A preliminary assessment for eastern Bering Sea snow crab

Cody Szuwalski and Grant Adams

May 5, 2025

Contents

A. Executive summary

The eastern Bering Sea snow crab population has been at historically low abundances for commercially preferred crab for nearly a decade. Federal reference points are in flux, given status quo proxies allow for removal of all commercially preferred crab. Abundances of small crab are showing signs of recovery after the 2018-2019 marine heatwave and subsequent population collapse. Several topics relevant to the assessment and management of snow crab are considered in this document, including:

GMACS assessment scenarios. Large changes to GMACS have occurred since the version for snow crab was updated. The state has also provided revised catch data for snow crab. Both of these changes produced changes in assessment output. Updated growth data were also included, nearly doubling the number of observation used in the assessment. Three other scenarios responsive to SSC requests related to maturity and selectivity are included here. None of those scenarios addressed issues with reference points. Ultimately the assessment is not contributing to the stock decline because it has never influenced removals, so the topics below were considered to facilitate discussion around how to identify management strategies to address the negative population trends and recent historic lows.

Mortality events drive Bering Sea crab populations. Boom and bust dynamics for Bering Sea crab are common; several assessments require estimating mortality events to fit the survey data. The median estimate of fishable abundance during 2023 across six Bering Sea crab populations was less than 5% of the maximum observed. Most large population declines appear unrelated to crab fisheries; episodic mortality and recruitment failure are the key drivers. Density dependent effects, temperature, and size explain a significant portion of the variability in mortality for all populations examined. Incorporating episodic mortality events into management projections lowers expected yields and biomasses. If these lower expectations for biomass are incorporated into reference points, more fishing would be allowed during periods of low abundance, which is a counter intuitive management response.

Is spatial depletion in the fishery a problem? Assessment of snow crab is performed on fishery-independent data and at the population level. However, spatial catch-per-unit-effort data from the fishery also exist and can be used to evaluate the impact of the fishery on the population from another perspective. We show that there is considerable spatio-temporal heterogeneity in the observed depletion rates, but do not find persuasive evidence to suggest that management should move to a spatially-explicit framework. Discussion around what would be persuasive evidence would be useful from the plan team.

Historical biases in biomass are not a problem. Recent analyses have questioned the use of stock assessment in fisheries management, noting 'historical biases' (i.e. management historically thought the population was larger than it is based on assessment outputs) are more common in stocks that are at relatively low biomasses than those that are not. Eastern Bering Sea crab were included in these analyses and large historical biases were reported. We show that these historical biases were spurious in some instances for EBS crab and largely a non-issue when actual removals were considered.

Where do we go from here? I continue to recommend focusing management for snow crab on the large males because without them, a fishery cannot occur. Effort should be made to bring the state and federal systems together so that 1) the downward trajectories of crab stocks can be addressed (if possible), 2) we no longer provide conflicting management advice, and 3) increasingly limited resources are used efficiently to provide desired management outcomes. Given struggles with sensible reference points within the assessment, working backwards from what has actually occurred (i.e. TACs determined by the state HCR) and attempting to identify reasonable targets for fishing mortality and biomass from there may be more fruitful than continued modeling with no new data to inform uncertain processes like reproductive dynamics.

B. SSC comments

The highest priority would be to continue to refine the Maximin analysis as requested by the SSC in June 2024, specifically using values of steepness of 0.50, 0.67, and 0.80, and considering both the Beverton-Holt and Ricker stock recruitment relationships. The yield analysis also indicated that fishing mortality rates much lower than F35% achieved a high percentage of MSY, indicating potential flexibility in specifying reference points. The SSC suggested that some type of collaborative work during the spring, perhaps including SSC members and/or others might facilitate additional progress on this topic. The SSC is interested in developing a wider range of options for reference points for snow crab for consideration in the next assessment cycle.

This analysis was not extended beyond what was presented last year and would be of limited use until we know more about reproductive contribution of large male snow crab. The specific values for steepness noted by the SSC are arbitrary and do not match observations, estimates, or available data on the reproductive dynamics of snow crab. Until then, we know that the fishery will not exist without large males and we know that the large males are at historical lows, so we should focus our efforts on understanding the dynamics of large males.

The SSC again requests an analysis of the probability of maturing/terminal molt which addresses the observation error in these data and the lack of a monotonically increasing curve. A hierarchical analysis that treats years as random effects might be a starting point. The SSC would also like to better understand the sampling design for the molt data and is concerned about the weighting of the spatial samples in the analysis; weighting should be based on abundance if the sampling rate differs by area (which it would, unless abundance were uniform and/or the targets were in direct proportion to abundance).

The Kodiak lab will give a presentation describing how these data are analyzed. Potential changes to this workflow produces different values for the probability of having undergone terminal molt. Two models runs are included within to explore the impact of different assumptions about the probability of terminal molt.

Investigate whether there is information outside the assessment model (e.g., larval or post-settlement data) or in the model, supporting estimated skewed sex-ratios at recruitment and the mismatch between recent large recruitments for males and females occurring in different years. Explore whether the estimated large differences in male and female recruitment years could be related to the lack of fit to molt-increment data.

No new information has come to light on why this might be the case. Previous discussion centered around the possibility that different spatial distributions by sex of small crab could expose them to different environments, which could alter what we perceive as recruitment five years after settlement. While this difference is an interesting scientific question to pursue, solving it will not address the central problem in management: commercially preferred males are at historic lows over the last decade. Removing the females from the model entirely might be a prudent step towards focusing on the most pressing management problem at hand given conflicting trends.

Geostatistical (e.g. VAST) modeling of trawl survey data including both the NBS and EBS should be prioritized. This could help understand some of the inconsistent recruitment/growth trends observed in recent years as well as prepare for potential changes in stock distribution or productivity under future warming of the Bering Sea. Geostatistical modeling should evaluate alternative error distributions and other model configurations as appropriate.

The Kodiak lab will present indices for mature females and large males.

C. GMACS Assessment scenarios

Models

Eight assessment variants were considered this year:

- 24.1: accepted model from last year
- 25.1: Same data from 24.1, but updated GMACS model
- 25.2: 25.1 + updated catch data from 1990-present provided by ADFG
- 25.3: 25.2 + updated growth data from Kodiak lab
- 25.3a: 25.3 + reference points calculated on commercial biomass instead of morphometrically mature
- 25.4: 25.3 + a single specified molting probability
- 25.4a: 25.3 + a single specified molting probability based on alternate analyses by Kodiak
- 25.5: 25.4 + a stacked logistic curve for survey selectivity

A large number of changes in both structure and formatting have occurred since 2023 when the GMACS version used for snow crab was last updated. For example, functionality was extended to include allowing indices for immature crab, the .CTL file and variable naming schemes were rewritten, the implementation of priors was standardized across parameters, and jitter protocols were revised. For an exhaustive list of changes and bug fixes, see the very end of the .TPL file for the most recent GMACS version on the github repo. All of these changes appear to be improvements, but they did result in some changes to the model fits and management advice. The precise reasons for these changes are unclear currently and will likely be difficult to determine.

In addition to change in GMACS, ADFG analysts provided revised time series for retained and total catches in the directed fishery for snow crab and for other crab fisheries in which snow crab are caught. Since 2017, one year of data at a time were provided from ADFG and then processed and appended to the existing data files for stock assessment. Revisions of the catch data presented here resulted in a number of changes to both the directed fishery data and non-directed fishery data that are input to the assessment (Figure 1). For example, the non-directed fishery data was missing a substantial year of catch of snow crab in the Tanner crab fishery in 2015. Discarded males in the directed fishery also were substantially revised (Figure 1). Part of the difference is the way that the data were input into GMACS. Historically the data were entered into the .DAT file with the assumed discard mortality already applied and the discard mortality specified in the GMACS .DAT file was set to 1. With the updated data, the raw discard data were entered into the .DAT file with a specified discard mortality of 30%. This will facilitate easier sensitivity analyses to specified discard mortalities in the future. While this change to data input explains much of the systematic scaling upward of the male discard data, there are were also differences earlier in the time series that cannot be explained by this change. The standardization of the catch time series is an outstanding step forward and we appreciate the ADFG team making it happen.

Models 25.4, 25.4a, and 25.5 are directed at SSC comments around maturity and survey selectivity. The SSC expressed concern around how the probability of terminally molting was being input as a yearly vector and the dip in selectivity for medium sizes estimated in recent GMACS models. To explore the impact of inputting yearly ogives for the probability of having undergone terminal molt, it was input as a single ogive for all years, derived from the median probability of terminally molting at size across all years of available data. This process was repeated for a preliminary revised data set of the probability of terminal molt provided by the Kodiak lab. To explore the impact of freely estimating a vector of parameters for selectivity at size informed by priors derived from the BSFRF data, we implemented a more 'stiff' stacked logistic selectivity curve that would not allow the dip at medium sizes. Unfortunately, this also makes it more difficult to place priors on the parameters using the BSFRF data.

The final model presented is identical to 25.3, except commercially preferred males are used as the currency of management in reference point calculations (model 25.3a). This is the author's preferred modeling scenario because it focuses management on large males. The author would also recommend further simplifying this model by removing females, similar to the models presented in the appendices used to explore time-varying mortality. The continued decline of large snow crab males is concerning and focusing efforts on understanding the decline in the portion of the population upon which the fishery depends (and reversing declines, if possible) should be the priority of assessment and management. Both changing the currency of management to large males and removing females from the assessment would serve to focus the assessment and management on large males.

Results

Model convergence and comparison

Issues with convergence occurred for all models beyond model 25.3 once additional growth data were incorporated (i.e. no Hessians in spite of small gradients and no parameters on bounds). It is not clear why this occurs, but this will hopefully be resolved by September. The total likelihoods across models are only comparable for a subset because of changes in data and likelihood components (Table 1). While the data for 24.1 and 25.1 were identical, the way that likelihoods, penalties, and priors differ in the updated version of GMACS. Model 25.2 cannot be compared to any other model, but models 25.3, 25.3a, 25.4, and 25.5 all have the same data and underlying likelihood structure. Of these models, 25.3(a) fit the data the best, owing primarily to improvements in fits to the index and size composition data (Table 2). Jittering and retrospective analyses will be provided in September.

Model fits

Fits to the survey indices of mature biomass were similar among models (Figure 2). Updating the version of GMACS resulted in a downward revision of the 2018 mature female biomass estimate. It also shifted the peak in mature male biomass from 2018 to 2019. Altering the probability of terminally molting and survey selectivity functional form impacted magnitudes in the early portion of the time series for MMB and resulted in the inability of the model to capture the sharp uptick in MMB during 2018-19.

Updating the growth data resulted in changes in estimated molt increments given pre-molt carapace width for both males and females (Figure 3). For females, the molt increment at size was systematically lower with updated data compared to historical data. Part of this change comes from the bulk of observations occurring at small pre-molt size, which swamp the signal from the observations occurring at larger sizes. Paring the dataset to be more balanced in the quantity of observations for a given size class might be a useful future exercise. Using a growth model more flexible than a linear model might also be useful. For males, the estimated molt increments were smaller than historical estimates for smaller pre-molt carapace width until approximately 60mm carapace width, when the estimated molt increments began to be larger than historical estimates.

Fits to the non-directed fishery catches were good for both the original and revised data sets (Figure 4). However, in the directed fishery, fits deteriorated somewhat with the updated dataset. Although most of the retained catch data were still fit well, the model was unable to capture several years of discarded data that historically were well fit (e.g. 1990, 1996 and 2019). Given the methodology for producing the discard data from the 1990s for the status quo assessment are unclear, it is not clear if this a bug or a feature of the updated models.

Only small differences were observed among models in the fits to size composition data for retained catch in the directed fishery, with the most recently occuring fishery (2021) showing the largest differences (Figure 5). Fits to the total size composition data from the directed fishery were distinctly different once the growth data were updated in the models (Figure 6). No visually discernible differences among models existed for female discards from the directed fishery (Figure 7). Fits to the non-directed fishery had some of the largest misfits of all data sources (particularly for male data), but this is unsurprising given a single estimated selectivity curve and changes in fishery behavior over time (Figure 8 & Figure 9).

Fits to the survey size composition data were broadly similar among models, except for model 25.5 in which the form of survey selectivity was a stacked logistic (Figure 10 - Figure 17). Model 25.5 produced pronounced peaks not apparent in the observed data around carapace widths of 115 mm, particularly in the early years of the survey time series (Figure 12).

Estimated population processes

The estimated abundance of the commercially preferred size of male crab (arguably the most important output of the assessment) varied among models (Figure 18). The early years during which differences in survey selectivity occurred were particularly different. All models agreed that the abundance in recent years have been the lowest in the time series. The estimates of the most recent era of survey selectivity were fairly similar across models for males, even when the stacked survey selectivity was used (Model 25.5; Figure 19). This is interesting given the prior information from BSFRF was not constraining the parameter estimates for the stacked logistic curve. Estimates for females between 25.5 and all other models were more markedly different.

Estimated fisheries selectivities were similar for all models (Figure 20). Estimated fishing mortality for the directed fishery was highly variable and relative patterns changed among model configurations, particularly as the catch data sets changed (Figure 21). In particular, with the introduction of the updated catch data, estimated fishing mortalities from 2015 to present were higher than estimates from previous models and higher than historical estimates within the same model. This change in relative magnitude of estimated fishing mortality changed the trend from a roughly decreasing trend for historical assessments to an increasing trend for assessment incorporating the updated catch data. Adopting modeling choices like a single molt or stacked selectivity exacerbated this increase.

For all but two models (25.4 and 25.4a), the calculated probability of undergoing terminal molt used to separate crab into immature and mature survey indices were input as data to the assessment (Figure 22). For the two scenarios where a single ogive was input, the median probability of terminal molt at size over years was used. The revised methodology the Kodiak lab will present resulted in somewhat lower probabilities over the mid-sized crab.

Trends in recruitment were very similar across models, with some variability in scale (Figure 23). Differences in estimated recruitment in the most recent years were also observed for males. Models in which a single ogive of probability of terminally molting were used had higher relative estimates of the last two years of recruitments compared to other models.

Base estimates of natural mortality were similar for mature animals of both sexes, but estimates of immature natural mortality were more variable (Figure 24). Similarly, the timing and magnitude of the estimated mortality events in 2018 and 2019 differed widely among model configurations, but all models estimated mortality events that removed at least 80% of crab by sex and maturity state at some point during the marine heatwave of 2018-19.

MMB and management quantities

Estimated MMB time series all had similar trends, but the scales differed by up to ~50% in some years (Figure 25). The scale of MMB when only males > 101 mm carapace width were used in the calculation (model 25.3a) was predictably much lower than when morphometrically mature males were used, but the trends were similar. Differences in estimates and currencies of MMB resulted in large differences in the management targets and advice (Table 3). All models that use morphometrically mature male biomass as the currency of management produced target fishing mortality rates that would allow for the removal of all of the largest male crab. Focusing management on the largest male crab by calculating management quantities based on crab larger than 101mm carapace width resulted in lower fishing mortality targets and would have resulted in a federally closed fishery during 2024.

Recommendations for GMACS in September

I continue to recommend focusing management for snow crab on the large males because without them, a fishery cannot occur and their recent trajectory is alarming. New growth data and revised catch data represent an improvement in the best available science for snow crab. Consequently, I recommend bringing forward model 25.3 and 25.3a for September. The convergence issues (i.e. small gradients, but no Hessian) will hopefully be solved by then.

More broadly, I suggest that effort should be made to bring the state and federal systems together so that 1) the downward trajectories of crab stocks can be addressed (if possible), 2) we no longer provide conflicting management advice, and 3) increasingly limited resources are used efficiently to provide desired management outcomes. Given struggles with sensible reference points within the assessment, working backwards from past management actions and their outcomes (i.e. TACs determined by the state HCR and the resulting stock trajectories) and attempting to identify reasonable targets for fishing mortality and biomass from there may be a useful alternative to the current management strategy.

Table 1: Likelihood components by category and model.

	24.1 Status quo	25.1 update_gmacs	25.2 update catch	25.3 update growth	25.4 single molt	25.5 stacked surv sel	25.3a commercial targets
Catch	-70.01	-30.15	343.97	340.06	297.89	249.02	339.01
Index	107.50	99.60	141.50	141.23	134.46	96.08	129.40
Size	-28170.90	-28158.21	-28019.31	-27881.61	-27802.65	-27136.42	-27861.82
SRR	78.61	168.45	176.15	175.19	189.67	186.00	180.05
Tag	1040.70	1038.96	1042.68	3695.44	3693.80	3701.18	3686.93
Penalties:	1028.53	744.70	749.26	747.11	750.78	508.71	748.94
Priors:	224.19	362.68	371.52	362.27	374.97	439.53	384.10
Total:	-25761.38	-25773.97	-25194.23	-22420.32	-22361.09	-21955.90	-22393.40

Table 2: Individual likelihood components by model.

	24.1 status quo	25.1 update gmacs	25.2 update catch	25.3 update growth	25.4 single molt	25.5 stacked surv_sel	25.3a commercial targets
Retained	-16.66	0.73	124.00	129.44	118.59	100.34	121.61
Discard (male)	69.72	92.19	342.34	333.01	301.68	271.06	339.78
Discard (female)	-69.66	-69.66	-69.66	-69.66	-69.66	-69.66	-69.66
Non-directed	-53.41	-53.41	-52.72	-52.72	-52.72	-52.72	-52.72
NMFS 1982-89 (m)	57.12	56.76	58.67	57.60	46.12	44.48	45.56
NMFS 1989-present(m)	-3.59	-2.03	-1.21	1.79	-1.45	-13.41	-1.18
NMFS 1982-88 (f)	50.98	44.94	81.56	78.59	72.74	39.21	79.29
NMFS 1989-present (f)	2.99	-0.07	2.48	3.25	17.04	25.80	5.73
Retained SC	-3634.65	-3642.00	-3612.16	-3567.45	-3546.24	-3533.01	-3578.06
Total SC	-2647.49	-2639.11	-2603.79	-2567.99	-2560.98	-2552.95	-2558.82
Discard SC (f)	-2270.66	-2269.03	-2270.43	-2269.81	-2267.92	-2269.14	-2267.21
Bycatch SC (f)	-2537.19	-2531.17	-2531.78	-2521.63	-2521.87	-2555.34	-2526.04
Bycatch SC (m)	-2418.24	-2435.81	-2370.17	-2351.62	-2340.47	-2332.57	-2369.11
NMFS 1982-89 (imm m)	-630.01	-629.74	-633.22	-630.89	-612.35	-576.37	-600.94
NMFS 1989-present(imm m)	-3135.12	-3143.32	-3126.33	-3115.62	-3128.02	-2908.10	-3129.50
NMFS 1982-88 (imm f)	-541.67	-540.85	-550.24	-550.05	-558.70	-497.39	-551.02
NMFS 1989-present (imm f)	-2915.90	-2855.88	-2866.97	-2866.30	-2871.32	-2768.44	-2859.54
NMFS 1982-89 (mat m)	-683.53	-684.08	-685.18	-684.31	-670.15	-657.39	-670.24
NMFS 1989-present(mat m)	-3312.09	-3314.12	-3314.17	-3315.02	-3313.45	-3117.85	-3314.09
NMFS 1982-88 (mat f)	-583.09	-584.21	-586.20	-584.37	-586.79	-568.48	-584.62
NMFS 1989-present (mat f)	-2861.27	-2888.90	-2868.66	-2856.54	-2824.41	-2799.40	-2852.63
Growth	1040.70	1038.96	1042.68	3695.44	3693.80	3701.18	3686.93

Table 3: Management quantities derived from maximum likelihood estimates by model using Tier 3 reference points. Reported natural mortality is for mature males, average recruitment is for males, and status and MMB were estimates for February 15 of the completed crab year.

	BMSY	Bcurr/BMSY	OFL(tot)	Fmsy	Fofl
24.1 Status quo	191.81	0.56	19.60	49.63	25.07
25.1 update_gmacs	173.74	0.54	21.97	38.37	18.90
25.2 update catch	158.37	0.49	20.76	49.46	21.38
25.3 update growth	143.24	0.45	20.20	54.12	20.97
25.4 single molt	157.14	0.52	19.00	62.78	29.07
25.5 stacked surv sel	141.96	0.59	14.14	37.32	20.21
25.3a commercial targets	79.43	0.20	0.23	0.80	0.00
25.4a single molt (rev)	131.02	0.49	14.44	18.58	8.08

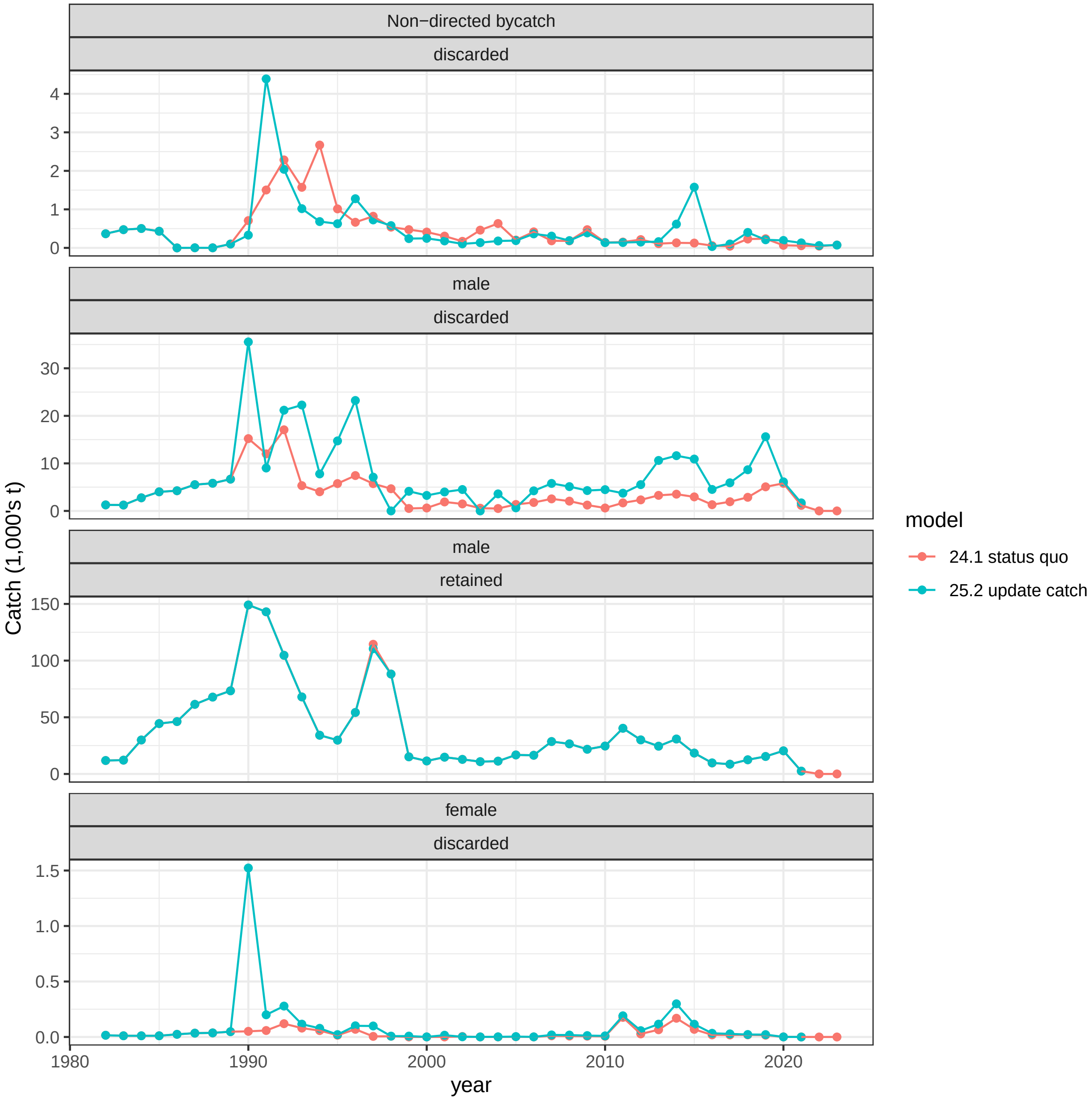

Figure 1: Model fits to catch data.

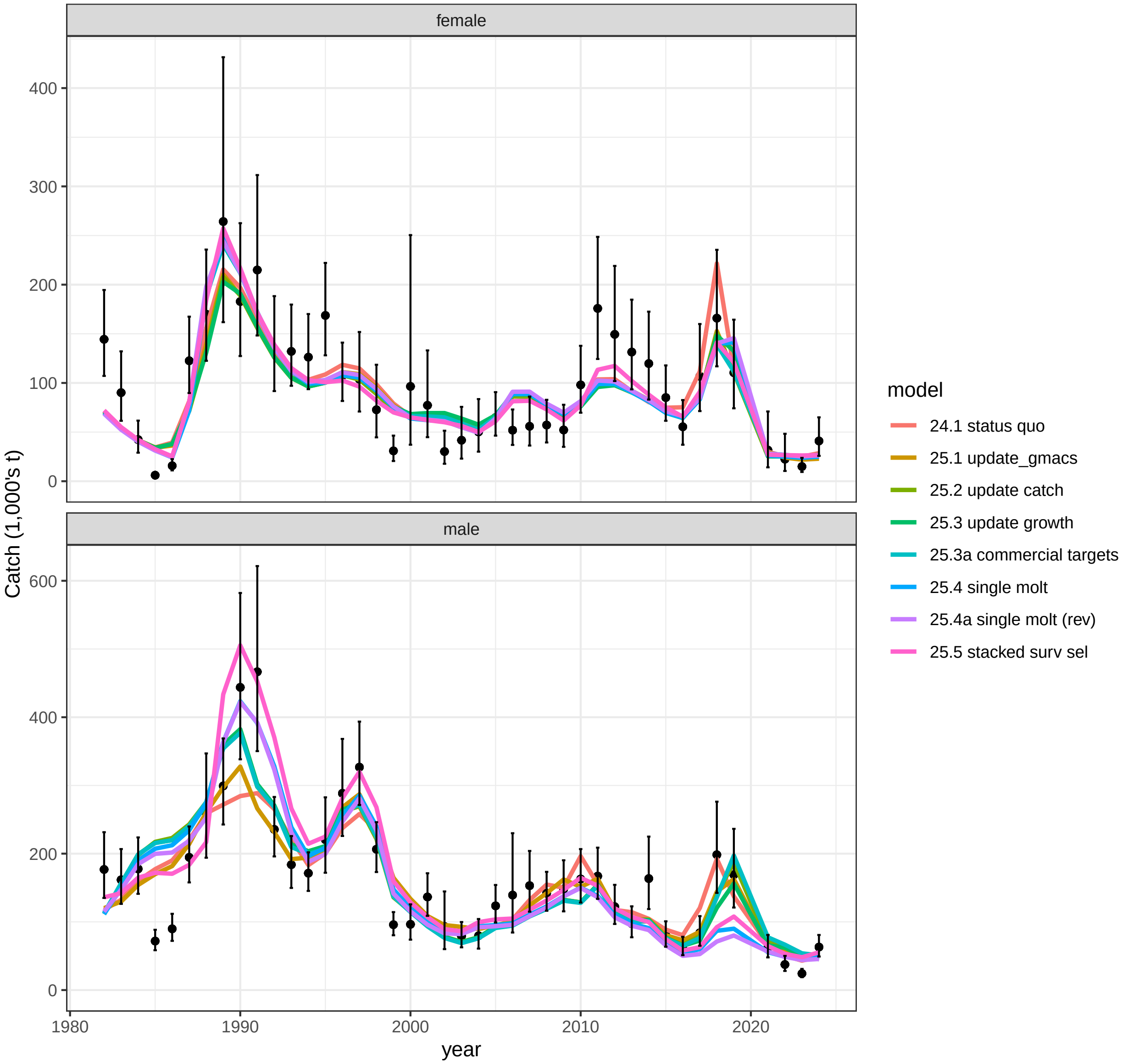

Figure 2: Model fits to the observed mature biomass at survey.

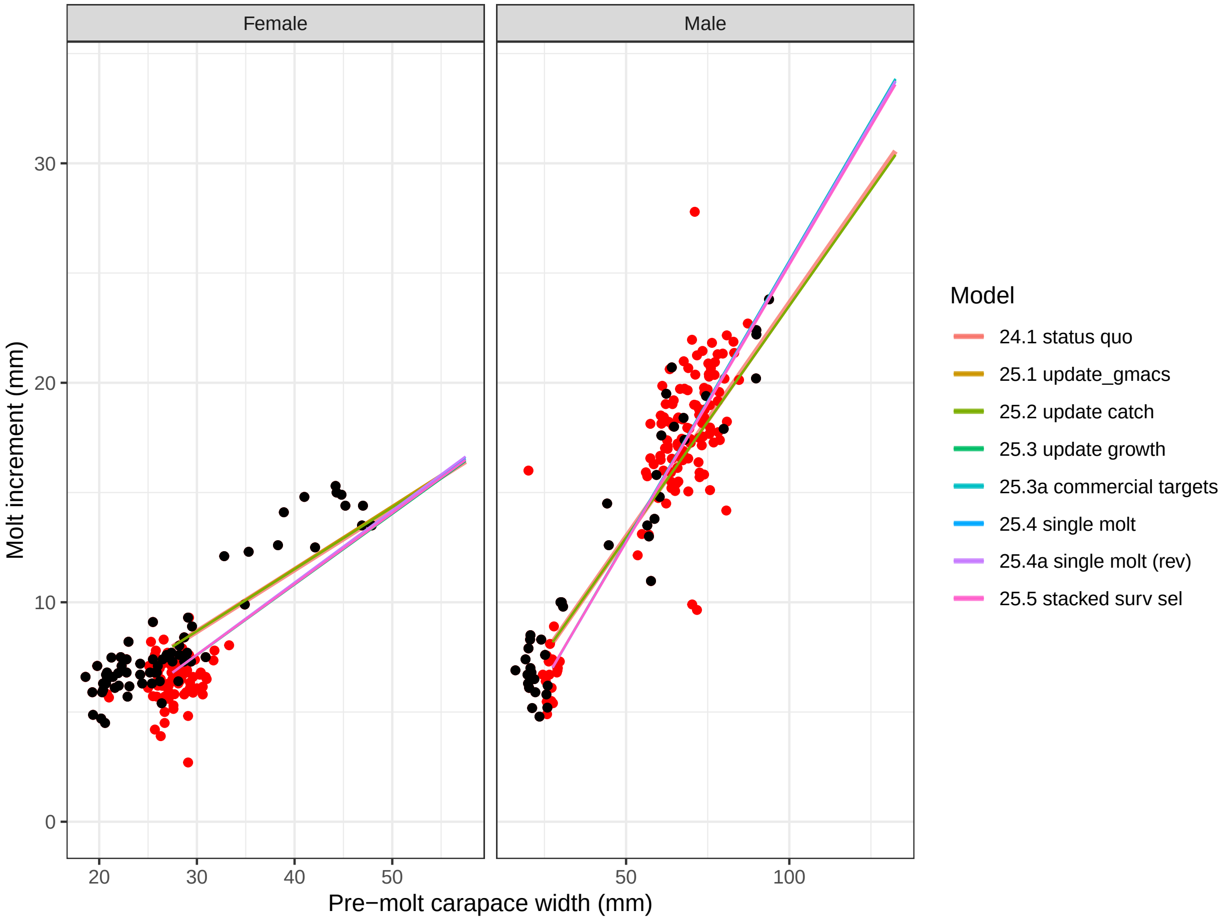

Figure 3: Model fits (colored lines) to the growth data (dots). Black dots are historical observations; red dots are new data for 2025.

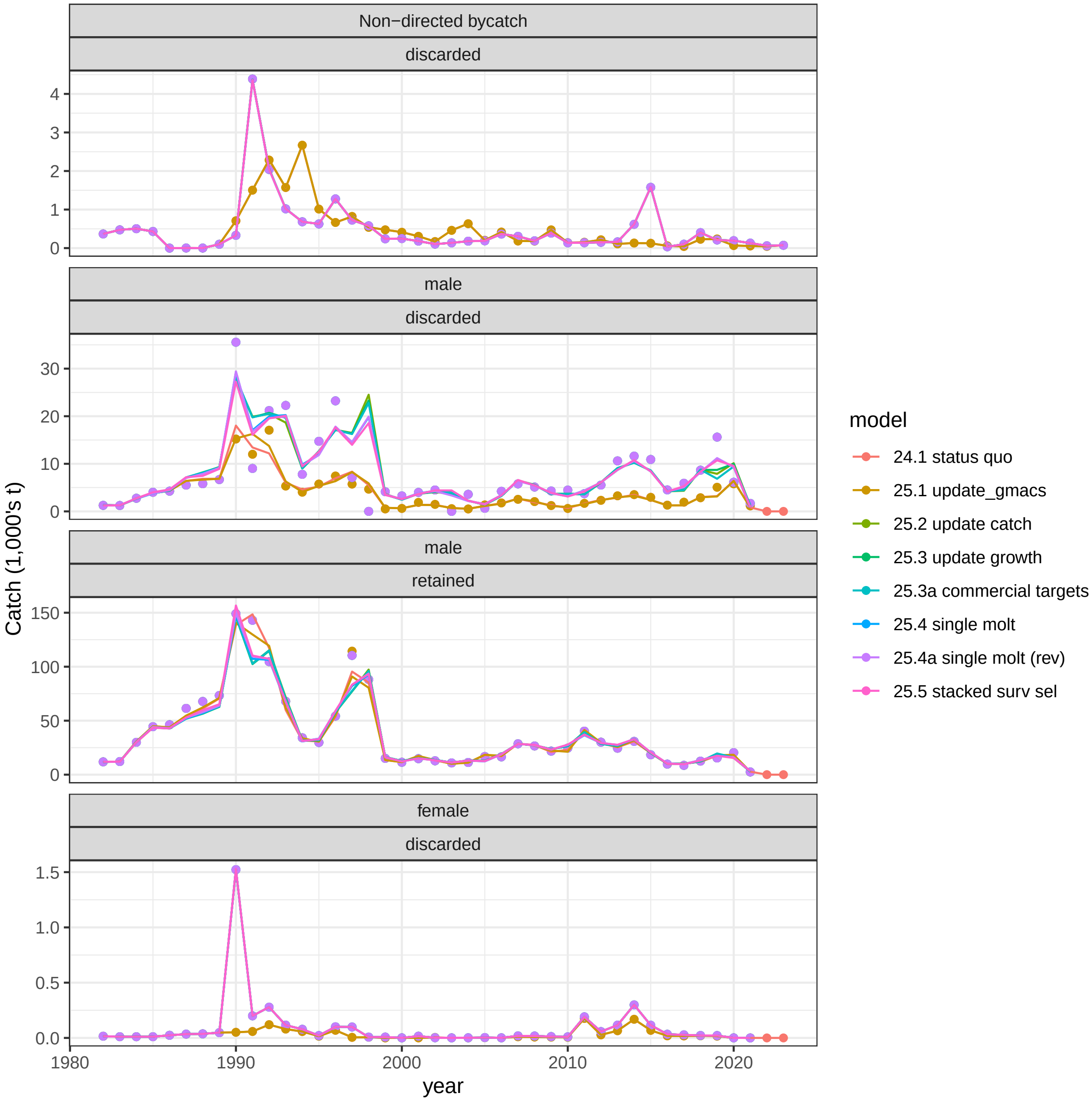

Figure 4: Model fits to catch data.

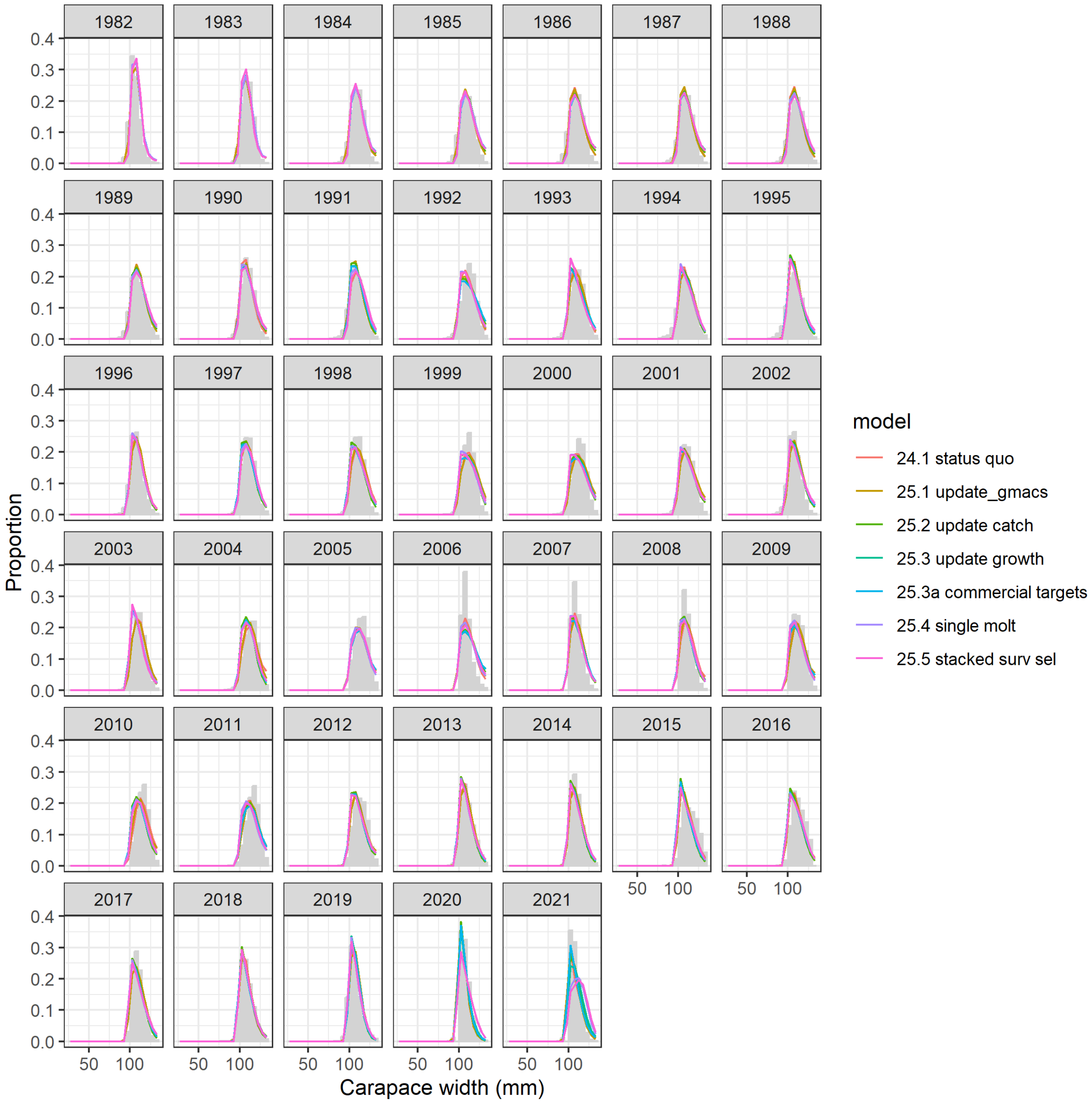

Figure 5: Model fits (lines) to the retained catch size composition data (grey bars).

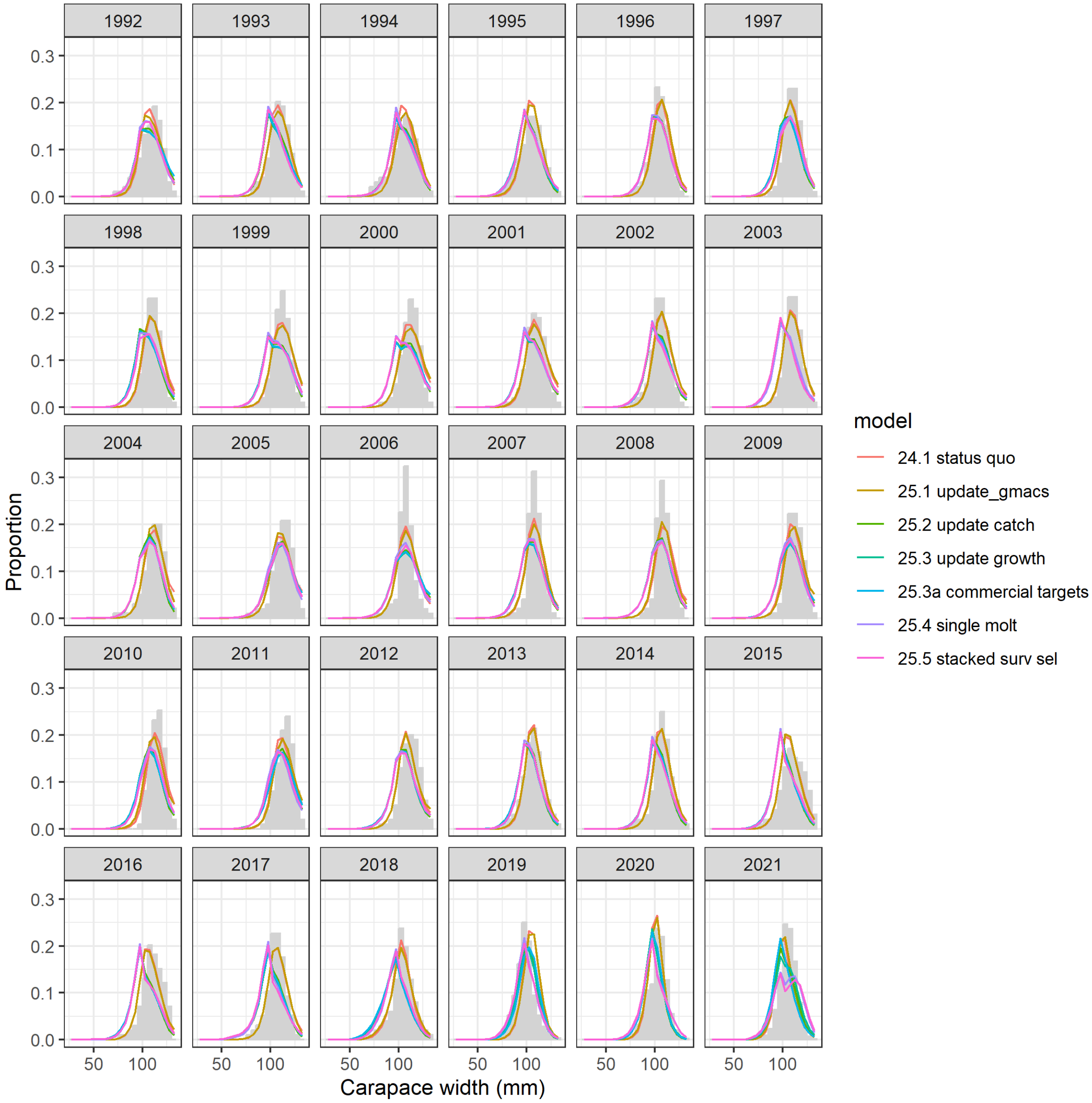

Figure 6: Model fits (lines) to the total catch size composition data (grey bars).

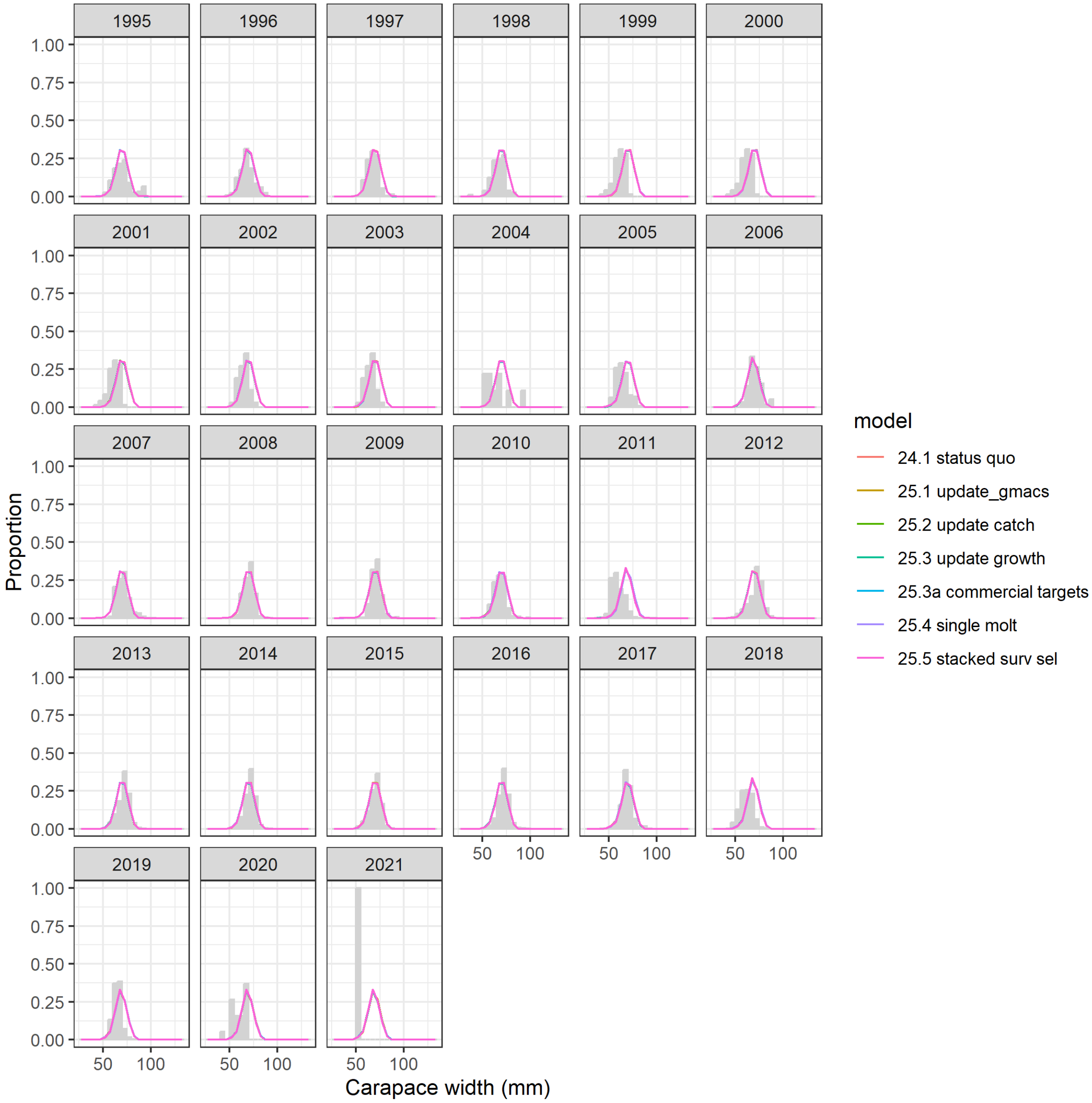

Figure 7: Model fits (lines) to the female discard size composition data (grey bars).

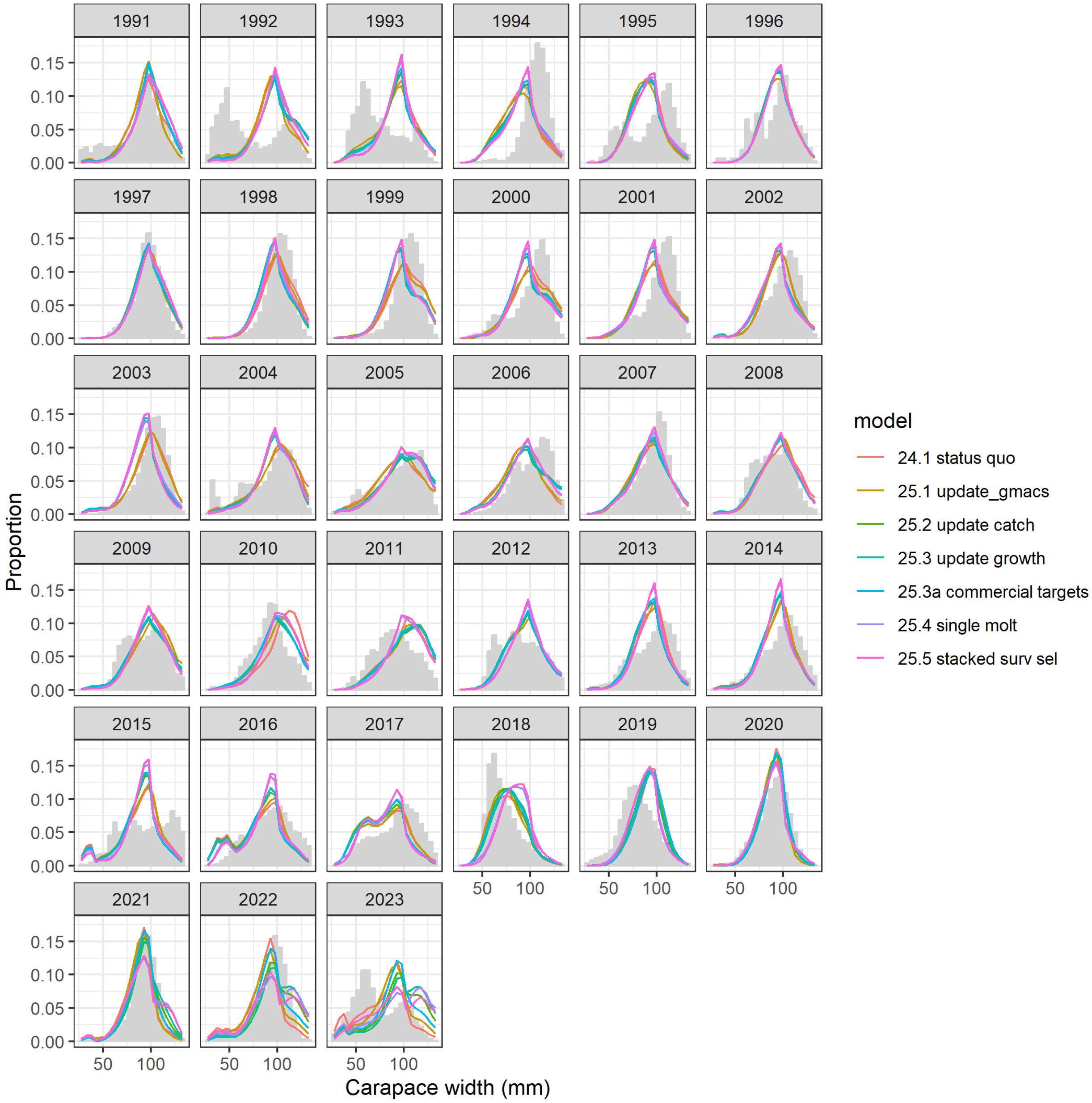

Figure 8: Model fits (lines) to the male non-directed fishery size composition data (grey bars).

