

Estimation of Juvenile M for Red Snapper Based on SEAMAP Survey Data

Scott Nichols, Gilmore Pellegrin Jr & G Walter Ingram Jr
NMFS Pascagoula

Introduction

The SEDAR7 data workshop was not successful in reaching consensus on a range for juvenile natural mortality rate (M for age 0 and 1). The best opportunity to do so now probably rests with further analysis of the SEAMAP trawl survey data and shrimping effort estimates. The most recent but still preliminary try was reported in SEDAR7-DW-2. The results were not very promising, being limited by the accuracy of the Z estimates and the minimal contrast in the effort data examined. However, that analysis did not attempt to age the survey data (which reduced reliability); and limited the analysis to the winter /spring season (thus with little effort contrast among years).

Size composition files from the trawl surveys (SEDAR7-DW-16) were being developed at same time as the Bayesian index analyses (SEDAR7-DW-2) were. Now that both products are available together, estimating age from size has been investigated, and it does appear feasible to estimate age from size for ages 0 and 1. Red snapper is a relatively abundant species in the SEAMAP trawl surveys. The surveys cover a major portion of stock, and use stratified random sampling. Given the apparent seasonal patterns of recruitment and escapement, the surveys may allow reliable estimation of Z fall to summer (end of age 0 to age 1), and summer to fall (age 1). By considering both seasons in single analysis, one can obtain a contrast in shrimping effort rates not available in single season analyses. There is a cost for the combined analysis: one must be willing to assume the same M and q hold both seasons. We have tentatively made that assumption, and report a statistical distribution describing our estimate of M, essentially for calendar age 1.

Methods

Details of the methods of the SEAMAP surveys were covered in SEDAR7-DW-1. Seasonal indexes of abundance (not aged) derived using a Bayesian approach were explained in SEDAR7-DW-2. Size data collection and pre-processing were covered in SEDAR7-DW-16. For this analysis, we included only those data from summer 1987 through summer 2003. For these years, both the summer and fall surveys used the complete SEAMAP design. In previous years the summer survey was less intensive, and had some design differences from the current SEAMAP design. The fall survey in earlier years was very restricted in space, covering only the immediate north central Gulf on a continuing basis. Also, although comprehensive size composition sampling was begun around 1985, size data prior to 1987 are considered less reliable than those collected since. We decided that these earlier surveys could not provide reliable estimates of Z to the level required for this analysis.

Shrimping effort estimates, as described in SEDAR7-DW-24, were obtained from NMFS Galveston Lab. We used the effort estimates from 'original' level of resolution (monthly x 21 alongshore areas x 10 offshore depth zones), so we could parallel Goodyear and consider effort only outside 5 fm. The Galveston data base now includes confidence intervals on effort, but the precision for effort is so tight compared to the variances found for the Z estimates that we elected to simplify our analysis by treating the point estimates of effort as if they were known without error. We did not make any adjustments to the nominal effort estimates supplied by Galveston, and simply summed the values, seasonally, over the range approximating the survey area (zones 10-21, 5-50 fm).

The size composition data introduced in SEDAR7-DW-16 were summarized to provide a single vector of size composition for each entire survey. The field sampling protocol establishes a maximum number of individuals to be measured at each station, so each station's size composition vector was weighted by the total catch per hour of red snapper caught at the station, and the resulting vectors were summed. We ignored any missing stations, or missing samples from otherwise successful stations in these estimates of size composition for the population.

Each survey's size composition vector was reinterpreted as an age composition vector by establishing age boundaries by eye. Histograms for each survey generally showed a clear gap between age 0 and 1, or at least a range of very sparse representation between two peaks (Figures in Appendix 1). The distributions differ enough among years to require separate boundaries each year. The distributions also have some variable skew, so parameterized fitting techniques to extract modes seemed less practical than simple inspection. Modes higher than age one were not apparent in the survey data. We retained the convention originally used by Goodyear: everything greater than the 0/1 boundary up to 300mm was presumed to be age 1. Everything above 300mm was considered 2+. Each mm bin was assigned an age based on the boundaries selected for each cruise (Appendix 1 Table 1), and the quantities in each bin were summed by age to produce estimates of the fraction of the population in each of the 3 age groups: 0, 1 and 2+.

We modeled each result as a multinomial distribution: one vector of age fractions per survey, with sample size set to the number of fish actually measured during the survey (Appendix 1 Table 2). This is a simplification from the sampling protocol, in that it treats each fish measured as if it were simple random selection from stock. As the variability is sure to be dominated by the total catch in numbers per unit effort, it seems unlikely that this approximation is limiting.

Age-specific abundance indexes were obtained by multiplying each age fraction vector by the CPUE value of the SEAMAP index from SEDAR7-DW-7 for each survey. The natural logs of the ratios of successive age-specific indexes for a cohort are used to derive estimates of year and season estimates of Z , which is tentatively interpreted as total mortality rate. The Z estimates are then regressed against values of shrimping effort occurring over the same periods. The slope will be an estimate of q , the catchability coefficient, the intercept on the Z axis will be an estimate of juvenile M . The surveys were treated as instant measurements taken on July 1 and November 1 without correction for their actual durations. The time intervals between surveys are of different length (4 month "summer;" 8 month of "winter/spring"), so both Z and effort rate are expressed on a 'per month' basis (dividing the logs of ratios and sums of efforts by 4 or 8). The Z 's are thus matched to sums of effort for the July through October and November through June of the next year. (By convention, the November through June Z 's and efforts are assigned a year designation based on the November date in the figures for this document.) Recruitment of young of year takes place during the summer, usually with most of it occurring after the summer SEAMAP survey. Recruitment is presumed complete prior to the fall SEAMAP survey. The first Z available for a cohort is thus the November 1 to July 1 period, as late age 0s to mid-age 1s (under the Goodyear calendar date convention). The second Z covers the July 1 to November 1 period (part of age 1). A third Z for a cohort cannot be derived reliably, as a mode for age 2 in the summer survey could not be discerned. With 16 years and 2 seasons, there are 32 points of Z and effort estimates available. As the November 1 to November 1 period of the two Z intervals is not far removed from the calendar age 1 designation (January through December), it seems acceptable to designate the derived M as an estimate of M at age 1.

The analyses sketched out in the preceding paragraph were programmed in BUGS software, the same software package used in the derivation of the Bayesian indexes of abundance in SEDAR7-DW-2. In this case, however, the Z vs f analysis is not in itself Bayesian – BUGS was just convenient vehicle for estimating statistical distributions of extended functions of input variables. The CPUEs and age compositions, and thus the Z 's described in the preceding paragraph actually enter the analysis as full statistical distributions. This approach allows our knowledge of the uncertainties of the inputs to be carried forward into the estimate of M in a very convenient manner. The process is actually a simulation program, as follows: at each iteration, draw CPUE index values (one for each survey) from the lognormal distributions reported in SEDAR7-DW-2, draw age composition vectors (again, one for each survey) from the multinomial distributions described in this paper, and multiply. Within each iteration, calculate every Z as the log of the ratios of successive abundances. Then, calculate a regression of Z on shrimping effort using the standard formula for single-variable regression. Over iterations, accumulate estimates of M , q , and various summary statistics. The actual BUGS program appears in Appendix 2. The program was run for 5500 iterations with 2 chains. We discarded first 500 iterations from each chain, in unlikely event that an unusual starting point (selected by BUGS program) influenced the results.

Results

Figure 1 shows the Z estimates in BUGS box graph format. We note immediately that two of the Z's have medians below zero (summer 1998 and summer 2001), although in both cases their distributions extend above zero, and indeed the upper portions overlap the lower portions of the confidence intervals for several summer Zs. Successive Z's are not independent, so we must expect that at least one of the two fall Z's neighboring each of the summer values with medians less than zero will be suspect as well. By inspection, the Z for fall 2000 is seen to be the highest fall Z, so the summer 2001 CPUE at age 1 suspected, but neither the fall 1987 nor fall 1998 Z's appear unusual. We concluded that although negative Z's are physically impossible, occasionally encountering such values is not out of line with the level of variation observed. We continued the analysis including the nominally negative Z's, but added an evaluation by running the analysis without them for sensitivity purposes.

Figure 2 shows the Z estimates plotted against shrimping effort rate. The fall and summer points separate completely on the effort axis, but overlap on the Z axis. A line based on the median results for M and q is shown. The distribution and line are clearly consistent with the relationship expected, but we also see enough scatter in Z among points close together on the effort axis to warn that the M estimate will not be very precise. The scatter among the Z values seems consistent with the precision descriptions for the individual Z's seen in Figure 1. Looking for pattern among the points within the scatter did not turn up anything intriguing.

Figure 3 shows the density plots from BUGS for the estimated values of q and M, both based on a monthly time interval. The BUGS quantiles are listed above each graph. Both distributions are seen to be highly symmetrical, and based on approximate agreement between the quantiles and their expectations under normal distributions with the means and standard deviations shown in the quantile lists, can be taken to be normally distributed. For convenience, we have also provided a tabulation and density plot for Manual, which is simply the monthly M multiplied by 12.

For comparison, we also ran a simple regression using SAS Proc Reg, using the median values for the Z's. The fitted lines are virtually identical, but the confidence interval on M is larger (s.e. 0.06 from SAS vs 0.03 from BUGS).

SAS Proc Reg was also used to examine the fit with the two Z's with negative medians removed. (It was quicker to knock out the offending points in SAS than in BUGS). As expected, the point estimate for M was quite a bit lower (about 0.1 per year), and the standard error on the intercept was reduced about 25%. A Bayesian interpretation of the confidence interval would put about 40% of the distribution below zero. Confidence intervals from a BUGS calculation would presumably be narrower, but the character would be much the same.

An additional distribution that might prove useful was calculated from the BUGS run: the average of annual Z's for the entire period (1987-2003). This was calculated as the sum of 8 average winter/spring Zs and 4 average summer Zs. The median value returned was 1.98 per year, with a 95% confidence interval 1.92 to 2.04 (Figure 4).

Discussion

This analysis of trawl survey data returned plausible M estimates, but failed to narrow the uncertainty much from a range that might have been considered realistic prior to this analysis. The primary analysis (no Z's set aside) tended to a higher point estimate than used in the past stock assessment, but the confidence interval is very broad, even extending slightly into negative territory. Setting aside the negative Z medians lowers the point estimate of M to a level that is probably lower than generally expected, with an uncertainty such that M=0 would be well supported by the data. There are no strong or independent reasons to reject the negative Z's as anything other than random noise, so the analysis with no Z's set aside should be preferred. However, the main message is probably in the width of the confidence band, not in the central tendency.

The limiting factor on the utility of the M analysis appears to be the precision of the Z estimates. Near neighbors on the effort axis can have very different Z's, but the spread is consistent with the confidence intervals on each Z. The effort contrast appears sufficient to establish a regression relationship.

There is a potential for bias associated with possible reduction in vulnerability to trawling with age. This would be expected to raise the summer Z estimates above their true values, producing an estimate of M lower than the true value. However, there is no evidence of such bias in the survey data (one would expect to see an ellipse imagined around the summer Z's in Figure 2 misaligned with an ellipse in the winter/spring Z's). Shrimp observer data show clear evidence of persistence of the fall age 1 cohort through the winter, and possibly beyond (Nichols 2004). We conclude that serious downward bias in M due to reduction in vulnerability is unlikely. If it exists at all, such bias would seem to be small relative to the confidence interval reported.

A second possible source of bias would be rapidly changing M with age. If M decreased with age but q did not change, this type of bias could lead to an estimate of M that was higher than the true values for both time intervals. If true q changed as well, the direction of bias in M could be unpredictable. In general, violations of the assumption of constant M and q over the two seasons (and over the changing ages) would again be expected to show up as "misaligned ellipses" in Figure 2. No evidence for different sub-ellipses is evident, but the uncertainties of the individual Z's imply little hope for detecting any. The best interpretation may be that the constant M and q assumptions are consistent to the order imposed by the precisions for the Z's and the small effort contrasts among years within each season.

