Appendix A: AIGKC Cooperative Pot Suvrey

Tyler Jackson, Ben Daly, Jared Weems, Andrew Nault
Alaska Department of Fish and Game, tyler.jackson@alaska.gov

September 2024

Background

The AIGKC cooperative pot survey was initiated in 2015 in the EAG and has continued every year since with the exception of 2020. The survey was extended to the WAG in 2018. The main purpose of the survey is to generate a cost effective data stream available to the stock assessment that is spatially representative and less susceptible to hyperstability than fishery catch-per-unit-effort (CPUE). The survey has occurred during the beginning of each season, with participating vessels setting pots strings at pre-determined stations and later hauling strings with ADF&G staff on board for collection of biological data. Survey data is available for 2015-2023 in the EAG and 2018 and 2019 in the WAG. Since only two years of data are available for the WAG, this appendix only details the preparation of assessment model inputs for the EAG.

Methods

Station Selection and Sampling

The survey area in the EAG consists of a 2x2 nmi grid overlaying the full range of historically fished areas. The survey grid is divided into strata to improve spatial representation. Strata have varied over the time series depending on the vessels participating, but since 2022 the survey design has included three strata (Figure 1). Vessels divided sampling efforts based on their respective fishing grounds. Survey data pre-2022 were stratified to the current design *post hoc*.

Grid cells were randomly selected among strata, and selections were removed or reallocated as necessary to avoid gear conflicts with groundfish trawl operations. The number of selected cells has varied by year. Since 2022 the initial selection process in the EAG included 25 grid cells per stratum. Vessels set pot strings on longlines intersecting grid cells at their digression and there are no specific requirements on string length or soak time (Figure 2). Biologists systematically sampled 5-7 pots per string, not beginning or ending on the first or last pot of a string. Due to logistical constraints, not all survey strings were sampled during biologist crewed trips.

Carapace lengths (CL) were recorded for legal and sublegal males, as well as females. Unmeasured crab were counted to obtain a catch number per size-sex catagory for each pot. On deck sampling procedure varied among vessels and crew as working conditions allowed, though biologists targeted CL records from at least ~ 30 crab per pot, with at least 10 of which being legal males. Females were given lowest priority and subsampled at a high rate. Sampling statistics from 2015 - 2023 are detailed in Table 1.

Survey CPUE

Nominal survey CPUE of all male golden king crab (U_y) was computed as

$$U_y = \frac{1}{k} \sum_{h=1}^{k} U_h \tag{1}$$

$$U_h = \frac{1}{n_h} \sum_{i=1}^{n} U_{i,h} \tag{2}$$

$$U_i, h = \frac{1}{m_i} \sum_{j=1}^{m} U_{j,i} \tag{3}$$

where

U_h = CPUE of stratum h of k;
$U_{i,h}$ = CPUE of string i of n, within stratum h; and
$U_{j,i}$ = CPUE of pot j of m, within string i.

Variance in U_y ($\sigma^2_{U_y}$) was estimated as

$$\sigma^2_{U_y} = \frac{1}{k^2} \sum_{h=1}^{k} \sigma^2_{U_h} \tag{4}$$

$$\sigma^2_{U_h} = \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2_{U_i} \tag{5}$$

Model based CPUE was estimated by a general additive mixed model (GAMM) in the form of

$$\ln(U_j) = \text{Year}_y + s(\text{soak time}) + s(\text{depth}) + (1|\text{Stratum}_h/\text{String}_{i,y}) + \epsilon \tag{6}$$

in which ϵ is negative binomial distributed error with θ estimated via maximum likelihood. GAMMs were estimated using the R library *gamm4* (Wood and Scheipl 2020). The main difference between this and previous analyses was that zero catches were included in the data set. Data on zero-catch pots were not previously available to previous analyses (Jackson 2024).

Simulated residuals were calculated using the R package *DHARMa* (Hartig 2020). *DHARMa* simulates a cumulative density function for each observation of the response variable for the fitted model and computes the residual as the value of the empirical density function at the value of the observed data. Residuals are standardized from 0 to 1 and distributed uniformly if the model is correctly specified.

Residual patterns (below) in GAMM fit were addressed by removing anomalous soak times (> 25 days) which were from 15 strings aboard a single vessel in 2022 (Figure 7). Soak times were longer than expected due to mechanical issues than caused vessel delays. The survey does not have formal bounds on acceptable soak times. After removing long soak times, estimated degrees of freedom for smoothed terms suggested a linear fit to depth and soak time. The model was refit using the R library *glmmTMB* (Brooks et al. 2017) to take advantage of Tweedie distribution implementation supported by the package *tweedie* (Dunn 2022).

The annual CPUE index was extracted from model results following the same method as used for observer CPUE standardization (Jackson 2024, Appendix A).

Results

The GAMM fit to the complete data set suggested a dome shaped fit to soak time (EDF = 3.72) with a slow positive increase followed by a sharp decline at the largest values (i.e., roughly at 25 days) (Figure 3). The marginal effect of depth was also dome shaped (EDF = 2.58), though nearly asymptotic (Figure 4). DHARMa residuals indicated heteroscedasticity and further analysis suggested it was due to soak time (Figure 5 and 6).

Removing anomalous soak times (> 25 days) resolved the heteroskedasticity in the negative binomial GAMM, but suggested an excess of zeros for the negative binomial distribution. The updated model also reduced smooth terms to linear fits. DHARMa residuals indicated the Tweedie GLMM ($p = 1.3$) with linear depth and soak time was a better fit to the data (Figure 8). Marginal effects of soak time and depth are in Figure 9 and 10, respectively. The resulting index follows a similar trend to the nominal CPUE (Table 2, Figure 11). The survey index is similar to the observer index, though with more extreme increase in 2018 and decrease in 2021 (Figure 11). Refitting and rescaling the observer standardization index with data from only 2015 - 2023 made little difference to the trend (Figure 11).

Survey size composition showed little variability in the prominent mode, with most individuals measured between 125 - 175 mm CL. Sublegal males were more abundant in 2015 - 2019 surveys (Figure 13).

Tables

Table 1: Number of strings and pots sampled, and total number of male crab caught and measured by legal status.

