

Executive Summary

The Groundfish Assessment Review Meeting (GARM) is a regional peer review process developed this year to provide assessment updates for the 20 stocks managed under the Northeast Multispecies Fishery Management Plan (Multispecies FMP). The meeting occurred during October 8-11, 2002, in Woods Hole, Massachusetts. The terms of reference were to:

- (a) provide updated catch information (landings and discards, where appropriate) for the stocks to be assessed. Catch-at-age data (based on port sampling) will be estimated, where applicable,
- (b) provide updated research vessel survey indices (through spring 2002) for all appropriate survey series, including NMFS spring and autumn series, Canadian series, and state surveys,
- (c) estimate 2001 fishing mortality rates (or appropriate proxies) for all 20 stocks, and provide estimates of 2001 stock sizes and measures of uncertainty,
- (d) evaluate stock status relative to applicable biological reference points (F_{MSY} and B_{MSY}),
- (e) provide updated estimates of F-Rebuild (the fishing mortality rate required to rebuild biomasses to B_{MSY} by 2009) for all applicable stocks,
- (f) evaluate and comment on the potential sensitivity of assessment results to trawl warp marking discrepancies that occurred in bottom trawl surveys conducted between winter 2000 and spring 2002.

Initial stock assessments were developed by the Northern and Southern Demersal Working Groups of the Stock Assessment Workshop (SAW), and the ASMFC Winter Flounder Technical Committee. These working groups and the Technical Committee met at various times before the GARM meeting to develop draft assessment documents. Additionally, work related to the trawl warp offset issue was coordinated through the SAW Assessment Methods Working Group.

Most stock assessments reviewed at the GARM were routine updates of assessments previously reviewed in the SAW or elsewhere. However, the Gulf of Maine winter flounder assessment was newly developed (by the ASMFC Technical Committee), and is scheduled to be peer reviewed at SAW 36 (December 2002). Accordingly, the details of the analytical stock assessment modeling are not incorporated herein, pending that “benchmark” review. The results are, however, summarized (Table 1; Figure 1), and input data are presented and evaluated.

The GARM meeting incorporated peer reviews by both regional stock assessment scientists (both NMFS and non-NMFS people) and external experts. The Center for Independent Experts (CIE, University of Miami) provided two individuals for the meeting. The roles of the CIE

experts were to comment on analyses presented at the GARM, and to provide written critiques; attached as appendices to this report.

Stock Assessment Results

Results of the stock assessment updates are summarized as fishing mortality rates and biomasses in 2001, relative to management reference points (Table 1; Figure 1). Of the 19 stocks for which 2001 fishing mortality (or its proxy) can be estimated, 10 were fished below F-MSY in 2001, and 9 above. Additionally, the biomass of eight of the stocks was at or above $\frac{1}{2}$ B-MSY, while 12 stocks were below their biomass thresholds. Stock biomasses have improved in 19 of the 20 stocks since 1995 (the exception being Mid-Atlantic yellowtail flounder), with a median percent increase in biomass for all stocks of 177% (range: -33 to 2430 percent). Landings of the complex of 20 groundfish stocks have increased by 40% since 1995, primarily driven by increases from four Georges Bank stocks (haddock, yellowtail flounder, cod and winter flounder). Fishing mortality (F) rates declined for 15 of 19 stocks between 1994 and 2001. In the case of Georges Bank yellowtail flounder, F has declined by about 90% since the mid-1990s. Numerous other stocks have experienced reductions in F of 20-50%, including Georges Bank and Gulf of Maine cod, Georges Bank haddock, witch flounder and American plaice. For several of the stocks where harvest rates are measured by landings to survey biomass ratios (exploitation index methods), relative Fs have been reduced by 50% or more (e.g., Gulf of Maine haddock, pollock and windowpane flounder). The four stocks showing increases in F since 1994 were Cape Cod and Mid-Atlantic yellowtail flounder, white hake and Southern New England/Mid-Atlantic winter flounder.

Two stocks continue to have extremely high fishing mortality rates (Mid-Atlantic yellowtail flounder and Cape Cod yellowtail flounder). In the former case, assessment scientists will present analyses to SARC 36 recommending that the Mid-Atlantic and Southern New England yellowtail resources be combined. The case of Cape Cod yellowtail flounder remains enigmatic, in that the apparent mortality rates on the stock remain exceptionally high despite the reductions in F seen in co-occurring stocks (e.g., Gulf of Maine cod, and winter flounder). The GARM recommended additional biological studies, including tagging, to better understand the relationships between Cape Cod yellowtail and adjacent stocks of the same species.

