

Estimating time-variation in confounded processes in population dynamics modeling: a case study for snow crab in the eastern Bering Sea

Cody Szuwalski

May 17, 2021

Population dynamics models used to provide management advice for harvested natural resources often estimate population processes that can be mutually confounded. These processes (e.g. natural mortality (M), catchability (q), growth) are often assumed to be time-invariant, but this assumption can be difficult to meet in real populations. Not allowing for time-variation in these processes can result in retrospective patterns in estimated biomass, which can result in errors in management advice. Allowing for time-variation in these processes can also result in errors in management advice if the estimated time-variation does not reflect the true underlying variation. Here I present a population dynamics model for snow crab in the eastern Bering Sea that tests different combinations of implementation of time-variation in M and q . The managing body for snow crab did not adopt any of the models with time-variation in M or q for management because the estimated variability in M and q was difficult to explain with any known mechanism and the management advice differed by nearly an order of magnitude among the models. This outcome underscores the importance of evidence-based approaches to incorporating time-variation into stock assessment models.

Introduction

Modeling the dynamics of populations often requires making simplifying assumptions to produce tenable models. For example, in fisheries stock assessment, population processes like natural mortality, catchability, growth, and maturity are often assumed to be time-invariant. However, this assumption has been shown to be false for some populations. For example, THIS, THIS, and THIS. Even for populations for which little evidence exists to demonstrate time variation in a particular process, time-variation can be suspected (REF).

Population dynamics models used to manage harvested natural resources need to strike a balance between simplicity and complexity in modeling assumptions that appropriately captures the processes important in determining sustainable harvests like natural mortality, growth, selectivity, catchability, and maturity. In general, three pieces of information are needed in harvest control rules that produce catch recommendations: biomass at the time of the next fishery, a target biomass, and a target fishing mortality. Broadly speaking, these pieces represent the current status of a fishery, the target status for a fishery, and the pathway for how to achieve the target from the current situation. Oftentimes the biomass and fishing mortality targets are based on maximum sustainable yield calculations (REF) or proxies thereof (REF).

Each of these pieces of information can be influenced by the assumptions about population processes made while developing a population dynamics model. Assuming time-invariance in the population processes that influence estimates of current biomass can produce retrospective patterns and lead to ineffective management advice if those processes vary over time (REF). Assuming time-invariance in the population processes that determine the management targets for biomass and fishing mortality may be reasonable if the forcing of these processes is stationary because the targets should represent the mean state of the system. However, if the environmental forcing of these processes is non-stationary (i.e. the mean changes over time), calculating management targets is more difficult.

Snow crab in the eastern Bering Sea present a good example of a population that appears to have time-variation in population processes and is at least at risk of non-stationarity in these population processes. The snow crab fishery only retains large male crab with carapace width greater than 101 mm. The reproductive dynamics of snow crab appear to be significantly related to ice cover and the Arctic Oscillation (Szuwalski et al., 2020). The National Marine Fishery Services (NMFS) summer trawl survey is a key information source in the stock assessment for snow crab. Large pseudocohorts have appeared three times since 1982 (??). The most recent large pseudocohort was spawned in 2010 and recruited to the survey gear in 2015. By 2018, it was the largest pseudocohort observed in the survey data (??). However, in 2019, it was much smaller than it should have been given estimated growth and natural mortality (??). This strongly implies either time-variation in catchability or natural mortality, but there are few data to decisively indicate which process is varying over time.

To further complicate the issue, there was no survey performed in 2020 as a result of the coronavirus pandemic. Here I present several iterations of the assessment method used for snow crab that differ in the way that time-variation in natural mortality and/or catchability are modeled. I then compare the resulting fits to the data sources, the character and credibility of the estimated time-variation in q and M , and the knock-on effects in other confounded processes like maturity and recruitment. I conclude by examining the differences in management-related quantities (e.g. F35% and B35%), summarizing the discussion the management body held concerning these models, and discussing potential paths forward.

Methods

The integrated size-structured model currently used (referred to here as the ‘status quo’ model) was developed following Fournier and Archibald’s (1982) methods, with many similarities to Methot (1990). The population dynamics in the status quo model tracks the number of crab of sex s , maturity state m , during year y at length l , $N_{s,m,y,l}$. A terminal molt occurs in which crab move from an immature to a mature state, after which no further molting occurs. The mid-points of the size bins tracked in the model span from 27.5 to 132.5mm carapace width, with 5 mm size classes. Parameters estimated within the assessment include

those associated with recruitment, growth, natural mortality (subject to a fairly informative prior), fishing mortality, selectivity (fishery and survey), survey catchability, and probability of maturing. Weight at length, discard mortality, bycatch mortality, and parameters associated with the variance in growth and proportion of recruitment allocated to size bin were estimated outside of the model or specified. Growth parameters are estimated outside of the model based measurements of pre-molt and post-molt carapace width for crab captured in the wild just prior to molting. Survey biomass indices and size composition data, directed fishery biomass and size composition data, and bycatch biomass and size composition data are fit to within the assessment. The assessment spans the years 1982 to 2020. See appendix A for a complete description of the population dynamics and visit the github repository for this paper to inspect the code.

The output of three models are compared to the status quo: “Vary M”, “Vary q”, and “Vary both”. Each of these models builds on the status quo model and alters the way in which natural mortality (“Vary M”), catchability (“Vary q”), or both processes (“Vary both”) are modeled. “Vary M” adds a vector of deviations to the existing estimated parameters for mature natural mortality of both sexes (i.e. a ‘dev_vector’ in ADMB). The dev_vector estimates an additional parameter for each year included in the assessment (1982-2020) for each sex. ‘Dev_vector’s are constrained to sum to zero in ADMB. “Vary q” estimates a catchability parameter for each year from 1989-2020 for each sex. The estimated vectors of q are bounded by 0.2 and 1.0 based information from experimental net efficiency work for Bering Sea snow crab that suggests catchability should be within this range (REF). “Vary both” combines both of these changes into a single model. A smoothing penalty is added to the negative log likelihood of each model in the form of the squared norm of the second difference of each vector of additionally estimated parameters multiplied by a user-defined weighting factor. The weights for the smoothing penalties were chosen by trial and error, with the aim of making them as small as possible while still avoiding convergence issues.

Retrospective analyses were performed in which the terminal year of data was removed sequentially from the model fitting, then estimated management quantities (e.g. MMB or the OFL) were compared between the model with the complete data set (i.e. 1982-2020) and successive ‘peels’ of the data to identify retrospective patterns. A retrospective pattern is a consistent directional change in assessment estimates of management quantities in a given year when additional years of data are added to an assessment. Mohn’s rho (which computes the average difference between the reference case and the peels over the period of retrospective analysis) was calculated for each model to quantify the retrospective patterns.

Reference point calculation. SBPR proxies (F35% and B35%). When time variation is included, the average over the past seven years is used in projections forward.

Results

All models produced a positive-definite Hessian and had maximum gradient components less than 0.004, except “Vary both”, which had a maximum gradient component of 0.01 for one of the recruitment deviations for males (the rest were < 0.004). The smallest viable smoothing penalties tested were 1 and 10 for natural mortality and catchability (respectively) when they were the only additional time-varying process in the assessment. When both processes were allowed to vary, the penalty for natural mortality had to be increased to 15 or the model did not converge.

Fits to the data

“Vary M”, “Vary q”, and “Vary both” all fit the survey MMB better than the status quo model (Figure 4), particularly in recent years. The status quo model missed the confidence intervals of 5 of the last 6 years of survey MMB, but additional time-variation predictably improved these fits. Fits to the catch data, growth data, and all size composition data sources can be seen in the appendix. In general, the fits were similar across models, except for the survey size composition data, in which there are several years in which models that allow for additional time-variation produce much better fits than models that do not.

Retrospective patterns in the status quo model (Mohn's $\rho = 0.36$) were improved substantially by allowing additional time-variation in either natural mortality or catchability (Figure 3).

Estimated population processes and derived quantities

Estimated population processes and derived quantities varied among models. Model "Vary both" produced the largest historical estimates of MMB (Figure 7).

All models estimated lower catchability in survey era 1 (1982-1988) relative to era 2 (1989-present) for males. The shapes of the NMFS selectivity curves were similar among all models; the largest changes were seen in the catchability coefficient (Figure 19). Status quo models that allowed at least one additional time-varying process (20.2q, 20.2m, 20.2mq) all had average catchabilities similar to the BSFRF implied catchability. However, the variability for 20.2q and 20.2mq was large, with estimated values ranging from ~ 0.2 to ~ 1 (Figure 20).

Predicted availability curves for the BSFRF experimental surveys were similar across assessments in years with similar configurations (Figure 21). The status quo assessment historically used a logistic curve for the availability for females in 2009, but this is likely overly restrictive. GMACS estimated a vector of availabilities for both years and sexes of BSFRF data, which more closely reflect the empirical availabilities.

The shape of the estimated curve representing the probability of maturing for both sexes were similar within sex, but the magnitude of the probabilities varied, most strongly for males in the 70-90 mm carapace width range (??). The estimated probability of maturing at smaller sizes was consistently higher for females in GMACS and this is related to the change from a kinked growth curve to a linear growth model. The 'hump' at 32.5 mm carapace width for females is likely related to the specified curve that determines what fraction of incoming recruitment is placed in which length bin, which has a peak at the same spot as the probability of maturing. Model 20.2mq had the highest fraction of sublegal and sub-industry-preferred crab maturing.

Estimated fishing mortality scaled with estimated population size across models (??). GMACS models generally estimated fishing mortality lower than the status quo models during survey era 1. Estimated fishery and discard selectivity were dissimilar between model type (i.e. GMACS vs. status quo), which is related to how selectivity and fishing mortality are treated in the code (see the May 2020 snow crab document for more discussion). GMACS estimates of female discard mortality were lower than the status quo, but, when balanced with changes in estimated selectivity, the estimated catches were similar to the status quo (Figure 10).

Patterns in estimated recruitment by sex were similar for both GMACS and status quo models, but GMACS estimates were more variable than the status quo estimates (??). There was a considerable amount of variability in recruitment estimates for 2015 among models, and GMACS had the highest estimates. Recruitment was larger in GMACS than the status quo model and the size of this recruitment is a strong driver of the current year MMB and OFL. When the status quo model is forced to fit the data in a similar manner to GMACS, the estimated recruitment from the status quo model is very similar to GMACS, as is the estimated OFL (175 kt; ??).

Estimated natural mortality from the GMACS model for immature crab was higher than the status quo models, in spite of identical priors (??). Estimated immature natural mortality was generally higher than mature natural mortality in GMACS, which was not seen in the status quo model for females. Estimates of mean natural mortality were smaller for models in which natural mortality was allowed to vary over time than for those in which natural mortality was constant over time. Natural mortality sharply increased in the last several years for 20.2m, but only rose sharply in the last three for 20.2mq and reaching higher mortality levels (Figure 22).

Estimated population processes Management advice

Discussion

Recap Simplicity vs. complexity Meaning of reference points in a changing environment Increasing prevalence of this sort of problem in natural resource management Strategies for confronting this issues

Changes in character of the estimated time variation in q and M when both are estimated. Changes and instability in the management advice.

Different modeling paradigms/platforms with the same data (but not identical assumptions) give different answers (Murphy). Needs for data that directly informs time variation in these processes. There are still other confounded processes: Maturing, growth, shape of survey selectivity curve.

Time-varying catchability and natural mortality are explored below in the status quo model and result in smaller retrospective patterns, but produce different management advice. Implementing any new time-varying process in an assessment with a retrospective pattern will improve the retrospective pattern, but management advice can be drastically in error if the incorrect process is allowed to vary (Szuwalski et al., 2019). Consequently, an understanding of what process is time-varying is recommended before implementation of time-variation in integrated assessments.

We know it is hard to estimate time variation in many processes with the data often available for population dynamics modeling. (Johnson et al. 2014)

Generally, fitting multiple data sources in an integrated model is preferable when data are available to parameterize the model such that variation in the modeled processes are informed sufficiently enough to be estimated. Fitting multiple data sources at once allows the uncertainty in each process to be propagated appropriately, but also requires that the model is appropriately structured to make inference about population processes. This can be difficult if processes are confounded (like growth, natural mortality, and catchability). If confounded processes are also varying over time, estimating this variation can be difficult (e.g. Johnson et al., 2014).

It is often thought that estimates of population processes can be improved by fitting multiple data sources because the information content in multiple data sources can corroborate and clarify the role of different processes in the population dynamics. This works well if the model is well-specified, but if it is not, estimating within an integrated assessment could result in incorrect inference about population processes.

One of the key uncertainties is that we do not know if the variability observed in the survey biomass is caused primarily by issues related to the survey methods (e.g. time-varying catchability or measurement error) or by issues related to population dynamics (e.g. time-varying natural mortality or other sorts of process error).

