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

Chantel R. Wetzel1,2, André Punt2

1Northwest Fisheries Science Center, National Marine Fisheries Service, 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

*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 United States, federally-managed stocks that fall below a minimum stock size threshold 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. A rebuilding stock should have harvests limited to levels below the replacement yield. However, being able to measure the rate of recovery is still crucial to management, and increased uncertainty due to limited data can impede this. Additionally, biological data are critical to improve estimates of key parameters within stock assessments (e.g. natural mortality, growth, steepness) and can indicate incoming poor or strong year-classes (recruitment), which can impact estimates of relative stock 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 on the U.S. west coast, catches of yelloweye rockfish were large and unsustainable during the 1980s and early 1990s. Catches of yelloweye rockfish were cut dramatically following the overfished declaration relative to the historical catches, where the allowable catch in the first year of rebuilding was reduced to approximately 10% of the catch observed four years earlier (Stewart *et al*., 2009).

Many species of rockfish are not sampled well by main fishery-independent survey off the U.S. west coast either due to the survey’s inability to sample rocky habitat using trawl gear or other restrictions on sampling locations (e.g. rockfish conservation areas or near-shore habitat). Yelloweye rockfish fall into the category of poorly sampled rockfish by the fishery-independent survey, resulting in the majority of historical information (e.g. index of abundance, length, and age data) available for assessment, albeit somewhat limited, coming primarily from recreational and commercial fishery samples. During rebuilding, retention has been prohibited in the recreational fishery, and the limited allowed catch has led to an avoidance behavior by the commercial fishery (Stewart *et al.*, 2009). The most recent 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). The limited fishery-independent data, along with the reduction of fishery data during rebuilding, presents a challenge for assessment and 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). Studies often focus on the ability to estimate either management quantities or biological parameters. However, to date, these studies have not addressed the long-term impact of reduced data on the ability to monitor a stock during rebuilding. This paper simulates 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 management target stock size (i.e. it is rebuilt), 2) are the limited data from the fishery able to detect a change 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 U.S. west coast was simulated (Table 1), based on the life history and fishery for yelloweye rockfish. Yelloweye rockfish are assumed to have very slow population dynamics based on very low natural mortality and recruitment compensation (steepness). However, the operating model was parameterized using higher natural mortality and steepness values to allow for shorter recovery periods (< 100 years) for computation 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-invariant versus time-variation in natural mortality and fishery selectivity. The first case, “time-invariant”, involved a fixed natural mortality over-time and a fixed fishery selectivity curve during the historical period, the overfished period, and after the stock rebuilt to target biomass levels (Fig. 1a). The historical period involved asymptotic fishery selectivity. The simulated stocks were reduced to an overfished state (below minimum stock size threshold) at the time of the first assessment in year 50. Subsequent to the stock being estimated to be overfished, the fishery selectivity shifted to dome-shaped selectivity while the stock rebuilt to the target biomass to reflect a change in fisher behavior due to harvest restrictions induced by an overfished declaration (e.g. avoidance behavior, closed-areas) (Fig. 1b). The fishery selectivity returned to asymptotic after the stock was estimated to be rebuilt.

The second case, “time-varying”, involved autocorrelated annual deviations in natural mortality and uncorrelated deviations in the parameters on which the fishery selection pattern was based during the historical, overfished, and rebuilt periods (Figs. 1c and 1d). Annual deviations in selectivity were applied to two fishery selectivity parameters: 1) the length at which maximum selectivity occurred, and 2) the width at maximum selectivity prior to the reduction in selectivity with increasing length during the years the stock was overfished. A standard error of 0.50 was applied annually about the size at maximum selectivity for all years (Fig. 1c) and a standard error of 0.20 was applied for the width at maximum selectivity (Fig. 1d).

The operating model was based on a single-sex age-structured model. An annual fishery catch-per-unit effort (CPUE) was observed with error and length- and age-composition data were collected for selected years, and used by the estimation method to estimate population size and a corresponding catch level. The catches were then removed without error from the simulated stock. The 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 of 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  is the number of fish of age *a* at the start of the year *t*,  is the number of age-0 animals at the start of year *t*,  is the selectivity during year *t* for animals of age *a*, *A* is the plus group,  is the instantaneous fishing mortality rate during year *t*, and *Mt* is the instantaneous rate of natural mortality during year *t*.

Autocorrelated annual deviations in natural mortality are defined as:

 (4)

where *M* is the mean value of natural mortality,  is the standard error of the annual variations 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,  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 an index of fishery catch-per-unit-effort (CPUE) for each year *t*:



where *Q*is the catchability coefficient,  is the standard deviation of catchability in log space, and the  expected vulnerable biomass defined as:

 (4)

where *wa* is the weight of a fish at 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 asymptotic (Fig. 1). The fishery selectivity shifted to a domed-shaped (compared to the historical asymptotic) 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 selectivity depended on the estimated status of the stock 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 of U.S. west coast groundfish varies as a consequence of commercial importance, which is used as 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.

*R0*, steepness, growth, annual recruitment deviations, initial age-structure deviations, the size at maximum selectivity, and the width at the maximum selectivity for the fishery were estimated. The relative stock biomass in the assessment year was estimated and the forecasted catches were determined using the Pacific Fishery Management Council (PFMC) rockfish harvest control rule. The catches were removed from the operating population without error, fishery CPUE and composition data were generated for the subsequent six years.

The PFMC harvest control rule applies a linear reduction in catch when a stock falls below 0.40*SB*0, with no fishing when the stock falls 0.10*SB*0. The overfishing level (OFL) was set equal to the target spawner-per-recruit harvest rate multiplied by *SBt*. The OFL was reduced by a management buffer value that accounts for the uncertainty about current biomass for well-assessed stocks (0.956; Ralston *et al.*, 2011) to determine the acceptable biological catch levelm ABC (i.e. ABC = 0.956 OFL). The annual catch limit was set equal to the ABC when the stock was above 0.40*SB*0 or reduced from the ABC according to the harvest control rule when the stock fell below 0.40*SB*0.

One major difference in this simulation design and actual management practice on the U.S. west coast was the omission of the rebuilding plans that are implemented when a stock is assessed to have fallen below a minimum stock size threshold (MSST, defined as 0.25*SB*­0 for U.S. west coast rockfish). Harvest for stocks below the MSST would no longer be determined based on the standard harvest control rule, but rather a rebuilding plan would be used to determine catches until the stock rebuilt to the target biomass. To reduce simulation complexity, a rebuilding plan was not imposed if a simulated stock was assessed to have fallen below the threshold and the catches were determined based on the harvest control rule for all estimated stock sizes.

## 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 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 data scenarios have different data availability based on stock status (e.g. overfished vs. rebuilt) in the assessment year relative to the management targets. The simulated stocks were in an overfished state when the first assessment was conducted and were classified as overfished until the relative spawning biomass surpassed the target biomass (0.40*SB*0), at which time they were declared rebuilt.

