Assessment model, biological reference points and projection development for Georges Bank yellowtail flounder

Alex Hansell[1], Jessie Kittel[2], Cole Carrano[2], Steve Cadrin[2]

1. Northeast Fisheries Science Center, Woods Hole, MA, USA
2. University of Massachusetts Dartmouth, School for Marine Science and Technology, New Bedford MA, USA

Abstract:

This working paper documents stock assessment model, biological reference points and projections explored during the 2024 yellowtail flounder research track stock assessment of Georges Bank yellowtail flounder. The research track working group used the Woods Hole Assessment Model (WHAM) to complete this work. The assessment incorporated combined US and Canadian fisheries data. The model was fit to US and Canadian trawl survey data. Model development explored alternative statistical distributions for age composition, random effects on recruitment, selectivity, numbers at age and catchability. Following recommendations from ToR 1, environmental covariates were explored as modulates of recruitment and natural mortality. A combination of stepwise selection and model diagnostics were used to choose the optimal model. The optimal model included random effects on commercial fleet selectivity and numbers at age (i.e., cohort survival process). Additionally, the model used a bottom water temperature covariate to help estimate deviations from a Beverton-Holt stock recruit relationship. A changepoint analysis on the bottom water temperature time series was used to determine recent conditions (2009 – 2022). MSY reference points and projections were calculated based on these recent conditions and an average of the most recent two years for weight at age.

Introduction:

This working paper documents stock assessment models, biological reference points and projections explored during the 2024 Yellowtail Flounder Research Track stock assessment. Historically, the Georges Bank stock of yellowtail flounder supported a productive fishery (Figure 1) and was one of the three principal groundfish stocks, with cod and haddock. The fishery is data rich with catch history starting in the 1930's and scientific sampling of age structure starting in the 1970s (Legault & Cadrin, 2014; TRAC, 2023). The stock has a long history of overfishing, with the stock being depleted in the 1990s, but then recovering in the early 2000s (Stone et al., 2004). However, in recent years, stock size has declined again, but this time, it has remained low despite reductions in fishing (TRAC, 2023).

Until this year, the stock was assessed as part of the Transboundary Resources Assessment Committee (TRAC) where scientists from the US and Canada worked together annually to update and revise the assessment to determine best available science (TRAC, 2023). Several analytical models have been used to assess the stock, with the primary platform being a virtual population analysis (VPA). However, over the course of multiple assessments, the retrospective patterns became larger, which ultimately resulted in the assessment being rejected in 2014. In response, an empirical benchmark assessment was developed (O'Brien and Clark, 2014). In 2021, a revised empirical approach called the limiter was developed (TRAC, 2021). The Limiter is based on an average area-swept biomass from the three surveys that cover the stock area (Northeast Fisheries Science Center (NEFSC) spring, NEFSC fall and Fisheries and Oceans Canada spring). The Limiter is designed to not chase noise in the surveys and produces a constant catch advice if the average survey biomass falls within certain bounds. The Limiter does not produce biological reference points or projections. However, the state of the resource is considered to be low and in poor condition (TRAC, 2023). Peer reviewers from past assessments have recommended re-examination of analytical assessment models for this stock to integrate all available information (TRAC, pers. comm.).

This working paper documents the yellowtail flounder research track working group's development of an analytical assessment model, reference points and projections. All model development applied the Woods Hole Assessment Model (WHAM), which is an age-structured model that allows for the inclusion of process error and environmental covariates (Stock and Miller, 2021). The working group hypothesized that the flexibility provided by WHAM would help to reduce diagnostics problems that led to previous analytical models being rejected.

Methods:

The terminal year of the assessment was 2022 and data inputs included: a combined fleet (USA and Canadian); catch rates from the Northeast Fisheries Science Center (NEFSC) fall and spring surveys as well as the Department of Oceans Canada (DFO) spring survey (Figure 2-4). Commercial catch at age data were available from 1973-2022 (Figure 5). Catch at age data were available since 1973 for the NEFSC surveys and since 1986 for the DFO survey (Figure 6-8). For all model runs, the maturity ogive was time invariant, while weight-at-age was time-varying (Figure 9-10). Alternative assumptions about natural mortality were explored during model fitting.

<u>Woods Hole Assessment Model (WHAM):</u>

Model exploration followed the process document outlined by the research track working group. Model exploration proceeded in an order of: 1) age distributions; 2) recruitment assumptions; 3) time varying selectivity; 4) full state-space models; and 5) environmental covariates. WHAM allows for 10 statistical distributions for age-composition and all were explored. Input effective sample sizes were assumed to be 200 for all data sources and years. High input effective sample sizes were needed for the Dirichlet age composition distributions.

For recruitment estimation, model runs using annual deviations from the time series mean estimated as fixed effects, IID or AR1_Y were explored. Beverton-Holt stock recruit relationships were also explored with annual deviations estimated as IID or Ar1_Y random effects. For selectivity, both age-based and logistic assumptions were examined for the commercial fleet and all three surveys. To account for changes in commercial selectivity, alternative selectivity periods were explored that accounted for fisheries regulations (1973-1993, 1994-2009) and to account for changes in fishing behavior (2010-2022, annual catch limits and primarily bycatch). Alternative process error assumptions (IID, AR1_Y, 2DAR1) were also explored for time varying fleet selectivity. All available process error correlation structures were explored for full state space models. Environmental covariates identified by the working group from ToR 1 (bottom water temperature and Atlantic Multi-decadal Oscillation Index) were explored as covariates on recruitment and natural mortality. For recruitment, environmental covariates were explored for annual deviations from both mean recruitment and the Beverton-Holt stock recruit relationship. On the Beverton-Holt stock recruit relationship the covariates were explored as influencing density independent mortality (controlling) or carrying capacity (limiting; Stock & Miller, 2021). These environmental covariates were also explored to estimate annual process errors for time varying natural mortality. Both constant and age varying natural mortalities were explored during model fitting. The assessment also attempted to freely estimate natural mortality and different process error assumptions were explored that allowed natural mortality to vary with time.

For reference points, MSY and MSY-proxy reference points were tested. Inputs for reference points and projections are: weights-at-age, natural mortality, fleet selectivity, recruitment and maturity. Typically, an average of the most recent years is assumed for each of these values and to calculate reference points and projections. We explored the sensitivity of using alternative time periods (2 year average,5 year average, time series average). Natural mortality and recruitment were assumed to be constant. Root mean squared error was used to determine the optimal number of years to calculate average weight at age for use in the reference points and projections. Changepoint analyses of the bottom water temperature time series tested if conditions have changed and are influencing recruitment. Changepoint analyses explored using mean and linear relationship to identify changepoints. Analyses also explored the number of changepoints. Sensitivity analyses examined how to treat environmental covariates in the projection years (e.g., average, linear, AR1_Y).

Results:

A total of 165 WHAM models were completed for the Georges Bank yellowtail flounder stock. All assessment model runs and projections are available on GitHub (https://github.com/achansell/yellowtail_gbk_wham).

<u>Woods Hole Assessment Model (WHAM):</u>

Model runs exploring alternative statistical distributions for age compositions supported using the logistic-normal-pool0 for the commercial fleet and all three surveys. The logistic-normal-pool-0 had low AIC, few residual patterns and small retrospective patterns (Table 1). Initial runs with fixed effects for recruitment deviations had the lowest AIC and similar SSB and F retrospectives as runs that included process error (Table 2). Initial runs exploring time vary selectivity on the commercial fleet did not support the inclusion or process error or blocking. Patterns in selectivity were not realistic and led to higher retrospective patterns (Table 3). However, in the final model configuration, an AR1_Y process error was used to estimate time varying selectivity in the commercial fleet. The inclusion of this process error improved residual diagnostics and led to realistic time varying selectivity that captured observed changes in the fleet (e.g., switch to annual catch limits and primarily bycatch in the 2010s).

In total, nine runs were explored to examine process error on numbers at age (i.e., the cohort survival process). AIC supported the use of either 2DAR1 or AR1_Y (coupled and decoupled) process error. Retrospective patterns and self-tests supported using IID process error (Table 4-5). Because of diagnostic tradeoffs among alternative runs, all four process errors were carried forward for model development.

Model runs with IID process error and fit to the bottom temperature covariate on recruitment had good convergence. AIC supported using a Beverton-Holt stock recruit relationship with IID deviations informed by bottom water temperature. Retrospective patterns were better when the environmental relationship influenced the stock recruit relationship by assuming density independent mortality (i.e., controlling effect of bottom temperature on recruitment). Allowing for the covariate to follow a random walk had slightly better retrospective patterns, but this method prevented using an AR1 process in the projection period. Fitting to the bottom water temperature covariate also reduced retrospective patterns in recruitment. Similar results were observed for the AR1_Y process error; but the AR1_Y decoupled and 2DAR1 had lower convergence rates (Table 6). Model runs exploring the Atlantic Multidecadal Oscillation (AMO) on recruitment had similar results as runs fitting to the bottom temperature time series (Table 7-10).

Model diagnostics were not improved by using environmental covariates (bottom temperature or AMO) to help estimate time varying natural mortality. Similarly, using process error or allowing the model to freely estimate natural mortality did not improve model diagnostics. Many of the models that explored alternative natural mortality assumptions had convergence problems (Tables 11-15). Thus, a constant value for M of 0.4 was used in subsequent runs (Cadrin, 2024).

The working group decided to move forward with the IID process error on numbers at age and include the bottom water temperature time series as a covariate on the Beverton-Holt stock recruit relationship. The bottom temperature time series followed an AR1 process because this could be explored in projections. A controlling process was used because it was selected by AIC and had improved retrospective patterns. The bottom water temperature covariate was chosen over AMO because it is more biologically relevant and at a more appropriate scale (i.e., AMO is derived from sea surface from the entire North Atlantic, and the bottom temperature series is for the GB yellowtail stock area).

Model runs that split the NEFSC survey time series to account for a vessel change in 2009 did not improve diagnostics and thus, a single calibrated series was used for the spring and fall. Additionally, including process error on survey catchabilities led to lower convergence rates and poor retrospective patterns. Thus, constant catchability was assumed for each of the three surveys.

In conclusion, the working group determined that the optimal model included: logistic-normal-pool0 age compositions on the commercial fleet and all three surveys, an AR1_Y correlation structure on commercial fleet selectivity, IID random effects on numbers at age, a Beverton-Holt stock recruit relationship with deviations estimated using a bottom temperature time series that had an AR1_Y correlation structure. This run met all of the model evaluation criteria outlined by the working group.

The optimal model successfully converged. The optimal run had good fits to the two NEFSC survey indices, but it had some nonrandom residual patterns for the DFO survey (e.g., negative residuals for most recent years; Figure 11-13). The model fit well to age composition from all three surveys (Figure 14-19). The model fit the commercial age data well (Figure 20-21). Additionally, the AR1_Y process error had a realistic fit to the commercial catch at age and accounted for known changes in the fishery (Figure 22). IID process error deviations were realistic for the numbers at age transitions (Figure 23). The model fit the bottom water temperature time series well (Figure 24). The Beverton-Holt stock-recruit relationship had a good fit and produced realistic results that showed recruitment decreased in recent years with the increase in bottom water temperature (Figure 25).

Retrospective patterns for fishing mortality (rho=-0.02) and spawning stock biomass (rho=0.06) were minimal. Retrospective patterns for recruitment were larger, but including the bottom water temperature time series improved retrospective consistency (rho=0.79; Figures 26-28). The model successfully passed a self test (mean F bias = -0.05; mean SSB bias = 0.045; mean R bias =0.099; Figures 29-31). The model successfully passed a jitter analysis with a convergence rate of 92% (Figure 32). Model results suggest that both fishing mortality and spawning stock biomass are low (Figure 33). Recruitment has been decreasing in recent years while the amount of older fish has increased (Table 16; Figure 34). A comparison of model estimated biomass and expanded NEFSC survey biomass displayed similar trends and scale (Figure 35).

<u>Biological Reference Points:</u>

The data inputs for biological reference points are: weight at age, natural mortality, fleet selectivity, recruitment and maturity. For GB yellowtail, natural mortality and maturity were assumed to be constant. An analysis using root mean squared error found that using an average from the most recent two years of weight at age led to the best prediction accuracy (Figure 36). For fleet selectivity, the AR1_y correlation process was used in the projections, over the long term it reverts back to the mean. The break point analyses found several different breakpoints; however, the working group decided to use a single breakpoint in 2009 (Figure 37). The working group decided to use MSY reference points because the Beverton-Holt stock recruit relationship had a decent fit to the data (Figure 25). To ensure that reference points and projections were consistent the stock was projected 200 years into the future at Fmsy (0.15) and the equilibrium SSBmsy and MSY were chosen as final reference points. Using the static SSBmsy and MSY reference points would not allow the stock to rebuild under current environmental conditions (Figure 37). Thus, the proposed reference points for GB yellowtail flounder are: MSY = 554 mt (54 – 5,661 mt); Fmsy = 0.15 (0.12-0.19); and SSBmsy = 4,942 mt (485-50,258 mt). Stock status will be updated with 2023-2024 data in the 2025 management track process, but these provisional reference points and WHAM estimates for the terminal year (2022) suggest that the stock was overfished, and overfishing was not occurring.

Projections:

The data inputs for projections are consistent with those used in reference points. Sensitivity analyses exploring the treatment of bottom water temperature in the projection years led the working group to recommend using the average bottom water temperature since 2009 in projection years. Using a linear increase was unrealistic and the working group felt using the AR1_Y process did not capture current conditions because it reverted to the mean after several years (Figures 38-40).

Discussion:

The stock assessment developed by the yellowtail research track working group and documented in this working paper is a major advancement for the GB yellowtail flounder fishery. This stock has not had an analytical assessment since 2014, so stock status was unknown and projections have not been possible. The empirical assessment that is being currently used to inform fishery management relies solely on survey data. At least one of the three surveys failed to sample the stock from 2019-2022 due to Covid-19, vessel changes or mechanical issues (TRAC, 2023). The candidate WHAM model is an integrated analysis of all available fishery and survey data, with a few minor exceptions for surveys that do not sample the entire stock area and fishery monitoring programs that sample a small portion of the catch. Additionally, it has been hypothesized that environmental drivers might be the cause of low stock size/productivity (Stone et al 2004). The results of this work support that hypothesis and suggest that increasing water temperature has led to decreases in recruitment in recent years.

The decision to include bottom temperature as a covariate on recruitment was supported by alternative process error models. An IID error structure was used to estimate process error on numbers at age and a bottom water temperature time series was included to help estimate deviations from a Beverton-Holt stock recruit relationship. The inclusion of an environmental covariate in the estimation model and the assumption of recent environmental conditions for projections and reference points are substantial improvements over previous analytical models that assumed stationary production.

The majority of diagnostics performed well for the optimal model, but future model development should explore improving fit to the DFO survey. The model has mostly negative residuals for the DFO survey in recent years. These residuals could indicate changes in survey coverage or vessels that have occurred in recent years. Assessments of other Georges Bank fisheries have also had poor fits to the DFO survey (Hansell, 2023). Future model development would benefit from collaborating with Canadian scientists to improve fit to their survey.

Future work should continue on the best methods for incorporating environmental covariates into the assessment, reference points and projections. The yellowtail assessments are the second research track to recommend using an environmental covariate in the Northeast US (NEFSC, 2023). More work is needed to determine how these relationships should be accounted for in projections (e.g., Figures 38-40), how they hold up over time, and a process for re-examining environmental drivers in future assessments. Further development of WHAM is needed to include survey values of zero in the model likelihood. Currently, the platform assumes that all true zeros in survey indices (i.e., the survey did not catch any yellowtail flounder) are modeled as missing values. This is currently not a problem for GB yellowtail, but given the low stock size and trends, it is possible that in the future surveys may not capture any yellowtail flounder on Georges Bank.

References:

Cadrin, S. X. 2024. APPROXIMATING NATURAL MORTALITY RATE FOR NEW ENGLAND YELLOWTAIL FLOUNDER STOCKS. Working paper in support of the Yellowtail flounder research track. Available here: https://apps-nefsc.fisheries.noaa.gov/saw/sasi.php

Hansell, A. C. 2023. From VPA to state-space: Georges Bank Winter flounder. Working paper in support of the State-space research track-working group. Available here: https://apps-nefsc.fisheries.noaa.gov/saw/sasi.php

Legault, C.M. and S.X. Cadrin. 2014. A Guided Tour through the Yellowtail Flounder Literature for the 2014 Georges Bank Yellowtail Flounder Diagnostic and Empirical Approach Benchmark. TRAC Ref Doc 2014/. https://repository.library.noaa.gov/view/noaa/27075

O'Brien, L., and K. Clark, editors. 2014. Proceedings of the Transboundary Resources Assessment Committee for Georges Bank Yellowtail Flounder Diagnostic and Empirical Approach Benchmark: Report of Meeting held 14–18 April 2014. TRAC Proceedings 2014/01.

Stock, B. C., and Miller, T. J. 2021. The Woods Hole Assessment Model (WHAM): A general state-space assessment framework that incorporates time- and age-varying processes via random effects and links to environmental covariates. Fisheries Research, 240: 105967. doi: 10.1016/j.fishres.2021.105967

Stone, H. H., Gavaris, S., Legault, C. M., Neilson, J. D., & Cadrin, S. X. (2004). Collapse and recovery of the yellowtail flounder (Limanda ferruginea) fishery on Georges Bank. *Journal of Sea Research*, *51*(3-4), 261-270.

TRAC. 2021. Georges Bank Yellowtail Flounder. TRAC Status Report 2021/XX.

TRAC. 2023. Georges Bank Yellowtail Flounder. TRAC Status Report 2022/01.

Northeast Fisheries Science Center (NEFSC), 2023. Working Group Report for the Black Seabass Research Track Stock Assessment Working Group. Availble here: https://apps-nefsc.fisheries.noaa.gov/saw/sasi.php

Table 1: Comparison of WHAM model runs exploring different age composition likelihoods.

Age comp	conv	NLL	dAIC	AIC	Rho R	Rho SSB	Rho Fbar
multinomial	TRUE	7489.378	17406.4	15216.8	1.0124	0.8392	-0.4141
dir-mult	TRUE	3785.416	10006.4	7816.8	2.0078	0.2031	-0.182
dirichlet-miss0	TRUE	-858.819	718	-1471.6	2.2583	0.1956	-0.179
dirichlet-pool0	TRUE	-850.234	735.1	-1454.5	2.2456	0.1992	-0.1844
logistic-normal-miss0	TRUE	-746.359	942.9	-1246.7	3.7946	0.158	-0.1313
logistic-normal-ar1-miss0	TRUE	-1221.78	0	-2189.6	3.2691	0.1102	-0.173
logistic-normal-pool0	TRUE	-745.244	945.1	-1244.5	3.5926	0.1494	-0.1283

Table 2: Comparison of WHAM model runs exploring different recruitment assumptions.

Type	conv	pdHess	dAIC	AIC	Rho R	Rho SSB	Rho Fbar
Fixed effects	TRUE	TRUE	0	-1206	5.068	0.3142	-0.2215
Ar1_y	TRUE	TRUE	85.5	-1120.5	1.6481	0.3487	-0.2398
Beverton-holt	TRUE	TRUE	175.6	-1030.4	6.7204	0.3627	-0.2613
iid	TRUE	TRUE	179.2	-1026.8	6.569	0.3166	-0.2397

Table 3: Comparison of WHAM model runs exploring different selectivity assumptions.

RE	conv	pdHess	dAIC	AIC	Rho R	Rho SSB	Rho Fbar
2dar1	TRUE	TRUE	0	-1496	2.7126	0.5645	-0.3032
Ar1_y	TRUE	TRUE	2.4	-1493.6	3.4697	0.6533	-0.314
block_1994	TRUE	TRUE	64.9	-1431.1	6.3254	1.5862	-0.4448
block_ 2010	TRUE	TRUE	225	-1271	4.0561	0.6666	-0.1304
None	TRUE	TRUE	375.5	-1120.5	1.6481	0.3487	-0.2398
iid	FALSE	FALSE	-	-	-	-	-

Table 4: Comparison of WHAM model runs exploring different process error assumptions on numbers at age.

RE Age-1	RE Age 2+	conv	pdHess	dAIC	AIC	Rho R	Rho SSB	Rho Fbar
2dar1		TRUE	TRUE	0	-2117	0.209	-0.0987	0.1054
Ar1_y	Ar1_y	TRUE	TRUE	1.3	-2115.7	0.1975	-0.1106	0.1101
Ar1_y		TRUE	TRUE	1.3	-2115.7	0.1975	-0.1106	0.1101
Ar1_y	2dar1	TRUE	TRUE	19.2	-2097.8	0.4492	-0.0956	0.1254
Ar1_y	IID	TRUE	TRUE	46	-2071	0.346	-0.115	0.1151
AR1_y	Ar1_a	TRUE	TRUE	47.6	-2069.4	0.4121	-0.0977	0.0998
Ar1_a		TRUE	TRUE	97.2	-2019.8	0.9292	-0.0061	0.0024
IID		TRUE	TRUE	138.3	-1978.7	1.1715	0.0054	-0.0235
none	none	TRUE	TRUE	996.5	-1120.5	1.6481	0.3487	-0.2398

Table 5: Simulation self-test results for WHAM model runs exploring different process error assumptions.

RE Age-1	RE Age-2	F bias	R bias	SSB bias	Mean bias
IID		17.3	-9.36	-14.56	13.7
Ar1_y		27.96	-0.06	-15.39	14.5
Ar1_y	Ar1_y	28.24	-0.1	-15.53	14.6
Ar1_y	IID	34.56	-13.55	-18.42	22.2
Ar1_a		30.57	-16.32	-21.42	22.8
2dar1		35.91	-11.98	-21.55	23.1
AR1_y	Ar1_a	40.07	-19.87	-22.46	27.5
Ar1_y	2dar1	106.01	-37.17	-37.36	60.2

Table 6: WHAM model runs exploring different process error assumptions on numbers at age and fit to bottom temperature as a modulate of recruitment.

