Fisheries and Oceans Canada Pêches et Océans Canada

NOAA FISHERIES
NATIONAL MARINE FISHERIES SERVICE

CERT

Comité d'évaluation des ressources transfrontalières

Document de travail 2017/??

Ne pas citer sans autorisation des auteurs

TRAC

Transboundary Resources Assessment Committee

Working Paper 2017/??

Not to be cited without permission of the authors

Empirical estimates of maximum catchability and minimum biomass of Georges Bank yellowtail flounder on the NEFSC bottom trawl survey

Author: David Richardson[1], John Hoey[1], John Manderson[1], Michael Martin[2,3], Chris Roebuck[4]

[1] National Marine Fisheries Service
Northeast Fisheries Science Center
28 Tarzwell Drive
Narragansett RI, 02882 USA

[2] National Marine Fisheries Service
Northeast Fisheries Science Center
166 Water Street
Wood Hole MA 02543 USA

[3] Current address: National Marine Fisheries Service
Alaska Fisheries Science Center
7600 Sand Point Way N.E., Building 4
Seattle, Washington 98115 USA

[4] F/V Karen Elizabeth

Narragansett RI 02882

ABSTRACT

For fisheries research surveys the catchability coefficient (q) defines the relationship between a survey index and population size. Typically, catchability is estimated within a stock assessment model. However, empirical estimates of survey gear catchability can provide important information to a stock assessment as: 1) a diagnostic measure of model accuracy, 2) a direct input into the assessment model or 3) a component of setting catch advice when there are major unresolved issues in the standard modelling approaches. For Georges Bank yellowtail flounder, bottom trawl catchability is used in the third manner, as one component in setting catch advice outside of a standard stock assessment model. This use of catchability in formulating management advice has prompted the Northeast Fisheries Science Center (NEFSC) to undertake a number of studies to understand the efficiency of the bottom trawl gear used on the shelfwide surveys. Here we present the results of a 2015-2016 gear comparison study using a twin trawl onboard the *F/V Karen Elizabeth*. We used the data from this study to develop an estimate of the maximum catchability of Georges Bank yellowtail flounder on the NEFSC bottom trawl survey. These results were then applied to the survey index to provide an estimate of minimum total stock biomass for the spring and fall survey.

RÉSUMÉ

Don’t worry about French translation, it will be handled later at the Res Doc stage.

INTRODUCTION

In multi-species fisheries-independent bottom trawl surveys, sampling gear is not optimized for any one species. Rather the gear and sampling protocols are designed to provide consistent and representative samples that allow indices of abundance, size and age to be developed for a suite of species. In order to provide indices of population abundance with minimal potential sources of bias, survey bottom trawls must use sweeps with discs or other modifications that allow the gear to be towed across as wide a variety of habitats as possible, including seafloor habitats with complex physical structures.

Indices of abundance at age and size derived from fisheries independent bottom trawl surveys are scaled to population size using the survey catchability (q) parameter. Catchability is typically estimated internally within stock assessment models that incorporate fisheries landings, survey-based estimates of population trends, age structure information and life history parameters. However, empirical estimates of survey catchability can also sometimes provide important information to a stock assessment. These empirical estimates can be used as a direct input into the assessment model, can serve as a diagnostic measure of model accuracy, or can contribute to an alternate means of providing catch advice when an assessment model is not considered acceptable.

Two independent processes are accounted for by the catchability (q) parameter. These are detectability (δ), the proportion of organisms present at a sampling site detected with the sampling method (also termed gear efficiency), and availability, (ρ), the proportion of a population falling within the footprint of the survey:

$$q=\delta*\rho.$$

Catchability (q) is used in equations to estimate population biomass (B_t)at time t based on tow swept area (a), total survey area (A), and an index of abundance (I_t):

$$B_t = \frac{A}{aq} \mathrm{I}_t$$

All else being equal, this equation specifies an inverse relationship between catchability and biomass. In other words, as the catchability estimate decreases the biomass estimate increases.

Absolute values of detectability are difficult to estimate for most species on most surveys because accurate field-based estimates of abundance are usually not available. However, maximum bounds on detectability can be established by comparing relative catch rates of organisms caught in a scientific survey gear versus a gear optimized for a given target species or group of species (Munroe and Somerton 2002). This approach can be particularly useful for species such as flatfish that have an escape response that involves associating with the bottom and diving underneath the gear (Ryer 2008). For these species the same features of a survey trawl that allow for a broad sampling of bottom habitats, also increase the likelihood of individuals escaping under the gear. Field comparisons of catch rates in survey trawls and trawls specialized for the capturing bottom tending fish, such as flatfish, can be used to estimate maximum bounds on detectability.

For many stocks and surveys the proportion of fish occupying habitat off the survey footprint can be notable and availability should thus be considered in the catchability estimates. However, for Georges Bank yellowtail flounder the stock boundaries do not encompass areas inshore of the Northeast Fisheries Science Center bottom trawl survey, nor is there any evidence that fish will occur in locations deeper than sampled by the bottom trawl survey. This means that availability of Georges Bank yellowtail flounder to the survey can reasonably be estimated as 100%, with catchability only affected by the detectability of the sampling gear.

In this study we:

1) Present results from a field experiment to estimate the relative detectability of yellowtail flounder in the same net fitted with the rockhopper sweep routinely used in the NEFSC bottom trawl survey and a chain sweep optimized to capture flatfish. These results are interpreted in the context of available NEFSC trawl survey data to provide an estimate of detectability on that survey.
2) Use these catchability estimates to provide an estimate of the scale of the population.

METHODS

Twin trawl field experiment

The field experiment was performed on the *FV Karen Elizabeth*, a 24 meter twin rigged fishing trawler, during November of 2015 and August of 2016. Sampling in 2015 was performed on two legs. The first leg sampled 35 tows November 9-13, 2015 in Southern New England. The second leg sampled 73 tows November 15-19, 2015 on Georges Bank. The experiment in August of 2016 used the same experimental approach, but targeted Witch Flounder, a species that tends to occur in deeper waters than yellowtail flounder. The first leg occurred from 7-13 August and sampled along the northern edge of Georges Bank. The second leg occurred from 16-20 August and sampled offshore of Cape Anne.

The twin trawl rig allowed the vessel to simultaneously tow two similar nets, allowing direct comparisons of gear efficiency (Fig. 1-3). A number of studies have used twin trawl rigs to measure the relative efficiencies of trawls and trawl configurations, including studies focused on flatfishes (Ryer et al, 2010; Engas and Godo, 1989; Walsh, 1992; Munro and Somerton, 2002; Somerton et al., 2007; Ingolfsson & Jorgensen, 2006; Krag et al., 2010). Two 400 x 12 4-Seam trawls as used on the Northeast Fisheries Science Center (NEFSC) fishery independent bottom trawl surveys (for details, see Politis, 2014), were fished on both sides of the twin rig. The inside wings and sweeps of each net were attached by 38 m bridles to the 1134 kg central clump that was towed along the bottom (Fig. 3). The outside wings and sweep of each net were attached by a 38 meter bridles to 84” type 4 Thyborøn doors (weight = 590 kg; area = 3.39 m^2). To ensure consistent trawl geometry for each net irrespective of gear differences or depth, 32 meter restrictor ropes (9/16” Polytron) were connected between each of the doors and the clump. The spreading force of the doors, which are larger than those used the spring and fall NEFSC trawl surveys (Poly_Ice oval - weight = 550 kg, area = 2.2 m^2), ensured that the restrictor rope remained consistently taut throughout each tow and ensured a wing spread of ~ 13 m for each

net. From the attachment points and the likely orientation of the gear during towing, we believe that the restrictor ropes were ~ 1 m off the seafloor at the clump and ~ 1.5 m at the doors when in fishing configuration. A trawl monitoring system was employed to monitor wing spread using spread sensors sewn into the outer portion of the wings.

While the nets used on each side of the twin trawl rig were the same, the sweeps and number of floats on the headrope differed. These changes were made to optimize the ability of the experimental gear to catch flatfish and to provide a measure of the proportion of fish passing under the standard survey gear. The net on one side of the twin trawl was a standard NEFSC bottom trawl fitted with the rockhopper sweep (Politis, 2014). The rockhopper sweep has 30, 40.64 cm x 2.54 cm, 6.35 cm thick discs in the center of the net (total length = 890 cm) and 27, 35.56 cm x 2.54cm, 6.35 cm thick disks in each of the wings (total length = 820 cm). The net is connected to the bolchline which is then connected to the sweep with a wire traveler. The net with the rockhopper sweep had 66, 20cm spherical floats positioned along the headrope and upper wing end extensions (Fig. 1). The net on the other side of the twin rig was fitted with a sweep specifically designed to maximize flatfish capture. The sweep was made with 2 rows of ½ inch (1.25 cm) steel chain (6cm) connected together every other link by 1/2 inch (1.25 cm) shackles (Fig. 2). The 3rd row of ⅝ inch (1.6 cm) chain of the sweep had 5 inch outer diameter x 2 inch hole diameter (12.7 cm x 5 cm) ⅜ inch (1 cm) thick disks at every second link. Every fourth link of the ⅝ chain was attached directly to the footrope of the net with a ½ inch shackle (Fig. 2). The headrope of the net with the chain sweep had 32, 20 cm spherical floats, roughly half the number of the standard survey trawl, to limit the buoyancy of the trawl and maximize seafloor contact.

