WIDOW ROCKFISH

STAR Panel Report

July 13-17, 2009

Southwest Fisheries Science Center
110 Shaffer Road, Santa Cruz, CA 95060

Panel Reviewers

Martin Dorn	Panel Chair, Scientific and Statistical Committee (SSC) Representative
Chris Francis	Center for Independent Experts (CIE)
Vladlena Gertseva	NMFS, Northwest Fisheries Science Center (NWFSC)
J. J. Maguire	Center for Independent Experts (CIE)

Panel Advisors

John DeVore	Pacific Fishery Management Council
Gerry Richter	Groundfish Advisory Subpanel (GAP) Representative
John Budrick	Groundfish Management Team (GMT) Representative

Stock Assessment (STAT) Team members present

Xi He	NMFS, Southwest Fisheries Science Center (SWFSC)
Alec MacCall	NMFS, Southwest Fisheries Science Center (SWFSC)
Stephen Ralston	NMFS, Southwest Fisheries Science Center (SWFSC)
John Field	NMFS, Southwest Fisheries Science Center (SWFSC)
E. J. Dick	NMFS, Southwest Fisheries Science Center (SWFSC)

Overview

A draft assessment of widow rockfish (*Sebastes entomelas*) off the United States West Coast was reviewed by the STAR panel during July 13-17, 2009. The assessment assumes a single coastwide stock given that there is currently no biological or genetic information to suggest the presence of multiple stocks. Northern and southern areas are modeled with different growth in each area. Recruits are produced from a common pool of spawners, and are assigned to each area according to year-invariant proportion (estimated by the model). No mixing is assumed to occur post recruitment.

This assessment uses the Stock Synthesis platform (version 3.03a) and incorporates a variety of data sources into the base model. Catch and age-frequency data from four fisheries are used in the assessment, including Washington and California fisheries (all gears combined), and Oregon midwater and bottom trawl fisheries. A fishery-dependent relative abundance (CPUE) index calculated from Oregon trawl fishery data was used in the model, and three bycatch indices were estimated from foreign, joint-venture and domestic Pacific whiting fisheries. Fishery-independent data included the midwater trawl pelagic juvenile survey, the triennial trawl survey and the NWFSC shelf-slope survey. In addition to survey indices, length-frequency data were available from the triennial and NWFSC shelf-slope surveys, and age data from the NWFSC shelf-slope survey.

The last full assessment of widow rockfish was conducted in 2005, and it was subsequently updated in 2007. Major changes made in this assessment, compared with the previous assessment include:

- Use of SS3 modeling framework instead of previously employed direct ADMB code;
- Extension of period modeled from 1958 to 1916;
- Use of revised catch history based on Lynde (1986), Tagart (1985) and Ralston et al. (2009);
- Addition of NWFSC shelf-slope trawl survey (referred to as NWFSC combo survey in the assessment report);
- Use of revised triennial survey estimates using a GLMM approach (instead of area-swept approach previously used);

The STAR panel concluded that the widow rockfish assessment constitutes the best available scientific information on the status of widow rockfish off the U.S. west coast and recommends that it be used for status determination and management in the Council process. The base model developed during the Panel meeting is considered reasonably well investigated. However further model evaluation and exploration of alternative model configurations is a priority for widow rockfish. There was insufficient time to accomplish this during the STAR Panel meeting because of difficulties in developing a suitable base model. The STAR panel thanks the STAT team members for their hard work and willingness to respond to panel requests.

Analyses requested by the STAR panel

A total of 14 requests were made by the STAR panel. While data and modeling requests were fulfilled in an order different than that in which they were made, data requests are discussed first in the list below, followed by modelling requests. The requests have been grouped and re-ordered for readability.

