

Draft Working Paper for pre-dissemination peer review only.

Appendix 1 (Primary Appendix) to the Wolffish Report
November 26, 2008

Appendix 1

Wolffish SCALE Model

By
Paul Nitschke
NEFSC

Data Poor Working Group Meeting
Woods Hole, MA
December 8-12, 2008

This information is distributed solely for the purpose of pre-dissemination peer review. It has not been formally disseminated by NOAA. It does not represent any final agency determination or policy.

Introduction

Incomplete or lack of age-specific catch and survey indices often limits the application of a full age-structured assessment (e.g. Virtual Population Analysis and many forward projecting age-structured models). Stock assessments will often rely on the simpler size/age aggregated models (e.g. surplus production models) when age-specific information is lacking. However the simpler size/age aggregated models may not utilize all of the available information for a stock assessment. Knowledge of a species growth and lifespan, along with total catch data, size composition of the removals, recruitment indices and indices on numbers and size composition of the large fish in a survey can provide insights on population status using a simple model framework.

The Statistical Catch At Length (SCALE) model, is a forward projecting age-structured model tuned with total catch (mt), catch at length or proportional catch at length, recruitment at a specified age (usually estimated from first length mode in the survey), survey indices of abundance of the larger/older fish (usually adult fish) and the survey length frequency distributions. The SCALE model was developed in the AD model builder framework. The model parameter estimates are fishing mortality and recruitment in each year, fishing mortality to produce the initial population (Fstart), logistic selectivity parameters for each year or blocks of years and Qs for each survey index.

The SCALE model was developed as an age-structured model that does NOT rely on age-specific information on a yearly basis. The model is designed to fit length information, abundance indices, and recruitment at age which can be estimated by using survey length slicing. However the model does require an accurate representation of the average overall growth of the population which is input to the model as mean lengths at age. Growth can be modeled as sex-specific growth and natural mortality or growth and natural mortality can be model with the sexes combined. The SCALE model will allow for missing data.

Model Configuration

The SCALE model assumes growth follows the mean input length at age with predetermined input error in length at age. Therefore a growth model or estimates of the average mean lengths at age is essential for reliable results. The model assumes static growth and therefore population mean length/weight at age are assumed constant over time. A depiction of model assumed population growth at age using the input mean lengths at age and variation can be seen in table 1 and Figure 1).

The SCALE model estimates logistic parameters for a flattop selectivity curve at length in each time block specified by the user for the calculation of population and catch age-length matrices or the user can input fixed logistic selectivity parameters. Presently the SCALE model can not account for the dome shaped selectivity pattern.

The SCALE model computes an initial age-length population matrix in year one of the model as follows. First the estimated populations numbers at age starting with age-1 recruitment get normally distributed at one cm length intervals using the mean length at age with the assumed standard deviation. Next the initial population numbers at age are calculated from the

previous age at length abundance using the survival equation. An estimated fishing mortality (F_{start}) is also used to produce the initial population. This F can be thought of as the average fishing mortality that occurred before the first year in the model. Now the process repeats itself with the total of the estimated abundance at age getting redistributed according to the mean length at age and standard deviation in the next age ($age+1$).

This two step process is used to incorporate the effects of length specific selectivities and fishing mortality. The initial population length and age distribution is constructed by assuming population equilibrium with an initial value of F , called F_{start} . Length specific mortality is estimated as a two step process in which the population is first decremented for the length specific effects of mortality as follows:

$$N_{a,len,y_1}^* = N_{a-1,len,y_1} e^{-(PR_{len}F_{start} + M)}$$

In the second step, the total population of survivors is then redistributed over the lengths at age a by assuming that the proportions of numbers at length at age a follow a normal distribution with a mean length derived from the input growth curve (mean lengths at age).

$$N_{a,len,y_1} = \pi_{len,a} \sum_{len=0}^{L_\infty} N_{a,len,y_1}^*$$

where

$$\pi_{len,a} = \Phi(len + 1 | \mu_a, \sigma_a^2) - \Phi(len | \mu_a, \sigma_a^2)$$

where

$$\mu_a = L_\infty (1 - e^{-K(a-t_0)})$$

Mean lengths at age can be calculated from a von Bertalanffy model from a prior study as shown in the equation above or mean lengths at age can be calculated directly from an age-length key. Variation in length at age $a = \sigma_s^2$ can often be approximated empirically from the growth study used for the estimation of mean lengths at age. If large differences in growth exist between the sexes then growth can be input as sex-specific growth with sex-specific natural mortality. However catch and survey data are still fitted with sexes combined.

This SCALE model formulation does not explicitly track the dynamics of length groups across age because the consequences of differential survival at length at age a do not alter the

mean length of fish at age $a+1$. However, it does more realistically account for the variations in age-specific partial recruitment patterns by incorporating the expected distribution of lengths at age.

In the next step the population numbers at age and length for years after the calculation of the initial population use the previous age and year for the estimate of abundance. Here the calculations are done on a cohort basis. Like in the previous initial population survival equation the partial recruitment is estimated on a length vector.

$$N_{a,len,y}^* = N_{a-1,len,y-1} e^{-(PR_{len} F_{y-1} + M)}$$

second stage

$$N_{a,len,y} = \pi_{len,a} \sum_{len=0}^{L_{\infty}} N_{a,len,y}^*$$

Constant M is assumed along with an estimated length-weight relationship to convert estimated catch in numbers to catch in weight. The standard Baranov's catch equation is used to remove the catch from the population in estimating fishing mortality.

$$C_{y,a,len} = \frac{N_{y,a,len} F_y PR_{len} (1 - e^{-(F_y PR_{len} + M)})}{(F_y PR_{len}) + M}$$

Catch is converted to yield by assuming a time invariant average weight at length.

$$Y_{y,a,len} = C_{y,a,len} W_{len}$$

The SCALE model results in the calculation of population and catch age-length matrices for the starting population and then for each year thereafter. The model is programmed to estimate recruitment in year 1 and estimate variation in recruitment relative to recruitment in year 1 for each year thereafter. Estimated recruitment in year one can be thought of as the estimated average long term recruitment in the population since it produces the initial population. The residual sum of squares of the variation in recruitment $\sum(Vrec)^2$ is then used as a component of the total objective function. The weight on the recruitment variation component of the objective function ($Vrec$) can be used to penalize the model for estimating large changes in recruitment relative to estimated recruitment in year one.

The model requires an age-1 recruitment index for tuning or the user can assume relatively constant recruitment over time by using a high weight on V_{rec} . Usually there is little overlap in ages at length for fish that are one and/or two years of age in a survey of abundance. The first mode in a survey can generally index age-1 recruitment using length slicing. In addition numbers and the length frequency of the larger fish (adult fish) in a survey where overlap in ages at a particular length occurs can be used for tuning population abundance. The model tunes to the catch and survey length frequency data using a multinomial distribution. The user specifies the minimum size (cm) for the model to fit. Different minimum sizes can be fit for the catch and survey data length frequencies.

The number of parameters estimated is equal to the number of years in estimating F and recruitment plus one for the F to produce the initial population (F_{start}), logistic selectivity parameters for each year or blocks of years, and for each survey Q . The total likelihood function to be minimized is made up of likelihood components comprised of fits to the catch, catch length frequencies, the recruitment variation penalty, each recruitment index, each adult index, and adult survey length frequencies:

$$L_{catch} = \sum_{years} \left(\ln(Y_{obs,y} + 1) - \ln \left(\sum_a \sum_{len} Y_{pred,len,a,y} + 1 \right) \right)^2$$

$$L_{catch_lf} = -N_{eff} \sum_y \left(\sum_{inlen}^{L_{\infty}} \left((C_{y,len} + 1) \ln \left(1 + \sum_a C_{pred,y,a,len} \right) - \ln(C_{y,len} + 1) \right) \right)$$

$$L_{vrec} = \sum_{y=2}^{N_{years}} (V_{rec}_y)^2 = \sum_{y=2}^{N_{years}} (R_1 - R_y)^2$$

$$\sum L_{rec} = \sum_{i=1}^{N_{rec}} \left[\sum_y^{N_{years}} \left(\ln(I_{rec_i,inage_i,y}) - \ln \left(\sum_{len}^{L_{\infty}} N_{y,inage_i,len} * q_{rec_i} \right) \right)^2 \right]$$

$$\sum L_{adult} = \sum_{i=1}^{N_{adult}} \left[\sum_y^{N_{years}} \left(\ln(I_{adult_i,inlen+i,y}) - \left(\sum_a \sum_{inlen_i}^{L_{\infty}} \ln(N_{pred,y,a,len} * q_{adult_i}) \right) \right)^2 \right]$$

$$\sum L_{lf} = \sum_{i=1}^{N_{lf}} \left[-N_{eff} \sum_y \left(\sum_{inlen_i}^{L_{\infty}} \left((I_{lf_i,y,len} + 1) \ln \left(1 + \sum_a N_{pred,y,a,len} \right) - \ln(I_{lf_i,y,len} + 1) \right) \right) \right]$$

In equation L_{catch_lf} calculations of the sum of length is made from the user input specified catch length to the maximum length for fitting the catch. Input user specified fits are indicated with the prefix “in” in the equations. LF indicates fits to length frequencies. In equation L_{rec} the input specified recruitment age and in L_{adult} and L_{lf} the input survey specified lengths up to the maximum length is used in the calculation.

$$Obj\ fcn = \sum_{i=1}^N \lambda_i L_i$$

Lambdas represent the weights to be set by the user for each likelihood component in the total objective function.

Wolffish SCALE Model Configuration and results

Mean lengths at age and variation in mean length at age were based on fish collected during the 1980s from Nelsen and Ross (1992). A Gompertz relationship had the best fit using all ages. We have re-estimated a von Bertalanffy relationship using data limited to fish older than 4 with L-infinity fixed at 110 cm (Figure 2). The mean lengths from Nelsen and Ross’s Gompertz relationship for fish younger than age 5 were also used in the SCALE model. The mean lengths from the younger fish does not have a large effect on the SCALE model results. In the final growth model we fixed L-infinity (110) at a slightly higher value than what was estimated by the Gompertz (98.9) model because few larger and older fish exist in Nelsen and Ross’s study and the SCALE model had difficulty predicting larger fish that are in seen in the catch length frequency distributions. A North Sea wolffish growth study estimated L-infinity at 111 for males and 115 for females (Liao and Lucas, 2000). Figure 3 shows the predicted catch length distribution under low Fs ($F=0.001$) assuming different L-infinities. A standard deviation of 6 was used for fish older than age-7. The assumed variation around the mean lengths at age can be seen in Table 1 and Figure 1. Nelson and Ross’s oldest fish was 22 years. The age matrix was dimensioned from ages 1 to 30 with an assumed natural mortality of 0.15.

