Estimating Catch at Age for Red Snapper in the Shrimp Fleet Bycatch

Scott Nichols
NMFS Pascagoula

Introduction

A Bayesian approach to estimating red snapper bycatch by the offshore shrimp fleet in the Gulf of Mexico was introducted at the SEDAR red snapper Data Workshop (SEDAR7-DW-3 and SEDAR7-DW-54). The approach allowed a more realistic characterization of uncertainty about the estimates than had been possible in the past – some of the consequences of the unbalanced sampling in the observer programs were accounted for. The analysis produced posterior distributions of bycatch in numbers by four-month "trimesters" in 8 geographical strata. The estimates were for total bycatch in number, without respect to age of the fish.

Discussion at the DW took note of the large amount of data now available on size composition of the bycatch. Examination of scatter plots of size vs date of capture showed that it would be possible to establish time dependent boundaries between age 1 from age 0 fish with virtually no ambiguity in second trimester, and minimal ambiguity in third trimester. Large differences in fractions of ages 0 and 1 among years were evident, which was not surprising, given year to year variations in year class strength. A preliminary run of a procedure to interpret sizes as ages was done at the DW using a single boundary set for all years. Results were promising, but the boundary set chosen did appear to 'nip' some modes some years. This was again not surprising, as both the timing of recruitment and growth surely have some interannual variations. This potential mis-ageing was expected to be corrected by considering year-specific boundaries between ages. The working group concentrating on bycatch at the DW also discussed a modeling approach, which was in essence multiplication of the statistical distributions of total catch from the Bayesian analysis by age vectors to be modeled as multinomials. The group adopted the spatial and temporal resolution of the Bayesian bycatch analysis, recommended a minimum sample size, and drafted a set of rules for substitutions if sampling were insufficient in any cell. I drew the charge to carry the analysis to completion.

Methods

The Bayesian procedure for total bycatch from SEDAR7-DW-54 was used as the source of bycatch in numbers. The procedures uses the BUGS program tabled in the SEDAR7-DW-54 appendix. There was one update – one highly unlikely data point from 1999 was identified after the DW, and removed. This removal changed the 1999 result (median reduce from 52 to 43). Elsewhere, changes were slight. The new results, as quantiles for annual catches, appear in Appendix 1. It was not practical, and probably not possible due to memory limits, to build the age composition analysis into the same BUGS program producing the total bycatch estimates. Instead, the lognormal parameterization for the total bycatch estimates suggested in SEDAR7-DW-3 was used input to this catch at age analysis. More details about this part of the procedure are described near the end of this section, after the age composition processing is introduced.

Size composition data suitable for estimating age composition are available for 1992-2003. The observer data base is actually a collection of several projects, with some differences in protocols. (See SEDAR7-DW-5; also SEDAR7-DW-3 and SEDAR7-DW-54). The problems caused by haphazard, unbalanced sampling discussed at length for the papers estimating total bycatch also potentially affect size composition data as well – there is no way to evaluate how representative the existing data are of the population of catches within cells.

Length composition observations from nets without BRDs were used for years prior to 1998. The year 1998 was split seasonally only, with trimesters 2 and 3 using BRD-net observations. Observer data for nets with BRDs were used for all years after 1998. The specific observer projects included up through

1998 are described in SEDAR7-DW-3. The projects included post-1998 were as described in SEDAR7-DW-54. To keep the 'E' data sets (data from all nets) on a 'per net' footing like the other sets, the catches in the 'E' data were divided by the number of nets reported.

Age designations in this paper continue the Goodyear convention of promoting all individuals on January 1. This is strictly a bookkeeping convention. On January 1, young of the year probably range in true age between 3 and 8 months.

The rapid growth of young of the year fish must be accounted for in determining a boundary between age 0 and 1. Choosing a single size as a boundary for any extended time interval could miss-assign fish rather seriously, especially in the summer. Fortunately, date and time are recorded for virtually all samples in the observer database. The final analysis here is based on graphs with time resolved to individual days.

There is no objective technique that will address all the issues to be considered in inferring age composition from length composition -- the analysis must rely on subjective decisions for interpreting visual evidence. I have gone to some length in what follows to document the subjective parts of this process, but upon rereading what I have written, I find that the detail may produce an unwanted consequence – it makes the age decisions look more uncertain than they really are. I would argue that any conscientious effort that incorporated the time/length resolution considered here would reach virtually identical 0:1 distinctions in summer, and would be unlikely to differ much in the fall. In the estimation of bycatch at age, uncertainty in ages of fish actually sampled will not be an important source of variation. The major quantified variation will remain uncertainty over numerical catch (all ages together). The major unquantified uncertainty will remain the doubts about whether the haphazard sampling of trips was truly representative of the behavior of the fleet, and this concern applies to both the catch rate and size composition data elements.

The age determinations here follow much the same procedures as those described by Goodyear (1985) and Nichols (1999), except that the analysis here is done for each year individually. The size information is plotted on a graph of length vs date of capture, and gaps between the age 0 and age 1 groups are often very evident. The DW was able to examine a plot of most of the data from all years plotted on a single graph, reproduced here, albeit at low quality, as Figure 1. The separations between zeros and ones are partially obscured by the shear quantity of data, but even in this dense plot, there were indications that straight line boundaries between zeros and ones should suffice, with one boundary line for each summer and one for each fall. In fact, a single pair of boundary lines applied all years would be reasonably accurate. However, the DW noted that accuracy could be improved by subdividing the data, with years as the first division. We expected there would be some interannual variation in growth rate at size, and could tell from inspection that there was considerable variation among years in the timing of the main influx of age zero fish.

Initially, this analysis followed another Goodyear convention, endorsed in the DW discussion, that all fish above the age 0:1 boundary but under 300mm would be assigned to age 1. After examining the annual breakouts of the data, I decided to make an attempt to extract more information, by evaluating a procedure to extract ages 0, 1, 2, and 3+. This attempt is described in later paragraphs, but I finally concluded that the partitioning I developed using four age groups would be less accurate than the original convention.

I separated the length frequency data files into annual subsets (combined over projects, areas, and depths), and made 12 plots like Fig 1. I recognized that I could later consider area subsets if separations were not evident in these Gulfwide figures (at the cost of a reduced number of measurements per plot). In the end, I concluded that further subsetting was quite unnecessary. I was concerned, however, that the one fish / one point perspective might be misleading – in high density regions it is hard to get any feel for the shape of the distribution of fish with length. I set the point by point figures aside temporarily, and plotted a series of histograms for all samples in 15 day intervals for the 12 years. I used this new set for preliminary decisions. I identified two classes of potential boundaries from the histograms – those with extended gaps between apparent modes; and those without a clear gap, but with local minimum in frequency. I recorded end points for the obvious gaps. For the histograms with potential overlap between age 0 and 1, I decided that the distributions did not appear regular enough for any formal distributional fitting to provide any advantage over patterns identified by eye. I instead recorded ranges determined by eye for the 'minimum zones,' with the objective of getting a nearly equal numbers of fish that might be overlappers from each year class on either side of bound, that is, a single bound somewhere in the recorded range should meet that objective exactly. I plotted these bounds in a format similar to Figure 1. Figure 2 is an example.

I set up an Excel spreadsheet with the capability to overplot boundaries on the plots of the gaps and minimum zones. Tentative boundary locations in the plots could be altered by changing the spreadsheet cells holding the parameters for the boundary lines. These parameters were manipulated by hand until a fit acceptable to the eye was obtained. The criterion used was to match as well as possible all the gaps and minima simultaneously, with the simplification that a single straight line be used for each year /season, with each year.summer and year.fall considered separately. It makes no difference where the boundary passes in a complete gap, and if a "interval of minimum" is well chosen, it should make little difference where the boundary passes within that interval (or else a better interval should have been chosen). However, note that in plots like Figure 2, there is no consideration of number of fish -- when choosing the best line satisfy all gaps simultaneously, there is no information in these plots about number of fish in each small time interval. I considered a linear boundary line for the summer months, a second linear boundary line for the fall months, and examined the appropriateness of the intersection of the summer and fall lines as a transition point. I also plotted von Bertalanffy boundaries between 0:1, 1:2, and 2:3. The vonBert bounds were a legacy of the Nichols 1999 analysis, and were based on Goodyear's curve in total length with +/- 0.5 yr on tzero, and Goodyear's converison to fork length.

I did not originally intend to use the vonBert boundary lines for much, but three things brought my attention to them. 1) In the years with winter observations, there was often a clear continuation of the previous fall's age ones as a separate mode, now designated age 2 by the Jan 1 convention. 2) in some year's the fall linear boundary lay almost atop the vonBert boundary line, indicating the vonBert line was as good an approximation as the straight line. I modified the spreadsheet to allow adjustment of the vonBert boundary, effectively by manipulating tzero, and started using that line as the fall boundary. 3) Upon adjusting the vonBert line for the fall boundary, I noted that the adjusted boundary also appeared to hold through the winter (in years with winter data), and for years with ample samples, that boundary seemed to separate some less obvious modes in the summer months as well. This last point suggested that the 300 mm upper bound for age 1 accepted by convention at the DW might be improved upon, and that perhaps age 0,1,2, and 3+ fish could be separated by this analysis.

Continuing the boundary determination procedure, I returned to the 12 plots of all points for each year (the Figure 1 format), and used the boundary parameters from the gap identification plots as starting points. Note that by going back to the individual fish format, temporal resolution improved from 15 days to 1 day. Some modes were still 'nipped' by the preliminary boundary lines, so the bounds could be improved. The plots with all fish also provide an indication of numbers along the time axis, so given the decision to simplify the boundary by use of single lines, I could evaluate the relative strengths of the evidence at different intervals on the time axis. I modified the parameters as needed to improve the acceptability of the fit to the eye, which usually required only a small adjustment, affecting a small number of fish. The lines thus established became the final choice of the boundaries between year classes. Note that at this point, I was carrying two versions. The 0:1 bound was the same in both versions. One version called everything above that bound but less than 300mm age 1. The other version uses the vonBert-shaped bounds to separate the fish into 1, 2, and 3+ classes. The final figures from these procedures are collected in Appendix 2.

I returned to the size composition data files, and programmed a SAS routine to assign an age to each fish based on its length and date of capture. The routine then weighted each sample (i.e. each vector of number of fish by age) by the number of fish caught in that trawl, and summed over all trawls in the year.trimester.area.depthzone combination. I also summed the age composition vector over year.trimester and trimester (ignoring the remain factors), which were the combinations chosen at the DW to be used as substitutes if individual cell sample sizes were less than 30 fish.

