Appendix A: Fishery CPUE Standardization

Tyler Jackson
Alaska Department of Fish and Game, tyler.jackson@alaska.gov

May 2026

Purpose

The AIGKC stock assessment has used catch per unit effort (CPUE) data collected by at-sea observers and fish ticket data as a primary index of stock abundance since model development began (Siddeek et al. 2017; Siddeek et al. 2016). The 2025 final assessment used general additive models (GAM) to standardize observer CPUE in two time periods: pre-rationalization (1995 - 2004) and post rationalization (2005 - present) (Jackson 2025). Several efforts have been made to account for spatial and spatiotemporal variability in fishing effort and CPUE including specifying large-scale blocks within subdistricts (here; Siddeek et al. 2016), interactions between blocks and year (Jackson 2024, Siddeek et al. 2023), non-parametric smooths of latitude and longitude, and interactions between year and smooth terms (Jackson 2024). Geostatistical models may provide better utility over previous approaches given their ability to estimate spatial correlations and account for varying spatial coverage (Maunder et al. 2020). This appendix details an update to the GAM CPUE standardization for the base model 23.1c, as well as an alternative CPUE standardization using spatiotemporal generalized additive mixed models (26.0, 26.0a, and 26.1).

Methods

Core Data Preparation

Observer data sets were limited to pots that represented core fishing effort in an attempt to remove observations that may not be indicative of overall fishing performance. Core vessels and permit holders during the pre-rationalized time series were those that participated in more than a single season. The fleet was consolidated enough in the post-rationalized time series that reductions on number of vessels and permit holders were not warranted. Following Siddeek et al. (2016; 2023) several gear types were combined, and pot types not typical to the directed fishery were removed. Pot sizes 4'x4', 10'x10', 8'x8' (EAG only), and 6.5'x7' (WAG only) were removed due to small sample sizes. Since many fishing seasons in the pre-rationalized era did not align with the crab year used in the post-rationalized era (July - June), crab year was assigned to pre-2005 data *post hoc*. Observer pots sampled on dates that fall after June 30 in a given season were assigned the next crab year (Siddeek et al. 2016, 2023). Soak time and depth data were truncated by removing the outer 5% and 1% of distributions, respectively.

GAM

Since there were no changes to the pre-rationalized time period, CPUE standardization was only updated for post-rationalization. Tweedie GAMs with a log link were estimated using the R package *mgcv* (Wood 2004). The power variable, p, that relates the Tweedie mean to its variance was estimated as a model parameter. All

splines were fit as thin plate regression splines, with smoothness determined by generalized cross-validation (Wood 2004).

Null models included only crab year as an explanatory variable

$$\ln(CPUE_i) = Year_{y,i} \tag{1}$$

The full scope of models evaluated included gear (i.e., pot size), vessel, permit holder (i.e., proxy for captain), month and block (i.e., discrete geographic subarea) as factorial variables. Prospective smoothed terms include soak time, depth, and slope angle. Sea floor slope angle (degrees) was computed in ArcGIS (Redlands 2011) from a 100-m resolution raster surface of Aleutian Islands bathymetry (Zimmermann 2013). Addition of new variables was considered significant if AIC decreased by at least two per degree of freedom lost and R^2 increased by at least 0.01. Variables were added (or subtracted) from the model until no candidate variables met AIC and R^2 criteria. Consistent AIC (CAIC; -2LogLik+(ln(n)+1)p; Bozdogan 1987) was used instead of the traditional AIC, in which n is the number of observations and p is the number of parameters (Siddeek et al. 2016, 2023). If forward and backward selection produced conflicting results, the best model was determined by CAIC.

$$R^2 = \frac{D_{Null} - D_{Resid}}{D_{Null}} \tag{2}$$

GLMM

Saptiotemporal GLMMs were constructed using the R package *sdmTMB* and *sdmTMBextra* (R Core Team 2024; Anderson et al., 2024; Anderson et al., 2025). This approach models spatial random effects as a series of Gaussian random fields, which are approximated using stochastic partial differential equation matrices (SPDE) (Lindgren et al., 2011). Correlation of spatial random effects is constrained by the Matérn covariance function (Anderson et al., 2022).

The underlying spatial domain of the model was represented by a triangular mesh constructed using k-means clustering with 150 knots. Pot locations were transforms to UTM coordinates using zone 2N. Spatial polygons of the Aleutian Islands were downloaded using the R package *geodata* (Hijmans et al., 2025) and used as barriers to spatial correlation within the mesh (Figure 8).

In addition to year, gear type was included as a factor covariate, soak time and depth were included as smooth splines as described above, and vessel was included as a random effect. Models assumed a Tweedie error distribution with log link function and estimated power variable, p. Spatiotemporal random fields were modeled as independent and identically distributed process. Pre- and post-rationalization periods were fit in a single model to leverage the most available data within the standardization model. A model with the same form was also fit without spatial random effects to allow for a more direct comparison with a non-spatial model.

Model Diagnostics

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. Several other diagnostics including tests of convergence and range of parameter estimates are built in to the `sanity()` function of $sdmTMB$ were used to evaluate GLMMs.

Spatial residual patterns were evaluated using a Moran's I clustering analysis within the spatial domain of each year (Moran 1950; Cacciapaglia et al. 2024). Moran's I was computed via Monte Carlo simulations in the R package *spdep* (Bivand 2022). Significant Moran's I values indicate non-random, spatial clustering.

Partial effects were plotted to view the relationship between CPUE and individual variables. Step plots that show the change the standardized index with addition of each explanatory variable were also examined to consider the influence of each variable (Bishop et al. 2008; Bentley et al. 2012).

Index Prediction

Following Siddeek et al. (2016, 2023) standardized CPUE index was extracted from the GAMs as the year coefficient (β_i) with the first level set to zero. Standardized CPUE index for GLMMs was estimated by constructing a 5 km^2 prediction gird within the boundary of the model mesh. Depth for each grid cell was taken from the nearest neighbor in a 100 m resolution raster of bathymetry of the Aleutian Islands (updated version of Zimmerman and Prescott 2021). Grid cells were limited to depths up to 1.5 times the maximum depth observed in EAG and WAG observer pots (333 fa and 301 fa, respectively). The bathymetry of the Aleutian Islands is highly complex, and it is worth noting that there is some inconsistency in bathymetric raster and depths of nearby observer pots. That being noted, these data are the best available and depth does not have a strong influence on the resulting CPUE index (Figure 23 and 25). Soak time was set to the mean value while gear type was set to the reference level. Vessel was not included in index prediction. Predictions were summed across the spatial domain using area weighting.

For both model types the resulting index was scaled to canonical coefficients (β_i') as

$$\beta_i' = \frac{\beta_i}{\bar{\beta}} \tag{3}$$

where

$$\bar{\beta} = \sqrt[n_j]{\prod_{j=1}^{n_j} \beta_{i,j}} \tag{4}$$

and n_j is the number of levels in the year variable. Nominal CPUE was scaled by the same method for comparison.

Results and Conclusions

GAM

Forward and backward selection resulted in different 'best' models in both subdistricts. The best forward selection model in the EAG included (in addition to year) vessel, permit holder, gear type, and soak time, whereas backwards selection did not include permit holder (Table 1). The backward selection model was chosen for index standardization (Δ AIC = -46). In the WAG, forward selection resulted in a model including vessel, gear type, and month, whereas backward selection included permit holder instead of month and was used for index standardization (Δ AIC = -38) (Table 2). Residual diagnostics did not indicate misspecification in best models for either subdistrict (Figure 1 and 2).

Partial effects of covariates are in Figure 3 and 4). Standardized indices differed minimally from null model indices. Vessel and permit holder appeared to have the largest influence on the resulting index in the EAG and WAG, respectively (Figure 5 and 6). Standardized indices track nominal indices, as expected. Relative to the 2023/4, both subdistricts underwent a decrease in CPUE (Figure 7).

Spatiotemporal GLMM

Full models successfully converged and DHARMa residuals did not identify misspecifications (Figure 9 and 11). Moran's I for the GLMM model ranged from -0.025 - 0.095 in the EAG and -0.022 - 0.067 in the WAG (Table 3 and 4). Despite apparently weak clustering, many tests in the EAG pre-2009 and many years in the WAG were significant, likely due to the very large sample size. Still, the spatiotemporal model resolved spatial autocorrelation somewhat better than the GAM models (Table 3 and 4, Figure 10 and 12).

The effect of gear type suggested increasing CPUE with large pots and considerably lower CPUE for round pots. Smoothed soak time was approximately dome shaped for both subdistricts. Depth had a weak positive linear realtionship with CPUE in the EAG and a weak negative relationship with CPUE in the WAG (Figure 13 and 14).

Spatial and spatiotemporal random effects are shown in Figures 15 and 19, and 16 and 20, respectively. Maps of predicted CPUE indicated that in years with higher CPUE, estimated CPUE was higher across the full model domain, not just the areas with data (Figures 17 and 21). Spatial estimates were fairly imprecise during 1995 - 1996 in the EAG and 1995 and 1999 in the WAG. Generally, CV's were low throughout much of the prediction grid though there were more 'hotspots' of high uncertainty in the WAG than the EAG (Figure 18 and 22).

Adding vessel as a random effect and smoothed soak time had the greatest impact on the standardized index in both subdistricts (Figures 23 and 25). Ultimately, standardized indices were greater than nominal indices during pre-rationalized years, and less than nominal indices during post-rationalized years. Indices resulting from spatiotemporal GLMMs followed similar trajectories to indices derived from a non-spatial GAMM (Figures 24 and 26).

Spatiotemporal models seem suitable for standardization of AIGKC CPUE data given clear spatial variability in the fishery footprint, though resulting indices were not drastically different than those estimated without spatiotemporal random effects. CPUE standardization has fit separate models to pre- (1995/96 - 2004/05) and post-rationalization (2005/06 - present) data since a length-based assessment model has been used for this stock (Siddeek et al. 2017). Naturally, indices derived from these models have been assigned different catchability parameters within the assessment model. Hoyle et al. (2024) argues that splitting the CPUE time series limits its utility as an index of abundance and that continuity should be preserved when possible. Time blocks for catchability are best addressed within the assessment framework. Here, I leveraged the full time series in index standardization, though it is possible than some effects could be revisited with consideration for rationalization (e.g. soak time).

Tables

Table 1: Residual degrees of freedom, AIC, and R^2 for the EAG post-rationalized period best legal CPUE model including year (Yr), vessel (Ves), gear type (Gr), and s(soak time).

Form ($p = 1.385$)	Residual DF (Δ DF)	AIC (Δ AIC)	R^2 (Δ R^2)
Yr + s(soak time, 4.44) + Mon + Ves + Gr	10,824.73	94,999	0.14
+ s(depth)	-4.02	15.54	0.002
+ s(slope)	-3.08	7.27	0.002
+ Block	-3.00	23.58	0.001
- Permit Holder	-13.4	26	0.008

Table 2: Residual degrees of freedom, AIC, and R^2 for the WAG post-rationalized period best legal CPUE model including year (Yr), gear type (Gr), permit holder (PH), and s(longitude, latitude).

Form ($p = 1.495$)	Residual DF (Δ DF)	AIC (Δ AIC)	R^2 (Δ R^2)
Yr + Mo + PH + Gr	18,008	142,306	0.097
+ s(soak time)	-8.01	-27.53	0.005
+ s(depth)	-4.17	-33.85	0.003
+ s(slope)	-2.98	4.17	0.001
+ Block	-5.00	16.10	0.002
- Vessel	-3	18.66	0.000

Table 3: Moran's I and associated p-value by year for the spatiotemporal GLMM, GLMM, and GAM fit to legal CPUE in the EAG.

