

## Original Article

# The influence of trawl efficiency assumptions on survey-based population metrics

Micah J. Dean  <sup>1,\*</sup>, William S. Hoffman <sup>1</sup>, Nicholas C. Buchan <sup>1</sup>, Steven X. Cadrian <sup>2</sup>, and Jonathan H. Grabowski  <sup>3</sup>

<sup>1</sup>Massachusetts Division of Marine Fisheries, Annisquam River Marine Fisheries Station, Gloucester, MA 01930, USA

<sup>2</sup>School for Marine Science and Technology, University of Massachusetts Dartmouth, 836 South Rodney French Blvd., New Bedford, MA 02744, USA

<sup>3</sup>Marine Science Center, Northeastern University, Nahant, MA 01908, USA

\*Corresponding author: tel: (+978) 282-0308 × 104; e-mail: [micah.dean@mass.gov](mailto:micah.dean@mass.gov)

Dean, M. J., Hoffman, W. S., Buchan, N. C., Cadrian, S. X., and Grabowski, J. H. The influence of trawl efficiency assumptions on survey-based population metrics. – ICES Journal of Marine Science, 0: 1–17.

Received 3 June 2021; revised 22 July 2021; accepted 10 August 2021.

Catch data from bottom trawl surveys are used in various ways (e.g. stock assessments, fisheries management, and ecosystem studies) to represent trends in fish populations across space, time, season, or size. Relative abundance indices assume constant capture efficiency, and area-swept abundance requires an estimate of capture efficiency. Therefore, it is important to develop a predictive understanding of the interaction between fish and survey gear. We conducted experiments to test two primary factors that influence the efficiency of survey trawls at capturing demersal groundfish: (1) footrope escapement—estimated by attaching a collection bag beneath the primary trawl, and (2) herding of the sweeps/doors—estimated by varying sweep length. Random forest models were used to disentangle the herding effect from patterns caused by environmental variables. Contrary to common assumptions, footrope efficiency was incomplete (< 100%) and herding was non-trivial (> 0%), which introduces a bias in population metrics that rely on such assumptions. This bias varied by species and depended upon the relative strength of the counteracting effects of footrope escapement and herding. Our findings suggest that trawl efficiency should be estimated (not assumed) to derive area-swept abundance, and relative abundance indices should account for size-based efficiency and changing size compositions.

**Keywords:** area swept biomass, Atlantic cod, Gulf of Maine, random forest, trawl efficiency

## Introduction

Catch data from bottom trawl surveys are used in a variety of ways to inform fish population and ecosystem dynamics, stock assessments, and fisheries management (Gunderson, 1993; Helser and Hayes, 1995; Pennington and Strømme, 1998). Nearly all applications of these data rely on the catch of a species being representative of its population trends over years (e.g. Collie *et al.*, 2008), seasons (e.g. Gibson *et al.*, 1993), areas (e.g. Swartzman *et al.*, 1992), or fish sizes (e.g. Bianchi *et al.*, 2000). In general, these trends are typically described using a relative abundance index, including to inform analytical stock assessment models (e.g. catch-at-age models). However, there are many instances where trawl survey data are used to estimate absolute population size and trends (e.g. an area-

swept biomass estimate—Gunderson, 1993). Relative abundance indices rely on an assumption of constant capture efficiency, and area-swept abundance requires an assumption or estimate of capture efficiency. Therefore, it is vital to account for the efficiency of the survey trawl at capturing a given species—because how fish interact with the survey gear affects the relationship between observed catches and local fish density. Simplistic assumptions are often made about survey trawl efficiency, yet are rarely empirically evaluated. If these assumptions do not match reality, they can introduce significant biases into the analyses of trawl survey catch data.

Describing how fish species interact with trawl gear and whether systematic bias exists in catch data first requires an understanding of the specific components of a trawl system (He, 2010; Hayes *et al.*, 2013), which include: (1) The “doors,” which provide the

horizontal spreading force to maintain consistent net geometry. These come in a variety of styles, but many are designed to drag along the seabed and create long trailing clouds of suspended sediment; (2) The “sweeps,” or ground cables that connect the doors to the net. Although there are numerous configurations, we refer to the entire length of cable closest to the seafloor between the doors and net, including sections commonly referred to as lower “bridles” or “legs”; (3) The net itself, which is essentially a large mesh bag that strains fish from water; (4) At the bottom of the net, there is a weighted footrope covered in some arrangement of rubber discs or bobbins; and (5) At the top of the net, a headrope is equipped with floats to maintain vertical spread. This simplification of the anatomy of a bottom trawl omits many features and sub-components that distinguish specific styles, yet it encompasses the primary components with which fish interact and that impact the trawl’s efficiency (Somerton *et al.*, 1999; Ryer, 2008).

Accounting for trawl efficiency requires an understanding of what fraction of fish within the path of this entire gear system end up in the codend (i.e. capture probability—Dickson, 1993a). The simplest approach is to assume that all fish within the net path are captured, so that the catch-per-unit-area swept by the net is a direct measure of the local density of the population. However, this assumption ignores that the trawl doors and sweeps “herd” some of the fish beyond the wings of the net into its path, increasing the effective area swept (Ramm and Xiao, 1995). Another approach is to assume that all fish between the doors are captured by the net. Both assumptions ignore the escapement of some fish beneath the trawl footrope, above the headrope, or through the meshes of the net, decreasing its efficiency (Weinberg *et al.*, 2002).

Several studies have attempted to account for the efficiency of each component of the trawl gear by analyzing field trials with mathematical models of varying complexity that describe the relationship between observed catches and local density (e.g. Dickson, 1993a; Ramm and Xiao, 1995; Somerton *et al.*, 2007). The approach of Somerton *et al.* (2007) is particularly useful for describing the efficiency of a bottom trawl at capturing demersal fish species (i.e. “groundfish”). Under this model, catch ( $C$ ) relates to fish density ( $D$ ) as a function of the area swept by the net ( $A_n$ ), the area swept between the doors ( $A_d$ ), the probability that fish outside the net path are herded into the net path ( $h$  or herding efficiency), and the probability that fish within the net path do not escape beneath the footrope ( $k$  or footrope efficiency):

$$C_l = D_l k (A_n + h (A_d - A_n)). \quad (1)$$

Both  $k$  and  $h$  can be represented as a function of auxiliary variables (e.g. species, size, light, and so on) to more accurately represent observed patterns in trawl efficiency (Somerton *et al.*, 2007). Although this approach ignores mesh selectivity (escapement of the smallest individuals through the meshes of the net) and escapement above the headrope, these components of trawl efficiency are considered to be relatively minor when using a fine-mesh survey trawl to sample demersal groundfish populations (Somerton *et al.*, 1999).

The areas swept by the net ( $A_n$ ) and between the doors ( $A_d$ ) are typically measured using the vessel’s geographic positioning system (for tow length,  $L$ ) and sensors that record the distance between trawl components: net width ( $W_n$ ) and door width ( $W_d$ ; Kotwicki *et al.*, 2011). In contrast, the two efficiency coefficients,  $k$  and  $h$ , need to be estimated through experiments that examine changes in catch resulting from perturbations to the components of the standard trawl apparatus. Footrope efficiency ( $k$ ) has typically been estimated by attaching one or more collection bags beneath the pri-

mary net, just behind the footrope (e.g. Engås and Godø, 1989a; Munro and Somerton, 2002; Brinkhof *et al.*, 2017). This “underbag” allows researchers to observe the fish that escape beneath the footrope of the net. Although cumbersome to operate and a logistical challenge to simultaneously process the catches from two nets for each experimental tow, the analysis of these data and estimation of  $k$  is straightforward because of the matched pairs experimental design. Both trawl and underbag are operated in unison over the same area ( $A_n$ ) and experience the same fish density ( $D_l$ ) and environmental factors.

Herding efficiency ( $h$ ) has been measured by systematically adding or removing lengths of ground cable to the sweeps, in order to alter the ratio of the width covered by the sweeps ( $W_s = W_d - W_n$ ) to the width of the net ( $W_n$ ; e.g. Engås and Godø, 1989b; Ramm and Xiao, 1995; Somerton and Munro, 2001). If a herding effect is present, the observed catch should increase with the ratio of sweeps width to net width ( $\emptyset = W_s / W_n$ ). Herding effects have also been detected by lifting the sweeps above the seafloor with rubber discs, floating line, or pelagic doors (Rose *et al.*, 2010; He *et al.*, 2015; Sistiaga *et al.*, 2015) or inferred through video observations of fish behavior interacting with trawls (Winger *et al.*, 2004; Ryer *et al.*, 2010; Bryan *et al.*, 2014). Compared to a footrope experiment, each tow in a herding experiment is relatively simple to conduct, because it is similar to a standard survey operation. From an analytical perspective, estimating  $h$  is far more difficult because each treatment samples a different area, fish density ( $D_l$ ), and environmental conditions. Previous studies have applied randomized block experimental designs, estimating  $h$  by assuming equal  $D_l$  within each block (e.g. Somerton and Munro, 2001; Somerton *et al.*, 2007). Although this approach may be valid in some cases (e.g. small blocks with little habitat variability), in many cases  $D_l$  can vary substantially over short distances due to the complex and subtle relationships that most fish species have with their environment (Hedger *et al.*, 2004; Maxwell *et al.*, 2009; Lenoir *et al.*, 2011). This introduces significant natural variation in observed catches from adjacent tows that can prevent the detection of a herding effect. Furthermore, collecting all experimental observations from a narrow subset of the possible habitat conditions encountered by the survey may not represent capture efficiency for the entire survey area (Somerton *et al.*, 1999).

An alternative approach to assuming equal  $D_l$  across blocks of tows is to record auxiliary variables and account for their influence on the observed catch, simultaneously estimating the effect caused by experimental manipulation of the gear (i.e. altering sweep length and thus  $\emptyset$ ). Machine learning approaches, such as random forest models, are particularly adept at describing the distribution of species in the wild because they can account for the many non-linear and interacting influences of these environmental variables (Stockwell and Peterson, 2002; Evans *et al.*, 2011). Random forests gain their predictive accuracy by averaging across many uncorrelated regression trees. Although the accuracy of each individual tree is quite low, the average predictions from a “forest” of these trees (made un-correlated or “random” through the bootstrap and subsampling algorithm) can be highly accurate, account for complex patterns, and are robust to outliers. If we also include  $\emptyset$  (i.e. the ratio of  $W_s$  to  $W_n$ ) as an explanatory variable in a random forest model of the catches from a herding experiment, then it is possible to measure the herding effect by controlling for the effects of the other variables.

Experimentally-derived estimates of trawl efficiency can be applied to a survey dataset to calculate total abundance or spawner

biomass of a stock, which in turn can help evaluate the accuracy of population models from a stock assessment (Somerton *et al.*, 1999). This is especially important for stocks that have assessment models with diagnostic problems and for which the scale of population size is uncertain (Cadrin *et al.*, 2016). In situations where no assessment model is possible, survey-based estimates of population size may be the *only* source of information available to managers for regulating the scale of fishery removals—e.g. winter flounder in the Gulf of Maine (NEFSC, 2011); Georges Bank yellowtail flounder (NEFSC, 2015); witch flounder in the US (NEFSC, 2017); Atlantic cod on Georges Bank (NEFSC, 2019); red hake in the US (NEFSC, 2020a), windowpane flounder in the Gulf of Maine (NEFSC, 2020b). Area-swept survey estimates are also used as a diagnostic confirmation of the scaling for stock assessments (e.g. NEFSC, 2017, 2019). Trawl efficiency is also inherent to other population and community metrics, such as species composition and size distribution, that are frequently used to inform ecosystem studies, multispecies trophic models, and hydroacoustic surveys. In the Gulf of Maine, off the northeast US coast, there are several stocks of demersal “ground-fish” species that could benefit from an independent survey-based perspective on their populations. To this end, the Massachusetts Division of Marine Fisheries conducted a three-year bottom trawl survey (November 2016–January 2019) that intensively sampled the demersal fish community of the Western Gulf of Maine. Known as the Industry-Based Survey because it used a commercial-style net and was operated from a chartered commercial fishing vessel, this program executed a stratified random fishery-independent sampling design that facilitates generalizing these results to estimates of the population (see Dean *et al.*, 2021, 2019 for details on survey design and methodology).

The objectives of this study were to estimate the efficiency of a bottom trawl survey and to evaluate the impact of trawl efficiency assumptions on population metrics. We describe separate experiments to estimate the footrope escapement and herding effect for six species of groundfish in the Gulf of Maine, USA: Atlantic cod (*Gadus morhua*), haddock (*Melanogrammus aeglefinus*), yellowtail flounder (*Pleuronectes ferruginea*), winter flounder (*Pseudopleuronectes americanus*), American plaice (*Hippoglossoides platessoides*), and witch flounder (*Glyptocephalus cynoglossus*). Specifically, we tested the common assumptions of 0% footrope escapement and either 0 or 100% herding efficiency. Through this example, we aim to offer both an incentive and a means for other groups to invest in their survey datasets through similar experiments, thereby enabling an independent perspective on demersal fish populations that is not subject to substantial biases.

