | 2.492 | 3.221 | 3 591 | 1.000 |
| sel[13] (trawl male) | 100.000 | 99.995 | 0.000 | 0.003 | 99.990 | 99.993 | 99.995 | 99.998 | 100.001 | 4 014 | 0.999 |
| sel[14] (trawl male) | 0.907 | 0.911 | 0.002 | 0.088 | 0.750 | 0.849 | 0.908 | 0.968 | 1.094 | 2 845 | 1.000 |
| sel[15] (trawl female) | 9.848 | 10.218 | 0.014 | 0.729 | 9.005 | 9.689 | 10.148 | 10.668 | 11.820 | 2 911 | 1.000 |
| sel[16] (trawl female) | 2.774 | 3.019 | 0.009 | 0.483 | 2.206 | 2.673 | 2.967 | 3.313 | 4.105 | 3 183 | 1.000 |
| sel[17] (trawl female) | 100.000 | 99.995 | 0.000 | 0.003 | 99.990 | 99.993 | 99.996 | 99.998 | 100.001 | 3 023 | 1.001 |
| sel[18] (line male) | 10.305 | 10.874 | 0.019 | 1.027 | 9.141 | 10.165 | 10.789 | 11.465 | 13.116 | 3 077 | 1.001 |
| sel[19] (line male) | 5.189 | 5.796 | 0.020 | 1.135 | 3.857 | 4.966 | 5.725 | 6.496 | 8.210 | 3 236 | 1.000 |
| sel[20] (line male) | 0.405 | 0.436 | 0.002 | 0.083 | 0.302 | 0.377 | 0.427 | 0.483 | 0.629 | 2 928 | 1.001 |
| sel[21] (line female) | 10.434 | 10.576 | 0.008 | 0.422 | 9.793 | 10.294 | 10.552 | 10.849 | 11.462 | 2 589 | 1.000 |
| sel[22] (line female) | 3.888 | 4.012 | 0.008 | 0.411 | 3.246 | 3.731 | 3.998 | 4.274 | 4.871 | 2 815 | 1.000 |
| Objective function | -2 614 | -2 705 | 0.115 | 5.885 | -2 717 | -2 708 | -2 704 | -2 701 | -2 694 | 2 639 | 1.002 |

**Table 6: Maximum *a posteriori* (MAP) and Markov chain Monte Carlo (MCMC) statistics for CASAL and Stan-ASD model runs for ling. Fixed parameters are indicated by grey shading. Although the right-hand limbs of the trawl selectivity ogives were fixed at 100 in CASAL, they were not fixed in Stan-ASD but were instead bounded to be between 99.99 and 100.01.**