Data quantity available for assessments differed depending on stock status following the first assessment among the three data scenarios. 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 reduced 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 above the target biomass. The “eliminated data” scenario had no fishery data during rebuilding (Fig. 2). The fishery CPUE index and composition data resumes at historical samples sizes when the stock was estimated to be rebuilt.

The operating model applied a shift from asymptotic to dome-shaped fishery selectivity when the stock was under rebuilding. The estimation method in the full and reduced data scenarios were allowed to estimate a potential change in selectivity during the rebuilding period. However, the eliminated data scenario was forced to assume constant asymptotic selectivity over all years because no fishery composition data were available to detect the shift in selectivity.

## Performance measures

The outcomes of the simulations for each harvest control rule were summarized using five metrics that were selected to evaluate the impact of data and 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* if the true value from the operating model.

1. The 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)

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 spawning biomass and relative spawning biomass was negatively biased while the stock rebuilt for both scenarios (Figs. 3a-b and 4a-b). As expected, the estimates of spawning biomass and relative spawning biomass for the full data scenario were less variable during the rebuilding period compared to the reduced data and eliminated data scenarios (Figs. 3a-c and 4a-c). However, by the end of the management period, the variance of estimated spawning biomass and relative spawning biomass were similar among the full and reduced data scenarios.

The eliminate data scenario resulted in median estimates of spawning biomass and relative spawning biomass with limited bias but were highly imprecise across the start of the management period (years 50-74) (Figs. 3c and 4c). As stocks begin to be estimated rebuilt and data collection resumed, the median estimates become negatively biased and remained imprecise. 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 precision by the end of the management period (Figs. 3c and 4c).

The eliminated data scenario had the lowest percentage of stocks that were estimated to rebuild to the target biomass during the projection period, resulting in a mixture of simulations that had resumed data collection after rebuilding and simulations still estimated to be rebuilding with only historical data were present. The estimated spawning biomass and relative spawning biomass were strongly negatively biased in the first assessment year (Figs. 3a-c and 4a-c shown in red) for the 35 simulations that were estimated not to be rebuilt by the end of the management period. The biased estimates were driven by negatively biased estimates of steepness in the first assessment (Fig. 5a-c). In the absence of new data, the estimates of the steepness did not improve over the management period. However, with either full or reduced data were present, estimated quantities for this subset of simulations were unbiased by the end of the management period (Figs. 5a-b and 5a-b shown in red). In contrast, the simulations that were estimated to have successfully rebuilt for the eliminate data scenario resulted in estimated steepness values that were positively biased relative to the operating model (Figs. 5a-c and 4a-c shown in green).

The RMSE, a measure of both bias and precision, 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. 6a). However, the eliminate data scenario resulted in the highest RMSE over the entire management period and fails to substantially reduce when data collection resumes for the recovered simulations (Fig. 6a). The lack of improvement in the RMSE for the eliminated data scenario was driven by the number of simulations that were never estimated to have recovered (35 out of 100 simulations) and dependent upon only the historical data (Table 2).

The median number of years estimated for the stocks to recover to the target biomass for the full data scenario was longer the median recovery year within the operating model (Table 2). In contrast, both the reduced and eliminated data scenarios estimated a shorter median recovery time 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 precision of estimates. The full data scenario estimated negatively biased relative spawning biomass with high precision across assessment years, resulting in estimates that predicted constant rebuilding but at a slower rate than the true stock (Fig. 4a-c). The reduced data scenario had a higher level of variance among estimates of the relative spawning biomass (Fig. 4b). The variability of estimates between assessments resulted in stocks being estimated recovered earlier than the operating model due to estimation error driven by the limited composition samples during rebuilding.

The variability in estimated stock status among assessment years resulted in the reduced data scenario having the lowest average catch during rebuilding (Table 3). The reduced data scenario would set catches during the early rebuilding years at appropriately low levels that would estimate rebuilding but assessment error increased as time progressed with limited data, with stocks being estimated recovered earlier than the true population. The eliminated data scenario which was entirely dependent upon historical data, essentially projecting the population forward with each assessment based on the initial parameter estimates, resulting in high average catches during rebuilding and the lowest AAV during rebuilding and across the entire management period (Table 3).

Reduction or elimination of data during rebuilding increased the variance about estimates of the size at maximum fishery selectivity, with the estimates generally equal to the true value in median terms for all data scenarios (Fig. 7a-c). The precision of the estimates for the reduced and eliminated scenarios only improved at the end of the management period, when the majority of the stocks had rebuilt, and fishery composition sample sizes returned to the higher historical levels (Fig. 6b-c). 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 positively biased estimates that were highly variable at start of the management period (Fig. 8a-b). The positively biased estimates 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 (a higher estimated value implies a less severe dome in selectivity, e.g. decreased slope for the selectivity descending limb) during rebuilding. The full data scenario resulted in markedly improved estimates in the shape of the dome over the management period, especially compared to the reduced data scenario (Fig. 8a-b).

## The impact of time-varying parameters

The case that assumed time-varying natural mortality and fishery selectivity generally resulted in increased variation in estimates compared to the time-invariant case. The initial estimates of spawning biomass at the time of the first assessment were positively biased and highly variable (Fig. 3d-f). The variance about spawning biomass decreased markedly for the full data scenario after the first assessment in year 50 (Fig. 3d). However, the variance of estimated spawning biomass remained high for approximately the first twenty-five years of the management period for both the reduced data 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 negatively biased (Fig. 3c-e). However, the median relative spawning biomasses were variable over the management period (Fig. 4c-e).

The estimates of relative spawning biomass for eliminated data scenario were positively biased at the start of the management period, but became negatively biased as stocks rebuilt to target biomass and data collection resumed (Fig. 4f). Similar to the results in the time-invariant case, the simulations that failed to be estimated rebuilt by the eliminated data scenario (32 simulations), had negatively biased estimates of spawning biomass and relative spawning biomass at the time of the first assessment (Figs. 3d-f and 4d-f shown in red) which were driven by negatively biased estimates of steepness (Fig. 5d-f shown in red) that did not improve in the absence of new data.

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. 6). 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. 6b). Similar to the time-invariant results, the RMSE of relative spawning biomass for the eliminate data scenario was the highest among the scenarios across the entire management period.

The inclusion of time-varying parameters in the operating model resulted in shorter estimated recovery time relative to the time-invariant case for all the data scenarios (except the eliminated scenario which were equal) (Table 2). However, the median number of rebuilding years 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 of estimates of the stock status between assessments (estimation error) (Fig. 4).

