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

Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass

Timothy J. Miller[1], Michael Martin[12], Phil Politis[1], Christopher M. Legault[1], Jessica Blaylock[3]

[1] National Marine Fisheries Service
Northeast Fisheries Science Center
166 Water Street
Woods Hole, MA 02543 USA

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

[3] Integrated Statistics, Inc.
16 Sumner Street
Woods Hole, MA 02543 USA

TABLE OF CONTENTS

ABSTRACT ii
RÉSUMÉ ii
Introduction 1
Methods 1
Data collection 1
relative efficiency of chain and rockhopper sweeps 1
LENGTH-WEIGHT RELATIONSHIPS 1
DIEL EFFECTS ON OBSERVABLE ABUNDANCE FOR NEFSC AND DFO SURVEYS . 2
MODEL ESTIMATION 2
ALTERNATIVE BIOMASS ESTIMATES 2
bootstrap-based estimation of precision 3
results 4
relative efficiency of chain and rockhopper sweeps 4
diel effects on observable abundance 4
Alternative biomass estimates 5
discussion 5
conclusions 6
LITERATURE CITED 6
TABLES 7
FIGURES 22
Appendix A: Analysis of paired trawl experiments using rockhopper and cookie sweep gear 33

ABSTRACT

We use the hierarchical modeling approach from Miller (2013) to estimate the relative efficiency of chain and rockhopper ground-gear (sweeps) for yellowtail flounder (*Limanda ferruginea*) from two studies carried out aboard a twin trawl vessel. As in Miller (2013), we compared a set of models with different assumptions about variation of relative efficiency between paired gear tows, size effects on the relative efficiency, and extra-binomial variation of observations within paired gear tows. However, we further expand the models to investigate diel effects on the relative efficiency. We also evaluate statistical evidence for stratum, size and diel effects on observable abundance of yellowtail flounder catches within the annual Northeast Fisheries Science Center (NEFSC) and Department of Fisheries and Oceans (DFO) surveys from fitting these data with a suite of negative binomial mixed models. Given the best performing models, we make abundance at size estimates for Georges Bank yellowtail flounder as measured by the Bigelow gear applying the estimated change in catch rates between the chain sweep and rockhopper sweep. We make these estimates using three alternative methods that assume no diel effects on relative catch efficiency or the observable abundance, diel effects on the relative catch efficiency only, or diel effects on both. We perform the estimation for all three surveys used to assess Georges Bank yellowtail flounder. We make aggregate biomass estimates by estimating length-weight relationships for each survey, annually. We also make average estimates for 2010-2017 of the 3 annual surveys (NEFSC fall survey from the previous year) corresponding to the current biomass estimation procedure in the TRAC. We estimated uncertainty in all results using bootstrap procedures for each data component.

The best estimates of relative catch efficiency of the rockhopper sweep and chain sweep gears depended on size, diel effects, random size effects by paired observation, and extra-binomial dispersion within paired observations at size. Effects on observable abundance depended on the survey, but size, stratum, and diel effects by year were important for all three surveys. The best models for the fall NEFSC and DFO surveys were the same and included random size effects across tow observations whereas the spring NEFSC survey only required random intercepts across tows. Effective efficiencies of observable biomass expected for the chain sweep gear for each survey were similar for the first two estimation approaches (0.30-0.34). The second method accounted for diel effects on the relative catch efficiency of the rockhopper to chain sweep estimated from the twin trawl study, but neither method attempts to account for diel effects of rockhopper sweep efficiency estimated from the survey data itself. The third estimation method which accounts for yearly diel effects on the rockhopper sweep efficiency, consequently resulted in more variable biomass efficiencies for the three surveys (0.09-0.42) with much larger uncertainty estimates. This greater variability and poorer precision of estimates using the third method are likely a result of a lack of a designed study to estimate diel effects on rockhopper sweep observable abundance and the potential confounding associated with estimating diel effects from the data available. These issues would indicate a preference for using one of the first two methods to estimate biomass of Georges Bank yellowtail flounder. Since the estimates from the first two methods are so similar and the lesser complexity involved in using the first method, we recommend using the first method. Furthermore, the first method can be sufficiently approximated by setting the survey catchability to 0.31 for all surveys and all years currently and may provide an easier to understand and apply approach than the more involved length-based calibration.

RÉSUMÉ

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

INTRODUCTION

This working paper provides estimates for Georges Bank yellowtail flounder of the relative efficiency of the trawl gear used by Northeast Fisheries Science Center to conduct annual bottom trawl surveys to the same gear except with a chain sweep rather than a rockhopper sweep. It also presents three alternative biomass estimates for the stock using the estimated relative efficiency and assuming the chain sweep is 100% efficient. It applies these efficiency estimates to both NEFSC surveys and the DFO survey as is current practice for assessing Georges Bank yellowtail flounder.

METHODS

DATA COLLECTION

Data were collected during two field experiments carried out in 2015 and 2016 aboard the F/V Karen Elizabeth, a 78ft stern trawler capable of towing two trawls simultaneously side by side. One side of the twin-trawl rig towed a NEFSC standard 400 x 12cm survey bottom trawl rigged with the NEFSC standard rockhopper sweep (Politis et al., 2014) (Figure 1). The other side of the twin-trawl rig towed a version the NEFSC 400 x 12cm survey bottom trawl modified to maximize the capture of flatfish. The trawl was modified by reducing the headline floatation from 66 to 32, 20cm, spherical floats, reducing the port and starboard top wing-end extensions by 50cm each and utilizing a chain sweep. The chain sweep was constructed of 1.6cm (5/8in) trawl chain covered by 12.7cm diameter x 1cm thick rubber discs on every other chain link (Figure 2). Two rows of 1.3cm (1/2in) tickler chains were attached to the 1.6cm trawl chain by 1.3cm shackles (Figure 2). To ensure equivalent net geometry of each gear, 32m restrictor ropes, made of 1.4cm (9/16in) buoyant, Polytron rope, were attached between each of the trawl doors and the center clump. 3.4m^2 Thyboron Type 4 trawl doors were used to provide enough spreading force to ensure the restrictor ropes remained taut throughout each tow. Each trawl used the NEFSC standard 36.6m bridles. All tows followed the NEFSC standard survey towing protocols of 20 minutes at 3.0 knots.

RELATIVE EFFICIENCY OF CHAIN AND ROCKHOPPER SWEEPS

In the 2015 study the F/V Karen Elizabeth made 108 (45 day, 63 night) paired tows in eastern Georges Bank and southern New England (Figure 3). In the 2016 study, the F/V Karen Elizabeth made 117 (74 day, 43 night) paired tows in western Gulf of Maine and northern edge of Georges Bank (Figure 4). Yellowtail flounder were caught in 93 of the 108 paired tows in 2015 and 46 of the 117 paired tows in 2016 for a total of 139 paired tow observations.

We fit the same set of 13 models described in Miller (2013, Tables 2 and 3) to all data combined and also to data for observations conducted during the day and night separately. Day and night tows were defined by whether the sun was above the horizon at the time of the tow.

LENGTH-WEIGHT RELATIONSHIPS

We used the length and weight observations for each survey to estimate a survey-specific length-weight relationship. We assumed weight observation *j* from survey *i*, was log-normal distributed,

$$\log W_{i,j} \sim \mathrm{N}\left(\log \alpha_i + \beta_i \log L_{i,j} - \frac{\sigma_i^2}{2}, \sigma_i^2\right) \tag{1}$$

We used a bias correction to ensure the expected $E\left(W_{i,j}\right) = \alpha_i L_{i,j}^{\beta_i}$.

DIEL EFFECTS ON OBSERVABLE ABUNDANCE FOR NEFSC AND DFO SURVEYS

We fit a suite of negative binomial models with and without stratum, diel and smooth size effects to each year of data from the NEFSC spring and fall and DFO surveys. These models considered random smoothers for size effects on observable abundance to account for variability in size composition across tows (Table 1). Like the models considered by Miller (2013), the types of models considered here include those that would be classified as generalized linear mixed models (GLMMs) and generalized additive mixed models (GAMMs), but there is no software readily available to estimate those we wished to consider. The first 6 models consider stratum and size effects on the mean whereas the other 5 models include diel (day/night) effects on the observable abundance. We also fit 4 models to all yearly observations by survey (Table 2). The observations we used were counts by length within each tow so we included random effects for each tow to account for correlation of observations within tows. For models with diel effects, the ratio of night-to-day efficiency in terms of abundance can be estimated by exponentiating the coefficient for the diel effect. For models including stratum effects, there are often strata with all zero tows, and the models will have difficulty estimating some parameters in such situations because the means are estimated on log scale and essentially the parameter estimates are reaching a lower bound at zero. This will result in flags indicating lack of convergence, but we assumed all convergence issues were due to this lack of positive values associated with a given coefficient in the model. The parameters of primary interest (size and diel effects and stratum effects where positive catches occurred) are assumed well estimated in such situations.

MODEL ESTIMATION

For all the models fitted to the twin trawl data, the length-weight data, and the survey data, parameter estimation was performed by maximizing the model likelihood programmed in TMB (Kristensen et al. 2016) and R (R Core Team 2016).

ALTERNATIVE BIOMASS ESTIMATES

We estimated biomass for each annual survey using three approaches. The first does not account for any diel effects on observable abundance of the surveys or diel effects on the relative efficiency of the rockhopper sweep to chain sweep gear. The second also does not account for any diel effects on observable abundance, but does account for diel differences in the relative efficiency of the rockhopper to chain sweep gear estimated from the twin trawl study. The third approach accounts for diel effects on both the observable densities for each survey and the relative efficiency of the rockhopper sweep to chain sweep gear. The main difference is how the numbers at length in each tow are scaled. In some tows for the NEFSC surveys there may be subsampling when a large number of a given species are caught, but we used the extrapolated numbers at length for such tows. Similarly, we used the standardized numbers at length in each of the DFO survey tows. For the first approach

$$\tilde{N}_{1,h,i}(L) = N_{h,i}(L)\hat{\rho}(L) \tag{2}$$

where $N_{h,i}(L)$ is the number at length *L* in tow *i* from stratum *h* and $\hat{\rho}(L)$ is the relative efficiency of the chain sweep to rockhopper sweep at length *L* estimated from the twin trawl observations in a model without diel effects on the relative efficiency. Note that we have omitted any subscripts denoting the year or survey. For the second approach

$$\tilde{N}_{2,h,i}(L) = N_{h,i}(L)\hat{\rho}(L,D)^{I_{h,i}(D)}\hat{\rho}(L,N)^{1-I_{h,i}(D)} \tag{3}$$

where $\hat{\rho}(L,D)$ and $\hat{\rho}(L,N)$ are the relative efficiency of the chain sweep to rockhopper sweep at length for daytime and nighttime tows, respectively, estimated from the twin trawl observations in a model with diel effects on the relative efficiency, and $I_{h,i}(D)$ is the indicator variable of whether tow *i* was conducted during the daytime. For the third approach,

$$\tilde{N}_{3,h,i}(L) = N_{h,i}(L)\hat{\kappa}^{I_{h,i}(D)}\hat{\rho}(L,N) \tag{4}$$

where $\hat{\kappa} = e^{\hat{\beta}_d}$ is the estimated ratio of rockhopper sweep efficiency in night and day from one of the models in Tables 1 and 2. The stratified abundance estimate is then calculated using the design-based estimator,

$$\hat{N}_g(L) = \sum_{h=1}^{H} \frac{W_h}{n_h} \sum_{i=1}^{n_h} \tilde{N}_{g,h,i}(L) \tag{5}$$

where *g* indicates which abundance at length per tow is used and W_h is the fraction of total area or number of possible tows in stratum *h*, and $n_h$is the number of tows that occurred in stratum *h*. The corresponding biomass estimate is then

$$\hat{B}_g = \sum_{l=1}^{n_L} \hat{N}_g(L=l)\hat{w}(L=l) \tag{6}$$

where $\hat{w}(L=l)$ is the estimated weight at length from fitting length-weight observations described above.

