Estimated Survey Catchability for Tanner Crab in the Eastern Bering Sea Survey using Side-by-Side Gear Comparisons

William T. Stockhausen, Scott Goodman, Madison Shipley, Robert McConaughey

1 Introduction

Since 2010, BSFRF and NMFS have conducted several cooperative "side-by-side" (SBS) tow studies during the NMFS EBS Shelf bottom trawl survey to better characterize the selectivity and catchability of the NMFS survey gear with respect to snow crab, Bristol Bay red king crab, and Tanner crab. In this report, data from studies conducted in 2013-2017 are analyzed to develop survey-level sex- and size-specific catchability curves for Tanner crab.

In the SBS studies, a trawl vessel chartered by the Bering Sea Fisheries Research Foundation (BSFRF) would accompany a NMFS-chartered trawl vessel on part of the NMFS EBS Shelf bottom trawl survey. At each SBS station, the NMFS vessel would perform a standard survey haul using an 83-112 Eastern trawl towed at 3 knots for 30 minutes (Stauffer, 2004). Following a parallel track with a separation of ~0.25 nmi, the BSFRF vessel would perform a haul for 5 minutes at 2 knots using a *Nephrops* trawl modified to increase bottom contact and capture all crabs in the path of the net (Somerton et al., 2013). Video observations suggested the contact of the modified *Nephrops* trawl was quite good, and no crab were seen to escape under the footrope—or to be herded into trawl path (Somerton et al., 2013). It is a working assumption that the BSFRF gear captures all crab in the path of the net, and so provides an absolute estimate of local crab density (although Kotwicki et al., 2017, suggest this may not be the case for large crab).

Somerton et al. (2013) analyzed data from the first (2010) SBS study to estimate the annual NMFS "survey catchability" for snow crab by first estimating the haul-level "trawl efficiency" of the NMFS 83-112 trawl gear and then scaling this up to the entire survey. They used the term "survey catchability" to refer to the size-specific ratio between the survey-estimated abundance of crab by size, expanded to the entire survey area, relative to the population abundance by size. When survey catchability varies with size (or is less than 1 uniformly), the survey provides estimates of *relative* population abundance—however, knowing the survey catchability allows one to infer estimates of *absolute* population abundance.

In contrast to "survey catchability", Somerton et al. (2013) used the term "trawl efficiency" to refer to the size-selective nature of the trawl gear for a given haul, which could depend on site-specific characteristics such as bottom depth, temperature, and sediment characteristics that modify aspects of how the gear fishes or the behavior of crab in relation to the gear (e.g., burrowing behavior). "Trawl efficiency" quantifies the size-dependent probability that a crab in the path of the gear is caught for a particular haul—knowing the trawl efficiency and the numbers of crab caught by size, one can infer the local density of crab (by size). In the paired-haul design, the BSFRF haul is assumed to provide an estimate of the absolute size-specific density of crabs at the location of the paired haul, so it can be used to estimate the location-specific trawl efficiency of the NMFS gear. By estimating how trawl efficiency varied with bottom depth and sediment grain size, Somerton et al. (2013) were able to estimate the trawl efficiency at all trawl stations, from which a catch-weighted average gave an estimate of the survey catchability.

Here, using the available SBS data for Tanner crab, we expand on Somerton et al.'s (2013) approach to estimate annual survey catchability for this species across the entire time period the NMFS EBS Shelf bottom trawl survey has used the 83-112 Eastern trawl as its standard gear (1982-present). We also present results for NMFS survey catchability from an alternative approach that does not make use of the paired-haul nature of the SBS data.

2 Data

2.1 Side-by-side tow studies

With an initial focus on Bristol Bay red king crab, SBS studies that sampled Tanner crab started in Bristol Bay in 2013 and were repeated in 2014 and 2015 (Figure 2.1). In 2016, the focus was much more on adequately sampling Tanner crab and stations at which SBS tows were conducted expanded to the west beyond Bristol Bay. In 2017, the study area shifted further to the west completely out of Bristol Bay. Stations ran from the southeast along the Alaskan Peninsula to the northwest beyond the Pribilof Islands.

Consistent with differences in area swept (the area swept is ~6 time larger for NMFS hauls), the total number of crabs sampled in the SBS studies was generally larger in the NMFS hauls than in the BSFRF hauls, except in 2017 when a large group of small crabs was caught by the BSFRF gear. Aggregated over all tows by 5-mm CW size bin, the number of crab caught was larger for the NMFS hauls than for the BSFRF hauls (Figure 2.2) except at the smallest crab sizes (e.g., < 50 mm CW in 2017). In contrast, the estimated crab abundance by size bin expanded to the associated SBS study area, and consequently in total, tended to be larger for the BSFRF gear than for the NMFS gear (Table 2.1, Figure 2.3).

2.2 Sediment data

To include sediment characteristics as potential environmental covariates of haul-level catchability, pointwise data at the level of individual samples collected at sites on the EBS shelf were obtained from dbSEABED to use as potential environmental covariates of haul-specific catchability (Figures 2.4 and 2.5). Sediment samples are generally characterized by the mass-weighted (not number-weighted) distribution of φ, the grain size on a $\log_2$ scale. Standard measures of grain size distribution include the mean φ and sorting coefficient (i.e., the standard deviation of the distribution). These parameters in particular encapsulate basic characteristics of the size distribution and have been shown to affect such basic sediment properties as porosity, permeability, and compaction (Selley 1988). In general, the sites at which sediment samples were collected did not coincide with the locations of stations in the SBS studies, so smooth surfaces for mean φ and the sorting coefficient were created using Empirical Bayes co-Kriging in ArcMap 10.7.1 (ESRI, 2020) to allow interpolation of these quantities to individual SBS haul locations, as well as to all EBS Shelf survey haul locations.

3 Survey-level catch ratio analysis

3.1 Introduction

As noted previously, "survey catchability" refers to the size-specific ratio between survey-estimated abundance of crab by size, expanded to the entire survey area, relative to the population abundance by size *within the area surveyed*. "Availability" is a related concept: it refers to the fraction of individuals in a population, by size, that are available to be captured in a survey (i.e., the fraction that could conceivably be caught). Availability is considered to be 1 when the survey area encompasses the entire population/stock area and no individuals occupy refuge habitats that cannot be surveyed (e.g., untrawlable rocky bottom). When the survey area does not encompass the entire population/stock area, then availability will be less than 1. Availability, like catchability, may vary with crab size. On the geographic scale of an entire survey, size-specific survey abundance is related to size-specific population abundance by

$$N_z^{survey} = A_z^{survey} \cdot C_z^{survey} \cdot N_z^{population} \tag{3.1}$$

where z represents size, N represents size-specific abundance, A represents size-specific availability, and C represents size-specific survey catchability.

Availability in the NMFS EBS bottom trawl survey is considered to be 1 for all Tanner crab in the EBS stock because the survey area encompasses the distribution of the stock. Because the SBS studies were conducted on areas smaller than the full survey, availability to the NMFS and BSFRF gear in these studies must be less than or equal to 1. The BSFRF gear is assumed to catch all crab within the footprint of a tow (i.e., it is non-selective and provides an estimate of absolute abundance), so the following relationships are assumed to hold:

$$N_z^{NMFS\ EBS} = 1 \cdot C_z^{NMFS} \cdot N_z^{EBS} \tag{3.2}$$

$$N_z^{NMFS\ SBS} = A_z^{SBS} \cdot C_z^{NMFS} \cdot N_z^{EBS} \tag{3.3}$$

$$N_z^{BSFRF\ SBS} = A_z^{SBS} \cdot 1 \cdot N_z^{EBS} \tag{3.4}$$

where the assumptions $A_z^{EBS} \equiv 1$ and $C_z^{BSFRF} \equiv 1$ have been inserted into the equations above.

Using Equations 3.3-4, it is possible to estimate NMFS survey catchability (C_z^{NMFS}) directly from the SBS study data. Dividing Equation 3 by Equation 4 yields:

$$C_z^{NMFS} = \frac{N_z^{NMFS\ SBS}}{N_z^{BSFRF\ SBS}} \tag{3.5}$$

which provides an empirical estimate of NMFS survey catchability outside the assessment model. Note, however, that catchability can only be estimated using this method for size classes that occupy the SBS area.

3.2 Methods

3.2.1 Annual estimates

Annual sex-specific estimates for C_z^{NMFS}, $C_{z,x,y}^{NMFS}$ (where *x* indicates sex and *y* indicates survey year), were obtained by first applying an unstratified design-based, area-swept estimator for $N_{z,x,y}^{S\ SBS}$, (*S* denotes the survey), to the number of sex *x* crab caught in each SBS tow by 5-mm bins to obtain sex-specific estimates, annual estimates of $N_z^{S\ SBS}$ for both surveys. Equation 3.5 was then used to estimate $C_{z,x,y}^{NMFS}$ for each year and sex in question.

One expects catchability to vary fairly smoothly with size, so sex-specific Generalized Additive Models (GAMs) were fit to the $C_{z,x,y}^{NMFS}$ using R package "mgcv" (Wood, 2011; Wood, 2017) to obtain annual estimates of catchability that varied smoothly with size. The "raw" $C_{z,x,y}^{NMFS}$ were fit separately by sex using a log link to a normal error distribution model with annual cubic spline "smooths" in size:

$$\ln\left(C_{z,y}^{EBS}\right) = s_y(z) + \epsilon_{z,y}, \quad where\ \epsilon_{z,y} \sim N(0, \sigma^2) \tag{3.6}$$

where $s_y(z)$ denotes a year-specific smooth function in *z*. When fitting the models, each $C_{z,x,y}^{NMFS}$ was weighted by the total number of crab from which it was derived.

Equation 3.6 implies that the likelihood distribution for the $C_{z,y}^{EBS}$ is lognormal with constant coefficient of variation. This is consistent with assumptions in the assessment model that the likelihood distribution for fitting survey biomass (or abundance) is lognormal, because the ratio of lognormally-distributed random variables is also lognormally-distributed.

3.2.2 Aggregated estimates

The assessment model estimates a single sex-specific catchability curve for the entire 1982-2019 time period, so sex-specific GAMs were also fit to the "raw" $C_{z,x,y}^{NMFS}$ using a single smooth term over crab size such that

$$\ln\left(C_{z,y}^{EBS}\right) = s(z) + \epsilon_{z,y}, \quad where\ \epsilon_{z,y} \sim N(0, \sigma^2) \tag{3.7}$$

Again, each $C_{z,x,y}^{NMFS}$ was weighted by the total number of crab from which it was derived.

3.3 Results

3.3.1 Annual curves

The sex-specific NMFS and BSFRF size compositions and resulting "raw" empirical catchability curves from the 2013-2017 SBS selectivity studies are shown in Figures 3.1 (males) and 3.2 (females). The empirical catchability curves estimated directly from the SBS data appear to imply a fair bit of interannual variability in size-specific catchability, as well as some indications that the shapes of these curves are not logistic (as is assumed in the assessment model). In addition, the empirical catchability is occasionally greater than 1 (e.g., 2015 in Figure 3.1), indicating the associated estimates of abundance were larger from the NMFS gear than from the BSFRF gear. Because the assumption underlying this analysis is that the BSFRF gear provides estimates of absolute abundance while the NMFS gear provides estimates of relative abundance, one might not expect the empirical catchability to be greater than 1. However, this sort of result is certainly possible given that estimate of absolute abundance from the BSFRF gear has uncertainty associated with it.

The GAM fits to the "raw" empirical values (upper plots in Figures 3.3 and 3.4) reduce some of the variability in the "raw" estimates, but still indicate a fair amount of interannual variability and non-logistic shape.

3.3.2 Aggregated curves

However, when the "raw" estimates are fit without regard to year, the resulting smooth curves are much more similar to logistic curves, with fully-selected catchability coefficients ("survey q's") of about 0.35 for females and 0.6 for males. This is, in part, due to standardizing the data weighting in the model fit across years, rather than within year as with the annual fits model.

4 Haul-level catch ratio analysis

4.1 Introduction

The approach in Section 3 does not take into account the paired-tow nature of the SBS studies. That approach can be equally well applied to two completely independent surveys conducted with different gear that sampled the same area at the same time, but not necessarily at the same tow locations. In this section, to take better advantage of the paired tow nature of the SBS data, we apply a modified version of the approach taken by Somerton et al. (2013) to the paired haul data.

Assuming that crab of size z on paired haul h become available to the gear following a spatial Poisson process with density $d_{z,h}$, the number of crabs sampled by NMFS in a paired haul, $N_{z,h}^N$, will follow a Poisson distribution (Fryer et al., 2003) with

$$N_{z,h}^N \sim Poisson(f_h^N \cdot E_{z,h}^N \cdot A_h^N \cdot d_{z,h}) \tag{4.1}$$

where f_h^N is the sub-sampling factor used to measure the crab (<1 if the whole haul was not sampled), A_h^N is the area swept by the NMFS gear, $E_{z,h}^N$ is what Somerton et al. (2013) referred to as "trawl efficiency" in size bin z on haul h, and $d_{z,h}$ is the average density of crab in size bin z across the location where the paired tow was conducted. As noted in the Introduction, trawl efficiency refers to catchability at the level of an individual haul, which is similar to, but may be different from, survey catchability if environmental variables such as depth or sediment type influence trawl performance. Trawl efficiency characterizes a single haul at a specific location and possibly under specific environmental effects, whereas survey catchability characterizes the group of tows constituting a survey.

Similar to Equation 4.1, the number of crabs sampled by BSFRF in the same paired tow will follow a Poisson distribution with

$$N_{z,h}^B = Poisson(f_h^B \cdot E_{z,h}^B \cdot A_h^B \cdot d_{z,h}) \tag{4.2}$$

where all superscripted quantities now refer to the BSFRF gear and sampling. Note that the variability in catch associated with 4.1 and 4.2 depends on the degree of patchiness of crab within the area of the paired tow: the Poisson distribution is correct if crabs are randomly distributed across the area sampled in paired tow, but an alternative distribution such as the negative binomial may be more appropriate if the degree of spatial patchiness is large.

Conditioned on the total number of crabs in size bin z measured in the paired haul, $N_{z,h}^N + N_{z,h}^B$, $N_{z,h}^N$ is thus binomially-distributed with probability $\phi_{z,h}$ (Fryer et al., 2003):

$$N_{z,h}^N | N_{z,h}^N + N_{z,h}^B \sim B\left(N_{z,h}^N + N_{z,h}^B, \tilde{\phi}_{z,h}\right) \tag{4.3}$$

The observed probability $\phi_{z,h}$, then, of crabs being sampled from the NMFS gear, given the total number of crabs sampled, is

$$\phi_{z,h} = \frac{N_{z,h}^N}{N_{z,h}^N + N_{z,h}^B} = \frac{f_h^N \cdot A_h^N \cdot E_{z,h}^N}{f_h^N \cdot A_h^N \cdot E_{z,h}^N + f_h^B \cdot A_h^B \cdot E_{z,h}^B} = \frac{q_h \cdot r_{z,h}}{1 + q_h \cdot r_{z,h}} \tag{4.4}$$

where $q_h = \left(f_h^N \cdot A_h^N\right) / \left(f_h^B \cdot A_h^B\right)$ is the haul-specific ratio of sampling factors $f_h \cdot A_h$ and $r_{z,h} = \left(E_{z,h}^N / E_{z,h}^B\right)$ is the size- and haul-specific *relative trawl efficiency*. Under the working hypothesis that the BSFRF gear captures all crabs in its path, $E_{z,h}^B = 1$ and $r_{z,h}$ then represents the absolute trawl efficiency of the NMFS gear.

Applying a logit transformation to both sides of Equation 4.4 yields

$$logit(\phi_{z,h}) = \ln\left(\frac{\phi_{z,h}}{1-\phi_{z,h}}\right) = ln(q_h \cdot r_{z,h}) = ln(q_h) + ln(r_{z,h}) \quad (4.5)$$

Note that the q_h's are known quantities and can be treated in the statistical analysis as haul-specific offsets, so any model for $logit(\phi_{z,h})$ is essentially also a model for $ln(r_{z,h})$.

4.2 Methods

4.2.1 Data culling

Before applying the statistical analysis, the paired haul data was examined for outliers. A paired haul was considered problematic if the ratio of the sampling factors, q_h, was too far from the nominal value (Figure 4-2.1). Typically, this was the result of one of the pair of hauls being foreshortened, but still considered a "valid" tow. For this analysis, paired tows with q_h's falling outside the range of $[\frac{1}{3}, 3] \cdot q_{nom}$ were eliminated from further consideration (Figure 4-2.2).

In addition, values of the catch ratio that were based on fewer than 5 individuals were eliminated from the analysis (Figures 4-2.3 and 4-2.4). This culling resulted in the final dataset illustrated in Figure 4-2.5.

4.2.2 Catchability estimation

Sex-specific models for the catch ratio data shown in Figure 4-2.5 were estimated using the R package "mgcv" (Wood, 2011; Wood, 2017). The number of crabs sampled in size bin *z* by NMFS on paired haul *h*, $N^N_{z,h}$, was assumed to follow the binomial distribution given in Equation 4.3 after conditioning on the total number of crabs sampled in the size bin, $N^N_{z,h} + N^B_{z,h}$. $\phi_{z,h}$ was parameterized on the logit-scale as a function, $f_h(z)$, of size *z* and (possibly) haul-specific covariates such as bottom depth using

$$logit(\phi_{z,h}) - \ln(q_h) = f_h(z) + \epsilon_h + \varepsilon_{h,z};\ \epsilon_h \sim N(0, \sigma^2_\epsilon),\ \varepsilon_{h,z} \sim N(0, \sigma^2_\varepsilon) \quad (4.6)$$

where ϵ_h is a haul-level random effect and $\varepsilon_{h,z}$ represents sampling error. Comparison with Equation 4.5 indicates that

$$ln(r_{z,h}) = f_h(z) \quad (4.7)$$

after integrating out the random effects.

Models were tested using haul-specific environmental variables as potential spatial covariates, including bottom depth, bottom temperature, mean grain size (φ), and sorting coefficient. The base model tested was a model which included all 2nd-order smooth terms among crab size and the covariates, together with the haul-level random effect. All smooth terms were implemented as "ti"-type tensor product smooths in mgcv using thin-plate regression splines with extra shrinkage. Starting with the base model, terms were evaluated for significance and dropped if found non-significant ($p>0.05$) using a backward stepping approach to model selection. Finally, a simple smooth function of size with no random effects: $f_h(z) = s(z)$ and $\epsilon_h \equiv 0$ was also evaluated. All models evaluated were compared by AIC and BIC to determine the most parsimonious model.

4.3 Results

4.3.1 Males

For males, the "best" model, as determined by AIC and BIC after eliminating non-significant terms, included first order smooth terms involving crab size, depth, and temperature (but not mean grain size or sorting coefficient) and second order terms involving size *x* depth, size *x* sorting coefficient, and depth *x*

sorting coefficient, as well as the haul-specific random effects (Tables 4-1 and 4-2). Although quantile-quantile (qq) plots of the deviance residuals appear reasonable, the histogram of residuals is somewhat non-normal and a plot of the residuals versus the linear predictor indicates somewhat positive bias at small values of the predictor, as well as heteroscedasticity (Figure 4.3-1).

The 1-d smooth term for crab size increases steeply over the range of 10-70 mm CW, then flattens out substantially above ~80 mm CW (Figure 4.3-2). The term involving bottom depth increases very gradually with depth in an almost linear fashion whereas the term involving bottom temperature exhibits almost the opposite trend.

The 2-d smooth term involving crab size and bottom depth increases with size at shallow depths, has little effect at moderate depths (~100 m), and decreases with size at deep depths (Figure 4.3-3). The term involving crab size and bottom temperature exhibits a qualitatively similar pattern, although the scale is different. Patterns for the other 2-d smooth terms (crab size *x* φ, size *x* sediment sorting coefficient, depth *x* sorting coefficient, and temperature *x* sorting coefficient) exhibit patterns with multiple minima and maxima (Figure 4.3-3).

The estimated haul-specific trawl efficiencies at all NMFS survey stations at which hauls were conducted in a given year, the annual survey catchabilities estimated using unweighted means of the haul-specific trawl efficiencies, and annual survey catchabilities estimated using inverse variance-weighted means are illustrated in Figures 4.3-4 through 4.3-10 by six-year groups. The haul-specific trawl efficiencies exhibit some unexpectedly large values (>>1) that reflect highly uncertain predictions based on the fitted model. The unweighted mean survey catchabilities, however, do not rise above 1 at any crab size, although the uncertainty increases substantially as crab size increases. Using inverse-variance weighting to estimate the annual survey catchability typically reduces the increasing uncertainty at large crab sizes, but unreasonably large estimates for catchability occasionally still occur when only a few measured crabs are associated with the estimate.

The annual survey catchabilities estimated using inverse variance-weighted means are plotted *en masse* in Figure 4.3-11 to emphasize the similarities and differences in the estimated survey catchabilities across years. All years show roughly the same trend, with survey catchability increasing with crab size until about 130 mm CW, then generally decreasing but becoming much more variable. Most of the extreme excursions at large crab size occur when only a few observations were available to include in the average.