| | MAP | | CASAL MCMC | | | | Stan-ASD MCMC | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Parameter | CASAL | Stan-ASD | Mean | 2.5% | 50% | 97.5% | Mean | 2.5% | 50% | 97.5% |
| $B_{2021}$ | 136 061 | 137 841 | 131 578 | 102 895 | 128 426 | 182 959 | 132 479 | 104 909 | 131 291 | 168 999 |
| $B_0$ | 189 661 | 190 563 | 185 397 | 159 568 | 182 922 | 233 756 | 184 952 | 160 653 | 183 856 | 217 151 |
| $B_{2021} / B_0$ | 0.717 | 0.723 | 0.707 | 0.634 | 0.704 | 0.789 | 0.714 | 0.646 | 0.713 | 0.785 |
| $q$ (CPUE longline) | 0.082 | 0.082 | 0.089 | 0.057 | 0.088 | 0.121 | 0.089 | 0.064 | 0.088 | 0.120 |
| $q$ (autumn) | 0.177 | 0.175 | 0.194 | 0.130 | 0.192 | 0.260 | 0.192 | 0.143 | 0.190 | 0.254 |
| $q$ (summer) | 0.204 | 0.203 | 0.218 | 0.152 | 0.216 | 0.290 | 0.220 | 0.164 | 0.217 | 0.293 |
| Process error CV (CPUE longline) | 0.156 | 0.158 | 0.156 | 0.156 | 0.156 | 0.156 | 0.156 | 0.156 | 0.156 | 0.156 |
| Process error CV (autumn) | 0.130 | 0.129 | 0.130 | 0.130 | 0.130 | 0.130 | 0.130 | 0.130 | 0.130 | 0.130 |
| Process error CV (summer) | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| sel[1] (summer survey male) | 4.180 | 4.177 | 4.131 | 3.548 | 4.145 | 4.571 | 4.102 | 3.420 | 4.133 | 4.570 |
| sel[2] (summer survey male) | 2.014 | 2.007 | 2.607 | 1.194 | 2.438 | 5.067 | 2.661 | 1.223 | 2.450 | 5.380 |
| sel[3] (summer survey male) | 0.570 | 0.570 | 0.558 | 0.456 | 0.559 | 0.661 | 0.557 | 0.453 | 0.556 | 0.663 |
| sel[4] (summer survey female) | 5.406 | 5.405 | 5.630 | 4.940 | 5.557 | 6.715 | 5.640 | 4.961 | 5.578 | 6.668 |
| sel[5] (summer survey female) | 3.926 | 3.929 | 4.668 | 2.904 | 4.514 | 7.535 | 4.721 | 2.864 | 4.571 | 7.486 |
| sel[6] (autumn survey male) | 5.357 | 5.350 | 5.648 | 4.635 | 5.582 | 7.014 | 5.649 | 4.612 | 5.574 | 7.119 |
| sel[7] (autumn survey male) | 2.592 | 2.584 | 3.260 | 1.698 | 3.062 | 6.127 | 3.306 | 1.659 | 3.060 | 6.402 |
| sel[8] (autumn survey male) | 0.927 | 0.926 | 0.964 | 0.667 | 0.949 | 1.353 | 0.958 | 0.646 | 0.944 | 1.364 |
| sel[9] (autumn survey female) | 5.875 | 5.869 | 6.100 | 4.921 | 6.047 | 7.624 | 6.132 | 4.930 | 6.068 | 7.676 |
| sel[10] (autumn survey female) | 3.286 | 3.280 | 3.864 | 2.060 | 3.743 | 6.498 | 3.920 | 2.100 | 3.774 | 6.644 |
| sel[11] (trawl male) | 8.451 | 8.472 | 8.716 | 7.743 | 8.669 | 10.031 | 8.755 | 7.741 | 8.678 | 10.164 |
| sel[12] (trawl male) | 2.038 | 2.043 | 2.229 | 1.534 | 2.183 | 3.190 | 2.243 | 1.550 | 2.186 | 3.221 |
| sel[13] (trawl male) | 100.000 | 100.000 | 100.000 | 100.000 | 100.000 | 100.000 | 99.995 | 99.990 | 99.995 | 100.001 |
| sel[14] (trawl male) | 0.916 | 0.907 | 0.925 | 0.762 | 0.923 | 1.106 | 0.911 | 0.750 | 0.908 | 1.094 |
| sel[15] (trawl female) | 9.783 | 9.848 | 10.100 | 8.863 | 10.034 | 11.854 | 10.218 | 9.005 | 10.148 | 11.820 |
| sel[16] (trawl female) | 2.751 | 2.774 | 2.967 | 2.142 | 2.917 | 4.170 | 3.019 | 2.206 | 2.967 | 4.105 |
| sel[17] (trawl female) | 100.000 | 100.000 | 100.000 | 100.000 | 100.000 | 100.000 | 99.995 | 99.990 | 99.996 | 100.001 |
| sel[18] (line male) | 10.341 | 10.305 | 10.853 | 9.086 | 10.765 | 13.307 | 10.874 | 9.141 | 10.789 | 13.116 |
| sel[19] (line male) | 5.253 | 5.189 | 5.803 | 3.795 | 5.713 | 8.377 | 5.796 | 3.857 | 5.725 | 8.210 |
| sel[20] (line male) | 0.412 | 0.405 | 0.441 | 0.306 | 0.432 | 0.626 | 0.436 | 0.302 | 0.427 | 0.629 |
| sel[21] (line female) | 10.425 | 10.434 | 10.542 | 9.759 | 10.527 | 11.411 | 10.576 | 9.793 | 10.552 | 11.462 |
| sel[22] (line female) | 3.909 | 3.888 | 4.030 | 3.265 | 4.022 | 4.914 | 4.012 | 3.246 | 3.998 | 4.871 |
| Objective function | -2 614 | -2 614 | -2 655 | -2 643 | -2 655 | -2 669 | -2 705 | -2 717 | -2 704 | -2 694 |

