Retrospective Coding of Dual Size Classes Of Size Frequency Data for Red Snapper Collected During SEAMAP Shrimp/BottomFish Surveys

Gilmore “Butch” Pellegrin
Nathaniel Sanders, Jr
Kim Johnson
André Debose

Introduction

Since the inception of the Fall Groundfish Program in 1972, species exhibiting two distinct size classes in trawl samples were separated into small and large categories. The small category was arbitrarily referred to as Young-of-the-Year (although the strict definition of Young-of-the-Year was not met) and the large category as Adults. The rationale for distinguishing the two size classes was to improve estimates of individual weight. Since individuals of respective species were weighed collectively and not individually, an estimate of individual weight was obtained by dividing the weight of all individuals of a species, by the number caught. This approach works well when individuals are of relatively uniform size; however, as size ranges increase, the estimated weight per individual becomes less representative. This shortcoming culminates in cases where two distinct size classes are sampled and the estimated weight per individual at times yields a result which does not accurately represent either size class. Separating distinct size classes into Young-of-the-Year and Adult categories allowed for the computation of separate estimated weights per individual, one for each size class.

Although numbers and weights of species were coded at-sea to represent the Young-of-the-Year and Adult size classes for catch data, this coding was not carried over to size measurement data. That is, as individuals were measured, there was no method of recording whether they were categorized into Young-of-the-Year or Adults. It thus became necessary at the end of each survey, to introduce a data field into the size frequency data to reflect instances where Young-of-the-Year and Adult categories were measured. The object of this report is to explain the method used to retrospectively assign Young-of-the-Year and Adult categories to size measurement data.

Method

Categorizing measurement data was problematic in that the individual performing the categorizations had to do so in retrospect. Thus, the individual was required to have extensive knowledge of the encountered fauna and procedures used during surveys. These requirements were best met by the Chief Scientists. The retrospective nature of the task required subjective decisions based on individual expertise, thus the repeatability of results may vary according to individual. Coding was straightforward except for occasions when Watch Leaders appeared to either misunderstand the Young-of-the-Year concept, or understood it but misapplied it. This became apparent when size frequency data were inspected and there was no obvious distinction between two size classes. It therefore became necessary to develop conventions for assigning Young-of-the-Year codes to size frequency data.

In general, retrospective coding of size frequency data was accomplished by comparing the number of individuals caught, to the number (and sizes) of individuals measured, with additional guidance provided by average weights of individuals (in cases where distinct differences between two size classes were not evident). The required data were stored in two files entitled Biological and Size Frequency. The Biological File contained numbers and weights of species sampled and provided an expected number of individuals measured (the convention is to measure a maximum of 20 individuals per category). These values were compared to the number of measurements found in the Size Frequency File with consideration given to noticeable differences between size classes. If the distinction between size classes was not immediately obvious, then individual weights were estimated for measured fish using a length-weight relationship, and estimated weights were compared to measured weights recorded in the Biological File. Length-weight relationships were derived from four surveys (two Summer and two Fall) for which individual lengths and weights were collected for red snapper.

$$\textit{Summer}: \hat{W} = (27.3\times10^{-8})\, L^{2.49116} \qquad (p<0.0001,\ n=108)$$
$$\textit{Fall}: \hat{W} = (9.5\times10^{-8})\, L^{2.69240} \qquad (p<0.0001,\ n=525)$$

The boundary between Young-of-the-Year and Adults was determined utilizing the observed differences in sizes and, if necessary, weights of individuals. Perhaps the best method of explaining the coding conventions is by way of examples.

Step 1

All occasions where red snapper were categorized into Young-of-the-Year and Adult categories were identified from the Biological File. An average weight per individual was computed for each category and categorizations were inspected for errors (average weights per individual for Young-of-the-Year categories should always be less than those for Adults, i.e., $\bar{W}_Y < \bar{W}_A$). Erroneous categorizations were corrected as necessary. This step also provided guidance for coding the corresponding size measurements in the Size Frequency File.