Model	Rec	Ecov process	Ecov how	NAA re	conv	pdHess	dAIC	Rho R	Rho SSB	Rho Fbar
m1	Random	rw	none	iid	TRUE	TRUE	180.1	1.3796	0.0516	-0.0672
m2	Random	rw	controlling	iid	TRUE	TRUE	141.6	0.6004	0.0314	-0.0431
m3	Random	ar1	controlling	iid	TRUE	TRUE	145.5	0.727	0.0352	-0.0476
m4	Bev-Holt	rw	none	iid	TRUE	TRUE	153.6	0.9851	0.0679	-0.078
m5	Bev-Holt	rw	controlling	iid	TRUE	TRUE	133.7	0.3168	0.0137	-0.0233
m6	Bev-Holt	rw	limiting	iid	TRUE	TRUE	133.5	0.8653	0.0261	-0.0283
m7	Bev-Holt	ar1	controlling	iid	TRUE	TRUE	135.9	0.4025	0.0196	-0.0294
m8	Bev-Holt	ar1	limiting	iid	TRUE	TRUE	136.7	0.8536	0.0287	-0.0311
m1	Random	rw	none	ar1_y	TRUE	TRUE	143.5	1.4054	0.0176	-0.0354
m2	Random	rw	controlling	ar1_y	TRUE	TRUE	100.3	0.6464	0.0129	-0.0207
m3	Random	ar1	controlling	ar1_y	TRUE	TRUE	104.5	0.7911	0.0178	-0.0257
m4	Bev-Holt	rw	none	ar1_y	TRUE	TRUE	112.2	0.9523	0.0317	-0.0435
m5	Bev-Holt	rw	controlling	ar1_y	TRUE	TRUE	91.7	0.3465	-0.0173	0.0123
m6	Bev-Holt	rw	limiting	ar1_y	TRUE	TRUE	91.6	0.9324	-0.0088	0.0163
m7	Bev-Holt	ar1	controlling	ar1_y	TRUE	TRUE	94	0.4193	-0.0115	0.0063
m8	Bev-Holt	ar1	limiting	ar1_a	TRUE	TRUE	95	0.9188	-0.0047	0.0115
m1	Random	rw	none	ar1_y	FALSE	FALSE	974.1	1.3753	2.8795	-0.2362
m2	Random	rw	controlling	ar1_y decoupled	TRUE	FALSE	2.8	-0.0334	-0.048	0.0627
m3	Random	ar1	controlling	ar1_y decoupled	TRUE	FALSE	6.4	0.3107	-0.0537	0.062
m4	Bev-Holt	rw	none	ar1_y decoupled	TRUE	FALSE	6.9	0.3102	-0.0849	0.0994
m5	Bev-Holt	rw	controlling	ar1_y decoupled	TRUE	FALSE	1.3	-0.0222	-0.0936	0.1128
m6	Bev-Holt	rw	limiting	ar1_y decoupled	TRUE	FALSE	0	0.1568	-0.0487	0.0629
m7	Bev-Holt	ar1	controlling	ar1_y	TRUE	FALSE	3.4	-0.0166	-0.0791	0.097

				decoupled						
m8	Bev-Holt	ar1	limiting	ar1_y decoupled	TRUE	FALSE	3.3	0.1367	-0.055	0.0699
m1	Random	rw	none	2dar1	TRUE	FALSE	10.7	0.3097	-0.0484	0.0575
m2	Random	rw	controlling	2dar1	TRUE	FALSE	4.3	0.0203	-0.0456	0.0593
m3	Random	ar1	controlling	2dar1	TRUE	FALSE	8	1.0035	-0.0665	0.0683
m4	Bev-Holt	rw	none	2dar1	TRUE	FALSE	8	0.3609	-0.0708	0.0844
m5	Bev-Holt	rw	controlling	2dar1	TRUE	FALSE	2.4	-0.0997	-0.0718	0.0894
m6	Bev-Holt	rw	limiting	2dar1	TRUE	FALSE	1.3	0.1958	-0.0231	0.0365
m7	Bev-Holt	ar1	controlling	2dar1	TRUE	FALSE	4.9	-0.2332	-0.0559	0.0762
m8	Bev-Holt	ar1	limiting	2dar1	TRUE	FALSE	4.7	0.3419	-0.0395	0.055

Table 7: WHAM model runs assuming AR1_y process error on numbers at age and fitting to AMO as a modulate on recruitment.

Recruitment	Ecov process	Ecov how	conv	pdHess	NLL	daic	aic	rho R	rho SSB	rho Fbar
Random	rw	---	TRUE	TRUE	-1170.85	5.6	-2177.7	0.2096	-0.1054	0.1043
Random	rw	Controlling	TRUE	TRUE	-1171.08	7.1	-2176.2	0.1713	-0.105	0.1058
Random	ar1	Controlling	TRUE	TRUE	-1169.66	12	-2171.3	0.1277	-0.1242	0.1243
Bev-Holt	rw	---	TRUE	FALSE						
Bev-Holt	rw	Controlling	FALSE	FALSE						
Bev-Holt	rw	Limiting	FALSE	FALSE						
Bev-Holt	ar1	Controlling	FALSE	FALSE						
Bev-Holt	ar1	Limiting	TRUE	FALSE						

Table 8: WHAM model runs assuming AR1_y decoupled process error on numbers at age and fitting to AMO as a modulate on recruitment.

Recruitment	Ecov process	Ecov how	conv	pdHess	NLL	daic	aic	rho R	rho SSB	rho Fbar
Random	rw	---	TRUE	TRUE	-1170.84	1.1	-2179.7	0.1976	-0.1106	0.1101
Random	rw	Controlling	TRUE	TRUE	-1170.94	2.9	-2177.9	0.1882	-0.1014	0.1011

Random	ar1	Controlling	TRUE	TRUE	-1169.62	7.6	-2173.2	0.1405	-0.1144	0.111
Bev-Holt	rw	---	TRUE	TRUE	-1172.41	0	-2180.8	0.1947	-0.0883	0.0991
Bev-Holt	rw	Controlling	TRUE	TRUE	-1172.46	1.9	-2178.9	0.1722	-0.0855	0.094
Bev-Holt	rw	Limiting	TRUE	TRUE	-1172.45	1.9	-2178.9	0.1454	-0.0878	0.0974
Bev-Holt	ar1	Controlling	TRUE	TRUE	-1171.17	6.5	-2174.3	0.093	-0.0968	0.1057
Bev-Holt	ar1	Limiting	TRUE	TRUE	-1171.17	6.5	-2174.3	0.1881	-0.0936	0.1055

Table 9: WHAM model runs assuming 2DAR1 process error on numbers at age and fitting to AMO as a modulate on recruitment.

Recruitment	Ecov process	Ecov how	conv	pdHess	NLL	daic	aic	rho R	rho SSB	rho Fbar
Random	rw	---	TRUE	TRUE	-1135.39	0	-2106.8	2.81	0.4525	-0.2086
Random	rw	Controlling	TRUE	TRUE	-1135.87	1.1	-2105.7	2.7929	0.4847	-0.2155
Random	ar1	Controlling	TRUE	TRUE	-1134.5	5.8	-2101	3.1308	0.5578	-0.2198

Table 10: WHAM model runs assuming IID process error on numbers at age and fitting to AMO as a modulate on recruitment.

Recruitment	Ecov process	Ecov how	conv	pdHess	NLL	daic	aic	rho R	rho SSB	rho Fbar
Random	rw	---	TRUE	TRUE	-1101.36	38.7	-2042.7	1.1715	0.0054	-0.0235
Random	rw	Controlling	TRUE	TRUE	-1106.04	31.3	-2050.1	1.0686	0.0036	-0.0193
Random	ar1	Controlling	TRUE	TRUE	-1104.55	36.3	-2045.1	1.1547	0.0065	-0.0227
Bev-Holt	rw	---	TRUE	TRUE	-1115.92	11.6	-2069.8	0.813	0.0188	-0.0303
Bev-Holt	rw	Controlling	TRUE	TRUE	-1122.34	0.7	-2080.7	0.3504	-0.0162	0.0097
Bev-Holt	rw	Limiting	TRUE	TRUE	-1122.7	0	-2081.4	0.0829	-0.0385	0.0227
Bev-Holt	ar1	Controlling	TRUE	TRUE	-1120.93	5.5	-2075.9	0.4362	-0.014	0.0084
Bev-Holt	ar1	Limiting	TRUE	TRUE	-1121.22	5	-2076.4	0.7276	-0.0028	-0.0057

Table 11: WHAM model runs assuming AR1_y process error on numbers at age and fitting to AMO as a modulate on natural mortality.

M model	M re	Ecov process	Ecov link	conv	pdHess	NLL	daic	aic	Rho R	Rho SSB	Rho Fbar

---	none	ar1	0	FALSE	FALSE						
---	none	ar1	1	TRUE	TRUE	-1179.95	38.1	-2193.9	0.368	-0.013	0.002
---	none	ar1	2	TRUE	FALSE						
age-specific	none	ar1	0	TRUE	FALSE						
weight-at-age	none	ar1	0	TRUE	TRUE	-1172.14	55.7	-2176.3	2.293	0.207	-0.079
constant	none	ar1	0	TRUE	TRUE	-1170.73	56.5	-2175.5	1.279	0.087	-0.035
constant	ar1_y	ar1	0	TRUE	TRUE	-1172.02	58	-2174	1.961	0.28	-0.142
age-specific	2dar1	ar1	0	TRUE	FALSE						
constant	none	ar1	1	TRUE	FALSE						
constant	none	ar1	2	TRUE	FALSE						
---	2dar1	ar1	0	TRUE	FALSE						

Table 12: WHAM model runs assuming AR1_y decoupled process error on numbers at age and fitting to AMO as a modulate on natural mortality.

M model	M_re	Ecov process	Ecov link	conv	pdHess	NLL	daic	aic	rho R	Rho SSB	Rho Fbar
---	none	ar1	0	TRUE	FALSE						
---	none	ar1	1	TRUE	TRUE	-1179.95	22.1	-2193.9	0.368	-0.013	0.002
---	none	ar1	2								
age-specific	none	ar1	0								
weight-at-age	none	ar1	0	TRUE	TRUE	-1171.1	39.8	-2176.2	0.629	-0.061	0.081
constant	none	ar1	0	TRUE	TRUE	-1169.6	40.8	-2175.2	0.197	-0.11	0.11
constant	ar1_y	ar1	0	TRUE	FALSE						

age-specific	2dar1	ar1	0	FALSE	FALSE						
constant	none	ar1	1	TRUE	TRUE	-1193	0	-2216	-0.74	-0.347	1.138
constant	none	ar1	2								
---	2dar1	ar1	0								

Table 13: WHAM model runs assuming 2DAR1 process error on numbers at age and fitting to AMO as a modulate on natural mortality.

M	re	Ecov process	Ecov link	conv	pdHess	NLL	daic	aic	Rho R	Rho SSB	Rho Fbar
---	none	ar1	0	TRUE	TRUE	-1170.71	70.8	-2175.4	0.0869	-0.1154	0.1212
---	none	ar1	1	TRUE	TRUE	-1181.26	51.7	-2194.5	0.3252	-0.0191	0.009
weight-at-age	none	ar1	0	TRUE	TRUE	-1175.04	66.1	-2180.1	4.1552	0.7748	-0.2376
constant	none	ar1	0	TRUE	TRUE	-1173.1	68	-2178.2	1.6118	0.2786	-0.1299
constant	ar1_y	ar1	0	TRUE	TRUE	-1174.3	69.6	-2176.6	12107.23	3754.377	-0.394
age-specific	2dar1	ar1	0	TRUE	FALSE						
constant	none	ar1	1	TRUE	FALSE						
---	2dar1	ar1	0	TRUE	FALSE						

Table 14:WHAM model rRuns assuming IID process error on numbers at age and fitting to AMO as a modulate on natural mortality.

M model	M re	Ecov process	Ecov link	conv	pdHess	NLL	daic	aic	Rho R	Rho SSB	Rho Fbar
---	none	ar1	0	TRUE	TRUE	-1090.83	182.1	-2019.7	1.077	0.018	-0.037

---	none	ar1	1	TRUE	TRUE	-1090.99	183.8	-2018	1.572	0.374	-0.267
Age-specfic											
weight-at-age	none	ar1	0	TRUE	TRUE	-1118.63	130.5	-2071.3	5.327	0.312	-0.11
constant	none	ar1	0	TRUE	TRUE	-1117.84	130.1	-2071.7	2.489	0.179	-0.099
constant	ar1_y	ar1	0	TRUE	TRUE	-1122.62	124.6	-2077.2	1982517	5170227	-0.886
constant	2dar1	ar1	0	TRUE	TRUE	-1185.92	0	-2201.8	0.046	-0.015	0.055
constant	none	ar1	1	TRUE	FALSE						
constant	none	ar1									
---	2dar1	ar1									

Table 15: WHAM model runs assuming different process error on numbers at age and fitting to bottom temperature as a modulate on natural mortality.

RE	M model	Ecov process	Ecov link	conv	pdHess	NLL	daic	aic	Rho R	Rho SSB	Rho Fbar
2dar1	---	ar1	0	TRUE	FALSE						
	---	ar1	1	TRUE	TRUE	-1094.91	-	-2019.8	1.095	-0.004	0.046
	constant	ar1	1	TRUE	FALSE						
Ar1_y	---	ar1	0	TRUE	FALSE						
	---	ar1	1	FALSE	FALSE						
	constant	ar1	1	TRUE	TRUE	-1092.36	-	-2014.7	1.511	0.028	-0.009
ar1_y decouple	---	ar1	0	FALSE	FALSE						
	---	ar1	1	FALSE	FALSE						

	constant	ar1	1	TRUE	TRUE	-1092.36	-	-2014.7	1.511	0.028	-0.009
IID	---	ar1	0	TRUE	TRUE	-1012.98	54	-1862	1.077	0.018	-0.037
	---	ar1	1	TRUE	TRUE	-1041	0	-1916	5941166	470184	-0.393
	constant	ar1	1	TRUE	TRUE	-1040.12	3.8	-1912.2	2.195	0.067	-0.016

Table 16: Trends from the optimal WHAM model.

Year	Rec	SSB	F
1973	84886	26848	1.052283
1974	116050	17620	1.0625
1975	109152	11930	1.467708
1976	30025	11220	1.371553
1977	19690	9864	1.268806
1978	164192	5872	1.074096
1979	29252	8012	0.9862
1980	30087	22680	0.362482
1981	31712	12559	0.794067
1982	44993	13161	0.80026
1983	14674	12065	0.858341
1984	18249	4451	1.464623
1985	30375	3354	0.853353
1986	9608	4465	0.781808
1987	7496	2536	1.462521
1988	8434	2275	1.109767

1989	11215	2609	0.653925
1990	8004	4228	0.969056
1991	58486	2970	1.301793
1992	14124	4672	0.776874
1993	38093	4366	0.820488
1994	9540	4803	1.642238
1995	13269	4767	0.532769
1996	31608	11677	0.151753
1997	23950	21637	0.124782
1998	60797	15511	0.298488
1999	60591	40222	0.138336
2000	34654	49978	0.153586
2001	42700	49728	0.186508
2002	64182	48235	0.192798
2003	57944	33410	0.302568
2004	35757	17928	1.125488
2005	51778	20423	0.328468
2006	138548	18163	0.558731
2007	72162	29363	0.122553
2008	28434	34728	0.127089
2009	45138	39076	0.064943
2010	25641	28923	0.051321
2011	24132	19532	0.112956
2012	23669	16424	0.065881
2013	27393	6561	0.047798
2014	12026	4902	0.074804

2015	5414	3744	0.083331
2016	2778	2650	0.028598
2017	2797	1581	0.07871
2018	3576	1020	0.070062
2019	13489	889	0.009969
2020	2612	510	0.219442
2021	607	1008	0.081157
2022	4395	747	0.020842

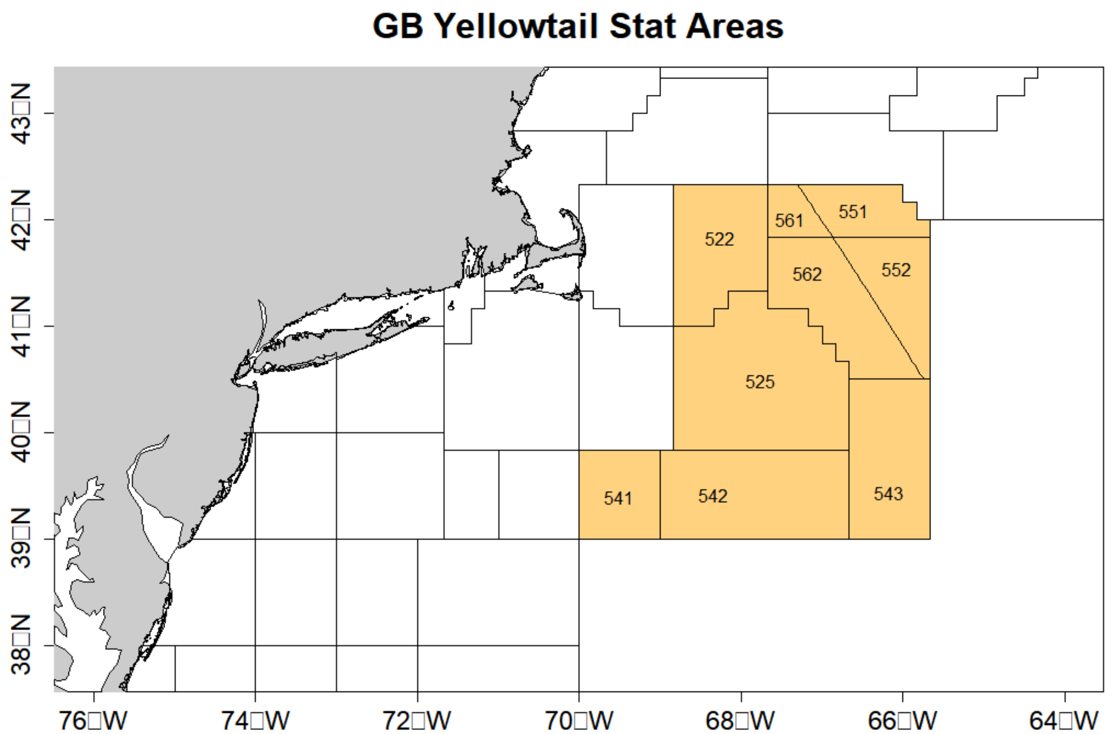

Figure 1: Commercial reporting areas that define the Georges Bank yellowtail flounder stock.

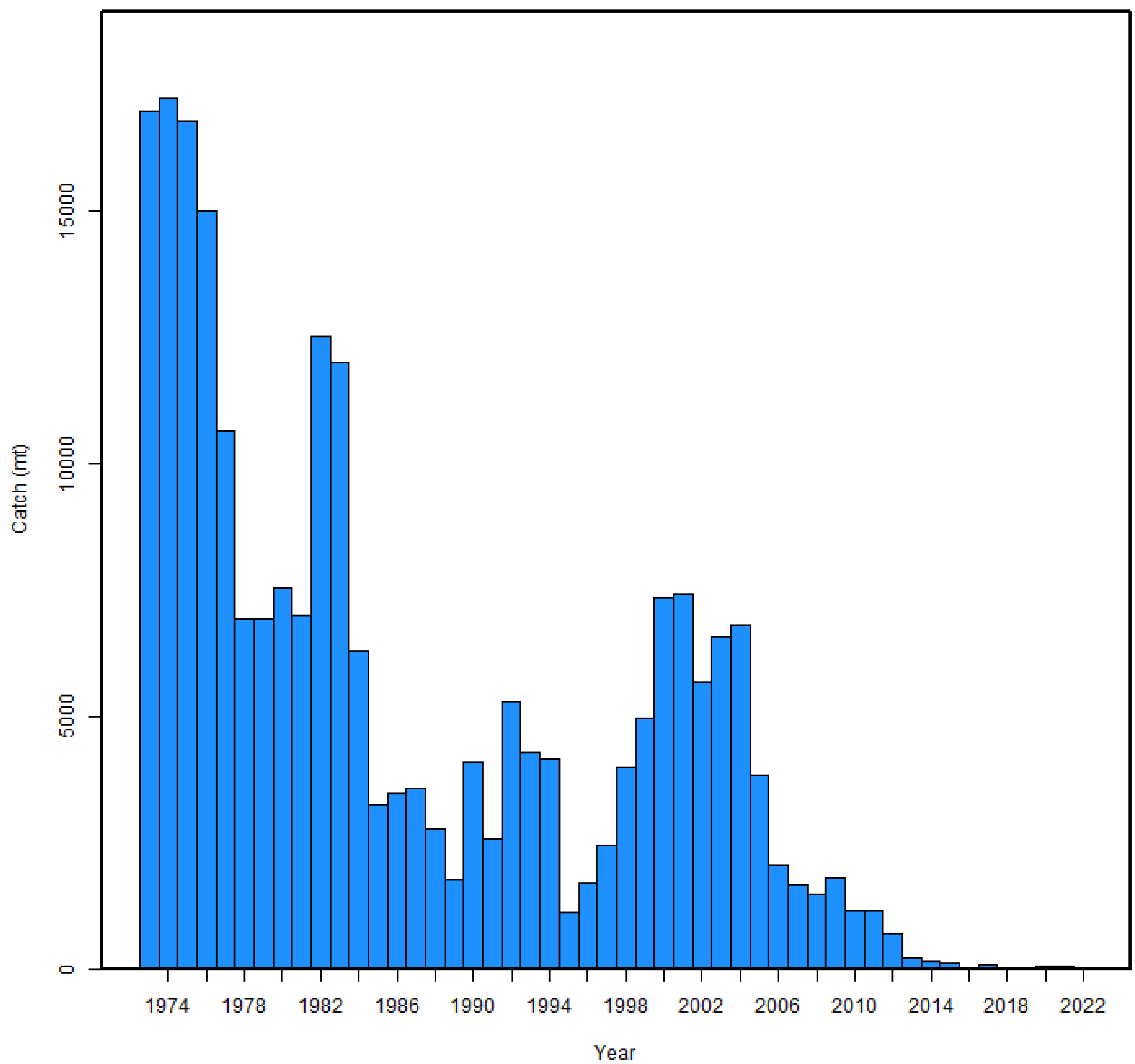

Figure 2: Combined US and Canadian catch used in the WHAM runs.