The Bayesian estimates of bycatch in numbers by year.season.area.depthzone and the age composition vectors available at the same resolution form one set of products that can input separately to some forms of stock assessment models. These products were also used here as the raw material for an additional step – developing a parameter-based description that modeled bycatch at age in another form that might be used in other stock assessment models. Catch at age estimates were obtained by multiplying the two items together, and summing over the geographic areas to be covered. The Bayesian estimates of catch in numbers incorporate the uncertainty created by the unbalanced sampling among cells, and the estimates are actually reported as statistical distributions. Age composition vectors have uncertainty – the observer data are draws from a much larger population of fish caught throughout the fleet. I modeled the uncertainty from that sampling process as a multinomial distribution. The catch at age estimates are in essence the "product" of those two distributions.

I commented in my SEDAR7-DW-3 paper that the distribution of bycatch in numbers remained skewed even on a log scale, but suggested it might be more practical to parameterize the results using a lognormal distribution, sacrificing any real information existing in the asymmetric tails. (I also argued that the upper tails extended into very biologically unrealistic regions, and probably should be discounted to some extent anyway.) Here, I did develop lognormal fits to the Markov Chain Monte Carlo (MCMC) results for the distribution of catch in number by year.season.area.depthzone. I used the median of the distribution on the log scale as a mean, and to develop a precision term, averaged the deviations of the two 95% confidence bounds. I also tried using a 2-parameter gamma distribution, using parameters that would return the mean and variance tabulated in the MCMC results. (The gamma can incorporate a skew explicitly.) I then ran BUGS simulations for the fitted lognormal and gamma distributions, and compared by eye the quantiles generated with the original MCMC results. Somewhat surprisingly, the lognormal did better at approximating the original quantiles than the gamma did, despite the skew. I chose to use the lognormal parameterization without further debate.

The multinomial portion of the estimation uses the age fractions for calculated for each cell vector, and number of fish actually measured as parameters. This structure is a simplification of the actual sampling procedure, in that it treats each fish as if it were a random selection from the population of fish caught by the entire fleet in each cell. As the underlying trip selection was haphazard, I did not believe exploring more complicated approximations would be productive.

The DW working group considering bycatch had settled on a minimum of 30 fish for any cell to be described. Failing to reach the minimum in any cell, a sequence of substitutions was proposed: 1) sum over all samples in that year.season (i.e. ignoring area and depth), 2) sum over all samples in a season (ignoring year, area, and depth). Using number of fish as the minimum criteria makes no recognition of how the measurements might be spread among trips, or over time and space within cells, but again, given the haphazard nature of trip selection, there would probably be little real gain from using anything more complicated.

Wishing to keep the multinomial structure even for 'substituted' cells (primarily for analytical convenience), I considered how to model variation in the substituted cells. Using total number of fish measured in the combined distributions used for substitution would be very misleading, as this would imply high certainty of the age composition for each individual 'prediction' made by the combined distribution. Some type of 'random effects' structure might be more rigorous, but I again opted for simplicity, and decided to use assign a sample size of 25 as a stand-in for the substituted cells. This approach generates some variation due to uncertainty in age composition in the substituted cells, but the amount of variation is arbitrary.

The procedure multiplying the lognormal catch distributions by the multinomial age vectors and summing to trimester totals, was written in BUGS code (Appendix 3). For each run, I used 5500 iterations of two chains, and discarded the first 500 of each, making available 10k iterations to characterize the distributions of the log of the trimester catches by age. To produce a parameterized version of these results, the medians from the BUGS quantile outputs were used as means, precision was calculated from the 95% points by averaging just as for the parameterization of the total catch, and output from the BUGS coda procedure was used to calculate the correlations among the age estimates for each trimester (done in SAS). This procedure resulted in a multivariate lognormal model for bycatch at age for each trimester.

Sometimes the catches at age will equal zero on an arithmetic scale, which must be addressed in the lognormal modeling. Some are structural zeros, like age 0 fish in the first trimester. These can be ignored. Others have zeros for only some iterations, when random draws from the multinomial returned zero. For those, I took the logs of the upper quantiles that were greater than zero, and projected a lognormal distribution from them. The effects on the catch estimates are small, as all catches so simulated are very small, but the procedure does allow retention of variability information for cells that would otherwise have to be set to a constant zero, or treated as missing.

Because estimates were also requested for potential east and west substocks, similar products were developed from spatial subsets of the data and the intermediate results. I decided that the problems with sparse, unbalanced, and haphazard observer CPUE data would be too severe to entertain separate east and west models for total bycatch in numbers; so I simply took the distributions for the regional totals from the overall Gulf model, and parameterized those as described for the Gulfwide estimates. The age composition data are quite a bit more 'stable.' As the Gulfwide age boundaries looked fully adequate, and should apply to any subset of the data without modification, I made no age boundary plots for the subsets. However, using those boundaries, I did repeat the age composition calculations as described for the Gulfwide analysis on the east and west subsets of the data.

Results

The 12 annual graphs of individual fish sizes vs date of capture, and the boundaries selected, are available in Appendix 2. At the point these figures were developed, I was still considering two versions of size to age conversions – a breakdown into 0, 1, and 2+, with a constant bound between 1 and 2+ at 300 mm, and a breakdown into 0, 1, 2, and 3+, with the 1:2 and 2:3+ bounds set by vonBertalanffy-shaped curves. One set of figures suffices, however, as the 0:1 boundary is the same in both schemes. The 0:1 boundaries seem quite definitive, with virtually no ambiguity in the summer, and very little in the fall. Examined collectively, the tentative 1:2 and 2:3+ boundaries were less convincing. There does seem to be a clear continuation of the age 1 mode into the winter months in years with winter samples. Impressions about possible continuation of that mode into later months vary among years, and among shorter time intervals within years. In some cases, fish above the vonBert upper age 1 line look to be members of a separate mode. In other cases, they appear to be just extensions of the age 1 distributions. In many intervals, the sizes are too scattered to form an impression.

I constructed a graphical test for the reliability of the tentative 'higher modes' by looking for consistency in relative year class strengths based on the 4 – age group scheme. I constructed graphs, plotting the ratios a season's age 2 : age 1 against the ratio of the previous fall's age 1: age 0 on a log scale. The sparse first trimester graph (Fig. 3) shows a positive relationship, but somewhat below the one-to-one line. The second trimester graph (Fig. 4) shows a flatter relationship, falling further away from one-to-one. The third trimester graph (Fig. 5) shows a negative relationship. A graphs of the third trimester 2:0 ratio against 1:0 ratios from the same year show a positive relationship (Fig. 6). I interpret this sequence of graphs to mean that the first trimester 2's are largely true age 2's, but that the second trimester "2's" are actually highly contaminated by large age 1's, and that the third trimester "2's" are virtually all large age 1 fish. I concluded the scheme for extracting four age groups with three time-varying boundaries was less reliable than the 0, 1, 2+ scheme with the constant 300 mm boundary between 1 and 2+, and only considered the 0, 1, 2+ scheme in what follows.

Tables of the age fractions and the estimates of bycatch at age are quite lengthy, and are not shown here. The tables have been forwarded to the Miami stock assessment group in several forms for use at the AW. For this paper, I have collected a few summary descriptors of age composition, mostly in graphical formats.

Year to year variation in the fraction of age 0's in the second and third trimesters are shown in Figures 7 and 8. (The first trimester fish are now almost all designated age 1 by the 300 mm convention, so no graph is shown.) Figure 9 shows the fraction of zeros in the third trimester plotted against fraction of zeros in the second. The flat pattern and the wide spread on the second trimester axis suggest that the variation in the second trimester arises more from variations in timing of the seasonal influx (with, as always, possible confounding by sample coverage), rather than from variations in year class strength. These 3 graphs are based on the fractions in the samples. To get a feel for how this translates into uncertainty at the end of the BUGS simulation to derive catch at age, I have included the quantiles for the fraction of age 1's in the second trimester as Table 1. (The medians can differ from sample values, because the sample combinations are calculated ignoring area and depth. The simulation carries through the spatial distribution among cells of the bycatch estimates.)

The figures in Appendix 2 can be examined to evaluate the year to year variation in seasonal timing of the main influx of age zeros. A few years do not have samples in critical periods, but an overall pattern still seems clear. A trace may show up even as early as June (2002) or early July, but then zeros might be absent or nearly so for a few weeks. The main influx, however, tends to be sudden, obvious, and persistent. This main influx appears most often in late August, but appeared as early as late July in 1995, and as late as late September (1997) – the interannual variation was substantial. However, I did not investigate the spatial distribution of the samples in the most extreme years, so there is still a chance of some confounding between interannual timing and spatial differences.

The final summary set to be presented is the age composition for each trimester established by combining all samples, ignoring year, area and depth zone (Table 2). This distribution was the ultimate 'substitution' distribution for insufficiently sampled cells (if more localized substitution options still did not total 30 fish), and as a method, derivation of this distribution most resembles the procedures of Goodyear 1995 and Nichols 1999. Under the substitution rules recommended by the DW, these would be the age compositions used for all years prior to 1992, although most of the assessment approaches being considered may not use them that way. Compared to the Goodyear 1995 and Nichols 1999 results, this distribution shows more age 0s in the 'average' bycatch, courtesy of the years with large age 0 contributions since 1998.

Discussion

There appear to be ample data to contribute to year class strength estimation in the assessment from 1992 forward, although only 93 of the 288 year, season, area, depthzone cells meet the 30 fish minimum criterion. This means that most of the cells are actually filled using the data from all areas and depth zones within each year.trimester. Tthe trimesters not meeting the minimum of n=30 in any year were: trimester 1 in 1992, 1997, 1998, 1999, and 2001; and trimester 3 in 1996. For substitution, the cells in these trimesters default to the overall fractions of Table 2. The estimates do have large uncertainties, generated primarily by the sparse and unbalanced data available for the estimation of total catch in numbers. Our confidence in results beyond the statistical uncertainty is most limited by the nature of the sampling of trips, but given the sampling, I believe this analysis has represented our knowledge -- and via the confidence intervals, the limits to our knowledge -- as realistically as practical.

There is certainly subjectivity in the determination of the age boundaries. That may cause concern, because subjective techniques are inherently not fully reproducible, but the subjectivity does not seem to be a limiting issue here. It seems unlikely that different investigators would come up with substantially different 0:1 boundaries, and the less convincing situation regarding 1:2+ boundary is not likely to be a major source of error.

The multinomial modeling using number of fish actually measured as sample size (rigorously appropriate only if each fish measured is independently and randomly drawn from the larger population of all fish caught by the fleet) was also the technique used in developing an M estimate for juvenile fish from trawl survey data (Nichols Pellegrin & Ingram 2004). We commented in that paper that the sampling actually used in the trawl surveys probably made the multinomial a very good approximation for that analysis. A similar comment cannot be made for the bycatch observer data, but I do not see another route that would be appreciably better. There is really no statistical model for the haphazard draw of vessels and trips (which are linked to area / depth cells only indirectly through proximity to the ports of departure). Within trips, data come from sequential tows, but that is probably not nearly as serious as a 'violation' of statistical assumptions as the trip selections. Fort the majority of the catches, all snapper were measured; subsampling was rarely invoked. The 30 fish minimum could be satisfied by a single tow, but in general, there were multiple trips contributing data in most cells that had data. In the end view, I consider the multinomial model as convenient way of incorporating a hopefully realistic contribution to a variation that is dominated by the uncertainties of the CPUE predictions. The approach used here certainly seems superior to treating the age classes as independent 'species.'

