NOAA Fisheries Reef Fish Video Surveys:
Yearly Indices of Abundance for gag (*Mycteroperca microlepis*)

Christopher T. Gledhill, G. Walter Ingram, Jr., Kevin R. Rademacher, Paul Felts, and Brandi Trigg
Southeast Fisheries Science Center
Mississippi Laboratories
Pascagoula, MS

I. SEAMAP Reef Fish Survey of Offshore Banks

INTRODUCTION

The objective of the annual Southeast Area Monitoring and Assessment Program (SEAMAP) offshore reef fish survey is to provide an index of the relative abundances of fish species associated with topographic features (banks, ledges) located on the continental shelf of the Gulf of Mexico (Gulf) in the area from Brownsville, TX to the Dry Tortugas, FL (Figure 1). The total reef area surveyed is approximately 1771 km²; 1244 km² in the eastern and 527 km² in the western Gulf. The offshore reef fish survey was initiated in 1992, with sampling conducted during the months of May to August from 1992-1997, and in 2001-2005. No surveys were conduced from 1998 to 2000 and in 2003. A survey was conducted in 2005. However, data edits were not completed for SEDAR 10. The 2001 survey was abbreviated due to ship scheduling.

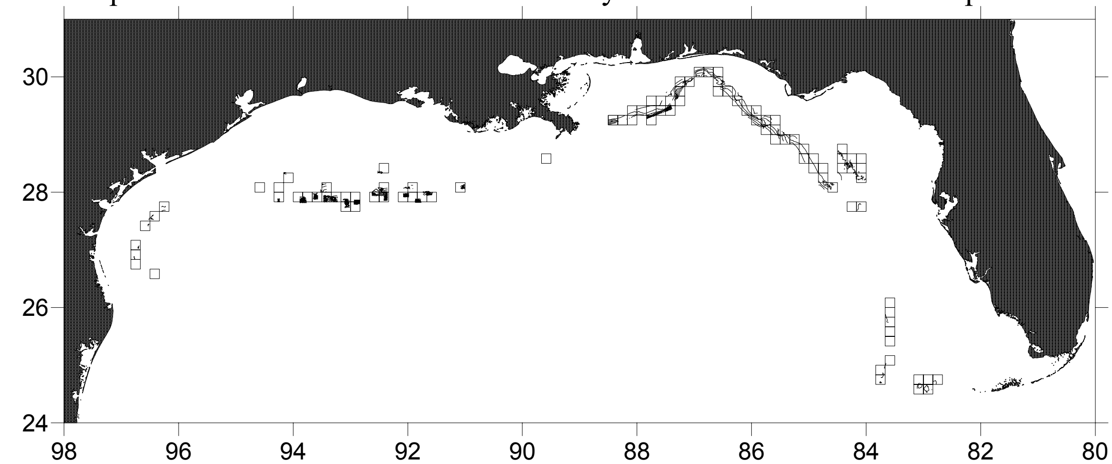

Figure 1. Gulf of Mexico shelf-edge banks sampled during SEAMAP offshore reef fish survey with sample blocks.

SAMPLE DESIGN

The survey area is large. Therefore, a two-stage sampling design is used to minimize travel times between sample stations. The first-stage or primary sampling units (PSUs) are blocks 10 minutes of latitude by 10 minutes of longitude (Figures 2 and 3). The first-stage units are selected by stratified random sampling. The blocks were stratified, with strata defined by geographic region (4 regions: South Florida, Northeast Gulf, Louisiana-Texas Shelf, and SouthTexas), and by reef habitat area (Blocks ≤ 20 km² reef, Block > 20 km² reef). There are a total of 7 strata. The ultimate sample sites (second stage units) within a block are selected randomly. However, stratum 1 (South Florida, small blocks) and stratum 7 (S. Texas, small blocks) were not consistently sampled. So, these were dropped from annual indices.

GEAR

The SEAMAP reef fish survey currently employs four Sony VX2000 DCR digital camcorders mounted in Gates PD150M underwater housings. The housings are rated to a maximum depth of 150 meters. The four Sony VX2000 camcorders are mounted orthogonally and a height of 30 cm above the bottom of the pod. A chevron (or arrow) fish trap with 1.5-inch vinyl-clad mesh is used to capture fish for biological samples. In its greatest dimensions, the trap is 1.76 m in length, 1.52 m in width and 0.61 m in depth. A 0.4 m by 0.29 m blow out panel is placed on one side and kept closed using 7-day magnesium releases. The magnesium releases are examined after each soak and replaced as needed. The trap is deployed at a randomly selected subset of video stations. Both the camera pod and fish trap are baited with squid.

VIDEO TAPE VIEWING PROCEDURES

One video tape from each station is selected out of the four for viewing. If all four video cameras face reef fish habitat and are in focus, the viewed tape is selected randomly. Tape viewers examine 20 minutes of the selected video tape, identify, and enumerate all species for the duration of the tape. Identifications are made to the lowest taxonomic level and the time when each fish enters and leaves the field of view is recorded. This is referred as a time in - time out procedure (TITO).

Tapes are viewed from the time when the view clears from any silt plume raised by the gear when it landed. Less than 20 minutes may be viewed if the duration when water is not clear enough to count fish is less than 20 minutes, or if the camera array is dragged. If a tape contains a large amount of fish, it is sub-sampled. There are four cases for sub-sampling: 1) when there is generally a large number of fish of a given species present throughout the tape so that following individual fish is difficult; 2) large number of fish occur in pulses periodically during the tape; 3) a single school of fish; and, 4) multiple schools of fish. Three estimators of relative abundance are available from the video data: 1) presence and absence; 2) maximum count (each fish of each taxon is counted each time it appears on the screen); and, 3) a minimum count (i.e., mincount: the greatest number of a taxon that appears on screen at one time). Presence and absence (frequency of occurrence) and mincount estimators are advantageous because they avoid the potential of multiple counting of fish, and are reported here.

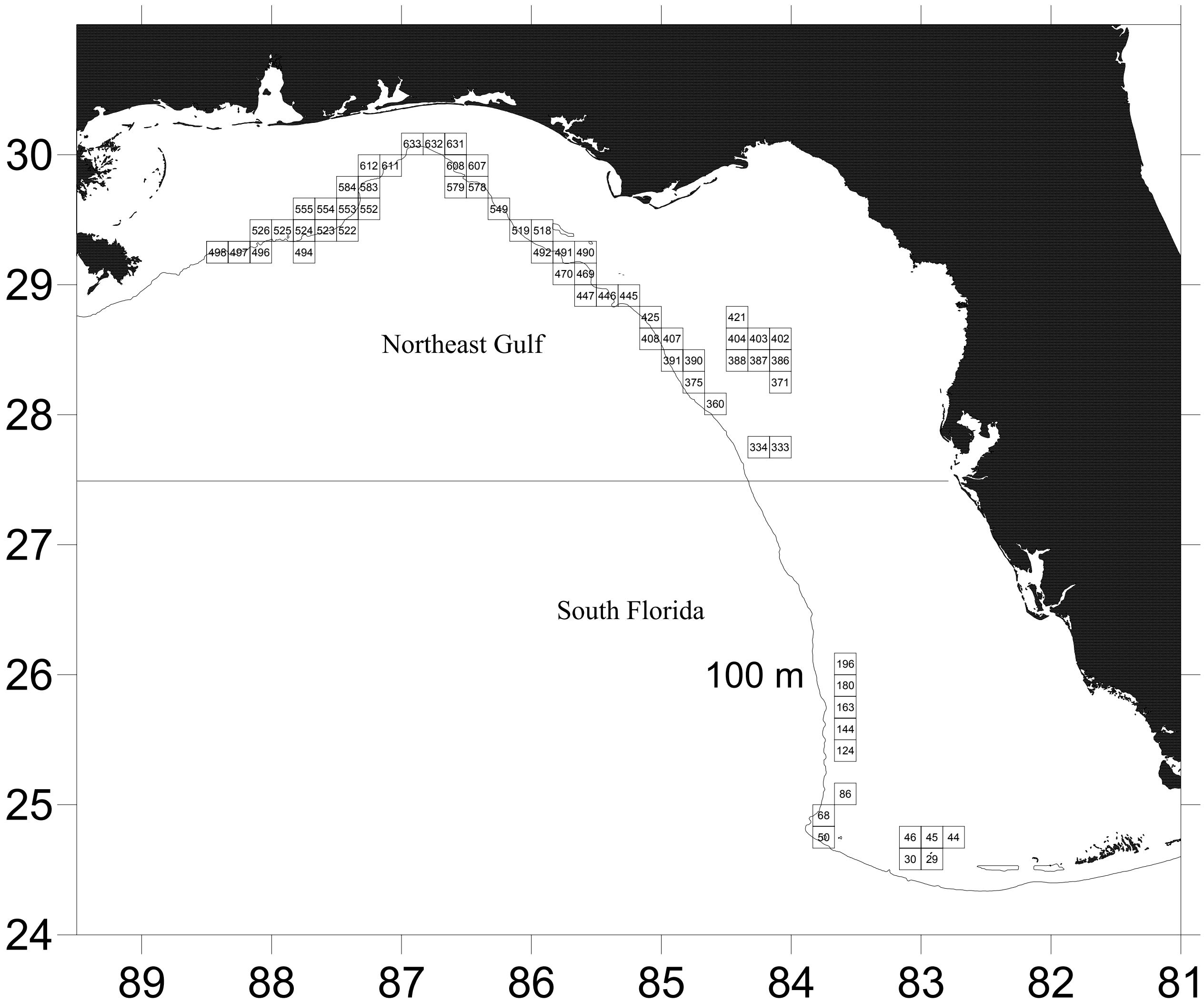

Figure 2. SEAMAP offshore reef fish survey sample blocks in the eastern Gulf of Mexico.

