IMPROVED UNDERSTANDING OF FISHERIES & ECOSYSTEMS FROM NOISY & DISPARATE DATA

Mark Scheuerell

Fish Ecology Division
Northwest Fisheries Science Center
Seattle, WA

#PSAW2019

@mark_scheuerell

Acknowledgments

Elizabeth Holmes, Jim Thorson, Eric Ward (NMFS)

Eric Buhle, Kevin See (Biomark)

Brice Semmens (Scripps)

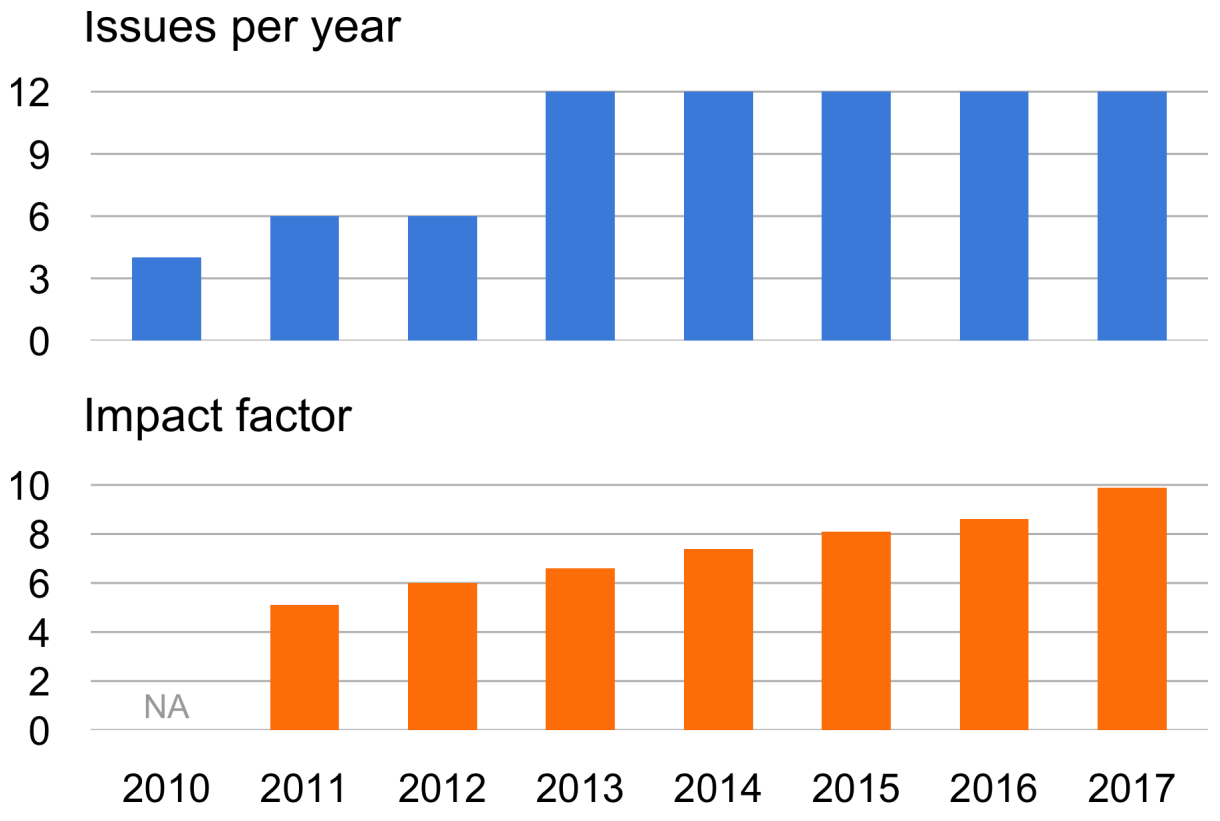
Steve Katz (Washington St Univ)

Q: Do you want clean data?