4.3.2 Females

For females, the most parsimonious model included first order smooth terms involving crab size and depth (but not temperature, mean grain size or sorting coefficient) and second order terms involving size *x* depth, size *x* grain size, as well as the haul-specific random effects (Tables 4-3 and 4-4). Observations regarding the diagnostic plots for the best fitting model (Figure 4.3-12) are similar to those for males: the quantile-quantile (qq) plots of the deviance residuals appear reasonable, the histogram of residuals is somewhat non-normal and a plot of the residuals versus the linear predictor indicates somewhat positive bias at small values of the predictor, as well as heteroscedasticity.

The 1-d smooth term for crab size increases steeply over the range of 5-40 mm CW, then increases much more gradually with size (Figure 4.3-13). Qualitatively, it looks very similar to the same term for males, but the both the x- (size) and y-axis scales are different. As with males, the term involving bottom depth increases gradually with depth in an almost linear fashion, although the slope is somewhat steeper.

The 2-d smooth term involving crab size and bottom depth (Figure 4.3-14) is also qualitatively similar to that for males (Figure 4.3-3): it increases with size at shallow depths, has little effect at moderate depths (~100 m), and decreases with size at deep depths (Figure 4.3-14). The other 2-d smooth term involving crab size and φ was not a significant term for males. At small φ, it first increases with crab size, reaches a

plateau at intermediate sizes, then decreases. At large φ, the opposite pattern occurs. The pattern at intermediate φ is generally similar to that at large phi, but with less overall variation. (Figure 4.3-14).

The estimated haul-specific trawl efficiencies at all NMFS survey stations at which hauls were conducted in a given year are illustrated in Figures 4.3-15 through 4.3-21 by six-year groups (plots in the upper righthand quadrant of each figure). Some of the individual curves suggest efficiencies >> 1 for large female crabs. Some of these are driven by small sample sizes, others involve extrapolating the smooth terms beyond values supported by the SBS data. In particular, the size range of females was poorly represented in the 2013-2015 SBS study areas. Averaging the individual trawl efficiency estimates to obtain annual estimates of survey catchability reduced the overall variability substantially (plot in upper left quadrant of each figure), but the predicted catchabilities are still biased to values much larger than 1 for the largest female crabs. However, weighting the haul-specific trawl efficiency estimates by the size-specific CPUE for the haul and inversely by the variance of the model estimate yielded much more reasonable estimates of annual survey catchability (plot in bottom row of Figures 4.3-15-4.3-21).

The annual survey catchabilities estimated using CPUE and inverse variance-weighted means are plotted *en masse* in Figure 4.3-22 to emphasize the similarities and differences in the estimated survey catchabilities across years. All years show roughly the same trend, with survey catchability increasing with crab size from 25 to 50 mm CW, followed by a plateau until 75 mm CW, after which it rises again until ~ 90 mm CW, followed by a gradual decline. As with males, a few extreme excursions from the general trend occur at large crab size where only a few observations were available to include in the average.

5 Discussion

This report presents estimates for catchability of Tanner crab in the NMFS EBS Shelf survey based on applying two different approaches to data collected in a series of side-by-side (SBS) haul studies performed by the Bering Sea Fisheries Research Foundation in collaboration with NMFS in conjunction with the annual NMFS EBS Shelf bottom trawl survey.

The "empirical catchability" approach (Section 3) is based on the ratio of size compositions from the NMFS and BSFRF hauls expanded to the appropriate study area. This yields an estimate of the NMFS survey catchability across the SBS study area under the assumption that the BSFRF gear captures all crabs in the path of its gear in each haul. For the study years 2013-2017, the resulting annual estimates of sex-specific survey catchability were fairly variable between adjacent size bins, in contrast to the expectation of a smoothly-varying curve. Generalized additive models (GAMs) were applied to obtain smooth annual curves in size, but the shapes varied fairly substantially among years, in contrast to the expectation that survey catchability should be fairly similar from year to year. However, applying a GAM to the annual estimates to obtain a single (sex-specific) smooth curve in size yielded what appear to be reasonable estimates for Tanner crab survey catchability: the curves for males and females are approximately ascending logistic curves in shape, suggesting fully-selected "survey q's" of 0.6 for males (Figure 3.3) and 0.35 for females (Figure 3.4). This approach, however, does not make any use of the paired-haul nature of the SBS studies.

The "catch ratio" approach (Section 4) does make use of the paired-haul nature of the SBS studies by fitting smooth functions to the proportion by size of the number of crabs caught and measured using the NMFS gear in a paired haul relative to the total number-at-size caught and measured using both gears. Incorporating potential haul-specific environmental covariates (bottom depth, temperature, mean sediment grain size, sediment sorting coefficient) into models for the haul-level "trawl efficiency" of the NMFS gear allows one to predict the haul-specific trawl efficiency for NMFS hauls that were not part of the SBS studies. Averaging, in some fashion, the predicted haul-level trawl efficiency over all hauls in an annual survey then provides an estimate of "survey catchability" for that survey. Using this approach, we

were able to estimate the annual survey catchability of the NMFS EBS Shelf bottom trawl survey for male and female Tanner crab (separately) for surveys from 1982 to 2019 (Figures 4.3-11 and 4.3-22). In contrast to the final sex-specific curves estimated using the "empirical catchability" approach, the catchability curves estimated using the catch ratio approach were not logistic and featured a descending limb at large crab sizes. In addition, although these all had the same general shape, there was fairly substantial annual variation in size-specific catchability for both males and females across the time series. It seems likely that this variability, particularly at large crab sizes, is driven by changes in the spatial distribution of Tanner crab abundance relative to survey station locations, although changes in bottom temperature due to cold pool variation contributes somewhat for males (temperature *x* crab size and temperature *x* sediment sorting coefficient interactions were significant for males in the "best" model describing male catch ratios, but no temperature effects were significant for females).

These results suggest it may be worthwhile to evaluate the use of alternatives to the logistic selectivity functions used to describe survey catchability in the Tanner crab assessment model (Stockhausen 2020). The estimated catchability curves can also be used in the assessment model, either as the bases for priors on estimated survey catchability in the model or more directly as assumed values for catchability in the model.

6 References

Buczkowski, B.J., Reid, J.A., & Jenkins, C.J., 2020. Sediments and the sea floor of the continental shelves and coastal waters of the United States—About the usSEABED integrated sea-floor-characterization database, built with the dbSEABED processing system. U.S. Geological Survey Open-File Report 2020–1046, 14 p., https://doi.org/10.3133/ofr20201046.

ESRI. 2020. ArcGIS Desktop: Release 10.7.1. Redlands, CA: Environmental Systems Research Institute.

Fryer, R.J., A.F. Zuur, and N. Graham. 2003. Using mixed models to combine smooth size-selection and catch-comparison curves over hauls. Can. J. Fish. Aquat. Sci. 60:448-459.

Kotwicki, S., R.R. Lauth, K. Williams, and S.E. Goodman. 2017. Selectivity ratio: a useful tool for comparing size selectivity of multiple survey gears. Fish. Res. 191:76-86. https://dx.doi.org/10.10.16/j.fishers.2017.02.012

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Richwine, K.A., K.R. Smith, and R.A. McConnaughey. 2018. 377Surficial Sediments of the Eastern Bering Sea Continental Shelf: EBSSED-2 Database Documentation. NOAA Technical Memorandum NMFS-AFSC-377. 57 pp.

Selley, R.C. 1988. Applied Sedimentology. Academic Press, San Diego, CA. 446 p.

Somerton, D.A., R.A. Mconnaughey, and S.S. Intelmann. 2017. Evaluating the use of acoustic bottom typing to inform models of bottom trawl sampling efficiency. Fish. Res, 185:14-16. https://dx.doi.org/10.10.16/j.fishers.2016.09.029

Somerton, D. A., and R. S. Otto. 1999. Net efficiency of a survey trawl for snow crab, *Chionoecetes opilio,* and Tanner crab, *C. bairdi.* Fish. Bull. 97:617-625.

Stauffer, G. 2004. NOAA protocols for groundfish bottom trawl surveys of the Nation's fishery resources[online]. US Dept. Commer. NOAA Tech. Memo. AFSC-SPO-65. Available from http://spo.nmfs.noaa.gov/tm65/tm65.pdf.

Stockhausen, W. 2020. 2020 Stock Assessment and Fishery Evaluation Report for the Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions. In: Stock Assessment and Fishery

Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands: 2020 Final Crab SAFE. North Pacific Fishery Management Council. Anchorage, AK.

Wood, S.N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B) 73(1):3-36.

Wood, S.N. 2017. Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC.

7 Table captions

8 Figure captions

9 Tables

Table 2.1. Sample sizes and estimated total abundance within the SBS study areas for BSFRF and NMFS surveys.

	number sampled				total abundance (millions)			
	female		male		female		male	
year	BSFRF	NMFS	BSFRF	NMFS	BSFRF	NMFS	BSFRF	NMFS
2013	267	538	640	1302	53.32	17.071	139.14	47.01
2014	92	207	441	1814	20.4	7.511	91.18	60.64
2015	108	198	264	998	17.37	6.254	49.05	33.42
2016	698	682	998	2281	114.76	20.665	171.19	67.09
2017	2625	1784	2556	3471	455.59	62.617	446.67	89.03

Table 3.1. Results for fits to male empirical catchability by GAM models.

model	df	AIC	BIC	% deviance explained
annual fits	25.44	-44.44	30.58	75.02
aggregated fit	5.02	25.41	40.21	48.32

Table 3.2. Results for fits to female empirical catchability by GAM models.

model	df	AIC	BIC	% deviance explained
annual fits	24.40	-32.80	27.66	74.76
aggregated fit	4.40	-28.61	-17.72	42.34

Table 4.1. Comparison of GAM models for males.

model	df	AIC	BIC
size only	136.79	3031.66	3696.11
all 2-way interactions	140.54	2936.95	3619.62
only significant 2-way interactions	137.71	2931.46	3600.40

Table 4.2. Diagnostics output for best catch ratio GAM model for males. *edf*: empirical degrees of freedom; *ti*: tensor interaction smooth function; *z*: crab size; *d*: bottom depth; *f*: φ, $\log_2$-scale mean grain size, *t*: bottom temperature; *s*: sediment sorting coefficient; *s*(*h*): smooth representing random effects.

```
Method: REML   Optimizer: outer newton
full convergence after 22 iterations.
Gradient range [-0.001232008,0.007510317]
(score 1533.706 & scale 1).
Hessian positive definite, eigenvalue range [2.109921e-05,32.51288].
Model rank =  314 / 314

Basis dimension (k) checking results. Low p-value (k-index<1) may
indicate that k is too low, especially if edf is close to k'.

             k'     edf k-index p-value
ti(z)     4.000   3.923    1.00    0.55
ti(d)     4.000   1.110    0.98    0.28
ti(f)     4.000   0.914    1.00    0.39
ti(z,d)  16.000   4.710    0.98    0.21
ti(z,t)  16.000   2.610    1.02    0.81
ti(z,f)  16.000   7.335    0.98    0.20
ti(z,s)  16.000   1.522    1.02    0.71
ti(d,s)  16.000   2.513    1.10    1.00
ti(t,s)  16.000   1.654    1.08    1.00
ti(f,s)  16.000   0.858    1.11    1.00
s(h)    189.000  99.648      NA      NA
```

Table 4.3. Diagnostics output for best catch ratio GAM model for females.

model	df	AIC	BIC
size only	95.37	1620.73	2023.05
all 2-way interactions	113.53	1500.95	1979.90
only significant 2-way interactic	102.51	1487.81	1920.28

Table 4.4. Diagnostics output for best catch ratio GAM model for females. *edf*: empirical degrees of freedom; *ti*: tensor interaction smooth function; *z*: crab size; *d*: bottom depth; *f*: φ, $\log_2$-scale mean grain size, $s(h)$: smooth representing random effects.

```
Method: REML   Optimizer: outer newton
full convergence after 7 iterations.
Gradient range [-2.203405e-07,1.33785e-07]
(score 820.9957 & scale 1).
Hessian positive definite, eigenvalue range [0.2495432,23.32384].
Model rank =  175 / 175

Basis dimension (k) checking results. Low p-value (k-index<1) may
indicate that k is too low, especially if edf is close to k'.

            k'    edf k-index p-value
ti(z)     4.00   3.82    0.99    0.48
ti(d)     4.00   1.11    1.13    1.00
ti(z,d)  16.00   6.44    1.06    0.96
ti(z,f)  16.00   7.92    1.08    0.98
s(h)    134.00  76.86      NA      NA
```

10 Figures

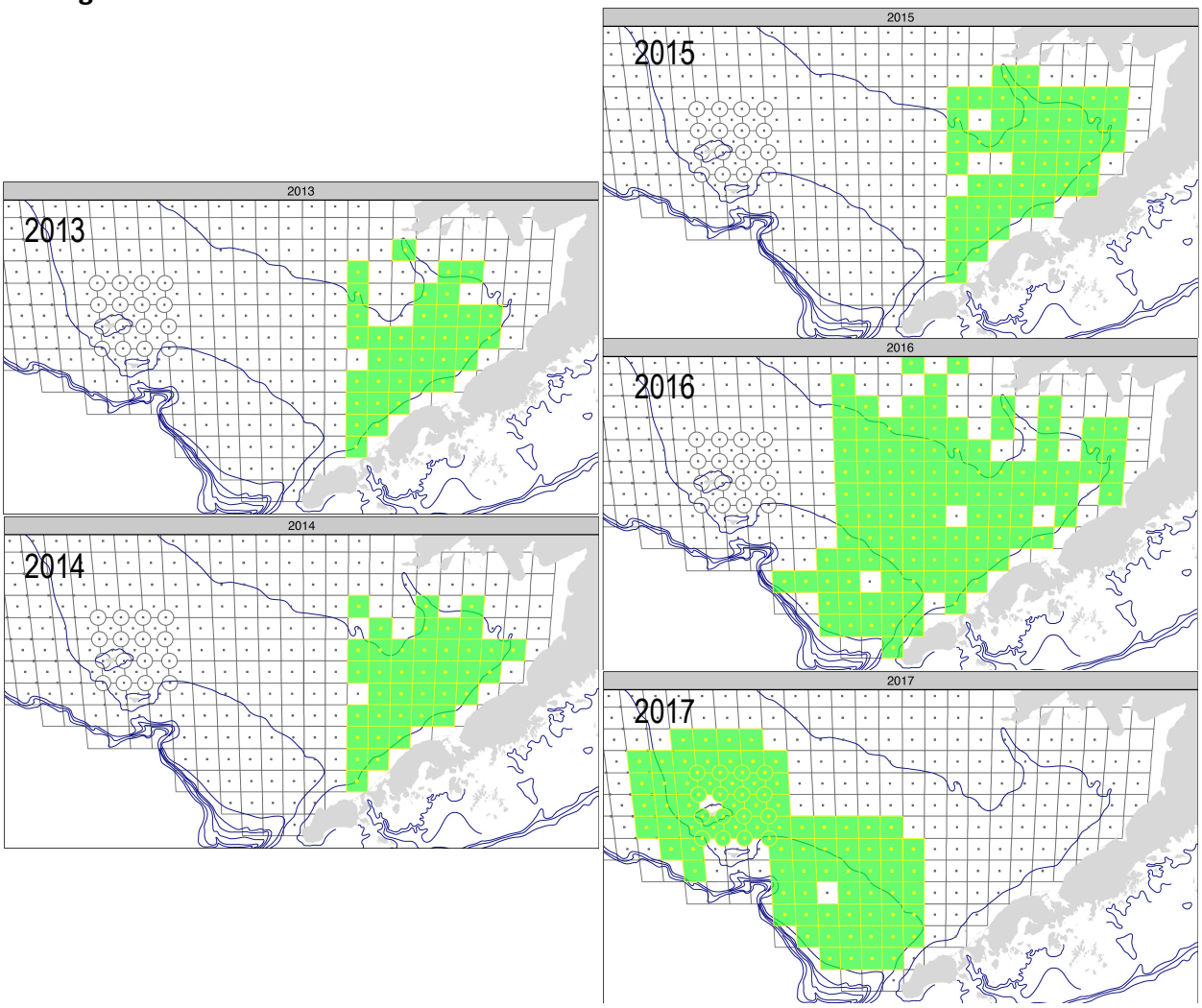

Figure 1. Joint NMFS-BSFRF side-by-side (SBS) haul study stations, by year. An SBS study was conducted in 2018 as well, but the BSFRF data have not been released yet.

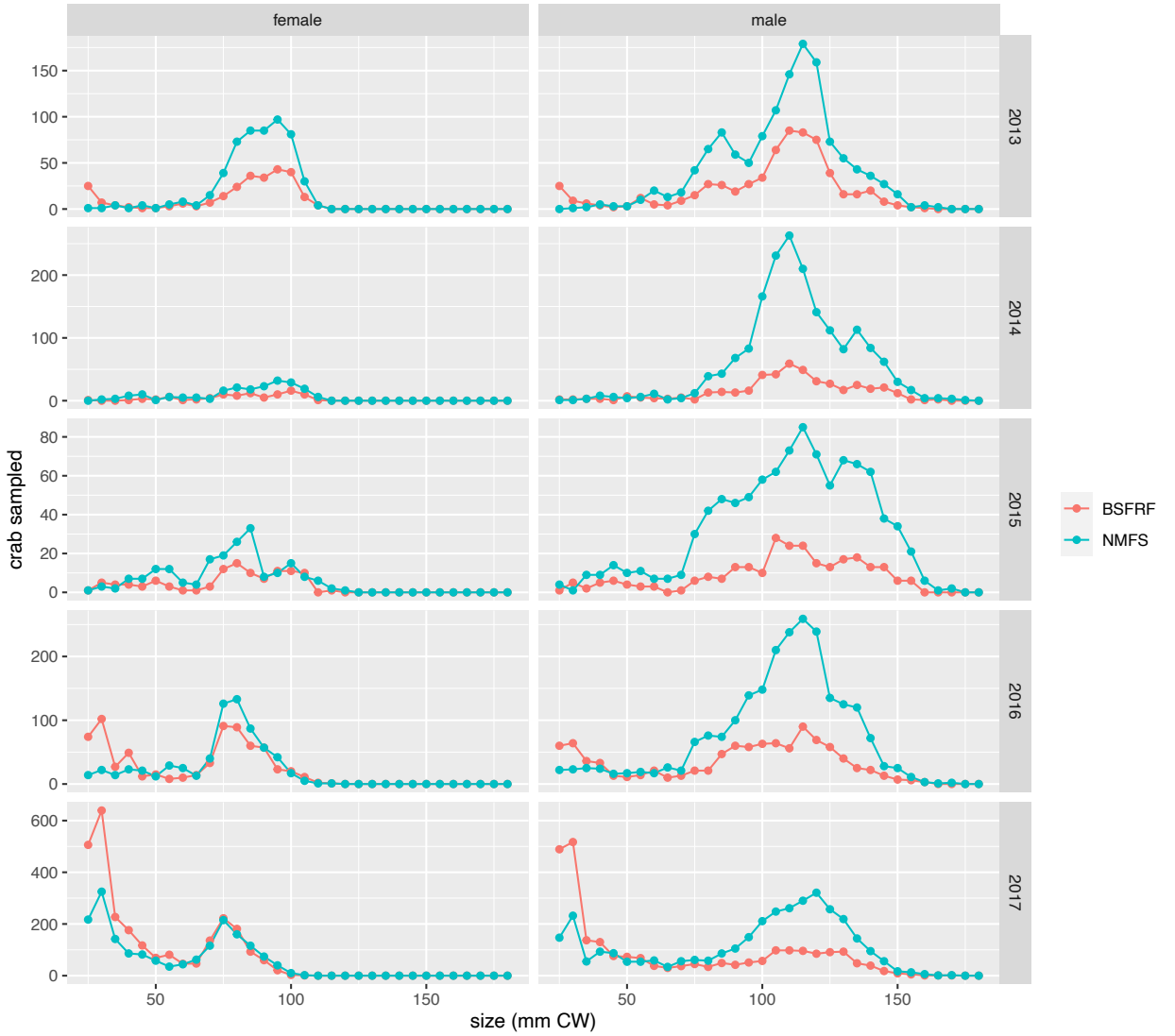

Figure 2. Total numbers sampled in the SBS studies, by crab size (5 mm bins). Color = organization.