Although the CPUE estimations of SEDAR7-DW-3 and SEDAR7-DW-54 were Bayesian, none of the analytical steps after that CPUE estimation were, despite the use of BUGS as a programming tool. BUGS was used here to simulate the distributional consequences of extended functions of the several intermediate products leading to catch at age. There is potentially some loss of information in the approach used, in the steps to parameterize the CPUE trimester totals as if they were independent estimates. This procedure wipes out any hidden correlations among years / seasons. There is nothing directly in the CPUE model structure that would lead to interdependence (indeed, one would expect the existence of the 'local' term would minimize any dependence), but some correlation structure could have been imposed by the very unbalanced data. However, an end-to-end Bayesian analysis to estimate catch at age within a single BUGS program would probably be well beyond the capacity of any computer we have available.

Several of the years with high fractions of zeros since 1998were particularly rich in sample size, and thus one may question the appropriateness of using the combined data, such as for substitution for years prior to 1992 in a procedure similar to Goodyear's. It is not the increase in fraction of zero's that might cause one to question using a combined sample, it is the increase in variability since the Goodyear analysis. In Goodyear's day, the data available nearly all came from years that happened to have low fractions of zeros, and thus there was no direct indication that the fraction was as variable among years as it has turned out to be since that time. Some types of assessment models introduced since the last assessment no longer need to have substitute size compositions matched to all catches, so the issue may be moot for those types of models.

It was disappointing that a reliable identification of age 2 ultimately could not be supported by the data (particularly considering how much time I invested in trying). Under the scheme ultimately accepted, there will be some error introduced by misidentifying age 2 fish as age 1. However, most of that error will occur in the (low bycatch) first trimester, and there should be little impact on the overall assessment. The flatness of a constant 300 mm upper boundary on age one is probably not an issue – a vonBert bound would also be fairly flat at the scale of size vs date figures. The choice of 300 mm was an arbitrary one, but there are so few fish at that size that another choice would rarely change the estimates for fraction of 1's appreciably.

There is one last statistic to introduce. Presuming the time series was long enough to average out recruitment variation, the log of the ratio of the fraction of 1's to the the fraction of 0's from the data with all years pooled can be taken as an estimate of average annual Z. The value calculated is 1.90, strikingly close to the independent estimate of 1.98 from the SEAMAP trawl survey data. The two estimates are independent, so their similarity is reassuring. The persistence of the age 1 fish into the winter samples suggests that these Z's may not be appreciably confounded with reduction in vulnerability to trawls.

Literature Cited

SEDAR7-DW-3
Nichols S. 2004. Some Bayesian approaches to estimation of shrimp fleet bycatch.

SEDAR7-DW-5.
Scott-Denton E. 2004. Observer coverage of the US Gulf of Mexico and Southeastern Atlantic shrimp fishery, February 1992-December 2003 – Methods.

SEDAR7-DW-54
Nichols S. 2004. Update for the Bayesian estimation of shrimp fleet bycatch.

Goodyear C P. 1995. Red snapper in the US waters of the Gulf of Mexico. Report tot the Gulf of Mexico Fishery Management Council MIA-95/96-05.

Nichols S. 1999. Estimating fishing mortality rate reduction by BRDs. Report from the Summer 1998 Red Snapper / Shrimp Research Program. Report to the Gulf of Mexico Fishery Management Council.

Nichols S, Pellegrin G & Ingram G W. 2004. Estimation of juvenile M for red snapper based on SEAMAP survey data. Report for the SEDAR7 Stock Assessment Workshop.

Table 1. BUGS quantiles for fraction of age 1's in the second trimester. (These are not sample totals. These are end products from the simulation program, thus incorporating catch variation and substitutions for cells with n<30.) The years are 1992-2003, designated by 21-32 in the first index for fr[ ].

node	mean	sd	MC error	2.5%	25.0%	median	75.0%	97.5%
fr[21,2,2]	0.6471	0.05861	6.068E-4	0.5283	0.6084	0.6481	0.6872	0.7588
fr[22,2,2]	0.9821	0.01406	1.378E-4	0.9479	0.9746	0.9847	0.9936	0.9975
fr[23,2,2]	0.9635	0.02047	1.95E-4	0.9141	0.9528	0.967	0.9783	0.992
fr[24,2,2]	0.6007	0.03879	3.422E-4	0.5244	0.5749	0.6008	0.6261	0.678
fr[25,2,2]	0.1886	0.05846	5.532E-4	0.0954	0.1476	0.1821	0.2209	0.3275
fr[26,2,2]	0.9947	0.005019	5.349E-5	0.9813	0.9927	0.9961	0.9981	1.0
fr[27,2,2]	0.5863	0.05801	5.155E-4	0.4545	0.5521	0.5922	0.6288	0.679
fr[28,2,2]	0.1465	0.03393	3.454E-4	0.0857	0.124	0.1439	0.1667	0.2198
fr[29,2,2]	0.7252	0.05709	5.869E-4	0.6119	0.6863	0.7264	0.7654	0.8328
fr[30,2,2]	0.6565	0.03827	3.473E-4	0.5904	0.6296	0.6531	0.6801	0.7415
fr[31,2,2]	0.681	0.02329	2.391E-4	0.6482	0.6643	0.6768	0.693	0.7387
fr[32,2,2]	0.9881	0.01006	1.054E-4	0.9648	0.9832	0.9919	0.9951	0.998

Table 2. Trimester age composition fractions with all data combined over years, areas, and depth zones (i.e. "ignoring").

Trimester	N	Fract 0	Fract 1	Fract 2+
1	6965	0	0.9975605665	0.0024394335
2	44494	0.368650179	0.6266714703	0.0046783507
3	85186	0.8672137733	0.1298364314	0.0029497953

Figure 1. Plot developed at the SEDAR7-DW with most available observer data overplotted in a figure of fork length vs time of capture. The amount of data exceeded Excel's limits, so some portion was omitted. The lines represent some preliminary tries as age boundaries. The original figure was too large to transmit to this word document, so this is a lower quality 'Print Screen' snapshot of an Excel page. The length scale was truncated at 300 mm.

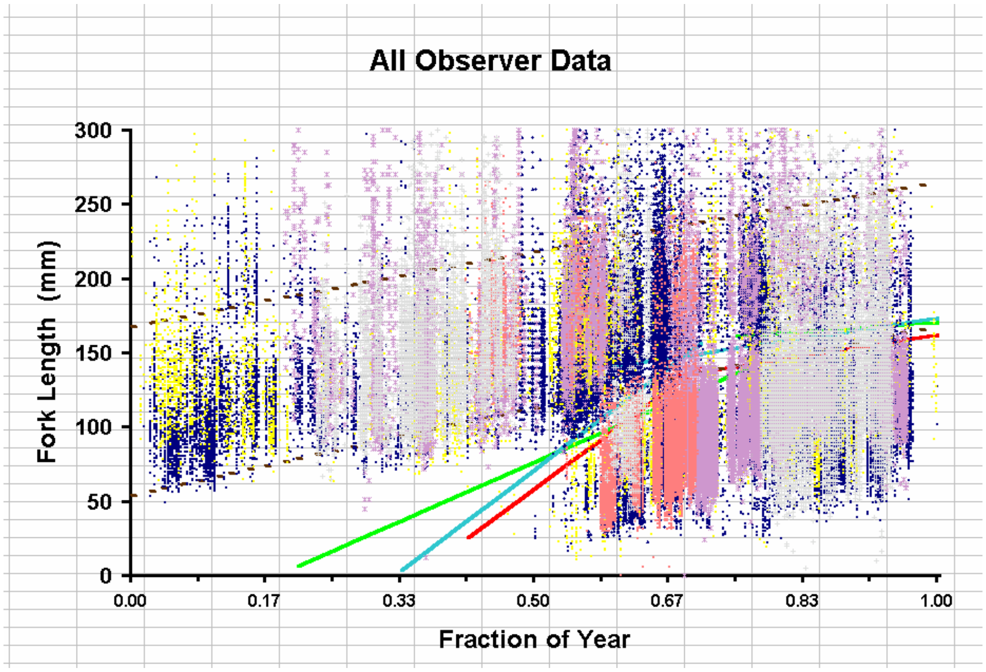


Figure 2. Example figure from the series used to place preliminary boundaries in the gaps and minimum intervals between modes. This, and all plots shown in the appendix are truncated at 400 mm. (Fish occur sporadically but reliably at size up to about 1 meter.)

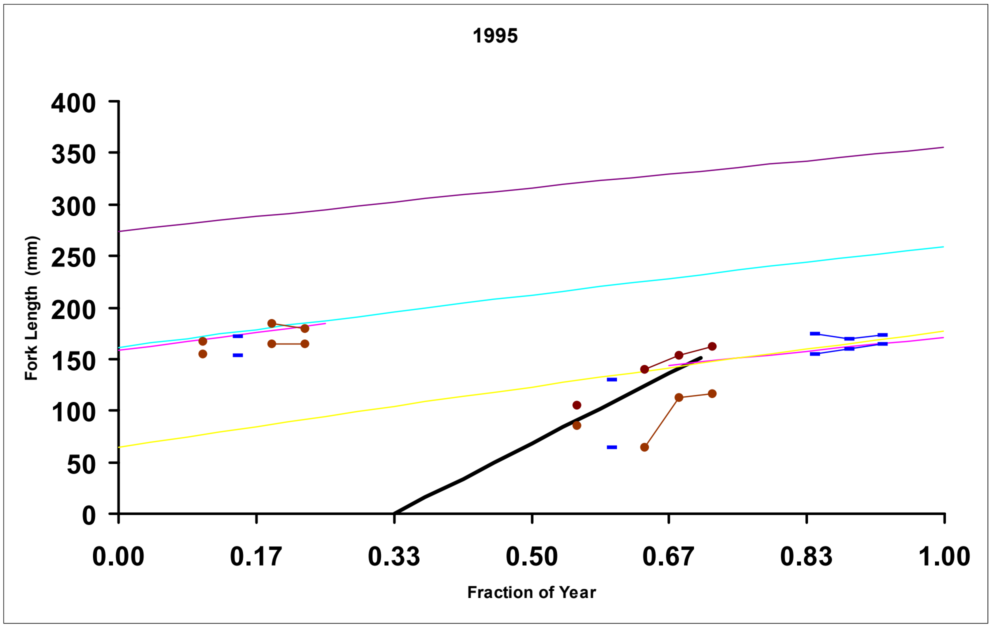


Figure 3. The ratio of age 2:age 1 in the 1st trimester vs the previous fall's (3rd trimester) age 1:age 0. With no differential losses, a 1:1 relationship would be expected. The fitted line falls somewhat below 1:1, although many of years with high 1:0 had no winter samples the following winter.

