Norton Sound red king crab stock assessment

Appendix B: spatiotemporal model-based abundance index

Caitlin Stern[1,2]
[1]Alaska Department of Fish and Game
[2]caitlin.stern@alaska.gov

November 2026

Summary

The Norton Sound red king crab (*Paralithodes camtschaticus*) stock assessment uses estimated abundance from three fishery-independent surveys as indices of abundance in the stock assessment model: the historical National Oceanic and Atmospheric Administration (NOAA) Norton Sound trawl survey, the NOAA Northern Bering Sea (NBS) trawl survey, and the Alaska Department of Fish and Game (ADF&G) Norton Sound trawl survey. These surveys cover different years and areas, and within a survey the set of locations sampled sometimes varies among years, meaning that annual estimates may not be directly comparable. While the complete lack of overlap in time means that the historical Norton Sound trawl survey cannot be combined with the other two surveys in a single model, the two ongoing surveys have both spatial and temporal overlap. Using a spatiotemporal modeling approach to combine the NBS and ADF&G survey data into a single index of abundance for inclusion in the stock assessment model offers the opportunity to produce directly comparable annual estimates of abundance, potentially improving the reliability of the abundance time series. This document details the development of a spatiotemporal model-based abundance index for use in the Norton Sound red king crab stock assessment model. In accordance with scientific review bodies' requests, two model-based indices are developed: one excluding bottom depth and one including bottom depth. The effects on stock assessment model outcomes of including these indices are evaluated in the main assessment document.

Introduction

Indices of abundance are key inputs to stock assessment models, providing information crucial for estimating stock size (Quinn and Deriso 1999; Patterson et al. 2001). However, design-based methods for producing these indices from fishery-independent surveys can yield biased results when practical constraints lead survey sampling to deviate from survey design. For example, financial resources may not be sufficient for surveys to occur as frequently as planned, or changes in survey priorities may limit vessels to sampling only a subset of survey stations. Spatiotemporal models including spatial and spatiotemporal random effects have the potential to mitigate the effects of inconsistent survey sampling by accounting for the spatiotemporal processes that underlie survey observations (Thorson et al. 2021; Yalcin et al. 2023). Simulation work suggests that index standardization using spatiotemporal models selected through a rigorous evaluation process can reduce the impact of changes in survey effort by providing a consistent method of inferring abundance in missing areas (Yalcin et al. 2023).

In addition to potentially producing biased results when survey sampling changes, design-based indices of abundance can inflate temporal variability by failing to account for spatial dependence (Shelton et al. 2014; Thorson et al. 2015). Due to the importance of indices of abundance in stock assessment models, this can lead to increased uncertainty in outputs used in fishery management (Cao et al. 2017). By accounting for spatial variation in survey catch, as well as incorporating habitat variables, spatiotemporal model-based index

standardization can produce abundance indices with reduced variability compared to design-based indices (Cao et al. 2017, Chen et al. 2024). This can lead to less extreme retrospective patterns for stock assessment models that use model-based rather than design-based indices (Cao et al. 2017, Chen et al. 2024).

The population of red king crab (*Paralithodes camtschaticus*) in Norton Sound, Alaska, supports locally important commercial and subsistence fisheries. The stock assessment for Norton Sound red king crab uses three fishery-independent trawl survey time series in design-based indices of abundance: the NOAA Norton Sound trawl survey (1976 - 1991), the ADF&G Norton Sound trawl survey (1996 - present), and the NOAA NBS trawl survey (2010 - present) (Stern and Palof 2025; Hamazaki 2024). The data sets from these surveys, all of which use a systematic sampling design, cover different years and spatial footprints, and within a survey the set of locations sampled sometimes varies among years (Figure 1). This history of both spatial and temporal variation in survey sampling is one of the factors motivating the development of spatiotemporal model-based indices for this stock, as standardization over both space and time is required.

Initial work on a model-based approach for deriving Norton Sound red king crab survey indices of abundance led to the development of three separate model-based indices of abundance, one for each of the survey time series (Stern 2025). However, replacing the design-based indices of abundance with these model-based indices of abundance in the stock assessment model did not improve key aspects of the stock assessment model, including the retrospective pattern (Stern and Palof 2025). Combining data from the trawl surveys into a single model-based index of abundance has been a long-standing goal for this stock assessment, and seemed the logical next step given the lack of improvement following the incorporation of separate model-based indices. Including data from the historical Norton Sound trawl survey led to failures in spatiotemporal model convergence, likely due to the complete lack of overlap in time between the historical Norton Sound trawl survey and the two ongoing surveys. This led to the current approach of developing spatiotemporal models combining data from the two ongoing surveys, while retaining the design-based estimates from the historical Norton Sound trawl survey as a separate index in the stock assessment model. This document details the development of spatiotemporal model-based abundance indices combining data from the NBS and ADF&G trawl surveys. The effects of these indices on stock assessment model outcomes are evaluated in the main stock assessment document.

Methods

Data sources

I used the R (R Core Team 2026) package **crabpack** (Hennessey 2025) to extract the NOAA NBS trawl survey data (2010 - 2025). I obtained the ADF&G trawl survey data (1996 - 2024) from Jenefer Bell of ADF&G. Survey effort is displayed in Figure 1. The ADF&G and NBS surveys have taken place in the same year four times: 2017, 2019, 2021, and 2023. For both surveys, I calculated crab density at a survey station as the number of male red king crab $\geq$ 64 mm in carapace length (CL) at the station divided by the area swept in km^2.

Both survey station sampling and the set of survey stations at which at least one male red king crab $\geq$ 64 mm in CL was recorded have changed over the time series (Figure 1). Red king crab were consistently detected within a core area in Norton Sound, and were occasionally detected in areas to the north or south of the core area. Exploratory sampling by the ADF&G trawl survey in 2006 around the eastern tip of St. Lawrence Island did not detect red king crab. The ADF&G trawl survey typically sampled at a finer spatial scale and observed greater densities of crab than did the NOAA NBS survey, with the highest crab densities usually occurring near Nome (Figure 1).

Both the ADF&G and NBS trawl surveys recorded bottom depth at sampled survey stations (Figure 2), allowing for exploration of the relationship between bottom depth and crab abundance. The ranges of both station bottom depths sampled and, as noted above, crab abundance per area vary between the surveys, with the ADF&G trawl survey typically sampling shallower depths and observing greater abundances per area than the NOAA NBS survey (Figure 3). Norton Sound is shallow, with bottom depth $<$ 30 m in most of the Sound (Prescott and Zimmerman 2015; Figure 4). For an explanation of how depth information was added

to the prediction grid used to make predictions from models containing depth information, please see the "Predictions" section, below.

Spatiotemporal models

I fit geostatistical GLMMs with spatiotemporally correlated random effects to survey data using the R package **sdmTMB** (Anderson et al. 2022). In this approach, spatial effects are modeled using the stochastic partial differential equation (SPDE) approximation to Gaussian random fields (Lindgren et al. 2011). The rate at which spatial covariance decays with distance is defined by the Matèrn covariance function, a derived parameter of which is the spatial (or Matèrn) range, the distance at which spatial correlation decays such that two points are effectively independent (Anderson et al. 2022). **sdmTMB** uses geostatistical time series data to estimate spatial and spatiotemporal generalized linear mixed effects models (GLMMs). This approach allows for index standardization when the set of stations surveyed is not consistent across years: one can generate a spatial grid that covers the area of interest, predict from the model onto that grid, and sum the predicted abundance to obtain an area-weighted index that is independent of sampling locations (Anderson et al. 2022).

Applying the SPDE approach to modeling random fields requires defining a triangulation mesh of the spatial domain (Anderson et al. 2022). To define the mesh area, I first designated water area by finding the part of a spatial polygon box (defined to encompass all sampled stations in the survey data sets) that did not intersect with the land spatial geometry. I then generated a custom mesh over this area using the **INLA** package (Lindgren and Rue 2015) with a maximum triangle edge length of 50 km for the inner domain, in which data points occur, and 100 km for the outer domain, in which data points do not occur. I used this custom mesh to construct a triangulation mesh for the combined NBS and ADF&G survey data set using the **sdmTMB** function make_mesh(), a wrapper for the triangulation mesh functions in the **fmesher** package (Lindgren 2023), and the **sdmTMBextra** (Anderson 2025) function add_barrier_mesh(). The add_barrier_mesh function permitted incorporation of correlation barriers based on the Norton Sound coastline (Figure 5). The barrier mesh approach effectively excluded from the mesh both vertices on land and vertices outside of the designated spatial polygon (Figure 6).

For all models, I estimated spatiotemporal random fields as independent and identically distributed (IID). The spatial random fields are estimated for each time slice, the period of which is specified in the model (e.g., time = year). Estimating spatiotemporal random fields as IID is likely most appropriate for the standardization of survey indices intended to be used in stock assessment models as this approach minimizes the estimation covariance among years, which is usually ignored in stock assessment models (Thorson et al. 2020, Chen et al. 2024). I compared models in which the only fixed effects were year and survey (NBS or ADF&G), both specified as factors, with models that also included survey station depth (m), as variation in the bottom depth of stations surveyed could influence estimated abundance. I explored three approaches to including depth in the model: as a covariate, in a linear interaction with survey, and in a smooth (nonlinear) interaction with survey. The model formulae for these methods of including depth are captured by abundance per area ~ year + survey + depth, abundance per area ~ year + survey + depth * survey, and abundance per area ~ year + survey + s(depth, by = survey), respectively. I ran models with the Tweedie, delta gamma, and delta lognormal distribution families.

Model evaluation and diagnostics

After fitting models, I checked model convergence using the **sdmTMB** sanity() function. Models were considered converged if they met the following criteria evaluated by the sanity() function: the non-linear minimizer suggested successful convergence, the Hessian matrix was positive definite, no extreme or very small eigenvalues were detected, all gradients with respect to fixed effects were < 0.001, no fixed-effect standard errors were NA, no standard errors looked unreasonably large, no sigma parameters were < 0.01; no sigma parameters were > 100, and the range parameter was of a reasonable size.

Following sanity checks, I calculated residuals using the **DHARMa** package (Hartig 2024) and examined diagnostics included in the package. I plotted the DHARMa residuals over space and time to visualize potential patterns of autocorrelation.

I evaluated model predictive skill (the predictive ability of the model for new observations; Anderson *et al.* 2024) using the cross validation function sdmTMB_cv(). This function measures model predictive skill by holding out subsets of the data in turn and using each as a test set. These subsets of data are termed "folds". Following the practices of similar investigations (e.g., Vihtakari *et al.* 2026), I compared models by performing cross validation with 10 randomly arranged folds for each model before extracting the model's total predicted out-of-sample negative log-likelihood (NLL). I calculated Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) based on the predictive values; smaller RMSE and MAE values indicate better model fit.

Predictions

To generate a prediction grid, I pulled the NBS trawl survey stratum for Norton Sound red king crab from the **akgfmaps** package (AFSC 2026; Figure 7) using the get_crab_strata() function, then produced a grid with a resolution of 5 km^2 that matched the Norton Sound red king crab survey stratum area (Figure 8). The NBS survey stratum covers a smaller area than the spatial mesh used in model fitting and includes nearly all survey stations at which crab have been observed; some tow locations that were surveyed prospectively in 2006 and 2008 fall outside the stratum but almost all of these tow locations had no observed crab and excluding them from the prediction area is not expected to impact estimated abundance (Figure 9). I used the **sdmTMB** predict() function to generate predicted crab abundance from each model over the prediction grid. In order to make predictions from models containing depth, I added depth information to the prediction grid using a georeferenced TIFF file from the Southern Alaska Coastal Relief Model (Lim et al. 2011; Figure 10), matching to prediction grid locations using a nearest neighbor spatial join (Figure 11). Because survey identity was included as a fixed effect in all models, survey identity had to be included in the prediction grid; I used the NBS survey as the survey factor level in the prediction grid because this survey covered a larger spatial area than the ADF& survey. Future work could evaluate the impact of instead using the ADF&G survey factor level in the prediction grid.

Indices of abundance

I used the **sdmTMB** get_index() function to calculate total abundance estimates and standard errors from each model, then plotted these model-based indices with the design-based indices currently used in the stock assessment model for visual comparison.

Results

Model diagnostics and selection

For each of the three distribution families (Tweedie, delta gamma, and delta lognormal), I ran a base model with year and survey as fixed effects. I then explored including depth using the following model formulae, where density is defined as number of crab per km^2: density ~ year + survey + depth, including depth as a separate fixed effect; density ~ year + survey + depth * survey, including an interaction between survey and depth and assuming that the effect of depth is linear; and density ~ year + survey + s(depth, by = survey), including an interaction between survey and depth and allowing the effect of depth to be nonlinear. This yielded a total of 12 models.

The only model that did not converge, showing a non-positive-definite Hessian, was the delta gamma family model with the linear effect of depth in the depth by survey interaction. This model was not considered further. All three of the models with a nonlinear effect of depth in the depth by survey interaction produced a warning in the sanity check, resulting from a large standard deviation for the smooth interaction term for the ADFG survey factor level. This large standard deviation was not unexpected given the substantial difference in scale between the ADFG and NBS surveys and the lack of complete overlap in the depth ranges covered by the two surveys, so these models were retained for further analysis. Examination of the conditional effect of depth by survey using the delta gamma family model with depth as a separate fixed effect showed a nonlinear effect of depth on abundance per area that differed between the surveys, supporting exploration of including a smooth of depth by survey (Figure 12).

Comparing diagnostics using DHARMa residuals among the models showed that the delta gamma family models performed best: all 3 models showed no evidence of significant deviation in the outlier, dispersion, or Kolmogorov-Smirnov tests, although they did show quantile deviations (Figures 13 - 15). The Tweedie and delta lognormal family models performed similarly to each other, with all 8 showing evidence of significant deviation in the dispersion test, quantile deviations, and no evidence of significant deviation in the outlier or Kolmogorov-Smirnov tests (Figures 20 - 19). The spatial distributions of residuals did not appear to show concerning patterning (Figures 24 - 34).

Because the SSC and CPT requested a comparison of stock assessment model runs using a model-based index with depth excluded versus included, we sought to select two models to use for producing the indices. Cross-validation analysis for all 4 of the delta lognormal distribution family models produced non-positive-definite Hessian matrices in at least 5 of the 10 folds, so these models were excluded from further consideration. Among the models for which cross-validation analysis was successful, the model using the delta gamma distribution family with year, survey, and a smooth of depth by survey had the smallest NLL value, indicating that this model had the best out-of-sample predictive skill of the models evaluated; this model also had the smallest MAE and RMSE values (Figure 35). We selected this model to generate abundance estimates for the model-based index with depth included. Of the 2 models with depth excluded, the delta gamma distribution family model had a NLL only slightly larger than the Tweedie distribution family model (4132 vs. 4131) and had substantially lower MAE and RMSE values, so we selected this model to generate abundance estimates for the model-based index with depth excluded.

The Matèrn range for the presence/absence component of the best-fitting model was 165 km, indicating spatial autocorrelation in crab presence/absence over a large distance, while the range was smaller for the positive catch component of the model, indicating spatial autocorrelation in crab abundance over a shorter distance (Table 1). Standard deviations were higher for spatial fields (σ_ω) than for spatiotemporal fields (σ_ϵ) for the presence/absence model components, indicating that spatial variability was larger than spatiotemporal variability for crab presence/absence, while the opposite was true for the positive catch component of the model, indicating that spatial variability was smaller than spatiotemporal variability for crab abundance (Table 1). In the Tweedie distribution models, the spatial ranges were more similar to the ranges for the positive catch components of the delta gamma distribution models than to the ranges for the presence/absence components (Table 2). The fixed effect coefficient estimate for the NBS survey was positive in the delta gamma distribution models but negative in the Tweedie distribution models, while the estimate for depth was positive whether depth was included as a separate fixed effect or in an interaction with survey (Table 3).

Model predicted abundance

The spatial distributions of model-predicted crab presence/absence and/or abundance for the models for which cross-validation analysis was successful appeared broadly concordant with spatial abundance patterns observed in the surveys, with the highest crab densities typically predicted for the area near or to the south of Nome (Figures 36 - 45).

Model-based abundance indices compared to the design-based abundance indices

Abundance estimates from the design-based indices are shown alongside model-based estimates in Figures 46 and 47. For the models using the delta gamma distribution family, the spatiotemporal model-based abundance estimates tended to be similar in scale to the design-based abundance estimates; including depth led to slightly greater estimated abundance (Figure 46). For the models using the Tweedie distribution family, including depth also typically led to greater estimated abundance (Figure 47). The relatively small impact on model-estimated abundance of including depth in the delta gamma family models could suggest that most of the variation explained by depth was encompassed by the survey effect, leaving a smaller amount of additional variation explained by the depth or depth interaction with survey term. However, the improved out-of-sample predictive skill and RMSE and MAE values of the models that included depth suggest that accounting for this additional variation did improve model performance. The model-based abundance estimates from the models using the delta gamma distribution family showed reduced interannual variability compared to the design-based estimates, a finding consistent with previous work (Cao et al. 2017, Chen et al. 2024).

Conclusions

The spatiotemporal model-based abundance indices for Norton Sound red king crab account for differences in survey stations sampled among years and across space, producing consistent time series that include data from both the NOAA NBS and ADF&G trawl surveys. Two model-based abundance indices, one excluding and one including depth, were evaluated for inclusion in the 2026 Norton Sound red king crab stock assessment model in the main document.

Acknowledgements

I thank the members of the North Pacific Fishery Management Council Crab Plan Team, as well as Chris Siddon, Alex Reich, Aaron Lambert, Tyler Jackson, and Andrew Whetten, all of ADF&G, for their helpful feedback on this work. I thank Lewis Barnett and Sean Hardison for providing invaluable advice on the development of the spatiotemporal models.

References

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, Ward EJ, English PA, Barnett LAK (2022) 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*. doi:10.1101/2022.03.24.485545.

Anderson SC, Ward EJ, English PA, Barnett LAK, Thorson JT (2024) Cross-validation for model evaluation and comparison. Retrieved from https://pbs-assess.github.io/sdmTMB/articles/cross-validation.html.

Cao J, Thorson JT, Richards RA, Chen Y (2017) Spatiotemporal index standardization improves the stock assessment of northern shrimp in the Gulf of Maine. Canadian Journal of Fisheries and Aquatic Sciences 74: 1781–1793 doi:10.1139/cjfas-2016-0137.

Cacciapaglia C, Brooks EN, Adams CF, Legault CM, Perretti CT, Hart D (2024) Developing workflow and diagnostics for model selection of a vector autoregressive spatiotemporal (VAST) model in comparison to design-based indices. Fisheries Research 275:107009 doi:10.1016/j.fishres.2024.107009.

Chen J, Gao J, Zhang F (2024) Spatiotemporal model improves survey indices for witch flounder stock assessment in the Grand Banks. Canadian Journal of Fisheries and Aquatic Sciences 81:459–487 doi:10.1139/cjfas-2023-0101.

DeFilippo L, Kotwicki S, Barnett L, Richar J, Litzow MA, Stockhausen WT, Palof K (2023) Evaluating the impacts of reduced sampling density in a systematic fisheries-independent survey design. Frontiers in Marine Science 10:1219283 doi:10.3389/fmars.2023.1219283.

Hamazaki T (2024) Norton Sound red king crab stock assessment for the fishing year 2024. In: Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands: 2024 Final Crab SAFE. North Pacific Fishery Management Council, Anchorage AK. PDF.

Hartig F (2024) DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.7, https://CRAN.R-project.org/package=DHARMa.

Hennessey S (2025) crabpack: Calculate Bering Sea Crab CPUE, Abundance, and Biomass. R package version 1.0.0, commit 65322a20b78cf157b69587eb3039e30ec8ee114d, https://github.com/AFSC-Shellfish-Assessment-Program/crabpack.

Lim E, Eakins BW, Wigley R (2011) Coastal Relief Model of Southern Alaska: Procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-43.

Lindgren F, Rue H (2015) Bayesian spatial modelling with R-INLA. Journal of Statistical Software 63:1-25. doi:10.18637/jss.v063.i19.

Lindgren F, Rue H, Lindström J (2011) An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Statistical Methodology Series B 73:423-498. doi:10.1111/j.1467-9868.2011.00777.x

Moran PA (1950) Notes on continuous stochastic phenomena. Biometrika 37:17–23. doi:10.1093/biomet/37.1-2.17.

Patterson K, Cook R, Darby C, Gavaris S, Kell L, Lewy P, Mesnil B, Punt A, Restrepo V, Skagen DW, Stefánsson G (2001) Estimating uncertainty in fish stock assessment and forecasting. Fish and Fisheries 2:125-157. doi:10.1046/j.1467-2960.2001.00042.x.

Prescott MM, Zimmermann M (2015) Smooth sheet bathymetry of Norton Sound. NOAA Technical Memorandum NMFS-AFSC-298. doi:10.7289/V5V69GJ9.

Quinn TJ, Deriso RB (1999) Quantitative fish dynamics. Oxford University Press, New York, New York, USA.

R Core Team (2026) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Shelton AO, Thorson JT, Ward EJ, Feist BE (2014) Spatial semiparametric models improve estimates of species abundance and distribution. Canadian Journal of Fisheries and Aquatic Sciences 71:1655–1666 doi:10.1139/cjfas-2013-0508.

Stern (2025) Spatiotemporal model-based survey indices of abundance for Norton Sound red king crab. North Pacific Fishery Management Council, Anchorage AK. PDF.

Stern and Palof (2025) Proposed models for the 2025 Norton Sound red king crab stock assessment. North Pacific Fishery Management Council, Anchorage AK. PDF.

Thorson JT, Shelton AO, Ward EJ, Skaug HJ (2015) Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES Journal of Marine Science 72:1297-1310. doi:10.1093/icesjms/fsu243.

Thorson JT, Cunningham CJ, Jorgensen E, Havron A, Hulson P-JF, Monnahan CC, von Szalay P (2021) The surprising sensitivity of index scale to delta-model assumptions: recommendations for model-based index standardization. Fisheries Research 233:105745. doi:10.1016/j.fishres.2020.105745.

Vihtakari M, Aune M, Assmann KM, Howell D, Hallfredsson EH, Ward EJ (2026) Demersal distribution and fisheries independent trends of beaked and golden redfish in the Barents and Norwegian Seas. ICES Journal of Marine Science 83:fsag009 doi:10.1093/icesjms/fsag009.

Yalcin S, Anderson SC, Regular PM, English PA (2023) Exploring the limits of spatiotemporal and design-based index standardization under reduced survey coverage. ICES Journal of Marine Sciences 80:2368-2379. doi:10.1093/icesjms/fsad155.

Tables

Table 1: Spatial parameter estimates of the delta gamma distribution family models. Delta (hurdle) models include one component that models presence/absence and one component that models positive catch (abundance); results from both components are shown. Terms include the Matèrn range (the distance at which spatial correlation decays such that two points are effectively independent), the standard deviation for spatial fields (σ_ω), and the standard deviation for spatiotemporal fields (σ_ϵ).