Survey Year	Strings Sampled	Pots Sampled	Legal Caught	Legal Measured	Sublegal Caught	Sublegal Measured
2015	65	361	14,290	3,630	4,746	1,459
2016	65	325	11,221	2,781	5,787	1,217
2017	48	224	9,308	2,682	4,669	1,167
2018	50	250	9,225	1,889	9,533	1,434
2019	47	352	9,582	3,870	6,195	2,320
2021	46	349	6,328	5,135	1,748	1,534
2022	43	263	19,352	8,101	4,619	2,175
2023	39	251	7,074	4,009	3,370	1,649

Table 2: Nominal CPUE of all male GKC and scaled index in comparison to GLMM standardized index and associated CV.

Survey Year	CPUE	Nominal Index	CV	GLMM Index	CV
2015	52	1.012	0.04	1.366	0.07
2016	48	0.944	0.03	1.060	0.09
2017	61	1.183	0.03	1.211	0.09
2018	75	1.459	0.04	1.503	0.07
2019	52	1.019	0.03	1.062	0.10
2021	26	0.501	0.04	0.415	0.28
2022	72	1.405	0.03	1.088	0.10
2023	43	0.845	0.04	0.791	0.15

Table 3: Time series of number of crab measured (N) by the survey and bootstrap estimated effective sample size (N_{eff}). Number of crab measured is restricted to those > 100 mm carapace length.

Year	N	Bootstrap N_{eff}		
		Min	Mean	Max
2015	4,987	136	408	1,305
2016	3,686	146	594	3,979
2017	3,655	241	834	5,555
2018	3,110	163	645	2,720
2019	5,876	206	837	3,327
2021	6,541	68	367	5,336
2022	7,022	238	1,364	7,206
2023	5,632	234	1,262	4,929

