

Challenges in understanding impacts of anthropogenic pressures on ecosystem recovery

Laura Uusitalo, PhD (fisheries science), PhD (computer science)

VELMU conference, 14 March 2023 in Helsinki, Finland



Suomen ympäristökeskus
Finlands miljöcentral
Finnish Environment Institute

Challenges in understanding

ecosystem recovery

Laura Uusitalo, PhD (fisheries science), PhD (computer science)

VELMU conference, 14 March 2023 in Helsinki, Finland



Suomen ympäristökeskus
Finlands miljöcentral
Finnish Environment Institute

Contents

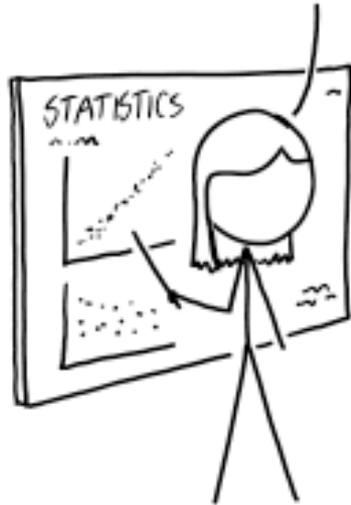
- Challenge 1: Understanding
- Challenge 2: Understanding ecosystems
- Challenge 3: Understanding ecosystem recovery
- Two more hopeful case studies: Tracking ecosystem change, Predicting recovered state
- Reflections and recommendations



Challenge 1: Understanding



IF YOU DON'T CONTROL FOR
CONFOUNDING VARIABLES,
THEY'LL MASK THE REAL
EFFECT AND MISLEAD YOU.

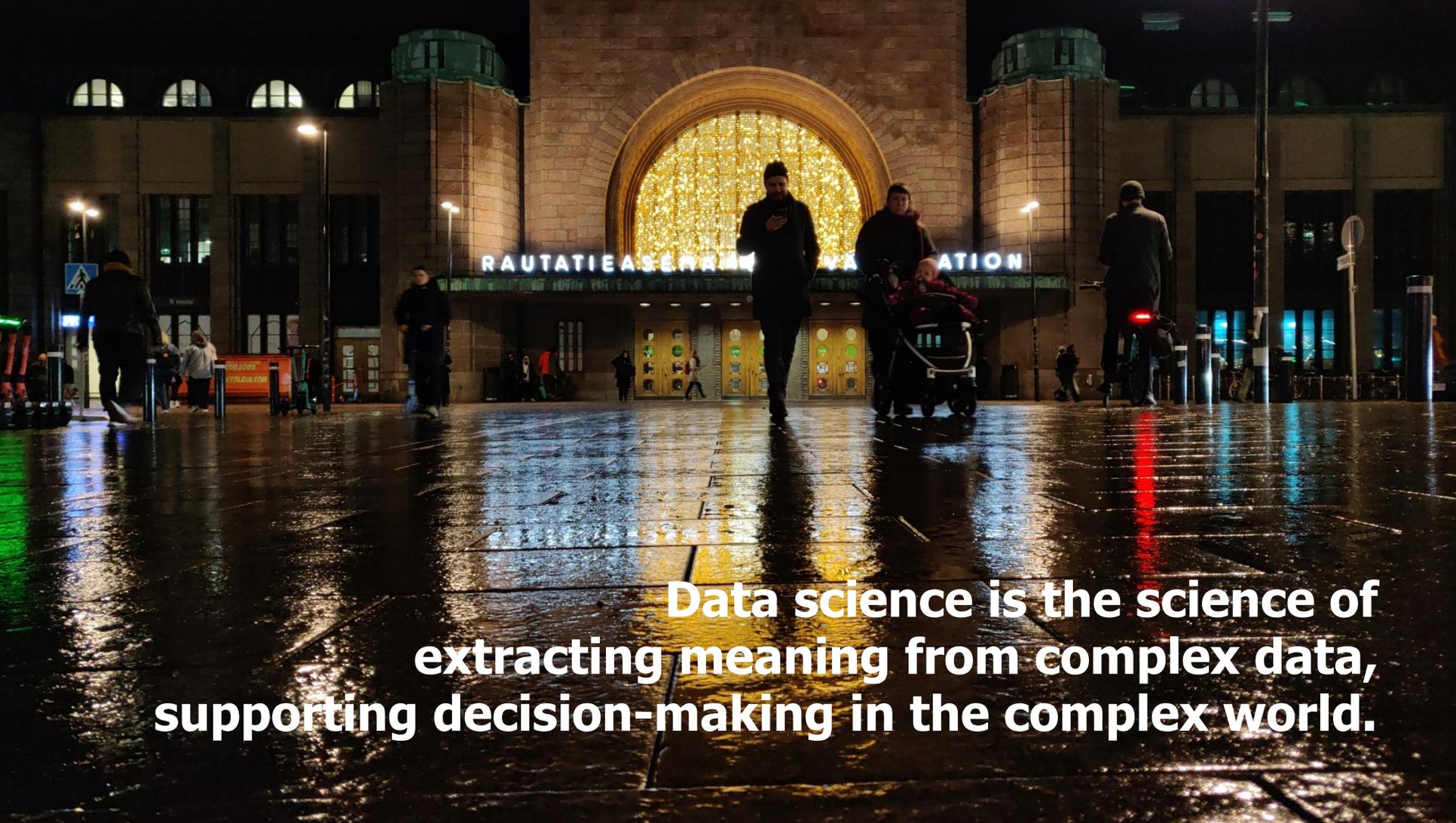


BUT IF YOU CONTROL FOR
TOO *MANY* VARIABLES,
YOUR CHOICES WILL SHAPE
THE DATA, AND YOU'LL
MISLEAD YOURSELF.



SOMEWHERE IN THE MIDDLE IS
THE SWEET SPOT WHERE YOU DO
BOTH, MAKING YOU DOUBLY WRONG.
STATS ARE A FARCE AND TRUTH IS
UNKNOWABLE. SEE YOU NEXT WEEK!