Data requests

1. **Split Oregon catches for midwater and bottom trawl fisheries prior to 1983 out of the Washington fishery.**
   Rationale: The draft assessment models four fisheries: a) Vancouver – Columbia, b) Oregon midwater trawl, c) Oregon bottom trawl, d) Eureka, Monterey and Conception. Because there was no age composition data prior to 1983, catches for the OR MWT and BT for 1983 and prior years, when widow catches were substantial, were lumped with the Vancouver – Columbia (Washington) catches. This is inaccurate in terms of fishery history and could result in model misspecification. Information is available for years prior to 1983 to partition Oregon catches into bottom trawl and midwater trawl and the panel requested that it be done.
   Response: Available information was used to parse OR MWT and BT out of the Vancouver-Columbia fishery, into OR MWT and BT, with little effects on the modeling results.
2. **Provide new estimates for catches in Vancouver – Columbia for 1916 to 1976.**
   Rationale: Some catches are known to have occurred prior to the late 1970s, but as there were no estimates, the draft assessment assumed that catches in Vancouver – Columbia for 1916 to 1976 were the same as in the Eureka, Monterey and Conception commercial fisheries for the same years. This seemed highly unlikely and the panel requested that real data be used or that catches in those years without information be assumed to be zero.
   Response: Widow rockfish catches were estimated from a spreadsheet prepared by Tagart (1985), total rockfish catches by state from 1938 to 1955 published in the Bulletin of Commercial Fisheries (red books), an Alaska Fisheries Science Center report on catch reconstruction (Lynde 1986) for 1956 to 1980, and on the percent widow rockfish in the total rockfish category. This should be seen as provisional reconstruction that should be re-evaluated in the next assessment. These new landings were assigned to the OR bottom trawl fishery for modeling as opposed to the Vancouver – Columbia fishery. This was not expected to make a large difference, but was felt to correspond more to reality.
3. **Provide discard rates by State and gears from the West Coast observer program for 2002-2008. Review Pikitch data to see if there is information on discard rates by gear (bottom vs midwater trawl). If the discard rates are different, total removals**

should be adjusted accordingly. For foreign fleets, remove the discards as discards are not believed to have occurred. Report what discard rates had been applied to what fishery in what year.

Rationale: It was unclear what discard rates had been used in the draft assessment. The panel asked that what had actually been done be documented and potentially revised using sources of information that might not have been used.

Response: No discard was applied to the foreign fishery, since Rogers (2003) indicates that the large capacity foreign vessels did not discard fish based on either size or species. For all domestic fisheries, discards were assumed to be 6% until 1981, and 19% from 1982-2001 (Pikitch study did not show significant differences in discard rates between bottom and midwater trawl fisheries for widow rockfish). Between 2002 and 2007, total catches (landings plus discard) were reconstructed from West Coast Groundfish Observer Program total mortality estimates and at-sea whiting fishery bycatch estimates, provided by the Northwest Fisheries Science Center. The only 2008 discard information for widow rockfish available at the time of STAR Panel was the discard estimate from the PacFIN quota-species monitoring report, and this value was minimal, therefore for 2008 landing data from PacFIN and the widow bycatch estimate in the at-sea whiting fishery were used, and no discard was added.

4. **Provide a graph of the abundance indices scaled to their own average for a common time period so that they are on a similar scale.**

   Rationale: A close examination of the abundance indices should be standard practice into assessments before any modeling takes place to identify what story the indices are telling if any.

   Response: As this request was not seen as high priority, there was insufficient time during the panel meeting to provide this graph. Such a graph should be included in the final assessment report.

Modeling requests

5. **Reset all effective Ns using Stewart's method with number of trips and number of fish for fisheries; for the surveys, use Stewart's method with the number of positive hauls. Do the re-tuning with effective N multipliers for the composition data sets, rather than adjusting each individual year.**

   Rationale: The draft assessment appears to have tuned each individual year of each data source rather than use a single weight by source of data. This is not appropriate, and single weight per series should be used rather than individual weights for each data point of each data series.

   Response: This was done, but the multipliers were used as new weights rather than as multipliers for the original weights. A new run was requested where the multipliers are used to modify the original weights rather than used as new estimates for the weights. This resulted in more stable model with a higher Bcurrent to B0 ratio (depletion).