Only one recruitment index exists in the SCALE model (Figure 4). The spring NEFSC survey shows a distinct mode between 1 and 7 cm. This index was tuned to age-1. The recruitment index suffers from zero catches in many years and at times in blocks of several years. A 40+ cm index was developed from the NEFSC spring, NEFSC fall and the MDMF spring survey (Figure 5). All three surveys show declining trends in abundance with the indices also suffering from zero catches at the end of the time series. The survey length frequency

distributions are limited due to the low numbers of wolffish caught in the surveys. There is concern that biomass may have fallen below detection in the surveys. Preliminary evidence suggests the Bigelow survey may also suffer from the same low catchability issue. Survey indices were scaled using the approximate area of survey coverage divided by the average coverage of a survey tow (Table 2). The area swept estimates can provide some insight from estimated survey efficiencies using the estimated Q_s in the SCALE model.

Zero catches were set to missing in the SCALE model. Setting zeros to the smallest value in the time series appears to have a large unsubstantiated influence on the model results. The age-1 recruitment series was given a relatively low weight (Table 3). Setting the weight to high on the recruitment index will force SCALE to fit the recruitment index very closely but the model is less constrained in estimating recruitment for years where recruitment information is missing which can produce unrealistic results. The age-1 index was used more as a guide with setting the penalty on recruitment variation. The penalty on recruitment variation was set high enough to produce recruitment variation within the bounds of what was observed in the recruitment index. The model has to estimate a declining trend in recruitment to fit the decline in the 40+ cm indices and the declining trend in the catch since 1983. The recruitment index was used as guidance on whether recruitment failure has occurred for the wolffish stock. Sensitivity of the model to the weighting on the recruitment index and the penalty on variation in recruitment can be seen in Figures 6 through 9.

The catch length frequency distributions are an important component of the SCALE model. Observer trawl kept length sampling and port samples were combined to characterize the catch size distributions. Catch length frequency information exists from 1982 to 1985 and from 2001 to 2007. A single selectivity block over the time series was used due to the lack of a distinct shift in the size distribution and due to the lack of size information in many years. There is no indication of size truncation in the catch length frequency distributions over time.

The lack of data prevents the SCALE model from estimating a reliable logistic selectivity curve. The SCALE model estimates a very flat selectivity curve that produces a L-50 at very large sizes. There is a tradeoff in the SCALE model between the estimated selectivity and fishing mortality rates. Two different selectivity regimes were chosen to determine its influence of stock status determination (Figure 10). Run one had a relatively flat selectivity curve which was allowed to hit the L-50 bound of 90 cm. Run two was setup to hit the slope parameter bound of 0.15 which produces a steeper selectivity function with a lower L-50 estimate. Results of the two selectivity runs are summarized in Figure 11-14 and Table 3.

The SCALE model time series starts in 1968 with the beginning of the NEFSC spring index. The SCALE model estimates virgin conditions at the beginning of the time series with a low F_{start} estimate (0.001) in 1968 when the catch was low. A strong retrospective pattern did not exist with the Slope = 0.15 run (Figure 15). The sensitivity of the assumed L-infinity for growth on the model estimated F_s and recruitment can be seen in Figure 16.

Non-parametric biological reference points (BRP) were developed for both the selectivity L-50 = 90 run (run 1) and the slope = 0.15 run (run 2) within the SCALE model using $F_{40\%}$ as a Proxy from F_{msy} (Table 4). A range of knife edge maturities values were used in estimating the BRPs. Maturity as 40+ cm fish was used to correspond to NEFSC survey maturity results, a 65+ cm and 75+ cm cutoffs were used as bounds taken from the Gunnarsson et al (2006) and Templeman (1986). Templeman found maturation occurring at larger sizes in lower latitudes. However Gunnarsson et al (2006) found maturation occurring at larger sizes in the colder waters

on the eastern side of Iceland compared to the western side. The working group suggests that F50% may be a better proxy of Fmsy for a species that is long lived, late maturing, and has low fecundity. F50% BRPs were developed for the Slope = 0.15 run (Table 4 and Figure 17). The F50% BRPs are based on run 3 incorporating some minor fixes to the catch and catch length frequency (1986) data which were found after the working group meeting (Figures 18-20 and Appendix 2). All SCALE model results suggest the stock in 2007 is at a low biomass (26% to 45% of Bmsy). The overfishing status determination was more uncertain with F₂₀₀₇ to Fmsy ratios ranging from 56% to 158%.

Table 2. Survey area coverage, estimated average survey tow coverage, total area divided by the survey footprint and the survey efficiency q estimates for run 1 and 2.

Wolffish	NEFSC			MDMF
	Spr Age 1	Spr 40+	Fall 40+	40+
survey area (nm ²)	25,911	25,911	25,911	1,833
Avg tow area swept	0.0112	0.0112	0.0112	0.003846
Tow duration	30 min	30 min	30 min	20 min
total area / tow area swept	2,313,482	2,313,482	2,313,482	476,573
Q L50 = 90	0.303	0.400	0.203	0.023
Q Slope = 0.15	0.305	0.387	0.196	0.022

Table 3 Wolffish working group SCALE runs. Run 1 was allowed to hit the L-50 bound on selectivity and run 2 hit the selectivity slope bound of 0.15. Run 3 fixed some minor catch and catch length frequency errors and was used to develop F50 BRPs.

Run	1			2			3		
	L ₅₀ = 90			slope = 0.15			slope = 0.15 (Updated F50% run)		
	weight	qs	Residuals or parameters	weight	qs	Residuals or parameters	weight	qs	Residuals or parameters
total objective function			250.06			253.55			252.57
total catch	10		0.22	10		0.22	10		0.20
catch len freq 1+	500		11.26	500		10.99	500		10.29
Variation in recruit penalty (Vrec)	2		14.12	2		14.80	2		14.85
NEFSC Spr 1 Age-1 1968-2007	2	0.303	8.60	2	0.305	8.83	2	0.310	8.80
NEFSC Spr 40+ 1968-2007	12	0.400	5.78	12	0.387	5.93	12	0.391	5.93
MDMF Spr 40+ 1978-2007	3	0.023	9.70	3	0.022	9.61	3	0.022	9.64
NEFSC Fall 40+ 1968-2007	3	0.203	26.62	3	0.196	26.78	3	0.198	26.70
NEFSC Spr 40+ len freq	5		12.83	5		12.82	5		12.82
Fstart			0.004			0.001			0.001
recruitment year 1 (1968, 000s)			171			175			172
Selectivity Alpha (L50) 1982-1984			90.00			72.89			72.93
Selectivity Beta (slope) 1982-1984			0.09			0.15			0.15

Table 4. Estimated biological reference points based on F40 and F50 for three wolffish SCALE runs. A range of knife edge maturity cutoffs were used (40, 65, and 75 cm).

SCALE run	1			2			3		
	L ₅₀ = 90			slope = 0.15			slope = 0.15		
Selectivity	40	65	75	40	65	75	40	65	75
Length of maturity	40	65	75	40	65	75	40	65	75
F _{MSY} proxy	F _{40%}	F _{40%}	F _{40%}	F _{40%}	F _{40%}	F _{40%}	F _{50%}	F _{50%}	F _{50%}
F _{MSY}	0.70	0.51	0.39	0.35	0.25	0.20	0.195	0.154	0.128
F _{max}	> 0.8	> 0.8	> 0.8	0.60	0.60	0.60	0.60	0.60	0.60
YPR	0.871	0.841	0.809	0.854	0.829	0.788	0.854	0.829	0.788
SSB per Recruit	5.987	5.247	4.686	5.792	5.166	4.548	5.792	5.166	4.548
Initial Recruits (000s)	171	171	171	175	175	175	172	172	172
MSY (mt)	149	144	138	149	145	138	135	125	117
SSB _{MSY} (mt)	1,024	898	802	1,011	902	794	1,314	1,171	1,042
SSB ₀₇ (mt)	405	293	209	457	339	249	447	330	242
F ₀₇	0.516	0.516	0.516	0.195	0.195	0.195	0.202	0.202	0.202
SSB ₀₇ /SSB _{MSY}	40%	33%	26%	45%	38%	31%	34%	28%	23%
F ₀₇ /F _{MSY}	74%	101%	132%	56%	78%	98%	104%	131%	158%

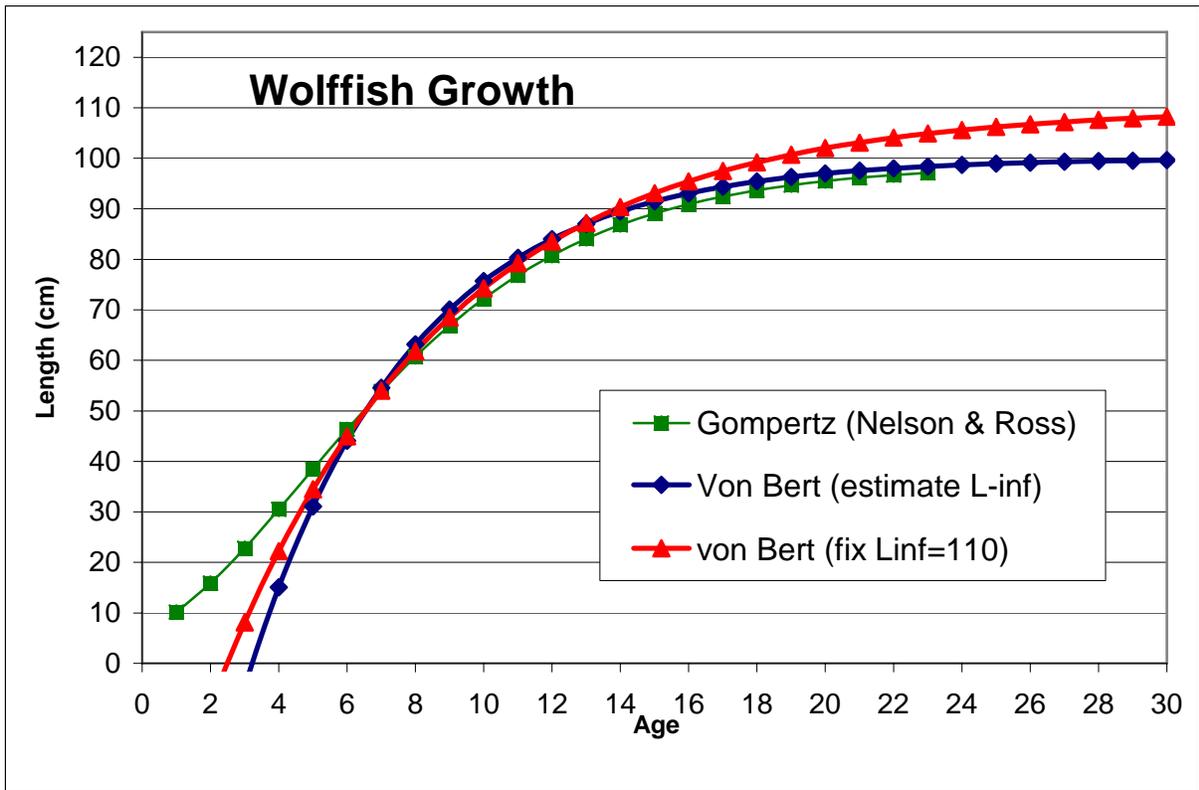


Figure 2. Wolffish estimated growth from Nelson and Ross (1992), von Bertalanffy model limited to 5+ fish, and von Bertalanffy model limited to 5+ fish with fixed L-infinity at 110 cm.

