Considering different currencies of management for eastern Bering Sea snow crab

Szuwalski, C.S.

Summary
No new operational models are presented here, this document is intended to clarify modeling choices and provide analyses to support decisions on management currency.

Problem:
Revisions to the assessment model to match available biological data resulted in $SBPR_{35\%}$ reference points that would allow complete removal of the large males from the population. This is a direct result of a larger number of morphometrically mature males being protected from the fishery when the probabilities of having undergone terminal molt were revised.

Goal:
Identify an appropriate currency of management (i.e. a definition of mature male biomass) and associated reference points that account for the best available science on the biology of eastern Bering Sea snow crab and achieve the goals of management.

Potential solutions:
There are at least three possible paths forward:
- estimate maturity parameters even though estimates don't match the observations
- specify maturity parameters, use morphometric maturity as currency, and adjust the SBPR%
- specify maturity parameters, change currency of management to reflect fishery selectivity, and retain $SBPR_{35\%}$

Pros and cons exist for each solution and discussion at the crab plan team will be useful for identifying motivating reasons to select one path over another. The author-preferred path is to specify maturity parameters because it is one of the few population processes informed by direct observations. Adjusting the SBPR% or changing the maturity definition after specifying maturity parameters have broadly similar outcomes, but both lack overwhelming information to support adoption of one over the other.

An analysis linking estimated mortality, recruitment, and probability of having undergone terminal molt to environmental covariates and density dependent effects is also included as an appendix. Projections under a receding ice scenario suggest a possible rebuilding in the short-term if environmental conditions are favorable, but large reductions in abundance of mature males in the eastern Bering Sea over the long-term.

<u>**SSC Comments and responses:**</u>
(*SSC comment in italics*, **response in bold**)

(Public comment) expressed a lack of understanding as well as lack of confidence in the stock assessment modeling. It was suggested that preparing a simple stock assessment narrative that documents recent history on model development for stakeholders could improve comprehension and buy-in. The SSC understands that assessment methods are technical and complex and agrees that more effort is needed

A description of the model, recent changes, and rationale for those changes are included below.

Alternative Snow Crab reference points - A major issue with Model 23.3a is that the value of F35% is extremely high and would effectively remove all the industry-preferred crab from the population. This occurs because snow crab mature at a smaller size than the size at which snow crab are retained by the fishery, so there is a component of mature males that are protected from fishing mortality. This was already an issue in the base model (23.1) but was further exacerbated in 23.3a as the effective maturity was moved to younger males and the length of full selection increased with the addition of the empirical terminal molt probabilities. The analysis that provided the basis for the F35% harvest rate (Clark 1991[1]) assumed that maturity and fishery selection curves were the same, but also considered scenarios where maturation occurred earlier than fishery selection. The extreme mismatch between maturation and selectivity seen for snow crab was not considered in Clark (1991). Furthermore, this mismatch was not found to be an issue when the analyses were done to support adoption of the tier system in the crab FMP. Therefore, potential alternatives to F35% should be considered for snow crab in the future. It is important to note that this flexibility is built into the crab FMP, which indicates that alternative values to the default reference points F35% and B35% can be recommended by the SSC based on best available information. The SSC did not support replacement of M for F35% in the Tier 3 OFL control rule. Due to the delayed fishery selectivity pattern relative to size composition of the exploitable males, natural mortality may be an extremely conservative harvest policy, and it would not be advisable to adopt without further evaluation. The SSC also finds weak support for moving to average MMB for the BMSY proxy as B35% provides a reasonable reference point. The stock has varied above and below B35% primarily due to recruitment variability that does not appear directly associated with harvesting. Indices of female reproduction have remained high, and the proportion of large males in the population has remained stable even while overall abundance has declined.

The idea of retaining some percentage of the reproductively important population is conceptually satisfying and relatively intuitive. The reproductively important part of the population (i.e. management currency) and appropriate percentage to be retained (i.e. reference points) need to be identified. SBPR analyses are performed below for different definitions of mature male biomass and at different percentages of unfished biomass as target to explore this question.

The SSC strongly supports the plans of the CPT to evaluate other metrics for reproductive output. The CPT may want to consider a multi-attribute measure of reproductive output. For example, both percent reduction in mature male biomass and percent reduction in large males could be evaluated as a function of fishing mortality.

Reproductive output appears to be strongly influenced by environmental conditions. Appendix 1 explores the implications of environmentally driven recruitment dynamics and receding ice in the Bering Sea. Short-term projections hold some possibility for rebuilding if conditions align; long-term projections suggest large-scale declines of mature male abundance in the eastern Bering Sea.

The SSC requests a yield analysis be done for snow crab, including the relationship between fishing mortality and catch, MMB, functional maturity, and the proportion of large males in the population. The stock production curve, i.e., yield as a function of MMB, should also be developed.

These were performed for a range of steepnesses and definitions of mature male biomass. Morphometrically mature male biomass could not be depleted to 35% of unfished levels over a wide range of steepnesses. Defining mature male biomass closer to the sizes impacted by the fishery (e.g. 95-100 mm carapace width) resulted in maximin solutions for $SBPR_{XX\%}$ closer to 35% of unfished biomass. See below for further analyses.

consider greater use of the modeling structure to diagnose problems in how the data are being interpreted as opposed to more generally viewing resulting models as potential options for management. Sensitivity and other exploratory approaches using the model should be conducted and presented diagnostically to inform a smaller set of self-consistent models for management considerations.

I think the SSC is asking me to delineate research vs. operational models more carefully and I will do my best.

One idea for statistical exploration regarding the shape of the within-model empirical smoothed estimate of selectivity would be to examine to what extent the spatial distribution of differences in availability of small and large crab (or males and females) would be sufficient to explain the anomalous shape of the survey selectivity curve.

I'm not clear what is 'anomalous' about the shape of the selectivity curve—the shape makes some intuitive sense to me. Very small crab would be poorly selected (they go under and through the gear), a range of medium sized crab would have similar selectivity higher than small crab (harder to go under and through the gear, but still possible) and then selectivity would increase to nearly one for the largest sized crab (the biggest crab do not escape the gear). This seems more reasonable than the historically used logistic curve that had the same selectivity for crab 50-150 mm carapace width. The SSC may also be referring to the small 'hump' at smaller sizes in the BSFRF data. Differences in aggregation behavior by size and maturity state could be related to this phenomenon.

The SSC still requests an analysis of the probability of maturing/terminal molt which treats years as random effects. A hierarchical fit to molt data might be better than annual independent GAMs.

I don't think I have explained this part of the assessment appropriately based on this comment and endeavor to do so more completely below. Reading Richar and Foy (2022; reference below) might also be helpful.

The SSC would like to better understand the sampling design for molt data and is concerned about the weighting of the spatial samples in the analyzing; 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). Hierarchical fit to molt data might be better than annual independent GAMs.

Sampling design and methodology for analysis of the chelae data to determine the probability of having undergone terminal molt at size by year is documented in Richar, J and Foy, R (2022) A novel morphometry-based method for assessing maturity in male tanner crab, Chionoecetes bairdi. FACETS. https://doi.org/10.1139/facets-2021-006

Figure 23 on page 73 of the SAFE report shows the decline in CPUE over a season by statistical area and year. This represents a kind of depletion experiment, suggesting that total mortality (Z) could be estimated from the linear parameters representing each line. This might help determine spatial patterns in F, indicate the natural bounds for F and M, and assist in determining stock status.

This will be explored at a later date.

Providing a clear crab specification narrative would help the SSC and the public navigate the tiers, models, and justifications for both. In addition, it would be helpful to clearly identify models that are being explored for diagnostic purposes as opposed to models that are directly relevant for use in decision making. Public testimony indicated that help and financial support for developing such a narrative might be available.

I will attempt to delineate research vs. operational models more effectively in September.

Regarding the missing socioeconomic indicators coded as NAs in Table 1b (page 123 of the SAFE report), while having all general socioeconomic indicators in one table would seem useful, the labels "low", "neutral" and "high" are not ordered consistently enough across categories to provide a quick good/bad assessment as the stoplight approach does in Table1a. So, in addition to considering remedies for that (is high good or bad?), consider separating quantitative metrics from qualitative metrics. Putting these two types of metrics in separate tables might help in interpreting the information, and if the stoplight coloring is used, putting the quantitative value within the colored box may also help succinctly present the information.

Erin has done this.

the SSC looks forward to a more complete document next year, including figures of composition residuals.

To be included in September.

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.

The differences in spatial distribution (and therefore environmental conditions) is probably the most useful piece of information to demonstrate differences in environmental conditions by the sexes. This has been shown in previous assessment documents, but will be shown again in September.

Avoid connecting 2019 and 2021 when plotting survey time-series (e.g., Figure 14) as there were no data in 2020.

Ok.

Report the scale of standardized residuals where plotted (e.g., Figures 42 and 43).

Ok.

Population dynamics synopsis and recent changes

I begin with a high-level description of the assessment model to provide context before describing recent changes made to the assessment. The assessment model year begins in July and runs through the end of June in the following year. Each year is divided into three parts: pre-fishery (~7 months), fishery (~instantaneous), post-fishery (~5 months). The survey occurs at the beginning of the pre-fishery phase. Crab of both sexes are captured with different probability based on their size and the selectivity of the survey gear. The only life history process that occurs pre-fishery is natural mortality. Then the fishery occurs, during which a fraction of the largest crab are removed from the ocean. From that total catch, a fraction of those crab are retained and a fraction of the crab returned to the ocean die. After the fishery, molting occurs and the amount a crab grows is determined by their pre-molt carapace width. Crab mature with some probability based on their post-molt carapace width and, when a crab matures, it no longer grows. Recruitment to the initial size bins in the model occur after growth.

All of the processes modeled in the assessment contribute to population dynamics in different ways and have different data sources that inform them (Table 1). Recruitment determines the number of crab entering the smallest size bins in the population and is informed primarily by survey abundance and size composition. Growth and maturity determine how the recruitment moves (or does not move) through the larger size bins. Growth increments are informed by measurements of pre- and post-molt carapace widths for 110 crab collected over several years (figure 1; see Szuwalski, 2023 for description of data). The probability of undergoing terminal molt determines at what size a crab ceases growth. Natural and fishing mortality (and selectivity) determine how crab that recruited to the population are removed from it. Fishing mortality is directly informed by observer data and is also influenced by estimated survey selectivity. Survey selectivity influences the interpretation of the indices of abundance and size composition data that inform other processes. Good estimates of survey selectivity are important to understanding the scale of the population (i.e. is every crab encountered by the survey gear captured?) and estimates of scale are in turn important to accurate estimates of fishing mortality. The BSFRF selectivity experiments collected data on the selectivity of the NMFS survey gear over several years, providing an informative data source that most fisheries do not have (figure 2).

Process	Data	Historical assumptions	Updated assumptions
Recruitment	Survey abundance + size composition	Equal sex ratio	Unequal sex ratios
Natural mortality	Longevity + survey data	Constant with strong priors	Strong priors and time-block in 2018-2019
Growth	Growth increment	Piece-wise	Linear
Maturity	Chelae height	Single estimated ogive	Input yearly observations
Fishing mortality	Observer data	Freely estimated	GMACS changed form
Fishery selectivity	Observer data	Freely estimated	GMACS changed form
Survey selectivity	BSFRF	Logistic, BSFRF as survey	Non-parametric, BSFRF as priors

Table 1. Population processes modeled in the stock assessment for snow crab, the data that informs each process, historical assumptions about a given process, and updated assumptions included in the most recent assessment model.

The assumptions about how to represent these processes within the model can impact inferred population scale and/or dynamics. Over the past several years, assumptions for a majority of the processes in the assessment have been updated to reflect new information, improve interpretation of existing information, or correct undesirable model behavior. Growth curves were changed from piece-wise linear models to linear models to address bimodality in management quantities. Blocks of natural mortality were added in 2018 and 2019 to address an abrupt population collapse attributed to a marine heatwave. Sex ratios of recruitment were allowed to vary over time to address retrospective patterns arising from forcing male recruitment to equal female recruitment.

Changing assumptions about the probability of having undergone terminal molt has had the largest impact on management quantities in the recent past. Prior to 2023, the probability of having undergone terminal molt was a time-invariant function of size. An ogive was freely estimated with a smoothing penalty applied for each sex that determined what fraction of the numbers of immature crab at size moved to the mature numbers at size after molting. The estimates of the probability of having undergone terminal molt at size for males were substantially smaller for mid-sized crab than the observed probabilities (figure 3). For years in which chelae height are available, observed probabilities of having undergone terminal molt are calculated by estimating a chelae height for each size class above which a crab is mature (see Richar and Foy (2022) for a complete description). These probabilities are directly used each year to divide new shell male crab data into mature and immature components, which are then used directly in the assessment. The inability of the assessment to estimate at least the median value of the probability of having undergone terminal molt at size over time suggests that misspecification exist elsewhere in the model that need to be identified. Until that misspecification can be located, the decision to use the observed probability of terminal molt data in the assessment stemmed from: 1) the importance of terminal molt in determining growth through the size classes, 2) the inability of the model to produce estimates of the probability of terminally molting similar to the observations and 3) the probability of having undergone terminal molt is one of the few processes for which direct measurements are available (growth increments being the other). It is difficult to defend continuing to use the estimated ogive when it directly conflicts with the data collected in the survey each year.

Another important change to the assessment model in 2023 was the shift from using the BSFRF data as an additional survey to using the data as direct priors on selectivity at size. This necessitated a shift from an assumed logistic shape of the selectivity curve (figure 4) to a non-parametric (i.e. more flexible) curve (figure 5). If survey selectivity was not allowed more flexibility, the estimated morphometrically mature male biomass increased dramatically when using the observed probabilities of having undergone terminal molt (figure 6). However, if survey selectivity was allowed more flexibility, the estimated morphometrically mature male biomass was similar in scale to historical estimates. This is because the medium selectivity for mid-sized crab and higher selectivity for large crab (a result that cannot be obtained easily with a logistic curve) provided some balance to the effect of more crab maturing at smaller sizes.

The change to probability of having undergone terminal molt and treating the BSFRF data as priors are both improvements to the interpretation of the available data, but also come with large knock-on effects. One of the most dramatic effects of replacing the estimated probability of having undergone terminal molt ogive with the observed values was an increase in the fishing mortality reference points. Fully selected target exploitation rates moved from 75-85% to ~100% (Table 2). This would allow for the complete

removal of the males >101 mm carapace width (discussed more below). The potential rationality of very high exploitation rates on large crab was discussed at Crab Plan Team meetings, centered on the fact that smaller male crab have been observed in the lab mating with larger female crab (Watson, 1972) and other crab fisheries have persisted for decades under very high harvest rates (e.g. Dungeness crab, Richerson et al., 2016). Counterpoints were centered on only larger males being observed in reproductive embraces in the wild (Conan and Comeau, 1986) and differences in life history (particularly longevity) between snow crab and Dungeness crab. Ultimately, the CPT and SSC were uncomfortable with reference points that would allow the complete removal of large male crab, so the analyses described below were undertaken to explore other options for the currency of management used for eastern Bering Sea snow crab.

Updating parameters associated with selectivity and maturity resulted in changes in the fits to the data. Model 23.1 (which estimated the probability of having undergone terminal molt and assumed logistic selectivity) fit the data better than the model that specified the probability of having undergone terminal molt and estimated a non-parametric selectivity curve with the BSFRF data as priors (Table 3). Poorer fits when changing the functional forms of survey selectivity and changing the shape of the probability of having undergone terminal molt suggests one of two things. Either the data we have on survey selectivity and probability of having undergone terminal molt do not reflect the underlying reality or there are other processes mis-specified in the model. Given these two processes are one of three that we have direct observations for (growth being the other), it is likely that other processes are mis-specified. Time- and/or size-varying natural mortality

Model	MMB	B35	F35	FOFL	OFL	M	avg_rec	Status
23.1	56.41	189.24	1.60	0.30	8.58	0.29	169.90	0.30
23.2	135.43	132.46	71.89	30.14	37.10	0.29	222.75	1.02
23.3a	92.39	155.91	53.25	14.96	15.44	0.29	141.66	0.59

Table 2. Management quantities from selected models in the September 2023 SAFE document. Model 23.1 was the status quo model in which maturity was estimated and survey selectivity was logistic, 23.2 specified the probability of maturity, 23.3a was 23.2 + estimating survey selectivity with BSFRF data as priors. The negative log likelihoods for the models were: -23714, -25638, and -25155, respectively.

Component	Description	23.1	23.2	23.3a	23.1 - 23.3a
catch	Retained	-7.08	-4.15	2.92	-10
catch	Discard (male)	140.7	130.44	88.95	51.75
catch	Discard (female)	-69.66	-69.66	-69.66	0
catch	Trawl	-52.03	-52.03	-52.02	-0.01
cpue	NMFS survey (era 1; females)	43.44	54.65	53.59	-10.15
cpue	NMFS survey (era 2, females)	-28.82	-13.3	-2.38	-26.44
cpue	NMFS survey (era 1, males)	32.15	35.28	46.62	-14.47
cpue	NMFS survey (era 2, males)	21.33	-0.98	31.23	-9.9
growth_inc	1	1020.3	1061.75	1038.86	-18.56
rec_dev	3	93.91	79.3	80.87	13.04
size_comp	Retained males	-3702.08	-3608.48	-3641.29	-60.79
size_comp	Survey mature females (1982-1988)	-688.42	-685.48	-678.99	-9.43
size_comp	Survey mature females (1989-present)	-3168.88	-3056.79	-3198.41	29.53
size_comp	Survey mature males (1982-1988)	-595.71	-581.17	-582.62	-13.09
size_comp	Survey mature males (1989-present)	-2828.43	-2694.15	-2800.92	-27.51
size_comp	Total males	-2708.96	-2630.54	-2638.41	-70.55
size_comp	Discard females	-2283.93	-2275.2	-2270.88	-13.05
size_comp	Non-directed bycatch (females)	-2539.63	-2481.05	-2460.89	-78.74
size_comp	Non-directed bycatch (male)	-2435.53	-2426.65	-2358.69	-76.84
size_comp	Survey immature females (1982-1988)	-624.89	-624.2	-613.95	-10.94
size_comp	Survey immature females (1989-present)	-2966.22	-2999.83	-3045.46	79.24
size_comp	Survey immature males (1982-1988)	-577.99	-569.18	-541.84	-36.15
size_comp	Survey immature males (1989-present)	-2828.2	-2815.83	-2791.83	-36.37
	Sums	-26754.6	-26227.3	-26405.2	-349.43

Table 3. Likelihood components for data sources for models listed in table 2, reproduced from September 2023 SAFE document . Last column is the difference between model 23.1 and 23.3a (i.e. specifying maturity parameters and estimating non-parametric survey selectivity with BSFRF as priors).

Exploring possibilities for alternate currencies of management

Reference point background
The Fisheries Management Plan (FMP) for snow crab in the Bering sea identifies the Council's management goal as maximizing 'the overall long-term benefit to the nation' and the historical methods used to pursue this goal for snow crab are harvest control rules based on proxies for maximum sustainable yield (Clark, 1991). The concept of maximum sustainable yield is predicated on the idea that there is a relationship between spawning biomass and recruitment. However, stock recruitment relationships are notoriously difficult to detect and can be noisy when observed (Szuwalski et al., 2015). Consequently, recruitment dynamics are a key uncertainty in stock assessment models and reference points derived from them. Proxies for maximum sustainable yield based reference points for fishing mortality and biomass were developed because of the difficulty associated with estimating stock recruit relationships. Clark (1991) produced yield curves over a range of steepnesses for a given life history, scaled both of the axes of the yield curve (biomass and yield) to have a maximum of 1, then identified the level of fishing that maximized the minimum yield ("maximin yield") across stock recruit relationships. In the example life history that Clark (1991) used, a fishing mortality rate that reduced the biomass to 35% of unfished biomass produced the maximin solution. Three quantities are required in the harvest control rule to set the overfishing level (OFL), which then is used to determine acceptable biological catch (ABC) each year: an estimate of biomass, a target biomass, and a target fishing mortality. The currently used targets for biomass and fishing mortality used for eastern Bering Sea snow crab are the SBPR proxies described in Clark's analysis.

Recent updates to the modeling of snow crab have resulted in these SBPR proxies not providing reasonable limits on the fishing mortality imposed by the fishery (Szuwalski, 2023). Part of this issue stems from the difference in size between the crab harvested in the fishery and the size of morphometrically mature crab. A male crab can mature at a size much smaller than that harvested by the fishery, so a large fraction of the morphometrically mature male biomass can be protected from the fishery (Figure 7). In some cases, essentially all the commercially preferred males can be removed from the population and the management targets for mature biomass can still be reached. The estimated $F_{35\%}$ for snow crab in 2023 when morphometrically mature male biomass was used as the currency of management was >300, which translates to an exploitation rate of 99.999% (Szuwalski, 2023). Clark (1991) was based on life history in which maturity was equal to fishery selectivity. Punt et al. (2015) revisited this work using population dynamics models and life histories specific to terminally molting crab and found that these proxies were roughly appropriate for snow crab under the existing modeling assumptions, which included estimating the probability of having undergone terminal molt. However, the observed probabilities used in the 2023 assessment are dramatically different than those used in Punt et al (2015), so it is unclear if the currently used proxies for management are appropriate given the differences in life history.

The definition of 'mature' is a key assumption in the management of snow crab is a key issue in the search for appropriate reference points. The current metric used for 'mature biomass' is the morphometrically mature male biomass (i.e. those crab with a large claw). Snow crab undergo a terminal molt, after which they no longer grow (Dionne, 2004). Any given crab can be distinguished as 'morphometrically' mature based on the relationship between the size of its claw and its carapace width (Richar and Foy, 2022). Whether or not all morphometrically mature crab are important in reproduction is an open question. In laboratory settings, a 60.8 mm carapace width male has been observed to mate with a 73.7 mm carapace

width female, suggesting small males can be important in reproductive dynamics (Watson, 1972). But in settings where large and small males had to compete for females, the large males always out competed smaller males (Comeau et al., 1998). Furthermore, in the few in situ observations during the mating season reported in the literature, only males larger than 96 mm carapace width were observed to grasp mature females in a mating embrace (Conan and Comeau, 1986). This lead some researchers to define 'functional' maturity as the size at which mature males are observed to actively partake in mating (>95 mm carapace width in this instance). Other regions managing snow crab populations also use this figure as their mature biomass (e.g. Gulf of Saint Laurence).

Here, I explore the appropriateness of morphometrically mature male biomass as the currency of management for EBS snow crab given recent model changes, quantify the impacts of changing the currency of management to range of size-based cutoffs on the perceived status of the stock and resulting overfishing levels (OFLs), and evaluate the currently used percentage (35%) for spawning-biomass-per-recruit proxies under a range of currencies of management and underlying population dynamics. I conclude with options modifying the management strategy for snow crab.

Population dynamics models

Two population dynamics models are used to explore these questions: the currently used assessment model and a research model. The currently used assessment model is the Generalized Model for Assessing Crustacean Stocks (GMACS) and the research model is similar to the GMACS model. The research model is fit to much of the same data as the GMACS model, but is simplified to more easily make the modifications required to answer our research question. Key differences include: the research model only considers male crab (GMACS includes females), excludes the bycatch fishery (which imposes a very small estimated mortality to the population each year; table 4), specifies the size transition matrix, and fits to an index of immature abundance.

Data component in GMACS	**Years**	Fit in RM?	Inform RM?
Retained male crab pot fishery size frequency by shell condition	1982 - 2022	X	X
Discarded Males and female crab pot fishery size frequency	1992 - 2022	X	X
Trawl fishery bycatch size frequencies by sex	1991 - 2022		
Survey size frequencies by, maturity, sex and shell condition	1982 – 2019 2021 - 2023	X	X
Retained catch estimates	1982 - 2022	X	X
Discard catch estimates from crab pot fishery	1992 - 2022	X	X
Trawl bycatch estimates	1993 - 2022		
Total survey abundance estimates and coefficients of variation	1982 - 2019, 2021 - 2023	X	X
2009 study area biomass estimates, CVs, and size frequency for BSFRF and NMFS tows	2009		X
2010 study area biomass estimates, CVs, and size frequency for BSFRF and NMFS tows	2010		X
Growth increment data	2003, 2016-18		X

Table 4. Data components in GMACS (left column), temporal availability (middle), and presence in research model (right).

The population dynamics for both models presented here tracks numbers of crab at size s by maturity state m over time t ($N_{t,s,m}$) with size bins ranging from 30-135 mm carapace width with 5 mm bin widths. The research model estimates natural mortality (M) by year y and maturity state m and fishing mortality (F) by year y. Other estimated parameters in the research model 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. Parameters determining growth and maturity are estimated outside of the model. Mortality is the only population process that occurs in the first seven months of a given year (the crab year begins in July with the bottom trawl survey):

$$N_{t=y+0.59,s,m} = N_{t=y,s,m} \, e^{-\frac{7}{12}M_{t,s,m}}$$

Fishing occurs as a pulse fishery on February 15 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 $C_{ret,y}$ and what fraction are discarded back into the ocean. A discard mortality d_{mort} of 25% is applied to the crab returned to the ocean.

$$C_{cap,y} = n_{t=y+0.59,s,m}1 - e^{-F_t*S_{cap}}$$

$$C_{ret,y} = C_{cap,y}S_{ret}$$

$$n_{t=y+0.59,s,m} = n_{t=y+0.59,s,m}e^{-F_tS_{cap}}$$

$$n_{t=y+0.59,s,m} = n_{t=y+0.59,s,m} + d_{mort}C_{cap,y}(1 - S_{ret})$$

Growth occurs for immature crab after the fishery 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. Snow crab undergo a 'terminal molt' to maturity after which growth ceases (Tamone et al., 2005). Accordingly, all immature crab are assumed to molt and no mature crab molt in the model. Newly molted crab are assigned to a maturity state based on observed ogives of the proportion of mature new shell males by size calculated from chelae height measured in the NMFS survey data, which varies over time $\rho_{y,s}$; figure 3). The average probability of having undergone terminal molt is used in years during which data were not collected. This process results in two temporary vectors of numbers at size:

$$n_{t=y+0.59,s,m=1} = (1 - \rho_{y,s})X_{s,s'}n_{t=y+0.59,s,m=1}$$

$$n_{t=y+0.59,s,m=2} = \rho_{y,s}X_{s,s'}n_{t=y+0.59,s,m=1} + n_{t=y+0.59,s,m=2}$$

The size transition matrix $X_{s,s'}$ was constructed using growth increment data collected over several years (see Szuwalski [2023] for a summary) to estimate a linear relationship between pre- and post-molt carapace

width (figure 1), (W_s and $\bar{W}_{s'}$, respectively). The variability around that relationship was characterized by a discretized and renormalized normal distribution, $Y_{s,s'}$, where s and s' represent entries in the rows and columns of the matrix. M_s is the midpoint of size bin s and σ_{grow} is a standard deviation around the growth increment specified based on observed variability in post-molt sizes.

$$X_{s,s'} = \frac{Y_{s,s'}}{\sum_{s'} Y_{s,s'}}$$

$$Y_{s,s'} = \frac{(\bar{W}_{s'} - M_s)^2}{2\sigma_{grow}^2}$$

$$\bar{W}_{s'} = \alpha + \beta_1 W_s$$

Recruitment by year was estimated as a vector of deviations τ_y around an estimated mean μ in log space and added to the first three size of classes of immature crab based on another estimated vector Δ_s with bounds 0-200 that determines the proportion allocated to each size bin φ_s.

$$tot_d = 20 + \Delta_1 + \Delta_2$$

$$\varphi_1 = \frac{20}{tot_d}$$

$$\varphi_{2,3} = \frac{\Delta_{1,2}}{tot_d}$$

$$n_{t=y+0.59,s=1-3,m=1} = \varphi_s e^{\mu+\tau_y}$$

For projections across uncertainty in recruitment, a spawner recruit relationship was coded based on the 'steepness' parameterization of the Beverton Holt curve, in which steepness h (i.e. the proportion of unfished recruitment supplied when spawning biomass is depleted to 20% of unfished levels) and unfished recruitment R_0 must be specified.