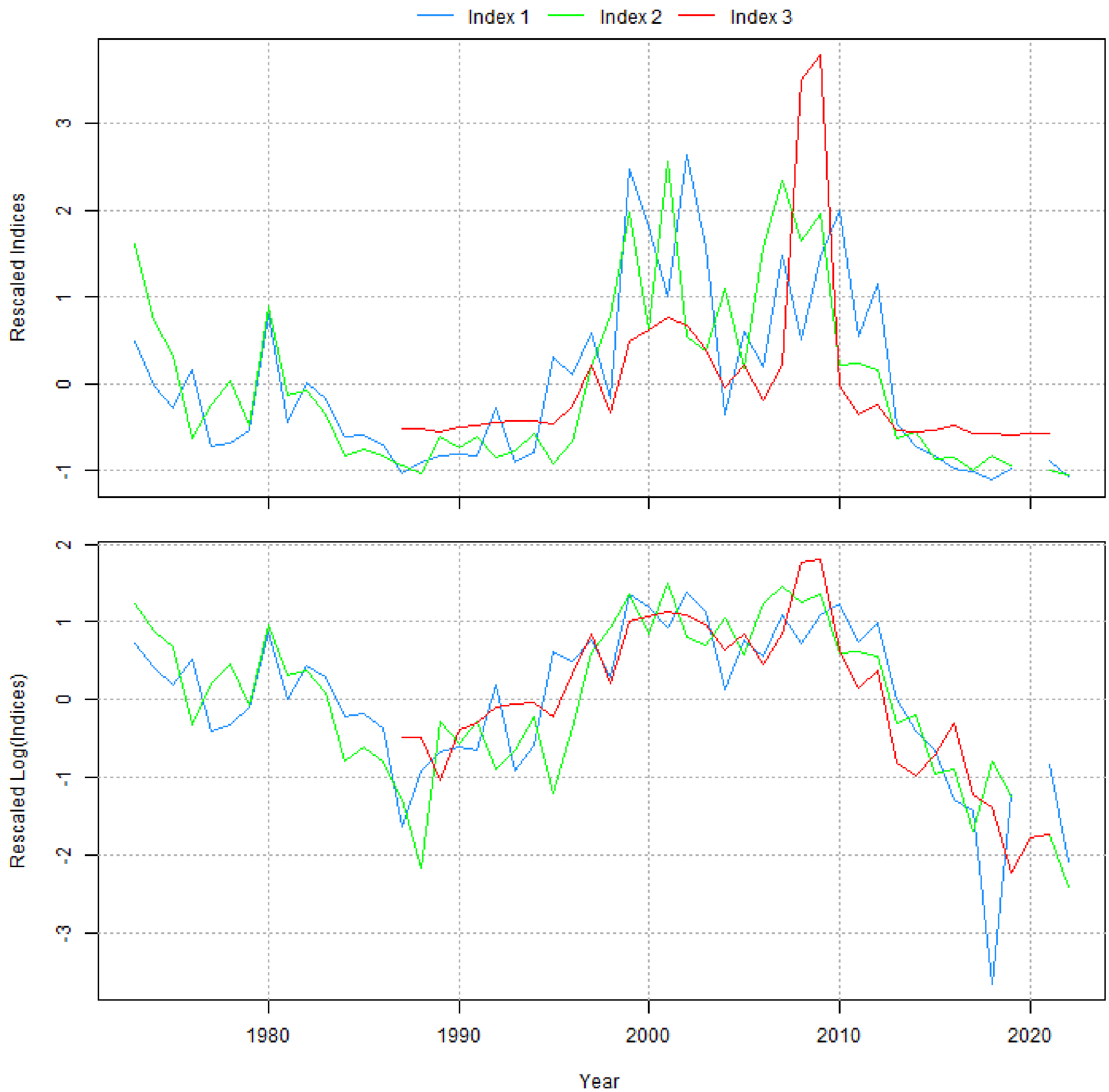

Figure 3: Trawl survey indices used in WHAM runs (Index 1 = NEFSC spring; Index 2 = NEFSC fall; Index 3 = DFO).

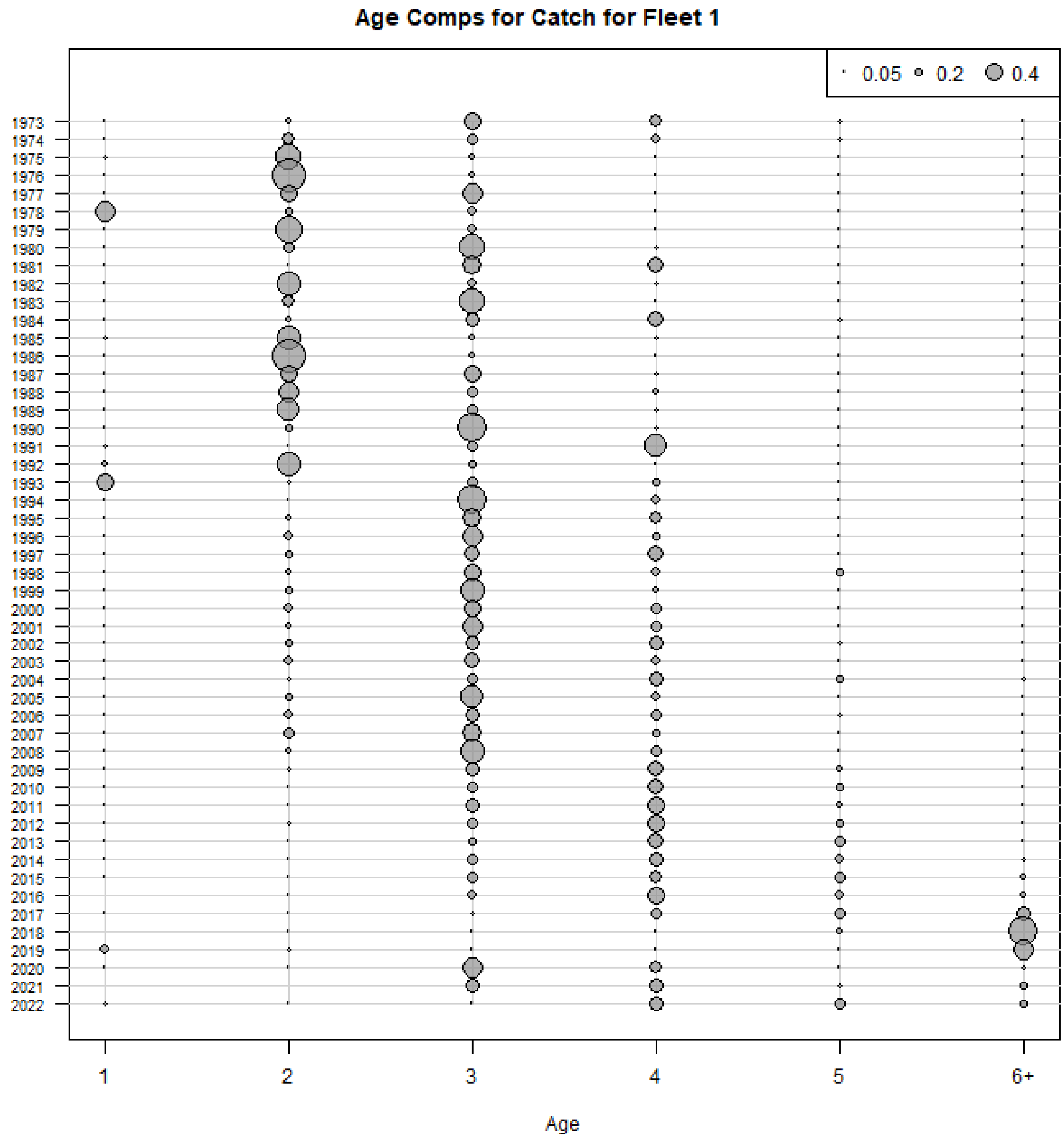

Figure 5: Commercial catch at age.

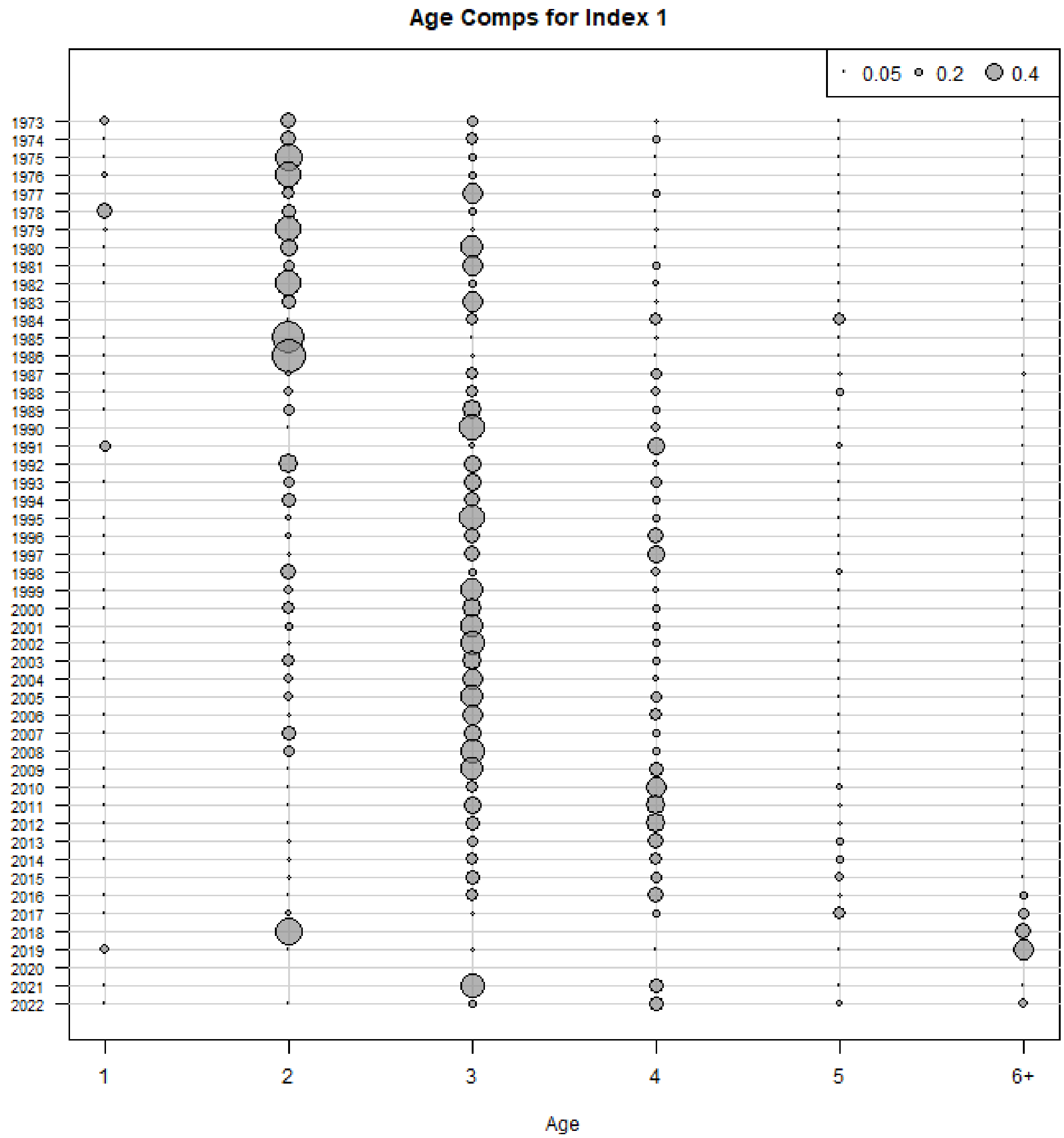

Figure 6: Catch at age from the NEFSC spring survey.

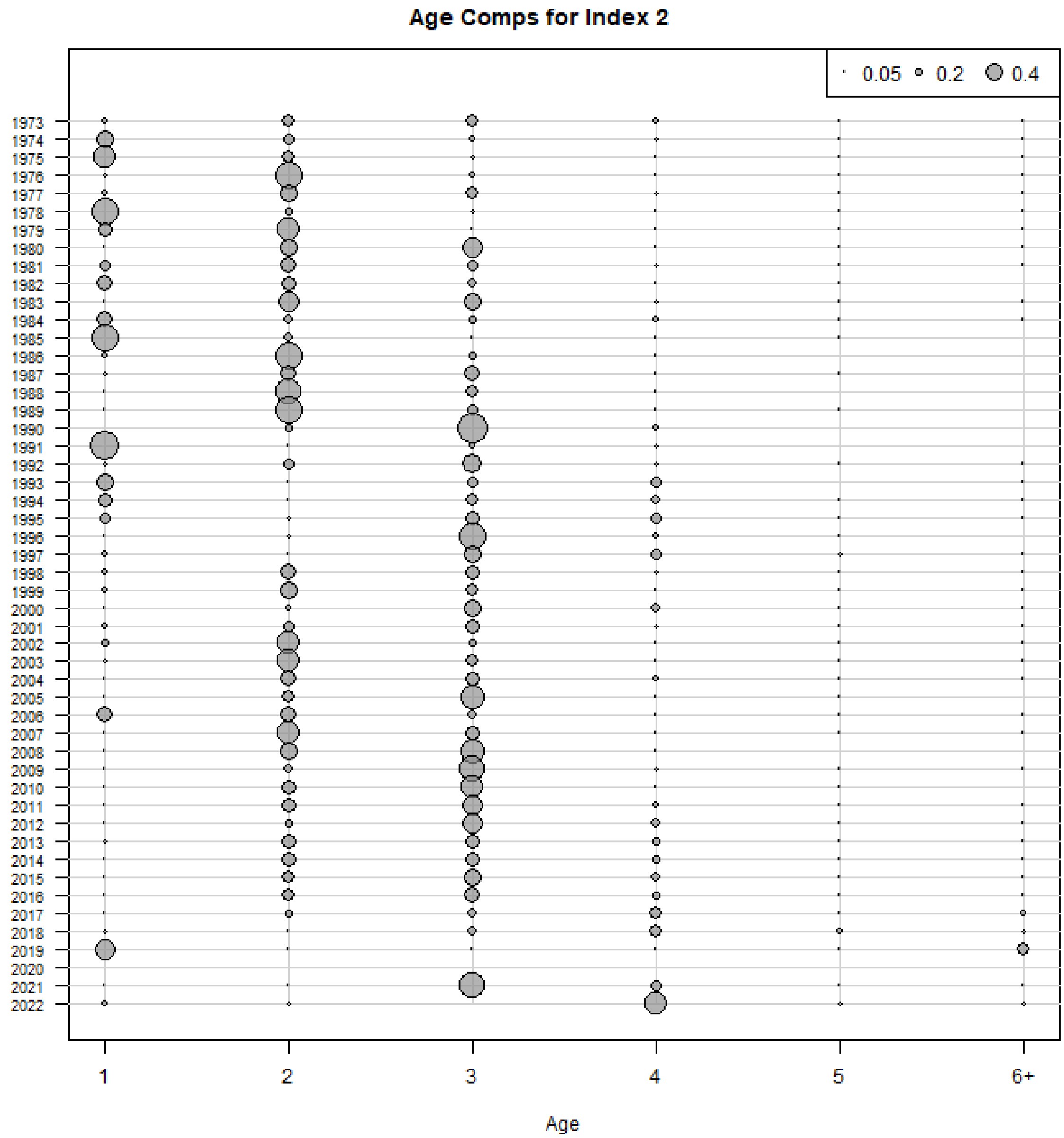

Figure 7: Catch at age for the NEFSC fall survey.

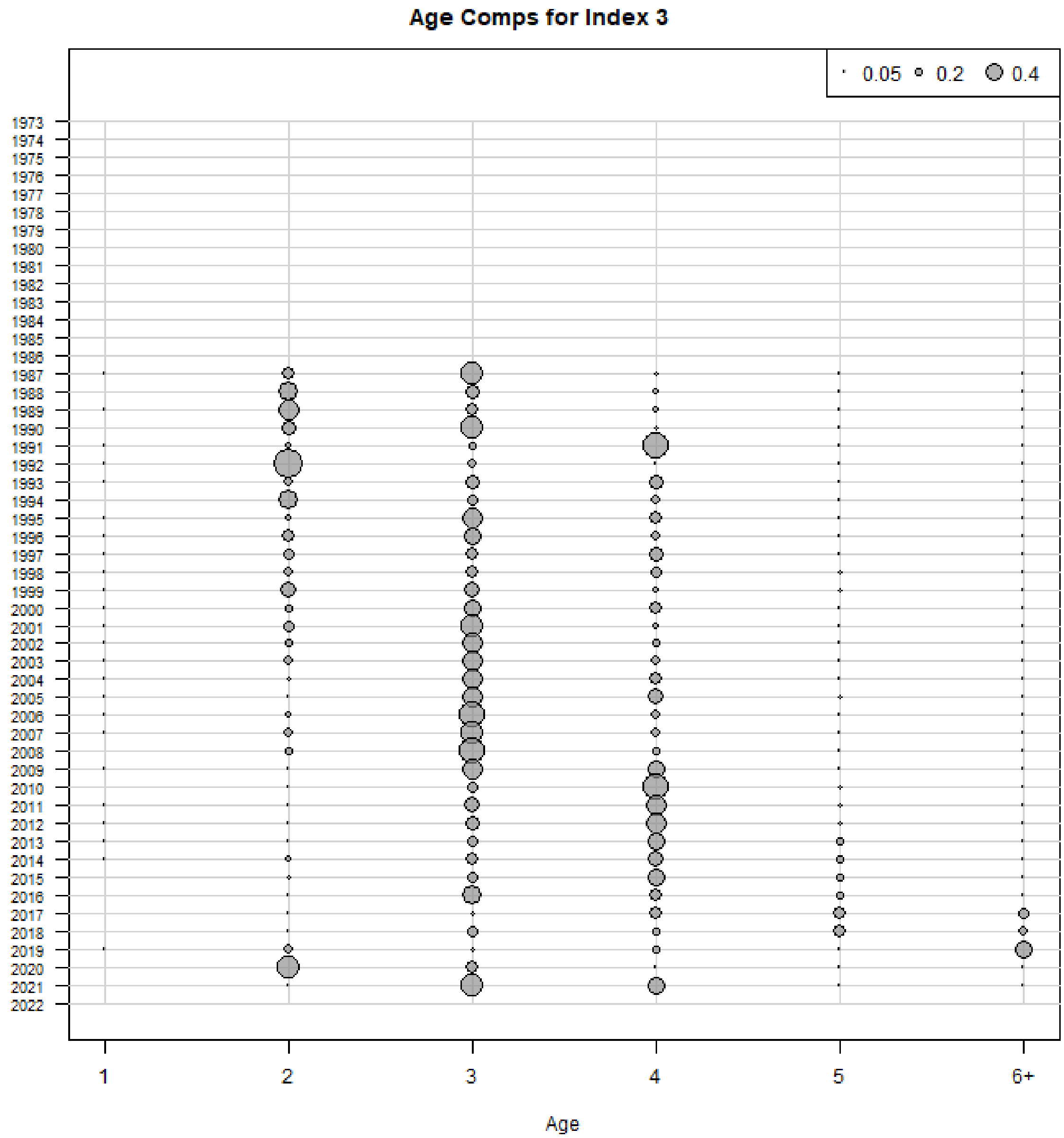

Figure 8: Catch at age from the DFO survey.