Figures
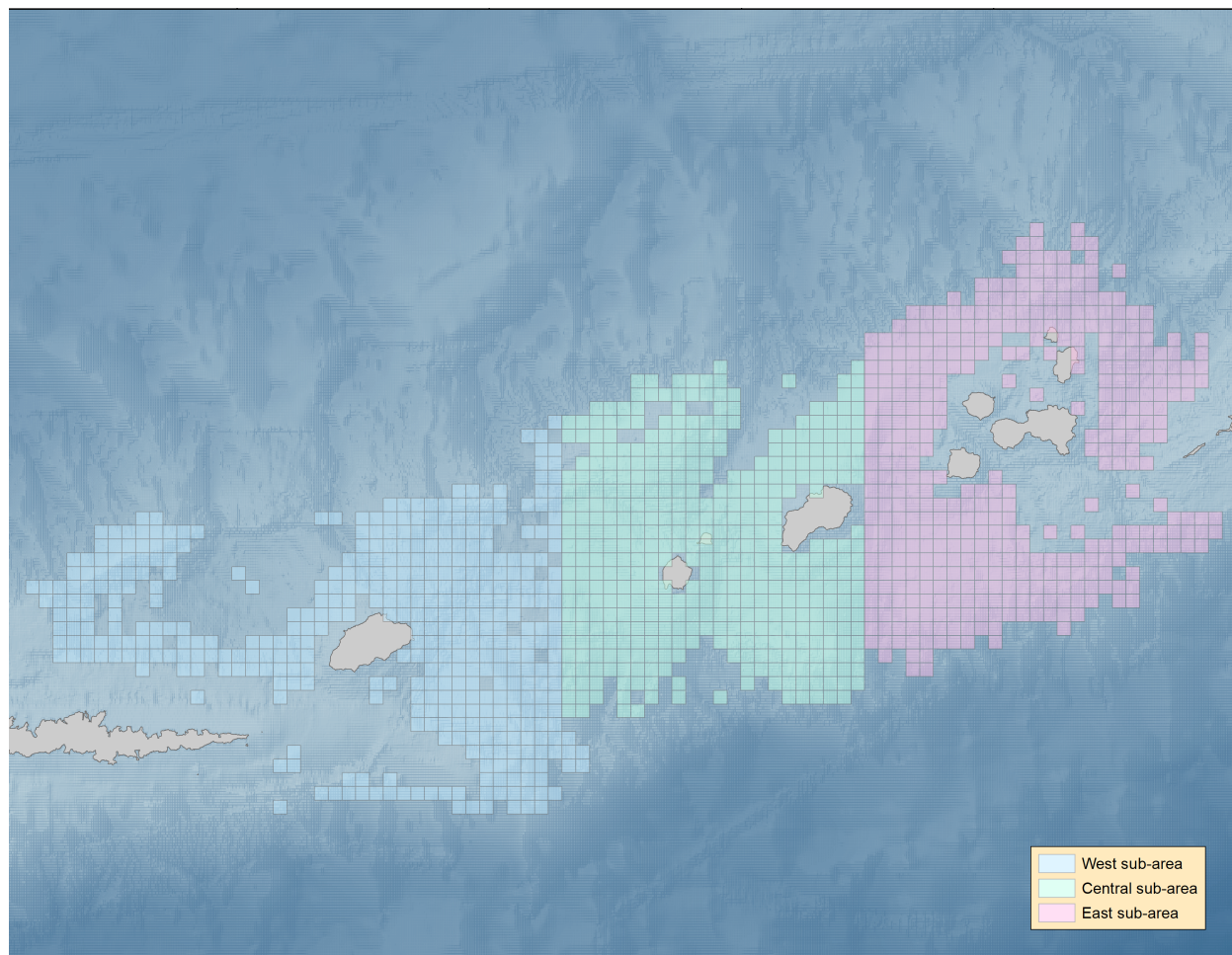
Figure 1: Map of survey grid in the EAG colored by stratum.
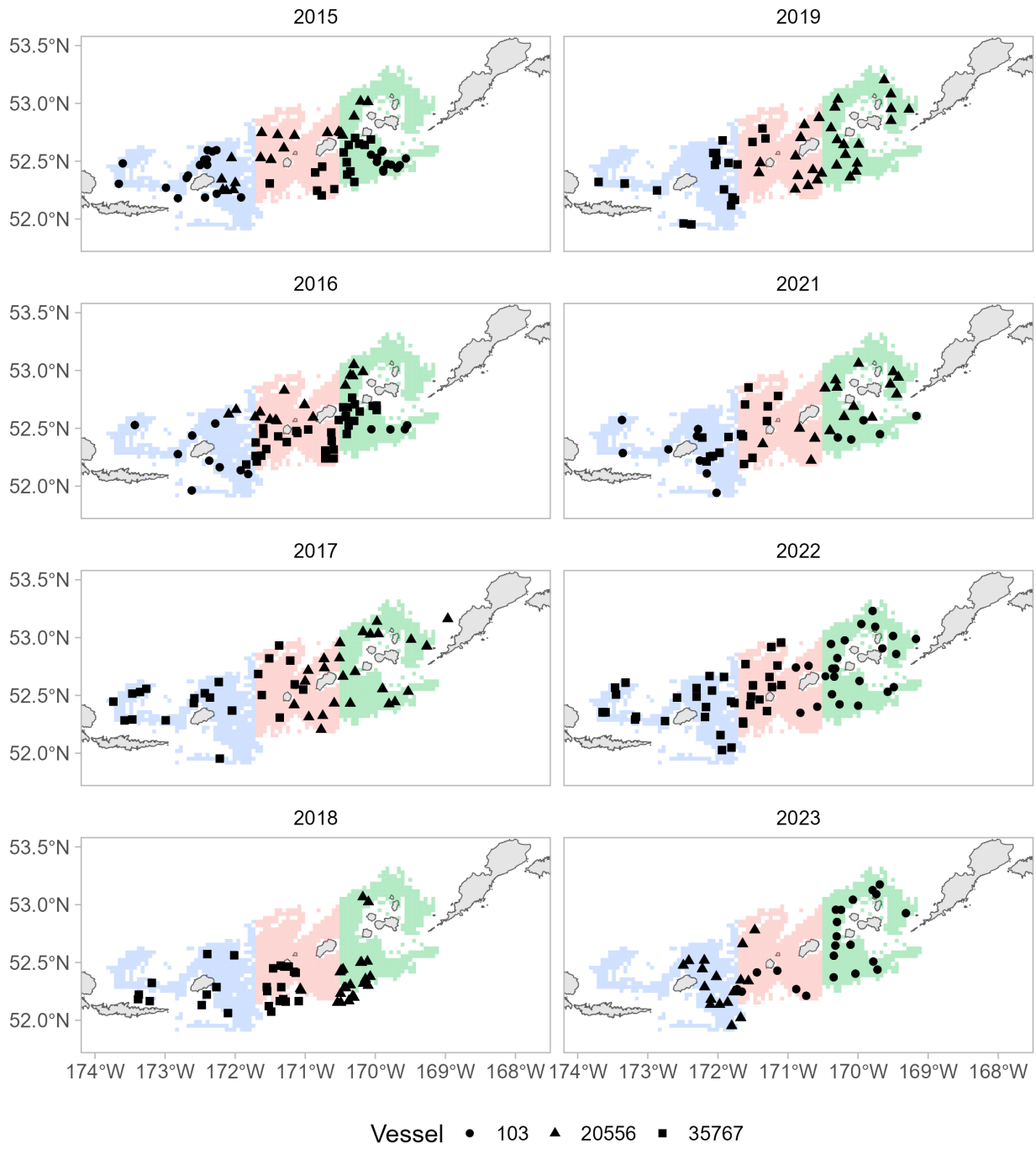
Figure 2: Map of stings sampled in the EAG by survey year.
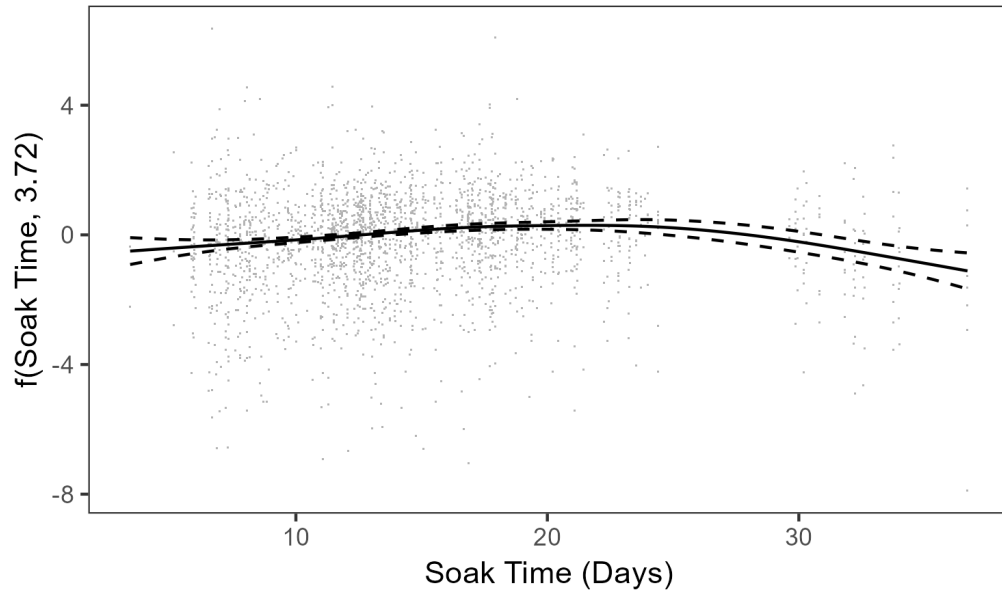
Figure 3: Marginal effect of soak time with associated partial residuals for the negative binomial GAMM fit to survey CPUE in the EAG.
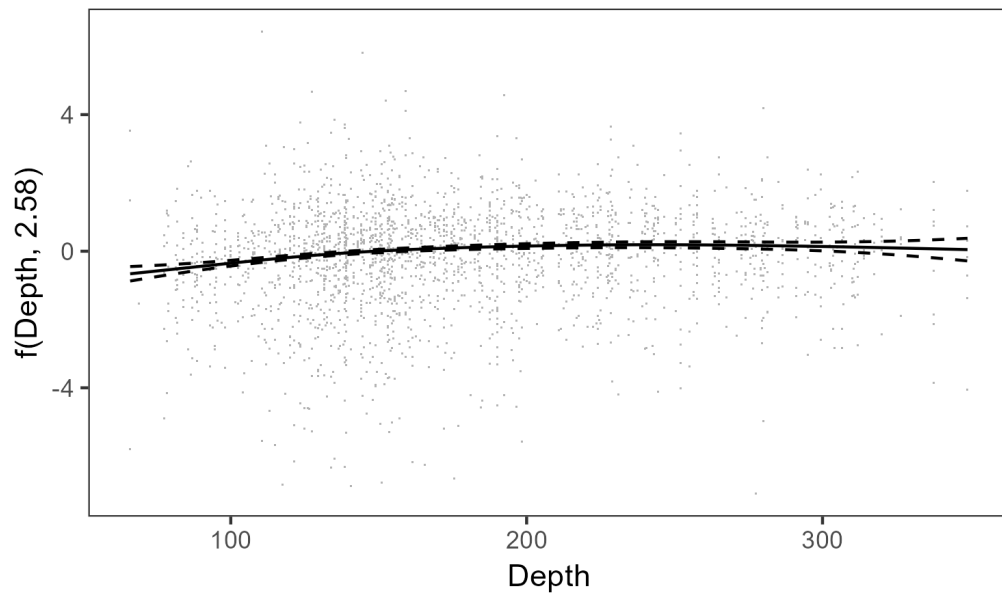
Figure 4: Marginal effect of depth with associated partial residuals for the negative binomial GAMM fit to survey CPUE in the EAG.

