Northeast Fisheries Science Center Reference Document 18-12

Visualizing Fishing Data as Rasters

by Sharon Benjamin, Min-Yang Lee, Geret DePiper

December 2018

Northeast Fisheries Science Center Reference Document 18-12

Visualizing Fishing Data as Rasters

by Sharon Benjamin[1], Min-Yang Lee[2], Geret DePiper[2]

[1]Integrated Statistics, 16 Sumner Street, Woods Hole, MA 02543
[2]NOAA Fisheries, Northeast Fisheries Science Center, Woods Hole Laboratory,
166 Water Street, Woods Hole, MA 02543

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Northeast Fisheries Science Center
Woods Hole, Massachusetts

December 2018

Northeast Fisheries Science Center Reference Documents

This series is a secondary scientific series designed to assure the long-term documentation and to enable the timely transmission of research results by Center and/or non-Center researchers, where such results bear upon the research mission of the Center (see the outside back cover for the mission statement). These documents receive internal scientific review, and most receive copy editing. The National Marine Fisheries Service does not endorse any proprietary material, process, or product mentioned in these documents.

To access the electronic version of a document in this series, go to *http://www.nefsc.noaa.gov/publications/*. The electronic version is available in PDF format to permit printing of a paper copy directly from the Internet. If you do not have Internet access, or if a desired document is one of the pre-April 2001 documents available only in the paper version, you can obtain a paper copy by contacting the senior Center author of the desired document. Refer to the title page of the document for the senior Center author's name and mailing address. If there is no Center author, or if there is corporate (*i.e.*, non-individualized) authorship, then contact the Center's Woods Hole Laboratory Library (166 Water St., Woods Hole, MA 02543-1026).

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.

This document may be cited as:

Benjamin S, Lee MY, DePiper G. 2018. Visualizing fishing data as rasters. NEFSC Ref Doc 18-12; 24 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at https://www.nefsc.noaa.gov/publications/

TABLE OF CONTENTS

Abstract....2
Introduction....2
Data and Methods Used to Build the Raster Dataset....3
Data....3
Methods....5
Raster Overview....6
Working with Raster Data....7
Examples of How Rasters Have Been Used....8
Socio-economic Impact Analysis of Offshore Wind Energy Installations....8
Understanding the Seafood Supply Chain....8
Index of Overlap among Large Scallop Ports....9
Management Actions....9
Herring Amendment 8....10
Discussion and Caveats....10
References Cited....12

ABSTRACT

The Social Sciences Branch (SSB) developed a fishing-intensity raster dataset to improve the spatial representation of self-reported Vessel Trip Report (VTR) fishing locations. This derived dataset allows us to describe the spatial footprint of fishing statistically, instead of treating fishing locations as points or as discrete areas of the ocean based on a priori assumptions on spatial precision that have previously been used. In this document, we describe how this dataset was constructed, how it can be linked to other commonly used data, caveats with using and interpreting the processed data, basic directions on how to work with rasters, and we describe applications for this data set that have been used in both research and fishery management support. We also describe the aggregated, processed dataset that has been constructed and is readily available for interested parties, including the regional Fisheries Management Councils, state fisheries agencies, and the general public.

INTRODUCTION

Fishing vessels with Northeast region permits, or those that are eligible to renew a limited access Northeast region permit, are required[1] to fill out Vessel Trip Reports (VTRs) (50 CFR 648.7), through which a vessel operator reports, among other things, a pair of coordinates[2] and the statistical area fished. These data are used extensively in science to support fisheries management following the NOAA mission "To conserve and manage coastal and marine ecosystems and resources." Using VTR data matched to observer data, DePiper (2014) illustrates that the observed location of fishing hauls may occur far from the VTR reported coordinates; these departures vary systematically with trip characteristics that are reported on the VTRs. In this CRD, we describe work built on the statistical relationship in DePiper (2014) used to develop a realistic spatial representation of a fishing trip.

The modifiable areal unit problem (MAUP) is a well-known issue in spatial analysis and quantitative geography: the results of a study are not independent of the choice of spatial unit employed (Gehlke and Biehl 1934; Openshaw 1983; Wong 2009). Therefore, the spatial unit should be chosen with care to match the scale of the underlying data-generating process. Stock assessments are primarily concerned with estimating the yearly biomass of fish in an area. Stock assessment scientists operating under the assumption that fish do not leave the boundaries of those stock regions therefore typically aggregate catch of fish to large stock areas (e.g., NEFSC 2013; NEFSC 2014). Other researchers, particularly those focused on assessing fishing locations and behavior, have aggregated metrics of fishing activity from latitude-longitude at 10- or 30-minute grid cells, or to statistical areas (Clay 1996; Murawski et al. 2005; Marcoul and Weninger 2008; Jin et al. 2013; Lucey and Fogarty 2013). Each of these aggregations carries the assumption that fishing activity is uniformly distributed within the chosen spatial unit. Further, the gridding is assumed to generate an unbiased representation of fishing that cancels out any error in location attribution. Analysis based off of this type of gridding is particularly vulnerable to the MAUP. Other analyses treat each trip as a point (GARFO 2014b) which DePiper (2014)

[1] Vessels that hold *only* an American Lobster (*Homarus americanus*) permit are not required to fill out VTRs. In addition, the Surfclam (*Spisula solidissima*) and Ocean Quahog (*Arctica islandica*) permits have a separate reporting system.

[2] Coordinates are reported in degrees latitude and longitude; when bearings are reported in Loran, they are converted to latitude-longitude.

shows will severely underrepresent the true spatial extent of fishing activity. A poor choice of spatial units could lead to inaccurate fisheries management advice, particularly when spatial management is considered. We attempt to overcome this aspect of the MAUP by using auxiliary data to probabilistically model the spatial extent of a fishing trip, essentially allowing these auxiliary data to determine the appropriate spatial scale of a fishing trip.

In this document, we describe the methods and primary data used to construct a derived dataset of fishing-intensity raster[3] that accounts for the spatial variability reported by DePiper (2014). Unlike traditional gridding approaches, in which point data are attributed in totality to whichever grid cell they happen to fall, our approach assigns probabilities of fishing to each grid based on their distance from a latitude-longitude point and the characteristics of the trip the point represents. We feel that the grid chosen (a 500x500 meter grid) is a sufficient approximation to continuous space within each confidence interval employed, in order to faithfully represent these probabilistically-weighted distances (Figure 1). The spatial unit is thus the trip-level statistically-derived footprint of fishing, as approximated by a fine-meshed grid for computational and memory efficiency. This processed dataset can be linked to any observation in the VESLOGyyyyG table using the GEARID field. Most importantly, this dataset can be used to map any metric that can be linked to the GEARID field, either directly or indirectly. These processed data are available for all GEARIDs in the VESLOG data from 1996-2015, inclusive, for all GEARIDs with non-null latitude and longitude. The dataset also includes data from the clam logbooks from 2003 onwards, using SERIAL_NUM to identify each latitude-longitude point[4]. For readability we will use the term VTR to describe both the VTR dataset, as well as the clam logbook dataset, acknowledging that it might add some confusion.

The derived data have been used to: analyze the socio-economic impact of building offshore wind turbines on the US Atlantic Coast (Section V.a); visualize the full seafood supply chain from fishing areas to location of final use (as food or other products) (Section V.b); build an index of overlap to understand competition among fishing ports in the scallop industry (Section V.c); identify fisheries potentially impacted by habitat management areas (Section V.d); build a "heat map" of where herring was retained by the midwater trawl fishery over a 5 year time period (Section V.e)[5]. In addition to describing this derived dataset we describe how to access, display, and interpret maps produced with this dataset. Just as importantly, we include a section of caveats. Finally, we include some sample maps and analysis steps with the hope of stimulating further use of these data.

DATA AND METHODS USED TO BUILD THE RASTER DATASET

Data

The raster dataset is assembled with the VTR data; the statistical model estimated by DePiper (2014); and spatial data describing closures that were gathered from GARFO's GIS

[3] General information on the raster data format can be found at the ESRI ArcGIS Resources webpage.

[4] From our experience it is impossible to generate unique identifiers for all observations in the clam logbooks prior to 2002, and we thus have excluded these data from the final dataset generated.

