Supporting analysis S-1: Fishery selectivity alternatives

RTMB technical comparison of alternative time-varying fishery selectivity forms

James Ianelli

August 17, 2026

This supporting analysis compares the bridge-reference fishery selectivity specification with a hierarchical time-varying double-logistic form and two year-by-age AR1 alternatives. It documents convergence, age-composition diagnostics, selectivity patterns, and effects on spawning biomass for the corrected RTMB EBS pollock configuration.

Table of contents

1 Purpose

This analysis supports the fishery-selectivity discussion in the main RTMB report. It compares fitted alternatives using the same corrected model and data configuration, with attention to convergence, residual behavior, selectivity patterns, and assessment effects.

Return to the September 2026 model-development overview

This report compares four RTMB runs based on the corrected full-age BTS multinomial treatment that differ **only** in the fishery selectivity form.

- Base: `fishery_sel_form` = 0
- Comparators: `fishery_sel_form` = 2, 5, 6

The base and 2D AR1 outputs were generated by `R/run_fishery_selectivity_forms.R`. The time-age varying double-logistic output was generated by `R/run_double_logistic_experiments.R` and is now maintained with the comparison in the primary repository. Results are loaded from:

- analysis/output/corrected_full_age_bts/fishery_sel_forms (base and 2D AR1)
- analysis/output/corrected_full_age_bts/double_logistic_experiments (stage-2 double logistic)

Table 1: Model optimization summary by fishery selectivity form. The double-logistic objective is Laplace marginal; the others are joint NLL values. Code 0 indicates normal nlminb completion.

Fit summary

Form	Scenario	Objective	Type	Code	Max |gradient|
0	Base	7,075.2	Joint NLL	0	2.31×10^{-5}
2	Double-logistic (30% CV)	7,592.6	Laplace marginal	0	6.43×10^{-5}
5	2D AR1	6,883.3	Joint NLL	0	2.40×10^{-3}
6	2D AR1 (age-11 plus)	6,917.9	Joint NLL	0	3.71×10^{-3}

! How to use this page

Optimization and likelihood results should be considered together with age-composition fits, selectivity patterns, spawning biomass, convergence, and biological plausibility. A lower objective value alone is insufficient to select an assessment configuration.

i Plan Team takeaways

The 30% CV double-logistic fit passed the optimizer, gradient, positive-Hessian, and finite-standard-error checks. Its fishery OSA residual SDNR is 1.22; the fit improves the conditional fishery age-composition likelihood but still shows more residual dispersion than either AR1 alternative. These results support continued evaluation of time-varying fishery selectivity; they do not identify a replacement operational specification.

1.1 Forms included

Form	Implementation	Time variation	Role in this report
0: Base coefficients	Estimated log-selectivity coefficients by age	Deviations at configured change years	Bridge-compatible reference and current base form
2: Time-age varying double logistic	Ascending and descending logistic limbs	Annual random deviations around a shared mean curve, with fixed 30% process CV	Parsimonious dome-shaped sensitivity with substantial interannual change
5: 2D AR1	Latent log-selectivity surface over year and age	Separable AR1 correlation across neighboring years and ages	Flexible time-varying sensitivity
6: 2D AR1, age-11 plus	Latent log-selectivity surface over years and ages 1–11; ages 12–15 copy age 11	Separable AR1 correlation across neighboring years and ages 1–11	Flexible time-varying sensitivity with an explicit old-age plus group

The fishery forms are the focus of the fitted comparison. The model also uses fleet-specific selectivity structures for BTS and ATS; the relevant code and parameter mapping are summarized in the implementation appendix.

2 Fit Summary

An internal report-versus-optimizer check was also conducted. The values agree for the fixed-effect fits. For the double-logistic fit, the optimized objective is Laplace marginal whereas the reported total is the joint objective at the fitted random-effect modes, so their difference is not a fit improvement or an internal mismatch.

2.1 Catch biomass fit (observed vs predicted)

2.2 Per-scenario assessment reports

For deeper inspection of each scenario (catch/SSB/selectivity surfaces/age-comp residuals), see the scenario-specific assessment pages:

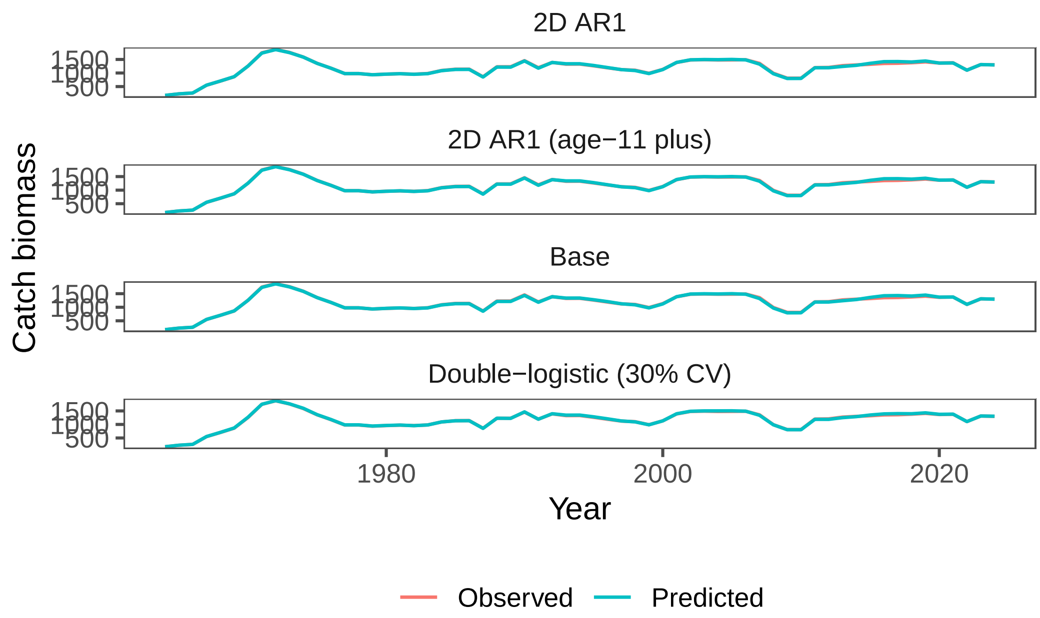

Figure 1: Observed vs predicted catch biomass by scenario (fishery).

Table 2: Age-composition NLL contributions (and deltas vs base) by fleet and scenario.

Fleet	Scenario	Age-composition NLL	Change from base
Fishery	Base	157.8	0.0
Fishery	Double-logistic (30% CV)	454.8	297.0
Fishery	2D AR1	97.2	-60.6
Fishery	2D AR1 (age-11 plus)	95.8	-62.0
BTS	Base	204.3	0.0
BTS	Double-logistic (30% CV)	249.2	44.9
BTS	2D AR1	202.2	-2.1
BTS	2D AR1 (age-11 plus)	198.8	-5.5
ATS	Base	30.1	0.0
ATS	Double-logistic (30% CV)	35.4	5.3
ATS	2D AR1	31.6	1.5
ATS	2D AR1 (age-11 plus)	31.2	1.1

- Base (form 0): `fishery_sel_form_0_assessment.html`
- Time-age varying double logistic (form 2): the stage-2 results documented on this page
- 2D AR1 (form 5): `fishery_sel_form_5_assessment.html`

3 Likelihood Component Comparison

3.1 Age-composition fit comparison

The model returns an age-composition negative log-likelihood contribution by fleet in `report$age_like`.

It also stores fishery observed (`oac_fsh`) and expected (`eac_fsh`) age-composition proportions by year×age in the saved run outputs.

3.2 One-step-ahead residual comparison

One-step-ahead (OSA) residuals provide a distribution-based check of the age-composition fits. The comparison uses `afscOSA` v0.0.1, multinomial residuals, the same observed compositions and effective sample sizes for every approach, and seed 202508. Each gear has one figure, with one column per fitted selectivity approach: base coefficients, the time-age varying double logistic