Data science is the science of extracting meaning from complex data, supporting decision-making in the complex world.

Humanity relies on ecosystems



**... but
environmental
data science is not
well-developed**

(Blair et al. 2019)

Image: TEEB Europe

Environmental data comes from a wide variety of sources

Field campaigns

Remote sensing

Autonomous monitoring systems

Citizen science

Historical records

Model outputs

Data mining from social media etc.

- Heterogeneous in
 - Temporal resolution
 - Spatial scales
 - Reliability
- Many crucial organisms / processes / variables may still be missing

Keep in mind: A note from critical data studies

- Data are always interpretation: different choices may be made due to context and tradition
- Data lose meaning without their context
 - What is context, what is data?
 - Relevant context is not self-evidently clear → never fully incorporated in metadata
- Data are influenced by software, hardware, protocols, documentation, social choices



Challenge 2:

Understanding ecosystems



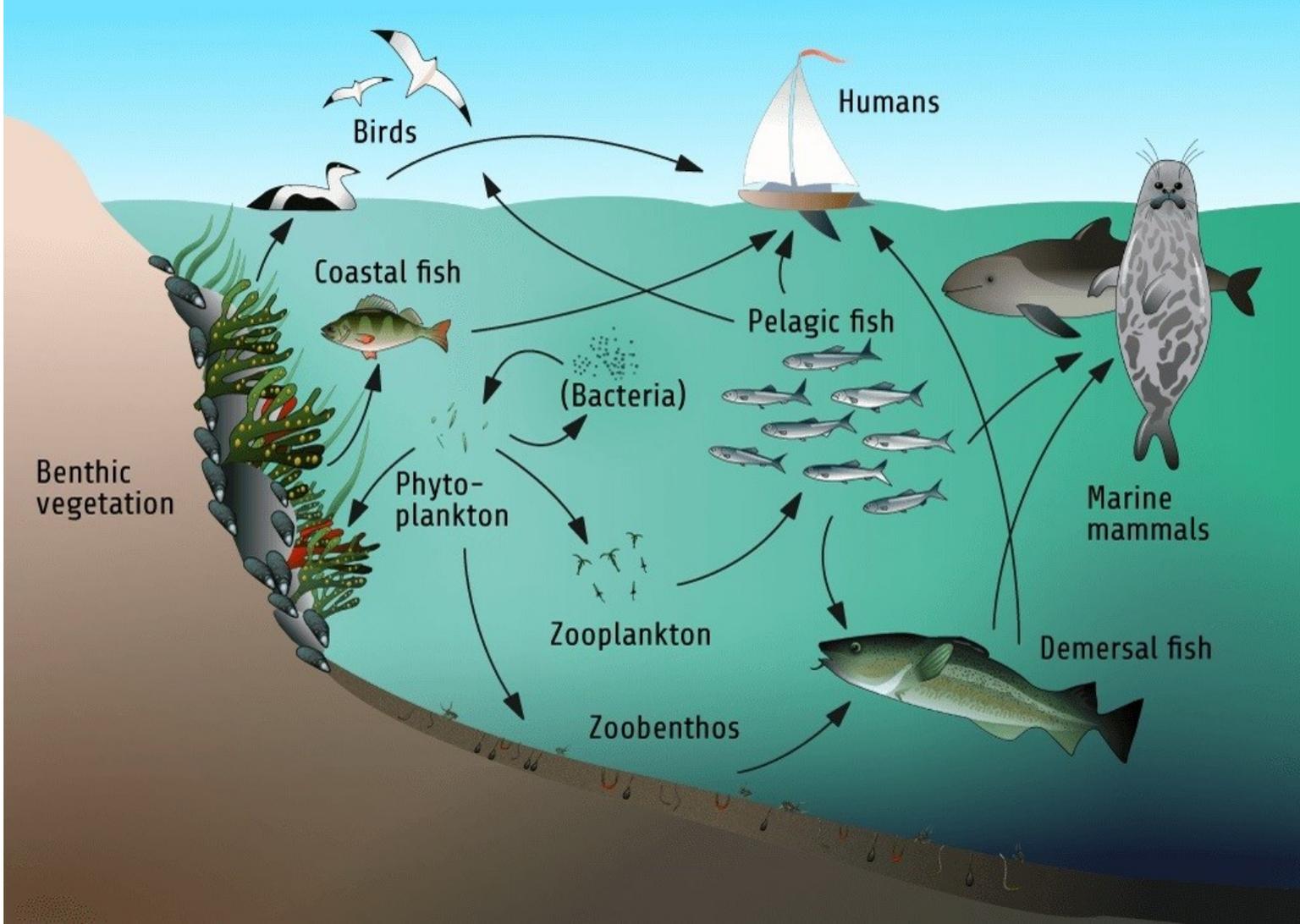


Image:
 HELCOM
 HOLAS II,
 Fig-1.3:
 Baltic Sea
 food web.
 Sebastian
 Dahlström.

