PICES-2026

Actionable Science for a Changing Ocean

Oct 24-30, 2026 | Nanaimo, Canada

Exploring capelin (*Mallotus villosus*) biomass dynamics on the Newfoundland Grand Banks

PhD. student

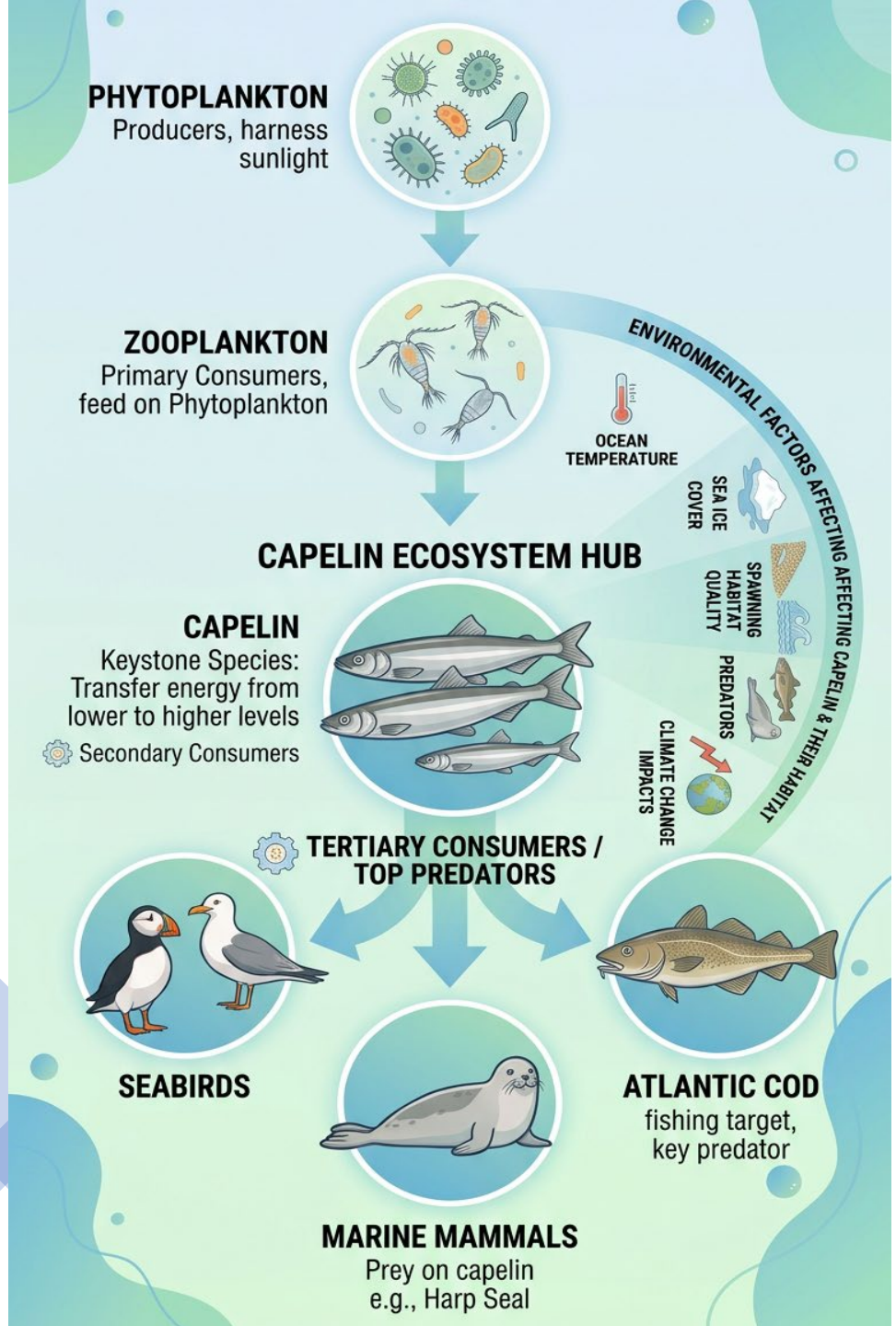
Jiaying Chen

Supervisor

Dr. Tyler Eddy

Dr. Jin Gao (DFO)