Finally, the last half of the year of natural 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.

$$N_{t=y+1,s,m} = n_{t=y+0.59,s,m}\, e^{-\frac{5}{12}M_{t,s,m}}$$

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 by fitting to 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 ($\widehat{N}_{y,s,NMFS}$ and $\widehat{N}_{y,s,BSFRF}$, respectively) can be used to generate priors for the selectivity of the NMFS gear in year y $S_{y,NMFS}$ as:

$$S_{y,NMFS} = \frac{\widehat{N}_{y,s,NMFS}}{\widehat{N}_{y,s,BSFRF}}$$

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 (figure 8). 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 (figure 9). A generalized additive model was fit to the inferred selectivity data to provide a prior mean and standard deviation of NMFS survey selectivity at size to be fit to in the population dynamics model.

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 prior 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 for immature and mature crab caught in the survey and retained and discarded crab in the fishery were fit using multinomial likelihoods and were implemented in the form:

$$L_x = \lambda_x \sum_y N_{x,y} \sum_l p_{x,y,l}^{obs} \ln \left(\frac{p_{x,y,l}^{pred}}{p_{x,y,l}^{obs}}\right)$$

L_x was the likelihood associated with data component x, where λ_x represented an optional additional weighting factor for the likelihood, $N_{x,y}$ was the sample sizes for the likelihood, $p_{x,y,l}^{obs}$ was the observed proportion in size bin l during year y for data component x, and $p_{x,y,l}^{pred}$ was the predicted proportion in size bin l during year y for data component x. Sample sizes were input as 150.

Observed indices of abundance for immature and mature males in the survey and retained and discarded catches in numbers in the fishery were fit with log normal likelihoods implemented in the form:

$$L_x = \lambda_x \sum_y \frac{(\ln(\hat{I}_{l,x,y}) - \ln(I_{l,x,y}))^2}{2\sigma_{x,y}^2}$$

L_x was the contribution to the objective function of data component x, λ_x was any additional weighting applied to the component, $\hat{I}_{l,x,y}$ was the predicted value of quantity l from data component x during year y, $I_{l,x,y}$ was the observed value of quantity I from data component x during year y and $\sigma^2_{x,y}$ was the observed or specified variance for data component x during year y.

Smoothing penalties were placed on estimated vectors of deviations for immature and mature natural mortality using normal likelihoods on the second differences of the vectors. Normal priors were also placed on the mean value of natural mortality and the deviation of the estimated mortality from that mean. A prior value of 0.27 is used for the average natural mortality based on an assumed maximum age of 20 and Hamel's (2015) empirical analysis of life history correlates with natural mortality. The priors used for survey selectivity were derived from the selectivity experiments described above. The normal priors were of the form:

$$P_x = \lambda_x \sum_y \frac{(\ln(\hat{I}_{l,x,y}) - \ln(I_{l,x,y}))^2}{2\sigma^2_{x,y}}$$

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, $\hat{I}_{l,x,y}$ was the predicted value of population process I relevant to penalty x during year y, $I_{l,x,y}$ was the prior value of process I relevant to penalty x during year y and $\sigma^2_{x,y}$ was the observed or specified variance for data component x during year y.

Model fits and some diagnostics for the research model can be seen in Appendix I. The GMACS model is similar to the research model and a more complete description can be found in Szuwalski (2023). The fits and diagnostics for the GMACS model presented here are identical to those presented in 2023 as no changes were made to the base model. One alteration was made to GMACS to allow for the replacement of morphometric maturity with a user-input functional maturity in calculation of management quantities (e.g. target fishing mortalities, target biomasses, and overfishing levels). This alteration does not change the fits to the data presented in September 2023.

Yield analyses and GMACS outputs with selected currencies of management

I repeated Clark's exercise with the research model described above with three definitions of mature male biomass: morphometric, >95 mm carapace width, and >100 mm carapace width. For each definition of mature male biomass, a steepness was specified (from 0.3 to 0.9 in increments of 0.1) and the population was fished for 100 years at a constant exploitation rate. This was performed for a wide range of exploitation rates for each definition of mature biomass and specified steepness, which then allowed the equilibrium yield as a function of biomass to be plotted. Both the raw yield curves and normalized yield curves are presented here.

When morphometric mature male biomass was used as the currency of management and steepness was high, the population could not be reduced to low levels of unfished biomass because a large amount of the morphometrically mature biomass was protected from the exploitation as a result of the fishery selectivity

(figure 10). Even under the lowest steepness considered (0.3), the mature male biomass could not be reduced to zero even under high fishing mortalities. The maximin solution was approximately $F_{50\%}$. Specifying mature male biomass as >95 mm carapace width resulted in maximin solutions that reduced biomass to ~28% of unfished levels (figure 11). Using >100 mm carapace width as the definition did not appreciably change the maximin solution, but the mature biomass could be reduced to lower levels than when >95 mm was used to denote maturity (figure 12).

A range of definitions of mature male biomass were also used to rerun the assessment model presented in September 2023. Target biomasses, target fishing mortalities, and corresponding overfishing levels were calculated based on the redefined mature male biomasses (Table 5). Target SBPR percentages were maintained at 35% across maturity scenarios. Different definitions of mature biomass produced different perceptions about the status of the stock (figure 13). The time series of estimated mature male biomass never declined beneath the MSST and the current status was 78% of $B_{35\%}$ when morphometric maturity based on the updated probabilities of having undergone terminal molt was used as the definition for 'mature'. The target fishing mortality, $F_{35\%}$, was 305.86, which would allow for complete removal of the commercially preferred crab. Increasing the definition of maturity to a cutline at 85 or 90 mm carapace width resulted in much smaller $F_{35\%}$, and stocks that declined beneath the minimum stock size thresholds (0.49 and 0.38). In spite of being beneath MSST, definitions of maturity at 85 and 90 mm carapace width were above the threshold for opening the fishery and small OFLs would have been recommended. Increasing the maturity cutoff to 95 mm carapace width and beyond resulted in federal fishery closures.

Maturity	MMB	B35	F35	FOFL	OFL	M	avg_rec	Status
Morphometric	128.51	165.03	305.86	167.34	41.78	0.29	155.67	0.78
85mm	51.27	103.91	8.29	2.57	8.90	0.29	155.67	0.49
90mm	34.83	92.12	4.31	0.93	4.59	0.29	155.67	0.38
95mm	20.96	80.44	2.48	0.00	0.06	0.29	155.67	0.26
100mm	11.76	67.97	1.59	0.00	0.06	0.29	155.67	0.17
105mm	7.32	54.14	1.12	0.00	0.06	0.29	155.67	0.14

Table 5. Management quantities from the 2023 stock assessment model (observed probabilities of having undergone terminal molt and BSFRF data treated as priors) for scenarios in which the mature biomass is defined as denoted in the column "Maturity". Measurements indicate the carapace width at which a cutline was drawn for inclusion in mature biomass.

Another potential way to address the problem of complete removal of large crab under a morphometric currency of management and SBPR$_{35\%}$ reference points is to adjust the percentage of unfished biomass set as a target while retaining morphometric maturity as the currency of management. The net result of increasing the target percentage is aiming to conserve more of the morphometrically mature biomass. So, similar to the results for changes in currency of management, as the target percent of unfished biomass to be conserved increases, the inferred status and OFL decrease and the probability of a fishery closure increases (Table 6). At 35% and 45%, the stock status is above the MSST and the target fishing mortalities are high (305 and 67, respectively). From 55% to 75%, the stock is beneath MSST, but above the closure threshold (figure 14). Target fishing mortalities range from 14.32 to 0.92. When the biomass target is 85%

of unfished morphometrically mature males, the fishery is closed (in spite of an overfished status slightly above the closure threshold—the projected status is the metric that opens or closes the fishery).

SBPR%	MMB	B_target	F_target	FOFL	OFL	M	avg_rec	Status
35%	128.51	165.03	305.86	167.34	41.78	0.29	155.67	0.78
45%	128.51	212.18	67.12	26.90	24.06	0.29	155.67	0.61
55%	128.51	259.34	14.32	4.41	11.63	0.29	155.67	0.50
65%	128.51	306.49	3.12	0.76	3.94	0.29	155.67	0.42
75%	128.51	353.64	0.92	0.18	1.14	0.29	155.67	0.36
85%	128.51	400.79	0.30	0.00	0.06	0.29	155.67	0.32

Table 6. Management quantities from the 2023 stock assessment model (observed probabilities of having undergone terminal molt and BSFRF data treated as priors) for scenarios in which the target percentage for biomass relative to unfished biomass were varied in the SBPR calculations (noted in the SBPR% column).

Conclusions

There are at least three potential paths forward for a currency of management for snow crab and reference points:
1. continue to estimate probability of terminal molt even though it doesn't match the observations used to apportion new shell crab into immature and mature indices used in the assessment
2. use a more biologically realistic model that specifies the probability of terminal molt to match observations, retain morphometrically mature biomass change currency of management, and then adjust the SBPR% with some concession for a cap on exploitation of large males
3. use a more biologically realistic model that specifies the probability of terminal molt to match observations, change currency of management to reflect fishery selectivity, retain SBPR$_{35\%}$ (or something close to the value identified from the yield analyses) as biomass target

Each of these paths have pros and cons. From a modeling perspective, ensuring that processes for which there are reliable data reflect those observations should be a priority, even if reflecting those observations come at a cost to model fit in the short term. Consequently, inputting the probabilities of having undergone terminal molt is the author-preferred action of those presented (#2 or #3 above). Another possibility for including the terminal molt data is inputting the observed probabilities of terminal molt as data and fitting to them as Buck does in the Tanner crab assessment. This would introduce many more parameters (21 size classes x 41 years of data x 2 sexes, if the yearly variability in maturity is modeled). This would require a revision to GMACS.

All that said, although estimating the probability of terminal molt results in outcomes that do not match what we think we know about the biology of the stock, continuing to estimate parameters associated with terminal molt (option #1 above) leads to larger OFLs and ABCs. Larger ABCs would be less constraining for the State process which may be a desirable outcome as other model issues are identified and addressed.

Although providing the State with more flexibility by setting higher ABCs may be operationally useful, the discrepancy between the Federal and State catch advice has been a source of confusion for stakeholders. The State's TAC is on average 40% of the ABC, and addressing this discrepancy could improve understanding of the interactions between the management systems.

Using the best available data to model snow crab biology is a relatively straightforward decision, but dealing with the knock-on effects on reference points is more difficult. Both changing the definition of maturity (e.g. from morphometric to >95 mm carapace width) and changing the targets for unfished biomass of morphometric maturity have similar outcomes. The size of OFLs and the probability of fishing scales in a similar way for both options. Identifying an appropriate percentage or size cutoff is difficult given a lack of in situ observations to inform the decision. The maximin SBPR% from the research model was ~55%, but it still had a high fully selected fishing mortality target (~14) which would still allow for removal of all commercially preferred crab. If changing the percentage used in SBPR targets is chosen, modifying GMACS to perform the yield analyses would be needed, which might change the SBPR% noted above. Choosing a currency of management more similar to that used by the State and reflects 'functional' maturity is likely the most straightforward path, but results in much smaller OFLs and ABCs, which could result in Federal fishery closures. A better understanding of how important large males are in reproduction and how the presence or absence of large males can impact the size at terminal molt could influence the selection of these points, but these data are difficult to collect.

References

Conan, GY and Comeau, M (1986) Functional maturity and terminal molt of male snow crab. Can J Aquat Fish Sci. 43: 1710-1719.

Comeau, M. Robichaud, G, Starr, M, Therriault, JC, Conan, GY (1998) Mating of snow crab in the fjord of bonne Bay, Newfoundland. Crustaceana 71(8): 925-941.

Richar, J and Foy, R (2022) A novel morphometry-based method for assessing maturity in male tanner crab, Chionoecetes bairdi. FACETS. https://doi.org/10.1139/facets-2021-006

Richerson, K, Punt, AE, Holland, DS. (2020) Nearly half a century of high, but sustainable exploitation in the Dungeness crab fishery. Fisheries Research. 226: 105528.

Clarke, WG, (1991) Groundfish exploitation rates based on life history parameters. Can J Fish Aquat Sci 48: 734-750.

Dionne, M., Sainte-Marie, B., Bourget, E., Gilbert, D. (2003) Distribution and habitat selection of early benthic stages of snow crab. MEPS 259: 117-128.

Szuwalski, C.S. 2023. 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 99501Punt et al 2015. Proxies good for crab?

Szuwalski, CS, Vert-Pre, KA, Punt, AE, Branch, TA, Hilborn, R. (2015) Examining common assumptions about recruitment: a meta-analysis or recruitment dynamics for worldwide marine fisheries. Fish and Fisheries 16(4): 633-648.

Tamone, S.L., M. Adams and J.M. Dutton. 2005. 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), p.166-171.

Hamel, O. 2015. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. ICES Journal of Marine Science. 72: 62-69.

Watson, J (1972) Mating behavior in the spider crab, Chionoecetes opilio. J fish Res Bd Canada 29: 447-449.
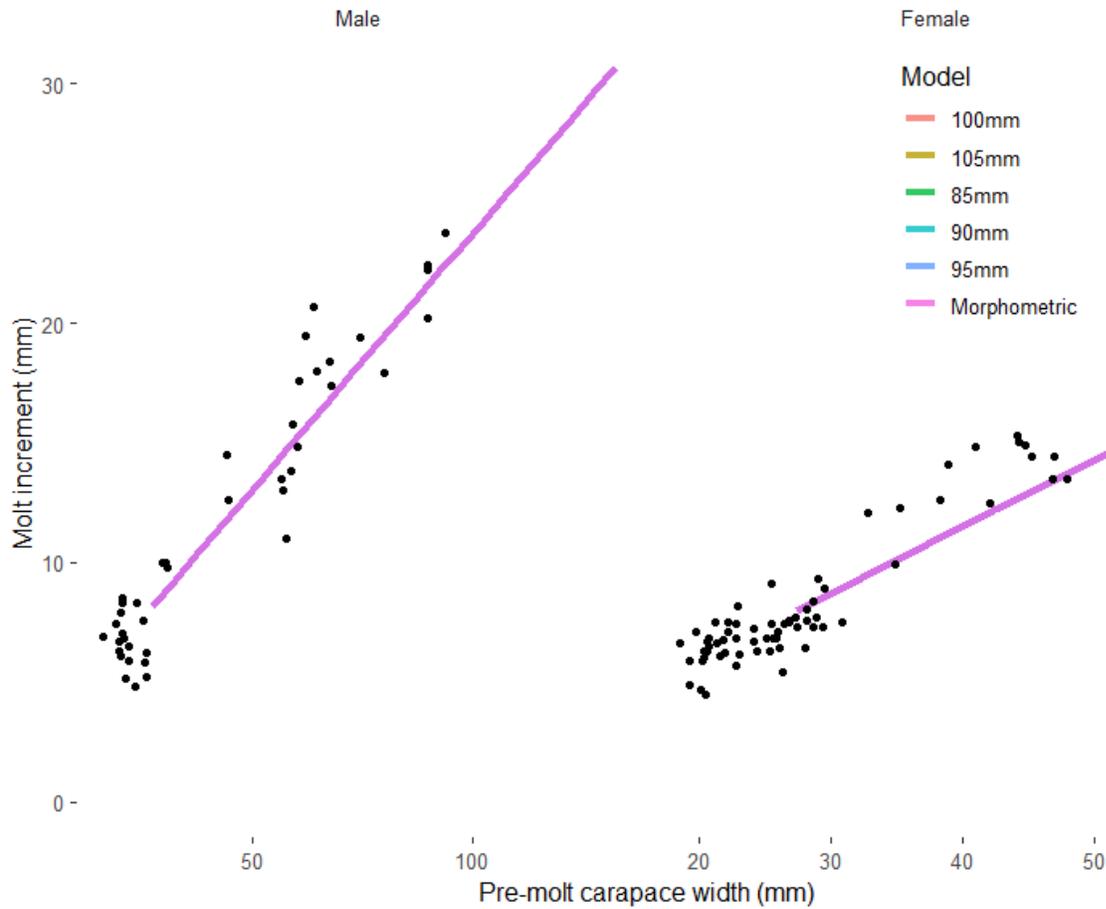
Figure 1. Growth data increment data included in the snow crab assessment for males (left) and females (right). Each black dot represents an observation from 1 crab, pre- and post-molt. Colored lines are model fits from the most recent assessment. Fits do not change as this was from a run exploring the impact of changing the currency of management on reference points. Model fits are only plotted over the range of sizes included in the model; all data are fit to within the model.
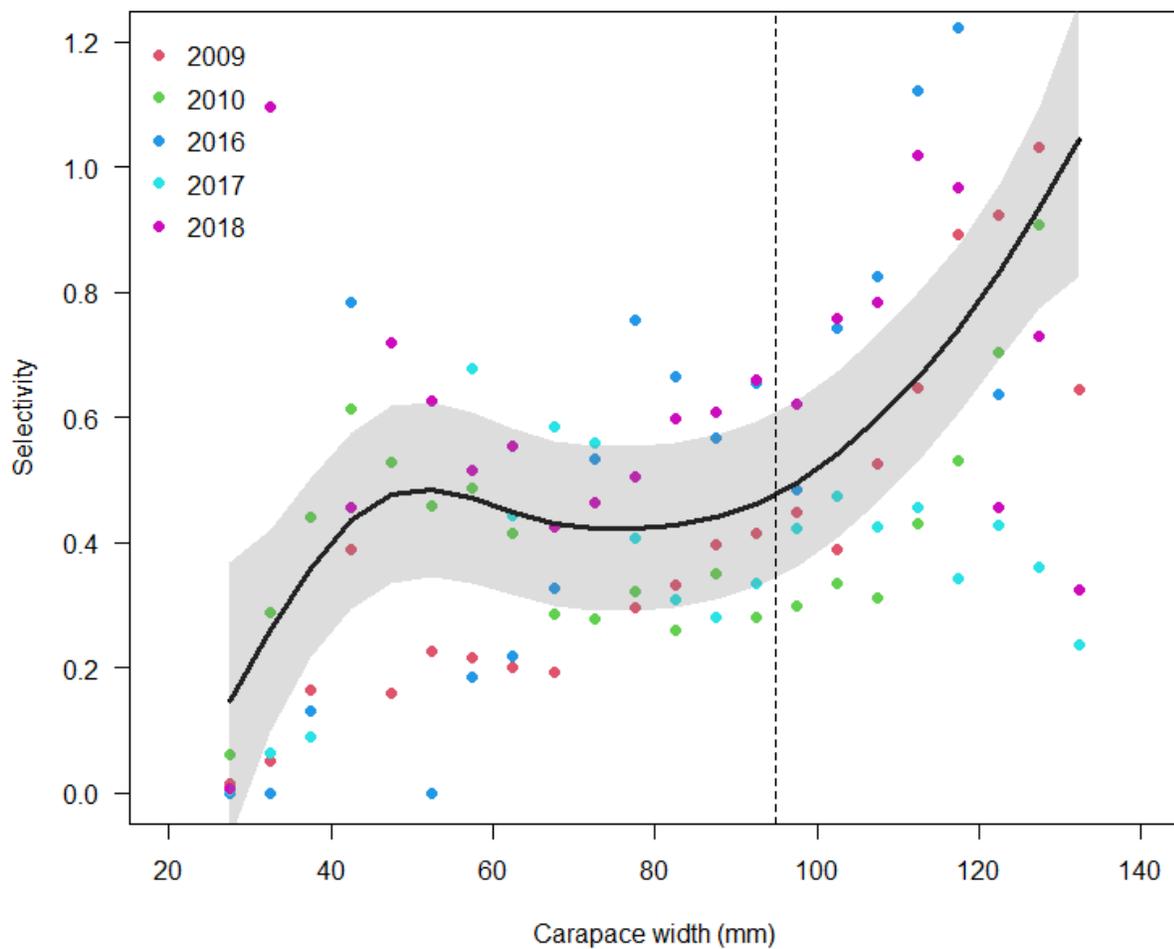
Figure 2. Inferred selectivity of NMFS gear from BSFRF experiments with fitted GAM. Each point is an inferred selectivity at size; color represents the year of the selectivity experiment. Black line is a GAM fit to the data and grey ribbon is the 95% confidence interval.
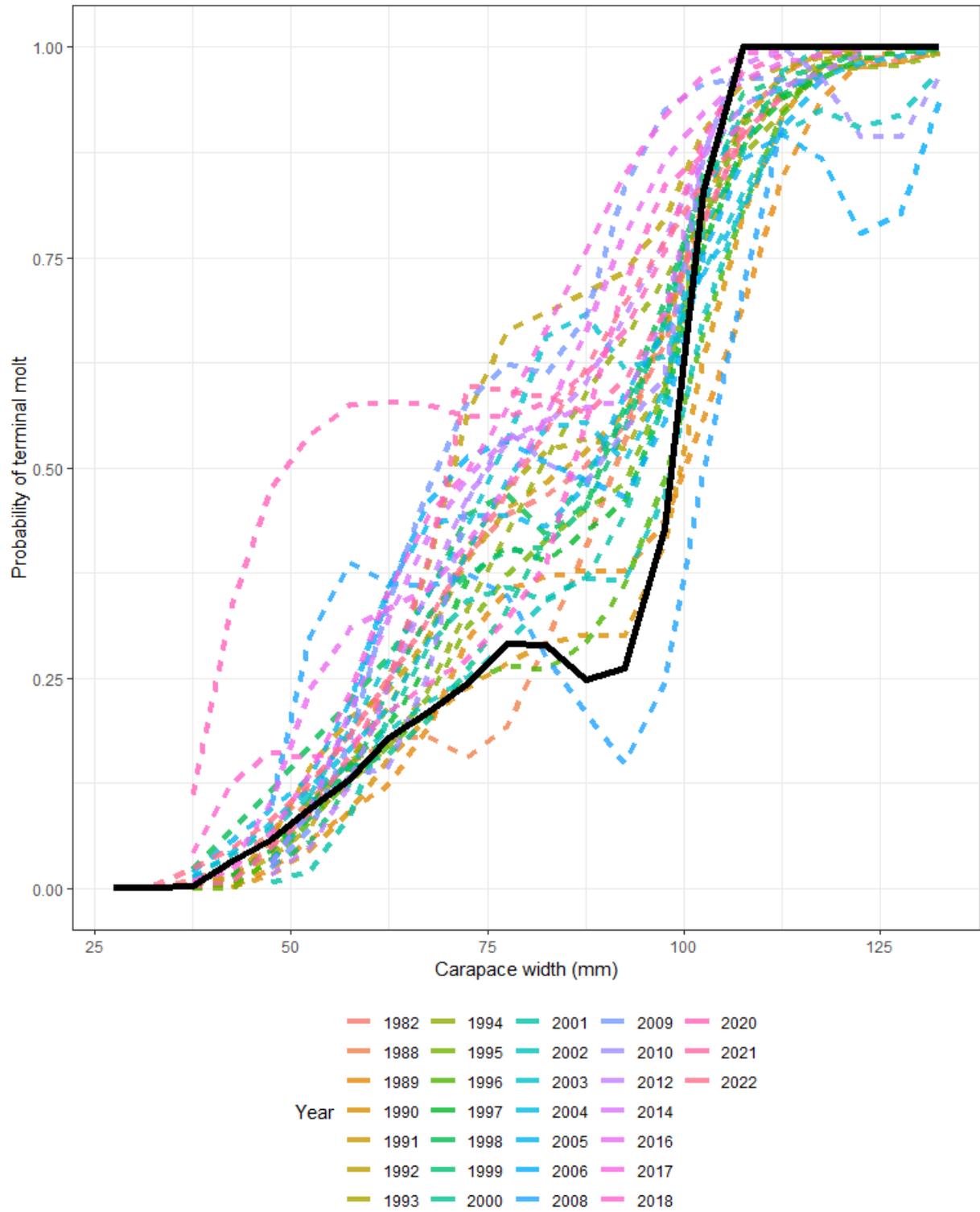
Figure 3. Estimated probability of having undergone terminal molt (black) vs. observed probabilities (colors) from the 2023 model.