Distribution	Effects	Component	Term	Estimate	SE
DG	year, survey	presence/absence	range	113.9287	30.57
DG	year, survey	presence/absence	sigma_O	1.7647	0.35
DG	year, survey	presence/absence	sigma_E	0.3475	0.17
DG	year, survey	positive catch	range	38.2545	8.02
DG	year, survey	positive catch	phi	1.8552	0.15
DG	year, survey	positive catch	sigma_O	0.6000	0.11
DG	year, survey	positive catch	sigma_E	0.6857	0.07
DG	year, survey, depth	presence/absence	range	125.1462	32.08
DG	year, survey, depth	presence/absence	sigma_O	1.9225	0.37
DG	year, survey, depth	presence/absence	sigma_E	0.3177	0.18
DG	year, survey, depth	positive catch	range	42.4174	8.99
DG	year, survey, depth	positive catch	phi	1.8539	0.15
DG	year, survey, depth	positive catch	sigma_O	0.5955	0.11
DG	year, survey, depth	positive catch	sigma_E	0.6711	0.07
DG	year, survey, s(depth, by=survey)	presence/absence	range	164.6761	56.17
DG	year, survey, s(depth, by=survey)	presence/absence	sigma_O	2.1307	0.50
DG	year, survey, s(depth, by=survey)	presence/absence	sigma_E	0.1738	0.40
DG	year, survey, s(depth, by=survey)	presence/absence	sd__s(depth):as.factor(survey.f)ADFG	2.4716	
DG	year, survey, s(depth, by=survey)	presence/absence	sd__s(depth):as.factor(survey.f)NBS	0.0006	
DG	year, survey, s(depth, by=survey)	positive catch	range	41.5902	9.16
DG	year, survey, s(depth, by=survey)	positive catch	phi	1.8638	0.15
DG	year, survey, s(depth, by=survey)	positive catch	sigma_O	0.5052	0.11
DG	year, survey, s(depth, by=survey)	positive catch	sigma_E	0.6565	0.07
DG	year, survey, s(depth, by=survey)	positive catch	sd__s(depth):as.factor(survey.f)ADFG	4.8657	
DG	year, survey, s(depth, by=survey)	positive catch	sd__s(depth):as.factor(survey.f)NBS	0.0000	

Table 2: Spatial parameter estimates of the Tweedie distribution family models. Terms include the Matèrn range (the distance at which spatial correlation decays such that two points are effectively independent), the standard deviation for spatial fields (σ_ω), and the standard deviation for spatiotemporal fields (σ_ϵ).

Distribution	Effects	Component	Term	Estimate	SE
Tweedie	year, survey		range	55.4263	8.91
Tweedie	year, survey		phi	32.4387	1.84
Tweedie	year, survey		sigma_O	1.4491	0.22
Tweedie	year, survey		sigma_E	1.1024	0.10
Tweedie	year, survey		tweedie_p	1.3257	0.02
Tweedie	year, survey, depth		range	59.9738	9.24
Tweedie	year, survey, depth		phi	32.3255	1.85
Tweedie	year, survey, depth		sigma_O	1.4757	0.21
Tweedie	year, survey, depth		sigma_E	1.1222	0.10
Tweedie	year, survey, depth		tweedie_p	1.3091	0.02
Tweedie	year, survey, depth*survey		range	61.7041	9.64
Tweedie	year, survey, depth*survey		phi	32.2192	1.85
Tweedie	year, survey, depth*survey		sigma_O	1.5221	0.22
Tweedie	year, survey, depth*survey		sigma_E	1.1113	0.10
Tweedie	year, survey, depth*survey		tweedie_p	1.3112	0.02
Tweedie	year, survey, s(depth, by=survey)		range	60.8638	9.60
Tweedie	year, survey, s(depth, by=survey)		phi	33.0265	1.85
Tweedie	year, survey, s(depth, by=survey)		sigma_O	1.4960	0.23
Tweedie	year, survey, s(depth, by=survey)		sigma_E	1.1302	0.10
Tweedie	year, survey, s(depth, by=survey)		tweedie_p	1.2889	0.02
Tweedie	year, survey, s(depth, by=survey)		sd__s(depth):as.factor(survey.f)ADFG	4.4858	
Tweedie	year, survey, s(depth, by=survey)		sd__s(depth):as.factor(survey.f)NBS	0.0003	

Table 3: Estimates of the survey and depth fixed effect coefficients from the models for which cross-validation analysis was successful. These models used either the delta gamma (DG) or Tweedie distribution families.

Distribution	Fixed effects	Term	Estimate	SE
DG	year, survey	survey.fNBS	0.77	0.30
DG	year, survey, depth	survey.fNBS	0.37	0.31
DG	year, survey, depth	depth	0.18	0.04
DG	year, survey, s(depth, by=survey)	survey.fNBS	0.40	0.31
DG	year, survey, s(depth, by=survey)	sdepth):as.factor(survey.f)ADFG	0.31	0.91
DG	year, survey, s(depth, by=survey)	sdepth):as.factor(survey.f)NBS	1.57	0.32
Tweedie	year, survey	survey.fNBS	-0.18	0.18
Tweedie	year, survey, depth	survey.fNBS	-0.46	0.18
Tweedie	year, survey, depth	depth	0.16	0.02
Tweedie	year, survey, depth*survey	survey.fNBS	-1.25	0.86
Tweedie	year, survey, depth*survey	depth	0.15	0.02
Tweedie	year, survey, depth*survey	survey.fNBS:depth	0.04	0.04
Tweedie	year, survey, s(depth, by=survey)	survey.fNBS	-0.27	0.20
Tweedie	year, survey, s(depth, by=survey)	sdepth):as.factor(survey.f)ADFG	0.11	1.24
Tweedie	year, survey, s(depth, by=survey)	sdepth):as.factor(survey.f)NBS	1.05	0.25

Figures

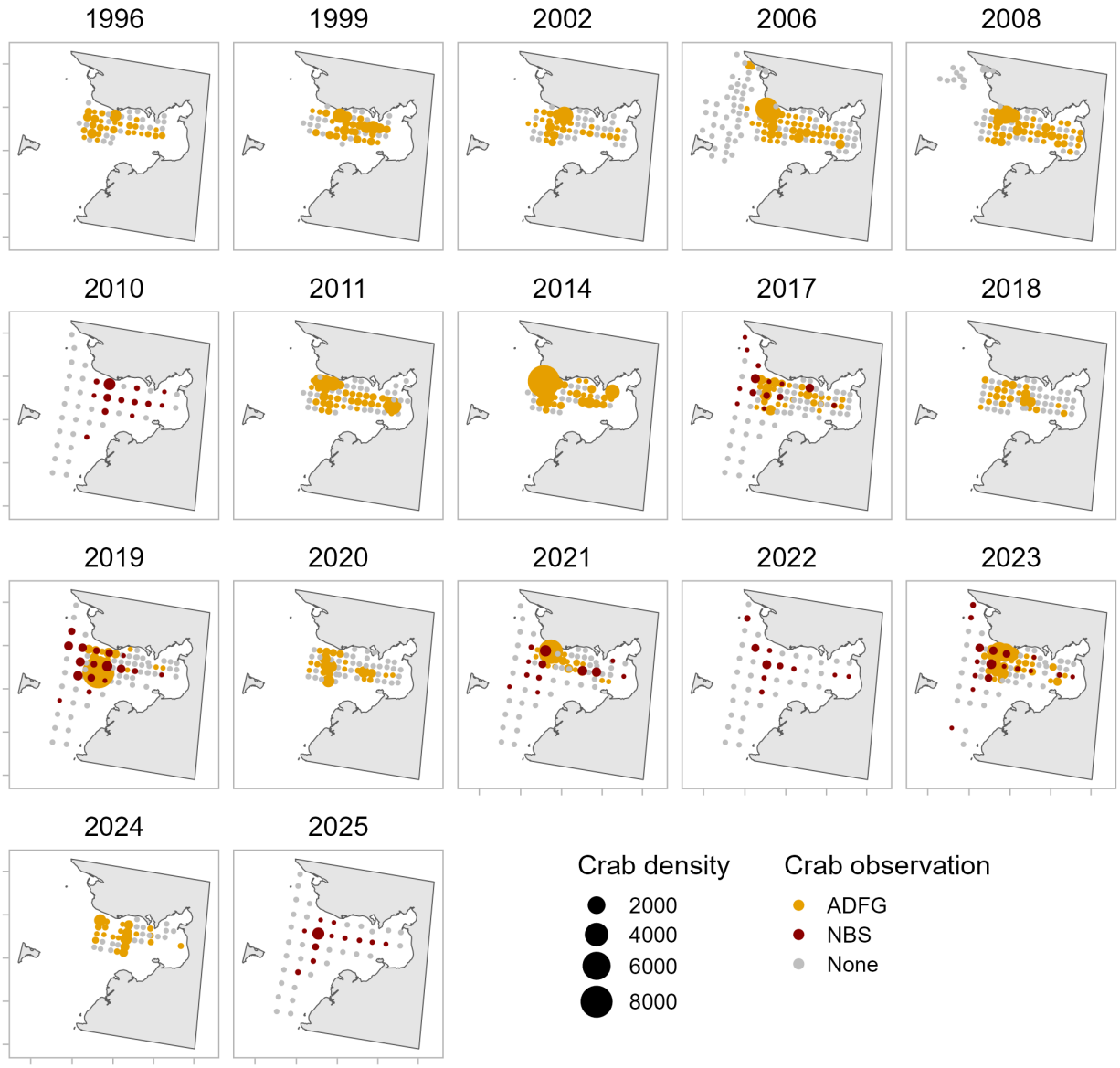


Figure 1: Norton Sound red king crab density at survey stations sampled in each year. Point size is proportional to the density of male red king crab $\geq$ 64 mm in carapace length at the station. Stations colored orange were sampled in the ADFG trawl survey and recorded at least one male red king crab $\geq$ 64 mm in carapace length, stations colored red were sampled in the NOAA NBS trawl survey and recorded at least one male red king crab $\geq$ 64 mm in carapace length, and stations colored grey recorded zero male red king crab $\geq$ 64 mm in carapace length.

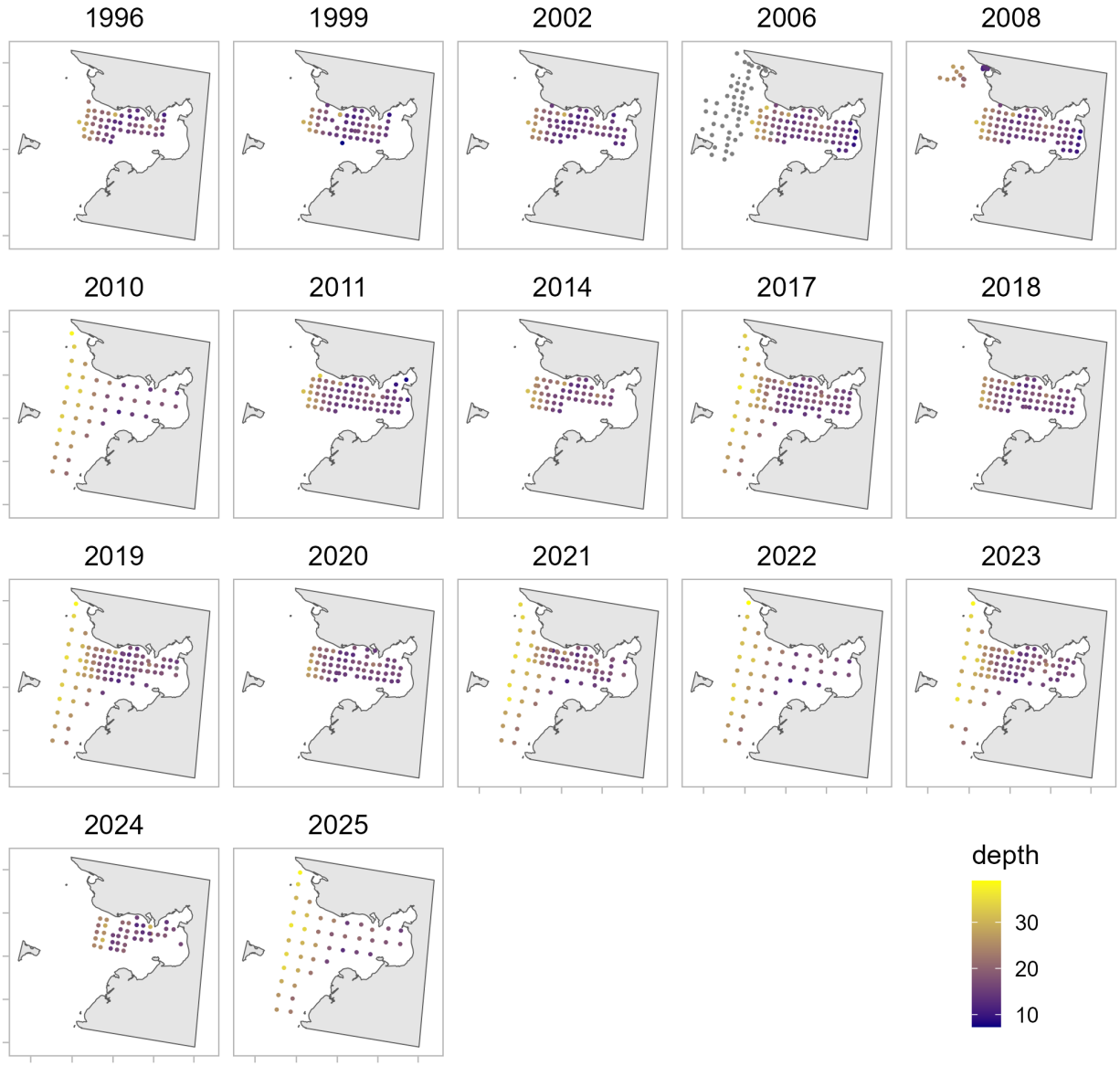


Figure 2: Measured depths (m) of stations sampled in the NOAA NBS and ADFG trawl surveys. Grey points are stations without depth measurements.

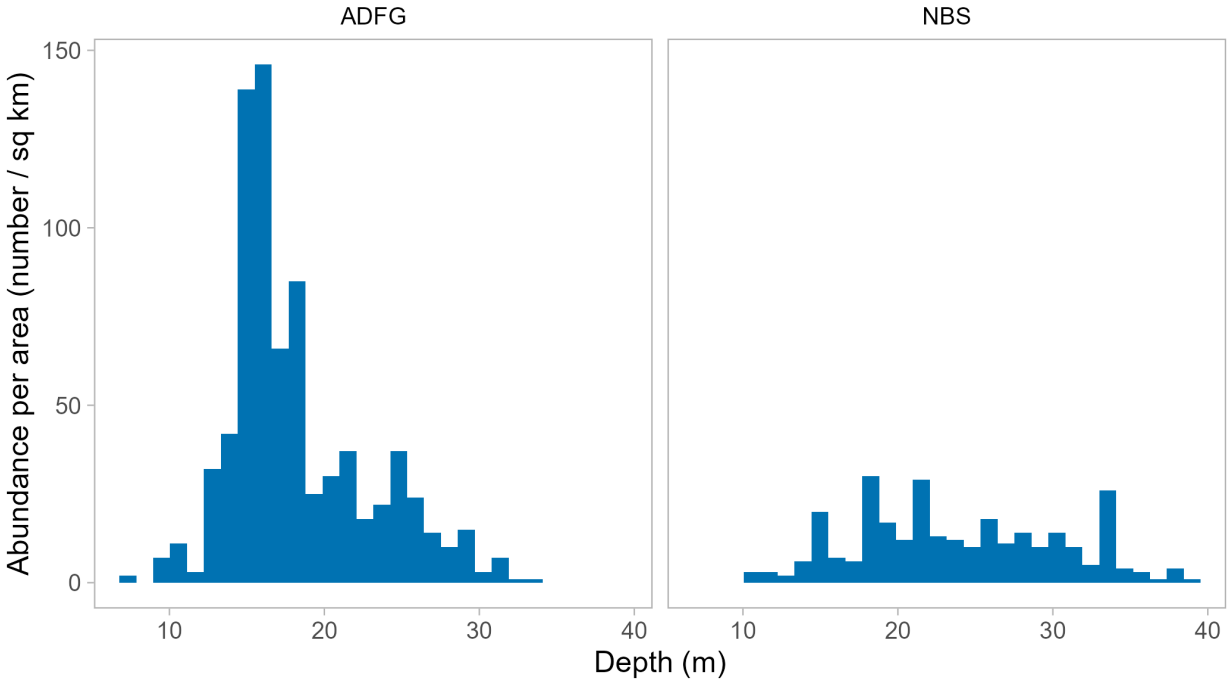


Figure 3: Histograms showing the differences between the ADFG trawl survey and NOAA NBS trawl survey in both the range of survey station depths sampled and the abundance per area (number per km^2 of male Norton Sound red king crab $\geq$ 64 mm in carapace length).

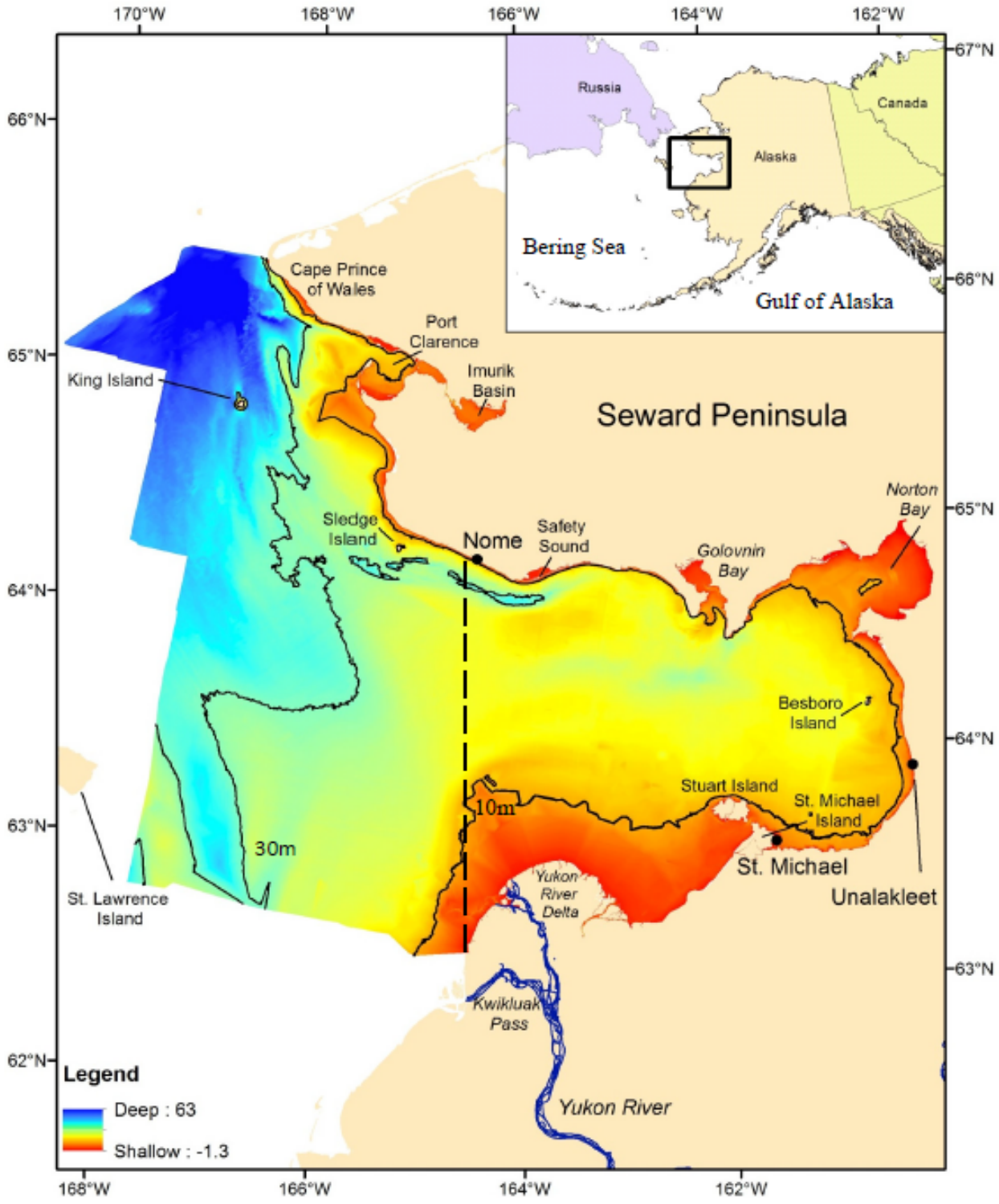


Figure 4: Bathymetry of Norton Sound, from Prescott and Zimmermann (2015). Depths of 10 m, and 30 m are marked by contour lines.

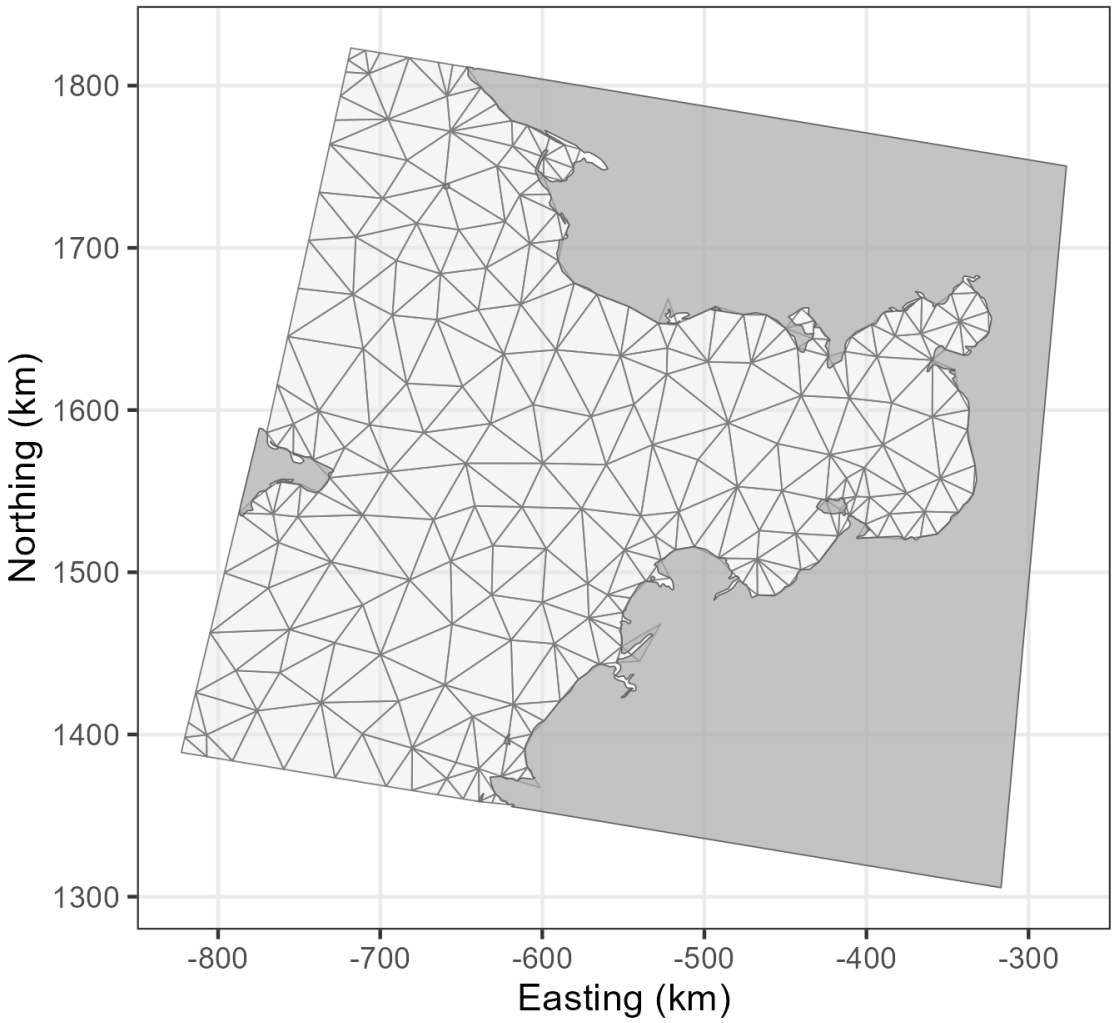


Figure 5: Spatial mesh used in spatiotemporal models, with land areas filled in dark grey. Vertices on land or outside the designated polygon were excluded from the mesh.

