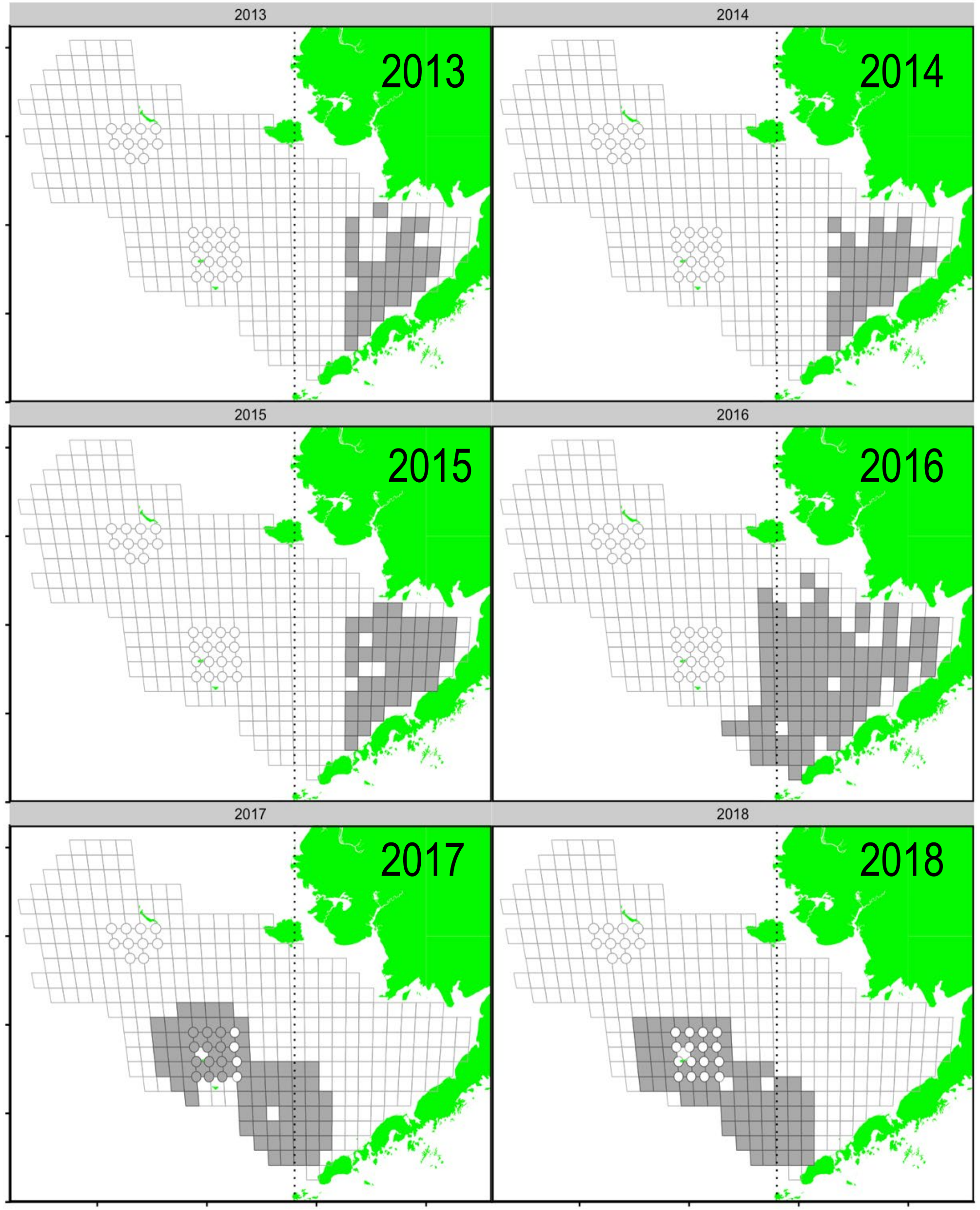
Estimated Selectivity Using Side-by-Side Haul Data

William Stockhausen

AFSC

January 13, 2026

BSFRF-NMFS "Side-By-Side" Selectivity Studies



The Math

- c: gear-specific CPUE (number caught/unit area)
- q: gear-specific catchability/selectivity
- d: crab density (number/unit area)
- z: size (sex, maturity state, shell condition, etc.)

$$c_z^N \equiv q_z^N \cdot d_z$$
$$c_z^B \equiv q_z^B \cdot d_z$$

$$\frac{c_z^N}{c_z^B} = \frac{q_z^N \cdot d_z}{q_z^B \cdot d_z} = \frac{q_z^N}{q_z^B} = R_z \equiv relative\ catchability$$

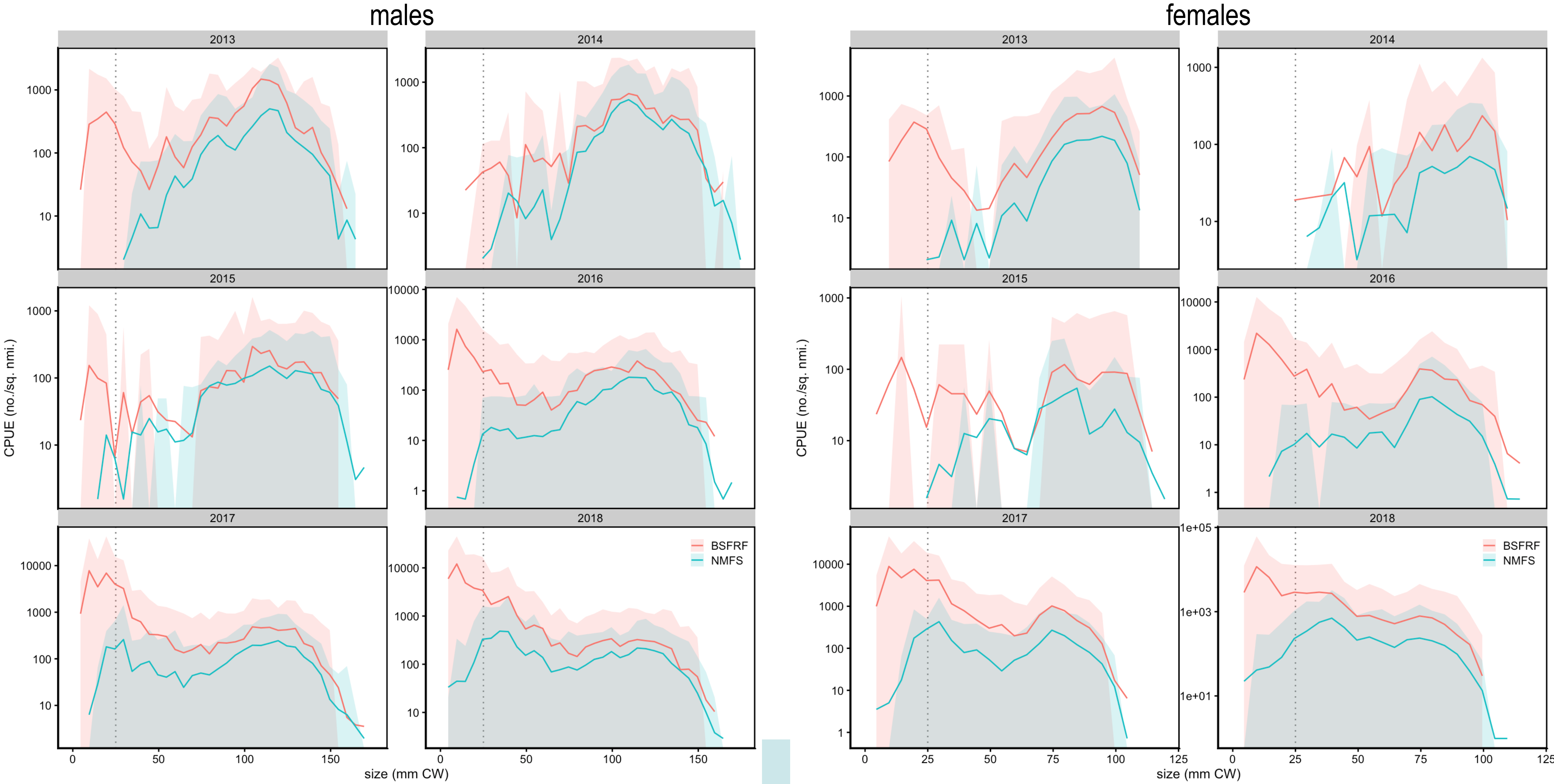
- ***BIG*** assumption 1: $q_z^B = 1$ $\Rightarrow$ $R_z = \frac{c_z^N}{c_z^B} = q_z^N \equiv absolute\ catchability$

