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.

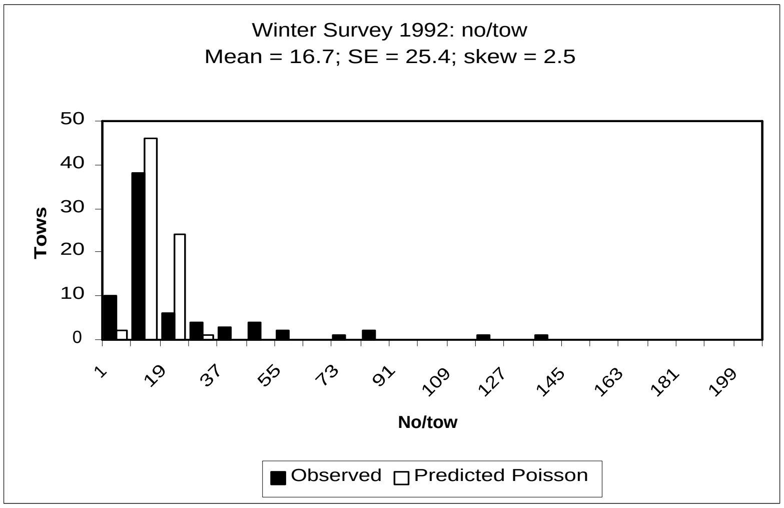
Rohlf FJ. 1981.Statistical tables, second edition. W.H.Freeman and Company. New York. 219 p.

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

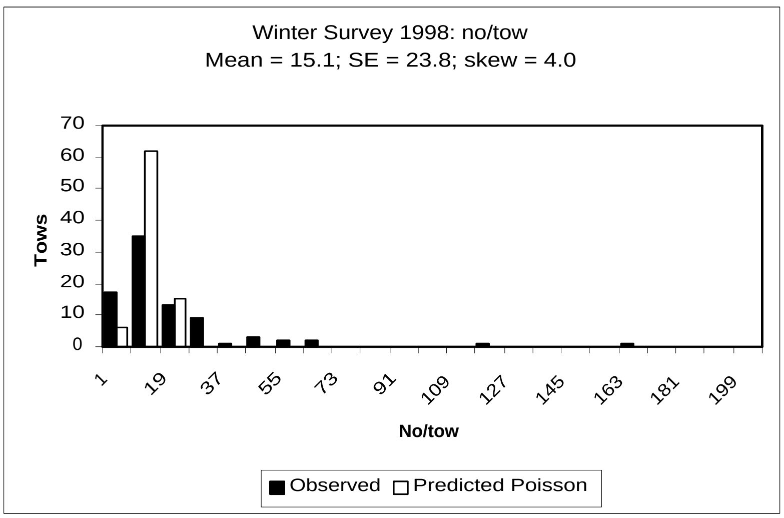
Stansbury DE, Shelton PA, Murphy EF, Healey BP, Brattey J. 2001. An assessment of the cod stock in NAFO divisions 3NO. NAFO SCR Doc. 01/72. Serial No. N4450.

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.

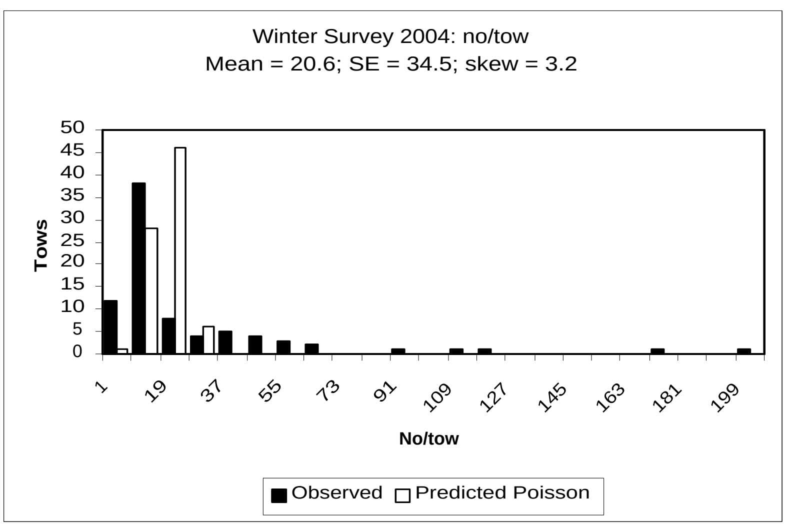
Terceiro M. 2006. Stock assessment of summer flounder for 2006. NEFSC Ref Doc. 06-17. 119 p.



Winter Survey 1992: no/tow
Mean = 16.7; SE = 25.4; skew = 2.5
Tows
No/tow
Observed
Predicted Poisson



Winter Survey 1998: no/tow
Mean = 15.1; SE = 23.8; skew = 4.0
Tows
No/tow
Observed
Predicted Poisson



Winter Survey 2004: no/tow
Mean = 20.6; SE = 34.5; skew = 3.2
Tows
No/tow
Observed
Predicted Poisson

Figure 1

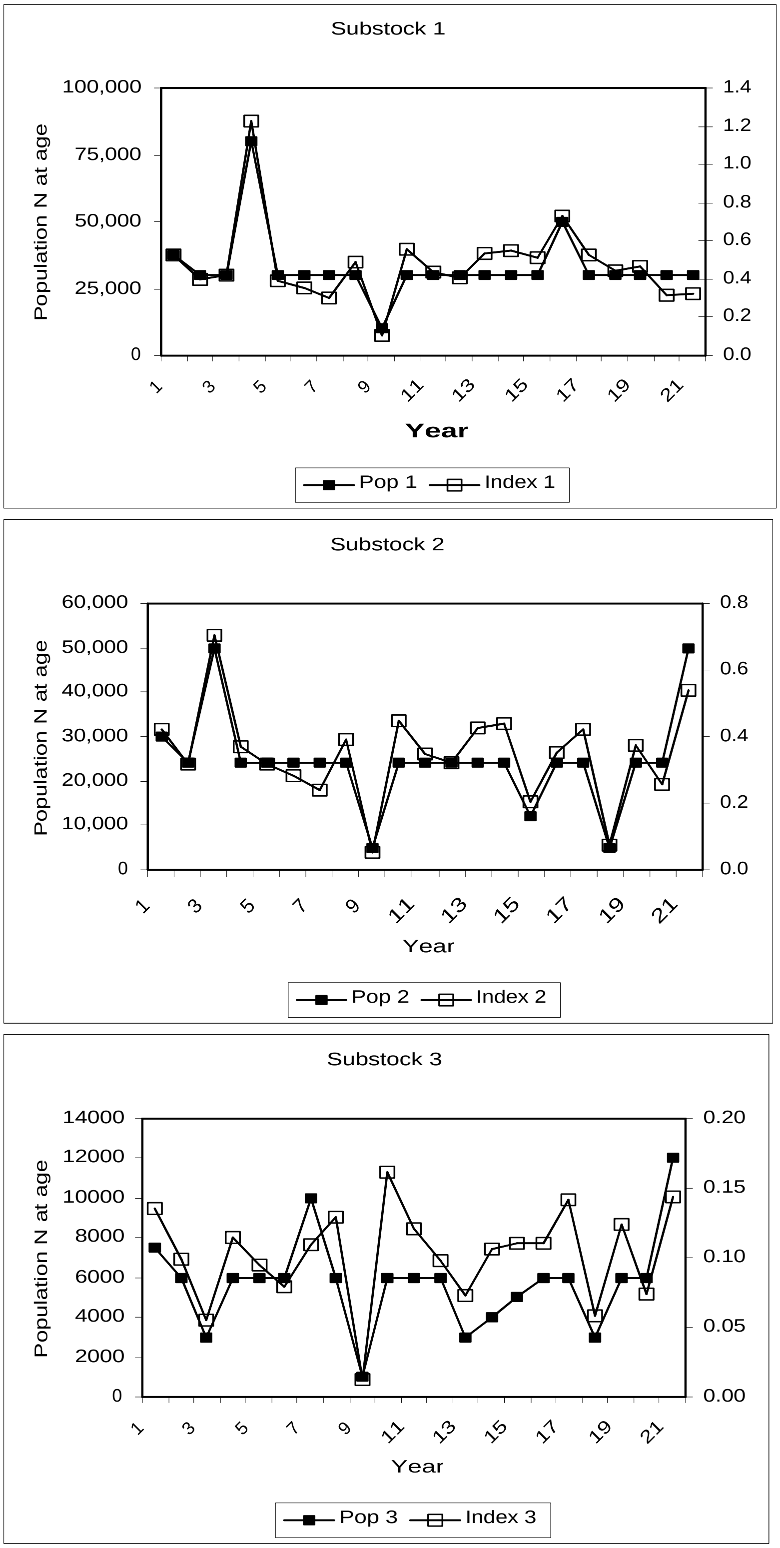
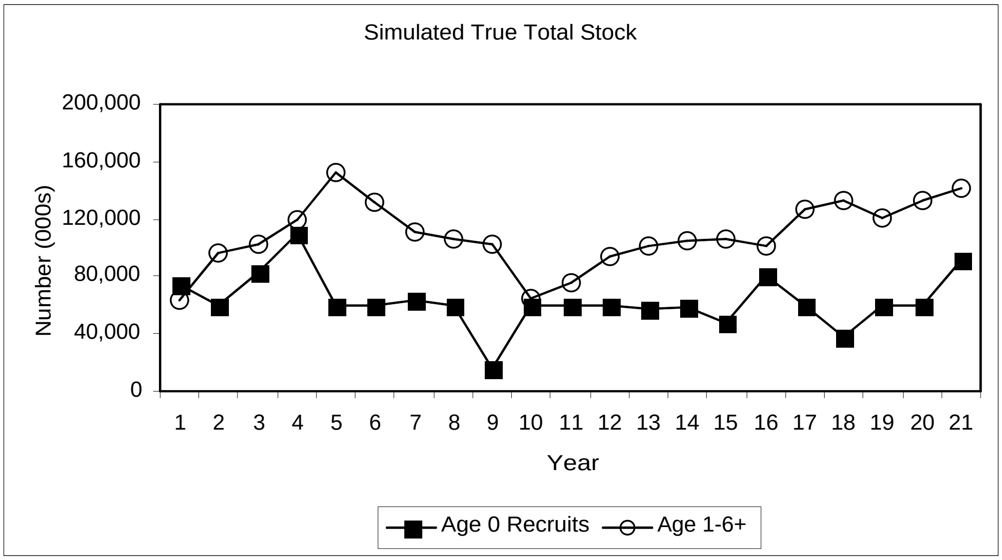

Substock 1
Population N at age
Year
Pop 1
Index 1
Substock 2
Population N at age
Year
Pop 2
Index 2
Substock 3
Population N at age
Year
Pop 3
Index 3

Figure 2.


Simulated True Total Stock
Number (000s)
200,000
160,000
120,000
80,000
40,000
0
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+

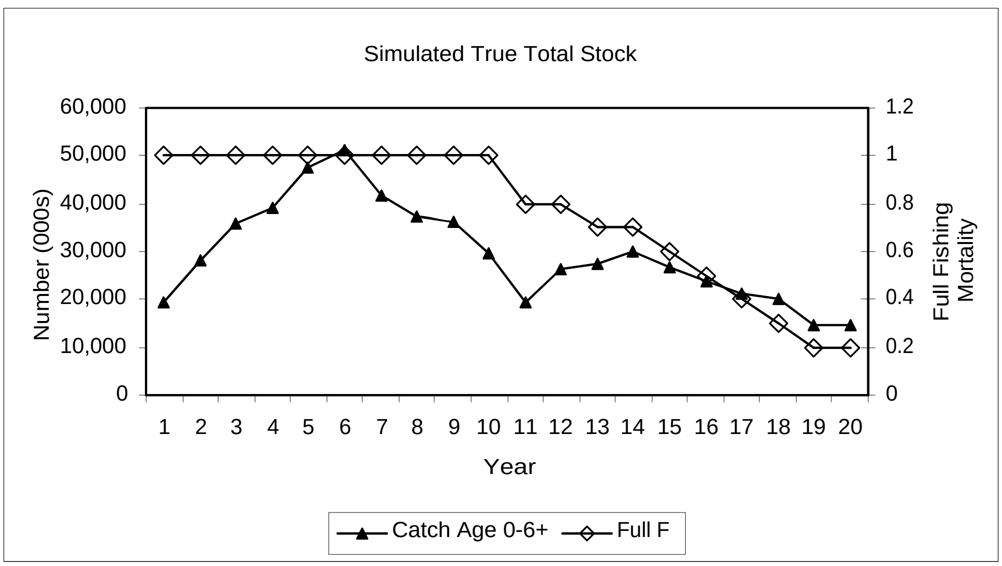

Simulated True Total Stock
Number (000s)
60,000
50,000
40,000
30,000
20,000
10,000
0
1.2
1
0.8
0.6
0.4
0.2
0
Full Fishing Mortality
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Year
Catch Age 0-6+
Full F

Figure 3.

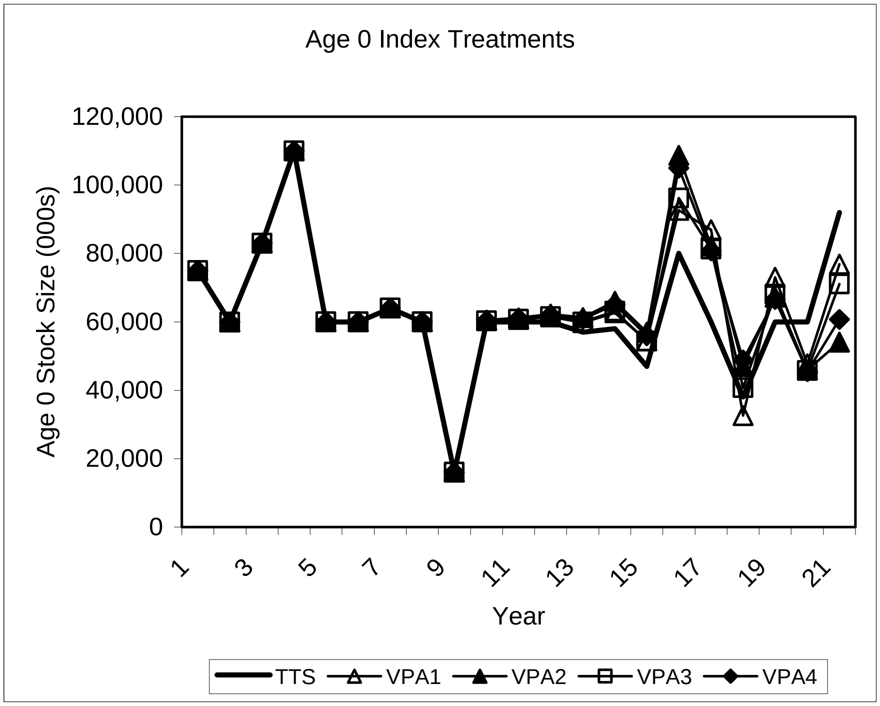
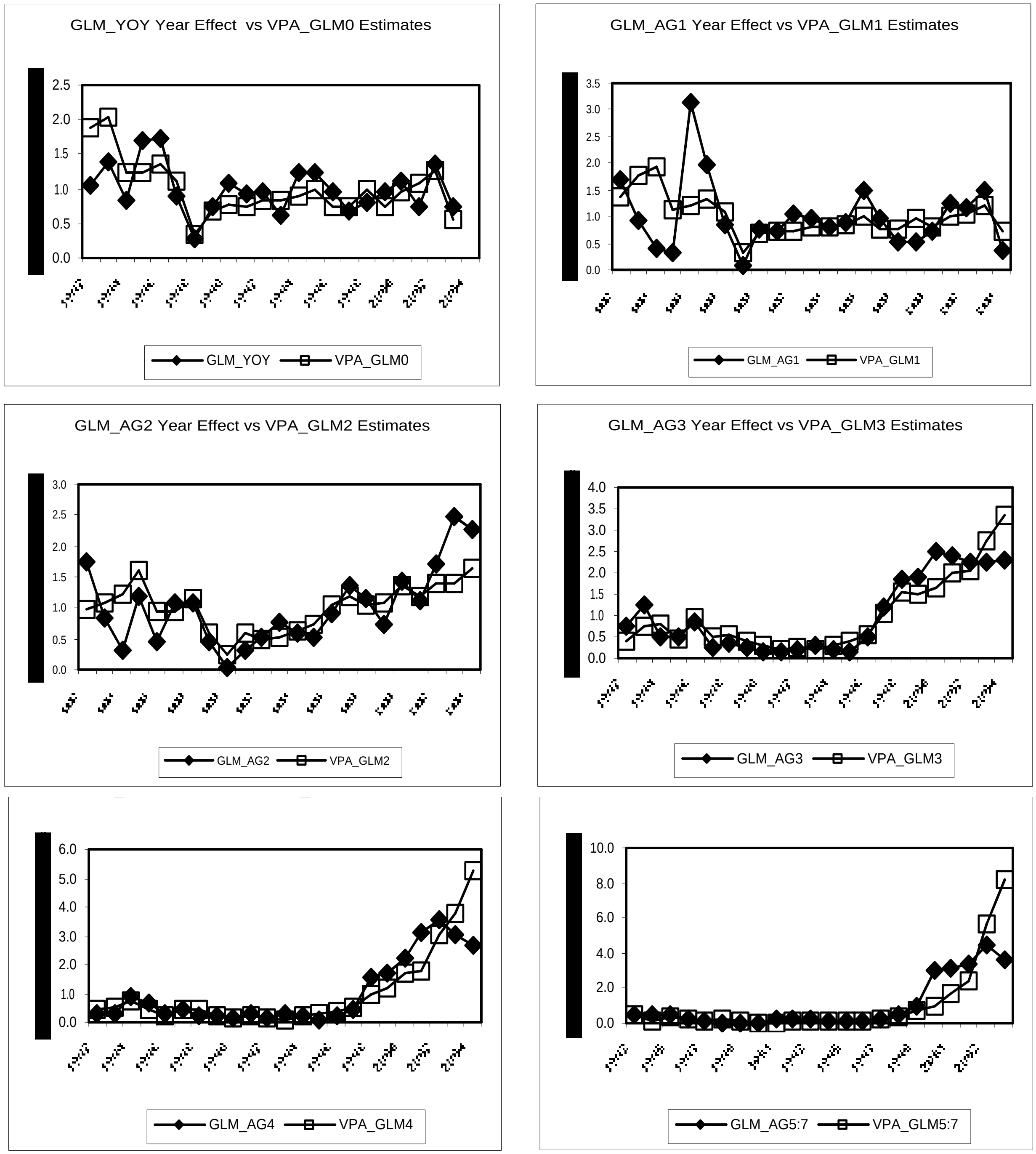

Age 0 Index Treatments
120,000
100,000
80,000
60,000
40,000
20,000
0
Age 0 Stock Size (000s)
1
3
5
7
9
11
13
15
17
19
21
Year
TTS
VPA1
VPA2
VPA3
VPA4

Figure 4.



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

Figure 5.