References

- Clark, W.G. 1991. Groundfish exploitation rates based on life history parameters. *Can. J. fish. Aquat. Sci.* 48: 734-750.
- Fonseca, D. B., B. Sainte-Marie, and F. Hazel. 2008. Longevity and change in shell condition of adult male snow crab *Chionoecetes opilio* inferred from dactyl wear and mark-recapture data. *Transactions of the American Fisheries Society* 137:1029-1043.
- Fournier, D.A. and C.P. Archibald. 1982. A general theory for analyzing catch-at-age data. *Can.J.Fish.Aquat.Sci.* 39:1195-1207.
- Methot, R. D. 1990. Synthesis model: An adaptable framework for analysis of diverse stock assessment data. *Int. N. Pac. Fish. Comm. Bull.* 50:259-277.
- Nevissi, A.E., J.M. Orensanz, A.J.Paul, and D.A. Armstrong. 1995. Radiometric Estimation of shell age in Tanner Crab, *Chionoecetes opilio* and *C. bairdi*, from the eastern Bering Sea, and its use to interpret indices of shell age/condition. Presented at the International symposium on biology, management and economics of crabs from high latitude habitats October 11-13, 1995, Anchorage, Alaska.
- NPFMC (North Pacific Fishery Management Council). 2007. Environmental Assessment for Amendment 24. Overfishing definitions for Bering Sea and Aleutian Islands King and Tanner crab stocks. North Pacific Fishery Management Council, Anchorage, AK, USA..
- Otto, R.S. 1998. Assessment of the eastern Bering Sea snow crab, *Chionoecetes opilio*, stock under the terminal molting hypothesis. In *Proceedings of the North Pacific Symposium on Invertebrate Stock Assessment and Management*. Edited by G.S. Jamieson and A. Campbell. *Can. Spec. Publ. Fish. Aquat. Sci.* 125. pp. 109-124.
- Paul, A.J., J.M. Paul and W.E. Donaldson. 1995. Shell condition and breeding success in Tanner crabs. *Journal of Crustacean Biology* 15: 476-480.
- Sainte-Marie, B., J. Sevigny and M. Carpentier. 2002. Interannual variability of sperm reserves and fecundity of primiparous females of the snow crab (*Chionoecetes opilio*) in relation to sex ratio. *Can.J.Fish.Aquat.Sci.* 59:1932-1940.
- Somerton, D.A. Weinberg, K.L., Goodman, S.E. 2013. Catchability of snow crab by the eastern Bering Sea bottom trawl survey estimated using a catch comparison experiment. *Can.J.Fish.Aquat.Sci.* 70: 1699-1708.
- Szuwalski, C.S., Ianelli, J.N., Punt, A.E. 2019. Reducing retrospective patterns in stock assessment and impacts on management performance. *ICES Journal of Marine Science.* 75(2): 596-609.

Appendix A: Population dynamics model description

Numbers of sex s of shell condition v and maturity state m at length l in the initial year of the assessment, $N_{s,v,m,y=1,l}$, were calculated from an estimated vector of numbers at length l by sex s and maturity state m for males, $\lambda_{s,m,l}$ and numbers at length l by sex s and shell condition v for females (i.e. 2 vectors for each sex were estimated). Estimated vectors of initial numbers at length by maturity for females were calculated by splitting the estimated vectors at length by the observed proportion mature in the first year of the survey.

$$N_{s,v,m,y=1,l} = \begin{cases} \Omega_{s,l}^{obs} \lambda_{s,1,l} & \text{if } v = \text{new}; m = \text{mat}, s = \text{fem} \\ 1 - \Omega_{s,l}^{obs} \lambda_{s,1,l} & \text{if } v = \text{new}; m = \text{imat}, s = \text{fem} \\ \lambda_{s,2,l} & \text{if } v = \text{old}; m = \text{mat}, s = \text{fem} \\ 0 & \text{if } v = \text{old}; m = \text{imat} \end{cases} \quad (1)$$

Initial numbers at length for males were all assumed to be new shell.

$$N_{s,v,m,y=1,l} = \begin{cases} \lambda_{s,1,l} & \text{if } v = \text{new}; m = \text{mat}, s = \text{male} \\ \lambda_{s,2,l} & \text{if } v = \text{new}; m = \text{imat}, s = \text{male} \\ 0 & \text{if } v = \text{old}; m = \text{mat}, s = \text{male} \\ 0 & \text{if } v = \text{old}; m = \text{imat}, s = \text{male} \end{cases} \quad (2)$$

The dynamics after the initial year were described by:

$$N_{s,v,m,y+1,l} = \begin{cases} \Omega_{s,l} \kappa_{s,l'} Q_{s,imat,y,l'} X_{s,l',l} & \text{if } v = \text{new}; m = \text{mat} \\ 1 - \Omega_{s,l} \kappa_{s,l'} Q_{s,imat,y,l'} X_{s,l',l} + Rec_y^\epsilon Pr_l & \text{if } v = \text{new}; m = \text{imat} \\ Q_{s,mat,y,l'} & \text{if } v = \text{old}; m = \text{mat} \\ (1 - \kappa_{s,l'}) Q_{s,imat,y,l'} & \text{if } v = \text{old}; m = \text{imat} \end{cases} \quad (3)$$

Where $\Omega_{s,l}$ was the probability of maturing at length l for sex s (a freely estimated vector for both males and females constrained by penalties on smoothness), $\kappa_{s,l'}$ was the probability of molting for an immature crab of sex s at length l' (set to 1 for all immature crab), and $X_{s,l',l}$ was the size transition matrix describing the probability of transitioning from size l' to size l for sex s . $Q_{s,m,y,l'}$ was the number of crab of sex s , maturity state m , and length l' surviving natural and fishing mortality during year y :

$$Q_{s,m,y,l} = \sum_v N_{s,v,m,y,l} e^{Z_{s,v,m,y,l}} \quad (4)$$

Where $N_{s,v,m,y,l}$ represented the numbers, N , of sex s during year y of shell condition v and maturity state m at length l . $Z_{s,v,m,y,l}$ represented the total mortality experienced by the population and consisted of the sum of instantaneous rates of natural mortality by sex and maturity state, $M_{s,m}$, and fishing mortality, $F_{s,f,y,l}$ from each fishery. Each fishing mortality was subject to selectivity by length l , which varied between sexes s and fisheries f (and by year y if specified). $M_{s,m}$ was specified in the model and a multiplier $\gamma_{natM,m}$ was estimated subject to constraints (see this formulation effectively specified a mean and standard deviation for a prior distribution for M).

$$Z_{s,v,m,y,l} = \gamma_{natM,m} M_{s,m} + \sum_f S_{s,f,y,l} F_{s,f,y,l} \quad (5)$$

Selectivities in the directed and bycatch fisheries were estimated logistic functions of size. Different selectivity parameters were estimated for females and males in the directed fisheries ($S_{fem,dir,l}$ and $S_{male,dir,l}$, respectively), a single selectivity for both sexes was estimated for bycatch in the groundfish trawl fishery ($S_{trawl,l}$), and a retention selectivity was estimated for the directed fishery for males ($R_{dir,l}$; all females were discarded).

$$S_{male,dir,l} = \frac{1}{1 + e^{-S_{slope,m,d}(L_l - S_{50,m,d})}} \quad (6)$$

$$S_{fem,dir,l} = \frac{1}{1 + e^{-S_{slope,f,d}(L_l - S_{50,f,d})}} \quad (7)$$

$$S_{trawl,l} = \frac{1}{1 + e^{-S_{slope,t}(L_l - S_{50,t})}} \quad (8)$$

$$R_{dir,l} = \frac{1}{1 + e^{-S_{slope,m,d}(L_l - S_{50,m,d})}} \quad (9)$$

Where $S_{slope,s,f}$ was the slope of the logistic curve for sex s in fishery f and $S_{50,s,f}$ was the length at 50% selection for sex s in fishery f . Catches for all fisheries were modeled as pulse fisheries in which all catch was removed instantaneously (i.e. no natural mortality occurred during the fishery). Catch in fishery f during year y was calculated as the fraction of the total fishing mortality, $F_{s,f,y,l}$, applied to a given sex s in a fishery f times the biomass removed by all fisheries for that sex.

$$C_{male,dir,y} = \sum_l \sum_v \sum_m w_{male,l} \frac{R_l F_{male,dir,y,l}}{F_{male,dir,y,l} + F_{trawl,y,l}} N_{male,v,m,y,l} e^{-\delta_y M_{s,m}} (1 - e^{-(F_{male,dir,y,l} + F_{trawl,y,l})}) \quad (10)$$

$$C_{male,tot,y} = \sum_l \sum_v \sum_m w_{male,l} \frac{F_{male,dir,y,l}}{F_{male,dir,y,l} + F_{trawl,y,l}} N_{male,v,m,y,l} e^{-\delta_y M_{s,m}} (1 - e^{-(F_{male,dir,y,l} + F_{trawl,y,l})}) \quad (11)$$

$$C_{fem,dir,y} = \sum_l \sum_v \sum_m w_{fem,l} \frac{F_{fem,dir,y,l}}{F_{fem,dir,y,l} + F_{trawl,y,l}} N_{fem,v,m,y,l} e^{-\delta_y M_{s,m}} (1 - e^{-(F_{fem,dir,y,l} + F_{trawl,y,l})}) \quad (12)$$

$$C_{m+f,trawl,y} = \sum_s \sum_l \sum_v \sum_m w_{s,l} N_{s,v,m,y,l} e^{-\delta_y M_{s,m}} (1 - e^{-(F_{trawl,y,l})}) \quad (13)$$

Where δ_y was the mid point of the fishery (all fisheries were assumed to occur concurrently and the midpoint was based on the directed fishery, which accounts for the vast majority of the fishing mortality) and $w_{s,l}$ was the weight at length l for sex s . Trawl data and discard data were entered into the model with an assumed mortality of 80% and 30%, respectively. Fully-selected fishing mortality parameters for fishery f were estimated as a logged average over a given time period (F_{avg}^{log}) with yearly deviations around that mean ($F_{dev,y}^{log}$).

$$F_{f,y} = e^{(F_{avg,f}^{log} + F_{dev,f,y}^{log})} \quad (14)$$

Selectivity for the survey was estimated for 2 eras in the base model: 1982-1988 and 1989-present. Selectivity was assumed to be logistic and separate parameters representing the length at which selection probability

equal 50% and 95% ($s_{50,s,e}$ and $s_{95,s,e}$, respectively) were estimated for males and females in the third era (1989-present). Separate catchability coefficients ($q_{s,e}$) were estimated for males and females in all eras.

$$S_{surv,s,l,e} = \frac{q_{s,e}}{1 + e^{-\log(19) \frac{L_l - s_{50,s,e}}{s_{95,s,e} - s_{50,s,e}}}} \quad (15)$$

Survey selectivity was informed by experimental surveys during the years 2009 and 2010. A portion of the NMFS summer survey tows were accompanied by an industry vessel using nephrops trawls with an assumed selectivity of 1 for all size classes. To represent the proportion of the population covered by the experiment, a vector was freely estimated for males, S_y^{free} (subject to a scaling parameter), and a logistic curve was estimated for females.

$$S_{ind,s,l,y} = \begin{cases} \frac{q_{ind,s,y}}{1 + e^{-\log(19) \frac{L_l - s_{50,s,y}}{s_{95,s,y} - s_{50,s,y}}}} & \text{if } s = \text{female} \\ q_{ind,s,y} S_y^{free} & \text{if } s = \text{male} \end{cases} \quad (16)$$

Based on this logic, after identifying the fraction of the crab at length covered by the experimental surveys, the length frequencies of the NMFS data collected simultaneously with the experimental trawls can be calculated by multiplying the numbers at length ‘available’ to the experimental trawls by the overall survey selectivity, $S_{surv,s,l,y}$. The predicted numbers at length for the NMFS and industry data from the selectivity experiment were calculated by multiplying the respective selectivities by the survey numbers at length.

$$S_{nmfs,s,l,y} = S_{ind,s,l,y} S_{surv,s,l,y} \quad (17)$$

Mature male and female biomass (MMB and FMB, respectively) were fitted in the objective function and were the product of mature numbers at length during year y and the weight at length, $w_{s,l}$:

$$MMB_y = \sum_{l,v} w_{male,l} N_{male,v,mat,y,l} \quad (18)$$

$$FMB_y = \sum_{l,v} w_{fem,l} N_{fem,v,mat,y,l} \quad (19)$$

$$w_{s,l} = \alpha_{wt,s} L_l^{\beta_{wt,s}} \quad (20)$$

Mature biomass can be calculated for different time through out the year, in which case the numbers at length are decremented by the estimated natural mortality. Parameters $\alpha_{wt,s}$ and $\beta_{wt,s}$ were estimated outside of the assessment model and specified in the control file.

Molting and growth occur before the survey. Immature crab were assumed to molt every year with an estimated probability of molting to maturity based on length l (in all the scenarios presented here, the probability of molting was 1 for all immature animals). For crab that do molt, the growth increment within the size-transition matrix, $X_{s,l,l'}$, was based on a piece-wise linear relationship between predicted pre- and post-molt length, ($\hat{L}_{s,l}^{pred}$ and $\hat{L}_{s,l}^{post}$, respectively) and the variability around that relationship was characterized by a discretized and renormalized gamma function, $Y_{s,l,l'}$.

$$X_{s,l,l'} = \frac{Y_{s,l,l'}}{\sum_{l'} Y_{s,l,l'}} \quad (21)$$

$$Y_{s,l,l'} = (\Delta_{l,l'})^{\frac{L_{s,l} - (\bar{L}_l - 2.5)}{\beta_s}} \quad (22)$$

$$\hat{L}_{s,l}^{post,1} = \alpha_s + \beta_{s,1} L_l \quad (23)$$

$$\hat{L}_{s,l}^{post,2} = \alpha_s + \delta_s(\beta_{s,1} - \beta_{s,2}) + \beta_{s,2}L_l \quad (24)$$

$$\hat{L}_{s,l}^{post} = \hat{L}_{s,l}^{post,1}(1 - \Phi(\frac{L_l - \delta_{a,x}}{stgr})) + \hat{L}_{s,l}^{post,2}(\Phi(\frac{L_l - \delta_{a,x}}{stgr})) \quad (25)$$

$$\Delta_{l,l'} = \bar{L}_{l'} + 2.5 - L_l \quad (26)$$

$\hat{L}_{s,l}^{post,1}$ and $\hat{L}_{s,l}^{post,2}$ were predicted post-molt lengths from each piece of the piece-wise relationship, and $\Phi()$ was a cumulative normal distribution in which $\delta_{a,x}$ was an estimated change point. The model in which linear growth was estimated removed equations 26 and 27 from the model.

An average recruitment for the assessment period (1982-present) and yearly deviations around this average were estimated within the assessment for models in which only a single vector of recruitment deviations was estimated. The sex ratio of recruitment was assumed to be 50/50 male to female. Each year's estimated recruitment was allocated to length bins based on a discretized and renormalized gamma function with parameters specified in the control file.

$$Rec_y = e^{(Rec_{avg} + Rec_{dev,y})} \quad (27)$$

$$Pr_l = \frac{(\Delta_{1,l})^{\alpha_{rec}/\beta_{rec}} e^{-\Delta_{1,l}/\beta_{rec}}}{\sum_{l'} (\Delta_{1,l'})^{\alpha_{rec}/\beta_{rec}} e^{(-\Delta_{1,l'}/\beta_{rec})}} \quad (28)$$

For models in which separate vectors of recruitment deviations were estimated for males and females, a separate average recruitment was also estimated (in log space). Each vector of deviations was also subject to a smoothing penalty, but were not linked directly in any way (e.g. priors on the ratio of estimated male to female average recruitment).