Given the broad confidence interval for M, it is an open question whether M should be considered nearly constant from year to year, or whether some of the source of variability is true variation in M among years. M could be modeled either way in the stock assessment, but there is no evidence here supporting the dominance of either interpretation. Beyond the two seasonal clusters, no subsetting is evident in Figure 2, and there doesn't appear to be any sequential pattern in the scatter. We are limited again by precision of Z, and perhaps additionally by the shortness of the time series.

The confidence interval implied by the distribution of M from the BUGS simulation was smaller than that calculated by the simple regression using median Zs and the standard formula for precision of the intercept. We believe this is a consequence of having separate precisions available for the individual Z's (which is not considered in simple regression), and interdependence of successive Z's. The BUGS confidence interval is derived from a procedure that actually simulates the usual interpretation of a frequentist confidence interval – many conceptual repeats of the same experiment, here done as a computer simulation with draws from the individual CPUE distributions. As the BUGS confidence interval includes extra information not available to the simple regression interval, we consider the smaller BUGS interval more reliable. This is a practical value to choosing the BUGS interval -- very little of distribution of M is in negative M range, which is not the case for the simple regression interval.

For analytical convenience, we did not introduce variance for the estimates of shrimping effort – the effect would be small compared to the spreads on the Z estimates. We also made no adjustments for any possible changes in the 'effectiveness' of shrimping effort, including possible effects of BRDS. The effect of BRDs was estimated to be small relative to year to year and season to season changes in nominal effort even in 1998 (Nichols 1999), and BRD effectiveness appears to have dropped following an unknown path in time since then (SEDAR7-DW-38). The lack of interannual pattern in Figure 2 supports the appropriateness of the simplifying assumptions regarding effort.

We consider two additional analyses, but chose not to complete them. The first considered was a partition of Z's and efforts spatially, in an attempt to enhance the effort contrast. We stopped consideration of that approach upon noting the effort contrast available in the full-survey analysis, and the importance of the low precisions on the Z's. Separate Z's for different geographic regions would be expected to be less precise, and effort would have to be calculated on a 'per unit area' basis, which would imply a large assumption about the local effectiveness of effort. The second analysis considered but dropped involved an alternative set of assumptions about age boundaries, based on a tentative analysis of bycatch observer data being

developed at the same time by Nichols 2004. Nichols (2004) attempted to identify varying upper bounds on age 1 and age 2, but concluded the results were not reliable.

We believe that the most realistic description of our knowledge, and lack thereof, about juvenile M is contained in the original run and its broad confidence interval. We recommend a normal distribution centered on 0.6 per year, with a standard error of 0.36 as the most realistic model for M at age 1, with the restriction that the distribution be truncated at zero if used directly as an assessment input. One might also consider a uniform distribution with bounds at approximately 0.2 and 1 per year as a reasonable alternative.

Literature Cited

SEDAR7-DW-1

Nichols S. 2004. Derivation of red snapper time series from SEAMAP and groundfish trawl surveys.

SEDAR7-DW-2

Nichols S. 2004. Calibration among the separate trawl survey programs to extend the time series for juvenile snapper indexes.

SEDAR7-DW-16

Pellegrin G, Sanders N, Johnson K & DeBose A. 2004. Retrospective coding of dual size classes of size frequency data for red snapper collected during the SEAMAP Shrimp/Bottomfish surveys.

SEDAR7-DW-24

Nance J. 2004. Estimation of effort in the offshore shrimp trawl fishery of the Gulf of Mexico.

SEDAR7-DW-38

Foster D & Scott-Denton E. 2004. Status of bycatch reduction device performance and research in the north-central and western Gulf of Mexico.

Nichols S. 1999. Overall effectiveness of BRDs. Report from the Summer 1998 red snapper / shrimp research program. Report to the Gulf of Mexico Fishery Management Council.

Nichols S. 2004. Estimating catch at age for red snapper in the shrimp fleet bycatch. Report for the SEDAR7 Stock Assessment Workshop

Figures

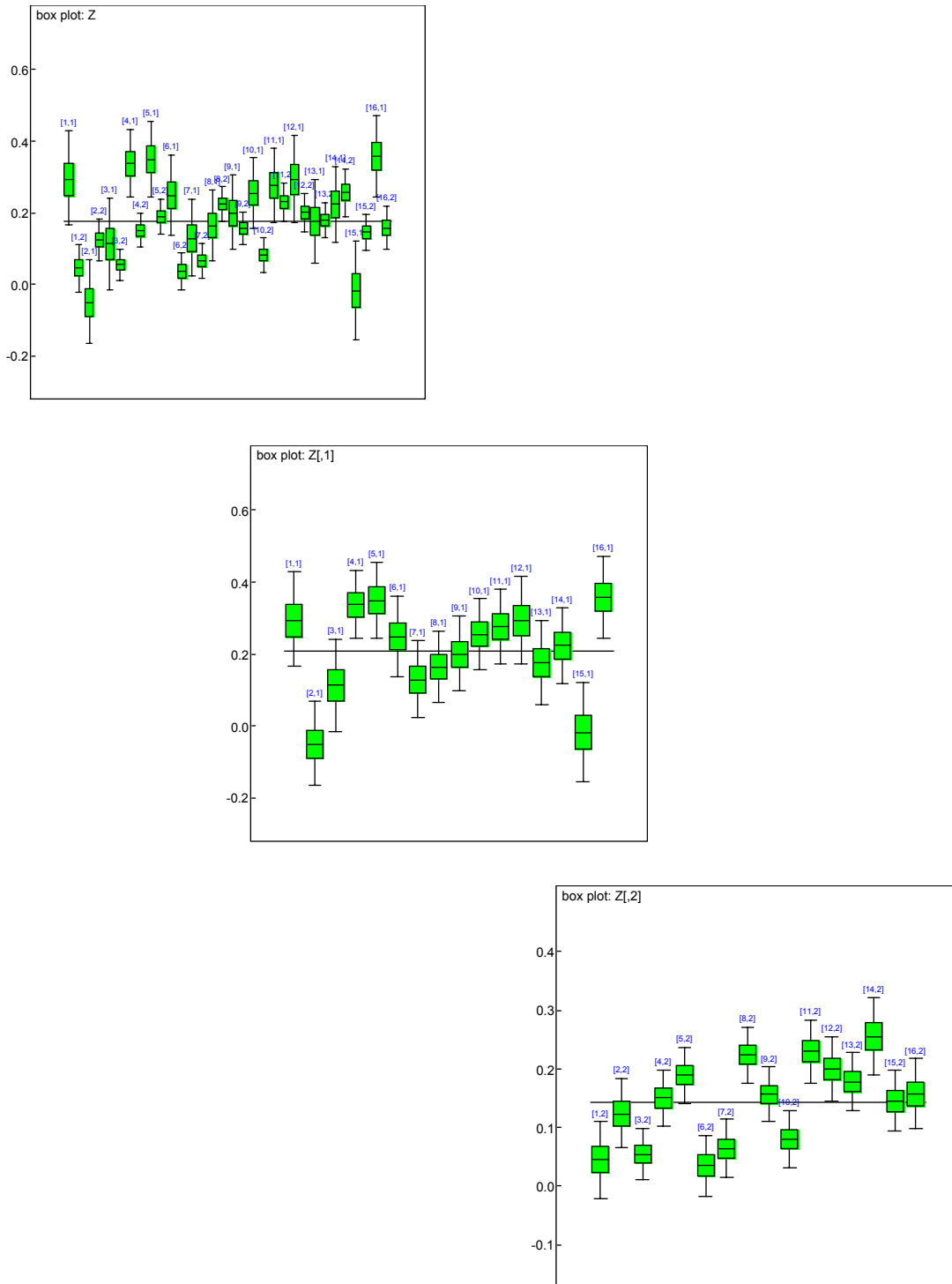


Figure 1. Sequential values of Z returned by ratios of successive age specific CPUEs. Top, all values in one plot. Middle, summer Z (July 1 - November 1). Bottom, winter/spring Z (November 1 – July 1). All units are per month.

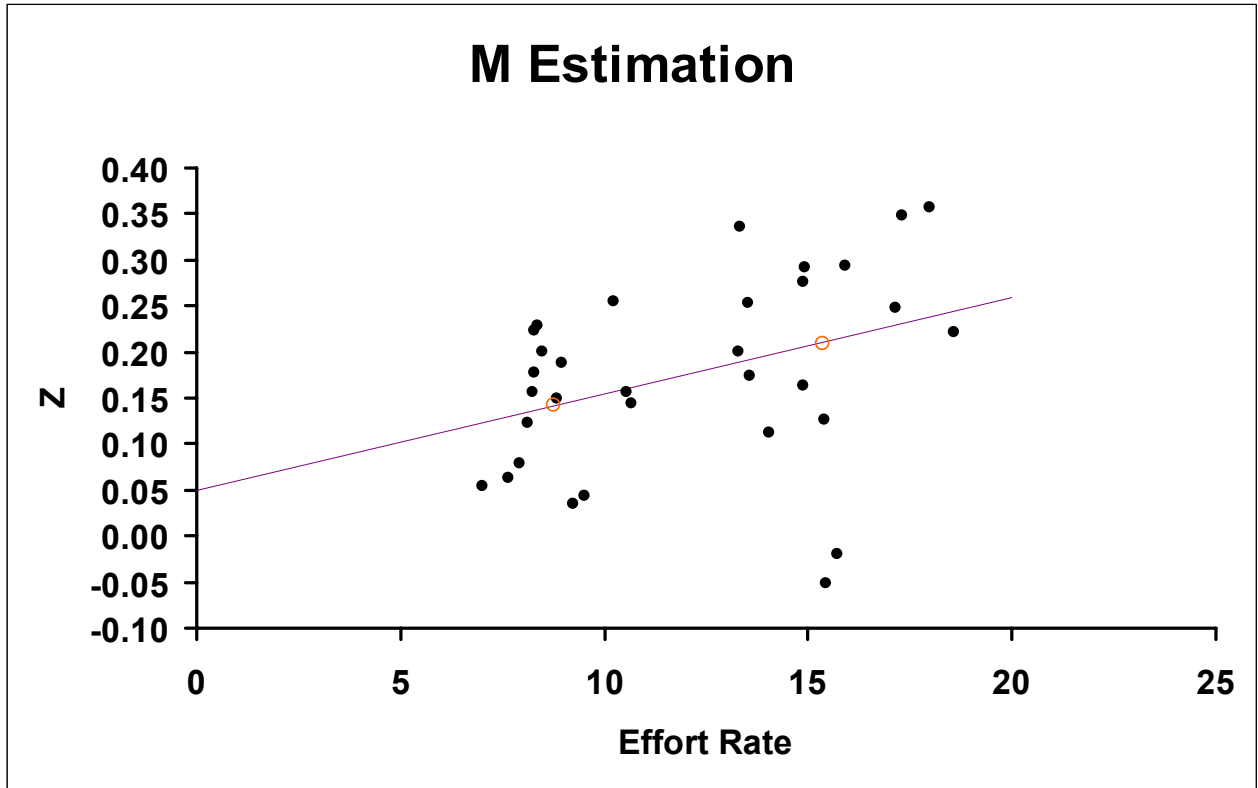
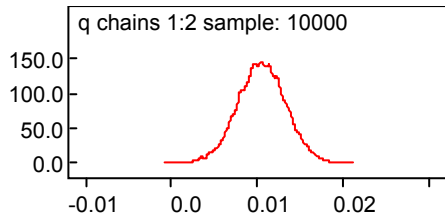
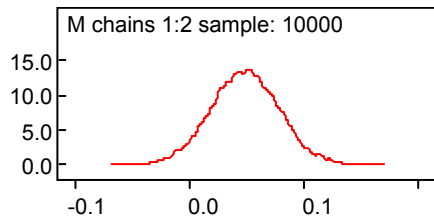