QQ plot residuals

Residual vs. predicted
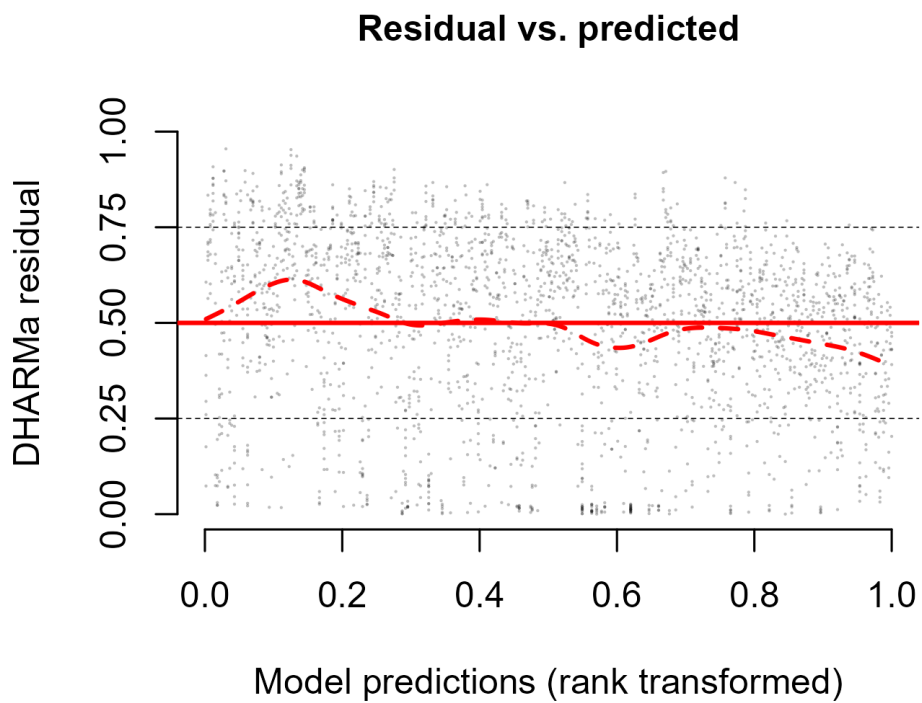
Figure 5: Diagnostic plots of DHARMa residuals for the negative binomial GAMM fit to survey CPUE in the EAG.

