Species Distribution Models and Applications to Fisheries Management

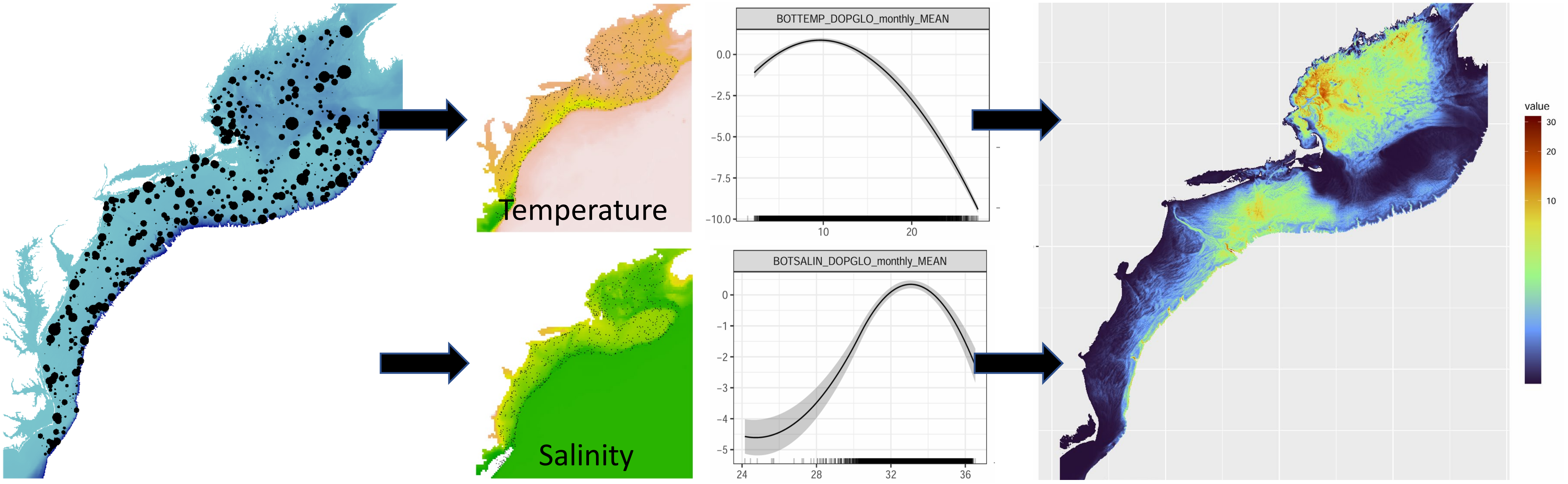
Dr. Chris Haak, Monmouth University

NEFMC Climate and Ecosystem Steering Committee Meeting

March 27, 2026

Species Distribution Models (SDMs)

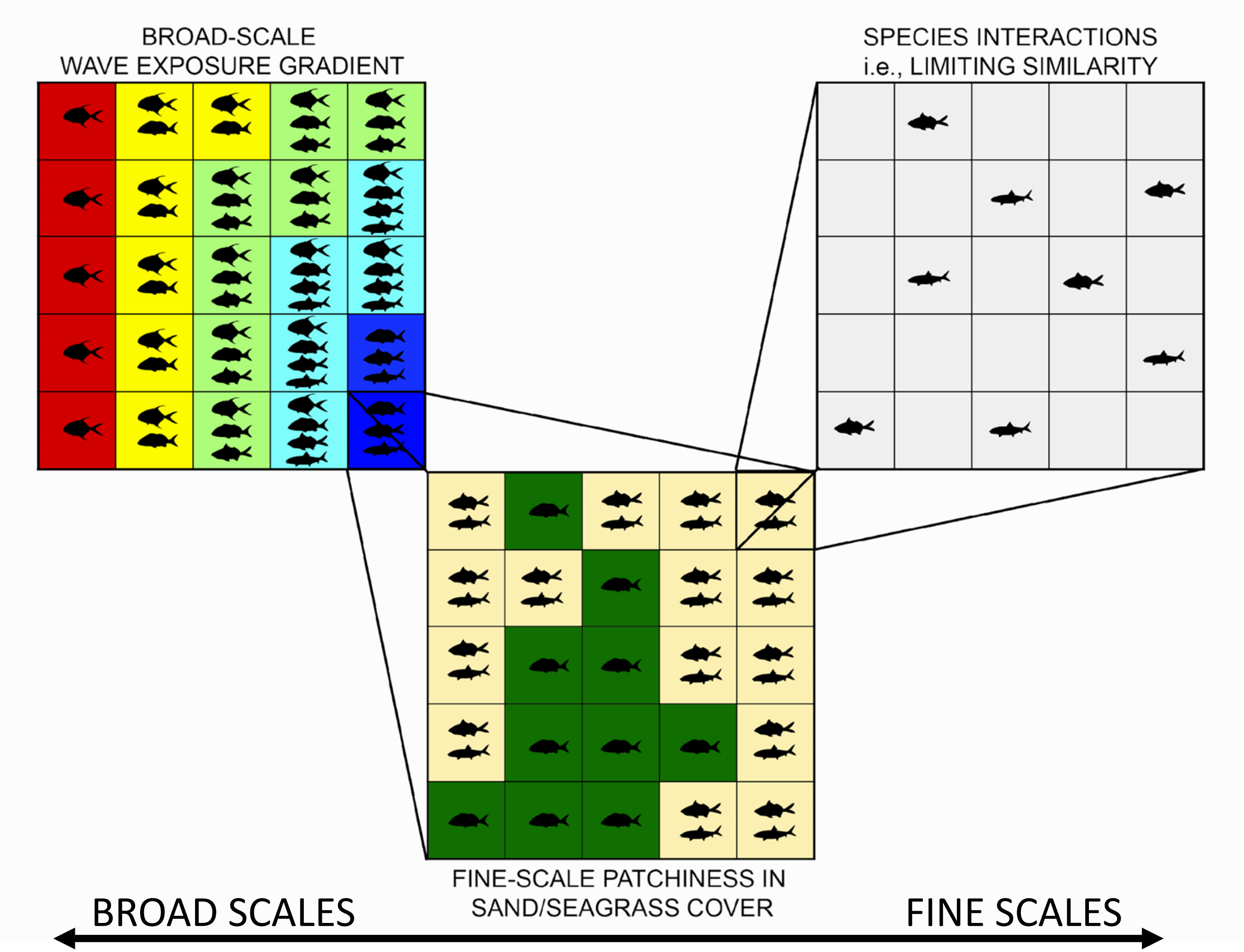
- SDMs estimate the habitat "niche" of organisms by relating observed densities to measured **environmental predictor variables**