- ***BIG*** assumption 2: d_z is (statistically) the same for both gears

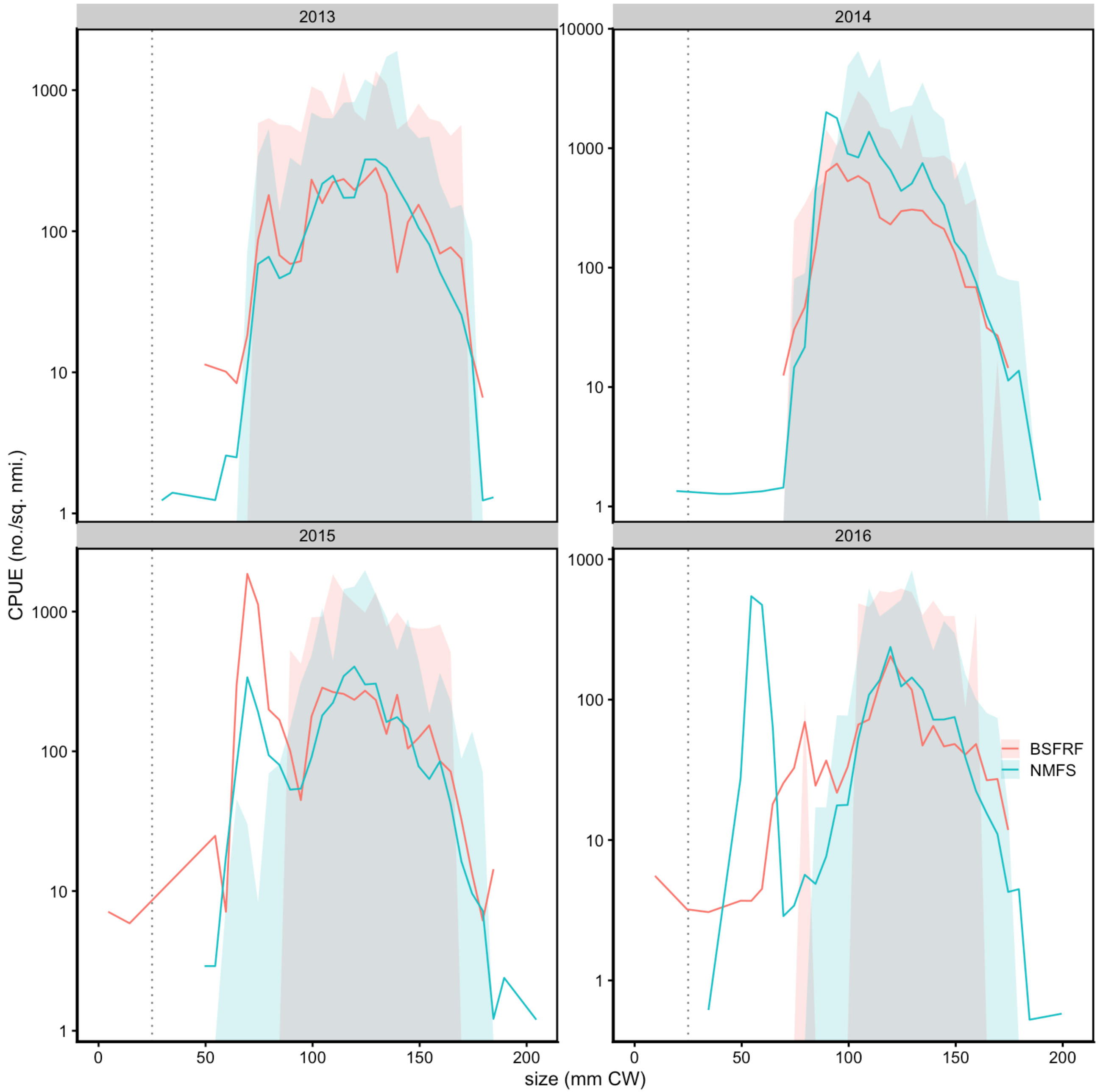
- Applications
  - haul level (location-specific) $c_z^{haul} = q_z^{haul} \cdot d_z^{station\ area}$
  - survey-level (area-aggregated) $c_z^{survey} = q_z^{survey} \cdot d_z^{survey\ area}$

Survey Level (Aggregated) Analysis

Area-expanded Bootstrapped Size Compositions: Tanner Crab



Area-expanded Bootstrapped Size Compositions: BBRKC



NOAA FISHERIES

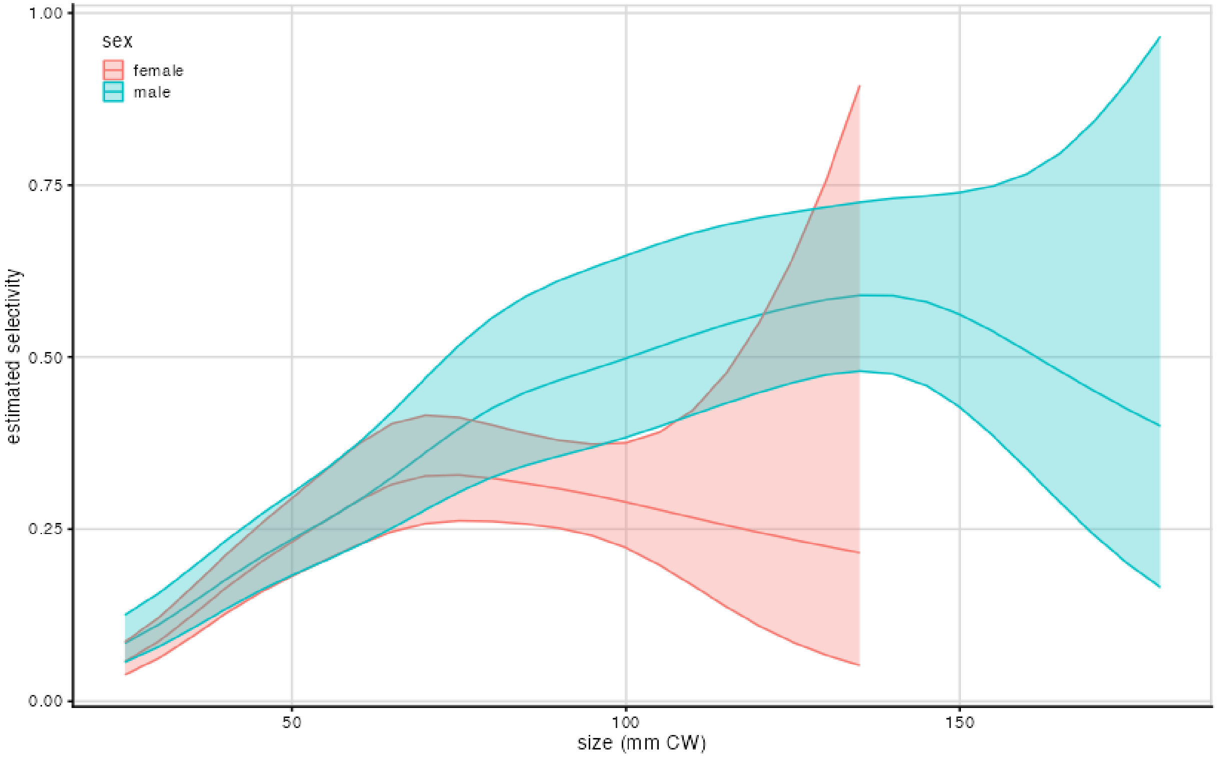
Estimated Catchability: Best Survey-Level (Area-aggregated) Models

- model 1: R~s(z,bs="tp",k=10)
- model 2: R~s(z,bs="tp",k=10) + ti(z,by=y,bs="sz") with year as fixed effect
- model 3: R~s(z,bs="tp",k=10) + s(z,y,bs="fs") with year as random effect

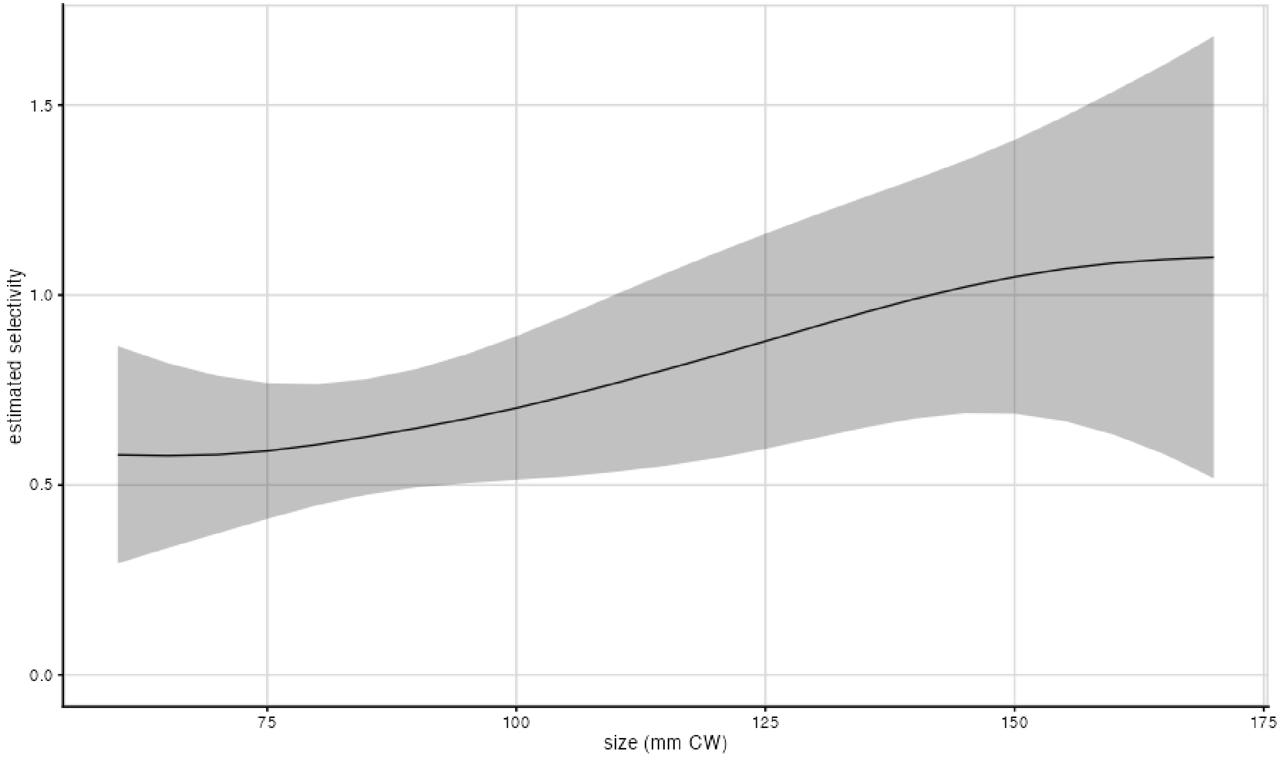
all models:
family: tweedie (log link)
weights: n^{B}_{z}/mean(n^{B}_{z})