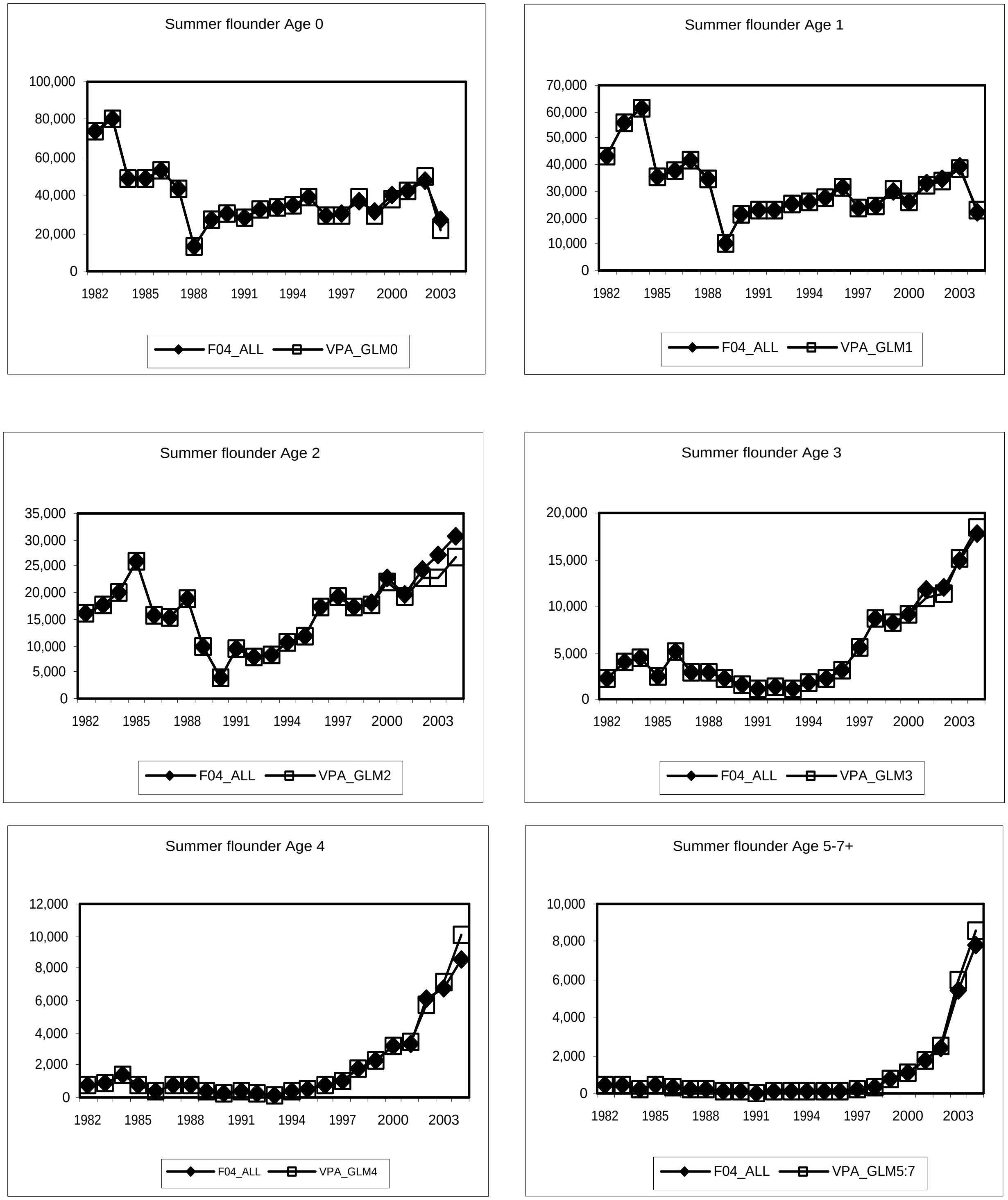
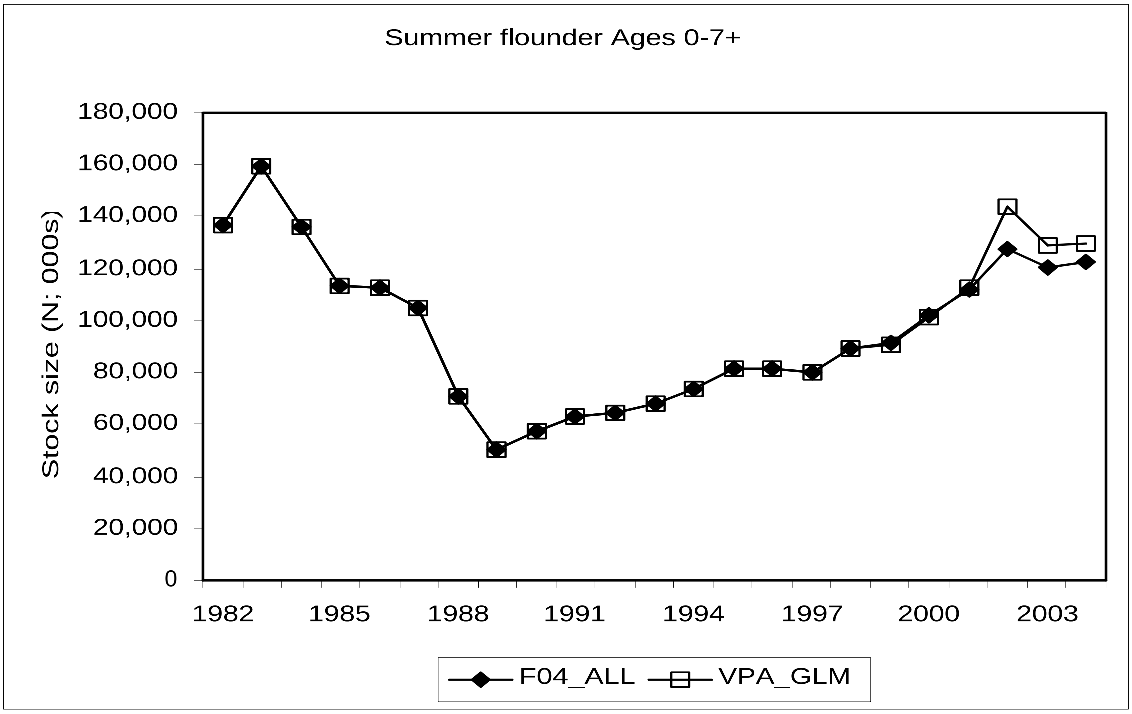

Summer flounder Age 0
100,000
80,000
60,000
40,000
20,000
0
1982 1985 1988 1991 1994 1997 2000 2003
F04_ALL VPA_GLM0
Summer flounder Age 1
70,000
60,000
50,000
40,000
30,000
20,000
10,000
0
1982 1985 1988 1991 1994 1997 2000 2003
F04_ALL VPA_GLM1
Summer flounder Age 2
35,000
30,000
25,000
20,000
15,000
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

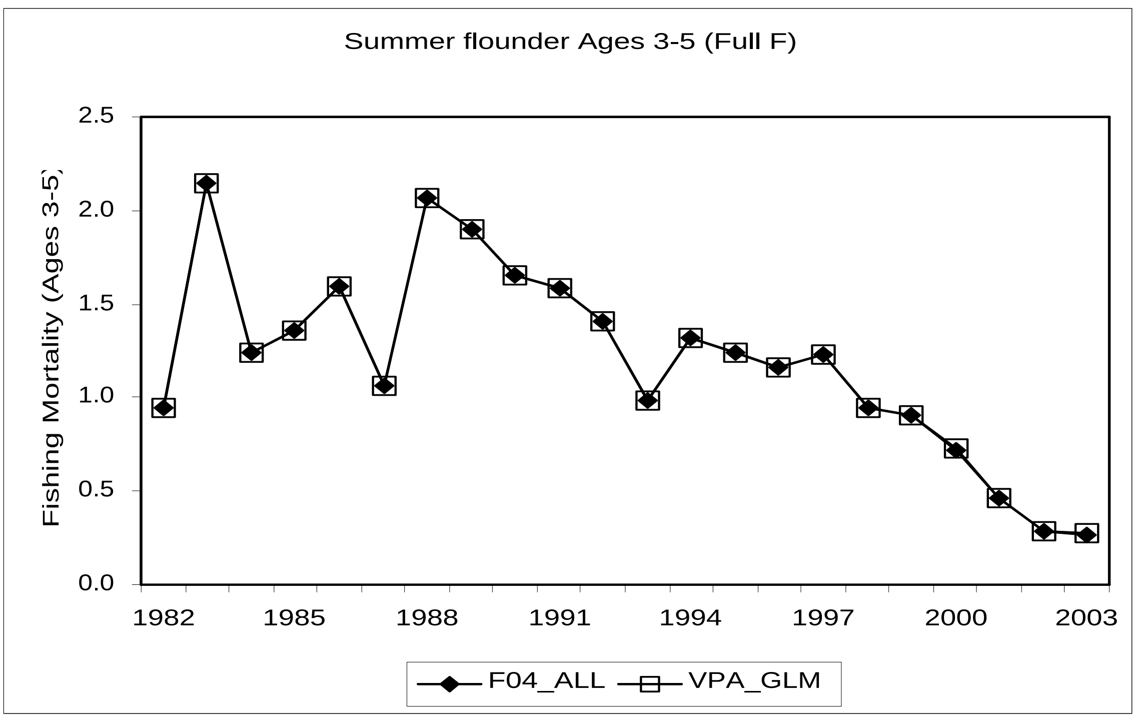

Summer flounder Ages 3-5 (Full F)
Fishing Mortality (Ages 3-5)
2.5
2.0
1.5
1.0
0.5
0.0
1982
1985
1988
1991
1994
1997
2000
2003
F04_ALL
VPA_GLM

Figure 7.

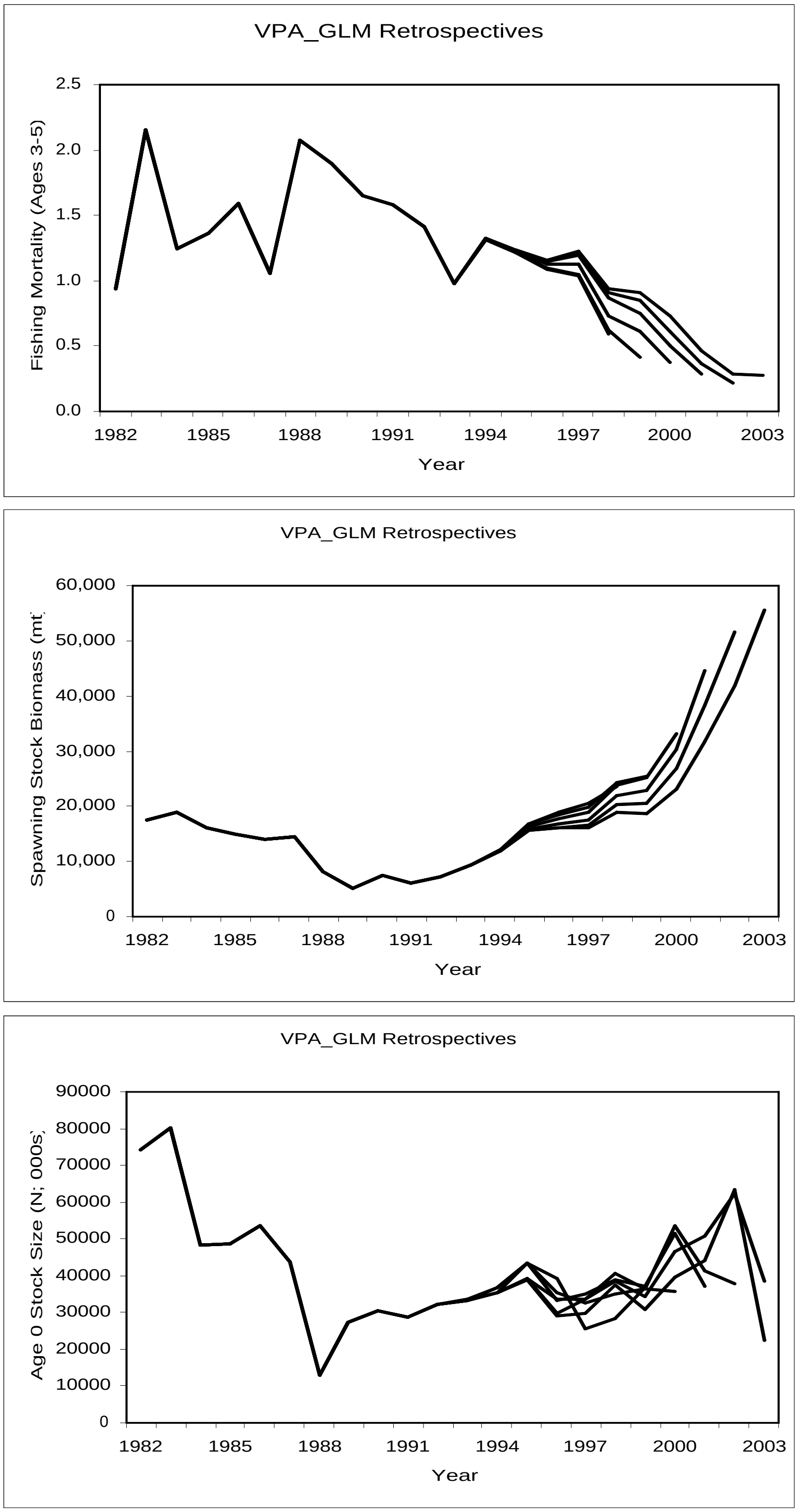

VPA_GLM Retrospectives
Fishing Mortality (Ages 3-5)
2.5
2.0
1.5
1.0
0.5
0.0
1982
1985
1988
1991
1994
1997
2000
2003
Year
VPA_GLM Retrospectives
Spawning Stock Biomass (mt
60,000
50,000
40,000
30,000
20,000
10,000
0
1982
1985
1988
1991
1994
1997
2000
2003
Year
VPA_GLM Retrospectives
Age 0 Stock Size (N; 000s)
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
1982
1985
1988
1991
1994
1997
2000
2003
Year

Figure 8.

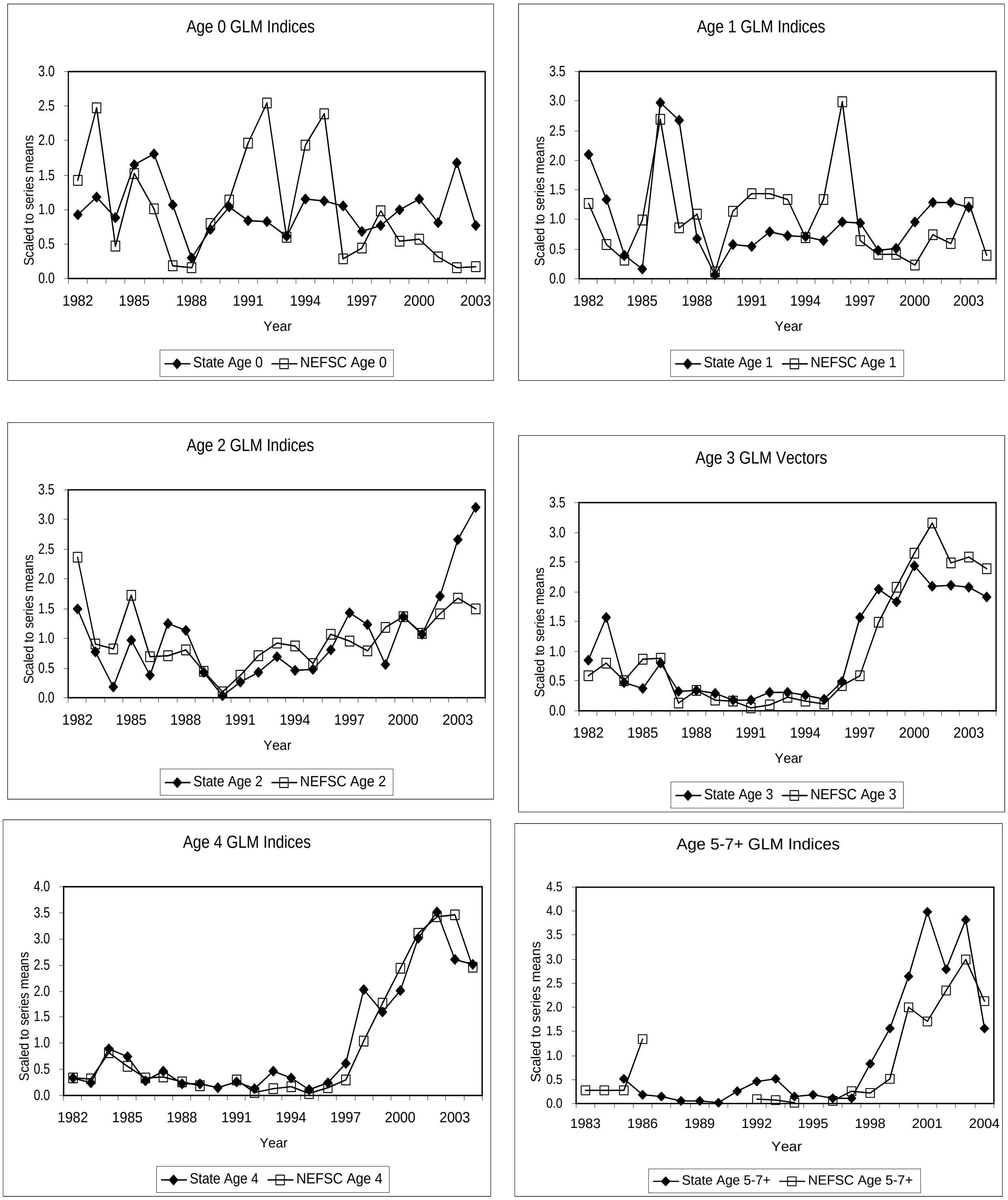
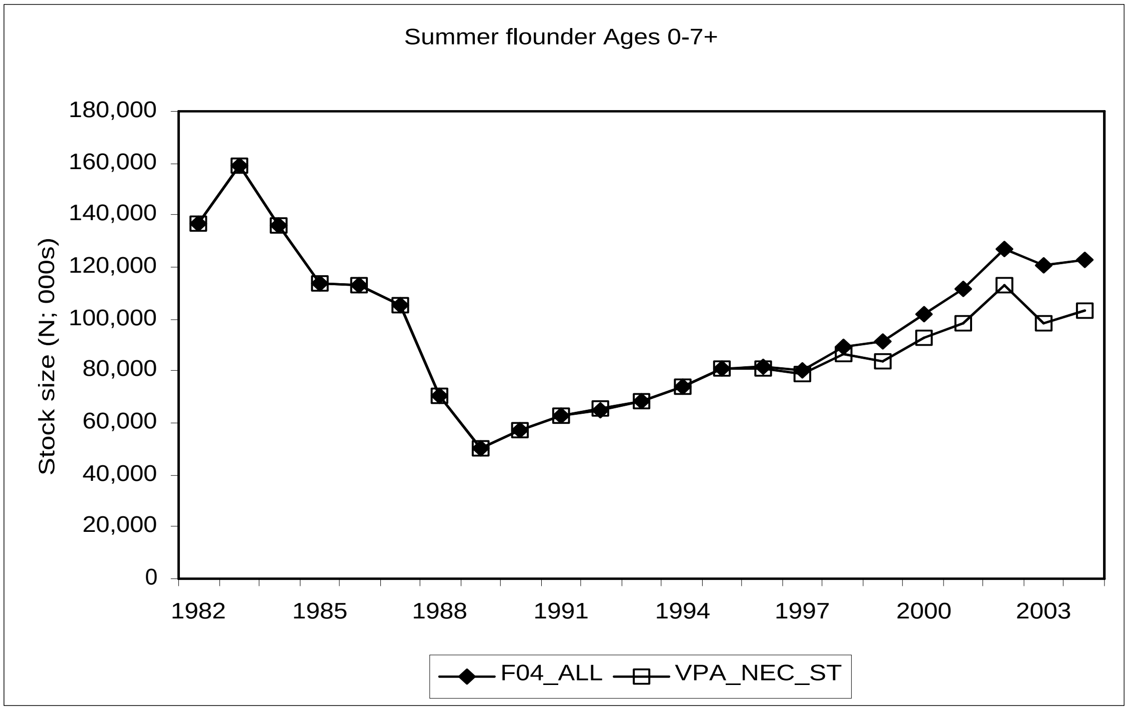