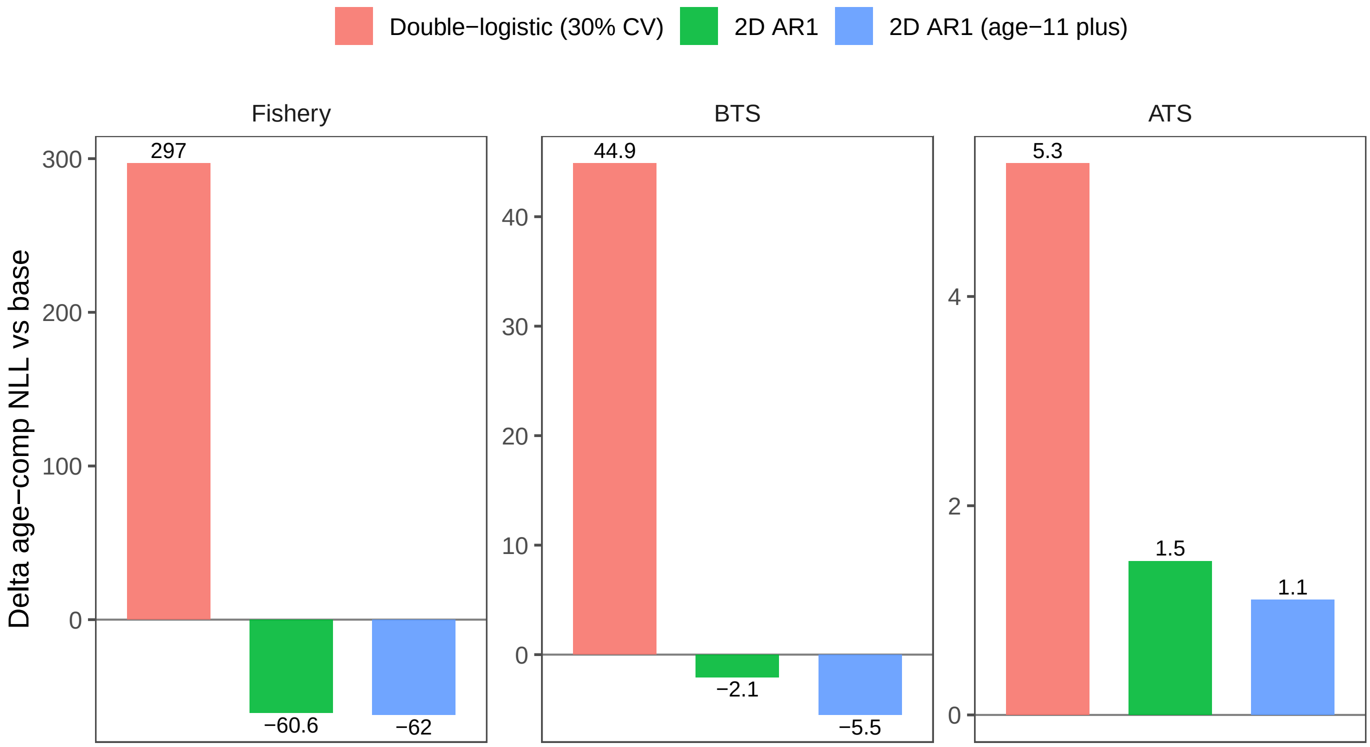

Figure 2: Change in age-composition NLL relative to the base scenario by fleet. Form 2 is the hierarchical double logistic with 30% process CV. These conditional fleet-specific likelihood contributions are comparable as data-fit diagnostics because they use the same observations and likelihood, but they are not marginal overall model objectives. Negative values indicate improved fit relative to base.

with 30% process CV, the full-age 2D AR1 surface, and the age-11-plus 2D AR1 surface. The stage-1 shared curve was only an initialization step and is not treated as a competing assessment configuration.

The acoustic-trawl survey calculation uses ages 2–15. Its 2020 composition is omitted because restricting the data to those ages produces zero rounded observations; this rule is applied identically to all four approaches. As in the standard `afscOSA` plotting method, residuals outside +/-6 are capped in the figures only so the main patterns remain visible; their uncapped finite values are retained in the summary table.

The fishery comparison is the clearest separation among methods (Figure 3). The time-age varying double logistic has SDNR 1.22 and 10% of residuals outside +/-2, while the base, full-age 2D AR1, and age-11-plus 2D AR1 approaches remain closer to the reference scale (SDNR 0.87, 0.81, and 0.8). Differences are smaller for BTS (Figure 4), while all four approaches give similar ATS residual distributions (Figure 5). These diagnostics favor the more flexible fishery selectivity forms; they do not by themselves establish which configuration should be used for assessment.

3.3 Age Composition Fits

These figures follow the observed-versus-expected age-composition layout used for figures 5-7 in the main assessment report. Points show the observed proportions; colored lines show the expected proportions from each fishery selectivity scenario.

The RTMB report object stores a set of likelihood (or penalty) components. The most relevant ones for this comparison are:

- `sel_like`, `sel_like_dev` (selectivity penalties)
- `age_like` (age composition likelihood)
- `cat_like` (catch fit)
- `cpue_like` (CPUE)

4 Fishery Selectivity Surfaces

Fishery selectivity is represented as a relative age pattern that multiplies annual fully selected fishing mortality. Let t index years, a index ages, and $x_a = a + 0.5$ denote the age value used in the selectivity functions. Each form is converted to log

Table 3: OSA residual summaries by gear and selectivity approach. The saved OSA metadata verifies that Form 2 is the hierarchical Stage 2 run with 30% process CV, random annual effects, and no old-age cap. SDNR is the standard deviation of the OSA residuals; values near 1 are consistent with the assumed residual scale. The last two columns give the percentage of absolute residuals exceeding 2 and 3.

Gear	Approach	n	SDNR	Mean	Abs. > 2 (%)	Abs. > 3 (%)
Fishery	Base coefficients	840	0.87	-0.06	3.1	0.6
Fishery	Double-logistic (30% CV)	840	1.22	-0.08	10.0	3.0
Fishery	2D AR1	840	0.81	-0.09	2.0	0.2
Fishery	2D AR1 (age-11 plus)	840	0.80	-0.04	2.3	0.2
BTS	Base coefficients	588	0.94	-0.05	3.7	0.7
BTS	Double-logistic (30% CV)	588	1.01	0.18	5.3	0.5
BTS	2D AR1	588	0.94	0.04	3.9	0.7
BTS	2D AR1 (age-11 plus)	588	0.93	-0.02	4.1	0.7
ATS	Base coefficients	234	0.89	0.03	5.1	0.4
ATS	Double-logistic (30% CV)	234	0.92	0.11	5.6	0.4
ATS	2D AR1	234	0.90	0.05	4.7	0.4
ATS	2D AR1 (age-11 plus)	234	0.90	0.04	4.7	0.4

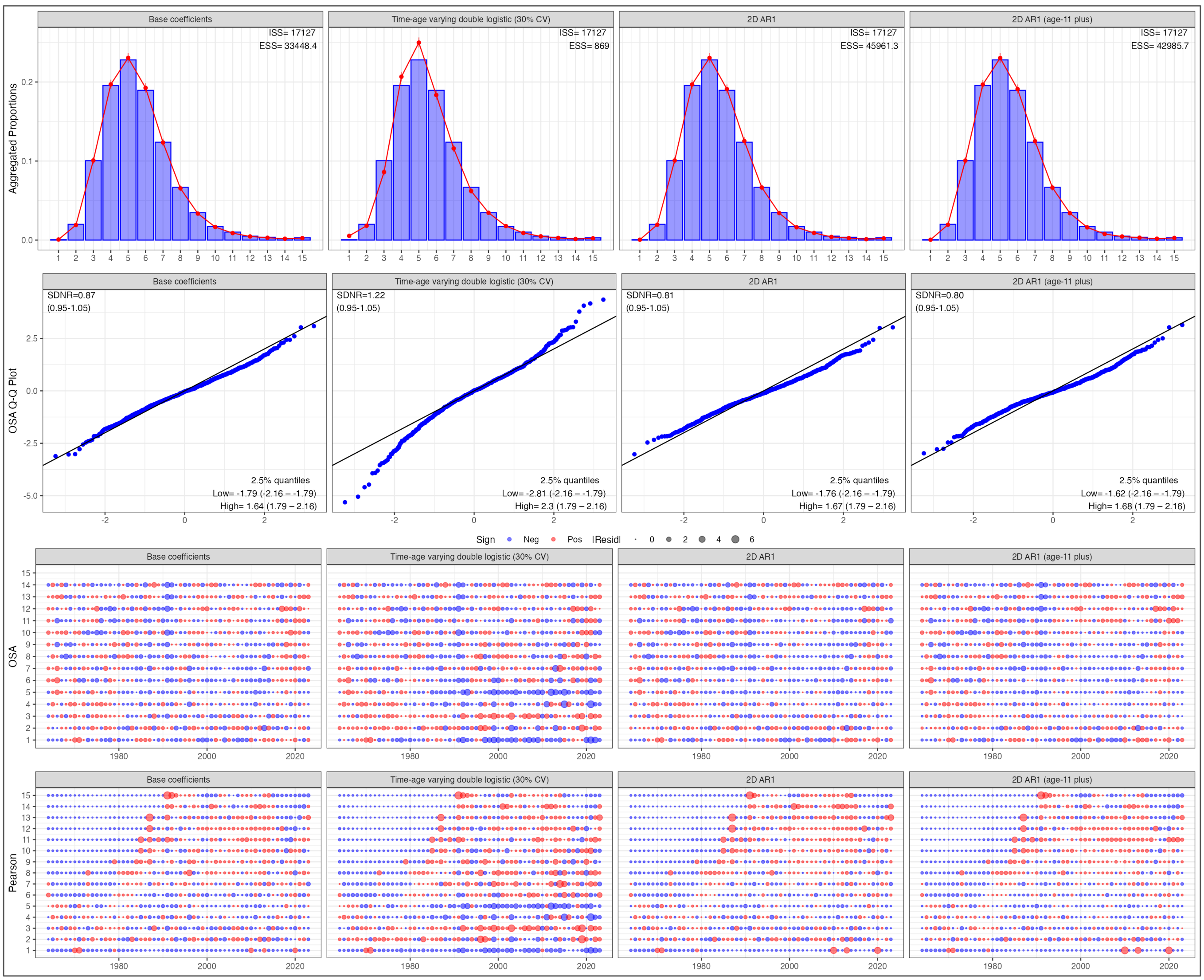

Figure 3: Fishery multinomial OSA comparison. Columns are the four fitted selectivity approaches; rows show aggregated fit, OSA Q-Q behavior, OSA residuals by age and year, and Pearson residuals. Values beyond +/-6 are capped in the figure only.

