2012 ROV Distance Analysis

***Methods***

*Distance Data*-Only adults and subadults greater than 350 mm were included in the analysis. If subadults did not have length information, then they were excluded from the analysis. A total of 118 yelloweye were included in the density analysis in 2012, 112 adults and 6 subadults. Both the left and right side of the transect was sampled and this data were pooled together for the distance analysis. The sampling fraction for the ROV survey was equal to 1; for the submersible survey the sampling fraction was 0.5 and a correction factor was applied in order to obtain the density estimate using data from only one side of the transect line.

*Histograms*- Frequency histograms of the binned distance data were created in excel. Binned data was examined for any patterns that may indicate avoidance or attraction behavior by yelloweye rockfish and to determine which bins would produce the best model in the Distance program, such as a model with a shoulder and a decreasing frequency of observations with distance from the ROV. The following bins were examined: 0.5-ft, 0.75-ft, 1-ft, 1.25-ft, 1.5-ft, 1.75-ft, 2-ft, and 2.5-ft bins.

*Key functions and adjustment terms-* I explored the models of the half-normal hermite polynomial, half normal cosine, and the hazard rate cosine in my distance analyses. The uniform and negative exponential models were not considered good models for our 2012 data, because the negative exponential is generally used for salvaging poorly collected data and the uniform assumes that there is no decrease in probability to the effective width of this key function (T. Quinn ‘s class).

*Data binning*- I performed analyses with and without data binning. If no binning was used then the analyses were performed with the exact distance data. Data binning was explored to determine if it would improve the results for the analyses. Data binning may increase the robustness of the results. We used data binning for the submersible survey because observers would often unknowingly round their distances and because our distance estimates were not exact and had measurement error. For the ROV data we have close to exact measurements for distances to yelloweye; however, there may still be some error in distances due to the clarity of the video and our ability to identify the same exact point in both the left and right stereo cameras. We explored the following data bins in the Distance Program to determine which bin would produce the best fit of the data: 1-ft, 1.5-ft, 1.75-ft, 2-ft, 2.5-ft. For analyses that were performed without data binning I was able to specify bins for the chi-square test diagnostics; these bins only affect the results of the chi-square tests and do not affect the density estimate.

*Data Truncation* - Distance data are often truncated in order to prevent the tail of the model from overly influencing the model fit, because it is most important for the model to fit near the origin of the transect line. Terry Quinn suggests truncating 5-10% for distance data. For the submersible survey data, I usually truncated the data at some value at a large distance, but for the submersible survey fish were seen out to 30 ft. For the 2012 ROV survey, yelloweye rockfish (that were included in the density estimate) were only seen out to around 10 ft. I truncated the data after 9 ft for models without any binning and for the model with 1-ft bins; the data were truncated after 9-ft for these models because there were no observations at 10 ft but observations at 9 and 11 ft. For the models with intervals of 1.5-ft bins or greater, there were no bins that did not have observations and there was no bins at large distances that seemed to deviate from the models so no truncation was performed for these models.

*Choosing a model*

I examined the results of the Distance analyses to determine if the 2012 ROV data were able to produce a valid density estimate. It was determined if the data fit a model well by examining the fit visually and by examining diagnostic tests, such as the Q-q plot and the Chi-square and Kolmogorov-Smirnov (K-S) goodness of fit tests. In addition, the coefficient of variation (CV) was examined to determine if we were able to produce data with good precision.

The preferred model key function and adjustment term was chosen against other models with the same binning and truncation scheme by comparing the AIC values, model fits, and CV values between models. To determine if binning or truncation improved the model results and determine what would be the best binning scheme for the data, we compared the CV values and model fit. However, the AIC cannot be used to compare between models without the same binning or truncation schemes.

For visual examination, a model is preferred with a good fit at the origin, a shoulder, and a shape that is biologically realistic. The K-S and Chi-square goodness of fit tests were used to determine if the data fit the model well. The K-S test is considered to be a better goodness of fit test; however, it is only provided in the diagnostics if no binning was performed. If the K-S or chi-square p-values are insignificant and >0.05 it indicates the model is a good fit and there are no significant deviations from the model. The Q-q plot can also be used to determine if the model fits the data well for models with no binning. If the points are plotted along the diagonal-45 degree line of the empirical distribution function and the fitted cumulative distribution function it indicates the model is a good fit.