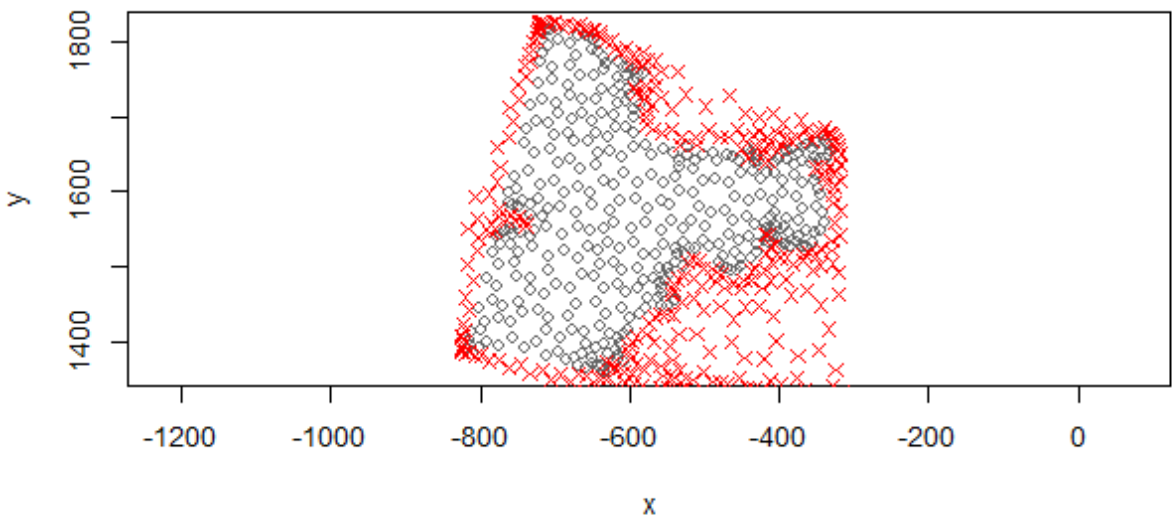


Figure 6: Vertices of the spatial barrier mesh used in spatiotemporal models, demonstrating that land as well as the area outside the designated polygon were excluded from the mesh (red crosses).

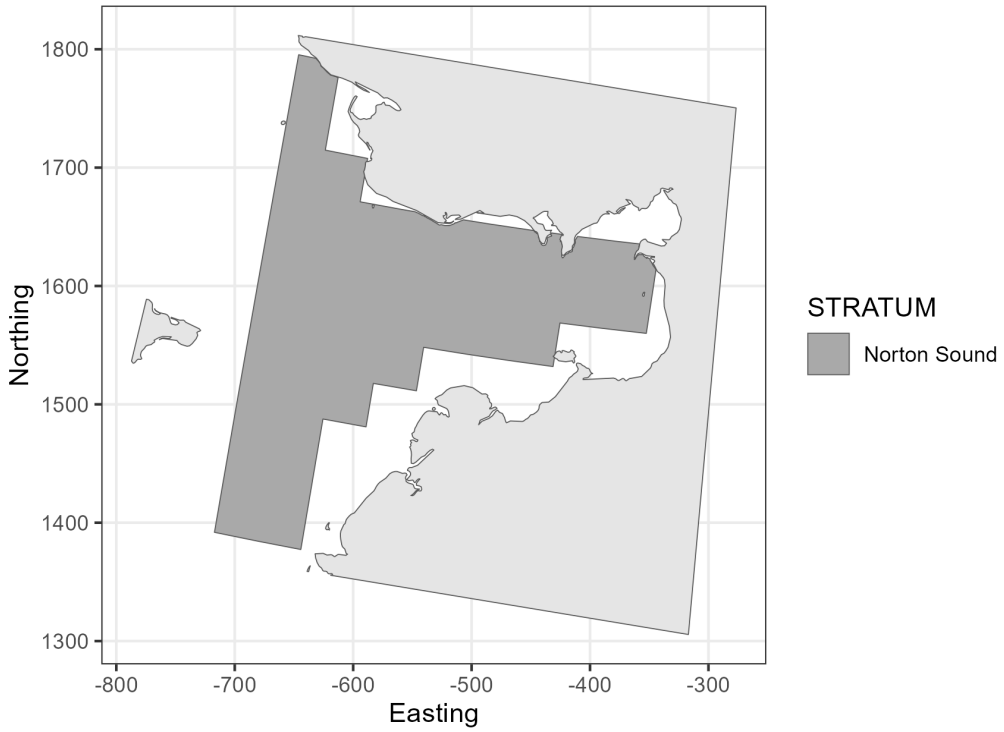


Figure 7: NOAA Northern Bering Sea bottom trawl survey stratum for the Norton Sound red king crab stock.

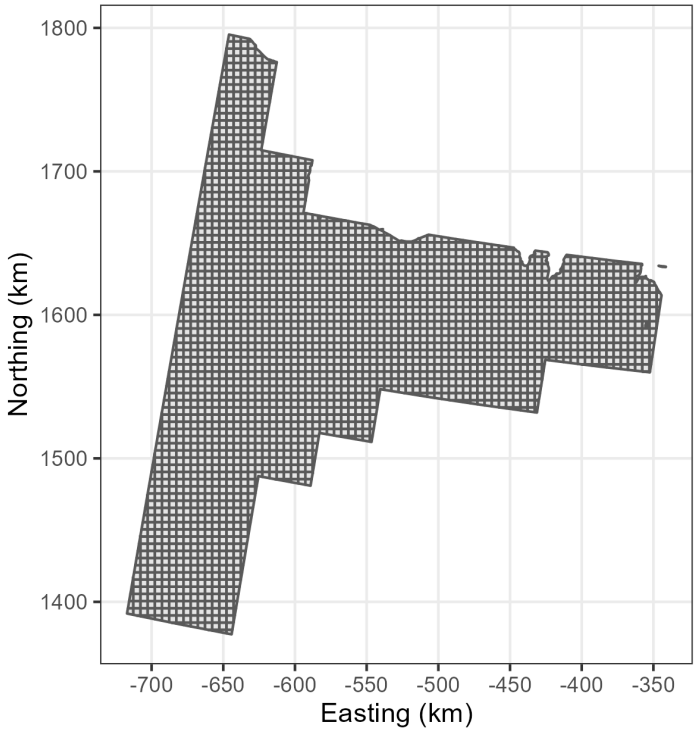


Figure 8: Prediction grid used to generate Norton Sound red king crab abundance estimates.

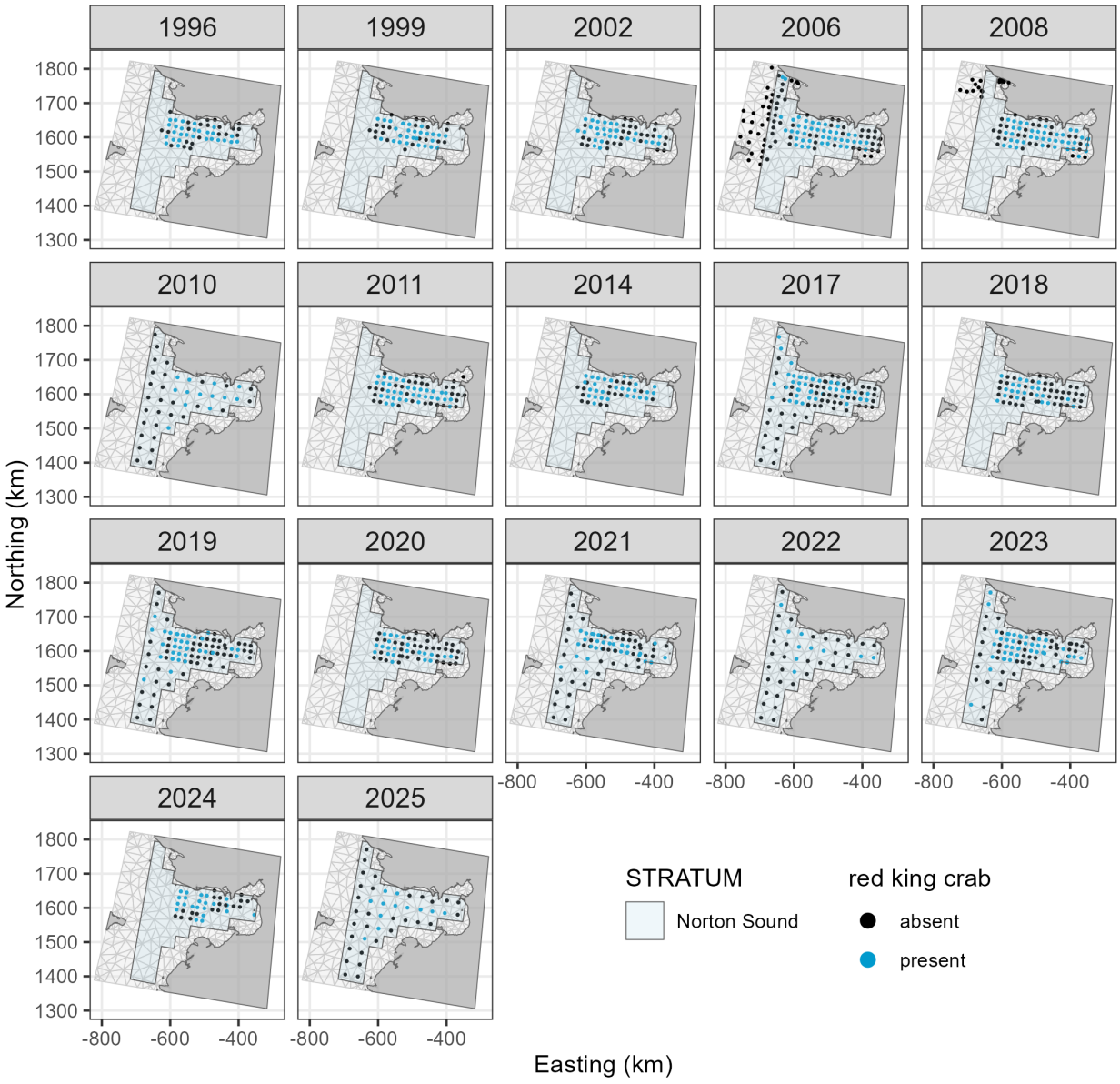


Figure 9: NOAA Northern Bering Sea bottom trawl survey stratum for the Norton Sound red king crab stock overlaid on the spatial mesh used in spatiotemporal models with ADFG and/or NOAA NBS survey tow locations by year; blue points indicate crab presence in the tow and black points indicate crab absence in the tow.

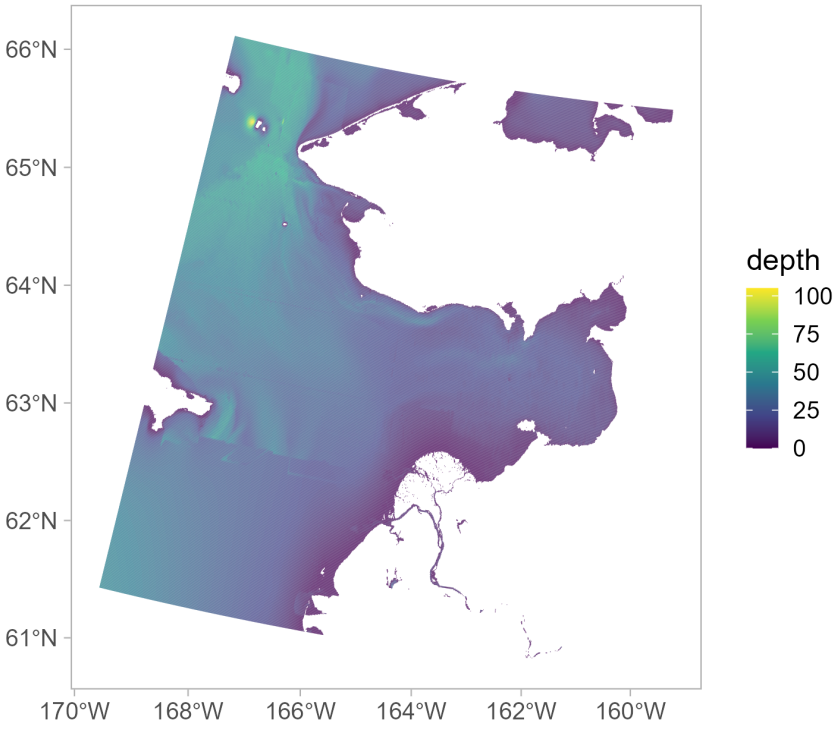


Figure 10: Depth information available for Norton Sound from the Southern Alaska Coastal Relief Model (Lim et al. 2011).

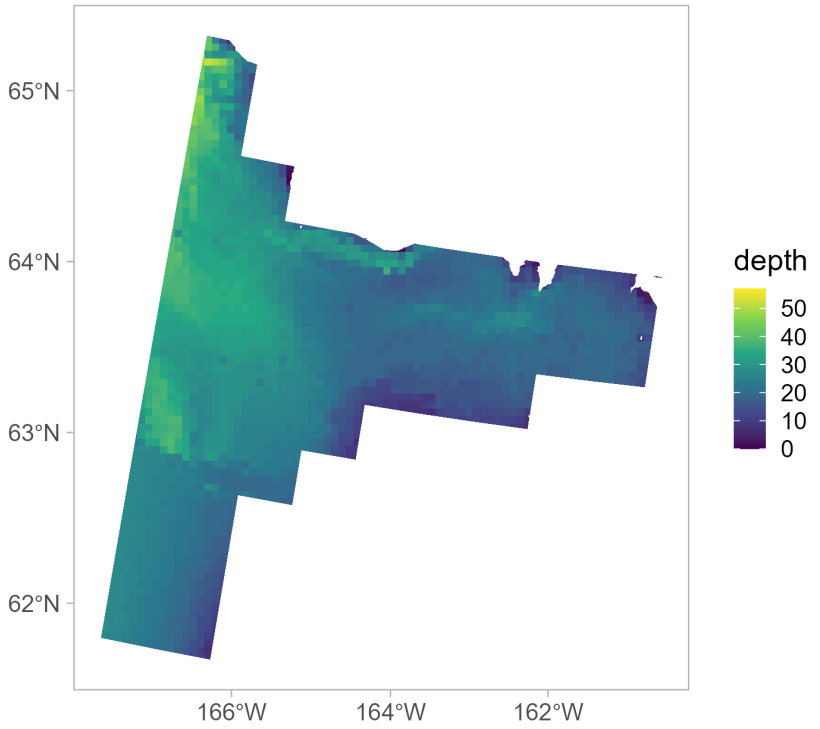


Figure 11: The depth information included in the prediction grid, generated by matching prediction grid locations to the Southern Alaska Coastal Relief Model data set using a nearest neighbor spatial join.

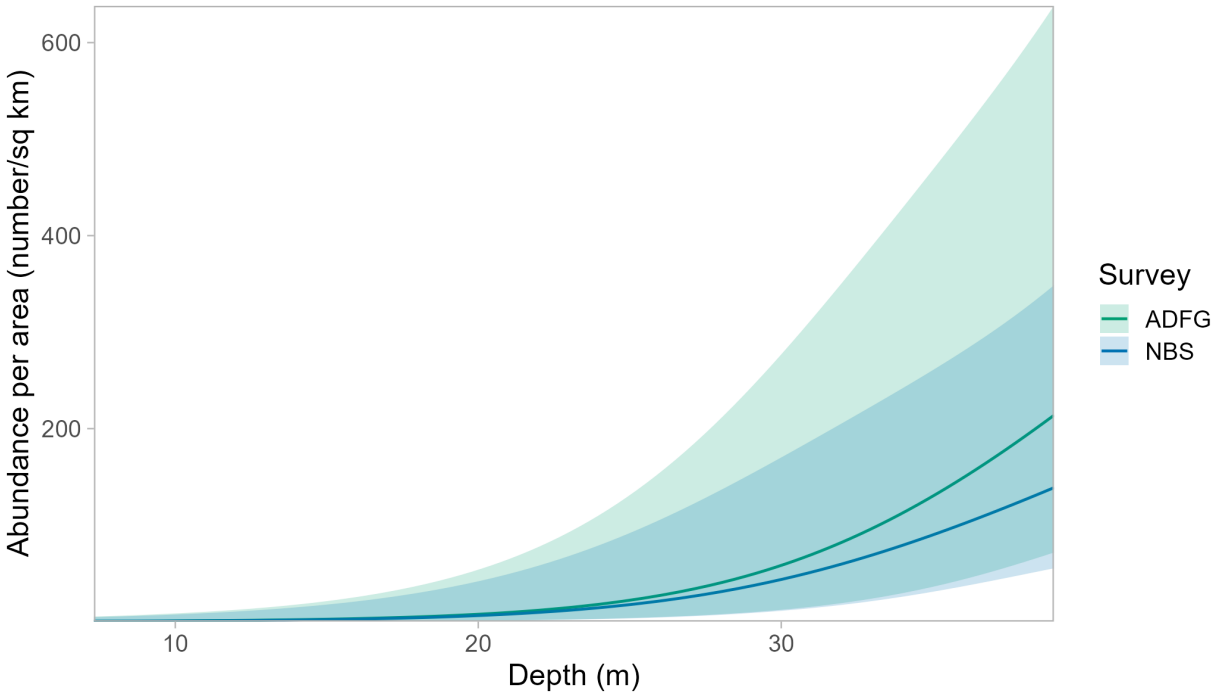


Figure 12: Conditional effect of depth on abundance per area by survey from the delta gamma distribution family model with year, survey, and depth as fixed effects.

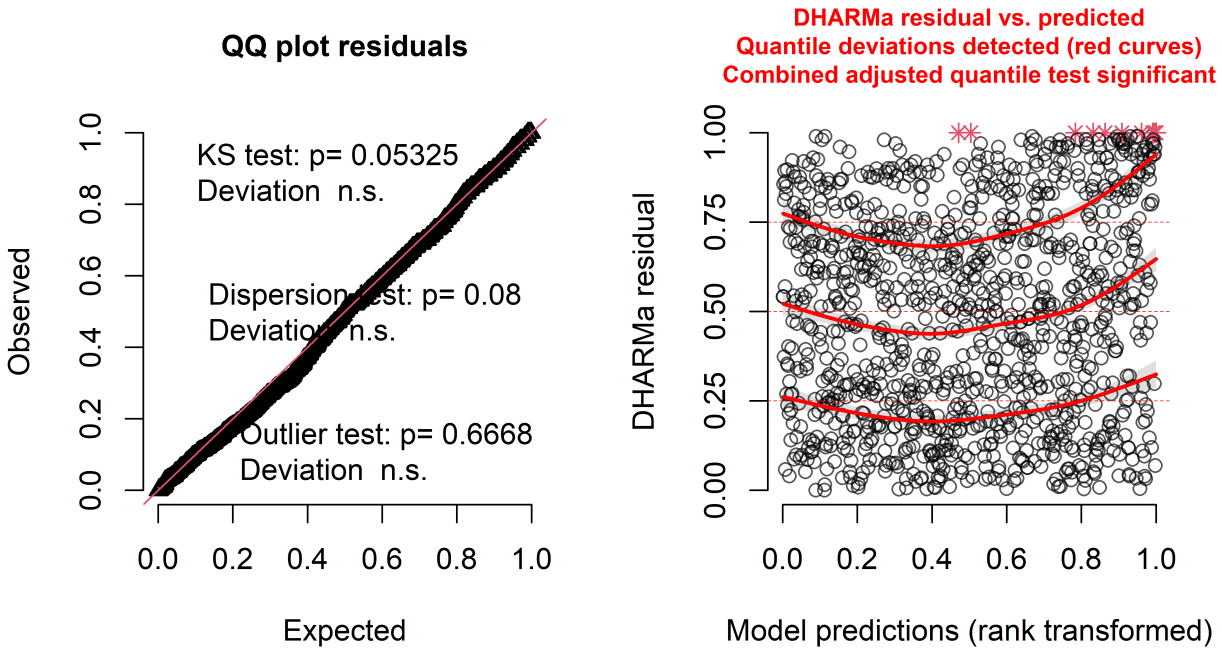


Figure 13: Diagnostics for the delta gamma distribution family model with year and survey as fixed effects, using DHARMa residuals.

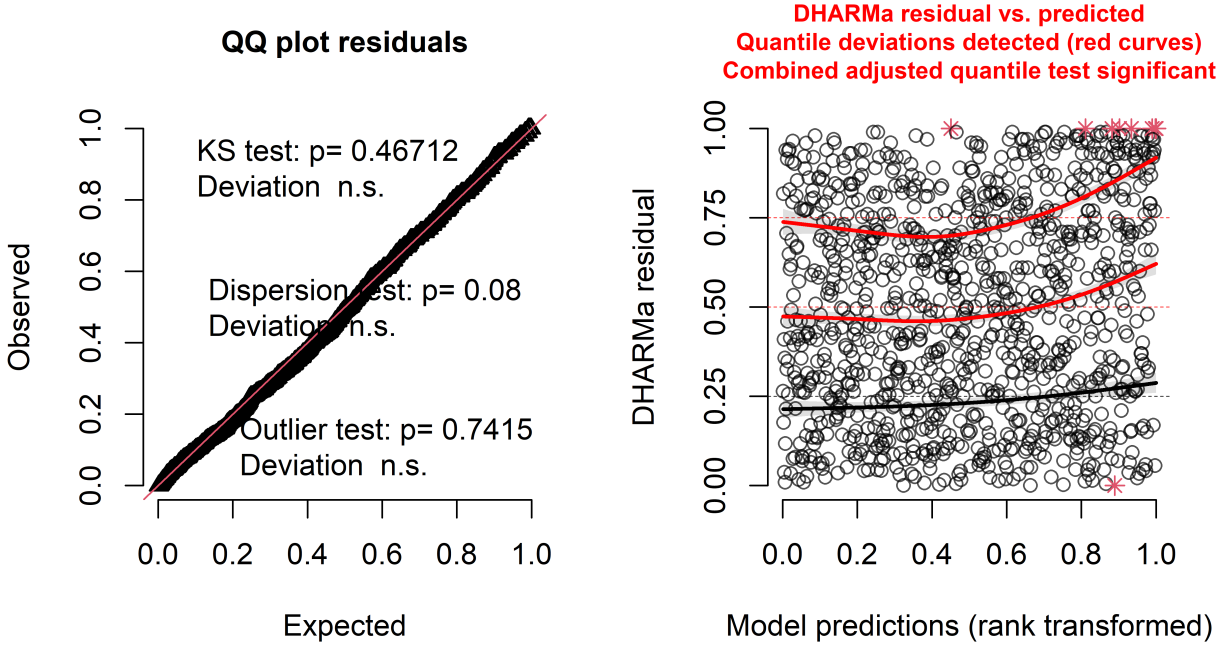


Figure 14: Diagnostics for the delta gamma distribution family model with year, survey, and depth as fixed effects, using DHARMa residuals.