For the remaining seven stocks where fishing mortality exceeded F-MSY, the average reduction necessary to reach that level was 52% (range: 37% for Southern New England/Mid-Atlantic winter flounder to 64% for witch flounder). Fishing mortality rates on the two cod stocks were below those projected based on 2000 assessment results. Maximum fishing mortality rates necessary to rebuild the stocks to B-MSY by the target dates (2009 for most stocks) were computed using medium-term projection methodologies. The percent reductions in F necessary to achieve B-MSY by the target dates varied by stock and were primarily dictated by the strength of incoming recruitment. For Gulf of Maine cod, F in 2001 declined, but F-rebuild also declined despite the presence of a strong 1998 year class, because of below average recruitment in 1999 and a poor 2000 year class.

Short-term projections of target TACs for the 2003-2004 fishing year and medium-term projections for calculating F-Rebuilds assumed that F in 2002 (calendar year) would be 85% of that in 2001, based on assumptions provided by the Multispecies Plan Development Team (PDT).

Evidence for Interventions in Trawl Survey Data Due to Warp Offsets

The GARM reviewed the results of a series of 10 different studies to evaluate evidence for an intervention in the NMFS trawl survey data associated with the use of mis-calibrated trawl warps (the wire ropes attaching the trawl doors to the vessel). There were eight affected surveys (winter 2000, 2001 and 2002; spring 2000, 2001 and 2002; and fall 2000 and 2001). Information collected from dockside warp measurements indicated that the warp mis-calibration was related to the initial biased marking of the 50 meter intervals on one warp and was not due to progressive wire stretch. Therefore, the degree of intervention was thought to be approximately equal in all surveys since winter 2000.

Information on the potential effects of the warp offset on trawl survey performance evaluated by the GARM included studies of rates of gear damage over time, calculations of trawl geometry as a function of the warp offsets, by depth; patterns in mean/variance relationships in trawl survey catch data by stock, and depth-at-capture information from pre- and post-warp misaligned cruises. Additionally, the GARM evaluated trends (directional changes from year-to-year) in abundance measures before and after the warp mis-marking. The results from side-by-side trawling experiments conducted by the *Albatross* and *Delaware* to estimate their relative fishing power, conducted before and after the warp mis-marking on the *Albatross* were also considered. Standardized catch-rates from surveys conducted with mismatched warps were compared to survey CPUEs from surveys with comparable spatial and temporal coverage, and unaffected by the problem (e.g., Canadian trawl surveys and USA sea scallop surveys). The GARM also examined evidence for differences in length distributions from survey catches pre- and post warp offset by evaluating the relative size composition in Canadian and USA spring surveys in overlapping survey areas (e.g. eastern Georges Bank). Monkfish size composition data collected on industry-based surveys and the winter 2001 *Albatross* survey were also compared, as were length compositions with data obtained in side-by-side trawling of the *Albatross* and *Delaware* in spring 2002.

The GARM examined information on wing-spread and headrope height measurements from experimental warp offsets as presented at the Trawl Warp Workshop conducted during October 2-3, 2002. These data were collected during the September 25-27 warp experiment. Additionally, the GARM examined video information collected in that same experiment.

It was postulated by gear experts at the Trawl Warp Workshop that the warp offset would induce changes in gear efficiency resulting from the “long” trawl wing being more prone to damage (as it would be potentially more susceptible to hang-ups). The GARM found no significant change in the frequency of trawl tows experiencing minor or major damage associated with the warp offset as compared to previous surveys.

It was postulated at the Trawl Warp Workshop that one effect of misaligned warps might be the differential loss of large fish in survey catches. Based on examinations of size distributions of cod and haddock, not only was there little difference in the proportions of large fish but there was little apparent difference in the entire size frequency, by survey series, of these stocks pre- and post-warp offset in comparisons of USA and Canadian survey series in areas they overlap (northeast Georges Bank). The small relative differences in USA mean length distributions of cod and haddock for the three years before and three years after the warp offset were similar to the differences in the Canadian series in pre- and post-warp periods. Differences in the size composition of large monkfish between industry and *Albatross* winter surveys were minimal. Size compositions from *Albatross-Delaware* paired towing experiments in spring 2002 also indicated no loss of large fish due to the *Albatross* warp mis-marking.

