APPENDIX 2

SAW 47 Working Paper 3 (TOR 2a) – Integrating Indices

May 9, 2007

Some approaches to the integration of survey abundance indices used in VPA calibration

by

Mark Terceiro
NOAA Fisheries NEFSC
166 Water Street
Woods Hole, MA 02563
1-508-495-2203 (Phone)
1-508-495-2393 (Fax)
Mark.Terceiro@noaa.gov

Keywords: abundance indices, virtual population analysis, general linear models, summer flounder

ABSTRACT

Northeast United States fish stock assessments typically incorporate multiple abundance indices at age from several state and federal research survey programs. Peer reviews of these assessments have recommended investigation of methods to better integrate trends in abundance provided by survey indices, prior to their use in population model calibration. Exercises were performed to explore different approaches to the integration of survey indices for use in virtual population analysis calibration. General linear modeling of integrated indices of abundance provides a useful summarization of mean survey trends. However, an empirical example for summer flounder shows that the use of integrated indices as input to virtual population analysis calibration does not guarantee substantially more accurate or precise results than using the original survey indices. The greatest potential utility for the integrated index approach is in simple index-based assessments.

INTRODUCTION

Many of the Northeast United States fish stock assessments conducted by Northeast Regional (NER) Stock Assessment Workshop (SAW) Working Groups and Atlantic States Marine Fisheries Commission (ASMFC) Technical Committees incorporate abundance indices from several state and federal agency research survey programs. Typically, these indices of abundance are provided to the assessment process as annual or seasonal indices at age. Use of these indices ranges from: a) isolated consideration in single-index based assessments (i.e., as measures of stock abundance in relation to an index based reference point; e.g., for scup, black sea bass, and the skate stocks); b) use of single indices in calibrated analyses (e.g., lobster in a Collie-Sissenwine model); and c) use of many indices from many surveys for a range of age classes in complex age-structured calibrated analyses (e.g., striped bass, bluefish, summer flounder, and winter flounder assessments).

Evaluation of the utility of indices for inclusion in complex models using many survey indcies has typically been accomplished by looking for common trends in abundance (i.e., signal) through: a) examination of time series plots; b) analysis of correlation (of lack thereof) between survey indices and between survey indices and population dynamics model results; c) outlier analysis; and d) consideration of the trend and magnitude of residuals (i.e., noise) when indices are included in population dynamics models. Multiple analyses with different sets of indices are often conducted to examine the sensitivity of results to inclusion of a given index series to determine the best analysis configuration to characterize stock status. Alternatively, all available abundance indices may be included in an analysis with the results most strongly influenced by those indices that statistically fit best within the analytical framework. Even given these approaches, with 50 or more indices of abundance at age to consider for inclusion in a complex age structured assessment, it can be difficult to discern general trends in abundance from the battery of available indices. The decision to include a given index time series at age can therefore often be subjective, based on a loose set of decision rules that may vary from one assessment to another.

Recent NER SAW peer reviews have recommended the investigation of methods to better integrate trends in stock abundance inferred from survey indices of abundance, prior to the inclusion of such indices in a population model calibration. For example, in the development of the NER index-based assessments for monkfish (NEFSC 1997), scup (NEFSC 1998) and skates (NEFSC 2000), extensive discussions occurred about which survey time series (i.e., NEFSC Spring or Fall) would best serve as the basis for biological reference points and the evaluation of stock status. A recent review of the NER summer flounder assessment (NEFSC 2002) included the following discussion:

- The SARC discussed the procedure for selecting survey indices used in the summer flounder VPA. The use of state surveys, which cover only a small component of the stock, was questioned. It was noted that YOY surveys may be variable due to the low numbers of fish caught per tow. The SARC requested that the standard error also be shown with the survey indices in the future. Whether differences in state surveys truly measure different trends in different components of the stock or whether differences are simply due to variation among survey was questioned.

and research recommendations:

- Explore the possibility of weighting survey indices used in VPA calibration by the areal coverage (i.e., in square kilometers) of the respective seasonal surveys.
- Evaluate trends in the regional components of the NEFSC surveys and contrast with the state surveys that potentially index components of the stock.@

A recent review of the NER black sea bass stock assessment (NEFSC 2004) also called for improved integration of survey indices to aid in the interpretation of stock abundance trends. That review recommended:

- More comprehensive evaluation of regional survey data is required to give more integrated indices of recruitment. For example, catch rates of recruits can be modeled as a function of location, time of year, and gear type in the surveys, to provide standardized indices, and
- Attempts should be made to extract as much information as possible from all time series considered using, for example, a GLM or GAM approach to combine the various surveys and gear types into a standardized index.

In a recent review of the NER bluefish stock assessment, the review panel (NEFSC 2004) recommended:

- There is a need for an integrated analysis of the many different research surveys for juvenile bluefish. The surveys cover different regions using different gear types and provide data on 0- and 1- group bluefish. It is recommended that serious consideration be given to... methods for standardizing and combining data from small scale intensive surveys with large scale less spatially intensive surveys, to give improved indices of recruitment.

Finally, another review of the NER summer flounder assessment (NEFSC 2005) recommended:

- Develop integrated survey indices by: combining the three NEFSC research trawl indices into a single annual abundance index, and combining state-run survey indices into a single annual abundance index.

In assessments like those for NER stocks of striped bass, bluefish, summer flounder, and winter flounder, the recommendations for development of integrated indices stem, in part, from the realization that the state agency survey data do not index trends for the entire stock, but merely components or substocks of the whole. While some state survey indices may, in fact, capture stock-wide trends, the peer-review panel research recommendations suggest that a method to statistically summarize and/or appropriately weight indices which are considered *a priori* to not adequately characterize stock-wide trends - to Aintegrate them@ - will provide more reliable and transparent results than if the indices were simply used in their original form in Virtual Population Analysis (VPA) calibration.

The integration of survey indices collected by different research sampling programs can be viewed as analogous to the standardization analysis of commercial fishing vessel rates in developing fishery-dependent indices of abundance. Viewed in that light, a General Linear Model framework (GLM; SAS Institute 1999) can be used in which deviations from the mean trend are modeled by defining various classification variables which are thought to account for the deviations. This general approach has been used in several NAFO groundfish stock assessments to integrate multiple fishery-independent survey indices of recruitment (e.g., Healey et al. 2001 and subsequent Greenland halibut assessments, and Stansbury et al. 2001 and subsequent Grand Banks cod assessments). In the current study, four exercises were constructed to explore and illustrate different approaches to the integration of indices of abundance in VPA calibrations of NER assessments.

MATERIALS AND METHODS

Exercise 1: simple, simulated survey data

As GLM modeling results can be strongly influenced by the assumed nature of the underlying error structure of the data (Terceiro 2003), the first step was to determine the appropriate error assumption to apply to research survey data. The statistical characteristics of *positive* catch data for summer flounder (*Paralichthys dentatus*) from the NEFSC Winter Trawl Survey for 1992, 1998, and 2004 were examined. Compiled on a total catch (numbers per tow) basis, the summer flounder data appear to resemble a Poisson or negative binomial distribution, although the majority of the catches (closest to the origin of the plots) reasonably approximate a lognormal distribution (Figure 1). A K-S test indicated that any of these distributions might be appropriate, with slightly better fit indicated for the Poisson (slightly smaller deviations from the expected). Terceiro (2003) indicated that inclusion of zero catch events (trips or tows) in such distributions increases the likelihood that the Negative binomial distribution will fit best. Most of the analytical models currently used in Northeast U.S. stock assessments, however, assume a normal or lognormal error structure, due mainly to variance estimation considerations.

The next step was to illustrate how Acombining indices into an integrated index@ should workgiven simulated survey data with known statistical characteristics and patterns. To this end, survey catch per tow data were simulated for 15 years and 2 seasons, with means ranging from 8 to 100 fish per tow and corresponding Coefficients of Variation (CVs) of 150% (standard errors ranging from 12 to 150), under a Poisson error distribution assumption. One hundred catch per tow values were randomly simulated for each year/season combination, for a total of 15 years * 2 seasons * 100 tows = 3,000 total tows. The annual sequence of the seasonal abundance indices was ordered to provide a time series pattern of a period of high abundance followed by a steady decline, followed by a relatively rapid increase, and then a short term decline. This exercise provided two realistic seasonal time series of survey abundance indices with: a) known statistical properties; b) slightly different annual rank orders; and c) generated a significant correlation ($r = 0.7$) between the series comparable to that between the actual NEFSC Winter and Spring survey 1992-2005 time series for summer flounder ($r = 0.66$; NEFSC 2005).

The 3,000 simulated individual tows were used an input to a GLM model with year of sampling and survey season as the main effects classification variables. The goal was to derive an Aintegrated abundance index@ from the two independent survey series - i.e., the GLM reproduction of the simple mean of two independent series with known characteristics. Models were run under lognormal, Poisson (true), and Negative binomial error assumptions. Normalized, retransformed year effect model coefficients served as the annual indices of abundance. This exercise was intended to demonstrate that if the assumption about the error distribution is correct, the GLM model should exactly extract the simple mean of two known series - i.e., a simple form of an Aintegrated@ survey index.

Exercise 2: Simulation of integrated indices at age

This simulation extended Exercise 1 to create integrated age-based indices, as might be used in an age-structured population model calibration (e.g., a VPA). Exercise 2 also explored the issue of weighting indices by the geographical coverage of individual surveys, as recommended by peer reviews of the NER stock assessments (see the Introduction). The intent was to simulate the averaging of multiple, individual survey indices at age into single, integrated indices of abundance, and compare the performance of four different index treatments in VPA calibration.