## Methods

### Standard survey gear

The survey trawl is based on a configuration commonly used in the Gulf of Maine commercial fishery and was designed collaboratively by a team of commercial fishermen, net builders, and survey scientists (Hoffman *et al.*, 2012). It is a two-seam net with a 27 m footrope, a 26 m headrope, and a 46 m fishing circle. The trawl footrope consists of wire covered in 15 cm rubber discs, interspersed with larger diameter discs at regular intervals: 36 cm diameter, every 30 cm at center, transitioning to 30 cm diameter every 60 cm in the wings. This footrope design is known colloquially as a “rockhopper” and allows the survey to operate in a wide variety of habitats, including complex substrate like cobble. The sweeps are

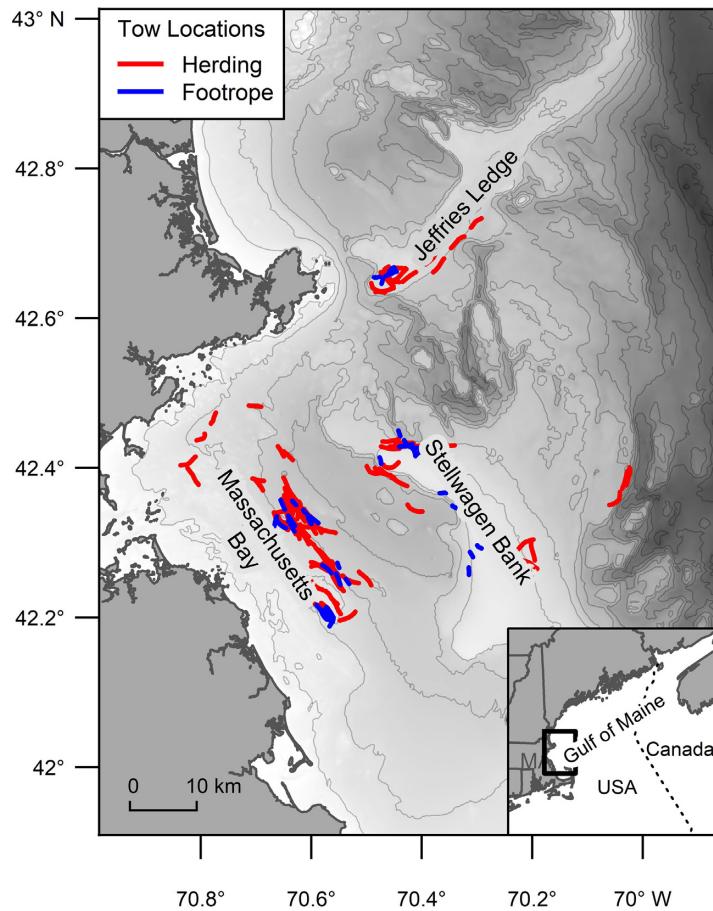
comprised of a 27 m ground cable attached to a 27 m bridle, both of which are covered by 5 cm rubber discs (“cookies”) along their entire length. The sweeps are attached to low-aspect rectangular “Bison” doors that partly utilize seabed contact to provide horizontal spreading force. This combination of net, doors, and sweep serve to maximize seafloor contact and suspension of sediment between the net and doors (Valdemarsen *et al.*, 2007), and is a commonly used design in the US commercial fishery to augment the herding effect for demersal fish (Lomeli *et al.*, 2019). To maximize the retention of a variety of species and sizes, the net was designed with a high headrope height (4.5 m) and a small (5.1 cm) codend mesh size. During the design process, flume trials were conducted on a scale model of the trawl to confirm that it operated as intended. Video observations of the trawl footrope were collected from a variety of seafloor types during standard survey operations to confirm that net maintain consistent bottom contact. In addition, sidescan sonar imagery was collected on several tows to observe all components of the trawl while underway. The protocol for standard survey operation was to tow the net for 30 min at a vessel speed of 1.5 m s<sup>-1</sup> (3.0 knots). Towing time began when the winches were locked, with the net on the seafloor, and vessel moving at towing speed. Towing time ended when the winches were re-engaged to initiate haulback. Vessel speed-over-ground, latitude, longitude, and seafloor depth were measured continuously via the survey vessel’s GPS and depth sounder. Total aggregate weight of each fish species was measured, and individual lengths were recorded for all ground-fish species. All survey tows were conducted during daylight hours aboard the commercial fishing vessel F/V Miss Emily (17 m length overall).

### Net mensuration

The trawl was equipped with mensuration sensors (Notus Inc.) to record the distance between the wings of the net ( $W_n$ ) and the distance between the doors ( $W_d$ ). When operating correctly, mensuration sensors transmitted data approximately once every 30–60 s. These sensors were sometimes unreliable and did not always transmit useable data, with the net sensors failing more frequently than the door sensors. For instance, individual transmitted values occasionally were implausible (e.g.  $W_n$  increasing by +100% for 2 s). To minimize the influence of these unrealistic values, mensuration data were filtered to remove outliers ( $> 1.5 \times$  interquartile range) prior to calculating tow-level median values. As with most trawls,  $W_d$  increased predictably with depth (Weinberg and Kotwicki, 2008). Therefore, a linear model of  $W_d$  as a function of  $\log(depth)$  was fit to the observed median values and used to estimate  $W_d$  for tows without mensuration data. Similarly, a linear model of net width ( $W_n$ ) as a function of  $W_d$  was developed and used to estimate  $W_n$  for tows without recorded mensuration data (ICES, 2019). The recorded  $W_n$  and  $W_d$  values (or expected values when not recorded) were used to calculate the median width of the sweeps ( $W_s$ ), sweep angle, and sweeps-to-net width ratio ( $\emptyset = W_s / W_n$ ) for each tow.

### Footrope efficiency

An underbag with the same codend mesh size was attached to the bolschline of the primary trawl and covered the entire width of the trawl net to observe the escapement of fish beneath the trawl footrope. The footrope of the underbag was constructed of 1.9 cm wire, weighted with 4.5 kg of lead per meter, and covered with



**Figure 1.** Survey tow locations for the herding experiment (red lines) and footrope experiment (blue lines). Water depth is indicated by the shaded background where darker colors indicate deeper water. Gray contour lines are shown at 20 m intervals.

5 cm diameter rubber disks. In addition, the underbag footrope had 13 cm diameter rubber disks spaced every 5 cm at center, tapering to 5 cm diameter in the wings. As such, there was a pattern of 4 cm  $\times$  5 cm gaps beneath the rubber discs at center (smaller toward the wings), making escapement unlikely for fish that would be retained by the 5 cm mesh of the cod end. The underbag footrope was also 1.5 m longer than that of the primary trawl, which allowed the underbag to trail a short distance behind the trawl footrope. This design represents a compromise between being able to tow over a variety of substrates vs. minimizing escapement beneath the underbag footrope.

Underbag tows were conducted in Massachusetts Bay, Stellwagen Bank, and Jeffries Ledge between May 2016 and June 2019 (Figure 1). These areas were chosen because previous survey efforts identified these as locations that are inhabited by a broad size distribution of several groundfish species, representing the populations of inference for the entire survey. In addition, our aim was to conduct the experiment in as broad and representative range of habitats as possible. However, given the experience of other groups that have conducted underbag studies to examine footrope escape, gear damage was also a concern (Ingólfsson and Jørgensen, 2006; Brinkhof *et al.*, 2017). Therefore, to minimize underbag damage and maximize the number of observations, initial tows were short (10 minutes) and occurred over relatively fine-grained sediment (i.e. sand and silt), but were later lengthened (20 min) and included more complex substrates (i.e. cobble and rock).

Footrope efficiency ( $k$ ) typically varies with fish size and was therefore calculated as the fraction of total catch ( $C$ ) within each 1 cm length bin ( $l$ ) that entered the trawl net:

$$k_l = \frac{C_{trawl,l}}{C_{trawl,l} + C_{underbag,l}}. \quad (2)$$

Similar to Somerton *et al.* (2007) and Munro and Somerton (2002), several parametric models were fit to pooled data via maximum likelihood methods using a binomial likelihood function (Table 1). The candidate model set included a variety of potential functional relationships, including no length effect (fixed), symmetrically asymptotic (logistic), asymmetrically asymptotic (Richards, log-logistic), and unimodal (double-logistic). The best fitting model was selected using AIC and confidence intervals were estimated by bootstrap re-sampling of the raw data by tow (1000 iterations).

### Herding efficiency

A separate set of experimental tows were conducted to address the herding efficiency of the trawl doors and sweeps. Lengths of cookie-covered ground cable were either added to or removed from the standard trawl sweeps in order to achieve three treatments: (1) standard: 27 m bridle + 27 m cable; (2) short: 27 m bridle + no cable; and, (3) long: 27 m bridle + 55 m cable. Each tow was conducted according to the standard survey protocols (target = 30 min at  $1.5 \text{ m s}^{-1}$ ). The order of each cable variant was randomized, and the

**Table 1.** Candidate models considered in the footrope efficiency experiment. The AIC for each model is provided, with the lowest AIC for each species is listed in bold.

Model	AIC					
	Atlantic cod	Haddock	Yellowtail flounder	Winter flounder	American plaice	Witch flounder
Fixed $\frac{1}{1+e^{-\alpha}}$	1035.4	219.6	4838.8	2171.2	1013.6	586.9
Logistic $\frac{\gamma}{1+e^{-(\alpha+\beta \cdot \text{len})}}$	690.9	<b>197.5</b>	4571.1	<b>1519.7</b>	<b>908.1</b>	590.2
Log-Logistic $\frac{\gamma}{1+(\text{len}/a)^{-\beta}}$	699.0	198.9	4577.6	1523.3	909	<b>583.3</b>
Richards $\gamma \left( \frac{1}{1+e^{-\beta(l-\alpha)}} \right)^{\frac{1}{\delta}}$	<b>685.1</b>	199.4	<b>4564.1</b>	1521.6	910.1	592.2
Double Logistic $\frac{\gamma}{(1+e^{-\beta(l-\alpha)})(1+e^{-\theta(l-\delta)})}$	690.3	227.6	4575.0	1526.2	911.2	587.5

captain was instructed to avoid overlapping the area swept by each tow on a given date.

This experiment was originally designed to estimate the herding coefficient ( $h$ ) via maximum likelihood methods by conditioning on the total catch per block, similar to the approach of Somerton and Munro (2001). However, habitat heterogeneity in the study area was too great to assume equal fish density within each block, and it was not possible to consistently complete a block of three tows at a similar depth or substrate type. Preliminary analyses indicated that there were complex yet persistent spatial patterns to the catch rate of each species. Therefore, we applied a random forest regression model (Breiman, 2001; Liaw and Wiener, 2002) to determine the influence of the sweeps-to-net width ratio ( $\emptyset$ ) on the catch ( $C$ ) of each species, while accounting for the influence of several other independent variables, including: longitude, latitude, depth, seafloor complexity, 10-minute “square” of latitude and longitude (as a factor), and tow date (as a factor). Latitude, longitude, and depth were calculated as the mean values per tow from the vessel’s GPS and depth sounder. Seafloor complexity was measured using a raster grid (100 m resolution) where each cell represented the average absolute difference in depth between itself and its surrounding eight neighbors (Wilson *et al.*, 2007).

The partial dependence of catch on each of  $j$  independent variables (i.e.  $C \sim F(j)$ ) was estimated from the species-specific random forest models and evaluated along a grid of 20 equally spaced values (Friedman, 2001). If a herding effect exists for a given species, there should be a positive linear relationship between expected catch ( $\hat{C}$ ) and  $\emptyset$ . As such, a simple linear model was fit to the partial dependence of catch on  $\emptyset$  (i.e.  $\hat{C} = \alpha\emptyset + \beta$ ), weighted by the number of observations in each  $\emptyset$  “bin.” The intercept of this relationship ( $\beta$ ) represents the average catch due to the net path alone, whereas the slope of this relationship ( $\alpha$ ) represents the average catch from just the sweeps path, when the width of the sweeps is equal to the net width (i.e.  $\emptyset = 1$ ). Thus, dividing the slope by the intercept yields an estimate of the herding coefficient:

$$h = \frac{\alpha}{\beta} = \frac{(\hat{C}|W_s = W_n) - (\hat{C}|W_s = 0)}{(\hat{C}|W_s = 0)}. \quad (3)$$

A value of  $h = 1$  indicates 100% herding efficiency (i.e. all fish in the sweeps path moved into the net path) and a value of  $h = 0$  indicates 0% herding efficiency (i.e. none of the fish in the sweeps path

moved into the net path). For some species, there were departures from the apparent linear relationship between catch and  $F(\emptyset)$  at extreme values of  $\emptyset$ . This is to be expected, given that the estimated partial dependence of a variable at the limits of its observed range is often imprecise (Auret and Aldrich, 2012). Therefore, fitting of the linear model was restricted to between the 5th and 95th quantiles of  $\emptyset$ .