A: Go into finance

The rest of us seek ways to deal with our noisy & disparate data

Methods in Ecology and Evolution



Lots of focus on *hierarchical models*

Hi·er·ar·chi·cal

adjective

1. Arranged in an order

A hierarchical model is simply
a model within a model

Hierarchical models also masquerade as

Nested data models

Mixed models

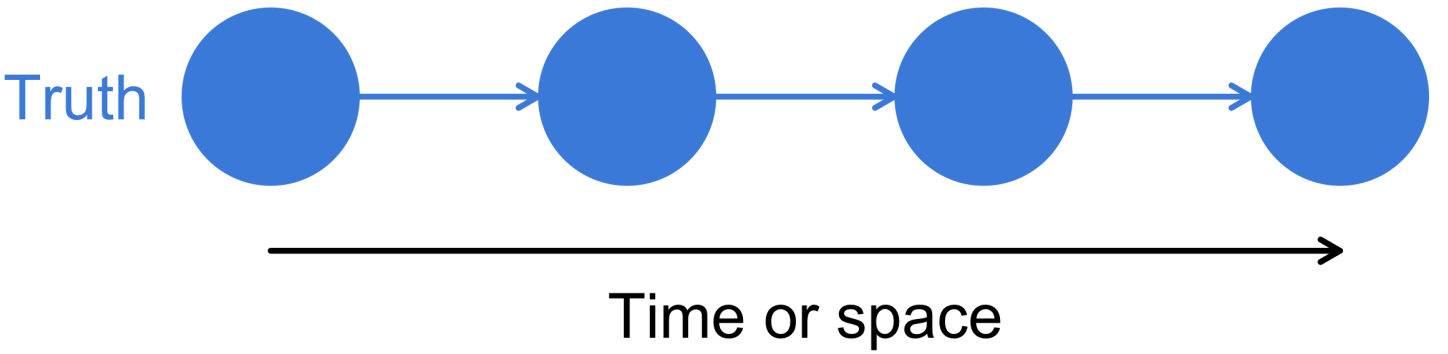
Random-effects models

State-space models

A state-space model has 2 parts

Part 1: State model

Describes the true state of nature over time or space



States of nature might be

Animal location

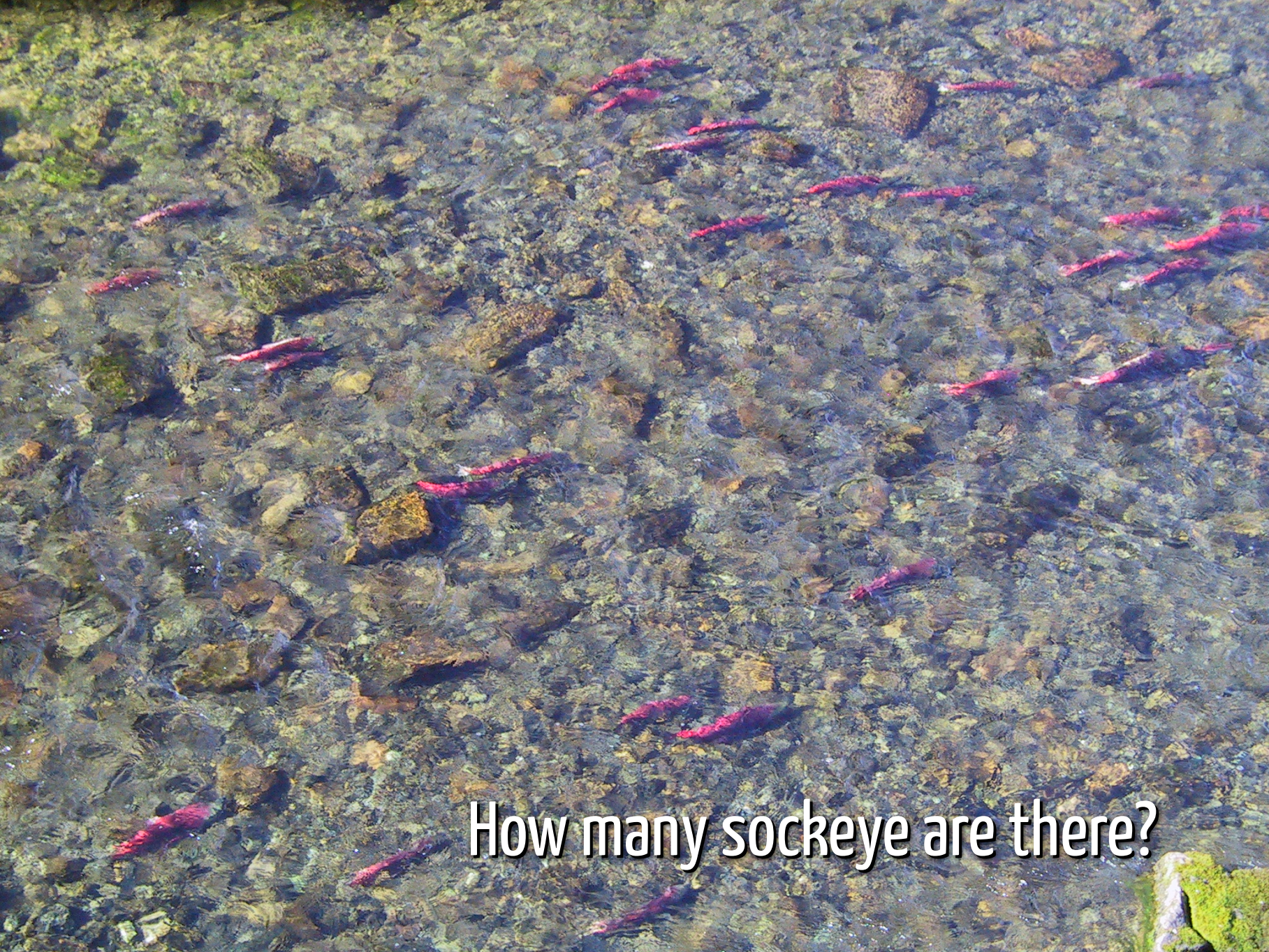
Species density

Age structure

Reproductive status


Revealing the true state requires observations

Observing nature can be easy


How many sockeye are there?