Historical data: a fickle friend

Hydrobiologia (2013) 707:109–133

DOI 10.1007/s10750-012-1414-4

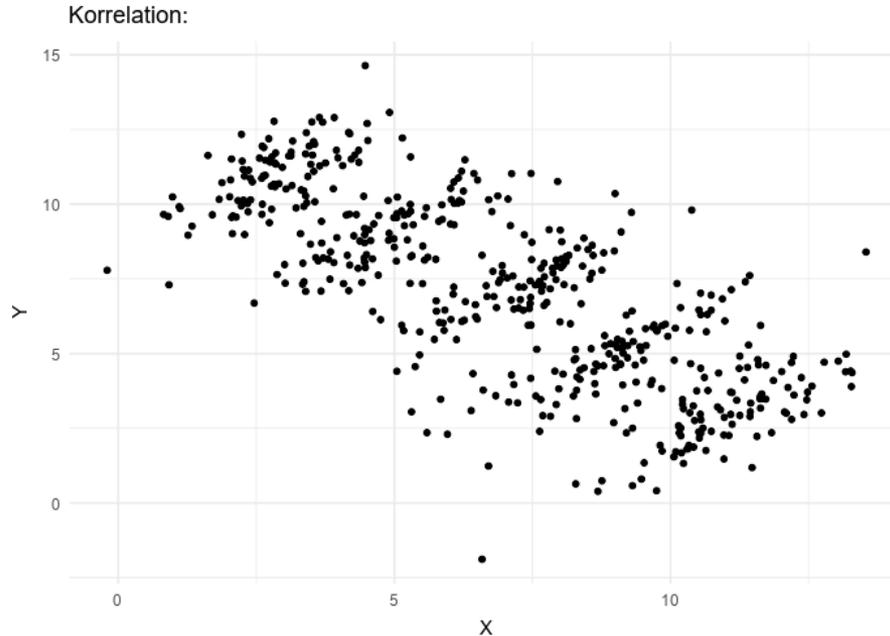
PRIMARY RESEARCH PAPER

The northern Baltic Sea phytoplankton communities in 1903–1911 and 1993–2005: a comparison of historical and modern species data

**Heidi Hällfors · Hermanni Backer ·
Juha-Markku Leppänen · Seija Hällfors ·
Guy Hällfors · Harri Kuosa**

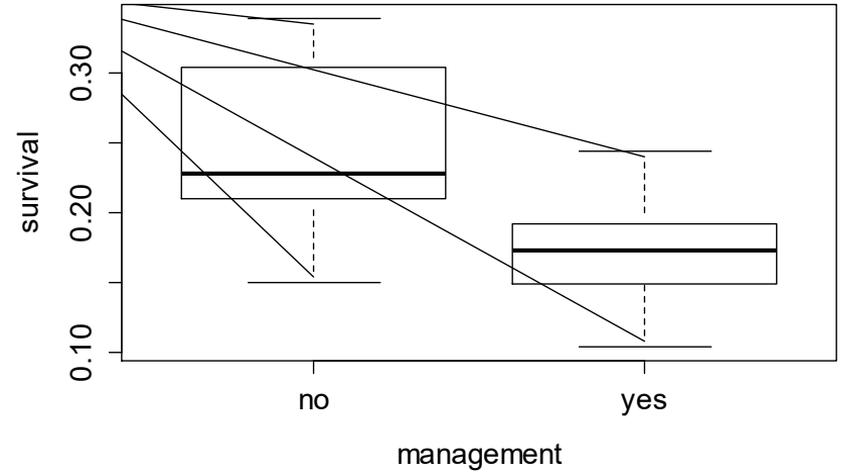
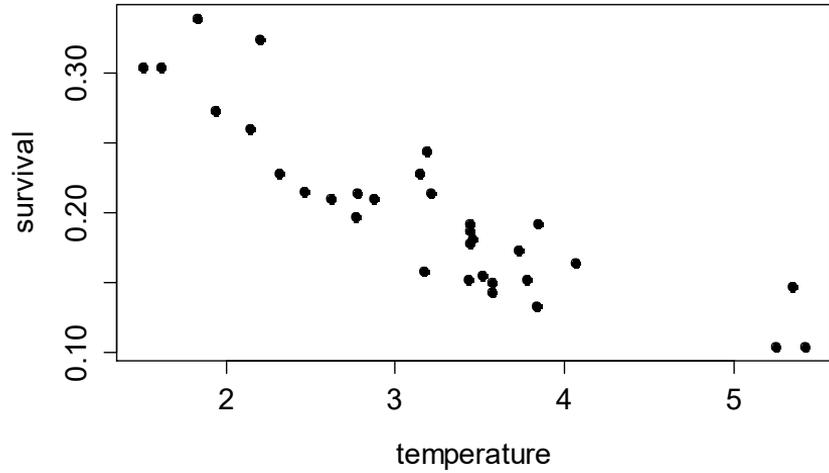
- Sampling equipment
- No information on preservation & analysis methods
- Nomenclature changes
- Identification issues