Residual vs. predictor

Residual vs. predictor

soaktime (rank transformed)
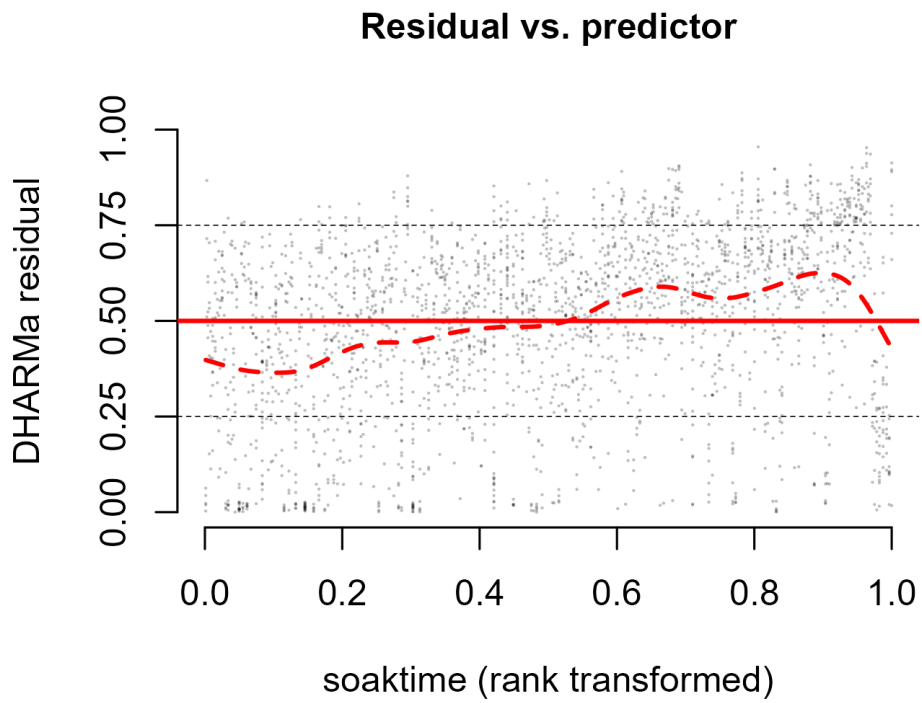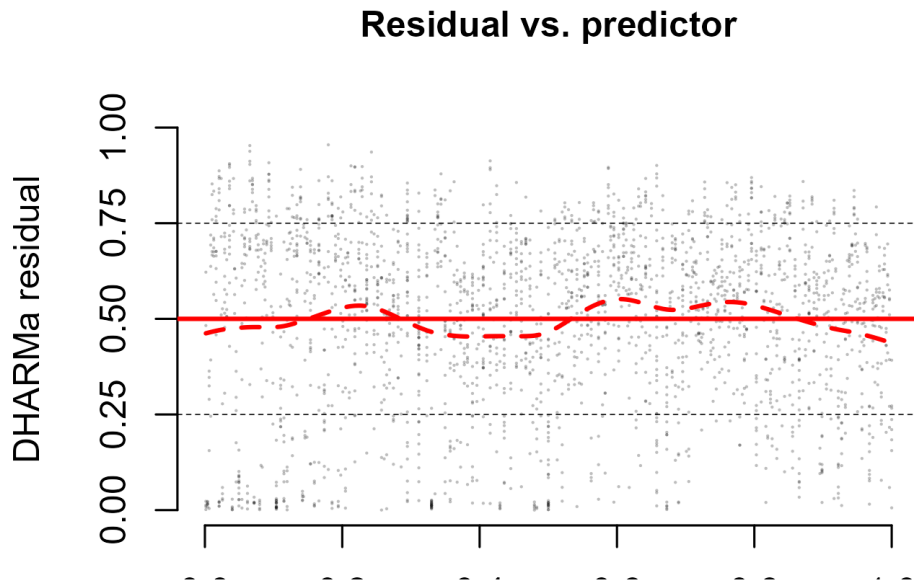
Figure 6:   DHARMa residuals plotted against smoothed terms for the negative binomial GAMM fit to survey CPUE in the EAG.
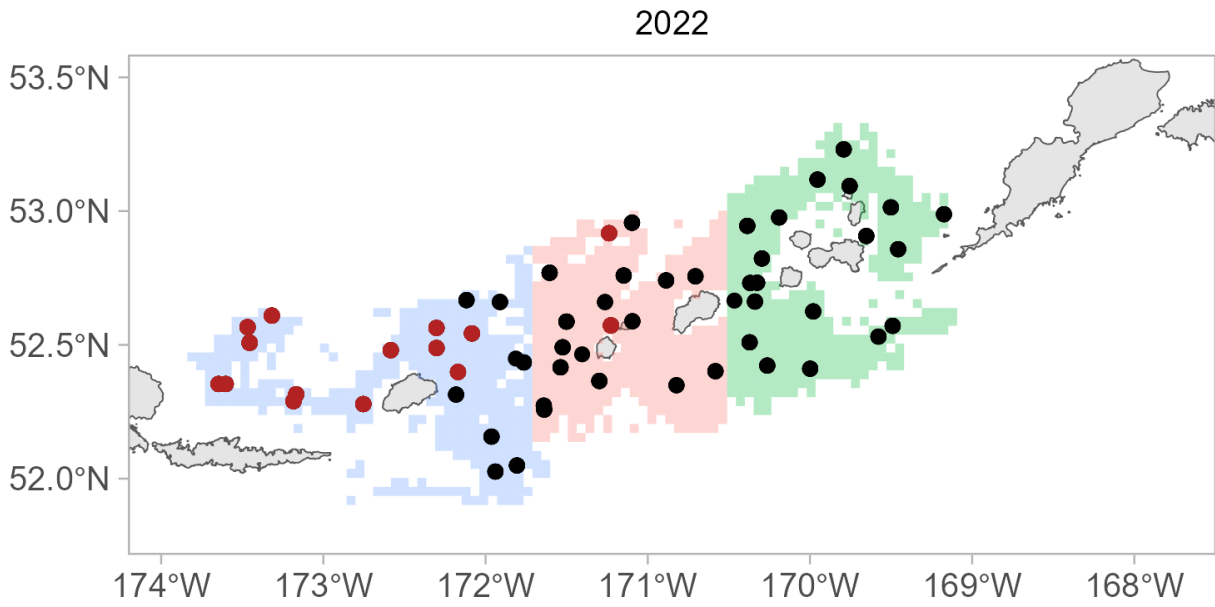
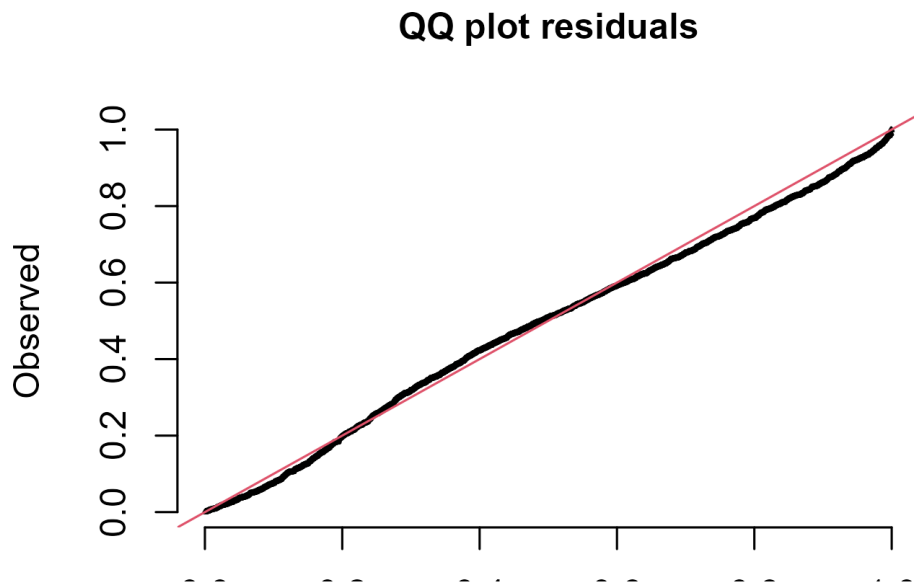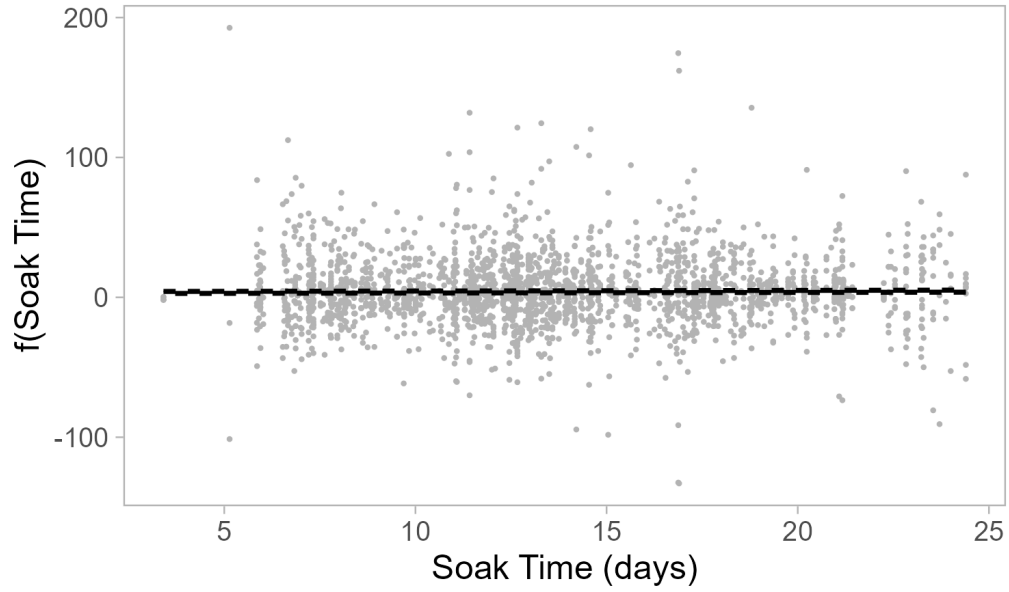Figure 7: Map of stings with soak times over 25 days (red) compared to location of all other pots (black) in the EAG.

