Making Fish Contaminant Data FAIR to Improve Fish Consumption Advisories

Pianpian Wu (Dartmouth College)
Caredwen Foley (Boston University School of Public Health)
June 3, 2021

Overview

Making Fish Contaminant Data FAIR to Improve Fish Consumption Advisories

Overarching goal:
To create a searchable data platform containing publicly-available fish tissue contaminant and environmental data

Research question:
Do fish contaminant data support risk evaluation of combined exposures to multiple contaminants for protective fish consumption advisories?

DARTMOUTH

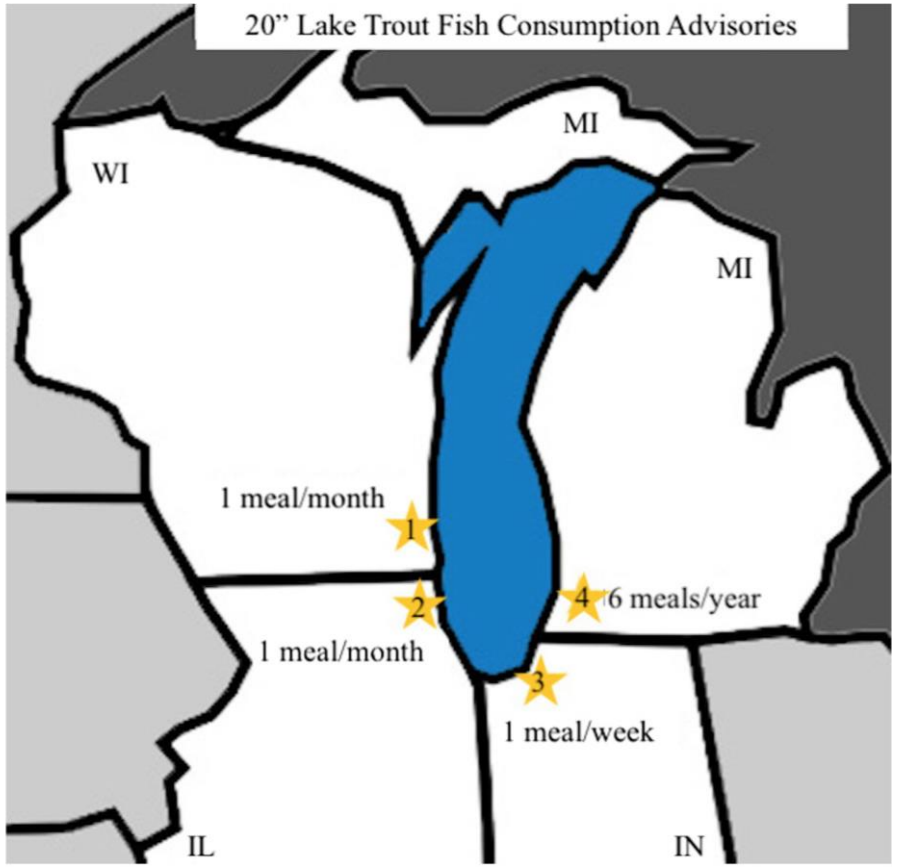
Celia Chen, Arnold Song, Scott Shumway, Pianpian Wu

BOSTON UNIVERSITY

Wendy Heiger-Bernays, Caredwen Foley

Fish Consumption Advisories

- Data collected across monitoring programs differ in chemicals measured; fish species, sizes, and tissues; and ecological data
- State data collection has primarily addressed mercury; assessment of polychlorinated biphenyls (PCBs), pesticides, and emerging contaminants like per- and polyfluorinated substances (PFAS) has lagged
- Advisories differ across states, bodies of water, and contaminants (Cleary et al., 2021)

