

Data in Brief

Historical fish survey datasets from productive aquatic ecosystems in Lithuania --Manuscript Draft--

Manuscript Number:	
Article Type:	Data Article
Keywords:	fish surveys; data standardization; catch per unit effort; fish size; data sharing; Curonian Lagoon; Kaunas Water Reservoir; gillnetting
Corresponding Author:	Eglė Jakubavičiūtė, PhD Nature Research Centre Vilnius, Vilnius LITHUANIA
First Author:	Eglė Jakubavičiūtė, PhD
Order of Authors:	Eglė Jakubavičiūtė, PhD Freddie Heather Giedrė Višinskienė Augustas Morkvėnas Harry Gorfine Žilvinas Pūtys Linas Ložys Asta Audzijonyte
Abstract:	<p>Many inland ecosystems (lakes, rivers, reservoirs, lagoons) around the world undergo regular biological monitoring surveys, including monitoring the abundance, biomass and size structure of fish communities. Yet, the majority of fish monitoring datasets for inland ecosystems remain inaccessible. This is especially true for historical datasets from the early and middle 20th century, despite their immense importance for establishing baselines of ecosystem status (e.g., prior to manifestations of climate change and intensive fisheries impacts), assessing the current status of fish stocks, and more generally determining temporal changes in fish populations. Here we present a newly digitized fish monitoring dataset for two major Lithuanian inland ecosystems – Curonian Lagoon and Kaunas Water Reservoir. The data comprises >60000 records from >800 fish surveys conducted during 1950s to 1980s, using a range of fishing gears and sampling methods. We introduce three different definitions for survey methods to describe the level of detail for each fish community study. Method 1 surveys include individual fish sizes and weights, Method 2 surveys record frequencies of fish in length or weight groups, whereas Method 3 only records the total catch biomass of a given species. The majority of historical and currently collected fish survey data can be attributed to one of these three methods and we present R codes to convert data from higher resolution methods into aggregated data formats, to facilitate data sharing. In addition, commercial fisheries catch data for years that were surveyed are also provided. The data presented here can facilitate ecological and fisheries analyses of baseline ecosystem status before the onsets of rapid warming and eutrophication, exploration of fish size structure, evaluation of different catch per unit effort standardization methods, and assessment of population responses to commercial fishing.</p>
Suggested Reviewers:	Magnus Huss Swedish University of Agricultural Sciences magnus.huss@slu.se Abigail Lynch ajlynch@usgs.gov Hiroshi Okamura okamura@fra.affrc.go.jp

Dear Editors,

Attached please find our publication “Historical fish survey datasets from productive aquatic ecosystems in Lithuania” for your consideration to be published in “Data in brief” journal. In this dataset we present a large and comprehensive data compilation from historical fish surveys in Lithuania, including fish lengths, weights, and biomasses through time. We believe this dataset will be of great value for ecological and fisheries research, including studies of size structure in freshwater ecosystems, recreational and commercial fishing impacts, warming impacts on fish biomass, size composition, and relative abundance.

All authors agree with the publication of dataset and reviewed the final manuscript. The authors declare no conflict of interest.

The manuscript has not been submitted to any other journal.

On behalf of all co-authors,

Eglė Jakubavičiūtė

Historical fish survey datasets from productive aquatic ecosystems in Lithuania

Authors

Eglė Jakubavičiūtė¹, Freddie Heather², Giedrė Višinskienė¹, Augustas Morkvėnas¹, Harry Gorfine¹, Žilvinas Pūtys¹, Linas Ložys¹, Asta Audzijonyte¹

Affiliations

¹ Nature Research Centre, Akademijos g. 2, Vilnius, Lithuania

² Institute for Marine and Antarctic Sciences, University of Tasmania, Australia

Corresponding author(s)

Egle Jakubavičiūtė (egle.jakubaviciute@gamtc.lt), Asta Audzijonyte (asta.audzijonyte@gamtc.lt)