Three general types of likelihood components were used to fit to the available data. Multinomial likelihoods were used for size composition data, log-normal likelihoods were used for indices of abundance data, and normal likelihoods were used for catch data, growth data, priors, and penalties. Multinomial likelihoods were implemented in the form:

$$L_x = \lambda_x \sum_y N_{x,y}^{eff} \sum_l p_{x,y,l}^{obs} \ln(\hat{p}_{x,y,l}/p_{x,y,l}^{obs}) \quad (29)$$

L_x was the likelihood associated with data component x , where λ_x represented an optional additional weighting factor for the likelihood, $N_{x,y}^{eff}$ was the effective 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 $\hat{p}_{x,y,l}$ was the predicted proportion in size bin l during year y for data component x .

Log normal likelihoods were implemented in the form:

$$L_x = \lambda_x \sum_y \frac{(\ln(\hat{I}_{x,y}) - \ln(I_{x,y}))^2}{2(\ln(CV_{x,y}^2 + 1))} \quad (30)$$

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

Normal likelihoods were implemented in the form:

$$L_x = \lambda_x \sum_y (\hat{I}_{x,y} - I_{x,y})^2 \quad (31)$$

L_x was the contribution to the objective function of data component x , λ_x was represents the weight applied to the data component (and can be translated to a standard deviation), $\hat{I}_{x,y}$ was the predicted value of quantity I from data component x during year y , $I_{x,y}$ was the observed value of quantity I from data component x during year y .

Smoothing penalties were also placed on some estimated vectors of parameters in the form of normal likelihoods on the second differences of the vector with a user-specified weight.

Table 1: Changes in management quantities for each scenario considered. Reported management quantities are derived from maximum likelihood estimates. Reported natural mortality is for mature males and average recruitment is for males.

Model	MMB	B35	F35	FOFL	OFL	M	avg_rec
Status quo	133.51	121.47	1.23	1.23	88.90	0.29	103.91
Vary q	121.61	137.56	1.94	1.94	77.08	0.30	132.86
Vary M	43.29	17.85	6.29	6.29	70.88	0.59	152.61
Vary both	92.20	28.06	12.46	1.86	14.72	0.58	241.96

Table 2: Contribution to the objective function by individual likelihood component by modeling scenario. Values in columns after Model 0 are the likelihood contribution of Model 0 minus the likelihood contribution of the model in the column. Positive values represent improvements in fit. Note that some of the model scenarios involve changing the weightings of data sources which invalidate the comparison of likelihoods for a data source among models.

Likelihood component	Status quo	Vary q	Vary M	Vary both
Recruitment deviations	71.12	70.7	62.97	69.83
Initial numbers old shell males	2.83	2.86	2.81	2.59
small length bins				
ret fishery length	194.07	189.63	182.27	180.54
total fish length (ret + disc)	560.36	554.67	546.02	536.71
female fish length	127.32	128.3	129.98	126.79
survey length	2298.23	2275.79	2208.05	2228.97
trawl length	169.34	164.05	151.31	165.36
2009 BSFRF length	-45.58	-45.9	-46.92	-44.82
2009 NMFS study area length	-36.62	-37	-37.86	-35.78
M multiplier prior	34.55	40.78	11.44	13.14
maturity smooth	45.6	42.6	29.55	34.71
growth males	0	0	0	0
growth females	0	0	0	0
2009 BSFRF biomass	0.49	0.03	0.09	0.15
2009 NMFS study area biomass	0.34	0.04	0.06	0.08
cpue q	0.4	0.48	0.19	0.56
retained catch	6.02	3.91	1.21	1.69
discard catch	106.68	116.94	41.42	46.54
trawl catch	11.64	5.62	7.85	5.53
female discard catch	4	4.67	5.71	7.03
survey biomass	215.85	126.53	79.96	77.08
F penalty	23.97	21.48	23.83	22.32
2010 BSFRF Biomass	5.53	2.93	3.4	1.56
2010 NMFS Biomass	6.42	1.85	3.55	4.14
Extra weight survey lengths	273.98	274.43	267.88	265.7
first year				
2010 BSFRF length	-24.2	-23.6	-22.6	-23.08
2010 NMFS length	-30.65	-29.33	-30.84	-31.08
smooth selectivity	1.22	1.3	1.26	1.22

Likelihood component	Status quo	Vary q	Vary M	Vary both
smooth female selectivity	0	0	0	0
init nos smooth constraint	32.25	32.84	31.03	31.17
Total	4055.16	3926.6	3653.62	3688.65

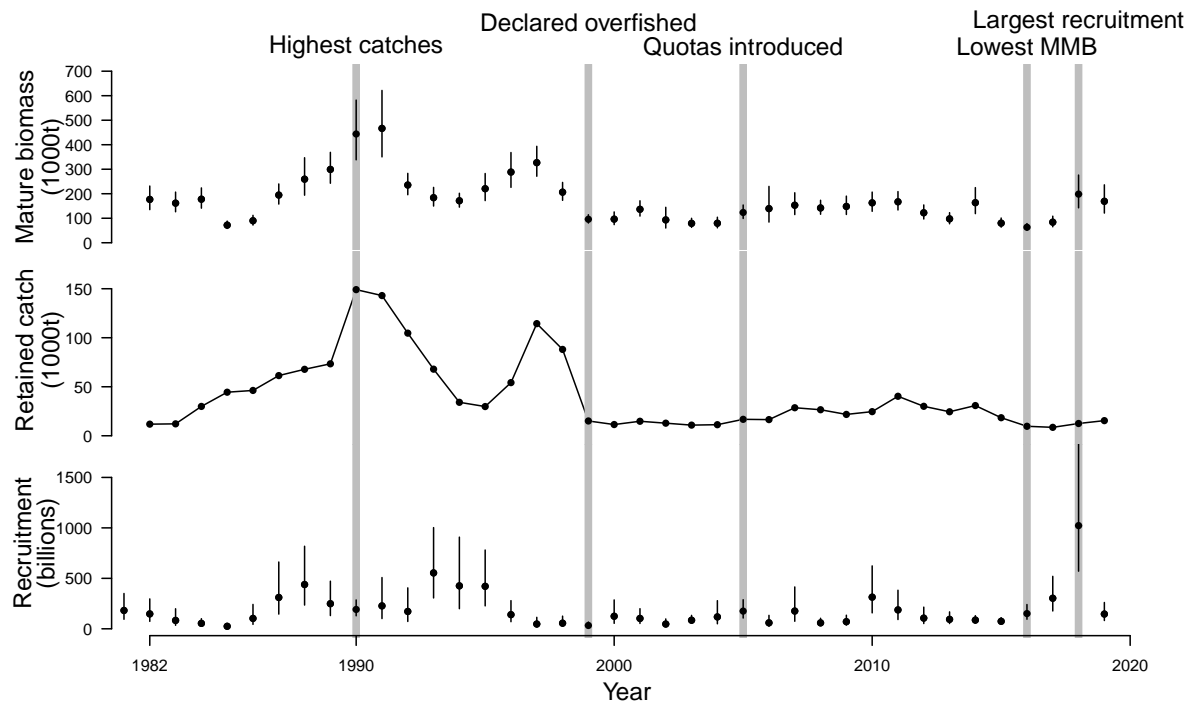


Figure 1: History of the snow crab fishery in the eastern Bering Sea.

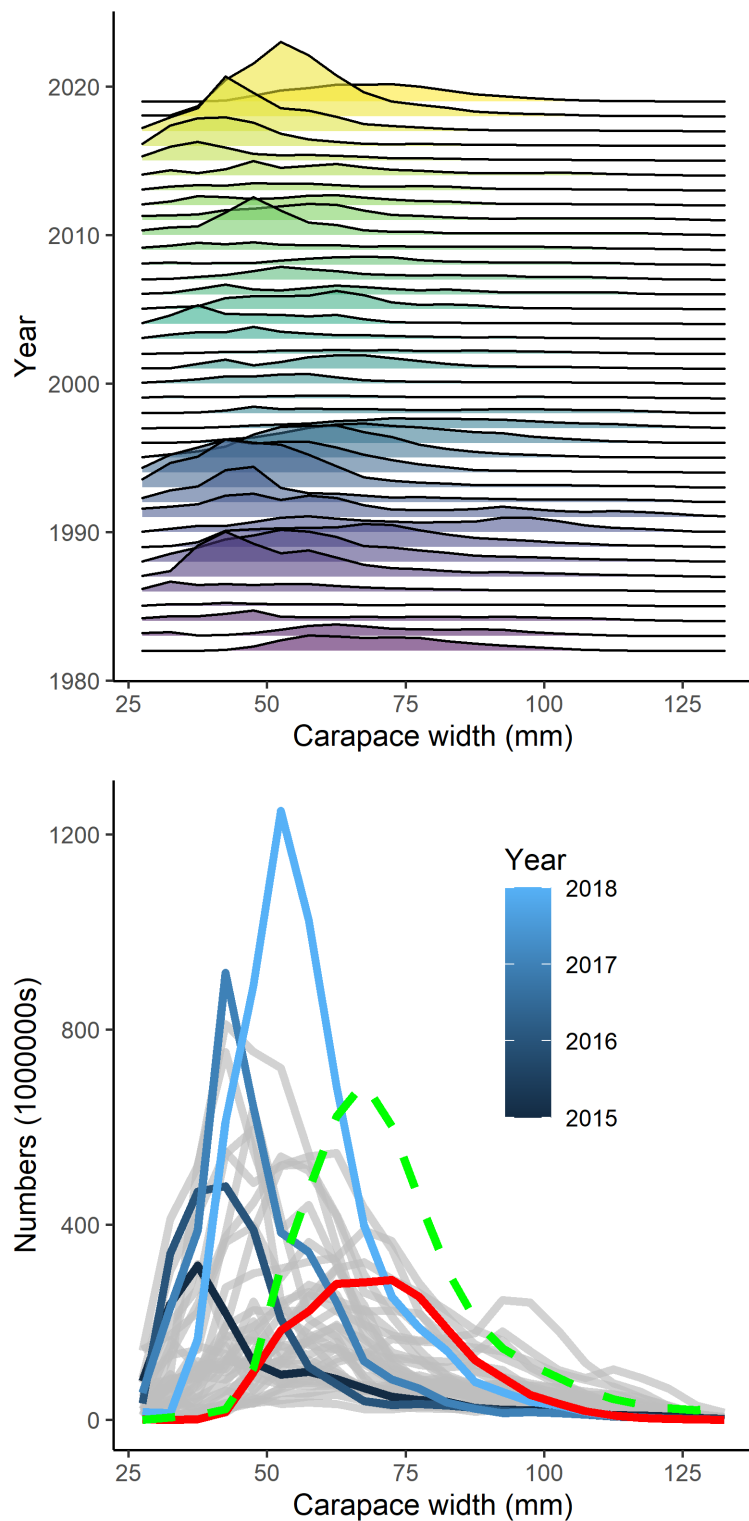


Figure 2: Numbers at length from the NMFS summer trawl survey. Top figure shows the relative numbers at length over time in which the height of the polygon for a given year represents the number of crab observed in that size class. Bottom figure shows the same data, but overlaid. Highlighted years show the recruitment of the 2010 cohort to the survey gear in 2015 and the subsequent 'disappearance' of the cohort in 2019 (red line). Dashed green line represents the approximate expected numbers at length given estimated growth and natural mortality.

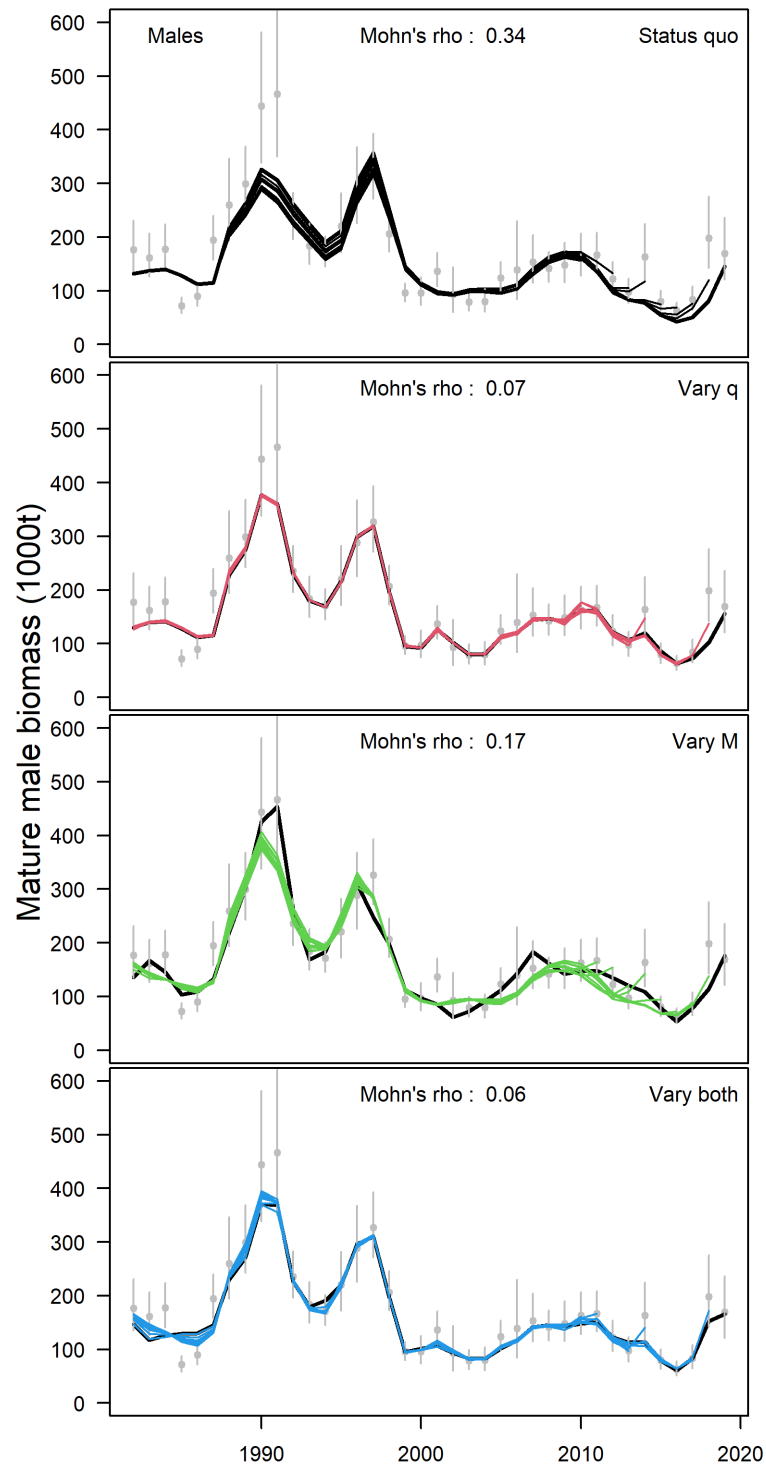


Figure 3: Retrospective patterns in mature male biomass for all models in the analysis.