Age 0 GLM Indices
Scaled to series means
Year
State Age 0
NEFSC Age 0
Age 1 GLM Indices
Scaled to series means
Year
State Age 1
NEFSC Age 1
Age 2 GLM Indices
Scaled to series means
Year
State Age 2
NEFSC Age 2
Age 3 GLM Vectors
Scaled to series means
Year
State Age 3
NEFSC Age 3
Age 4 GLM Indices
Scaled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

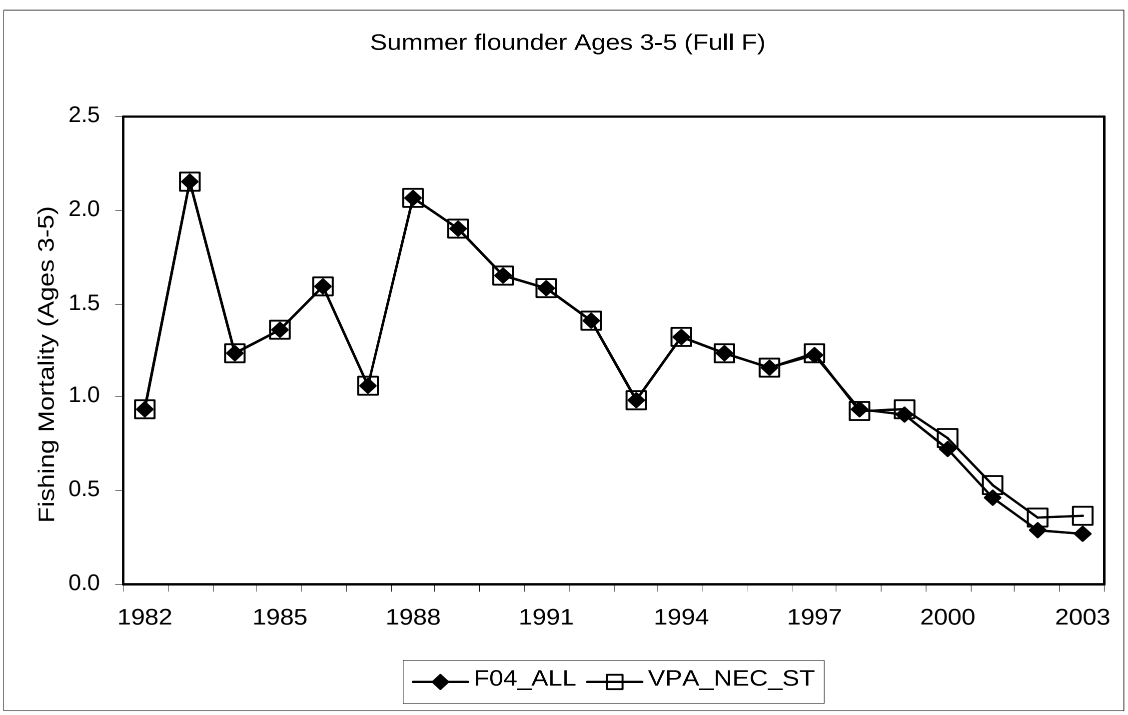
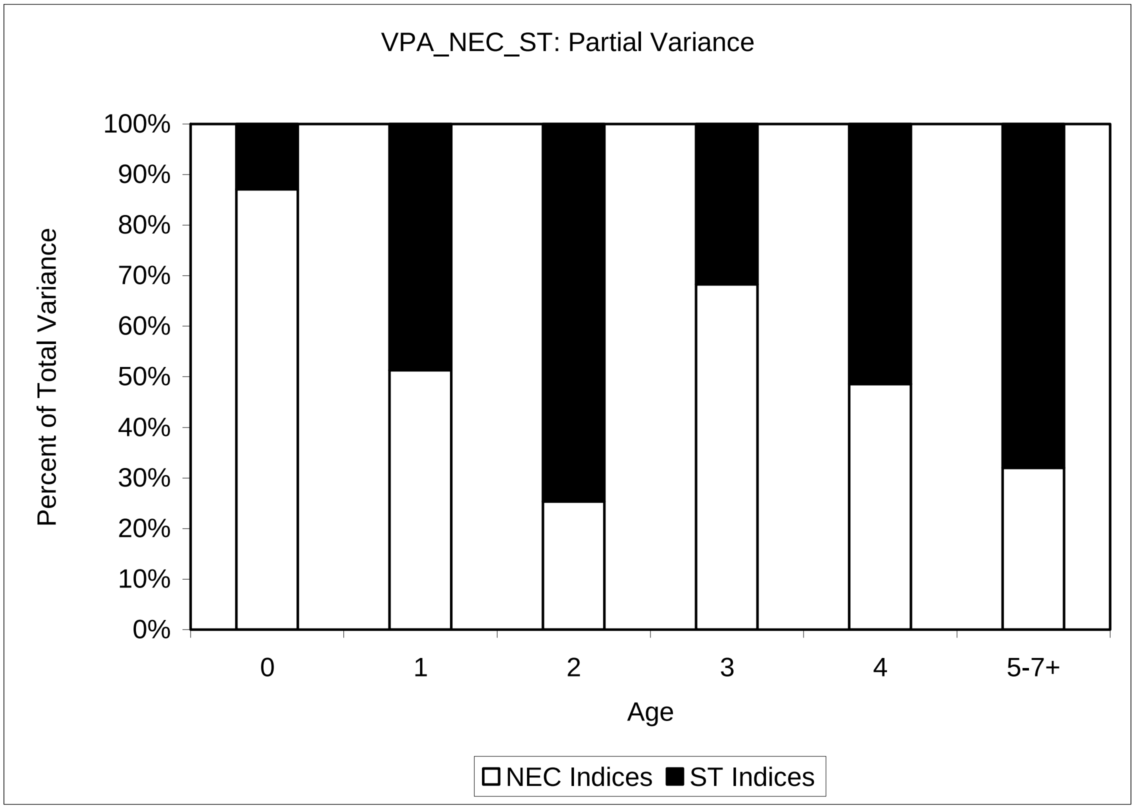

Summer flounder Ages 3-5 (Full F)
Fishing Mortality (Ages 3-5)
2.5
2.0
1.5
1.0
0.5
0.0
1982
1985
1988
1991
1994
1997
2000
2003
F04_ALL
VPA_NEC_ST

Figure 10.



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

Figure 11.

SAW 47 Working Paper 4 (TOR 2b) – Zero Filling

8 November 2006

Simulation Studies of Issues Associated with Filling Zeros in VPA Tuning Indices

by

Chris Legault and Al Seaver

Introduction

Recently, the 2006 assessment of summer flounder (Terceiro 2006) was subject to a NMFS Office of Science and Technology (S&T) Peer Review (Methot 2006). Among the recommendations made by the S&T Peer Review panel was the following:

The Panel finds that one immediate modification of the VPA is justifiable and reduces the retrospective pattern in stock size during 2003-2005. The VPA model currently treats survey observations of zero as missing values. An observation of zero for a particular age of fish in a particular survey year does not mean that there are no fish of that age in the stock, only that the number of survey samples was not sufficient to detect any fish of that age. This VPA model, as with most assessment models, tunes to the logarithm of the survey observations so cannot explicitly deal with observations of zero. However, treating these zeroes as missing values can result in a bias because time periods of low abundance are underrepresented in the data input to the assessment model. In the case of summer flounder, the result may be an underestimate of the degree to which the stock has rebuilt since the low levels that occurred around 1990. The committee did not discuss this issue during the Sept 14-15 meeting, so is not prepared to present a definitive solution. An interim approach would use a small value in place of the zeroes. A value equal to one sixth of the smallest observed positive value would be reasonable until a more complete statistical solution can be developed.

This recommendation departs from the standard practice in NEFSC assessments of treating zero values in tuning indices as missing values. To more fully understand the implications of this suggested change, two types of simulation analyses were conducted. The first is a simple spreadsheet example of how a single time series is impacted by different levels of fish detection and the implications for a full VPA. The second is a full simulation that generates many random sets of data for VPA from a known case, creates zeros for some of the indices in some years, and compares different methods for dealing with these zeros, including treating them as missing values, replacing the zeros with a fixed small value, and the one sixth of the smallest observation rule.

First Study: Impact of Zeros on One Time Series

A population that declined and then increased was created artificially. A catchability coefficient was applied to generate a survey time series exactly from the data. The values in the time series were rounded to two, one, and zero decimal places creating observations of zero for 2, 4, and 7 years, respectively (Table 1). A series of constants was added to the time series ranging from 0.0001 to 10 so that the holes were filled. A new catchability coefficient was calculated that minimized the difference between the true population and the observed survey time series which had been modified to fill the holes. This was done to show how a model would need to change the predicted values to more closely match the observed series. In this study, treating the index values as missing results in an exact match between the observed and predicted values, due to the formulation of the problem and so are not considered further.

The differences between observed and predicted values depend strongly on the constant added to the time series (Figs 1-3). Adding a large value, such as 10, causes the survey time series to flatten relative to the true population. A model would try to reduce the change in the population in this case. Conversely, adding a very small value, such as 0.0001, causes the survey time series to exhibit a stronger decline and recovery than the true population. In this case, a model would try to increase the changes in the population. Adding one sixth of the minimum observed value appears to be an objective way to determine a value that is not too big or too small for the round 2 case where only two zeros are replaced.

However, the more disturbing result seen in these simulations is that the addition of a constant value to replace the zeros in a survey time series artificially imposes a pattern that may not match the actual pattern in the population. This is most clearly seen in the round 0 case where seven zeros are filled with the same value even though the true population declines then increases during the seven year period.

Second Study: Simulation Analysis of Different Methods of Treating Zeros

A comparative study was performed using the POPCOMP length based population simulator tool and VPA version 2.3.3. The objective was to examine the effects of using indices of abundance with some portion of the index data treated as missing or alternatively replaced with an imputed value. Four scenarios were examined. In each case the simulated data were sampled to create 100 realizations of VPA input data and the results of the multiple realizations were compared in their ability to recover the true stock numbers and fishing mortality at age. The test was performed in such a manner as the VPA files created for each realization would be the same for each scenario except in the specified removal and alternative replacement of index data based on an input cut point.

The simulated population was loosely based on the summer flounder assessment with the population initially declining due to high F (>2) and then rebuilding as F was lowered to <0.5). The simulated population spans 24 years starting in 1982 and consisting of 8 age classes with the last age class acting as a plus group. Natural mortality was 0.2 for all ages and years. Both recruitment and fishing mortality vary widely over the time series. The growth projection matrix was created using von Bertalanffy growth coefficients and length bins ranging from 10 to 84 cm. A logistic equation was used for fishery selectivity at length. Catch was removed from the population based on the true F but samples were collected from four market categories based on size (sample sizes 65-133 per 100 metric tons) to introduce variability in the catch at length. Age-length keys were created based on sampling 25% of the observed lengths and an ageing error matrix was included to introduce variability in the catch at age (mis-aged proportions ranged from 4% to 17%). The length-weight equation coefficients supplied to allow expansion of sampled catch to total landings, which had a small amount of variability relative to the true landings (CV = 0.01). Discards were not included in this simulation. This level of uncertainty in the catch at age is thought to be representative of the level associated with the summer flounder assessment. However, there is not a retrospective pattern observed when the simulated data are analyzed with VPA, so not all sources of uncertainty have been captured.

There was only one index generated for each age. The catchability for each index was chosen so that catchability increased with age (Table 2). The uncertainty was higher for the indices at younger ages than older ages (Table 2). The coefficients of variation were used to generate lognormally distributed error in the observed indices. The population trends, catchability coefficients, and coefficients of variation combined to produce different probabilities that a given index value would fall below 1.0 (Table 3). Index values below 1.0 were treated in four different ways:

- Case 1 - Actual values used
- Case 2 - Replaced with 0.0 and treated as missing
- Case 3 - Replaced with the arbitrary constant 0.01
- Case 4 - Replaced with 0.0 then a constant of 1/6 times the smallest non-zero element in the index vector added to all index vector elements including zeros.

The VPA input files generated for each realization were identical excepting that indices of abundance were altered by case.

The median values of F and N at age from the 100 realizations of the VPA model under the four cases of treating index values below 1.0 were compared with the true values from the simulated population (Tables 4-5 and Figs 4-5). Due to the convergence properties of VPA, the medians from the 4 cases are essentially identical for years 1982-1994, as seen in Figures 4-5, and so are not shown in Tables 4-5. The most striking feature seen in the tables and figures is the poor performance of Case 3 (replacing zeros with the arbitrary constant 0.01). The fishing mortality rates in Case 3 were well below the true values while the estimated population abundances were well above the true values. Case 3 clearly demonstrates the potential for introducing bias by replacing zeros in tuning index time series with an arbitrary constant. While not as clear, generally the Case 4 (add 1/6 of smallest non-zero element) estimates were more biased than the Case 2 (treat zeros as missing) estimates. The exception to this generality is seen in age 1 results where the VPA formulation had to be modified slightly to estimate only ages 3-8 in the terminal year +1 due to the lack of information for age 2 in the terminal year +1 when the index was zero. For older ages, Case 2 actually outperformed Case 1 (all data used) relative to the truth. It is not clear why this happened and may be an artifact of the bias introduced by the mis-ageing matrix used to generate the catch data. However, even if Case 1 is used as the basis for comparison, instead of the true values, Case 2 performs at least as well as Case 4 for all ages except age 1.

Discussion

An alternative method to determining the constant to use in place of zeros that was not considered in this exercise is provided by Berry (1987). The Berry approach consists of finding the constant that minimizes a function of the skewness plus kurtosis of the raw data. This approach is not appropriate for use with tuning index data because the residuals are assumed to follow a lognormal distribution, not the raw observations.

While the 1/6 of the smallest non-zero approach appears to provide reasonable results in some cases, it is an arbitrary rule. In some situations, 1/5 or 1/7 of the smallest non-zero index value would perform better than 1/6. The main problem remains however. Filling zeros with a constant value, no matter how that constant is selected, creates a pattern that may not match reality. These simulations show that this approach can produce results further from the truth than treating zeros as missing values.

Of course, in reality the zeros do have information. Results should be checked to ensure that predicted values are not high when index is zero. If an assessment model predicts high abundance for a year-age combination that had a zero index, the model results should be questioned. However, adding incorrect information arbitrarily has the potential to bias the results, as demonstrated in these simulations.

Conclusions

The two simulation studies have demonstrated problems that can arise when tuning indices with zero values are replaced with arbitrary constants. This practice assumes that the correct magnitude can be chosen to fill the zeros and that it is better to provide the model with information that the index is low rather than treat the data as missing. Results demonstrate that this premise is not always correct. Thus, we recommend the NEFSC standard approach of treating zero values in tuning indices for VPA as missing values.

References Cited

Berry DA. 1987. Logarithmic transformations in ANOVA. Biometrics 43:439-456.

Methot R. 2006. Review of the 2006 Summer Flounder Assessment Update. Chair's Report. NMFS Office of Science and Technology. 6 p.

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

Table 1. Artificial time series for a population and the associated time series of indices given a catchability of 0.000002 when the values are rounded to two, one, and zero decimal places. Highlighted cells are years when the tuning index has an observed zero.

		Index		
Year	Population	Round 2	Round 1	Round 0
1980	2000000	4.00	4.0	4
1981	1500000	3.00	3.0	3
1982	1300000	2.60	2.6	3
1983	1000000	2.00	2.0	2
1984	500000	1.00	1.0	1
1985	300000	0.60	0.6	1
1986	200000	0.40	0.4	0
1987	10000	0.02	0.0	0
1988	5000	0.01	0.0	0
1989	1000	0.00	0.0	0
1990	2000	0.00	0.0	0
1991	50000	0.10	0.1	0
1992	100000	0.20	0.2	0
1993	300000	0.60	0.6	1
1994	400000	0.80	0.8	1
1995	700000	1.40	1.4	1
1996	1200000	2.40	2.4	2
1997	1500000	3.00	3.0	3
1998	1100000	2.20	2.2	2
1999	1200000	2.40	2.4	2
2000	1700000	3.40	3.4	3

Table 2. Catchability coefficients (q) and coefficients of variation (CV) by age for the tuning indices used in the second study. The q values multiplied the true populations at age to generate the expected values for the indices by year. The CV values describe the amount of lognormally distributed error used to create the random VPA input data.

Param	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8
q	0.00000003	0.0000001	0.0000001	0.00001	0.00001	0.00001	0.00001	0.00001
CV	0.5	0.5	0.5	0.3	0.3	0.3	0.3	0.3

Table 3. Probability that an index value will be below 1.0 and thus set to zero given the true population, catchability coefficient, and uncertainty associated with each index and year.

Year	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8
1982	6.9	<1.0	16.4	<1.0	<1.0	6.4	9.5	94.6
1983	4.9	<1.0	2.1	<1.0	<1.0	<1.0	98.5	89.6
1984	28.0	<1.0	4.4	<1.0	<1.0	46.6	>99.0	>99.0
1985	27.8	<1.0	1.2	<1.0	<1.0	<1.0	>99.0	>99.0
1986	21.6	<1.0	10.9	<1.0	<1.0	<1.0	85.8	>99.0
1987	35.5	<1.0	14.0	<1.0	<1.0	<1.0	63.4	>99.0
1988	98.6	1.0	5.8	<1.0	<1.0	<1.0	6.6	>99.0
1989	73.8	64.1	26.5	<1.0	<1.0	<1.0	58.5	>99.0
1990	65.9	11.2	97.6	<1.0	<1.0	1.5	>99.0	>99.0
1991	70.0	6.9	59.9	<1.0	<1.0	0.8	>99.0	>99.0
1992	60.7	8.5	47.8	<1.0	<1.0	1.5	>99.0	>99.0
1993	58.4	4.9	48.1	<1.0	<1.0	84.2	>99.0	>99.0
1994	53.2	3.8	27.8	<1.0	<1.0	<1.0	>99.0	>99.0
1995	45.3	3.2	31.0	<1.0	<1.0	<1.0	92.5	>99.0
1996	70.8	1.9	26.3	<1.0	<1.0	<1.0	64.5	>99.0
1997	68.8	7.9	18.7	<1.0	<1.0	<1.0	51.2	>99.0
1998	64.2	7.3	42.7	<1.0	<1.0	<1.0	36.4	>99.0
1999	69.5	5.3	34.4	<1.0	<1.0	<1.0	10.8	>99.0
2000	53.7	7.0	29.1	<1.0	<1.0	<1.0	1.7	96.7
2001	69.0	2.8	31.7	<1.0	<1.0	<1.0	<1.0	71.5
2002	56.2	6.3	15.0	<1.0	<1.0	<1.0	<1.0	8.4
2003	80.3	2.9	21.5	<1.0	<1.0	<1.0	<1.0	<1.0
2004	55.1	11.4	12.2	<1.0	<1.0	<1.0	<1.0	<1.0
2005	97.5	2.7	31.9	<1.0	<1.0	<1.0	<1.0	<1.0
2006	97.5	46.6	12.4	<1.0	<1.0	<1.0	<1.0	<1.0

Table 4. Comparison of true fishing mortality at age with medians from 100 realizations under the four cases of treating index values less than 1.0.