# 4. DISCUSSION

Good matches were achieved between CASAL and Stan-ASD for both the hake and ling models. The two examples tested different model structures: for example, the hake model included ages 1–30+ and the ling model ages 3–25+; the hake model included three time steps within the annual cycle, and the ling model had two time steps; three selectivity ogives were estimated in the hake model (November survey series, April–May survey series, and fishery selectivity), whereas eight selectivity ogives were estimated in the ling model (summer survey, winter survey, trawl fishery, and line fishery all by sex); the hake model was fitted to proportions at age for males and females combined (i.e., numbers of males and females were summed), whereas the ling model was fitted to males and females separately (i.e., male and female proportions were concatenated into a single vector and then fitted to using the multinomial likelihood); and natural mortality was an estimated parameter in the hake model but was fixed for ling.

The process of implementing the Stan-ASD model also identified a potential, but minor, bug in CASAL relating to the treatment of $q$ (see Figure 11 and the description in Section 3.2). Although this is an issue that is easily remedied, it is good to identify such issues and understand how they arise.

Although each MCMC iteration is much slower in Stan-ASD when compared to CASAL, far fewer MCMC iterations are required to be run to obtain adequately mixed MCMC chains. This is due to the improved efficiency of the algorithms implemented by Stan (NUTS) from those by CASAL (random walk Metropolis-Hastings). This results in Stan-ASD providing posterior samples in a fraction of the time and is beneficial because it allows the stock assessment scientist to test alternative model structures, using Bayesian inference, much more rapidly.

One of the outputs from the Stan-ASD model is the posterior predictive distribution of survey biomass for each survey. The posterior predictive distribution is an excellent diagnostic tool that allows stock assessment scientists to visually check if the predicted uncertainty about an index of abundance is sensible. We recommend this as a standard output for any stock assessment model.

We see the Stan-ASD model as being complementary to the CASAL model because it allows us to rapidly code new ideas/concepts and test various aspects of age-structured stock assessment models. If worthwhile, then these new concepts can be implemented into CASAL.

Finally, we suggest that further development is necessary if the Stan-ASD model is to become more easily accessible to other users. Future development should include:

- further documentation and streamlining of the code;
- implementing an alternative recruitment parameterisation (see Appendix II);
- implementing an alternative exploitation rate parameterisation (see Appendix III);
- implementing projections, standard yield calculations, and fishery performance metrics;
- implementing standard diagnostic plot function calls;
- bundling the code up into an easily distributable R package;
- implementing Travis CI unit test functionality for future-proofing code updates and changes (i.e., each time the Stan-ASD code is updated and pushed to GitHub, Travis CI can automatically run a series of unit tests to ensure that the code still runs and that model outputs are consistent).

# 5. ACKNOWLEDGEMENTS

We thank the Fisheries New Zealand Data Management Team for the data extracts used in these analyses, and the members of the Deepwater Fisheries Assessment Working Group for their input and discussions of this work. This work was funded by Fisheries New Zealand projects HAK2020-01, LIN2020-01, and SEA2020-15.

# 6. REFERENCES

Beverton, R.; Holt, S. (1957). On the Dynamics of Exploited Fish Populations. Series 2, Vol. 19, Ministry of Agriculture, Fisheries and Food, London.

Bull, B.; Dunn, A.; McKenzie, A.; Gilbert, D.J.; Smith, M.H.; Bian, R.; Fu, D. (2012). CASAL (C++ algorithmic stock assessment laboratory) User Manual v2.30-2012/03/21. *NIWA Technical Report 135*. 280 p.

Dunn, A.; Mormede, S.; Webber, D.N. (in press). Stock assessment of hake (*Merluccius australis*) in the Sub-Antarctic (HAK 1) for the 2020–21 fishing year. Draft New Zealand Fisheries Assessment Report.

Gelman, A.; Carlin, J.B.; Stern, H.S.; Dunson, D.B.; Vehtari A.; Rubin, D.B. (2013). Bayesian Data Analysis. Third Edition. London: Chapman & Hall / CRC Press.

Geyer, C.J. (1992). Practical Markov Chain Monte Carlo. *Statistical Science 7 (4)*: 473–483.

Geyer, C.J. (2011). Introduction to Markov Chain Monte Carlo. *In*: Brooks, S.; Gelman, A.; Jones, G.L.; Meng, X.-L. (eds), pp. 3–48. *Handbook of Markov Chain Monte Carlo*, Chapman; Hall/CRC.