The AIC value was used to determine the model with the best key function and adjustment term for models with the same binning and truncation schemes. A lower AIC score is preferred and the AIC score incorporates the number of parameters giving a penalty for more parameters. The ∆AIC indicates whether the model with the lowest AIC is preferred over other models with a ∆AIC<2 indicating no credible evidence, ∆AIC 2–4 weak evidence, 4–7 definite evidence, 7–10 strong evidence, and >10 very strong evidence.

The precision of the density estimate was examined using the CV and the variance components. The CV was used to determine if a model had good precision. If a model has a CV<20% then the model is considered to have good precision. The variance of the density estimate is composed of the variance due to the detection function and the variance due to the encounter rate. As the model fit to the data improves then the variance due to the detection function decreases. The variance due to the encounter rate is due to the variability in the number of observations between transects.

***Results***

*Choosing a model*

The preferred model for our 2012 ROV data had the hazard rate key function with a cosine adjustment term and no binning or truncation. This model has a good fit to the data; visually the probability detection function has a wide shoulder and declining probability of detection at larger distances (Figure 3). In addition, the K-S and chi-square test P-values were not significant indicating a good model fit. Also, the Q-q plot had generally a good fit with only a little deviation in the middle of the plot. The model also had a CV=0.125 indicating there is high precision in the density estimate.

*Key function*-The hazard rate cosine models were preferred over the half normal cosine or half normal hermite models. The hazard rate cosine models had the lowest CV values, lowest AICs, highest *X*2 p-values, and best model fits (Tables 1–8).

The hazard rate model had the lowest AIC score when compared to the half normal cosine or the half normal hermite models for models with no binning and all models explored with binning (Tables 1–8). The ∆AIC was between 2-4, indicating weak evidence that the model with the lowest AIC is preferred, for the hazard rate cosine compared to the next best model for models with no binning and models with the binning schemes of 1-ft, 1.5-ft, and 2.5-ft. For the models with the 2-ft binning scheme, the ∆AIC= 5, indicating definite evidence to choose the hazard rate model over the model with the next lowest AIC. In addition, the hazard rate model had one less parameter generally than the half normal cosine model. The half normal hermite had two parameters as well but generally ran with errors that the parameters were highly correlated.

The hazard rate cosine models fit the data the best visually and using other goodness of fit tests compared to the half normal hermite or half normal cosine models. The hazard rate model has a larger shoulder at the origin that includes the observations that were made around 5 ft; whereas, the other models also have shoulders but have a decrease in probability which doesn’t include all of the observations that were made around 5 ft. For the model using exact data, Q-q plots and the K-S test were produced for comparison between the models (Table 1 and 6). The q-q plots were OK for the hazard rate cosine and for the half normal cosine but had quite a bit of deviation for the other models. The K-S P-value was also highest for the hazard rate model.

*Binning*- A model using exact data for the analyses was preferred over a model with binning. Our results indicate with the ROV data that it may not be necessary to bin since we are using distance measurements with little observer or measurement error. The CV was lowest and the *X*2 highest for the model with no binning; however, the model with 1-ft bins had very similar results (Table 1–6). Consequently a model with no binning would be preferred because exact data could be included in the analyses with no masking of the data. However, the difference in the density estimates was only 17 yelloweye/km2 for the preferred hazard rate model with no binning and the model with 1-ft bins.