		F at Age					Percent bias in Medians vs Truth			
Age	Year	True	Case 1	Case 2	Case 3	Case 4	Case 1	Case 2	Case 3	Case 4
1	1995	0.084	0.084	0.084	0.082	0.084	0	0	-3	0
1	1996	0.079	0.078	0.078	0.073	0.078	-1	-1	-7	-1
1	1997	0.085	0.082	0.082	0.073	0.082	-3	-3	-14	-3
1	1998	0.066	0.065	0.066	0.049	0.064	-1	0	-25	-3
1	1999	0.067	0.064	0.065	0.039	0.061	-5	-3	-42	-9
1	2000	0.060	0.058	0.060	0.031	0.055	-3	0	-49	-8
1	2001	0.046	0.044	0.045	0.038	0.044	-3	-1	-17	-3
1	2002	0.033	0.031	0.032	0.026	0.032	-4	-3	-20	-3
1	2003	0.032	0.031	0.031	0.029	0.032	-1	-2	-10	0
1	2004	0.032	0.031	0.033	0.025	0.029	-3	3	-21	-10
1	2005	0.036	0.030	0.024	0.057	0.034	-16	-34	59	-5
2	1995	0.378	0.380	0.381	0.377	0.380	1	1	0	1
2	1996	0.356	0.370	0.370	0.361	0.369	4	4	1	4
2	1997	0.380	0.385	0.385	0.358	0.382	1	1	-6	1
2	1998	0.299	0.299	0.301	0.256	0.297	0	1	-14	-1
2	1999	0.306	0.305	0.308	0.223	0.298	0	1	-27	-3
2	2000	0.272	0.266	0.270	0.155	0.253	-2	-1	-43	-7
2	2001	0.210	0.208	0.213	0.104	0.196	-1	1	-51	-7
2	2002	0.151	0.150	0.151	0.129	0.150	-1	0	-15	-1
2	2003	0.147	0.144	0.146	0.119	0.142	-2	0	-19	-3
2	2004	0.147	0.152	0.151	0.137	0.154	3	3	-6	5
2	2005	0.167	0.170	0.173	0.135	0.156	2	4	-19	-6
3	1995	0.730	0.709	0.710	0.705	0.709	-3	-3	-3	-3
3	1996	0.688	0.669	0.669	0.657	0.668	-3	-3	-4	-3
3	1997	0.737	0.732	0.733	0.700	0.729	-1	-1	-5	-1
3	1998	0.578	0.562	0.564	0.502	0.558	-3	-2	-13	-3
3	1999	0.597	0.573	0.577	0.457	0.563	-4	-3	-23	-6
3	2000	0.529	0.506	0.509	0.329	0.486	-4	-4	-38	-8
3	2001	0.410	0.378	0.388	0.188	0.357	-8	-5	-54	-13
3	2002	0.297	0.281	0.295	0.124	0.261	-5	-1	-58	-12
3	2003	0.289	0.275	0.278	0.227	0.275	-5	-4	-22	-5
3	2004	0.290	0.287	0.284	0.226	0.277	-1	-2	-22	-4
3	2005	0.329	0.336	0.335	0.291	0.338	2	2	-11	3
4	1995	0.973	0.913	0.913	0.909	0.913	-6	-6	-7	-6
4	1996	0.913	0.848	0.849	0.836	0.848	-7	-7	-8	-7
4	1997	0.980	0.910	0.911	0.877	0.908	-7	-7	-11	-7
4	1998	0.765	0.726	0.727	0.662	0.721	-5	-5	-13	-6
4	1999	0.790	0.725	0.729	0.595	0.714	-8	-8	-25	-10
4	2000	0.701	0.637	0.643	0.431	0.617	-9	-8	-39	-12
4	2001	0.542	0.488	0.499	0.262	0.465	-10	-8	-52	-14
4	2002	0.390	0.337	0.351	0.138	0.310	-14	-10	-65	-21
4	2003	0.381	0.340	0.361	0.125	0.308	-11	-5	-67	-19
4	2004	0.383	0.349	0.355	0.271	0.351	-9	-7	-29	-8

4	2005	0.435	0.410	0.408	0.303	0.392	-6	-6	-30	-10
5	1995	1.110	1.051	1.051	1.046	1.051	-5	-5	-6	-5
5	1996	1.041	0.974	0.975	0.960	0.973	-6	-6	-8	-7
5	1997	1.115	1.041	1.042	1.004	1.038	-7	-7	-10	-7
5	1998	0.870	0.820	0.822	0.750	0.814	-6	-5	-14	-6
5	1999	0.894	0.849	0.855	0.710	0.838	-5	-4	-21	-6
5	2000	0.793	0.730	0.737	0.502	0.706	-8	-7	-37	-11
5	2001	0.611	0.548	0.564	0.292	0.522	-10	-8	-52	-15
5	2002	0.438	0.390	0.407	0.166	0.361	-11	-7	-62	-18
5	2003	0.427	0.374	0.396	0.122	0.333	-12	-7	-71	-22
5	2004	0.429	0.375	0.405	0.108	0.328	-13	-6	-75	-23
5	2005	0.488	0.449	0.453	0.312	0.442	-8	-7	-36	-9
6	1995	1.177	1.133	1.133	1.128	1.133	-4	-4	-4	-4
6	1996	1.106	1.028	1.028	1.017	1.027	-7	-7	-8	-7
6	1997	1.185	1.107	1.107	1.069	1.102	-7	-7	-10	-7
6	1998	0.923	0.848	0.851	0.777	0.844	-8	-8	-16	-8
6	1999	0.947	0.867	0.875	0.727	0.856	-8	-8	-23	-10
6	2000	0.838	0.778	0.788	0.535	0.742	-7	-6	-36	-11
6	2001	0.644	0.573	0.591	0.308	0.542	-11	-8	-52	-16
6	2002	0.461	0.400	0.415	0.169	0.370	-13	-10	-63	-20
6	2003	0.447	0.385	0.403	0.126	0.341	-14	-10	-72	-24
6	2004	0.449	0.374	0.413	0.093	0.319	-17	-8	-79	-29
6	2005	0.511	0.424	0.474	0.090	0.350	-17	-7	-82	-32
7	1995	1.209	1.133	1.133	1.128	1.133	-6	-6	-7	-6
7	1996	1.136	1.028	1.028	1.017	1.027	-10	-10	-10	-10
7	1997	1.218	1.107	1.107	1.069	1.102	-9	-9	-12	-10
7	1998	0.948	0.848	0.851	0.777	0.844	-11	-10	-18	-11
7	1999	0.971	0.867	0.875	0.727	0.856	-11	-10	-25	-12
7	2000	0.859	0.778	0.788	0.535	0.742	-9	-8	-38	-14
7	2001	0.660	0.573	0.591	0.308	0.542	-13	-10	-53	-18
7	2002	0.471	0.400	0.415	0.169	0.370	-15	-12	-64	-21
7	2003	0.456	0.385	0.403	0.126	0.341	-16	-12	-72	-25
7	2004	0.458	0.374	0.413	0.093	0.319	-18	-10	-80	-30
7	2005	0.520	0.490	0.503	0.369	0.474	-6	-3	-29	-9
8	1995	1.227	1.133	1.133	1.128	1.133	-8	-8	-8	-8
8	1996	1.150	1.028	1.028	1.017	1.027	-11	-11	-12	-11
8	1997	1.234	1.107	1.107	1.069	1.102	-10	-10	-13	-11
8	1998	0.961	0.848	0.851	0.777	0.844	-12	-11	-19	-12
8	1999	0.984	0.867	0.875	0.727	0.856	-12	-11	-26	-13
8	2000	0.870	0.778	0.788	0.535	0.742	-11	-9	-38	-15
8	2001	0.667	0.573	0.591	0.308	0.542	-14	-11	-54	-19
8	2002	0.476	0.400	0.415	0.169	0.370	-16	-13	-64	-22
8	2003	0.461	0.385	0.403	0.126	0.341	-17	-12	-73	-26
8	2004	0.462	0.374	0.413	0.093	0.319	-19	-11	-80	-31
8	2005	0.525	0.490	0.503	0.369	0.474	-7	-4	-30	-10

Table 5. Comparison of true population numbers at age (thousands) with medians from 100 realizations under the four cases of treating index values less than 1.0.

		F at Age					Percent bias in Medians vs Truth			
Age	Year	True	Case 1	Case 2	Case 3	Case 4	Case 1	Case 2	Case 3	Case 4
1	1995	35236	35204	35180	35943	35249	0	0	2	0
1	1996	25724	26354	26310	27872	26458	2	2	8	3
1	1997	26449	26916	26811	30592	27112	2	1	16	3
1	1998	28054	28554	28367	36913	29142	2	1	32	4
1	1999	26207	27566	27102	44094	28562	5	3	68	9
1	2000	31907	32711	31866	61032	34452	3	0	91	8
1	2001	26383	27216	27231	31739	27606	3	3	20	5
1	2002	30976	31174	31460	37429	31911	1	2	21	3
1	2003	22272	21688	22217	24795	21886	-3	0	11	-2
1	2004	31379	31477	30866	38880	33836	0	-2	24	8
1	2005	13176	14426	20116	8650	13929	9	53	-34	6
1	2006	13176	13461	27312	3659	11610	2	107	-72	-12
2	1995	24025	24166	24162	24367	24182	1	1	1	1
2	1996	26523	26475	26454	27088	26520	0	0	2	0
2	1997	19463	19941	19918	21154	20006	2	2	9	3
2	1998	19900	20228	20157	23206	20440	2	1	17	3
2	1999	21504	21908	21761	28759	22399	2	1	34	4
2	2000	20057	21123	20803	34764	22028	5	4	73	10
2	2001	24611	25247	24559	48407	26662	3	0	97	8
2	2002	20634	21314	21398	24997	21670	3	4	21	5
2	2003	24546	24731	24986	29795	25386	1	2	21	3
2	2004	17667	17198	17604	19741	17318	-3	0	12	-2
2	2005	24890	24958	24448	31007	26906	0	-2	25	8
2	2006	10407	11427	16081	6674	11061	10	55	-36	6
3	1995	12645	12600	12599	12651	12604	0	0	0	0
3	1996	13490	13521	13517	13659	13529	0	0	1	0
3	1997	15224	14967	14959	15492	15024	-2	-2	2	-1
3	1998	10904	11082	11050	12111	11133	2	1	11	2
3	1999	12087	12282	12214	14715	12447	2	1	22	3
3	2000	12966	13234	13057	18672	13585	2	1	44	5
3	2001	12517	13289	12992	24479	13977	6	4	96	12
3	2002	16330	16799	16227	35675	17965	3	-1	118	10
3	2003	14523	15014	15026	17976	15280	3	3	24	5
3	2004	17356	17479	17544	21764	18001	1	1	25	4
3	2005	12489	12140	12403	14081	12114	-3	-1	13	-3
3	2006	17254	17285	16870	22139	18867	0	-2	28	9
4	1995	4938	4959	4959	4970	4960	0	0	1	0
4	1996	5000	5099	5100	5141	5102	2	2	3	2
4	1997	5559	5670	5668	5796	5677	2	2	4	2
4	1998	5974	5904	5899	6279	5933	-1	-1	5	-1
4	1999	5015	5191	5169	5967	5241	4	3	19	5
4	2000	5454	5672	5621	7670	5803	4	3	41	6
4	2001	6259	6555	6417	11078	6850	5	3	77	9
4	2002	6802	7408	7157	16538	7980	9	5	143	17
4	2003	9941	10365	9909	25846	11260	4	0	160	13
4	2004	8905	9327	9348	11763	9506	5	5	32	7

4	2005	10635	10737	10873	14259	11289	1	2	34	6
4	2006	7361	7048	7273	8649	7071	-4	-1	17	-4
5	1995	1354	1488	1488	1490	1488	10	10	10	10
5	1996	1529	1639	1639	1649	1640	7	7	8	7
5	1997	1645	1772	1771	1809	1775	8	8	10	8
5	1998	1710	1867	1865	1977	1875	9	9	16	10
5	1999	2277	2343	2335	2659	2368	3	3	17	4
5	2000	1865	2047	2037	2688	2098	10	9	44	12
5	2001	2216	2451	2418	4064	2576	11	9	83	16
5	2002	2982	3290	3203	6966	3519	10	7	134	18
5	2003	3770	4345	4133	11817	4791	15	10	213	27
5	2004	5560	6003	5660	18664	6770	8	2	236	22
5	2005	4971	5379	5373	7295	5482	8	8	47	10
5	2006	5636	5852	5934	8603	6258	4	5	53	11
6	1995	355	394	394	395	394	11	11	11	11
6	1996	366	427	427	429	427	17	17	17	17
6	1997	442	505	505	515	506	14	14	16	14
6	1998	442	510	509	542	512	15	15	23	16
6	1999	587	679	676	760	684	16	15	30	16
6	2000	762	815	808	1076	833	7	6	41	9
6	2001	691	812	798	1342	852	17	16	94	23
6	2002	985	1157	1136	2464	1258	18	15	150	28
6	2003	1575	1825	1750	4824	2010	16	11	206	28
6	2004	2014	2445	2271	8516	2823	21	13	323	40
6	2005	2964	3397	3073	13723	3991	15	4	363	35
6	2006	2498	2812	2784	4372	2903	13	11	75	16
7	1995	66	83	83	83	83	27	27	27	27
7	1996	90	106	106	106	106	18	18	19	18
7	1997	99	122	122	125	122	23	23	26	24
7	1998	111	135	135	145	136	22	22	31	23
7	1999	144	180	179	201	181	25	25	40	26
7	2000	186	232	230	305	237	24	23	64	27
7	2001	270	307	303	514	322	14	12	90	19
7	2002	297	369	358	804	405	24	20	171	36
7	2003	509	633	609	1692	703	24	20	233	38
7	2004	825	1010	963	3467	1175	23	17	320	42
7	2005	1053	1376	1234	6359	1679	31	17	504	60
7	2006	1456	1809	1562	10263	2309	24	7	605	59
8	1995	9	13	13	13	13	48	48	48	48
8	1996	18	24	24	24	24	34	33	34	34
8	1997	28	36	36	36	36	27	27	29	27
8	1998	31	41	41	44	41	34	34	42	35
8	1999	45	60	59	68	60	33	32	52	34
8	2000	58	73	73	95	75	26	25	64	29
8	2001	85	110	108	182	113	30	28	115	34
8	2002	150	184	180	388	198	23	20	159	32
8	2003	228	287	277	803	321	26	22	252	41
8	2004	382	487	456	1705	559	28	19	347	46
8	2005	624	690	666	845	710	11	7	35	14
8	2006	814	1057	937	14806	997	30	15	1718	22

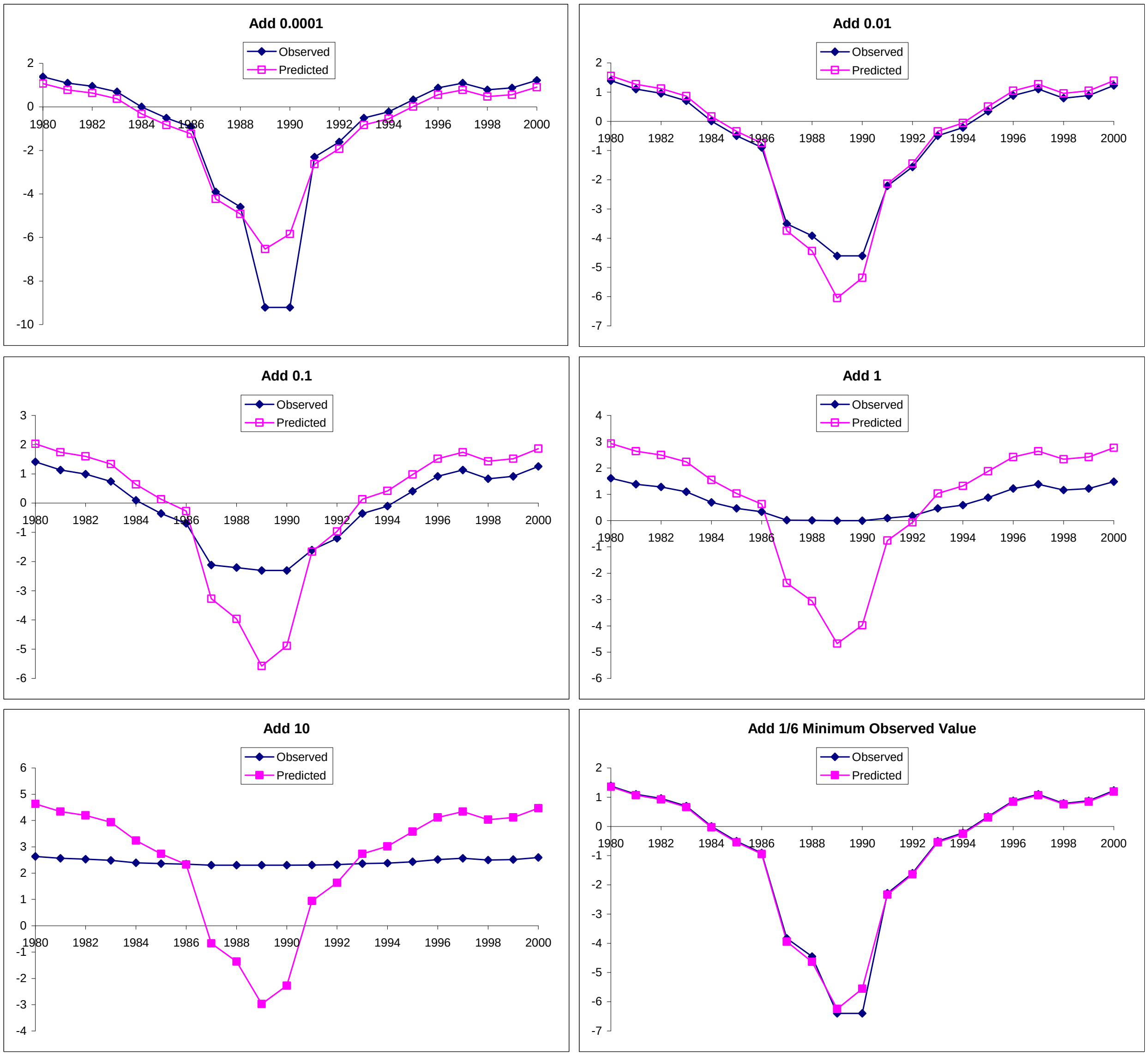


Figure 1. Comparison of observed and predicted indices when observed values are rounded to two decimal places and resulting zeros are replaced by different constants. The predicted indices follow the true population pattern and are scaled by a catchability coefficient to minimize the natural logarithm of the squared residuals. Note the y-axes are log scale.