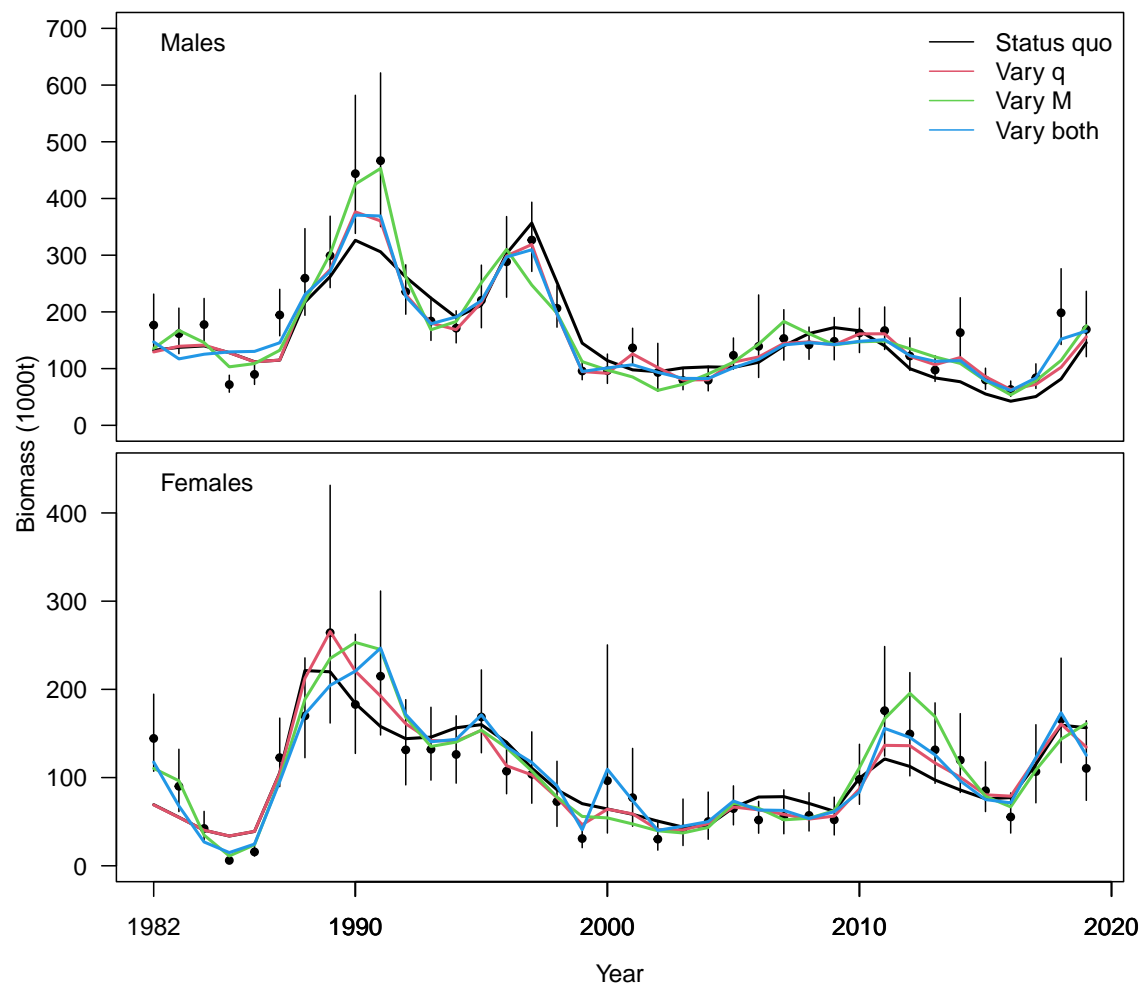


Figure 4: Model fits to the observed mature biomass at survey

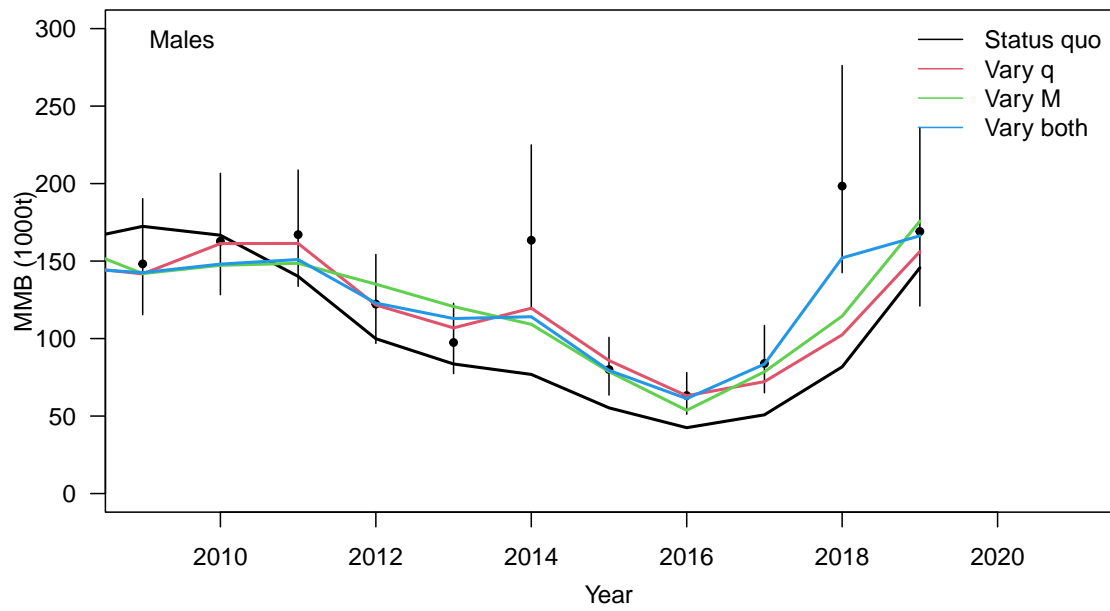


Figure 5: Model fits to the observed mature biomass at survey 2009-present

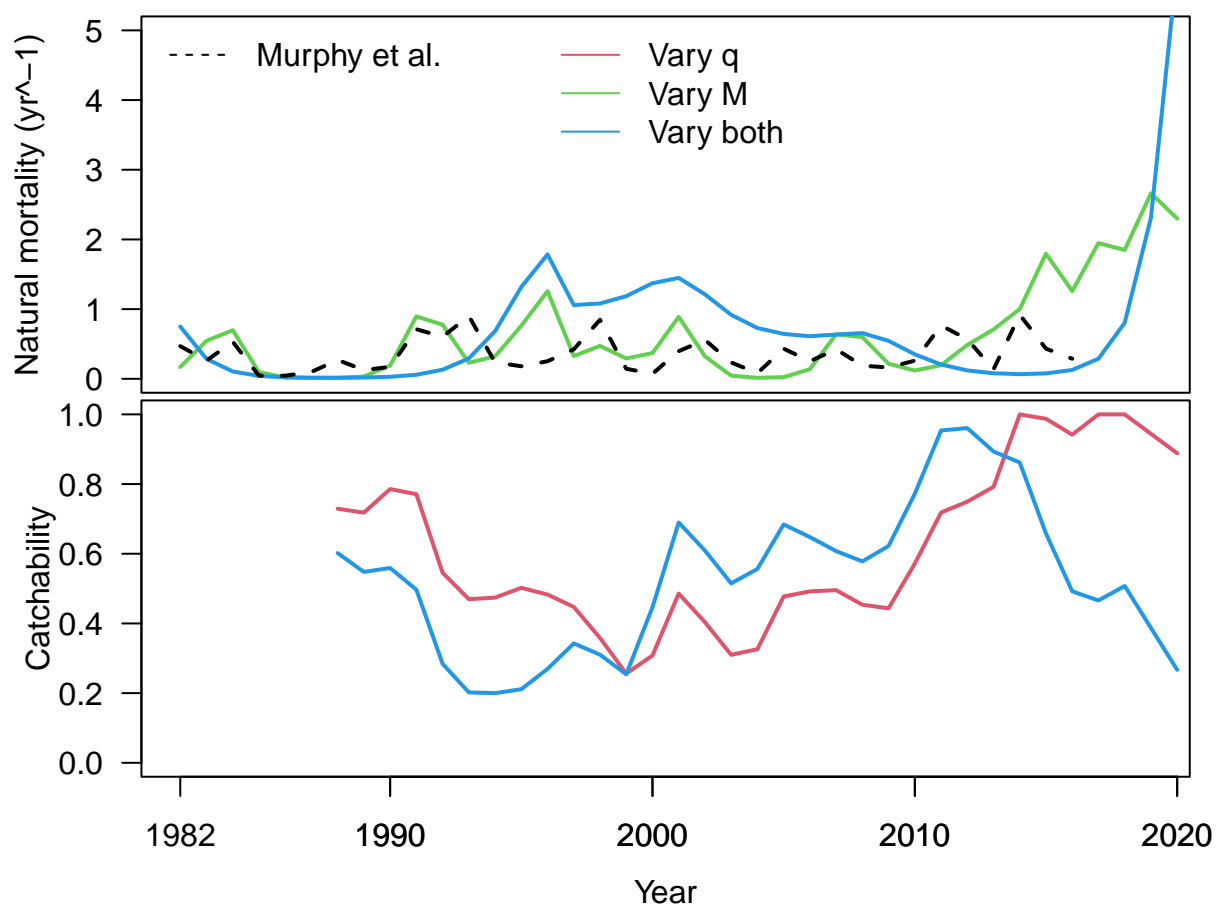


Figure 6: Estimated time varying natural mortality and catchability for mature males.

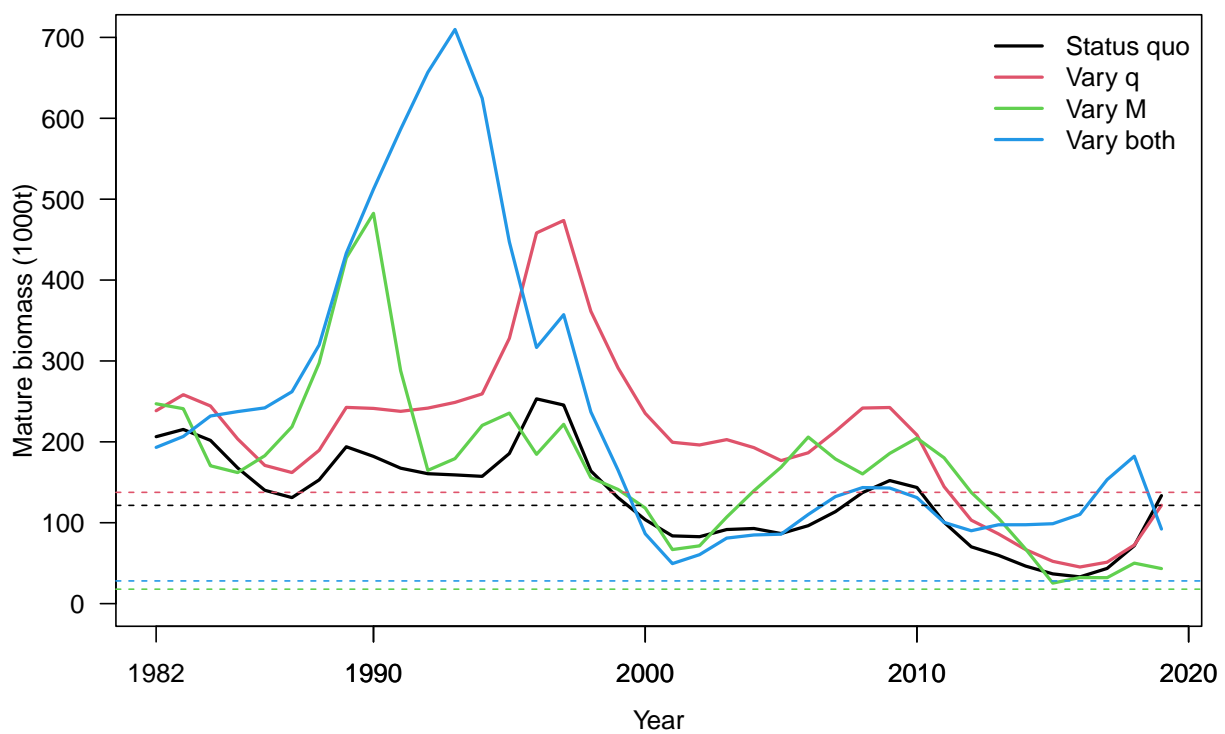


Figure 7: Model predicted mature biomass at mating time. Dotted horizontal lines are target biomasses.

