Fisheries and Oceans Canada | Pêches et Océans Canada
Ecosystems and Oceans Science | Sciences des écosystèmes et des océans

Canadian Science Advisory Secretariat (CSAS)

Research Document 2023/017

Newfoundland and Labrador Region

A State-Space Assessment Model for 3Ps Cod (3PsSSAM)

N. Cadigan

Centre for Fisheries Ecosystem Research
Fisheries and Marine Institute - Memorial University of Newfoundland
PO Box 4920
St. John's, NL A1C 5R3

March 2023

Foreword

This series documents the scientific basis for the evaluation of aquatic resources and ecosystems in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/
csas-sccs@dfo-mpo.gc.ca

ISSN 1919-5044
ISBN 978-0-660-47468-7 Cat. No. Fs70-5/2022-017E-PDF

Correct citation for this publication:

Cadigan, N. 2023. A State-Space Assessment Model for 3Ps Cod (3PsSSAM). DFO Can. Sci. Advis. Sec. Res. Doc. 2023/017. iv + 68 p.

Aussi disponible en français :

Cadigan, N. 2023. Un modèle espace-état pour l'évaluation des stocks de morue de la sous-division 3Ps. Secr. can. des avis sci. du MPO. Doc. de rech. 2023/017. iv + 71 p.

TABLE OF CONTENTS

ABSTRACT

I describe an implementation of a state-space stock assessment model (SSM) for 3Ps cod. An assessment challenge for this stock is that there are uncertainties and possible biases in fishery landings information that are difficult to quantify. I use the censored likelihood approach based on the best available information on potential inaccuracies in landings to address this uncertainty. Hence, the SSM only uses information on lower and upper bounds of what the real fishery landings were. Information on the age-composition of the catches is included in the model fitting using a likelihood based on the continuation ratio logits of catch proportions-at-age.

Otherwise, the SSM I describe is formulated like the typical SAM used in ICES assessments, with some differences in the stochastic model for variation in fishing mortalities (F's) and the likelihood function for survey indices. I assume that survey indices have a normal distribution with a constant coefficient of variation. An advantage of this approach is that indices with zero values can be used for estimation, whereas they cannot be used with the more commonly used lognormal distribution. This is an important issue when there are many zeros that are not distributed at random throughout the survey ages and years, which is the case for 3Ps cod.

Many model formulations were presented and compared in terms of AIC/BIC and other model results. Retrospective patterns were presented for some model formulations. The retrospective patterns were large enough to be a concern. The model formulations had trouble finding a good fit to the DFO RV and Sentinel indices. I did not find an acceptable assessment model formulation, and I concluded that further research on assessment inputs is needed before I could recommend a reliable assessment model formulation.

INTRODUCTION

State-space stock assessment models (SSMs) involve stochastic process equations that describe stock dynamics, and stochastic observation equations that relate data to stock quantities and account for observation error. These stochastic equations involve random effects with assumed statistical distributions that involve mean and covariance parameters that must be estimated or somehow specified. A good way to estimate model parameters is maximum likelihood based on the marginal likelihood of observations in which random effects have been integrated out. The Template Model Builder (TMB) package (Kristensen et al. 2016) is a state-of-the-art tool for this purpose, and I use this software to implement a state-space assessment model for 3Ps Cod (3PsSSAM).

Ideally in an integrated stock assessment framework the data would be used in “raw” form. However, in practice there are at least two reasons why this is not possible or pragmatic:

1. some of the raw data are not available and
2. the statistical distributions and associated likelihood equations of the raw data may be highly complex and involve spatial aspects of stock abundance etc.

Hence, aggregation of data as summary statistics (e.g., survey mean number per tow) is required. However, the statistical distribution of aggregated data may still be complicated and pragmatic assumptions and simplifications are required to make progress. The validity of these assumptions must be examined during model fitting which creates an additional layer of model building (i.e., the observation likelihoods) in addition to the stochastic process equations to describe stock dynamics.

A challenge for 3Ps cod is that there are uncertainties and possible biases in fishery landings information that are difficult to quantify. I use the censored likelihood approach based on the best available information on potential inaccuracies in landings to address this uncertainty. Hence, 3PsSSAM only uses information on lower and upper bounds (derived from expert knowledge) of what the real fishery landings were. This is the primary reason why 3PsSSAM is based on landings bounds and the age-composition of the catches rather than a catch-at-age matrix that could be derived from landings estimates and age compositions. Another reason to use landings and age-compositions is the integrated modelling philosophy where different data sources have different observation likelihoods. However, other SSMs such as SAM (Nielsen and Berg 2014), which is commonly used for ICES assessments, use the catch-at-age directly as observations with only measured errors that may be correlated. Albertsen et al. (2016) concluded that using the catch-at-age data directly was a better approach than modelling total catch and age compositions separately. However, in the 3Ps cod context in which we think there are directional (i.e., under-reporting) errors in landings, the censored approach directly addresses the issue. As far as I am aware, expert knowledge about the accuracy of landings cannot be accounted for using SAM.

Otherwise, the 3PsSSAM I propose is formulated like the typical SAM used in ICES assessments. The 3PsSSAM formulation is described in the next section, with special emphasis on the rationale for a few other differences with SAM. Many fixed-effect mean and variance parameters and random effects in 3PsSSAM are initially specified to have high dimension, i.e., a parameter or effect for each year and/or age, and then the TMB map argument is used to constrain most of the parameters to be the same, where appropriate. This is part of the stock assessment model building process, and can occur on the “R-side” rather than the “TMB-side” which makes implementation easier. I refer to this dimension reduction as mapping, so that a high-dimensional parameter vector is “mapped” to a lower dimension.

METHODS

PROCESS EQUATIONS

The 3PsSSAM stochastic cohort model with a plus group is

$$\log(N_{a,y}) = \begin{cases} \log(N_{a-1,y-1}) - Z_{a-1,y-1} + \delta_{a,y}, & a < A, \\ \log\{N_{a-1,y-1}\exp(-Z_{a-1,y-1}) + N_{a,y-1}\exp(-Z_{a,y-1})\} + \delta_{a,y}, & a = A, \end{cases} \quad y = 1, \ldots, Y. \quad (1)$$

where $N_{a,y}$ is stock abundance at age *a* in year *y*, $Z_{a,y}$ is the total mortality rate, and $\delta_{a,y}$ is process error (PE). The ages are 1–14+ and years are 1959–2019. The total mortality rate is separated into a fishing mortality rate (*F*) and a natural mortality rate (*M*), $Z_{a,y} = F_{a,y} + M_{a,y}$, where

$$M_{a,y} = \begin{cases} 0.5, & a = 1 \\ 0.3 & a = 2 \\ 0.2 & a = 3+ \end{cases}, y = 1959, \ldots, 2019$$

is an initial assumption, and $F_{a,y}$ are model random effects. The value of *M=0.2* for ages 3+ is commonly used in cod assessments, and I chose higher values of M at younger ages to account for the likely higher M experienced by younger and smaller cod. Other choices for M may also be reasonable. Equation (1) is the typical cohort model used in fish stock assessment. The recruitment vector, $R = (N_{1,1}, \ldots N_{1,Y})$, is assumed to be a lognormal random vector variable,

$$\log(R) \sim MVN(\mu_R, \Sigma_R),$$

where the *Yx1* parameter vector μ_R is mapped to either a constant value for all years, or two values split at 1990, with justification given in the **Results** section. Σ_R is the stationary covariance matrix of an AR(1) process defined by σ_R and φ_R. The correlation between $\log(R_i)$ and $\log(R_j)$ is $\varphi_R^{|i-j|}$. The numbers at age's 2–14+ in the first year are treated as unknown and free parameters to estimate.

The δ process errors are assumed to have a normal distribution with zero mean but possible autocorrelation over ages and years because process errors should be more similar for fish that are close together in age and time. These errors are assumed to have a stationary distribution derived from a lag 1 autoregressive process in both age and year so that

$$Cov\{\delta_{a,y}, \delta_{a-j,y-k}\} = \frac{\sigma_\delta^2 \varphi_{\delta,a}^j \varphi_{\delta,y}^k}{(1-\varphi_{\delta,a}^2)(1-\varphi_{\delta,y}^2)}, \; Corr\{\delta_{a,y}, \delta_{a-j,y-k}\} = \varphi_{\delta,a}^j \varphi_{\delta,y}^k. \quad (2)$$

However, the age and year autocorrelations can be difficult to estimate reliably. Autocorrelated process errors may be confounded with recruitment and F patterns. Hence, similar to SAM, I usually assume that these correlations are zero (i.e., i.i.d process error) and only consider estimating these correlations in exploratory models to reduce retrospective patterns.

Catches are modelled using the Baranov catch equation,

$$C_{a,y} = N_{a,y}\{1 - \exp(-Z_{a,y})\}F_{a,y}/Z_{a,y}. \quad (3)$$

Fishing mortalities are modelled as a stochastic process about some mean values similar to recruitment. There are no commercial catches at ages 1 and very few at age 2 so I fixed F's to be zero for these ages. 3Ps cod are not targeted at age 3 and temporal patterns in *F* at this age are likely different than temporal patterns in *F* at older ages. Hence, I modelled *F* at age 3 separately from *F* at ages 4–14+. If ***F*** is an *(A-3)Yx1* vector of all $F_{a,y}$'s for ages 4–14+, then

$$\log(\boldsymbol{F}) \sim MVN(\mu_F, \Sigma_F),$$

where μ_F is mapped to 6 values, for ages 4 and 5+ crossed with time periods 1959–1993, 1994–1996, and 1997–2019. This allows for a stochastic *F* process that accounts for shifts in mean *F* because of the three-year moratorium on fishing that started in 1994. Similar to the process errors, $\log(\boldsymbol{F})$ is modelled as an AR(1) stochastic process in age and year, and the elements of Σ_F are based on

$$Cov\{\log(F_{a,y}), \log(F_{a-j,y-k})\} = \frac{\sigma_F^2 \varphi_{F,a}^j \varphi_{F,y}^k}{(1-\varphi_{F,a}^2)(1-\varphi_{F,y}^2)}. \quad (4)$$

F's at age 3 have a different $\varphi_{F,y}$ parameter.

The SAM log(F)-process is an age-correlated random walk that implicitly has a year autocorrelation fixed at one. The SAM log(F)-process does not include changes in mean log(F), μ_F. This differs from 3PsSSAM where μ_F is modelled separately for coarse blocks of ages and years, and the year correlation is estimated for age 3 and ages 4–14+.

OBSERVATION EQUATIONS

In my 3Ps cod model this involves specifying the covariance matrix of observations, and I use model residual diagnostics for this purpose.

The index observation model is like the traditional approaches used for NW Atlantic fish stocks. Let $I_{s,a,y}$ denote the observed age-based abundance index for survey s. Let t be the midpoint of the survey dates which is expressed in a fraction of the year. The model predicted index is

$$\mu_{s,y,a} = E\left(I_{s,a,y}\right) = q_{s,a} N_{y,a} \exp^{-t_{s,y} Z_{y,a}}. \quad (5)$$

The $\exp^{-t_{s,y} Z_{y,a}}$ term projects beginning-of-year abundance to the time of the survey. I use Equation (5) for all survey ages and years, including the plus group age for plus group survey indices. The $q_{s,a}$'s are catchability parameters to estimate, possibly with constraints. The index observation equation is

$$I_{s,a,y} = \mu_{s,y,a} + \varepsilon_{s,a,y}. \quad (6)$$

Additive index observation errors are unusual in stock assessment models, and the rationale for this is described below. Let $\varepsilon_{s,y}$ be the vector of index errors for all ages in a year. The index likelihood is based on

$$\varepsilon_{s,y} \sim MVN(0, \Sigma_{s,y}),$$

where

$$\Sigma_{s,y} = Diag(\sigma_{s,a_{min},y}, \ldots, \sigma_{s,a_{max},y}) \Pi_s Diag(\sigma_{s,a_{min},y}, \ldots, \sigma_{s,a_{max},y}), \quad (7)$$

a_{min} and a_{max} are the minimum and maximum ages of the index used for model estimation, Π_s is an AR(1) correlation matrix with parameter φ_s , and $\sigma_{s,a,y} = \gamma_{s,a,y}\, \mu_{s,y,a}$. This is a constant coefficient of variation (CV; γ) model although in 3PsSSAM these CV's can be modelled separately by survey, age, and year but will be mapped to a small number of groups, usually the same for all ages and years, but separate for each survey. Residual diagnostics are examined to determine ages and years that may require separate γ's.

In preliminary models it was obvious that index residuals were correlated across ages within years. This is common in stock assessments. Correlations could be caused by variations in the fraction of the stock covered by the survey (i.e., year effects in catchability), temporal changes in growth rates, etc. We model this correlation in the observation equation. However, if the patterns of residuals are similar for multiple indices then this may indicate mis-specification of the stock assessment model, and correlated index errors may only mask this problem. This is a particularly difficult problem when there is only a single index available for model estimation. For model diagnostic purposes we also compute a vector of residual diagnostics (e_{sz}) for each survey that standardize for this expected correlation in the index raw residuals (e_s), using their covariance matrix ($\hat{\Sigma}_s = Diag(\hat{\Sigma}_{sy_{min}}, \dots, \hat{\Sigma}_{sy_{max}})$. We solve for e_{sz} in the linear system of equations $\Sigma_s^{1/2\prime} e_{sz} = e_s$, where $\Sigma_s^{1/2}$ is the Cholesky factorization of $\hat{\Sigma}_s$. The standardized e_{sz} residuals should be approximately independent with mean zero and variance 1.

SAM, like most other stock assessment packages, models log indices based on the assumption that the variance of the log index does not depend on the mean. If $Var\{log(I)\} = \sigma^2$ then $CV(I) \approx \sigma$ and $STD(I) \approx \sigma\mu$. Hence, the 3PsSSAM index variance model is approximately the same as SAM. A difference in the approaches is that 3PsSSAM can use indices with zero values, whereas SAM cannot. SAM assumes any zero index is a missing value. This a reasonable approach when there are not many zeros, but many Northwest Atlantic stocks have declined to low stock sizes and zero indices are common and indicate low abundance of the corresponding age classes. Treating these indices as missing could result in biased estimates of the size of these age classes. Strictly speaking the $\varepsilon_{s,a,y}$ terms in Equation (6) should have a bounded distribution, $\varepsilon_{s,a,y} \in (-\mu_{s,y,a}, \infty)$, so that the predicted indices are positive. Hence, strictly speaking a truncated normal distribution is more appropriate. However, Perreault et al. (2020) discuss that the difference in a constant CV normal distribution and a truncated normal distribution will usually be negligible except when the CV > 0.5, and even then, it only effects the weighting indices get when they have different CV terms (i.e., σ) in the model formulation.

3PsSSAM uses reported landings and estimates of the catch age-compositions to internally infer the catch at age. The total fishery catch is derived from landings and age compositions; hence, total catch is not independent of the age compositions which is why I prefer to model landings and age compositions, which are more independent data sources. 3PsSSAM predicted landings may differ from reported landings, depending on how accurate the reported landings are.

The reliability of the landings is quantified by lower and upper bounds that must be provided by experts with knowledge about the current and historical 3Ps cod fishery. Ideally there should be consensus agreement from scientists, managers, and harvesters that the real landings are contained within the bounds with high probability. 3PsSSAM uses a censored likelihood (e.g., see Cadigan 2016; Van Beveren et al. 2017) to ensure that model predicted landings are highly unlikely to exceed the bounds. However, within the bounds, the likelihood surface will be almost flat so that 3PsSSAM will have approximately no preference about values of landings within the bounds. If L_y denotes the true but unknown landings in year *y*, and L_{Ly} and L_{Uy} are the lower and upper bounds (i.e., the data), then the censored negative loglikelihood (nll) for the stock assessment model parameters (collected in a vector *θ*) is

$$nll(\theta|L_{Ly}, L_{Uy}) = -\sum_{y=1}^{Y} \log\left[\phi_N\left\{\frac{\log(L_{Uy}) - \log(L_y)}{\sigma_l}\right\} - \phi_N\left\{\frac{\log(L_{Ly}) - \log(L_y)}{\sigma_l}\right\}\right]. \quad (8)$$

The σ_l parameter controls the sharpness of the bounds and is set at $\sigma_l = 0.02$ in 3PsSSAM. The censored nll is illustrated in Fig. 1.

The time-series of catch abundance proportion at age, which I more simply refer to as catch age compositions, are modelled using the multiplicative logistic multivariate normal distribution, based on the continuation ratio logits (crl's), which are computed as follows.

1. For each age and year, compute $P_{a,y} = \frac{C_{a,y}}{\sum_{a=1}^{A} C_{a,y}}$.

2. Compute $\pi_{a,y} = Prob(age = a | age \geq a) = \frac{P_{a,y}}{P_{a,y} + \ldots + P_{A,y}}$.

3. Compute the continuation-ratio logit, $X_{a,y} = \log\left(\frac{\pi_{a,y}}{1-\pi_{a,y}}\right), a = 1, \ldots, A-1$.

This is done for both the observed and the model predicted catches. Note that there are only *A-1* crl's derived from *A* catch proportions because catch proportions only contribute *A-1* independent observations since $\sum_{a=1}^{A} P_{a,y} = 1$. If an observed catch at age is zero then the crl is not defined. For 3Ps cod, estimated catches at ages 10–14+ in 1995 were zero, and these values were replaced with half the minimum non-zero estimated catch at these ages.

The observation equation for the vector X_{oy} of observed crl's in year *y* is

$$X_{oy} = X_y + \varepsilon_{X,y}, \varepsilon_{X,y} \sim MVN(0, \Sigma_X),$$

where Σ_X is AR(1) in form, with variance parameter σ_X^2 and correlation φ_X.

The DFO RV survey was extended to include inshore strata in 1997, and indices for all strata are called the RV_IO indices, while indices from the strata surveyed prior to 1997 are called the RV_OFF indices. The age-distribution of 3Ps cod varies spatially, such that younger fish are found closer to shore. In most surveys since 1997 the mean number per tow (mnpt) for the RV_IO are slightly greater than RV_OFF at young ages, and slight lower than RV_OFF at older ages. One approach to deal with this change in survey design is to treat the RV_OFF and RV_IO as separate indices with different catchability *q* parameters, which is what is done in 3PsSSAM. However, this will involve some loss of stock trend information before/after 1997 since we expect that the differences in q's for RV_OFF and RV_IO are small. This is compounded by not having other indices that extend many years before/after 1997. For example, the Sentinel indices start only in 1995. In 3PsSSAM, I also include the age-based differences in the RV_IO minus RV_OFF since 1997 to constrain differences in the *q*'s for these two indices. The expected log-difference in the survey indices is

$$E\{log(I_{RV\_IO,a,y}) - log(I_{RV\_OFF,a,y})\} = log(q_{RV\_IO,a}) - log(q_{RV\_OFF,a}).$$

The final 3PsSSAM observation equation for $D_{RV,a,y} = log(I_{RV\_IO,a,y}) - log(I_{RV\_OFF,a,y})$ is

$$D_{RV,a,y} = log(q_{RV\_IO,a}) - log(q_{RV\_OFF,a}) + \varepsilon_{D,a,y}, \varepsilon_{D,a,y} \overset{iid}{\sim} N(0, \sigma_D^2). \qquad (9)$$

This provides information about the potential magnitude and uncertainty about the log differences in q's for the RV_OFF and RV_IO indices. If either RV_OFF or RV_IO indices are zero then the log difference is not used. Some of the RV log differences have long tails so I also explore using a t-distribution with 3 degrees of freedom for $\varepsilon_{D,a,y}$.

BASIC MODEL INPUTS

Landings estimates for 1959–2018 (see Fig. D1) and estimates of catch at age during 1959–2017 and ages 1–14+ (Figs. D2–5) were available as model inputs. This defined the model's years (1959–2019) and ages (1–14+). Note that the final model year is 2019 because we can estimate beginning of year stock size in 2019 based on landings in 2018.

The landings estimates are assumed to have errors that varied over time. I used the censored catch approach described above. The bounds were derived as lower (L) and upper (U) multipliers of reported landings. Models were developed for two scenarios of uncertainty about landings:

Good landings multipliers (L/U):

Year range	1959–75	1976–77	1978–86	1987–93	1994–96	1997–2009	2010+
L/U	0.99/1.1	0.7/1.3	0.99/1.1	0.5/1.5	0.99/1.05	0.99/1.3	0.99/1.1

Uncertain landings multipliers (L/U):

Year range	1959–93	1994–96	1997–99	2000+
L/U	0.5/1.5	0.9/1.1	0.35/1.75	0.75/1.25

I chose the uncertain multipliers to reflect my impression of the potential errors in the catches. However, the good multipliers were those recommended during the framework process.

Survey indices available for model estimation are:

- Canadian research vessel (RV) bottom-trawl index for inshore+offshore strata (Can_RV_IO), 1997–2018 (not 2006); ages 1–14+ (see Figs. D6–8).
- Canadian RV index for offshore strata (CAN_RV_OFF), 1983–1996; ages 1–14+ (see Figs. D9–11).
- French ERHAPS RV survey, 1978–1992, ages 2–14+ (see Figs. D12–14).
- GEAC industry survey, 1998–2005, ages 2–13. A 14+ index for this survey was not available (see Figs. D15–17).
- Sentinel gillnet index (SENT_GN), 1995–2017, ages 3–10 (see Figs. D18–20).
- Sentinel linetrawl index (SENT_LT), 1995–2017, ages 3–10 (see Figs. D21–23).

The difference between the Can_RV_IO and the CAN_RV_OFF during 1997–2018, and for ages 2–14+ (not 2006), is also used as a likelihood component to provide information about the ratio of catchabilities of these two survey indices. Note: age 1's are not used for this purpose because of uncertain Engel-Campelen conversion at that age. Hence, age 1 q's are not constrained by the difference likelihood component. The RV differences are illustrated in Figs. D24 and D25.

Other assessment model inputs were values for natural morality rates: age 1 M=0.5, age 2 M=0.3, and M=0.2 for other ages. M is assumed to be constant across years. Stock weights from a simple model with cohort effects (see Cadigan in press[1]). Catch weights from the simple model (see Cadigan in press[1]) were used in pre-FW (framework meeting) runs, but during the framework process it was decided to use raw catch weights in FW runs. Catch weight at age in 2018 was assumed to be equal to weight at age in 2017.

[1] Cadigan, N.G. In Press. A Simple Random-Effects Model to Smooth and Extrapolate Weights-at-Age for 3Ps Cod. DFO Can. Sci. Advis. Sec. Res.

FORMULATIONS

Many preliminary models were examined and presented at a pre-framework meeting on October 3, 2019. Descriptions of these models can be found in Table 1; these formulations were based on the "uncertain landings multipliers" shown above. It was recommended at the pre-framework meeting that I also pursue models using the landings multipliers based on expert input, which are the "good landing bounds" described above. The latter bounds imply much more precision in reported landings. I also used reported catch weights at age rather than modelled values for the framework meeting. This led to the following new model formulations:

1. **M17.** M10 formulation from pre-FW meeting, but with the good landings bounds (GL).
2. **M18.** M17, but no year correlation in crl's. This correlation was estimated to be only 0.19 in M10, and removing the temporal correlation simplifies the model and may improve run times.
3. **M19.** M18 but with age-invariant and time-varying M's based on indices of M derived from cod condition (see Ings et al. In Press[2]).
4. **M20.** M18 but with AR(1) process errors, to examine effects on retrospective patterns.
5. **M21.** M18 but a T distribution for differences in the RV survey indices.
6. **M22.** M21 but with AR(1) process errors (autocorrelations fixed at 0.5) and only ages 5–14+ for differences in the RV survey indices.

RESULTS

PRE-FRAMEWORK

These preliminary models are summarized in Table 1. A summary of AIC/BIC for these models is provided in Figure 2. The largest improvement in fit occurred for model 4 (i.e., M4) in which AR(1) age-correlated measurement errors were assumed for the CRL's and survey indices, although including IID process error (PE) in the model (i.e., M3) also resulted in a substantial improvement in model fit, as measured by AIC and BIC. M11, with a year/age separable AR(1) correlation structure and heterogeneous PE variances, also resulted in a substantial improvement in fit, although this type of model formulation is not common in ICES-SAM assessments and resulted in substantially different assessment results compared to formulations with independent (ID) PE's. Models with correlated PE's are implemented for diagnostics purposes only.

The M5–M7, M9, M10, and M16 models with IID PE's all resulted in similar assessments. A concern with these models is the cluster of positive residuals at young ages in the most recent years for the RV_IO index (e.g., see M7 results in Fig. 3). All these models had this residual pattern. There is also a cluster of negative residuals for the SENT_LT age 3 index in recent years, and two large negative CRL residuals at ages 3 and 4 in 2017 (Fig. 4). It seems that the RV_IO index recruitment trends in recent years are somewhat inconsistent with the SENT_LT index and the catch age compositions.

The M6 model resulted in some retrospective patterns (Fig. 5) that indicate assessment bias. Including correlated PE's improved the retrospective patterns (Figs. 6 and 7) but resulted in

[2] Ings, D.W., Varkey, D., Regular, P., Rideout, R.M., and Vigneau, J. In Press. Assessing the status of the cod (*Gadus morhua*) stock in NAFO Subdivision 3Ps in 2019. DFO Can. Sci. Advis. Sec. Res. Doc.

substantially different SSB and F scales (compared Figs. 5–7). M14 predicted process error are mostly negative (Fig. 8), especially at older ages since the late 1990’s.

FRAMEWORK

M17

This model incorporated more narrow bounds on landings than in the pre_FW runs. This has a large effect on stock assessment results (Fig. 9), more so for stock abundance and biomass than F’s. The tighter landing bounds in M17 forced estimated landings to be closer to reported landings, and more variation in F was required to achieve this. M10 has wider landings bounds with less constraints on estimated landings. In this case, the autocorrelation in F’s was estimated to be higher and this results in smoother estimates of F with less between-year and between-age variations. This is considered further in the Discussion. Model predicted landings for M17 are much larger overall than for M10 (Fig.10).

When there is much uncertainty about landings then the model acts like a survey-based assessment model (SURBA; Cadigan 2010; Cook 1997, 2013; Needle 2002), in which estimates of F, conditional on the assumed M, are absolute but estimates of stock size are relative. This is the basic behavior we see when comparing M10 and M17 in Fig. 9. F’s are more similar in scale for M10 and M17, but SSB and recruitment are more different. The CV on the 1959 estimate of SSB is 72% for M10, but only 18% for M17 (Fig. 11). The 1970 CV is 29% for M10 and 12% for M17. This reflects the greater precision in landings used in M17. However, retrospective patterns are pronounced (see M18).

M18

3PsSSAM runs slowly and retrospective runs take many hours. One reason for this may be a lack of sparseness in the random effects. The CRL correlation structure is in both age and year, although the M17 estimate of the year autocorrelation is low (0.18; see Table 2). In M18 I set the year correlation to zero to simplify the model and improve run times by increasing the sparseness of the random effects. AIC for M18 is slightly higher than M17, and BIC is lower (Table 2). Stock assessment results (Table 3) are nearly identical (Fig. 12) as were residual diagnostics (not shown). I consider M18 a useful simplification. However, M18 run times are only slightly faster than M17.

M18 has pronounced retrospective patterns (Fig. 13), such than terminal year confidence intervals do not cover subsequent estimates. The retrospective patterns for M17 are almost identical to M18.

Note: The slow run-time for 3PsSSAM was resolved after the framework meeting (see Discussion). The issue was the way process errors were implemented. However, this does not affect model results.

M19

Mis-specification of M is one possible cause of retrospective patterns. In M19 I used the age-aggregated condition M estimates I was provided to see if this improved model fit. I assumed M’s at ages 3–14+ were equal to annual condition M (cM) estimates, and M’s at age 1 were cM + 0.2, and M’s at age 2 were cM + 0.1. The differences for ages 1 and 2 were chosen to be consistent with my other model formulations in which M at these ages was higher than M at ages 3–14+.

M19 did not result in improved fit (Table 2) but did result in substantially different stock assessment results (Fig. 14a). However, retrospective patterns for SSB and average F ages 6–9 (Fig. 14b) were improved compared to M18 (Fig. 13), although the M19 pattern for recruitment was no better and opposite M18. However, the reduced retrospective patterns in M19 provides an indication that realistic changes in M that are consistent with changes in the fraction of cod in very poor condition (and the implied mortality due to starvation) can improve retrospective patterns. However, more investigation is required about the reliability of the condition M estimates and how they contribute to total M, so that this information could be used to provide an improved assessment model formulation.

M20

An alternative way to account for retrospective patterns is to use correlated PE's. If the retrospective patterns are caused by mis-specification of M or catch bounds, then the process errors, which are included to account for model mis-specification, may correct for this. However, IID process errors may not provide sufficient flexibility to account for model mis-specification that varies smoothly over time and ages (e.g., M18; Fig. 13). M20, with auto-correlated process errors, may provide improved retrospective patterns which could be a good thing if the type of processes errors the model predicts provide a realistic correction of the model mis-specification, or provides a reliable quantification of the uncertainty due to the model mis-specification. The latter may be the more realistic advantage of using autocorrelated process errors.

M20 results in a much better model fit compared to M17–19 (Table 2), and index CV's are generally slightly smaller which indicates a better mode fit to the indices. However, CV's for M20 assessment estimates are substantially higher than M18 (compare Tables 3 and 4). Overall, M20 results in the best fits to the indices, although residual patterns are very similar to M7 (compare Figs. 3 and 15) and subsequent models. Recent stock assessment results from M20, except for recruitment, are more similar to M19 than M18 (Fig. 16a), but less so historically. Retrospective patterns are much improved compared to M18 (Fig. 16b), but confidence intervals are much wider to reflect the additional uncertainty related to the process errors. Predicted process errors from M20 are mostly negative since around 2008 (Fig. 17), similar to M14, implying decreased survival of cohorts.

M20 estimates of stock size may not be more accurate than M18 estimates, but M20 confidence intervals may cover true stock values more reliably. For example, M18 confidence intervals often do not include retrospective estimates which demonstrates that these confidence intervals are too narrow and not reliable. This problem does not occur with M20.

M21

There are some fairly large residuals in the fit to the differences in the RV_OFF and RV_IO indices (i.e., equation 9), and the distribution of residuals had a long tail. This information on the relationship between the RV_OFF and RV_IO *q*'s could be important in the assessment because it links these two surveys, which together provide the longest index time-series. In M21 the observation error model for these data was changed from the normal to a t-distribution with 3 degrees of freedom, which will be less affected by a small number of outliers than the normal distribution. Otherwise, the M21 formulation was identical to M18.