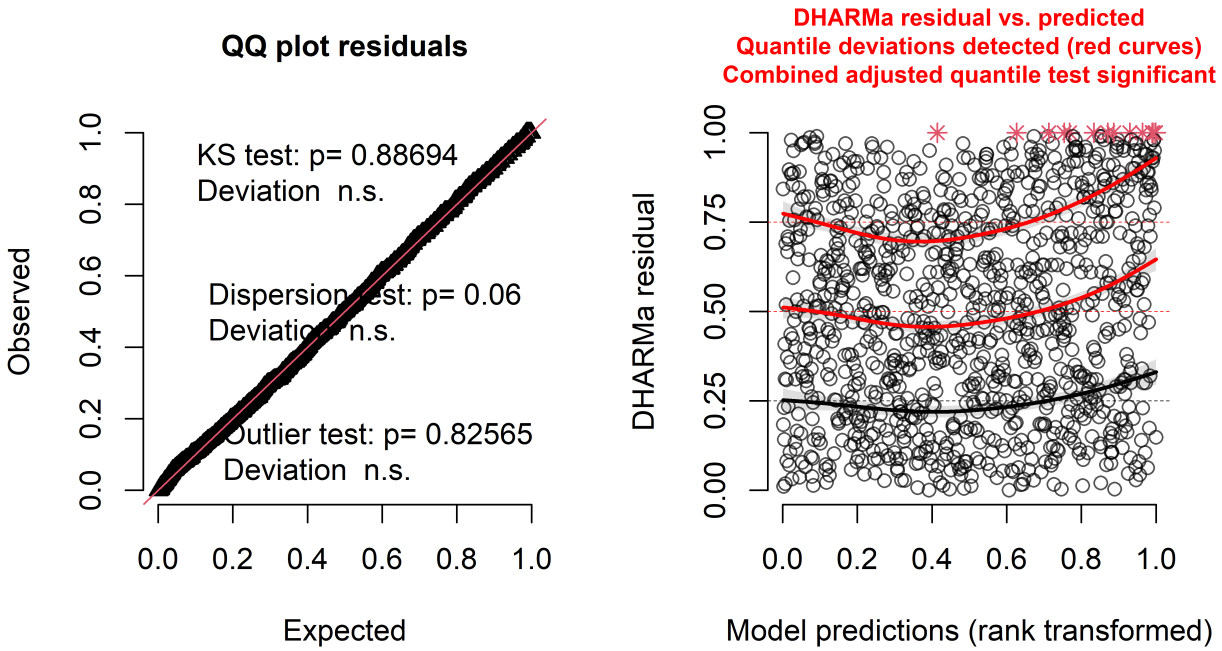


Figure 15: Diagnostics for the delta gamma distribution family model with year and survey as fixed effects and a smooth of depth by survey, using DHARMa residuals.

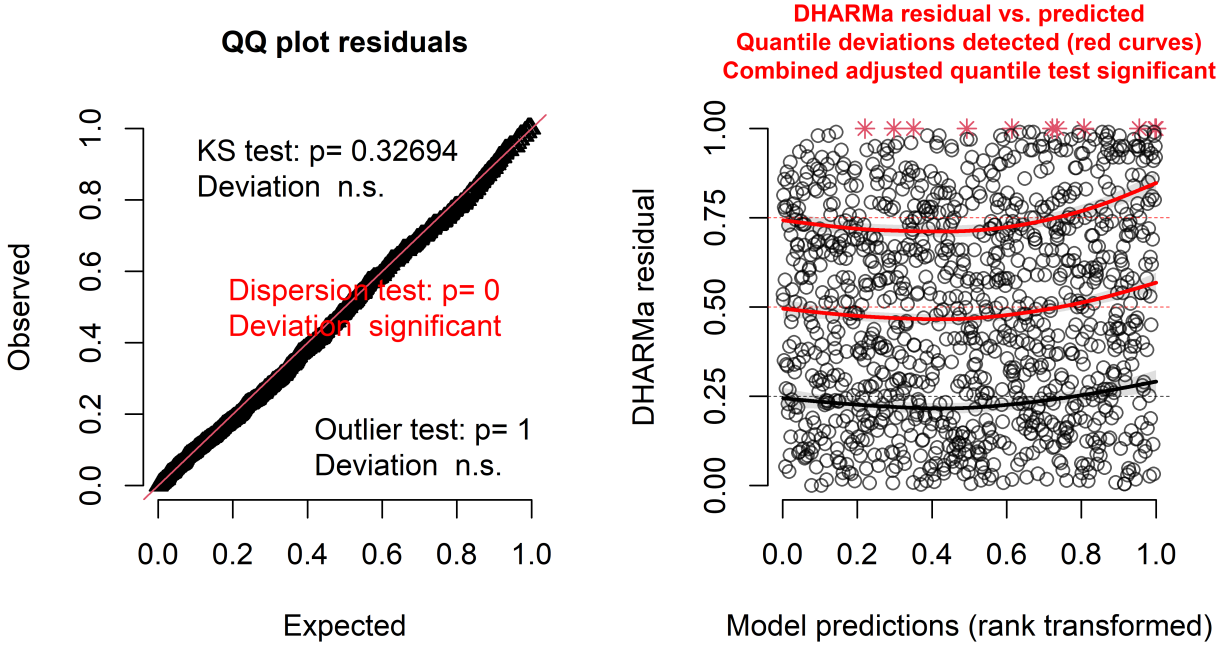


Figure 16: Diagnostics for the delta lognormal distribution family model with year and survey as fixed effects, using DHARMa residuals.

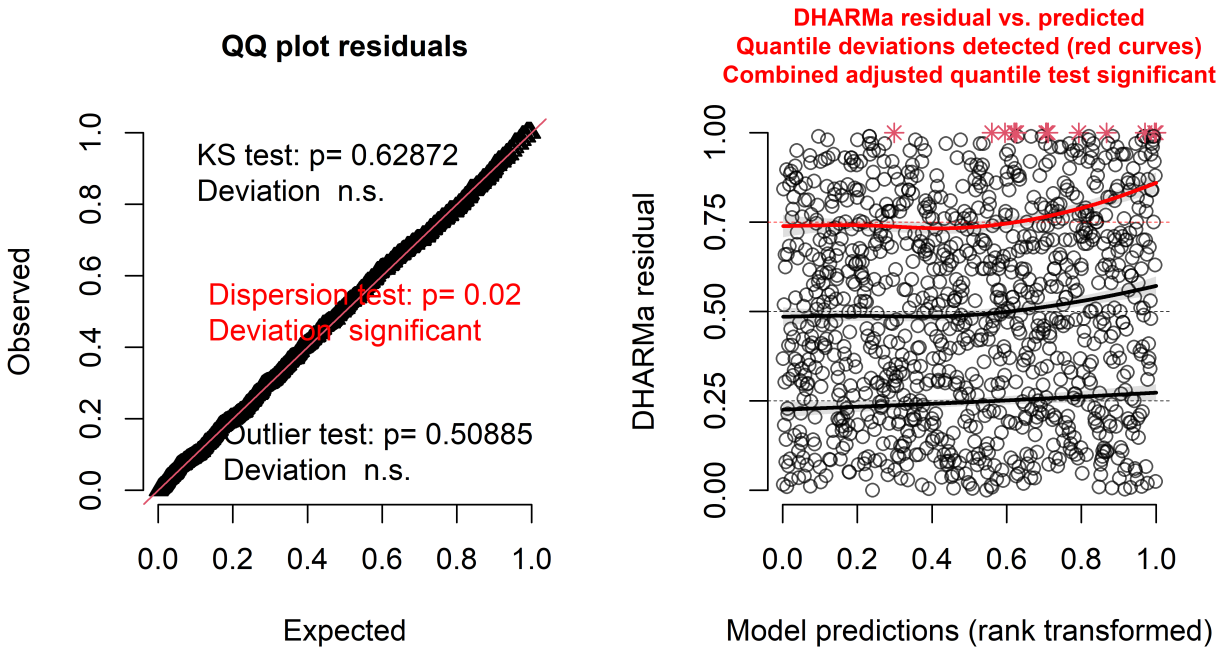


Figure 17: Diagnostics for the delta lognormal distribution family model with year, survey, and depth as fixed effects, using DHARMa residuals.

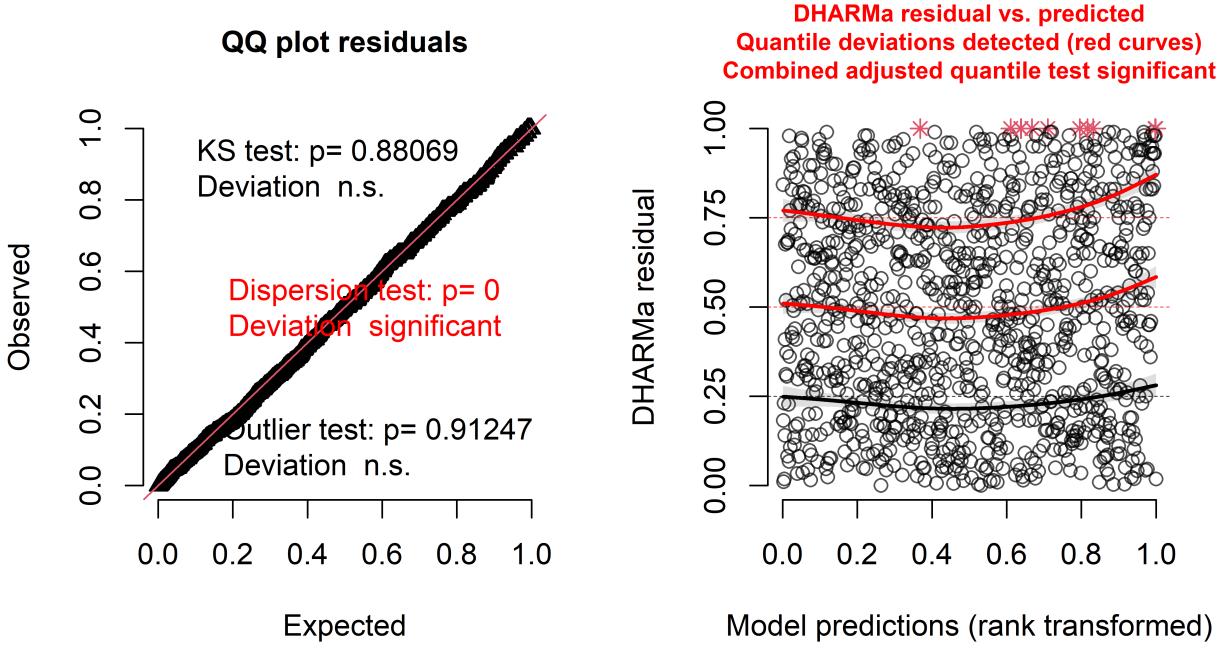


Figure 18: Diagnostics for the delta lognormal distribution family model with year and survey as fixed effects and a linear interaction between depth and survey, using DHARMa residuals.

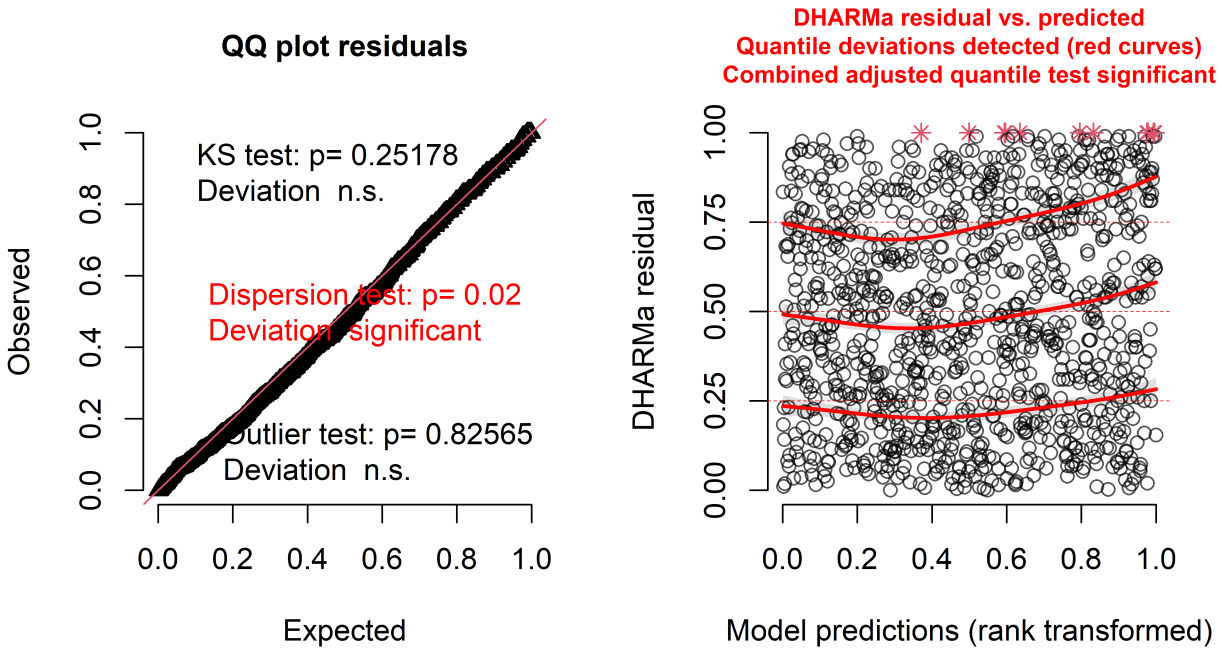


Figure 19: Diagnostics for the delta lognormal distribution family model with year and survey as fixed effects and a smooth of depth by survey, using DHARMa residuals.

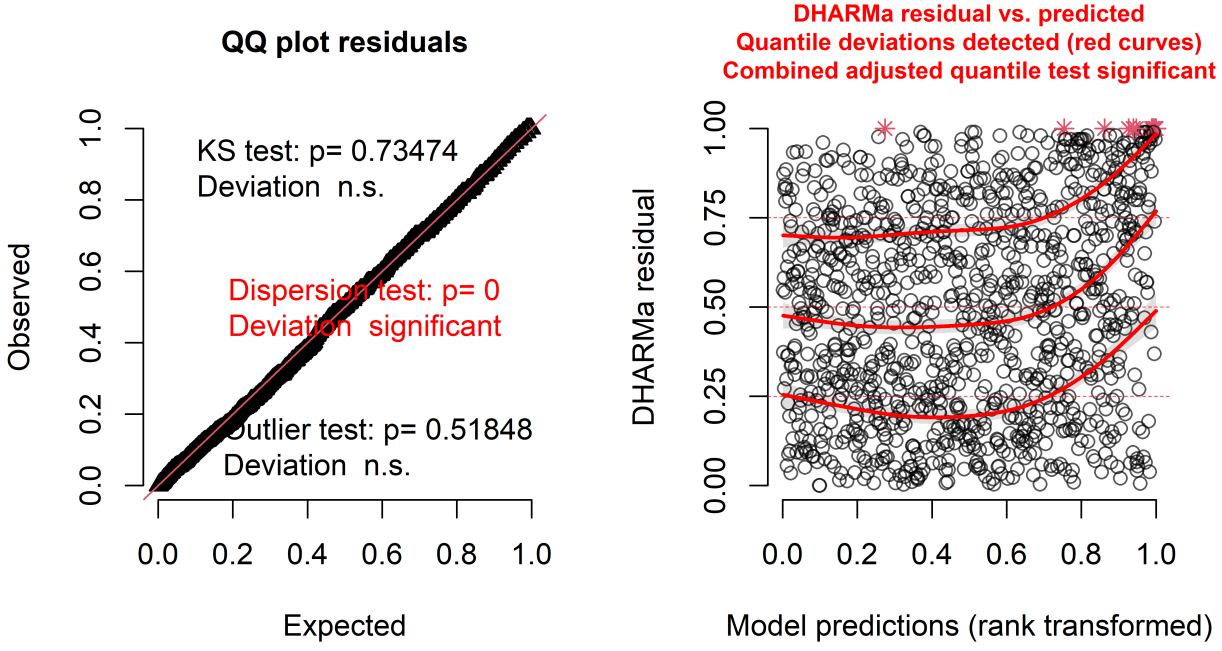


Figure 20: Diagnostics for the Tweedie distribution family model with year and survey as fixed effects, using DHARMa residuals.

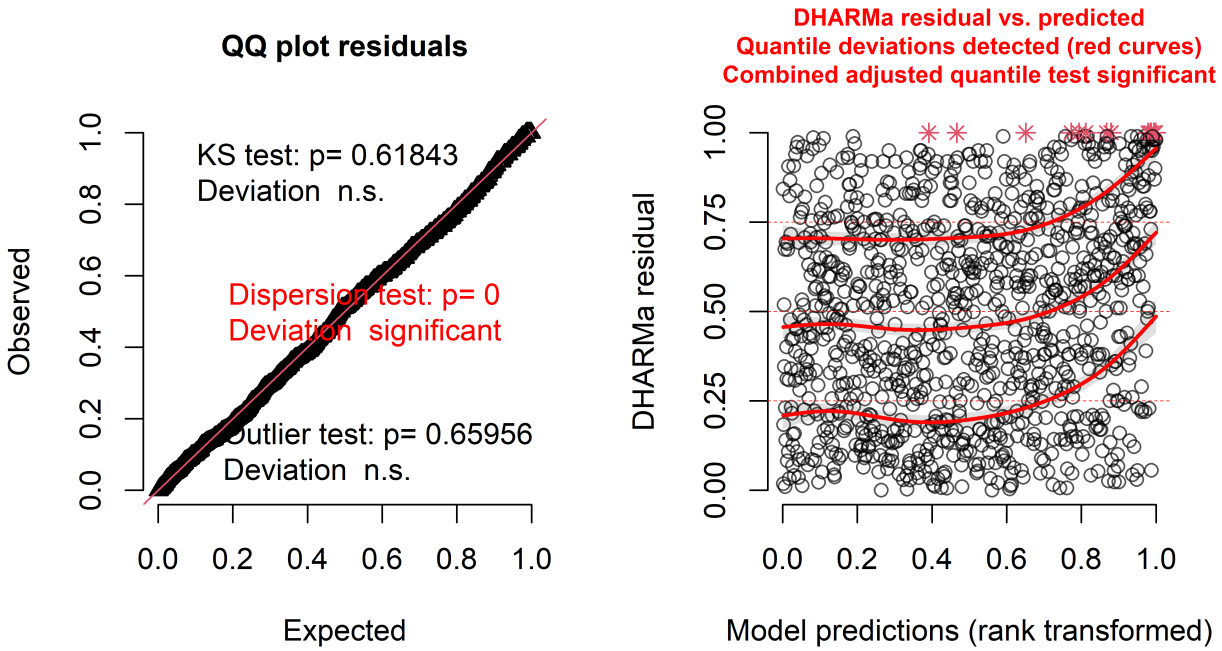


Figure 21: Diagnostics for the Tweedie distribution family model with year, survey, and depth as fixed effects, using DHARMa residuals.

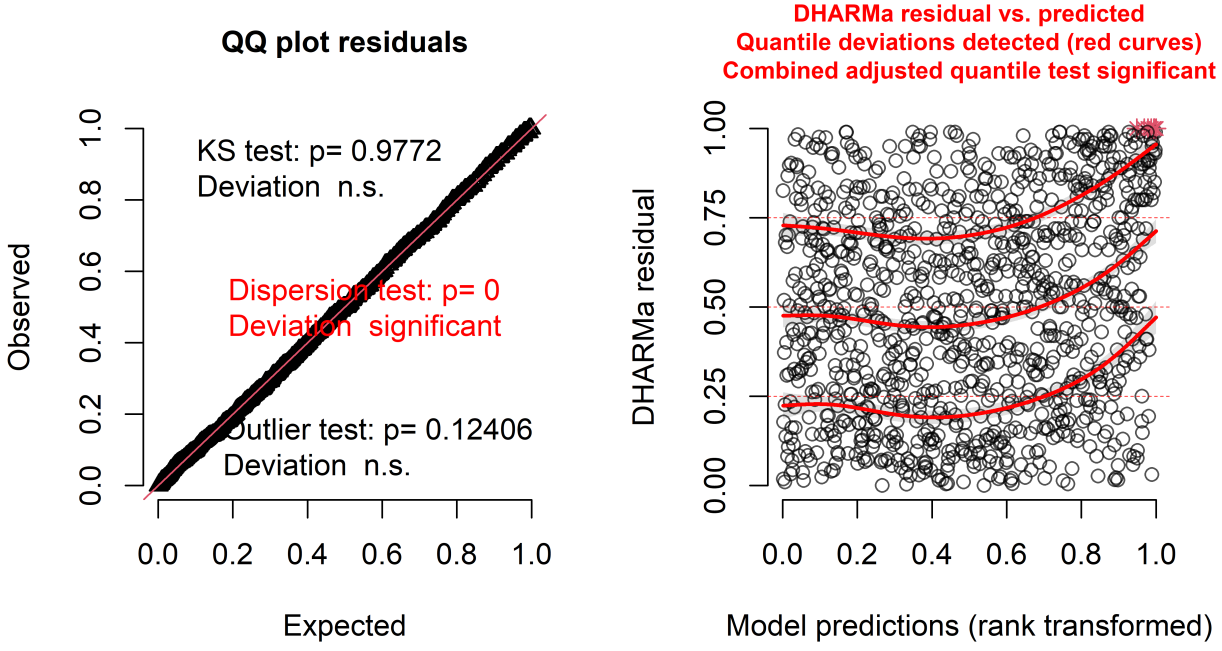


Figure 22: Diagnostics for the Tweedie distribution family model with year and survey as fixed effects and a linear interaction between depth and survey, using DHARMa residuals.

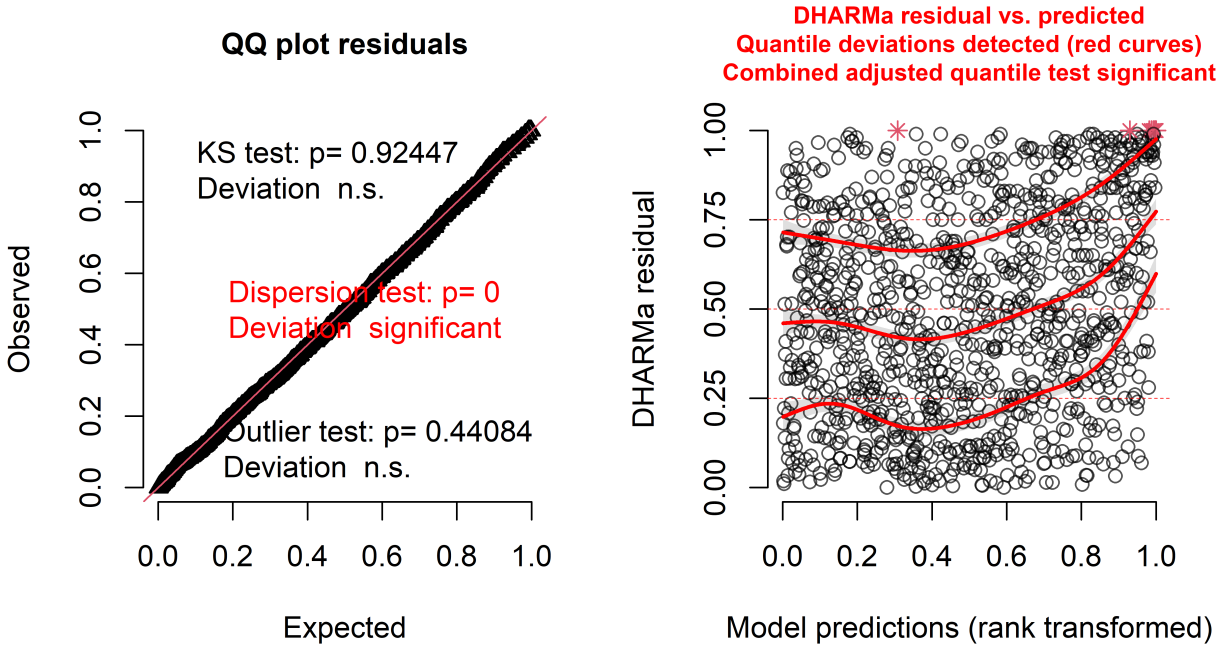


Figure 23: Diagnostics for the Tweedie distribution family model with year and survey as fixed effects and a smooth of depth by survey, using DHARMa residuals.