Abstract

Many inland ecosystems (lakes, rivers, reservoirs, lagoons) around the world undergo regular biological monitoring surveys, including monitoring the abundance, biomass and size structure of fish communities. Yet, the majority of fish monitoring datasets for inland ecosystems remain inaccessible. This is especially true for historical datasets from the early and middle 20th century, despite their immense importance for establishing baselines of ecosystem status (e.g., prior to manifestations of climate change and intensive fisheries impacts), assessing the current status of fish stocks, and more generally determining temporal changes in fish populations. Here we present a newly digitized fish monitoring dataset for two major Lithuanian inland ecosystems – Curonian Lagoon and Kaunas Water Reservoir. The data comprises >60000 records from >800 fish surveys conducted during 1950s to 1980s, using a range of fishing gears and sampling methods. We introduce three different definitions for survey methods to describe the level of detail for each fish community study. Method 1 surveys include individual fish sizes and weights, Method 2 surveys record frequencies of fish in length or weight groups, whereas Method 3 only records the total catch biomass of a given species. The majority of historical and currently collected fish survey data can be attributed to one of these three methods and we present R codes to convert data from higher resolution methods into aggregated data formats, to facilitate data sharing. In addition, commercial fisheries catch data for years that were surveyed are also provided. The data presented here can facilitate ecological and fisheries analyses of baseline ecosystem status before the onsets of rapid warming and eutrophication, exploration of fish size structure, evaluation of different catch per unit effort standardization methods, and assessment of population responses to commercial fishing.

Keywords

fish surveys, data standardization, catch per unit effort, fish size, data sharing, Curonian Lagoon, Kaunas Water Reservoir, gillnetting

Specifications Table

Subject	Ecology
Specific subject area	Fisheries ecology Studies of exploited organisms in marine and freshwater systems
Type of data	Table
How data were acquired	Fish abundance, biomass and sizes were surveyed by Nature Research Centre (Lithuania) using a range of gillnets and trawls during 1950–1980s. For each survey fish abundances, sizes and biomasses were recorded and stored in archival journals, currently available at the Nature Research Centre (Lithuania). In this dataset, the archival journals were digitized and data supplemented with all available information on gear types and fishing effort. Details of catches for each fishing survey were provided using three alternative survey methods (Method 1–3).
Data format	Raw
Parameters for data collection	All records of fish surveys, where at least some gear information was available were included in the digitization. Gear information, such as mesh size and gear length or type, was collected as accurately as possible. When specific details were not provided in the original archival journals, expert opinion was used to make the best guess and recorded in separate columns, defining broad categories (e.g., of the mesh size, gear length).
Description of data collection	Fish surveys were conducted using a range of fishing gears, including nets, different trawls and traps, different mesh sizes and gear lengths, and different times of the day. For each fishing survey fish analyses ranged from a general estimate of the total catch biomass only, length frequencies of each species, or individual length and weight measurements. Sometimes different levels of precision were done for each survey, which in the database are identified with Methods 1–3.
Data source location	Institution: Nature Research Centre City: Vilnius Country: Lithuania Latitude and longitude for collected samples/data: Curonian Lagoon – 55.0236° N, 20.8904° E Kaunas Water Reservoir - 54.8604° N, 24.1394° E

Data accessibility	In a public repository: Repository name: Dryad Direct URL to data (currently for private viewing, until the publication of this manuscript): https://datadryad.org/stash/share/utksFXHnn3U9wiwbc3QPTMeOeKlq9OJU7fuPkovHE8Y
---------------------------	---

Value of the Data

- Freshwater and coastal ecosystems are highly vulnerable to global warming, pollution, intensive fishing, and other human impacts. Detailed fish monitoring data from the mid-20th century, as presented here, is essential to establish baselines of ecosystem status and to assess human impacts in terms of the extent of change in ecological parameters and indicators.
- The data are useful for ecological and fisheries research, including studies of size structure in freshwater ecosystems, recreational and commercial fishing impacts, warming impacts on fish biomass, size composition, and relative abundance.
- The data presented here could also be used for exploring various approaches to catch-per-unit-effort standardization, assessments of stock responses to fishing, determining species interactions and changes in relative biomasses, as well as measuring interactions between body size and abundance for different species.