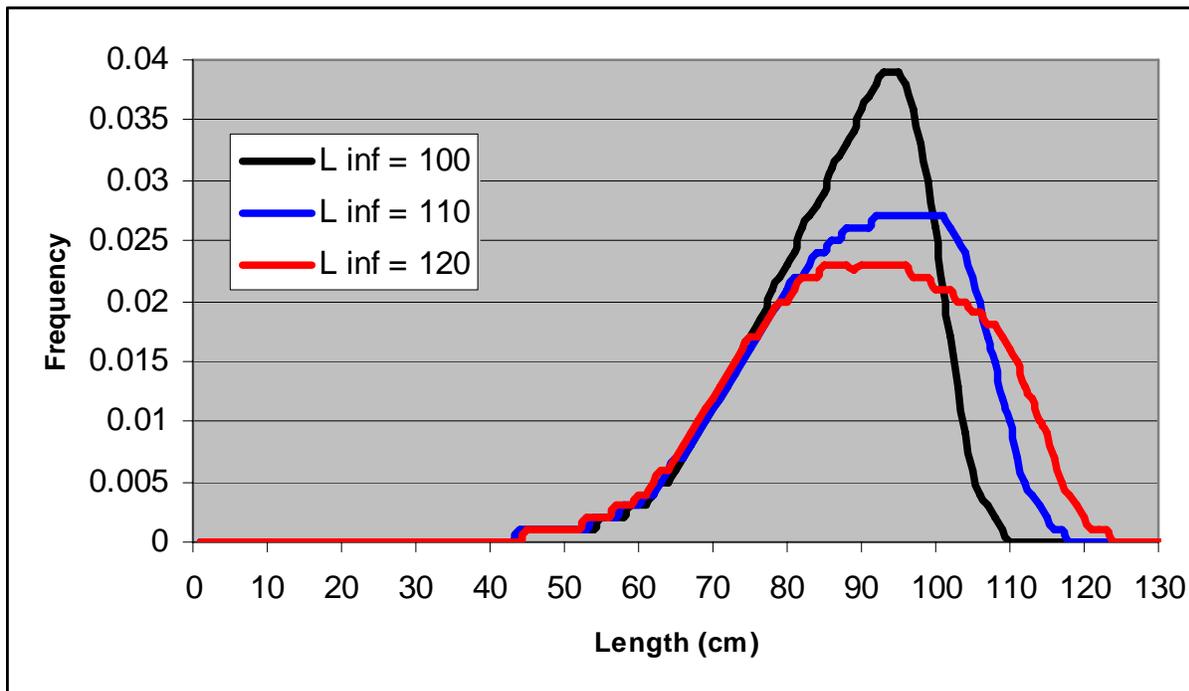


Figure 3. Predicted catch length frequency distributions at low fishing mortality ($F = 0.001$) with different assumed L -infinity values for growth.

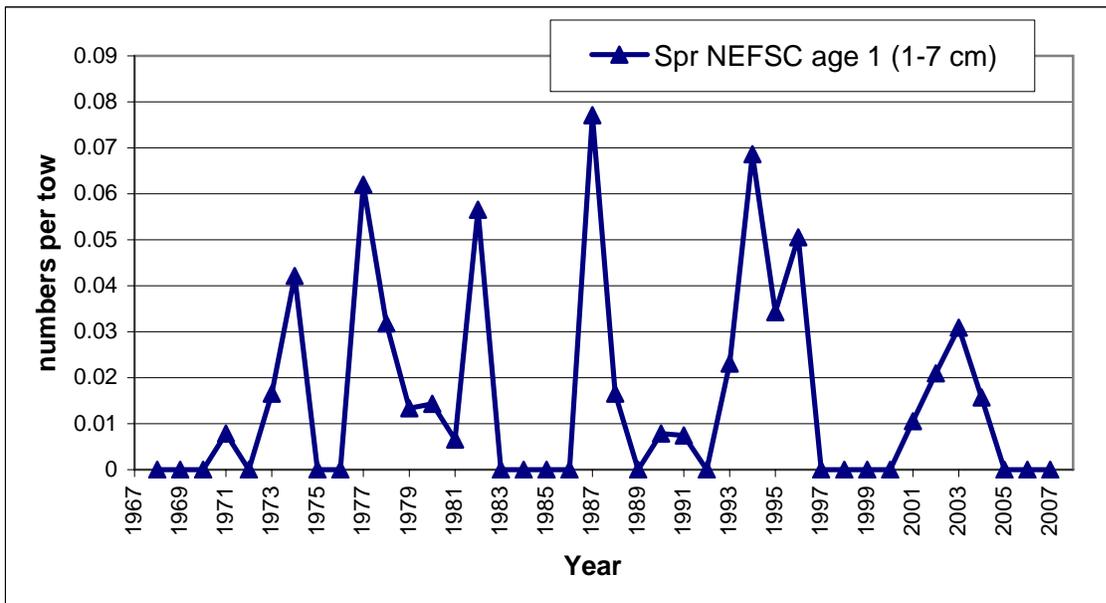


Figure 4. NEFSC spring age-1 stratified mean numbers per tow index. Lengths 1-7 cm was used as a proxy for age-1.

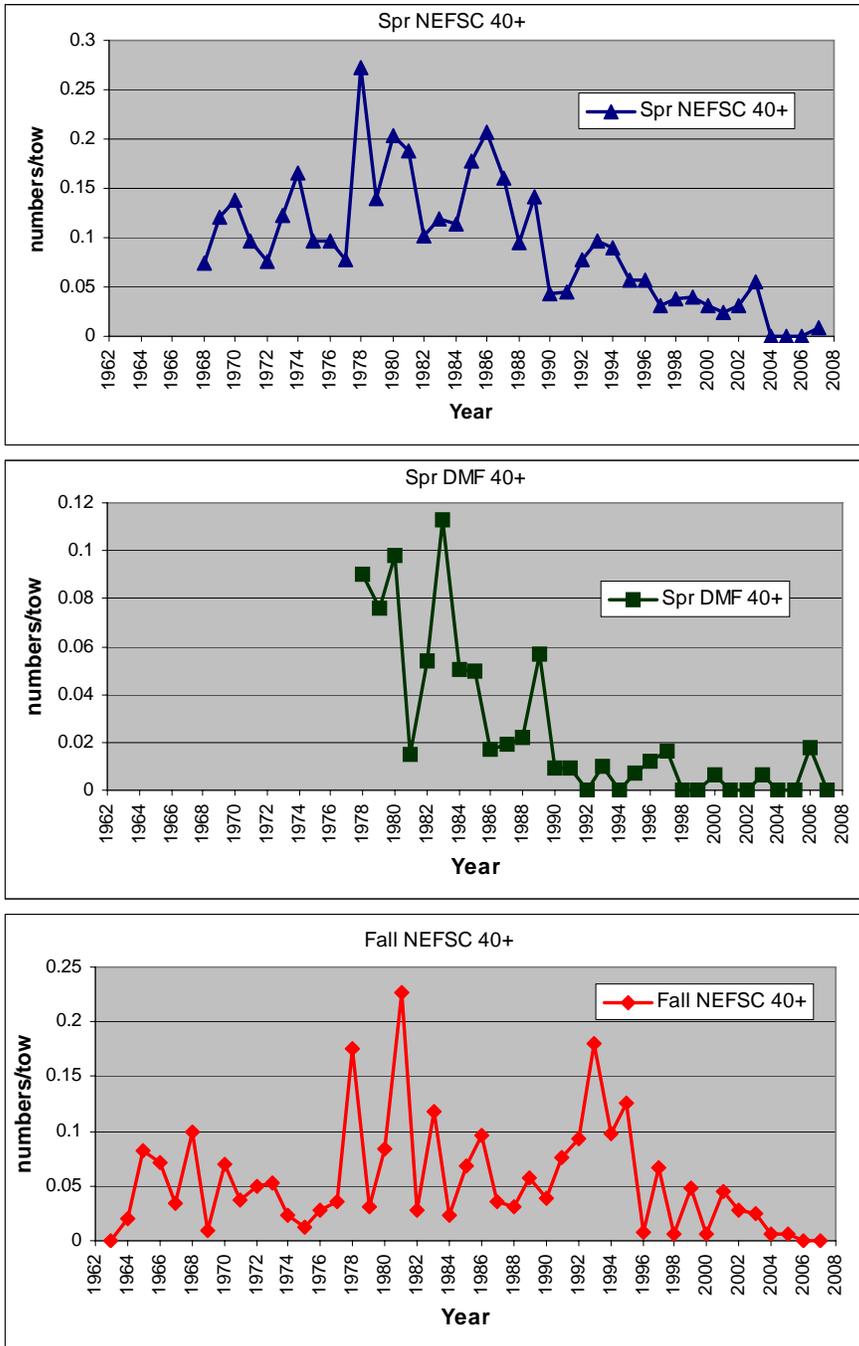


Figure 5. NEFSC spring 40+ cm, MDMF spring 40+ cm, and NEFSC fall 40+ cm stratified numbers per tow survey indices for wolffish.