6. **Provide a run with one asymptotic length-based selectivity for all surveys, genders and areas combined. Also provide a run with separate asymptotic length-based selectivity curves for each survey, but combining genders and areas.**
   Rationale: The draft assessment estimated selectivities by gender for each of the survey (2), genders (2) and areas (2) for a total of 8 selectivities. The resulting estimates showed strange patterns and were based on few observations.
   Response: Estimating a single selectivity for the 2 surveys, the 2 genders and the 2 areas deteriorated the fit and resulted in unreasonable selectivities. It seemed more sensible to estimate one selectivity per survey with genders combined. This improved the fit, but the results were counterintuitive in that the two survey curves shifted substantially to the left of the combined survey selectivity seen earlier. This indicates the survey selectivities are not well estimated, which is not surprising given the sparseness of the survey length composition data. There could also be issues associated with forcing the survey selectivities to be asymptotic.
   The Panel thought that taking two steps to change these selectivities should have been done one at a time. The panel requested two survey selectivity curves and allowing these curves to be dome-shaped using a double-normal function. These curves were more reasonable and the panel considered this configuration to be suitable for the base case model.
7. **Provide a run with the abundance index for juveniles up-weighted and the age and length compositions down-weighted.**
   Rationale: As a result of management measures, age and length compositions have become increasingly sparse since the early 2000s. In addition, length and age compositions provide only indirect evidence on the size of year-classes while the juvenile index does provide direct evidence.
   Response: As expected, a closer fit to the juvenile abundance index was obtained, and this model appeared to be more stable in comparison to other model configurations.
8. **Provide a new base case using the Francis multipliers applied to the N-multipliers from the last re-weighted run (Reweight 2). For the age compositions, four Francis weights will be used: 1) average WA and EM fisheries = 0.4 , 2) average for the ORM and ORB trawl fisheries = 0.22, and 3) average for north and south NWFSC survey = 0.20. For the length compositions, use 0.25 for the triennial survey and 0.10 for the NWFSC survey.**
   Rationale: The panel considered that the weights given to age and length compositions is too high because of autocorrelation in the compositions. A method was proposed to estimate appropriate weights (see Appendix) and a subjective judgement, based on the results, was made for the weights to be applied to each age and length composition data source.

Response: The panel considered this model to be an improvement given concerns about overweighting age and length composition data. The STAR panel and the STAT team agreed that this model was appropriate as a base case.

9. **For the new base case with the Francis weights, use h=0.30 and h=0.50 to bracket uncertainties. If these provide too narrow a range to bracket the uncertainties, use h=0.25 and h=0.55.**
   Rationale: Steepness is poorly known and changes in steepness results in substantial changes in depletion and recent population trends.
   Response: Runs with the four steepness were provided and the panel concluded that those with steepness of 0.25 and 0.55 were bracketing the uncertainties in this assessment.
10. **Provide spawning output trends for proportion of recruits in northern area at 0.68 and 0.80 and a table of the key parameters. Provide a graph of spawning output as depletion for the two areas to better see what is happening in the two areas for all three scenarios, including base case.**
    Rationale: The panel questioned the estimate of the proportion of recruitment in the north (0.72). The estimated coastwide depletion varies substantially as that proportion of northern recruitment varies, and the depletion was the greatest when the proportion was at the estimated value. The panel requested sensitivity analyses where the recruitment proportion in the north is fixed at lower (0.68) and higher (0.80) values.
    Response: When profiling on proportion of recruits in the north with the final base case (estimated propN = 0.717), steepness was near one when the proportion north was 0.65. The panel asked that the final assessment report provide runs with the proportion north of 0.6875, 0.717 and 0.75 with the requested graphs and key parameters.
11. **Provide a likelihood profile on steepness to be completed after the meeting.**
    Rationale: Steepness has considerable influence on the assessment results, particularly on recruitment and depletion estimates. It is important to know if the steepness used in the assessment is optimal in terms of likelihood.
    Response: Because of the time required to sort out the landings and discard data, and the time devoted to finding an acceptable model configuration, there was insufficient time during the panel meeting to complete this request. This request should be addressed in the final assessment report.
12. **Examine the contribution of age 1 male of OR MWT in 2007 to the total likelihood.**
    Rationale: This observation had a large residual and the panel was concerned that it could have a large influence on the total likelihood.
    Response: Because of the time required to sort out the landings and discard data, and the time devoted to finding an acceptable model configuration, there was insufficient time during the panel meeting to complete this request. This request should be addressed in the final assessment report.

