Are we there yet? The impact of reduced data on the ability to monitor rebuilding for overfished fish stocks

Chantel R. Wetzel1,2, André E. Punt2, Jason. M. Cope1

1Fishery Resource Analysis and Monitoring Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Boulevard East, Seattle, WA 98112, United States

2School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195-5020, United States

ABSTRACT

Select rockfish stocks off the U.S. west coast are below target biomasses and are managed under rebuilding plans that limit the allowable harvest, reducing the opportunity to collect fishery-dependent data, even though this is the primary source of information on changes in abundance for many rockfish stocks. A simulation study was conducted using operating models that involved time-invariant or time-varying parameters to evaluate the impact of reduced data to estimate spawning biomass and biological parameters during rebuilding. Decreased data during rebuilding resulted in increased among-simulation variation in estimates of spawning biomass in absolute terms and relative to unfished spawning biomass, reduced average catch, and increased inter-annual variation in catch during rebuilding compared to when data collection was maintained. The presence of time-varying parameters in the operating model that were not accounted for within the estimation method resulted in increased among-simulation variability about spawning biomass and relative spawning biomass compared to the time-invariant case, with the largest increase in variability occurring during rebuilding when data were reduced or eliminated. Retaining data collections at historical levels allowed for improved parameter estimation during rebuilding which resulted in reduced the variability in estimated stock size with larger average catch during rebuilding.

*Keywords*:

Rebuilding, overfished, simulation, limited data

Corresponding author: C.R. Wetzel

Email: [Chantel.Wetzel@noaa.gov](mailto:Chantel.Wetzel@noaa.gov)

Phone: 1-206-302-1753

Fax: 1-206-860-6792

# Introduction

Rebuilding overfished stocks requires a reduction in fishing mortality to a level that allows stock biomass to increase (in the absence of a run of above average recruitment allowing for stock growth despite overfishing). In the United States, federally-managed stocks that fall below a minimum stock size threshold (MSST) are declared overfished and are mandated to be rebuilt to target biomass levels in the shortest amount of time, accounting for biological and environmental conditions (SFA, 1996). This can lead to substantial reductions in fishing effort relative to historical levels. The severity of restrictions during rebuilding can, for some stocks, lead to a situation where the ability to collect data becomes limited over the period when managers are likely most concerned about stock size (i.e. when the stock is under the rebuilding plan).

Data are necessary to determine the extent to which a stock is on track to rebuild. The ability to measure the rate of recovery is crucial to management, and increased uncertainty due to limited data can impede monitoring to determine if a stock is on target to rebuild in a specified timeframe. Additionally, biological data are critical to improve estimates of key parameters within stock assessments (e.g. natural mortality, growth, recruitment compensation termed steepness) and can indicate incoming poor or strong year-classes (recruitment), which can impact estimates of relative stock biomass (the ratio of current biomass to unfished biomass) and rebuilding rates. Potential improvements in parameter estimates, and the ability to detect incoming fluctuations in recruitment during rebuilding could be restricted when collection of new biological data is limited due to harvest restrictions.

Overfished rockfish species off the U.S. west coast have experienced large reductions in harvest during rebuilding. One example is yelloweye rockfish (*Sebastes ruberrimus*), which was declared overfished in 2002 (Methot and Piner, 2002). Similar to other rockfish species off the U.S. west coast, catches of yelloweye rockfish were large and unsustainable during the 1980s and early 1990s. Catches of yelloweye rockfish were reduced dramatically relative to historical catches following the overfished declaration, where the allowable catch during the first year of rebuilding was reduced to approximately 10% of the catch four years earlier (Stewart *et al*., 2009).

The reduction of fishery catch, and resulting fishery data during rebuilding, presents a challenge for assessment and management. Many species of rockfish, such as yelloweye rockfish, are not reliably sampled by the main fishery-independent survey off the U.S. west coast, either due to the inability of the fishery-independent survey to sample rocky habitat using trawl gear or other restrictions on sampling locations (e.g. rockfish conservation areas or near-shore habitat). Because these species are not well sampled, the majority of historical information (e.g. index of abundance, length, and age data) available for assessment derives primarily from recreational and commercial fishery samples. Yet, because of retention restrictions triggered by the rebuilding plan, recreational and commercial fishery behavior is profoundly altered (Stewart *et al.*, 2009). The most recent yelloweye rockfish assessment cited limited fishery data during rebuilding as a challenge to ‘produce conclusive information about the stock for the foreseeable future’ (Stewart *et al.*, 2009).

Understanding the long-term impact of reduced data on the ability to monitor a stock during rebuilding would provide insight and guidance for management. There have been numerous simulation studies evaluating the impact of data quality and quantity on the performance of stock assessment methods (e.g. Hilborn, 1979; Chen *et al.*, 2003; Yin and Sampson, 2004; Magnusson and Hilborn, 2007; Wetzel and Punt, 2011; Lee *et al.*, 2012). However, studies often focus on the ability to estimate either management quantities or biological parameters. The simulation performed here evaluates the ability to accurately monitor rebuilding of an overfished long-lived rockfish stock, such as those that are common to the U.S. west coast, where harvest and the collection of fishery data are restricted during rebuilding. The simulation study addresses three main questions; 1) does limited data result in increased uncertainty impacting the ability to detect when an overfished stock has recovered to the management target stock size (i.e. it is rebuilt), 2) are limited data from the fishery able to detect a shift in fishery selectivity resulting from changing fishing behavior during rebuilding, and 3) how are model estimates of stock size and biological parameters affected during periods of limited data?

# Material and Methods

## General approach

A rockfish life history type common to the U.S. west coast was simulated (Table 1), based on the life history for yelloweye rockfish. Yelloweye rockfish exhibit very low natural mortality and recruitment compensation (steepness) (even relative to other U.S. west coast rockfish species) and, therefore, are assumed to have slow population dynamics. The operating model was parameterized using higher natural mortality and steepness values to be more similar to other U.S. west coast rockfish species, and to allow for shorter recovery periods (< 100 years) for computational efficiency while still maintaining the characteristics of a rockfish life history.

Two alternative cases were simulated using the operating model to account for the potential impacts of time-varying natural mortality and fishery selectivity. The first case, referred to as “time-invariant”, involved a single fixed natural mortality rate over the entire time period. The fishery selectivity was assumed (and fixed) to be asymptotic during historical period, dome-shaped during the overfished period, and then again asymptotic after the stock was rebuilt (Fig. 1). The simulated stocks were reduced to an overfished state (below MSST) at the time of the first assessment in year 50.

The second case, referred to as “time-varying”, involved annual deviations in natural mortality and in the parameters on which the fishery selectivity pattern was based during the historical, overfished, and rebuilt periods (Fig. 1C and 1D). Annual deviations in fishery selectivity were applied to two selectivity parameters: 1) the length (in cm) at which the ascending limb of selectivity curve reached maximum selectivity (termed ‘size at maximum selectivity’, Fig. 1C), and 2) the width of the plateau for maximum selectivity (defined as a logistic function between peak and the maximum length) resulting in dome-shaped selectivity curve (termed ‘width at maximum selectivity’, Fig. 1D) during the years the stock was overfished. A standard error of 0.05 was applied annually about the size at maximum selectivity parameter for all years and a standard error of 0.20 was applied for the width at maximum selectivity parameter during the years the stock was estimated to be overfished. The level of variation about each parameter was selected to ensure that the ascending limb of the selectivity curve was greater than the 50% length at maturity (37cm) within the operating model, and the width of maximum selectivity (creating dome-shaped curve) was small enough to allow potential detection by the estimation method (a detectable portion of the population with reduced selectivity due to dome-shaped curve). Annual deviations in natural mortality were autocorrelated.

The operating model was based on a single-sex age-structured model. An annual fishery catch-per-unit effort (CPUE) index was observed with error, length- and age-composition data were collected for selected years, and used by the estimation method to estimate population size and a catch level. The catches were then removed without error from the simulated stock. Data generation, catch estimation and stock updating was conducted in an iterative fashion for 100 years (termed the management period), a length of time that would allow for recovery (stock growth to at or greater than the target biomass) to occur.

## The operating model

The numbers-at-age at the start of the year are computed using the equation:

 (1)

where *Nt,a*  is the number of fish of age *a* at the start of the year *t*, *Rt* is the number of age-0 fish at the start of year *t*, *St,a* is the selectivity during year *t* for fish of age *a*, *A* is the plus group, *F*t is the instantaneous fishing mortality rate during year *t*, and *Mt* is the instantaneous rate of natural mortality during year *t*.

Natural mortality for year *t* is defined as:

 (4)

where *M* is the mean value of natural mortality,  is the standard error of the annual deviations in natural mortality, and  is the autocorrelated lognormal deviation in natural mortality for year *t*:

 (5)

where  is the level of autocorrelation associated with natural mortality and  is the deviation in natural mortality for year *t*. The time-invariant natural mortality case assumed  and hence .

The number of age-0 fish is related to spawning biomass according to the Beverton-Holt stock recruitment relationship:

 (2)

where *SB0* is the unfished spawning biomass, *SBt* is the spawning biomass at the start of the spawning season in year *t*,  is the standard deviation of recruitment in log space, and *h* is steepness.

A non-equilibrium starting condition was created by applying equations (1) and (2) for the number of years equal to the maximum age prior to the start of fishing, with variation in recruitment. Historical catches for years 1-50 were generated so that the populations were at 0.15*SB*0 in year 50. This ratio of spawning biomass to unfished spawning biomass (relative spawning biomass) was selected to allow for correct detection by the estimation method that the stocks were in an overfished state, and that it would require an extended number of years for the stock to rebuild to the target biomass (where the loss of data could impact the long-term performance of the estimation method). The catch of fish of age *a* during year *t* in numbers is given by:

 . (3)

The observation model was used to generate a fishery CPUE index for each year *t*:



where *Q*is the catchability coefficient,  is the standard deviation of catchability in log space, and the  is the selected biomass in the middle of year *t*:

 (4)

where *wa* is the weight of a fish of age *a*. The length- and age-composition data for the fishery were assumed to be multinomially distributed (see *Data Scenarios* for details). Ageing error was assumed to be normally distributed with ages subject to a 5% standard deviation by age.

The fishery selectivity during the historical period (years 1-50) were assumed to be asymptotic (Fig. 1A and 1C). Fishery selectivity shifted to a dome-shaped (compared to the historical asymptotic) form (Fig. 1B and 1D) within the operating model during the period that the stock was estimated to be below the target biomass (0.40*SB*0). Once the population was estimated to have recovered to above the target biomass, fishery selectivity reverted to the asymptotic form. The shift in selectivity was designed as a way to mimic a change in fisher behavior resulting from an overfished declaration. The change in shape of the selectivity curve depended on the estimated stock status rather than the true operating model status, i.e. changes in fisher behavior modeled by a change in selectivity were assumed to be driven by management restrictions based on the estimation method’s perception of the stock rather than the true unobservable state of the simulated stock.

## The estimation method

Stock synthesis (SS), an integrated statistical catch-at-age model (Methot and Wetzel, 2013), was the estimation method used to assess the simulated stocks. SS was applied for the first time in year 50 and then every 6th year thereafter. Assessment frequency for U.S. west coast groundfish varies as a consequence of commercial importance (an indicator of exploitation), the time since last assessment, and dynamics of the stock (Methot, 2015). Long-lived rockfish species generally have slow dynamics, resulting in minimal fluctuations in biomass from year to year (assuming non-extreme harvesting). Assessing the stock every 6th year was selected as an interval of time that would balance the need for re-evaluation based upon the slow dynamics of the stock and to minimize computational time per simulation.

Parameters determining unfished recruitment (*R0*), steepness, growth, annual recruitment deviations, initial age-structure deviations, and the size and width at maximum selectivity for the fishery were estimated. Natural mortality, the variation of length-at-age, weight-at-length, the fecundity relationship, and the variation of recruitment  were assumed known. The relative spawning biomass in the assessment year was estimated and the forecasted catches were determined using the harvest control rule adopted by the Pacific Fishery Management Council (PFMC) for rockfish. The catches were removed from the operating population without error, fishery CPUE index, length- and age-composition data were then generated for the subsequent six years.