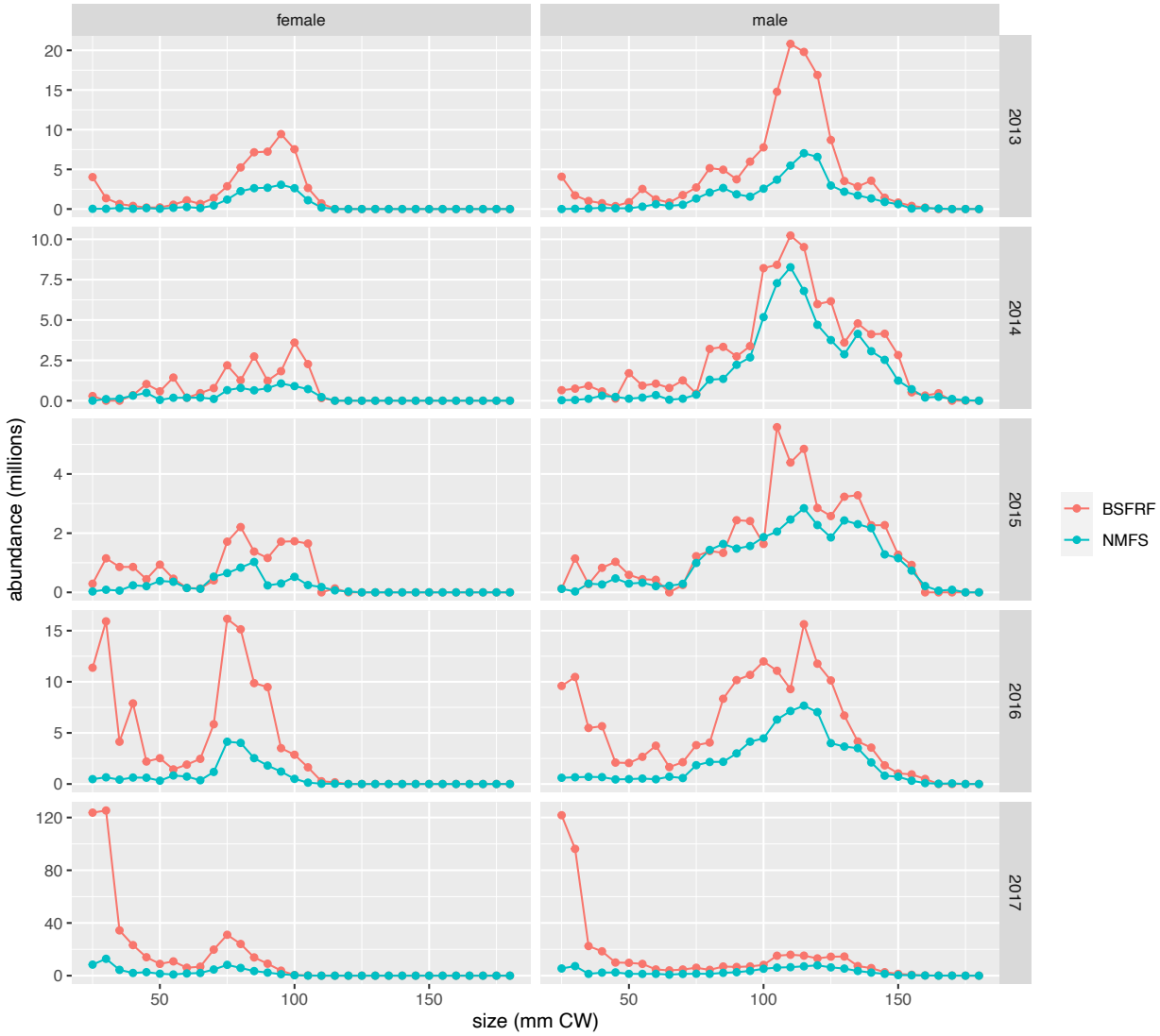

Figure 3. Area-swept abundance in the SBS studies, by crab size (5 mm bins). Color = organization.

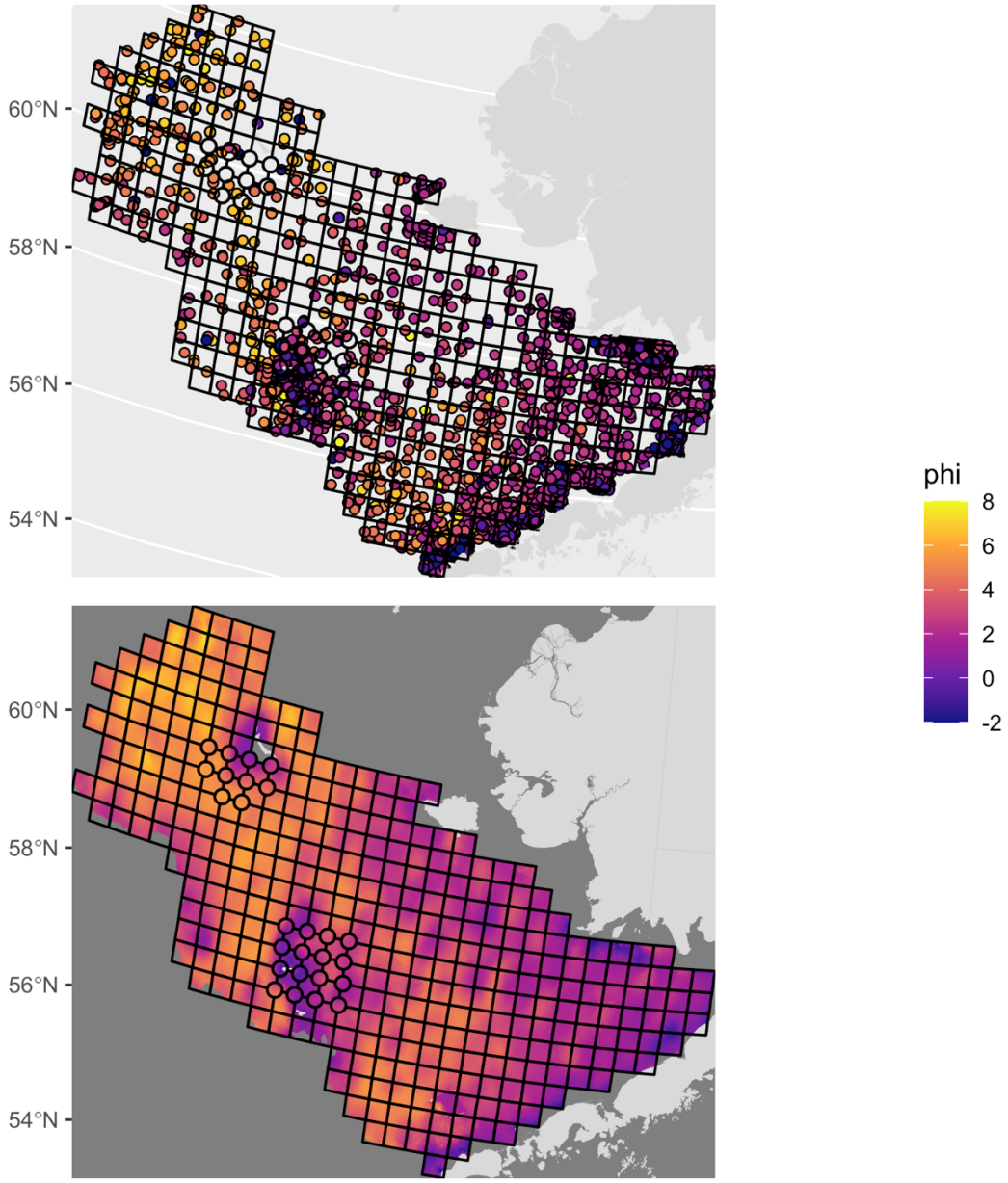

Figure 4. Upper: mean grain size (φ) data from dbSEABED; lower: interpolated 1km x 1km raster. Grain size is expressed on a logarithmic scale, with representing the negative log2 grain diameter in millimeters. Mean grainsize data at sediment sampling sites were kriged in ArcGIS using Empirical Bayes co-kriging and converted to a raster to allow interpolation to NMFS survey haul locations.

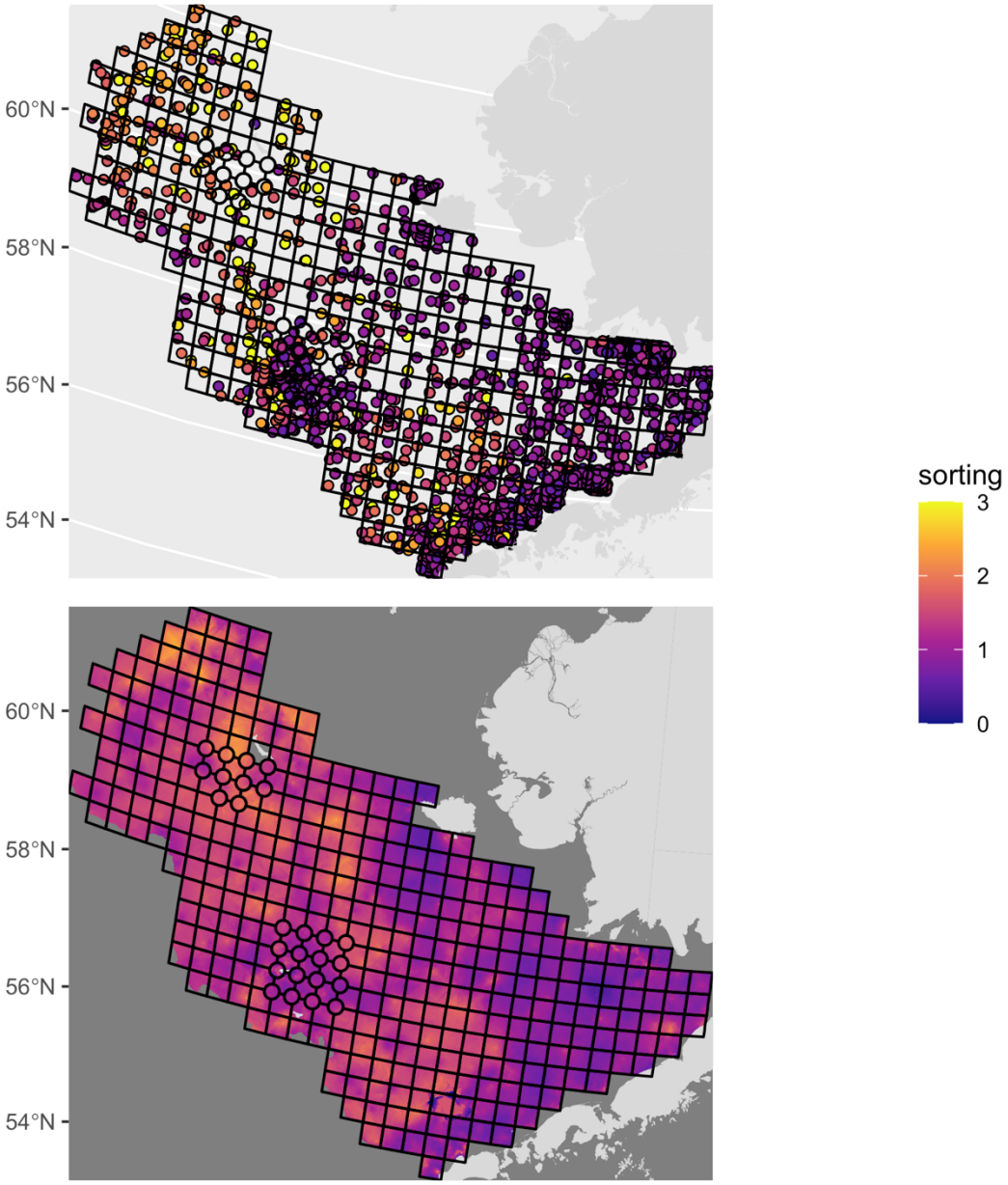

Figure 5. Upper: Sorting coefficient data from dbSEABED; lower: interpolated 1km x 1km raster. The sorting coefficient is a measure of size variation in a sediment sample. Sorting coefficients at sediment sampling sites were kriged in ArcGIS using Empirical Bayes co-kriging and converted to a raster to allow interpolation to NMFS survey haul locations.

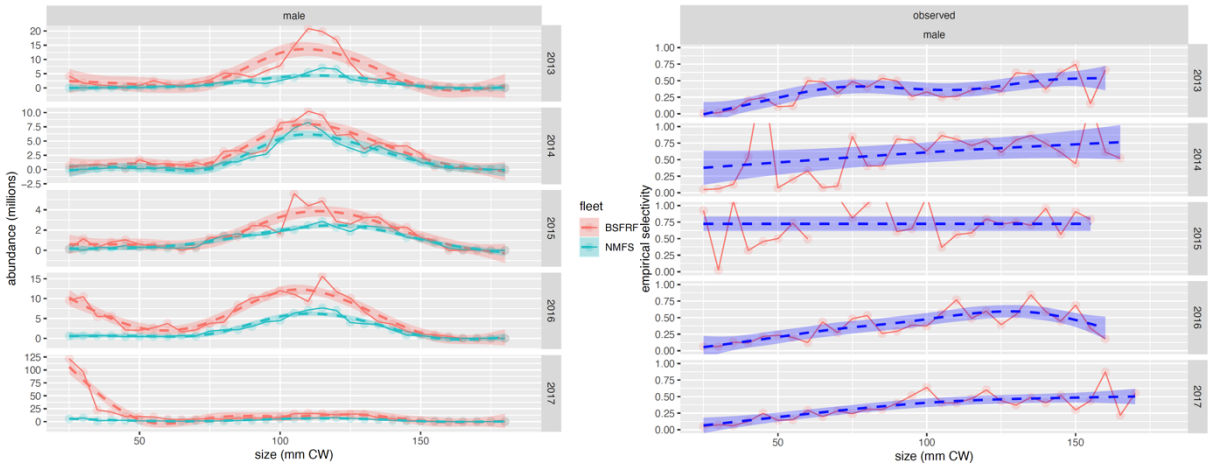

Figure 3.1. Lefthand plot: BSFRF (orange) and NMFS (green) size compositions for male Tanner crab from the SBS study areas for 2013-2017. Solid lines and dots are "raw" estimates, dashed lines are smoothed fits using cubic splines. Righthand plot: empirical catchability curves calculated using Equation 6. Red lines and dots: "raw" curves, dashed lines and fills: smoothed fits. All smooth fits are illustrative only.

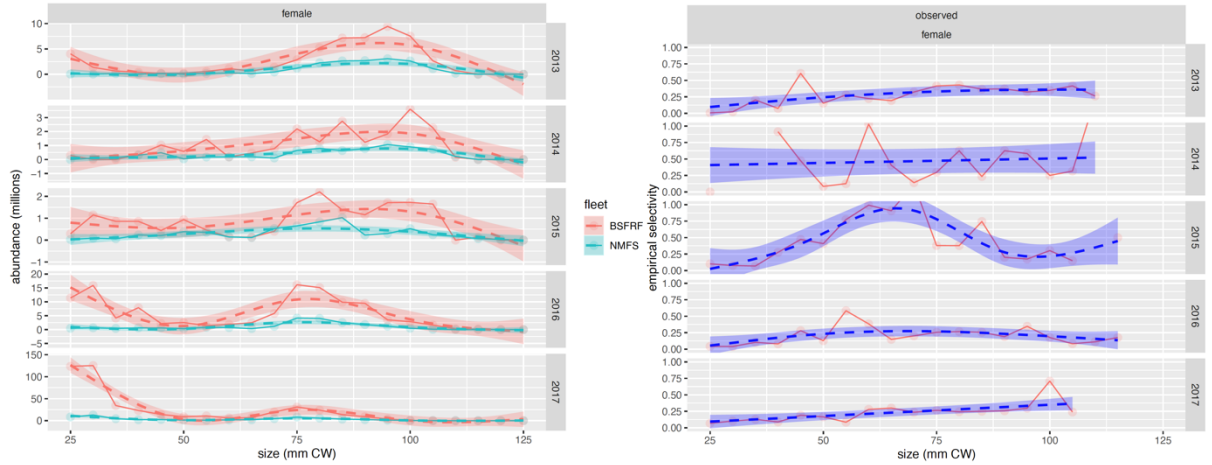

Figure 3.2. Lefthand plot: BSFRF (orange) and NMFS (green) size compositions for female Tanner crab from the SBS study areas for 2013-2017. Solid lines and dots are "raw" estimates, dashed lines are smoothed fits using cubic splines. Righthand plot: empirical catchability curves calculated using Equation 6. Red lines and dots: "raw" curves, dashed lines and fills: smoothed fits. All smooth fits are illustrative only.

`

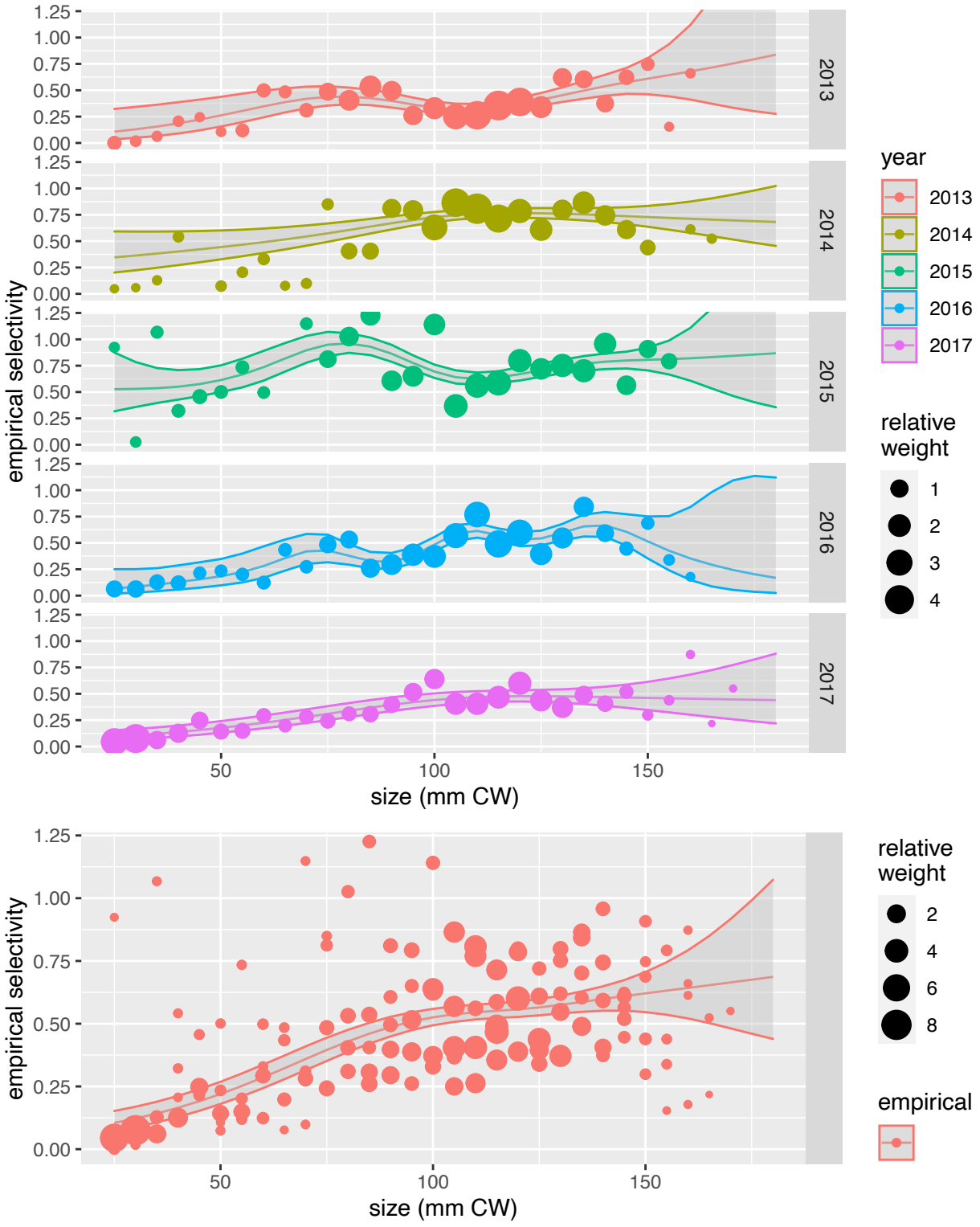

Figure 3.3. Estimated empirical catchability curves for males based on fitting the "raw" empirical catchability estimates in Figure 3.1 using generalized additive models (GAMs) with log link functions and normal distributions. Upper plot: fits to annual empirical catchability estimates. Lower plot: "global" fit to all estimates. The center line in each graph represents the model fit, the fills represent 80% confidence intervals.

