



**NOAA Technical Memorandum NMFS-NE-229**

# **Statistically Assessing the Precision of Self-reported VTR Fishing Locations**

**US DEPARTMENT OF COMMERCE  
National Oceanic and Atmospheric Administration  
National Marine Fisheries Service  
Northeast Fisheries Science Center  
Woods Hole, Massachusetts  
June 2014**



## **NOAA Technical Memorandum NMFS-NE-229**

This series represents a secondary level of scientific publishing. All issues employ thorough internal scientific review; some issues employ external scientific review. Reviews are transparent collegial reviews, not anonymous peer reviews. All issues may be cited in formal scientific communications.

# **Statistically Assessing the Precision of Self-reported VTR Fishing Locations**

by Geret Sean DePiper

National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543

**US DEPARTMENT OF COMMERCE**  
Penny Pritzker, Secretary  
**National Oceanic and Atmospheric Administration**  
Kathryn Sullivan, Acting NOAA Administrator  
**National Marine Fisheries Service**  
Samuel D Rauch III, Acting Assistant Administrator for Fisheries  
**Northeast Fisheries Science Center**  
Woods Hole, Massachusetts

**June 2014**

## Editorial Notes

**Information Quality Act Compliance:** In accordance with section 515 of Public Law 106-554, the Northeast Fisheries Science Center completed both technical and policy reviews for this report. These predissemination reviews are on file at the NEFSC Editorial Office.

**Species Names:** The NEFSC Editorial Office's policy on the use of species names in all technical communications is generally to follow the American Fisheries Society's lists of scientific and common names for fishes, mollusks, and decapod crustaceans and to follow the Society for Marine Mammalogy's guidance on scientific and common names for marine mammals. Exceptions to this policy occur when there are subsequent compelling revisions in the classifications of species, resulting in changes in the names of species.

**Statistical Terms:** The NEFSC Editorial Office's policy on the use of statistical terms in all technical communications is generally to follow the International Standards Organization's handbook of statistical methods.

**Internet Availability:** This issue of the NOAA Technical Memorandum NMFS-NE series is being published as a Web document in PDF format and can be accessed at:  
<http://www.nefsc.noaa.gov/publications/>.

**Editorial Treatment:** To distribute this report quickly, it has not undergone the normal technical and copy editing by the Northeast Fisheries Science Center's (NEFSC's) Editorial Office as have most other issues in the NOAA Technical Memorandum NMFS-NE series. Other than the four covers and first two preliminary pages, all writing and editing have been performed by – and all credit for such writing and editing rightfully belongs to – those so listed on the title page.

## TABLE OF CONTENTS

|                        |   |
|------------------------|---|
| Abstract .....         | 1 |
| Introduction .....     | 1 |
| Data and Methods ..... | 1 |
| Conclusion .....       | 6 |
| References Cited ..... | 7 |

## **ABSTRACT**

The precision of self-reported VTR points often comes into question, despite its importance in both fisheries management and stock assessment. This manuscript uses a novel statistical approach to assess the spatial precision of these points in order to generate a better understanding of how the data can best be used.

## **INTRODUCTION**

By merging Vessel Trip Reports (VTR) with Northeast Fisheries Observer Program data at the trip level, statistical models can be developed to rigorously assess the spatial precision of VTR through the comparison of VTR self-reported fishing locations with observed haul locations. Furthermore, we parameterize and estimate cumulative distribution functions for the distance between VTR points and observed sets/hauls by using only variables reported directly on the VTR. This method allows a flexible framework from which to generate out-of-sample predictions for the spatial footprint of fishing, covering the universe of VTR data available and filling a hole in the current understanding of the spatial precision of the data in question. This approach also allows precision to be assessed for periods in which a Vessel Monitoring System (VMS) does not exist.

The precision issue associated with VTR is inherent in the attempt to represent the entirety of a trip's effort by a single set of latitude and longitudinal points for each gear and statistical area fished, regardless of the length of that trip. The problem is further compounded by issues such as an underreporting bias associated with the number of gear and statistical areas fished (Palmer and Wigley 2007, 2009).

Traditionally, studies utilizing VTR fall back on 1 of 2 approaches to the spatial data: either using the raw latitude and longitude (lat-lon) points, or spatially joining the data to predetermined grids, often composed of 10 min squares. Both of these approaches rely on a priori assumptions regarding VTR spatial precision. However, the VTR instructions state that fishermen must "Enter a single set of latitude [longitude] bearings (degree and minutes) where most of your effort occurred," which provides no obvious guidance for the precision of self-reported fishing location (Northeast Regional Office 2014). There have been some important strides in utilizing secondary data sources to more rigorously account for fishing location in the last decade (Palmer and Wigley 2007, 2009; Records and Demarest 2014). However, these tend to rely on VMS data, which only cover a subset of vessels for which a realistic spatial footprint might be of interest.

## **DATA AND METHODS**

### **Data Sources**

This study used 13 years of observations (2000 – 2012). A dataset including permit, vessel hull number, date sailed, date landed, area fished, gear code, latitude, and longitude was compiled from the VTR database's trip and gear tables, for the years of interest. This dataset was processed to remove records for which missing values of latitude, longitude, or hull number existed and to generate a variable for trip length, in days, by differencing the date landed and date sailed variables and the rounding up. A second dataset comprising link1, link3, hull number

(hullnum1), date sailed, area fished, date landed, and haul beginning and endpoints (lat-lon) was generated for the relevant years by querying the haul and trip tables of the Observer database. The 2 datasets were then joined through a hierarchical matching algorithm. Table 1 outlines the variables used to match VTR and observed haul records in each round of the algorithm. The algorithm resulted in 488,251 hauls in the OBDBS being matched to 27,358 VTR records. The matched observations represent 87.5% of all hauls with either a beginning or end point of a haul recorded. Distance between haul beginning and end points were calculated by using a haversine function, and the top 1% of the distribution was dropped to remove hauls whose length indicated data errors in the observer dataset.

The joined VTR-Observer dataset was then exported into ArcGIS. For hauls with both beginning and endpoints, the haul path was imputed by using a straight line between these 2 points. Minimum distances between the VTR position and either the haul path, or haul beginning and end points if only 1 was recorded, was then calculated in nautical miles. These data were then exported into Stata for the statistical model estimation. In order to allow for out of sample predictions, 10% of the hauls were randomly selected and held back from the modeling endeavor.

Tables 2 – 4 present descriptions of variables from the VTRs within the matched dataset which are of interest in the current study. Bottom trawl and scallop dredge trips make up the vast majority of observed trips within the dataset, as evident in Table 2. Some of the gear categories contain too few observations to use in the model. The purse seine observations were therefore folded into the bottom trawl gear category, and the other dredge recoded as scallop dredge for modeling purposes, while the 5 observations with either unknown or harpoon gear classifications were dropped.<sup>1</sup> Table 3 details the diversity of trip lengths within the dataset, with only 34% of the observations occurring on trips reported to be less than 7 days long. Table 4 details the areas of the ocean in which fishing on these observed trips was reported within the VTR, with the majority coming from southern New England and the Mid-Atlantic Bight. In the next section we detail how this information will be used to model the distance between observed hauls and self-reported fishing location.