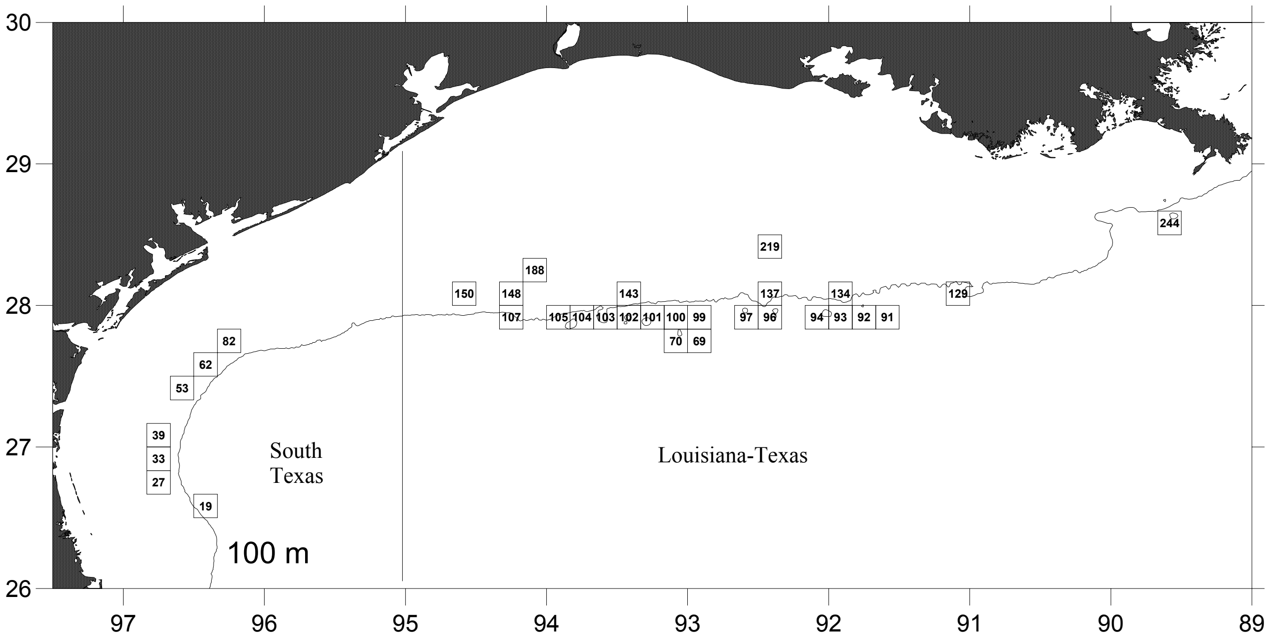

Figure 3. SEAMAP offshore reef fish survey sample blocks in western Gulf of Mexico.

STATISTICS

Design-based Estimator

The design-based estimators of abundance are those for stratified, two-stage sampling (Cochran, 1977). The number of strata and number of blocks sampled in the eastern and western Gulf of Mexico during SEAMAP reef fish survey are shown in Table 1.

1. Block means

$$\bar{x}_{hi} = \frac{\sum_{j=1}^{m_{hi}} x_{hij}}{m_{hi}},$$

where x_{hij} is the number of fish observed at the j-th site in the i-th block within the h-th stratum, and m_{hi} in the number of sites sampled in the i-th block and h-th stratum.

2. Stratum means

$$\bar{x}_h = \frac{\sum_{j=1}^{n_h} \bar{x}_{hi}}{n_h},$$

where $\bar{x}_{hi}$ is the i-th block mean in the h-th stratum and n_h is the number of blocks sampled in the h-th stratum.

3. Stratified mean

$$\bar{x}_{st} = \sum_h w_h \bar{x}_h ,$$

where w_h is the stratum weight estimated as the area of the stratum divided by the total survey area (A_h/A).

4. Variance of the stratified mean ($V(\bar{x}_{st})$), ignoring finite population correction

$$V_{\bar{x}_{st}} = \sum_h w_h^2 \left[\frac{s_{1h}^2}{n_h} + \frac{s_{2h}^2}{n_h m_h} \right]$$, where w_h is the stratum weight, s^2_{1h} and s^2_{2h} are the variances among

the first-stage and second-stage units, n_h and m_h are the number of first stage and second-stage units sampled.

5. Variance among first-stage units, s^2_{1h}

$$s_{1h}^2 = \frac{\sum_h (\bar{x}_{hi} - \bar{x}_h)^2}{n_h - 1}.$$

6. Variance among second-stage units, s^2_{2h}

$$s_{2h}^2 = \frac{\sum_i \sum_j (x_{hij} - \bar{x}_{hi})^2}{n_h (m_{hi} - 1)}.$$

The estimates for the frequency of occurrence of each species were calculated using the same equations where x_{hij} was either 0 or 1. The final estimate is a stratified mean proportion.

Model-based Index

In addition to the calculations of stratified means, a delta-lognormal modeling approach (Lo et al., 1992) was used to develop abundance indices. In order to develop standardized indices of annual average mincount for gag in the U.S. Gulf of Mexico, a delta-lognormal model, as described by Lo et al. (1992), was employed. This index is a mathematical combination of yearly mincount estimates from two distinct generalized linear models: a binomial (logistic) model which describes proportion of positive mincounts (i.e., presence/absence) and lognormal model which describes variability in only the nonzero mincount data. The GLMMIX and MIXED procedures in SAS were employed to provide yearly index values for both the binomial and lognormal sub-models, respectively. A backward stepwise selection procedure was employed to develop both sub-models. Type 3 analyses were used to test each parameter for inclusion or exclusion into the sub-model. Both variable inclusion and exclusion significance level was set at an $\alpha = 0.05$. The parameters tested for inclusion in each sub-model were region, year, stratum, and block nested within stratum, station depth (scaled to a mean of one). Also, the estimates from each model were weighted using the stratum area, and separate covariance structures were developed for each survey year. For the binomial sub-models, a logistic-type mixed model was employed. The fit of each model was evaluated using the fit statistics provided by the GLMMIX macro. When all annual frequencies of occurrence for a time series were less than 0.1, indicating a zero-inflated binomial distribution, a zero-inflated binomial regression model was also employed in place of the binomial sub-model using the methodology of (Tyre et al., 2003) and the NLMIXED procedure in SAS in an attempt to better model those data sets. Initially, several sub-model types were used to describe the nonzero mincount data. These included lognormal, Poisson and negative binomial. Based on analyses of residual scatter and QQ plots, the lognormal sub-model was more fitting than the others in describing the variability in the nonzero data in most of the models. Those models where a lognormal approach did not fit as well were those where there were very few data points. In those cases, the other two model types did not perform any better. As with the design-based analyses, model-based estimators were developed for gag Gulfwide, East Gulf only and West Gulf only. Also, all videos not lost during Hurricane Katrina were reviewed and mincounts were estimated for those gag with a black or copper-colored ventral surface (see Gilmore and Jones, 1992, and Bullock and Smith, 1991, for descriptions of gag with a darkened ventral surface).

Fish Sizes

The size of gag observed during the SEAMAP survey comes from fish measured on video tape with lasers. Lasers were first introduced in 1995. We developed a length frequency histogram for gag measured by laser on video.

RESULTS

Design-based Results and Conclusions

Abundance data from five strata were included for analysis during all years except 2001 (Tables 1-3). Stratum 1 was sampled only in 1994, 1996, 1997, and 2004. This stratum was 62.847 km^2 in area. Stratum 7 was sampled only in 1996, 1997, 2002, and 2004, and was 13.030 km^2 in area. Since these strata were not sampled during all years of the survey, they were excluded from design-based and model-based estimates of annual mean mincount. However, when included for those years, the stratified means and variances changed very little since their stratum weights were small. The 2001 survey was abbreviated. Only one stratum was sampled in the eastern Gulf of Mexico. We recommend that the 2001 estimates of abundance for the entire Gulf of Mexico, and for the eastern Gulf of Mexico not be used for estimating trends in fish abundance. However, the two strata in the western Gulf of Mexico were sampled, and provide a real estimate of zero catch.

Model-based Results and Conclusions

Three models converged: a Gulfwide model for gag, a model for gag in the East Gulf, and model for copper-belly gag (CBG) in the East Gulf. Table 4 lists yearly numbers of video stations, numbers of stations with gag and CBG, and those stations whose videos were lost during the storm. Based on Table 4 and the above conclusions, we reason that 2001 data should be dropped from analyses due to its low sample size compared to those of all other years. Also, due to issues of model convergence and index calculation we dropped 1992 data, and began the model-based analyses with 1993 data.

Gag Gulfwide Model-based Index

Tables 5-7 summarize the model-building process by which the Gulfwide gag index was developed. Table 5 summarizes the results of Type 3 analyses for those variables retained in the binomial sub-model. Table 6 summarizes the results of Type 3 analyses for those variables retained in the lognormal sub-model. Table 7 and Figure 4 summarize the index values for gag Gulfwide.

Due to the lack of any significant variables in the lognormal model, we suggest an alternate index based on the modeled frequency of occurrence. As seen in Table 8, modeled nonzero mincounts do not vary significantly (Table 6). Therefore this sub-model may contribute undescribed variability to the final index values. As seen in Table 8, a new standardized index based upon modeled frequency of occurrence provides lower CVs and smaller confidence intervals. For both of the Gulfwide gag abundance indices described above, there is an increase in index values in later survey years.

Gag East Gulf Model-based Index

Tables 9-11 summarize the model-building process by which the index was developed for gag from the East Gulf. Table 9 summarizes the results of Type 3 analyses for those variables retained in the binomial sub-model. Table 10 summarizes the results of Type 3 analyses for those variables retained in the lognormal sub-model. Table 11 and Figure 5 summarize the index values for gag from the East Gulf.

Due to the lack of any significant variables in the lognormal model, we suggest an alternate index based on the modeled frequency of occurrence. As seen in Table 12, modeled nonzero mincounts do not vary significantly (Table 10). Therefore this sub-model may contribute undescribed variability to the final index values. As seen in Table 12, a new standardized index based upon modeled frequency of occurrence provides lower CVs and smaller confidence intervals. For both of the East Gulf gag abundance indices described above, there is an increase in index values in later survey years.

CBG East Gulf Model-based Index

Tables 13-15 summarize the model-building process by which the index was developed for CBG from the East Gulf. Table 13 summarizes the results of the significance levels of parameters for those variables in the zero-inflated binomial sub-model. Table 14 summarizes the results of Type 3 analyses for those variables retained in the lognormal sub-model. Table 15 and Figure 6 summarize the index values for CBG from the East Gulf.