QQ plot residuals

Residual vs. predicted
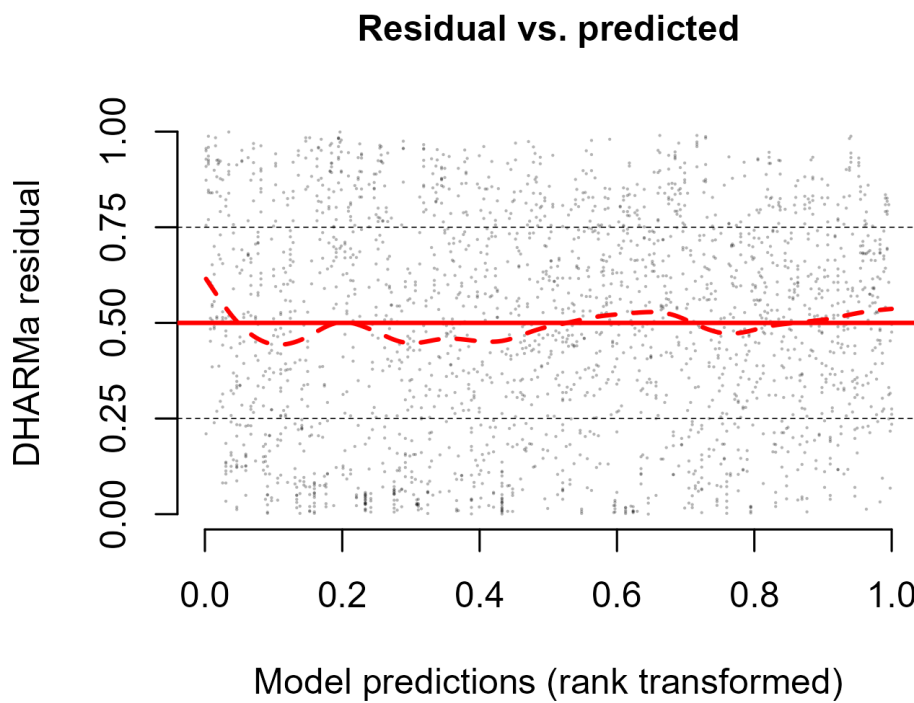
Figure 8: Diagnostic plots of DHARMa residuals for the Tweedie GLMM fit to survey CPUE in the EAG.

Figure 9: Marginal effect of soak time with associated partial residuals for the Tweedie GLMM fit to survey CPUE in the EAG.
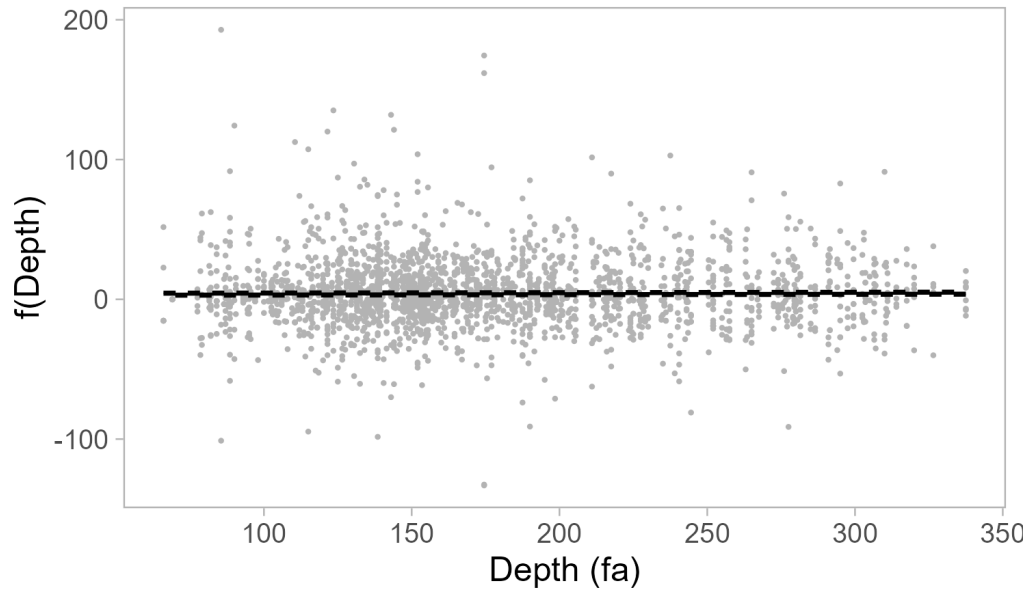