Slope = 0.15 run (Spr age 1 = 2)

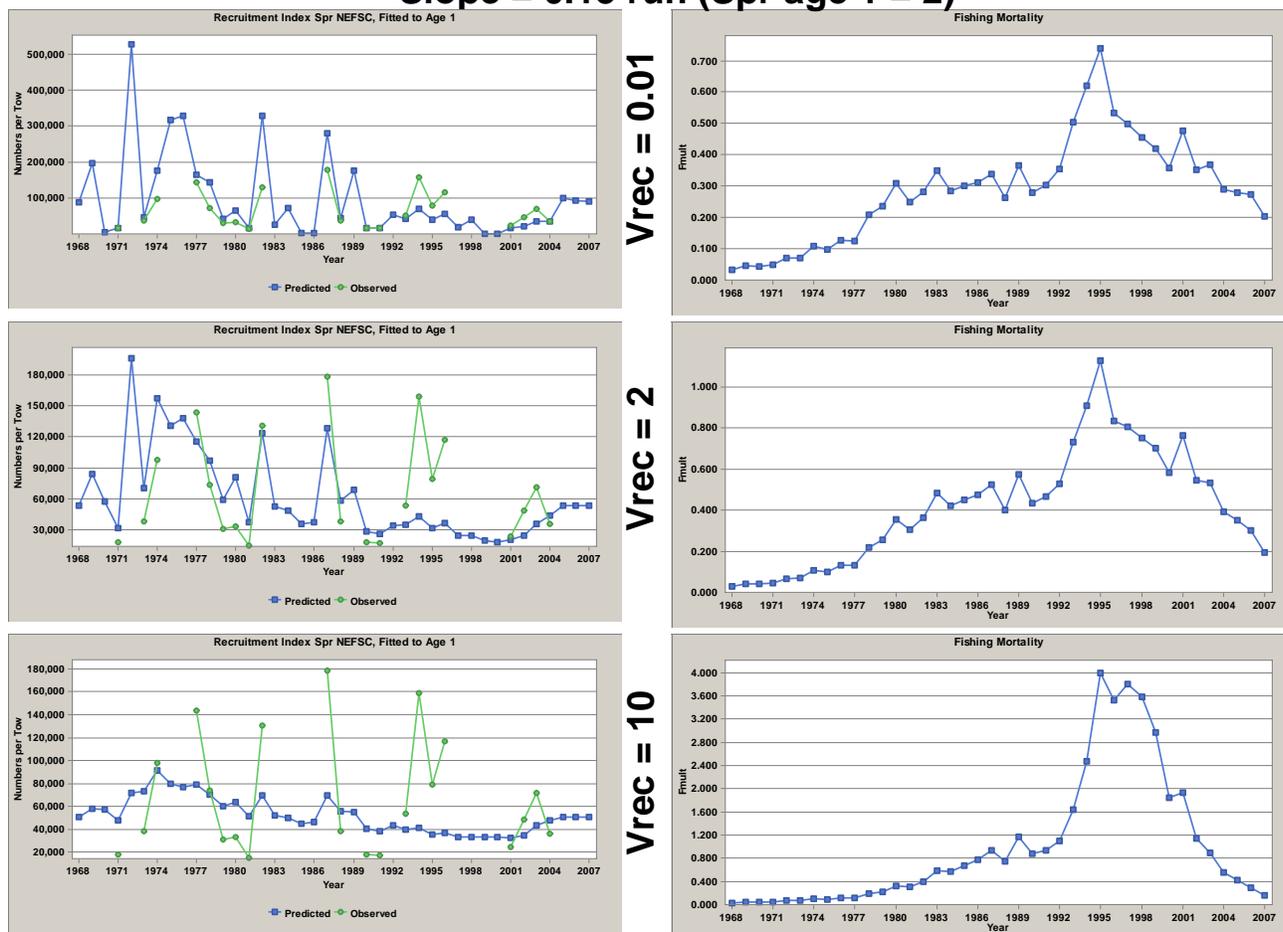


Figure 6. SCALE model sensitivity of fitting the recruitment index and the estimated fishing mortality with different penalty weights on recruitment variation (0.01, 2, 10). The weight on the age-1 recruitment index was fixed at 2.

Slope = 0.15 run (Spr age 1 = 2)

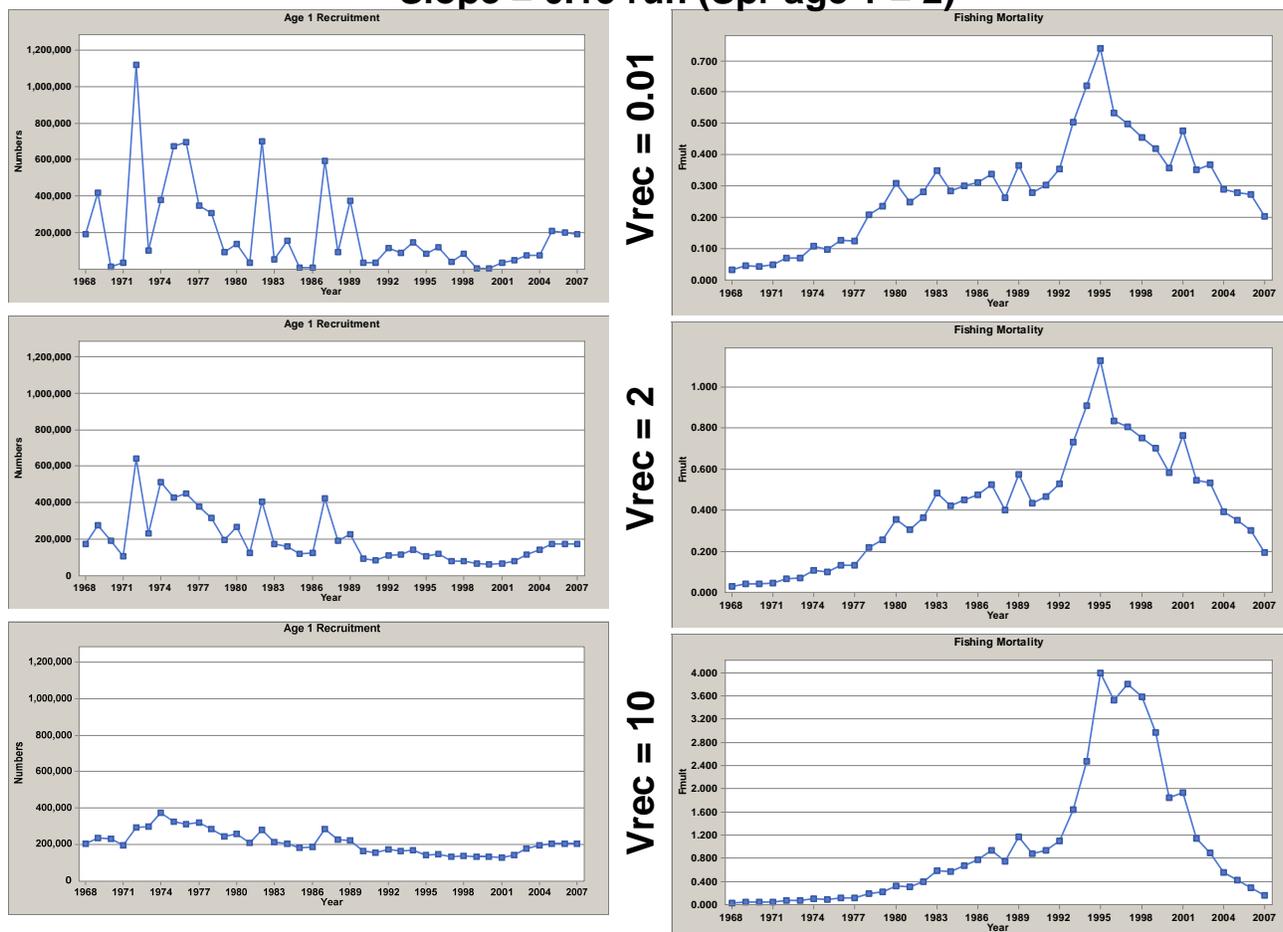


Figure7. SCALE model sensitivity of estimated recruitment and fishing mortality with different penalty weights on recruitment variation (0.01, 2, 10). The weight on the age-1 recruitment index was fixed at 2.

Slope = 0.15 run (Vrec = 2)

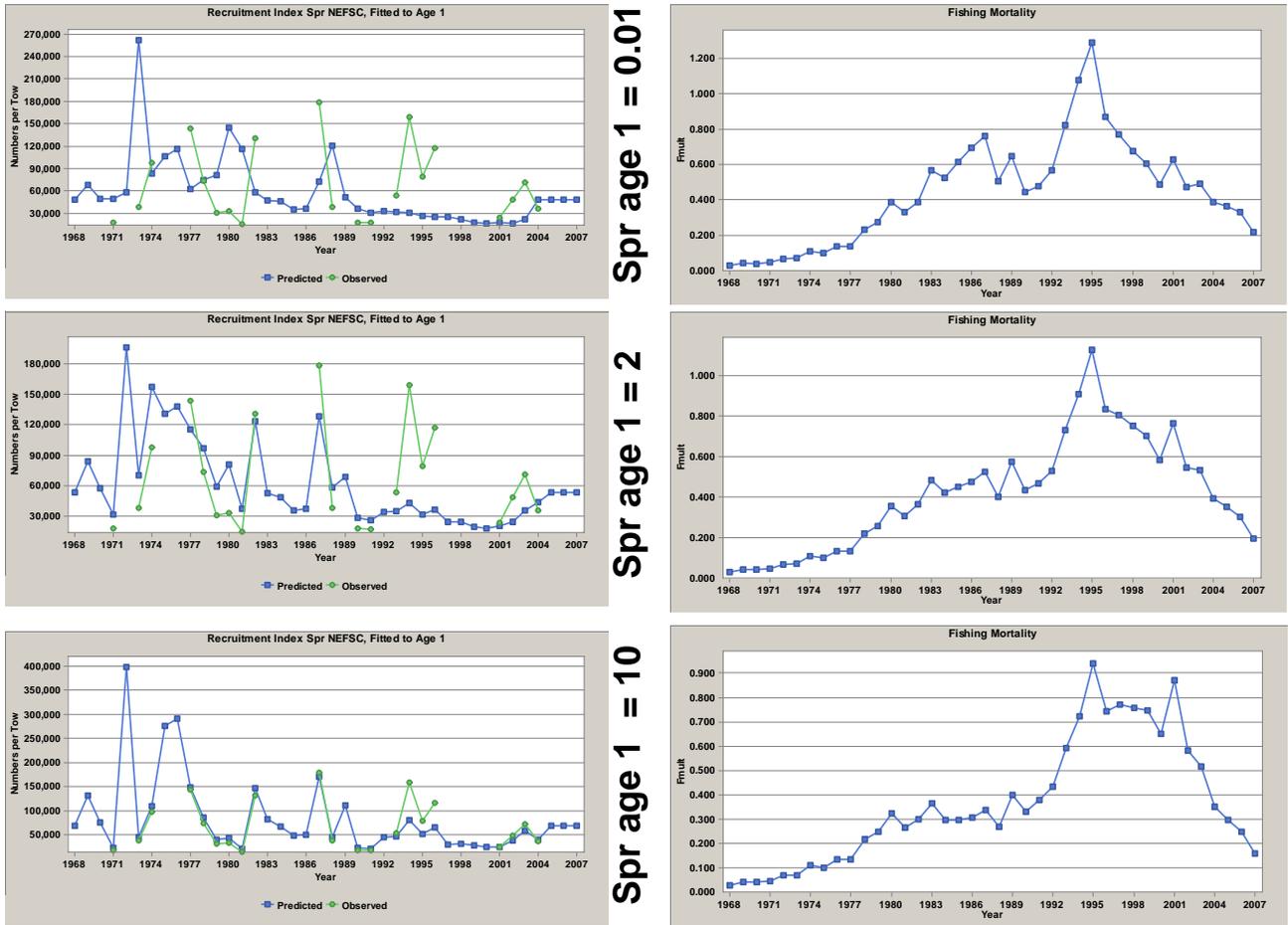


Figure 8. SCALE model sensitivity of fitting the recruitment index and the estimated fishing mortality with different weights on the recruitment index (0.01, 2, 10). The weight on recruitment variation penalty was fixed at 2.