	ST GLMM		GLMM		GAM	
Year	Moran's I	p	Moran's I	p	Moran's I	p
1995	0.051	0.001	0.121	0.001	0.120	0.001
1996	0.074	0.001	0.169	0.001	0.129	0.001
1997	0.031	0.001	0.117	0.001	0.088	0.001
1998	0.029	0.001	0.116	0.001	0.099	0.001
1999	0.048	0.001	0.135	0.001	0.116	0.001
2000	0.084	0.001	0.200	0.001	0.160	0.001
2001	0.075	0.001	0.202	0.001	0.186	0.001
2002	0.053	0.001	0.157	0.001	0.160	0.001
2003	0.061	0.001	0.158	0.001	0.139	0.001
2004	0.037	0.001	0.146	0.001	0.138	0.001
2005	0.019	0.071	0.211	0.001	0.159	0.001
2006	0.035	0.003	0.138	0.001	0.129	0.001
2007	0.070	0.001	0.181	0.001	0.194	0.001
2008	0.031	0.039	0.181	0.001	0.162	0.001
2009	-0.007	0.542	0.012	0.237	-0.013	0.634
2010	0.010	0.244	0.156	0.001	0.174	0.001
2011	0.011	0.277	0.111	0.001	0.131	0.001
2012	0.001	0.384	0.157	0.001	0.160	0.001
2013	0.030	0.044	0.153	0.001	0.132	0.001
2014	0.002	0.376	0.132	0.001	0.107	0.001
2015	0.050	0.006	0.140	0.001	0.120	0.001
2016	0.015	0.172	0.123	0.001	0.114	0.001
2017	-0.005	0.572	0.067	0.001	0.044	0.005
2018	-0.016	0.767	0.063	0.002	0.067	0.001
2019	-0.024	0.897	0.113	0.001	0.138	0.001
2020	0.043	0.013	0.081	0.001	0.083	0.001
2021	0.010	0.284	0.214	0.001	0.234	0.001
2022	-0.019	0.685	0.106	0.001	0.118	0.001
2023	0.025	0.133	0.069	0.014	0.073	0.003
2024	0.012	0.271	0.184	0.001	0.126	0.001
2025	0.028	0.110	0.142	0.001	0.167	0.001

Table 4: Moran's I and associated p-value by year for the spatiotemporal GLMM, GLMM, and GAM fit to legal CPUE in the WAG.

	ST GLMM		GLMM		GAM	
Year	Moran's I	p	Moran's I	p	Moran's I	p
1995	0.048	0.001	0.126	0.001	0.099	0.001
1996	0.043	0.001	0.119	0.001	0.096	0.001
1997	0.037	0.001	0.104	0.001	0.045	0.001
1998	0.032	0.001	0.088	0.001	0.073	0.001
1999	0.047	0.001	0.097	0.001	0.086	0.001
2000	0.032	0.001	0.106	0.001	0.093	0.001
2001	0.035	0.001	0.082	0.001	0.105	0.001
2002	0.026	0.001	0.093	0.001	0.075	0.001
2003	0.052	0.001	0.156	0.001	0.136	0.001
2004	0.055	0.001	0.180	0.001	0.175	0.001
2005	0.008	0.242	0.114	0.001	0.109	0.001
2006	0.002	0.375	0.066	0.001	0.076	0.001
2007	0.013	0.152	0.121	0.001	0.104	0.001
2008	0.013	0.165	0.140	0.001	0.114	0.001
2009	0.003	0.388	0.208	0.001	0.164	0.001
2010	0.040	0.002	0.164	0.001	0.153	0.001
2011	-0.014	0.789	0.077	0.001	0.093	0.001
2012	0.020	0.059	0.147	0.001	0.150	0.001
2013	0.002	0.413	0.093	0.001	0.101	0.001
2014	0.021	0.043	0.122	0.001	0.100	0.001
2015	0.009	0.199	0.105	0.001	0.077	0.001
2016	-0.001	0.485	0.136	0.001	0.134	0.001
2017	0.037	0.016	0.179	0.001	0.158	0.001
2018	-0.012	0.723	0.122	0.001	0.123	0.001
2019	0.016	0.093	0.133	0.001	0.073	0.001
2020	0.007	0.243	0.165	0.001	0.148	0.001
2021	0.053	0.001	0.144	0.001	0.136	0.001
2022	0.010	0.251	0.103	0.001	0.112	0.001
2023	0.050	0.001	0.108	0.001	0.109	0.001
2024	0.024	0.098	0.118	0.001	0.141	0.001
2025	0.009	0.290	0.221	0.001	0.241	0.001

Figures

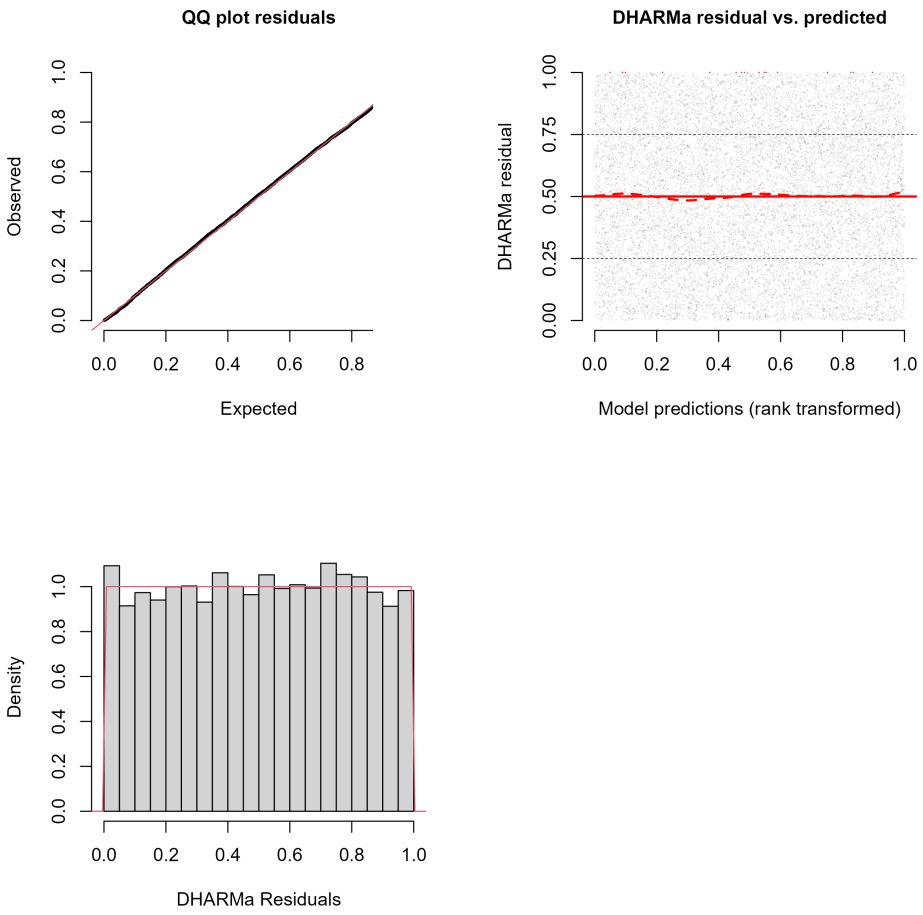

Figure 1: DHARMa residual plots for the final Tweedie GAM fit to legal CPUE during the post-rationalized period in the EAG.

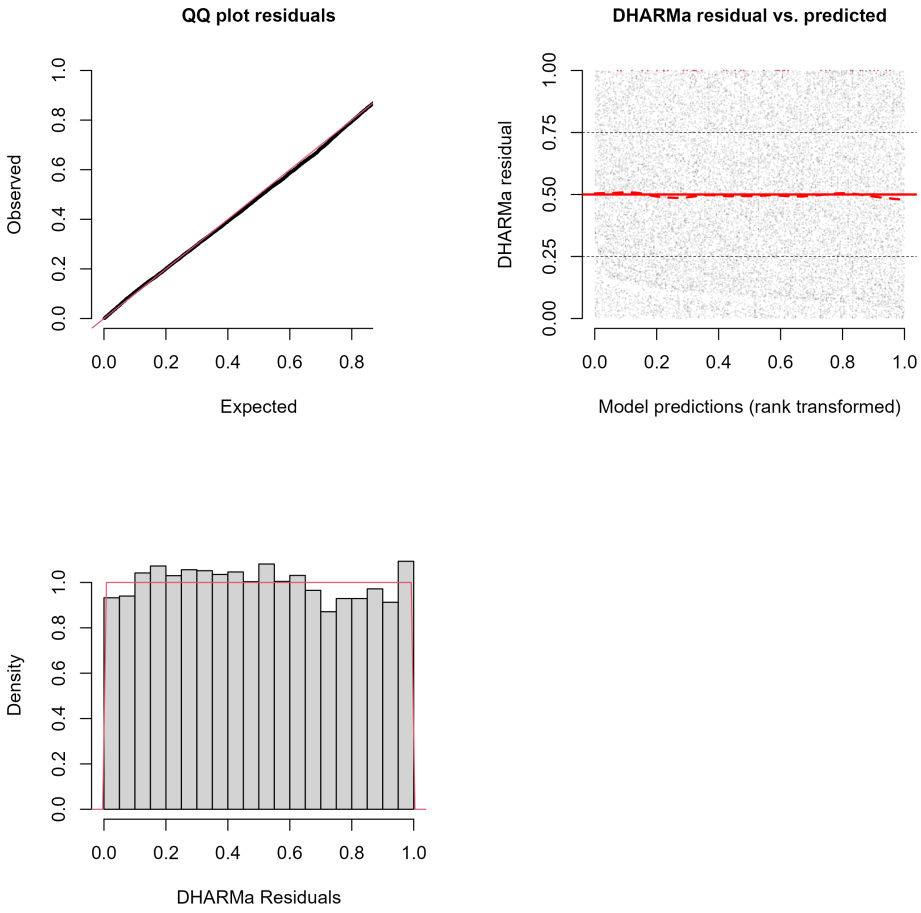

Figure 2: DHARMa residual plots for the final Tweedie GAM fit to legal CPUE during the post-rationalized period in the WAG.

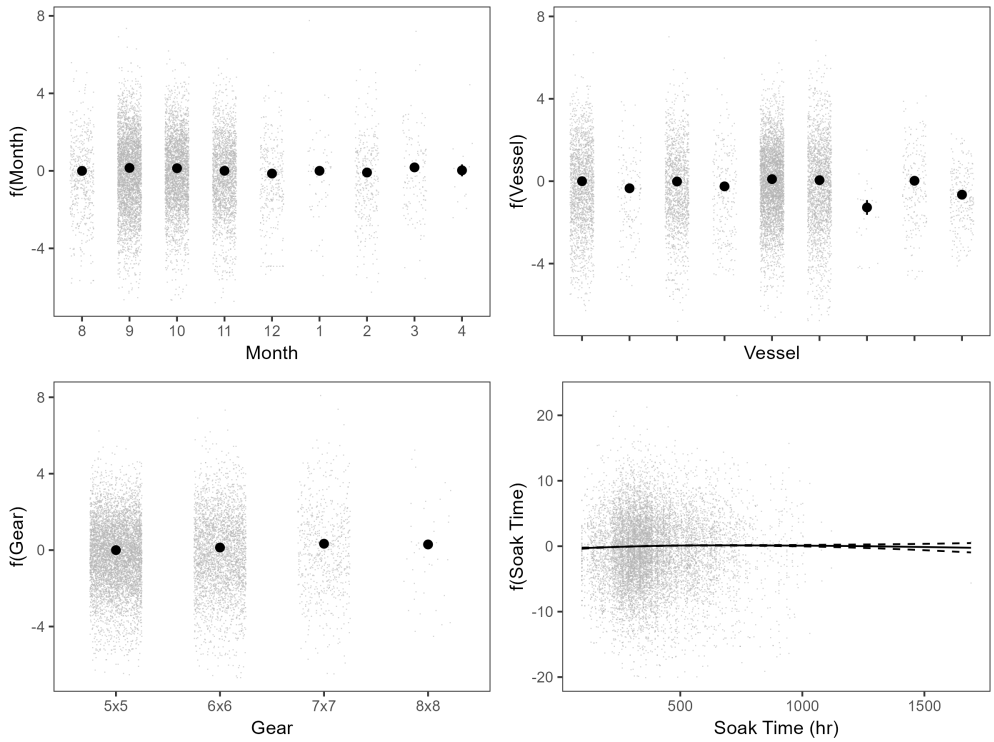

Figure 3: Marginal effects of month, permit holder, and gear type with associated partial residuals for the GAM fit to legal CPUE during the post-rationalized period in the EAG.