The median average catch obtained during the recovery period was the highest for the eliminated data scenario due to the positively biased estimates of the relative biomass at the start of the management period (Table 3, Fig. 4f). Additionally, the eliminated data scenario had the lowest median AAV during the overfished period (Table 3). In contrast, the eliminated data scenario, resulted in the highest number of stocks that were never estimated to exceed the target biomass (Table 3).

Inclusion of time-varying selectivity resulted in a persistent positive bias for the estimated size at maximum selectivity across all data scenarios (Fig. 7d-f), although the full data scenario resulted in the lowest variation. The full data and the reduced data scenarios which were allowed to estimate dome-shaped selectivity (width at maximum selectivity) during the recovery period, resulted in highly imprecise estimates at the start of the management, period with the precision of the estimates improving earlier for the full data scenario (Fig. 8c-d).

# Discussion

Maintaining fishery data at historical levels during rebuilding improved the predictability of estimates between assessments. While the full data scenario had an increase in precision, the estimates of spawning biomass and stock status were consistently negatively biased 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 is survery that provides an index and length and age-composition data determined that this negative bias was eliminated if survey composition data were available along with fishery composition data. The observed bias was driven by two key factors; the fishery selectivity curve and data quantity. The fishery selectivity curve was specified to be greater than the maturity curve, resulting in only mature larger fish being selected. The fishery data were informative about recruitment, but with a lag between recruitment to population and being observed in the fishery. However, a fishery independent survey that selected fish at smaller sizes yields information about recruitment earlier. Additionally, increased additional composition data from multiple data sources that are consistent improve estimates of recruitment, spawning biomass, and relative stock size (Yin and Sampson, 2004; Wetzel and Punt, 2011).

The negative bias in estimated relative spawning biomass by the full data scenario resulted in longer estimated rebuilding periods than required, an outcome that would produce extended harvest restrictions that were not warranted, a situation fishery managers would like to avoid. However, the reduced estimation variability 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 serious negative consequences for fisheries management. This is because overly optimistic estimates of stock size can result in overfishing when catch limits are set too high, reducing biomass further, potentially leading to a subsequent overfished declaration by a future assessment.

The high number of simulations that failed to be estimated rebuilt by the eliminated data scenario highlighted the importance of continued data collection during rebuilding. In the absence of new data, the first and subsequent assessments are entirely dependent on the quality of the historical data to inform parameter estimates. The simulations that failed to be estimated rebuilt were driven by negatively biased estimates of steepness at the time of the first assessment. Estimating a stock less productive than the true population resulted in reduced estimates of spawning biomass and relative spawning biomass, with assessment setting harvest at levels well below the true ABC. 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) 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 with limited data may not to 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 of steepness.

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 estimates were more variable across all data scenarios. The estimation method assumed a single natural mortality across all years equal to the mean value which 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 natural 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 estimated a positively biased 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 of the size at maximum selectivity. However, the estimation method was able to identify the change in the selectivity form (asymptotic to domed through a reduction in the width at peak selectivity) during the rebuilding years with similar bias to the time-invariant case. Each case led to estimates that under estimated the decline in the descending limb, defining the dome in selectivity. This evaluation applied time blocks defined by the status of the stock to allow for shifts in selectivity, ignoring the annual shifts 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 here and how data quantity and quality impacts these estimates.

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) that will allow for better approximations to reality through models. Specifically, when stocks are driven to depleted stock sizes there could be long-term impacts to the population that could negatively affect 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 due, for example, to 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 (Legaut and Palmer, 2015) that needs to be accounted for when setting harvest limits.

The generally well behaved results from each of the data scenarios were related to some key assumptions and choices. The strategic decision was made to construct the operating model population with the estimation method with similar assumptions concerning critical biological and fishery properties (e.g. stock-recruitment relationship) to focus on the role of data in the ability to monitor a rebuilding stock rather than the results created by various model misspecifications. Additionally, simulating the messiness that is inherent in fisheries data through observation and process error alone is insufficient. The patterns and trends in real fisheries data are likely caused by highly complex processes that are not yet completely understood and hence difficult to recreate through simulation. The benefits of data quality and quantity is likely underestimated here relative to real-world assessments which require informative data to identify trends in biomass, population scale, biology, and recruitment.

# References

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

Conn, P.B., Williams, E.H., Shertzer, K.W. 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., Sogard, S.M. 2014. BOFFFs: on the importance of conserving oldh-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 Peck, M. A. 2011. Effects of climate change on fish and fisheries: forecasting impacts, assessing ecosystem responses, and evaluating management strategies. ICES Journal of Marine Science, 68: 984–985.

Ianelli, J. N., Hollowed, A. B., Haynie, A. C., Muter, F. J., and Bond, N. A. 2011. Evaluating management strategies for eastern Bering Sea walleye pollock (*Theragra chalcogramma*) in a changing environment. ICES Journal of Marine Science, 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., Punt, A.E. 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 Palmer, M. C. 2015. In what direction should the fishing mortality target change when natural mortality increases within an assessment? Canadian Journal of Fisheries and Aquatic Sciences, 73: 349–357.

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

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

Martell, S., Stewart, I. 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., Piner, K., 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., Wetzel, C.R., 2013. Stock Synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142, 86-99.

Nielsen, A., Berg, C.W. 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., Conser, R.J. 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., *et al*. 2014. Linking Northeast Pacific recruitment synchrony to environmental variability. Fisheries Oceanography, 23: 389–408.

Stewart, I.J., Wallace, J.R., McGilliard, C., 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 Benoit, H. P. 2015. Extreme increases in natural mortality prevent recovery of collapsed fish populations in a Northwest Atlantic ecosystem. Marine Ecology Progress Series, 519: 165–182.

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

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

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

# Tables

Table 1. Life-history and observation parameters used in the operating 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 Solved |

Table 2. The median and 90% simulation interval of the estimated number of years needed to recover to target biomass, the operating model number of years needed to recover, 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 data scenario and case. The median recovery times (estimated and operating model) were based on the stocks that successfully rebuilt to the target biomass.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Selectivity / data scenario | Estimated num. of rebuilding years | | Operating model num. of rebuilding years | | Num. of stocks that failed to recover | |
|  | 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 of the average catch during rebuilding, the AAV during rebuilding, and the AAV over all years for each data scenario and case.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 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