Note that in this working paper we have assumed that wingspread is the appropriate width to use when calculating the area swept by a single tow. This differs from the doorspread assumption agreed to during the 2014 Empirical Benchmark for Georges Bank Yellowtail Flounder. Since the wingspread is approximately half of the doorspread, the estimates presented here will be approximately double those in recent TRAC Reference Documents if the same survey catchability is applied. The decision about whether to change from doorspread to wingspread, or not, rests with the TRAC. If the TRAC decides to continue use of doorspread, then the biomass estimates presented in this working paper will be reduced approximately in half.

BOOTSTRAP-BASED ESTIMATION OF PRECISION

We used bootstrap methods for all components of the average biomass estimation. For the estimates of relative catch efficiency of the chain sweep to rockhopper sweep, we made 1000 bootstrap data sets by randomly drawing with replacement the paired observations from the original twin trawl data. We refit the best performing model to each bootstrap data set and retained the predicted relative efficiency (at size) for each bootstrap.

Similarly, bootstrap predictions of weight at length were made by sampling with replacement the length-weight observations within each annual survey and refitting the length-weight relationship to each of the bootstrap datasets.

Bootstrap data sets for each of the annual surveys respected the stratified random designs by resampling with replacement within each stratum. The best performing models for diel and size effects on observable abundance were refitted to each bootstrap data set, and the diel effect and abundance at size for each bootstrap fit as well as the bootstrap survey data sets were retained. Note, that this method of resampling rather than resampling all of the tows in a given annual survey independently will provide a more appropriate and potentially greater variability in the estimates among bootstraps. There were 1, 2, and 4 bootstraps omitted for years 2015, 2017, and 2016, where one of the diel effects of the surveys was estimated extremely large

resulting in average biomass estimates for that year of greater than 1 million mt. Including these would result in extremely large CV estimates for those years, but the point estimates are not affected and confidence intervals are negligibly affected.

For each of the 1000 combined bootstrap results, survey observations for bootstrap *b* were scaled with the corresponding bootstrap estimates of relative cookie sweep to rockhopper sweep efficiency, diel effects (for $\hat{B}_3$), and predicted weight at length, using Eqs. 2-6 for the respective biomass estimation method.

RESULTS

RELATIVE EFFICIENCY OF CHAIN AND ROCKHOPPER SWEEPS

As measured by AIC, the best performing model for the paired gear observations of Georges Bank yellowtail flounder (BB_6) included size effects on the relative efficiency of chain sweep to rockhopper sweep that varied between pairs and extra-binomial dispersion of observations within pairs (Table 3, Figure 5). When all parameters differ by whether twin trawl observations occurred by day or night, but the same model is fitted to both types of observations, the same model performed best (BB_6) and allowing parameters to differ by observation type performed better than when diel differences in observations were ignored (Figure 6).

DIEL EFFECTS ON OBSERVABLE ABUNDANCE

The model that performed best in terms of AIC for observable abundance differed by year and survey. For the NEFSC fall survey (2009-2016), the best performing model included diel effects (NB_9) in 4 of the 8 years (NB_5 otherwise, Table 4). For all yearly NEFSC fall results, size and stratum effects were important. For the NEFSC spring survey (2009-2016), the best performing model included diel effects in 4 (NB_8 in 1 and NB_9 in 3) of the 8 years (NB_2, NB_4, or NB_5 otherwise, Table 5). For all yearly NEFSC spring results, size effects were important and stratum effects were important in all years except 2014. For the DFO survey (2010-2016), the best performing model included diel effects in 4 (NB_8 in 1 and NB_9 in 3) of the 7 years (NB_5 otherwise, Table 6). For all yearly DFO results, size and stratum effects were important.

When fitting each annual set of data for a given survey together with the same model, the best performing model for each survey included, size, stratum, and diel effects on observable abundance (last column of Tables 4-6). The best performing model across all years for the NEFSC fall and DFO surveys (NB_9) included random size effects by tow whereas the best model for the NEFSC spring survey (NB_8) just required random intercepts. The less parameterized 4 models (NB_A to NB_D) we fitted to all data for a given survey never fit as well as when all parameters differed by year (Tables 4-6).

The estimated ratio of night-to-day efficiency of the rockhopper sweep for abundance varied substantially by year for each survey, and the confidence intervals for the yearly estimates indicate substantial uncertainty (Table 7). For the NEFSC fall survey, the annual estimated ratios ranged from 0.51 in 2014 to 13.07 in 2013. The annual 95% confidence intervals imply significant differences from 1 (no diel effect) in the same 4 of the 8 years as AIC indicated models including diel effects to perform best. For the NEFSC spring survey the annual estimated ratios ranged from 0.74 in 2011 to 5.02 in 2012. The annual 95% confidence intervals imply a significant difference from 1 only for 2012. For the DFO survey, the annual estimated ratios ranged from 1.02 in 2010 to 5.64 in 2014. The annual 95% confidence intervals imply a significant difference from 1 only for 2014 and 2017.

ALTERNATIVE BIOMASS ESTIMATES

Using $\hat{B}_1$ from Eqs. 2, 5, and 6, the NEFSC fall biomass estimates declined from about 74000 mt in 2009 to 5300 mt in 2016 (Table 8, Figure 7). The CVs of the estimates ranged from 0.28 in 2009 to 0.6 in 2015. The NEFSC spring biomass estimates declined from 64400 mt in 2010 to 2600 mt in 2017 (Table 9). The CVs of the estimates ranged from 0.19 in 2014 to 0.48 in 2012. The DFO biomass estimates declined from 27100 mt in 2010 to 1460 mt in 2017 (Table 10). The CVs for the estimates ranged from 0.23 in 2013 and 2017 to 0.58 in 2016. When all three surveys were averaged annual biomass estimates declined from 55000 mt in 2010 to 3100 mt in 2017 (Table 11). The CVs of the estimates ranged from 0.16 in 2011 to 0.33 in 2016.

Using $\hat{B}_2$ from Eqs. 3, 5, and 6, the NEFSC fall biomass estimates declined from about 69000 mt in 2009 to 5000 mt in 2016 (Table 12, Figure 7). The CVs of the estimates ranged from 0.26 in 2011 to 0.62 in 2015. The NEFSC spring biomass estimates declined from 62500 mt in 2010 to 2600 mt in 2017 (Table 13). The CVs of the estimates ranged from 0.20 in 2014 to 0.51 in 2012. The DFO biomass estimates declined from 26600 mt in 2010 to 1460 mt in 2017 (Table 14). The CVs for the estimates ranged from 0.22 in 2013 to 0.56 in 2016. When all three surveys were averaged annual biomass estimates declined from 52700 mt in 2010 to 3000 mt in 2017 (Table 15). The CVs of the estimates ranged from 0.16 in 2011 to 0.33 in 2016.

Using $\hat{B}_3$ from Eqs. 4, 5, and 6, the NEFSC fall biomass estimates declined from about 76000 mt in 2009 to 4900 mt in 2015, then increased slightly to 5800 mt in 2016 (Table 16, Figure 7). The CVs of the estimates ranged from 0.39 in 2010 to 2.9 in 2014. The NEFSC spring biomass estimates declined from 88400 mt in 2010 to 3200 mt in 2017 (Table 17). The CVs of the estimates ranged from 0.24 in 2013 to 1.85 in 2016. The DFO biomass estimates declined from 25300 mt in 2010 to 3200 mt in 2017 (Table 18). The CVs for the estimates ranged from 0.34 in 2012 and 2013 to 0.90 in 2017. When all three surveys were averaged annual biomass estimates declined from 63300 mt in 2010 to 4000 mt in 2017 (Table 19). The CVs of the estimates ranged from 0.24 in 2011 to 1.6 in 2015.

The effective efficiency of the biomass estimates can be inferred by dividing the scaled biomass estimates derived using Eqs. 2-6 with the minimum swept-area biomass estimated without any rescaling of the abundance observations. The relative biomass efficiency when using $\hat{B}_1$ ranged from 0.31-0.32, 0.31-0.32, and 0.30-0.32 for the NEFSC fall and spring, and DFO surveys, respectively (Table 20, Figure 8). The relative biomass efficiency when using $\hat{B}_2$ ranged from 0.33-0.34, 0.31-0.33, and 0.30-0.34 for the NEFSC fall and spring, and DFO surveys, respectively (Table 21). The relative biomass efficiency when using $\hat{B}_3$ ranged from 0.0.12-0.42, 0.09-0.38, and 0.13-0.34 for the NEFSC fall and spring, and DFO surveys, respectively (Table 22). The size of the 95% confidence intervals indicates greater uncertainties in the biomass estimates and ratios using than the other biomass estimators (Figures 7 and 8).

DISCUSSION

We found that there were diel effects on both the relative efficiency at size of the rockhopper and chain sweep gear in the twin trawl study and the observable abundance of yellowtail flounder for each survey. However, this diel effect on observable abundance varied annually for each survey, with higher observable abundance estimated in the daytime in some years and vice versa in other years. Furthermore, the ratio of efficiencies in night and day were not significantly different from 1 (no day/night effect) in several years, and there was large uncertainty in the estimated diel effects and resulting biomass estimates.

The NEFSC spring and fall surveys are not stratified by time of day, rather the stations are occupied according to a transit path that is as short as possible while incorporating weather and

start/end of leg issues. This can and does result in highly unbalanced number of tows within a particular season, year, stratum combination (Figure 9). For example, the fall 2009 survey in stratum 1160 occurred almost entirely during the night with only one tow occurring during the day. There is not a consistent ratio of catch rates between the day and night observed when year and stratum is taken into account. This is why the four models that estimated single day-night effects across all years and stratum performed so poorly (models NB_A-NB_D). Yellowtail flounder abundance has declined from 2009 through 2016, resulting in lower overall catches in all strata, but the highest density areas continue to occur in and around the southern portion of Closed Area II (Figure 10).

CONCLUSIONS

We recommend using the estimates from the first method (Table 11) over those from the second or third method. This recommendation is based on the lack of a scientific study to estimate diel effects and the potential (and almost certain) confounding of diel effects in the data collected during the twin trawl study and surveys relative to day-night timing.

LITERATURE CITED

Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., and Bell, B. M. 2016. TMB: Automatic differentiation and Laplace approximation. Journal of Statistical Software 70(5): 1–21.

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316.

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

R Core Team. 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

TABLES

Table 1. Mean (μ) submodels for negative binomial mixed models fitted to annual survey abundance at length observations. A single, constant overdispersion parameter (τ) was estimated in all models. A subscript d indicates that the parameter depends on whether the tow was conducted during the day or night and $I_i(D)$ is an indicator variable equal to one if the tow occurred during day time.