The harvest control rule adopted by the PMFC for rockfish involves a linear reduction in catch when a stock falls below 0.40*SB*0, with no fishing when the stock falls below 0.10*SB*0. The overfishing level (the catch corresponding to the proxy for the fishing mortality at which maximum sustainable yield is achieved) was set equal to the target relative spawning biomass-per-recruit harvest rate (*F*0.50) multiplied by *SBt*. The overfished level was reduced by a management buffer (0.956) that accounts for the uncertainty about current biomass for well-assessed stocks to determine the acceptable biological catch level (i.e. acceptable biological catch = 0.956 overfishing level, Ralston *et al.*, 2011). The annual catch limit was set equal to the acceptable biological catch when the stock was above the target biomass, 0.40*SB*0,or reduced from the acceptable biological catch according to the harvest control rule when the stock fell below 0.40*SB*0.

One major simplification in this simulation design and actual management practice of U.S. west coast groundfish was the omission of the rebuilding plans that are implemented when a stock is assessed to have fallen below the MSST (defined as 0.25*SB*­0 for U.S. west coast rockfish). In reality, harvest for stocks below the MSST is not based on the standard harvest control rule, but rather a rebuilding plan that determines catches until the stock is rebuilt to the target biomass (see Wetzel and Punt, 2016 for additional details on PFMC rebuilding plans).

## Data scenarios

Three data scenarios were created to explore the impact of data availability on the ability to monitor rebuilding of an overfished stock (Fig. 2). The data scenarios were designed to emulate a stock, like yelloweye rockfish, that is infrequently encountered by a fishery-independent survey (e.g. due to depth or habitat) and only fishery data were available. The historical sample sizes were generally based on the effective sample sizes observed for yelloweye rockfish. Following the first assessment in year 50, the three scenarios have different data availability based on estimated stock status (e.g. overfished vs. rebuilt) in the assessment year.

The “full data” scenario maintained the fishery CPUE index and length- and age-composition data at the historical levels (prior to the stock being declared overfished in year 50) during rebuilding (Fig. 2). The “reduced data” scenario decreased the amount of data available from the fishery during rebuilding (Fig. 2). The length and age-composition data were reduced to 20% of the historical sample sizes during rebuilding and the fishery CPUE index was eliminated during the rebuilding period. The CPUE index resumes and composition sample sizes reverted to historical levels when the stock was estimated to have rebuilt to the target biomass. The “eliminated data” scenario had no fishery data during rebuilding (Fig. 2). The fishery CPUE index and composition data resumed at historical samples sizes when the stock was projected to be rebuilt.

The estimation method in the full and reduced data scenarios were allowed to estimate a change in selectivity from asymptotic to dome-shaped during the rebuilding period through the application of a time block on selectivity. However, the eliminated data scenario was forced to assume constant asymptotic selectivity over all years because no fishery composition data were available to detect a potential shift in selectivity.

## Performance measures

The outcomes of the simulations for each case and data scenario were summarized using five metrics that were selected to evaluate the impact of data on estimation of indicators of stock status (e.g. relative spawning biomass) and management quantities (e.g. rebuilding catch):

1. The relative errors (REs) for estimated parameters, calculated as:

 (5)

where *E* is the estimated quantity of interest and *T* is the true value from the operating model.

1. The percent root mean square error (RMSE), a measure of precision and bias, was calculated to assess the overall level of error given the amount of data available:

 (6)

where *N* is the number of simulations (*N* = 100).

1. The average (over simulations) of the total catch while the stock was recovering to the target biomass.
2. The annual average variability of the catches (abbreviation AAV), defined as:

(7)

where C*y* is the catch during year *y*.

# Results

## Assessment performance with time-invariant parameters

The trends of the relative error about spawning biomass and relative spawning biomass were generally consistent among the full and reduced data scenarios (Figs. 3A-B and 4A-B). The median estimates of spawning biomass and relative spawning biomass were less than the true values during rebuilding for both scenarios (Figs. 3A-B and 4A-B). As expected, there was less among-simulation variability in the difference between operating model and estimated spawning biomass and relative spawning biomass for the full data scenario during the rebuilding period compared to the reduced and eliminated data scenarios (Figs. 3A-C and 4A-C). However, by the end of the management period, the among-simulation variability of errors in biomass metrics were similar between the full and reduced data scenarios. The eliminated data scenario resulted in median (across simulations) estimates of spawning biomass and relative spawning biomass errors that were similar to the true values, but were highly imprecise at the start of the management period (years 50-74) (Figs. 3C and 4C). The eliminated data scenario, in the absence of new data during rebuilding, projected stocks based on the historical data and new catches until rebuilt, at which time data collected resumed allowing the estimation method to estimate population status. The median estimates for the eliminated data scenario were less than the true values, with high among-simulation variability in error as stocks began to be projected to be rebuilt and data collection resumed. In contrast to the full and reduced data scenarios, the estimates of spawning biomass and the relative spawning biomass for the eliminated data scenario showed little improvement in the among-simulation variability in error estimates by the end of the management period (Figs. 3C and 4C).

The RMSE for the estimated relative spawning biomass for each assessment year shows the increased precision of the full data scenario during the rebuilding period compared to the reduced and eliminated data scenarios (Fig. 5A). The eliminated data scenario resulted in the highest RMSE over the entire management period (Fig. 5A). However, the RMSE for the reduced data scenario showed improvement over the management period as stocks began to be assessed rebuilt to the target biomass, and sample sizes returned to historical levels. The limited improvement in the RMSE for the eliminated data scenario was driven by the simulations that were never projected to have rebuilt to the target biomass (35 out of 100 simulations).

Examining the eliminated data scenario closer revealed a pattern in the performance of the estimation method based on the estimation of steepness in the first assessment year. The eliminated data scenario simulations were divided and plotted based on whether the estimation method projected the simulation rebuilt (65 simulations) or failed to rebuild (35 simulations) by the end of the management period. To allow comparison, the estimates from the full data scenario were also divided into the same two groups and plotted. The estimated spawning biomasses were considerably less that the true values in the first assessment year (Figs. 6B [white]) for the 35 simulations that were estimated not to be rebuilt by the end of the management period. The underestimates of spawning biomass (Fig. 6B [white]) were driven by estimates of steepness that were much less than the true value in the first assessment (Fig. 6D [white]). In the absence of new data, the underestimates of steepness resulted in the estimation method perceiving a less productive stock requiring an extended period to rebuild to the target biomass. However, with full data present, estimated quantities (spawning biomass and steepness) improved for this subset of simulations and were median unbiased (the term "median unbiased" will be used to define cases in which the median of the relative errors equals zero) by the end of the management period (Fig. 6A and 6C [white]).

The estimates of steepness varied across data scenarios. The full data scenario resulted in median unbiased estimates by the end of the management period (Fig. 7A). In contrast, the median of the estimates of steepness for the reduced data scenario were greater than the true steepness during the management period (Fig. 7B). The eliminated data scenario had the highest among-simulation variability among estimates of steepness during the management period (Fig. 7C) due to the mixture of rebuilt and not rebuilt stocks.

The median number of years estimated for the stocks to recover to the target biomass for the full data scenario was longer than the median recovery year within the operating model (Table 2). In contrast, both the reduced and eliminated data scenarios had shorter median recovery times compared to the operating model (Table 2). The contrast in estimated recovery times across the data scenarios was related to the average catch obtained during rebuilding along with the bias and variability of estimates. The median error associated with relative spawning biomass for the full data scenario was less than zero, with low among-simulation variability (compared to the other data scenarios) for all assessment years, resulting in estimates that predicted constant rebuilding but at a slower rate than the true stock (Fig. 4A). In contrast, the reduced data scenario had higher variability over time (i.e. within-simulation) across the estimates of error associated with relative spawning biomass (Fig. 4B). The variability of estimates between assessments resulted in stocks being estimated recovered earlier than was the case in the operating model due to estimation error driven by the limited composition samples during rebuilding.

The reduced data scenario had the lowest median average catch during rebuilding (Table 3), with the median rebuilding time estimated shorter than the true time to recovery within the operating model (Table 2). The eliminated data scenario which was entirely dependent upon historical data until the stocks were projected to be rebuilt, essentially projected the population forward with each assessment based on the initial parameter estimates from the historical data, resulting in high median average catches during rebuilding, and the lowest median AAV during rebuilding and across the entire management period (Table 3).

Reduction or elimination of data during rebuilding increased the among-simulation variability about estimates of the size at maximum fishery selectivity which defined the ascending limb of the selectivity curve (see Fig. 1A), with the median estimates generally equal to the true value for all data scenarios (Fig. 8A-C). The among-simulation variability of the estimates for the reduced and eliminated scenarios improved when the majority of the stocks were estimated to be rebuilt, and fishery composition sample sizes returned to historical levels. The full and reduced data scenarios were allowed to estimate dome-shaped selectivity during the rebuilding period (the eliminated data scenario did not allow for estimation of dome-shaped selectivity due to the absence of fishery composition data), and resulted in median estimates that exceeded the true values and were highly variable among simulations at start of the management period (Fig. 9A-B). The estimates that exceeded the true values for this parameter indicated that the data available were not sufficient to inform the estimation method about the severity of the dome-shape in selectivity during rebuilding (a higher estimated value implies the dome in selectivity occurs at larger sizes with a higher proportion of the population relative to the operating model at full selectivity). The full data scenario resulted in markedly improved estimates of the shape of the dome over the management period, compared to the reduced data scenario (Fig. 9A-B).

## The impact of time-varying parameters

The case that assumed time-varying annual deviations in natural mortality and fishery selectivity generally resulted in increased among-simulation variation in estimation errors compared to the time-invariant case. The median error of estimates of spawning biomass at the time of the first assessment exceeded the true values and were highly variable among simulations (Fig. 3D-F). The among-simulation variance in errors of estimates of spawning biomass decreased markedly for the full data scenario after the first assessment (Fig. 3D). However, this variability remained high for approximately the first twenty-five years of the management period for both the reduced and eliminated data scenarios, until approximately 50% of the simulated stocks were estimated recovered and the fishery sample sizes increased to historical levels (Fig. 3E-F). The full and reduced data scenarios resulted in median spawning biomass estimates that were generally smaller than the operating model values (Fig. 3D-E). However, the medians of the errors for relative spawning biomasses were variable over the management period (Fig. 4D-E). The medians of the estimates of relative spawning biomass for eliminated data scenario were larger than operating model values at the start of the management period, but became smaller than these values as stocks rebuilt to target biomass and data collection resumed (Fig. 4F).

Compared to the case with time-invariant parameters the RMSE was higher for all data scenarios when time-varying parameters were present within the operating model (Fig. 5). The RMSE about the estimated relative spawning biomass for the full data scenario was lower relative to the other scenarios for the entire management period (Fig. 5B). Similar to the time-invariant results, the RMSE of relative spawning biomass for the eliminated data scenario was the highest among the scenarios across the entire management period, peaking in assessment year 68 at 221% (a single simulation for the eliminated data scenario, with extreme outliers for two assessment years, was removed for a more informative summary of the RMSE).

The time-varying results for the eliminated data scenario were qualitatively similar to time-invariant case, where a large number of simulations failed to be projected rebuilt (32 simulations). As was observed in the time-invariant case, the simulations that failed to be projected rebuilt had median estimates of spawning biomass and relative spawning biomass below the operating model values at the time of the time of the first assessment, which were driven by estimates of steepness that were considerably lower than the true value (not shown).

The inclusion of time-varying parameters in the operating model resulted in shorter median estimated recovery times relative to the time-invariant case for the full and reduced data scenarios (Table 2). However, the median number of years to rebuild for the operating model stocks were similar between the time-varying and -invariant cases. The estimation method estimated earlier recovery times for the time-varying case due to the increased variability in the estimates of relative spawning biomass, resulting in the estimation method having an increased frequency of erroneously estimating the biomass to be above the target stock size (Fig. 4).