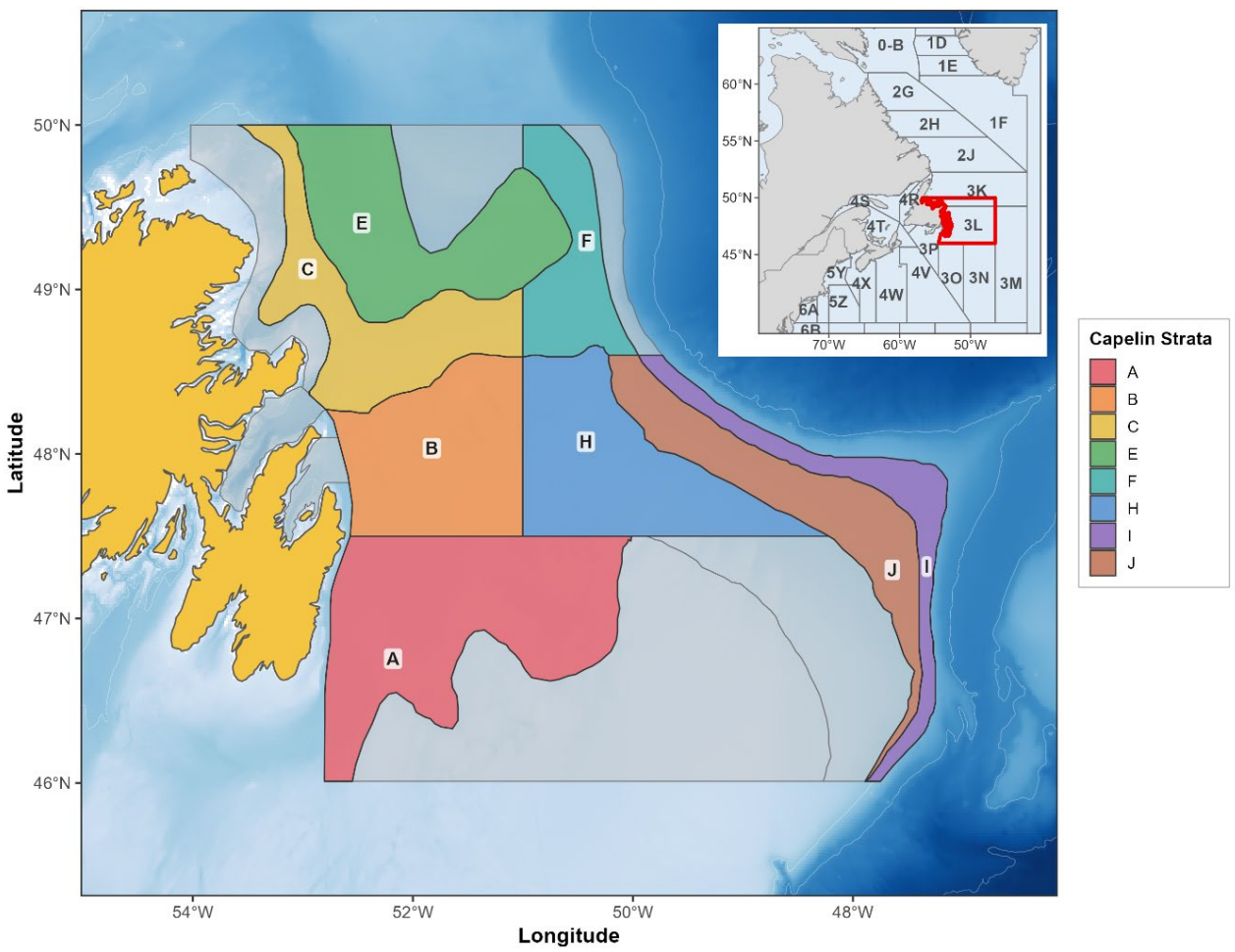


Atlantic capelin (*Mallotus villosus*)

- ❑ **Short-lived forage fish**, rarely living beyond 5 years
- ❑ **Matures at age 2 or 3,** with spawning populations now dominated by age-2 and age-3 fish
- ❑ **Key forage species**, linking zooplankton to predators, such as cod, seabirds, and marine mammals
- ❑ **Environmentally influenced**, with biomass dynamics associated with temperature, sea ice, wind, and other oceanographic conditions
- ❑ **Good candidate for empirical dynamic modelling** the population turns over quickly and it responds nonlinearly to the environment