Model	$\log \mu$	Across-station effects on mean	Station-specific random effects on mean	Random effects Variance
NB_0	$\beta_0 + \delta_{0,i}$	Constant intercepts	Intercept	Σ
NB_1	$\beta_0 + \delta_{0,\boldsymbol{i}} + \boldsymbol{X_f^T}\beta_L + \boldsymbol{X_r^T}\boldsymbol{b}$	Smoother for size	Intercept	Σ
NB_2	$\beta_0 + \delta_{0,\boldsymbol{i}} + \boldsymbol{X_f^T}(\beta_L + \delta_{L,i}) + \boldsymbol{X_r^T}(\boldsymbol{b} + \varepsilon_{\boldsymbol{i}})$	Smoother for size	Smoother for size	Σ
NB_3	$\beta_{0,\boldsymbol{h}} + \delta_{0,\boldsymbol{i}}$	Stratum-specific intercepts	Intercept	Σ
NB_4	$\beta_{0,\boldsymbol{h}} + \delta_{0,\boldsymbol{i}} + \boldsymbol{X_f^T}\beta_L + \boldsymbol{X_r^T}\boldsymbol{b}$	Stratum-specific intercepts and smoother for size	Intercept	Σ
NB_5	$\beta_{0,\boldsymbol{h}} + \delta_{0,\boldsymbol{i}} + \boldsymbol{X_f^T}(\beta_L + \delta_{L,i}) + \boldsymbol{X_r^T}(\boldsymbol{b} + \varepsilon_{\boldsymbol{i}})$	Stratum-specific intercepts and smoother for size	Smoother for size	Σ
NB_6	$\beta_0 + \beta_{\boldsymbol{d}} I_i(D) + \delta_{0,\boldsymbol{i}}$	Day/night-specific intercepts	Intercept	Σ
NB_7	$\beta_{0,\boldsymbol{h}} + \beta_{\boldsymbol{d}} I_i(D) + \delta_{0,\boldsymbol{i}}$	Stratum-specific intercepts and day/night effects	Intercept	Σ
NB_8	$\beta_{0,\boldsymbol{h}} + \beta_{\boldsymbol{d}} I_i(D) + \delta_{0,\boldsymbol{i}} + \boldsymbol{X_f^T}\beta_L + \boldsymbol{X_r^T}\boldsymbol{b}$	Stratum-specific intercepts, day/night effects and smoother for size	Intercept	Σ
NB_9	$\beta_{0,\boldsymbol{h}} + \beta_{\boldsymbol{d}} I_i(D) + \delta_{0,\boldsymbol{i}} + \boldsymbol{X_f^T}(\beta_L + \delta_{L,i}) + \boldsymbol{X_r^T}(\boldsymbol{b} + \varepsilon_{\boldsymbol{i}})$	Stratum-specific intercepts, day/night effects and smoother for size	Smoother for size	Σ
NB_{10}	$\beta_{0,\boldsymbol{h},\boldsymbol{d}} + \delta_{0,\boldsymbol{i}} + \boldsymbol{X_f^T}(\beta_{L,\boldsymbol{d}} + \delta_{L,i}) + \boldsymbol{X_r^T}(\boldsymbol{b_d} + \varepsilon_{\boldsymbol{i}})$	Stratum- and day/night-specific intercepts and day/night-specific smoother for size	Smoother for size	$\Sigma_{\boldsymbol{d}}$

Table 2. Mean (μ) submodels for negative binomial mixed models fitted to all annual abundance at length observations from a specific survey (NEFSC fall or spring or DFO). A subscript d indicates that the parameter depends on whether the tow was conducted during the day or night and $I_i(D)$ is an indicator variable equal to one if the tow occurred during day time. Subscripts h and y indicate that the parameter depends on stratum and year, respectively.

Model	$\log \mu$	Across-station effects on mean	Station-specific random effects on mean	Random effects Variance
NB_A	$\beta_{0,d} + \delta_{0,i}$	Day/night-specific intercepts	Intercept	Σ
NB_B	$\beta_{0,y,h} + \beta_d I_i(D) + \delta_{0,i}$	Year- and stratum-specific intercepts and day/night effect	Intercept	Σ
NB_C	$\beta_{0,y,h} + \beta_d I_i(D) + \delta_{0,i} + X_f^T \beta_L + X_r^T b$	Year- and stratum-specific intercepts, day/night effect, and smoother for size	Intercept	Σ
NB_D	$\beta_{0,y,h} + \beta_{d,y} I_i(D) + \delta_{0,i} + X_f^T \beta_L + X_r^T b$	Year- and stratum-specific intercepts, year-specific day/night effects, and smoother for size	Intercept	Σ

Table 3. Difference in AIC for each model from that with the lowest AIC. See Miller (2013) for descriptions of each model.

	Model	n_p	ΔAIC
All data	BI_0	1	221.4
	BI_1	2	35.3
	BI_2	3	211.5
	BI_3	4	22.9
	BI_4	7	8.0
	BB_0	2	165.4
	BB_1	3	24.2
	BB_2	4	159.4
	BB_3	6	162.7
	BB_4	5	13.4
	BB_5	7	17.1
	BB_6	8	1.8
	BB_7	10	5.8
Day/night-specific models	BI_0	2	156.1
	BI_1	4	27.9
	BI_2	6	139.3
	BI_3	8	10.6
	BI_4	14	5.6
	BB_0	4	120.5
	BB_1	6	18.2
	BB_2	8	110.2
	BB_3	12	114.4
	BB_4	10	2.7
	BB_5	14	7.9
	BB_6	16	0.0
	BB_7	20	5.7

Table 4. The AIC for models fit to NEFSC fall survey data in each year and the AIC for all years of data, fitting the same model annually.

Model	2009	2010	2011	2012	2013	2014	2015	2016	Total
NB_0	995.2	695.0	596.7	448.6	349.6	400.9	165.6	190.0	3835.9
NB_1	123.4	78.6	67.4	47.9	47.3	75.5	35.4	46.9	516.7
NB_2	22.0	35.0	34.8	18.6	25.6	26.1	22.7	12.6	191.9
NB_3	972.5	661.0	572.0	428.0	337.8	372.8	141.9	177.3	3657.8
NB_4	100.9	44.9	42.7	27.6	35.0	47.6	10.8	33.7	337.3
NB_5	0.0	4.9	7.3	3.3	12.6	0.0	0.0	0.0	22.5
NB_6	988.0	696.4	594.3	450.3	347.8	402.7	165.7	192.0	3831.6
NB_7	973.4	655.3	569.0	425.2	323.9	374.3	143.9	178.2	3637.5
NB_8	101.7	39.6	39.6	24.8	21.3	49.1	12.7	34.6	317.7
NB_9	0.9	0.0	0.0	0.0	0.0	1.4	2.0	1.3	0.0
NB_{10}	20.7	26.2	24.2	19.9	26.5	16.3	27.2	23.7	178.8
NB_A									3853.7
NB_B									3632.1
NB_C									343.9
NB_D									343.6

Table 5. The AIC for models fit to NEFSC spring survey data in each year and the AIC for all years of data, fitting the same model annually.

Model	2009	2010	2011	2012	2013	2014	2015	2016	Total
NB_0	872.8	1075.6	764.6	748.0	452.3	357.0	827.0	130.2	4797.8
NB_1	199.5	70.7	47.1	111.1	70.4	118.5	557.9	5.2	750.5
NB_2	36.4	29.7	36.5	35.3	41.3	0.0	562.7	8.4	320.4
NB_3	835.6	1044.4	729.1	721.7	409.0	341.9	814.0	125.5	4591.5
NB_4	163.0	39.8	11.9	84.7	27.2	103.6	544.8	0.0	545.1
NB_5	0.4	2.2	0.0	8.0	0.0	93.5	549.5	2.0	225.8
NB_6	873.9	1077.5	763.6	748.9	452.3	358.6	828.8	129.9	4803.7
NB_7	836.0	1043.8	730.7	716.6	410.7	343.7	815.0	125.8	4592.6
NB_8	163.3	39.2	13.6	79.2	28.8	105.4	0.0	0.3	0.0
NB_9	0.0	0.0	2.0	0.0	1.1	94.9	550.3	3.5	222.0
NB_{10}	18.6	27.5	28.4	22.9	34.0	114.6	566.1	28.5	410.7
NB_A									4830.4
NB_B									4594.5
NB_C									551.1
NB_D									552.6

Table 6. The AIC for models fit to DFO survey data in each year and the AIC for all years of data, fitting the same model annually.

Model	2009	2010	2011	2012	2013	2014	2015	2016	Total
NB_0		1142.6	993.9	1626.5	425.3	244.6	257.2	347.1	5027.0
NB_1		35.0	28.2	47.8	29.2	8.0	6.8	22.7	167.4
NB_2		22.1	18.6	20.7	2.6	8.4	12.7	4.9	79.8
NB_3		1115.4	973.7	1604.2	420.3	241.8	251.5	342.8	4939.6
NB_4		8.1	7.3	25.7	23.8	4.2	1.0	15.0	74.8
NB_5		0.0	1.7	0.0	0.0	5.3	7.0	0.1	3.9
NB_6		1144.5	995.4	1625.5	427.0	243.3	256.9	348.1	5030.5
NB_7		1117.4	973.0	1605.5	421.1	238.9	250.4	342.4	4938.7
NB_8		10.1	6.6	27.1	24.5	1.2	0.0	14.6	73.9
NB_9		2.0	0.0	1.7	0.6	0.0	6.0	0.0	0.0
NB_{10}		23.3	19.1	29.6	13.2	27.1	16.7	25.6	144.4
NB_A									5084.6
NB_B									4956.5
NB_C									135.5
NB_D									144.2

Table 7. Annual estimated diel ratio from the best fitting model for each survey with bootstrap-based 95% confidence intervals in parentheses.

Year	NEFSC fall (NB_9)	NEFSC spring (NB_8)	DFO (NB_9)
2009	2.30 (0.40, 22.37)		
2010	3.80 (1.23, 11.79)	2.33 (0.92, 6.04)	1.02 (0.17, 6.91)
2011	8.35 (2.93, 28.43)	0.74 (0.28, 2.59)	3.11 (0.99, 11.22)
2012	9.57 (1.56, 55.89)	5.02 (1.75, 12.30)	1.37 (0.36, 5.50)
2013	13.07 (2.39, 47.27)	1.25 (0.61, 2.39)	1.93 (0.57, 4.86)
2014	0.51 (0.02, 8.04)	1.30 (0.39, 4.12)	5.64 (1.37, 23.12)
2015	0.92 (0.22, 3.63)	1.64 (0.72, 4.65)	2.37 (0.89, 6.32)
2016	2.01 (0.30, 10.47)	2.49 (0.95, 13.35)	2.91 (0.69, 18.12)
2017		1.54 (0.49, 4.84)	4.66 (1.71, 34.32)

Table 8. Biomass estimates ($\hat{B}_1$) from observable abundance at length and length-weight observations from the NEFSC fall survey scaled by day and night relative efficiency estimates from twin trawl studies using Eqs. 2, 5, and 6. Coefficients of variation (CVs) and 95% confidence intervals (CI) are based on 1000 bootstrap data sets for each set of observations.

Year	Biomass (mt)	CV	CI
2009	73980	0.28	38973, 120612
2010	26220	0.29	12898, 41533
2011	27998	0.26	15030, 43120
2012	29211	0.46	7145, 59832
2013	10279	0.37	4115, 18739
2014	11902	0.30	5486, 19512
2015	5419	0.60	1366, 12808
2016	5305	0.36	1945, 9393

Table 9. Biomass estimates ($\hat{B}_1$) from observable abundance at length and length-weight observations from the NEFSC spring survey scaled by day and night relative efficiency estimates from twin trawl studies using Eqs. 2, 5, and 6. Coefficients of variation (CVs) and 95% confidence intervals (CI) are based on 1000 bootstrap data sets for each set of observations.

Year	Biomass (mt)	CV	CI
2010	64351	0.26	36953, 103547
2011	28444	0.23	16383, 41278
2012	44112	0.48	16190, 93877
2013	11954	0.23	7408, 17872
2014	7821	0.19	5055, 10874
2015	5467	0.21	3477, 8003
2016	3270	0.25	1822, 4917
2017	2584	0.24	1423, 3776

Table 10. Biomass estimates ($\hat{B}_1$) from observable abundance at length and length-weight observations from the DFO survey scaled by day and night relative efficiency estimates from twin trawl studies using Eqs. 2, 5, and 6. Coefficients of variation (CVs) and 95% confidence intervals (CI) are based on 1000 bootstrap data sets for each set of observations.

Year	Biomass (mt)	CV	CI
2010	27117	0.28	14401, 44074
2011	16128	0.26	8932, 25422
2012	25016	0.25	14221, 38919
2013	3339	0.23	1965, 4911
2014	1818	0.30	946, 3053
2015	2713	0.35	1143, 4764
2016	5693	0.58	1274, 12801
2017	1461	0.23	857, 2150

Table 11. Average of biomass estimates ($\hat{B}_1$) from the NEFSC fall and spring surveys and the DFO surveys (Tables 8 to 10). Coefficients of variation (CVs) and 95% confidence intervals (CI) are based on 1000 bootstrap data sets for each set of observations.