All flatfish, skates, and scallops collected in each net in each tow were independently sorted and weighed. The lengths of all flatfish, thorny skate, cod, and monkfish captured in each net were measured when fewer than approximately 150 individuals were present in the catch for that species. If greater than approximately 150 individuals were present for a species the catch was randomly sub-sampled for length.

Analytical Approach

Our objective was to developing an estimate of the minimum total biomass of Georges Bank yellowtail flounder using empirical estimates of catchability. Catchability is known to vary with fish length and light levels (Somerton et al. 2007, Ryer 2008). We present a range of approaches to account for diel and length-specific catchability, as described below and in Figure 4.

1) **Length-specific or aggregate biomass**. It is possible to calculate catchability values working with only the total biomass sampled. Alternately, length-specific catchabilities can be calculated and applied, and then a length-weight relationship can be used to convert abundance at length back to total biomass. The results of these two approaches can diverge when the catchability values are applied to survey data with a different length distribution than the experimental data used to derive these catchability values.
2) **Use day-night differences or aggregate day-night tows**. A number of different studies have shown that flatfish have different reactions to a bottom trawl in light conditions and dark conditions (Ryer and Barnett 2006, Ryer 2008, Ryer et al. 2010). In the light flatfish

have a tendency to exhibit a herding response, remain close to the seafloor and also to dive under the gear. At night herding breaks down but fish tend to rise more off the bottom when disturbed making them more susceptible to some types of survey gear (Ryer 2008). Daytime and night time catches are thus expected to differ for both survey gear and gear optimized for flatfish. We evaluated three approached to account for day-night behavior in evaluations of catchability:

a. **Do not separate day and night tows in analyses**
b. **Assume 100% experimental gear efficiency during both daytime and nighttime. Use day-night differences in experimental catchability to correct for any differences in the proportion of day:night tows between the experiment and survey.** The NEFSC trawl survey works over a 24-hour period. The fall survey typically occurs after the fall equinox (i.e. longer nights than days), resulting in more nighttime tows, whereas the spring survey on Georges Bank typically occurs around or after the spring equinox (longer days than nights). The experimental work also operated day and night, but no attempt was made to keep the balance of day-night tows the same as for the survey, generating a potential source of bias if day and night tows are not accounted for. As an example, for witch flounder there was a prominent difference in the proportion of daytime tows (≈60%) in the experiment versus the survey (43%). Daytime catch ratios of the chain:rockhopper sweep were higher than at night. For the witch flounder estimate, failing to account for differences in the proportion of daytime tows in the 2016 experiment versus the survey would have resulted in an underestimate of catchability and an overestimate of biomass (Hare et al. 2016).
c. **Assume 100% experimental gear efficiency at night and calculate daytime survey efficiency using nighttime twin trawl study data and survey day:night ratio.** This approach calculates the 24-hour survey catchability on the assumption that the experimental chain sweep is 100% efficient at night. The rockhopper sweep nighttime efficiency is based on the twin-trawl experiment and the daytime rockhopper efficiency is based on the day:night detectability differences estimated from the survey data. This approach can result in a daytime chain sweep efficiency that is less than or greater than 1 (i.e. daytime herding). This approach ensures that Night:Day survey catch ratios are maintained; assuming 100% daytime and nighttime efficiencies of the chain sweep and calculating the rockhopper sweep efficiencies can lead to rockhopper day and night efficiencies that are inconsistent with survey day:night catch ratios.

3) **Apply catchability tow-by-tow or to the survey index**
   a. **Apply day night detectability based on the average proportion of nighttime tows for the survey across years.** This approach uses a single q estimate across all years. This q estimate is readily applied to new data and the trend in the index and biomass estimate do not diverge.
   b. **Apply day night differences on a station-by-station basis.** The q estimate varies by year based on the allocation of nighttime and daytime tows.

Integrating NEFSC trawl survey data

All analyses incorporating the NEFSC trawl survey were done using offshore strata 13-21. Analyses of the twin trawl data used the "suncycle" function in Matlab to designate tows as daytime or nighttime. Trawl survey stations were assigned to day:night based on the solar zenith angle (90.83) calculated in SAGA using Jacobson (2011). Length-specific calculations were done on 1 cm intervals with the exception of individuals ≤25 cm and ≥42cm which were aggregated in these two length bins. For length-specific analyses the length-weight equation ($W=\exp(-11.7897)*L^{3.0249}$) was used to convert abundance-at-length to biomass-at-length.

Below we show the set of equations that were used in the calculations that assumed 100% experimental gear efficiency at night and with daytime survey efficiency calculated using nighttime twin trawl study data and survey day-night catch ratios. Other scenarios modified these equations, but are not show for simplicity

To calculate the maximum detectability of the 24-hour standard trawl survey on the *Henry Bigelow* ($\delta_{Bigelow,24h}$) we quantified the proportion of tows made during the night and proportion of tows made during the day during the fall survey from 2009-2016. The 24 hour detectability of the NEFSC trawl survey on the Bigelow is thus:

$$\partial_{Bigelow,24h} = Proportion\ night\ tows * \partial_{RH,night} + Proportion\ Day\ tows * \partial_{RH,day}$$

The detectability of the Rockhopper sweep at night ($\delta_{RH,night}$) was estimated using the catch data from the twin trawl experiment, with the detectability of the chain sweep at night assumed to be 1 ($\delta_{CH,night}$=1):

$$\partial_{RH,night} = \frac{Catch_{RH}}{Catch_{CH}} * \partial_{CH,night} = \frac{Catch_{RH}}{Catch_{CH}}$$

The detectability of the Rockhopper sweep during the daytime was estimated using the ratio of the mean catch during the nightime and daytime for 2009-2016 on the NEFSC trawl survey aboard the *Henry B. Bigelow*.

$$\partial_{RH,day} = \frac{Mean\ Catch_{Big,day}}{Mean\ Catch_{Big,night}} * \partial_{RH,night}$$

Calculation of minimum swept area biomass

Swept area biomass can be calculated using the following equation:

$$B_t = \frac{A}{aq} \mathrm{I}_t$$

Where: B_t=Swept area biomass
A = Total stock area covered by survey (km^2)
a = area covered by a tow (km^2 tow^{-1})
q= catchability→for GBYT = $\delta_{Bigelow,24h}$
I_t = index.

The survey on the Henry Bigelow on averages samples 0.024 km^2 per tow (i.e. a=0.024 km^2). For the Georges Bank yellowtail flounder, the designated strata (NEFSC strata: offshore 13-21) encompass a total strata area (A) of 37,773 km^2. We present calculations for both the spring and fall survey.

Confidence intervals

All confidence intervals were calculated using bootstrapping. For the twin trawl data paired tows were maintained in the bootstrapping (n=1000) analyses. We report two confidence intervals on the total biomass. The first set of confidence intervals accounts only for the uncertainty of catchability (q). The second set of confidence intervals, in addition to the uncertainty in catchability, also accounts for the uncertainty in the index, using bootstrapped CIs on either the stratified mean biomass of the index or stratified mean number at length (Smith 1997).

Potential Habitat Based Biases

A concern raised in previous work is that the results of gear comparison experiments, which are often logistically constrained to certain seafloor types, may not be broadly applicable to other bottom types. One important piece of information in evaluating the potential importance of this issue for a specific stock is the relative contribution of tows on different bottom types to the total trawl survey swept area biomass of that stock. We used a two-step process to calculate the proportion of the total swept area biomass by bottom type for yellowtail flounder.

The first step involved assigned a bottom habitat to every station of the trawl survey. The data for this step was a bottom habitat map developed by the Nature Conservancy as part of the Northwest Atlantic Marine Ecoregional Assessment (Nature Conservancy 2016). A full list of data sources, data quality control, and processes for developing the underlying map can be found in the associated documentation. The output resolution of the map is ≈500 m. For each trawl station we assigned a bottom habitat type based on the grid cell for the nearest neighboring point on the underlying map. Each trawl station was assigned to one of eight bottom habitat categories based on the sediment grain size (Clay <0.002 mm; Silt 0.002-0.06 mm; Very Fine Sand 0.06-0.125; Fine Sand 0.125-0.25 mm; Medium Sand 0.25-0.50 mm; Coarse Sand 0.50-1.0; Very Coarse Sand 1.0-2.0 mm; Gravel/Granule >2.0mm).

The second step involved calculating the proportion of the total swept area biomass that can be attributed to each of the habitat types each year. Each tow on a random stratified survey contributes to the total swept area biomass based on the catch for that tow, the area of the strata for that tow and the number of stations allocated to that strata. The proportion of swept area biomass on any one habitat type can thus be calculated using the following equation.

$$\boldsymbol{Proportion\ on\ Fine\ Sand} = \frac{\sum_{i=1}^{Fine\ sand\ stations} Catch_i \frac{Strata\ area\ for\ station\ i}{N\ stations\ in\ strata}}{\sum_{i=1}^{all\ stations} Catch_i \frac{Strata\ area\ for\ station\ i}{N\ stations\ in\ strata}}$$

We calculated, for each year of the survey for each stock the proportion of the swept area biomass that can be assigned to each bottom habitat category. A mean proportion for each bottom type category is also provided and used in sensitivity analyses.