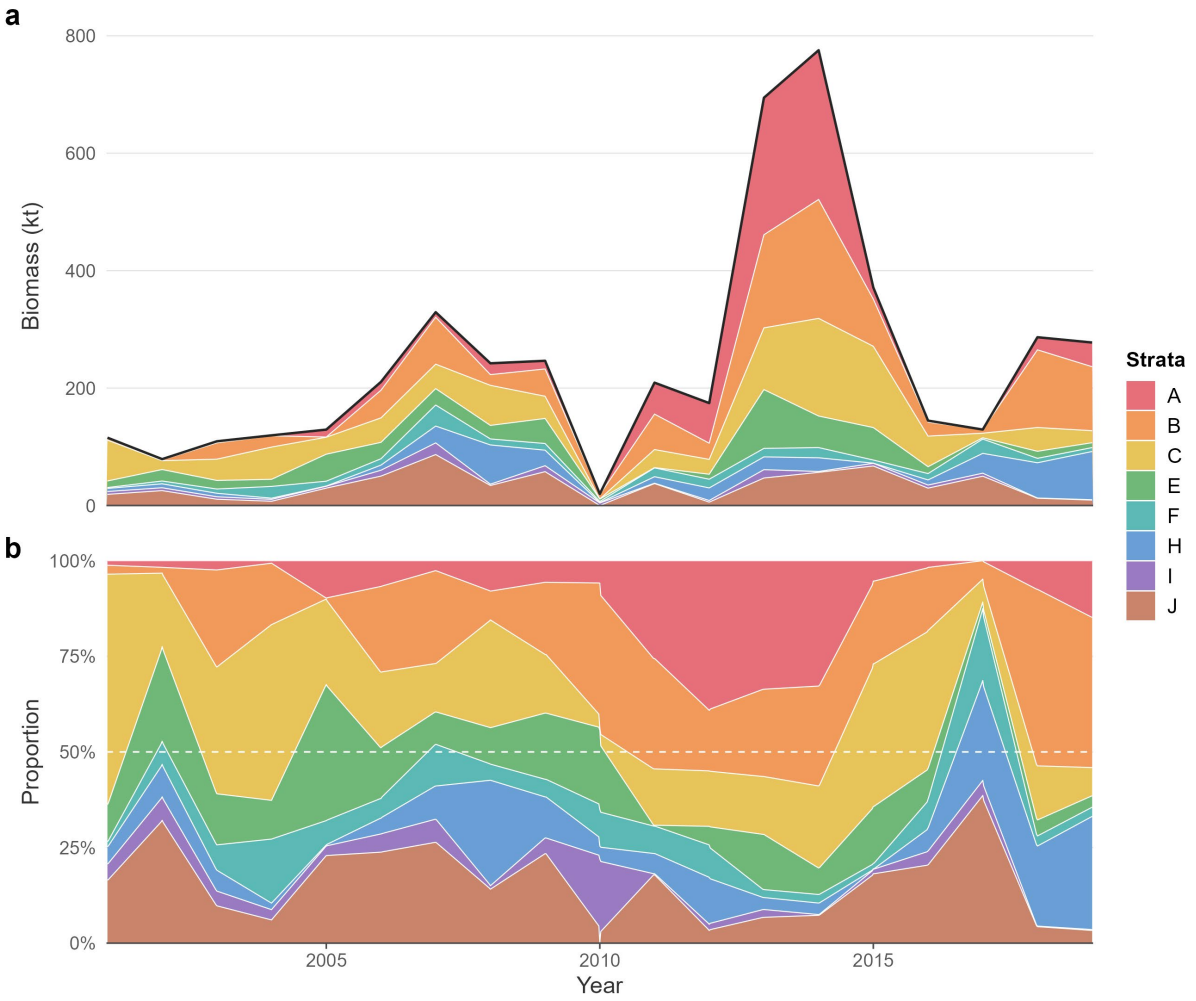
Objective: to explore which variables are relevant to capelin biomass dynamics