- best models: model 3 in all cases

Tanner crab



BBRKC



Haul Level Analysis

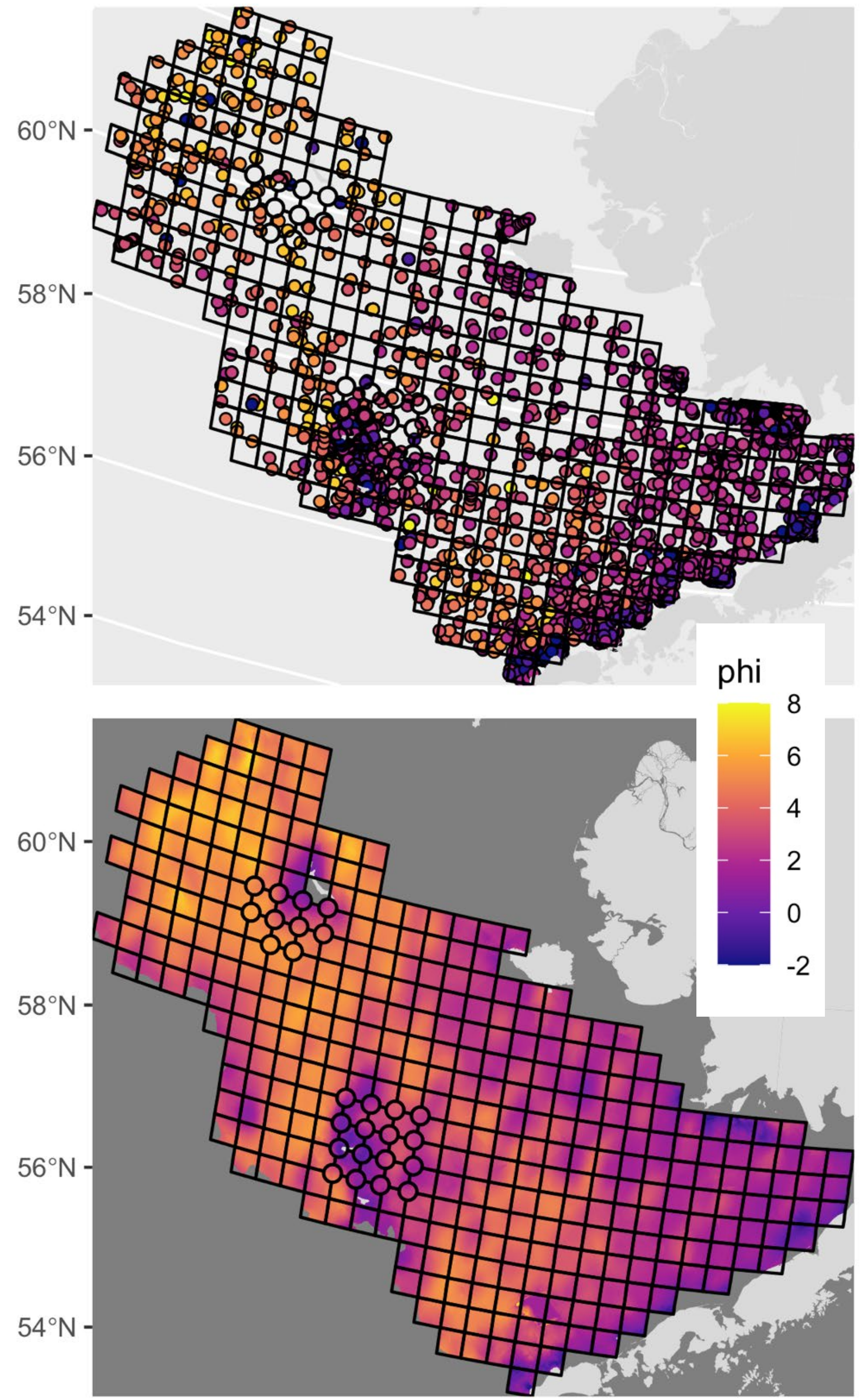
Potential environmental covariates

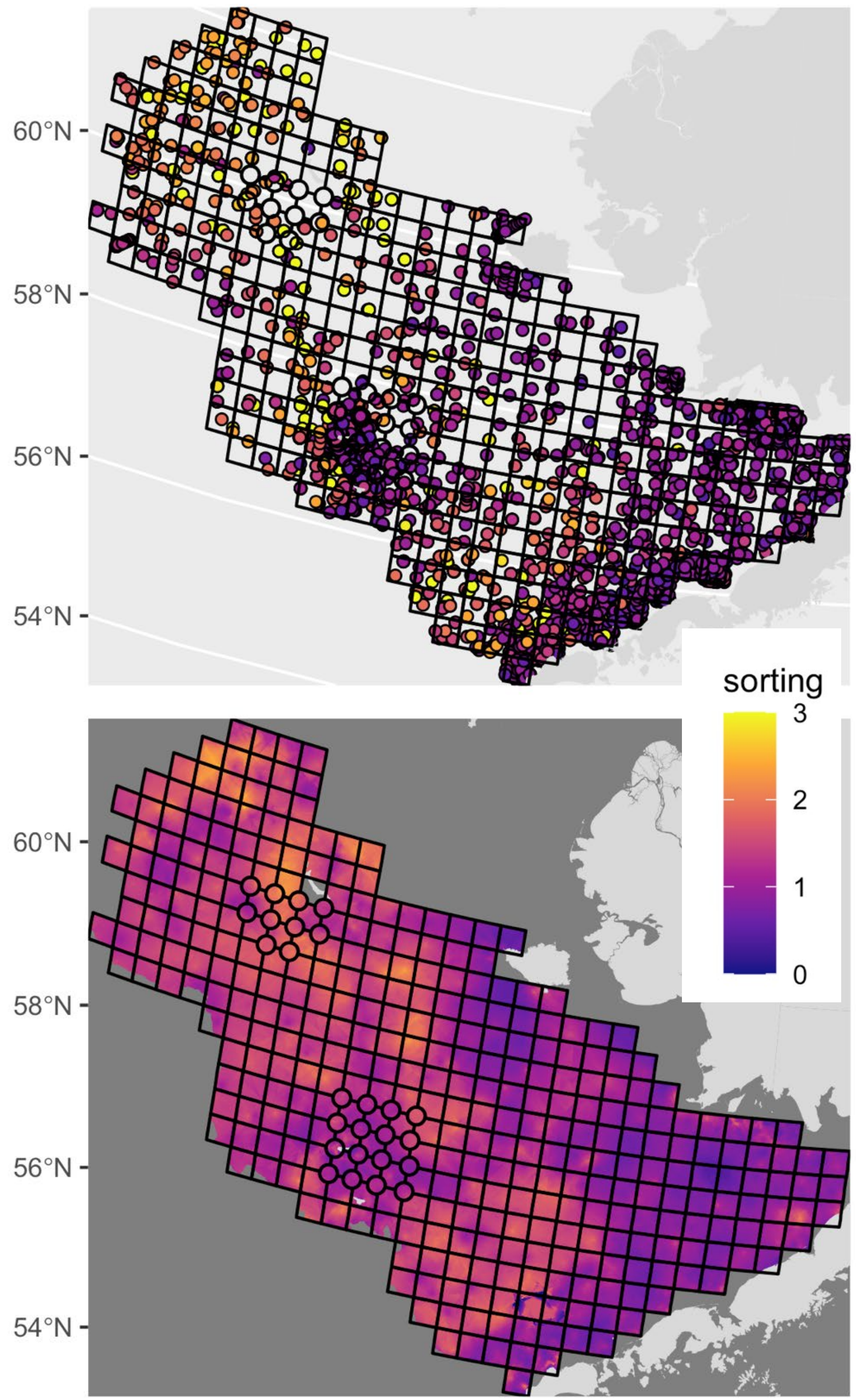
phi (f): measure of
ln-scale mean grain size

sorting (s): measure of
ln-scale grain size variance

depth (d)

bottom temperature (t)





Potential haul-specific environmental covariates

based on NMFS haul

bottom depth (d)

bottom temp. (t)