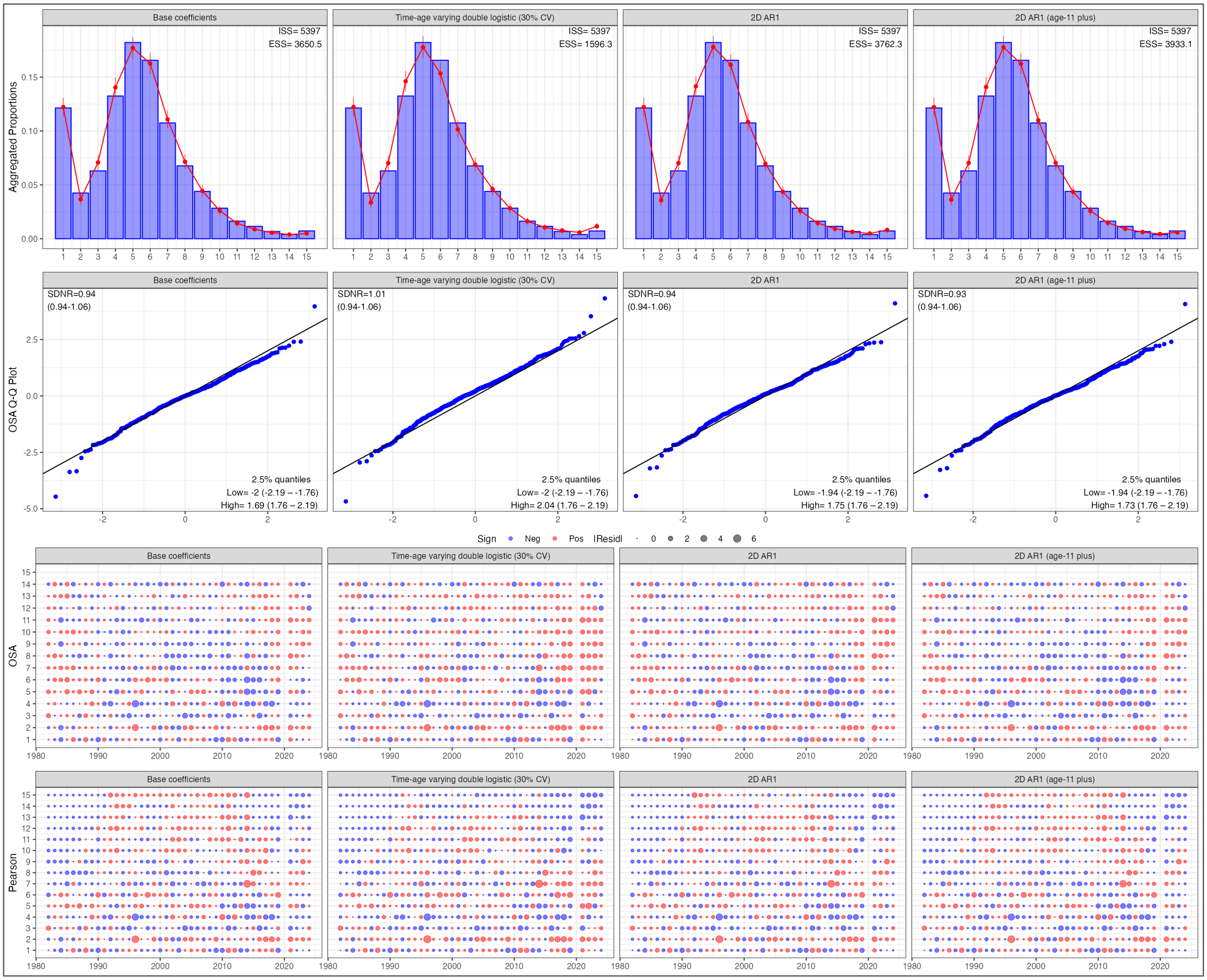

Figure 4: Bottom trawl survey multinomial OSA comparison, with one selectivity approach per column.

Table 4: Selected likelihood / penalty components returned in the RTMB report.

NLL components

fishery_sel_form	label	sel_like	sel_like_dev	age_like_total	cat_like	cpue_
0	Base (coefficients + change-year deviations)	17	192	392	3	
2	Time-age varying double logistic (30% CV)	0	86	739	1	
5	2D AR1 year×age (all years)	0	79	331	2	
6	2D AR1 ages 1–11; ages 12–15 fixed at age 11	0	121	326	2	

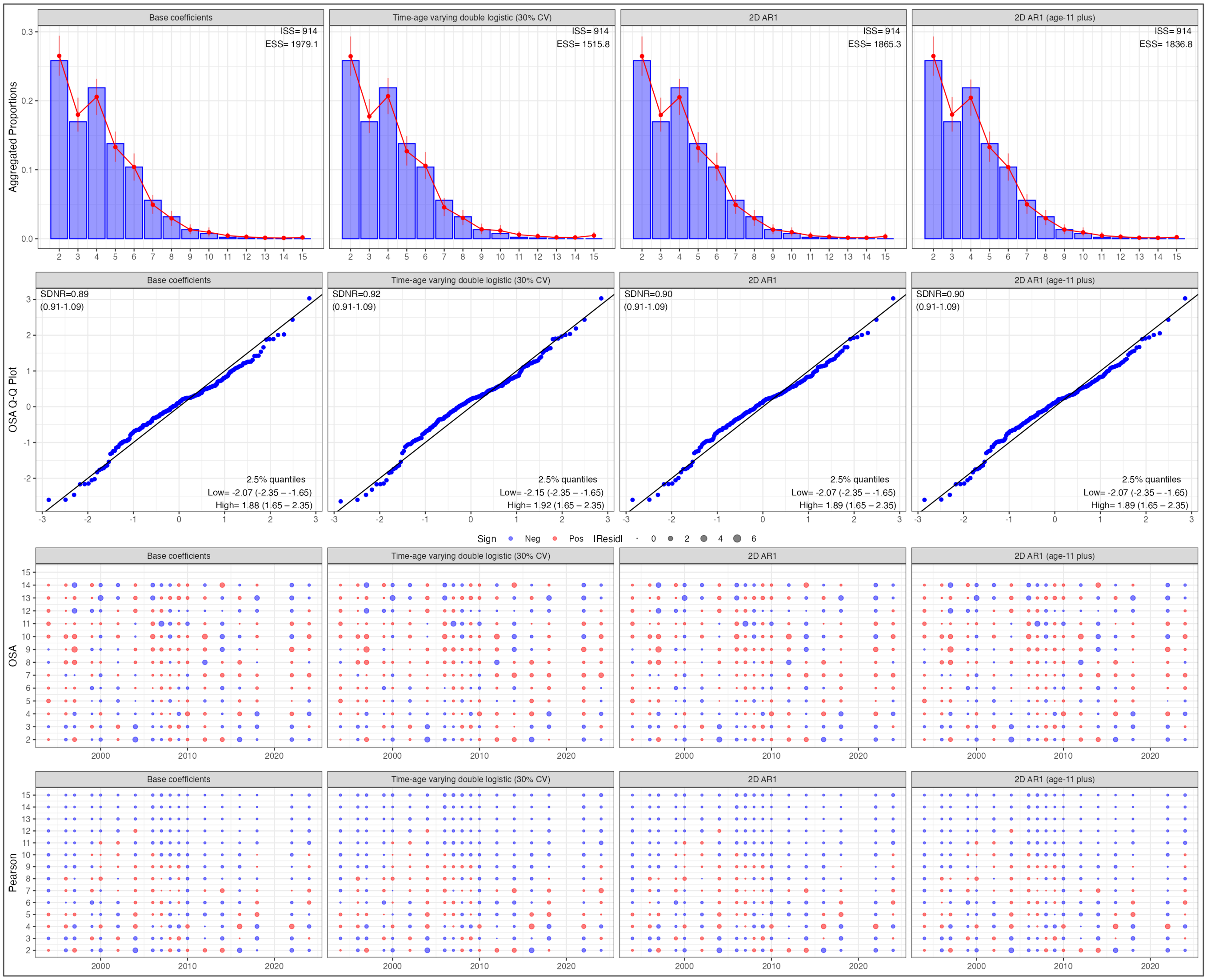

Figure 5: Acoustic-trawl survey multinomial OSA comparison for ages 2–15, with one selectivity approach per column. The 2020 composition has zero rounded observations after restricting to these ages and is omitted consistently.