Although M21 resulted in a much better fit compared to M18 (Table 2), the stock assessment results were very similar (Fig. 18). The fit to the differences in the RV_OFF and RV_IO indices (Fig. 19) was good overall, but there are a few years (2005, 2007, and 2017) when the differences were substantially larger than other years, perhaps because more cod migrated inshore of the survey earlier in the season in these anomalous years. There are a small number of large residuals (i.e., >±5) which is expected due to the robust nature of fitting with a t-distribution with 3 degrees of freedom.

M22

A decision made during the framework meeting was to estimate separate survey *q*'s for the RV_OFF and RV_IO at younger ages. I did not understand the rationale for this decision and I was uncomfortable with it. However, to partially address this issue I removed ages 2–4 in the likelihood component for the differences in the RV_OFF and RV_IO indices. This essentially means that the *q*'s at ages 2–4 are estimated freely for the RV_OFF and RV_IO indices. Note that because this model change involves a change in the data, then AIC and BIC are not comparable with other formulations. I also included age x year correlated process errors, but fixed the correlations at 0.5 because I am concerned that these correlation parameters are difficult to estimate, and when the correlations are high then PE could be confounding with change in F. Fixing the correlations at 0.5 may be a compromise solution to this problem.

The index CV's were estimated to be slightly lower than **M18** (Table 2) which indicates that this formulation resulted in slightly better fits to the survey indices. Assessment results (Fig. 20) were intermediate between **M18** and **M20**. Retrospective patterns for the M22 formulation were improved compared to **M18** (compare Figs. 21 and 13); however, large and systematic process errors (Fig. 22), especially at ages 10–14, are required to achieve this improvement in retrospective patterns. There is some association between predicted process error and F's (Fig. 23). Low negative process errors are often associated with F's close to zero, especially at ages 6–9, and higher values of F are often associated with process errors closer to 0. This requires further research to better understand potential confounding between process errors and F's. Residuals at ages 1–3 in recent years are positive for the RV IO (Fig. 24), similar to other model formulations. The **M22** formulation does not fit this index very well (Fig. 25) at ages 1–3. However, the fit is better at ages 4–8, but the fit is worse at the poorly sampled ages 9–14+.

DISCUSSION

I did not find an assessment model formulation that provided a reasonably good fit to the survey indices and fishery age composition information. The various survey indices often have high between-year variability which is far greater than potential between-year variation in stock size. These indices also disagree with each other to some extent. It seems difficult for any model to fit all the indices well, at the same time. Some lack of fit may be unavoidable. However, most assessment model formulations I considered had serious retrospective patterns in which confidence intervals for current stock size in retrospective peels do not cover final year values. This indicates the potential that the assessment model may provide mis-leading management advice. I was able to substantially reduce retrospective patterns using correlated process errors but the efficacy of this approach requires further research before I could recommend it as a preferred option for stock assessment. I also concluded that further research on assessment inputs is needed before I could recommend a reliable assessment model formulation.

The model formulations I investigated resulted in long run times related to a lack of sparseness in the state-space model estimation. After the framework meeting I determined that this was because of the specific way I modelled process errors. There is an alternative way that will give the same model estimates but with much faster run times.

Tighter landing bounds resulted in estimates of F with more between-year variation. In general, mixed-effect models with zero-mean random effects will estimate or predict these effects to be as close to zero as possible but at the same time to fit the data well. The estimation involves a trade-off between fitting the data well and having small random effects. This is analogous to treating the random effects as fixed effect parameters to estimate, with a penalty function to limit the (e.g., between-year) variation in the estimates. If there is more data with smaller measurement errors then this will tend to result in more variation in estimates of random effects. 3PsSSAM was based on a stochastic F process that included temporal autocorrelation. Hence, when landing bounds are wide then the estimation gives more weight to achieving smooth F estimates over time. Conversely, when landings bounds are tight then landings estimates are forced to be within the bounds, or very close to the bounds, because the negative log-likelihood increases rapidly outside of the bounds. In this case the estimation gives much less weight to achieving smooth F estimates.

ACKNOWLEDGEMENTS

Research funding was provided by the Ocean Frontier Institute, through an award from the Canada First Research Excellence Fund. Research funding was also provided by the Ocean Choice International Industry Research Chair program at the Marine Institute of Memorial University of Newfoundland. I thank the stock assessment science staff of Fisheries and Oceans Canada at the Northwest Atlantic Fisheries Science Center and scientists at the Center for Fisheries Ecosystems Research of Memorial University of Newfoundland for many discussions and contributions that assisted in the development of this paper.

REFERENCES CITED

Albertsen, C.M., Nielsen, A., and Thygesen, U.H., 2016. Choosing the observational likelihood in state-space stock assessment models. Can. J. Fish. Aqua. Sci. 74: 779–789.

Cadigan, N. 2010. Trends in Northwest Atlantic Fisheries Organization (NAFO) Subdivision 3Ps Cod (*Gadus morhua*) stock size based on a separable total mortality model and the Fisheries and Oceans Canada Research Vessel survey index. DFO Can. Sci. Advis. Sec. Res. Doc. 2010/015. iv + 43 p.

Cadigan, N.G. 2016. A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. Can. J. Fish. Aqua. Sci. 73: 296–308.

Cook, R.M. 1997. Stock trends in six North Sea stocks as revealed by an analysis of research vessel surveys. ICES J. Mar. Sci. 54(5): 924–933.

Cook, R.M. 2013. A fish stock assessment model using survey data when estimates of catch are unreliable. Fish. Res. 143: 1–11.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., and Bell, B.M. 2016. TMB: Automatic Differentiation and Laplace Approximation. J. Stat. Soft. 70: 1–26.

Needle, C. 2002. Preliminary analyses of survey indices for whiting in IV and VIId. Working Document WD2 to the ICES Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak, Copenhagen.

Nielsen, A., and Berg, C.W. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fish. Res. 158: 96–101.

Perreault, A.M.J, Wheeland, L.J., Morgan, M.J., and Cadigan, N.G. 2020. A state-space stock assessment model for American plaice on the Grand Bank of Newfoundland. J. Northw. Atl. Fish. Sci. 51: 45–104.

Van Beveren, E., Duplisea, D., Castonguay, M., Doniol-Valcroze, T., Plourde, S., and Cadigan, N. 2017. How catch underreporting can bias stock assessment and advice in northwest Atlantic mackerel and a possible resolution using censored catch. Fish. Res. 194: 146–154.

APPENDIX I – TABLES

Table 1. Summary of preliminary models fit before the pre-assessment framework meeting. Note: IID – independent and identically distributed; ID – independent.

M	Description of Base Model and Changes	Impact/comparison model	NLL	AIC	BIC
1	Simple: No PE. AR1(year, age) logF_dev. CRL sd constant all ages 1 mean rec	-	1896.86	3957.72	4371.54
2	2 CRL sd time-blocks (95–97 and O/W)	little / M1	1833.37	3832.74	4251.61
3	IID Process error (PE)	moderate / M2	1750.91	3669.81	4093.73
4	AR1(age) index and AR1(age, year) CRL errors	moderate / M3	1442.30	3068.61	3532.90
5	F age 3 not correlated with F ages 4+	little / M4	1434.43	3056.85	3531.24
6	PE last year = PE previous year, by age	none / M5	1434.43	3056.85	3531.24
7	CRL sd ages 3–5, 6–8, and 9+	little / M6	1386.85	2969.71	3464.28
8	AR1(year, age) PE	large / M7	1328.65	2855.30	3354.92
9	ID PE, sd by age	little / M7	1337.04	2894.07	3449.21
10	ID PE sd ages 1–4, 5–9, 10+	little / M7	1345.47	2890.94	3395.60
11	AR1(year, age) PE; sd 1–4, 5–9, 10+	moderate / M8	1299.89	2803.78	3318.54
12	AR1(year) PE; sd 1–9, 10+	large / M11	1322.57	2845.14	3349.81
13	AR1(year, age) PE with age/year corr=0.5 fixed; sd 1–9, 10+	moderate / M11	1315.11	2828.23	3327.85
14	AR1=0.5 PE;sd 1–9, 10+ and <=1977/>1977	little / M13	1307.02	2816.04	3325.75
15	AR1=0.5 PE with sd blocks + mean rec blocks, <=1990/>1990	large / M14	1301.56	2807.12	3321.88
16	ID PE with sd blocks + mean rec blocks	minor / M10	1330.41	2862.81	3372.52

Table 2. Estimates of model standard deviation and autocorrelation parameters. nll is the negative log-likelihood, and AIC is Akaike's Information Criterion and BIC is the Bayesian Information Criterion.

Quantity	M17		M18		M19		M20		M21		M22	
-	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
	2988	3491	2991	3489	3022	3521	2912	3421	2812	3296	2931	3430
	EST	CV	EST	CV	EST	CV	EST	CV	EST	CV	EST	CV
std_log_F_3	0.95	23.5	0.99	25.6	1.02	23.1	1.01	31.1	0.99	22.1	0.98	22.9
std_log_F_4p	1.08	27.1	0.94	22.9	0.92	19.3	0.93	20.7	0.90	22.1	0.80	24.8
std_pe_1–4	0.17	18.8	0.17	18.3	0.20	15.8	0.29	14.7	0.13	25.2	0.20	17.7
std_pe_5–9	0.12	14.8	0.12	14.8	0.18	9.5	0.30	9.7	0.13	13.9	0.21	12.2
std_pe_10+	0.48	8.7	0.49	8.4	0.44	8.4	0.68	9.5	0.50	8.4	0.63	8.5
std_log_R	0.37	14.2	0.37	14.1	0.49	16.4	0.46	21.1	0.37	13.6	0.40	18.9
cv_Can_RV_IO	0.78	12.9	0.77	12.6	0.65	10.0	0.71	11.7	0.79	13.0	0.69	11.6
cv_CAN_RV_OFF	0.66	13.9	0.67	14.2	0.78	16.2	0.62	13.9	0.68	14.2	0.64	16.3
cv_ERHAPS	0.65	14.5	0.66	14.6	0.66	14.2	0.60	13.8	0.66	14.7	0.62	13.9
cv_GEAC	0.97	22.0	0.97	21.8	0.98	21.9	0.91	19.8	0.97	22.0	0.99	22.0
cv_SENT_GN	0.89	14.9	0.89	15.0	0.91	15.5	0.86	15.8	0.71	14.4	0.85	15.3
cv_SENT_LT	0.38	11.2	0.38	11.3	0.40	10.8	0.36	10.7	0.35	10.6	0.37	11.6
ar_logF_age	0.98	1.2	0.97	1.5	0.95	1.7	0.96	1.7	0.96	1.6	0.95	2.5
ar_logF_year	0.97	1.7	0.96	2.0	0.95	2.1	0.94	2.6	0.95	2.1	0.94	2.9
ar_logF1_year	0.80	14.6	0.83	13.1	0.81	13.5	0.85	13.8	0.78	15.1	0.79	15.2
ar_pe_year	-	-	-	-	-	-	0.06	187	-	-	0.50	-
ar_pe_age	-	-	-	-	-	-	0.81	4.4	-	-	0.50	-
ar_logRec	0.13	161.4	0.13	163	0.64	22.4	0.50	36.5	0.16	128.1	0.36	58.3
ar_ERHAPS	0.80	4.6	0.80	4.6	0.73	6.4	0.76	5.5	0.82	4.2	0.77	5.3
ar_GEAC	0.81	5.4	0.81	5.5	0.82	4.9	0.76	6.7	0.80	5.6	0.83	5.7
ar_SENT_LT	0.76	6.2	0.77	6.1	0.76	5.7	0.73	7.5	0.77	6.0	0.74	6.9
ar_Can_RV_IO	0.64	12.6	0.64	12.8	0.63	12.9	0.58	15.2	0.64	12.6	0.62	13.4
ar_ SENT_GN	0.63	10.2	0.64	10.1	0.68	8.1	0.67	9.8	0.73	7.8	0.65	10.4
ar_CAN_RV_OFF	0.68	10.9	0.68	10.8	0.66	10.7	0.62	12.9	0.62	12.6	0.63	13.8
ar_crl_year	0.19	42.3	-	-	-	-	-	-	-	-	-	-

Quantity	**M17**		**M18**		**M19**		**M20**		**M21**		**M22**	
-	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
	2988	3491	2991	3489	3022	3521	2912	3421	2812	3296	2931	3430
	EST	CV	EST	CV	EST	CV	EST	CV	EST	CV	EST	CV
ar_crl_age	0.44	12.9	0.42	14.3	0.44	13.9	0.42	15.3	0.42	14.5	0.43	14.1

Table 3. M18 stock size and average F's. rssb is ssb relative to the 1994 value.

-	biomass	cv	ssb	cv	Rec	cv	rssb	cv	aveF_46	cv	aveF_69	cv
1959	340.0	0.07	122.5	0.16	115.9	0.25	3.28	0.19	0.33	0.11	0.35	0.16
1960	343.3	0.07	119.4	0.14	115.2	0.25	3.19	0.17	0.40	0.11	0.44	0.16
1961	316.6	0.07	120.9	0.13	114.9	0.25	3.23	0.17	0.51	0.10	0.56	0.14
1962	275.8	0.08	116.0	0.14	154.8	0.26	3.10	0.18	0.36	0.11	0.40	0.16
1963	263.6	0.08	111.1	0.14	172.7	0.27	2.97	0.18	0.35	0.11	0.39	0.15
1964	259.6	0.07	106.4	0.14	155.0	0.28	2.85	0.17	0.40	0.10	0.44	0.14
1965	260.2	0.07	110.5	0.13	188.5	0.27	2.95	0.17	0.39	0.11	0.44	0.14
1966	276.0	0.07	111.9	0.12	144.4	0.26	2.99	0.16	0.49	0.10	0.55	0.14
1967	280.0	0.07	106.2	0.12	116.0	0.25	2.84	0.16	0.48	0.10	0.53	0.14
1968	304.7	0.07	122.3	0.14	95.6	0.25	3.27	0.17	0.54	0.10	0.60	0.13
1969	275.9	0.06	102.2	0.12	126.3	0.25	2.73	0.16	0.48	0.09	0.53	0.13
1970	258.6	0.06	99.1	0.10	93.4	0.25	2.65	0.15	0.61	0.09	0.65	0.13
1971	213.9	0.06	86.9	0.10	87.9	0.26	2.32	0.14	0.58	0.09	0.63	0.13
1972	175.9	0.06	78.7	0.10	107.1	0.25	2.10	0.14	0.46	0.09	0.53	0.13
1973	156.0	0.06	69.9	0.10	97.7	0.25	1.87	0.14	0.66	0.09	0.68	0.13
1974	142.4	0.06	62.3	0.10	112.1	0.25	1.67	0.14	0.73	0.09	0.75	0.12
1975	136.5	0.07	59.6	0.10	133.2	0.26	1.59	0.14	0.64	0.10	0.64	0.13
1976	143.5	0.08	57.5	0.11	84.4	0.26	1.54	0.15	0.71	0.12	0.63	0.14
1977	136.7	0.09	48.6	0.13	63.4	0.27	1.30	0.16	0.59	0.18	0.53	0.19
1978	126.8	0.07	40.9	0.10	80.2	0.26	1.09	0.14	0.41	0.10	0.40	0.13
1979	155.7	0.06	52.1	0.09	123.0	0.25	1.39	0.13	0.43	0.10	0.40	0.13
1980	170.4	0.06	67.9	0.08	74.6	0.26	1.82	0.13	0.43	0.10	0.41	0.12
1981	162.7	0.06	79.7	0.08	124.8	0.23	2.13	0.12	0.41	0.09	0.40	0.12
1982	173.7	0.06	85.7	0.09	116.5	0.21	2.29	0.13	0.38	0.10	0.36	0.13
1983	172.5	0.06	82.8	0.08	109.9	0.22	2.22	0.12	0.40	0.09	0.35	0.12
1984	180.5	0.06	93.6	0.08	66.0	0.23	2.50	0.12	0.30	0.09	0.29	0.11
1985	203.4	0.06	108.3	0.08	81.9	0.23	2.90	0.12	0.44	0.10	0.40	0.12
1986	191.1	0.07	89.7	0.09	110.0	0.21	2.40	0.13	0.52	0.10	0.48	0.12
1987	164.7	0.10	78.1	0.11	113.7	0.21	2.09	0.14	0.57	0.17	0.53	0.18
1988	136.2	0.10	66.6	0.12	132.1	0.21	1.78	0.14	0.53	0.18	0.53	0.19
1989	124.0	0.10	60.7	0.12	56.0	0.25	1.62	0.14	0.43	0.17	0.46	0.18
1990	113.3	0.09	49.9	0.11	149.3	0.21	1.34	0.14	0.51	0.17	0.55	0.17
1991	102.7	0.08	49.7	0.09	42.8	0.24	1.33	0.13	0.92	0.15	1.04	0.15
1992	80.0	0.09	40.0	0.11	23.9	0.26	1.07	0.13	0.87	0.16	1.02	0.17
1993	70.4	0.08	35.1	0.09	31.9	0.25	0.94	0.06	0.49	0.11	0.62	0.14
1994	66.3	0.10	37.4	0.10	29.9	0.24	1.00	0.00	0.01	0.12	0.02	0.16
1995	87.3	0.08	62.5	0.09	24.3	0.24	1.67	0.06	0.01	0.11	0.01	0.14
1996	95.7	0.08	74.0	0.08	26.5	0.23	1.98	0.08	0.01	0.09	0.01	0.13
1997	109.3	0.07	89.1	0.08	21.3	0.24	2.38	0.09	0.12	0.10	0.13	0.13
1998	110.0	0.08	93.1	0.08	33.6	0.22	2.49	0.10	0.26	0.23	0.27	0.25
1999	102.4	0.08	84.9	0.09	49.6	0.23	2.27	0.11	0.41	0.11	0.42	0.13

-	biomass	cv	ssb	cv	Rec	cv	rssb	cv	aveF_46	cv	aveF_69	cv
2000	95.8	0.10	75.9	0.11	33.3	0.20	2.03	0.12	0.42	0.17	0.45	0.18
2001	74.4	0.08	48.9	0.10	18.8	0.19	1.31	0.12	0.40	0.18	0.49	0.20
2002	67.8	0.07	45.3	0.08	25.7	0.21	1.21	0.11	0.38	0.10	0.46	0.12
2003	66.2	0.07	49.4	0.08	24.0	0.22	1.32	0.11	0.37	0.09	0.43	0.12
2004	65.5	0.08	51.8	0.09	21.1	0.24	1.39	0.11	0.38	0.18	0.40	0.19
2005	65.6	0.07	51.6	0.08	28.8	0.23	1.38	0.10	0.36	0.10	0.42	0.12
2006	58.1	0.08	45.4	0.08	20.0	0.23	1.21	0.10	0.37	0.10	0.43	0.13
2007	52.3	0.08	38.9	0.09	41.9	0.18	1.04	0.10	0.40	0.10	0.45	0.13
2008	44.5	0.08	30.8	0.10	17.8	0.22	0.82	0.11	0.43	0.10	0.48	0.13
2009	43.2	0.09	28.0	0.11	24.0	0.19	0.75	0.11	0.40	0.18	0.43	0.20
2010	42.3	0.08	26.5	0.10	29.6	0.21	0.71	0.11	0.33	0.10	0.37	0.14
2011	37.4	0.08	25.4	0.10	21.4	0.24	0.68	0.11	0.33	0.11	0.35	0.15
2012	36.9	0.09	25.5	0.10	61.6	0.17	0.68	0.11	0.23	0.10	0.26	0.16
2013	35.8	0.09	22.5	0.10	28.6	0.18	0.60	0.11	0.20	0.10	0.24	0.18
2014	36.4	0.09	24.1	0.10	23.2	0.20	0.64	0.11	0.28	0.12	0.32	0.19
2015	37.0	0.10	23.6	0.11	12.7	0.25	0.63	0.12	0.24	0.12	0.29	0.18
2016	35.7	0.11	23.9	0.12	16.6	0.23	0.64	0.12	0.28	0.14	0.30	0.20
2017	34.2	0.12	25.8	0.14	26.2	0.24	0.69	0.14	0.27	0.16	0.29	0.20
2018	37.0	0.15	29.3	0.16	25.2	0.28	0.78	0.16	0.19	0.18	0.21	0.22
2019	38.1	0.18	29.4	0.20	27.4	0.39	0.79	0.20	-	-	-	-

Table 4. M20 stock size and average F's. rssb is ssb relative to the 1994 value.

-	biomass	cv	ssb	cv	Rec	cv	rssb	cv	aveF_46	cv	aveF_69	cv
1959	360.3	0.24	136.5	0.33	118.1	0.40	2.77	0.41	0.32	0.28	0.32	0.35
1960	363.5	0.22	133.4	0.31	115.7	0.40	2.70	0.38	0.37	0.27	0.40	0.35
1961	325.8	0.22	130.3	0.29	113.4	0.40	2.64	0.37	0.50	0.28	0.52	0.34
1962	269.5	0.24	117.2	0.32	160.6	0.41	2.38	0.39	0.38	0.28	0.39	0.37
1963	262.6	0.23	114.4	0.32	175.3	0.41	2.32	0.39	0.37	0.27	0.37	0.35
1964	255.6	0.22	106.5	0.30	162.0	0.42	2.16	0.38	0.43	0.26	0.43	0.34
1965	268.3	0.23	123.9	0.30	183.4	0.41	2.51	0.37	0.38	0.26	0.39	0.34
1966	279.5	0.21	119.2	0.27	147.7	0.41	2.41	0.35	0.50	0.25	0.51	0.33
1967	305.8	0.21	117.5	0.27	128.9	0.40	2.38	0.35	0.44	0.25	0.45	0.32
1968	335.7	0.20	134.9	0.26	108.1	0.41	2.73	0.35	0.49	0.24	0.51	0.31
1969	274.2	0.20	104.3	0.26	142.1	0.40	2.11	0.34	0.48	0.24	0.50	0.31
1970	288.7	0.19	112.1	0.24	113.2	0.40	2.27	0.33	0.54	0.24	0.55	0.31
1971	240.0	0.20	98.7	0.24	104.7	0.41	2.00	0.32	0.51	0.24	0.52	0.33
1972	188.6	0.20	89.7	0.25	127.7	0.41	1.82	0.33	0.43	0.24	0.44	0.33
1973	177.1	0.19	83.2	0.24	116.4	0.41	1.69	0.32	0.58	0.23	0.53	0.33
1974	165.5	0.20	75.5	0.24	119.8	0.42	1.53	0.31	0.63	0.24	0.59	0.32
1975	146.2	0.20	67.0	0.23	122.1	0.44	1.36	0.30	0.58	0.24	0.52	0.32
1976	155.6	0.20	64.0	0.22	68.0	0.46	1.30	0.29	0.64	0.24	0.50	0.31
1977	144.9	0.23	55.0	0.26	45.3	0.49	1.12	0.32	0.60	0.26	0.47	0.30
1978	102.2	0.18	39.5	0.21	51.8	0.48	0.80	0.25	0.51	0.21	0.42	0.26
1979	131.1	0.16	55.1	0.19	87.7	0.42	1.12	0.24	0.51	0.21	0.41	0.25
1980	153.0	0.17	71.2	0.20	63.2	0.40	1.44	0.24	0.49	0.22	0.42	0.25
1981	162.6	0.18	86.2	0.20	127.3	0.37	1.75	0.24	0.42	0.21	0.35	0.26
1982	193.5	0.19	107.3	0.22	140.7	0.35	2.17	0.23	0.35	0.22	0.27	0.29
1983	197.7	0.19	106.3	0.22	148.3	0.36	2.15	0.21	0.35	0.23	0.26	0.29
1984	206.9	0.19	114.8	0.22	91.9	0.36	2.33	0.21	0.27	0.22	0.22	0.27
1985	258.4	0.19	149.6	0.22	124.3	0.36	3.03	0.22	0.37	0.23	0.28	0.29
1986	276.8	0.20	134.6	0.23	175.0	0.36	2.73	0.22	0.35	0.23	0.28	0.29
1987	235.8	0.21	116.1	0.24	193.1	0.37	2.35	0.22	0.45	0.32	0.37	0.36
1988	186.7	0.22	97.3	0.24	192.2	0.37	1.97	0.22	0.49	0.32	0.44	0.37
1989	155.1	0.23	76.2	0.25	104.2	0.40	1.54	0.23	0.42	0.30	0.41	0.34
1990	172.2	0.21	73.8	0.24	249.1	0.38	1.50	0.22	0.41	0.32	0.44	0.34
1991	152.3	0.20	74.7	0.21	93.7	0.39	1.51	0.23	0.54	0.39	0.58	0.41
1992	76.9	0.23	38.3	0.23	30.2	0.40	0.78	0.22	0.48	0.41	0.54	0.44
1993	75.9	0.25	39.4	0.26	35.6	0.39	0.80	0.20	0.41	0.36	0.47	0.41
1994	84.4	0.26	49.3	0.28	29.6	0.39	1.00	0.00	0.01	0.29	0.01	0.33
1995	178.2	0.23	133.9	0.24	29.2	0.40	2.71	0.17	0.00	0.26	0.00	0.29
1996	177.3	0.24	137.5	0.25	33.9	0.39	2.79	0.20	0.01	0.25	0.01	0.29
1997	163.8	0.24	129.0	0.25	33.0	0.40	2.61	0.22	0.08	0.26	0.09	0.30
1998	150.3	0.22	124.2	0.22	48.2	0.39	2.52	0.20	0.17	0.25	0.18	0.28
1999	113.7	0.20	89.5	0.21	67.1	0.38	1.81	0.20	0.37	0.23	0.38	0.26

-	biomass	cv	ssb	cv	Rec	cv	rssb	cv	aveF_46	cv	aveF_69	cv
2000	154.0	0.21	121.3	0.21	65.0	0.38	2.46	0.21	0.32	0.22	0.33	0.25
2001	108.6	0.21	66.8	0.22	35.2	0.36	1.35	0.21	0.31	0.28	0.36	0.31
2002	93.1	0.21	59.1	0.21	41.5	0.37	1.20	0.21	0.30	0.22	0.37	0.25
2003	92.9	0.21	69.0	0.21	33.4	0.37	1.40	0.21	0.30	0.27	0.34	0.29
2004	101.3	0.21	80.2	0.21	33.2	0.40	1.63	0.21	0.29	0.23	0.31	0.25
2005	88.0	0.21	68.3	0.21	45.2	0.39	1.38	0.21	0.28	0.23	0.32	0.27
2006	97.7	0.22	75.1	0.22	42.8	0.39	1.52	0.23	0.26	0.28	0.29	0.31
2007	78.2	0.21	57.5	0.21	74.3	0.35	1.17	0.21	0.27	0.23	0.29	0.26
2008	71.8	0.21	49.9	0.23	33.5	0.38	1.01	0.22	0.27	0.26	0.29	0.30
2009	76.6	0.22	51.0	0.23	41.9	0.36	1.03	0.21	0.24	0.27	0.23	0.31
2010	69.5	0.22	42.6	0.23	55.2	0.39	0.86	0.21	0.20	0.23	0.21	0.28
2011	65.4	0.22	42.8	0.22	52.9	0.41	0.87	0.21	0.18	0.24	0.18	0.28
2012	67.1	0.23	46.5	0.23	119.7	0.36	0.94	0.21	0.12	0.24	0.12	0.28
2013	76.2	0.24	43.9	0.24	77.7	0.38	0.89	0.21	0.10	0.24	0.10	0.29
2014	57.4	0.23	37.2	0.22	43.6	0.36	0.75	0.21	0.17	0.24	0.17	0.29
2015	53.4	0.22	32.8	0.22	25.3	0.38	0.66	0.21	0.17	0.23	0.18	0.28
2016	51.5	0.21	34.1	0.21	27.6	0.36	0.69	0.21	0.19	0.23	0.18	0.28
2017	46.4	0.22	34.4	0.22	40.5	0.36	0.70	0.21	0.19	0.25	0.19	0.29
2018	49.4	0.25	38.0	0.26	41.6	0.40	0.77	0.25	0.14	0.28	0.14	0.32
2019	49.6	0.35	36.2	0.38	44.1	0.52	0.73	0.36	-	-	-	-

APPENDIX II – STOCK ASSESSMENT MODEL FIGURES

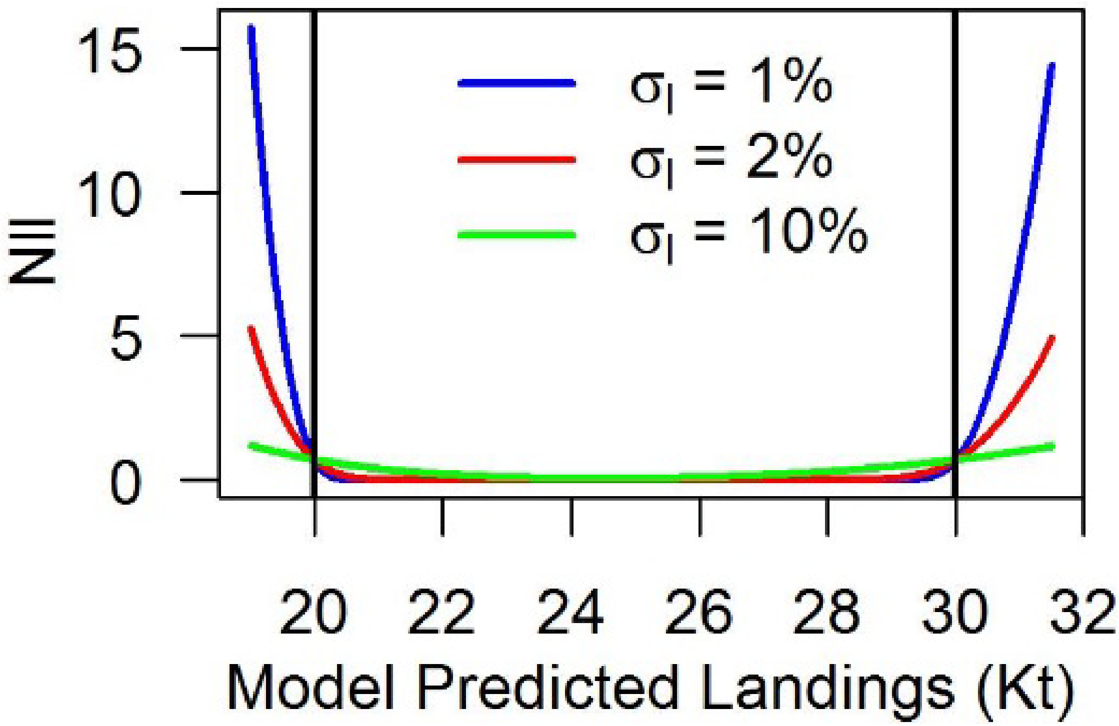

Figure 1. Censored nll for a range of model predicted landings, using three choices of σ_l. Vertical solid black lines indicate the lower and upper bounds.