s
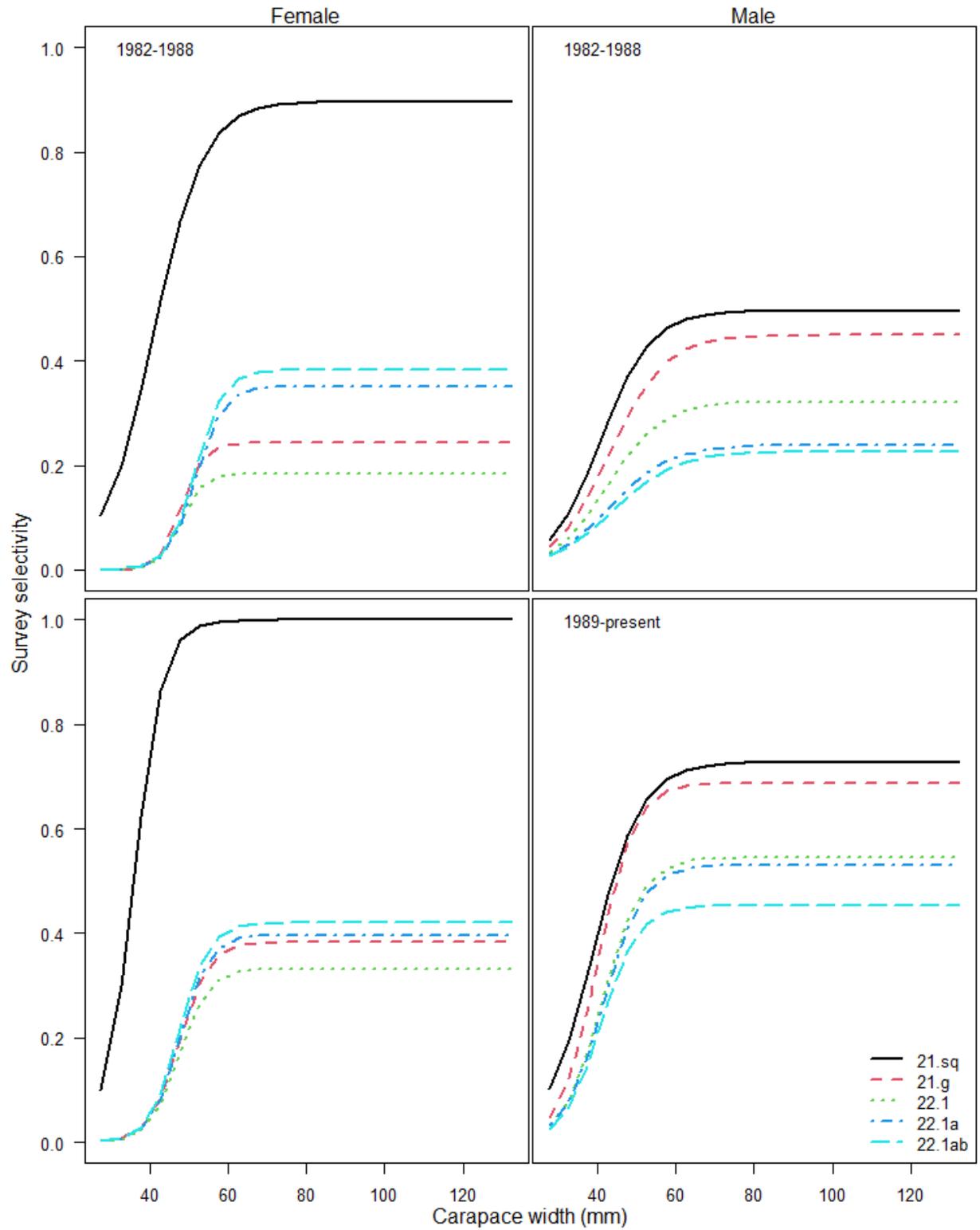
Figure 4. Estimated selectivities from 2022 assessment. The different models are unimportant, the point is to contrast the logistic shape with figure 5.
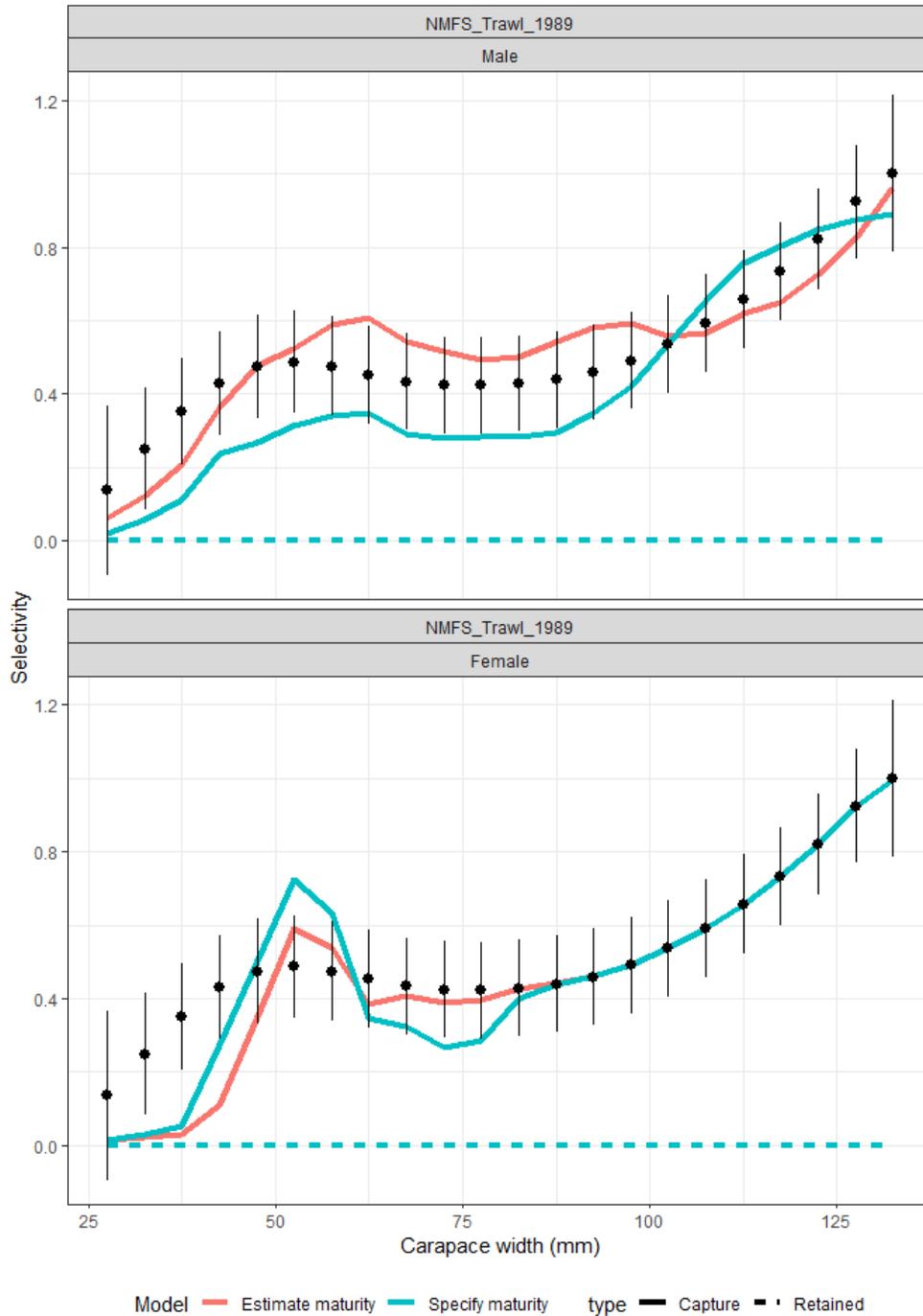
Figure 5. Estimated survey selectivity when BSFRF data are used as direct priors (black dots with input uncertainty) on selectivity at size. Blue lines come from a model in which the probability of having undergone terminal molt was specified at observed values; the probability of having matured was estimated for the red lines.
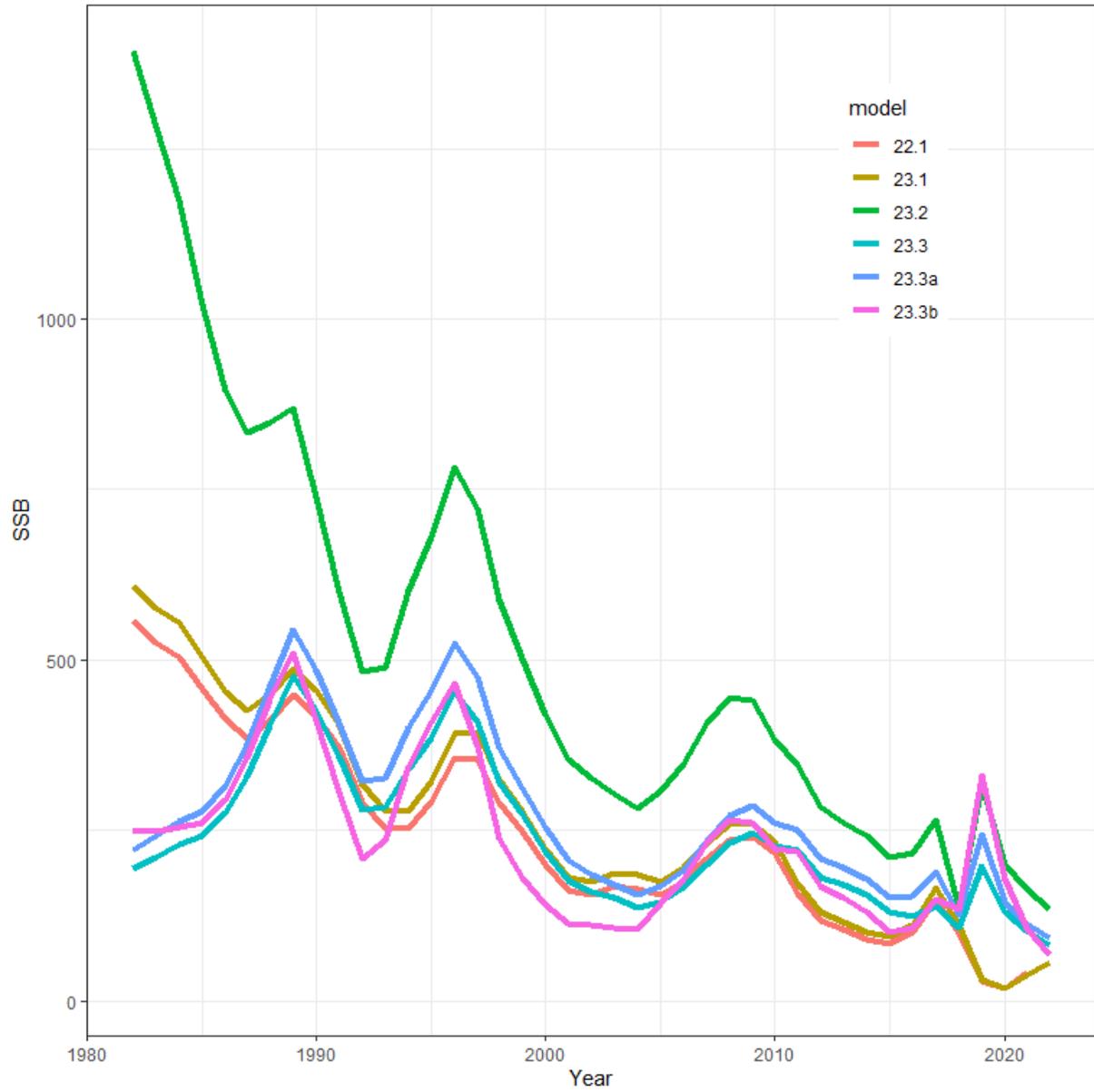
Figure 6. Estimates of mature male biomass from 2023 under different model scenarios outlined in table 2.
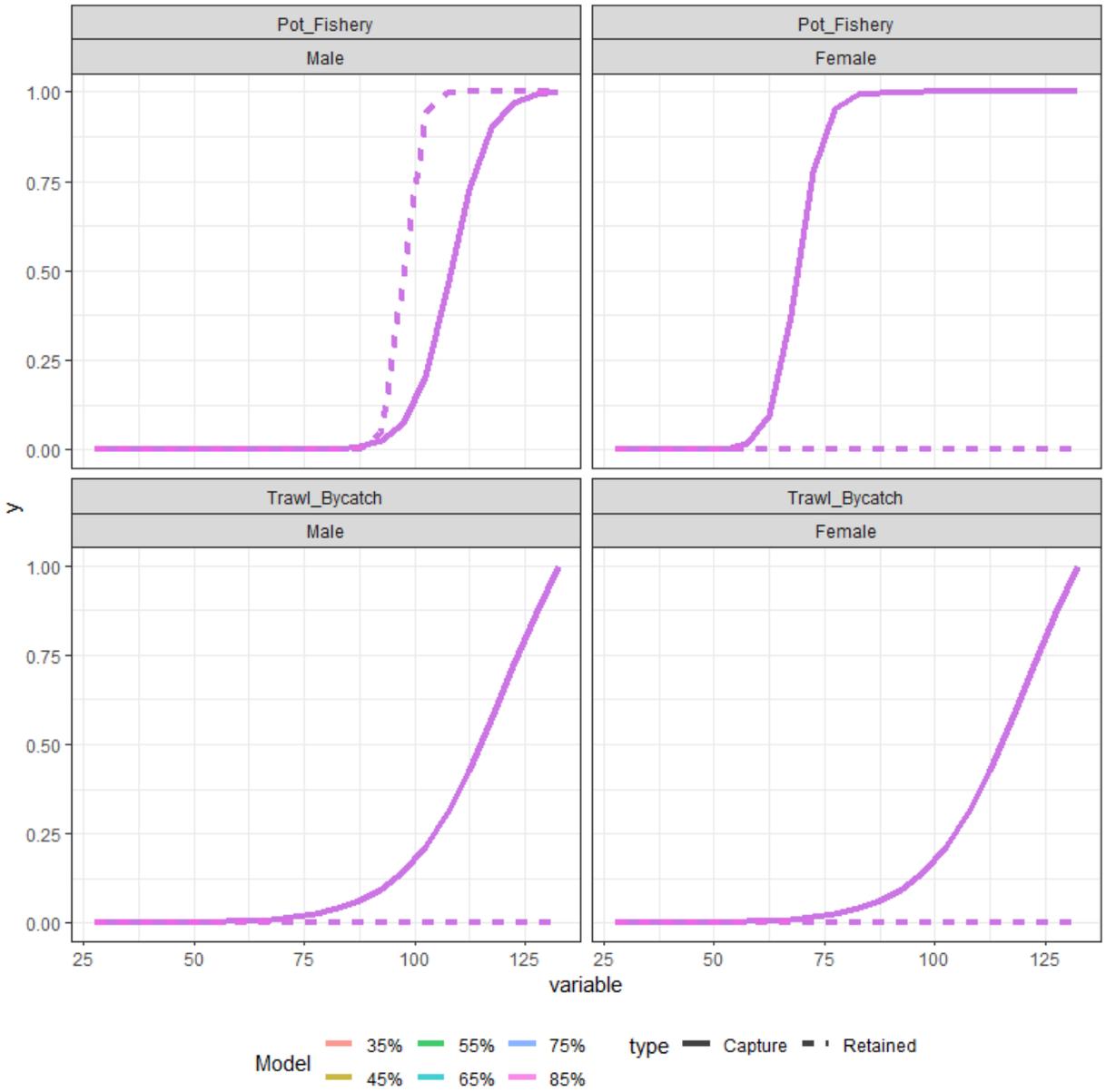
Figure 7. Estimated fishery selectivity for the 2023 assessment model. Estimates do not change among models because this output was from runs in which the percentage used in SBPR calculations was varied.
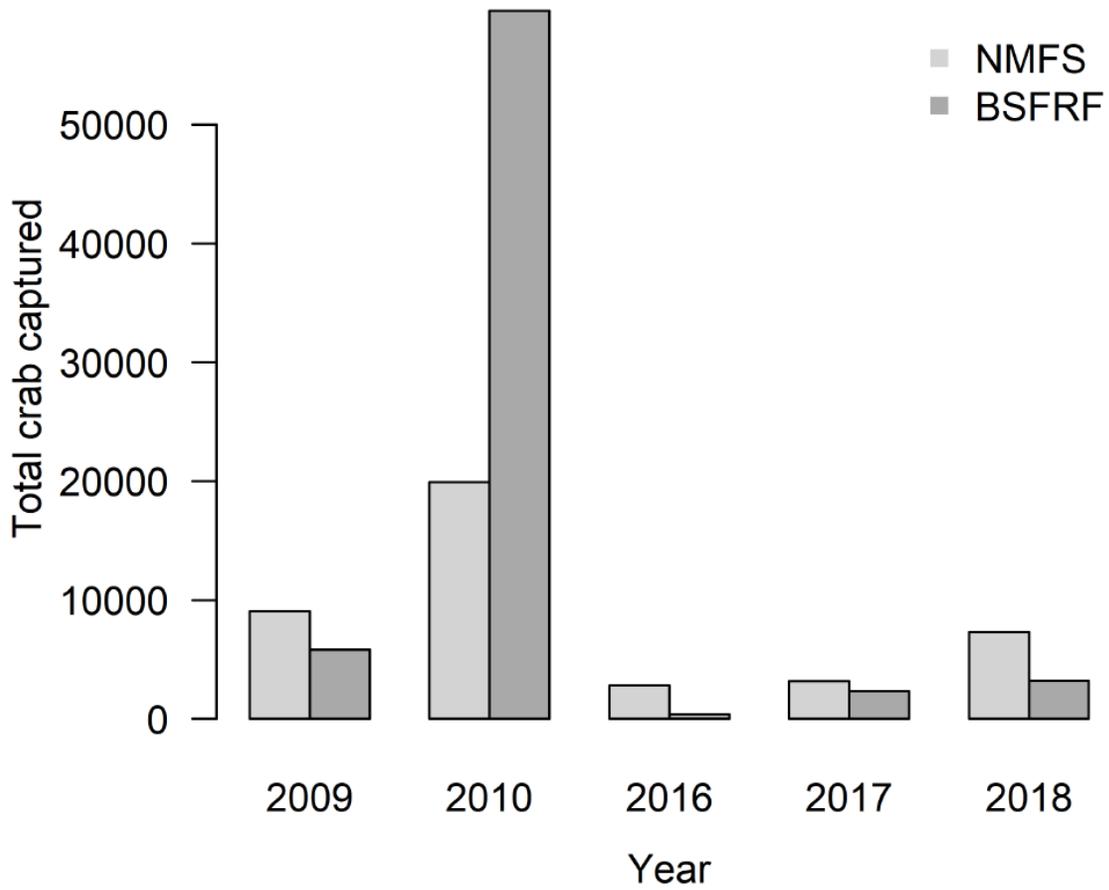
Figure 8. Sample sizes for the BSFRF survey selectivity experiments by year.
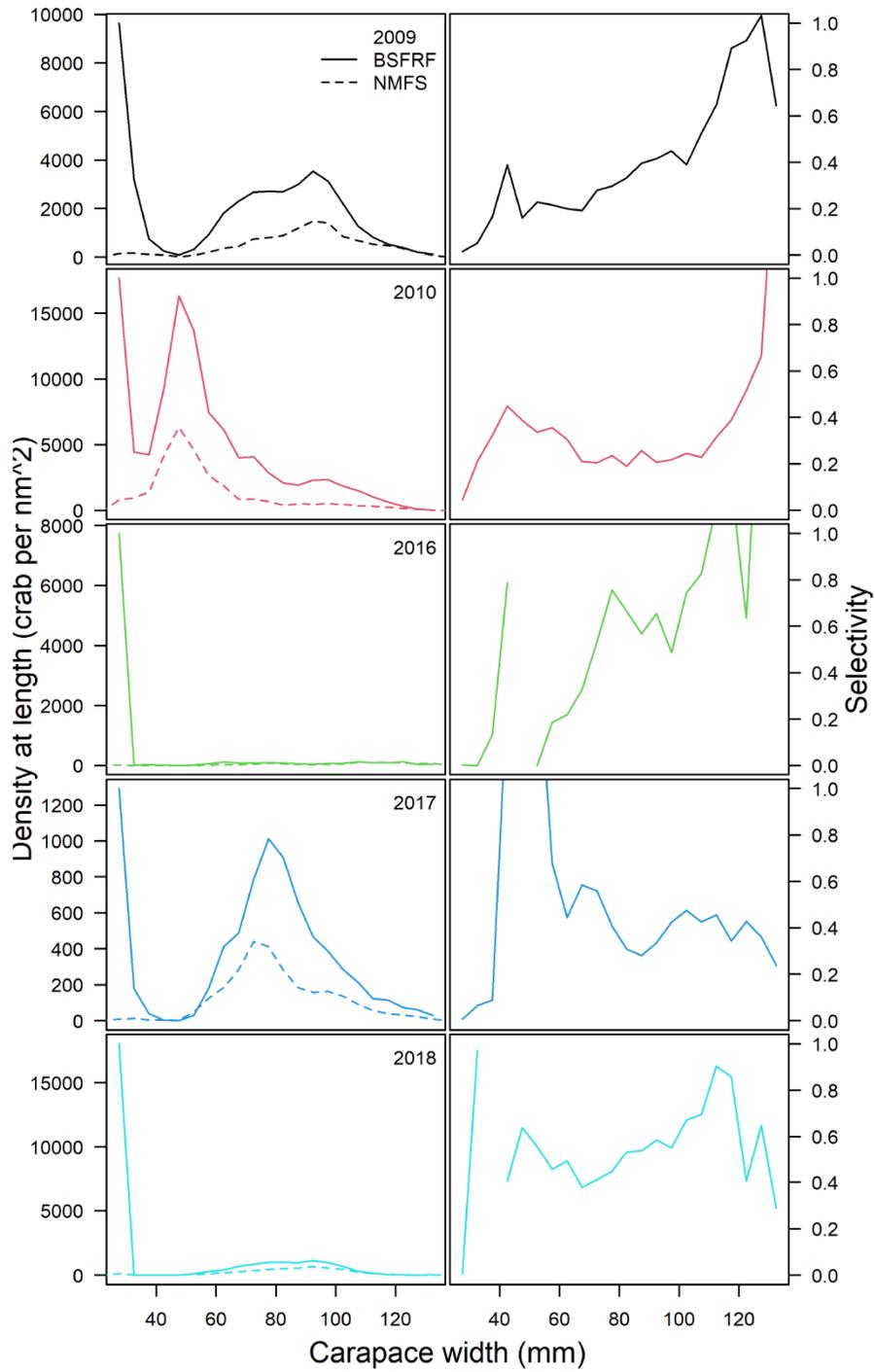
Figure 9. BSFRF sample sizes at size by year (left column) and inferred NMFS selectivity by year (right column).
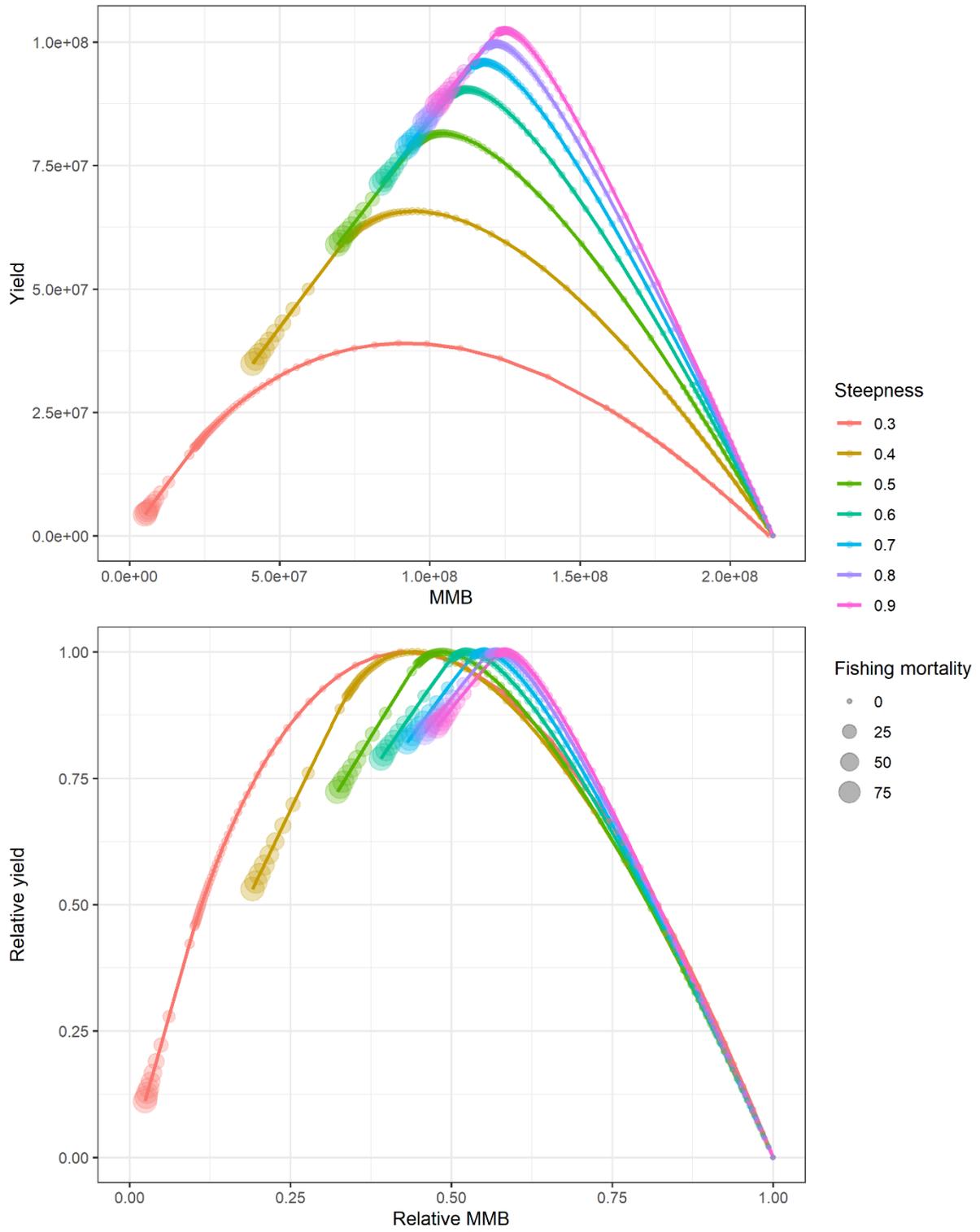
Figure 10. Clark's reanalysis using morphometrically mature biomass as the currency of management.
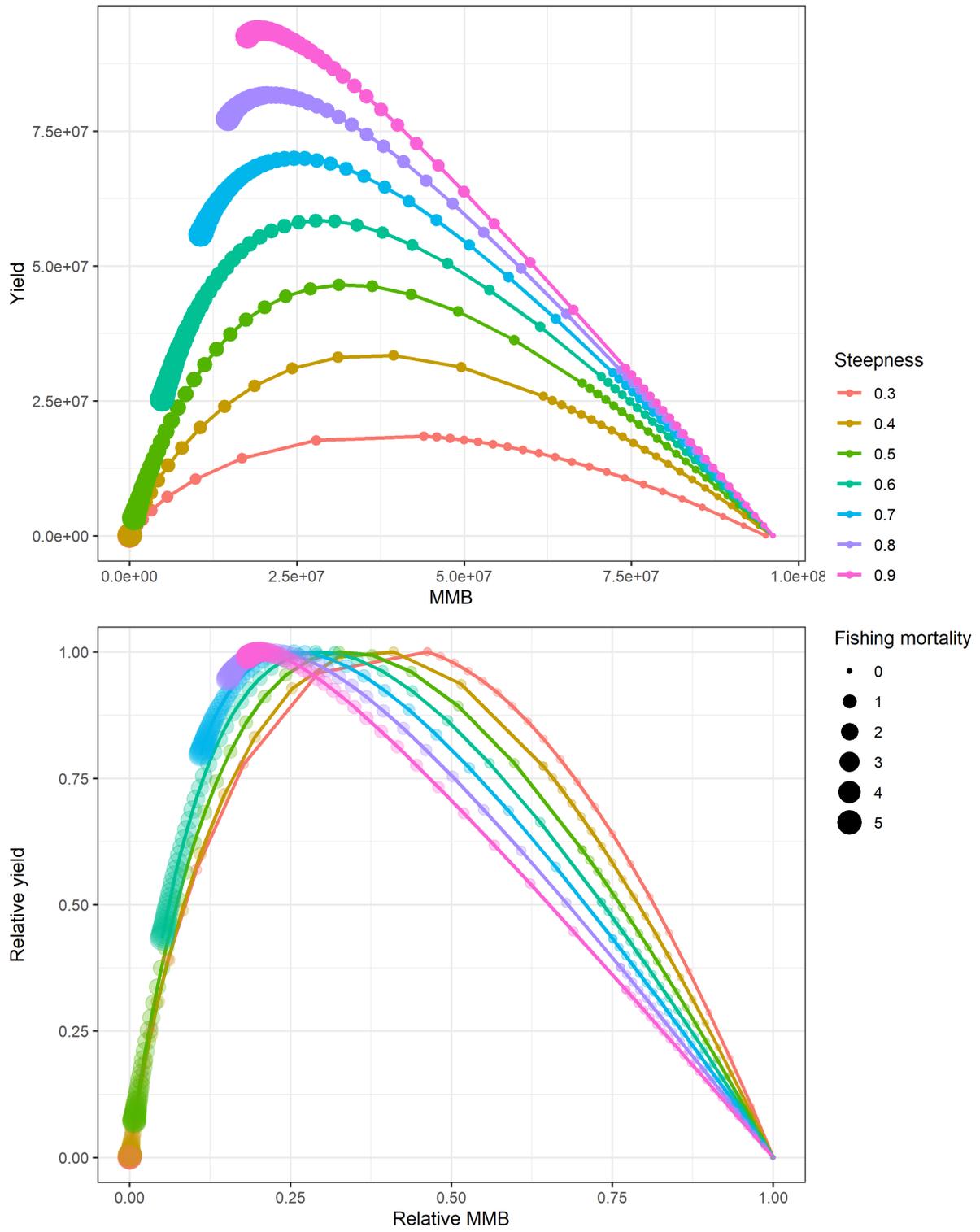
Figure 11. Clark's reanalysis using >95mm male crab as the currency of management.
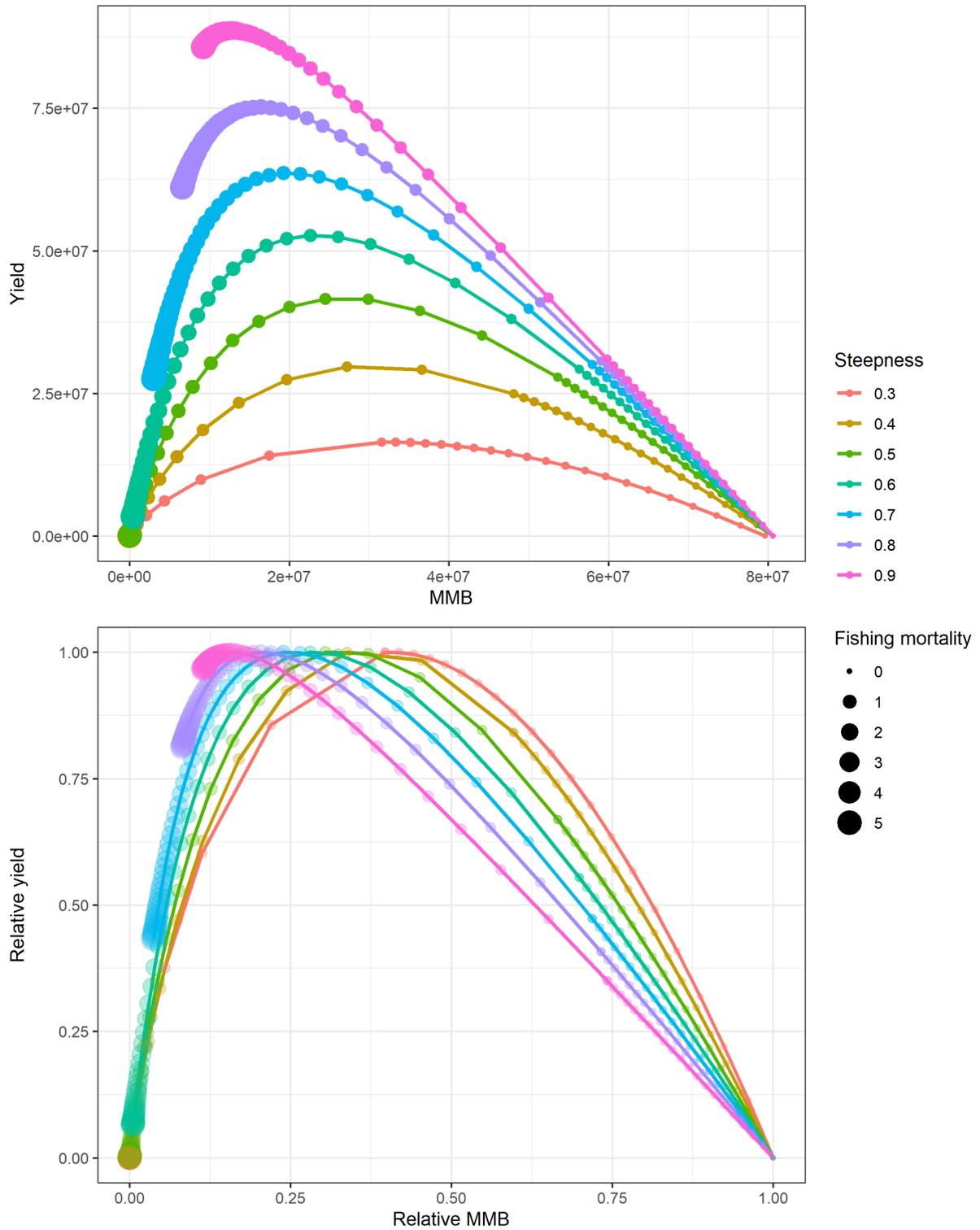
Figure 12. Clark's reanalysis using >100mm male crab as the currency of management.
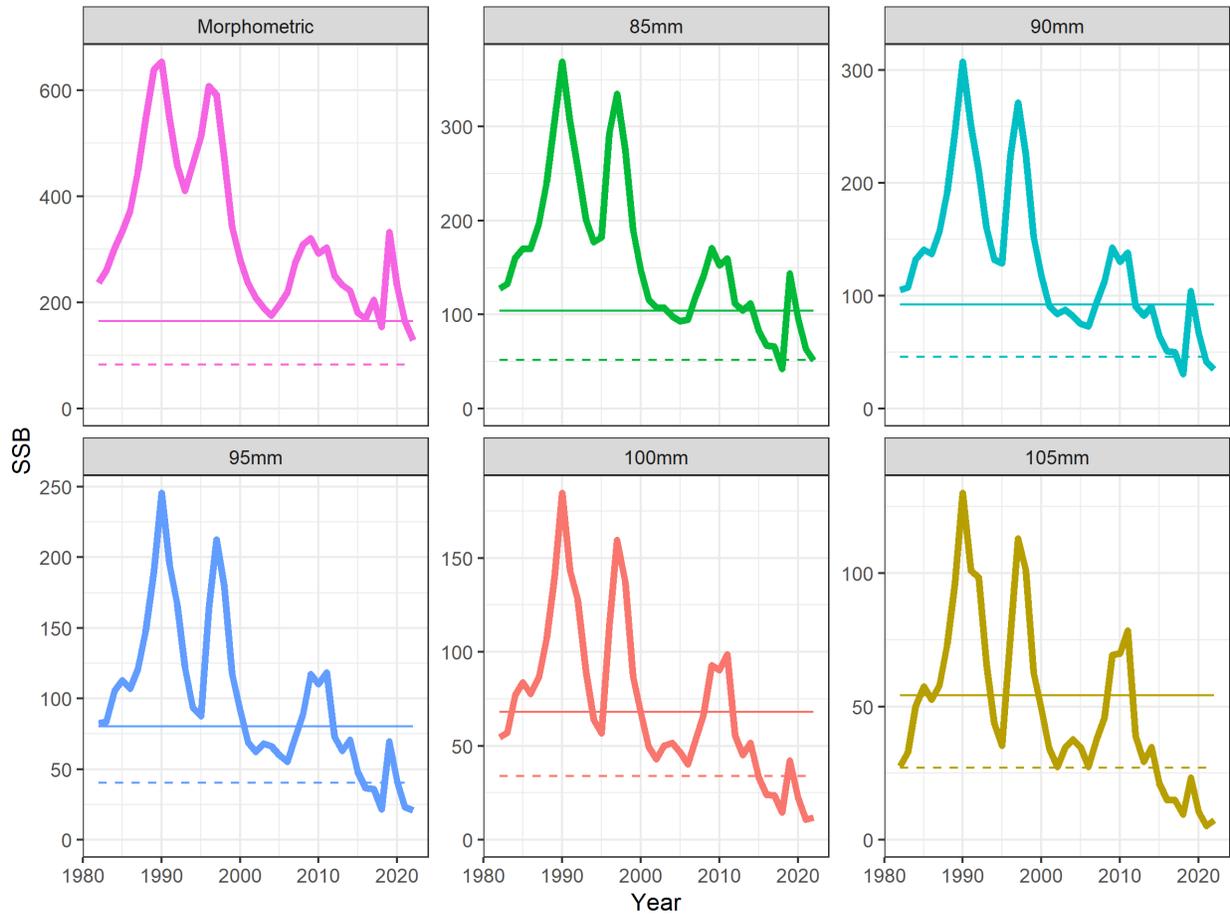
Figure 13. Trajectories of mature male biomass under different definitions of maturity (top label). The solid lines are $B_{35\%}$ for each definition and the dashed lines are the minimum stock size thresholds.
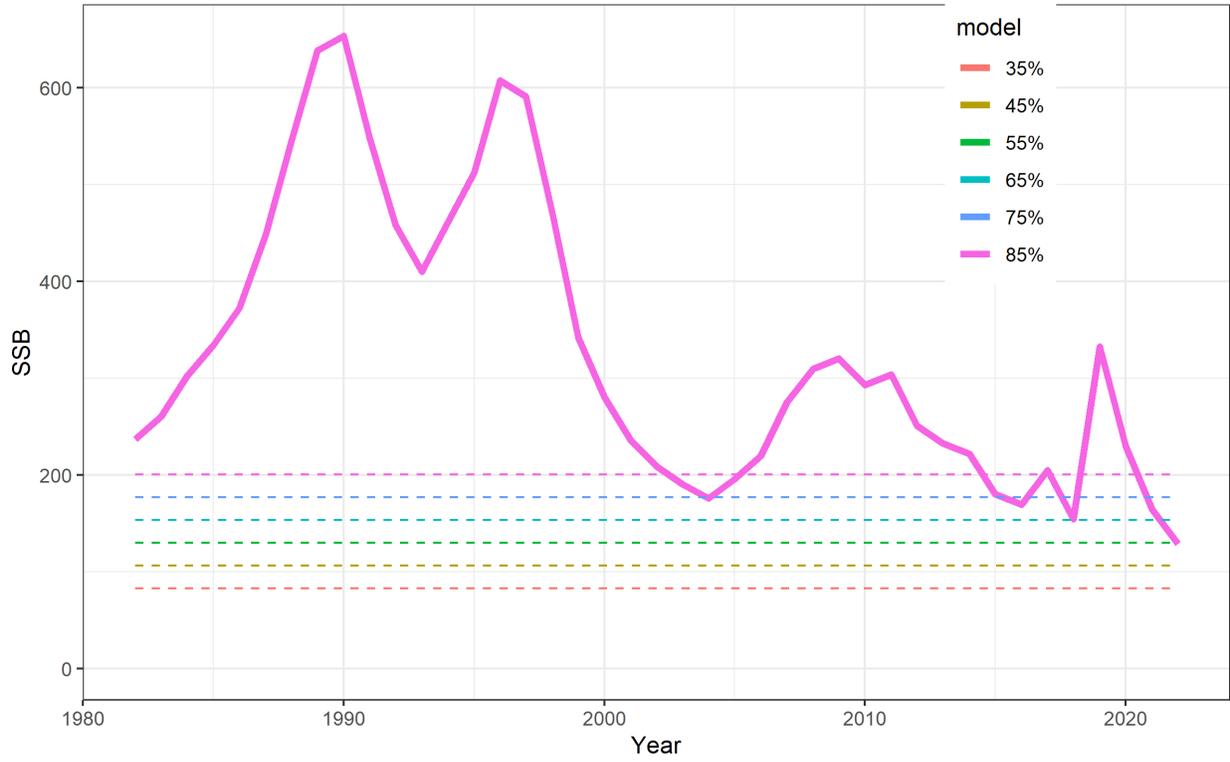
Figure 14. Estimated morphometrically mature male biomass relative to minimum stock size thresholds for a range of SBPR target percentages (colored horizontal dashed lines).