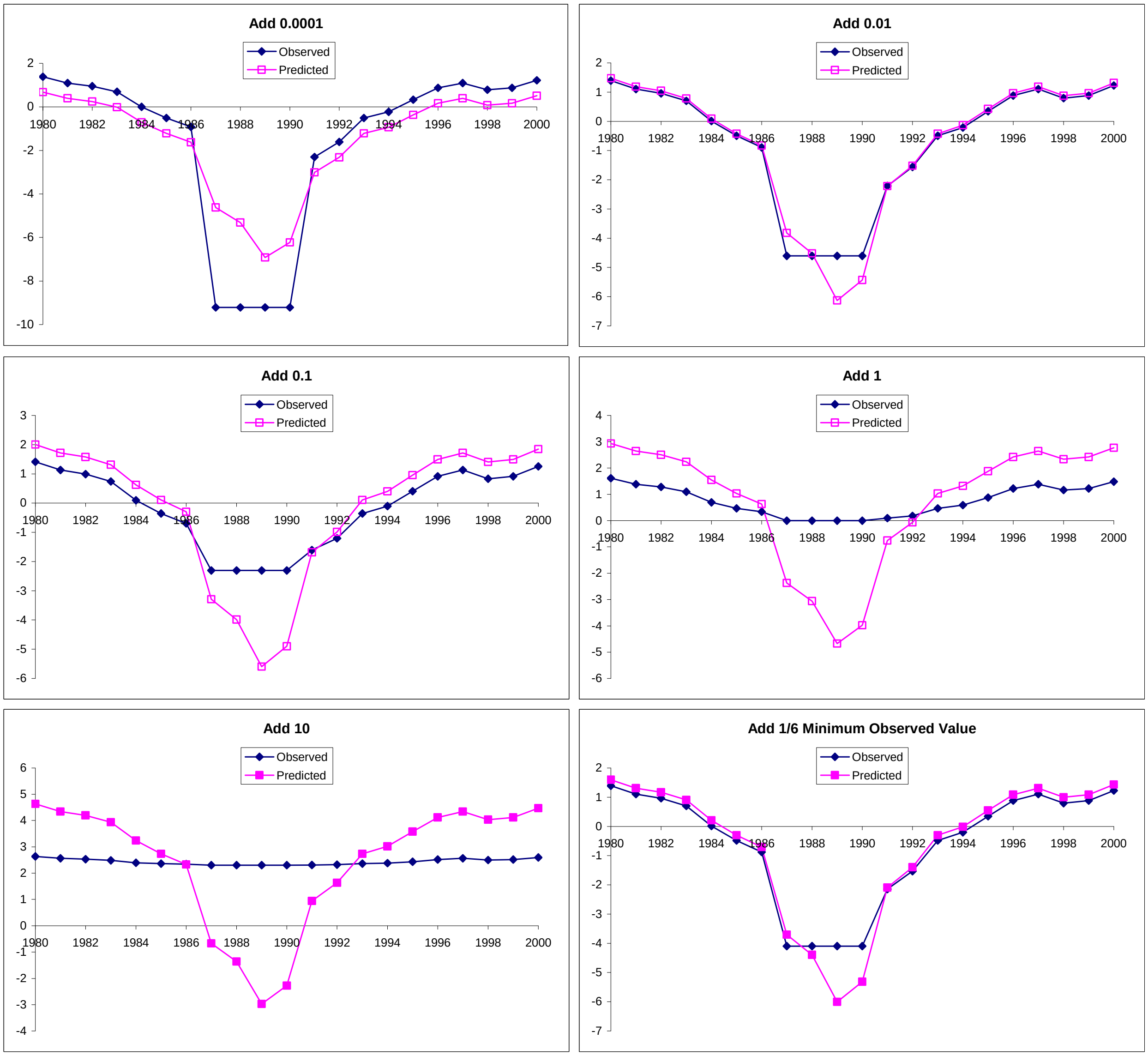


Figure 2. Comparison of observed and predicted indices when observed values are rounded to one decimal place and resulting zeros are replaced by different constants. The predicted indices follow the true population pattern and are scaled by a catchability coefficient to minimize the natural logarithm of the squared residuals. Note the y-axes are log scale.

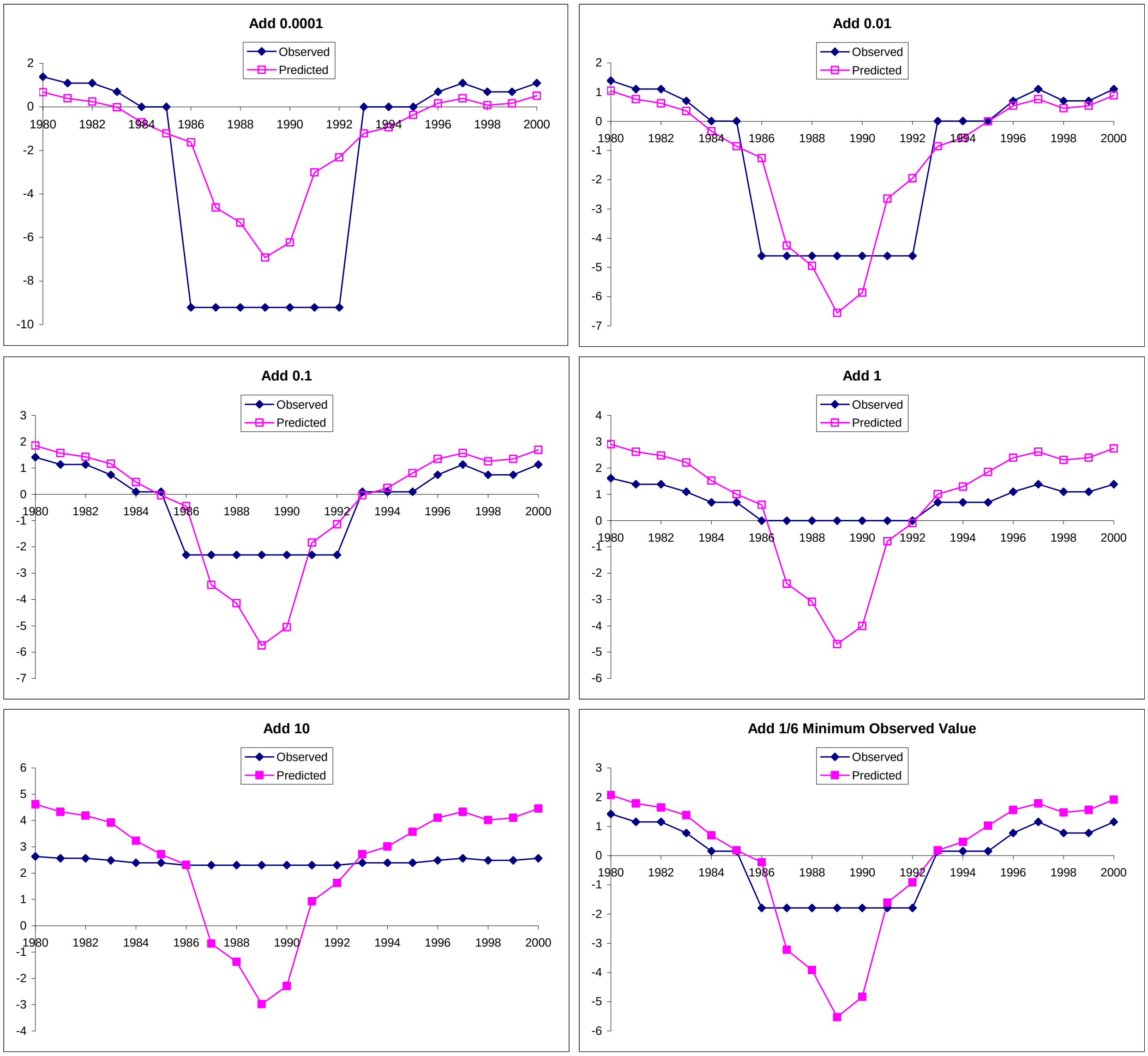


Figure 3. Comparison of observed and predicted indices when observed values are rounded to zero decimal places and resulting zeros are replaced by different constants. The predicted indices follow the true population pattern and are scaled by a catchability coefficient to minimize the natural logarithm of the squared residuals. Note the y-axes are log scale

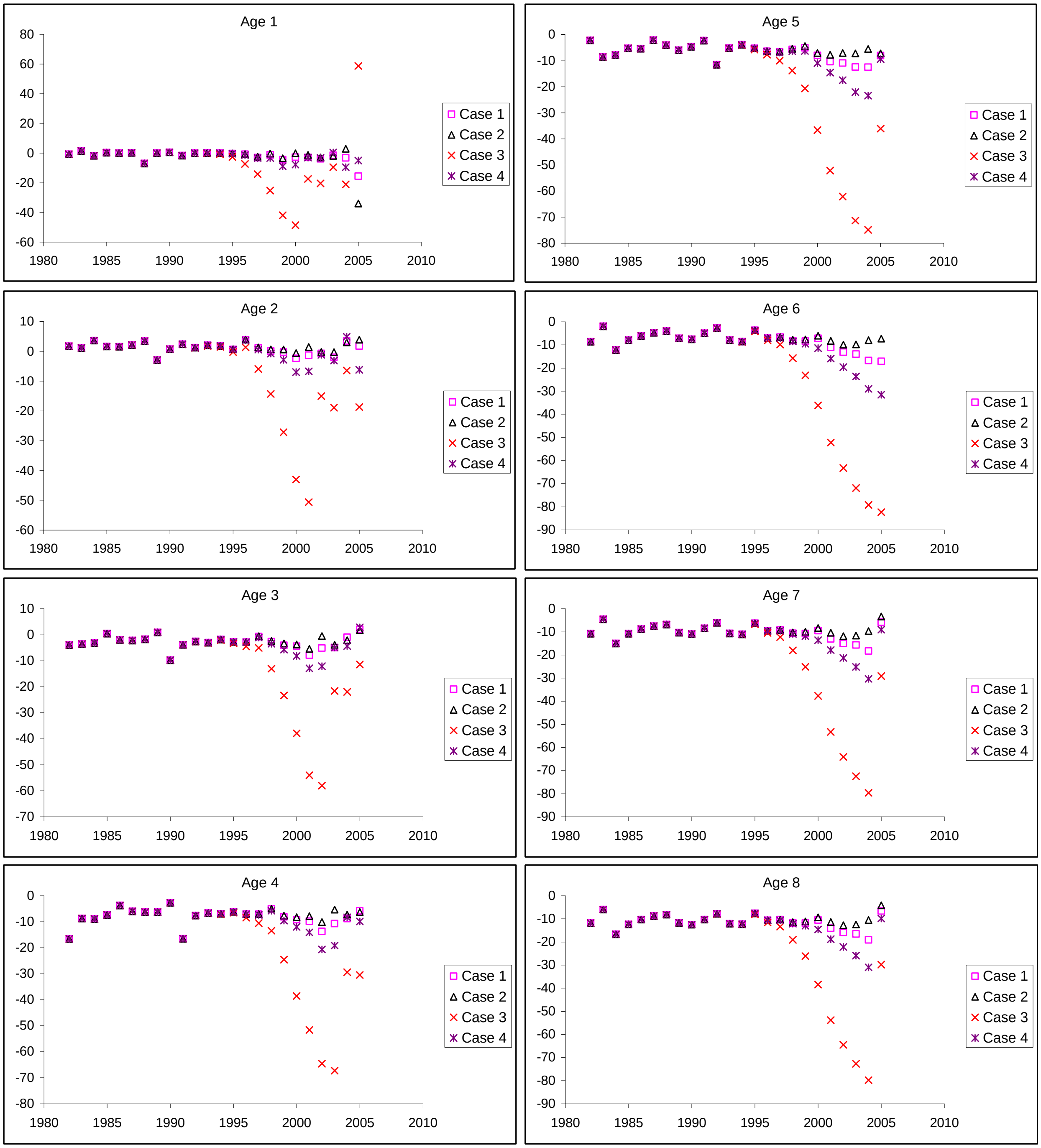


Figure 4. Percent bias in the medians of fishing mortality by age and year for the four cases of how index values less than one are treated

Figure 5. Percent bias in the medians of population numbers by age and year for the four cases of how index values less than one are treated. Note a number of case 3 values are too large to be shown on the plots, values are given in Table 5.

SAW 47 Working Paper 5 (TOR 2b) – Zero-fill Regression Example

15 November 2006

Some More Thoughts on Filling Zeros in Tuning Indices: A Simple Regression Example

by
Chris Legault

Introduction

The problem of zeros in tuning indices is that a lognormal error distribution is assumed. Since the logarithm of zero is undefined these zero tuning indices must be either treated as missing data or else be replaced by a positive value. One objective method to do this is to add 1/6 of the smallest non-zero value in the series to all values. The consequences of these two approaches are considered in a simple regression example.

Methods

A 26 year population time series was simulated with each value varying uniformly between zero and 50,000 fish. A catchability coefficient of 0.0001 was applied to generate the predicted index value. Lognormal noise with exp(std dev) of 0.2 was applied to the predicted values to generate the observed indices. If an observed index was below 0.5, then it was set to zero to mimic the problem of low abundance not being detected. The constant c was determined for each realization as 1/6 of the smallest non-zero value in the observed time series. Four time series of values were created, ln(obs), ln(obs+c), ln(pred), and ln(pred+c) where ln(obs) was missing when the observed value was zero. Two slopes were computed, one for ln(obs) vs ln(pred) denoted "missing" and the other for ln(obs+c) vs ln(pred+c) deonted "add c." Since in both cases the only source of error is the lognormal error assumed around the observed values, the expectation is that both lines will have slope equal to one. Random series of populations and observation errors were drawn 10,000 times and the two slopes computed for each realization.

Results

When zero observations were treated as missing, the slope was slightly negatively biased with mean 0.983 and 90% confidence interval (0.864, 1.109). When a constant of 1/6 the smallest non-zero value was added to all observed and predicted values, the slope was highly positively biased with mean 1.261 and 90% confidence interval (1.018, 1.483). Note that the 90% confidence interval for the "add c" case does not overlap one and has a range nearly twice as large as the "missing" case.

Discussion

The reason for this large disparity between the “missing” and “add c” results can be seen by examining an extreme example of the data used in the regressions (Figure 1). There were five observations that were replaced with c=0.094 causing the five ln(obs+c) values to all be -2.364 even though the associated ln(pred+c) values ranged from -1.265 to -0.307. These points do not fall on the line that would have been fit to the remaining data and are the source of the bias in the results. More typical results followed the same pattern but with less difference between the two slopes.

The constant was added to both observed and predicted data because to ensure an appropriate comparison. In a separate simulation I did not replace values less than 0.5 with zero and found nearly identical distributions for the “missing” and “add c” slopes. This demonstrates that filling of zeros causes the problem, not the addition of a constant.

In order for the “add c” approach to be unbiased, the constant would have to be selected for each realization such that the average of the ln(pred+c) was the same as ln(obs+c) for the values when obs=0. This cannot happen because the predicted values are positive while the observed values are by definition set to zero. Thus, adding a constant to all values when a zero is in the time series will always bias the results.

Conclusion

Filling observed zeros in tuning indices causes a bias relative to the true population that is much greater than the bias introduced by treating the zeros as missing in this simple regression example.

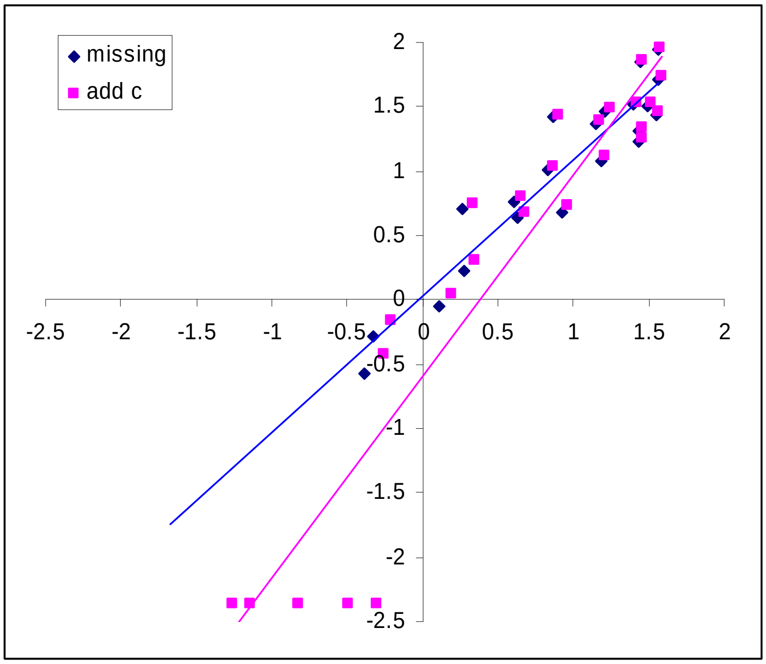


Figure 1. One realization of the "missing" and "add c" regressions. This example is extreme with "add c" slope slightly larger than the upper 90% confidence interval. The x-axis is either ln(pred) or ln(pred+c) and the y-axis is either ln(obs) or ln(obs+c).

SAW 47 Working Paper 6 (TOR 2b) – Treatment of Zeros

October 30, 2006

The treatment of "zero" observations in the Summer Flounder ADAPT VPA calibration

by
Mark Terceiro

Introduction

The issue of how to treat "zero" observations in ADAPT VPA calibrations was addressed in a previous Southern Demersal Working Group (SDWG) working paper used in preparing the 2004 summer flounder assessment (SDWG 2003; beginning on page 8). That work responded to the 2002 SAW 35 (NEFSC 2002) summer flounder assessment Research Recommendation: *Explore the sensitivity of the VPA calibration to the addition of 1 and/or a small constant to values of survey series with "true zeros."* This recommendation stemmed from the nature of the ADAPT VPA calibration (tuning) algorithm, which includes natural logarithm (ln) transformation (i.e., assumption of a lognormal error distribution) of the input survey abundance indices prior to calibration. Some of the tuning series in the assessment include several "true zero" observations (as contrasted with years for which no sampling was performed) in their time series. Since "zeros" are treated as missing values in the ADAPT computations, a possible solution would be to add a constant to every value in these series to enable use of these "true zeros" as observations.