Study region



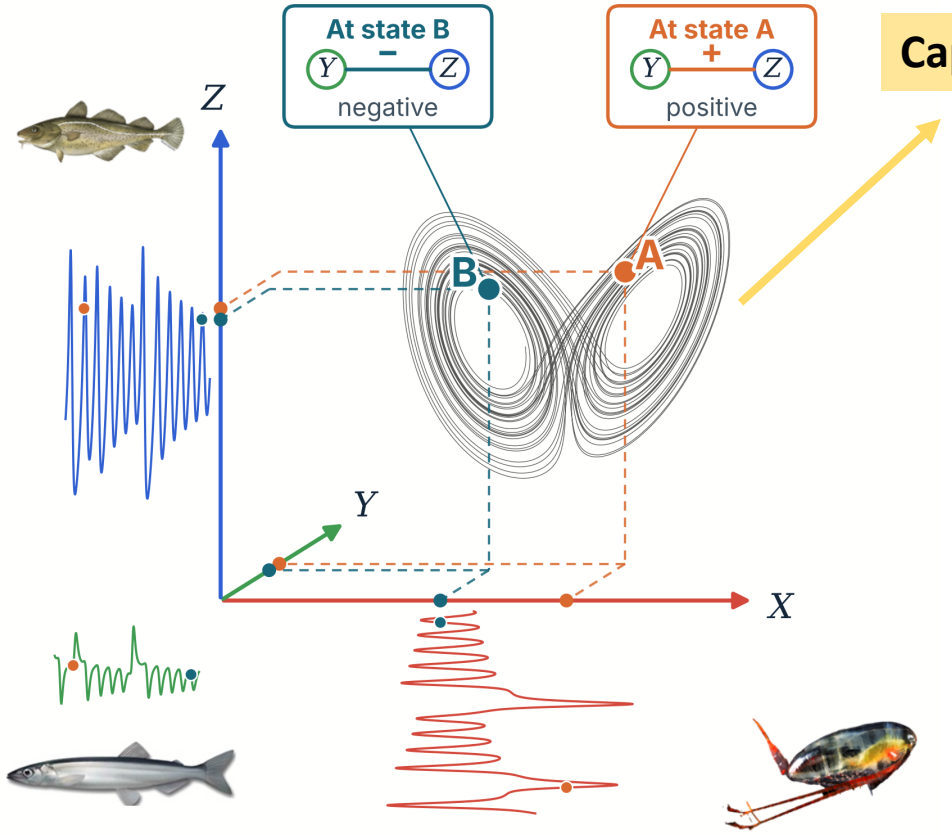
NAFO Division 3L, Newfoundland Grand Banks

Capelin Biomass

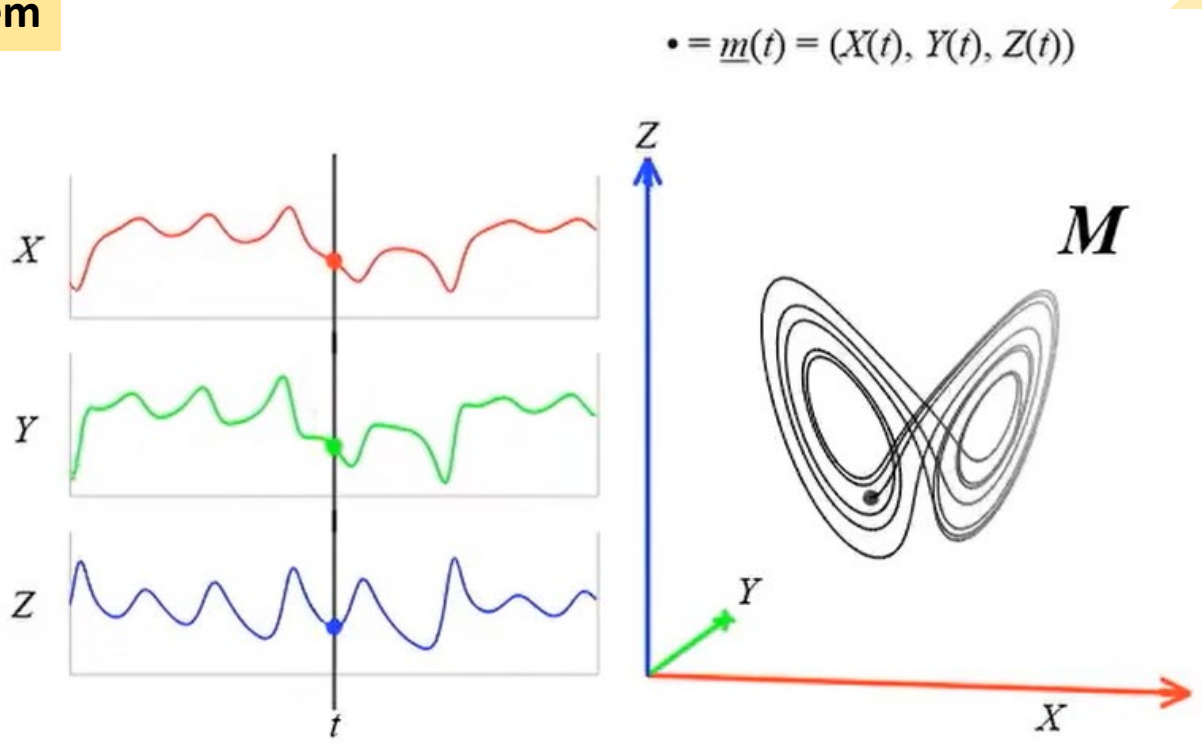


- **Data source:** DFO spring acoustic survey
- **Age composition:** Primarily age-2 and age-3 capelin
- **Study period and coverage:** 2001–2019; 8 strata

Multivariate Spatial Empirical Dynamic Modelling via Gaussian Process regression (GP-sEDM)



Capelin system



Based on Takens's embedding theorem

(Sugihara Lab)