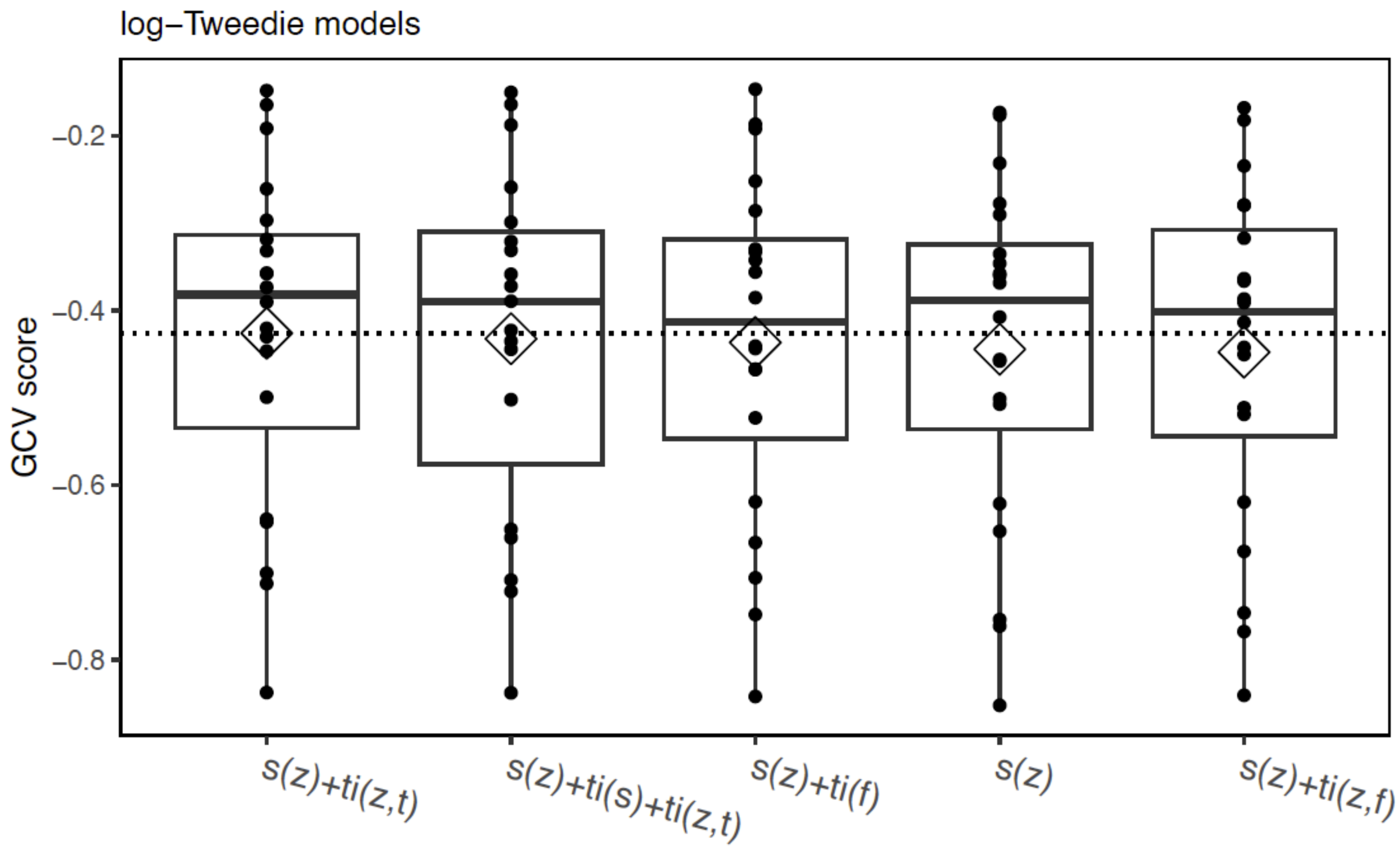
phi (f): mean ln-scale grain size

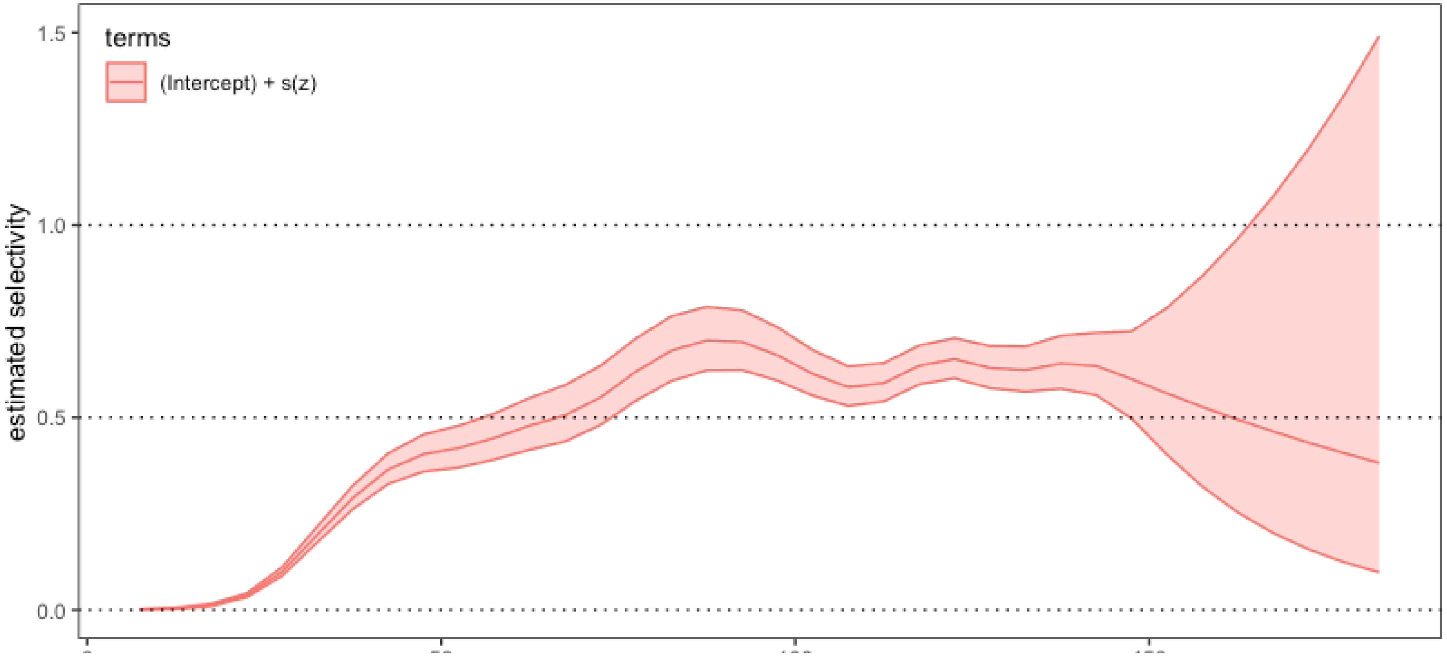
sorting coefficient (s): measure of ln-scale grain size variance

Tanner Crab

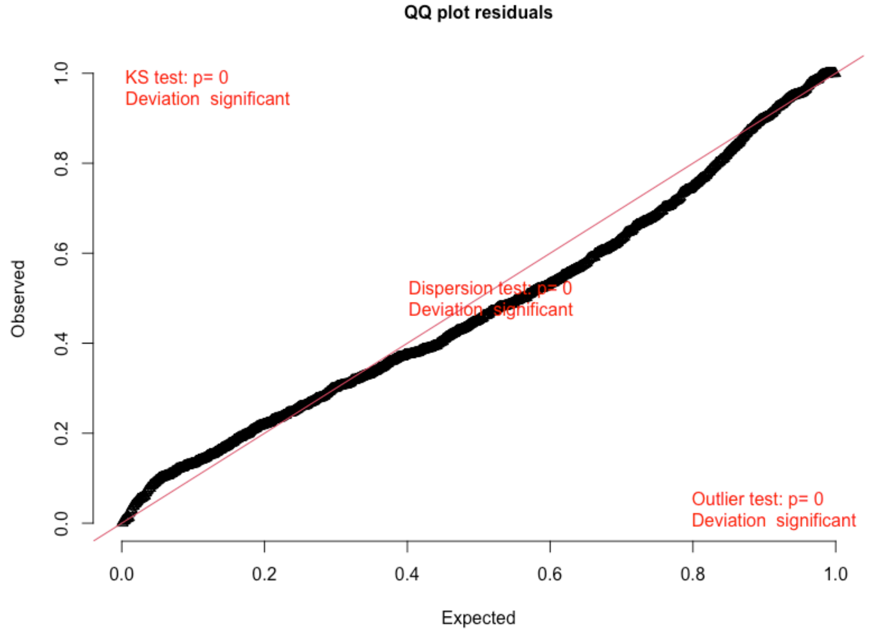
```
R ~ s(z,bs="ts",k=k1)   +
        ti(d,bs="ts",k=k2)   + ti(t,bs="ts",k=k2)   + ti(f,bs="ts",k=k2)   + ti(s,bs="ts",k=k2) +
        ti(z,d,bs="ts",k=c(k1,k2)) + ti(z,t,bs="ts",k=c(k1,k2)) + ti(z,f,bs="ts",k=c(k1,k2)) + ti(z,s,bs="ts",k=c(k1,k2))
with tw(link="log")
```

Model selection: GCV
- 20 folds
- Best GCV score

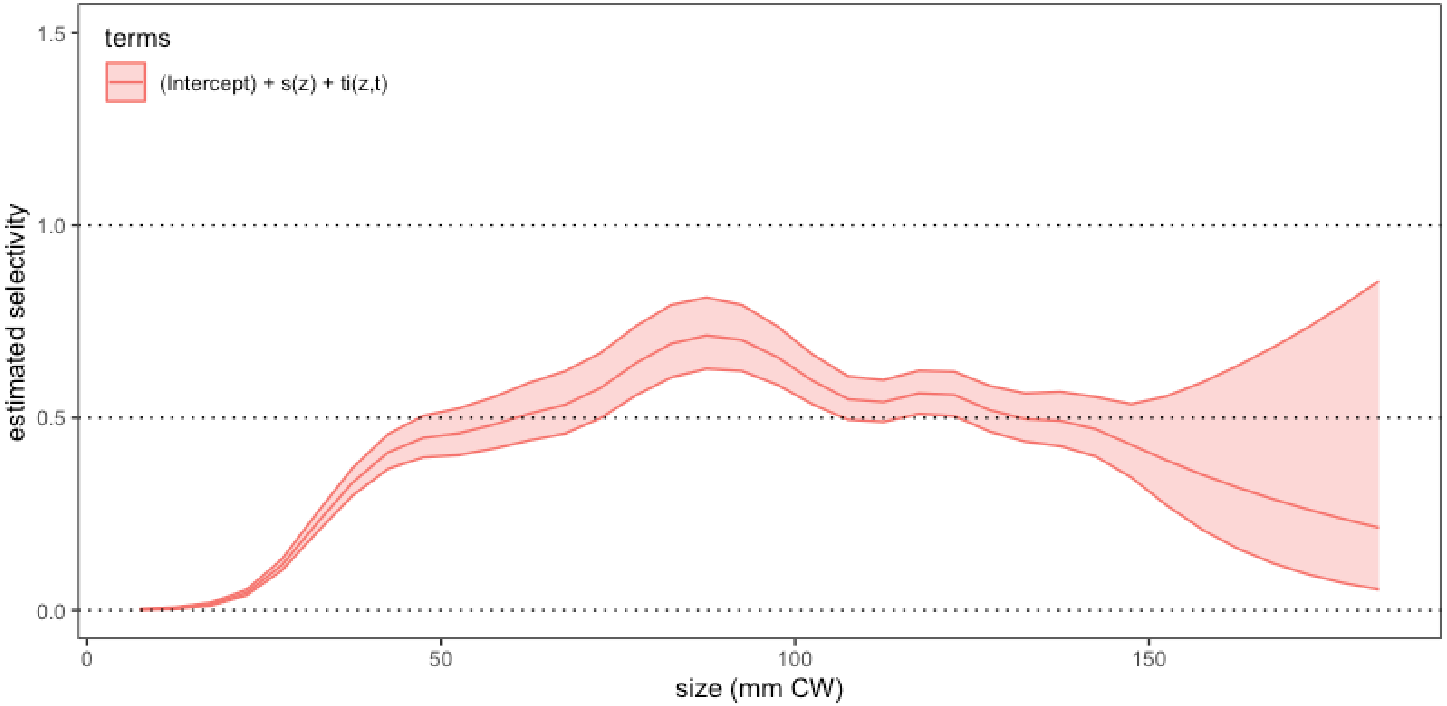




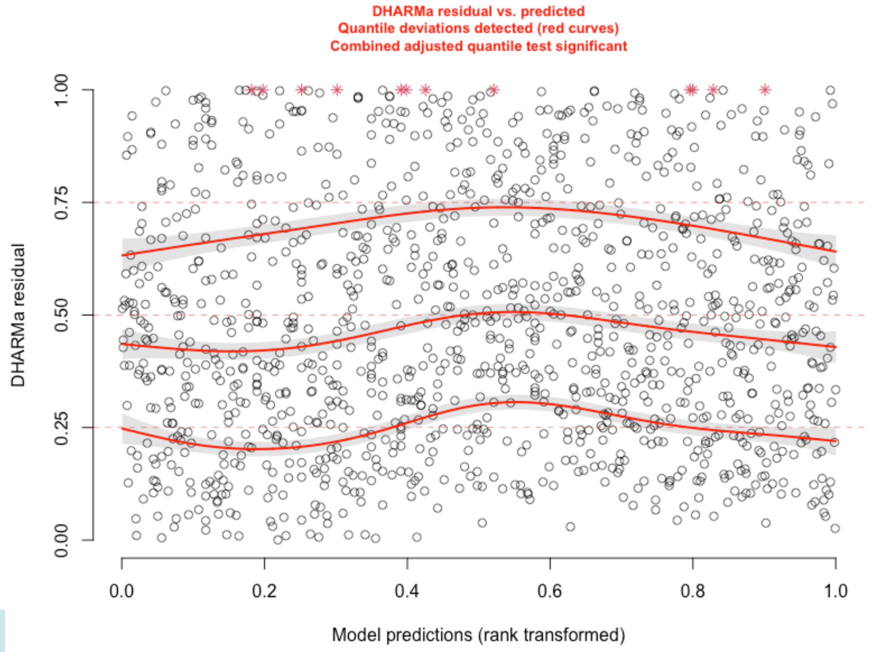
terms
(Intercept) + s(z)
estimated selectivity
1.5
1.0
0.5
0.0



QQ plot residuals
KS test: p= 0
Deviation significant
Dispersion test: p= 0
Deviation significant
Outlier test: p= 0
Deviation significant
Observed
Expected