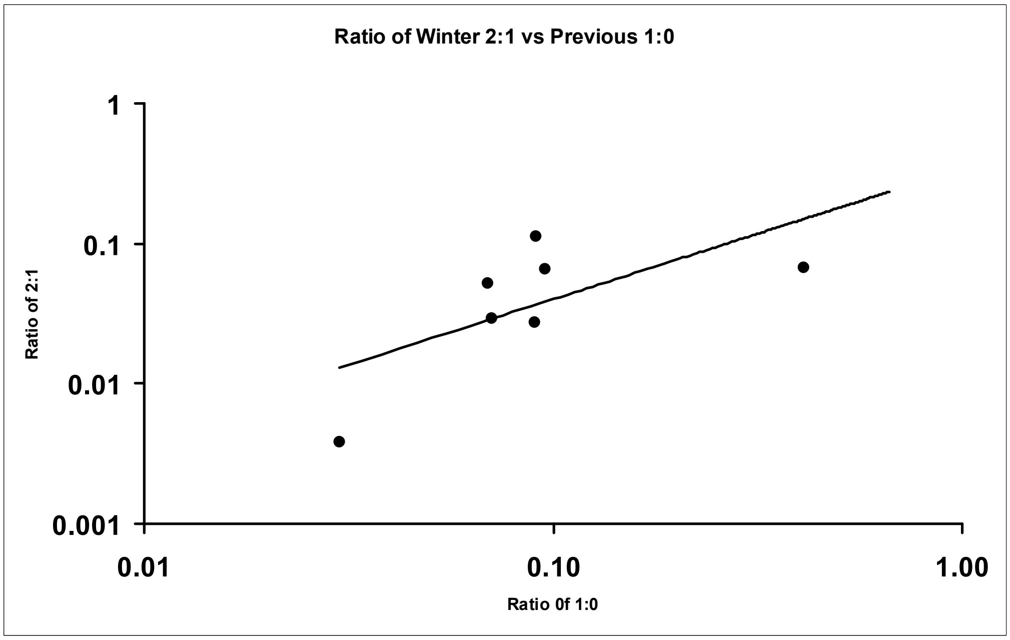


Figure 4. The ratio of "age 2":age 1 in the 2nd trimester vs the previous 3rd trimester age 1:age 0. Compared to Figure 3, the relationship appears to have fallen away further from the 1:1 line.

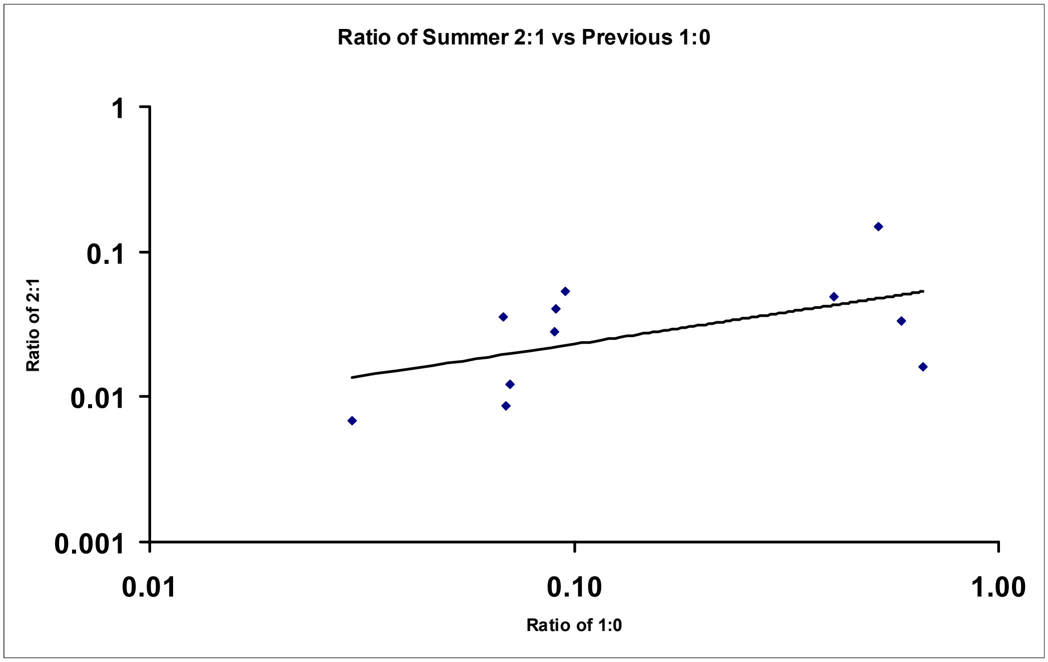


Figure 5. The ratio of "age 2":age1 in the 3rd trimester vs the previous 3rd trimester age 1: age 0.

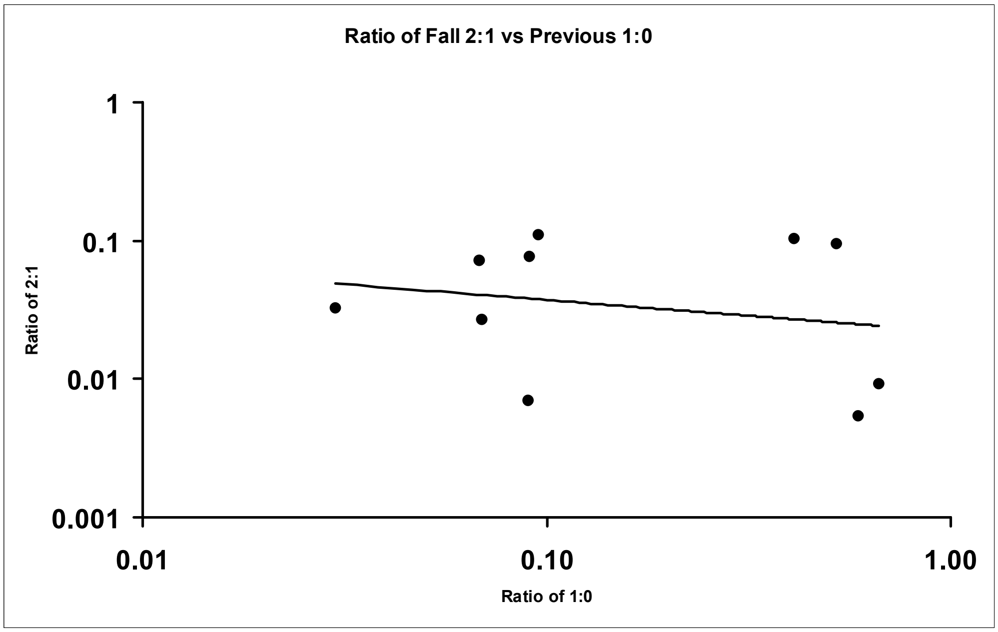


Figure 6. The ratio of "age 2" :age 0 in the 3rd trimester vs the age 1:age 0 ratio in the <u>same</u> trimester. This plot shows that putative "age 2" abundances actually closely follow the age 1 abundances, suggesting the "2's" may in fact be particularly large 1's.

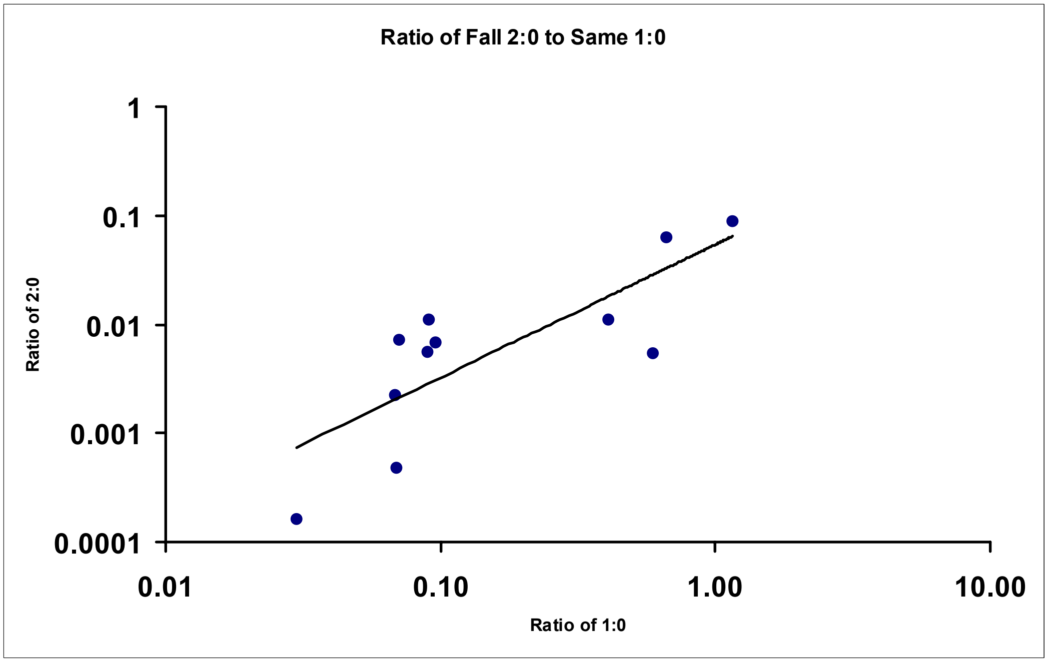


Figure 7. Fraction of age 0's in the second trimester by year (area and depth zone ignored).