Slope = 0.15 run (Vrec = 2)

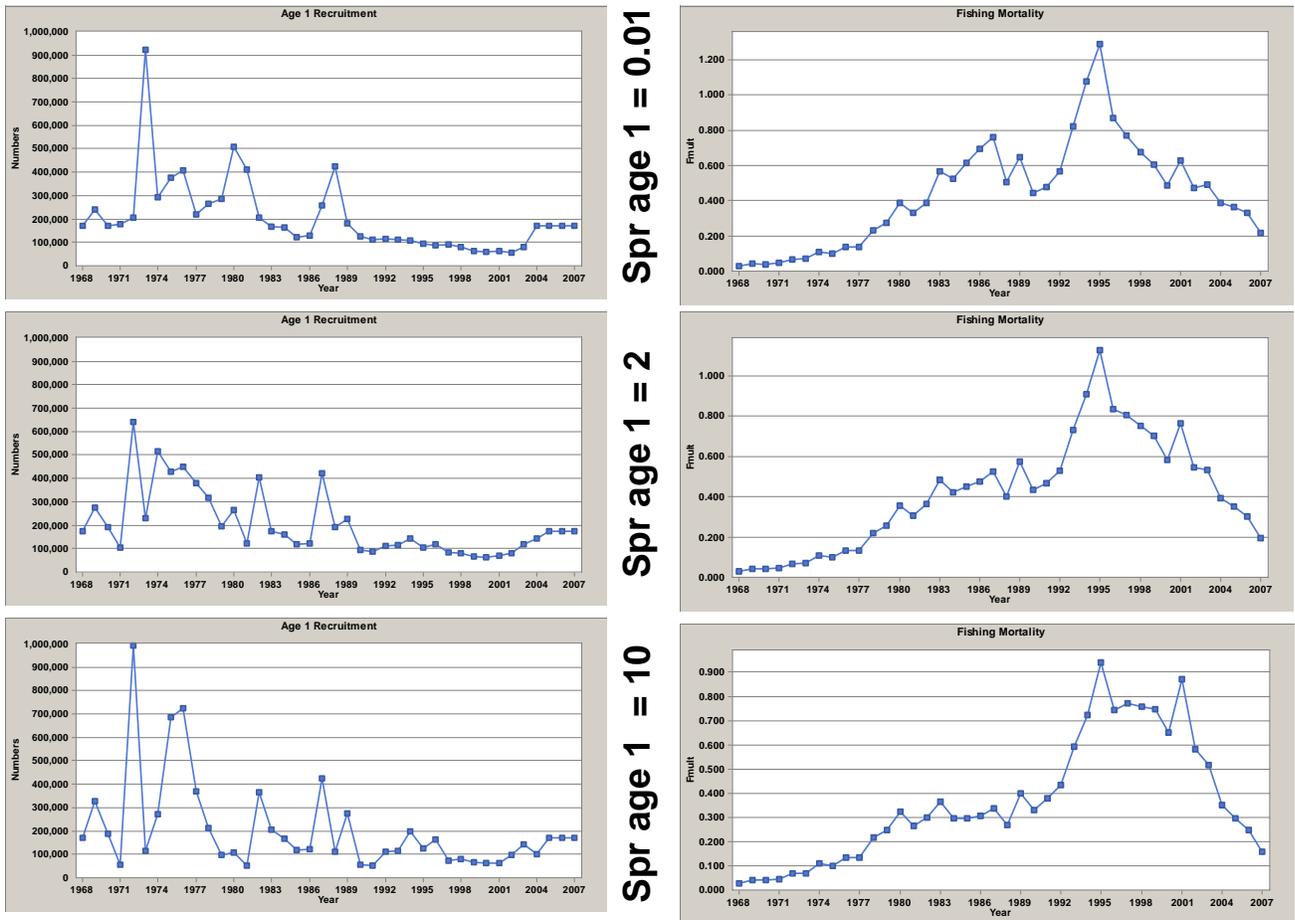


Figure 9. SCALE model sensitivity of estimated recruitment and fishing mortality with different weights on the recruitment index (0.01, 2, 10). The weight on recruitment variation penalty was fixed at 2.

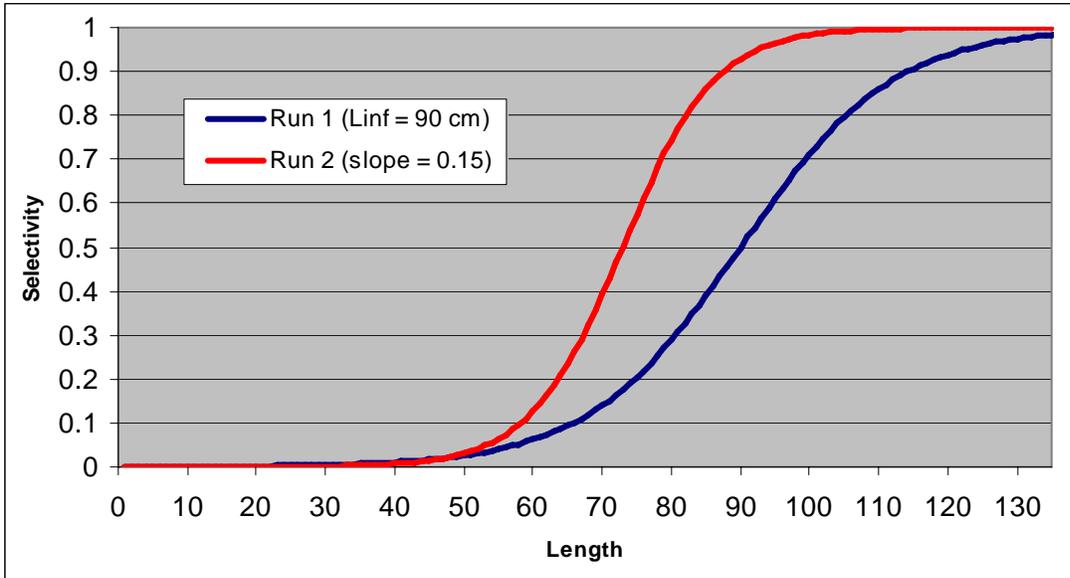


Figure 10. SCALE run 1 selectivity was allowed to hit the L-infinity bound of 90 cm which estimates a relatively flat selectivity curve. SCALE run 2 hits the slope bound of 0.15 which estimated a lower L-infinity.

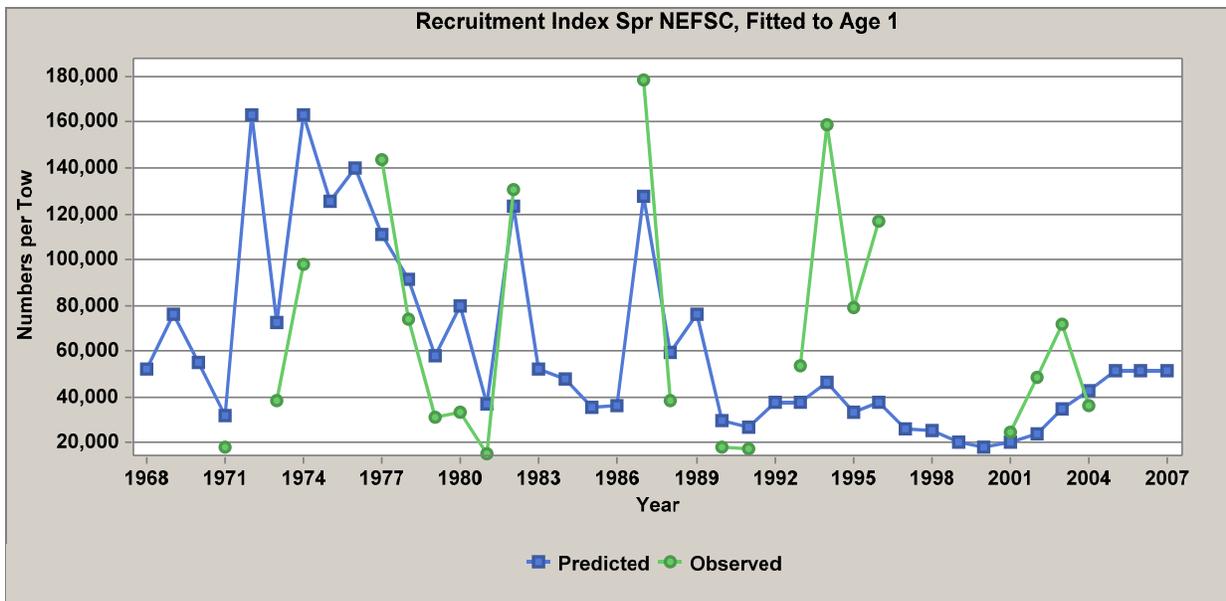


Figure 11. SCALE run 1 (L-infinity = 90 cm) fit to the NEFSC spring age-1 recruitment index.