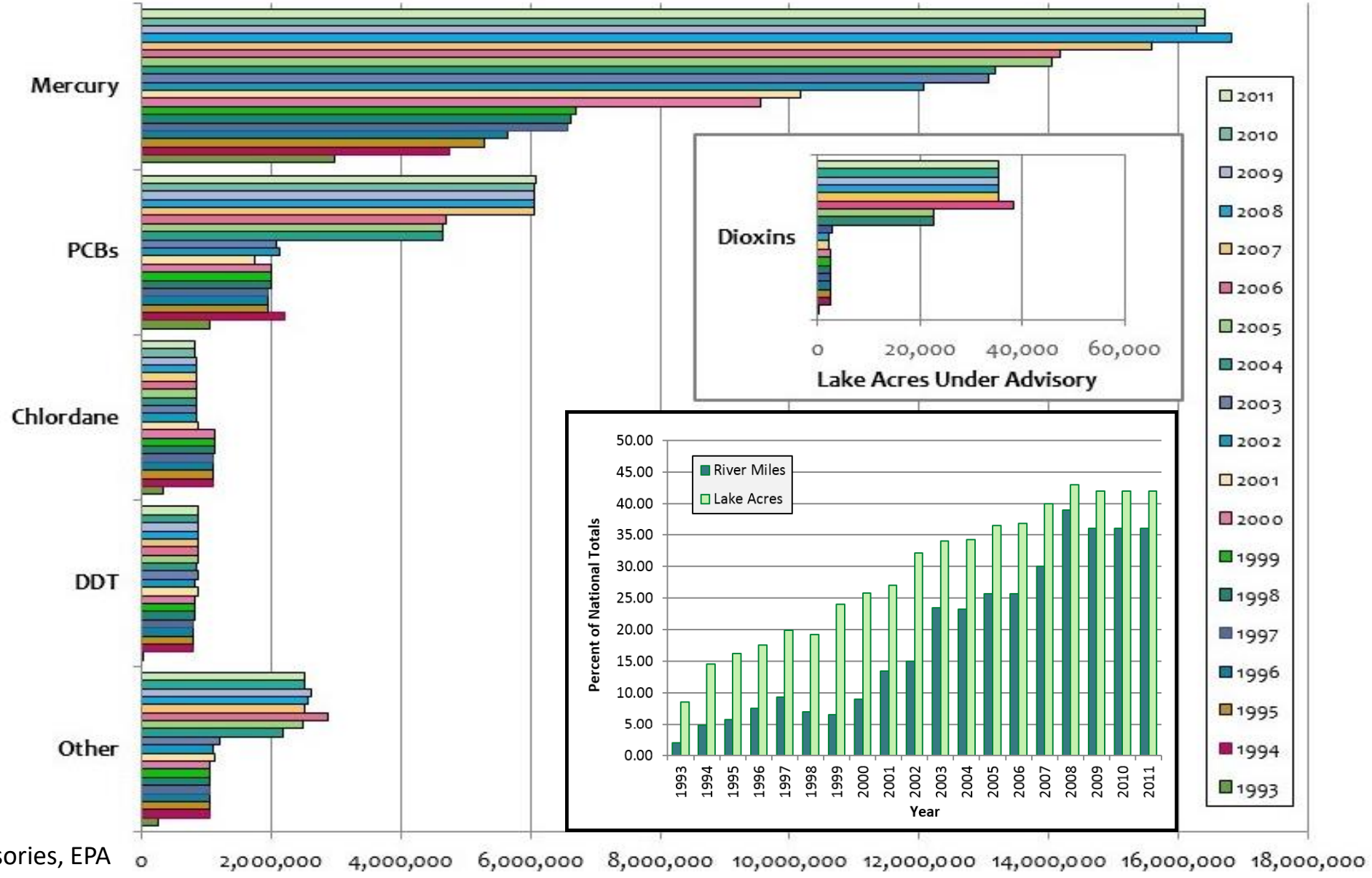


Differences in advisories issued by states bordering Lake Michigan (Cleary et al., 2021)

Cleary, B.M., Romano, M.E., Chen, C.Y., Heiger-Bernays, W., Crawford, K.A., 2021, Comparison of recreational fish consumption advisories across the USA, Current Environmental Health Reports

Fish Consumption Advisories

Total lake acres under advisory for mercury, PCBs, chlordane, dioxins, DDT, and other contaminants from 1993 to 2011.



Source: 2011 National Listing of Fish Advisories, EPA

Project Significance

- Create a more comprehensive assessment of the totality of risks where multiple contaminants are present
- Increase data transparency to engage communities
- Use fish consumption advisories to inform hazardous waste clean-up and source reduction efforts with the end goal of protecting public health