Appendix I: Climate-linked projections of eastern Bering Sea snow crab

This appendix is a draft of a manuscript to be submitted to a peer-reviewed scientific journal and should not be cited until after publication.

Density dependence can modulate climate change impacts on populations

Cody Szuwalski[1,2], Baptiste Alglave[3], Laurinne Balstad[4], Maxime Olmos[5], Andre Punt[2], Jon Richar[6], Matthieu Veron[2]

[1]Alaska Fisheries Science Center, National Oceanic and Atmospheric Administration, Seattle, WA, 98105
[2]University of Washington, School of Aquatic and Fishery Sciences, Seattle, WA
[3]University of South Brittany
[4]University of California, Davis
[5]IFREMER, Brest, France
[6]Alaska Fisheries Science Center, National Oceanic and Atmospheric Administration, Kodiak, AK

Population dynamics are the sum of many processes, each of which may be influenced by density dependence and climate. Models of population dynamics are important in conservation and economic planning, but incorporation of these drivers varies. Here we demonstrate the impacts of different assumptions about climate and density dependent drivers of reproduction, mortality, and maturity for projections of the eastern Bering Sea snow crab (*Chionoecetes opilio*) population. Feedbacks among processes result in different short- and long-term projections. For example, including density dependence in mortality slows population declines driven by climate change in the short-term, but as ice recedes the harmful effects of warming outweigh the protective effect of lower density. Our results illustrate the importance of considering climate and density effects on a range of population processes in projections and the need for management that addresses the impacts of non-stationarity in population dynamics.

Populations grow and shrink over time as a result of the net effect of many biological processes. Reproductive processes determine recruitment, or the number of young individuals joining the population. Growth and maturity processes determine the size of the animals in the population and their contribution to reproduction. Predation, disease, and metabolic needs contribute to mortality rates, which determine the number of individuals exiting the population. Each of these processes can be influenced by external climate factors (e.g. sea surface temperature or rainfall) or internal factors (e.g. density dependence). The feedbacks among these processes can provide both stabilizing (e.g. density dependent growth, Lorenzen and Enberg, 2002; Stenevik et al., 2022; Matthias et al.; 2018) and destabilizing effects (e.g. Allee effects that destabilize dynamics at low population densities; Stephens et al., 1999; Berec et al., 2007).

Climate change is altering biological processes. Recruitment of marine species has been linked to a range of climate variables (e.g. sea surface temperature (Britten, 2015; Le Bris et al., 2018), large scale indices of climatic variation (Faillettaz et al., 2019), plankton (Beaugrand et al., 2003), and ice extent (Szuwalski et al., 2021). North Sea Atlantic cod growth rates and size-at-age have been linked to changes in temperature (Neuheimer and Gronkjaer, 2012). Bering Sea groundfish growth rates show strong temporal variation related to environmental change (Stawitz et al., 2015). Changes in maturity have been related to density dependent effects for silver hake (Helser & Almeida, 1997), exploitation in herring (Engelhard & Heino, 2004; Olsen et al., 2004), and environmental drivers like temperature and prey-availability for fresh water perch (Kaemingk et al. 2014). Changes in mortality have been linked to heatwaves (e.g. snow crab (Szuwalski et al., 2023) and blue petrel (Barbraud and Weimerskirch, 2003)). Rain-on-snow events increase mortality of vulnerable ages of reindeer in the Arctic (Hansen et al., 2019). Increased hurricane activity has been linked to increased mortality in dolphins in the Bahamas (Coxon et al., 2022).

Projections of populations under climate change can support effective conservation and economic planning. Knowledge of plausible futures of the populations under management can be used to anticipate and adjust for potential problems like population collapse or explosion. Populations are often projected by fitting models to historical data to estimate parameters that determine processes influencing population dynamics and then projecting the population forward based on those parameters (Hollowed et al., 2009; Holsman et al., 2020, Punt et al. 2024). Sometimes a process is linked to projections of a climate variable derived from a global climate model to inform future variation in that process (Szuwalski and Punt, 2013; Amar et al. 2009). However, most commonly only single processes (like recruitment or growth) are linked to climate and/or population density in projections (Punt et al. 2014), in spite of the strong possibility that multiple processes may be affected by a changing environment.

Here we demonstrate how considering density dependence and climate drivers in multiple processes can change perceptions about the future of a population using eastern Bering Sea snow crab (*Chionoecetes opilio*) as a case study. Snow crab is a useful motivating example because an economically costly collapse in 2021 resulted in the first ever fishery closure (Szuwalski et al., 2023) and initiated a rebuilding plan that requires population projections under different climate scenarios and fishing strategies. The stock is widely distributed in the Bering Sea (Figure 1) and longitudinal datasets are available to estimate variability in key population processes and their relationship to climate variables.

We fit an integrated population model to abundance and size-composition data for immature and mature male snow crab collected in the National Marine Fisheries Service bottom trawl survey from 1982-2022. A fishery occurs on the largest individuals in the population and the retained and discarded numbers-at-size of snow crab from this fishery were fit to in our model to account for fishery removals. Time-varying natural

mortality (separately for immature and mature crab), recruitment, and fishing mortality were estimated within the model. Then we used generalized additive models (GAMs; Wood, 2011) to explain the variability in natural mortality, recruitment, and the probability of maturing using climate variables and estimates of crab density. Finally, we projected the snow crab population forward to 2060 under two potential climate futures while varying which processes were linked to density and climate variables (see SM for complete methods, simulation testing, and diagnostics).

Previous studies examining snow crab recruitment identified a positive relationship with ice extent and a negative relationship with Arctic oscillation (a measure of the volatility in the jet stream; Szuwalski et al., 2021). Szuwalski et al., (2023) identified positive relationships between natural mortality of mature snow crab and bottom temperature and crab density. Snow crab exhibit a terminal molt to maturity after which they no longer grow (Tamone et al, 2005). The probability of maturing has been linked to the density of large males in eastern Canadian populations (Mullowney, 2021). Based on these studies, we tested for relationships between ice extent (which can be a proxy for bottom temperature), the Arctic Oscillation, and our model estimates of population density against our estimates of mortality, recruitment, and the probability of maturing (see SM for details).

The population dynamics model fitted both the survey and catch data well, which is perhaps unsurprising given the estimation of time-varying mortality (Figure 1). Estimated natural mortality varied over time, with mature natural mortality higher on average and more variable than immature natural mortality (Figure 2; see SM for discussion). Estimated fishing mortality was lower than mature natural mortality on average, though several years of high estimated fishing mortality occurred. Estimated recruitment was punctuated by two large pseudocohorts (i.e. groups of similarly sized-crab; Figure 2).

GAMs including climate variables and mature male density explained the variability in mortality, recruitment, and maturity better than those without based on AIC (Figure 2; SM for details). GAMs using density and ice cover (as a proxy for temperature) explained 50% and 51% of the deviance in immature and mature mortality, respectively. Less ice was associated with higher mortality for both mature and immature crab. More mature males were associated with higher mature mortality, but the relationship for immature mortality was dome-shaped (Figure 3). GAMs using the Arctic Oscillation, ice cover, and density of mature crab explained 67% of the deviance in estimated recruitment. Negative phases of the AO and more ice were associated with higher recruitment (Figure 3). GAMs using carapace width, population density, and ice cover explained 93% of the deviance in the probability of maturing. Colder temperatures, larger numbers of small and large crab were associated with smaller probabilities of undergoing terminal molt (Figure 3).

We simulated two climate scenarios including a future in which environmental conditions were sampled from historical conditions and a future in which ice extent in the Bering Sea declined linearly from the average conditions seen in 2022 to the minimum observed ice extent (seen in 2018; figure s1). Projections to 2060 were repeated 100 times with variability added to the linear decline in ice comparable to historical variance. The two climate scenarios represent a future in which climate change mitigation is successful and future climate conditions look similar to the past and a future in which climate change is not successfully mitigated and future climate conditions do not resemble the past. Projected fishing mortality rates were set to the average over the last 15 years, which was a period of stable management. Four combinations of variability in population processes driven by environment and density effects from the best fitting GAMs from above were considered: variable recruitment (Rec), variable recruitment and mortality (Rec + M),

variable recruitment and maturity (Rec + Maturity), and variable recruitment, mortality, and maturity (Rec + Maturity + M).

When environmental conditions were sampled from historical values, the median trajectories of the number of mature males in the population and other population processes were similar to historical dynamics (Figure s2). Ensuring this is the case is an important check on the model before projecting under directional change in climate. When ice declined over the projection period, the stock rapidly declined after a brief rebound. The magnitude of this rebound was almost 50% larger when mortality included density-dependence and climate drivers compared to when it did not (Figure 4). Declines in recruitment related to receding ice drove the decline of the stock in projections and increases in mortality related to disappearing ice compounded the decline after ~2040. The probability of maturing at size increased in the short term, but ultimately fell below historical averages as a result of fewer large males in the population. Dramatically lower abundances under receding ice would affect the fishery given potential fishery closures.

Feedbacks in our models were determined by the directionality and strength of climate relationships and density-dependence on population processes. Mortality increased with population size for mature crab, recruitment was dome-shaped with respect to population size, and the probability of maturing depended on both the numbers of small and large mature crab. If climate drivers weren't considered in mortality, low projected population density would suggest that mortality should remain low. However, when accounting for increasing temperatures, mortality increased towards the end of the projections even while abundance remained low. The opposite directions of the relationships to mortality between density and temperature resulted in a stable mortality rate for more of the projection than if only one process was considered and that rate was lower than the median of the historical estimates. The probability of terminally molting also exhibits a number of feedbacks: it is affected by the number of mature males, which is in turn influenced by mortality. Linking mortality to density and the environment resulted in an increase in the number of mature males in the short term, which was then related to an increase in the probability of terminally molting.

Global climate models and current carbon mitigation efforts suggest an ice-free future in the eastern Bering Sea (Cheng et al., 2021; Crawford et al., 2021), which bodes poorly for the long-term future of snow crab in this region. However, our projected short-term prospects for snow crab changed depending on which processes were assumed to vary with the environment and density and the difference in short-term projections among scenarios could have large implications for management decisions. In the U.S., when a managed population declines beneath a specified biomass, a rebuilding plan is legally required to return the population to a target biomass. The decisions required when a stock is deemed unable to return to that biomass even in the absence of harvest can be different than those when the stock has the potential to reach the target biomass, so distinguishing between the two scenarios is important to management.

Decisive management actions based on projections like these can be difficult due to uncertainty associated with the relationships identified between the environment and population processes. Recruitment processes and their relationship to the environment have a long history of study in fisheries, yet recruitment-environment relationships are infamously unstable. Myers (1998) found that only 28 of 74 significant recruitment-environment relationships reported in the literature retained their significance when new data were added to the time series. When more than one process is linked to the environment and density, uncertainties are compounded, not only because there are more relationships to breakdown, but also because changing the relative strengths of the climate drivers and density dependence on the processes alters the way in which they feedback on one another and can result in different population trajectories.

Even if these uncertainties are ignored, incorporating non-stationary projections of populations can be difficult to incorporate into management systems dependent on equilibrium assumptions. Maximum sustainable yield is the prevailing management paradigm in fisheries science and many management plans have provisions to consider the current climate conditions in setting management targets (e.g. the Magnuson-Stevens Act and the Common Fisheries Policy: MSA, 2006; Lado, 2016). Under climate change-driven changes in productivity, adapting management targets to incorporate current climate conditions would have the effect of placing higher anthropogenic pressure on populations experiencing climate stress than if not adapting targets (Szuwalski et al., 2023). This seems like a counter intuitive management response, yet can be a consequence of attempting to apply equilibrium-based management targets in a non-stationary environment (Szuwalski and Hollowed, 2016).

In addition to difficulties identifying appropriate management targets, projections that suggest a continued decline of a population raise difficult management questions for which science-based answers are scarce. If we place confidence in an abrupt decline of snow crab sometime in the next 20 years, how should that influence management today? One possible management response is maintaining fishery closures to allow the population every chance to adapt and survive in new climates. Another possible management response is harvesting the commercially viable biomass that would otherwise die in the water as a result of inhospitable environmental conditions. Our current management frameworks do not have quantitative methods for addressing these questions, but we will need to address these problems. One of the key uncertainties in these decisions is the adaptive capacity of species experiencing climate stress.

Attempting to predict the future trajectories of populations given observed relationships between population processes and the environment comes with many challenges, but ignoring potential climate drivers can be even more undesirable. Unrealistic expectations of population futures can confound conservation planning, resulting in chasing unreachable targets in management, towards which a lack of progress can be seen as a failure of management. Ignoring climate drivers can result in economic plans that are not sufficiently flexible to respond to the coming changes in ecosystems (e.g. capitalization in fishing fleets designed to capture species that will no longer be present in the ecosystem). One of the difficulties in managing populations responding to climate stress is that resource managers do not possess a lever for the fundamental process driving change in their systems: atmospheric CO_2 (Mullowney et al., 2023). The lack of a clear picture of what to expect for future mitigation confounds efforts of management and clarity does not appear forthcoming given uncertain and unfulfilled promises by governments of key CO_2 producing nations (Brown et al., 2019).

Longitudinal data collection targeted at identifying relationships such as those described here for snow crab is needed to produce plausible projections and should be a key priority in institutions charged with tracking climate-driven changes in population dynamics. Data collection at a wider spatial scale than the population currently resides is particularly important, so that distributional shifts (or lack of shifts) can be captured. Analysis of population dynamics at range edges and along spatiotemporal gradients may be useful for understanding population responses to novel environmental conditions (Fredston et al., 2021). Populations will be more likely to approach extreme abundances (both high and low) that have been rarely observed in the past under the coming changes in climate. Allee effects may also become more prevalent and density dependence may become stronger under less hospitable environmental conditions (Berec, 2007). Some processes (e.g. growth and recruitment) can be easier to study because they can be directly observed, while other processes (e.g. mortality and catchability) can be more difficult because they are confounded and must be inferred from commonly available data, rather than directly observed (Szuwalski, 2022). More

focused data collection may also be required to identify changes in processes that are difficult to characterize with current data sources. Estimating annual mortalities, for example, could be improved by implementing large-scale mark-recapture studies (Queroue et al., 2021) and other important (and potentially changing) life history information like maturity could be collected alongside these efforts.

Our example here is a harvested marine population, but the importance of considering density dependence and climate forcing of population processes is applicable to the modeling of all populations experiencing the effects of climate change. Identifying three processes influenced both by climate and density dependence as we have done here for snow crab is unusual in the literature, but this may become a more common occurrence as species are subject to environmental conditions that are outside of historical ranges (Berec, 2007). It is important to note that density-dependence will not universally stabilize mortality as occurred here; population level outcomes depend on the direction and relative strengths of processes. Baltic blue mussels are an example of the opposite outcome in which density dependence appears to exacerbate the harmful effects of climate change (Jaatinen et al., 2021).

Identifying climate drivers of populations and considering them in projections of populations will need to be a priority as climate change becomes more pronounced (Hollowed et al., 2023). The identification of these drivers can be difficult to incorporate into currently used management systems that rely on equilibrium assumptions, but examples of complex intrinsic and extrinsic drivers are needed as testing beds to develop new methodologies. Feedback between climate drivers and density-dependence could play an outsized role in the dynamics of populations under historically unobserved conditions. Considering these drivers and feedbacks may help us understand the future trajectory of a population better, but this understanding is only the first step in answering the difficult questions posed above. Models and projections like the one we present here will likely be a center point for the cooperative, transparent, and adaptive decision making that will be needed to address the complex and contentious natural resource problems on the horizon.