In the 2002 (NEFSC 2002) and 2003 (Terceiro 2003a) summer flounder assessments, the addition of the constant value of 1 was made for five age 0 recruitment indices: the MA DMF Seine, CT DEP fall trawl, RI DFW fall trawl, RI DFW monthly trawl, and DE DFW 16 foot bay trawl survey series (note that the latter series was not included in the final ADAPT VPA tuning configuration). No constant was added to survey series with "zero" observations for other age classes. The choice of the value of 1 as the additive constant was based on recommendations from statistical texts (e.g., Snedecor and Cochran 1967, Sokal and Rohlf 1981) for the ln-transformation of data.

Berry (1987) provides guidance on the objective selection of the appropriate value of the additive constant based on the statistical properties (skewness and kurtosis) of data series to be ln-transformed. Briefly, the method consists of 1) addition of a range of constants from very large (e.g., 100) to very small (e.g., 0.0001) to the original values in the series, 2) ln-transformation of the modified series, 3) calculation of the skewness and kurtosis of the modified series, and 4) summation of the absolute value of the skewness and kurtosis (providing the statistic g) of the modified series. The additive constant that minimizes g for a given series of data is the one that best minimizes the effect of outliers and normalizes residuals from the lognormal error distribution, hence best adhering to the assumption of the lognormal distribution. Work using the procedures suggested by Berry (1987) with recreational fishery catch rates as indices of abundance indicated that the additive constant of 1 was an appropriate value for those data, typically with values between zero and 50 (Terceiro 2003b).

The SDWG (2003) work applied the method suggested by Berry (1987) to summer flounder age 0 surveys with “zero” observations. Of the five age 0 series with “zero” observations, the MA DMF series varies between 0 and 70, while the other four series contained small values that varied between 0 and 1. The 2003 work (SDWG 2003) found that for the MA DMF series, the additive constant of 1 minimized the value of *g*. For the other four series, *g* was minimized by small values of the additive constant ranging from 0.001 to 0.1, with an “average” best additive constant of 0.1. The SDWG (2003) therefore recommended use of the revised, varying (1 or 0.1) additive constants in future assessments, and this revision was made in the 2004-2006 assessment, for age 0 survey series only. No constant was added for survey series of other age classes, pending further research.

Recently, the 2006 assessment of summer flounder (Terceiro 2006a) was subject to a NMFS Office of Science and Technology (S&T) Peer Review (Methot 2006). Among the recommendations made by the S&T Peer Review panel was the following:

The Panel finds that one immediate modification of the VPA is justifiable and reduces the retrospective pattern in stock size during 2003-2005. The VPA model currently treats survey observations of zero as missing values. An observation of zero for a particular age of fish in a particular survey year does not mean that there are no fish of that age in the stock, only that the number of survey samples was not sufficient to detect any fish of that age. This VPA model, as with most assessment models, tunes to the logarithm of the survey observations so cannot explicitly deal with observations of zero. However, treating these zeroes as missing values can result in a bias because time periods of low abundance are underrepresented in the data input to the assessment model. In the case of summer flounder, the result may be an underestimate of the degree to which the stock has rebuilt since the low levels that occurred around 1990. The committee did not discuss this issue during the Sept 14-15 meeting, so is not prepared to present a definitive solution. An interim approach would use a small value in place of the zeroes. A value equal to one sixth of the smallest observed positive value would be reasonable until a more complete statistical solution can be developed.

As a result, a revised 2006 ADAPT VPA for summer flounder was developed for which the previous treatment of “zero” observations for age 0 indices was retained (additive constant of 1 for MA DMF seine survey, 0.1 for the CT DEP fall trawl, RI DFW fall trawl, RI DFW monthly trawl, and DE DFW 16 foot bay trawl surveys). For ages 1-7+ survey series with “zero” observations, a value equal to one-sixth of the minimum value in each series was used in place of the “zero” observations. Typically, the minimum non-zero value in these series was 0.01, and so the additive constant was 0.001667 (Terceiro 2006b).

Summer flounder 2006 ADAPT VPA

In this work, the Berry (1987) approach is applied to the summer flounder survey series for all ages with observed “zeros” to determine the best additive constant to use to remove these “zero” observations from the ADAPT VPA calibration data. Table 1 summarizes the statistical properties of the 24 survey series that were examined. The distributions of the surveys are characterized by non-zero values between 0.001 and 70, CVs that generally exceed 100%, positive skewness (long right hand tail), and significant kurtosis (high degree of peak, or contagion, near the mean). The proportion of “zeros” in the time series ranged from 1 of 31 = 3% (NEFSC Spring Age 3 index) to 13 of 28 = 46% (MA Fall 4).

Table 2 summarizes the results of the exercise for each group of age-specific indices. Values of *g* were minimized for constants between 0.001 and 100 (minimum values in ***bold italics***), for the age 0, 1, 2, 3, 4, and 5-7+ (aggregate) survey indices (number per tow or haul). There is no statistically significant correlation (Table 3) between the value of the additive constant that minimizes g and the statistical parameters listed in Table 1.

Age 0 Indices

For the five age 0 series, the *g* statistic was minimized by values of the additive constant ranging from 0.001667 to 1. The constant equated to one-sixth of the minimum non-zero observed value for 2 of the 5 series . The relationships between the additive constants and calculated values of *g* for the age 0 indices are shown in Figure 1.

Age 1 Indices

For the three age 1 series, the additive constant of 0.01 minimized the absolute value of *g*. The relationships between the additive constants and calculated values of *g* for the age 1 indices are shown in Figure 2.

Age 2 Indices

For the single age 2 series, the additive constant of 0.1 minimized the absolute value of *g*. The relationships between the additive constants and calculated values of *g* for the age 2 indices are shown in Figure 3.

Age 3 Indices

For the six age 3 series, the absolute value of the *g* statistic was minimized by values of the additive constant ranging from 0.001 to 100. The constant equated to one-sixth of the minimum non-zero observed value for 1 of the 6 series . The relationships between the additive constants and calculated values of *g* for the age 3 indices are shown in Figure 4.

Age 4 Indices

For the six age 4 series, the absolute value of the *g* statistic was minimized by values of the additive constant ranging from 0.001 to 1. The constant equated to one-sixth of the minimum non-zero observed value for 1 of the 6 series . The relationships between the additive constants and calculated values of *g* for the age 4 indices are shown in Figure 5.

Age 4/5-7+ Indices

For the three age 4/5-7+ series, the absolute value of the *g* statistic was minimized by values of the additive constant ranging from 0.001667 to 100. The constant equated to one-sixth of the minimum non-zero observed value for 1 of the 3 series . The relationships between the additive constants and calculated values of *g* for the age 4/5-7+ indices are shown in Figure 6.

Conclusion

There is no consistent pattern in the identification of the additive constant that minimizes the absolute value of Berry's (1987) *g* statistic. There is no strong relationship between the absolute magnitude of the index values, the length of the time series, the number of zeros, the magnitude of the smallest observed value, or any of the usual statistical moments of the series (mean, maximum, non-zero minimum, CV, skewness, kurtosis), and the value of the additive constant that minimizes *g*. Further, while the "one-sixth" of the minimum observed value was identified as the "best" additive constant in 5 of the 24 (21%) cases examined, this level is not high enough to justify this approach as a reliable rule-of-thumb. In fact, the additive constant of 0.01 was identified as "best" for a higher percentage of series (6 of 24 = 25%). Given the inability to identify a constant that consistently minimizes *g*, the best rule is to maintain the current approach of making no adjustment and continue to treat "zero" observations as "missing."

References Cited

Berry DA 1987. Logarithmic transformations in ANOVA. Biometrics 43:439-456.

Methot R. 2006. Review of the 2006 Summer Flounder Assessment Update. Chair's Report. NMFS Office of Science and Technology. 6 p.

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

Snedecor GW, Cochran WG. 1967. Statistical methods (6th ed). Iowa State University Press. Ames IA. 593 p.

Sokal RR, Rohlf FJ. 1981. Biometry (2nd ed). W.H. Freeman and Company. New York, NY. 859 p.

Southern Demersal Working Group (SDWG). 2003. SAW Southern Demersal Working Group (WG) Responses to 2002 SAW 35 and 2003 Summer Flounder Assessment Research Recommendations (numbered as in the 2003 assessment): December 22, 2003. 16 p.

Terceiro M. 2003a. Stock assessment of summer flounder for 2003. NEFSC Ref Doc. 03-09. 179 p.

Terceiro M. 2003b. The statistical properties of recreational catch rate data for some fish stocks off the northeast U.S. coast. Fish Bull. 101(3): 653-672.

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

Terceiro M. 2006b. Summer flounder assessment and biological reference point update for 2006. 64 p.

Table 1 . Statistical properties of summer flounder ADAPT VPA survey calibration series with "zero" observations.

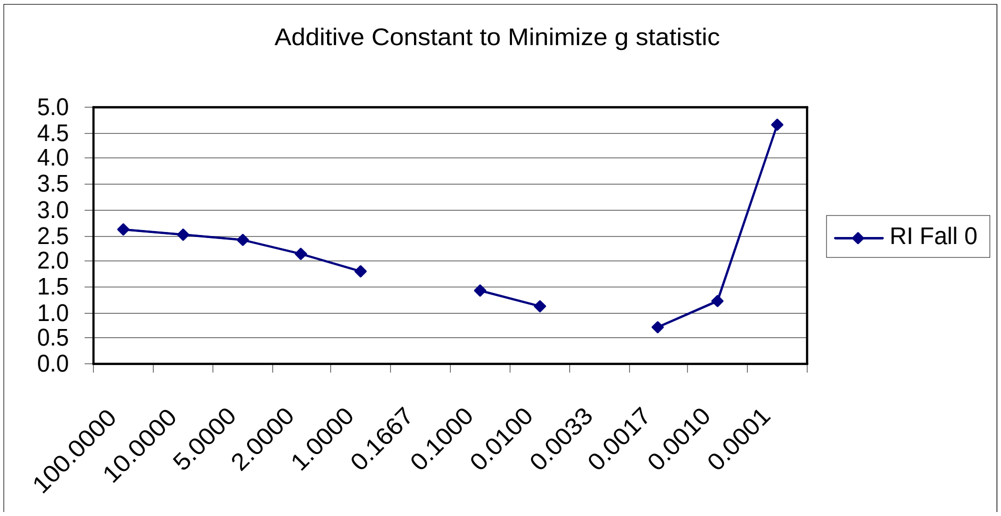
Survey Name	N Obs	N "Zeros"	Mean	Max	Non-zero Min	CV (%)	Skew	Kurt
Age 0 Indices								
RI Fall 0	26	2	0.130	0.550	0.01	118	1.422	1.212
RI Monthly 0	16	3	0.037	0.110	0.01	95	1.044	0.152
CT Fall 0	22	4	0.085	0.442	0.013	128	2.078	4.733
MA Seine 0	24	3	8.292	70.000	1	170	3.904	17.141
DE 30 0	15	1	0.487	2.280	0.02	133	1.874	3.463
Age 1 Indices								
MA Spring 1	28	2	0.260	1.770	0.025	140	3.105	11.161
MA Fall 1	28	2	0.761	2.907	0.011	109	1.280	0.686
RI Fall 1	26	1	0.535	2.470	0.05	120	2.197	4.724
Age 2 Indices								
MA Fall 2	28	2	0.759	2.235	0.047	85	0.884	-0.224
Age 3 Indices								
NEC Spring 3	31	1	0.302	1.020	0.01	99	0.865	-0.536
NEC Fall 3	24	2	0.168	0.660	0.01	111	1.076	0.334
MA Fall 3	28	2	0.132	0.756	0.010	132	2.086	5.191
RI Monthly 3	16	2	0.199	0.530	0.01	95	0.786	-0.916
NJ Trawl 3	18	3	0.340	1.280	0.01	112	1.141	0.828
DE 30 3	15	2	0.155	0.470	0.01	105	0.991	-0.137
Age 4 Indices								
NEC Spring 4	31	5	0.092	0.310	0.01	111	0.985	-0.404
NEC Fall 4	24	8	0.043	0.190	0.01	144	1.444	0.762
MA Spring 4	28	5	0.086	0.317	0.010	116	1.187	0.242
MA Fall 4	28	13	0.019	0.186	0.01	196	3.484	14.026
RI Fall 4	26	5	0.035	0.280	0.01	179	2.961	9.516
RI Monthly 4	16	4	0.060	0.240	0.01	122	1.257	0.856
Age 5-7+ Indices								
NEFSC Spring 5-7+	31	10	0.060	0.210	0.01	121	0.892	-0.793
NEFSC Winter 5-7+	15	1	0.803	2.600	0.01	106	0.698	-0.636
NJ Trawl 4-7+	18	4	0.172	0.810	0.01	129	1.715	2.946

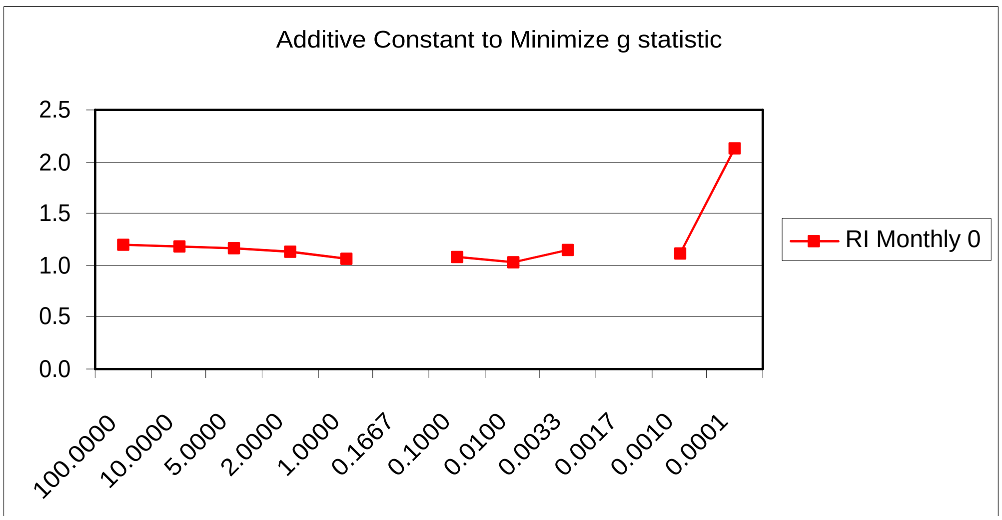
Table 2. Values of the additive constants that minimize the statistic *g*. Values that are one-sixth the minimum observed in the series are in bold.

Survey Name	**Constant**
Age 0 Indices	
RI Fall 0	**0.001667**
RI Monthly 0	0.01
CT Fall 0	0.01
MA Seine 0	1
DE 30 0	**0.003333**
Age 1 Indices	
MA Spring 1	0.01
MA Fall 1	0.01
RI Fall 1	0.01
Age 2 Indices	
MA Fall 2	0.1
Age 3 Indices	
NEC Spring 3	0.01
NEC Fall 3	**0.001667**
MA Fall 3	0.001
RI Monthly 3	0.01
NJ Trawl 3	2
DE 30 3	100
Age 4 Indices	
NEC Spring 4	0.001
NEC Fall 4	0.1
MA Spring 4	1
MA Fall 4	0.001
RI Fall 4	**0.001667**
RI Monthly 4	0.1
Age 5-7+ Indices	
NEFSC Spring 5-7+	100
NEFSC Winter 5-7+	**0.001667**
NJ Trawl 4-7+	0.1

Table 3. Correlation analysis (value of *r*) of various statistical properties of the 24 summer flounder index series with "zero" observations. For n = 24, the critical value of *r* at the 0.05 significance level is about 0.4.

	N	Nzero	Mean	Max	Min	CV	Skew	Kurt	*g*
N	1.00								
Nzero	0.32	1.00							
Mean	0.01	-0.12	1.00						
Max	0.02	-0.08	1.00	1.00					
Min	0.18	-0.28	-0.01	-0.05	1.00				
CV	0.20	0.57	0.31	0.35	-0.19	1.00			
Skew	0.21	0.26	0.50	0.53	0.08	0.88	1.00		
Kurt	0.20	0.26	0.57	0.60	0.05	0.86	0.99	1.00	
g	-0.02	0.25	-0.08	-0.07	-0.12	-0.13	-0.24	-0.22	1.00


Additive Constant to Minimize g statistic
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
100.0000
10.0000
5.0000
2.0000
1.0000
0.1667
0.1000
0.0100
0.0033
0.0017
0.0010
0.0001
RI Fall 0


Additive Constant to Minimize g statistic
2.5
2.0
1.5
1.0
0.5
0.0
100.0000
10.0000
5.0000
2.0000
1.0000
0.1667
0.1000
0.0100
0.0033
0.0017
0.0010
0.0001
RI Monthly 0

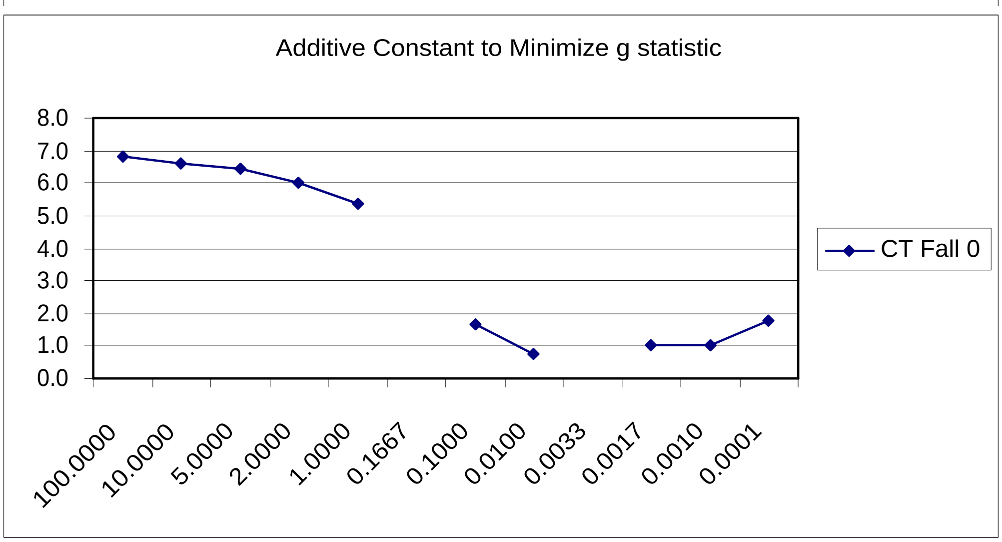
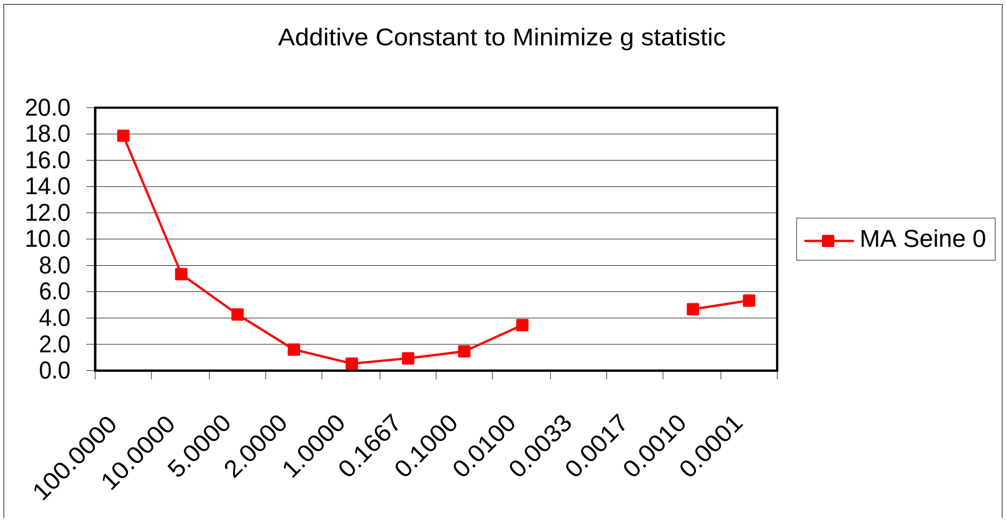

Additive Constant to Minimize g statistic
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
100.0000
10.0000
5.0000
2.0000
1.0000
0.1667
0.1000
0.0100
0.0033
0.0017
0.0010
0.0001
CT Fall 0

Figure 1.