Example 1:

---------- Cruise Number=197 ----------

Obs	Station Number	YOY	Number Caught	Weight Per Individual (kgs)
1	53378	T	5	0.020
2	53378		13	0.023
3	53384	T	5	0.010
4	53384		6	0.017
5	53397	T	8	0.012
6	53397		2	0.023
7	53406	T	11	0.009
8	53406		96	0.021
9	53409	T	42	0.031
10	53409		3	0.567
11	53410	T	25	0.028
12	53410		8	0.225
13	53417	T	1	0.020
14	53417		3	0.033
15	53418	T	6	0.010
16	53418		4	0.050
17	53430	T	7	0.011
18	53430		5	0.020
19	53435	T	19	0.008
20	53435		16	0.025
21	53436	T	5	0.020
22	53436		2	0.050
23	53441	T	32	0.006
24	53441		6	0.033
25	53445	T	2	0.040
26	53445		1	0.020

The above data shows 13 instances where red snapper were categorized into Young-of-the-Year (coded as "T" in the YOY column) and Adults (YOY field left blank). Note the error at Station 53445 where the estimated weight per individual for Young-of-the-Year was greater than that for Adults. In this case, the Young-of-the-Year/Adult codes appeared to be erroneous in the Biological File and the codes were reversed. The above information also provided expected results for size measurement data. At Station 53378 there should be 18 total measurements, 5 Young-of-the-Year and 13 Adults.

Step 2

At each sampling station, the expected numbers of measured Young-of-the-Year (N_Y) and Adult (N_A) individuals obtained from the Biological File were summed and compared to the total number of individuals measured (M_T) in the Size Frequency File. If the two values were equal (i.e., $N_Y + N_A = M_T$), then the M_T measurements were sorted in ascending order and the first N_Y measurements in the Size Frequency File were coded as Young-of-the-Year.

Example 2:

---------- Cruise Number=197 ----------

Obs	Station Number	YOY	Number Caught	Weight Per Individual (kgs)
1	53378	T	5	0.020
2	53378		13	0.023

The Biological File indicated 18 red snapper were caught, 5 Young-of-the-Year and 13 Adults.

---------- Cruise Number=197 Station Number=53378 CODE=01 ----------

Obs	Size	YOY
1	52	T
2	75	T
3	79	T
4	83	T
5	84	T
6	106	
7	107	
8	111	
9	111	
10	112	
11	112	
12	119	
13	119	
14	126	
15	128	
16	131	
17	141	
18	148	

Since 18 measurements were found in the Size Frequency File, they were sorted in ascending order and the first 5 were coded as Young-of-the-Year and the remaining 13 as Adults.

Step 3

If the number of measurements contained in the Size Frequency File was not equal to the expected number of measurements derived from the Biological File, measurements were sorted in ascending order and differences between adjacent measurements were computed. The first large difference was initially interpreted as a boundary between Young-of-the-Year and Adult categories. Assuming this boundary was correct, the numbers of measured Young-of-the-Year (M_Y) and Adults (M_A) were counted. These values were compared to the expected number of measurements, N_Y and N_A.

1. If the number of measured Young-of-the-Year or Adults was less than the expected numbers measured (i.e., $M_{Y\ or\ A} < N_{Y\ or\ A}$), then,

   A. The Comments section of the field data sheet was inspected for references to individuals that were damaged and thus rendered immeasurable, individuals that were not measured due to oversight, or electronic measuring board malfunctions. If missing measurements were accounted for, then no remedial action was possible.

B. Although Watch Leaders should have measured a maximum number of 20 individuals of each category, occasionally it appeared that a random sample was taken from a mixture of Young-of-the-Year and Adult individuals. In such instances, the percentage of Young-of-the-Year individuals measured should be similar to the percentage of Young-of-the-Year individuals occurring in the Biological File. If this condition was satisfied, the appropriate number of measurements were coded as Young-of-the-Year.

Example 3:

Cruise Number=197

Obs	Station Number	YOY	Number Caught	Weight Per Individual (kgs)
97	53511	T	26	0.050
98	53511		13	0.246

Twenty Young-of-theYear and thirteen Adults should have been measured.