*Truncation-* A model with no truncation was preferred over a model with truncation at the right tail of the model. For the hazard rate model with no binning and the model with 1-ft bins the *X*2 p-values were higher and the CV lower than with models with truncation after 9 ft (Table 1, 2, 6 and 7). However, for the model with no binning the K-S p-value was higher for the model with truncation. The K-S is a better goodness of fit test compared to the *X*2 test but the differences between the K-S, CV, and *X*2 were minimal between the model with truncation and the model with no truncation. In addition, there was only a difference in the density estimate of 2 yelloweye/sq km between these models. Consequently the model with no truncation would be preferred in order to retain all of the data in the analyses. We only observed yelloweye out to around 10 ft for the ROV survey so it seems appropriate to not truncate the data compared to the submersible survey where yelloweye were observed out to 30 ft. In addition, it doesn’t appear that the tail of the model is overly influencing the model fit for the ROV survey data, so truncation seems unnecessary.

*Comparison between 2007 and 2012 CSEO data results-*

The 2012 ROV survey results indicate that using the ROV to conduct line transect sampling is a valid method to estimate yelloweye rockfish density and may be a good replacement for the submersible. A model was produced that fit the data well and a density estimate was obtained with a low CV estimate, indicating good precision. The CV estimate for the 2012 ROV survey had low precision with a CV of 13% which was similar to that obtained by the 2007 submersible survey. In addition, the variance composition for the density estimates for both years was similar with the majority of the variance due to the encounter rate and the rest of the variance due to the detection probability. The variance due to the detection probability decreases as the fit of the model to the data improves. In 2012 only 18% of the variance was composed of the detection probability variance compared to the similar proportion of 15% in 2007; this indicates that a good model fit was obtained with the ROV survey data. The low CV value is an indication that the sample size (118 yelloweye) we obtained for the distance analysis and the number of transects (46 transects) performed were suitable to obtain a valid density estimate.

The behavior of yelloweye rockfish appears to be suitable for line transect sampling with the ROV. The majority of yelloweye rockfish appeared to be milling, hovering, or resting on the bottom when encountered by the ROV; consequently, there is likely no attraction or avoidance behavior influencing the distance estimate. The probability detection function did not indicate any attractive or avoidance behavior. There was no peak at the origin of the transect line in the probability detection function indicating that there probably wasn’t any attraction behavior (Figure 3). No avoidance behavior was indicated by the pattern of yelloweye observations with distance; if avoidance behavior occurred, then there would be lower detections near the origin of the transect line and then an increase in detection to some distance away from the origin. Instead the frequency histogram of distances binned in 0.5-ft intervals indicated that there was a large shoulder until around 6 ft when the probability of detection appears to decline; the shoulder is composed of big peaks at 2, 4.5, and 5 ft and smaller peaks at 0.5, 2.5, 3, 4 and 6 ft (Figure 1). If more data were collected, then it is likely that the frequency of observations across this shoulder would be more uniform. In comparison, there is a shoulder to around 4 ft for the 2007 CSEO submersible survey data (Figure 2); however, there was not an evident shoulder in the model chosen for the probability detection function chosen as the best model in 2007 (Figure 4). There may be a larger shoulder for the ROV survey, because we were often backing up the video and getting our distance measurements at the earliest point in time the fish appeared in the video. This may result in more observations further from the ROV and less close to the ROV than if we were to make our observations when the fish are easily discernible. With the submersible survey, fish were identified at the point in time the fish was noticeable to the observer, which would probably be at a distance closer to the camera than if the video was backed up to the first appearance of the fish.

The assumption of distance sampling that all yelloweye rockfish along the transect line were observed appeared to be met. The stereo camera field of view was similar to that of the forward-facing camera and indicated that fish close to the ROV and along the transect line were able to be observed.

The ROV is appropriate for performing line transect sampling even though there is a more limited field of view compared to the submersible. With the ROV yelloweye rockfish were only observed to a perpendicular distance of about 10 ft compared to 33 ft with the submersible; however, this is accounted for because the probability detection function is scaled to the effective width in order to estimate density (Figure 6). The effective width is the distance at which the probability of being detected is the same before or after that particular distance and is calculated from the probability of detection within the width of the line transect multiplied by the total width of the line transect (Figure 6); the width of the line transect is the distance at which the largest observation occurred. In addition, if the probability of detection at the origin is assumed to be 100%, then line transect estimation is equivalent to estimating the probability of detection at the origin. Consequently, it is most important to have a good model fit near the transect line and a smaller field of view should not affect the density estimation as long as an appropriate sample size is obtained.