Herding

Calculations were all done assuming wing swept area as the appropriate measure of the area swept by the trawl. An alternate measure is door swept area. We present an analysis of the change biomass that results from different assumptions of bridle efficiency from 0 (no fish herded by the bridles) to 1 (all fish herded by the bridles to the net where they are subject to capture.

RESULTS

Twin Trawl Results

The two twin trawl experiments sampled yellowtail flounder at a total of 139 stations, with 121 and 119 of those catching yellowtail flounder with the chain sweep and rockhopper sweeps respectively. For total weight the chainsweep to rockhopper sweep catch ratios was 4.92 (95% CI: 4.24-5.71) for daytime tows. For nighttime tows the ratio was 3.05 (95% CI: 2.79-3.33). For all tows the ratio was 3.62 (3.30 -4.02). With increasing length, a slight pattern of decreasing chain sweep to rockhopper sweep catch ratios was evident.

Total biomass

We used ten different approaches to estimate total biomass using NEFSC trawl survey data and the results of the twin trawl study (Figure 5). These ten approaches were developed based on conversations among the internal NEFSC “yellowtail flounder catchability” working group tasked with overseeing this analysis. The intent of these various iterations was to evaluate a suite of factors that may be important to estimating Georges Bank yellowtail flounder biomass. These results do not necessarily extend to other species and stocks. Table 3-6 provides the biomass from each of these approaches for 2009-2016 for the Fall (Tables 3-4) and Spring (Tables 5-6) and with confidence intervals based only on uncertainty in catchability (Tables 3 & 5) and based on uncertainty in catchability and the survey index (Table 4 and 6). Figure 6 illustrates how the resulting biomass of each of these approaches differs from the simplified “weight-only” approach (Approach 1 in tables 3-6).

The simplest approach was to develop a catchability value for the rockhopper sweep for weight per tow based on the results of the twin-trawl study. This approach resulted in 2014, 2015, 2016 total biomasses of 9.6, 7.6, and 3.2 kt respectively for the spring survey and 13.5 kt, 6.0 kt, and 7.0 kt respectively for the fall survey.

A second approach used length specific catchabilities and applied these to abundance at length data. A weight length relationship was then used to calculate biomass at length, which was summed to give total biomass. The difference between this approach and the weight only approach was limited for the fall and the spring (3-15%).

The first way to account for day-night differences in the twin trawl study was to calculate catchabilities separately for daytime and nighttime tows. To calculate 24-hour catchability and a weighting factor based on the survey proportion of nighttime daytime tows was applied to the daytime and nighttime catchabilities. This approach can account for a mismatch between the proportion of catch and daytime tows in the twin trawl experiment for a species, and the proportion of daytime tows on the survey. The underlying assumption of this approach is that daytime and nighttime tows with the chain sweep are both 100% efficient. For the fall the proportional Day-Night approach differed little from aggregating nighttime and daytime tows (<1%). The Spring has a higher proportion of daytime tows and thus the resulting biomass was about 7% higher.

A second way to account for day-night differences is to apply daytime and nighttime catchability values calculated from the twin trawl study on a tow by tow basis, and then to calculate the stratified mean and swept area biomass. This approach generates total catchability estimates that differ among years.

The third way to account for Day-Night differences is to only assume that nighttime tows with the chain sweep are 100% efficient but not to constrain daytime efficiency. The Rockhopper sweep nighttime catchabilities are thus calculated from twin trawl study, while the survey data is used to estimate the daytime catchability of the rockhopper sweep relative to the nighttime catchability. This approach had a notable effect on the end results for the fall, raising the total biomass 29% when applied in isolation of the other techniques. For the spring, this approach resulted in a slightly lower (2%) biomass than assuming 100% efficiency during both the day and night.

Combinations of the different approaches (i.e. day:night & length-based) are shown in the tables, but are not described further.

Bottom habitat sensitivity analysis

For Georges Bank Yellowtail flounder on average 86% of the swept area biomass was sampled on bottom types that were coarse sand or finer. The highest percentage (63%) of the swept area biomass came from sampling on medium sand bottom habitat. No obvious time trend was apparent in the bottom habitats contributing to the swept area biomass for Georges Bank Yellowtail Flounder. Assuming the rockhopper gear was 100% efficient on the very coarse sand and gravel habitat types, rather than 28% efficient as calculated by the twin trawl experiment the total biomass values calculated would be a 12% overestimate (Figure 7).

Herding

Herding by the bridles is discussed further in the next session. Figure 2 shows a sensitivity analysis of the biomass estimates assuming different levels of herding. Herding is defined as the proportion of fish encountered by the bridles that are driven to the path of the net where they are subject to capture by the net, outswim the net or can pass under or over the net. The door swept area is 2.54-fold the wing swept area; the area subject to herding is thus 1.54-fold the wing swept area.

DISCUSSION

With some notable exceptions (e.g. acoustic surveys, camera surveys of scallops, egg surveys), the primary goal of research surveys is to provide an index of abundance rather than an absolute estimate of abundance. One function of a stock assessment model is to integrate a diverse set of data to determine the scale of the population. However, a variety of factors can make for imprecise and inaccurate scaling of population levels including inaccurate catch data, low fishing mortality rates over the time series, uncertain and time-varying natural mortality, and time-varying catchability on the research survey leading to biased estimates of trend. In these cases, external information can serve a role in population scaling, even if only to establish minimum estimates on stock biomass. The purpose of this analysis was to establish a maximum bound on catchability of the NEFSC fall trawl survey for Georges Bank yellowtail flounder and a minimum bound on the estimate of biomass.

The primary data used in this analysis came from an experiment using a twin trawler and many of the standard tow protocols for the NEFSC survey on the R/V Bigelow. The experimental net used on one side of twin trawler was the same as the standard survey trawl used on the R/V Bigelow except that it contained roughly half number of floats and the sweep was modified to optimize flatfish catch by minimizing the ability of flatfish to pass under the net. On the other side the standard trawl and ground gear used on the R/V Bigelow was fished. The towing of the standard survey bottom trawl on the twin trawl experiment differed in a few ways from its deployment on the spring and fall bottom trawl surveys, but we believe that these differences did not have a significant effect on the results. The use of larger doors and the restrictor rope served to fix the net geometries which is likely the biggest potential source of variability in comparative trawl catches. This setup also allowed us to avoid many of the potential problems due to the rather large size difference of the Bigelow and the Karen Elizabeth. The use of the restrictor rope is not believed to have influenced flatfish behavior in front of the trawl as flatfish have been shown to generally not react to trawling induced stimuli until they are in very close proximity or even contacted by the fishing gear (Ryer et al, 2010). The spread data indicated that the restrictor rope remained taut throughout the towing process (setting, towing, hauling back), so we believe it likely that the restrictor rope was almost always at least 1 m off the bottom at all times. All information indicates that bridle herding was very minimal for flatfish, and therefore the possibility of an increase in herding due to an increase in mud clouds from the use of larger doors was likely a non-factor in the catches that we observed. Similarly, the use of the clump instead of a door was likely not a factor for similar reasons. The use of a modified scope table which ensured that the net ground gears on the twin trawler were in contact with the seabed in deep water probably resulted in a very small difference as the amount of wire out never differed (+/-) by more than 13 fm from what the Bigelow scope table prescribes.

The analyses presented here assume that bridle herding did not occur at night with the chain sweep, and thus that wing swept area, rather than the door swept area, is the appropriate measure of the area swept by a tow. Herding is a known phenomena for flatfish and many other species when certain types of gear are used (Ramm and Xiao 1995, Somerton and Munro 2001, Somerton et al. 2007, Rose et al. 2009). In calculating bridle efficiency, Somerton and Munro considered two factors. The first was W_{off}, or "the width of the path in which the bridle is not in contact with the bottom." The second factor is the herding coefficient, h, which is the "proportion of fish in the bridle contact path that is herded into the net path." These two factors

encompass the design of the gear (W_{off}) and the response of fish to the gear (h), providing a useful framework to evaluate herding.

The bottom trawl used on the NEFSC Bigelow survey was designed to minimize herding. On the second set of experimental work, the bridles used during the tows were painted to evaluate whether they were making contact with the bottom; abrasion from bottom contact over the many nautical miles that this gear was towed during this leg would result in the paint wearing off. The paint was not observed to wear off during the tows consistent with a high value of W_{off} or low bridle contact with the bottom.

The second factor underlying our decision to use wing swept area was based on a review of published studies of flatfish herding and behavior in relation to trawl gear. We characterized these studies based on whether only daytime tows were evaluated or both day and nighttime tows were evaluated. The studies consistently found that herding behavior occurred during the daytime (Glass and Wardle 1989, Somerton and Munro 2001, Ryer and Barnett 2006, Somerton et al. 2007, Bryan et al 2014, Ryer et al. 2010) with some studies indicating high herding coefficients (h) along the sections of the bridles in contact with the bottom. The form of the herding behavior was directional swimming near to the seafloor away from the bridles and towards the center of the gear. In contrast, the subset of studies that evaluated herding at night or in low light conditions did not find evidence for a directional herding response (Glass and Wardle 1989, Ryer and Barnett 2006, Ryer 2008, Ryer et al. 2010). Rather, under low light conditions fish would rise off the bottom when disturbed making individuals encountered between the wings more vulnerable to capture than during the daytime, but limiting the capture of fish encountered by the bridles. Both laboratory and field studies were consistent in these findings. These behavioral studies can explain why many flatfish species experience higher nighttime catch rates in survey gear (designed not to herd) while having higher daytime catch rates in commercial gear (optimized to herd and minimize fish passage under the net). Our calculations in all approaches assume that at night the chain sweep was 100% efficient between the wings but that herding did not occur.

