Fisheries and Oceans Canada Pêches et Océans Canada

NOAA FISHERIES
NATIONAL MARINE FISHERIES SERVICE

CERT	TRAC
Comité d'évaluation des ressources transfrontalières	**Transboundary Resource Assessment Committee**
Document de travail 2014/17	**Working Paper 2014/17**
Ne pas citer sans autorisation des auteurs	Not to be cited without permission of the authors

Evaluating age and length composition data for inference about selectivity shape

Elizabeth N. Brooks[1], Philip J. Politis[2]

[1]Population Dynamics Branch; [2]Ecosystem Survey Branch
Northeast Fisheries Science Center
166 Water Street
Woods Hole, MA 02543

ABSTRACT

Age and length composition data for Georges Bank Yellowtail Flounder from three fisheries independent surveys and from the commercial catch at age were examined to determine whether there was sufficient information to infer selectivity shape outside of the VPA model application. Age and length composition supported the same conclusions about relative selectivity between the NMFS spring, NMFS fall, and DFO surveys. Specifically, the DFO survey appears to have dome-shaped selectivity, the NMFS spring survey has higher selectivity than the DFO survey at the oldest ages, and the NMFS fall survey has the greatest selectivity (over all 3 surveys) at the youngest ages. The NMFS fall survey does not always observe the oldest age (6), so the relative selectivity pattern was not as informative about this age class. The DFO survey had relative selectivity patterns that were most similar to the fishery. This would suggest that there might be some doming expected for the fishery selectivity. Two alternative VPA configurations were explored to determine if configurations with doming in the fishery could help resolve the severe retrospective pattern; the dome models reduced the retro slightly, but the magnitude was still very large.

Introduction

The Georges Bank (GB) Yellowtail Flounder assessment has exhibited strong retrospective patterns for a decade. Retrospective patterns are characterized by a consistent bias in model estimates as subsequent years of data are added to the model. In the case of GB Yellowtail Flounder, previous assessment estimates of spawning stock biomass (SSB) were found to be consistently overestimated each time another year of data was added to the model. Similarly, previous estimates of fishing mortality (F) were found to be consistently underestimated each time another year of data was added.

The specific cause for this pattern of bias is not known for GB Yellowtail Flounder. General research on retrospective patterns has concluded that misspecification of model assumptions and unrecognized bias in data (catch, e.g.) can produce the same pattern of bias observed in the GB yellowtail assessment (ICES 2008; Legault 2009). Approaches that have been shown to “fix” retrospective patterns include: increasing catch or natural mortality (M) in recent years or reducing catch of M early in the time series; allowing survey q to increase in recent years, often by artificially splitting surveys into two distinct time series. While sensitivity analysis can identify when in the time series a change is needed, it is not possible to determine which of the multitude of possible factors is the underlying cause, or whether the retrospective pattern is due to more than one factor.

In the case of GB Yellowtail Flounder, and several other New England groundfish assessments, extensive sensitivity analyses identified the years 1994-1995 as the most likely period when a change in data or model assumptions would induce a retrospective pattern (Legault 2009). It is known that substantial changes occurred in this period for both data collection and fishery management, however, the operation of the Northeast Fisheries Science Center (NEFSC) research bottom trawl survey and the Canadian Department of Fisheries and Oceans (DFO) bottom trawl survey did not change. Despite this fact, the accepted fix for the retrospective pattern for GB Yellowtail Flounder has been to split the survey time series between 1994 and 1995, treat the split surveys as independent, and freely estimate catchability at age (q_a) for both sets of survey time series (TRAC 2005). This retrospective fix worked for a number of years, but recently another retrospective pattern with the same direction of bias has appeared. As this pattern has worsened, concerns have been raised about the appropriateness of the assessment model to estimate population trend and abundance, and the resulting management advice.

In this working paper, we draw attention to the pattern in q_a that results from splitting the survey time series. When the series are not split, the model estimates of q_a suggest that both the spring and the fall NEFSC bottom trawl surveys have a selectivity shape that could be described as “flat topped,” meaning that the selectivity of the gear at age increases to a point, and then is more or less constant (flat with respect to the y-axis). The DFO bottom trawl survey, however, is estimated to have a “dome shaped” pattern in selectivity, meaning that the selectivity of the gear increases to some intermediate age and then decreases for the oldest ages. When the survey series are split into pre- and post-1995, then the pattern in q_a is dramatically different in the two periods. For the NEFSC spring survey, the pre-1995 q_a look convex rather than flat-topped, while the post-1995 q_a now display a dome-shaped selectivity with a peak at age 4. In the fall NEFSC survey, both the pre- and post-1995 q_a are dome-shaped, with the peak age shifting from age-4 to age-3 after the split. For the DFO survey, the q_a post-1995 matches the unsplit time series, and the pre-1995 series merely shifts the peak age by one year from 4 to 5 (Fig. 1). On the other hand, the selectivity of the fishery is flat-topped whether or not the survey series are split; this is primarily a constraint in the model specification for F on the oldest ages (Fig. 2a-d).

In what follows, we review the literature on gear performance to understand factors that may be important to interpreting selectivity patterns. Next, we explore the raw yellowtail age and length composition data for all three surveys (NEFSC spring and fall, DFO) and also the catch series. We evaluate whether relative selectivities derived from age and length composition data support the pattern of model estimated selectivities among the different data sets. In addition, we explore whether a different model specification for F on the oldest ages—one which allows more flexibility in the selectivity at the oldest ages—has any impact on the selectivity patterns estimated for the surveys, and on the retrospective pattern. We conclude with research recommendations that could provide a framework for experimentally evaluating gear selectivity.

The fish capture process by bottom trawl is complex with species and size specific selectivity occurring well in front of the actual net. Behavioral responses to the approaching vessel and gear influence the horizontal and vertical distribution of some species, while other species react to the bottom trawl doors and bridles, and the sand clouds they generate (Main and Sangster, 1981a, 1981b; Wardle, 1993; Engas, 1994; Walsh, 1996). Some of these reacting fish move outwards and avoid capture while others move inwards towards the path of the approaching trawl, repeatedly encountering and reacting to the bridles while moving further inward toward the path of the trawl in a process termed herding (Hemmings, 1969; Wardle, 1993). Commercial trawls seek to exploit this herding response by extending the distance between the wing-ends and doors to increase the area swept by the gear and increase catch rates and often cover the cables with rubber discs to maximize bottom contact and visual stimuli (Dickson, 1974; Strange, 1984). Herding is influenced by many factors including species, size, light, temperature, gear configuration and individual fish condition, thus increasing variability in the overall process (Wardle, 1993; Winger et al., 1999; Ryer and Barnett, 2006; Ryer, 2008; Kotwicki et al., 2009). Flatfish species tend to react to bottom trawl gear at shorter distances than roundfish and roundfish tend to have stronger burst and sustained swimming capabilities, suggesting differences in the herding or avoidance response mechanisms (Wardle, 1993; Winger et al., 1999; Winger et al., 2004; Ryer, 2008). Studies have shown that the length of the bridles are an influential factor for herding efficiency and that herding efficiency is both size and species specific (Harden Jones et al., 1977; Engas and Godo, 1989a; Dickson, 1993a, 1993b; Ramm and Xiao, 1995; Somerton and Munro, 2001). However, herding efficiency is not well defined for species in the Northwest Atlantic, particularly flatfish species. Engas and Godo (1989a) examined the effects of different bridle lengths on the length composition of survey trawl catches of Atlantic Cod and Haddock and found a general increase in both numbers and size of both species with increasing bridle length. Somerton and Munro (2001) estimate bridle efficiency of a survey trawl based on three different bridle lengths for seven species of flatfish in the Eastern Bering Sea. Their results showed significant herding effects for all species of flatfish with no length effect for five of the species and a significant decrease in bridle efficiency with increasing fish length for two species. The speed of the tow and the angle at which the bridle extends away from the wing-end are thought to influence herding efficiency (Engas, 1994). If the bridle angle is too large or the tow speed too fast, iterative encounters with the bridle will occur too quickly and fish will be unable to react and thus overtaken by the bridle and not captured. Once fish reach the mouth of the trawl they may react again in an attempt to avoid the gear. Some species are known to turn and swim in the mouth of the trawl or directly in front of the ground gear, at the same speed of the trawl until they become exhausted and fall back into the net or attempt to escape either over the headrope or dive under the ground gear (Hemmings, 1973; Wardle, 1993). Escapement under the ground gear is a function of the ground gear size and configuration and is also known to be species and size dependent (Engas and Godo, 1989b; Walsh, 1992; Munro and Somerton, 2002). Survey bottom trawls typically utilize small mesh webbing and codend liners so escapement through the meshes is likely minimal for all but the smallest fish (Engas, 1994; Walsh, 1996). Vision is considered the primary sense used by fish for the avoidance of gear (Glass and Wardle, 1989; Wardle, 1993). Herding and avoidance behaviors are therefore influenced by bottom light levels and studies have shown significant diel differences in catches of certain species as a result (Walsh, 1991; Ryer and Barnett, 2006; Kotwicki et al., 2009).

Research survey bottom trawl designs are typically based on common commercial trawls utilized in the survey region and modified by reducing mesh size and installing codend liners to retain juvenile fish (Walsh, 1996). Survey bottom trawls seek to obtain a representative sample of the species, age and sex composition in the area sampled (Godo and Walsh, 1992; Engas, 1994). Given the inherent variability in the capture process due to herding variability, some multispecies bottom trawl surveys utilize the shortest possible bridle lengths necessary to achieve the desired net mouth opening. The survey trawl designs utilized for the NEFSC multispecies groundfish surveys follow the logic of using minimal wing-end to door distance to limit variability from herding. The NEFSC Yankee #36 (primarily used during 1963-2008, except for the Yankee #41 for spring surveys in 1973-1981) was a low vertical opening trawl and used a total wing-end to door distance of 12.2 m (Table 1). In contrast, the current NEFSC standard 4-seam, 3-bridle trawl (2009-present) is designed to achieve greater vertical opening of 4.0 m and uses a total wing-end to door distance of 48.9 m (Table 1). The DFO Maritimes region operates a multispecies bottom trawl survey that overlaps areas with the NEFSC survey for several species. The DFO survey bottom trawl, the Western IIA, has 27.2 m bridles to achieve a vertical opening of 4 m; however, the standard configuration includes an additional 27.4 m of ground cable length with a total wing-end to door distance of 64.3 m (Table 1). Each survey must sample in hard bottom habitats and use large ground gear accordingly, although the specific design characteristics vary (Table 1). The NEFSC Yankee 36 used 35.6 cm roller gear with 60 cm spacing between them, the NEFSC standard 4-seam, 3-bridle uses a rockhopper ground gear with 40.6 cm rockhopper discs in the center and 35.6 cm rockhopper discs on the wing sections, and the Western IIA uses 48.2 cm spherical bobbins in the center section and 29.2 cm bobbins on the wing sections (Table 1). Given the differences in gear configuration, there are likely species and size catchability differences between these survey bottom trawls due to different ground gear and bridle efficiencies.

In an effort to maximize profits, commercial fishing operations target high abundance areas with gear configured to capture the most marketable fish for the least amount of cost. This leads to fairly wide range of gear configurations and towing procedures based on the target species, bottom type, and captain’s preference. In general, commercial bottom trawls are typically larger than survey bottom trawls and utilize larger mesh sizes to select for adult, legal sized fish. To take advantage of herding behavior and increase the area swept by the gear, bridle lengths are extended, especially when targeting flatfish on smooth bottom, where ground cable lengths used may exceed 300 m. Ground cable lengths must be shorter in hard bottom to limit hangs and gear damage. The ground gear designs and sizes used are based primarily on bottom type. In order to fish hard bottom, larger ground gears must be used, such as bobbins or rockhoppers, compared to soft bottom areas where chain and flat cookie ground gears are often utilized to maximize bottom contact and minimize areas of escapement. Representative characterizations of commercial gear employed on trips where the captain identified “flatfish” as a target were extracted from the Northeast Fisheries Observer Program (NEFOP) database for years 1994-2013 (Table 1).

Methods

Relative Selectivity

To motivate the methods used to explore the age composition data, we first illustrate relative selectivity for data that was generated from known population numbers at age and known selectivity patterns. For the values specified in Table 2, three cases were explored: 1) flat-topped survey and domed fishery selectivity; 2) domed survey and flat-topped fishery selectivity; 3) domed selectivity in the survey and the fishery. In all three cases, the peak age in the survey was 5 while in the fishery it was age 4.

For the specified numbers at age (N), F, and selectivity values, we calculated annual catch at age and numbers at age at the beginning of the next year, for 5 years. In addition, a "survey" was estimated based on the beginning of year numbers. Next, the numbers at age in the survey and the catch were used to calculate the proportion at age (age composition) for each year. To calculate relative selectivity, the age composition in each year of the survey was divided by the age composition of the catch in the same year. The average of all 5 years is then plotted in Figure 3.