Observing nature can also be hard

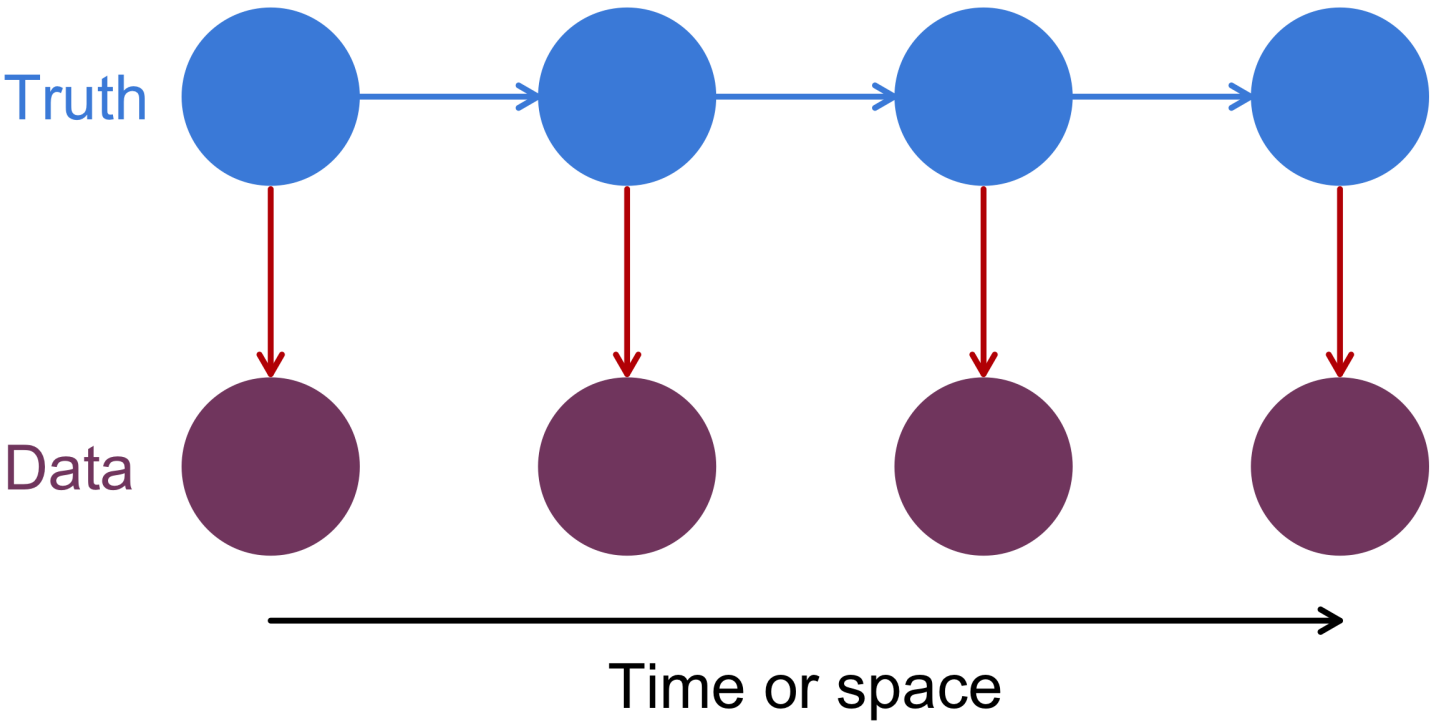

How many mayflies are there?

Part 2: Observation model

Data = Truth ± Errors

Part 2: Observation model

Data = Truth ± Errors



OK, but why bother?

Advantages of hierarchical models

1. Can combine many different data types

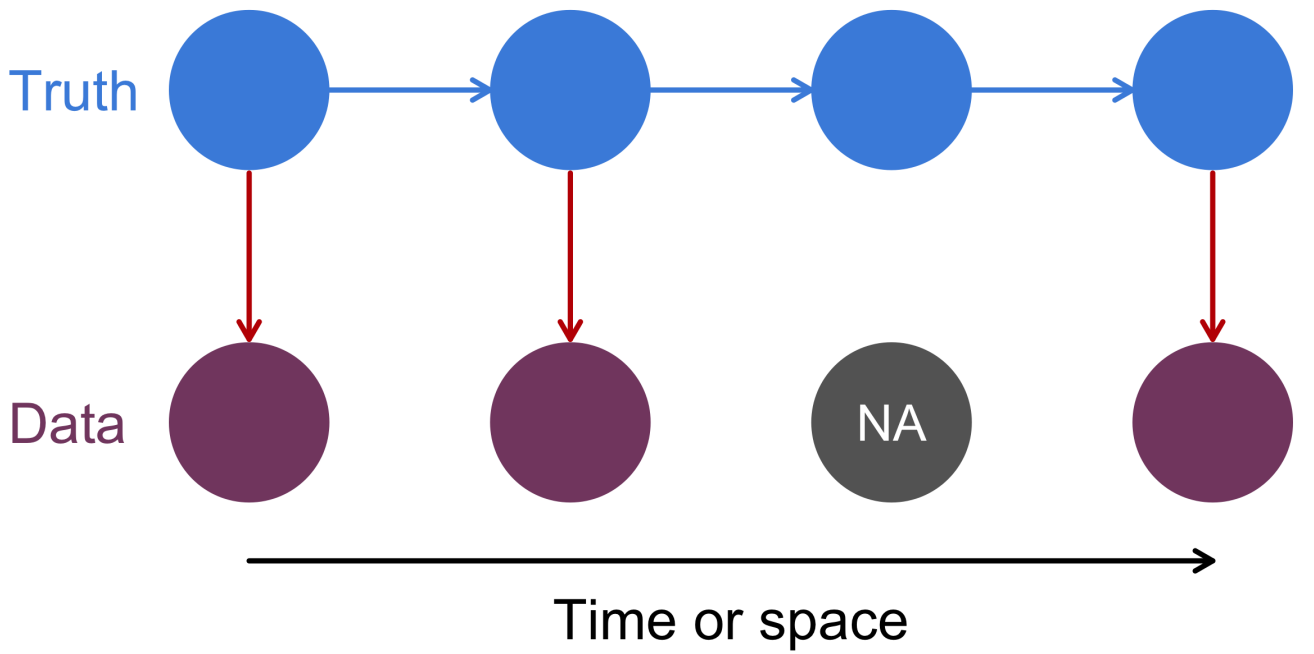
Changes in observers or sensors

Varying survey locations & effort

Direct & remote sampling

Advantages of hierarchical models

2. Missing data are easily accommodated



Advantages of hierarchical models

3. Improved accuracy & precision

Article | OPEN | Published: 08 February 2016

Joint estimation over multiple individuals improves behavioural state inference from animal movement data

Ian Jonsen

Scientific Reports **6**, Article number: 20625 (2016) | Download Citation

Advantages of hierarchical models

4. Data-poor benefit from data-rich

Advantages of hierarchical models

5. Rather flexible

This simple model can be used for 5+ unique applications

$$\mathbf{x}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{w}_t$$
$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{v}_t$$

How do I actually do this?

Many software options

Canned **R** packages (dlm, vars, MARSS*)