By virtue of their construction, random forest models provide a convenient representation of prediction error (out-of-bag error rate–Zhang *et al.*, 2020). Unfortunately, this measure of model uncertainty does not naturally extend to partial dependence relationships (Cafri and Bailey, 2016), which are a property of the entire forest and not the individual trees. Instead, the original herding experiment dataset was re-sampled by tow with replacement (i.e. bootstrapped) 1000 times. At each iteration, a random forest model was fit to the re-sampled data and a linear model was used to estimate  $h$  from the  $C \sim F(\emptyset)$  partial dependence relationship. The distribution of these outcomes was used to characterize uncertainty in herding efficiency.

In order to evaluate whether the herding effect varied by either fish length or habitat variables (depth, seafloor complexity), the interaction strength between  $\emptyset$  and each of these variables was evaluated. Interaction strength was measured using the  $H$ -statistic proposed by Friedman and Popescu (2008), which can be interpreted as the % of variance in the 2-way partial dependence function  $F(\emptyset, j)$  that cannot be explained by the sum of the individual partial dependence functions:  $F(\emptyset) + F(j)$ . Therefore, a low  $H$ -value (i.e. close to 0) indicates predominantly additive effects and little to no interaction between  $\emptyset$  and  $j$ ; conversely, a high  $H$ -value (i.e. close to 1) indicates a strong  $\emptyset : j$  interaction, and that the degree of herding varied as a function of variable  $j$ . However, even random sample variation of independent variables often results in some level of non-additivity ( $H > 0$ ; Lampa *et al.*, 2014). Therefore, the observed  $H$ -values were compared to a null distribution of  $H$ -values ( $H^0$ ) assembled via the parametric bootstrap method suggested by Friedman and Popescu (2008; see section 8.3), using 100 iterations. If the observed  $H$ -value was outside the range of  $H^0$ , the interaction was interpreted as significant. Random forest models were built using the R package *randomForest* with 2000 trees and 2 randomly-selected variables at each node (i.e. *mtry* = 2); partial dependence plots and variable interactions were created using the R package *iml*.

## Impact of efficiency on population metrics

Many applications of trawl survey data are derived from a measure of mean fish density, either in terms of abundance or biomass, including estimates of total population size, survey catchability ( $q$ ), size composition, and species composition. To illustrate the impact of trawl efficiency on these population metrics, estimates of  $h$  and  $k$  were used to calculate mean abundance (number  $\text{km}^{-2}$ ) and mean spawner biomass ( $\text{kg km}^{-2}$ ) for each species from the broader survey dataset, which is comprised of 1001 standard configuration tows conducted throughout the Western Gulf of Maine from 2016–2019 (see Dean *et al.*, 2019, 2021 for additional details). Mean density estimates were then compared to calculations using the common assumptions of no footrope escapement and either no herding ( $k = 1, h = 0$ ) or complete herding ( $k = 1, h = 1$ ). The survey did not routinely collect individual weights or maturity observations for all groundfish species. Therefore, in order to calculate mean spawner biomass, weight-at-length and maturity-at-length functions were derived from data collected in the Gulf of Maine by the Northeast Fisheries Science Center (NEFSC) bottom trawl surveys over the past decade (2010–2019). Biological samples collected from the NEFSC Spring and Fall survey cruises from within the Gulf of Maine (NOAA statistical areas 511–515) were used to fit a logistic regression of the proportion mature as a function of length, and individual weight as an exponential function of length (i.e.  $W = aL^b$ ).

## Results

### Net mensuration

Valid door width ( $W_d$ ) measurements were collected from 86% of the experimental tows ( $n = 123$ ), including both footrope and herding experiments. Door width data from an additional 621 standard tows from the broader survey were combined with these data to fit the model of door width as a function of depth and treatment ( $W_d \sim \log(\text{depth}) : \text{treatment}$ ;  $R^2 = 0.80$ ). Increasing sweep length caused the door width to increase more as a function of depth in comparison to the standard configuration (Supplementary Figure S1); likewise, decreasing sweep length caused the door width to increase less as a function of depth, as compared to the standard configuration. Recorded door width from underbag tows (footrope efficiency experiment) were slightly narrower at depth than standard survey tows. The predictive relationship of  $W_d$  and depth was used to impute an expected  $W_d$  value for the 14% of experimental tows that were missing door width data.

Of the herding experiment tows, 35% ( $n = 33$ ) had valid measurements of both  $W_n$  and  $W_d$ ; these were combined with data from an additional 67 tows with  $W_n$  and  $W_d$  data from the broader survey dataset and used to fit a model to predict net width for tows without recorded  $W_n$  data ( $W_n \sim W_d : \text{treatment}$ ;  $R^2 = 0.93$ ). The short sweep treatment caused  $W_n$  to increase more as a function of  $W_d$ , while the longer sweep treatment caused  $W_n$  to increase less as a function of  $W_d$  (Figure S1). Unfortunately, no net width measurements were collected during the footrope experiment, but sidescan imagery of underbag tows suggest that the net was operating normally (Figure 2).

On average,  $W_d$  was 10% wider with the longer sweeps and 19% narrower with the shorter sweeps, as compared to the standard configuration (Figure 2). Conversely, the longer sweeps yielded an 8% narrower  $W_n$ , on average, while the shorter sweeps achieved a 5% wider  $W_n$ . The sweeps-to-net ratio ( $\emptyset$ ) was 27% higher with the

long sweeps and 32% lower with the short sweeps, as compared to the standard configuration. Combining both trawl geometry relationships revealed that  $\emptyset$  increased with depth, essentially mirroring the shape of the  $W_d \sim \log(\text{depth})$  curve, and ranged from 1.8 at a depth of 20 m to 2.6 at a depth of 200 m, under the standard configuration. The sweep angle decreased by 3° with longer sweeps and increased by 6° with shorter sweeps, as compared to the standard configuration.

### Footrope efficiency

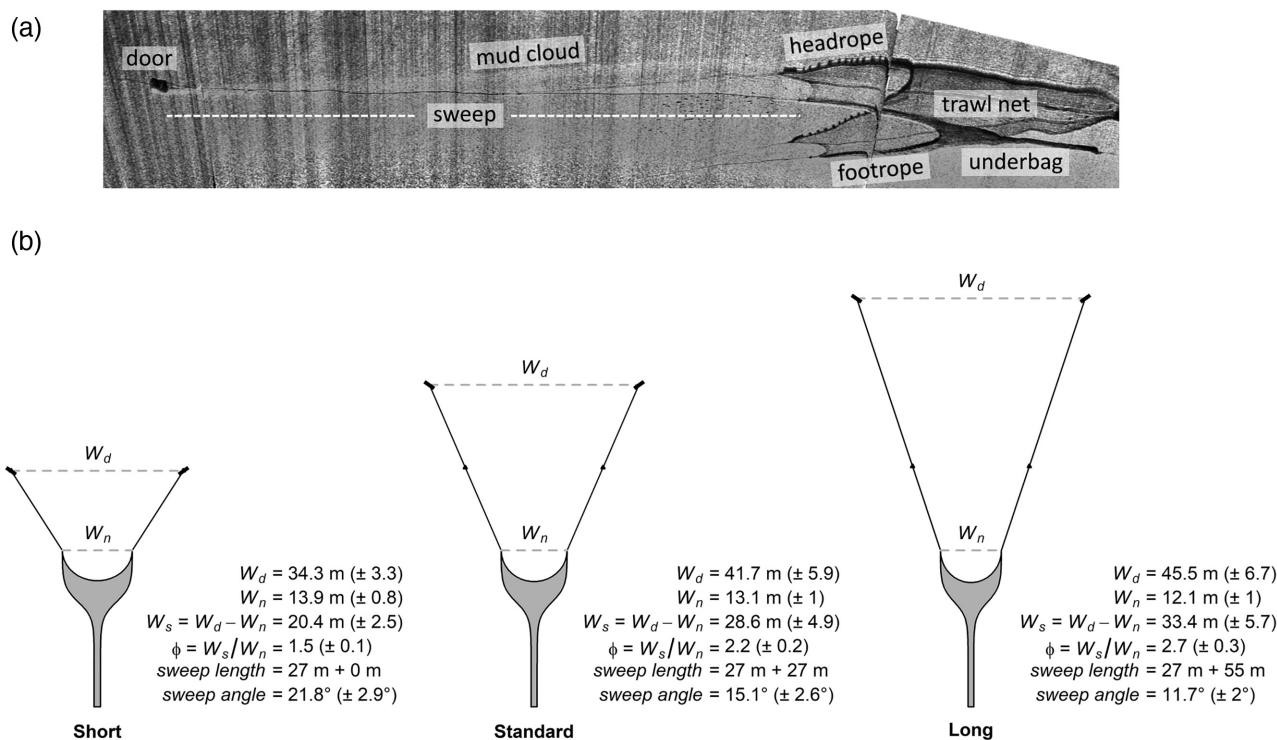
A total of 50 tows were conducted as part of the footrope efficiency experiment (Table 2 and Figure 1). Seafloor complexity and tow speed were similar between these experimental tows and the primary survey ( $t$ -tests:  $P$ -value = 0.954,  $df = 981$ ;  $P$ -value = 0.218;  $df = 1048$ ; respectively), but experimental tows were shorter on average and conducted at shallower depths ( $t$ -test:  $P$ -value < 0.0001;  $df = 1049$ ). These differences resulted from intentional choices to minimize underbag damage and to avoid deeper areas with generally low groundfish abundance. A total of 22 308 fish were caught from the six focal species, 40% of which were captured by the underbag. The portion of total catch coming from the underbag varied by species and ranged from 0.5% (haddock) to 69% (witch flounder). The mean size of fish caught in the underbag was smaller than in the trawl net for all species except for witch flounder (Table 2).

The asymmetric Richards curve best described the footrope efficiency for Atlantic cod, with efficiency rising from 0.5 at 28 cm to an asymptote of 0.96 at 50 cm (Figure 3 and Table 3). Yellowtail flounder was also best described by Richards curve, but with footrope efficiency approaching an asymptote of 0.44 at 37 cm. A logistic curve best described the footrope efficiency for haddock, although there was little difference between the efficiency of the smallest ( $k_{16cm} = 0.87$ ) and largest individuals ( $k_{58cm} = 1.00$ ). Winter flounder and American plaice were also best described by logistic models, with footrope efficiency approaching a maximum asymptote of 0.87 and 0.72, respectively. Witch flounder was the only species for which footrope efficiency decreased with length, best described by the asymmetric log-logistic curve.

### Herding efficiency

A total of 93 tows were conducted as part of the herding experiment, split equally across the three treatments, between 25 May and 8 June of 2018 ( $n = 46$ ) and 2019 ( $n = 46$ ; Table 2). These tows were conducted over a range of depths (range: 30–122 m) and habitats in Massachusetts Bay, Stellwagen Bank, and Jeffries Ledge (Figure 1). A total of 51 225 fish were caught from the six focal species (Table 2). The short sweep treatment captured the lowest number of individuals for all species, while the long sweep treatment captured the greatest number for all except for winter flounder and witch flounder.

The random forest model of catch-at-length for each species revealed that between 83 and 96% of the variance in the  $F(\emptyset, L)$  partial dependence function could be explained by the additive effects of the individual partial dependence functions  $F(\emptyset)$  and  $F(L)$  (i.e. interaction strength,  $H = 0.04$  to 0.17). All of these  $H$ -values were within the null distribution of interaction strength ( $H^0$ ). Therefore, the herding effect was assumed to be unaffected by the size of fish and all further analysis was conducted on length-aggregated data (i.e. total catch by species). The interaction strength between  $\emptyset$  and habitat variables (depth and seafloor complexity) in the length-



**Figure 2.** [Top] Sidescan sonar image of the survey trawl underway during the footrope efficiency experiment. A school of fish being herded can be seen directly in front of the net. [Bottom] Schematic of trawl geometry during the herding experiment. Mean values are reported for each treatment, with the standard deviation shown in parentheses.  $W_d$  = door width;  $W_n$  = net width;  $W_s$  = sweeps width;  $\phi$  = sweeps-to-net width ratio; sweep length = distance from net to door; sweep angle = orientation of sweep in degrees of deviation from trawl path.

aggregated analysis was also low ( $H < 0.12$  in all cases) and was within the null distribution for each species, suggesting that herding was unaffected by these variables as well.

The random forest models explained between 30% (cod) and 54% (American plaice) of the observed variation in catches. Partial dependence plots of habitat and spatio-temporal variables revealed that cod and winter flounder were generally more abundant in the shallower and more complex habitat of Massachusetts Bay and Jeffries Ledge (Figure 4). Conversely, haddock, witch flounder, and American plaice were more abundant in the deeper and less complex habitat between Massachusetts Bay and Stellwagen Bank. Yellowtail flounder were caught in relatively equal abundance in all but the shallowest areas.