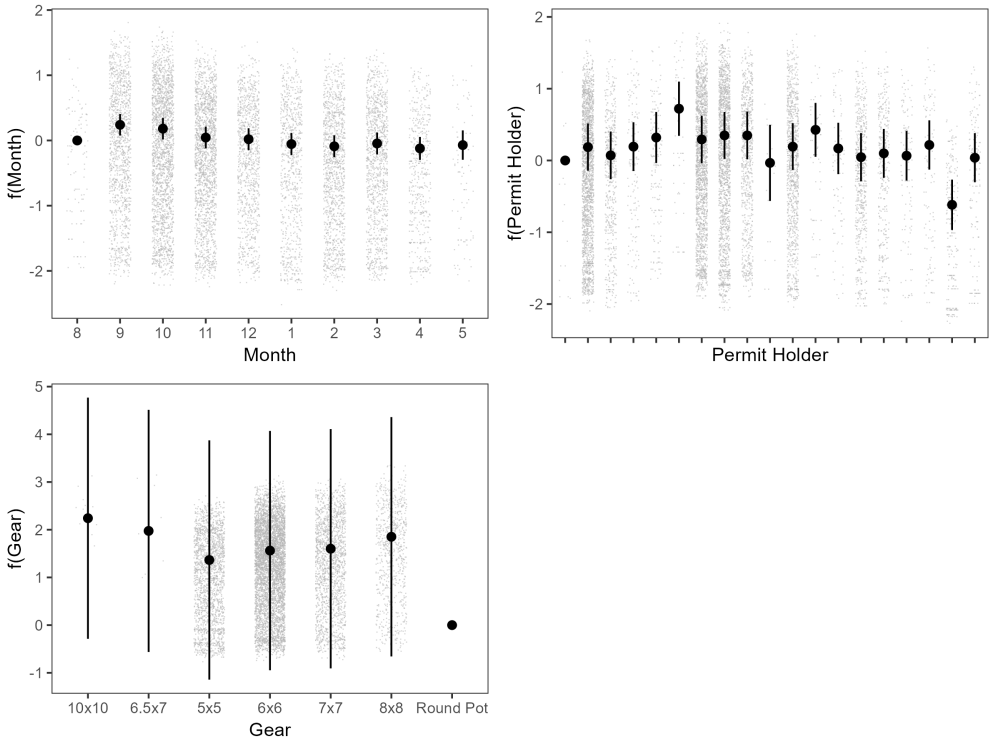

Figure 4: Marginal effects of month, permit holder, and gear type with associated partial residuals for the GAM fit to legal CPUE during the post-rationalized period in the WAG.

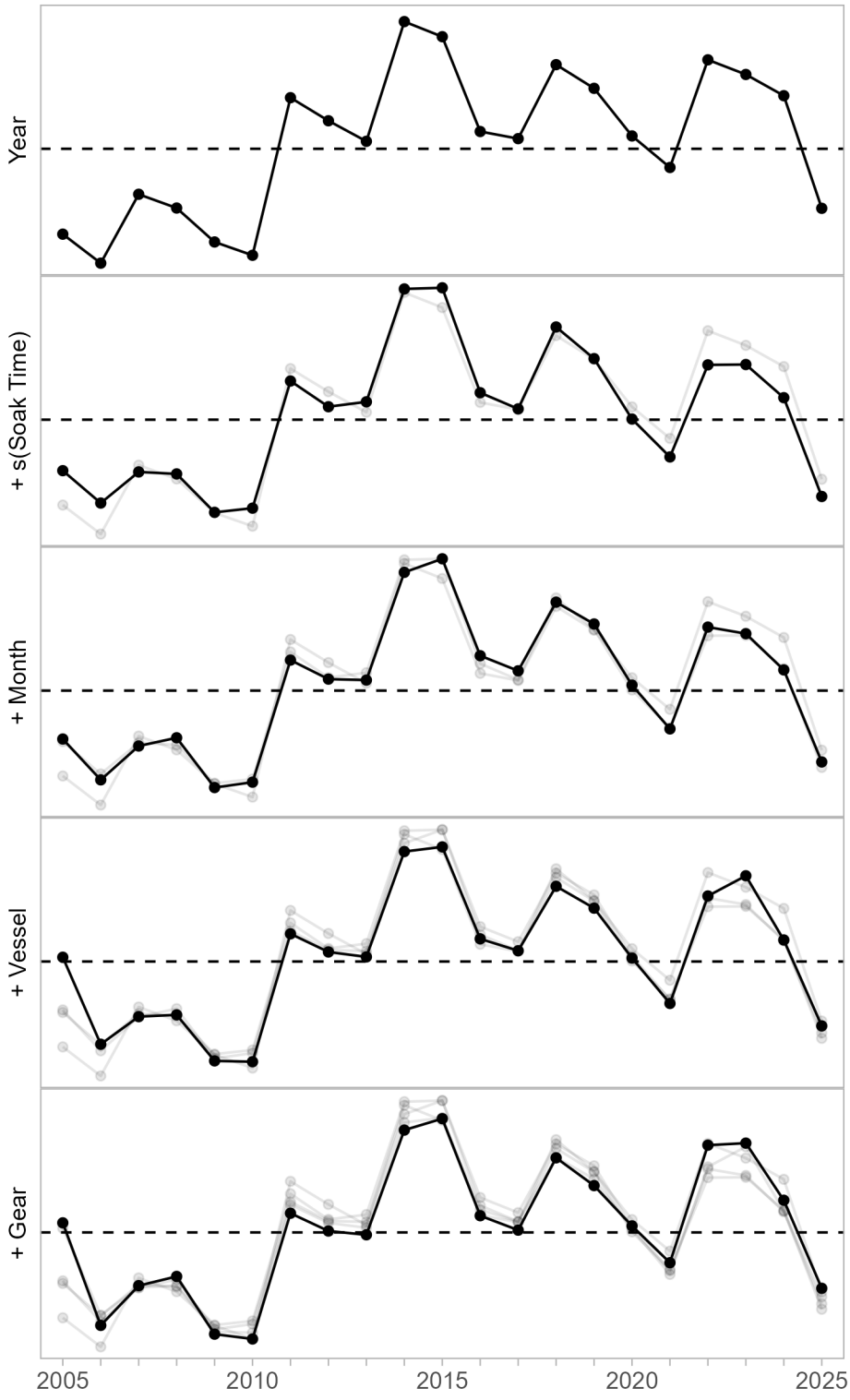

Figure 5: Step plot of CPUE index for the GAM fit to legal CPUE during the post-rationalized period in the EAG.

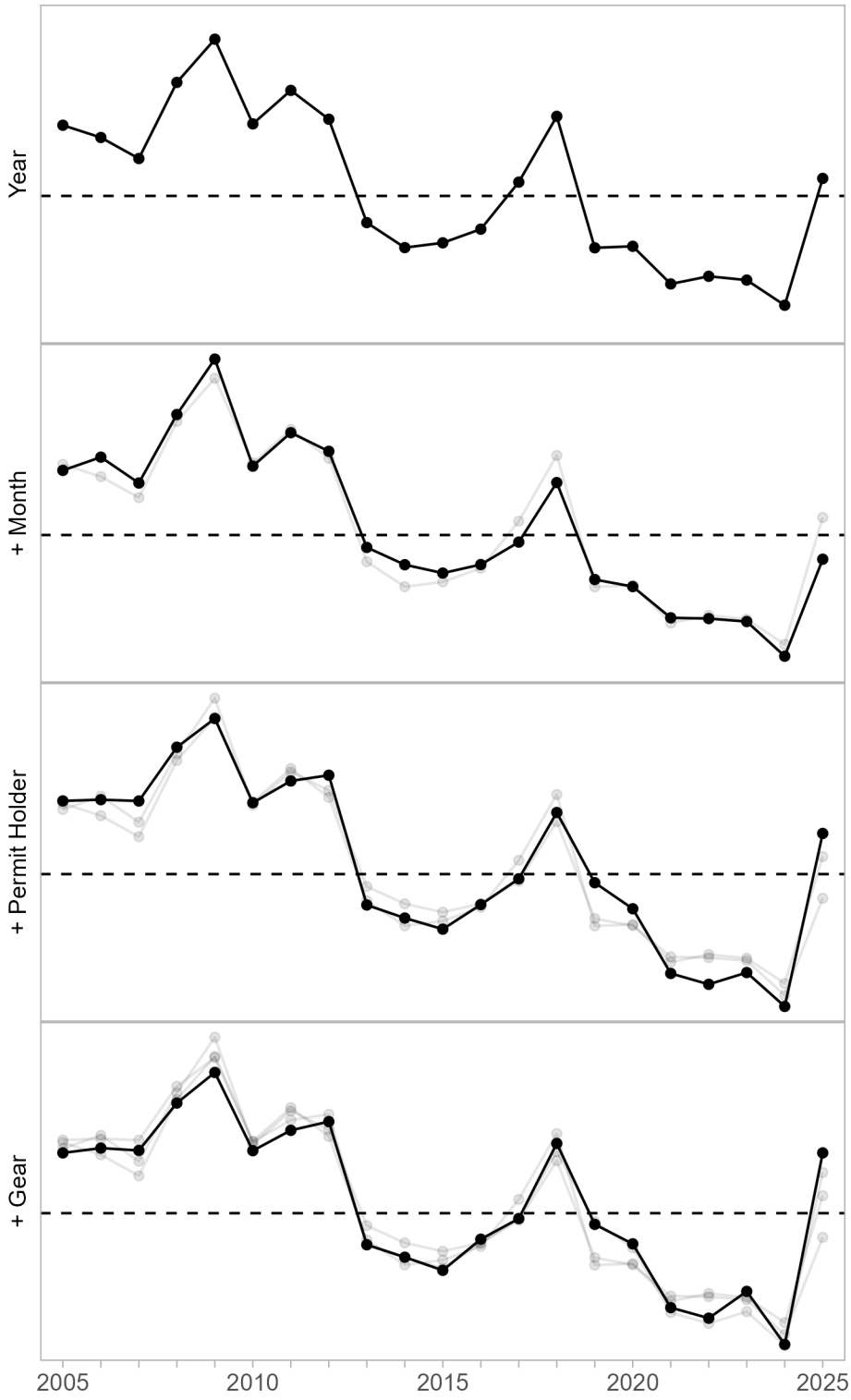

Figure 6: Step plot of CPUE index based on the GAM fit to legal CPUE during the post-rationalized period in the WAG.

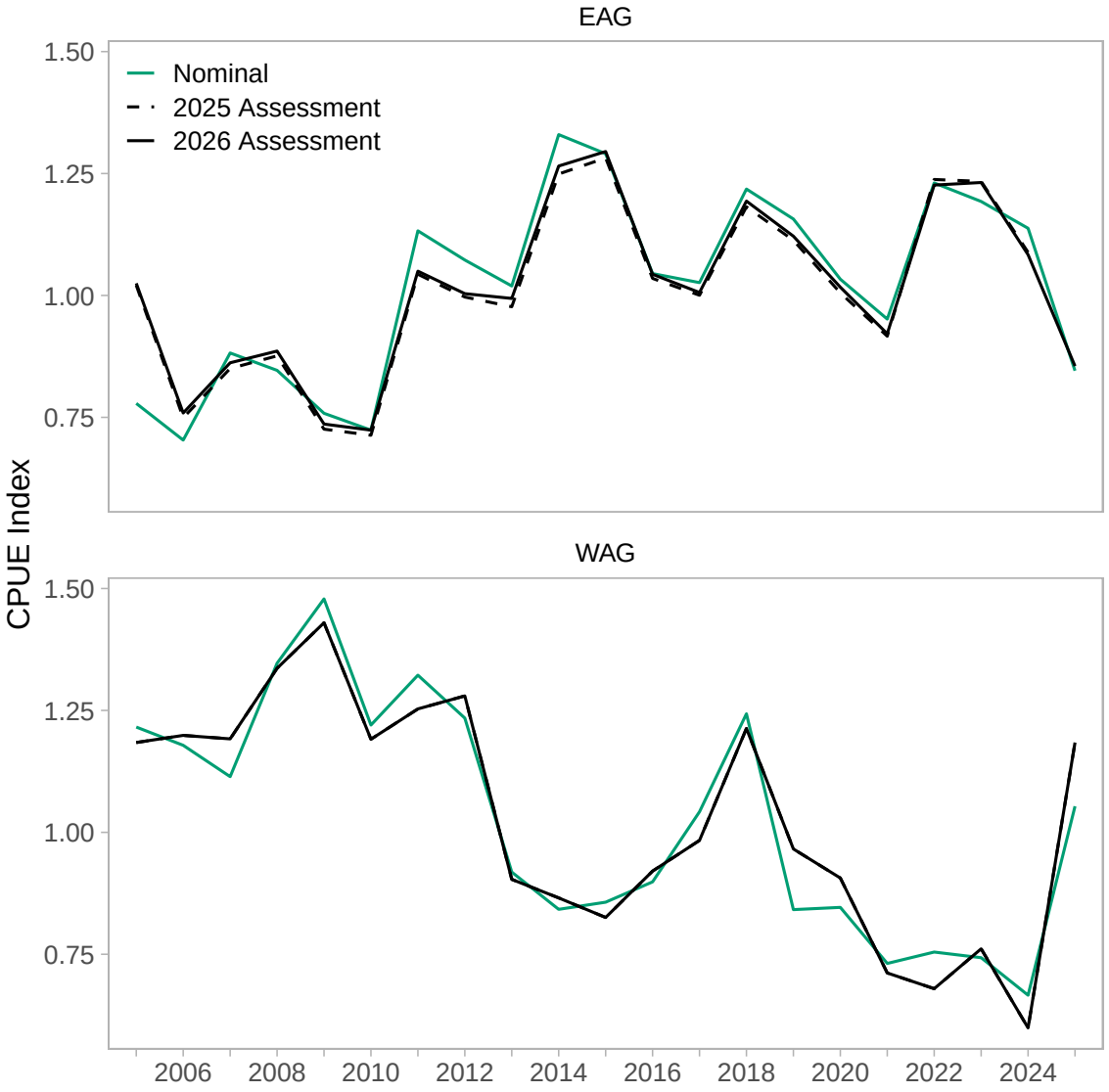

Figure 7: Plot of CPUE index based on the GAM fit to legal CPUE during the post-rationalized period in the WAG.