Trawl mensuration data indicate that wing spread and head rope height did not vary appreciably with offsets that occurred in depths where groundfish typically occur (warp offset up to about 9 feet), and the net remained open with warp offsets up to 18 feet. Consistent trawl performance within this range of warp offsets is supported by the absence of detectable effects as indicated by the other information reported herein. The GARM noted that catching efficiency might be related to other factors such as bottom contact by the foot rope and vibrations associated with the offset gear. Video information on the former was equivocal (as concluded at the Trawl Warp Workshop where some participants thought the foot rope contact changed with offsets while others did not). Measurements on vibrations and pressure waves in relation to warp offsets were not made.

Calculations based on geometry of the trawl in the offset condition (a worst-case scenario) and the postulated increase in the potential problem in relation to species catches-at-depth indicate that reductions on the order of 50% in trawl survey catches are implausible.

It was postulated by the GARM that if there were a trawl warp effect, more variable catches might result from a misaligned net, influencing the relationship between the variance and the mean. Empirical plots of catch data indicated no apparent differences in the variance compared to mean relationships for the species examined, and plots of the coefficient of variation (standard deviation divided by the mean) of catches in numbers by survey stratum over time showed no obvious differences pre- and post warp offsets.

Since the warp offset increased proportionally with depth, it was postulated that if the catch efficiency of the trawl decreased accordingly, then this would result in a shallower apparent depth of capture for the deeper-dwelling species in the post-offset period as compared with the pre-offset surveys. There were no detectable differences in the catch-weighted depth of capture of any species examined relative to the warp offset, however (Figure 2).

There was no evidence for a trend in the direction of abundance index changes associated with the warp offset, when comparing pairs of adjacent years. For each pair of years (e.g., 1998 vs. 1999, 1999 vs. 2000, etc.), the direction of the abundance index change was evaluated. While the evaluation of the changes in abundance indices is potentially confounded by underlying changes in resource abundance, the number of stock/index combinations showing positive

increases in abundance was virtually identical between 1998-1999 and 1999-2000 (when the intervention was made). The abundance indices for the deepest dwelling stocks did not show differential reductions between years pre- and post-warp offsets.

Albatross trawl survey data were compared to independent surveys conducted by other vessels (e.g. Canadian trawl survey and sea scallop dredge surveys aboard *Albatross* but using a single warp). The frequency of species showing positive relative changes in abundance in *Albatross* surveys was nearly the same in the three years before (50%) and the three years after (54%) the warp change. For all species, the relative fishing power of *Albatross* post-warp change was slightly, but not statistically significantly, greater than the comparison vessels.

In examining the stock assessments, there was no obvious improvement in VPA residual patterns (e.g., reduced serial correlation) or tightness of the fit when trawl survey catches were arbitrarily increased by 10%, 25% and 100%. In fact, VPA model fits showed, on average, a 4% decrease in model fit when survey indices in 2000-2002 were arbitrarily increased by 100%. Similarly, retrospective patterns that occur in some VPA models persisted even with the arbitrarily increased survey catches. The stock assessment models integrate catch-at-age information and the full time series from the surveys, thereby damping the influence of variation in recent survey indices.

Fishing power studies were conducted between the *Albatross* and the *Delaware* in 2002 (after the warp change on the *Albatross*) and in 1982, 1983, and 1988. Estimates of fishing power coefficients (ratio of *Albatross* to *Delaware* catches) were similar between vessels in experiments before and after the warp change on the *Albatross IV* (Figure 3). There was only one statistically significant change in this ratio after the warp change in 10 species examined. In this one case, the ratio of *Albatross* to *Delaware* catch of yellowtail flounder increased between the 1980s and 2002. These paired comparison tests (although not intended for that purpose at the time) provide a robust means to test the warp effects (and include any other systematic changes in the fishing system since 1988). Specifically, because these paired trawl studies were conducted simultaneously before and after the warp offset they are not confounded by underlying changes in the abundance of the groundfish stocks. Based on information from 2002, the catch ratio test can detect differences of between 12% and 35%, depending on species. Therefore hypothesized large reductions (greater than 40-50%) in catchability of the *Albatross* survey during the period of the warp offset are highly unlikely. For all species combined, the ratio of *Albatross-Delaware* catches was 0.88 before the warp offset and 0.91 after, suggesting negligible change.

Based on the evidence cited above, there is no indication of a systematic reduction in trawl survey fish catch efficiency due to the trawl warp offsets.

Sensitivity of Stock Assessment Calculations to Potential Warp Offsets

Given the absence of measurable intervention effects associated with the warp offsets, the GARM endorsed the nominal assessment calculations as the basis for management decision-making. However, in order to examine the robustness of the management advice to potential

variations in the survey catches, the GARM also carried out a series of sensitivity analyses examining survey catchability.