A concern raised in previous work is that the results of gear comparison experiments, which were logistically constrained to certain seafloor types (sand, mud, cobble, shell hash), may not be broadly applicable to all bottom types. We assumed 100% efficiency of the chain sweep in the analyses with the goal of obtaining a maximum detectability for the survey. A chain sweep to rockhopper sweep catch ratio higher than what we observed in another bottom type would suggest even lower rockhopper catchabilities in that area. In this case the maximum efficiency would be an overestimate and biomass an underestimate. If the catch ratios were lower in another unsampled area but the catchabilities of the rockhopper survey gear were assumed to be constant across space, the resulting biomass would not change. The scenario in which the results of the twin trawl work could lead to an underestimate of detectability (and an overestimate of biomass) is 1) if the detectabilities of the survey rockhopper gear vary with bottom type on the NEFSC trawl survey, and 2) if the gear comparison experiment occurred in a low detectability area for the rockhopper sweep, that is not an equally (relative to other areas) low detectability area for the chain sweep. Analyses of spatially varying survey catchabilities of the rockhopper sweep on the NEFSC trawl survey are not available, nor to our knowledge has bottom type been considered in other paired gear experiments with this gear. As a sensitivity analysis, we quantified the percentage of the swept area biomass for Georges Bank yellowtail

flounder by habitat type. For Georges Bank yellowtail flounder 86% of the swept area biomass was estimated to have been caught on coarse sand bottom habitats or finer sediments. Assuming the rockhopper gear was 100% efficient on the larger grain size habitat types, rather than the 28% efficiency (for the simplest model) calculated by the twin trawl experiment the total biomass values calculated in this manuscript for the Bigelow would be a 12% overestimate.

Most of the approaches to estimating stock biomass provided comparable results for Georges Bank yellowtail flounder. Length-specific analyses are appropriate if there is a mismatch between the length structure in the experimental work and the survey. For Georges Bank yellowtail flounder, across the years analyzed, using length structure did not result in prominent differences in the biomass estimate. Day and night should be dealt with separately if 1) the experiment indicates day night catchability differences and 2) there is a mismatch between the survey and experiment in the proportion of catch in nighttime tows. This weighting of day and night tows in calculating a catchability value made a minimal difference for Georges Bank yellowtail flounder in the fall and a slight difference in the Spring which has a higher proportion of daytime tows. Tow-by-tow analyses resulted in the expected year to year variability in the effect on the index as day tows were scaled up by a higher margin than night tows.

The most prominent difference among techniques resulted from only constraining nighttime efficiency of the chainsweep at 100%, versus constraining both nighttime and daytime efficiency at 100%. Conceptually, this is the difference between the expected swept area biomass of a survey using only nighttime tows of the chain sweep versus 24 hour operations with the chain sweep. In this approach daytime rockhopper sweep efficiency is calculated using a two step process of using the nighttime twin trawl study results and the survey day night ratios, instead of directly through the daytime twin trawl study results. However, applying the daytime twin trawl results to the calculated daytime rockhopper efficiency values allows for an estimate of daytime chain sweep efficiency. For the fall survey, the estimated daytime efficiency of the chain sweep is <100%, explaining the higher biomass estimates using this technique. In contrast for the spring survey the estimated daytime efficiency is ≈100% explaining the minimal difference in applying this approach. Notably, this approach does allow daytime chainsweep efficiencies to exceed 100%, as would occur if there is some level of daytime herding not offset by lower net efficiency. This approach is thus consistent with only assuming herding does not occur during the nighttime.

LITERATURE CITED

Bryan, D.R., Bosley, K.L., Hicks, A.C., Haltuch, M.A., Wakefield, W.A. 2014. Quantitative video analysis of flatfish behavior and impact on effective swept area of a survey trawl. Fisheries Research. 154: 120-126.

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

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

Hare JA, Hoey J, Manderson JP, Martin M, Politis P, Richardson DE, Roebuck C. 2016. Empirical estimates of maximum catchability of Witch Flounder *Glyptocaphalus*

cynoglossus L. on the Northeast Fisheries Science Center Fall bottom trawl survey. SAW 62 Data meeting.

Ingolfsson O. and T. Joergensen. 2006. Escapement of gadoid fish beneath a commercial bottom trawl: Relevance to the overall trawl selectivity. Fish. Res. 79:303-312.

Jacobson, L., A. Seaver, and J. Tang. 2011. AstroCalc4R: Software to Calculate Solar Zenith Angle; Time at sunrise, Local Noon, and Sunset; and Photosynthetically Available Radiation Based on Date, Time, and Location. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 11-14; 10 p.

Krag, L.A., Holst,R., Madsen, N., Hansen, K., and Frandsen, R.P. 2010. Selective haddock (Melanogrammus aeglefinus) trawling—avoiding cod (Gadus morhua) bycatch. Fish. Res. 101:20–26

Politis, P.J., Galbraith, J.K., Kostovick,P. Brown, R.W.. 2014. Northeast Fisheries Science Center bottom trawl survey protocols for the NOAA Ship Henry B. Bigelow. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 14-06; 138 p.

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

Nature Conservancy. 2016. Soft sediments by grain size (in mm) Northwest Atlantic United States. Prepared for Northeast Regional Ocean Council, Northeast Ocean Data. www.northeastoceandata.org

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

Rose, C.S., Gauvin, J.R. and Hammond, C.W. 2009. Effective herding of flatfish by cables with minimum seafloor contact. Fishery Bulletin. 108: 136-144.

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

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

Ryer, C. H., Rose, C. S., and Iseri, P. J. 2010. Flatfish herding behavior in response to trawl sweeps: a comparison of diel responses to conventional sweeps and elevated sweeps. Fish. Bull. (Seattle) 108:145-154.

Smith S. J. 1997. Bootstrap confidence limits for groundfish trawl survey estimates of mean abundance. Can. J. Fish. Aquat. Sci. 54:616-630.

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

Somerton, D. A., Munro, P. T., and Weinberg, K. L. 2007. Whole-gear efficiency of a benthic survey trawl for flatfish. Fish. Bull. 105:278-291.

Walsh, S. J., and Hickey, W. M. 1993. Behavioural reactions of demersal fish to bottom trawls at various light conditions. ICES Marine Science Symposium, 196: 68–76.

Table 1. Summary of yellowtail flounder catches during the day, night and all tows during the twin trawl experiment. Median and 95% confidence intervals of the ratio of the chain to rockhopper sweep are presented.

		Day	Night	All Tows
Weight	Chain	1516.7	2176.83	3693.5
	Rockhopper	307.8	712.44	1020.2
	Ratio (95% CI)	4.93 (4.18 -5.74)	3.05 (2.80 -3.32)	3.62 (3.30 -4.02)
Number	Chain	4202	6062	10264
	Rockhopper	812	1964	2776
	Ratio (95% CI)	5.17 (4.43 -5.94)	3.08 (2.82 -3.35)	3.70 (3.36 -4.10)
Positive Stations	Chain	59	62	121
	Rockhopper	56	63	119

Table 2. Chain sweep and Rockhopper sweep catch numbers and catch ratios during the daytime, nighttime and for all tows. Median and 95% confidence intervals are presented.

	Day			Night			All Tows		
Length (cm)	CH	RH	Median Ratio (95 % CI)	CH	RH	Median Ratio (95% CI)	CH	RH	Median Ratio(95% CI)
≤25	115	15	7.67 (3.95-12.55)	99	52	1.89 (1.16-2.79)	214	67	3.20 (2.00-5.59)
26	39	11	3.56 (1.89-7.20)	37	26	1.38 (0.76-2.37)	76	37	2.05 (1.38-3.29)
27	36	7	5.20 (2.58-17.42)	32	21	1.53 (0.99-3.17)	68	28	2.50 (1.54-4.62)
28	56	10	5.58 (2.91-13.37)	89	26	3.42 (2.14-6.06)	145	36	4.06 (2.76-6.39)
29	132	17	7.89 (4.50-14.62)	207	40	5.21 (3.48-8.08)	339	57	6.05 (4.29-8.62)
30	270	35	7.56 (5.18-10.98)	404	95	4.23 (3.36-5.29)	674	130	5.20 (4.24-6.33)
31	353	51	6.75 (4.72-10.54)	497	154	3.23 (2.60-4.01)	850	205	4.14 (3.42-5.08)
32	427	88	4.83 (3.61-6.63)	698	187	3.72 (3.07-4.66)	1125	275	4.08 (3.44-4.95)
33	494	94	5.24 (3.97-7.39)	720	246	2.90 (2.50-3.44)	1214	340	3.58 (3.08-4.24)
34	398	64	6.27 (4.89-8.33)	612	208	2.95 (2.47-3.44)	1010	272	3.72 (3.20-4.37)
35	447	98	4.54 (3.45-5.84)	641	241	2.65 (2.28-3.12)	1088	339	3.21 (2.73-3.82)
36	290	55	5.29 (3.61-8.13)	403	156	2.60 (2.17-3.12)	693	211	3.29 (2.74-4.06)
37	315	71	4.43 (2.98-6.40)	451	138	3.25 (2.74-3.94)	766	209	3.68 (3.05-4.44)
38	201	48	4.19 (2.98-6.56)	336	94	3.59 (2.87-4.49)	537	142	3.80 (3.13-4.74)
39	171	38	4.48 (2.96-7.58)	294	82	3.58 (2.81-4.80)	465	120	3.86 (3.09-5.10)
40	168	28	6.03 (3.91-8.99)	199	62	3.23 (2.33-4.48)	367	90	4.09 (3.09-5.40)
41	100	37	2.70 (1.58-4.73)	118	53	2.25 (1.53-3.32)	218	90	2.40 (1.75-3.37)
≥42	190	45	4.31 (2.84-6.85)	225	83	2.73 (2.09-3.53)	415	128	3.27 (2.57-4.25)

Table 3. Total biomass estimates (kt)for the Fall Survey calculated using different approaches for addressing catchability. The Median and 95% Confidence intervals of the 1000 bootstraps are tabled. The median value and confidence intervals only account for uncertainty in catchability. The different approaches are presented in Figure 5.