[5] Rasters were also used in the following projects: NEFMC Habitat Omnibus Amendment 2, NEFMC Habitat Deep-Sea Coral Amendment, NEFMC Clam Exemption framework, MAFMC Deep Sea Corals Amendment, MAFMC New Jersey Special Management Zones, LAGC-IFQ Report

portal[6], the Federal Register, and the Code of Federal Regulations. We describe each of these in turn.

VTRs are required for nearly all federally managed fisheries in the Northeast Region. VTRs include start and end dates of the trip, identification of the fishing vessel, type of gear used on that trip, and a detailed record of what was kept and discarded on that trip. A single point location and a Greater Atlantic Region statistical area is reported for each trip, theoretically creating a unique combination of gear and statistical area fished during a trip. In addition, fishermen are required to submit a separate logbook page when fishing with different gear and/or in a different stock area, while on the same trip (GARFO 2014a). However, previous research has indicated a downward bias in the number of statistical areas and gears actually reported, meaning individuals tend to report fewer fishing locations than required (Palmer and Wigley 2007, 2009). Furthermore, the single reported fishing location does not account for the likely differences between a trip that takes 1 day or 2 weeks. Similarly, a single point will not differentiate fishing trips using gear that is deployed in very specific locations (as in trap or hand gear) from those trips employing highly mobile gear (e.g., trawl or dredge) which sweep an expansive area.

In an effort to improve on these limitations, DePiper (2014) examines the degree of correspondence between the VTR coordinate with observed hauls on the corresponding trip. This work formed the basis for constructing the raster dataset. DePiper (2014) constructs the great circle (haversine) distance between the VTR coordinate and all observed hauls on that trip. A duration model is then estimated to explain distance from the self-reported VTR to observed fishing locations as a function of VTR characteristics, and finds that gear, trip length, and broad ocean area explain this distance. The model results can be used to construct *c*-confidence intervals, defined as the smallest distance in which we expect to find *c*% of observed hauls around a VTR point can be constructed. DePiper (2014) illustrates these for various métiers as a function of distance in New England and Mid-Atlantic waters. We reprint DePiper's (2014) model results that were used subsequent analysis in Table 1. We carry DePiper's (2014) analysis to the logical next step: mapping all of the VTR data with the results of the statistical model. The c-confidence intervals are calculated by using the following equation:

$$(1)\quad d = \exp\left(\frac{\ln\left(\frac{u}{\gamma}\right)\sigma}{\kappa} + X\beta\right)$$

For each probability band *c*, the standardized distance u can be estimated by using an inverse gamma function, with $\ln(\sigma)$, κ, and $\beta's$ as presented in Table 1, and $\gamma = |\kappa|^{-2}$.

The final dataset that we employed in this process is the historical data describing areas that were closed to fishing. Depending on the fishery permit, gear used, or the time of year, vessels are prohibited from fishing in certain areas of the ocean. From the Federal Register, Code of Federal Regulations, and NEFMC documents, we constructed shapefiles containing those areas and the dates in which they were closed. The spatial data files representing regulatory areas were built from the legal spatial definition of each area, as described in the Code of Federal Regulations. Most closed area regulations can be found in 50 CFR § 648.81 ("NE multispecies closed areas and measures to protect EFH"). See Table 3 for the complete list of referenced

[6] https://www.greateratlantic.fisheries.noaa.gov/educational_resources/gis/data/index.html. Accessed October 30, 2015. We thank Dean Szumylo for providing archived historic closure shapefiles for some of the management regions.

federal regulations. We included regulations related to the year-round closed areas (Closed Area I, Closed Area II, and Nantucket Lightship), Groundfish Rolling Closures, Gulf of Maine Cod Protection Areas, Sector Rolling Closures, and Scallop Rotational Closed Areas.

Methods

All work was conducted in the R programming language (R Core Team 2015). Constructing the raster dataset is a 4-step process that is depicted graphically in Figure 2. In Step 1, we combine VTR and clam logbooks with the results of DePiper (2014). We extract sail date, landing date, latitude, longitude, statistical area, and gearcode from the VESLOGyyyyG and T tables for all GEARIDs. Gearcode is recoded to the broader categories used in DePiper's (2014) original analysis; statistical area is also used to classify a trip as Southern New England/Mid-Atlantic (SNE/MA) or not. Sail date and land date are used to construct trip duration, rounded up to the nearest integer. Negative trip durations were assigned an absolute value of the calculated trip duration, while trips with lengths > 365 days were assigned a trip length of the reported length minus 365 days, on the assumption that both issues result from transcription errors. We construct 4 concentric rings at the 25th, 50th, 75th, and 90th percentiles based on reported VTR data. We then assign 25% of the fishing trip[7] to each of these rings[8]. This generalized approach provides flexibility to map any trip-level metric, such as revenue, pounds landed, or crew-hours. Implicit in this step is the assumption that the results of the statistical model in DePiper (2014) can be used to make out-of-sample predictions.

In Step 2, we eliminate portions of those rings that overlap impossible or illegal fishing locations. Implicit in this step is the assumption that fishing vessels are not catching fish on land or in Canadian waters. We also eliminate portions of rings that overlap an area of the ocean that is closed to fishing because of that vessel's gear, permit, or time of year. Fishing activity often occurs on the borders of closed areas. We assume fishermen abide by the area closures and do not fish where it is prohibited.

In Step 3, we reallocate the trip value proportionally in the remaining fishing footprint, after cropping the closures out of the potential fishing footprint. In Step 4, we discretize fishing activity to a 500x500m raster grid. The resolution of the fishing data is the trip-specific rings; the 500m grid was used to approximate a continuous ring, while capitalizing on a raster's ease and speed of data processing. Each cell in the grid is assigned a fraction (share) of the fishing trip. The resulting raster files are composed of a GRI and GRD file, 1 for each GEARID. They are named GEARID.GRI and GEARID.GRD for every commercial trip that occurred between calendar year 1996 and calendar year 2017; exceptions include surfclam (*Spisula solidissima*) and ocean quahog (*Arctica islandica*) logbooks before 2003. The ".gri" files have sequential binary values, while the ".grd" file is a header file, meaning that it contains georeferencing and other metadata about the raster file[9]. More information about this file format may also be found in Hijmans' (2016) R-code vignette that explains the raster file format.

[7] Technically, these are "sub-trips" when there are multiple GEARIDs per TRIPID.

[8] Moving from the 90 to 95 confidence interval nearly doubles the radius of the circle circumscribed, a result deemed too costly for 5% more coverage.

[9] The individual trip .gri and .grd raster files are currently located at *net://socialsci/Geret_Rasters/Data/individualrasters* and are organized by subfolders for state landed, and then in subfolders by year. An R dataset containing the names and folders of all rasters can be found in *net://socialsci/Geret_Rasters/Data/Rasterfilenames.Rdata*.

Examination of any single trip may not be interesting; however, aggregating a metric of fishing effort (for example, landed quantities in the herring fishery) at a reasonable temporal scale (for example: fishing year 2010) is much more useful to fisheries managers and researchers. After constructing rasters for all valid GEARIDs in the VESLOG database, we can easily select, subset, and aggregate to produce maps of fishing effort for various variables. Table 2 contains a detailed list of readily available maps.

RASTER OVERVIEW

Raster data provide a model of the world defined as a regular set of cells, usually squares in a grid pattern. This data format is used to represent continuous spatial data[10] such as elevation, and the attribute value of each cell represents the value of that variable at a specific location (Bolstad 2008). VTR data are commonly aggregated to 10 minute squares of latitude and longitude, which range from 70-80 square nautical miles and vary because of the curvature of the Earth. In contrast, rasters in this dataset are in a projection that conserves area to minimize the difference in area of cells at different latitudes. The individual trip rasters are aggregated to represent subsets of all reported fishing activity, and saved in a GeoTIFF (.tif) format. The rasters are built in the Geographic Coordinate System North American Datum 1983 by using a customized Albers Conic Equal Area projection, which minimizes the distortion of raster cells over the large study region (spanning most of the US East Coast). Conic projections use 2 standard parallels, defined by degrees latitude; while no projection completely reflects the actual shape of the earth, distortion of area is minimized in the region between the standard parallels. The customized parallels were adjusted from 20 and 60 degrees to 28 and 42, respectively, based on the total extent of the study area.