Year	Biomass (mt)	CV	CI
2010	55150	0.17	38048, 76236
2011	23597	0.16	16459, 31024
2012	32375	0.24	19850, 49772
2013	14824	0.31	7525, 25117
2014	6639	0.21	4225, 9868
2015	6694	0.20	4389, 9811
2016	4794	0.33	2168, 8305
2017	3117	0.22	1882, 4583

Table 12. Biomass estimates ($\hat{B}_2$) from observable abundance at length and length-weight observations from the NEFSC fall survey scaled by day and night relative efficiency estimates from twin trawl studies using Eqs. 3, 5, and 6. Coefficients of variation (CVs) and 95% confidence intervals (CI) are based on 1000 bootstrap data sets for each set of observations.

Year	Biomass (mt)	CV	CI
2009	68998	0.27	(38574, 111189
2010	24719	0.28	(12513, 39299
2011	26086	0.26	(14460, 39496
2012	27237	0.45	(7385, 55362
2013	9637	0.36	(4303, 17793
2014	11571	0.32	(5586, 19992
2015	5060	0.62	(1317, 12282)
2016	5003	0.35	(1772, 8339)

Table 13. Biomass estimates ($\hat{B}_2$) from observable abundance at length and length-weight observations from the NEFSC spring survey scaled by day and night relative efficiency estimates from twin trawl studies using Eqs. 3, 5, and 6. Coefficients of variation (CVs) and 95% confidence intervals (CI) are based on 1000 bootstrap data sets for each set of observations.

Year	Biomass (mt)	CV	CI
2010	62532	0.27	(35478, 100697)
2011	27888	0.24	(15997, 41632)
2012	45182	0.51	(15783, 100506)
2013	11790	0.24	(7251, 18123)
2014	7937	0.20	(5111, 11408)
2015	5486	0.22	(3470, 8213)
2016	3204	0.25	(1876, 5009)
2017	2627	0.25	(1525, 4079)

Table 14. Biomass estimates ($\hat{B}_2$) from observable abundance at length and length-weight observations from the DFO survey scaled by day and night relative efficiency estimates from twin trawl studies using Eqs. 3, 5, and 6. Coefficients of variation (CVs) and 95% confidence intervals (CI) are based on 1000 bootstrap data sets for each set of observations.

Year	Biomass (mt)	CV	CI
2010	26605	0.28	(13435, 42980)
2011	15908	0.25	(8652, 24640)
2012	23592	0.24	(14257, 35814)
2013	3246	0.22	(2019, 4871)
2014	1770	0.30	(899, 3000)
2015	2588	0.34	(1128, 4550)
2016	5299	0.56	(1241, 11881)
2017	1458	0.21	(930, 2129

Table 15. Average of biomass estimates ($\hat{B}_2$) from the NEFSC fall and spring surveys and the DFO surveys (Tables 12 to 14). Coefficients of variation (CVs) and 95% confidence intervals (CI) are based on 1000 bootstrap data sets for each set of observations.

Year	Biomass (mt)	CV	CI
2010	52712	0.17	(36495, 74663)
2011	22838	0.16	(16110, 30839)
2012	31620	0.26	(19421, 50418)
2013	14091	0.30	(7089, 23403)
2014	6448	0.20	(4269, 9219)
2015	6549	0.21	(4281, 9499)
2016	4521	0.33	(2118, 7651)
2017	3029	0.21	(1798, 4264)

Table 16. Biomass estimates ($\hat{B}_3$) from observable abundance at length and length-weight observations from the NEFSC fall survey scaled by day and night relative efficiency estimates from twin trawl studies using Eqs. 4-6. Coefficients of variation (CVs) and 95% confidence intervals (CI) are based on 1000 bootstrap data sets for each set of observations.

Year	Biomass (mt)	CV	CI
2009	76184	0.90	37143, 158891
2010	35319	0.39	15696, 75367
2011	49008	0.45	22275, 115068
2012	39925	0.58	9568, 106470
2013	26715	0.63	8527, 62082
2014	9003	2.92	3136, 35334
2015	4935	0.65	1115, 12578
2016	5763	0.51	1473, 13201

Table 17. Biomass estimates ($\hat{B}_3$) from observable abundance at length and length-weight observations from the NEFSC spring survey scaled by day and night relative efficiency estimates from twin trawl studies using Eqs. 4-6. Coefficients of variation (CVs) and 95% confidence intervals (CI) are based on 1000 bootstrap data sets for each set of observations.

Year	Biomass (mt)	CV	CI
2010	88384	0.36	42531, 176685
2011	23232	0.31	12230, 41829
2012	149142	0.75	29701, 406373
2013	11943	0.24	6971, 17713
2014	8454	0.45	3271, 18576
2015	6599	0.42	3514, 14567
2016	4802	1.85	2272, 13875
2017	3189	0.64	1241, 7926

Table 18. Biomass estimates ($\hat{B}_3$) from observable abundance at length and length-weight observations from the DFO survey scaled by day and night relative efficiency estimates from twin trawl studies using Eqs. 4-6. Coefficients of variation (CVs) and 95% confidence intervals (CI) are based on 1000 bootstrap data sets for each set of observations.

Year	Biomass (mt)	CV	CI
2010	25335	0.63	8901, 87579
2011	26088	0.46	11108, 58423
2012	24361	0.34	12327, 47802
2013	3943	0.34	1858, 6773
2014	4251	0.74	1398, 12112
2015	3169	0.36	1259, 5900
2016	6529	0.63	1433, 16710
2017	3206	0.90	1501, 16431

Table 19. Average of biomass estimates ($\hat{B}_3$) from the NEFSC fall and spring surveys and the DFO surveys (Tables 16 to 18). Coefficients of variation (CVs) and 95% confidence intervals (CI) are based on 1000 bootstrap data sets for each set of observations.

Year	Biomass (mt)	CV	CI
2010	63301	0.42	39869, 108899
2011	28213	0.24	18554, 45945
2012	74170	0.49	31902, 162057
2013	18604	0.43	8142, 41596
2014	13140	0.44	6566, 25984
2015	6257	1.64	3802, 15580
2016	5422	0.68	2601, 11029
2017	4052	0.41	2451, 10034

Table 20. Ratios of observable to scaled biomass estimates for the NEFSC and DFO surveys using $\hat{B}_1$ *and Eqs. 3, 6, and 7.*

Year	Fall NEFSC	Spring NEFSC	DFO
2009	0.32		
2010	0.31	0.31	0.32
2011	0.31	0.31	0.31
2012	0.32	0.31	0.32
2013	0.31	0.31	0.31
2014	0.32	0.32	0.30
2015	0.32	0.31	0.30
2016	0.31	0.32	0.31
2017		0.32	0.30

Table 21. Ratios of observable to scaled biomass estimates for the NEFSC and DFO surveys using $\hat{B}_2$ *and Eqs. 4, 6, and 7.*

Year	Fall NEFSC	Spring NEFSC	DFO
2009	0.34		
2010	0.33	0.32	0.32
2011	0.33	0.32	0.32
2012	0.34	0.31	0.34
2013	0.33	0.31	0.32
2014	0.33	0.31	0.31
2015	0.34	0.31	0.32
2016	0.33	0.33	0.33
2017		0.31	0.30

Table 22. Ratios of observable to scaled biomass estimates for the NEFSC and DFO surveys using $\hat{B}_3$ *and Eqs. 5, 6, and 7.*

Year	Fall NEFSC	Spring NEFSC	DFO
2009	0.31		
2010	0.23	0.23	0.34
2011	0.18	0.38	0.20
2012	0.23	0.09	0.32
2013	0.12	0.31	0.27
2014	0.42	0.29	0.13
2015	0.35	0.26	0.26
2016	0.29	0.22	0.27
2017		0.26	0.14

FIGURES

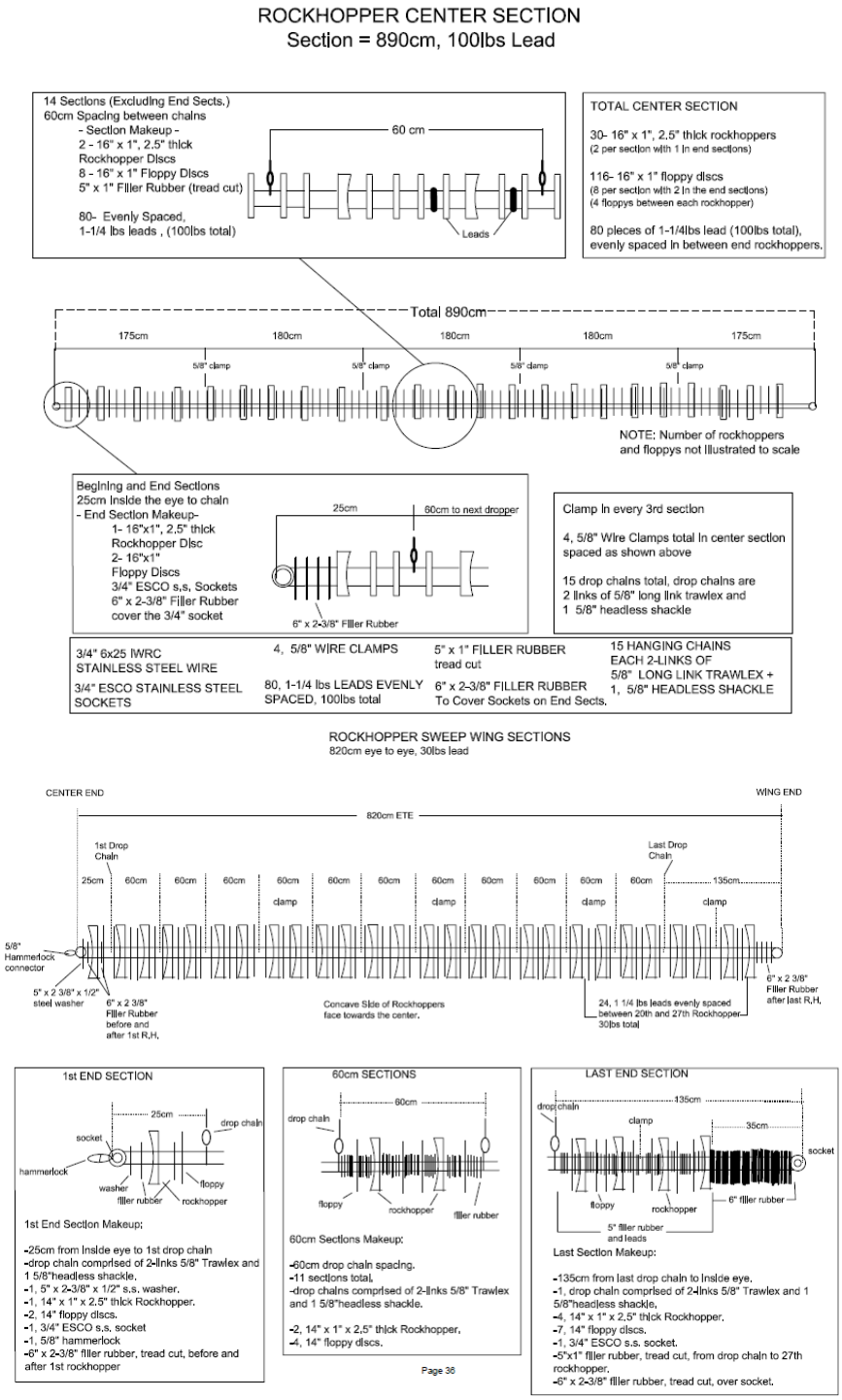

Figure 1. Diagram of the standard Northeast Fisheries Science Center rockhopper sweep center and wing sections.