Three substock populations were simulated using NFT Popsim (NFT 2005a). The substocks were simulated with common biological and fishery characteristics (e.g., partial recruitment to the fishery and magnitude and time series patterns of fishing and natural mortality), but with different initial proportions of the additive, total stock numbers in Year 1 at ages 0 (recruits) through age 6. The magnitude of the correlation between the three simulated indices (ranging from 0.3 to 0.4, or borderline significance at the alpha = 10% level for degrees of freedom of about 20 observations; Rohlf 1981) and between the three simulated indices and the true substock sizes (ranging from 0.5 to 0.7) was made comparable to that observed in recent summer flounder assessments (NEFSC 2005, Terceiro 2006) so as to lend realism to the simulation. In actual assessments, indices with a poorer correlation than these are generally excluded from the VPA calibrations in preliminary screening work (NEFSC 2005, Terceiro 2006). Error was incorporated into the catchability coefficient (q) of each of the three simulated substock abundance indices at recruiting age 0 for the 21 years (random error with CV = 100%, 100%, and 150%) to ensure a realistic degree of deviation from the True Total Stock (TTS) sizes. The catch from each substock was simulated without error, to isolate the effects in the VPA calibration caused solely by the treatment of the age 0 indices. The percentages that each substock accounted for of the TTS numbers was set at 50%, 40%, and 10%. The simulated catch and population numbers were summed to provide the TTS catch and population numbers.

To create an integrated index for use in the four VPA calibration treatments, the three 21 year time series of simulated age 0 indices were averaged to single integrated age 0 index series within GLM models. Both simple (unweighted) and stratified (area-weighted) integrated indices were compiled. This step was intended to reconfirm the conclusion of exercise 1, but on a index-at-age basis: to establish that the GLM can exactly extract the means, simple or stratified, of multiple input time series of indices of abundance to create an integrated index of age 0 abundance. The areal coverages of the respective surveys were set at (63%, 31%, and 6%) [different from the TTS percentages in numbers (50%, 40%, and 10%)] to explore the impact of such differences (i.e., what if the assumption that survey area coverage = percentage of total stock is wrong?) on integrated index modeling and VPA calibration.

In the final step of Exercise 2, the use of the four index treatments was explored in an ADAPT VPA (NFT 2005b) calibration for the TTS catch at age and the differences summarized. The normalized versions of all indices (each value divided by its= time series mean) were used to remove scale effects prior to calibration. Only the age 0 index treatments were used as VPA calibration indices. Stock sizes for ages 1-6 were calculated using the known, input fishing mortality rates and a partial recruitment vector. Both deterministic (one-time run) and stochastic (1000 bootstrap iterations of the age 0 index calibration residuals) VPA calibrations were explored.

The four age 0 index treatments were:

1) three age 0 substock indices, simple (unweighted)
2) three age 0 substock indices, stratified (area weighted in the VPA calibration)
3) one GLM integrated age 0 index, simple (unweighted)
4) one GLM integrated age 0 index, stratified (area weighted in the GLM)

Exercise 3: Real data GLM integrated indices of abundance at age

In Exercise 3, the GLM approach was used with actual research survey data to calculate integrated indices of abundance at age for use in VPA calibration. Data from a recent NER assessment (NEFSC 2005) for summer flounder were used as an empirical test case. The time series of years for the fishery catch and research survey indices was 1982-2003/2004; the VPA calibration used survey indices at age (0-7+) from three seasonal NEFSC trawl survey series and 12 seasonal state surveys. As previously noted in the Introduction, the analytical approach is analogous to a GLM standardization analysis of commercial fishing vessel catch per unit effort data: the Ayear@ main effect classification variable serves as the index of abundance, while the Asurvey@ classification variable is analogous to a Avessel@ classification variable, each with its= own time series of catch per unit effort that has some relationship to the underlying true abundance of the stock. The mean index of abundance is modeled as a log-linear function of the classification variables. The analysis could be expanded by including additional classification variables, such as the sampling gear type or tow duration, temporal variables (e.g., spring/fall; day/night) or environmental variables (e.g., water temperature anomalies). However, such details typically are not available for most assessments, and indices are most often presented as aggregate annual or seasonal indices at age. As configured here, the analysis provides average, or integrated, annual indices of abundance at age.

Examination of the observed distribution of the normalized summer flounder age 0 survey indices suggested that the indices were best characterized by either a lognormal or Poisson/Negative binomial distribution. The standard error of the indices is slightly less than the mean (mean = 1.0, standard error = 0.8, skew = 2.8), with a single data point accounting for the high skewness. K-S tests indicated that the Poisson and negative binomial expected distributions were the same, and slightly better than the expected lognormal distribution in fitting the observed mean and variance. Visual differences (observed minus expected) were similar for the three expected distributions. Since the indices were to be lognormal-transformed in the ADAPT VPA calibration (NFT 2005b), and the age 0 indices represent the largest group of indices with the greatest absolute value range (and hence provide the best age for which to reliably examine statistical properties), it was concluded that GLM modeling under a lognormal error distribution would be reasonable for all ages in this exercise.

GLM models were constructed for ages 0, 1, 2, 3, 4, and 5-7+. Main effects were limited to the year of sampling (1982, 1983...2004) and the identity of the survey (NEFSC age 1, NEFSC age 2...NEFSC age 5-7+). The resulting year effect coefficients, corrected for lognormal-transformation bias and re-transformed to the original scale, were used as a single index of abundance at age 0 input to the VPA calibration in place of the twelve original survey series. The input GLM age 0 vector was called GLM_YOY. The corresponding VPA run using this vector was called VPA_GLM0. In VPA_GLM0, for example, all of the original indices for all the other ages (1, 2, 3, 4, and 5-7+) were retained so that the effect of using the GLM_YOY vector could be isolated. The pattern was repeated as GLM vectors (GLM_1, GLM_2...GLM_5:7) for the other ages tested. A run using only the GLM vectors at age

(VPA_GLM) was also constructed. Results from these seven GLM integrated index run configurations were compared to the trends in stock size at age provided by the VPA calibration run (F04_ALL) using the original, full suite of indices at age.

Exercise 4: Real data GLM integrated indices at age, NEFSC vs. State

The 2005 SARC 41 Panel review of the NER summer flounder assessment (NEFSC 2005) recommended the development of integrated survey indices by combining the various seasonal NEFSC research trawl survey series indices at age into single annual abundance indices at age (e.g., NEFSC age 0 index, age 1 index, etc.), and likewise combining the state survey indices into a single annual abundance indices at age. In Exercise 4, the GLM approach was used with the same data as in Exercise 3 to construct integrated indices at age from the three seasonal NEFSC surveys (winter, spring and fall) and from the state surveys (MA, RI, CT, NJ, MD, VA, NC), for a total of twelve GLM integrated indices at age (NEFSC ages 0-5:7+; State ages 0-5:7+). Considering the series in this manner resulted in more inconsistent data in terms of the length of the series, and more frequent occurrence of >zero@ observations. Therefore, the resulting GLM integrated ANEFSC@ and AState@ indices exhibit a greater number of missing observations for some year and age combinations than did the six GLM integrated indices at age for all surveys combined constructed in Exercise 3. Given the extent of Exercise 3, comparisons in Exercise 4 were limited to a VPA calibration using the 12 GLM integrated indices (VPA_NEC_ST) and the VPA calibration (F04_ALL) using the original suite of indices at age.

RESULTS

Exercise 1: simple, simulated survey data

Table 1 shows the time series of annual means of the two simulated seasonal survey indices, the combined annual means of the two simulated series, and retransformed GLM year effect coefficients (annual indices of abundance) under lognormal, Poisson, and negative binomial error assumptions. As the two seasonal series were simulated with Poisson error, the expected result was that the retransformed Poisson coefficients would exactly match the combined mean of the two input series, while the lognormal and negative binomial results would differ slightly. For ease of comparison, all results were rescaled to the means of the respective series in the bottom of Table 1. The results demonstrate that if the error distribution is correctly specified, the GLM model can exactly reproduce the combined mean of averaged survey series.

Exercise 2: Simulation of integrated indices at age

The Year 1 numbers at age of the three substock populations (SS1, SS2, SS3) simulated using NFT Popsim (NFT 2005a) are presented in Table 2. The catch and population numbers were summed to provide the true total stock (TTS) catch and population numbers. The panels in Figure 2 show the relationship between the simulated age 0 population sizes in the three substocks and the respective simulated age 0 survey indices over the 21 year time series. Figure 3 presents the trends in age 0 stock size, age 1-6+ stock size, and age 0-6+ catch in the simulated VPA used to explore the sensitivity of the VPA calibration to different treatments of the age 0 indices.

The simulated integrated indices of age 0 abundance using simple arithmetic averaging and GLM modeling are presented in the upper section of Table 3. As noted earlier, Astratified@ equates to area weighted; rescaled indices (divided by the time series means) are presented in the lower section of Table 3. The results re-confirm those of Exercise 1, that given the correctly specified error distribution, the GLM model can exactly reproduce the combined mean of averaged survey series.

VPA calibration results for the four treatments of the age 0 indices (VPA1 = simple mean, VPA2 = stratified mean, VPA3 = simple GLM, and VPA4 = stratified GLM) differ from the True Total Stock (TTS) and also from each other mainly in the Auncoverged@ part of the VPA in Years 14-21 (Table 4). Because the area weights intentionally did not match the true substock percentages, the stratified mean (VPA2) and stratified GLM (VPA4) treatments generated calibration results that deviated more (both on an aggregated deviation and mean deviation basis), and correlated less well (Pearson r) than the simple mean (VPA1) and simple GLM (VPA3) treatments. Since the area weight was highest for SV1 (63%), the weighted index treatments (VPA2 and VPA4) correlated best with the SV1 index, and poorest with SV2 and SV3, than the simple mean (VPA1) and simple GLM (VPA3) treatments. In this exercise, the smoothing effect of the simple GLM model of the indices produced VPA calibrated stock sizes that deviated from the TTS sizes slightly less, on both aggregate and mean bases, than the simple mean treatment (Table 4, Figure 4).