## Statistical Model and Estimation

The final dataset can be understood as a repeated measure of the distance on a single trip between observed hauls and the self-reported location of fishing. As such, it is analogous to what is known as either duration or survival modeling in economics, in which the researcher models the amount of time until an event occurs. Duration models are often employed in order to assess the conditional time until an event occurs. By recasting these models from time into distance, a realistic spatial footprint can be assessed as a function of trip characteristics likely to affect precision of reported fishing locations. Formally, we are interested in estimating the cumulative distribution function (CDF) for the distance between the self-reported centroid of fishing and observed hauls, conditional on observed characteristics of that trip. With censoring of observations not an issue, this can be represented mathematically as:

$$(1) \quad Pr(D \leq d) = F(d) \equiv \int_0^d f(x)dx,$$

---

<sup>1</sup> Although other gear types are fished in a manner more similar to purse seine, the same does not necessarily hold for the manner in which reporting occurs. The conservative decision was therefore made to combine purse seine with bottom trawl, the gear representing the largest spatial footprint.

where  $F(\cdot)$  represents a generic CDF,  $f(\cdot)$  is the corresponding probability density function,  $d$  is a given distance, and  $\Pr(\cdot)$  is the probability of a haul occurring within distance  $d$  from the reported centroid of fishing. Although the final form depends on the parametric distribution, the CDF can be estimated through maximum likelihood, or equivalently maximizing the log-likelihood  $\ln L(\theta) = \sum_{i=1}^n \ln f(d_i|\theta)$ . In this log-likelihood,  $\theta$  represents a vector of parameters to be estimated, and  $n$  is the number of observations. The choice of distribution is not trivial, as different distributions will impose substantially different restrictions on the dependence between the rate of haul occurrence and the distance from the self-reported fishing location. See, for example Van den Berg (2000), Kiefer (1988), or chapter 22 in Greene (2003). For the purposes of this paper, the 3-parameter (shape, scale, and location) generalized gamma distribution is adopted. The gamma distribution is a flexible functional form, with the log-normal, exponential, and Weibull distributions as special cases, which provides appeal. However, criteria such as Akaike's Information Criterion (AIC) or Bayesian Information Criterion can be used to assess the appropriateness of different distributional assumptions. The location is parameterized such that  $\mu_i = X_i\beta$ , with  $X_i$  being a vector of observed characteristics on trip  $i$ , and  $\beta$  representing a parameter vector to be estimated, while the shape  $\kappa$  and scale  $\sigma$  are estimated as free parameters. Following the notation used in StataCorp (2011), the exact specification of the 3-parameter gamma distribution depends on the value of the shape parameter:

$$(2) \quad f(d) = \begin{cases} \frac{\gamma^\gamma}{\sigma d \sqrt{\gamma} \Gamma(\gamma)} \exp(z\sqrt{\gamma} - u) & \text{if } \kappa \neq 0 \\ \frac{1}{\sigma d \sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) & \text{if } \kappa = 0 \end{cases}$$

where:

$$(3) \quad \gamma = |\kappa|^{-2},$$

$$(4) \quad z = \text{sign}(\kappa)(\ln(d) - X\beta)/\sigma,$$

$$(5) \quad u = \gamma \exp(|\kappa|z).$$

Here  $\Gamma(\cdot)$  is the gamma function,  $u$  is the standardized distance, and all other arguments are as previously defined. Note that if  $\kappa = 0$ , the log-normal distribution results.

By parameterizing the model as a function of observed trip characteristics, the model can control for variability between trips that would theoretically be expected to impact the precision of the self-reported fishing locations. For purposes of this paper, we parameterize the CDF as a function of characteristics solely reported on the VTR and likely to explain a substantial amount of variance in precision across trips. These trip characteristics include gear employed, trip length (coded as a discrete categorical variable), and subarea fished.<sup>2</sup> Gears are fished in different manners, and a highly mobile gear such as bottom trawls chasing mobile fish are likely to cover more ground on a fishing trip than a gear such as scallop dredge which harvests sessile organisms highly concentrated in known beds. Furthermore, hauls on a 7 day trip would be expected to be more diffuse than those of a 1 day trip. Even with trips using a single gear and

---

<sup>2</sup> Additional variables including month, season, and year indicators, as well as vessel length and horsepower, proved to be insignificant in alternate specifications.

fully encompassed within a single statistical area, a longer trip would be expected to cover more area of the ocean than a shorter trip, all else equal.

Estimation results can be found in Table 5. All variables are binary indicators, with a 1 day bottom trawl trip to southern New England/Mid-Atlantic forming the baseline for both model specifications. The majority of the parameter estimates from the full model specification are significant at the .01 level. The exceptions are the indicator for a pot trip, the Area 521 and Gulf of Maine indicators, and the estimate for the shape parameter,  $\kappa$ .<sup>3</sup> The Wald test for the null hypothesis that  $\kappa = 0$  is rejected at the .1 significance level (p-value = 0.083). This test suggests that the gamma distribution fits the model better than the log-normal. An additional Wald test was conducted to test the null hypothesis that  $\kappa = 1$ , which was also rejected at the .1 significance level (p-value = 0.0000).<sup>4</sup>

Wald tests were performed on the full model specification in order to better understand the parameter estimates. The null hypothesis that the parameter estimates associated with the Area 521, Georges Bank, and Gulf of Maine were equivalent could not be rejected at any conventional level (p-value = 0.8192). Likewise, the null hypothesis of equality between the parameter estimates for 4 – 6 day trips (p-value = 0.9964), 9 – 10 day trips (p-value = 0.7691), 13 – 14 day trips (p-value = 0.8942), and 11 – 14 day trips (p-value = 0.6484) could not be rejected at conventional levels. The variables were therefore collapsed in the parsimonious specification, in order to facilitate the further investigation of the distances between observed hauls and self-reported fishing location. Further investigation is particularly important given that the marginal effects in the gamma distribution are not linear, and the magnitude and sign of parameter estimates do not lend themselves readily to interpretation.