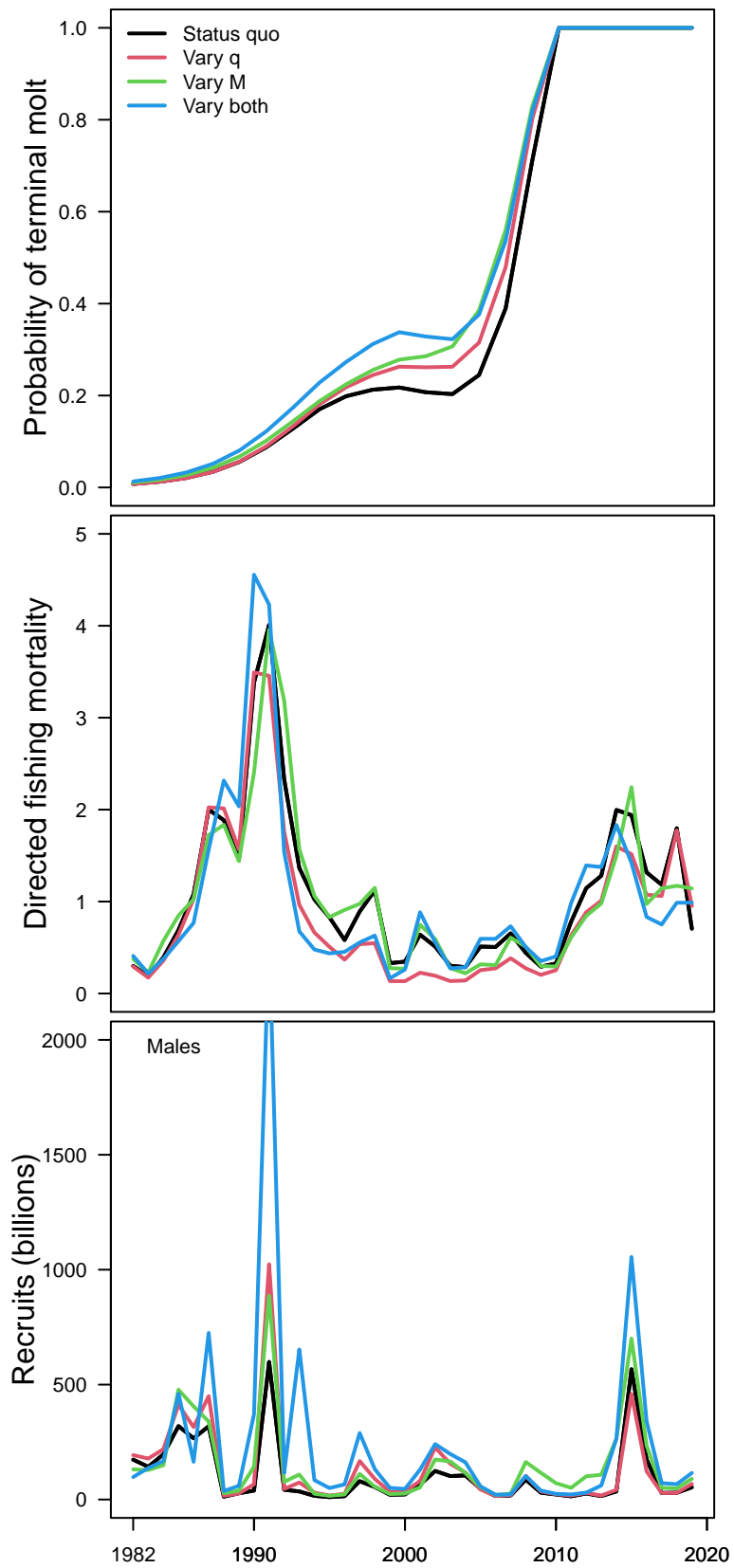


Figure 8: Estimated probability of maturing, directed fishing mortality, and recruitment for males.

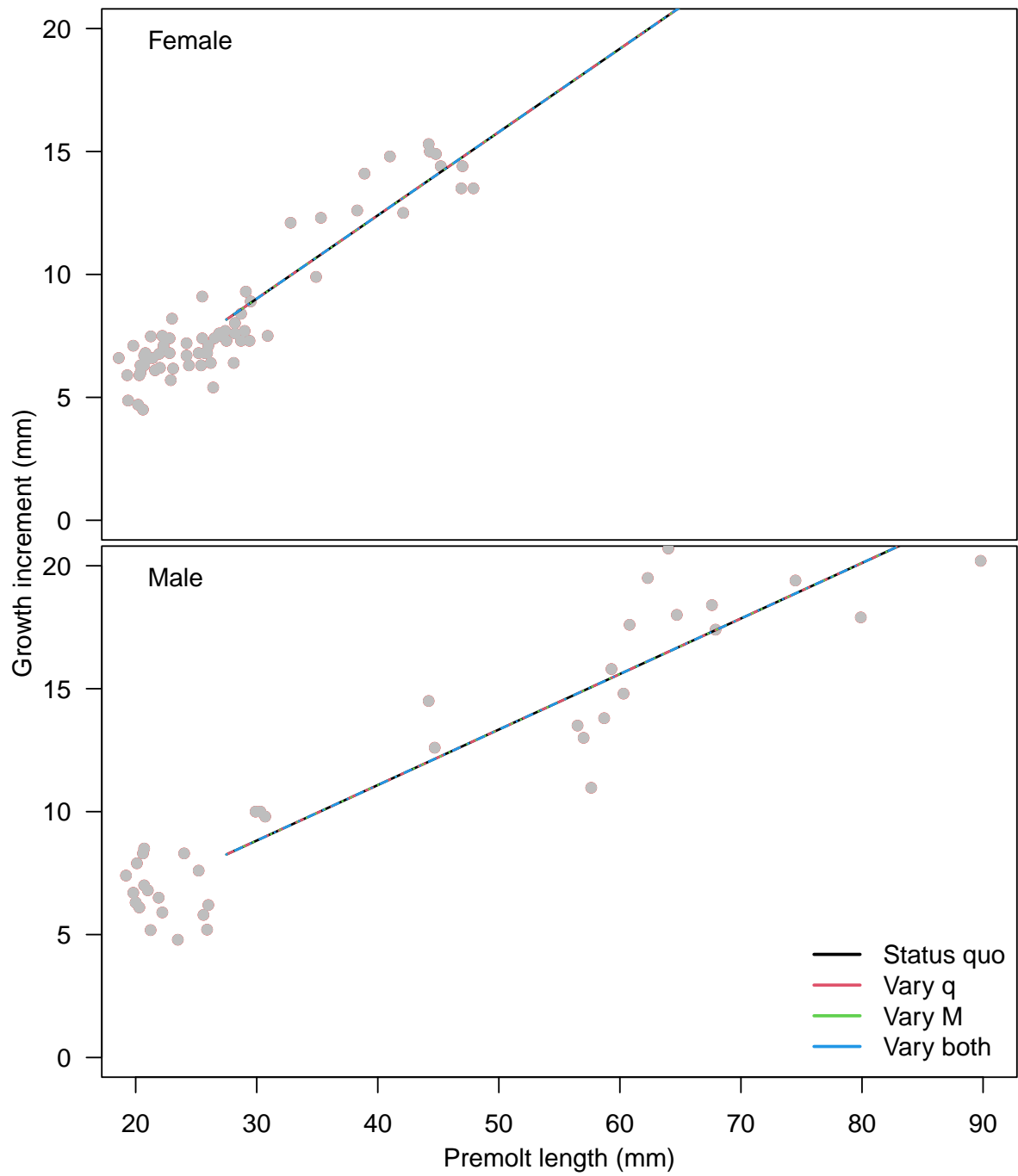


Figure 9: Model fits to the growth data. These data were fit outside of the assessment model with a linear regression and the associated parameters were specified within the assessment.