VPA bootstrap results were qualitatively similar to the one-time runs, and with the focus on the Year 21 age 0 stock size estimates, show how the incorrect assumption of survey area coverage as a proxy for stock size percentages can provide inaccurate results. As in the one-time runs, the simple (unweighted) VPA1 and VPA3 bootstrap estimates more closely match the TTS size for age 0 in Year 21 than the stratified (area weighted) VPA2 and VPA4 estimates (Table 5). The smoothing effect of the GLM on the integrated age 0 index in the VPA3 calibration produces a larger deviation from the TTS size than the simple three index VPA1 calibration. Finally, while the VPA2 estimate (54,107) is most precise (CV = 0.11; due to the good correlation of estimated stock sizes and the SV1 index), it deviates most from the TTS size (92,000).

Exercise 3: GLM using real indices of abundance at age

The results for the age 0 indices in exercise 3 are first provided in the upper left panel of Figure 5, where the pattern of age 0 stock sizes indicated by the GLM model estimated year effect vector (the integrated index at age 0, GLM_YOY) is compared with the estimates from the VPA (VPA0_GLM0) when this single, integrated age 0 index is used in place of the 12 original indices to calibrate age 0 stock size. The overall patterns are similar, with highest recruitment at the start of the series and a poor year class in 1988. The major difference is in the rank order of the 1982/1983 and 1985/1986 year classes. The upper left panel of Figure 6 compares estimates of age 0 stock size from VPA_GLM0 (using the integrated index for age 0, and the original indices for ages 1 and older) with the VPA using the original survey series for all ages (F04_ALL). In the two VPAs, the estimates of age 0 abundance are nearly identical. Exercise 3 results for ages 1, 2, 3, 4 and 5-7+ are provided in the successive panels of Figures 5-6. A VPA calibration was also conducted using only the GLM integrated indices (i.e., 6 index series at ages 0-5:7+), and these results are presented in Figure 7.

In general, the GLM integrated indices at age diverged somewhat from the VPA_GLM estimates, due to the smoothing effect of the GLM and possibly due to process error caused by mis-specification of the true error structure. However, the VPA_GLM and F04_ALL estimates are nearly the same for all ages, diverging only in the most recent 2-3 years in the unconverged part of the VPA. This last finding reflects the substantial influence on stock size estimates of the input catch at age data and the convergence properties of the VPA model. The F04_ALL VPA calibration is characterized by a substantial retrospective pattern, with F underestimated and stock size overestimated over the unconverged part of the analysis. The retrospective patterns for F, SSB, and recruitment at age 0 (R) are nearly identical for the VPA_GLM calibration, indicating no improvement in this characteristic of the analysis by using integrated indices of abundance (Figure 8).

Exercise 3 results suggest that using GLM integrated indices may increase the difficulty of interpreting the uncertainty of the VPA estimates. VPA calibrations based on a limited number of externally derived GLM integrated indices are likely to have less total absolute variance than calibrations with multiple sets of indices; an example is the difference in the Residual Sums of Squares (RSS) and Mean Squared Residual (MSR) between the F04_ALL and VPA_GLM runs. The VPA_GLM MSR is about one-third of the F04_ALL MSR, indicating an overall Abetter fitting@ model (Table 6). However, the number of potential calibration residuals in the VPA_GLM run is also much lower (130 versus 937 in the F04_ALL run), and so estimates of individual stock sizes, and subsequently derived quantities such as Average F and Biomass, are less precise (a Adegrees of freedom@ phenomenon). This is evident in both the deterministic Nonlinear Least Squares (NLLS) and Bootstrap (1000 iterations; BOOT) results for the F04_ALL and VPA_GLM runs (Table 6).

Exercise 4: GLM Integrated Indices, NEFSC vs. State

Figure 9 compares the GLM integrated indices at age derived from the NEFSC survey indices at age are compared with those derived from state survey indices at age. Consistency in trend and rank order between the NEC and ST indices at age is poorest for ages 0 and 1, and best for ages 3 and 4, but overall is very similar across all ages. The use of the twelve GLM integrated indices at age in a VPA calibration, the VPA_NEC_ST run, produced estimates of stock size at age that were generally slightly lower than the F04_ALL run in the unconverged (most recent) years of the analysis. As a result, the estimated total stock size is slightly lower for the VPA_NEC_ST run compared to the F04_ALL estimate, and correspondingly the average fishing mortality rate (F) is slightly higher (Figure 10). Finally, as with the Exercise 3 comparison, the total variance (RSS) for the integrated index VPA_NEC_ST run is smaller than for the F04_ALL run. However, the VPA_NEC_ST run Mean Squared Residual (MSR) is higher than the F04_ALL run, indicating a slightly poorer fit. As well, the VPA_NEC_ST run provided larger CVs on the estimated parameters, in both the deterministic (NLLS) and bootstrap (BOOT) runs (Table 7). The overall fit of the NEC integrated indices was similar to the ST integrated indices, with the NEC integrated indices accounting for 51.5% of the total variance in the fit and the ST integrated indices 48.5%. NEC integrated indices at age fit better (i.e., smaller partial variance) for ages 2, 4, and 5-7+; the ST indices fit better for ages 0, 1, and 3 (Figure 11).

CONCLUSIONS

The results of Exercises 1 and 2 suggest that use of a lognormal error distribution in constructing integrated indices in a GLM framework can introduce some degree of process (model) error to these indices because of mis-specification of the true error distribution of survey catch data. Future developments should consider alternative error distribution assumptions. Such alternative assumptions, or use of non-parametric approaches such as General Additive Modeling (GAM), would also permit the use of Azero@ observations in the calibration which are generally treated as missing observations in lognormal models.

The inclusion of auxiliary information (e.g., environmental data) in the GLM modeling could theoretically improve the accuracy and utility of integrated indices. If fine-scale auxiliary data are available and have predictive utility, integration of indices at the tow or stratum level could be easily accomplished. Further research should consider alternative modeling frameworks such as GAM or ordination approaches such as Principal Components Analysis (PCA) which could incorporate non-parametric assumptions and smoothing. Results of exercises using real summer flounder data indicate that without the inclusion in the GLM model of significant main effects (beyond year of sampling and survey identity) that account for a large proportion of the variance of survey series at age from the simple overall means, use of a GLM to develop integrated indices at age provides no clear advantage over using the original indices as input to the VPA calibration. While the GLM integrated indices provide a useful summarization of mean survey trends, the use of integrated indices as VPA calibration input does not guarantee substantially more accurate or precise results than calibration using the original survey indices.

A number of stock assessments in the Northeast United States rely on a single, seasonal time series of survey indices, selected from among several candidate series, as the sole means of evaluating the status of the stock with respect to an index-based reference point. For those situations, the construction of an integrated index of abundance from several different time series could provide a more robust approach to the evaluation of the status of a stock. Therefore, the greatest potential utility for the integrated index approach may be for simple index-based assessments.

LITERATURE CITED

Healey BP, Cadigan NG, Brodie WB. 2001. Analysis of pre-recruit data from surveys for Greenland halibut in NAFO subarea 2 and divisions 3KLMNO. NAFO SCR Doc. 01/44. Serial No. N4422.

Northeast Fisheries Science Center (NEFSC). 1997. Report of the 23rd Northeast Regional Stock Assessment Workshop (23rd SAW): Stock Assessment Review Committee (SARC) Consensus Summary of Assessments . NEFSC Ref Doc. 97-05. 191 p.

Northeast Fisheries Science Center (NEFSC). 1998. Report of the 27th Northeast Regional Stock Assessment Workshop (27th SAW): Stock Assessment Review Committee (SARC) Consensus Summary of Assessments . NEFSC Ref Doc. 98-15. 350 p.

Northeast Fisheries Science Center (NEFSC). 2000. Report of the 30th Northeast Regional Stock Assessment Workshop (30th SAW): Stock Assessment Review Committee (SARC) Consensus Summary of Assessments . NEFSC Ref Doc. 00-03. 477 p.

Northeast Fisheries Science Center (NEFSC). 2002. Report of the 35th Northeast Regional Stock Assessment Workshop (35th SAW): Stock Assessment Review Committee (SARC) Consensus Summary of Assessments . NEFSC Ref Doc. 02-14. 259 p.

Northeast Fisheries Science Center (NEFSC). 2004. 39th Northeast Regional Stock Assessment Workshop (39th SAW) Assessment Summary Report & Assessment Report. NEFSC. Ref Doc. 04-10. 211 p.

Northeast Fisheries Science Center (NEFSC). 2005. 41st Northeast Regional Stock Assessment Workshop (41st SAW) Assessment Summary Report & Assessment Report. NEFSC Ref Doc. 05-14. 259 p.

NOAA Fisheries Toolbox (NFT). 2005a. Population simulator, Version 3.1. http://nft.nefsc.noaa.gov.

NOAA Fisheries Toolbox (NFT). 2005b. Virtual population analysis, Version 2.5. http://nft.nefsc.noaa.gov.