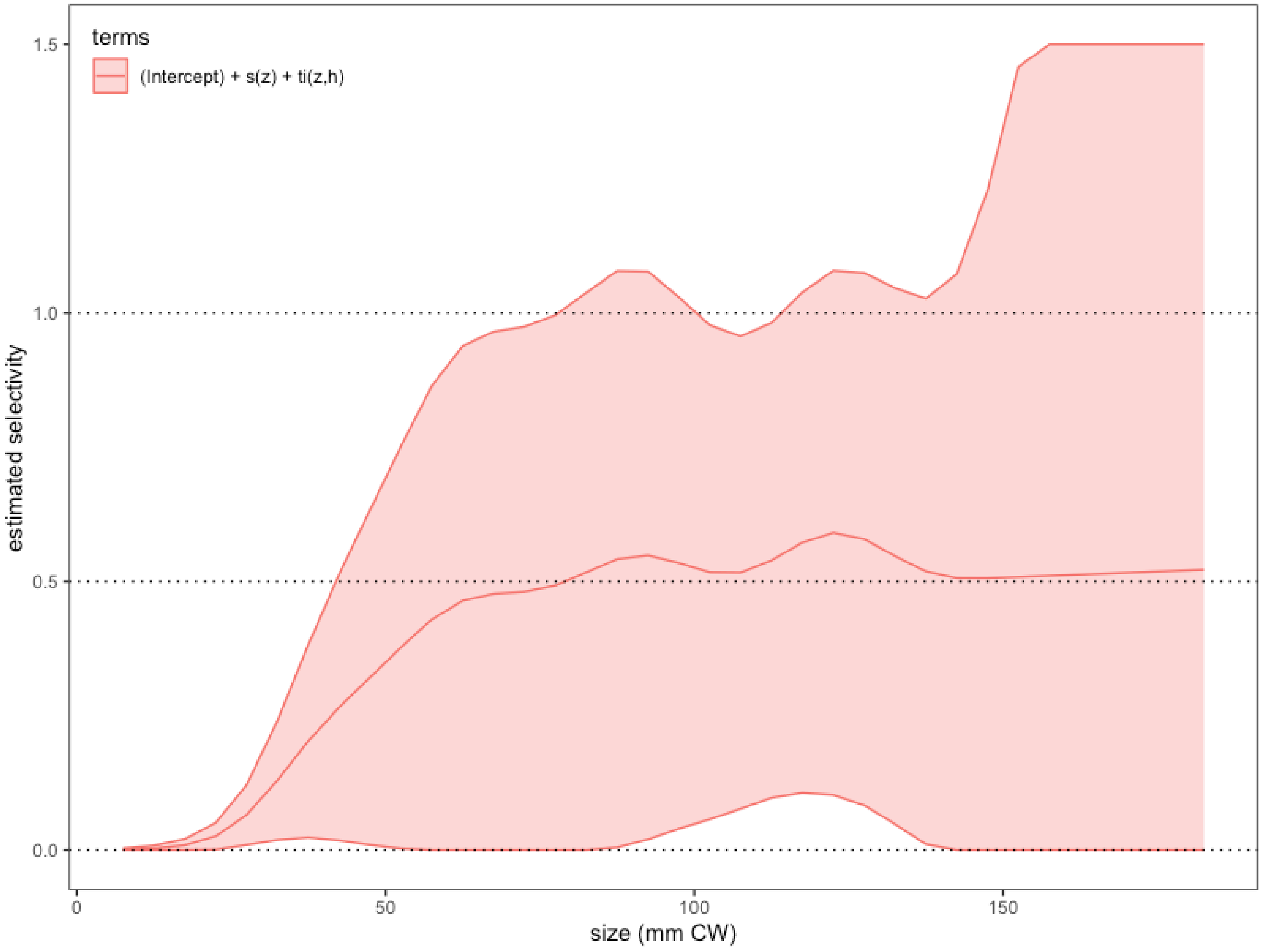


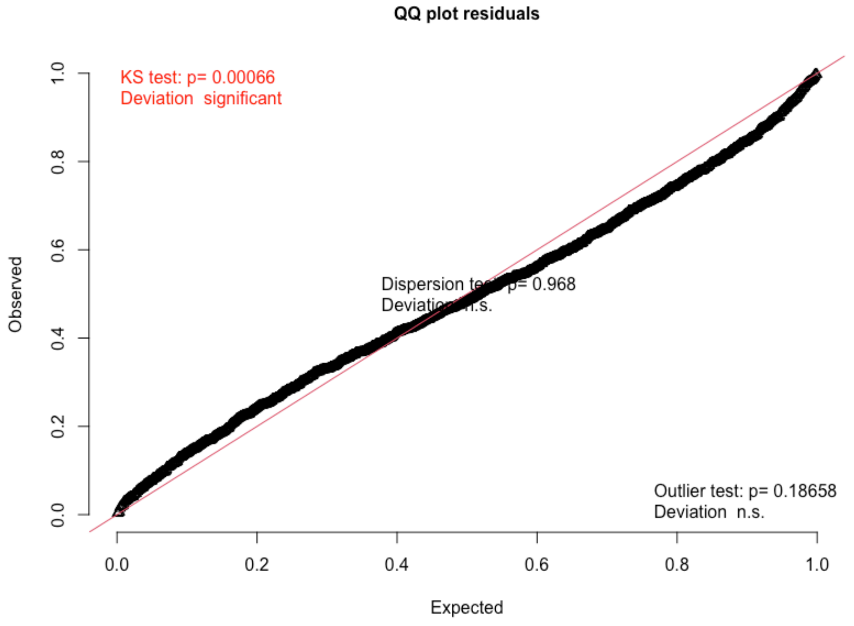
terms
(Intercept) + s(z) + ti(z,t)
estimated selectivity
size (mm CW)
0
50
100
150

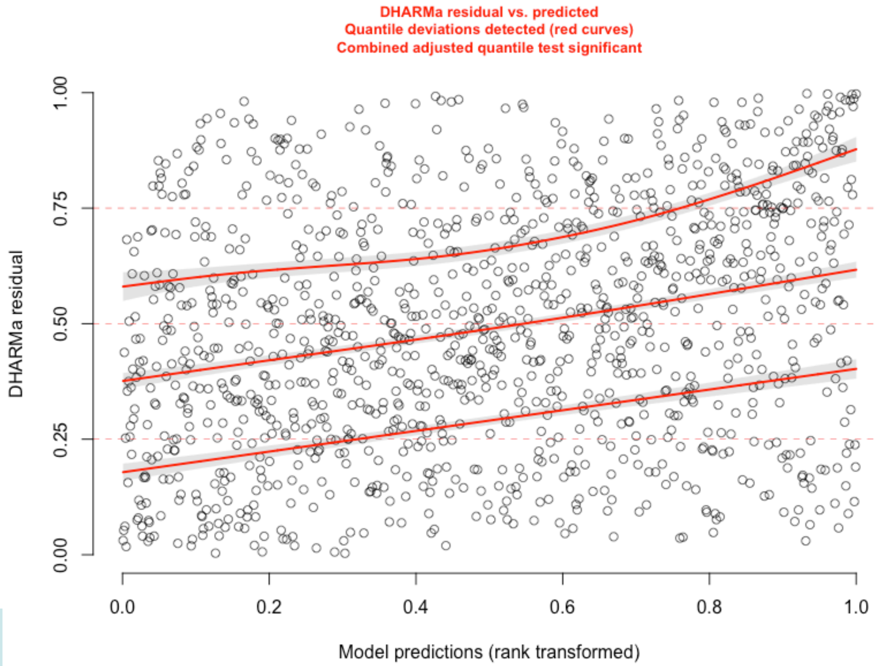


DHARMa residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
DHARMa residual
Model predictions (rank transformed)

R ~ s(z,bs="ts",k=k1) + ti(z,h,bs="fs",k=k1)
with tw(link="log");