Correlations: Simpson's paradox

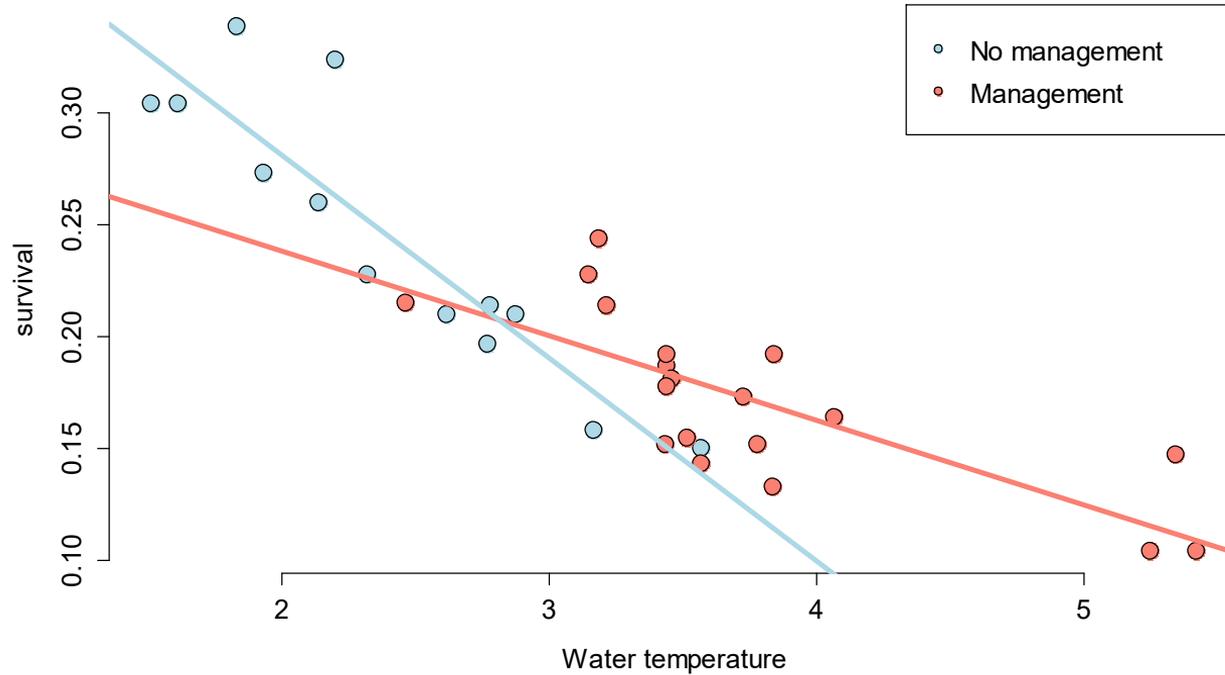


By Pace~swiki - Own work, CC BY-SA 4.0,
<https://commons.wikimedia.org/w/index.php?curid=62007681>

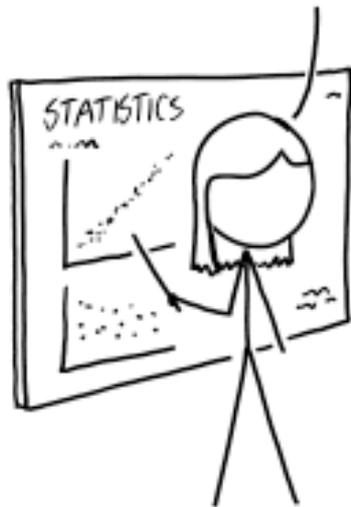
Simpson's paradox in ecology (1/2)



Simpson's paradox in ecology (2/2)



IF YOU DON'T CONTROL FOR
CONFOUNDING VARIABLES,
THEY'LL MASK THE REAL
EFFECT AND MISLEAD YOU.



BUT IF YOU CONTROL FOR
TOO *MANY* VARIABLES,
YOUR CHOICES WILL SHAPE
THE DATA, AND YOU'LL
MISLEAD YOURSELF.



SOMEWHERE IN THE MIDDLE IS
THE SWEET SPOT WHERE YOU DO
BOTH, MAKING YOU DOUBLY WRONG.
STATS ARE A FARCE AND TRUTH IS
UNKNOWABLE. SEE YOU NEXT WEEK!

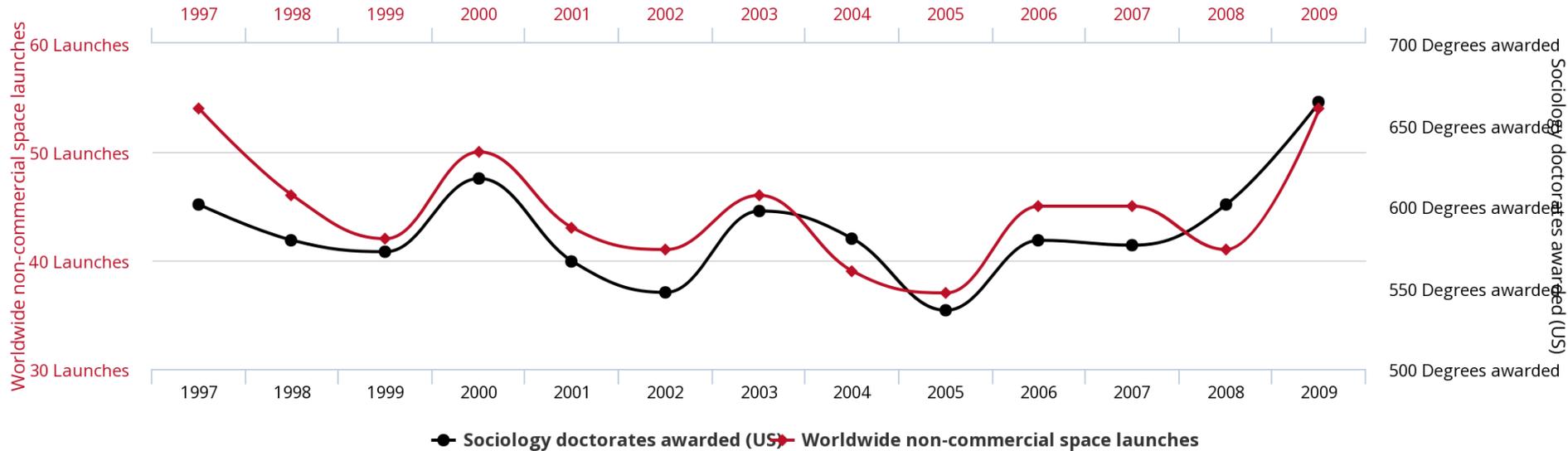


Spurious correlations

Worldwide non-commercial space launches

correlates with

Sociology doctorates awarded (US)



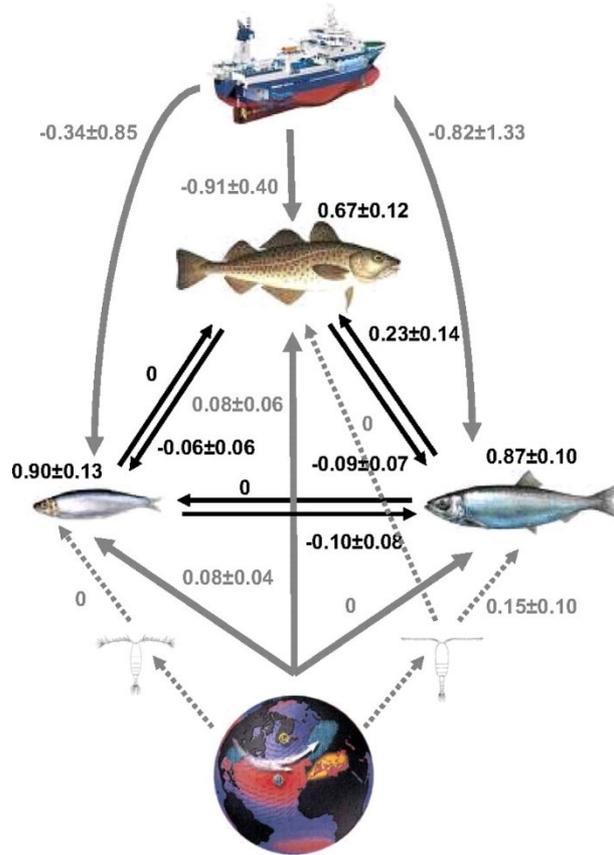
tylervigen.com

Challenge 3:

Understanding ecosystem recovery



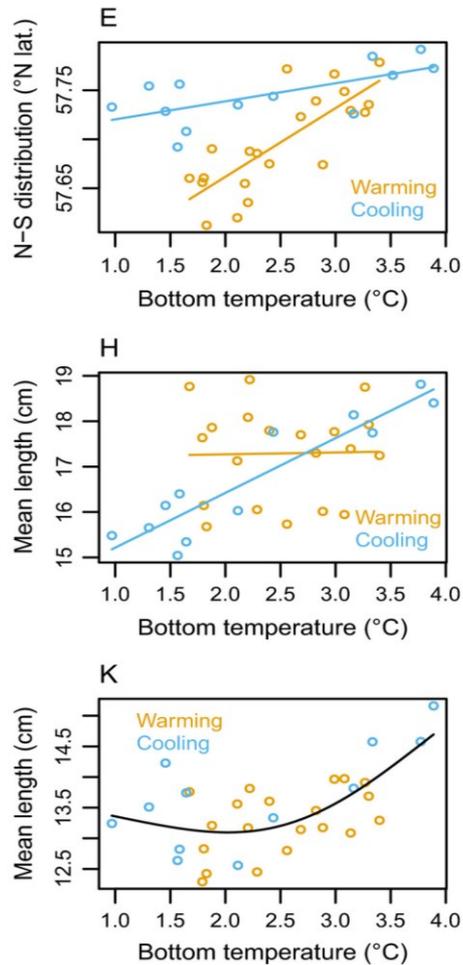
We try to understand ecosystems by modelling them



...but can we predict the future based on the past?

Image from:
Martin Lindegren, Christian Möllmann,
Anders Nielsen, and Nils C. Stenseth
PNAS August 25, 2009 106 (34) 14722-14727;
<https://doi.org/10.1073/pnas.0906620106>

©2009 by National Academy of Sciences



Hysteresis

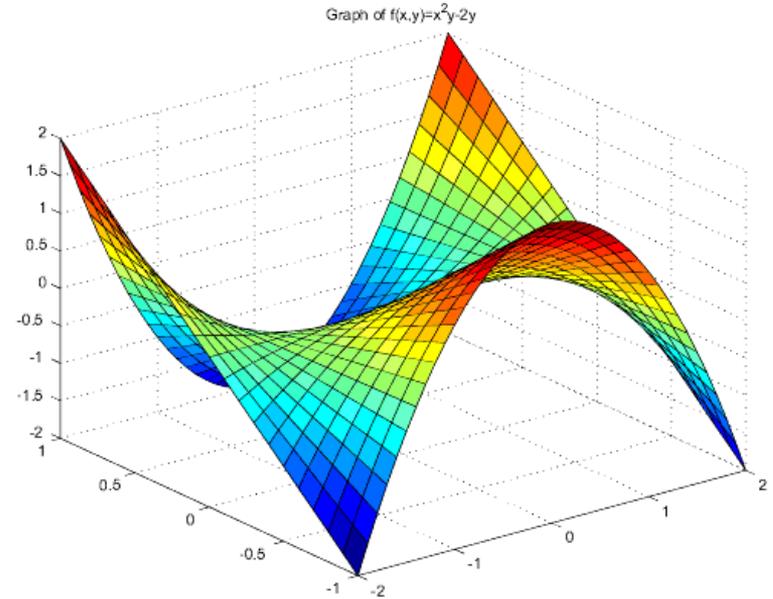
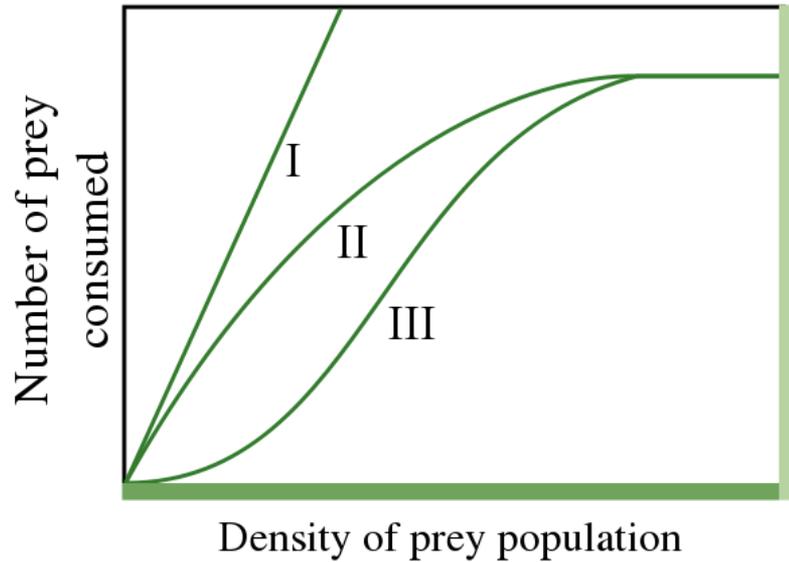
Response does not reverse its initial path after a perturbation is reversed

→ **Recovery does not follow the path of deterioration!**

May depend on feedback loops or different variables controlling the response in different states.