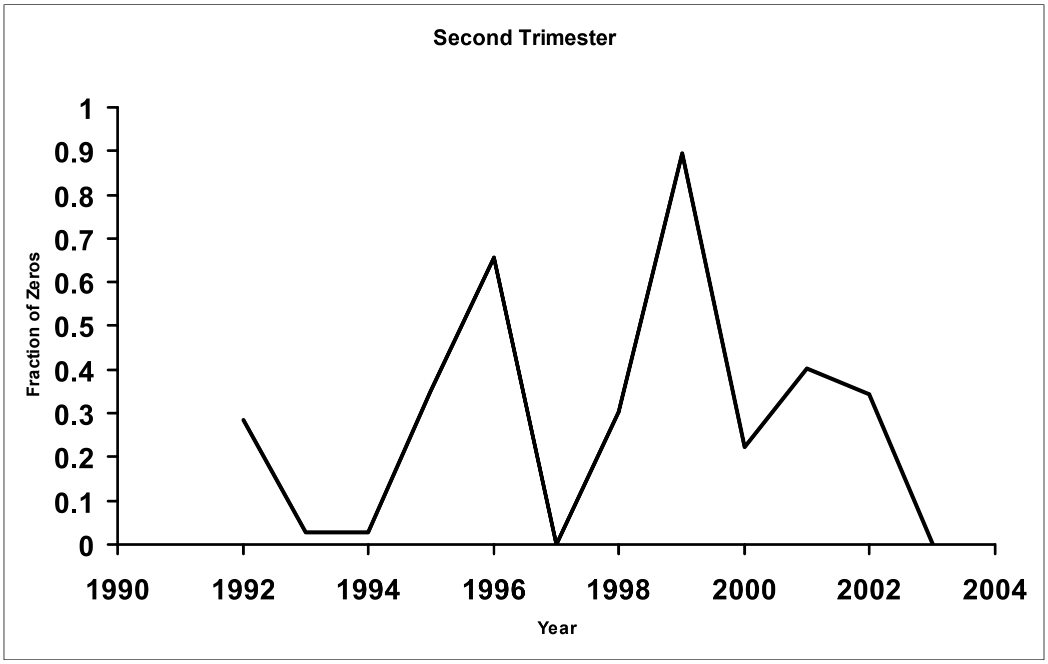


Figure 8. Fraction of age 0's in the third trimester by year (area and depth zone ignored).

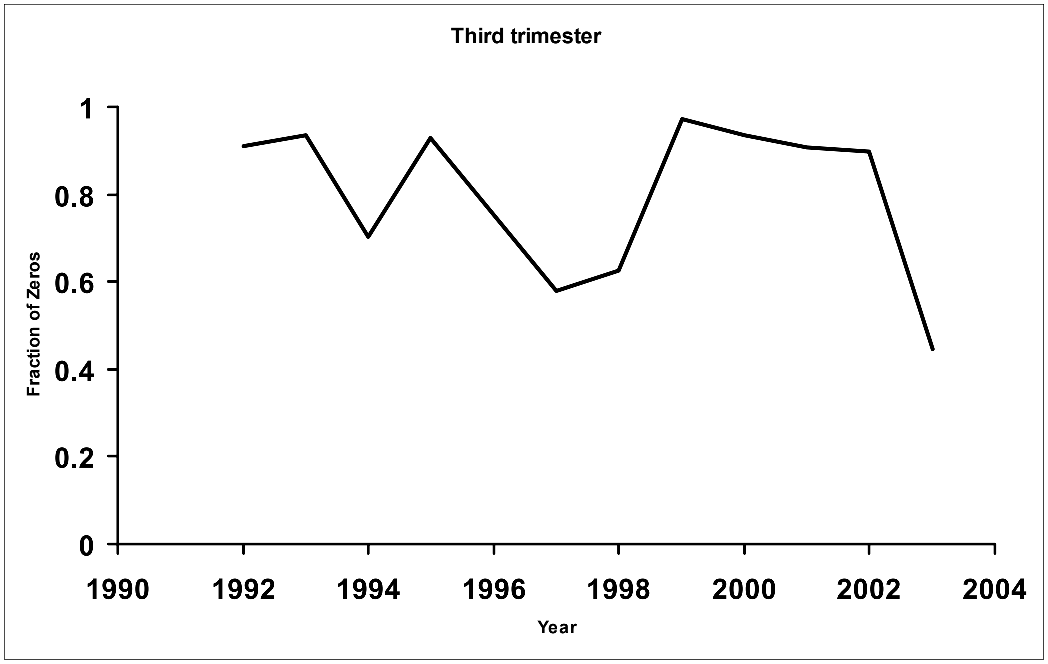
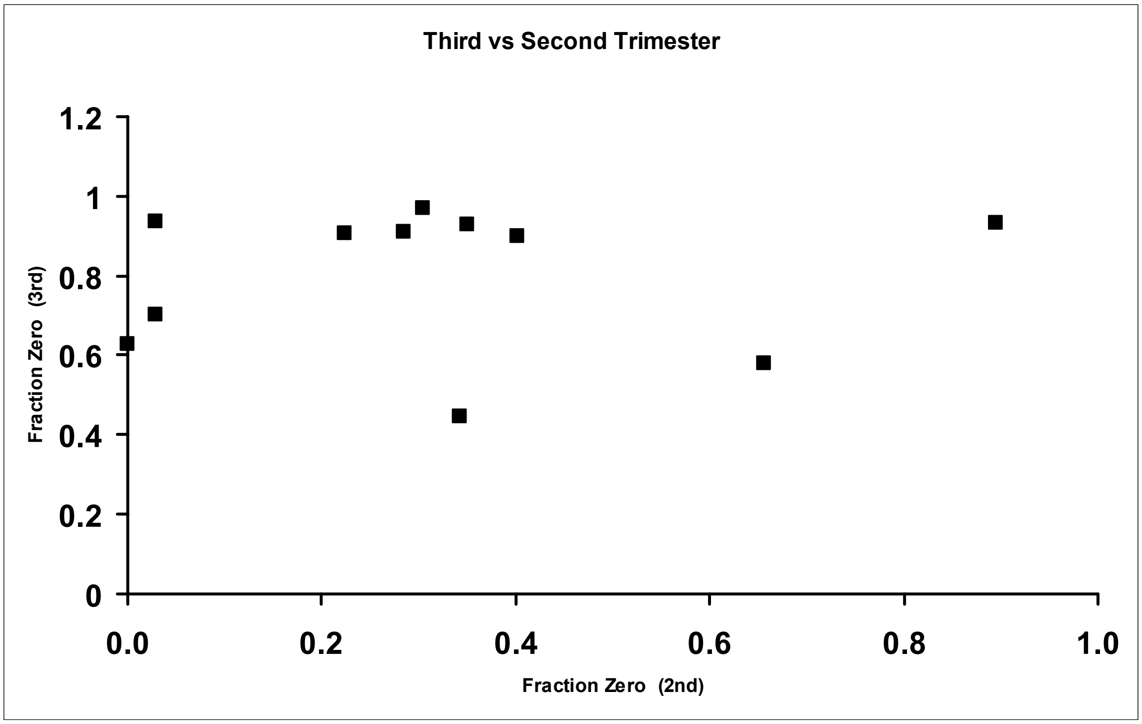


Figure 9. Fraction of zeros in the 3rd trimester vs fraction of zeros in the 2nd trimester.

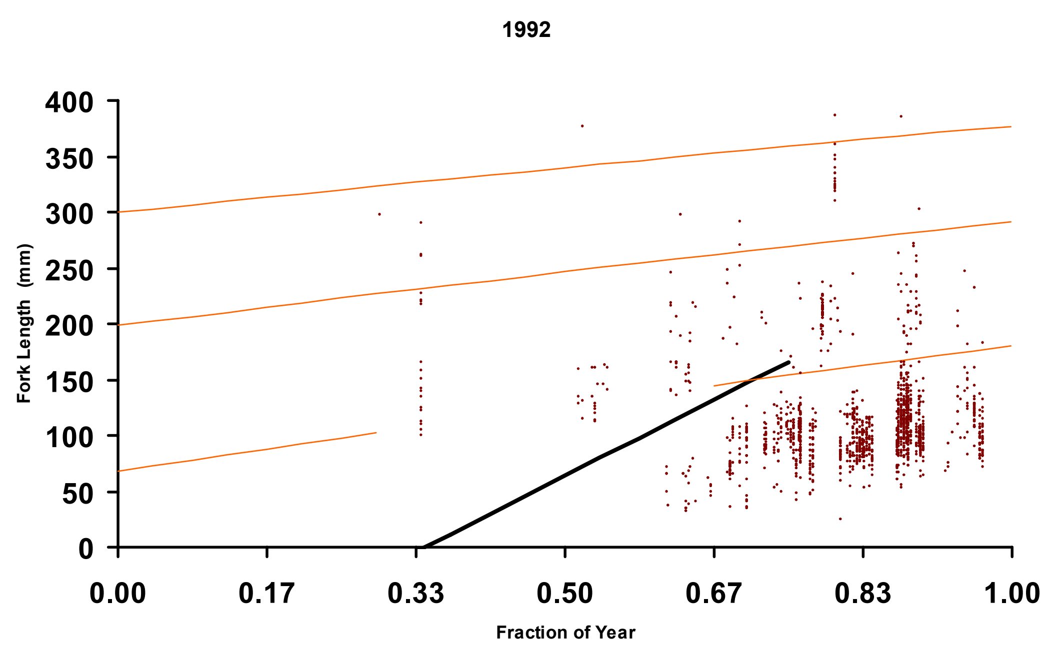


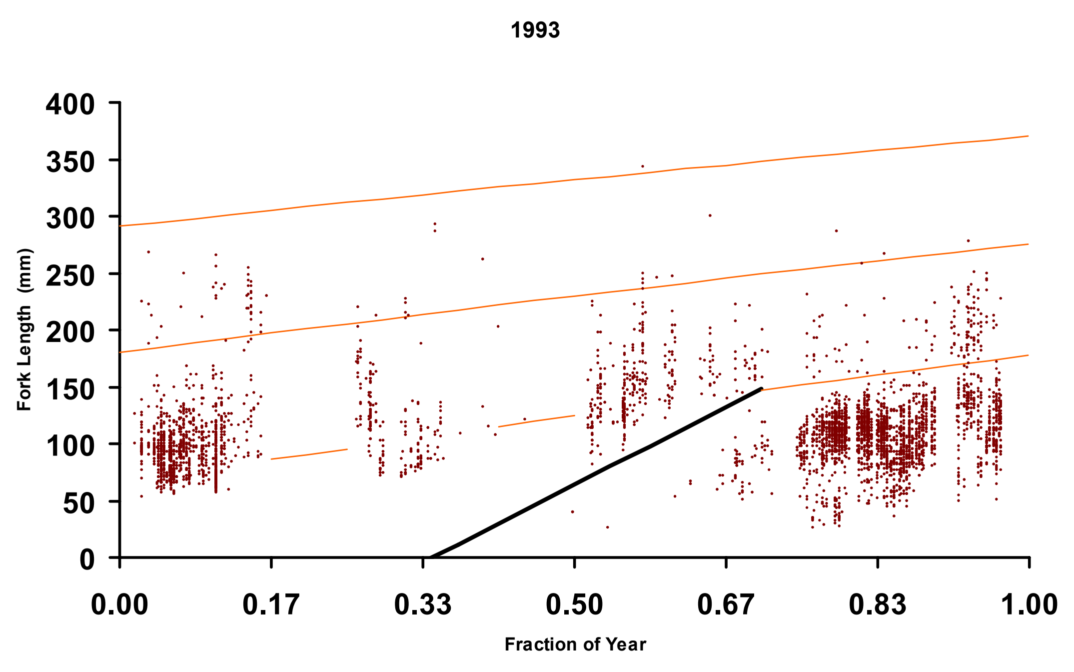
Appendix 1. Quantiles from the repeat run of SEDAR7-DW-54 results, with one bogus data point identified after the DWremoved. There main difference due to the removal was in 1999 (median dropped from 52 to 43). There were small differences in other years, but most would not even show on graphs like those in SEDAR7-DW-54.