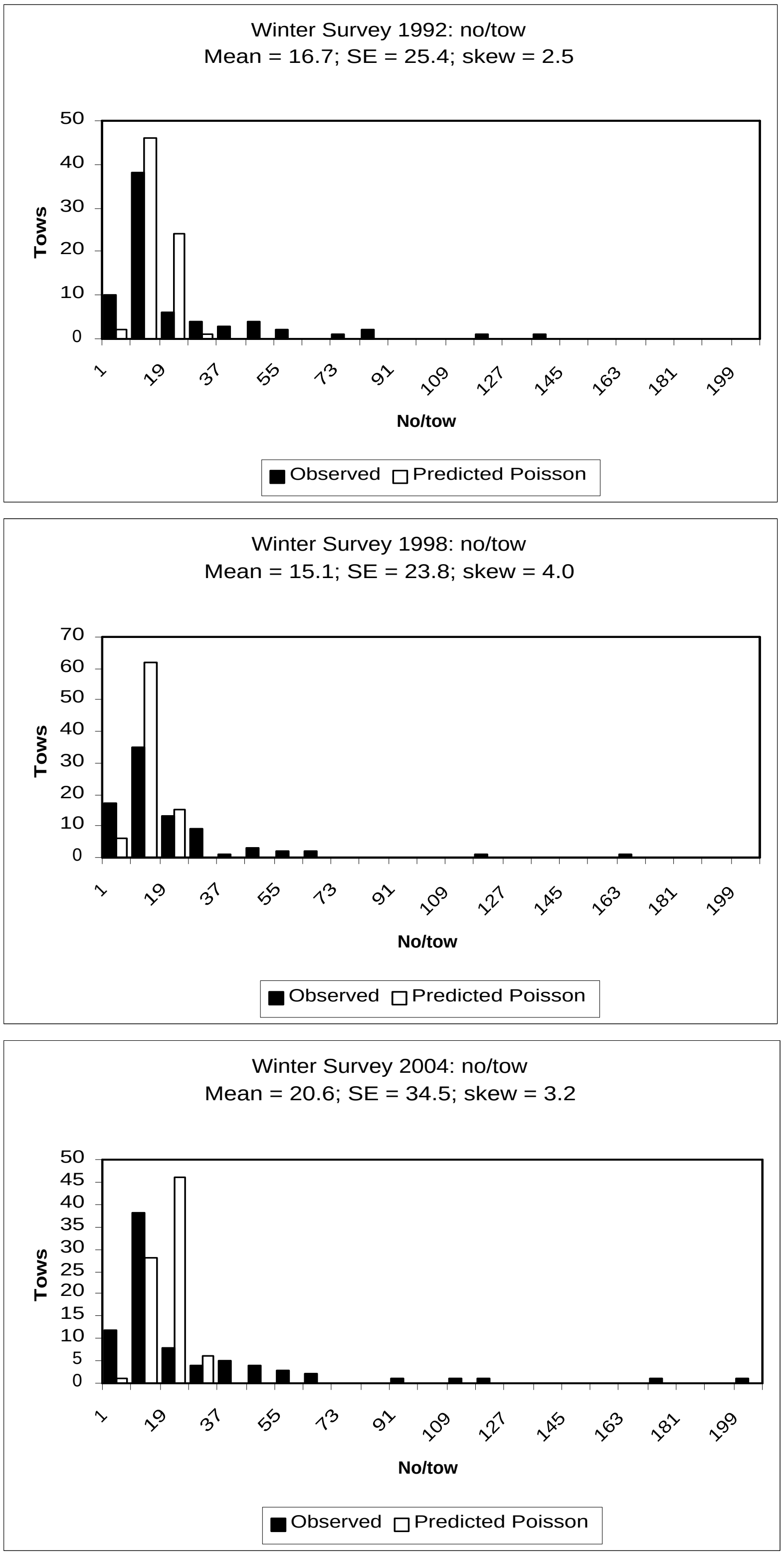
Rohlf FJ. 1981.Statistical tables, second edition. W.H.Freeman and Company. New York. 219 p.

SAS Institute. 1999. SAS OnlineDoc. version 8. SAS Institute Inc. Cary NC.

Stansbury DE, Shelton PA, Murphy EF, Healey BP, Brattey J. 2001. An assessment of the cod stock in NAFO divisions 3NO. NAFO SCR Doc. 01/72. Serial No. N4450.

Terceiro M. 2003. The statistical properties of recreational catch rate data for some fish stocks off the northeast U.S. coast. Fish Bull. 101:653-672.

Terceiro M. 2006. Stock assessment of summer flounder for 2006. NEFSC Ref Doc. 06-17. 119 p.



Winter Survey 1992: no/tow
Mean = 16.7; SE = 25.4; skew = 2.5
Tows
50
40
30
20
10
0
1
19
37
55
73
91
109
127
145
163
181
199
No/tow
Observed
Predicted Poisson
Winter Survey 1998: no/tow
Mean = 15.1; SE = 23.8; skew = 4.0
Tows
70
60
50
40
30
20
10
0
1
19
37
55
73
91
109
127
145
163
181
199
No/tow
Observed
Predicted Poisson
Winter Survey 2004: no/tow
Mean = 20.6; SE = 34.5; skew = 3.2
Tows
50
45
40
35
30
25
20
15
10
5
0
1
19
37
55
73
91
109
127
145
163
181
199
No/tow
Observed
Predicted Poisson

Figure 1

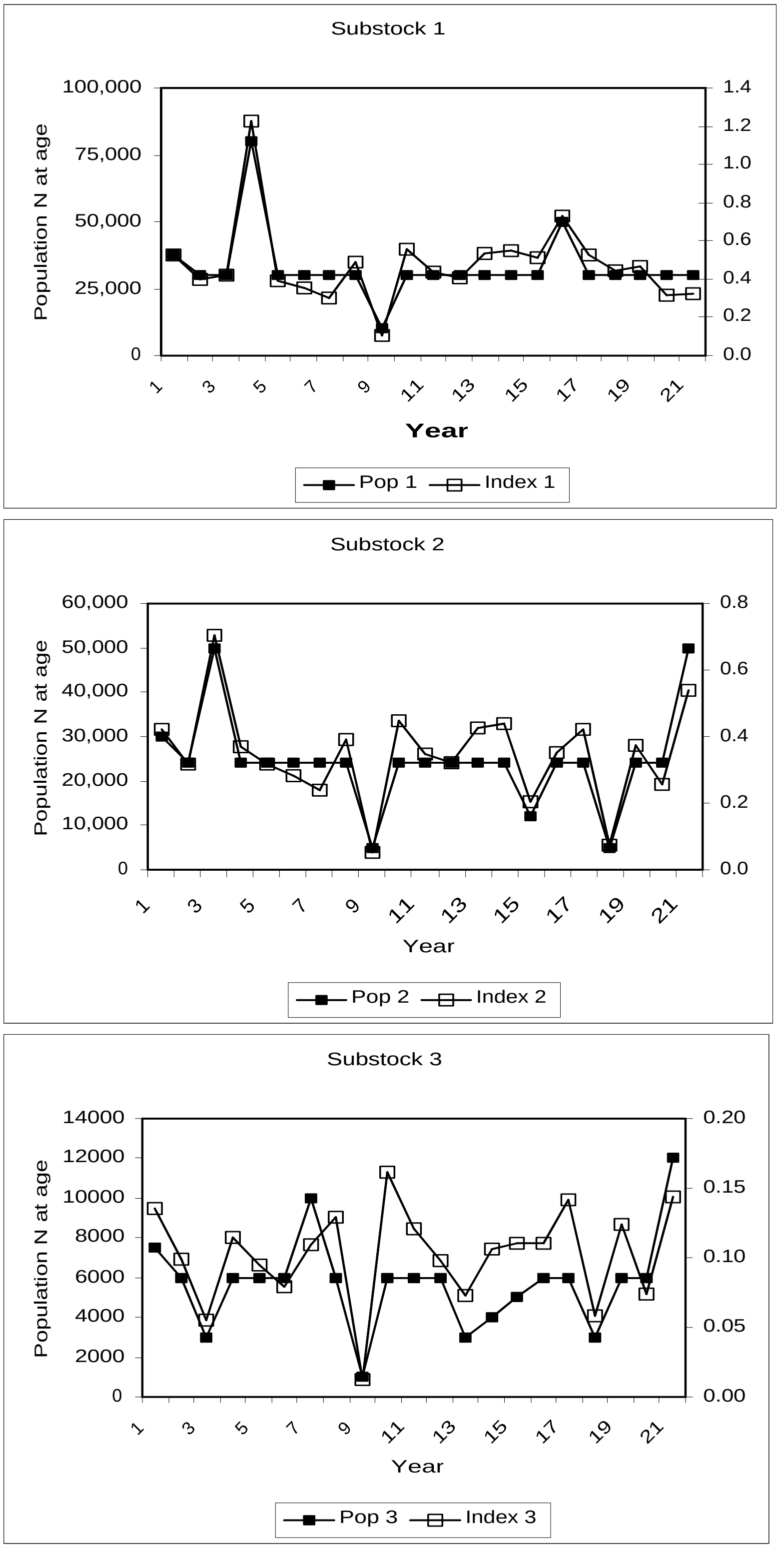
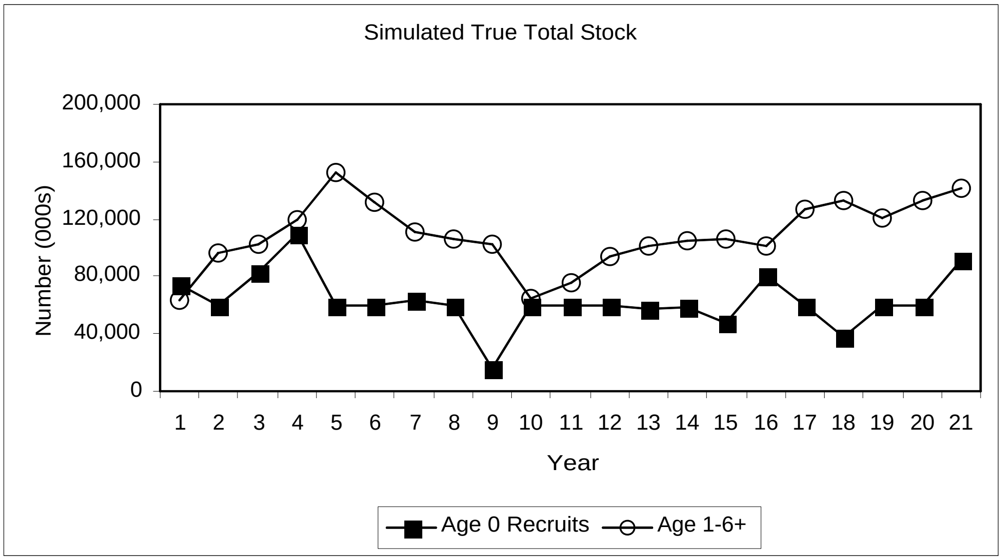

Substock 1
Population N at age
100,000
75,000
50,000
25,000
0
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1
3
5
7
9
11
13
15
17
19
21
Year
Pop 1
Index 1
Substock 2
Population N at age
60,000
50,000
40,000
30,000
20,000
10,000
0
0.8
0.6
0.4
0.2
0.0
Year
Pop 2
Index 2
Substock 3
Population N at age
14000
12000
10000
8000
6000
4000
2000
0
0.20
0.15
0.10
0.05
0.00
Year
Pop 3
Index 3

Figure 2.