Sensitivity runs conducted for the various assessments included arbitrary increases in trawl survey catches for affected surveys of 10%, 25% and 100%. The first two scenarios consider decreases in survey catch rates that are at or below the limits of detection of the analyses of offset effects carried out at the GARM. The 100% increase is not supported by results of analyses carried out at the meeting, the increase is only included for illustrative purposes. An effect of this magnitude would likely have been detectable in the various exploratory data analyses. It should be noted that these arbitrary increases in survey catches were used in assessment calculations across all species, including those found in shallow depths (and thus less likely to be negatively influenced by warp offsets, e.g., yellowtail flounder, winter flounder, windowpane flounder).

The confidence intervals from the +10% and +25% sensitivity runs overlapped the nominal assessment results for all stocks, thus changes of this magnitude have no statistically significant impact on estimates of F and SSB. The stock assessment models integrate unaffected catch information from commercial and recreational fisheries and the full time series from the research vessel surveys, reducing the influence of variations in recent survey indices.

In only three of 20 stocks did the qualitative status determination for overfished (i.e., $B < 1/2 B_{MSY}$) change from overfished to not overfished by adding arbitrary increases in survey abundance indices (Table 2). In two cases (American plaice, and Gulf of Maine haddock), the stocks were near $1/2 B_{MSY}$ based on nominal assessment results. In these cases the hypothesized 10% increases in survey catches were sufficient to change biomass status determination. Of the 18 other cases, arbitrary increases in recent survey catches of 100% changed only the biomass status for white hake (from overfished to not overfished).

The status determination with respect to overfishing (fishing mortality rate) did not change under this sensitivity analysis in 19 of 20 stocks. The only instance of a change from 'overfishing' to 'not overfishing' was for Southern New England yellowtail flounder under the assumption of a 100% increase in survey catchability.

The overall management advice is robust to variations in recent survey catch rates.

Recommendations

The GARM evaluated the level of port sampling used for catch-at-age estimation for all stocks assessed with age-based models. Port sampling provides samples of the length distribution of landings (by market category), and sub-samples for age determination. Overall, the level of port sampling increased in 2000 and 2001 as compared to previous years, in some cases substantially. For example, for several of the most important stocks (Georges Bank and Gulf of Maine cod, Georges Bank and Southern New England yellowtail flounder), the numbers of samples/lengths/ages obtained from the ports in recent years were as follows:

Stock	1999			2000			2001		
	Samples	Lengths	Ages	Samples	Lengths	Ages	Samples	Lengths	Ages
GM Cod	15	1305	350	61	4687	1300	113	7326	2436
GB Cod	68	5987	1503	155	12219	2951	108	8389	2389
GB YTF	11	3066	300	11	3678	605	30	3768	814
SNE YTF	9	834	333	28	1146	984	18	1454	1224

Sustaining relatively high levels of port sampling is considered a priority for these assessments.

The GARM considered short-and medium-term projection methodologies used to estimate target TACs and F-rebuild. In general, it was concluded that all sources of uncertainty are not adequately addressed in such projections, and the GARM recommended a retrospective analysis to evaluate the performance of past projections.

The GARM was concerned about the adequacy of sea sampling to estimate discarded portions of the catch-at-age. Increased sea sampling coverage, initiated in 2002, should allow more precise estimation of discards for inclusion in catch at age estimates.

Numerous recommendations and comments pertaining to individual assessments are provided in the stock-specific chapters of the report.

Table 1. Summary of fishing mortality rate and biomass status for 20 Northeast groundfish stocks in 2001. Projections of maximum F to achieve B-MSY (F-Rebuild) assume F in 2002 = $0.85 * F$ in 2001, and stocks should be rebuilt by 2009, unless otherwise noted.