With real data, the shape of selectivity is unknown, however by examining the ratio of proportion at age between different data sources, one gains inference into the relative selectivity at age, i.e. whether one data source has greater or lesser selectivity at age than another data source can be inferred based on whether the ratio at age is greater or less than one (Brooks 2011, unpublished data; Clark 2014). For the three cases defined in Table 2, the true selectivity at age, true relative selectivity at age, and the ratio of proportion at age is plotted (Figure 3). Looking only at the bottom two rows of Figure 3, it is not possible to determine whether the original data sources exhibited flat-topped or domed selectivity, it is only possible to infer the ages for which the survey exhibited greater selectivity than the fishery (those ages where the ratio is above 1.0). These example cases have no error associated with the age composition, so interpreting plots of real data would probably require that ratios be some margin above or below 1 to infer a meaningful difference in selectivity at age.

The exercise of calculating relative selectivities is repeated for research survey and commercial catch data. First, pair-wise comparisons of the research surveys are presented. An underlying assumption of research surveys is that selectivity is constant for the whole time period, so one expects that patterns in the relative selectivity should be stable. A caveat to that expectation is that the number of fish observed in the surveys is a lot less than for the catch, so the relative selectivity pattern may be noisier.

A second set of pair-wise comparisons is between each survey and the commercial catch at age. For the time series of catch, which spans 1973-present, various management actions, gear modifications, and other factors can influence the realized selectivity in the fishery. Thus, we do not expect constant selectivity for the fishery in all years. Another consideration is the fact that the surveys occur at relatively the same time of year for the entire time series, while the catch is taken throughout the year and the proportion caught per quarter can vary (this would potentially affect length frequency comparisons, but is not expected to affect age composition comparisons).

After evaluating the relative selectivity of the age composition data, we look to length frequency data to verify that the same patterns exist. Specifically, we compare the NEFSC spring length composition to the DFO survey length composition for all years that data are available (1987-present). We also compare the NEFSC spring survey and DFO survey length composition to that of the catch in calendar quarters 1&2 for years 2007-2012 (data from earlier years exist but were not processed in time to be included in this analysis).

When comparing the various data sources, the value of 1.0 is not interpreted as the absolute determinant for greater selectivity at age. An arbitrary margin of +/- 0.1 is allowed for the threshold in recognition of the fact that this is real data.

Finally, after evaluating the data to see if it suggests selectivity patterns, we test a few different VPA configurations using the 2012 TRAC Single series model. Because there is a severe retrospective pattern in the Single series model, the aim of these new configurations is to see if a different specification for fishery selectivity can reduce any of the model tension that might be responsible for creating the retrospective bias.

Results

Comparison between survey age composition

The number of years of age composition available for comparison depends on the survey. Age composition for Yellowtail Flounder is available since 1982 (using the Yankee 36 net on the former NOAA ship *Albatross IV*) for the NMFS spring survey, since 1973 for the NMFS fall survey, and since 1987 for the DFO survey. For years 2009-2012, the NMFS surveys were conducted on a new vessel with new gear; those surveys were calibrated to *Albatross IV* units based on an extensive calibration study (Miller et al. 2009), using length based estimates that were estimated for the 2010 TRAC assessments (Brooks et al., 2010). Proportions at age for each survey were compared for years where data were available. Relative selectivities were calculated annually, and the time series average was summarized (Table 3a-c).

The NMFS spring survey observed zero age 1 fish in about 1/3 of the years compared to only 1 out of 39 years in the fall survey. For the years where both surveys observed age 1, the NMFS fall survey appears to have higher selectivity at age 1 (Table 3a). Similarly, the fall survey has higher selectivity at age 2, while the spring survey probably has slightly higher selectivity at age 3. Selectivity in the spring survey is higher than in the fall survey at ages 4 and 5. In about 1/3 of the years, the fall survey did not observe any age 6 fish, while there were only two years that the spring survey did not observe age 6. For years where both spring and fall surveys observed age 6, the spring survey appears to have higher selectivity for age 6.

The DFO survey did not observe age 1 fish in about 20% of the years, and that index is not used in the VPA model. For the remaining ages, the NMFS spring survey has higher selectivity at age 2, similar selectivity at ages 3-4, and higher selectivity than DFO at ages 5-6 (Table 3b).

Comparing the DFO and the NMFS fall surveys (for ages>1), the NMFS fall survey has higher selectivity at age 2. The pattern at age 3 is variable, but the selectivity is probably similar for the two surveys. For ages 4 and 5, the DFO survey has higher selectivity than the fall survey. At age 6, there are many years that the fall survey did not observe that age group; for the years where both surveys observed age 6, it appears that the DFO survey has higher selectivity.

Comparison between survey and catch age composition

Similar to the comparisons between surveys, relative selectivities between each survey and the catch were calculated annually, and the time series average was summarized (Table 4a-c). In addition, summaries were made for different blocks of years (Table 5).

There are no years without observations of age 1 fish in the catch. For years where both the NMFS Spring survey and the catch observed age 1, it appears that selectivity at age 1 may be slightly higher in the catch (Table 4a). The pattern at age 2 is variable, but on the whole the selectivity is probably similar between the catch and the NMFS spring survey for ages 2-4. For ages 5-6, there is a notable temporal pattern, where the NMFS spring survey appears to have higher selectivity before 2000, similar selectivity from 2001-2003, and then the catch has higher selectivity from 2004-present (Table 5).

The NMFS Fall survey has higher selectivity than the catch at age 1 (Table 4b). At age 2, the pattern is variable up to year 2000 and the selectivity may be more or less similar; however, after 2000, selectivity in the Fall survey appears to be higher than in the catch. The pattern at age 3 is variable, not indicating strongly whether the Fall survey or the catch has higher selectivity. At ages 4 and 5, the catch appears to have somewhat higher selectivity than the Fall survey. While there are a number of years where the Fall survey did not observe age 6, for the years where it did the pattern is variable up to year 2002, from which point the catch has higher selectivity (Table 5).

The DFO survey did not observe age 1 fish in about 20% of the years, and that index is not used in the VPA model. Comparing the DFO survey and catch (for ages>1), the pattern at age 2 is variable but no strong trend is indicated for higher selectivity in the DFO survey or the catch (Table 4c, Table 5). For ages 3-6, the DFO survey and the catch appear to have similar selectivity, possibly with the exception of the years since about 2003 where the catch appears to have somewhat higher selectivity.

Comparison between surveys and catch length composition

The NMFS Spring and DFO survey length compositions were compared for lengths between 23 and 51 cm. There were insufficient samples outside of this length range to make meaningful comparisons. Across all years (1987-2013), at the smallest (less than 30 cm) and largest (above 40 cm) sizes, the NMFS Spring survey had a higher average relative selectivity than the DFO survey (Figure 4). Between 30-40 cm, which corresponds to ages 3 to 4 or 5, the surveys had very similar selectivity (Figure 4). This confirms the same pattern observed in the age composition data.

Each survey was compared to the catch length composition for lengths between 30 and 50 cm. Both surveys had higher relative selectivity than the catch below 35 cm, and both surveys had lower relative selectivity than the catch between 35-45 cm. Beyond 45 cm, the NMFS Spring survey had higher relative selectivity than the catch, while the DFO survey had lower relative selectivity (Figure 4). The pattern at the smaller sizes is expected given minimum size regulations and net differences. The age composition data suggested that out of all three surveys, the DFO survey was likely to have a selectivity pattern that was most similar to the fishery, although it was noted that for all of the comparisons there was a temporal shift in the pattern around the years 2000 or 2003 to the present. The catch length composition data is from 2007-2012, the period where this shift was detected.

Alternative VPA selectivity specification for fishery

The current VPA configuration fixes selectivity at age 5 (the oldest true age) to be the same as age 4, and then fixes the ratio between the plus group (ages 6) and the oldest true age (5) to be 1. Thus, selectivity at ages 4, 5, and 6 is specified to be the same in the Single Series (and Split Series) model. Two alternative configurations were explored, referred to as Dome-1 and Dome-2. In Dome-1, F on age 5 was specified to be the average of the F on ages 3 and 4, and the F ratio between age 6+ and 5 was specified to be 1. This configuration forces the F on ages 5 and 6 to be the same, but allows them to be different from age 4. In the Dome-2 configuration, the F on age 5 is also specified to be the average of the F on ages 3 and 4; however the ratio between F on age 6+ and age 5 is specified to be 0.8. This configuration allows the F on age 5 to be different than on age 4, and then the F on age 6 is calculated as $0.8*F_{age5}$. The resulting fishery selectivity at age for the two dome scenarios is plotted in Figure 5a-5d.

Retrospective peels were compared for the Single Series, Split Series, Dome-1 and Dome-2 for peels of 16 years (back to 1995). The value of Mohn's rho for SSB is smaller for both dome configurations for all peels compared to the Single Series, however the values of Mohn’s rho are still extreme and exceed 1 for peels of more than 2 years (Figure 6). The Split Series has the “best” overall rho, but it is still unacceptably high.

The exploration of the alternative VPA configurations did not resolve the retrospective pattern. Allowing for a small amount of doming affected the absolute scale of the catchability at age for the indices and had a minor impact on the index selectivity (Table 6).

Discussion/Conclusions

Some of the gear configuration work indicated decreasing efficiency for longer bridle length. For some species, there was a length effect, meaning that longer bridles were less efficient at catching larger fish. While that finding did not hold for all flatfish studied by Somerton and Munro (2001), it does provide a potential mechanism for interpreting some of the patterns observed between surveys and the commercial catch. Compared to the NMFS surveys conducted on the *Albatross IV*, the DFO survey had much longer bridle length (3 times longer). Between the NMFS sSpring and the DFO February survey, both of which are on Georges Bank within a month or two of each other, there was clear evidence that selectivity at the older ages and larger lengths was greater in the NMFS spring survey than the DFO survey. It is not expected that growth was substantial enough in that month or two between the surveys to explain the difference. Furthermore, the difference cannot be blamed on different ageing protocols, because the NMFS age reader has been reading scales from DFO for the last decade and before that, the NMFS spring age length key was applied to DFO length frequencies (Stone and Perley, 2002). Thus, one explanation that cannot be ruled out at this point is that the efficiency, and hence selectivity, of the DFO gear at the larger sizes (older ages) is less than for the NMFS spring survey. Although the relative selectivities do not inform as to the true shape of the NMFS spring survey selectivity, the relative selectivity results would imply that the DFO survey has some doming in its selectivity.

When comparing the relative selectivities between surveys and the catch, the DFO survey appeared to have selectivity that was the most similar to the catch. Although there was a wide variety of gear configurations identified in the observer database, in general the bridle lengths for the commercial gear was longer than the NMFS surveys. Therefore, if the inferred selectivity pattern between the NMFS and DFO surveys can be attributed to decreasing efficiency with bridle length, then one might expect that, of the surveys examined, the DFO survey would be most similar in selectivity pattern to the fishery. While there are many other differences between the NMFS and DFO surveys (Table 1), the bridle length is one that has been investigated on the west coast for flatfish and some of those results support the findings in this analysis.

If we accept that there is a possibility for some doming in the DFO survey, then we should expect some dome shaped selectivity for the fishery—at least before 2003, when the pattern between the fishery and the surveys seems to have changed. Comparing these expectations with the VPA estimates of survey selectivity (Single Series model), we do indeed find that the DFO survey is estimated to have peak selectivity at age 4 and lower selectivity at ages 5 and 6 (Figure 1). Conversely, the VPA estimated selectivity in the fishery is flat-topped for nearly all years (Figure 2).

Allowing for a dome did not resolve the retrospective problem. It reduced the scale of q, and for the Dome-2 configuration, nearly all q estimates were less than 1 (there was a single value of 1.1, all others were < 1.0). Some view model results where q estimates are <1 to be a reasonable diagnostic, however there are ample examples where this diagnostic is overinterpreted.

There was a strong temporal pattern indicating a change in fishery relative selectivities after 2000, especially after 2003-present. It would be interesting to understand the cause of this change.

Recommendations

With respect to recommendations moving forward, we leave it to a future benchmark model meeting to determine if a dome-shaped selectivity configuration is a better representation of the fishery. It did not “fix the retro” and did not have a major impact on the scale of abundance for the population. In this regard, it does not change the advice regarding the status of the stock or the fact that catches in the immediate future must remain low.