Figure 2. Z estimates vs shrimping effort rate. Clear circles are mean values for summer and winter/spring. Winter/spring and summer points are separate by the gap near an effort rate of 12. Z's are per month, effort rates are thousands of 'days fished' per month.

node	mean	sd	MC error	2.5%	25.0%	median	75.0%	97.5%
q	0.01053	0.002774	2.7E-5	0.004992	0.008665	0.01053	0.01239	0.01598



node	mean	sd	MC error	2.5%	25.0%	median	75.0%	97.5%
M	0.04906	0.02971	2.899E-4	-0.009278	0.029	0.04906	0.06898	0.1083



node	mean	sd	MC error	2.5%	25.0%	median	75.0%	97.5%
Mannual	0.5888	0.3565	0.003479	-0.1113	0.348	0.5887	0.8278	1.3

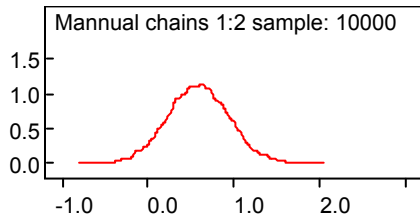


Figure 3. BUGS parameter estimates and distributions for q and M in monthly units, and M expressed as an annual rate.

node	mean	sd	MC error	2.5%	25.0%	median	75.0%	97.5%
Zannual	1.98	0.03179	3.03E-4	1.919	1.959	1.98	2.001	2.043

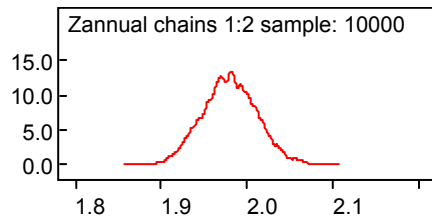
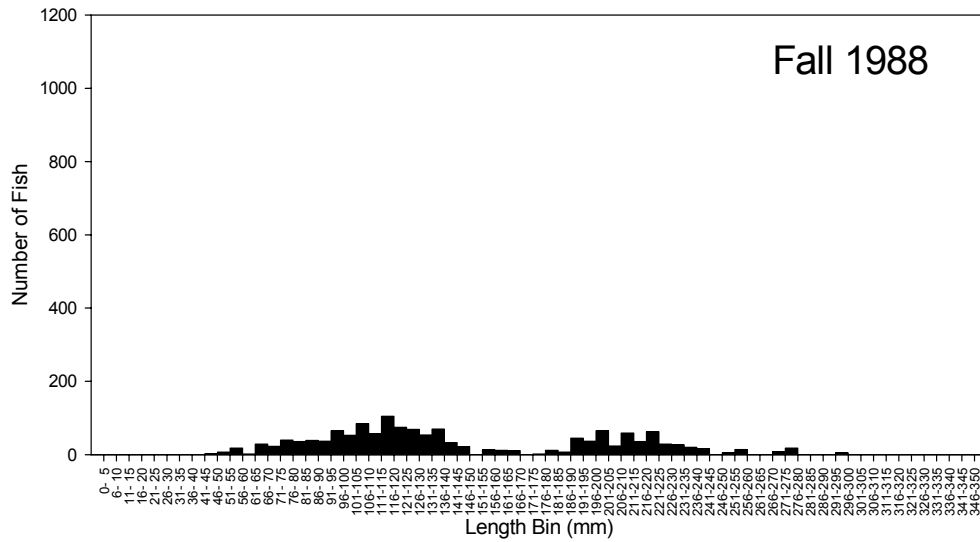
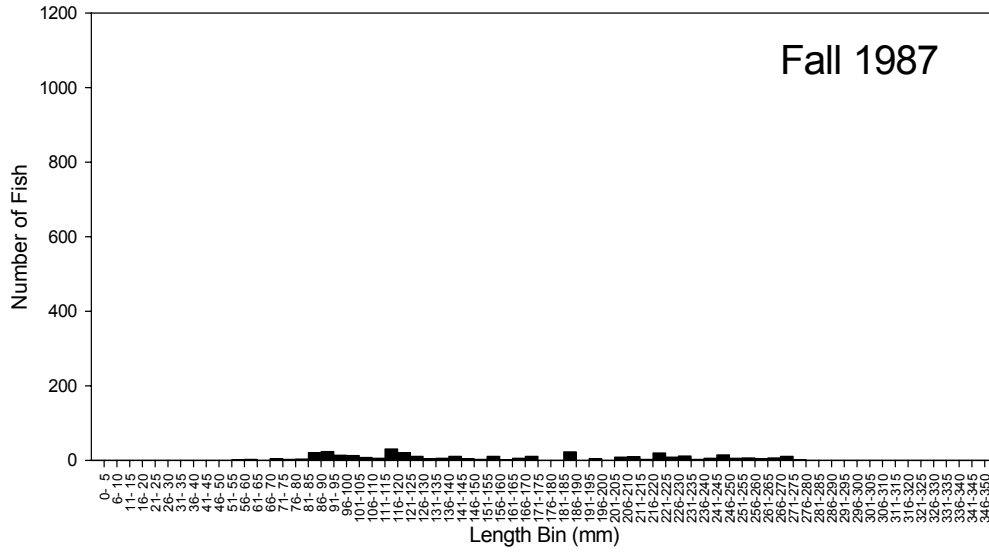
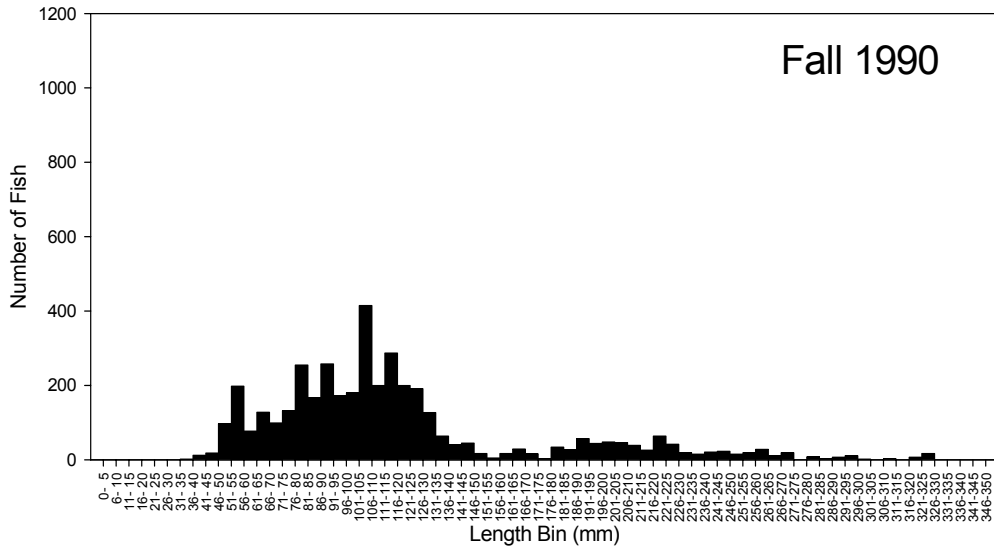
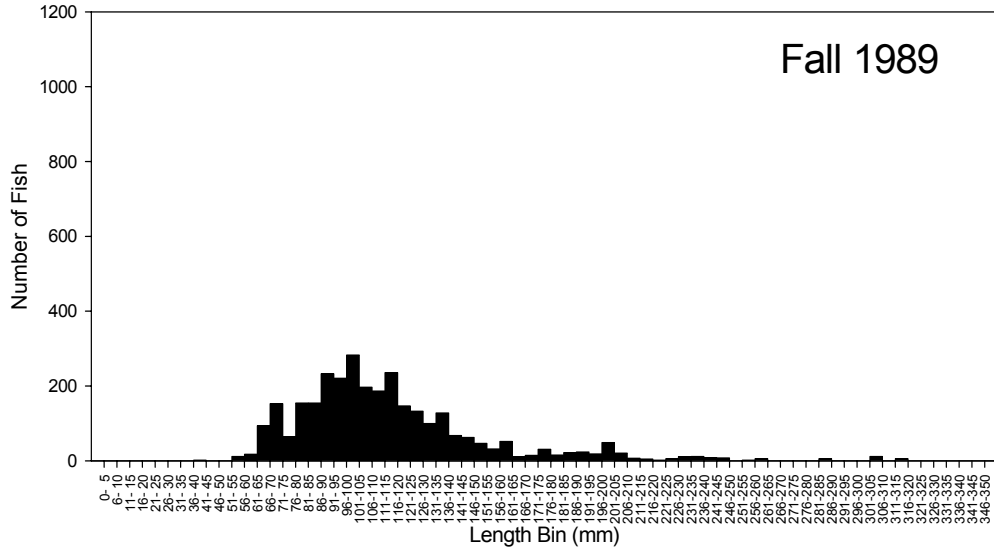
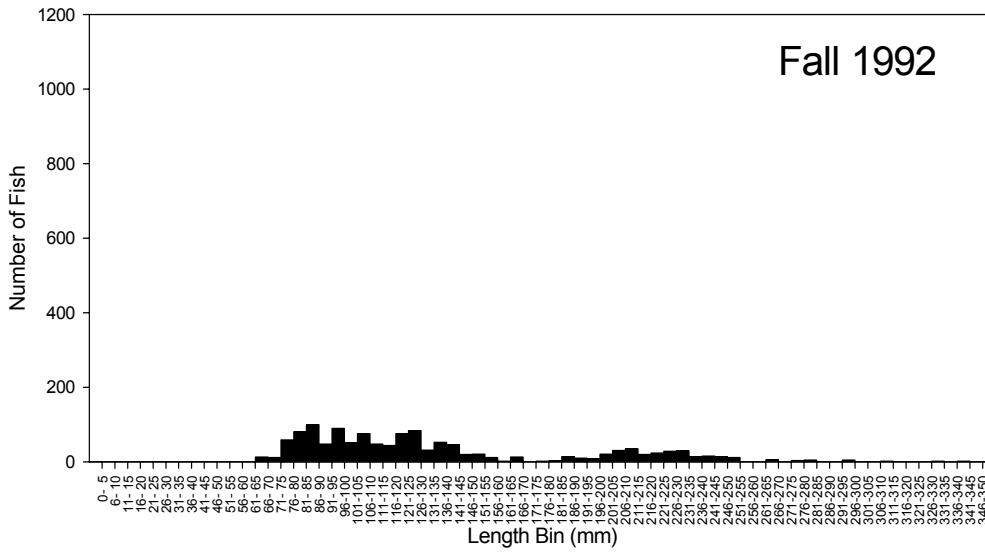
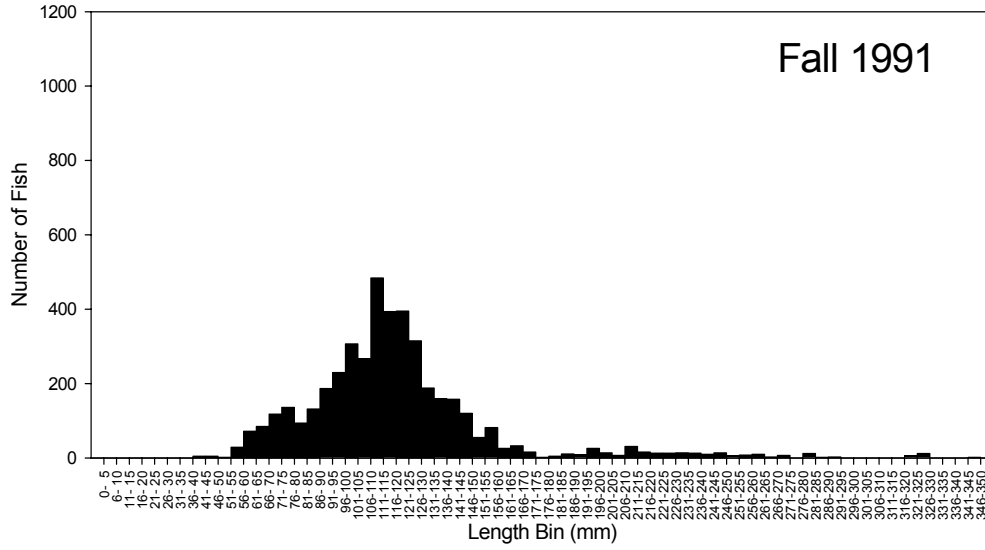