References (still to be formatted)

1. Lorenzen, K., Enberg, K., 2002. Density-dependent growth as a key mechanism in the regulation of fish populations: evidence from among-population comparisons. Proceedings of the Royal Society of London. Series B: Biological Sciences 269, 49–54. https://doi.org/10.1098/rspb.2001.1853

2. Stenevik, E.K., Hølleland, S., Enberg, K., Høines, Å., Salthaug, A., Slotte, A., Vatnehol, S., Aanes, S., 2022. Predicting density-dependent somatic growth in Norwegian spring-spawning herring. ICES Journal of Marine Science 79, 1340–1352. https://doi.org/10.1093/icesjms/fsac057

3. Matthias, B. G., Ahrens, R. N., Allen, M. S., Tuten, T., Siders, Z. A., & Wilson, K. L. (2018). Understanding the effects of density and environmental variability on the process of fish growth. *Fisheries Research*, *198*, 209-219.

4. Stephens, PA, Sutherland, WJ, Freckleton, RP. (1999) What is the Allee Effect? Oikos. 87(1): 185-190.

5. Berec, L, Angulo, E., Courchamp, F. (2007) Multiple Allee effects and population management. Trends in Ecology and Evolution. 22(4): 185-191.

6. Britten, G, Dowd, M, Worm, B. (2015) Changing recruitment capacity in global fish stocks. Proc Nat Acad Sci. 113(1): 134-139.

7. Le Bris, A., Mills, K.E., Wahle, R.A., Chen, Y., Alexander, M.A., Allyn, A.J., Schuetz, J.G., Scott, J.D., Pershing, A.J., 2018. Climate vulnerability and resilience in the most valuable North American fishery. Proc Natl Acad Sci USA 115, 1831–1836. https://doi.org/10.1073/pnas.1711122115

8. Faillettaz, R., Beaugrand, G., Goberville, E., Kirby, R.R., 2019. Atlantic Multidecadal Oscillations drive the basin-scale distribution of Atlantic bluefin tuna. Science Advances 5, eaar6993. https://doi.org/10.1126/sciadv.aar6993

9. Beaugrand, G., Brander, K.M., Alistair Lindley, J., Souissi, S., Reid, P.C., 2003. Plankton effect on cod recruitment in the North Sea. Nature 426, 661–664. https://doi.org/10.1038/nature02164

10. Szuwalski, C., Cheng, W., Foy, R., Hermann, A.J., Hollowed, A., Holsman, K., Lee, J., Stockhausen, W., Zheng, J., (2021)\. Climate change and the future productivity and distribution of crab in the Bering Sea. ICES Journal of Marine Science 78, 502–515. https://doi.org/10.1093/icesjms/fsaa140

11. Neuheimer, A.B., Grønkjær, P., 2012. Climate effects on size-at-age: growth in warming waters compensates for earlier maturity in an exploited marine fish. Global Change Biology 18, 1812–1822. https://doi.org/10.1111/j.1365-2486.2012.02673.x

12. Stawitz, CC, Essington, TE, Branch, TA, Haltuch, MA, Hollowed, AB, Spencer PD. (2015) A state-space approach for detecting growth variation and application to North Pacific groundfish. Can J Fish Aquat Sci. 72(9): 1316-1328.

13. Helser, T. E., & Almeida, F. P. (1997). Density-dependent growth and sexual maturity of silver hake in the north-west Atlantic. *Journal of Fish Biology*, *51*(3), 607-623.

14. Kaemingk, M. A., Graeb, B. D., & Willis, D. W. (2014). Temperature, hatch date, and prey availability influence age-0 yellow perch growth and survival. *Transactions of the American Fisheries Society*, *143*(4), 845-855.

15. Szuwalski, C. S., Aydin, K., Fedewa, E. J., Garber-Yonts, B., & Litzow, M. A. (2023). The collapse of eastern Bering Sea snow crab. *Science*, *382*(6668), 306-310.

16. Barbraud, C and Weimerskirch, H. (2003) Climate and density shape population dynamics of a marine top predator. Pro R Soc Lond B. 270: 2111-2116.

17. Hansen, BB, gamelon, M, Albon, SD, Lee, AM, Stien, A, Irvine, RJ, Saether, BE, Loe, LE< Ropstad, E, Veiberg, V, Grotan, V. (2019) more frequent extreme climate events stabilize reindeer population dynamics. Nature Comm. 1616. https://doi.org/10.1038/s41467-019-09332-5

18. Coxon, J., Arso Civil, M., Claridge, D. *et al.* Investigating local population dynamics of bottlenose dolphins in the northern Bahamas and the impact of hurricanes on survival. *Mamm Biol* 102, 1133–1148 (2022). https://doi.org/10.1007/s42991-021-00208-0

19. Hollowed, A. B., Bond, N. A., Wilderbuer, T. K., Stockhausen, W. T., A'mar, Z. T., Beamish, R. J., ... & Schirripa, M. J. (2009). A framework for modelling fish and shellfish responses to future climate change. *ICES Journal of Marine Science*, *66*(7), 1584-1594.

20. Punt, A. E., A'mar, T., Bond, N. A., Butterworth, D. S., de Moor, C. L., De Oliveira, J. A., ... & Szuwalski, C. (2014). Fisheries management under climate and environmental uncertainty: control rules and performance simulation. *ICES Journal of Marine Science*, *71*(8), 2208-2220.

21. Holsman, K.K., Haynie, A.C., Hollowed, A.B., Reum, J.C.P., Aydin, K., Hermann, A.J., Cheng, W., Faig, A., Ianelli, J.N., Kearney, K.A., Punt, A.E., 2020. Ecosystem-based fisheries management forestalls climate-driven collapse. Nat Commun 11, 1–10. https://doi.org/10.1038/s41467-020-18300-3

22. A'mar, Z. T., Punt, A. E., & Dorn, M. W. (2009). The evaluation of two management strategies for the Gulf of Alaska walleye pollock fishery under climate change. *ICES Journal of Marine Science*, *66*(7), 1614-1632.

23. Szuwalski, CS and Punt, AE (2013) Fisheries management for regime-based systems: a management strategy evaluation for snow crab in the eastern Bering Sea. ICES J Mar Sci.

24. Wood, SN. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. Royal Stat. Soc. (B)*, 73(1), 3-36.

25. Tamone, S.L., M. Adams and J.M. Dutton. (2005) 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), p.166-171.

26. Mullowney, DRJ and Baker, KD. (2021) Size-at-maturity shift in a male-only fishery: factors affecting molt-type outcomes in Newfoundland and Labrador snow crab (Chionoecetes opilio). ICES J Mar Sci. 78(2): 516-533.

27. Crawford, A., Stroeve, J., Smith, A. *et al.* (2021) Arctic open-water periods are projected to lengthen dramatically by 2100. *Commun Earth Environ* 2, 109. https://doi.org/10.1038/s43247-021-00183-x

28. Cheng, W, Hermann, AJ, Hollowed, AB, Holsman, KK, Kearney, KA, Pilcher, DJ, Stock, CA, Aydin, KY. (2021) Eastern Bering Sea shelf environmental and lower trophic level responses to climate forcing: Results of dynamical downscaling from CMIP6. Deep Sea Research Part II: topical studies in oceanography. 193: 104975.

29. Myers, RA. (1998) When do environment-recruitment correlations work? Reviews in Fish Biology and Fishereies. 8: 285-305.

30. Magnuson-Stevens Fishery Conservation and Management Reauthorization Act of 2006. (2007). Pub. L. No. 109-479, 120 Stat. 3575.

31. Lado, EP. (2016) The common Fisheries Policy: The question for sustainability. John Wiley & Sons. ISBN 978-1-119-08566-9.

32. Szuwalski, CS et al. Unintended consequences of climate adaptive fisheries management targets. *Fish Fish.* (2023) Fish and Fisheries. doi.org/10.1111/faf.12737

33. Szuwalski, CS and Hollowed, AB. (2016) Climate change and non-stationary population processes in fisheries management. ICES J Mar Sci 73(5): 1297-1305.

34. Mullowney et al. Subarctic no more.

35. Brown, C, Alexander, P, Arneth, A., Holman, I, Rounsevell, M. (2019) Achievement of Paris climate goals ulikely due to time lags in the land system. Nature climate change. 9: 203-208.

36. Fredston, A, Pinsky, M, Selden RL, Szuwalski, C, Thorson, JT, Gaines, SD, Halpern, BS. (2021) Range edges of North American marine species are tracking temperature over decades. Global Change biology. 27(13): 3145-3156.

37. Szuwalski, CS (2022) Estimating time-variation in confounded processes in population dynamics modeling: a case study for snow crab in the eastern Bering Sea. Fisheries research. 251: 106298.

38. Queroue, M, Barbraud, C, Barrauqnad, F, Turek, D, Delord, K, Pacoureau, N, Gimenez, O. (2021) Multispecies integrated population model reveals bottom-up dynamics in a seabird predator-prey system. Ecological Monographs. 91(3): e01459. https://doi.org/10.1002/ecm.1459

39. Jaatinen, K, Westerbom, M, norkoo, A, Mustonene, O., Koons, DN (2021) Detrimental impacts of climate change may be exacerbated by density-dependent population regulation in blue mussels. Journal of Animal Ecology. 90(3): 562-573.

40. Hollowed, AB, et al. (2020) Integrated modeling to evaluate climate change impacts on coupled social-ecological systems in Alaska. Fron Mar Sci. 6: https://doi.org/10.3389/fmars.2019.00775

41. L.S. Zacher, J.I. Richar, E.J. Fedewa, E.R. Ryznar, M.A. Litzow. (2023) The 2023 Eastern Bering Sea Continental Shelf Trawl Survey: Results for Commercial Crab Species. NOAA Technical Memorandum NMFS-AFSC.

42. Richar, JI and Foy RJ (2022) A novel morphometry-based method for assessing maturity in male Tanner crab, Chinoecetes bairdi. FACETS. 7: 1598-1616.

43. D.A. Somerton, K.L. Weinber, S.E. Goodman. Catchability of snow crab (*Chionoecetes opilio*) by the eastern Bering Sea bottom trawl survey estimated using a catch comparison experiment. *Can. J Fish Aquat. Sci*. **70,** 1699-1708. (2013).

44. O. Hamel. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. *ICES J. Mar. Sci*. **72,** 62-69. (2015).

45. H. Akaike. A new look at the statistical model identification. *IEEE Trans. Auto. Con*. **19(6),** 716-723. (1974).

46. F.J. Mueter, M.A. Litzow. Sea ice retreat alters the biogeography of the Bering Sea continental shelf. *Eco. App*. **18(2),** 309-320. (2008).

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

Figures
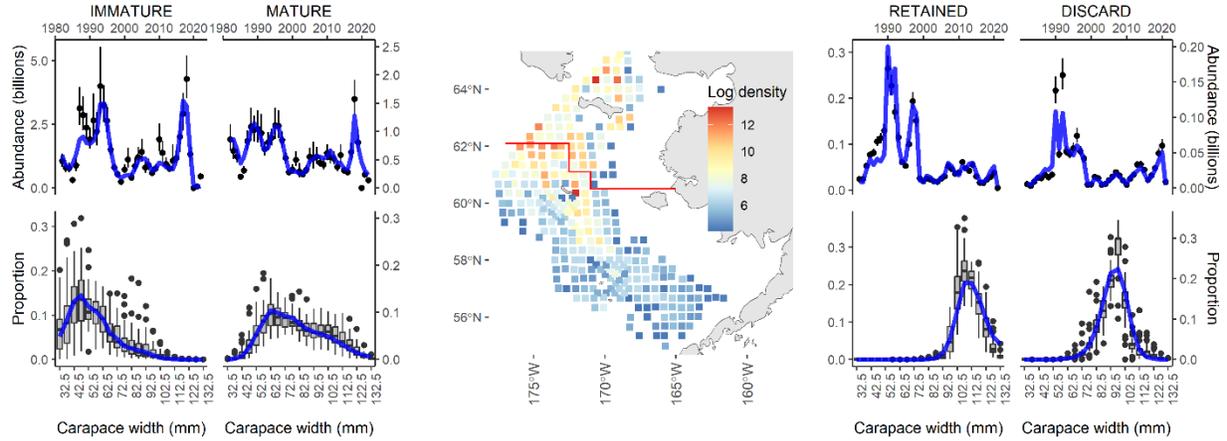
Figure 1. Data from the survey and fishery with model fits. Left panel are the immature and mature indices of abundance and size composition data from the survey. Middle panel is the observed density of snow crab in the eastern Bering Sea during the 2019 survey; each square is a survey station and the color is the log density in numbers per km^2. Right panel are the retained and discarded numbers and size-composition data in the directed fishery for snow crab. Observations are black dots with 95% confidence intervals for abundances and box plots for size-composition data, which are aggregated over time. Blue lines represent the model fits to the data.
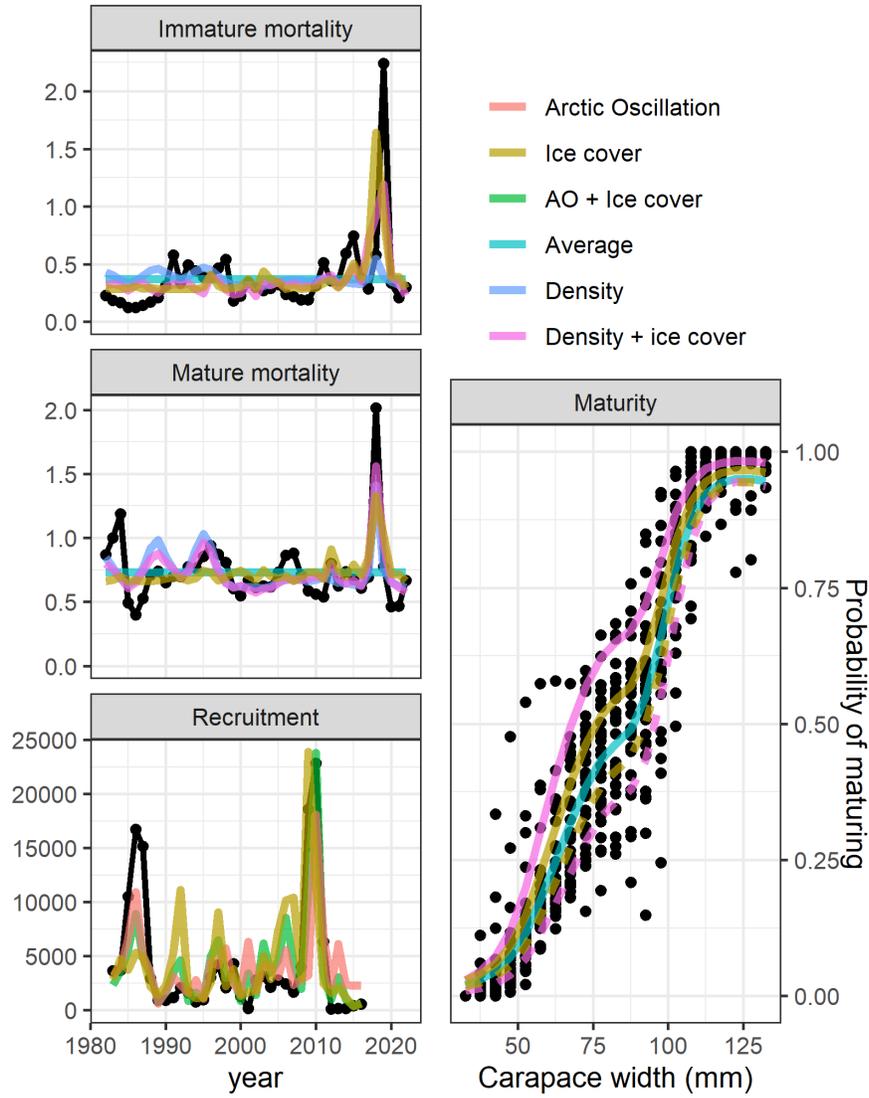
Figure 2. Identifying density and climate effects in population processes. Black dots and/or lines are estimated population processes from the population dynamics model (recruitment and mortality) or observations from the NMFS survey (probability of terminally molting). Colored lines are the fits from GAMs incorporating a given set of covariates. Colored lines for the probability of maturing in which environmental covariates are included represent the probability given the combination of variables that produce the highest and lowest observed values (solid and dashed, respectively).
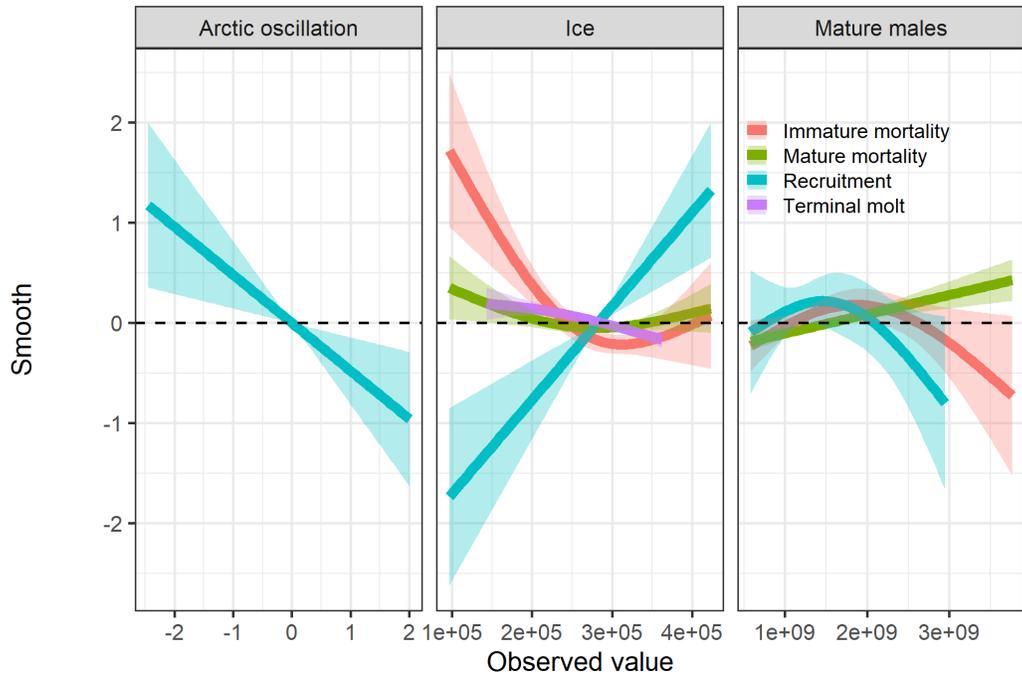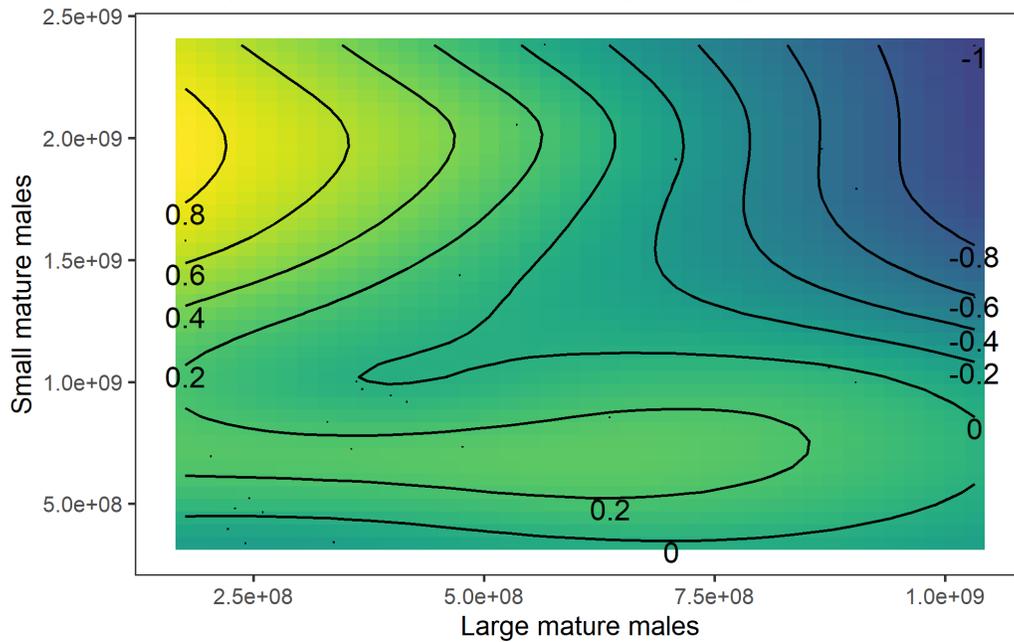
Figure 3. Shapes of estimated smooths relating population process to covariate (top). Each panel represents a covariate, each color corresponds to a population process, and shading represents one standard error. Bottom plot shows the surface relating the probability of terminally molting to the density of both large and small males. More yellow indicates higher probability of terminally molting; more blue indicates lower probability.
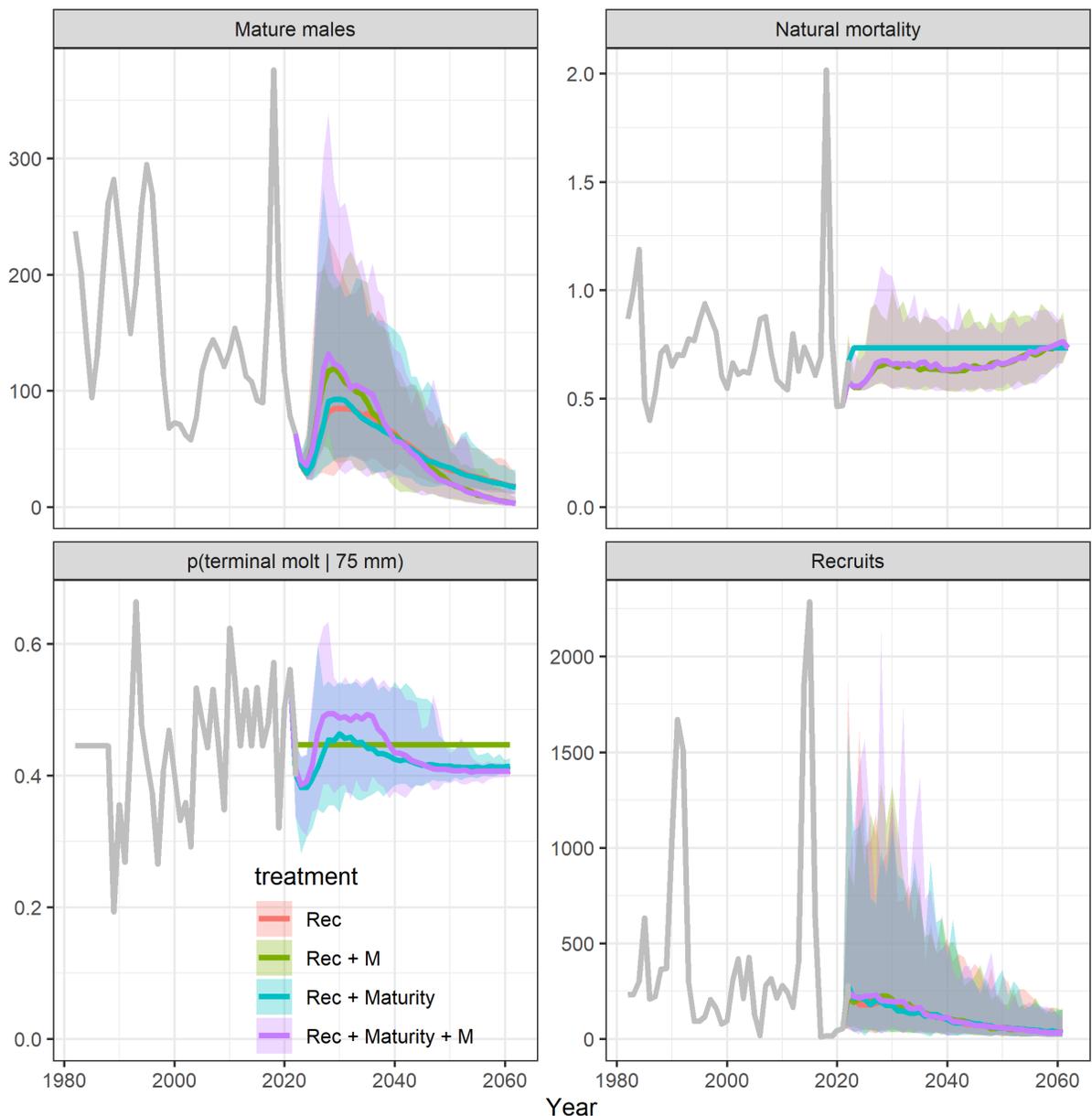
Figure 4. Projections of population dynamics under declining ice cover. Grey lines are the historical maximum likelihood estimates (or observations in the case of maturity) and colored lines are scenarios based on the biological processes allowed to vary over the projection based on the best fitting model in figure 2. Solid lines are the medians over 100 simulations; shaded areas represent the inter-simulation variability in populations in either the numbers resulting from the population processes or the population process linked to density and environment. Units for recruits and mature males are 10,000,000 individuals, natural mortality is a yearly rate, and maturity is the probability of having undergone terminal molt at 75 mm carapace width.

Supplementary materials for 'Density dependence can modulate climate change impacts on populations'

Methods overview
We used an integrated population dynamics model to estimate variation in mortality (partitioned into 'natural' and fishing) and recruitment over time for snow crab in the eastern Bering Sea. The population dynamics model was fit to abundance and size composition data by maturity state 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 for snow crab. We then used generalized additive models (GAMs) to relate the estimated variation in mortality and recruitment to environmental covariates. GAMs were also used to relate variation in observed probability of having undergone terminal molt to environmental covariates. The population was then projected under two climate scenarios and four scenarios for variability in population processes, all while being fished at the average estimated fishing mortality over the last 15 years.

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. 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 2 (37 x 37 km 2) 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 logcarapace 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.

Directed fisheries data
Fisheries catches are reported by Alaska Department of Fish and Game (ADFG) statistical area in logbooks and were aggregated across 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 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. Information on discarded crab are collected via on-board observers who measure, count, and weigh samples of crab brought aboard fishing vessels. Bycatch data are also available for non-directed fleets, but the mortality attributed to removals estimated by the stock assessment

by these fleets is so small the bycatch fisheries were not included in this analysis. More information on the fisheries data collection process can be seen in the stock assessment (Szuwalski, 2023).

Population dynamics model
See the main text for population dynamics model equations.