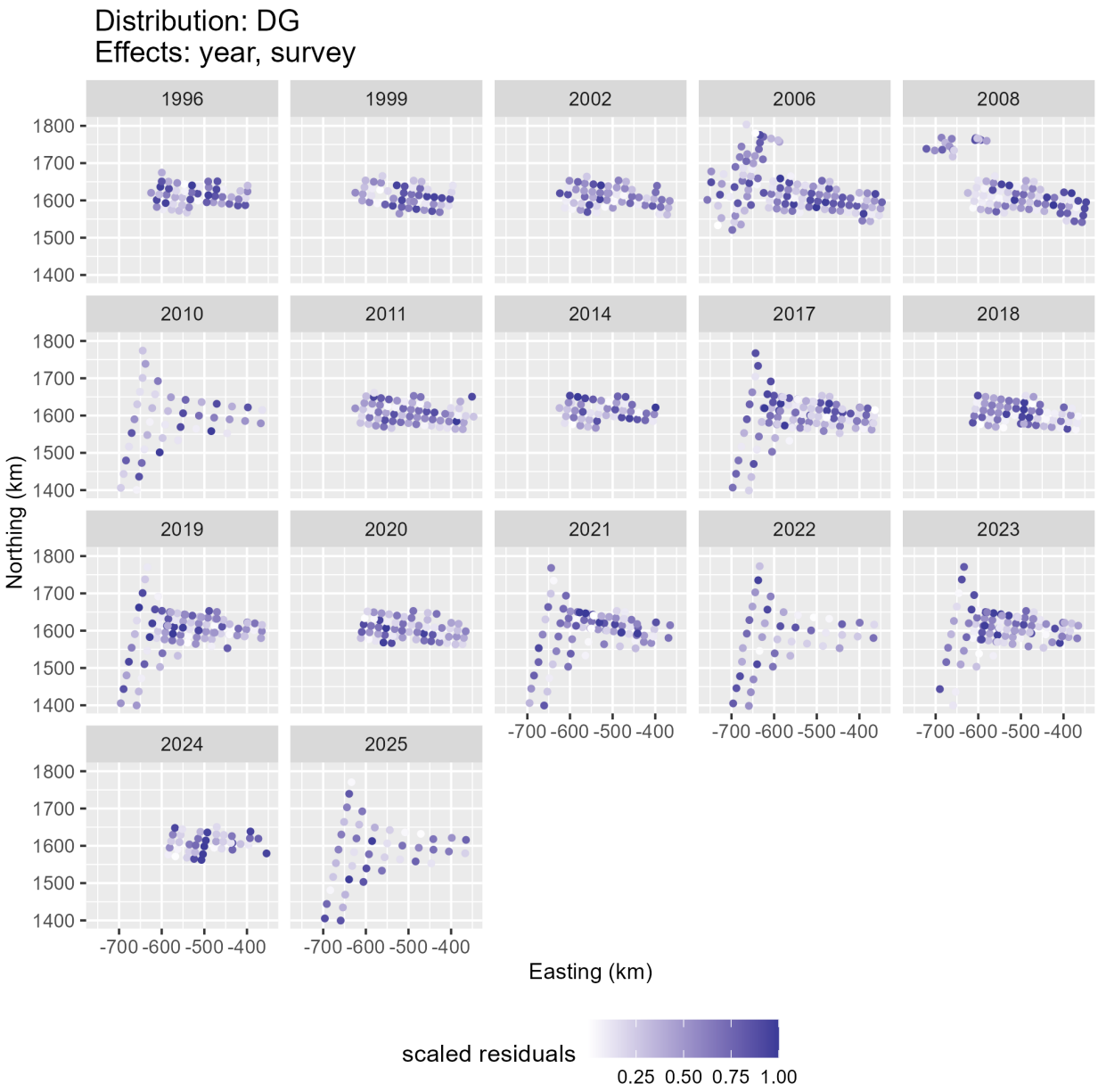


Figure 24: Spatial distribution of DHARMa residuals for the model using the delta gamma distribution with year and survey as fixed effects.

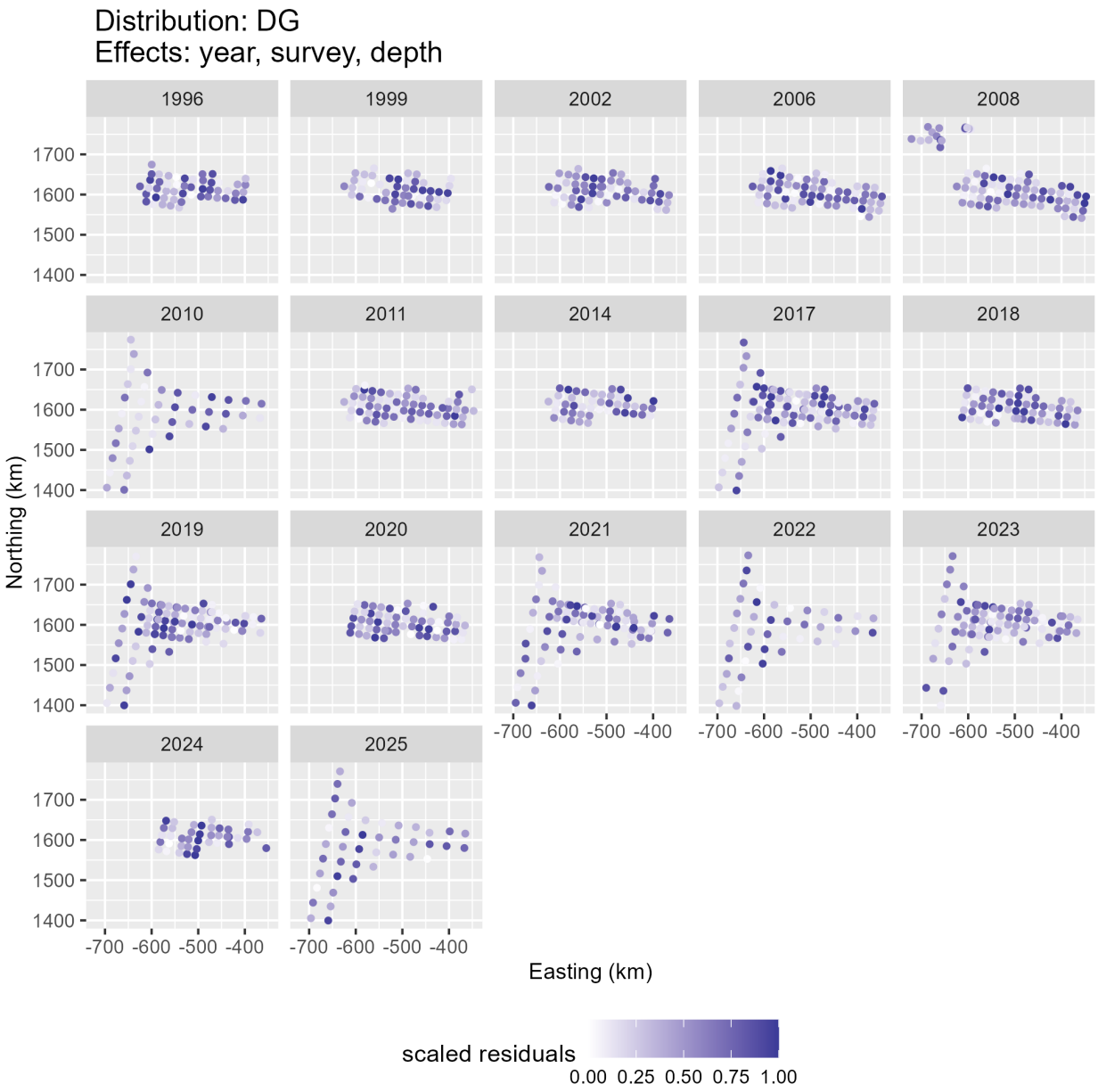


Figure 25: Spatial distribution of DHARMa residuals for the model using the delta gamma distribution with year, survey, and depth as fixed effects.

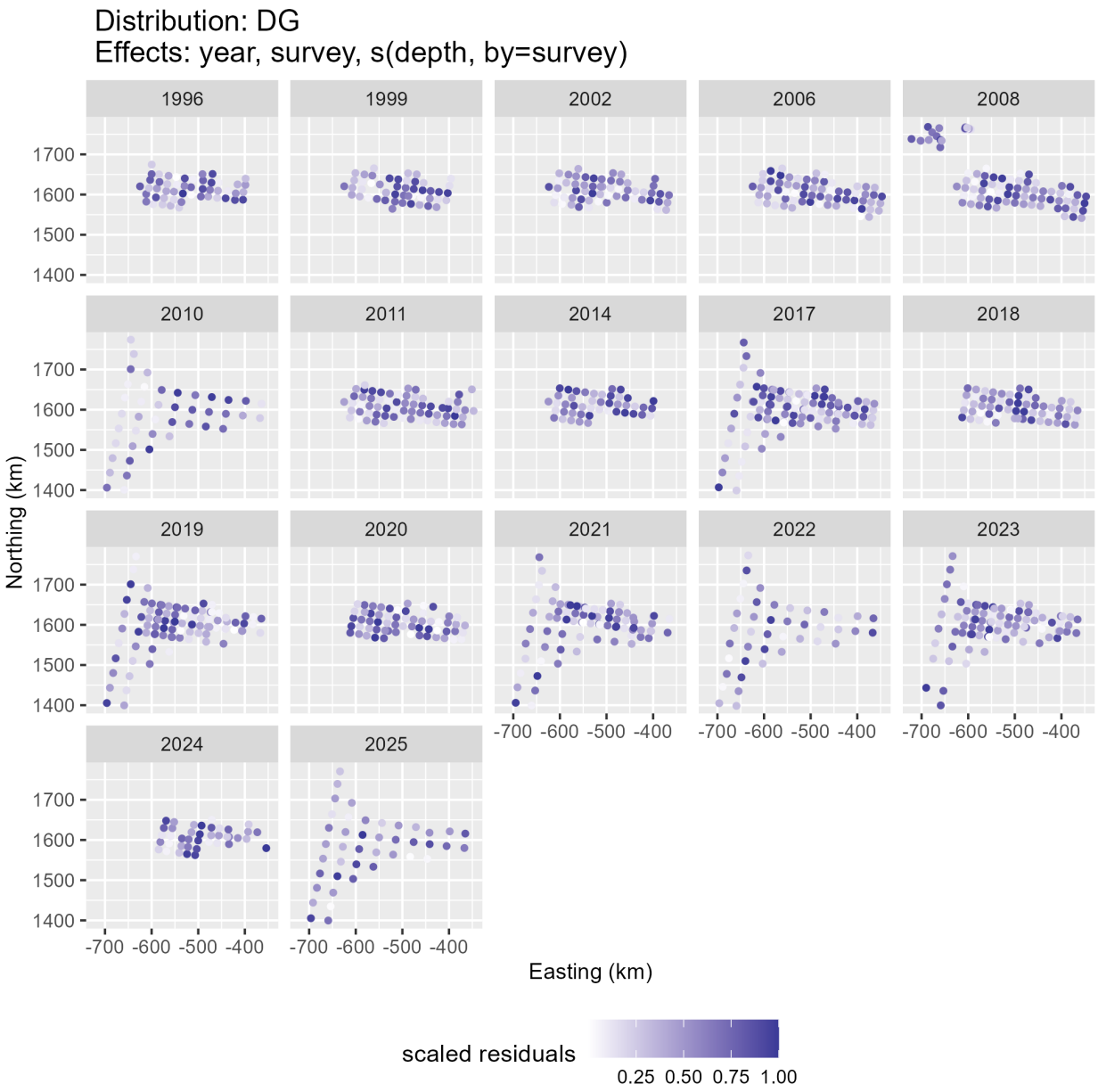


Figure 26: Spatial distribution of DHARMa residuals for the model using the delta gamma distribution with year and survey as fixed effects and a smooth of depth by survey.

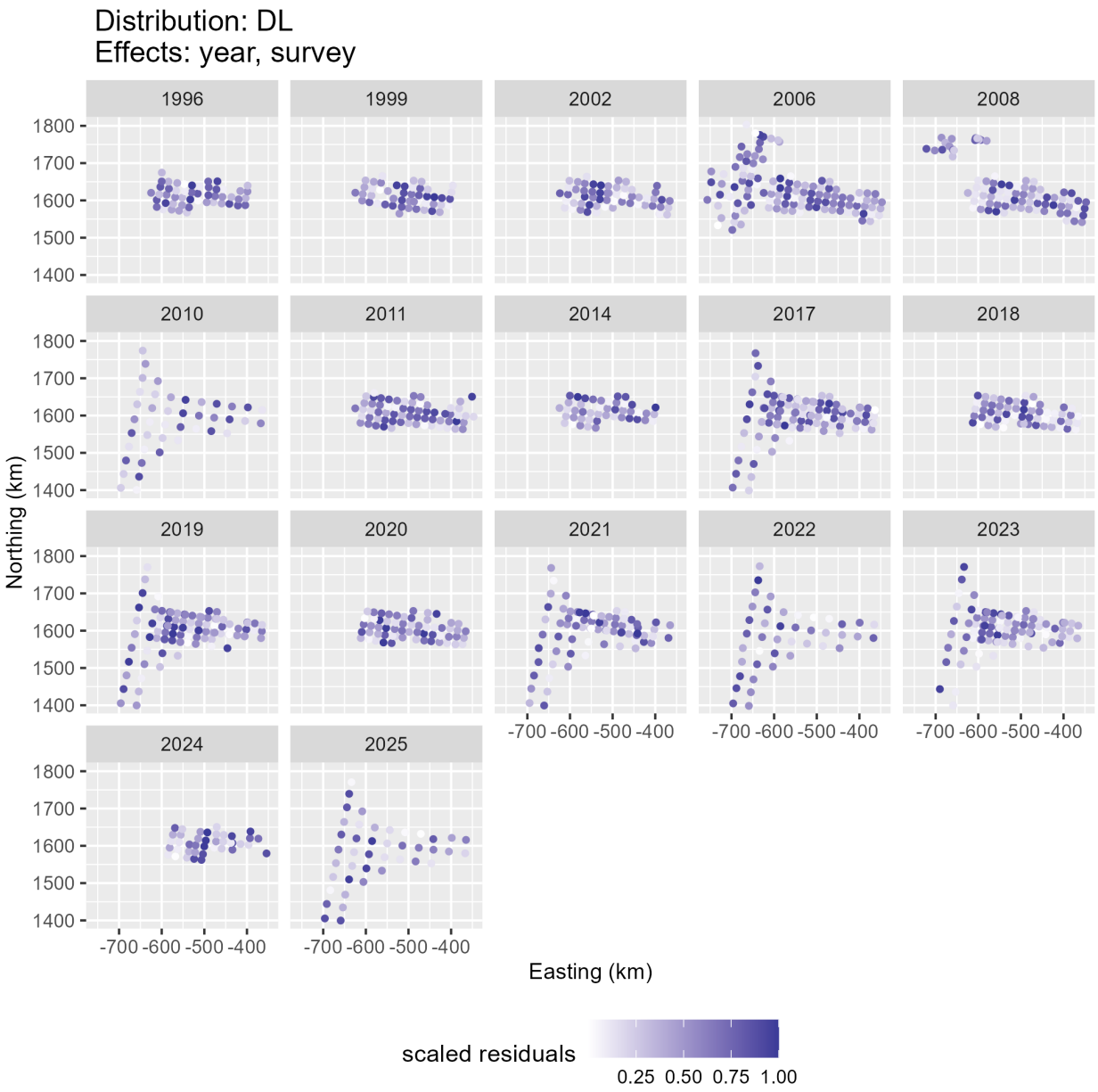


Figure 27: Spatial distribution of DHARMa residuals for the model using the delta lognormal distribution with year and survey as fixed effects.

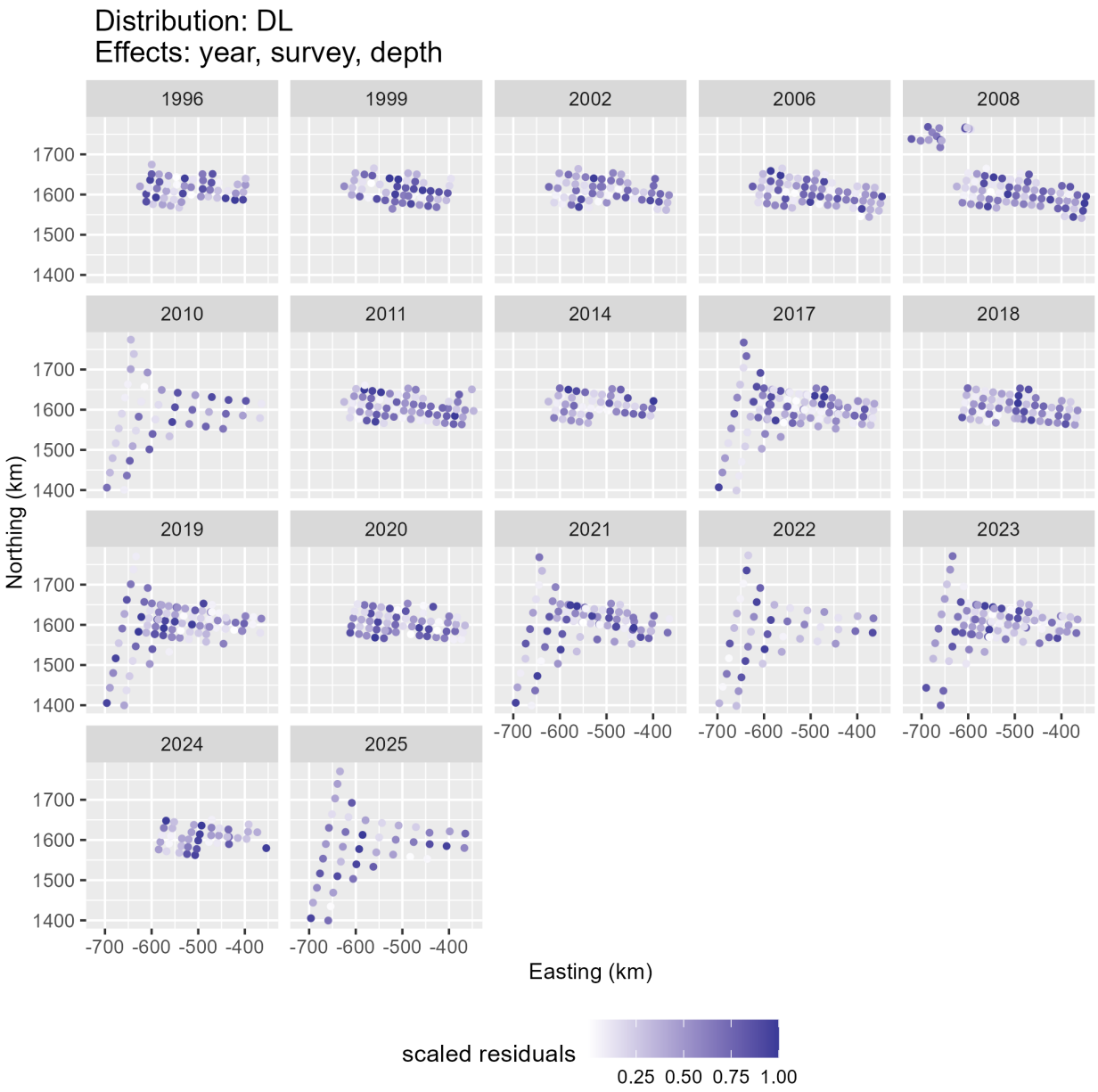


Figure 28: Spatial distribution of DHARMa residuals for the model using the delta lognormal distribution with year, survey, and depth as fixed effects.

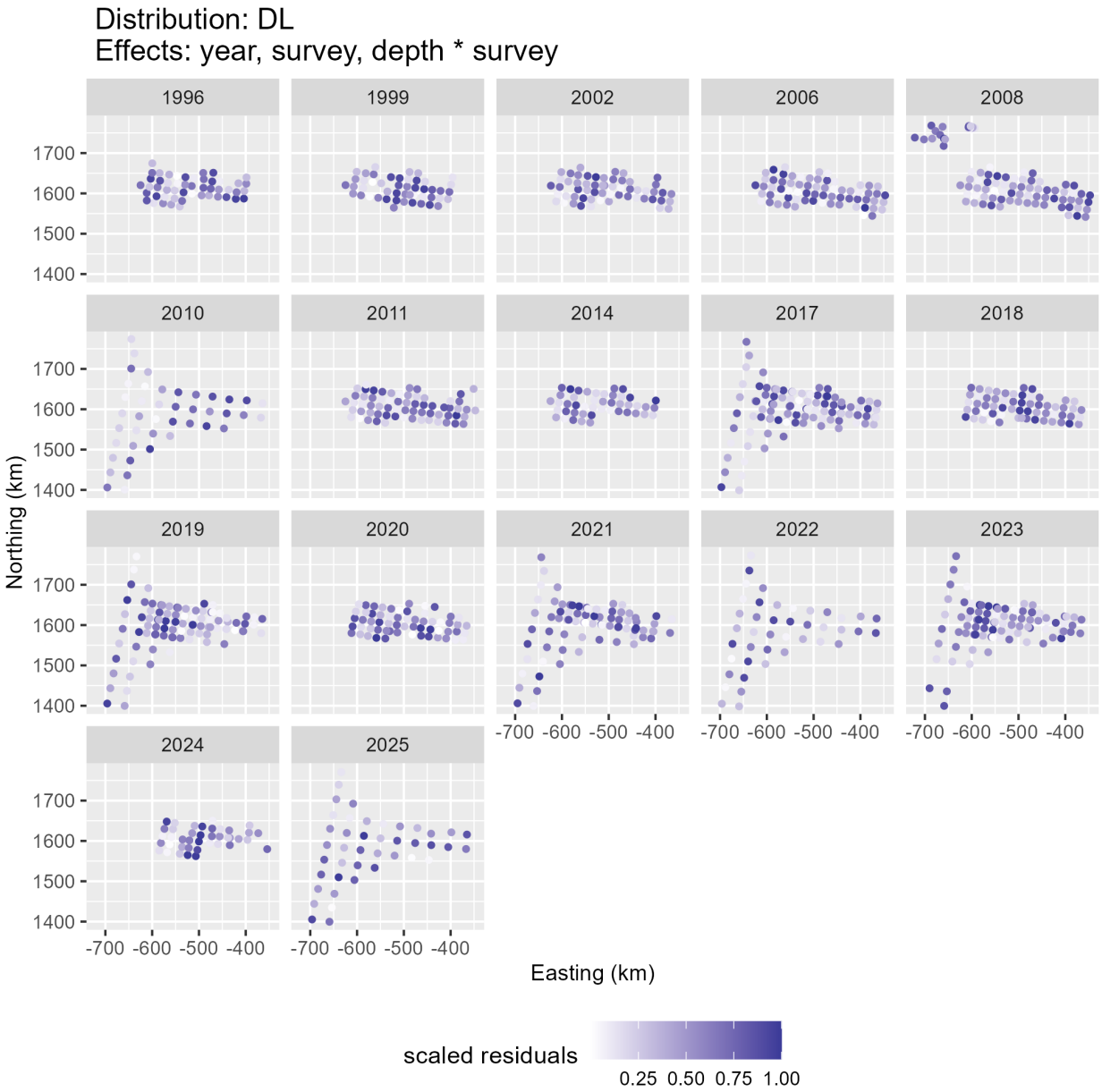


Figure 29: Spatial distribution of DHARMa residuals for the model using the delta lognormal distribution with year and survey as fixed effects and a linear interaction between depth and survey.