Due to the lack of any significant variables in the lognormal model except for year, we suggest an alternate index based on the modeled frequency of occurrence. As seen in Table 16, modeled nonzero mincounts do vary significantly, but the magnitude of the differences between nonzero mincounts is small (Table 14). Therefore the results of the lognormal sub-model may not improve the final index values and contribute undescribed variability to the final index values. As seen in Table 16, a new standardized index based upon modeled frequency of occurrence provides lower CVs and smaller confidence intervals. For both of the East Gulf CBG abundance indices described above, there is an increase in index values in later survey years.

GAG SIZE RESULTS

Fifty gag were hit by lasers during SEAMAP surveys. Since gag are rare, very few fish were hit by lasers. A single gag was measured in 1996. In 2002 and 2004, 7 and 42 gag were measured, respectively. The mean total length (±SE) was 690 (±12). Figure 7 shows the length frequency histogram of these fish.

Table 1. The number of strata and number of blocks sampled in the eastern and western Gulf of Mexico during SEAMAP reef fish survey.

	Eastern Gulf of Mexico		Western Gulf of Mexico	
Year	Number of Strata	Number of Blocks	Number of Strata	Number of Blocks
1992	3	12	2	11
1993	3	18	2	9
1994	3	14	2	9
1995	3	13	2	10
1996	3	21	2	11
1997	3	20	2	17
2001	1	5	2	9
2002	3	19	2	14
2004	3	18	2	10

Table 2. Gag grouper stratified means for the entire Gulf of Mexico video survey area, and by region of the Gulf (n=number of blocks).

	Gulf of Mexico			Eastern Gulf of Mexico			Western Gulf of Mexico		
YEAR	n	MEAN	SE	n	MEAN	SE	n	MEAN	SE
1992	23	0.000	0.0000	12	0.000	0.0000	11	0.000	0.0000
1993	27	0.036	0.0622	18	0.052	0.0886	9	0.000	0.0000
1994	23	0.039	0.0853	14	0.056	0.1214	9	0.000	0.0000
1995	23	0.047	0.0806	13	0.066	0.1147	10	0.000	0.0000
1996	32	0.060	0.0918	21	0.083	0.1254	11	0.004	0.0123
1997	37	0.067	0.1074	20	0.075	0.0989	17	0.047	0.1275
2001	14	-	-	5	-	-	9	0.000	0.000
2002	33	0.145	0.1378	19	0.206	0.1962	14	0.000	0.0000
2004	28	0.183	0.1866	18	0.243	0.2316	10	0.040	0.0803

Table 3. Gag Grouper stratified means standardized to the mean of the survey time series (n=number of blocks, 2001 was not included in time series mean for the Gulf of Mexico and Eastern Gulf, but was used in the Western Gulf).

	Gulf of Mexico			Eastern Gulf of Mexico			Western Gulf of Mexico		
YEAR	n	Standardized Mean	SE	n	Standardized Mean	SE	n	Standardized Mean	SE
1992	23	0.000	0.0000	12	0.000	0.0000	11	0.000	0.0000
1993	27	0.504	0.8643	18	0.529	0.9071	9	0.000	0.0000
1994	23	0.545	1.1850	14	0.572	1.2437	9	0.000	0.0000
1995	23	0.647	1.1192	13	0.679	1.1747	10	0.000	0.0000
1996	32	0.830	1.2745	21	0.852	1.2842	11	0.429	1.2140
1997	37	0.929	1.4914	20	0.770	1.0124	17	4.659	12.5412
2001	14	-	-	5	-	-	9	0.000	0.000
2002	33	2.009	1.9143	19	2.109	2.0093	14	0.000	0.000
2004	28	2.536	2.5916	18	2.489	2.3718	10	3.916	7.897

Table 4. Numbers of video stations viewed and summed mincounts of gag and CBG observed during SEAMAP surveys.

Year	*Total Number of Video Stations*	*Number of Videos with Gag*	*Summed Mincounts of Videos with Gag*	*Number of Videos Reviewed for CBG*	*Number of Videos Lost During Hurricane*	*Summed Mincounts of Videos with Gag*	*Summed Mincounts of Videos with CBG*	*Proportion of CBG*
1992	133	0	.	.	.	.	.	.
1993	167	6	8	6	0	8	7	0.88
1994	121	3	3	3	0	3	3	1.00
1995	238	6	8	6	0	8	4	0.50
1996	289	12	14	12	0	14	6	0.43
1997	258	14	17	8	6	8	3	0.38
2001	77	0	.	.	.	.	.	.
2002	261	28	44	20	8	33	15	0.45
2004	196	31	45	17	14	26	16	0.62

Table 5. Results of Type 3 analyses for those variables retained in the binomial sub-model.

Type 3 Tests of Fixed Effects

Effect	*Num DF*	*Den DF*	*Chi-Square*	*F Value*	*Pr > ChiSq*	*Pr > F*
Year	6	595	32.12	5.32	<.0001	<.0001
Region	1	1150	34.71	34.71	<.0001	<.0001
Station Depth	1	1113	27.40	27.40	<.0001	<.0001

Table 6. Results of Type 3 analyses for those variables retained in the lognormal sub-model.

Type 3 Tests of Fixed Effects

Effect	*Num DF*	*Den DF*	*F Value*	*Pr > F*
Year	6	90	0.48	0.8245
Region	1	90	0.53	0.4667
Station Depth	1	90	2.56	0.1129

Table 7. Index values for gag Gulfwide. Frequency is the nominal frequency of occurrence, N is the number of video stations, Lo Index is the index in mincounts, Standardized Index is that same index scaled to a mean of one, CV is the coefficient of variation on the mean, and LCL and UCL are lower and upper 95% confidence limits.

Survey Year	*Frequency*	*N*	*Lo Index*	*Standardized Index*	*CV*	*LCL*	*UCL*
1993	0.03593	167	0.020334	0.67757	1.05936	0.11953	3.84071
1994	0.02479	121	0.011495	0.38302	1.61989	0.03959	3.70539
1995	0.02521	238	0.011797	0.39312	1.23670	0.05725	2.69930
1996	0.04152	289	0.022372	0.74548	0.75569	0.19437	2.85916
1997	0.05039	258	0.029651	0.98803	0.75345	0.25844	3.77729
2002	0.10728	261	0.048218	1.60674	0.45148	0.67894	3.80240
2004	0.15816	196	0.066204	2.20604	0.43907	0.95262	5.10869

Table 8. Index values for gag Gulfwide based on modeled frequency of occurrence. Standardized Index is based on the Modeled Frequency of Occurrence index scaled to a mean of one, CV is the coefficient of variation on the mean, and LCL and UCL are lower and upper 95% confidence limits.

Year	*Modeled Average Nonzero Mincount*	*Modeled Frequency of Occurrence*	*Standardized Index*	*CV*	*LCL*	*UCL*
1993	1.37532	0.014896	0.611709	0.449852	0.251238	1.470632
1994	1.04587	0.011322	0.464942	0.501413	0.172228	1.239845
1995	1.21379	0.009810	0.402851	0.384098	0.188983	0.853708
1996	1.08164	0.020528	0.84299	0.347623	0.424246	1.658137
1997	1.18124	0.024974	1.025566	0.389325	0.474428	2.182582
2002	1.28991	0.036822	1.512108	0.288116	0.854447	2.642638
2004	1.25062	0.052108	2.139833	0.316804	1.138619	3.934759

Table 9. Results of Type 3 analyses for those variables retained in the binomial sub-model.

Type 3 Tests of Fixed Effects

Effect	*Num DF*	*Den DF*	*Chi-Square*	*F Value*	*Pr > ChiSq*	*Pr > F*
YEAR	6	373	24.08	3.98	0.0005	0.0007
Station Depth	1	683	20.01	20.01	<.0001	<.0001

Table 10. Results of Type 3 analyses for those variables retained in the lognormal sub-model.

Type 3 Tests of Fixed Effects

Effect	*Num DF*	*Den DF*	*F Value*	*Pr > F*
YEAR	6	84	0.51	0.8020
Station Depth	1	84	2.26	0.1365

Table 11. Index values for gag from the East Gulf. Frequency is the nominal frequency of occurrence, N is the number of video stations, Lo Index is the index in mincounts, Standardized Index is that same index scaled to a mean of one, CV is the coefficient of variation on the mean, and LCL and UCL are lower and upper 95% confidence limits.

Survey Year	*Frequency*	*N*	*Lo Index*	*Standardized Index*	*CV*	*LCL*	*UCL*
1993	0.05000	120	0.07400	0.73316	0.66580	0.21830	2.46230
1994	0.03947	76	0.04266	0.42263	0.99930	0.08002	2.23225
1995	0.03871	155	0.04401	0.43600	0.71152	0.12122	1.56822
1996	0.08029	137	0.07854	0.77819	0.45304	0.32794	1.84660
1997	0.06757	148	0.08758	0.86769	0.48217	0.34773	2.16514
2002	0.16471	170	0.17065	1.69077	0.25315	1.02709	2.78330
2004	0.19310	145	0.20908	2.07155	0.24110	1.28773	3.33246

Table 12. Index values for gag from the East Gulf based on modeled frequency of occurrence. Standardized Index is based on the Modeled Frequency of Occurrence index scaled to a mean of one, CV is the coefficient of variation on the mean, and LCL and UCL are lower and upper 95% confidence limits.

Year	*Modeled Average Mincount*	*Modeled Frequency of Occurrence*	*Standardized Index*	*CV*	*LCL*	*UCL*
1993	1.37017	0.05391	0.662867	0.424467	0.281574	1.49603
1994	1.04498	0.04169	0.512612	0.528376	0.17583	1.41869
1995	1.21343	0.03627	0.445969	0.360877	0.216898	0.8987
1996	1.08033	0.07152	0.879396	0.28771	0.492816	1.529967
1997	1.14162	0.0758	0.932022	0.30967	0.498964	1.687107
2002	1.28984	0.12904	1.58665	0.189639	1.081539	2.279765
2004	1.26581	0.16107	1.980485	0.186472	1.354506	2.821026

Table 13. Summary of parameters for those variables in the zero-inflated binomial sub-model.