The median average catch during the recovery period was the highest for the eliminated data scenario due to estimates of the relative spawning biomass that were higher than the true values at the start of the management period (Table 3 and Fig. 4F). Additionally, the eliminated data scenario had the lowest median AAV during the rebuilding period (Table 3). In contrast, the eliminated data scenario, resulted in the highest number of stocks that were never projected to exceed the target biomass (Table 3).

Inclusion of time-varying selectivity resulted in the median estimates of maximum selectivity across all data scenarios exceeding the mean of the operating model values (Fig. 8D-F), although the full data scenario resulted in the lowest among-simulation variation. The full and reduced data scenarios, which were allowed to estimate dome-shaped selectivity (width at maximum selectivity) during the recovery period, resulted in highly variable among-simulation estimates at the start of the management, period with the variability for the estimates decreasing earlier for the full data scenario (Fig. 9C-D).

# Discussion

Maintaining fishery data at historical levels during rebuilding reduced the variation in estimates between assessments (i.e. over time within a simulation). While the full data scenario had less variation, the median (over simulations) estimates of spawning biomass and relative spawning biomass were consistently below the operating model values for much of the management period. This result is contrary to what might be expected when additional data are available. Explorations with simulations where there was a fishery independent survey that provided an index of abundance and composition data (length and age), determined that this underestimation of the true spawning biomass was eliminated if survey composition data were available along with fishery composition data (see *Supplementary information*). The underestimation was driven by two key factors; the shape of fishery selectivity curve and data quantity. The fishery selectivity curve was specified to be greater than the maturity-at-length curve, resulting in only mature larger fish being selected by the fishery. The fishery data were informative about recruitment, but with a lag between recruitment to population and recruitment to the fishery. However, a fishery-independent survey selecting fish at smaller sizes yields information about recruitment to population earlier. Additionally, an increase in the length- and age-composition samples from multiple data sources can improve estimates of recruitment, spawning biomass, and relative spawning biomass (Yin and Sampson, 2004; Wetzel and Punt, 2011).

Median relative errors for relative spawning biomass below zero for the full data scenario resulted in the estimation method failing to determine that the operating model population was at or above the target biomass (median number of rebuilding years greater than the operating model, Table 2), an outcome that would lead to extended harvest restrictions that were not warranted given the true state of the population, a situation fishery management would like to avoid. However, the reduced estimation variability (within- and among-simulations) offered by the full data scenario resulted in an improvement in the consistency of estimates by subsequent assessments, which offers a level of stability for fisheries managers and stakeholders. In contrast, the higher between-assessment variation in estimates of spawning biomass for the reduced data scenario resulted in stocks being estimated rebuilt when the true population was still below the target biomass which could have undesirable outcomes for fisheries management. Overly optimistic estimates of relative spawning biomass can result in overfishing when catch limits are set too high, leading to further reductions in biomass, potentially requiring an overfished declaration by a future assessment.

Loss of data during rebuilding resulted in a high number of simulations that failed to rebuild due to poor initial estimates of steepness, a key parameter, controlling how quickly a stock can rebuild from low biomass levels. In the absence of new data, the first and subsequent assessments were entirely dependent on the quality of the historical data to inform parameter estimates. The simulations that failed to correctly detect rebuilt stocks were driven by erroneously low estimates of steepness at the time of the first assessment. Estimating a stock to be less productive than the true population resulted in lower estimates of spawning biomass and relative spawning biomass, with the assessment setting harvest at levels well below the true acceptable biological catch. The reduced harvest allowed the population in the operating model to rebuild to or above the target biomass. However, in the absence of new (and informative) data, the estimation method was unable to detect the correct stock size. The operating model population represented a two-way trend of abundance (decline and increase in biomass) with fishery data available during the fishing down and recovery period, which previous studies have found informative in estimating steepness (Magnusson and Hilborn, 2007; Conn *et al.*, 2010). This work showed that a one-way trip scenario in stock size with limited data may not be adequate to correctly estimate steepness, but the inclusion of even limited data can, with contrast in stock size, improve the estimation of steepness even if the initial assessment produced a poor estimate (Figs. 6C and 7).

The general trend in results when the operating model included time-varying natural mortality and fishery selectivity were similar to the time-invariant case, although the among-simulation estimates were more variable across all data scenarios. The estimation method assumed a single natural mortality across all years equal to the mean value that was used to generate the autocorrelated annual deviations in the operating model. This setup was a strategic choice that allowed variation in the composition data that the estimation method would not be able to account for, but would not be anticipated to result in strongly biased estimates due to model misspecification. Time-varying natural mortality is likely more complex with extended periods of high or low mortality based upon external factors (e.g. predator abundance, climate conditions), which could result in large biases in estimated quantities if not accounted for in an assessment (Johnson *et al.* 2014).

Shifts in the form of selectivity over time and the impact of annual deviations in selectivity led to mixed results. The estimation method consistently overestimated the mean size at maximum selectivity for all data scenarios with time-varying selectivity. The operating model selectivity involved normally distributed deviations to generate the annual shifts in selectivity and one would not *a priori* predict the estimation method to have a consistent bias in estimates. However, the estimation method was able to identify the change in the selectivity form (asymptotic to dome-shaped through a reduction in the width at peak selectivity) during the rebuilding years with similar error to the time-invariant case. Each case led to estimates that overestimated the width at maximum selectivity, defining the dome in selectivity (dome-shaped selectivity occurring at larger sizes with increased sizes subject to full selectivity compared to the operating model). This evaluation applied time blocks defined by the status of the stock to allow for shifts in selectivity, ignoring the annual deviations in the selectivity curve. Studies have evaluated other ways of estimating time-varying selectivity using state-space models (Nielsen and Berg, 2014), or the implications of applying time blocks vs. allowing a random walk component in selectivity parameters or catchability (Wilber and Bence, 2006; Martell and Stewart, 2014). Further exploration should be conducted to evaluate if allowing a random walk or applying an alternative estimation method eliminates the bias observed in the estimated selectivity observed here and how data quantity and quality impacts these estimates. Additionally, if shifts in fishery selectivity are anticipated due to management actions, increased data collections may be required in order to achieve a similar level of precision in estimates during rebuilding.

This work highlights the benefits of continued data collection during rebuilding on the precision of estimates, but there are many additional reasons why retaining data streams is important. The first benefit of continued data collection is the potential ability to identify misspecification in the model assumptions. The estimation method and operating models applied here, generally made similar structural assumptions. However, the true state of nature is never known with confidence in real-world assessments. Continued data collection may allow for the identification of model misspecification in the structural assumptions (e.g. growth, recruitment), which will allow for better approximations to reality through models. Specifically, there could be long-term impacts to a population that is depleted, which could negatively impact the ability of the stock to rebuild (e.g. Hixon *et al*., 2014) that additional data would be required to detect. The second benefit of continued data collections is that ability to identify potential long-term changes in stock dynamics (e.g. changes in climate-driven forces). There has been much work identifying potential links between climate and recruitment (Hollowed *et al.*, 2011; Ianelli *et al.*, 2011; Mueter *et al.*, 2011; Strachura *et al.*, 2014). Additionally, long-term varying climate conditions may result in changes in biological parameters (Swain and Benoit, 2015), which will impact the productivity of stocks (Legault and Palmer, 2015) that needs to be accounted for when setting harvest limits.

# Acknowledgements

This work has benefitted from feedback provided by Ian Taylor (NWFSC), Vladlena Gertseva (NWFSC), Ian Stewart (IPHC), and Allan Hicks (IPHC).

# References

Chen, Y., Chen, L., and K.I. Stergiou. 2003. Impacts of data quantity on fisheries stock assessment. Aquat. Sci. 65: 92-98.

Conn, P.B., Williams, E.H., and K.W. Shertzer. 2010. When can we reliably estimate the productivity of fish stocks? Can. J. Fish. Aquat. Sci. 67: 511-523.

Hilborn, R. 1979. Comparison of fisheries control systems that utilize catch and effort data. J. Fish. Res. Board. Can. 36: 1477-1489.

Hixon, M.A., Johnson, D.W., and S.M. Sogard. 2014. BOFFFs: on the importance of conserving old-growth age structure in fishery populations. ICES J. Mar. Sci. 71: 2171-2185.

Hollowed, A. B., Barange, M., Ito, S., Kim, S., Loeng, H., and M.A. Peck. 2011. Effects of climate change on fish and fisheries: forecasting impacts, assessing ecosystem responses, and evaluating management strategies. ICES J. Mar. Sci. 68: 984–985.

Ianelli, J. N., Hollowed, A. B., Haynie, A. C., Muter, F. J., and N.A. Bond. 2011. Evaluating management strategies for eastern Bering Sea walleye pollock (Theragra chalcogramma) in a changing environment. ICES J. Mar. Sci. 68: 1297–1304.

Johnson, K.F., Monnahan, C.C., McGilliard, C.R., Vert-pre, K.A., Anderson, S.C., Cunningham, C.J., Hurtado-Ferro, F., Licandeo, R.R., Muradian, M.L., Ono, K., Szuwalski, C.S., Valero, J.L., Whitten, A.R., and A.E. Punt. 2014. Time-varying natural mortality in fisheries stock assessment models: identifying a default approach. ICES J. Mar. Sci. 72: 137-150.

Legault, C. M., and M.C. Palmer. 2015. In what direction should the fishing mortality target change when natural mortality increases within an assessment? Can. J. Fish. Aquat. Sci. 73: 349–357.

Lee, H.H., Maunder, M.N., Piner, K.R., and R.D. Methot. 2012. Can steepness of the stock-recruitment relationship be estimated in fishery stock assessment models? Fish. Res. 125-126: 254-261.

Magnusson, A., and R. Hilborn. 2007. What makes fisheries data informative? Fish Fish. 8: 337-358.

Martell, S., and I. Stewart. 2013. Towards defining good practices for modeling time-varying selectivity. Fish. Res. 158: 84-95.

Methot. Jr., R.D. (editor) 2015. Prioritizing fish stock assessments. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-F/SPO-152, 31 pp.

Methot, R.D., and K. Piner. 2002. Rebuilding analysis for yelloweye rockfish: Update to incorporate results of coastwide assessment in 2002. Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.12 pp.

Methot, R.D., and C.R. Wetzel. 2013. Stock Synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142: 86-99.

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

Ralston, S., Punt, A.E., Hamel, O.S., DeVore, J.D., and R.J. Conser. 2011. A meta-analytic approach to quantifying scientific uncertainty in stock assessments. Fish. Bull. 109: 217-231.

Sustainable Fisheries Act of the Magnuson-Stevens Fishery Conservation and Management Act (SFA). 1996. Silver Spring: U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service. 170 pp.

Stachura, M. M., Essington, T. E., Mantua, N. J., Hollowed, A. B., Haltuch, M. A., Spencer, P. D., Branch, T. A., and M.J. Doyle. 2014. Linking Northeast Pacific recruitment synchrony to environmental variability. Fish. Oceanogr. 23: 389–408.

Stewart, I.J., Wallace, J.R., and C. McGilliard. 2009. Status of the U.S. yelloweye rockfish resource in 2009. Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220. 435 pp.

Swain, D. P., and H.P. Benoit. 2015. Extreme increases in natural mortality prevent recovery of collapsed fish populations in a Northwest Atlantic ecosystem. Mar. Ecol. Prog. Ser. 519: 165–182.

Wilberg, M.J., and J.R. Bence. 2006. Performance of time-varying catchability estimators in statistical catch-at-age analysis. Can. J. Fish. Aquat. Sci. 63: 2275-2285.

Wetzel, C.R., and A.E. Punt. 2011. Performance of a fisheries catch-at-age model (Stock Synthesis) in data-limited situations. Mar. Fresh. Res. 62: 927-936.