Cruise Number=197 Station Number=53511 CODE=01

Obs	Size	Delta	YOY
1	110	0	T
2	120	10	T
3	120	0	T
4	145	25	T
5	150	5	T
6	151	1	T
7	154	3	T
8	155	1	T
9	160	5	T
10	161	1	T
11	163	2	T
12	165	2	T
13	219	54	
14	224	5	
15	225	1	
16	230	5	
17	245	15	
18	248	3	
19	248	0	
20	253	5	

However, there were only 20 measurements in the Size Frequency File. The largest difference between adjacent measurements (Delta) occurred between observations 12 and 13 (i.e., 65% Young-of-the-Year, 35% Adults). This composition agreed closely with the data contained in the Biological File (67% Young-of-the-Year and 33% Adults). Therefore measurements of 165 mm and smaller were coded as Young-of-the-Year.

C. Occasionally Watch Leaders measured 20 red snapper even though Young-of-the-Year and Adults were present (recall that a maximum of 20 of each category should have been measured). The 20 individuals measured appeared to consist of all members of one category and the remaining complement consisted of the other. If this was suspected, an additional check was performed. Individual weights of measured fish were estimated using a length-weight relationship. Average weights per individual were then computed for the two categories and compared to the average weights computed from the Biological File. If the averages compared favorably then the appropriate number of measurements were coded as Young-of-the-Year.

Example 4:

---------- Cruise Number=243 ----------

Obs	Station Number	YOY	Number Caught	Weight Per Individual (kgs)
543	00108	T	35	0.051
544	00108		4	0.146

Twenty Young-of-the-Year and four Adults should have been measured.

---------- Cruise Number=243 Station Number=00108 CODE=01 ----------

Obs	Size	Delta	YOY	Estimated Weight (kgs)
1	119	0	T	0.037
2	121	2	T	0.038
3	127	6	T	0.044
4	135	8	T	0.052
5	135	0	T	0.052
6	140	5	T	0.057
7	140	0	T	0.057
8	140	0	T	0.057
9	144	4	T	0.061
10	144	0	T	0.061
11	145	1	T	0.063
12	146	1	T	0.064
13	151	5	T	0.070
14	151	0	T	0.070
15	151	0	T	0.070
16	158	7	T	0.079
17	187	29		0.124
18	192	5		0.133
19	199	7		0.147
20	231	32		0.219

But only 20 were measured, possibly 16 Young-of-the-Year and all 4 of the Adults (it is also possible that only Young-of-the-Year were measured). The average estimated weight per individual for fish 158 mm and smaller was 0.058 kgs, and 0.156 for the remaining four large fish. Since these values agreed closely to 0.051 and 0.146 kgs for the Young-of-the-Year and Adults, respectively, from the Biological File, it appeared that all 4 Adults were measured and the remaining 16 were Young-of-the-Year. The data were then coded accordingly.

D. Occasionally Watch Leaders did not measure a sufficient number of both categories of fish. Estimated weights were compared to average weights per individuals from the Biological File as an aid in determining the desired boundary between categories.

Example 5:

---------- Cruise Number=232 ----------

Obs	Station Number	YOY	Number Caught	Weight Per Individual (kgs)
473	00217	T	3	0.022
474	00217		2	0.040

At Station 217, the Biological File showed 3 Young-of-the-Year and 2 Adults were caught.

---------- Cruise Number=232 Station Number=00217 CODE=01 ----------

Obs	Size	Delta	YOY	Estimated Weight (kgs)
1	84	0	T	0.014
2	149	65		0.067

However, the Size Frequency File contained 2 measured fish instead of the expected 5. Since a large Delta was observed, the estimated weight for the 84 mm fish was less than the recorded weight per individual for Young-of-the-Year in the Biological File, and the estimated weight for the 149 mm fish exceeded the recorded weight per individual for Adults in the Biological File, the smaller fish was coded as Young-of-the-Year.