Data Description

The dataset comprises information about fish sampled during scientific surveys. Specifically, data columns include:

- Year, day, and month when the sampling was performed
- Water body (Curonian Lagoon or Kaunas Water Reservoir), and location where the survey sampling took place. If known, precise coordinates of the sampling location are provided.
- Fish species (common English and Latin names), total length (in cm), standard length (in cm), individual weight (in g), sex (1 – male, 2 – female, 3 - juvenile), gonad weight (in g) gonad stage (from immature to spawning, I to VI), age (as inferred from scales), fat weight (in g)
- Gut fullness. Three-digit number reflects gut fullness at the beginning of the gut (first number), in the middle (second) and at the end (third number). 0 – empty, 5 – full. Sometimes only a general indication with one- or two-digit numbers were given.
- 'exclude_from_weight_est' column indicates that a specific record (usually for Method 2, where number of fish i.e., frequency in sequential size groupings is given) should not be used for estimating total weight of the catch for that specific day and location, because the total catch weight is indicated separately, where all the individually measured fish were weighed in bulk.
- Number of fish
- Catch weight (in g), available when fish were not measured individually

- Catch_length_min and catch_length_max – size bins (in cm) of the fish caught, when fish were not measured individually. This refers to Method 3 surveys.
- Gear type (gillnet, trawl, beach seine, fyke-net), total gear length (in m).
- Total_gear_length_broad defines general length categories of the gear used because specific lengths were not always available. “Very short”: ≤ 20 m, “short”: 20-60 m, “medium”: 61-120 m, “long”: 121-300 m, “very long”: > 300 m.
- Mesh size of the gear (in mm) indicates the mesh size of the gear used. In some cases, a range of mesh sizes were used and fish catches were recorded for all meshes combined. In this case mesh_size_min and mesh_size_max indicate ranges of mesh sizes. Mesh_size_broad (in mm) – based on expert judgement, rough categories of the mesh sizes are defined, when exact values were missing: “small” ≤ 38 mm, “big” > 38 mm, “full” – full range, from small to big mesh sizes, e.g., 17-70 mm.
- Soak time (in hours). Soak_time_broad – soak time categories defined based on expert opinion, when exact values (in h) were not available: “short” ≤ 5 h, “medium” 6-12 h, “long” ≥ 12 h.
- Depth (in m) – depth of the sampling location.
- Method (method1, method2, method3) – indicates whether (which) individual fish data category is available (see Materials and Methods section), enabling easy filtering for future analysis.