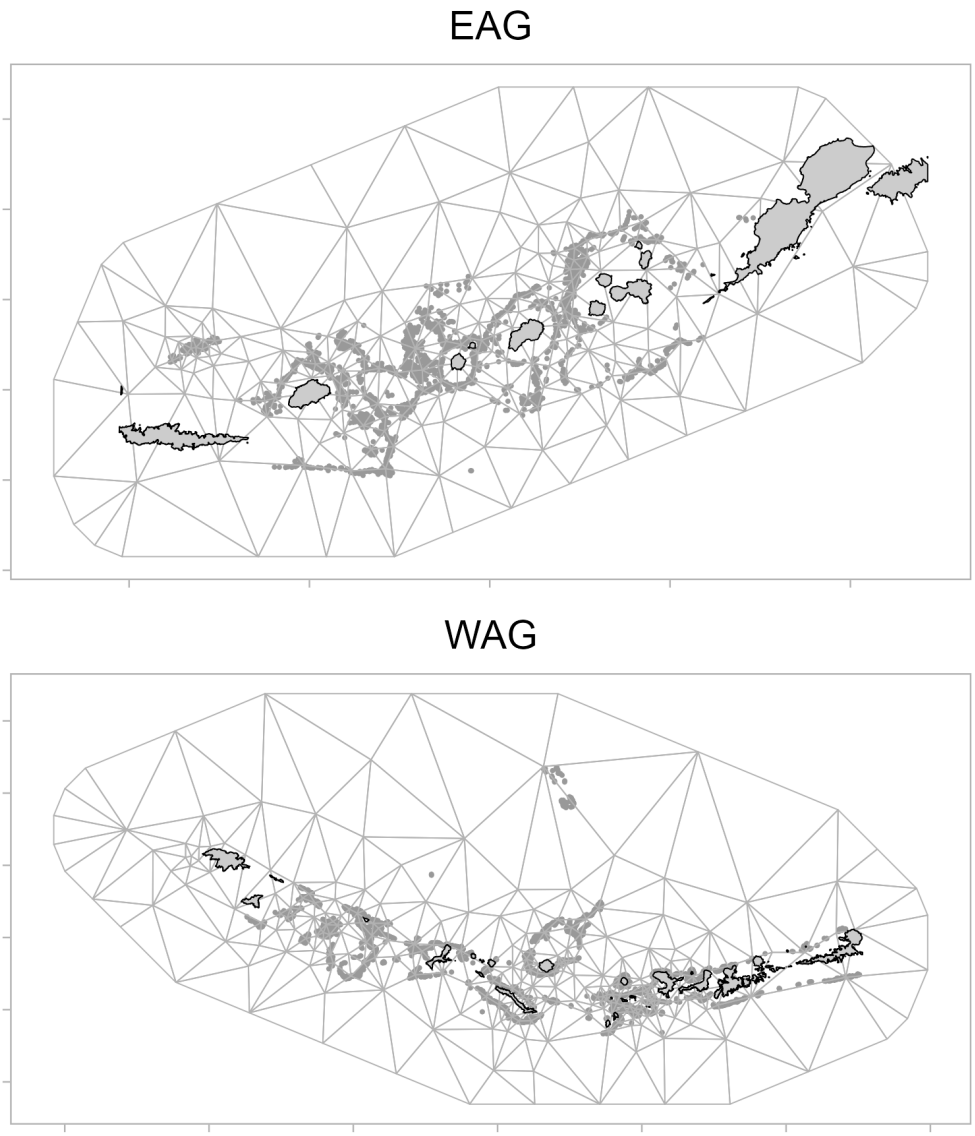

Figure 8: Triangular mesh based on k-means clustering with 150 knots used in spatiotemporal GAMMs for EAG and WAG. Grey points are observer pot locations for which confidentiality waivers have been obtained.

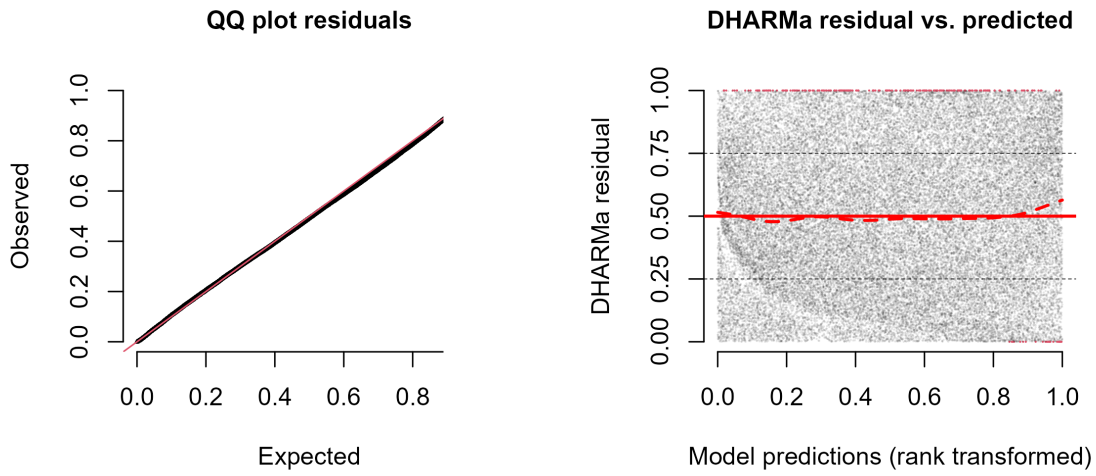

Figure 9: DHARMa residual QQ plots for the final Tweedie spatiotemporal GLMM fit to legal CPUE in the EAG.

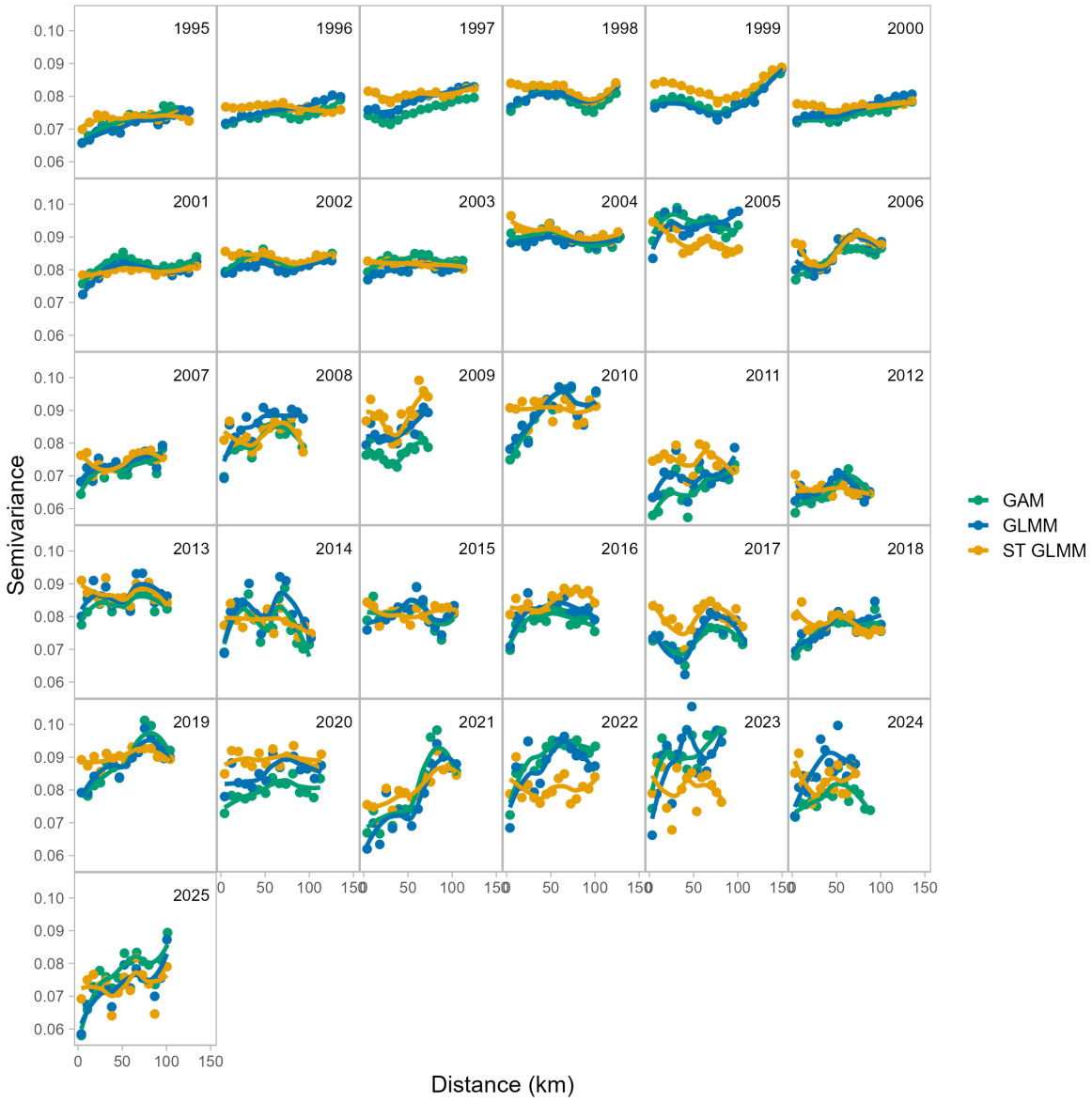

Figure 10: Semivariance of the GAM, GLMM, and spatiotemporal GLMM fit to legal CPUE data in the EAG

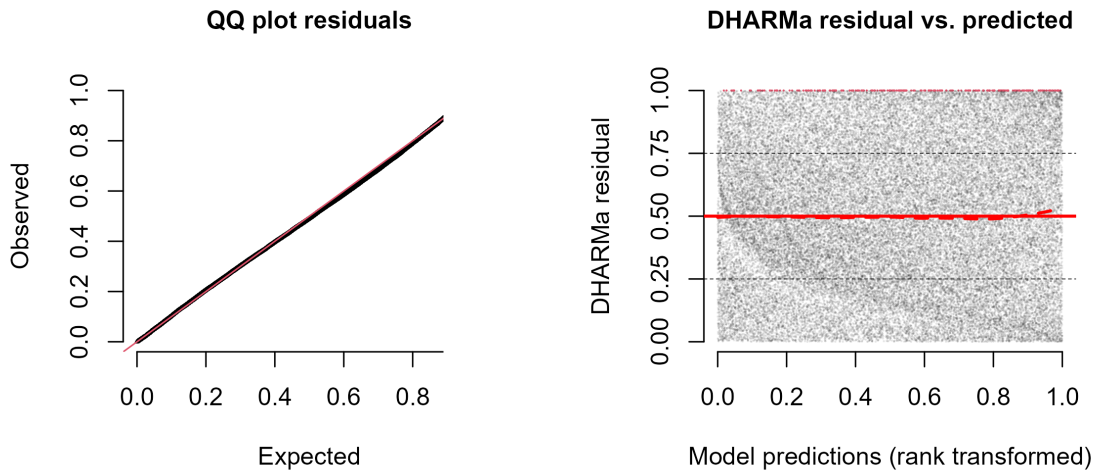

Figure 11: DHARMa residual QQ plots for the final Tweedie spatiotemporal GLMM fit to legal CPUE in the WAG.