2. If the number of measured individuals exceeded the expected number of measurements (i.e., $M_{Y\ or\ A} > N_{Y\ or\ A}$), then measurements were randomly deleted until the number of measurements equaled the expected number of measurements ($M_{Y\ or\ A} = N_{Y\ or\ A}$). Originally, this condition was considered to be an error in counting the number of individuals (i.e., the data in the Biological File). But upon further deliberation, it was considered equally possible that errors occurred in the measuring process. Fish were measured on an electronic measuring board which measures fish length by electronically sensing the position of a magnet placed over the fish's tail. Consequently, it is possible to accidently measure a fish twice by allowing the wand to linger over a fish, or record an incorrect measurement by carelessly laying the wand on the board. Although Watch Leaders verified at-sea that numbers caught equaled numbers measured, a few mistakes occurred due to inattention or fatigue. We considered increasing the numbers caught so that it equaled the numbers measured, but this approach required the estimation of weights corresponding to the increased number of fish. The increase in species weight then necessitated an adjustment to the sample weight which was in turn used to expand sample numbers and weights to total catch values. Thus, this approach impacted density and biomass estimates of all species in a catch, not just red snapper. We therefore chose a solution which minimized the overall impact to the data. The chosen solution affected only red snapper and was based on a random approach thereby minimizing bias.

Step 4

If there was no distinct difference between 2 adjacent measurements then the number of individuals measured was compared to the expected number of individuals measured for both categories. Estimated average weight per individual was again used as an aid in determining if two size classes were measured.

Example 6:

---------- Cruise Number=237 ----------

Obs	Cruise Number	Station Number	YOY	Number Caught	Weight Per Individual (kgs)
489	237	00123	T	1	0.001
490	237	00123		57	0.029

Twenty one fish should have been measured, one Young-of-the-Year and twenty Adults.

---------- Cruise Number=237 Station Number=00123 CODE=01 ----------

Obs	Size	Delta	YOY	Estimated Weight (kgs)
1	88	0		0.016
2	96	8		0.021
3	102	6		0.024
4	102	0		0.024
5	108	6		0.028
6	112	4		0.031
7	112	0		0.031
8	116	4		0.034
9	116	0		0.034
10	117	1		0.035
11	120	3		0.038
12	122	2		0.039
13	125	3		0.042
14	126	1		0.043
15	126	0		0.043
16	128	2		0.045
17	128	0		0.045
18	130	2		0.047
19	132	2		0.049
20	132	0		0.049

Only 20 fish were measured and it was not obvious as to whether all Adults were measured, or 19 Adults and 1 Young-of-the-Year were measured. All measurements were considered to be Adults since there was no relatively large Delta and the estimated weight for the 88 mm fish (0.016 kgs) did not compare favorably to the 0.001 kg weight found in the Biological File. Additionally, the average estimated weight per individual for the 20 fish (0.036 kgs) was similar to the average weight per individual for the Adult category (0.029 kgs).

Summary

The method of retrospectively assigning Young-of-the-Year codes to size frequency data was simple when the protocol was used correctly. The proper procedure was to measure a maximum number of 20 individuals per category. In such cases, information provided in the Biological File dictated the coding of individuals in the Size Frequency File. However, there were occasions where the Young-of-the-Year concept was either misunderstood or applied inappropriately. Misuse occurred when Watch Leaders, 1) took a random sample of 20 from a mixture of Young-of-the-Year and Adult fish, 2) measured only 20 individuals, all individuals from one category and the remaining complement from the other, 3) did not measure the required number of individuals from each category, and 4) measured only one of two size classes. The size distributions of fish were inspected for distinct size classes to determine which of the above four situations occurred. In general, retrospective coding of size frequency data was accomplished by comparing the number of individuals caught, to the number (and sizes) of individuals measured, with additional guidance provided by average weights of individuals (procedure summarized in Appendix).

Fifty-seven surveys were processed. The correct procedure was used in 89.6% of the tows where two size classes of red snapper were identified and measured. Insufficient numbers per category were measured at 5.2% of the tows, all of one category (remaining complement of the other) at 2.8%, categories mixed at 1.7%, damaged fish at 0.4%, and meauring one size class and not the other at 0.3% (Table 1).

It is also important to note that the Mississippi Laboratory now uses an automated data recording system called the Fishery Scientific Computing System and the need for retrospective coding of size data no longer exists as data are now coded at-sea.

Table 1. Summary of retrospective coding of young-of-the-year red snapper size measurements for Shrimp/Bottomfish surveys conducted by the Mississippi Laboratory (1972-2001).