Formal tests for hysteresis are extremely difficult in observational systems (Schröder et al. [2005](#), Dudgeon et al. [2010](#), Faassen et al. [2015](#))

Ecosystem change

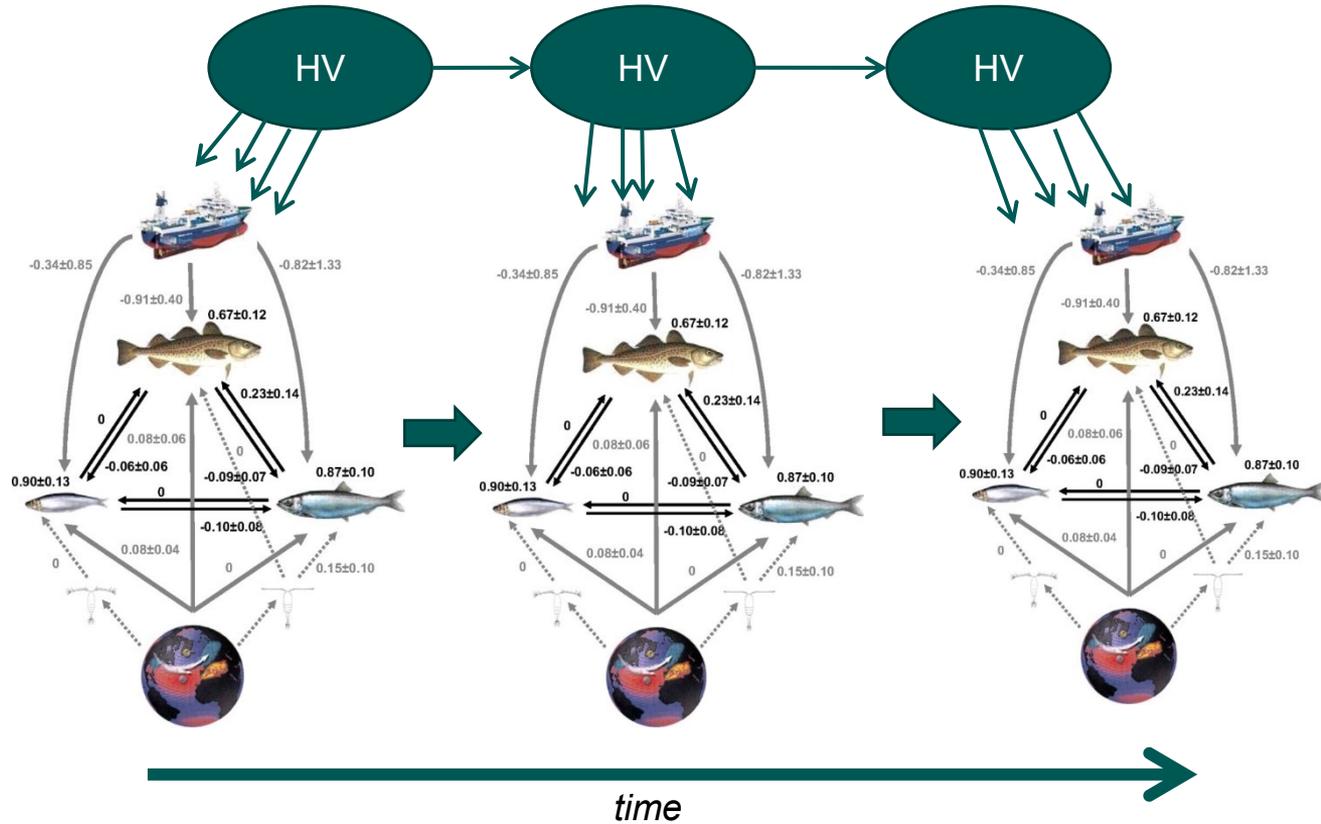


Case study: Tracking change

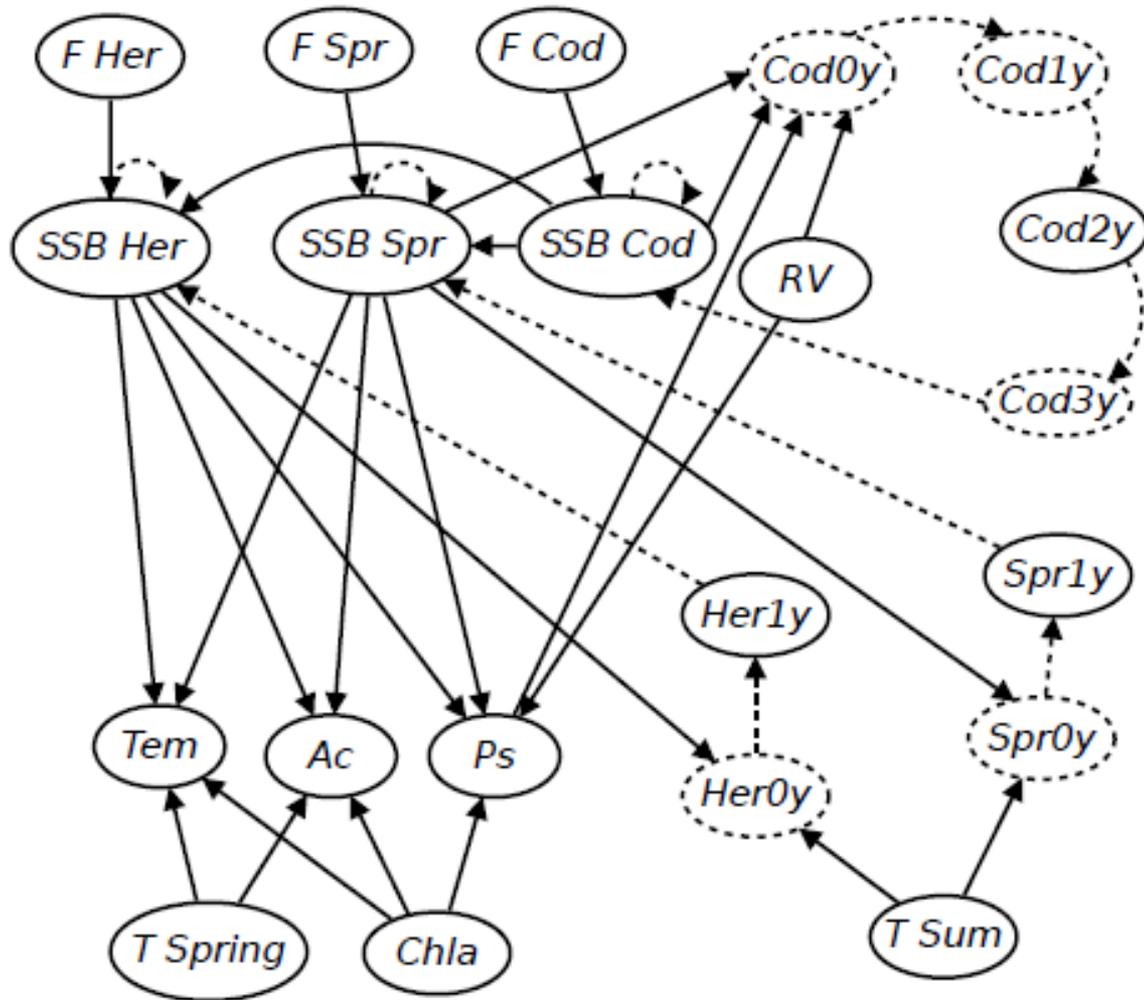
Uusitalo, L. Tomczak, M.T., Müller-Karulis, B., Putnis, I., Trifonova, N., Tucker, A. 2018. Hidden variables in a Dynamic Bayesian Network identify ecosystem level change. *Ecological Informatics* 45: 9-15.
<https://doi.org/10.1016/j.ecoinf.2018.03.003>