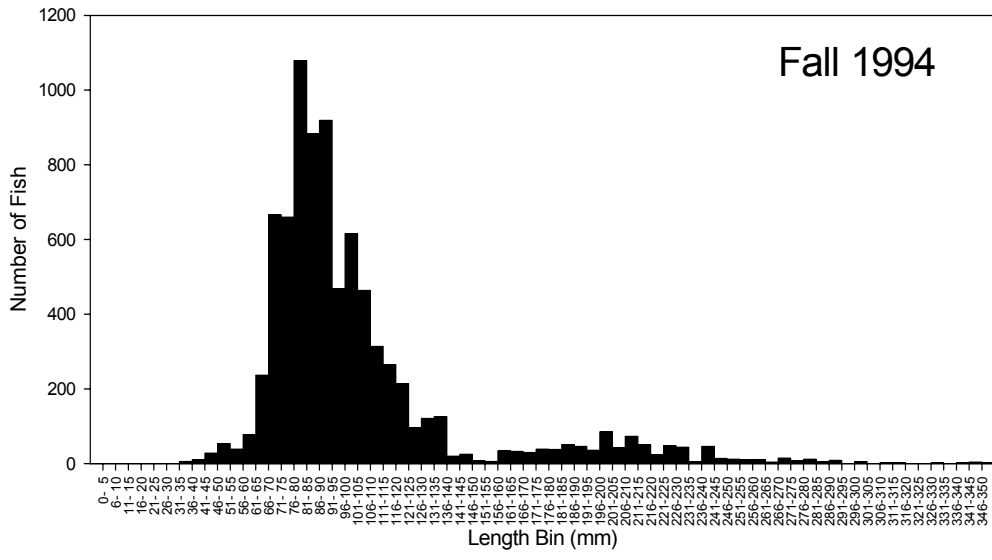
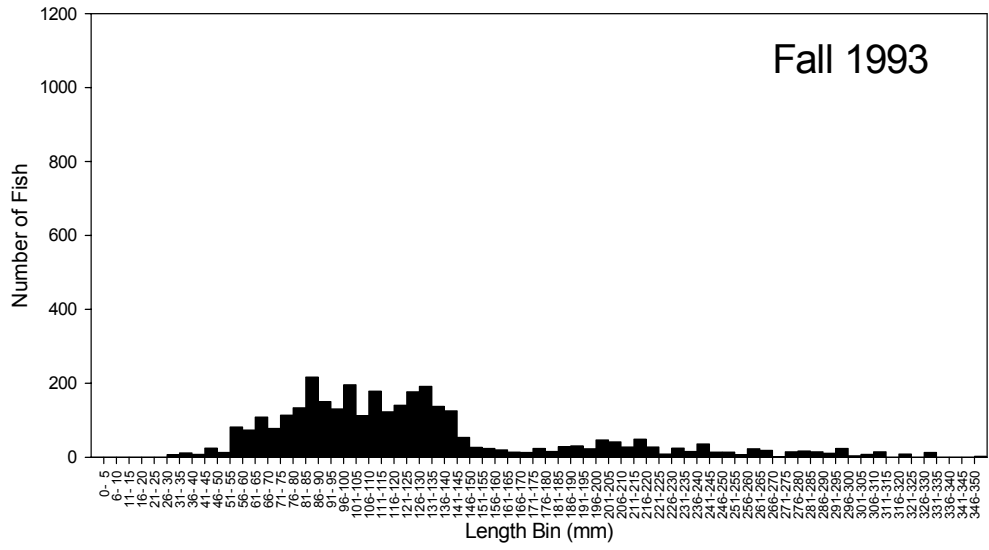
Figure 4. Average value of the annual Z's for the 1987-2003 period.

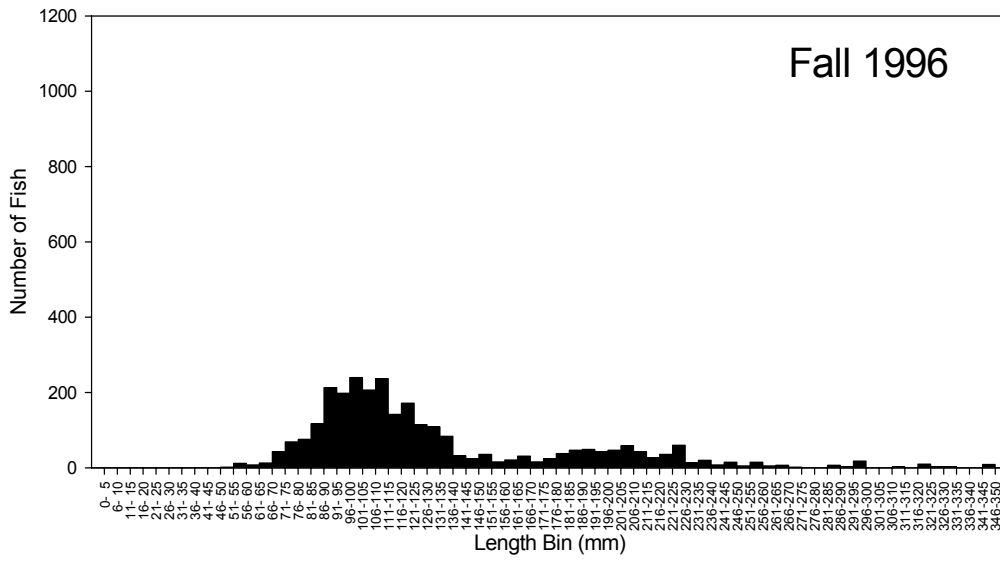
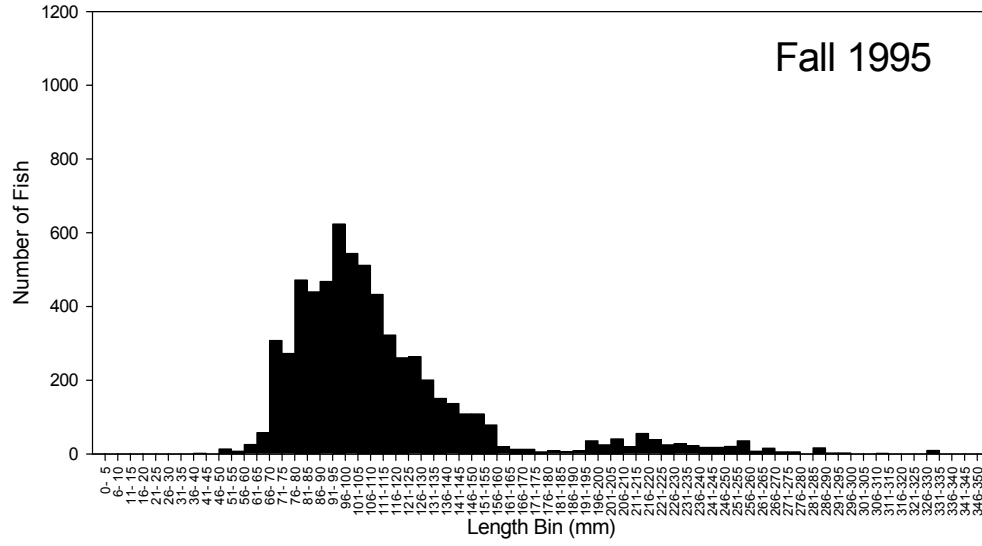
Appendix 1. Histograms of size composition for each survey. Table of age boundaries, and table of age fractions.

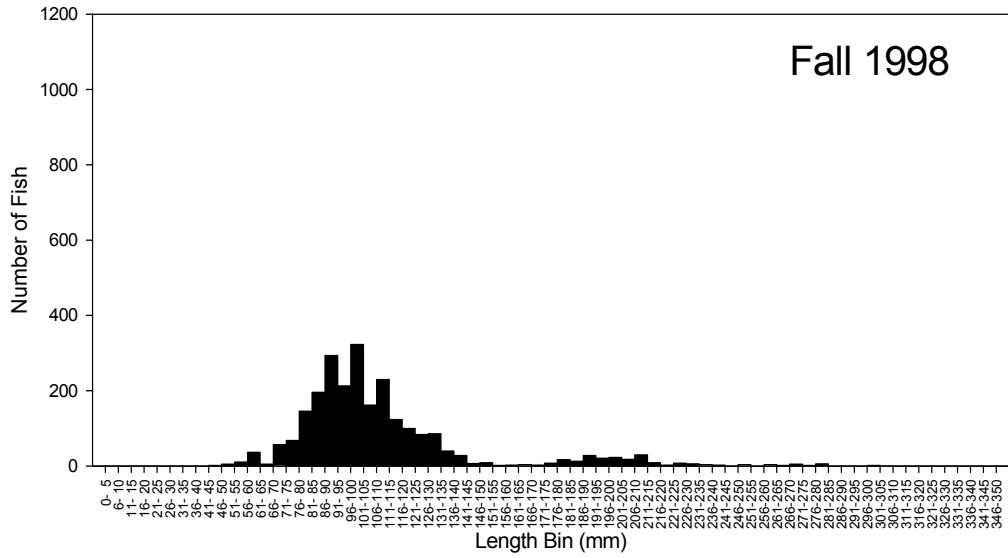
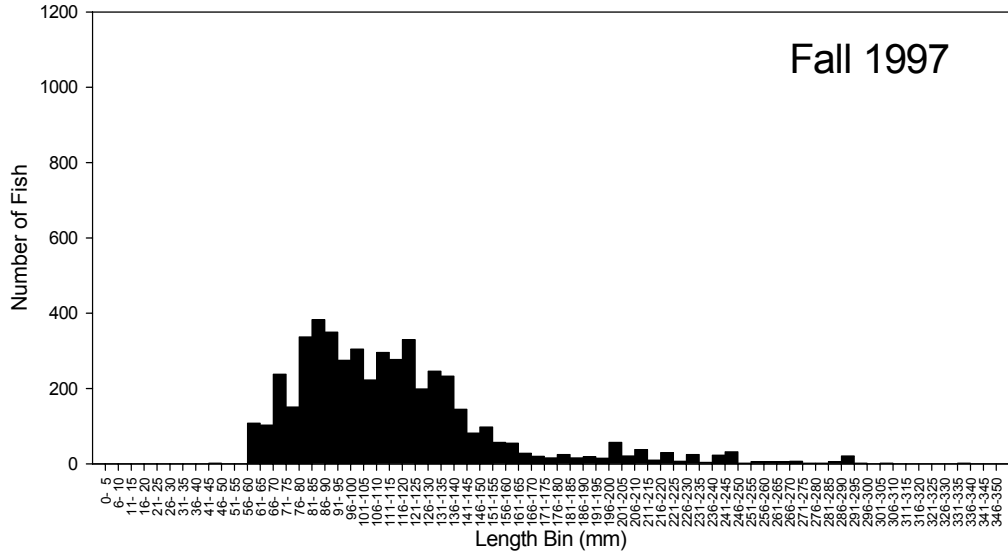