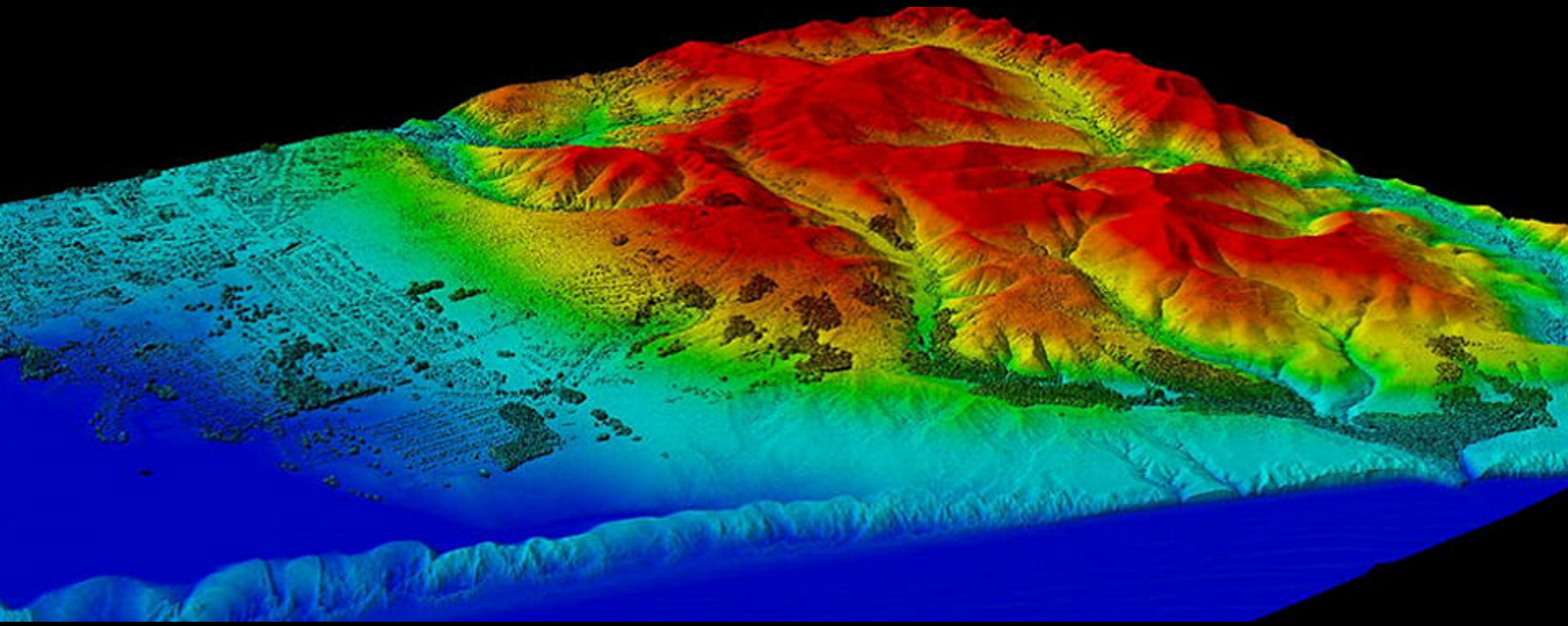
Code-your-own (JAGS, Stan, greta)

*Holmes, Ward, Scheuerell (2018) *Analysis of multivariate time-series using the MARSS package*