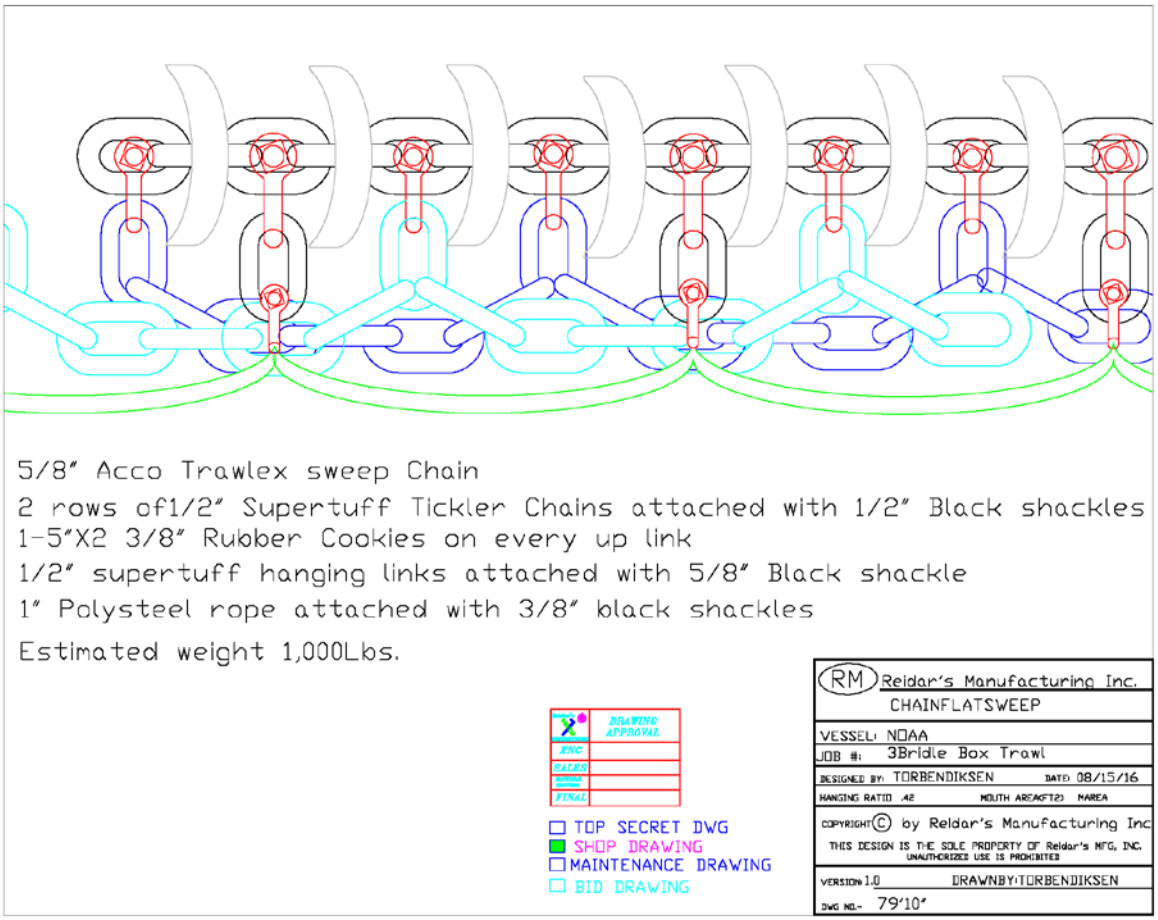

Figure 2. Diagram of the chain sweep designed maximize bottom contact and flatfish capture.

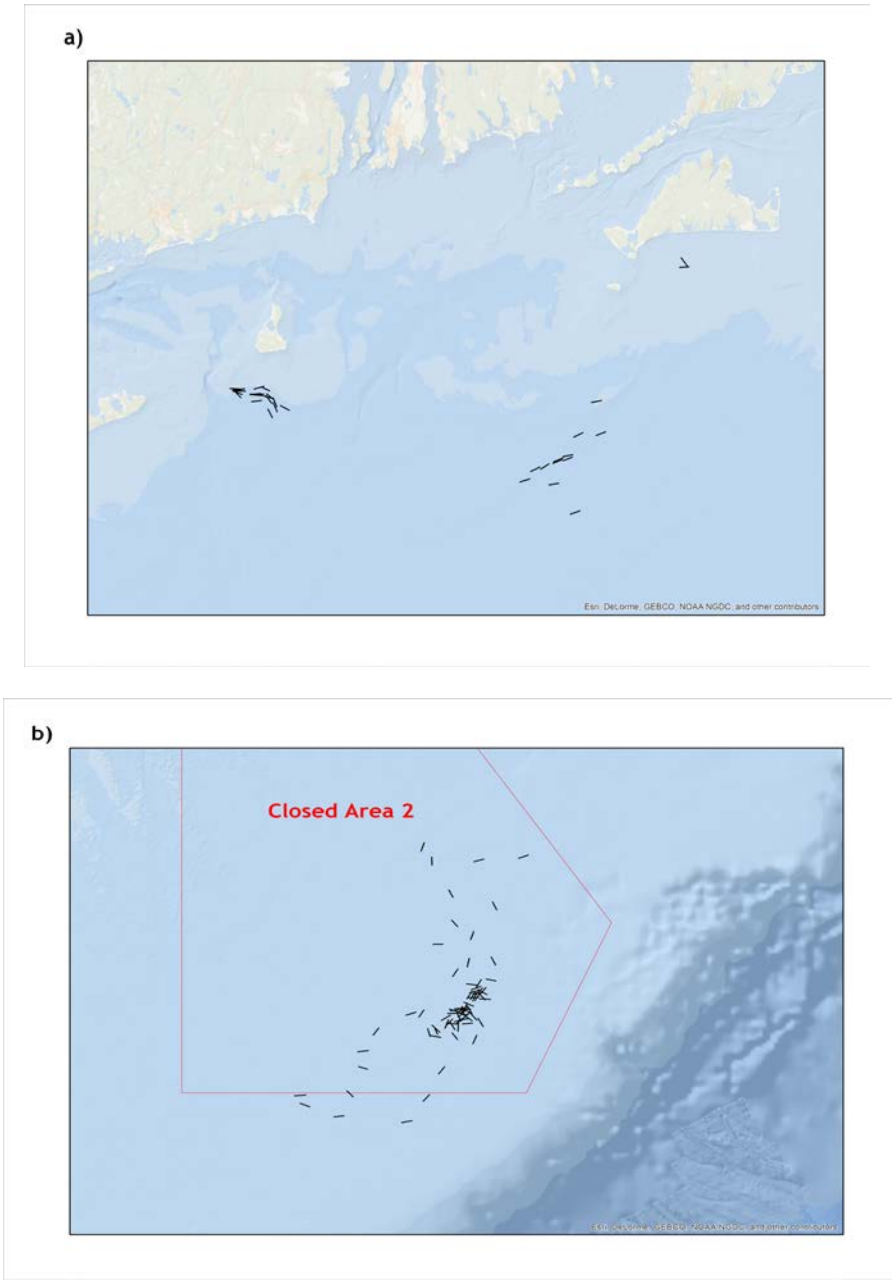

Figure 3. Twin-trawl tow locations conducted in 2015. a) Southern New England b) Georges Bank.

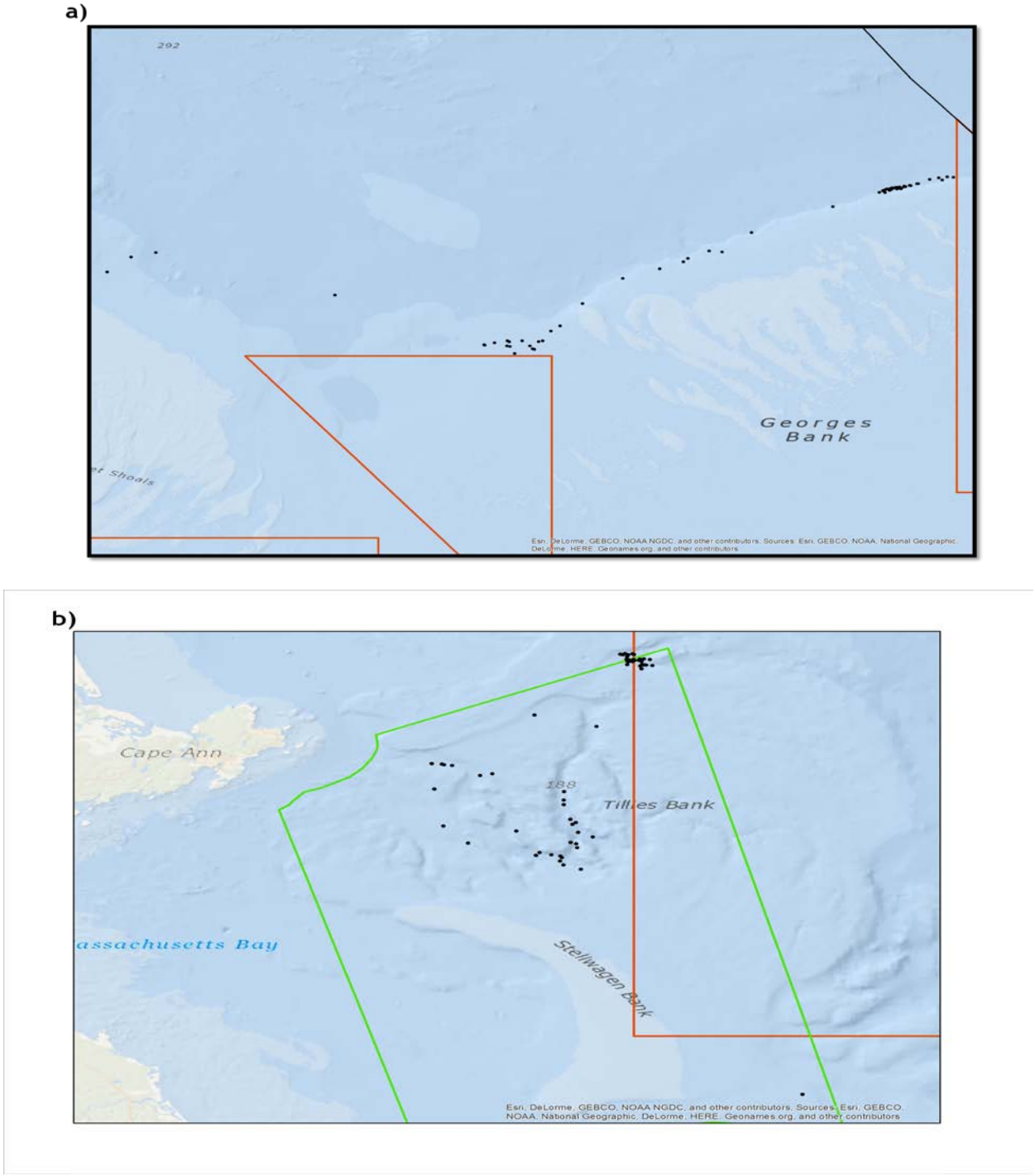

Figure 4. Twin-trawl locations conducted in 2016. a) Northern Edge of Georges Bank b) Western Gulf of Maine.