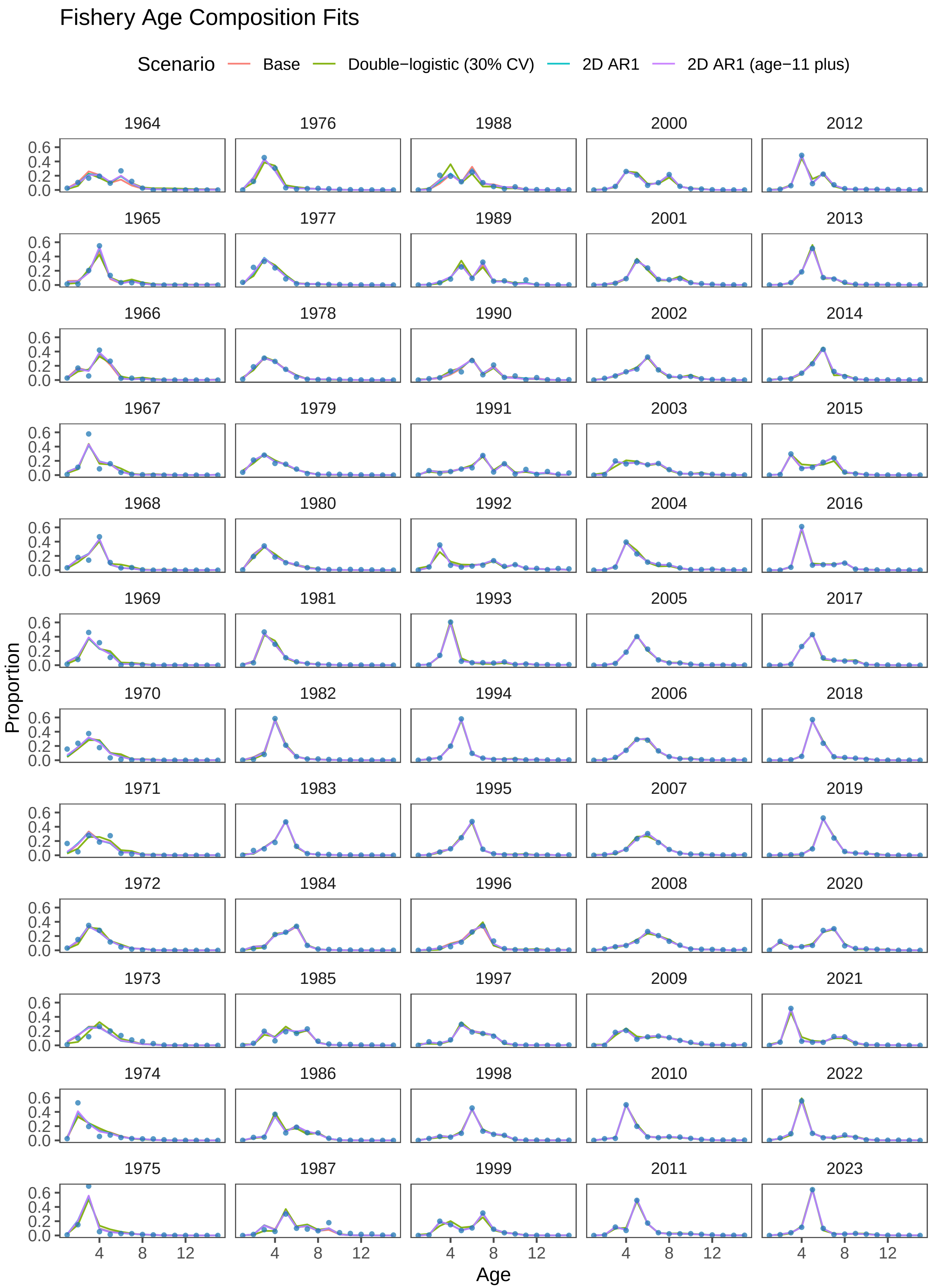

Figure 6: Observed and expected fishery age-composition patterns by fishery selectivity scenario, faceted by year. Points are observed proportions; lines are scenario-specific expected proportions.

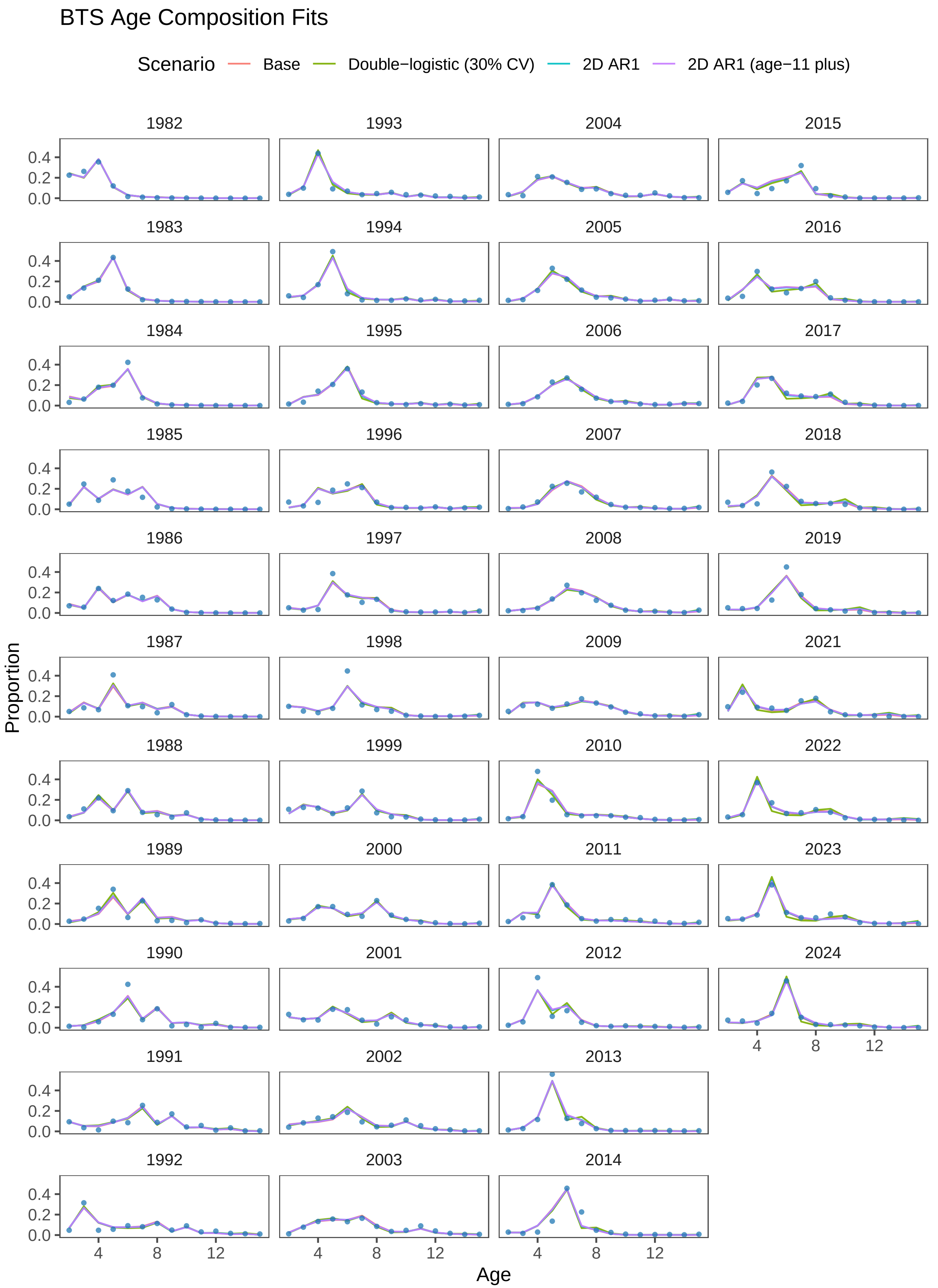

Figure 7: Observed and expected BTS age-composition patterns by fishery selectivity scenario, faceted by year. Points are observed proportions; lines are scenario-specific expected proportions.