Fits and diagnostics
The fits to the immature and mature indices of abundance were good, with a few runs of positive residuals in the 1980s for immature crab (Figure 1; main text). This could be related to small changes in survey foot print between 1982-1988. Aggregate survey size compositions were very well fit (Figure 1), though there were some individual years for which some issues in fits were apparent. For example, predictions of smaller immature crab in 1998 were not able to match the observations well (Figure S8) and there were fewer predicted large crab in 1991 than observed (Figure S9). The fits to the abundance of retained and discarded catch were acceptable, though two years of high discards early in the time series were poorly fit (Figure 1). The aggregate size compositions of retained and discarded crab were generally well fit, but the model predicted more of the largest crab in the retained catch than were observed. This can also be seen in the fits to the yearly size composition data (Figure S10). Discarded size composition data were fit very well in some years (e.g. 1992 and 2003), but there were runs of years in which the observed data were markedly more peaked than the predictions (2007-2011; Figure S11). This could be related to a change in selectivity over time related to rationalization that was not modeled in our analysis.

How well can the model estimate population processes given simulated data?
An important exercises to perform with any model before using its output is to perform a 'self-test' in which data are simulated from the population dynamics model with appropriate error and then fit to by the model. The goal of this test is to determine whether or not a model can return the parameter values underlying the simulated data with the available quantity and quality of data. For our analysis, the ability of the model to estimate mortality and recruitment are of particular interest because they will be used as input into GAMs to attempt to link the estimates to environmental covariates.

Log-normal error was added to the true underlying abundance from the simulation model with three different coefficients of variation (CV): 0.05, 0.10, and 0.30. Log-normal error was added to the true underlying catches from the simulation model with three different coefficients of variation: 0.03, 0.05, and 0.1. Catches are thought to be better observed given observation dockside, so the range of CVs used were smaller than the survey data. Simulated data sets were generated 100 times under each observation error scenario and the population dynamics models were fit to them. The model presented in the text was fit to the simulated data.

Estimates of recruitment and fishing mortality tracked the simulated values very closely in all scenarios, with correlations ranging from 0.93 in the most uncertain data quality scenario to 0.99 in the best (Figure S12 & Figure S13). Estimates of immature mortality were tracked the simulated values more closely than mature mortality, with a correlation of 0.98 for the best data quality scenario and a correlation of 0.93 for the worst (Figure S14). Mature natural mortality was the most poorly estimated quantity in the simulations with a correlation between estimated and simulated values of 0.55 in the worst data quality scenario (Figure S15). However, when data quality was better, the correlations were more similar to those of the other processes (i.e. ~0.96). Ultimately the high correlations between the simulated and estimated values for scenarios that are similar to the data quality available for snow crab provide some justification for the use of the estimates in the second part of our analysis: relating variability in population processes to environmental and population density.

The estimate of mortality for mature males from our model (mean of ~0.7) was higher than is often assumed in management (0.28). One of the key differences between our model and the model used in management is that an index of abundance for immature crab is also fit in our model, whereas only size composition data are fit in the management model. Reconciling the immature and mature indices of abundance is difficult given informative priors on selectivity produced from the BSFRF experiments. When the mean natural mortality for both immature and mature male crab is not estimated, but rather specified as 0.271 (based on an assumed longevity of 20 years and using Hamel's (2015) empirical estimates for M), the model is unable to fit the indices of abundance (Figure S16). Immature animals are underestimated and mature animals are overestimated. The negative log likelihood decreases from 4680 to 4189 when allowing the estimation of just two additional parameters (mean natural mortality for mature and immature crab). The idea of a higher natural mortality for mature crab may be justified when it is thought of as 'longevity' in a particular maturity state. A crab lives perhaps 7 years after maturing, which would suggest a relatively high natural mortality (around 0.7) to ensure that the crab that mature are no longer in the population 7 years later. This observation is corroborated by observations of old shell mature crab in the survey after large pseudocohorts enter the population (Szuwalski, 2023).

Generalized additive models

Structure

Generalized additive models (GAMs) were used in the R programming language (package *mgcv*; Wood, 2011) to relate changes in estimated 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. Models were first fitted in which all potential relevant covariates were included in the model of the form:

$$m_{p,y} = s(\phi_{m,y}) + \epsilon_i \tag{1}$$

where 's()' is a smoothing function based on thin-plate splines, ϕ is a matrix of environmental covariates, and ϵ is normally distributed error. Then the relevant covariates to each process were included in a factorial design and models were compared via Akaike's Information Criterion (AIC; Akaike; 1974). One of the smooth terms for the probability of terminally molting was two-dimensional to capture the interaction between the numbers of small and large males in the population. The family used as a link function for each process specified when fitting the GAMs were different. A Tweedie distribution was used for mortality, a negative binomial distribution was used for recruitment, and a beta distribution was used for the probability of having undergone terminal molt. Model diagnostics were performed by simulating residuals using the R package 'DHARMa' which is useful when using link functions other than Gaussian.

Covariate construction

A wide range of covariates were considered in previous efforts at identifying relationships between population processes and environmental drivers. For recruitment, Szuwalski et al. (2021) considered ice cover, cod biomass, sea surface temperature, bottom temperature, the Alaskan Index, the Aleutian Low, the Arctic Oscillation, and the Pacific Decadal Oscillation. For mortality, Szuwalski et al. (2023) considered cannibalism, fisheries bycatch, temperature occupied, cod predation, population density, and disease. We

narrowed our range of covariates considered based on these analyses and adjusted them contingent on our ability to project them. For example, temperature occupied was an important variable in mortality models, but a spatial distribution of crab and the temperature they will occupy is not available for our projections. However, ice cover is correlated to bottom temperature and projecting declining sea ice is simple and well-supported by global climate models.

Arctic oscillation The Arctic oscillation is an index of climate variability related to the winds circulating around the Arctic. In positive phases, the band of wind encircling the Arctic is more stable, preventing excursions of cold Arctic air southward. Szuwalski et al. (2021) linked snow crab recruitment to the Arctic oscillation. The Arctic Oscillation is represented by the first mode of empirical orthogonal function analysis of the mean 1000 millibar height field over 20-90 degrees N latitude. Here we downloaded the monthly index using the R package "rsoi" and used the average over the months January to April to represent the winter conditions in a given year. Recruitment entering the model is related to the Arctic oscillation five years prior, which corresponds to the pelagic larval phase of snow crab life history.

Ice cover Temperature is one of the key physical variables that structures the benthic ecosystem of the EBS (Mueter and Litzow, 2008). The cold pool, a mass of water <2 degrees Celsius, can act as a barrier to species interaction based on temperature preferences of different species. Snow crab are a stenothermic species, preferring cold water and juvenile snow crab in particular are associated with the cold pool (Dionne, 2003). The cold pool is directly related to the winter ice extent in the Bering Sea and has varied dramatically over time as the ecosystem moves between cool and warm stanzas (e.g. 2006-2010 vs. 2014-2019; Figure 1b of the main text and Figure S17). Ice cover was linked to estimated recruitment in Szuwalski et al. (2020). Temperature occupied by snow crab was linked to estimated mortality in Szuwalski et al. 2023 based on increased metabolic demands associated with higher temperatures. Consequently, ice cover was tested as a proxy for temperature occupied in predicting mortality. A time series of historical ice cover was downloaded from the Alaska Marine Ecosystem Status Reports website (https://apps-afsc.fisheries.noaa.gov/refm/reem/ ecoweb/). Lags between ice and recruitment and maturity were selected to represent the early juvenile life history (e.g. ~3 years lagged from entry to the model). Ice was not lagged for mortality, given it is used as a proxy for bottom temperature and its impacts on metabolism.

Population density The mature population size (or mature density, given a constant area) was used in all models as a covariate. Specifically, the number of morphometrically mature males estimated in the population dynamics model not subject to survey selectivity was used for recruitment and mortality; whereas morphometrically mature male crab were split into two categories (i.e. small [<80 mm carapace width] and large [>=80 mm carapace width]) for use in comparison to the probability of having undergone terminal molt. Mature crab were split by size in the terminal molt models based on the assumption that size-mediated competition for mates could be associated with the 'decision' to molt to maturity or remain immature. If there are many large, mature males, delaying terminal molt until a size large enough to compete with the existing large males may hold an evolutionary advantage. On the other hand, if there were very few large, mature males, maturing at a smaller size may provide more opportunities to mate. If there were few large, mature males, and also many small males, waiting to mature in order to avoid competition with a large number of other small males might be advantageous. The use of mature males in modeling the probability of terminally molting was spurred by studies in eastern Canada on snow crab which showed an effect of the density of large males on terminal molt (Mullowney and Baker, 2022).

Fits and diagnostics

Each model that included environmental and density covariates outperformed the models without based on AIC (Table S1). The full model for mature mortality only performed marginally better than the model with only mature population size in it (Δ 1.5 AIC), but ice was retained in the model for projection given the potential for large changes in ice cover compared to historical measurements. The diagnostics associated with each of the best fitting models were acceptable, save the lowest quantiles for the residuals of the immature mortality model (Figure S18, Figure S19, Figure S20, and Figure S21). This is related to the low values in the early 1980s that are poorly captured by the model. The change in survey footprint from 1987 to 1988 is an unaddressed issue in this model, which may be the root of the issue here. This will be explored before publication. The estimated smooths for the best fitting models via AIC can be seen in figure 3 of the main text. Given the consistency with previous studies that performed extensive robustness testing (e.g. leave-one out crossvalidation, randomization testing, and incorporating uncertainty in estimates of recruitment and mortality into the models), we did not perform further robustness testing for this analysis.

Projected processes

For each year in each projection, a random value for the Arctic oscillation (AO) was generated by sampling the historical values. This was done because there is no obvious trend in the AO and global climate models do not suggest non-stationarity in this process. A value for ice cover was also generated in each year in each projection, but draws were performed in two different ways. The first sampled from historical values; the second generated values by interpolating a line from 2024 to 2060 between the points represented by the median ice cover and the minimum ice cover. Then, in each year that value was multiplied by a random normal variable with mean 1 and standard deviation of 0.2, which was chosen to produce similar variability to the historical time period in the initial years of the projections. The numbers of mature and immature males were summed from the numbers at size matrix for a given projection.

The appropriate simulated environmental covariates were then input to each of the GAMs that were fit to the historical data. New values of immature mortality, mature mortality, recruitment, or probability of terminally molting were predicted based on those inputs and then used to project the population one more step into the future. this process was repeated for each year of the projection. The projected values for each process were drawn directly from the estimated curve and do not incorporate additional uncertainty related to a lack of fit in the environment-density-population process models. When the family used in the GAMs is Gaussian, adding uncertainty in predictions is just a matter of extracting the standard error and sampling from the normal distribution associated with that mean and variance. However, when the family used in the GAMs is not Gaussian, the process is less straight forward. Consequently, the uncertainty in the projected scenarios presented here under-represents the actual uncertainty in the projection.

A word on methods and interpretation

Attribution of changes in population processes in ecology is a difficult problem, particularly for wild populations that are difficult to directly observe and next to impossible to experiment on in situ. There are a large range of methodologies that claim to identify causality in observational data (structural equation modeling, empirical dynamic modeling, etc.). Some of the difficulties in determining causality in ecological time series are related to the generally short time series that are available, non-linear dynamics, and departures of populations or covariates into unexplored parameter space. These issues can present problems for any modeling framework and we have tried to address these to the best of our ability with the models used here. In-depth interpretation of the mechanisms behind the relationships observed here is beyond the scope of this manuscript (and should be approached with caution), but discussions can be seen in the cited manuscripts that were used to identify potential covariates.

	Immature mortality	Mature mortality	Maturity	Recruitment
Ice	-30.18	-15.39	-998.72	
Large + Small pop, Ice			-1115.96	
Mature pop	-10.38	-16.29	-1003.76	667.85
Mature pop, AO				658.69
Mature pop, AO, Ice				649.47
Mature pop, Ice	-30.31	-19.84	-1028.55	655.72