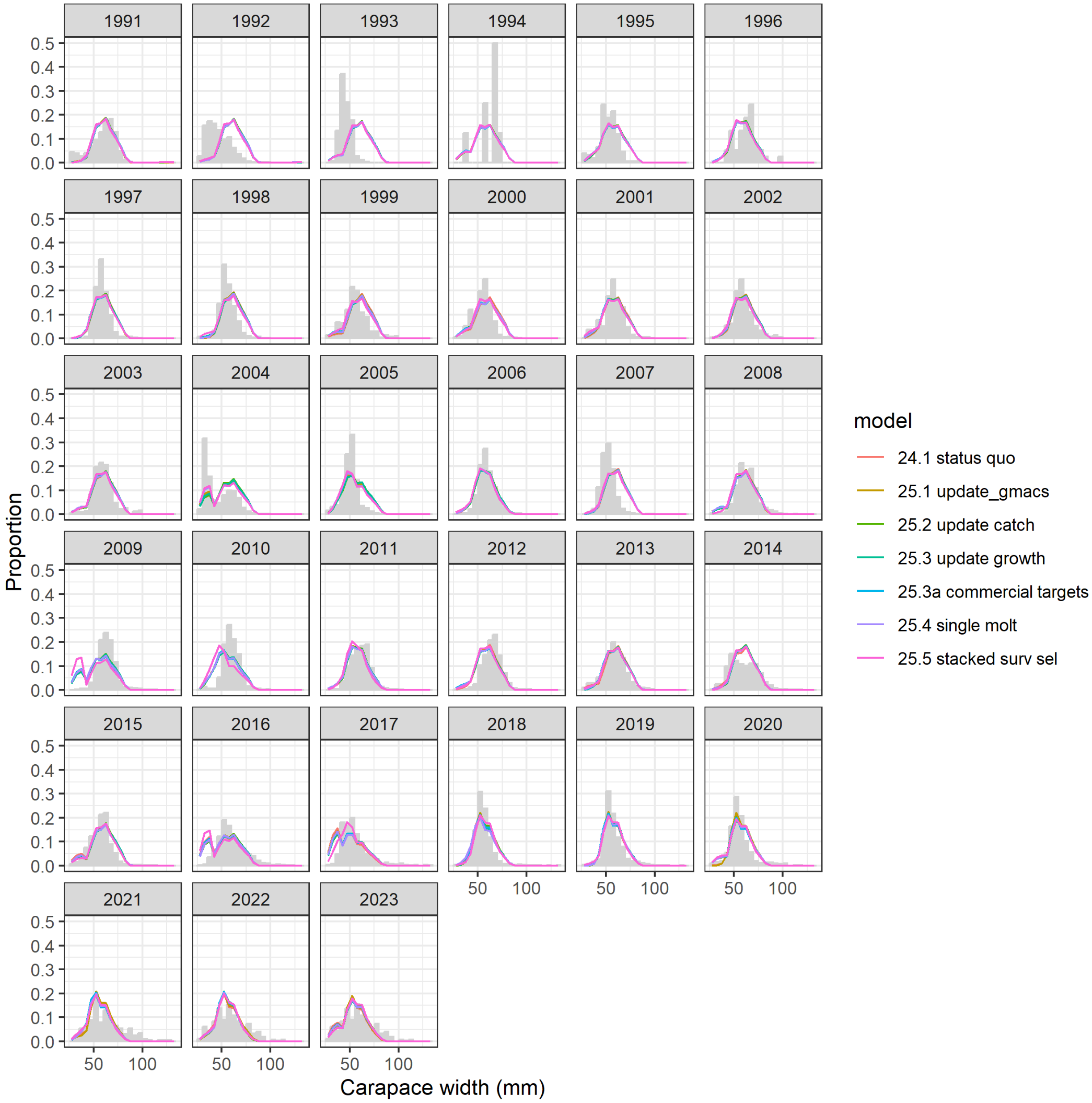

Figure 9: Model fits (lines) to the female non-directed size composition data (grey bars).

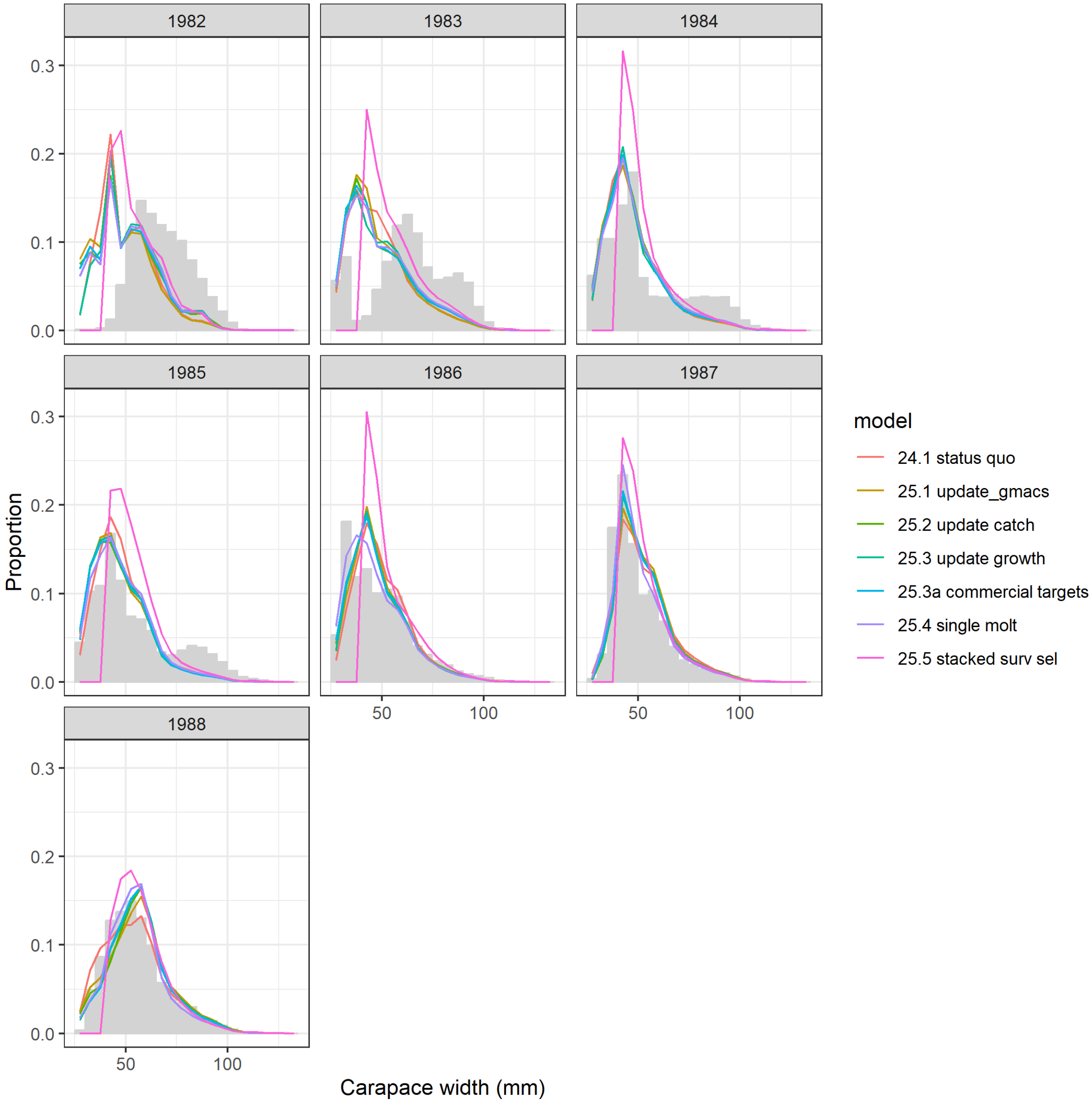

Figure 10: Model fits to immature male survey size composition data from 1982-1988.

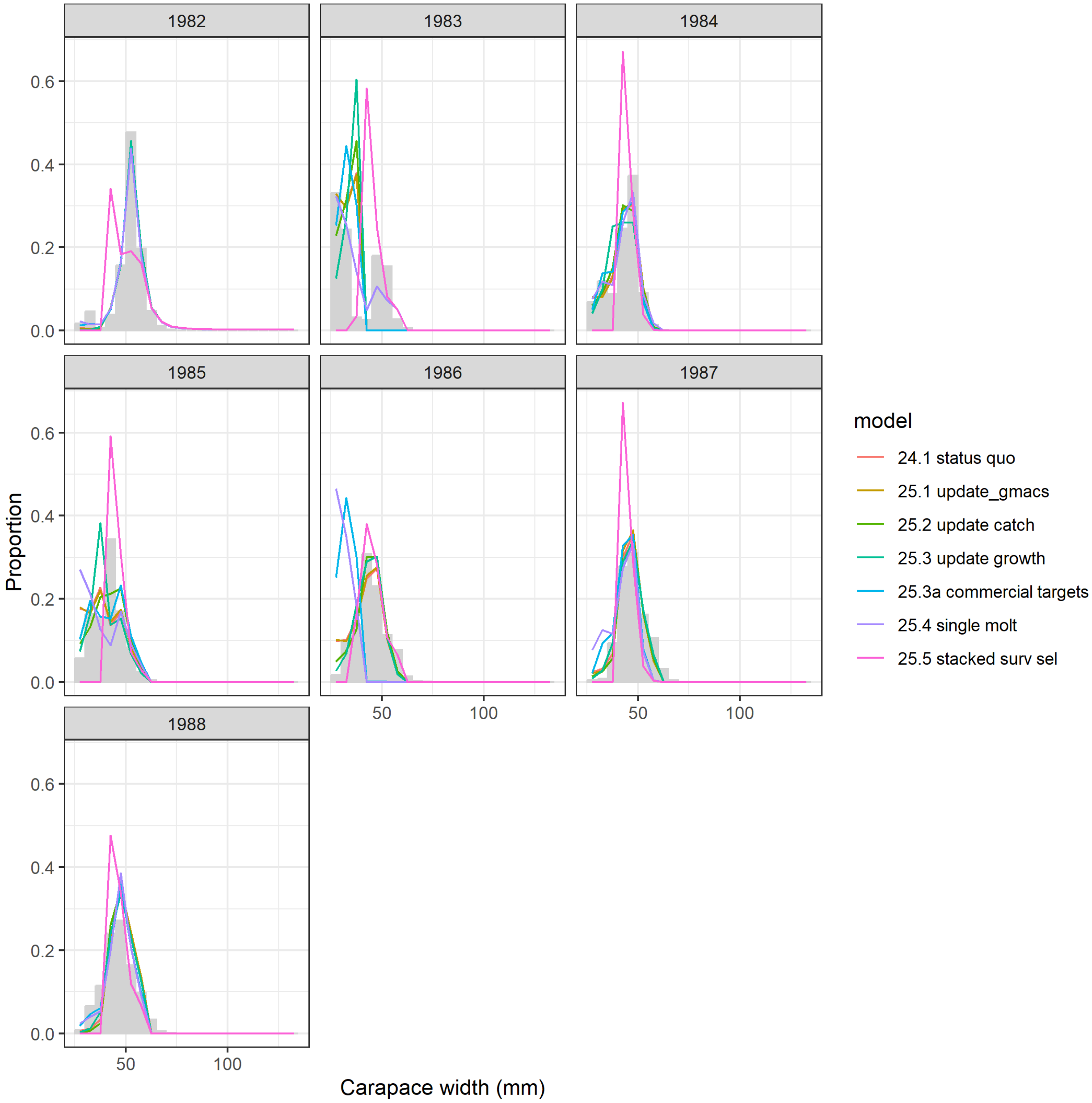

Figure 11: Model fits to immature female survey size composition data from 1982-1988.

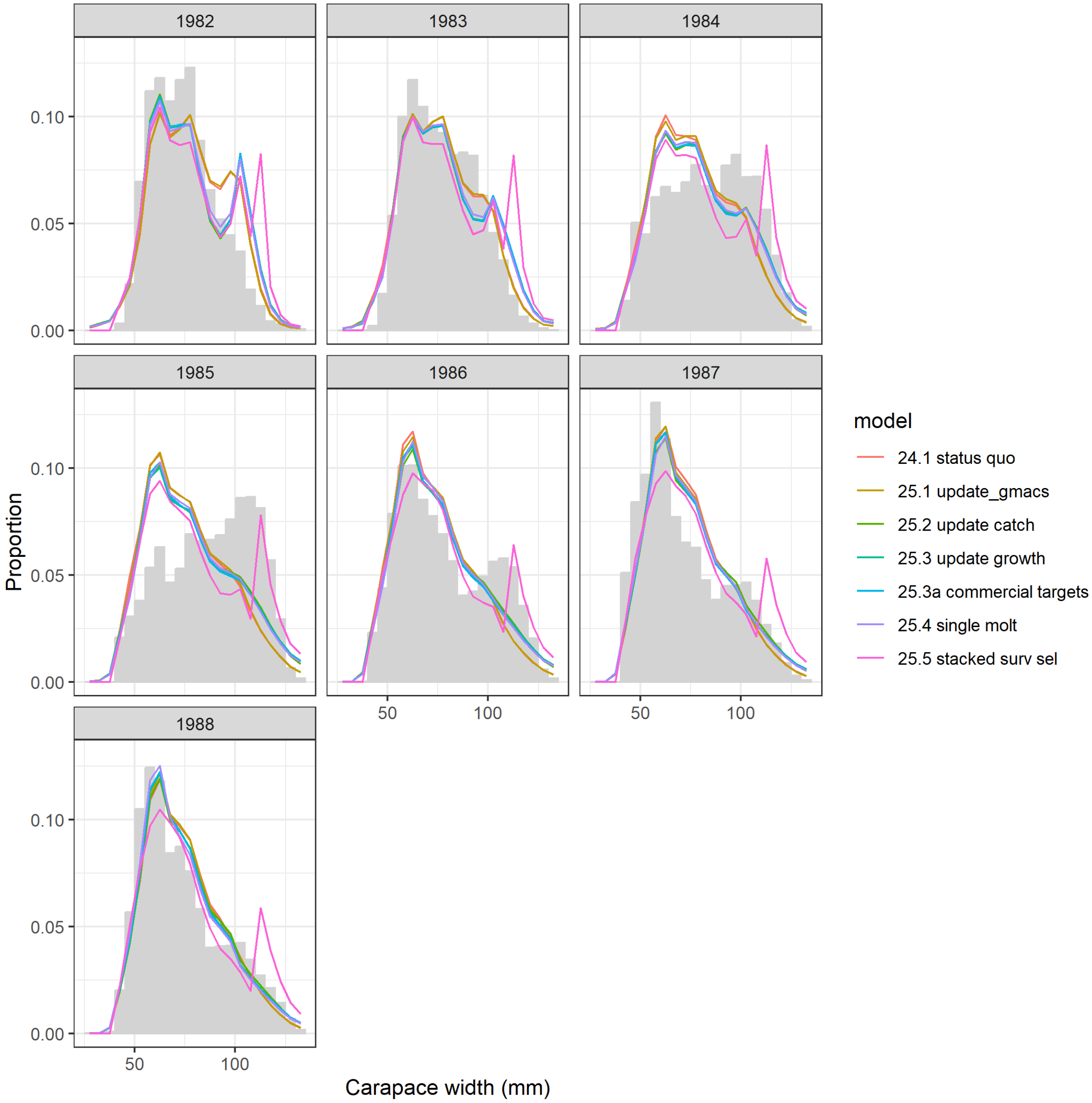

Figure 12: Model fits to mature male survey size composition data from 1982-1988.

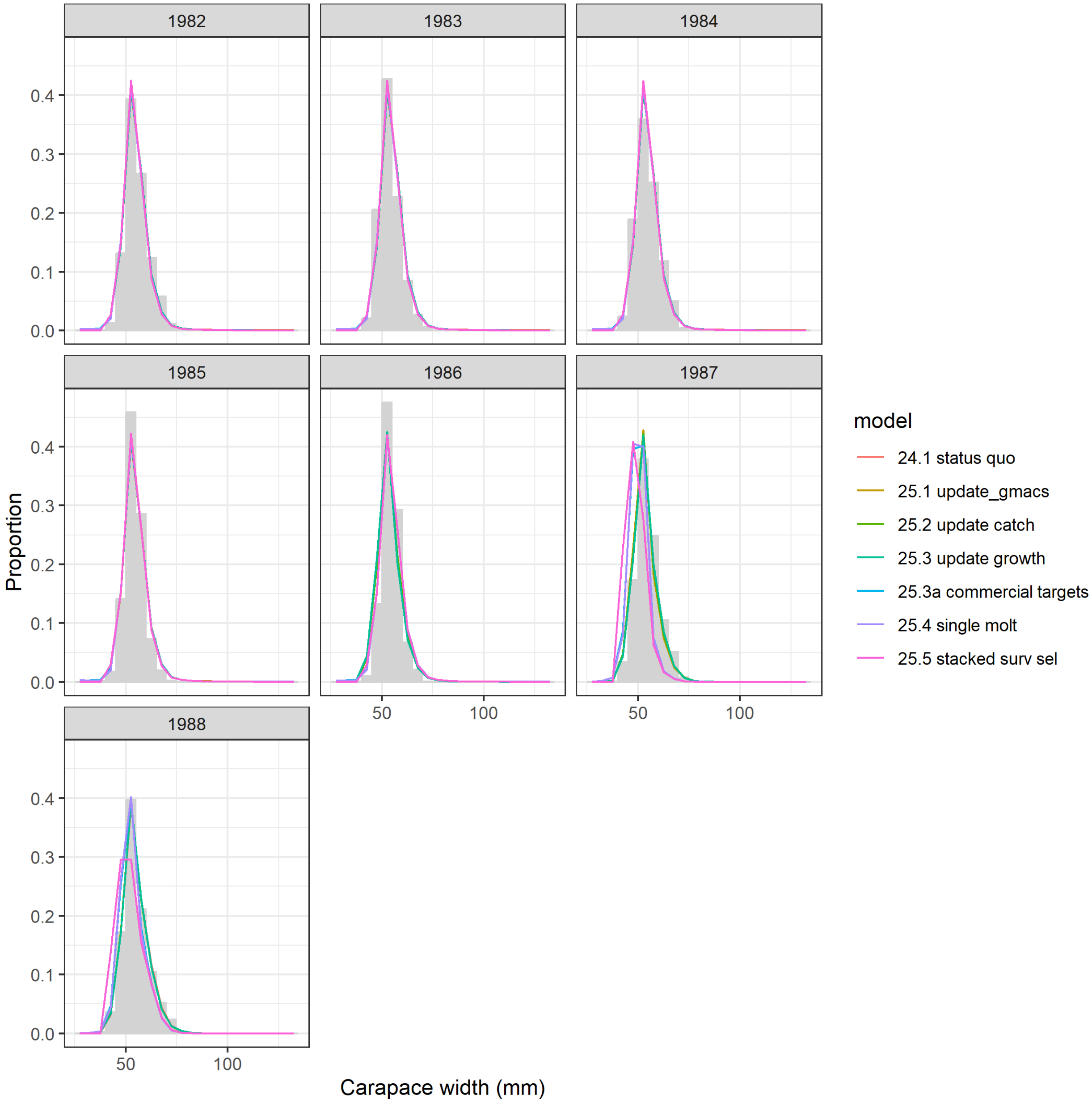

Figure 13: Model fits to mature female survey size composition data from 1982-1988.

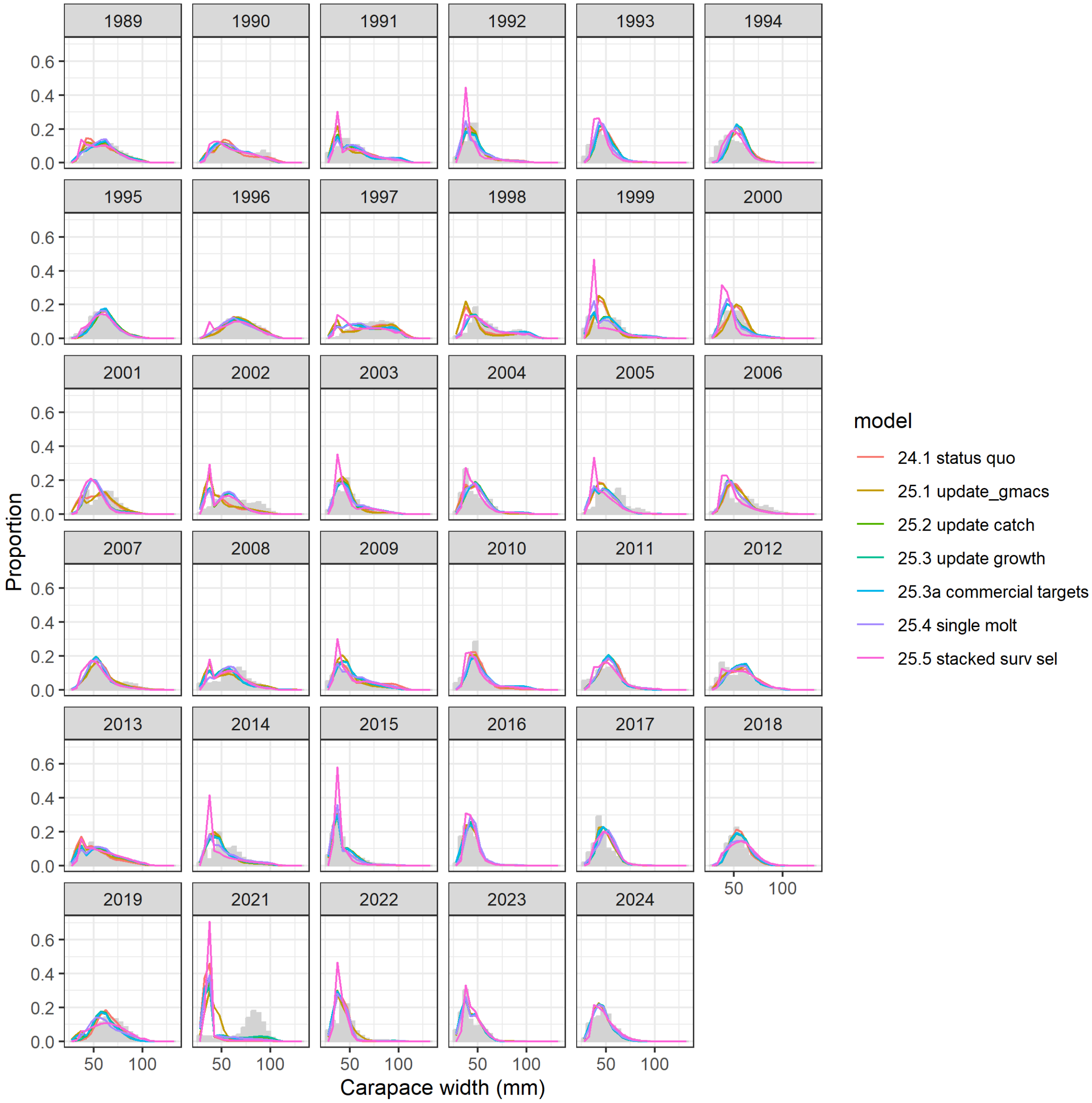

Figure 14: Model fits to immature male survey size composition data from 1989-present.

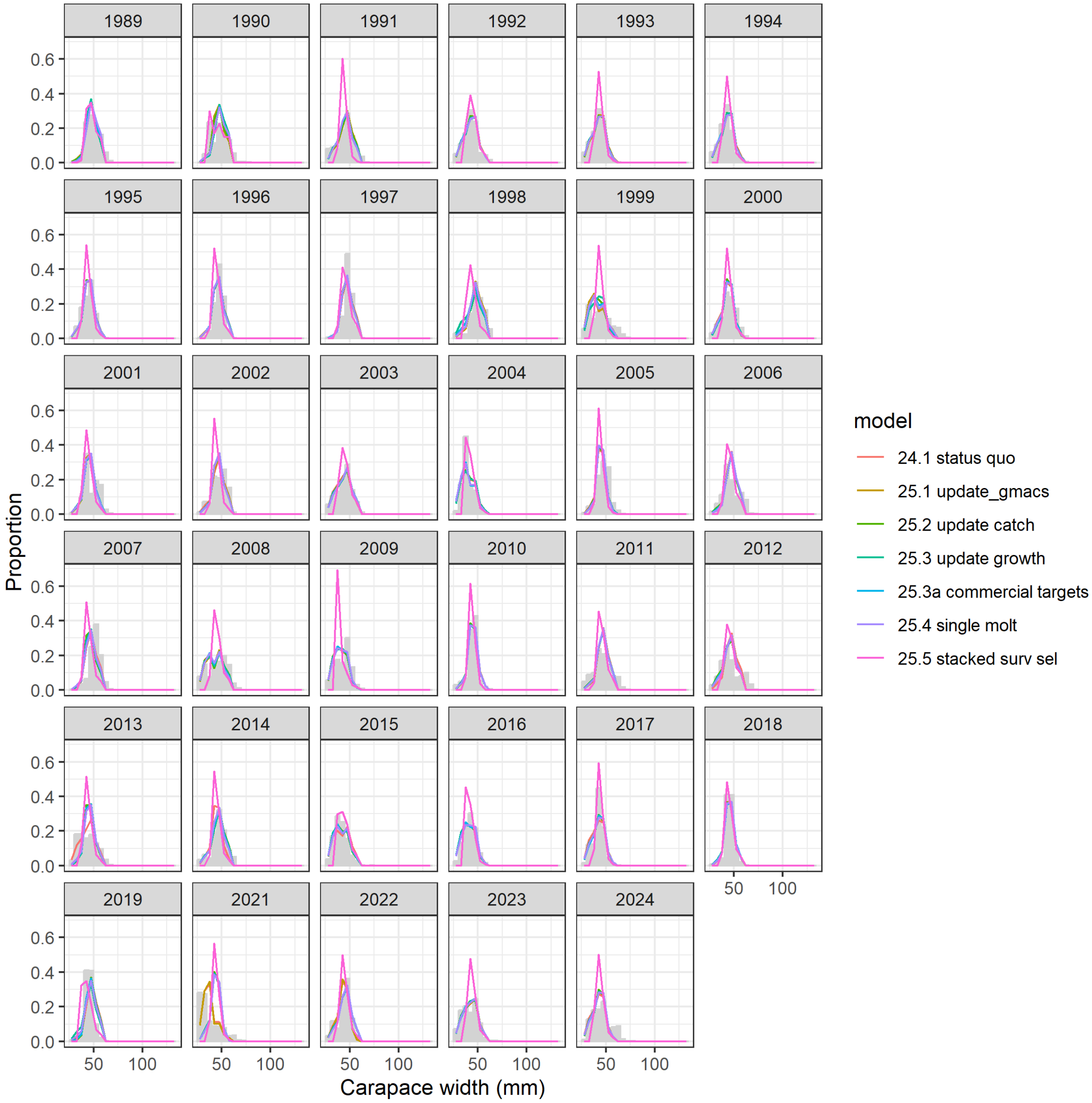

Figure 15: Model fits to immature female survey size composition data from 1989-present.

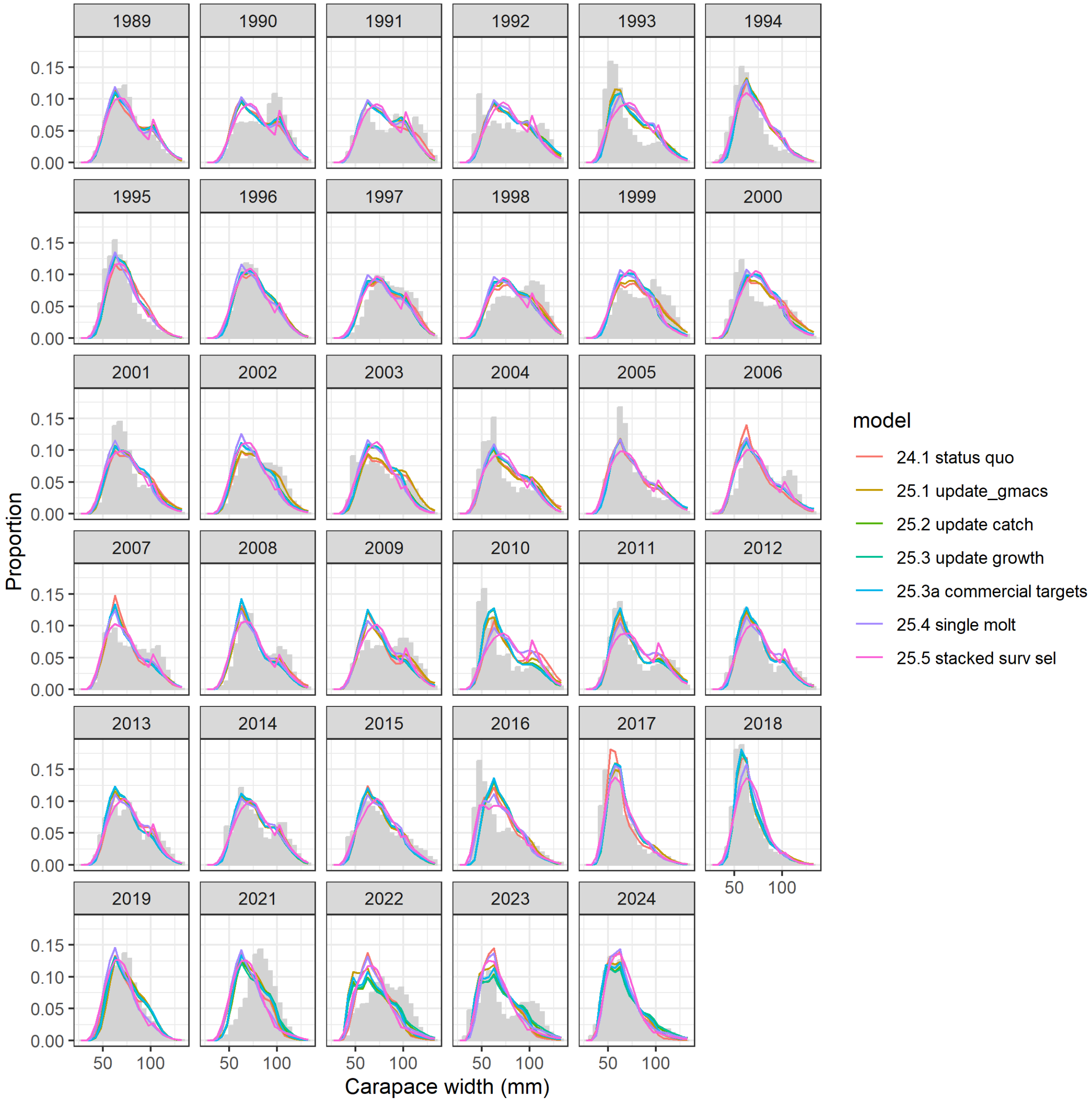

Figure 16: Model fits to mature male survey size composition data from 1989-present.

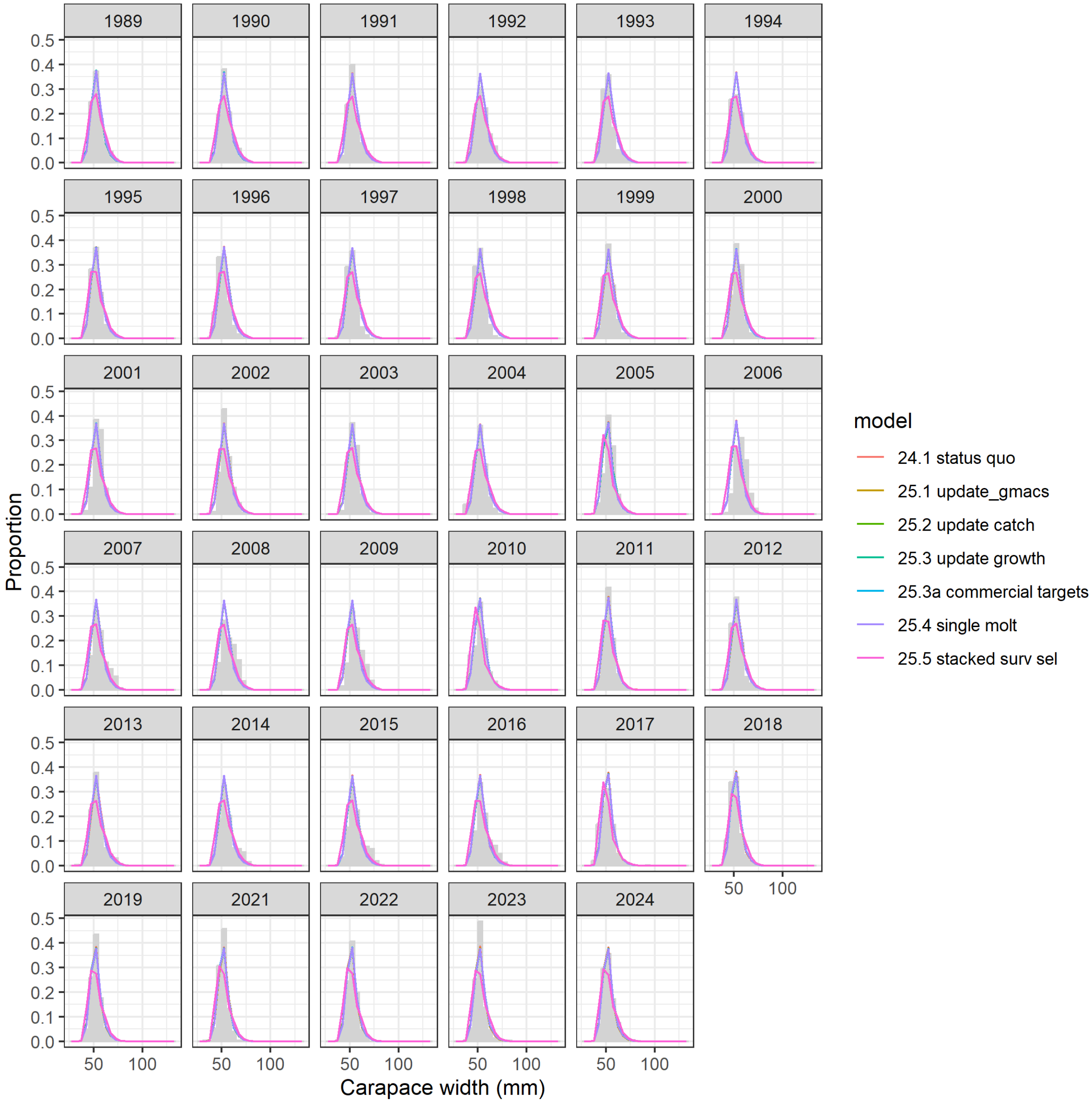

Figure 17: Model fits to mature female survey size composition data from 1989-present.

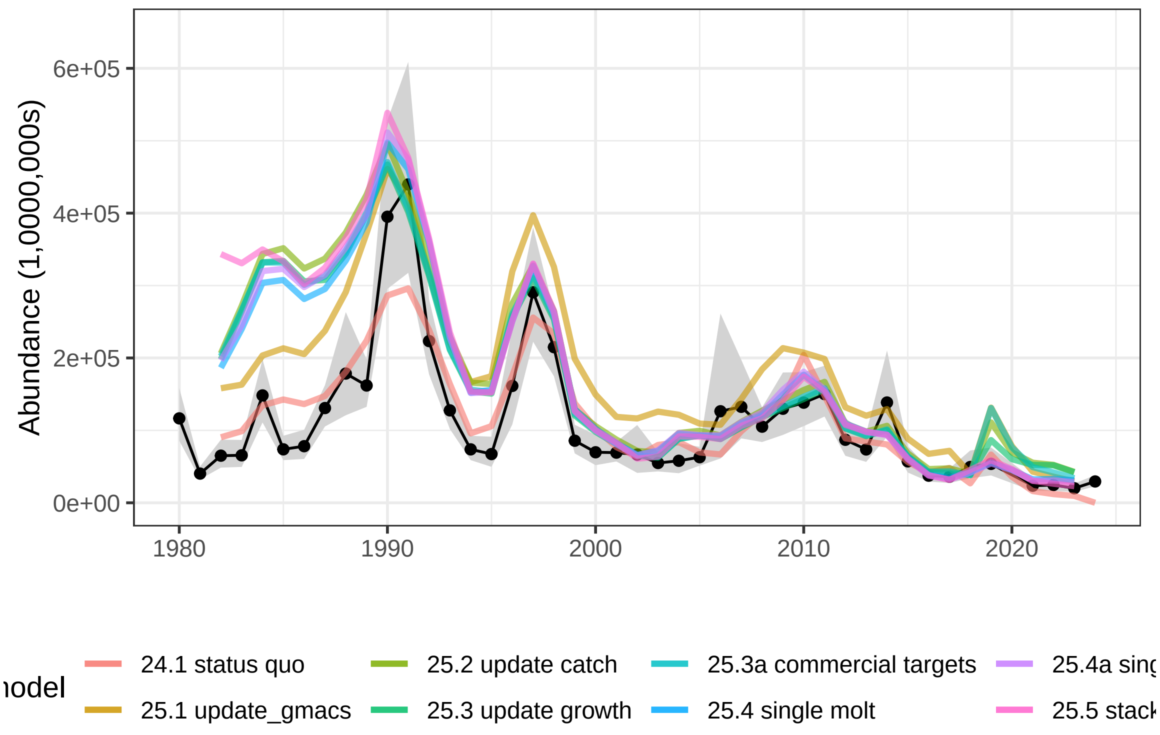

Figure 18: Estimated biomass of male crab >101mm carapace width from the survey (black line and dots with gray 95th CI) and from each model in the assessment (colored lines).

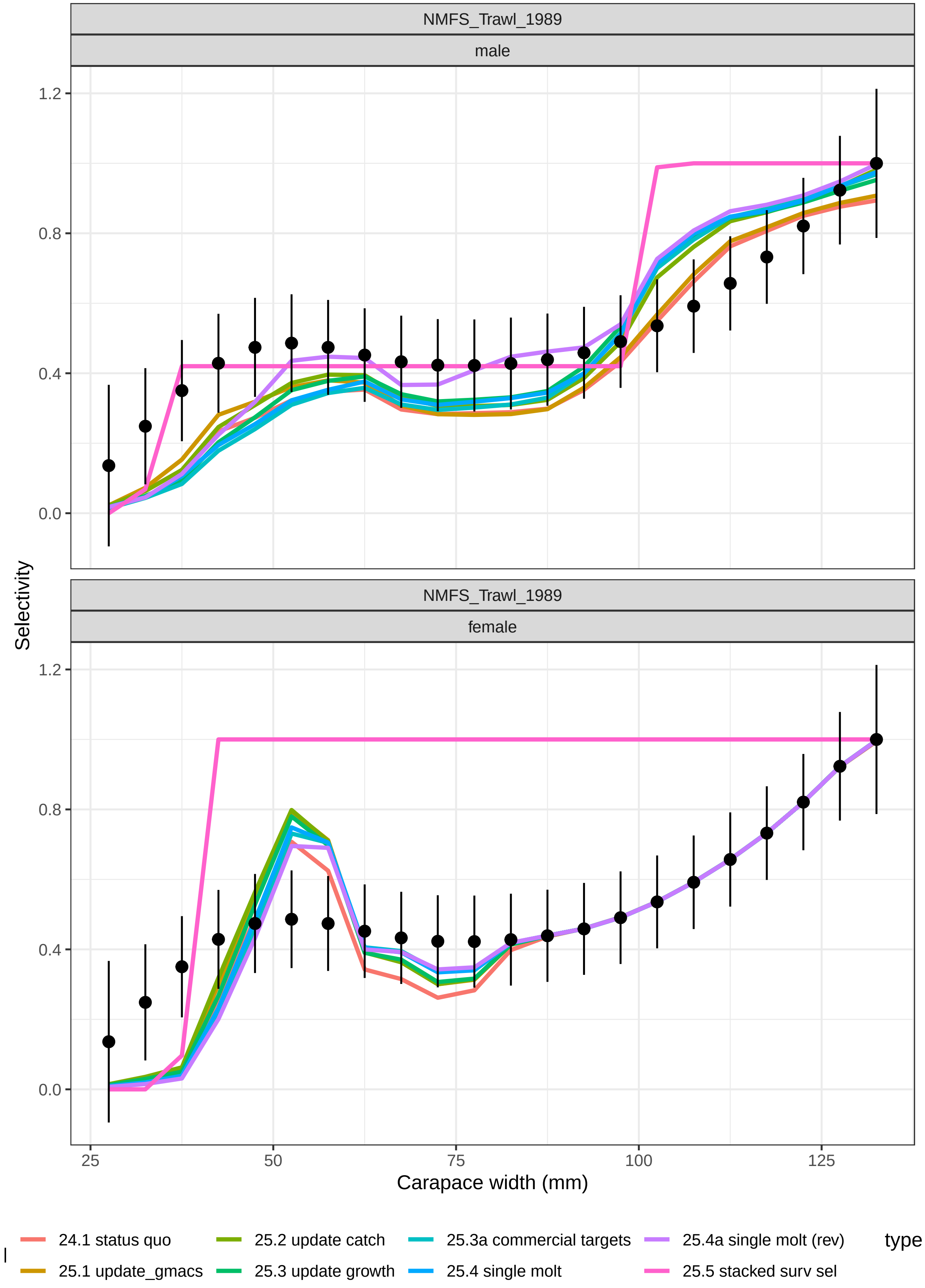

Figure 19: Estimated survey selectivity (lines) with normal priors derived from BSFRF selectivity experiment data. Points are the mean of the prior at a given size; intervals are 95th quantiles based on input CVs. Model 25.5 not estimated with priors.

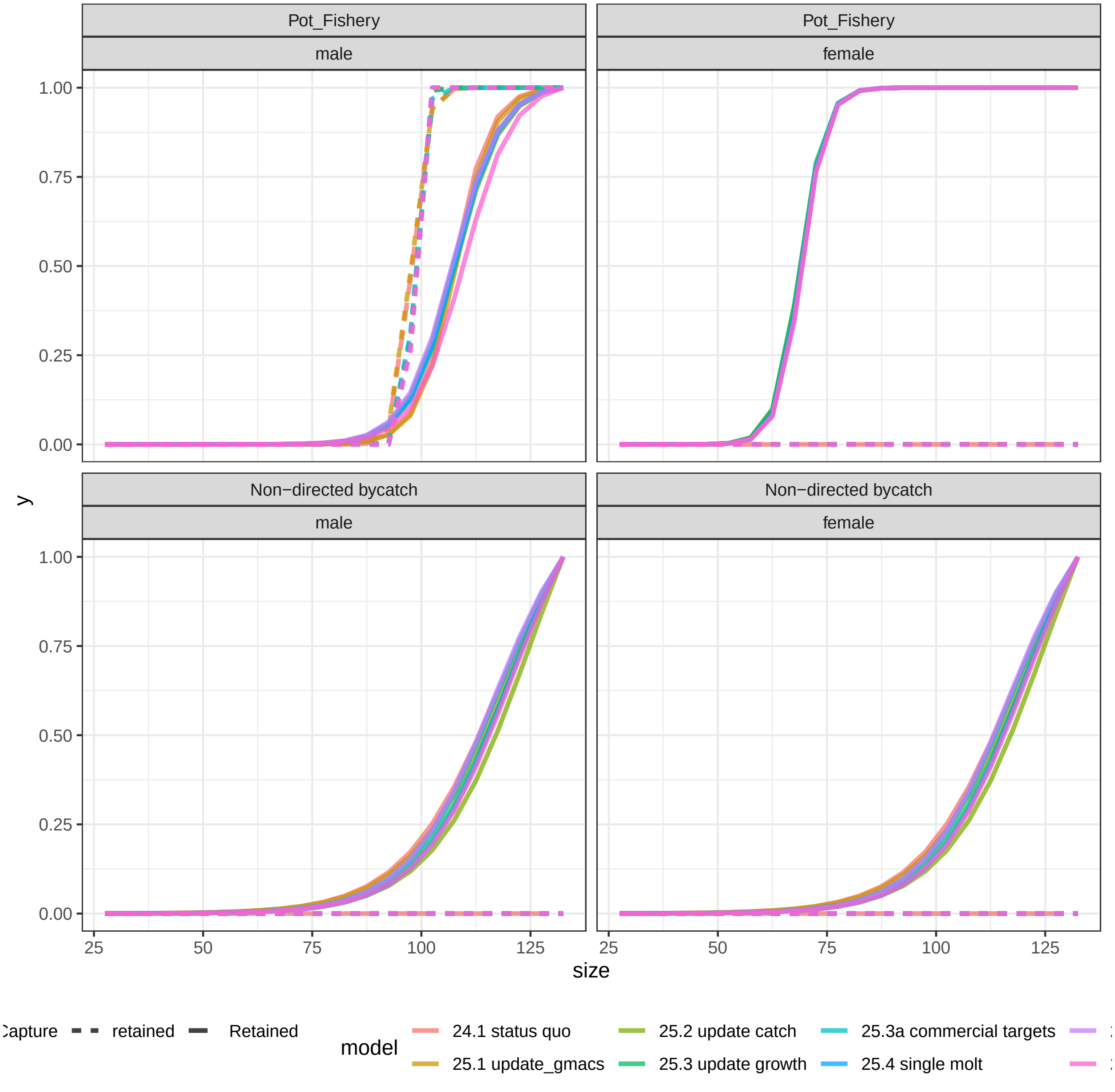

Figure 20: Estimated selectivities by fishing fleet and sex for capture and retained catches.

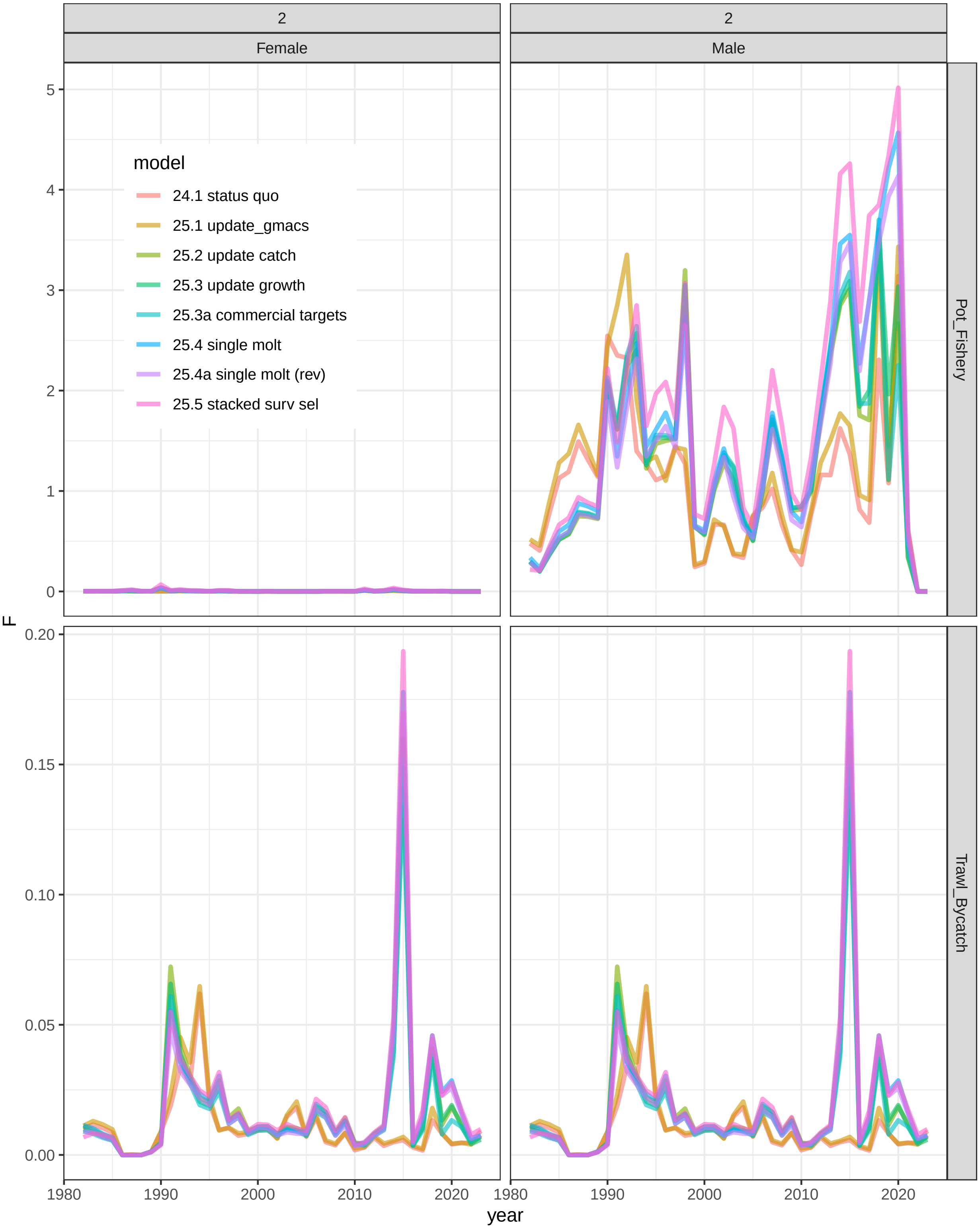

Figure 21: Estimated fishing mortalities for the directed and non-directed fisheries.