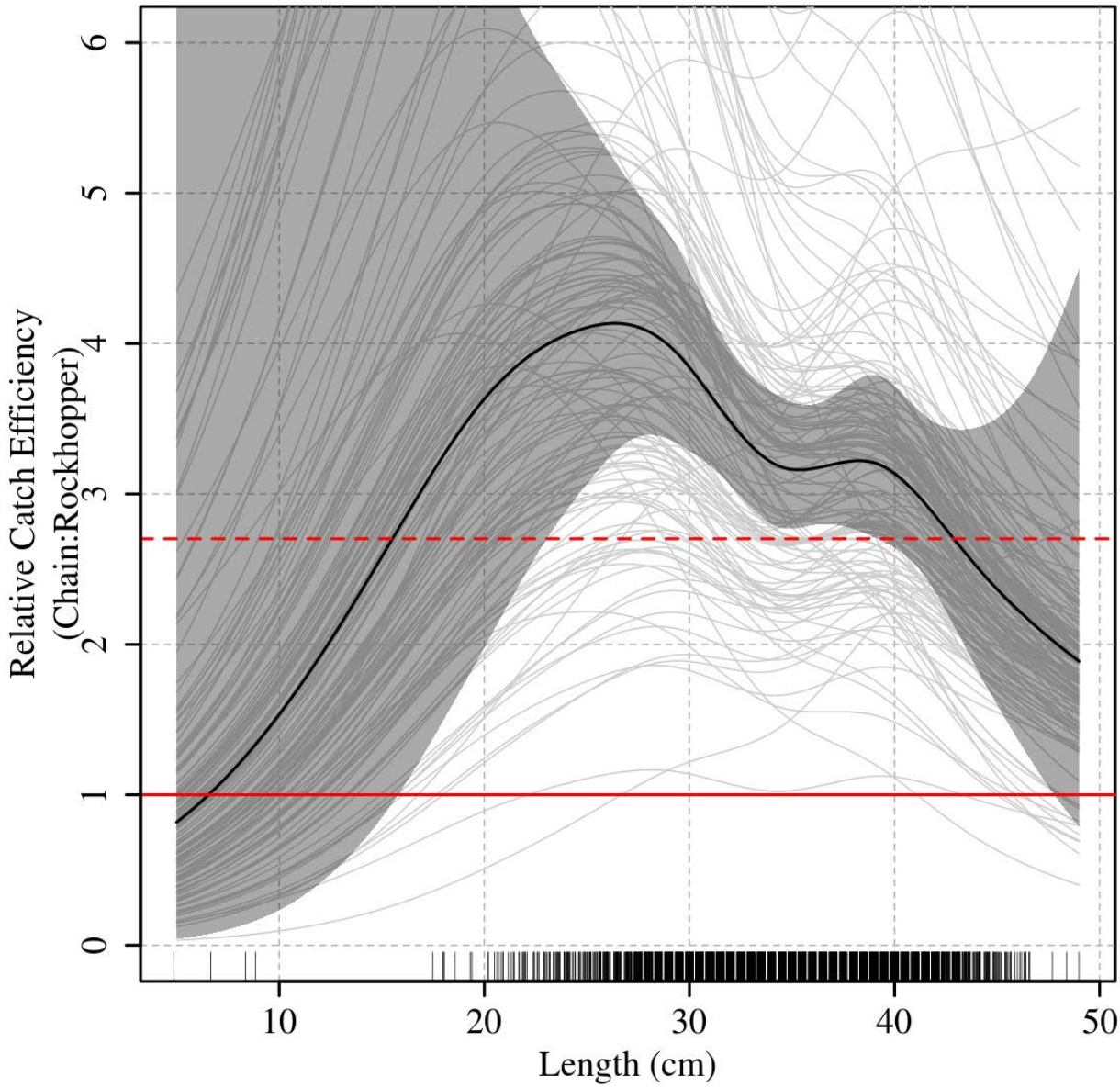

Figure 5. Relative chain sweep to rockhopper efficiency at size for Georges Bank yellowtail flounder from the best performing model (BB_6) not considering day-night differences in model parameters. Light gray lines are pair-specific estimates of relative efficiency, red line indicates equal efficiency, gray polygon represent 95% confidence interval for the mean relative efficiency (black line) and rug indicates observations at length.

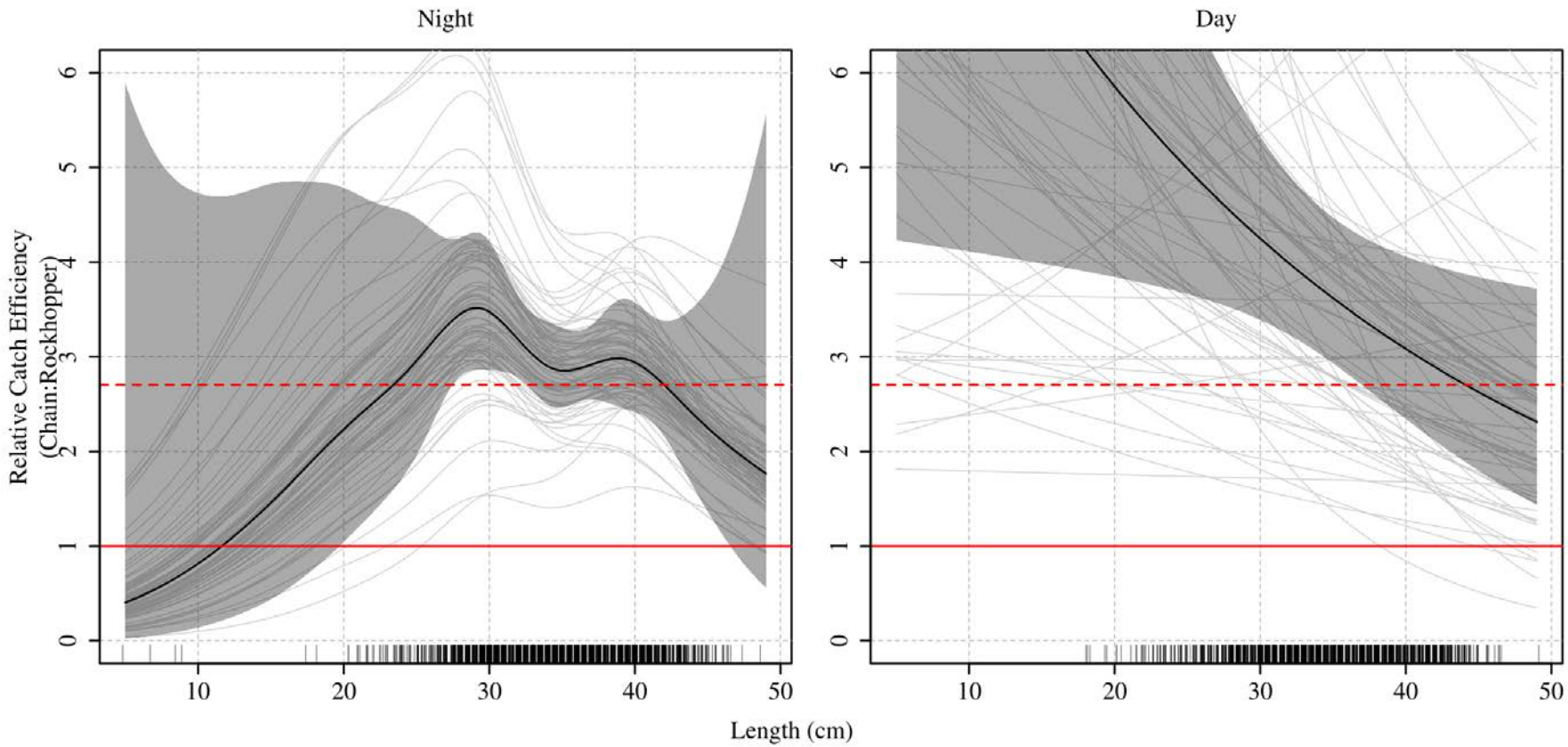

Figure 6. Relative chain sweep to rockhopper efficiency at size for Georges Bank yellowtail flounder from the best performing model (BB_6) fitted separately to day and night observations. Light gray lines are pair-specific estimates of relative efficiency, red line indicates equal efficiency, red dashed line indicates 0.37 efficiency currently used for biomass estimation in the TRAC, and gray polygon represent 95% confidence interval for the mean relative efficiency (black line) and rug indicates observations at length.

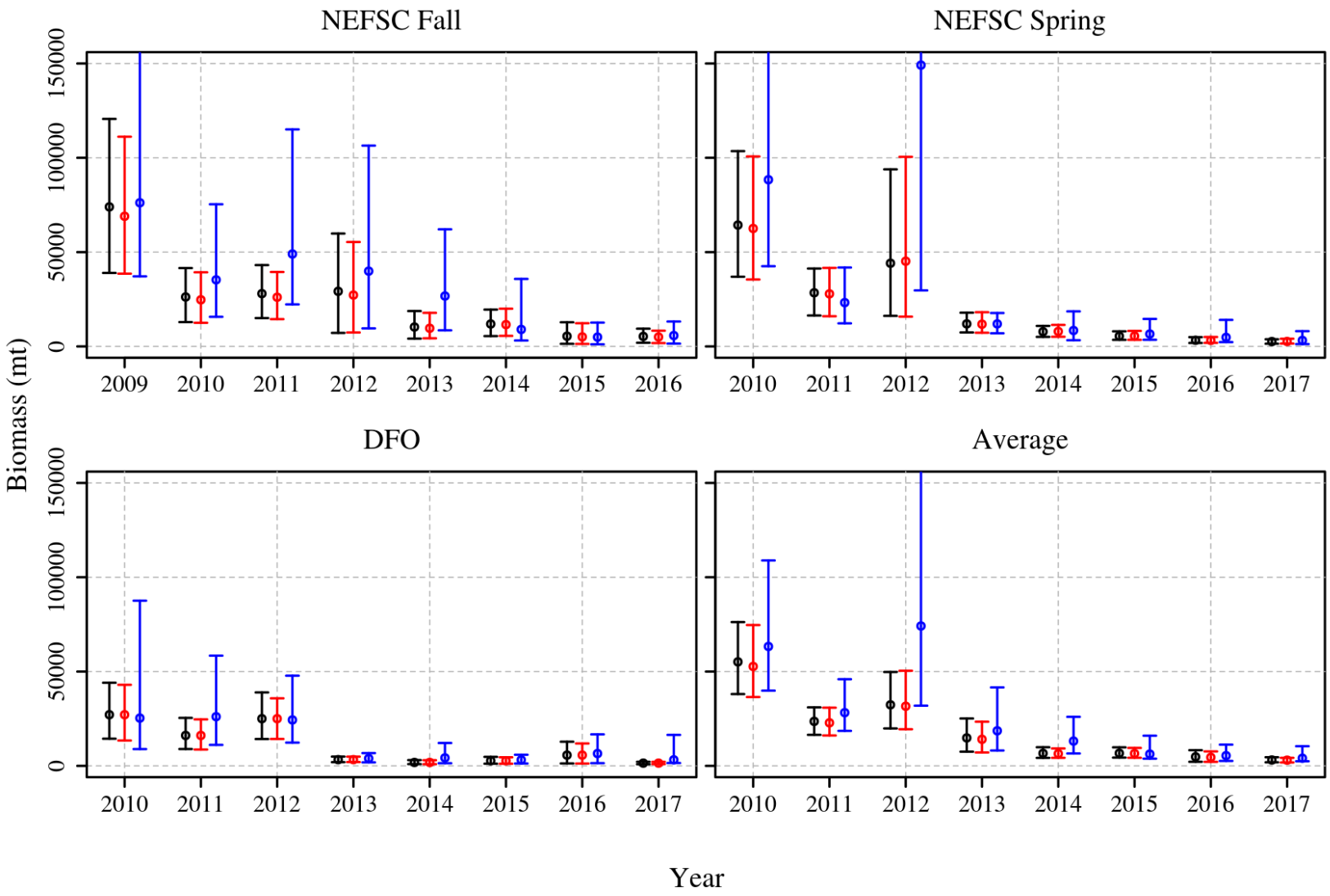

Figure 7. Annual biomass estimates for each survey and for the average across the surveys corresponding to that used in assessing Georges Bank yellowtail flounder. Black, red and blue results represent using three different estimators of abundance at size ($\hat{B}_1$, $\hat{B}_2$, and $\hat{B}_3$, respectively defined in Eqs. 2-6). Vertical bars represent 95% confidence intervals from quantiles of the bootstrap estimates.