13. **Slide 26 of the original widow rockfish presentation on July 14 shows stacked fishing mortalities by area. Provide different graphs for northern and southern areas.**
    Rationale: These data were from a two area model and the fishing mortality in the two areas are not additive.
    Response: Because of the time required to sort out the landings and discard data, and the time devoted to finding an acceptable model configuration, there was insufficient time during the panel meeting to complete this request. This request should be addressed in the final assessment report.
14. **Provide a run with single area model, re-combining the Triennial and NWFSC surveys into one series each with corresponding age composition data. If results are significantly different, try to explain the differences.**
    Rationale: The draft assessment used two-areas mostly because of difference in growth. As the assessment is largely age-based, difference in growth are less important than the reduction in sample size of stock size indices and composition information caused by having to break the series in two areas.
    Response: Because of the time required to sort out the landings and discard data, and the time devoted to finding an acceptable model configuration, there was insufficient time during the panel meeting to complete this request. This request should be addressed in the next assessment.

Description of base case model and alternative models to bracket uncertainty

Start year of the model =1916; two area model with time-invariant proportion of recruits assigned to each area; discard incorporated into total catches;
M fixed at 0.125yr^{-1} for both females and males; h estimated (but with Dorn's prior); σ_R = 0.6;
Von Bertalanffy growth parameters - all fixed for females and males.

Fisheries

- Washington fishery (all gears combined)
- Oregon midwater trawl fishery
- Oregon bottom trawl fishery
- California fishery (all gears combined)

Abundance indices:

- Oregon trawl CPUE (1984-1999)
- Foreign Pacific whiting fishery bycatch CPUE (1977-1988)
- Joint-venture Pacific whiting fishery bycatch CPUE (1983-1990)
- Domestic Pacific whiting fishery bycatch CPUE (1991-1998)
- Midwater trawl pelagic juvenile survey (2001-2008)
- Triennial trawl survey south of 43^{o} N (1980-2004)
- Triennial trawl survey north of 43^{o} N (1980-2004)

NWFSC shelf-slope trawl survey south of 43° N (2003-2008)
NWFSC shelf-slope trawl survey north of 43° N (2003-2008)

Age frequencies:

Washington fishery (all gears combined) (1980-2008)
Oregon midwater trawl fishery (1984-2008)
Oregon bottom trawl fishery (1984-2008)
California fishery (all gears combined) (1978-2008)
NWFSC shelf-slope trawl survey north of 43° N (2003-2008)
NWFSC shelf-slope trawl survey south of 43° N (2003-2008)

Length frequencies:

Triennial trawl survey index north of 43° N (1980-2004)
Triennial trawl survey index south of 43° N (1983-2004)
NWFSC shelf-slope trawl survey north of 43° N (2003-2008)
NWFSC shelf-slope trawl survey south of 43° N (2003-2008)

Uncertainty was bracketed by fixing steepness at 0.25 (low biomass) and at 0.55 (high biomass) to contrast with the base model where steepness was estimated to be 0.4.

Technical merits of the assessment

There were problems with the data and structure of the model in the draft document. Through fruitful interchanges between the STAT and the STAR Panel, considerable improvements in both model structure and data were achieved. The base case available at the end of the Panel meeting is considered reasonably well investigated and a substantial improvement on the original base case. There remain problems, however, particularly with the sharply-peaked selectivity patterns. Because the panel spent time trying to reconstruct catch history and because an inappropriate tuning algorithm had been used, there was insufficient time to fully examine models fits and model results. Because of these concerns, the Panel recommends that the next widow rockfish assessment be a full assessment rather than an update.

Explanation of areas of disagreement regarding STAR panel recommendations

A. Among STAR panel members (including concerns raised by the GAP and GMT representatives

There were no areas of disagreement among STAR panel members.