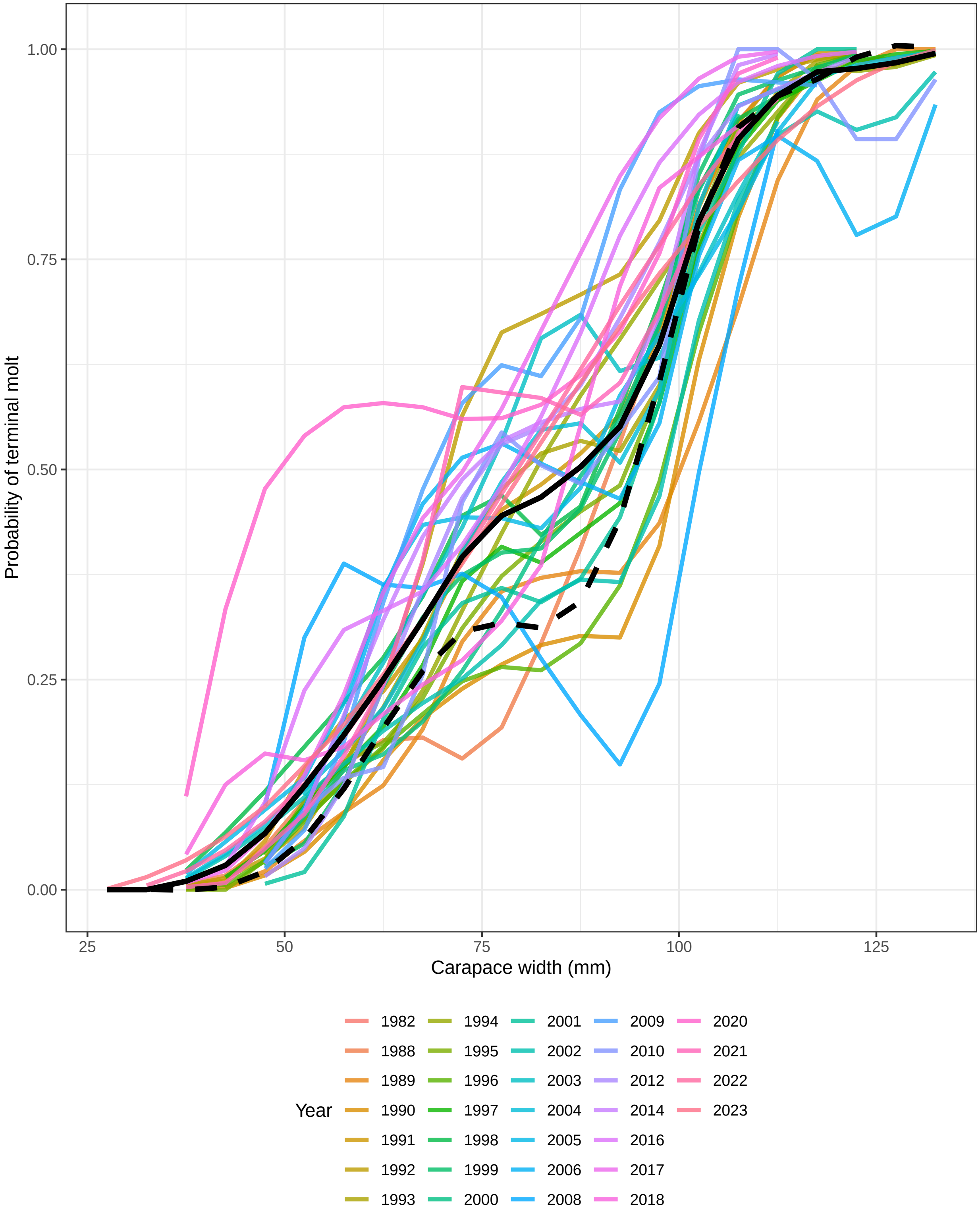

Figure 22: Estimated (black line) or specified (colored lines) probability(s) of maturing for male crab.

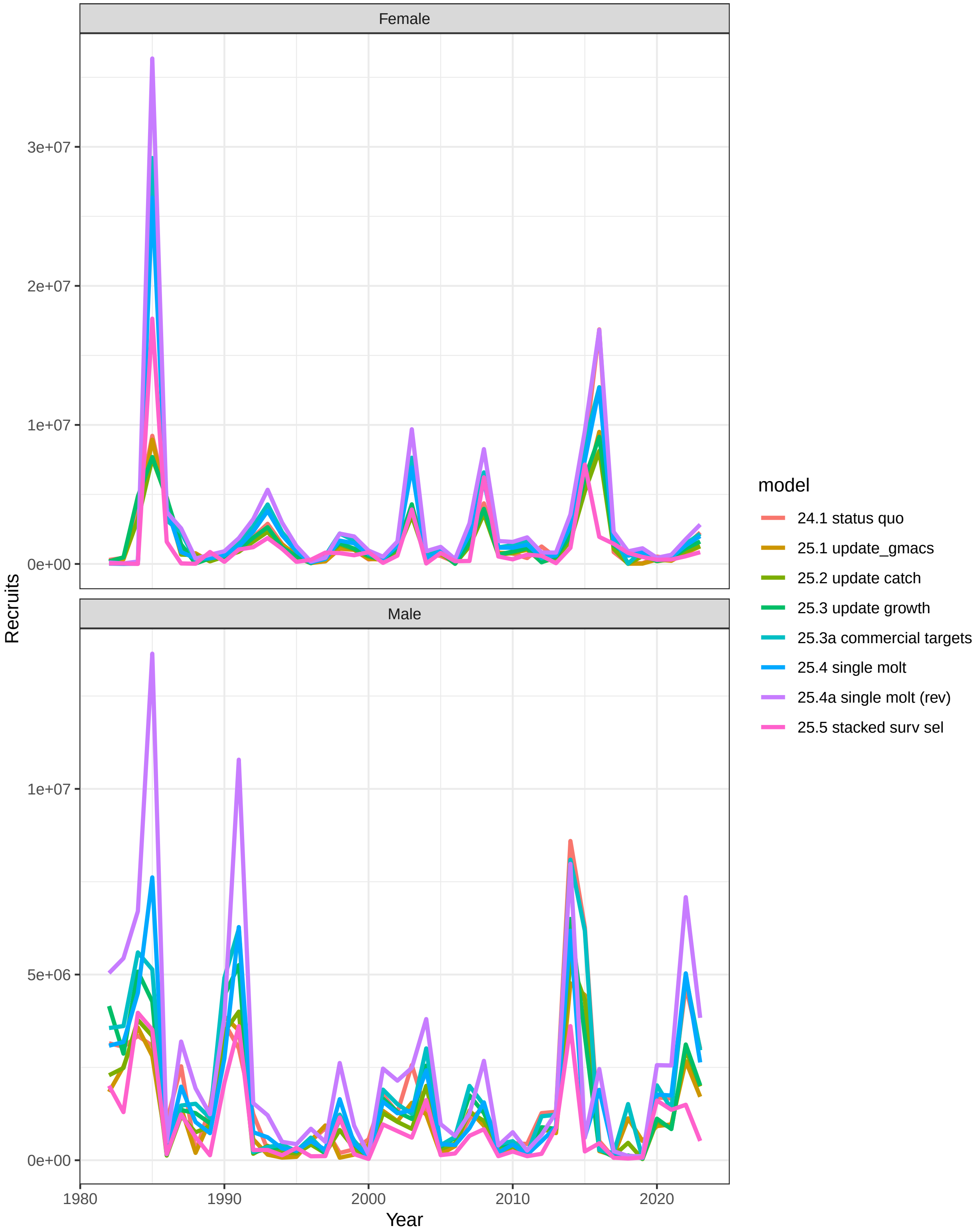

Figure 23: Estimated recruitment by sex (bottom) and proportions recruiting to length bin (top) by model.

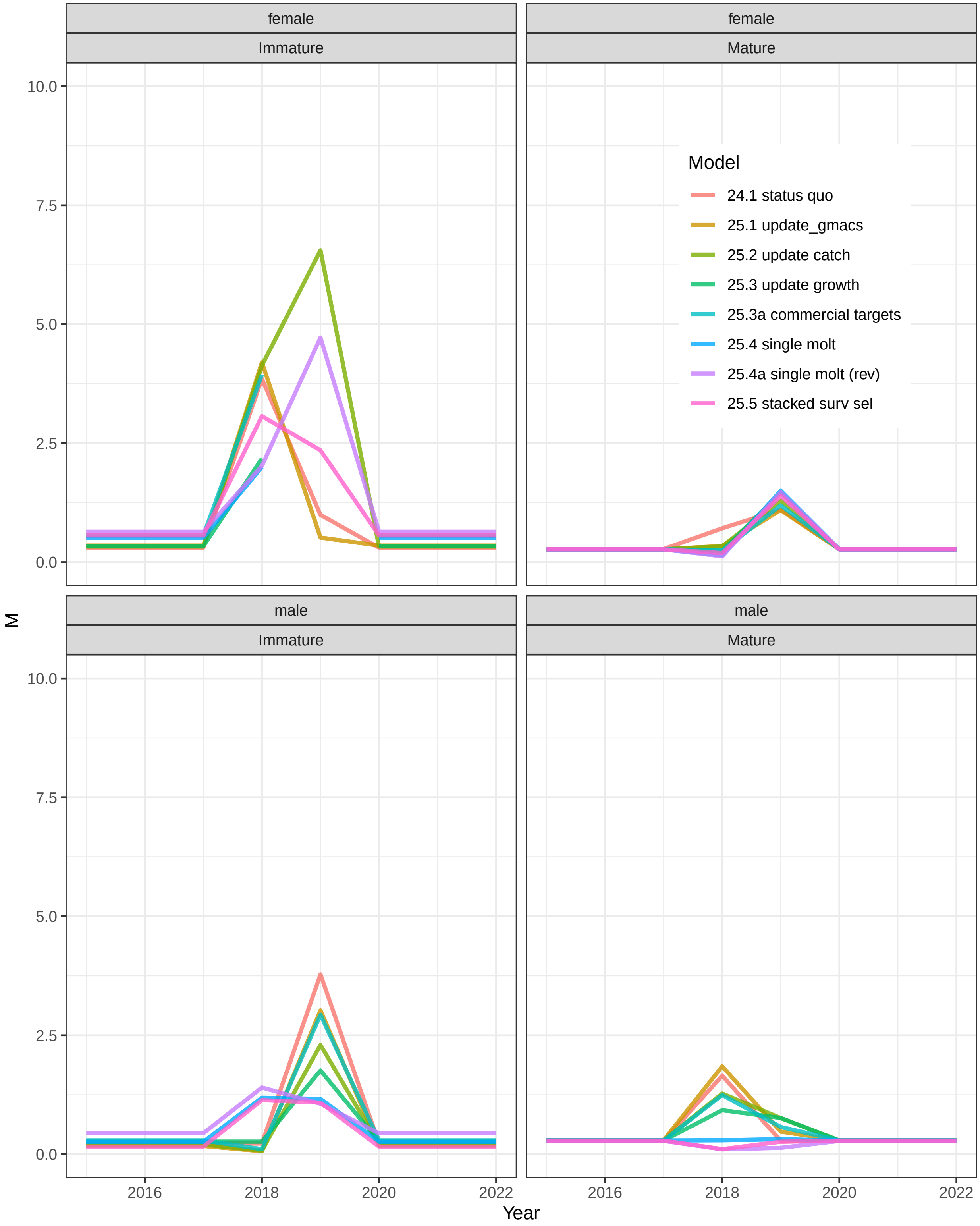

Figure 24: Estimated natural mortality by sex and maturity state. Natural mortality in all years previous to 2018 and after 2019 are equal to the estimated M in 2017.

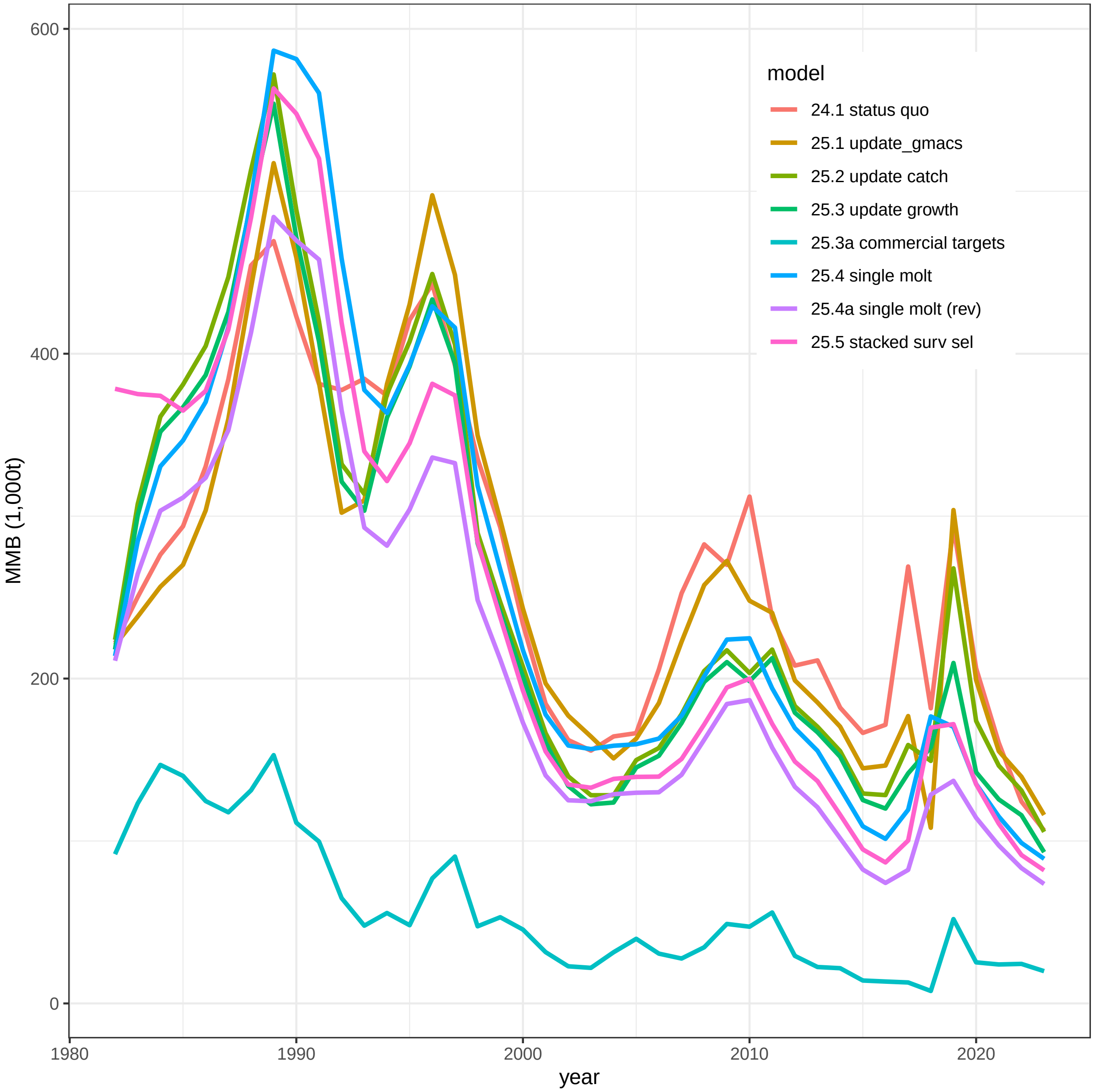

Figure 25: Model predicted mature male biomass at mating time in 1,000 tonnes.

D. Mortality events in Bering Sea crab

This manuscript is presented in draft form to facilitate discussion at plan team. It is not fully referenced yet and should not be cited anywhere until it has undergone peer-review.

Mortality events drive Bering Sea crab populations

Cody Szuwalski[1,2] other authors to come

[1]National Oceanic and Atmospheric Administration, Alaska Fisheries Science Center, Seattle, W
[2]University of Washington, School of Aquatic and Fishery Sciences, Seattle, WA

Abstract

Crab are some of the most abundant species in the Bering Sea and support fisheries historically valued at over $300 million ex-vessel per year. Boom and bust dynamics are common, and we use population dynamics models to identify drivers of change for six crab populations in the Bering Sea. Most large population declines appear unrelated to crab fisheries; recruitment variability and episodic mortality are the key drivers. Density dependent effects, temperature, and size explain a significant portion of the variability in mortality for all populations in our study. Correlated population processes among lightly fished and heavily fished populations suggest environmental change is an important driver of dynamics. Incorporating episodic mortality events into management projections lowers expected yields and biomasses. Our results underscore the need to model historical and future mortality events in fisheries management particularly when there are positive links to temperature or density, in spite of the complications to traditional management strategies it can present.

Body

Crab are some of the most abundant species in the Bering Sea ecosystem and historically have supported some of the most valuable fisheries in the Bering Sea (with salmon and pollock). In particular, snow crab and Bristol Bay red king crab support large and valuable fisheries, but several other more lightly fished populations also exist in the Bering Sea (Figure 1). The trajectories for most of these populations are volatile: booms and busts are common. Most recently, the snow crab population collapsed to unprecedented levels after a marine heatwave in 2018-2019 (Szuwalski et al., 2023) and the fishery was closed for the first time ever. This coincided with a closure of Bristol Bay red king crab, which put communities dependent upon crab fisheries under considerable financial strain and precipitated a ~$150 million disaster relief effort (REF). Large variability in crab abundance occurs in spite of intense management and some of the best data in the world (Hilborn et al., 2021). Population surveys for Bering Sea crab exist stretching back to 1975 (Zacher et al., 2024), catch data from the fisheries are carefully collected both on the docks and at sea (REF), and laboratory experiments are routinely conducted to understand population processes like growth and maturity (REF).

Management for Bering Sea crab is similar to most other closely managed fisheries with extensive data. Each year acceptable biological catches are set based on three outputs from a stock assessment: current biomass, target biomass, and a target fishing mortality. Colloquially, these pieces represent 'where we are, where we want to be, and how we get there' (respectively). Targets for biomass are based on equilibrium projections that assume stationarity in recruitment (i.e. young crab entering the population) and a constant natural mortality. However, periods of elevated natural mortality have been required to fit the survey data for crab in the assessments used in management, suggesting natural mortality is not constant. Furthermore, there is a possibility of unobserved mortality in groundfish fisheries of unknown magnitude as a result of trawl gear fatally impacting, but not capturing crab on the sea floor (Rose et al., 2013).

Given the observed boom and bust dynamics and a potential violation of assumptions made in management, we ask the following questions: 1) how frequent are large declines in Bering Sea crab populations, 2) what drives these declines and are there common linkages, 3) if population collapse is a common phenomenon, how can this information be used in management, and 4) if collapses are linked to temperature, how can management respond under a changing climate?

To answer these questions, we fit population dynamics models to the available data for six crab populations in the eastern Bering Sea with a key difference from the stock assessment models used in management: the estimation of an annually varying 'other mortality' that accounts for processes like predation, unobserved mortality in non-directed fisheries, cannibalism, disease, and starvation. We then used generalized additive models (GAMs; Wood, 2011) to explain variability in our estimates of other mortality using population abundance, temperature occupied, and average size as covariates. We also identified periods during which a population declined more than 40% and used our model estimates to identify drivers for each of these population declines. Finally, we explored the impacts of incorporating the expectation of mortality events into management. Based on these analyses, we end with suggestions for Bering Sea crab fisheries and the larger natural resource management community.

We modeled six populations: snow crab, Bristol Bay red king crab (BBRKC), Tanner crab, Pribilof Islands red king crab (PIRKC), Pribilof Islands blue king crab (PIBKC), and Saint Matthews blue king crab (SMBKC). The models fit the available data well with acceptable diagnostics for all populations (see supplementary materials). Large correlations existed among estimated population processes among populations, which suggest shared drivers of dynamics. For example, Tanner crab, immature snow crab, and PIRKC all had significantly correlated estimates of other mortality, as did the two populations around the Pribilof Islands (Figures S1). Other mortality for mature snow crab and BBRKC, the two most valuable commercial populations, were also significantly positively correlated. Abundance trajectories were positively correlated for the blue king crab stocks and BBRKC, reflecting the large declines seen in these populations.

GAMs incorporating abundance, average size, and temperature occupied explained the variability in estimates of other mortality significantly better than not including the variables for 7 out of 8 population/maturity combinations (Figure 2). Higher abundances were associated with higher mortality for all populations except Tanner crab and SMBKC (Figure 2). Larger average size was associated with higher mortality for all populations except Tanner crab and BBRKC, which had U-shaped relationships. Higher temperatures were associated with higher estimated mortality for all populations and the effect was most dramatic for immature snow crab and BBRKC.

The combined contributions of changes in recruitment, fishing mortality, and other mortality resulted in volatile total abundance trajectories for Bering Sea crab (figure 3). Abundance declines of 40% or greater happened 19 times across all populations, but only 4 of these were associated with fishing mortalities that were higher than the median estimated rates. This, coupled with the relatively small fraction of the population vulnerable to fishing (figure S2), suggests crab fishing is rarely a large driver of decline for populations other than blue king crab. Caution must be taken when using above average fishing mortality as an indicator of fishery-driven decline, though. For example, the decline in snow crab abundance after the 2018-2019 heatwave was primarily driven by declines in the proportion of the population not vulnerable to the fishery; only XX% of the population was large enough to be retained in the fishery (figure S2). So, even though estimated fishing mortality was high, it was clearly not the driver of the population-wide collapse. The remaining large drops in abundance for Bering Sea crab were related to increased mortality, low recruitment, or a combination of the two. Clusters of population declines occurred during three periods: 1980-1984, 1996-1999, and 2015-2020. These all marked periods of ecosystem change and were relatively warm (REF).

The output of our models allows us to begin developing descriptive histories of the volatility for each population. Snow crab has gone through two periods of boom and bust. The largest population abundances occurred during 1990s, but the population declined sharply in 1999. Neither fishing mortality nor other mortality were above average during this period: the declines were driven by a sustained collapse in recruitment. The management response was strong, the population was declared 'overfished' (which, in spite of referencing 'fishing' is a term used to describe a stock that is at low relative abundance regardless of cause; REF), and the stock slowly rebuilt. In the mid-2010s, the largest recruitment event observed in the population occurred and the population was projected increase to abundances not seen since the 1990s. However, the population collapsed again before those projections could come to fruition and this collapse was the result of a mortality event, not a recruitment failure. A model similar to those used within this manuscript was developed for snow crab to try to understand what drove the mortality behind the collapse. Predation, cannibalism, starvation, disease, non-directed fishery bycatch, and THIS were considered as potential drivers. Ultimately, the model output suggested that the 2018-2019 marine heatwave increased the metabolic requirements for snow crab and that starvation likely contributed to the collapse (Szuwalski et al., 2023).

The other large commercial fishery, BBRKC, declined precipitously during the early 1980s as a result of a confluence of unfortunate events: recruitment collapsed in 1979, fishing mortality spiked in 1980, and a period of elevated ‘other mortality’ occurred from 1980-1984. BBRKC experienced relatively high temperatures during this period of other mortality, but there have also been questions about bycatch in non-crab fisheries during this period of time (REF). The management response to the collapse was strong and fishing mortality was reduced, which resulted in a slow rebuilding to another peak in abundance during the mid-2000s. Poor recruitment occurred throughout this time period, but ‘other mortality’ was also low, which allowed the rebuilding. Unfortunately, estimated other mortality increased again in the 2010s, which, coupled with poor recruitment and increasing fishing mortality lead to the lowest estimates of abundance on record in the 2020s. Analyses have been published for BBRKC that link declines in recruitment to ocean acidification (REFS), but more comprehensive consideration of hypotheses and processes that drive dynamics (e.g. ‘other mortality’ and drivers like predation, bycatch, disease, cannibalism, or starvation) could be useful in directing management effort for BBRKC and the other crab stocks in the Bering Sea. Developing appropriate indices to represent those drivers is time- and resource-intensive and will depend on continued funding for surveys and laboratory studies.

The descriptive histories above are only possible when using models that estimate a time-varying ‘other mortality’ and we argue that this sort of history-building will be crucial to management, so that population changes for which management has a ‘lever’ to enact change (e.g. directed fishing) can be distinguished from those management does not (e.g. climate change). However, dangers also exist in building ‘just-so-stories’ around ecological models like this (REF). Determining causality from observational data can be difficult, particularly in complex systems for which relatively short time series are available (REF). Nonetheless, there are considerable risks in not attempting to understand the population changes in a broader context than just fishery-driven dynamics. For example, assuming that population dynamics are driven by fishing when they are actually driven by climate or disease (for example) can result in misdirected (and ultimately wasted) management resources. Experiments, surveys, and analyses aimed at identifying causal drivers of ecosystem change are likely the best hope for designing effective management strategies as climate change reorganizes ecosystems.

Regardless of the true underlying drivers of population change for Bering Sea crab, Bering Sea crab populations are in crisis. The median estimate of commercial abundance during 2023 across all populations was less than 5% of the maximum observed, with the two most economically important populations (BBRKC and snow crab) estimated at 1.7% and 7.2% of maximum, respectively (figure 4). Given the currently poor status of Bering Sea crab populations, it is natural to ask if incorporating the expectation of episodic mortality events into biomass and fishing mortality targets could improve management outcomes for Bering Sea crab. To explore this question, we parameterized a model based on BBRKC and then calculated equilibrium management targets and expected yields in two ways. In the first model, natural mortality was assumed to be constant. In the second model, we sampled yearly other mortality from historical estimates and simulated 1000 populations to develop a distribution for the expected yield and biomass for a given fishing mortality under the possibility of episodic mortality events.

Expected yield and biomass were nearly halved when allowing for the possibility of mortality events and variability in yield and biomass was high (figure 4). Depending on the timing and magnitude of mortality events and their relationship to the timing of recruitment events, biomasses could range from 5-200% of the equilibrium value. Even though there were large differences in expected yields and biomass, the fishing mortality that produced the maximum long-term yield did not change. The key change to the harvest control rule affected by incorporating expected mortality events would lowering the target biomass. This would allow fishing to occur at much lower abundances than is currently allowed. Unintended consequences of incorporating environmentally driven variability into management targets (i.e. increasing the fishing pressure on populations undergoing 'stress') happens for other population processes as well. For example, incorporating climate change effects into recruitment for snow crab results in lowering biomass targets as ice recedes, allowing higher exploitation rates to occur as the population is undergoing climate stress (Szuwalski et al., 2021).

The current management goals for Bering Sea crab are to provide 'optimal yield' based on the equilibrium concepts of maximum sustainable yield (Crab FMP). Firmly negative population trends culminating in all-time lows of commercial biomass for many populations in the last decade suggest that management may not be achieving those goals, in spite of what appear to be relatively conservative harvests. Multiple violated assumptions of the underlying theory behind status quo management strategies suggest that working backwards from what has occurred rather than what 'ought' to have occurred based on theory might yield more positive results. In practice, this could mean taking a more experimental approach to exploitation rates to identify management actions that might improve stock trajectories (e.g Walters). Experimental management would likely need to be coupled with institutional change to provide fishers and communities dependent on crab fisheries more flexibility to support their livelihoods from the Bering Sea ecosystem.

Our observations and modeling for Bering Sea crab may have important implications for other managed living resources. Mortality events can occur for a range species for a range of reasons (REFs) and management response might change for each. Broadly, resource management operates at two levels: strategic (i.e. goal setting) and tactical (i.e. actions to accomplish goals).

From a strategic perspective, protecting populations from (and after) collapse and supporting resource-dependent communities are often central goals and should guide the type of management actions considered. Protecting populations from/after collapse can be challenging in the current equilibrium management paradigm if adverse drivers are built into the calculation of expectations of the population under no fishing, as shown above. The challenge arises from the downward shift of targets from incorporating expected collapses into the projection, which would allow for harder fishing at lower abundances compared to status quo rules. It is possible that alternate forms of harvest control rules could provide better management outcomes when populations are prone to mortality events. The shape of these rules would likely depend on the drivers of mortality events. Based on our analyses, consideration of population abundance, size composition, and temperature may be useful covariates to consider.

The problem of boom-bust dynamics is not new: forage fish are the classic example in fisheries. Anchovy (*Engraulis mordax*) and sardine (*Sardinops sagax*) off southern California have gone through boom-bust cycles over the last 500 years (much of which was in the absence of fishing) according to scales collected in the sediment around the Channel Islands (McClatchie et al., 2017). Crustacean populations can also vary considerably and withstand high fishing mortalities. For example, Dungeness crab (*Metacarcinus magister*) populations in the Pacific Northwest have persisted in spite of very high exploitation rates (>90% in some areas) on the large males (Richerson et al., 2016). These examples suggest that fishing is not a dominant driver of population change for some species and might suggest a higher tolerance for low abundances when specifying allowable harvests.

However, a key difference between forage fish, Dungeness crab, and the Bering Sea crab in our study is life span. Anchovies live 7 years and mature at 1-2 years old (Baxter, 1967), sardines can live 25 years, but mature at 1-3 years old (REF), and Dungeness crab live ~10 years and mature at 2-3 years old (REF). In contrast, snow crab live up to 20 years and mature at ~6-8 years old (REF), red king crab lives up to 30 years and mature at ~5-6 years old (REF), and blue king crab live ~20-30 years and mature at ~5-6 years old (REF). These differences in life history can be important for how responsive a population can be to favorable environmental conditions. These differences also might be related to why PIBKC has not come back in spite of not being fished for a quarter of a century and why the king crab populations in the Gulf of Alaska collapsed and did not return (REF). Longer times to maturity can make it more difficult for populations to be responsive to environmental conditions.

From a tactical perspective, models with time-varying other mortality are not often used in fisheries science, and for what seem like good reasons in stationary systems where fishing is a primary driver of dynamics. First, change in other processes alias for change in mortality—a change in growth or catchability can look very similar to a change in mortality ('hide or die' REFs; Szuwalski et al., 2019). Furthermore, the data to which population dynamics models are fit are also subject to observation error, so there is some risk that estimates of mortality are reflecting noise in the data rather than a signal from the population (REF). In addition to difficulties reliably estimating variable mortality, small changes in the mortality used to calculate reference points can result in large changes to management advice (e.g. Thorson et al., 2015). Because of these points, the prevailing advice for population dynamics models is to estimate a time-invariant natural mortality (Johnson et al., 2014), often informed by a prior that considers the longevity of the modeled species (Hamel, 2014).

Our modeling choices are a departure from current 'best practices' in stock assessment, but we argue that modeling 'other mortality' and attributing that mortality to some source(s) (or at least attempting to) will likely be an increasingly important issue as climate change progresses and when fishing is one of many drivers of dynamics. We also recommend that scientists and managers build models that have alternate assumptions about the drivers of dynamics (including time-varying mortality), simulation test all population dynamics models, methods of attribution and management strategies, and check for historical biases and retrospective patterns.

Collapses of harvested natural resources will continue to happen; knowing when and/or why is a hard scientific problem, but one that needs to be undertaken. If the causes of collapse can be identified, developing appropriate management strategies can be relatively straightforward. For example, if predation is a key driver of a harvested population, incentivizing fishing for the predator in areas where there is the most overlap between predator and prey may reduce pressure on the prey population (REF). Regardless of whether we understand the causes, management systems should be developed that have the flexibility to promptly address sudden collapses, which may be more frequent as climate changes (Litzow et al., 2024). These systems should be built on the principles that collapses should be met with decisive management action (i.e. closures) and implementing social institutions that soften the economic blow to those dependent upon the resource (e.g. providing diverse fishing opportunities). If we can understand why something occurred, decisive action can be taken. If we cannot understand why something happened, action still must be taken, so modeling exercises should be undertaken to ensure decisions are made sensibly.

Acknowledgements to come

Github.com/szuwalski/ebs_crab

References to come

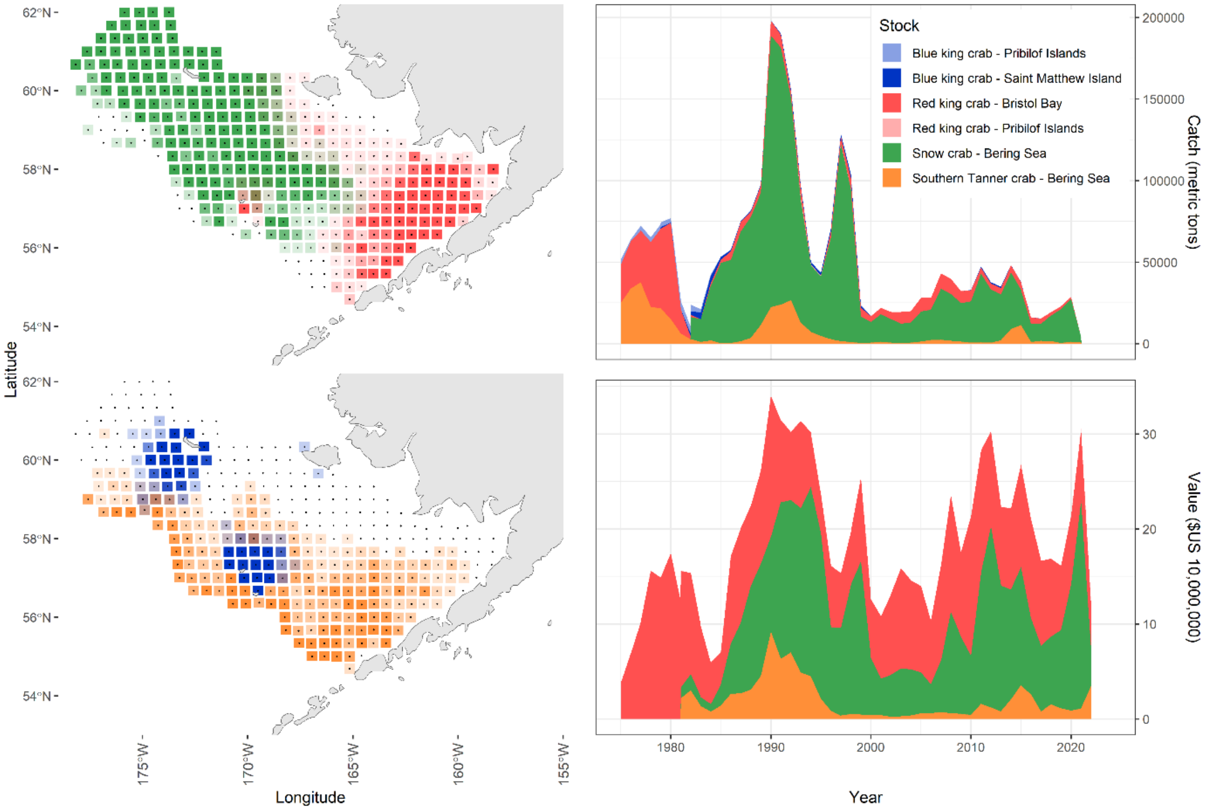

Figure 1. Location of crab populations in the Bering Sea (left) and the catch and ex-vessel value of that catch for crab populations (right).

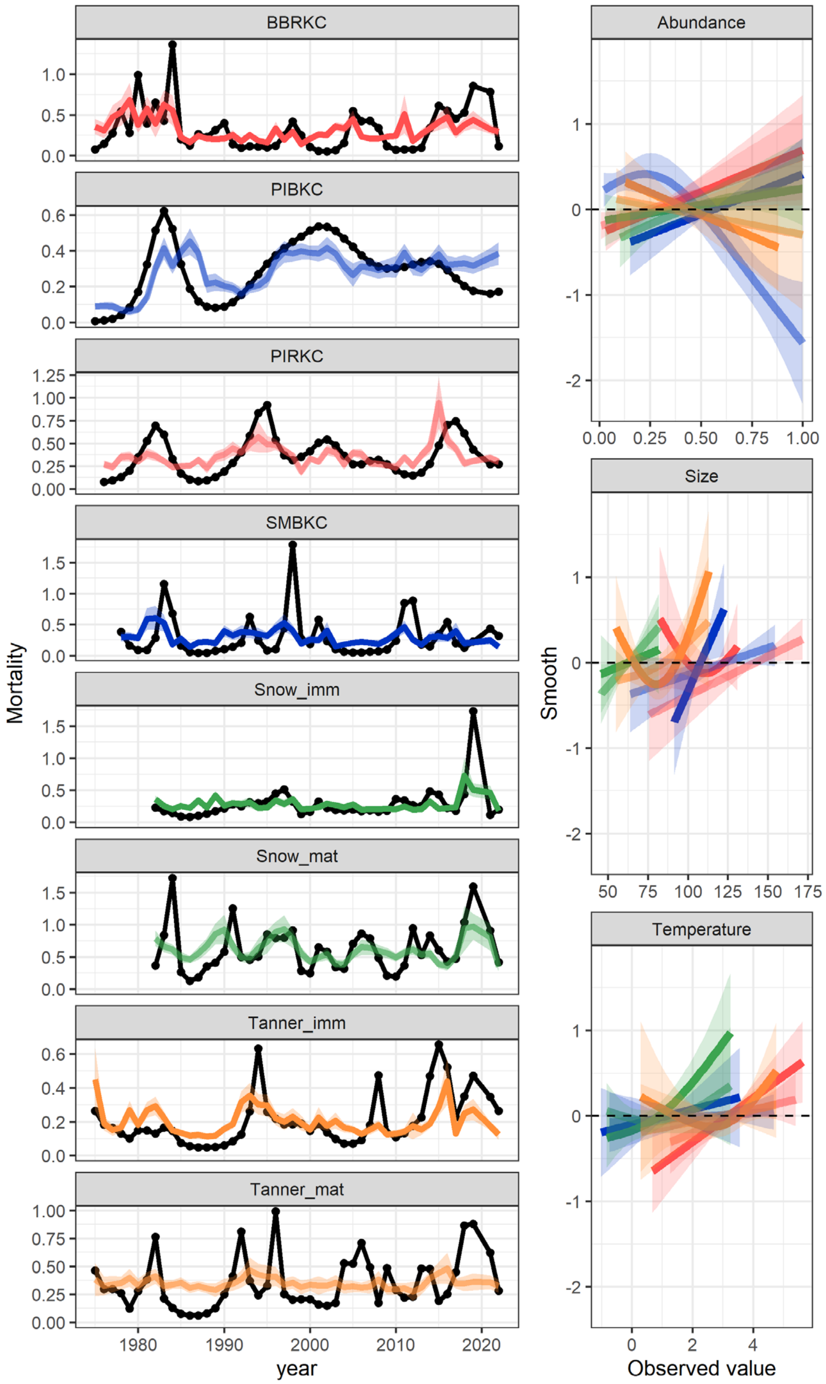

Figure 2. Estimated mortality from population dynamics models (left; black lines) with GAM estimates of those time series (colored lines). Smooth relationships estimated within the GAMs for three covariates. Colors represent a population and are consistent among figures.

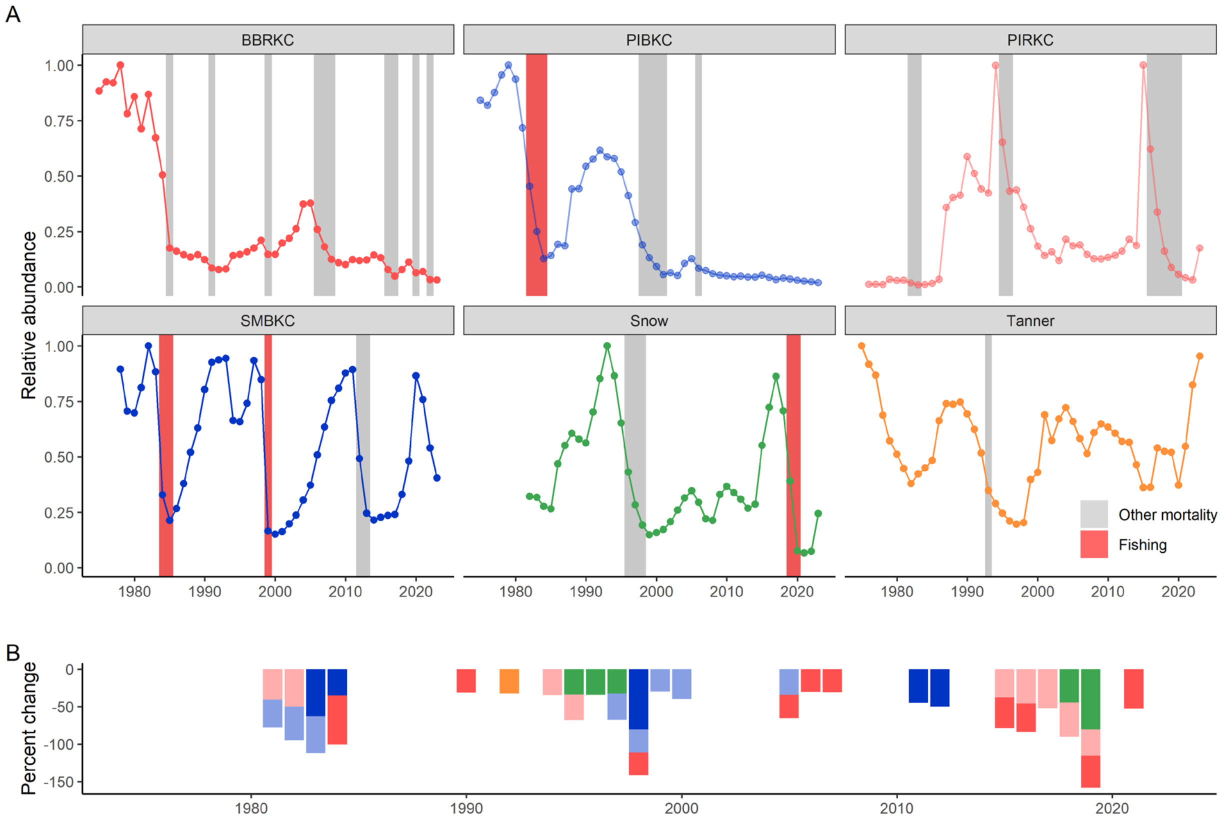

Figure 3. Estimated relative total abundance (add uncertainty) by crab population (A) with periods during which the abundance declined by at least 40% (vertical rectangles). Red rectangles correspond to periods of decline during which estimated fishing mortality was greater than the median; grey bars indicate periods when fishing mortality was below the median. Panel B collates the declines by year and magnitude.

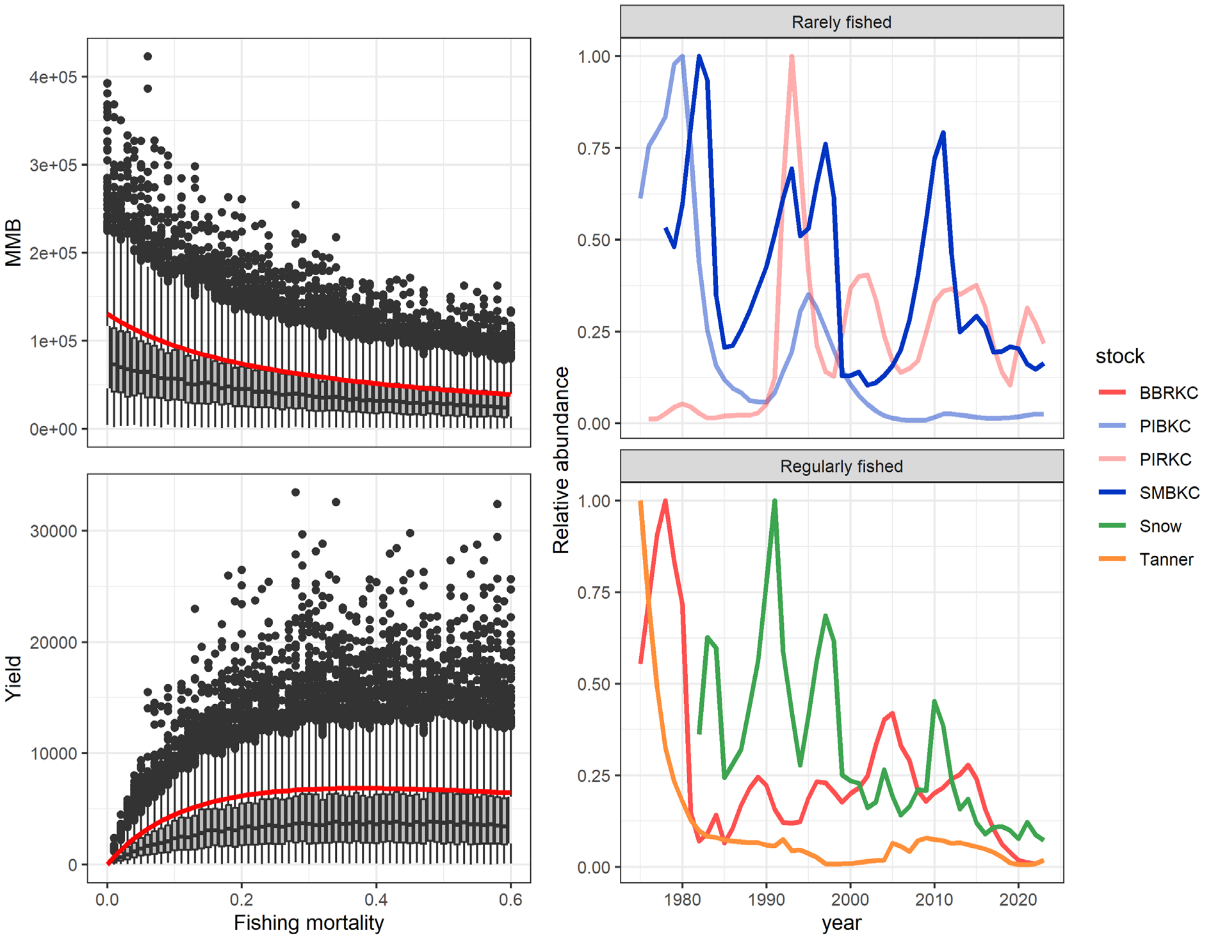

Figure 4. Equilibrium mature biomass (top left) and yields (bottom left) for a given fishing mortality when natural mortality is constant (red line) and when mortality events are considered via stochastic simulations (box plots). Trajectories of the abundance of industry preferred sizes of crab from rarely fished populations (top right) and regularly fished populations (bottom right)[add uncertainty].

Supplementary information for 'Episodic mortality drives Bering Sea crab populations'

Contents

Supplementary materials

Methods overview

We used an integrated population dynamics model to estimate variation in mortality (partitioned into 'other' and 'fishing') and recruitment over time for six crab populations in the eastern Bering Sea. The population dynamics models were fit to abundance and size composition data for male crab from the National Marine Fisheries Service (NMFS) summer bottom trawl survey on the eastern Bering Sea shelf. The model was also fit to retained and discarded catch from the directed fishery when available. We then used generalized additive models (GAMs) to relate the estimated variation in 'other' mortality to population metrics and temperature occupied.

These supplementary materials are incomplete and will be finished at a later date. In spite of their unfinished state, they were included here to facilitate discussion about the manuscript, particularly about the model framework.

Assessment data

NMFS Bottom trawl survey

The eastern Bering Sea trawl survey has been conducted annually since the mid-1970s although no survey was conducted during 2020 due to the COVID-19 pandemic (Zacher et al. 2023). During the first decade of the survey there was a gradual expansion of the survey area leading to variability in both the number of stations sampled, and survey extent, primarily in the northwest of the survey grid (i.e. in the vicinity of St. Matthew Island). The survey extent was standardized in 1988, and 376 stations have been consistently sampled since. A 0.5-hr tow is made near the center of each 20 x 20 nmi (37 x 37 km) square on a sampling grid. Since 1982, the standardized sampling gear has been the 83-112 eastern otter trawl (Zacher et al. 2023). Prior to 2006 all carapace width and chela height measurements (used to assess male maturity) were measured to 1 mm. During the period 2006-2015 the majority were measured were measured to 1 mm, and a subsample specifically for chela height measurements were measured to 0.1 mm using Vernier calipers. Beginning in 2016 all crab have been measured to 0.1 mm using digital calipers. Individual crab are measured to the nearest 2 g using a digital scale. In cases where subsampling is required due to catch size, crab are divided into kept and discarded categories and weighed to the nearest 10 g. The ratio of kept/discarded is then used to calculate a sampling factor by which the sampled catch may later be expanded.

Male crab maturity was determined using the procedure outlined in Richar and Foy (2022). Briefly, paired, log-transformed, carapace width and chela height data are divided into a series of intervals along the log-carapace width (x) axis, within the log-carapace width range over which two overlapping data clouds occur. Data from each interval are used to calculate a distribution, and the minima of the region separating the 2 dominant peaks is obtained. The set of midpoints for each log-carapace width interval is then combined with the set of extracted minimas to create an (x,y) data set. A linear model is then run using these points as inputs, and the calculated regression line is the final maturity cutline.

Temperature occupied for each stock was calculated from the measured bottom temperature from each tow, weighted by the CPUE for a given species at that tow (figure SX). Average size was calculated as a weighted mean of the vector of size, with weights given as the abundance of crab in a given size bin estimated from the fitted population dynamics models (figure SX).

Directed fisheries data

Fisheries catches are reported by Alaska Department of Fish and Game (ADFG) statistical area in logbooks and were aggregated by year to be used as input in this analysis. The fishery generally starts at the end of October and closes at the beginning of May, and most of the catches occur during a short period in winter. Catch of retained crab from the directed snow crab pot fishery from survey year 1982 to the 2021 were used in this analysis and are based on dockside landings. The fishery was closed in 2022 and 2023 due to low abundances. Information on discarded crab are collected via on-board observers who measure, count, and weigh samples of crab brought aboard fishing vessels. More information on the fisheries data collection process can be seen in the individual stock assessments (Szuwalski, 2023; Palof et al., 2023; Sterns, et al., 2023; Stockhausen 2023).

Population dynamics model

Size-structured population

The population dynamics model presented here incorporates the best available information on relevant population processes. The model tracks numbers of male crab at size (s) (and maturity state (m) for Chionoecetes) over time ($N_{t,s,m}$) with 5 mm size bins over different ranges of sizes for each population (**??**). Other mortality (M) is estimated by year (y) (and maturity state (m) for Chionoecetes); fully-selected fishing mortality (F) is estimated by year (y). Other estimated parameters include the initial numbers at size by maturity state, an average recruitment, yearly recruitment deviations, a vector of scalars that determine the proportions of estimated recruitment split into the first three size bins, fishery selectivity, and survey selectivity (see **??** for a list of what parameters are estimated for each stock. Parameters determining growth and maturity are estimated outside of the model based on available data from stock assessments. The timing of the fishery is different for each population and is denoted by the month in which it occurs (q below). Mortality is the only population process that occurs during the first several months of a given year (the crab year begins in July with the bottom trawl survey):

$$N_{t=y+q/12,s,m} = N_{t=y,s,m}e^{-(\frac{q}{12})M_{t,s,m})} \tag{1}$$

Fishing occurs as a pulse fishery in which the crab captured and brought on deck of fishing vessels (C_{cap}) are a function of capture selectivity (S_{cap}), the number of crab in the Bering Sea at the time of fishing, and the fishing mortality applied (F_t). A retention ogive (S_{ret}) is applied to the captured crab to determine what fraction of crab at a given size are retained for sale (C_{ret}) and what fraction are discarded back into the ocean (C_{disc}). A discard mortality (d_{mort}) of 25% is applied to the crab returned to the ocean.

$$C_{cap,y} = n_{t=y+q/12,s,m}(1 - e^{-F_t * S_{cap}}) \tag{2}$$

$$C_{ret,y} = C_{cap,y} * S_{ret} \tag{3}$$

$$n_{t=y+q/12,s,m} = n_{t=y+q/12,s,m}e^{-F_t * S_{cap}} + n_{t=y+q/12,s,m} + (1 - d_{mort})C_{cap,y}(1 - S_{ret}) \tag{4}$$

Capture selectivity and retention selectivity are logistic functions of size in which the slope and size at which 50% selectivity are estimated (two time-invariant parameters for each ogive).

$$S_{cap,s} = \frac{1}{1 + exp(-S_{slope,cap}(size(s) - S_{50,cap}))} \tag{5}$$

$$S_{ret,s} = \frac{1}{1 + exp(-S_{slope,ret}(size(s) - S_{50,ret}))} \tag{6}$$

Growth occurs after fishing and is represented in the model by multiplying the vector of immature crab at size by a size-transition matrix $X_{s,s'}$ that defines the size to which crab grow given an initial size. this process results in two temporary vectors of numbesr at siez for Chionoecetes species where immature crab are denoted by the subscript m=1 and mature by m=2.

$$n_{t=y+q/12,s,m=immature} = z_s(1 - \rho_{y,s})X_{s,s'}n_{t=y+q/12,s,m=1} \tag{7}$$

$$n_{t=y+q/12,s,m=mature} = z_s\rho_{y,s}X_{s,s'}n_{t=y+q/12,s,m=1} + (1 - z_s)n_{t=y+q/12,s,m=2} \tag{8}$$

A single equation can describe king crab speices given the lack of a terminal molt:

$$n_{t=y+q/12,s} = (1 - z_s)X_{s,s'}n_{t=y+q/12,s} + z_s X_{s,s'}n_{t=y+q/12,s} \tag{9}$$

The size transition matrices $X_{w,w'}$ used here were constructed using growth increment or tagging data collected over several years (see assessments for a summary) to estimate a linear relationship between pre- and post-molt carapace width, ($\hat{W}_w^{pre}$ and $\hat{W}_w^{post}$, respectively) and the variability around that relationship was characterized by a discretized and renormalized normal distribution, $\text{Y}_{w,w'}$, where w and *w'* represent entries in the rows and columns of the matrix.

$$X_{w,w'} = \frac{Y_{w,w'}}{\sum_{w'} Y_{w,w'}} \tag{10}$$

$$Y_{w,w'} = (\bar{W}_{w'} + 2.5 - W_w)^{\frac{\hat{W_w} - (\bar{W}_w - 2.5)}{\beta}} \tag{11}$$

$$\hat{W}_w^{post} = \alpha_+ \beta_1 \hat{W}_w^{pre} \tag{12}$$

Where α, β, and β_1 are parameters estimated outside of the population dynamics model.