Observed Density (in trawls) **Environmental variables** **Estimated Niche** **Predicted Density**

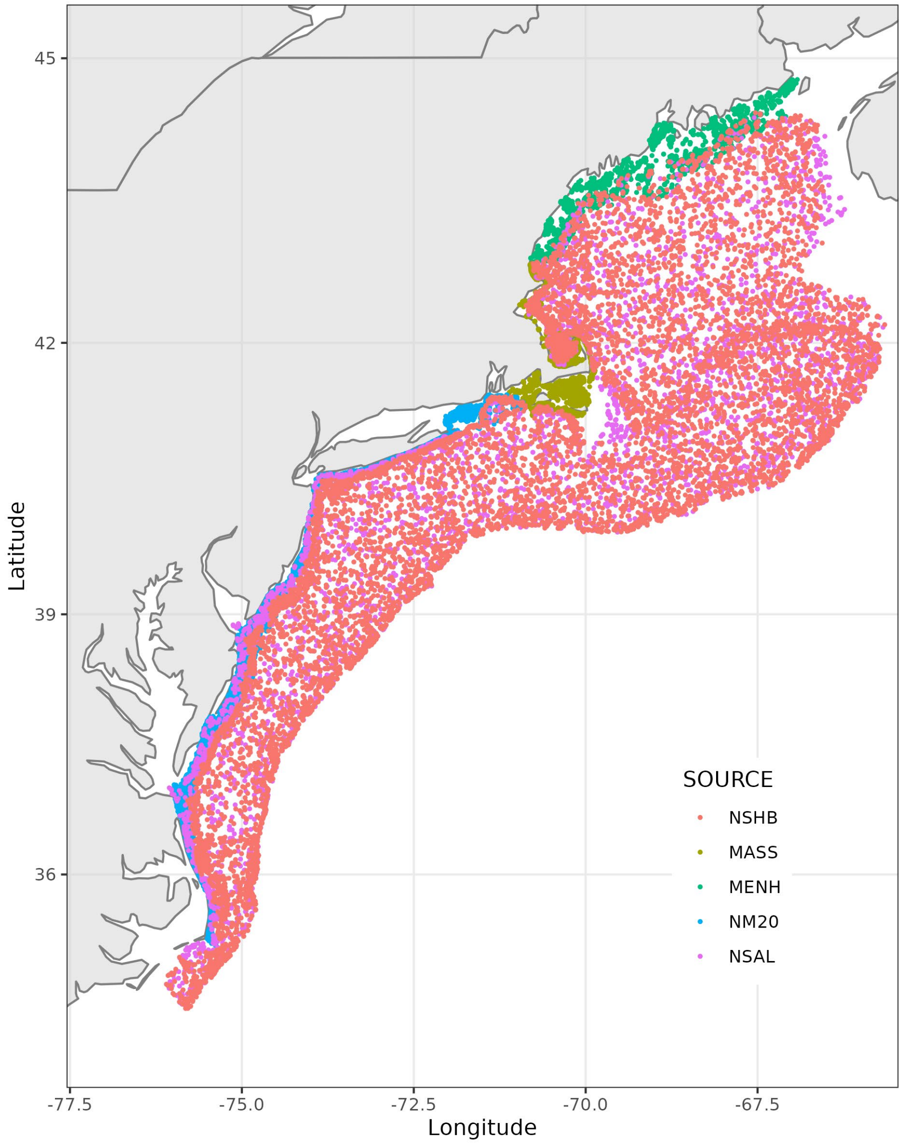
Species Distribution Models: Joint SDMs

- **Species interactions** also shape distributions
  - Competition, predation, facilitation, etc.
  - **Induce (+) or (-) correlations** in abundance
- **JSDMs model these correlations**
  - Increase precision of estimated species-environment relationships
  - Ecological insights?