B. Between the STAR panel and the STAT team

There were no areas of disagreement between the STAR panel and the STAT team.

Unresolved problems and major sources of uncertainty

Whether it is preferable to do a single area stock assessment or a two area assessment remains one of the major unresolved problems for the widow rockfish assessment. There also remain problems with age determination and the reliability of age readings which needs to be better evaluated. As a result of substantially lower allowable catch due to fisheries management measures, the widow rockfish assessment has moved from being relatively data rich to being data poor because low allowable catches have resulted in sparse age and length compositions, and have rendered useless fishery-dependent stock size indices that were used in previous assessments. Furthermore, there is no reliable fishery independent monitoring.

Management, data, or fishery issues raised by the GAP and the GMT representatives

GAP and the GMT representatives (as well as representatives of fishing industry attending the meeting) provided assistance in identifying appropriate discard rates to apply to different fisheries in different time periods.

Prioritized recommendations for future research and data collection

- For the next assessment of widow rockfish, reconsider the overall structure of the model including the definition of fisheries, assignment of catches, age or size based selectivities, one or two area model. Development of a one area model (at least for comparative purposes) should be a priority. Estimates of growth, maturity, and fecundity used in a one area model should be representative of widow rockfish throughout its range off the West Coast.

- Do an interagency aging comparison to more reliably estimate the aging error matrix for widow rockfish, possibly using the Punt et al. (2008) aging error program.

- Do a comparative analysis of break and burn and surface age reading of otoliths to evaluate the reliability of ages in the early years of the assessment. If the comparative analysis indicates that age composition estimates can be improved, re-read the otoliths using the best available method.

- Low allowable catches have resulted in sparse age and length compositions in recent years. Sampling protocols for widow rockfish should be re-evaluated to ensure that adequate samples are obtained.

- There is currently no effective monitoring program for this stock. There is a need for the development of alternative survey methods for widow rockfish.

- Investigate the usefulness of blocking time periods into presumptive environmental regimes as indicated by the PDO (Pacific Decadal Oscillation) to reflect possible climate effect.

- Bias adjustment of stock recruitment relationship in SS3 seems to have been widely implemented with little or no peer review. Simulation testing is needed to confirm that bias adjustment is justified in all cases, and guidelines are needed on how to configure bias adjustment settings to reflect the information available and the biological characteristics of the stock.

- Develop methods to incorporate uncertainty in natural mortality and/or steepness in model configurations in which these parameters are fixed. The delta method for propagating uncertainty (McCall in prep.) is promising approach that warrants further evaluation.

Acknowledgements

The Panel thanks staff at the SWFSC Santa Cruz laboratory their exceptional support and provisioning during the STAR meeting.

References

Lynde, M.V. H. 1986. The historical annotated landings (HAL) database: documentation of annual harvest of groundfish from the northeast Pacific and eastern Bering Sea from 1956-1980. NOAA Tech. Memo. NMFS-F/NWC103.

Punt, A.E., D.C. Smith, K. KrusicGolub, S. Robertson. 2008. Quantifying age-reading error for use in fisheries stock assessments, with application to species in Australia’s southern and eastern scalefish and shark fishery. Can. J. Fish. Aquat. Sci. 65: 1991-2005.

Ralston, S., D. Pearson, J. Field, M. Key. 2009. Documentation of California catch reconstruction project (draft). NOAA Fisheries, SWFSC, Santa Cruz, California.

Rogers, J.B. 2003. Species allocation of Sebastes and Sebastolobus sp. caught by foreign countries from 1965 through 1976 off Washington, Oregon, and California, USA. U.S. Department Commerce, NOAA Tech. Memo. NMFS-NWFSC-57.

Tagart, J.V. 1985. Estimated domestic trawl rockfish landings, 1963-1980. Washington Department of Fisheries (unpublished manuscript and data).

Appendix: Multinomial sample sizes

This Appendix describes the method that was used to correct the multinomial sample sizes for compositional data in the widow rockfish assessment, shows the results of its application to these data, and provides a brief rationale for the use of this method. The description that follows concerns age data, but the approach is exactly the same for length data. Also, for simplicity, the following description ignores sex, but the extension to include sex is straightforward.