Figure 10: Marginal effect of depth with associated partial residuals for the Tweedie GLMM fit to survey CPUE in the EAG.
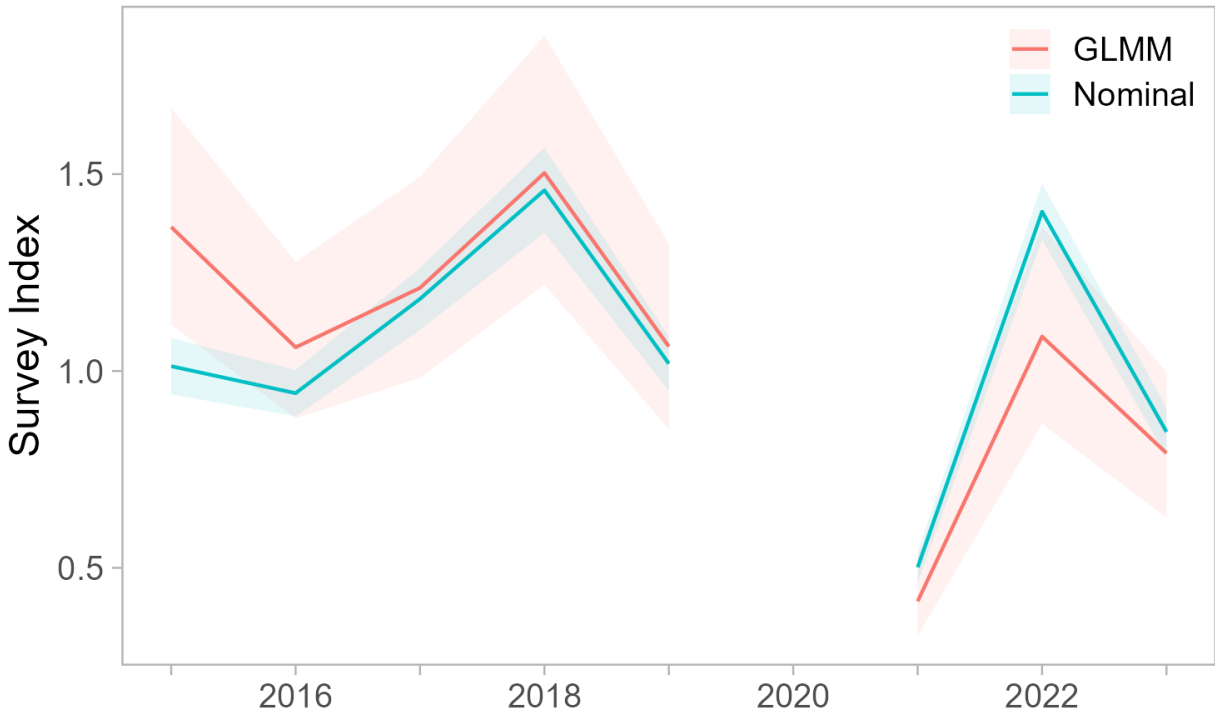
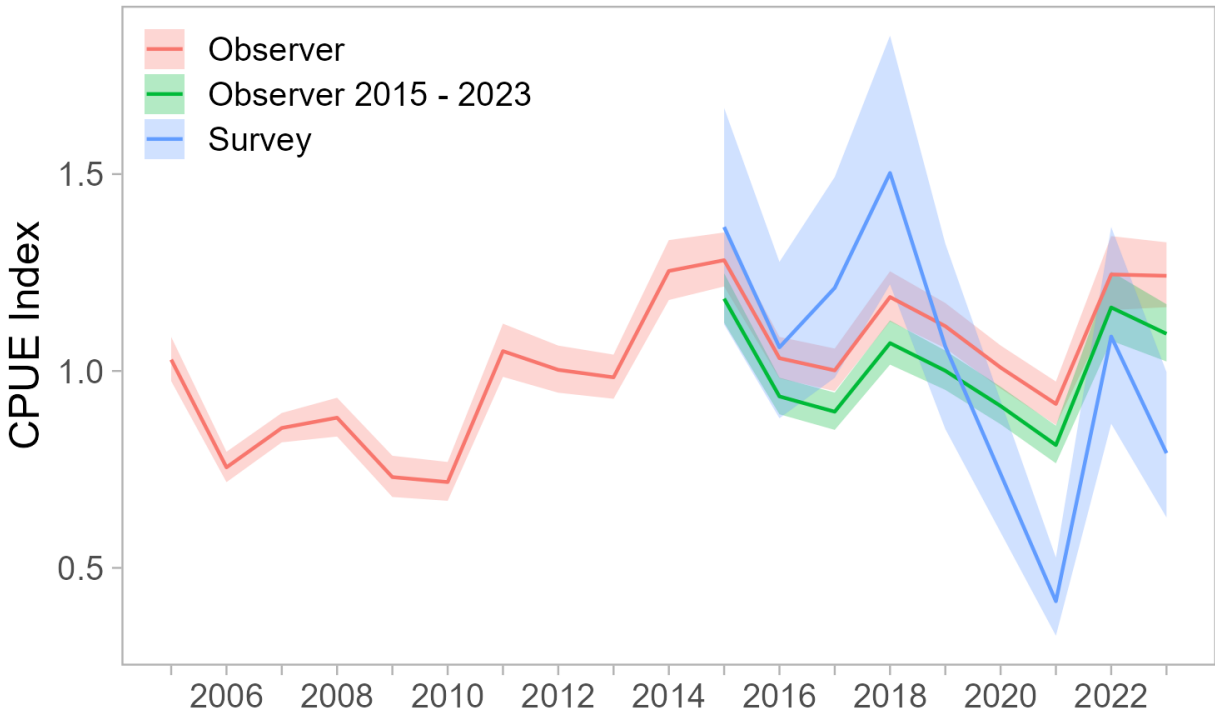Figure 11: Time series of standardized survey CPUE indices estimated for the EAG.