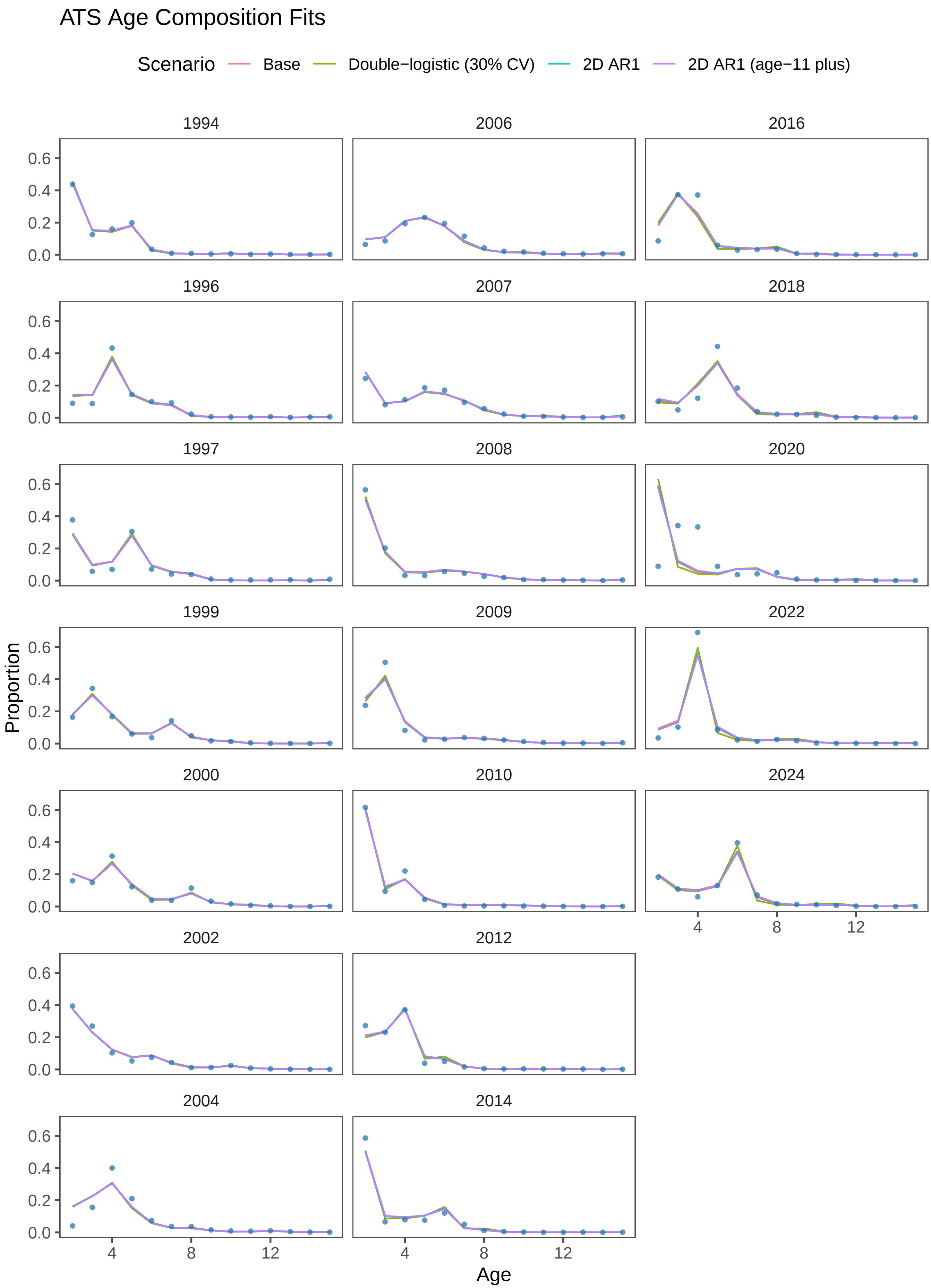

Figure 8: Observed and expected ATS age-composition patterns by fishery selectivity scenario, faceted by year. Points are observed proportions; lines are scenario-specific expected proportions.

selectivity $\ell_{t,a}$, centered within year, and exponentiated. The base and age-11-plus 2D AR1 runs constrain ages 11 and older to share the age-11 value:

$$\ell_{t,a} = \ell_{t,11}, \quad a \geq 11,$$

The time-age varying double logistic and full-age 2D AR1 do not impose this equality; their age effects are evaluated separately through age 15.

$$\ell^*_{t,a} = \ell_{t,a} - \log\left\{A^{-1}\sum_{a=1}^{A}\exp(\ell_{t,a})\right\}, \qquad s_{t,a} = \exp(\ell^*_{t,a}).$$

The centered selectivity $s_{t,a}$ enters the population dynamics as $F_{t,a} = F_t s_{t,a}$. The HTML edition includes an implementation appendix with code references for reproducibility.

4.1 Literature context

Sampson (2014) provides a useful framing for this comparison by separating contact selection, availability, and population-level fishery selectivity. That distinction matters here because the RTMB selectivity surface is not simply a gear-retention curve; it is the relative fishing mortality pattern implied by the combined fishery, fleet behavior, fish distribution, and assessment model structure. The paper also illustrates that population selectivity can be temporally variable and can take diverse shapes, which motivates comparing the base change-year coefficient form against more flexible alternatives.

Punt et al. (2014) focuses on the assessment-model problem: selectivity forms reduce variance when they are parsimonious, but they can bias assessment outputs when the assumed functional form is wrong. The paper reviews how residual diagnostics, information criteria, and simulation studies can be used to evaluate selectivity choices. This report follows that logic by comparing model likelihood components, age-composition fits, selectivity surfaces, and spawning biomass trajectories rather than interpreting one selectivity form from the objective value alone.

Martell and Stewart (2014) is the most direct precedent for the time-varying selectivity comparison. Martell and Stewart used simulation-estimation experiments to evaluate fixed, blocked, and continuously changing selectivity patterns and emphasized that both statistical fit and policy consequences should be evaluated. The base model here resembles a blocked approach through change-year deviations, whereas the 2D AR1 option is closer to a continuous time-varying surface that borrows strength across neighboring years and ages.

Butterworth et al. (2014) asks how much selectivity uncertainty affects management advice across a set of applied assessment and management-procedure examples. The study is a caution against treating selectivity sensitivity as automatically dominant: selectivity can strongly affect reference points and biomass scale in some cases, but other uncertainties may be equally or more influential. For this report, that means SSB and likelihood differences should be read as one part of a broader sensitivity analysis rather than as a complete basis for model choice.

Cadrin et al. (2016) connects assessment selectivity and catchability to field observations of fishing gear, survey gear, efficiency, and availability. The paper is relevant because assessment-estimated selectivity can absorb several processes that field studies may measure separately. For EBS pollock, any interpretation of the fishery selectivity surface should therefore distinguish model-based population selectivity from direct observations about gear performance, fishing behavior, or fish availability.

Vasilakopoulos et al. (2020) shifts the focus from estimating selectivity curves to operational metrics for management advice. The authors evaluate candidate metrics for sensitivity to selectivity changes and robustness to recruitment and fishing-mortality variation, identifying the ratio of fishing mortality on the first recruited age to average fishing mortality as a useful summary. That perspective suggests a practical extension of this comparison: after fitting alternative selectivity forms, summarize whether the implied exploitation pattern changes in a way that managers could track, not only whether the full selectivity surface looks different.

4.2 Form 0: base coefficients with change-year deviations

The base EBS pollock form follows the statistical approach developed by Butterworth et al. (2003), which placed catch-at-age fitting in a likelihood framework and allowed age-specific selectivity to change over time within a specified structure. That paper is the key methodological reference for the assessment's use of age coefficients with time-varying deviations.

The implemented base model uses estimated age coefficients plus deviations at configured fishery-selectivity change years:

$$\ell_{t,a} = c_a + d_{k(t),a},$$

where c_a is the baseline age coefficient and $d_{k(t),a}$ is the deviation for the change-year block $k(t)$. This form is flexible at the configured break points but does not estimate a continuous year-to-year process for fishery selectivity.

4.3 Form 2: time-age varying double logistic

The double-logistic curve is generated from positive parameters $p_{1,t}$, $p_{2,t}$, and $p_{3,t}$. The model estimates their logarithms, so $p_{j,t} = \exp(\theta_{t,j})$. The initial values on the natural scale are 1.5, 3, and 4. These parameters define the ascending and descending transition ages:

$$\gamma_{1,t} = p_{1,t} + p_{2,t}, \qquad \gamma_{2,t} = 2p_{1,t} + p_{2,t} + p_{3,t}.$$

The two logistic limbs are

$$u_{t,a} = \left[1 + \exp\left\{-\log(19)\frac{x_a - \gamma_{1,t}}{p_{1,t}}\right\}\right]^{-1},$$

$$v_{t,a} = 1 - \left[1 + \exp\left\{-\log(19)\frac{x_a - \gamma_{2,t}}{p_{3,t}}\right\}\right]^{-1},$$

Following the EBS pollock prior-elicitation formulation, the raw curve is

$$s_{t,a}^{\text{hard}} = \min\left\{1,\ u_{t,a}v_{t,a}0.95^{-2}\right\}.$$

The literal minimum has a kink where its two arguments meet. The RTMB implementation therefore uses the differentiable approximation

$$s_{t,a} = 1 - \frac{\log\left[1 + \exp\{\kappa(1 - r_{t,a})\}\right]}{\kappa} + \frac{\log\{1 + \exp(-\kappa)\}}{\kappa}, \qquad r_{t,a} = u_{t,a}v_{t,a}0.95^{-2}, \quad \kappa = 50.$$

This transform is zero at zero, approaches one smoothly, and closely follows the hard cap except immediately around $r_{t,a} = 1$. Define $b_{t,a} = \log(s_{t,a} + \epsilon)$.

