

DATA MANAGEMENT PLAN:

1.) Types of data created

Produced data will consist of three main types. 1.) Fish collection locations and dates and associated biometric data (standard lengths, weights); 2.) tissue chemistry data, including otoliths, eye lenses, and muscle tissues (trace element ratios, stable isotope ratios, mercury concentrations); and 3.) modelling code for food web analyses and output data on ecosystem services from scenario analyses. Chemical data will be archived in the Biological & Chemical Oceanography Data Management Office (BCO-DMO); code and output will be archived on GitHub (see below).

Fish & chemical data

Field-collected fish will be measured for length and weight. Measurements recorded by hand in notebooks and later converted to electronic format. Tissues will be dissected and weighed prior to freeze-drying and packaging for analysis. Freeze-dried material will be weighed in tin capsules for bulk isotopic analysis. Otoliths and eye lenses will be cleaned, air-dried, stored in labeled envelopes until they are prepared for analysis. Preparations will be to embed the otolith and lens samples in epoxy and then make thin sections for chemical analysis (below). We will digitally archive all raw chemical data on the acquisition computers, using the original output format. Data will be calibrated using certified reference materials and reports will be generated (using MS Excel) to assess accuracy and precision. The calibrated results and reports on accuracy and precision will be stored on computer in Excel file format. Laser ablation ICP-MS analysis will be used to produce individual otolith and eye lens transects of elemental concentrations across each sample. ICP-MS results will be output first as raw counts (uncorrected) as well as molar ratios calculated after conversion using accompanying analyses of certified reference materials (CRMs). Raw counts, corrected molar ratios, and associated CRMs will all be contributed to public databases. ICP-MS data are output in electronic format by instrument software, and remain in electronic form through correction processes. Each data sheet will be checked for quality control and augmented with metadata regarding sample identification and cross-referenced with fish collection databases (through individual fish ID numbers).

Food web modelling code and outputs

The code for running and analyzing the novel food web model will be written in the form of r-files to be executed in R. During development, draft code files will be stored on local computers as well as on GitHub (github.com; a code development platform). These will be accompanied by an explanatory instructions file, a so-called Vignette, which will be stored and shared alongside the code files. Separate r-files for regenerating the outputs from scenario analyses of the model will be created, for reproducibility.

2.) Standards for data & metadata format and content

Metadata will be created by taking careful notes in laboratory notebooks that refer to specific data files and describe all columns, units, abbreviations and missing value identifiers. Notebooks will be scanned weekly to make .pdf backup files of handwritten notes. Notes will also be transferred weekly to .txt documents that will be stored with the associated data file. These notes will be used to create metadata for all uploaded data.

Field and laboratory data will be documented by both manual entry into lab notebooks and direct electronic quantification. All data will then be copied into Excel spreadsheets weekly and checked for accuracy after entry by having one person enter data and a second person check entries. Excel spreadsheets will be saved as comma-separated value files (.csv) daily and backed up to servers in-house and off site. Metadata .txt files associated with each .csv spreadsheet will be stored with the associated document. PI Limburg will be responsible for otolith and muscle tissue chemistry data management for Baltic and Lake Erie species, PI Walther will manage otolith and

muscle tissue chemistry data for Lavaca Bay, co-PI Razavi will manage mercury concentration data, and co-PI Gårdmark will be responsible for food web modelling code and scenario outputs. PI Limburg will be responsible for coordinating overall data management and consolidation.

3.) Data stewardship & preservation standards

Short-term data preservation will occur by storing (1) data in Excel spreadsheet saved as .csv files, (2) metadata saved as .txt files, and (3) code saved as .r files. Files will be saved to computers automatically and continuously backed up to on-site servers. These computers are also backed up nightly to off-site cloud server storage for redundancy. Hard copies of laboratory notebooks will be also be kept and scanned weekly as .pdf files and entered digitally into spreadsheet form.

Data and associated metadata will be submitted to the Biological & Chemical Oceanography Data Management Office (BCO-DMO) database for long-term preservation and storage within 2 years of the creation of the data. An accession number will be obtained from the BCO-DMO for persistent identification of the data set. PI Limburg will be responsible for uploading data and metadata to BCO-DMO and serve as the primary contact for the BCO-DMO.

4.) Previous data sharing experience

All PIs have many years of experience collecting, storing, preserving and managing data. This includes experience training students and other personnel in proper data management techniques. PIs Limburg and Walther are currently the primary contact for an NSF funded project (OCE-1633020, OCE-1433759) that is archiving data outputs at BCO-DMO. All PIs have had a history of providing data on request to interested parties, sharing data outputs with relevant management bodies, or sharing data with open access databases. PI Walther has attended workshops on Best Practices for Data Management conducted by the DataONE consortium and the California Digital Library. Policies in this document follow recommended procedures by those entities.

5.) Policies for security, access, and sharing

Security of data during storage of produced data will be maintained by strict control of access to relevant personnel only (limited to PIs, students and technicians directly involved in the project). All computers and servers housing associated data will be password protected and data files will be inaccessible by non-involved personnel.

The authors will retain rights to the data until the resulting publication is produced or within two years after the project ends, whichever is first. At this time, authors will provide data for public use. This will consist of uploading data to BCO-DMO and allowing open access to any users. Interested parties will not be required to contact authors prior to downloading, but any use of data will require acknowledgement of authors by citing a BCO-DMO accession number in any resulting publication based on re-use of provided data. Data will also be available for sharing via direct electronic requests to the authors by interested parties; acknowledgement of authors will also be required.