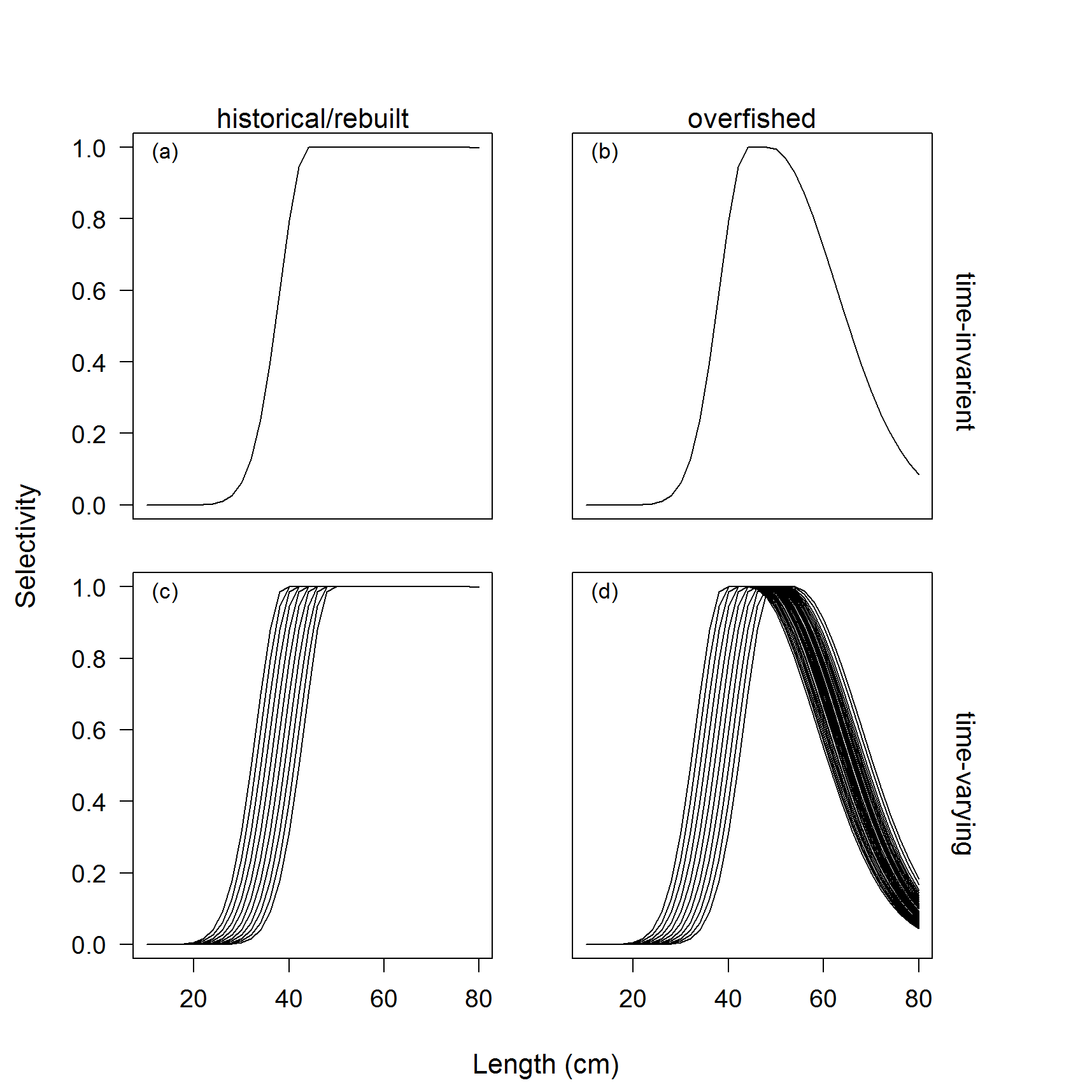


Figure 1. The fishery selectivity for either time-invariant or time-varying selectivity for the historical/rebuilt and overfished periods. A standard error of 0.50 was applied annually about the size at maximum selectivityand a standard error of 0.20 was applied for the width at maximum selectivity while the stock was estimated overfished (c) and (d) (see Methot and Wetzel, 2013 for additional details on double normal selectivity).

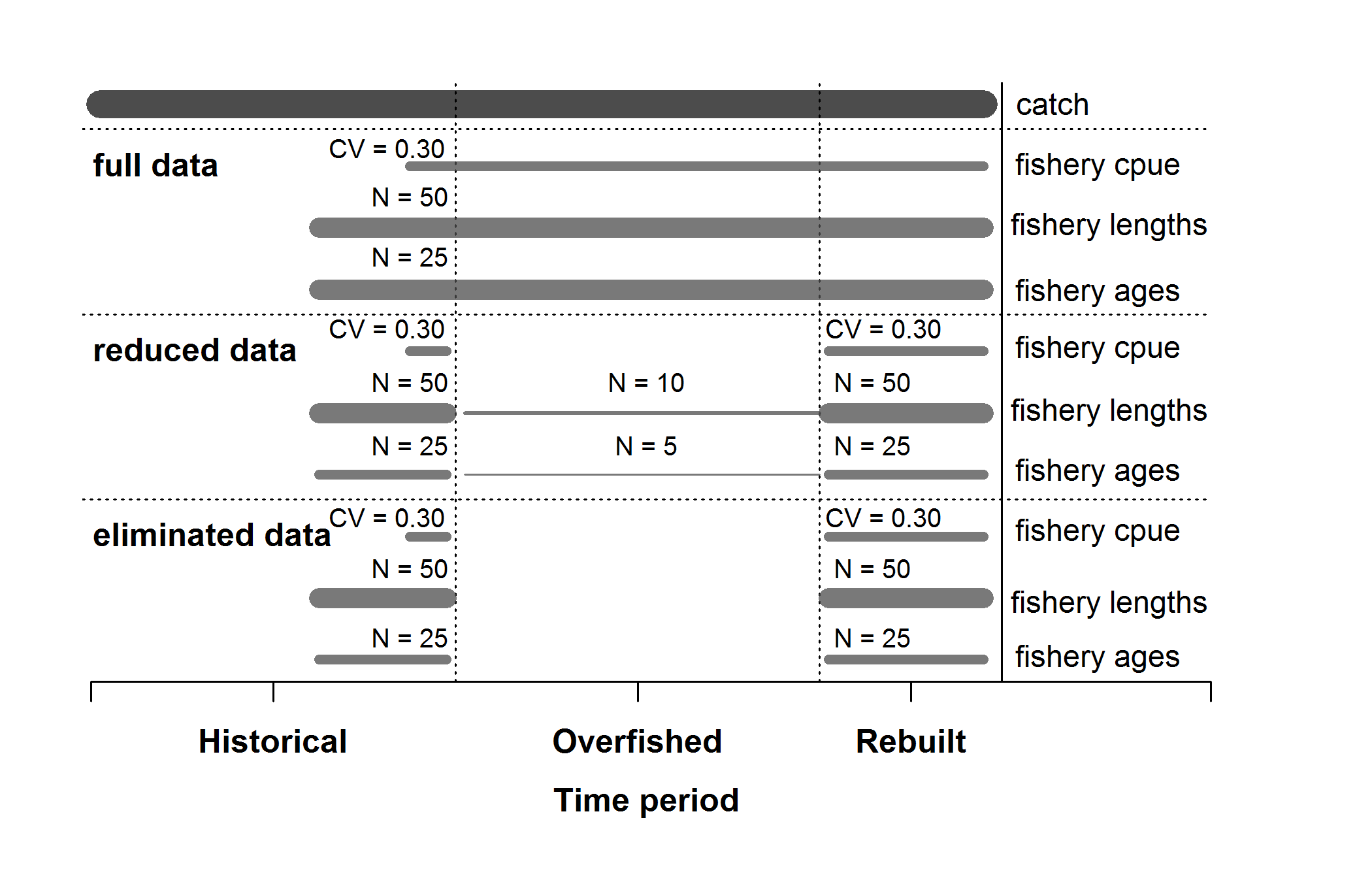


Figure 2. Summary of the data available for each of the data scenarios. Catches were known without error and were available across all data scenarios.