It is then centered so arithmetic mean selectivity within each year is one:

$$\ell_{t,a} = b_{t,a} - \log\left\{\frac{1}{A}\sum_{a'=1}^{A}\exp(b_{t,a'})\right\}.$$

An earlier implementation subtracted $\max_a b_{t,a}$ before this centering. That subtraction was unnecessary: centering removes any common additive constant, so $\text{center}(b - \max b) = \text{center}(b)$ apart from numerical roundoff. It was removed because `max()` is not differentiable when the age attaining the maximum changes. The standard implementation then sets ages 11 and older equal to the age-11 value and centers the result again; the uncapped sensitivity below omits that equality constraint.

The original annual fixed-effect implementation estimated one log-scale triplet $\theta_t = (\log p_{1,t}, \log p_{2,t}, \log p_{3,t})$ for every year. It used shrinkage toward $\mu_0 = \log(1.5, 3, 4)$ and a first-difference penalty:

$$Q_{\text{fixed}} = \sum_t \sum_j \left(\frac{\theta_{t,j} - \mu_{0,j}}{\sigma_{p,j}}\right)^2 + \frac{1}{2}\sum_{t=2}^{T}\sum_j \left(\frac{\theta_{t,j} - \theta_{t-1,j}}{\sigma_{\Delta,j}}\right)^2,$$

where $\sigma_p = (0.75, 0.75, 0.45)$ and $\sigma_\Delta = (0.60, 0.60, 0.40)$. This is the 183-parameter formulation that failed the convergence checks; it is retained here to document that attempt, not as the formulation used for the corrected stage-2 results.

The successful staged formulation does not use the first-difference term above. Stage 1 estimates a single mean triplet μ shared across years, with

$$Q_1 = \sum_j \left(\frac{\mu_j - \mu_{0,j}}{\sigma_{p,j}}\right)^2.$$

Stage 2 defines the annual log-scale parameters as

$$\theta_t = \mu + \delta_t, \qquad \delta_{t,j} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_{\text{CV}}^2), \qquad \sigma_{\text{CV}} = \sqrt{\log(1 + 0.30^2)} = 0.2936.$$

Ignoring normalizing constants, the corresponding penalty is

Table 5: Selected 30% CV hierarchical option-2 fit. The optimizer objective is Laplace marginal; the evaluated total is the conditional joint NLL at the fitted random-effect modes.

Configuration	Marginal objective	Conditional joint NLL	Optimizer code	Maximum gradient	Positive
30% CV annual random effects	7592.6	7305.6	0	6.4e-05	TRU

$$Q_2 = \sum_j \left(\frac{\mu_j - \mu_{0,j}}{\sigma_{p,j}} \right)^2 + \frac{1}{2\sigma_{\mathrm{CV}}^2} \sum_t \sum_j \delta_{t,j}^2.$$

The code evaluates the random-effect contribution with the full normal log density and integrates the annual deviations using RTMB's Laplace approximation. The current section-2.6 implementation retains the configured across-age curvature penalty but omits the historical old-age shape penalty. The smaller prior scale for the third mean parameter provides stronger shrinkage on the descending limb, which is otherwise weakly informed by older ages.

4.3.1 Development history and reproducibility

The initial option-2 fit should be treated as a convergence diagnosis rather than as a candidate assessment result. It estimated 61 annual parameter triplets, or 183 double-logistic parameters. The saved run ended with optimizer code 1 (`false convergence (8)`) and a maximum absolute gradient of 1.8366. Its objective value was 8299.2 and agreed with the independently evaluated reported total to numerical precision. Thus, the objective was evaluated consistently, but the parameter estimates do not yet satisfy the convergence standard needed for biological or assessment interpretation.

> **!** Result of the staged tests
>
> The originally implemented annual fixed-effect option failed, but the corrected two-stage section-2.6 formulation converged. A shared three-parameter curve was estimated first, followed by zero-centered annual random effects with a fixed 30% CV. The selected stage-2 fit had a marginal gradient of 0.000064. After two unused parameters were removed, the stage-2 Hessian was positive definite and finite standard errors were obtained.

The selectable-form implementation and the exact September 2025 ADMB–RTMB bridge now coexist on `main`, with the bridge behavior protected by regression tests. The former experiment branch remains on GitHub as a recovery point, but a separate worktree is no longer needed to reproduce the accepted fit.

4.3.1.1 Section-2.6 formulation without the age-11-and-older equality

The preferred option-2 experiment now follows the three-parameter function on the selectivity prior-elicitation page: the product of the ascending and descending limbs is smoothly capped near one, ages 11–15 are not forced to share the age-11 value, and the associated nonsmooth old-age shape penalty is omitted.

The selected 30% CV random-effects fit returned optimizer code 0, a maximum gradient of 0.000064, a positive-definite Hessian with minimum eigenvalue 0.1354, and finite standard errors. Its conditional joint NLL was 7305.6 and its fishery age-composition contribution was 454.8.

The process-CV ladder started from the previously accepted 5% fit and increased the allowance through 7.5%, 10%, 12.5%, 15%, 20%, and 30%. Every rung returned optimizer code 0, a maximum gradient below 0.0001, a positive Hessian, and finite standard errors. At 30%, the mean across-age temporal standard deviation of selectivity was 0.298, compared with 0.068 at 5%; the largest age-specific temporal standard deviation increased from 0.198 to 0.602. Fishery age-composition NLL declined from 790.3 at 5% to 454.8 at 30%. The gains persist through 30%, although the OSA diagnostics show that the double-logistic family still leaves more aggregate structure than the base and 2D AR1 forms.

This is time-and-age variation through a parsimonious curve rather than a free year-by-age surface: the three annual parameters move the ascending limb, peak, and descending limb, thereby changing selectivity across ages each year. There is no additional independent random effect for every age; that more flexible structure is represented by the 2D AR1 alternative.

The stage-2 mean parameters were $p_1 = 1.40$, $p_2 = 2.86$, and $p_3 = 5.98$. Selectivity declined smoothly after age 11; in 2024 its values at ages 11–15 were 1.169, 0.961, 0.753, 0.564, and 0.406 after within-year centering.

The earlier zero-eigenvalue warning was not created by these three selectivity parameters. Two unused temperature-residual parameters, `resid_temp_x1` and `resid_temp_x2`, had inadvertently remained in the experimental parameter vector even though neither appears in the objective. Perturbing either from -10 to +10 changed the objective by exactly zero. The 623-parameter Hessian therefore had two numerical-zero eigenvalues and no substantively negative eigenvalues. After mapping those unused

Table 6: Option-2 comparison. The current section-2.6 implementation uses a smooth approximation to the cap at one, does not impose the age-11-and-older equality, and maps two unused temperature-residual parameters out of the optimization. Earlier capped rows are retained only as historical context. Stage-2 objectives are Laplace marginal objectives and should be compared only with other stage-2 rows.

Configuration	Stage	Objective	Evaluated total	Fishery age NLL	Maximum gradient
Earlier ages-11+ formulation	Shared curve	8659.1	8659.1	1220.6	0.00072
Earlier ages-11+ formulation	Annual random effects	8728.2	8293.4	790.3	0.00027
Smooth section-2.6 formulation	Shared curve	8653.0	8653.0	1222.7	0.00045
Smooth section-2.6 formulation	Annual random effects	8482.6	8173.4	498.8	0.00007

Table 7: Selected option-2 convergence experiments. Objective is the minimized total negative log likelihood; max gradient is the largest absolute element of the objective gradient; a positive Hessian is required for a locally stable solution.

Scenario	Objective	Optimizer code	Max gradient	At bounds	Positive Hessian
Original reproduction	8299.2	1	1.8366	0	FALSE
Pooled three-parameter fit	9101.7	0	0.0005	0	FALSE
Annual fit from pooled start	8299.0	1	1.7937	0	FALSE
Proposed stronger prior	8383.5	1	14.5524	0	FALSE
Proposed scales × 0.75	8453.8	1	15.0113	0	FALSE
Proposed scales × 1.25	8338.6	1	2.7645	0	FALSE
Proposed scales × 1.50	8306.0	1	4.9645	0	FALSE

parameters out, the current 621-parameter stage-2 Hessian is positive definite: its minimum eigenvalue is 0.1349 and finite standard errors are obtained.