The Cox-Snell residuals are used to assess goodness of fit for duration models. These residuals are defined as follows:

$$(6) \quad \hat{r}_j = -\ln(\hat{S}_j(d_j)),$$

in which  $\hat{S}_j(d_j) = 1 - \hat{F}_j(d_j)$  is the estimated survival function and is calculated from the estimated model and observed trip characteristics. Cox and Snell (1968) determine that a correctly specified model leads to these residuals following an exponential distribution, with a mean of 1. The fit of the model can then be assessed visually by graphing the Nelson-Aalen empirical estimator of the residual's cumulative hazard function against the Cox-Snell residuals. Again following the notation in StataCorp (2011), the Nelson-Aalen estimator is defined as:

$$(7) \quad \hat{H}_j(\hat{r}_j) = \sum_{j|\hat{r}_j < r} \frac{y_j}{n_j},$$

---

<sup>3</sup> Area 521, off the eastern coast of Cape Cod, was pulled out separately because it is treated differently by different stock assessments. For example, winter flounder from 521 is attributed to the southern New England/Mid-Atlantic Bight Stock, while yellowtail flounder from this same area is attributed to the Cape Cod/Gulf of Maine stock. This area was thus delineated as a separate area of interest because of the potential for differing fishing practices from a mix of Southern New England and Gulf of Maine stocks and species.

<sup>4</sup> These assertions are reinforced by the AIC for the two models, which suggests that the generalized gamma distribution fits the data better than the Weibull, exponential, and log-normal distributions. The log-normal scored closest to the generalized gamma distribution, although the AIC difference between the two models was still 80 points.

with  $y_j$  representing a binary variable equal to 1 if a haul occurred and 0 otherwise,  $n_j$  the total number of trips for which a haul could have occurred, and  $\hat{r}_j$  as defined in equation 6. See, for example, Stata’s reference manual for Survival Analysis and Epidemiological Tables (StataCorp 2011) for a more thorough discussion on the topic. A correctly specified model’s residuals will fall along a 45° line on the graph.

Figure 1 graphs the Nelson-Aalen estimator versus the Cox-Snell residuals, along with a histogram of the Nelson-Aalen results. Although there is deviation within the right tail of the distribution, the vast majority of the estimates fall very near the 45° line. The substantial length of the tail, as illustrated in Figure 2, is likely symptomatic of problems in parametrically fitting the longest distances between observed hauls and reported fishing location. Regardless, 95% of the observations diverge from the 45° line by no more than 3%. Overall, the tight alignment of the majority of the points along the 45° suggests that the model is correctly specified.

Figure 2 compares the modeled survival function (1-CDF) versus the empirical distribution of the 10% of hauls reserved for out-of-sample validation of the model. The modeled survival function was estimated by using each variable’s mean value as calculated from the in-sample observations. As Figure 2 illustrates, the predicted outcome very closely matches the empirical distribution, aside from some deviation indicating differences in density in the 30 – 60 nautical mile range.

In order to better understand the correlation between trip characteristics and spatial precision, confidence intervals conditioned on observable characteristics of trips can be generated. Substituting equations 3 and 4 into 5 and solving for the distance  $d$  leads to the following:

$$(7) \ d = \exp\left(\frac{\ln\left(\frac{u}{\gamma}\right)\sigma}{\kappa} + X\beta\right).$$

An inverse gamma function can be used to estimate the standardized distance,  $u$ , for each probability band of interest. The conditional distance can then be calculated from equation 7. For the purposes of this paper, the distance defining the 25, 50, 75, 90, and 95 percentiles were calculated for each combination of gear, area, and trip length. The results of these estimates, for observed combinations, are presented in Figures 3 – 8. The trip length and gear effects dominate the distance associated with any particular confidence level. This is intuitive, given the theoretical reasoning that longer trips submitting a single VTR are likely to be less precise in location than shorter trips, and that mobile gear is likely to cover more ground than static gear.

In essence, the distance estimated can be interpreted as a radius of a circle centered around the self-reported fishing location within which there is a certain confidence of all a trip’s hauls falling. As an example, a 1 day trip employing scallop dredge in the Mid-Atlantic has a 25% confidence interval extending .43 nautical miles from the self-reported centroid of the circle. This means that on average we would expect 25% of a 1 day scallop dredge trip’s hauls to fall within .43 nautical miles of a self-reported fishing location. Looking again at Figures 3 – 8, it becomes evident that the ability of a 10 minute square to effectively represent the spatial footprint of a fishing trip depends greatly on the length and type of gear employed on the trip in question. For example, a 1 day scallop dredge trip in the Mid-Atlantic has a 90% confidence interval extending 5.87 nautical miles from a VTR point, a distance for which a 10 min square might be a realistic representation of effort for that trip. However, a 2 day bottom trawl trip in the same region has a 90% confidence interval of 19.01 nautical miles, a footprint much less likely

to be represented effectively by a 10 min square. Given the distribution of distances presented in Table 3, the results are not promising for the ability of 10 min squares to reflect fishing locations across all trips of interest.

## **CONCLUSION**

This paper suggests a statistical approach that can be used to ascertain the precision of self-reported VTR fishing location. Modeling results indicate that the gear employed and length of a fishing trip greatly impact the spatial precision of self-reported fishing locations from VTR. These results suggest that more care in the selection of spatial aggregation is likely warranted, and a trip's spatial resolution depends on the gear and type of trip being investigated. Future work will look at the ability of this statistical approach to replicate the distribution of effort on observed hauls, with a comparison to raw VTR points and aggregations to the 10 min square, currently 2 common treatments of this spatial data.

## **ACKNOWLEDGEMENTS**

This study was funded in part by the US Department of the Interior, Bureau of Ocean Energy Management Environmental Studies Program through Interagency Agreement M12PG00028 with the US Department of Commerce, National Oceanic and Atmospheric Administration.

## REFERENCES CITED

- Cox DR, Snell EJ. 1968. A general definition of residuals. *Journal of the Royal Statistical Society, Series B (Methodological)*, 30(2): 248 – 275.
- Greene WH. 2003. *Econometric analysis* (5<sup>th</sup> ed.). Upper Saddle River, NJ: Prentice Hall.
- Kiefer NM. 1988. Economic duration data and hazard functions. *Journal of Economic Literature*, 26(2): 646-679.
- Northeast Regional Office. 2014. Fishing Vessel Trip Report (VTR) Reporting Instructions. Gloucester, MA:NOAA National Northeast Regional Office.  
[http://www.nero.noaa.gov/ro/fso/vtr\\_inst.pdf](http://www.nero.noaa.gov/ro/fso/vtr_inst.pdf) (Accessed May 8, 2014).
- Palmer MC, Wigley SE. 2007. Validating the stock apportionment of commercial fisheries landings using positional data from vessel monitoring systems. U.S. Department of Commerce Northeast Fisheries Science Center Reference Document 07-22, 44 p.
- Palmer MC, Wigley SE. 2009. Using positional data from vessel monitoring systems to validate the logbook-reported area fished and the stock allocation of commercial fisheries landings. *North American Journal of Fisheries Management*, 29(4), p. 928 – 942.
- Records D, Demarest C. 2014. Producing high resolution spatial estimates of fishing effort using a VMS-based statistical model. Mimeo. Woods Hole, MA:NOAA Northeast Fisheries Science Center.
- StataCorp. 2011. *Stata: Release 12. Statistical Software*. College Station, TX: StataCorp LP.
- Van den Berg GJ. 2000. Duration models: Specification, identification, and multiple durations. Munich Personal RePEc Archive Paper No. 9446, 106 p.