SNAKE
OIL

Emergence of high-dimension data

Remote sensing



Genetics

Citizen Science

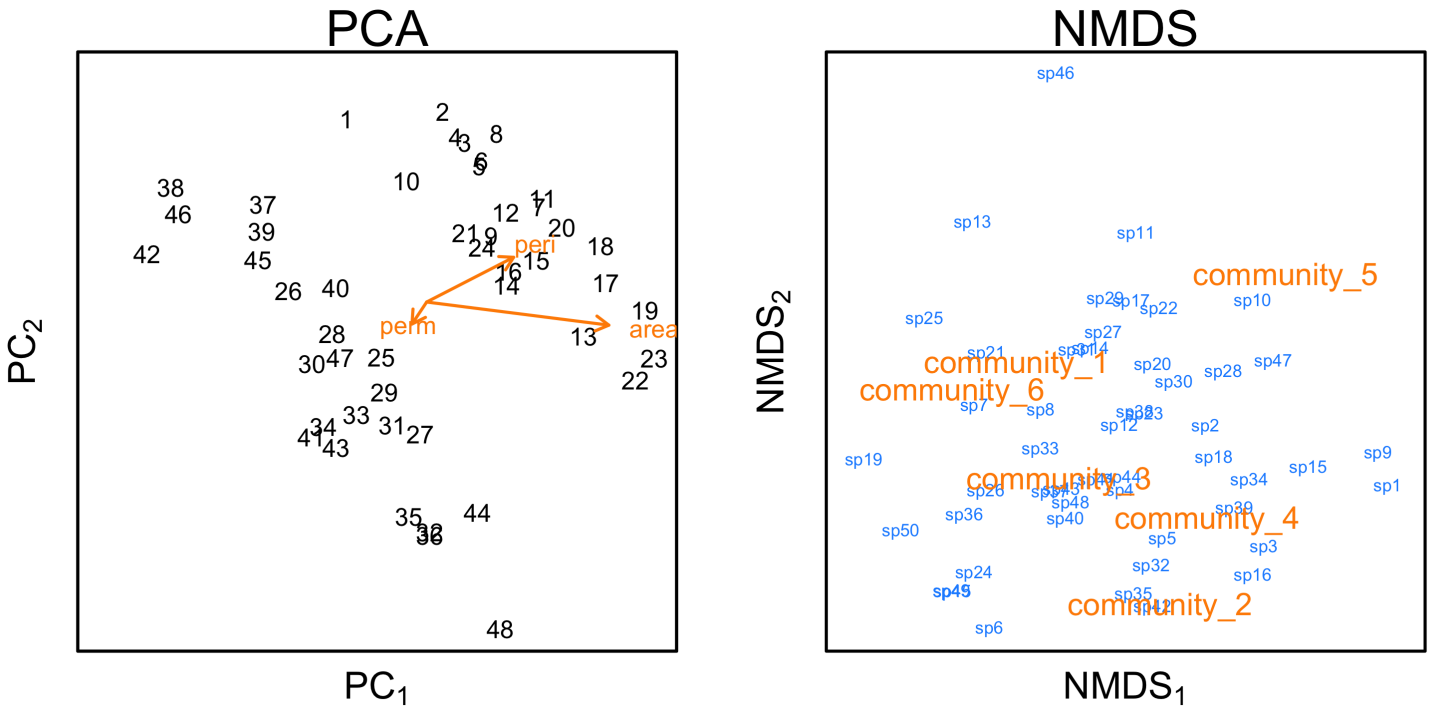
Audubon

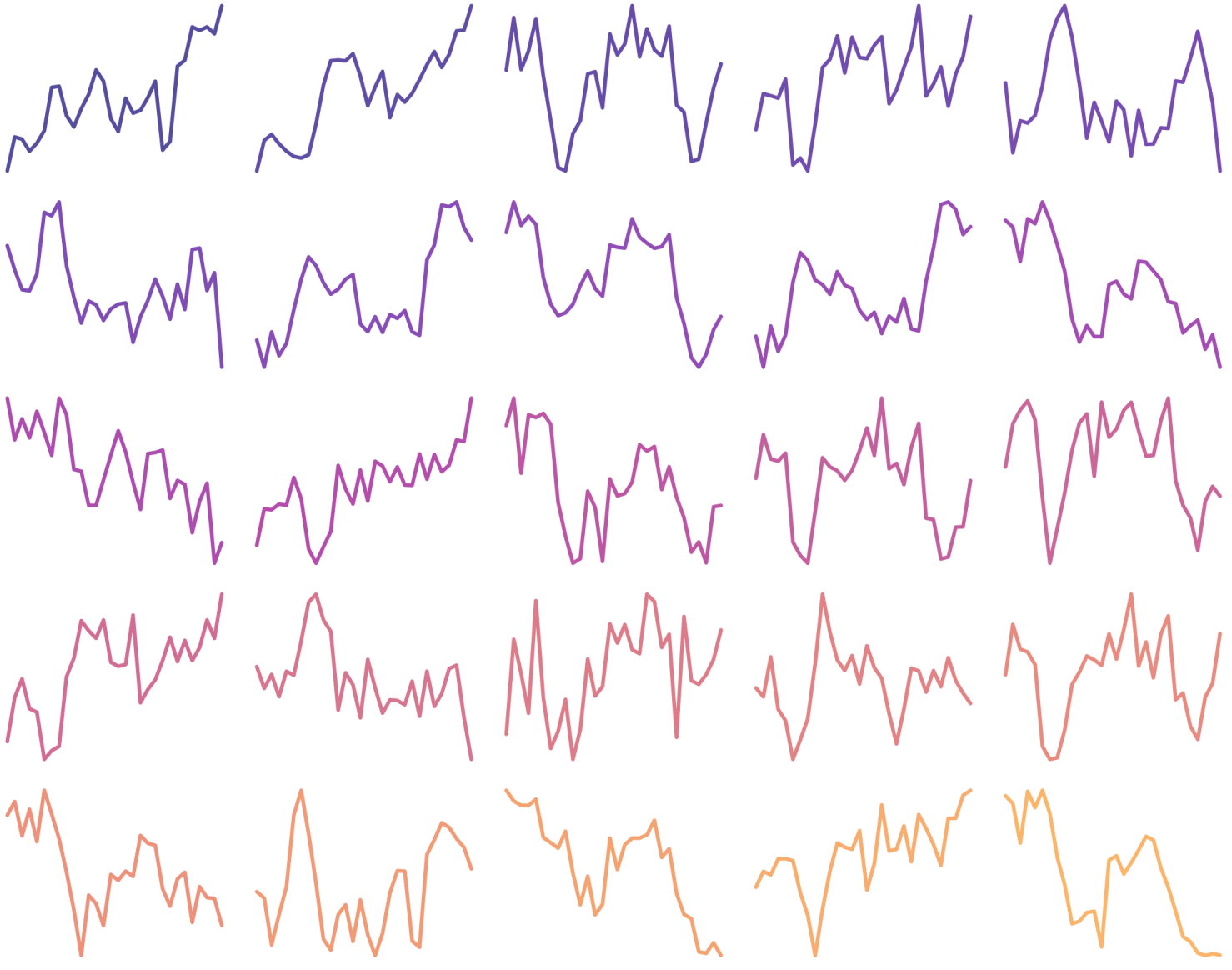
CHRISTMAS BIRD COUNT

How can we make sense of all of this?


Classification
History
Tragedy
Comedy

Ordination


PCA
PC2
PC1
perm
peri
area
NMDS
NMDS2
NMDS1
community_1
community_2
community_3
community_4
community_5
community_6



Just some combination of these

Dynamic Factor Analysis

State model

$$\text{Trends}_t = \text{Trends}_{t-1} + \text{errors}_t \qquad \text{(A few)}$$

Observation model

$$\text{Data}_t = \text{Loadings} \times \text{Trends}_t + \text{errors}_t \qquad \text{(Many)}$$