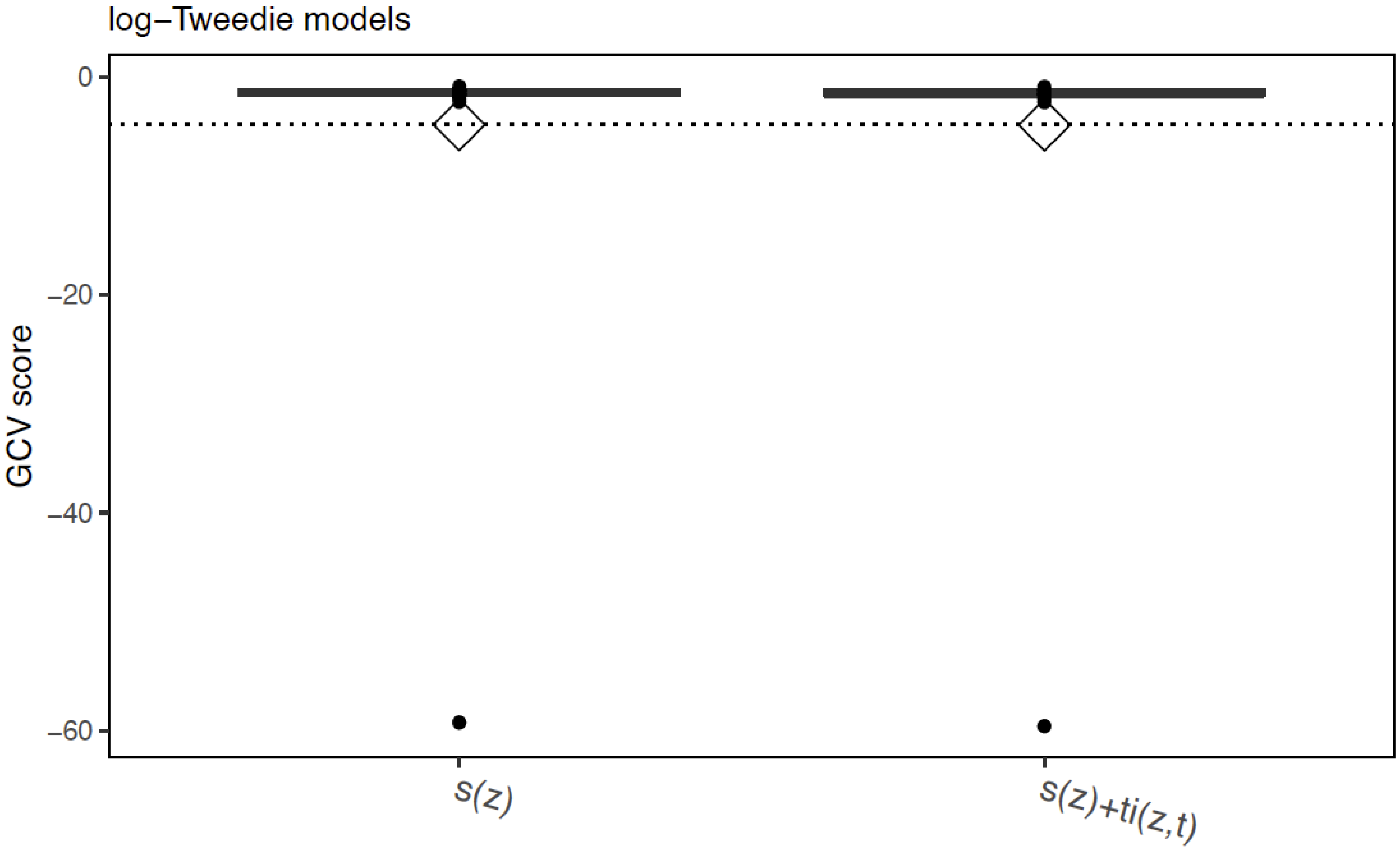


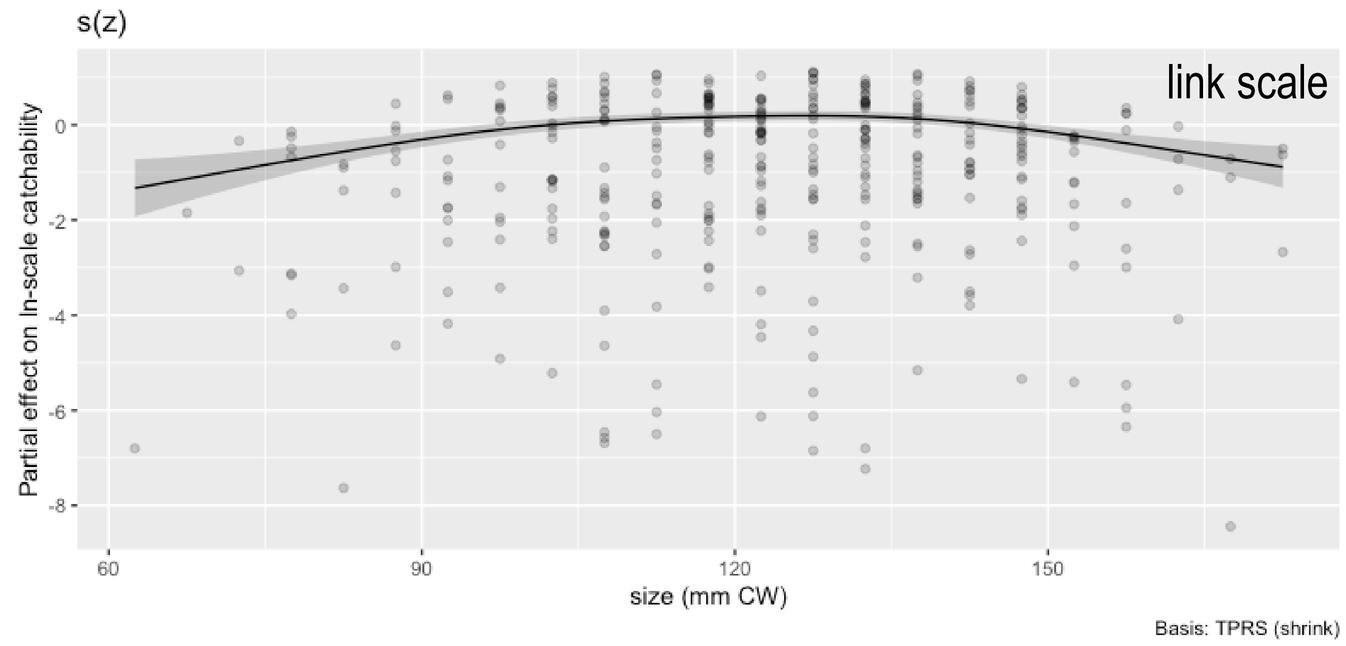


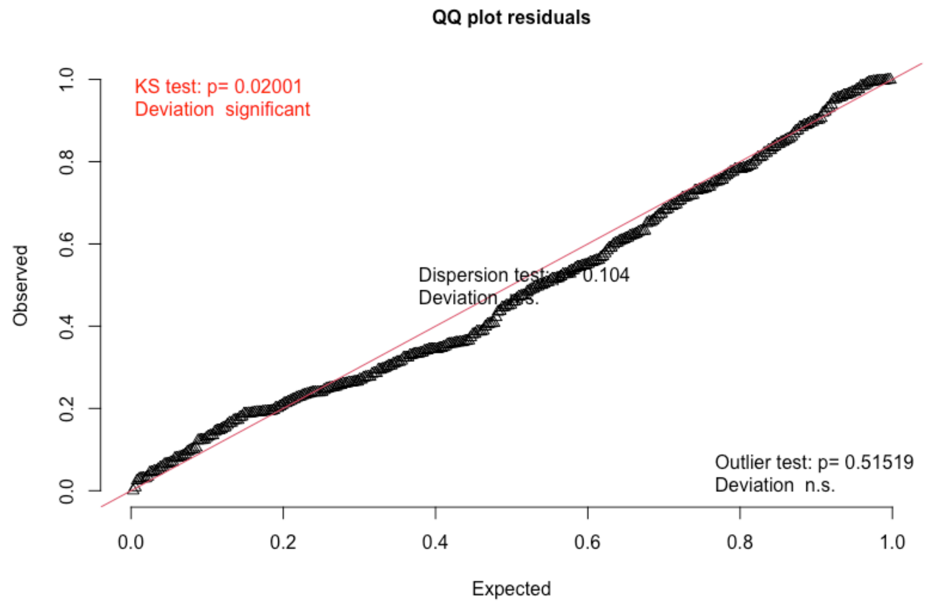


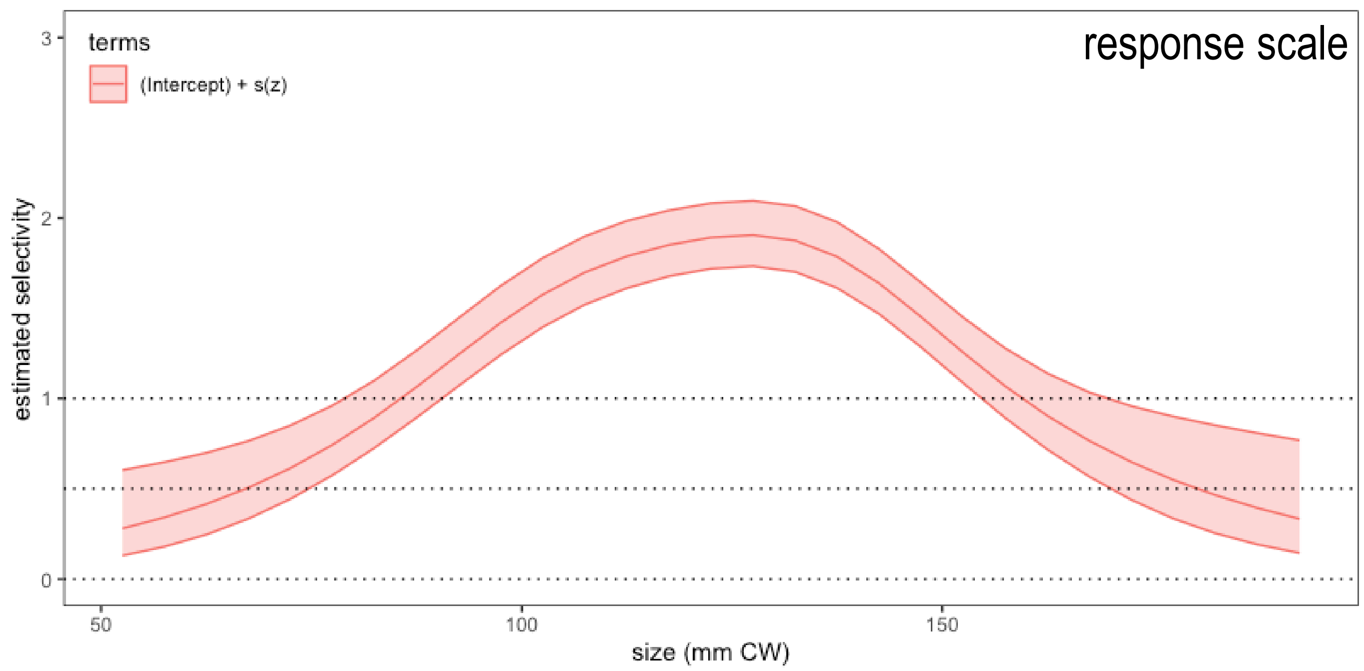
BBRKC

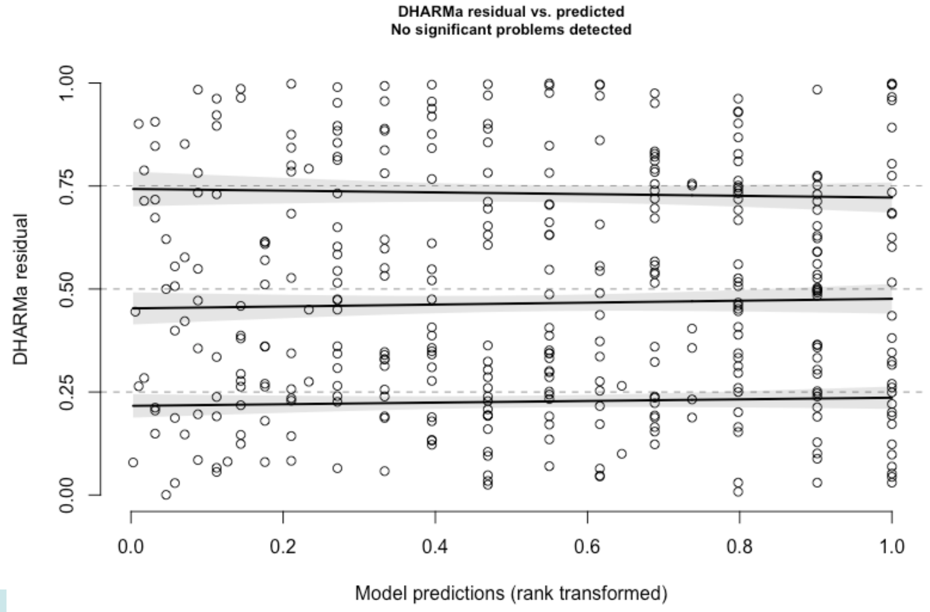
```
obsR~s(z,bs="ts",k=k1) +
        ti(d,bs="ts",k=k2) + ti(t,bs="ts",k=k2) + ti(f,bs="ts",k=k2) + ti(s,bs="ts",k=k2) +
        ti(z,d,bs="ts",k=c(k1,k2)) + ti(z,t,bs="ts",k=c(k1,k2)) + ti(z,f,bs="ts",k=c(k1,k2)) + ti(z,s,bs="ts",k=c(k1,k2))
with tw(link="log");
```