Aside from the model decision of how best to represent fishery selectivity and selectivity for the research surveys, one could approach the problem with a designed field experiment. For example, experimental work could be conducted to directly estimate the catchability of the NEFSC survey bottom trawl gear, specifically focused on bridle efficiency and ground gear efficiency of flatfish. To estimate bridle efficiency a study comparing the length-based catch differences between the standard NEFSC bridle length and increased bridle lengths is recommended. To estimate ground gear efficiency a bag experiment attached under the trawl net and behind the ground gear is recommended. The objectives of these studies should be:

1) Define the species that are herded by the NEFSC sampling gear, focusing primarily on flatfish.
2) Estimate the length-based bridle efficiency for the herded species.
3) Estimate the length-based ground gear efficiency by comparing the catch ratios between the trawl codend and bag codends.

4) Examine the relationship between herding efficiency and ground gear efficiency and bottom light levels from archival light sensors attached to the trawl during all experimental tows.
5) Assimilate the information from the experiments to improve the estimate of the overall catchability for the NEFSC research survey bottom trawl.

Acknowledgements

The authors thank Chris Legault (NEFSC) for providing the commercial length frequency and Heath Stone (DFO) for providing the length frequency for the DFO survey.

Literature Cited

Brooks, E.N., T.J. Miller, C.M. Legault, L. O'Brien, K.J Clark, S. Gavaris, and L. Van Eeckhaute. 2010. Determining length-based calibration factors for cod, Haddock, and Yellowtail Flounder. TRAC Ref. Doc. 2010/08.

Clark, W.G. 2014. Direct calculation of relative fishery and survey selectivities. Fish. Res. 158:135-137.

Dickson, W. 1974. A review of the efficiency of bottom trawls. Fiskeriteknologisk Forskningsinstittut, FTFI Report, Bergen, 14-17pp.

Dickson, W. 1993a. Estimation of the capture efficiency of trawl gear. I: Development of a theoretical model. Fish. Res. 16: 239–253.

Dickson, W. 1993b. Estimation of the capture efficiency of trawl gear. II: Testing a theoretical model. Fish. Res. 16:255–272.

Engas, A., 1994. The effects of trawl performance and fish behaviour on the catching efficiency of demersal sampling trawls. In: Ferno, A., Olsen, S. (Eds.), Marine Fish Behaviour. Blackwell Science Inc., Massachusetts, pp. 45–68.

Engas, A., and O.R. Godo. 1989a. The effect of different sweep lengths on the length composition of bottom-sampling trawl catches. J. Cons. Int. Explor. Mer. 45: 263-268.

Engas, A., and O.R. Godo. 1989b. Escape of fish under the fishing line of a Norwegian sampling trawl and its influence on survey results. J. Cons. Int. Explor. Mer. 45: 269-276.

Glass, C.W., and C.S. Wardle. 1989. Comparison of the reactions of fish to a trawl gear, at high and low light intensities. Fish. Res. 7: 249-266.

GODØ, O. R., and S. J. WALSH. 1992. Escapement of fish during bottom trawl sampling – implications for resource assessment. Fish. Res. 13: 281–292.

Harden Jones, F.R., A.R. Margetts, M.G. Walker, and G.P. Arnold. 1977. The efficiency of the Granton otter trawl determined by sector-scanning sonar and acoustic transponding tags. Rapp. P-v. Reun. Cons. Explor. Mer. 170: 45-51.

Hemmings, C.C. 1969. Observations on the behaviour of fish during capture by the Danish seine net and their relation to herding by the trawl bridles. FAO Fish. Rep. No. 62, 3: 645-655.

Hemmings, C.C. 1973. Direct observations of the behaviours of fish in relation to fishing gears. Helgol. Wiss. Meeresunters. 24: 348-360.

ICES (International Council for the Exploration of the Sea). 2008. Report of the Working Group on Methods of Fish Stock Assessments (WGMG), 7-16 October 2008,Woods Hole, USA. ICES CM 2008/RMC:03. 147 pp.

Kotwicki, S., A. De Robertis, P. von Szalay, and R. Towler. 2009. The effect of light intensity on the availability of walleye pollock (*Theragra chalcogramma*) to bottom trawl and acoustic surveys. Can. J. Fish. Aquat. Sci. 66: 983-994.

Legault, C.M. (Chair) 2009. Report of the Retrospective Working Group, January 14-16, 2008, Woods Hole, Massachusetts. Northeast Fish. Sci. Center Ref. Doc. 09-01; 30p.

Legault, C.M., L. Alade, H.H. Stone, and W.E. Gross. 2012. Stock assessment of Georges Bank
Yellowtail Flounder for 2012. TRAC Ref. Doc. 2012/02. 133 p.

Main, J., and G. I. Sangster 1981a. A study of the sand clouds produced by trawl boards and their possible effect on fish capture. Scott. Fish. Res. Rep. 20:1−20.

Main, J., and G. I. Sangster 1981b. A study of the fish capture process in a bottom trawl by direct observations from a towed underwater vehicle. Scott. Fish. Res. Rep. 23:1−23.

Miller TJ, Das C, Politis PJ, Miller AS, Lucey SM, Legault CM, Brown RW, Rago PJ. 2010. Estimation of Albatross IV to Henry B. Bigelow calibration factors. Northeast Fish Sci Cent Ref Doc. 10-05; 233 p.

Munro, P.T., and D.A. Somerton. 2002. Estimating net efficiency of a survey trawl for flatfishes. Fish. Res. 55: 267-279.

Ramm, D. C., and Y. Xiao 1995. Herding in groundfish and effective pathwidth of trawls. Fish. Res. 24:243−259.

Ryer, C.H. 2008. A review of flatfish behavior relative to trawls. Fish. Res. 90: 138-146.

Ryer, C.H., and L.A.K. Barnett. 2006. Influence of illumination and temperature upon flatfish reactivity and herding behavior: potential implications for trawl capture efficiency. Fish. Res. 81: 242-250.

Somerton, D.A., and P. Munro. 2001. Bridle efficiency of a survey trawl for flatfish. Fish. Bull. 99: 641-652.

Strange, E.S. 1984. Review of the fishing trials with Granton and Saro deep sea trawl gear 1963-1967. Scottish Fisheries Working Paper, 8-18: 1-59.

Stone, H.H., and P. Perley. 2002. An evaluation of Georges Bank yellowtail flounder age determination based on otolith thin-sections. Can. Sci. Advis. Sec. Res. Doc. 2002/076. 32p.

TRAC (Transboundary Resources Assessment Committee). 2005. Proceedings of the Transboundary Resources Assessment Committee (TRAC) benchmark review of stock assessment models for the Georges Bank Yellowtail Flounder stock. S. Gavaris, R.O'Boyle, and W. Overholtz [eds.]. TRAC Proc. Ser. 2005/01. 65p.

Walsh, S.J. 1991. Diel variation in availability and vulnerability of fish to a survey trawl. J. Appl. Icthyol. 7: 147-159.

Walsh, S.J. 1992. Size-dependent selection at the footgear of a groundfish survey trawl. N. Am. J. Fish. Mgmt. 12: 626-633.

Walsh, S.J. 1996. Efficiency of bottom sampling trawls in deriving survey abundance indices. NAFO Sci. Coun. Studies, 28: 9-24.

Wardle, C.S., 1993. Fish behaviour and fishing gear. In: Pitcher, T.J. (Ed.), Behavior of Teleost Fishes. Champman & Hall, London, pp. 609–643.

Winger, P.D., He, P.,Walsh, S.J., 1999. Swimming endurance of American plaice (*Hippoglossoides platessoides*) and its role in fish capture. ICES J. Mar. Sci. 56: 252–265.

Winger, P.D., Walsh, S.J., Pingguo, H., Brown, J.A., 2004. Simulating trawl herding in flatfish: the role of fish length in behaviour and swimming characteristics. ICES J. Mar. Sci. 61: 1179–1185.

Table 1: Comparison of gear for survey gear and a representative range of commercial gear targeting “flatfish” (Information supplied by Philip J. Politis, Don S. Clark)

	Bigelow 4-Seam	Albatross Yankee 36	DFO Western 2A	Commercial Bottom Trawl Gear
Trawl Design	3-bridle, 4-seam 3 wings/jibs	2-bridle, 2-seam Full lower wing	2-bridle, 4-seam, flying wing (no lower wing)	Various
Ground Gear Design	Rockhopper (40.6cm/35.6cm)	Roller (35.6cm)	Bobbin (48.2cm/29.2cm)	Flat, chain, rockhopper, bobbin
Trawl Door	2.2m², 550kg Poly-Ice Oval	450kg Euronete	990kg Euronete	Various
Headrope Length	21.6m	18.3m	22.7m	15m-60m
Ground Gear Length	25.3m	24.3m	32m	15m-80m
Bridle Length	36.6m	9.1m	27.4m	10m – 60m
Ground Cable Length	0m	0m	27.4m	0m – 360m
Total Door-Wing Dist (extension+backstrap)	48.9m	12.2m	64.3m	10m – 400m
Mesh Size	12cm – 6cm – 2.54cm	12.7cm – 1.27cm	13cm – 3.2cm – 1.9cm	16.5cm
Avg. Door Spread	33.5m	22m	45m	30m – 100m
Avg. Wing Spread	12.6m	11m	13.5m	10m – 30m
Avg. Bridle Angle	12.3°	26.8º	14.2º	10º - 20º
Avg. Height	3.7m	1.5m	3.2m	1.5m – 8m
Standard Tow Speed Over Ground	3.0kts	3.8kts	3.5kts	2kts – 4kts
Standard Tow Duration	20min	30min	30min	60min – 120min

Table 2. Values used to generate relative selectivity examples. Starting numbers at age in year 1, and number of recruits at age 1 in each year, are indicated in the numbers at age matrix. For each case, these starting population numbers, annual F(year), and the selectivities specified for each case were used to complete the numbers at age matrix and to generate age composition for the survey and fishery.

	Numbers at age						
	1	**2**	**3**	**4**	**5**	**6**	**F(year)**
Year1	1500	2200	1895	1650	1230	1545	0.4
Year2	1800						0.6
Year3	1100						0.7
Year4	900						0.45
Year5	1325						0.35

	Selectivity at age					
Case 1	**1**	**2**	**3**	**4**	**5**	**6**
Survey	0.15	0.28	0.55	0.85	1	1
Fishery	0.05	0.15	0.67	1	0.88	0.75
Case 2						
Survey	0.15	0.28	0.55	0.85	1	0.75
Fishery	0.05	0.15	0.67	1	1	1
Case 3						
Survey	0.15	0.28	0.55	0.85	1	0.7
Fishery	0.05	0.15	0.67	1	0.8	0.5

Table 3a. Relative selectivity between NMFS spring and NMFS fall surveys. Blank cells indicate no data for one or more surveys, zero cells indicate a value of zero for the numerator (NMFS spring), NaN cells indicate a value of zero in the denominator (NMFS fall).Green cells are >1.1; red cells are <0.9.

	NMFS Spring/NMFS Fall					
Year	**1**	**2**	**3**	**4**	**5**	**6**
1973						
1974						
1975						
1976						
1977						
1978						
1979						
1980						
1981						
1982	0.02	1.78	0.73	2.49	5.29	NaN
1983	0.00	0.73	1.27	0.93	3.57	4.44
1984	0.00	0.14	1.66	2.17	6.80	2.38
1985	0.06	2.96	1.12	4.07	1.34	NaN
1986	0.07	1.25	0.64	0.58	NaN	NaN
1987	0.58	0.40	0.74	5.39	1.42	NaN
1988	1.28	0.40	1.05	2.72	NaN	NaN
1989	0.37	0.40	1.93	7.20	2.16	NaN
1990	NaN	0.22	0.91	1.52	NaN	NaN
1991	0.40	0.00	1.11	3.66	NaN	NaN
1992	0.00	1.73	0.89	1.32	0.38	0.21
1993	0.10	3.74	1.54	1.06	NaN	0.00
1994	0.00	5.37	1.23	0.95	1.05	1.35
1995	0.02	1.30	1.83	0.71	1.83	0.10
1996	0.07	1.16	0.58	2.50	3.36	NaN
1997	0.01	1.92	0.95	1.71	0.57	1.42
1998	0.00	1.00	0.58	2.25	5.88	10.94
1999	0.02	0.55	1.86	1.87	1.11	2.94
2000	0.19	1.90	1.07	0.81	0.58	0.42
2001	0.00	0.73	1.61	1.84	0.56	0.44
2002	0.05	0.21	2.92	3.46	2.88	16.02
2003	0.12	0.55	1.52	3.21	2.78	2.84
2004	0.18	0.59	1.44	1.07	0.91	5.10
2005	0.00	0.68	0.95	2.76	1.03	NaN
2006	0.19	0.35	2.24	4.35	2.89	4.19
2007	0.13	0.67	1.22	2.18	2.23	2.11
2008	0.00	0.61	1.00	3.25	NaN	NaN
2009	0.69	0.19	0.86	3.04	2.32	4.55
2010	0.07	0.12	0.57	5.22	4.42	NaN
2011	0.12	0.09	0.84	3.30	6.54	10.62
average	0.23	1.09	1.23	2.59	2.58	4.12
median	0.12	0.67	1.09	2.37	2.19	2.84

Table 3b. Relative selectivity between NMFS spring and DFO surveys. Blank cells indicate no data for one or more surveys, zero cells indicate a value of zero for the

numerator (NMFS spring), NaN cells indicate a value of zero in the denominator (DFO).Green cells are >1.1; red cells are <0.9.