Table S1:AIC by model by process (columns) and covariates (rows).
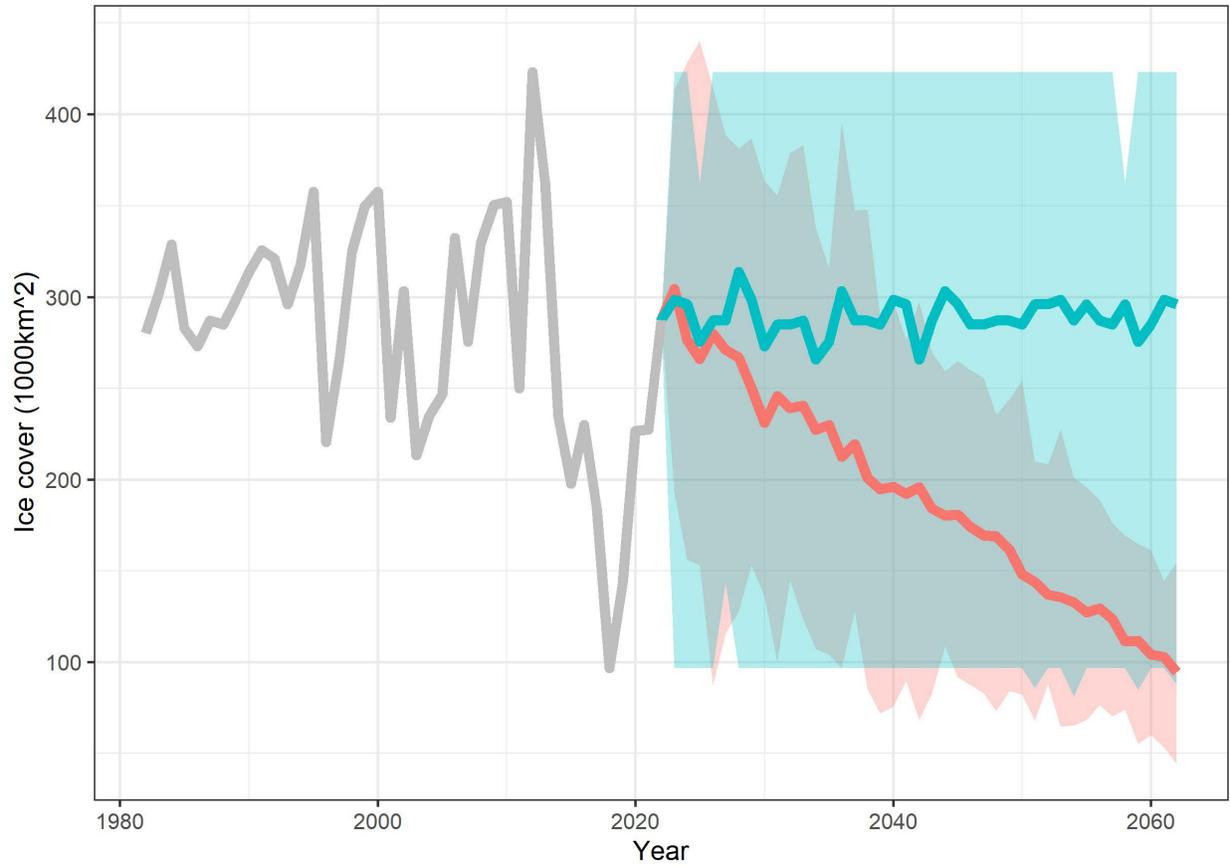
Figure S1: Historical (grey) and future (red and blue) ice scenarios in the Bering Sea. Solid lines are the median condition over simulations; shading are the inter-simulation ranges.
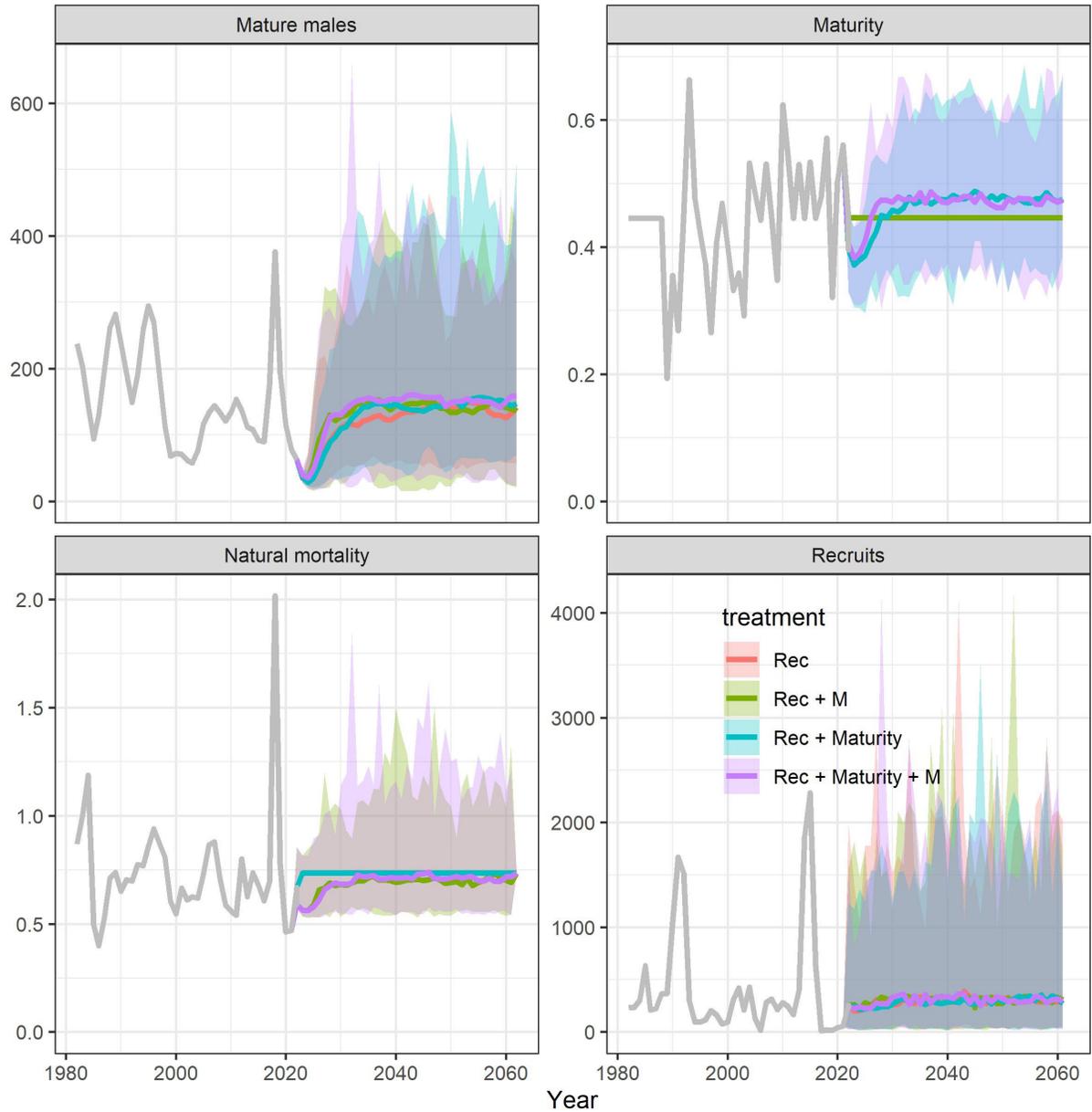
Figure S2: Projections of population dynamics under resampled historical environmental conditions. Grey lines are the historical maximum likelihood estimates (or observations in the case of maturity) and colored lines represent scenarios based on the biological processes allowed to vary over the projection. Solid colored lines are the medians over 100 simulations; shaded areas represent the inter-simulation variability in populations in either the numbers resulting from the population processes or the population process linked to density and environment. Units for recruits and mature males are 10,000,000 individuals, natural mortality is a yearly rate, and maturity is the probability of having undergone terminal molt at 75 mm carapace width..
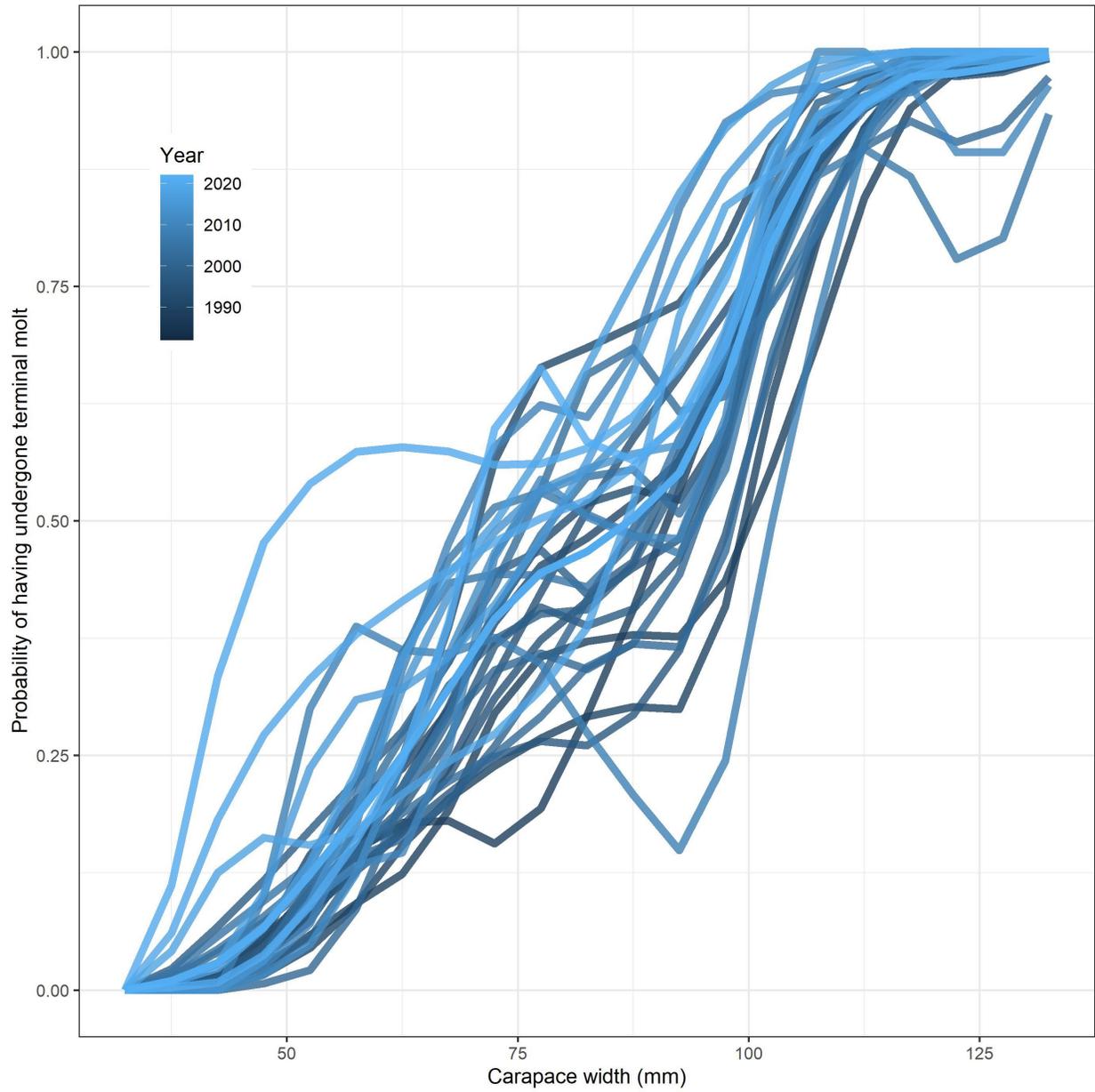
Figure S3: Observed probability of having undergone terminal molt (blue lines) in a given year, plotted over the average over time (red line).
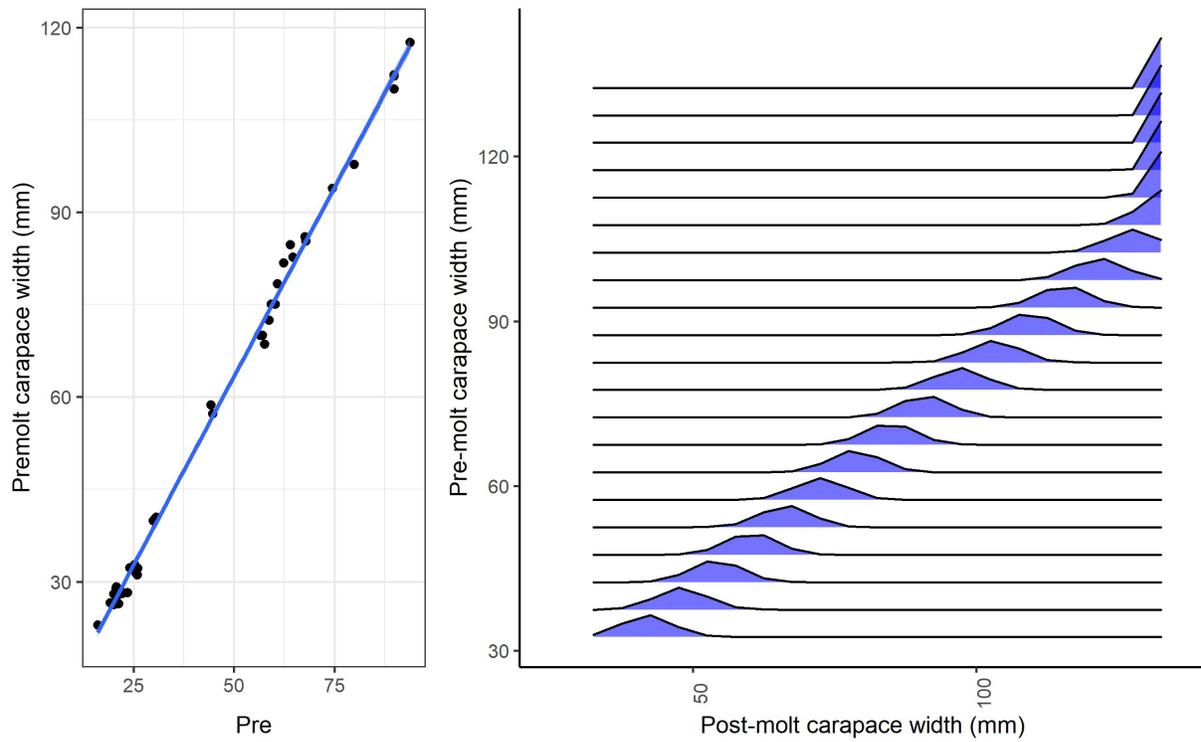
Figure S4: Observed relationship between pre- and post-molt size for male crab (left; dots are observations of individual crab, line) and the resulting size transition matrix used in the population dynamics model (right).
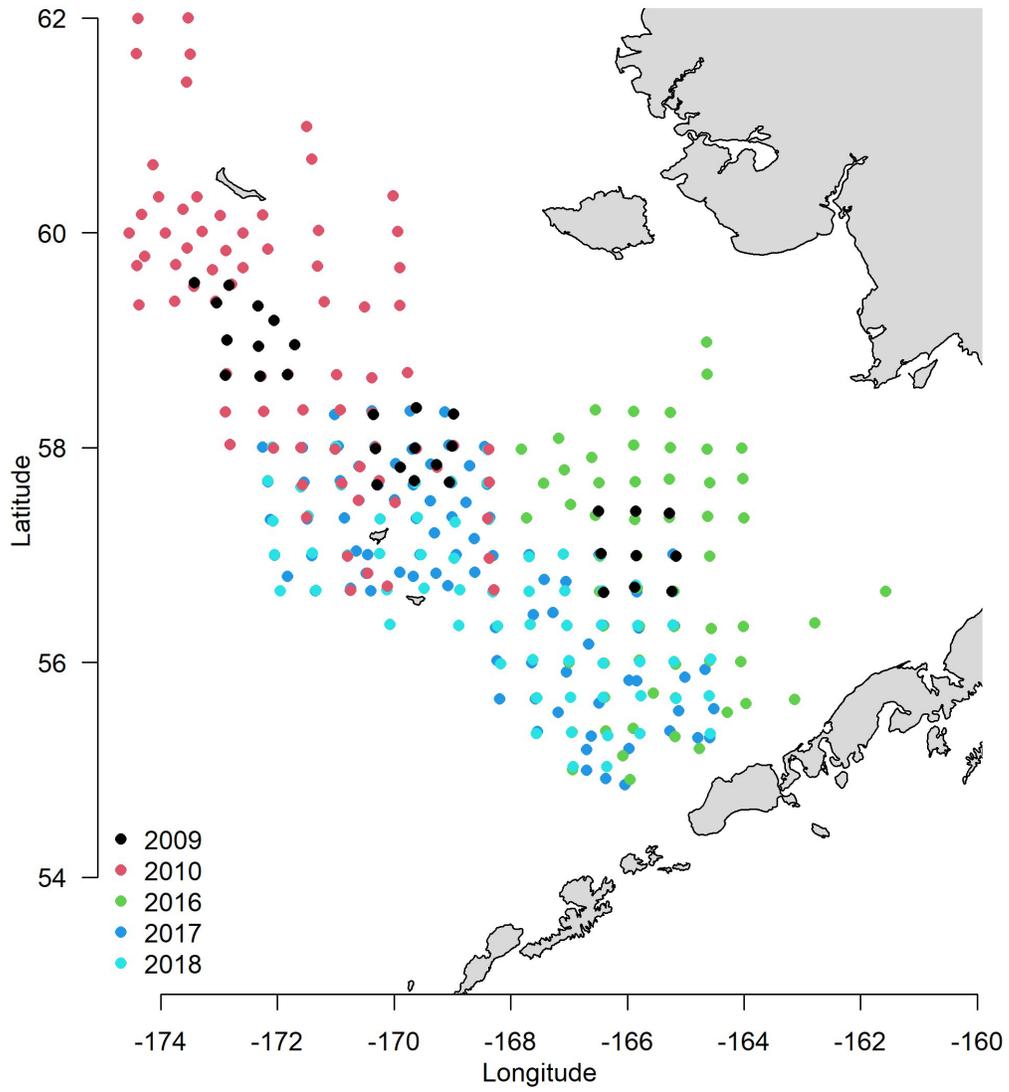
Figure S5: Map showing the survey location and timing of BSFRF experimental survey data collection.
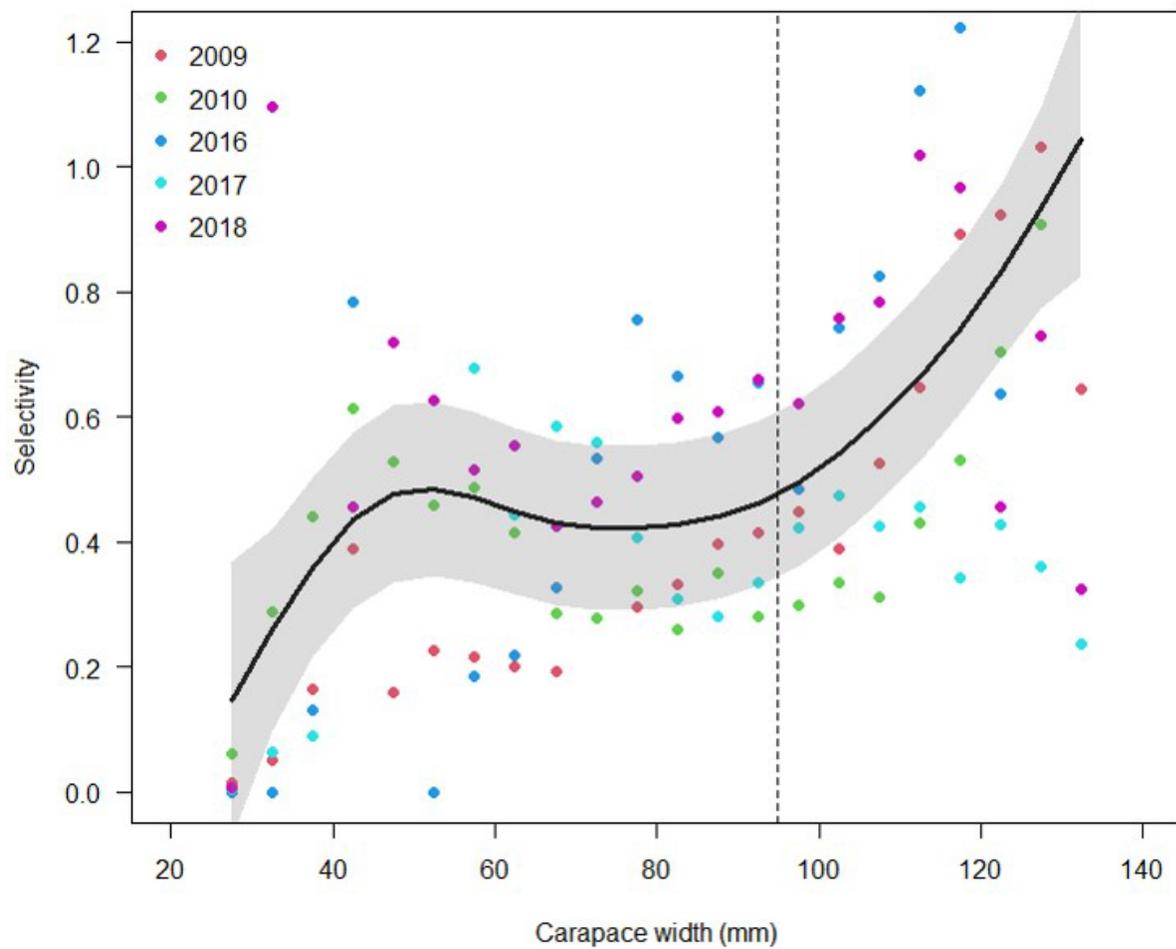
Figure S6: Inferred selectivity at size by year from the BSFRF selectivity experiments. Black line and grey shading are a GAM fit to the data.
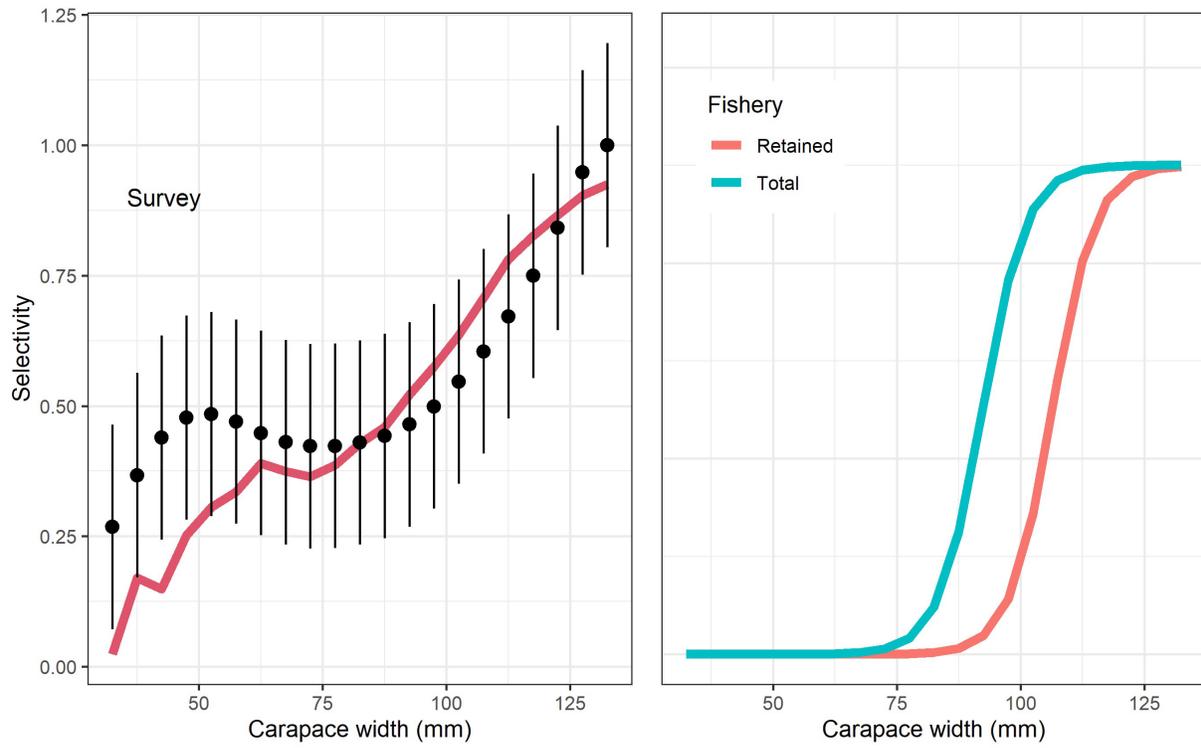
Figure S7: Estimated survey selectivity and fishery selectivity.
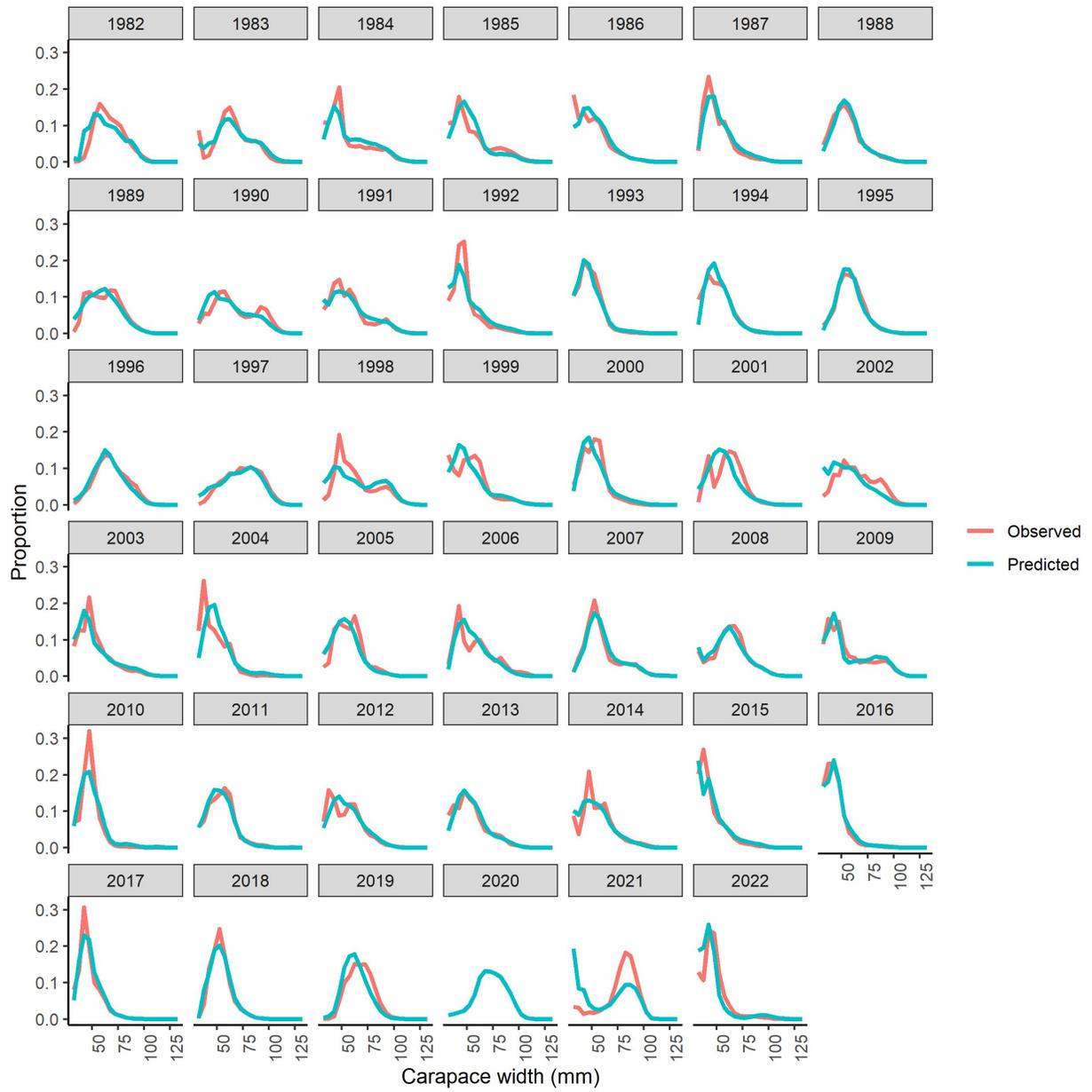
Figure S8: Model fits to individual years of size composition data for immature males captured in the survey.
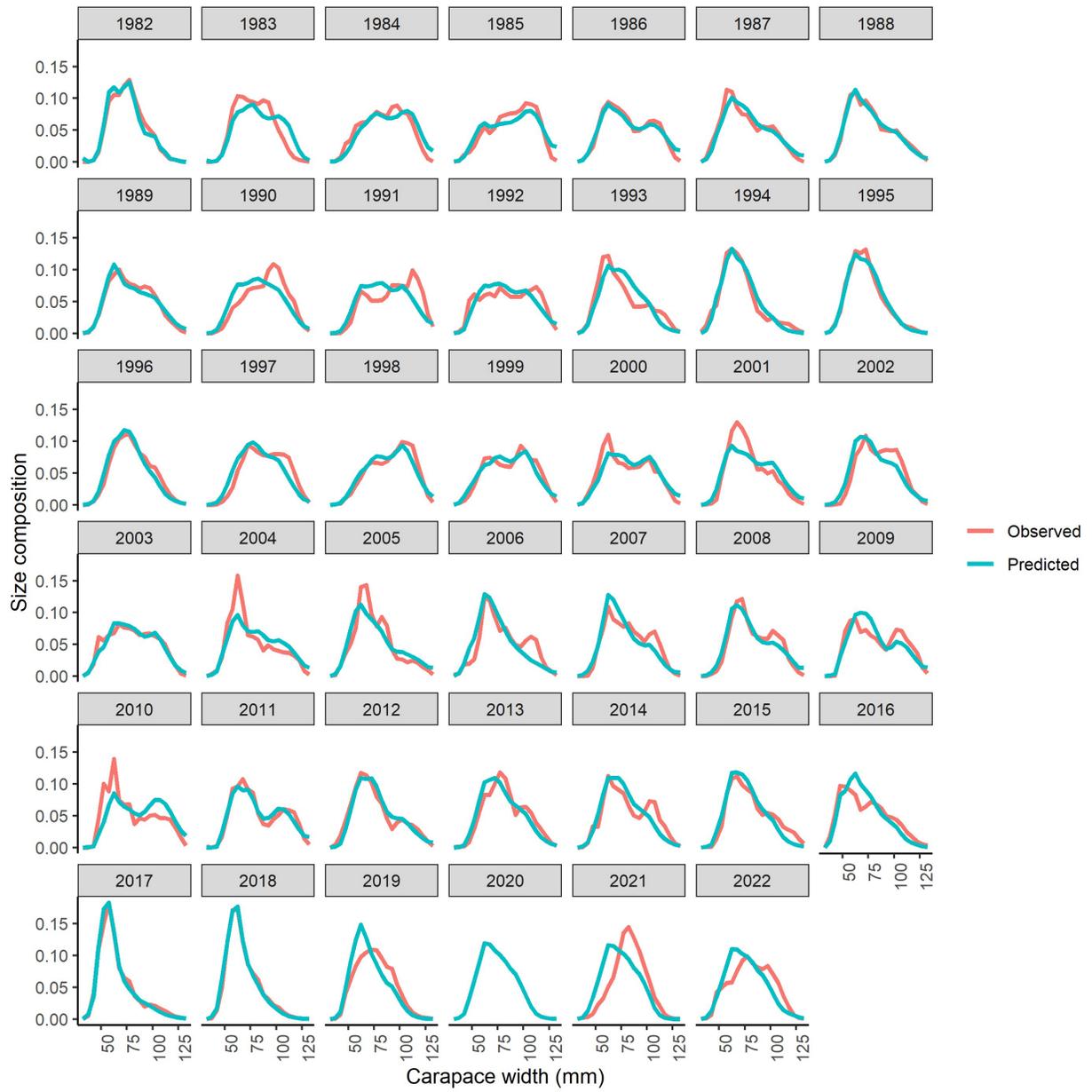
Figure S9: Model fits to individual years of size composition data for mature males captured in the survey.
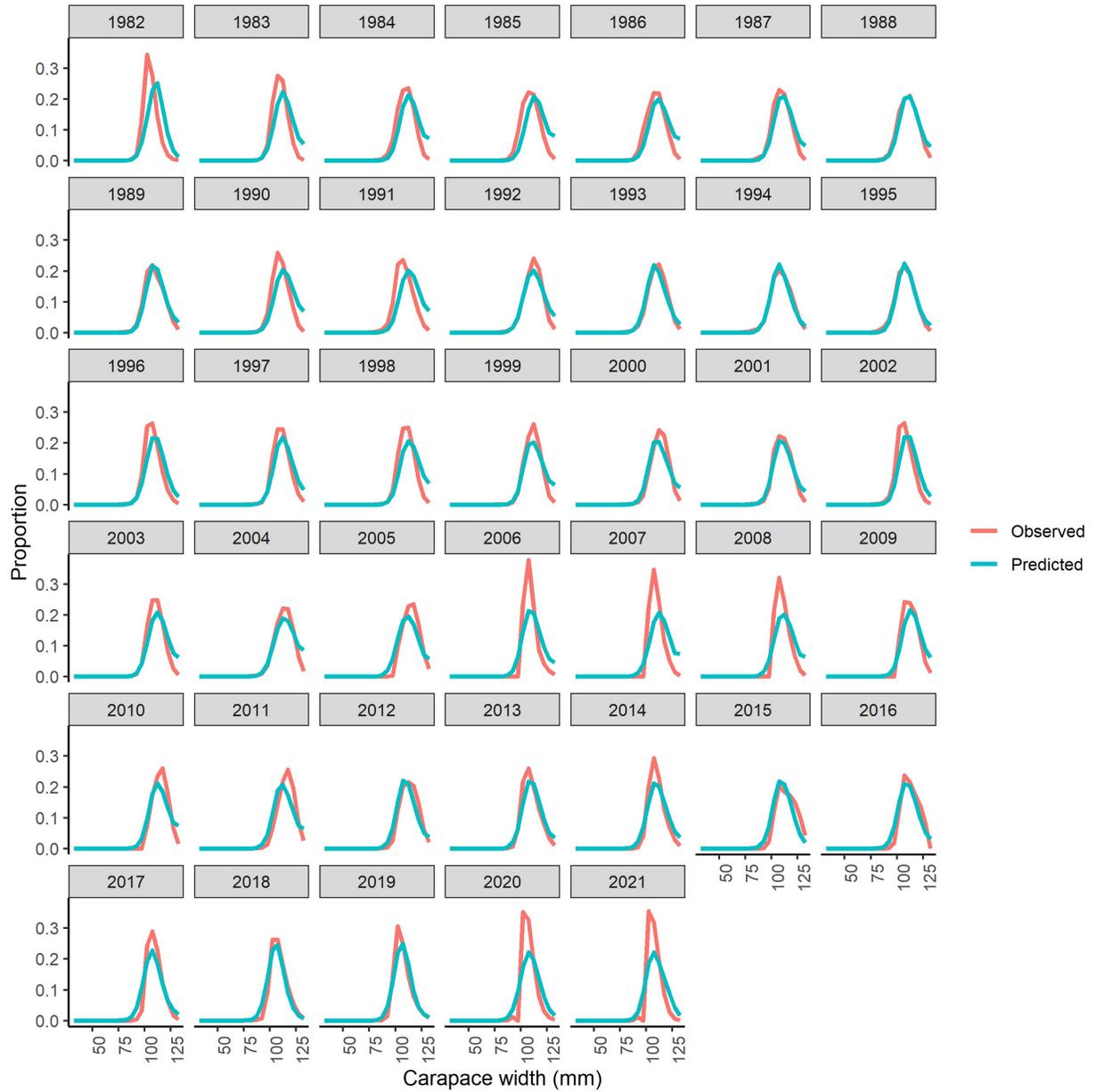
Figure S10: Model fits to individual years of size composition data for retained crab in the directed fishery.
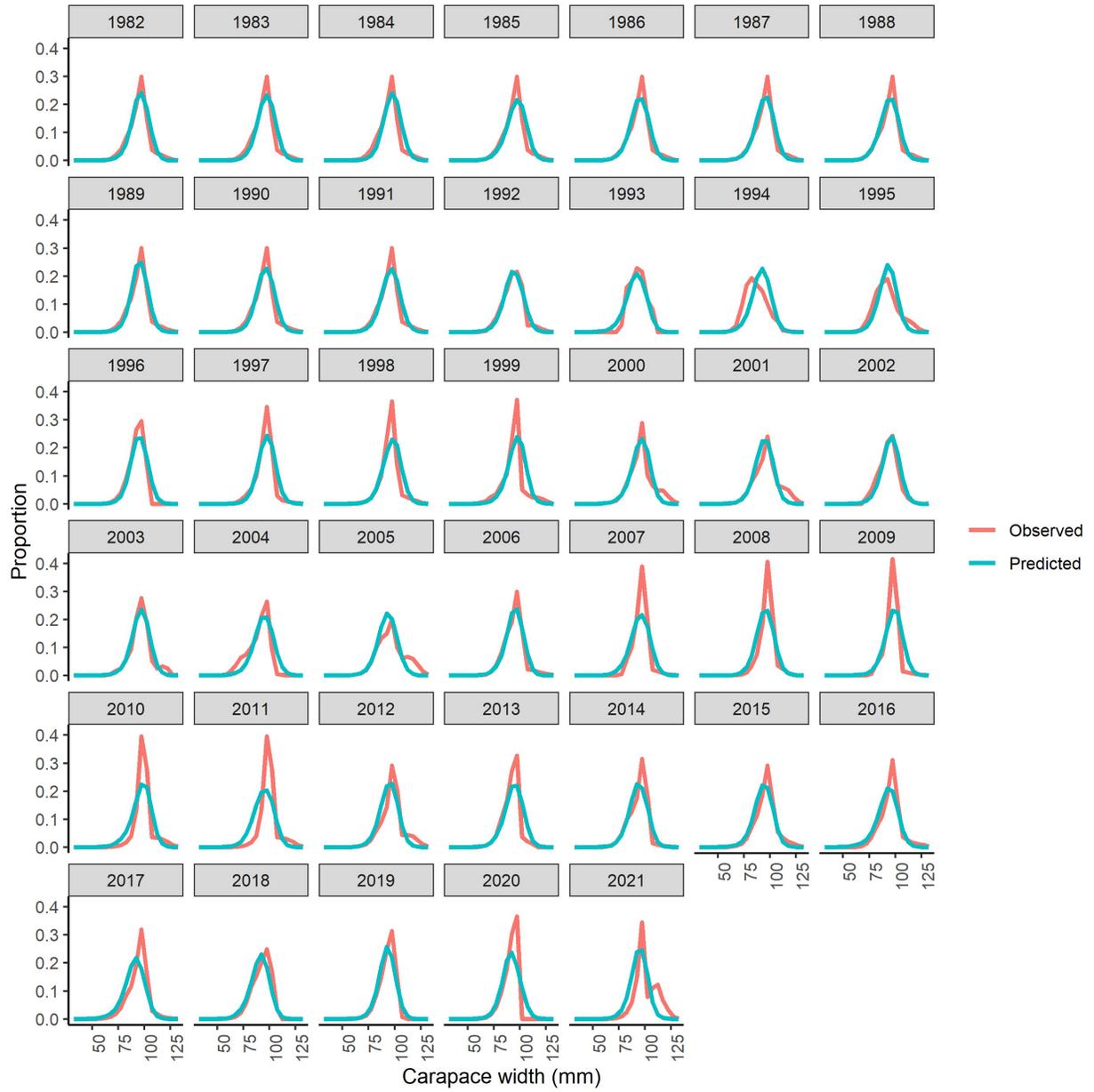
Figure S11: Model fits to individual years of size composition data for discarded crab in the directed fishery.
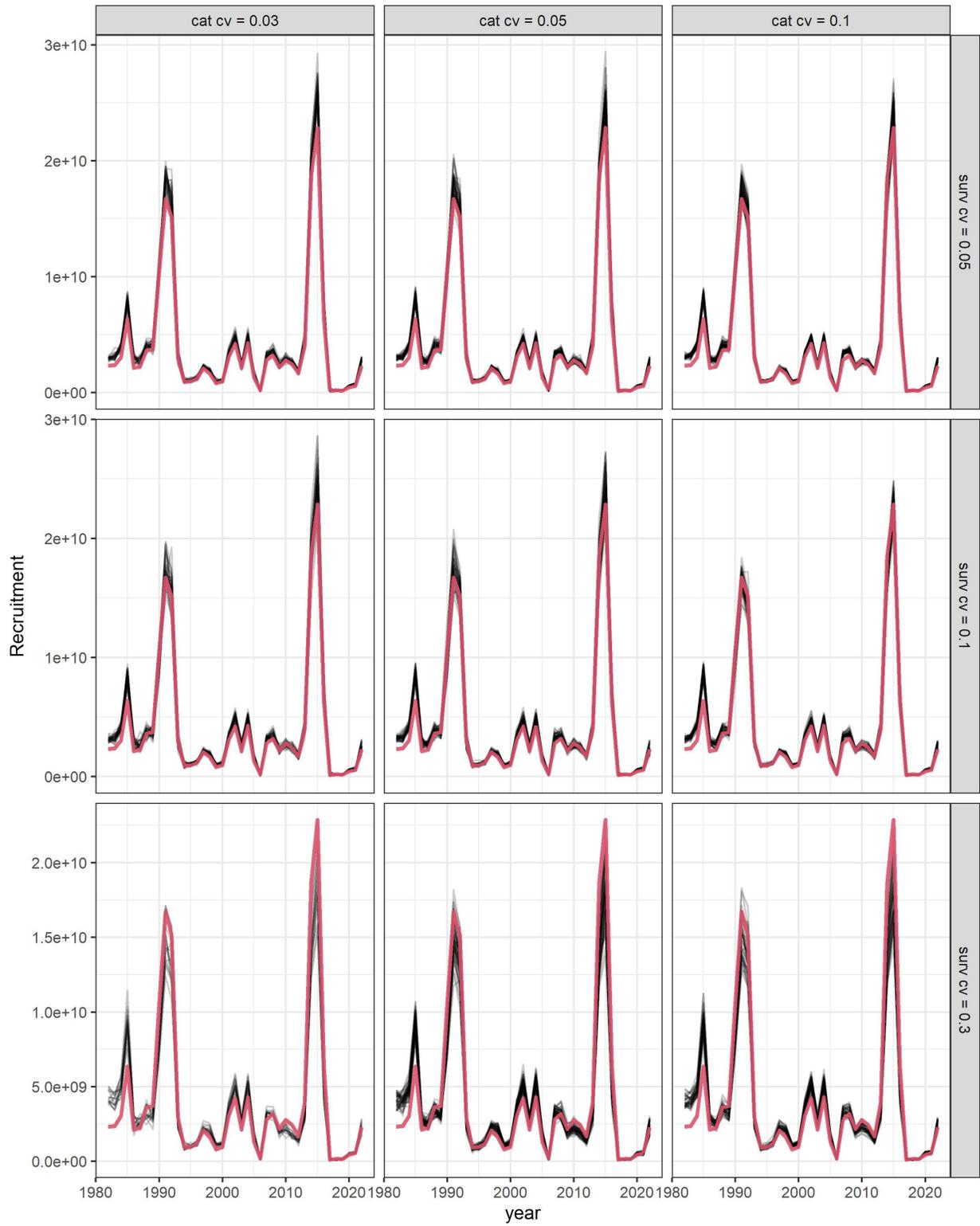
Figure S12: Estimated recruitment (black lines) vs. simulated (red line).
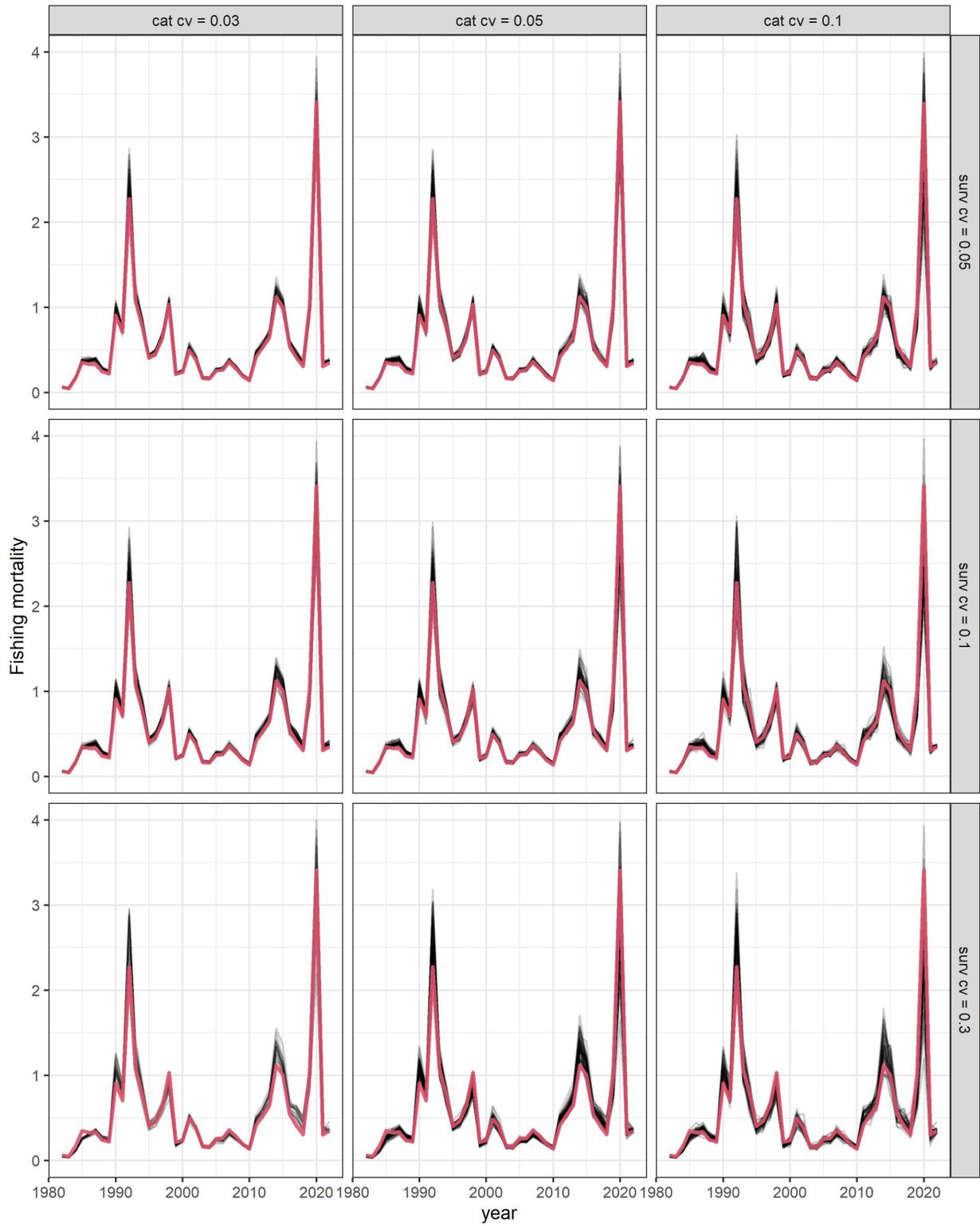
Figure S13: Estimated fishing mortality (black lines) vs. simulated (red line).
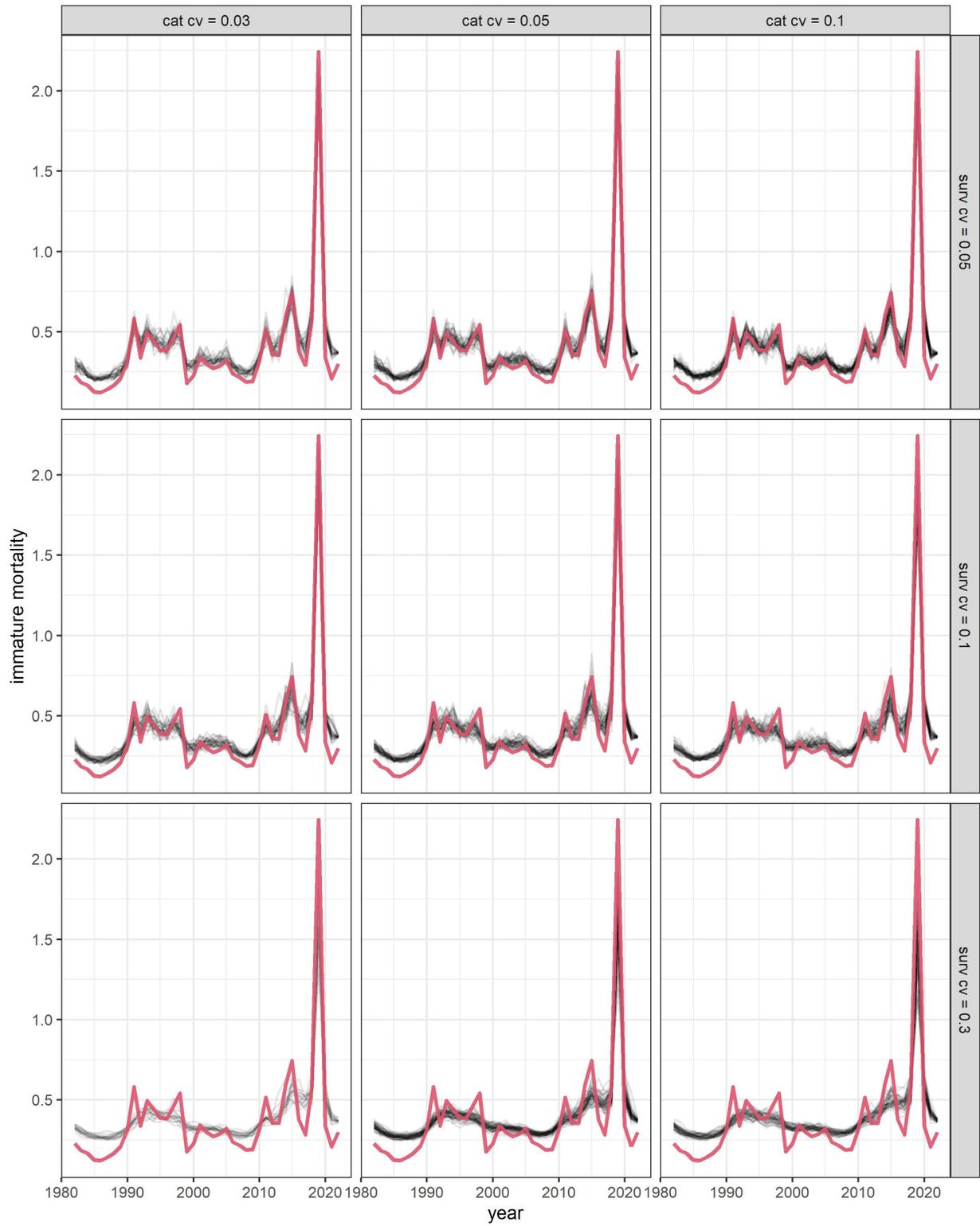
Figure S14: Estimated immature maturity (black lines) vs. simulated (red line).
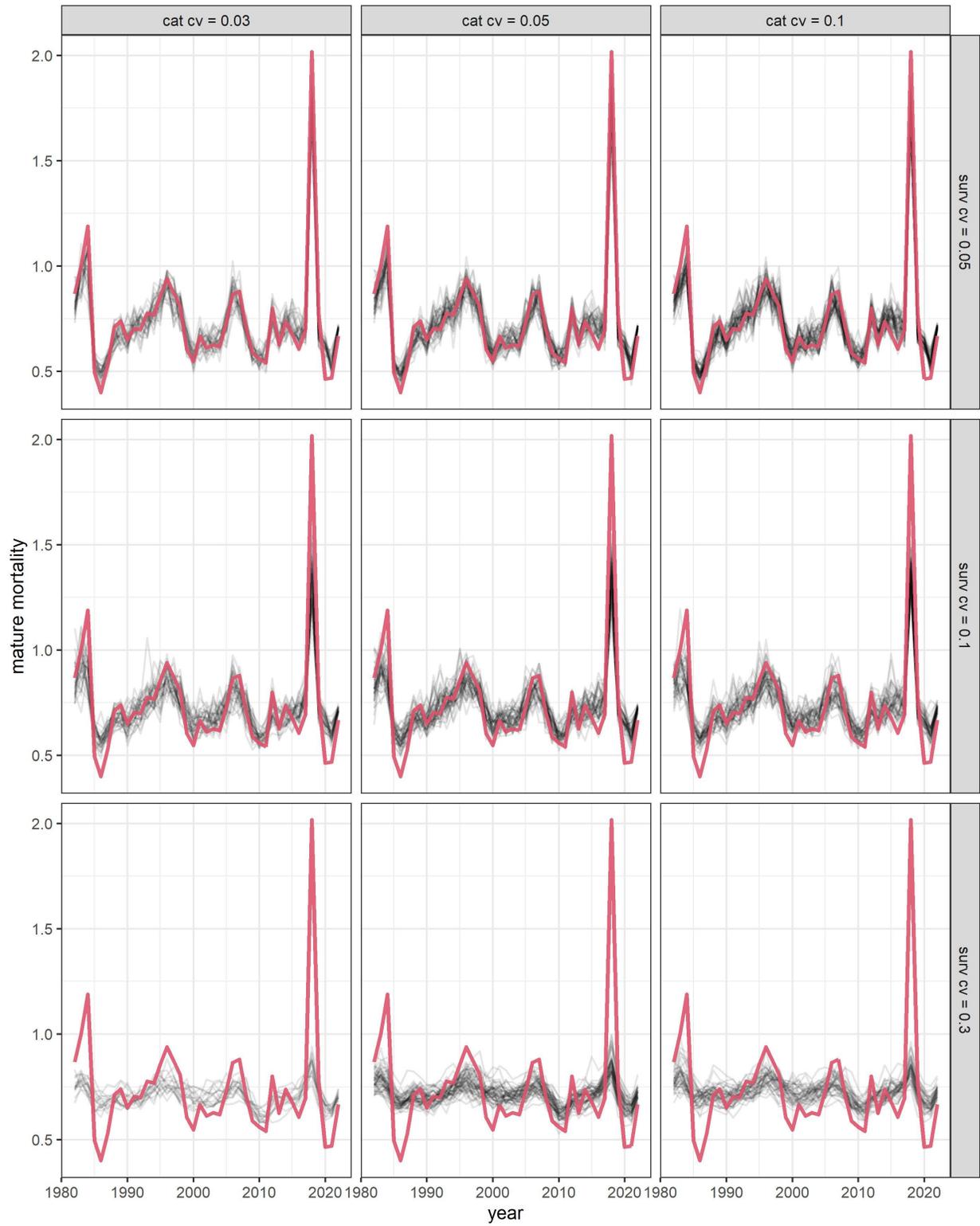
Figure S15: Estimated mature maturity (black lines) vs. simulated (red line).
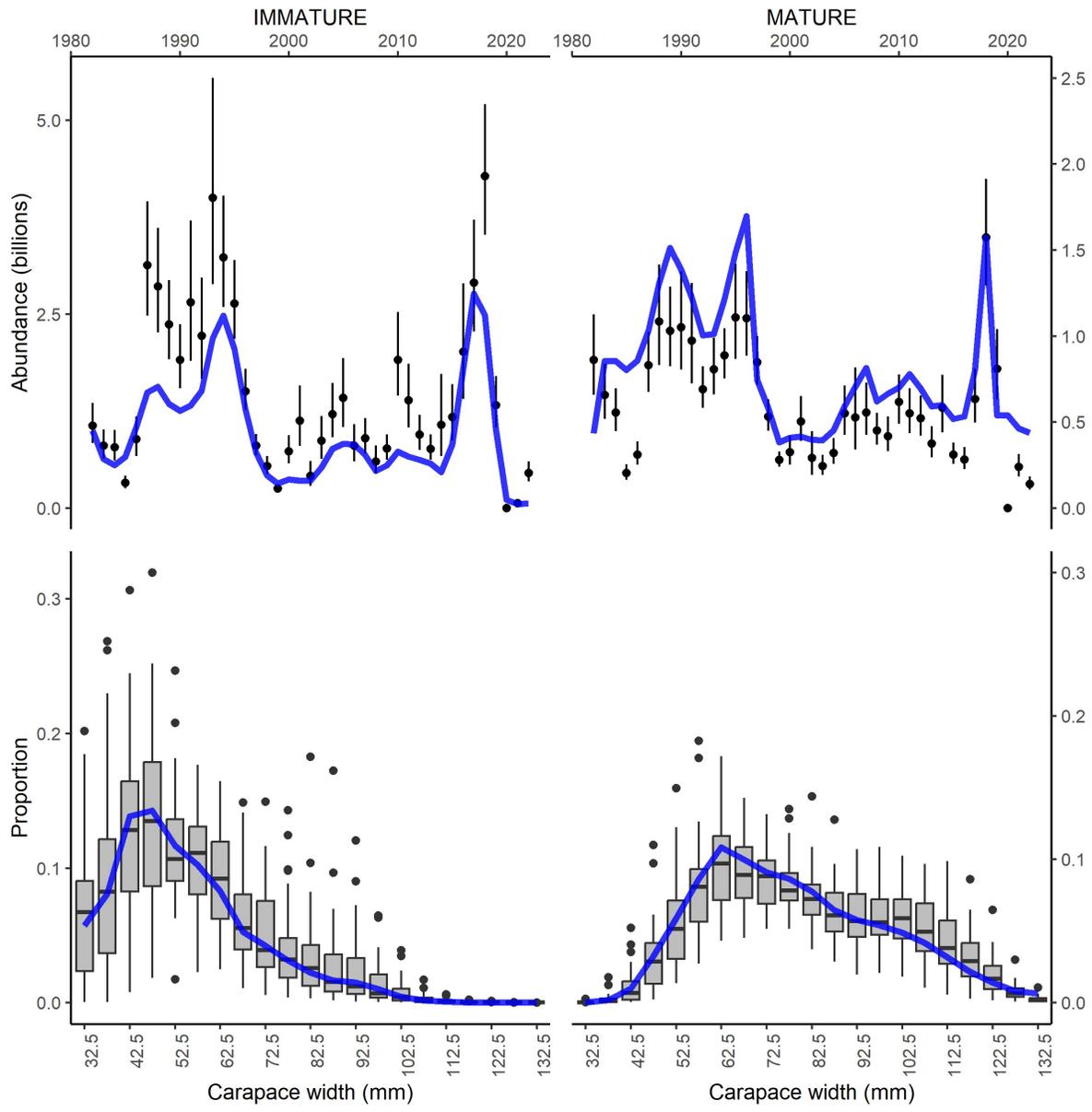
Figure S16: Fits to the survey indices of abundance and size composition data when the mean average natural mortality is not estimated.
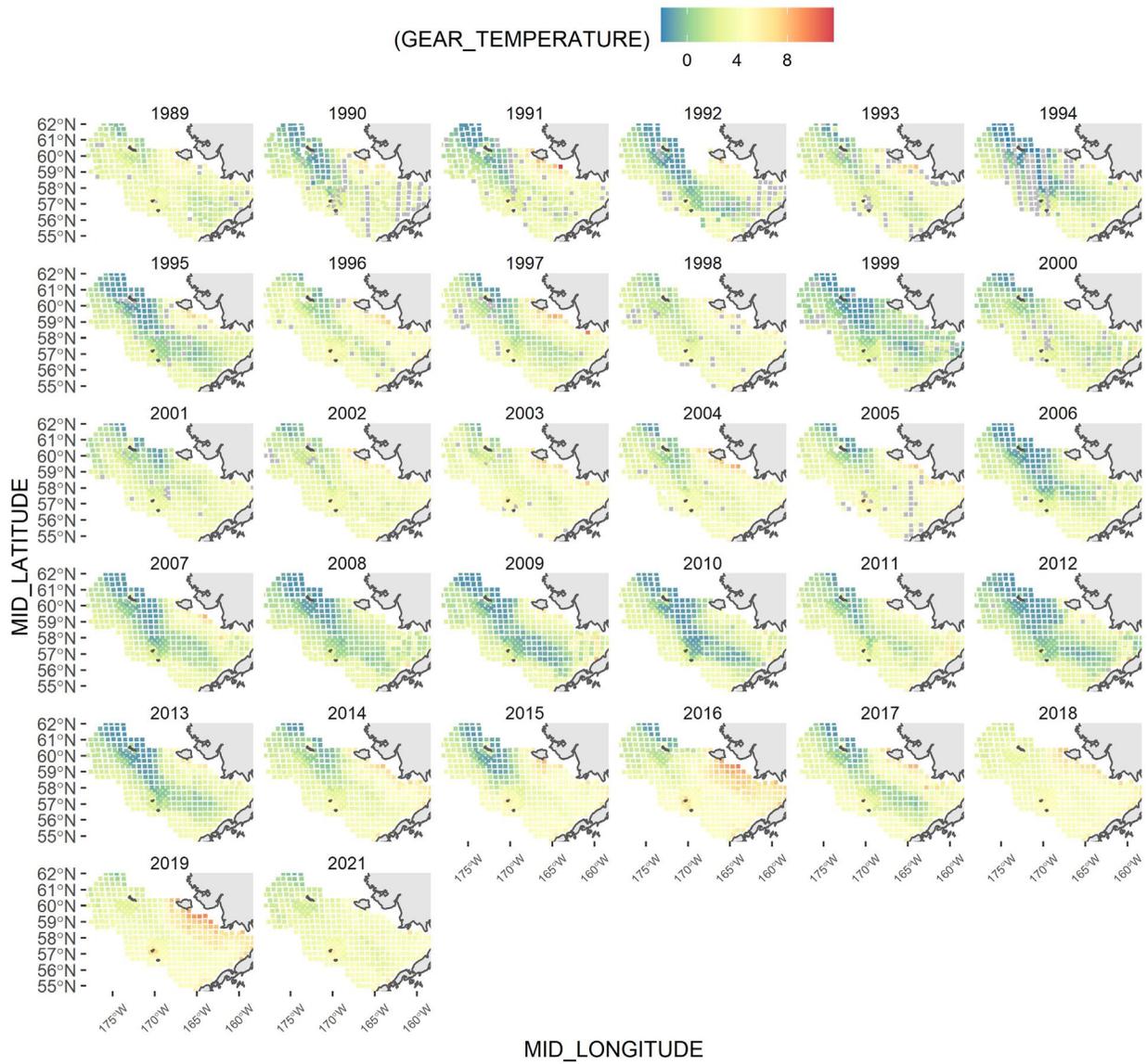
Figure S17: Bottom temperature recorded in the NMFS summer trawl survey.
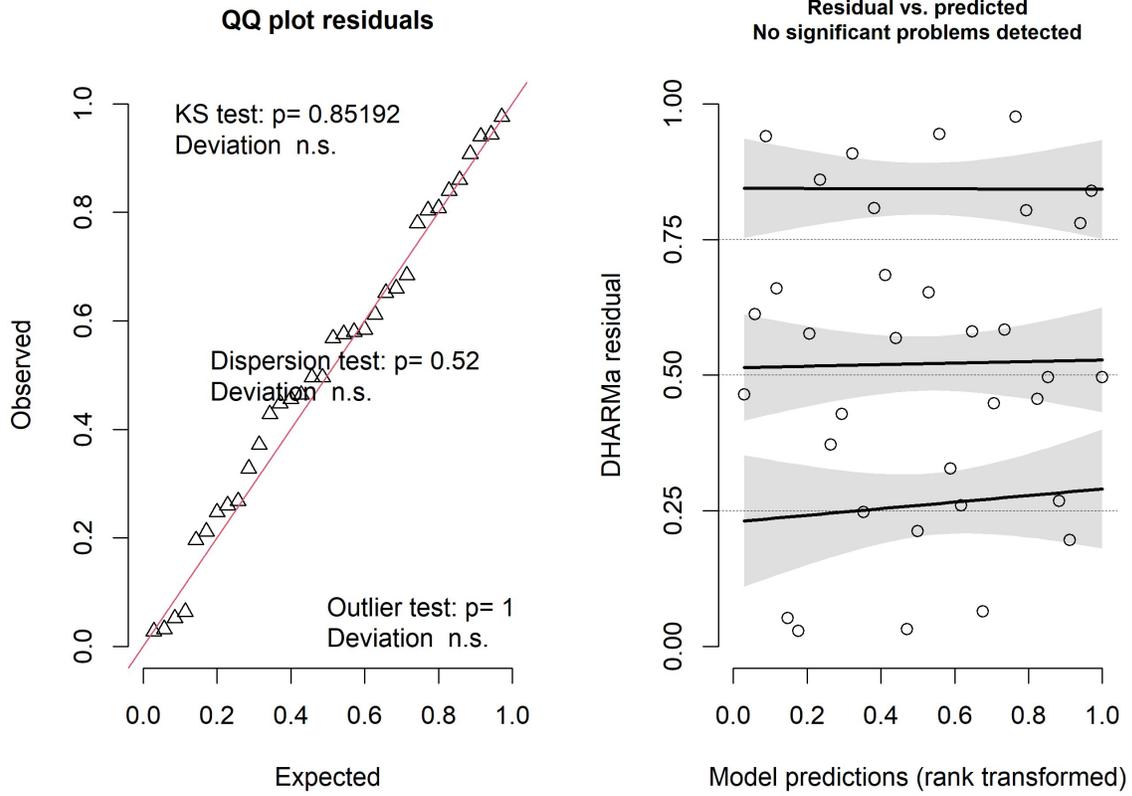
Figure S18: Diagnostics for GAMs fit to estimated recruitment. QQ plot for residuals simulated using DHARMa and quantiles of residuals vs. predicted.