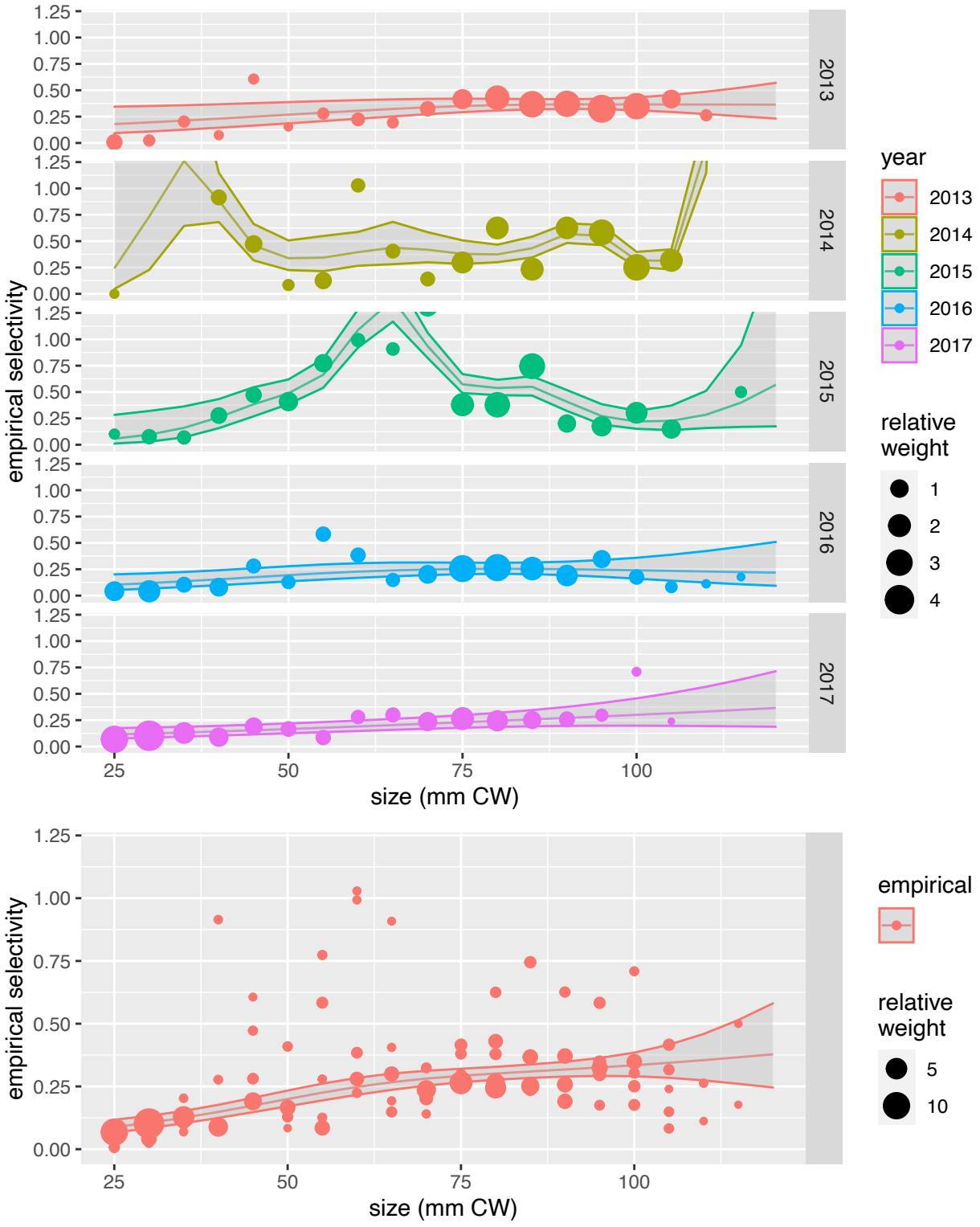

Figure 3.3. Estimated empirical catchability curves for females based on fitting the "raw" empirical catchability estimates in Figure 3.1 using generalized additive models (GAMs) with log link functions and normal distributions. Upper plot: fits to annual empirical catchability estimates. Lower plot: "global" fit to all estimates. The center line in each graph represents the model fit, the fills represent 80% confidence intervals.

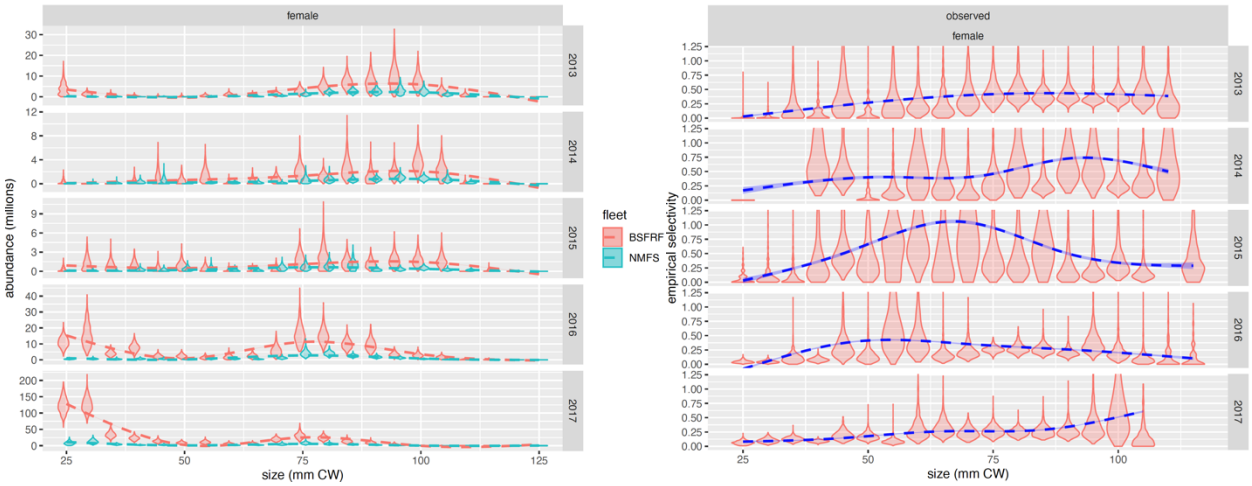

Figure 3.4. Lefthand plot: Bootstrapped BSFRF (orange) and NMFS (green) size compositions for female Tanner crab from the SBS survey areas for 2013-2017. Violin plots summarize "raw" estimates, dashed lines are smoothed fits using cubic splines. Righthand plot: bootstrapped empirical selectivity curves calculated using Equation 5. Red violin plots: "raw" estimates, dashed lines: smoothed fits.

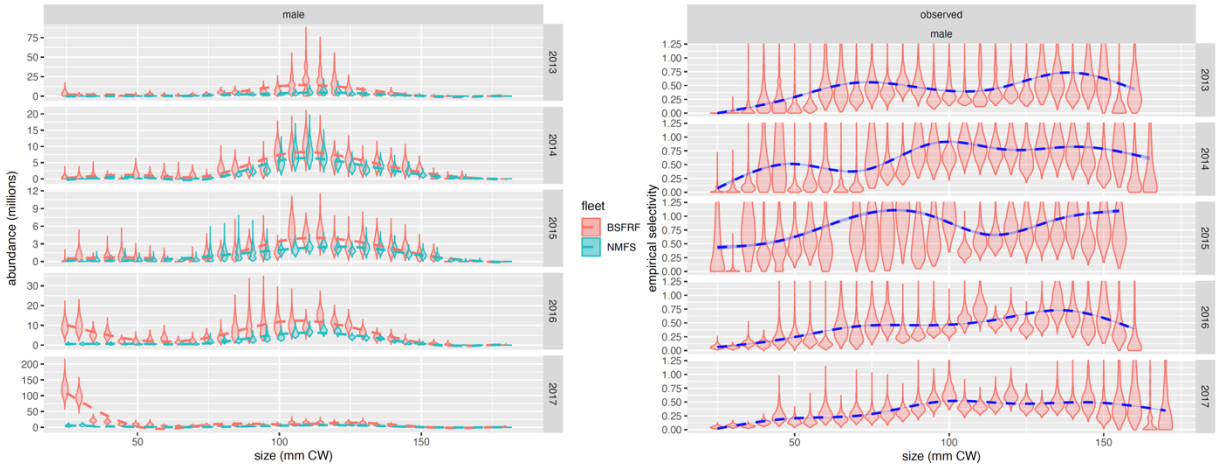

Figure 3.5. Lefthand plot: Bootstrapped BSFRF (orange) and NMFS (green) size compositions for female Tanner crab from the SBS survey areas for 2013-2017. Violin plots summarize "raw" estimates, dashed lines are smoothed fits using cubic splines. Righthand plot: bootstrapped empirical selectivity curves calculated using Equation 5. Red violin plots: "raw" estimates, dashed lines: smoothed fits.

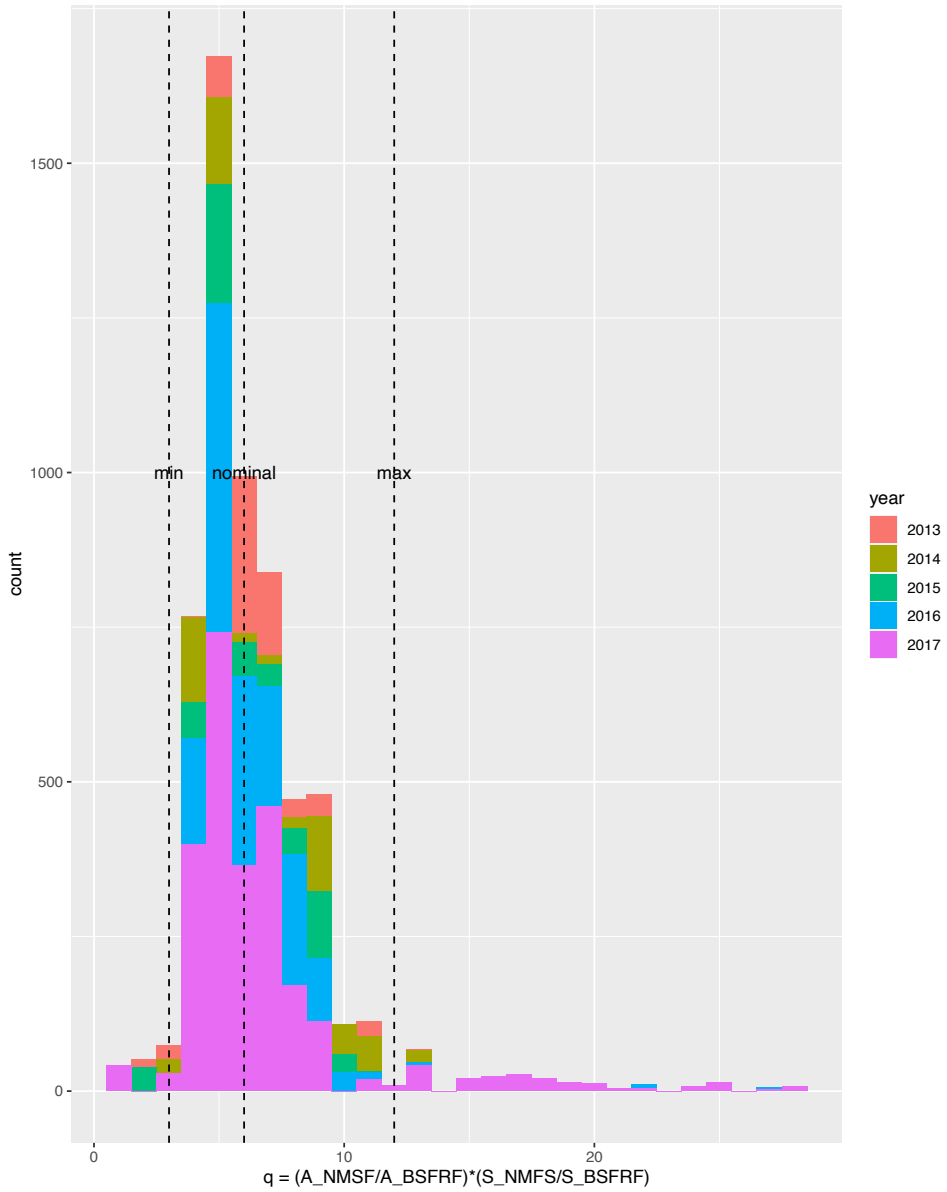

Figure 4-2.6. Histogram of q values from SBS data. The nominal value ($q_{nom} \approx 6$) is based on standard haul practice on both NMSF and BSFRF survey vessels (no sub-sampling; NMFS: 30-minute tow at 3 knots; BSFRF: 5-minute tow at 2 knots). Hauls with q values outside the min/max range were considered problematic and removed from further consideration in the haul-level catch ratio analysis.

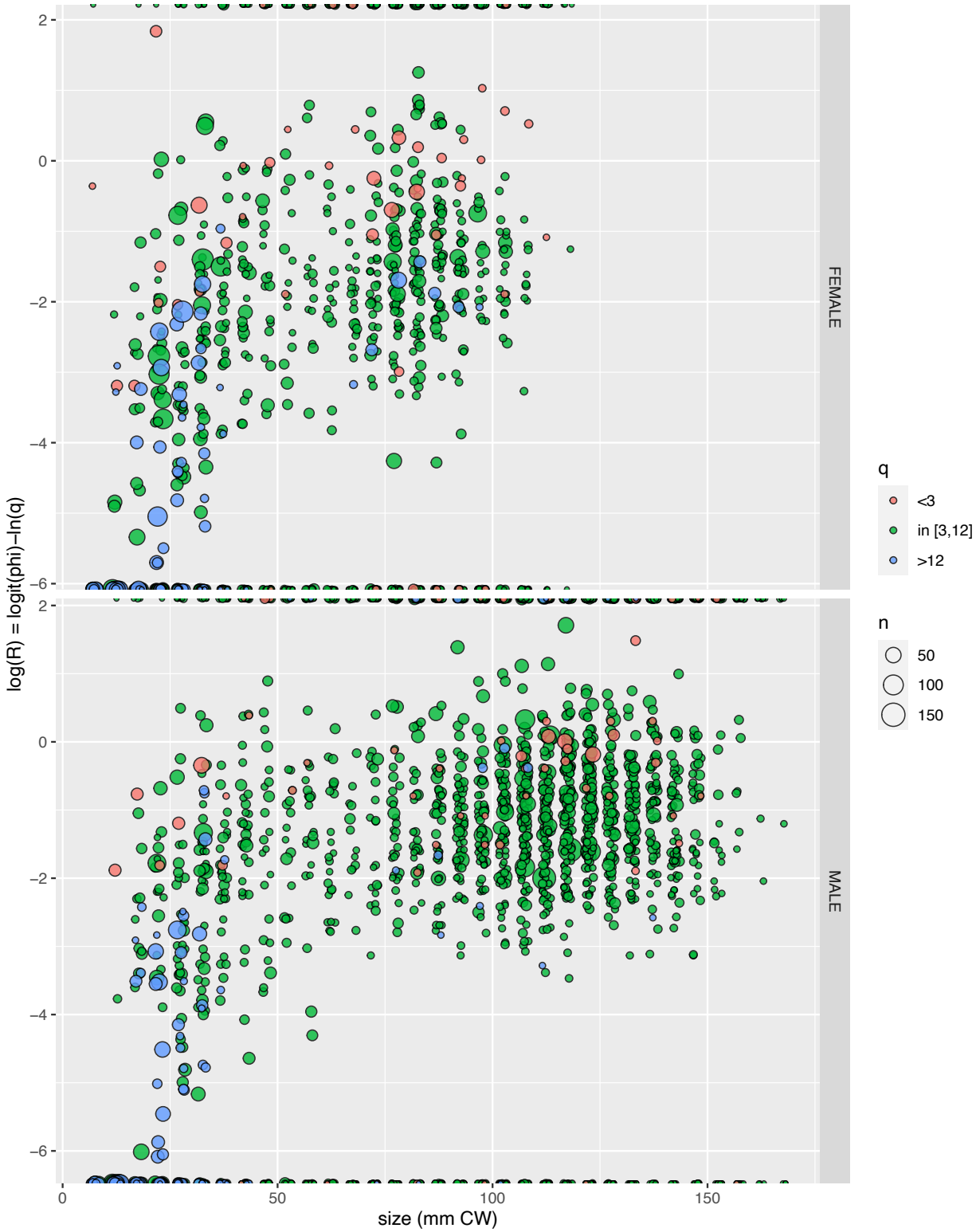

Figure 4.2-7. Observed values of log(*R*). Colors indicate associated values of *q*. Symbol area represents number of the total number of individuals. Symbols at the top and bottom of each plot represent values of log(*R*) at +∞ and -∞, respectively.

Figure 4.2-8. Observed values of log(R). Colors indicate year of survey. Black dots indicate observations with a value of q outside the acceptable range ($[\frac{1}{3}, 3] \cdot q_{nom}$). Symbols at the top and bottom of each plot represent values of log(R) at +∞ and -∞, respectively.

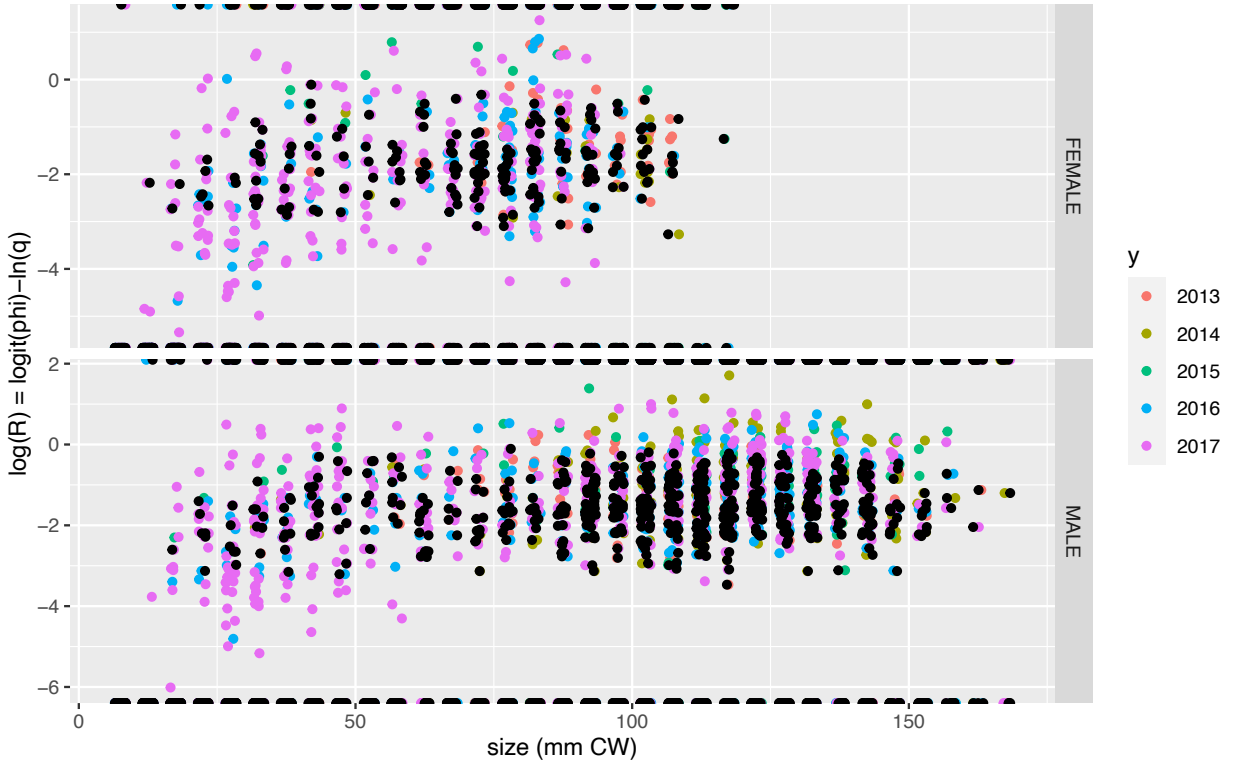

Figure 4.2-9. Observed values of log(*R*). Observations with < 5 total individuals are colored black.

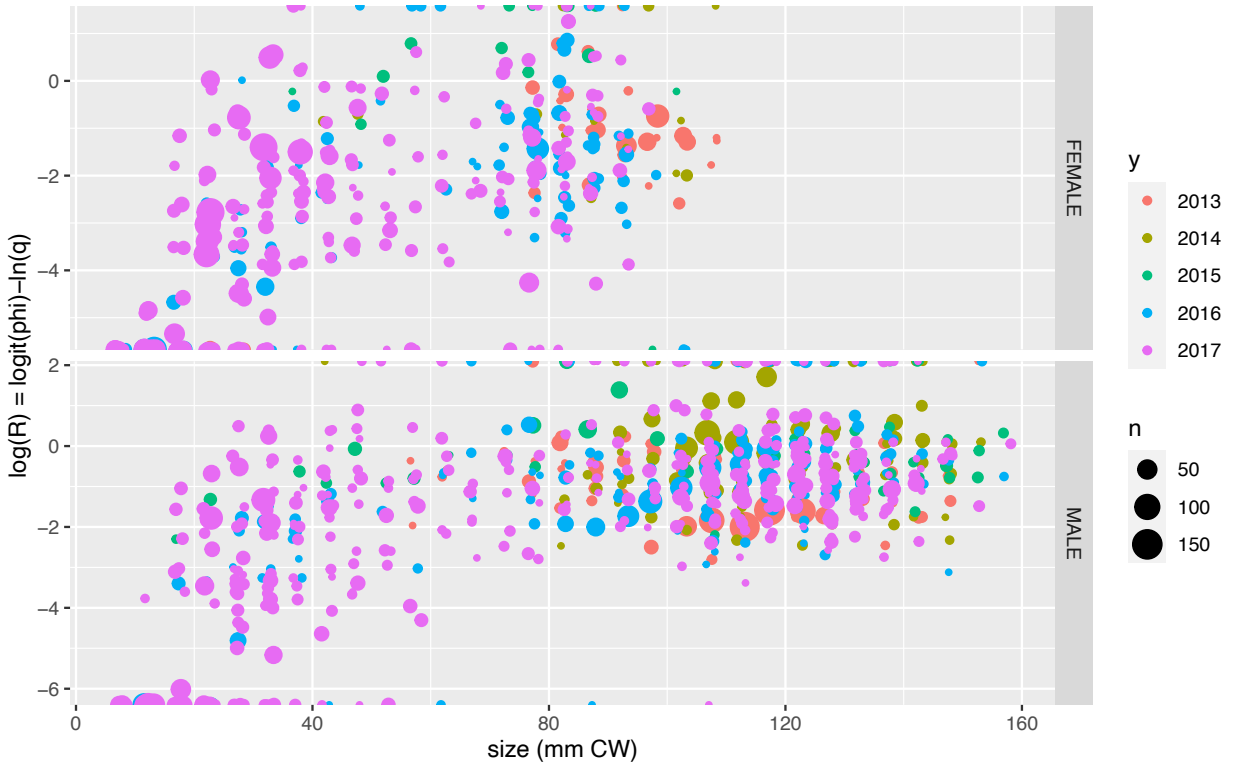

Figure 4.2-10. Subset of observed values of log(R) to be fit. Observations with < 5 total individuals were removed, as were hauls with q outside the acceptable limits ($[\frac{1}{3}, 3] \cdot q_{nom}$). Colors indicate year of survey. Symbol area reflects the total number of individuals captured. Symbols at the top and bottom of each plot represent values of log(R) at +∞ and -∞, respectively.