Wetzel, C.R. and A.E. Punt. 2016. The impact of alternative rebuilding strategies to rebuild overfished stocks. ICES J. Mar. Sci. 73: 2190-2207.

Yin, Y., and D.B. Samson. 2004. Bias and precision of estimates from an age-structured stock assessment program in relation to stock and data characteristics. North. Am. Jour. Fish. Manage. 24: 865-879.

# Tables

Table 1. Life history and observation parameters used in the operating model and their treatment within the estimation model.

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | Time-invariant | Time-varying | Treatment in estimation model |
| Natural mortality(*M*) yr-1 | 0.08 | 0.08 | Fixed |
| Natural mortality standard error (σm) | 0 | 0.10 |  |
| Natural mortality autocorrelation (ρ) | 0 | 0.707 |  |
| Steepness(*h*) | 0.65 |  | Estimated |
| Maximum length (*L∞*)(cm) | 64 |  | Estimated |
| Growth coefficient (*kt*) (yr-1) | 0.05 |  | Estimated |
| Weight at length (kg) | α =1.50x10-5β = 3 |  | Fixed |
| Length at 50% maturity (cm) | 37 |  | Fixed |
| Recruitment variation (σR) | 0.50 |  | Fixed |
| Fishery CPUE standard error (σf) | 0.30 |  | Fixed |
| Fishery CPUE catchability coefficient (*Qf)* | 0.01 |  | Analytically estimated |

Table 2. The median and 90% simulation interval for the estimated number of years needed to rebuild to the target biomass, the operating model number of years needed to rebuild to target biomass, and the number of stocks that failed to rebuild to the target biomass determined by the estimation method (EM) and the operating model (OM) for each case and data scenario.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Selectivity / data scenario | Estimated num. of rebuilding years | | Operating model num. of rebuilding years | | Num. of stocks that failed to rebuild | |
|  | Median | 90% SI | Median | 90% SI | EM | OM |
| Time-invariant |  |  |  |  |  |  |
| full data | 43 | (13 - 87) | 34 | (16 - 73) | 7 | 4 |
| reduced data | 31 | (19 - 61) | 34 | (14 - 83) | 1 | 5 |
| eliminated data | 25 | (14 - 72) | 37 | (14 - 87) | 35 | 4 |
| Time-varying |  |  |  |  |  |  |
| full data | 31 | (13 - 91) | 35 | (13 - 85) | 13 | 4 |
| reduced data | 25 | (13 - 79) | 32 | (12 - 74) | 8 | 2 |
| eliminated data | 25 | (13 - 77) | 36 | (12 - 79) | 32 | 5 |

Table 3. The median and 90% simulation intervals for the average catch during rebuilding, the AAV during rebuilding, and the AAV over-all years for each case and data scenario.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Selectivity / data scenario | Average catch during rebuilding | | AAV during rebuilding | | AAV all years | |
|  | Median | 90% SI | Median | 90% SI | Median | 90% SI |
| Time-invariant |  |  |  |  |  |  |
| full data | 44.0 | (15.3 - 78.9) | 6.0 | (3.7 - 11.5) | 3.2 | (2.1 - 4.7) |
| reduced data | 28.1 | (14.6 - 57.9) | 7.7 | (4.0 - 14.5) | 3.5 | (2.3 - 5.3) |
| eliminated data | 41.3 | (19.9 - 83.8) | 2.6 | (1.3 - 4.4) | 2.2 | (1.3 - 3.9) |
| Time-varying |  |  |  |  |  |  |
| full data | 31.7 | (11.0 - 75.4) | 7.3 | (4.4 - 17.5) | 4.2 | (2.7 - 5.9) |
| reduced data | 25.1 | (15.6 - 68.0) | 8.9 | (4.5 - 20.7) | 4.5 | (2.6 - 9.8) |
| eliminated data | 36.3 | (15.7 - 79.4) | 2.3 | (1.2 - 4.8) | 2.8 | (1.3 - 5.3) |

# Figures

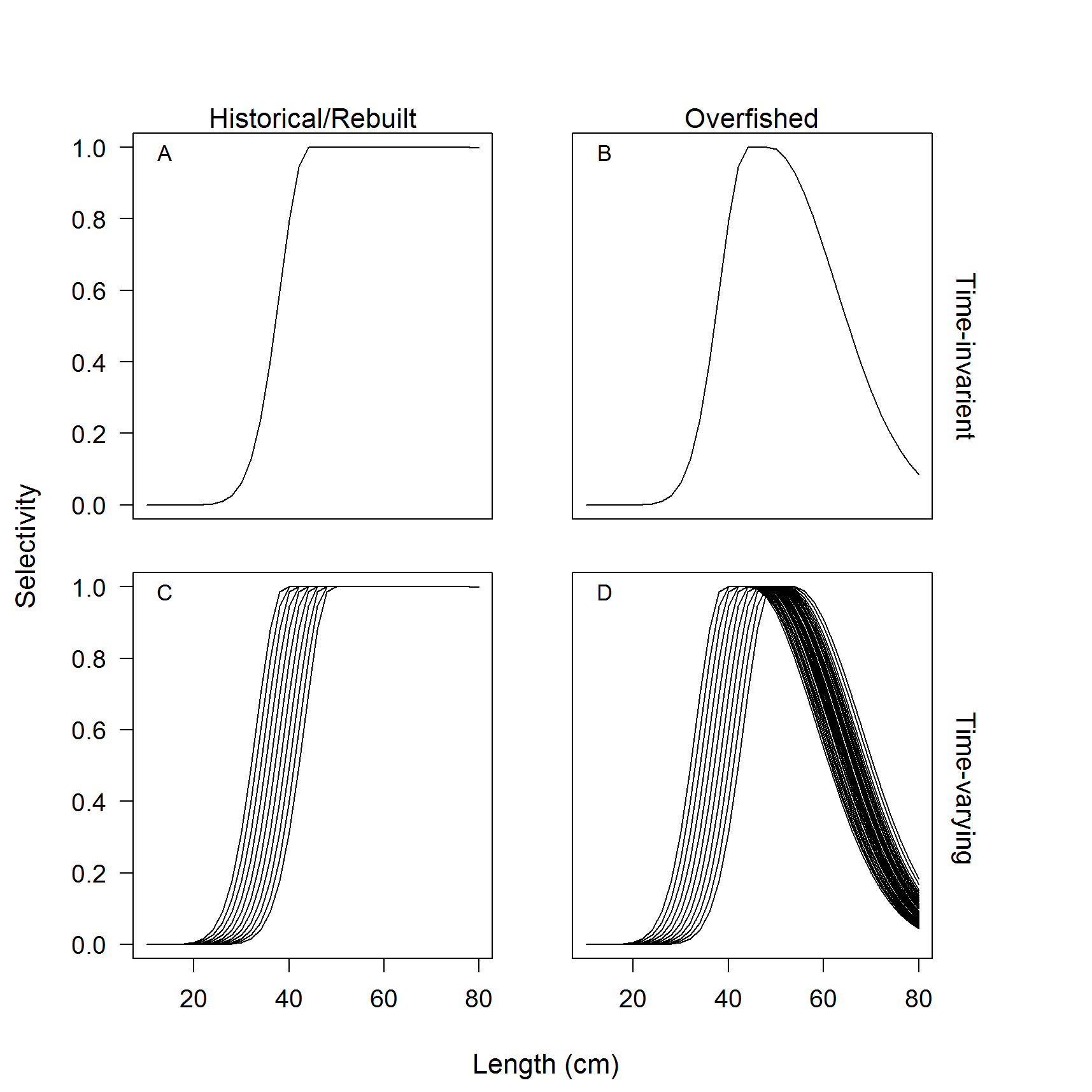


Figure 1. Fishery selectivity for either time-invariant or time-varying selectivity for the historical/rebuilt (A and C) and overfished (B and D) periods. A standard error of 0.05 was applied annually about the size at maximum selectivity, which defined the variability among the ascending limb of the selectivity curve (C and D), and a standard error of 0.20 was applied for the width at maximum selectivity that defined the length at which the dome in selectivity began while the stock was estimated overfished (D) (see Methot and Wetzel, 2013, for additional details on double normal selectivity).

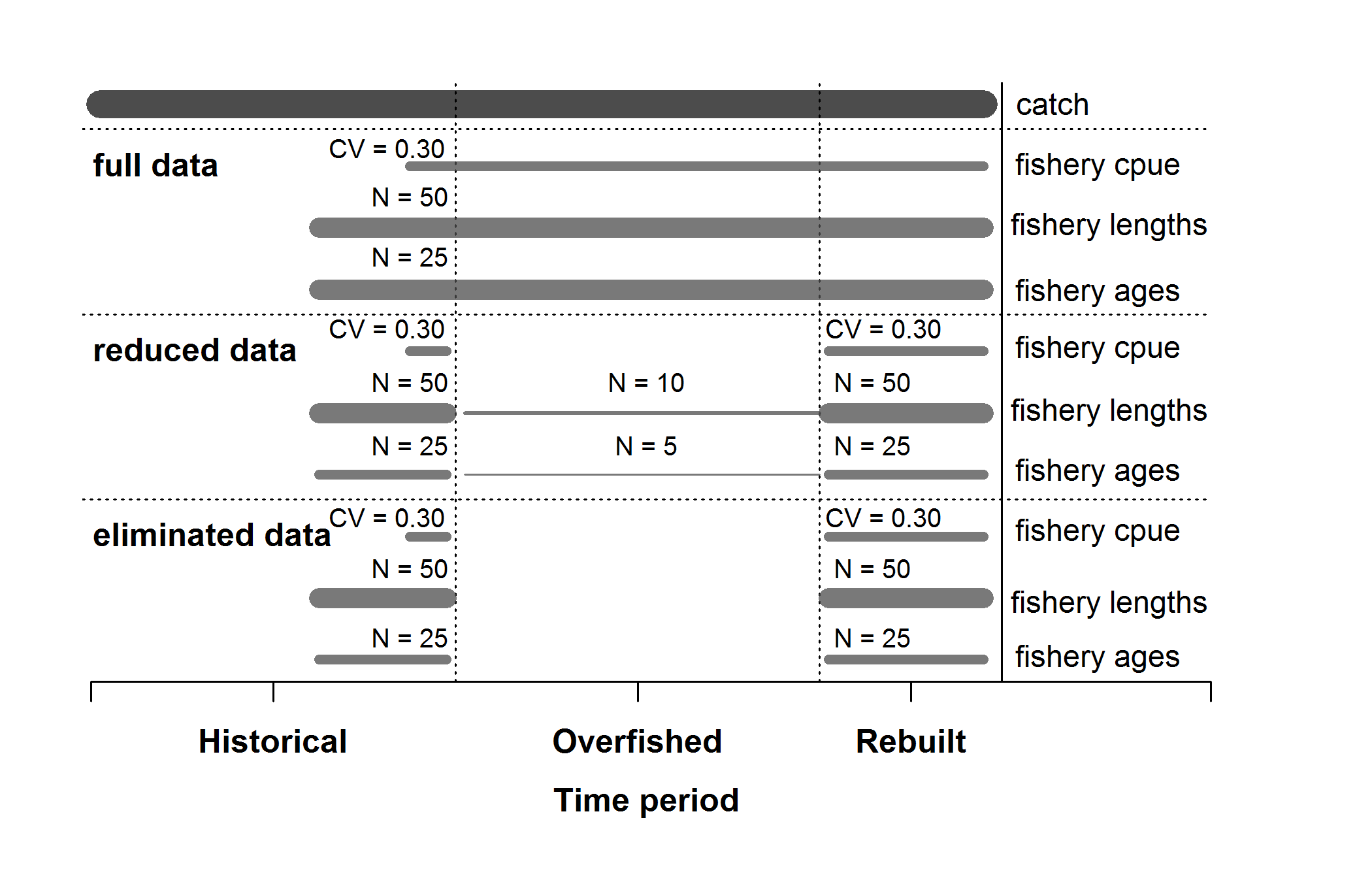


Figure 2. Summary of the data available for each of the data scenarios. Catches were known without error and were available for all data scenarios.