A valid density estimate was produced with the 2012 ROV data regardless of the smaller sample size than the 2007 submersible survey. Compared to the 2007 submersible survey, about half the number of yelloweye rockfish were observed for the ROV survey and on average about half were observed per transect with the ROV (Table 8). In 2012 the encounter rate was lower with 0.003 ye/ total transect length surveyed (L) compared to 0.005 ye/L in 2007. Fewer yelloweye rockfish were observed along the transect line on the ROV survey even though we sampled from both sides of the transect line compared to only the right side for the submersible survey. The reduced field of view and the decline in density form 2007 to 2012 probably contributed to this lower encounter rate. The video quality was not great for the stereo cameras for the ROV survey, and as a result we may have not been able to identify all yelloweye rockfish in the camera field of view that were not close to the cameras. Some yelloweye rockfish may have been blurry or washed out and been identified instead as “unknown rockfish”. These fish may have reduced the total number of yelloweye observed and the encounter rate. In addition, there would be somewhat a reduction in the ability to identify yelloweye rockfish from the video compared to the submersible where the observer is actually in the water viewing the fish.

The 2012 yelloweye rockfish density estimate of 752 yelloweye/km2 was a decline from the 2007 estimate of 1068 yelloweye/km2. It is likely that this difference in the density estimate represents a real decline in the population. During the 5 years between the 2007 submersible and 2012 ROV surveys, fishing pressure occurred on the yelloweye rockfish population in CSEO with yelloweye rockfish captured in the sport fisheries, from both charter fishermen as well as unguided fishermen, as bycatch in the commercial halibut fishery, and in the directed DSR fishery which was open in 2012 and took around 86,000 pounds. Yelloweye rockfish are late maturing and slow growing; consequently, replacement of recruits to the fisheries-and those assessed in our Distance analysis- would be slow. In addition, because yelloweye rockfish have high site fidelity, we would not expect replacement of yelloweye rockfish from other management areas. The 2012 density estimate fits into the trend of a steady decline in yelloweye rockfish density in CSEO which has been occurring since the submersible survey was first conducted in 1995 (Figure 5). In 2007 all subadult yelloweye rockfish were included in the density estimate; however, for 2012 only subadults >350 mm were included in the estimate because this is the smallest size of fish observed in the commercial halibut and directed DSR fisheries. No length data were available prior to the 2012 survey; consequently, there were no criteria to determine whether it was appropriate to include subadult yelloweye rockfish in the density estimate. However it was probably easier to judge from inside the submersible the size of the yelloweye rockfish and whether it appeared to be a subadult, adult, or juvenile. A total of nine more subadult yelloweye rockfish were observed in the 2007 submersible survey. This change in criteria for including subadults into the density analysis is not enough to solely cause the decline in the density estimate from 2007 to 2012.

Table 1. Results from models with exact data with no binning.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Hazard rate cosine** | **Half normal cosine** | **Half normal hermite** |
| AIC | 498 | 501 | 502 |
| Density | 752 | 770 | 831 |
| D LCL | 586 | 493 | 578 |
| D UCL | 966 | 1201 | 1193 |
| CV of D | 0.125 | 0.228 | 0.185 |
| Judgement | Good fit with wide shoulder out to end of 5 ft bin | Large shoulder but declines in middle of 5 ft bin | Fit OK but declines before spike at 5 ft |
| X2 P-value | 0.72 | 0.39 | 0.21 |
| # parameters | 2 | 3 | 2 |
| Q-q plot | Pretty good some deviation in middle | OK deviation in the middle | Lots of deviation |
| K-S P-value | 0.76 | 0.66 | 0.20 |
| warnings |  |  | parameters very highly correlated |