The probability of maturing (ρ) and molting probability (z) are handled differently for king crab vs. Chionoecetes species. Chionoecetes species have a terminal molt to maturity, after which they do not grow again (Tamone et al. 1995). The probability of maturing is based on the observed ogives of the proportion of mature new shell males by size calculated from chelae height measured in the NMFS survey data (Richar and Foy, 2022). These observed probabilities of having undergone terminal molt are input as data in the model. All immature individuals of Chionoecetes species are assumed to molt (i.e. z = 1). King crab do not have a molt to maturity, but the probability of molting varies with size. Here, a declining logistic function was estimated for each king crab species within the model:

$$z_s = 1 - \frac{1}{1 + exp(-m_s(size(s) - m_{50}))} \tag{13}$$

Where the size at 50% molting probability is *m_{50}* and the slope of the logistic function is *m_{s}*. Recruitment by year was estimated as a vector of deviations (τ_y) around a mean (μ) in log space and added to the first three size of classes of crab based on another vector that is determined by the estimated parameters δ_s with bounds 0-200 and determines the proportion allocated to each size bin ψ_s.

$$tot_d = 20 + \delta_1 + \delta_2 \tag{14}$$

$$\psi_1 = \frac{20}{tot_d} \tag{15}$$

$$\psi_{2,3} = \frac{\delta_d}{tot_d} \tag{16}$$

$$n_{(t=y+q/12,s=1-3,m=1)} = n_{(t=y+q/12,s=1-3,m=1)} + \psi_s e^{\mu+\tau_y} \tag{17}$$

Finally, the remaining other mortality is applied to the population after growth, molting, recruitment, and fishing occurs. Note that this allows a crab to experience two different mortalities within a given year as it undergoes terminal molt.

$$N_{t=y+1,s,m=1} = n_{t=y+q/12,s,m=1}e^{-\frac{12-q}{12}M_{t,s,m}} \tag{18}$$

$$N_{t=y+1,s,m=2} = n_{t=y+q/12,s,m=2}e^{-\frac{12-q}{12}M_{t,s,m}} \tag{19}$$

Survey selectivity priors

The observed numbers of crab at size by year in the NMFS survey reflect the ability of the trawl gear to capture the crab, also known as 'selectivity'. The selectivity of trawl gear can change according to size, and consequently needs to be accounted for in the population dynamics model when fitting to the survey data. Survey selectivity was estimated for snow crab and Tanner crab by fitting the model while accounting for priors based on the inferred selectivity from experimental *Nephrops* trawls (a small trawl net designed to maintain bottom contact), operated by the Bering Sea Fisheries Research Foundation in collaboration with the NMFS summer survey. The experimental trawls were performed at the same time and location as the NMFS summer survey tows to evaluate the efficiency of the NMFS survey trawl gear at capturing snow crab (Somerton et al., 2013). The *Nephrops* gear used by the BSFRF was assumed to capture all crab in its path given strong bottom contact. The resulting area-swept estimates of numbers of crab at size from the BSFRF and NMFS surveys ($\hat{N}_{y,s,NMFS}$ and $\hat{N}_{y,s,NMFS}$, respectively) can be used to generate priors for the selectivity of the NMFS gear in year y ($S_{y,NMFS}$) as:

$$S_{y,NMFS} = \frac{\hat{N}_{y,s,NMFS}}{\hat{N}_{y,s,BSFRF}} \tag{20}$$

The experimental trawls captured snow crab in the years 2010, 2011, 2016, 2017, and 2018, but the spatial foot print and sample sizes varied by year. The calculated selectivities by size and by year were fairly consistent for snow crab of carapace widths 40 - 95 mm, but the signal was less consistent for crab larger than ~100 mm carapace width given the more sparse nature of the data in those size ranges.

Survey selectivity for snow crab and Tanner crab is estimated based on a functional form that allows for a stair-stepped pattern in survey selectivity like that suggested by the BSFRF data. Five parameters are estimated: two that determine separate logistic curves (α and β parameters) and one that determines where the stair-step occurs with respect to size (ω).

$$\hat{S}_{NMFS} = \frac{\omega}{(1 + exp(-\alpha_1(size - \beta_1)))} + \frac{1-\omega}{(1 + exp(-\alpha_2(size - \beta_2)))} \tag{21}$$

Estimated selectivity for the survey fit these prior data reasonably given the uncertainty assigned to the data source. All king crab stocks have an estimated logistic survey selectivity with no priors imposed.

Objective function

The objective function for the population dynamics model consists of likelihood components (representing the fit of the model to the data) and penalty components (which incorporate constraints in the fitting based on auxiliary information) that are summed and minimized in log space to estimate parameters within the model. Several data sources were fit to using the following likelihoods. Observed size composition data crab caught in the survey and retained and discarded crab in the fishery were fit using multinomial likelihoods and were implemented in the form:

$$L_x = \lambda_x \sum_y N_{x,y} \sum_l p^{obs}_{x,y,l} ln(\hat{p}_{x,y,l}/p^{obs}_{x,y,l}) \tag{22}$$

Where L_{x} was the likelihood associated with data component x, λ_x represented an optional additional weighting factor for the likelihood, $N_{x,y}$ was the sample sizes for the likelihood, $p^{obs}_{x,y,l}$ was the observed proportion in size bin l during year y for data component x, and $\hat{p}_{x,y,l}$ was the predicted proportion in size bin l during year y for data component x. Sample sizes were input as 40 to balance the contributions of the size composition data and the indices.

Observed indices of abundance in the survey and retained and discarded catches (in numbers) in the fishery were fit with log normal likelihoods implemented in the form:

$$L_x = \lambda_x \sum_y \frac{(ln(\hat{I}_{x,y}) - ln(I_{x,y}))^2}{2\sqrt{ln(1+CV_{x,y}^2)}} \tag{23}$$

L_x was the contribution to the objective function of data component x, λ_x was any additional weighting applied to the component, $\hat{I}_{x,y}$ was the predicted value of quantity I from data component x during year y, I$_{x,y}$ was the observed value of quantity I from data component x during year y and CV$_{x,y}$ was the coefficient of variation for data component x during year y. See data files in the github repo for values.

Penalties and priors

Smoothing penalties were placed on estimated vectors of deviations for other mortality using normal likelihoods on the second differences of the vectors. Normal priors were also placed on the mean value of other mortality (see **??**) and the deviation of the estimated mortality from that mean. Prior values for the average other mortality based are generally based on assumed maximum ages and Hamel's (2015) empirical analysis of life history correlates with mortality. The normal priors were of the form:

$$P_x = \lambda_x \sum_y \frac{((\hat{I}_{x,y}) - (I_{x,y}))^2}{sqrt2\sigma_{x,y}^2} \tag{24}$$

P_x was the contribution to the objective function of the penalty associated with model estimate x, λ_x was any additional weighting applied to the component, $\hat{I}_{x,y}$ was the predicted value of population process I relevant to penalty x during year y, I$_{x,y}$ was the prior value of process I relevant to penalty x during year y and $\sigma_{x,y}$ was the input sigma for penalty x during year y. See data files in the github repo for values.

Fits and diagnostics

These models were fit using AD Model Builder (ADMB; Fournier et al., 2012). Variances associated with estimated time series of mortality were fixed at relatively large values to allow the model flexibility to capture changes in observed abundance (see data files in github repository for values used). Points to discuss: desired diagnostic performance, simulation testing, and model complexity.

Snow crab Figure S7 Figure S8 Figure S9 Figure S10 Figure S11 Figure S12 Figure S13

Tanner crab Figure S14 Figure S15 Figure S16 Figure S17 Figure S18 Figure S19 Figure S20

Bristol Bay red king crab Figure S21 Figure S22 Figure S23 Figure S24 Figure S25 Figure S26

Saint Matthews blue king crab Figure S27 Figure S28 Figure S29 Figure S30 Figure S31 Figure S32

Pribilof Islands red king crab Figure S33 Figure S34 Figure S35 Figure S36

Pribilof Islands blue king crab Figure S37 Figure S38 Figure S39 Figure S40

The models presented here can be sensitive to small changes in weighting of data components. Changing the weights to size composition data in the survey, for example, can result in a model that produces nearly identical parameter estimates and derived quantities, but does not produce an invertible Hessian (which is a criteria of convergence). It is not clear if this sensitivity is a function of the model structure or ADMB's optimization algorithms.

Generalized additive models

Structure

Generalized additive models (GAMs) were used in the R programming language (package *mgcv*; Wood, 2011) to relate changes in estimated other mortality by maturity state, $m_{m,y}$, and recruitment r_y to a vector including environmental covariates and mature population density, ϕ_y. GAMs were also fit to observed maturity data using relevant covariates to explain variability over time. GAMs are useful for our purposes because of their flexibility in fitting potential non-linear relationships, given we do not have an a priori expectation of the functional form between responses and covariates.

$$m_{m,y} = s(\phi_{m,y}) + \epsilon_i \tag{25}$$

where ‘s()’ is a smoothing function based on thin-plate splines, ϕ is a matrix of environmental covariates (each of which have a separate smooth estimated), and ϵ is unexplained error. Models were compared via Akaike’s Information Criterion (AIC; Akaike; 1974). The specified families used as link functions when fitting the GAMs were a Tweedie distribution. The number of knots allowed per smooth was constrained to 4 to minimize overfitting and facilitate biological interpretation. Model diagnostics were performed by simulating residuals using the R package ‘DHARMa’ which is useful when using link functions other than Gaussian.

Fits and diagnostics

AIC, DHARMa, crossvalidation to come.

Table S1: Specifications for models by population.

Population	Sizes	M prior	Maturity	Years	Fish time
Snow	30-135	0.27	Chela-based	1982-2023	Feb
Tanner	25-185	0.27	Chela-based	1975-2023	Feb
BBRKC	45-170	0.18	>120 mm	1975-2023	Sep
PIRKC	45-220	0.18	>120 mm	1976-2023	Sep
SMBKC	45-170	0.18	>105 mm	1978-2023	Sep
PIBKC	45-200	0.18	>120 mm	1975-2023	Sep

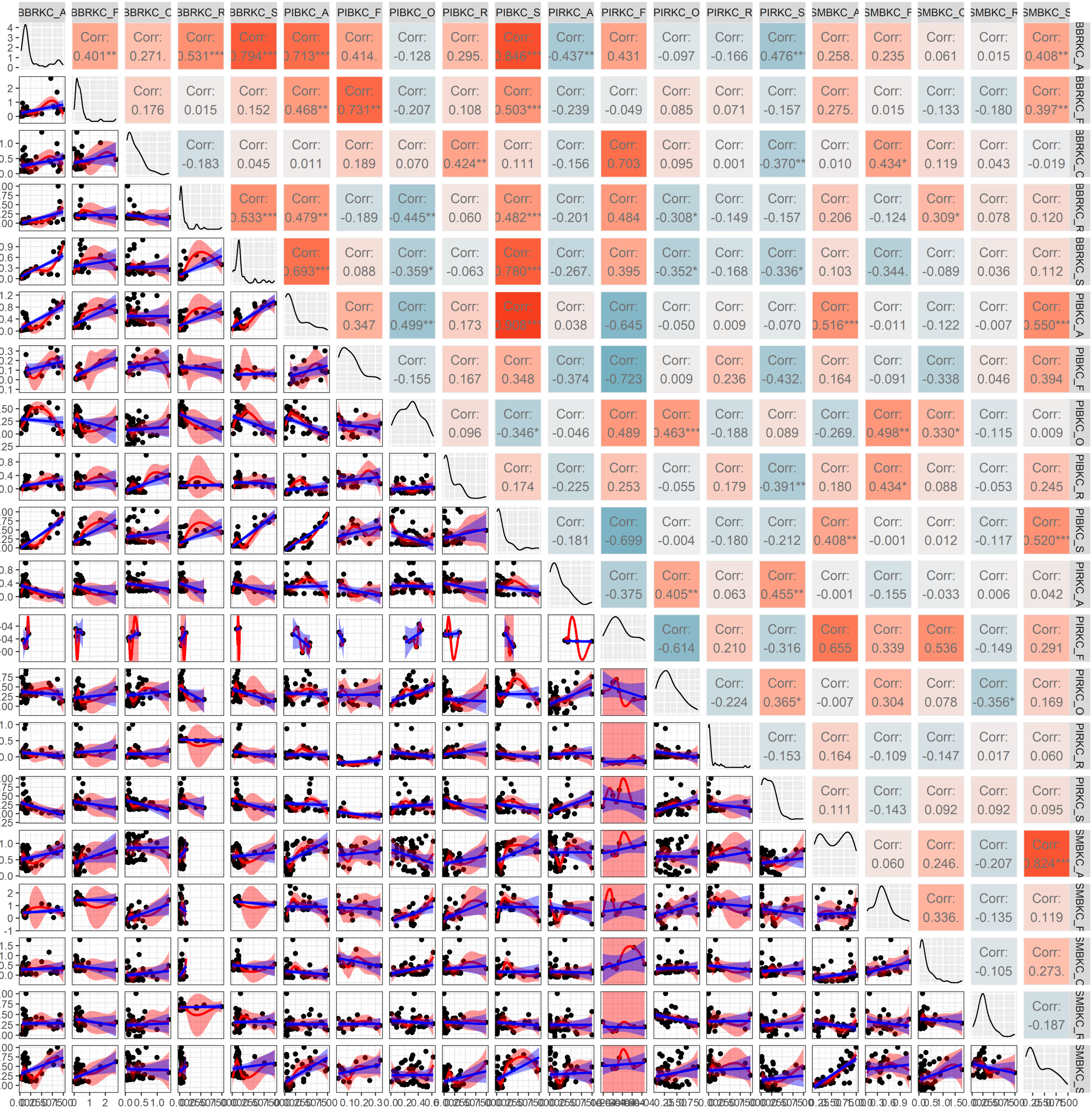

Figure S1: Correlations between estimated population processes for king crab species. The letter after a stock name (A,F,O,R,S) represent Total abundance, Fishing mortality, Other mortality, Recruitment, and Spawning abundance, respectively.

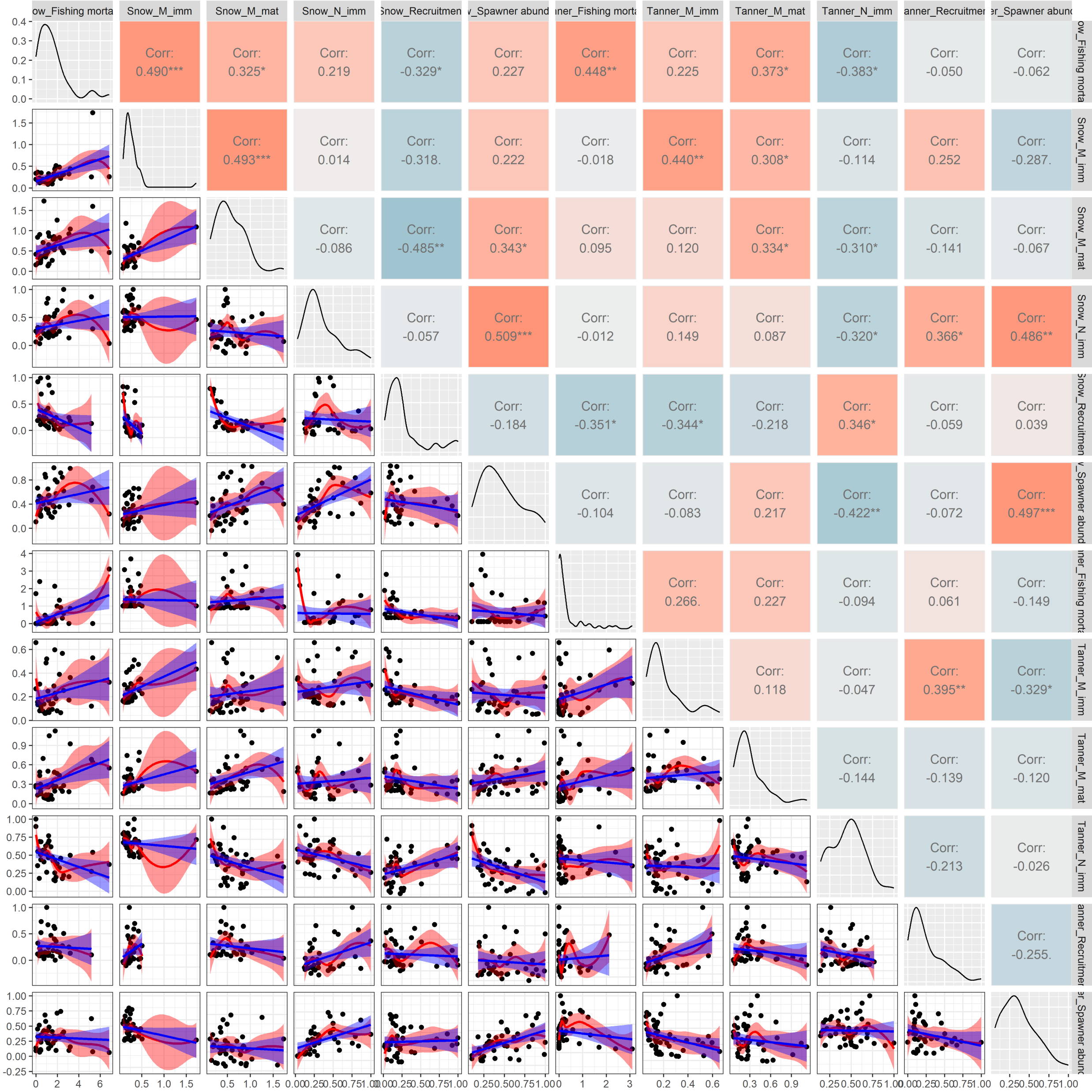

Figure S2: Correlations between estimated population processes for Chionoecetes crab species.

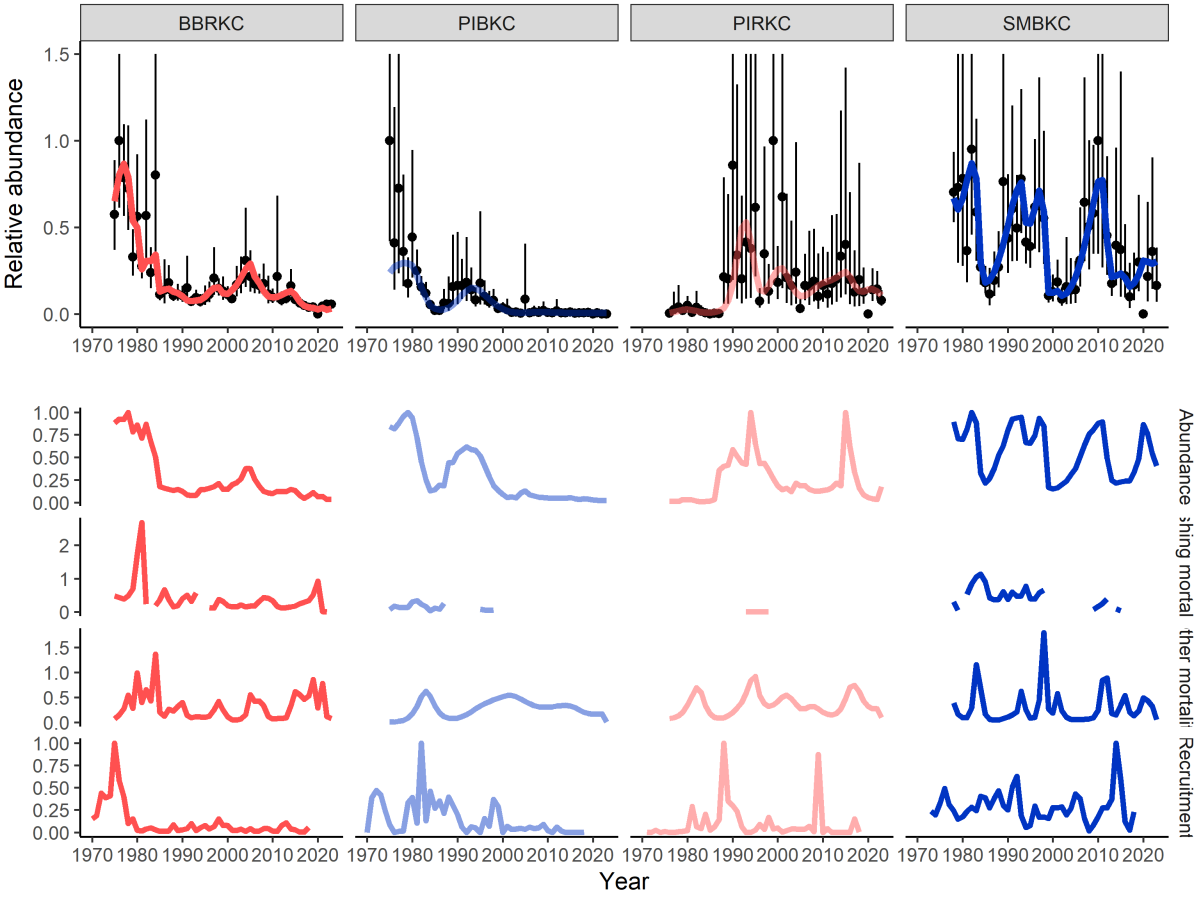

Figure S3: Comparison of model fits and estimated recruitment, other mortality, and fishing mortality for king crab species.

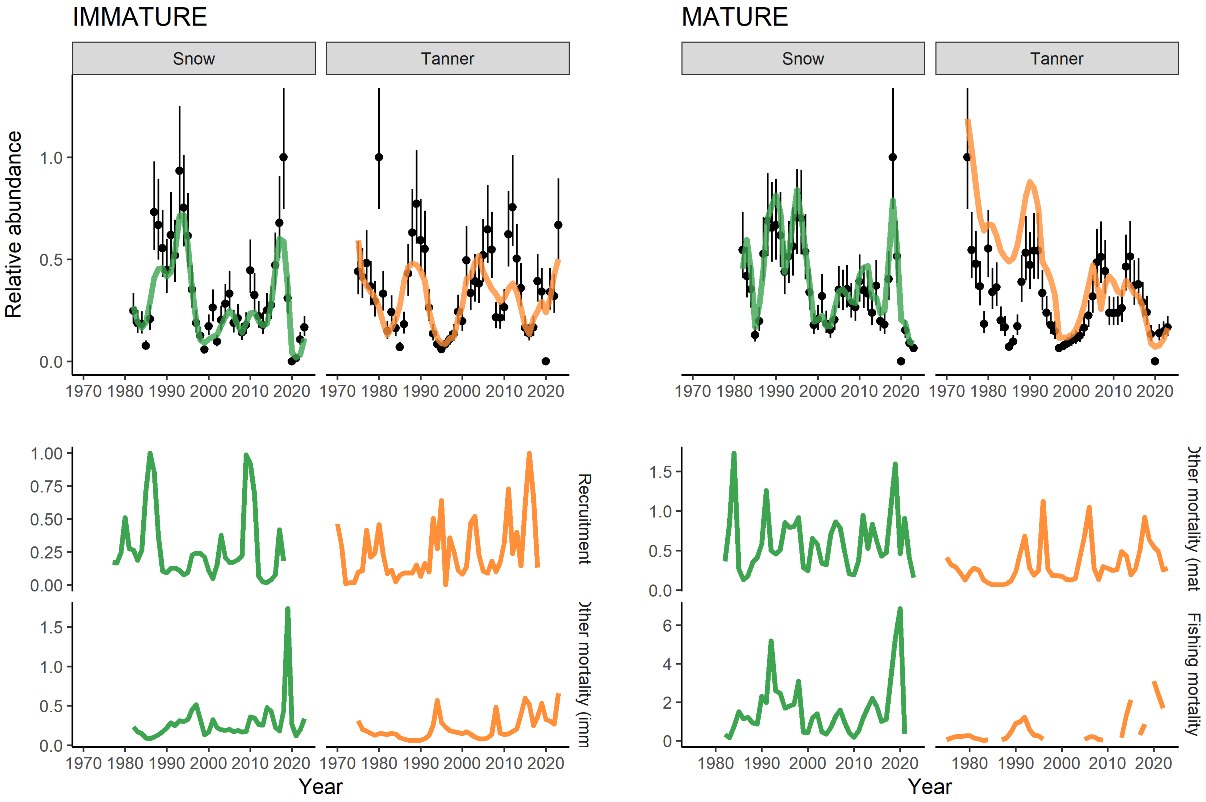

Figure S4: Comparison of model fits and estimated recruitment, other mortality, and fishing mortality for Chionoecetes crab species.

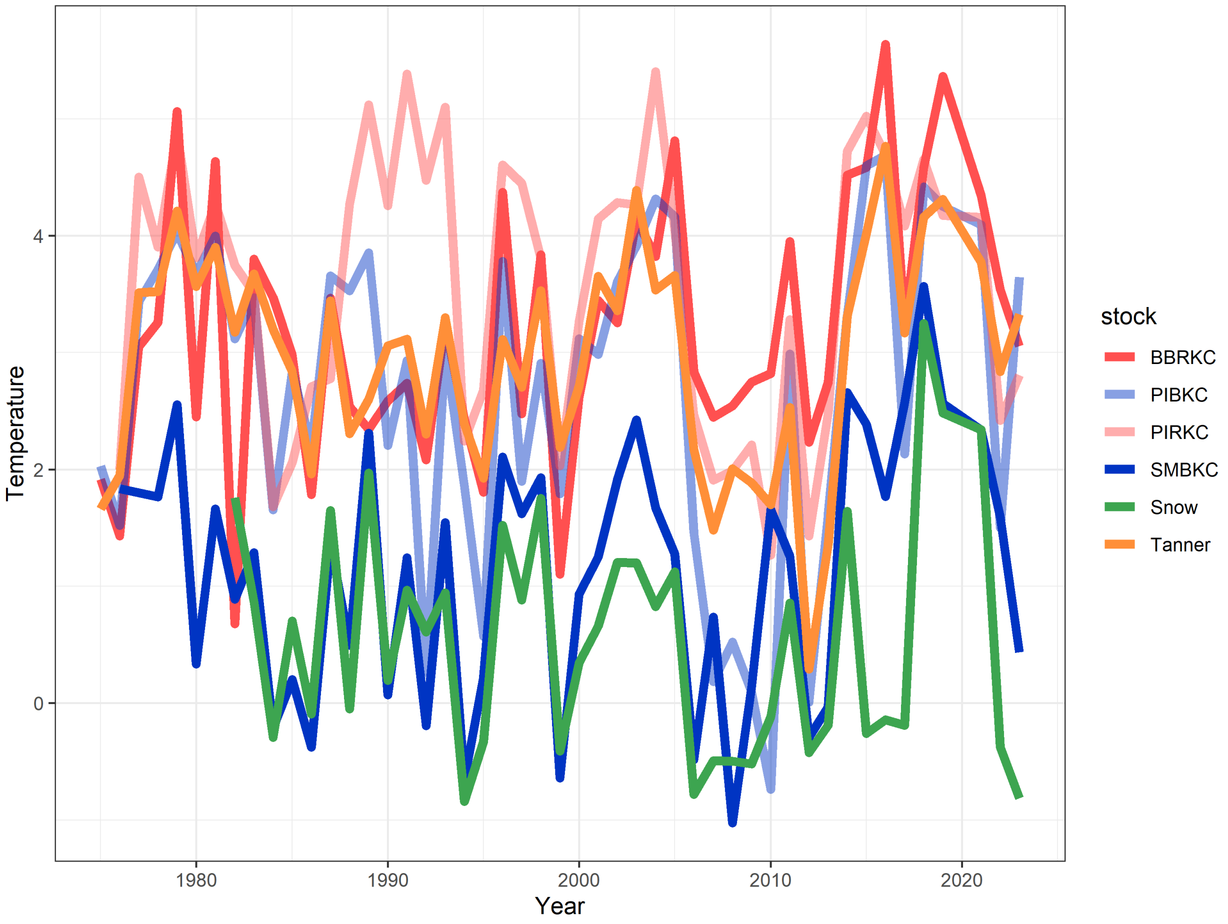

Figure S5: Temperature occupied in degrees celcius by population.

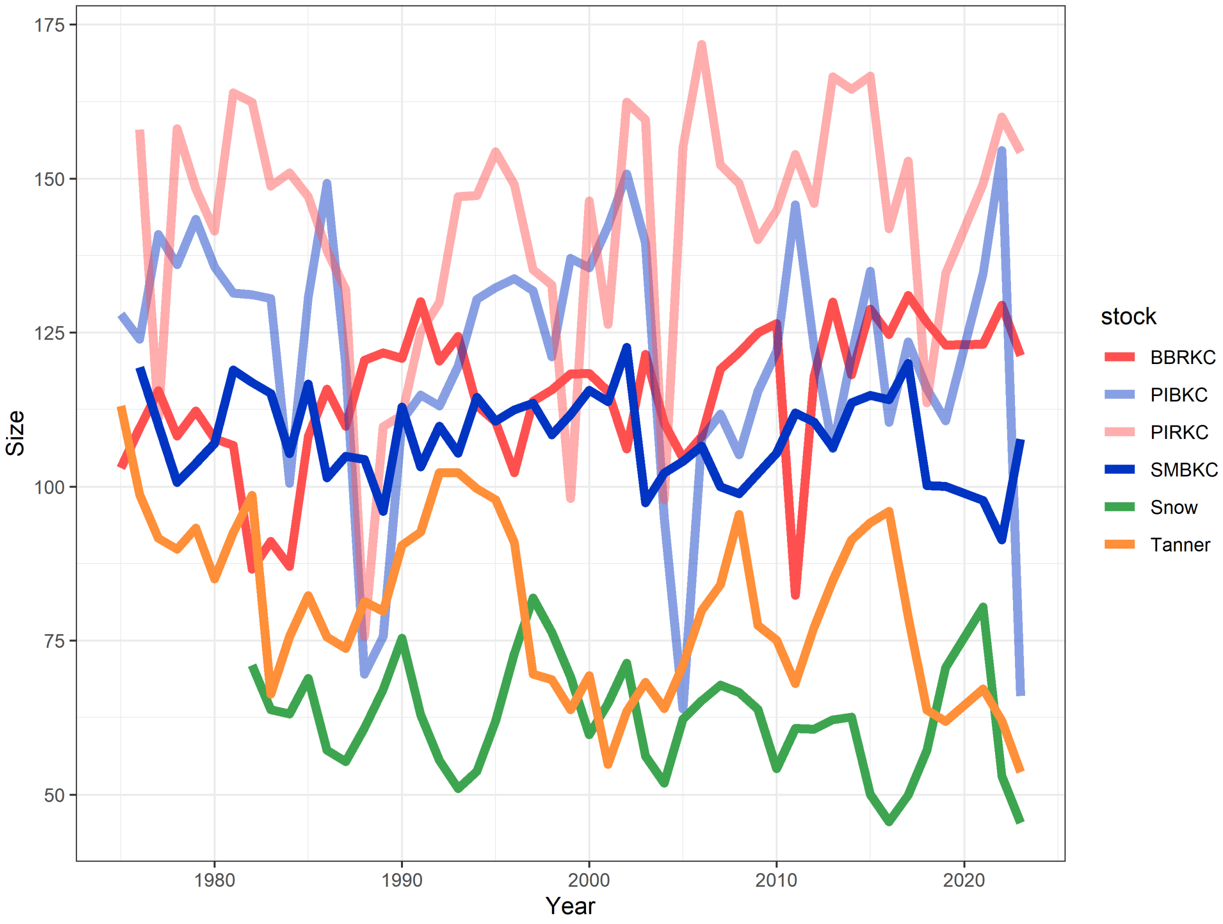

Figure S6: Average modeled size by population.

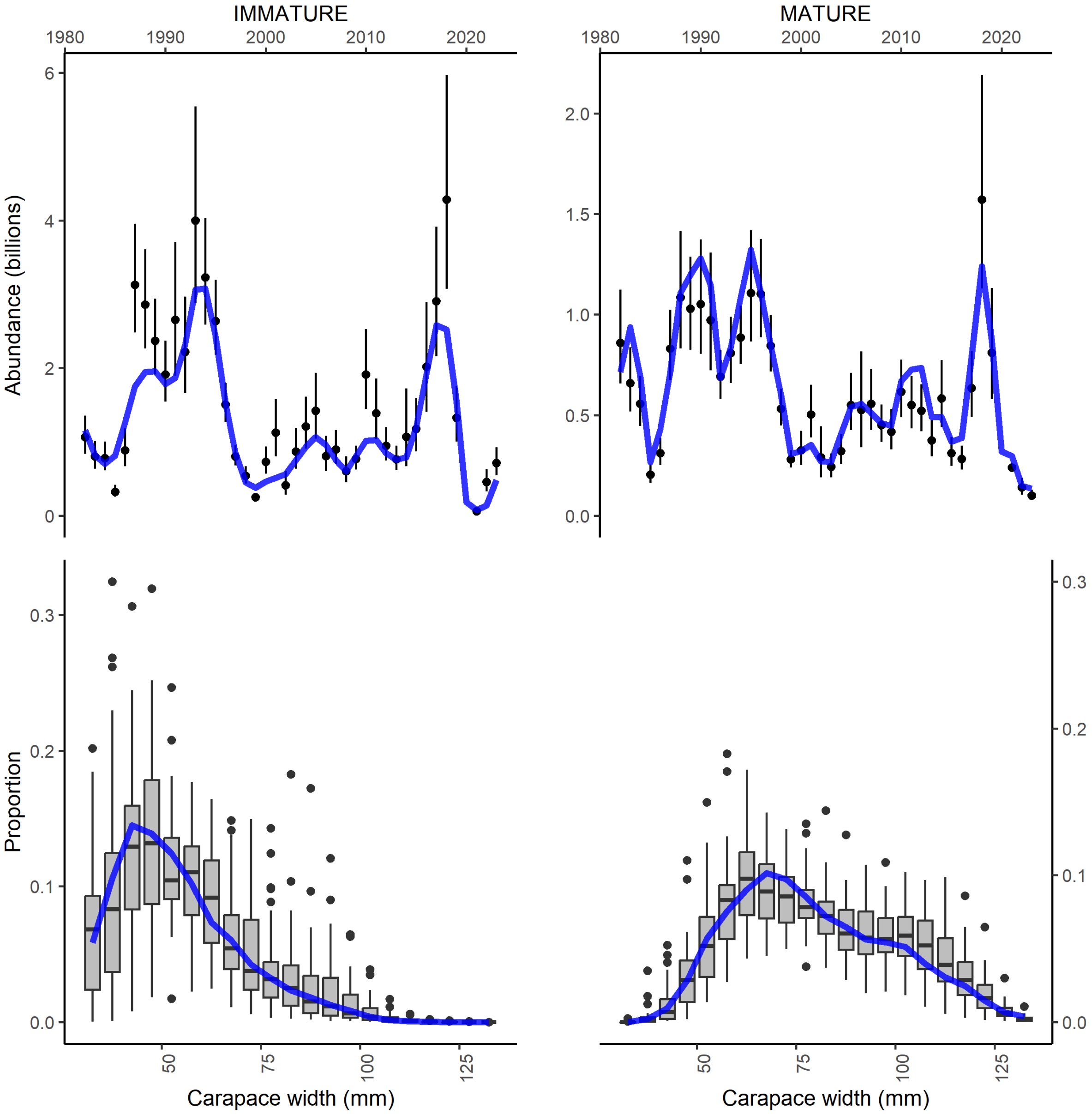

Figure S7: Fits to the survey indices of abundance and size composition data for snow crab.

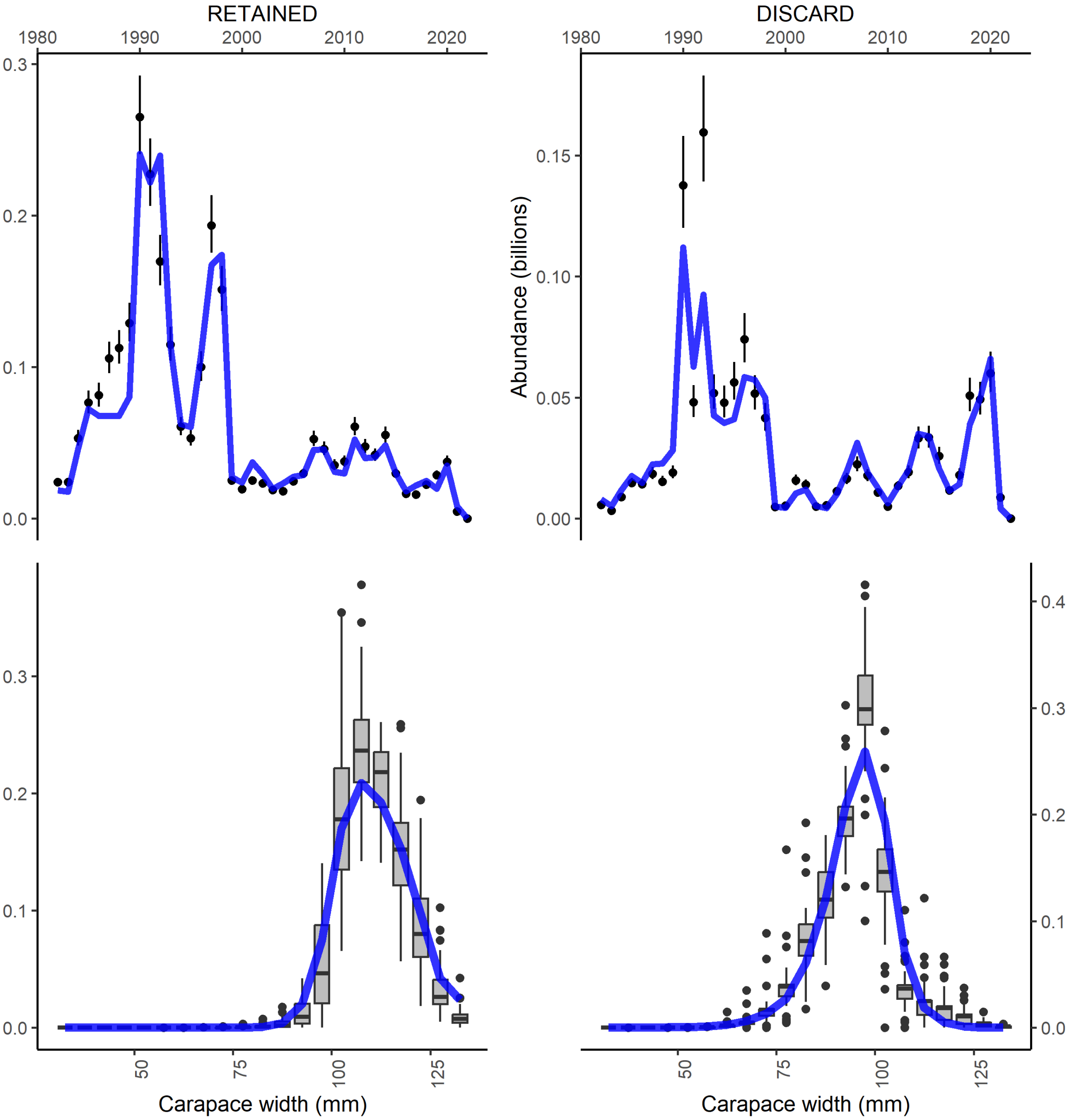

Figure S8: Fits to the retained and discarded abundances in the directed snow crab fishery.

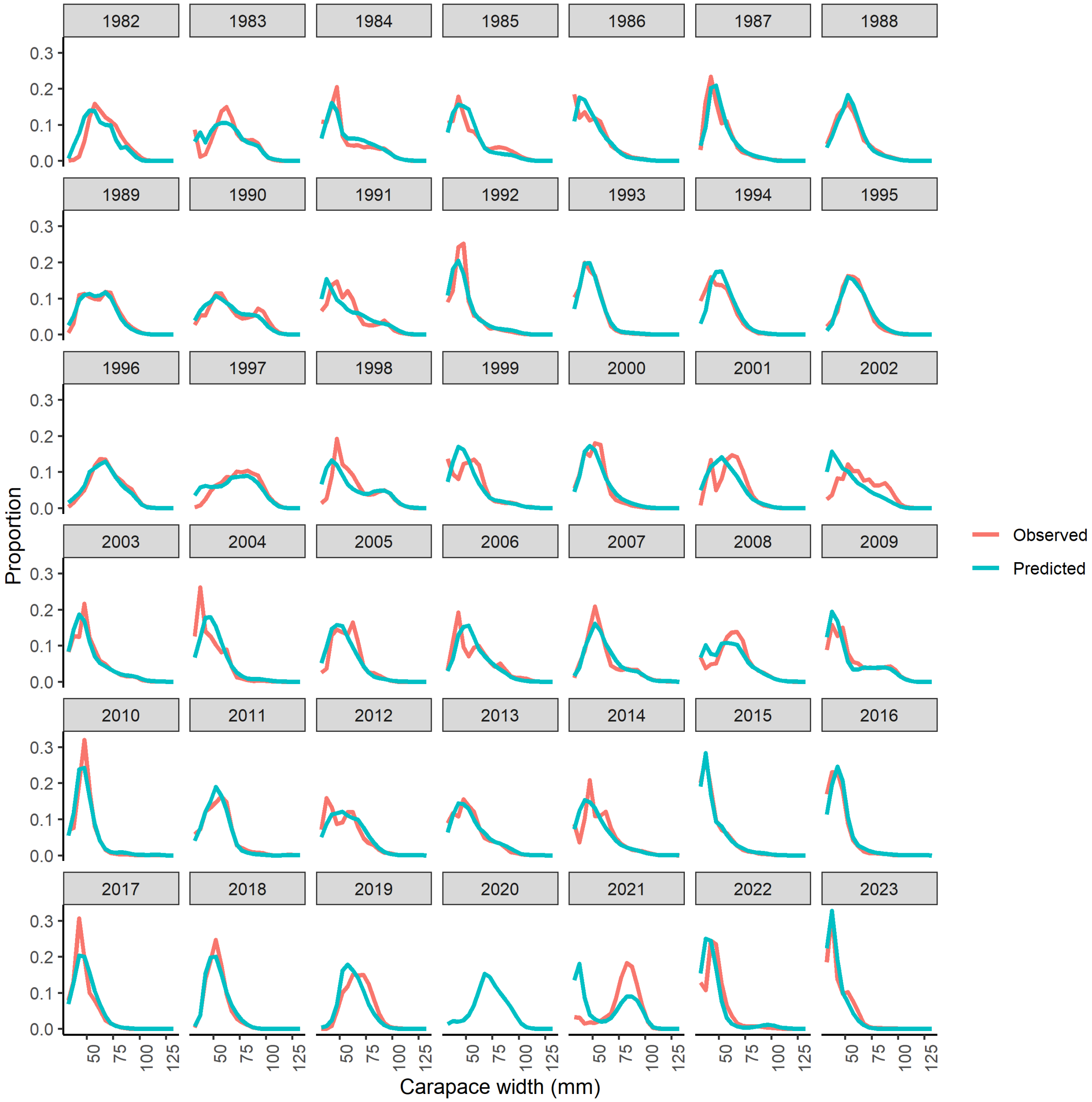

Figure S9: Model fits to individual years of size composition data for immature male snow crab captured in the survey.

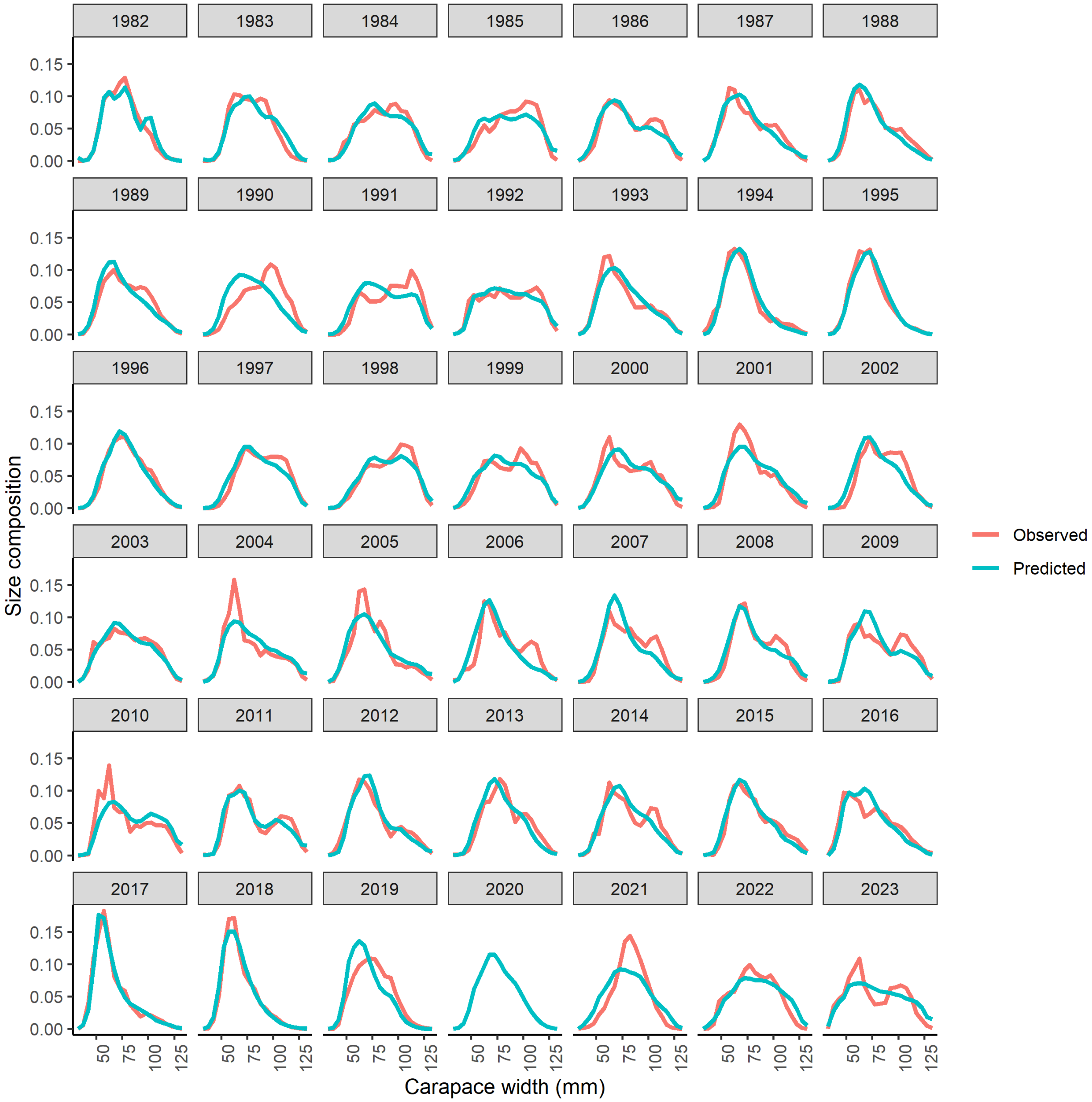

Figure S10: Model fits to individual years of size composition data for mature male snow crab captured in the survey.

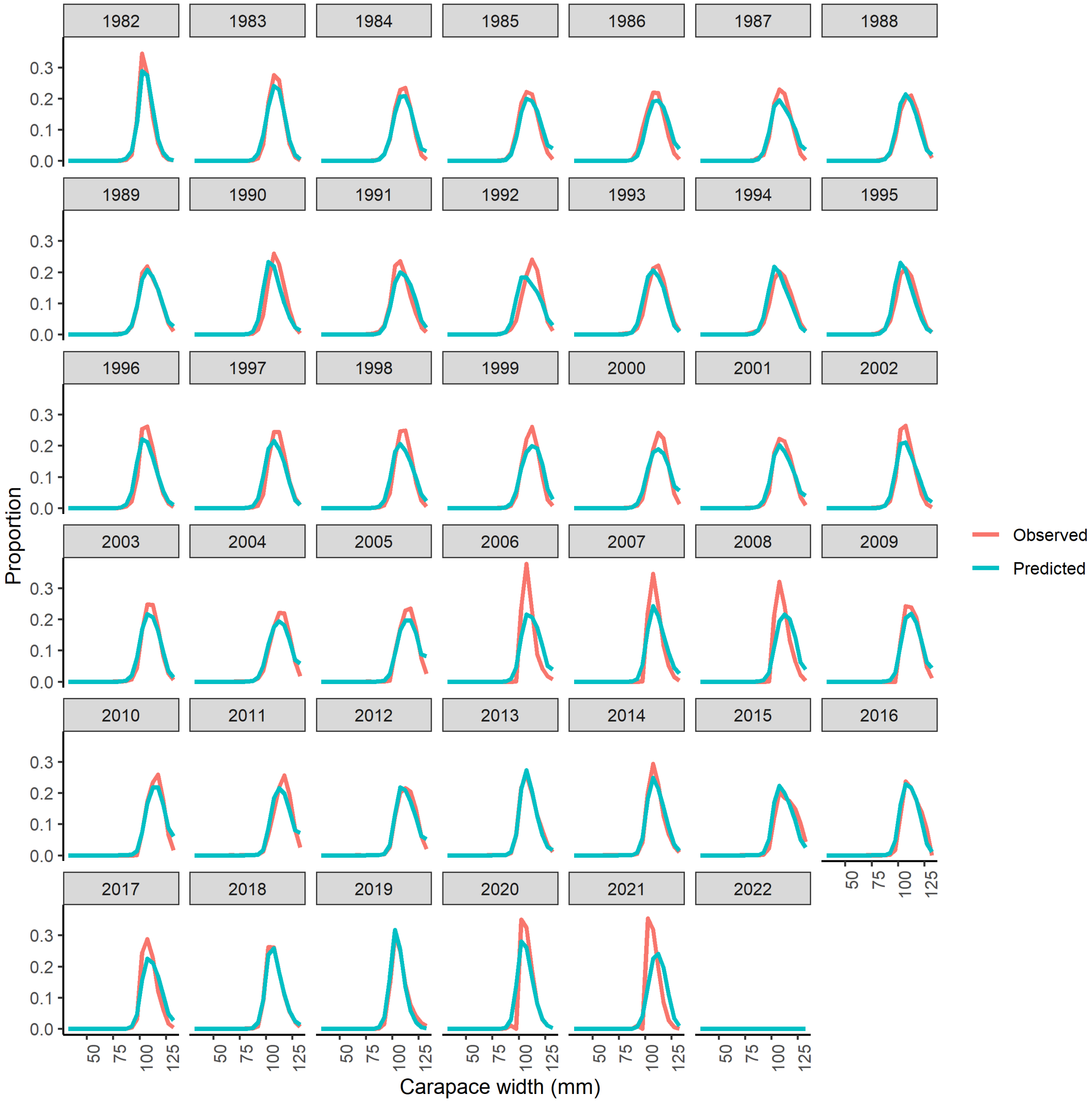

Figure S11: Model fits to individual years of size composition data for retained snowcrab in the directed fishery.