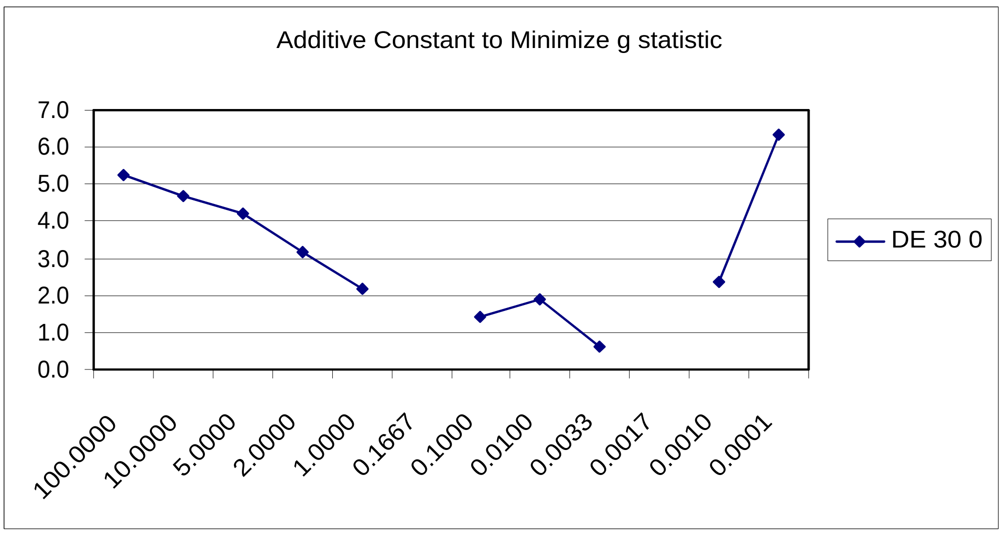
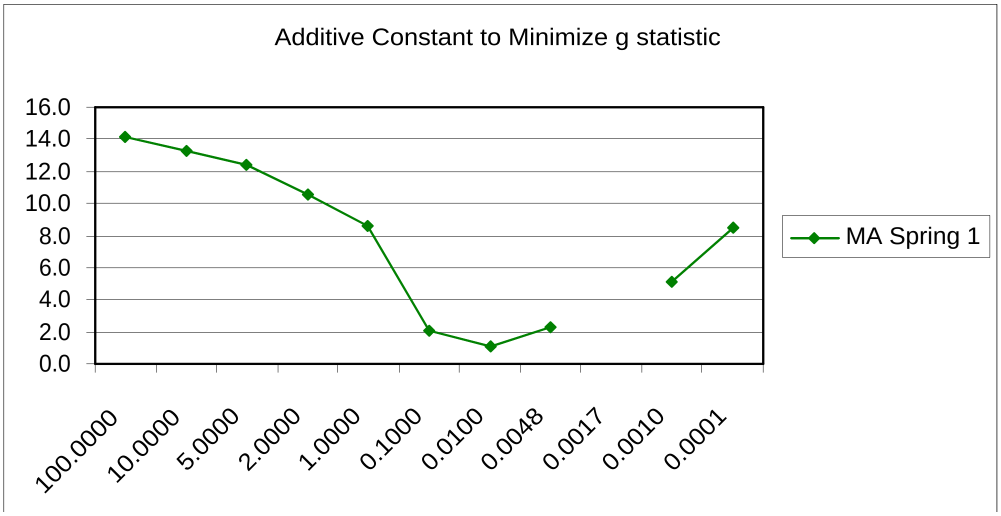
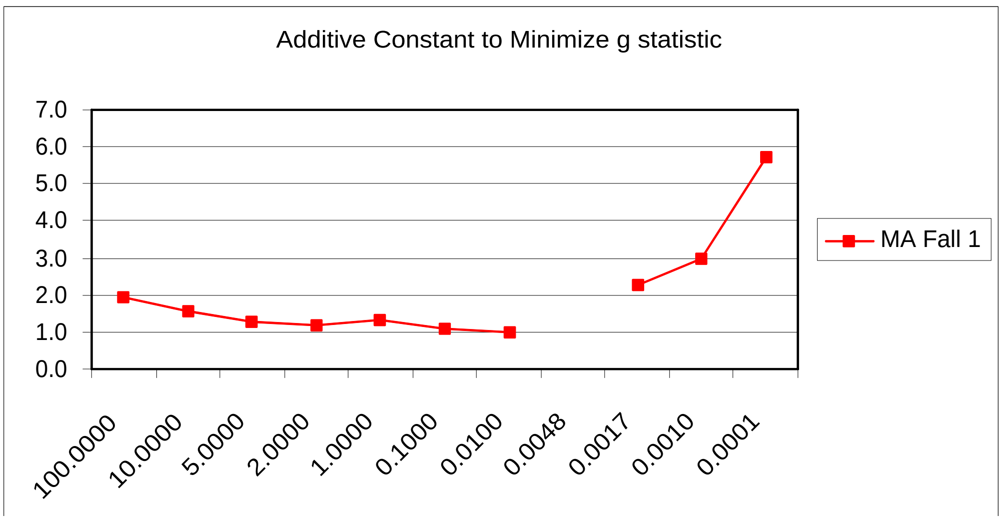


Figure 1 continued.


Additive Constant to Minimize g statistic
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
100.0000
10.0000
5.0000
2.0000
1.0000
0.1000
0.0100
0.0048
0.0017
0.0010
0.0001
MA Spring 1


Additive Constant to Minimize g statistic
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
100.0000
10.0000
5.0000
2.0000
1.0000
0.1000
0.0100
0.0048
0.0017
0.0010
0.0001
MA Fall 1

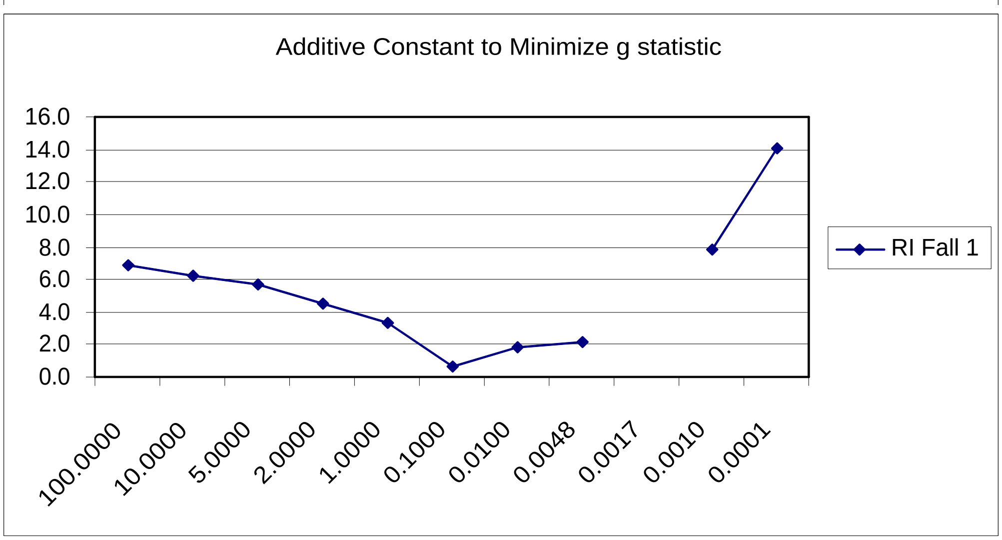

Additive Constant to Minimize g statistic
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
100.0000
10.0000
5.0000
2.0000
1.0000
0.1000
0.0100
0.0048
0.0017
0.0010
0.0001
RI Fall 1

Figure 2.

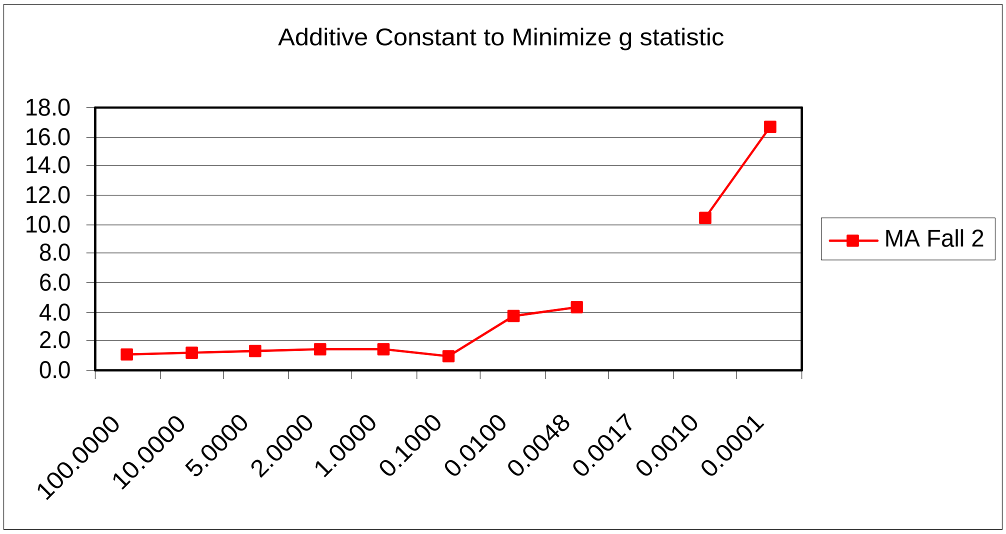
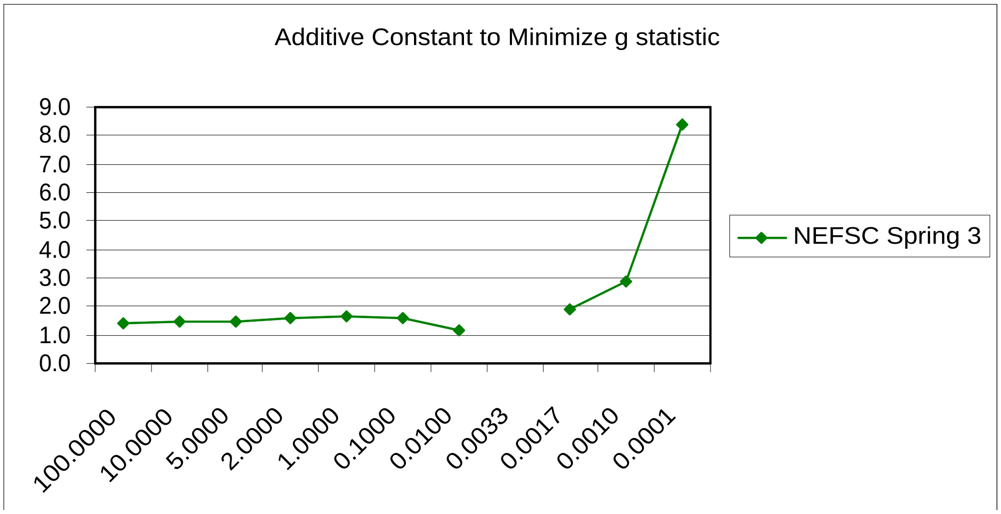
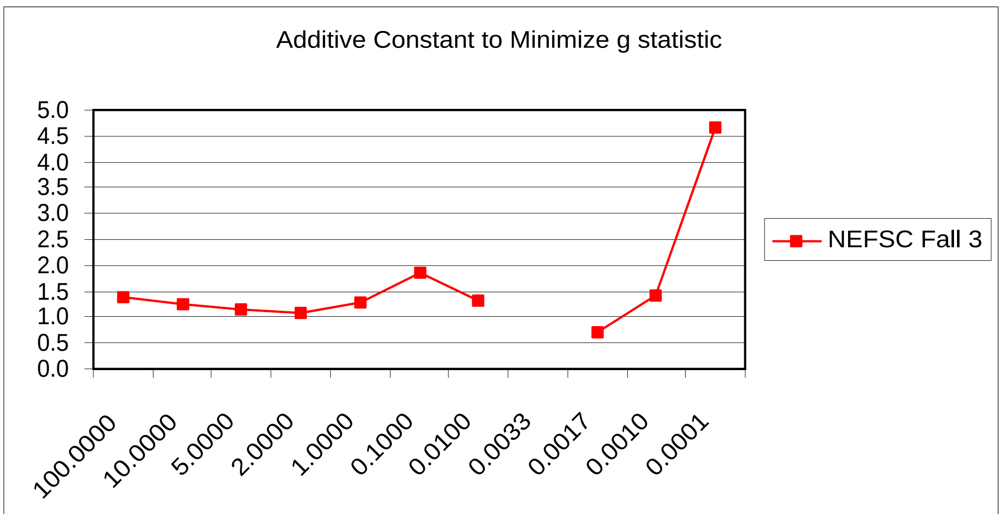

Additive Constant to Minimize g statistic
18.0
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
100.0000
10.0000
5.0000
2.0000
1.0000
0.1000
0.0100
0.0048
0.0017
0.0010
0.0001
MA Fall 2

Figure 3.


Additive Constant to Minimize g statistic
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
NEFSC Spring 3
100.0000
10.0000
5.0000
2.0000
1.0000
0.1000
0.0100
0.0033
0.0017
0.0010
0.0001


Additive Constant to Minimize g statistic
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
NEFSC Fall 3
100.0000
10.0000
5.0000
2.0000
1.0000
0.1000
0.0100
0.0033
0.0017
0.0010
0.0001

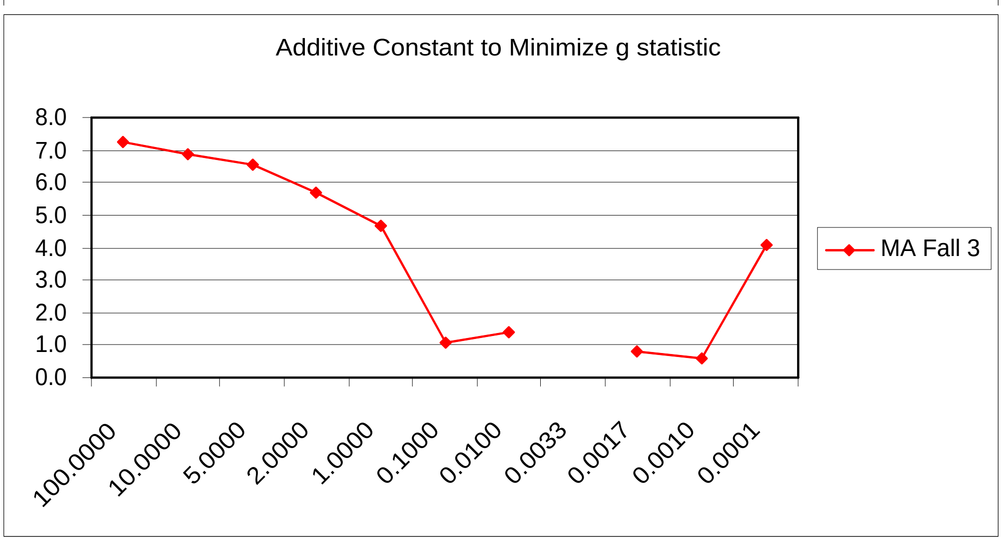
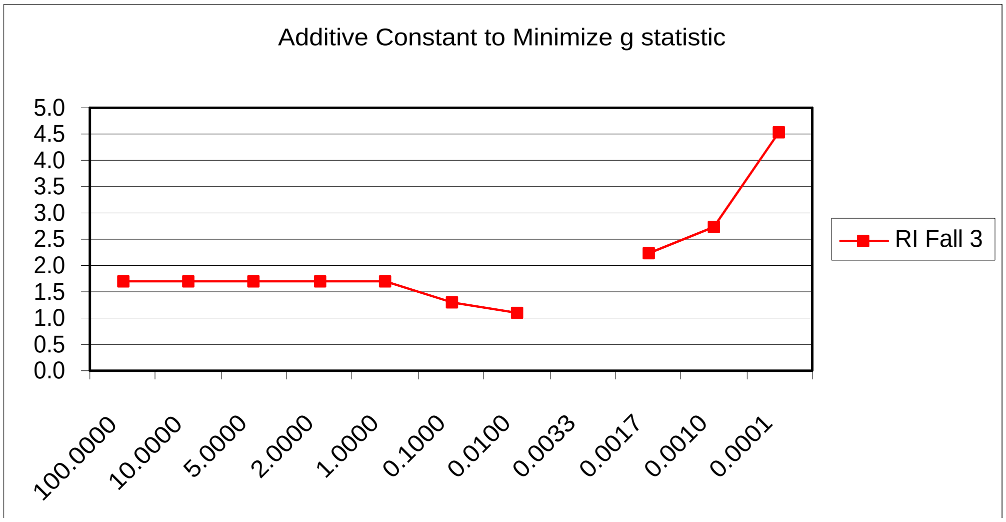
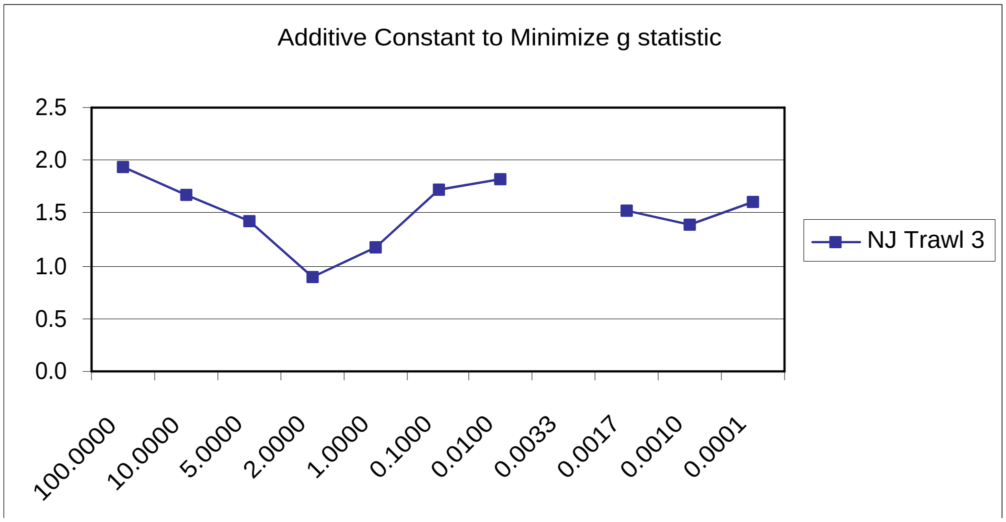

Additive Constant to Minimize g statistic
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
MA Fall 3
100.0000
10.0000
5.0000
2.0000
1.0000
0.1000
0.0100
0.0033
0.0017
0.0010
0.0001

Figure 4.