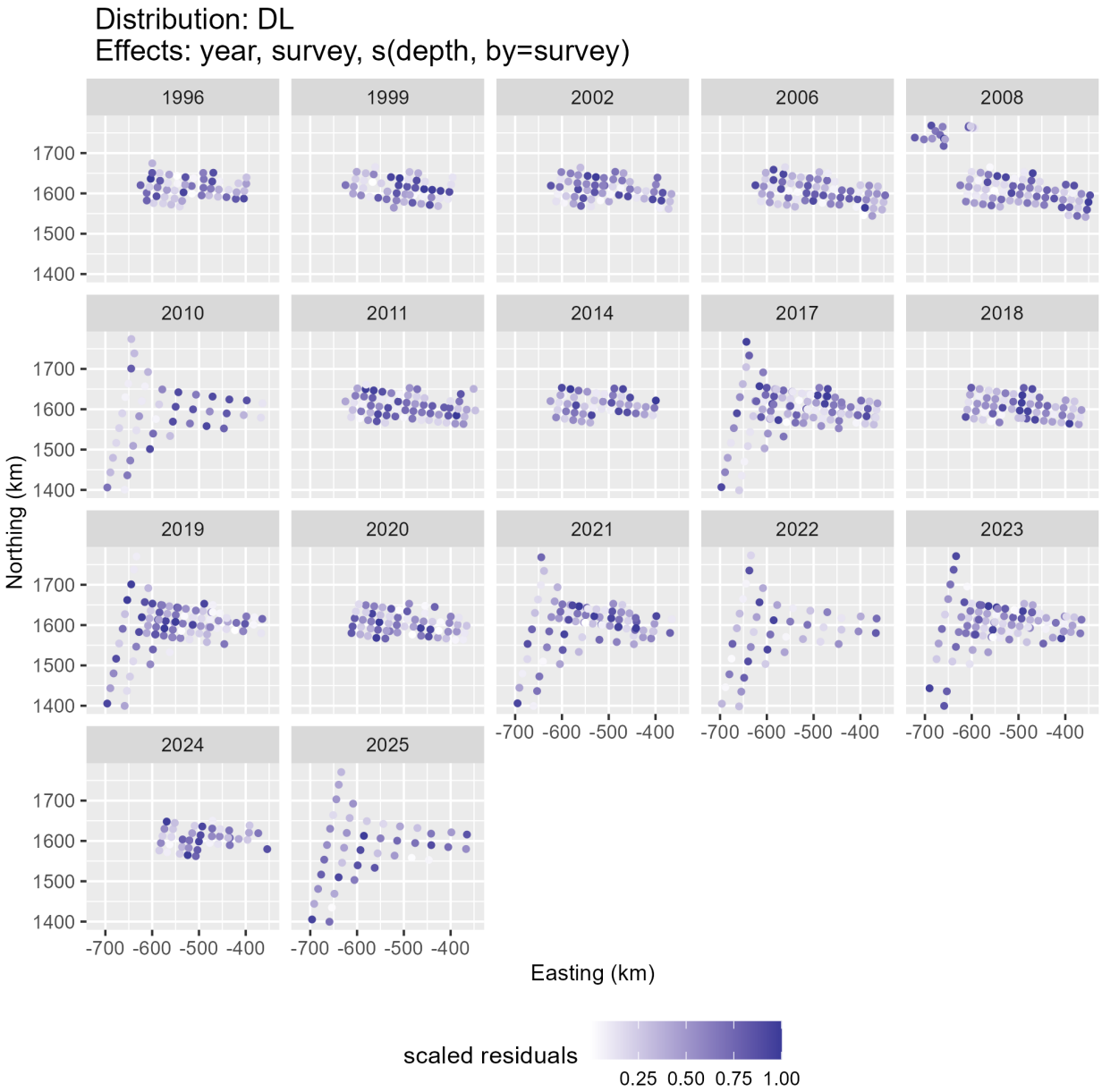


Figure 30: Spatial distribution of DHARMa residuals for the model using the delta lognormal distribution with year and survey as fixed effects and a smooth of depth by survey.

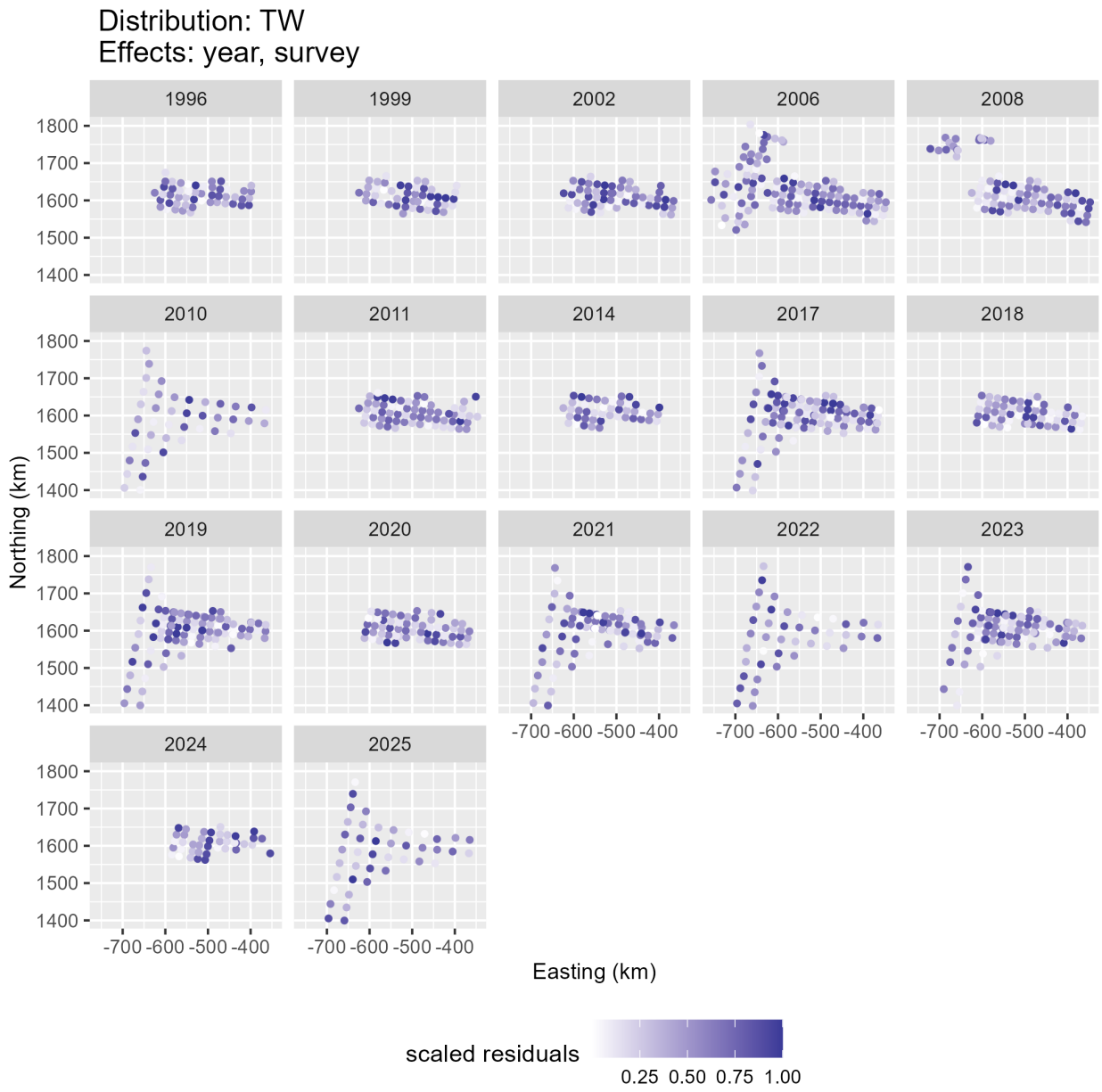


Figure 31: Spatial distribution of DHARMa residuals for the model using the Tweedie distribution with year and survey as fixed effects.

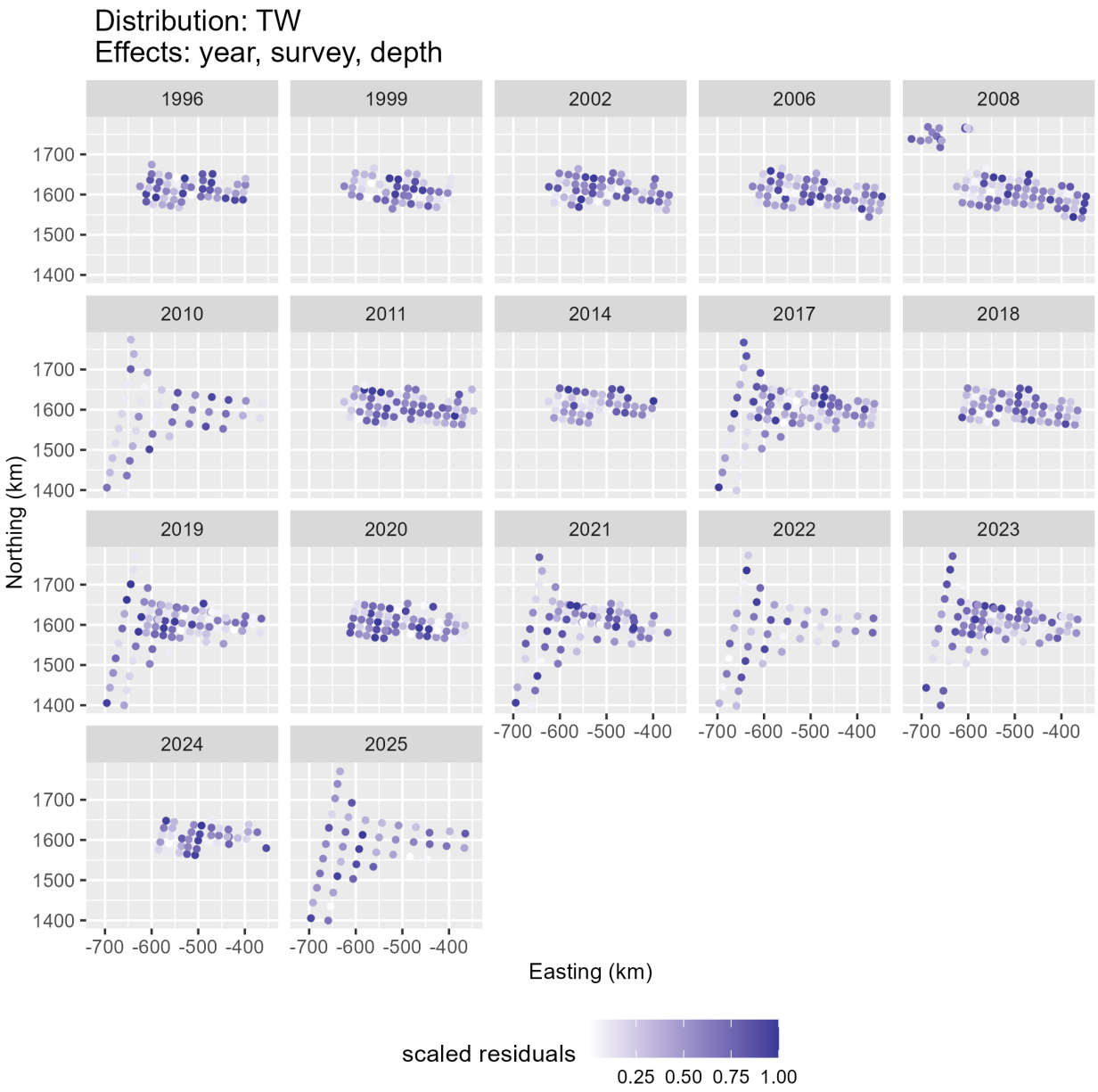


Figure 32: Spatial distribution of DHARMa residuals for the model using the Tweedie distribution with year, survey, and depth as fixed effects.

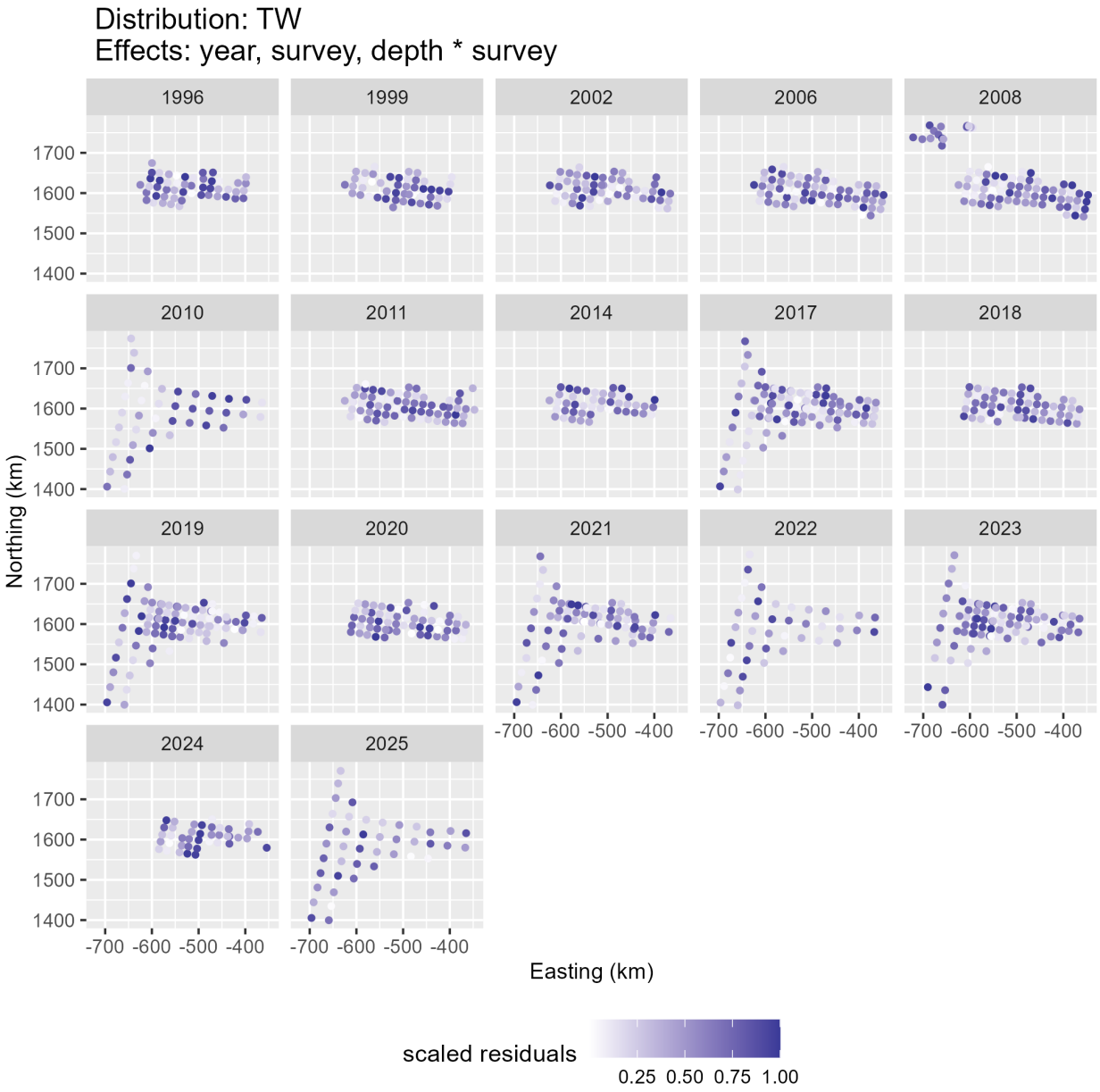


Figure 33: Spatial distribution of DHARMa residuals for the model using the Tweedie distribution with year and survey as fixed effects and a linear interaction between depth and survey.

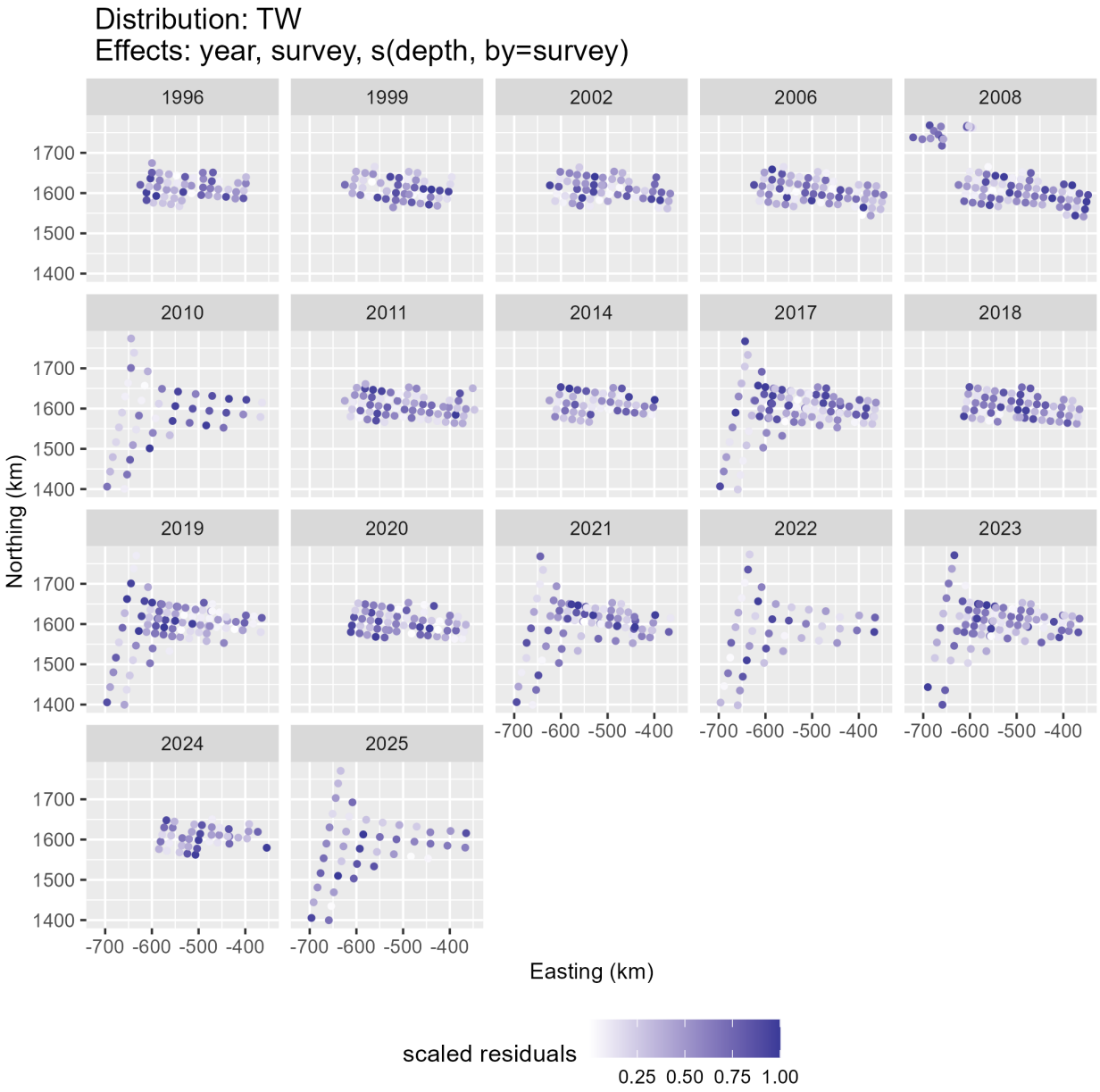


Figure 34: Spatial distribution of DHARMa residuals for the model using the Tweedie distribution with year and survey as fixed effects and a smooth of depth by survey.

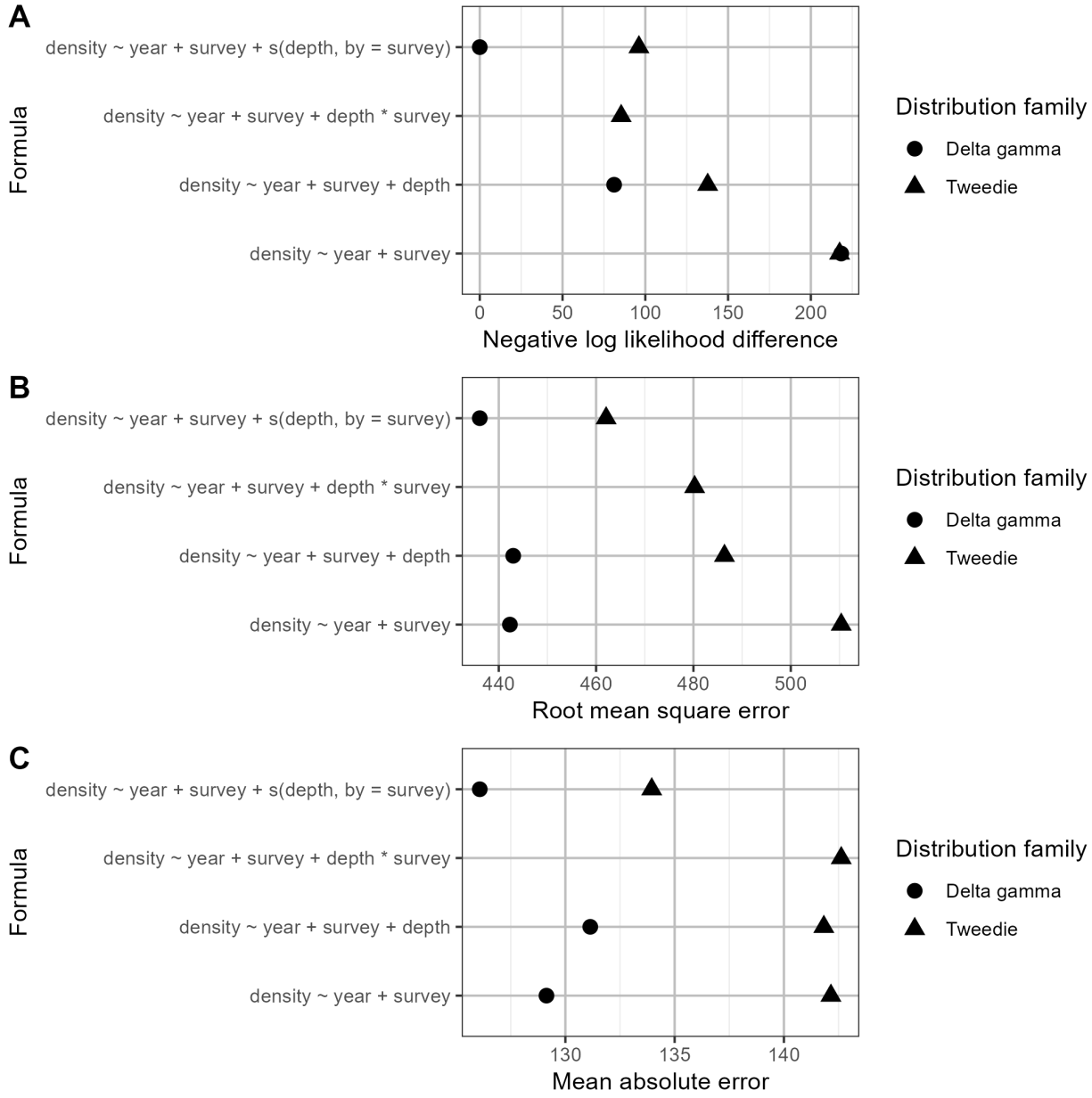


Figure 35: Metrics used in model selection, based on 10-fold cross validation. Note that models using the delta lognormal distribution family are not shown because, for each of those models, not all of the ten folds in the cross-validation analysis had non-positive-definite Hessian matrices. A) Negative log likelihood difference compared to the best-fitting model, which has NLL difference = 0, where smaller values indicate better model fit. B) Root mean square error, a measure of the average difference between observed and predicted values, where a lower value indicates better model fit. C) Mean absolute error, a measure of the average absolute value of the difference between observed and predicted values, where a lower value indicates better model fit.