**Table 1. Hierarchical matching between Vessel Trip Report and Observer datasets**

| Match Round | Variables matched                           | Hauls Matched |
|-------------|---|---------------|
| 1           | Hull number, date sailed, date landed, area | 351,481       |
| 2           | Hull number, date landed, area              | 22,408        |
| 3           | Hull number, date sailed, area              | 93,189        |
| 4           | Hull number, date sailed, date landed       | 11,599        |
| 5           | Hull number, date landed                    | 6,035         |
| 6           | Hull number, date sailed                    | 3,533         |

**Table 2. Gear representation in dataset**

| Gear           | Observations | Cum.<br>% | Northeast Gear Codes                                  |
|----------------|--------------|-----------|---|
| Bottom Trawl   | 129,150      | 27.7      | 050, 051, 052, 053, 056, 058, 059, 350, 360, 054, 057 |
| Drift Gillnet  | 1,271        | 28.1      | 100, 105, 117   |
| Harpoon        | 2            | 28.1      | 030, 031  |
| Longline       | 4,233        | 29.0      | 010, 020, 021   |
| Midwater Trawl | 3,626        | 29.8      | 170, 370  |
| Other Dredge   | 790          | 27.8      | 381, 386, 400   |
| Pot            | 1,235        | 30.1      | 181, 183, 186, 200, 300                               |
| Purse Seine    | 42           | 30.1      | 120, 121  |
| Scallop Dredge | 290,928      | 92.8      | 132   |
| Sink Gillnet   | 34,462       | 100       | 100, 105, 117   |
| Unknown        | 3            | 100       |   |
| Total          | 465,742      |           |   |

**Table 3. Trip length within dataset**

| Trip Length  | Observations   | Cum. % |
|--------------|----------------|--------|
| 1            | 48,676         | 10.5   |
| 2            | 14,091         | 13.5   |
| 3            | 10,376         | 15.7   |
| 4            | 12,620         | 18.4   |
| 5            | 17,281         | 22.1   |
| 6            | 22,698         | 27     |
| 7            | 31,520         | 33.8   |
| 8            | 41,463         | 42.7   |
| 9            | 47,314         | 52.8   |
| 10           | 47,946         | 63.1   |
| 11           | 42,895         | 72.3   |
| 12           | 32,062         | 79.2   |
| 13           | 28,218         | 85.3   |
| 14           | 23,411         | 90.3   |
| 15           | 24,208         | 95.5   |
| 16           | 11,560         | 98     |
| 17 plus      | 9,403          | 100    |
| <b>Total</b> | <b>465,742</b> |        |

**Table 4. Areas fished within the Northeast Fisheries Science Center (NEFSC) dataset**

| Areas                                   | Observations   | Cum. % | NEFSC Statistical Areas                                   |
|---|----------------|--------|---|
| Stat Area 521                           | 41,765         | 9      | 521   |
| Georges Bank                            | 120,883        | 35     | 522, 525, 542, 543, 561, 562                              |
| Gulf of Maine                           | 34,007         | 42.3   | 511 – 515   |
| Southern New England/Mid-Atlantic Bight | 268,195        | 100    | 526, 534, 537 – 539, 541, 611 – 616, 621 – 629, 631 – 639 |
| <b>Total</b>                            | <b>464,850</b> |        |   |

**Table 5. Regression of distance between observed hauls and self-reported centroid of fishing reported on Vessel Trip Reports, as a function of observed trip characteristics**

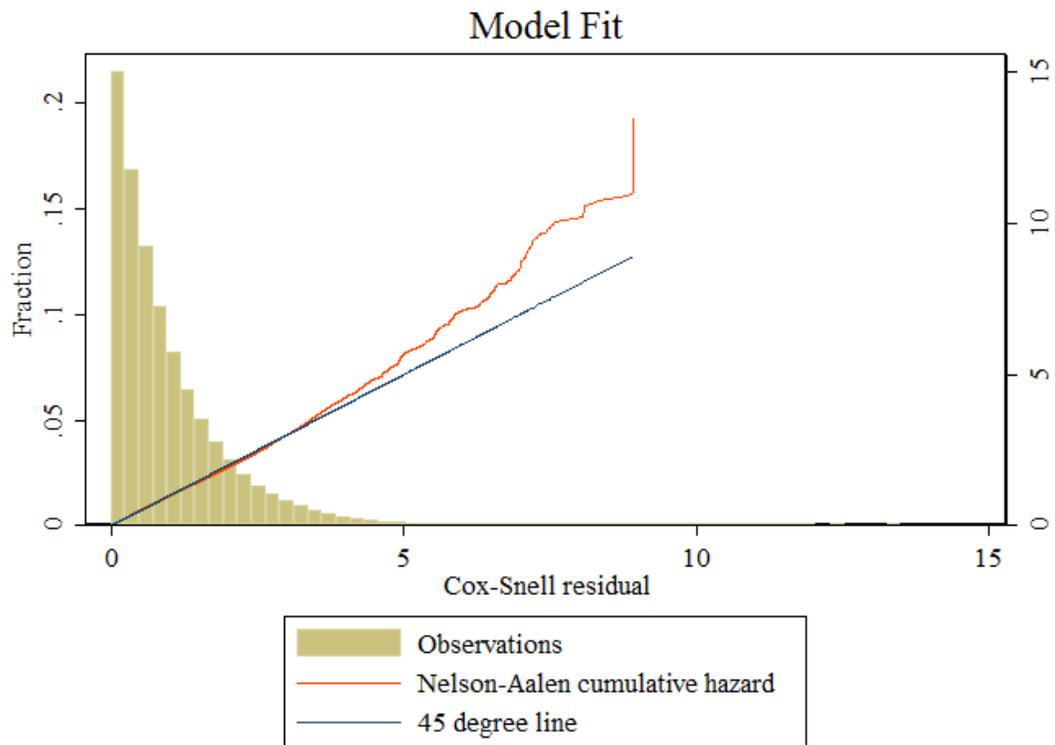
| Variables      | Full                  | Parsimonious          | All Observations      |
|----------------|-----------------------|-----------------------|-----------------------|
| Drift Gillnet  | -0.509***<br>(0.119)  | -0.521***<br>(0.118)  | -0.529***<br>(0.119)  |
| Longline       | -0.658***<br>(0.0964) | -0.663***<br>(0.0950) | -0.666***<br>(0.0941) |
| Midwater Trawl | -0.332***<br>(0.0892) | -0.333***<br>(0.0886) | -0.327***<br>(0.0865) |
| Pot            | 0.263<br>(0.258)      |                       |                       |
| Scallop Dredge | -0.870***<br>(0.0418) | -0.870***<br>(0.0415) | -0.871***<br>(0.0415) |
| Sink Gillnet   | -0.383***<br>(0.0530) | -0.382***<br>(0.0526) | -0.378***<br>(0.0526) |
| 2 day trip     | 0.312***<br>(0.0522)  | 0.304***<br>(0.0519)  | 0.304***<br>(0.0511)  |
| 3 day trip     | 0.723***<br>(0.0571)  | 0.717***<br>(0.0562)  | 0.722***<br>(0.0561)  |
| 4 day trip     | 1.014***<br>(0.0582)  |                       |                       |
| 5 day trip     | 1.015***<br>(0.0628)  |                       |                       |
| 6 day trip     | 1.018***<br>(0.0613)  |                       |                       |
| 4-6 day trip   |                       | 1.008***<br>(0.0495)  | 1.013***<br>(0.0492)  |
| 7 day trip     | 1.139***<br>(0.0646)  |                       |                       |
| 8 day trip     | 1.214***<br>(0.0608)  |                       |                       |
| 7-8 day trip   |                       | 1.171***<br>(0.0522)  | 1.175***<br>(0.0519)  |
| 9 day trip     | 1.372***<br>(0.0653)  |                       |                       |
| 10 day trip    | 1.389***<br>(0.0685)  |                       |                       |
| 9-10 day trip  |                       | 1.370***<br>(0.0569)  | 1.375***<br>(0.0566)  |
| 11 day trip    | 1.581***<br>(0.0706)  |                       |                       |
| 12 day trip    | 1.676***<br>(0.0766)  |                       |                       |
| 13 day trip    | 1.625***<br>(0.0948)  |                       |                       |
| 14 day trip    | 1.610***              |                       |                       |