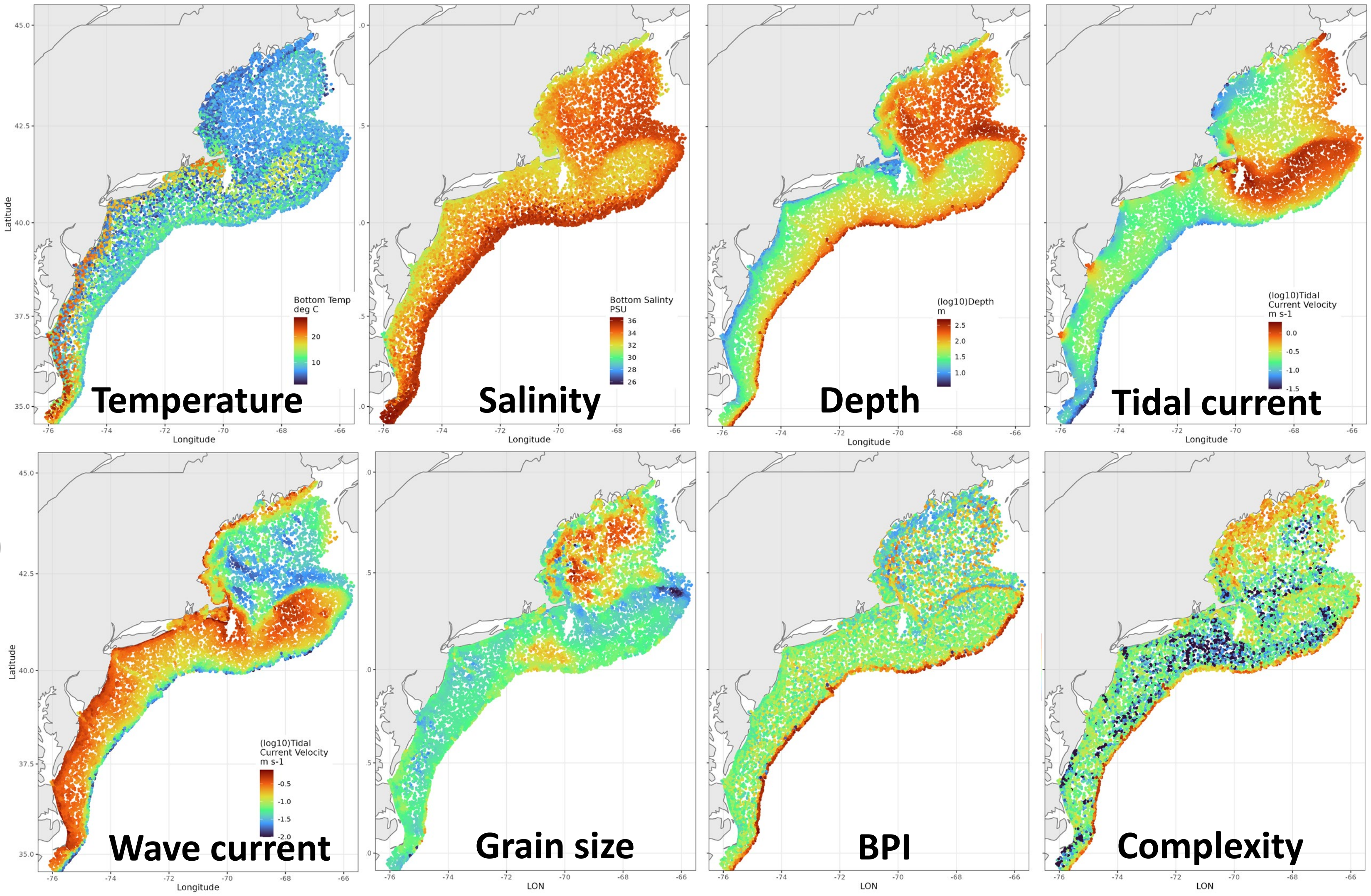
Modeling Framework: Response Data

- Bottom trawl surveys (Spring & Fall)
  - NEFSC (Bigelow & Albatross)
  - NEAMAP
  - Mass DMF
  - Maine/New Hampshire DMR
- Years 2002 – 2022 (omitting 2020)
  - ≈ 24,000 observations
- Lifestage-specific counts
  - Adult or juvenile/subadult stages
    - Based on length at maturity (L_{50})
    - Recruit/pre-recruit (e.g., *Illex squid*)



Modeling Framework: Environmental Data

- Temp & Salinity
  - Time-varying (30-day mean)
- Hydrodynamic (waves, tides & currents)
  - Static or 30-day climatology
- Benthic habitat (Sediment, topography)
  - Static
- Evaluated at Bottom (demersal spp) or surface (pelagic spp)



Modeling framework: Community Basis Function Model

- Joint SDM
- GAMs model fixed effects/ environmental responses (via R package mgcv)
- Hierarchical random effects capture species spatial & temporal covariance via community-level (i.e., species-common) basis functions (via R package TMB)

Received: 10 June 2022 | Accepted: 25 June 2023

DOI: 10.1111/2041-210X.14184

Methods in Ecology and Evolution
BRITISH ECOLOGICAL SOCIETY

RESEARCH ARTICLE

Spatiotemporal joint species distribution modelling: A basis function approach

Francis K. C. Hui[1] | David I. Warton[2] | Scott D. Foster[3] | Christopher R. Haak[4]

[1]Research School of Finance, Actuarial Studies and Statistics, The Australian National University, Canberra, Australian Capital Territory, Australia

[2]School of Mathematics and Statistics, and Evolution & Ecology Research Centre, The University of New South Wales, Sydney, New South Wales, Australia

[3]Data61, Commonwealth Scientific and Industrial Research Organisation, Hobart, Tasmania, Australia

[4]Urban Coast Institute, Monmouth University, West Long Branch, New Jersey, USA

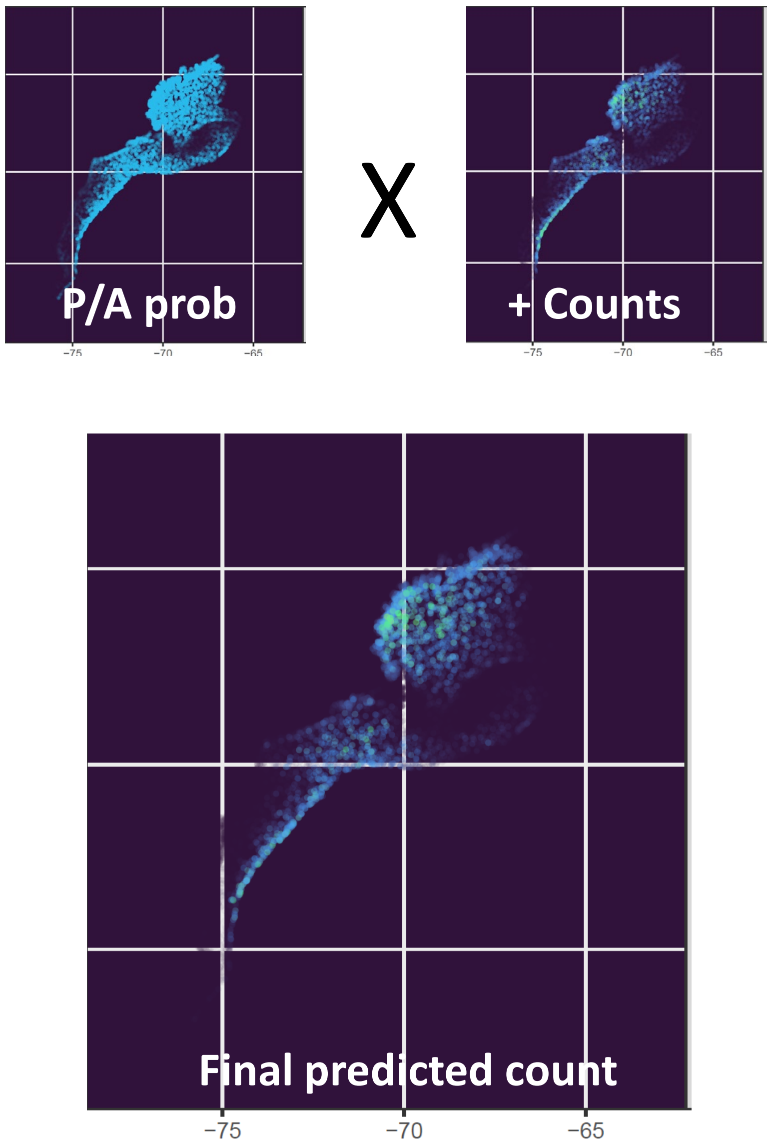
Correspondence
Francis K. C. Hui
Email: francis.hui@anu.edu.au