The aggregated rasters represent estimated fishing locations for the subset of fishing activity summed together, and each cell holds an estimated value. Certain gear category groups may include several gear types that are differentially affected by area closures. It may be a visualization problem (rather than a data error) if a raster indicates fishing activity in a place where the user did not expect any; for this reason, it is important to consider how the data are aggregated when interpreting the plotted raster data. For example, rasters aggregated by the landed species or Fishery Management Plan (FMP) will include any record of those species caught using any gear type. Certain species are generally caught with 1 gear type but may be caught, less frequently, with other gears – for example, lobster pots occasionally catch groundfish species like Atlantic cod (*Gadus morhua*).

It is impossible to extract information about any particular individual fishing trip from an aggregated raster. However, it is important to take into consideration the number of fishing vessels, unique permit holders, and unique dealers that were involved in any subset of fishing trips that are aggregated from the raw raster dataset. When rasters are aggregated along multiple parameters there is a risk that a very small group's fishing trips will be represented, and thus break the "rule of three" and will indicate fishing locations from identifiable vessels. For example, aggregating a raster from trips by vessels of 1 size category, from a single port-group, targeting a single species could result in a very small subset of trips, and risk revealing personally identifiable information (PII). Code has been developed to ensure adherence to the

[10] See ArcGIS help page on discrete and continuous data for more information on this data format ("Discrete and Continuous Data" (2016)).

rule of three for discrete areas of the ocean of management interest, and this code can be made available to interested parties.

Working with Raster Data

GeoTIFF raster files should only be opened in spatial analysis software such as ESRI ArcGIS[11] or QGIS[12] or by using the R programming language[13]. It is inappropriate to attempt to view or analyze this type of file in programs that are not meant for spatial analysis, such as Microsoft Window's "Preview" or Windows Explorer.

Raster data can be plotted where cell values are rendered with a "stretched" or "classified" visualization. A stretched rendering applies a "color ramp" or range of colors over the continuous cell values; ArcGIS for Desktop includes several histogram stretch methods including Minimum-Maximum, Percent Clip, and Standard Deviation, which is the default. Colors can be defined numerically in several ways; a common method is under the "additive RGB model" which simply refers to the mixing of red, green, and blue light in equal intensities creating white light.

The "Minimum-Maximum" method stretches the raster cell values to the color value range. The Standard Deviation method determines the minimum and maximum color ramp values based on the standard deviation of the raster's values; similarly the percent clip applies the stretch between the minimum and maximum defined by the percentage "clipped" from raster values. The advantage of using the histogram stretch method is that it allows the user to better show variation in a continuous dataset. The disadvantage of this method is that various histogram stretch methods each highlight particular aspects of the data, resulting in wildly different-looking renderings of the same dataset. The other "classification" visualization method categorizes cells based on value ranges, and assigns colors to those value classes or "bins." Quantile classification in particular is a method in which each value class has the same number of cells; Brewer and Pickles (2002) advocate this method for general use, based on a study of undergraduate students interpreting epidemiological data.

The visualization method should be chosen with care; each method has advantages and disadvantages depending on what the researcher is looking to communicate. Figure 3 shows 3 different ways to visualize the same data: a raster of fishing activity associated with trip landings during 2011 in the port of New Bedford, MA. On the left, the raster is visualized with the "Minimum-Maximum" histogram stretch, so that the black-and-white color ramp used in this map has been matched to the raster cell values range. The lowest value rendered on this map is the smallest value greater than zero (0.002187) and is matched to the color white. The darkest color (black) is matched to the largest raster cell value (33,603.74). For comparison, the center panel includes the same data, where the raster cell values are classified into 5 classes. In the third panel, the raster cell values are classified into similar groups, so as to maximize differences between groups' values. In addition, to further emphasize the importance of "knowing one's data," the center and right panels both exclude values less than 10, while the left panel does not assign a color to represent cells with a value of 0. There is a clear difference in the perceived footprint of activity in these panels, depending on whether the lowest-value cells are included.

The different rendering methods emphasize different aspects of the data; the first panel gives a sense of how heavily skewed the data are towards low-value cells. The center panel most

[11] More information on ESRI ArcGIS at http://www.esri.com/software/arcgis/arcgis-for-desktop
[12] More information on QGIS at http://www.qgis.org/en/site/index.html
[13] More information on R at https://www.r-project.org/

clearly shows a "heat map" style rendering of fishing intensity, although their values-classification creates an enormous range for the highest-value bin (ranging from cells of about $1,800 to $33,600 value). The right panel's classification method reflects natural subsets within the data, and shows that the majority of cells are of lower value, surrounding small "hotspots" of fishing activity. This figure is meant to illustrate the importance of carefully examining raster data and understanding its value range in order to choose the best method to visualize those data in a map.

When raster data are loaded into ArcMap, the raster is plotted by using the default rendering settings where the cell values are "stretched" along a black and white (low/high value) color ramp[14]. When working with R, the default renderer setting will depend on which data visualization package is used.

EXAMPLES OF HOW RASTERS HAVE BEEN USED

This section provides brief descriptions of several ways in which rasters of fishing activity have been used in research at the Social Sciences Branch, in the hopes of providing inspiration and motivation for expanding this work by researchers both within, and external to, the Northeast Fisheries Science Center.

Socio-economic Impact Analysis of Offshore Wind Energy Installations

The Social Sciences Branch (SSB) developed the first iteration of the fishing-intensity raster dataset as part of a socio-economic impact analysis of building wind turbine facilities in offshore areas, commissioned by the Bureau of Ocean Energy Management (BOEM). The fishing-intensity rasters were aggregated by port-group and were summed to indicate the revenue-intensity of fishing activity to inform a fishing location-choice model (Kirkpatrick et al. 2017). Revenue-intensity rasters for the years 2007-2012 were aggregated by several variables (FMP, Gear, Port, Species, State) and published on the BOEM website (https://www.boem.gov/Renewable-Energy-GIS-Data/), as well as the Marine Cadastre website (www.marinecadastre.gov), which is collaboratively managed by BOEM and NOAA's Office of Coastal Management. The revenue-intensity raster dataset has been improved and expanded to span the years 1996-2015. In particular, rolling groundfish closures in effect from 1996 through 2002 were used to refine where fishing activity was actually occurring (see Methods section for details on building historic regulatory area spatial data files).

Understanding the Seafood Supply Chain

The raster dataset provides a powerful method to intuitively visualize fishing behavior by creating intensity "heat maps" of activity at sea. Fishing-intensity rasters illustrate the differences in major fishing ground locations among different targeted species. Rasters were used to demonstrate the inextricable link between the marine resources at sea and the onshore seafood infrastructure in a study examining fisheries supply chains. Stoll et al. (2015) illustrated the full supply chain for the herring (*Clupea harengus*) and dogfish (*Squalus acanthias*) fisheries, starting from where both species are caught at sea, and following the catch to where it is landed,

[14] Additional background on ArcGIS raster data rendering options can be found at this ArcGIS for Desktop Help Page.

processed, and eventually reaches its final use as bait or human food (Figures 4A & B). The rasters provide valuable context in understanding the complete geography of where these 2 fisheries occur at sea.

Index of Overlap among Large Scallop Ports

Shared fishing grounds are a mechanism through which distant ports may be closely related. Benjamin et al. (2016) use the Czekanowski (Cz) index (1909) to quantify the overlap in fishing activity across ports[15]. We first construct revenue-intensity rasters that represent scallop fishing activity landed in all ports for each fishing year from 1996-2015. We convert these rasters to port-year "shares" by dividing each cell in the raster by the total fishing revenue for that port during that year. The resulting "share" raster cells represent the proportion of fishing revenue value from that port-year in that specific location in the ocean.

Share rasters for individual ports are compared to each other by year; for example, the share raster for New Bedford fishing activity in 2000 would be compared to the share raster of fishing activity in Atlantic City for the same year. The comparison of port pairs is the foundation for constructing the Cz index (1909) for each fishing year:

(2) $Cz_{AB} = \sum_{j}^{N} \min(s_j^A, s_j^B)$

$\min(\boldsymbol{s_j^A}, \boldsymbol{s_j^B})$ refers to the smaller share in patch *j*, only comparing port A's and port B's shares. The Cz index measures "overlap" or "similarity," where $0 \leq Cz \leq 1$; 0 indicates no overlap, and 1 indicates exact overlap.