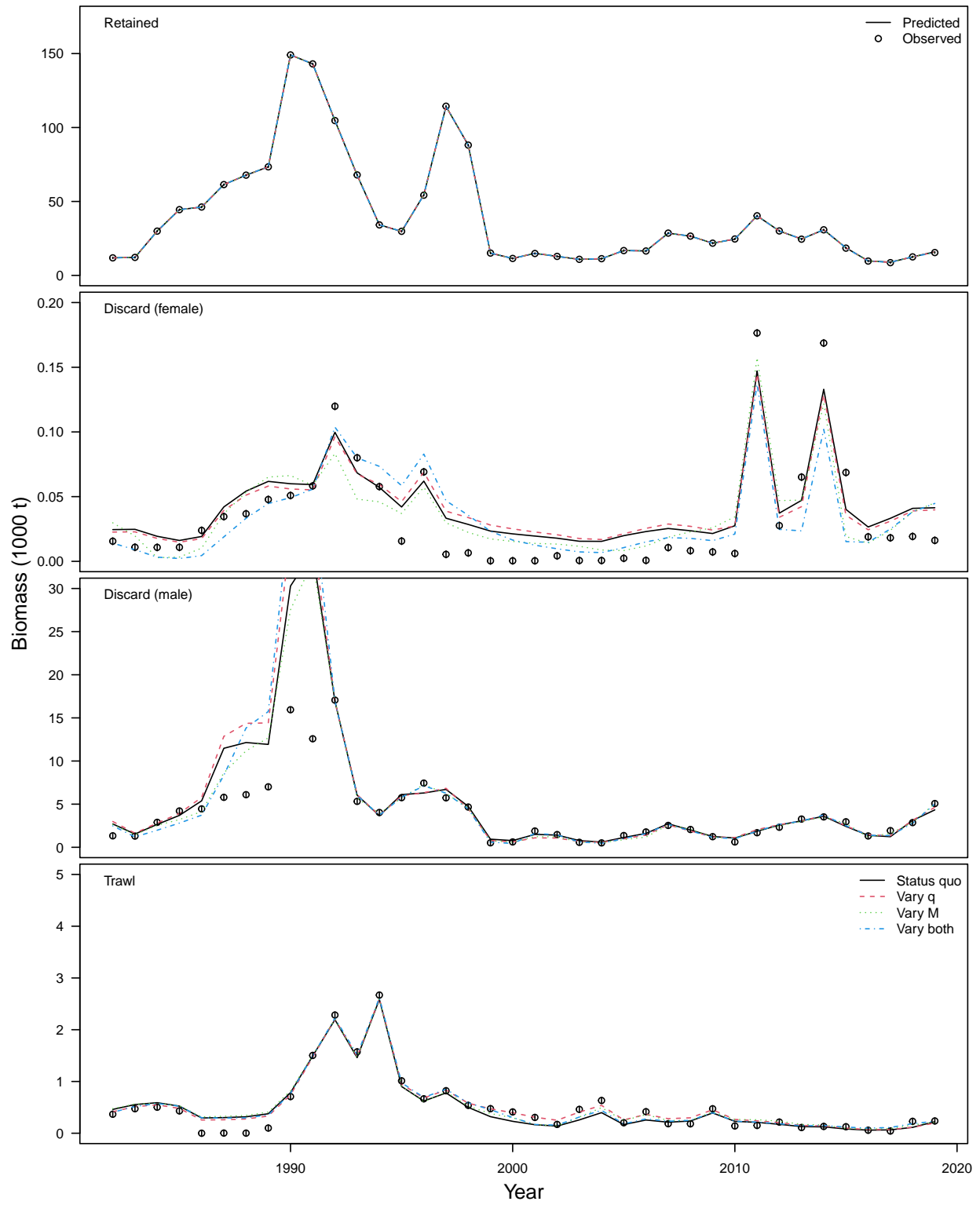


Figure 10: Model fits to catch data

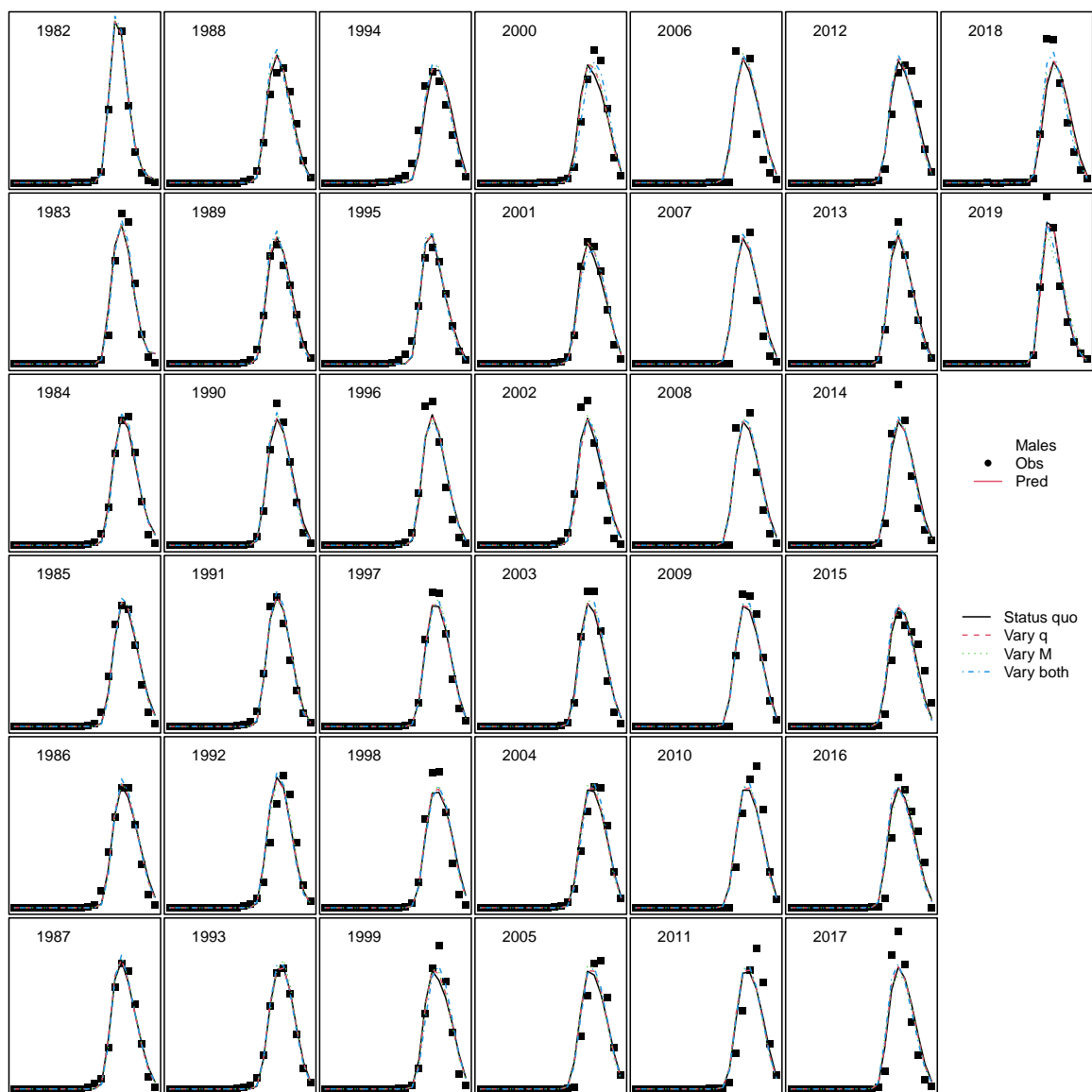


Figure 11: Model fits to retained catch size composition data

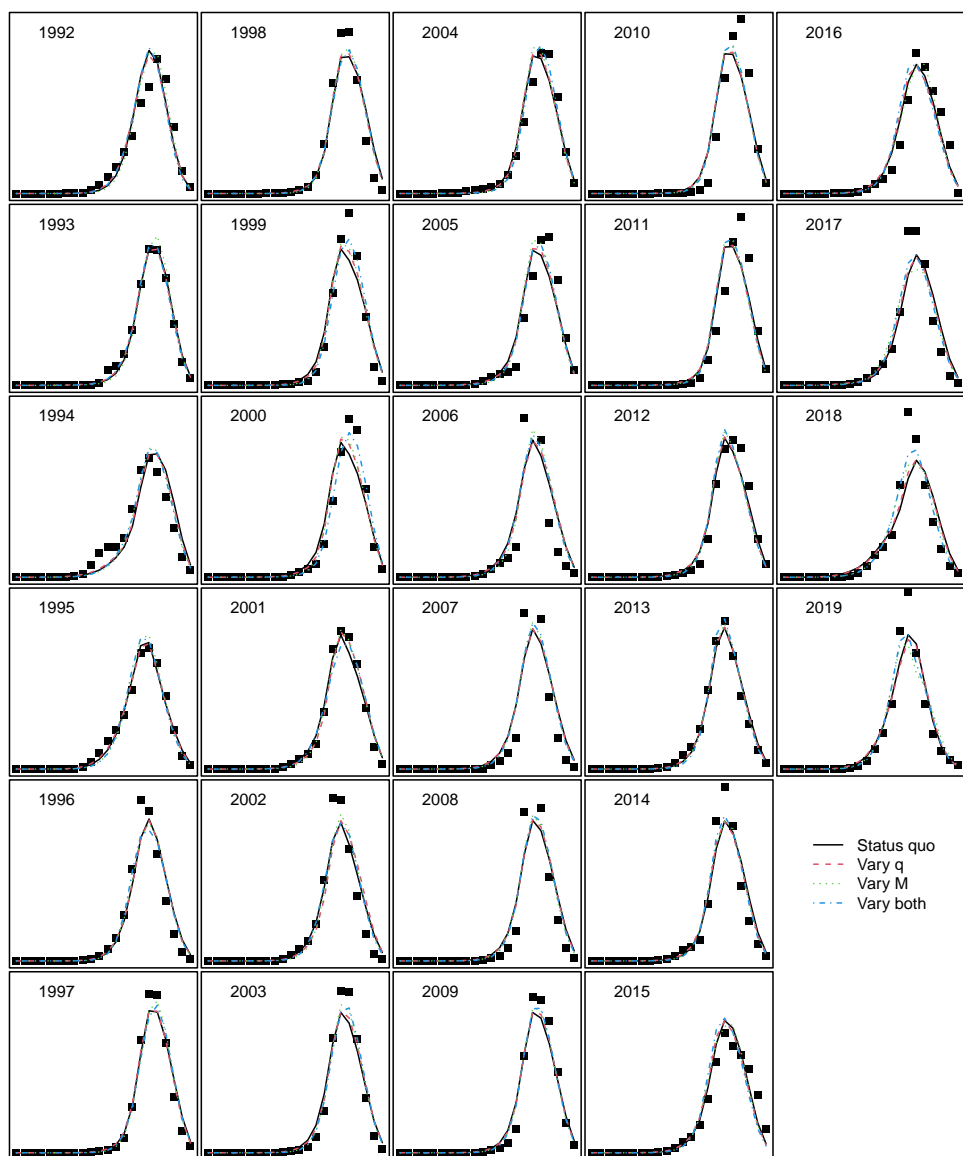


Figure 12: Model fits to total catch size composition data

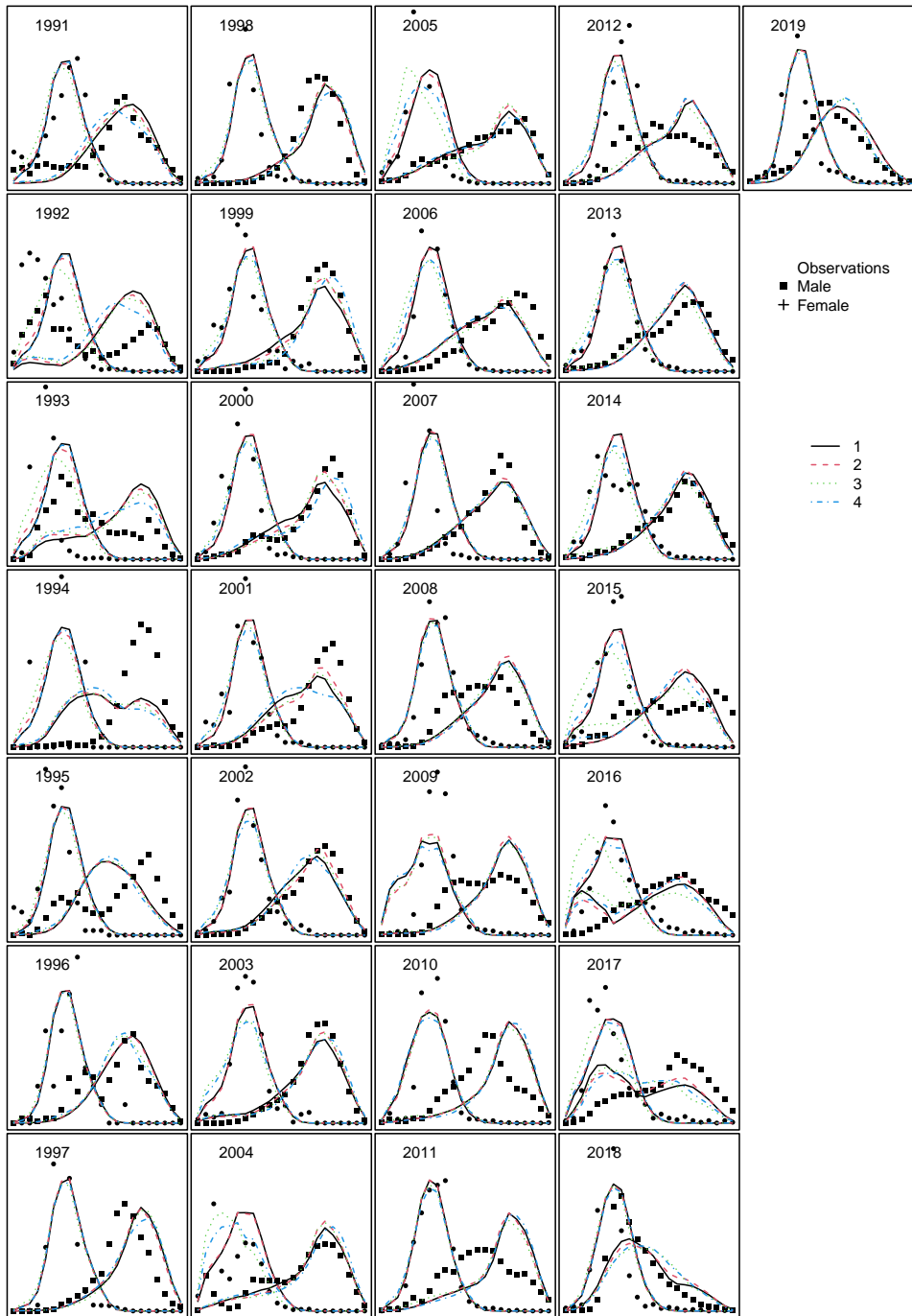


Figure 13: Model fits to trawl catch size composition data

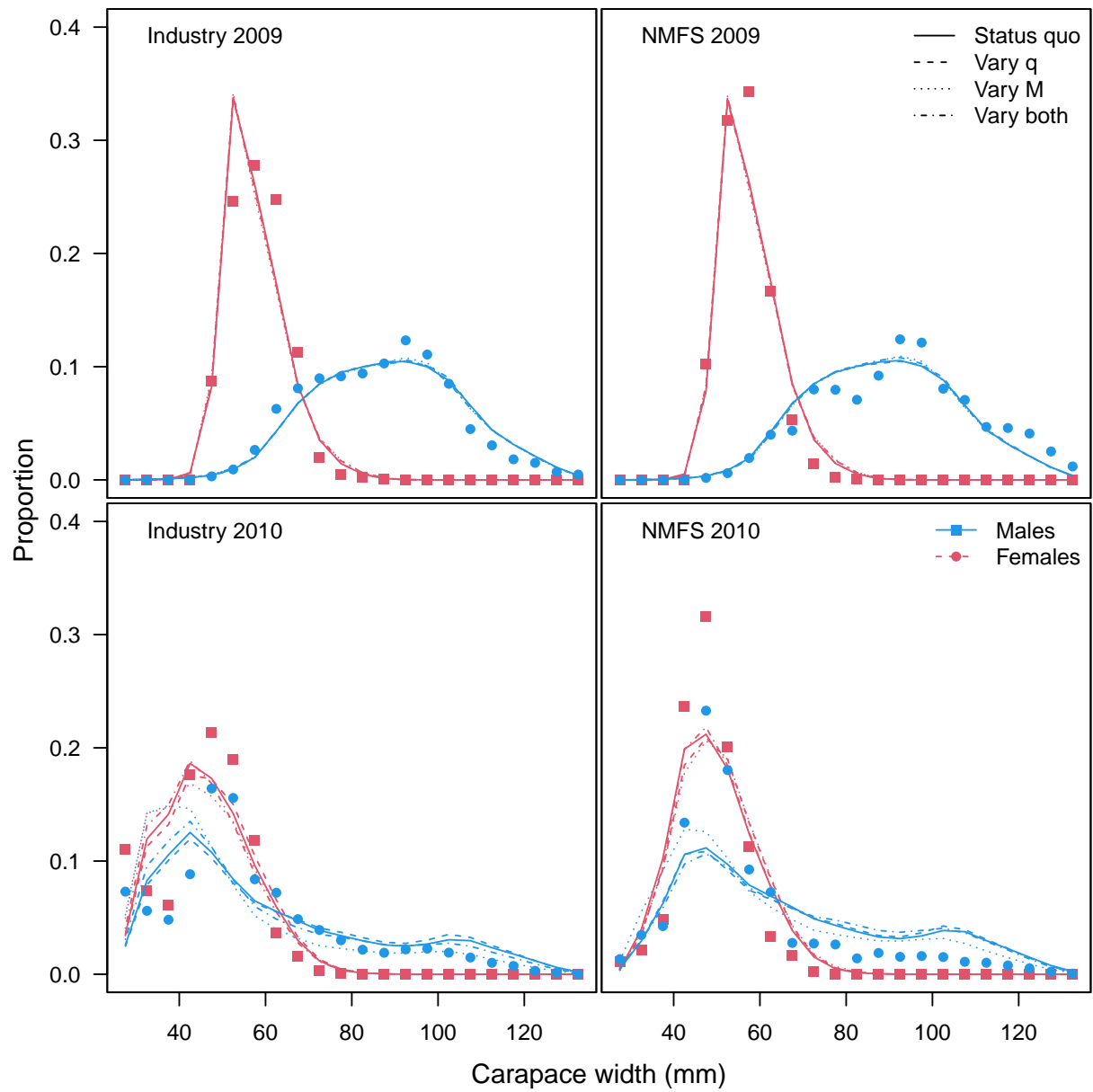


Figure 14: Model fits to size composition data from summer survey experiments (2009 & 2010)

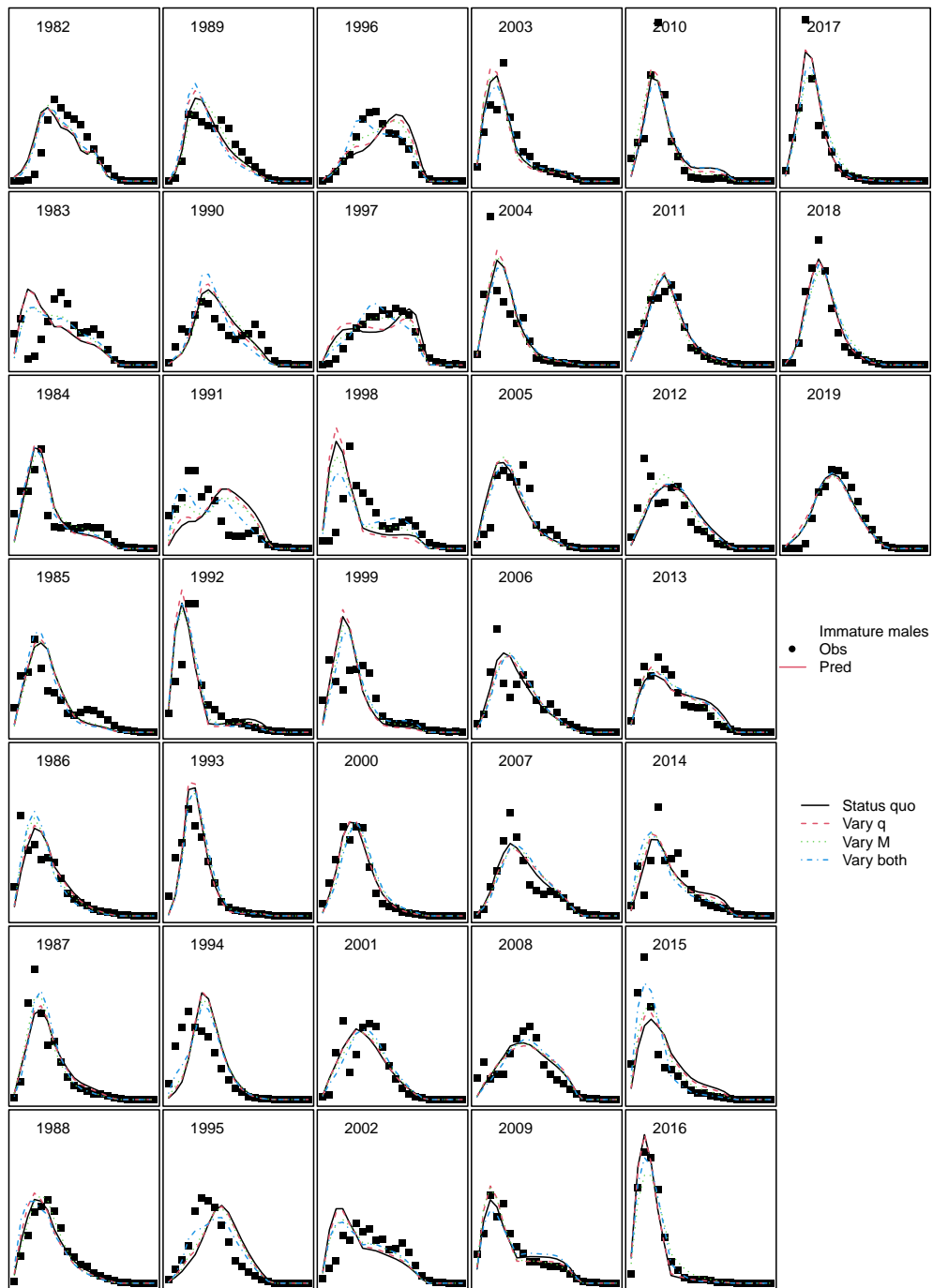


Figure 15: Model fits to immature male survey size composition data. Note that male and female survey selectivity proportions at length in a given year sum to 1. Consequently, the integral of predicted length compositions may appear to be different than the integral of the observed length composition data.