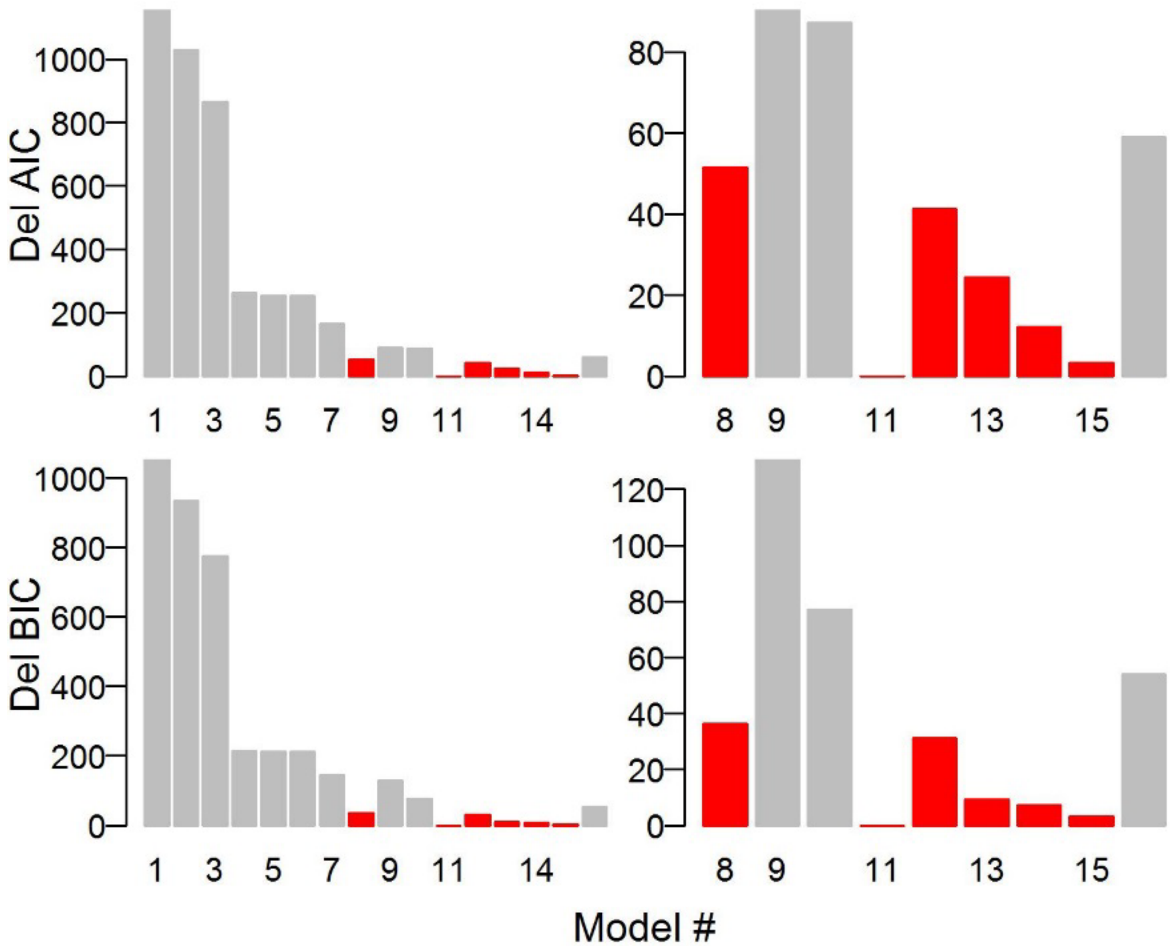

Figure 2. Summary of changes in AIC/BIC for preliminary models fit before the pre-assessment framework meeting. Grey bars indicate models with iid process errors, and red bars indicate models with autocorrelated process errors. Left column panels are for all models, and the right column panels are for models 8–16.

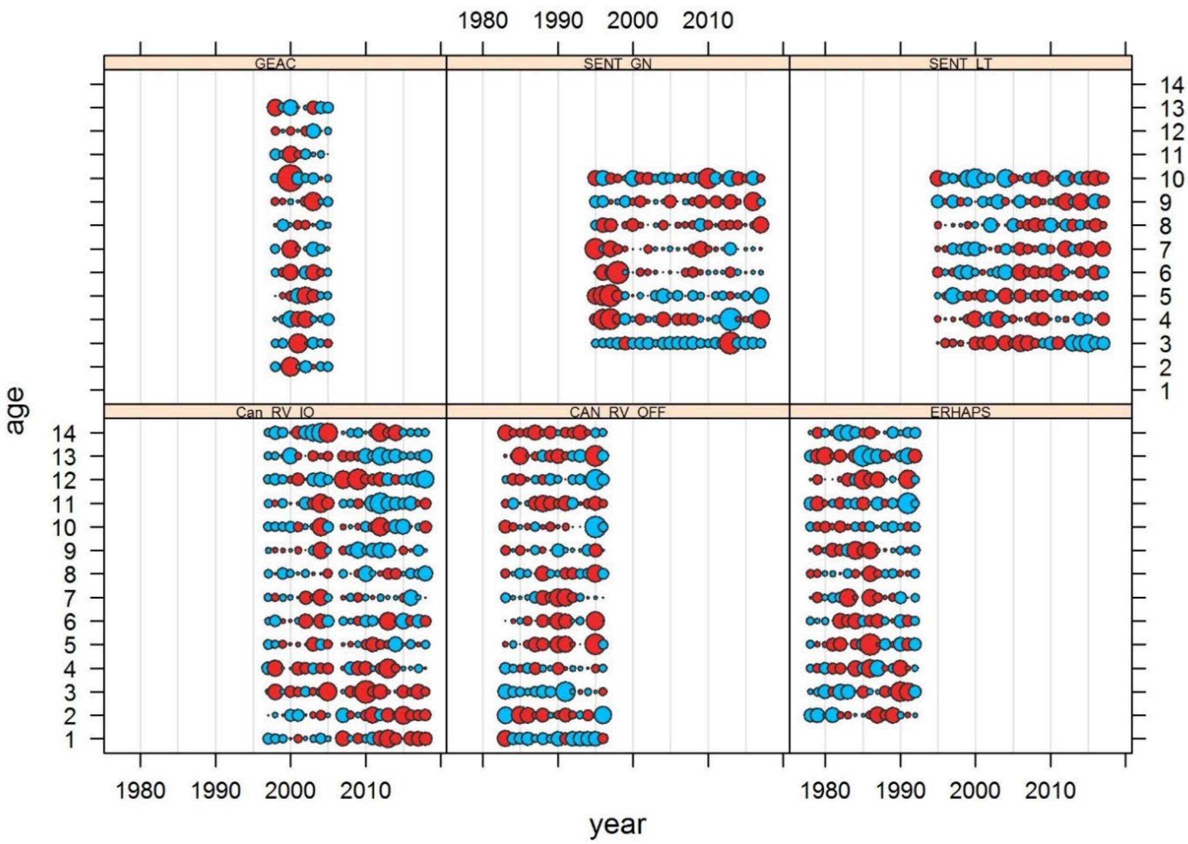

Figure 3. Standardized index residuals, e_z. from M7. Red circles are positive and blue are negative. The size of a circle is proportional to $|e_z|^{1/2}$. These residuals should be approximately uncorrelated across ages, within surveys and years. Each panel is for a separate survey.

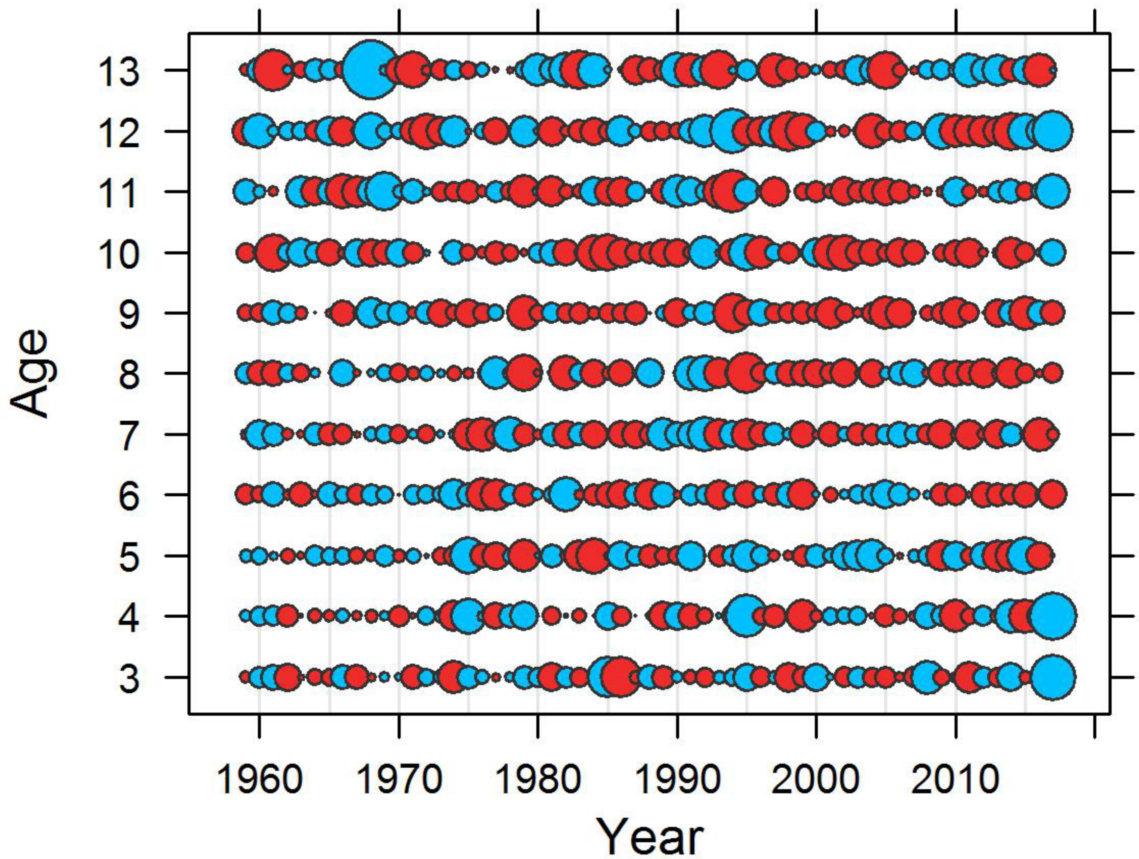

Figure 4. M7 standardized CRL residuals.

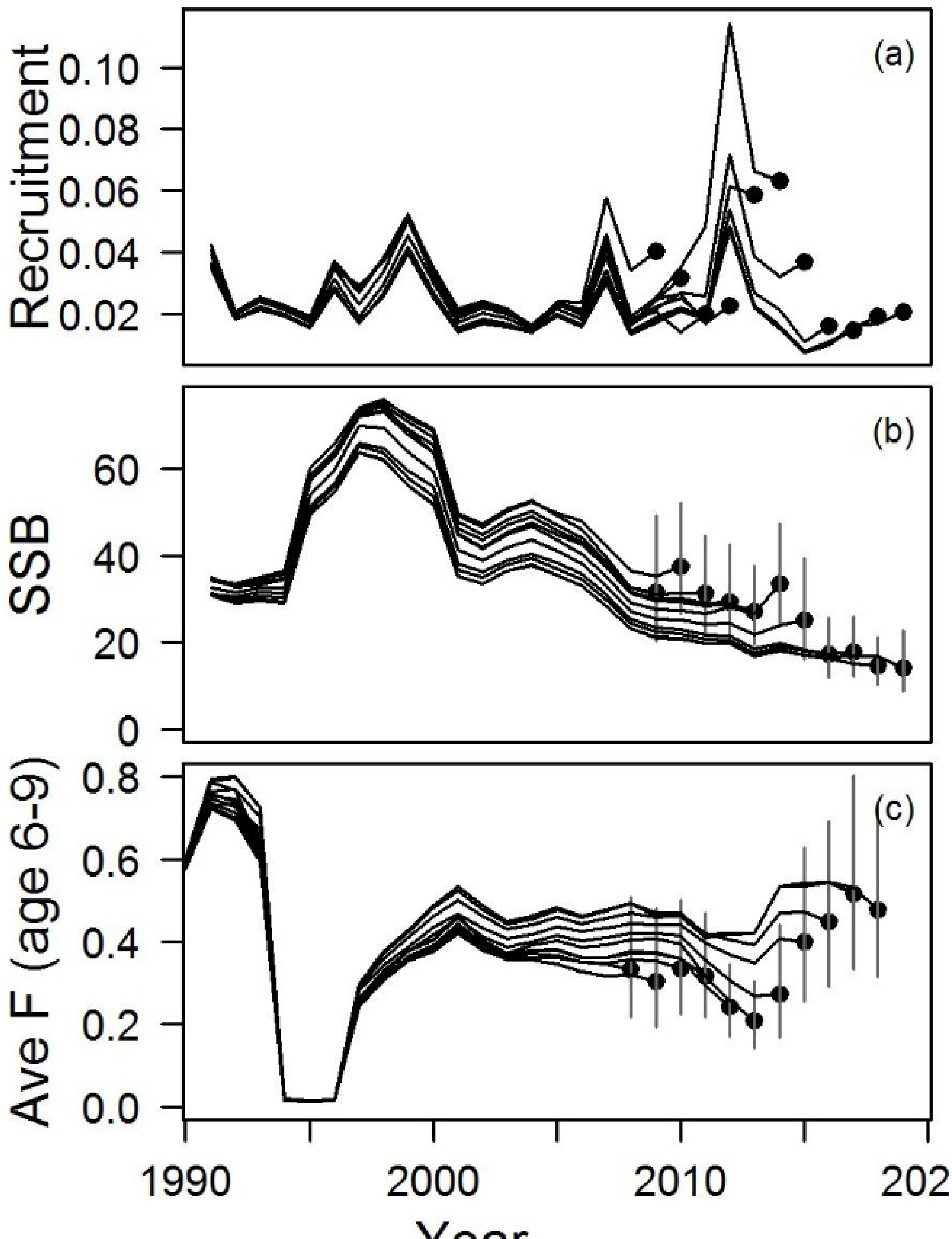

Figure 5. M6 retrospective estimates of recruitment, SSB, and average F at ages 6–9. Solid points indicate end retrospective year estimates, and solid vertical line segments indicate 95% confidence intervals.

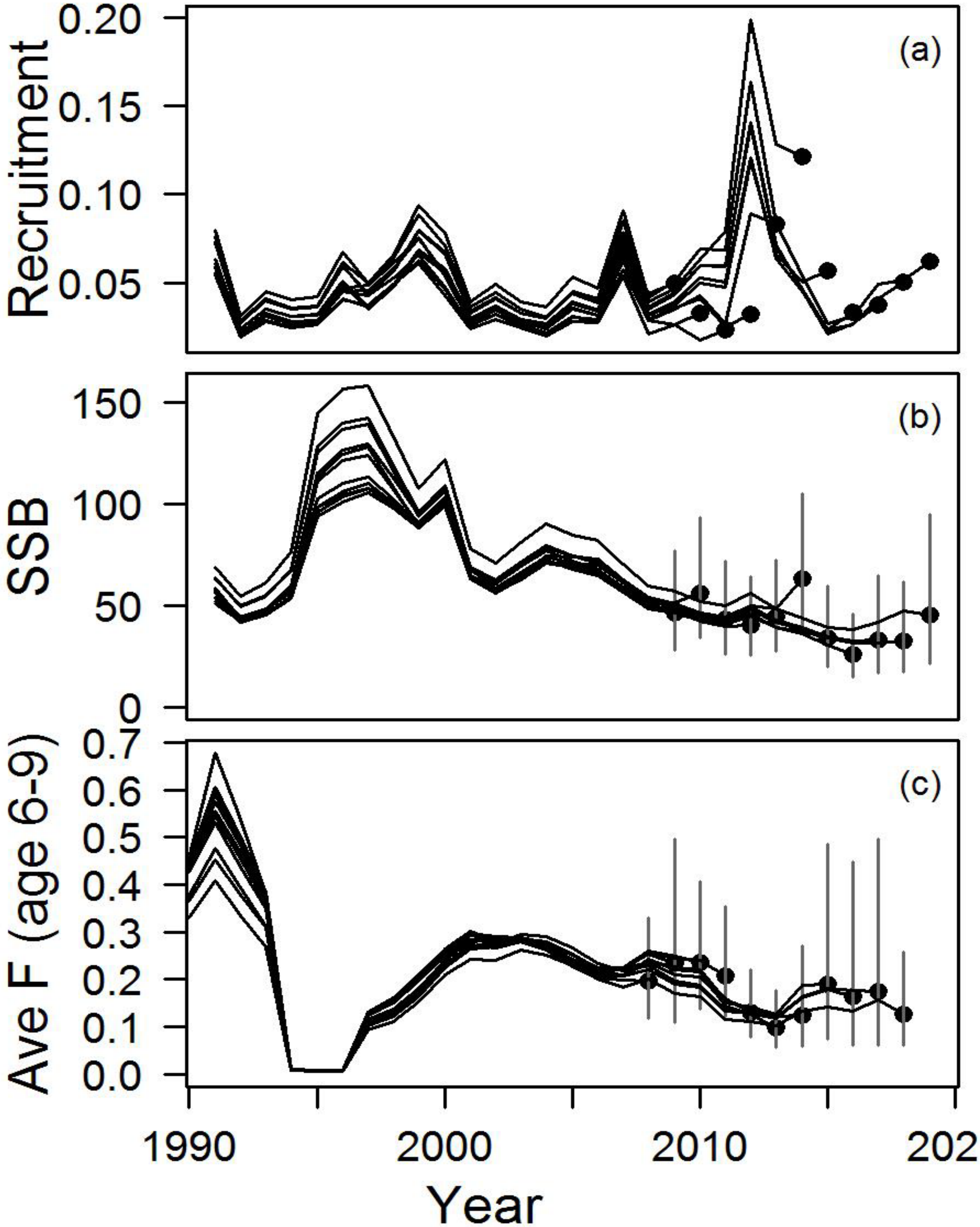

Figure 6. M13 retrospective model estimates.

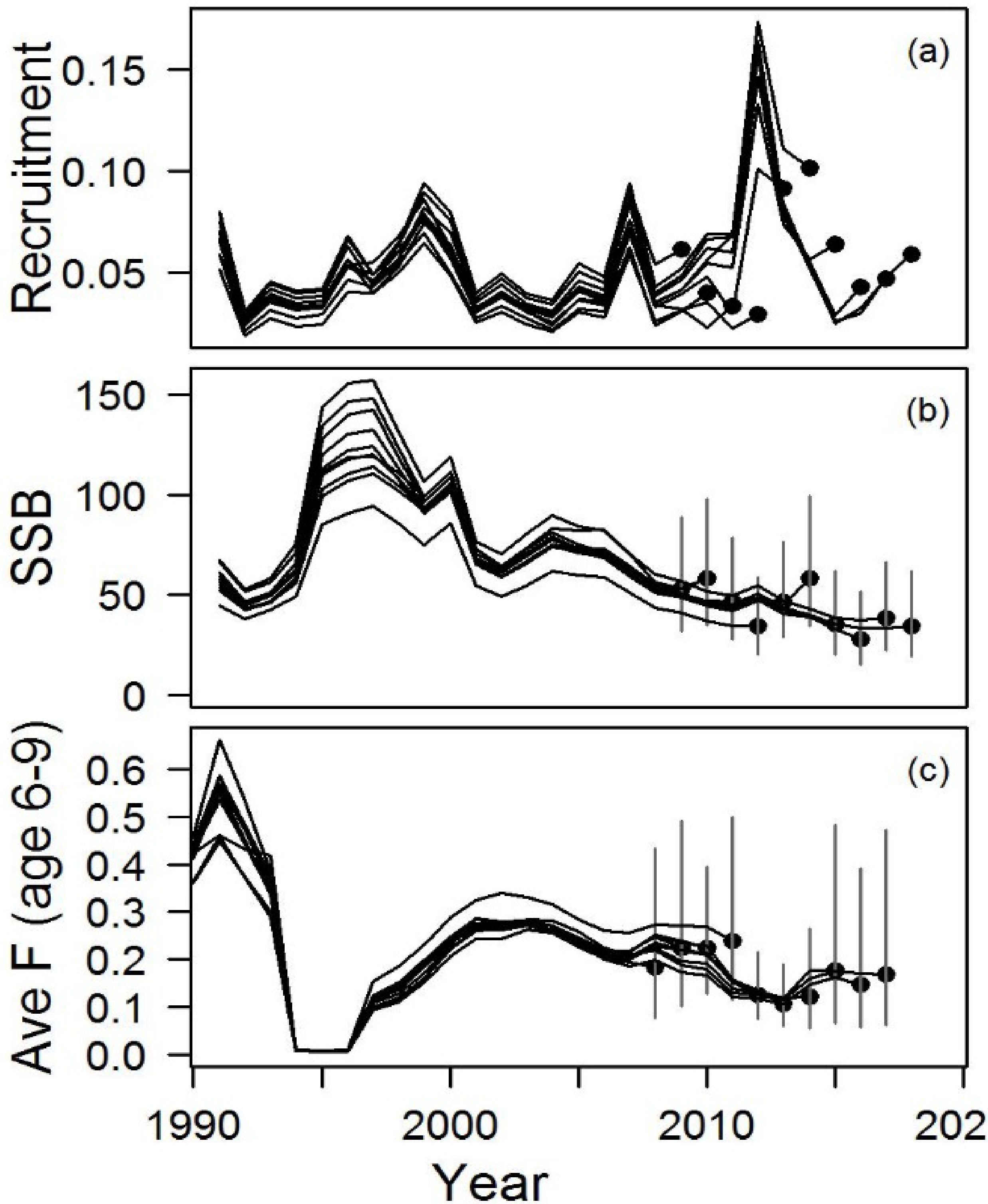

Figure 7. M14 retrospective model estimates.

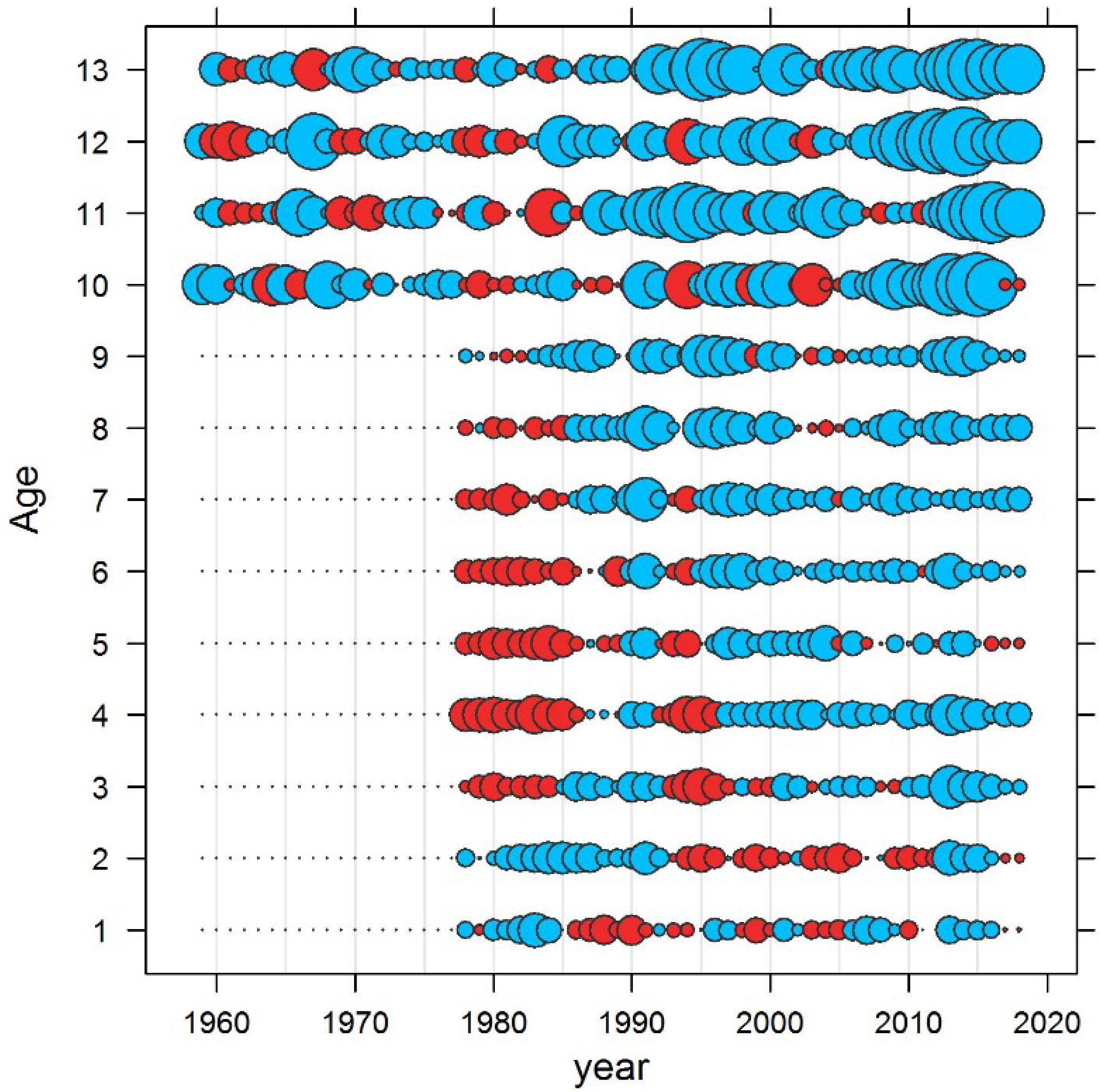

Figure 8. M14 predicted process errors (δ). Red circles are positive and blue are negative. The size of a circle is proportional to $|\hat{\boldsymbol{\delta}}|^{1/2}$.

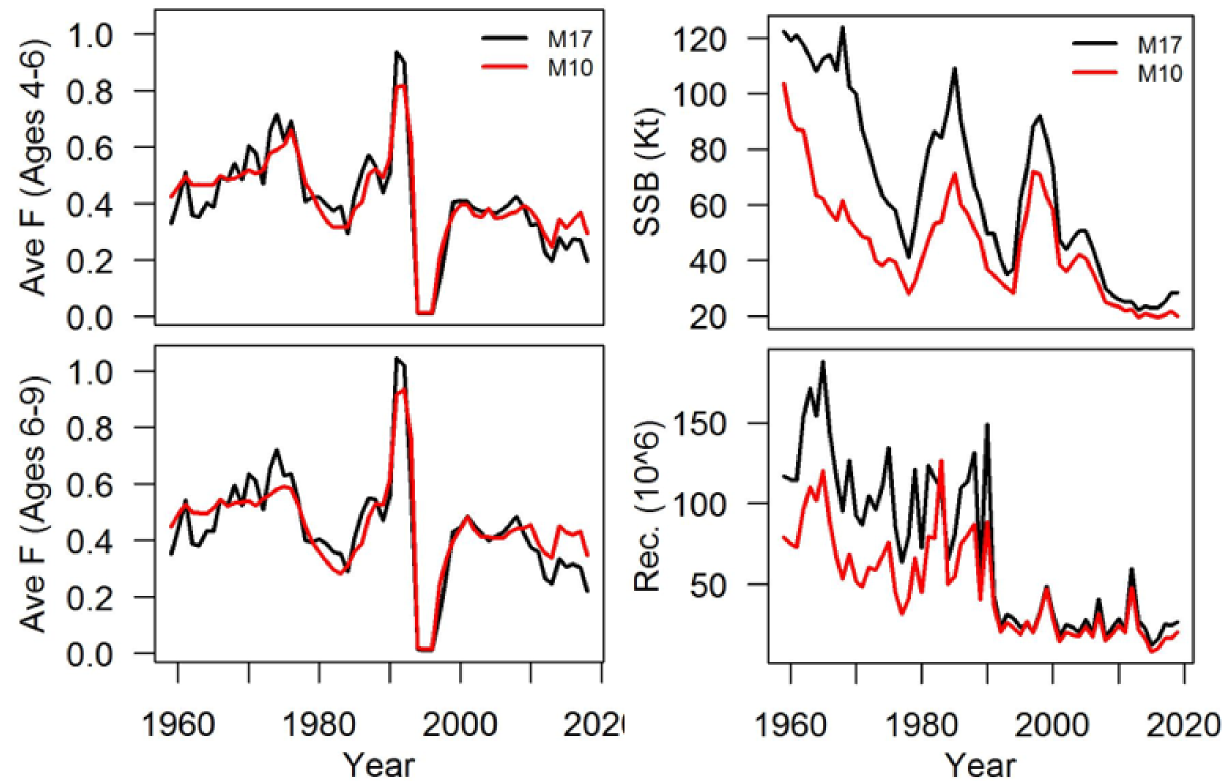

Figure 9. A comparison of average F's, SSB, and recruitment for M17 and M10.

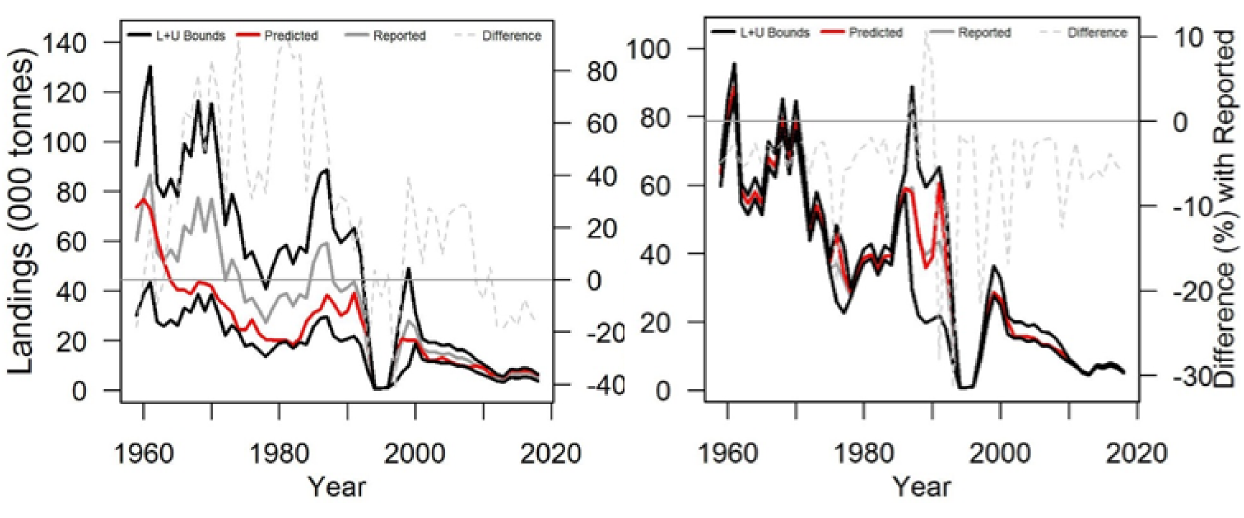

Figure 10. A comparison of model predicted and assumed landings bounds for M10 (left) and M17 (right).