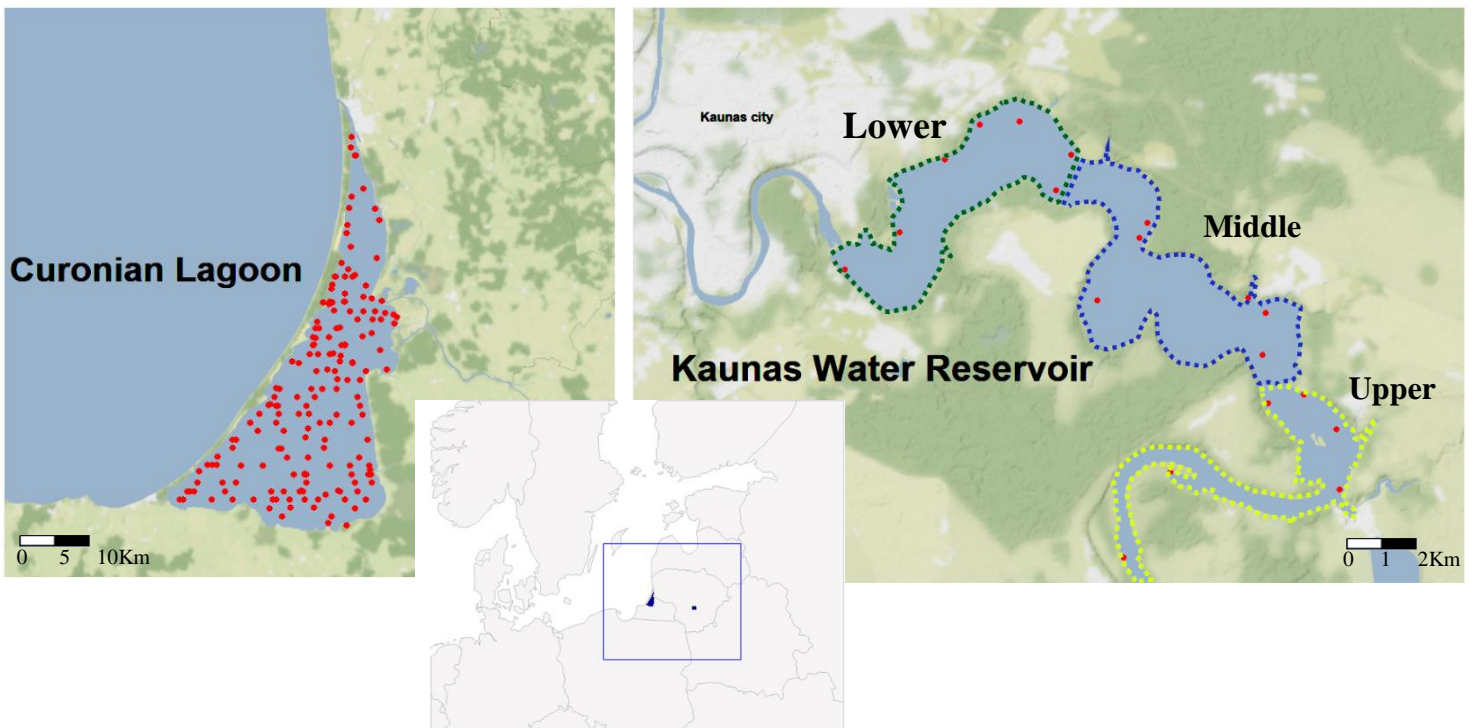


Figure 1. Map of the locations for which the fish monitoring datasets are provided – Curonian Lagoon, where red dots show sampling areas for which coordinates are known. Kaunas Water Reservoir, where dotted lines show division into three sections and sampling locations for which coordinates are known.

Supplementary material 1. Historical commercial fish catches in Curonian Lagoon and Kaunas Water Reservoir.

Supplementary material 2. R code to convert data from Method 1 (most detailed), into Method 2 and Method 3.

Experimental Design, Materials and Methods

Over ~40 years monitoring of fish communities has been undertaken at >150 locations in Curonian Lagoon (1951–1990) and 29 locations in Kaunas Water Reservoir (1961–1990), grouped into three main areas – Upper, Middle and Lower Reservoir (see **Figure 1**). Surveys were mainly performed using traditional gears: nets and trawls. A methodological shift from using predominately trawls to nets took place in the 80s (mostly in Curonian Lagoon).

Fish caught were often measured and weighed individually. In some cases, only frequencies in fish length bins (in cm) and/or weight of total catch (in g) of a given species were recorded. To distinguish between these different types of data and to make it easy for filtering and future analyses, in each data row we assign the method of measurement:

- Method1 applies to surveys where fish were individually measured (total length and/or standard length) and weighed.
- Method2 indicates that fish were individually measured, often within 1 cm precision only, but not weighed.
- Method3 – only total catch weight is available, sometimes also the minimum and maximum lengths of the weighted catch are provided. For this method length bins are larger than 1cm.

Ethics Statement

Not applicable, because data does not involve working with humans or animal experiments.

CRedit author statement

Eglė Jakubavičiūtė: data standardization, conceptualization, methodology, writing; **Freddie Heather**: R scripts for data conversion and analysis; **Giedrė Višinskienė**: data entry and standardization; **Augustas Morkvėnas**: data entry and standardization; **Harry Gorfine**: methodology, writing; **Žilvinas Pūtys**: methodology, standardization; **Linas Ložys**: methodology, data curation, standardization; **Asta Audzijonyte**: supervision, conceptualization, original idea, methodology, writing, funding.

Acknowledgments

This study was funded by the European Regional Development Fund (project No 01.2.2-LMT-K-718-02-0006) under grant agreement with the Research Council of Lithuania (LMTLT).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article.



Click here to download Research Data

<https://datadryad.org/stash/share/utksFXHnn3U9wiwbc3QPTMeOeKlq9>





Click here to access/download
Supplementary material
example_data.csv





Click here to access/download
Supplementary material
Supplementary material_1.xlsx



R code to convert data from Method 1 (most detailed), into Method 2 and Method 3.



Click here to access/download
Supplementary material
method_conversion.R