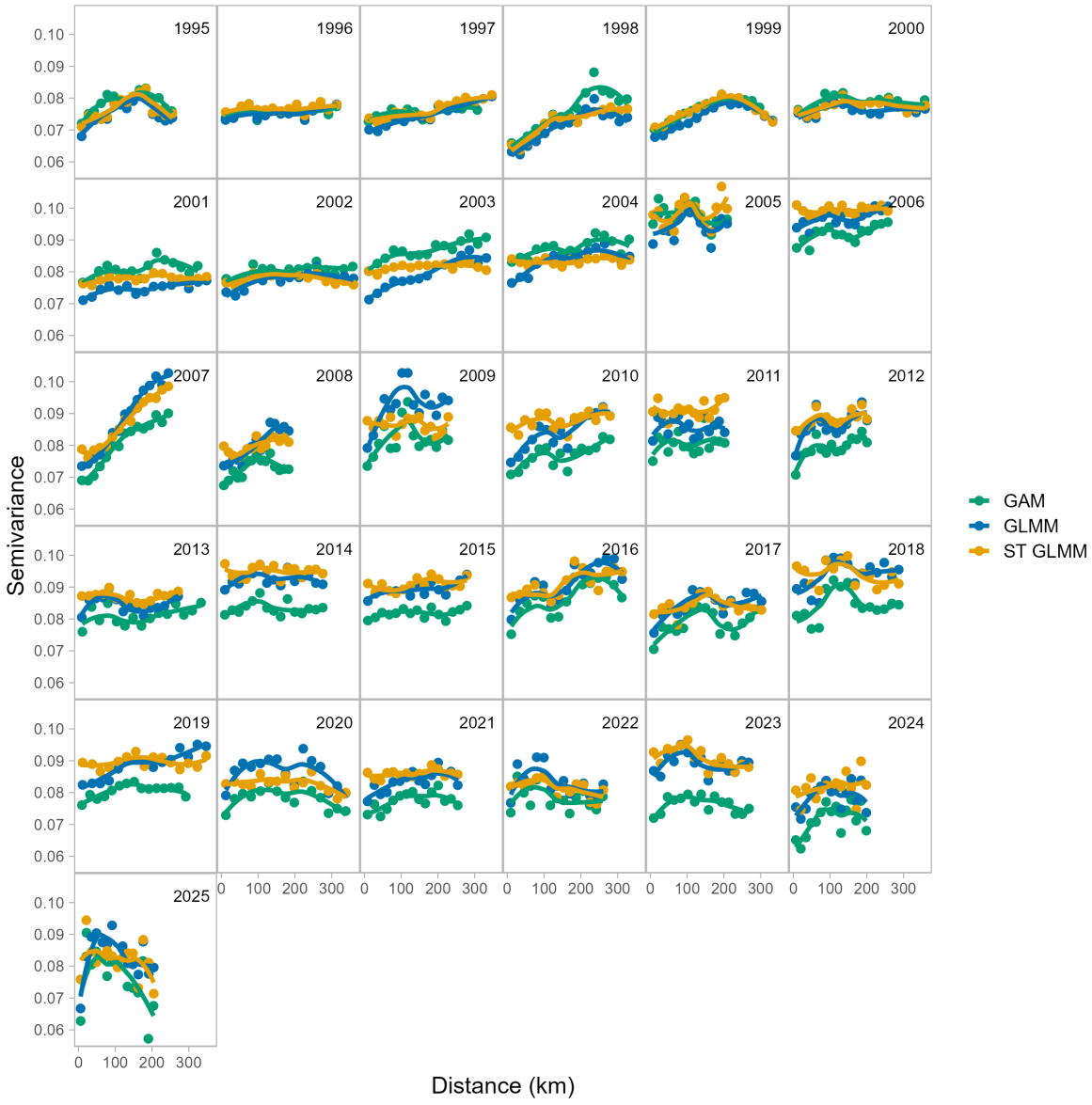

Figure 12: Semivariance of the GAM, GLMM, and spatiotemporal GLMM fit to legal CPUE data in the WAG

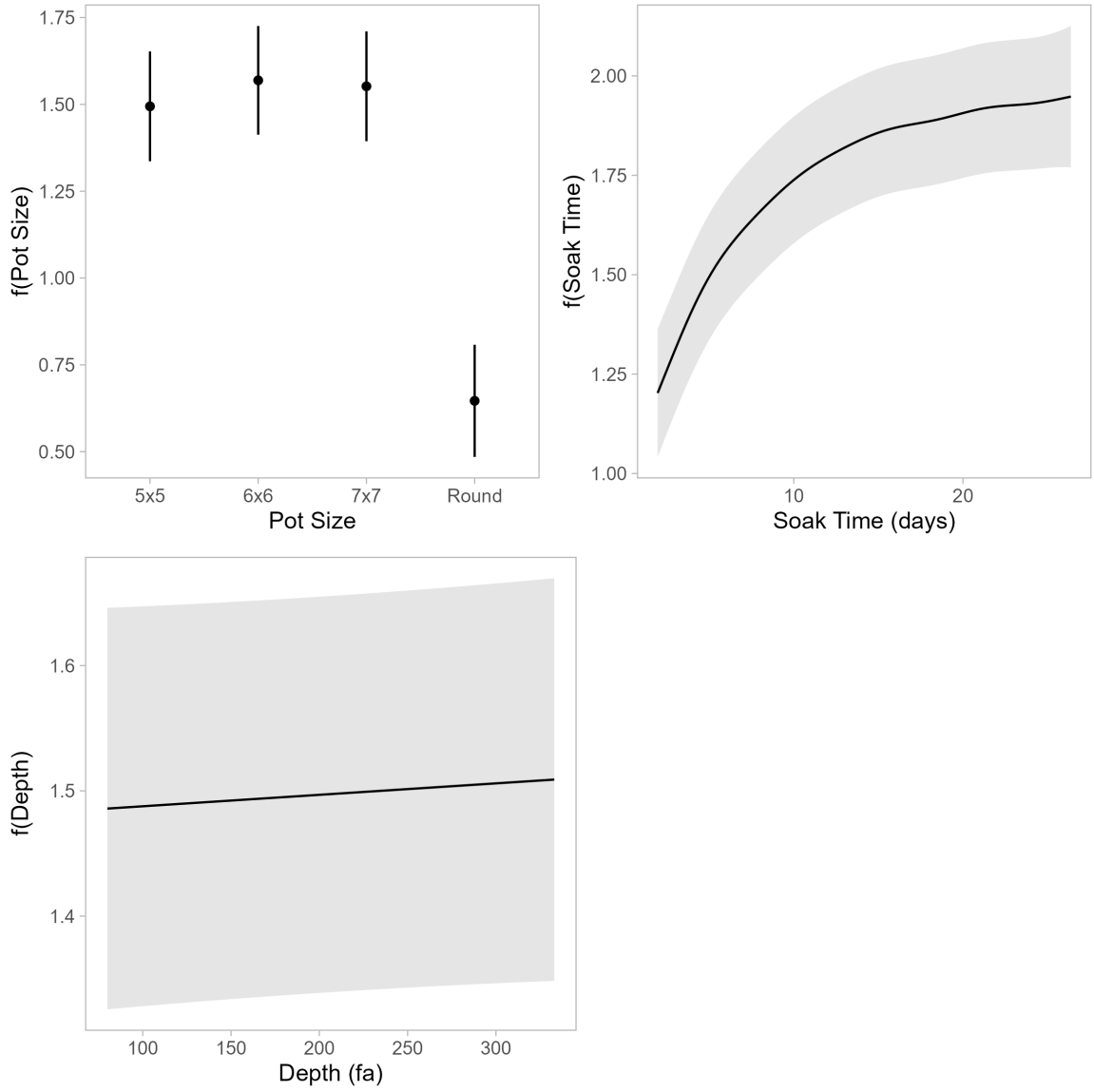

Figure 13: Marginal effects of gear type, soak time, and depth for the spatiotemporal GLMM fit to the EAG.

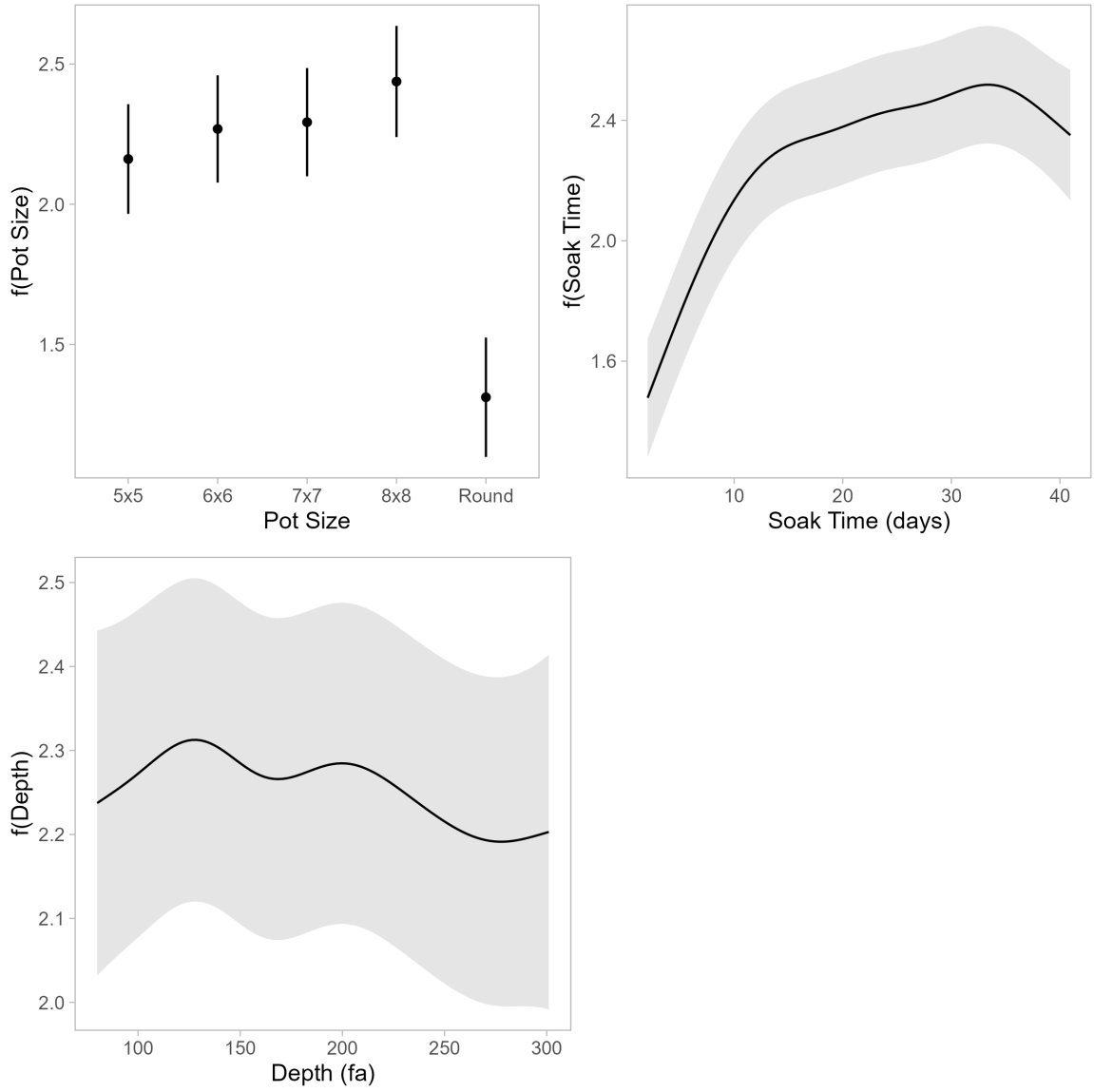

Figure 14: Marginal effects of gear type, soak time, and depth for the spatiotemporal GLMM fit to the WAG.