Abstract

1. We introduce community-level basis function models (CBFMs) as an approach for spatiotemporal joint distribution modelling. CBFMs can be viewed as related to spatiotemporal latent variable models, where the latent variables are replaced by a set of pre-specified spatiotemporal basis functions which are common across species.
2. In a CBFM, the coefficients that link the basis functions to each species are treated as random slopes. As such, the CBFM can be formulated to have a similar structure to a generalised additive model. This allows us to adapt existing techniques to fit CBFMs efficiently.
3. CBFMs can be used for a variety of reasons, such as inferring patterns of habitat use in space and time, understanding how residual covariation between species varies spatially and/or temporally, and spatiotemporal predictions of species- and

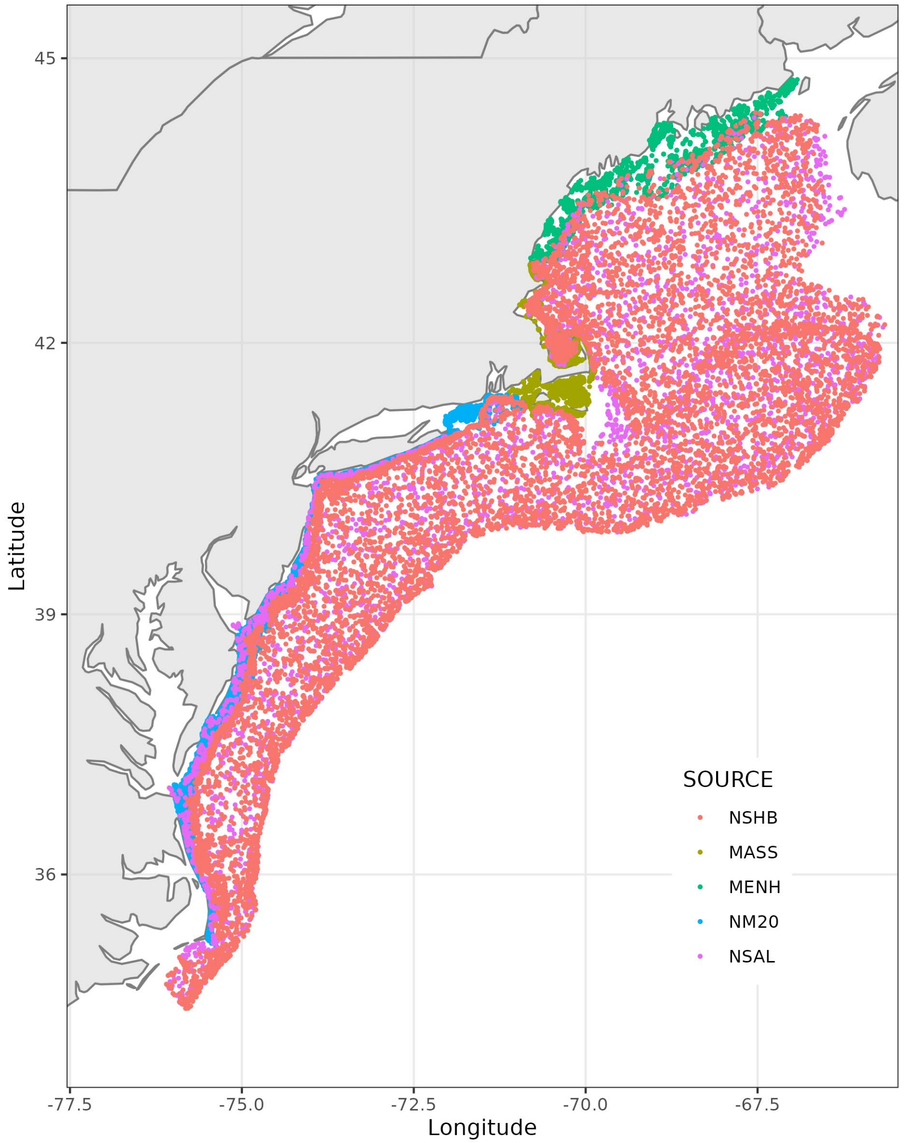
Modeling Framework: Basic Structure

- Hurdle-Negative Binomial
  - Presence/Absence (binomial)
  - Zero-truncated negative binomial
- Environmental Predictors (Fixed)
  - Estimated jointly across seasons: via thin plate regression splines (TPRS) smooths with shrinkage
- Spatial Effects (Random)
  - Season-specific 2D smooths: spatial patterns not explained by environs
- Temporal Effects (Random)
  - 1st order random walk: interannual variation not explained by environs