- **Empirical Dynamic Modelling :** Equation-free, data-driven, state-space reconstruction
- **Multivariate:** Multiple variables and lags → reconstructed state
- **Spatial:** Incorporate information from **spatially neighboring strata**
- **Gaussian Process:** Flexible, similarity-based forecasting framework, uncertainty estimation

Spatially Varying Variables

Capelin
Capelin historical dynamics indices

Spring Lag 1-4



Univariate GP-sEDM

Biological variables

Calanus finmarchicus (Copepod)
Key prey indices

Spring Lag 0 Fall Lag 1

Atlantic Cod
Key predator indices

Fall Lag 1-2

Environmental variables

Sea Surface & Bottom Temp
SST and Bottom Water Temperature Indices

Spring Lag 0 Annual Lag 1

Sea Water Salinity
Salinity indices

Spring Lag 0 Annual Lag 1

Sea-Ice Area Fraction
Ice cover dynamics indices

Spring Lag 0 Annual Lag 1

Chlorophyll-a
Primary bloom amplitude & timing indices

Spring Lag 0 Annual Lag 1

Nutrients (Nitrate & Silicate)
Nutrient availability indices

Spring Lag 0 Annual Lag 1

Spatially Constant Variables

Biological variables

Atlantic Harp Seal

Key Predator Indices

Annual Lag 1-2

Environmental variables

CIL Area

Cold intermediate layer area indices

Annual Lag 0 Annual Lag 1

Sea ice duration and area

Season duration, and maximum ice cover area indices

Annual Lag 0 Annual Lag 1

Time of Sea Ice Retreat

Sea ice retreat day-of-year indices

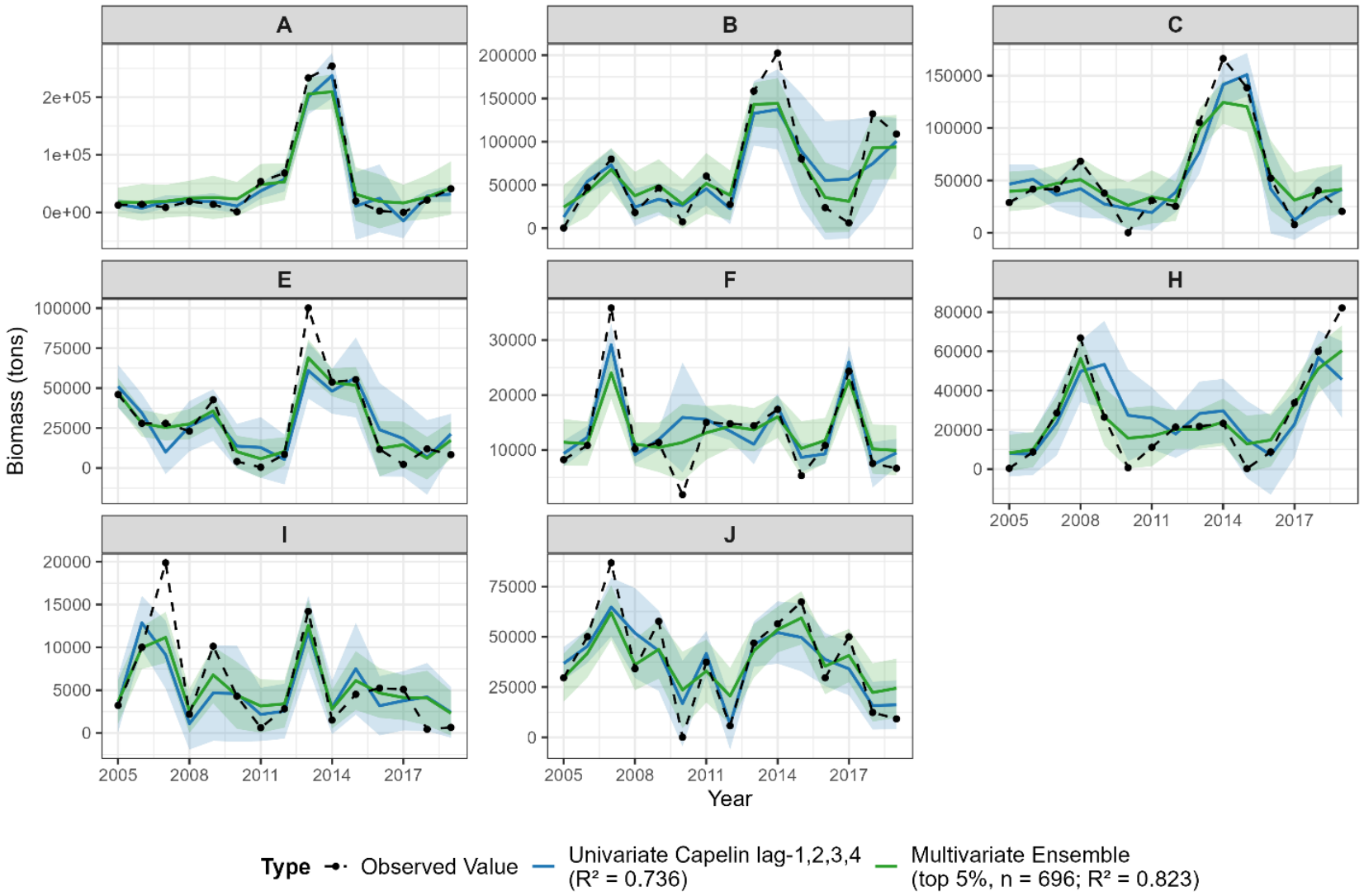
Annual Lag 0 Annual Lag 1

Winter NAO

North Atlantic Oscillation Large-scale indices

Annual Lag 1

Multivariate ensemble improves forecast over univariate (capelin alone)



How the ensemble was built

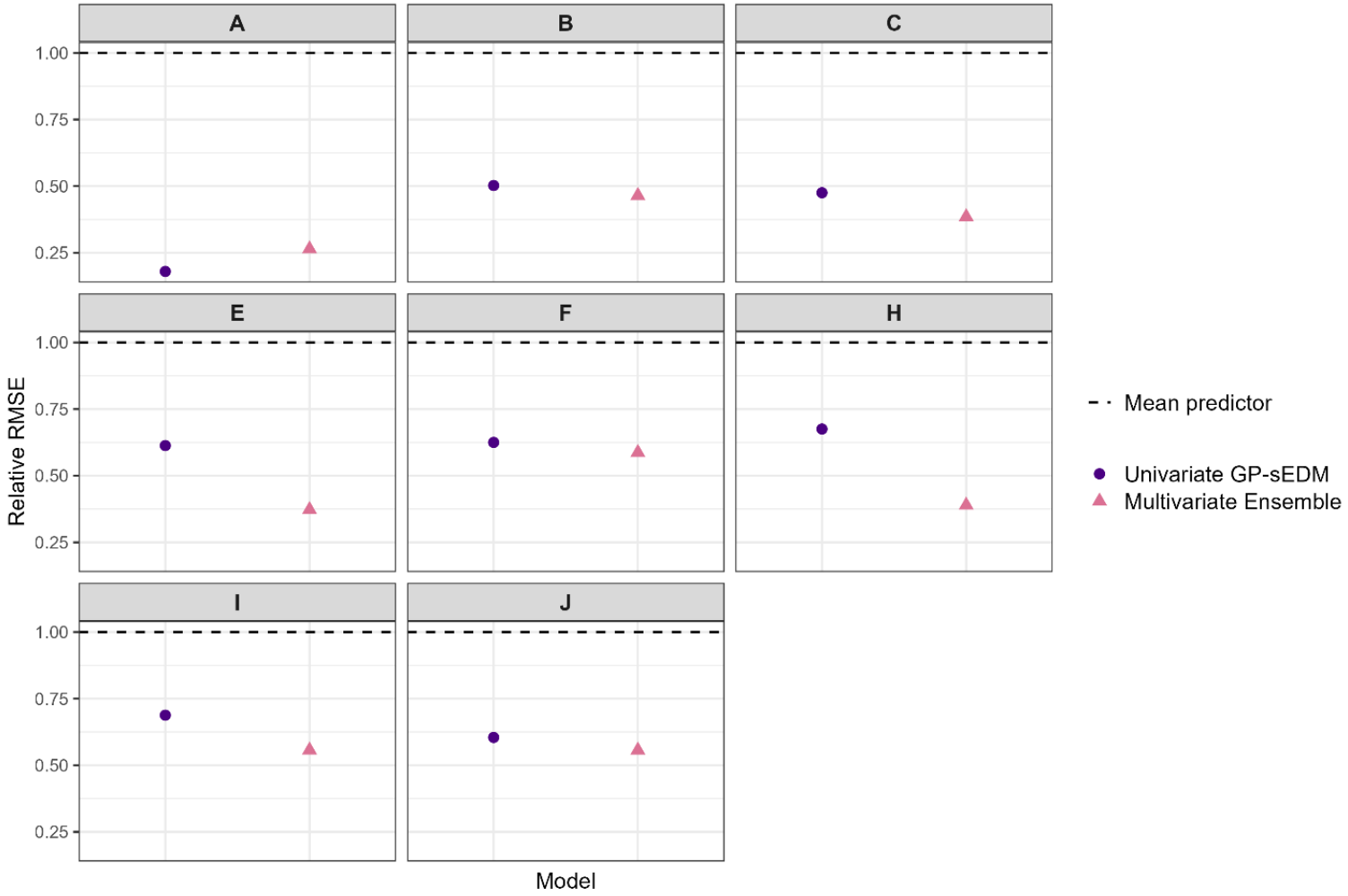
1. **13,915 four-variable models**
   each includes at least one capelin lag
2. **Top 5% by leave-one-out R^2**
   n = 696 views
3. **R^2-weighted ensemble**
   predictions and uncertainty combined