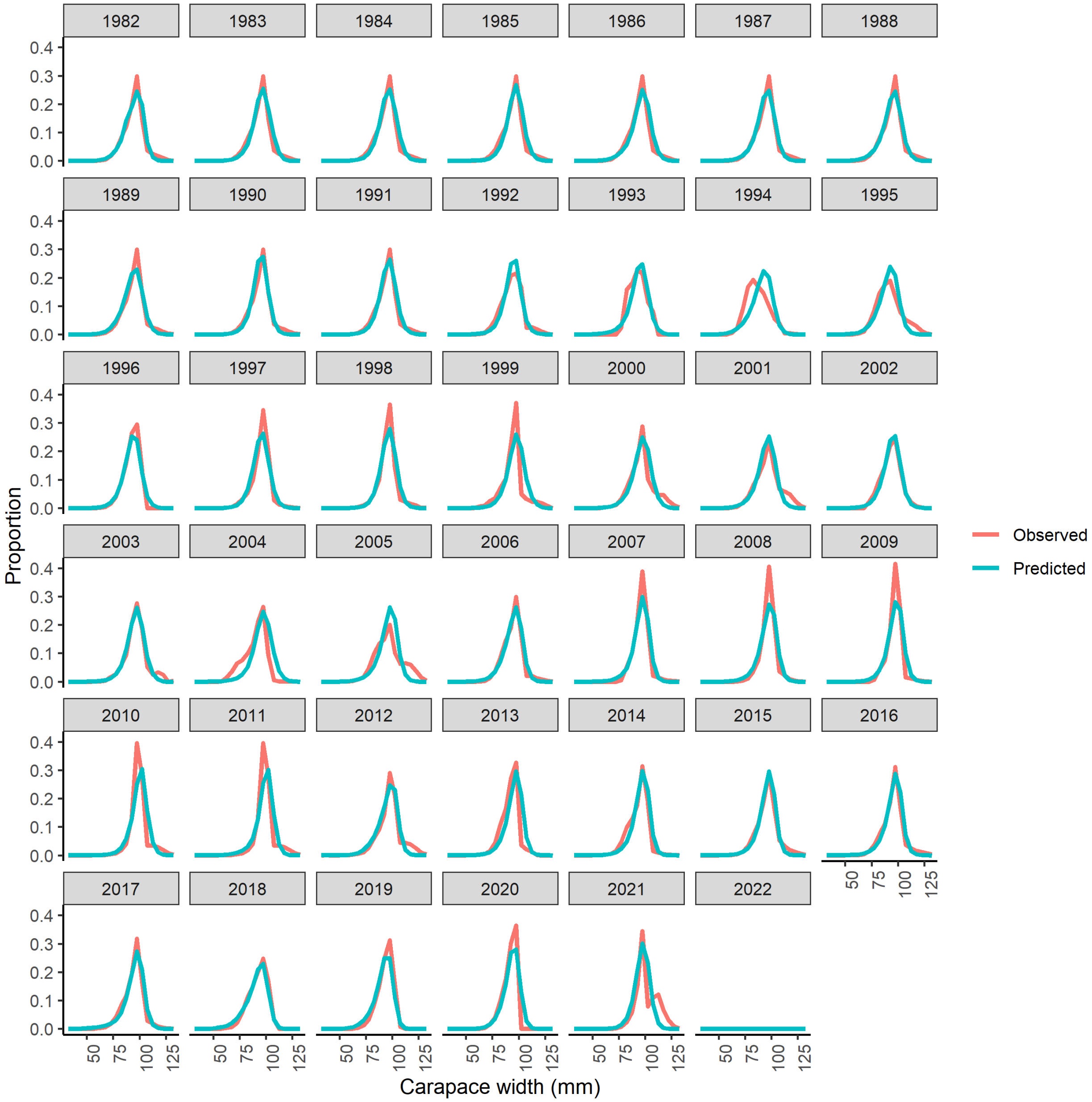

Figure S12: Model fits to individual years of size composition data for discarded snow crab in the directed fishery.

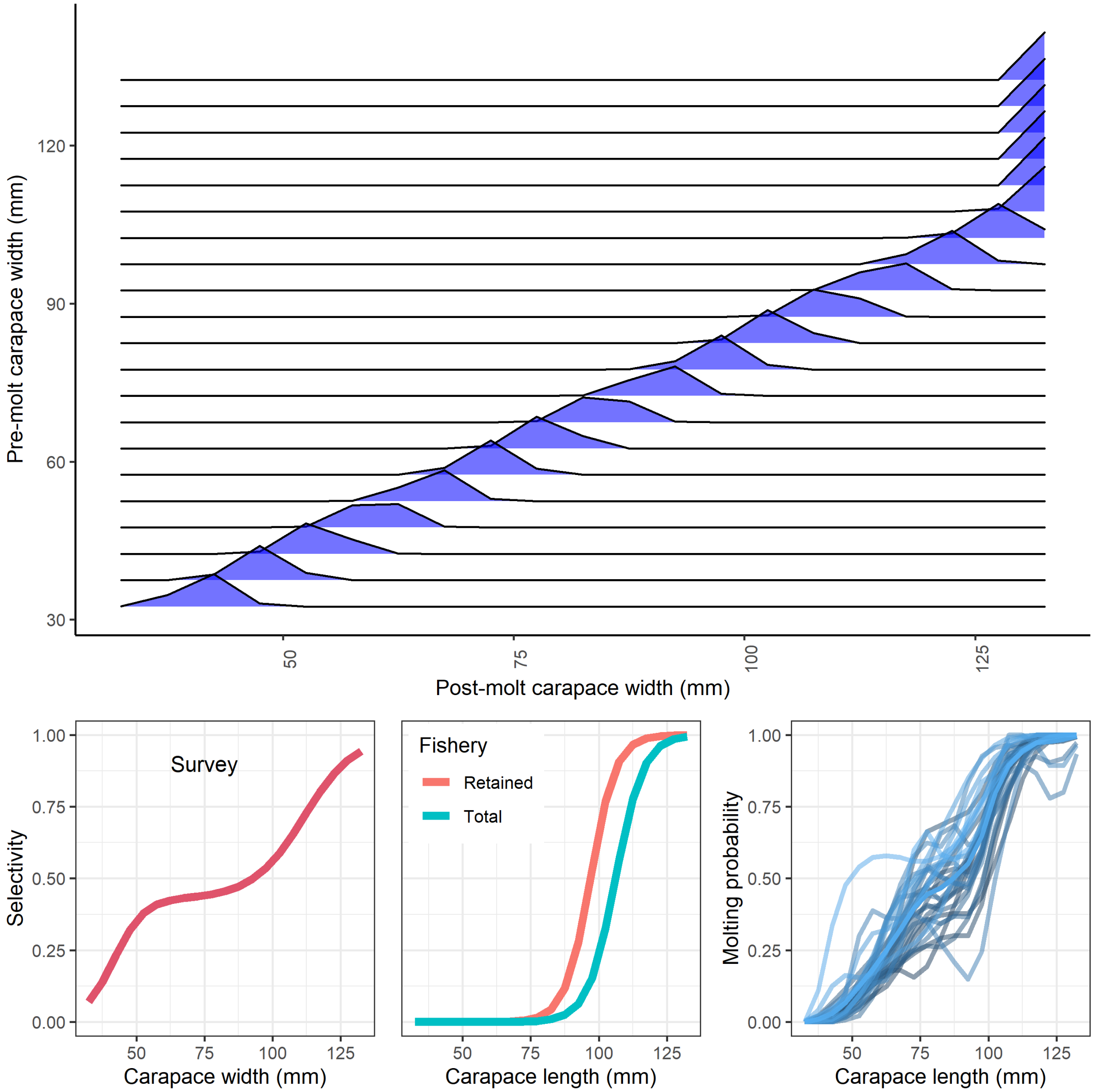

Figure S13: Estimates of population processes for snow crab.

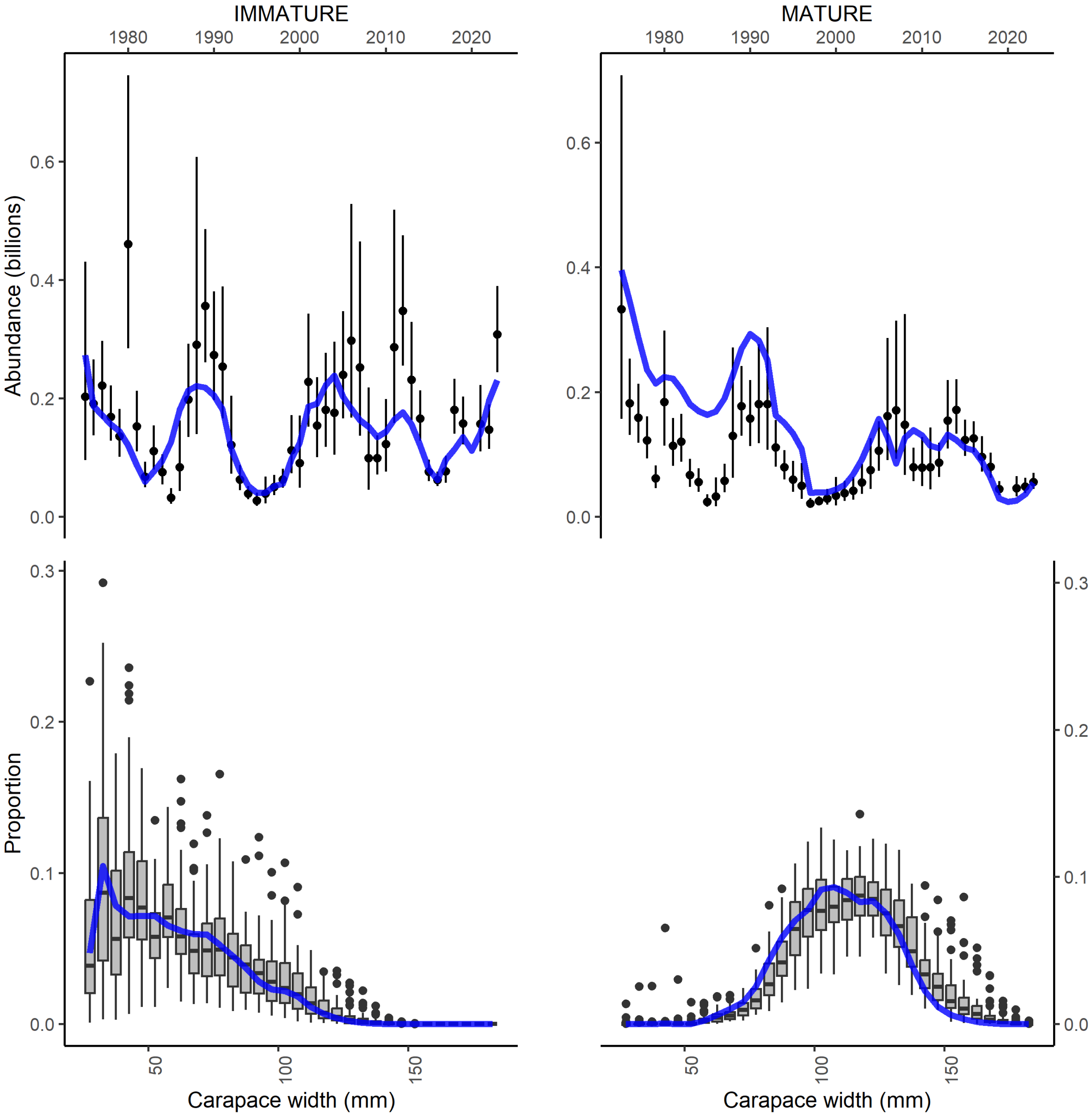

Figure S14: Fits to the survey indices of abundance and size composition data for Tanner crab.

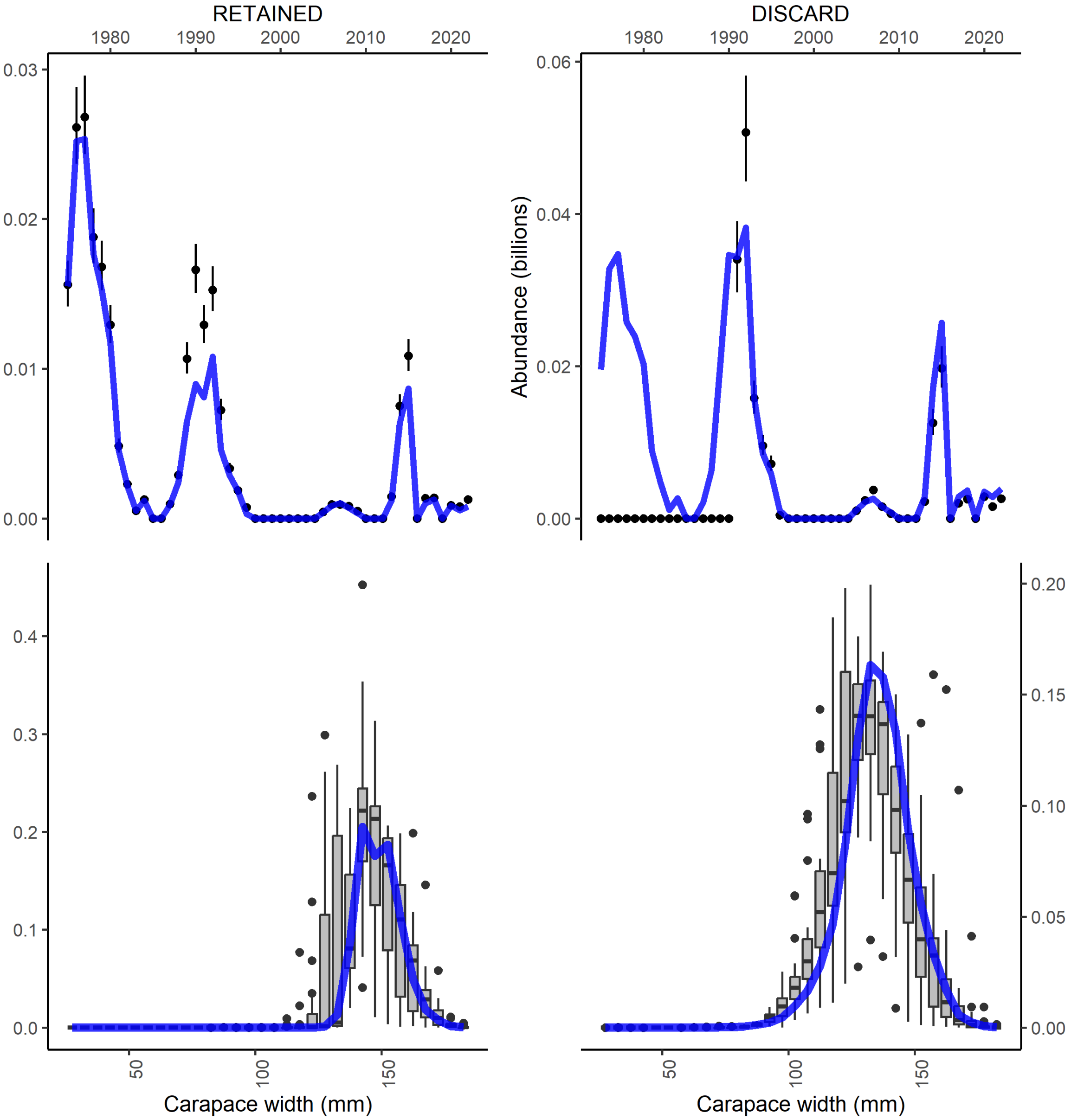

Figure S15: Fits to the retained and discarded abundances in the directed Tanner crab fishery.

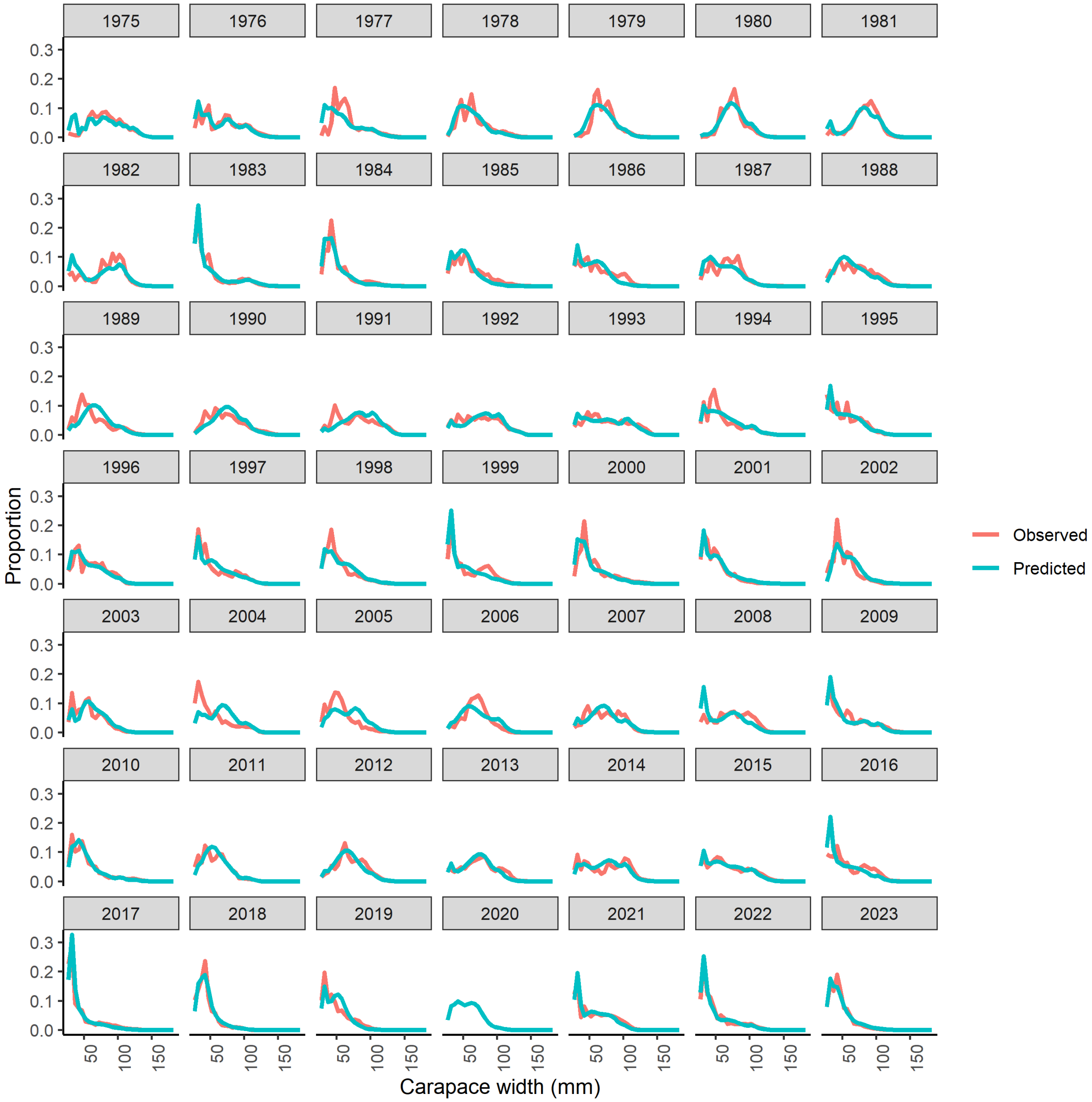

Figure S16: Model fits to individual years of size composition data for immature male Tanner crab captured in the survey.

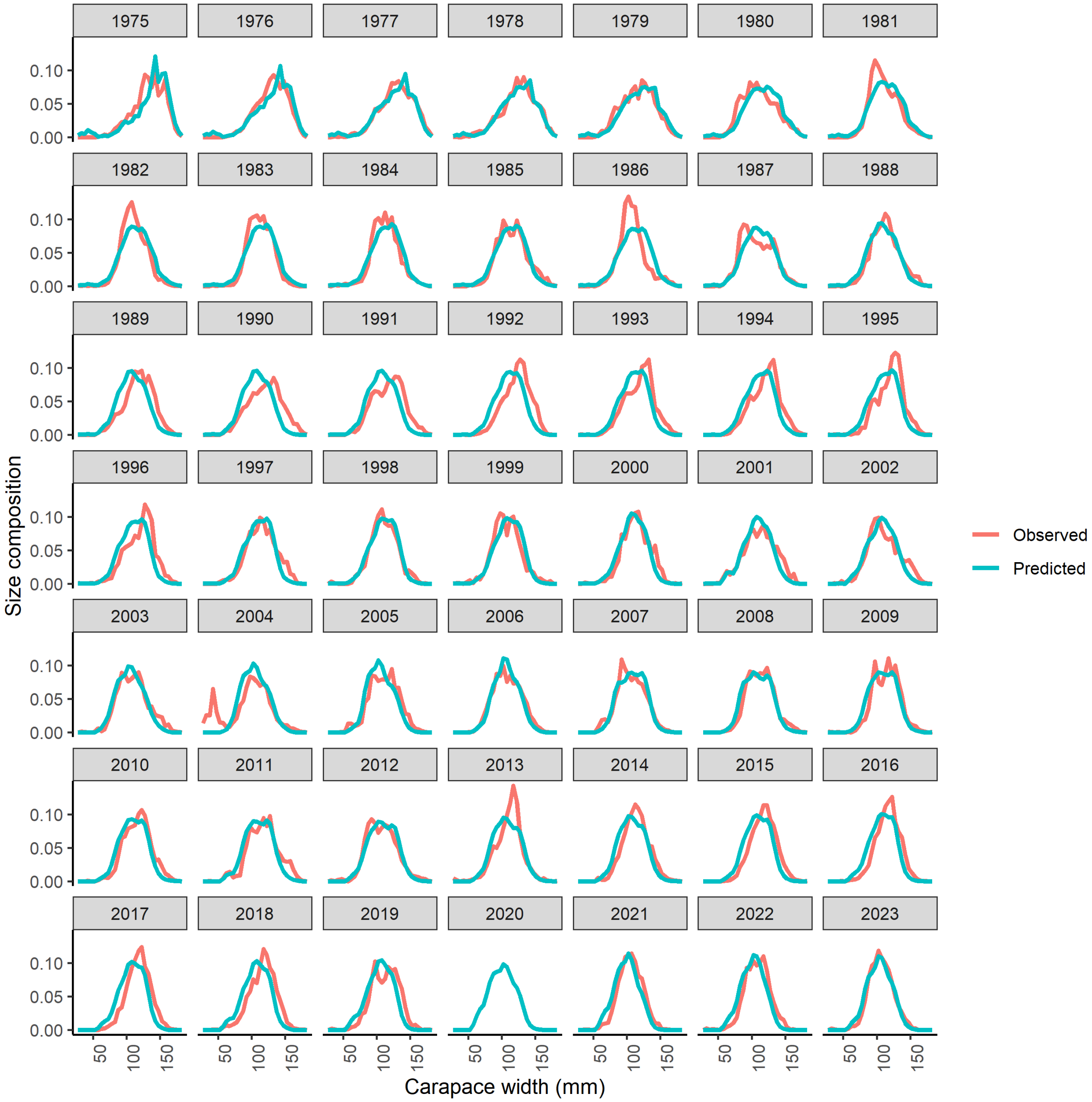

Figure S17: Model fits to individual years of size composition data for mature male Tanner crab captured in the survey.

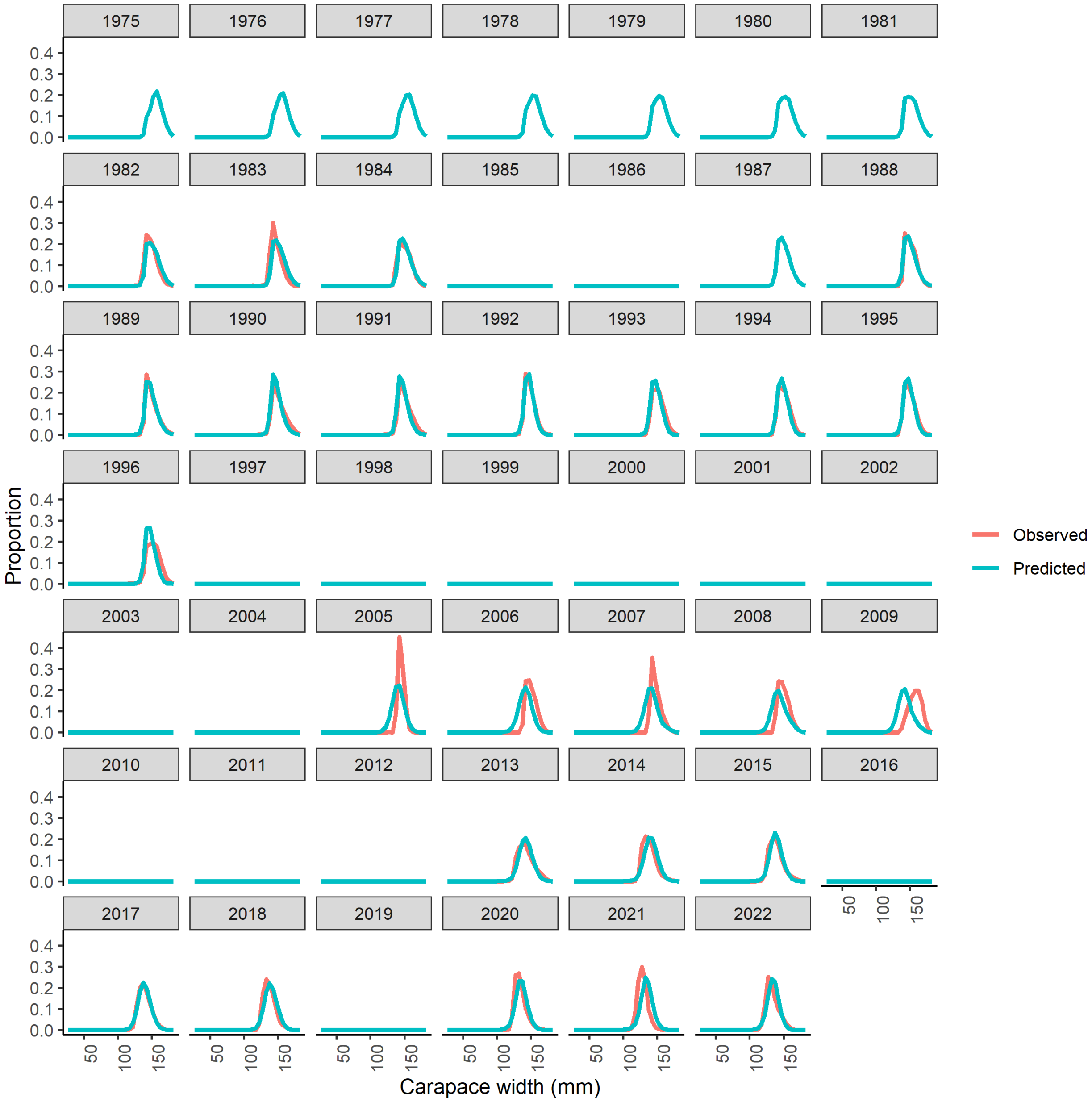

Figure S18: Model fits to individual years of size composition data for retained Tanner crab in the directed fishery.

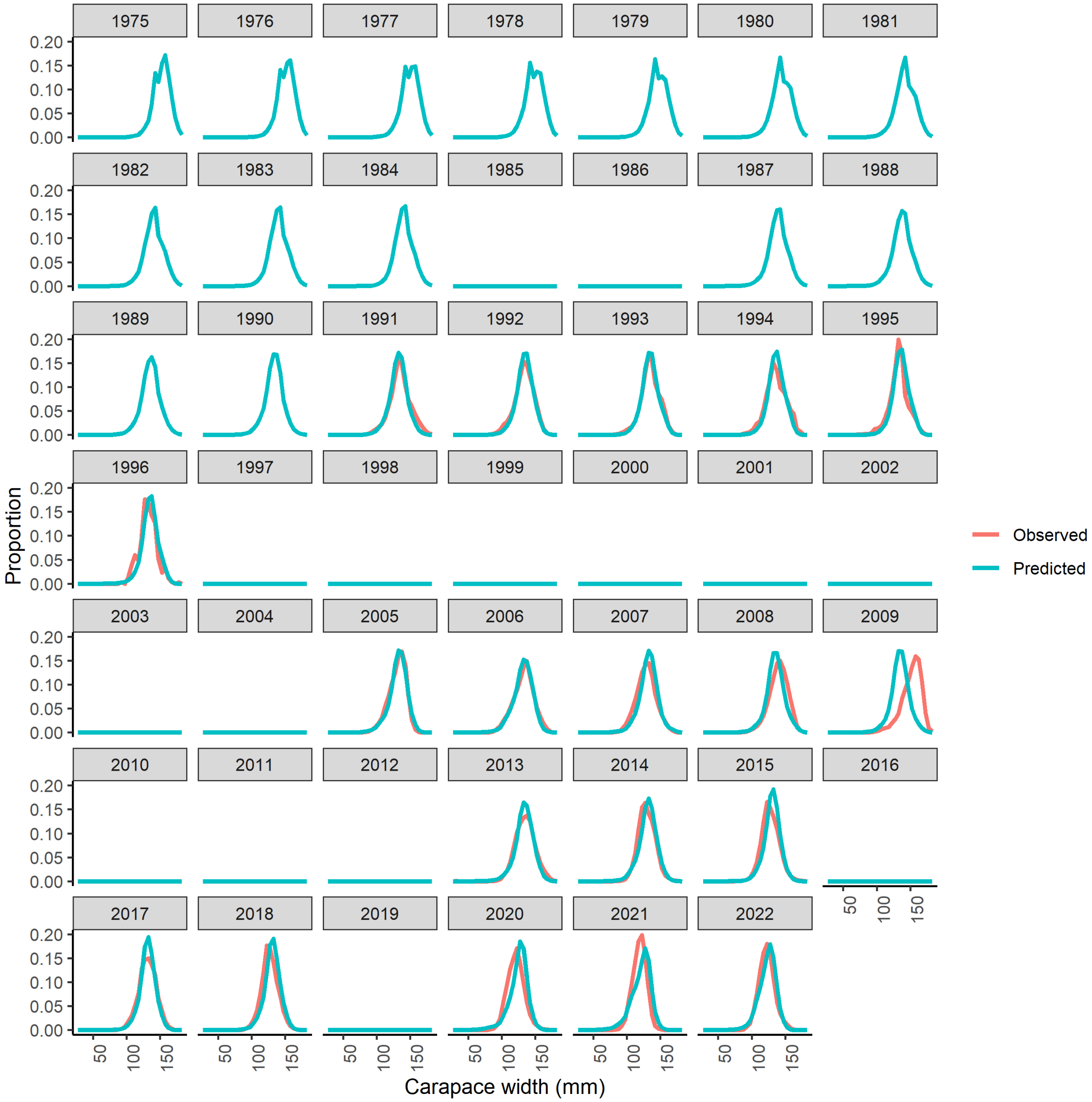

Figure S19: Model fits to individual years of size composition data for discarded Tanner crab in the directed fishery.

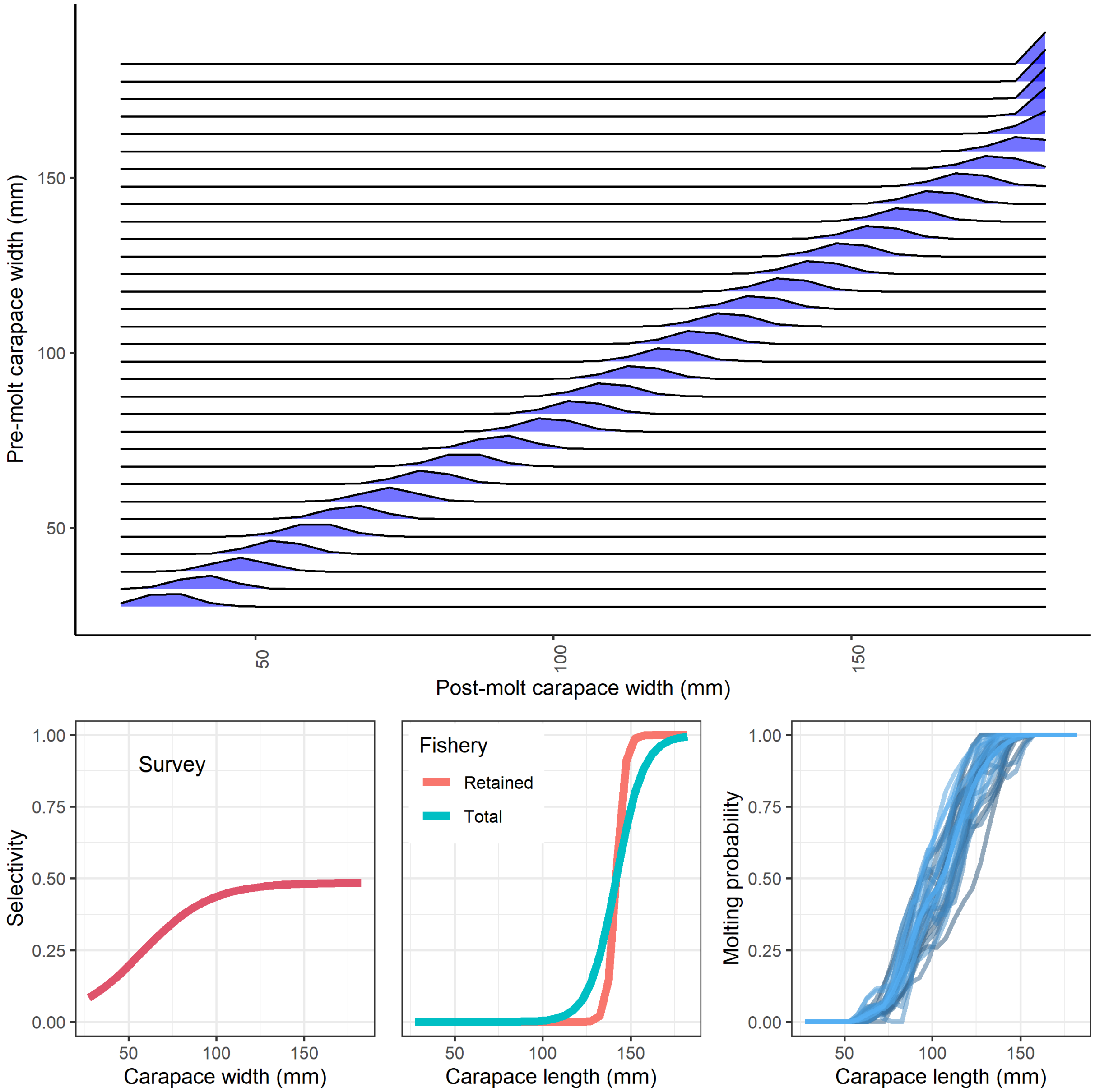

Figure S20: Estimates of population processes for Tanner crab.

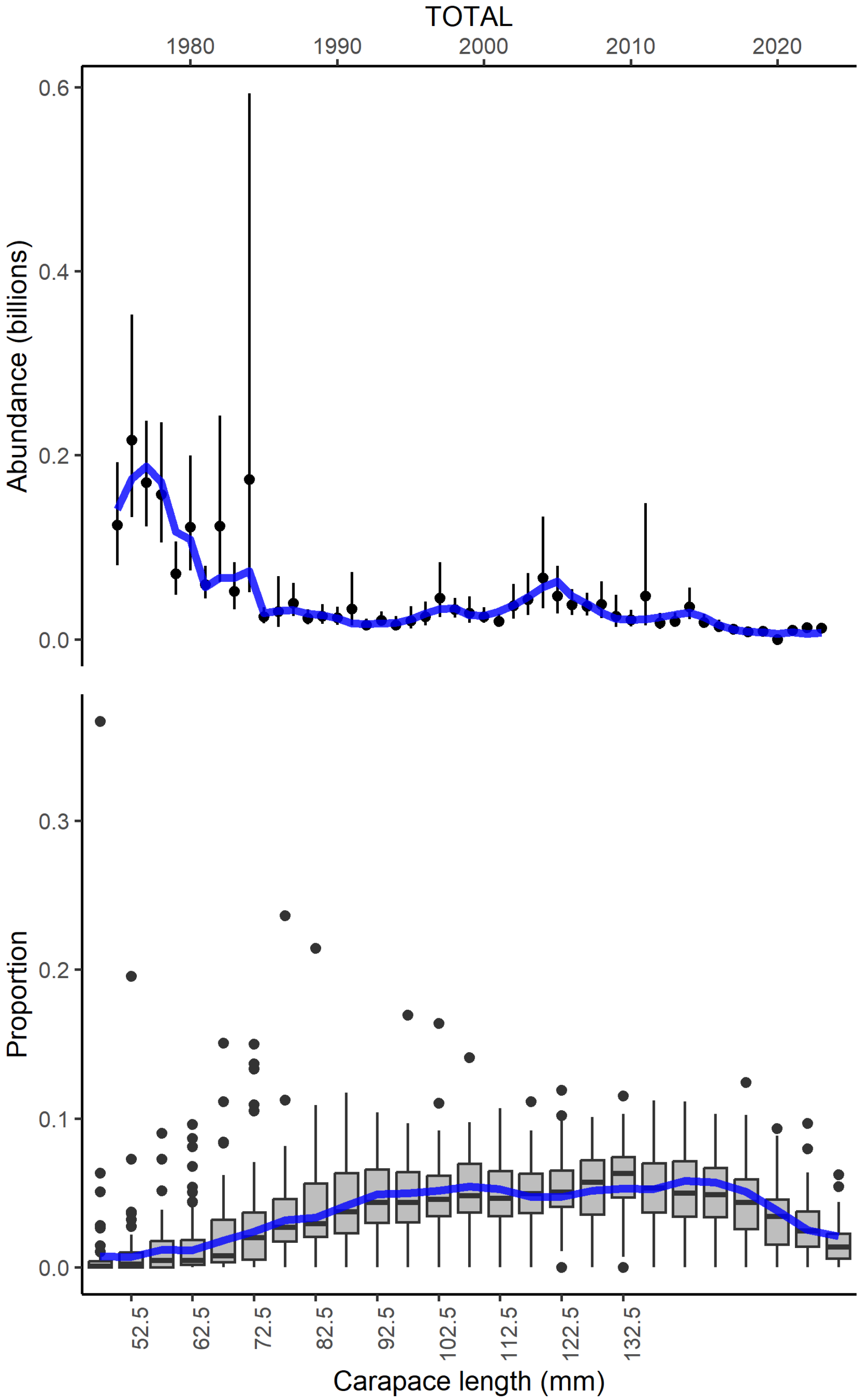

Figure S21: Fits to the survey indices of abundance and size composition data for Bristol Bay red king crab.

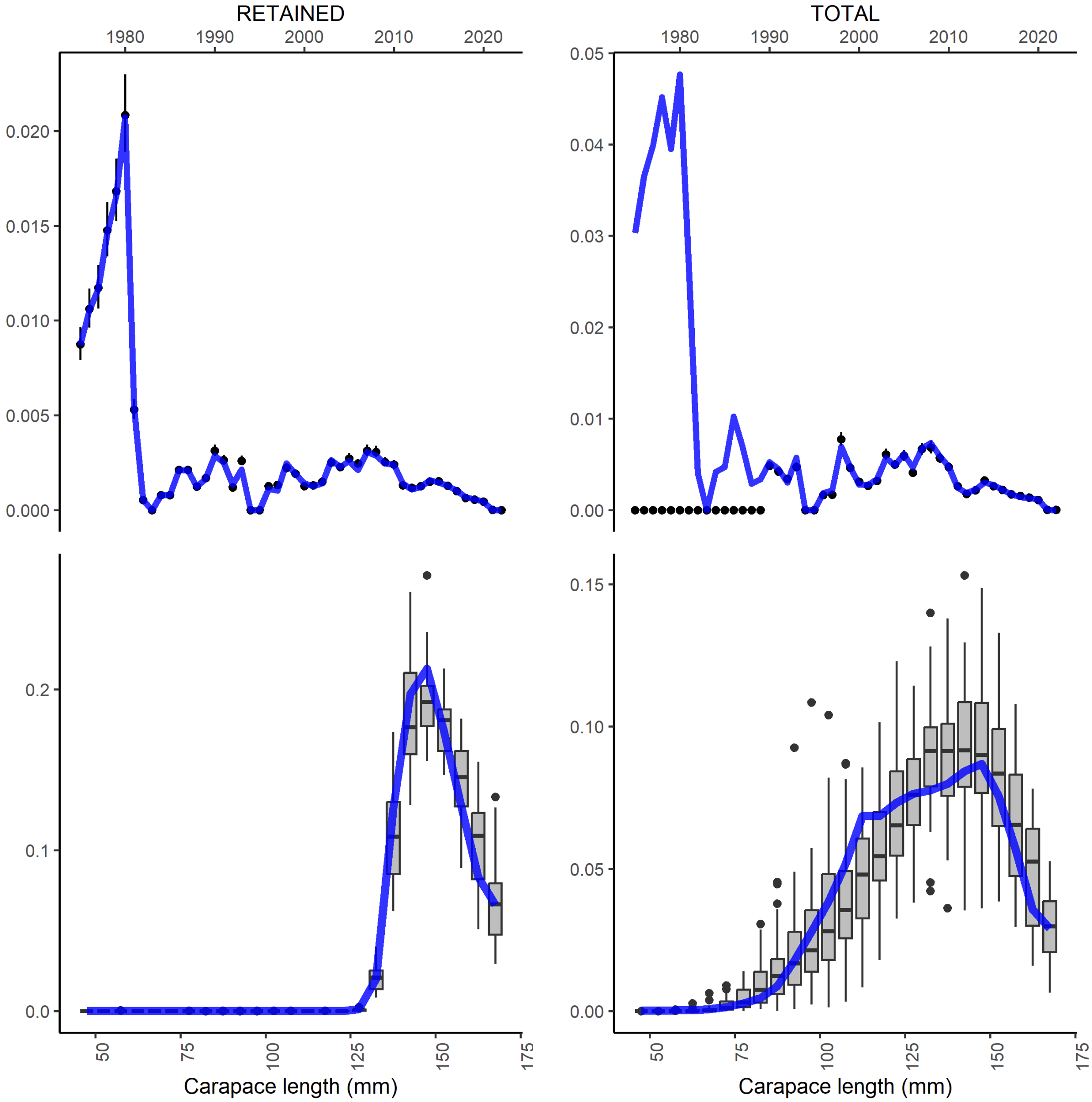

Figure S22: Fits to the retained and discarded abundances in the directed Bristol Bay red king crab fishery.

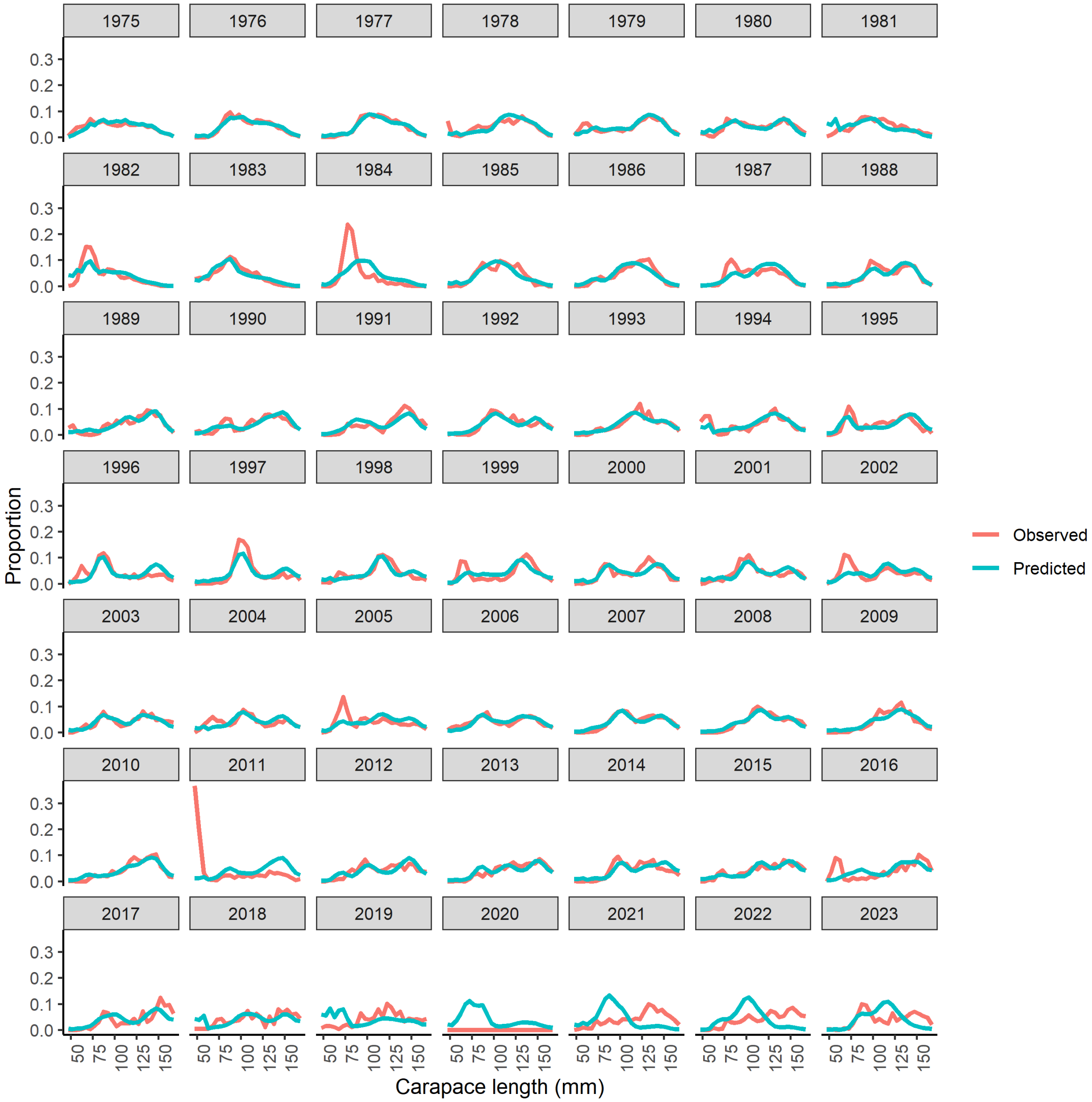

Figure S23: Model fits to individual years of size composition data for male Bristol Bay red king crab captured in the survey.

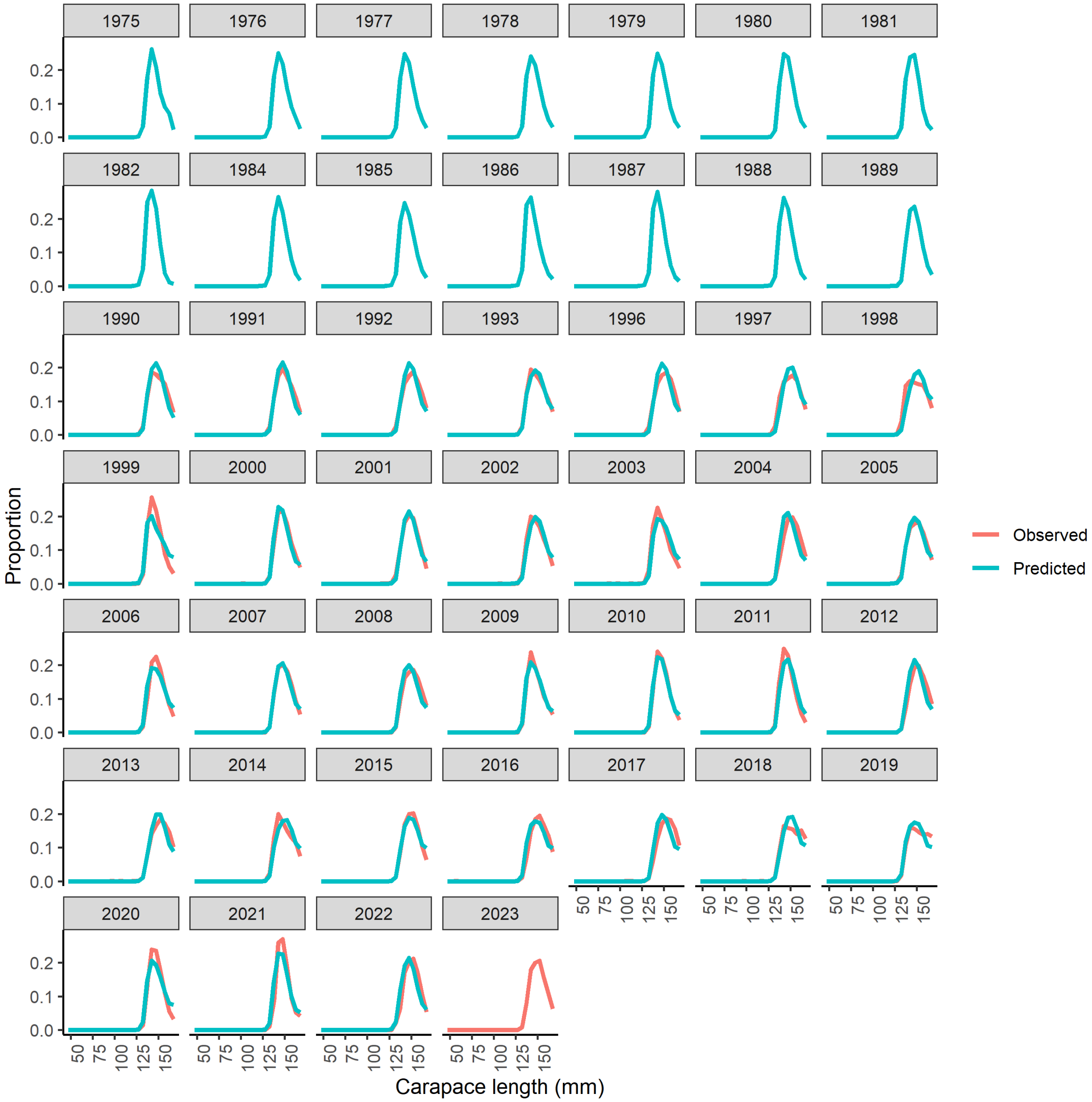

Figure S24: Model fits to individual years of size composition data for retained Bristol Bay red king crab in the directed fishery.