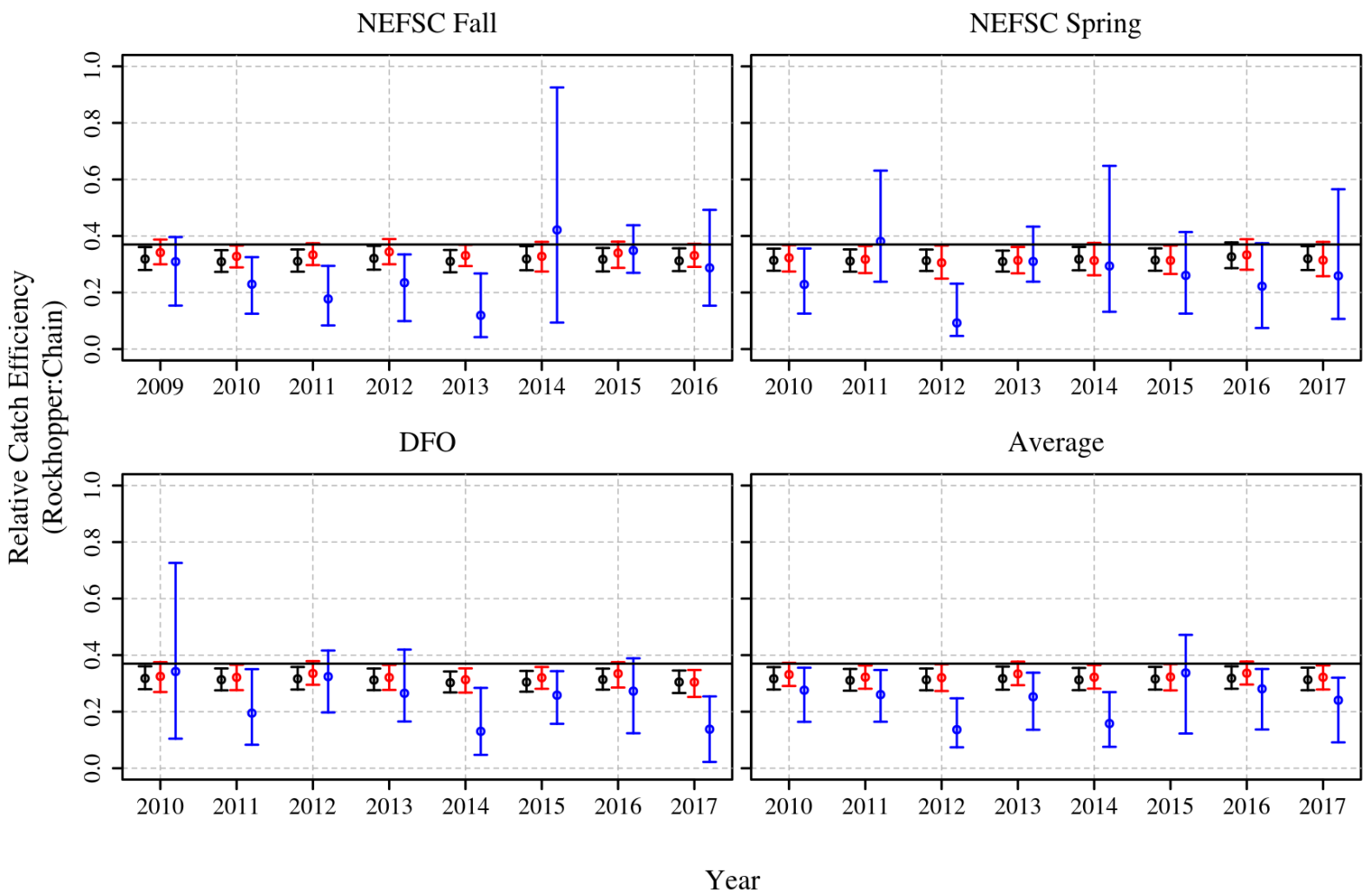

Figure 8. Relative efficiency of biomass estimates annually and by survey and for the average across the surveys corresponding to that used in assessing Georges Bank yellowtail flounder. Black, red and blue results represent using three different estimators of abundance at size $\hat{B}_1$, $\hat{B}_2$, *and* $\hat{B}_3$, *respectively defined in Eqs. 3-7). Vertical bars represent 95% confidence intervals from quantiles of the bootstrap estimates. Horizontal black line represents the 0.37 efficiency currently used in assessing the stock.*

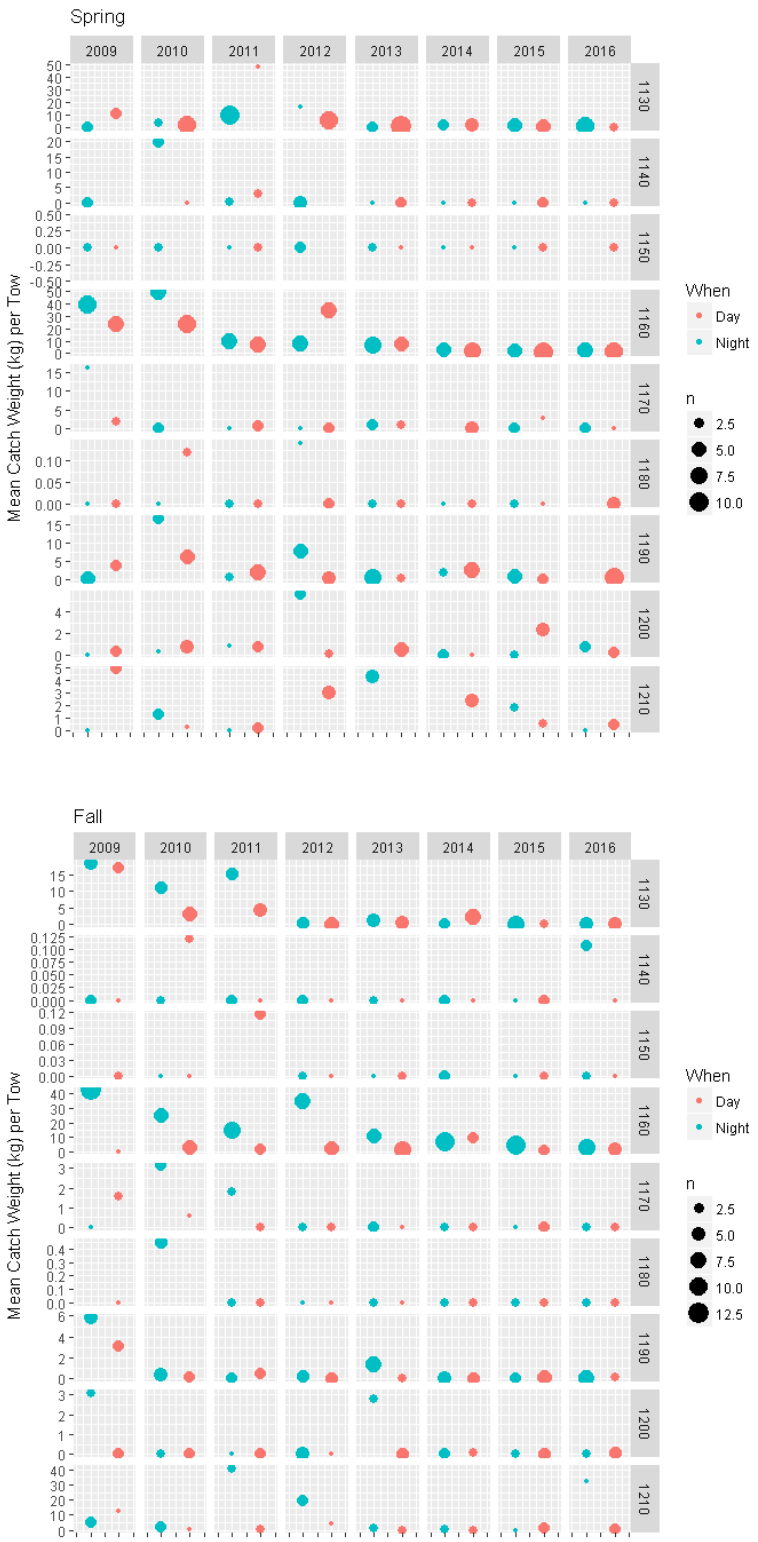

Figure 9. Mean catch per tow (kg) of yellowtail flounder by year and stratum and by day versus night. The color of the symbol indicates day versus night while the size of the symbol indicates the number of tows in that cell.

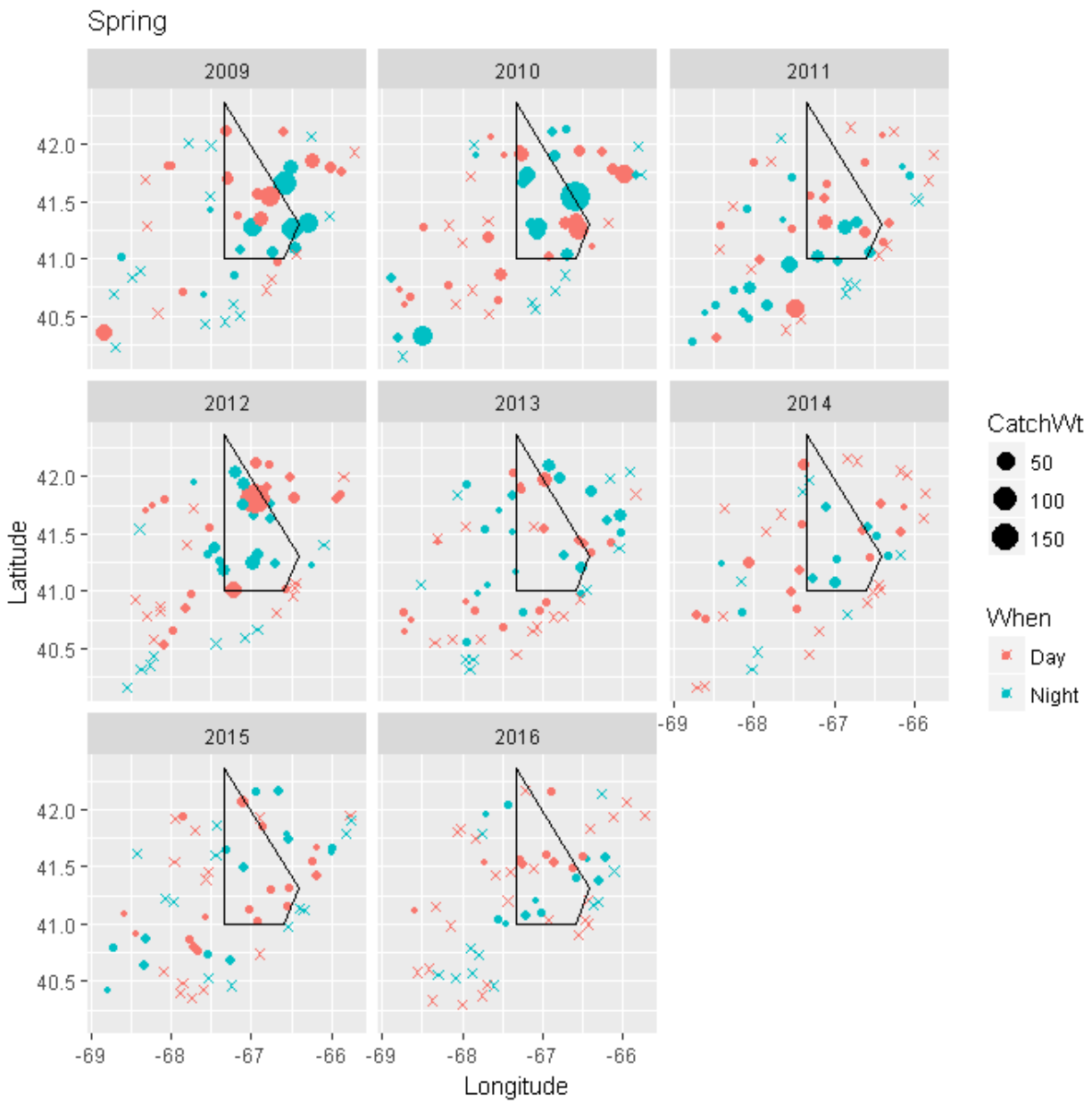

Figure 10a. Catch per tow (kg) of yellowtail flounder during the NEFSC spring survey by year showing location of tows and whether the tow occurred during the day or night. Size of the bubble is proportional to the weight of yellowtail flounder caught, an "x" indicates no catch of yellowtail flounder in that tow. The polygon indicates the location of Closed Area II.