	Day-Night	Tow-by-Tow	Length Specific	2009	2010	2011	2012	2013	2014	2015	2016
1	No	-	No	96.1 (87.4-106.9)	27.3 (24.8-30.3)	34.4 (31.3-38.3)	38.8 (35.3-43.1)	9.7 (8.8-10.8)	13.5 (12.3-15.1)	6.0 (5.5-6.6)	7.0 (6.4- 7.8)
2	No	-	Yes	98.6 (88.8-111.2)	31.9 (28.9-35.6)	36.6 (33.2-41.0)	39.1 (35.1-44.1)	9.7 (8.8-10.9)	14.0 (12.6-15.7)	5.8 (5.3-6.5)	7.4 (6.7- 8.3)
3	Day q=1 & Night q=1	No	No	96.7 (89.9-104.4)	27.4 (25.5-29.6)	34.6 (32.2-37.4)	39.0 (36.3-42.1)	9.7 (9.1-10.5)	13.6 (12.7-14.7)	6.0 (5.6-6.5)	7.1 (6.6- 7.6)
4	Day q=1 & Night q=1	No	Yes	98.7 (90.7-108.0)	32.0 (29.7-34.7)	36.9 (34.2-40.0)	39.0 (35.7-42.7)	9.8 (9.1-10.7)	14.0 (13.0-15.3)	5.8 (5.4-6.3)	7.4 (6.9- 8.0)
5	Day q=1 & Night q=1	Yes	No	85.8 (79.8-92.5)	25.2 (23.5-27.1)	30.8 (28.6-33.1)	33.7 (31.2-36.4)	9.1 (8.5- 9.9)	13.9 (12.8-15.2)	5.4 (5.0-5.8)	6.4 (6.0- 6.9)
6	Day q=1 & Night q=1	Yes	Yes	90.0 (82.5-98.5)	29.8 (27.7-32.4)	33.2 (30.8-36.1)	34.7 (31.5-38.2)	9.3 (8.6-10.1)	14.3 (13.1-16.0)	5.4 (4.9-5.7)	6.8 (6.3- 7.3)
7	Night q=1	No	No	124.2 (109.8-138.5)	35.2 (31.1-39.3)	44.5 (39.3-49.6)	50.1 (44.3-55.9)	12.5 (11.1-13.9)	17.5 (15.5-19.5)	7.8 (6.8-8.6)	9.1 (8.0-10.1)
8	Night q=1	No	Yes	130.9 (114.5-148.4)	42.2 (37.5-47.4)	48.5 (43.2-54.2)	51.2 (44.8-58.2)	12.8 (11.3-14.3)	18.2 (16.1-20.4)	7.6 (6.8-8.5)	9.7 (8.6-10.8)
9	Night q=1	Yes	No	117.0 (94.6-156.9)	39.1 (29.5-57.0)	42.3 (34.0-57.0)	40.3 (34.8-49.2)	15.4 (11.1-23.6)	30.0 (19.0-50.8)	7.8 (6.1-10.7)	9.8 (7.5-14.1)
10	Night q=1	Yes	Yes	123.5 (98.2-165.0)	48.2 (36.7-68.2)	46.0 (36.7-60.9)	41.9 (35.5-51.7)	17.1 (12.1-24.3)	30.3 (19.9-49.1)	8.2 (6.3-11.2)	10.6 (8.1-15.3)

Table 4. Total biomass estimates (kt) for the Fall Survey calculated using different approaches for addressing catchability. The Median and 95% Confidence intervals of the 1000 bootstraps are tabled. The median value and confidence intervals account for uncertainty in catchability and the index. The different approaches are presented in Figure 5.

	Day-Night	Tow-by-Tow	Length Specific	2009	2010	2011	2012	2013	2014	2015	2016
1	No	-	No	92.2 (41.5-162.7)	29.8 (12.6-51.9)	33.2 (13.4-57.8)	35.6 (5.5-79.7)	9.7 (3.5-22.6)	14.1 (6.3-26.8)	6.7 (1.4-15.93)	5.7 (1.6-12.7)
2	No	-	Yes	94.3 (40.7-159.6)	36.0 (15.5-63.7)	35.3 (15.9-63.3)	36.2 (6.0-78.9)	9.8 (3.4-22.0)	14.8 (6.5-26.3)	6.6 (1.4-16.61)	6.1 (1.9-13.0)
3	Day q=1 & Night q=1	No	No	92.5 (41.2-160.4)	30.2 (12.5-51.7)	33.5 (13.1-58.9)	35.7 (5.5-78.4)	9.8 (3.6-22.1)	14.1 (6.3-27.5)	6.7 (1.4-16.01)	5.6 (1.6-12.9)
4	Day q=1 & Night q=1	No	Yes	94.6 (42.0-159.2)	36.2 (15.6-63.1)	35.9 (15.7-63.5)	35.8 (5.9-80.6)	9.8 (3.4-22.3)	15.0 (6.6-26.3)	6.6 (1.4-16.56)	6.0 (1.8-13.5)
5	Day q=1 & Night q=1	Yes	No	83.0 (39.2-142.3)	27.1 (13.0-47.8)	30.6 (13.5-51.8)	30.9 (7.8-70.3)	9.4 (3.6-19.2)	15.2 (6.5-25.8)	6.1 (1.4-15.36)	5.4 (1.8-10.9)
6	Day q=1 & Night q=1	Yes	Yes	87.5 (39.1-153.1)	32.7 (15.6-54.9)	31.8 (13.7-53.5)	32.7 (7.6-70.8)	9.1 (3.6-19.5)	14.9 (6.5-27.0)	5.8 (1.5-14.39)	5.7 (2.0-12.2)
7	Night q=1	No	No	119.2 (53.1-206.4)	38.6 (16.2-66.8)	43.2 (16.9-75.9)	46.3 (7.2-104.2)	12.6 (4.5-28.4)	18.3 (8.2-35.8)	8.5 (1.8-20.50)	7.3 (2.0-16.4)
8	Night q=1	No	Yes	124.9 (54.0-210.5)	47.7 (20.8-82.7)	46.9 (20.1-83.8)	47.1 (7.8-107.2)	12.8 (4.5-28.7)	19.6 (8.6-35.3)	8.5 (1.9-21.66)	7.8 (2.4-17.6)
9	Night q=1	Yes	No	112.6 (46.2-246.6)	44.4 (23.4-79.4)	42.0 (18.8-85.2)	36.8 (12.0-74.6)	16.1 (6.0-32.5)	32.1 (9.3-71.7)	7.9 (2.4-18.22)	8.9 (3.2-19.6)
10	Night q=1	Yes	Yes	116.5 (47.4-266.3)	54.8 (29.9-93.7)	46.7 (20.3-98.4)	38.8 (12.6-77.3)	17.3 (6.3-32.1)	31.9 (10.9-70.9)	8.0 (2.3-18.99)	9.6 (3.6-19.2)

Table 5. Total biomass estimates (kt) for the Spring survey calculated using different approaches for addressing catchability. The Median and 95% Confidence intervals of the 1000 bootstraps are tabled. The median value and confidence intervals only account for uncertainty in catchability. The different approaches are presented in Figure 5.