Species	Stock	F-MSY	F-2001	% F Reduction to achieve F-MSY	F-Rebuild	% F Reduction to achieve F-Rebuild	B-MSY ('000 mt)	B-2001 ('000 mt)	B-2001 % of B-MSY
Cod	GM	0.23	0.47	51	0.11	76	82.8	22.0	27
	GB	0.18	0.38	53	0.15*	61	216.8	29.2	14
Haddock	GM	0.23+	0.12	none	0.20	none	22.17#	10.31	47
	GB	0.26	0.22	none	0.20	10	250.3	74.4	30
Yellowtail	CC	0.21	1.97	89	0.12	94	8.4	1.9	23
	GB	0.25	0.13	none	0.22	none	58.8	38.9	66
	SNE	0.27	0.46	41	0.10**	78	45.2	1.9	4
	MA	0.33+	2.17	85	0.30	86	12.91#	0.21	2
Witch Flounder		0.16	0.45	64	-	none	19.9	11.3	57
American Plaice		0.17	0.43	60	0.10	77	28.6	13.8	48
Winter Flounder	GM	0.26	0.14	none	-	none	5.4	5.37	99
	GB	0.32	0.25	none	-	none	9.4	9.8	104
	SNE-MA	0.32	0.51	37	0.12	76	30.1	7.6	25
White Hake		0.55+	1.36	60	0.50	63	7.70#	2.35	31
Pollock		5.88+	3.55	none	4.83	none	3.0#	1.60	53
Redfish		0.04	0.01	none	0.01***	none	236.7	119.6	51
Ocean Pout		0.31+	0.007	none	n/a	n/a	4.90#	2.46	50
Windowpane	Northern	1.11+	0.1	none	-	none	0.94#	0.79	84
	Southern	0.98+	0.69	none	0.73	none	0.92#	0.21	23
Atlantic Halibut		0.06	unknown	unknown	unknown	unknown	5.4	0.2	4

+ = fishing mortality rate proxy is catch divided by the survey abundance index

= biomass target based on survey abundance index

* = rebuilding period is 2019 for GB cod

** = the SNE YT stock cannot be rebuilt to long-term biomass target by 2009 even if $F=0.0$ (using recruitment from last 10 years)

*** = rebuilding period is 2051 for redfish

Table 2. Summary of status determinations for 20 New England groundfish stocks. Sensitivity of status determination to arbitrary increases in trawl survey abundance indices for 2000 to spring 2002 are given for three levels of increase (+10%, +25% and +100%). Overfishing refers to the current fishing mortality rate relative to F-MSY. Overfished refers to the current biomass relative to B-MSY. Asterisks (*) indicate cases where the 80% bootstrap confidence interval for a particular criterion does not overlap that from the nominal assessment run. Shaded cells are where status determination changes from the nominal assessment when survey catch data are increased. SSB is spawning stock biomass, TSB is total stock biomass.

Species	Stock	Status Criterion	Nominal Status	Status +10%	Status +25%	Status +100%
Atlantic Cod	Gulf of Maine	F	overfishing	overfishing	overfishing	overfishing
		SSB	overfished	overfished	overfished	overfished *
	Georges Bank	F	overfishing	overfishing	overfishing	overfishing *
		SSB	overfished	overfished	overfished	overfished *
Haddock	Gulf of Maine	F	no overfishing	no overfishing	no overfishing	no overfishing
		TSB	overfished	not overfished	not overfished	not overfished
	Georges Bank	F	no overfishing	no overfishing	no overfishing	no overfishing
		SSB	overfished	overfished	overfished	overfished *
Yellowtail Flounder	Cape Cod	F	overfishing	overfishing	overfishing	overfishing
		SSB	overfished	overfished	overfished	overfished
	Georges Bank	F	no overfishing	no overfishing	no overfishing	no overfishing *
		SSB	not overfished	not overfished	not overfished	not overfished *
	S. New England	F	overfishing	overfishing	overfishing	no overfishing *
		SSB	overfished	overfished	overfished	overfished *
	Mid-Atlantic	F	overfishing	overfishing	overfishing	overfishing
		TSB	overfished	overfished	overfished	overfished
Witch Flounder	F	overfishing	overfishing	overfishing	overfishing *	
	SSB	not overfished	not overfished	not overfished	not overfished *	
American Plaice	F	overfishing	overfishing	overfishing	overfishing	
	SSB	overfished	not overfished	not overfished	not overfished *	

Table 2 (continued).

Species	Stock	Criterion	Nominal	+10%	+25%	+100%
Winter Flounder	Gulf of Maine	F	no overfishing	no overfishing	no overfishing	no overfishing
		SSB	not overfished	not overfished	not overfished	not overfished
	Georges Bank	F	no overfishing	no overfishing	no overfishing	no overfishing
		TSB	not overfished	not overfished	not overfished	not overfished
	S. New England-Mid-Atlantic	F	overfishing	overfishing	overfishing	overfishing *
		SSB	overfished	overfished	overfished	overfished *
White Hake		F	overfishing	overfishing	overfishing	overfishing
		SSB	overfished	overfished	overfished	not overfished
Pollock		F	no overfishing	no overfishing	no overfishing	no overfishing
		TSB	not overfished	not overfished	not overfished	not overfished
Acadian Redfish+		F	no overfishing	no overfishing	no overfishing	no overfishing
		SSB	not overfished	not overfished	not overfished	not overfished
Ocean Pout		F	no overfishing	no overfishing	no overfishing	no overfishing
		TSB	Not overfished	not overfished	not overfished	not overfished
Windowpane Flounder	Northern	F	no overfishing	no overfishing	no overfishing	no overfishing
		TSB	not overfished	not overfished	not overfished	not overfished
	Southern	F	no overfishing	no overfishing	no overfishing	no overfishing
		TSB	overfished	overfished	overfished	overfished
Atlantic Halibut		F	unknown	unknown	unknown	unknown
		SSB	overfished	overfished	overfished	overfished