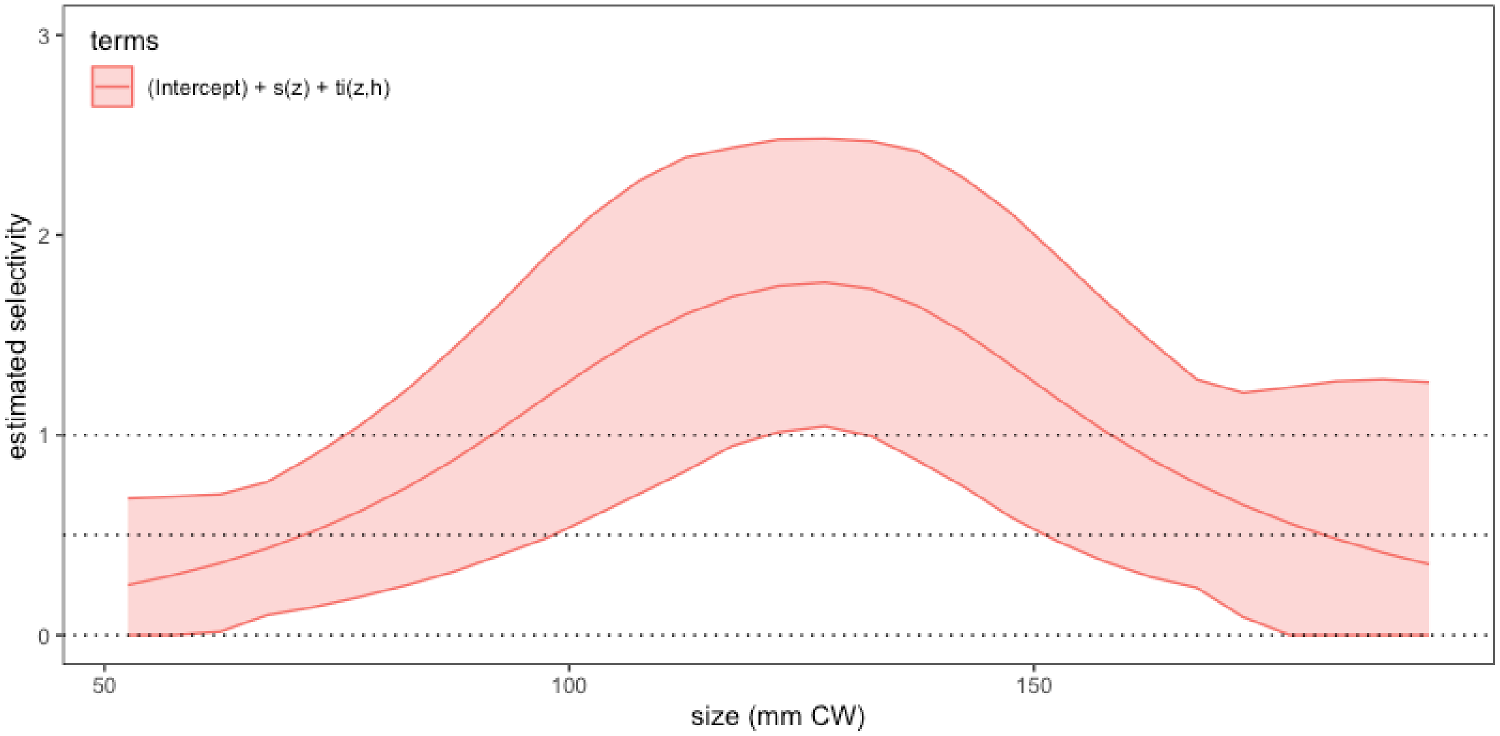

s(z)
link scale
Partial effect on ln-scale catchability
size (mm CW)
Basis: TPRS (shrink)

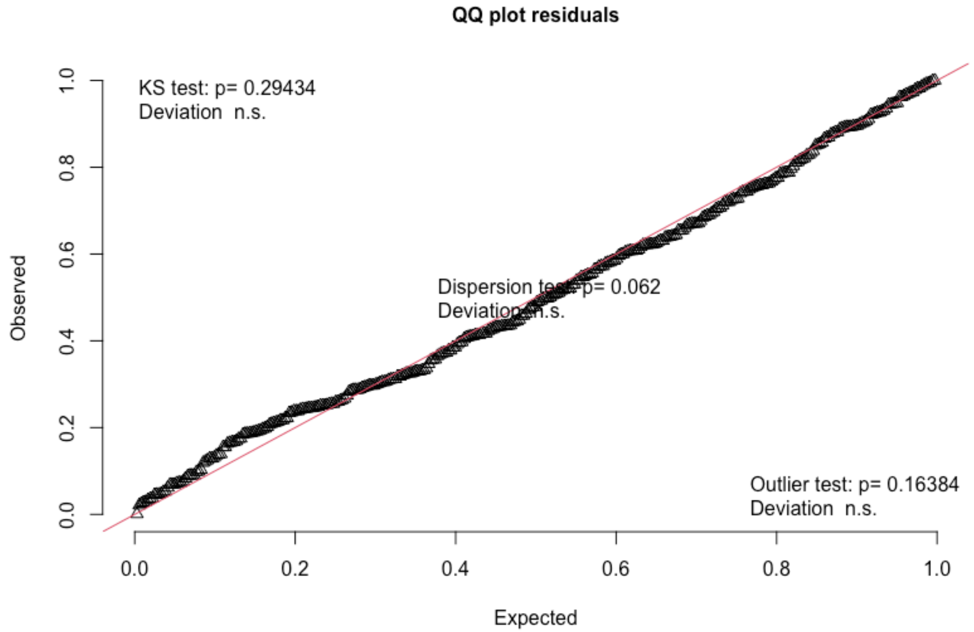

QQ plot residuals
KS test: p= 0.02001
Deviation significant
Dispersion test: p= 0.104
Deviation n.s.
Outlier test: p= 0.51519
Deviation n.s.
Observed
Expected

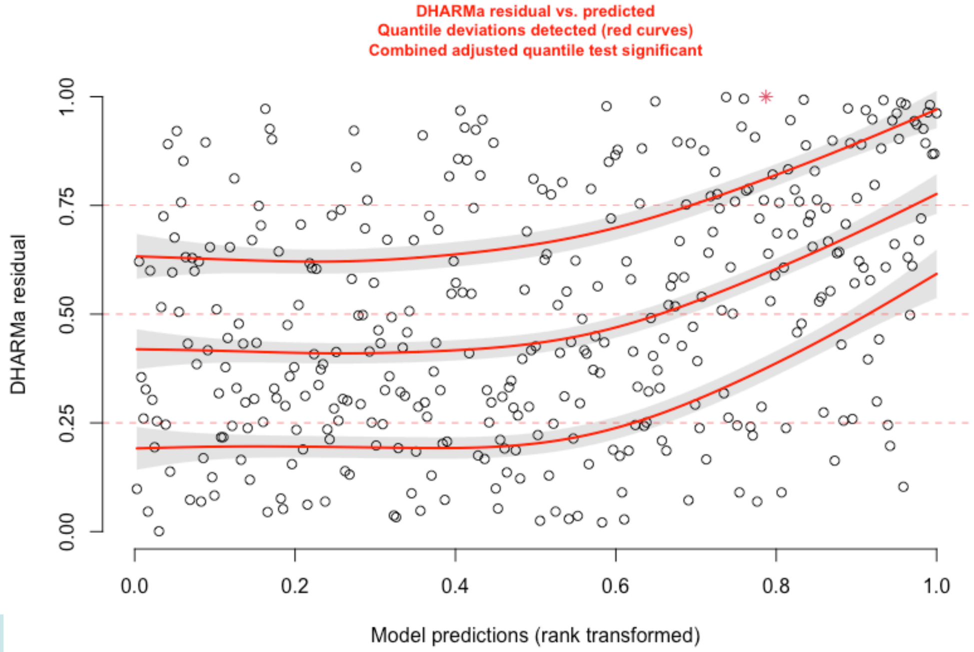

terms
(Intercept) + s(z)
response scale
estimated selectivity
size (mm CW)


DHARMa residual vs. predicted
No significant problems detected
DHARMa residual
Model predictions (rank transformed)

obsR~s(z,bs="ts",k=k1) + ti(z,h,bs="fs",k=k1)
with tw(link="log")







Thygesen et al Approach

Thygesen, U. H., Kristensen, K., Jansen, T., and Beyer, J. E. Intercalibration of survey methods using paired fishing operations and log-Gaussian Cox processes. – ICES Journal of Marine Science, 76: 1189–1199.

Ratio *R* modeled as a Log-Gaussian Cox process

- $N_{ijk} | \Phi, R, S \sim \text{Poisson}(A_{ij} \cdot \exp(S_{jk} + \Phi_{ik} + R_{ijk}))$ numbers at station *i* by gear *j* in size bin *k*
- Φ_{ik} characterizes the ***In-scale*** size distribution at station *i*, size bin *k*
- R_{ijk} characterizes ***In-scale*** haul/gear-specific variations in size distribution at station *i*
- S_{jk} characterizes the gear-specific ***In-scale*** selectivity for size bin *k*
- A_{ij} is the area swept by gear *j* at station *i* (known)

- No environmental covariates included in this formalism
- BSFRF: gear = 2. BSFRF selectivity = 1 => S_{2k} = 0 (ln-scale!)
- in absence of R_{ijk},

$$\exp(S_{1k}) = \frac{\mathbb{E}\{N_{i1k}/A_{i1} | \Phi, S\}}{\mathbb{E}\{N_{i2k}/A_{i1} | \Phi\}}$$

Statistical Modeling

- Statistical assumptions

$S_{1(k+1)} - S_{1k} \sim N(0, \sigma_S^2) \quad \text{for} \quad k = 1, \ldots, n_1 - 1$ In-scale change is RW in size