Parameter Estimates

Parameter	*Estimate*	*Standard Error*	*DF*	*t Value*	*Pr > |t|*	*Alpha*	*Lower*	*Upper*	*Gradient*
a0_est	-3.7063	0.5440	929	-6.81	<.0001	0.05	-4.7739	-2.6388	0.008792
adepth_est	1.2111	0.3336	929	3.63	0.0003	0.05	0.5563	1.8658	0.018058
a1993_est	-0.4915	0.5224	929	-0.94	0.3470	0.05	-1.5168	0.5337	-0.00564
a1994_est	-0.8386	0.6708	929	-1.25	0.2116	0.05	-2.1551	0.4779	-0.00235
a1995_est	-1.1154	0.5960	929	-1.87	0.0616	0.05	-2.2851	0.05438	-0.0054
a1996_est	-1.0025	0.5471	929	-1.83	0.0672	0.05	-2.0762	0.07108	0.008468
a1997_est	-1.2856	0.6668	929	-1.93	0.0542	0.05	-2.5943	0.02300	0.006935
a2002_est	-0.7599	0.4717	929	-1.61	0.1075	0.05	-1.6856	0.1657	-0.00097
b0_est	-2.3178	0.08062	929	-28.75	<.0001	0.05	-2.4760	-2.1596	-0.00078
b1993_est	-0.6292	0.1901	929	-3.31	0.0010	0.05	-1.0024	-0.2560	-0.00755
b1994_est	-0.9340	0.3830	929	-2.44	0.0149	0.05	-1.6855	-0.1824	0.008138
b1995_est	-1.3325	0.2755	929	-4.84	<.0001	0.05	-1.8732	-0.7918	-0.00251
b1996_est	-0.9623	0.2223	929	-4.33	<.0001	0.05	-1.3986	-0.5259	0.000235
b1997_est	-1.6222	0.3802	929	-4.27	<.0001	0.05	-2.3683	-0.8761	-0.00164
b2002_est	-0.6908	0.1514	929	-4.56	<.0001	0.05	-0.9880	-0.3936	-0.00877

Table 14. Results of Type 3 analyses for those variables retained in the lognormal sub-model.

Type 3 Tests of Fixed Effects

Effect	*Num DF*	*Den DF*	*F Value*	*Pr > F*
YEAR	6	35	2.79	0.0253

Table 15. Index values for CBG from the East Gulf. Frequency is the nominal frequency of occurrence, N is the number of video stations, Lo Index is the index in mincounts, Standardized Index is that same index scaled to a mean of one, CV is the coefficient of variation on the mean, and LCL and UCL are lower and upper 95% confidence limits.

Survey Year	*Frequency*	*N*	*Lo Index*	*Standardized Index*	*CV*	*LCL*	*UCL*
1993	0.050000	120	0.052102	1.18406	0.77120	0.30198	4.64267
1994	0.039474	76	0.032578	0.74037	1.32096	0.09923	5.52415
1995	0.025806	155	0.025937	0.58944	1.26441	0.08348	4.16165
1996	0.036496	137	0.029350	0.66701	1.10515	0.11173	3.98203
1997	0.020548	146	0.021034	0.47802	1.51108	0.05400	4.23171
2002	0.049383	162	0.061178	1.39032	0.68568	0.40171	4.81194
2004	0.097744	133	0.085840	1.95078	0.47902	0.78601	4.84159

Table 16. Index values for CBG from the East Gulf based on modeled frequency of occurrence. Standardized Index is based on the Modeled Frequency of Occurrence index scaled to a mean of one, CV is the coefficient of variation on the mean, and LCL and UCL are lower and upper 95% confidence limits.

Year	*Modeled Average Mincount*	*Modeled Frequency of Occurrence*	*Standardized Index*	*CV*	*LCL*	*UCL*
1993	1.08483	0.047922	1.243730	0.40344727	0.258961	2.228603
1994	1.0109	0.032506	0.843636	0.58592260	0.000000	1.813871
1995	1.00893	0.025829	0.670346	0.49734794	0.016065	1.324653
1996	1.00756	0.029188	0.757523	0.45720844	0.077808	1.437290
1997	1.00930	0.020955	0.543850	0.57380100	0.000000	1.156216
2002	1.64338	0.037126	0.963539	0.37100684	0.261868	1.665159
2004	1.11828	0.076190	1.977376	0.29729623	0.823755	3.129959

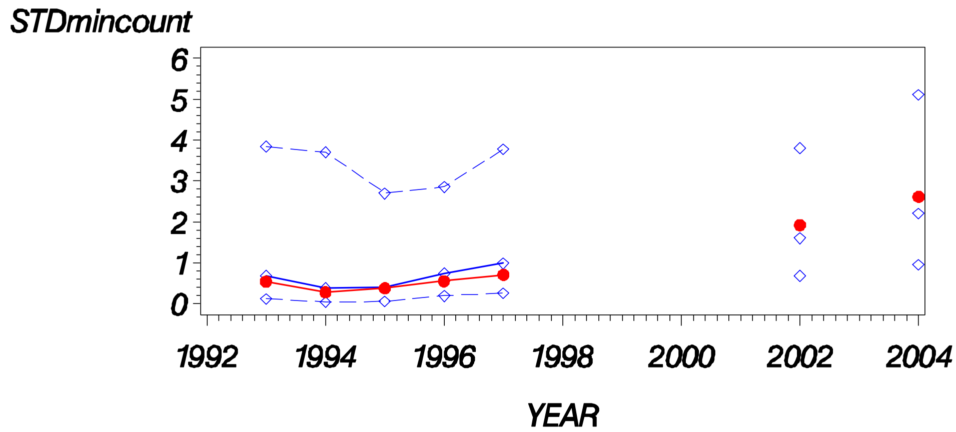

Figure 4. Gulfwide gag index values. STDmincount is synonymous with Standardized Index in Table 7. Solid circles indicate nominal index values, while open circles indicate model-based index values with corresponding lower and upper 95% confidence limits.

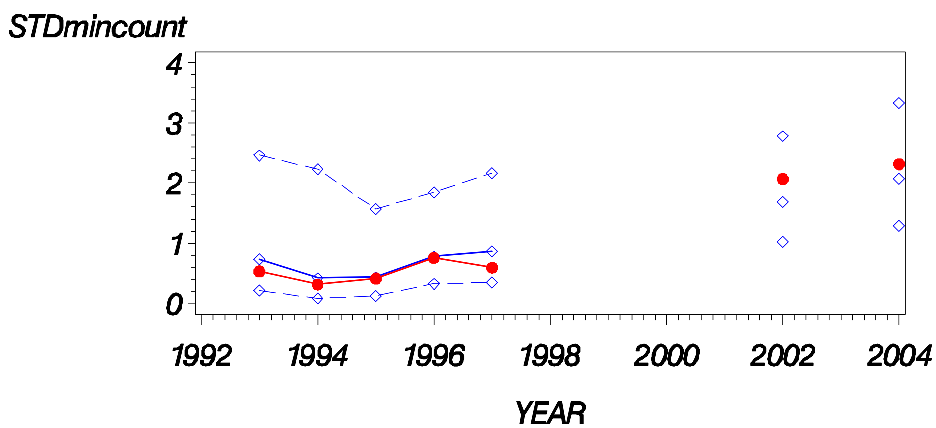

Figure 5. East Gulf gag index values. STDmincount is synonymous with Standardized Index in Table 11. Solid circles indicate nominal index values, while open circles indicate model-based index values with corresponding lower and upper 95% confidence limits.

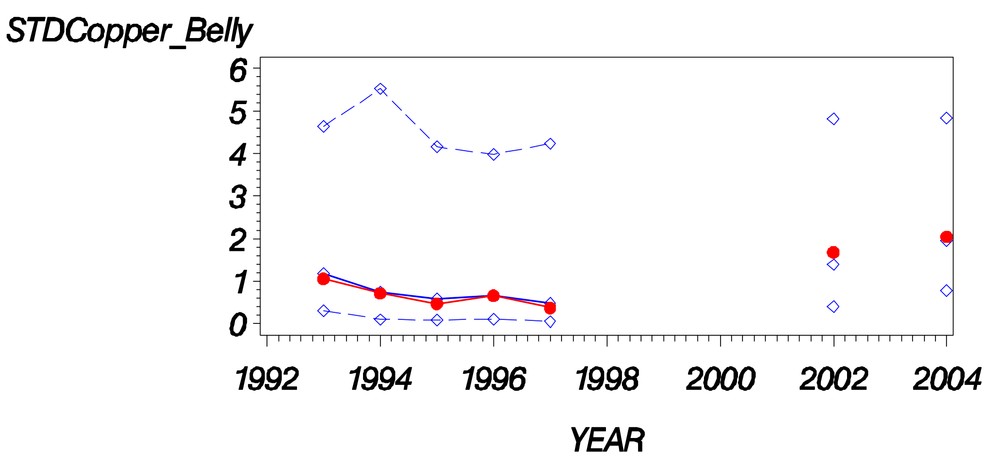

Figure 6. East Gulf CBG index values. STDCopper_Belly is synonymous with Standardized Index in Table 15. Solid circles indicate nominal index values, while open circles indicate model-based index values with corresponding lower and upper 95% confidence limits.

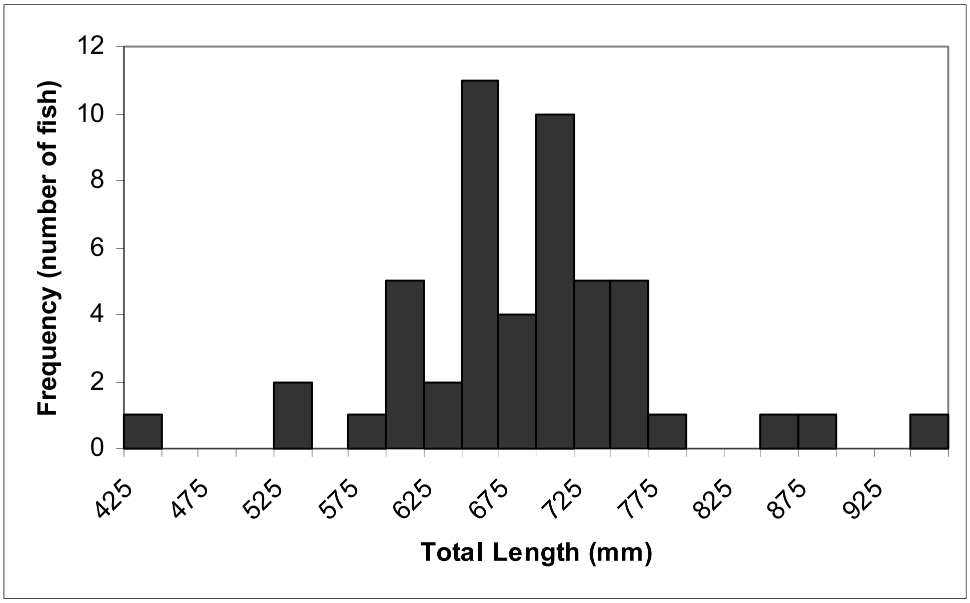

Figure 7. Length frequency histogram for gag hit by lasers during SEAMAP surveys (N = 50). One gag was measured in 1996, 7 in 2002 and 42 in 2004.