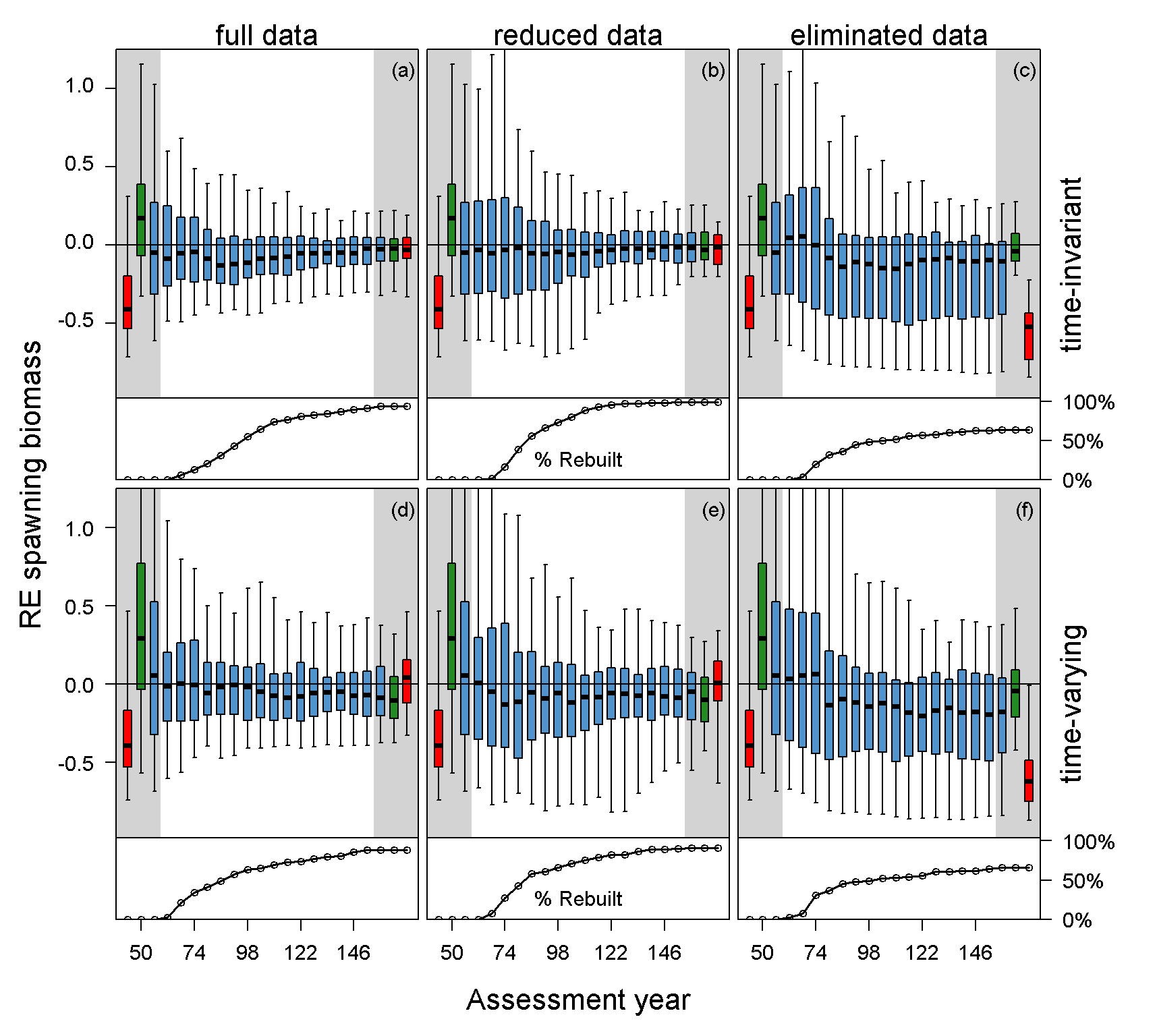


Figure 3. The relative error of estimated spawning biomass in each assessment year for each data scenario and case summarized in each assessment year for all simulations (top of each panels: blue bars) and the percentage of stocks that had rebuilt over time, with data collection consequently returning to historical levels, shown along the bottom of each panel. The eliminated data scenario resulted in a large portion of stocks that failed to rebuild (time-invariant: 35 out of 100, and time-varying: 32 out of 100). The relative error of estimated spawning biomass for the simulations that either successfully (green) or failed to rebuild (red) under the eliminated data scenario were plotted for each data scenario dependent upon the case (time-invariant vs. time-varying) in the first assessment and final assessment. The grey area indicates the subset results for the first and last assessment by simulation that either failed or were estimated rebuilt by the eliminated data scenario and all simulations combined. The median is denoted by the black line and the boxplot whiskers cover the 95% simulation interval for each assessment year.