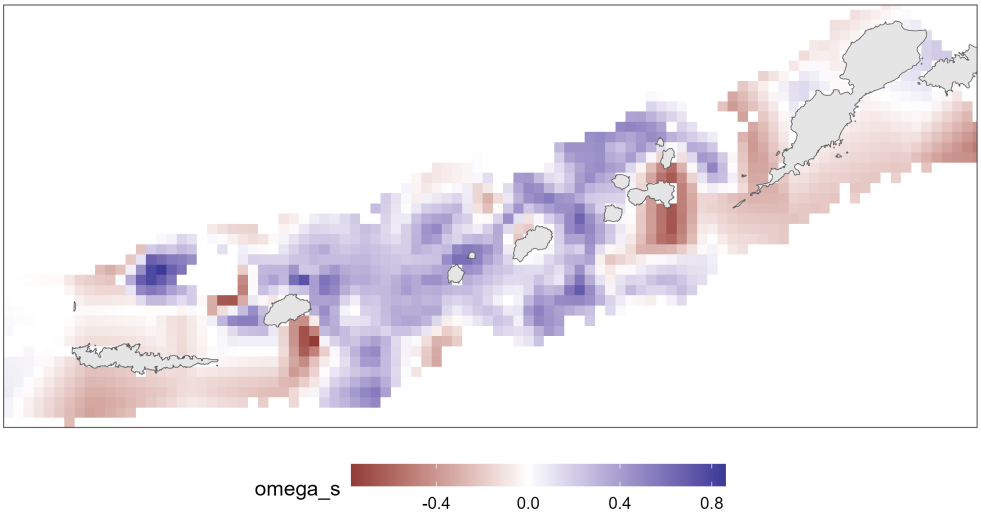

Figure 15: Spatial random effect for the spatiotemporal GLMM fit to the EAG.

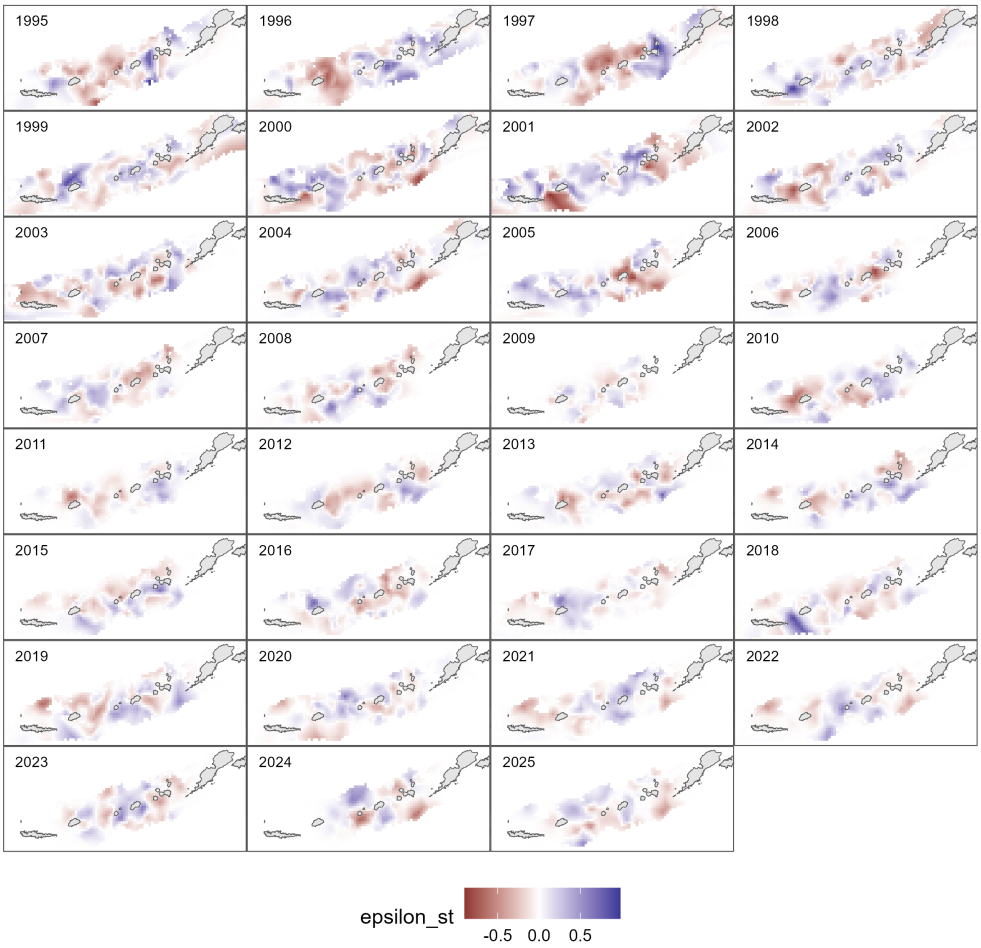

Figure 16: Spatiotemporal random effect for the spatiotemporal GLMM fit to the EAG.

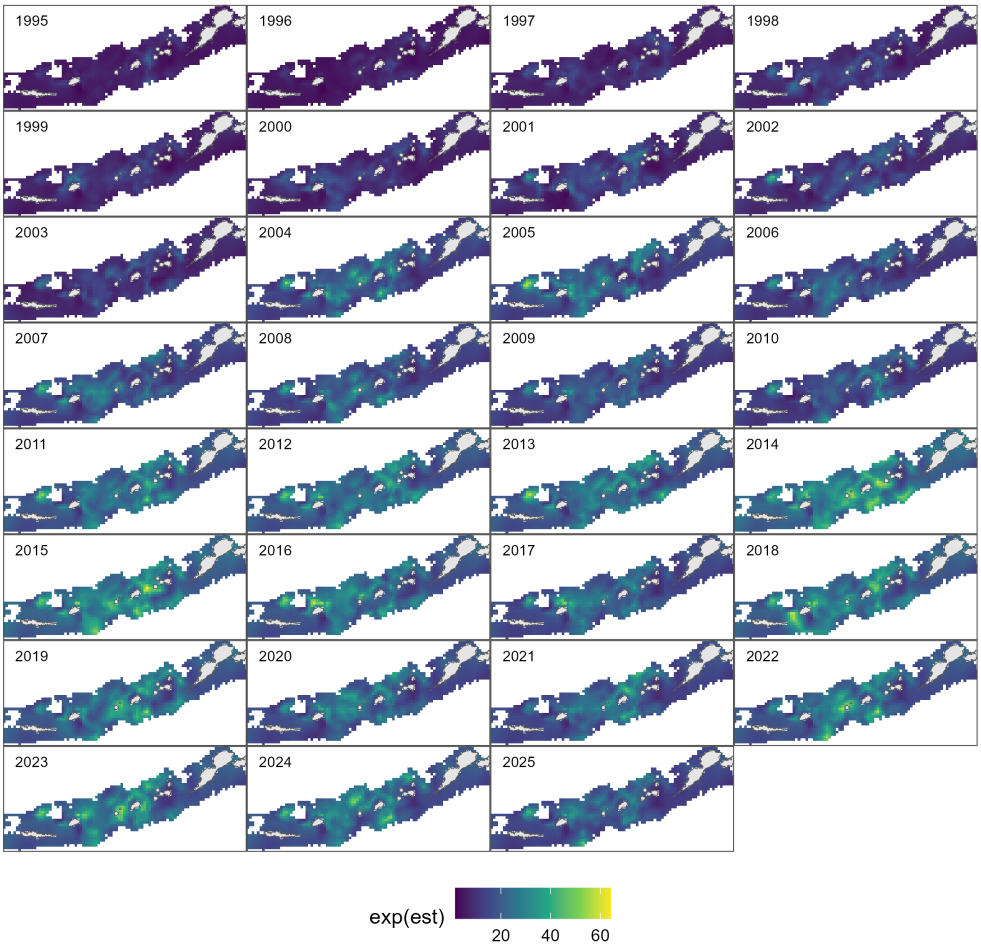

Figure 17: Predicted CPUE for the spatiotemporal GLMM fit to the EAG.

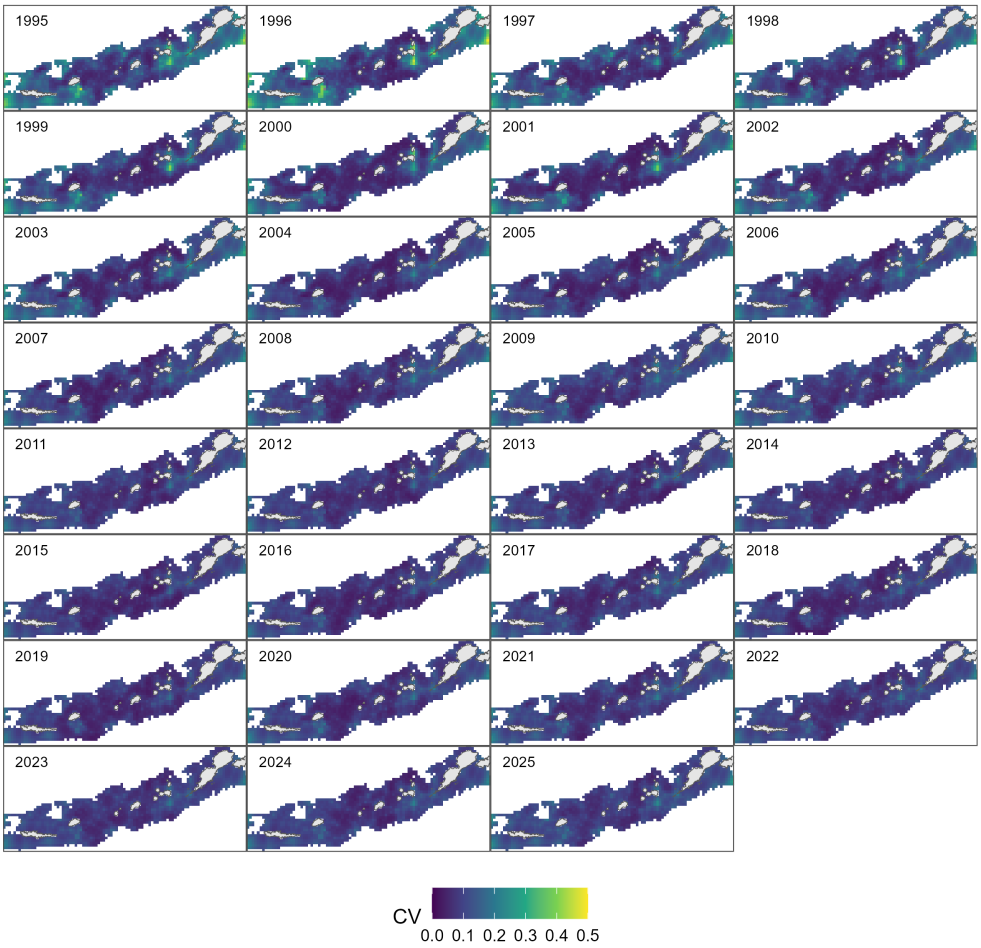

Figure 18: Coefficient of variation on predicted CPUE for the spatiotemporal GLMM fit to the EAG.