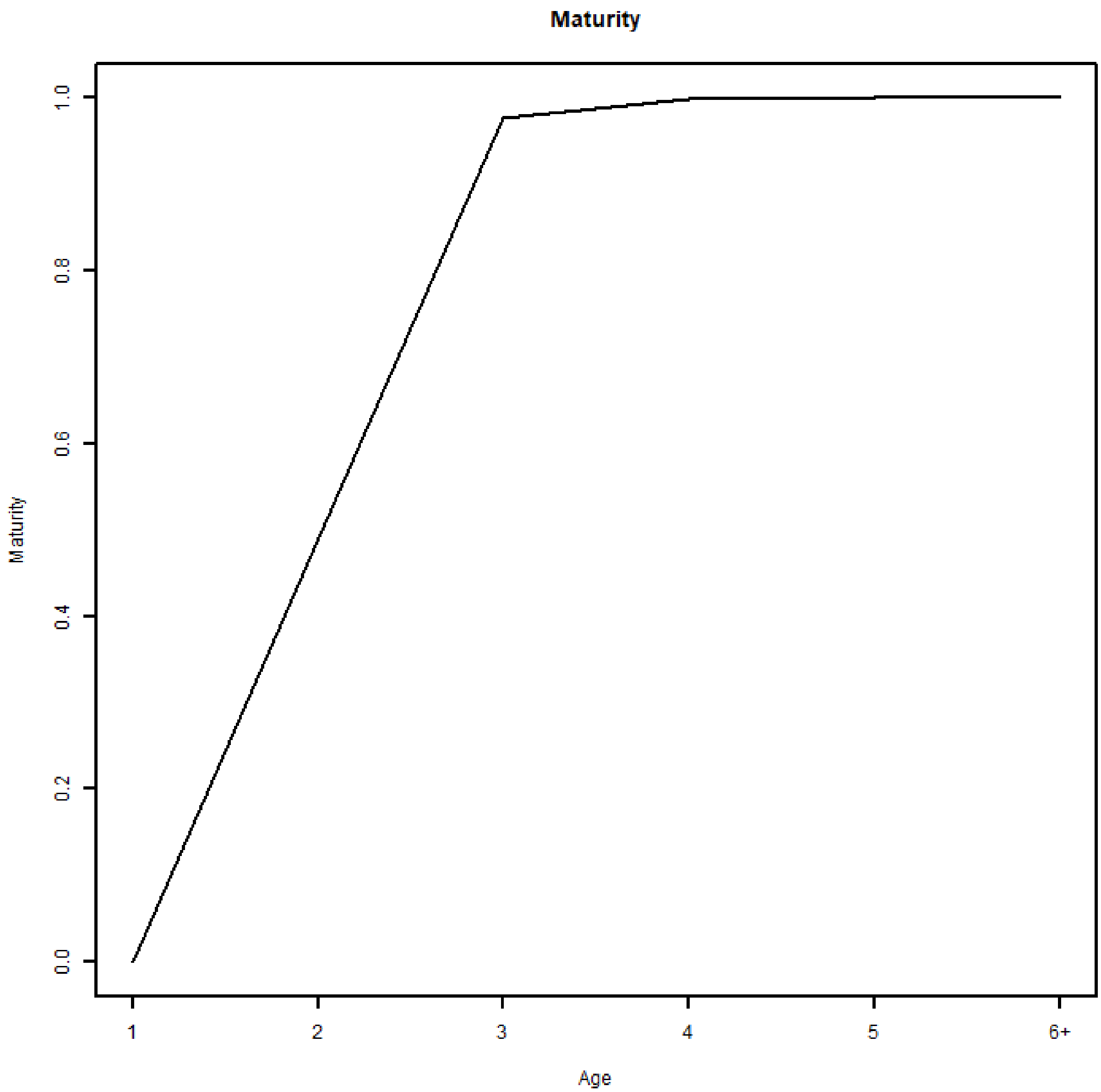

Figure 9: Maturity at age assumed for all WHAM runs.

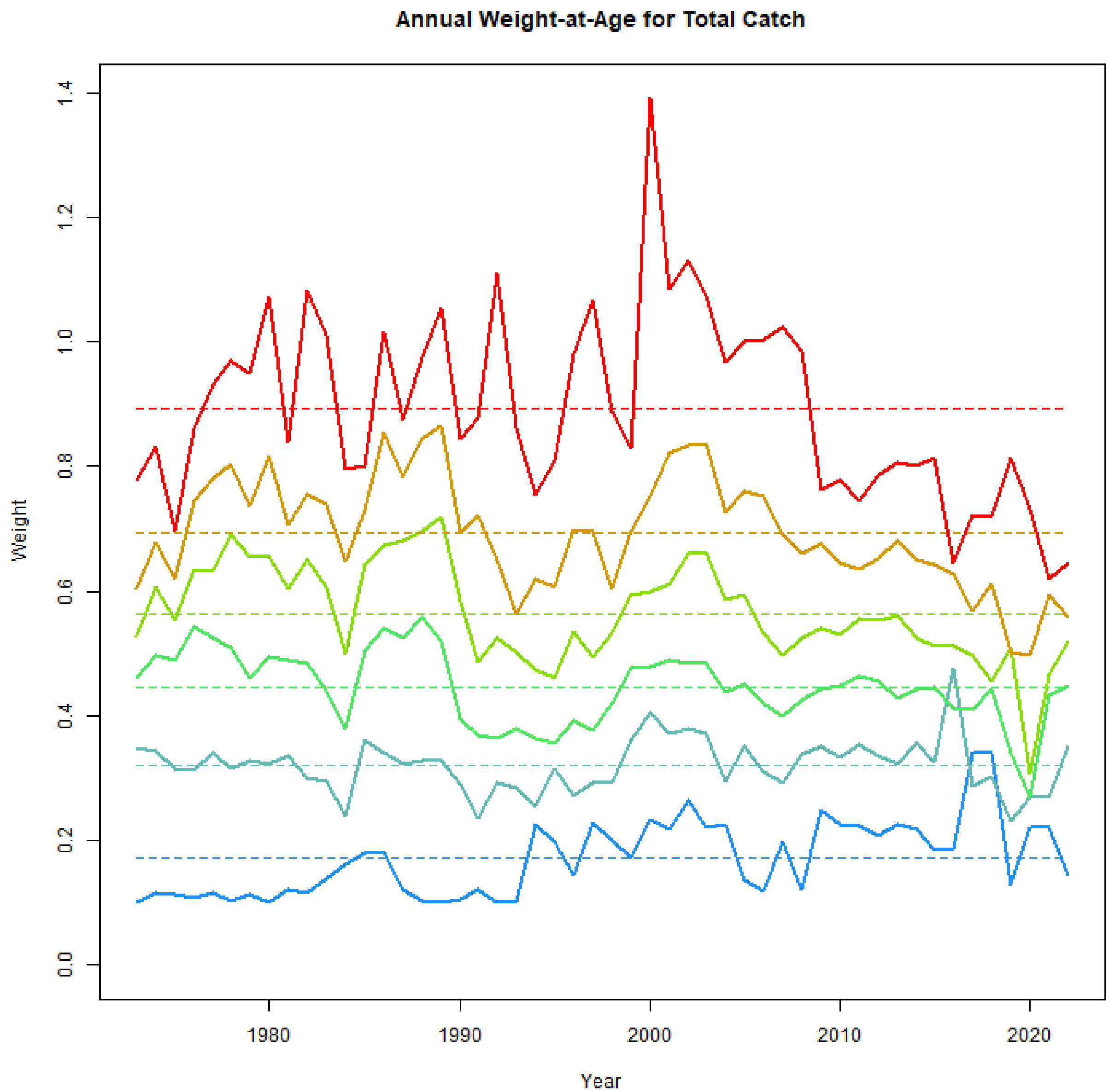

Figure 10: Weight at age for the commercial catch.

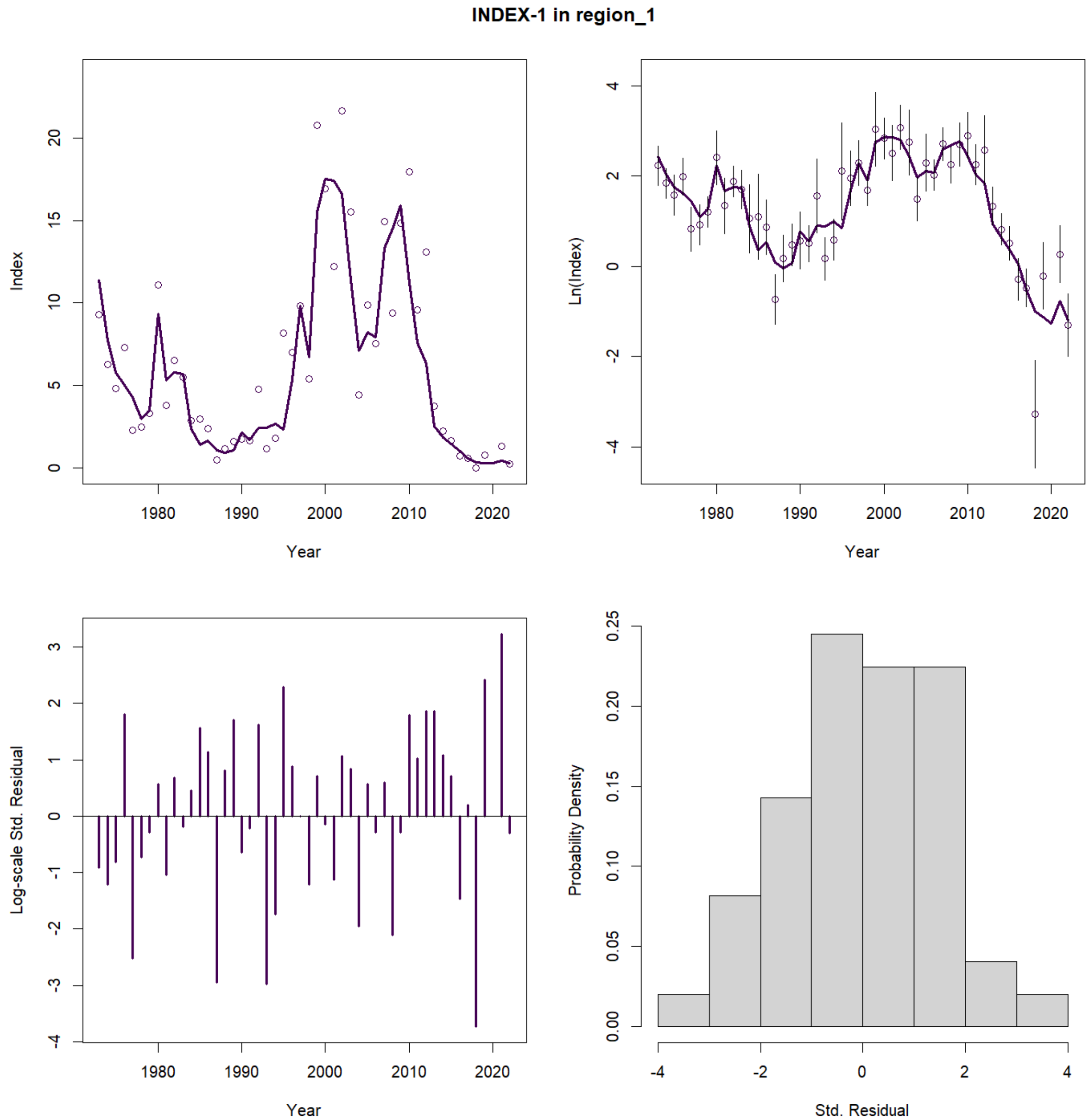

Figure 11: WHAM model fit to the NEFSC spring survey.

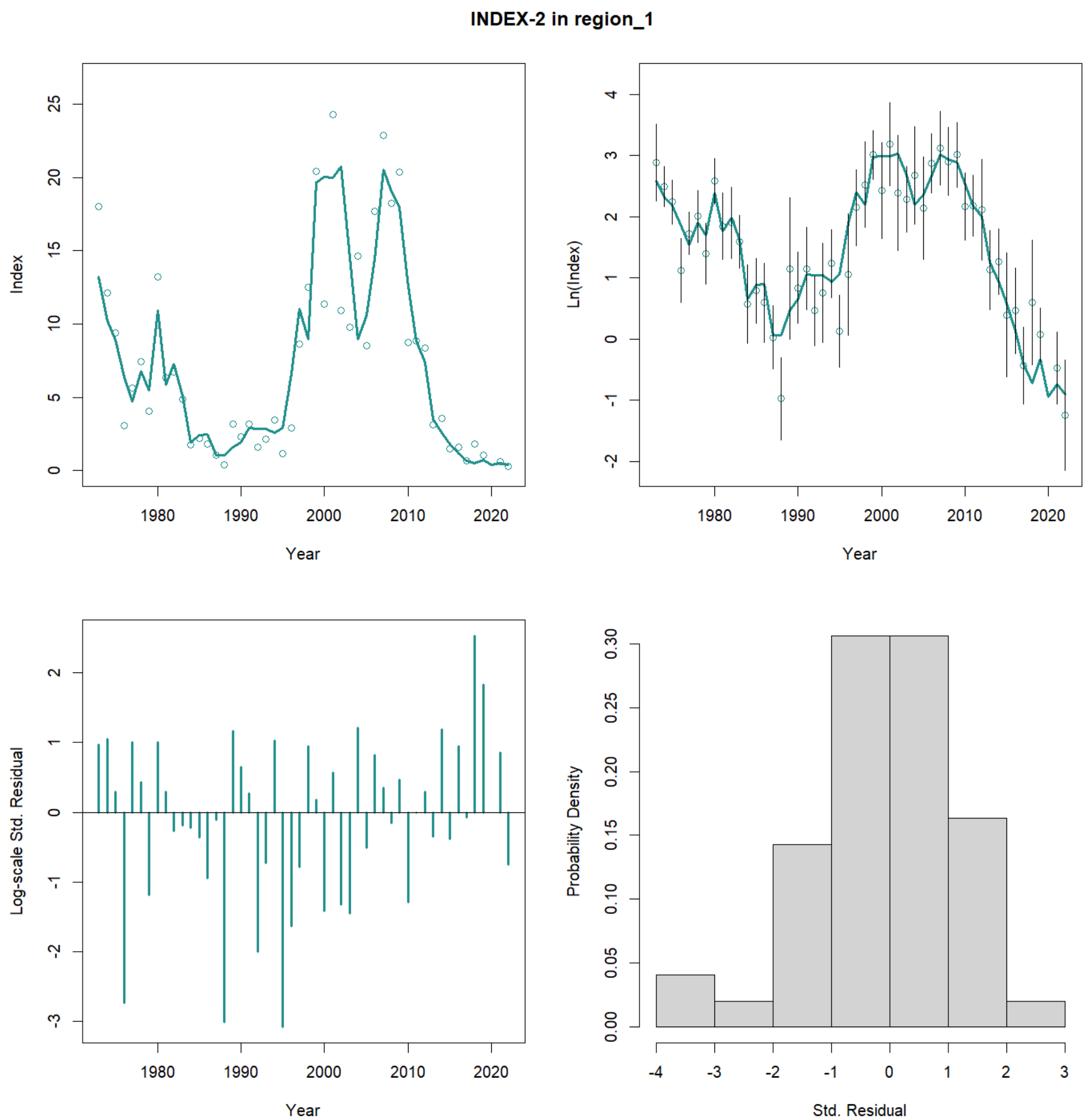

Figure 12: WHAM model fit to the NEFSC fall survey.

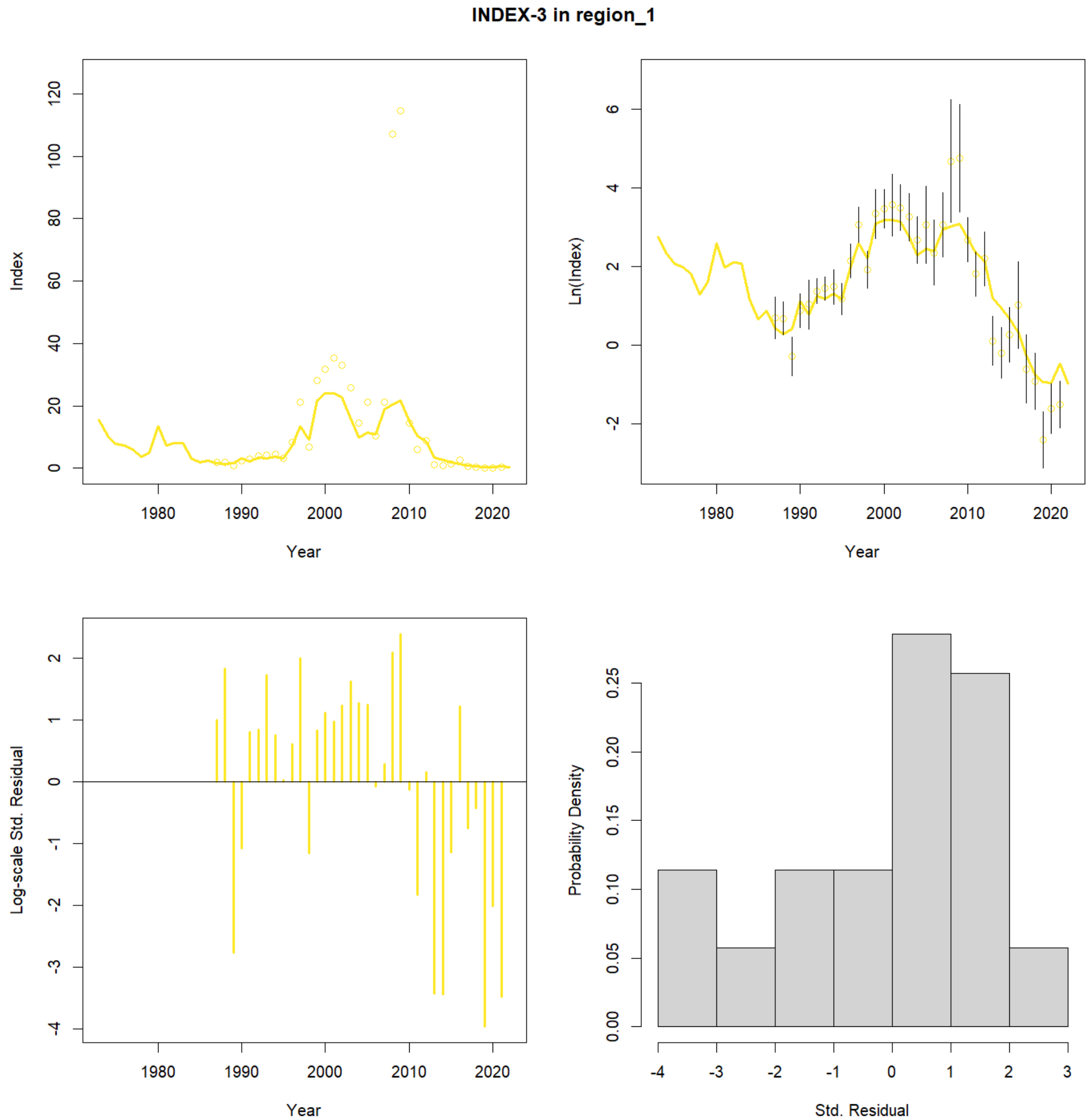

Figure 13: WHAM model fit to the DFO survey.

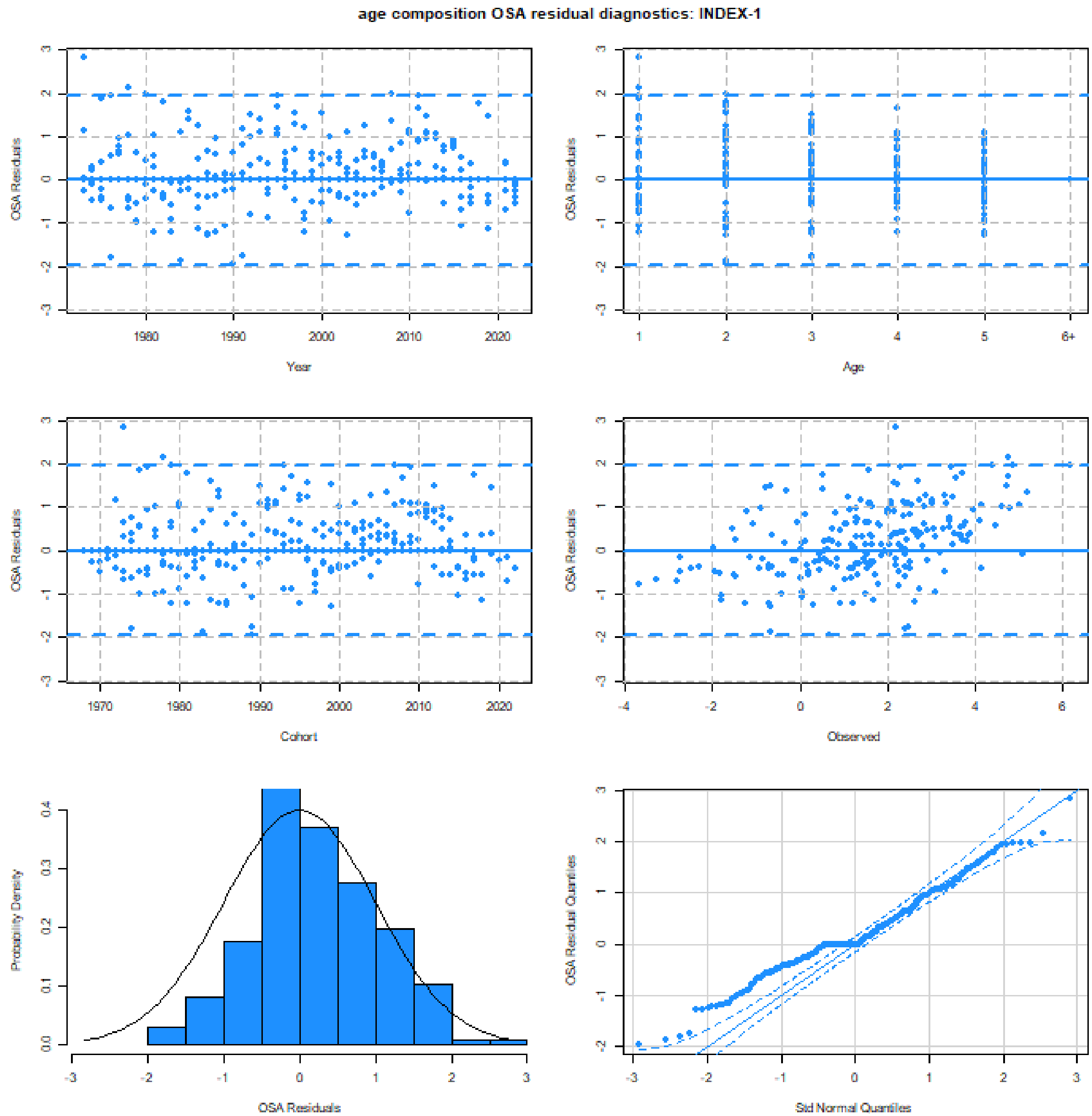

Figure 14: WHAM model fit to age data from the NEFSC spring survey.