The partial dependence plots of the sweeps-to-net-width ratio ( $\phi$ ) revealed a linear increase for each species, suggesting the presence of a herding response (Figure 4). The rate of this increase ( $\alpha$ ) with respect to its intercept ( $\beta$ ) and thus the herding coefficient ( $h$ ) varied by species, from a low of 0.07 for winter flounder to a high of 0.55 for yellowtail flounder (Table 1 and Figure 5). In general, uncertainty was positively correlated with the estimate of herding efficiency (i.e. higher efficiency, higher uncertainty—Figure 5). However, species for which the random forest model explained a larger amount of the variation in catch (winter flounder:  $R^2 = 0.51$ ; plaice:  $R^2 = 0.54$ ) had relatively low uncertainty in  $h$ . Likewise, species that had less variance explained by the random forest model (yellowtail flounder:  $R^2 = 0.34$ ; cod:  $R^2 = 0.30$ ) had relatively high uncertainty in  $h$ . In all cases, there was a high probability that  $h > 0$  (82–100%) and that  $h < 1$  (98–100%).

### Impact of efficiency assumptions on population metrics

When compared to the mean density calculated from estimates of footrope and herding efficiency, an assumption of no herding and 100% efficiency within the net path (i.e.  $k = 1$ ,  $h = 0$ ) resulted in a positive bias for both gadids (+7 to +53%) and a negative bias (−14 to −37%) for all four flatfishes (Figure 6). A greater bias (−54 to −81%) occurred for all six species when mean density was estimated assuming 100% efficiency between the doors (i.e.  $k = 1$ ,  $h = 1$ ). A similar pattern was found when mean abundance was translated to mean spawner biomass using functions of maturity-at-length and weight-at-length (Supplementary Table S1), but bias for spawner biomass was shifted toward the positive direction for all species except witch flounder, which shifted in the opposite direction (Figure 6). In most cases, bias in mean density and spawner biomass was similar in each of the three years covered by the primary survey dataset (2016–2018) and within +/- 4% of the overall bias for that species across the entire time series. The only exception was for annual levels of bias in the mean abundance of Atlantic cod (2016: +18%; 2017: +13%; 2018: −9%).

The species composition without correcting for species-specific differences in efficiency exaggerated the importance of haddock relative to other species (+16%) and diminished the relative importance of the four flatfish species (−2 to −8% individually; −16% collectively; Figure 7). As such, the portion of the catch attributed to haddock was approximately five times greater than the next most abundant species, American plaice. However, this haddock: plaice ratio decreased to only two when species-specific efficiencies were implemented. Bias in mean length was highest for

**Table 2.** Summary of tows and catches from footrope experiment, herding experiment, and the primary survey.

<b>Tow summary</b>		<b>Herding experiment</b>			<b>Primary survey</b>	
<b>Footrope experiment</b>		<b>Short</b>	<b>Standard</b>	<b>Long</b>		
n tows	50	31	31	31	1001	
mean depth (m)	50.1	59.1	59.2	58.4	109.8	
mean ground speed (m/s)	1.46	1.50	1.51	1.50	1.48	
mean tow length (km)	1.22	2.57	2.61	2.62	2.52	
mean seafloor complexity (m)	22.3	17.3	21.0	20.8	22.2	
<b>Number caught</b>		<b>Herding experiment</b>				
		<b>Footrope experiment</b>		<b>Primary survey</b>		
Species	Underbag	Trawl	Short	Standard	Long	
Atlantic cod	274	935	241	265	311	5774
Haddock	28	5926	8339	13 394	14 151	211 245
yellowtail flounder	6290	3678	1867	2776	3266	26 066
winter flounder	982	2061	1115	1615	1342	21 247
American plaice	653	623	666	778	837	47 989
witch flounder	594	264	61	102	100	12 460
<b>Mean size</b>		<b>Herding experiment</b>				
		<b>Footrope experiment</b>		<b>Primary survey</b>		
Species	Underbag	Trawl	Short	Standard	Long	
Atlantic cod	30.2	51.4	38.3	37.1	47.3	56.4
Haddock	31.7	37.0	39.3	39.3	38.8	38.0
yellowtail flounder	28.7	30.0	30.2	30.1	29.5	30.3
winter flounder	22.5	27.0	26.6	26.9	27.2	28.5
American plaice	25.6	29.1	28.5	28.7	28.9	30.6
witch flounder	36.8	35.9	37.9	37.5	38.2	32.7

cod (+4.1 cm for mean length), lowest for haddock (+0.1 cm), and ranged from +1.2 cm to -1.0 cm for flatfish species (Figure 8). Annual levels of bias in mean length were similar to the overall values for the 3-year period in all cases except cod (2016: +1.2 cm; 2017: +1.8 cm; 2018 +8.0 cm).

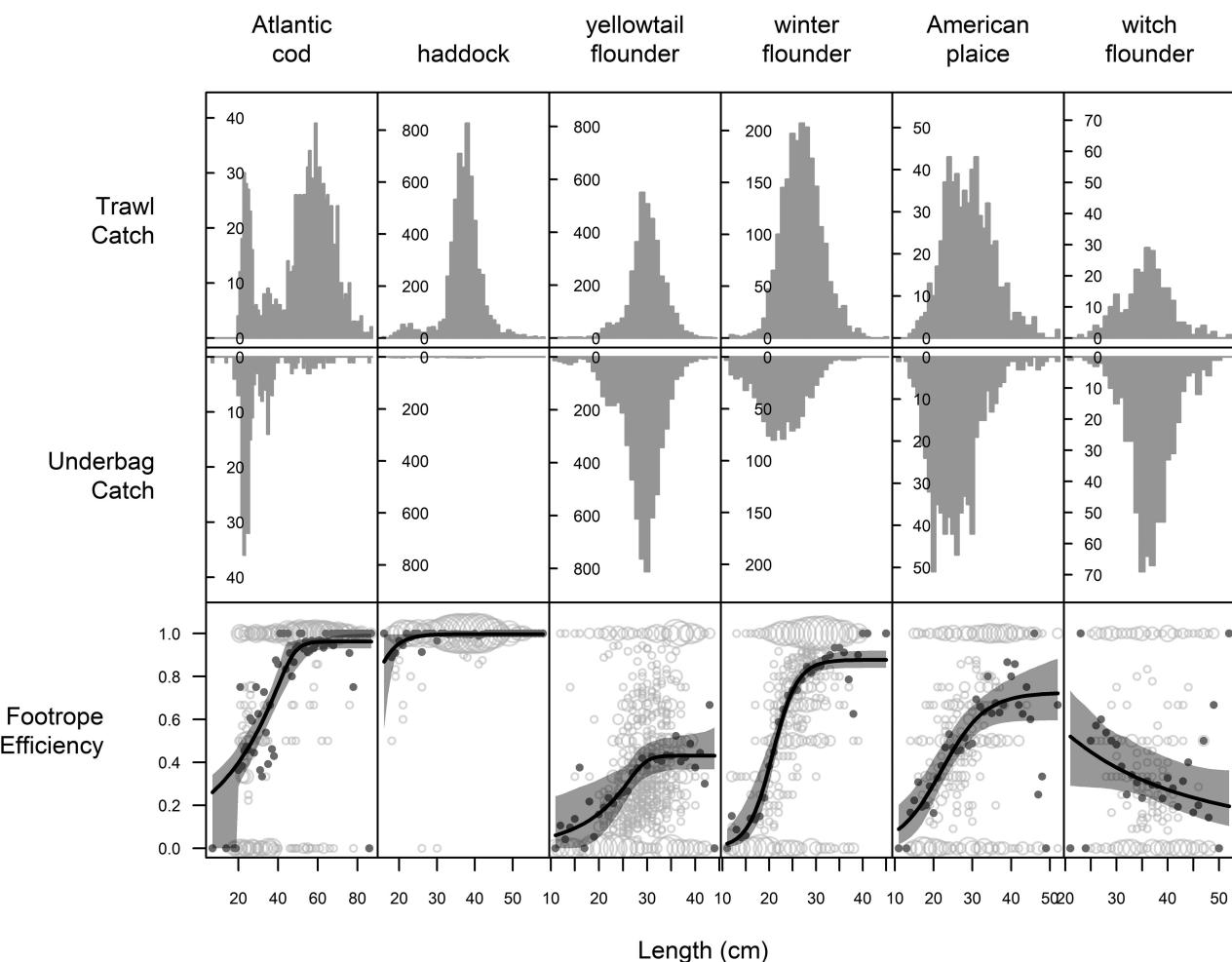
## Discussion

Bottom trawl surveys offer an invaluable perspective on the status and trajectory of demersal fish populations (Harley *et al.*, 2001). Yet, efficiency factors such as the escapement of fish beneath the trawl footrope and the herding effect of the doors and sweeps need to be accounted for to achieve an unbiased view. Several previous studies have conducted experiments to estimate either footrope efficiency (e.g. Engås and Godø, 1989; Munro and Somerton, 2002; Ingólfsson and Jørgensen, 2006) or herding efficiency (e.g. Engås and Godø, 1989; Ramm and Xiao, 1995; Somerton and Munro, 2001) of a survey trawl; relatively few have attempted to estimate both (e.g. Dickson, 1993b; Somerton *et al.*, 2007). In absence of such empirical efficiency estimates, it is common for escapement beneath the footrope to be ignored in analyses of trawl survey datasets (e.g. von Szalay *et al.*, 2007; Keller *et al.*, 2012; Rooper and Martin, 2012). Likewise, herding is typically either ignored or assumed to be complete (e.g. Walsh, 1996; Bull *et al.*, 2001; O'Driscoll, 2003; Fraser *et al.*, 2007; Godbold *et al.*, 2013). Our results suggest that commonly used "default" assumptions of trawl efficiency that assume no footrope escapement (i.e.  $k = 1$ ) and no herding (i.e.  $h = 0$ ) cannot adequately explain the relationship between observed survey catches and actual density. Assuming complete herding (i.e.  $h = 1$ ) introduced

the greatest bias in all cases, which suggests that door width is an unrealistic measure of the effective area swept by the trawl, even for species with a strong herding response (e.g. yellowtail flounder) captured with gear specifically designed to elicit such a response (e.g. our survey trawl).

While some studies have demonstrated a herding response in both gadids (Engås and Godø, 1989b; Sistiaga *et al.*, 2015) and flatfishes (Somerton and Munro, 2001; He *et al.*, 2015), others have been unable to detect or measure a herding effect, particularly for non-flatfish species (Ramm and Xiao, 1995; Somerton, 2004). These authors also noted large spatiotemporal variation in the abundance of the species examined, which hindered their analyses examining the change in catch resulting from modifications to the standard trawl gear. Similarly, many groundfish species in the Gulf of Maine have complex and fine-scale patterns that are often associated with environmental variables, like depth and seafloor type (Nye *et al.*, 2009; Ames and Lichter, 2013; Richardson *et al.*, 2014). Species distribution models, such as random forests, can utilize auxiliary variables to explain a significant portion of the variation in catches from a bottom trawl survey (e.g. Dean *et al.*, 2021), thereby augmenting the ability to observe a treatment effect. Fitting weighted parametric functions to partial dependence from a random forest is an emerging technique to quantify the nature and strength of relationships between explanatory variables and response (e.g. Cafri and Bailey, 2016; Zeng *et al.*, 2017), and represents a powerful alternative to relying on the often-unrealistic assumption that fish densities are equal across adjacent tows.