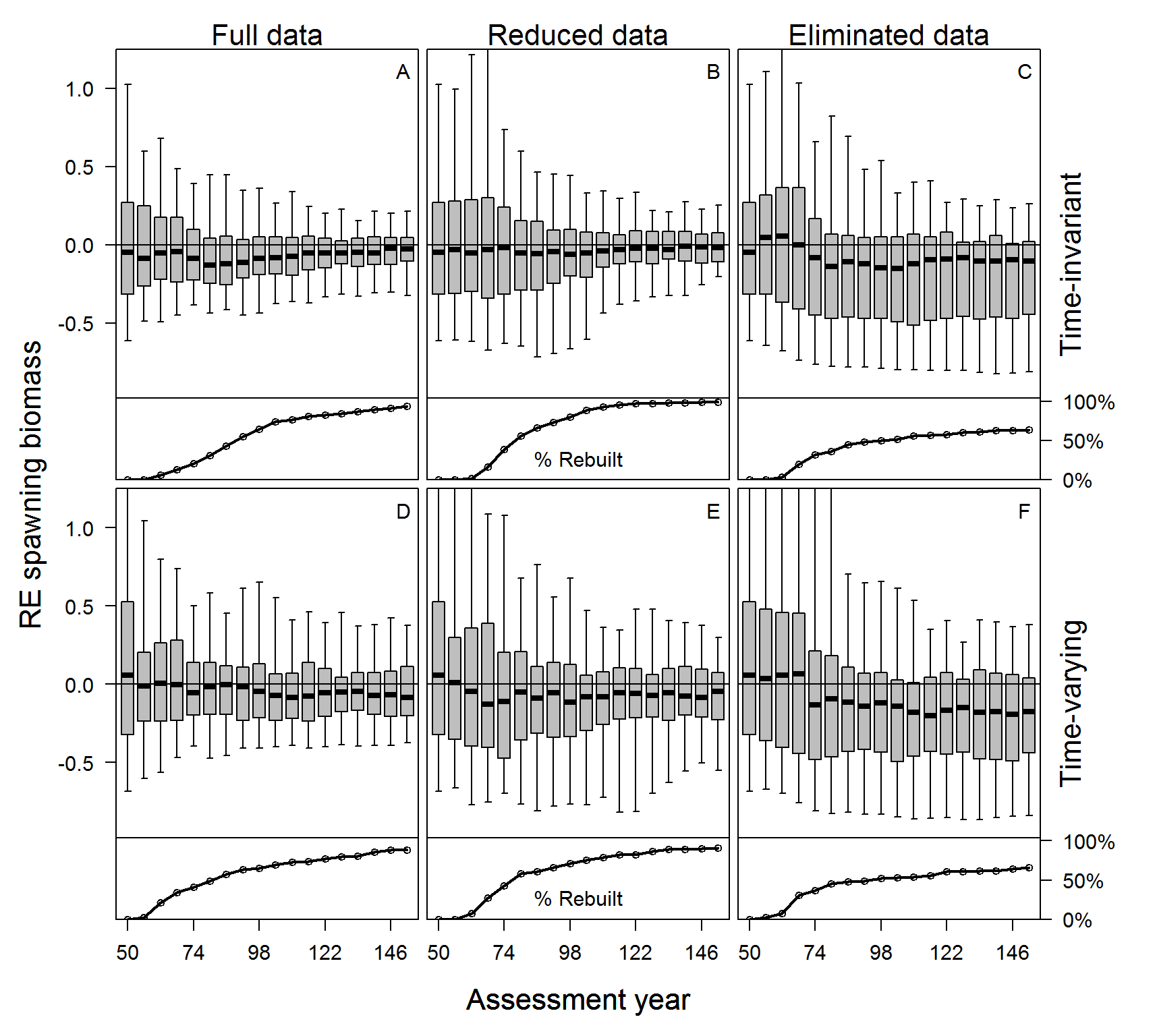


Figure 3. Relative error of estimated spawning biomass in each assessment year for each case and data scenario for all simulations (top panel) and the percentage of stocks that had rebuilt to the target biomass during the management period (bottom panel), with data collection consequently returning to historical levels. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

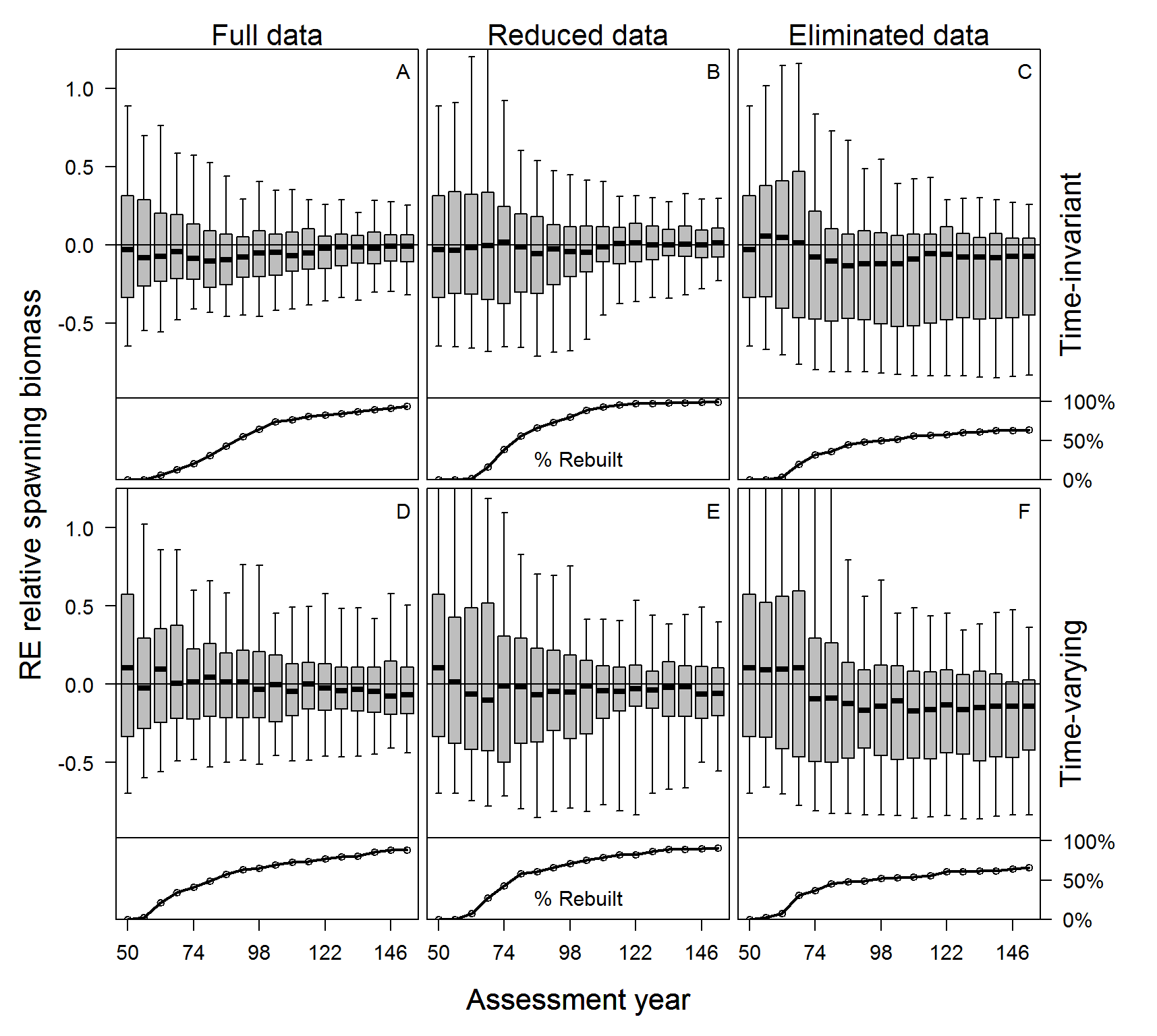


Figure 4. Relative error of estimated relative spawning biomass in each assessment year for each case and data scenario for all simulations (top panel) and the percentage of stocks that had rebuilt to the target biomass during the management period, with data collection consequently returning to historical levels (bottom panel). The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

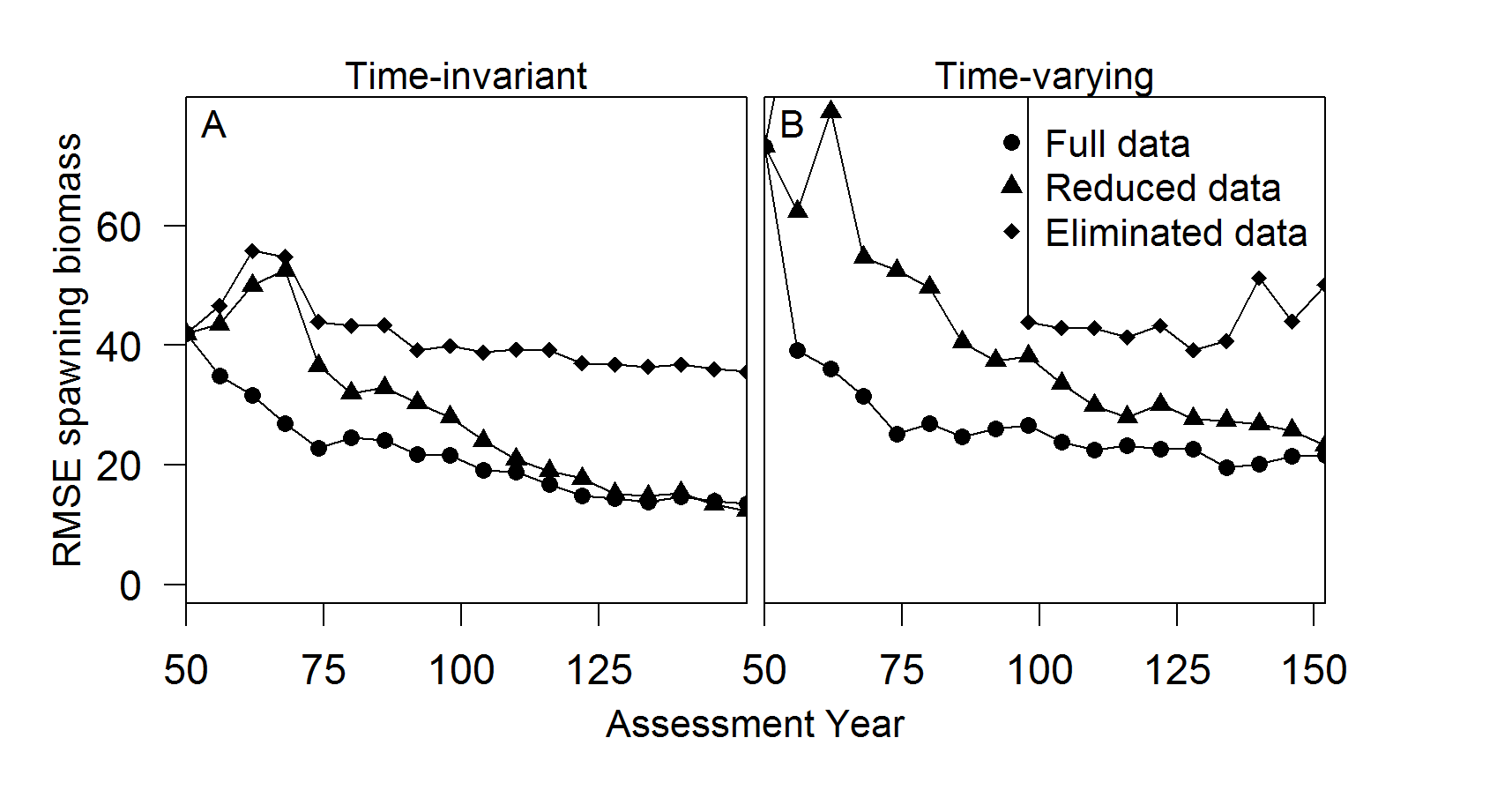


Figure 5. The root mean square error about relative spawning biomass by assessment year for each case and data scenario. The scale of the y-axis is the same for comparability of results between the time-invariant and the time-varying simulations. The time-varying eliminated data scenario peaked in year 68 at 221% RMSE.