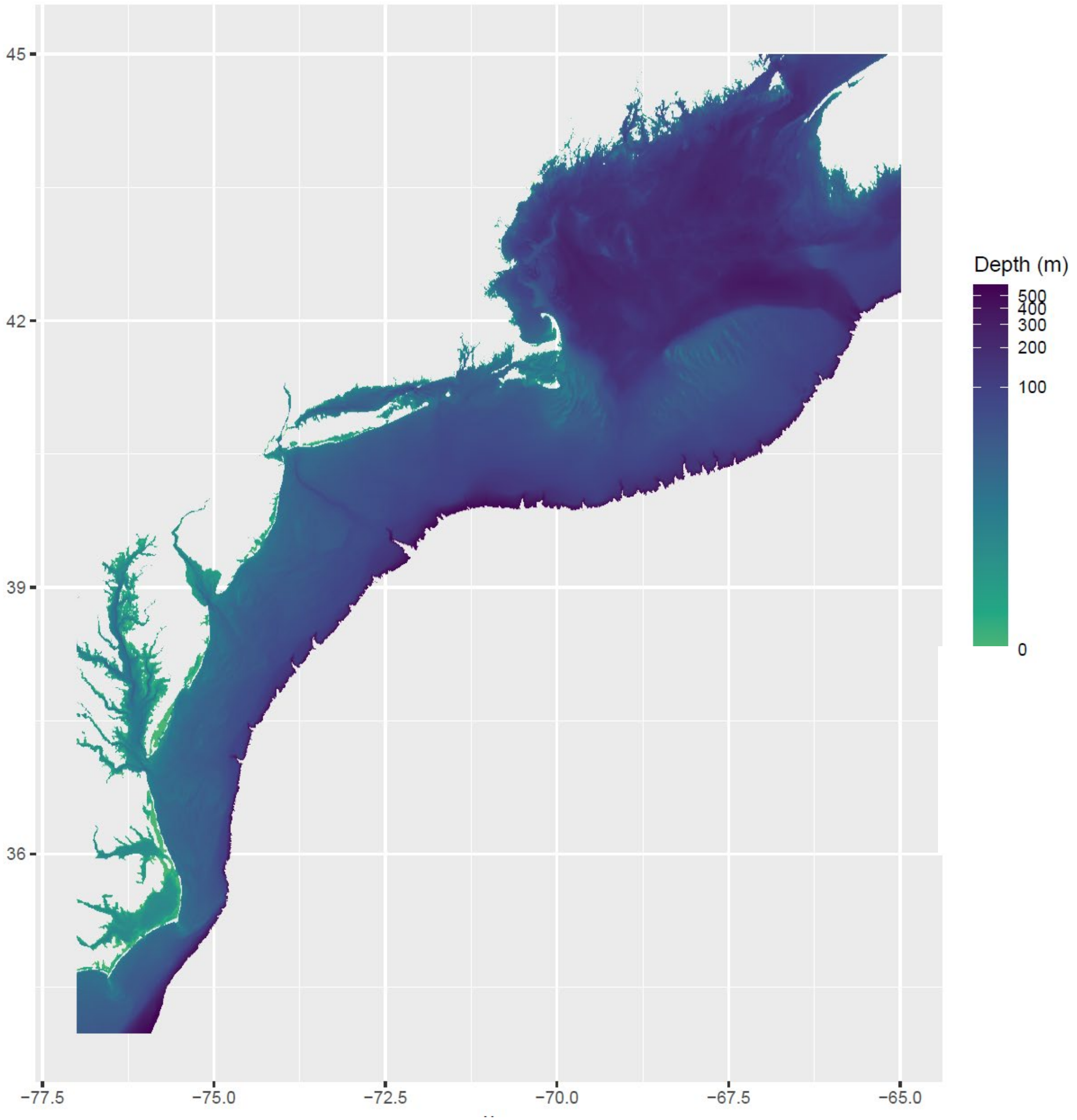
Modeling Framework: Survey Integration

- Survey Effects (Random)
  - 10 levels (5 surveys x 2 seasons)
  - ‘Footprints’ confounded with environment, space & time
    - Inflated coefficients and SEs
  - De-correlate survey from continuous predictors (orthogonal residuals via QR decomposition)
    - Remove environmental/spatial/temporal ‘signal’ from the survey effects;
  - ‘Zeroed out’ for prediction (marginal), or a selected survey (conditional)
- Swept Area offset (tow-level)
  - Survey counts share a common ‘per unit area’



Modeling Framework: Prediction Grid

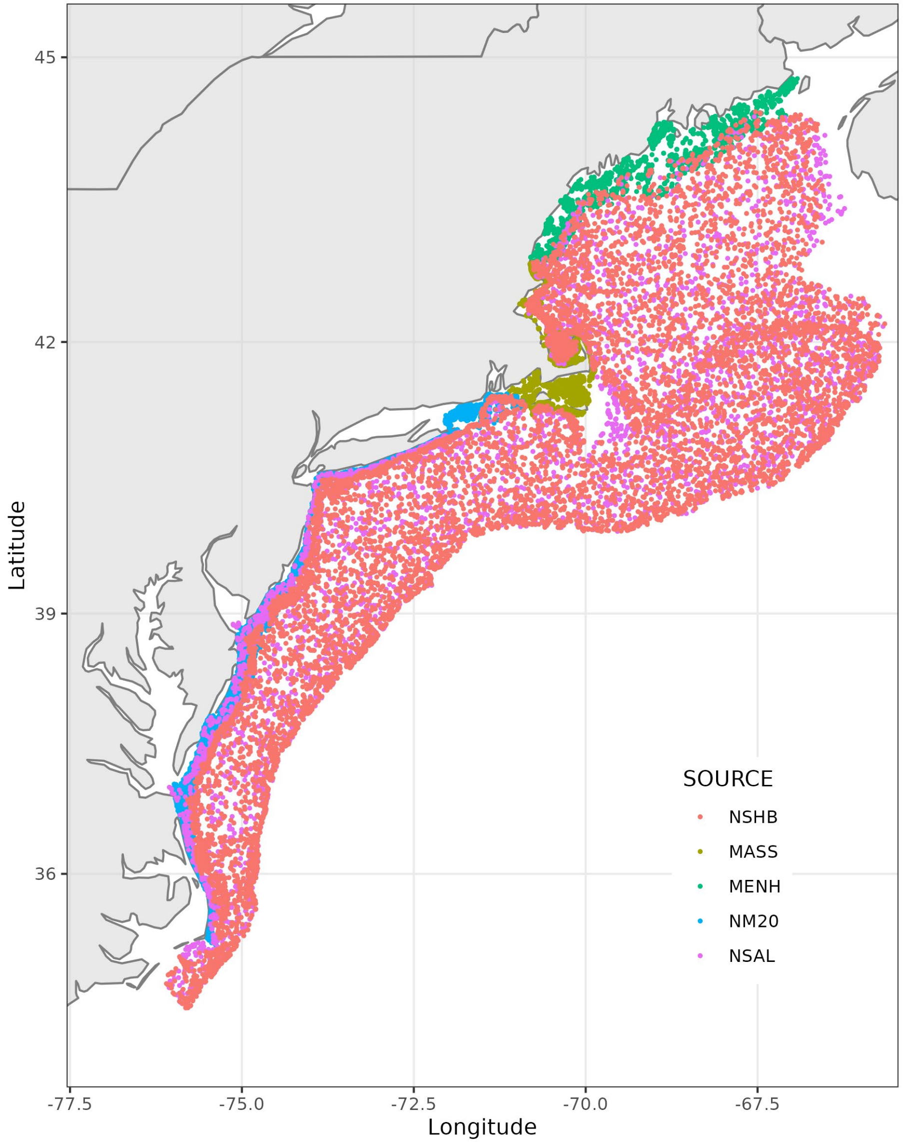
- ≈ Bounds of NEFSC trawl surveys
  - 34° to 45° N , 77° to 65° W
  - 5 - 500 m depth
- 1 km (≈ 0.01°) grid resolution
  - ≈ 300,000 'active' cells
- Predicting to monthly means
  - Mar-May (Spring)
  - Sept-Nov (Fall)
- 20 Years
  - 2002 – 2022 (omitting 2020)