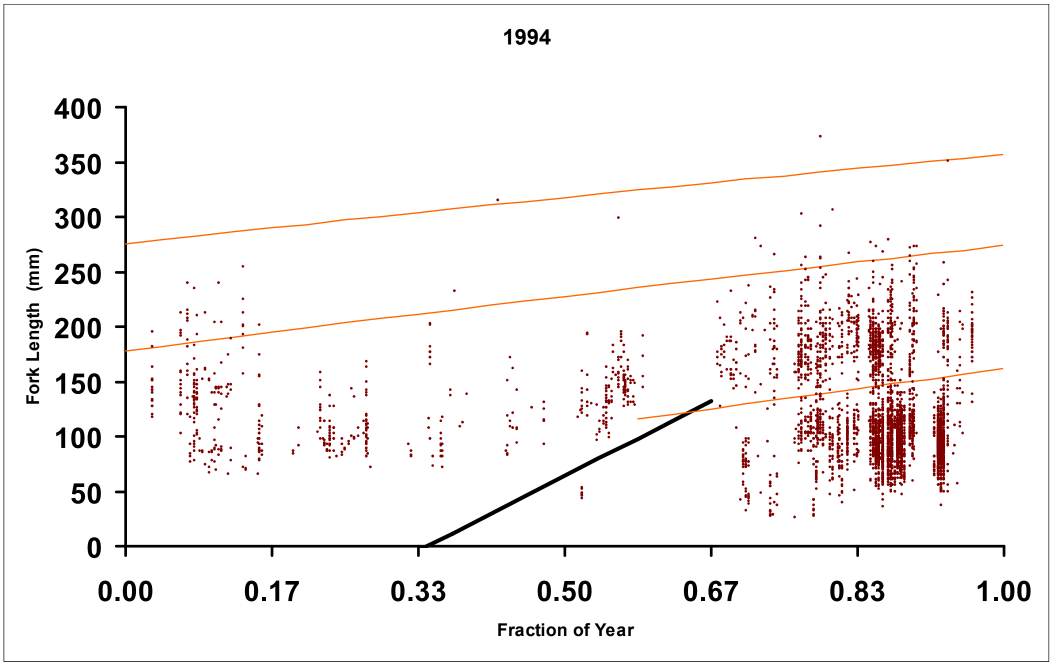
Node statistics

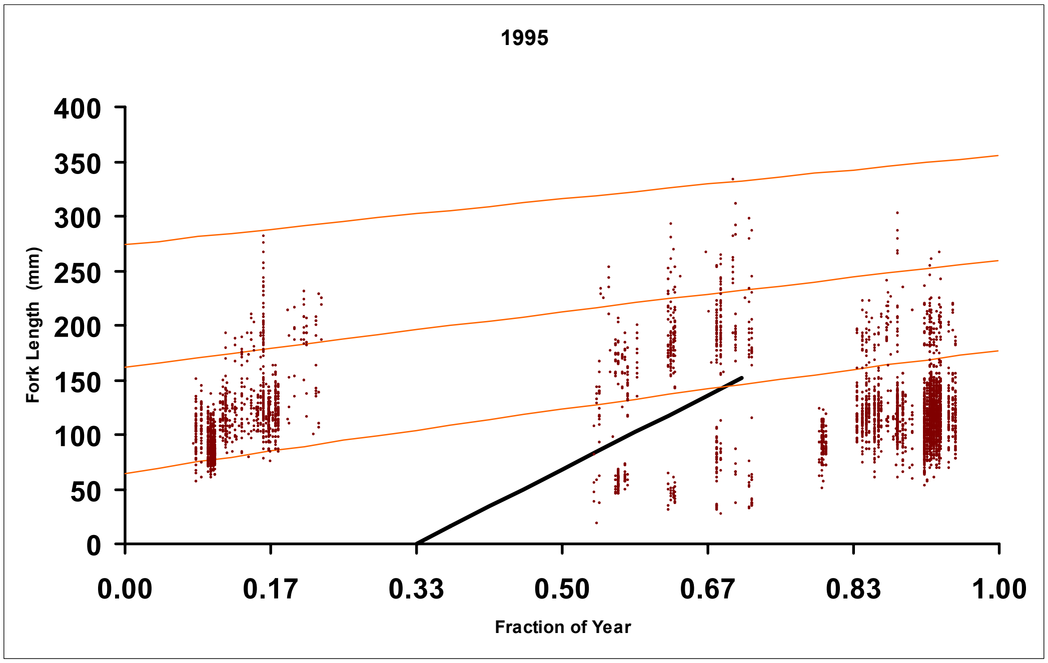
node	**mean**	**sd**	**MC error**	**2.5%**	**25.0%**	**median**	**75.0%**	**97.5%**
annual[1]	208.9	273.9	5.814	37.59	84.14	136.1	235.0	829.3
annual[2]	23.97	18.6	0.3849	7.453	13.28	18.8	28.1	71.8
annual[3]	21.76	10.19	0.2157	10.27	15.39	19.48	25.26	46.46
annual[4]	14.43	16.92	0.3018	3.088	6.286	9.852	16.42	54.25
annual[5]	38.34	9.661	0.1624	24.6	31.64	36.75	43.03	61.96
annual[6]	19.78	6.115	0.1008	12.19	15.99	18.64	22.19	33.61
annual[7]	11.84	5.527	0.1038	5.69	8.397	10.6	13.83	25.24
annual[8]	57.17	69.65	1.803	7.599	22.22	38.74	67.77	221.1
annual[9]	42.19	18.14	0.2848	21.88	30.65	37.8	48.44	87.73
annual[10]	99.25	102.1	1.325	35.3	53.79	71.9	106.6	336.5
annual[11]	40.64	51.38	0.834	9.369	17.97	27.36	44.58	153.8
annual[12]	26.71	32.36	0.6639	4.797	11.15	17.98	30.33	102.7
annual[13]	25.46	33.92	0.6958	4.226	10.3	16.83	28.7	97.85
annual[14]	24.53	33.6	0.5913	4.507	10.22	16.35	27.87	92.62
annual[15]	13.25	15.85	0.2889	2.33	5.486	9.002	15.19	50.23
annual[16]	27.54	34.65	0.5788	4.885	11.42	18.76	31.82	102.9
annual[17]	21.03	26.25	0.4494	3.639	8.691	14.16	24.09	80.69
annual[18]	27.92	35.27	0.588	5.139	11.74	18.94	31.82	104.7
annual[19]	108.5	141.2	2.583	18.83	44.24	72.64	123.7	414.6
annual[20]	78.16	102.6	1.737	13.32	31.5	51.84	89.11	297.8
annual[21]	31.46	7.585	0.09953	20.61	26.32	30.31	35.19	48.98
annual[22]	31.12	4.393	0.0531	23.94	28.09	30.65	33.59	40.95
annual[23]	41.54	9.096	0.1415	29.4	35.83	40.17	45.56	61.25
annual[24]	54.3	15.8	0.2369	32.51	43.52	51.5	61.7	92.25
annual[25]	53.2	53.1	0.7855	17.89	29.22	39.91	58.27	174.3
annual[26]	31.05	16.06	0.3007	13.59	20.94	27.24	36.44	71.27
annual[27]	54.93	31.82	0.4485	24.13	36.04	47.26	63.8	133.4
annual[28]	46.91	19.14	0.4219	25.72	35.58	43.06	53.59	90.14
annual[29]	17.19	6.602	0.1032	11.3	13.86	15.8	18.57	31.8
annual[30]	27.38	9.755	0.156	17.52	21.92	25.32	30.04	49.61
annual[31]	22.52	3.109	0.03958	17.5	20.32	22.15	24.29	29.7
annual[32]	10.58	3.781	0.07072	6.285	8.152	9.698	11.94	20.36

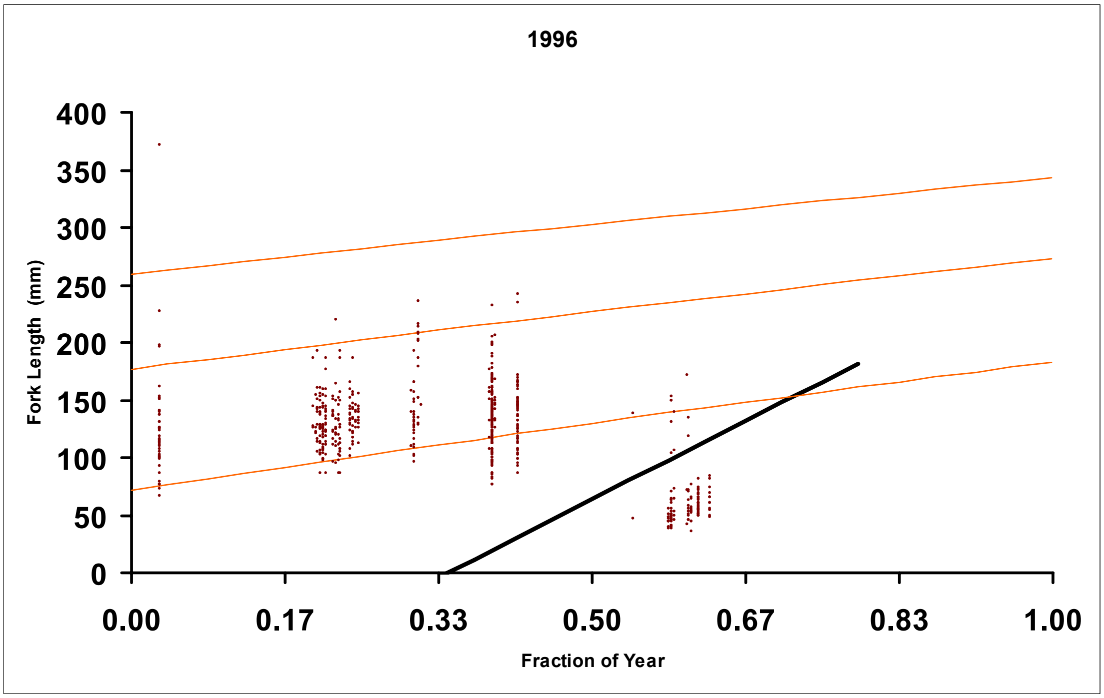
Appendix 2. Length vs time of capture plots with the age boundaries under development. For the age 0 : age 1 boundary, the final choice was to use the straight line boundary (heavy line) in the summer months, and the adjusted vonBert line (lowest fine line) in the fall months, starting in timeat the intersection of that vonBert line with the summer boundary line. The age 1 upper boundary was taken to be 300 mm at all times. The two upper vonBert lines, and the lowest vonBert line prior to intersection with the summer line were considered, but not used in the final age assignmests.