Inputs

Datasets

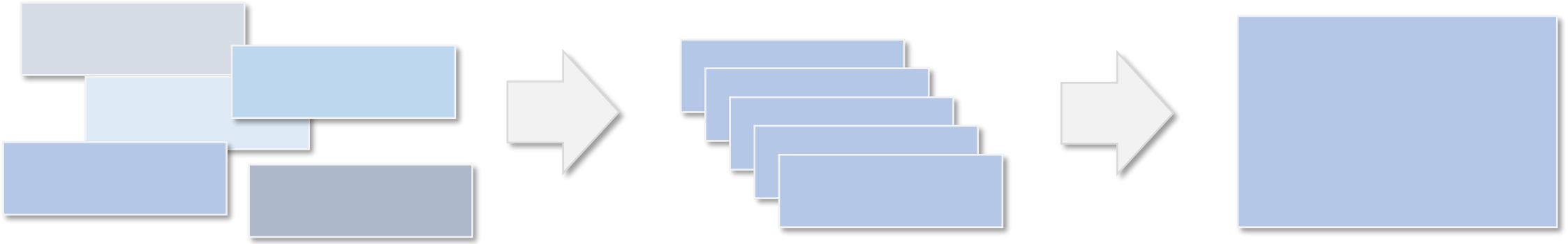
- USEPA fish contaminant datasets
    - National Rivers and Streams Assessment (NRSA) (2008-9, 2013-4)
    - National Lake Fish Tissue Study (NLFTS) (1999-2003)
- Great Lakes Fish Management and Surveillance Program (GLFMSP) (NCCA Great Lakes) (1999 - 2018)

Variables

- Species: freshwater fish
- Matrices: filet vs. whole body
- Contaminant concentrations: metals (Hg), organics (cyclodienes, PCBs, PAHs, DDXs, dioxins, PFAS, PBDEs)

Integrating Data: Repository Development

- Built relational database combining data from US Government datasets (e.g. NCCA, NRSA, NLFTS) including SRP data, allowing users to access these data all at once
- Repository underpins an interactive map visualization to provide a broad view of contamination nationwide (PCBs, mercury, other organic and inorganic pollutants)
- Defined schema for column mapping, data types and lengths
- Open-source reproducible design, including hosting all Extract, Transform, Load (ETL) codes in GitHub, easy to use and share on-premises or cloud hosting solutions



Disparate data sources

Sanitized inputs

Centralized repository

Data source:

NCCA

Download

#	UID	SITE_ID	DATE_COL	STATE	MATRIX	TAXA_NAME	PARAMETER	PARAMETER_NAME	PARAMETER_CAT	CAS_NO
36	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	BETAHCH	Beta-BHC	PESTICIDE	319-85-7
37	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	PCB180	2,2',3,4',5,5',6-Heptachlorobiphenyl	PCB	35065-29-3
38	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	OPDDT	2,4'-DDT	PESTICIDE	789-02-6
39	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	METHOXY	NaN	NaN	NaN
40	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	PPDDE	4,4'-DDE	PESTICIDE	72-55-9
41	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	PCB101	2,2',4,5,5'-Pentachlorobiphenyl	PCB	37680-73-2
42	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	TL	Thallium	METAL	7440-28-0
43	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	ZN	Zinc	METAL	7440-66-6
44	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	PCB156	NaN	PCB	NaN
45	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	PCB170	2,2',3,3'4,4',5-Hexachlorobiphenyl	PCB	35065-30-6
46	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	PCB153	2,2',4,4',5,5'-Hexachlorobiphenyl	PCB	35065-27-1
47	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	GAMMACHL	Gamma-Chlordane	PESTICIDE	5566-34-7
48	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	PCB149	NaN	PCB	NaN
49	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	NI	Nickel	METAL	7440-02-0
50	59	NCCA10-1111	7/1/2010	CA	TISSUE	Umbrina roncador	PCB128	2,2',3,3',4,4'-Hexachlorobiphenyl	PCB	38380-07-3

Addressing Data Integration Challenges

- A uniform ontology for contaminant monitoring in fish tissue data did not previously exist
- Used metadata from source datasets to map between data sources to normalize inputs
- Aggregated and extended several existing ontologies that encompass the following sets of parameters:
  - Ecological and physiological - ENVO
  - Environmental - ENVO
  - Contaminant - ChEBI

Towards an ontology for contaminant measurement in fish tissue samples

Caredwen Foley (1), Wendy Heiger-Bernays (1), Arnold Song (2), Celia Chen (3)