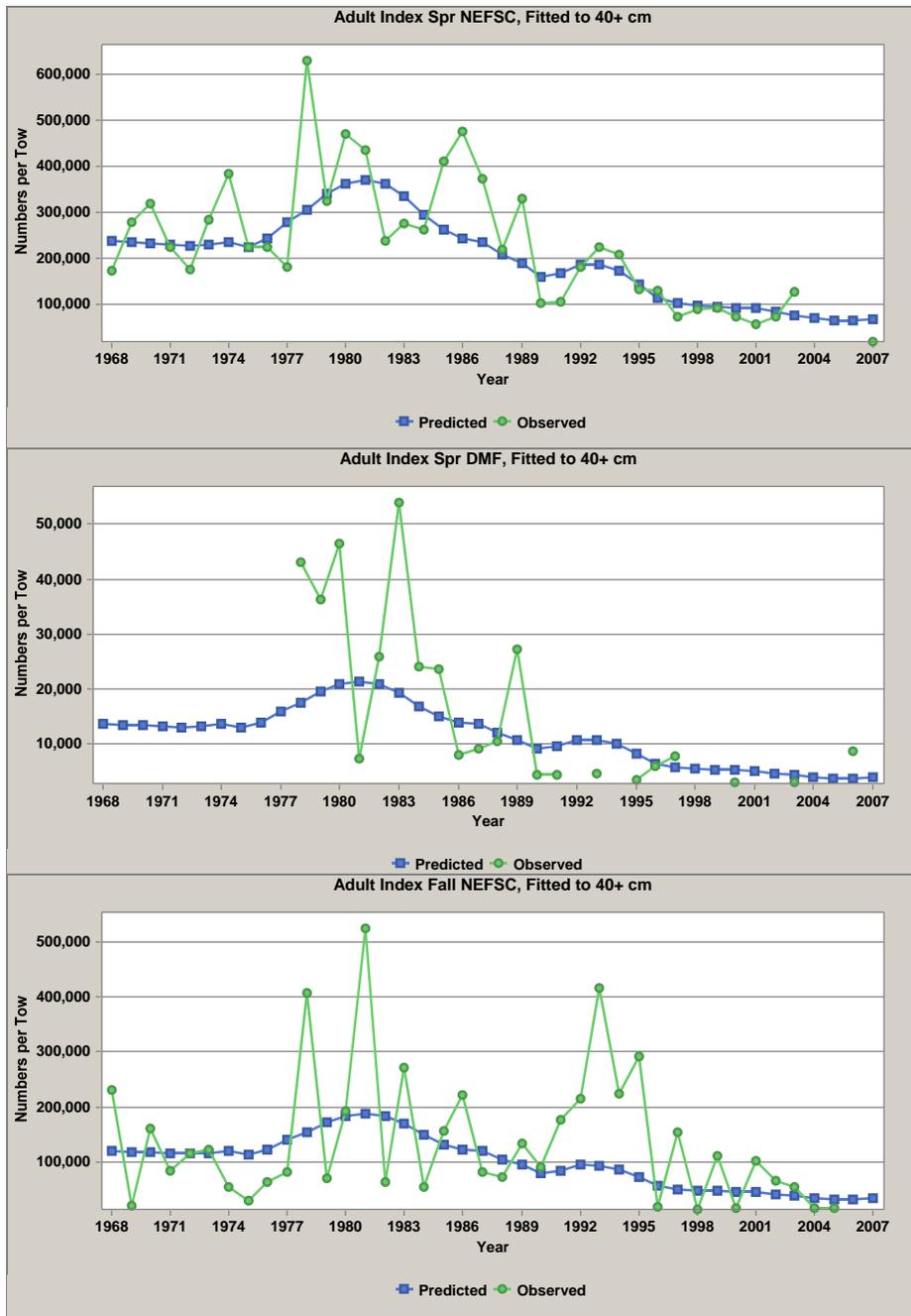


Figure 12. SCALE run 1 ($L\text{-infinity} = 90\text{ cm}$) fit to the NEFSC spring 40+ cm, MDMF 40+ cm, and NEFSC fall 40+ cm indices.

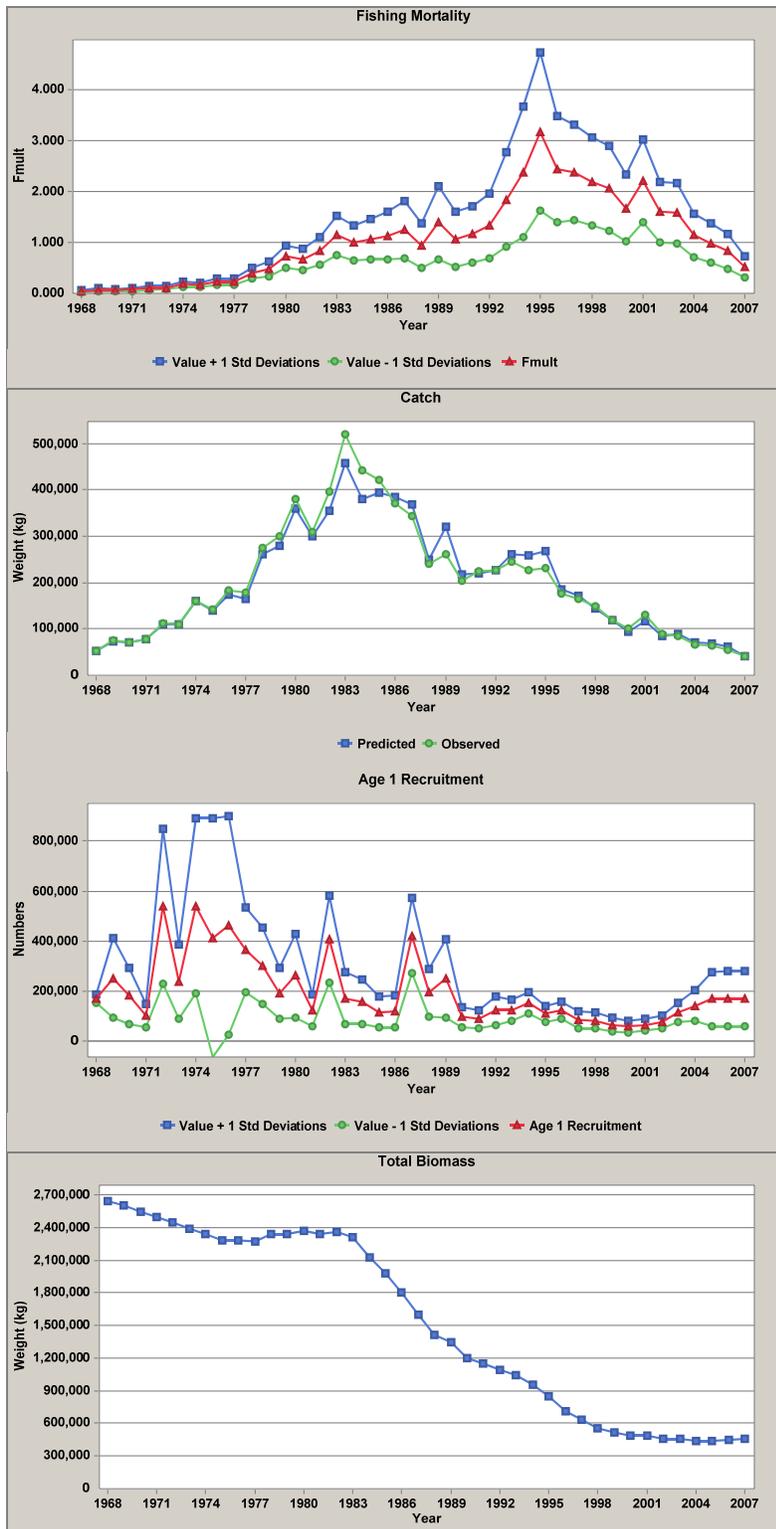


Figure 13. Run 1 ($L_{\infty} = 90$ cm) F, fit to the catch, recruitment and total biomass. Plus 1 and minus 1 standard deviations are shown on F and recruitment.

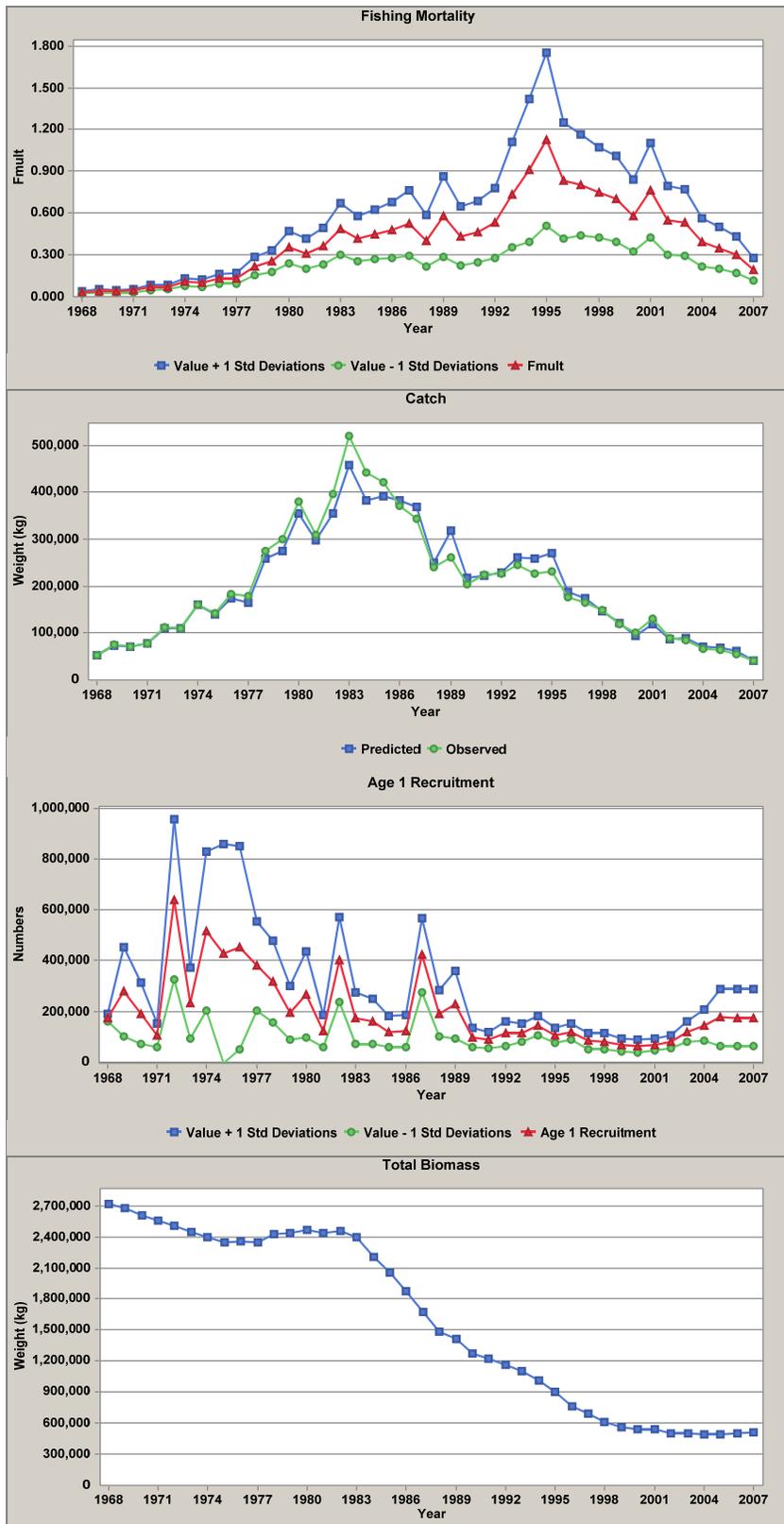


Figure 14. Run 2 (Slope = 0.15) F, fit to the catch, recruitment and total biomass. Plus 1 and minus 1 standard deviations are shown on F and recruitment.