Modeling Framework: Uncertainty

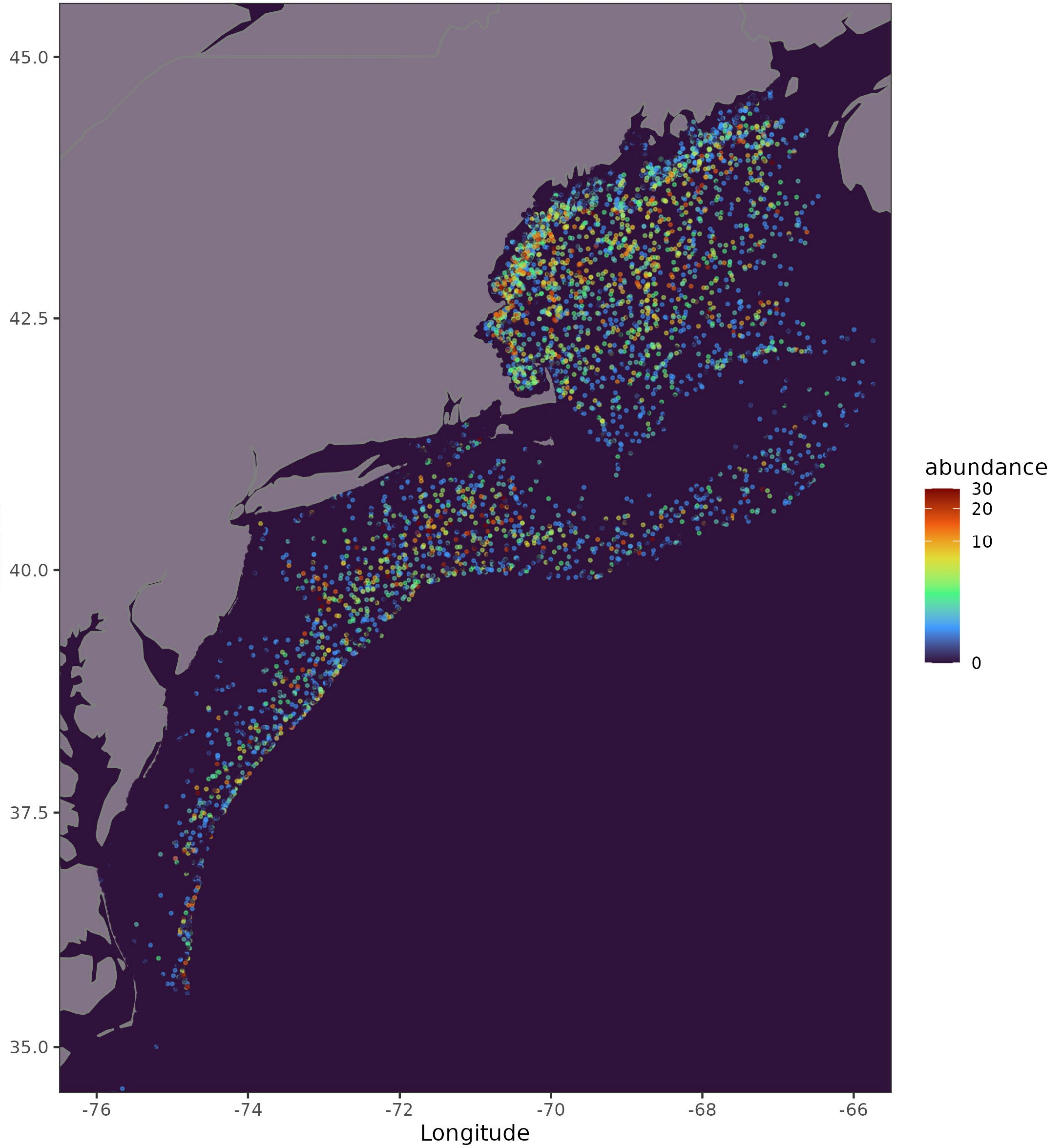
- **10-fold Cross Validation** to assess sensitivity to the training data
  - Re-fit 10 times, each time omitting a distinct 10% of the training data
  - Predict from each of 10 fits to assess across-fold variability
    - Coefficient of variation (CV) of predicted count (across 10 folds).
- **Out-of-sample prediction** for all data points to assess performance on ‘new’ data using several metrics
  - RMSE (Residual Mean Squared Error)
  - R (Spearman rank correlation)
  - AUC (Area Under the ROC Curve)
  - Log Score, CRPS (Continuous Ranked Probability Score)?