Suppose $p_{ay,\mathrm{obs}}$ is the observed proportion at age a in year y in a set of age composition data that is assumed to have a multinomial error structure, and let $N_{\mathrm{init},y}$ denote the initial sample sizes. We run our assessment model using these sample sizes and obtain, from the model fit, estimates of the expected proportions at age, $p_{ay,\mathrm{exp}}$. Our aim is to use the $p_{ay,\mathrm{obs}}$, $p_{ay,\mathrm{exp}}$, and $N_{\mathrm{init},y}$ to calculate a correction factor f so that the size of the model residuals is consistent with the corrected sample sizes, $N_{\mathrm{corr},y} = fN_{\mathrm{init},y}$.

This aim is the same as for the method currently used in SS3 to correct multinomial sample sizes. Where the two methods differ is that the SS3 method is based on the residuals for individual proportions, $r_{ay} = p_{ay,\mathrm{obs}} - p_{ay,\mathrm{exp}}$, whereas the present method is based on residuals for mean age, $r_y = (m_{y,\mathrm{obs}} - m_{y,\mathrm{exp}})$, where $m_{y,\mathrm{obs}} = \Sigma_a(a\ p_{ay,\mathrm{obs}})$, and $m_{y,\mathrm{exp}} = \Sigma_a(a\ p_{ay,\mathrm{exp}})$. The reason for using mean-age residuals is discussed below.

For the multinomial distribution, the expected variance of the mean age, $m_{y,\mathrm{obs}}$, and thus of the residual r_y, is $v_y/(fN_{\mathrm{init},y})$ [i.e., $v_y/N_{\mathrm{corr},y}$], where v_y is the variance of the age frequency in year y, given by $v_y = \Sigma_a(a^2\ p_{ay,\mathrm{obs}}) - m_{y,\mathrm{obs}}^2$. Therefore, the expected variance of $r_y(N_{\mathrm{init},y}/v_y)^{0.5}$ is $1/f$, and we estimate f as $1/\mathrm{Var}(r_y(N_{\mathrm{init},y}/v_y)^{0.5})$.

Figure 1 shows the application of this method to the widow rockfish age and length composition data. In this application, the sample sizes, $N_{\mathrm{init},y}$, were those obtained after correction (or tuning) using the SS3 method. That is to say, the residuals for individual proportions, r_{ay}, should be consistent with the size expected given these sample sizes. What Figure 1 shows is that the mean-age residuals are still too large (note that many of the confidence intervals for the observed values do not overlap the expected values). Thus, according to the present method, the sample sizes are too large, and so the estimated f is less than 1 for all data sets (range 0.04 to 0.45).

The reason the mean-age residuals, r_y, can be too large, while the individual proportion residuals, r_{ay}, are not, is that there is substantial correlation between the individual residuals. This is shown in Figure 2, in which the observed age frequency flips from one side of the expected frequency to the other from one year to the next. Lateral shifts of this magnitude would not be possible if the individual residuals were uncorrelated. This correlation could be caused by either observation or process error (or a combination of both). One explanation of how this could occur derives from the fact that size (and thus age) distribution of fish often varies spatially. Between-year lateral shifts in age frequencies could occur if the spatial distribution of fishing changes substantially from year to year. In a model like that for widow rockfish, in which selectivities are assumed to be time-invariant, changes in the spatial pattern of fishing would add process error to the observations. If the spatial distribution of catch sampling varied from year to year, in a way that did not reflect the movement of fishing activity, this would produce observation error, with a correlation between individual proportions. More information on the generation of correlations in composition data are given by Hrafnkelsson & Stefánsson (2004) and section 3 of Francis (2006).

Note that there is no intention to suggest that mean age (or length) is a quantity of particular interest in stock assessments. The only reason mean age (or length) is used in calculating the correction factor is that it is sensitive to the sort of correlations shown in Figures 1 and 2. Because of these correlations the composition data are less informative than is suggested by the multinomial sample size correction method used in SS3 (which assumes no correlation). Ideally, we should include these correlations in the likelihood function for composition data. However, that is not straightforward to do. The method proposed here is a simpler pragmatic alternative approach which adjusts sample sizes to compensate for the correlations.