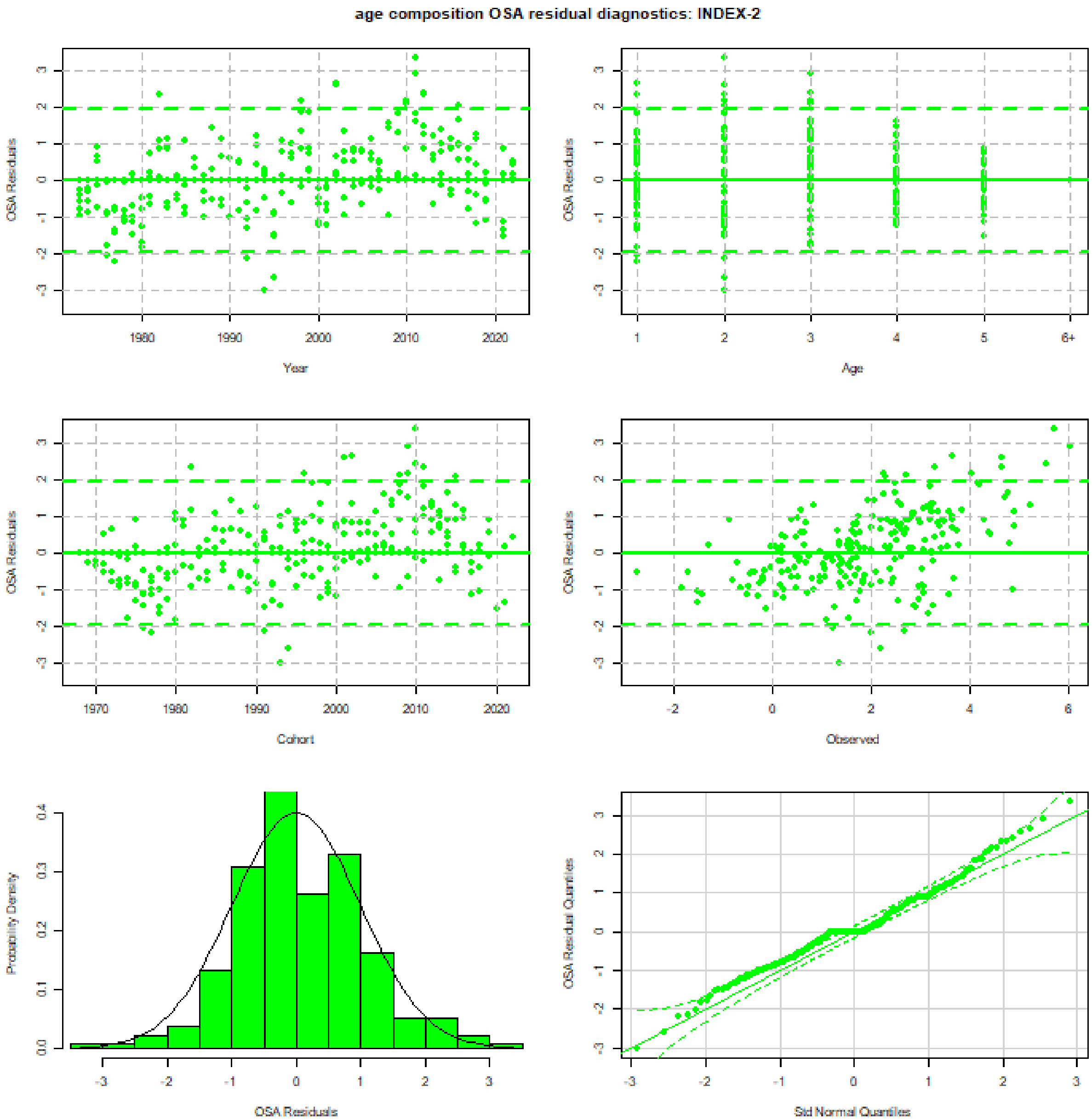

Figure 15: WHAM model fit to age data from the NFESC fall survey.

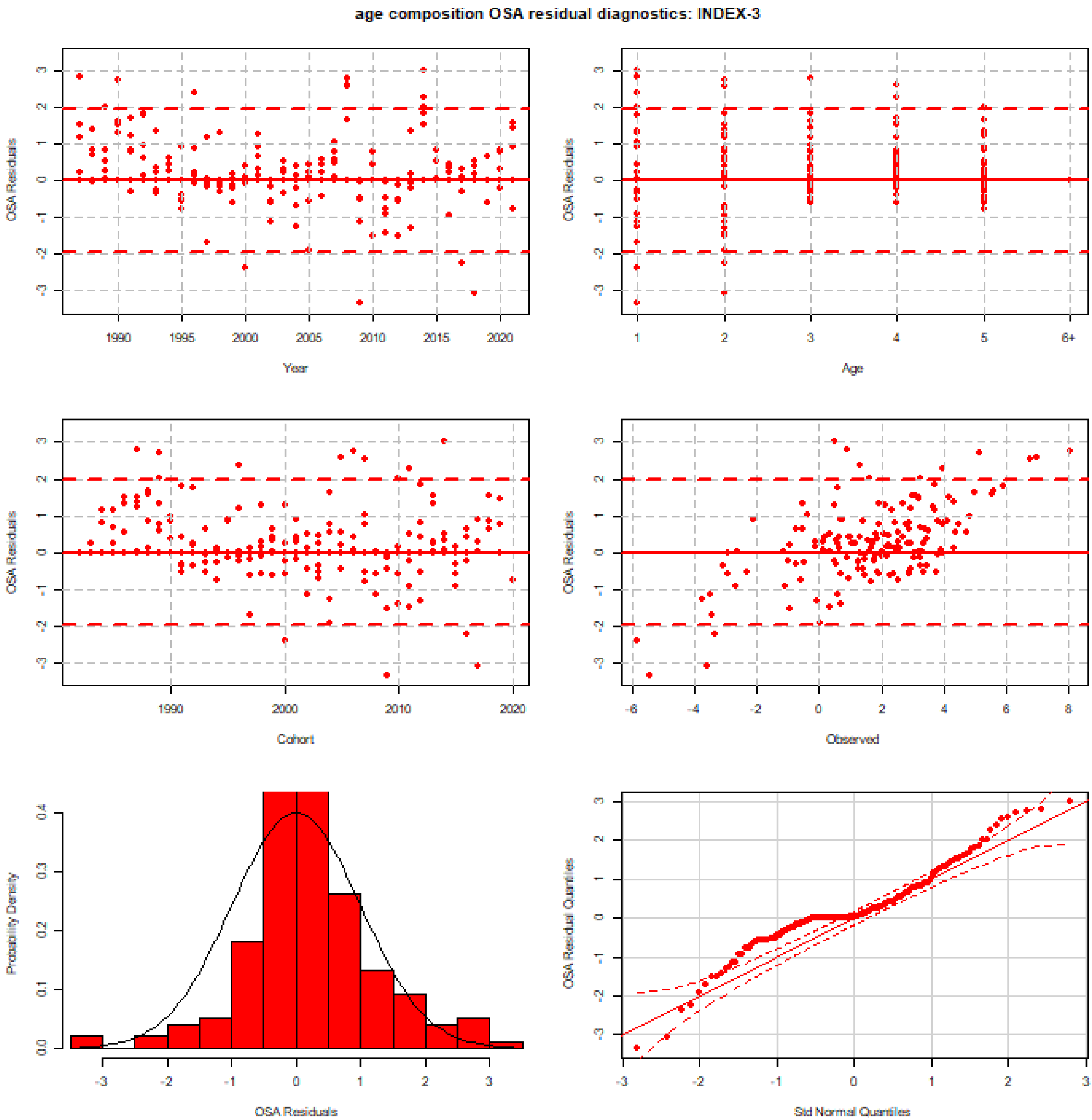

Figure 16: WHAm model fit to age data from the DFO survey.

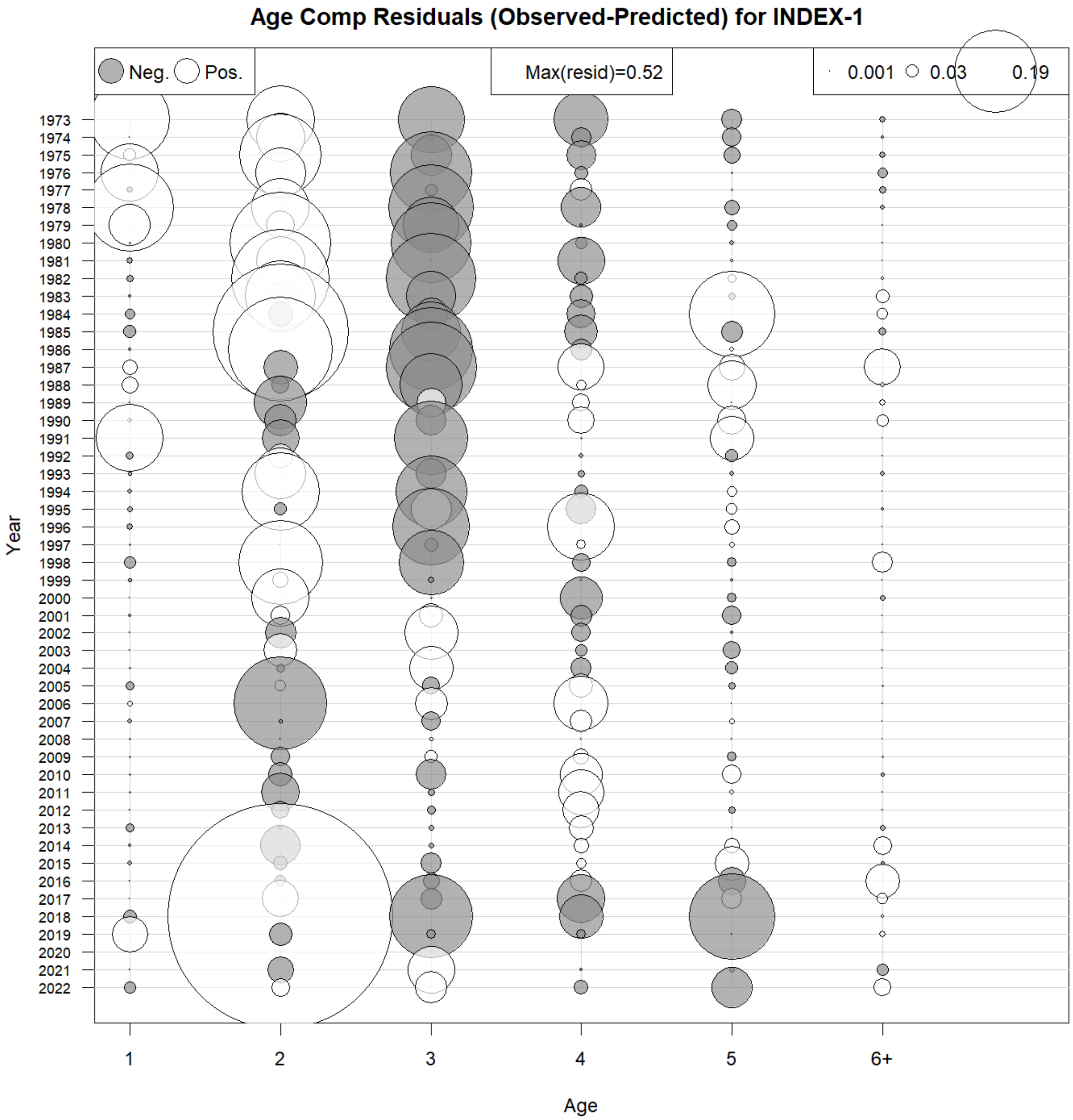

Figure 17: WHAm model fit to age composition from the NEFSC spring survey.

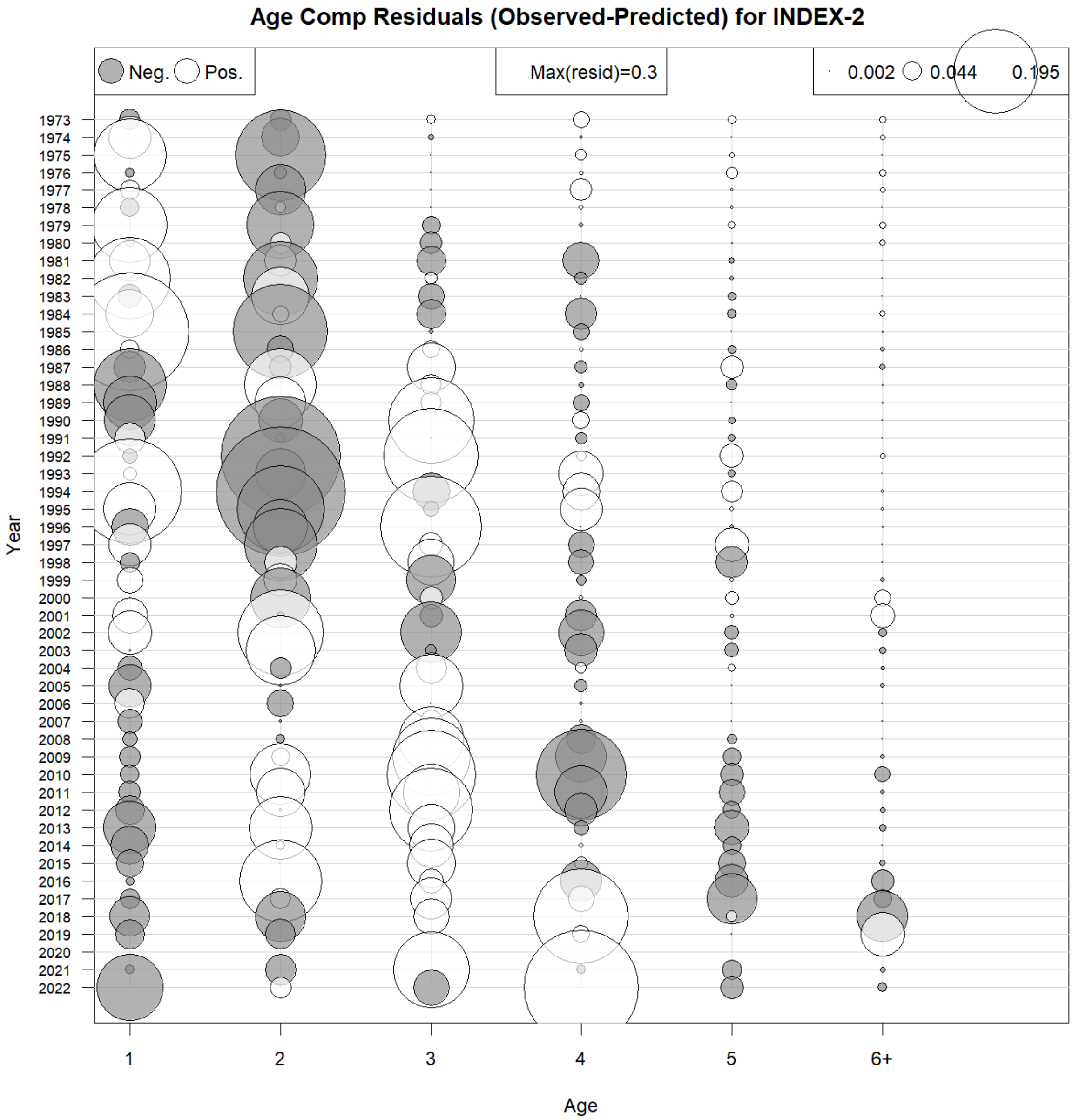

Figure 18: WHAM model fit to age composition data from the NEFSC fall survey.

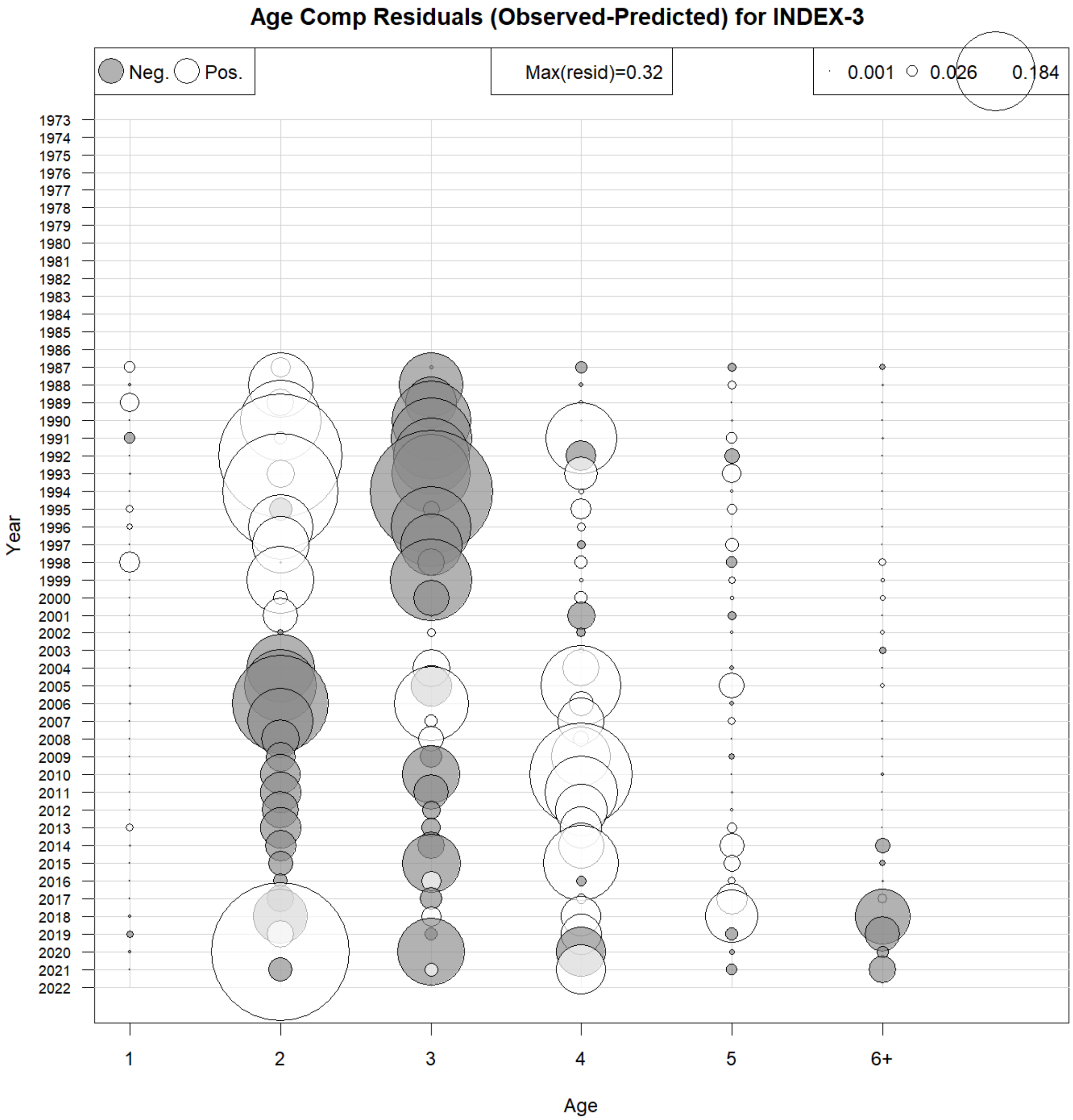

Figure 19: WHAM model fit to age composition from the DFO survey.

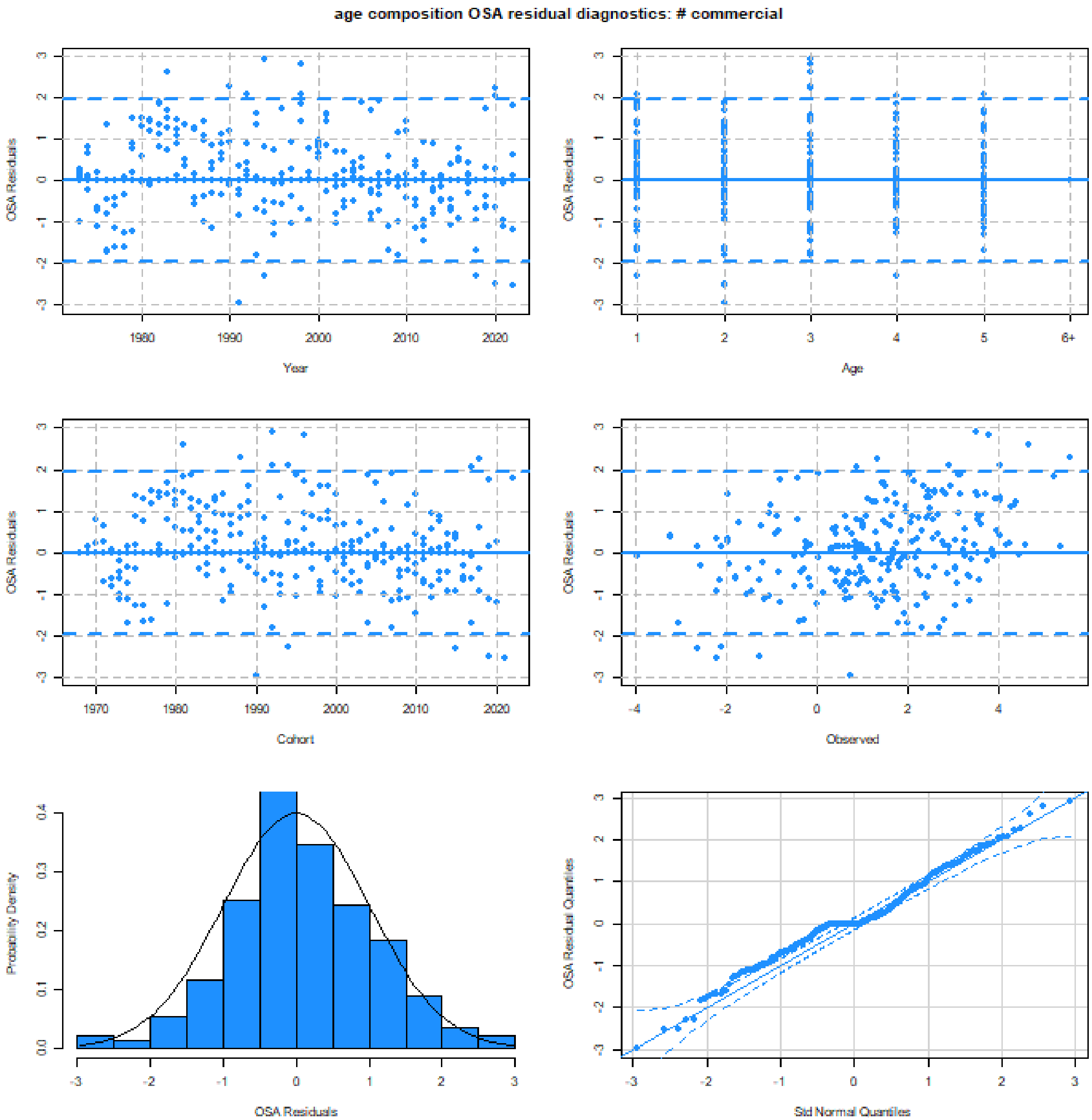

Figure 20: WHAM model fit to age composition from the commercial fleet.