0.736
Capelin lags 1–4 only

0.823
Ensemble of top 5%

Leave-one-out R^2, 8 strata, 2005–2019

Multivariate ensemble lowers error in 7 of 8 strata



8/8 **Beat the mean**
relative RMSE < 1, both models

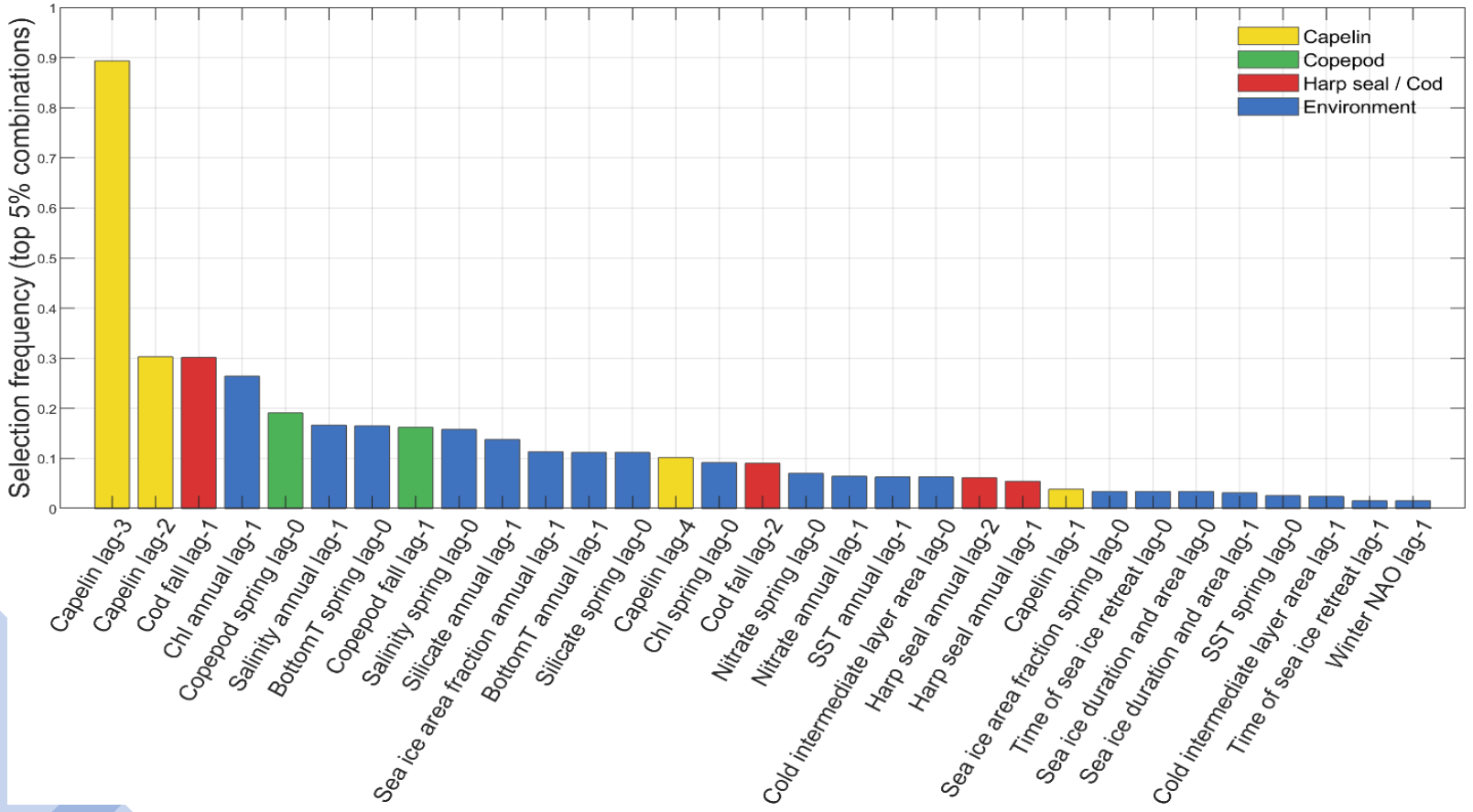
7/8 **Ensemble lower**
error than capelin lags alone

Exception

A: capelin lags alone do better (0.18 vs 0.27)

Relative RMSE = RMSE of the leave-one-out predictions / RMSE of predicting the stratum mean.