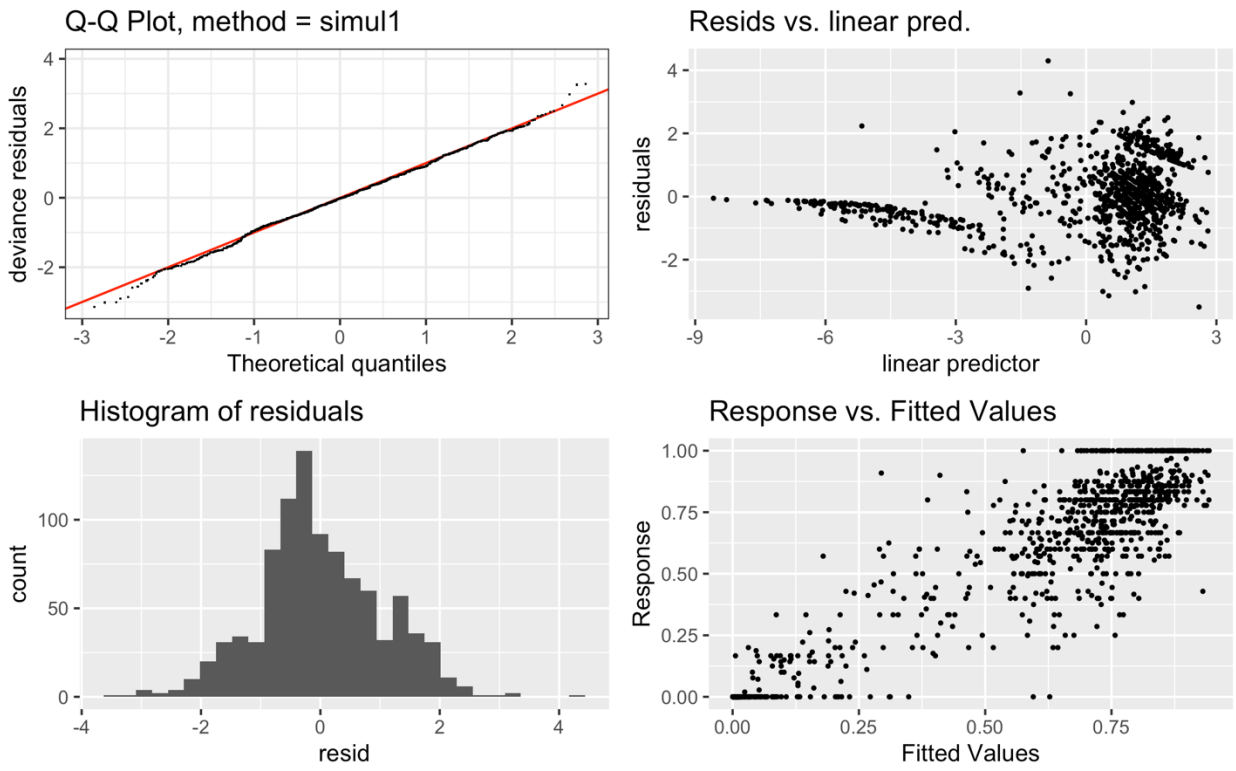

Figure 4.3-1. Diagnostic plots based on results from the “best” model for males.

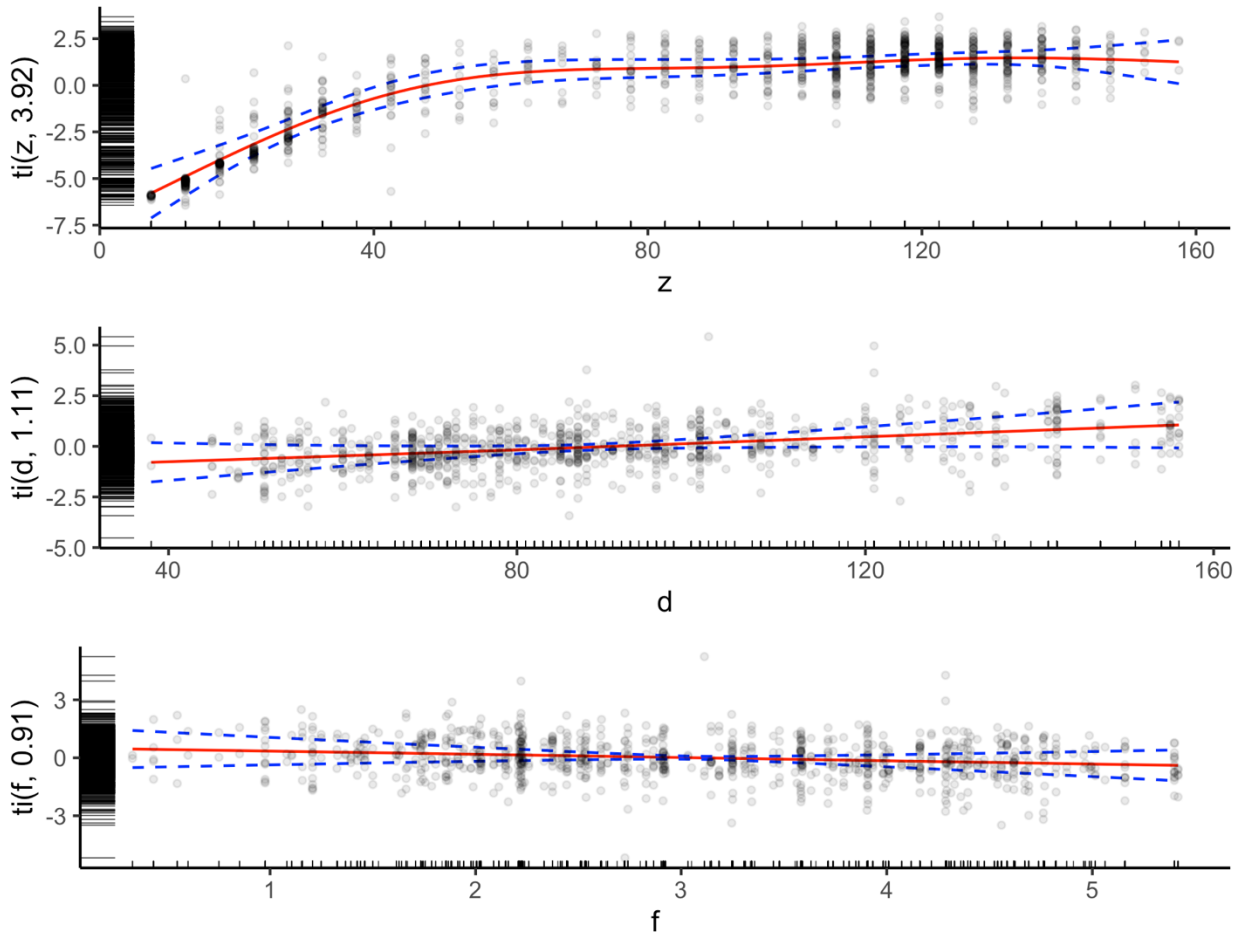

Figure 4.3-2. Component 1-d "smooths" contributing to model for males. "*z*": crab size (mm CW), "*d*": bottom depth (m), "*f*": φ, $\log_2$-scale mean grain size. Circles indicate covariate values contributing to the model fit.

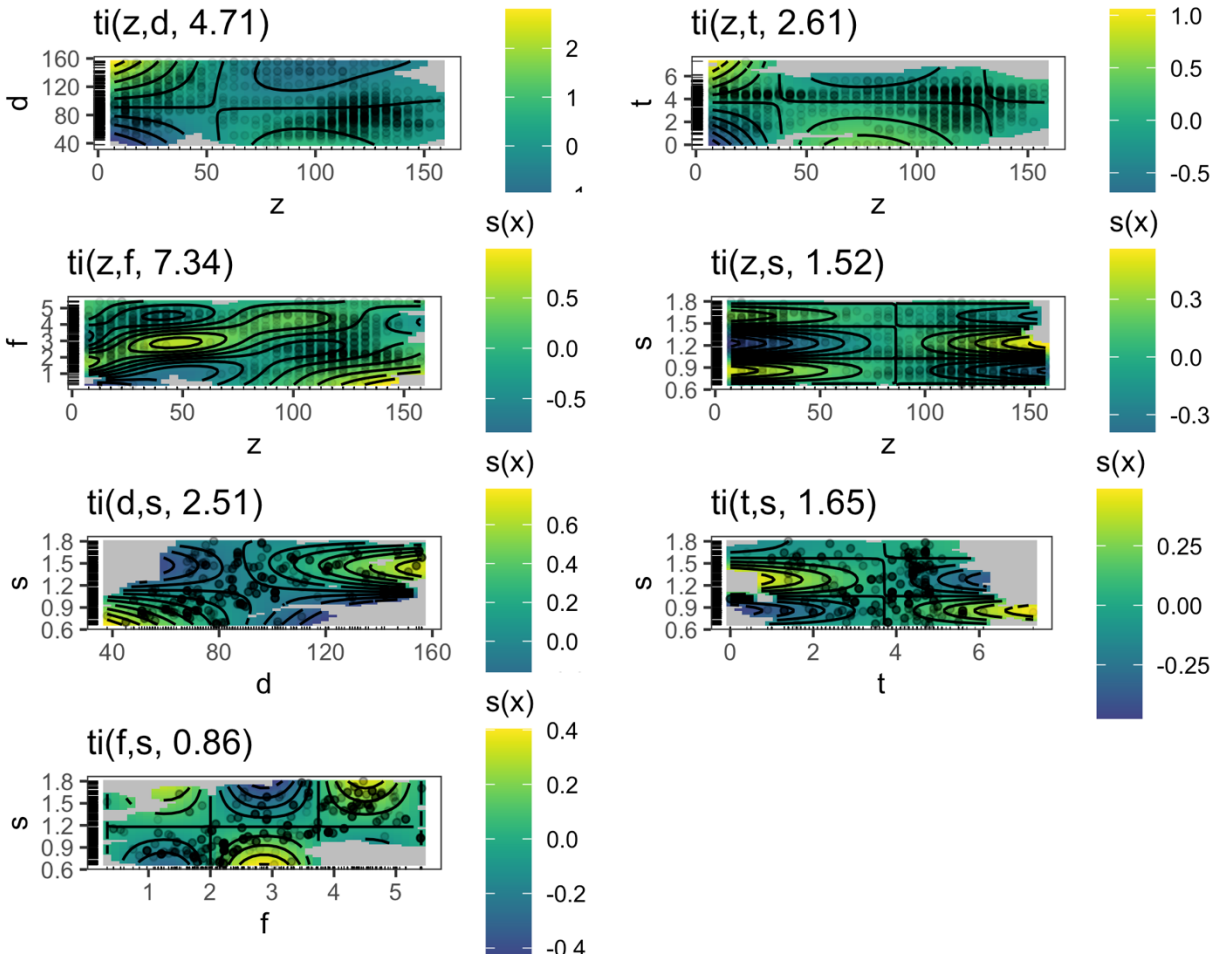

Figure 4.3-3. Component 2-d tensor interaction "smooths" (*ti*) contributing to model for males. *z*: crab size, *d*: depth, *t*: bottom temperature, *f*: φ , $\log_2$-scale mean grain size, *s*: sediment sorting coefficient. Circles indicate covariate values contributing to the model fit.

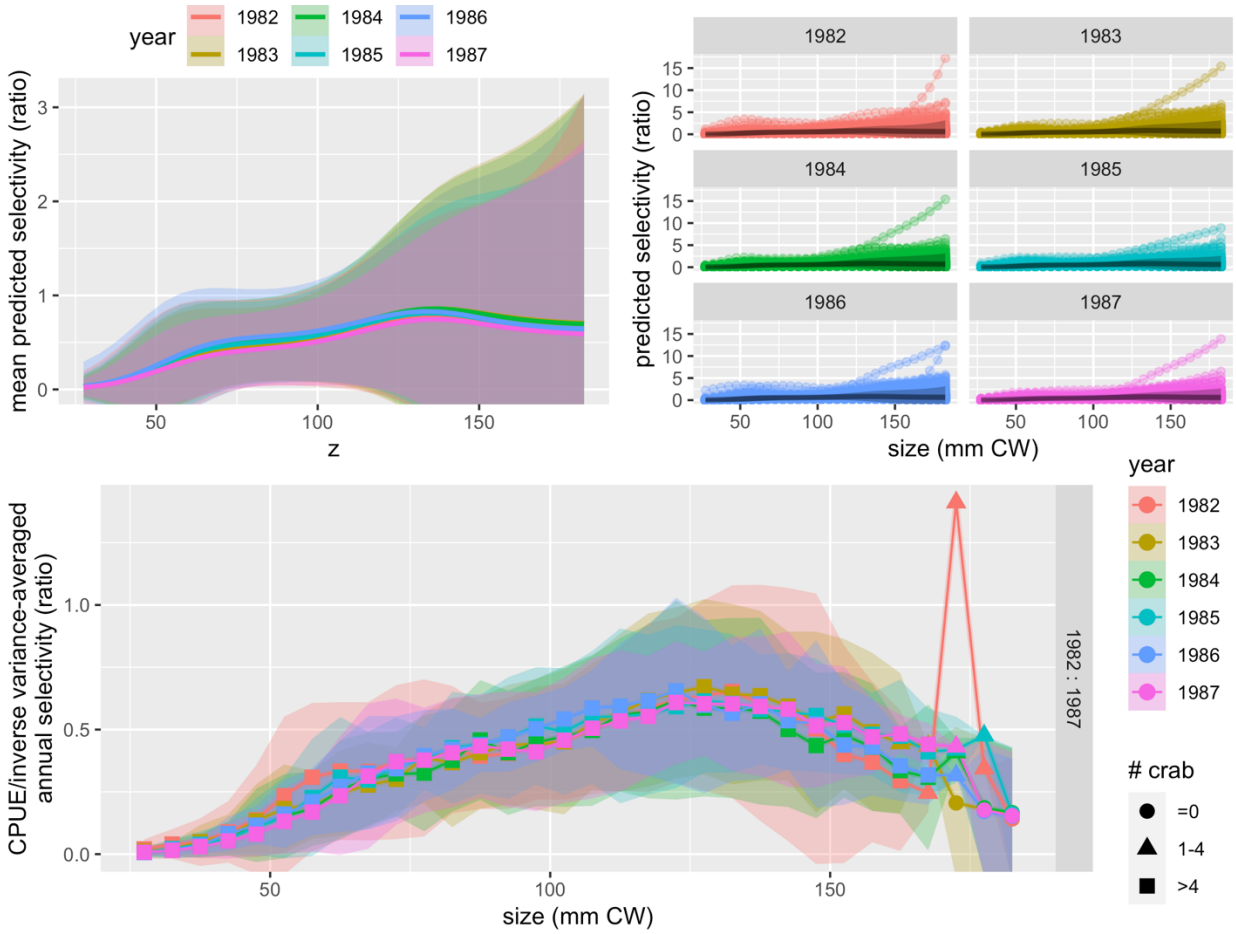

Figure 4.3-4. Predicted NMFS survey male catchability for 1982-1987. Upper row: unweighted mean and 95% confidence intervals (left); individual haul-level estimates (right). Lower row: Inverse-variance weighted estimates of annual NMFS survey catchability, with 95% confidence intervals.

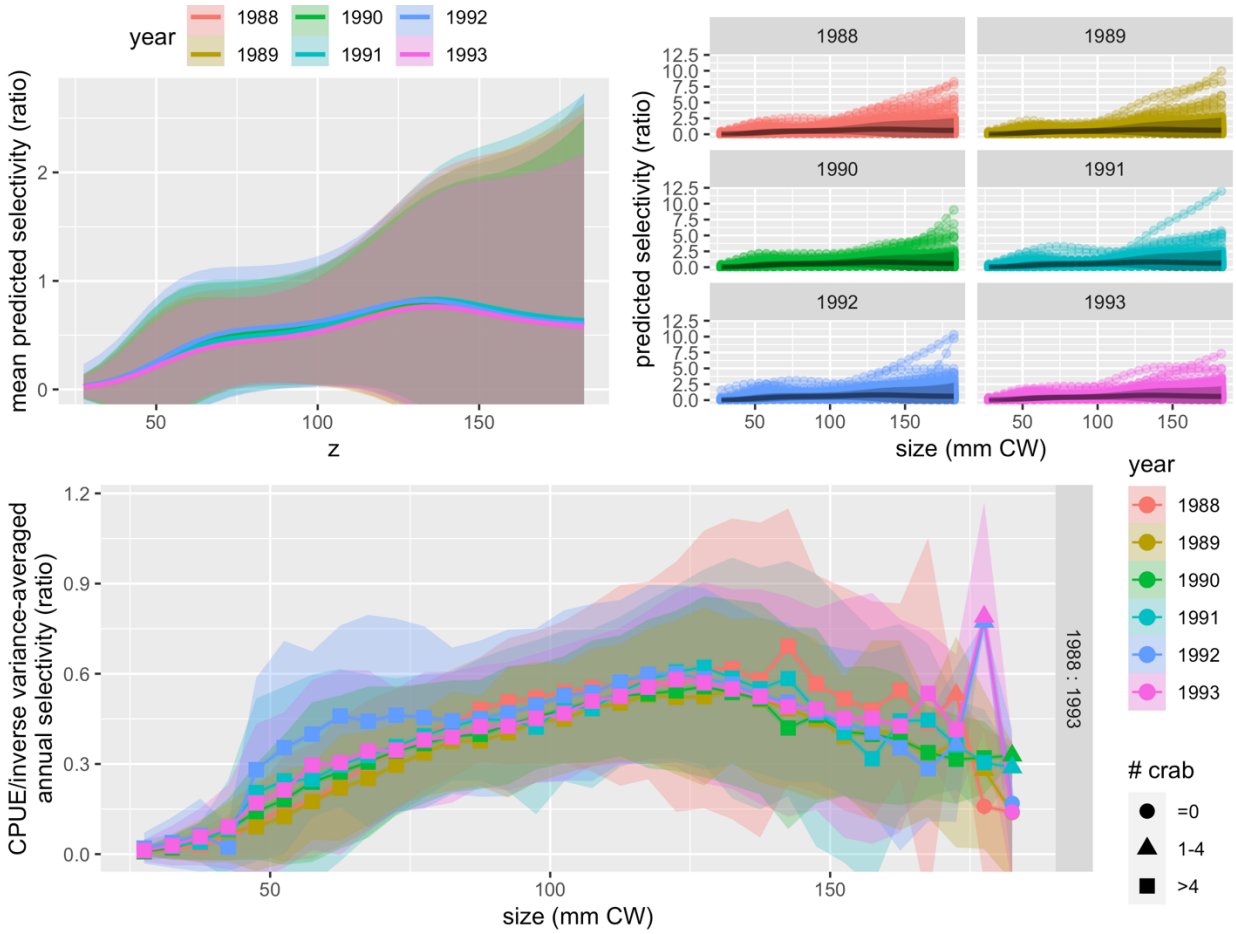

Figure 4.3-5. Predicted NMFS survey male catchability for 1988-1987. Upper row: unweighted mean and 95% confidence intervals (left); individual haul-level estimates (right). Lower row: Inverse-variance weighted estimates of annual NMFS survey catchability, with 95% confidence intervals.