Maldonado, A.D., Uusitalo, L., Tucker, A., Blenckner, T., Aguilera, P.A., Salmerón, A. 2019. Prediction of a complex system with few data: Evaluation of the effect of model structure and amount of data with dynamic bayesian network models. *Environmental Modelling & Software* 118: 281-297. <https://doi.org/10.1016/j.envsoft.2019.04.011>



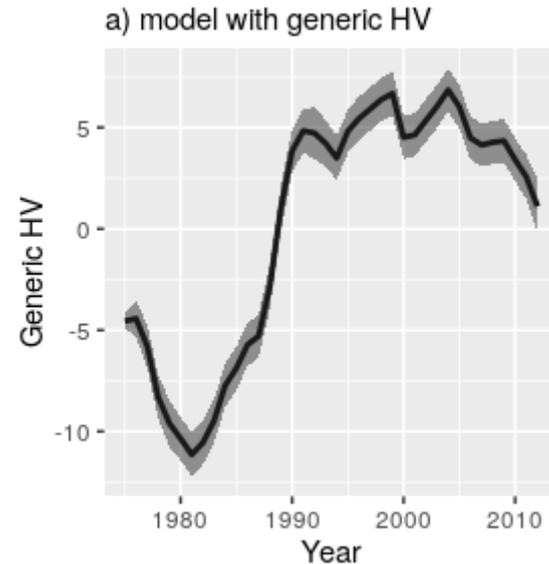


Food web image from:
 Martin Lindegren, Christian Möllmann, Anders Nielsen, and Nils C. Stenseth
 PNAS August 25, 2009 106 (34) 14722-14727; <https://doi.org/10.1073/pnas.0906620106>
 ©2009 by National Academy of Sciences



The hidden variables robustly detected the systemic change

- The result did not change with the exact model setup or when adding more data
- Hidden variable can be continuous (as in here) or discrete (alternate stable states)
- A lot of assumptions go also into this model...
- We need to be very careful not to under- or overparameterise!



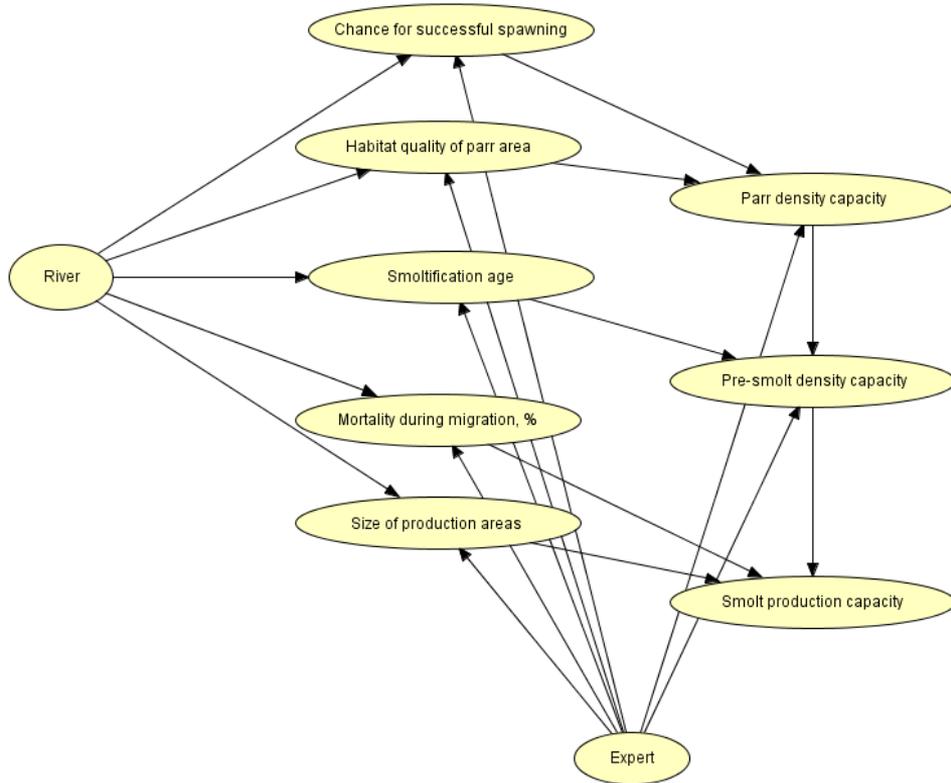
Case study: Predicting recovered state

- Uusitalo, L., Kuikka, S., Romakkaniemi, A. 2005. Estimation of Atlantic Salmon Smolt Carrying Capacity of Rivers Using Expert Knowledge. ICES Journal of Marine Science 62(4):708-722.
<https://doi.org/10.1016/j.icesjms.2005.02.005>