	NMFS Spring/DFO					
Year	**1**	**2**	**3**	**4**	**5**	**6**
1973						
1974						
1975						
1976						
1977						
1978						
1979						
1980						
1981						
1982						
1983						
1984						
1985						
1986						
1987	1.87	0.51	0.57	2.20	4.90	9.07
1988	NaN	0.52	0.96	1.32	2.39	3.89
1989	0.54	0.52	1.66	1.31	1.22	2.65
1990	NaN	0.11	1.24	1.64	2.25	30.79
1991	68.24	0.00	0.82	0.70	1.92	5.46
1992	0.00	0.63	1.75	2.19	1.60	1.95
1993	5.81	1.18	1.21	0.77	0.30	0.00
1994	NaN	0.77	1.38	0.89	1.40	1.46
1995	0.21	1.05	1.23	0.65	1.10	0.11
1996	0.19	0.46	0.96	1.67	1.77	1.58
1997	4.18	0.45	1.30	1.20	0.84	0.96
1998	0.00	1.85	0.64	0.79	1.13	1.86
1999	0.87	0.63	1.43	0.98	0.83	0.97
2000	69.38	1.43	1.14	0.62	0.66	0.58
2001	0.00	0.79	1.05	1.22	0.71	1.28
2002	6.27	0.59	1.17	0.98	0.92	0.91
2003	7.72	1.26	0.94	0.89	0.44	1.48
2004	11.01	2.03	0.99	0.55	0.83	1.42
2005	0.00	7.37	1.07	0.69	0.28	0.43
2006	3.44	0.80	0.80	1.39	1.63	1.46
2007	1.92	1.54	0.84	0.82	0.90	1.93
2008	NaN	1.42	0.90	0.91	1.16	15.08
2009	221.64	1.38	1.09	0.82	1.05	0.84
2010	NaN	2.82	1.15	0.80	1.52	1.00
2011	2.62	0.78	1.10	0.92	1.16	0.66
average	25.37	1.29	1.10	1.08	1.32	3.66
median	3.81	0.80	1.09	0.91	1.13	1.46

Table 3c. Relative selectivity between NMFS fall and DFO surveys. Blank cells indicate no data for one or more surveys, zero cells indicate a value of zero for the numerator (NMFS fall), NaN cells indicate a value of zero in the denominator (DFO).Green cells are >1.1; red cells are <0.9.

	NMFS Fall/DFO					
Year	**1**	**2**	**3**	**4**	**5**	**6**
1973						
1974						
1975						
1976						
1977						
1978						
1979						
1980						
1981						
1982						
1983						
1984						
1985						
1986						
1987	3.23	1.27	0.76	0.41	3.45	0.00
1988	NaN	1.30	0.91	0.49	0.00	0.00
1989	1.47	1.31	0.86	0.18	0.57	0.00
1990	NaN	0.52	1.36	1.08	0.00	0.00
1991	172.17	0.64	0.74	0.19	0.00	0.00
1992	27.70	0.36	1.96	1.66	4.17	9.31
1993	58.80	0.32	0.79	0.72	0.00	0.93
1994	NaN	0.14	1.12	0.93	1.32	1.08
1995	11.14	0.81	0.67	0.92	0.60	1.05
1996	2.70	0.40	1.65	0.67	0.53	0.00
1997	355.54	0.24	1.37	0.70	1.46	0.68
1998	2.73	1.86	1.10	0.35	0.19	0.17
1999	54.54	1.13	0.77	0.52	0.75	0.33
2000	356.64	0.75	1.07	0.77	1.14	1.37
2001	41.12	1.07	0.65	0.67	1.28	2.91
2002	138.56	2.81	0.40	0.28	0.32	0.06
2003	65.38	2.30	0.62	0.28	0.16	0.52
2004	60.18	3.44	0.69	0.51	0.91	0.28
2005	4.34	10.88	1.12	0.25	0.27	0.00
2006	17.97	2.31	0.36	0.32	0.56	0.35
2007	14.80	2.28	0.69	0.38	0.41	0.91
2008	NaN	2.31	0.90	0.28	0.00	0.00
2009	319.61	7.46	1.28	0.27	0.45	0.19
2010	NaN	24.32	2.00	0.15	0.34	0.00
2011	22.50	9.11	1.31	0.28	0.18	0.06
average	86.56	3.17	1.01	0.53	0.95	1.26
median	34.41	1.30	0.90	0.41	0.56	0.60

Table 4a. Relative selectivity between NMFS spring survey and the catch. Blank cells indicate no data for one or more surveys, zero cells indicate a value of zero for the

numerator (NMFS spring), NaN cells indicate a value of zero in the denominator (catch).Green cells are >1.1; red cells are <0.9.

	NMFS Spring/Catch					
Year	**1**	**2**	**3**	**4**	**5**	**6**
1973						
1974						
1975						
1976						
1977						
1978						
1979						
1980						
1981						
1982	0.10	1.03	0.75	1.49	2.08	4.71
1983	0.00	1.22	0.86	1.12	0.95	8.26
1984	0.00	0.23	0.91	0.85	2.32	3.20
1985	0.29	1.31	0.65	0.90	0.75	0.00
1986	0.54	1.00	0.92	0.51	2.38	1.71
1987	2.88	0.38	0.72	2.22	4.48	6.56
1988	0.81	0.48	1.23	1.49	4.56	2.67
1989	0.62	0.47	1.71	1.71	1.40	3.72
1990	0.00	0.18	0.94	2.25	4.91	23.67
1991	2.28	0.00	0.59	0.79	2.39	1.01
1992	0.00	0.78	2.23	1.52	0.96	1.43
1993	0.10	2.26	1.63	1.48	0.71	0.00
1994	0.00	2.63	0.57	1.10	1.72	1.57
1995	0.30	1.33	1.18	0.60	1.33	0.09
1996	0.14	0.82	0.72	1.72	1.71	1.19
1997	0.12	0.56	1.11	1.20	1.01	0.81
1998	0.00	1.68	0.50	0.99	1.09	3.86
1999	0.39	0.80	1.20	0.82	0.95	1.20
2000	1.17	1.05	1.13	0.77	0.83	0.83
2001	0.00	0.95	1.11	0.98	0.57	1.02
2002	0.45	0.29	1.81	1.07	1.29	0.91
2003	1.03	0.90	1.16	1.01	0.46	1.01
2004	2.10	2.26	1.81	0.47	0.34	0.49
2005	0.00	1.03	1.01	1.13	0.55	0.50
2006	2.04	0.44	1.32	1.40	0.53	0.23
2007	0.48	0.93	1.01	1.18	1.12	0.64
2008	0.00	1.68	0.97	0.74	0.90	0.41
2009	3.31	0.51	1.45	0.86	0.49	0.67
2010	1.07	0.62	1.04	1.17	0.75	0.80
2011	0.70	0.41	1.15	1.09	0.83	0.37
average	0.70	0.94	1.11	1.15	1.48	2.45
median	0.35	0.86	1.08	1.10	0.99	1.01

Table 4b. Relative selectivity between NMFS fall survey and the catch. Blank cells indicate no data for one or more surveys, zero cells indicate a value of zero for the numerator (NMFS fall), NaN cells indicate a value of zero in the denominator (catch).Green cells are >1.1; red cells are <0.9.

	NMFS Fall/Catch					
Year	**1**	**2**	**3**	**4**	**5**	**6**
1973	13.22	2.02	0.72	0.59	0.61	0.71
1974	5.91	0.91	0.55	0.42	0.38	0.76
1975	5.07	0.48	0.60	0.81	0.77	0.26
1976	7.21	0.80	1.15	1.09	2.83	1.45
1977	10.04	0.99	0.62	1.39	1.02	0.89
1978	1.34	1.01	0.48	0.60	0.70	0.48
1979	17.42	0.83	0.29	0.35	0.94	1.46
1980	2.94	1.56	0.77	0.54	0.76	2.80
1981	30.83	4.44	0.57	0.25	0.19	0.78
1982	5.40	0.58	1.02	0.60	0.39	0.00
1983	0.91	1.67	0.67	1.21	0.27	1.86
1984	10.51	1.62	0.55	0.39	0.34	1.35
1985	4.95	0.44	0.58	0.22	0.56	0.00
1986	7.31	0.80	1.44	0.88	0.00	0.00
1987	4.98	0.93	0.97	0.41	3.16	0.00
1988	0.63	1.20	1.17	0.55	0.00	0.00
1989	1.68	1.18	0.89	0.24	0.65	0.00
1990	0.00	0.86	1.03	1.48	0.00	0.00
1991	5.76	3.40	0.54	0.22	0.00	0.00
1992	0.66	0.45	2.51	1.15	2.49	6.83
1993	0.97	0.60	1.06	1.39	0.00	1.46
1994	49.64	0.49	0.46	1.16	1.63	1.16
1995	15.77	1.02	0.65	0.84	0.73	0.91
1996	2.02	0.71	1.24	0.69	0.51	0.00
1997	10.62	0.29	1.17	0.70	1.77	0.57
1998	7.86	1.69	0.86	0.44	0.19	0.35
1999	24.36	1.45	0.65	0.44	0.86	0.41
2000	6.00	0.55	1.06	0.96	1.44	1.96
2001	11.91	1.29	0.69	0.54	1.02	2.32
2002	9.89	1.38	0.62	0.31	0.45	0.06
2003	8.68	1.65	0.77	0.31	0.16	0.36
2004	11.51	3.82	1.26	0.44	0.37	0.10
2005	5.98	1.51	1.07	0.41	0.54	0.00
2006	10.66	1.26	0.59	0.32	0.18	0.06
2007	3.74	1.38	0.83	0.54	0.50	0.30
2008	1.09	2.73	0.98	0.23	0.00	0.00
2009	4.78	2.74	1.70	0.28	0.21	0.15
2010	16.06	5.34	1.82	0.22	0.17	0.00
2011	5.98	4.83	1.37	0.33	0.13	0.04
average	8.83	1.56	0.92	0.61	0.69	0.76
median	5.98	1.20	0.83	0.54	0.50	0.35

Table 4c. Relative selectivity between DFO survey and the catch. Blank cells indicate no data for one or more surveys, zero cells indicate a value of zero for the numerator (DFO),

NaN cells indicate a value of zero in the denominator (catch).Green cells are >1.1; red cells are <0.9.

	DFO/Catch					
Year	**1**	**2**	**3**	**4**	**5**	**6**
1973						
1974						
1975						
1976						
1977						
1978						
1979						
1980						
1981						
1982						
1983						
1984						
1985						
1986						
1987	1.54	0.73	1.27	1.01	0.92	0.72
1988	0.00	0.92	1.29	1.12	1.91	0.69
1989	1.14	0.90	1.03	1.31	1.15	1.40
1990	0.00	1.66	0.76	1.37	2.18	0.77
1991	0.03	5.28	0.72	1.13	1.24	0.18
1992	0.02	1.24	1.28	0.69	0.60	0.73
1993	0.02	1.92	1.34	1.92	2.36	1.57
1994	0.00	3.40	0.41	1.24	1.23	1.08
1995	1.42	1.27	0.96	0.92	1.20	0.87
1996	0.75	1.77	0.75	1.03	0.97	0.75
1997	0.03	1.25	0.86	1.00	1.21	0.85
1998	2.88	0.91	0.79	1.25	0.96	2.08
1999	0.45	1.28	0.84	0.84	1.14	1.23
2000	0.02	0.73	0.99	1.25	1.26	1.44
2001	0.29	1.21	1.06	0.80	0.80	0.80
2002	0.07	0.49	1.54	1.09	1.39	1.00
2003	0.13	0.72	1.24	1.13	1.04	0.68
2004	0.19	1.11	1.82	0.86	0.41	0.34
2005	1.38	0.14	0.95	1.65	1.99	1.16
2006	0.59	0.55	1.66	1.00	0.33	0.16
2007	0.25	0.60	1.20	1.43	1.24	0.33
2008	0.00	1.18	1.08	0.81	0.78	0.03
2009	0.01	0.37	1.33	1.05	0.47	0.79
2010	0.00	0.22	0.91	1.46	0.49	0.80
2011	0.27	0.53	1.05	1.19	0.72	0.57
average	0.46	1.21	1.08	1.14	1.12	0.84
median	0.13	0.92	1.05	1.12	1.14	0.79

Table 5. Summary of relative selectivity between surveys and catch by different blocks of years. The year range is indicated on the left; e.g. “73-80” refers to the year range 1973-1980. Green cells are >1.1; red cells are <0.9.