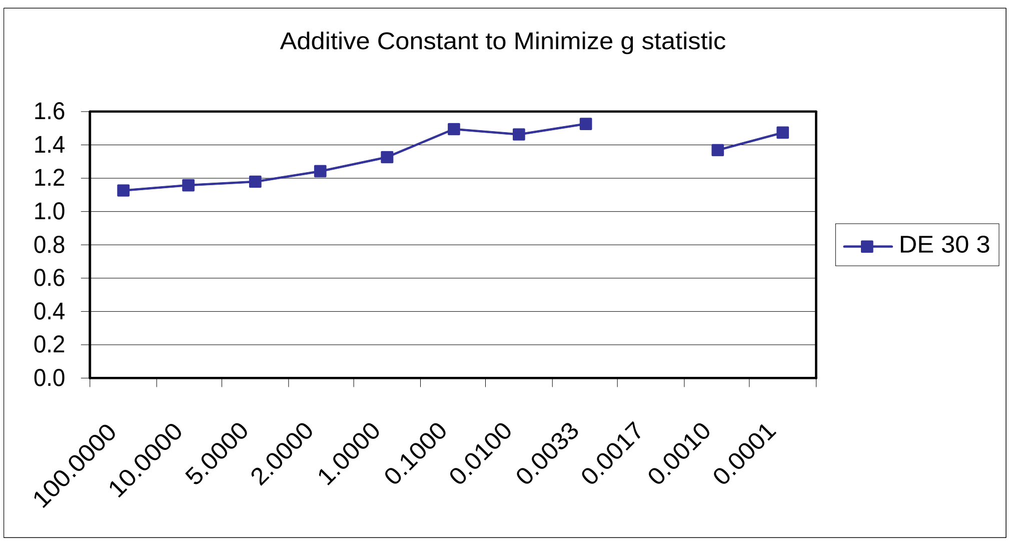
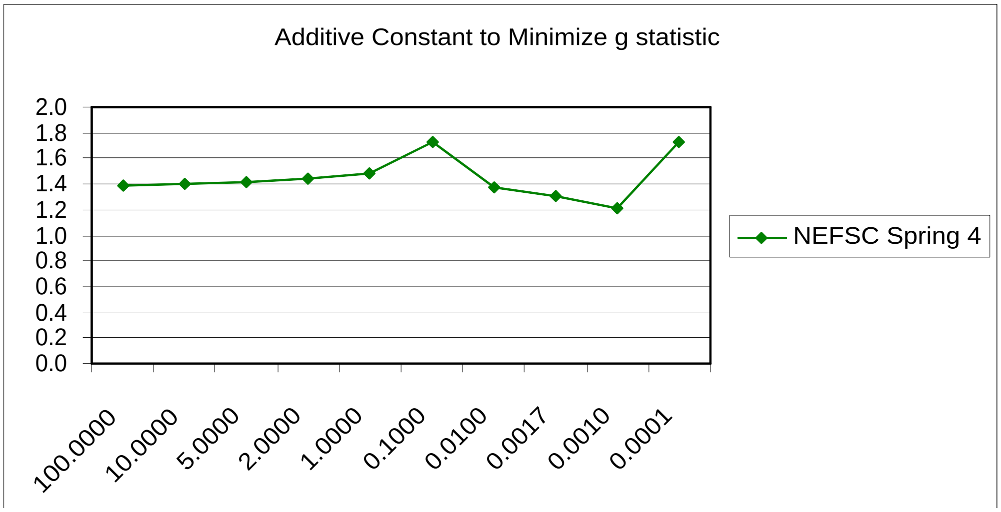
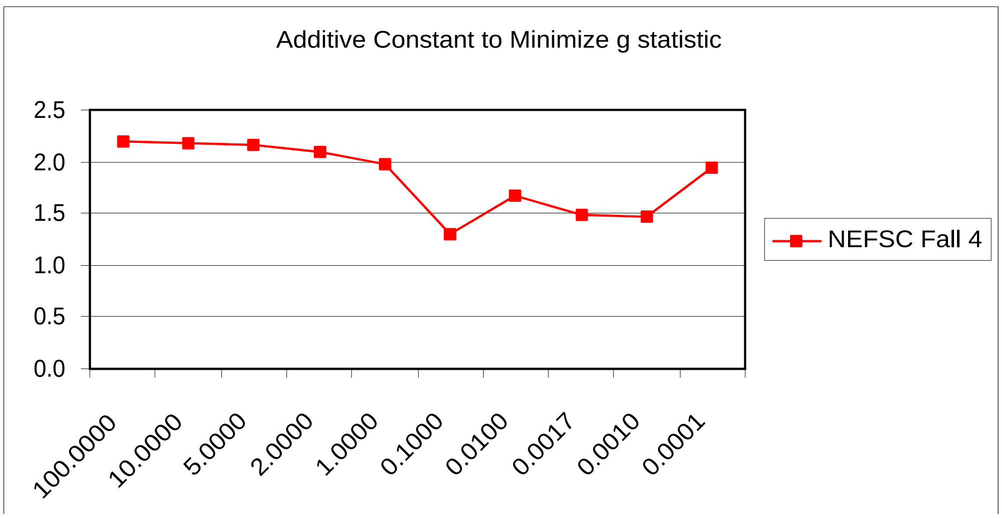


Figure 4 continued.


Additive Constant to Minimize g statistic
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
100.0000
10.0000
5.0000
2.0000
1.0000
0.1000
0.0100
0.0017
0.0010
0.0001
NEFSC Spring 4


Additive Constant to Minimize g statistic
2.5
2.0
1.5
1.0
0.5
0.0
100.0000
10.0000
5.0000
2.0000
1.0000
0.1000
0.0100
0.0017
0.0010
0.0001
NEFSC Fall 4

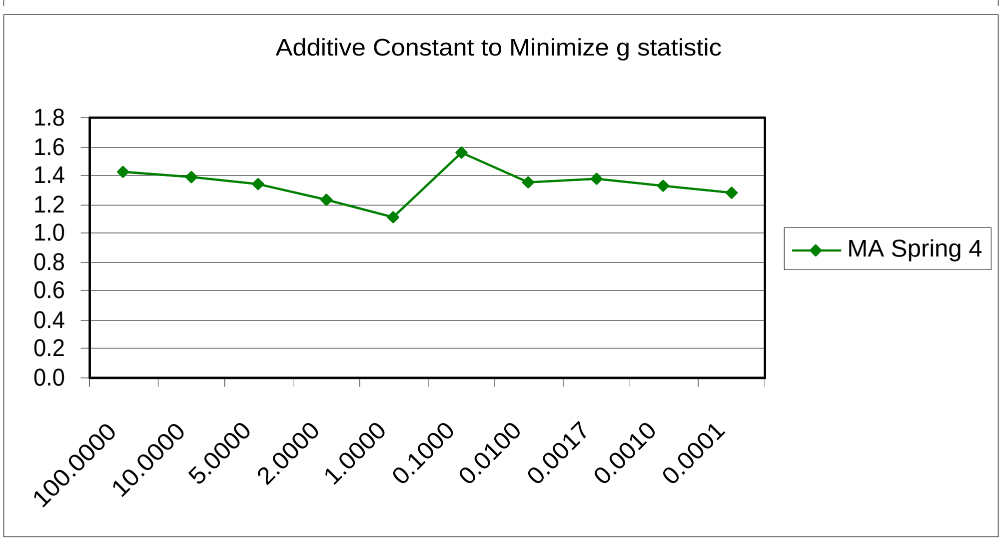
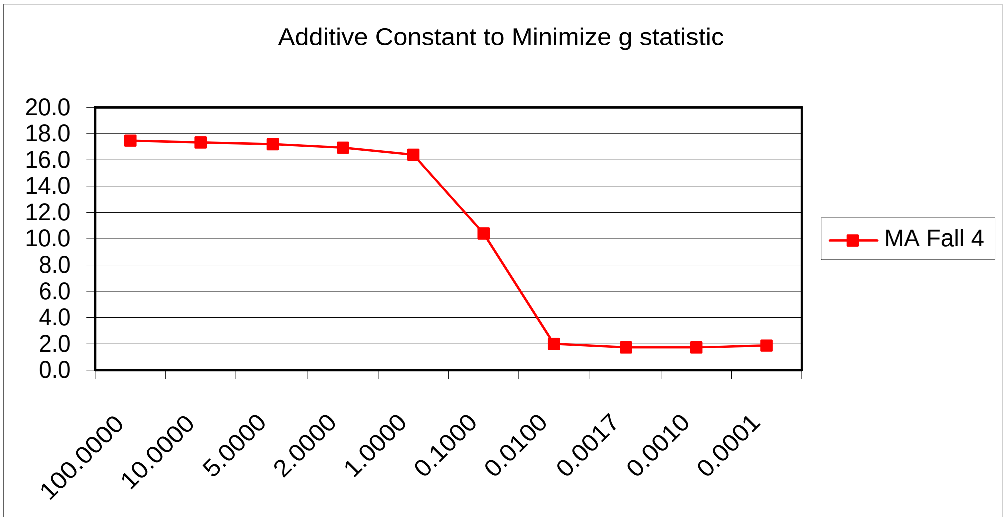
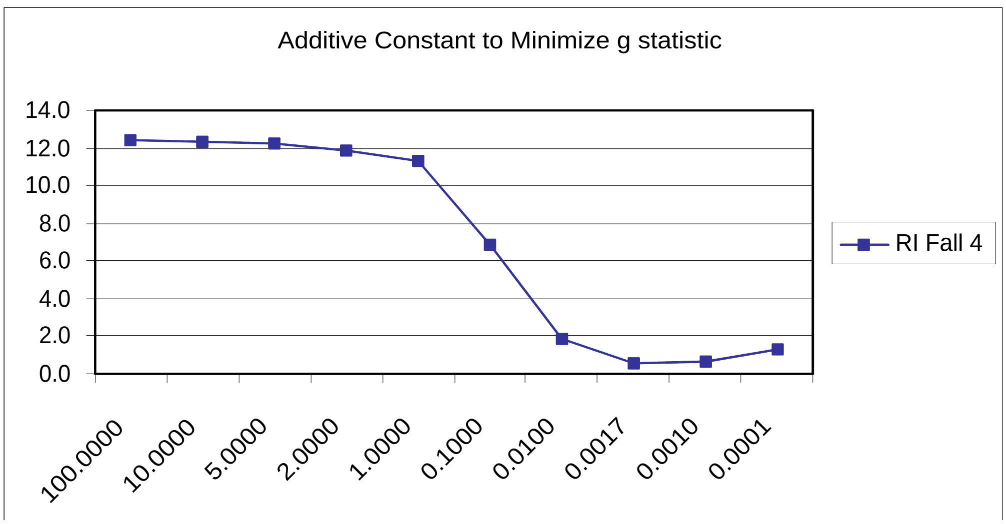

Additive Constant to Minimize g statistic
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
100.0000
10.0000
5.0000
2.0000
1.0000
0.1000
0.0100
0.0017
0.0010
0.0001
MA Spring 4

Figure 5.





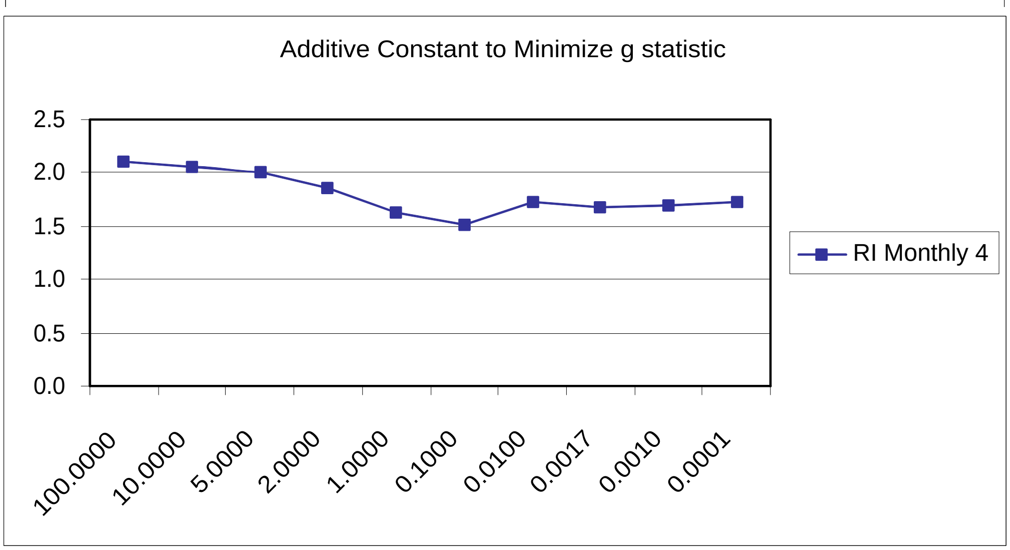
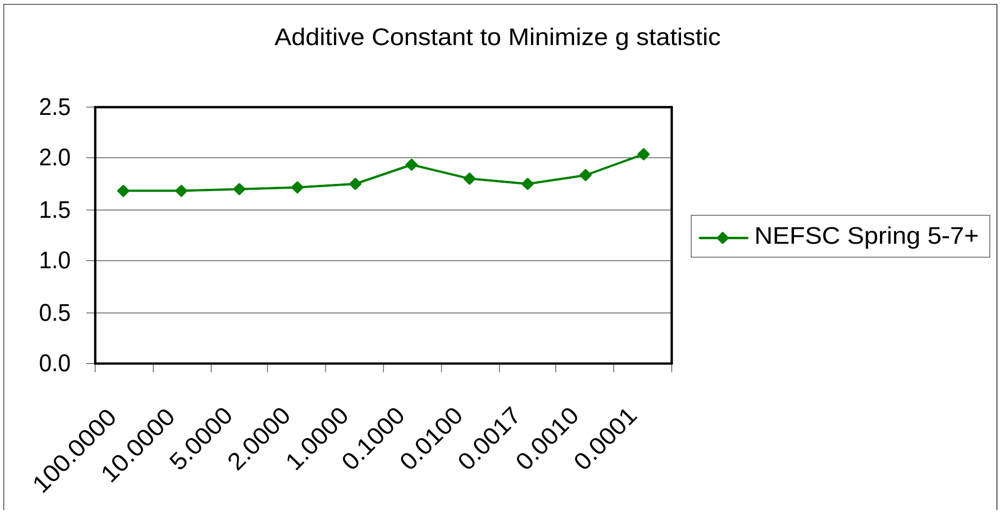
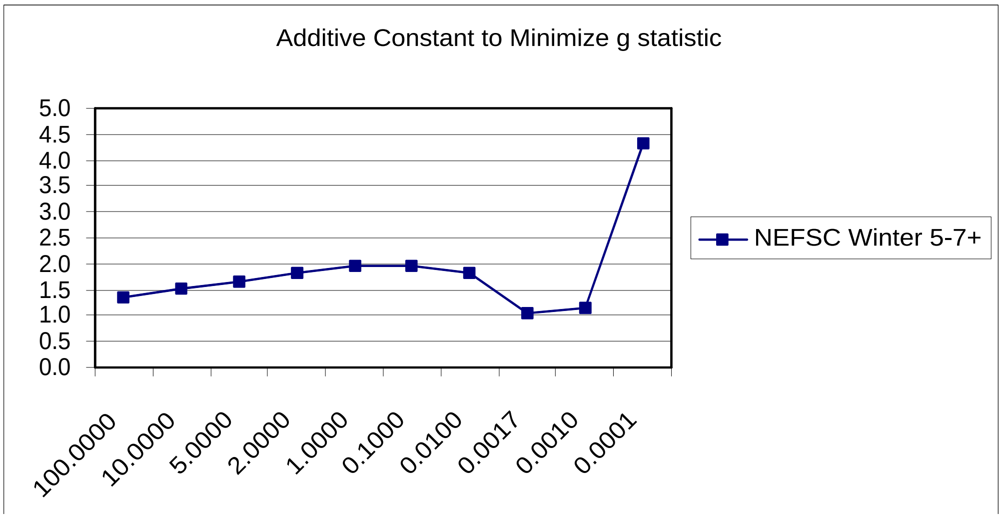


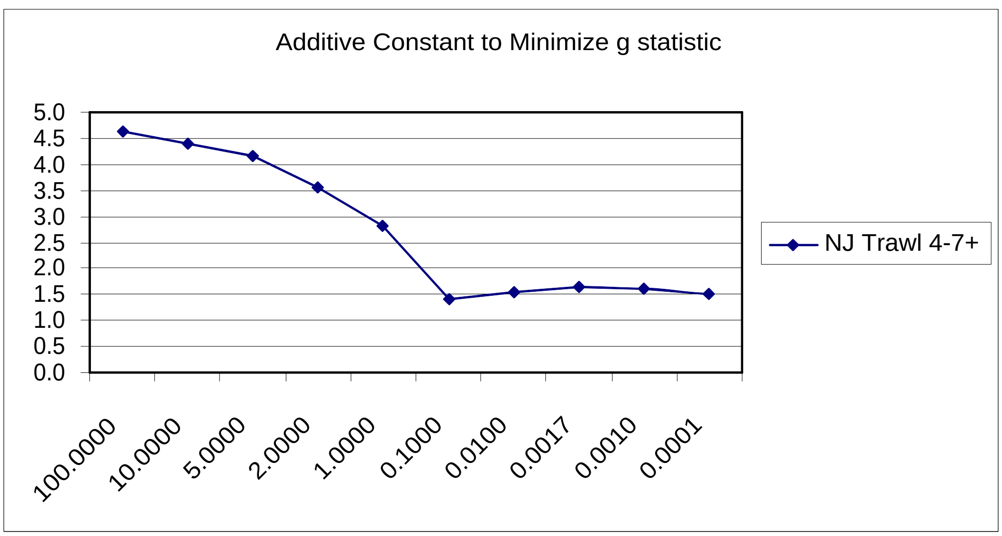
Figure 5 continued.



Additive Constant to Minimize g statistic
2.5
2.0
1.5
1.0
0.5
0.0
100.0000
10.0000
5.0000
2.0000
1.0000
0.1000
0.0100
0.0017
0.0010
0.0001
NEFSC Spring 5-7+



Additive Constant to Minimize g statistic
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
100.0000
10.0000
5.0000
2.0000
1.0000
0.1000
0.0100
0.0017
0.0010
0.0001
NEFSC Winter 5-7+



Additive Constant to Minimize g statistic
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
100.0000
10.0000
5.0000
2.0000
1.0000
0.1000
0.0100
0.0017
0.0010
0.0001
NJ Trawl 4-7+

Figure 6.