Trawl efficiency most directly affects *absolute* measures of fish abundance or biomass. These metrics are derived from mean fish



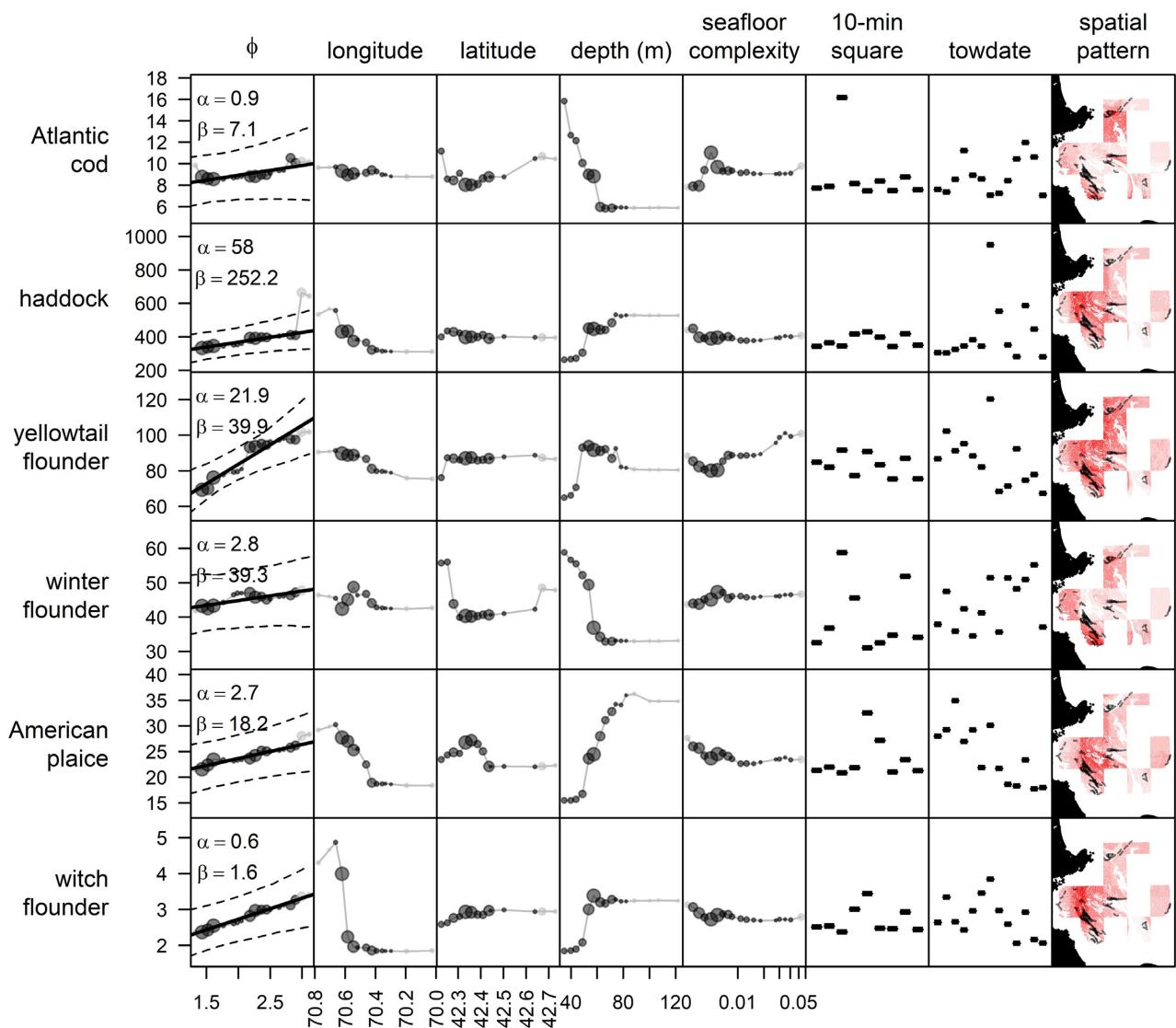
**Figure 3.** Results from footrope efficiency experiment, by species. [Top row] Aggregate length frequency of fish captured in the primary trawl net during footrope efficiency tows. [Middle row] Aggregate length frequency of fish captured in the underbag during the same tows. [Bottom row] Open circles represent raw tow-level footrope efficiency values per 1 cm length bin. Filled circles represent mean empirical values per length bin. Solid lines represent the best-fitting parametric model. Shaded regions represent the 95% prediction interval.

**Table 3.** Parameter estimates for length-dependent footrope efficiency ( $k_l$ ) and herding efficiency ( $h$ ).

Species	Form	Footrope efficiency ( $k_l$ )				Herding efficiency ( $h$ )
		$\alpha$	$\beta$	$\gamma$	$\delta$	
Atlantic cod	Richards	47.05	0.31	0.96	9.44	0.13
Haddock	logistic	9.45	0.29	1.00		0.23
Yellowtail flounder	Richards	28.94	0.73	0.43	6.47	0.55
Winter flounder	logistic	20.73	0.38	0.88		0.07
American plaice	log-logistic	22.03	0.18	0.72		0.15
witch flounder	logistic	22.07	-1.65	1.00		0.37

density (in numbers or weight) within the area “swept” by the trawl and are scaled up to the population by multiplying against the area of the stock or study (Gunderson, 1993). Although various other factors (e.g. depth, habitat and diel period) may be accounted for in the estimation of absolute population size (e.g. Godø and Engås, 1989; Kotwicki *et al.*, 2011), area-swept abundance estimates are directly influenced by bias in the sample of mean density. The

magnitude and direction of this bias varied by species and reflects behavioral differences that modulate the balance between footrope escapement (causing negative bias, assuming  $k = 1$ ) and the herding effect (causing positive bias, assuming  $h = 0$ ). For example, the combination of a moderate herding response ( $h = 0.25$ ) and negligible footrope escapement ( $k \sim 1$ ), resulted in the overestimation of haddock abundance by 53%. In contrast, yellowtail flounder abun-



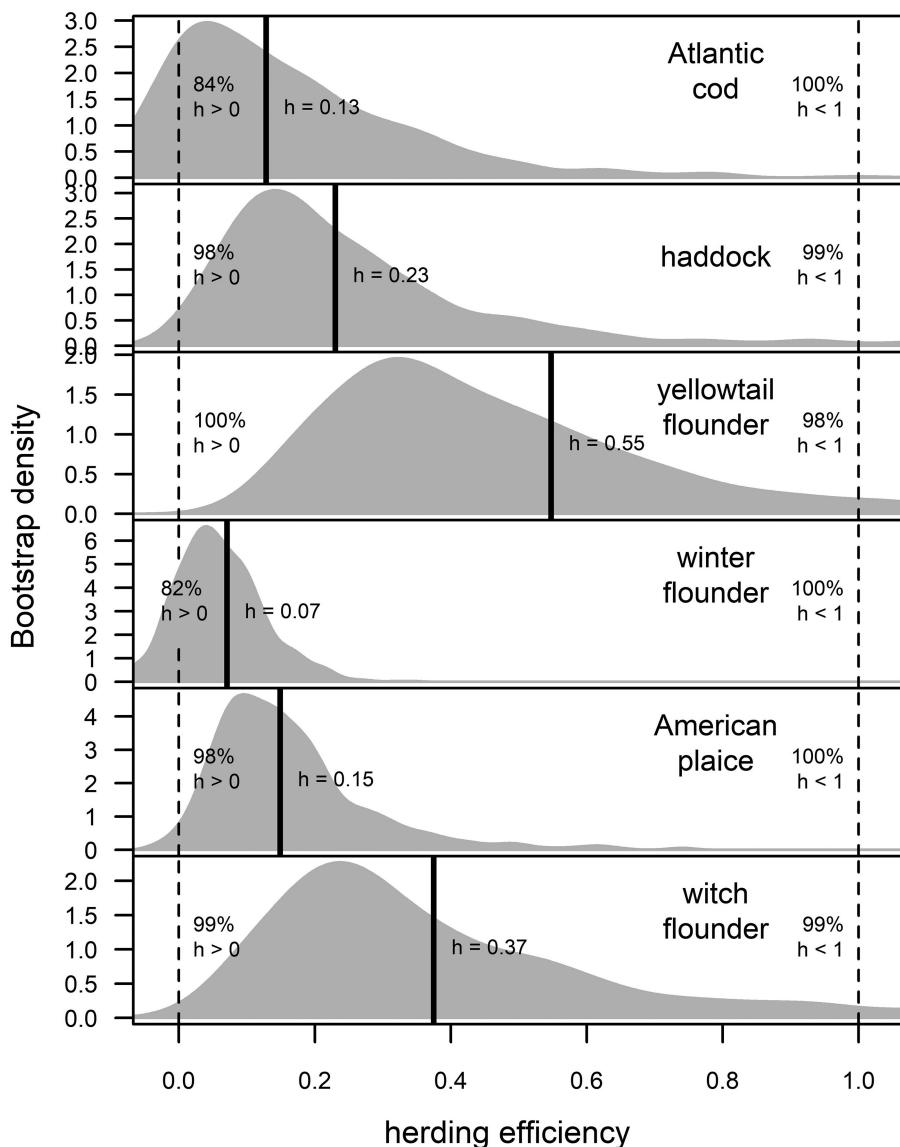
**Figure 4.** Partial dependence of catch ( $C$ , vertical axis) on independent variables from the random forest model applied to the data from the herding experiment. The variable  $\phi$  is the ratio of sweeps-width to netwidth. Latitude, longitude, depth, and seafloor complexity represent mean values along the tow path. Tow date and 10-minute square (of latitude and longitude) are treated as factors. For continuous variables, the size of bubbles is proportional to the number of observations, and lighter gray bubbles fall outside the [0.05, 0.95] sample quantiles. The solid black line represents a linear model fit to the partial dependence of catch on  $\phi$ , weighted by the number of observations, and restricted to the [0.05, 0.95] sample quantiles. Dashed lines represent bootstrapped 95% confidence intervals. The slope ( $\alpha$ ) and intercept ( $\beta$ ) of the  $C \sim \phi$  linear model are shown for each species. The last column depicts the spatial pattern of expected catch (darker red = higher catches), resulting from the combined effect of latitude, longitude, depth, seafloor complexity, and 10-minute square; short black lines indicate the tow paths.

dance was underestimated by 14% because of high footrope escapement, despite having a herding efficiency more than double that of haddock.

Several other studies have identified significant differences in efficiency between closely related species (e.g. Engås and Godø, 1989; Bublitz, 1996). Such species-specific behaviors have been exploited by designers of fishing gear, as in “separator” trawls (He *et al.*, 2008; Krag *et al.*, 2010), which allow flounders and cod to escape, while retaining haddock which are less likely to dive toward the seafloor in response to mobile fishing gear. This lack of a dive-response in haddock is further supported by our data, with < 1% of haddock being caught in the underbag, compared to 20% for cod. Collectively, these results and our study suggest that assuming similar

gear efficiencies across species is likely unmerited (e.g. Walker *et al.*, 2017).

The species-specific nature of trawl efficiency also alters the perception of community composition calculated directly from survey catches. Trawl survey datasets are frequently used to describe the species composition of the demersal fish community without accounting for the effect of trawl efficiency, both for inferring ecological relationships (e.g. Godbold *et al.*, 2013; Farriols *et al.*, 2017; Maureaud *et al.*, 2019) and for apportioning total backscatter (fish “echos”) from hydroacoustic surveys (e.g. Rose and Porter, 1996; McQuinn *et al.*, 2005; Zwolinski *et al.*, 2009). Our results suggest that care should be taken when inferring the apparent dominance (or lack thereof) of particular taxa from raw survey catch data.

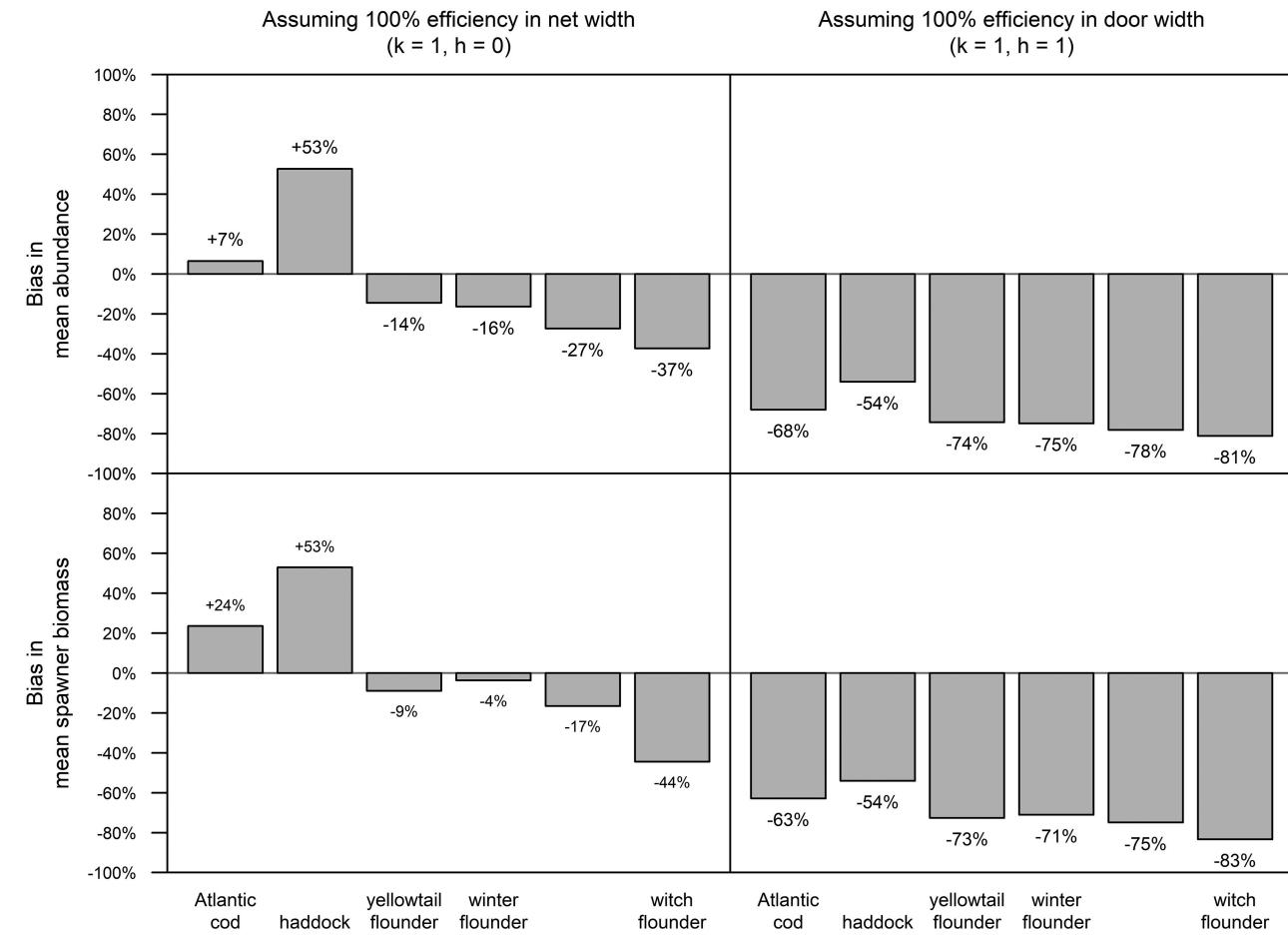


**Figure 5.** Bootstrapped probability distributions for herding efficiency ( $h$ ) by species. The solid vertical line indicates the herding efficiency estimate for each species. The dashed vertical lines represent  $h$  values of 0 and 1, with the probability of  $h > 0$  and  $h < 1$  shown.