+ = Assessment models were not updated for Acadian redfish
 unknown = estimates of F or proxy are not available for Atlantic halibut

Groundfish Stock Status - 2001

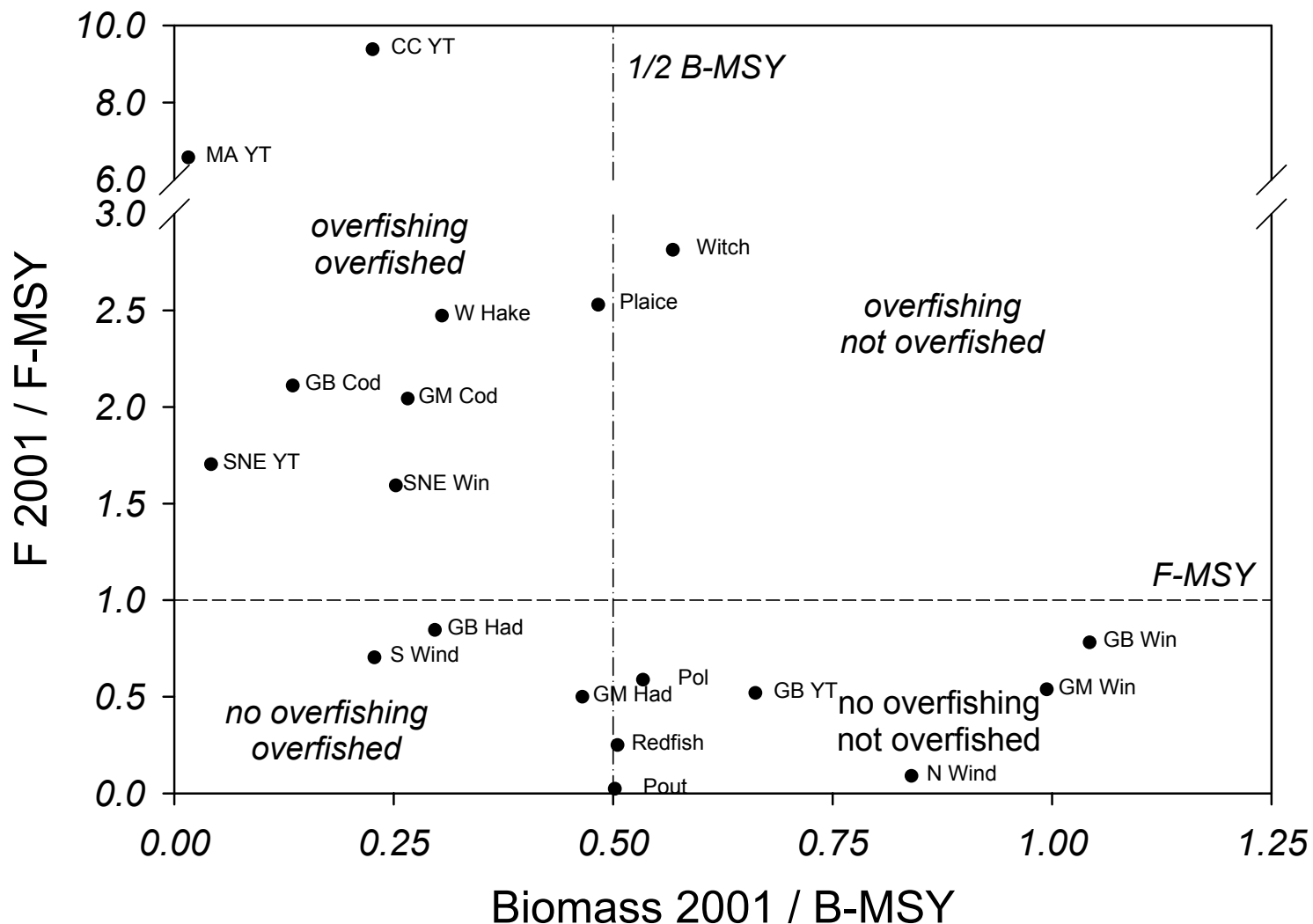


Figure 1. Fishing mortality rate (F) and biomass status of 19 groundfish stocks managed under the Northeast Multispecies FMP. Fishing mortality and biomass in 2001 are expressed as a proportion of F-MSY and B-MSY. Status determination statements are given for each quadrant: overfishing refers to fishing mortality greater than F-MSY; overfished refers to biomasses < 1/2 B-MSY.

