Creation of a gamefish occurrence dataset from public-focused informational newsletters

Rebecca Hedreen, MLIS, MAEd; Dr. Sean Grace
Department of Biology, Southern Connecticut State University

ABSTRACT

Fish and wildlife agencies produce a bounty of information aimed at the public. Under the right circumstances, that information can be compiled into scientifically useful data to complement full scientific studies. This poster describes some preliminary results from a project to compile mentions of gamefish species, locations, and sizes throughout the Long Island Sound and surrounding waters from the Weekly Fishing Report (2006, 2008-2018) and the Trophy Fish Report (2009-2017), both produced by the Connecticut Dept. of Energy and Environmental Protection. The dataset consists of more than 20,000 entries from the reports collected weekly by DEEP employees from tackle shops and charter companies. The current portion of the analysis is to determine the characteristics of the dataset, such as entry types, species counts, and some general trends.

INTRODUCTION

Fish and Wildlife agency data represents a wealth of knowledge, but because of the public service aims of those agencies, that knowledge is not always presented in a scientifically useful way. This project showcases how such information can be compiled into an open dataset of ecologically important information which includes seasonal data unavailable in other sources.

METHODS

Mentions of specific species were compiled from the marine fisheries portions of the Weekly Fishing Report (2006, 2008-2018), with location, length, and weight data when available. Anonymized data from the Trophy Fish Report (2009-2018) is also included. Clean up and clustering were done in OpenRefine.

DESCRIPTION OF THE DATASET

- 22,487 records
- 21,269 records with geo-codable locations
- 2,775 records with length measurements
- 3,314 records with weight measurements
- 1,118 records with geo-codable locations, lengths, and weights
- 63 species
- 329 geo-codable locations
- Most common species: Striped bass, Bluefish, Summer flounder, and Scup

Figure 4. Heatmap of geo-coded locations in the dataset

Figure 5. Common species included in the data set.

Figure 6. Rare species included in the data set.

LIMITATIONS

- Data has not been cross checked, either by the original compiler or independently
- Approximate locations, unknown locations
 - “Shoreline locations,” “throughout the Sound,” and other general location statements mean fewer specific location entries for commonly occurring species
- Where is Rugville Reef?
- Assumptions about clustering
 - Gardeners, Gardners, and Gardiners - all the same place?
- Mix of true occurrence data and predicted occurrence data

ANALYSIS

While this portion of the project did not include data analysis, a few trends showed up in preliminary work:

- Juvenile bluefish sizes fit a linear trend model at $p < 0.0001$
- Max weights of several common species are flat - no significant variations across the 11 year span of the current data

Figure 7. Reported length of juvenile bluefish by month.

Figure 8. Maximum weights showed no significant increase or decrease

FUTURE

- Additional 10-12 years of CT data, published in the Hartford Courant
- NY data published in the NYTimes back to 1965
- Similar newsletters from other states (Chesapeake Bay, possibly in the Carolinas and Florida)
- Combining with additional data: hypoxia, temperature, storms/disturbances
- Comparing data with currently available data, such as the trawl surveys
- The dataset and related materials will be openly available on the project website

Figure 9. Graphic from the New York Times fishing report (1970).

REFERENCES

1. CT DEEP Weekly Fishing Reports, https://www.ct.gov/deep/cwp/view.asp?a=2696&q=322752&deepNav_GID=1630, archived at <http://cdmresolver.worldcat.org/oclc/270710053/viewonline>
2. CT DEEP Saltwater Trophy Fish Award Program Reports, https://www.ct.gov/deep/cwp/view.asp?a=2696&q=322722&deepNav_GID=1630
3. Weekend Fishing and Boating. (1970, July 17). New York Times, pg 24.

CONTACT

Rebecca Hedreen
Southern Connecticut State University
501 Crescent St., New Haven, CT 06515
HedreenR1@SouthernCT.edu
203-392-5753
Project Website: <https://osf.io/sm8bv/>

Figure 3. Counts of species and geo-coded places in the dataset

Figure 2. Counts of records with geo-coding, lengths, and weights.