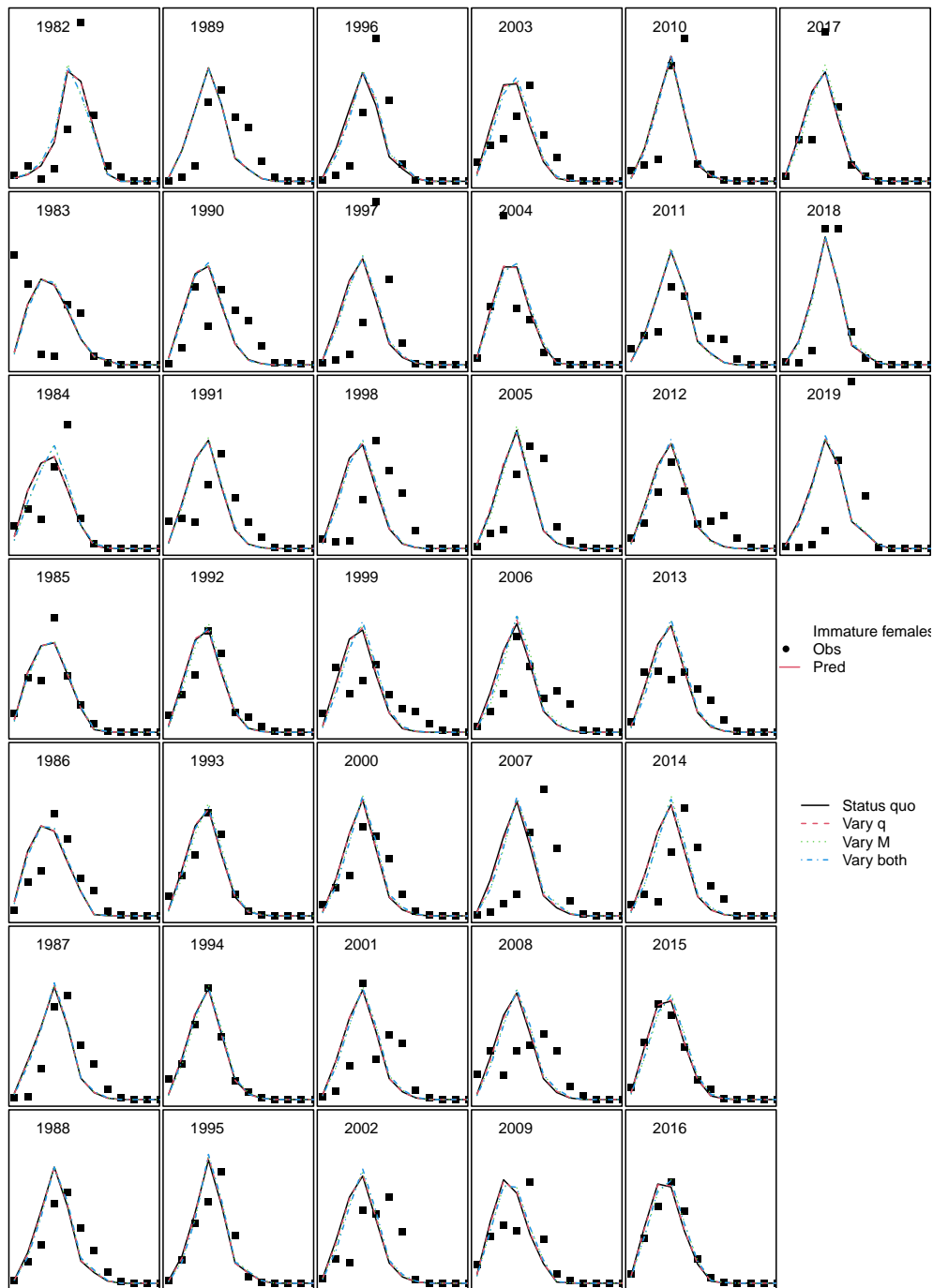


Figure 16: Model fits to immature female survey size composition data. Note that male and female survey selectivity proportions at length in a given year sum to 1. Consequently, the integral of predicted length compositions may appear to be different than the integral of the observed length composition data.

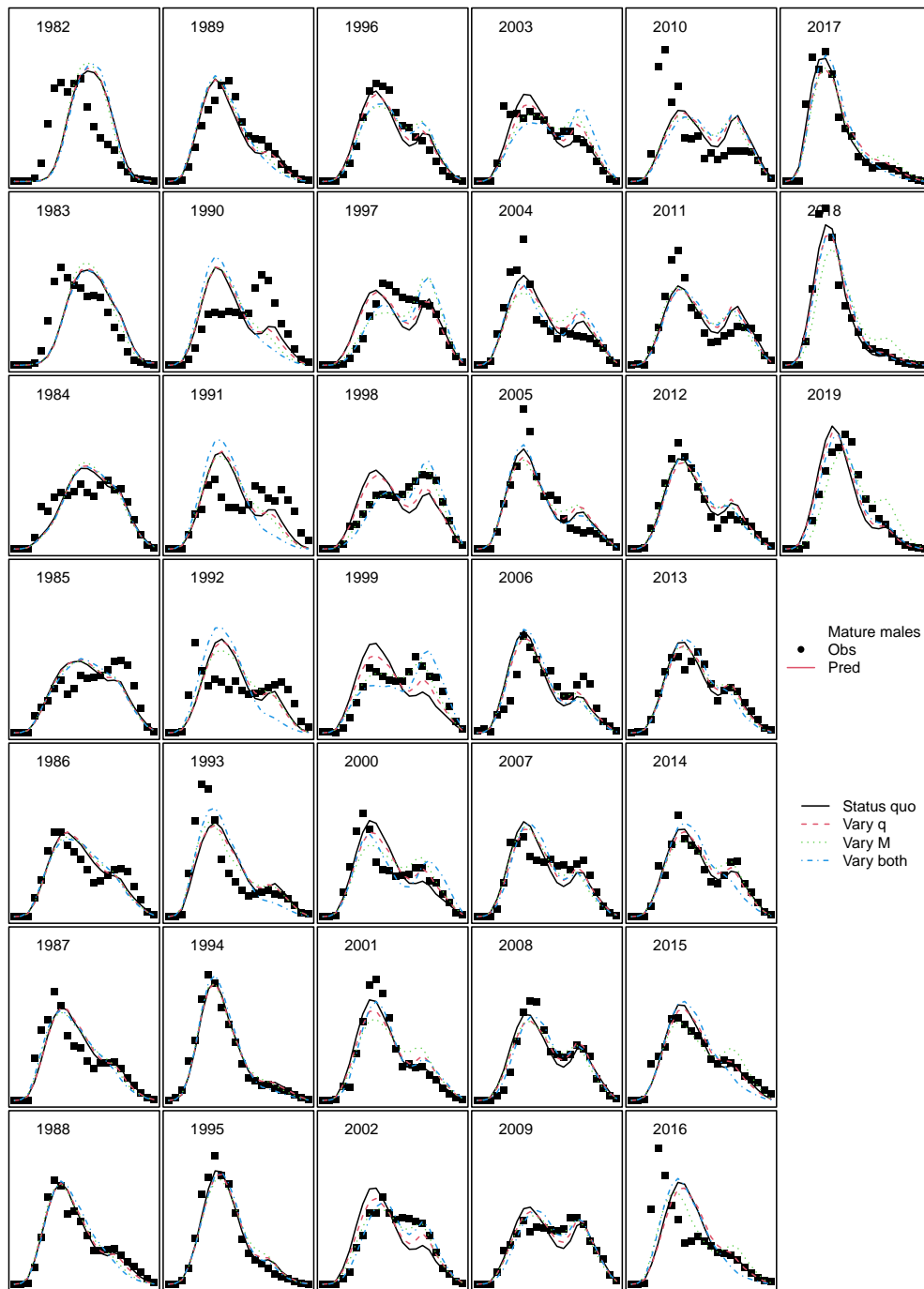


Figure 17: Model fits to mature male survey size composition data. Note that male and female survey selectivity proportions at length in a given year sum to 1. Consequently, the integral of predicted length compositions may appear to be different than the integral of the observed length composition data.

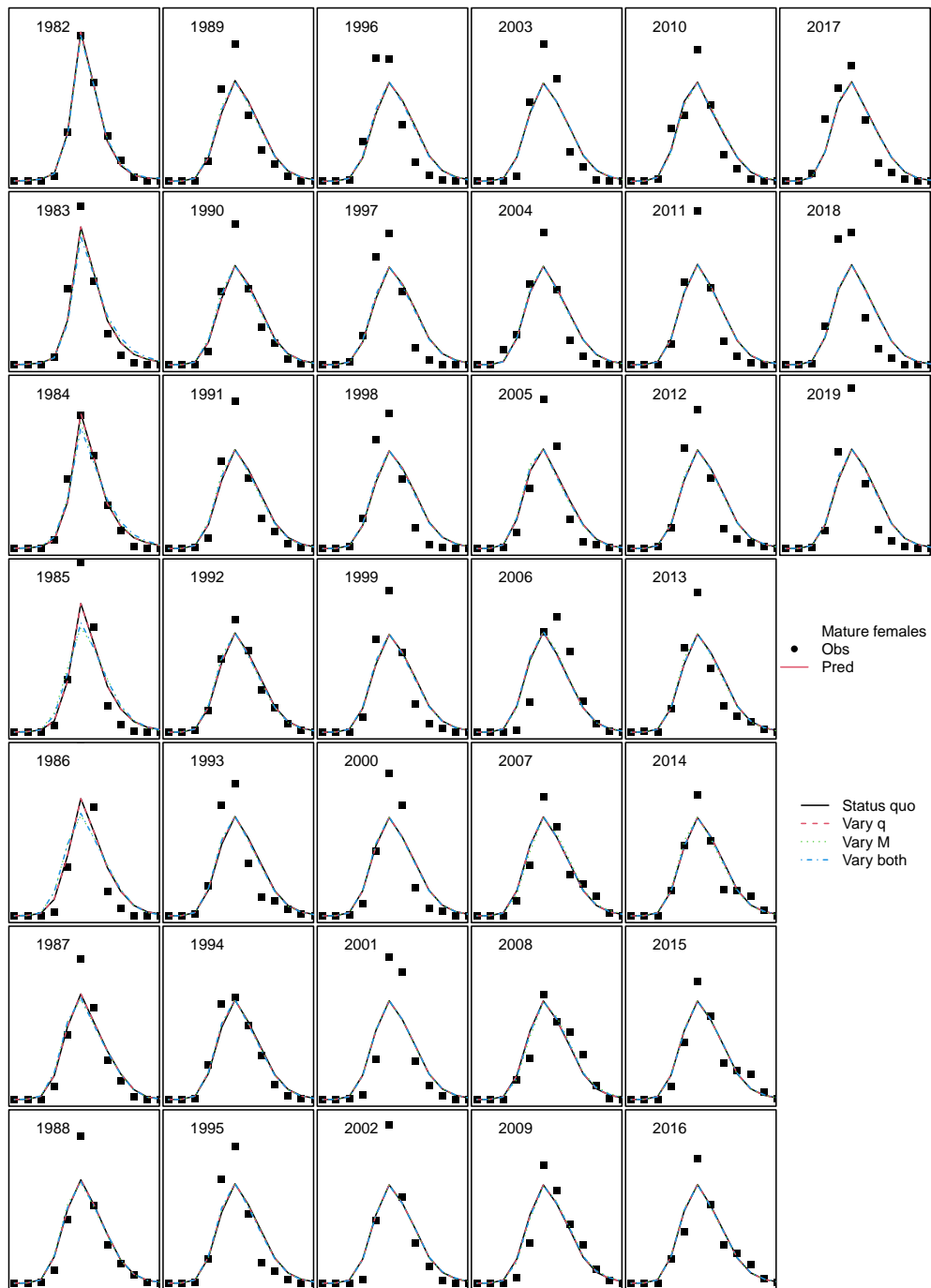


Figure 18: Model fits to mature female survey size composition data. Note that male and female survey selectivity proportions at length in a given year sum to 1. Consequently, the integral of predicted length compositions may appear to be different than the integral of the observed length composition data.