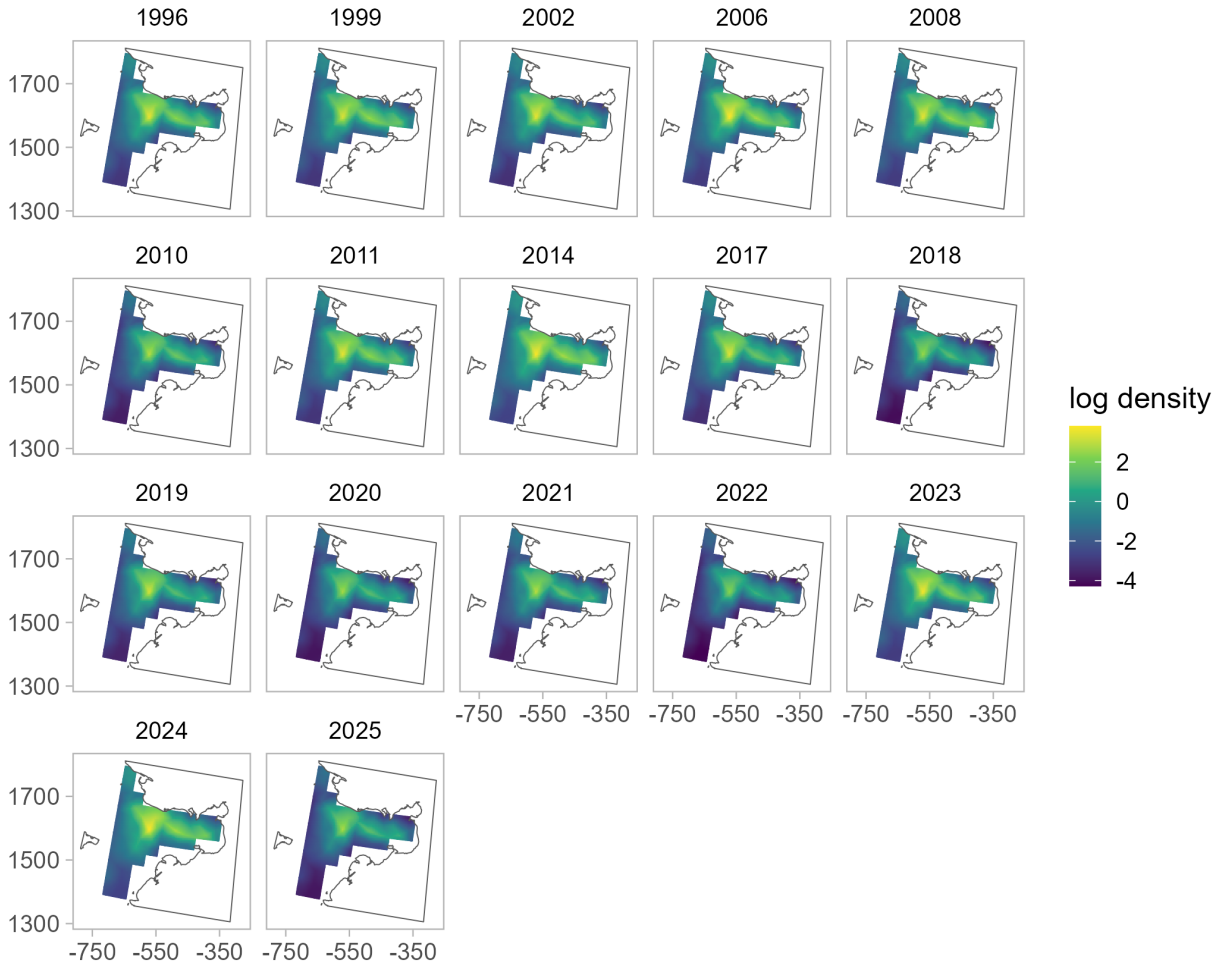


Figure 36: Spatial predictions of Norton Sound red king crab males $\geq$ 64 mm in carapace length by year from the model using the delta gamma distribution with year and survey as fixed effects. These predictions are from the presence/absence component of the hurdle model.

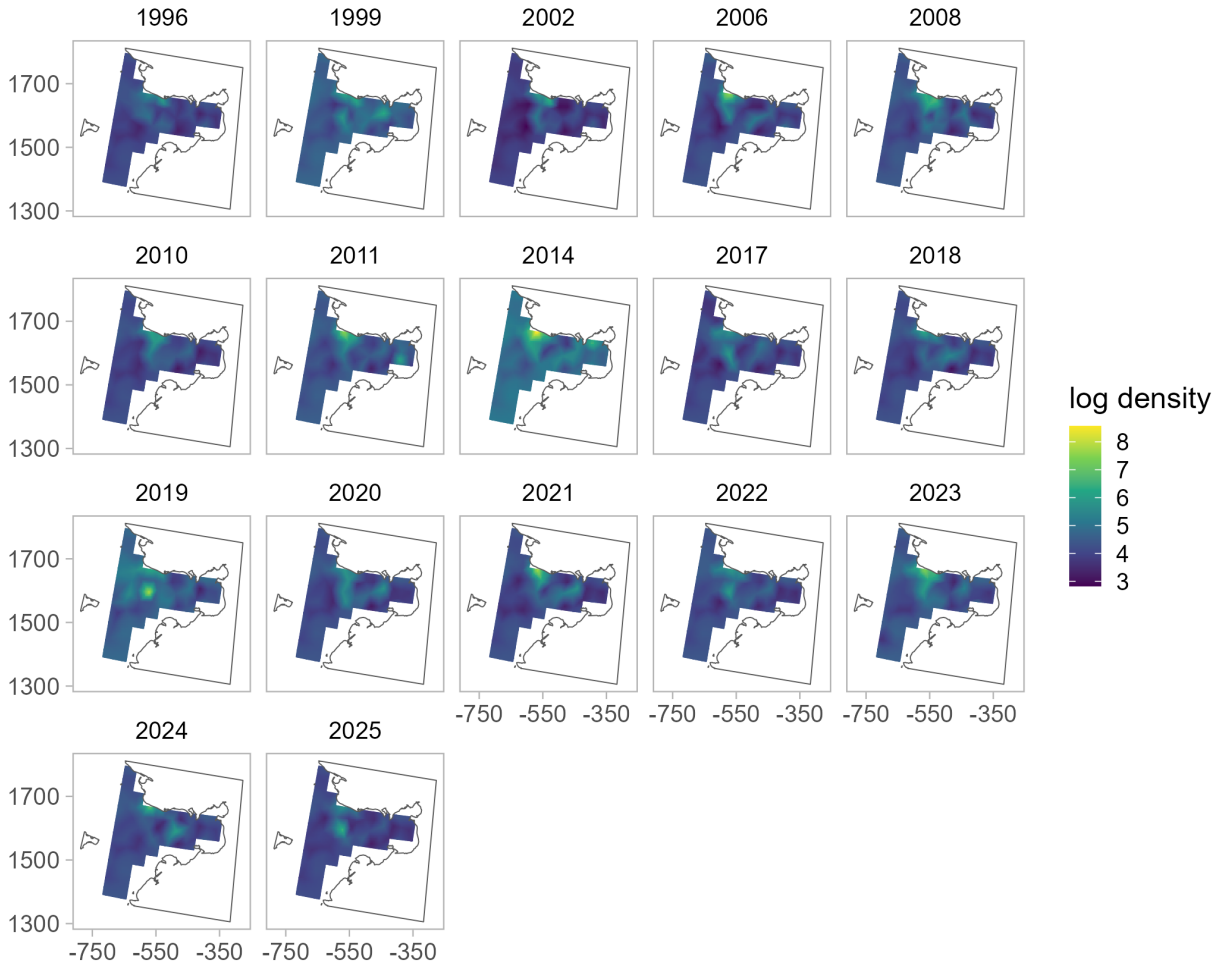


Figure 37: Spatial predictions of Norton Sound red king crab males $\geq$ 64 mm in carapace length by year from the model using the delta gamma distribution with year and survey as fixed effects. These predictions are from the positive catch (abundance) component of the hurdle model.

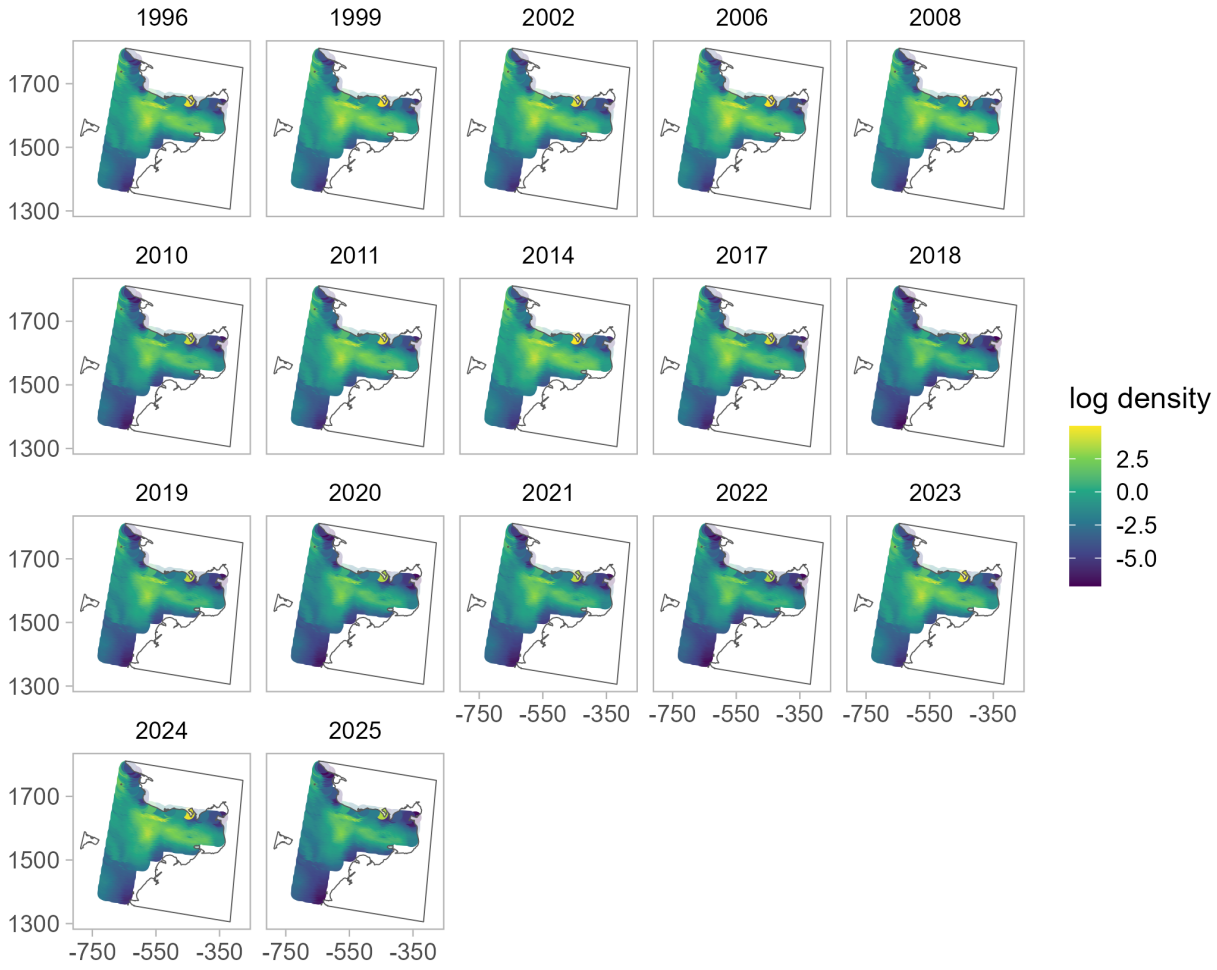


Figure 38: Spatial predictions of Norton Sound red king crab males $\geq$ 64 mm in carapace length by year from the model using the delta gamma distribution with year, survey, and depth as fixed effects. These predictions are from the presence/absence component of the hurdle model.

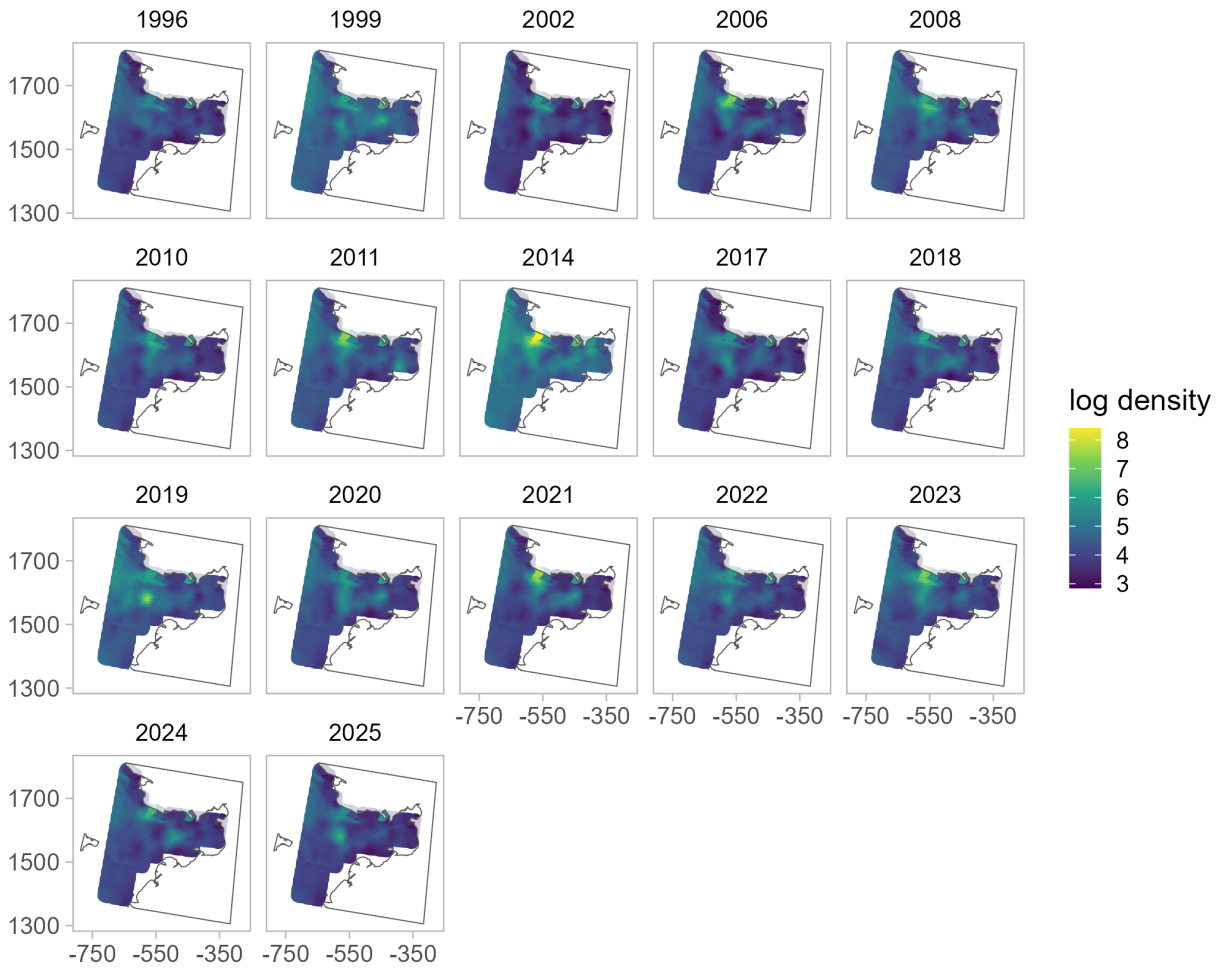


Figure 39: Spatial predictions of Norton Sound red king crab males $\geq$ 64 mm in carapace length by year from the model using the delta gamma distribution with year, survey, and depth as fixed effects. These predictions are from the positive catch (abundance) component of the hurdle model.

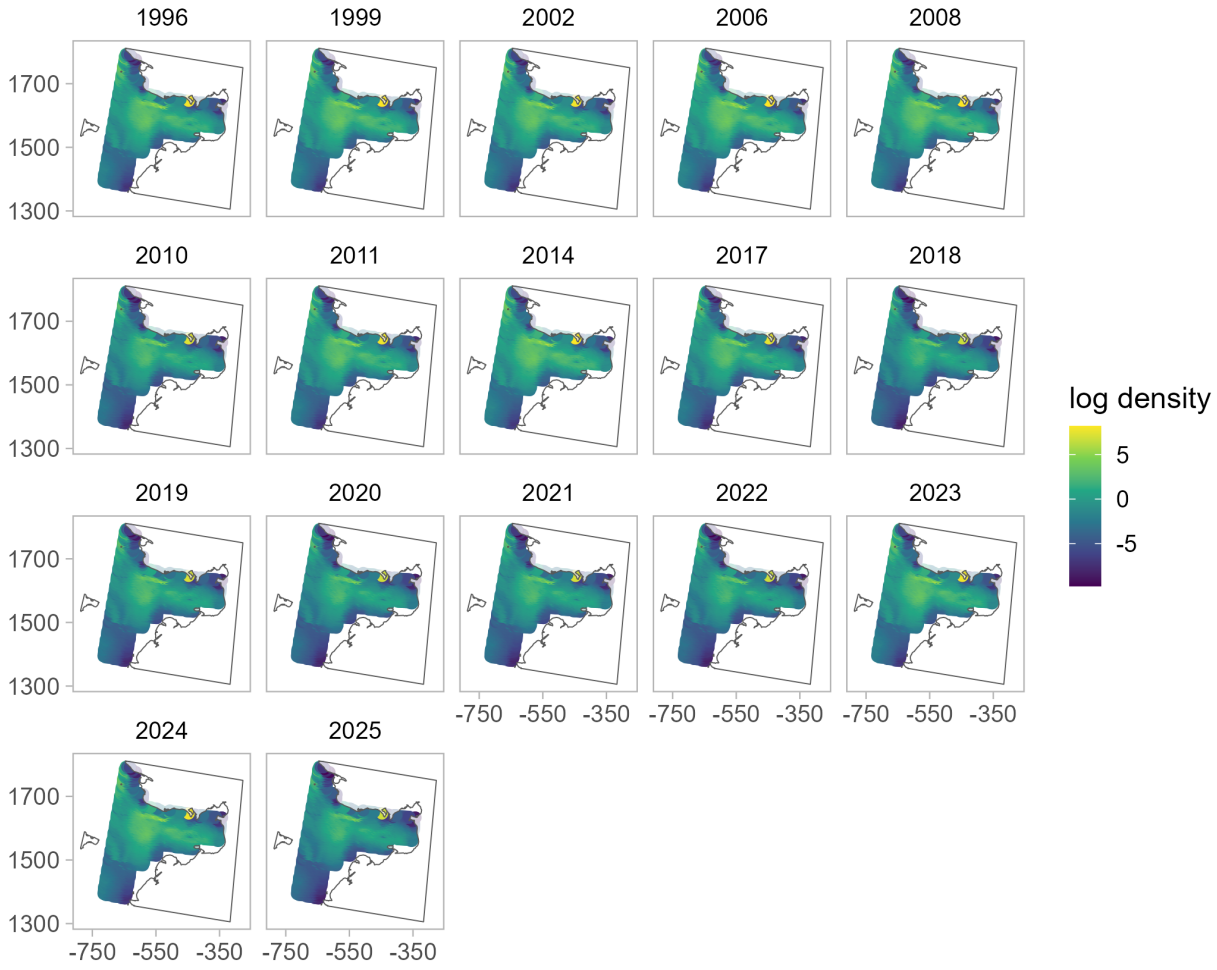


Figure 40: Spatial predictions of Norton Sound red king crab males $\geq$ 64 mm in carapace length by year from the model using the delta gamma distribution with year and survey as fixed effects and a smooth of depth by survey. These predictions are from the presence/absence component of the hurdle model.

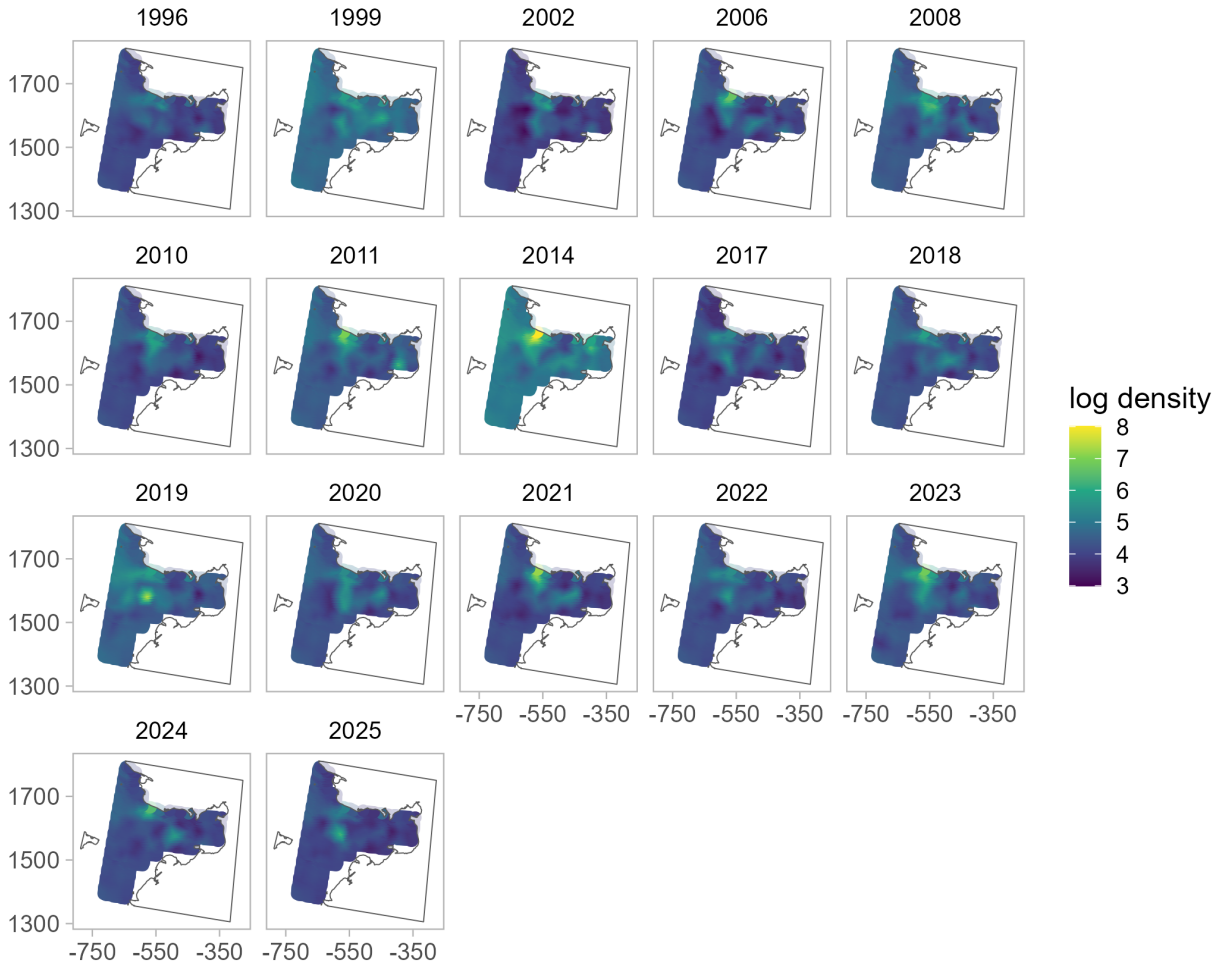


Figure 41: Spatial predictions of Norton Sound red king crab males $\geq$ 64 mm in carapace length by year from the model using the delta gamma distribution with year and survey as fixed effects and a smooth of depth by survey. These predictions are from the positive catch (abundance) component of the hurdle model.