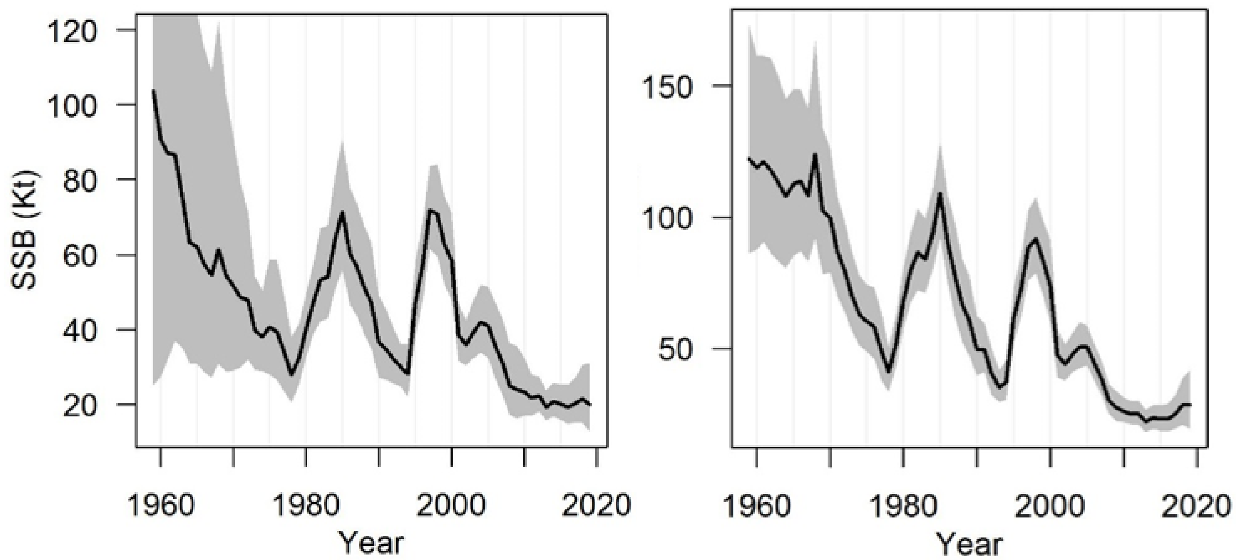

Figure 11. A comparison of SSB uncertainty for M10 (left) and M17 (right).

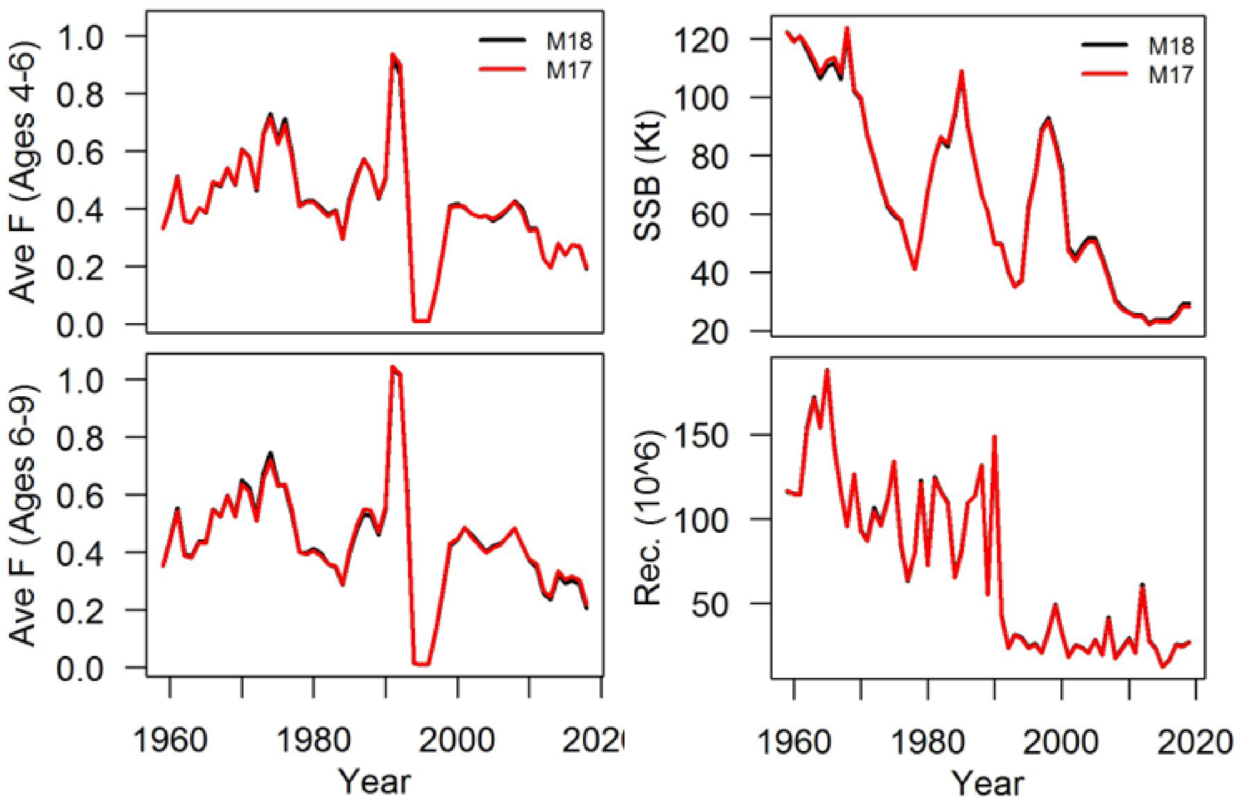

Figure 12. A comparison of average F's, SSB, and recruitment for M17 and M18.

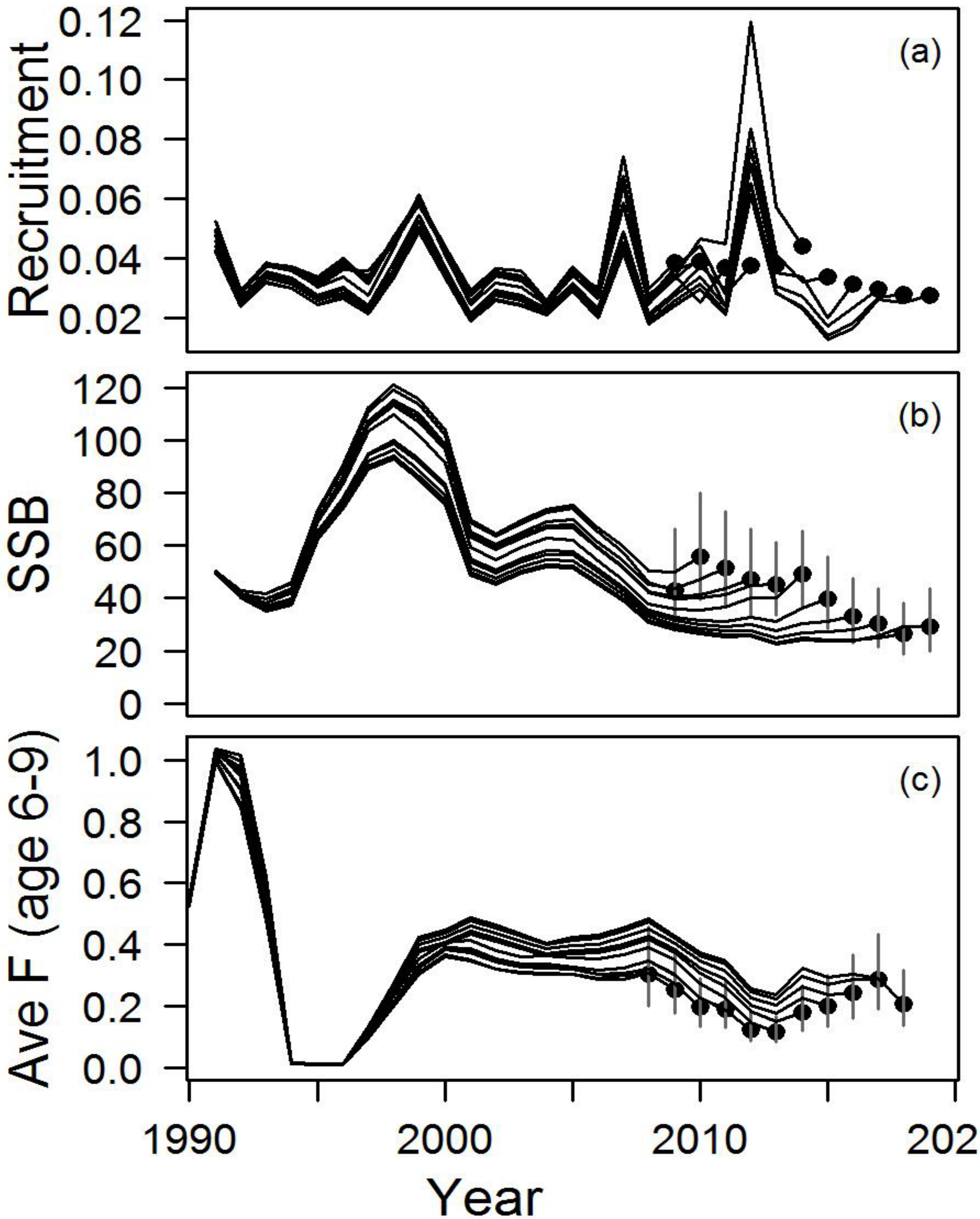

Figure 13. M18 retrospective model estimates.

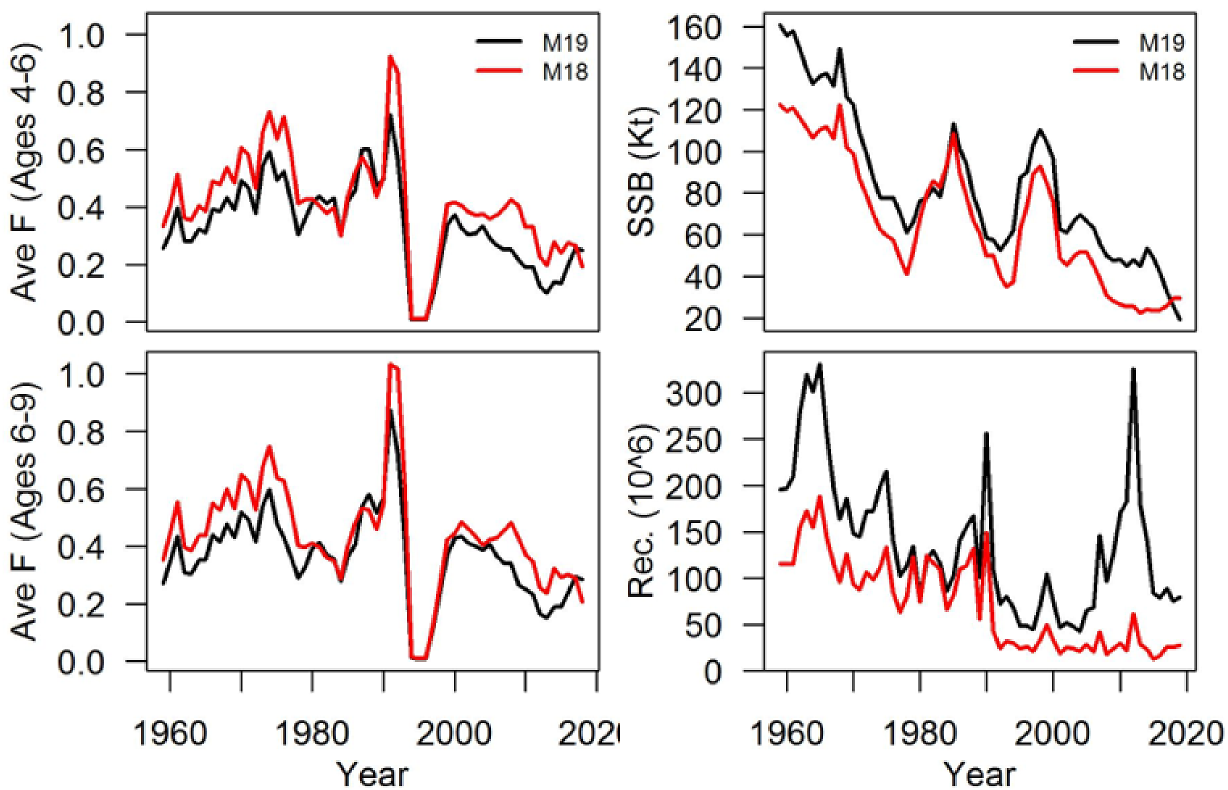

Figure 14a. A comparison of average F's, SSB, and recruitment for M18 and M19.

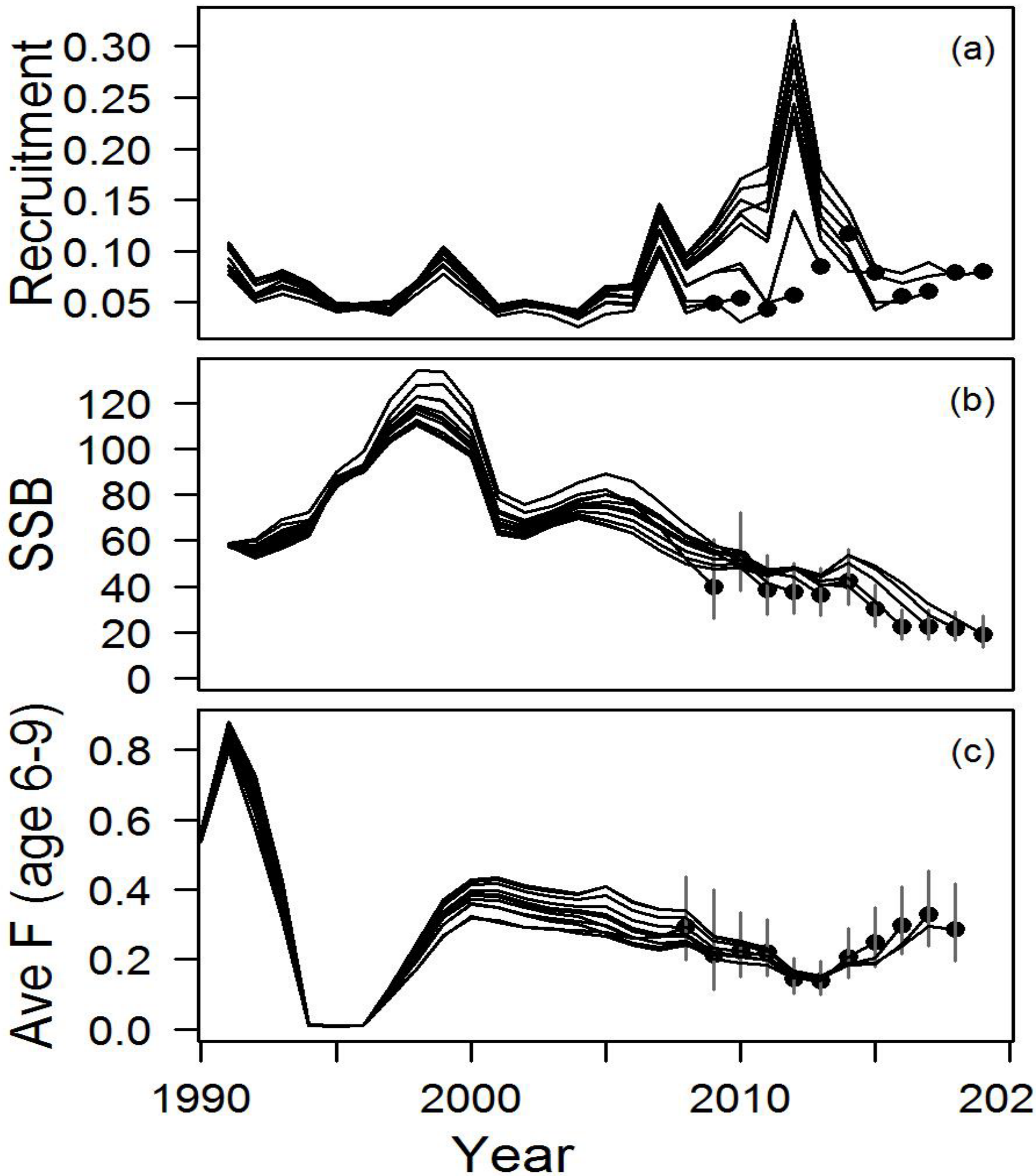

Figure 14b. M19 retrospective model estimates.

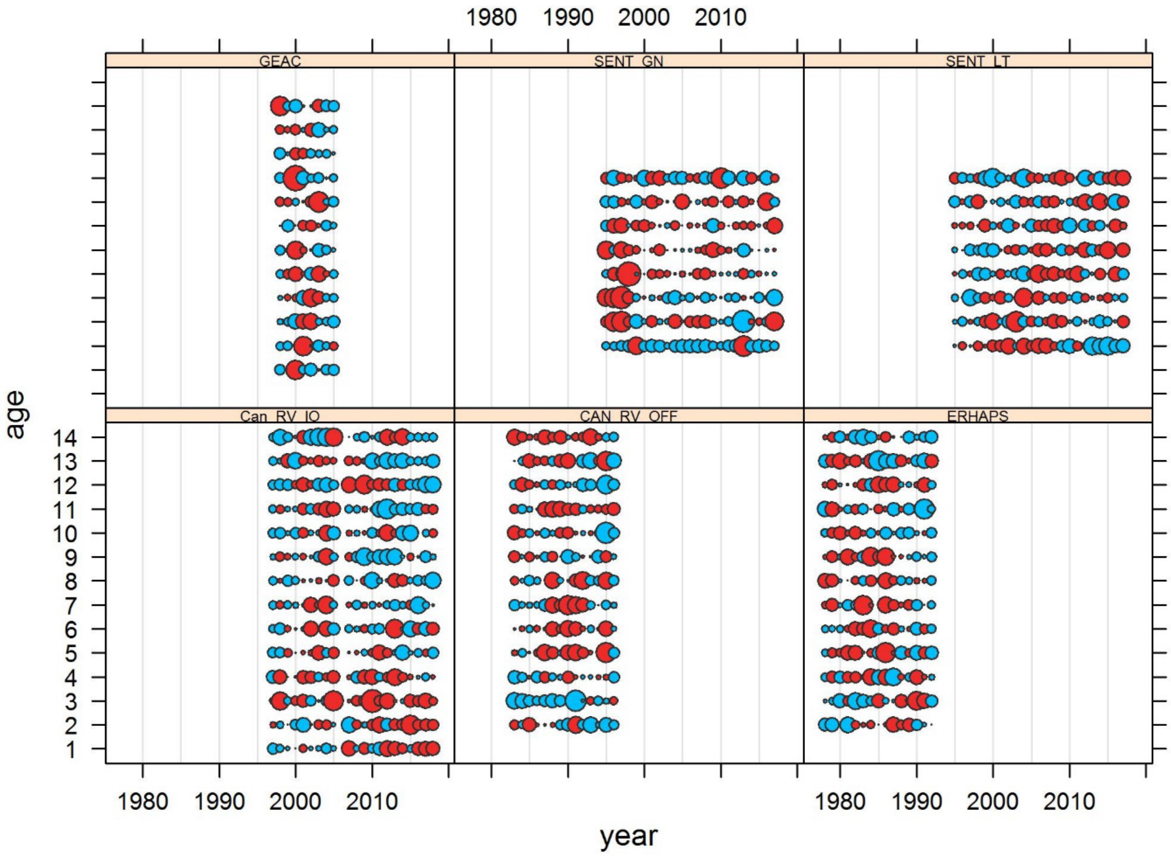

Figure 15. Standardized index residuals, e_z. from M20. Red circles are positive and blue are negative. The size of a circle is proportional to $|e_z|^{1/2}$. These residuals should be approximately uncorrelated across ages, within surveys and years. Each panel is for a separate survey.

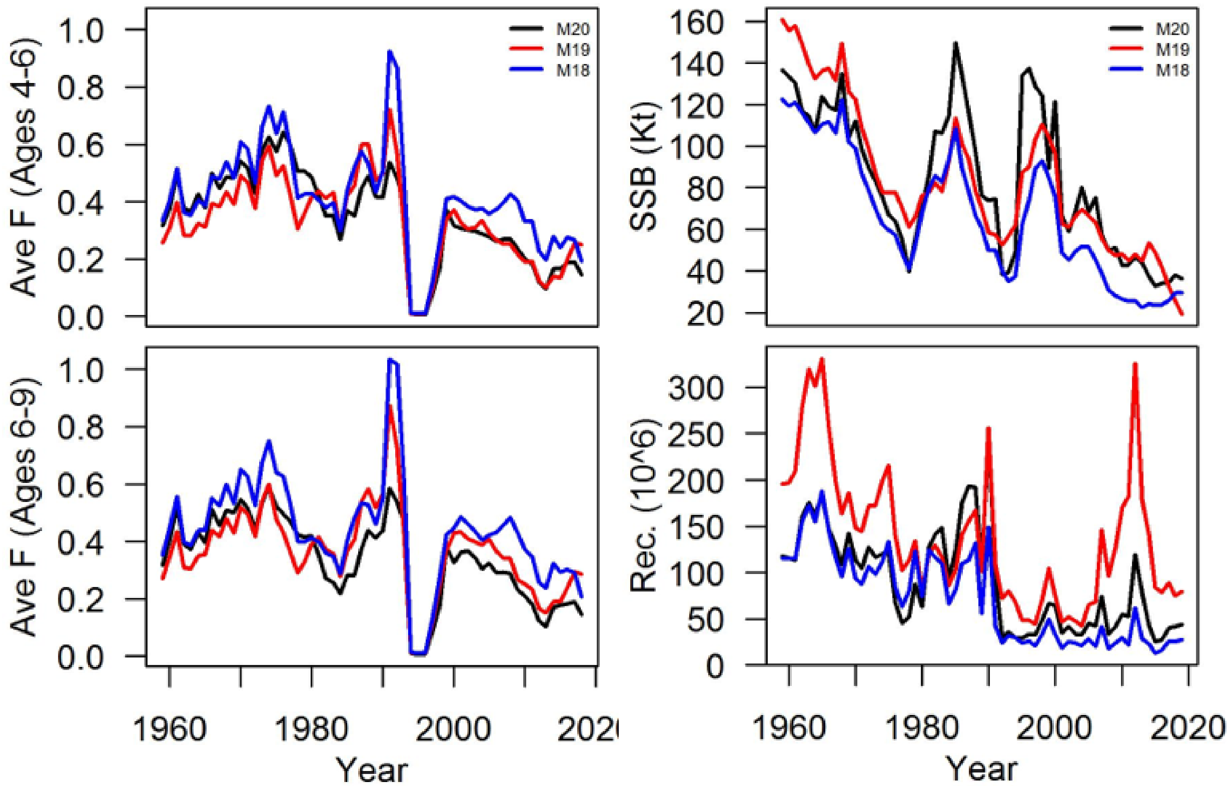

Figure 16a. A comparison of average F's, SSB, and recruitment for M18–20.

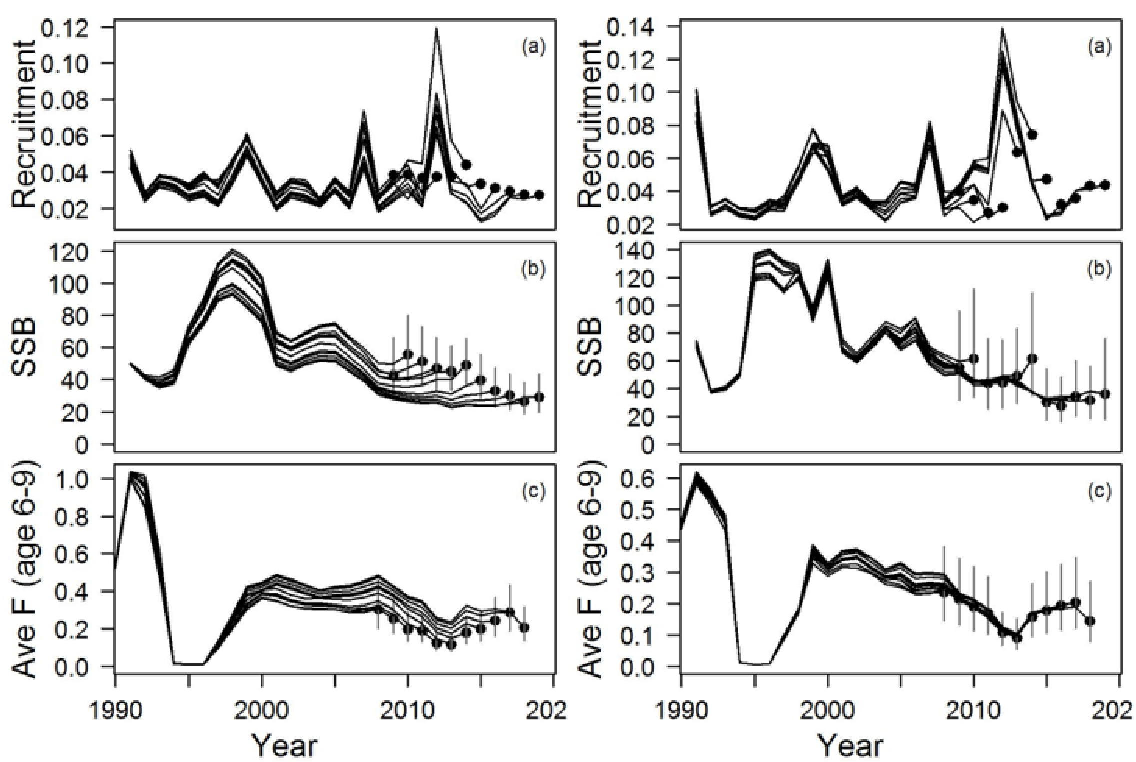

Figure 16b. A comparison of retrospective patterns from M18 (left column) and M20 (right column).

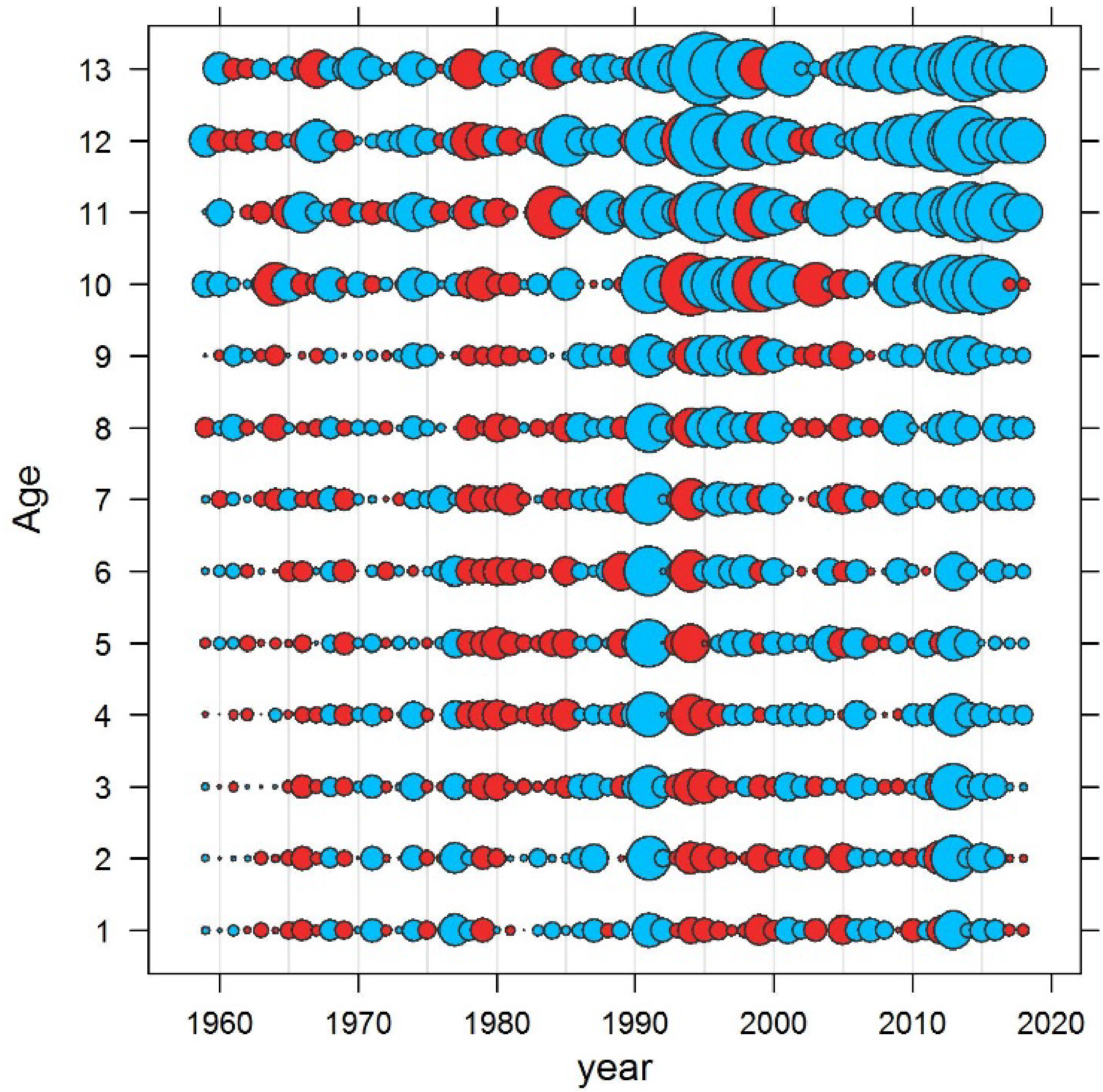

Figure 17. M20 predicted process errors (δ). Red circles are positive and blue are negative. The size of a circle is proportional to $|\hat{\boldsymbol{\delta}}|^{1/2}$.