The earlier smallest Hessian eigenvalues, about -1.2×10^{-16} and -1.9×10^{-16}, were numerical zeros rather than meaningful negative curvature. They corresponded to the unused `resid_temp_x1` and `resid_temp_x2` parameters. Once those were mapped out, the current stage-2 minimum eigenvalue was 0.1349 and all standard errors were finite.

4.3.1.2 Step 7: decide whether a fit is usable

Each experiment should pass all of the following checks before its assessment results are compared:

- optimizer convergence code 0;
- maximum absolute gradient below 0.001, with 0.0001 preferred;
- no important parameters at or very near their bounds;
- positive-definite Hessian and finite standard errors;
- objective value matching the independently evaluated reported total;
- stable results when optimization is restarted from the fitted values;
- similar solutions across reasonable starting points;
- smooth, biologically plausible annual parameter paths and selectivity curves;
- understandable changes in likelihood components, age-composition residuals, spawning biomass, recruitment, and fishing mortality; and
- conclusions that remain similar across modest changes in prior strength.

The lower objective value is only one diagnostic. Convergence, stability, biological plausibility, fit to the observed compositions, and effects on assessment quantities should be considered together.

Stage	Configuration	Status	Result
Original fit	Original starts and scales	Completed	Code 1; gradient 1.8366; objective 8299.2; not usable for interpretation
Reproduction and restart	Unchanged model, then restart from saved solution	Completed	Reproduced false convergence; restart did not improve it
Common-curve initialization	Pooled triplet followed by annual fit	Completed	Pooled fit passed the gradient criterion; annual fit failed
Modestly informative scales	$\sigma_p = (0.50, 0.50, 0.35)$, $\sigma_\Delta = (0.25, 0.25, 0.20)$	Completed	False convergence; gradient 14.5524

Stage	Configuration	Status	Result
Prior sensitivity	Scale multipliers 0.75, 1.25, and 1.50	Completed	All three failed convergence and Hessian checks
Multiple starts	10 original-prior and 10 proposed-prior perturbations	Completed	All 20 returned false convergence
Preliminary random-effects sensitivity	Existing absolute annual matrix integrated directly	Completed	Failed in RTMB's inner Newton optimization; this was not the intended hierarchy
Corrected stage 1	Smooth section-2.6 curve shared across all years	Completed	Code 0; gradient 0.00045; positive Hessian
Corrected stage 2	Smooth section-2.6 mean triplet plus annual random deviations with 30% CV	Completed	Code 0; marginal gradient 0.000064
Uncertainty check	Hessian and standard errors after mapping unused parameters out	Completed	Minimum eigenvalue 0.1354; finite standard errors obtained

Taken together, the tests reject the 183-parameter annual fixed-effect version but support continued evaluation of the smooth section-2.6 30% CV random-effects model. The next step is to compare its residuals and assessment trajectories with the base and 2D AR1 alternatives.

> **i** Double-logistic estimability
>
> The failed run described above estimated 183 unrelated annual fixed effects. It has been superseded by the two-stage hierarchical fit shown here. In the selected Stage 2 sensitivity, the annual deviations are random effects with a fixed 30% process CV around the shared three-parameter mean curve. This fit passed the optimizer, gradient, Hessian, and finite-standard-error checks.

4.4 Forms 5 and 6: 2D AR1 time-varying surfaces

The full-age 2D AR1 form estimates a year-by-age latent field $z_{t,a}$ for ages 1–15. The age-11-plus configuration estimates the same field only for ages 1–11 and deterministically sets ages 12–15 equal to age 11 before within-year centering. Before centering, the estimated-age log selectivity is

$$\ell_{t,a} = z_{t,a},$$

The latent field is penalized with a separable AR1 process over year and its estimated ages:

$$z \sim \mathrm{AR1}_{\mathrm{year}}(\rho_y) \otimes \mathrm{AR1}_{\mathrm{age}}(\rho_a),$$

with fitted correlations reported in the caption for Figure 10. The scenario recipe down-weights the AR1 penalty relative to the default, allowing more interannual variation while still borrowing strength across neighboring years and ages.

5 Spawning Biomass Comparison

Butterworth, D. S., J. N. Ianelli, and R. Hilborn. 2003. "A Statistical Model for Stock Assessment of Southern Bluefin Tuna with Temporal Changes in Selectivity." *African Journal of Marine Science* 25 (1): 331–61. https://doi.org/10.2989/18142320309504021.

Butterworth, Doug S., Rebecca A. Rademeyer, Anabela Brandão, Helena F. Geromont, and Susan J. Johnston. 2014. "Does Selectivity Matter? A Fisheries Management Perspective." *Fisheries Research* 158: 194–204. https://doi.org/10.1016/j.fishres.2014.02.004.

Cadrin, Steven X., Gregory R. DeCelles, and David Reid. 2016. "Informing Fishery Assessment and Management with Field Observations of Selectivity and Efficiency." *Fisheries Research* 184: 9–17. https://doi.org/10.1016/j.fishres.2015.08.027.

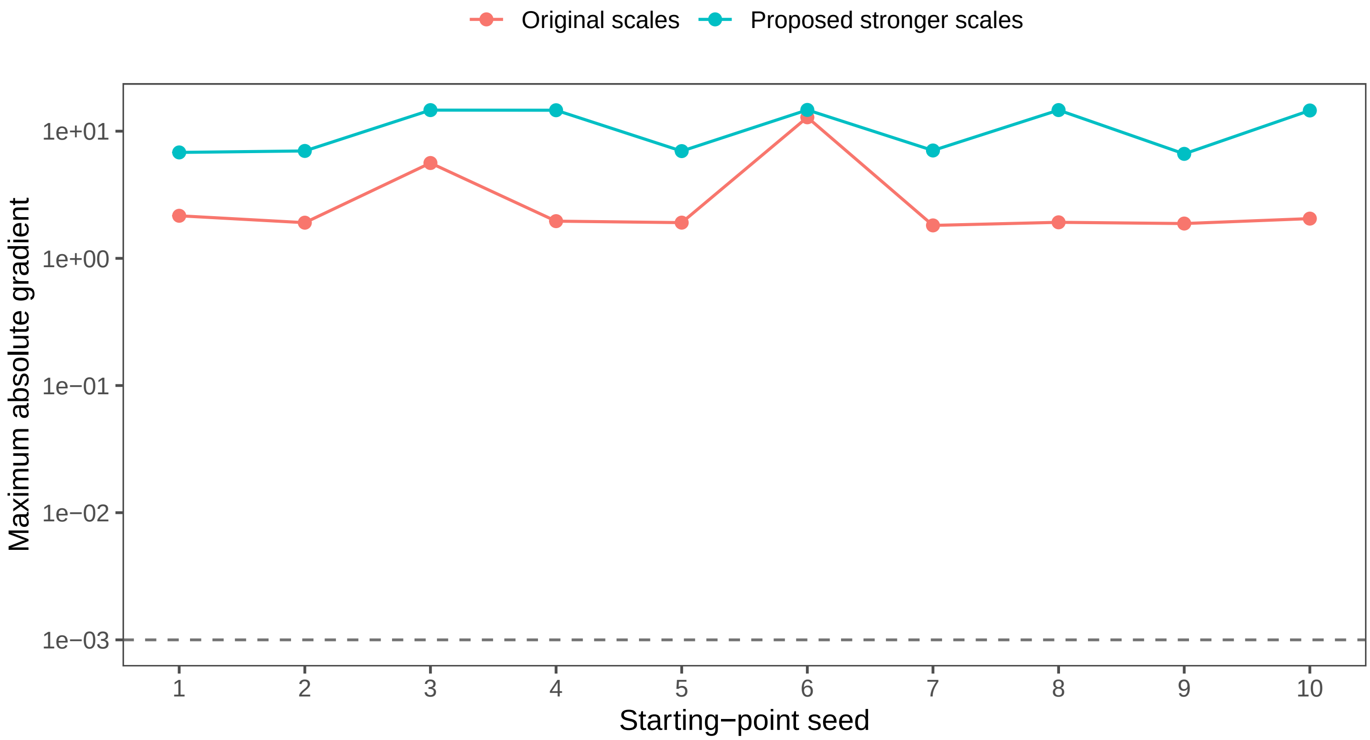

Figure 9: Maximum absolute gradients from 10 perturbed starts under each regularization. The dashed line marks the 0.001 acceptance threshold; every annual fit remained well above it.

Table 8: SSB differences relative to base run at key years.