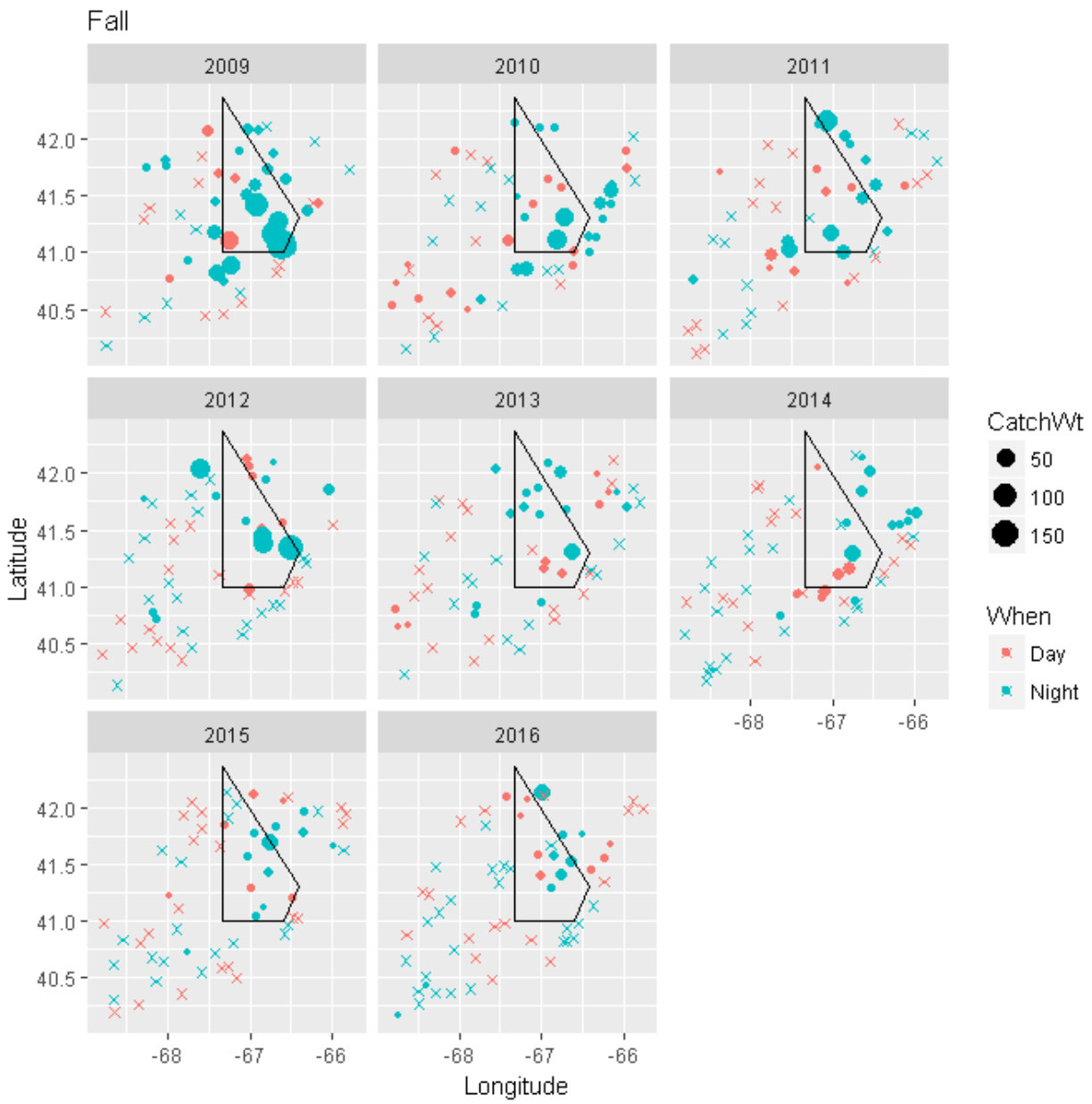

Figure 10b. Catch per tow (kg) of yellowtail flounder during the NEFSC fall survey by year showing location of tows and whether the tow occurred during the day or night. Size of the bubble is proportional to the weight of yellowtail flounder caught, an "x" indicates no catch of yellowtail flounder in that tow. The polygon indicates the location of Closed Area II.

APPENDIX A: ANALYSIS OF PAIRED TRAWL EXPERIMENTS USING ROCKHOPPER AND COOKIE SWEEP GEAR

Summary

We performed analyses of paired tow data collected in 2009 and 2010 that used rockhopper and cookie sweep gear. The relative efficiency of these two gears is estimated using the same models as Miller (2013) like the main working paper. However, we did not investigate diel effects on the relative catch efficiency of these gears. The best model estimated a mean relative efficiency across paired observations essentially without size effects, but random size effects were important at the pair level. The mean relative efficiency matched our expectation that the efficiency of the cookie sweep is intermediate to that of the chain sweep and the rockhopper sweeps. Across most of the well-observed sizes of yellowtail flounder, the relative efficiency of the cookie sweep to the rockhopper sweep is generally less than that of the chain sweep to rockhopper estimated in the main document.

Methods

The paired trawl data were collected during 2009 and 2010 by 3 vessels (FV Endurance, FV Moragh K, and FV Mary K). The study consisted in 6 10-day efforts in 3 regions (Southern New England, Georges Bank, and Gulf of Maine), where 2 vessels fished side-by-side using the same net (standard three-bridle, four-seam trawl used on NEFSC spring and fall bottom trawl surveys) but with different sweeps (standard rockhopper and cookie sweep with 3” discs). The study completed over 430 paired tows of which 242 were usable tows that caught yellowtail flounder in at least one of the paired tows:

Southern New England		Georges Bank		Gulf of Maine	
Spring	Fall	Spring	Fall	Spring	Fall
2010	2009	2010	2010	2010	2010
35	42	58	47	34	26

We fit the same set of conditional binomial and beta-binomial models described in Miller (2013) to the paired tow made with the cookie sweep and rockhopper sweep gears. These models make various assumptions about random effects across pairs and smooth size effects on the relative efficiency and beta-binomial dispersion parameters. We compared the relative performance of these models with AIC.

Results

The best model () was a beta-binomial model that allowed random smooth effects of size for each paired tow observation (Table A1). Notably, the best fit for this model results in a mean relative efficiency without size effects, but where the smoother for each paired observation was important. The model with the same mean assumptions but that assumes that the pairs are conditionally binomial distributed performed virtually identically because the estimated dispersion parameter for the beta-binomial model is extremely large where the beta-binomial and binomial variances are approximately equal. The estimated mean relative efficiency of the cookie sweep (approximately 2) is less than the relative efficiency of the chain sweep (approximately 3-4) over most of the well sampled sizes of fish.

Table A1. Difference in AIC for each model from that with the lowest AIC. See Miller (2013) for descriptions of each model.

	Model	n_p	ΔAIC
All data	BI_0	1	1345.7
	BI_1	2	209.0
	BI_2	3	1319.7
	BI_3	4	206.6
	BI_4	7	0.2
	BB_0	2	672.7
	BB_1	3	161.2
	BB_2	4	673.5
	BB_3	6	642.6
	BB_4	5	161.4
	BB_5	7	110.2
	BB_6	8	0.0
	BB_7	10	2.1

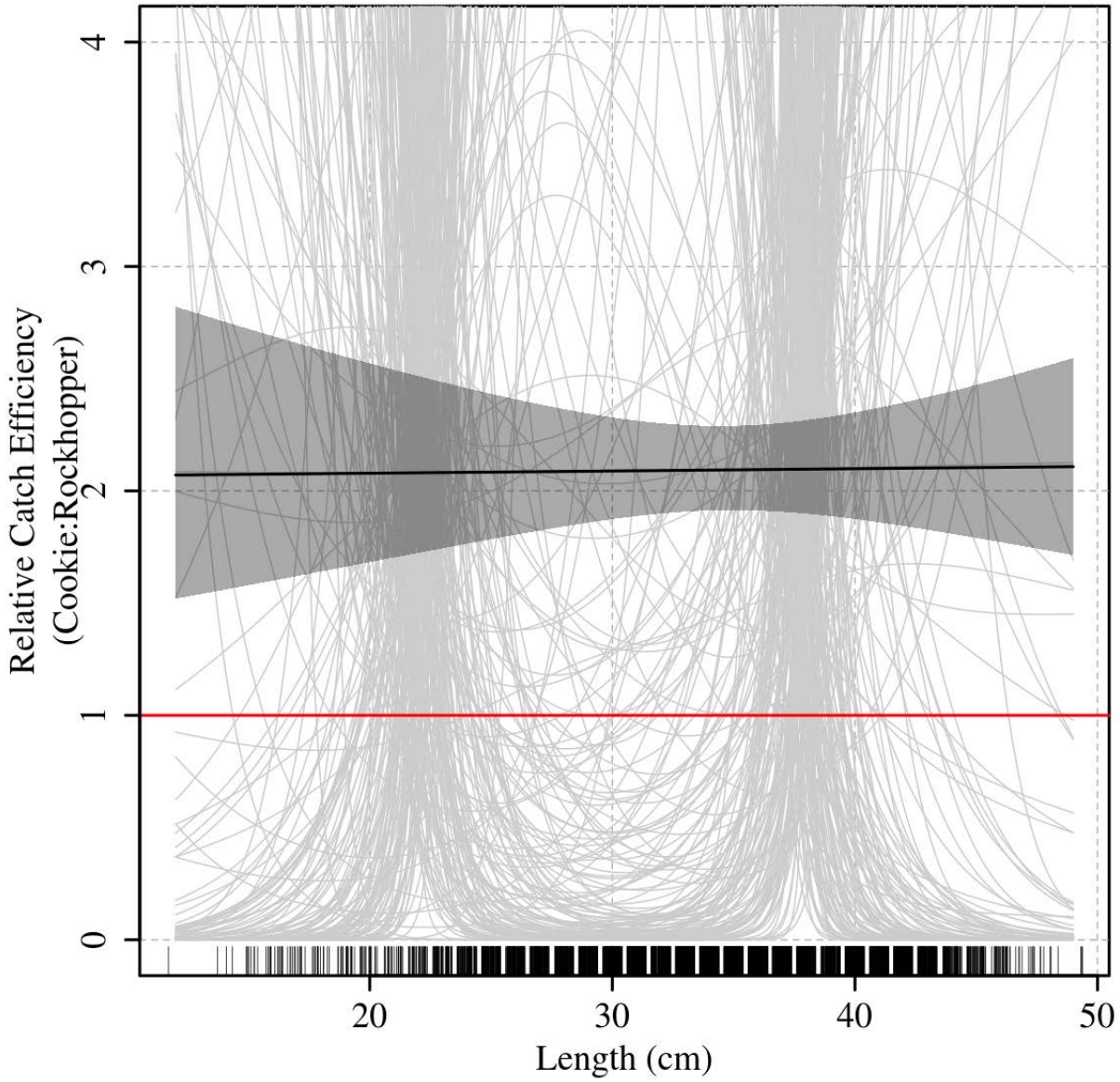

Figure A1. Relative cookie sweep to rockhopper sweep efficiency at size for Georges Bank yellowtail flounder from the best performing model (). Light gray lines are pair-specific estimates of relative efficiency, red line indicates equal efficiency, gray polygon represent 95% confidence interval for the mean relative efficiency (black line) and rug indicates observations at length.