The variation in Cz value for port pairs is particularly interesting when examined in the regional context. For example, in examining overlap between Mid-Atlantic ports and Newport News, VA, it is clear the ports tend to follow similar trends in terms of how much these ports overlap in activity (Figure 5). The 5 ports included have overlap-index values that reflect their proximity to Newport News; Hampton, an adjacent port, has the overall greatest overlap index value with Newport News' port activity, while the study port (Point Pleasant, NJ) at the greatest distance generally maintains the lowest overall level of overlap with Newport News. In contrast, New England ports paired with Newport News, VA, show somewhat different trends in overlap (Figure 6). There is a big jump in overlap from 2004 to 2005, while there is a less clear trend in those years for the Mid-Atlantic ports paired with Hampton. These jumps are associated with changes in available biomass in several key scallop areas, both in terms of the volume of harvestable scallops, and in terms of the regulatory closure schedule. These plots highlight the connections among ports that may not immediately or obviously be connected.

Management Actions

The rasters detailed in this reference document have been used in a number of management actions in both New England and the Mid-Atlantic. Examples include the NEFMC's Habitat Omnibus Amendment 2; Habitat Deep-Sea Coral Amendment; Habitat Clam Framework; Ecosystem Based Fishery Management Fishery Ecosystem Plan; the MAFMC's Deep Sea Corals Amendment; New Jersey Special Management Zones; and Herring Amendment 8 (described below). Typically, the aggregated raster maps are used to assess where patterns of

[15] This index was also used by Bray and Curtis (1957) and Finger and Kreinin (1979) in the ecological and economics fields, respectively.

fishing occur around areas being assessed for management purposes, while the individual GEARID.GRI and GEARID.GRD files are used to assess the numbers of individuals exposed by management alternatives and the degree of impact expected. The sum of the individual GEARID rasters within an area of interest provides a probability weighted fraction of a trip falling within that area. This percentage can then be used to assess the probability weighted revenue, effort, or other metric generated from or expected within an area of interest at the trip and individual level. This approach explicitly accounts for the uncertainty in location associated with VTR points, as opposed to using the raw point data.

Herring Amendment 8

In 2015, the New England Fisheries Management Council (NEFMC) was interested in management measures to address localized depletion of herring (80 Federal Register 50825)[16]. The Council was particularly interested in understanding fishing effort by the midwater-trawl gear in areas relatively close to shore and at spatial scales smaller than a statistical area. Providing information to the Council according to latitude-longitudes could be misleading because these vessels move extensively while fishing. Information that was gridded to the 30 minute square could also be misleading because of the MAUP issue described earlier. Therefore, the Council (through the Herring Plan Development Team) was provided with raster maps that illustrated where the midwater-trawl fleet is active. Figure 6 illustrates fishing activity for the years 2010-2014, aggregated to the months of February, June, and October. These panels illustrate the change in fishing location with the change in seasons.

Later in the process of developing fishery management alternatives, the Herring Advisory Panel recommended closure to all gear of the portion of 30 minute square #114 that is within 6 miles from shore. The raster data were used to compute the share of monthly landings inside this region relative to the total landings for the fishing year. These data suggest that there would be relatively small impacts on the herring fishery from this closure during the part of summer the fishery is using other parts of the ocean in addition to the near-shore area east of Cape Cod, MA.

The Herring Committee was also interested in the effects of closures to midwater -trawl gear within 12, 25, and 50 nautical miles of shore within Herring Management Areas 1B, 2, and 3. The raster data were again used to construct this fleet's share of landings within those areas, relative to the activity of that fleet within Herring Management Areas 1B, 2, and 3, averaged over the 2011-2015 time period. The midwater-trawl fleet mostly uses the inshore regions during the winter (November-February), in contrast to minimal use of inshore regions during the summer (May-August). The share of landings does not vary much by distance from shore during certain months (May-August and November-January), in contrast to the months of February-April and September, when the share of landings does vary by distance from shore.

DISCUSSION AND CAVEATS

We wrote this document to provide context and information to the NEFSC and external researchers on the work we have completed towards improving use and access to fishing data. The VTR raster dataset will allow researchers to explore new questions about the New England

[16] Fisheries of the Northeastern United States; Atlantic Herring Fishery; Supplemental Notice of Intent to Prepare an Environmental Impact Statement; Scoping Process; Request for Comments, 80 Fed. Reg. 50825 (August 21, 2015). Available from https://www.federalregister.gov/d/2015-20798

and Mid-Atlantic fisheries. There are potential new avenues to explore, such as examining interactions at sea among vessels, species, and gear types; studying changes in fishing location over time; and visualizing changes in fishing behavior among fishing ports.

In this section we list some important assumptions and caveats for raster data. Please note that this section only deals with caveats stemming from the data processing outlined in this paper and not general caveats for VTR data quality controls or the low level of lobster trip coverage.

- The manner in which existing management areas and Canadian waters are used to buffer trips relies on the assumption that fishermen obey current fishing regulations. This assumption might prove problematic for certain research and management actions. This approach also fails to consider the Special Access Programs (SAP) for groundfish which have historically allowed seasonal fishing for groundfish within the northern Hook Gear Haddock SAP Area of Closed Area I, the Yellowtail Flounder-Haddock SAP, and Eastern US-Canada Haddock SAP Areas of Closed Area II.

- Although the individual GEARID rasters are buffered in a manner that reflects scallop rotational access areas, in reality fishing during an access trip likely occurs fully within that access area and nowhere else because of the higher productivity within the access areas and time constraint associated with these trips. Thus, for these access trips the rasters likely provide a more diffuse footprint than what actually occurs at sea.

- The relatively low historical observer coverage for clam dredges necessitated the lumping of this gear in with scallop dredges, which fish in a very different manner, although it is unclear how different these two fishing gears are in terms of reporting precision.

- Certain species are caught by positioning and moving gear along bathymetric contours, providing a directionality to fishing which is not reflected in the rasters.

REFERENCES CITED

Benjamin S, Lee M, DePiper G. 2016. Examining changes in connection and competition among ports through shared fishing grounds. Presented at: Challenges of Natural Resource Economics and Policy, 5th National Forum on Socioeconomic Research in Coastal Systems; New Orleans, LA. http://www.cnrep.lsu.edu/2016/index.htm

Bolstad P. 2008. GIS fundamentals: A first text on Geographic Information Systems. 3rd ed. White Bear Lake (MN): Eider Press.

Bray JR, Curtis JT. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecol Monogr. 27:325–349.

Brewer, CA, Pickle L. 2002. Evaluation of methods for classifying epidemiological data on choropleth maps in series. Ann Am Assoc Geogr. 92(4):662-681.

Clay PM. 1996. Management regions, statistical areas and fishing grounds: Criteria for dividing up the sea. J Northw Atl Fish Sci. 19:103-126.

Czekanowski J. 1909. Zur differential Diagnose der Neandertalgruppe. Korrespondenzblatt der deutschen Gesellschaft für Anthropologie, Ethnologie und Urgeschichte. 40:44–47.

DePiper G. 2014. Statistically assessing the precision of self-reported VTR fishing locations. NOAA Technical Memorandum NOAA Tech Memo NMFS-NE-229. National Marine Fisheries Service. Available from http://www.nefsc.noaa.gov/publications/tm/tm229/

Discrete and Continuous Data. [Internet]. 2016. Environmental System Research Institute, Inc.: ArcGIS for Desktop. Available from http://desktop.arcgis.com/en/arcmap/latest/manage-data/raster-and-images/discrete-and-continuous-data.htm.

Finger J, Kreinin ME. 1979. A measure of 'export similarity' and its possible uses. Econ J (London). 89(356):905–912.

Gehlke CE, Biehl K. 1934. Certain effects of grouping upon the size of the correlation coefficient in census tract material. Proceedings of the American Statistical Journal. J Am Stat Assoc. 29(185):169-70.

Greater Atlantic Regional Fisheries Office (GARFO). 2014a. Fishing Vessel Trip Report (VTR) Reporting Instructions. Gloucester, MA:NOAA National Northeast Regional Office. https://www.greateratlantic.fisheries.noaa.gov/aps/evtr/ (Accessed Jan 3, 2017.)