**Table 6, continued. Regression of distance between observed hauls and self-reported centroid of fishing reported on Vessel Trip Reports, as a function of observed trip characteristics**

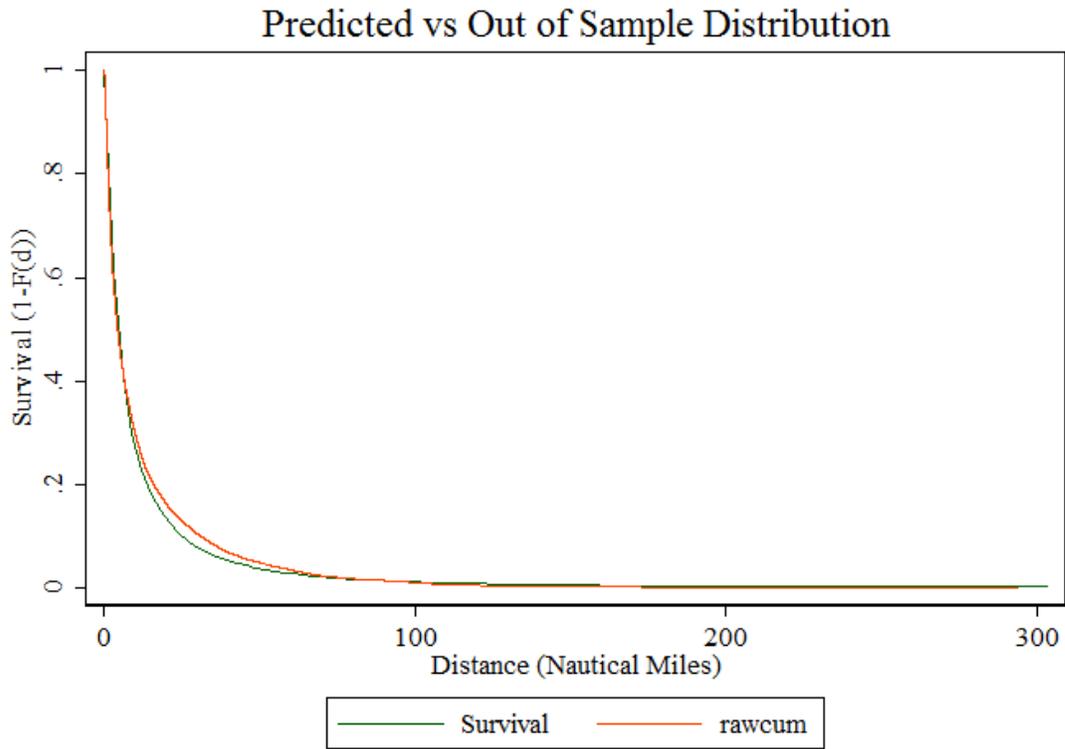
| Variables                | Full                  | Parsimonious          | All Observations      |
|--------------------------|-----------------------|-----------------------|-----------------------|
|                          | (0.103)               |                       |                       |
| 11-14 day trip           |                       | 1.609***<br>(0.0602)  | 1.611***<br>(0.0598)  |
| 15 day trip              | 1.709***<br>(0.115)   |                       |                       |
| 16 day trip              | 1.723***<br>(0.140)   |                       |                       |
| 15-16 day trip           |                       | 1.702***<br>(0.0954)  | 1.704***<br>(0.0951)  |
| 17 + day trip            | 1.898***<br>(0.128)   | 1.887***<br>(0.127)   | 1.891***<br>(0.126)   |
| Area 521                 | -0.127**<br>(0.0539)  |                       |                       |
| Georges Bank             | -0.125***<br>(0.0383) |                       |                       |
| Gulf of Maine            | -0.0964**<br>(0.0479) |                       |                       |
| Non - S. NE/Mid-Atlantic |                       | -0.124***<br>(0.0343) | -0.125***<br>(0.0344) |
| Constant                 | 0.894***<br>(0.0502)  | 0.906***<br>(0.0492)  | 0.902***<br>(0.0488)  |
| ln(sigma)                | 0.286***<br>(0.00624) | 0.287***<br>(0.00625) | 0.287***<br>(0.00625) |
| kappa                    | -0.0368*<br>(0.0214)  | -0.0371*<br>(0.0214)  | -0.0375*<br>(0.0213)  |
| Observations             | 417,535               | 417,535               | 463,943               |

Robust standard errors in parentheses, clustered at the permit level to account for correlation in disturbances across observations of single permit holder's trips

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1



**Figure 1. Histogram of observations and line graph illustrating the Nelson-Aalen empirical cumulative hazard function's divergence from the Cox-Snell residual.**



**Figure 2. Comparison of the model distribution predicted from variable mean values compared to the out of sample empirical distribution of haul distance from Vessel Trip Report centroids.**