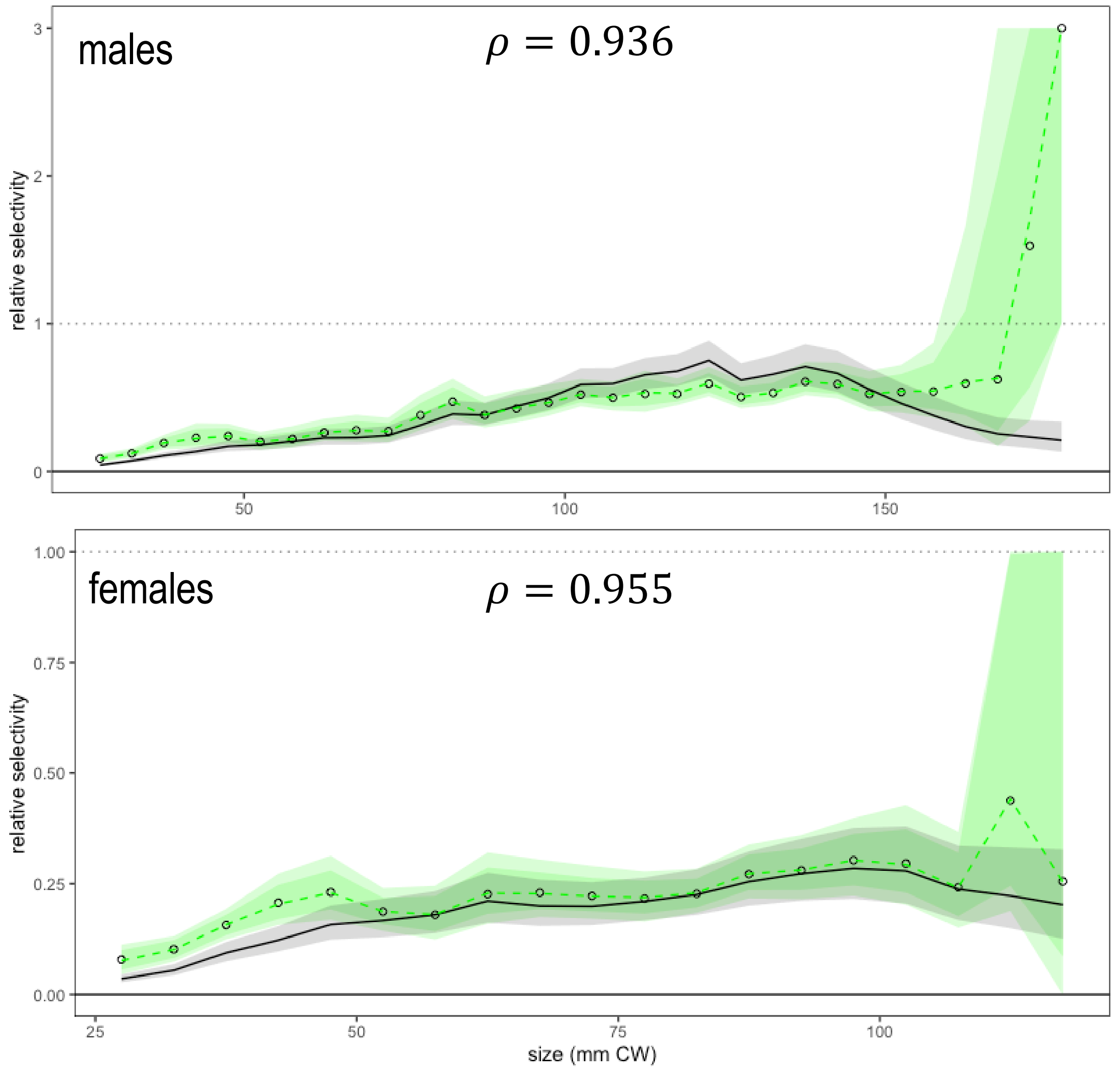
$\Phi_{i(k+1)} - \Phi_{ik} \sim N(0, \sigma_\Phi^2) \quad \text{for} \quad k = 1, \ldots, n_1 - 1,$ In-scale change is RW in size, independent between stations

$R_{ijk} = R_{ijk}^{WN} + R_{ijk}^{AR} \quad R_{ijk}^{WN} \sim N(0, \sigma_{WN}^2) \quad R_{ijk}^{AR} \sim N(0, \sigma_{AR}^2) \quad \mathbb{E}(R_{ijk}^{AR} R_{ijk'}^{AR}) = \sigma_{AR}^2 \rho^{|k-k'|}$ In-scale changes are AR1 in size independent WN between stations, gear

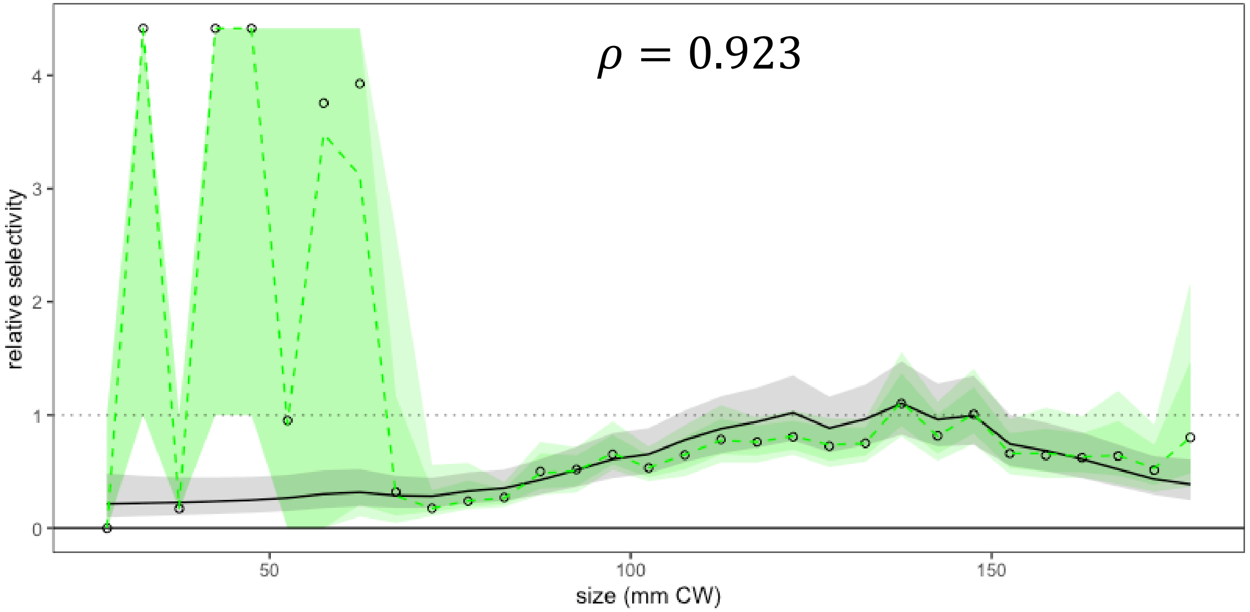
- model for joint distribution of catch numbers
- implemented in TMB
- estimates variance-covariance parameters: $\sigma_S^2, \sigma^2, \sigma^2, \sigma^2, \rho$
- many random effects associated with: Φ, R, S,
- likelihood maximized to estimate var-covar parameters
- posterior modes for S and other random effects R, S,

- green: bootstrapping estimate and confidence intervals
- black/grey: log-Gaussian Cox Type equation here.model estimate

Tanner crab



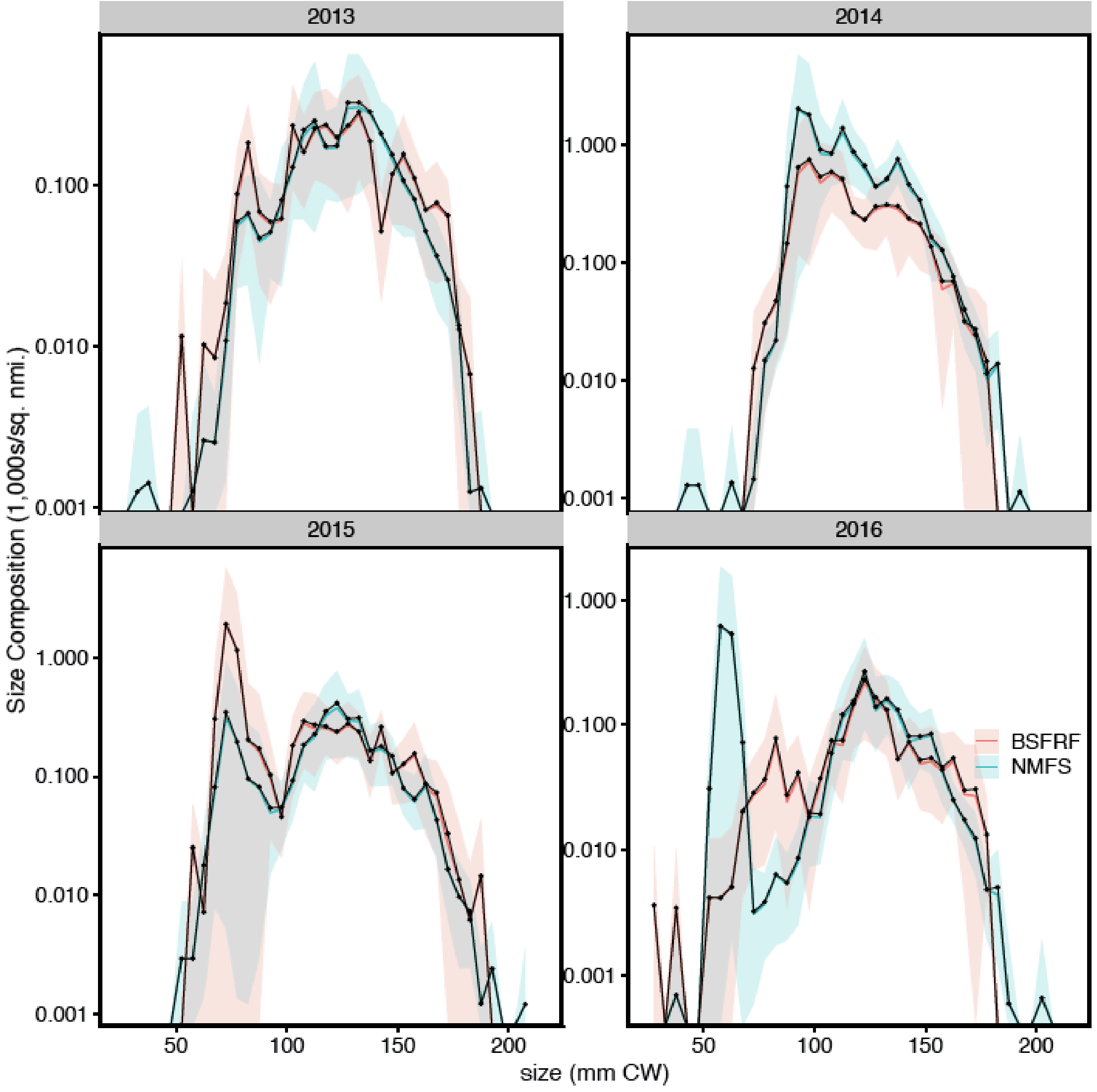
BBRKC



Thoughts?

- similar analyses for snow crab TBD

Area-expanded, Bootstrapped Size Compositions



Bootstrapped Empirical Selectivity Estimates