Table 2. Results from models with 1-ft bins. The Q-q plots and K-S test are not performed for binned data.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Hazard rate cosine** | **Half normal cosine** | **Half normal hermite** |
| AIC | 497 | 501 | 501 |
| Density | 769 | 777 | 845 |
| D LCL | 598 | 501 | 587 |
| D UCL | 988 | 1206 | 1218 |
| CV of D | 0.126 | 0.225 | 0.187 |
| Judgement | Good fit with wide shoulder out to end of 5 ft bin | Large shoulder but declines in middle of 5 ft bin | Fit OK but declines before spike at 5 ft |
| X2 P-value | 0.74 | 0.38 | 0.20 |
| # parameters | 2 | 3 | 2 |

Table 3. Results from models with 1.5 ft bins.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Hazard rate cosine** | **Half normal cosine** | **Half normal hermite** |
| AIC | 400 | 403 | 404 |
| Density | 755 | 786 | 848 |
| D LCL | 588 | 499 | 586 |
| D UCL | 970 | 1237 | 1229 |
| CV of D | 0.126 | 0.233 | 0.189 |
| Judgement | Wide shoulder fits data well | Fit looks good generally but weird dip in shoulder | Fit looks good |
| X2 P-value | 0.657 | 0.282 | 0.125 |
| # parameters | 2 | 3 | 2 |
| Warnings of concern |  |  | parameters highly correlated convergence failure |

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Hazard rate cosine** | **Half normal cosine** | **Half normal hermite** |
| AIC | 338 | 343 | 343 |
| Density | 749 | 763 | 837 |
| D LCL | 582 | 480 | 576 |
| D UCL | 964 | 1211 | 1215 |
| CV of D | 0.127 | 0.238 | 0.190 |
| Judgement | Good fit at shoulder | OK fit at shoulder but starts slopping down before end of 5 ft interval | Not great fit at shoulder |
| X2 P-value | 0.618 | 0.077 | 0.034 |
| # parameters | 2 | 3 | 2 |
| Warnings of concern |  |  | Parameters are very highly correlated |

Table 4. Results from models with 2 ft bins.

Table 5. Results from models with 2.5 ft bins.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Hazard rate cosine** | **Half normal cosine** | **Half normal hermite** |
| AIC | 285 | 287 | 287 |
| Density | 797 | 862 | 859 |
| D LCL | 616 | 594 | 587 |
| D UCL | 1033 | 1250 | 1256 |
| CV of D | 0.130 | 0.190 | 0.194 |
| Judgement | Fit looks good | Fit is Ok but not great at shoulder | Fit is OK but not great at shoulder |
| X2 P-value | 0.337 | 0.117 | 0.123 |
| # parameters | 2 | 2 | 2 |
| Warnings of concern |  |  | Parameters are highly correlated |

Table 6. Results from model with no binning but includes truncation after 9-ft.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Hazard rate cosine** | **Half normal cosine** | **Half normal hermite** |
| AIC | 484 | 488 | 488 |
| Density | 750 | 810 | 824 |
| D LCL | 582 | 511 | 568 |
| D UCL | 966 | 1285 | 1195 |
| CV of D | 0.127 | 0.237 | 0.190 |
| Judgement | Good fit wide shoulder | Ok fit, some decline before 5ft spike, weird wave in shoulder | Not a good shoulder, declines too quickly for data |
| X2 P-value | 0.578 | 0.422 | 0.233 |
| # parameters | 2 | 3 | 2 |
| Q-q plot | Good fit but some deviation in middle | Quite a bit of deviation in middle | Only fits OK at ends |
| K-S P-value | 0.7884 | 0.5047 | 0.2275 |
| Warnings of concern |  |  | Some parameters are highly correlated |

Table 7. Results from model with 1-ft bins and truncation after 9-ft.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Hazard rate cosine** | **Half normal cosine** | **Half normal hermite** |
| AIC | 483 | 486 | 487 |
| Density | 771 | 823 | 836 |
| D LCL | 596 | 526 | 574 |
| D UCL | 998 | 1288 | 1217 |
| CV of D | 0.130 | 0.230 | 0.192 |
| Judgement | Good fit with wide shoulder, good at tail | OK fit but some decline before 5 ft bin and weird dip in shoulder | OK fit but declines before 5 ft bin |
| X2 P-value | 0.616 | 0.301 | 0.234 |
| # parameters | 2 | 3 | 2 |
| Warnings of concern |  |  |  |

