SYNTHETIC ECOLOGY ACROSS SCALES: A GULF OF ALASKA CASE STUDY

Rachael E. Blake, Jessica Couture, Colette L. Ward

NCEAS, Univ. of California Santa Barbara, Santa Barbara, CA

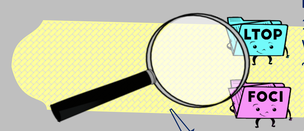
Does scale influence ecological synthesis?

- Spatial, taxonomic, temporal scales; 2 zooplankton datasets

Can a zooplankton community index be built using a taxa subset?

- Are there temporal correlations of abundance? Can we hind cast community abundance?

A tale of two datasets…



LTOP dataset (community):
14 yrs ('98 – '12)
Whole community (56 – 140 taxa)
~250 km; shelf and pelagic (up-current)

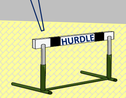
FOCI dataset (taxa subset):
27 yrs ('85 – '12)
Only juvenile Pollock prey (56 taxa)
~50 km; shelf (down-current)

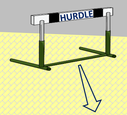
Difficulty discovering data:
web searches not complete, local knowledge often required

Solutions:
- **searchable global repositories such as knb**
- **knowledgeable people working in the system**

Difficulty acquiring data:
scientists don't respond, don't send data, etc.

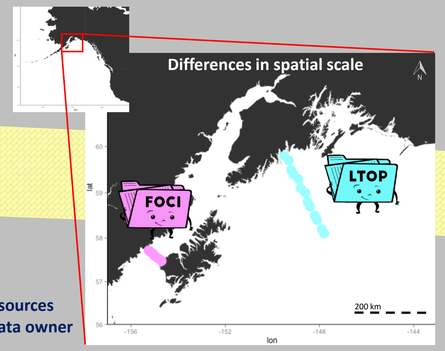
Solutions: be persistent and creative





Data in inaccessible format:
old database format, locked PDF, etc.

Solutions:
- **PDF/html - scraping, curl tools**
- **Old data formats - use institution resources**
- **Request more usable format from data owner**



Poor or missing metadata:
column headings or units undefined

Solutions:
- **contact data custodian for details**
- **search reports for metadata**
- **write complete metadata in EML**

Incomplete data:
don't have all years or samples

Solutions:
- **verify what *should* be included**
- **request missing data from custodian**
- **document complete dataset assembly in script**

Document work by sharing and versioning code online

Ecological Metadata Language to construct metadata

Online data repository with versioning

Open science tools:
Github - www.github.com
KNB - https://knb.ecoinformatics.org/#data/page/0
RStudio - https://www.rstudio.com/
DataONE - https://www.dataone.org/
Other tools - https://knb.ecoinformatics.org/#tools

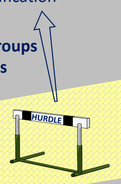
Incomplete data:
analysis limited by available data

Solutions:
- **request additional data**
- **choose appropriate statistical analysis**

Document work by sharing and versioning code online

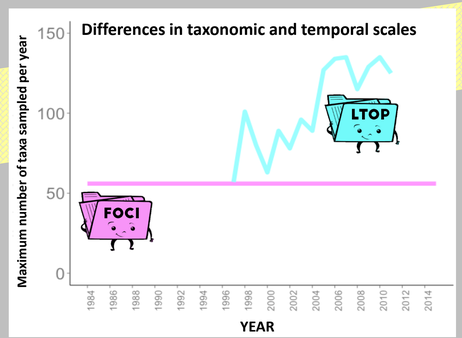
Inconsistent units / reporting between datasets:
different taxonomic or other classification

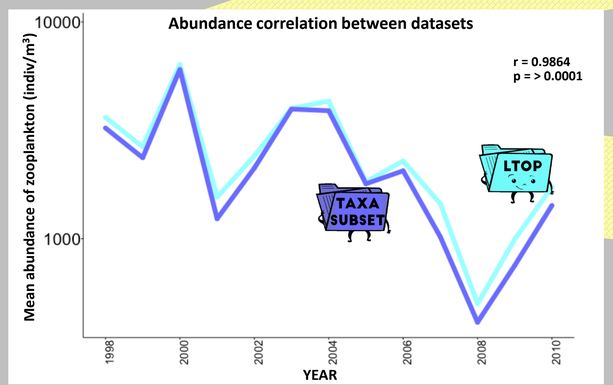
Solutions:
- **aggregate to comparable groups**
- **convert to comparable units**



Spatial or temporal differences between datasets:
conclusions limited by data mis-match

Solutions:
- **find complementary data**
- **find data representative of locations or eras**
- **create index to represent missing data**





Yes! We can build an index for zooplankton.

- High temporal correlation between community and taxa subset
- Taxa subset only comprised of numeric dominants
- Environmental conditions assumed similar across datasets
- Next steps: hind cast abundance based on index, apply to larger ecosystem synthesis

Tips for overcoming data hurdles:

- Metadata, Metadata, Metadata!
- Know specifics of available data (completeness, format, source, etc.)
- Be the squeaky wheel (persistent and specific)!

Yes! Scale influences ecological synthesis.

- Spatial scale may not influence analysis techniques or synthesis results
- Temporal scale limits analysis and conclusions; longer term data better
- Taxonomic scale and resolution limit conclusions; influenced by targeted taxa

Take-home messages:

- Be aware of limitations and caveats of data and analyses
- Document each step in data transformation via scripting
- Use open science tools for reproducibility

Thanks!

- Funding: Exxon Valdez Trustee Council grant to Matt Jones, NCEAS
- Janet Duffy-Anderson, NOAA; Cheryl Hopcroft, UAF; Stacy Rebich Hespanha, NCEAS ; Matt Jones, NCEAS; Colleen Harpold, NOAA

- Fisheries-Oceanography Coordinated Investigations. 2015. EcoFOCI project, NOAA. http://www.afsc.noaa.gov/Publications/ProcRpts_intro.htm
- Hopcroft R., Coyle K. 2015. Long-Term Observation Program (LTOP) Seward Line ZOOPLANKTON data, Gulf of Alaska (1997-2009). http://gulfwatch.nceas.ucsb.edu/#view/df35b.55.17