Greater Atlantic Regional Fisheries Office (GARFO). 2014b. Special Management Zone (SMZ) Designation for Artificial Reef Sites in the EEZ: A Final Environmental Assessment Including an Initial Regulatory Flexibility Analysis. National Marine Fisheries Service, Gloucester, MA. https://www.greateratlantic.fisheries.noaa.gov/regs/2015/June/15desmzea.pdf

Hijmans RJ. Raster: Geographic Data Analysis and Modeling. R package version 2.5-2. [Internet]. Available from http://CRAN.R-project.org/package=raster.

Jin D, Hoagland P, Wikgren B. 2013. An empirical analysis of the economic value of ocean space associated with commercial fishing. Mar Policy. 42:74-84.

Kirkpatrick AJ, Benjamin S, DePiper G, Murphy T, Steinback S, Demarest C. 2017. Socio-economic impact of outer continental shelf wind energy development on fishing in the U.S. Atlantic. U.S. Dept. of the Interior, Bureau of Ocean Energy Management, Atlantic OCS Region, Washington, D.C . OCS Study BOEM 2017-012.

Lucey, SM, Fogarty MJ. 2013. Operational fisheries in New England: Linking current fishing patterns to proposed ecological production units. Fish Res. 141:3-12.

Marcoul P, Weninger Q. 2008. Search and active learning with correlated information: empirical evidence from mid-Atlantic clam fishermen. J Econ Dyn Control. 32:1921-1948.

Murawski SA, Wigley SE, Fogarty MJ, Rago, Mountain DG. 2005. Effort distribution and catch patterns adjacent to temperate MPAs. ICES J Mar Sci 62: 1150-67.

NE multispecies closed areas and measures to protect EFH, 50 CFR § 648.81 (2017) Available from http://www.ecfr.gov/cgi-bin/text-idx?SID=3e1929cbd991e771bf9516bb6d00e41c&mc=true&node=se50.12.648_181&rgn=div8

Northeast Fisheries Science Center. 2013. 55th Northeast Regional Stock Assessment Workshop (55th SAW) Assessment Summary Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 13-01; 41 p.

Northeast Fisheries Science Center. 2014. 59th Northeast Regional Stock Assessment Workshop (59th SAW) Assessment Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 14-09; 782 p.

Openshaw S. The modifiable areal unit problem. Concepts and Techniques in Modern Geography No. 38. Geobooks, Norwich, England.

Palmer MC, Wigley SE. 2007. Validating the stock apportionment of commercial fisheries landings using positional data from Vessel Monitoring Systems (VMS). US Dept Commer, Northeast Fish Sci Cent Ref Doc. 07-22; 35 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): 928-42.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Stoll JS, Pinto da Silva P, Olson J, Benjamin S. 2015. Expanding the ‘geography’ of resilience in fisheries by bringing focus to seafood distribution systems. Ocean Coast Manage. 116:185-192.

Wong D. 2009. The Modifiable Areal Unit Problem (MAUP). In: Fotheringham/Rogerson (Eds.) The Sage Handbook of Spatial Analysis. Thousand Oaks (CA) Sage Publication.

Table 1. Three parameter gamma duration model of distance between self-reported Vessel Trip Report points and observed fishing hauls.

Variables	Parameters
Drift Gillnet	-0.5208564***
	(0.1183027)
Longline	-0.6629677***
	(0.0950136)
Midwater Trawl	-0.3332272***
	(0.0885561)
Scallop Dredge	-0. 8704655***
	(0.0415119)
Sink Gillnet	-0.3816576***
	(0.0526281)
2 day trip	0.3041691***
	(0.0519063)
3 day trip	0.7167141***
	(0.056231)
4-6 day trip	1.00832***
	(0.0495002)
7-8 day trip	1.170695***
	(0.0522018)
9-10 day trip	1.370113***
	(0.0568946)
11-14 day trip	1.60882
	(0.0602231)
15-16 day trip	1.702254
	(0.0953657)
17 + day trip	1.886611
	(0.1266852)
Non - S. NE/Mid-Atlantic	-0.1237827
	(0.0343461)
Constant	0.9056303
	(0.0491966)
ln(sigma)	0.286685
	(0.0062503)
kappa	-0.037053
	(0.0213887)
Observations	417535

Table 2. Aggregated raster maps listed here are available in the folder *Y:/socialsci/Geret_Rasters/FINAL MAPS.*

Metric	**Aggregation level**	**Years Mapped**	Include:
Revenue	FMP	1996 – 2015, Calendar	Bluefish Surfclam and ocean quahogs Herring Large mesh multispecies Mackerel-squid-butterfish Monkfish Atlantic deep sea red crab River herring Sea scallop Skate (spp) Small mesh multispecies Spiny dogfish Summer flounder-scup-BSB Unmanaged
Revenue	Gear Type	1996 – 2015, Calendar	Bottom longline Bottom trawl Clam dredge Drift gillnet Hand gear Lobster pot Midwater trawl Other dredge Other gear Other gillnet Other pot Pelagic longline Purse seine SAP trawl Scallop dredge Scallop trawl Shrimp trawl Sink gillnet
Revenue	All annual	1996 – 2015, Calendar	
Revenue	Individual Species	1996 – 2015, Fishing year	Atlantic cod Haddock Yellowtail flounder Pollock Plaice Witch flounder White hake Winter flounder Skates (spp) Goosefish Herring

Table 3. List of the rolling closure areas used to clean the raster data of fishing activity estimated in prohibited areas. The Code of Federal Regulations and the Federal Register were used to build a geodatabase of historic rolling closure areas.

Area Name	CFR - Paragraph	FR Citation	FR Date	Effective Date	Source	Start Date	End Date^
GOM Inshore Closure Area I	648.81(g)(1)(i)	63 FR 15331	31-Mar-98	1-May-98	FW 25	1-Mar	31-Mar
GOM Inshore Closure Area II	648.81(g)(1)(ii)	63 FR 15331	03/31/98	1-May-98	FW 25	1-Apr	30-Apr
GOM Inshore Closure Area III	648.81(g)(1)(iii)	63 FR 15331	03/31/98	1-May-98	FW 25	1-May	31-May
GOM Inshore Closure Area IV	648.81(g)(1)(iv)	63 FR 15331	03/31/98	1-May-98	FW 25	1-Jun	30-Jun
GOM Rolling Closure Area V	648.81(g)(1)(v)	64 FR 24075	05/05/99	1-May-99	FW 27	1-Oct	30-Nov
GOM Rolling Closure Area VI	648.81(g)(1)(vi)	65 FR 379	01/05/00	31-Jan-00	FW 31	1-Feb	28-Feb
Closed Area I Access Area*	648.58(b)(3)	64 FR 37914	06/19/00	06/15/00	Scallop FW 13	6/19/00	
Georges Bank Access Area	648.58(b)	64 FR 31150	06/10/99	06/15/99	Scallop FW 11	15-Jun-99	12/31/99
Closed Area II Access Area	648.58(b)(1)	64 FR 37914	06/19/00	06/15/00	Scallop FW 13	15-Jun-00	
Nantucket Lightship Access Area*	648.58(b)(2)	64 FR 37914	06/19/00	06/15/00	Scallop FW 13	15-Jun-00	
Delmarva Access Area	648.59(a)(2)	71 FR 33228	06/08/06	06/15/06	Scallop FW 18	15-Jun-06	
Virginia Beach Closed Area	648.57(b)	63 FR 15324	03/31/98	04/03/98	Secretarial Action	3-Apr-98	03/01/01
Elephant Trunk Access Area*	648.58(a)	69 FR 35220	06/23/04	07/23/04	Scallop Amend 10	23-Jul-04	
Elephant Trunk Closed Area	648.58(a)	71 FR 33228	06/08/06	06/15/06	Scallop FW 18	15-Jun-06	
Hudson Canyon Access	648.57(a)	63 FR 15324	03/31/98	04/03/98	Secretarial	3-Apr-98	

Area*					Action		
Closed Area I North EFH*	648.61(a)	69 FR 35223	06/23/04	07/23/04	Scallop Amend 10	23-Jul-04	
Closed Area I South EFH*	648.61(a)	69 FR 35223	06/23/04	07/23/04	Scallop Amend 10	23-Jul-04	
Closed Area II EFH*	648.61(b)	69 FR 35223	06/23/04	07/23/04	Scallop Amend 10	23-Jul-04	
Nantucket Lightship EFH*	648.61(c)	69 FR 35223	06/23/04	07/23/04	Scallop Amend 10	23-Jul-04	
Western Gulf of Maine EFH	648.61(d)	69 FR 35223	06/23/04	07/23/04	Scallop Amend 10	23-Jul-04	
Cashes Ledge EFH	648.61(a)(2)	76 FR 43770	07/21/11	07/21/11	Scallop Amend 15	21-Jul-11	

^ When no year is specified for the end-date, this closure is recurring. When no end date is specified, the closure is year-round.