Figure 12: Time series of standardized survey CPUE index in comparison to standardized observer CPUE index (full post-rationalization time series and 2015 - 2023).
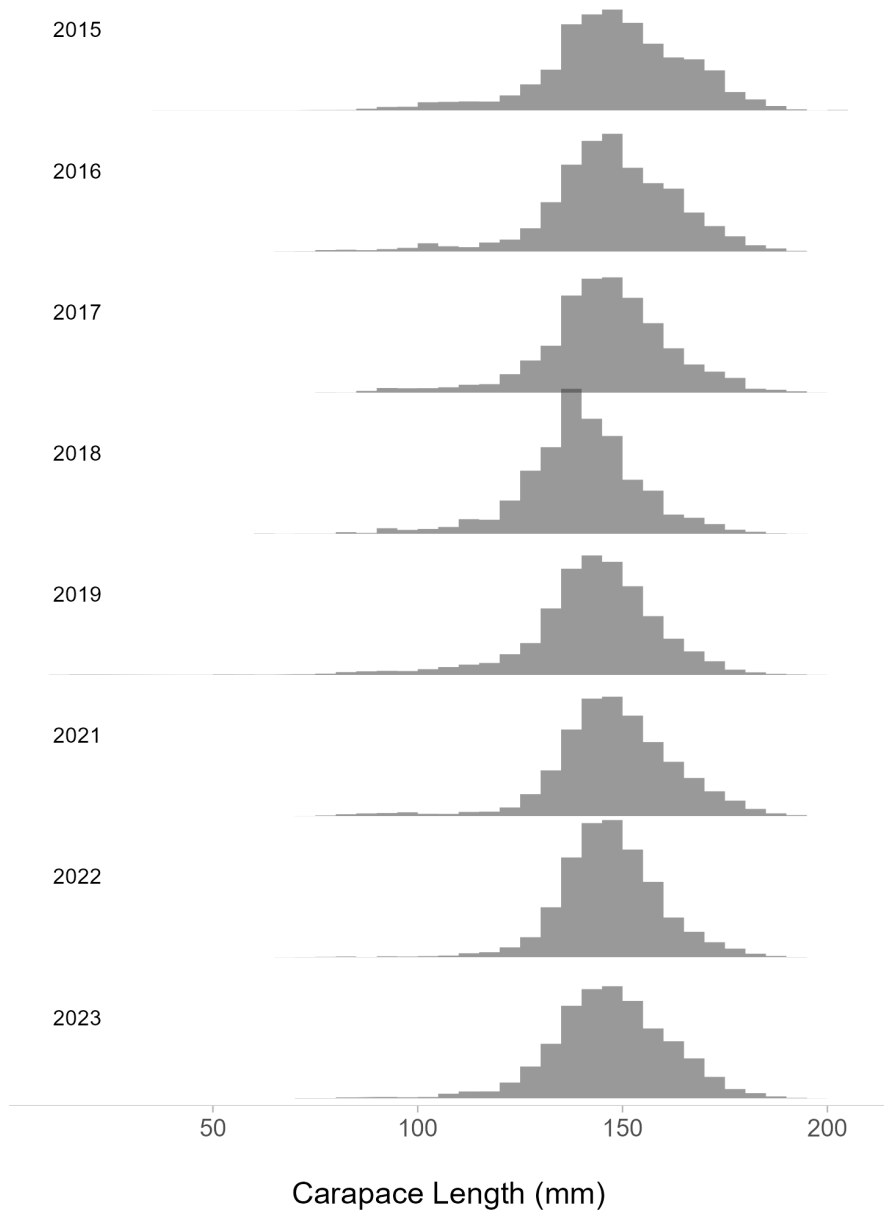
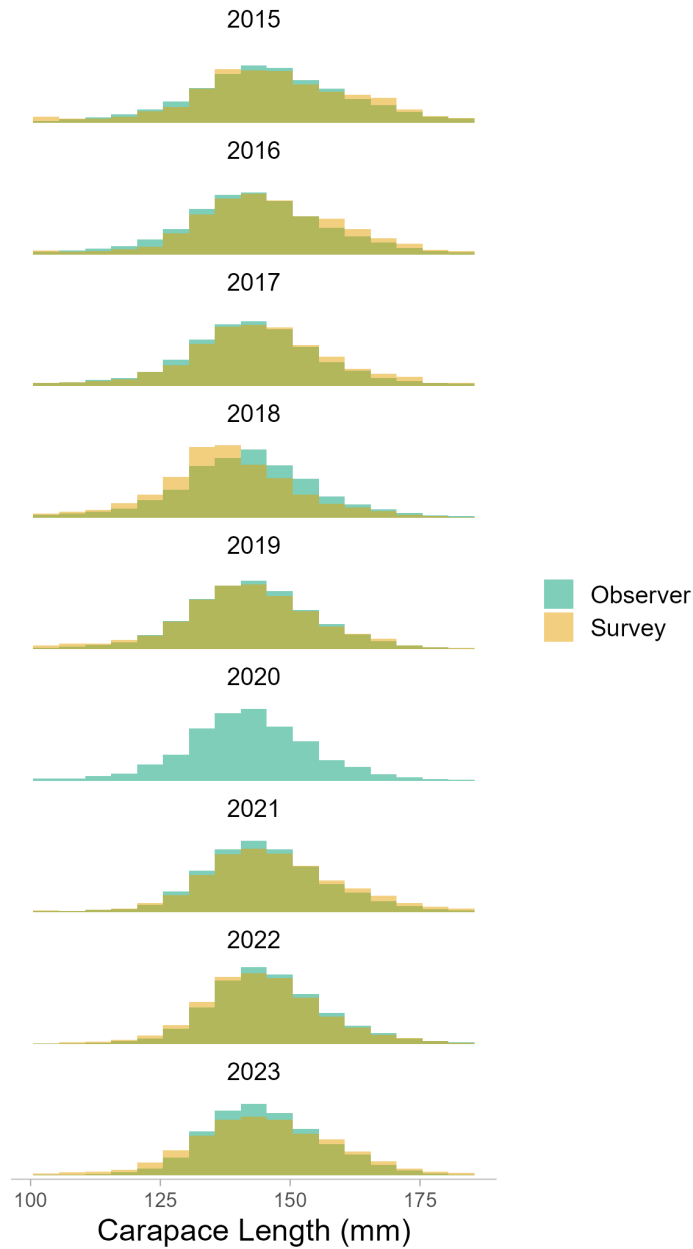Figure 13: Time series of carapace length composition estimated for the EAG.

Figure 14: Comparison of survey and observer size composition truncated to model size bins for the EAG.

Literature Cited

Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, Skaug HJ, Maechler M, Bolker BM (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R Journal 9(2): 378 - 400.

Dunn, P. K. (2022). Tweedie: Evaluation of Tweedie exponential family models. R package version 2.3.

Hartig, F. 2020. DHARMa: Residual Diagnostics for Hierarchical (multi-Level / Mixed) Regression Models. https://CRAN.R-project.org/package=DHARMa.

Jackson, T. 2024. Aleutian Islands golden king crab stock assessment 2024. North Pacific Fishery Management Council, Anchorage, AK.

Wood S, Scheipl F (2020). gamm4: Generalized Additive Mixed Models using 'mgcv' and 'lme4'. R package version 0.2-6. https://CRAN.R-project.org/package=gamm4.