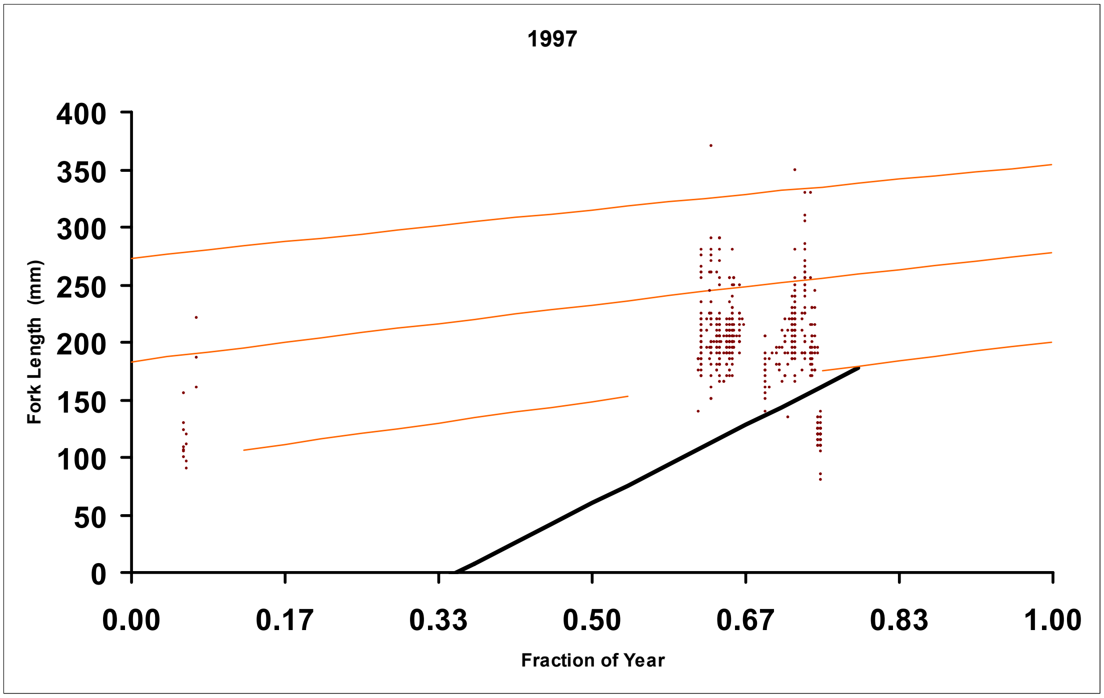


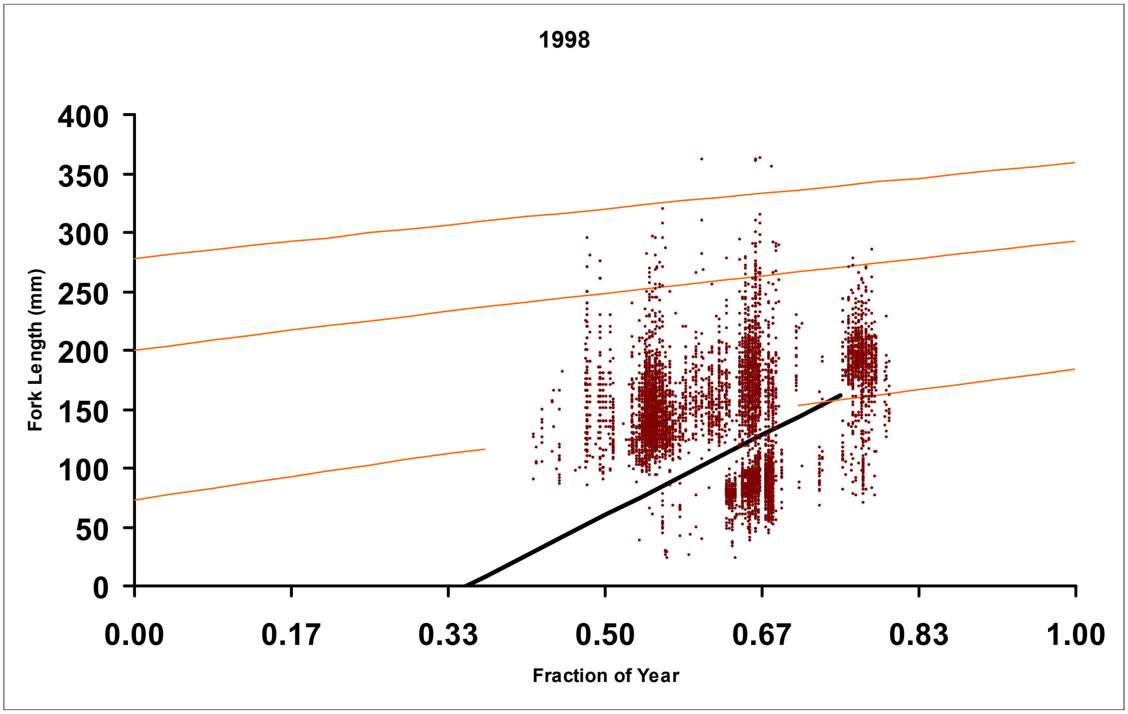


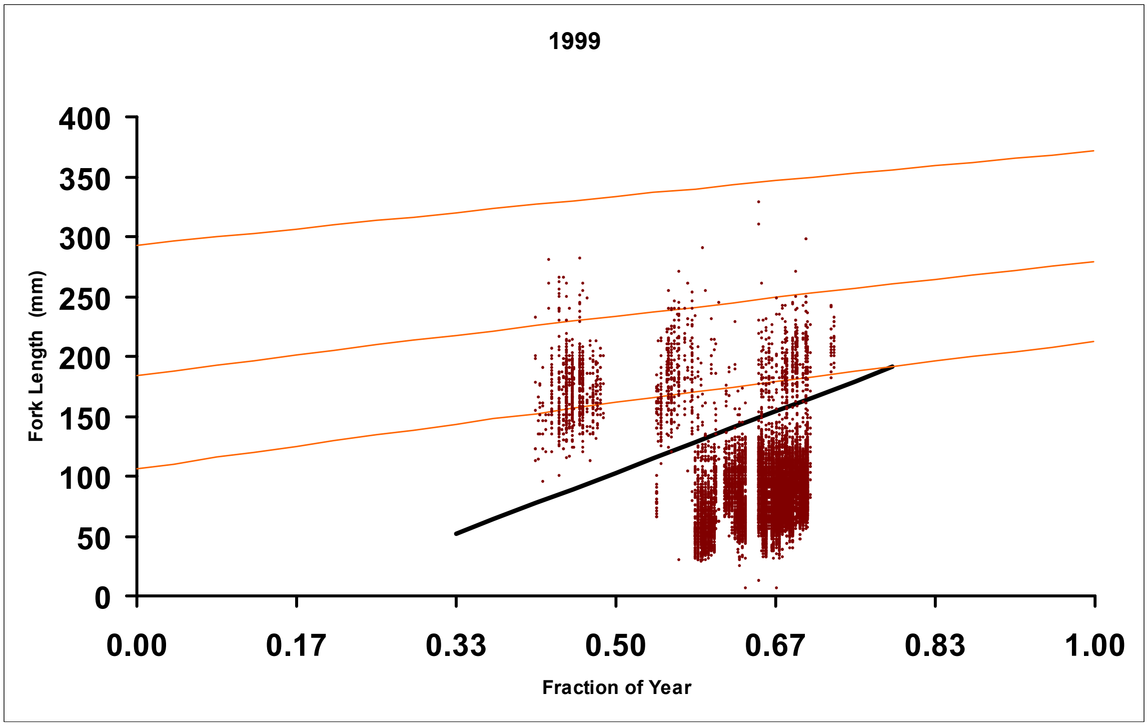

1994
Fork Length (mm)
400
350
300
250
200
150
100
50
0
0.00
0.17
0.33
0.50
0.67
0.83
1.00
Fraction of Year

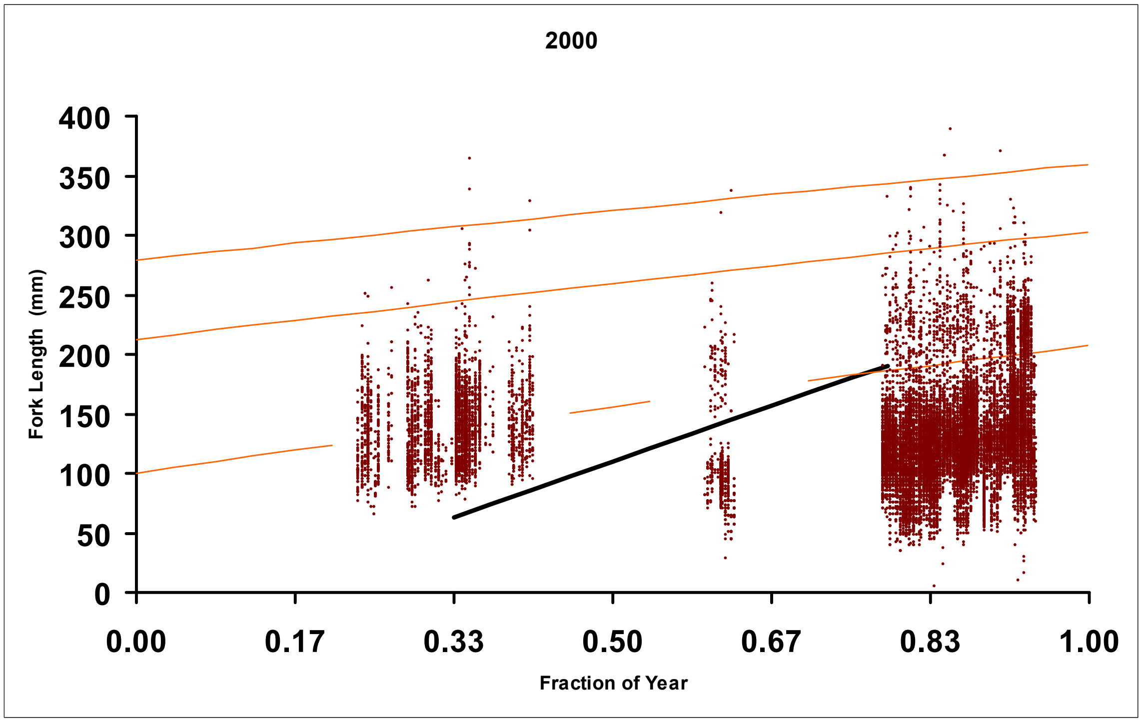

1995
Fork Length (mm)
400
350
300
250
200
150
100
50
0
0.00
0.17
0.33
0.50
0.67
0.83
1.00
Fraction of Year

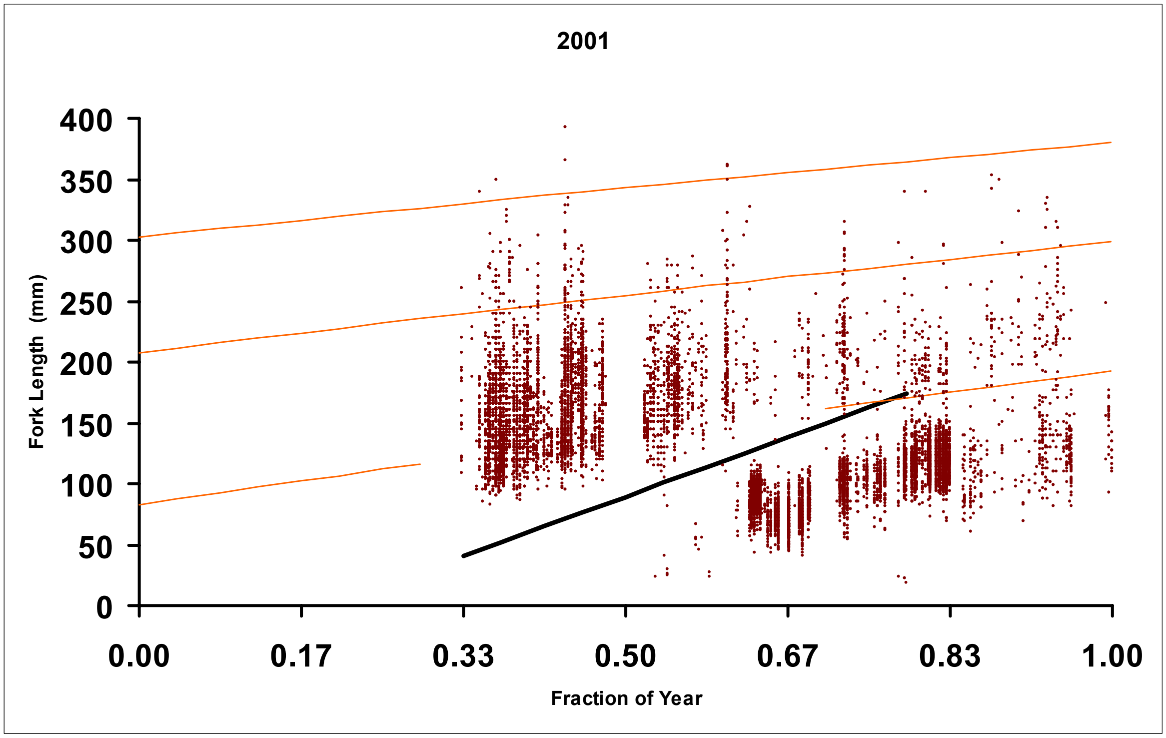

1996
Fork Length (mm)
400
350
300
250
200
150
100
50
0
0.00
0.17
0.33
0.50
0.67
0.83
1.00
Fraction of Year

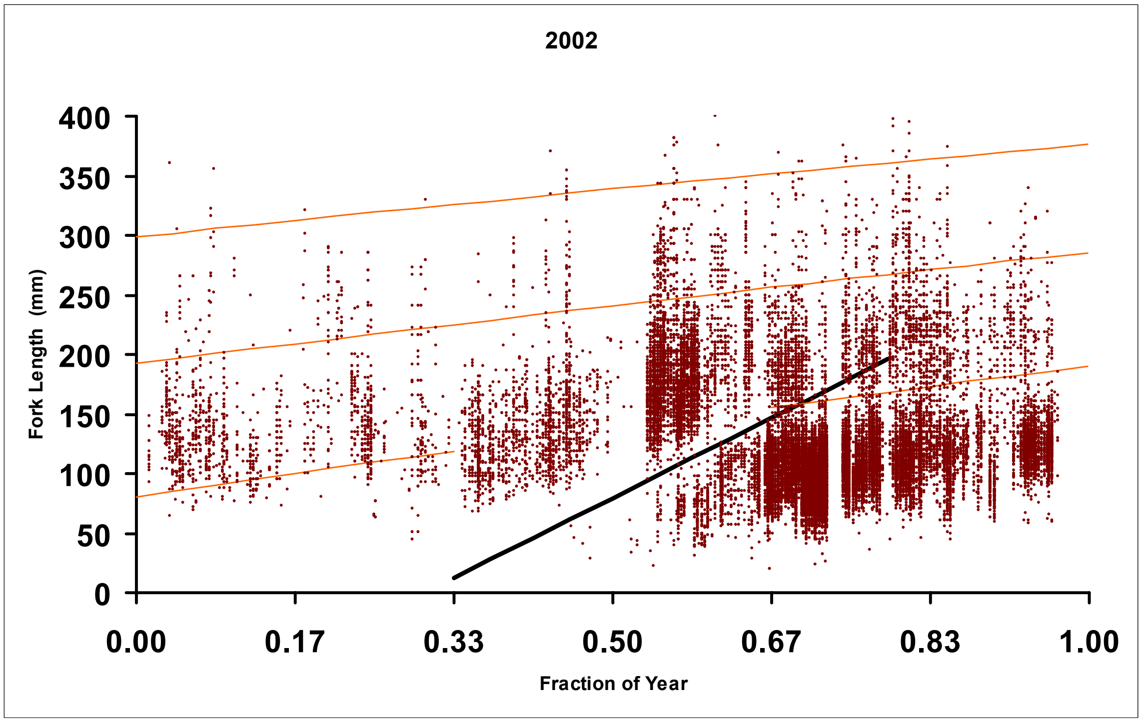

1997
Fork Length (mm)
400
350
300
250
200
150
100
50
0
0.00
0.17
0.33
0.50
0.67
0.83
1.00
Fraction of Year

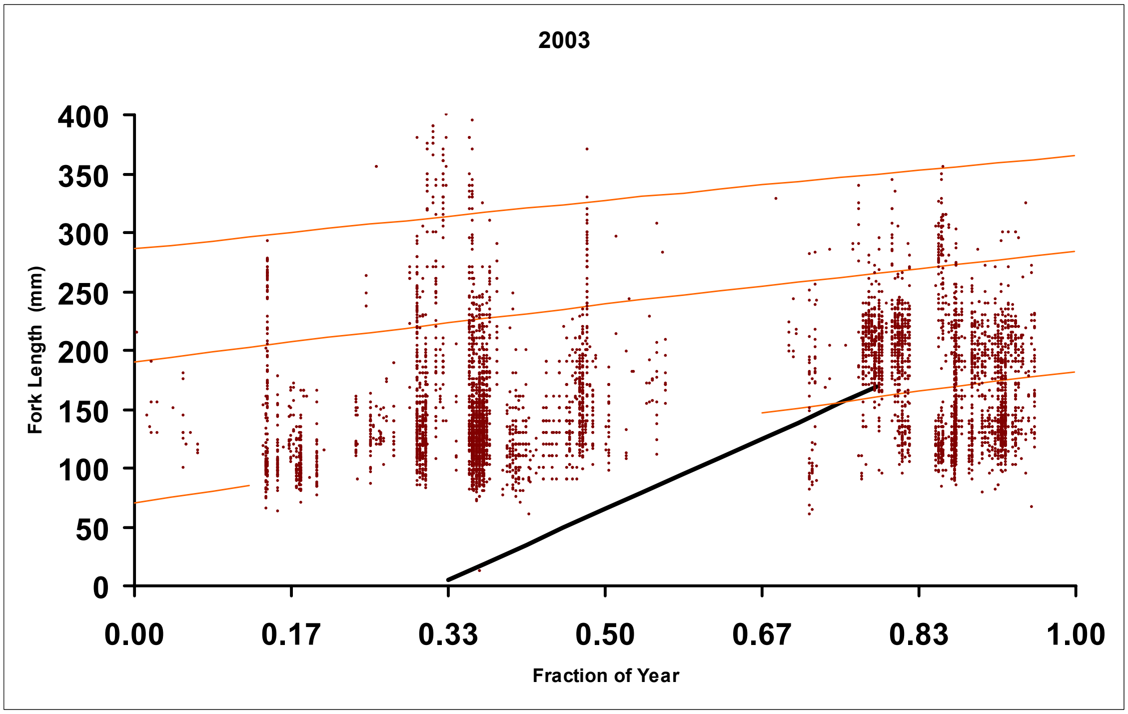

1998
400
350
300
250
200
150
100
50
0
Fork Length (mm)
0.00
0.17
0.33
0.50
0.67
0.83
1.00
Fraction of Year


1999
400
350
300
250
200
150
100
50
0
Fork Length (mm)
0.00
0.17
0.33
0.50
0.67
0.83
1.00
Fraction of Year


2000
Fork Length (mm)
400
350
300
250
200
150
100
50
0
0.00
0.17
0.33
0.50
0.67
0.83
1.00
Fraction of Year


2001
Fork Length (mm)
400
350
300
250
200
150