II. Gag video index at the Madison-Swanson and Steamboat Lumps Protected Areas and the Twin Ridges control site.

INTRODUCTION

The Madison-Swanson and Steamboat Lumps marine reserves on the West Florida Shelf were established by the Gulf of Mexico Fishery Management Council (GMFMC) in 1999 primarily to protect spawning aggregations of gag grouper (*Mycteroperca microlepis*) (Figure 8). The area closures went into effect on June 19, 2000. At the July, 2003 of the GMFMC, the reserve designation was extended until June 2010. However, fishing by surface trolling for coastal pelagic and highly migratory species will be allowed from May to October. The National Marine Fisheries Service has conducted stationary underwater video surveys to established baseline estimates of abundance of all fish associated with the major bottom features of each MPA. Special attention has been placed on all commercially important species, especially groupers and snappers.

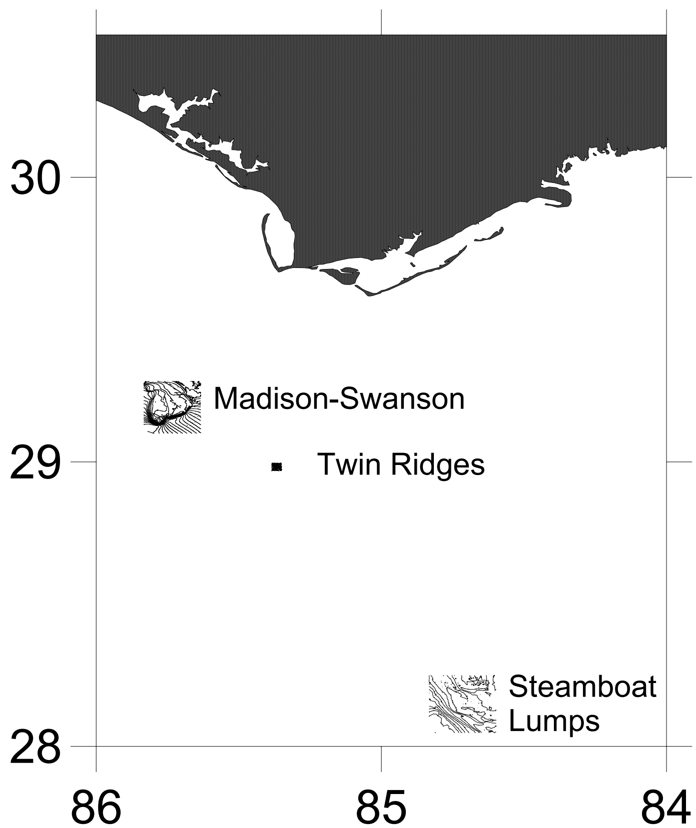

Figure 8. Location of Madison-Swanson, Steamboat Lumps and Twin Ridges.

METHODS

The Madison-Swanson and Steamboat Lumps MPAs were mapped using side-scan sonar (Scanlon et. al. USGS, Woods Hole, unpublished data.), and multibeam sonar (Gardner et al., 2002; Naar and Donahue, 2002). The MPAs were stratified based on bathymetry, acoustic backscatter, and features apparent on multibeam and side-scan mosaics. Madison-Swanson was divided into five strata: 1-Northeast (51 km²), 2-Ridge (4 km²), 3-Central (140 km²) , 4-Pinnacles (18 km²), 5- Mounds (1 km²); and Steamboat Lumps divided into five strata: 1-Northeast (105 km²), 2-Pits (1 km²), 3-Central (200 km²), 4-South Paleo-shore (12 km²), and 5- Ridge (3 km²) (Figures 9 and 10). The Twin Ridges feature has been sampled as a control site. This area is similar to the ridge feature within Madison-Swanson (Figure 11). Video gear and sampling methodology follows that employed for the SEAMAP survey. Sample sites within the two closed areas were selected using stratified random sampling. Sites within the Twin Ridges were selected randomly. Stations were sampled from February through July.

Following the design-based approach described in the previous section, gag abundance was estimated using either stratified or simple random estimators (Cochran, 1977). Also, following similar methodologies as described above, model-based estimators were developed using year, feature, stratum, and habitat variables. For model based estimators, we dropped data from June and July (42 stations) in order to better develop abundance indices of spawning aggregations in the MPAs. This was based on gonosomatic index data from Collins et al. (1998), which indicated a spawning period from January through May for gag in the eastern Gulf of Mexico. Habitat variables were developed for each sub-model using canonical linear discriminate function analyses (PROC CANDISC, SAS v. 9.1) following the methods of Kshirsagar (1972) and Rao (1973). In order to develop a linear combination of habitat variables [i.e., station depth, relief, proportion of bottom that is sand and clay, proportion of bottom that is rock, proportion of bottom that is attached epifauna (i.e., sponge, hardcoral, unknown sessiles, algae, softcoral, and seawhips)] to be used in the binomial submodel, the CANDISC was run using occurrence (i.e., presence/absence) as a class variable. In order to develop a linear combination of habitat variables to be used in the lognormal submodel, the CANDISC was run using nonzero mincount as a class variable. Also, CANDISC was run using all mincount data (including zero data) to discern associations between gag mincount and habitat type.

Finally, the size data of gag observed during the MPA surveys come from fish measured on video tape with lasers. We developed a length frequency histogram for gag measured by laser on video.

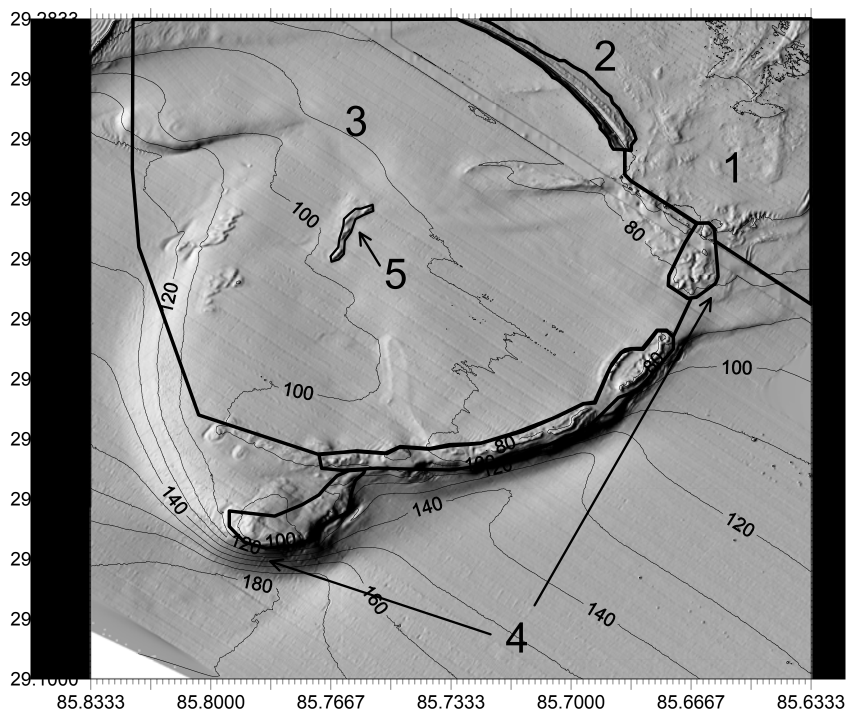

Figure 9. Madison-Swanson closed area with stratum boundaries. Isobaths are in meters.

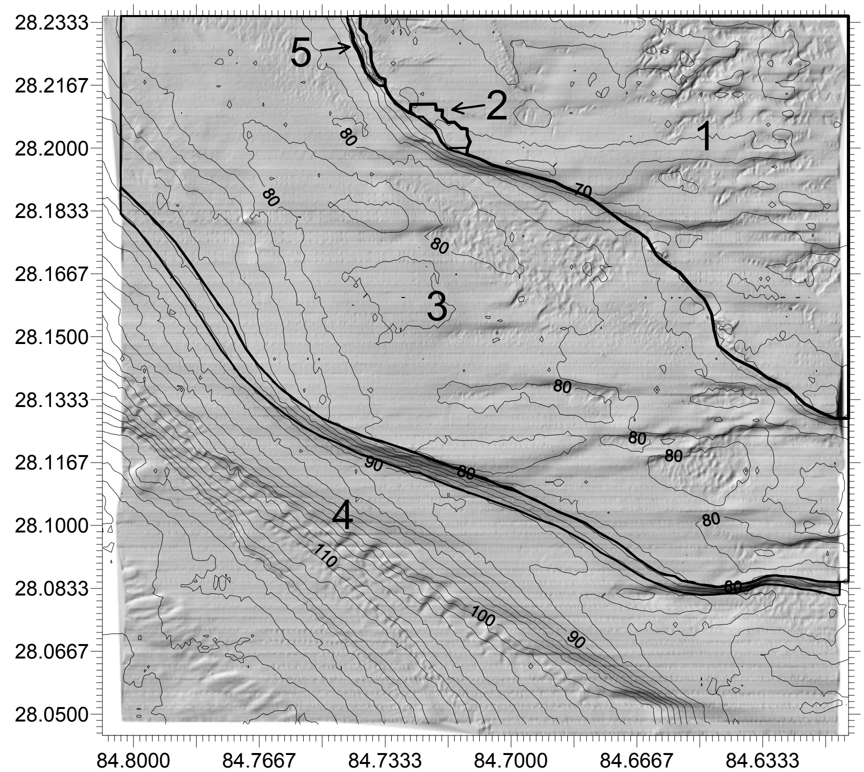

Figure 10. Steamboat Lumps closed area with stratum boundaries. Isobaths are in meters.