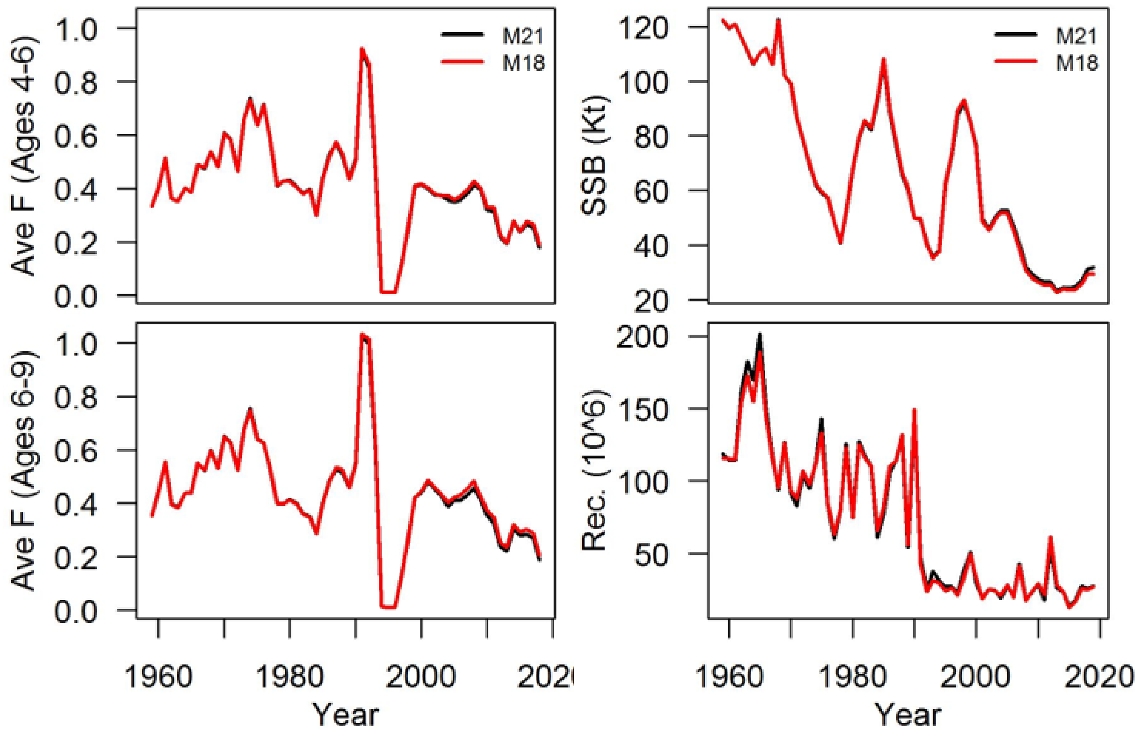

Figure 18. A comparison of average F's, SSB, and recruitment for M18 and M21.

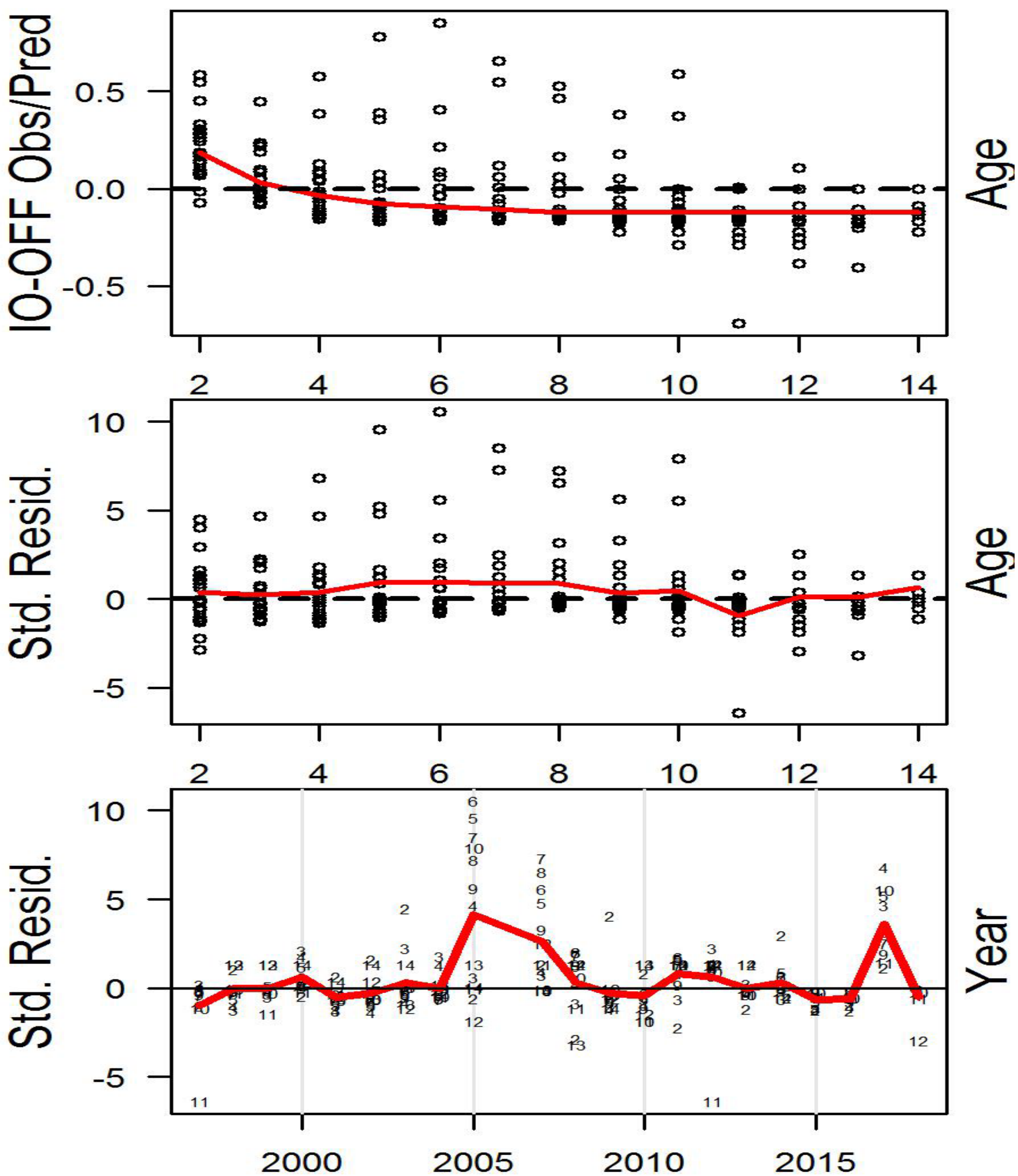

Figure 19. M21 standardized residuals for the difference in RV_IN and RV_OFF indices during 1997–2018. The top panel shows the observed RV differences (points) and model predictions (lines). The model predictions are the differences in log(q) for the two sets of strata. Residuals are plotted versus age (middle panel) and year (bottom panel).

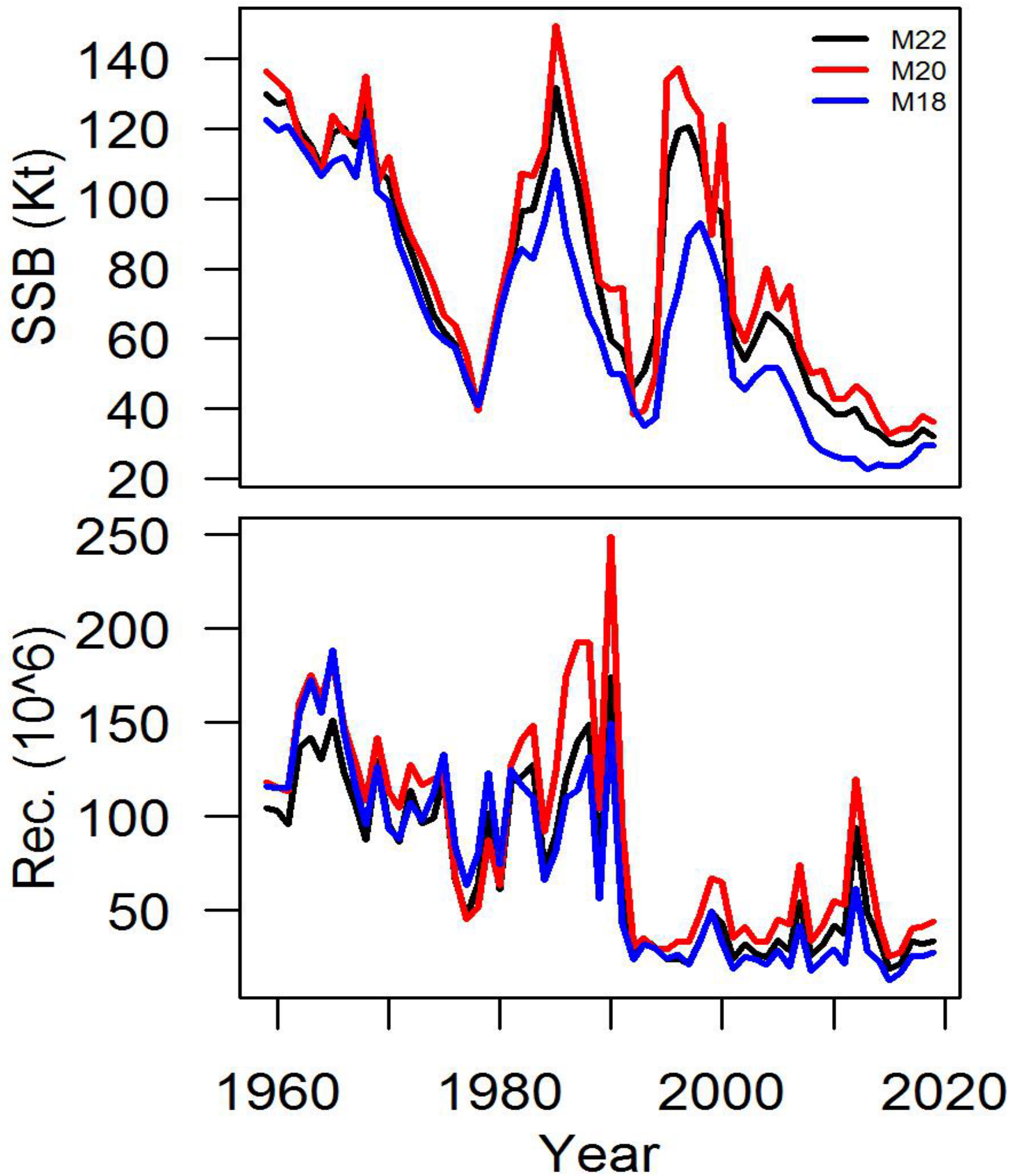

Figure 20. A comparison of average SSB, and recruitment for M18, M20, and M22.

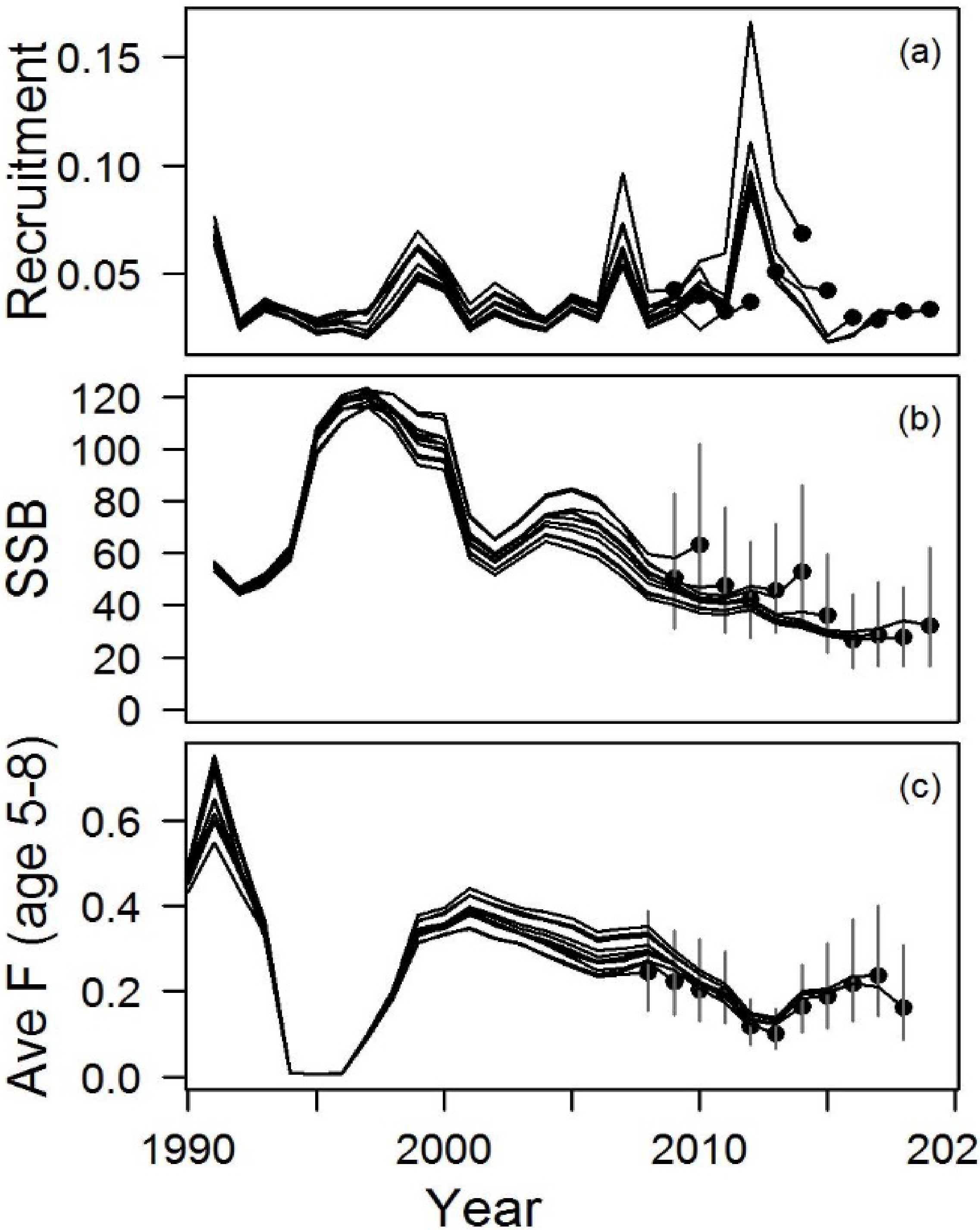

Figure 21. M22 retrospective model estimates.

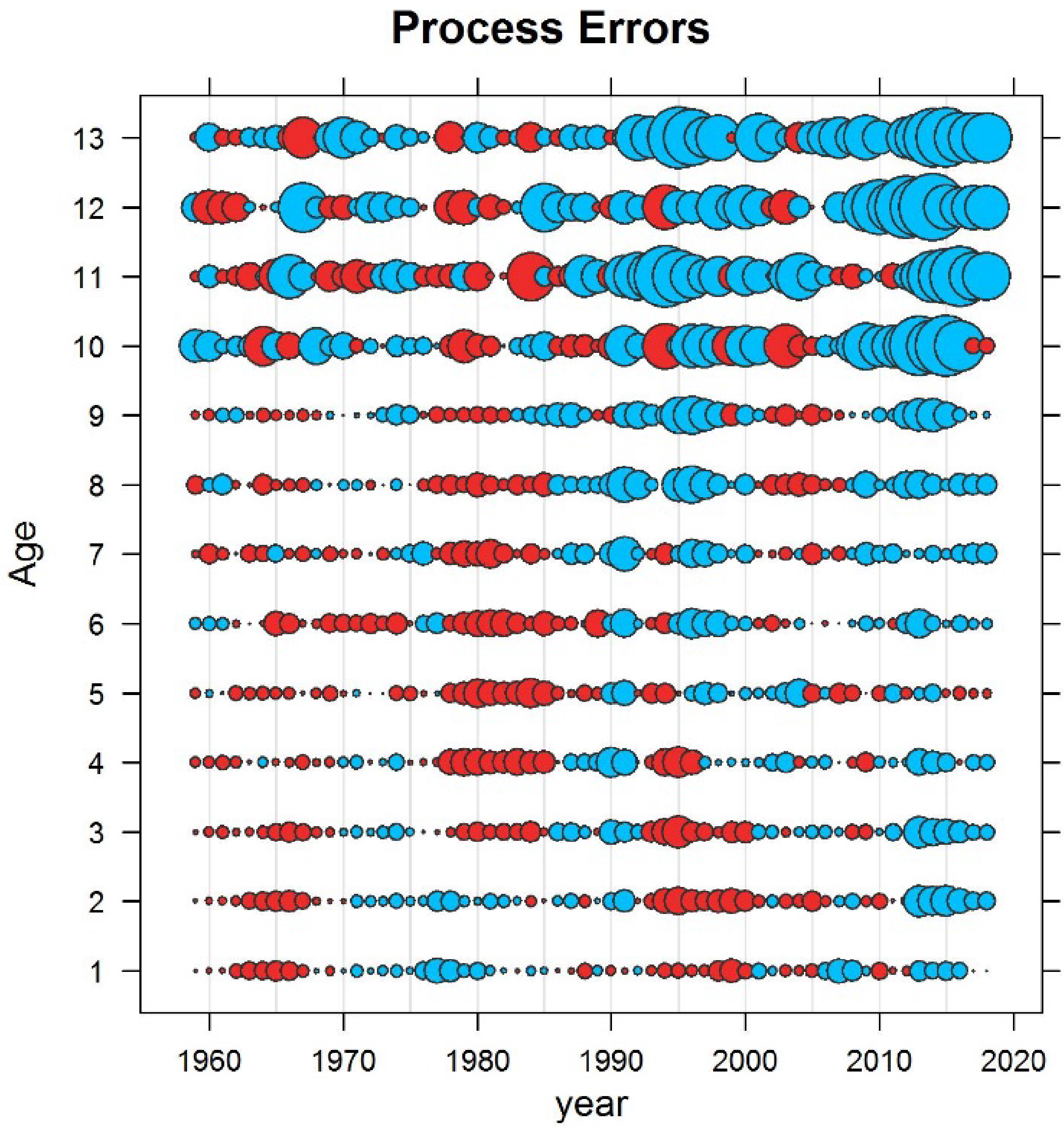

Figure 22. M22 predicted process errors (δ). Red circles are positive and blue are negative. The size of a circle is proportional to $|\hat{\delta}|^{1/2}$.

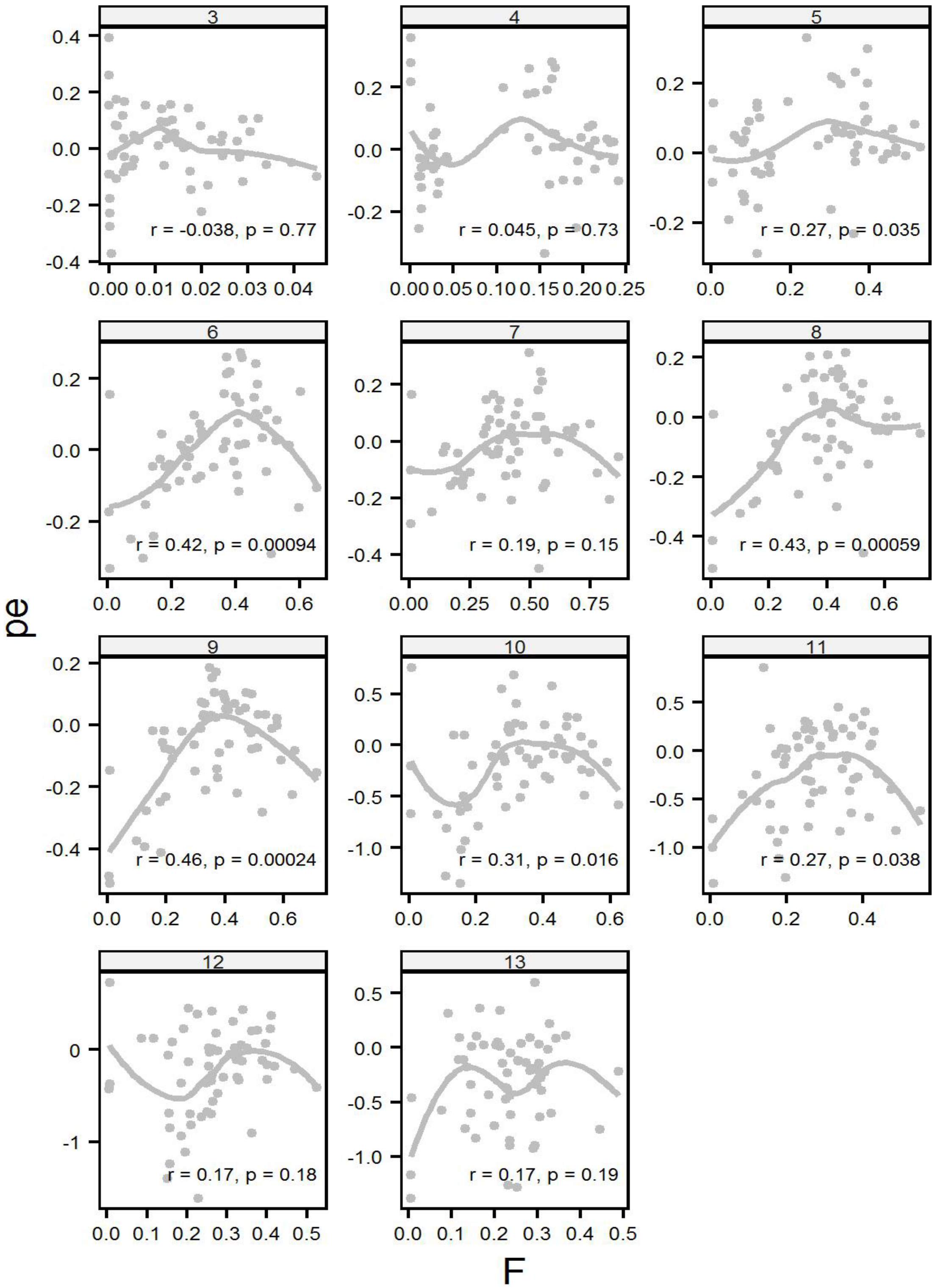

Figure 23. M22 predicted process errors versus F's, with correlations coefficients (r) and p-values. Grey curves indicate loess smoother results.

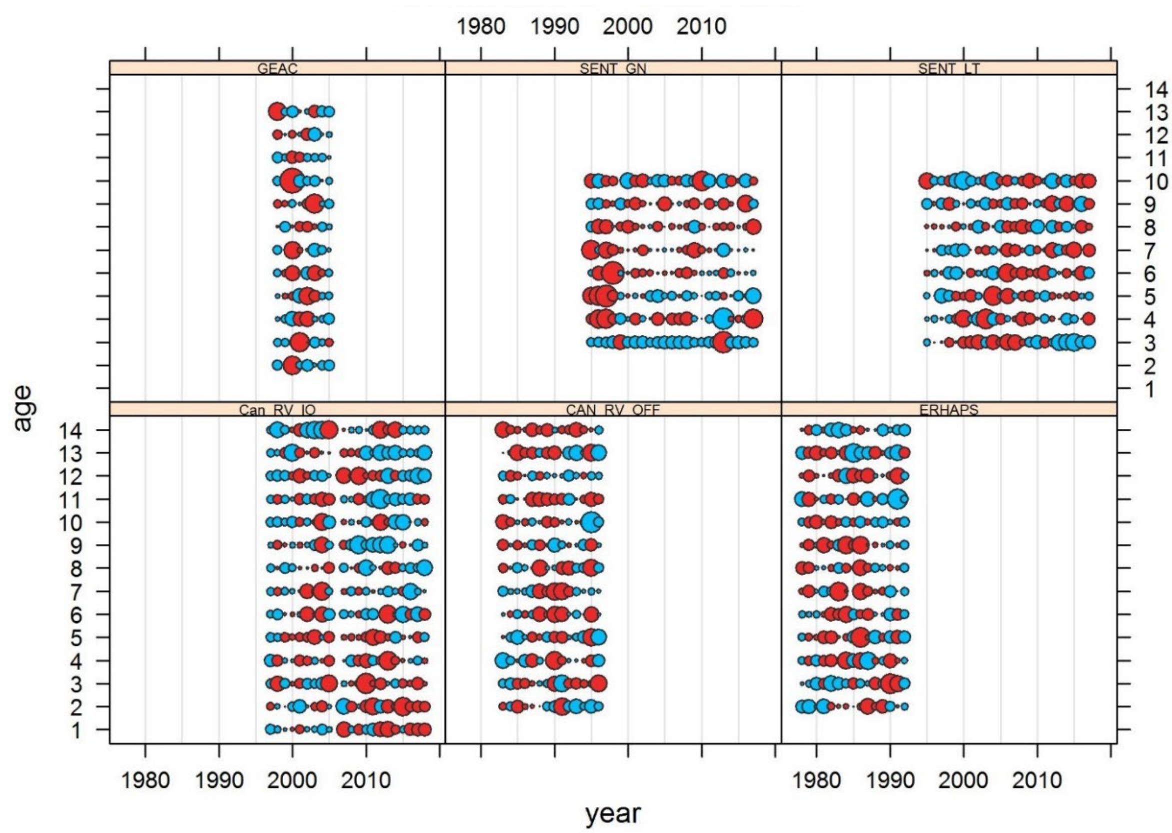

Figure 24. Standardized index residuals, e_z. from M22. Red circles are positive, and blue are negative. The size of a circle is proportional to $|e_z|^{1/2}$. These residuals should be approximately uncorrelated across ages, within surveys and years. Each panel is for a separate survey.

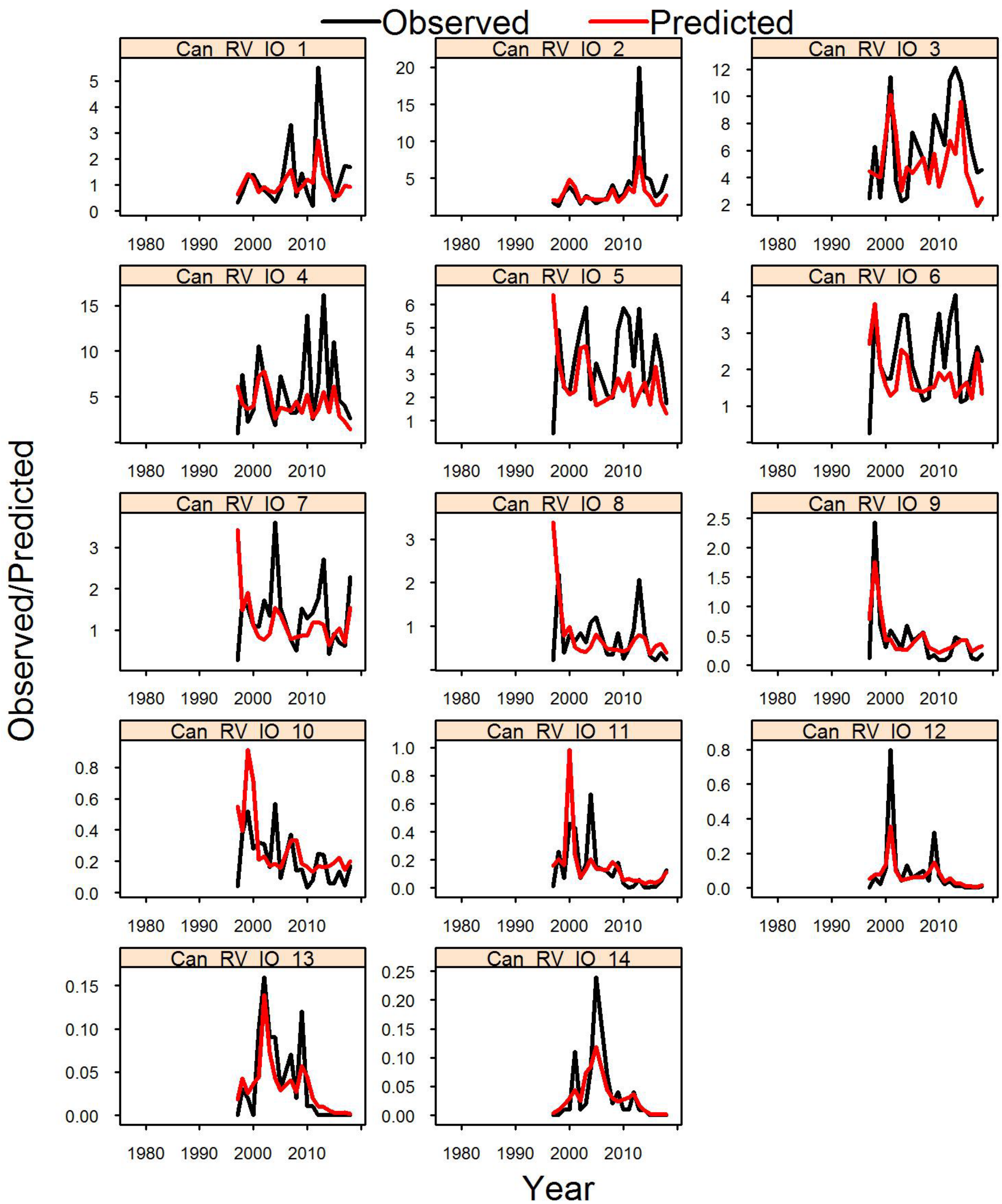

Figure 25. Observed (black) M22 predicted (red) RV IO indices.