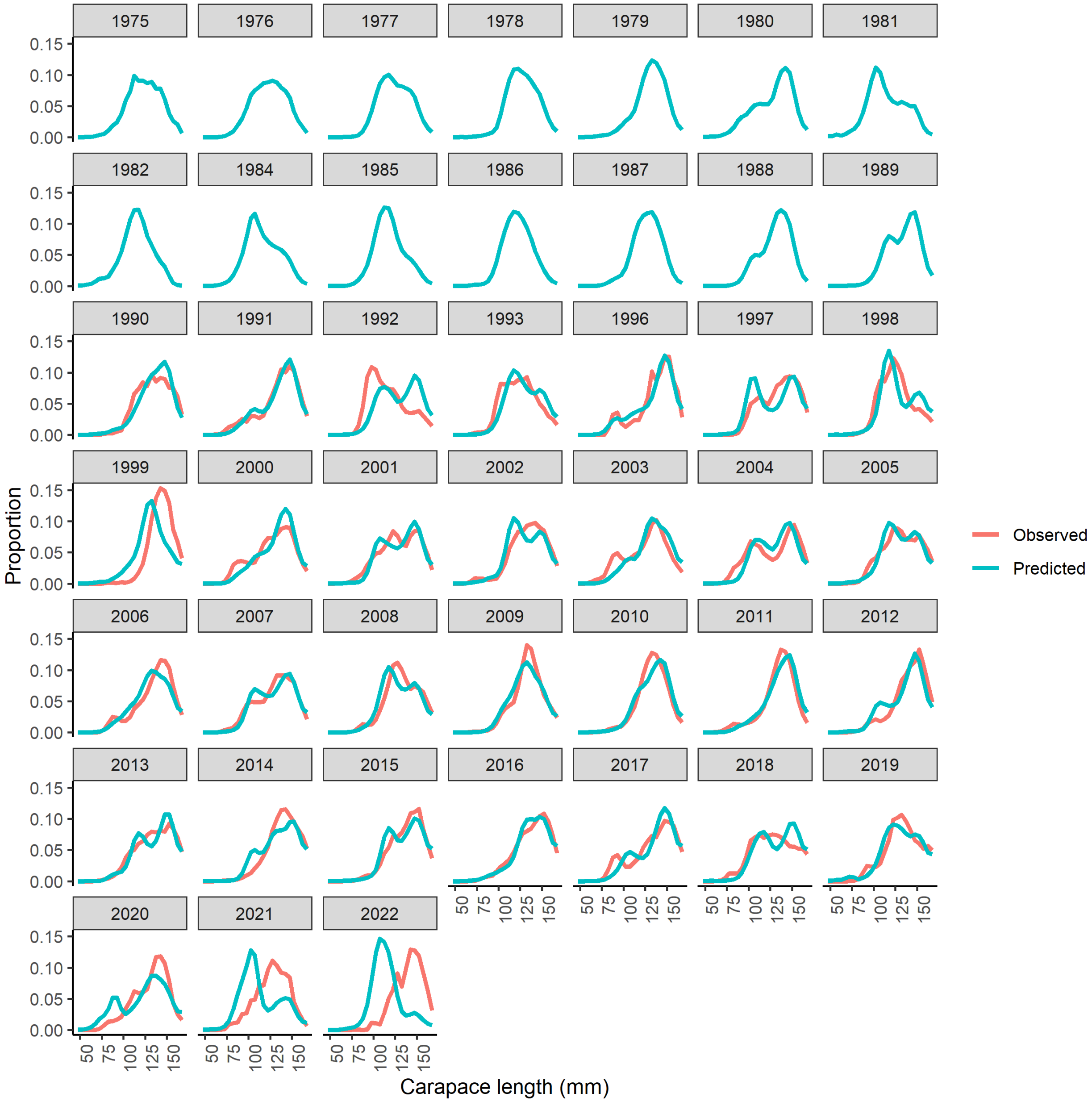

Figure S25: Model fits to individual years of size composition data for total Bristol Bay red king crab in the directed fishery.

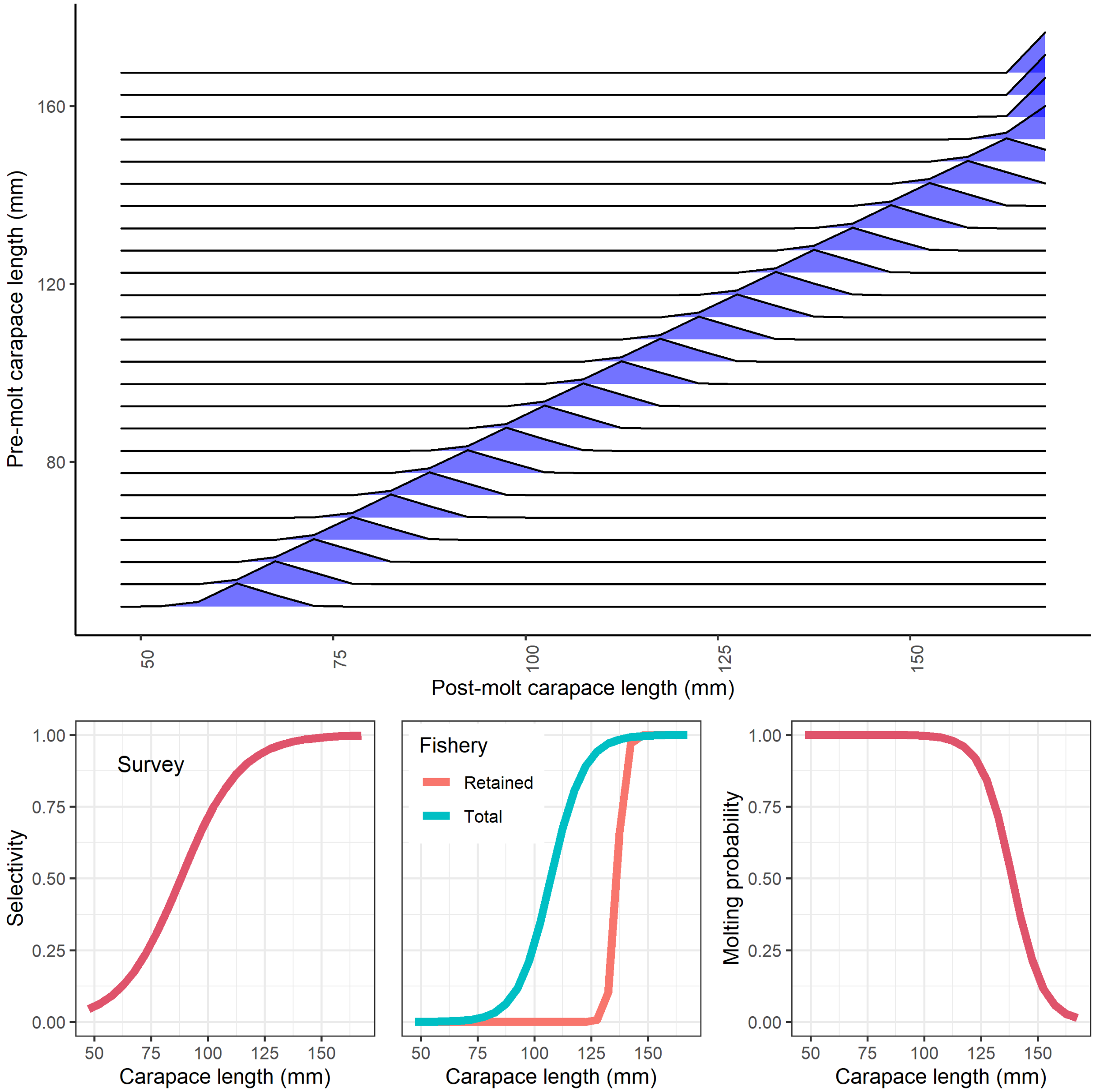

Figure S26: Estimates of population processes for Bristol Bay red king crab.

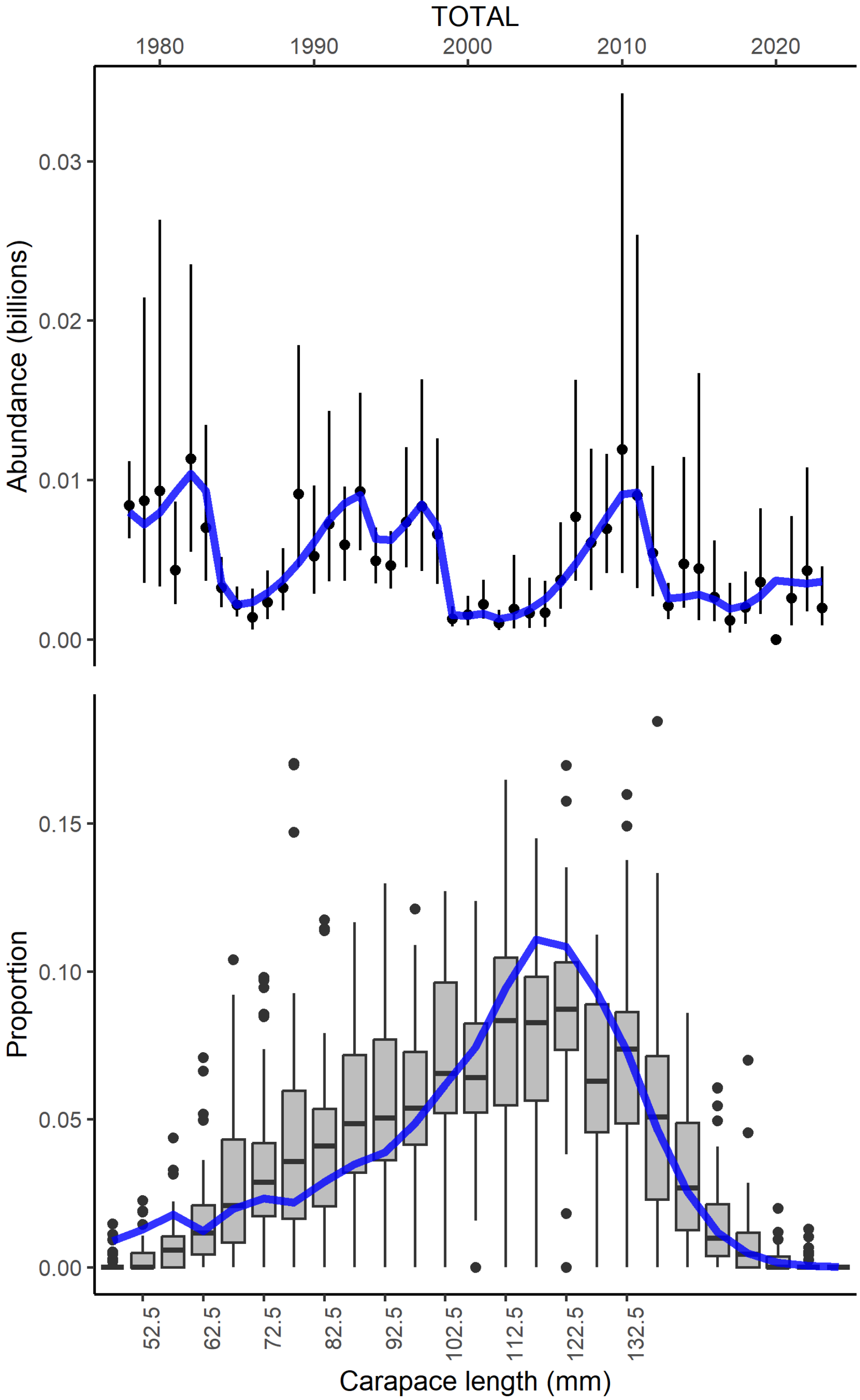

Figure S27: Fits to the survey indices of abundance and size composition data for Saint Matthews blue king crab.

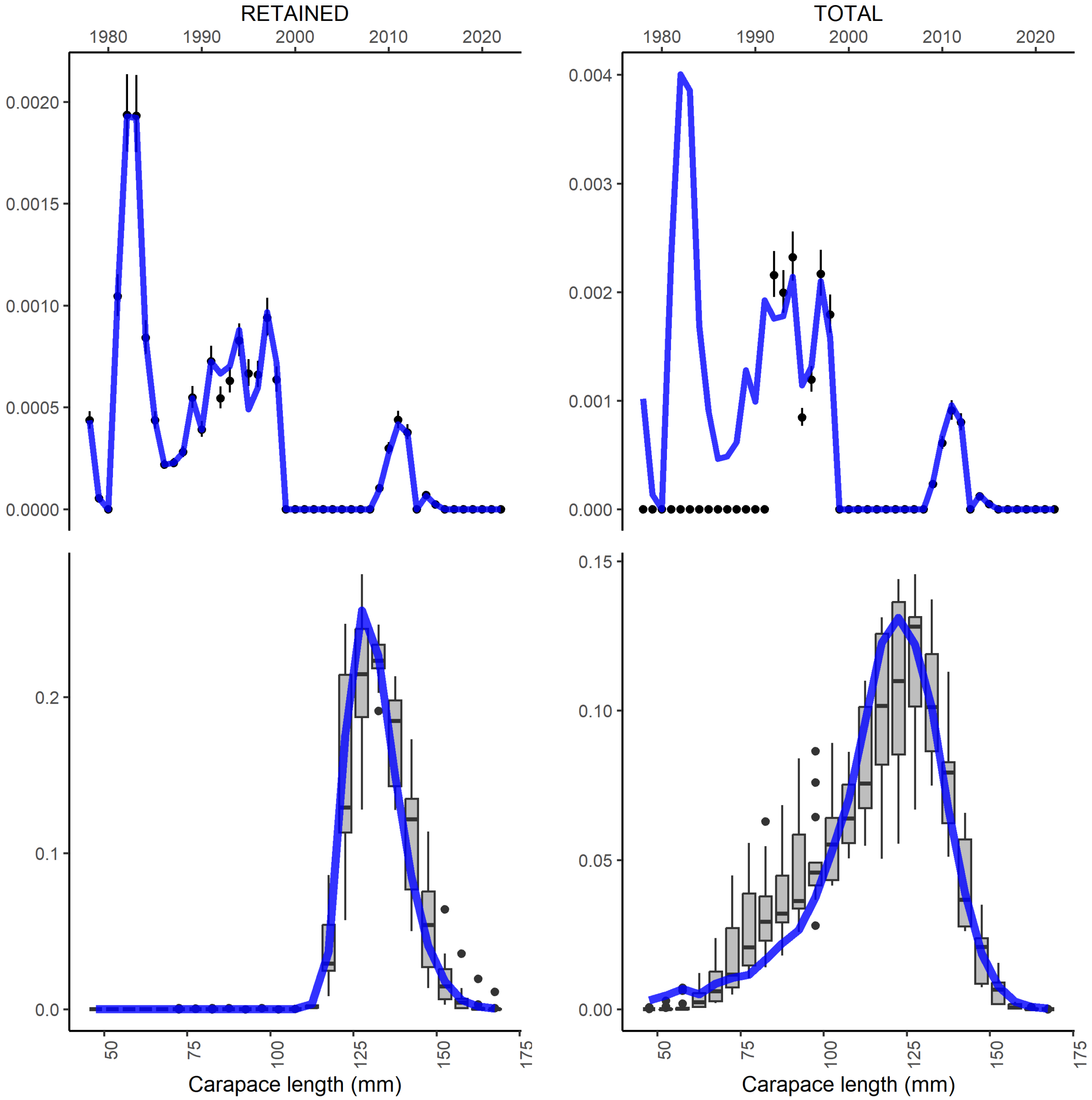

Figure S28: Fits to the retained and discarded abundances in the directed Saint Matthews blue king crab fishery.

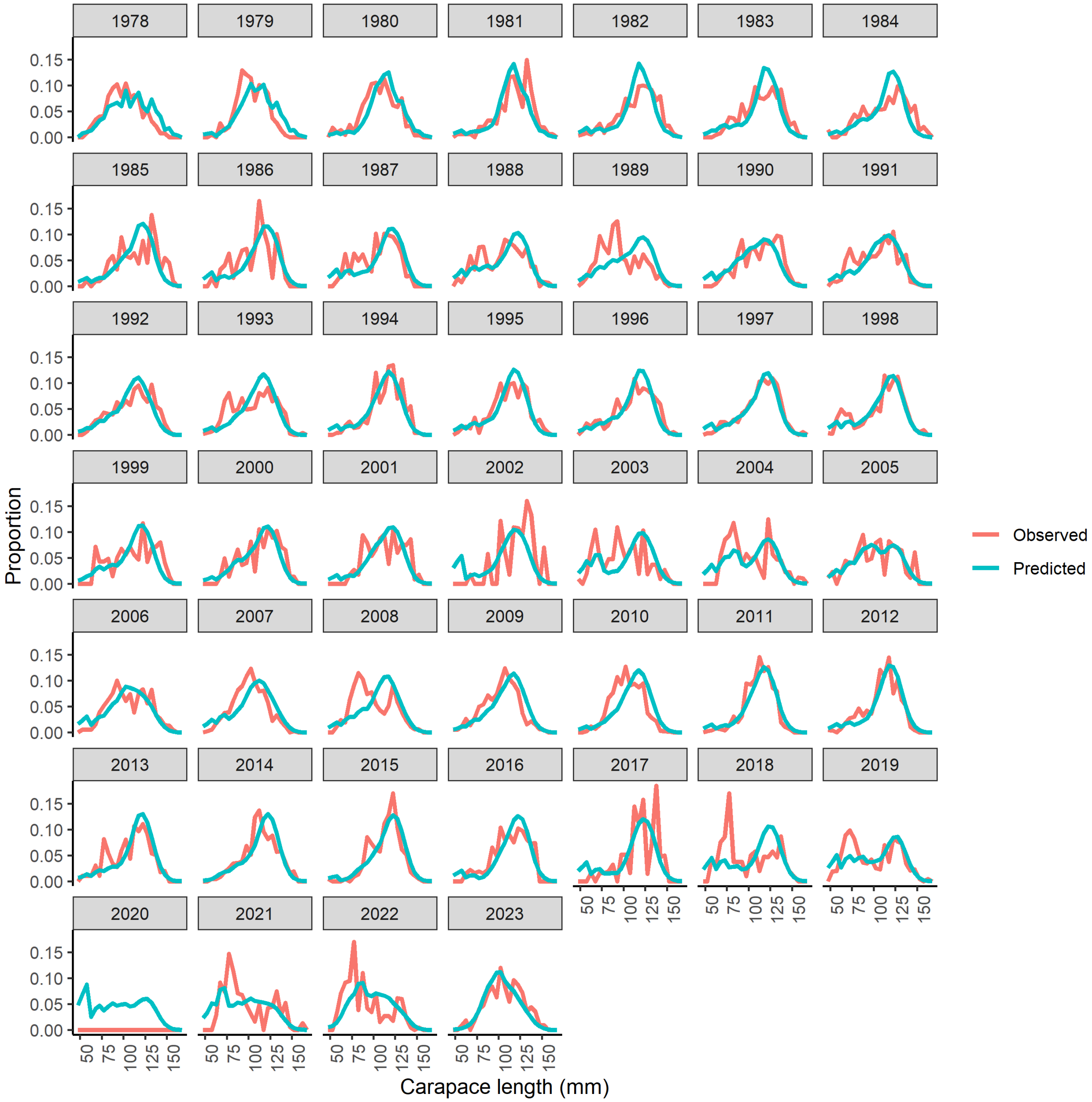

Figure S29: Model fits to individual years of size composition data for male Saint Matthews blue king crab captured in the survey.

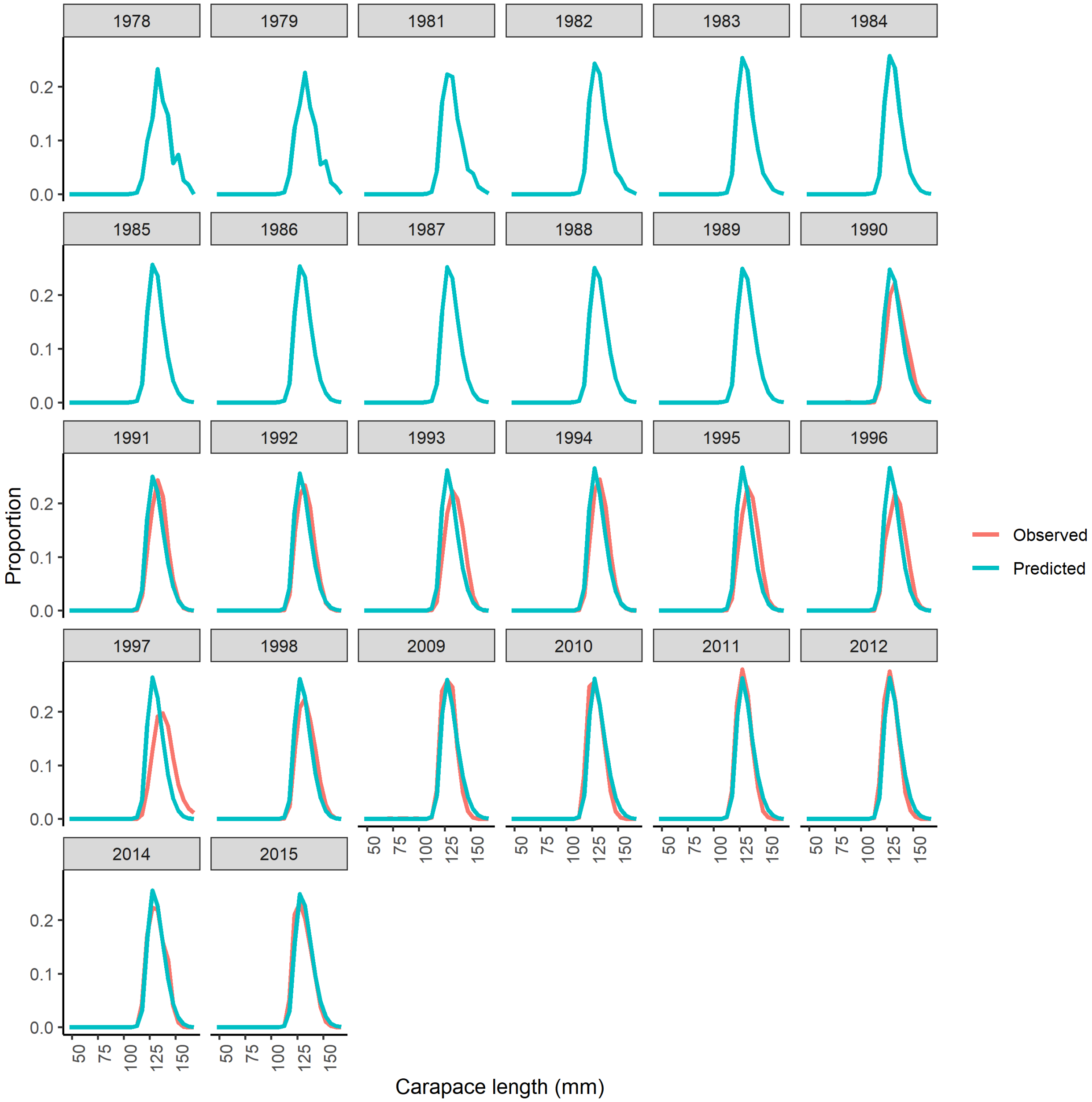

Figure S30: Model fits to individual years of size composition data for retained Saint Matthews blue king crab in the directed fishery.

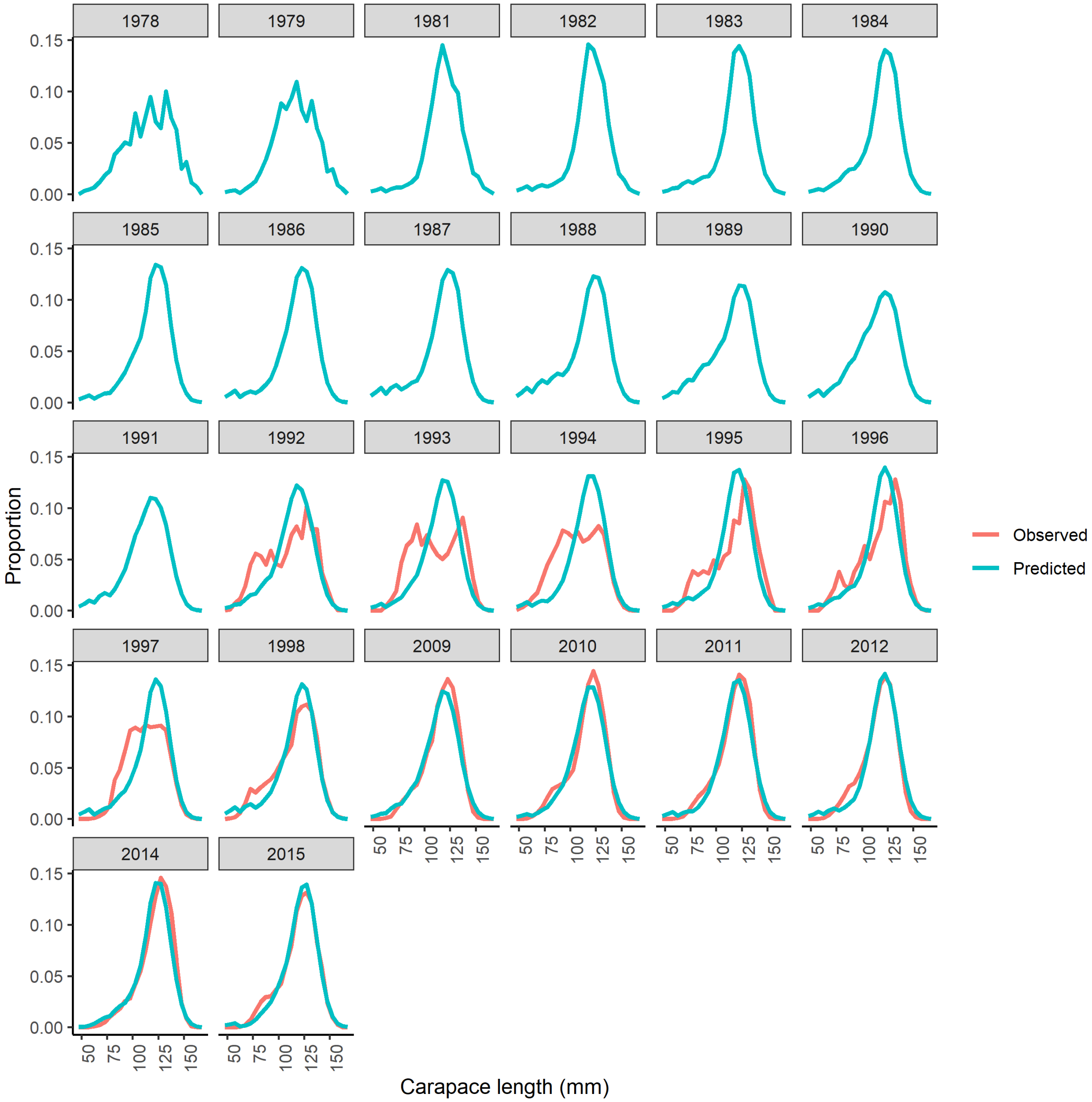

Figure S31: Model fits to individual years of size composition data for total Saint Matthews blue king crab in the directed fishery.

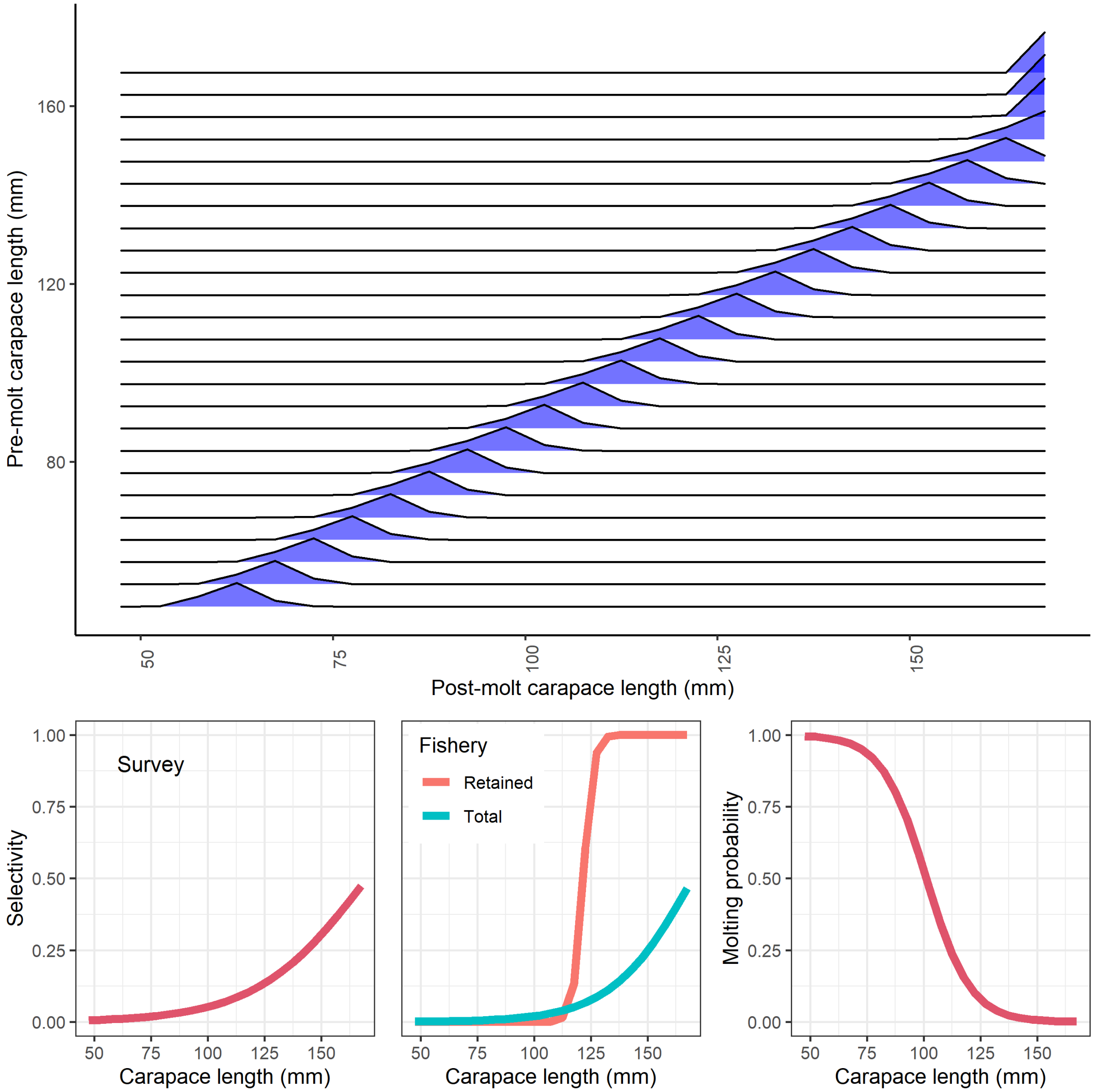

Figure S32: Estimates of population processes for Saint Matthews blue king crab.

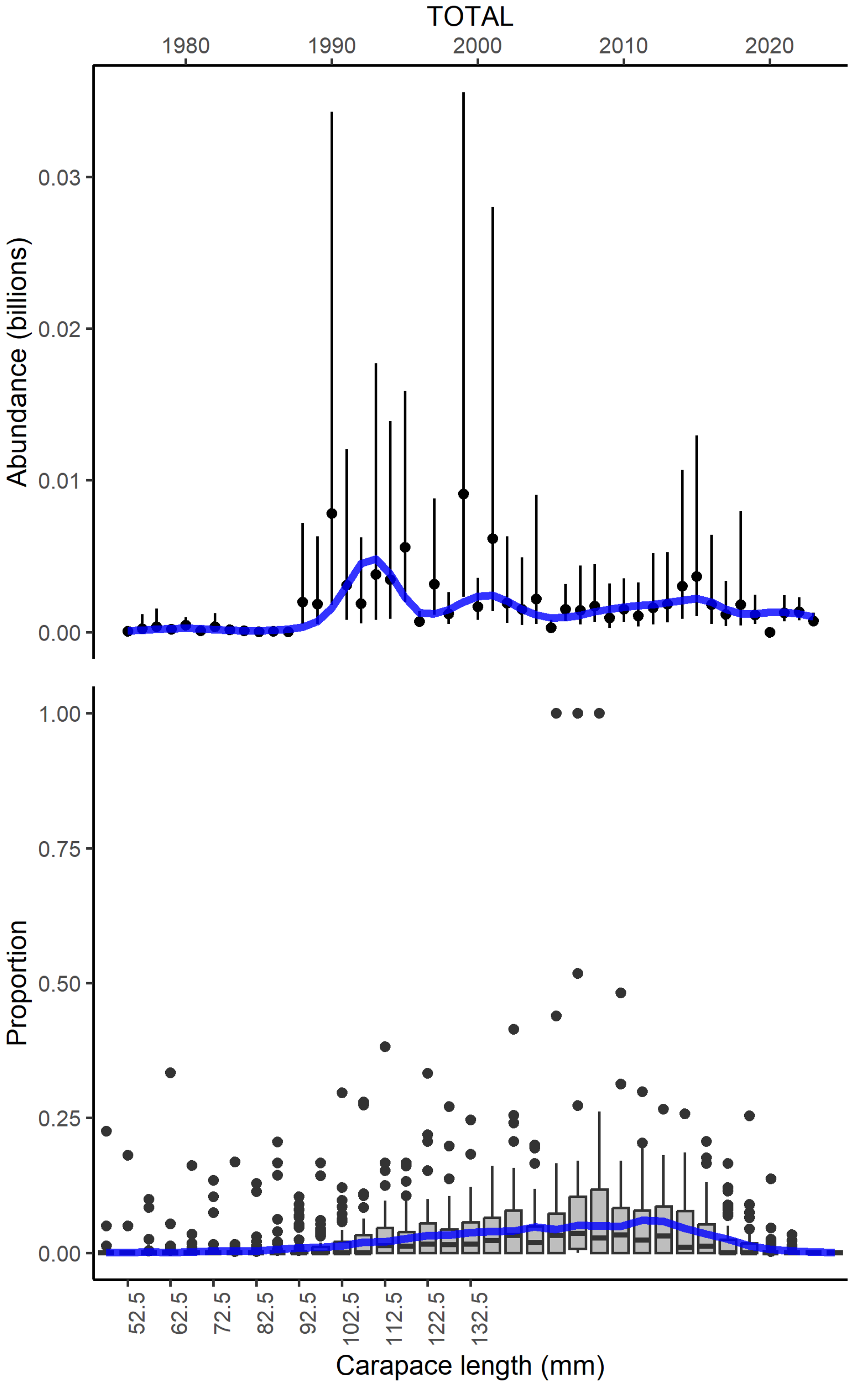

Figure S33: Fits to the survey indices of abundance and size composition data for Pribilof Islands red king crab.

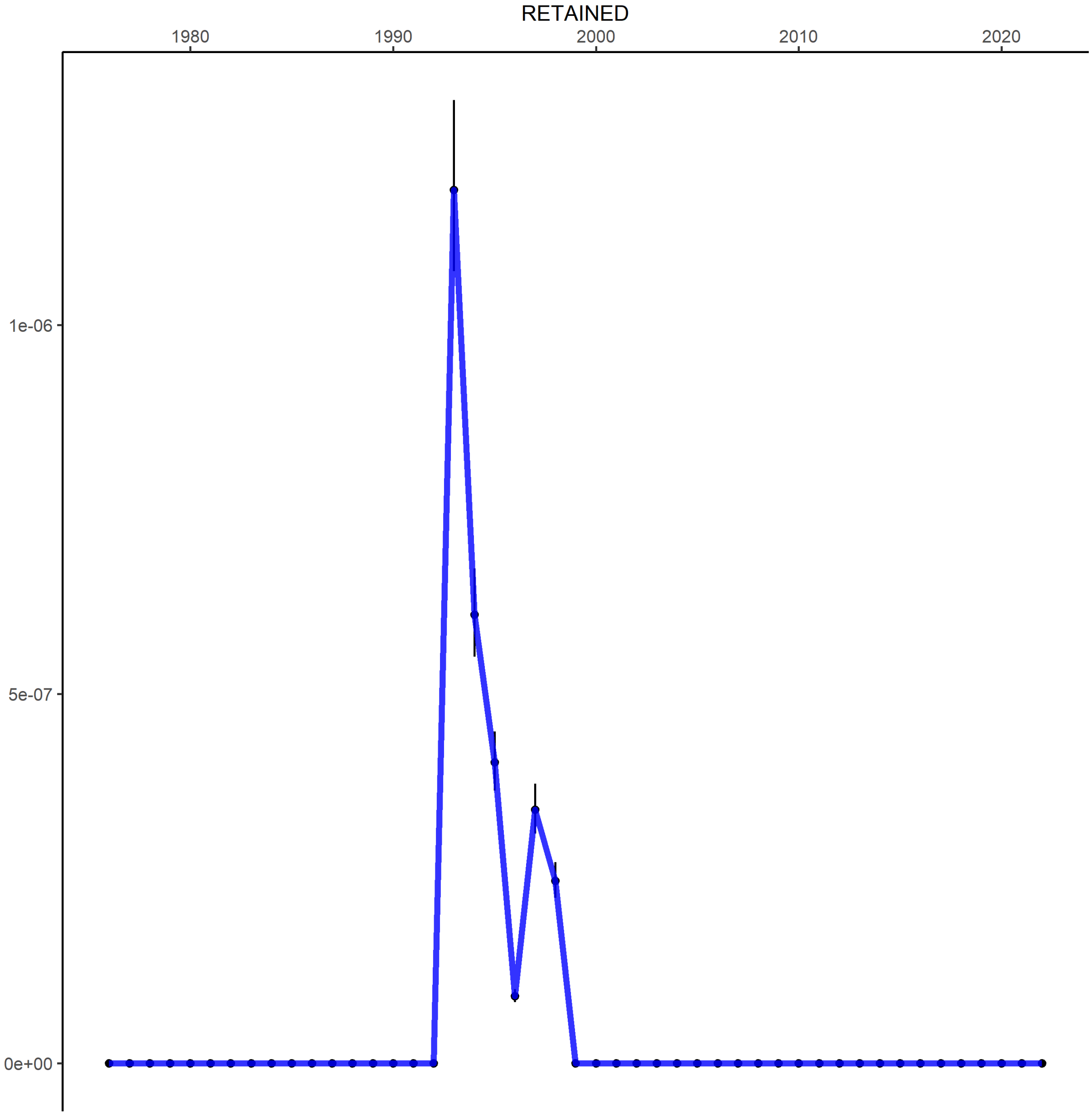

Figure S34: Fits to the retained and discarded abundances in the directed Pribilof Islands red king crab fishery.

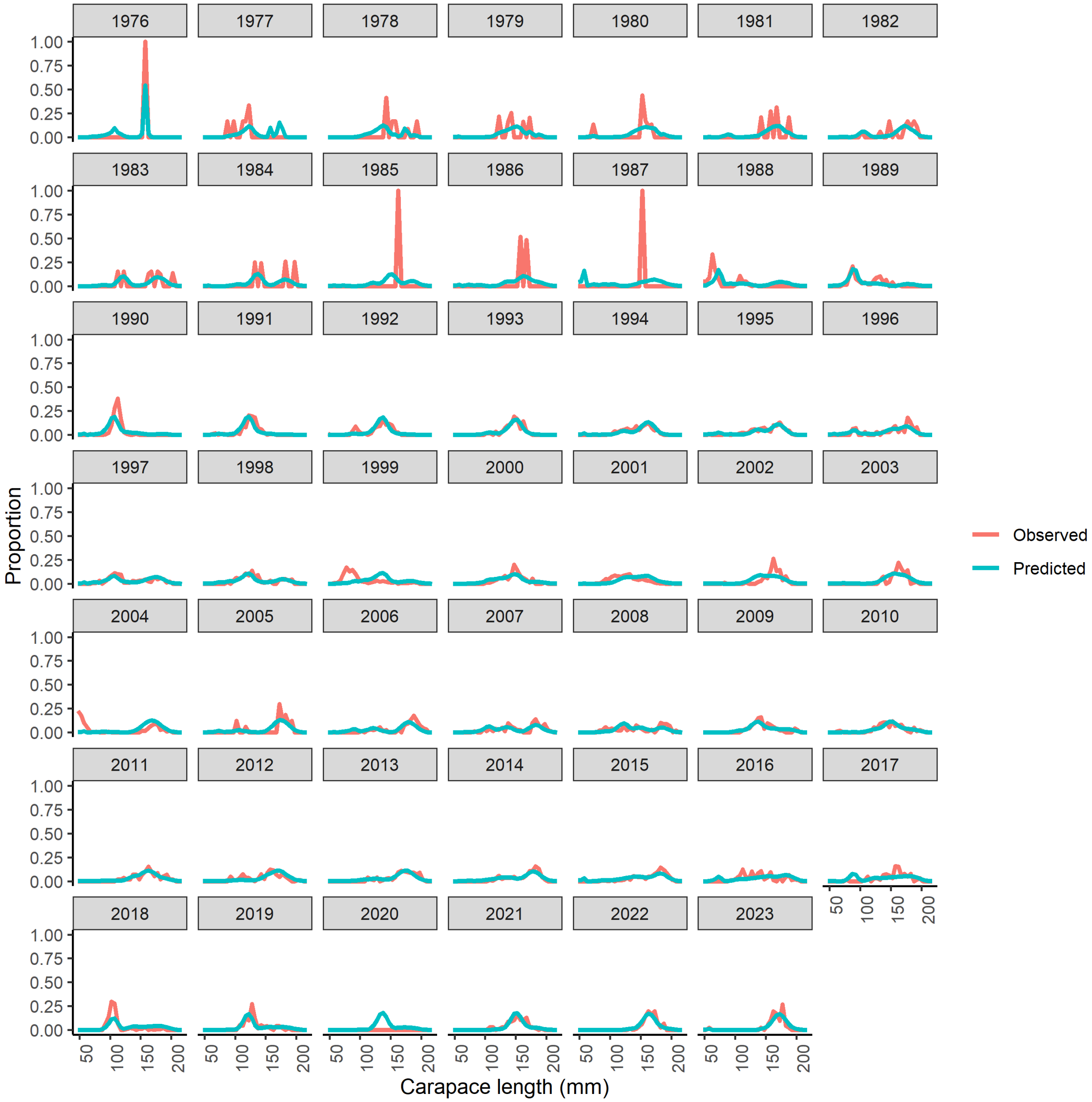

Figure S35: Model fits to individual years of size composition data for male Pribilof Islands red king crab captured in the survey.

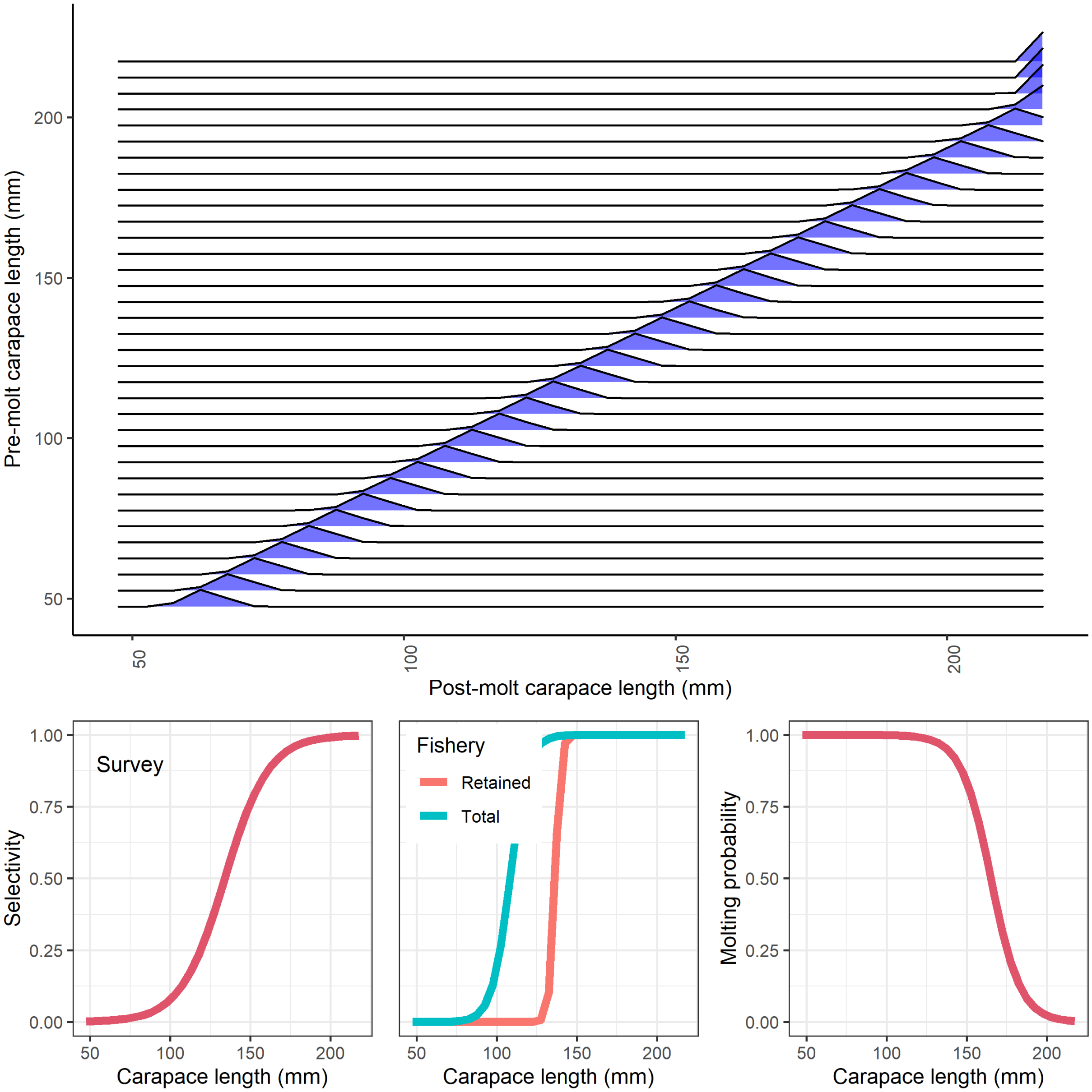

Figure S36: Estimates of population processes for Pribilof Islands red king crab.

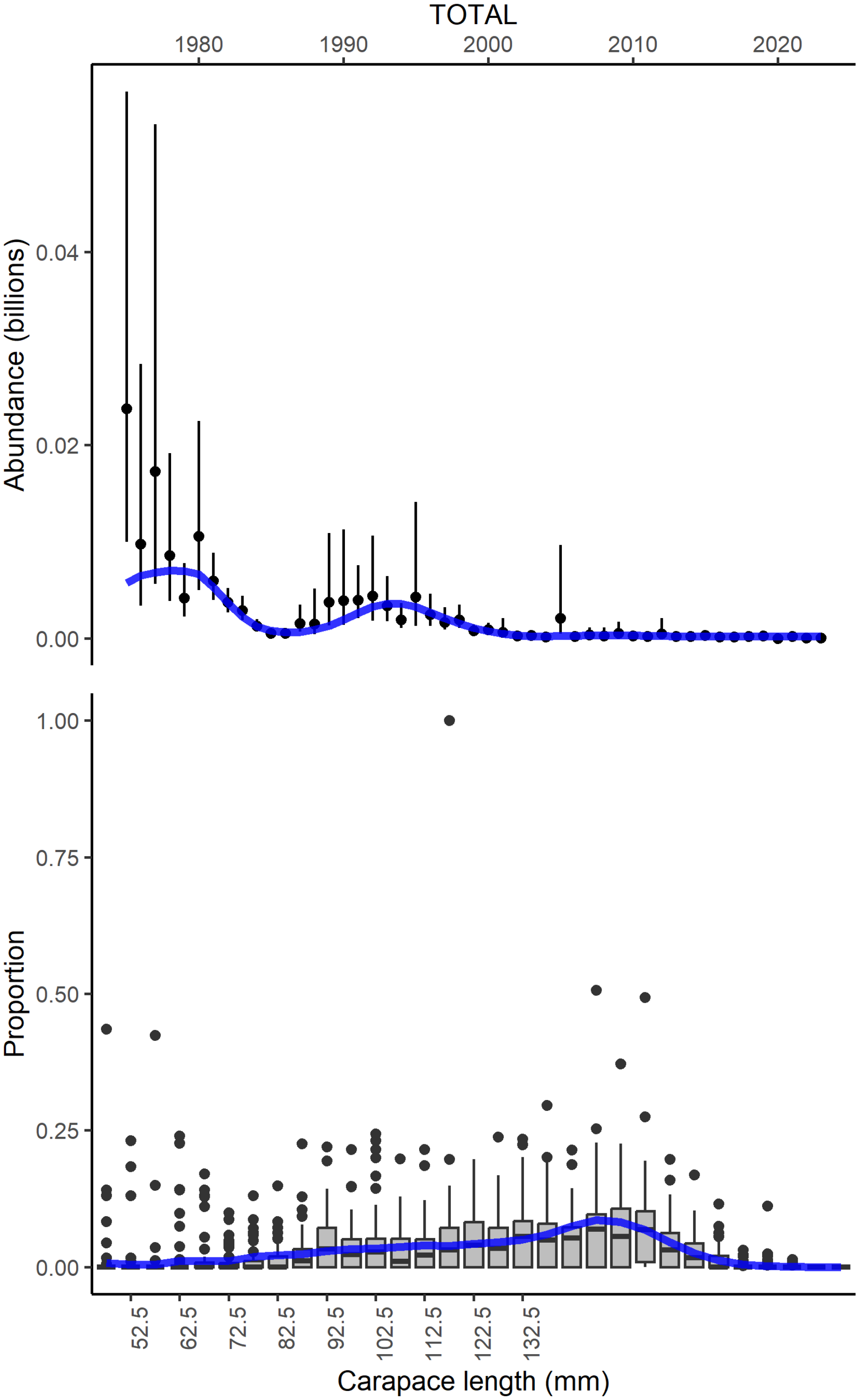

Figure S37: Fits to the survey indices of abundance and size composition data for Pribilof Islands blue king crab.