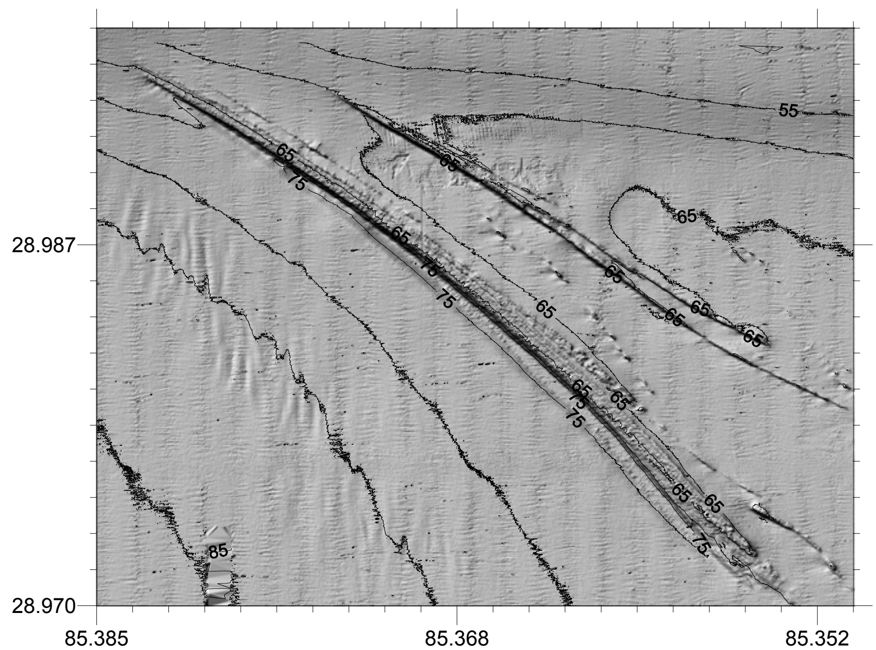

Figure 11. Twin Ridges control area.

RESULTS

Table 17 summarizes the results from the design-based and random estimators. Results indicate a low occurrence of gag in videos for both Twin Ridges and Steamboat Lumps areas, while the Madison-Swanson closed area had more occurrences of gag for all years but 2003 when no gag were observed. All design-based indices were highly variable.

Table 18 summarizes the multivariate statistics that indicate there was no significant difference between centroids of different mincounts (i.e., 0, 1, 2, 3, 4, 5), and Table 19 summarizes the eigen values and correlation statistics for the CANDISC procedure ran with all mincount and habitat data. Table 20 summarizes the total canonical structure of the first two canonical variables, which indicate gradients between low relief sandy bottom and high relief live bottom. Table 21 provides canonical coefficients for each habitat variable, and Table 22 class means for each level of mincount with each canonical variable. Figure 12 is a plot of the first two canonical variables at its corresponding mincount value. This plot indicates that the majority of gag were seen over shallower live bottom habitats than over deeper sandy bottom habitats.

Due to the large numbers of zeros in most years at other areas, only data from the Madison-Swanson closed area was used to develop model-based indices of abundance for MPAs. In order to produce a linear combination of habitat variables to be used in the binomial protion of the model, the CANDISC procedure was run with occurrence data (i.e., presence/ absence). Table 23 indicates there was a significant difference between centroids of stations where gag occurred versus those where they did not. Tables 24-27 summarize the CANDISC statistics as previously described. Figure 13 indicates as before that gag are observed more over high relief, live bottom than low relief, sandy bottom. Next, in order to produce a linear combination of habitat variables to be used in the lognormal protion of the model, the CANDISC procedure was run with nonzero mincount data (i.e., mincount = 1, 2, 3, 4, or 5). Table 28 summarizes the multivariate statistics that indicate there was no significant difference between centroids of different nonzero mincounts. Tables 24-27 summarize the CANDISC statistics as previously described. Figure 14 is a plot of the first two canonical variables at its corresponding mincount value. The canonical variables produced by the two latter procedures described above were used in the corresponding sub-models to develop a delta-lognormal gag MPA index of abundance.

Tables 33-35 summarize the model-building process by which the index was developed for gag from the Madison-Swanson MPA. Table 33 summarizes the results of Type 3 analyses for those variables retained in the binomial sub-model. Table 34 summarizes the results of Type 3 analyses for those variables retained in the lognormal sub-model. Table 35 and Figure 15 summarize the index values for gag from the Madison-Swanson MPA.

Forty-eight gag were hit by lasers during MPA surveys (4 in 2002, 39 in 2003, and 5 in 2004). The mean total length (±SE) was 673 (±17). However, the length frequency histogram in Figure 16 appears to be bimodal with modes at approximately 550 and 750 mm total length.

Table 17. Estimates of mean gag abundance at Madison-Swanson, Steamboat Lumps and Twin Ridges by stratum ($\bar{x}_h$) and stratified means ($\bar{x}_{st}$).

Madison-Swanson

	2001			2002			2003			2004		
Stratum	n_h	$\bar{x}_h$	S^2_h	n_h	$\bar{x}_h$	S^2_h	n_h	$\bar{x}_h$	S^2_h	n_h	$\bar{x}_h$	S^2_h
1	6	0.000	0.0000	3	0.000	0.0000	5	0.000	0.0000	10	0.000	0.0000
2	10	0.000	0.0000	15	0.533	0.4095	21	0.000	0.0000	9	0.111	0.1111
3	4	0.250	0.2500	1	1.000	-	6	0.000	0.0000	7	0.000	0.0000
4	10	0.178	0.1778	24	1.3750	1.0810	18	0.000	0.0000	18	0.556	1.0850
5	4	0.000	0.0000	2	3.500	4.5000	8	0.000	0.0000	6	0.500	0.7000
		$\bar{x}_{st}$	se_{st}		$\bar{x}_{st}$	se_{st}		$\bar{x}_{st}$	se_{st}		$\bar{x}_{st}$	se_{st}
Stratified		0.180	0.1748		0.7962	0.0332		0.000	0.0000		0.0511	0.0243

Steamboat Lumps

	2001			2002			2003			2003		
Stratum	n_h	$\bar{x}_h$	S^2_h	n_h	$\bar{x}_h$	S^2_h	n_h	$\bar{x}_h$	S^2_h	n_h	$\bar{x}_h$	S^2_h
1	14	0.000	0.0000	3	0.000	0.0000						
2	4	0.000	0.0000	5	0.000	0.0000				3	0.000	0.0000
3	17	0.000	0.0000	4	0.000	0.0000						
4	5	0.000	0.0000	2	0.000	0.0000						
5	3	0.333	0.1111	5	0.200	0.0400	5	0.000	0.0000	2	0.000	0.0000
		$\bar{x}_{st}$	se_{st}		$\bar{x}_{st}$	se_{st}		$\bar{x}_{st}$	se_{st}		$\bar{x}_{st}$	se_{st}
Stratified		0.0031	0.0031		0.0019	0.0019						

Twin Ridges

	2001			2002			2003			2004		
	n	$\bar{x}$	SE	n	$\bar{x}$	SE	n	$\bar{x}$	SE	n	$\bar{x}$	SE
	15	0.000	0.0000	7	0.429	0.2020	14	0.000	0.000	12	0.000	0.0000

Table 18. Multivariate statistics and F approximations for the CANDISC procedure ran with all mincount and habitat data.

S=5 M=2 N=48.5

Statistic	*Value*	*F Value*	*Num DF*	*Den DF*	*Pr > F*
Wilks' Lambda	0.58109196	1.15	50	454.87	0.2325
Pillai's Trace	0.50115530	1.15	50	515	0.2342
Hotelling-Lawley Trace	0.59036501	1.15	50	305.82	0.2355
Roy's Greatest Root	0.26759820	2.76	10	103	0.0047

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

Table 19. Eigen values and correlation statistics for the CANDISC procedure ran with all mincount and habitat data.

	*Eigenvalues of Inv(E)*H = CanRsq/(1-CanRsq)*						
	Eigenvalue	*Difference*	*Proportion*	*Cumulative*	*Canonical Correlation*	*Approximate Standard Error*	*Squared Canonical Correlation*
1	0.2676	0.1067	0.4533	0.4533	0.459463	0.074213	0.211106
2	0.1609	0.0697	0.2725	0.7258	0.372267	0.081035	0.138582
3	0.0912	0.0410	0.1544	0.8802	0.289029	0.086213	0.083538
4	0.0501	0.0295	0.0849	0.9651	0.218466	0.089582	0.047727
5	0.0206		0.0349	1.0000	0.142131	0.092172	0.020201

Table 20. Total canonical structure for the CANDISC procedure ran with all mincount and habitat data.

Variable	*Can1*	*Can2*
ST_DEPTH	0.005349	0.686833
SANDCLAY	-0.643102	0.225444
ROCK	0.410963	0.091164
MAXIMUM_RELIEF	0.897737	-0.098071
SPONGE	0.221485	-0.025813
HARDCORAL	0.320631	-0.558874
UKSESSILES	0.335990	-0.370790
ALGAE	0.271566	0.039672
SOFTCORAL	0.362246	-0.427518
SEAWHIPS	0.291156	-0.484766

Table 21. Total-sample standardized canonical coefficients for the CANDISC procedure ran with all mincount and habitat data.

Variable	*Can1*	*Can2*
ST_DEPTH	0.406617869	0.731770746
SANDCLAY	-0.240879777	0.123770808
ROCK	-0.199300722	0.371039643
MAXIMUM_RELIEF	1.017639184	0.213009852
SPONGE	0.079567413	0.037178354
HARDCORAL	0.209580498	-0.526923683
UKSESSILES	0.081982667	-0.252190176
ALGAE	0.115387146	0.204607553
SOFTCORAL	-0.084811519	-0.087522336
SEAWHIPS	-0.003591455	-0.159502790

Table 22. Class means on canonical variables for the CANDISC procedure ran with all mincount and habitat data.