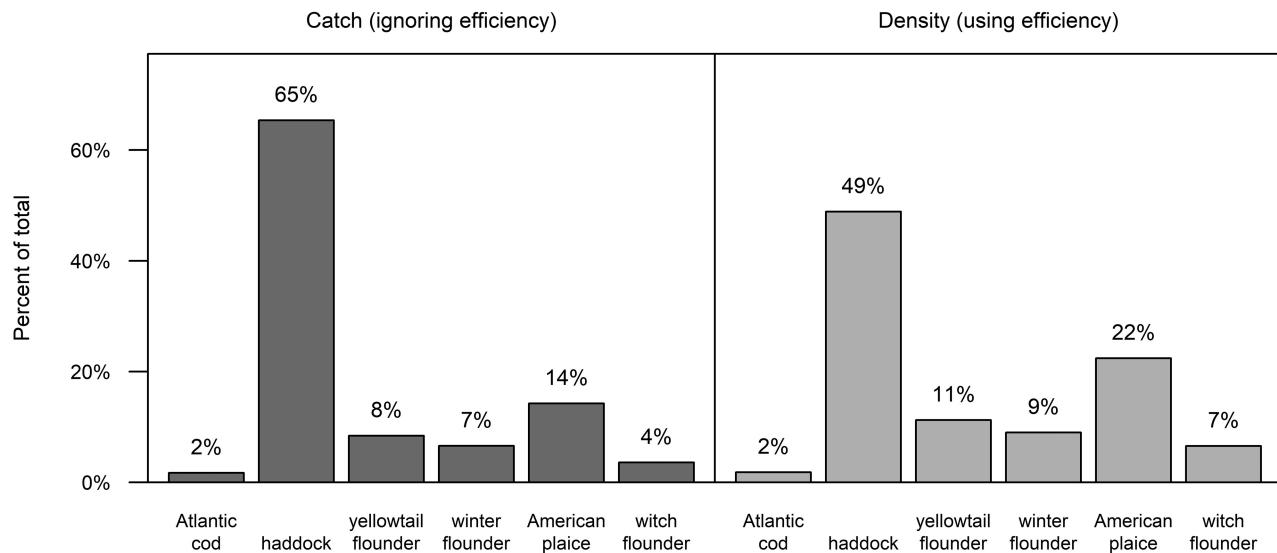
For many uses of trawl survey data, such as when calculating indices of relative abundance or biomass for species individually, the bias caused by improper efficiency assumptions may not substantially alter our perception of the stock, if the entire time series is subject to the same bias. However, if the distribution of a species shifts towards deeper water over time (e.g. through climate change—Dulvy *et al.*, 2008; Nye *et al.*, 2009), the herding effect will have a progressively greater influence on the observed catch, given that  $\emptyset$  increases with depth. If relative indices represent a wide range of sizes or ages in aggregate (e.g. Godø and Wespestad, 1993; Beare *et al.*, 2004), the length-dependent efficiency of the footrope could introduce a bias that affects some years or areas more than others (e.g. a large recruitment event). For example, the relatively high number of juvenile cod observed in 2018 caused the bias in abundance to change by  $-22\%$  in a single year (a similar effect was found for bias in mean size). Thus, stock assessments that utilize survey-based in-

dices of abundance should exercise caution to avoid these potential sources of bias.

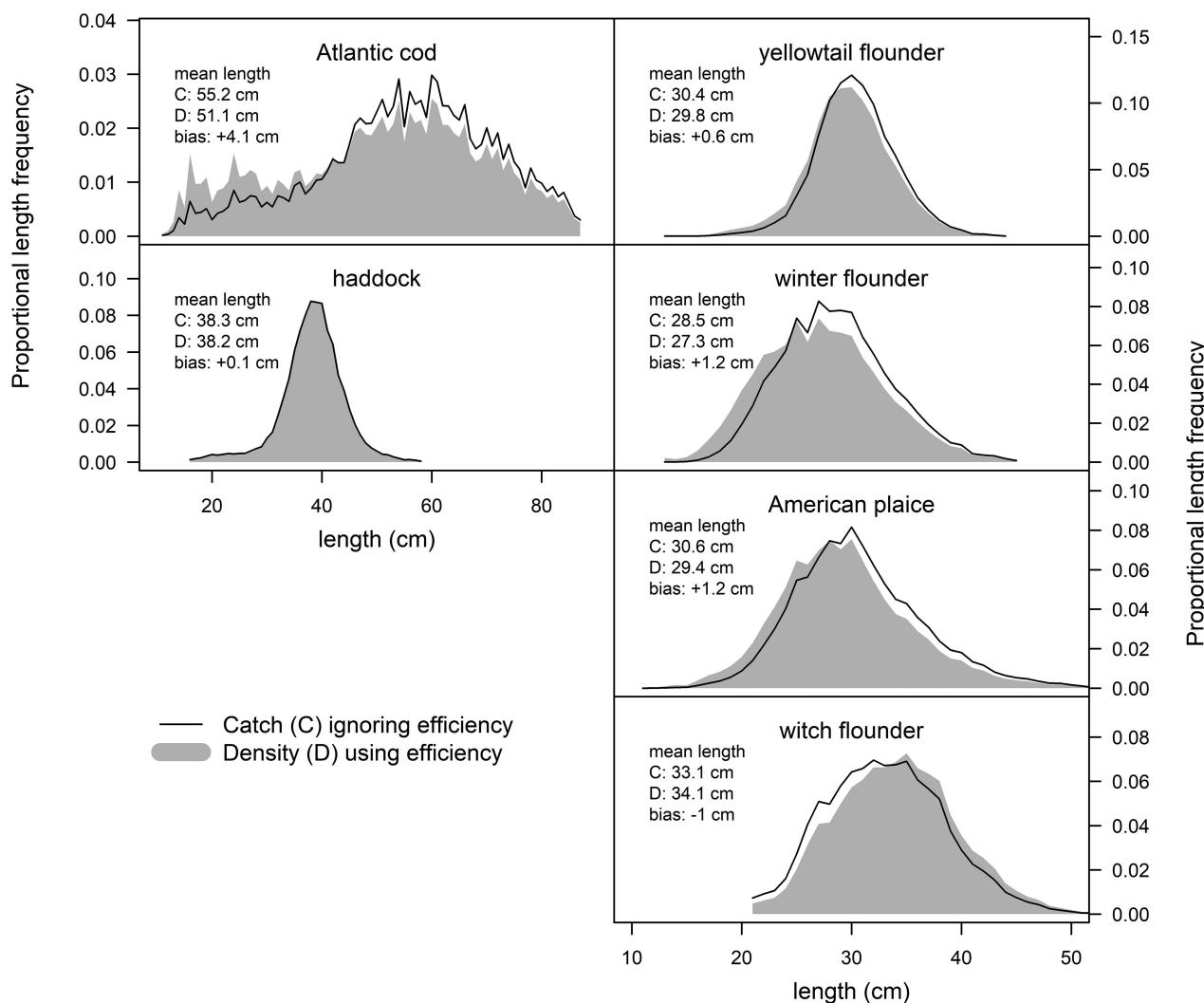
There are several other factors that can influence the efficiency of a bottom trawl that may need to be considered. Given their strong association with the seafloor, it is unlikely that many flatfish escaped above the approximately 4.5 m headrope height of the survey trawl (Ryer, 2008; Ryer *et al.*, 2010). We also expect that few Atlantic cod passed over the headrope, given that acoustically-tagged cod in the Gulf of Maine have been observed to remain within 2 m of the seafloor during the day (Dean *et al.*, 2014) and typically dive toward the seafloor when startled by mobile fishing gear (Handegard *et al.*, 2003; Handegard and Tjøstheim, 2005). Haddock had the highest potential for headrope escapement because of a tendency to swim upward when confronted with a trawl (He *et al.*, 2008), and ignoring this factor may add a negative bias to survey-based estimates of haddock abundance. Handegard and Tjøstheim (2005) also found



**Figure 6.** [Top row] Bias in mean abundance when assuming no footrope escapement and no herding [left column] and when assuming no footrope escapement and complete herding [right column], calculated from the standard survey during 2016–2019. [Bottom row] Same as top row, only with mean abundance translated to mean spawner biomass via functions of maturity-at-length and weight-at-length.



**Figure 7.** [Left] Species composition of the raw survey catch. [Right] Species composition in terms of predicted density, after accounting for the effects of footrope and herding efficiency.



**Figure 8.** Length frequency distribution of the survey catch (gray area) as compared to the predicted density at length of the population, calculated using estimates of footrope efficiency. The mean length of the catch (C) and in terms of density (D) is shown.

that the approaching survey vessel and trawl warps (connecting vessel to doors) could alter the horizontal distribution of fish prior to the arrival of the doors, although this effect was variable and may be due to vessel noise or the agitation of bioluminescent organisms. Trawl efficiency could also depend on the density of fish encountered (Godø *et al.*, 1999; Kotwicki *et al.*, 2014), because individuals may react to an approaching trawl differently when in a school. For some fish species, the reaction to an approaching trawl may be based primarily on visual stimuli (Kim and Wardle, 2003), thus creating diel patterns in trawl efficiency (Engas and Soldal, 1992; Korsbrekke and Nakken, 1999). However, the importance of this factor should be relatively minor in the present study, given that all experimental and primary survey tows were conducted during daylight hours. Similarly, water temperature may impact the swimming endurance of some fishes, creating a seasonal bias in trawl efficiency (Smith and Page, 1996), although the swimming endurance of cod is temperature-independent (Winger *et al.*, 2000). The mesh size of the trawl directly influences capture efficiency for the smallest individuals, and an unknown fraction of these young-of-year escaped through the meshes of the survey net. However, many assessment

models (including those for all stocks considered here) do not estimate abundance of fish younger than age 1 (NEFSC, 2017).

Although the survey trawl was designed to sample a variety of habitats, there are areas which cannot effectively be sampled by this gear (e.g. exposed bedrock crags; large boulder fields). The resulting non-random sampling can introduce a source of bias in survey-based population metrics if the mean density or size of a species in these un-trawlable areas is significantly different from the surveyed areas. More subtle habitat effects may also occur wherein the efficiency of the different trawl components vary with seafloor type (Jagiello *et al.*, 2003; Clarke *et al.*, 2009; Jones *et al.*, 2012). Although we did not directly address the influence of habitat on footrope escapement, experimental tows sampled a variety of seafloor types that were representative of the broader survey dataset. For the herding experiment, the random forest model was designed to account for the influence of habitat variables on catch. The generally low interaction strength between  $\emptyset$  and either depth or seafloor complexity suggests that the observed patterns in catch were driven by patterns in fish density (i.e. habitat preference), not due to a habitat influence on herding. In other words, our estimates of herding

efficiency should be applicable to the broad range of habitat types covered by this study.

It is important to note that  $\emptyset$  increased predictably with depth (Supplementary Figure S1), a pattern which is likely common to most trawl surveys (Godø and Engås, 1989). This depth effect creates spatial variation in bias by altering the relative area swept by each trawl component and therefore modulates the balance between the competing influences of herding and footrope escape. For this reason, using an expected area swept *per depth*, as opposed to a survey-wide average or nominal value, will reduce bias in survey-based population metrics. Some survey groups have used a “restrictor cable” between the warps to limit the variation in trawl geometry (Rose and Nunnallee, 1998; Weinberg and Kotwicki, 2015). We chose to allow door width and net width vary “naturally” with depth, to maintain consistency with the standard survey procedure. One consequence of this decision is that the sweep angle also varied with depth. Some researchers have suggested that a wider sweep angle increases the swim distance needed to reach the trawl path, making it more likely that some individuals (particularly small ones) will lack sufficient endurance to be effectively herded (Weinberg and Kotwicki, 2015). Following the calculations of Winger *et al.* (1999), the maximum swim distance to enter the trawl path for a standard survey tow at the deepest depths (widest sweep angle) is 20 m. Laboratory studies of swimming behaviour and endurance suggest that flatfish can cover this distance in 1–2 swimming bouts with a low probability of fatigue (Winger *et al.*, 2004). Given that the average swim distance to the trawl path under typical conditions will be far less ( $\sim 7$  m), and that the endurance of gadids greatly exceeds that of flatfish (Winger *et al.*, 2000), it is unlikely that fatigue played a large role in the herding efficiency of the survey trawl. Therefore, we consider the width of the sweeps in relation to the net width ( $\emptyset$ ) and not sweep angle to be the primary factor affecting herding efficiency.