Simulated True Total Stock
200,000
160,000
120,000
80,000
40,000
0
Number (000s)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Year
Age 0 Recruits
Age 1-6+

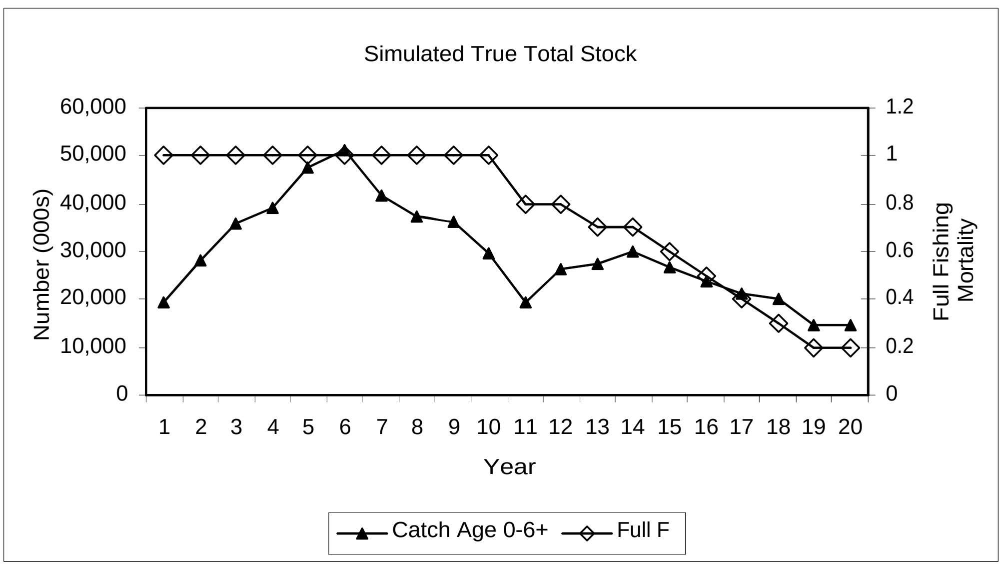

Simulated True Total Stock
60,000
50,000
40,000
30,000
20,000
10,000
0
Number (000s)
1.2
1
0.8
0.6
0.4
0.2
0
Full Fishing Mortality
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Year
Catch Age 0-6+
Full F

Figure 3.

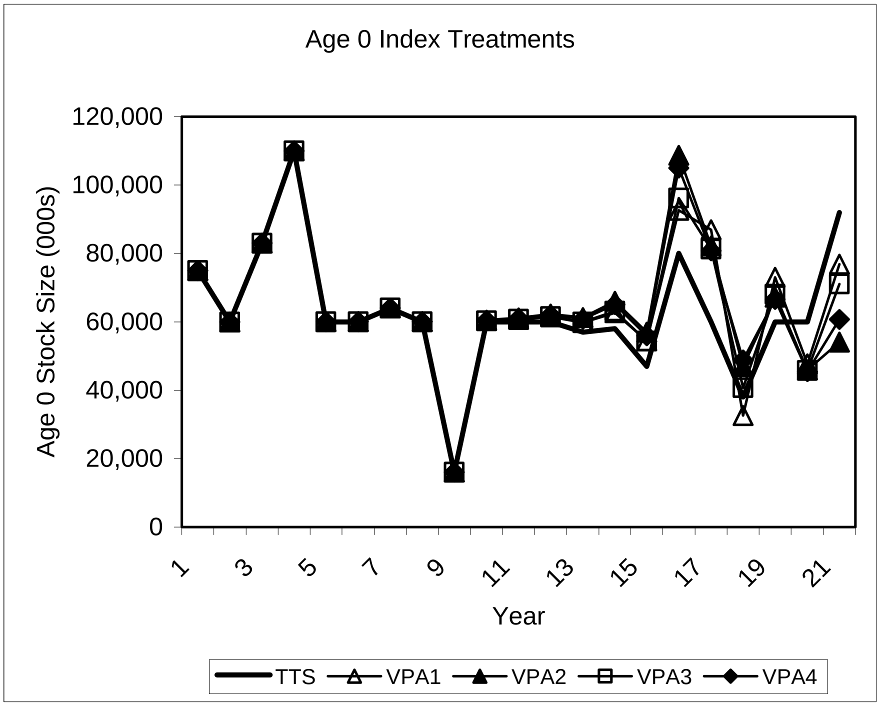

Age 0 Index Treatments
120,000
100,000
80,000
60,000
40,000
20,000
0
Age 0 Stock Size (000s)
1
3
5
7
9
11
13
15
17
19
21
Year
TTS
VPA1
VPA2
VPA3
VPA4

Figure 4.

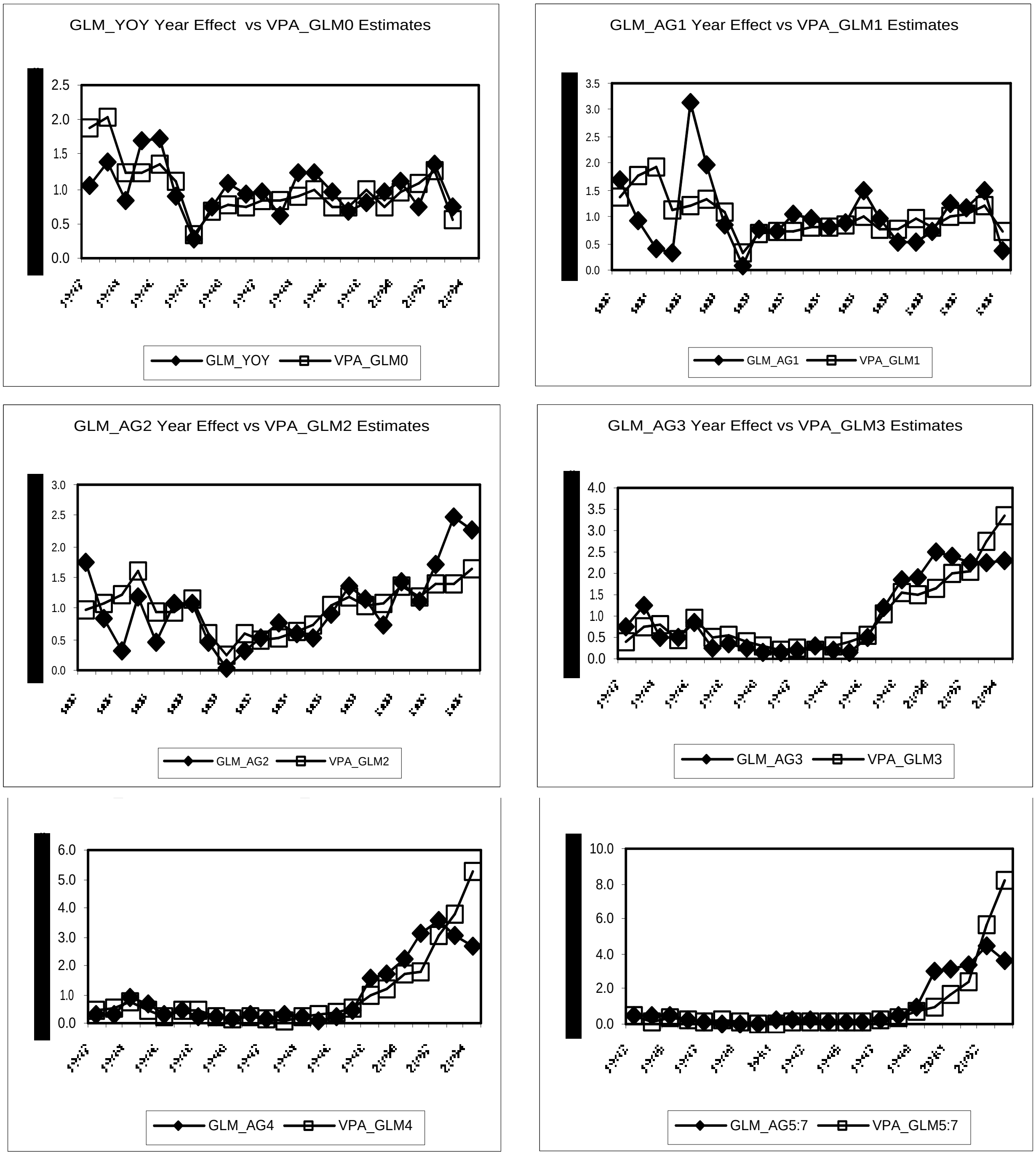

GLM_YOY Year Effect vs VPA_GLM0 Estimates
2.5
2.0
1.5
1.0
0.5
0.0
GLM_YOY
VPA_GLM0
GLM_AG1 Year Effect vs VPA_GLM1 Estimates
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
GLM_AG1
VPA_GLM1
GLM_AG2 Year Effect vs VPA_GLM2 Estimates
3.0
2.5
2.0
1.5
1.0
0.5
0.0
GLM_AG2
VPA_GLM2
GLM_AG3 Year Effect vs VPA_GLM3 Estimates
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
GLM_AG3
VPA_GLM3
6.0
5.0
4.0
3.0
2.0
1.0
0.0
GLM_AG4
VPA_GLM4
10.0
8.0
6.0
4.0
2.0
0.0
GLM_AG5:7
VPA_GLM5:7

Figure 5.