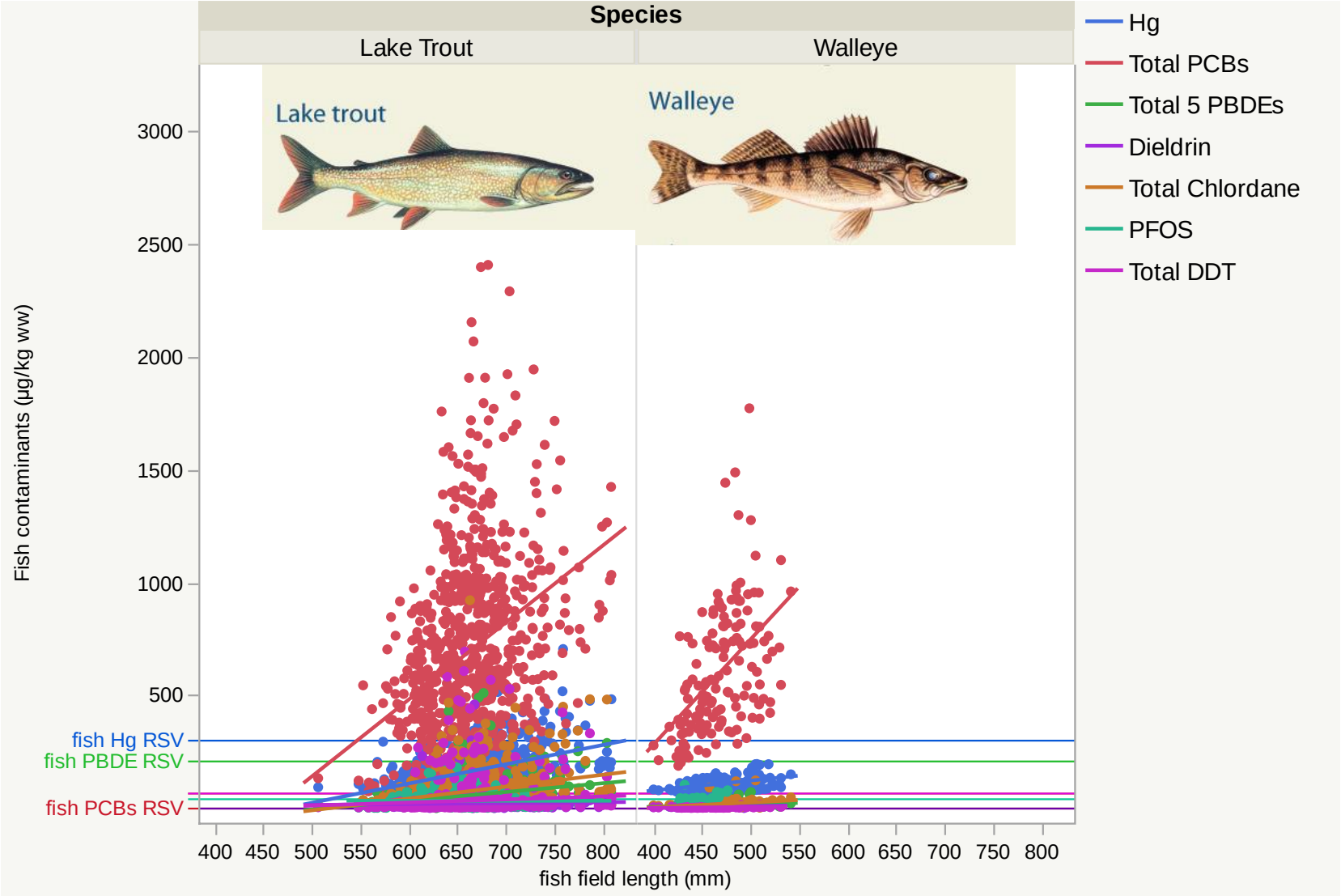
1. Boston University Superfund Research Program, Boston University School of Public Health;
2. Advanced Computing Lab, Dartmouth College;
3. Dartmouth Toxic Metals Superfund Research Project, Dartmouth College

Presented at SRP Annual Meeting, December 14, 2020

Preliminary Mixtures Analysis

Great Lakes Fish Management and Surveillance Program (GLFMSP) data:

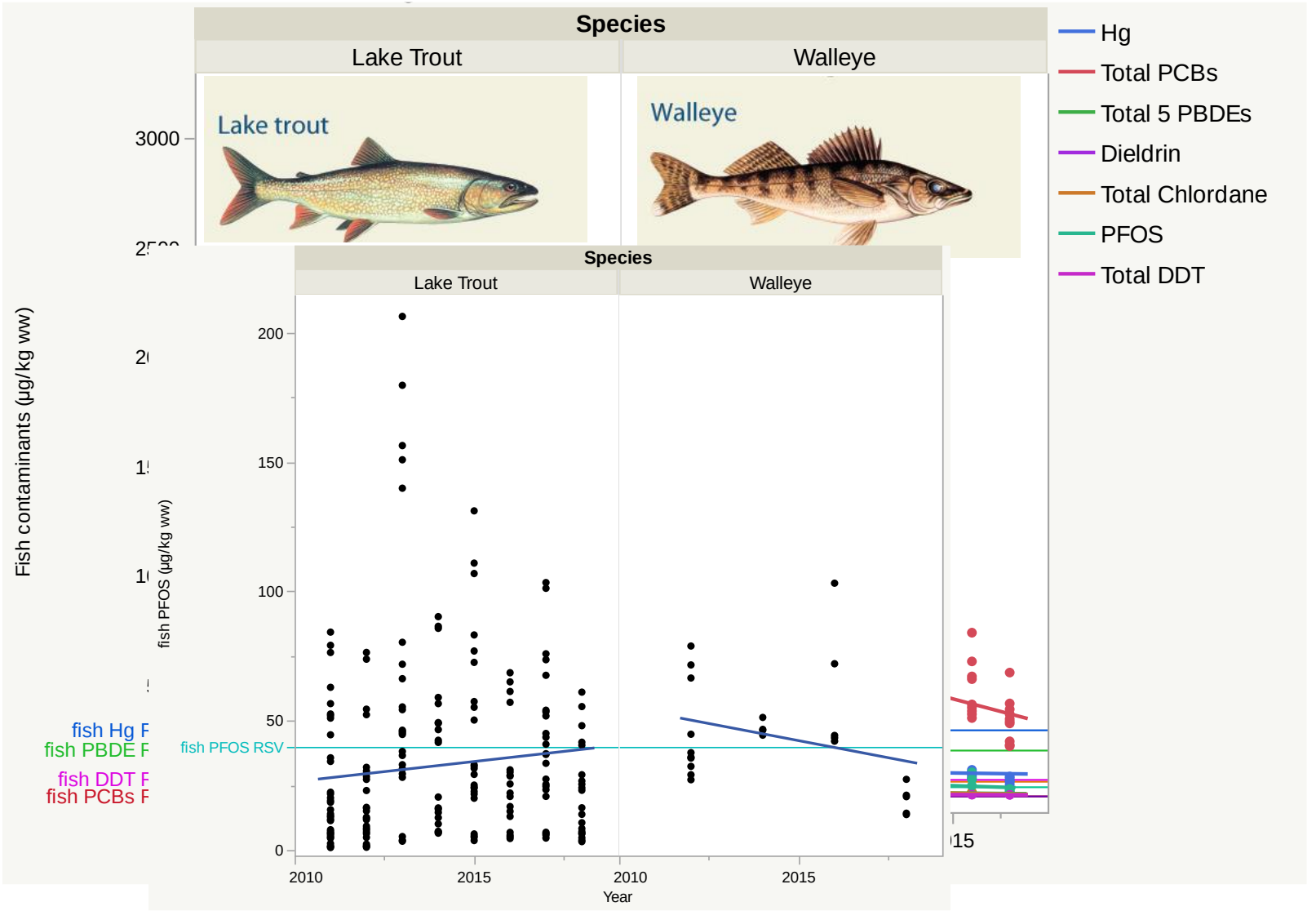
- Mainly lake trout and walleye collected during 1999-2018
- Fish contaminant levels increase with fish sizes, e.g., PCBs and mercury (Hg)



Preliminary Mixtures Analysis

Great Lakes Fish Management and Surveillance Program (GLFMSP) data:

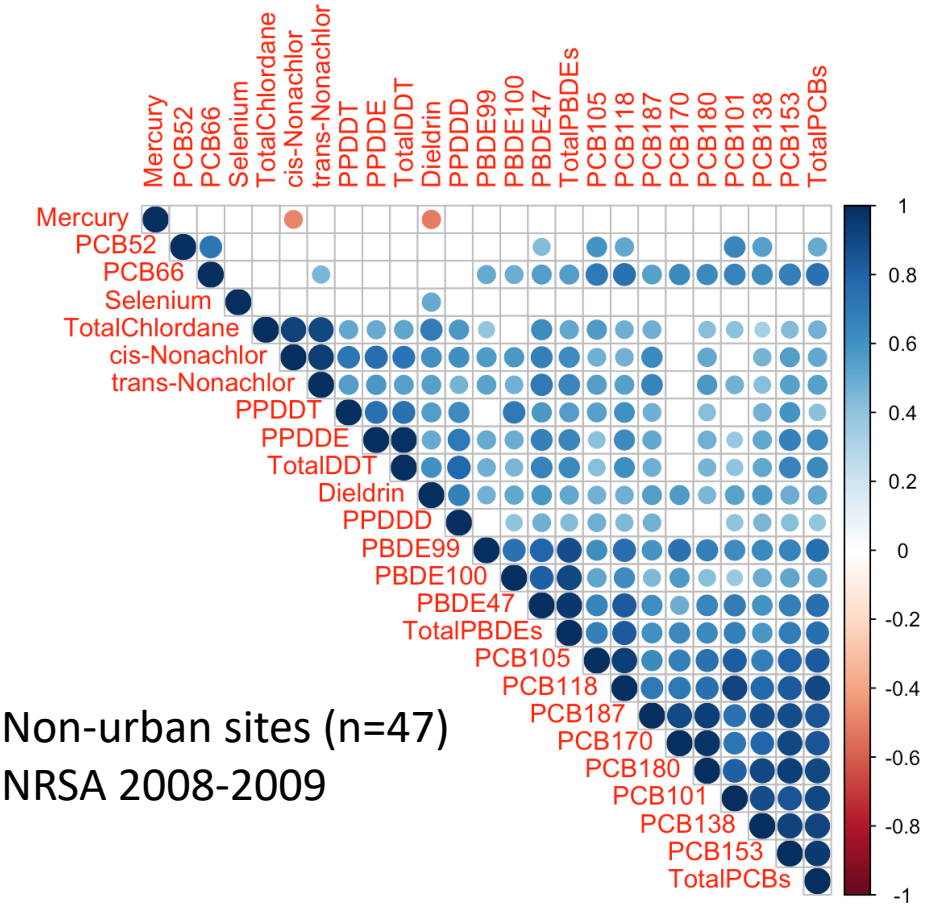
- Fish PCBs levels have decreased over time
- Fish mercury (Hg) levels, as well as pesticides, have remained stable
- Fish polyfluoroalkyl substances (PFOS) still exceed recommended risk screening values for safe fish consumption



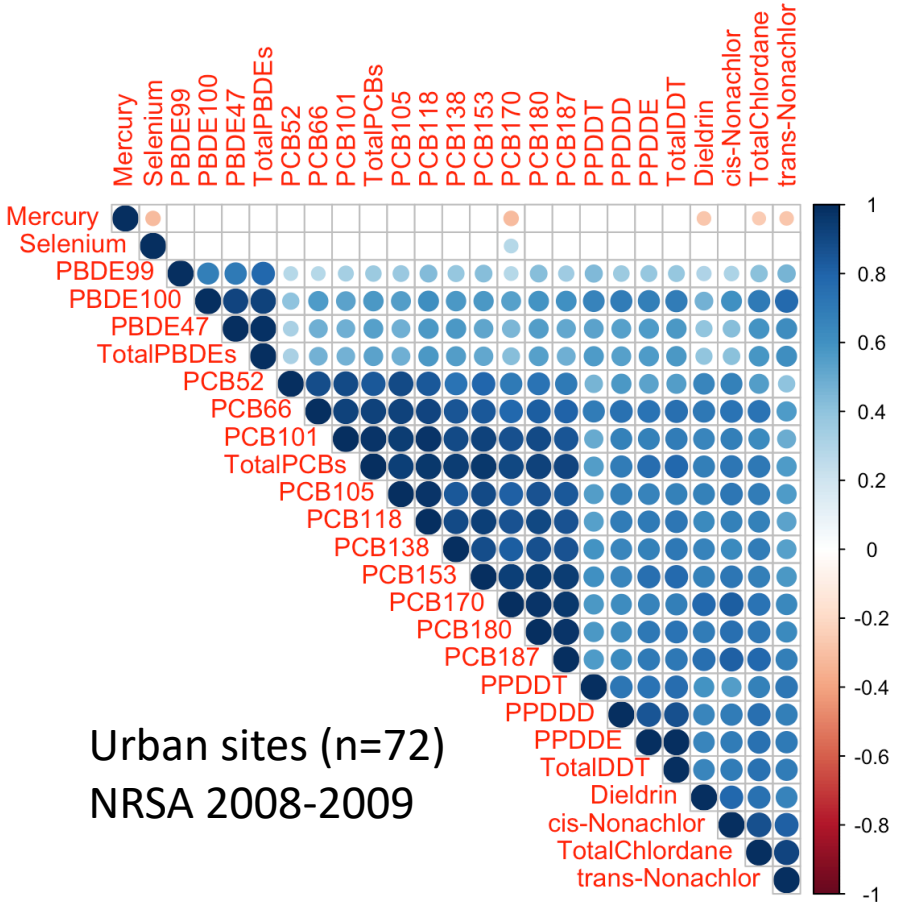
Preliminary Mixtures Analysis

National Rivers and Stream Assessment (NRSA)

