Llopiz Lab EDI: Diet Data

**Proposed Title:** Diet Composition for Small Pelagic Fishes across the Northeast U.S. Continental Shelf, 2013-2015.

**Proposed Keywords:** Fish, Zooplankton, Predation, Northeast U.S. Shelf.

**Proposed Abstract:** *The abstract should provide a description of the dataset, summarizing what are these data, where, when (is the dataset complete or ongoing?), and an overview of why and how collected (put details how in the Methods). Mention the taxonomic standard that you used (if applicable). Please do not use special characters, symbols, or formatting.*

These data represent the diet composition of small pelagic fishes across the Northeast U.S. Continental Shelf as reported by Suca, J.J., et al. (2018) Feeding dynamics of Northwest Atlantic small pelagic fishes. Progress in Oceanography, 165, 52-62, https://doi.org/10.1016/j.pocean.2018.04.014. These data were analyzed to derive the data products provided in the Supplementary Material to this article. The five species of fish in this dataset represent a subset of the species collected in bottom trawls conducted by the NOAA NEFSC Ecosystems Survey Branch from Cape Hatteras to the Gulf of Maine for years 2013-2015. Sampling occurred in the Spring and Fall seasons. Fish were frozen and stomach content analyses were conducted by members of the Llopiz Lab at the Woods Hole Oceanographic Institution. Data are counts for prey items examined under a dissecting microscope. Fish and prey species were matched to the lowest taxonomic level in the Integrated Taxonomic Information System (ITIS) for scientific name and taxonomic serial number. The dataset was supplemented with geospatial and temporal information from NOAA Fisheries trawl databases.

**Proposed Methods:** *Separate your methods into steps (e.g., sample collection, data processing). Be specific, include instrument descriptions, provide a citation to a paper or protocol if applicable. Please do not use special characters, symbols, or formatting.*

## Fish collection

Alewife, blueback herring, Atlantic mackerel, Atlantic herring, and Atlantic butterfish were collected from four NOAA Northeast Fisheries Science Center (NEFSC) trawl surveys: spring 2013, spring 2014, fall 2014, and fall 2015. Each survey spanned the continental shelf from the northern Gulf of Maine to Cape Hatteras, with spring sampling encompassing March through May and fall sampling extending from September through early November. The fish we analyzed were selected randomly from those available within each of 4 regions: Mid-Atlantic Bight, Southern New England, Georges Bank, and Gulf of Maine. The number of fish analyzed per species per station ranged from 1 to 5. Fish were frozen shipboard at −80 °C and stored at −80 °C until processed in the laboratory.

## Diet Data Collection

Fish were thawed to near room temperature before dissection. Fork lengths were recorded for each fish. The entire gastrointestinal tract was used due to the lack of a defined stomach in clupeids. The gastrointestinal tract was then opened, and contents were placed in 95% ethanol for preservation. Gut contents were examined under a dissecting microscope and identified to the lowest taxonomic level practical. In cases of extremely high gut fullness, a known subsample of prey items was taken with a pipette and enumerated, and this value was divided by the fraction of total volume that the subsample represented to yield an approximation for total stomach contents. (We do not recognize the rows in the data table for which the latter applies. The assumption is that all rows represent totals.)

## Data Cleaning

We gathered the original spreadsheet with prey items in wide-format into a long-format table with preyTaxon column and corresponding preyCount column. GMT date and time were merged from NOAA Fisheries trawl databases [provide citation for NOAA Fisheries trawl databases]. Average depth was also obtained from fisheries datasets, with the exception of Cruise 201504. Local time was converted to hh:mm:ss format. Another column was added which contained a unique identifier for each fish examined. Additional columns were added which contain the vernacular name for each fish, the best matched scientific name for fish and prey species from ITIS, along with their corresponding taxonomic serial numbers (TSN). The data cleaning and metadata template assembly was performed in R Markdown. Further documentation can be found on GitHub, at https://github.com/WHOIGit/nes-lter-fish-diet-isotope.

## Quality Assurance

We assured that the geographic and temporal coverage of the clean data table were within expected ranges. We added a quality assurance flag to the decimalLongitude column. We assured that the total counts of prey taxa per fish species aligned with the frequencies reported in Suca et al. (2018) Supplementary Material.