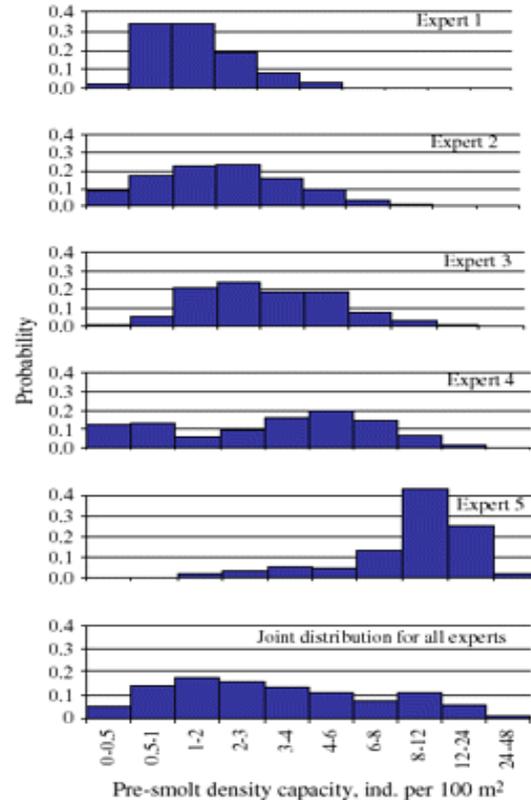


How many salmon can the rivers rear?

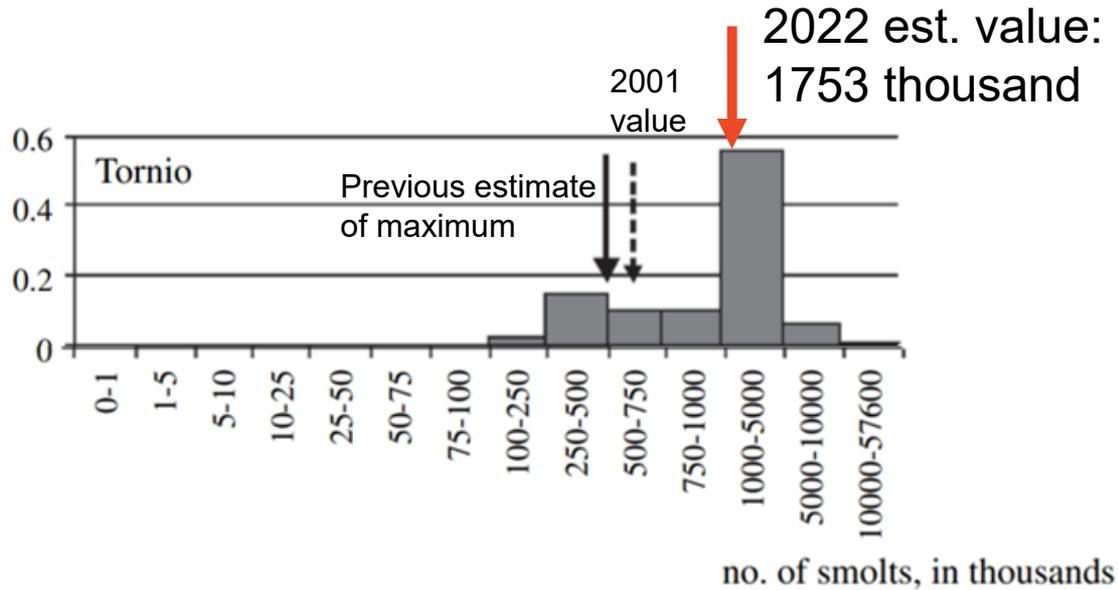
- Task: Estimate how many salmon juveniles could, at maximum, be reared in each of the rivers
- Problem: the maximum had not been reached during historical records
 - Extrapolating using equations was practically impossible
- Solution: Expert judgement!



The experts disagreed...



20 years later:

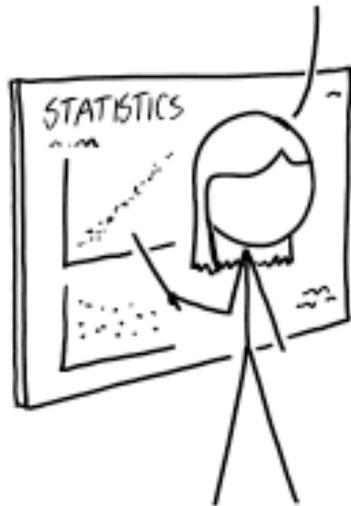


Reflections and recommendations

- Ecological knowledge and data analysis are both needed, and **collaboration** between experts in these domains is crucial!
- We need to be careful not to over- or underparameterise our models!
- Keep in mind what isn't in the data: Variables, habitats, context of the data collection, ...



IF YOU DON'T CONTROL FOR
CONFOUNDING VARIABLES,
THEY'LL MASK THE REAL
EFFECT AND MISLEAD YOU.



BUT IF YOU CONTROL FOR
TOO *MANY* VARIABLES,
YOUR CHOICES WILL SHAPE
THE DATA, AND YOU'LL
MISLEAD YOURSELF.



SOMEWHERE IN THE MIDDLE IS
THE SWEET SPOT WHERE YOU DO
BOTH, MAKING YOU DOUBLY WRONG.
STATS ARE A FARCE AND TRUTH IS
UNKNOWABLE. SEE YOU NEXT WEEK!



Reflections and recommendations

- Ecological analyses are often used in management
 - → they affect people's opportunities (income, well-being)
 - → **Transparency and accountability are needed throughout the process!**
 - Systems have to be understandable, explainable, and verifiable
- **Clear communication about starting points, assumptions**



Thank you!