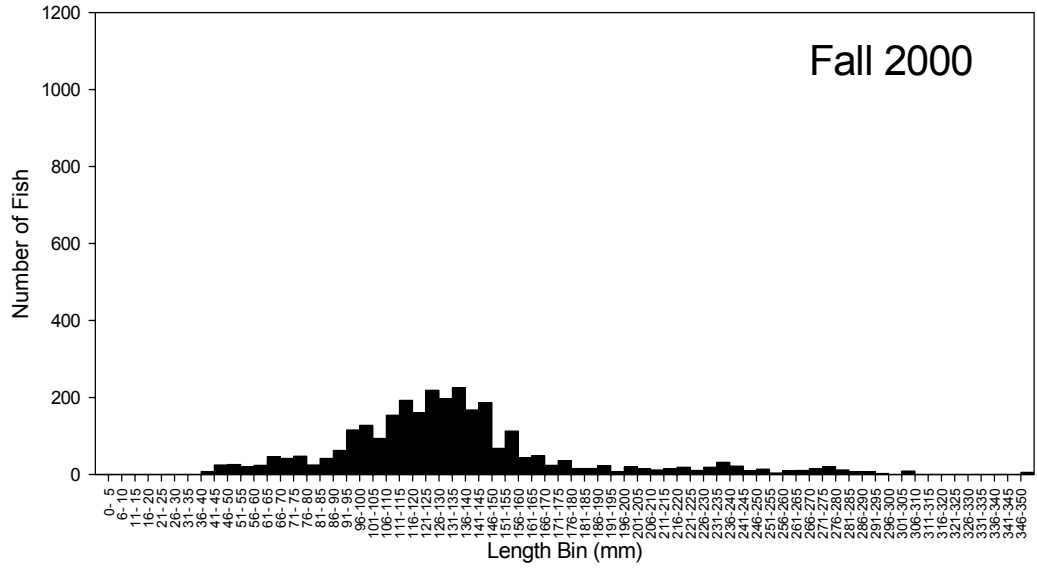
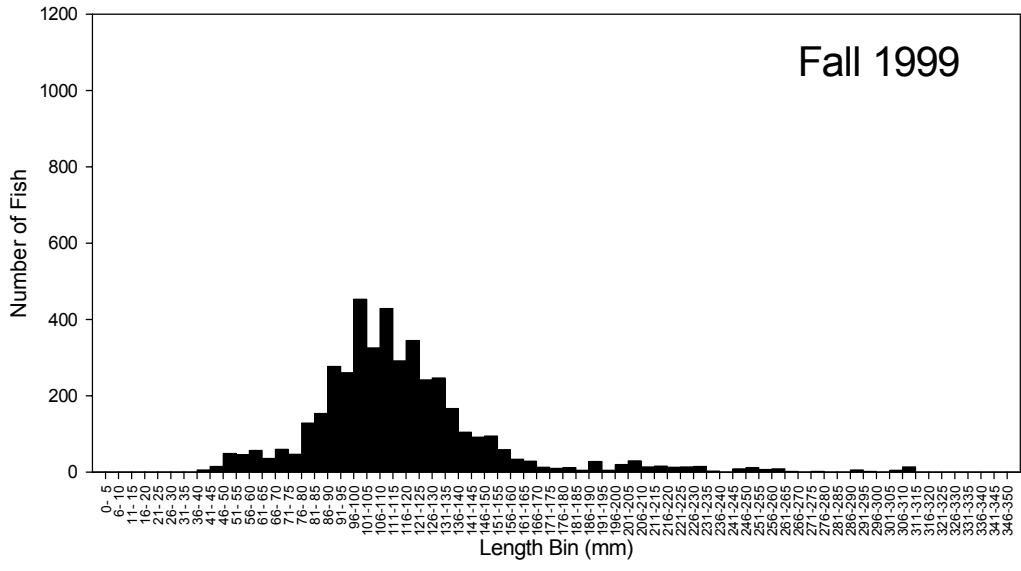


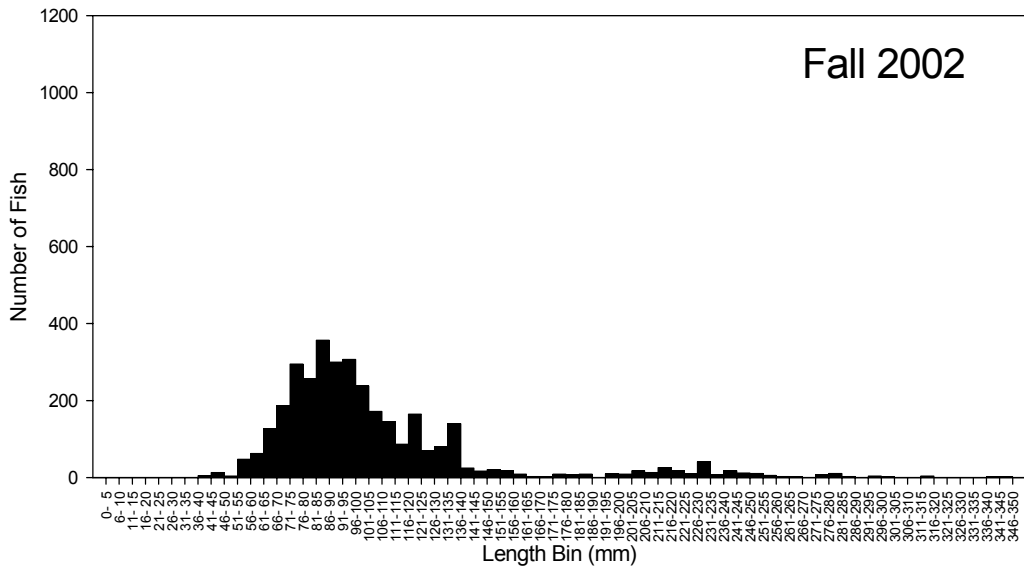
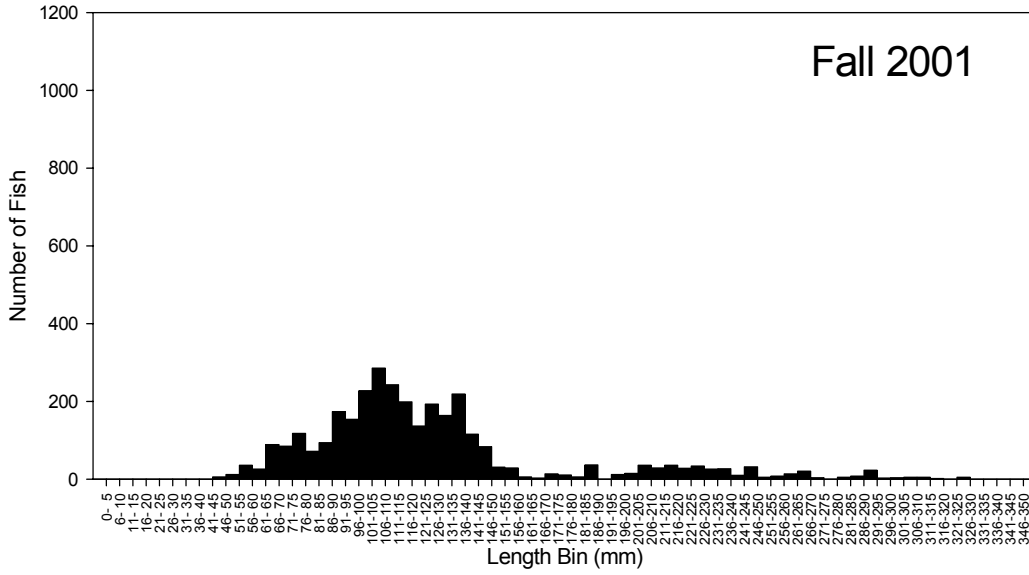


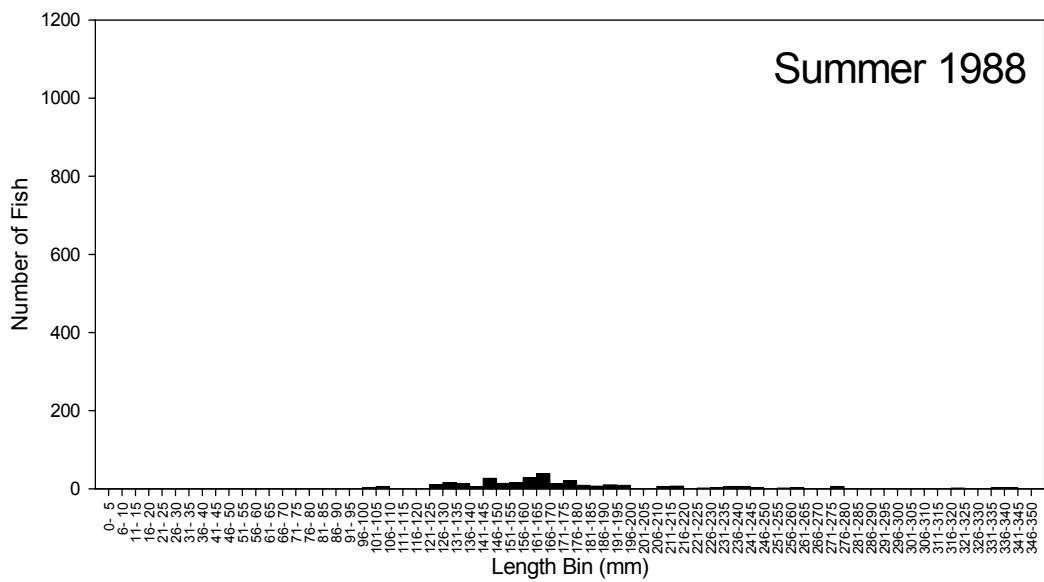
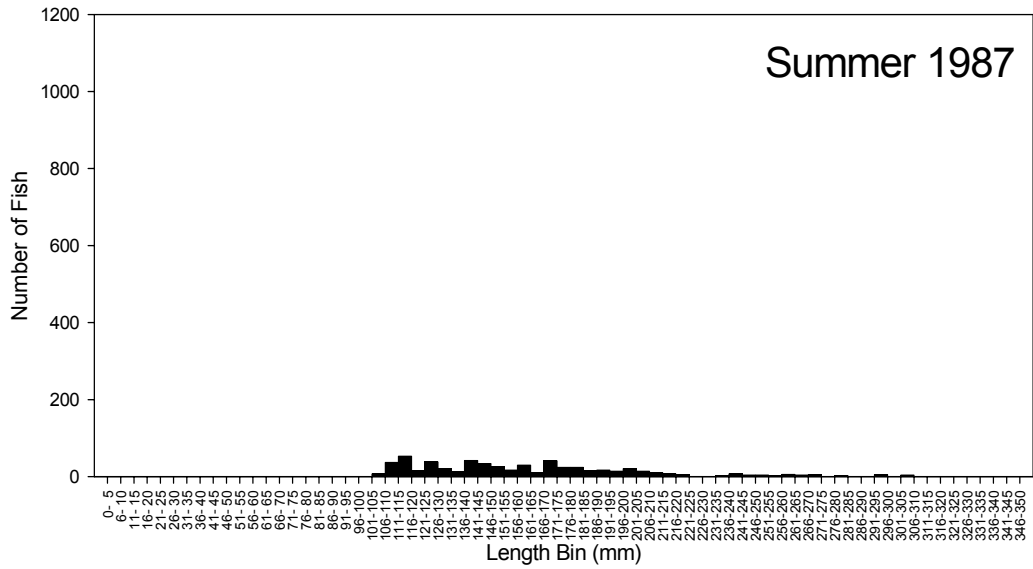