Mormede, S.; Dunn, A.; Webber, D.N. (in press). Stock assessment of ling (*Genypterus blacodes*) in the Sub-Antarctic (LIN 5&6) for the 2020–21 fishing year. Draft New Zealand Fisheries Assessment Report.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Stan Development Team. (2021). Stan Modeling Language Users Guide and Reference Manual, 2.21.2. https://mc-stan.org

Vehtari, A.; Gelman, A.; Simpson, D.; Carpenter, B.; Bürkner P.-C. (2019). Rank-normalization, folding, and localization: An improved R-hat for assessing convergence of MCMC. arXiv preprint arXiv:1903.08008.

Webber, D.N.; Haist, V.; Starr, P.J.; Edwards, C.T.T. (2018). A new model for the assessment of New Zealand rock lobster (*Jasus edwardsii*) stocks and an exploratory multi-area CRA 4 assessment. *New Zealand Fisheries Assessment Report 2018/53*. 111 p.

# 7. APPENDIX I: CASAL PLOTS

## 7.1 Hake

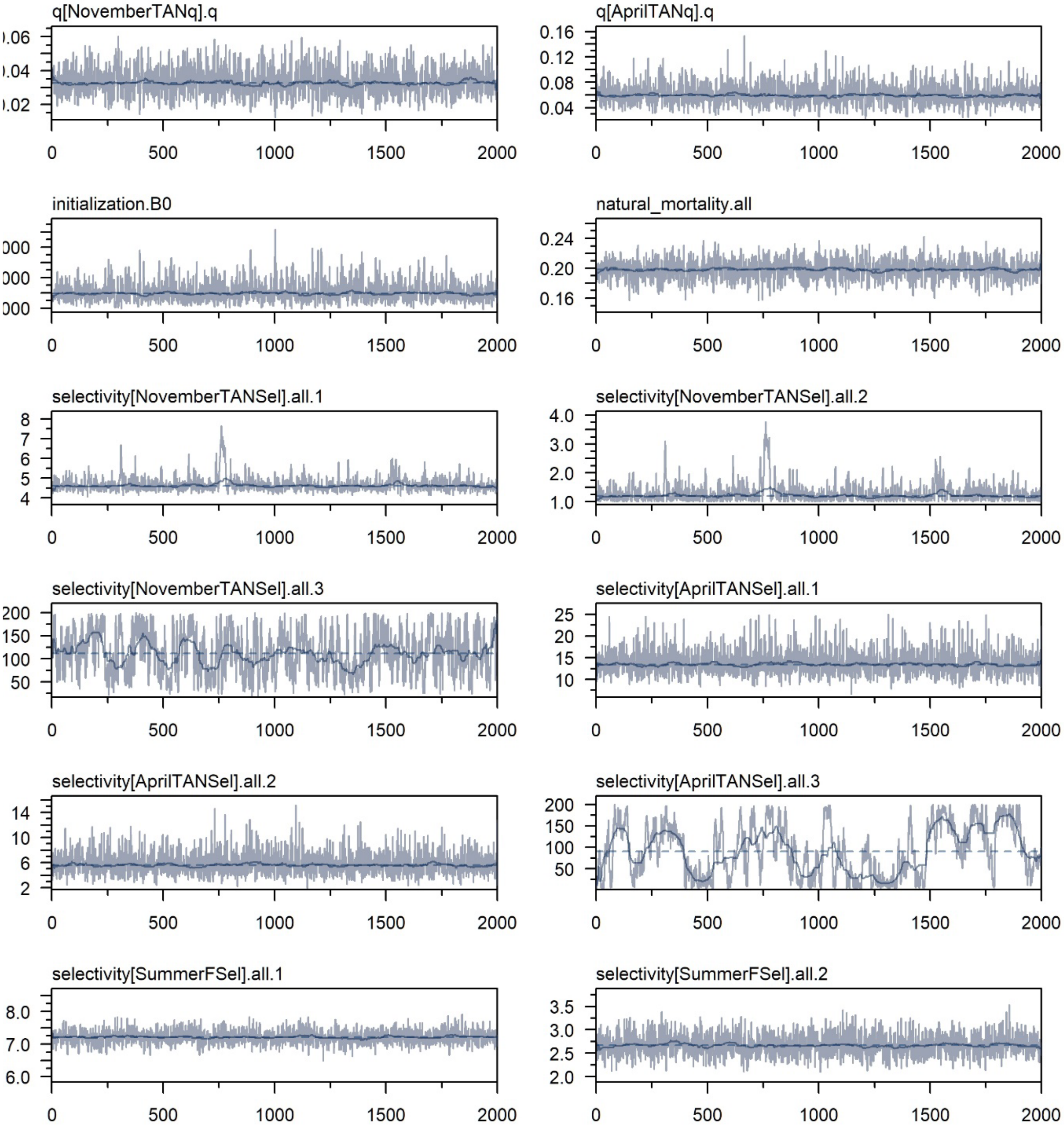

**Figure 23: MCMC trace plots for key model parameters by chain in the CASAL hake model.**

## 7.2 Ling

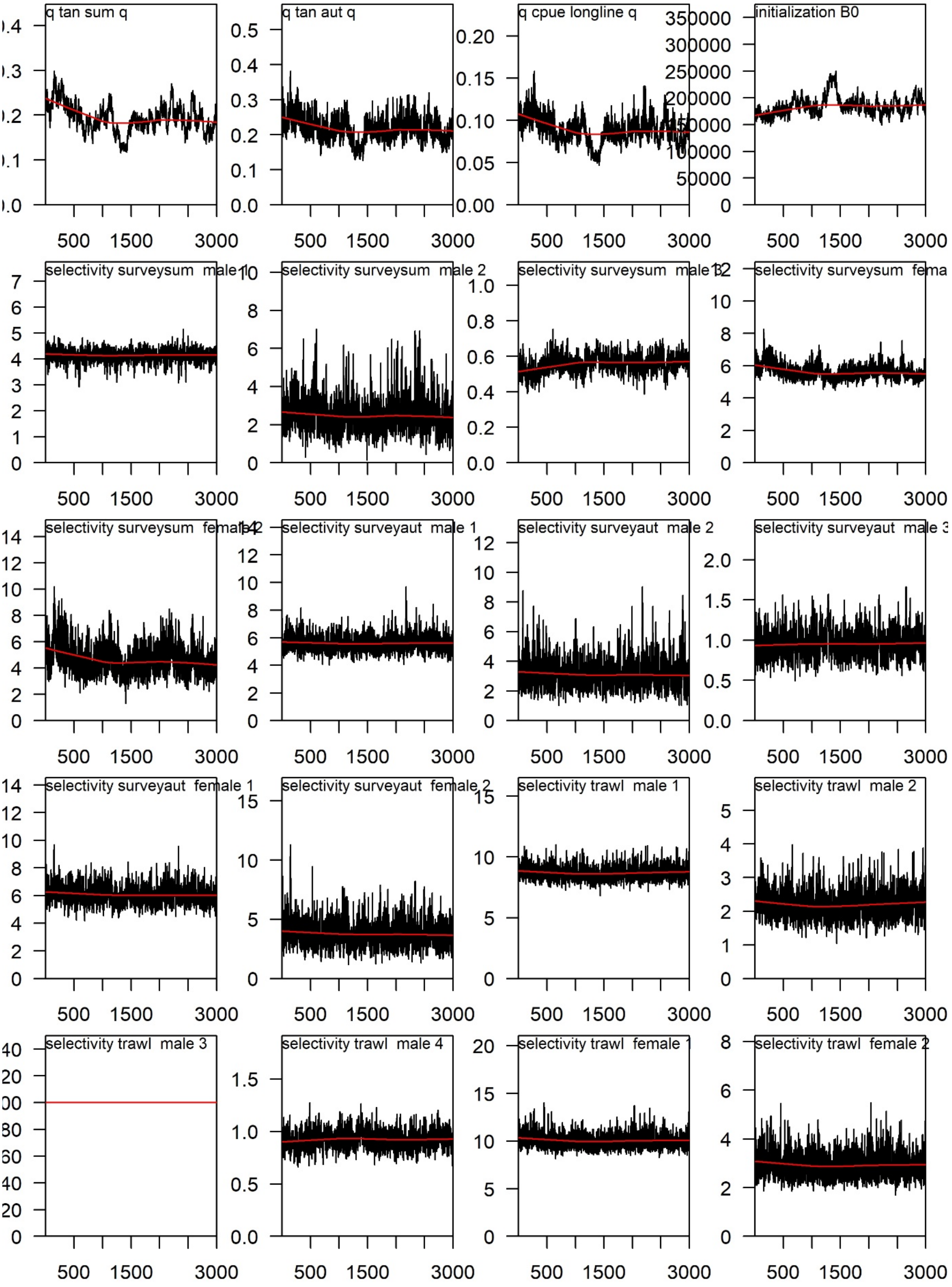