* These boundaries have changed over time.

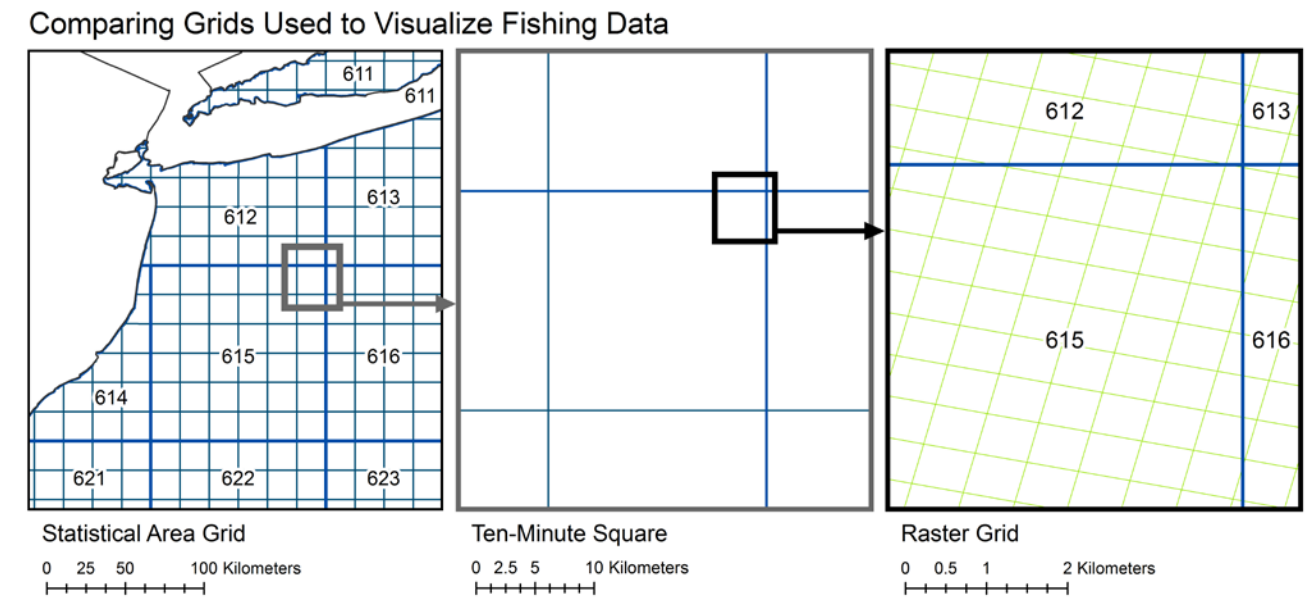

Figure 1. Comparison of statistical area, 10 minute square, and raster grids used in visualizing fishing activity. Fishing data are aggregated to the first 2 grids; in contrast, the raster grid (shown in the third panel) is applied to fishing activity which is estimated for intensity in ocean space. Reprinted from Stoll et al. 2015, with permission from Elsevier.

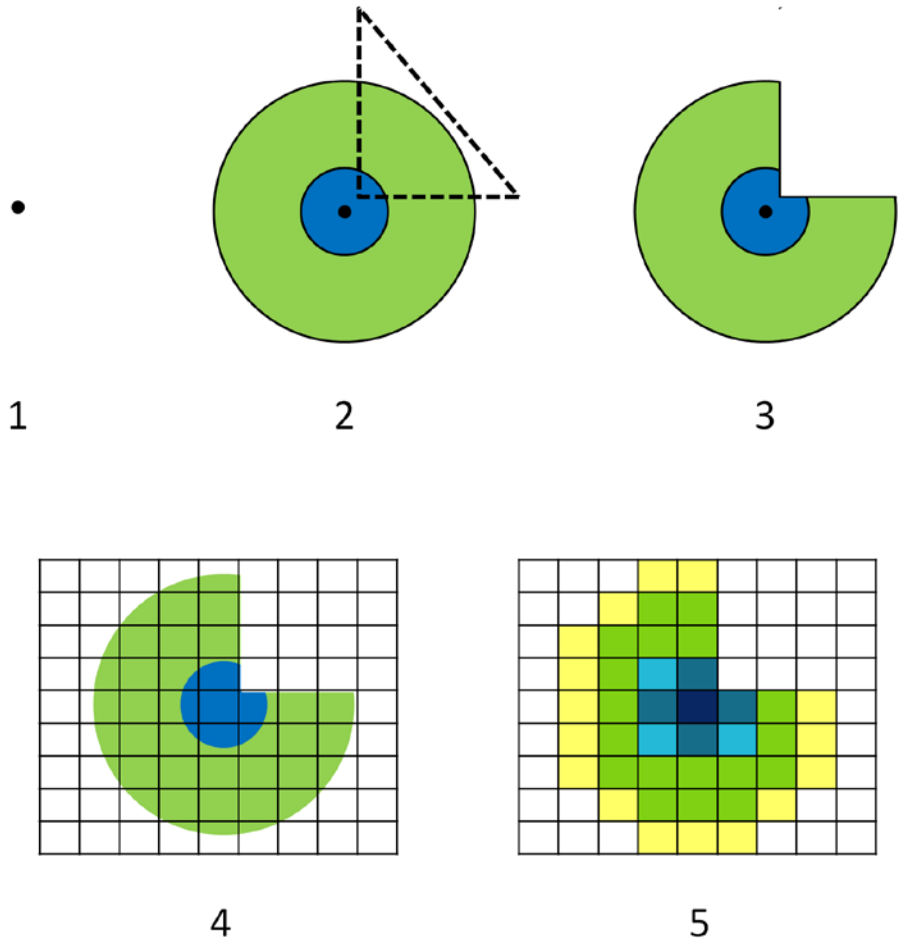

Figure 2. (Panel 1) Vessel Trip Reports (VTR) provide a single point to represent where fishing occurred on a trip.
(Panel 2) Statistical estimate of the radial distance within which fishing activity is likely to occur. The value of the trip is distributed in 25th, 50th, 75th, and 90th percentile rings, which may overlap regulatory closed areas. We show only 2 rings for simplicity in the illustration.
(Panel 3) If the estimated trip area overlaps a closed area (represented by triangle with dashed-lines), the overlapping section is "cut" from the fishing area, and the trip value from that area is redistributed.
(Panel 4) A 500m resolution grid (raster file) is created for each trip record.
(Panel 5) All cells that overlap the estimated trip area sum to a value of 1; thus, each cell represents a portion of the value of that trip.
Reprinted from Stoll et al. 2015, with permission from Elsevier.