SSB deltas vs base

label	fishery_sel_form	year	SSB	delta_SSB	pct_delta
Base (coefficients + change-year deviations)	0	2000	2791.918	0.0000	0.0000
Time-age varying double logistic (30% CV)	2	2000	3251.809	459.8914	16.4722
2D AR1 year×age (all years)	5	2000	2914.053	122.1349	4.3746
2D AR1 ages 1–11; ages 12–15 fixed at age 11	6	2000	2845.919	54.0014	1.9342
Base (coefficients + change-year deviations)	0	2010	1626.073	0.0000	0.0000
Time-age varying double logistic (30% CV)	2	2010	2135.417	509.3437	31.3235
2D AR1 year×age (all years)	5	2010	1736.572	110.4982	6.7954
2D AR1 ages 1–11; ages 12–15 fixed at age 11	6	2010	1667.581	41.5081	2.5527
Base (coefficients + change-year deviations)	0	2020	2381.126	0.0000	0.0000
Time-age varying double logistic (30% CV)	2	2020	3041.561	660.4354	27.7363
2D AR1 year×age (all years)	5	2020	2592.103	210.9767	8.8604
2D AR1 ages 1–11; ages 12–15 fixed at age 11	6	2020	2545.285	164.1591	6.8942
Base (coefficients + change-year deviations)	0	2023	3238.676	0.0000	0.0000
Time-age varying double logistic (30% CV)	2	2023	3754.709	516.0322	15.9334
2D AR1 year×age (all years)	5	2023	3427.318	188.6413	5.8246
2D AR1 ages 1–11; ages 12–15 fixed at age 11	6	2023	3328.523	89.8464	2.7742
Base (coefficients + change-year deviations)	0	2024	3269.154	0.0000	0.0000
Time-age varying double logistic (30% CV)	2	2024	3681.787	412.6333	12.6220
2D AR1 year×age (all years)	5	2024	3433.442	164.2887	5.0254
2D AR1 ages 1–11; ages 12–15 fixed at age 11	6	2024	3327.225	58.0716	1.7763

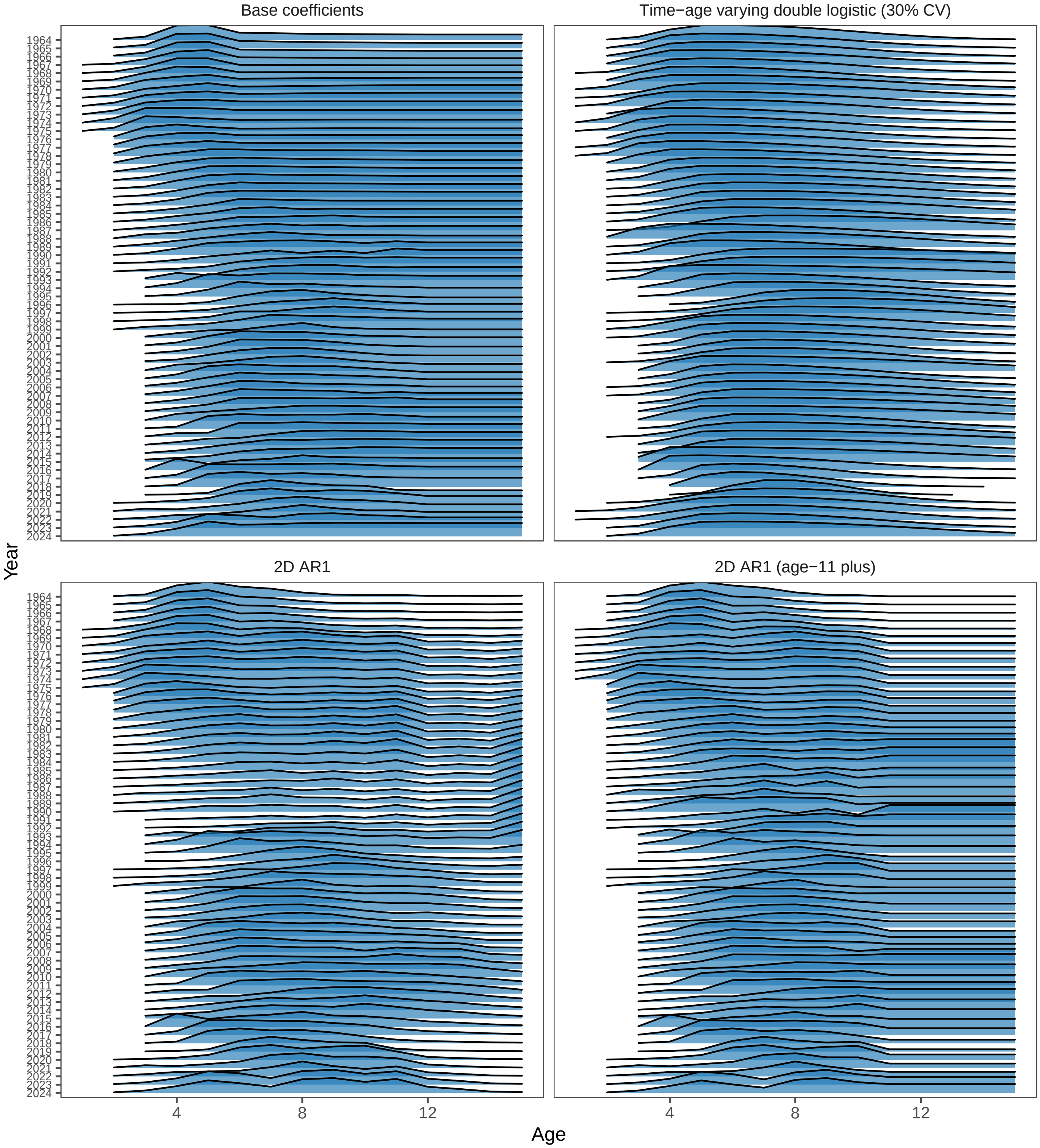

Figure 10: Fishery selectivity through time (ggridges), normalized within year. The time-age varying double logistic is the stage-2 fit: annual random deviations around a shared mean curve with 30% process CV, a differentiable cap near one, and no age-11-and-older equality. Stage 1 was used only to initialize this fit and is not displayed. The stage-2 fit passed the optimizer, gradient, positive-Hessian, and finite-standard-error checks. For the 2D AR1 form, fitted correlations are rho_y = 0.990 and rho_a = 0.366.

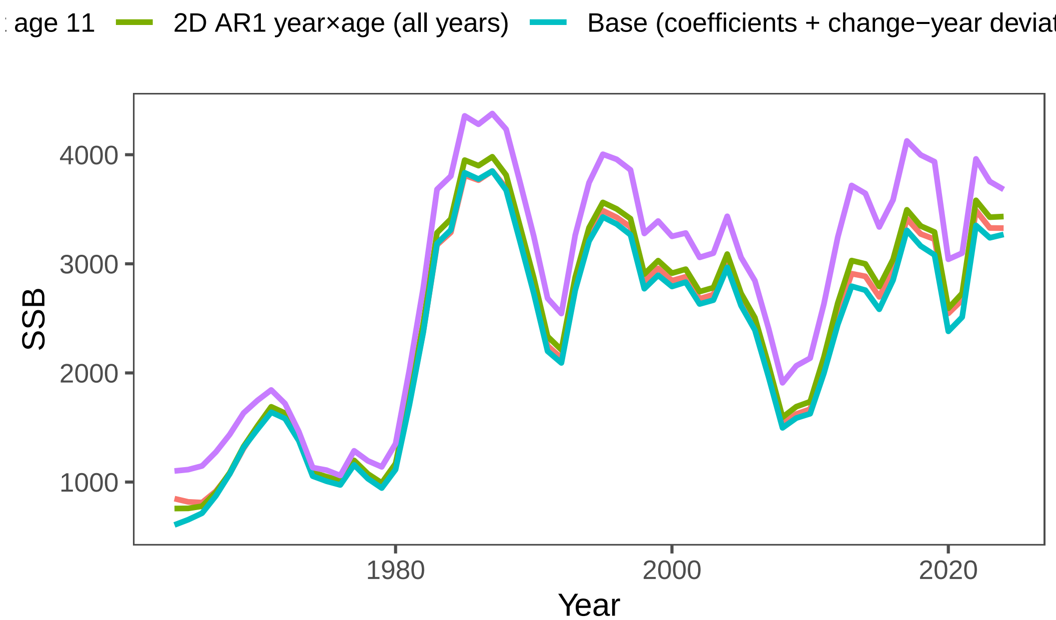

Figure 11: Spawning biomass (SSB) trajectories by fishery selectivity form.

Martell, S., and I. Stewart. 2014. "Towards Defining Good Practices for Modeling Time-Varying Selectivity." *Fisheries Research* 158: 84–95. https://doi.org/10.1016/j.fishres.2013.11.001.

Punt, André E., Felipe Hurtado-Ferro, and Athol R. Whitten. 2014. "Model Selection for Selectivity in Fisheries Stock Assessments." *Fisheries Research* 158: 124–34. https://doi.org/10.1016/j.fishres.2013.06.003.

Sampson, David B. 2014. "Fishery Selection and Its Relevance to Stock Assessment and Fishery Management." *Fisheries Research* 158: 5–14. https://doi.org/10.1016/j.fishres.2013.10.004.

Vasilakopoulos, Paraskevas, Ernesto Jardim, Christoph Konrad, et al. 2020. "Selectivity Metrics for Fisheries Management and Advice." *Fish and Fisheries* 21: 621–38. https://doi.org/10.1111/faf.12451.