Year	Cruise Number	Number Of Tows	Number of Tows With Red Snapper	Number of Tows With YOYs	$\bar{w}_A < \bar{w}_Y$	$M_T = N_A + N_Y$	$M_{A \text{ or } Y} < N_{A \text{ or } Y}$				One Size Class
							Damaged	Categories Mixed	All of One Category	Measured Insufficient Numbers	
1972	40	126	69	0							
	42	58	48	0							
1973	48	272	139	0							
1974	51	465	79	0							
	52	702	73	0							
	55	243	95	0							
1975	60	331	36	0							
	62	280	104	0							
1976	67	285	31	0							
	71	307	123	0							
1977	81	230	70	2							
	83	242	103	0							
1978	92	319	109	0							
1979	101	273	107	0							
1980	112	233	141	0							
1981	118	131	79	0							
	122	279	150	0							
1982	127	188	54	1							
	130	273	163	3							
1983	135	191	35	0							
	138	222	95	0							
1984	145	219	32	0							
	148	226	77	1							
1985	153	134	35	0							
	156	306	90	0							
1986	160	145	18	0							
	163	262	92	1						1	
1987	167	223	59	1		1					
	171	167	66	1				1			
1988	174	188	46	1		1					
	177	210	101	0							
1989	180	167	29	1		1					

Table 1. (Continued)

Year	Cruise Number	Number Of Tows	Number of Tows With Red Snapper	Number of Tows With YOYs	$\bar{W}_A < \bar{W}_Y$	$M_T = N_A + N_Y$	$M_{A \text{ or } Y} < N_{A \text{ or } Y}$				One Size Class
							Damaged	Categories Mixed	All of One Category	Measured Insufficient Numbers	
1989	184	220	121	6		4			1	1	
1990	189	216	92	0							
	191	238	151	5		2			2	1	
1991	195	225	76	5		5					
	197	245	163	37	1[1]	31		4		2	
1992	200	215	64	2		2					
	202	221	115	27		23				4	
1993	205	230	70	4		4					
	208	250	134	27	1[2]	26	1				
1994	210	236	89	4	1[3]	4					
	214	242	153	41		39				2	
1995	217	212	87	15		15					
	219	223	162	32		31				1	
1996	221	223	95	3		3					
	224	231	142	7		6				1	
1997	226	208	81	3		3					
	229	226	140	4		4					
1998	230	186	67	1		1					
	232	201	98	3		2				1	
1999	235	231	76	4		4					
	237	242	158	10		8				1	1
2000	240	236	104	13		13					
	243	233	160	17	1[4]	12			5		
2001	246	64	25	2		2					
	248	251	150	11		11					
	Sums	13,401	5,320	295	4	258	1	5	8	15	1
	Percents		39.7[5]	5.5[6]	1.4[7]	89.6[7]	0.4[7]	1.7[7]	2.8[7]	5.2[7]	0.3[7]

[1] Young-of-the-year/Adult coding appeared to be erroneous in the Biological File therefore the coding was reversed.
[2] The weight of one young-of-the-year was changed from 0.030 to 0.003 kgs.
[3] The weight of two adults was changed from 0.045 to 0.450 kgs.
[4] The weight of two young-of-the-year was changed from 0.106 to 0.016 kgs.
[5] Percent of total number of tows containing red snapper.
[6] Percent of tows with red snapper having two size classes.
[7] Expressed as percents of the number of tows containing two size classes. Note: 7 tows were subtracted from 295 since no fish were measured for 7 tows.
Note: Table does not contain a column for occasions where too many fish were measured per category because data were deleted upon survey completion and were unavailable for summary.

Appendix.

A. Identify samples where red snapper were divided into Young-of-the-Year and Adult categories. Compute average weight per individual for each category, inspect for errors and correct as necessary.
B. Determine the expected number of measured fish per category from the Biological File. Sum the number of fish measured and compare to the sum of expected measurements. If equal, then code accordingly.
C. If unequal and there were two distinct size classes,
    1. If number measured was less than expected number,
        A. Inspect Comments section for explanation of unmeasured fish.
        B. Were 20 fish randomly taken from a mixture of Young-of-the-Year and Adult fish?
        C. Were 20 fish measured, all members of one category and the rest of the other?
        D. Were insufficient numbers of each category measured?
    2. If number measured was more than expected number, randomly delete measurements until number measured equaled expected number.
D. Unequal and no two distinct size classes, code according to estimated weights per individual.