Previous studies identified that the herding effect of trawl doors and sweeps is limited by the portion of the sweeps that are in actual contact with the seafloor (Somerton and Munro, 2001; Somerton, 2003; Somerton *et al.*, 2007). Although some amount of bottom separation is necessary near the doors, beyond this point, the trawl gear can be designed to maximize the portion of the sweeps in close contact with the seafloor. Our survey trawl, as well as many commercial trawls, utilize this “herded area” design (Eigaard *et al.*, 2011; Lomeli *et al.*, 2019). Regardless, full seabed contact may not be necessary to elicit a herding response for many species (Lomeli *et al.*, 2019). In our analysis,  $h$  represents the *average* efficiency along the entire sweep length and therefore does not require full seafloor contact; instead, we assume that the *proportion* of the sweeps in contact with the seafloor was consistent among treatments. If this on-bottom proportion were to change with the addition or removal of lengths of sweep cable, we would expect to find non-linearities in the partial dependence of catch on  $\emptyset$ .

The lack of length-based herding that we found is supported by a general similarity in mean fish size across the three sweep-length treatments (Table 2). Other studies of the herding efficiency of bottom trawls for flatfish (Somerton and Munro, 2001; He *et al.*, 2015), gadids (Sistiaga *et al.*, 2015), and tropical groundfish (Ramm and Xiao, 1995) similarly found a general lack of length-dependent herding in most cases. In contrast, Engås and Godø (1989b), and later Dickson (1993b)—analyzing the same dataset), concluded that a longer sweep length led to less small cod and haddock. However, their results were statistically insignificant in many cases, and Dickson’s analysis included an arbitrary and *a priori* assumption of

length-based herding. Given our results, as well as those from prior experiments, we expect that fish size has a relatively minor influence on the herding effect, and that it is reasonable to assume that herding efficiency is length-independent for the species considered here.

In conclusion, our results demonstrate that the efficiency of a specific trawl at capturing a particular species should be estimated and not assumed, whenever possible. Trawl efficiency is a function of the species, the fish size, and the fishing gear, which limits the relevance of assuming values from other studies. Although developing efficiency estimates requires dedicated experiments, these represent a relatively minor investment in the context of the benefits they provide for the utility of an entire survey dataset. For example, our efficiency experiments required 20 out of the total 282 days of vessel time used during the entire three-year survey, yet now enable the calculation of population metrics not subject to the substantial biases caused by simplistic assumptions that would otherwise be necessary. Given the importance of fishery-independent surveys to stock assessments, fishery management, and ecosystem studies this small investment seems highly merited.

## Data availability statement

The data underlying this article will be shared on reasonable request to the corresponding author.

## Funding

This research was supported by the Commonwealth of Massachusetts, USA and Federal Groundfish Disaster Relief funds (NOAA Award # NA14NMF4540357).

## Supplementary data

*Supplementary material* is available at the ICES/JMS online version of the manuscript.

## Acknowledgements

We are grateful to the captains, crew, and scientific staff of the Industry-Based Survey for being capable and cooperative research partners. The collection of sidescan sonar imagery would not have been possible without the efforts of Steve Voss and David Chosid. We would also like to acknowledge the valuable input provided by the primary author’s dissertation committee throughout the development of this manuscript: Tarik Gouhier, Katie Lotterhos, and Steven Scyphers.

## References

- Ames, E. P., and Lichter, J. 2013. Gadids and Alewives: structure within complexity in the Gulf of Maine. *Fisheries Research*, 141: 70–78.
- Auret, L., and Aldrich, C. 2012. Interpretation of nonlinear relationships between process variables by use of random forests. *Minerals Engineering*, 35: 27–42.
- Beare, D. J., Burns, F., Greig, A., Jones, E. G., Peach, K., Kienzle, M., McKenzie, E. *et al.* 2004. Long-term increases in prevalence of North Sea fishes having southern biogeographic affinities. *Marine Ecology Progress Series*, 284: 269–278.
- Bianchi, G., Gislason, H., Graham, K., Hill, L., Jin, X., Koranteng, K., Manickchand-Heileman, S. *et al.* 2000. Impact of fishing on size

- composition and diversity of demersal fish communities. *ICES Journal of Marine Science*, 57: 558–571.
- Breiman, L. 2001. Random forests. *Machine Learning*, 45: 5–32.
- Brinkhof, J., Larsen, R. B., Herrmann, B., and Grimaldo, E. 2017. Improving catch efficiency by changing ground gear design: case study of Northeast Atlantic cod (*Gadus morhua*) in the Barents Sea bottom trawl fishery. *Fisheries Research*, 186: 269–282.
- Bryan, D. R., Bosley, K. L., Hicks, A. C., Haltuch, M. a., and Wakefield, W. W. 2014. Quantitative video analysis of flatfish herding behavior and impact on effective area swept of a survey trawl. *Fisheries Research*, 154: 120–126.
- Bublitz, C. G. 1996. Quantitative evaluation of flatfish behavior during capture by trawl gear. *Fisheries Research*, 25: 293–304.
- Bull, B., Livingston, M. E., Hurst, R., and Bagley, N. 2001. Upper-slope fish communities on the Chatham Rise, New Zealand, 1992–99. *New Zealand Journal of Marine and Freshwater Research*, 35: 795–815.
- Cadrin, S. X., DeCelles, G. R., and Reid, D. 2016. Informing fishery assessment and management with field observations of selectivity and efficiency. *Fisheries Research*, 184: 9–17.
- Cafri, G., and Bailey, B. A. 2016. Understanding variable effects from black box prediction: quantifying effects in tree ensembles using partial dependence. *Journal of Data Science*, 14: 67–96.
- Clarke, M. E., Tolimieri, N., and Singh, H. 2009. Using the seabed AUV to assess populations of groundfish in untrawlable areas. In *The Future of Fisheries Science in North America*, pp. 357–372. Springer.
- Collie, J. S., Wood, A. D., and Jeffries, H. P. 2008. Long-term shifts in the species composition of a coastal fish community. *Canadian Journal of Fisheries and Aquatic Sciences*, 65: 1352–1365.
- Dean, M. J., Elzey, S. P., Hoffman, W. S., Buchan, N. C., and Grabowski, J. H. 2019. The relative importance of sub-populations to the Gulf of Maine stock of Atlantic cod. *ICES Journal of Marine Science*, 76: 1626–1640.
- Dean, M. J., Hoffman, W. S., Buchan, N. C., Cadrin, S. X., and Grabowski, J. H. 2021. Deconstructing size selectivity to evaluate the influence of fishery management. *Fisheries Research*, 234: 105782.
- Dean, M. J., Hoffman, W. S., Zemeckis, D. R., and Armstrong, M. P. 2014. Fine-scale diel and gender-based patterns in behaviour of Atlantic cod (*Gadus morhua*) on a spawning ground in the Western Gulf of Maine. *ICES Journal of Marine Science*, 71: 1474–1489.
- Dickson, W. 1993a. Estimation of the capture efficiency of trawl gear. I: development of a theoretical model. *Fisheries Research*, 16: 239–253.
- Dickson, W. 1993b. Estimation of the capture efficiency of trawl gear. II: testing a theoretical model. *Fisheries Research*, 16: 255–272.
- Dulvy, N. K., Rogers, S. I., Jennings, S., Stelzenmüller, V., Dye, S. R., and Skjoldal, H. R. 2008. Climate change and deepening of the North Sea fish assemblage: a biotic indicator of warming seas. *Journal of Applied Ecology*, 45: 1029–1039.
- Eigaard, O. R., Rihan, D., Graham, N., Sala, A., and Zachariassen, K. 2011. Improving fishing effort descriptors: Modelling engine power and gear-size relations of five European trawl fleets. *Fisheries Research*, 110: 39–46.
- Engås, A., and Godø, O. R. 1989a. Escape of fish under the fishing line of a Norwegian sampling trawl and its influence on survey results. *ICES Journal of Marine Science*, 45: 269–276. <http://icesjms.oxfordjournals.org/content/45/3/269.short>.
- Engås, A., and Godø, O. R. 1989b. The effect of different sweep lengths on the length composition of bottom-sampling trawl catches. *ICES Journal of Marine Science*, 45: 263–289.
- Engås, A., and Soldal, A. 1992. Diurnal variations in bottom trawl catch rates of cod and haddock and their influence on abundance indices. *ICES Journal of Marine Science*, 48: 89–92.
- Evans, J. S., Murphy, M. A., Holden, Z. A., and Cushman, S. A. 2011. Modeling species distribution and change using random forest. In *Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications*, pp. 139–159. Springer, New York, NY, USA.
- Farriols, M. T., Ordines, F., Somerfield, P. J., Pasqual, C., Hidalgo, M., Guijarro, B., and Massutí, E. 2017. Bottom trawl impacts on Mediterranean demersal fish diversity: not so obvious or are we too late? *Continental Shelf Research*, 137: 84–102.
- Fraser, H. M., Greenstreet, S. P. R., and Piet, G. J. 2007. Taking account of catchability in groundfish survey trawls: implications for estimating demersal fish biomass. *ICES Journal of Marine Science*, 64: 1800–1819.
- Friedman, J. H. 2001. Greedy function approximation: a gradient boosting machine. *The Annals of Statistics*, 29: 1189–1232.
- Friedman, J. H., and Popescu, B. E. 2008. Predictive learning via rule ensembles. *The Annals of Applied Statistics*, 2: 916–954.
- Gibson, R. N., Ansell, A. D., and Robb, L. 1993. Seasonal and annual variations in abundance and species composition of fish and macrocrustacean communities on a Scottish sandy beach. *Marine Ecology Progress Series*, 98: 89–105.
- Godbold, J. A., Bailey, D. M., Collins, M. A., Gordon, J. D. M., Spallek, W. A., and Priede, I. G. 2013. Putative fishery-induced changes in biomass and population size structures of demersal deep-sea fishes in ICES Sub-area VII, Northeast Atlantic Ocean. *Biogeosciences*, 10: 529–539.
- Godø, O. R., and Engås, A. 1989. Swept area variation with depth and its influence on abundance indices of groundfish from trawl surveys. *Journal of Northwest Atlantic Fishery Science*, 9: 133–139.
- Godø, O. R., Engås, A., and Walsh, S. J. 1999. Investigating density-dependent catchability in bottom-trawl surveys. *ICES Journal of Marine Science*, 56: 292–298.
- Godø, O. R., and Wespestad, V. G. 1993. Monitoring changes in abundance of gadoids with varying availability to trawl and acoustic surveys. *ICES Journal of Marine Science*, 50: 39–51. <http://icesjms.oxfordjournals.org/content/50/1/39.abstract>.
- Gunderson, D. R. 1993. *Surveys of Fisheries Resources*. John Wiley & Sons, New York, NY. 1–256pp.
- Handegard, N. O., Michelsen, K., and Tjøstheim, D. 2003. Avoidance behaviour in cod (*Gadus morhua*) to a bottom-trawling vessel. *Aquatic Living Resources*, 16: 265–270.
- Handegard, N. O., and Tjøstheim, D. 2005. When fish meet a trawling vessel: examining the behaviour of gadoids using a free-floating buoy and acoustic split-beam tracking. *Canadian Journal of Fisheries and Aquatic Sciences*, 62: 2409–2422.
- Harley, S. J., Myers, R. a., and Dunn, A. 2001. Is catch-per-unit-effort proportional to abundance? *Canadian Journal of Fisheries and Aquatic Sciences*, 58: 1760–1772.
- Hayes, D. B., Ferreri, C. P., and Taylor, W. W. 2013. Active fish capture methods. In *Fisheries Techniques*, 3rd edn, pp. 267–304. American Fisheries Society. doi: [10.47886/9781934874295.ch7](https://doi.org/10.47886/9781934874295.ch7).
- He, P. 2010. *Behavior of Marine Fishes: Capture Processes and Conservation Challenges*. Wiley-Blackwell. 1–400pp.
- He, P., Rillahan, C., and Balzano, V. 2015. Reduced herding of flounders by floating bridles: application in Gulf of Maine Northern shrimp trawls to reduce bycatch. *ICES Journal of Marine Science*, 72: 1514–1524.
- He, P., Smith, T., and Bouchard, C. 2008. Fish behaviour and species separation for the Gulf of Maine multispecies trawls. *Journal of Ocean Technology*, 3: 59–77.
- Hedger, R., McKenzie, E., Heath, M., Wright, P., Scott, B., Gallego, A., and Andrews, J. 2004. Analysis of the spatial distributions of mature cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*) abundance in the North Sea (1980–1999) using generalised additive models. *Fisheries Research*, 70: 17–25.
- Helser, T. E., and Hayes, D. B. 1995. Providing quantitative management advice from stock abundance indices based on research surveys. *Fishery Bulletin*, 93: 290–298.
- Hoffman, W., Correia, S., and Pierce, D. 2012. Results of an Industry-Based Survey for Gulf of Maine Cod, November 2003 — May 2005. Massachusetts Division of Marine Fisheries Technical Report: 37.
- ICES. 2019. Workshop on Methods to develop a swept-area based effort index (WKSABI). *ICES Scientific Reports*, 1: 24pp.
- Ingólfsson, Ó. A., and Jørgensen, T. 2006. Escapement of gadoid fish beneath a commercial bottom trawl: relevance to the overall trawl selectivity. *Fisheries Research*, 79: 303–312.