Bottom Trawl VTR - Observed Haul distance  
Over Trip Length and Area

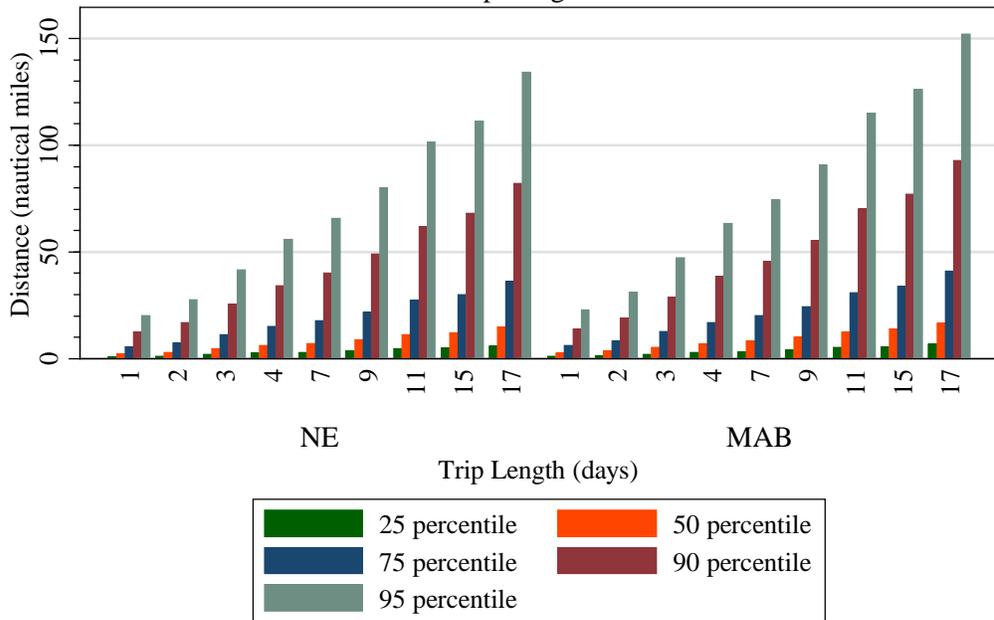


Figure 3. Confidence intervals for bottom trawl hauls, as distance from self-reported Vessel Trip Report fishing location. Confidence intervals conditioned on area and trip length.

Scallop Dredge VTR - Observed Haul distance  
Over Trip Length and Area

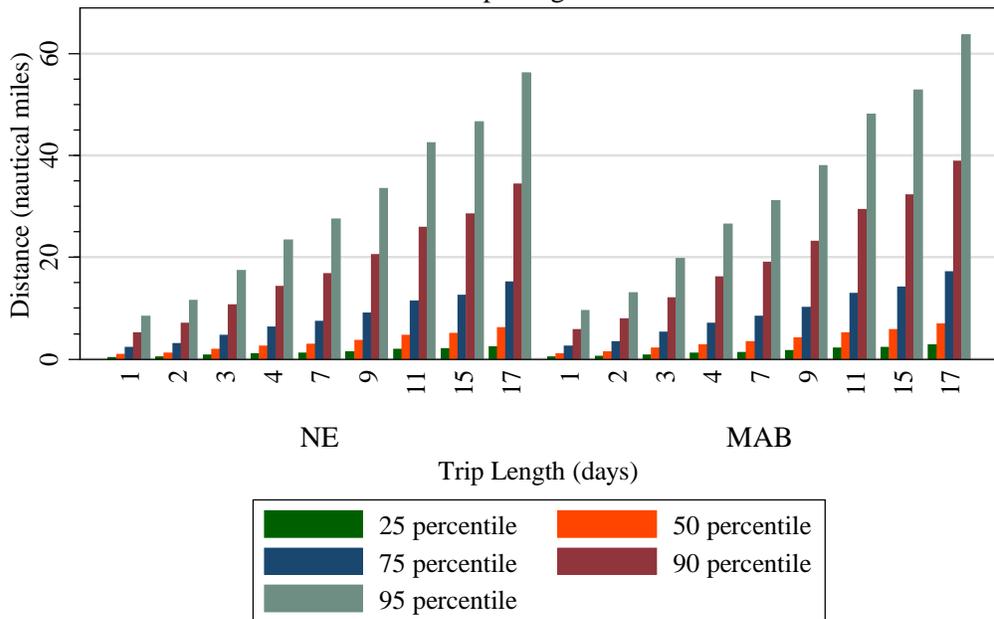


Figure 4. Confidence intervals for scallop dredge hauls, as distance from self-reported Vessel Trip Report fishing location. Confidence intervals conditioned on area and trip length.

### Drift Gillnet VTR - Observed Haul distance Over Trip Length and Area

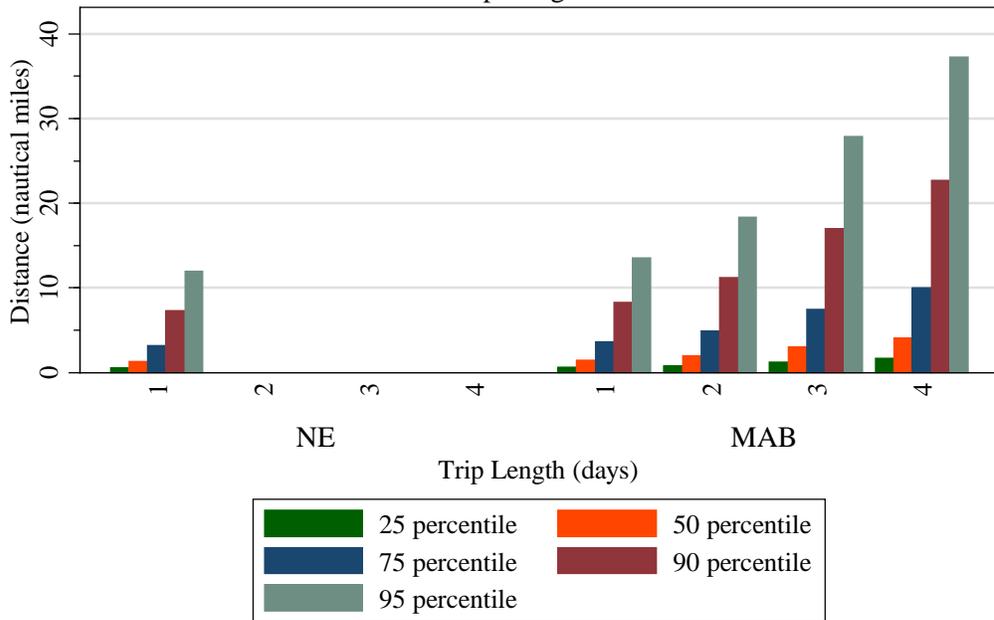


Figure 5. Confidence intervals for drift gillnet hauls, as distance from self-reported Vessel Trip Report fishing location. Confidence intervals conditioned on area and trip length.