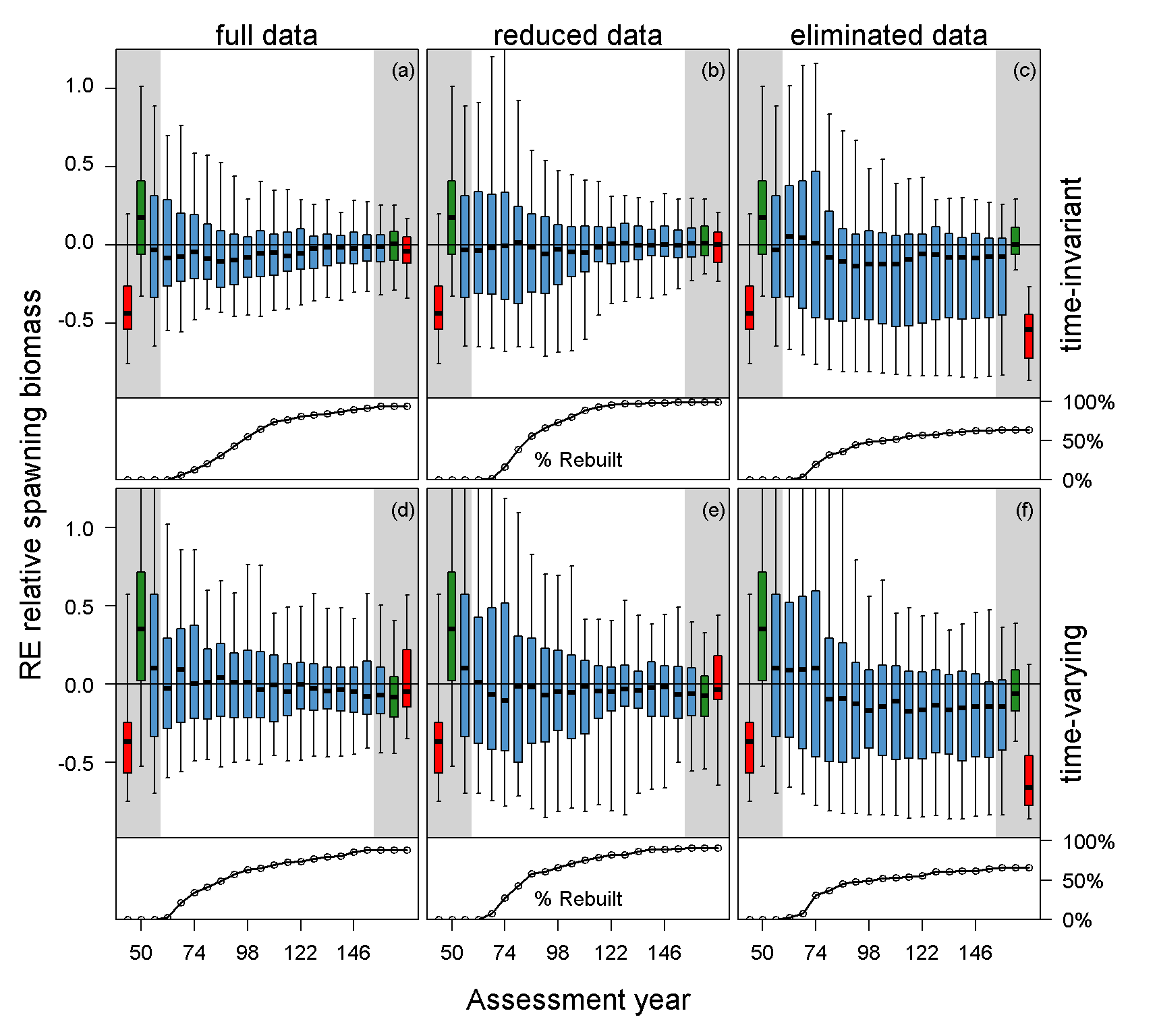


Figure 4. The relative error of estimated relative spawning biomass in each assessment year for each data scenario and case summarized in each assessment year for all simulations (top of each panels: blue bars) and the percentage of stocks that had rebuilt over time, with data collection consequently returning to historical levels, shown along the bottom of each panel. The eliminated data scenario resulted in a large portion of stocks that failed to rebuild (time-invariant: 35 out of 100, and time-varying: 32 out of 100). The relative error of estimated relative spawning biomass for the simulations that either successfully (green) or failed to rebuild (red) under the eliminated data scenario were plotted for each data scenario dependent upon the case (time-invariant vs. time-varying) in the first assessment and final assessment. The grey area indicates the subset results for the first and last assessment by simulation that either failed or were estimated rebuilt by the eliminated data scenario and all simulations combined. The median is denoted by the black line and the boxplot whiskers cover the 95% simulation interval for each assessment year.