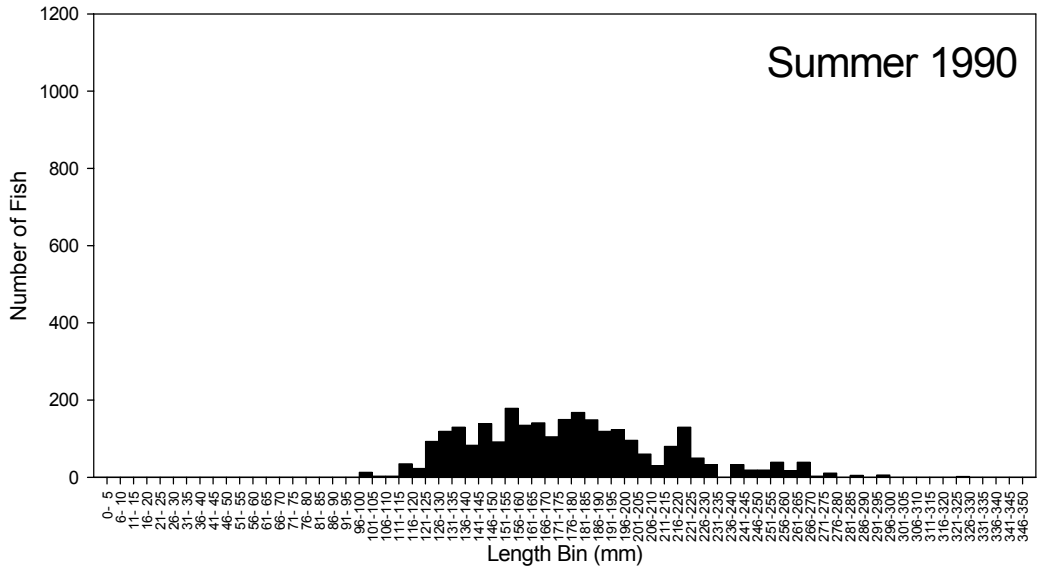
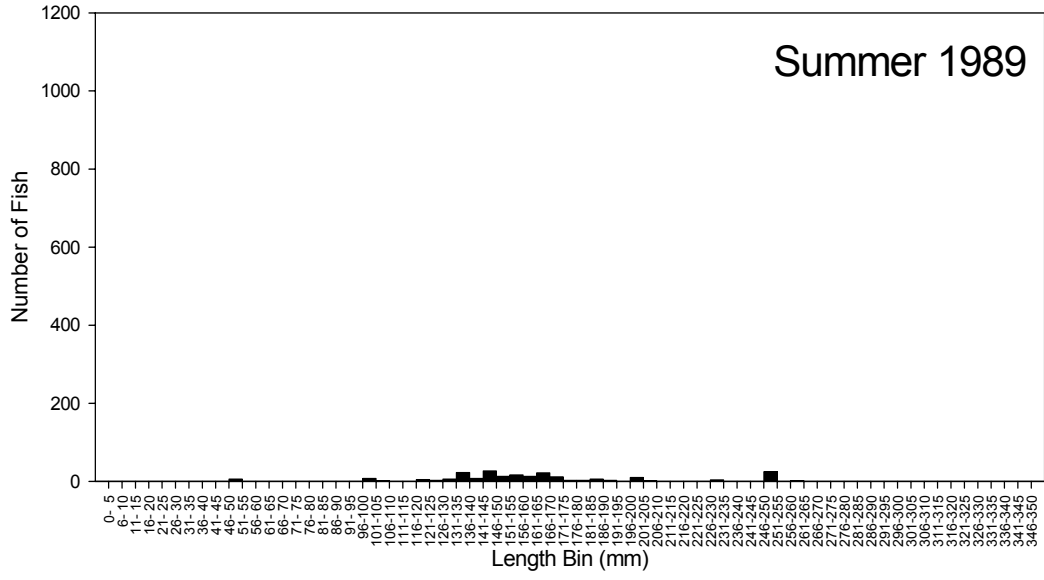


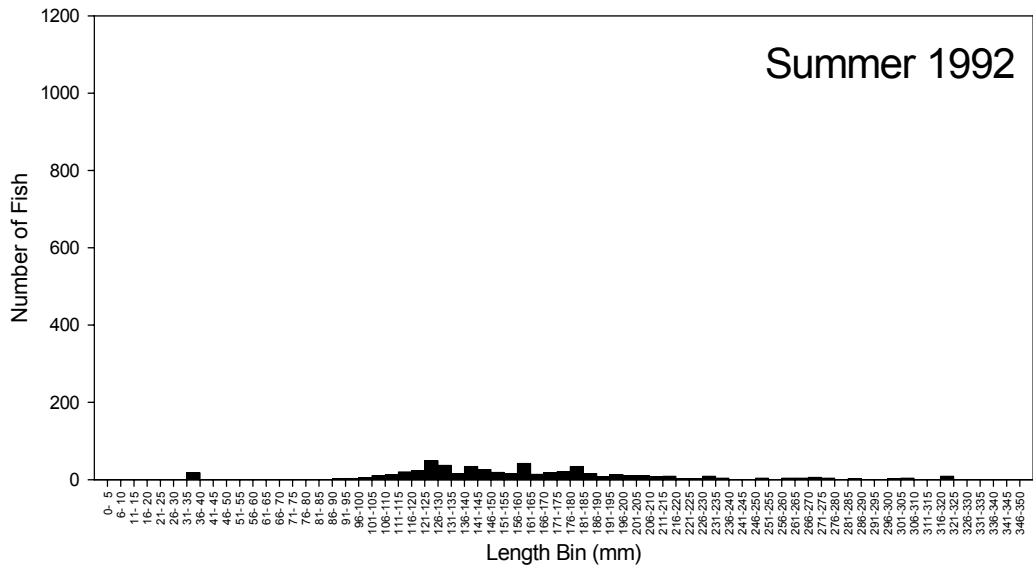
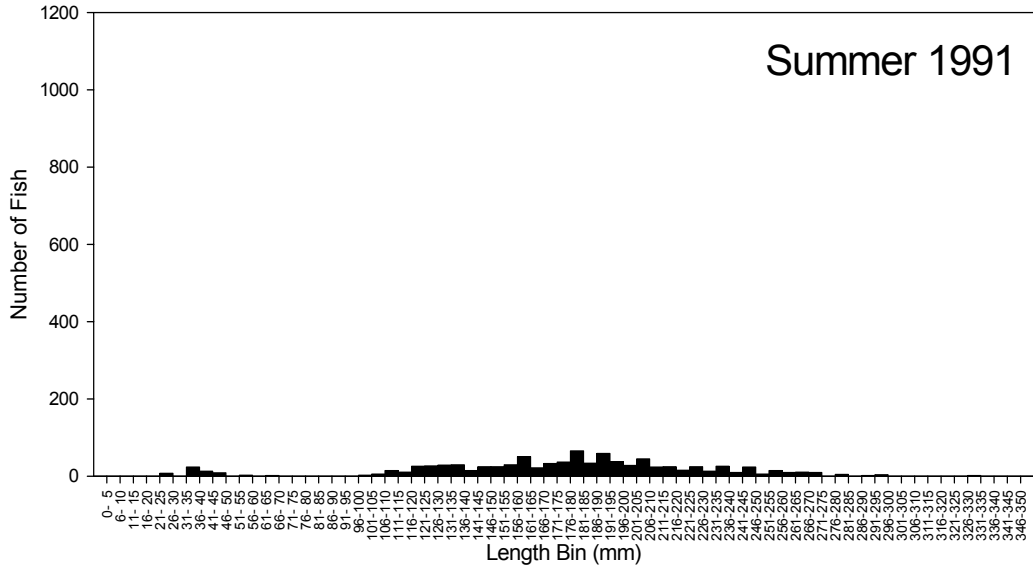


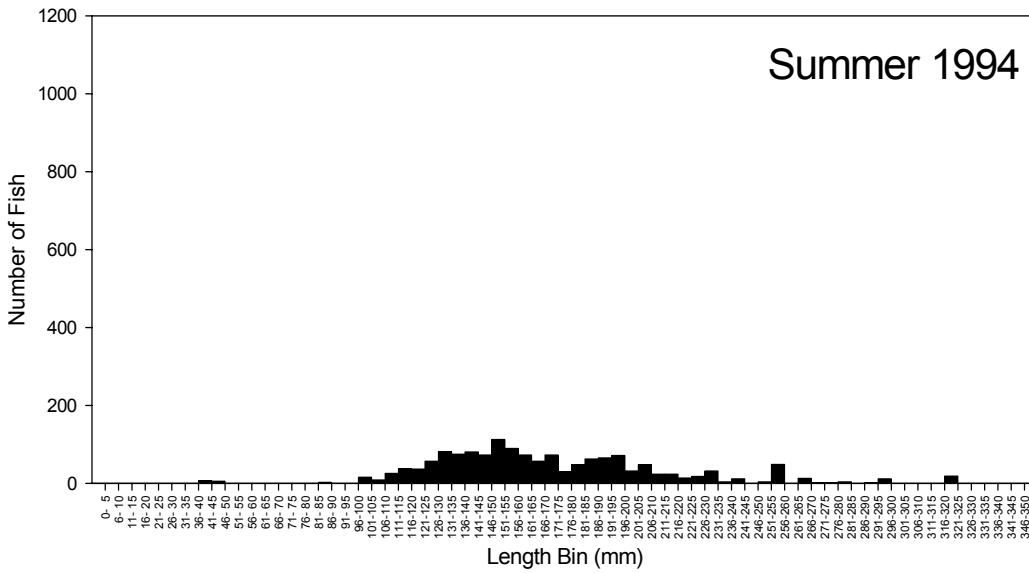
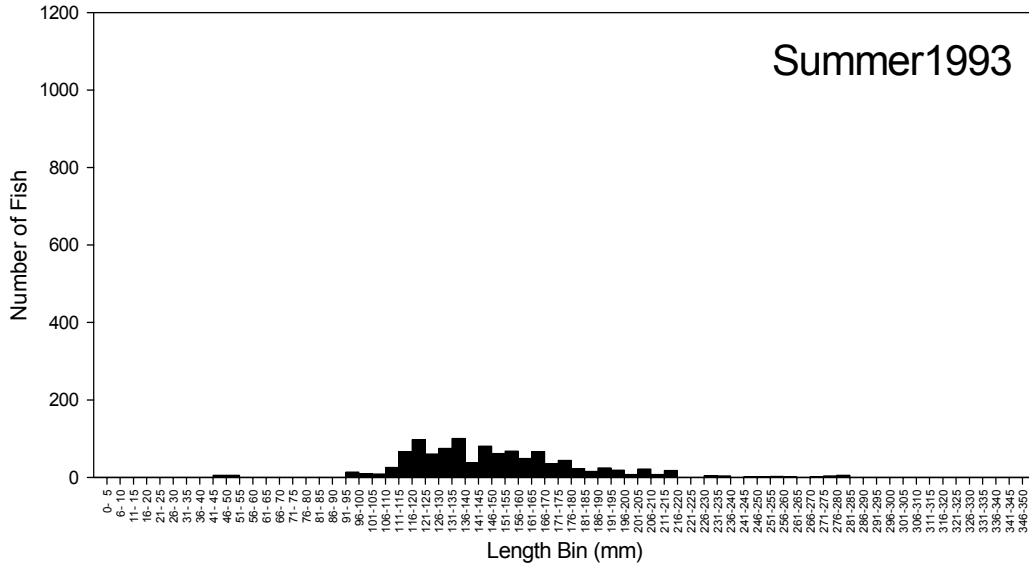