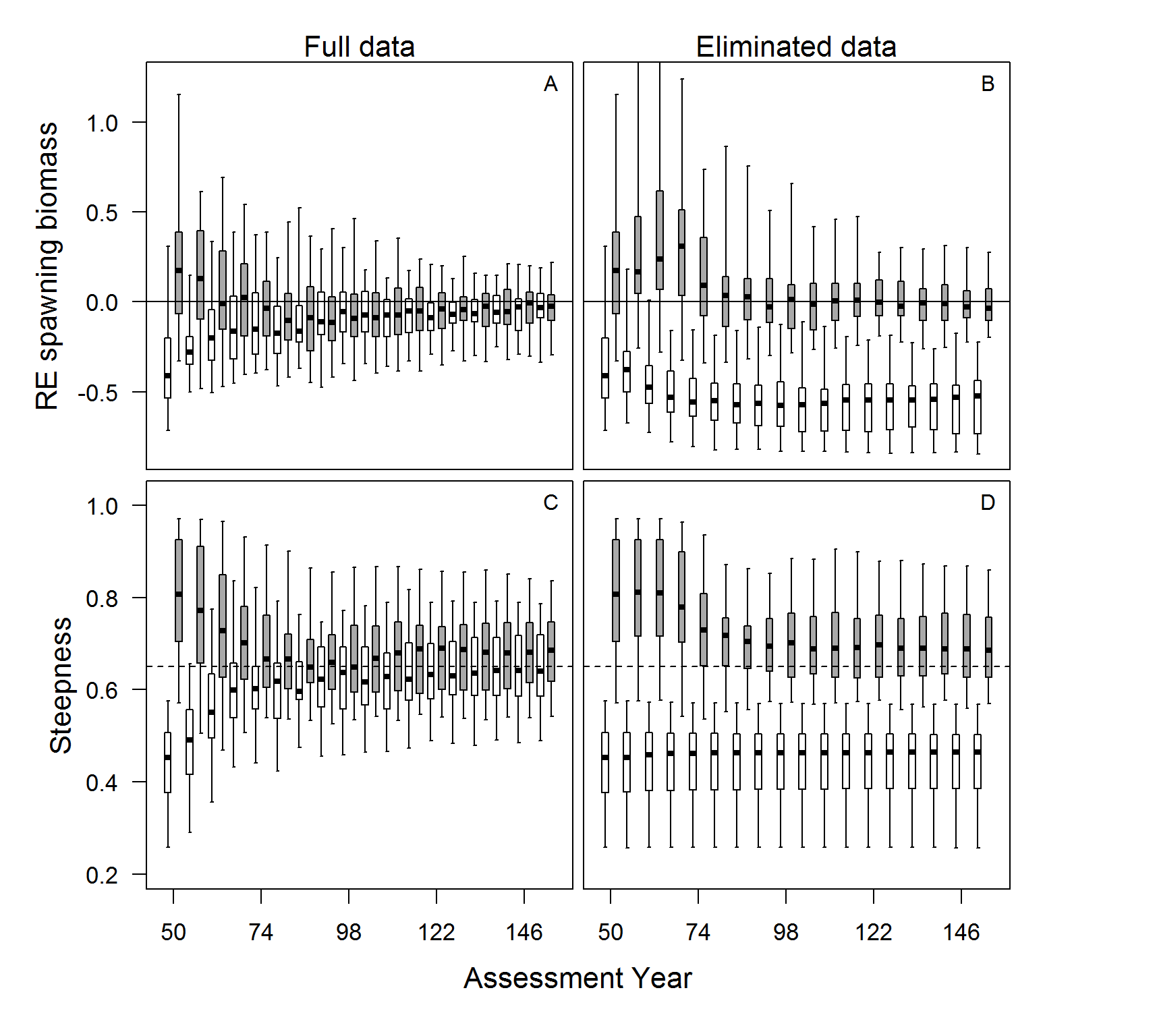


Figure 6. Relative error of spawning biomass and the estimates of steepness for the full and eliminated data scenarios for the time-invariant case, with the results divided by whether the simulated stock was estimated to be rebuilt (35 simulations [white]) or not (65 simulations [grey]) for the eliminated data scenario. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

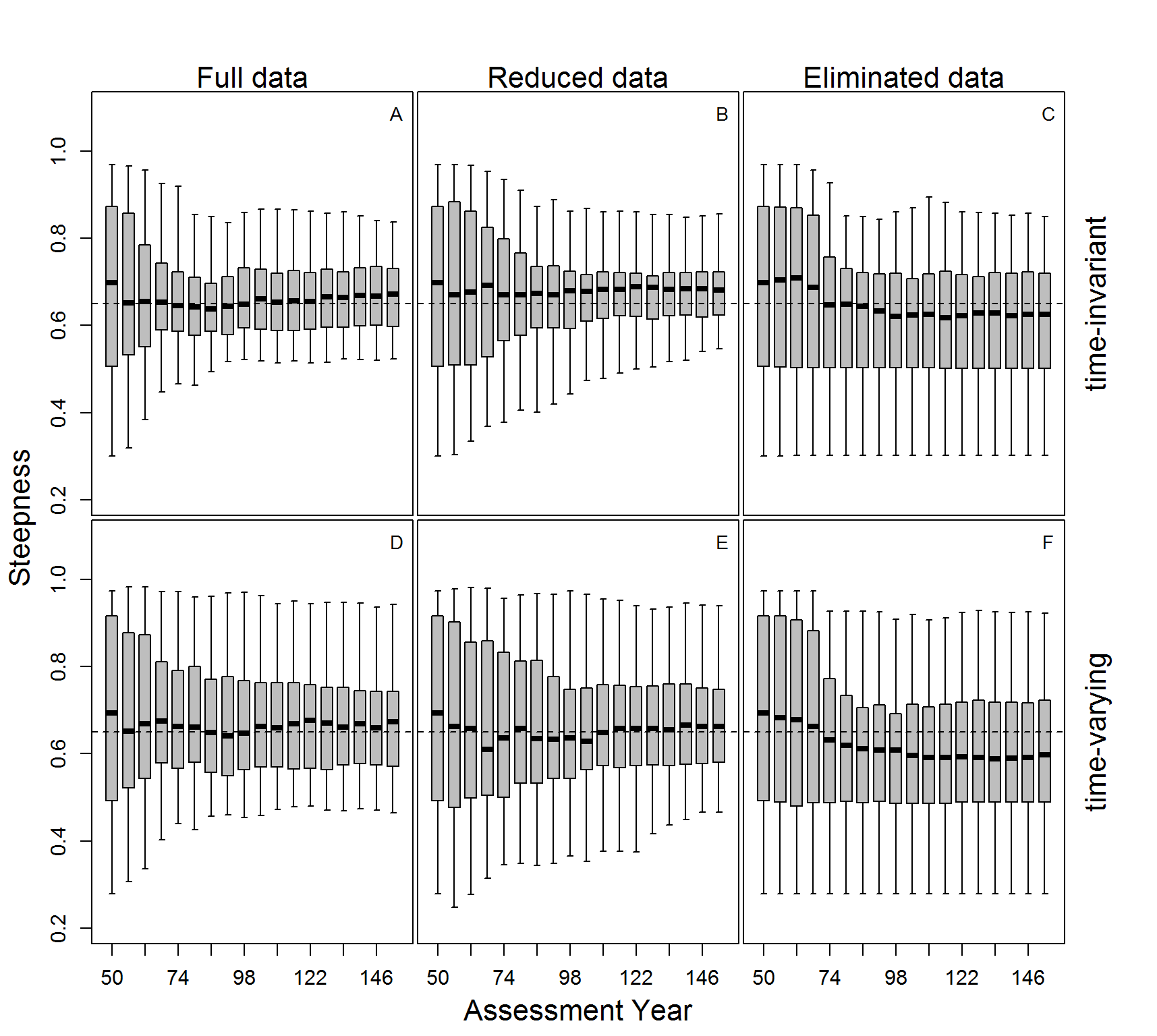
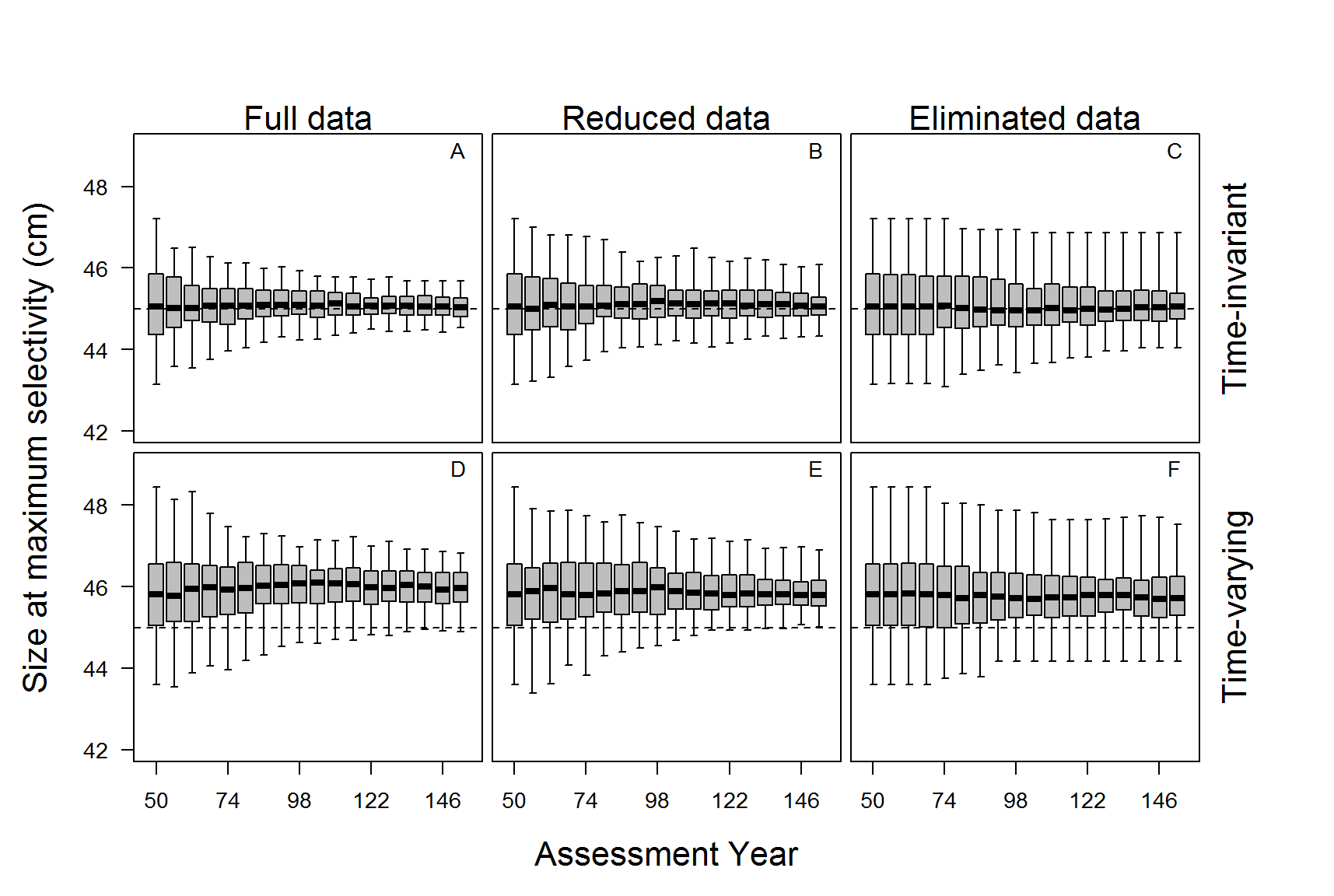