Worked example: Goosefish Juveniles

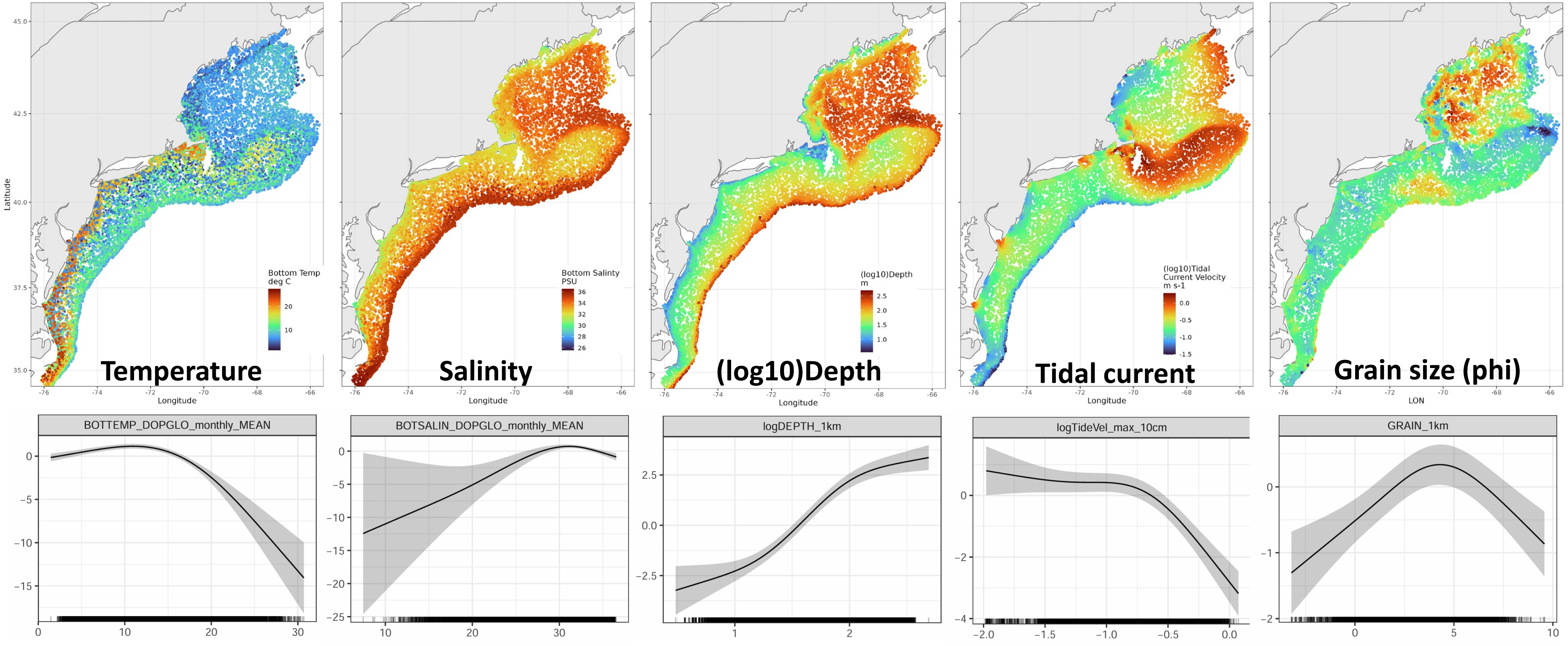


Lophius americanus

L_{50} = 37 cm



Worked example: Goosefish Juveniles



Lower to moderate bottom temps

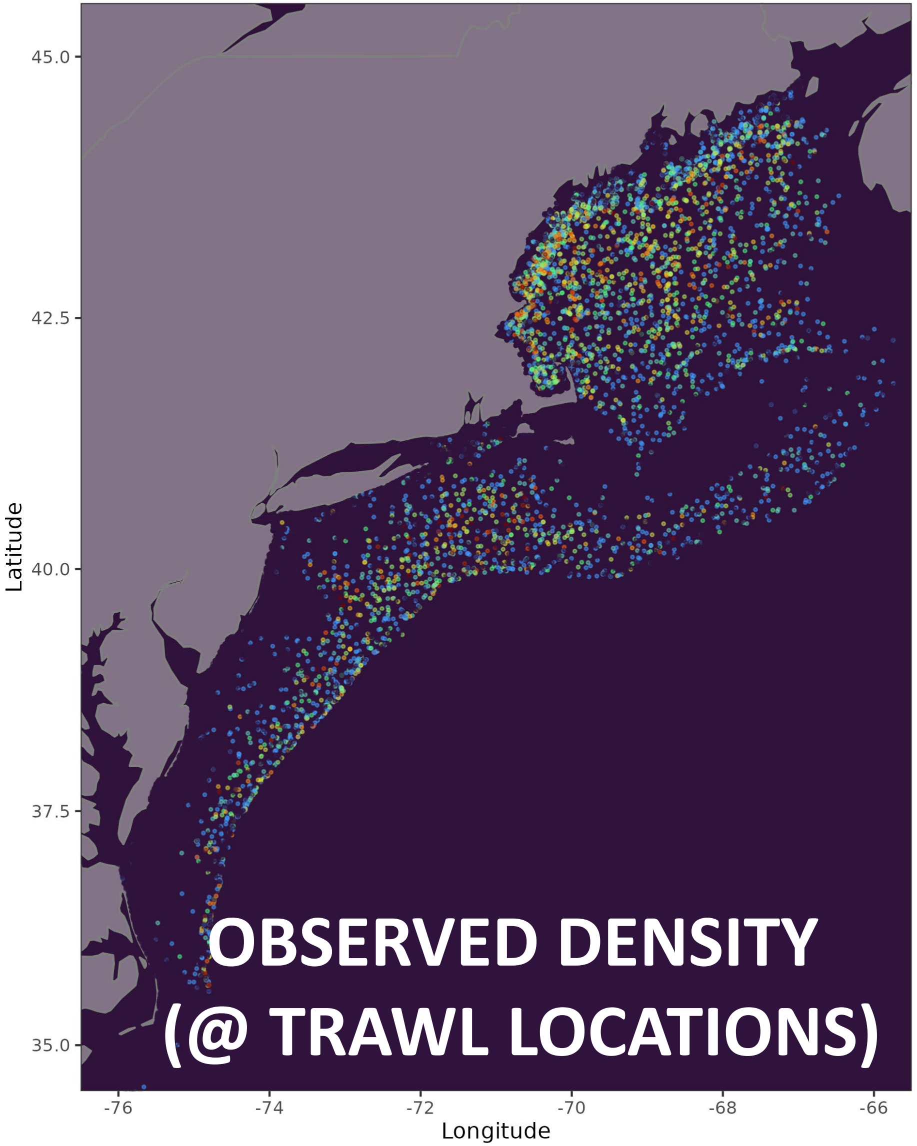
Higher salinities

Greater depths

Low to moderate tidal currents