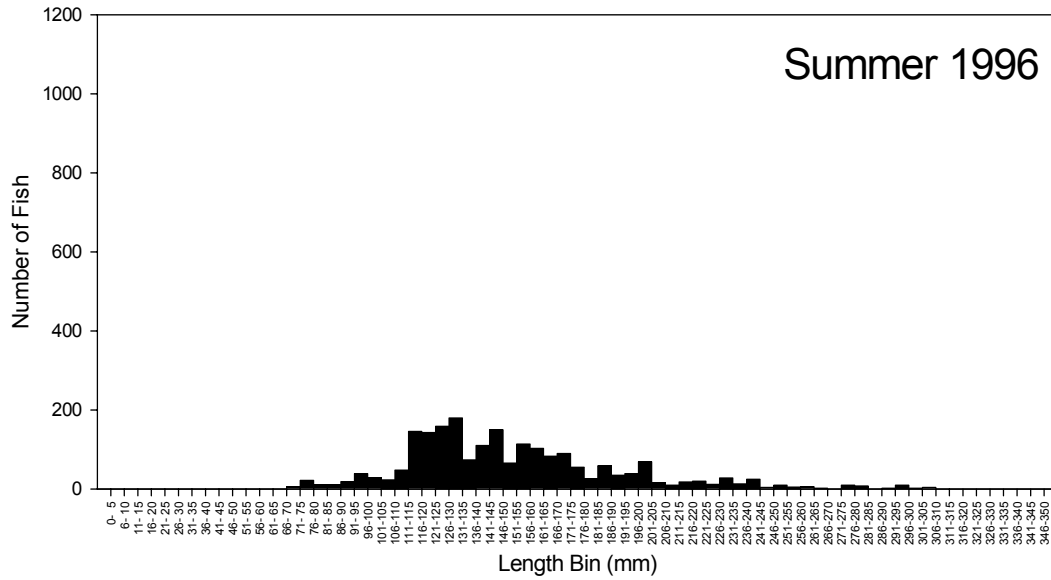
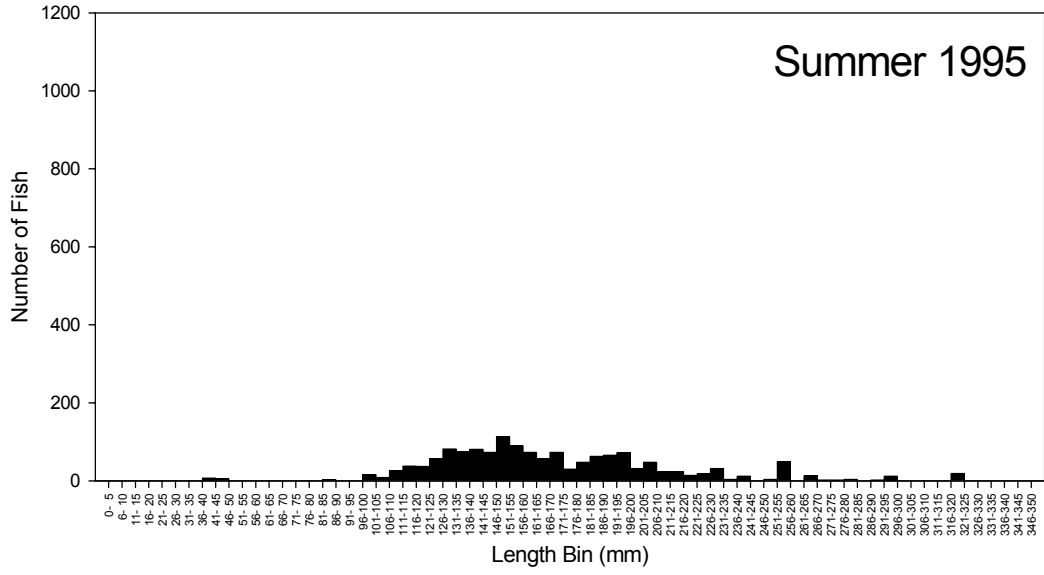


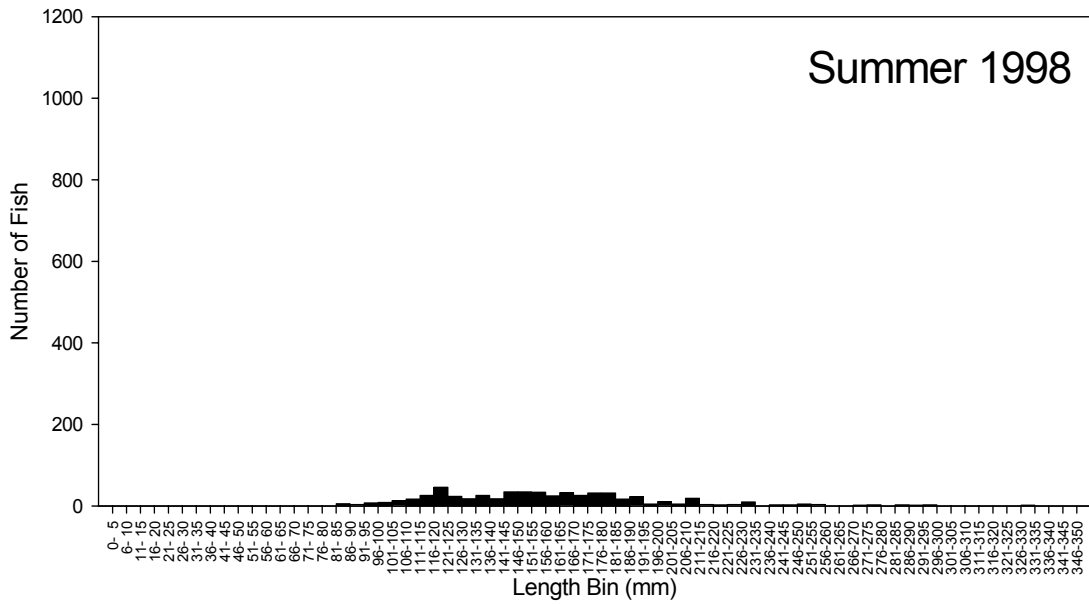
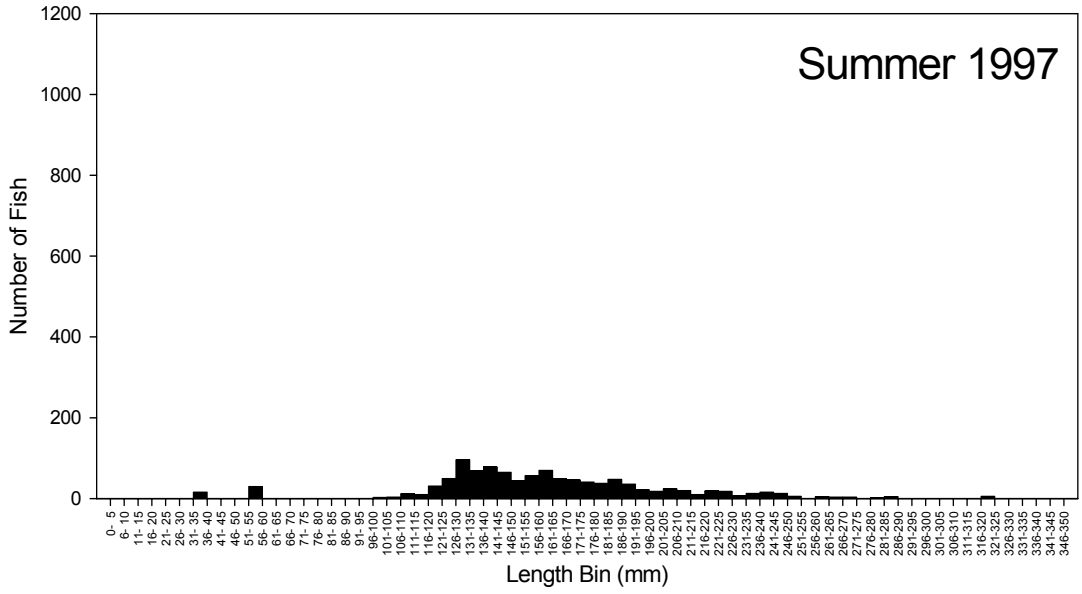


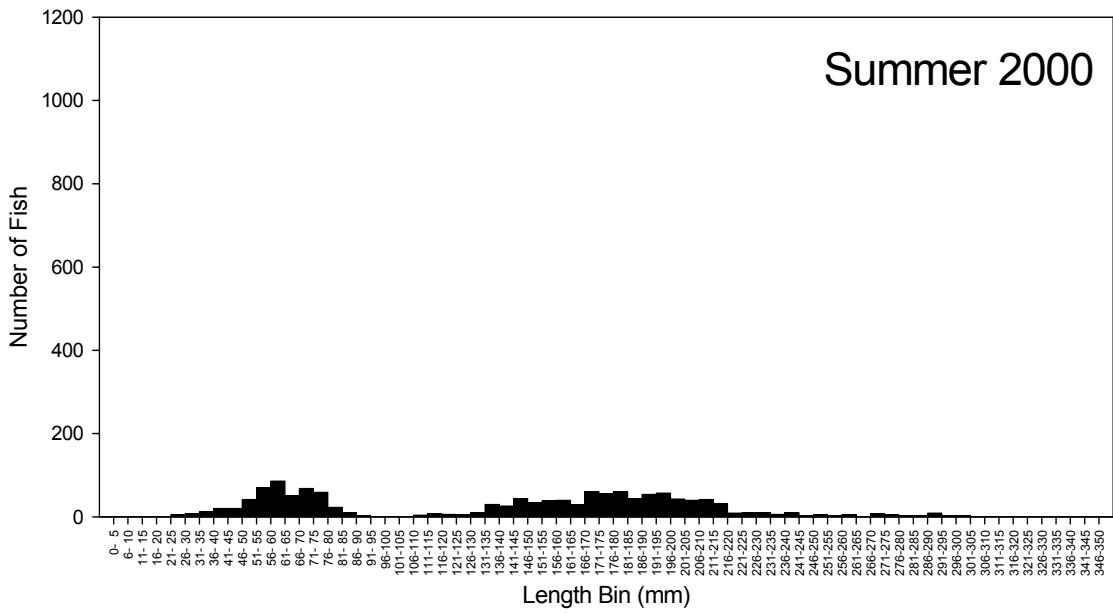
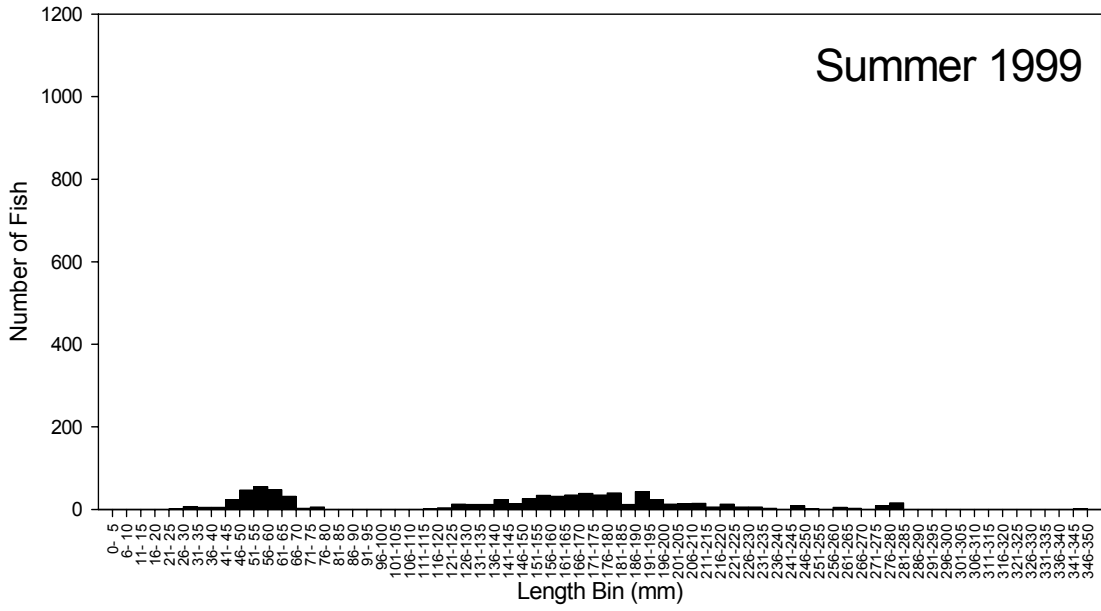