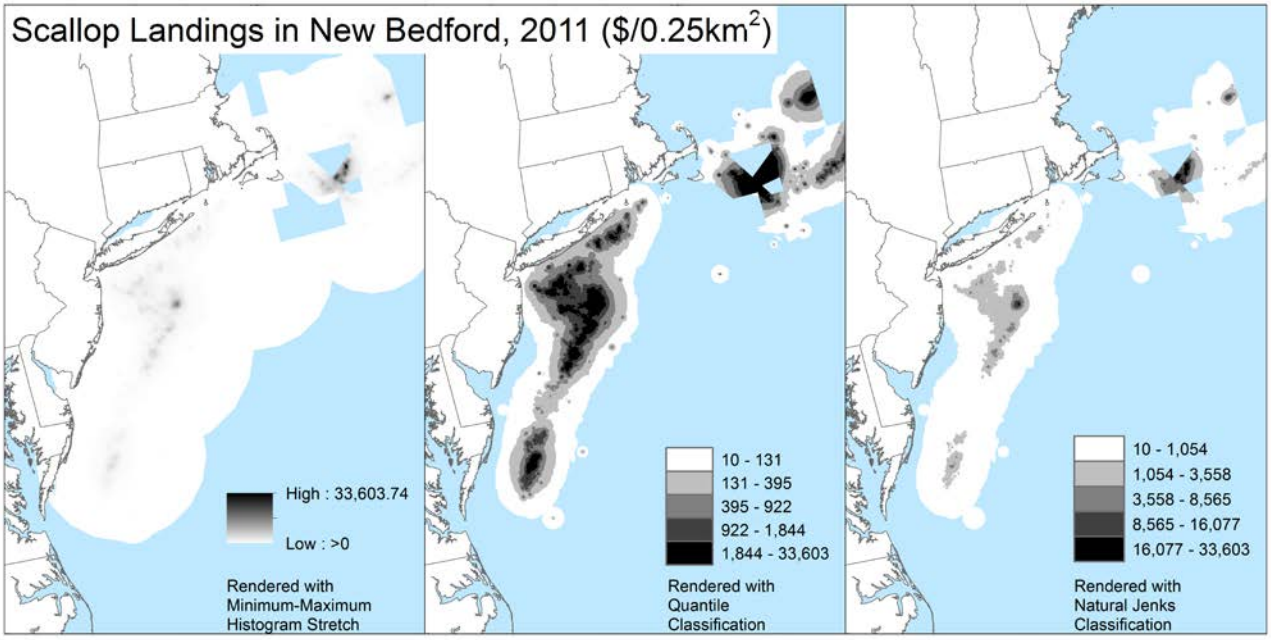

Figure 3. The same raster, depicting scallop fishing activity for trips landing their catch in New Bedford, MA, in the year 2011, is rendered by using 3 different methods in ArcGIS for desktop.

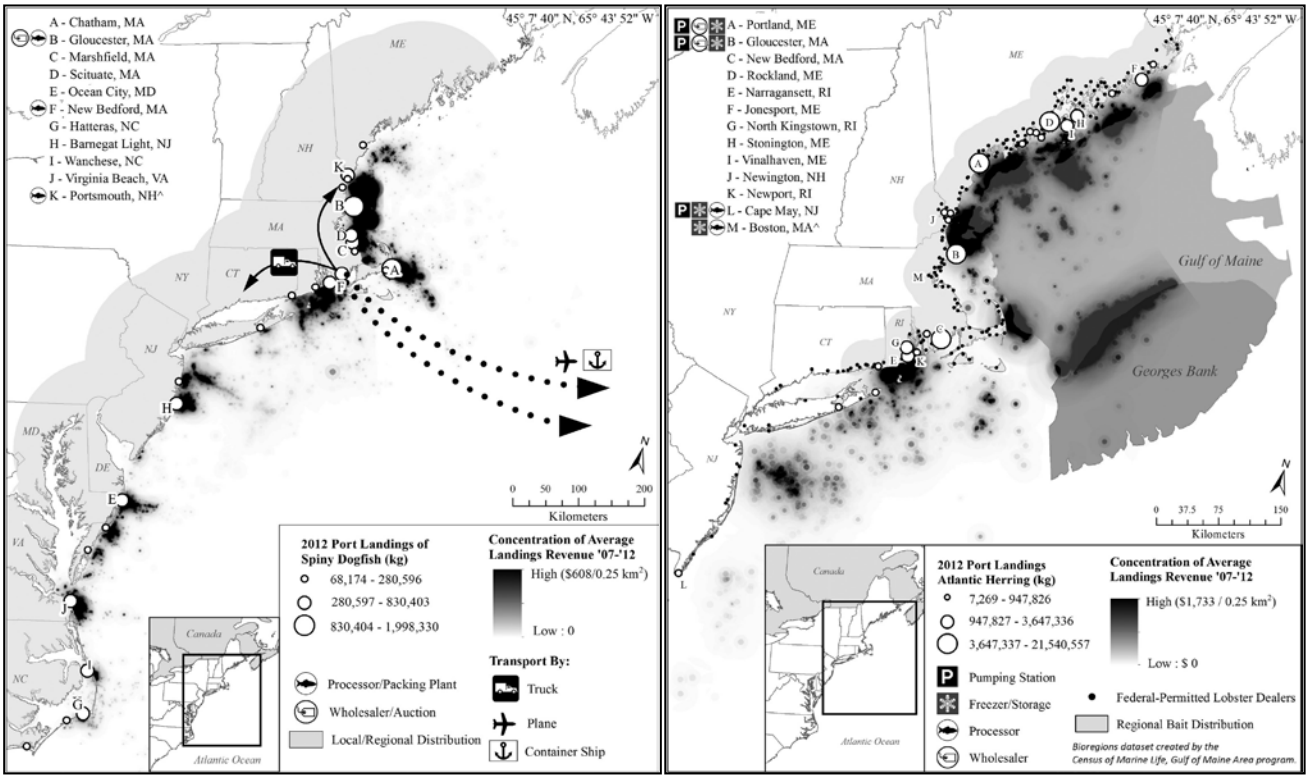

Figure 4. Maps used to illustrate connectivity between fishery and seafood supply chain in the dogfish (*Squalus acanthias*) (left) and Atlantic herring (*Clupea harengus*) (right) fisheries and were featured in Stoll et al. (2015).

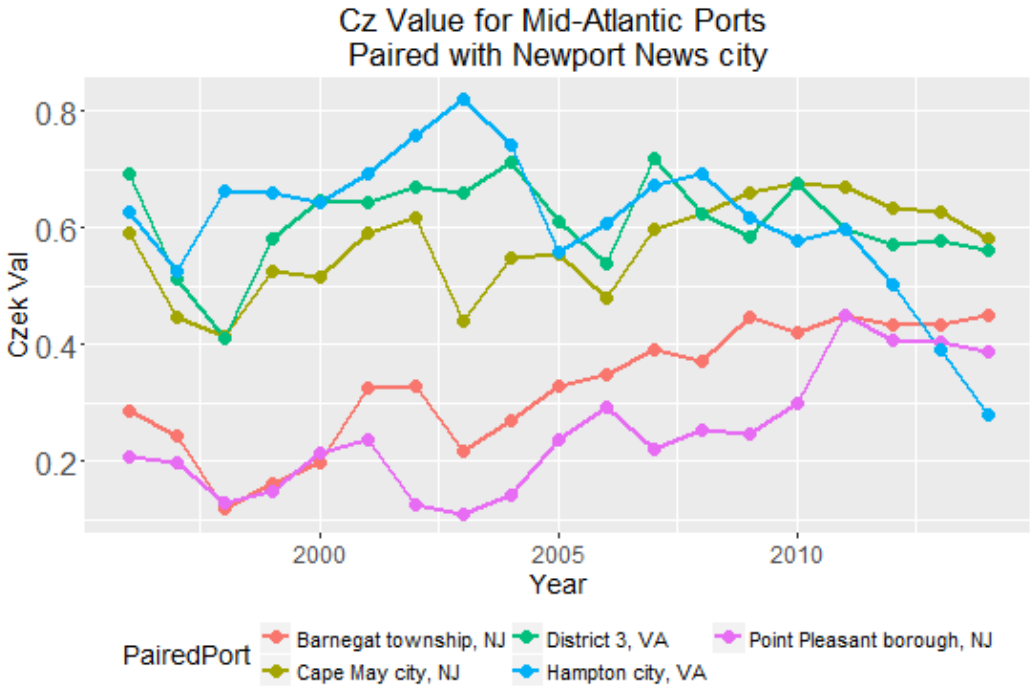

Figure 5. Comparing the overlap index values for Newport News, VA, paired with ports in the Mid-Atlantic, taken from Benjamin et al. (2016). Note: Cz refers to the Czekanowski index, as defined in equation 2.