APPENDIX III – DATA FIGURES

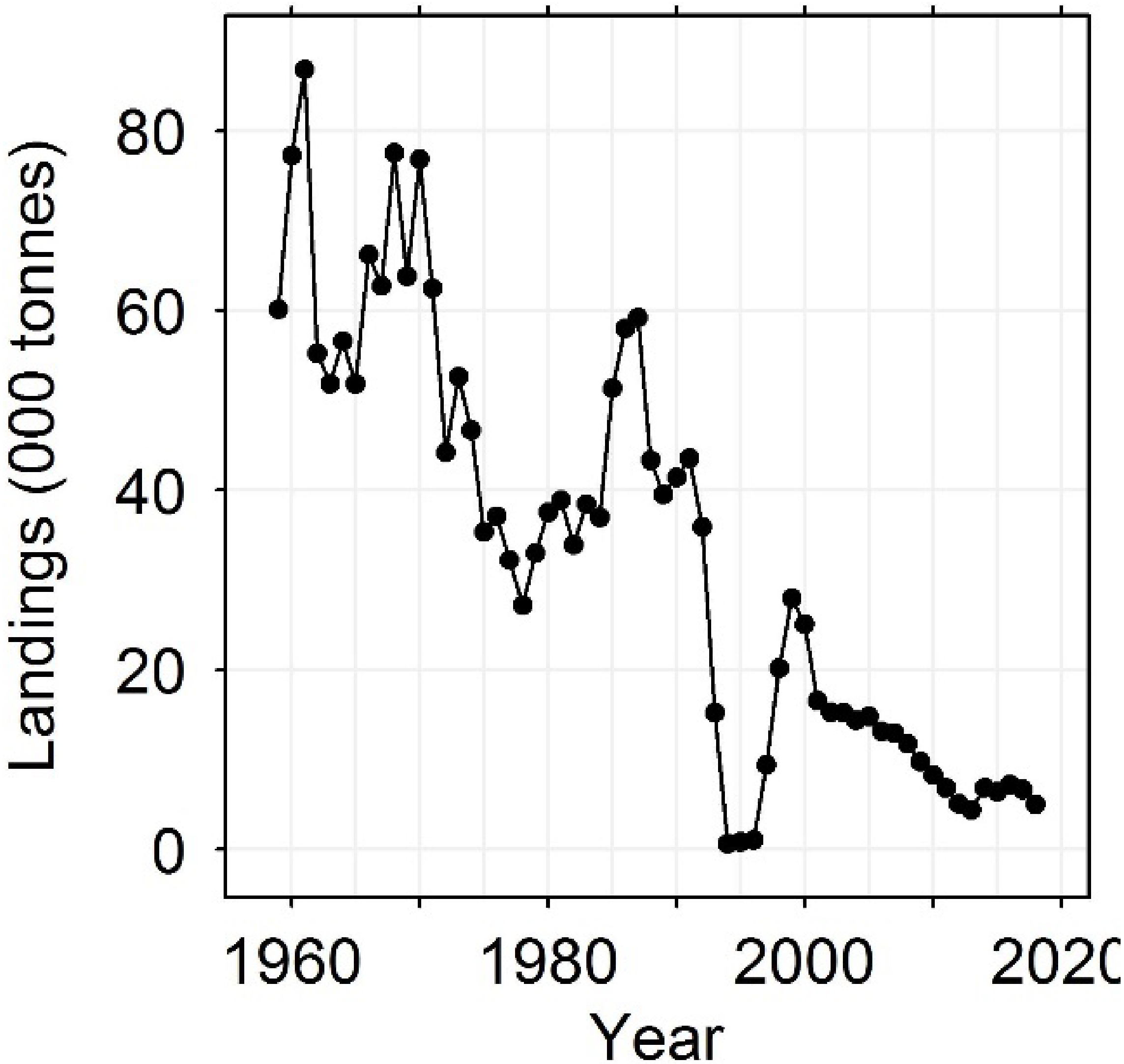

Figure D1. Time series of 3Ps cod reported landings.

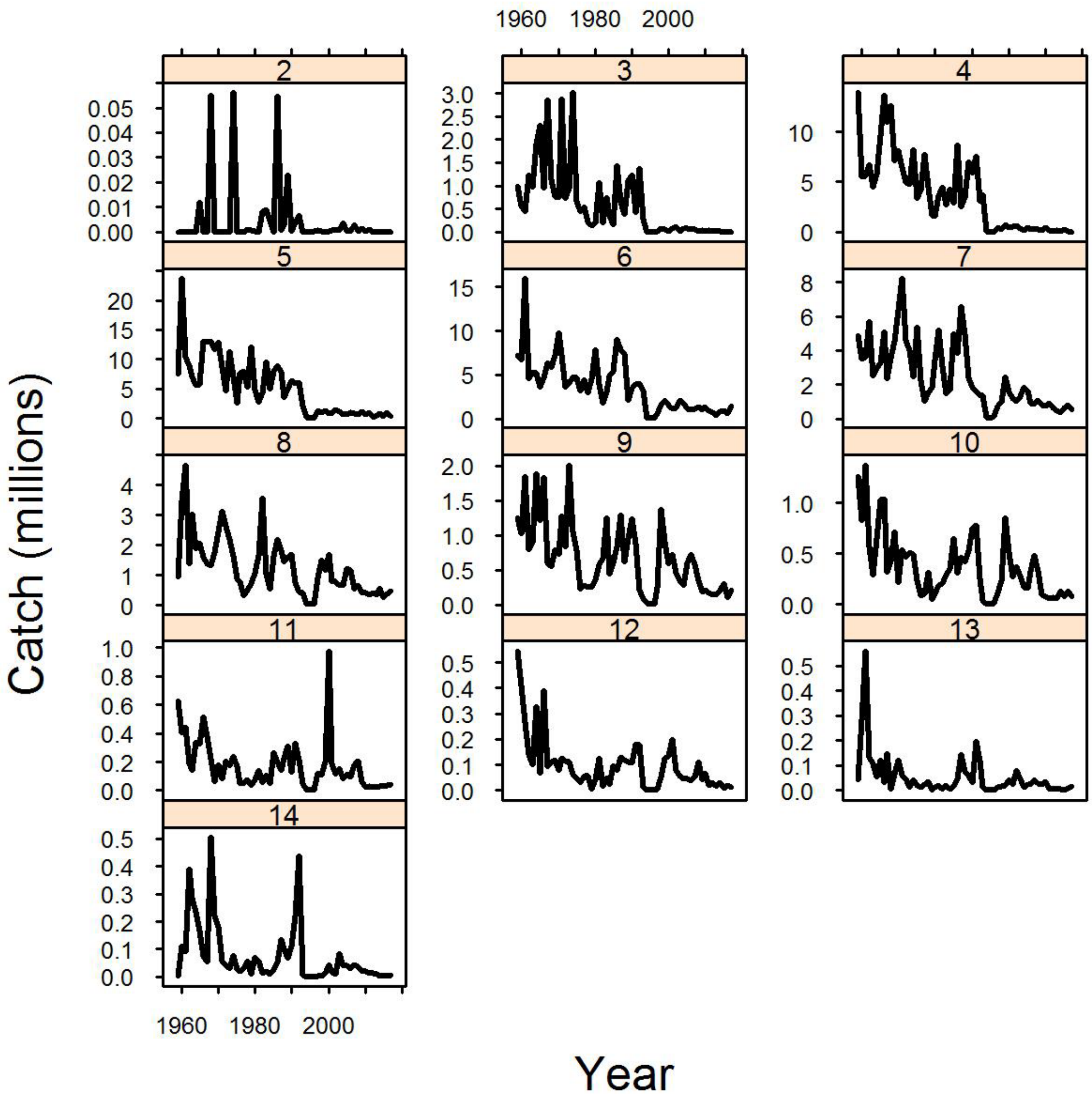

Figure D2. Time series of estimated catch numbers at ages 2–14+.

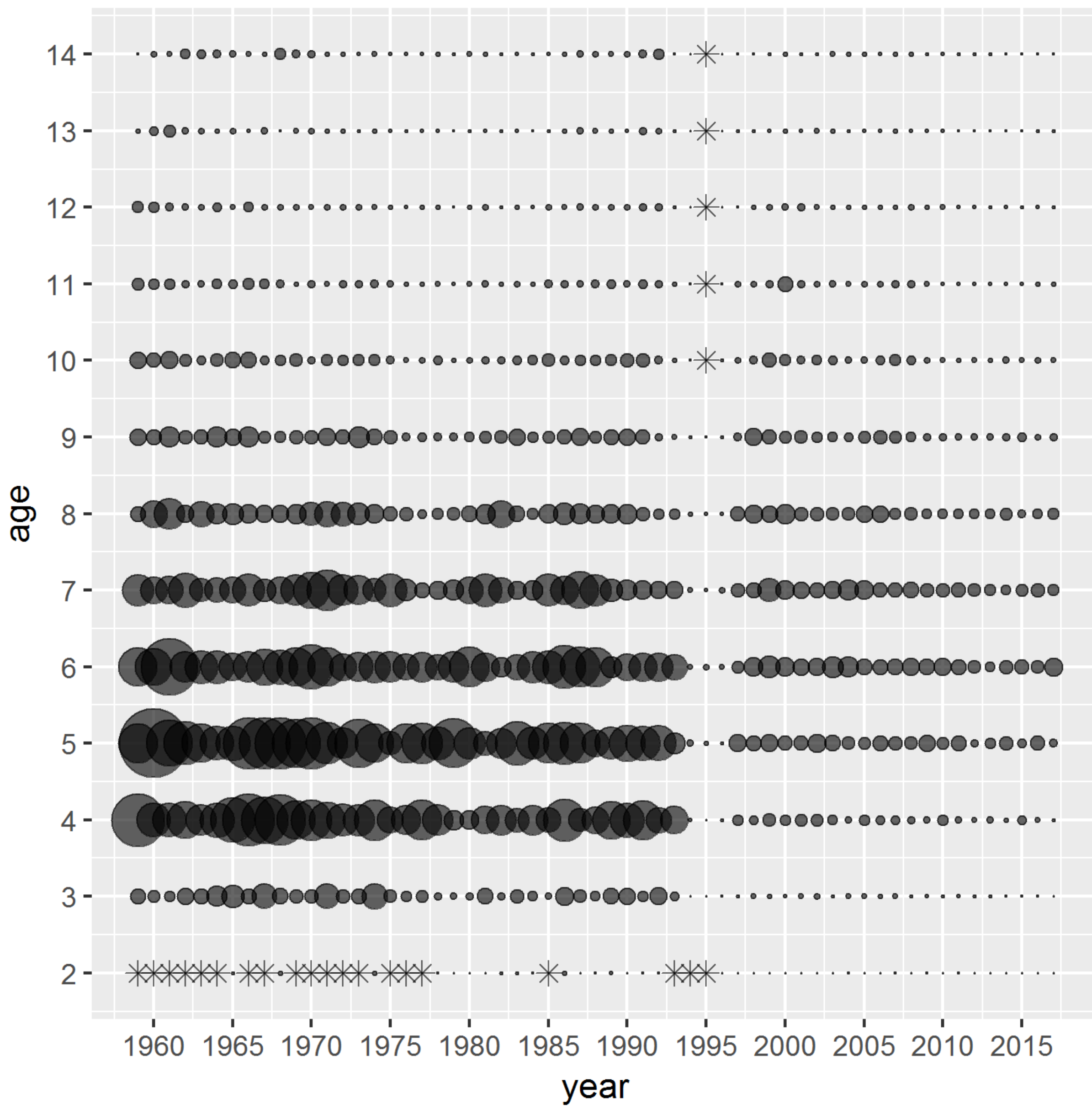

*Figure D3. Catch-at-age bubble plot. The bubble area is proportional to catch. A * indicates a zero catch.*

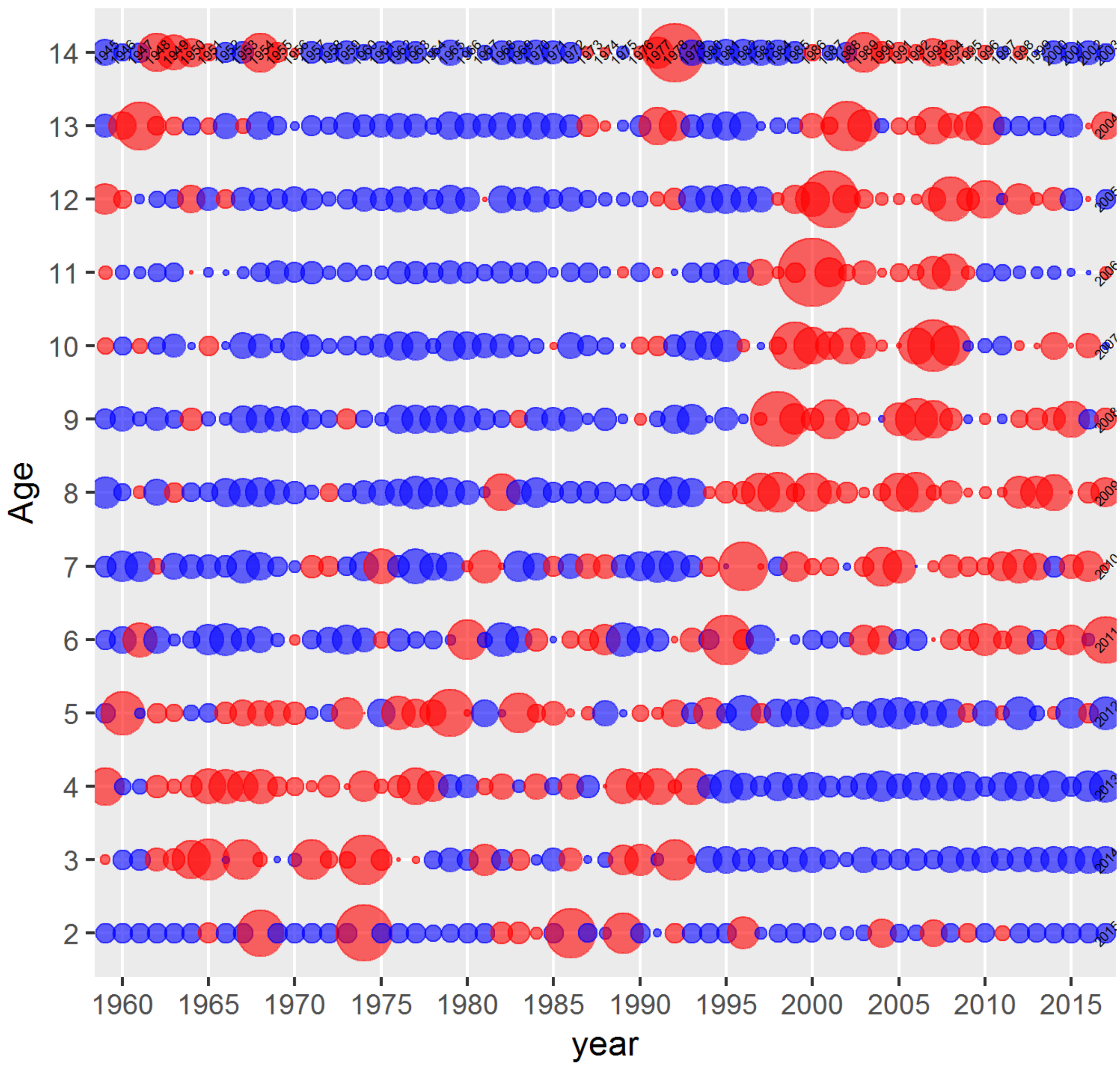

Figure D4. Catch-at-age standardized proportion at age (SPAY) bubble plot. Cohorts are indicated along the upper and right-hand margins. Red is positive and blue is negative. The bubble area is proportional to the absolute value of the standardized proportion.

Figure D5. SPAY plot since 2008. See Figure D4 for more details.

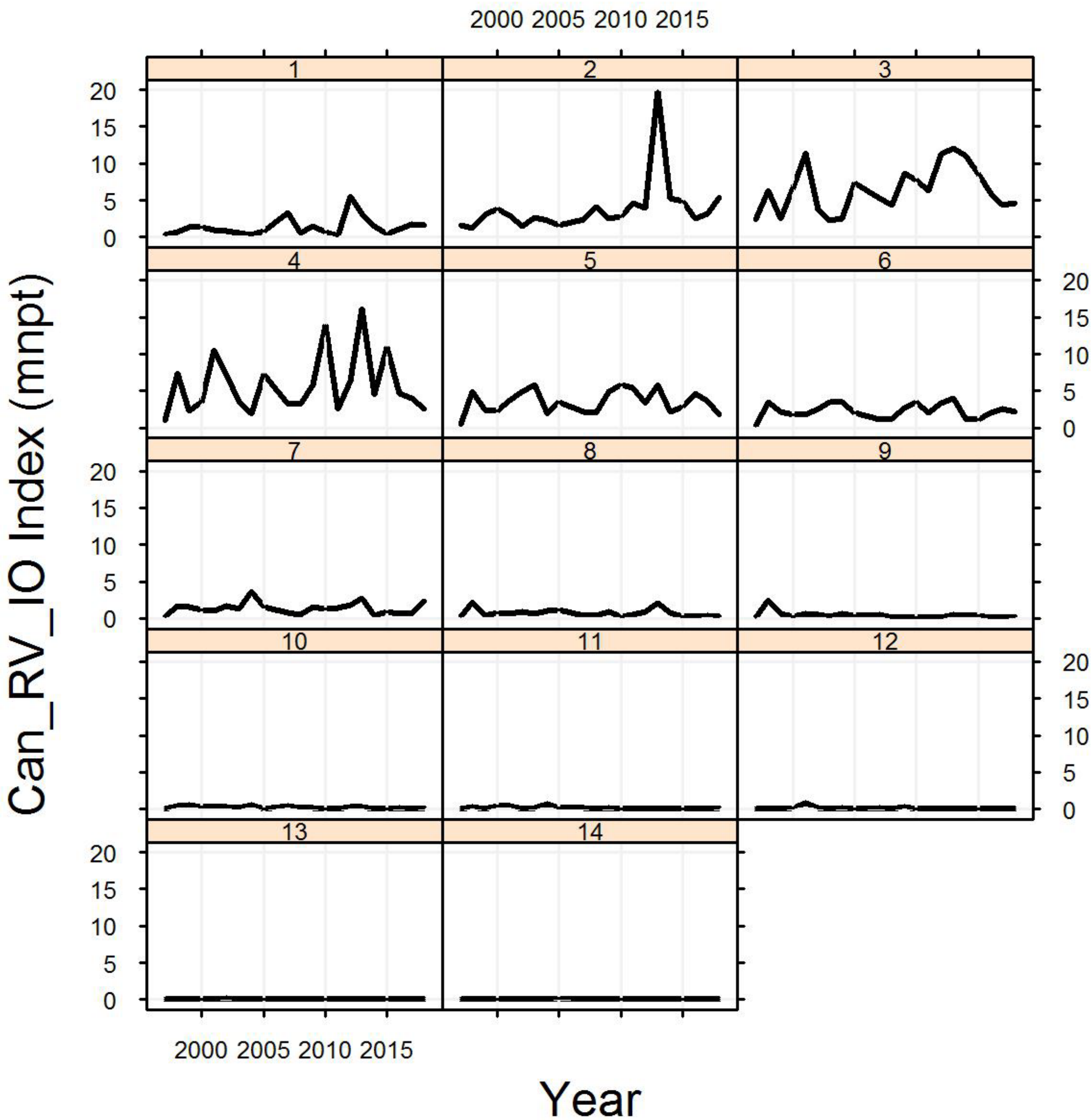

Figure D6. Time-series of the Canadian RV index (number per tow, mnpt) for inshore+offshore strata, since 1997.

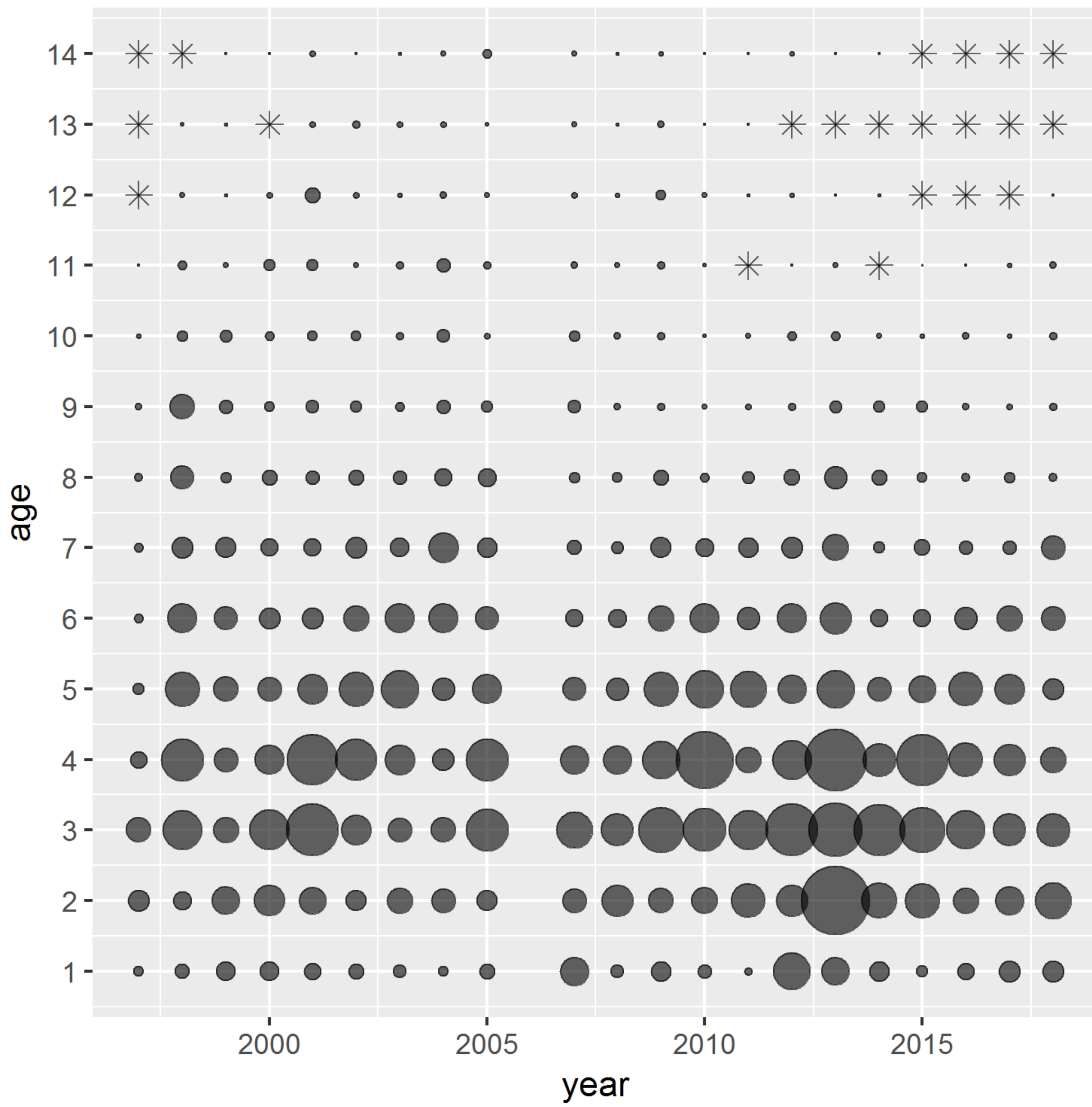

Figure D7. Bubble plot for the Canadian RV index for inshore+offshore strata. See Figure D3 for more details.

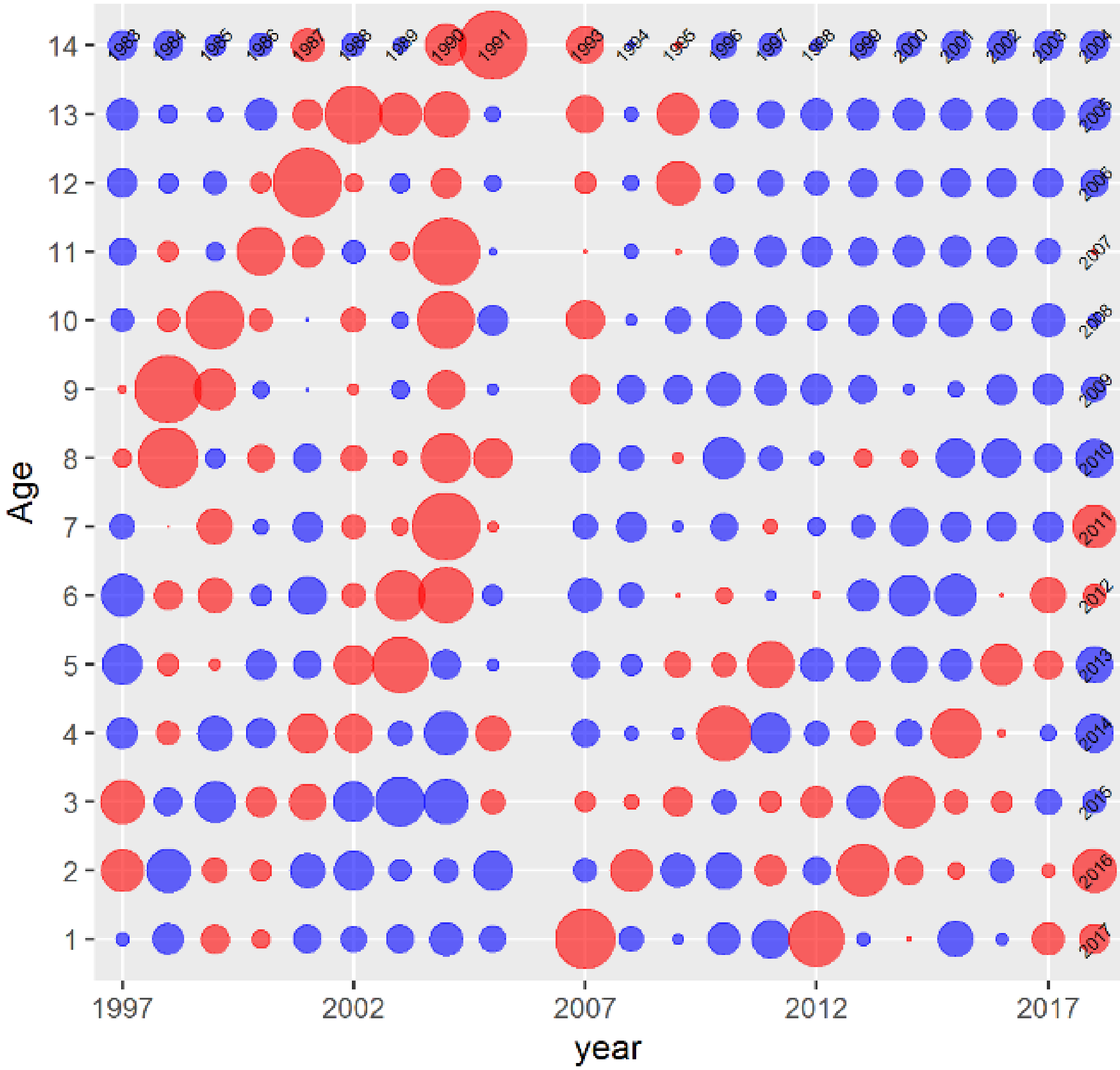

Figure D8. SPAY plot for the Canadian RV index for inshore+offshore strata. See Figure D4 for more details.

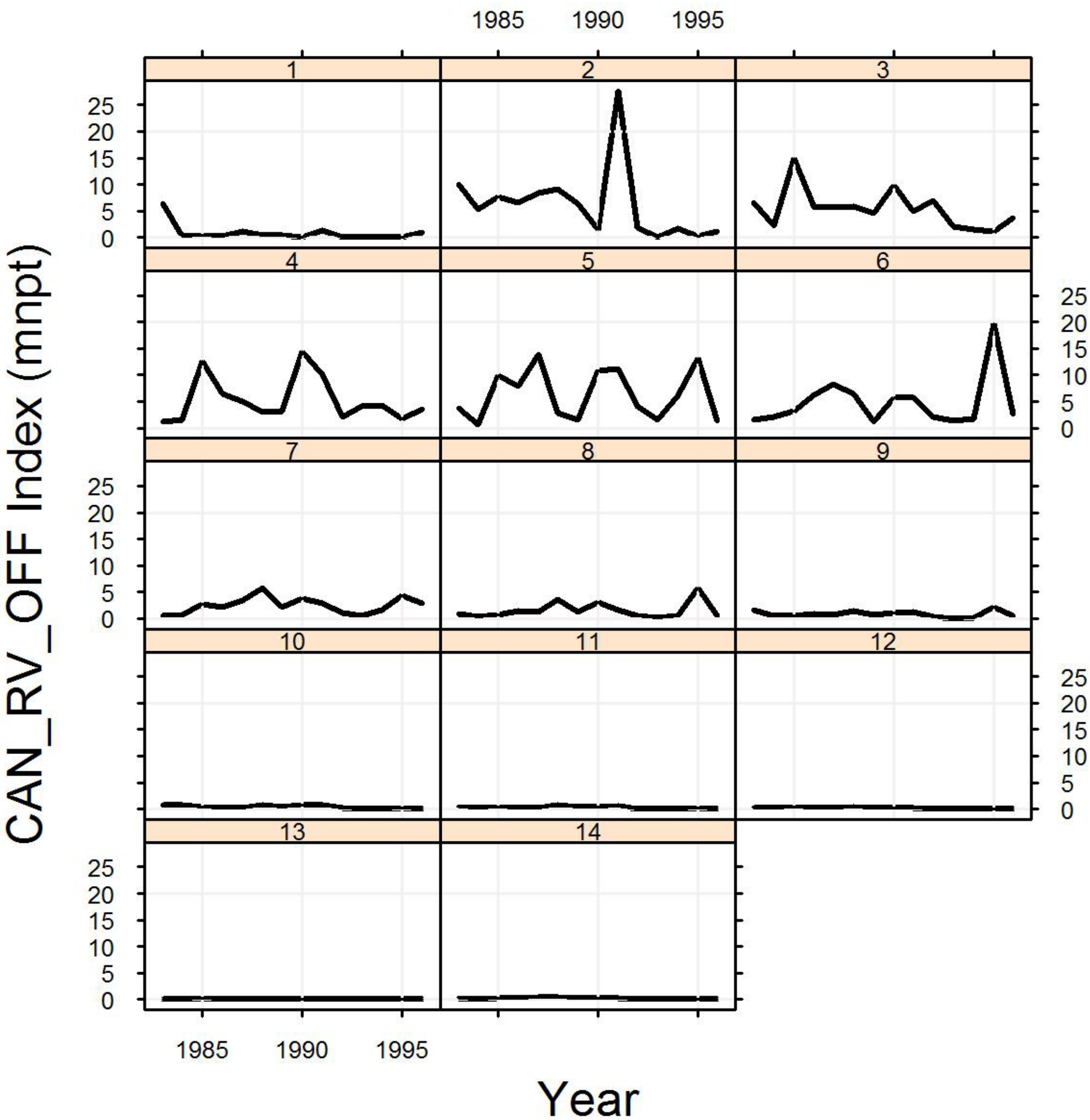

Figure D9. Time-series of the Canadian RV index (number per tow, mnpt) for offshore strata, during 1983–96.