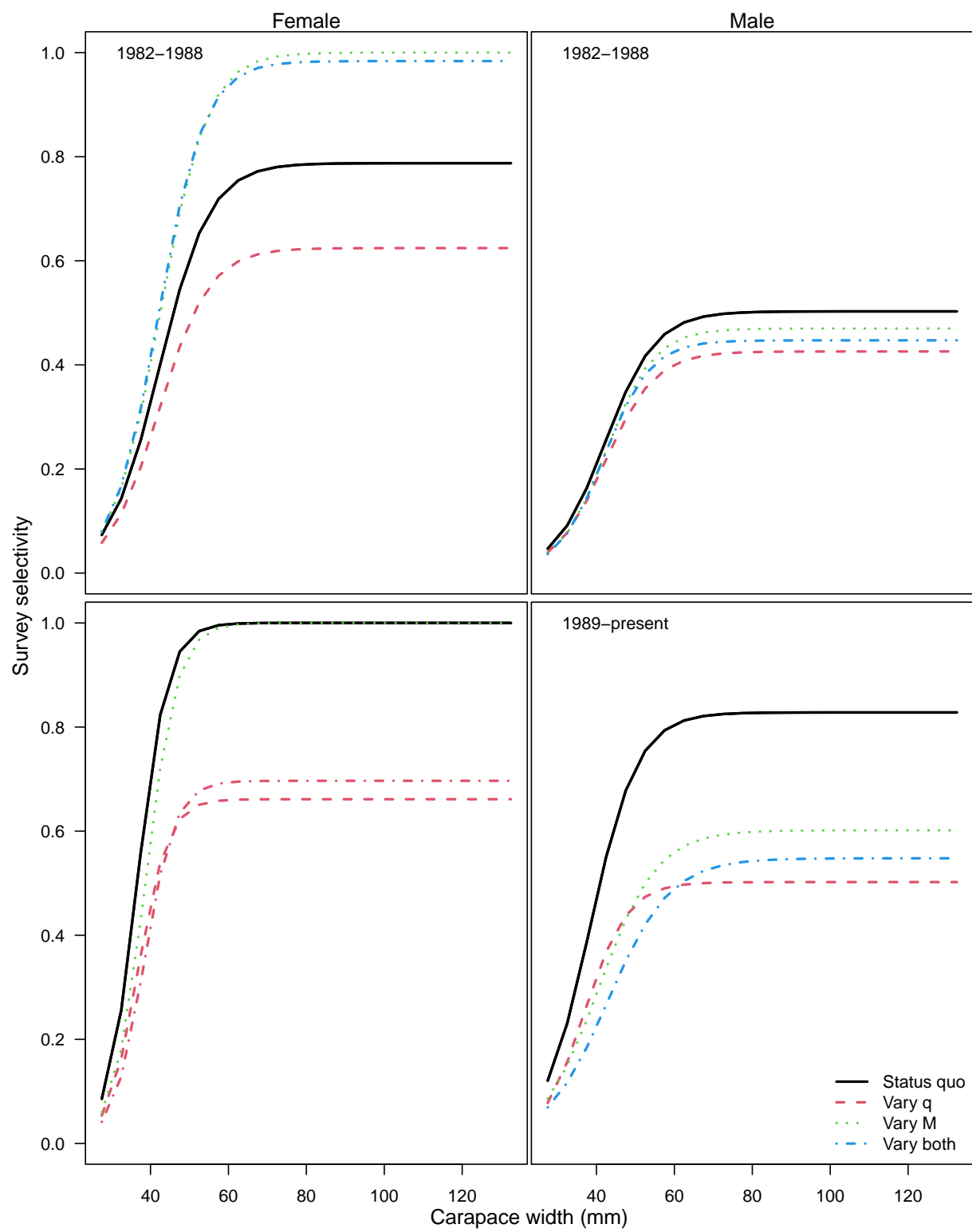


Figure 19: Estimated survey selectivity

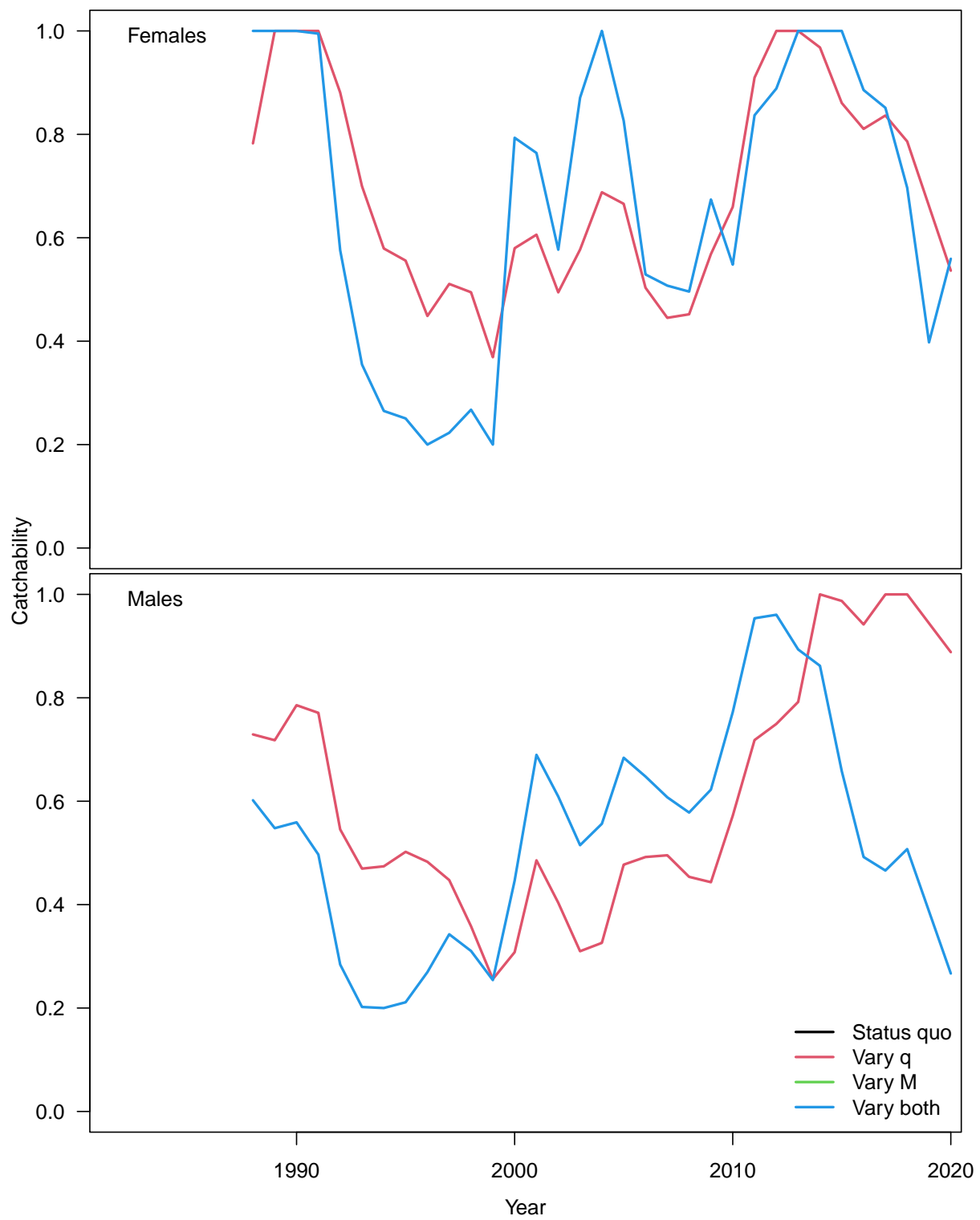


Figure 20: Estimated time-varying survey catchability (era 3).

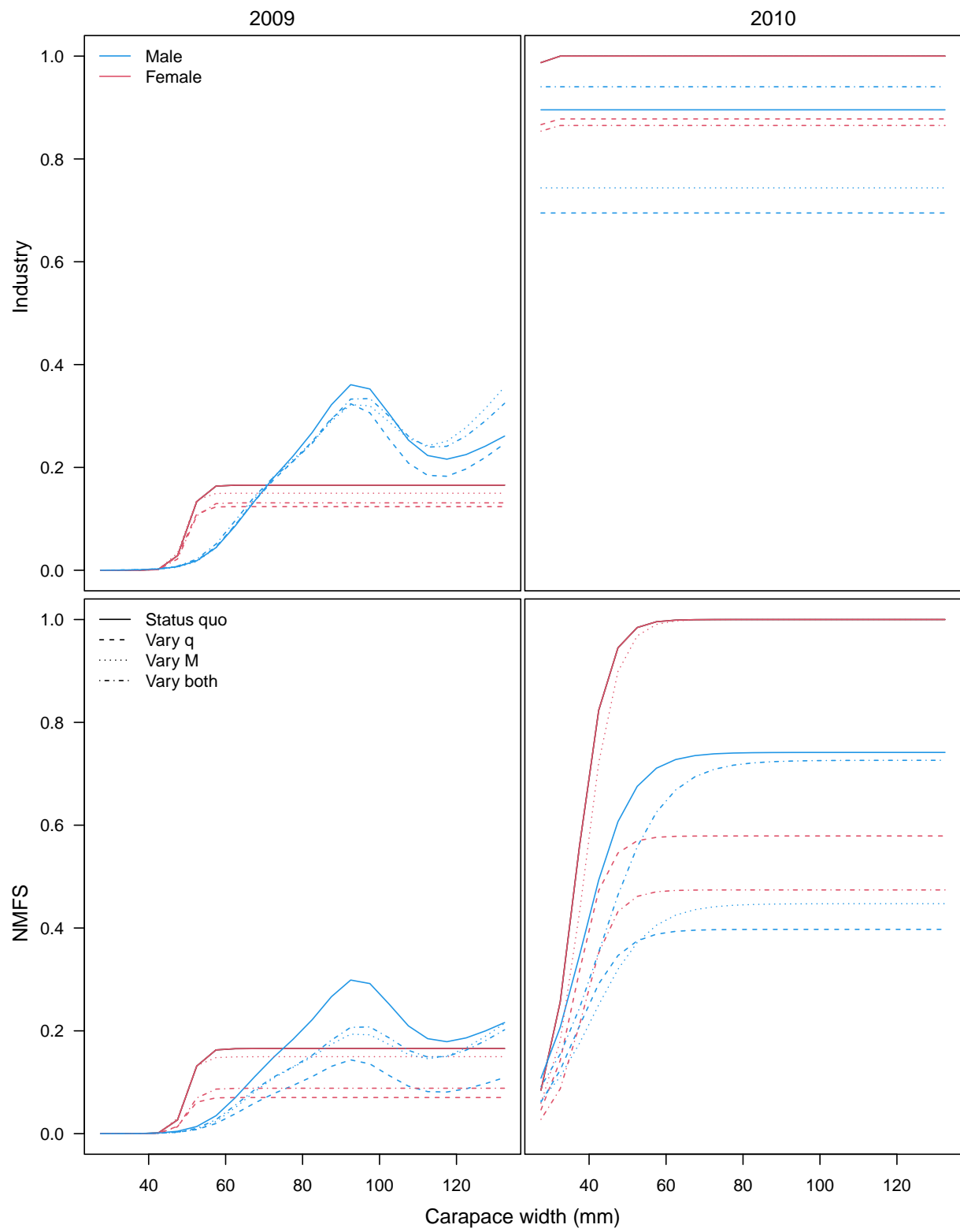


Figure 21: Estimated experimental survey selectivity (availability * survey selectivity)

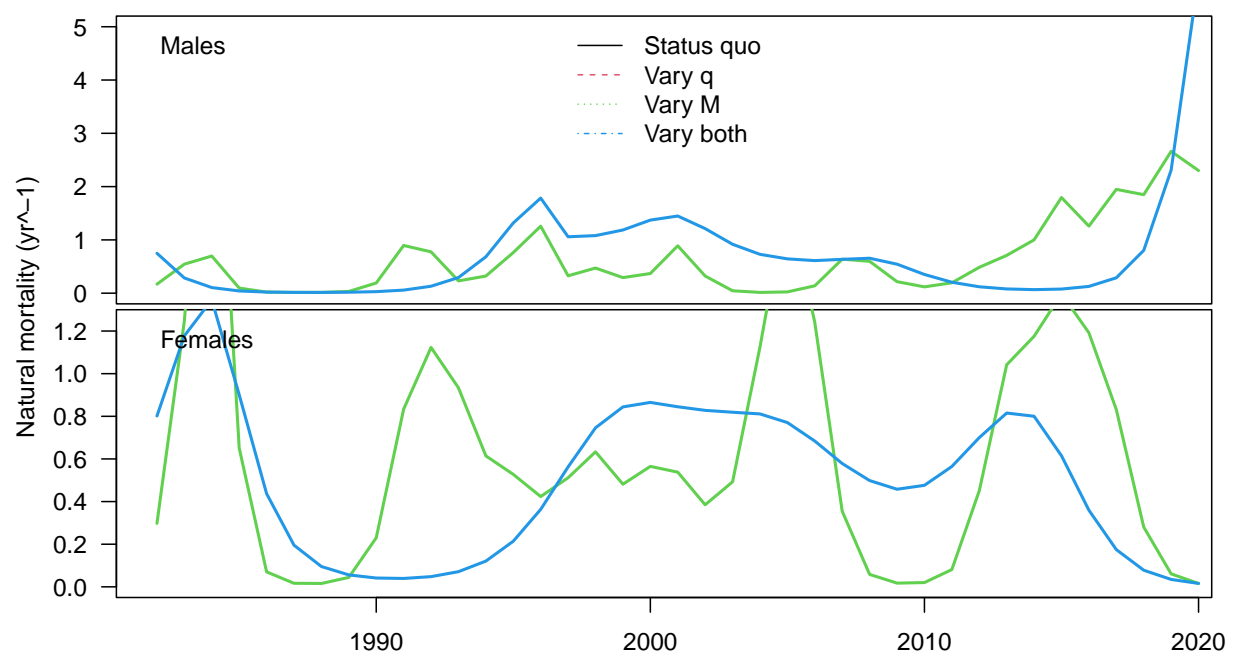


Figure 22: Estimated time varying natural mortality by sex and maturity state. Note different scales between sexes.