### Longline VTR - Observed Haul distance Over Trip Length and Area

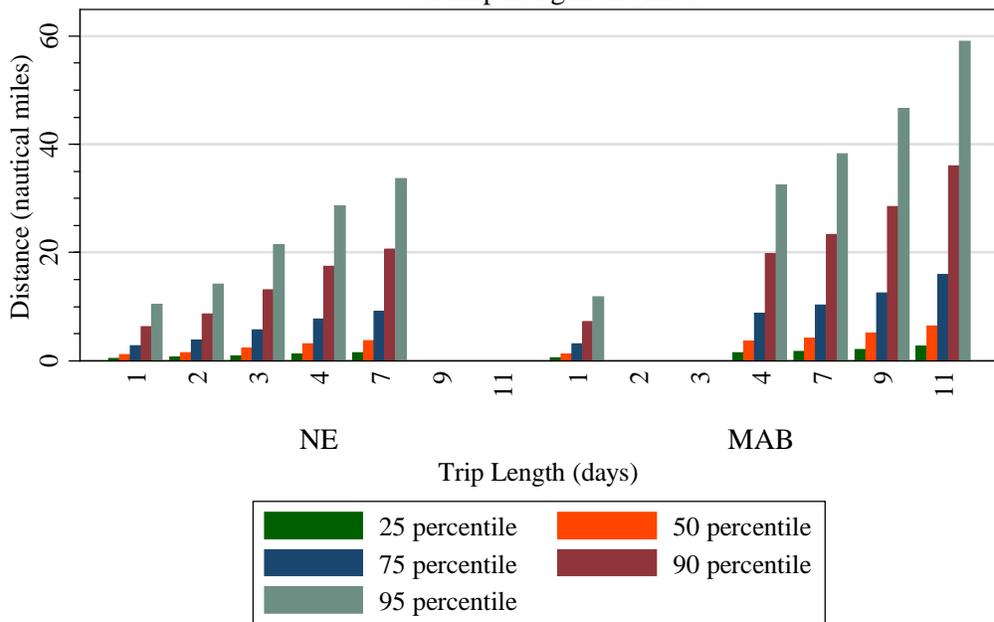


Figure 6. Confidence intervals for longline hauls, as distance from self-reported Vessel Trip Report fishing location. Confidence intervals conditioned on area and trip length.

### Midwater Trawl VTR - Observed Haul distance Over Trip Length and Area

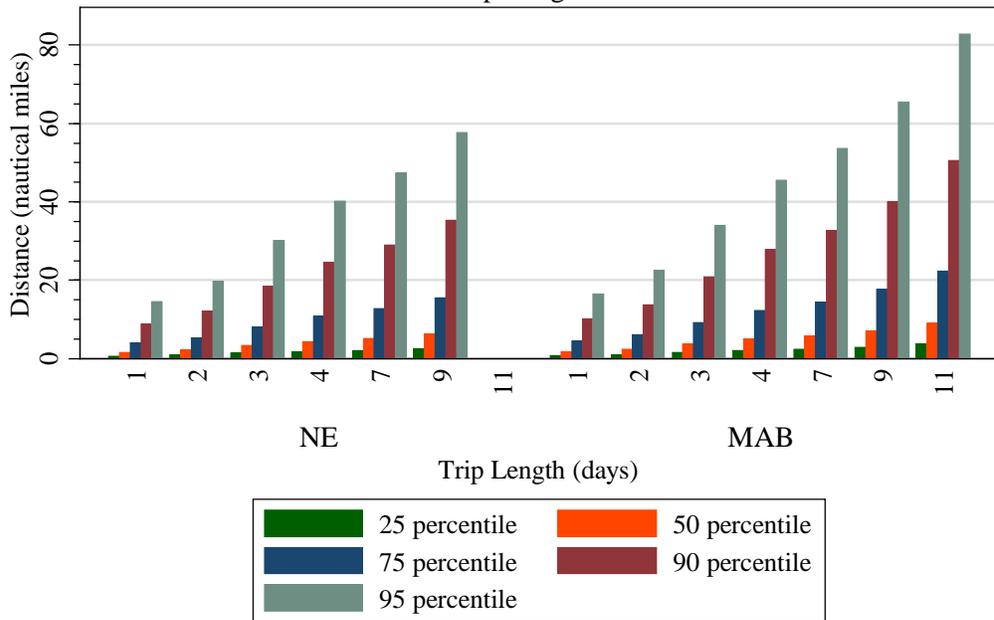


Figure 7. Confidence intervals for midwater trawl hauls, as distance from self-reported Vessel Trip Report fishing location. Confidence intervals conditioned on area and trip length.

### Sink Gillnet VTR - Observed Haul distance Over Trip Length and Area

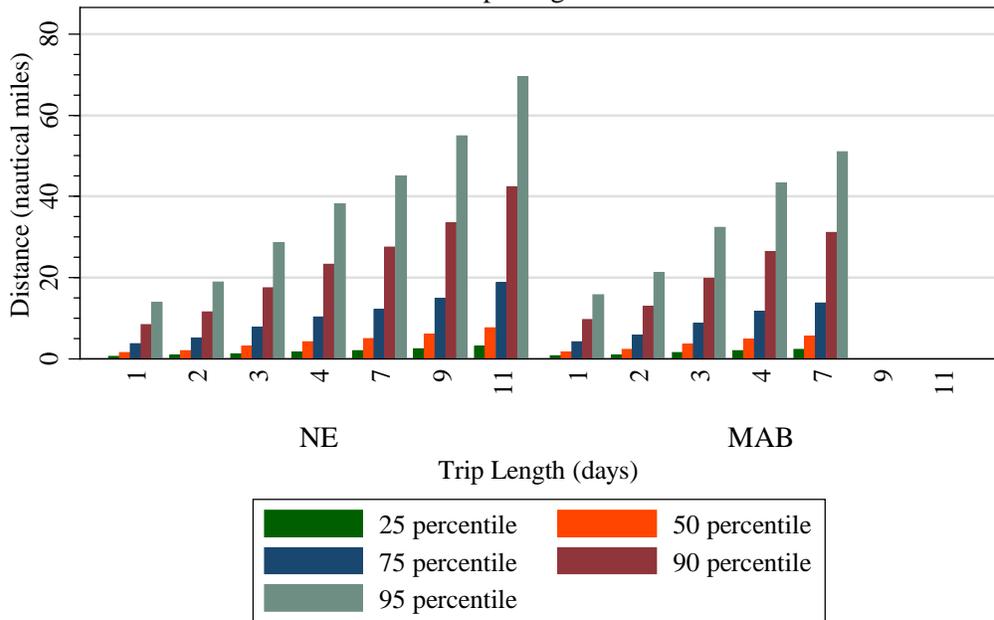


Figure 8. Confidence intervals for Sink Gillnet hauls, as distance from self-reported Vessel Trip Report fishing location. Confidence intervals conditioned on area and trip length.