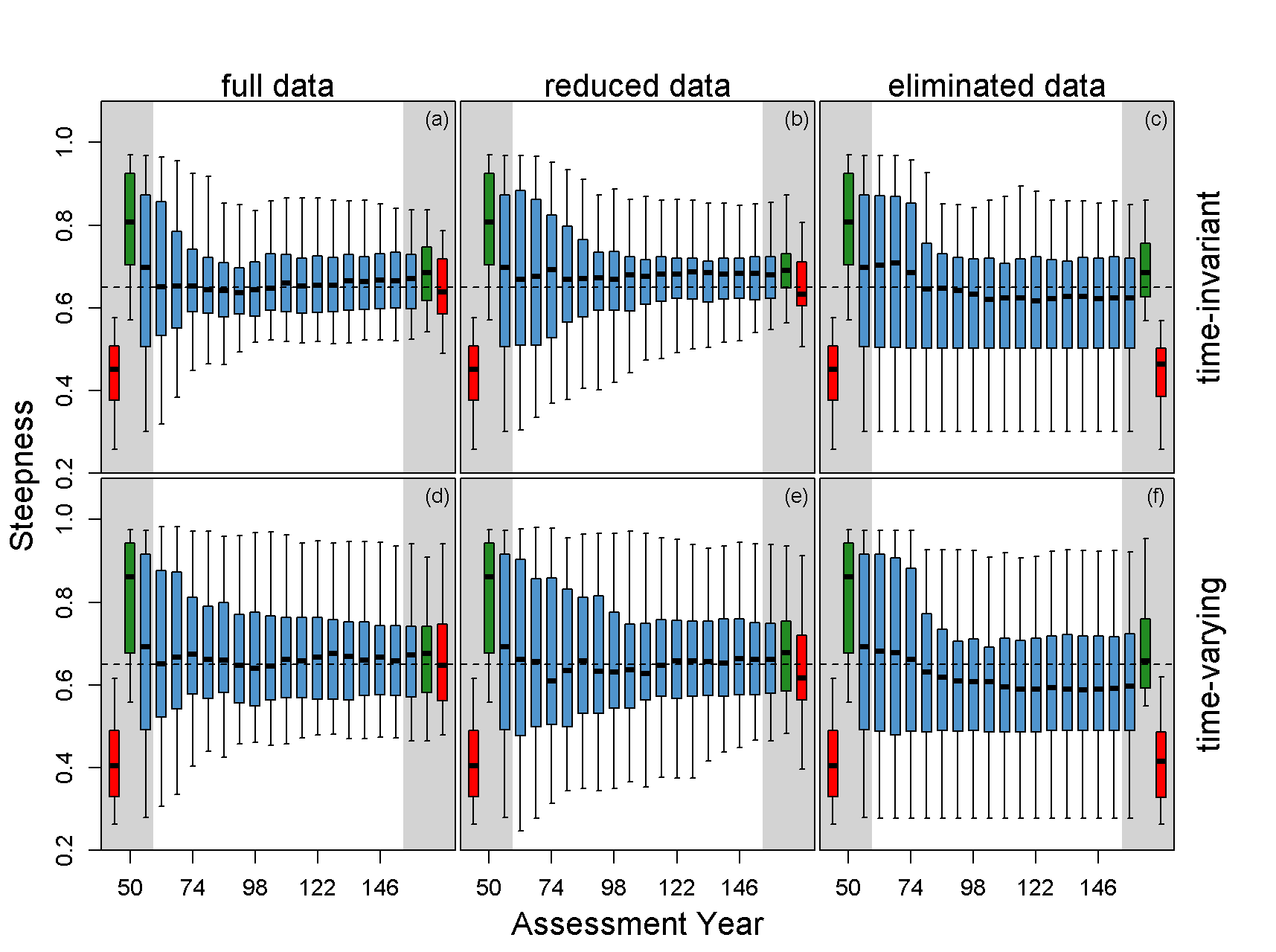


Figure 5. The estimate of steepness (*h*)in each assessment year for each data scenario and case summarized in each assessment year for all simulations (blue bars). The eliminated data scenario resulted in a large portion of stocks that failed to rebuild (time-invariant: 35 out of 100, and time-varying: 32 out of 100). The estimates of steepness for the simulations that either successfully (green) or failed to rebuild (red) under the eliminated data scenario were plotted for each data scenario dependent upon the case (time-invariant vs. time-varying) in the first assessment and final assessment. The grey area indicates the subset of results for the first and last assessment by simulation that either failed or were estimated rebuilt by the eliminated data scenario and all simulations combined. The median is denoted by the black line and the boxplot whiskers cover the 95% simulation interval for each assessment year.

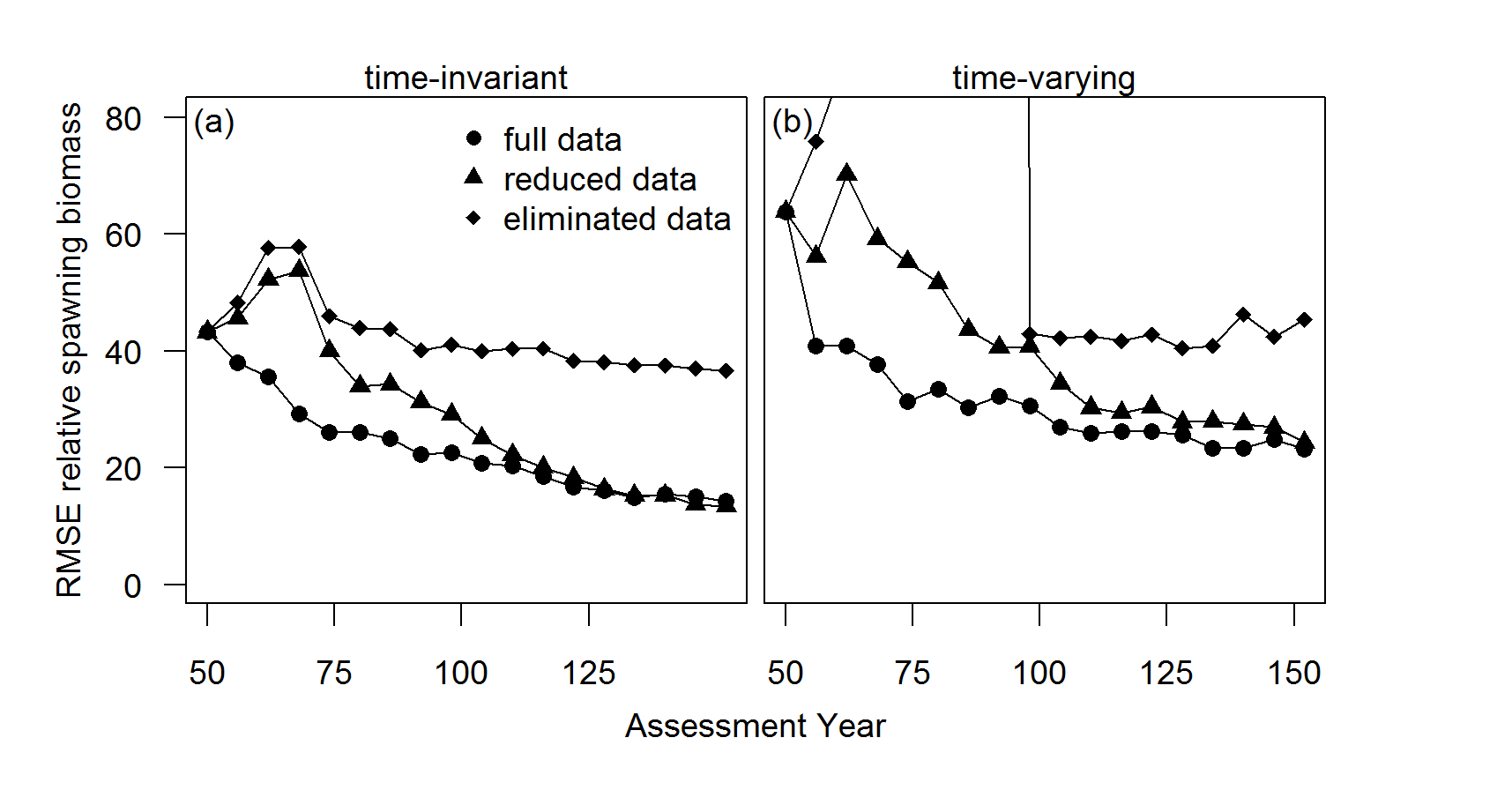


Figure 6. The root mean square error about relative spawning biomass in the assessment year for each data scenario and case.