Table 8. Comparison of 2007 and 2012 distance parameters and models.

|  |  |  |
| --- | --- | --- |
|  | ***2007*** | ***2012*** |
| ***Density estimate*** | 1068 ye/km2 | 752 ye/km2 |
| ***Density estimate CV*** | 0.127 | 0.125 |
| ***Variance components*** | 15% DP/85% ER | 18% DP/82% ER |
| ***Encounter rate (n/L)*** | 0.005 ye/L | 0.003 ye/L |
| ***Average number ye observed per transect*** | 5 | 2.6 |
| ***Number of fish observed*** | 301 | 118 |
| ***Range of ye observed per transect*** | 0-26 | 0-9 |
| ***Number of transects performed*** | 60 | 46 |
| ***ESW (ft)*** | 16.02 | 6.67 |
| ***Number of adults*** | 286 | 112 |
| ***Number of subadults*** | 15 | 6 |
| ***Transects no ye for distance*** | 6 | 7 |
| ***Greatest perpendicular distance ye observed (ft)*** | 33 | 10.39 |

Figure 1. 2012 yelloweye rockfish distance observations from the ROV in 0.5-ft bins.

Figure 2. 2007 yelloweye rockfish distance observations from the submersible.

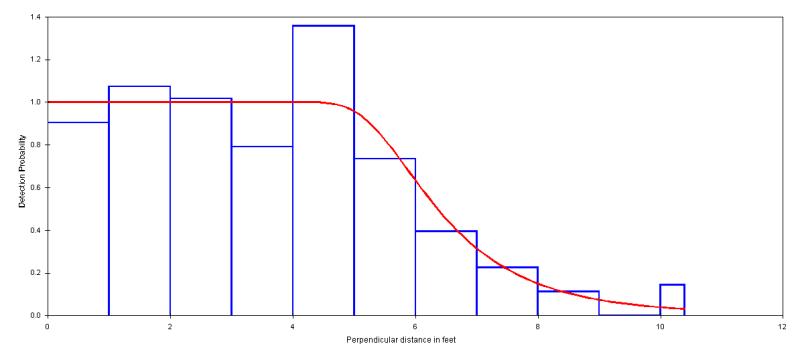


Figure 3. 2012 ROV probability detection function for yelloweye rockfish in CSEO with a hazard rate cosine model; 1-ft bins are used for the histogram but no bins or truncation is used for the estimation of distance.

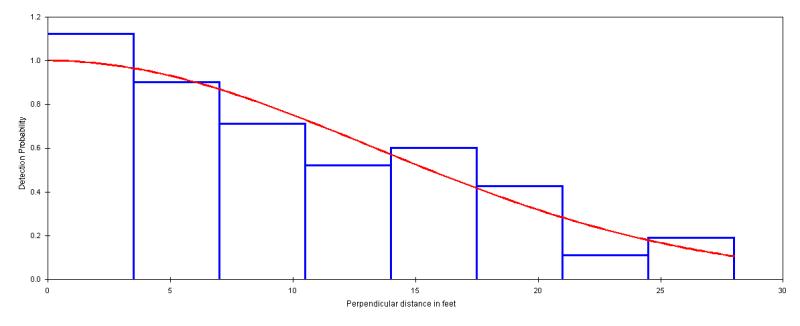


Figure 4. 2007 Sub probability detection function for yelloweye rockfish in CSEO with a half normal cosine model with 3.5-ft bins and truncation after 28 ft.

Figure 5. Density of yelloweye rockfish estimated from the submersible (1995, 1997, 2003, and 2007) and ROV (2012) surveys.

*n* = total number of yelloweye rockfish adults and subadults >350 mm observed,

*f(*0) = the probability density function evaluated at the origin of the transect line

*L* = total line length

*µ =* the effective width

*w* = width of line transect

*p* = probability of observing an object in the defined area

Figure 6. The equations used to estimate density using line transect sampling.