	Day-Night	Tow-by-Tow	Length Specific	2009	2010	2011	2012	2013	2014	2015	2016	2017
1	No	-	No	65.6 (59.6-72.7)	72.2 (65.6-80.0)	28.8 (26.2-32.0)	44.4 (40.3-49.2)	10.6 (9.6-11.7)	8.6 (7.9- 9.6)	6.1 (5.6-6.78)	3.5 (3.2- 3.9)	2.6 (2.4- 2.9)
2	No	-	Yes	64.3 (58.1-71.6)	74.9 (67.8-83.3)	32.4 (29.3-36.2)	46.0 (41.7-51.0)	10.8 (9.8-12.0)	8.4 (7.6- 9.3)	6.2 (5.6-6.81)	3.2 (2.9- 3.6)	2.5 (2.3- 2.8)
3	Day q=1 & Night q=1	No	No	70.1 (64.2-76.0)	77.2 (70.6-83.7)	30.8 (28.2-33.4)	47.5 (43.4-51.5)	11.3 (10.4-12.3)	9.2 (8.5-10.0)	6.5 (6.0-7.09)	3.7 (3.4- 4.1)	2.8 (2.6- 3.1)
4	Day q=1 & Night q=1	No	Yes	68.1 (61.8-75.0)	79.5 (72.1-87.2)	34.4 (31.3-37.8)	48.9 (44.7-53.4)	11.5 (10.6-12.5)	8.9 (8.1- 9.7)	6.6 (6.0-7.18)	3.4 (3.0- 3.8)	2.7 (2.4- 2.9)
5	Day q=1 & Night q=1	Yes	No	67.0 (61.4-72.8)	74.7 (68.1-81.2)	30.7 (27.9-33.5)	54.1 (47.4-61.5)	11.3 (10.2-12.3)	10.2 (9.0-11.4)	6.9 (6.2-7.63)	3.7 (3.4- 4.0)	3.2 (2.8- 3.7)
6	Day q=1 & Night q=1	Yes	Yes	66.2 (60.3-73.7)	78.1 (71.0-87.1)	34.7 (31.4-38.7)	56.8 (49.9-65.6)	11.6 (10.7-12.7)	10.1 (8.9-11.5)	7.2 (6.5-8.02)	3.4 (3.0- 3.8)	3.1 (2.7- 3.6)
7	Night q=1	No	No	68.6 (47.5-88.3)	75.5 (52.2-97.1)	30.2 (20.9-38.8)	46.4 (32.1-59.7)	11.1 (7.7-14.3)	9.0 (6.3-11.6)	6.4 (4.4-8.23)	3.7 (2.5- 4.7)	2.8 (1.9- 3.6)
8	Night q=1	No	Yes	68.2 (47.3-89.2)	80.1 (55.8-104.2)	34.6 (23.9-45.2)	48.7 (33.9-63.5)	11.5 (8.2-14.8)	8.8 (6.3-11.4)	6.6 (4.7-8.41)	3.3 (2.3- 4.4)	2.6 (1.9- 3.4)
9	Night q=1	Yes	No	65.6 (50.2-90.7)	72.9 (54.9-102.7)	29.8 (21.7-43.6)	51.8 (31.2-88.1)	11.0 (8.0-16.0)	9.8 (6.2-16.0)	6.6 (4.5-10.32)	3.6 (2.6- 5.2)	3.1 (1.9- 5.3)
10	Night q=1	Yes	Yes	65.0 (49.8-91.0)	78.8 (59.2-112.7)	34.2 (24.9-50.8)	56.3 (33.5-96.1)	11.6 (8.6-17.0)	9.6 (6.3-15.4)	7.5 (5.1-11.67)	3.2 (2.4- 4.6)	3.1 (1.9- 5.3)

Table 6. Total biomass estimates (kt) for the Spring Survey calculated using different approaches for addressing catchability. The Median and 95% Confidence intervals of the 1000 bootstraps are tabled. The median value and confidence intervals account for uncertainty in catchability and the index. The different approaches are presented in Figure 5.

	Day-Night	Tow-by-Tow	Length Specific	2009	2010	2011	2012	2013	2014	2015	2016	2017
1	No	-	No	53.9 (28.8-85.0)	72.7 (38.5-128.6)	32.1 (16.4-52.0)	49.9 (14.8-118.2)	13.4 (7.4-22.2)	9.1 (5.2-13.5)	6.4 (3.7-10.2)	3.8 (1.9- 6.5)	3.0 (1.7- 4.8)
2	No	-	Yes	53.1 (28.0-80.6)	76.3 (40.8-136.1)	35.9 (18.8-57.5)	51.9 (15.1-122.0)	13.6 (7.5-22.6)	9.0 (5.3-13.2)	6.5 (3.9-10.1)	3.6 (1.8- 6.2)	2.9 (1.6- 4.5)
3	Day q=1 & Night q=1	No	No	57.8 (31.0-90.5)	78.0 (40.6-136.3)	34.2 (17.5-56.5)	52.9 (15.7-129.7)	14.4 (8.0-23.1)	9.8 (5.6-14.4)	6.8 (4.1-10.8)	4.0 (2.0- 6.8)	3.3 (1.8- 5.1)
4	Day q=1 & Night q=1	No	Yes	56.6 (28.9-87.4)	80.3 (45.0-143.8)	38.3 (20.0-59.7)	54.1 (16.0-130.0)	14.6 (8.0-23.9)	9.5 (5.6-14.4)	6.9 (4.2-10.7)	3.8 (2.0- 6.5)	3.1 (1.7- 4.7)
5	Day q=1 & Night q=1	Yes	No	55.8 (32.2-84.0)	78.3 (41.8-124.5)	33.6 (16.4-60.0)	63.0 (15.5-159.0)	14.4 (7.6-25.6)	10.5 (5.8-16.2)	7.0 (4.1-12.1)	4.1 (2.0- 6.8)	3.7 (1.9- 5.9)
6	Day q=1 & Night q=1	Yes	Yes	55.1 (30.9-82.7)	80.9 (44.2-131.7)	38.0 (19.6-66.7)	66.7 (16.2-162.3)	14.1 (7.5-26.9)	10.6 (5.8-16.5)	7.3 (4.2-12.1)	3.7 (2.0- 6.3)	3.5 (1.9- 5.7)
7	Night q=1	No	No	55.2 (26.9-91.7)	75.4 (36.5-135.0)	33.3 (15.4-57.0)	48.6 (14.6-127.7)	13.7 (7.0-24.1)	9.3 (4.9-15.2)	6.5 (3.6-11.5)	3.9 (1.8- 7.2)	3.1 (1.6- 5.3)
8	Night q=1	No	Yes	55.1 (27.2-93.8)	79.7 (40.3-146.9)	38.1 (19.3-64.7)	50.6 (15.7-136.9)	14.4 (7.5-25.3)	9.3 (4.8-15.1)	6.9 (3.9-11.4)	3.7 (1.7- 6.6)	3.0 (1.5- 5.0)
9	Night q=1	Yes	No	54.4 (27.8-90.9)	77.1 (37.4-134.3)	32.8 (16.3-68.0)	51.4 (14.1-172.9)	13.0 (7.2-29.2)	10.0 (5.1-19.8)	6.7 (3.6-13.7)	4.0 (1.8- 7.8)	3.4 (1.6- 7.0)
10	Night q=1	Yes	Yes	54.4 (28.7-94.1)	81.9 (43.3-149.5)	38.4 (19.3-75.7)	57.6 (16.6-192.9)	14.4 (7.7-30.7)	10.0 (5.1-19.5)	7.5 (4.1-14.6)	3.5 (1.7- 6.6)	3.4 (1.6- 6.7)

Figure 1. Schematic Drawing of the center section of the rockhopper sweep used on the standard NEFSC bottom trawl