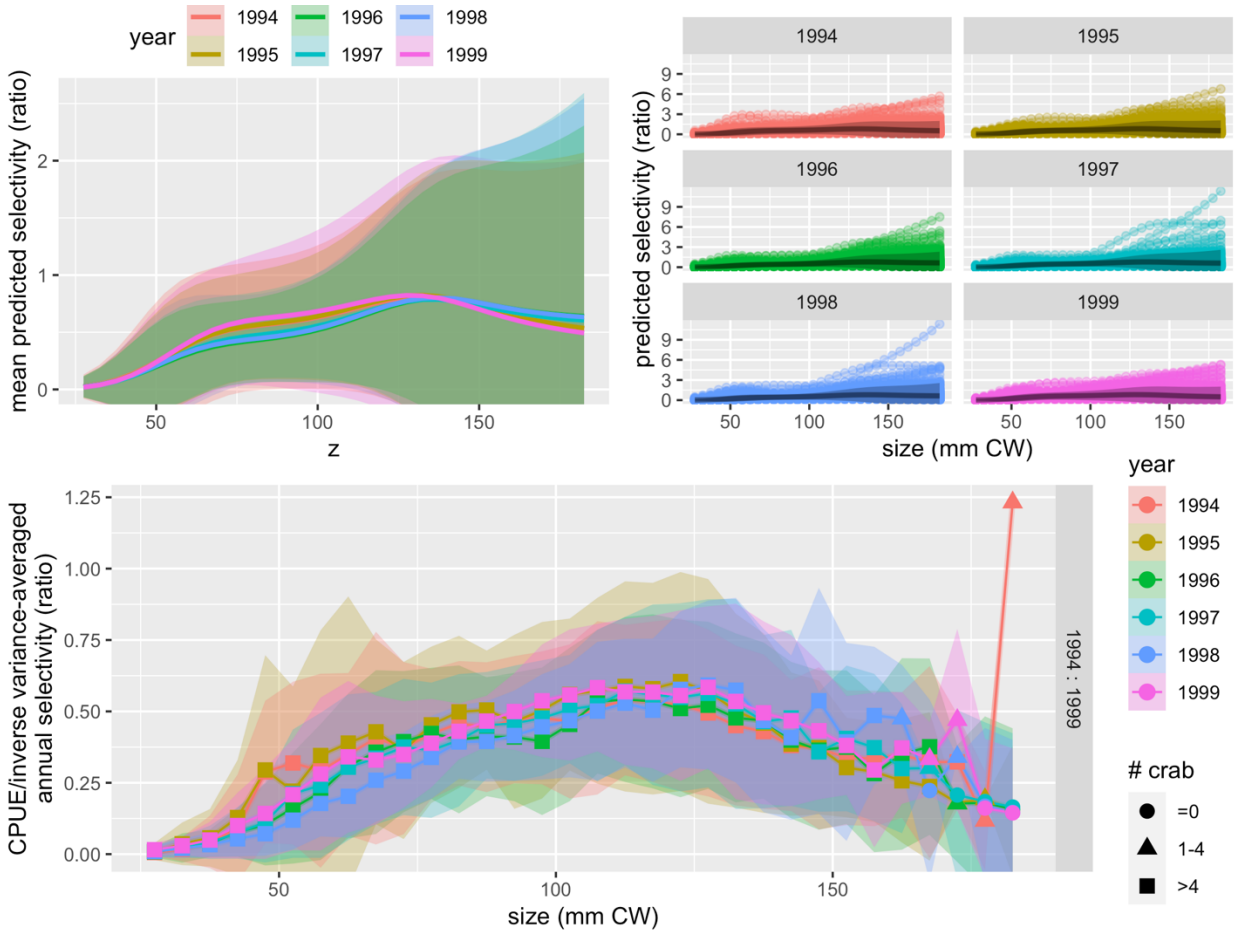

Figure 4.3-6. Predicted NMFS survey male catchability for 1994-1999. Upper row: unweighted mean and 95% confidence intervals (left); individual haul-level estimates (right). Lower row: Inverse-variance weighted estimates of annual NMFS survey catchability, with 95% confidence intervals.

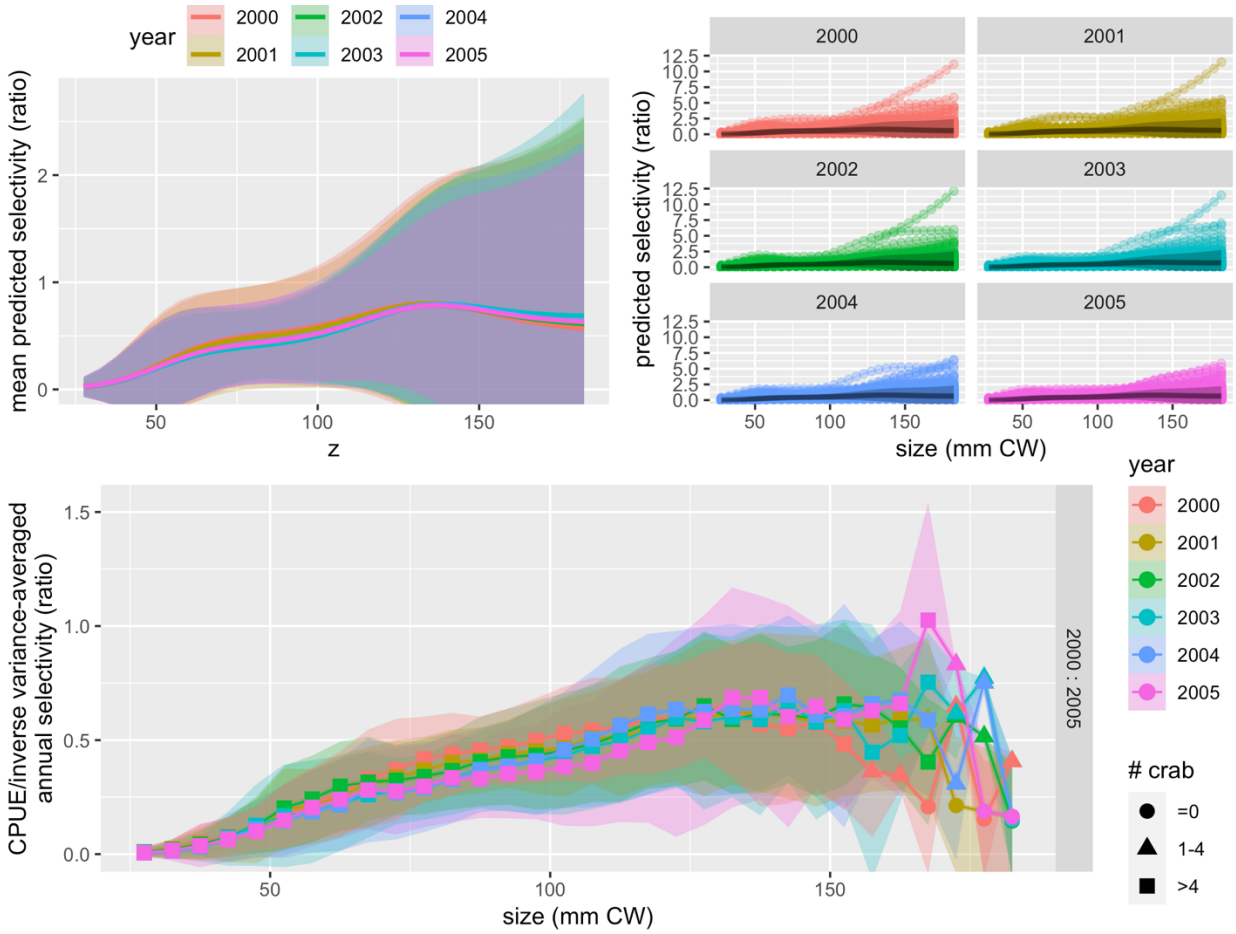

Figure 4.3-7. Predicted NMFS survey male catchability for 2000-2005. Upper row: unweighted mean and 95% confidence intervals (left); individual haul-level estimates (right). Lower row: Inverse-variance weighted estimates of annual NMFS survey catchability, with 95% confidence intervals.

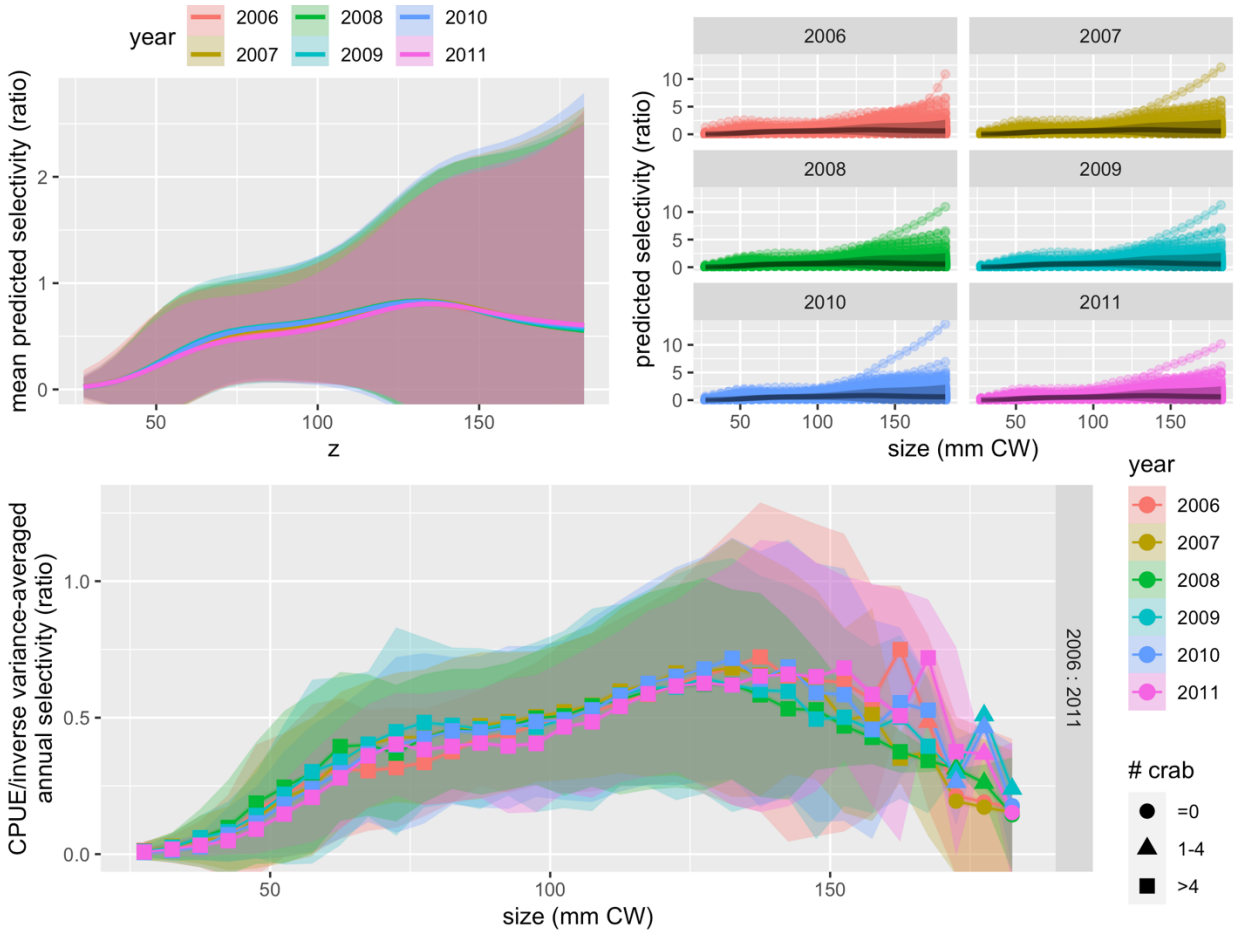

Figure 4.3-8. Predicted NMFS survey male catchability for 2006-2011. Upper row: unweighted mean and 95% confidence intervals (left); individual haul-level estimates (right). Lower row: Inverse-variance weighted estimates of annual NMFS survey catchability, with 95% confidence intervals.

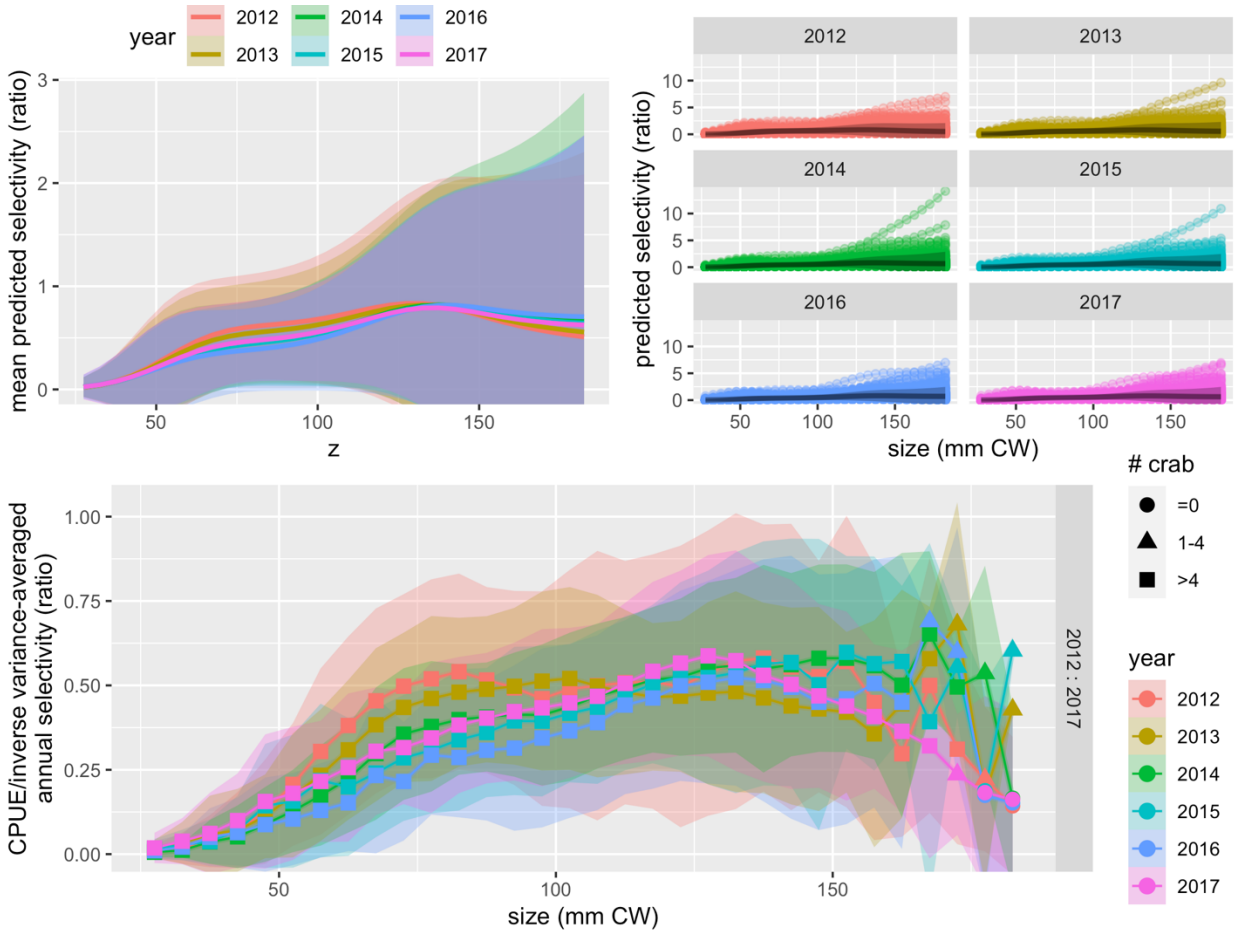

Figure 4.3-9. Predicted NMFS survey male catchability for 2012-2017. Upper row: unweighted mean and 95% confidence intervals (left); individual haul-level estimates (right). Lower row: Inverse-variance weighted estimates of annual NMFS survey catchability, with 95% confidence intervals.

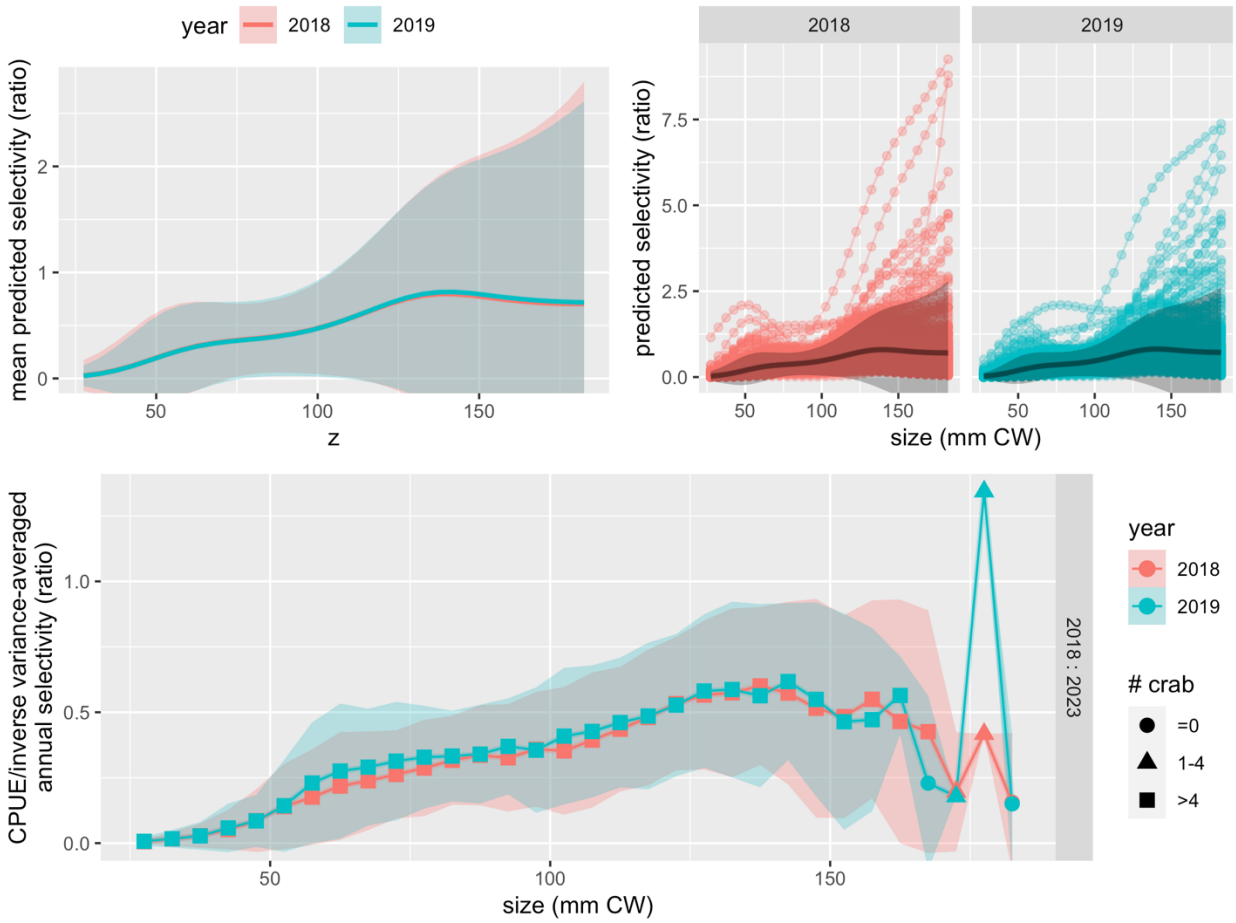

Figure 4.3-10. Predicted NMFS survey male catchability for 2018-2019. Upper row: unweighted mean and 95% confidence intervals (left); individual haul-level estimates (right). Lower row: Inverse-variance weighted estimates of annual NMFS survey catchability, with 95% confidence intervals.

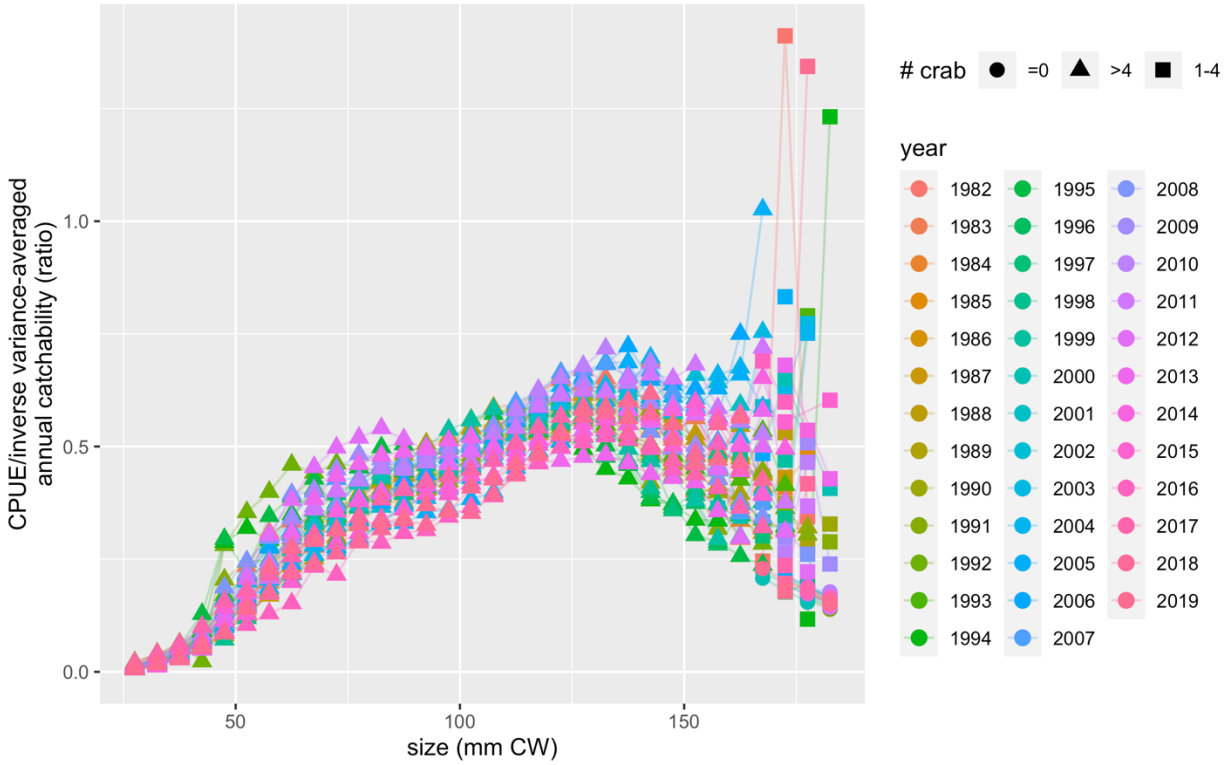

Figure 4.3-11. Predicted NMFS survey male catchability for all years.