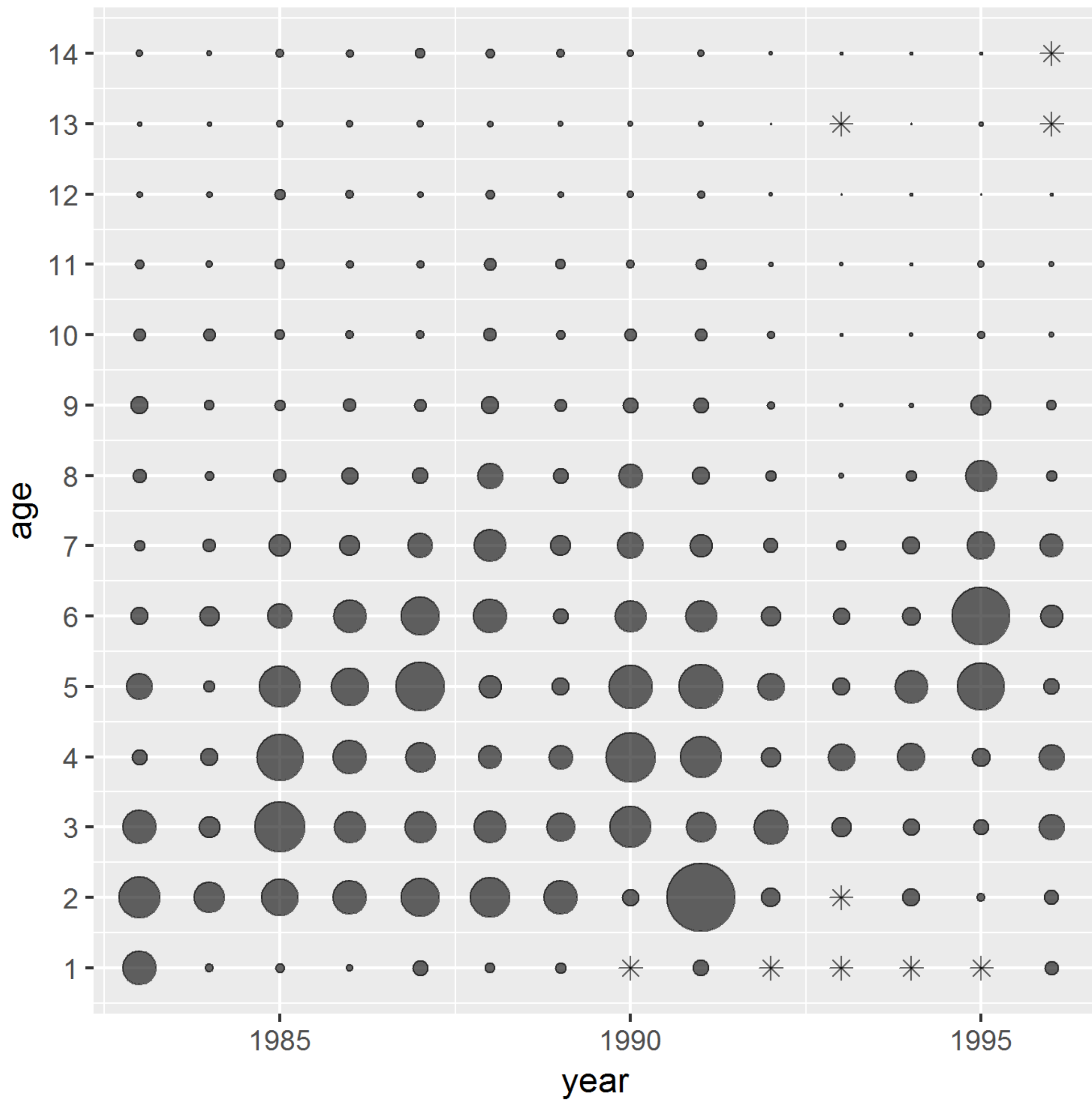

Figure D10. Bubble plot for the Canadian RV index (number per tow, mnpt) for offshore strata, during 1983–96. See Figure D3 for more details.

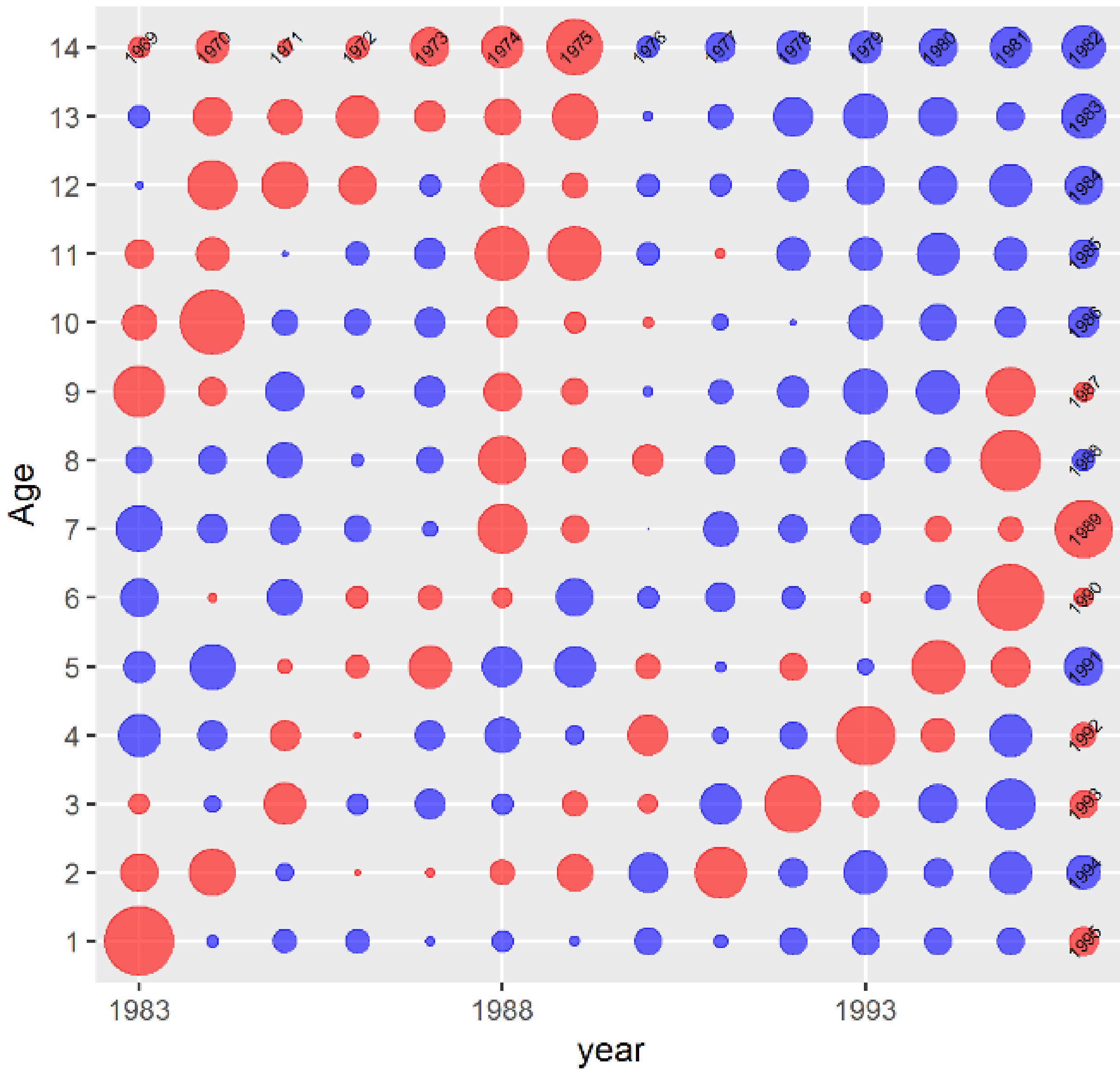

Figure D11. SPAY plot for the Canadian RV index (number per tow, mnpt) for offshore strata, during 1983–96. See Figure D4 for more details.

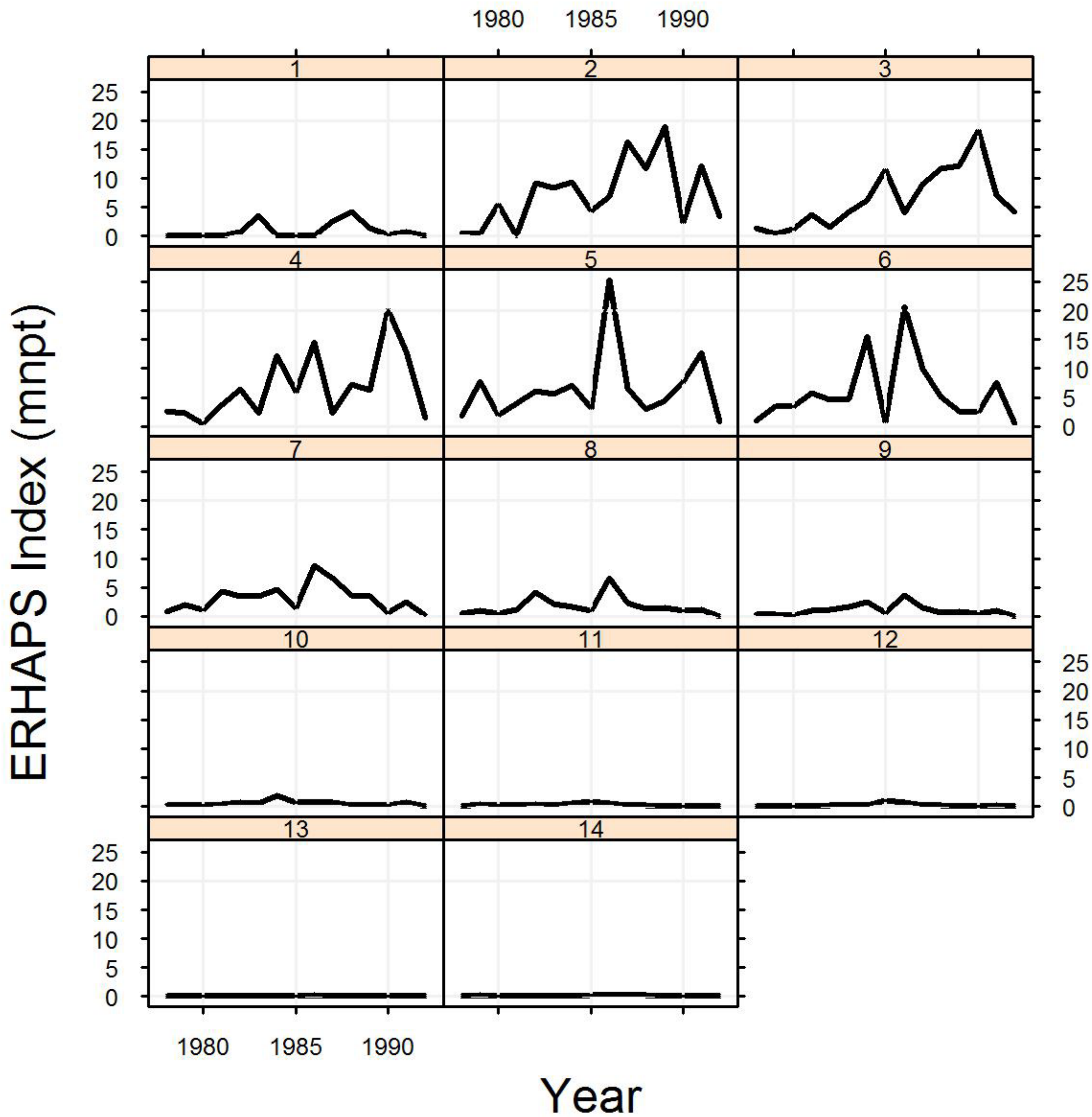

Figure D12. Time-series of the France (ERHAPS) RV survey (number per tow, mnpt) during 1978–92.

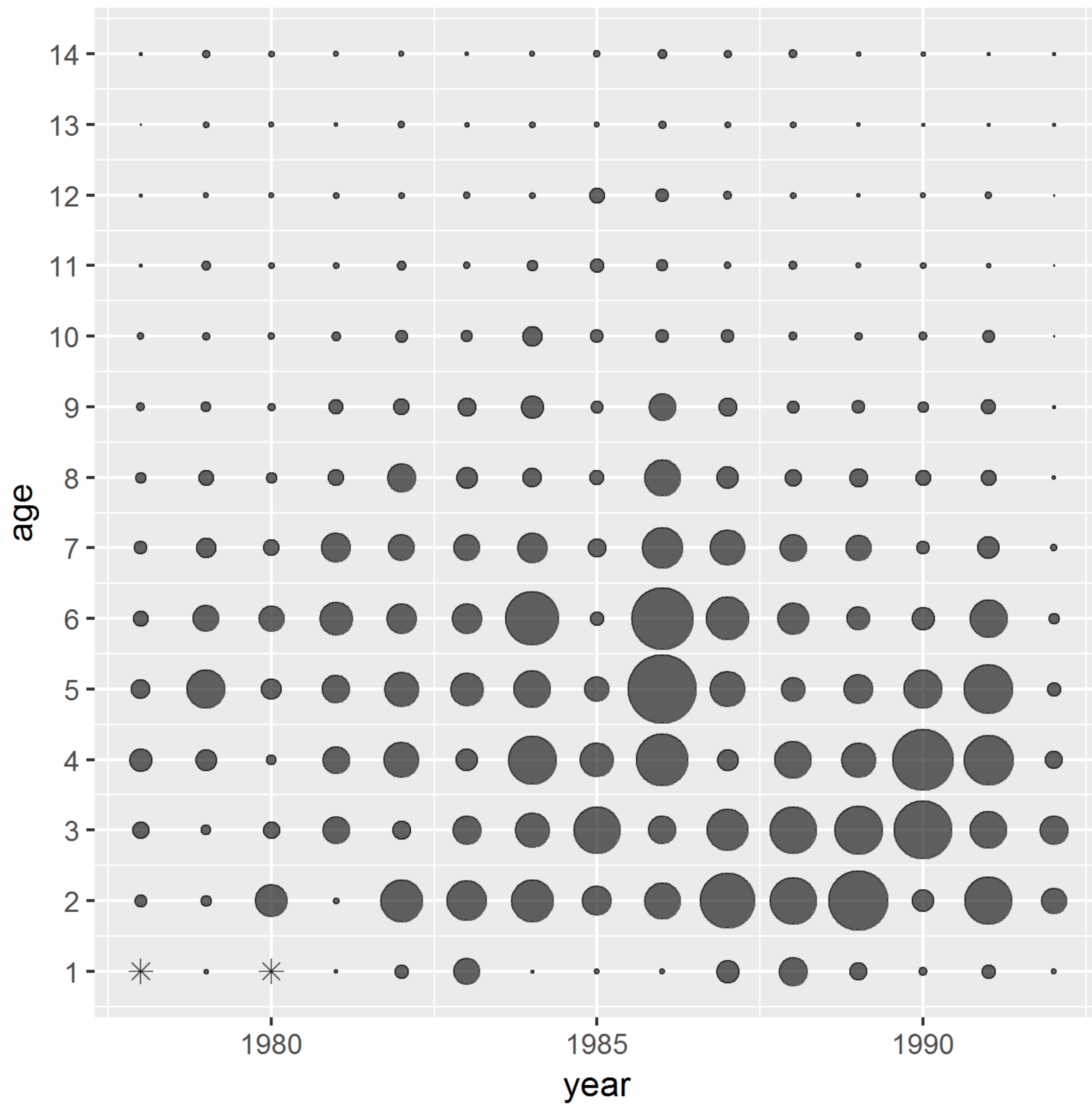

Figure D13. Bubble plot for the France (ERHAPS) RV survey (number per tow, mnpt) during 1978–92. See Figure D3 for more details.

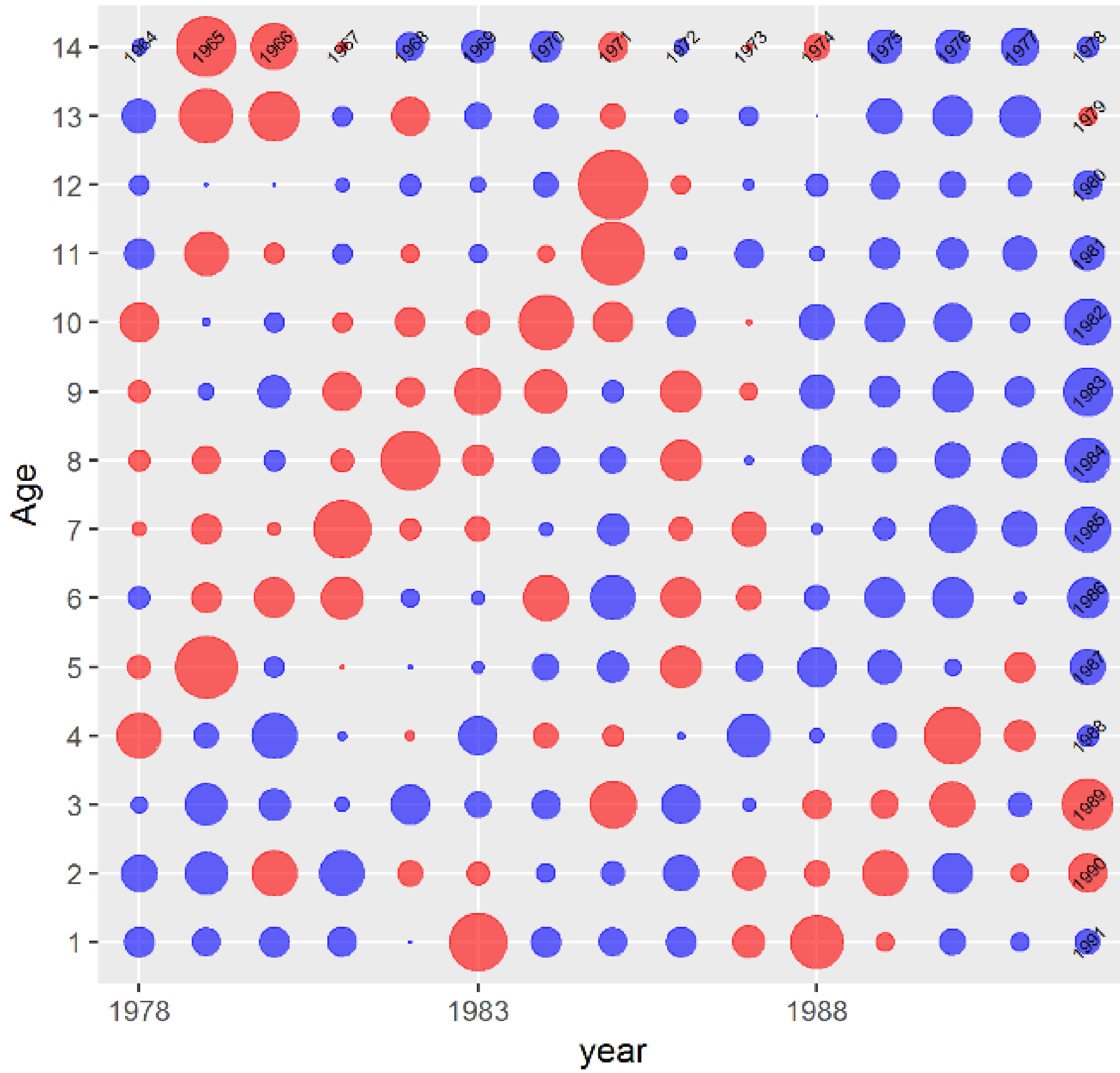

Figure D14. SPAY plot for the France (ERHAPS) RV survey (number per tow, mnpt) during 1978–92. See Figure D4 for more details.

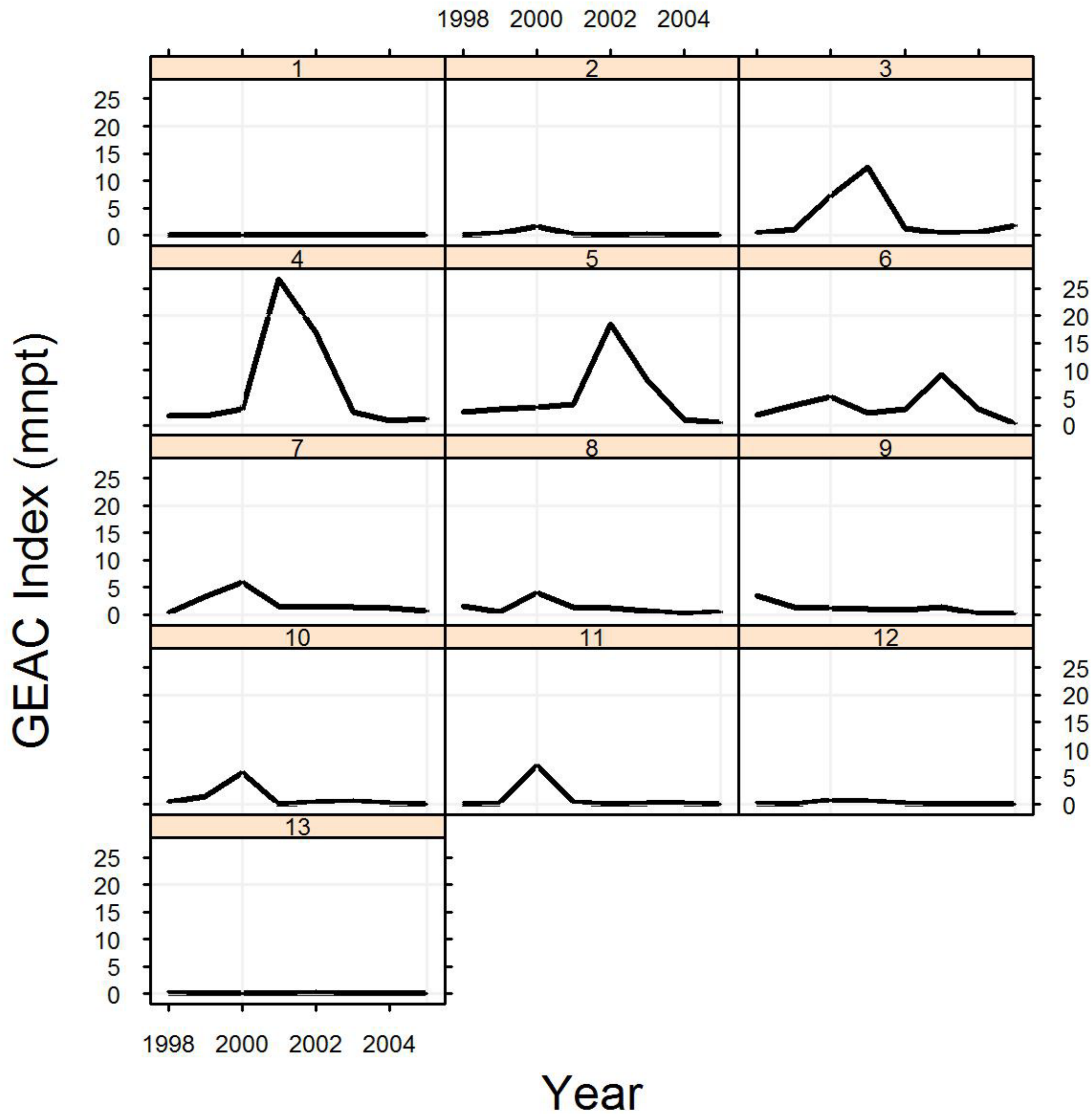

Figure D15. Time-series of the GEAC industry survey (number per tow, mnpt) during 1998–2005.

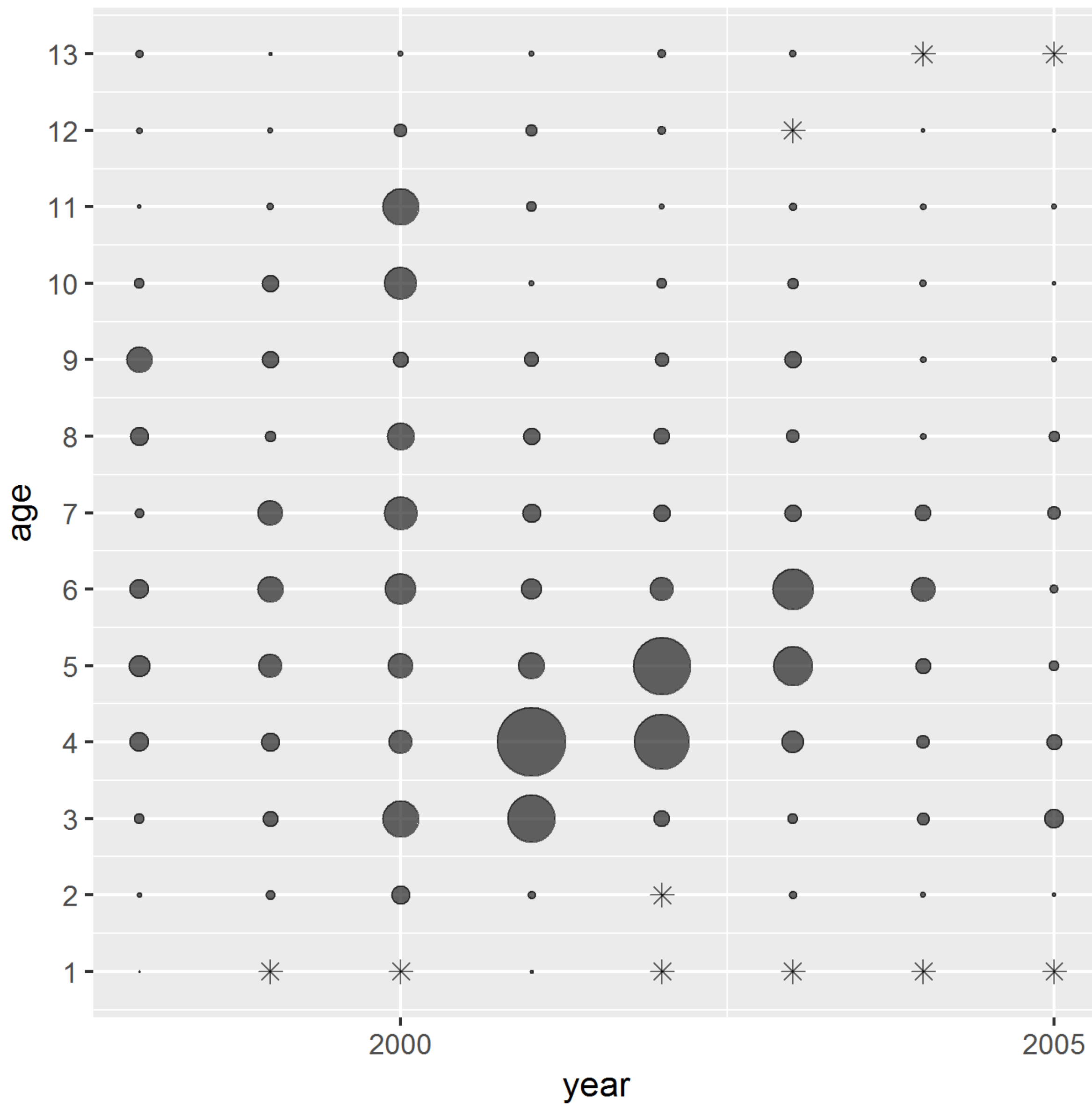

Figure D16. Bubble plot for the GEAC industry survey (number per tow, mnpt) during 1998–2005. See Figure D3 for more details.

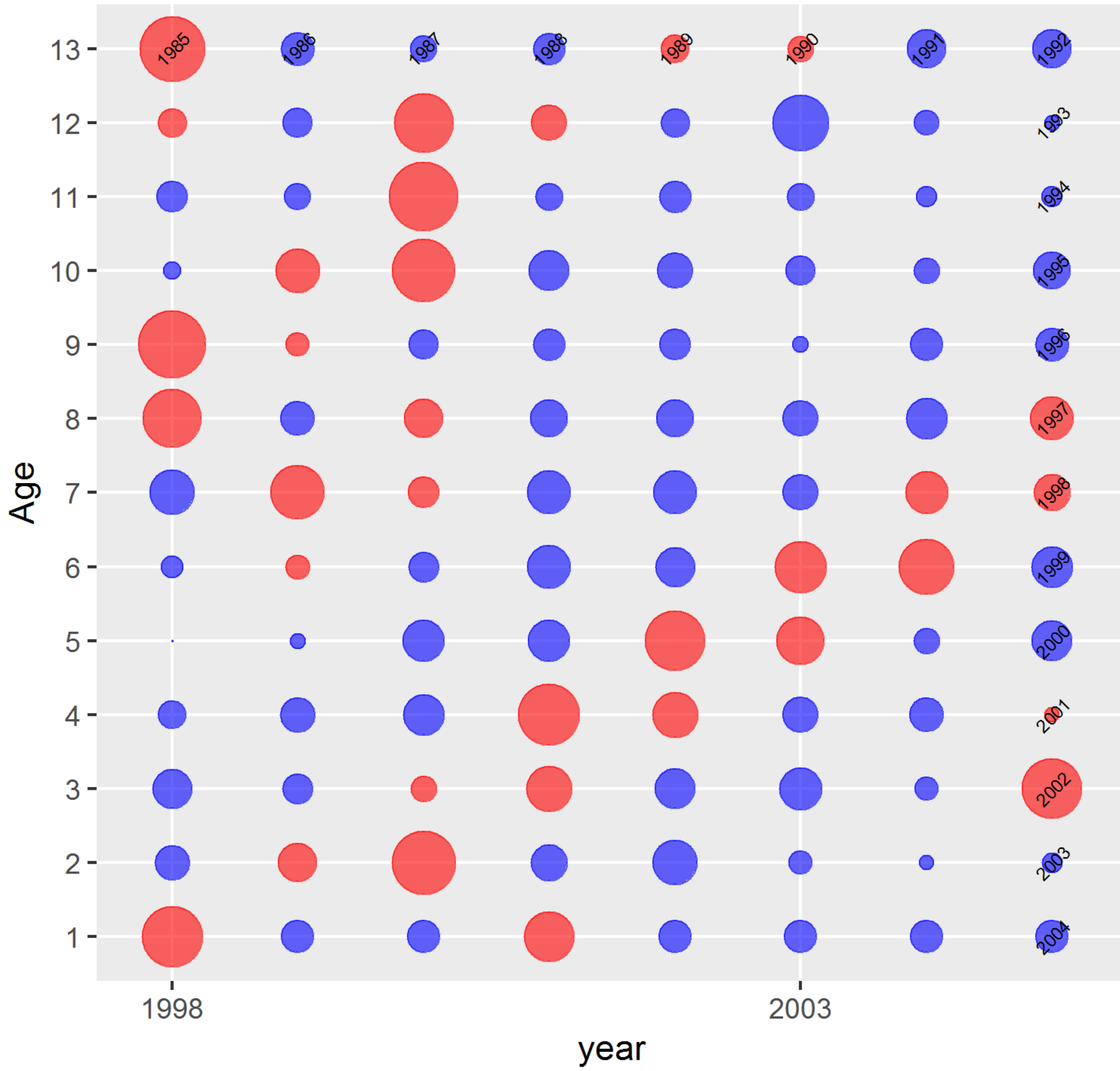

Figure D17. SPAY plot for the GEAC industry survey (number per tow, mnpt) during 1998–2005. See Figure D4 for more details.