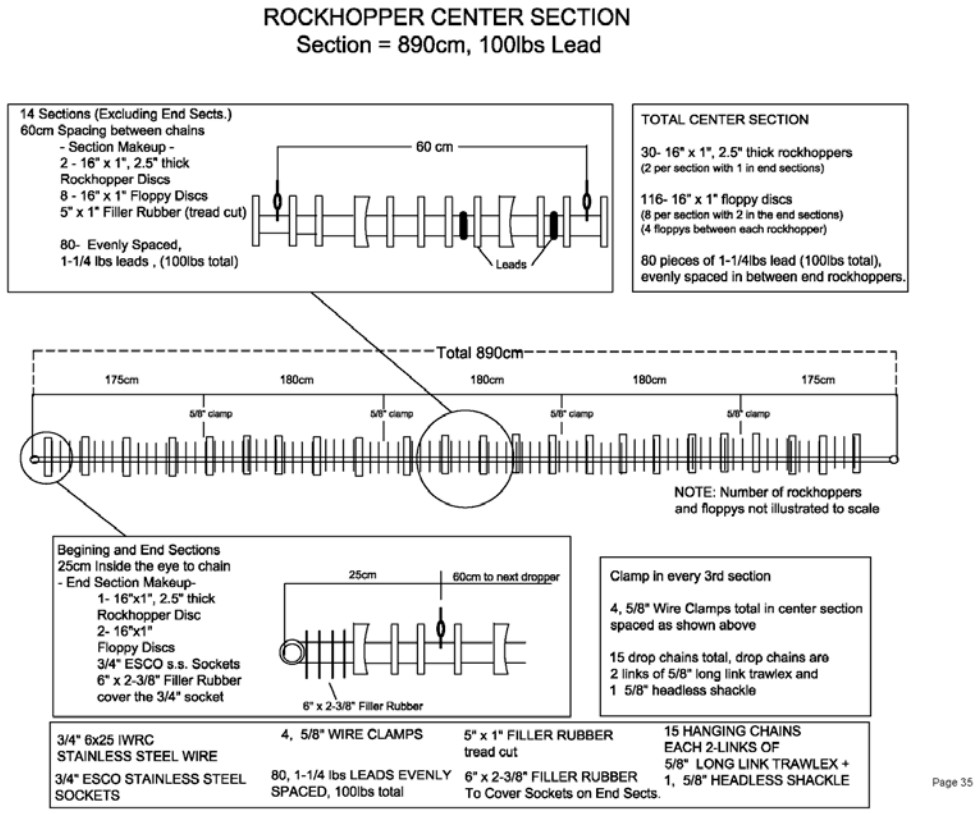

Figure 2. Schematic drawing of the chain sweep used on the experimental net

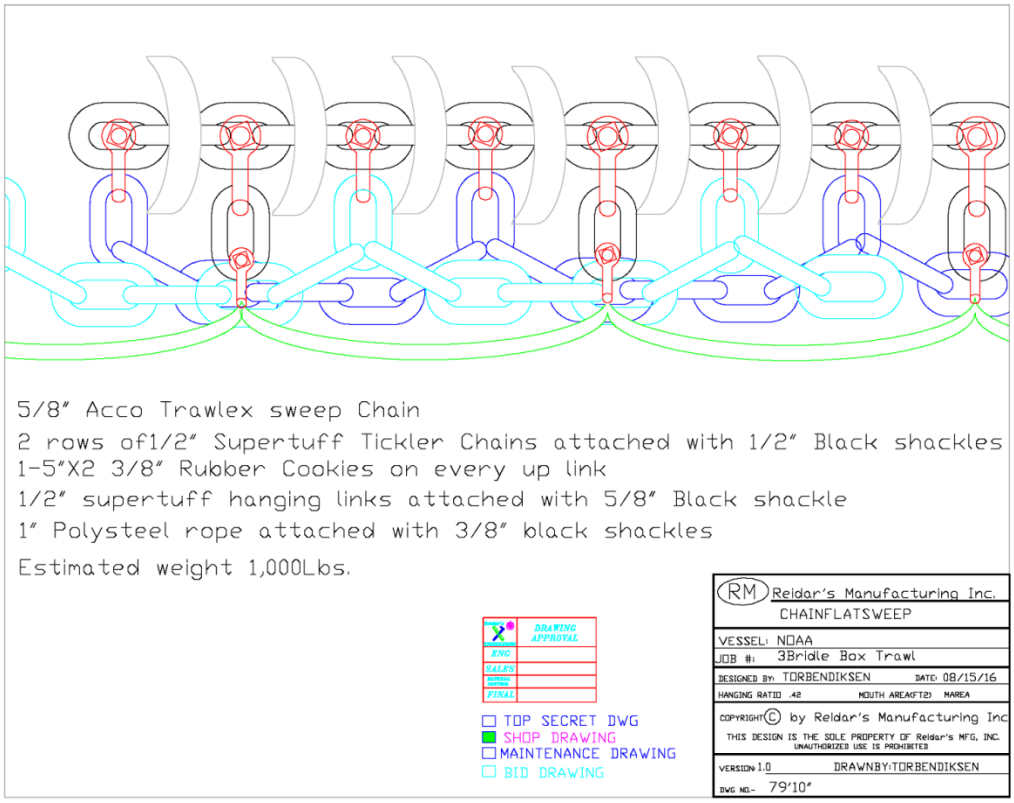

Figure 3. Photo of the rockhopper sweep (left), chain sweep (right) and the clump (center) during the twin trawl experiment.

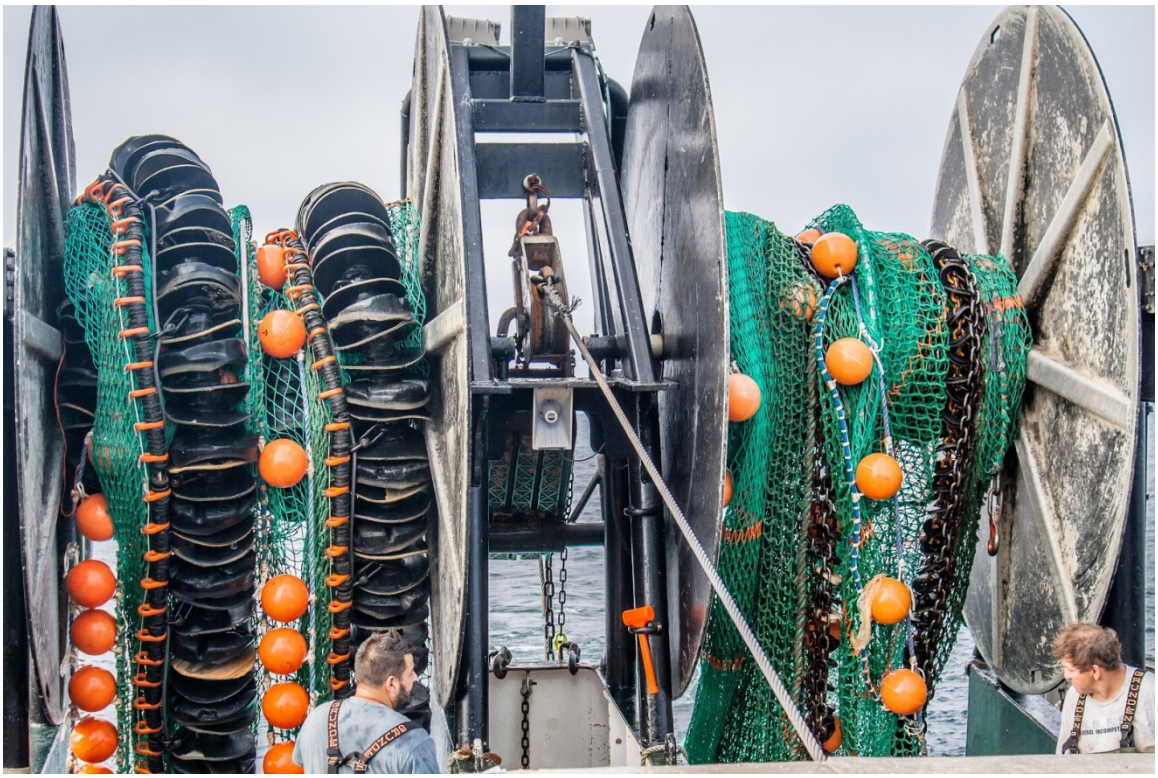

Figure 4. Tow locations during the two sweep studies. Maps on the left correspond to the 2015 sweep study focused on Yellowtail Flounder and Windowpane while maps on the right correspond to the 2016 study focused on Witch flounder.

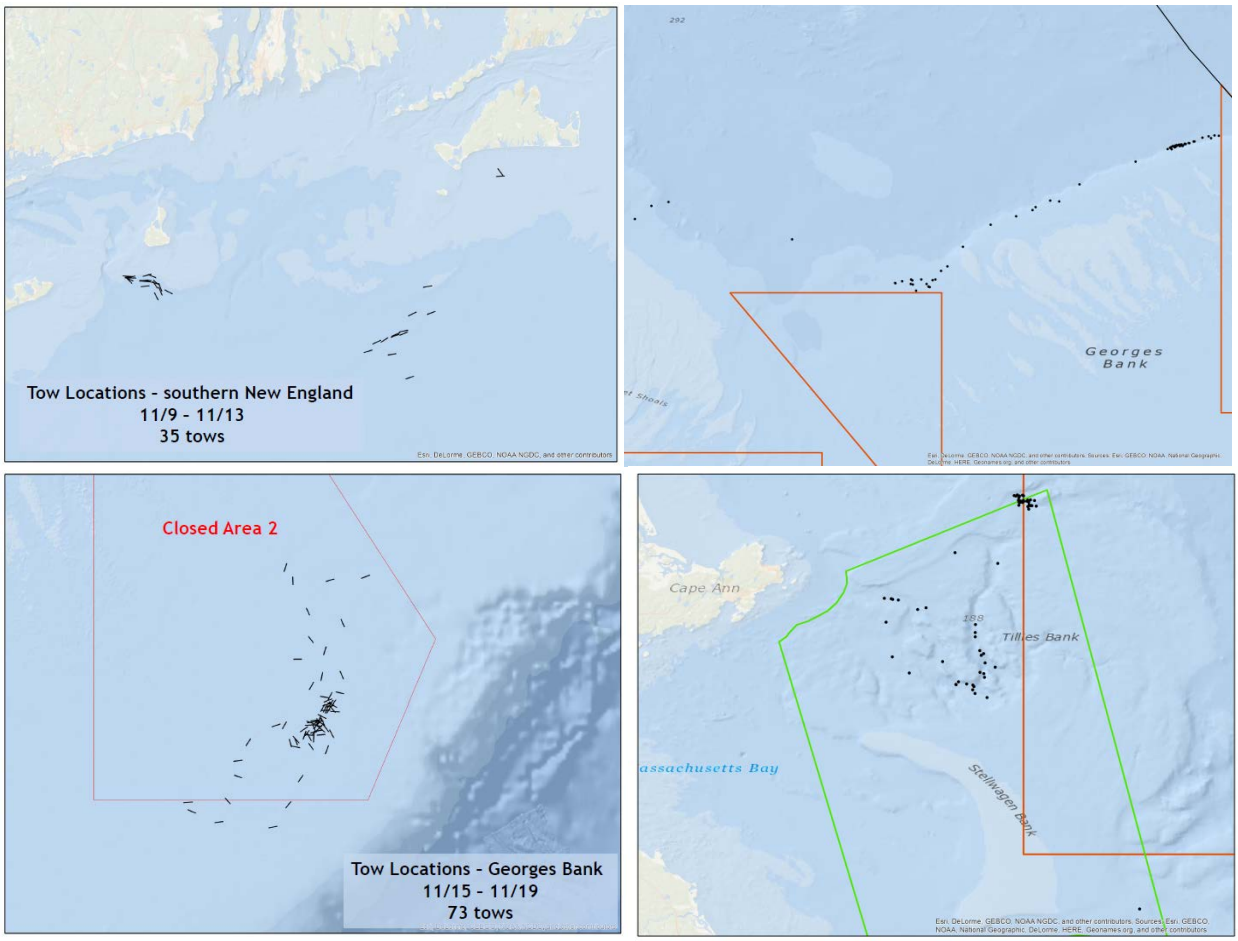

Figure 5. Flow chart of the different analytical steps used to estimate swept area biomass. In total 10 different approaches were used and compared