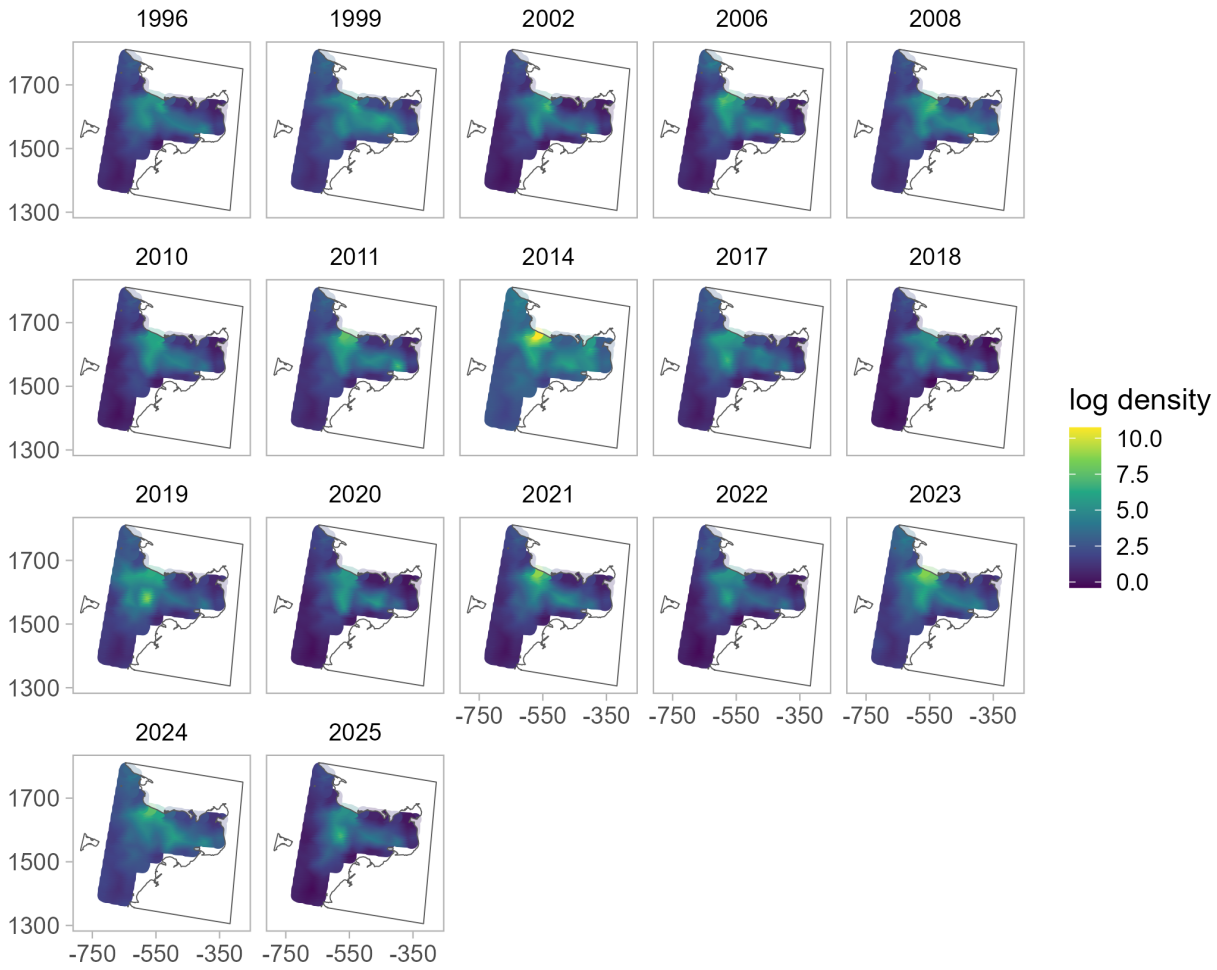


Figure 42: Spatial predictions of Norton Sound red king crab males $\geq$ 64 mm in carapace length by year from the model using the Tweedie distribution with year and survey as fixed effects.

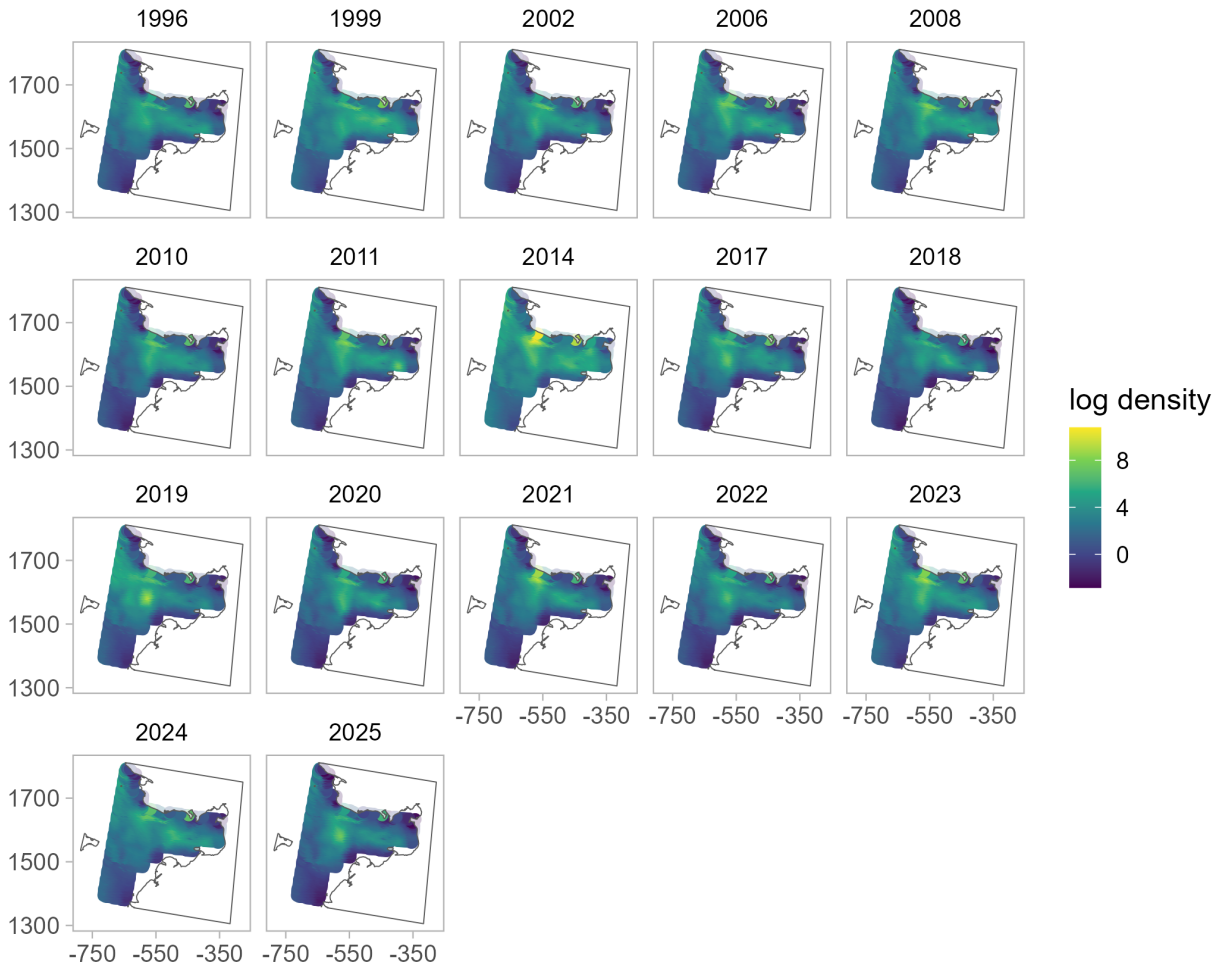


Figure 43: Spatial predictions of Norton Sound red king crab males $\geq$ 64 mm in carapace length by year from the model using the Tweedie distribution with year, survey, and depth as fixed effects.

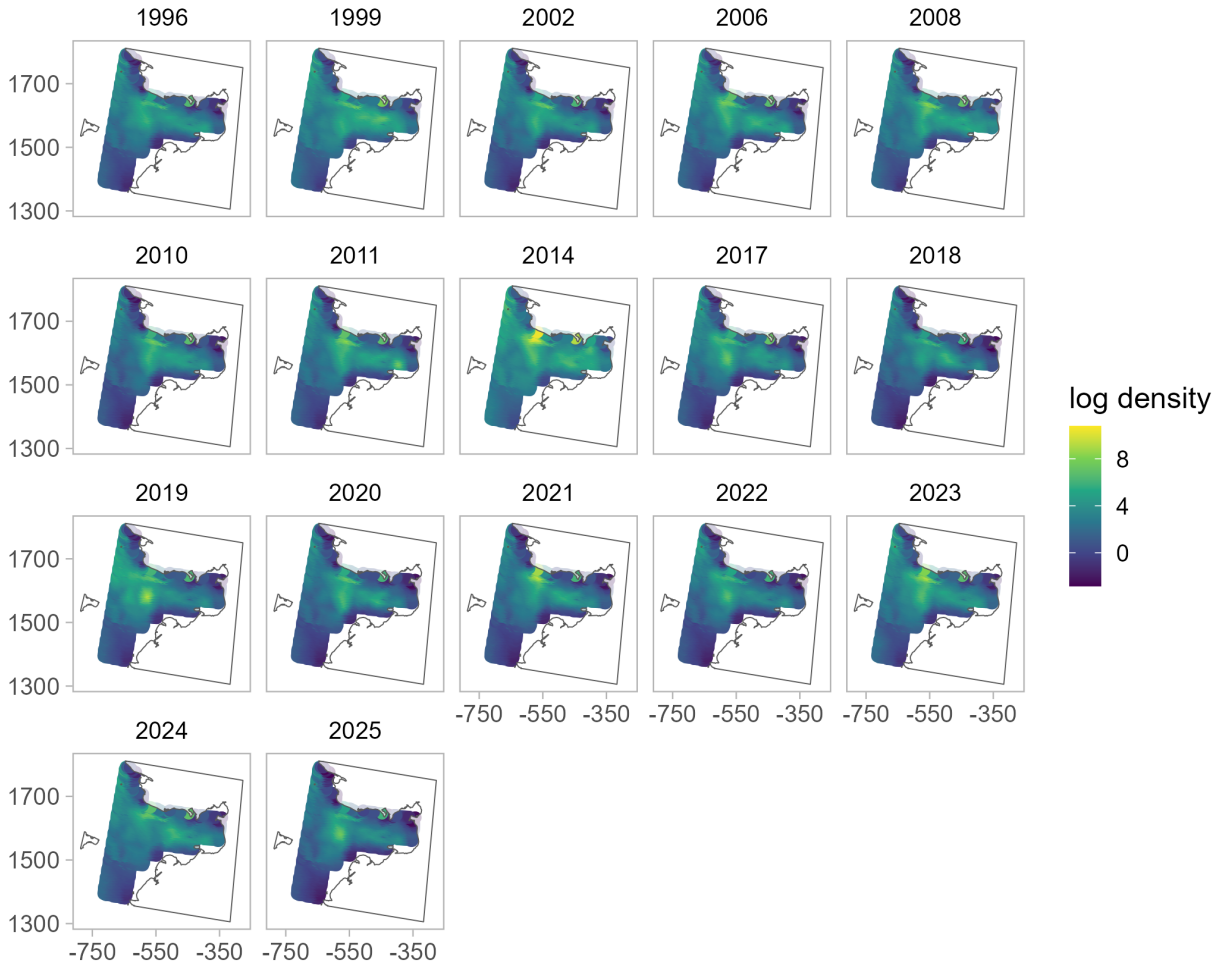


Figure 44: Spatial predictions of Norton Sound red king crab males $\geq$ 64 mm in carapace length by year from the model using the Tweedie distribution with year and survey as fixed effects and a linear interaction between depth and survey.

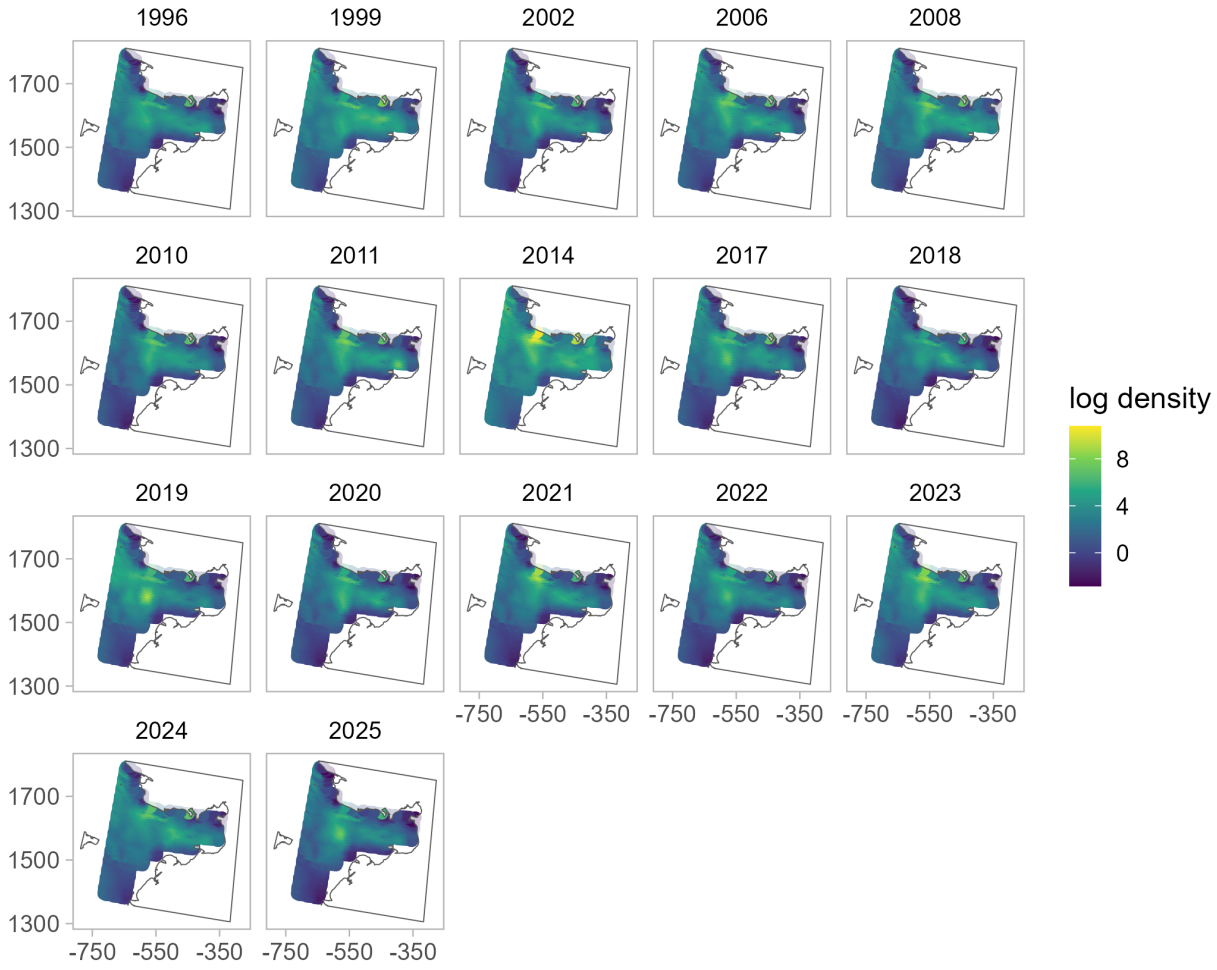


Figure 45: Spatial predictions of Norton Sound red king crab males $\geq$ 64 mm in carapace length by year from the model using the Tweedie distribution with year and survey as fixed effects and a smooth of depth by survey.

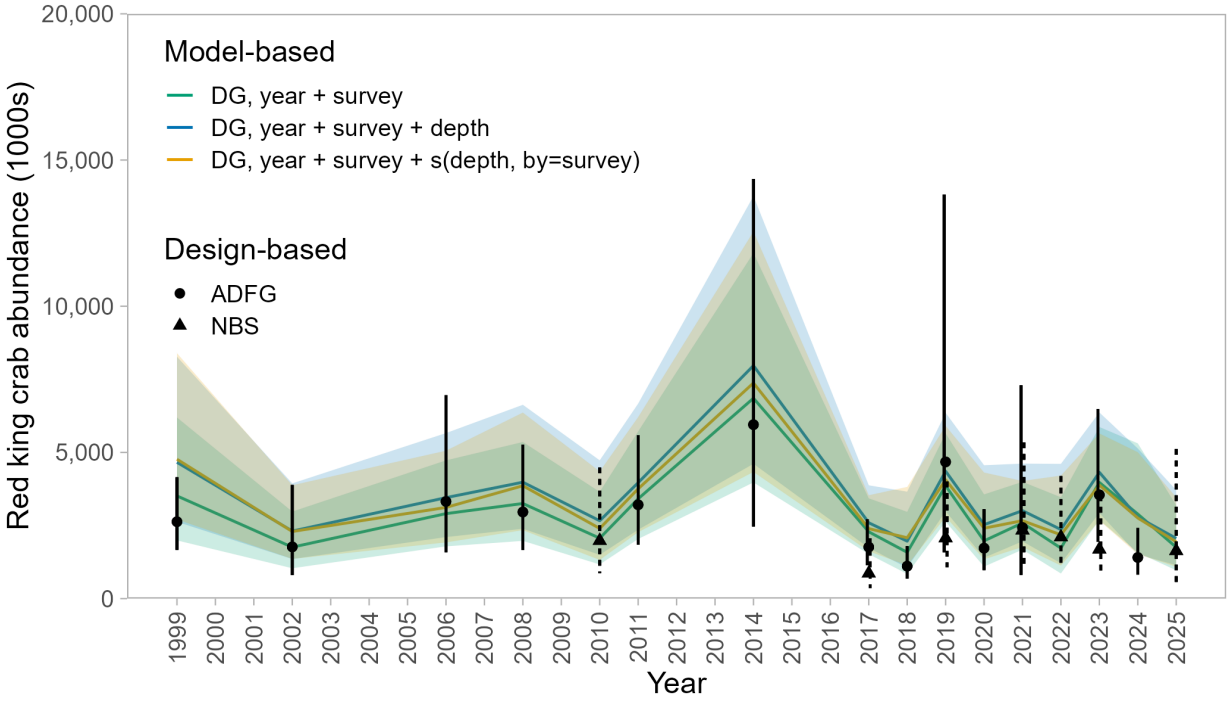


Figure 46: Estimated Norton Sound red king crab abundance for males with carapace length $\geq$ 64 mm. Colored lines represent abundance ($\pm$95% CI) estimated using spatiotemporal models with the delta gamma distribution family and fixed effects year and survey (green); fixed effects year, survey, and depth (blue); or fixed effects year and survey and a smooth of depth by survey (gold). Black points represent design-based abundance ($\pm$95% CI) estimates currently used in the stock assessment model from the NOAA Northern Bering Sea trawl survey (triangles, dashed error bars) and the ADFG trawl survey (circles, solid error bars).

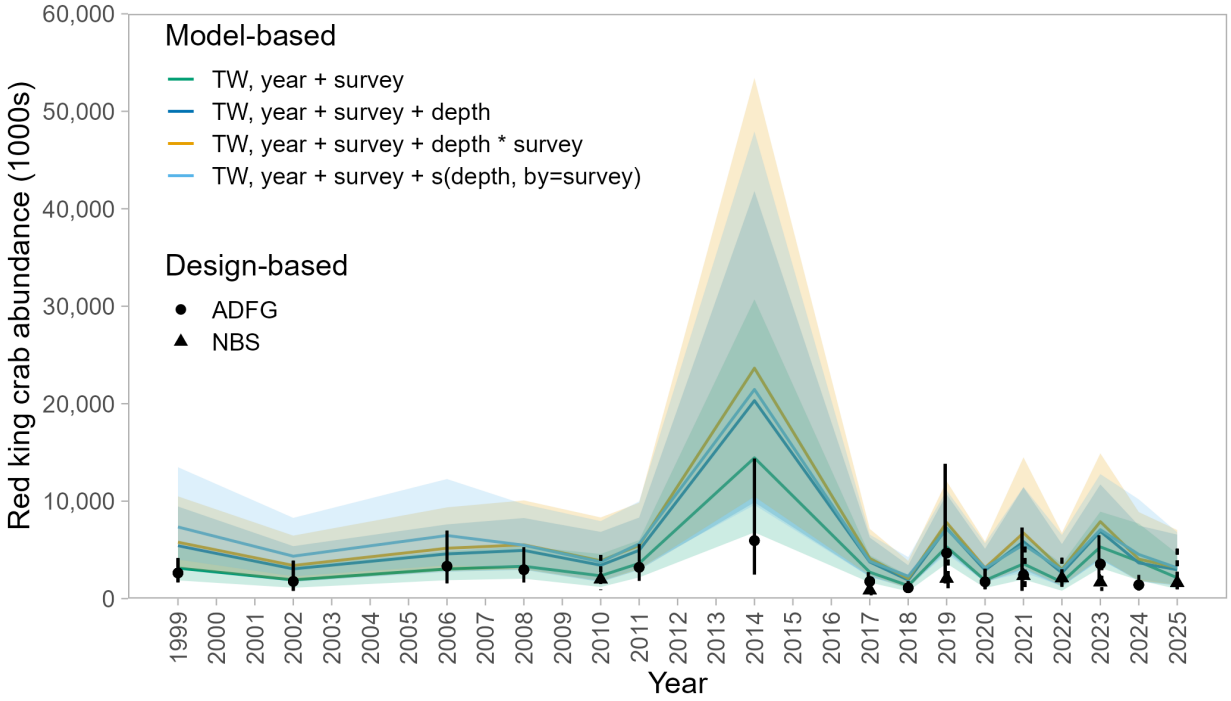


Figure 47: Estimated Norton Sound red king crab abundance for males with carapace length $\geq$ 64 mm. Colored lines represent abundance ($\pm$95% CI) estimated using spatiotemporal models with the Tweedie distribution family and fixed effects year and survey (green); fixed effects year, survey, and depth (blue); fixed effects year and survey and a linear interaction between depth and survey (gold); or fixed effects year and survey and a smooth of depth by survey (light blue). Black points represent design-based abundance ($\pm$95% CI) estimates currently used in the stock assessment model from the NOAA Northern Bering Sea trawl survey (triangles, dashed error bars) and the ADFG trawl survey (circles, solid error bars).