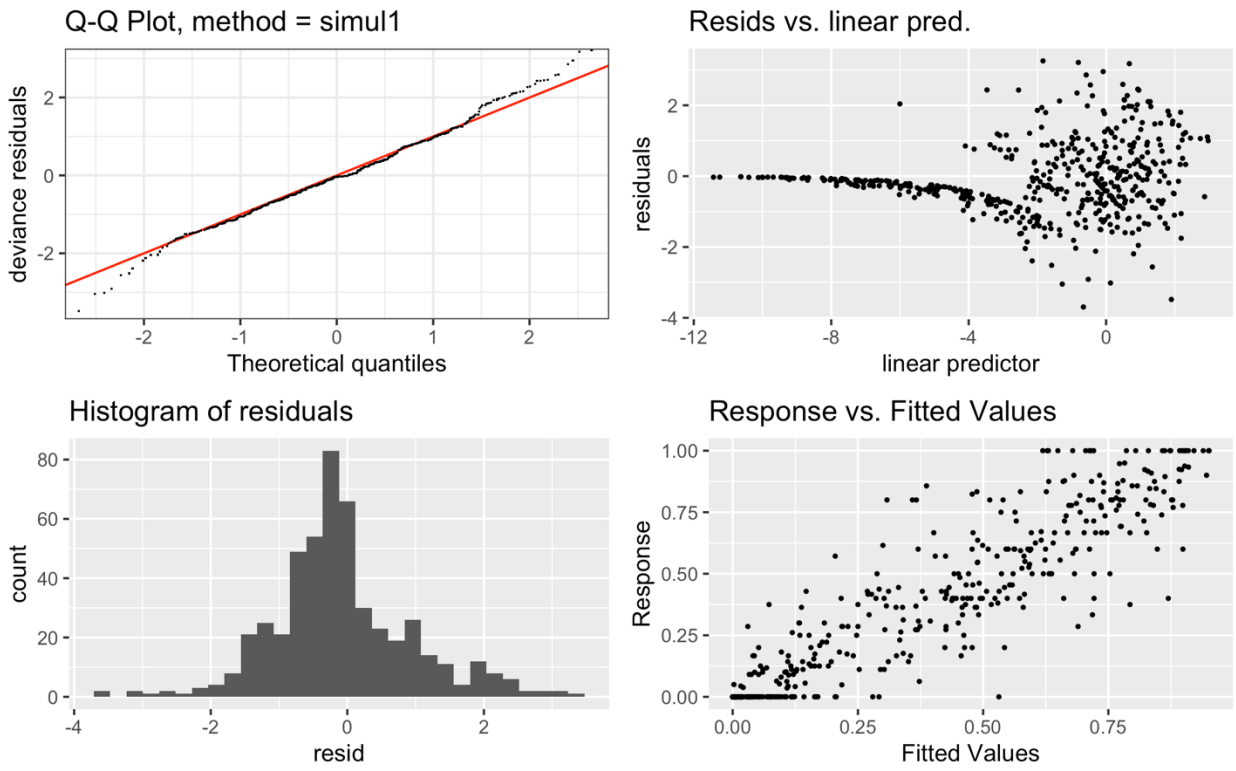

Figure 4.3-12. Diagnostic plots based on results from the “best” model for females.

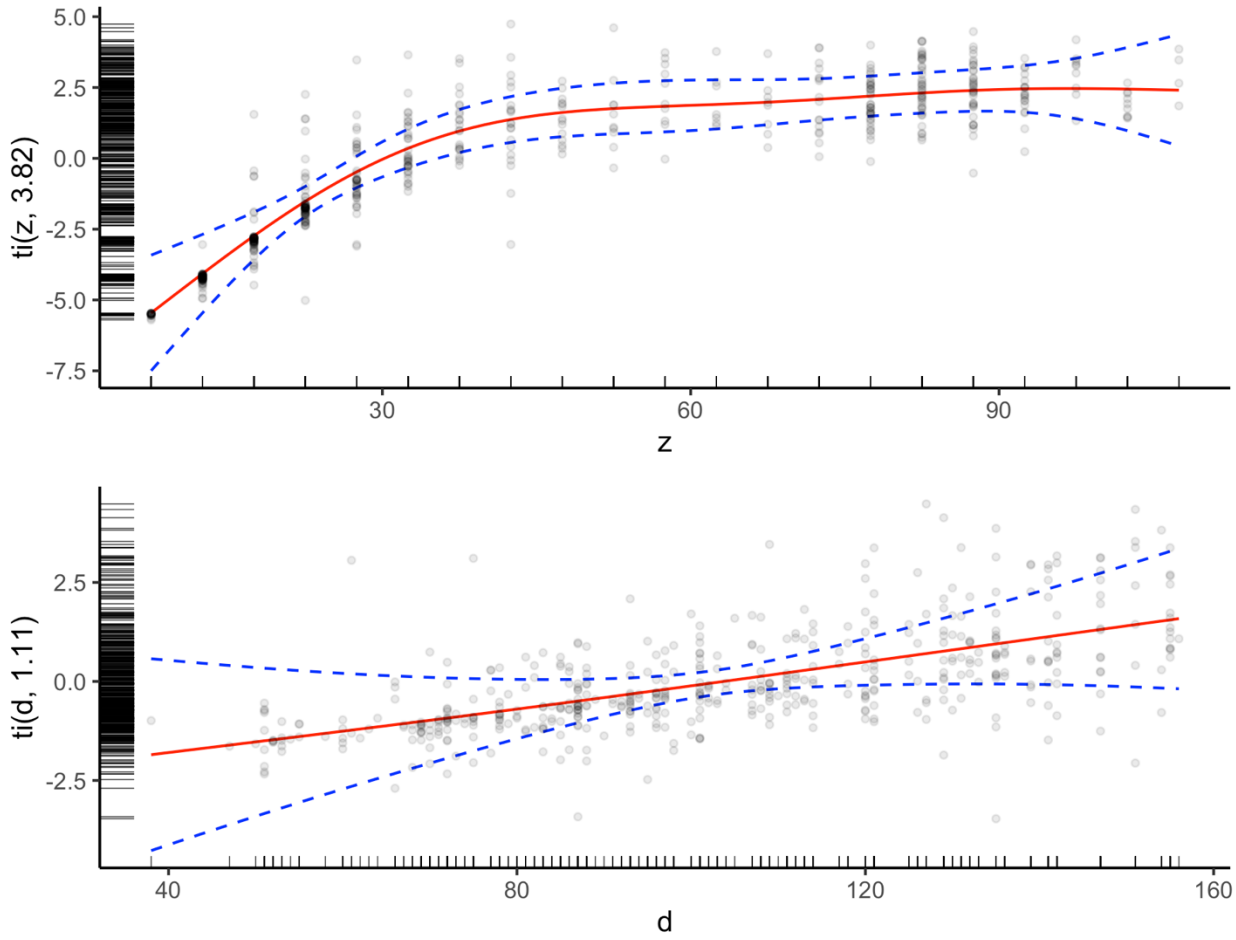

Figure 4.3-13. Component 1-d "smooths" contributing to model for females. "*z*": crab size (mm CW), "*d*": bottom depth (m). Circles indicate covariate values contributing to the model fit.

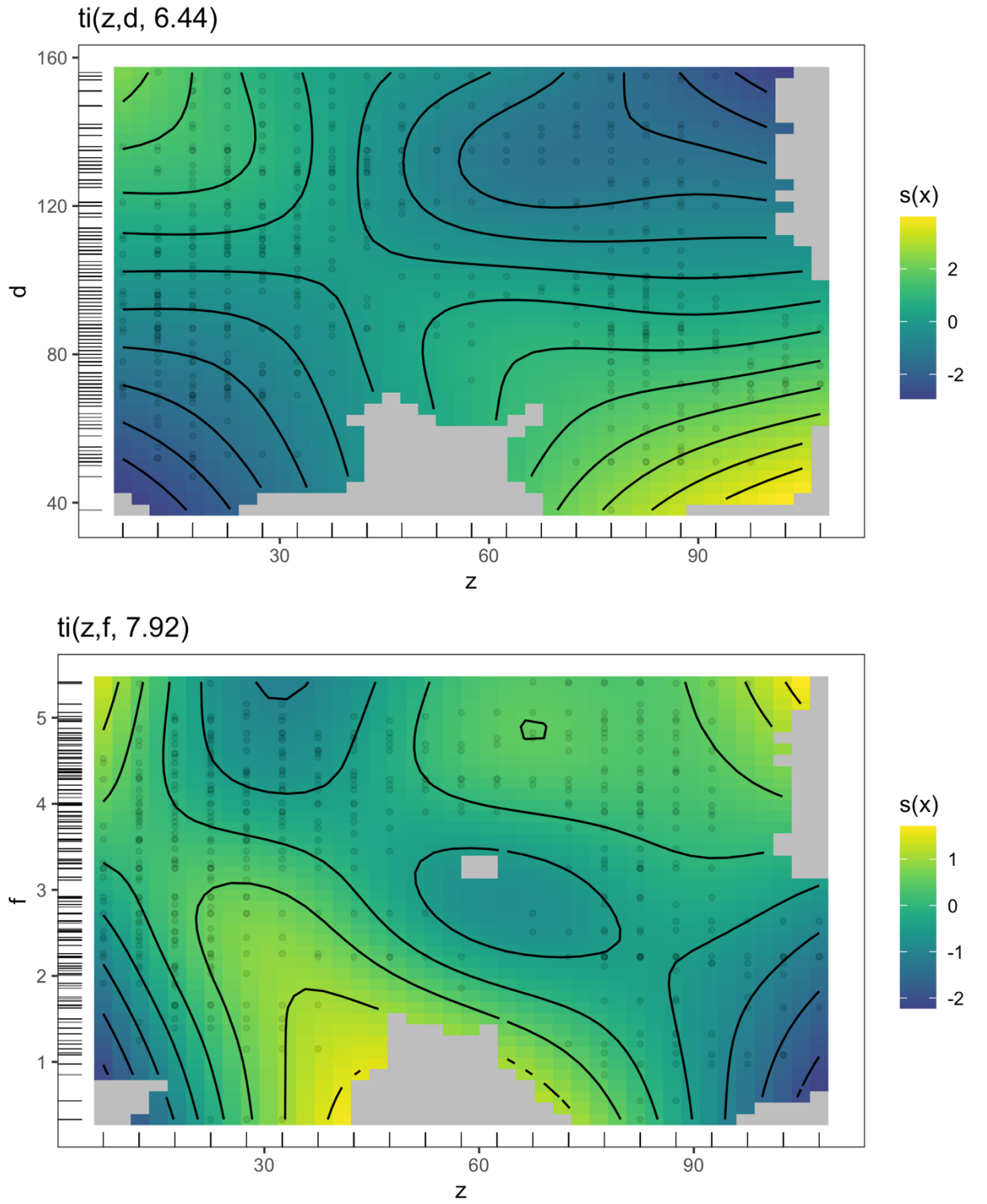

Figure 4.3-14. Component 2-d tensor interaction "smooths" contributing to model for females. "*z*": crab size (mm CW), "*d*": bottom depth (m), "*f*": φ , $\log_2$-scale mean grain size. Circles indicate covariate values contributing to the model fit.

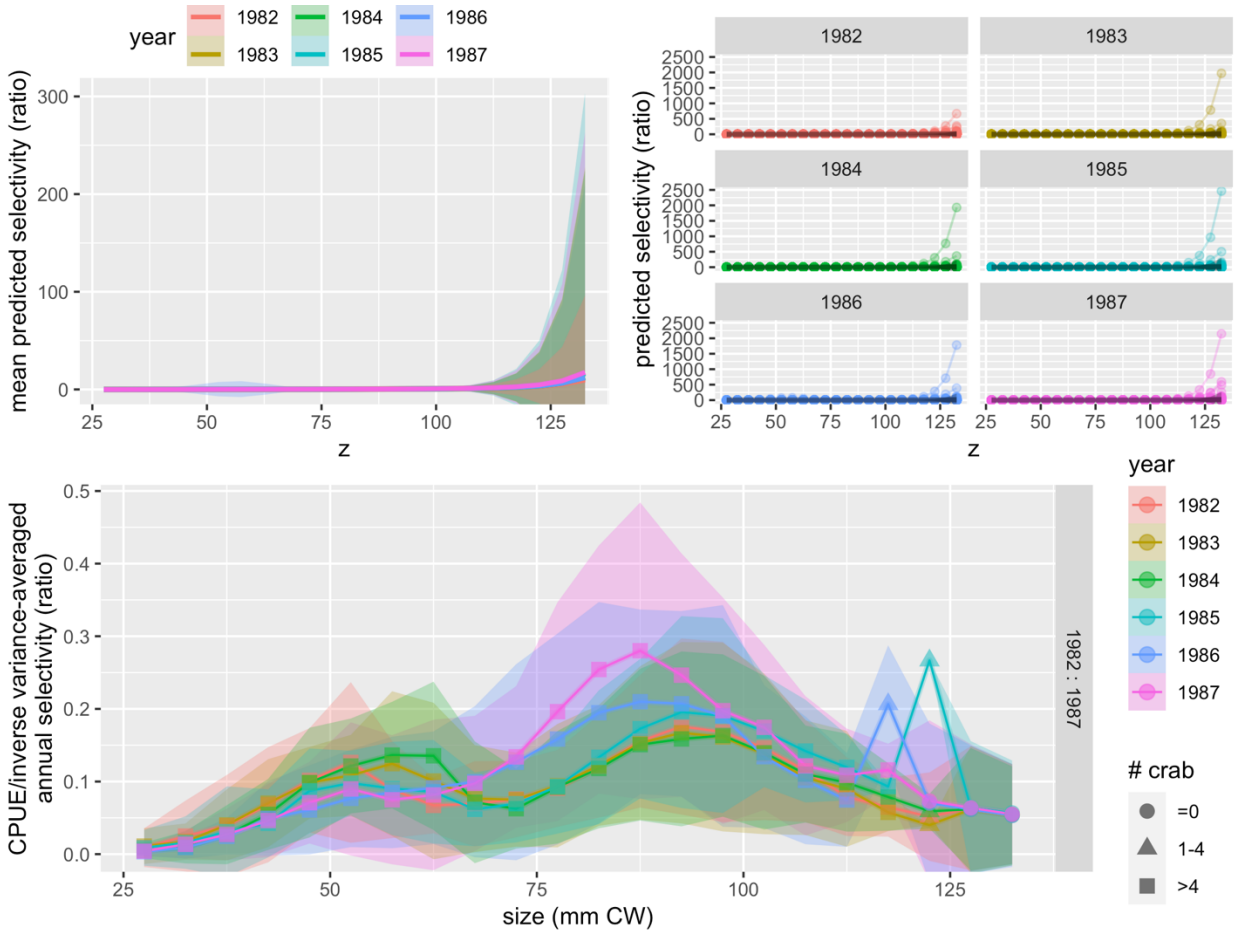

Figure 4.3-15. Predicted NMFS survey female catchability for 1982-1987. Upper row: unweighted mean and 95% confidence intervals (left); individual haul-level estimates (right). Lower row: Inverse-variance weighted estimates of annual NMFS survey catchability, with 95% confidence intervals.

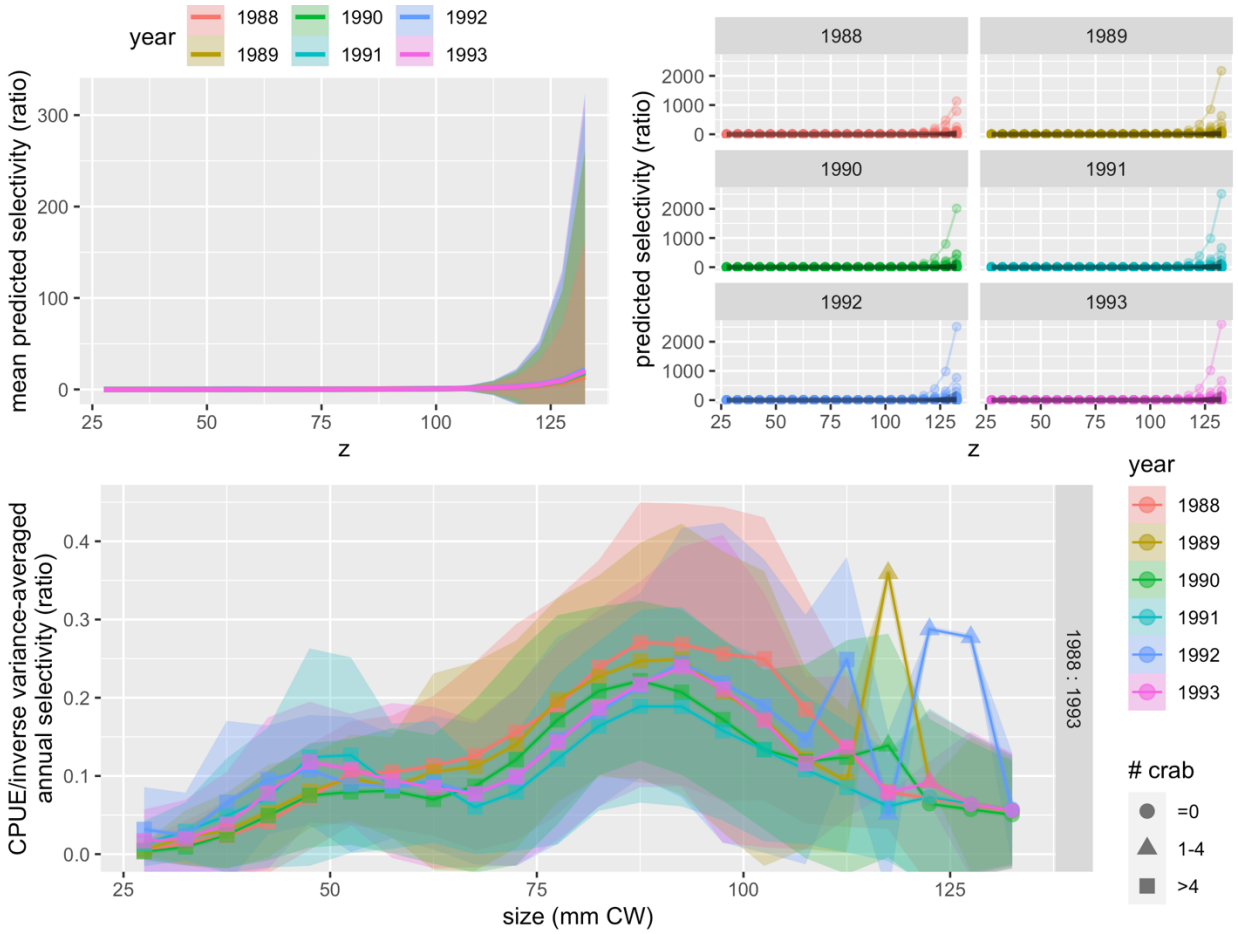

Figure 4.3-16. Predicted NMFS survey female catchability for 1988-1993. Upper row: unweighted mean and 95% confidence intervals (left); individual haul-level estimates (right). Lower row: Inverse-variance weighted estimates of annual NMFS survey catchability, with 95% confidence intervals.