## Publishing in NOAA Technical Memorandum NMFS-NE

---

### Manuscript Qualification

This series represents a secondary level of scientific publishing in the National Marine Fisheries Service (NMFS). For all issues, the series employs thorough internal scientific review, but not necessarily external scientific review. For most issues, the series employs rigorous technical and copy editing. Manuscripts that may warrant a primary level of scientific publishing should be initially submitted to one of NMFS's primary series (*i.e.*, *Fishery Bulletin*, *NOAA Professional Paper NMFS*, or *Marine Fisheries Review*).

Identical, or fundamentally identical, manuscripts should not be concurrently submitted to this and any other publication series. Manuscripts which have been rejected by any primary series strictly because of geographic or temporal limitations may be submitted to this series.

Manuscripts by Northeast Fisheries Science Center (NEFSC) authors will be published in this series upon approval by the NEFSC's Deputy Science & Research Director. Manuscripts by non-NEFSC authors may be published in this series if: 1) the manuscript serves the NEFSC's mission; 2) the manuscript meets the Deputy Science & Research Director's approval; and 3) the author arranges for the printing and binding funds to be transferred to the NEFSC's Research Communications Branch account from another federal account. For all manuscripts submitted by non-NEFSC authors and published in this series, the NEFSC will disavow all responsibility for the manuscripts' contents; authors must accept such responsibility.

The ethics of scientific research and scientific publishing are a serious matter. All manuscripts submitted to this series are expected to adhere -- at a minimum -- to the ethical guidelines contained in Chapter 2 ("Publication Policies and Practices") of the *Scientific Style and Format: the CSE Manual for Authors, Editors, and Publishers*, seventh edition (Reston VA: Council of Science Editors). Copies of the manual are available at virtually all scientific libraries.

### Manuscript Preparation

**Organization:** Manuscripts must have an abstract, table of contents, and -- if applicable -- lists of tables, figures, and acronyms. As much as possible, use traditional scientific manuscript organization for sections: "Introduction," "Study Area," "Methods & Materials," "Results," "Discussion" and/or "Conclusions," "Acknowledgments," and "References Cited."

**Style:** All NEFSC publication and report series are obligated to conform to the style contained in the most recent

edition of the *United States Government Printing Office Style Manual*. That style manual is silent on many aspects of scientific manuscripts. NEFSC publication and report series rely more on the *CSE Style Manual*, seventh edition.

For in-text citations, use the name-date system. A special effort should be made to ensure that the list of cited works contains all necessary bibliographic information. For abbreviating serial titles in such lists, use the guidance of the International Standards Organization; such guidance is easily accessed through the various Cambridge Scientific Abstracts' serials source lists (see <http://www.public.iastate.edu/~CYBERSTACKS/JAS.htm>). Personal communications must include date of contact and full name and mailing address of source.

For spelling of scientific and common names of fishes, mollusks, and decapod crustaceans from the United States and Canada, use *Special Publications* No. 29 (fishes), 26 (mollusks), and 17 (decapod crustaceans) of the American Fisheries Society (Bethesda MD). For spelling of scientific and common names of marine mammals, use *Special Publication* No. 4 of the Society for Marine Mammalogy (Lawrence KS). For spelling in general, use the most recent edition of *Webster's Third New International Dictionary of the English Language Unabridged* (Springfield MA: G. & C. Merriam).

**Typing text, tables, and figure captions:** Text, tables, and figure captions should be converted to Word. In general, keep text simple (*e.g.*, do not switch fonts and type sizes, do not use hard returns within paragraphs, do not indent except to begin paragraphs). Also, do not use an automatic footnoting function; all notes should be indicated in the text by simple numerical superscripts, and listed together in an "Endnotes" section prior to the "References Cited" section. Especially, do not use a graphics function for embedding tables and figures in text.

Tables should be prepared with a table formatting function. Each figure should be supplied in digital format (preferably GIF or JPG), unless there is no digital file of a given figure. Except under extraordinary circumstances, color will not be used in illustrations.

### Manuscript Submission

Authors must submit separate digital files of the manuscript text, tables, and figures. The manuscript must have cleared NEFSC's online internal review system. Non-NEFSC authors who are not federal employees will be required to sign a "Release of Copyright" form.

Send all materials and address all correspondence to: Jarita A. Davis (Editor), Editorial Office, NMFS Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA 02543-1026.

National Marine Fisheries Service, NOAA  
166 Water St.  
Woods Hole, MA 02543-1026

**MEDIA  
MAIL**

## **Publications and Reports of the Northeast Fisheries Science Center**

The mission of NOAA's National Marine Fisheries Service (NMFS) is "stewardship of living marine resources for the benefit of the nation through their science-based conservation and management and promotion of the health of their environment." As the research arm of the NMFS's Northeast Region, the Northeast Fisheries Science Center (NEFSC) supports the NMFS mission by "conducting ecosystem-based research and assessments of living marine resources, with a focus on the Northeast Shelf, to promote the recovery and long-term sustainability of these resources and to generate social and economic opportunities and benefits from their use." Results of NEFSC research are largely reported in primary scientific media (*e.g.*, anonymously-peer-reviewed scientific journals). However, to assist itself in providing data, information, and advice to its constituents, the NEFSC occasionally releases its results in its own media. Currently, there are three such media:

*NOAA Technical Memorandum NMFS-NE* -- This series is issued irregularly. The series typically includes: data reports of long-term field or lab studies of important species or habitats; synthesis reports for important species or habitats; annual reports of overall assessment or monitoring programs; manuals describing program-wide surveying or experimental techniques; literature surveys of important species or habitat topics; proceedings and collected papers of scientific meetings; and indexed and/or annotated bibliographies. All issues receive internal scientific review and most issues receive technical and copy editing.

*Northeast Fisheries Science Center Reference Document* -- This series is issued irregularly. The series typically includes: data reports on field and lab studies; progress reports on experiments, monitoring, and assessments; background papers for, collected abstracts of, and/or summary reports of scientific meetings; and simple bibliographies. Issues receive internal scientific review, but no technical or copy editing.

*Resource Survey Report* (formerly *Fishermen's Report*) -- This information report is a quick-turnaround report on the distribution and relative abundance of selected living marine resources as derived from each of the NEFSC's periodic research vessel surveys of the Northeast's continental shelf. There is no scientific review, nor any technical or copy editing, of this report.

**OBTAINING A COPY:** To obtain a copy of a *NOAA Technical Memorandum NMFS-NE* or a *Northeast Fisheries Science Center Reference Document*, or to subscribe to the *Resource Survey Report*, either contact the NEFSC Editorial Office (166 Water St., Woods Hole, MA 02543-1026; 508-495-2228) or consult the NEFSC webpage on "Reports and Publications" (<http://www.nefsc.noaa.gov/nefsc/publications/>).

**ANY USE OF TRADE OR BRAND NAMES IN ANY NEFSC PUBLICATION OR REPORT DOES NOT IMPLY ENDORSEMENT.**