- Jagiello, T., Hoffmann, A., Tagart, J., and Zimmermann, M. 2003. Demersal groundfish densities in trawlable and untrawlable habitats off Washington: implications for the estimation of habitat bias in trawl surveys. *Fishery Bulletin*, 101: 545–565.
- Jones, D. T., Wilson, C. D., De Robertis, A., Rooper, C. N., Weber, T. C., and Butler, J. L. 2012. Evaluation of rockfish abundance in untrawlable habitat: combining acoustic and complementary sampling tools. *Fishery Bulletin*, 109: 332–343.
- Keller, A. A., Wallace, J. R., Horness, B. H., Hamel, O. S., and Stewart, I. J. 2012. Variations in eastern North Pacific demersal fish biomass based on the U.S. west coast groundfish bottom trawl survey (2003–2010). *Fishery Bulletin*, 110: 2003–2010.
- Kim, Y. H., and Wardle, C. S. 2003. Optomotor response and erratic response: quantitative analysis of fish reaction to towed fishing gears. *Fisheries Research*, 60: 455–470.
- Korsbrekke, K., and Nakken, O. 1999. Length and species-dependent diurnal variation of catch rates in the Norwegian Barents Sea bottom-trawl surveys. *ICES Journal of Marine Science*, 56: 284–291.
- Kotwicki, S., Ianelli, J. N., and Punt, A. E. 2014. Correcting density-dependent effects in abundance estimates from bottom-trawl surveys. *ICES Journal of Marine Science*, 71: 1107–1116.
- Kotwicki, S., Martin, M. H., and Laman, E. A. 2011. Improving area swept estimates from bottom trawl surveys. *Fisheries Research*, 110: 198–206.
- Krag, L. A., Holst, R., Madsen, N., Hansen, K., and Frandsen, R. P. 2010. Selective haddock (*Melanogrammus aeglefinus*) trawling: avoiding cod (*Gadus morhua*) bycatch. *Fisheries Research*, 101: 20–26.
- Lampa, E., Lind, L., and Lind, P. M. 2014. The identification of complex interactions in epidemiology and toxicology: a simulation study of boosted regression trees.. *Environmental Health*, 13: 1–17.
- Lenoir, S., Beaugrand, G., and Lecuyer, É. 2011. Modelled spatial distribution of marine fish and projected modifications in the North Atlantic Ocean. *Global Change Biology*, 17: 115–129.
- Liaw, A., and Wiener, M. 2002. Classification and regression by randomForest. *R News*, 2: 18–22.
- Lomeli, M. J. M., Wakefield, W. W., and Herrmann, B. 2019. Evaluating off-bottom sweeps of a U.S. West Coast groundfish bottom trawl: effects on catch efficiency and seafloor interactions. *Fisheries Research*, 213: 204–211.
- Maureaud, A., Hodapp, D., Daniël Van Denderen, P., Hillebrand, H., Gislason, H., Dencker, T. S., Beukhof, E. et al. 2019. Biodiversity-ecosystem functioning relationships in fish communities: biomass is related to evenness and the environment, not to species richness. *Proceedings of the Royal Society B: Biological Sciences*, 286: 20191189. doi: [10.1098/rspb.2019.1189](https://doi.org/10.1098/rspb.2019.1189).
- Maxwell, D. L., Stelzenmüller, V., Eastwood, P. D., and Rogers, S. I. 2009. Modelling the spatial distribution of plaice (*Pleuronectes platessa*), sole (*Solea solea*) and thornback ray (*Raja clavata*) in UK waters for marine management and planning. *Journal of Sea Research*, 61: 258–267. Elsevier B.V. doi: [10.1016/j.seares.2008.11.008](https://doi.org/10.1016/j.seares.2008.11.008).
- McQuinn, I. H., Simard, Y., Stroud, T. W. F., Beaulieu, J. L., and Walsh, S. J. 2005. An adaptive, integrated ‘acoustic-trawl’ survey design for Atlantic cod (*Gadus morhua*) with estimation of the acoustic and trawl dead zones. *ICES Journal of Marine Science*, 62: 93–106.
- Munro, P. T., and Somerton, D. A. 2002. Estimating Net efficiency of a survey trawl for flatfishes. *Fisheries Research*, 55: 267–279.
- Northeast Fisheries Science Center (NEFSC). 2011. 52nd Northeast Regional Stock Assessment Workshop (52nd SAW) Assessment Report. 1–962pp.
- Northeast Fisheries Science Center (NEFSC). 2015. Georges Bank yellowtail flounder 2015 assessment update report. 1–7pp.
- Northeast Fisheries Science Center (NEFSC). 2017. 62nd Northeast Regional Stock Assessment Workshop (62nd SAW) Assessment Report. 1–822pp.
- Northeast Fisheries Science Center (NEFSC). 2019. Operational assessment of 14 northeast groundfish stocks, updated through 2018. 1–208pp.
- Northeast Fisheries Science Center (NEFSC). 2020a. Northern red hake 2020 assessment update report 1–9pp.
- Northeast Fisheries Science Center (NEFSC). 2020b. Northern windowpane flounder 2020 assessment update report 1–7pp.
- Nye, J. A., Link, J. S., Hare, J. A., and Overholtz, W. J. 2009. Changing spatial distribution of fish stocks in relation to climate and population size on the Northeast United States continental shelf. *Marine Ecology Progress Series*, 393: 111–129.
- O'Driscoll, R. L. 2003. Determining species composition in mixed-species marks: an example from the New Zealand hoki (*Macruronus novaezealandiae*) fishery. *ICES Journal of Marine Science*, 60: 609–616.
- Pennington, M., and Strømme, T. 1998. Surveys as a research tool for managing dynamic stocks. *Fisheries Research*, 37: 97–106.
- Ramm, D. C., and Xiao, Y. 1995. Herding in groundfish and effective pathwidth of trawls. *Fisheries Research*, 24: 243–259.
- Richardson, D. E., Palmer, M. C., Smith, B. E., and Cooper, A. 2014. The influence of forage fish abundance on the aggregation of Gulf of Maine Atlantic cod (*Gadus morhua*) and their catchability in the fishery. *Canadian Journal of Fisheries and Aquatic Sciences*, 14: 1–14..
- Rooper, C. N., and Martin, M. H. 2012. Comparison of habitat-based indices of abundance with fishery-independent biomass estimates from bottom trawl surveys. *Fishery Bulletin*, 110: 21–35.
- Rose, C. S., Gauvin, J. R., and Hammond, C. F. 2010. Effective herding of flatfish by cables with minimal seafloor contact. *Fishery Bulletin*, 108: 136–144.
- Rose, C. S., and Nunnallee, E. P. 1998. A study of changes in groundfish trawl catching efficiency due to differences in operating width, and measures to reduce width variation. *Fisheries Research*, 36: 139–147.
- Rose, G., and Porter, D. 1996. Target-strength studies on Atlantic cod (*Gadus morhua*) in Newfoundland waters. *ICES Journal of Marine Science*, 53: 259–265. <http://icesjms.oxfordjournals.org/content/53/2/259.short>.
- Ryer, C. H. 2008. A review of flatfish behavior relative to trawls. *Fisheries Research*, 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. *Fishery Bulletin*, 108: 145–154.
- Sistiaga, M., Herrmann, B., Grimaldo, E., Larsen, R. B., and Tatone, I. 2015. Effect of lifting the sweeps on bottom trawling catch efficiency: a study based on the Northeast arctic cod (*Gadus morhua*) trawl fishery. *Fisheries Research*, 167: 164–173.
- Smith, S. J., and Page, F. H. 1996. Associations between Atlantic cod (*Gadus morhua*) and hydrographic variables: implications for the management of the 4VsW cod stock. *ICES Journal of Marine Science*, 53: 597–614.
- Somerton, D. A. 2003. Bridle efficiency of a survey trawl for flatfish: measuring the length of the bridles in contact with the bottom. *Fisheries Research*, 60: 273–279.
- Somerton, D. A. 2004. Do Pacific cod (*Gadus macrocephalus*) and wall-eye pollock (*Theragra chalcogramma*) lack a herding response to the doors, bridles, and mudclouds of survey trawls? *ICES Journal of Marine Science*, 61: 1186–1189.
- Somerton, D. A., and Munro, P. 2001. Bridle efficiency of a survey trawl for flatfish. *Fishery Bulletin*, 99: 641–652.
- Somerton, D. A., Munro, P. T., and Weinberg, K. L. 2007. Whole-gear efficiency of a benthic survey trawl for flatfish. *Fishery Bulletin*, 105: 278–291.
- Somerton, D., Ianelli, J., Walsh, S., Smith, S., Godø, O. R., and Ramm, D. 1999. Incorporating experimentally derived estimates of survey trawl efficiency into the stock assessment process: a discussion. *ICES Journal of Marine Science*, 56: 299–302.
- Stockwell, D. R. B., and Peterson, A. T. 2002. Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, 148: 1–13.
- Swartzman, G., Huang, C., and Kaluzny, S. 1992. Spatial analysis of Bering Sea groundfish survey data using generalized additive models. *Canadian Journal of Fisheries and Aquatic Sciences*, 49: 1366–1378.

- Valdemarsen, J. W., Jorgensen, T., and Engas, A. 2007. Options to mitigate bottom habitat impact of dragged gears. FAO Fisheries Technical Paper, 506: 1–29.
- von Szalay, P. G., Somerton, D. A., and Kotwicki, S. 2007. Correlating trawl and acoustic data in the eastern Bering Sea: a first step toward improving biomass estimates of walleye pollock (*Theragra chalcogramma*) and Pacific cod (*Gadus macrocephalus*)? *Fisheries Research*, 86: 77–83.
- Walker, N. D., Maxwell, D. L., Le Quesne, W. J. F., and Jennings, S. 2017. Estimating efficiency of survey and commercial trawl gears from comparisons of catch-ratios. *ICES Journal of Marine Science*, 74: 1448–1457.
- Walsh, S. J. 1996. Efficiency of bottom sampling trawls in deriving survey abundance indices. *NAFO Scientific Council Studies*, 28: 9–24.
- Weinberg, K. L., and Kotwicki, S. 2008. Factors influencing net width and sea floor contact of a survey bottom trawl. *Fisheries Research*, 93: 265–279.
- Weinberg, K. L., and Kotwicki, S. 2015. Reducing variability in bottom contact and net width of a survey trawl by restraining door movement and applying a constant ratio of warp length to depth. *Fishery Bulletin*, 113: 180–190.
- Weinberg, K. L., Somerton, D. a., and Munro, P. T. 2002. The effect of trawl speed on the footrope capture efficiency of a survey trawl. *Fisheries Research*, 58: 303–313.
- Wilson, M. F. J., O'Connell, B., Brown, C., Guinan, J. C., and Grehan, A. J. 2007. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy*, 30: 3–35.
- Winger, P. D., He, P., and Walsh, S. J. 1999. Swimming endurance of American plaice (*Hippoglossoides platessoides*) and its role in fish capture. *ICES Journal of Marine Science*, 56: 252–265.
- Winger, P. D., He, P., and Walsh, S. J. 2000. Factors affecting the swimming endurance and catchability of Atlantic cod (*Gadus morhua*). *Canadian Journal of Fisheries and Aquatic Sciences*, 57: 1200–1207.
- Winger, P. D., Walsh, S. J., He, P., and Brown, J. A. 2004. Simulating trawl herding in flatfish: the role of fish length in behaviour and swimming characteristics. *ICES Journal of Marine Science*, 61: 1179–1185.
- Zeng, C., Yang, L., and Zhu, A.-X. 2017. Construction of membership functions for soil mapping using the partial dependence of soil on environmental covariates calculated by random forest. *Soil Science Society of America Journal*, 81: 341–353.
- Zhang, H., Zimmerman, J., Nettleton, D., and Nordman, D. J. 2020. Random forest prediction intervals. *The American Statistician*, 74: 392–406.
- Zwolinski, J., Fernandes, P. G., Marques, V., and Stratoudakis, Y. 2009. Estimating fish abundance from acoustic surveys: calculating variance due to acoustic backscatter and length distribution error. *Canadian Journal of Fisheries and Aquatic Sciences*, 66: 2081–2095.

Handling Editor: Finbarr O'Neill