Figure 7. The estimates of steepness in each assessment year for each case and data scenario. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

Figure 8. The estimated size at maximum fishery selectivity for each data scenario and case. The black dashed line indicates the operating model parameter value. The estimates from the data scenarios with time-varying selectivity were compared against the mean of the distribution from the operating model. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

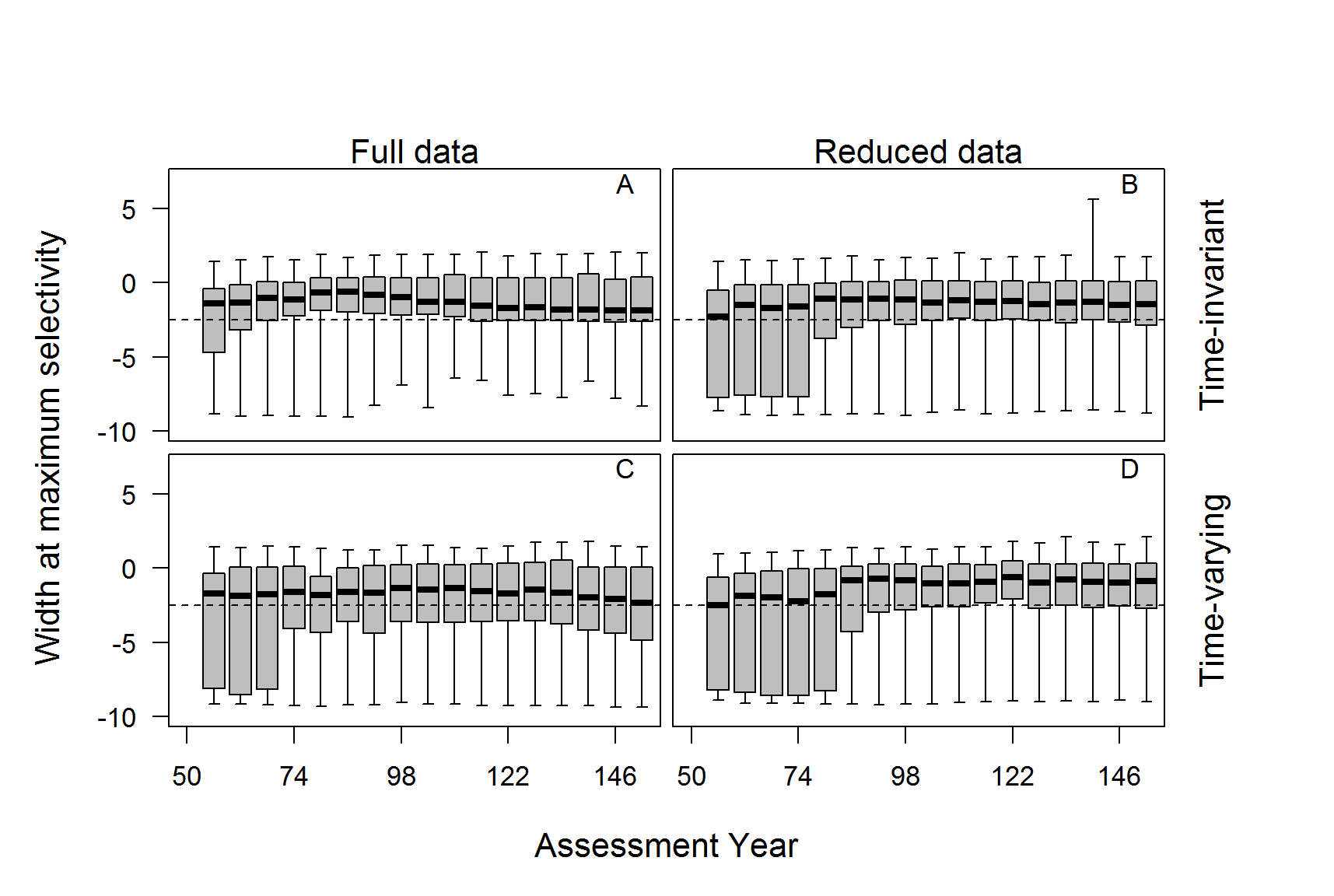


Figure 9. The estimated width at maximum selectivity (i.e. defines the extent of dome-shaped selectivity) for the fishery that occurred while the stock was overfished for each data scenario and case. The black dashed line indicates the operating model parameter value. The estimates from the data scenarios with time-varying were compared against the mean of the distribution from the operating model. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

# Supplementary information

Additional simulations were conducted to evaluate the impact of only having fishery information versus indices of abundance and length- and age-composition data available from both a fishery-independent survey and a fishery. The operating model generated a highly uncertain survey (CV = 0.40) that was conducted on a biennial basis with low length- and age-composition sample sizes, representative of a survey that poorly sampled the simulated stock (e.g. due to habitat and gear restrictions or restricted sampling areas) (Fig. Supp.1). The survey selectivity was assumed to be fixed at an asymptotic shape, selecting fish at smaller sizes relative to the fishery selectivity. All other specifications for the fishery within the operating model and the assumptions applied by the estimation method were the same as those detailed in the Material and Methods section.

The estimates of spawning biomass (Fig. Supp.2A-C) and relative spawning biomass (Fig. Supp.3a-c) for the time-invariant case were median unbiased at the time of the first assessment in year 50. The addition of a survey index and composition data for all data scenarios led to less among-simulation variability and reduced median bias over the management period relative to the simulations without survey data (Figs. 4A-C and 5A-C). The presence of survey data when fishery data were eliminated (eliminated data scenario) allowed the majority of the simulated stocks being estimated rebuilt by the end of the management period (Fig. Supp.2C) compared to the large fraction of simulations that failed to be estimated rebuilt when only historical data were available from the fishery (Fig. 4C).

The inclusion of survey data, in addition to fishery data when time-varying parameters were present led reduced the among-simulation variability in the estimates of spawning biomass and relative spawning biomass (Figs. Supp.2D-F and Supp.3D-F). The full data scenario had the lowest RMSE for relative spawning biomass during the early portion of the management period for both cases (time-invariant and time-varying), when the majority of simulations were rebuilding for both cases (Fig. Supp.4). However, midway through the management period, after a majority of the simulated stocks had rebuilt and data restrictions were removed, the data scenarios resulted in similar RMSEs (Fig. Supp.4). The inclusion of survey data for all data scenarios resulted in similar estimates of the median number of years to recover to the target biomass, which were similar to the median rebuilding time with the operating model (Table Supp.1).

Table Supp.1 The median and 90% simulation interval for the estimated number of years needed to rebuild to the target biomass, the operating model number of years needed to rebuild to target biomass, and the number of stocks that failed to rebuild to the target biomass determined by the estimation method (EM) and the operating model (OM) for each case and data scenario.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Selectivity / data scenario | Estimated num. of rebuilding years | | Operating model num. of rebuilding years | | Num. of stocks that failed to rebuild | |
|  | Median | 90% SI | Median | 90% SI | EM | OM |
| Time-invariant |  |  |  |  |  |  |
| full data | 37 | (18 - 79) | 39 | (18 - 79) | 2 | 2 |
| reduced data | 37 | (19 - 67) | 38 | (16 - 78) | 0 | 4 |
| eliminated data | 37 | (18 - 74) | 38.5 | (16 - 84) | 2 | 4 |
| Time-varying |  |  |  |  |  |  |
| full data | 31 | (13 - 85) | 41 | (14 - 85) | 15 | 3 |
| reduced data | 31 | (13 - 85) | 39 | (13 - 80) | 9 | 3 |
| eliminated data | 31 | (13 - 88) | 39 | (13 - 80) | 8 | 3 |

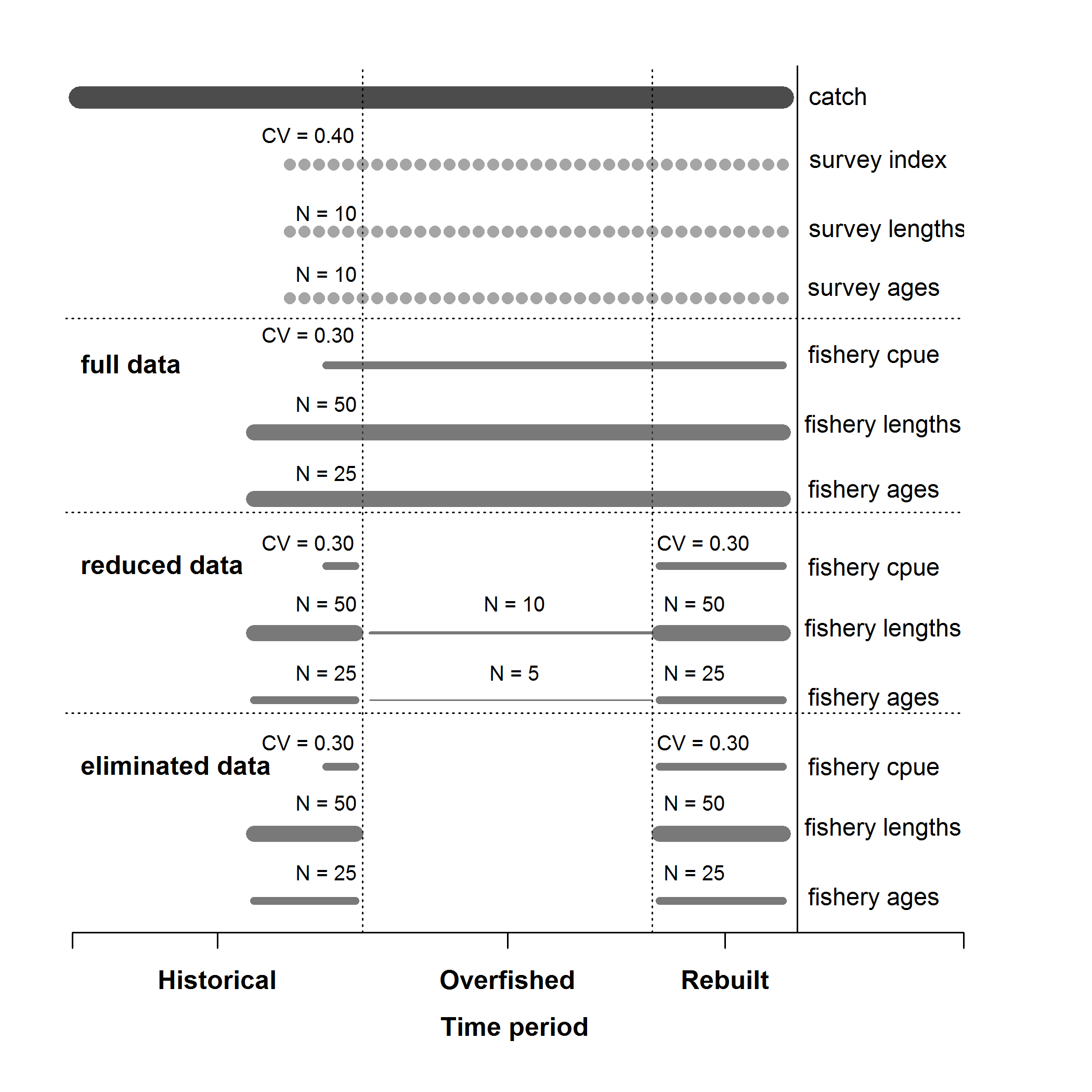


Figure Supp.1. Summary of the data available for each of the data scenarios. Catches, a fishery independent survey with length- and age-composition data were available for all data scenarios.