Salmon returns & productivity

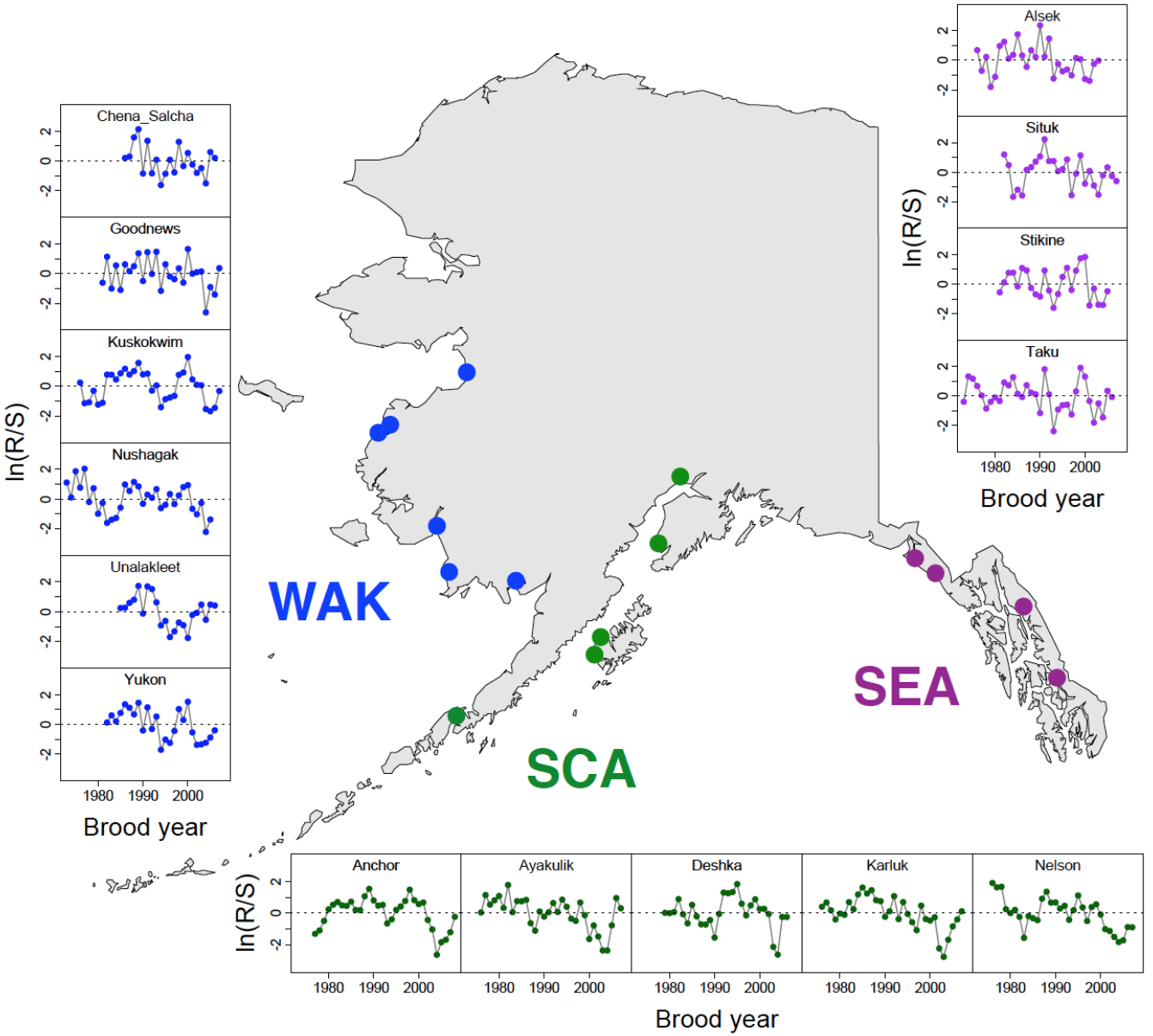
Trying to understand when & where stocks covary

Stachura et al. (2014) *CJFAS*

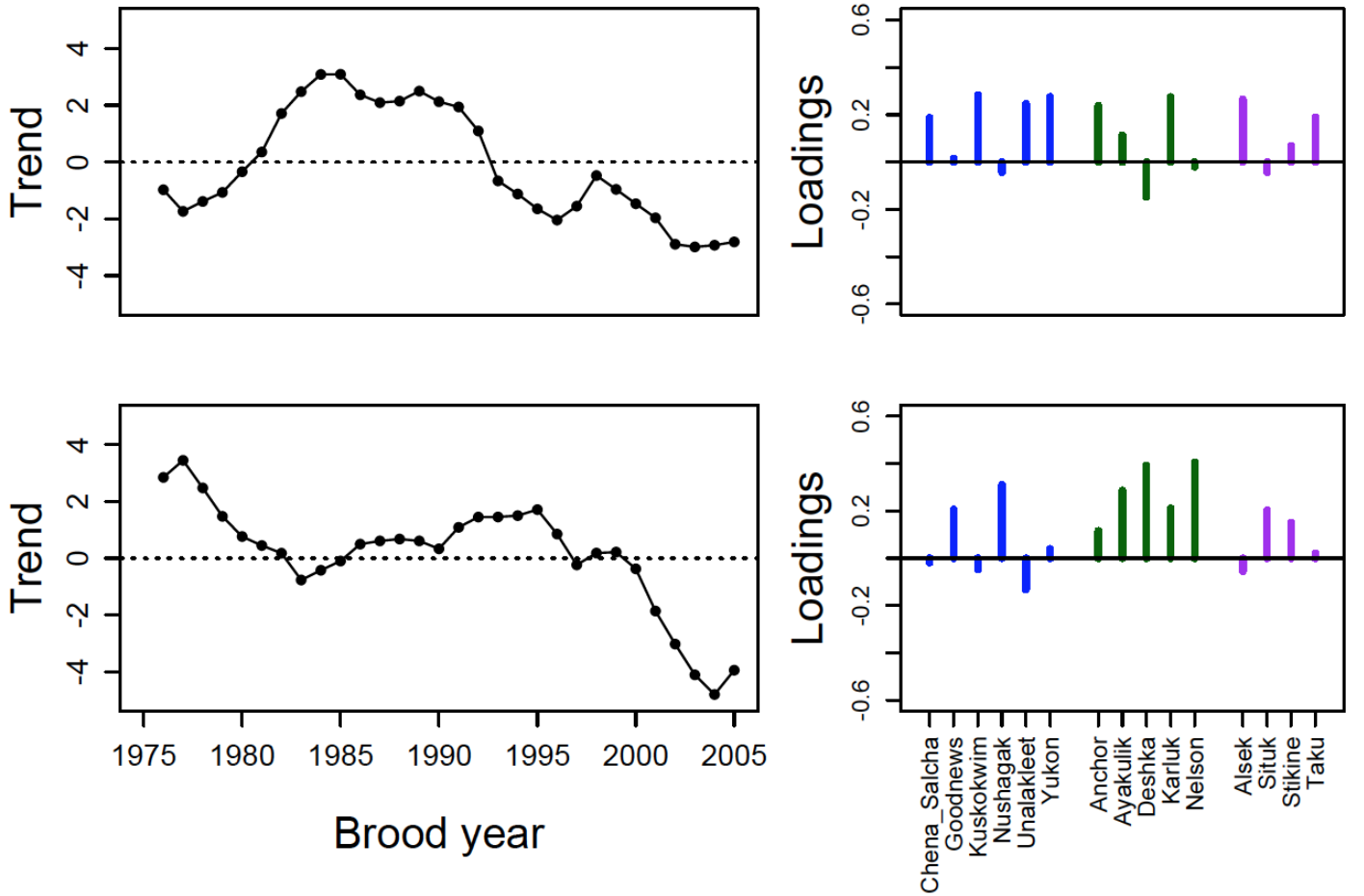
Goertler et al. (2016) *PLoS One*

Ohlberger et al. (2016) *Ecosphere*

Jorgensen et al. (2016) *Ecology & Evolution*


Chena_Salcha
Goodnews
Kuskokwim
Nushagak
Unalakleet
Yukon
ln(R/S)
1980 1990 2000
Brood year
WAK
SCA
SEA
Alsek
Situk
Stikine
Taku
ln(R/S)
1980 1990 2000
Brood year
Anchor
Ayakulik
Deshka
Karluk
Nelson
ln(R/S)
1980 1990 2000
Brood year