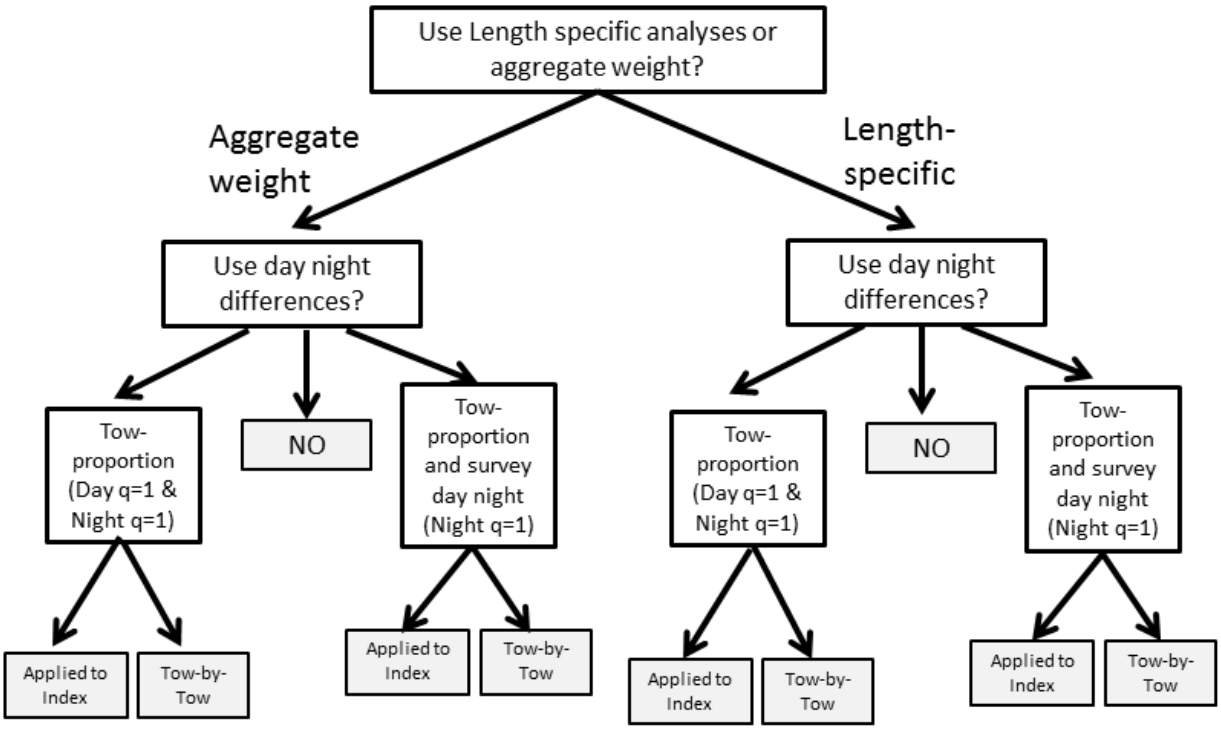

Figure 6. Estimated biomass values of the different approaches (Tables 3-6; Fig. 5) relative to weight only not accounting for day night differences approach. Each point is a year of sampling. Approach 3 and 7 apply a consistent q to weight per tow across all years.

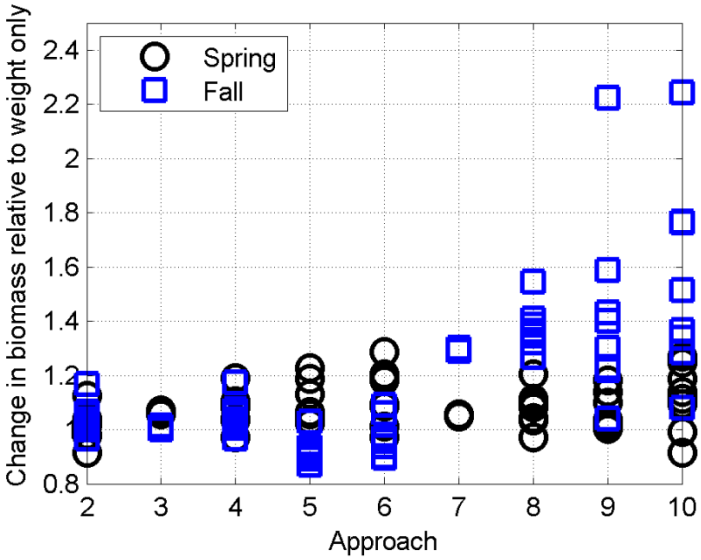

Figure 7. Sensitivity analysis of the effect of difference in catchability with bottom habitat on biomass estimates. The proportion of the swept area biomass attributed to 8 different bottom habitat types was calculated for each year of the survey (top panel) based on assigning each station a bottom habitat type from the Nature Conservancy sediment layer (bottom left panel; Nature Conservancy 2016). The sensitivity analysis was designed to test the scale of bias in the total biomass estimates that would result if the twin trawl results were only valid for smaller grain size bottom types and the rockhopper was more efficient on larger versus smaller grain size bottoms. A value >1 indicates an overestimate of biomass.

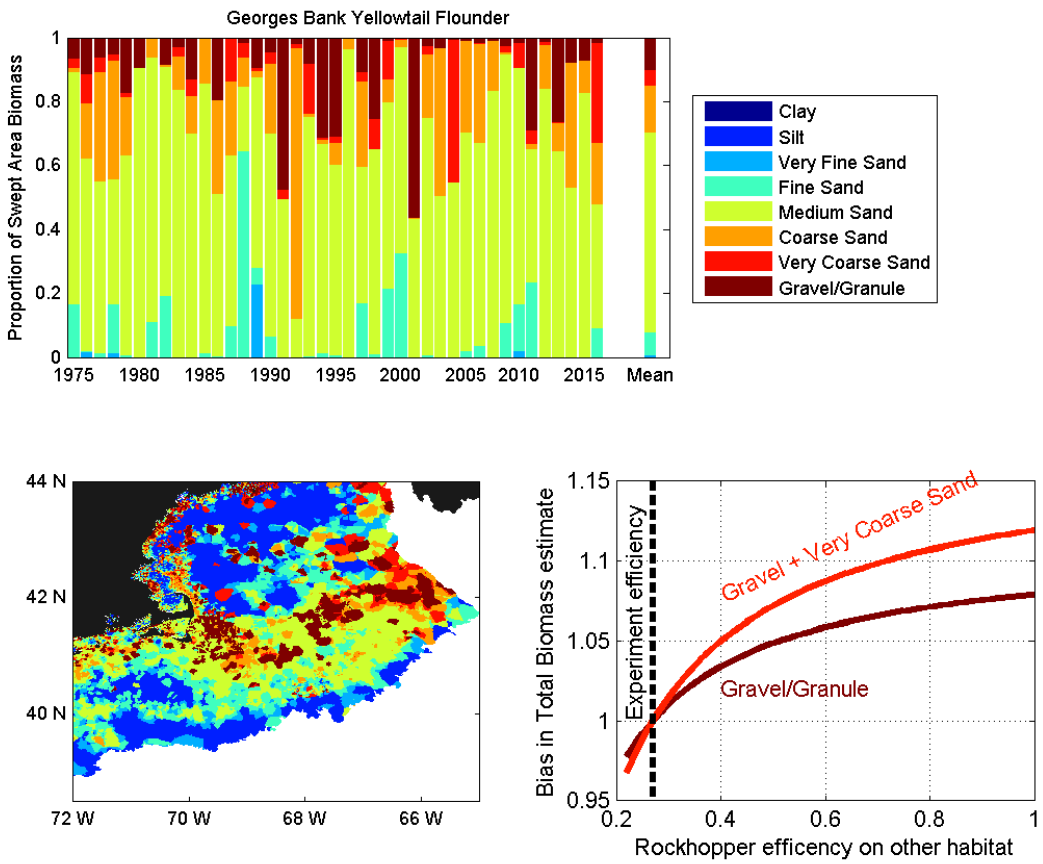

Figure 8. Sensitivity analysis of the effects of using different bridle efficiencies on the total biomass estimate. Past studies have defined bridle efficiency as a function of the width of the bridles in contact with the bottom (or off bottom W_{off}) and the average efficiency in the portion of the bridles in contact with the bottom (h) (Somerton and Munro 2001).