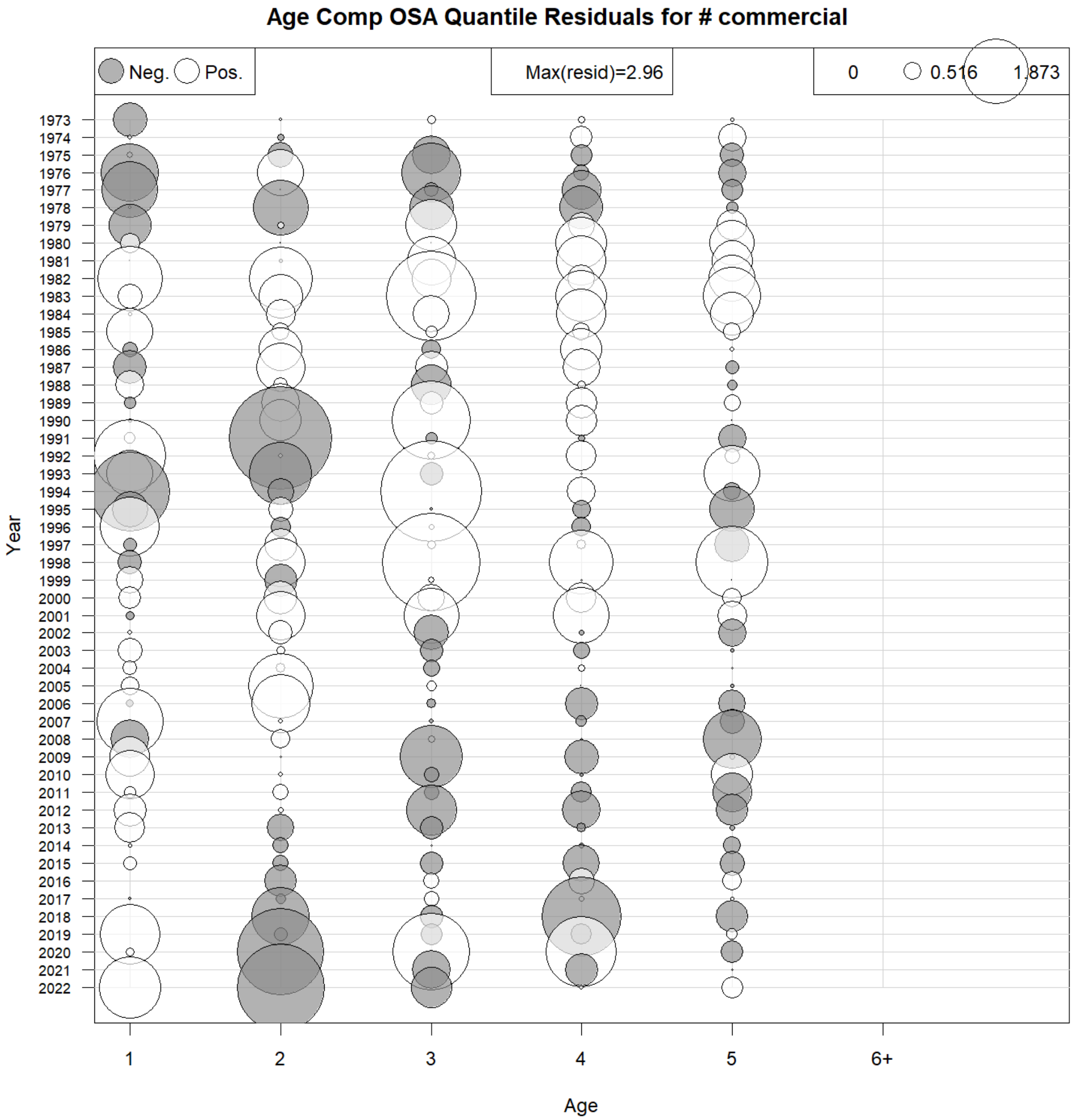

Figure 21: WHAM model fit to age composition from the commercial fleet.

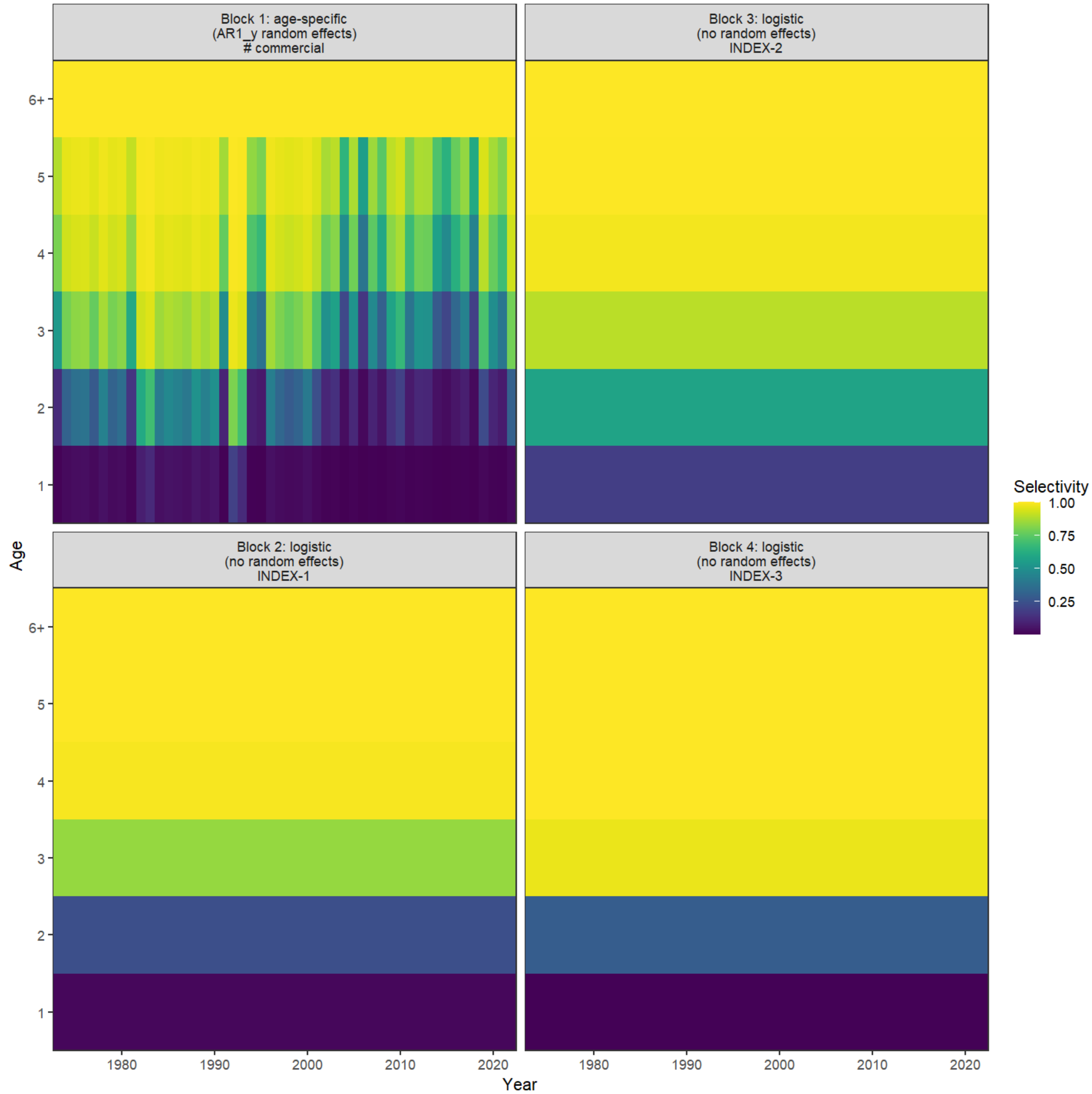

Figure 22: Selectivity estimates for the commercial fleet (Block 1) the NEFSC spring (Block 2), the NEFSC fall (Block 3) and DFO survey (Block 4).

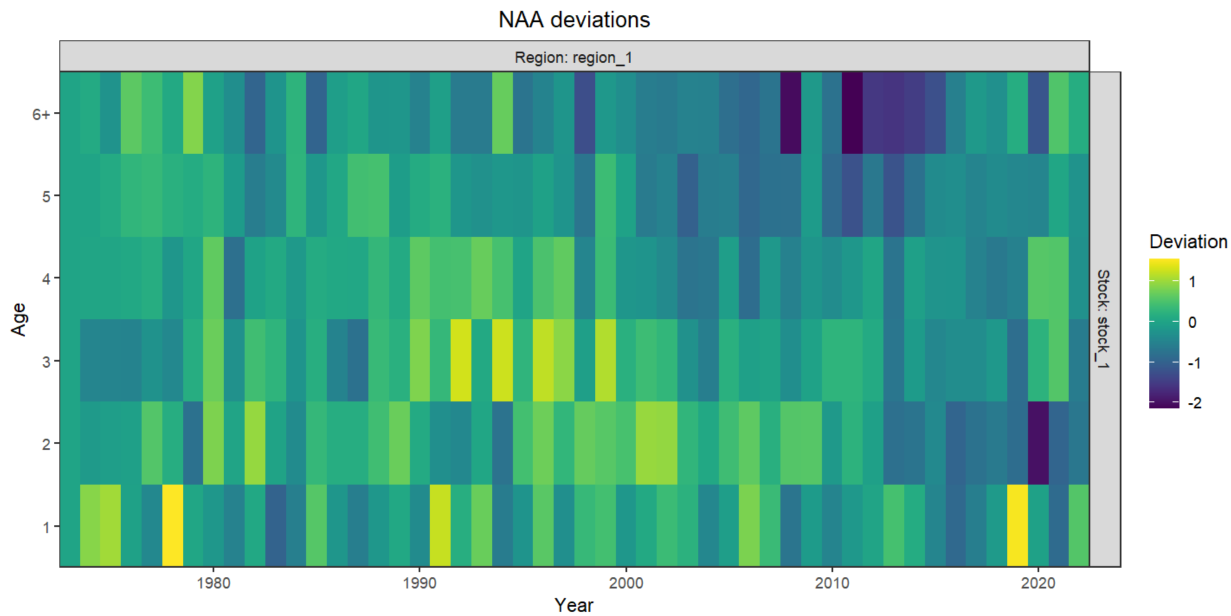

Figure 23: IID process error deviation on numbers at age from optimal WHAM run.

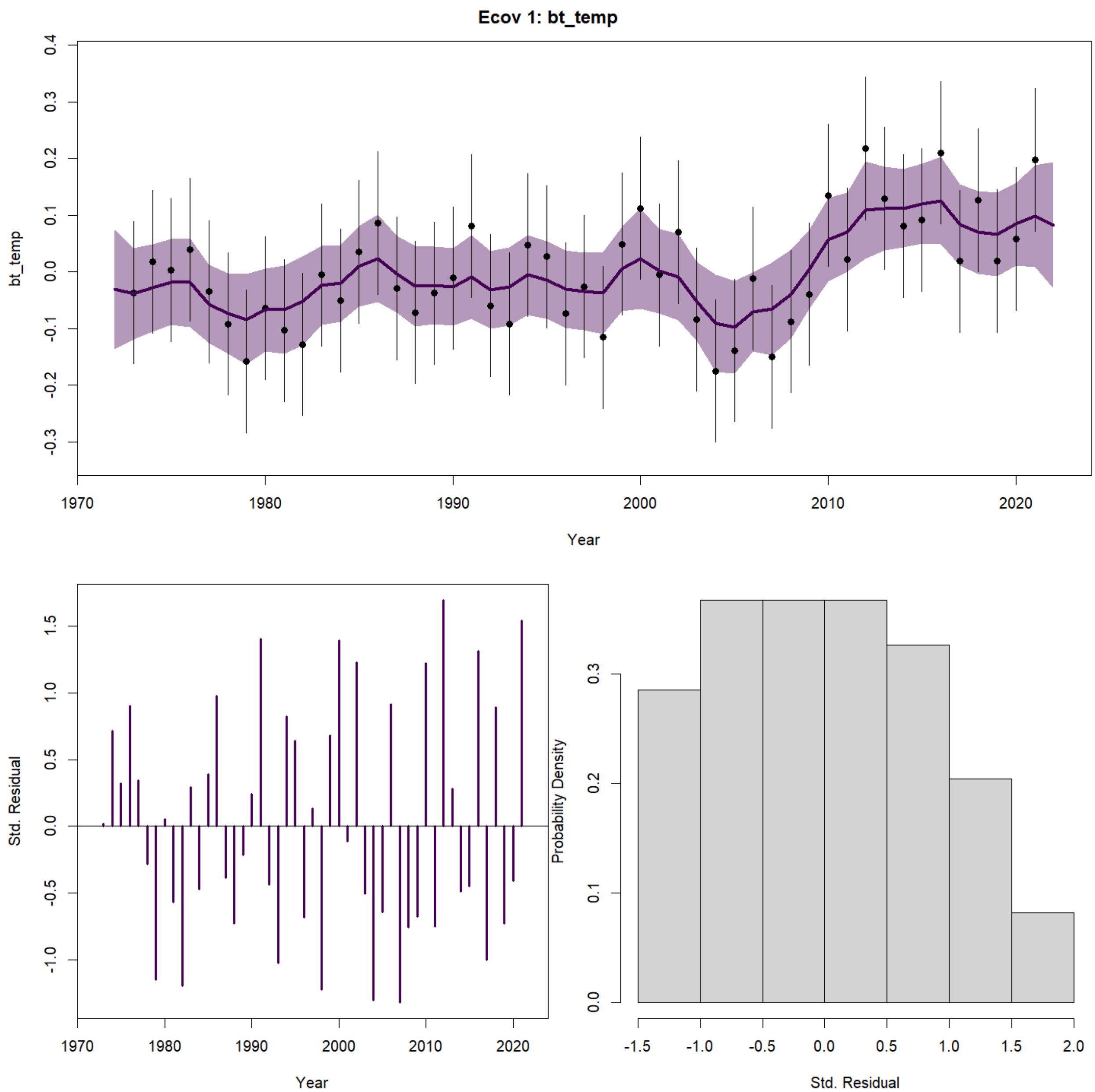

Figure 24: WHAM fit to the bottom water temperature covariate on recruitment.

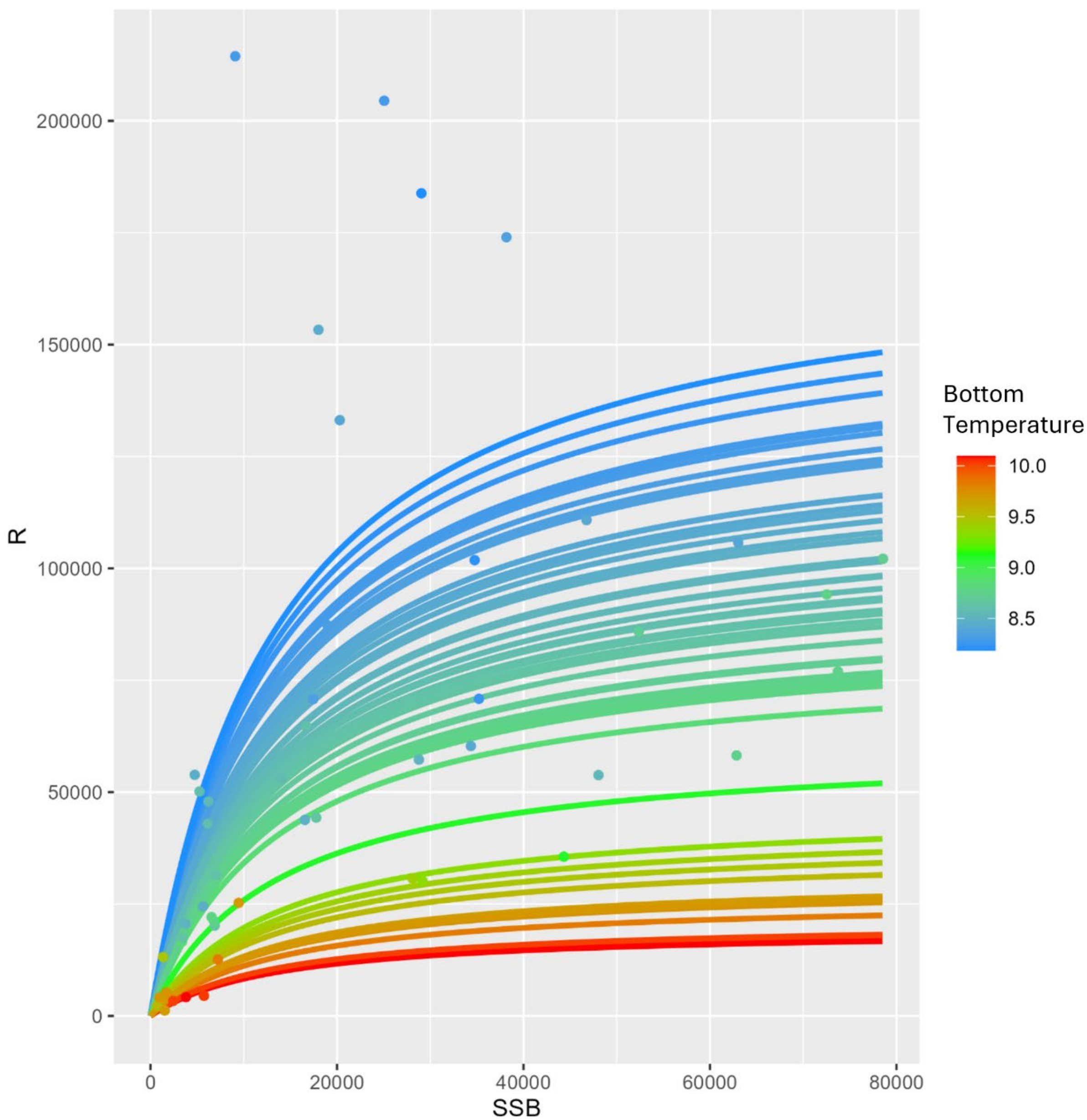

Figure 25: Beverton-Holt stock recruit relationship fit in WHAM.

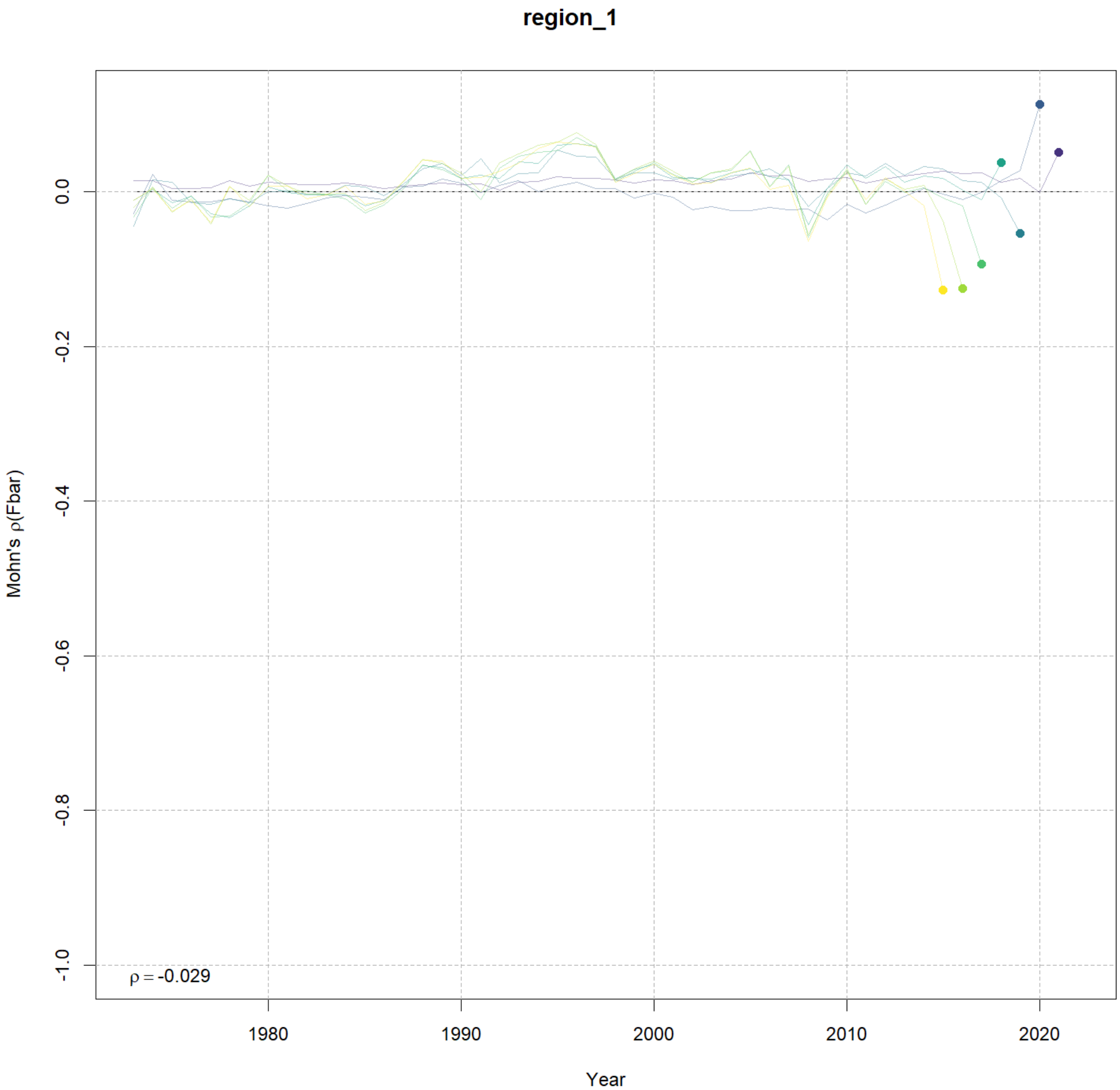

Figure 26: Mohn’s rho for retrospective estimates of fishing mortality.