	NMFS Spring/Catch					
	1	**2**	**3**	**4**	**5**	**6**
73-80						
81-90	0.58	0.70	0.96	1.39	2.65	6.06
91-00	0.45	1.19	1.09	1.10	1.27	1.20
01-11	1.02	0.91	1.26	1.01	0.71	0.64
73-94	0.59	0.92	1.05	1.34	2.28	4.50
95-00	0.35	1.04	0.97	1.02	1.15	1.33
01-11	1.02	0.91	1.26	1.01	0.71	0.64
	NMFS Fall/Catch					
	1	**2**	**3**	**4**	**5**	**6**
73-80	7.89	1.08	0.65	0.73	1.00	1.10
81-90	6.72	1.37	0.89	0.62	0.56	0.40
91-00	12.37	1.07	1.02	0.80	0.96	1.37
01-11	8.21	2.54	1.06	0.36	0.34	0.31
73-94	8.52	1.24	0.85	0.72	0.80	1.01
95-00	11.11	0.95	0.94	0.68	0.91	0.70
01-11	8.21	2.54	1.06	0.36	0.34	0.31
	DFO/Catch					
	1	**2**	**3**	**4**	**5**	**6**
73-80						
81-90						
91-00	0.56	1.90	0.89	1.13	1.22	1.08
01-11	0.29	0.65	1.26	1.13	0.88	0.61
73-94	0.34	2.01	1.01	1.22	1.45	0.89
95-00	0.92	1.20	0.86	1.05	1.12	1.20
01-11	0.29	0.65	1.26	1.13	0.88	0.61

Table 6. Estimated catchability (q) at age for the three surveys, and relative catchability (selectivity) for the Single Series VPA model and two alternative VPA configurations where the fishery selectivity had some doming (Dome-1 and Dome-2).

q at age						
NMFS Spring survey	**1**	**2**	**3**	**4**	**5**	**6**
Single Series	0.0051	0.0864	0.2702	0.3622	0.3632	0.4075
Dome-1	0.0048	0.0822	0.2551	0.3320	0.3076	0.3436
Dome-2	0.0048	0.0822	0.2552	0.3322	0.3077	0.2935
NMFS Fall survey	**1**	**2**	**3**	**4**	**5**	**6**
Single Series	0.0472	0.1451	0.2823	0.2439	0.2877	0.3031
Dome-1	0.0454	0.1392	0.2668	0.2219	0.2365	0.2515
Dome-2	0.0454	0.1392	0.2669	0.2220	0.2365	0.2012
DFO survey	**1**	**2**	**3**	**4**	**5**	**6**
Single Series	NA	0.2380	0.8528	1.2151	1.0223	0.6578
Dome-1	NA	0.2248	0.7988	1.0986	0.8446	0.5424
Dome-2	NA	0.2249	0.7991	1.0990	0.8450	0.4625
Relative q at age (selectivity)						
NMFS Spring survey	**1**	**2**	**3**	**4**	**5**	**6**
Single Series	0.01	0.21	0.66	0.89	0.89	1.00
Dome-1	0.01	0.24	0.74	0.97	0.90	1.00
Dome-2	0.01	0.25	0.77	1.00	0.93	0.88
NMFS Fall survey	**1**	**2**	**3**	**4**	**5**	**6**
Single Series	0.16	0.48	0.93	0.80	0.95	1.00
Dome-1	0.17	0.52	1.00	0.83	0.89	0.94
Dome-2	0.17	0.52	1.00	0.83	0.89	0.75
DFO survey	**1**	**2**	**3**	**4**	**5**	**6**
Single Series	NA	0.20	0.70	1.00	0.84	0.54
Dome-1	NA	0.20	0.73	1.00	0.77	0.49
Dome-2	NA	0.20	0.73	1.00	0.77	0.42

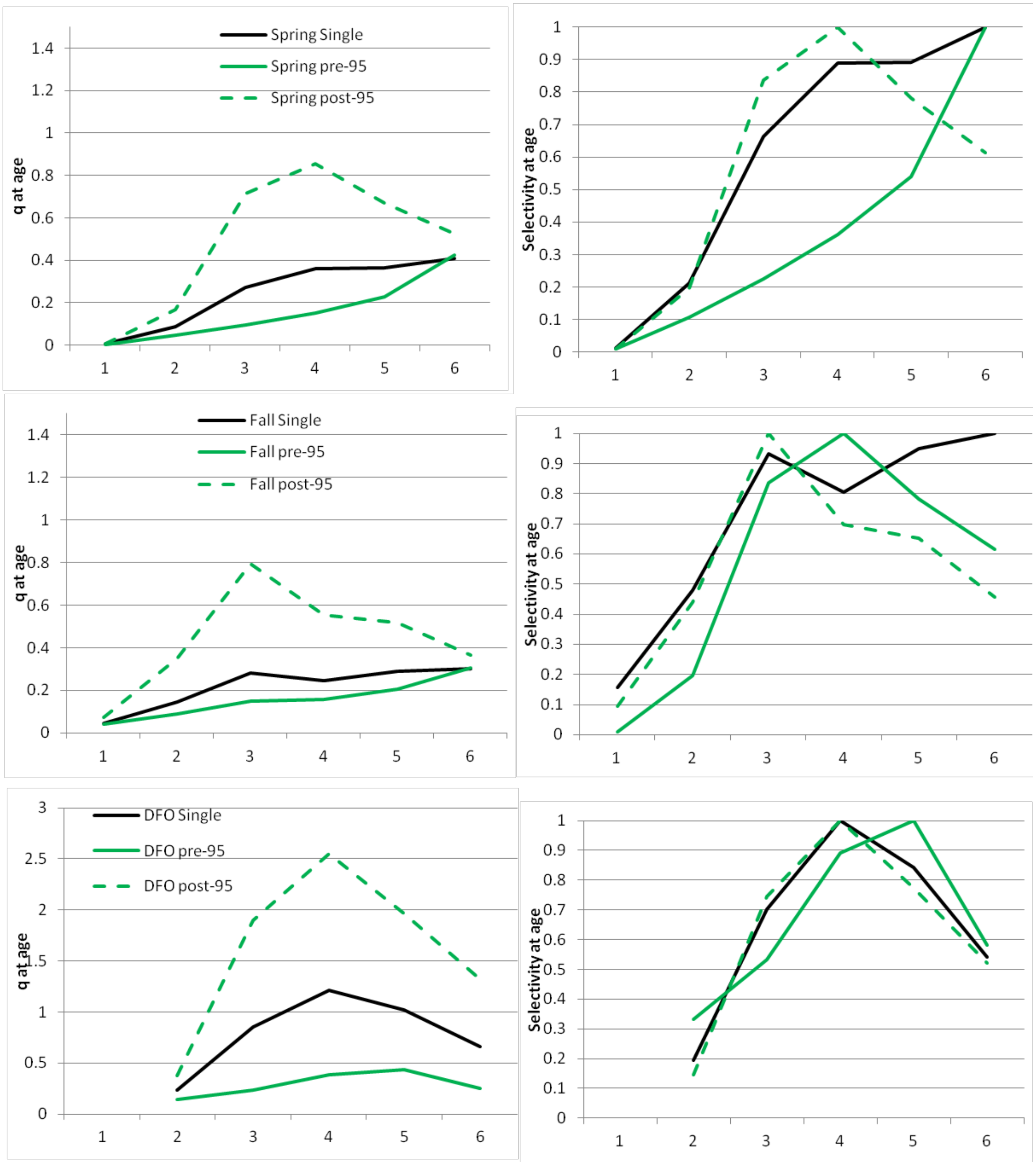

Figure 1. Estimated catchability at age (left) and selectivity (right) from the 2012 VPA for Georges Bank Yellowtail Flounder (Legault et al. 2012). The solid black line is the estimate when the time series are not split. The solid and dashed green are the model estimates pre- and post-1995, respectively. Selectivity is estimated from catchability by scaling q at age by the maximum value of the series.

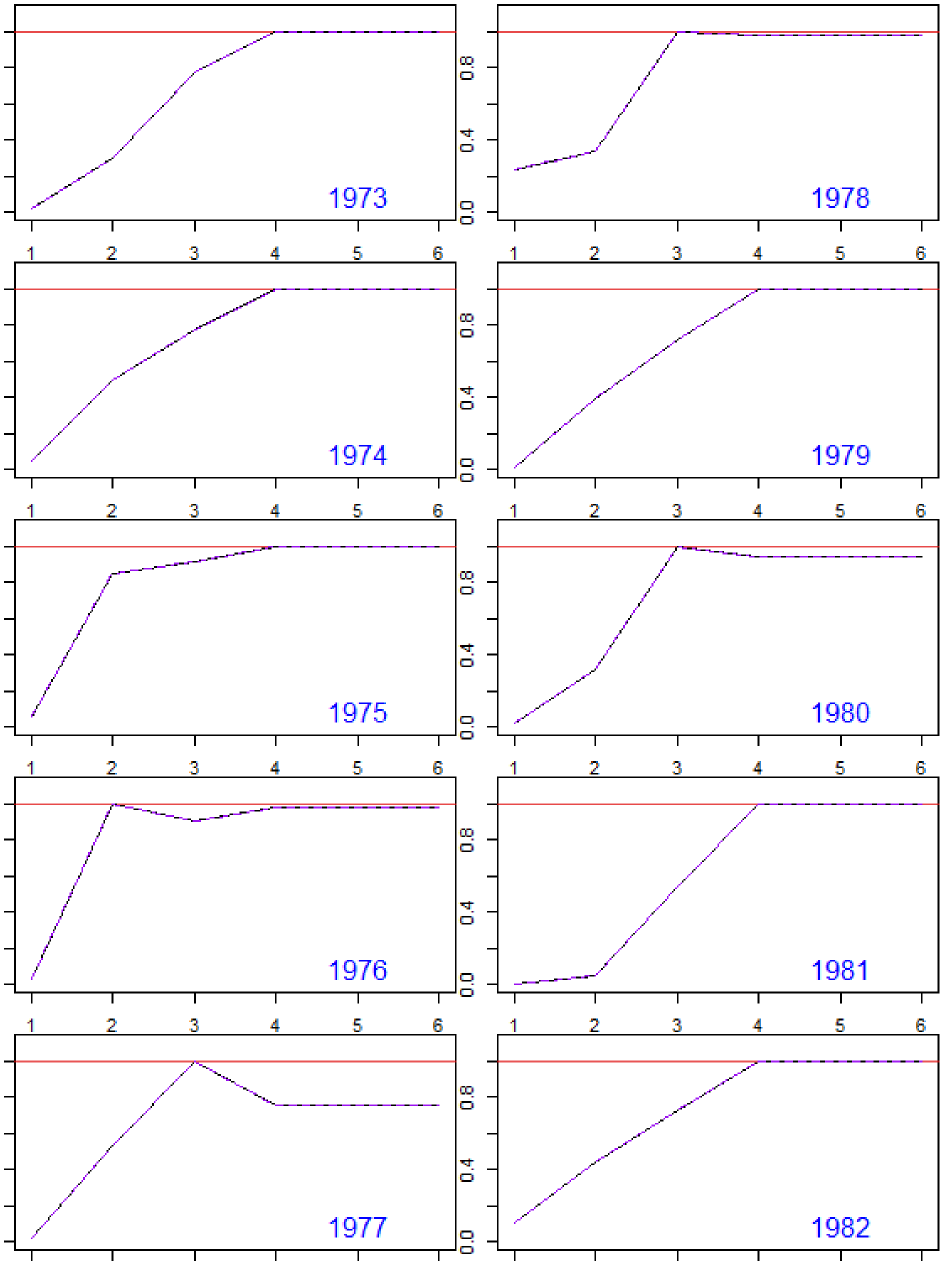

Figure 2a. Selectivity at age in the fishery, for each year, 1973-1982. Solid black line is for the Single Series VPA, while the dashed purple line is the the Split Series VPA.