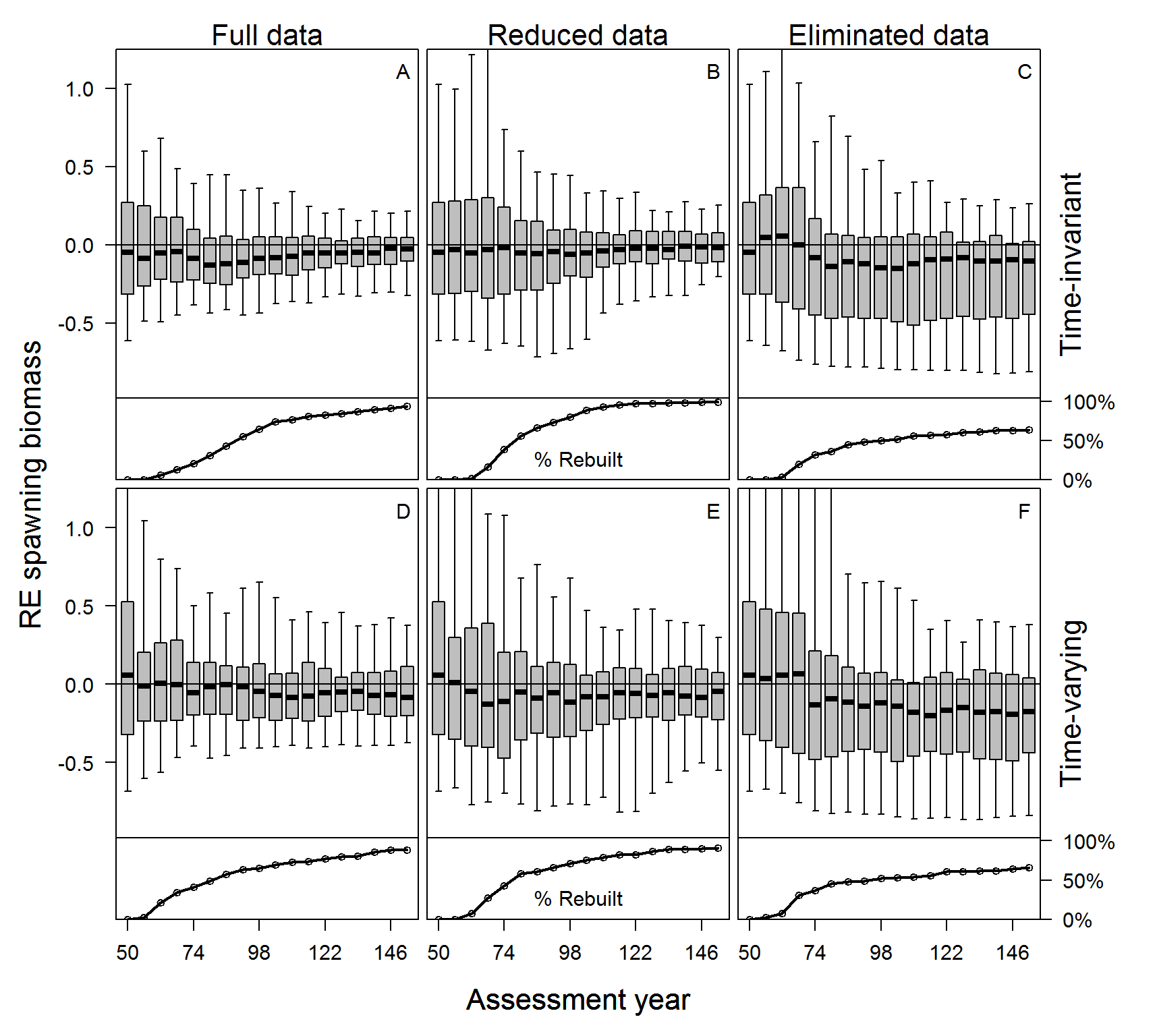


Figure Supp.2. Relative error of estimated spawning biomass in each assessment year for each case and data scenario for all simulations (top panel) and the percentage of stocks that had rebuilt to the target biomass during the management period (bottom panel), with data collection consequently returning to historical levels. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

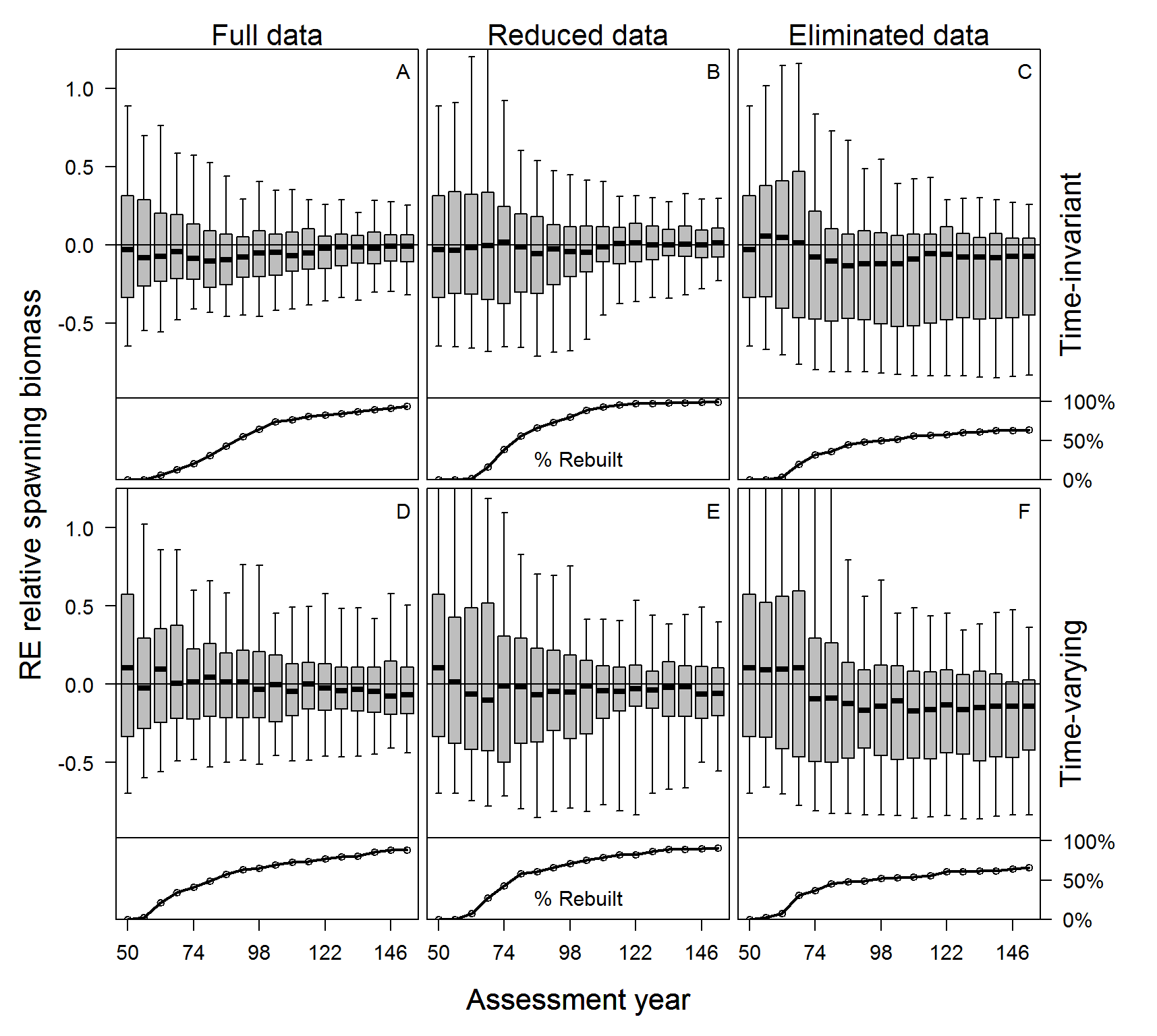


Figure Supp.3. Relative error of estimated relative spawning biomass in each assessment year for each case and data scenario for all simulations (top panel) and the percentage of stocks that had rebuilt to the target biomass during the management period (bottom panel), with data collection consequently returning to historical levels. The median is denoted by the black lines, the grey boxes cover the 25-75% simulation interval, and the boxplot whiskers cover the 95% simulation interval for each assessment year.

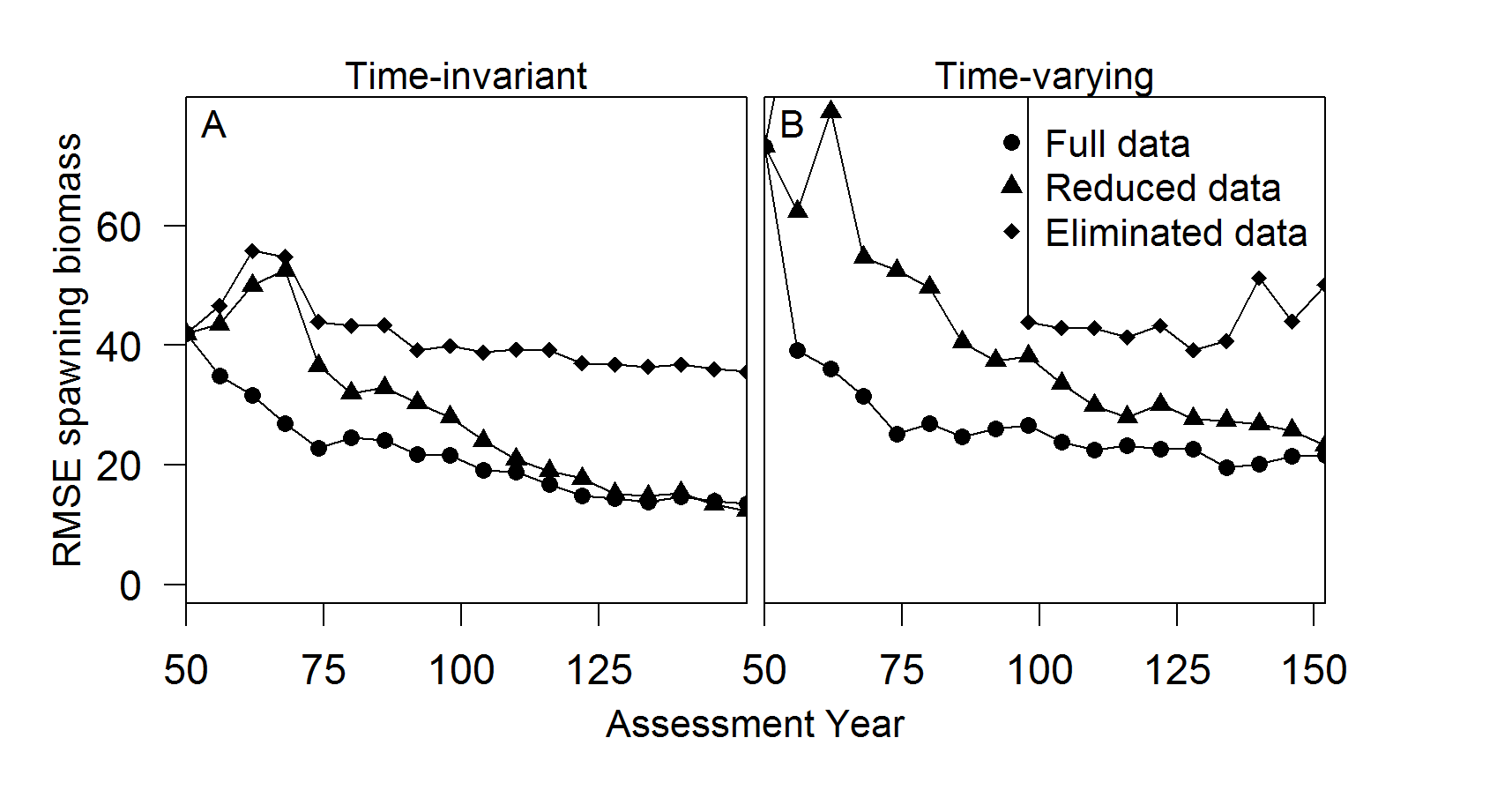


Figure Supp.4 The root mean square error about relative spawning biomass in the assessment year for each case and data scenario.