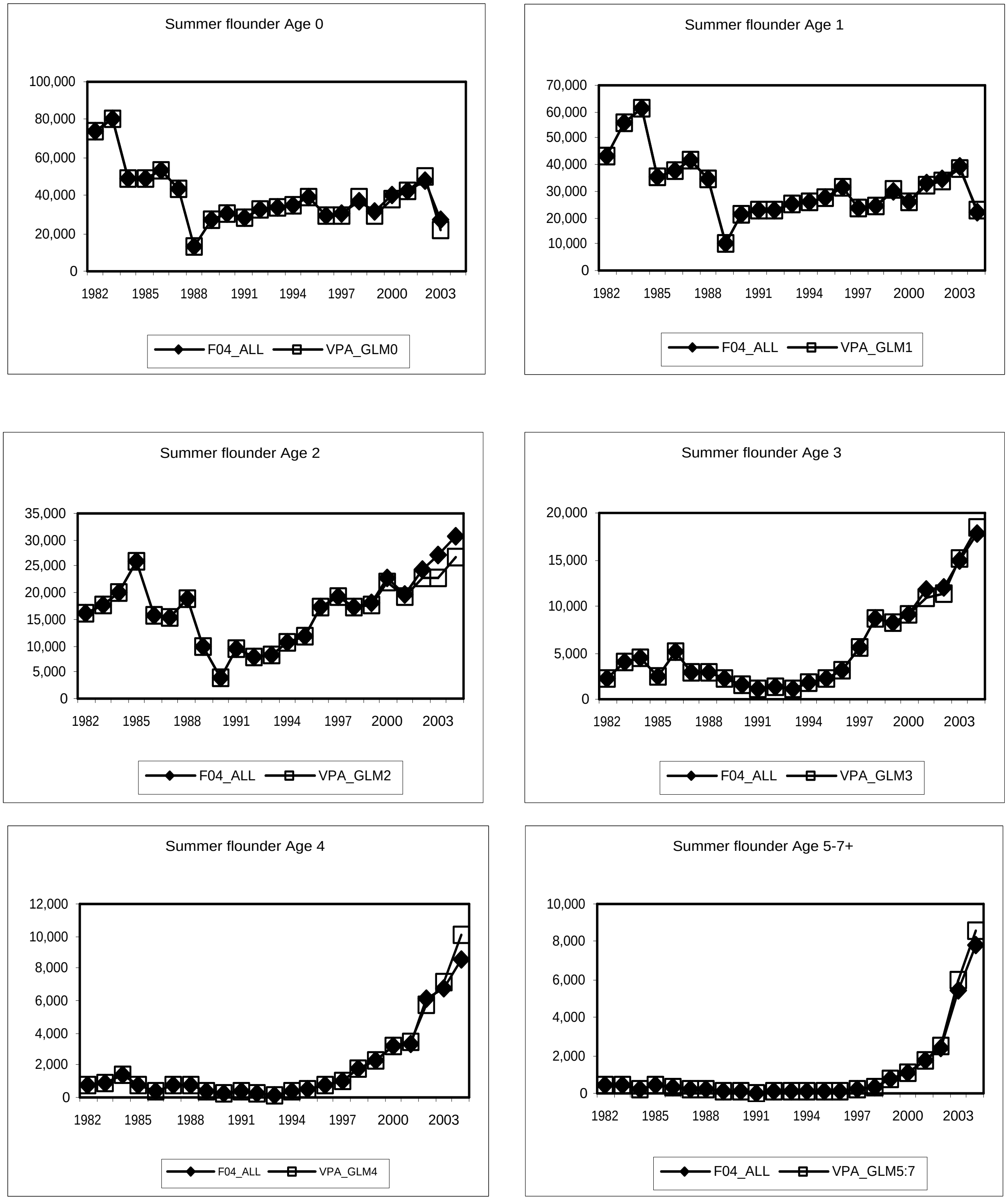
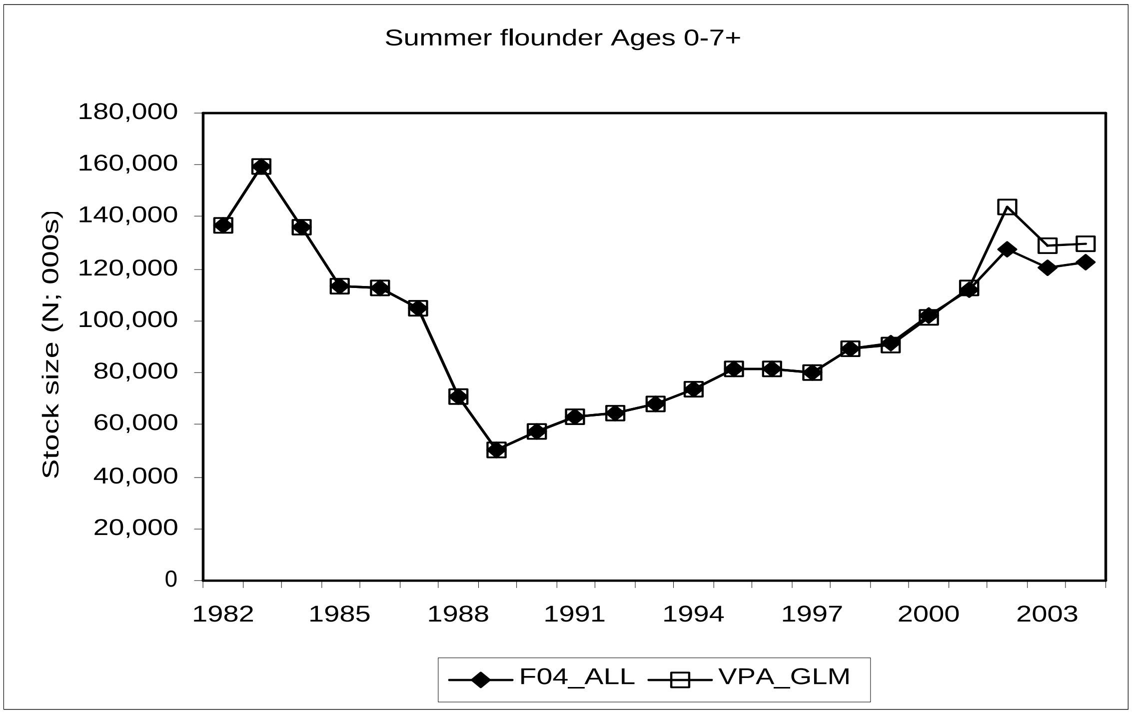

Summer flounder Age 0
100,000
80,000
60,000
40,000
20,000
0
1982 1985 1988 1991 1994 1997 2000 2003
F04_ALL VPA_GLM0
Summer flounder Age 1
70,000
60,000
50,000
40,000
30,000
20,000
10,000
0
1982 1985 1988 1991 1994 1997 2000 2003
F04_ALL VPA_GLM1
Summer flounder Age 2
35,000
30,000
25,000
20,000
15,000
10,000
5,000
0
1982 1985 1988 1991 1994 1997 2000 2003
F04_ALL VPA_GLM2
Summer flounder Age 3
20,000
15,000
10,000
5,000
0
1982 1985 1988 1991 1994 1997 2000 2003
F04_ALL VPA_GLM3
Summer flounder Age 4
12,000
10,000
8,000
6,000
4,000
2,000
0
1982 1985 1988 1991 1994 1997 2000 2003
F04_ALL VPA_GLM4
Summer flounder Age 5-7+
10,000
8,000
6,000
4,000
2,000
0
1982 1985 1988 1991 1994 1997 2000 2003
F04_ALL VPA_GLM5:7

Figure 6.


Summer flounder Ages 0-7+
Stock size (N; 000s)
180,000
160,000
140,000
120,000
100,000
80,000
60,000
40,000
20,000
0
1982
1985
1988
1991
1994
1997
2000
2003
F04_ALL
VPA_GLM