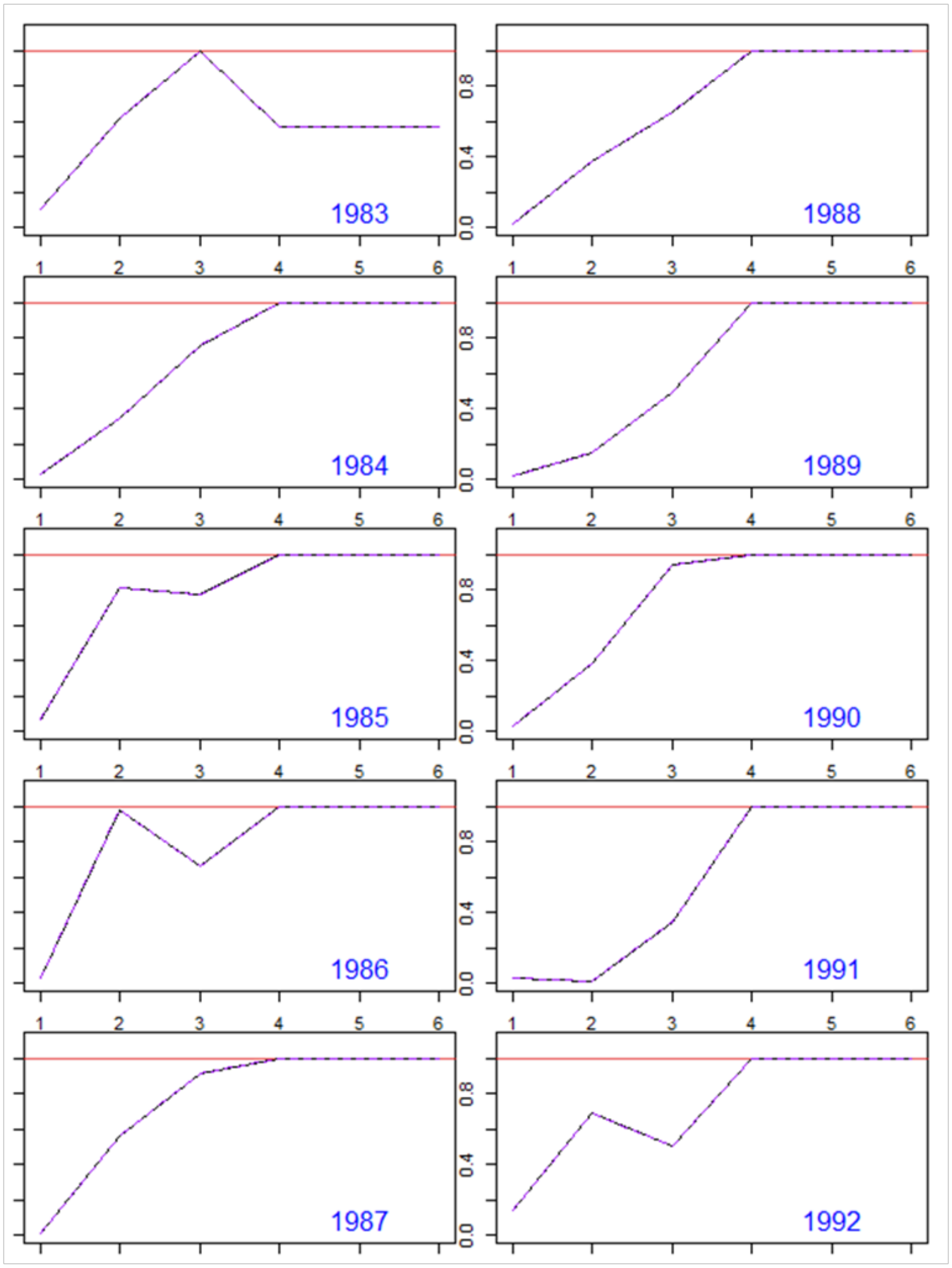

Figure 2b. Selectivity at age in the fishery, for each year,1983-1992. Solid black line is for the Single Series VPA, while the dashed purple line is the the Split Series VPA.

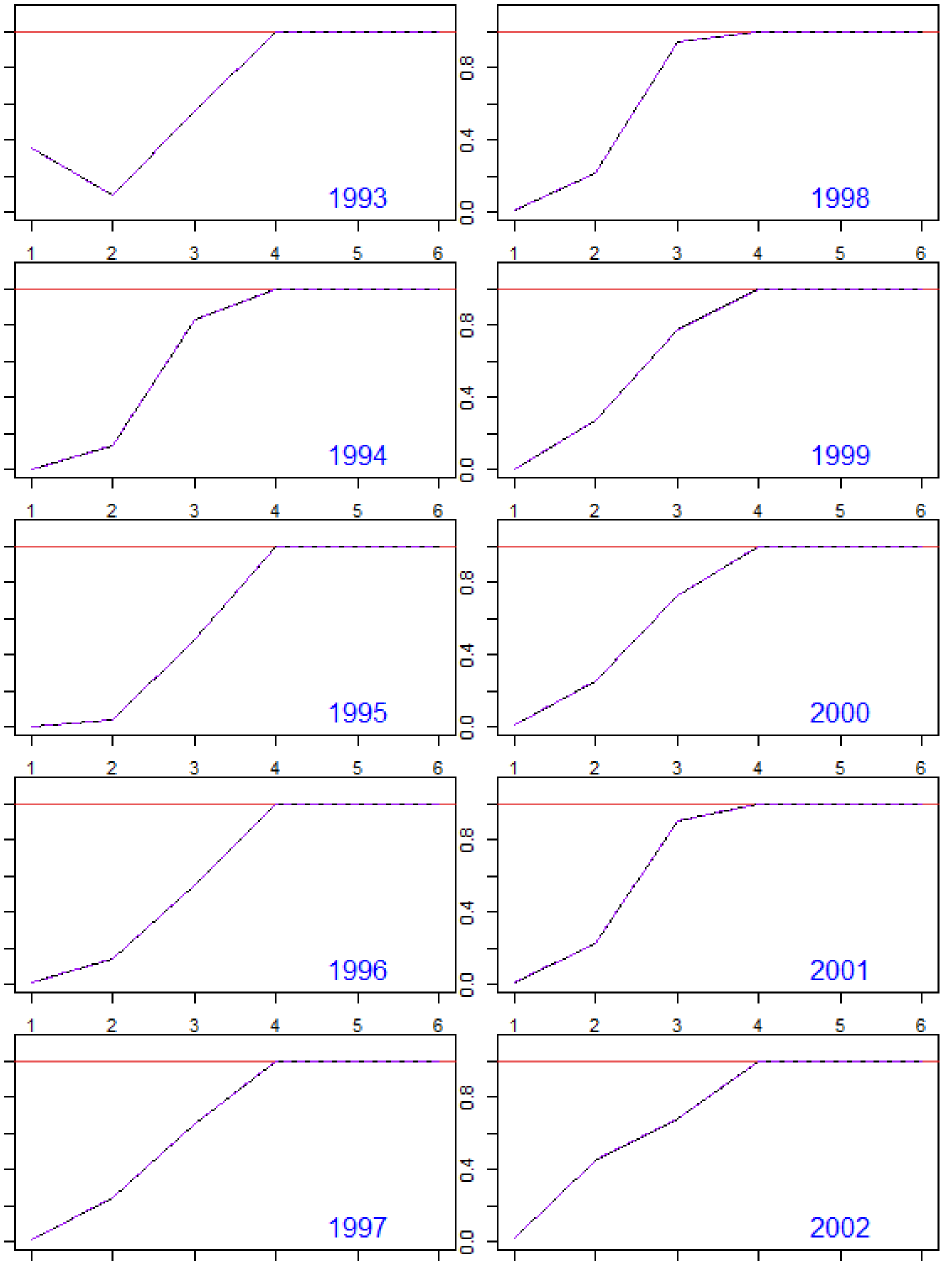

Figure 2c. Selectivity at age in the fishery, for each year. 1993 - 2002. Solid black line is for the Single Series VPA, while the dashed purple line is the the Split Series VPA.

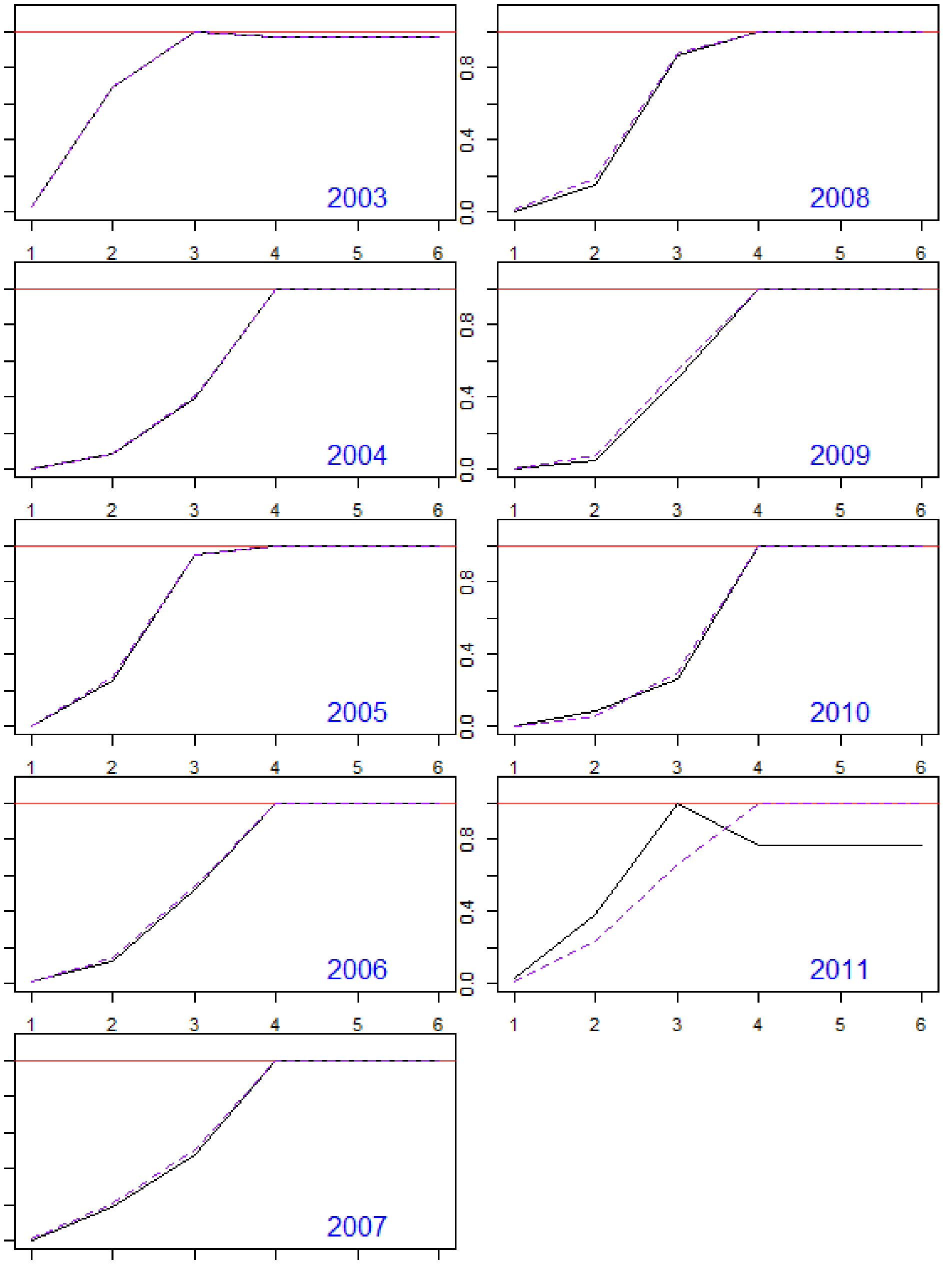

Figure 2d. Selectivity at age in the fishery, for each year, 2003-2011. Solid black line is for the Single Series VPA, while the dashed purple line is the the Split Series VPA.