Draft Working Paper for peer review only

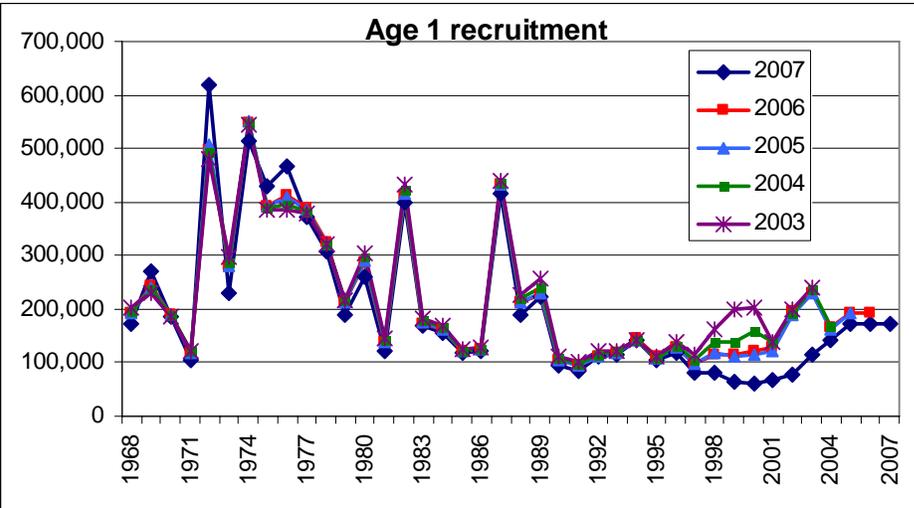
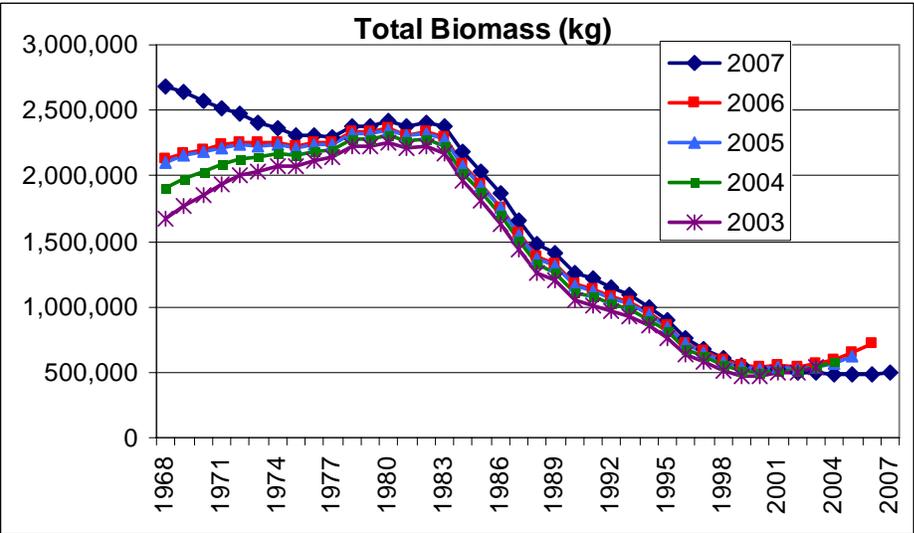
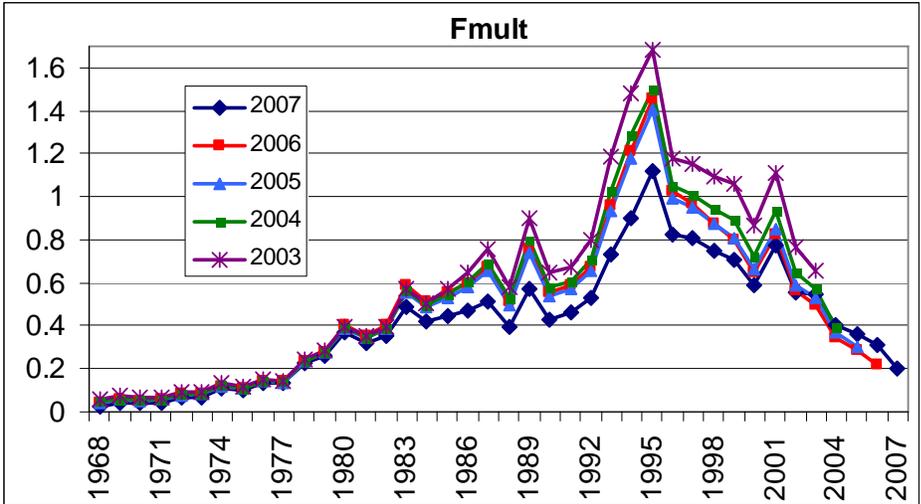


Figure 15. Run 2 (slope = 0.15) retrospective on F, total biomass and age-1 recruitment.

Slope = 0.15 run (Vrec = 2, Spr age 1 = 2)

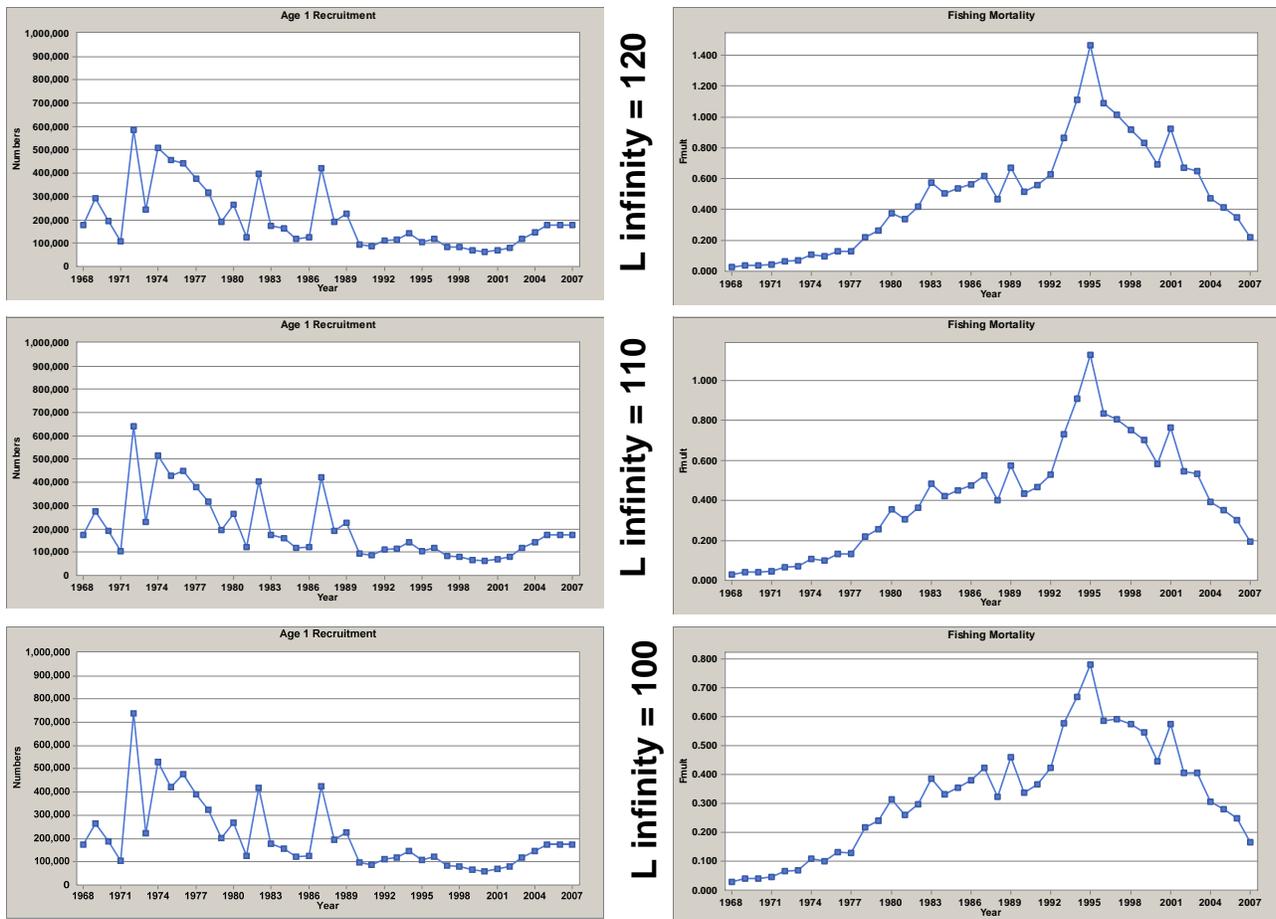


Figure 16. Run 1 (slope = 0.15) sensitivity of recruitment and fishing mortality using three different assumed L_{∞} values (100, 110, 120) on growth.

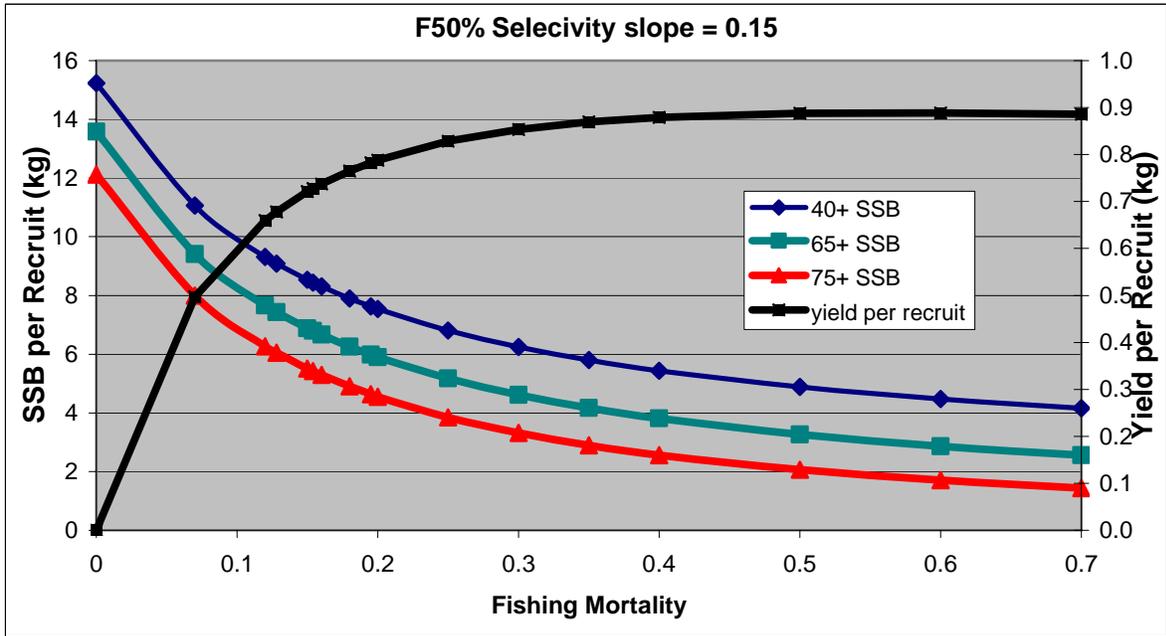


Figure 17. Updated Run 3 SCALE model F50% yield per recruit and spawn stock biomass per recruit curves.