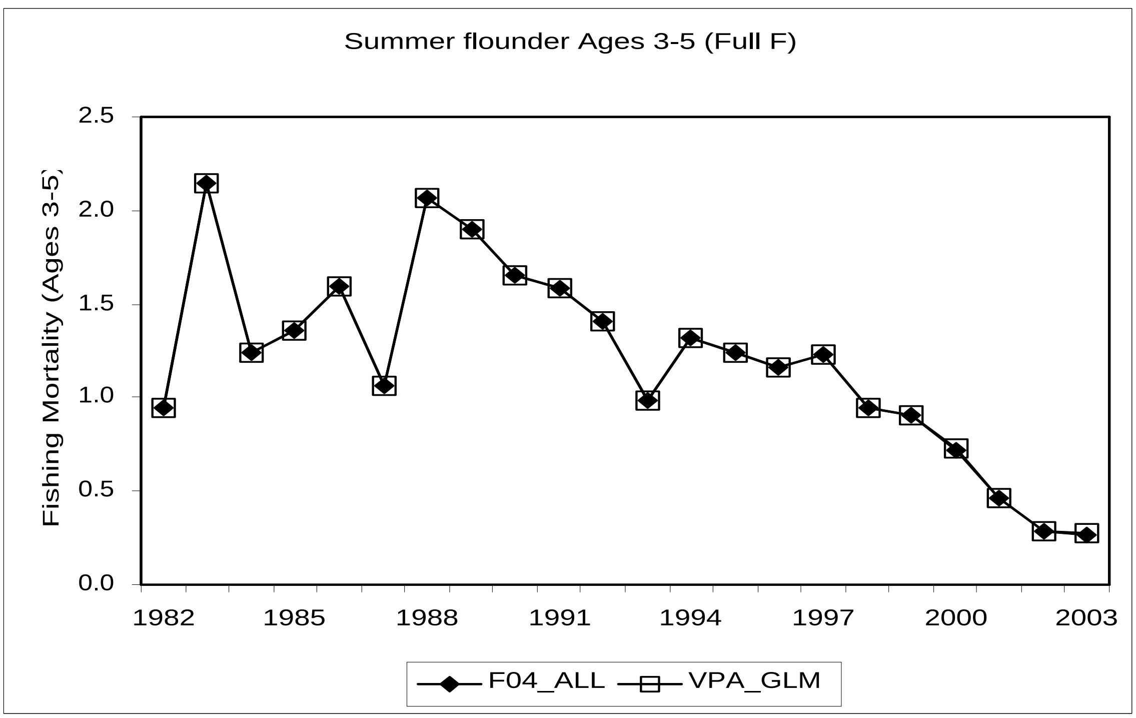
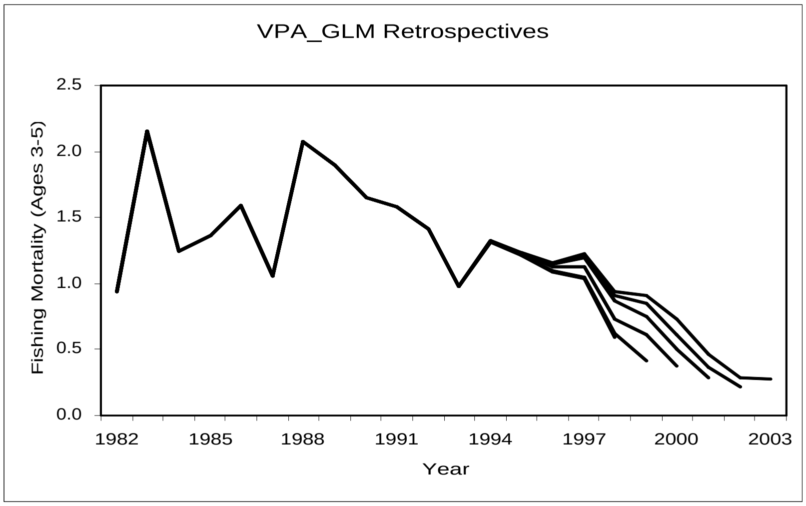
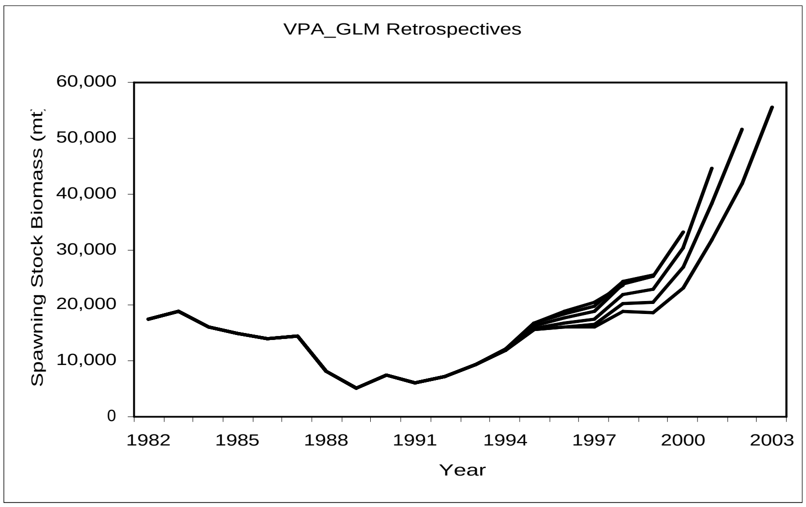

Summer flounder Ages 3-5 (Full F)
Fishing Mortality (Ages 3-5)
2.5
2.0
1.5
1.0
0.5
0.0
1982
1985
1988
1991
1994
1997
2000
2003
F04_ALL
VPA_GLM

Figure 7.


VPA_GLM Retrospectives
Fishing Mortality (Ages 3-5)
2.5
2.0
1.5
1.0
0.5
0.0
1982
1985
1988
1991
1994
1997
2000
2003
Year


VPA_GLM Retrospectives
Spawning Stock Biomass (mt
60,000
50,000
40,000
30,000
20,000
10,000
0
1982
1985
1988
1991
1994
1997
2000
2003
Year

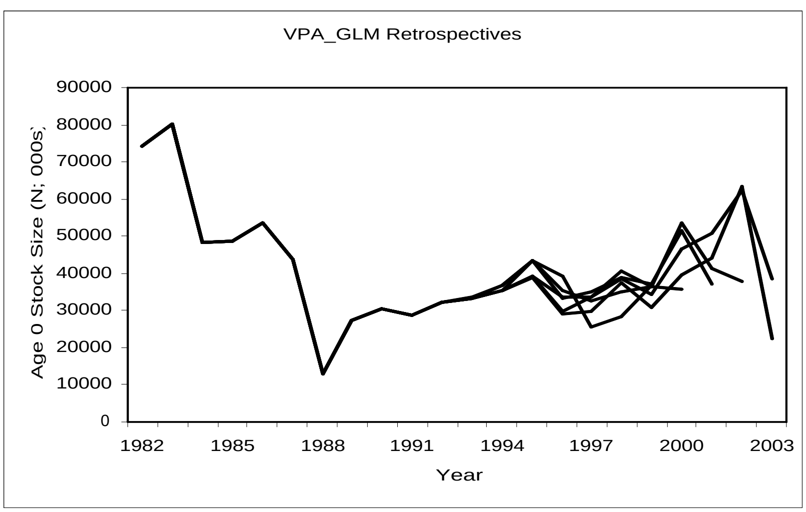

VPA_GLM Retrospectives
Age 0 Stock Size (N; 000s
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
1982
1985
1988
1991
1994
1997
2000
2003
Year

Figure 8.

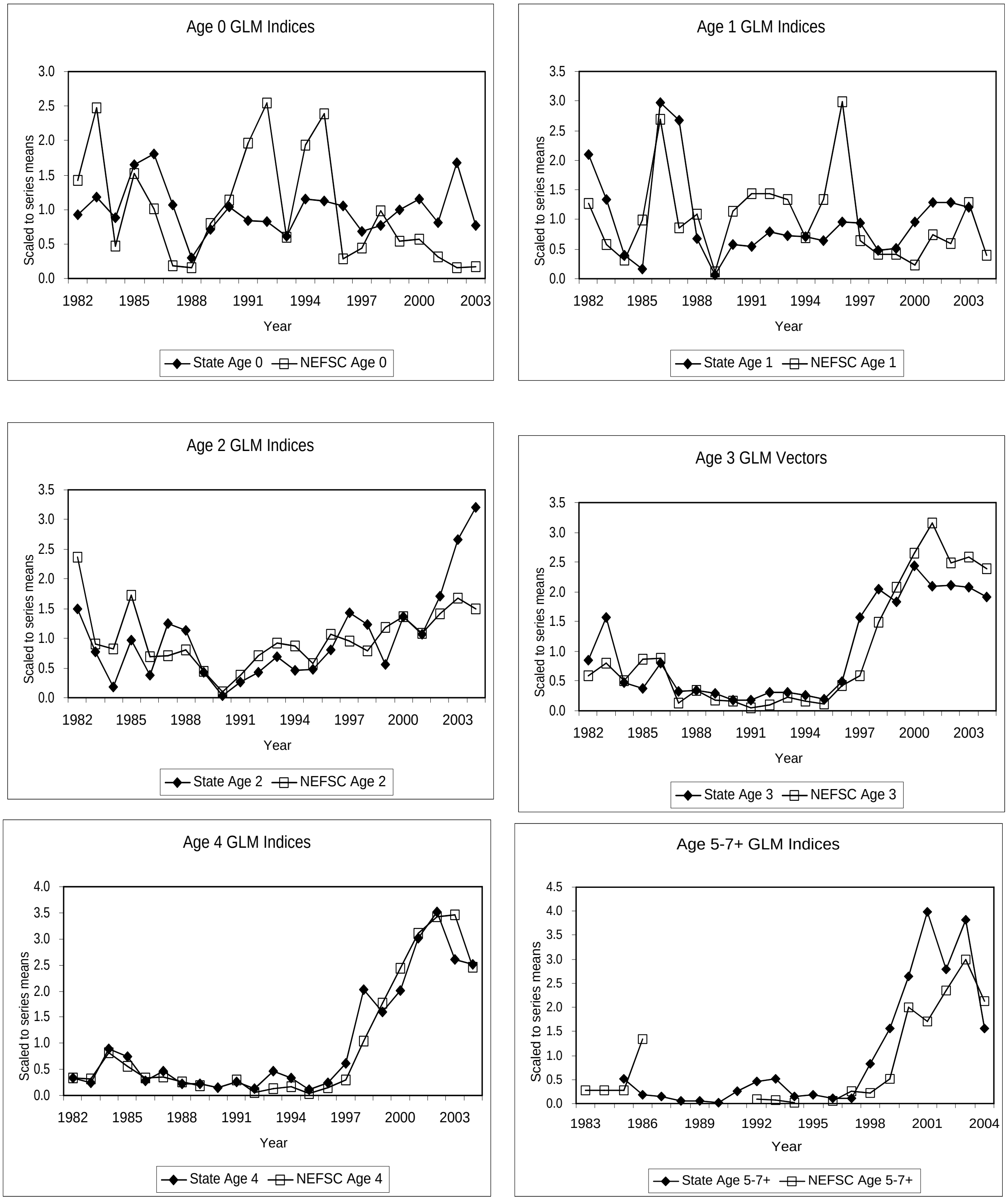
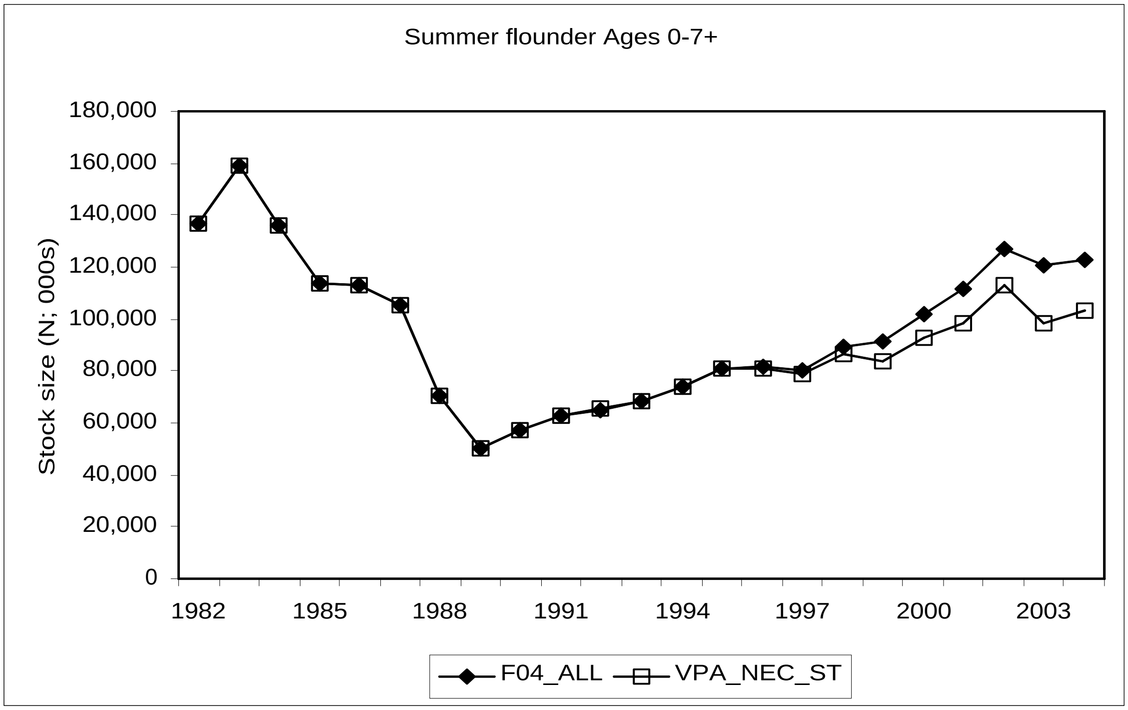