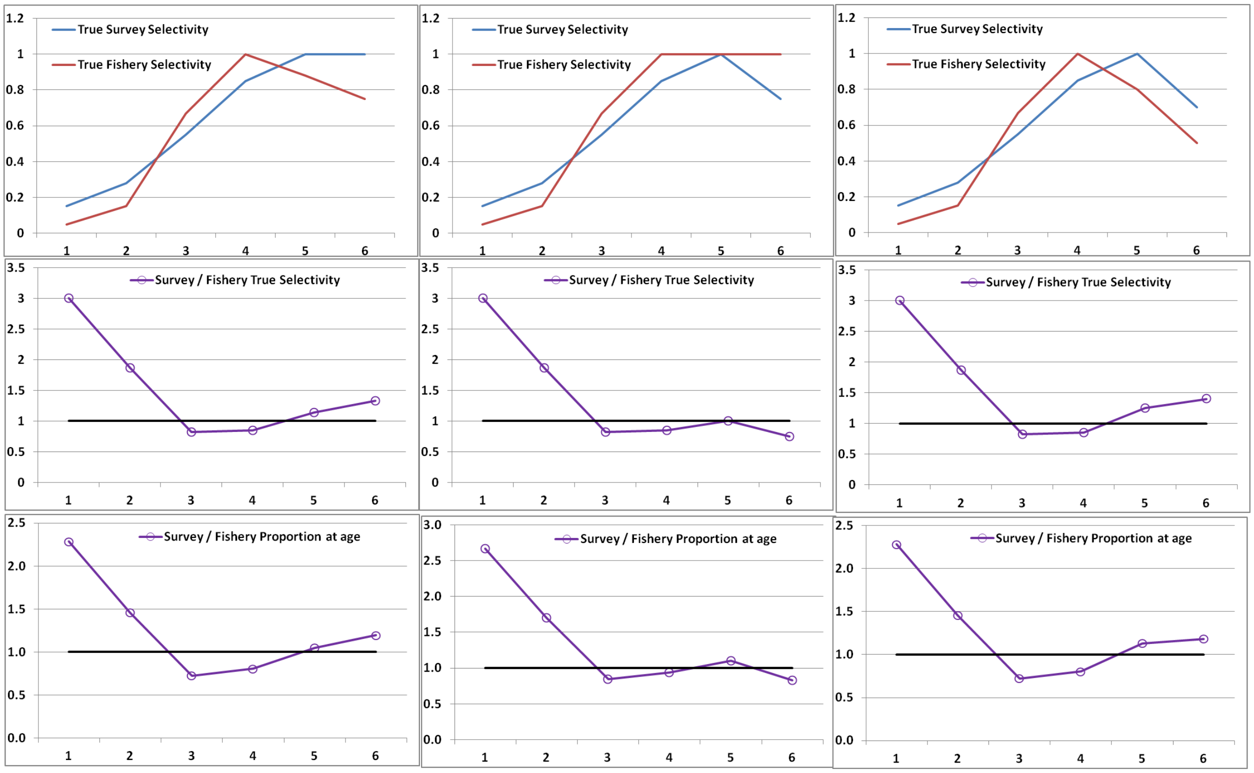

Figure 3. Illustration of relative selectivities for three cases: (left) flat-topped survey selectivity and domed fishery selectivity; (center) domed survey selectivity and flat-topped fishery selectivity; (right) domed selectivity in the fishery and the survey.

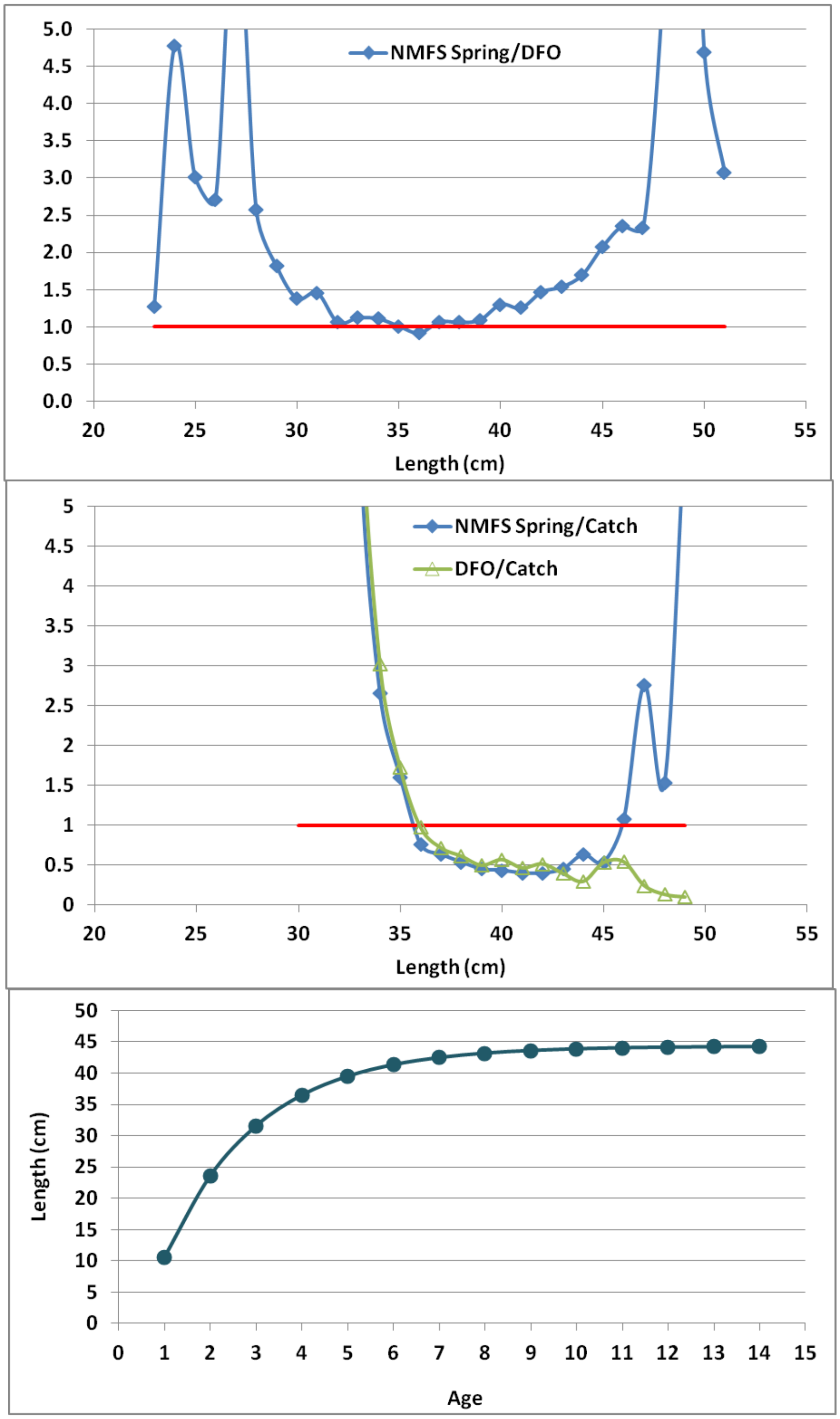

Figure 4. Relative selectivity at length between the NMFS spring survey and the DFO survey (top), between both surveys and the catch (center), and a von Bertalanffy growth curve based on NMFS spring and fall survey data (bottom).

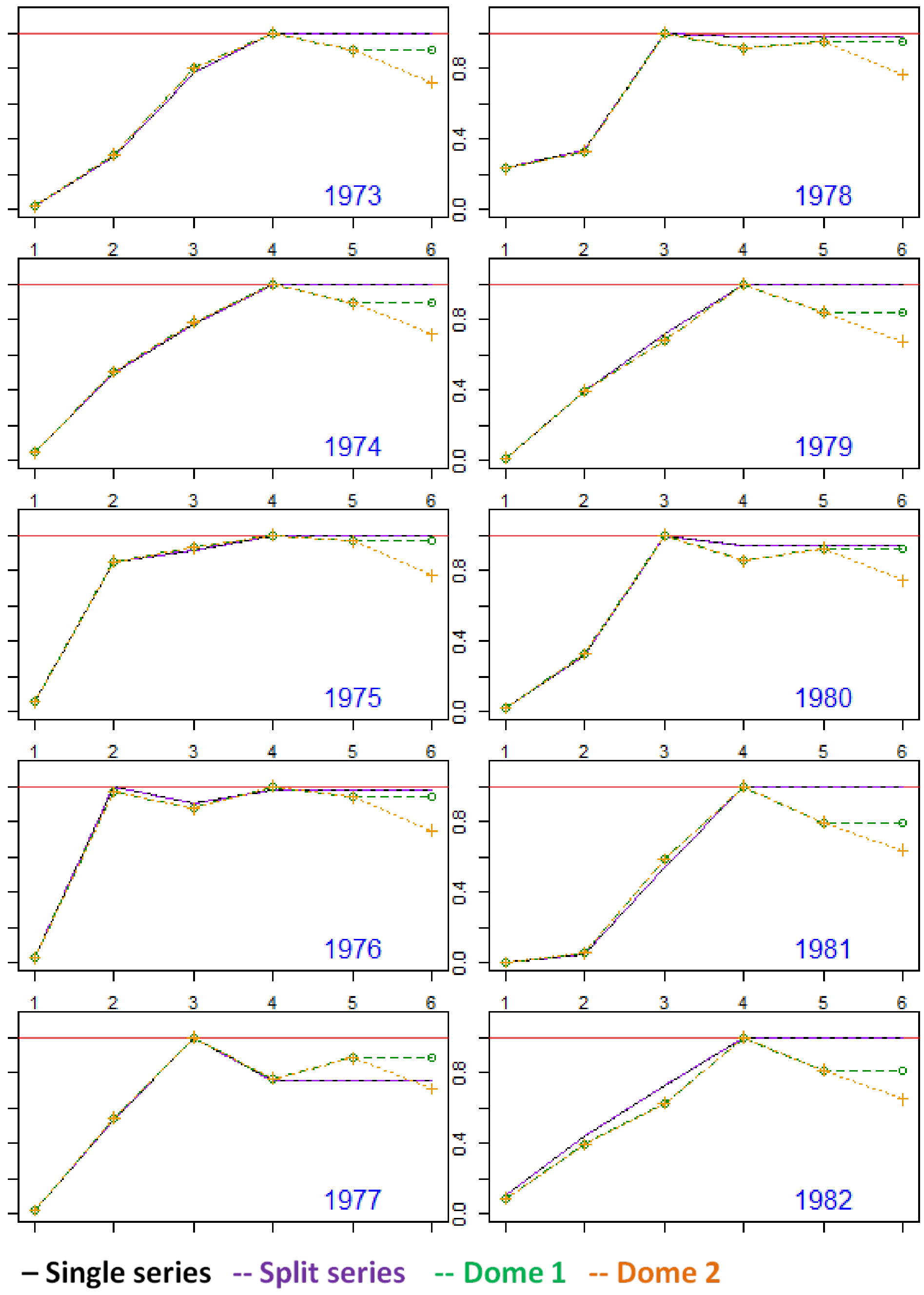

Figure 5a. Selectivity at age in the fishery, for each year, for four different VPA model configurations: Single Series, Split Series, Dome-1, and Dome-2.

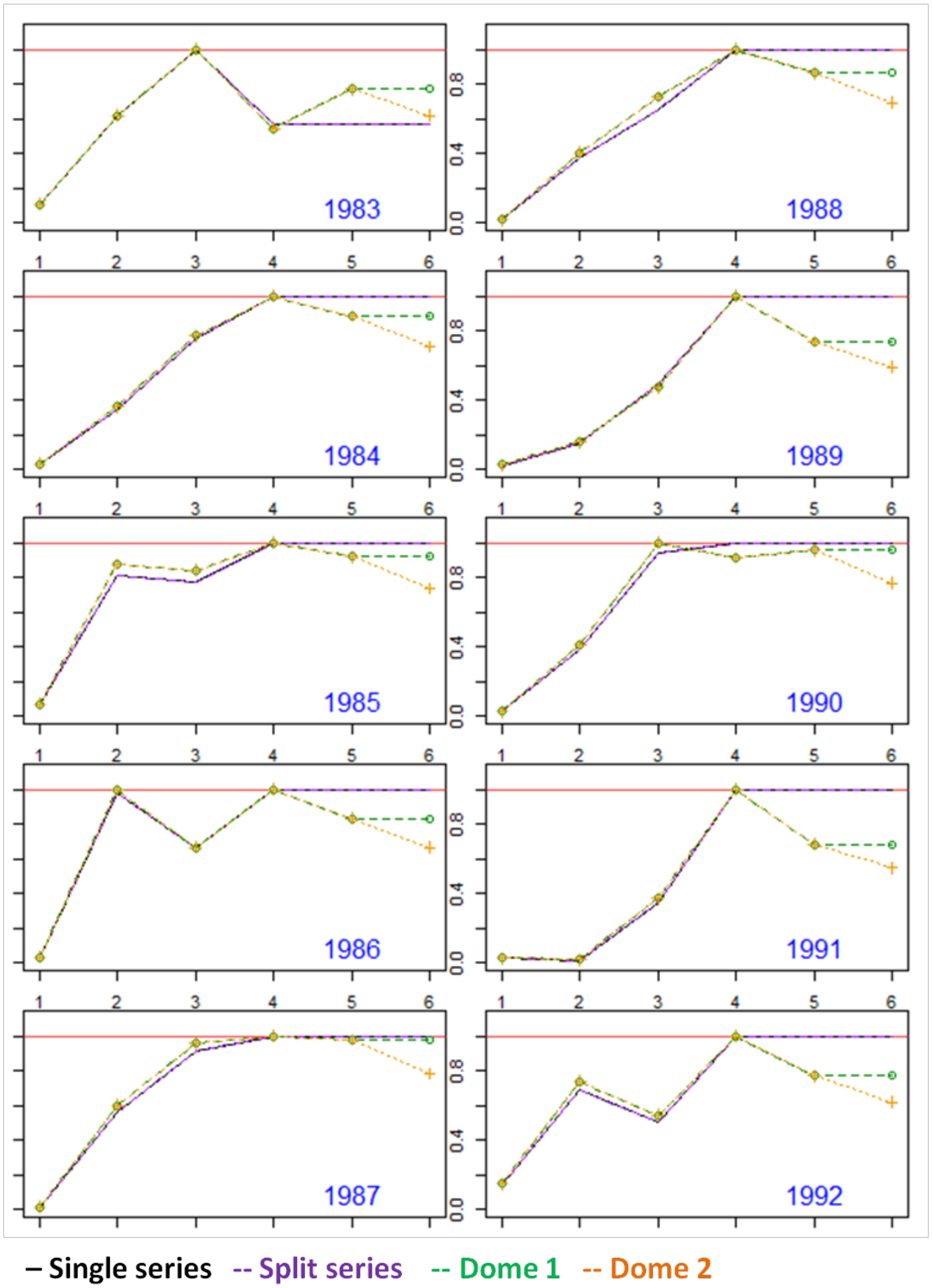

Figure 5b. at age in the fishery, for each year, for four different VPA model configurations: Single Series, Split Series, Dome-1, and Dome-2.