- Data collected during 2008-2009 from streams and rivers in 48 states
- Higher correlations of PCB congeners in urban sites



Non-urban sites (n=47)
NRSA 2008-2009



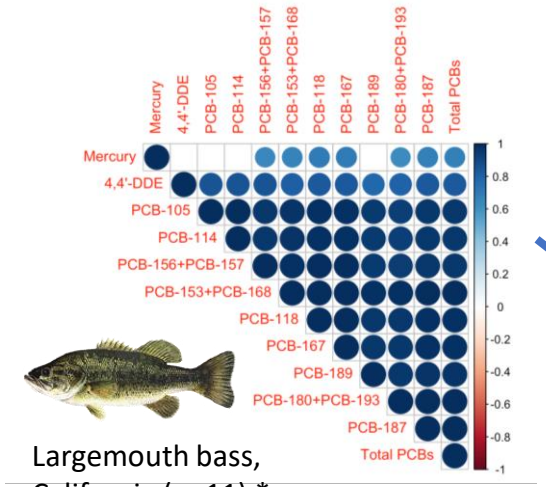
Urban sites (n=72)
NRSA 2008-2009

*Circles in grids represent significant correlation at α=0.05. Size of circle scales with the size of correlation coefficient. Color gradient illustrates whether correlations are positive (blue) or negative (red).