**Figure 24: MCMC trace plots for key model parameters by chain from the modified CASAL ling model with free *q*.**

## 8. APPENDIX II: AN ALTERNATIVE PARAMETERISATION FOR ANNUAL RECRUITMENT VALUES

In CASAL and other similar software, recruitment each year has been defined as

$$R_y = R_0 \times YCS_{y-e} \times SR\left(SSB_{y-e}\right)$$

where $R_0$ average recruitment, $YCS_y$ is the year class strength for each year, and $SR\left(SSB_{y-e}\right)$ is the stock recruitment multiplier which is a function of the spawning stock biomass in some years prior $e$ (e.g., Bull et al. 2012). To ensure that $R_0$ has the correct meaning, the YCSs are standardised (i.e., transformed by dividing the YCS vector by the mean of that vector) to have mean 1, and a penalty is applied to reduce the impact of this standardisation. We show the actual YCS parameters and the standardised YCSs for the ling example below which illustrates the impact of this standardisation (Figure 25).

Instead, we suggest defining recruitment strength as a group-level effect (i.e., a random-effect in frequentist terminology)

$$\log\left(Y_{y \in S}\right) \sim N(\mu_R, \sigma_R^2)$$

where $\mu_R$ is the mean recruitment strength, $\sigma_R^2$ is the variance, and $S$ is the set of years for which we want to estimate recruitment strengths. We can calculate $R_0$ as

$$R_0 = \frac{1}{n_S} \sum_y Y_y$$

where $n_S$ is the number of years for which we want to estimate annual recruitment. Recruitment each year is then defined as

$$R_y = \begin{cases} Y_{y-e} \times SR\left(SSB_{y-e}\right) \text{ for } y \in S \\ R_0 \times SR\left(SSB_{y-e}\right) \text{ for } y \notin S \end{cases}.$$

Note that this parameterisation estimates the same number of parameters that is estimated under the standard framework but does not require the YCSs to be standardised and therefore does not require the application of a penalty. This should result in improved MCMC performance and will potentially allow estimation of $\sigma_R^2$ in MCMC, allow for natural extensions including random walk YCS constraints and autocorrelation between YCSs, and allow implementing and evaluating evidence for regime shifts in recruitment within a stock assessment.