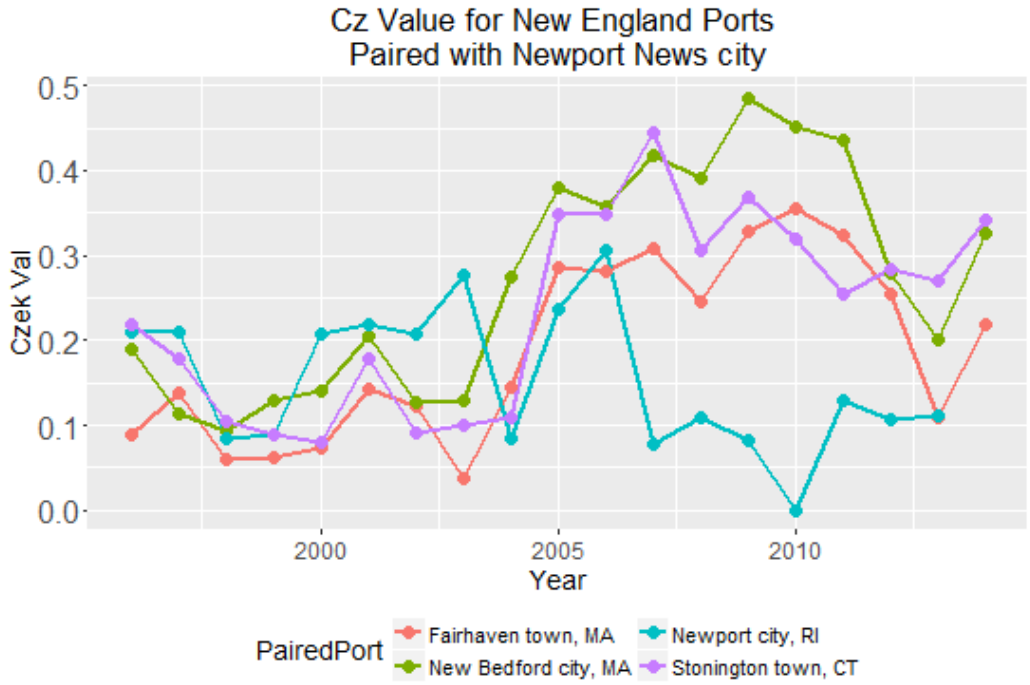

Figure 6. Comparing the overlap index values for Newport News, VA, paired with ports in New England, taken from Benjamin et al. (2016). Note: Cz refers to the Czekanowski index, as defined in equation 2.

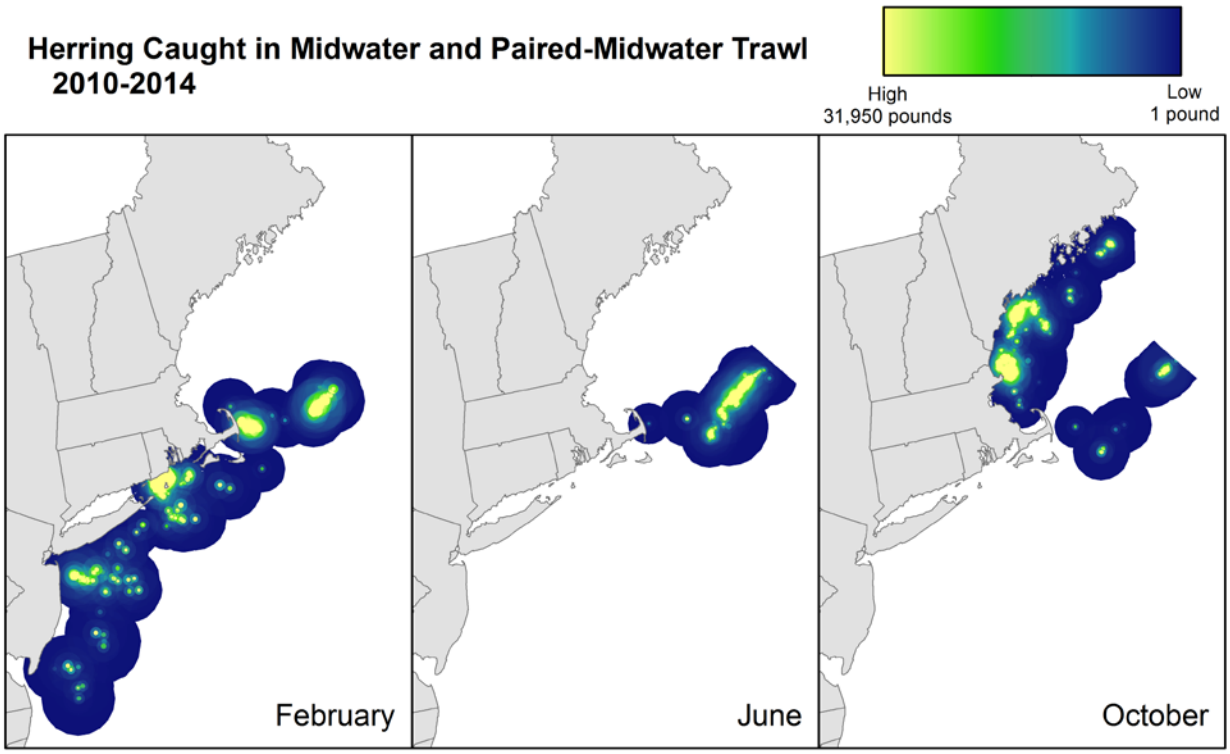

Figure 7. Midwater herring (*Clupea harengus*) trawl catch, 2010-2014.

Procedures for Issuing Manuscripts in the *Northeast Fisheries Science Center Reference Document* (CRD) Series

Clearance

All manuscripts submitted for issuance as CRDs must have cleared the NEFSC's manuscript/abstract/webpage review process. If any author is not a federal employee, he/she will be required to sign an "NEFSC Release-of-Copyright Form." If your manuscript includes material from another work which has been copyrighted, then you will need to work with the NEFSC's Editorial Office to arrange for permission to use that material by securing release signatures on the "NEFSC Use-of-Copyrighted-Work Permission Form."

For more information, NEFSC authors should see the NEFSC's online publication policy manual, "Manuscript/abstract/webpage preparation, review, and dissemination: NEFSC author's guide to policy, process, and procedure," located in the Publications/Manuscript Review section of the NEFSC intranet page.

Organization

Manuscripts must have an abstract and table of contents, and (if applicable) lists of figures and tables. As much as possible, use traditional scientific manuscript organization for sections: "Introduction," "Study Area" and/or "Experimental Apparatus," "Methods," "Results," "Discussion," "Conclusions," "Acknowledgments," and "Literature/References Cited."

Style

The CRD series is obligated to conform with the style contained in the current edition of the United States Government Printing Office Style Manual. That style manual is silent on many aspects of scientific manuscripts. The CRD series relies more on the CSE Style Manual. Manuscripts should be prepared to conform with these style manuals.

The CRD series uses the American Fisheries Society's guides to names of fishes, mollusks, and decapod crustaceans, the Society for Marine Mammalogy's guide to names of marine mammals, the Biosciences Information Service's guide to serial title abbreviations, and the ISO's (International Standardization Organization) guide to statistical terms.

For in-text citation, use the name-date system. A special effort should be made to ensure that all necessary bibliographic information is included in the list of cited works. Personal communications must include date, full name, and full mailing address of the contact.

Preparation

Once your document has cleared the review process, the Editorial Office will contact you with publication needs – for example, revised text (if necessary) and separate digital figures and tables if they are embedded in the document. Materials may be submitted to the Editorial Office as files on zip disks or CDs, email attachments, or intranet downloads. Text files should be in Microsoft Word, tables may be in Word or Excel, and graphics files may be in a variety of formats (JPG, GIF, Excel, PowerPoint, etc.).

Production and Distribution

The Editorial Office will perform a copy-edit of the document and may request further revisions. The Editorial Office will develop the inside and outside front covers, the inside and outside back covers, and the title and bibliographic control pages of the document.

Once both the PDF (print) and Web versions of the CRD are ready, the Editorial Office will contact you to review both versions and submit corrections or changes before the document is posted online.

A number of organizations and individuals in the Northeast Region will be notified by e-mail of the availability of the document online.

Research Communications Branch
Northeast Fisheries Science Center
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 and most issues receive copy editing.

Resource Survey Report (formerly *Fishermen's Report*) -- This information report is a regularly-issued, 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. This report undergoes internal review, but receives no technical or copy editing.

TO OBTAIN A COPY of a *NOAA Technical Memorandum NMFS-NE* or a *Northeast Fisheries Science Center Reference Document*, either contact the NEFSC Editorial Office (166 Water St., Woods Hole, MA 02543-1026; 508-495-2350) or consult the NEFSC webpage on "Reports and Publications" (http://www.nefsc.noaa.gov/nefsc/publications/). To access *Resource Survey Report*, consult the Ecosystem Surveys Branch webpage (http://www.nefsc.noaa.gov/femad/ecosurvey/mainpage/).

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