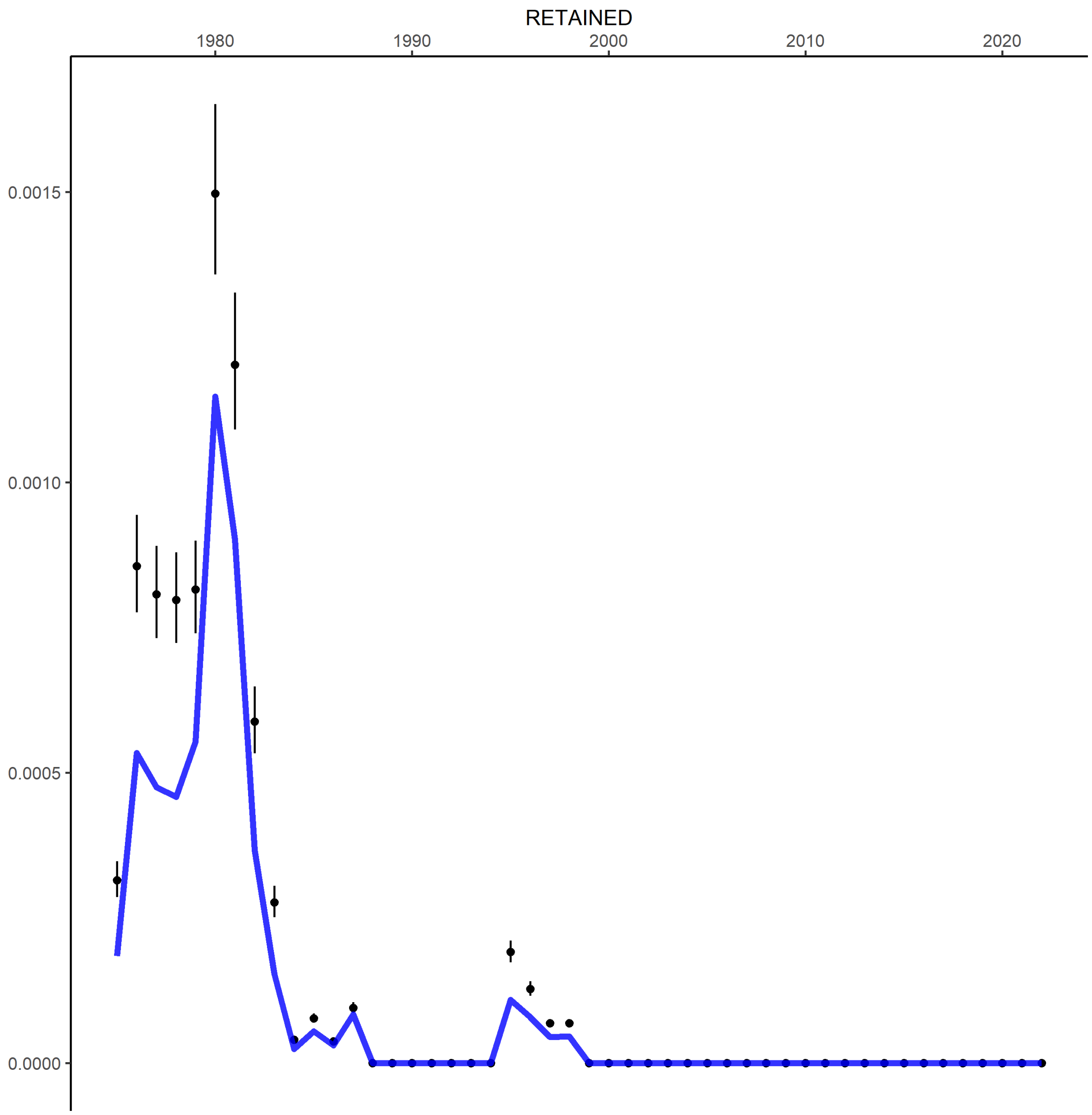

Figure S38: Fits to the retained and discarded abundances in the directed Pribilof Islands blue king crab fishery.

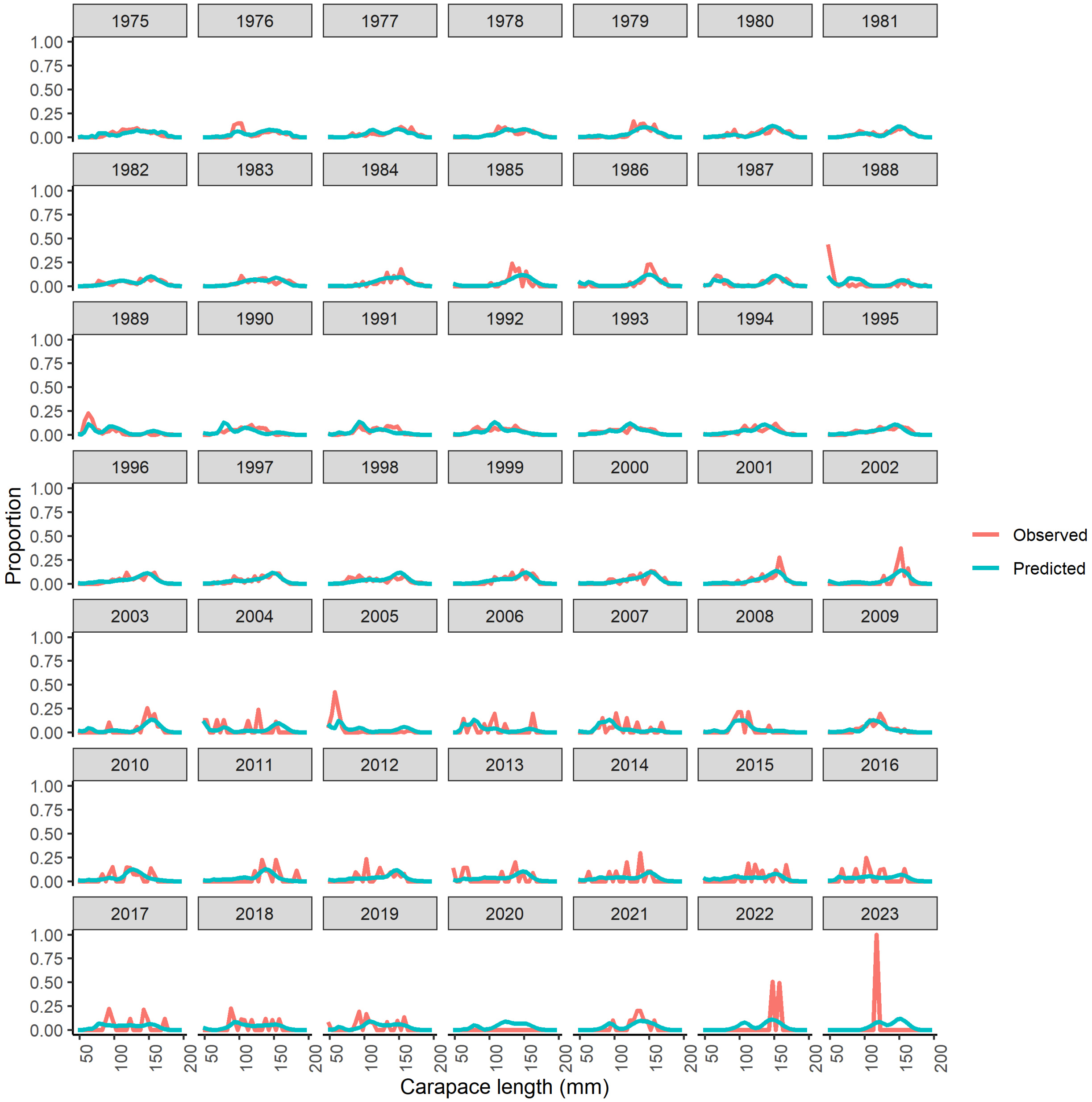

Figure S39: Model fits to individual years of size composition data for male Pribilof Islands blue king crab captured in the survey.

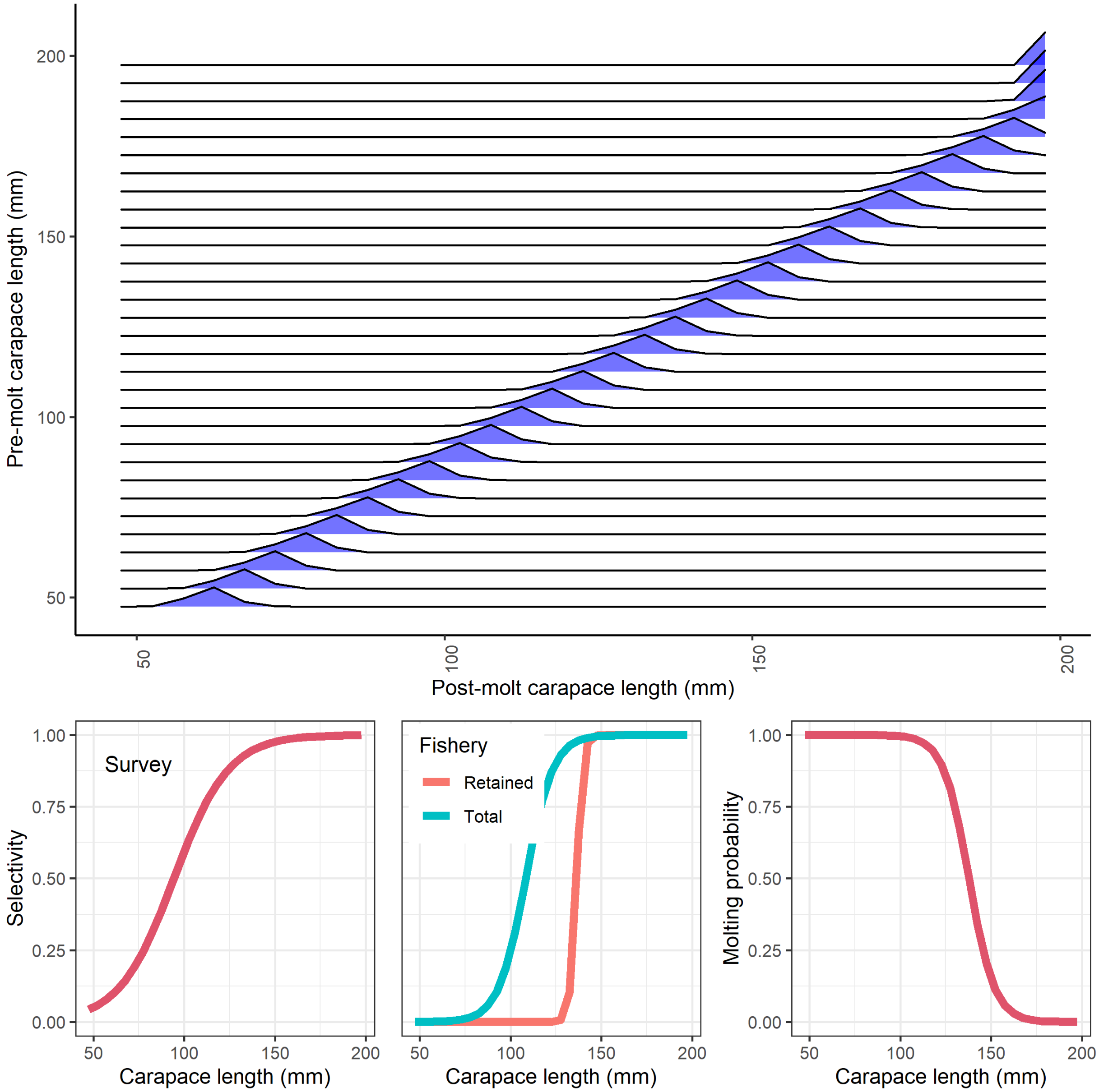

Figure S40: Estimates of population processes for Pribilof Islands blue king crab.

E. Spatial depletion in the eastern Bering Sea snow crab fishery

This manuscript is presented in draft form to facilitate discussion at plan team. It is not fully referenced yet and should not be cited anywhere until it has undergone peer-review.

Estimating local depletion rates in the eastern Bering Sea snow crab fishery

Cody Szuwalski[1,2] and others to come

[1]National Oceanic and Atmospheric Administration, Alaska Fisheries Science Center, Seattle, WA
[2]University of Washington, School of Aquatic and Fishery Sciences, Seattle, WA

Abstract

The eastern Bering Sea snow crab population supports an iconic and economically important fishery in Alaska. Management shifted from a derby-style fishery to an individual transferrable quota system in the mid-2000s. The stock is widely and heterogeneously distributed over the Bering Sea shelf, but quotas are not apportioned spatially. This results in the possibility that portions of the population have been harvested more intensely than the maximum permissible exploitation rates. To quantify this possibility, we calculated spatial depletion rates based on in-season catch-per-unit-effort data from 1990-present from the snow crab fishery. We found that depletion rates ranged from <1-99% and 13% of the areas/year depletion estimates were higher than the maximum permissible exploitation rates. Depletion rates declined sharply after the first overfished declaration for snow crab in 1999 and have gradually increased since the introduction of quotas. Estimates of median depletion rates were significantly correlated with the estimated exploitation rates from the assessment and higher depletions were weakly associated with lower CPUEs in the following year. Although our estimates of depletion were sometimes quite high and large male abundance has strongly declined over time, locally high depletion rates have not appeared to impact the reproductive dynamics of the stock. Consequently, it is difficult to recommend altering management to a spatial strategy to counter the decline of the fished portion of the stock without further information.

Introduction

The snow crab population in the eastern Bering Sea has historically supported a lucrative fishery. Catches exceeded 300 million pounds in 1990 and over 11 million pots have been dropped from 1990 to present over a wide portion of the Bering Sea shelf (figure 1). The snow crab fishery was declared overfished when the total biomass declined beneath a minimum stock size threshold in 1999 and more stringent management measures were implemented, restricting catches. The stock slowly increased and was declared rebuilt in 2011 (REF). Shortly after rebuilding the biomass declined again and a marine heatwave during 2018-2019 resulted in a collapse to biomass levels previously unseen (Szuwalski et al., 2023). The population collapse resulted in the first closure of the fishery in 2022-2023 (Szuwalski, 2024).

Snow crab are a stenothermic species (Dionne et al., 2003) associated with the cold pool, a mass of water less than 2 degree C in the Bering Sea. They are sexually dimorphic, with males being larger than females (REF). Snow crab undergo an ontogenetic migration during which they migrate from the settlement grounds on the NE portion of the eastern Bering Sea shelf towards deeper, warmer waters on the SW portion of the shelf (Ernst et al., 2005). Snow crab undergo a 'terminal molt' to maturity, after which they are morphometrically mature and do not molt again (Tamone et al., 2005). Only males are retained in the fishery and the size at which they undergo terminal molt determines whether or not they are large enough to be commercially harvested (>101 mm carapace width). Snow crab larvae undergo a pelagic period during which they are transported by ocean currents from wherever they are released by female crab to settle to the seafloor 4-5 months later (Lovrich et al., 1995). Some areas of release may therefore be more important for sustaining the population given the prevailing currents in the Bering Sea. Parada et al. (2010) suggested that larvae in the middle domain (an area between the 50 and 100 m isobaths) are more likely to be retained on the eastern Bering Sea shelf. Consequently, this might suggest that spawning occurring in the middle domain between large males and mature females plays an outsize role in reproductive dynamics.

Snow crab are managed jointly by the state of Alaska and the United States federal government via the North Pacific Fisheries Management Council (REF). The federal system establishes the overfishing level (OFL), the amount of catch beyond which overfishing would be deemed to occur and the acceptable biological catch (ABC) which is the OFL decremented by some amount to account for scientific uncertainty. The State sets the total allowable catch (TAC) which must be no higher than the ABC. The State of Alaska's harvest control rule includes a maximum permissible exploitation rate on large males (>101 mm carapace width) of 58% (REF). An annual stock assessment is performed at the federal level that incorporates a wide range of relevant data, including: survey indices of abundance and size composition data, retained and discarded catch and size composition from the directed fishery, discarded catch and size composition from non-directed fishery, laboratory data on growth, and survey efficiency data based on paired trawling (Szuwalski, 2024). The assessment output has not be used to set the TAC for the past decade, in part because of discomfort with the differences in scale between the survey data and the assessment output.

Given the spatial heterogeneity of the fishery over time and wide distribution of the population of snow crab in the eastern Bering Sea, it is possible that local fishing pressure can exceed the maximum permissible exploitation rates intended by the management strategies used for snow crab. This, coupled with the possibility that certain areas of the Bering Sea shelf may be more important in reproductive dynamics than others, creates the possibility that spatial heterogeneity in harvest could adversely influence the dynamics of the snow crab population. Furthermore, given questions about the validity of the stock assessment output, evaluating the impact of the fishery on the population with a separate dataset could be useful to corroborate trends estimated in the assessment. Given these observations, we calculated spatial depletion rates for the eastern

Bering Sea snow crab fishery from 1990-2021 using within season fisheries catch-per-unit-effort (CPUE) data to evaluate the potential impacts of the fishery on the population.

Methods

Data

We obtained data on the locations of effort and catch of the fishery from fish tickets reported to the Alaska Department of Fish and Game (ADF&G; REF). Fish tickets record effort in number of pots dropped and catch in weight for each vessel by fish ticket area every week. The spatial resolution of the fish ticket data are approximately 0.5 decimal degrees latitude by 1 decimal degree longitude (Fig. 1). Catch per unit effort by area, year, and week ($\text{CPUE}_{a,y,w}$) was calculated by summing the catch from all vessels in an area for a given year and week, then dividing it by the total effort (pots dropped) in that time period and area. In total, 815 area/week/year pairs spanning the years 1990-2021 were included in this analysis.

Depletion rate estimation

We fit generalized linear models regressions with a Gaussian family and log link to the CPUE data in areas (a) for which at least 4 data points were available during each year (y) using week (w) as a predictor and the number of pots deployed as weights (see figure 2 for an example year).

$$CPUE_{a,y,w} = \alpha_{a,y} + \beta_{a,y}$$

For area/year combinations that had 2 or 3 data points, the models were fit without a log link to allow convergence. The regressions provided a trend line for CPUE over week by area and year (figure 3). A depletion rate ($d_{a,y}$) for each area (a) during each year (y) was calculated as 1 minus the model-predicted CPUE from the last week (t=n) that area was fished divided by the model-predicted CPUE from the first week (t=1) that area was fished:

$$d_{a,y} = 1 - \frac{\widehat{CPUE}_{a,y,t=n}}{\widehat{CPUE}_{a,y,t=1}}$$

A spatially aggregated yearly depletion rate was calculated based on a catch-weighted mean to better capture the impacts of fishery removals. This 'depletion rate' is based only on CPUE; if CPUE were directly proportional to the abundance in an area, the depletion rate would be equivalent to an exploitation rate if no natural mortality took place during the fishing season and no migration into or out of the area occurred. There are other fishery-related factors that could also violate the assumption of proportionality (e.g. hyperstability, REF), so we use 'depletion rate' rather than 'exploitation rate' when referring to the above statistic (see discussion for more). In addition to depletion rates by area and year, we calculated the median depletion rate by spatial area over 5 year blocks to illustrate changes in the location and intensity of fishing over time. We also used linear regression to compare the depletion estimates derived from fisheries dependent CPUE observations to the estimates of exploitation rates derived from the stock assessment model for eastern Bering Sea snow crab (Szuwalski, 2024).

Measuring population impacts

Quantifying the local impacts of fishery removals on population dynamics is difficult. For snow crab, one of the major hurdles in understanding this impact is the movement of crab between the fishery-independent survey that occurs during the summer and the fishery that occurs during the winter. It is consequently difficult to use the survey data to quantify spatial impacts of the fishery because the movement impedes our ability to match local changes in the population seen in the survey with depletion occurring in the fishery. Nonetheless, attempting to understand the impacts of spatially heterogeneous fishery removals is important.

We attempt to quantify the local effects of the fishery on the snow crab population by comparing the change in CPUE in a given area between successive years to the depletion in that area. The rationale behind this comparison is that the following year of CPUE should decline compared to the previous year if depletion was high, depletion is an effective measurement of fishing impact, and CPUE is proportional to abundance. We quantify the relationship between depletion and change in CPUE by first calculating the percent change in CPUE by area by year:

$$\Delta CPUE_{a,y} = 100 * \frac{CPUE_{a,y} - CPUE_{a,y-1}}{CPUE_{a,y-1}}$$

Then we used generalized additive models (GAMs; Wood, 2011) to predict $\Delta CPUE_{a,y}$ using $d_{a,y}$ as a covariate:

$$\Delta CPUE_{a,y} = s(d_{a,y}) + \varepsilon_{a,y}$$

Given this analysis requires effort to occur in an area in adjacent years, the data set used for this portion of the analysis was only a fraction of the dataset used to calculate depletion above.

Results

The average number of weeks an area each year was fished ranged from 1 to 25, with an average of 4.5 weeks fished in an area (sd = 4.4 weeks). CPUE declined by 11.4% (median) per week fished, with a maximum observed decline in CPUE of 54% in one week (figure 4). Estimates of depletion rates ranged from 0 to 1, with a median depletion of 13.7% over all years and areas (figure 5). Median depletion estimates were highest in the early portion of the time period (1990-1999), after which an overfished declaration in 1999 resulted in drastically reduced allowable catches from 2000-2005. Allowable catches increased as the stock rebuilt during the 2010s and depletion rates increased in turn. The average depletion rate from 2010-2020 was 75% of the depletion rate estimated during the 1990s. Across all years, 13% of areas had depletion rates that were higher than the maximum permissible exploitation rate. The top ten areas in terms of catch production (which produced on average 65% of catch on a yearly basis) were fished in excess of the maximum permissible exploitation rate 16% of the time.

The location of fishing changed over time in addition to the intensity of depletion. In the 1990s, the fishery covered a larger area and had the highest estimated depletion rates (figure 6). An average of 38 areas were consistently fished during the 1990s and the highest depletion rates were estimated west and south of the Pribilof Islands. The average number of areas fished shrank to 24 when the stock was declared overfished and depletion rates were the low in all areas. A further contraction occurred to 18 areas fished after the rationalization of the fishery in 2005, and depletion rates increased fairly uniformly in space. During the 2010s, the number of areas fished trended slightly upward with a return of effort farther north (figure 6). The depletion rates estimated in the north during this period were more intense than they were in the 1990s when last fished. Furthermore, in 2020 and 2021, there was very little effort south of the Pribilof Islands, which had historically been an area of more intense and productive fishing.

The estimated exploitation rates from the 2024 stock assessment for eastern Bering Sea snow crab and the catch-weighted mean depletion rate were significantly and positively correlated (figure 7). The estimated slope between the two quantities, however, was 0.37, reflecting relatively small estimated exploitation rates from the assessment when depletion estimates were high. The years 2011 and 1990 were notable outliers during which high exploitation rates were coupled with relatively low catch-weighted mean depletion rates (and vice versa).

Depletion estimates were available for 383 year/area pairs in adjacent years (47% of the area/year combinations used in the above analyses). A significant, but weak negative relationship was found between the change in CPUE in an area and the depletion occurring in that area (figure 7). When depletion was less than ~0.21 in a given area and year, CPUE was higher in that area in the next year on average. However, when depletion was above 0.21, CPUE was lower in the following year on average. For example, CPUE was 13% and 23% lower the following year in a given area (on average) when depletion in that area was 50% and 75%, respectively. While there was a weak relationship between depletion and the following year's change in CPUE positive and negative changes in CPUE were seen at almost all values of depletion.

Discussion

We found that estimates of median depletion rates are roughly consistent with the estimates of exploitation rates from the stock assessment for eastern Bering Sea snow crab, but depletion is heterogeneous in space. Consequently, the snow crab population in some areas have experienced higher depletion rates than the maximum permissible exploitation rate at some point in time. Areas that have produced the most catch historically have been consistently subject to higher rates of depletion.

If higher depletion rates have impacted population dynamics, alternate management strategies might be used to address this. For example, it could be possible to monitor CPUE by area within a year and impose cutoffs for depletion estimates beneath which an area would be closed to fishing. Or, as a more intermediate step, we could subdivide the Bering Sea into fewer areas than the statistical areas used in this analysis, the apportion the yearly total allowable catches to different areas. This is similar to what is done for Newfoundland stocks (REF). Each of these methods come with additional overhead in terms of data collection, analysis, and enforcement and their effectiveness at affecting changes in the population trajectories depends on whether or spatial heterogeneity in depletion is driving population dynamics.

Unfortunately, evaluating whether or not local depletion rates higher than the maximum permissible exploitation rates have impacted the population dynamics of the stock is a hard question to answer. Previous attempts at understanding the spatial dynamics of eastern Bering Sea snow crab include a spatial stock assessment that incorporated the scientific surveys and it mirrored our conclusions about the spatial heterogeneity of fishing pressure over time (OImos et al., 2023). That spatial assessment adjusted for the movement between the survey and fishery shifting distribution of the catch to match the distribution of the large males at the time of the survey. Our depletion estimates suggest that, although there is a significant negative relationship between depletion and change in CPUE in a given area, the relationship is very weak. For any given amount of depletion in an area, the change in CPUE can be anywhere from doubling to halving. The difficulty in predicting the change in CPUE based on depletion is likely related to movement of large snow crab and a lack of site fidelity through the migrations that take place between the summer and winter.

Although it is difficult to compare the spatial survey data to the spatial depletion estimates because of this movement, we can examine the aggregated survey data to try to explore the impacts of the fishery on snow crab population dynamics. The survey abundance of commercially sized males (>101 mm carapace width) has strongly trended downwards over the study period (figure 8). In chronological order, the abundance shifted from the highest average in the study period (1990-1998) to the second lowest (1999-2005), then a small rebound occurred with rationalization (2005-2014), after which the abundance of large males entered a period during which the average is the lowest observed (2015-present). The current average abundance is just 16% of the average of the early period. These trends could suggest that fishery effects have had a large impact on the population upon which the fishery depends, and is corroborated by fairly high depletion rates and exploitation rates estimated in the assessment.

However, females (which are not retained by the fishery and are generally small enough to escape the pots of their own volition) have a much more stable trend (figure 8). Females have experienced periods of low abundance (e.g. 1995-2009 and the collapse during the 2018-2019 marine heatwave), but the females have shifted from high to low regimes with a stable trend, rather than declining over time like the large males. The difference between the high period and the low periods are also much less extreme for the females. The lack of a consistent downward trend in females could be interpreted as the fishery having small effects on the fundamental reproductive potential of the population (i.e. females). That is, even if there are localized hotspots of depletion, the remaining males are mobile and plentiful enough to inseminate the females in the population.

There are a number of caveats to working with fishery dependent CPUE data to understand population dynamics in space and time. Movement and/or mortality during the season among areas would be confounded with ‘depletion’—that is, a decline in CPUE in one area could be a result of fishery removals or because there are fewer crab available to be caught in an area because they moved or died. CPUE could also be ‘hyperstable’ as a result of trap saturation or strong schooling behavior (REF). All of these caveats are reasons that stock assessments are performed with scientific surveys at the population level. Our intention was not to provide analyses to replace the measure of fishery impact provided by the stock assessment, but to supplement the information contained in the assessment. Our analysis provides a direct, measurable impact in a given year of the snow crab fishery in space and provides corroboration for the trends in estimated fishing mortality from the assessment using a data set not included in the assessment (i.e. the effort data). This is useful because it provides a link from what the crabbers are seeing directly on the fishing grounds to the output of the assessment.

The lack of conclusive evidence suggesting local depletion rates have negatively impacted the snow crab population and conflicting signals in the population makes it difficult to recommend changes to the current spatially-aggregated management strategy. The continued downward trajectory of the large males upon which the fishery depends is a concerning fact. The more level trajectory of females suggests the fishery may not dramatically impact the reproductively active portion of the population, but that something is preventing small males progressing through the size classes to become the large males caught in the fishery. Size-based mortality and time-varying probability of terminally molting (either as a result of density dependent processes or environmental drivers) are potential mechanisms that could influence the progression of small males to large males. Even if research were able show that, for example, the density of large males is strongly related to the probability of terminally molting and, if the density of large males is depressed to low levels through fishing, many fewer crab will terminally molt at sizes above the commercially preferred size of >101 mm carapace width.

Acknowledgements (to come)

References (to come)

1. C.S. Szuwalski. Stock assessment of Eastern Bering Sea snow crab. Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions. 2021 Crab SAFE. North Pacific Fishery Management Council, 1007 West 3rd Ave., Suite 400, L92 Building, 4th floor. Anchorage, AK 99501. (2024).

2. S.L. Tamone, M. Adams, J.M. Dutton. Effect of eyestalk ablation on circulating ecdysteroids in hemolymph of snow crab *Chionoecetes opilio*: physiological evidence for a terminal molt. *Integr. Comp. Biol.* **45(120)**, 166-171. (2005).

3. LovrichG. A. Sainte-Marie B. Smith B. D. 1995 Depth distribution and seasonal movements of *Chionoecetes opilio* (Brachyura: Majidae) in Baie Sainte-Marguerite, Gulf of Saint Lawrence *Canadian Journal of Zoology* 73 : 1712–1726

4. C. Parada, D.A. Armstrong, B. Ernst, S. Hinckley, J.M. Orensanz. Spatial dynamics of snow crab (*Chionoecetes opilio*) in the eastern Bering Sea–Putting together the pieces of the puzzle. *Bulletin of Marine Science*. **86(2),** 413-437. (2010).

5. B. Ernst, J.M. (Lobo) Orensanz, D.A. Armstrong. Spatial dynamics of female snow crab (*Chionoecetes opilio*) in the eastern Bering Sea. *Can. J. Fish. Aquat. Sci*. **62,** 250-268. (2005).

6. C.S. Szuwalski, K. Aydin, E.J. Fedewa, B. Garber-Yonts, M.A. Litzow. The collapse of eastern Bering Sea snow crab. Science. 382: 306-310. (2023)

7. Dionne, M., Sainte-Marie, B., Bourget, E., Gilbert, D. 2003. Distribution and habitat selection of early benthic stages of snow crab (*Chionoecetes opilio*). Mar. Eco. Pro. Ser. 259: 117-128.

8. M. Olmos, J. Cao, J.T. Thorson, A.E. Punt, C.C. Monnahan, B. Alglave, C.S. Szuwalski. A step towards the integration of spatial dynamics in population dynamics models: eastern Bering Sea snow crab as a case study. Ecological Modelling. 485: 110484.

9. Darrell R J Mullowney, Krista D Baker, Size-at-maturity shift in a male-only fishery: factors affecting molt-type outcomes in Newfoundland and Labrador snow crab (*Chionoecetes opilio*), *ICES Journal of Marine Science*, Volume 78, Issue 2, March 2021, Pages 516–533

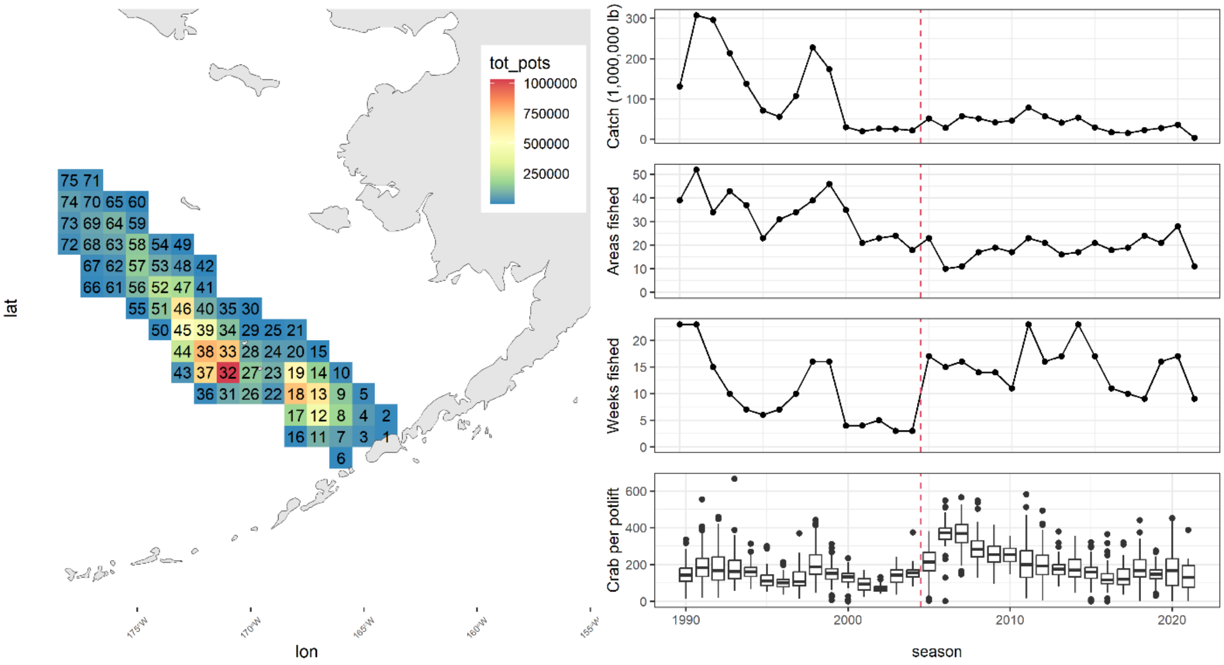

Figure 1. ADFG statistical areas numbered with heatmap of total pots dropped from 1990-2021 (left). Aggregate statistics over time (right) including (from top down): total catch in millions of pounds, number of areas fished in a given season, the number of weeks fished in a given season, survey estimates of male crab >101 mm carapace width, and boxplots of the catch per unit effort in crabs per potlift.

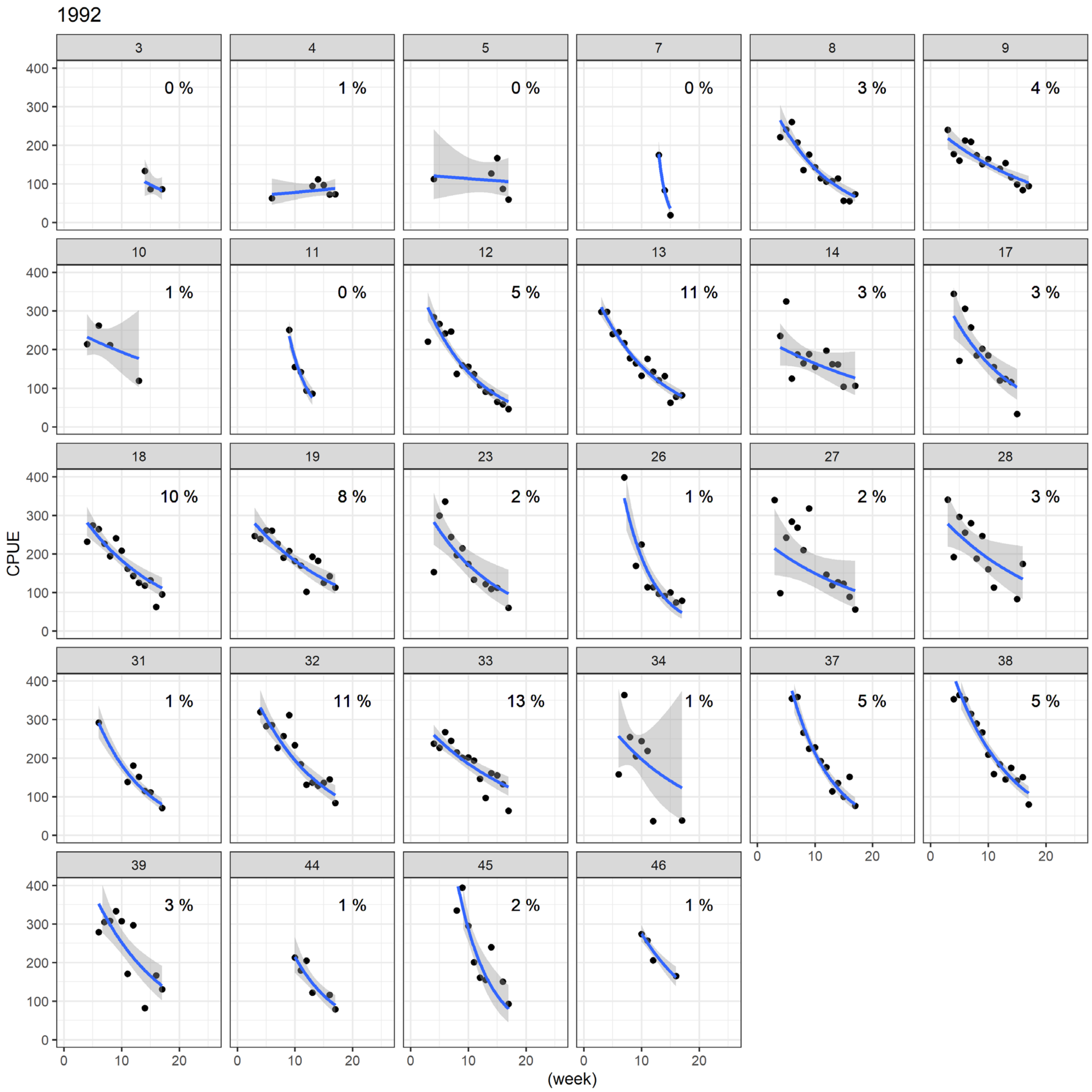

Figure 2. Area specific regressions for an example year (1992). Each panel represents an area (see figure 1 for locations), each black dot represents the average number of crab per potlift in a given week in a given area. Blue lines are fits of a linear regression to the data with the grey shading representing one standard deviation. Percentages on each panel represent the percent of the total catch in that year caught in a given area.

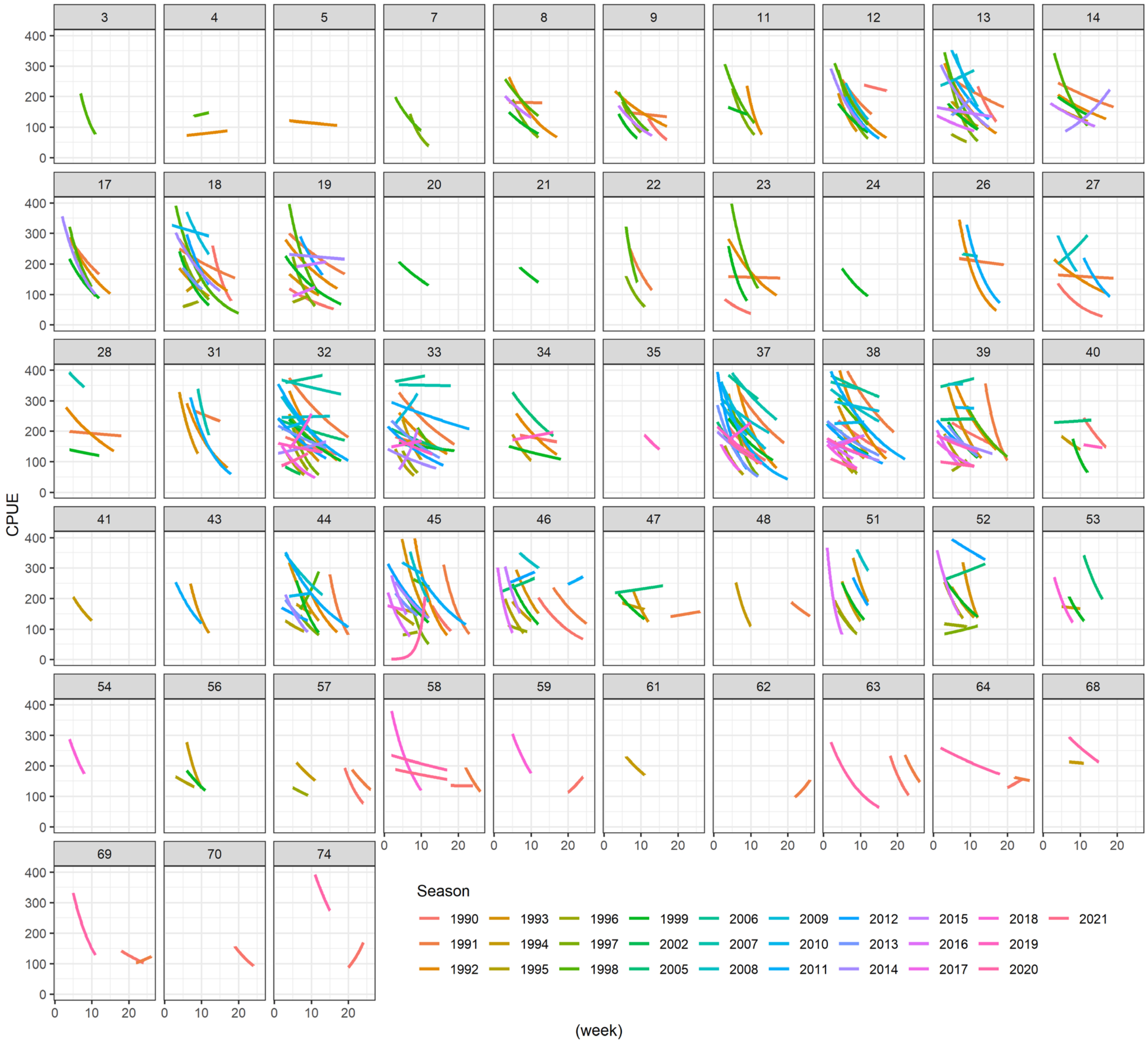

Figure 3. Area-specific regressions for an all years. Each panel represents an area (see figure 1 for locations), each colored line represents the fit of a regression for a given year to the cpue data from that area.

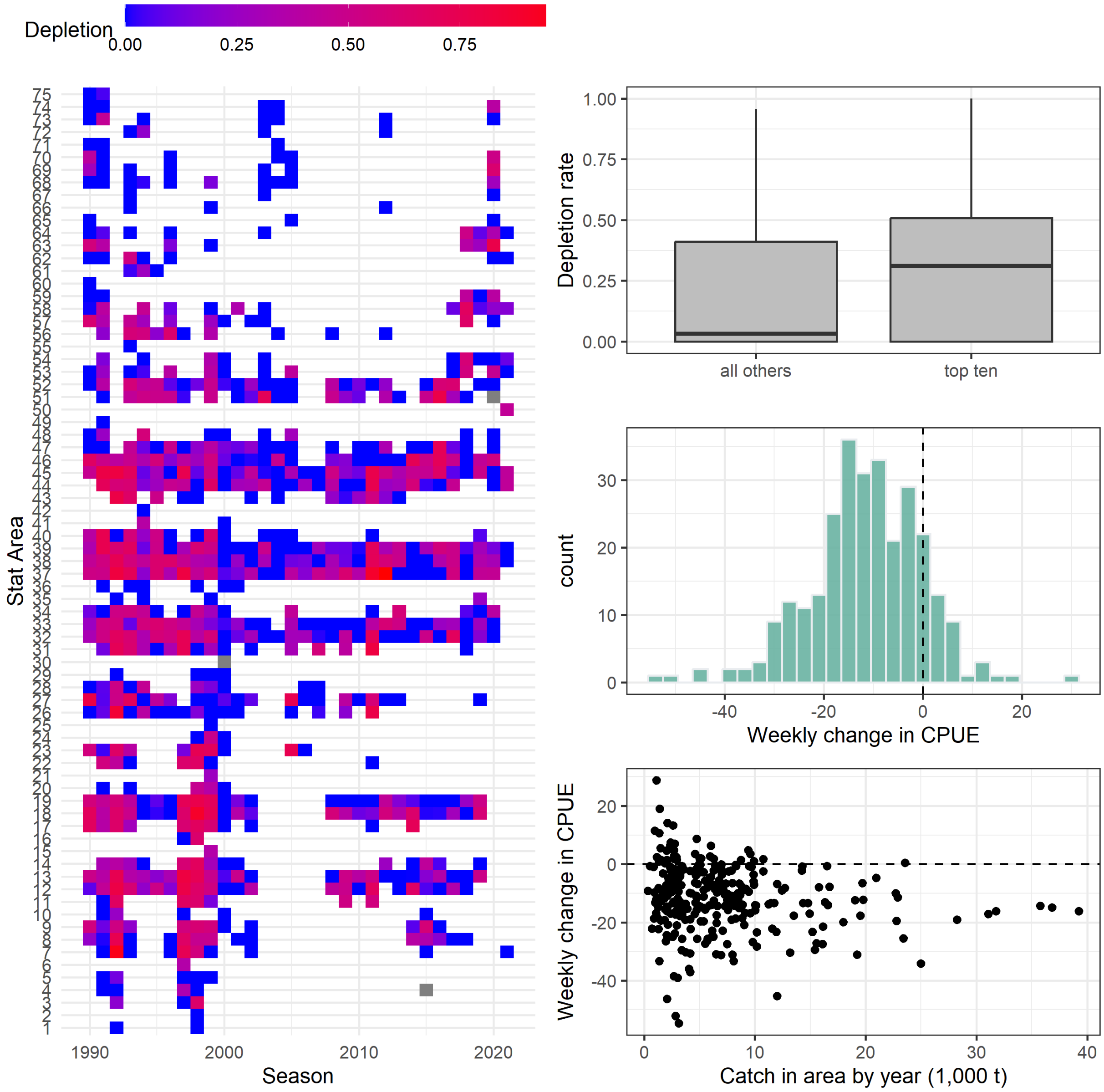

Figure 4. Depletion summary statistics. Depletion by area by season (left), boxplot of depletion rates by top ten producing areas and all others (top right), histogram of weekly change in CPUE over all seasons and areas (middle right), and relationship between weekly change in CPUE and the catch coming from that area (bottom right).

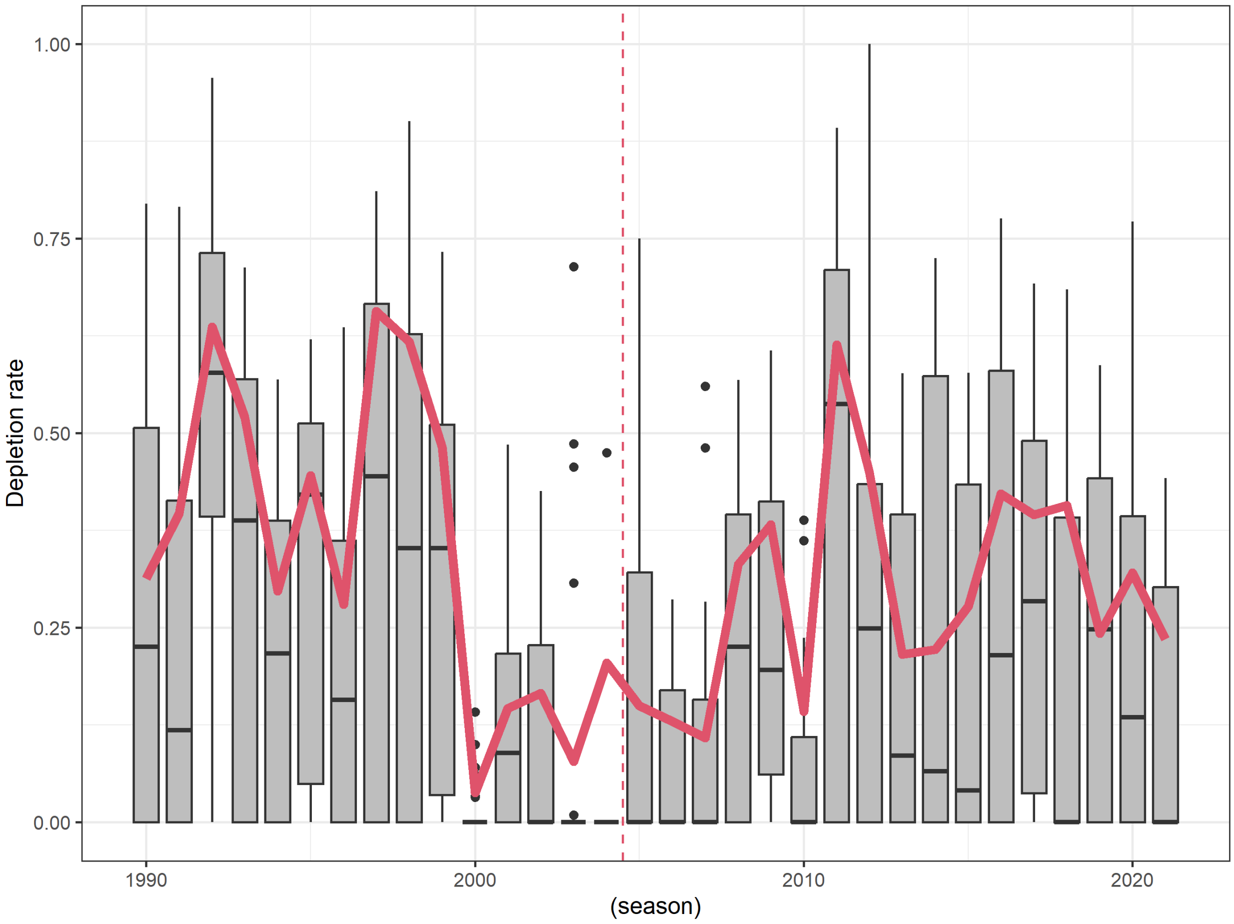

Figure 5. Boxplots over area of depletion by season. The red line represents the mean depletion weighted by the catch coming from an area.