MINCOUNT	*Can1*	*Can2*
0	-0.347136824	0.036784775
1	0.654754447	-0.545293824
2	0.698292174	0.956525719
3	0.755153773	0.514458969
4	0.143378122	-0.395585579
5	1.938822689	1.259614954

Table 23. Multivariate statistics and exact F statistics for the CANDISC procedure ran with occurrence and habitat data.

S=1 M=4 N=50.5

Statistic	*Value*	*F Value*	*Num DF*	*Den DF*	*Pr > F*
Wilks' Lambda	0.80258734	2.53	10	103	0.0090
Pillai's Trace	0.19741266	2.53	10	103	0.0090
Hotelling-Lawley Trace	0.24597031	2.53	10	103	0.0090
Roy's Greatest Root	0.24597031	2.53	10	103	0.0090

Table 24. Eigen values and correlation statistics for the CANDISC procedure ran with occurrence and habitat data.

	*Eigenvalues of Inv(E)*H = CanRsq/(1-CanRsq)*						
	Eigenvalue	*Difference*	*Proportion*	*Cumulative*	*Canonical Correlation*	*Approximate Standard Error*	*Squared Canonical Correlation*
1	0.2460		1.0000	1.0000	0.444311	0.075501	0.197413

Table 25. Total canonical structure for the CANDISC procedure ran with occurrence and habitat data.

Variable	*Can1*
ST_DEPTH	-0.066419
SANDCLAY	-0.637496
ROCK	0.386269
MAXIMUM_RELIEF	0.899187
SPONGE	0.224626
HARDCORAL	0.356031
UKSESSILES	0.421216
ALGAE	0.283087
SOFTCORAL	0.409485
SEAWHIPS	0.371613

Table 26. Total-sample standardized canonical coefficients for the CANDISC procedure ran with occurrence and habitat data.

Variable	*Can1*
ST_DEPTH	0.3333058415
SANDCLAY	-0.0264932024
ROCK	-0.0978200398
MAXIMUM_RELIEF	0.9998150439
SPONGE	0.0968346490
HARDCORAL	0.2360675433
UKSESSILES	0.1946662425
ALGAE	0.1379970305
SOFTCORAL	0.0054777005
SEAWHIPS	0.0704757116

Table 27. Class means on canonical variables for the CANDISC procedure ran with occurrence and habitat data.

success	*Can1*
0	-0.3544864205
1	0.6817046547

Table 28. Multivariate statistics and F approximations for the CANDISC procedure ran with nonzero mincount and habitat data.

S=4 M=2.5 N=11.5

Statistic	*Value*	*F Value*	*Num DF*	*Den DF*	*Pr > F*
Wilks' Lambda	0.26506158	1.01	40	96.652	0.4658
Pillai's Trace	1.08109675	1.04	40	112	0.4283
Hotelling-Lawley Trace	1.67292797	1.00	40	58.698	0.4983
Roy's Greatest Root	0.81114976	2.27	10	28	0.0427

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

Table 29. Eigen values and correlation coefficients for the CANDISC procedure ran with nonzero mincount and habitat data.

	*Eigenvalues of Inv(E)*H = CanRsq/(1-CanRsq)*						
	Eigenvalue	*Difference*	*Proportion*	*Cumulative*	*Canonical Correlation*	*Approximate Standard Error*	*Squared Canonical Correlation*
1	0.8111	0.2932	0.4849	0.4849	0.669227	0.089568	0.447865
2	0.5180	0.3130	0.3096	0.7945	0.584155	0.106865	0.341237
3	0.2050	0.0662	0.1225	0.9170	0.412438	0.134627	0.170105
4	0.1388		0.0830	1.0000	0.349128	0.142448	0.121890

Table 30. Total canonical structure for the CANDISC procedure ran with nonzero mincount and habitat data.

Variable	*Can1*	*Can2*
ST_DEPTH	-0.648204	0.272371
SANDCLAY	0.037608	0.383210
ROCK	-0.220399	0.054445
MAXIMUM_RELIEF	-0.190857	-0.204630
SPONGE	0.075444	0.167113
HARDCORAL	0.270572	-0.140239
UKSESSILES	0.346039	-0.350104
ALGAE	0.007329	0.235162
SOFTCORAL	0.333352	-0.414761
SEAWHIPS	0.332946	-0.099992

Table 31. Total-sample standardized canonical coefficients for the CANDISC procedure ran with nonzero mincount and habitat data.

Variable	*Can1*	*Can2*
ST_DEPTH	-1.116321809	-0.036521563
SANDCLAY	1.561810809	3.072291424
ROCK	0.328655622	2.183106715
MAXIMUM_RELIEF	-0.083883967	0.231132219
SPONGE	0.499656982	0.439176013
HARDCORAL	0.452470303	0.031012093
UKSESSILES	0.614705079	0.241671625
ALGAE	0.362691996	1.012395765
SOFTCORAL	0.534255220	0.698904855
SEAWHIPS	-0.019562004	0.061564782

Table 32. Class means on canonical variables for the CANDISC procedure ran with nonzero mincount and habitat data.

MINCOUNT	*Can1*	*Can2*
1	0.490464589	-0.177321693
2	-0.964725499	0.889119870
3	-0.256721484	0.517541934
4	0.395940847	-0.777979026
5	-3.723818158	-2.770790692

Table 33. Results of Type 3 analyses for those variables retained in the binomial sub-model.

Type 3 Tests of Fixed Effects

Effect	*Num DF*	*Den DF*	*Chi-Square*	*F Value*	*Pr > ChiSq*	*Pr > F*
YEAR	67.6	12.73	6.30	0.0017	0.0031	67.6
Can1	110	11.41	11.41	0.0007	0.0010	110

Table 34. Results of Type 3 analyses for those variables retained in the lognormal sub-model.

Type 3 Tests of Fixed Effects

Effect	*Num DF*	*Den DF*	*F Value*	*Pr > F*
YEAR	2	35	0.71	0.4989
Can2	1	35	7.59	0.0093

Table 35. Index values for gag from the Madison-Swanson MPA. Frequency is the nominal frequency of occurrence, N is the number of video stations, Lo Index is the index in mincounts, Standardized Index is that same index scaled to a mean of one, CV is the coefficient of variation on the mean, and LCL and UCL are lower and upper 95% confidence limits.

Survey Year	*Frequency*	*N*	*Lo Index*	*Standardized Index*	*CV*	*LCL*	*UCL*
2001	0.12500	32	0.15942	0.34960	0.81103	0.08434	1.44913
2002	0.61905	42	0.87289	1.91421	0.23459	1.20491	3.04106
2003	0.00000	48	0.00000	0.00000			
2004	0.22500	40	0.33571	0.73620	0.52445	0.27469	1.97312

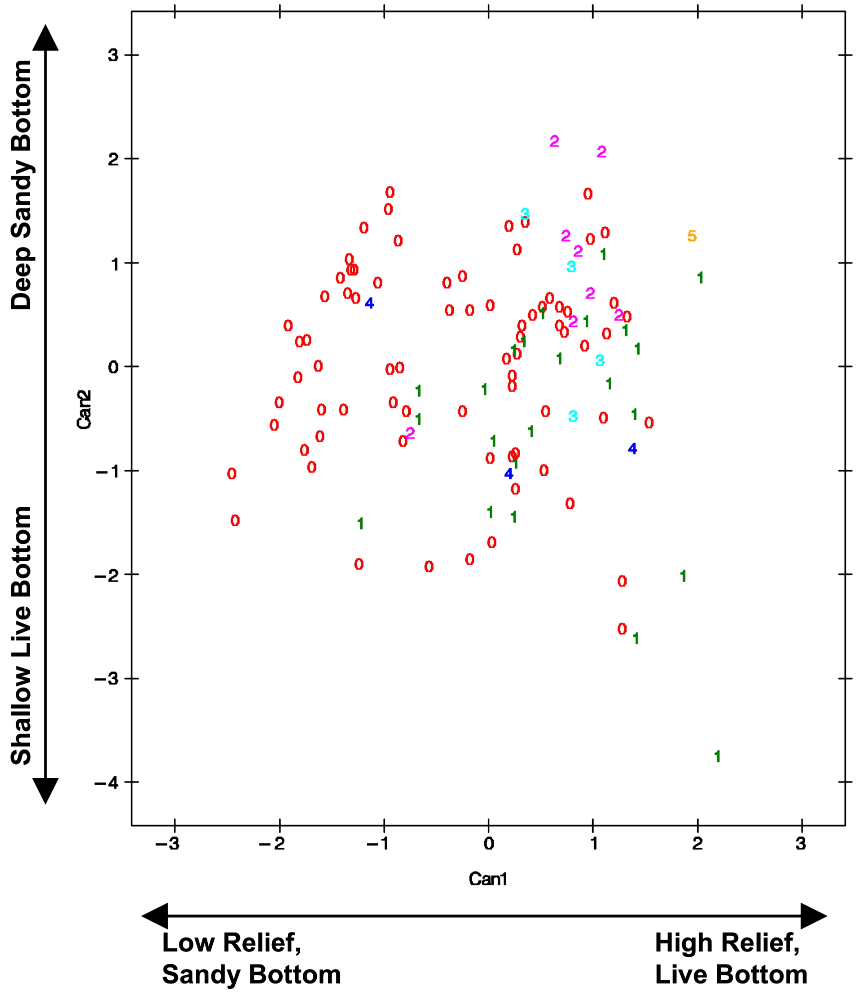

Figure 12. Plot of mincount by canonical variables 1 and 2.

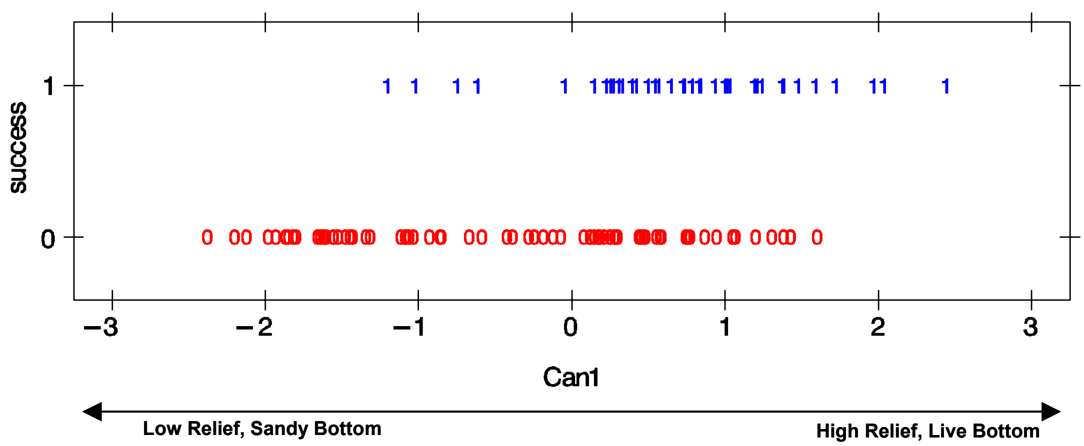

Figure 13. Plot of occurrence by canonical variable 1.