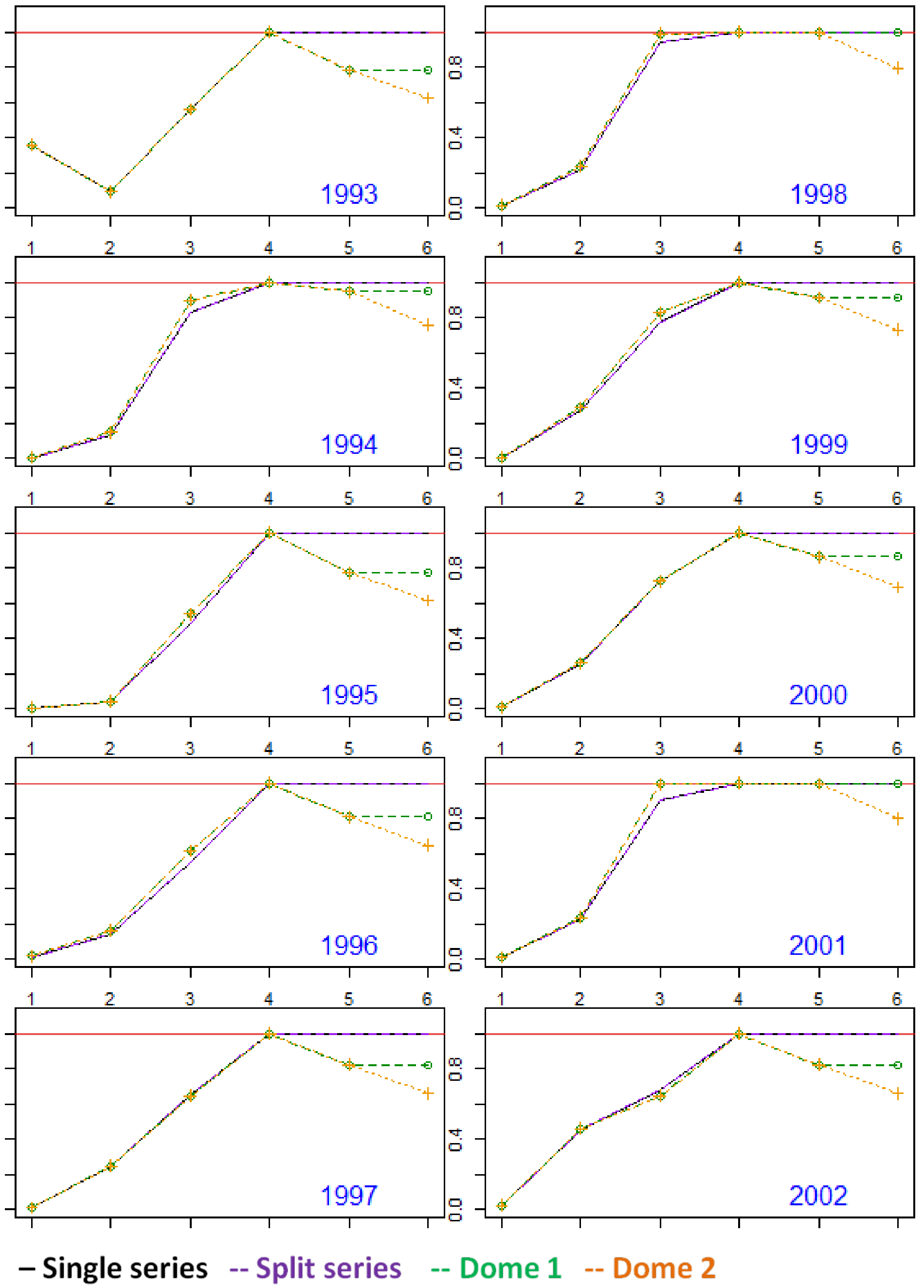

Figure 5c. at age in the fishery, for each year, for four different VPA model configurations: Single Series, Split Series, Dome-1, and Dome-2.

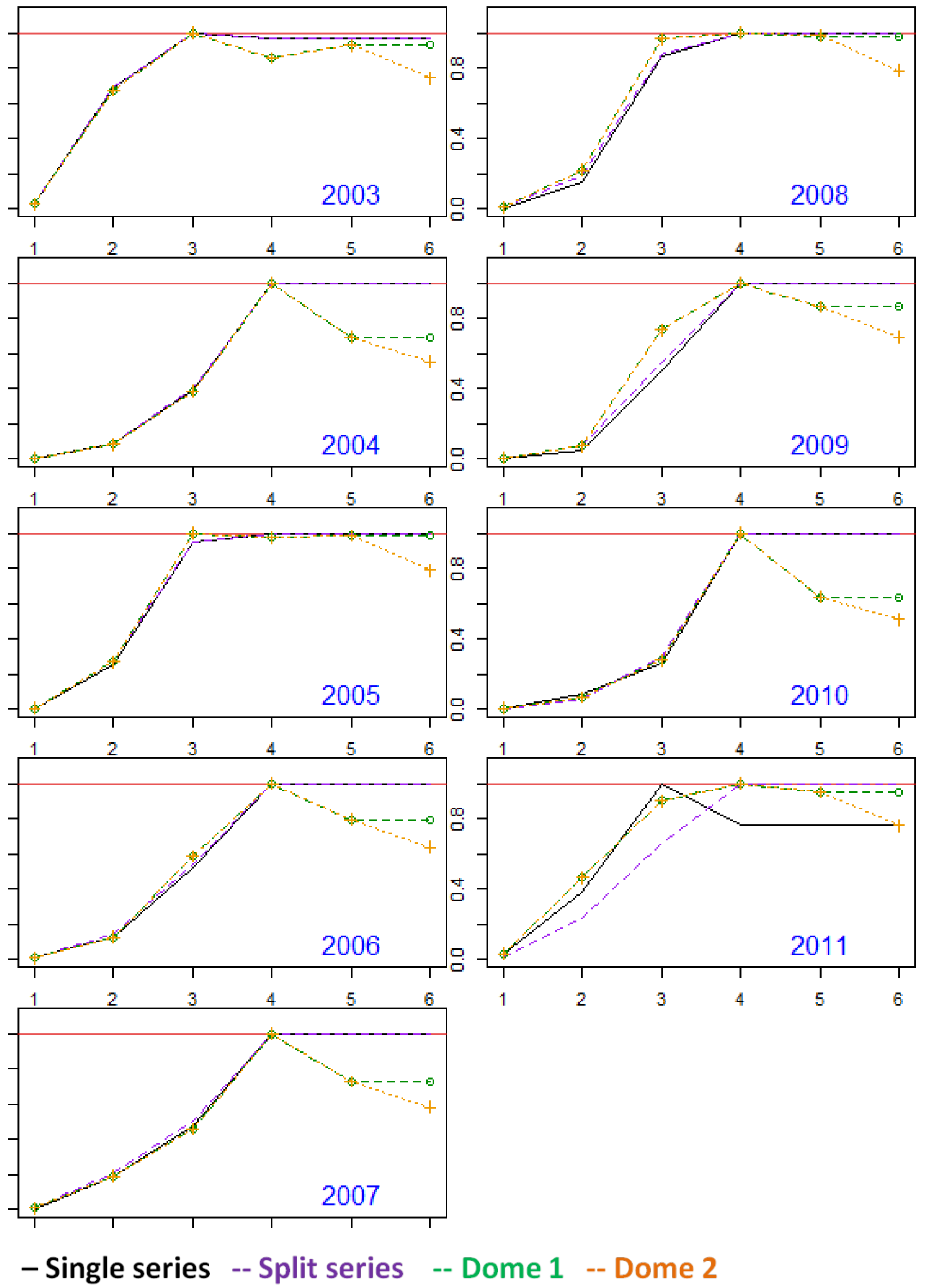

Figure 5d. at age in the fishery, for each year, for four different VPA model configurations: Single Series, Split Series, Dome-1, and Dome-2.

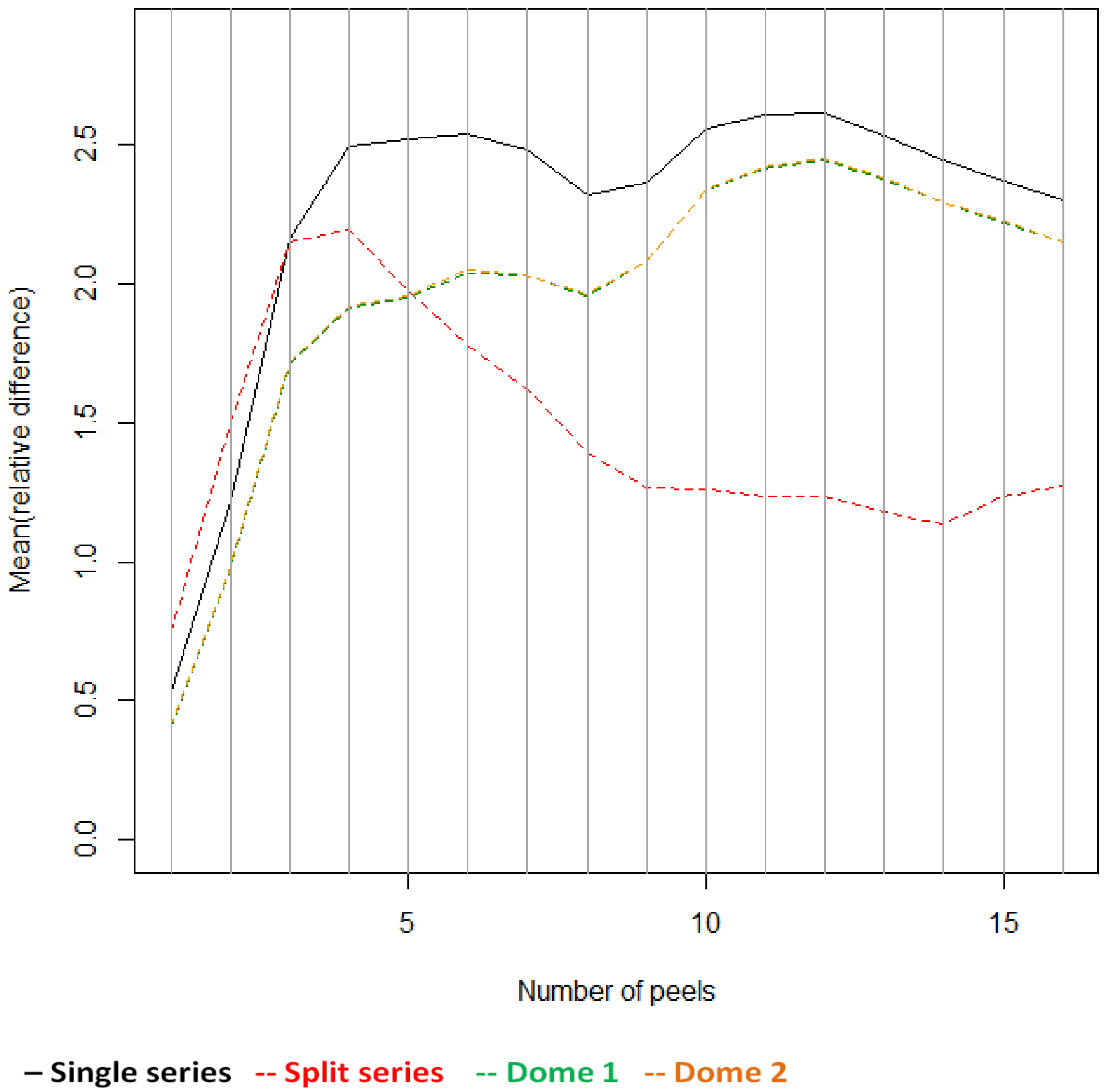

Figure 6. Mohn’s rho calculated for SSB for four different VPA model configurations. Sixteen total peels were made, and a value for Mohn’s rho was calculated for peels of length 1-16.