Median Catch-Weighted Average Depths: '63-99 v '00-02

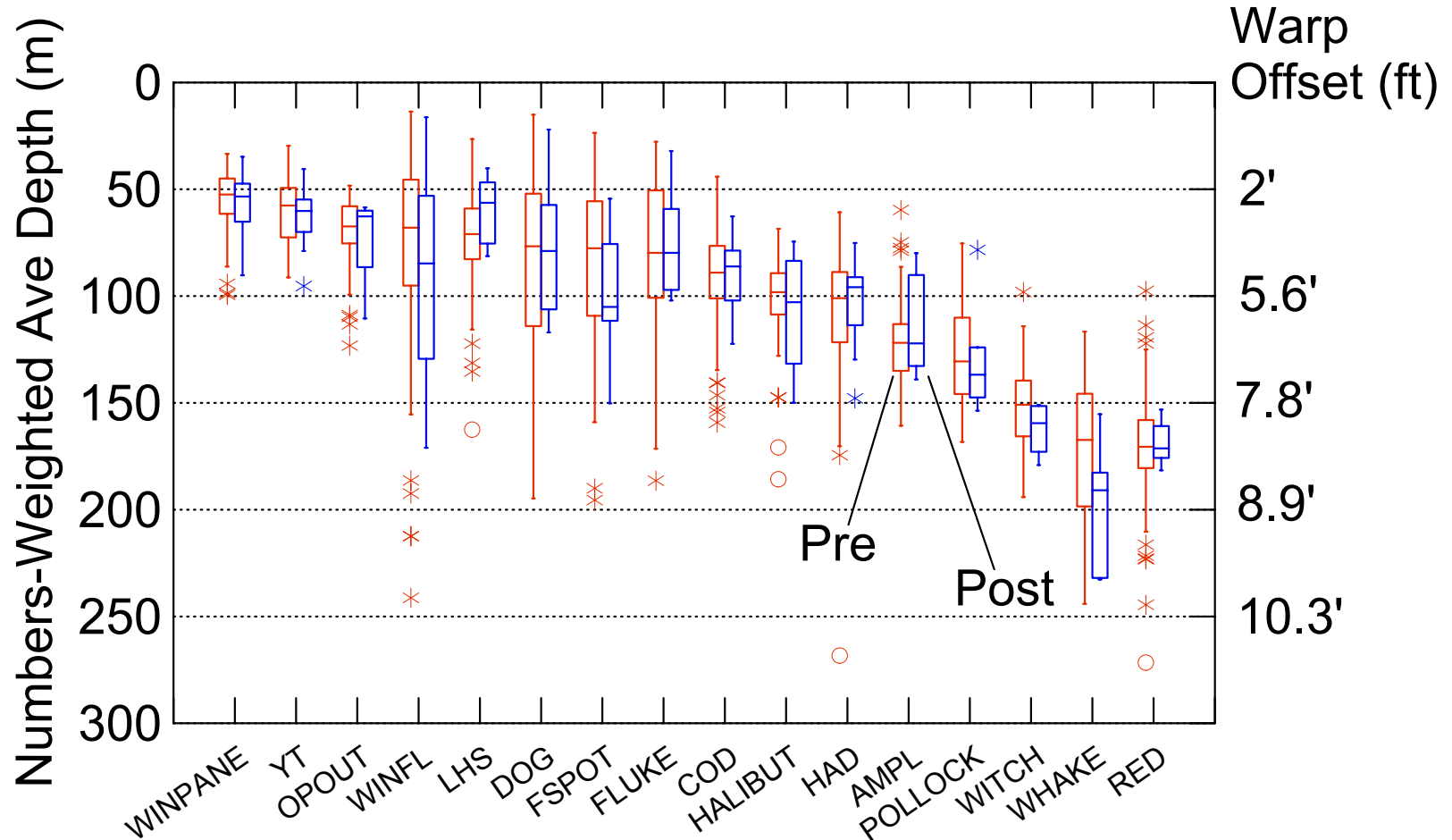


Figure 2. Median catch-weighted depth at capture of various groundfish species in NMFS bottom trawl surveys pre and post trawl warp offset problems. The amount of warp offset, as measured dockside, is also given. Note that most catches of these species are made in depths where the offset was less than about 9 feet. The box plots give the median value surrounded by the 25th and 75th percentiles of the distributions of depths of occurrence.