100
50
0
0.00
0.17
0.33
0.50
0.67
0.83
1.00
Fraction of Year


2002
400
350
300
250
200
150
100
50
0
0.00
0.17
0.33
0.50
0.67
0.83
1.00
Fork Length (mm)
Fraction of Year


2003
400
350
300
250
200
150
100
50
0
0.00
0.17
0.33
0.50
0.67
0.83
1.00
Fork Length (mm)
Fraction of Year

Appendix 3. BUGS program to perform the catch at age calculations.

```
model bycatchbyage {

for (i in 1:32)  {
for (j in 1:3)    {
  for (k in 1:4)   {
  for (l in 1:2)    {

    bycatch[i,j,k,l]~dlnorm(YMEAN[i,j,k,l],TAU[i,j,k,l])
    agecomp[i,j,k,l,1:3]~dmulti(fraction[i,j,k,l,1:3], nn[i,j,k,l])    # these will be 1:4 for ann3 version

    for (m in 1:3)  {
                #  reset for other ann version
      agefract[i,j,k,l,m]<-agecomp[i,j,k,l,m]/nn[i,j,k,l]
      catch[i,j,k,l,m]<-bycatch[i,j,k,l]*agefract[i,j,k,l,m]
      }
  }}
  for (m in 1:3)  {
    tritotal[i,j,m]<-sum(catch[i,j,,,m])
    tritolog[i,j,m]<-max(tritotal[i,j,m],exp(-99))   # will make mean, sd useless, but quantiles should be ok
    logtri[i,j,m]<-log(tritolog[i,j,m])
    }
 }}
for (i in 1:32)  { anntotal[i]<-sum(catch[i,,,,])  }
}
```