Two declining trends



Ohlberger et al. (2016) *Ecosphere* 7:e01333

WCGBTS: West Coast Groundfish Bottom Trawl Survey

1998-2018: 20 years of dependability, efficiency & excellence

NOAA FISHERIES
Northwest Fisheries Science Center

BIOLOGICAL DATA

- Age structures collected: 360,467
- Diet samples collected: 8,495
- Genetic samples collected
  - fin clip: 15,271
  - coral: 3,000
- Lengths taken: 1,800,324
- Maturity samples collected: 6,729

INTEGRATED ECOSYSTEM ASSESSMENT (IEA)

- Environmental data
- Food web
- Habitat mapping

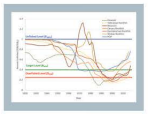
SURVEY LINKS

https://tinyurl.com/surveyteamGF
https://tinyurl.com/WCGBTS
https://tinyurl.com/surveyblogGF

VESSEL DATA

- Boats chartered: 15
- Captains: 26
- Days at sea: 2,727
- Miles traveled per year: ~2,400 per vessel
- Stations sampled: 11,805
- Volunteers: 245

MANAGEMENT & SCIENTIFIC CONTRIBUTIONS



- MSC certification
- New species: 2
- PFMC assessments: 52 (69% of total)
- Publications: 51
- Species range extensions: 5
- Recovered/rebuilt species: 8

NEAR BOTTOM OXYGEN

- Min O_2 = 0.02 ml/L
- Max O_2 = 5.96 ml/L

TOTAL FISH SALES

- 2017 pounds and dollars delivered:
  - 121,119 lbs
  - $68,461
- 2016 total revenue of groundfish fishery (not shrimp or crab): $62,517,873

EK-60

- Habitat mapping
- Max depth fished: 4920 ft
- Survey depth range: 180 - 4200 ft

CATCH DATA

- Fishing time: ~18 million seconds
- Pounds sampled: 9,748,073 lbs
- Sampled fish species: 1,073
- Sampled starfish species: 130

SEAFLOOR TEMPERATURE

- Min Seafloor temp = 2.9 °C
- Max Seafloor temp = 14.7 °C

Westport
Astoria
Newport: homeport *F/V* Excalibur, *F/V* Last Straw
Coos Bay: homeport *F/V* Ms. Julie
Brookings: homeport *F/V* Noah's Ark
Eureka
San Francisco
Santa Barbara
Long Beach
San Diego