Relevance: selection frequency in the top 5% models



0.89 Capelin lag-3

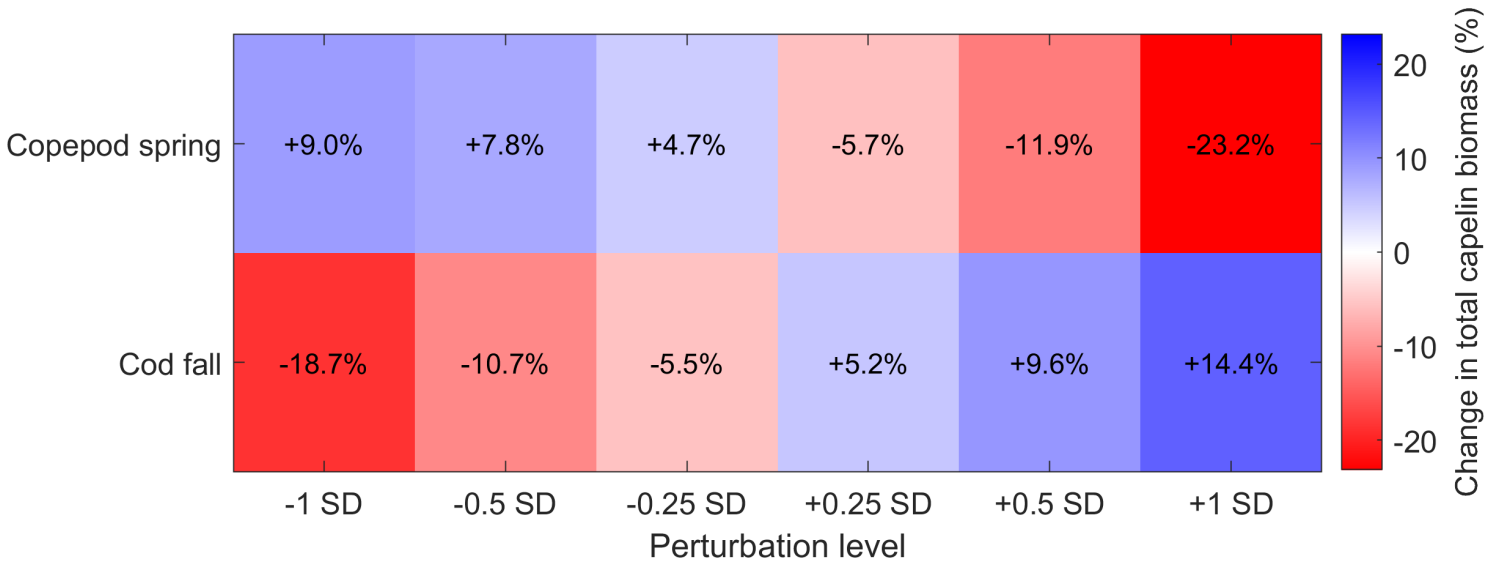
0.303 Capelin lag-2

0.301 Fall cod lag-1

0.26 Chlorophyll-a lag-1

0.19 Copepod spring lag-0

Scenario exploration: predicted capelin response to copepods and cod



−23.2% **Spring copepods +1 SD**
total capelin biomass

+14.4% **Fall cod +1 SD**
total capelin biomass

- ❑ Spring copepods are measured in the same year as capelin: the negative response is consistent with capelin predation reducing copepod abundance.
- ❑ Cod feed on capelin: the positive response is consistent with cod tracking capelin biomass.

View [2 3 6 7]: capelin lags 2 and 3, spring copepods (lag 0), fall cod (lag 1).

Take-home messages

1 Long records: forecast next year

If your ecological time series is long enough and its dynamics are roughly stationary, GP-EDM can address **the million-dollar** question: how many fish will there be next year?

2 Short records: add spatial information

Spatial GP-EDM can incorporate information from nearby locations that share similar dynamics to improve the forecast.

3 Potential drivers: identify and improve

Multivariate GP-EDM can identify the relevance of potential drivers and improve forecasts when those drivers carry real information.

4 Capelin: own history matters most

Capelin biomass three and two years earlier, carried most of the information; environmental covariates added only small gains.

Thank you!

Email: Jiaying.Chen@mi.mun.ca