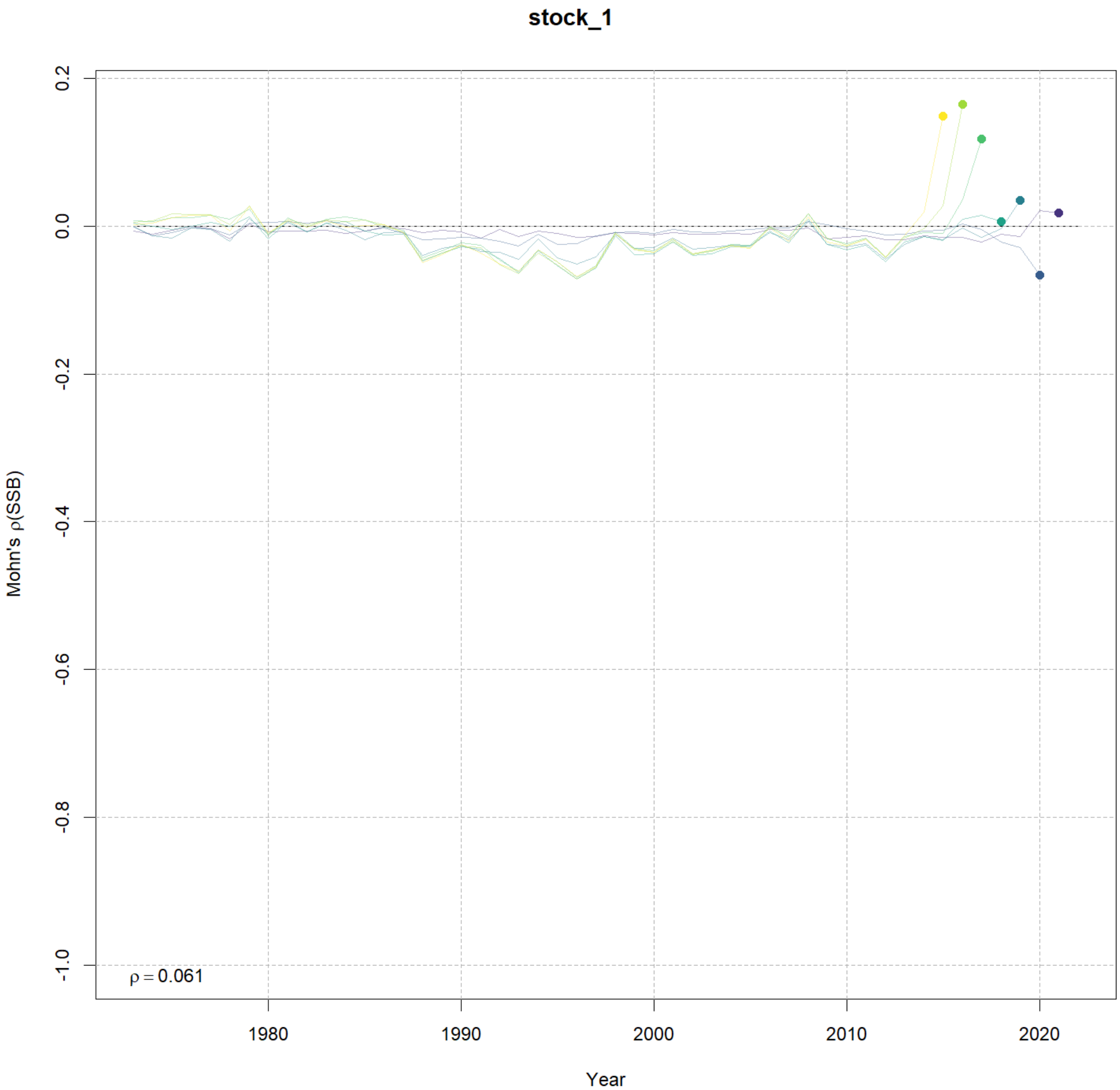

Figure 27: Mohn’s rho for retrospective estimates of spawning stock biomass.

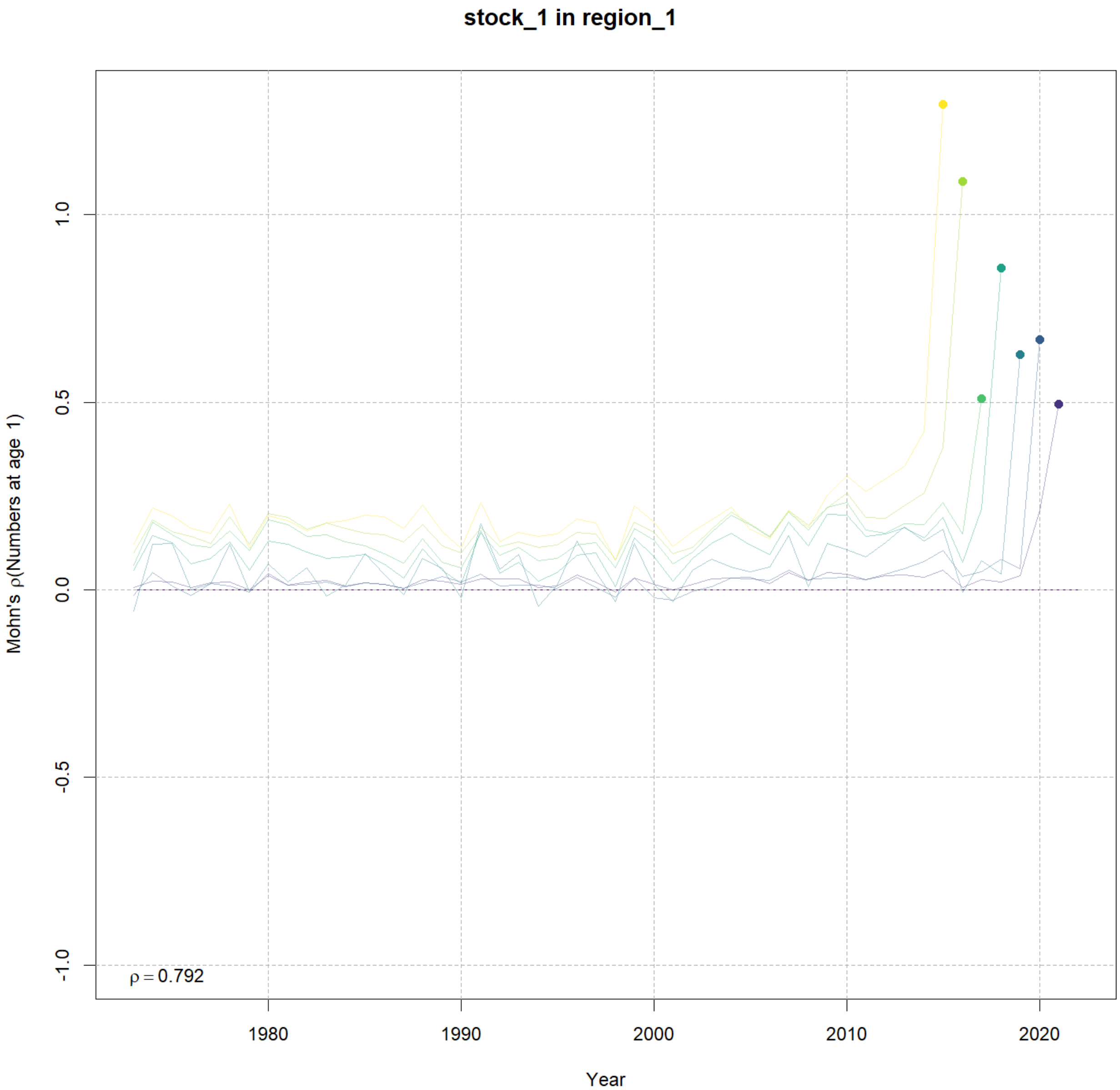

Figure 28: Mohn's rho for retrospective estimates of recruitment.

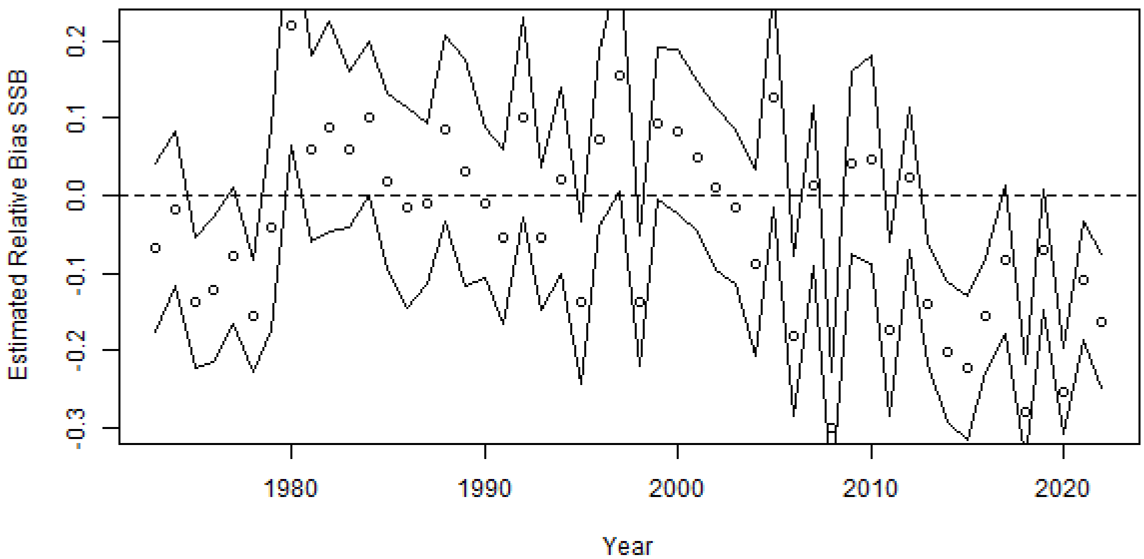

Figure 29: Simulation self-test for WHAM estimates of fishing mortality.

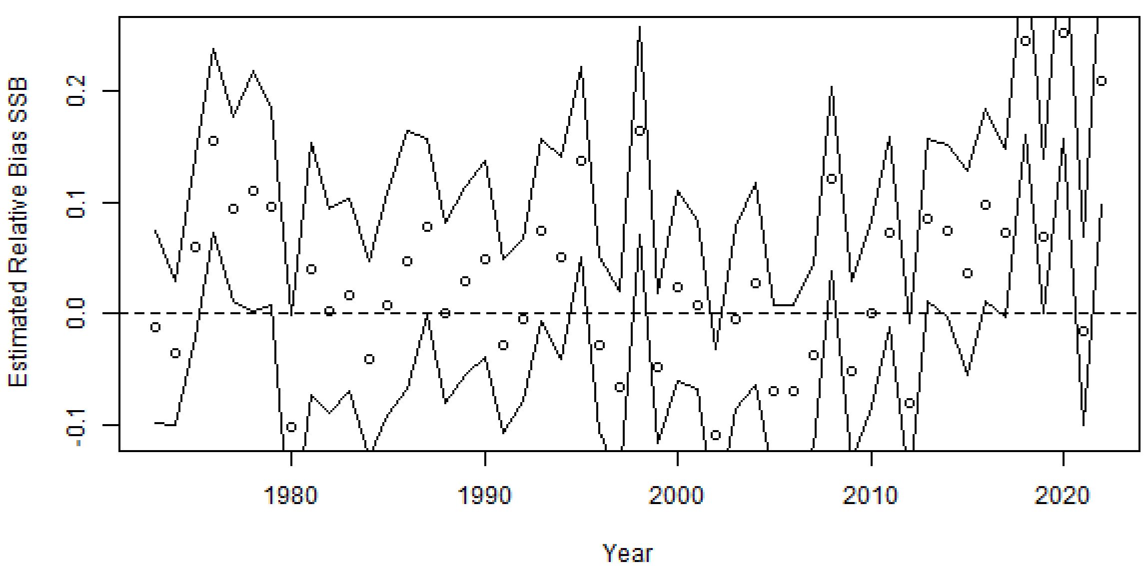

Figure 30: Simulation self-test for WHAM estimates of spawning stock biomass.

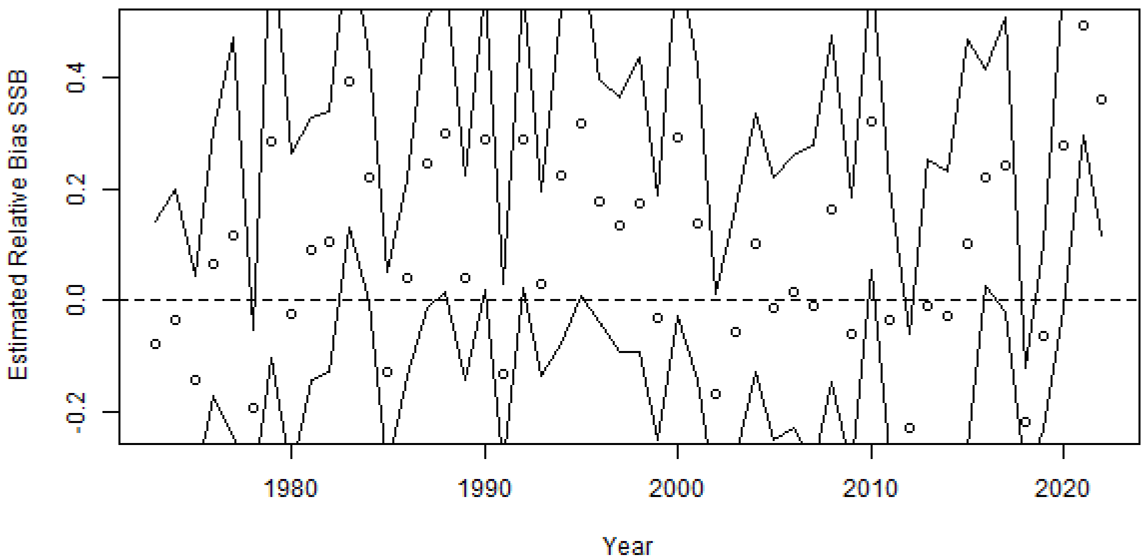

Figure 31: Simulation self-test for WHAM estimates of recruitment.

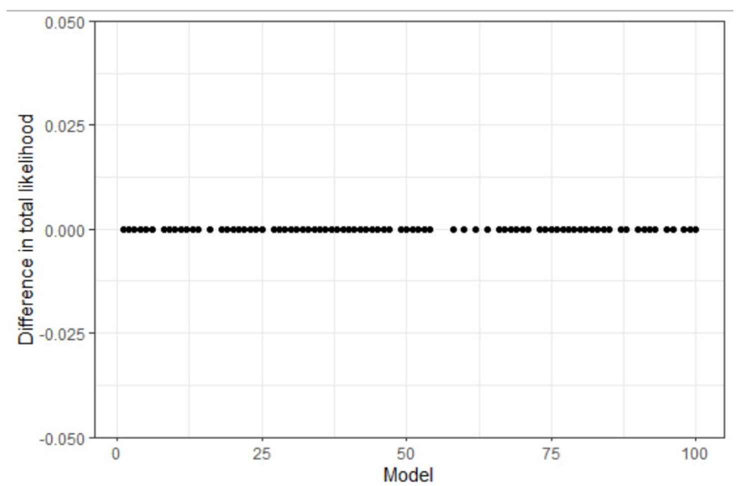

Figure 32: Jitter results from optimal WHAM run.

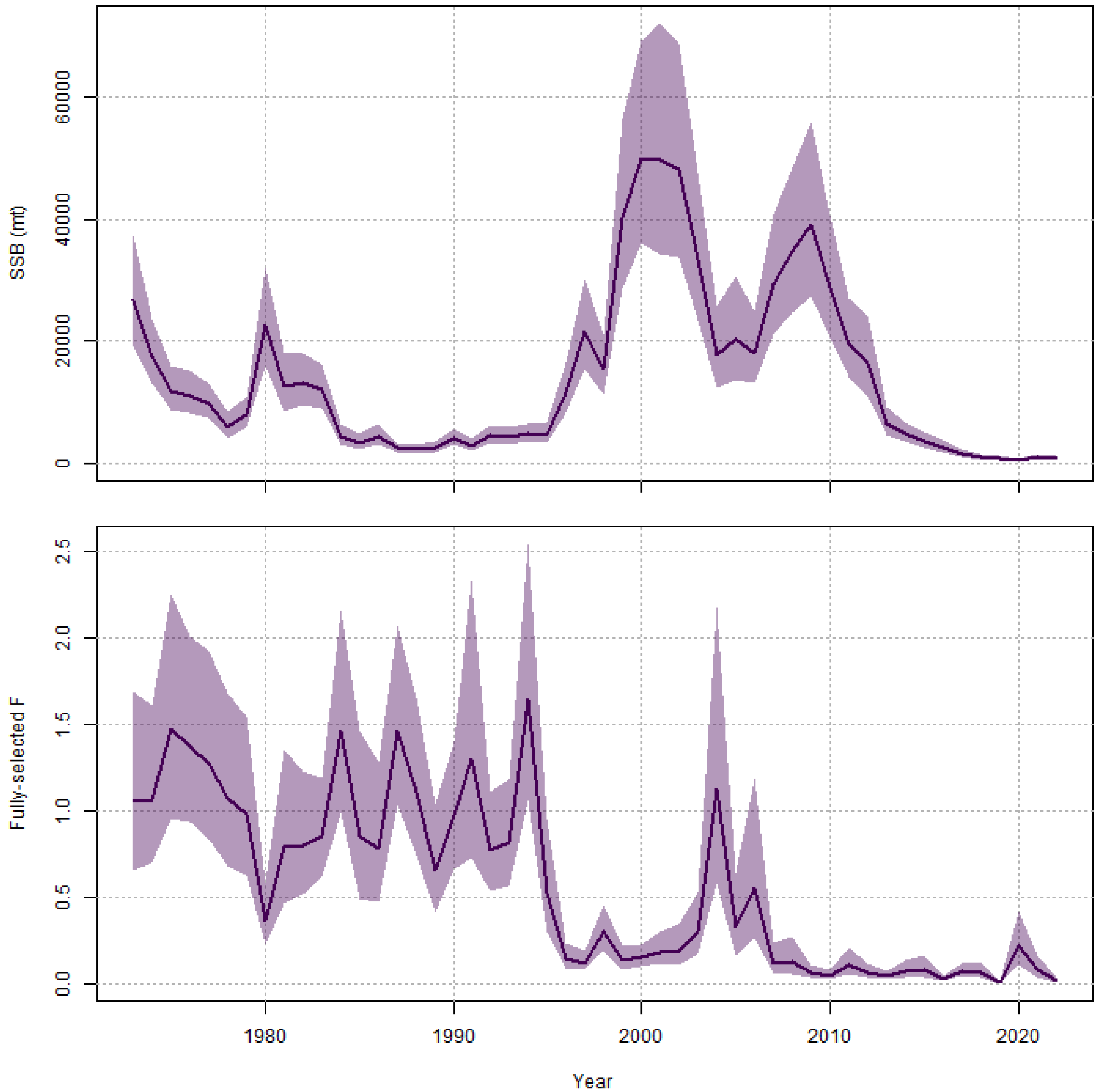

Figure 33: Fishing mortality and spawning stock biomass estimates from WHAM.