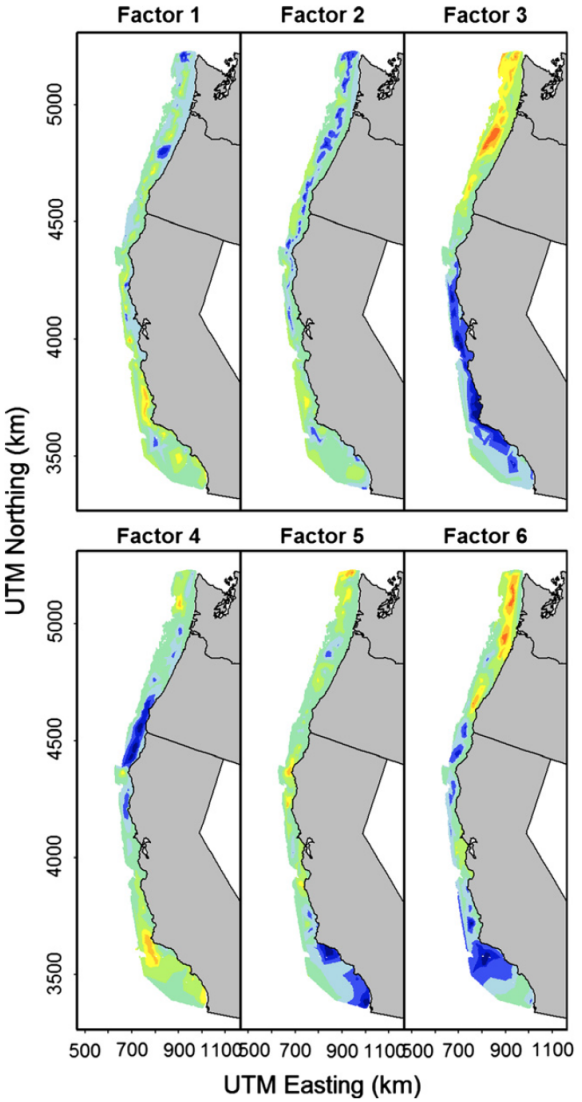
Spatial Factor Analysis

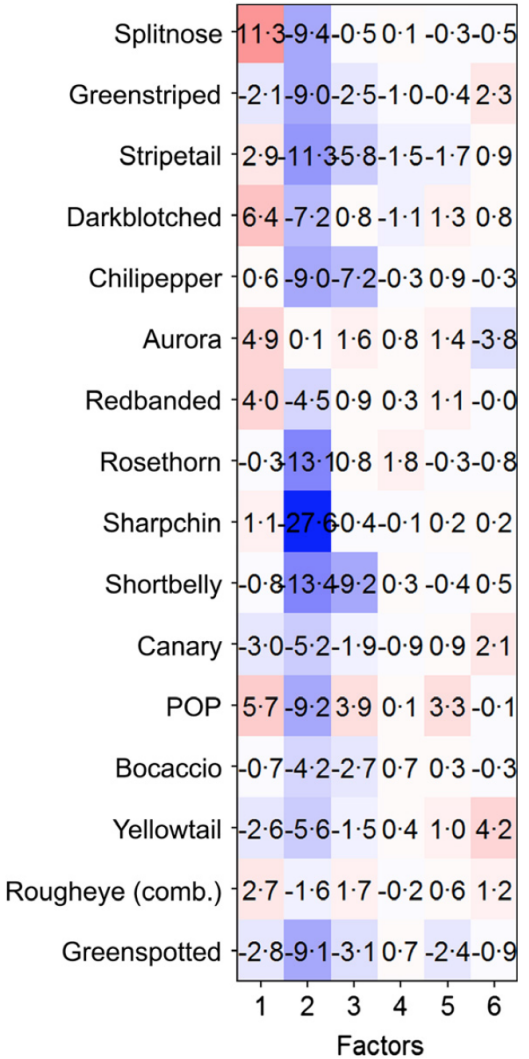
State model

Spatial maps$_i$

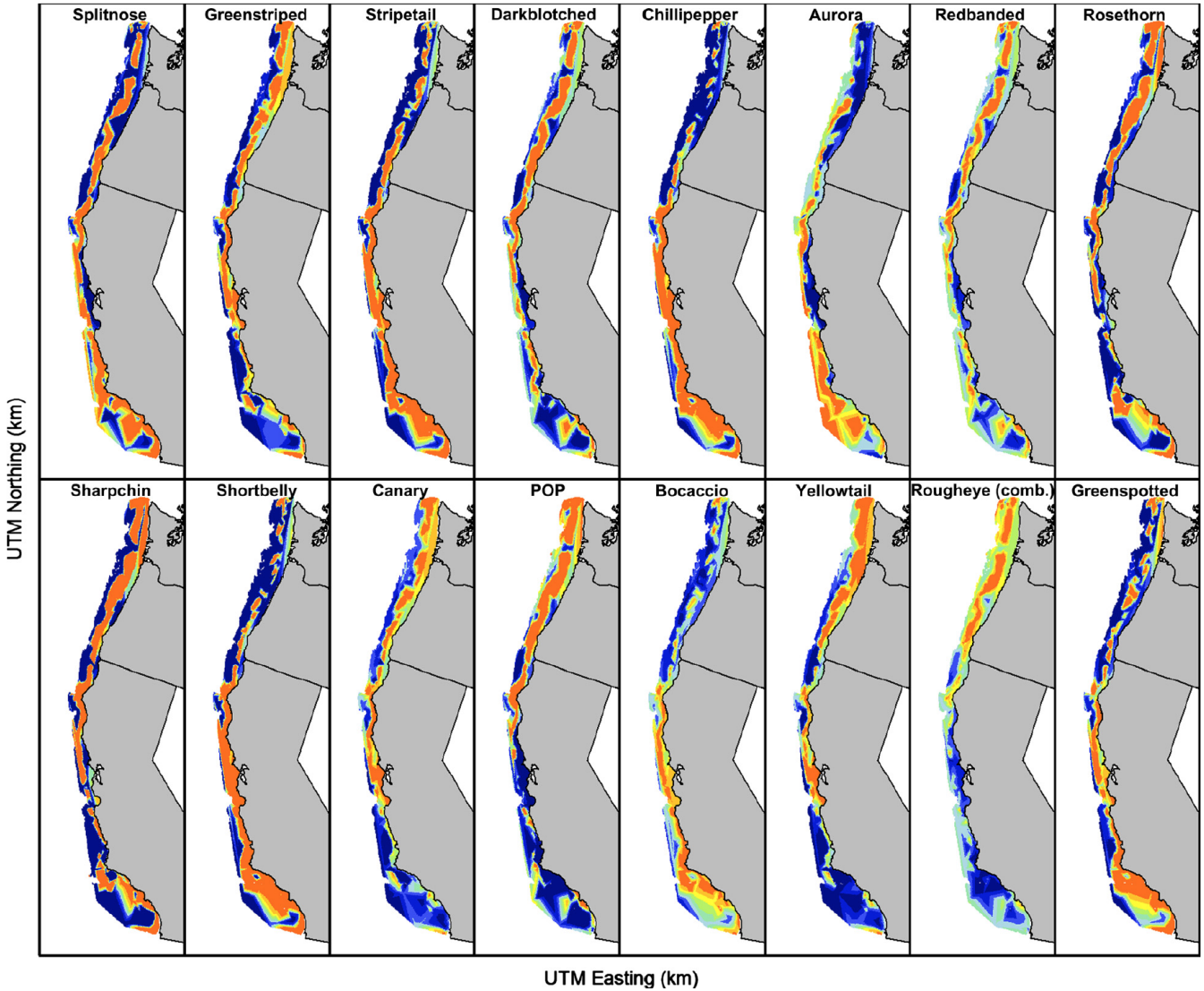
Observation model

$\text{Data}_i = \text{Loadings} \times \text{Spatial maps}_i + \text{error}_i$





Thorson et al. (2015) *Meth Ecol & Evol* 6:627–637



Thorson et al. (2015) *Meth Ecol & Evol* 6:627–637

Other applications of hierarchical models

Improving precision in species extinction forecasts

See & Holmes (2015) *Ecol Appl* 25:1157–1165

Other applications of hierarchical models

How do spatial patterns change over time?

Thorson et al. (2016) *Global Ecol & Biogeogr* 25:1144–1158

Other applications of hierarchical models

Relative importance of species vs spatial diversity in ecological portfolios

Thorson et al. (2018) *Proc Royal Soc B* 285:20180915

In summary

Much of our data is noisy and disparate

Hierarchical models offer a means for addressing these kinds of data

Software & hardware improvements open new doors

Slide deck

https://github.com/mdscheuerell/PSAW2/talk

Image sources

Drinking fountain: *Massachusetts Inst Tech* LIDAR: *NOAA*
DNA: *fizzgig (2016)*
Robin Hood: *John Escott*
M Caulkin: *20th Century Fox*
Carnival: *Frank Kovalchek (2010)*
Snake oil: *The Register*