Appendix References

Hrafnkelsson, B.; Stefánsson, G. (2004). A model for categorical length data from groundfish surveys. Canadian Journal of Fisheries & Aquatic Sciences 61: 1135-1142.

Francis, R.I.C.C. (2006). Some recent problems in New Zealand orange roughy assessments. New Zealand Fisheries Assessment Report 2006/43. 64 p.

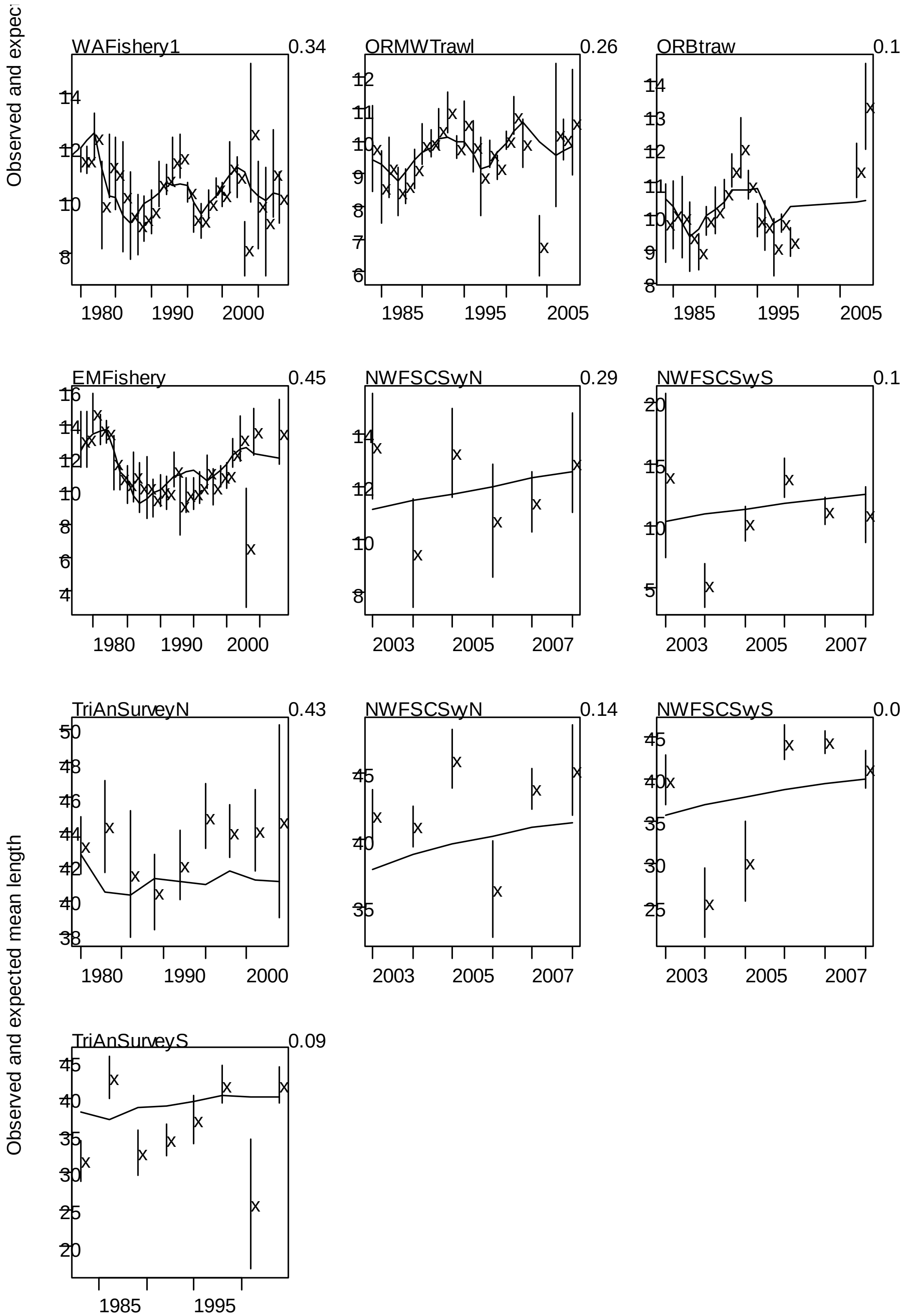

Figure 1. Observed and expected mean ages (upper panels) and lengths (lower panel) for the widow rockfish composition data in the (former) base case model. Vertical bars are approximate 95% confidence intervals based on the multinomial sample sizes used in that model. The number printed above each panel is the correction factor *f*, calculated as described above.

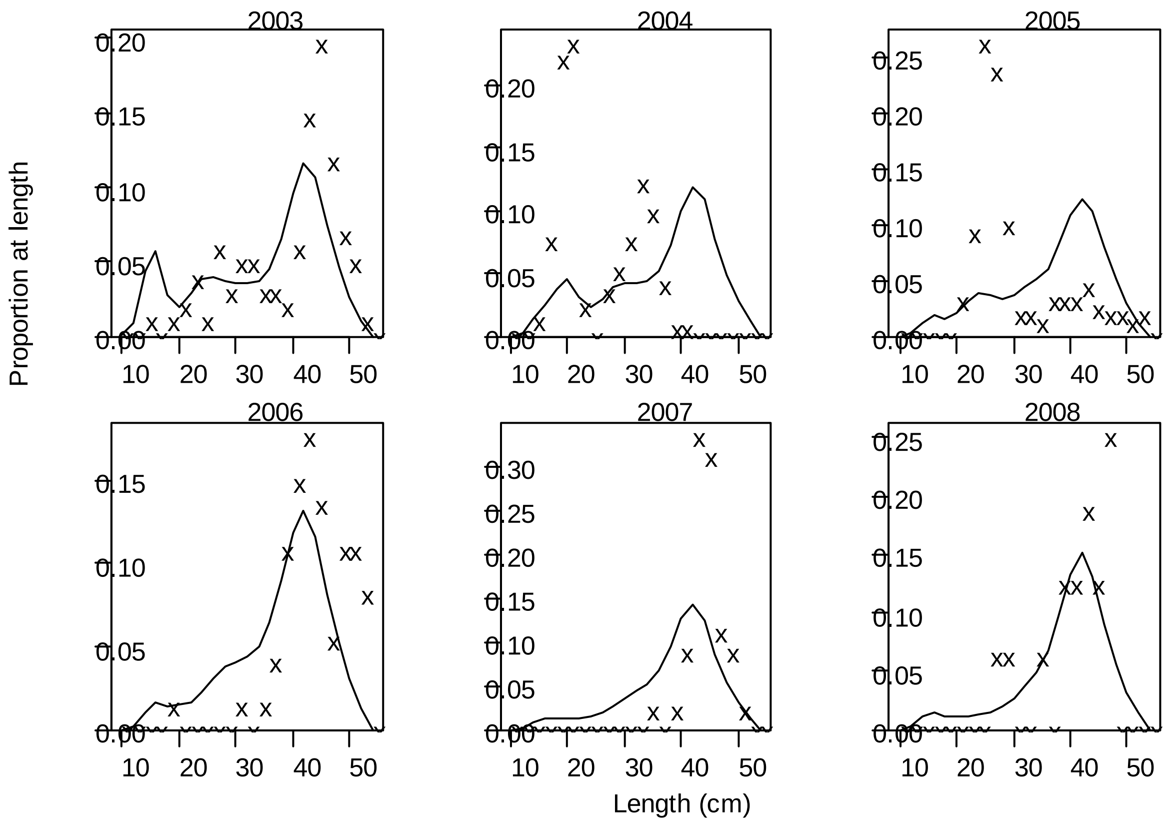

Figure 2. Observed ('x') and expected (line) proportions at length (sexes combined) for data set NWFSCSvyS.