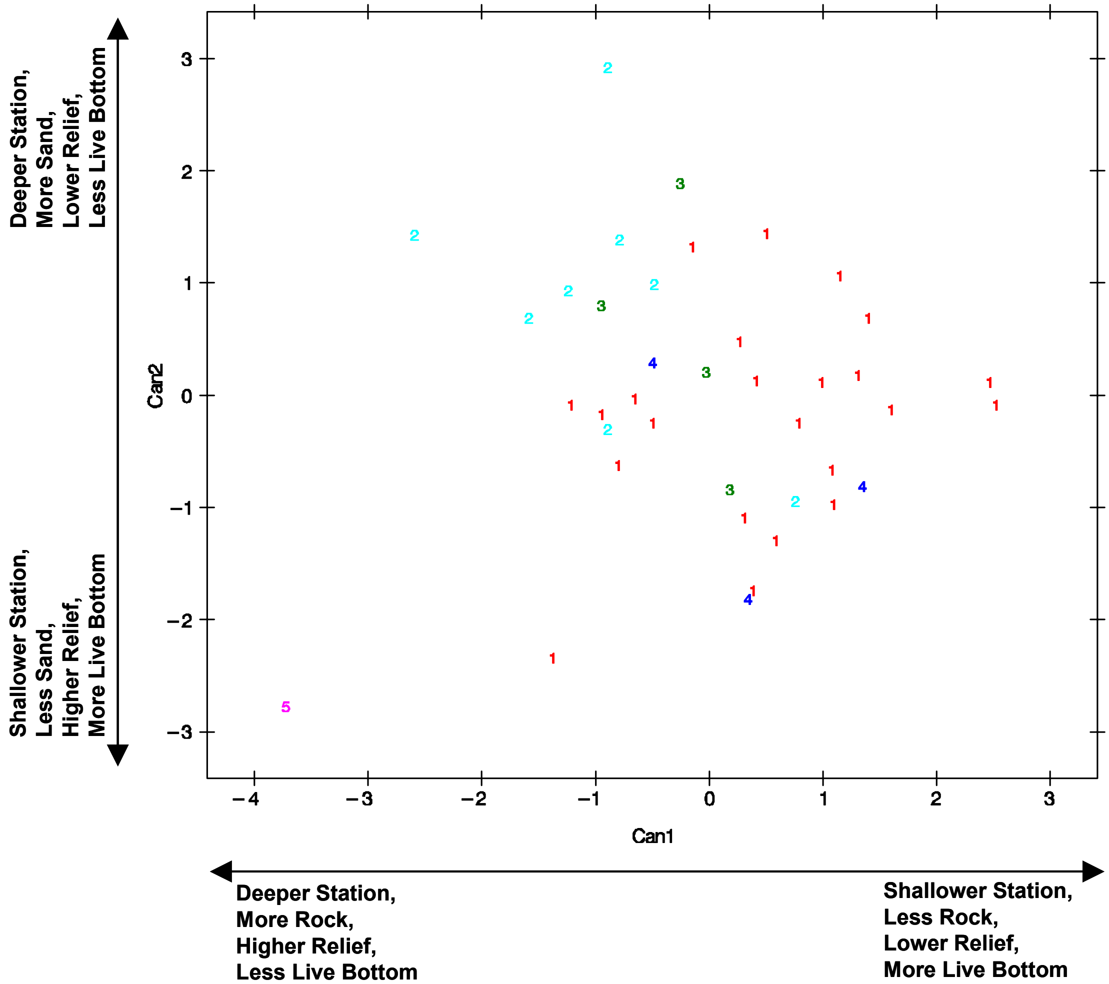

Figure 14. Plot of nonzero mincount by canonical variables 1 and 2.

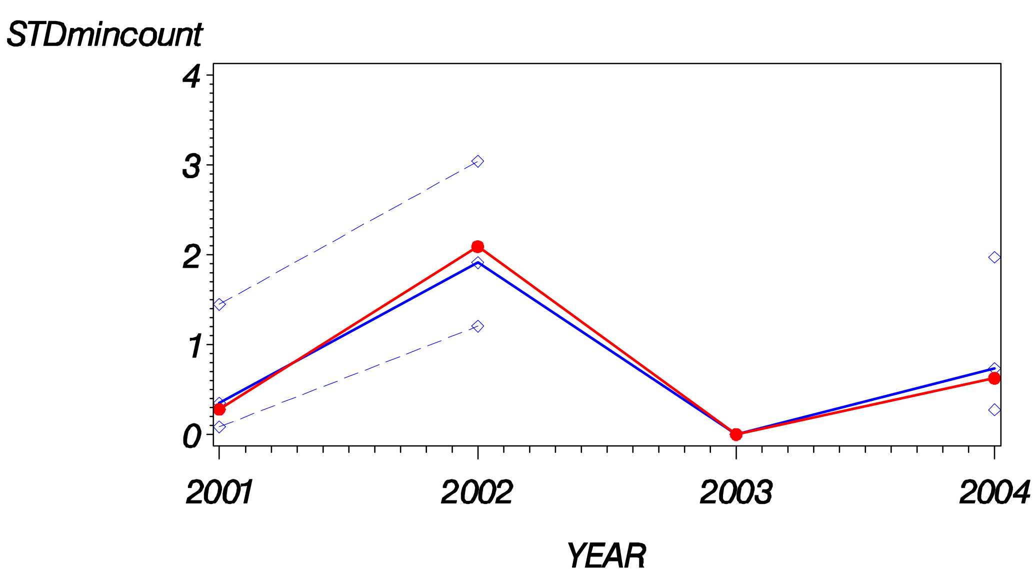

Figure 15. Madison-Swanson MPA gag abundance index values. STDmincount is synonymous with Standardized Index in the above table. Solid circles indicate nominal index values, while open circles indicate model-based index values with corresponding lower and upper confidence 95% limits.

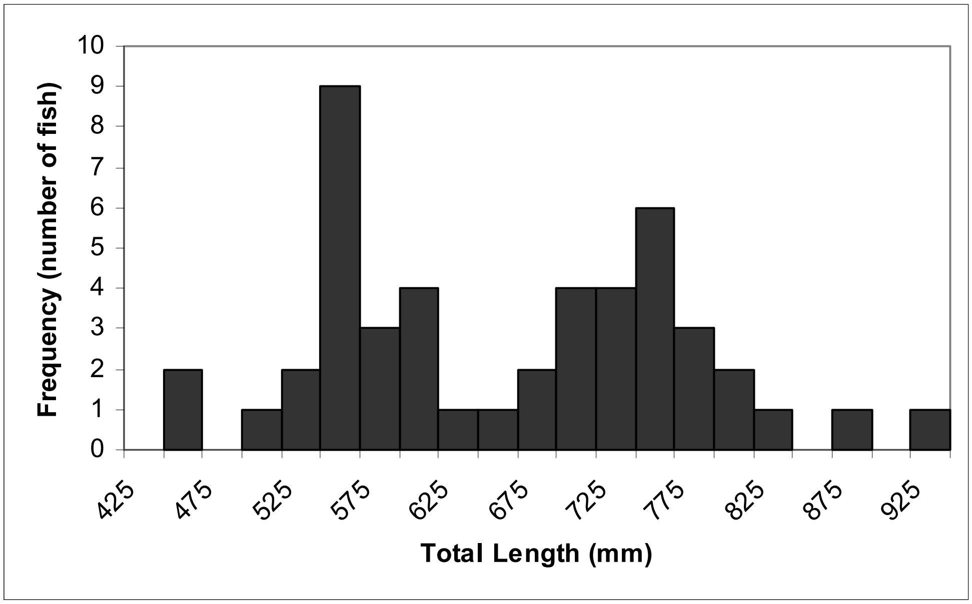

Figure 16. Length frequency histogram for gag hit by lasers during MPA surveys (N = 48). Four gag were measured in 2002, 39 in 2003, and 5 in 2004.

LITERATURE CITED

Bullock, L.H. and G.B. Smith. 1991. Memoirs of the Hourglass Cruises. Volume VIII. October 1991. Part II Seabasses (Pices: Serranidae). Fl. Mar. Res. Inst., Dept. of Nat. Res.

Cochran, W.G. 1977. *Sampling Techniques*. John Wiley & Sons. New York, NY. 428 p.

Collins, L.A., A.G. Johnson, C.C. Koenig, and M. S. Baker, Jr. 1998. Reproductive patterns, sex ratio, and fecundity in gag, *Mycteroperca microlepis* (Serranidae), a protogynous grouper from the northeastern Gulf of Mexico. Fish. Bull. 96: 415-427.

Gilmore, R.G. and R.S. Jones. 1992. Color variation and associated behavior in the Epinephelinae groupers, *Mycteroperca microlepis* (Goode and Bean) and *M. phenax* (Jordan and Swain). Bull. Mar. Sci. 51(1): 83-103.

Kshirsagar, A.M. 1972. *Multivariate Analysis*. Marcel Dekker, Inc., New York.

Lo, N. C. H., L.D. Jacobson, and J.L. Squire. 1992. Indices of relative abundance from fish spotter data based on delta-lognormal models. Can. J. Fish. Aquat. Sci. 49: 2515-1526.

Rao, C.R. 1973. *Linear Statistical Inference*. John Wiley & Sons. New York, NY.

Tyre, A.J., B. Tenhumberg, S.A. Field, D. Niejalke, K. Parris, and H.P. Possingham. 2003. Improving precision and reducing bias in biological surveys: estimating false-negative error rates. Ecological Applications 13: 1790-1801.