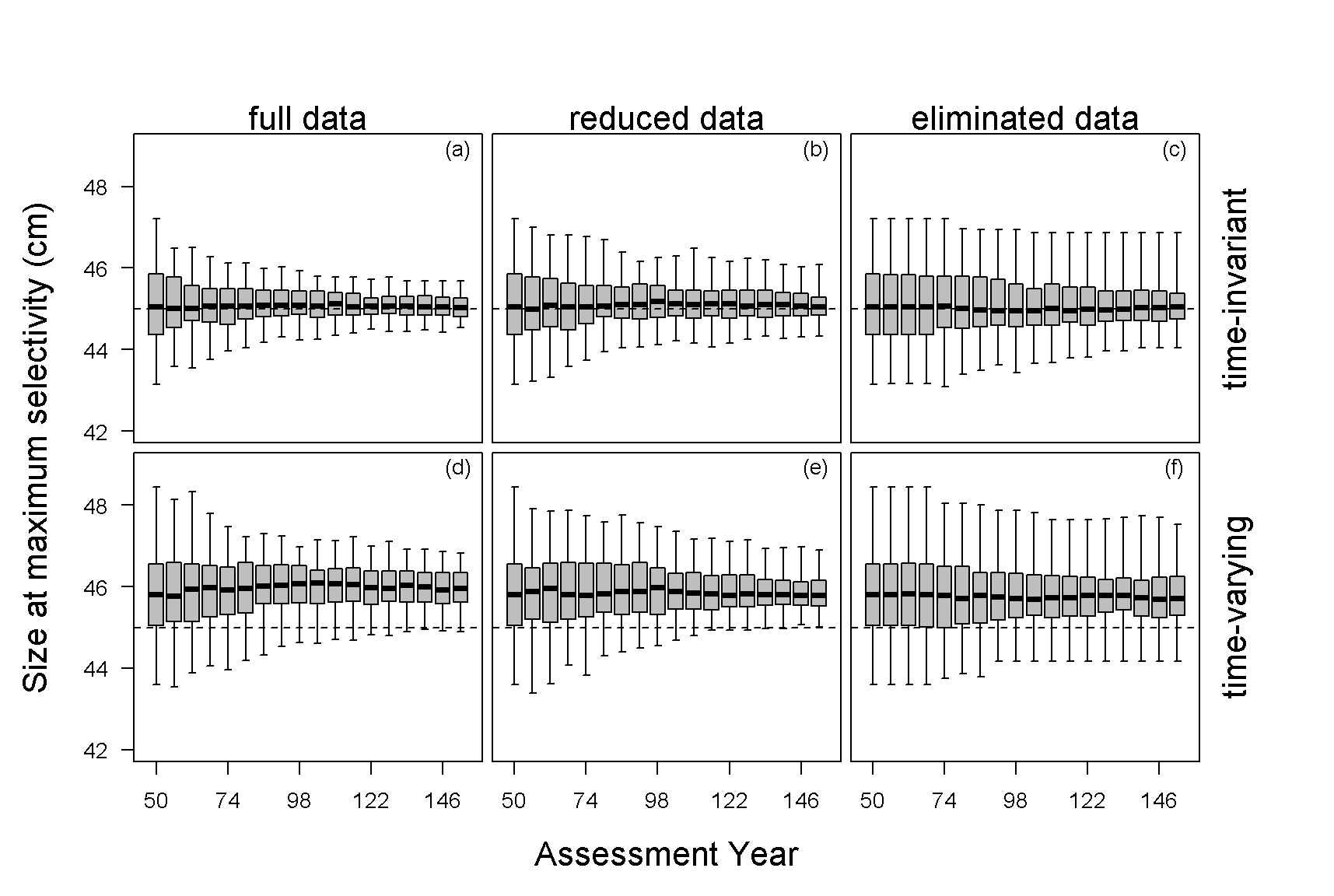


Figure 7. The estimated size at maximum fishery selectivity for each data scenario and case. The black dashed line indicates the operating model parameter value. The estimates from the data scenarios with as time progressed with limited data selectivity were compared against the mean of the distribution from the operating model. The median is denoted by the black line (45cm) and the boxplot whiskers cover the 95% simulation interval for each assessment year.

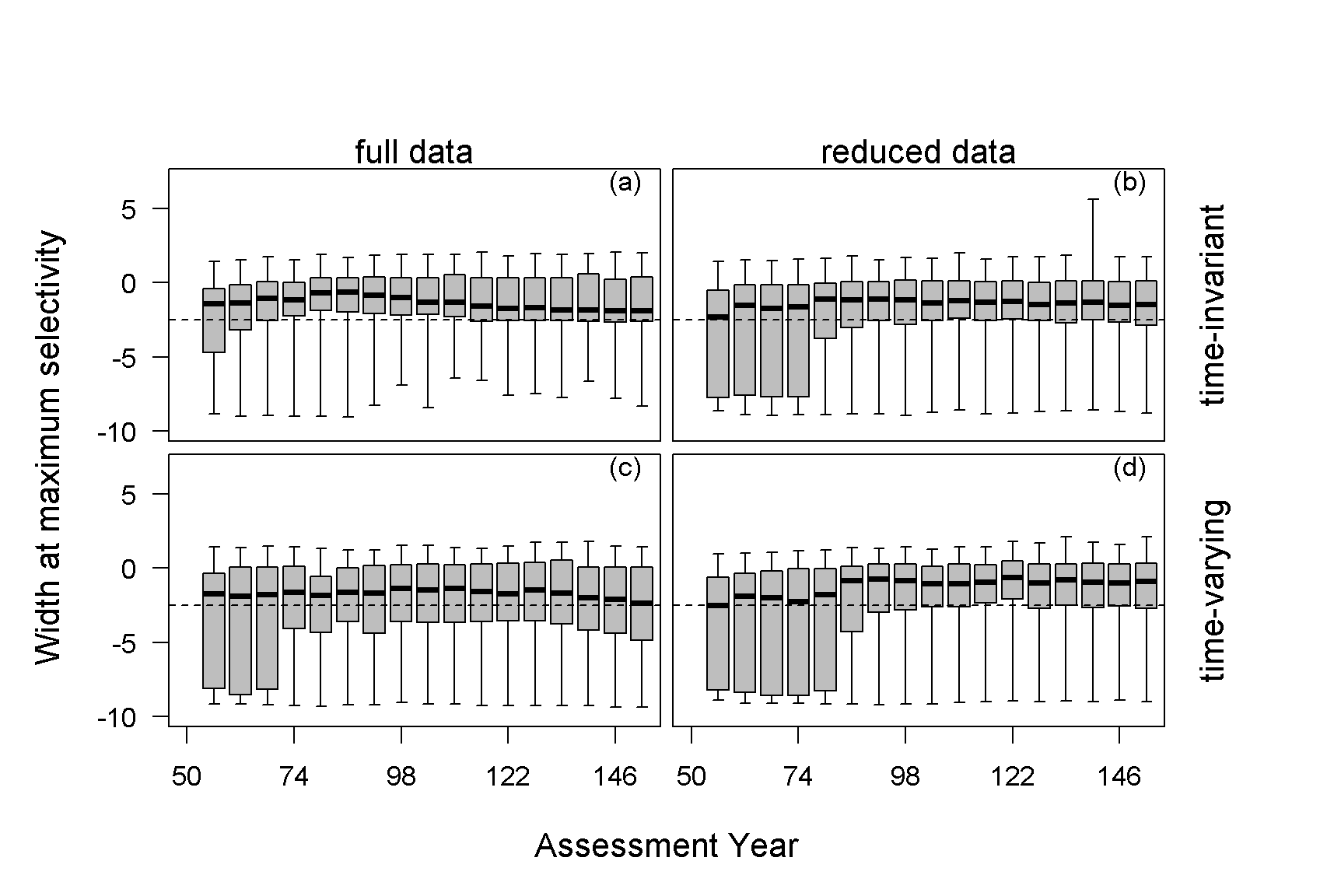


Figure 8. The estimated width at maximum selectivity (i.e. defines the extent of dome-shaped selectivity) for the fishery that occurred while the stock was overfished for each data scenario and case. The black dashed line indicates the operating model parameter value. The estimates from the data scenarios with as time progressed with limited data selectivity were compared against the mean of the distribution from the operating model. The median is denoted by the black line and the boxplot whiskers cover the 95% simulation interval for each assessment year.