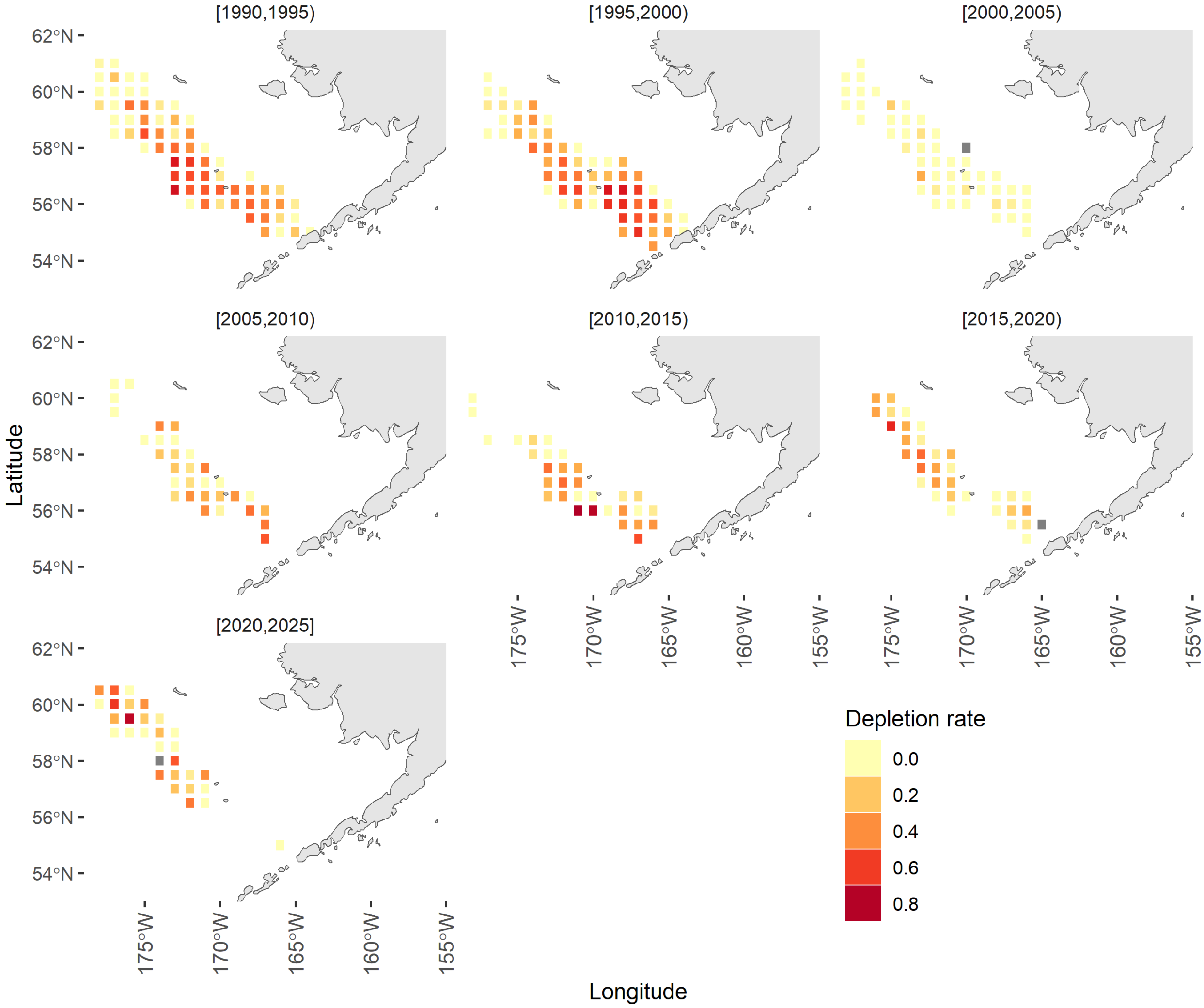

Figure 6 Maps of median depletion rates over five year intervals.

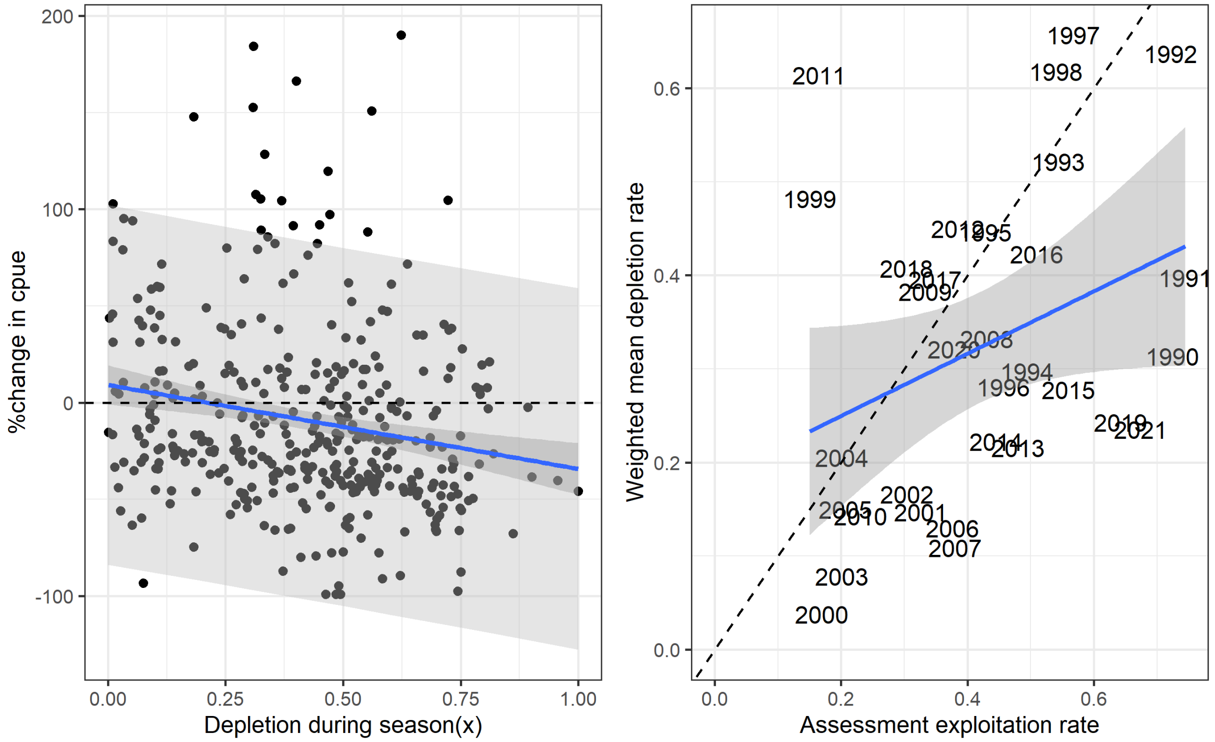

Figure 7. Comparison of % change in CPUE between seasons and depletion (left) and weighted mean depletion rate to the exploitation rates estimated in the 2024 stock assessment for eastern Bering Sea snow crab by year (right). Dashed black line is the zero change on left and 1 to 1 line on right; blue line is a regression fit to the estimates (GAM on left; GLM on right). Grey shading is 1 standard deviation; light grey shading is 95% prediction interval.

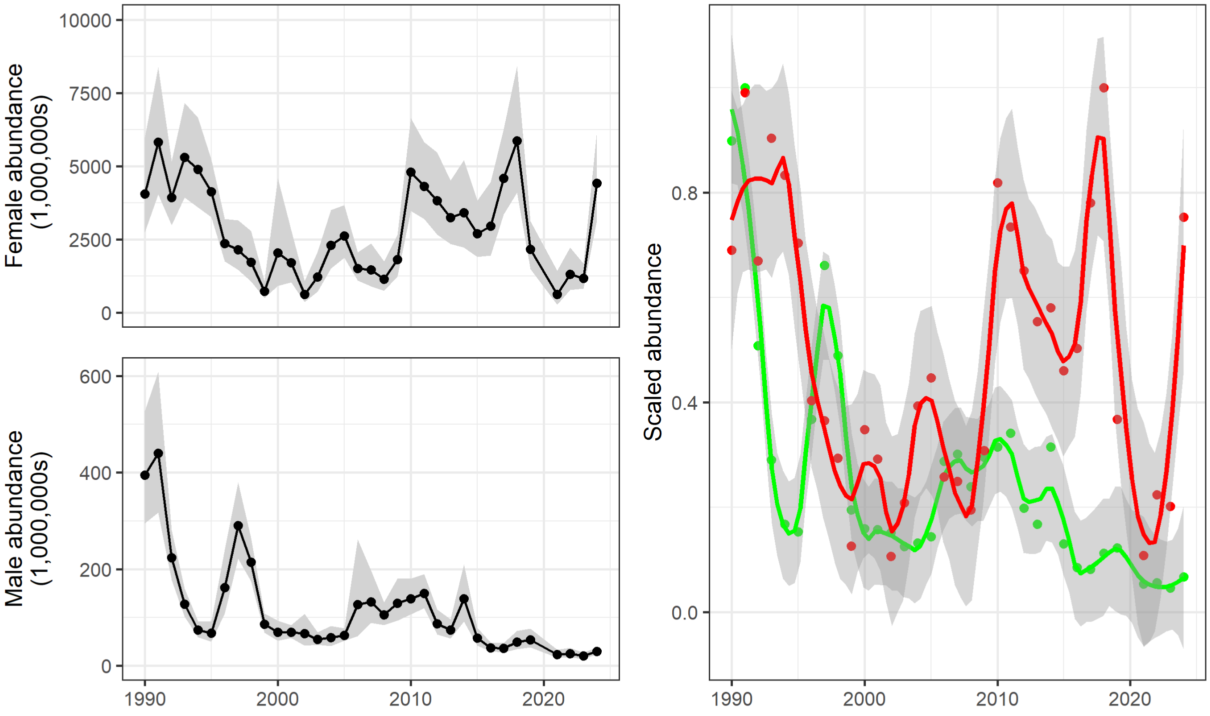

Figure 8. Survey estimates of female abundance (top left); large male abundance (bottom left) and both time series scale to their maximum and overlaid with a smoother (right; red = female, green = male).

F. Historical biases did not impact sustainability of Bering Sea crab fisheries

This manuscript is presented in draft form to facilitate discussion at plan team. It is not fully referenced yet and should not be cited anywhere until it has undergone peer-review.

Historical biases in Bering Sea crab assessments did not impact sustainability

Cody Szuwalski[1] and others to come

[1]Alaska Fishery Science Center, National Oceanic and Atmospheric Administration, Seattle, WA 98155

Stock assessments are a central tool of modern fisheries management and provide an estimate of the ‘status’ of a fishery. Stock assessments are updated regularly with new information and the present and historical status of a fishery may change based on the addition of new data or changes in assumptions. Recently, a paper by Edgar et al. claimed historical assessments were biased high for assessments of stocks that are currently at low levels, suggesting that the biases were responsible for overfishing. Here, we present an in-depth analysis of historical assessments and identified biases for Bering Sea and Aleutian Islands crab in Edgar et al. We found serious errors in the data used by Edgar et al. and major shortcomings of the methods used, which calls into question their results for other stocks. In spite of these errors, we still identify changes in perception of stocks over time in Bering Sea crab based on changes in assessment. Fortunately, the management system for Bering Sea crab fisheries (like many others) includes provisions that aim to account for this sort of scientific uncertainty. Historical biases in biomass were much smaller than the buffers for uncertainty subtracted from the overfishing level when recommending "acceptable biological catches" to guard against overfishing. Historical biases are an important quantity to track over time, and our analysis underscores the importance of collaboration with regional scientists in global analyses so that data are interpreted correctly, testing methods so they are performing as expected, and understanding the context of a system so that judgments about its management are grounded in reality.

Introduction

Stock assessments are statistical models that can use many sources of data (e.g. surveys, fishery catch, growth data, experimental data, mark/recapture data) to estimate the abundance or biomass of a population (Maunder and Punt, 2013). Stock assessments can also be used to calculate biological reference points that are often used in management to determine target fishing mortalities to set catches consistent with an agreed upon harvest policy aimed at achieving 'sustainable' catches (Gabriel et al., 1999). The target fishing mortalities take into account how fast individuals in the population grow, what fraction die each year, how much they reproduce, and how they interact with the fishery based on their size or age. The 'status' of a fishery relative to overfishing can also be calculated using reference points derived from the stock assessment. The data and assumptions used in stock assessment models are usually updated on a regular basis to provide management advice that reflects the best available science.

The widespread application of stock assessments and use of reference points based on maximum sustainable yield coincided with a reversal of overfishing for many marine fisheries in the 1990s (Hilborn et al., 2023). Recently, however, Edgar et al. (2024) questioned the utility of stock assessments in accurately portraying the sustainability of fisheries. They attempted to identify historical biases in assessments by comparing the estimates from historical assessments to those from a more recent assessment. If the historical assessments had different estimates than the most recent assessment, the difference was considered a bias. Importantly, they noted that stocks that are currently at low relative levels had higher levels of positive historical bias (i.e., historical estimates were larger than current estimates of the historical period). The implication of this piece of their analysis is that a manager's understanding of the size of the stock based on the assessment estimates was overestimated in the past, which led to catches that were too high, which consequently led to the current overexploited state.

Here we provide an analysis of historical biases for Bering Sea and Aleutian Islands crab identified by Edgar et al. We are scientists who conduct the assessments used for these stocks and/or members of the North Pacific Fishery Management Council's Crab Plan Team, which is the first body of peer review for the science used in management of these crab. We evaluate the data used by Edgar et al., and discuss the ability of their methods to capture the impacts of historical biases on management outcomes. We examine the actual harvests taken over the period of Edgar et al.'s analyses to evaluate if historical biases in assessment output would have led to unsustainable management. Finally, we conclude with recommendations on methods for evaluating the impact of historical biases in stock assessments, a discussion about the underlying causes of the trajectories of Bering Sea and Aleutian Islands crab stocks, and the central importance of stock assessment in fisheries management.

Stock-by-stock methods and data

Edgar et al (2024) used records of historical assessments from NOAA's StockSmart system (NOAA, 2024). Nine putative stocks from the Bering Sea and Aleutian Islands were included in their analysis, two of which were defined as ‘collapsed’, three designated “overfished”, and four labelled “lightly fished” or “sustainable” (table 1). The assessments spanned the years 2005-2019, with a total of 48 individual assessments in their analysis (figure 1). The basic methodology Edgar et al. used to identify historical biases in stock assessments was accomplished in two steps: 1) translate each time series of estimated spawning biomass (SSB) from a given assessment to “depletion” by dividing the time series by the maximum estimated value from that particular assessment (figure 2), and 2) calculate the difference between the terminal year of a historical assessment and the estimate of depletion in that same year from most recent assessment to represent the ‘historical bias’. It is important to note that the year having the maximum estimated value of biomass can change from assessment to assessment in their methodology. The decision to use ‘depletion’ rather than SSB can dramatically change the perception of historical bias for a variety of reasons discussed below (figure 3).

In addition to discussing the data used by Edgar et al. for each Bering Sea and Aleutian Islands crab stock, we collated the overfishing levels (OFLs; the amount of catch beyond which overfishing would be deemed to occur), the acceptable biological catches (ABC; which is the OFL decremented by some fraction to account for scientific uncertainty), and the actual harvests, which are closely related to the total allowable catches (TAC; the catches that were set by the State of Alaska and given as quota to stakeholders). Management of Bering Sea and Aleutian Islands crab is shared by the U.S. Federal government and the State of Alaska. The Federal system is tasked with setting the OFL and ABC which then acts as a ceiling for the State of Alaska TAC setting process. The harvest control rules used at the Federal and State level can be different and often are based on different fractions of the total biomass of crab in the ocean and use different definitions for target fishing mortalities and biomasses. For example, federal harvest control rules for snow crab are based on morphometrically mature male biomass, but the state harvest control rules are based on the sum of male and female biomass.

To understand how the buffers built into the management of crab influence the harvests relative to the management advice and subsequent ‘sustainability’, we calculated the percent difference between the OFL and the harvest and compared that to the percent difference between the terminal year of SSB for each historical assessment and the most recent assessment. Importantly, we examining change in SSB, not depletion, as it is the quantity that drives the absolute value of removals and depletion has some quirks (described below) that make interpretation difficult. A decrease in the OFL that is larger than the historical difference in SSB would generally suggest that the ‘historical bias’ would not have resulted in overfishing. Below, we discuss the data and the effectiveness of the methods used by Edgar et al. to identify the impact of historical biases on the sustainability of the fisheries in the context of the actual removals for each Bering Sea and Aleutian Islands crab fishery.

Bristol Bay red king crab

Bristol Bay red king crab (*Paralithodes camtschaticus*) is historically the most valuable crab stock in the Bering Sea. The Bristol Bay red king crab population declined abruptly from high levels during the early 1980s and has remained at comparatively low abundances since (Palof et al., 2024). The fishery has been closed five times in 1983, 1994-1995, and 2020-2021. Assessments from 2007-2018 (excluding 2010 and 2017) were included in Edgar et al.'s analysis and the stock was placed in their 'overfished' category (Table 1).

Curiously, Edgar et al. only included data from the period of high abundance in their analysis of the most recent assessment (figure 1). This exclusion has important consequences for the rest of the analysis for historical biases for this stock. When each time series was translated to 'depletion', all assessments apart from the most recent were rescaled to a year in the 2000s (figure 2), however the most recent assessment was rescaled to a year in the 1970s with much higher abundance. The net result of treating the data in this manner is a very large historical bias (i.e. historical assessments estimated much higher depletion than the most recent), even when the differences in the actual estimates of biomass were very small over the assessed time period (figure 3).

It is unclear why the early years of the assessments were excluded from Edgar's analysis. They are available in the database from which the assessment estimates were drawn. An additional issue with the data used for this stock is a difference in units reported. All years of assessment biomass used by Edgar et al. were reported in mature male biomass except for 2007, which was reported in total biomass (i.e. the sum of male and female mature biomass). This results in an inappropriate comparison, but is a relatively small issue compared to the exclusion of early years from all but the most recent assessment.

A potentially important bias was apparent in the 2008 and 2009 assessments compared to the most recent assessment. The estimates of absolute biomass in the terminal year from 2008 and 2009 were markedly higher than that estimated by the most recent assessment, which could have translated to higher OFLs in that year. However, the harvest is 45.7% smaller than the OFL (figure 4 and 5), which more than compensates for the median difference in SSB of -12% (Figure 6). A negative median difference in SSB means that historical assessments indicated biomass was lower in a given year than the most recent assessment. The negative median historical biases and the large buffers suggest that the historical biases seen in Bristol Bay red king crab assessments did not impact the sustainability of the fishery.

Snow crab

The snow crab (*Chionoecetes opilio*) fishery is the largest by volume in the Bering Sea, and the second most valuable. The snow crab population recently collapsed after a marine heatwave occurred in the Bering Sea in 2018 and 2019 (Szuwalski et al., 2023). After the collapse, the stock was declared overfished and the fishery was closed for the first time ever (Szuwalski, 2024). Eight assessments spanning 2005-2016 were included in Edgar et al. (2024) and the stock was designated as 'sustainable' by their methods (table 1).

The untransformed data for Bering Sea snow crab used by Edgar et al. seem to imply very large historical biases for this stock. The biomass estimates for the first three assessments have maximum biomasses nearly 3x the more recent assessments. Unfortunately, this is because the units were not accounted for appropriately. The assessments from the earliest three assessments were in total biomass, whereas the more recent assessment estimates were only for mature males. This change coincides with the use of mature male biomass as the 'currency' of management and was noted in the NOAA database used in Edgar et al.'s analysis.

Interestingly, the relative scale of the assessments with different definitions of biomass flips when the time series are transformed to depletion (figure 2). The more recent assessments have a higher relative depletion than the older assessments, implying that the historical perception of the stock was worse than it actually was (based on Edgar et al.'s interpretation of the historical biases they calculate). This is a good example of why the methods used by Edgar et al. should not be used to identify historical biases. Management and harvest advice is based on absolute biomass estimates from the assessment. If one only looks at the absolute biomass estimates from the 2005 assessment, the catch advice in 2005 would likely be more than double that provided by the estimated of biomass in 2005 from the 2019 assessment. However, when translating this to depletion, the estimate of 2005 from the 2005 assessment is actually lower than the estimate of 2005 from the 2019 assessment.

Based on the data used in Edgar et al., there was a median difference of 78.1% in SSB relative to the most recent assessment in their database (figure 6). This difference in SSB is larger than the median 62.8% decrease in the OFL to the harvest seen for snow crab (figure 5), which might present some concerns with respect to the impact of historical biases on sustainability. However, if a more recent assessment is used (2023), the median difference in SSB was -39.6%, implying that historical assessments suggested that there was less biomass in the water than the current assessment suggests.

Tanner crab

Tanner crab (*Chionoecetes bairdii*) is the third largest crab stock in the Bering Sea and the fishery has been intermittently open over the past several decades (Stockhausen, 2024). The stock declined from very high abundances in the 1970s and abundance has remained volatile since the 1980s, just at a lower amplitude (figure 1). Six assessments spanning 2008-2018 were included for Tanner crab. In Edgar's analysis, the assessments for this stock were divided into two different "stocks" for unknown reasons (i.e. "Southern Tanner crab_Bering Sea" and "Tanner crab Bering Sea and Aleutian Islands"). "Southern Tanner crab_Bering Sea" was placed in the 'sustainable' category and "Tanner crab Bering Sea and Aleutian Islands" was placed in the 'overfished' category. The historical biases for the 'sustainable' pseudo-stock were all negative (i.e. historical assessment estimates were lower than the most recent estimates for the terminal year), whereas the historical biases for the 'overfished' pseudo-stock were positive (i.e. historical assessment estimates were higher than the most recent estimates for the terminal year). Selectively splitting the stock assessments and excluding years from the assessments like this would result in the perception that 'overfished' stocks have higher historical biases, but this is only a result of the splitting and exclusion of data.

Historically, the median differences in SSB were small for both 'stocks' (-4% and 6%; figure 5) and the median difference between the OFL and the realized harvest was 97% (figure 4). This is a result of the fishery being closed by the State of Alaska for many years in which a federal OFL was still greater than zero. The frequent closures of the fishery and large reductions of the OFL negate any impact the small changes in estimated SSB could have had on the ability to achieve fishery management goals.

Pribilof Islands blue king crab

Pribilof islands blue king crab (*Paralithodes platypus*) is a small stock associated with St. Paul and St. George islands. This stock has no population dynamics model from which a 'bias' could arise because available biomass estimates are either raw survey estimates or survey estimates with a smoothing algorithm applied (e.g. moving average; Stockhausen, 2024). A fishery has not been open for this stock since the 1990s and even then it was very small compared to other fisheries. Four 'assessments' spanning 2010-2019 were included in Edgar et al.'s analysis for this stock and it was designated 'collapsed'. The raw biomass estimates look very similar over time, the biggest difference being that the more recent assessments began applying a smoothing algorithm to the survey data. In spite of these similarities, when transformed to depletion, the time series appear quite different as a result of different starting years. Ultimately, given there is no actual assessment model and there has been no fishery for this stock for three decades, its inclusion in an analysis of the impact of historical biases from assessment models on management outcomes is probably not justifable.

Saint Matthew Island blue king crab

Saint Matthew Island blue king crab is another small, island-associated stock. A fishery has only been open for this stock in six years since 2006. Eight assessments spanning 2007-2019 were included in Edgar et al.'s analysis for this stock and it was designated "collapsed". This stock presents concerning historical biases, due in large part to a recent increase in survey biomass that the assessment model has a difficult time fitting given the current assumption of a constant, relatively low natural mortality (Stern & Palof, 2024). The upward 'excursions' of the terminal years of assessment estimated biomass during the late 2000s and early 2010s are following a sharp increase in observed survey biomass (figure 1). Given that a constant natural mortality is used in this assessment, there is no other way for the model to remove large amounts of crab from the population except the fishery. So, ultimately, as low values of survey biomass were added to the assessment, the model was forced to ignore the large observed survey biomasses during the 2010s in the model fitting process and this resulted in the historical biases observed. So, although there is a very large median difference in SSB (figure 5), these differences are an artifact of an inflexible model and the frequent closure of the fishery and the resulting 100% median difference between the OFL and realized harvest suggest that the historical biases have not greatly impacted the sustainability of the fishery (figure 6).

Norton Sound red king crab

Norton Sound red king crab is the most northern of the crab stocks fished in the Bering Sea. It is the smallest consistently open fishery, with no closures during the past two decades, although little to no commercial harvest occurred in 2020 and 2021 due to lack of a commercial buyer (Hamazaki 2024). Six assessments spanning the years 2008-2017 were included in Edgar et al.'s analysis and the fishery was designated as "sustainable". The time series of raw assessment-estimated biomass were fairly similar (figure 1), with the earliest assessments having lower values than the most recent. However, when the time series were transformed to depletion, the most recent assessment was lower than many of the historical assessments. This transposition of time series comes primarily from changes in the estimates of the first 2-3 years of biomass from the stock assessment. The first years of biomass estimates from stock assessment models can be more uncertain than estimates in the middle of the time series because they do not have data on both sides (in time) that can corroborate estimated trends. Ultimately, the median difference in SSB was -8%, suggesting that biomass was underestimated historically. This coupled with the 40.8% median decrease of the OFL to realized harvest suggests that historical biases have not impacted the sustainability of this fishery.

Aleutian Islands golden king crab

No large issues were identified for the four Aleutian islands golden king crab (*Lithodes aequispinus*) assessments included in Edgar et al.'s analysis. The median decrease of the OFL to realized harvest was 47.4% which is much larger than the median difference in estimated SSB (-5% and 12%, depending on the stock). It should also be noted that assessments exist for Pribilof islands red king crab, but this stock was not included in Edgar's analyses.

Discussion

Historical biases are a potentially important management problem and should be regularly evaluated given the potential for them to result in overfishing. Edgar et al. attempted to identify historical biases in a selection of global stock assessments, including Bering Sea and Aleutian Islands crab. However, we identified several problems with their interpretation of the data and their methodology. There were serious problems with the data for 5 of the 9 'stocks' included in their analyses. A lack of familiarity with the region and stock assessments resulted in splitting a single stock into two stocks, using time series that were of different units for the same stock (and therefore not appropriate for comparisons), and including a stock for which there is no model-based stock assessment. Selective inclusion/exclusion of years within assessments for a given stock resulted in large exaggerations of historical biases.

Even if all of the data were correct, the methods used do not appropriately capture the potential for historical biases to impact the sustainability of fisheries for several reasons. First, translation to depletion can result in the identified bias changing from negative to positive (and vice versa) when compared to the bias observed from the untransformed biomass (see snow crab for example). Recommendations of catches are the primary method by which managers influence the sustainability of these crab fisheries and total allowable catches are most strongly influenced by the absolute estimate of biomass produced by a stock assessment. The act of scaling to the maximum estimate of biomass from a given assessment is fraught with potential pitfalls. One of the biggest can arise from uncertain initial conditions like in Norton Sound red king crab or exclusion of periods of data as in Bristol Bay red king crab.

A thought experiment might more clearly illuminate the problem. Imagine a stock for which all of the estimates of biomass are nearly identical for the last three decades and assessments began ten years ago (figure 6). However, the year in which the assessment begins (the timing of which is identical for all assessments available) has estimates of biomass that have been revised downward over time, but are always the maximum value in the time series. Scaling each of the time series to that year would result in very different 'depletions' and ultimately large historical biases. These biases would be present even though historical bias is close to zero when looking at absolute biomass.

Even if one ignores the inability of Edgar et al.'s methods to effectively identify historical biases, a historical bias is only a problem if it resulted in 'unsustainable' management advice, and Edgar et al. made no attempt to evaluate how historical biases actually impacted management. We demonstrated here that often there were no fisheries occurring during the periods in which historical biases were identified for Bering Sea and Aleutian Islands crab. Further, when there were potentially concerning historical biases, the actual harvests were much smaller than the recommended catches because of buffers for uncertainty built into the management systems and the difference between the federal and state harvest control rules (Crab FMP REF).

For all of these reasons, it is difficult to place much confidence in Edgar et al.'s assertions that historical biases in stock assessments suggest a misrepresentation of the sustainability of fisheries. Nonetheless, historical biases need to be considered in the management process. Stock assessment is an iterative process and perceptions of stocks can (and do) change based on new information and updates to assumptions about fishery and biological processes. Our analyses underscore the importance of considering the absolute biomass used instead of translating to depletion and evaluation of the biases in light of whether or not they resulted in 'unsustainable' management advice. In future analyses of historical biases, emphasis should be placed on how knowledge of a bias would change management advice because the largest 'lever' available to management is changing fishing removals.

Many Bering Sea and Aleutian Islands crab stocks are at low levels compared to historical levels, but historical biases in stock assessments are not to blame for this. For example, snow crab collapsed after a marine heatwave in 2018 and 2019 (Szuwalski et al., 2023). Pribilof Islands blue king crab collapsed 23 years ago and has not been fished for decades but still has not recovered

(Stockhausen, 2023). Bristol Bay red king crab collapsed in the early 1980s and in spite of what appears to be conservative management has still not rebuilt to previous levels (Palof et al., 2024). In some cases we have built plausible explanations for the declines and in every case for which a plausible explanation has been developed, stock assessments were a primary tool in understanding past dynamics. Stock assessments are also the only tool we have to attempt to determine appropriate management responses to rebuild these stocks. One of the primary utilities a stock assessment model provides is the ability to ask 'what if' questions. We generally do not perform large scale experiments in fisheries given the potential for adverse outcomes, so models are the primary method for 'experimentation'.

Finally, as stock assessment scientists, it is frustrating to see an analysis published in a high profile journal that disparages the work that we do without adequately understanding it. The journal Science has discontinued the 'technical comment' so that it is not possible directly address inaccuracies in the material presented so that there is a permanent link between the research in question. Stock assessment scientists spend their careers trying to incorporate a wide range of data into models to present as comprehensive a picture of a population as possible given the best available science. In this process, we routinely identify knowledge gaps and collaborate with other scientists to design experiments to fill these gaps. Updated information can change our perception of the dynamics of a stock, and this updating is an important part of the scientific process. Decisions on acceptable biological catches need to be made yearly even when we do not have all of the information that we would like. Fisheries managers do not blindly follow assessment output and fisheries management plans have provisions included that allow for buffers for scientific uncertainty like historical biases. Our job entails understanding the uncertainties, presenting the stock assessments and their uncertainties to stakeholders, and building trust in the management process.

Given our common interest in producing the best possible assessments of exploited populations and contributing to the achievement of management goals, we welcome critical evaluations of the performance of stock assessments over time. We hope that, in future, the authors of Edgar et al. and others interested in the contributions of stock assessment models to sustainable fisheries will engage in a more collaborative approach to understanding and analyzing the data sets and assessments used to manage the Bering Sea and Aleutian Islands crab stocks as well as other stocks around the globe.

Acknowledgements

We would like to thank the authors of Edgar et al. for providing a repository for their code.

References (to be completed)

Edgar, GJ, Bates, AE, Krueck, NC, Baker, SC, Stuart-Smith, RD, Brown, CJ. 2024. Stock assessment models overstate sustainability of the world's fisheries. Science. 385: 860-865.

Hamazaki, T (2024) Norton Sound red king crab stock assessment for the fishing year 2025. North Pacific Fishery Management Council, Anchorage, AK. Available from https://www.npfmc.org/library/safe-reports.

Hilborn, R, et al. 2024. Effective fisheries management instrumental in improving fishery status. PNAS. 117 (4), 2218-2224.

NOAA StockSmart. 2024

Palof, K. 2024 Bristol Bay red king crab assessment.

Stern, CA, Palof, K. 2024. Saint Matthew Island blue king crab stock assessment 2024. North Pacific Fishery Management Council, Anchorage, AK. Available from https://www.npfmc.org/library/safe-reports.

Stockhausen, W. 2023. PIBKC assessment.

Stockhausen, W. 2024. Tanner crab assessment.

Szuwalski, C. 2024. Eastern Bering Sea snow crab assessment.

Maunder MN and Punt AE (2013) A review of integrated analysis in fisheries stock assessment. Fisheries Research 142:61-74.

Gabriel W. L. Mace P. M.Restrepo V. R. (1999) A review of biological reference points in the context of the precautionary approach. In: Proceedings of the Fifth National NMFS Stock Assessment Workshop: Providing Scientific Advice to Implement the Precautionary Approach under the Magnuson-Stevens Fishery Conservation and Management Act, USA National Marine Fisheries Service (pp. 34-45). NOAA Technical Memorandum NMFS-F/SPO-40.

Stock	**Status (Edgar et al.)**	**Status (NPFMC)**	Error in analysis
Blue king crab (Pribilof Islands)	Collapsed	Overfished	No assessment
Blue king crab (Saint Matthew Island)	Collapsed	Overfished	
Golden king crab (Eastern Aleutian Islands)	Lightly fished	Not overfished	
Golden king crab (Western Aleutian Islands)	Lightly fished	Not overfished	
Red king crab (Bristol Bay)	Overfished	Not overfished	Mismatched units Excluded data
Red king crab (Norton Sound)	Overfished	Not overfished	
Snow crab (Bering Sea)	Sustainable	Overfished	Mismatched units Excluded most recent assessment
Southern Tanner crab (Bering Sea)	Sustainable	Not overfished	Not a distinct stock
Tanner crab Bering Sea and Aleutian Islands	Overfished	Not overfished	Not a distinct stock

Table 1. Bering Sea crab “stocks” used in Edgar et al. (2024) analysis on historical biases in stock assessments. Status based on Edgar et al. and the errors found in the data used for the

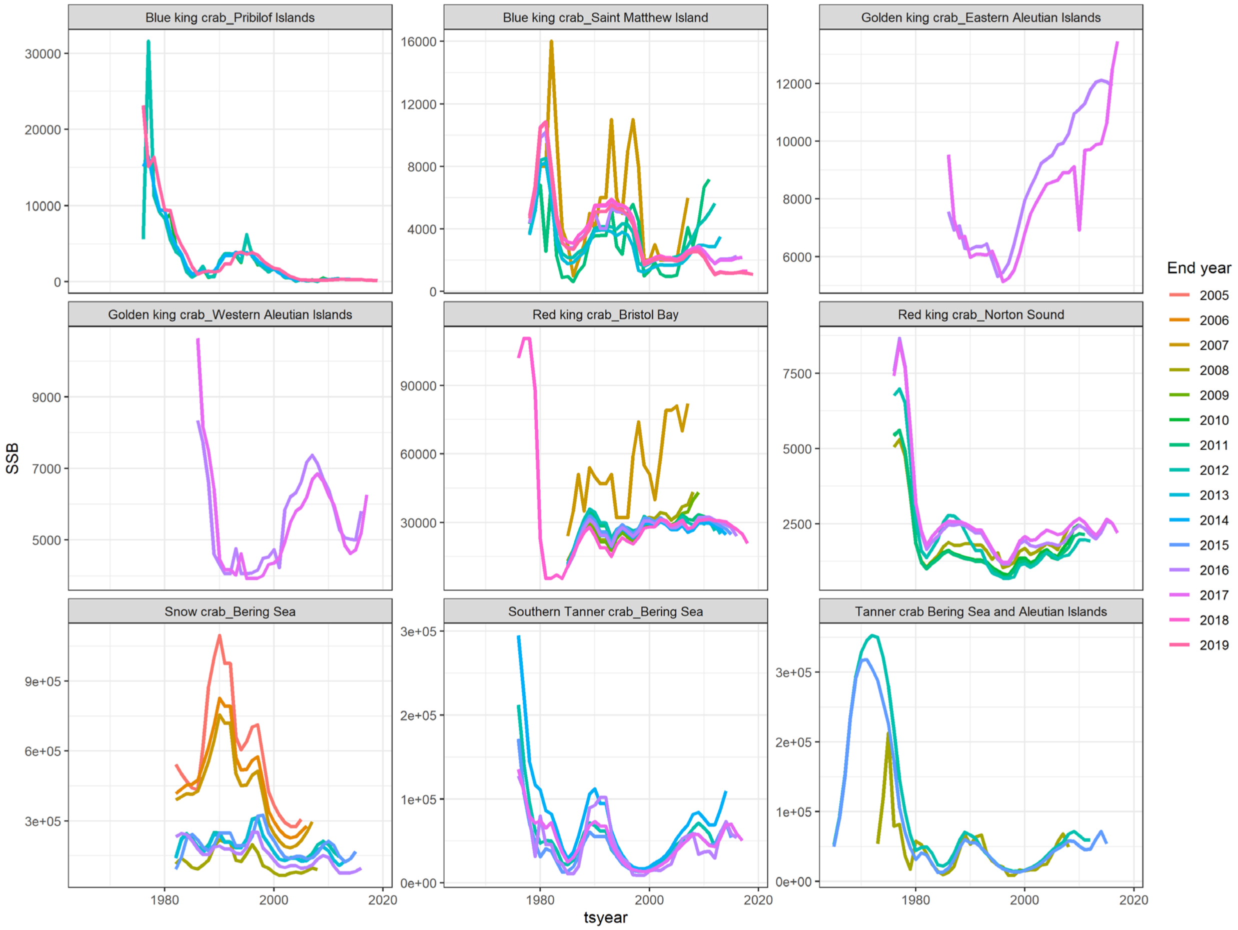

Figure 1. Raw spawning biomass time series for Bering Sea crab used in Edgar et al.'s (2024) analysis on historical biases in stock assessments.

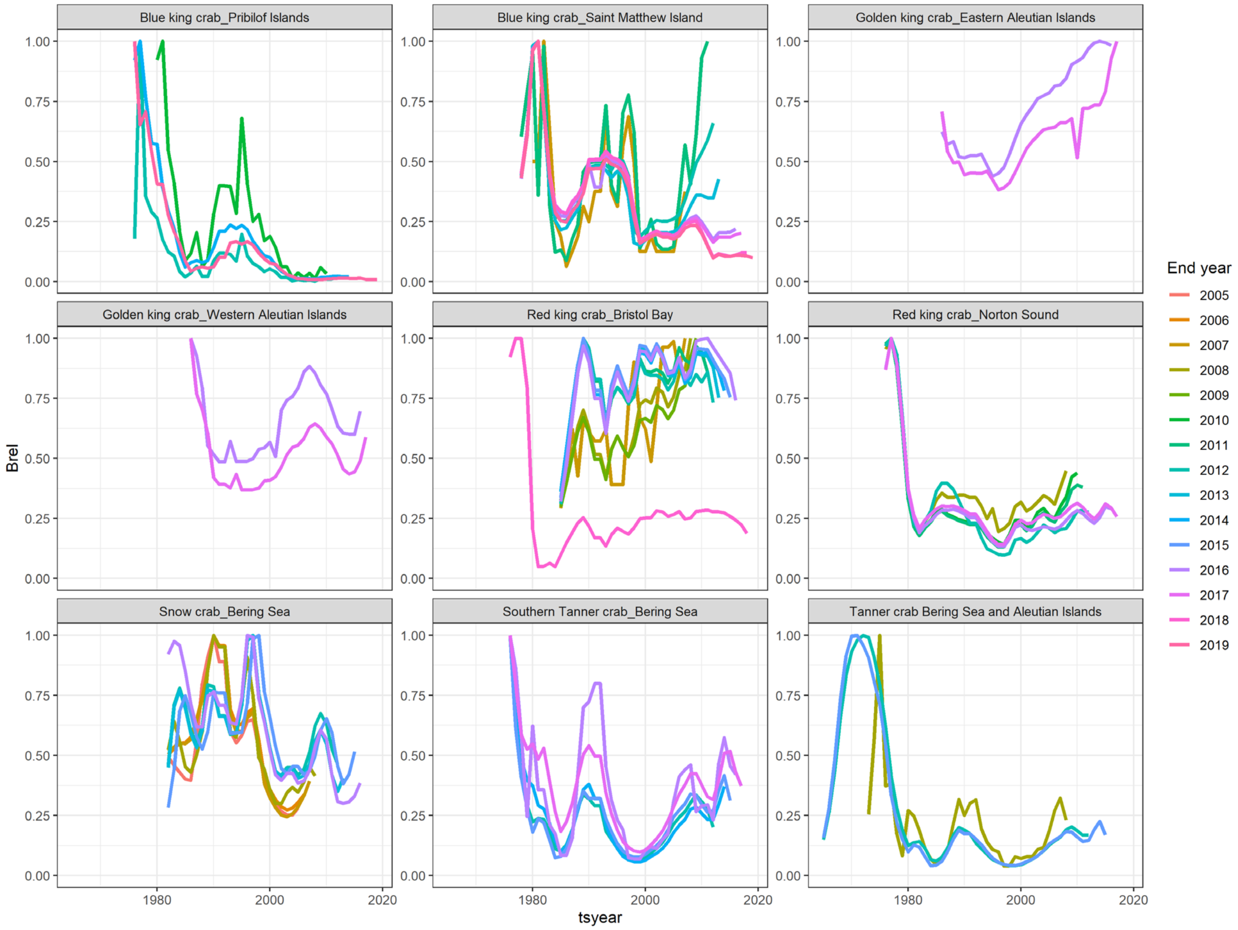

Figure 2. Spawning biomass time series transformed to 'depletion' (by dividing each time series by its maximum) for Bering Sea crab used in Edgar et al.'s (2024) to identify historical biases in stock assessments.

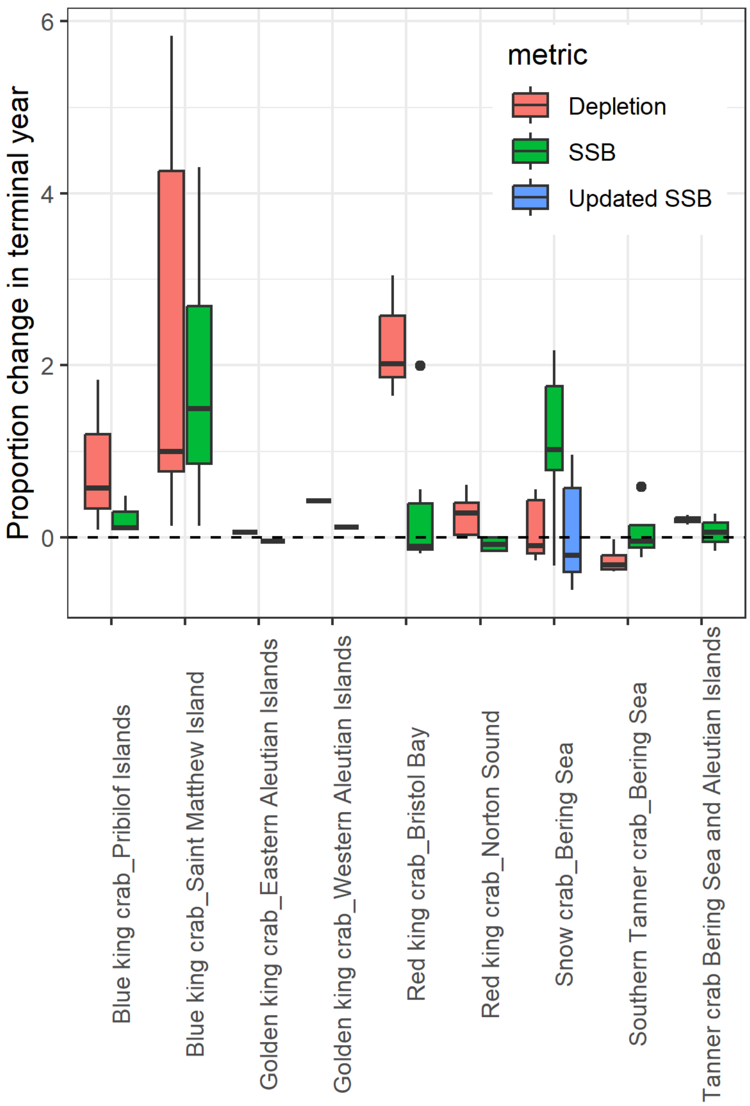

Figure 3. Proportion change in the terminal year of depletion and spawning biomass when compared to the most recent assessment used in Edgar et al. (2024). A proportion greater than zero indicates that historical assessment estimates were higher than the most recent assessment. The final assessment for snow crab included in Edgar et al. (2024) ended in 2019, so the calculation was performed with the 2023 assessment which is shown above in blue and noted in the legend as 'updated SSB'.

Figure 4. Overfishing level (blue), Acceptable Biological Catch (red) and actual harvest (green) for crab stocks in the eastern Bering Sea.

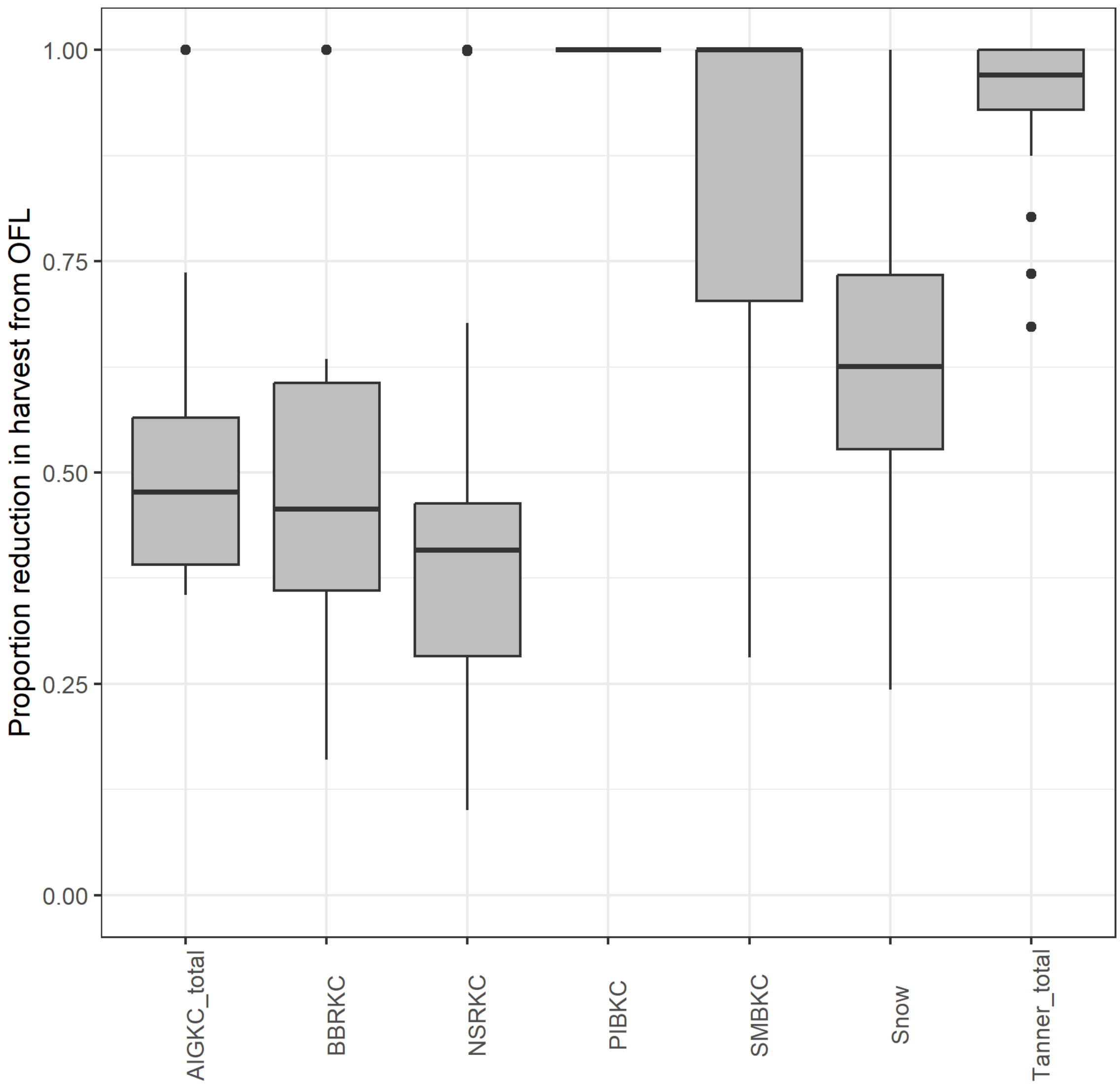

Figure 5. Proportion reduction from the OFL to the actual harvest for crab stock in the eastern Bering Sea over the period represented in figure 3.

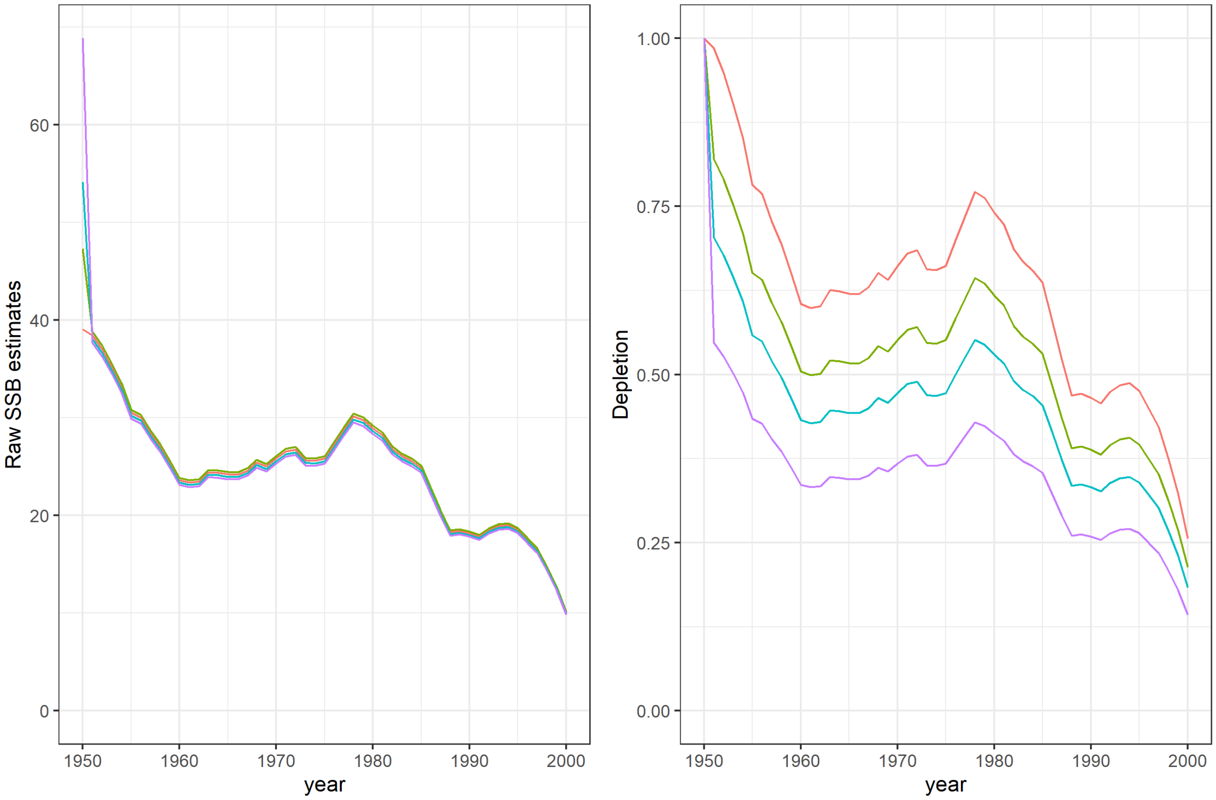

Figure 6. Hypothetical example demonstrating one statistical pitfall of the depletion method used by Edgar et al. Left panel shows four simulated assessments that differ very little apart from the estimates of the first year of biomass. Right panel shows the inferred 'depletion' when Edgar's methods are applied to the time series.