Age 0 GLM Indices
Scaled to series means
Year
State Age 0
NEFSC Age 0
Age 1 GLM Indices
Scaled to series means
Year
State Age 1
NEFSC Age 1
Age 2 GLM Indices
Scaled to series means
Year
State Age 2
NEFSC Age 2
Age 3 GLM Vectors
Scaled to series means
Year
State Age 3
NEFSC Age 3
Age 4 GLM Indices
Scaled to series means
Year
State Age 4
NEFSC Age 4
Age 5-7+ GLM Indices
Scaled to series means
Year
State Age 5-7+
NEFSC Age 5-7+

Figure 9.


Summer flounder Ages 0-7+
Stock size (N; 000s)
180,000
160,000
140,000
120,000
100,000
80,000
60,000
40,000
20,000
0
1982
1985
1988
1991
1994
1997
2000
2003
F04_ALL
VPA_NEC_ST

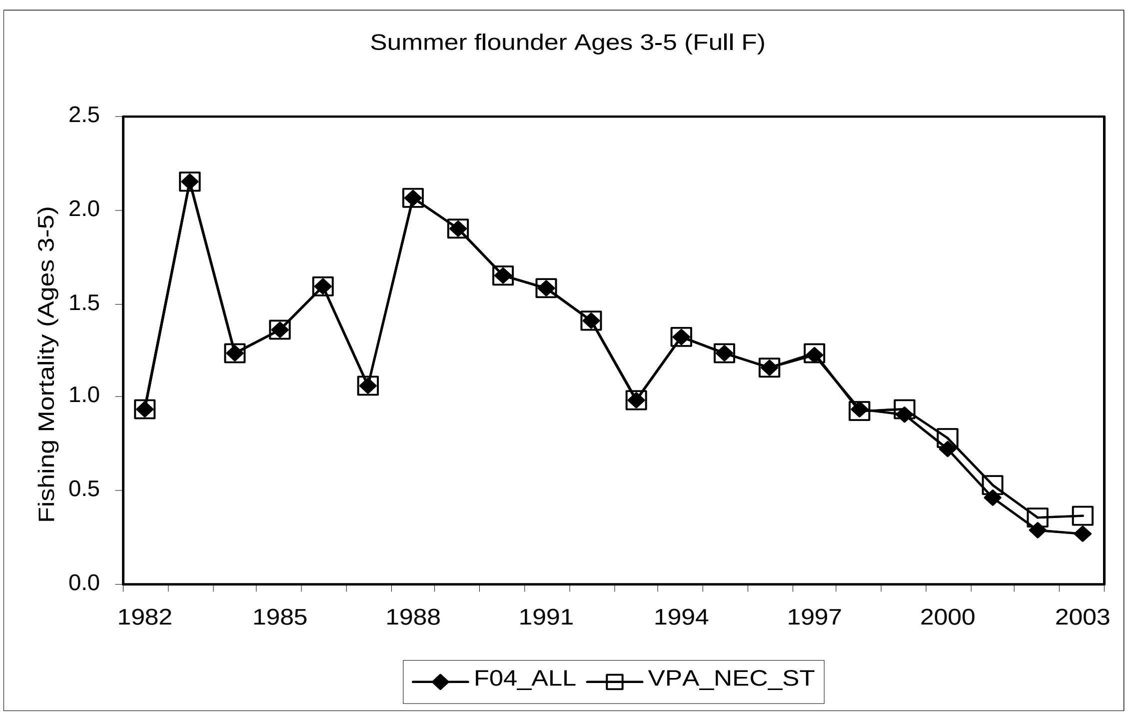

Summer flounder Ages 3-5 (Full F)
Fishing Mortality (Ages 3-5)
2.5
2.0
1.5
1.0
0.5
0.0
1982
1985
1988
1991
1994
1997
2000
2003
F04_ALL
VPA_NEC_ST

Figure 10.

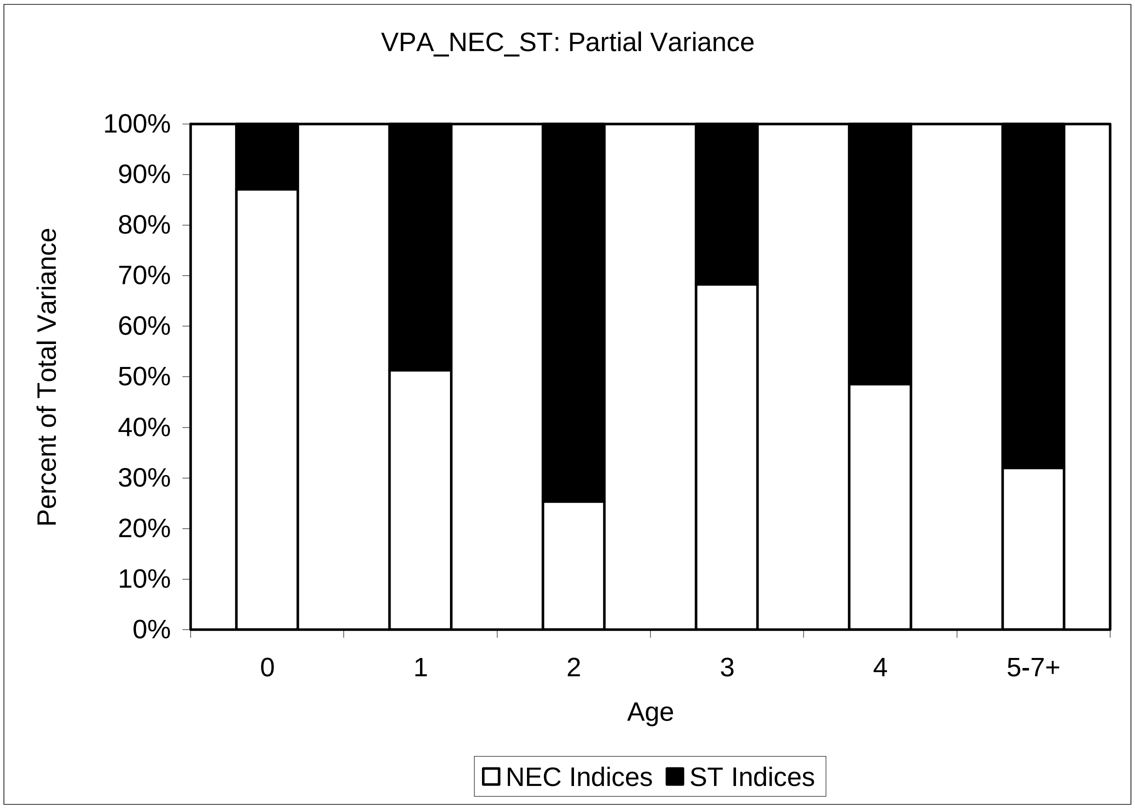

VPA_NEC_ST: Partial Variance
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Percent of Total Variance
0
1
2
3
4
5-7+
Age
NEC Indices
ST Indices

Figure 11.