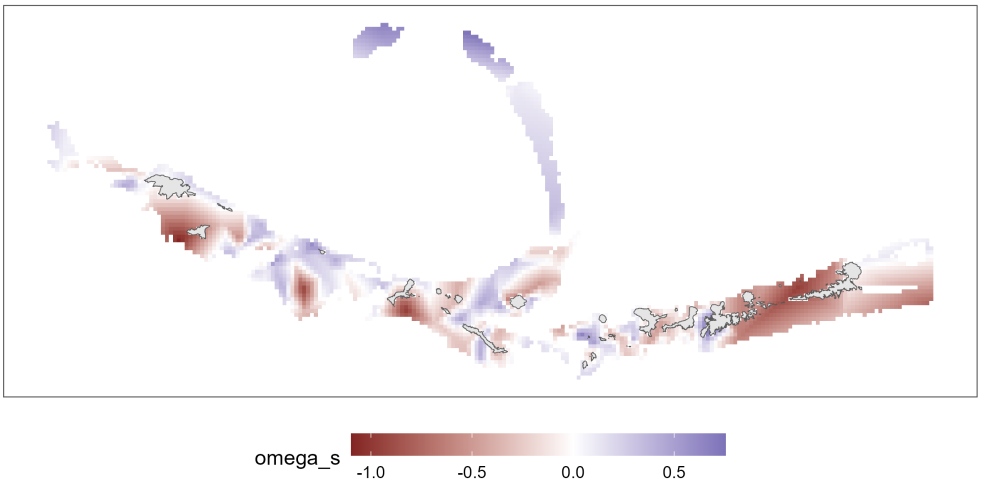

Figure 19: Spatial random effect for the spatiotemporal GLMM fit to the WAG.

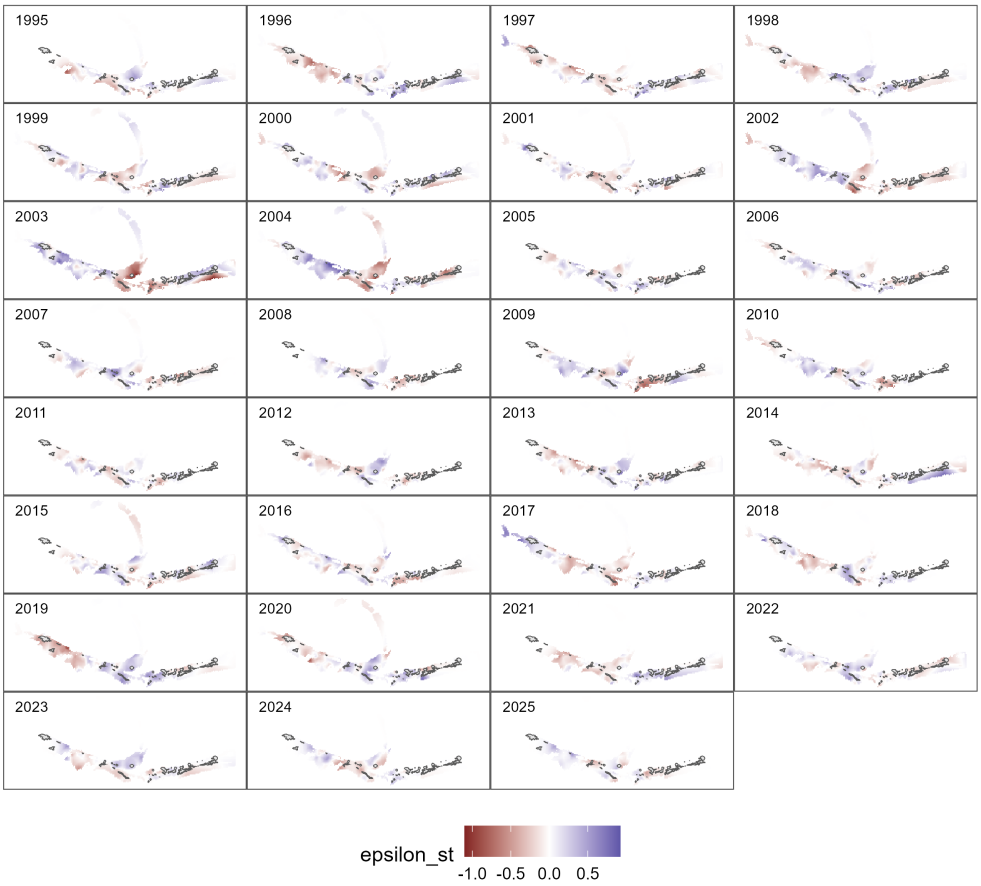

Figure 20: Spatiotemporal random effect for the spatiotemporal GLMM to the WAG.

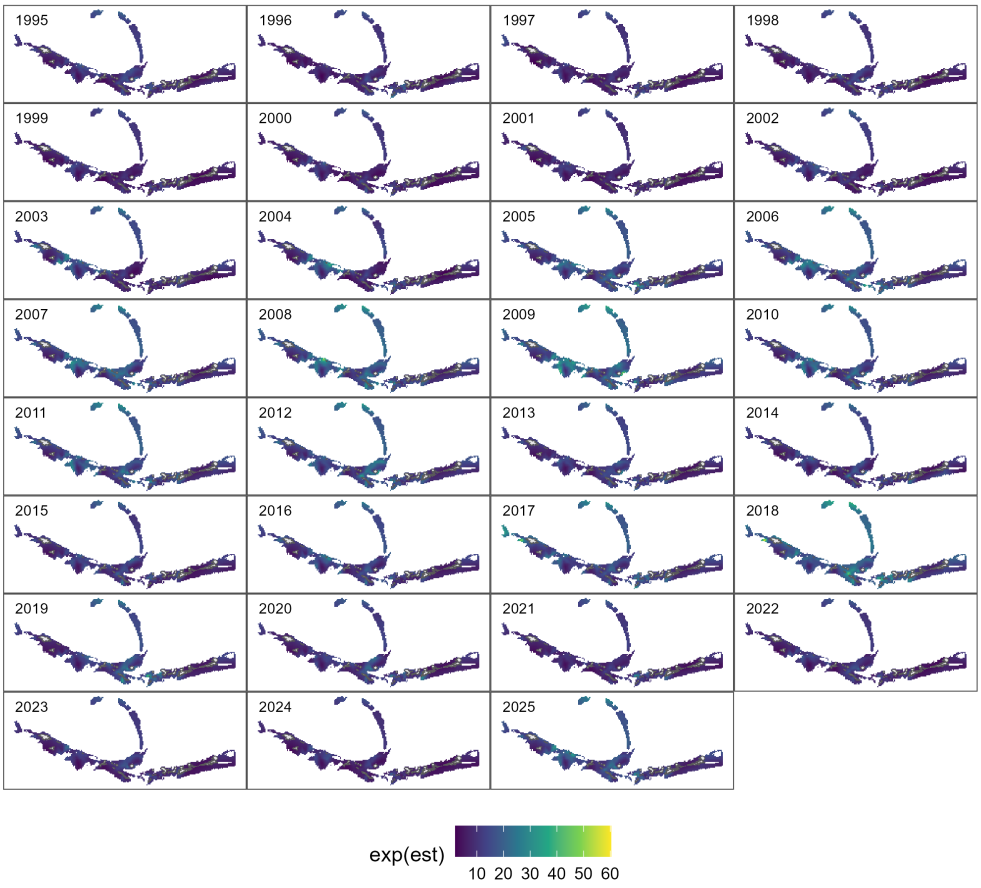

Figure 21: Predicted CPUE for the spatiotemporal GLMM to the WAG.

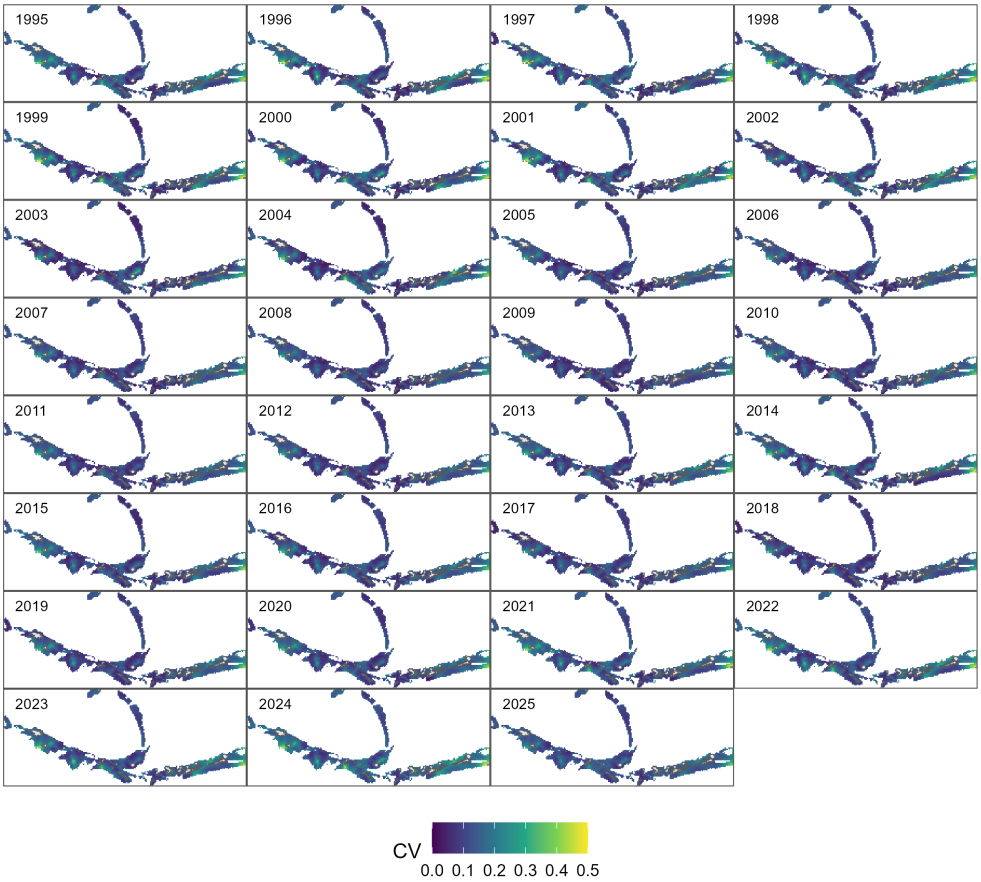

Figure 22: Coefficient of variation on predicted CPUE for the spatiotemporal GLMM fit to the WAG.

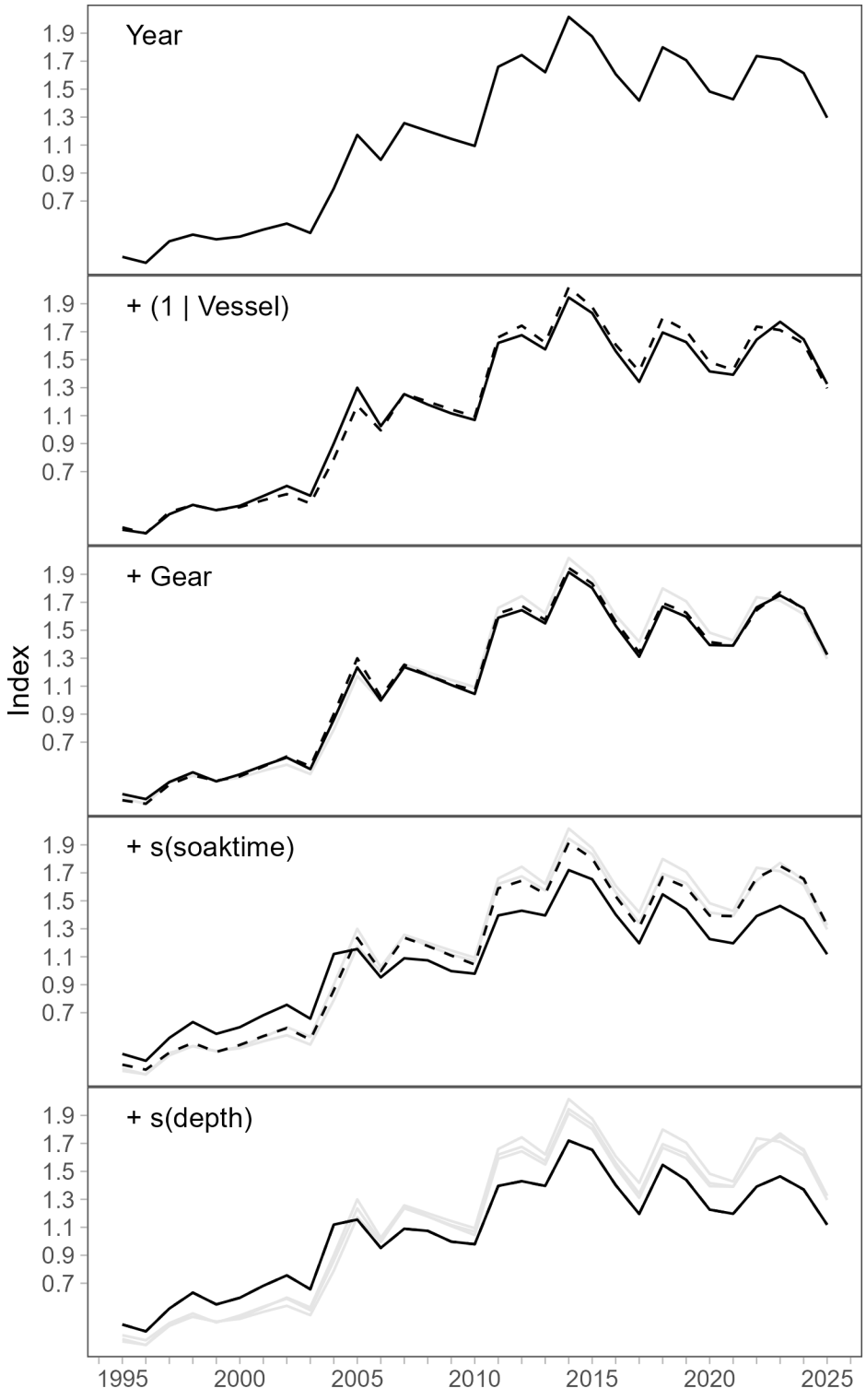

Figure 23: Step plot of CPUE index for the spatiotemporal model fit to the EAG.