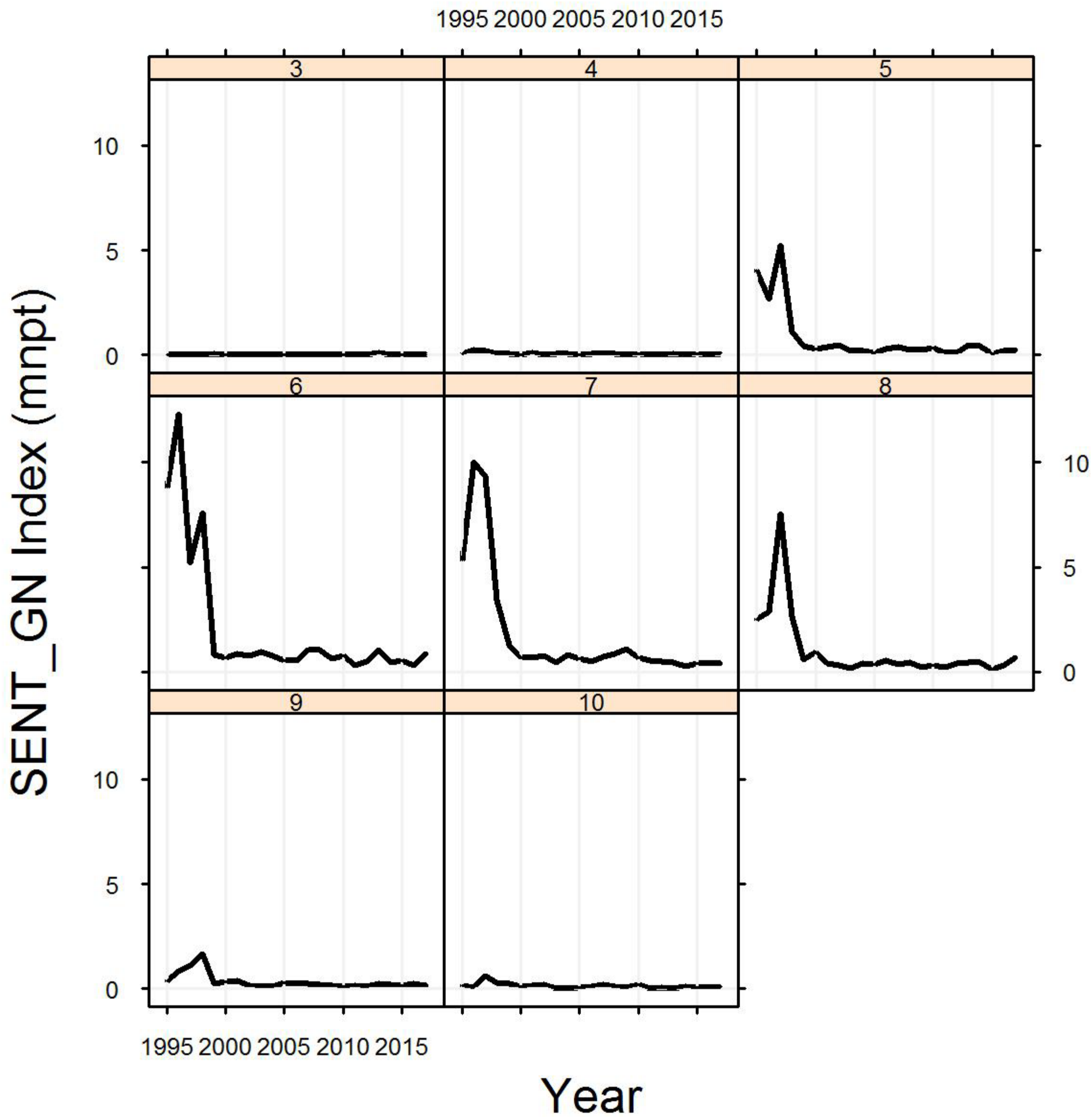

Figure D18. Time-series of the Sentinel gillnet index (number per tow, mnpt) during 1995–2017.

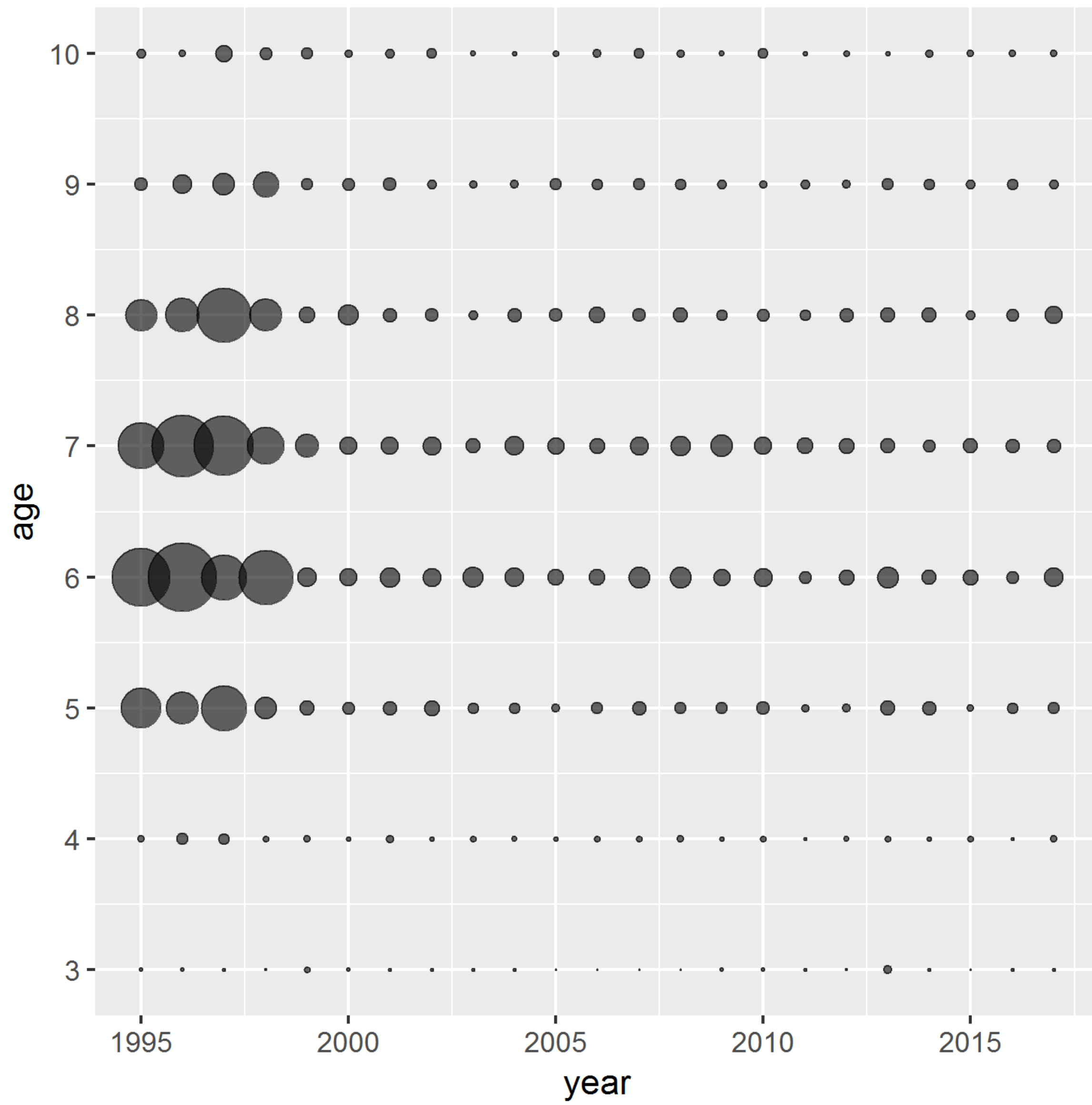

Figure D19. Bubble plot for the Sentinel gillnet index (number per tow, mnpt) during 1995–2017. See Figure D3 for more details.

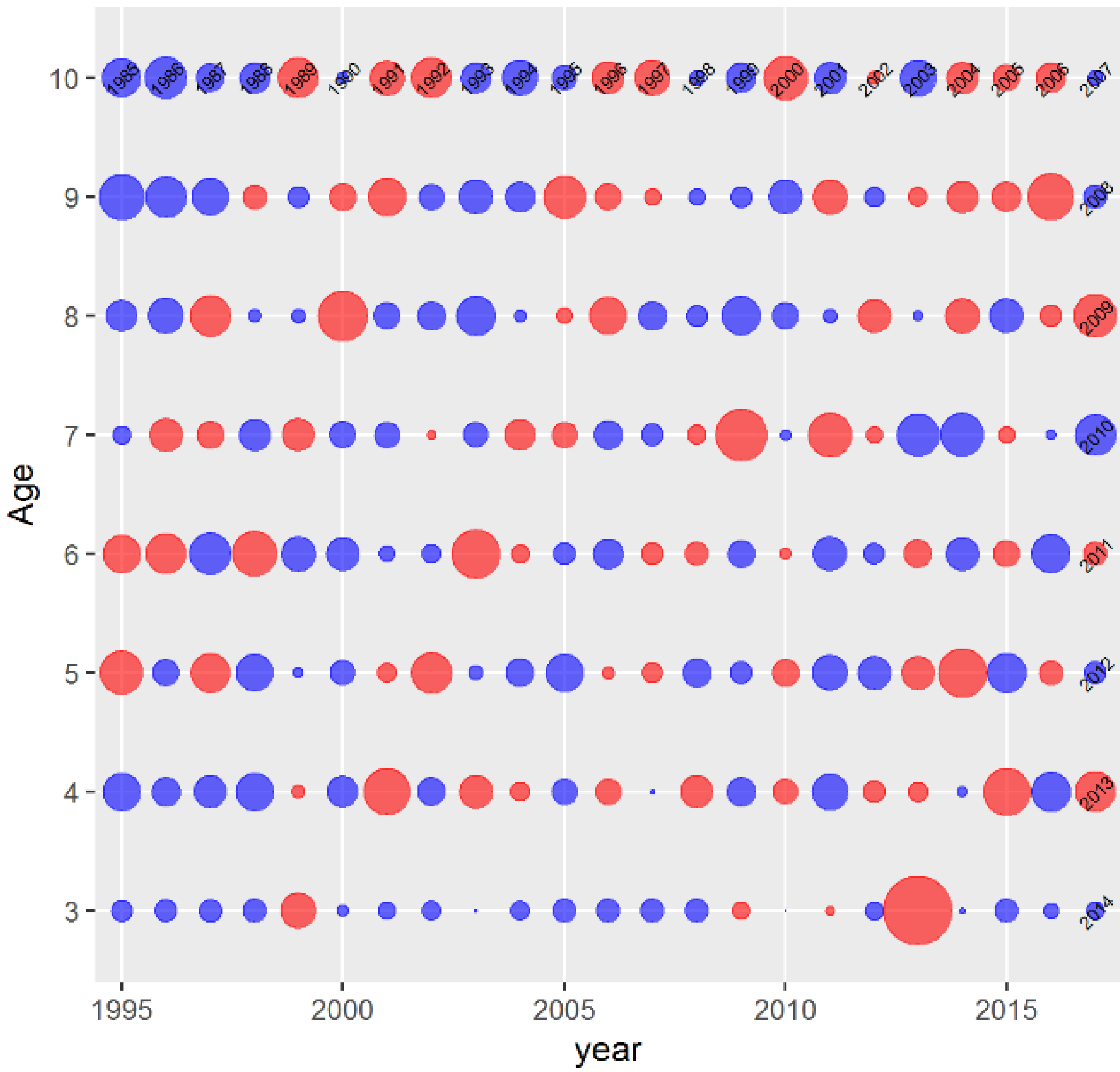

Figure D20. SPAY plot for the Sentinel gillnet index (number per tow, mnpt) during 1995–2017. See Figure D4 for more details.

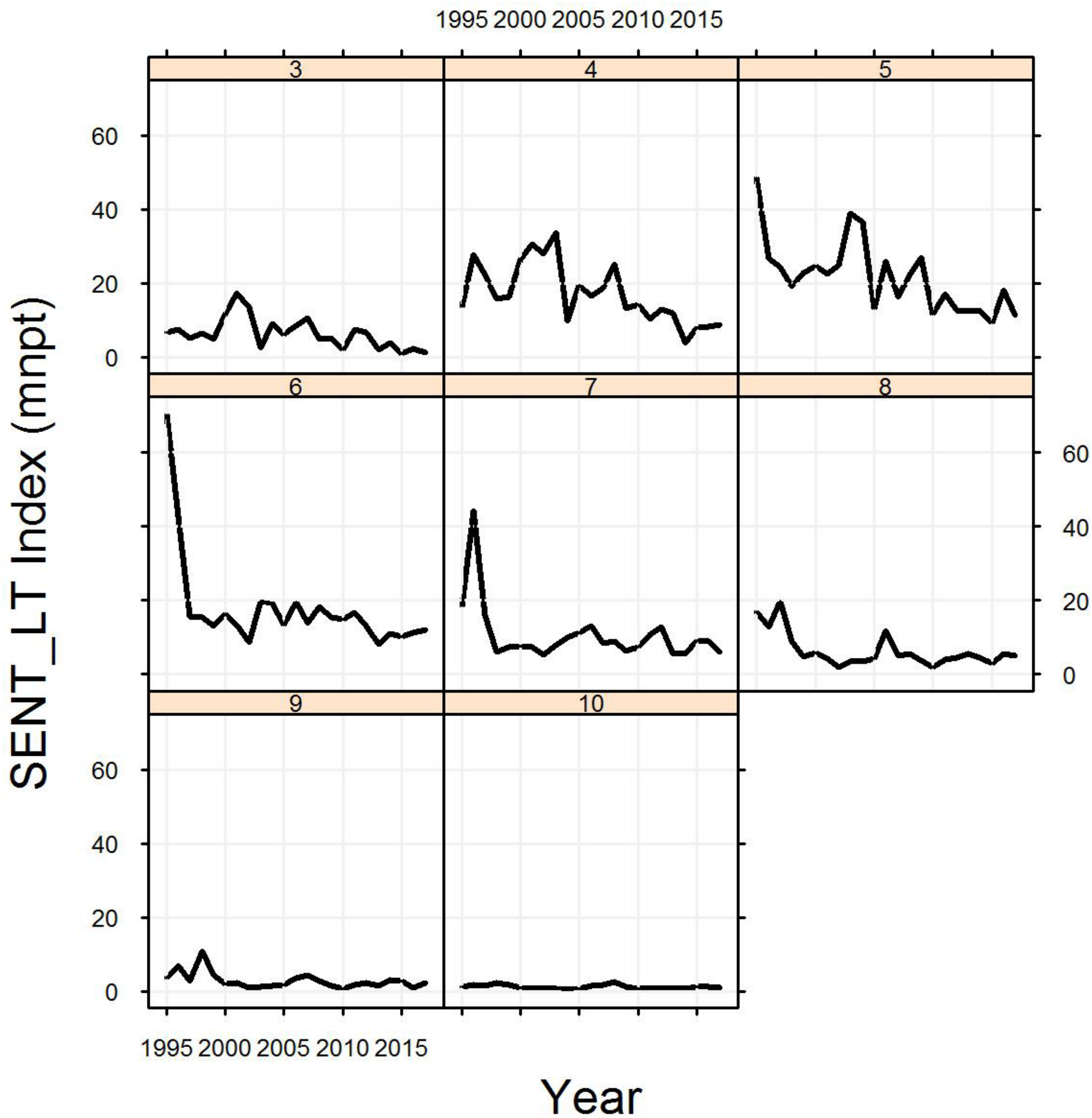

Figure D21. Time-series of the Sentinel linetrawl index (number per tow, mnpt) during 1995–2017.

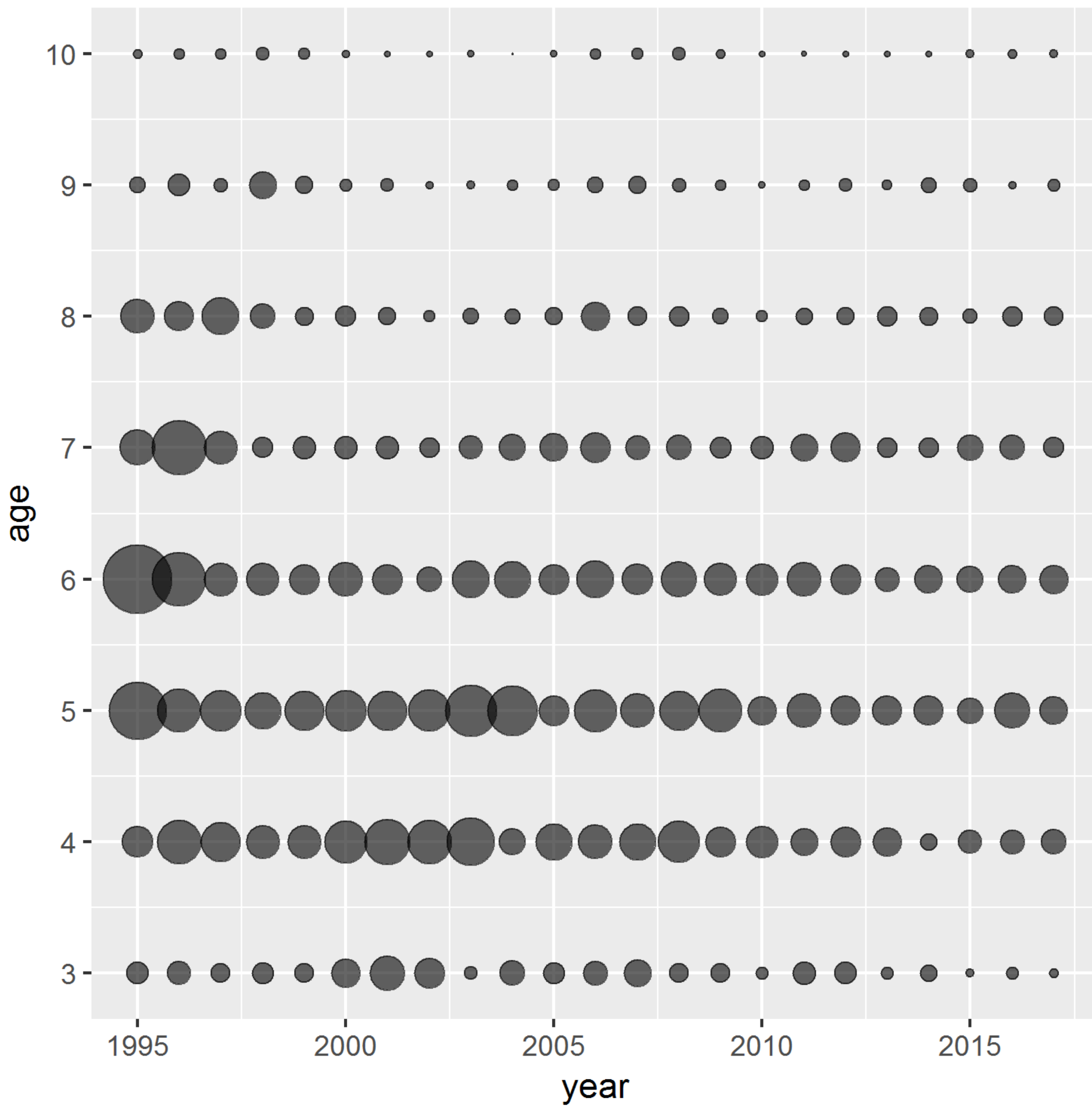

Figure D22. Bubble plot for the Sentinel linetrawl index (number per tow, mnpt) during 1995–2017. See Figure D3 for more details.

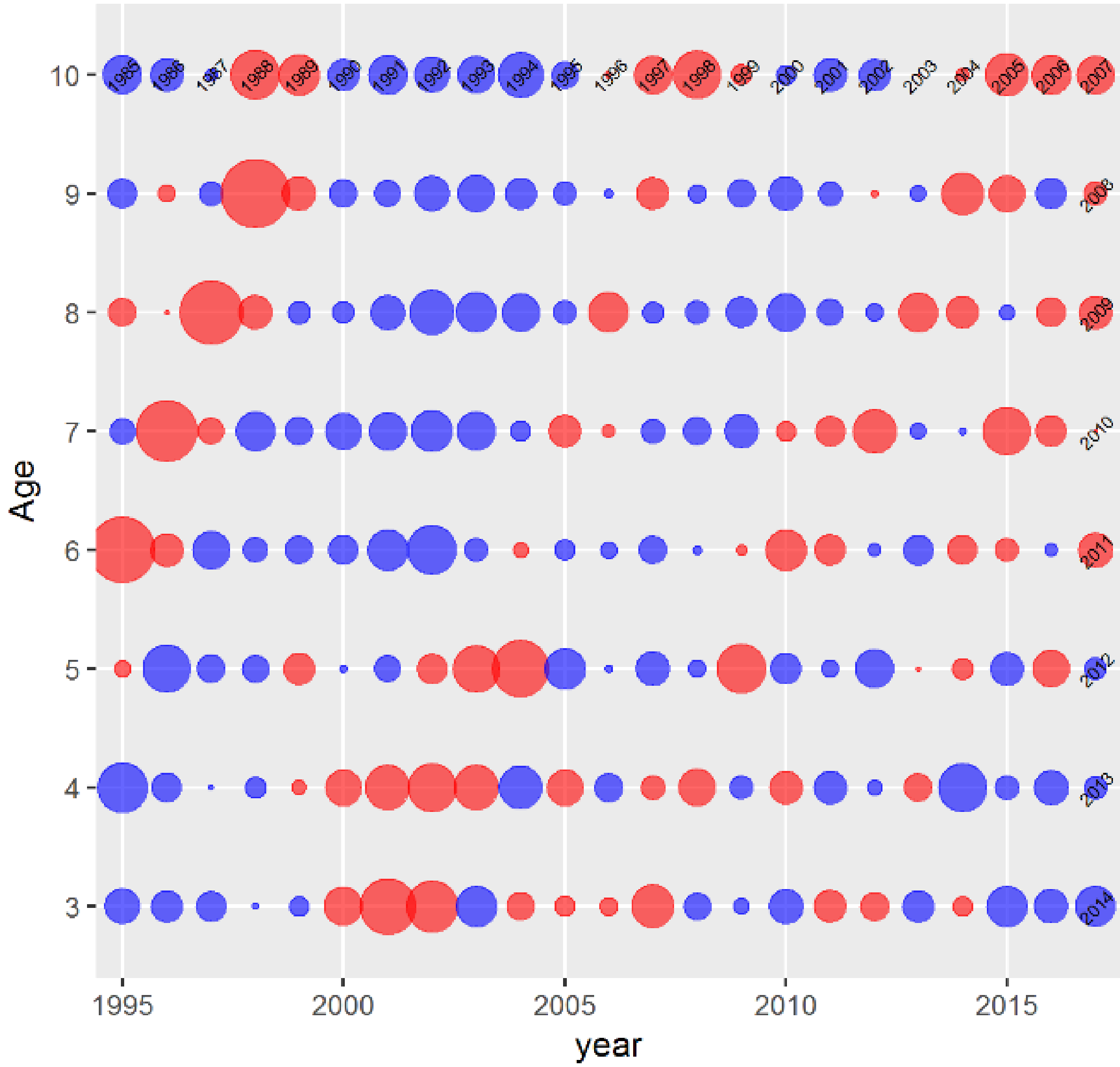

Figure D23. SPAY plot for the Sentinel linetrawl index (number per tow, mnpt) during 1995–2017. See Figure D4 for more details.

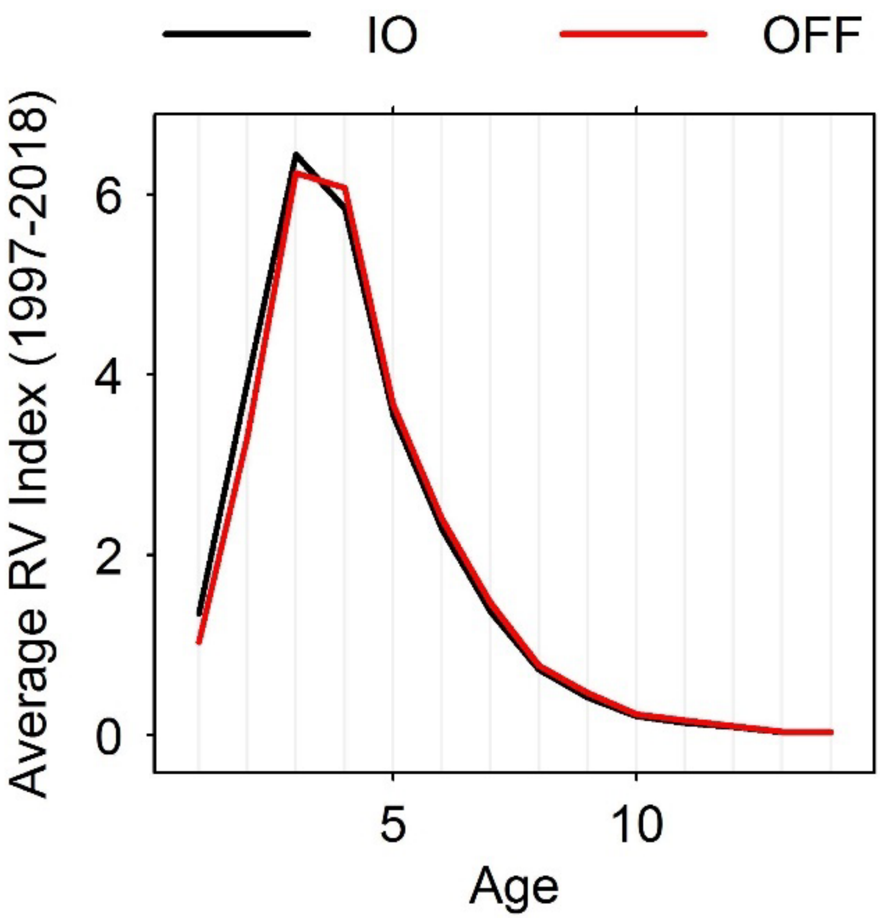

Figure D24. A comparison of the Canadian RV index, averaged over 1997–2018), for inshore+offshore strata (IO) versus the index for offshore strata only (OFF).

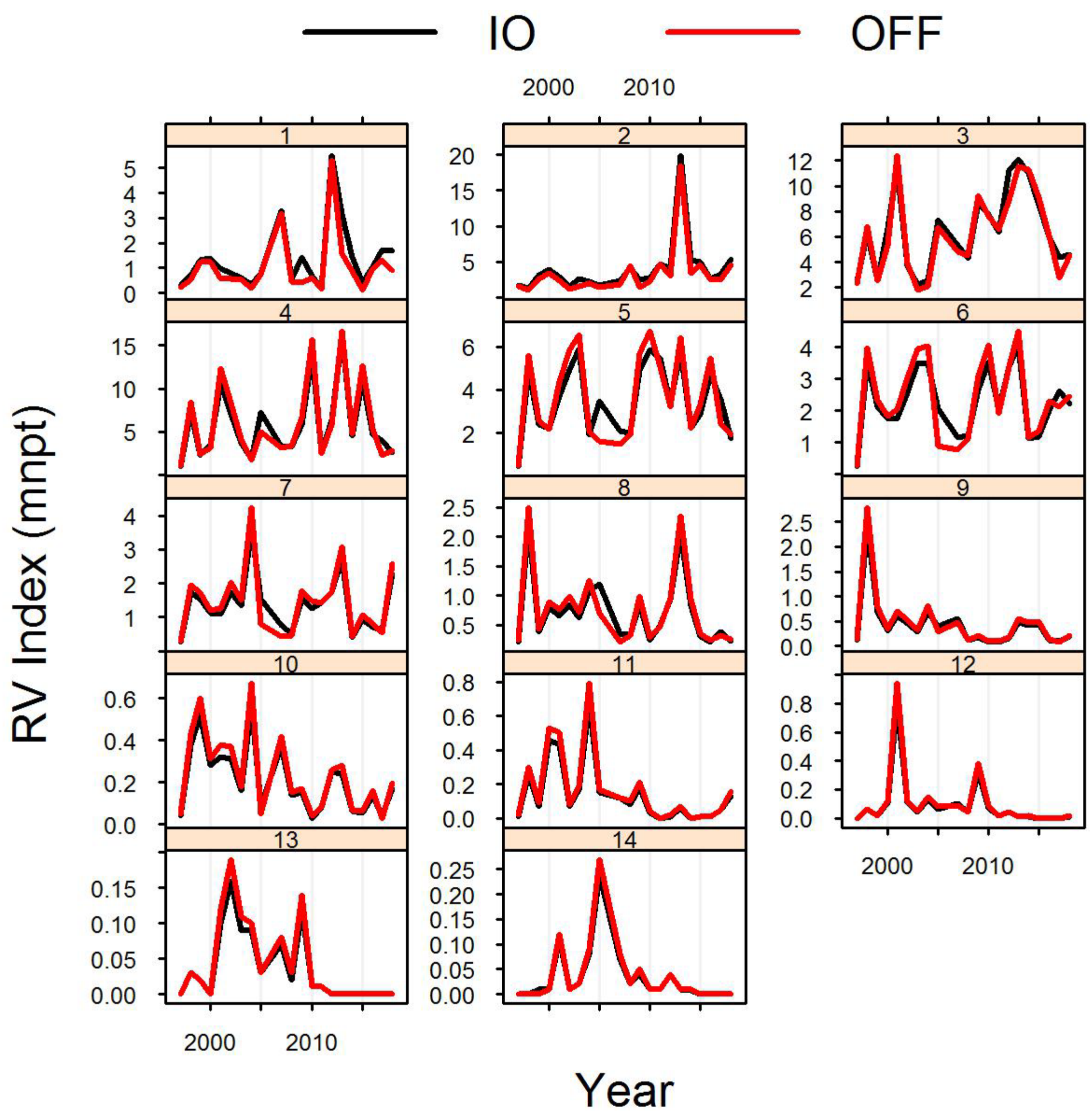

Figure D25. A comparison of the Canadian RV index for inshore+offshore strata (IO) versus the index for offshore strata only (OFF). Each panel shows results for an age.