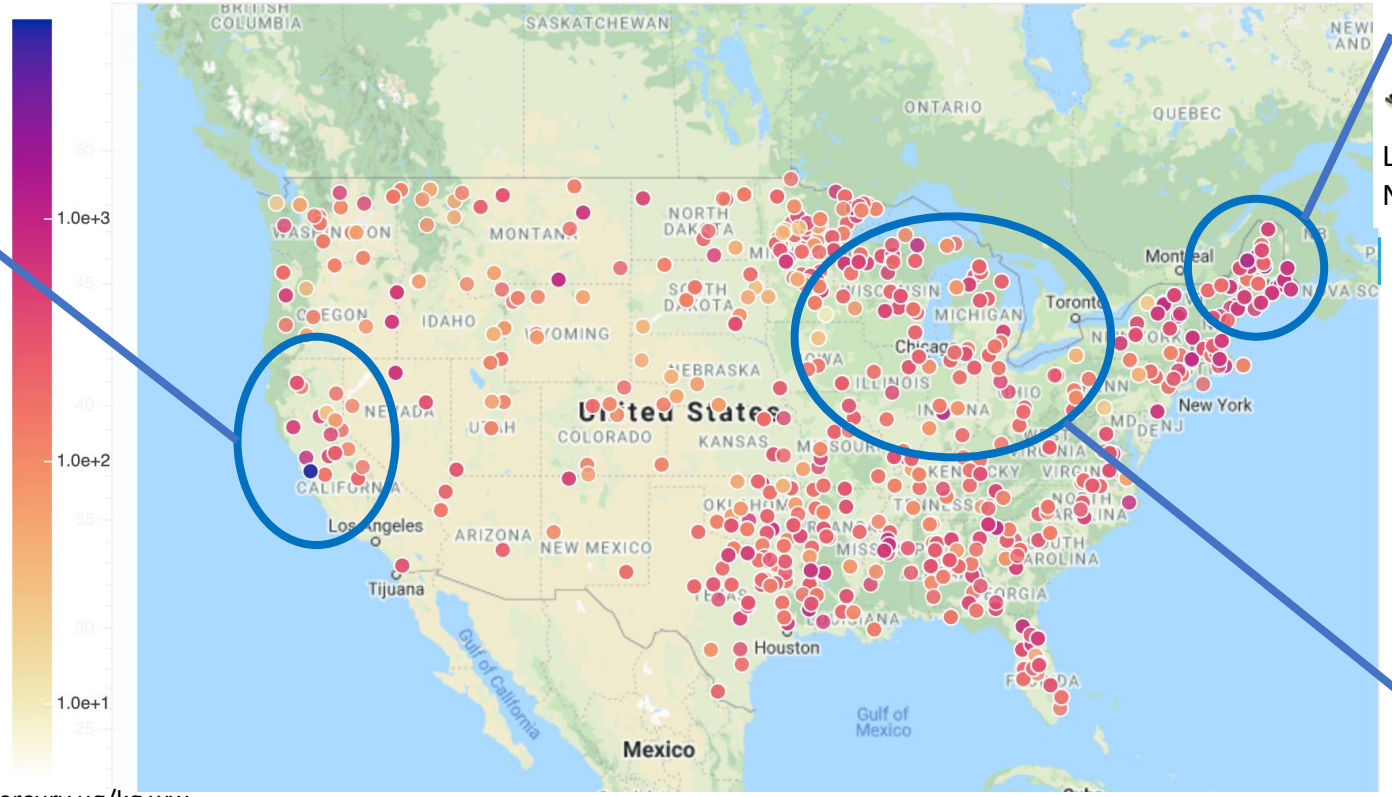
Preliminary Mixtures Analysis

National Study of Chemical Residuals in Lake Fish Tissue

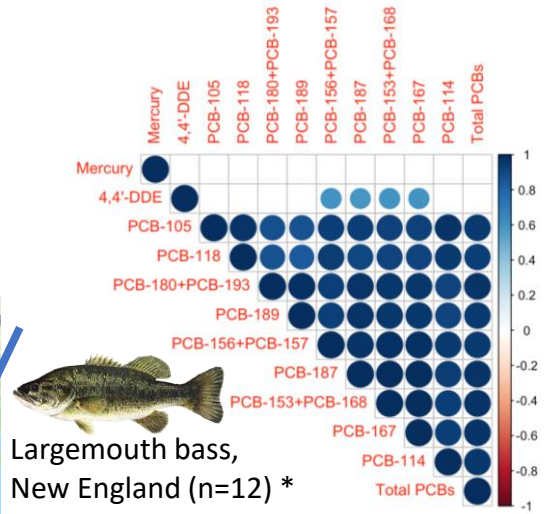
- Data collected during 1999-2003 from lakes in 47 states (n=72)
- Regional differences in correlations of PCB congeners with mercury, esp. California



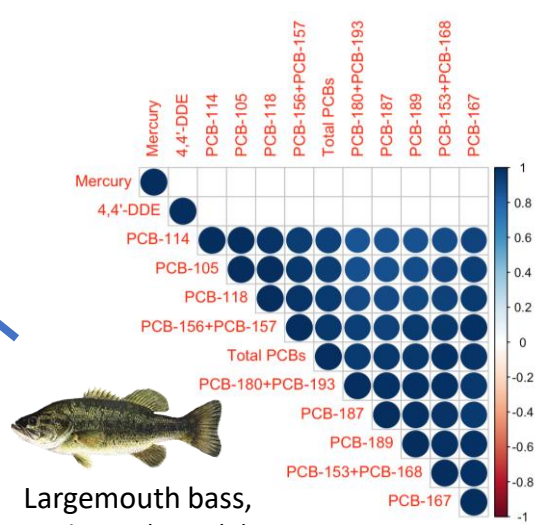
Largemouth bass, California (n=11) *



Fish mercury µg/kg ww



Largemouth bass, New England (n=12) *



Largemouth bass, Midwest (n=59) *

*Circles in grids represent significant correlation at α=0.05. Size of circle scales with the size of correlation coefficient. Color gradient illustrates whether correlations are positive (blue) or negative (red).

Communication and Collaboration Opportunities

This repository creates opportunities for scientific collaboration in:

Data access:

- SRP/other academic environmental health researchers
- “Citizen” (resident) scientists and community groups
- Federal, state, and local researchers

Data collection:

- Explore opportunities for external partners to submit data
  - Will require investigation into quality control and security
  - Consult with other researchers aggregating data to establish QA/QC protocols

What's next?

- **Truly protective consumption advisories require more consistent, comprehensive data collection—especially for multiple contaminants**
  - Extrinsic (water chemistry, land-use activities) and intrinsic (species, trophic level, tissue type, feeding behavior, food availability) variables
  - Chemical characteristics, e.g., persistent/degradable, hydrophobic/hydrophilic, biomagnifying chemicals
- **More funding and support needed to continue efforts to protect public health**
  - Funding for data scientists and infrastructure
  - Funding for addressing research questions relevant to community engagement
- **More collaboration across disciplines (e.g., public health, contaminant chemistry, and ecology), borders (states, federal, and international), and organizations (private sector, government, NGOs)**

Acknowledgement

Funding from National Institute of Environmental Health Sciences: Superfund Hazardous Substance Research and Training Program

- NIEHS/NIH, P42 ES007381-21S1
- NIEHS/NIH, P42 ES007373-23S3

Thank you!

Contact us if you have more questions!

- pianpian.wu@dartmouth.edu
- caredwen@bu.edu