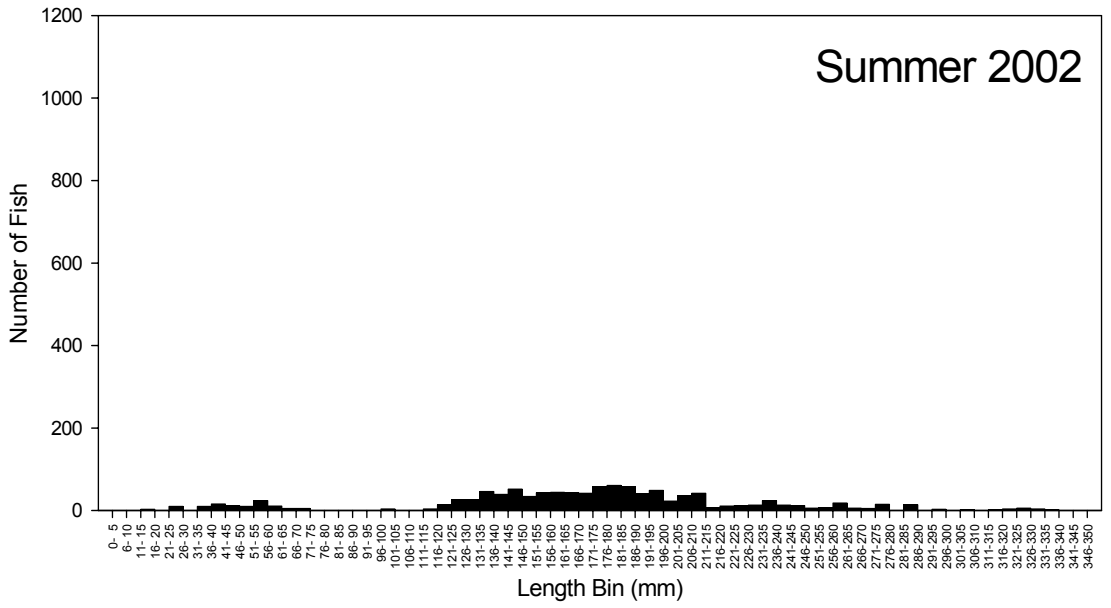
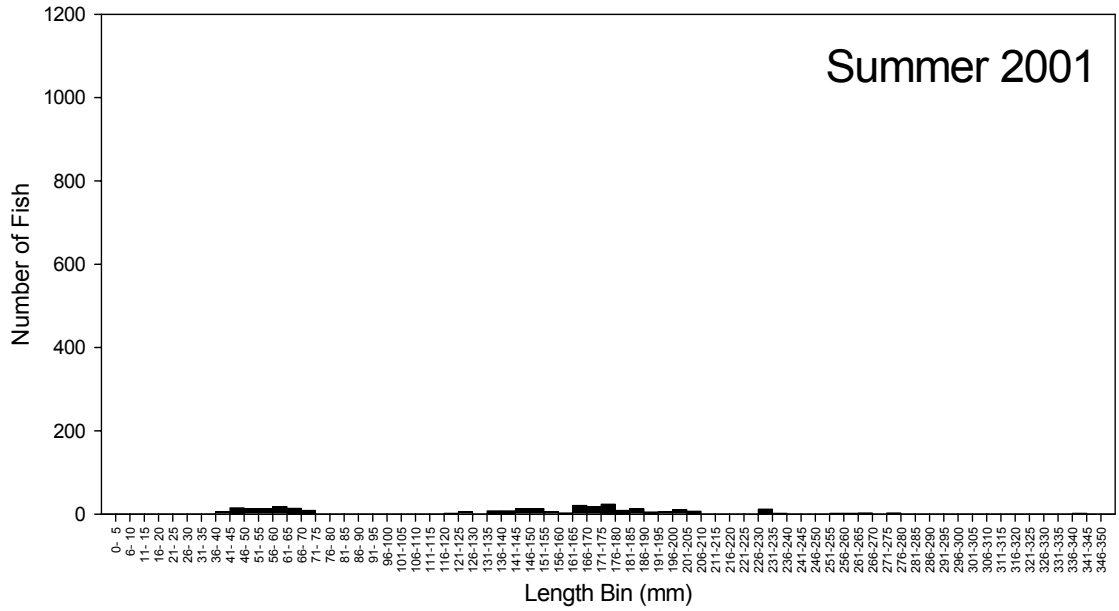


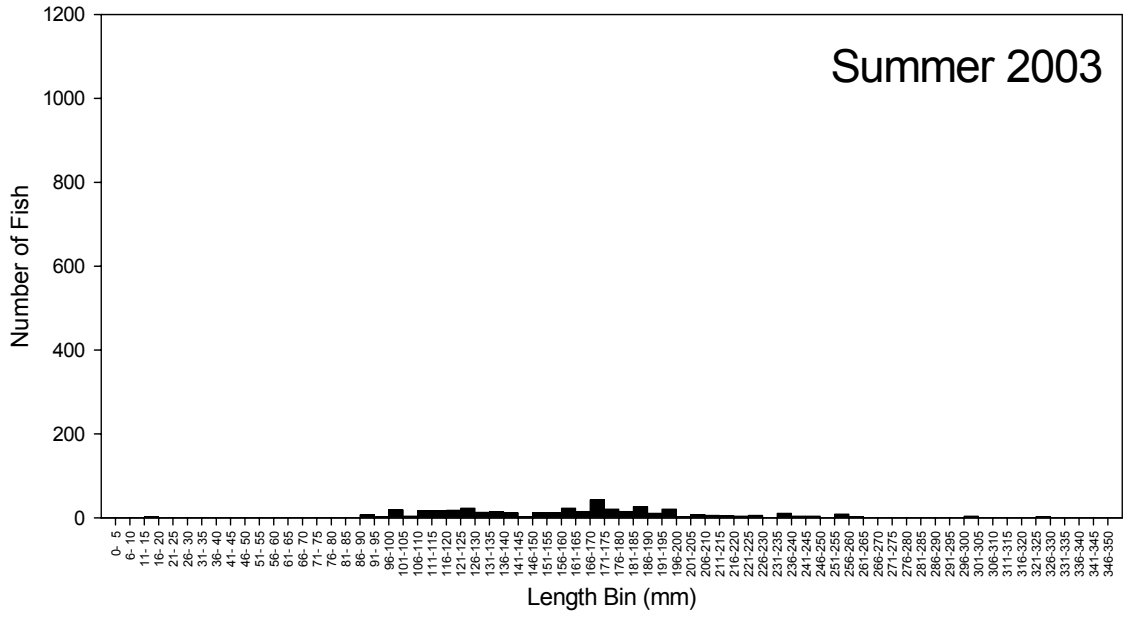












Appendix 1. Table 1. Boundaries used to separate age 0 from age 1 each year. (mm).

Year	Summer	Fall
1987	85	185
1988	85	180
1989	85	180
1990	85	175
1991	85	175
1992	85	170
1993	85	170
1994	85	155
1995	85	175
1996	65	170
1997	85	175
1998	80	155
1999	85	175
2000	95	180
2001	85	165
2002	85	170
2003	85	

Appendix 1. Table 2. Age fraction estimates for the trawl-vulnerable red snapper population.

	Sample size	Fractions							
		Fall	Summer	zero	Fall one	two plus	zero	Summer one	two plus
1987	159	203	0.697222	0.288889	0.013889		0	0.992579	0.007421
1988	460	146	0.667586	0.328966	0.003448		0	0.979798	0.020202
1989	850	91	0.925097	0.068093	0.006809	0.026316	0.973684		0
1990	1160	696	0.840643	0.152778	0.006579		0	0.998879	0.001121
1991	1490	418	0.937972	0.056764	0.005264	0.061975	0.928571		0.009454
1992	615	250	0.756361	0.235929	0.00771	0.03255	0.940325		0.027125
1993	1181	355	0.810412	0.17093	0.018658	0.010999	0.989001		0
1994	1632	546	0.893758	0.100447	0.005795	0.010568	0.974901		0.014531
1995	1809	568	0.920326	0.073871	0.005803	0.112583	0.883738		0.003679
1996	1363	641	0.773925	0.205964	0.020111	0.000939	0.992492		0.00657
1997	1502	446	0.919339	0.077032	0.00363	0.041554	0.95122		0.007227
1998	995	378	0.904221	0.092532	0.003247	0.008224	0.981908		0.009868
1999	1776	350	0.939926	0.052218	0.007856	0.294344	0.672237		0.033419
2000	1332	654	0.880069	0.112371	0.00756	0.354383	0.632987		0.01263
2001	1196	146	0.851548	0.136005	0.012447	0.305842	0.683849		0.010309
2002	1165	487	0.917318	0.07727	0.005412	0.096804	0.869406		0.03379
2003		215				0.004673	0.976636		0.018692

Appendix 2. Bugs code for the analysis described in this paper. Parameters describing the trawl survey indexes (CPUE's) are input as mu and tau. Parameters describing the age compositions are input as alpha and N. Shrimping effort is input as eff.

```

model seasonal {
  # one is summer, two is fall, year index is calendar
  for (i in 1:17) {
    for (j in 1:2) {
      index[i,j]~dlnorm(mu[i,j],tau[i,j])
      count[i,j,1:3]~dmulti(alpha[i,j,1:3],N[i,j])
      for (k in 1:3) {
        fract[i,j,k]<-count[i,j,k]/N[i,j]
        cpue[i,j,k]<-index[i,j]*fract[i,j,k]
      }
    }
  }
  for (i in 1:16) {
    Z[i,1]<-(-0.25)*log(cpue[i,2,2]/cpue[i,1,2])
    Z[i,2]<-(-0.125)*log(cpue[i+1,1,2]/cpue[i,2,1])
  }
  for (i in 1:16) {
    for (j in 1:2) {
      zsq[i,j]<-Z[i,j]*Z[i,j]
      xy[i,j]<-Z[i,j]*eff[i,j]
      xsq[i,j]<-eff[i,j]*eff[i,j]
    }
  }
  q<-(sum(xy[,,]) - sum(Z[,,])*sum(eff[,,])/32) / (sum(xsq[,,]) - sum(eff[,,])*sum(eff[,,])/32)
  M<-(sum(Z[,,])-q*sum(eff[,,]))/32

  tc<-sum(zsq[,,])-sum(Z[,,])*sum(Z[,,])/32
  srn<-sum(xy[,,])-sum(eff[,,])*sum(Z[,,])/32
  ssm<-pow(srn,2)
  ssrd<-sum(xsq[,,])-sum(eff[,,])*sum(eff[,,])/32
  sssr<-ssm/ssrd
  s2<-(tc-ssr)/30
  s<-sqrt(s2)
  sdintfact<-sum(xsq[,,])/(32*ssrd)
  sdint<-s*sqrt(sdintfact)

  ZbarS<-mean(Z[,1])
  ZbarF<-mean(Z[,2])
  Zannual<-4*ZbarS+8*ZbarF
  Mannual<-M*12
}

```