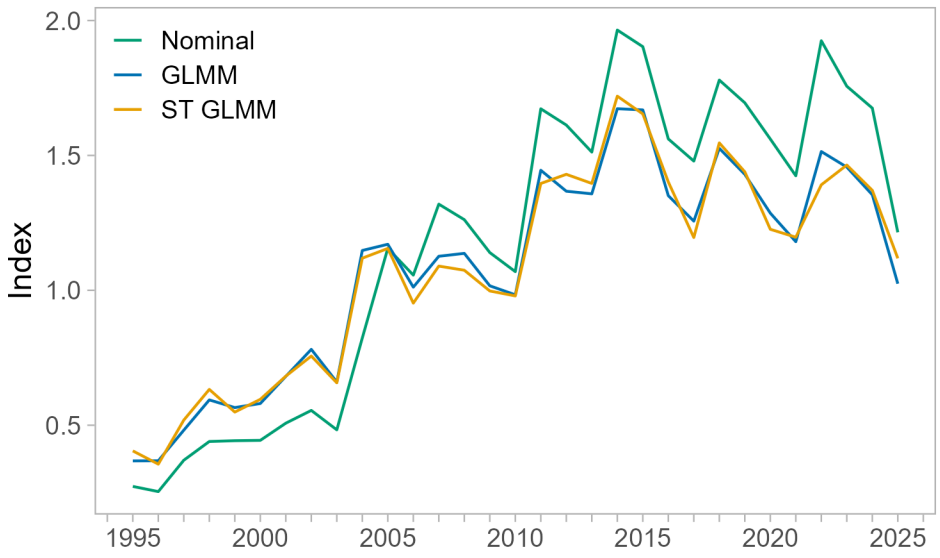

Figure 24: Comparison of nominal CPUE and standardized indices from a GLMM without spatial effects for the EAG.

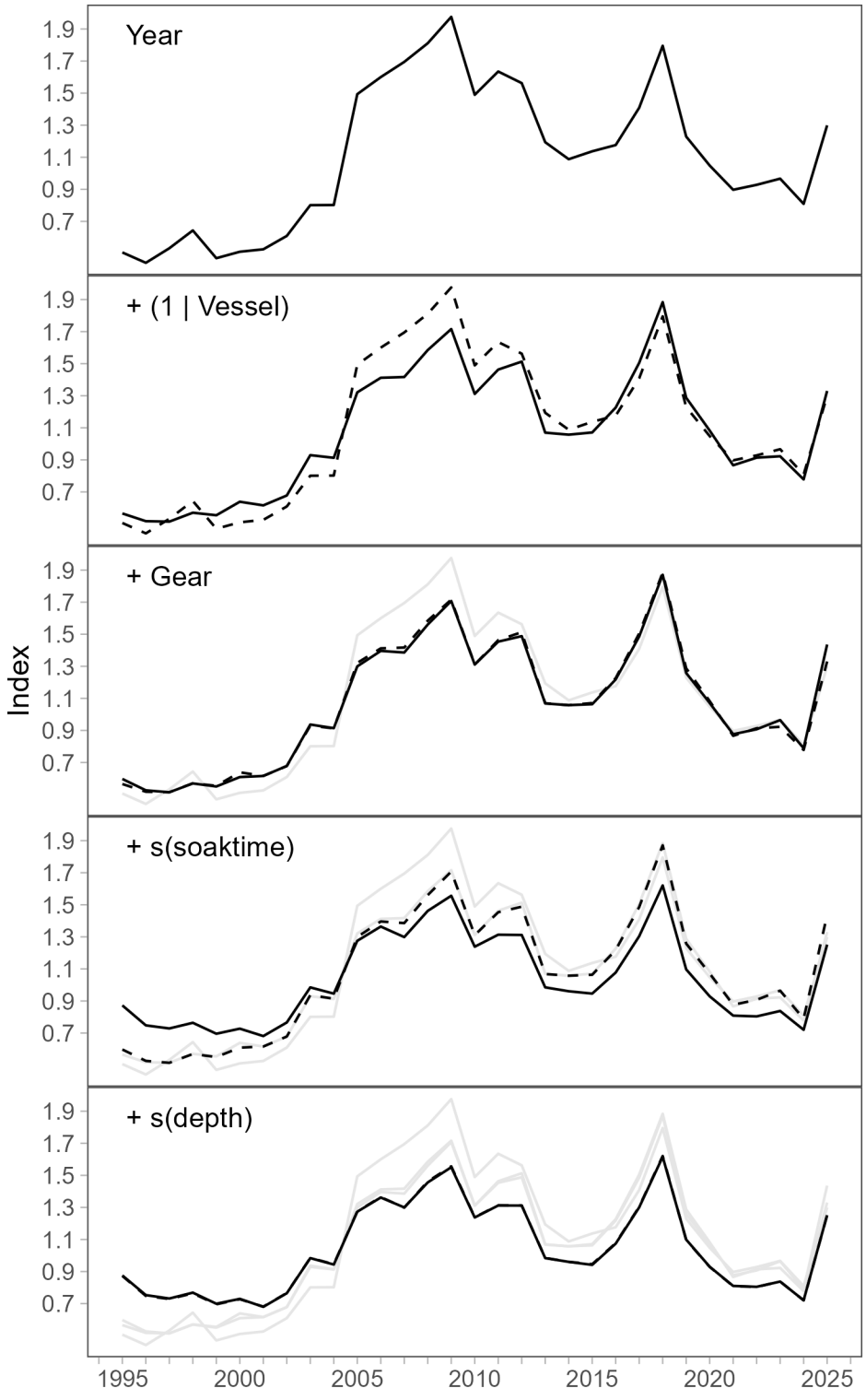

Figure 25: Step plot of CPUE index for the spatiotemporal model fit to the WAG.

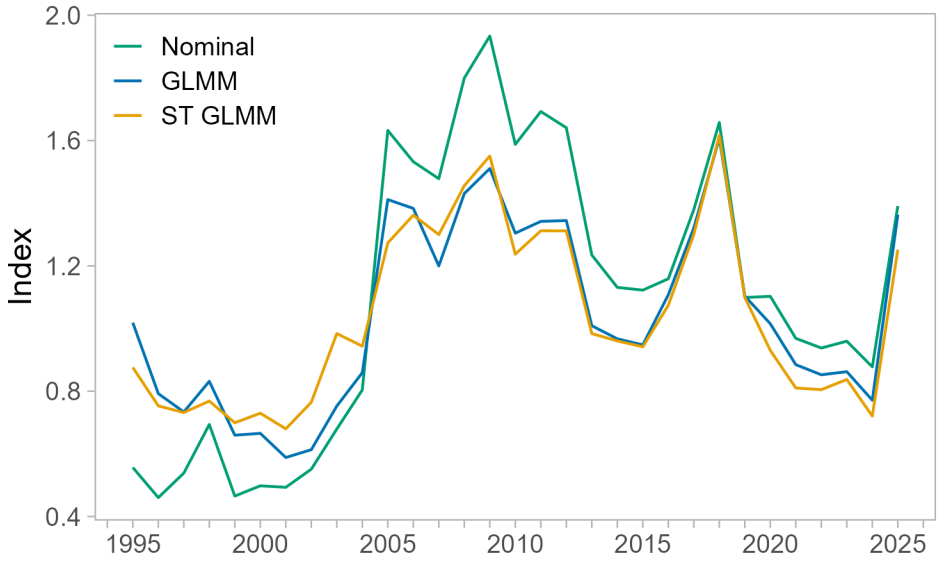

Figure 26: Comparison of nominal CPUE and standardized indices from a GLMM without spatial effects for the WAG.

Literature Cited

Anderson SC. 2025. sdmTMBextra: Extra Functions for Working with 'sdmTMB' Models_. R package version 0.0.4, commit 63f236912e12ce78b5a0529eedf1e11cb93d0a10, https://github.com/pbs-assess/sdmTMBextra.

Anderson, SC, EJ Ward, PA English, LAK Barnett, JT Thorson. 2024. sdmTMB: an R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. bioRxiv 2022.03.24.485545.

Bentley, N, TH Kendrick, PJ Starr, and PA Breen. 2012. Influence plots and metrics: tools for better understanding fisheries catch-per-unit-effort standardizations. ICES Journal of Marine Science, 69: 84-88.

Bishop, J, WN Venables, CM Dichmont, and DJ Sterling. 2008. Standardizing catch rates: is logbook information by itself enough? ICES Journal of Marine Science 65: 255-266.

Bivand, RS. 2022. spdep: Spatial Dependence: Weighting Schemes, Statistics. R Package Version 1.2-1, https://CRAN.R-project.org/package=spdep.

Lindgren F, H Rue, and J Lindström. 2011. An explicit link between Gaussian fields and Gaussian markov random fields: The stochastic partial differential equation approach. Journal of the Royal Statistical Society B 73: 423 - 498.

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

Hijmans RJ, M Barbosa, A Ghosh, and A Mandel. 2025. geodata: Download Geographic Data. R package version 0.6-3, https://github.com/rspatial/geodata.

Hoyle SD, RA Campbell, ND Ducharme-Barth, A Grüss, BR Moore, JT Thorson, L Tremblay-Boyer, H Winker, S Zhou, MN Maunder. 2024. Catch per unit effort modelling for stock assessment: A summary of good practices. Fisheries Research 269: 106860.

Jackson, TM. 2024a. Aleutian Islands golden king crab stock assessment 2024. North Pacific Fishery Management Council, Anchorage, Alaska.

Maunder MN, JT Thorson, H Xu, R Oliveros-Ramos, SD Hoyle, L Tremblay-Boyer, HH Lee, M Kai, SK Chang, T Kitakado, CM Albertsen, CV Minte_vera, CE Lennert-Cody, AM Aires-da-Silva, and KR Piner. 2020. The need for spatio-temporal modeling to determine catch-per-unit effort based indices of abundance and associated composition data for inclusion in stock assessment models. Fisheries Research 229: 105594.

Moran, PA. 1950. Notes on continuous stochastic phenomena. Biometrika 37:17–23.

Siddeek, MSM, J Zheng, and D Pengilly. 2016. Standardizing CPUE from the Aleutian Islands golden king crab observer data. Pages 97–116 in TJ Quinn II, JL Armstrong, MR Baker, J Heifetz, and D Witherell (eds.), Assessing and Managing Data-Limited Fish Stocks. Alaska Sea Grant, University of Alaska Fairbanks, Alaska.

Siddeek, MSM, J Zheng, C Siddon, and B Daly. 2017. Aleutian Islands golden king crab (*Lithodes aequispinus*) model-based stock assessment in spring 2017. North Pacific Fishery Management Council, Anchorage, Alaska.

Siddeek, MSM, T Jackson, B Daly, C Siddon, MJ Westphal, and L Hulbert. 2023. Aleutian Islands golden king crab model scenarios for May 2023 assessment. North Pacific Fishery Management Council, Anchorage, Alaska.

Zimmermann, M, MM Prescott. Passes of the Aleutian Islands: First detailed description. Fisheries Oceanography 30: 280 - 299.