DHARMa residual

QQ plot residuals

KS test: p= 0.87327
Deviation n.s.

Dispersion test: p= 0.08
Deviation n.s.

Outlier test: p= 1
Deviation n.s.

Observed

Expected

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant

DHARMa residual

Model predictions (rank transformed)
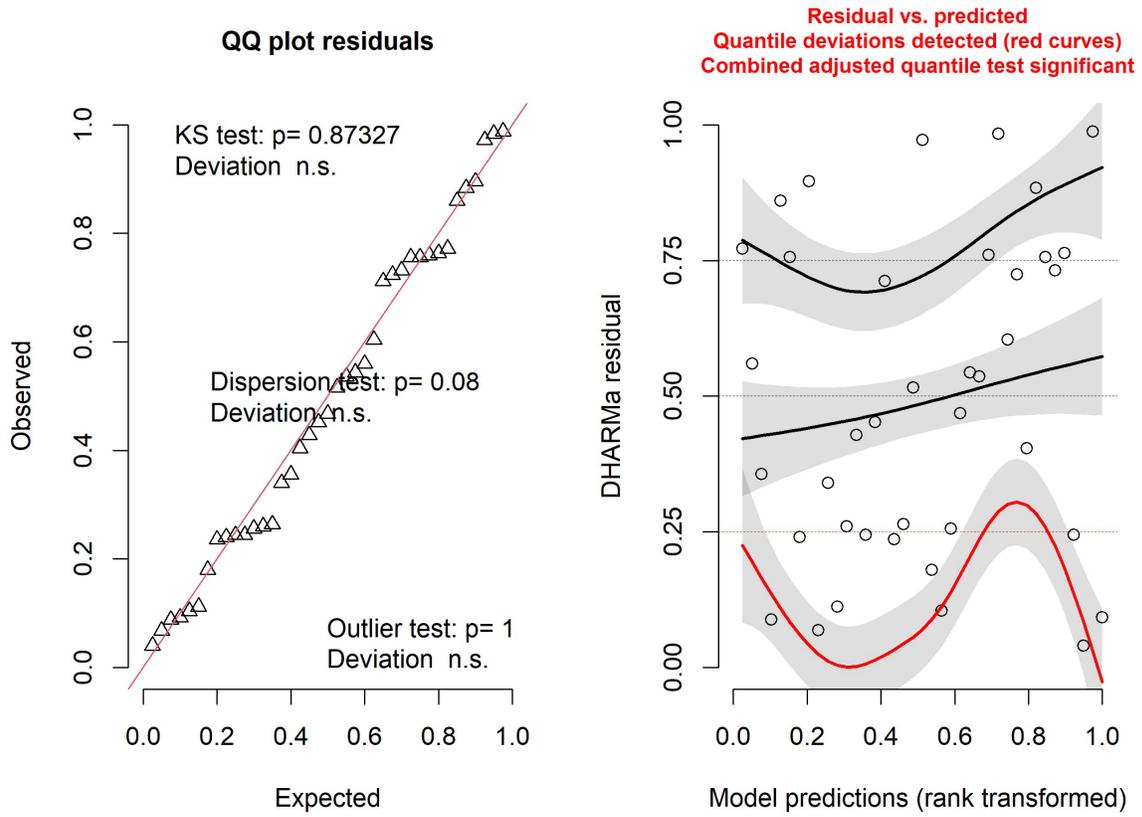
Figure S19: Diagnostics for GAMs fit to estimated immature mortality. QQ plot for residuals simulated using DHARMa and quantiles of residuals vs. predicted.

Figure S20: Diagnostics for GAMs fit to estimated mature mortality. QQ plot for residuals simulated using DHARMa and quantiles of residuals vs. predicted.

Figure S21: Diagnostics for GAMs fit to the observed probability of having undergone terminal molt. QQ plot for residuals simulated using DHARMa and quantiles of residuals vs. predicted.