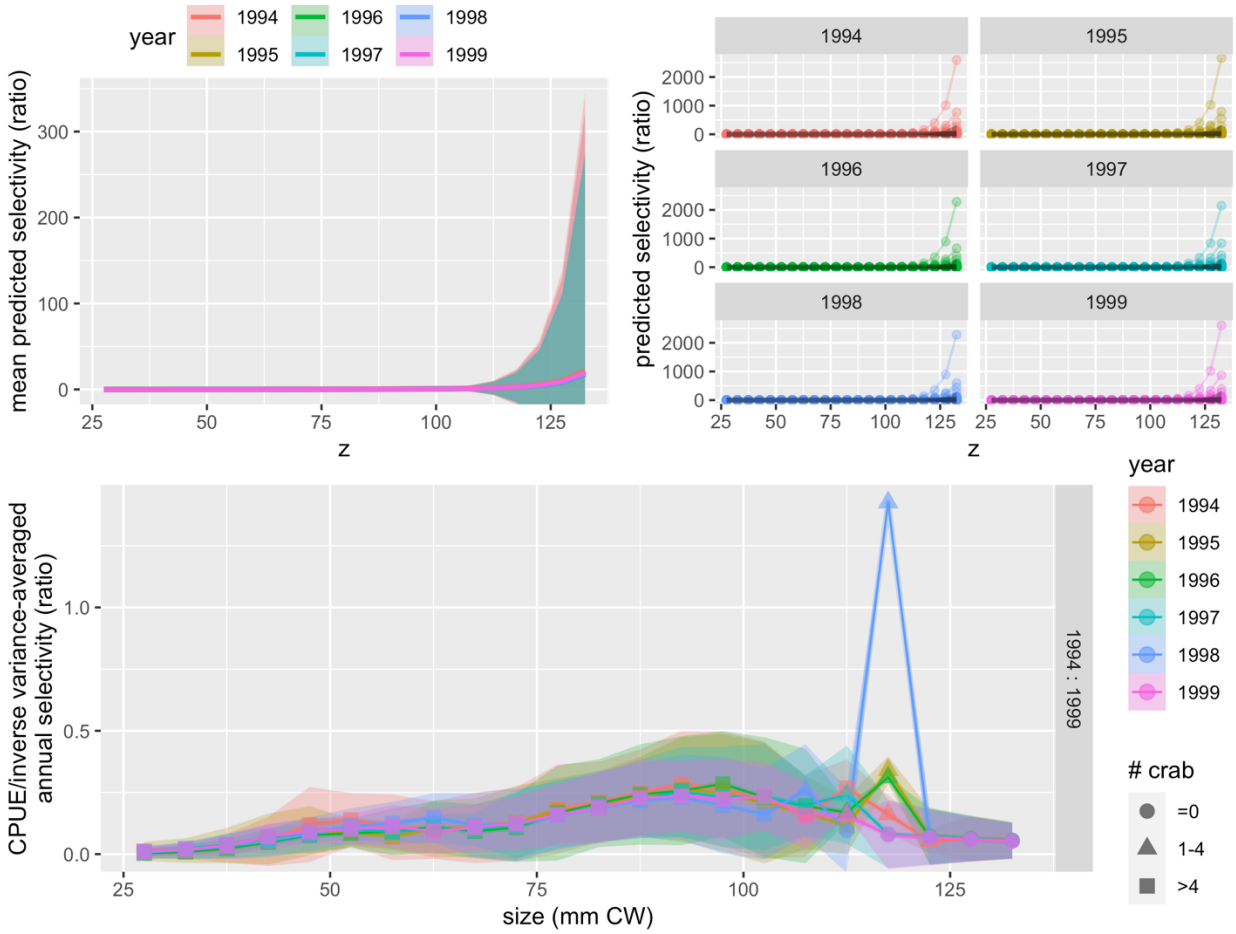

Figure 4.3-17. Predicted NMFS survey female catchability for 1994-1999. Upper row: unweighted mean and 95% confidence intervals (left); individual haul-level estimates (right). Lower row: Inverse-variance weighted estimates of annual NMFS survey catchability, with 95% confidence intervals.

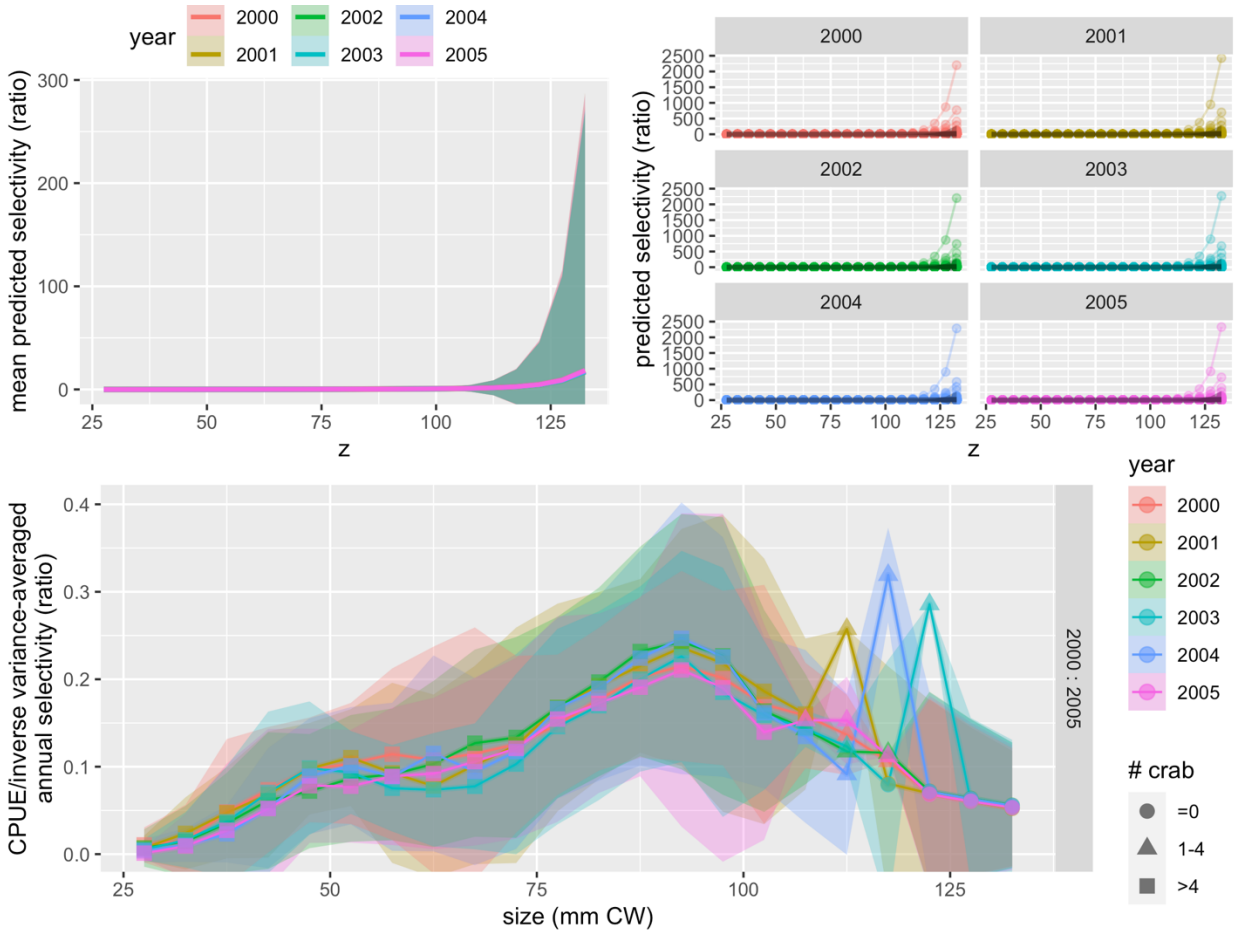

Figure 4.3-18. Predicted NMFS survey female catchability for 2000-2005. Upper row: unweighted mean and 95% confidence intervals (left); individual haul-level estimates (right). Lower row: Inverse-variance weighted estimates of annual NMFS survey catchability, with 95% confidence intervals.

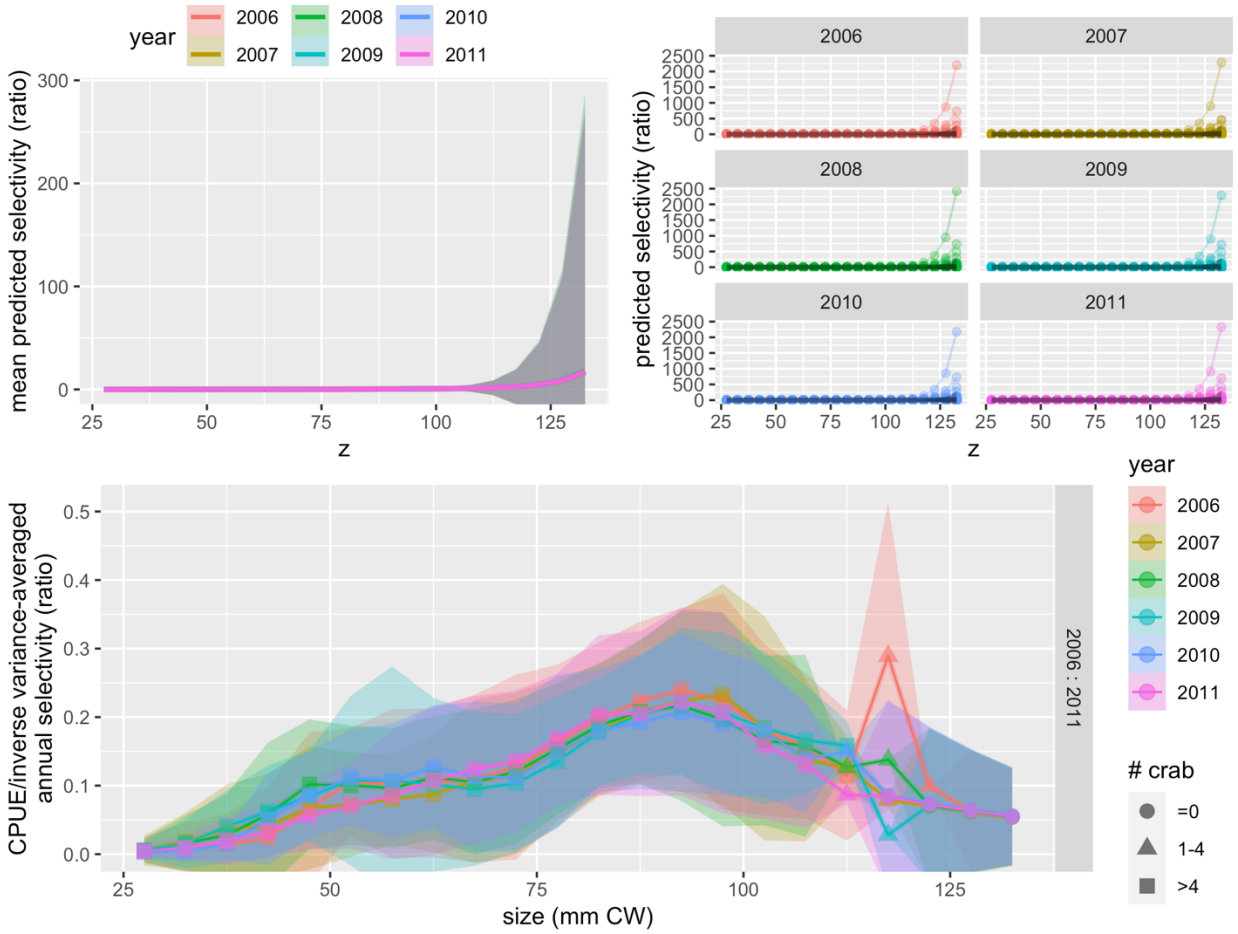

Figure 4.3-19. Predicted NMFS survey female catchability for 2006-2011. Upper row: unweighted mean and 95% confidence intervals (left); individual haul-level estimates (right). Lower row: Inverse-variance weighted estimates of annual NMFS survey catchability, with 95% confidence intervals.

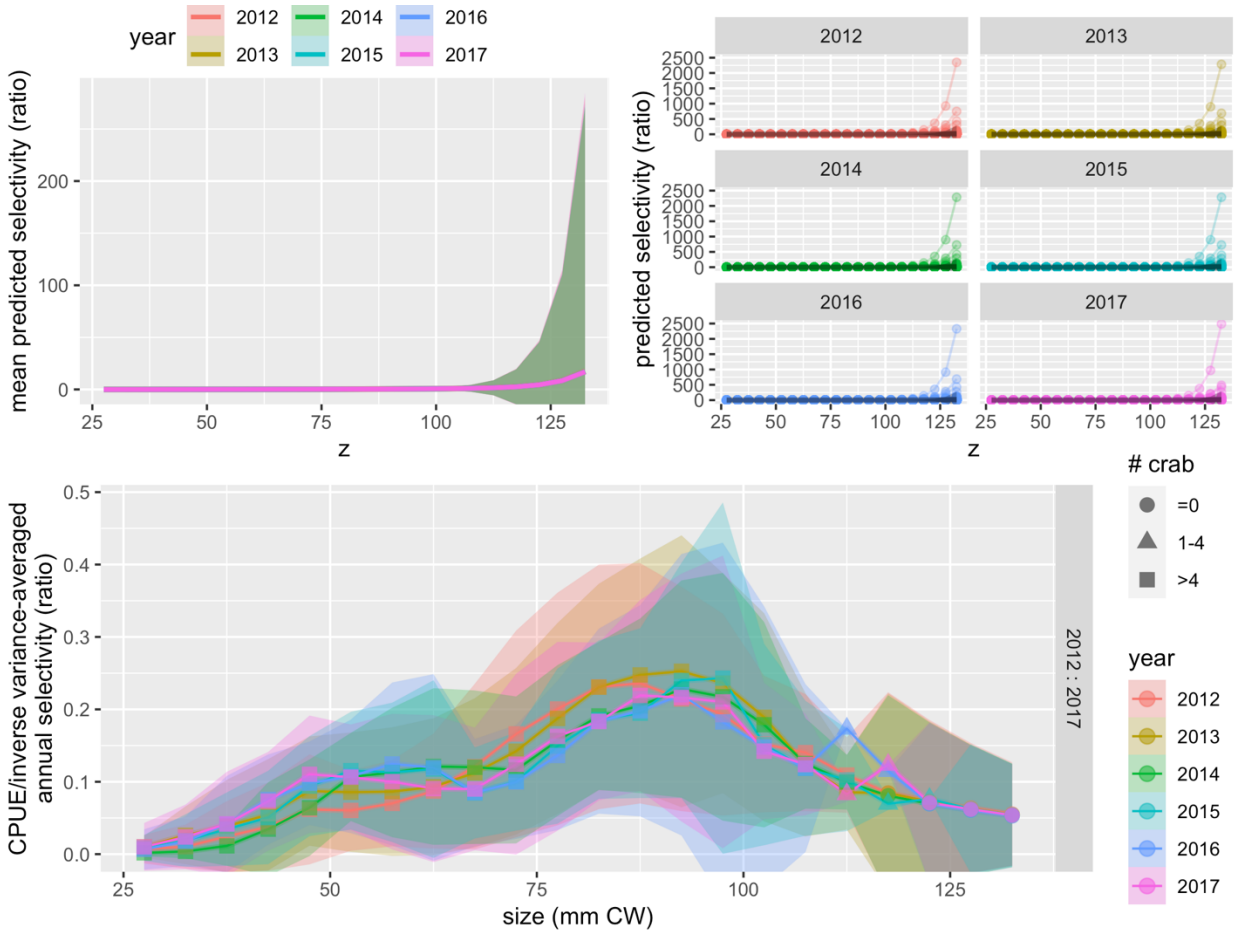

Figure 4.3-20. Predicted NMFS survey female catchability for 2012-2017. Upper row: unweighted mean and 95% confidence intervals (left); individual haul-level estimates (right). Lower row: Inverse-variance weighted estimates of annual NMFS survey catchability, with 95% confidence intervals.

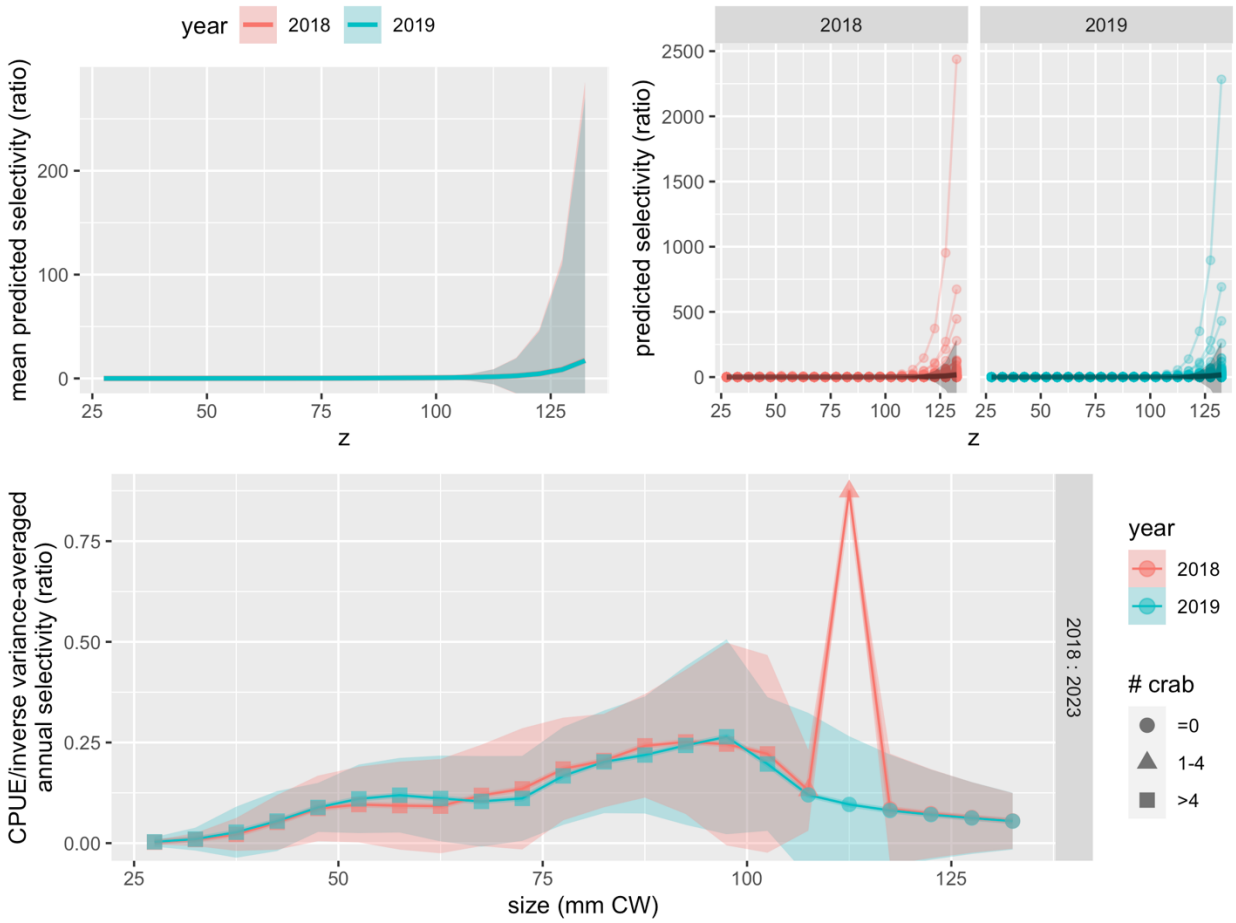

Figure 4.3-21. Predicted NMFS survey female catchability for 2018-2019. Upper row: unweighted mean and 95% confidence intervals (left); individual haul-level estimates (right). Lower row: Inverse-variance weighted estimates of annual NMFS survey catchability, with 95% confidence intervals.

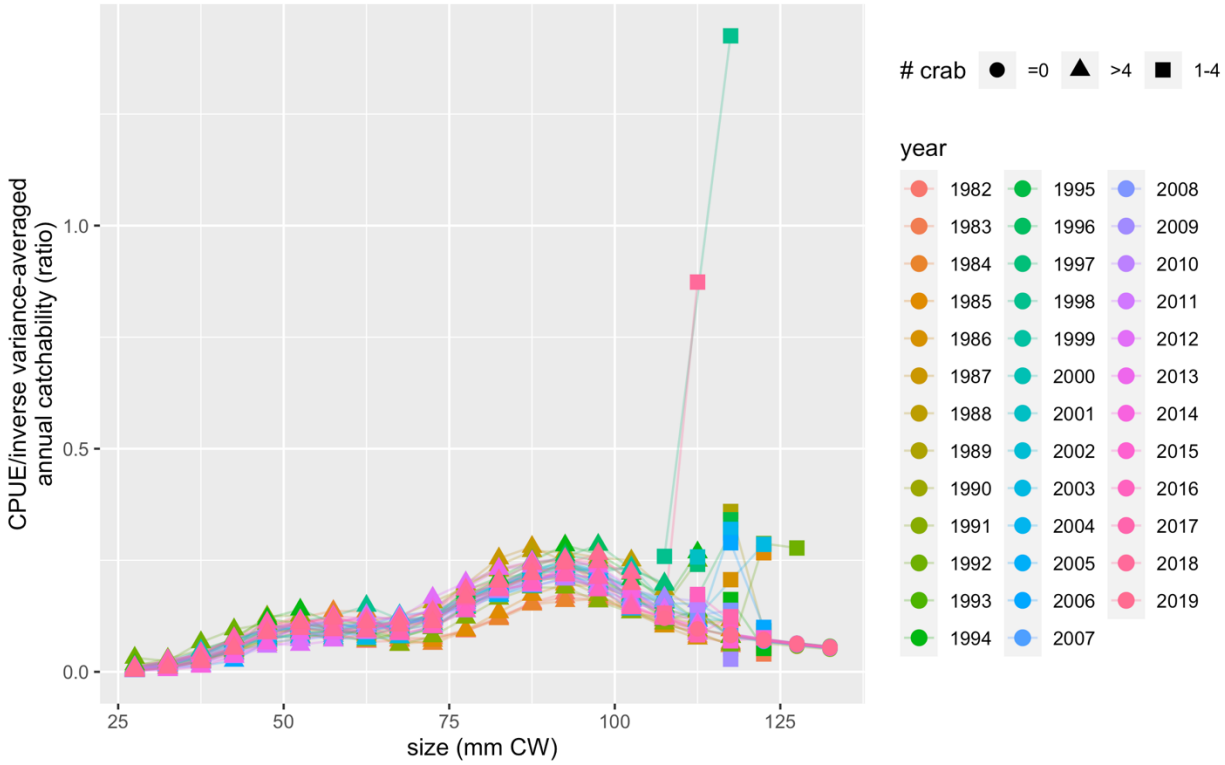

Figure 4.3-22. Predicted NMFS survey female catchability for all years.