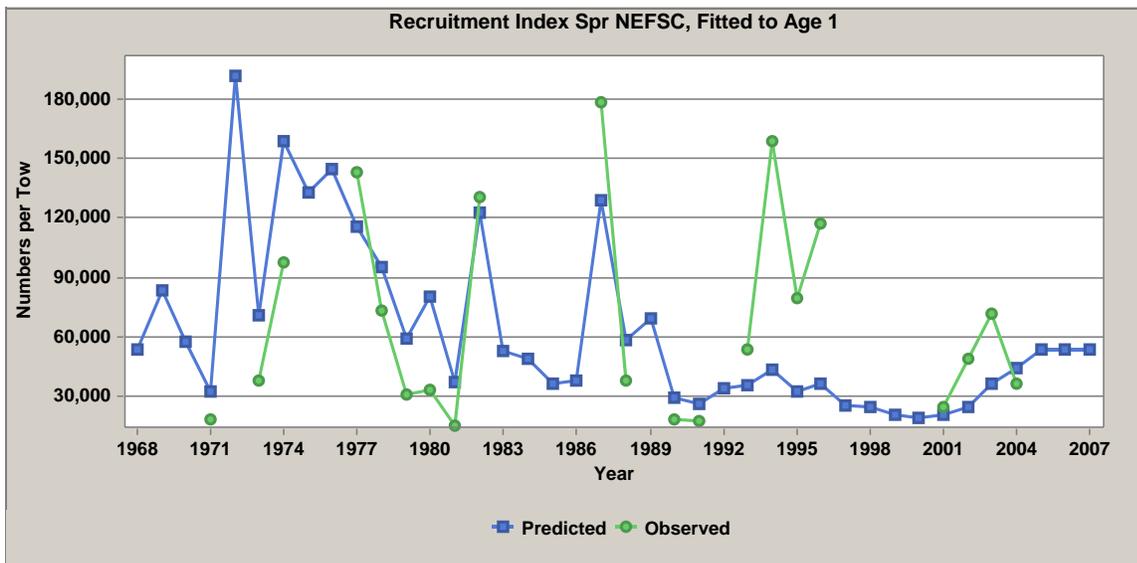


Figure 18. Updated Run 3 (slope = 0.15) SCALE model fit to the NEFSC spring age-1 recruitment index.

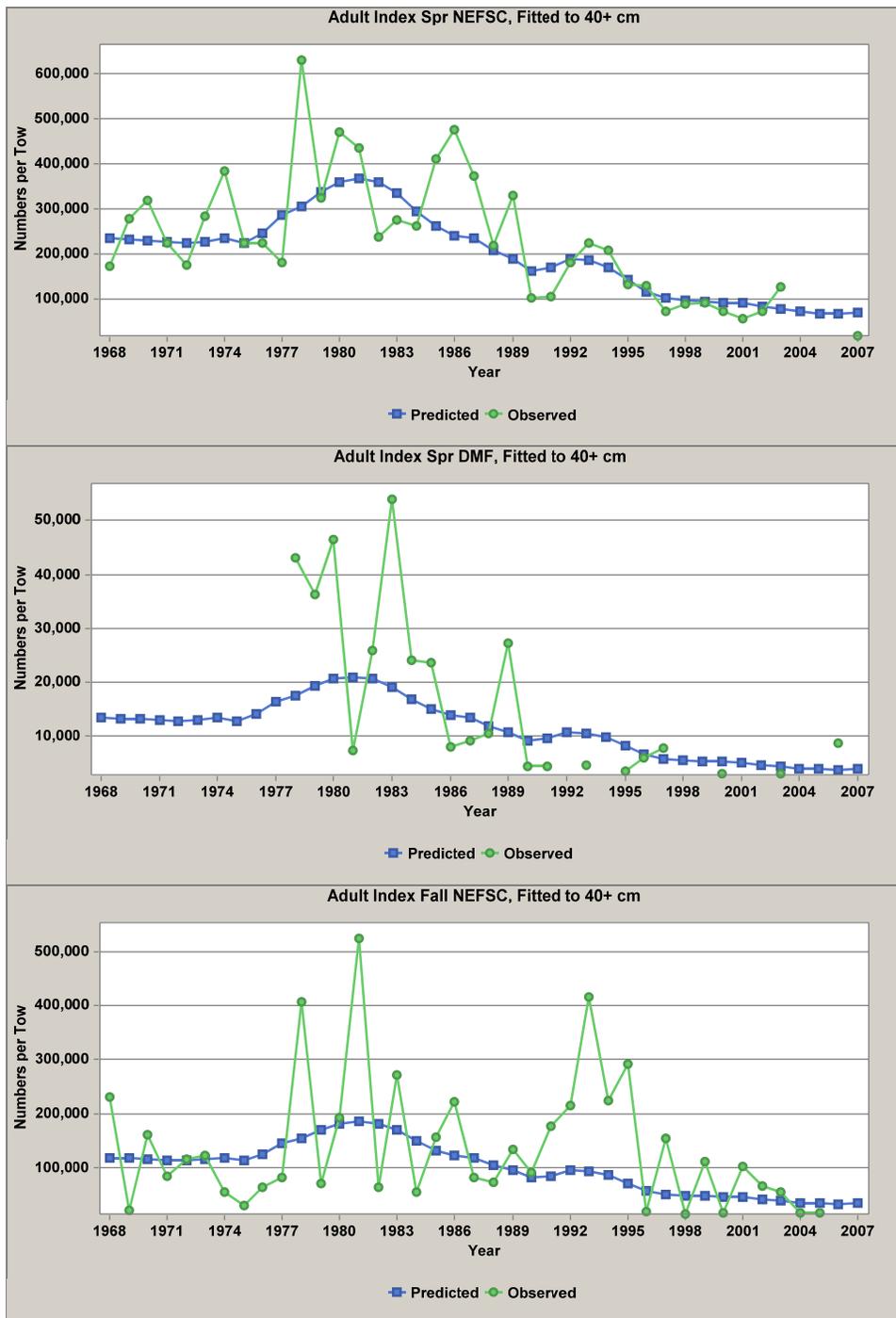


Figure 19. Updated Run 3 (slope = 0.15) SCALE model fit to the NEFSC spring 40+ cm, MDMF 40+ cm, and NEFSC fall 40+ cm indices.

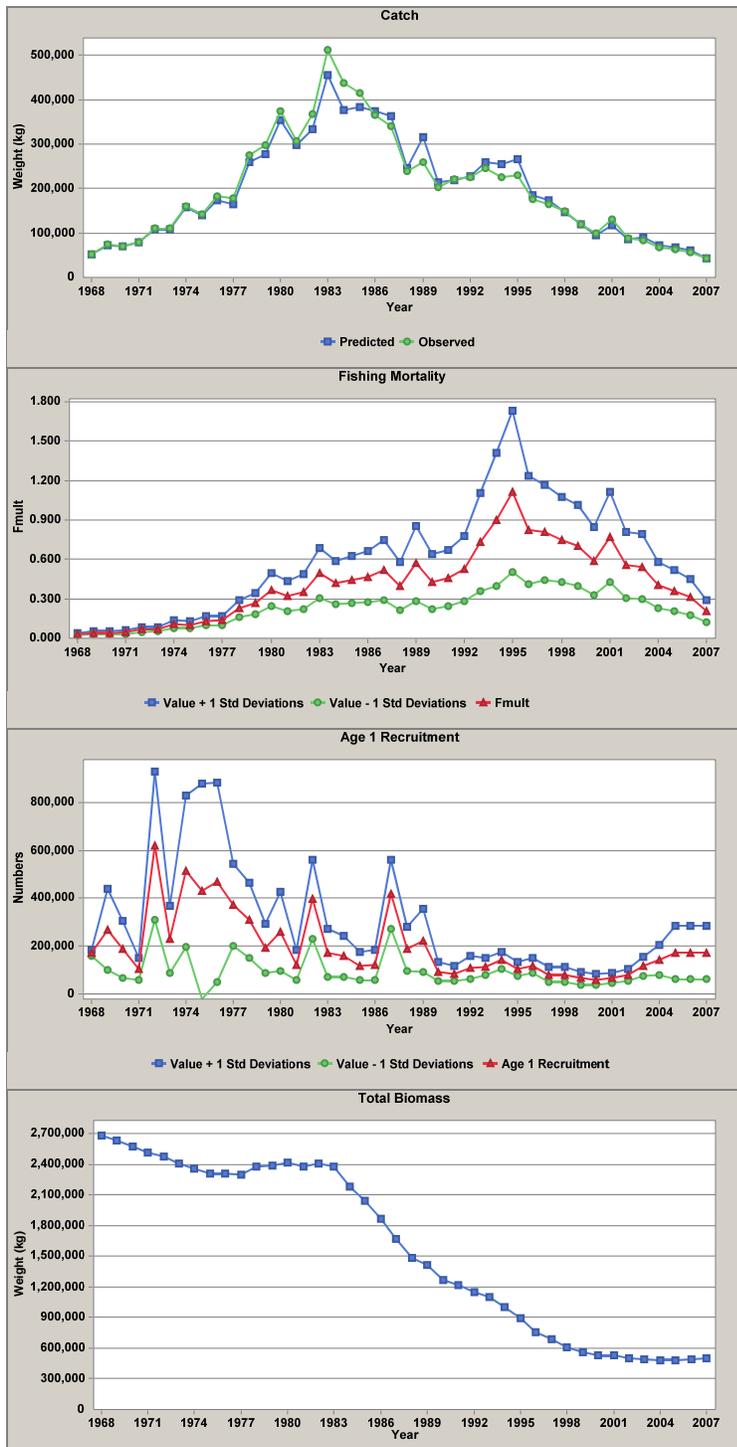


Figure 20. Run 3 (Slope = 0.15) F, fit to the catch, recruitment and total biomass. Plus 1 and minus 1 standard deviation are shown on F and recruitment.