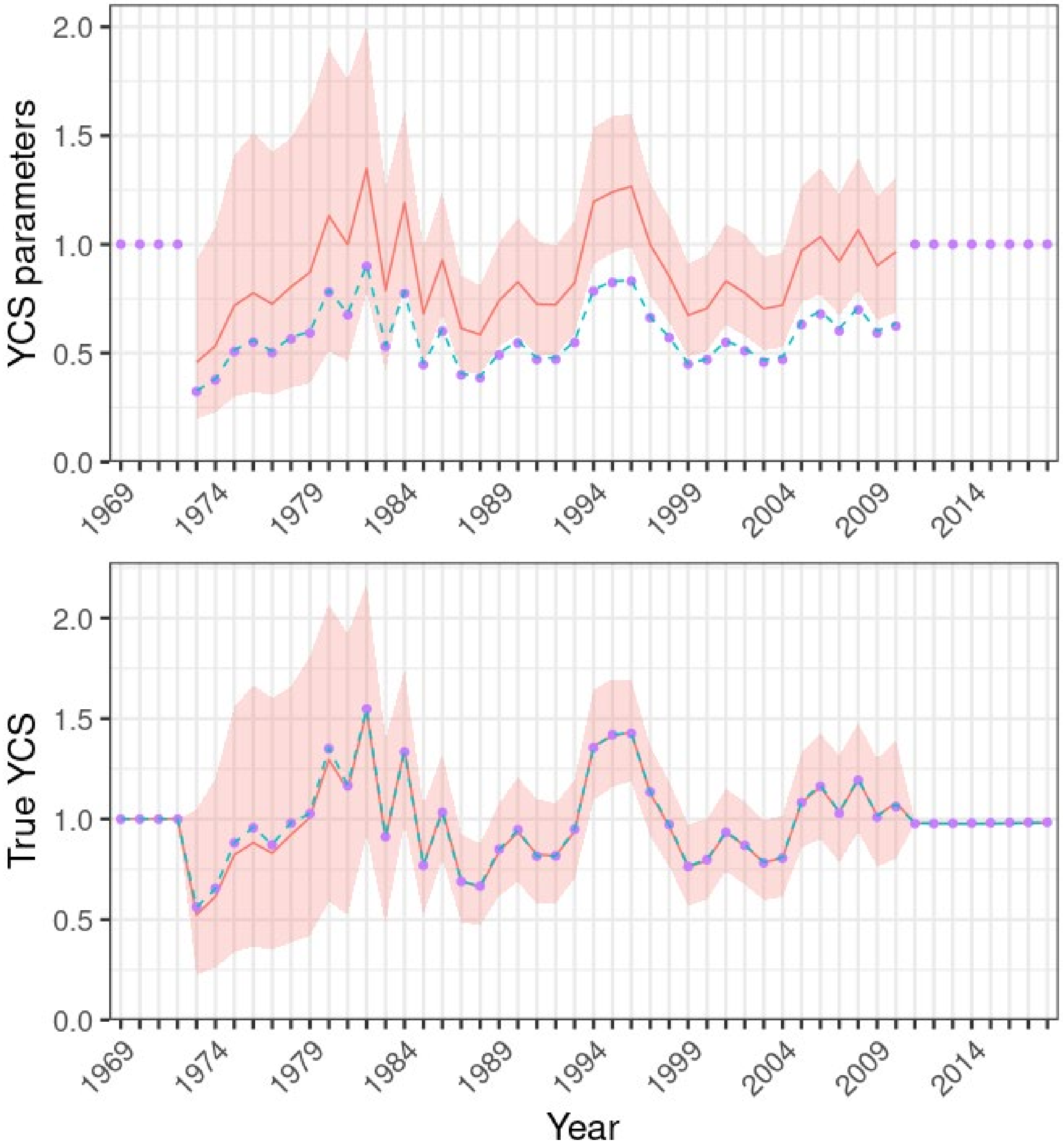

**Figure 25:** **YCS parameters [top] and YCSs after standardisation (i.e., rescaled to have mean 1) by year for the ling CASAL models MPD (purple points), the Stan-ASD MAP (dashed blue line), and the Stan-ASD posterior distribution (median is red line and red shading is the 90% credible interval).**

# 9. APPENDIX III: AN ALTERNATIVE PARAMETERISATION FOR FISHERIES EXPLOITATION

Generally, vulnerable biomass (tonnes) for each year $y$ and fishery $f$ has been defined as

$$V_{y,f} = \sum_{a} N_{a,y} w_a S_{a,f}$$

where $N_{a,y}$ is the numbers at age $a$ each year, $w_a$ is the mean weight at age, and $S_{a,f}$ is the selectivity at age for each fishery. In this case, the selectivity for each fishery is a function of parameters that define the shape of the selectivity ogive (e.g., two parameters for a logistic ogive or three parameters for a double normal ogive) and a parameter that defines the height of the selectivity ogive (i.e., the maximum proportion selected). For example, a model that includes a different selectivity ogive for males and females may fix the maximum height parameter to 1 for males but estimate the maximum height parameter for females (which may be greater than or less than 1). This allows the stock assessment model to define the proportions of ages selected and the proportion of males versus females that are selected.

Instead, we can define vulnerable biomass as

$$V_{y,f} = \sum_{a} N_{a,y} w_a S_{a,f} v_f \quad \text{where} \quad \sum_{f} v_f = 1 \quad \text{and} \quad S_{a,f} \in [0,1]$$

where $v_f$ is the vulnerability for each fishery $f$ and the maximum height for selectivity is always 1. The vulnerability variable replaces the maximum height parameter, essentially splitting the two processes (e.g., proportions and age and proportions of each sex selected). This helps with the model interpretation and parameterisation, and estimation. As usual, exploitation can then be defined as

$$U_{y,f} = \frac{C_{y,f}}{V_{y,f}} \quad \text{where} \quad U_{y,f} \in [0,1]$$

where $C_{y,f}$ is the catch (tonnes) during each year associated with each fishery. Here, exploitation is bounded to be between 0 and 1. Catch is then removed from the population using

$$N'_{a,y} = N_{a,y} \left( 1 - \sum_{f} U_{y,f} S_{a,f} v_f \right)$$

where $N_{a,y}$ is bounded to be greater than zero. And hence, a maximum exploitation rate is no longer required.