Paired Tow Experiments

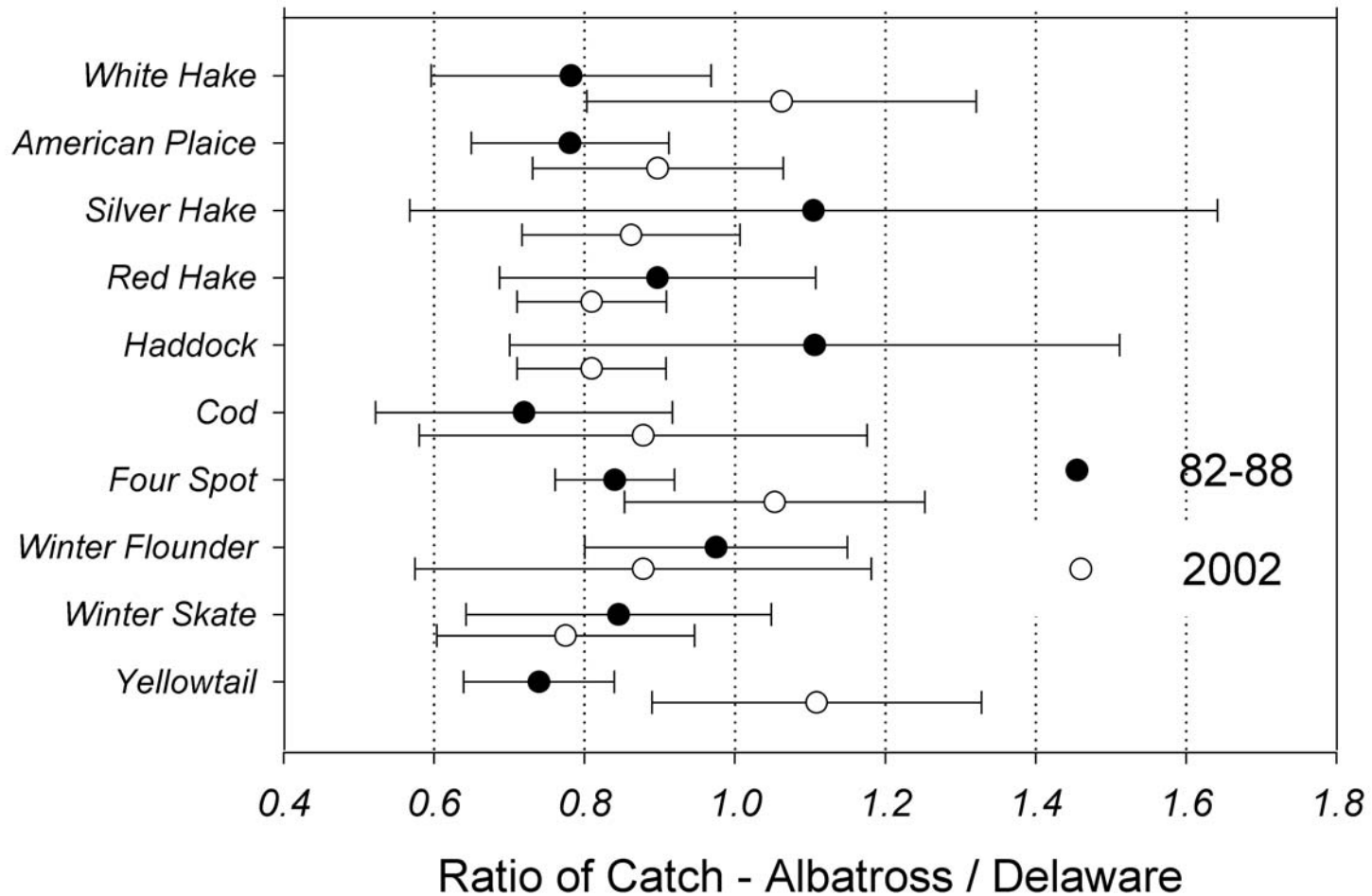


Figure 3. Results of side-by-side tows made by NOAA R/Vs *Albatross IV* and *Delaware II* in paired towing in the 1980s and 2002. Data are mean and 95% confidence intervals of the ratio (*Albatross to Delaware*) of catch rates by species. In only one case (yellowtail flounder) was there a significant change between time periods, and that difference was a positive change in the post warp period.