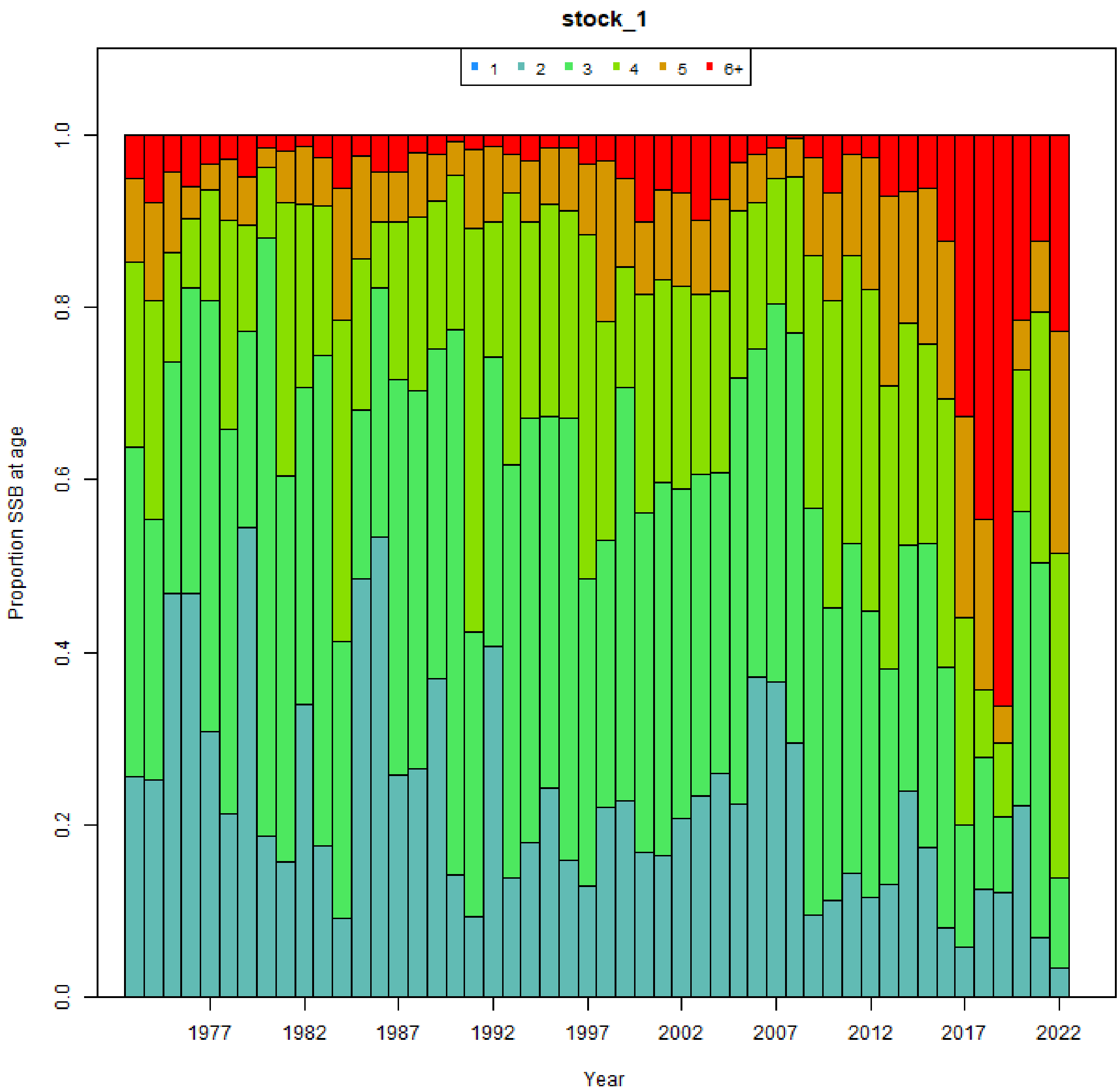

Figure 34: Spawning stock biomass at age estimates from WHAM.

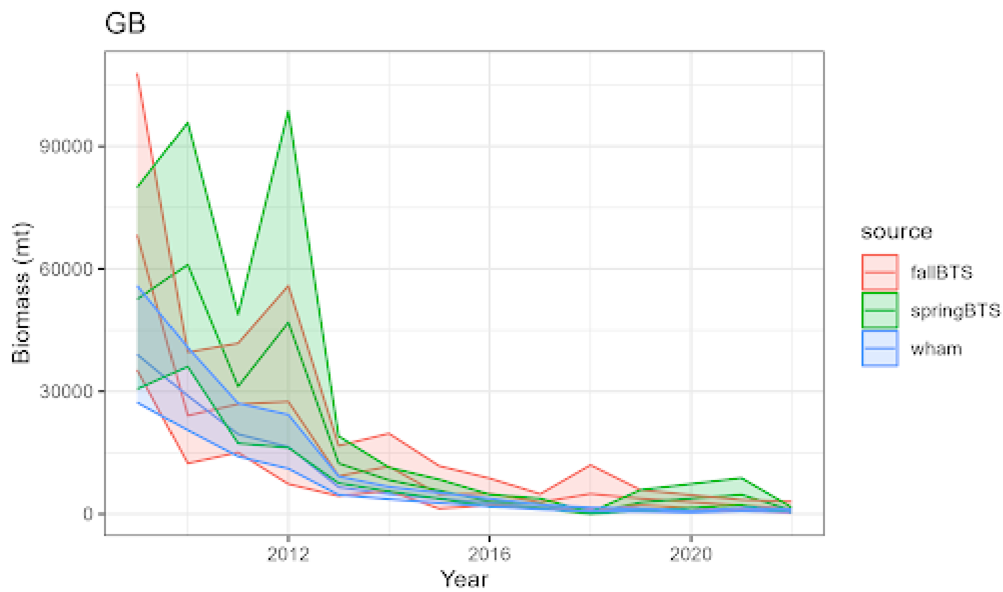

Figure 35: Comparison of biomass estimates from the optimal WHAM model and expanded NEFSC survey biomass.

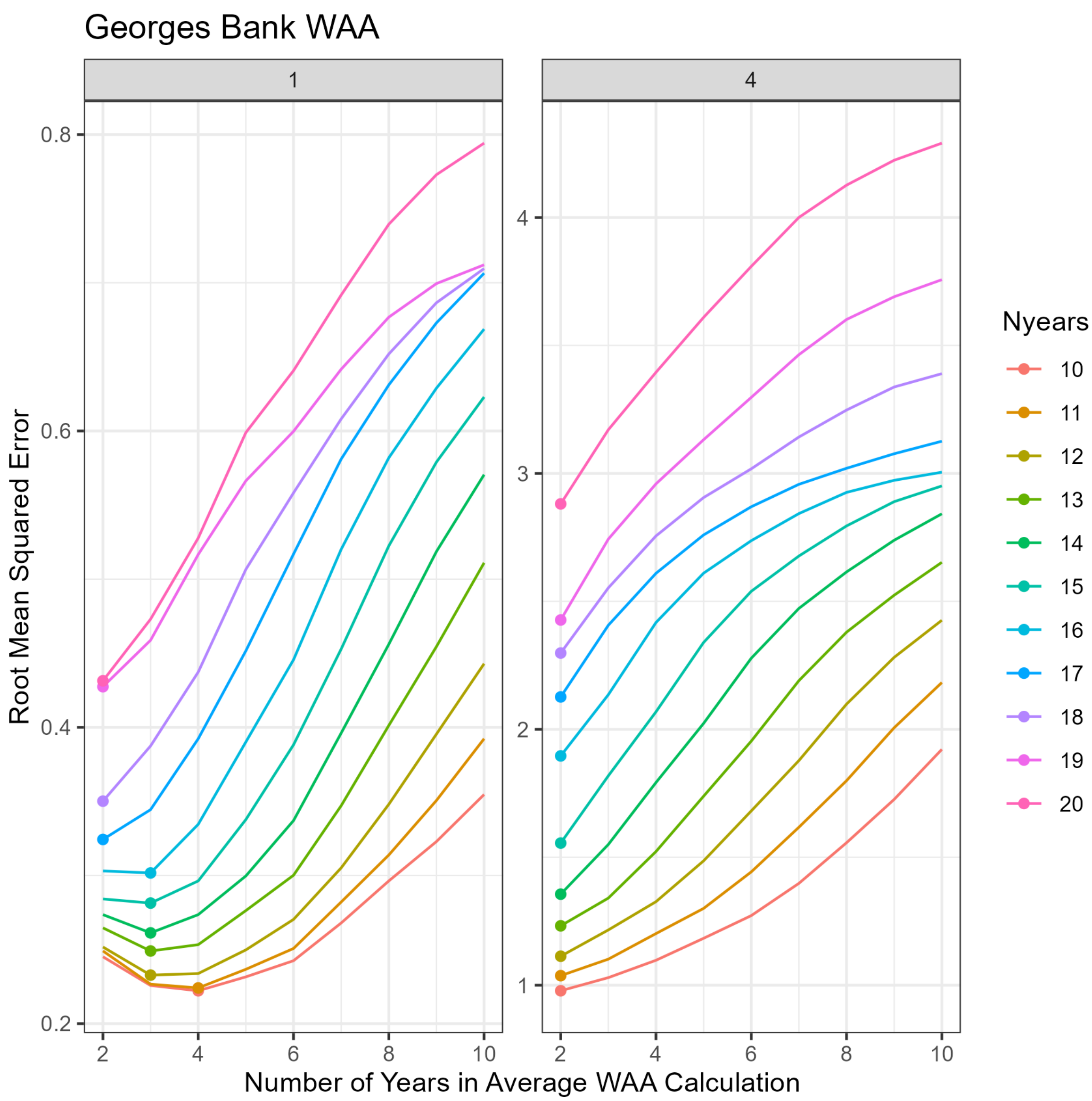

Figure 36: Root mean squared error for the numbers of years used to calculate weight at age.

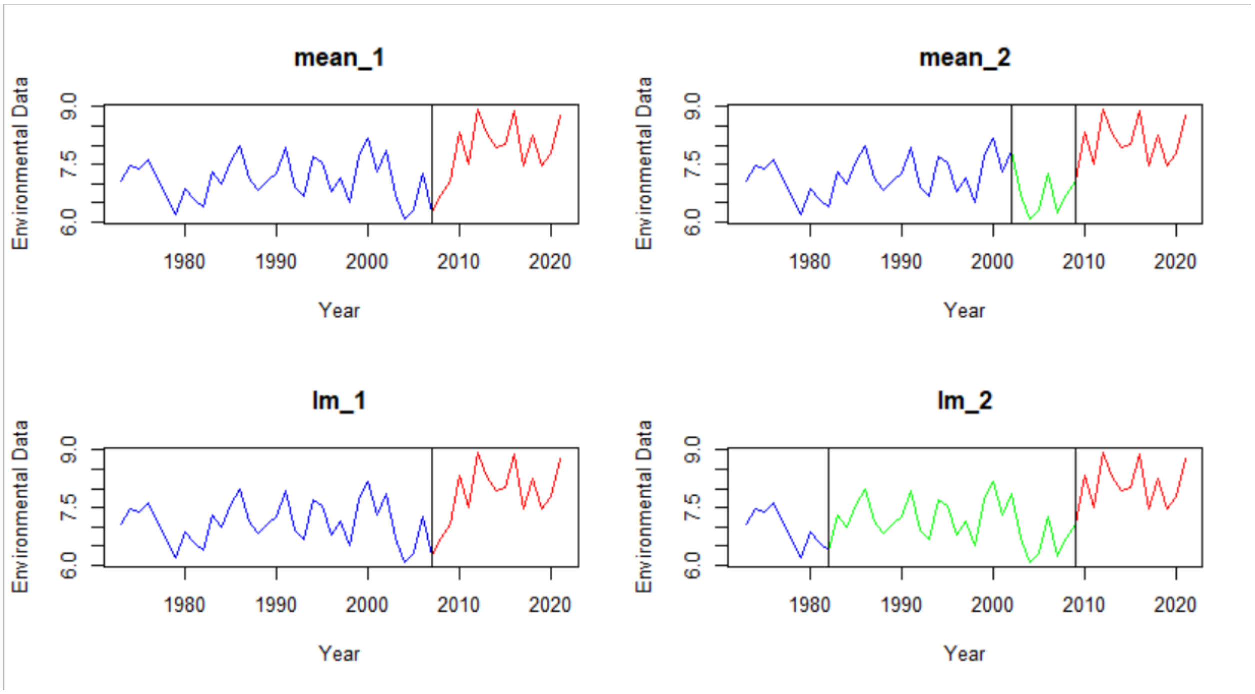

Figure 37: Change point analyses conducted on the bottom water temperature covariate.

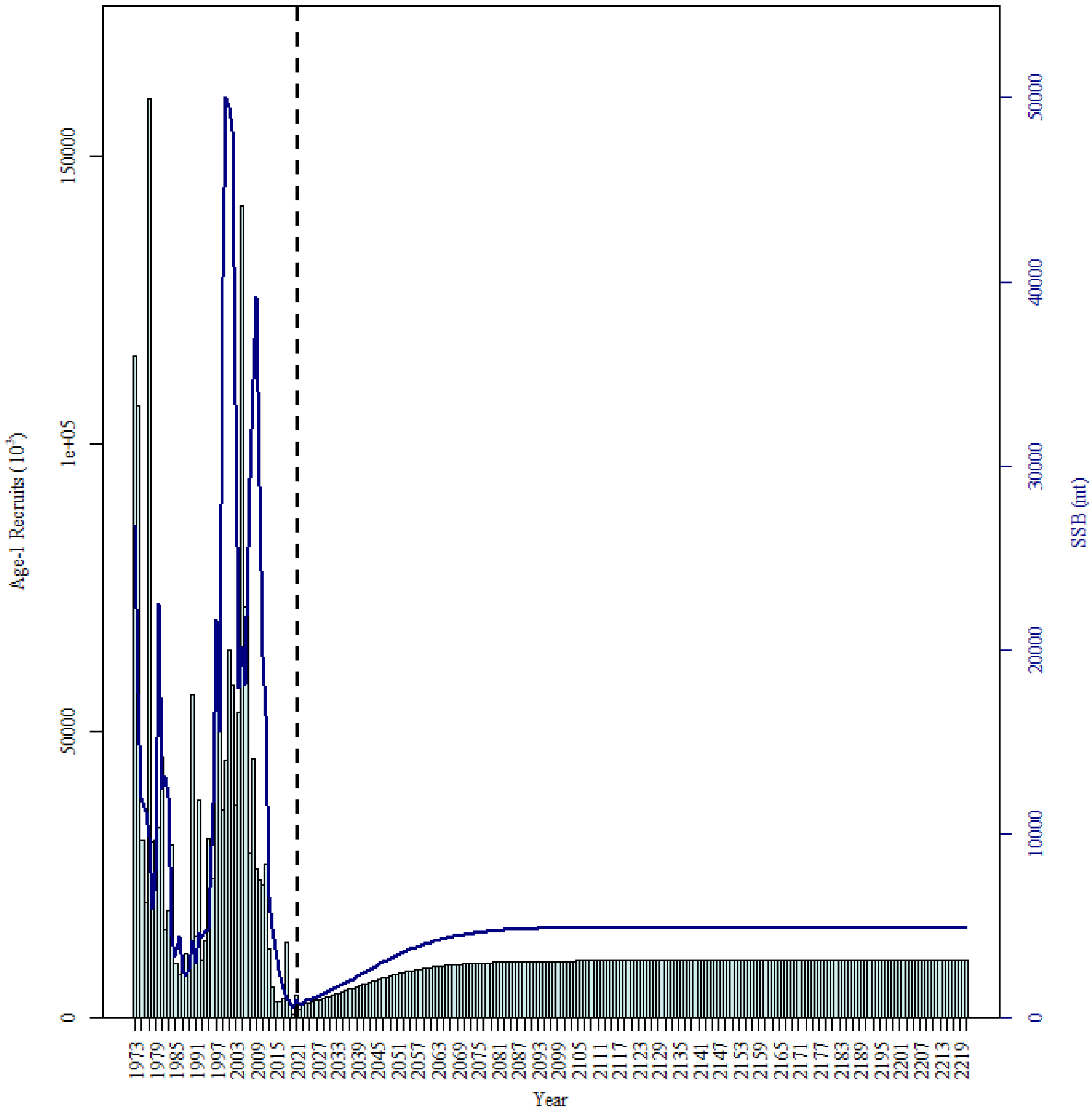

Figure 38: Long-term projections (200 years) of the stock under current environmental conditions and fishing at Fmsy (0.15).

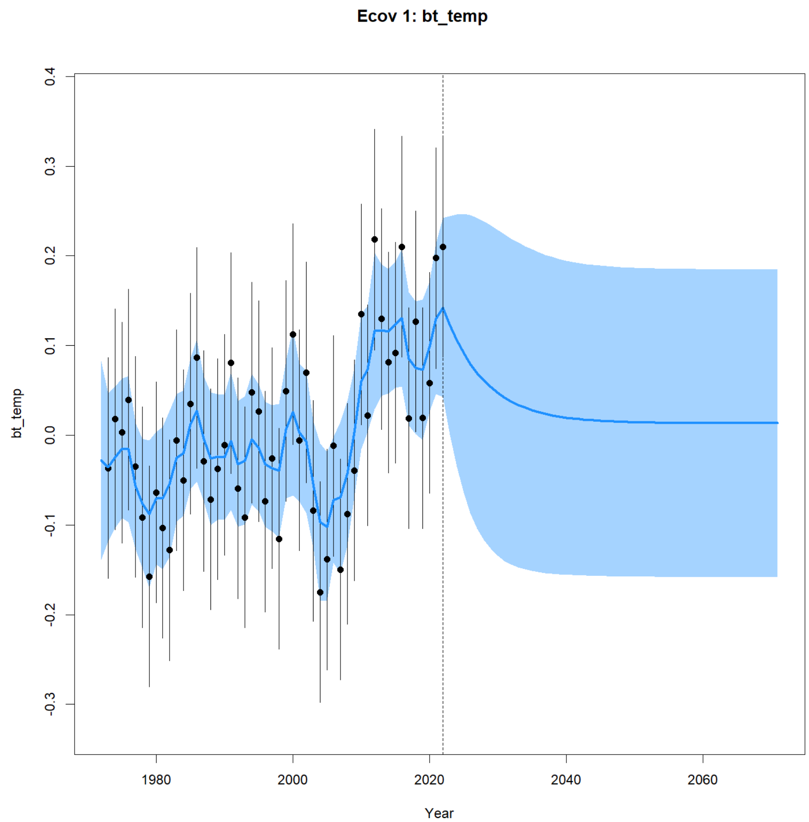

Figure 39: Example of using an AR1 process on bottom temperature in the projection years.

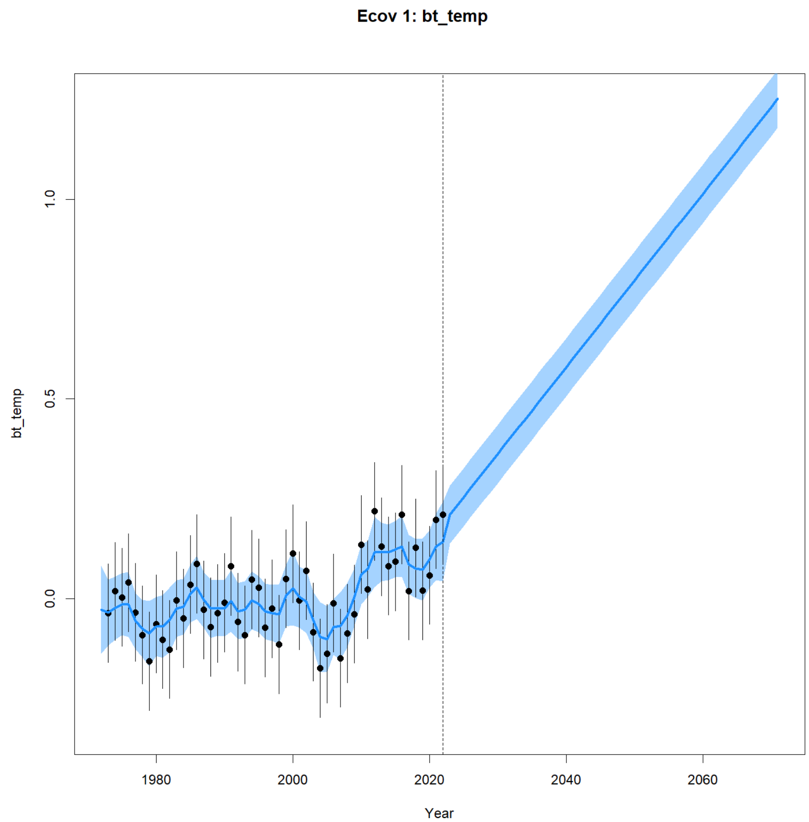

Figure 40: Example of using a linear increase on bottom temperature in the projection years.

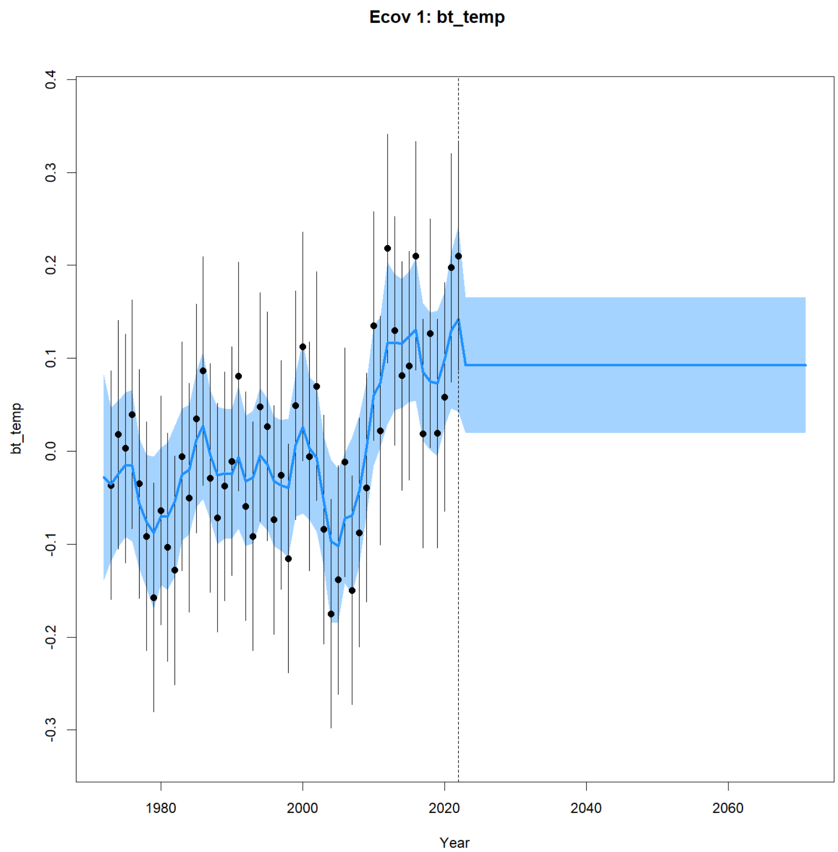

Figure 41: Example of using a AR1 process on bottom temperature in the projection years.