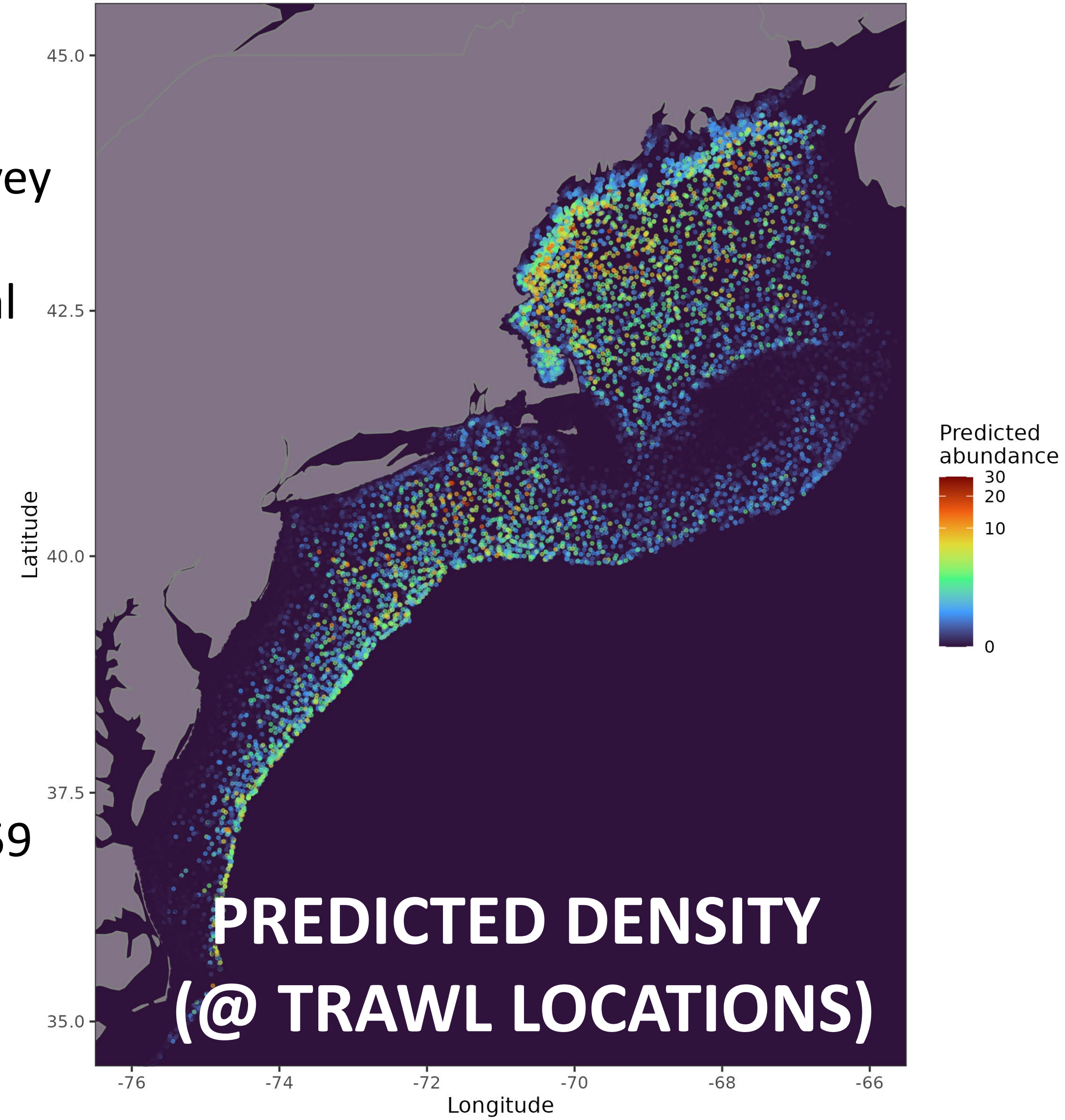
Intermediate grain sizes (fine sand-mud)

Worked example: Goosefish Juveniles



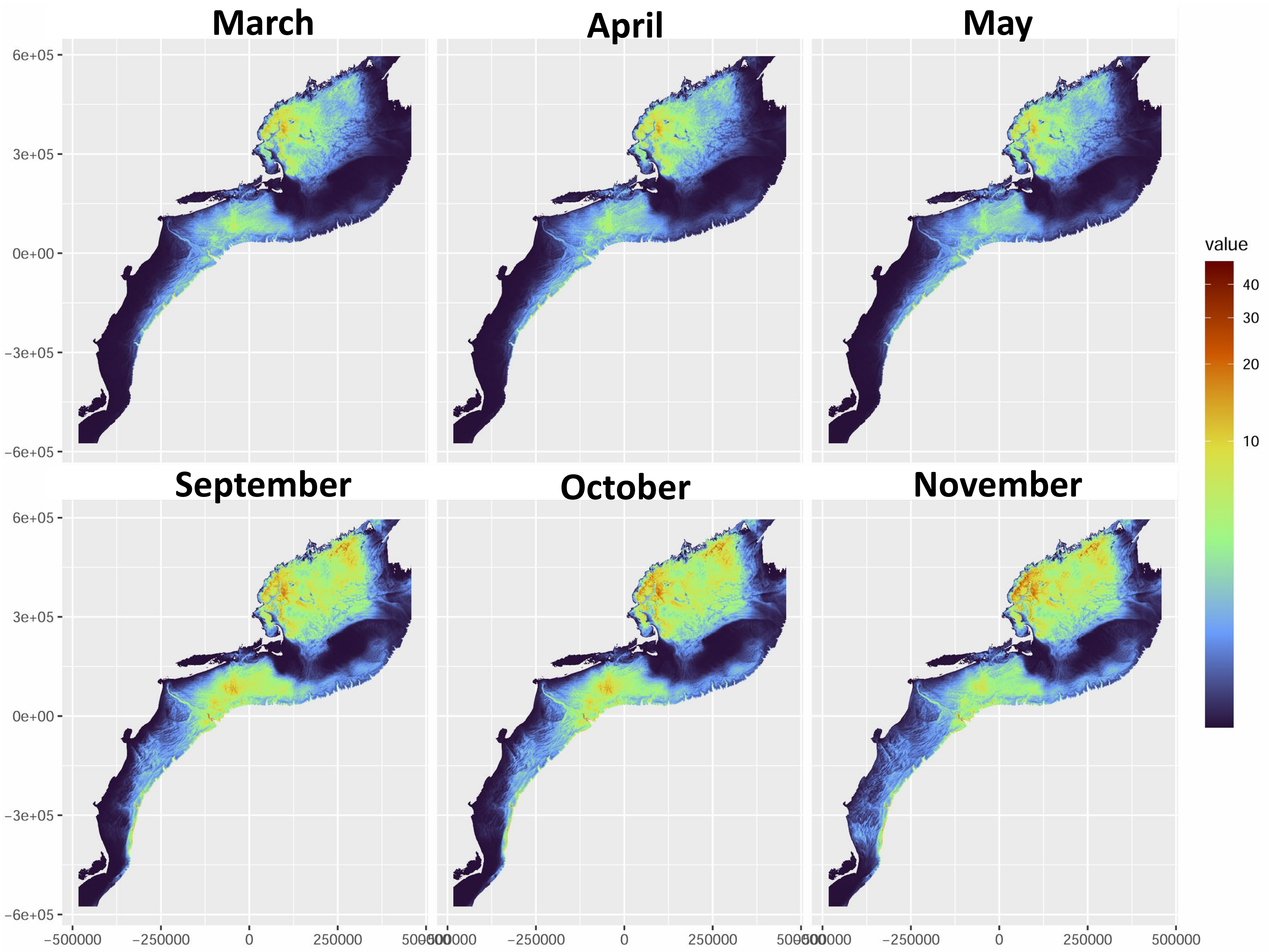
Environmental responses + Survey effects + Spatial fields & Temporal effects

Out-of-sample Performance:
RMSE = 2.92
Spearman R = 0.59
AUC = 0.91



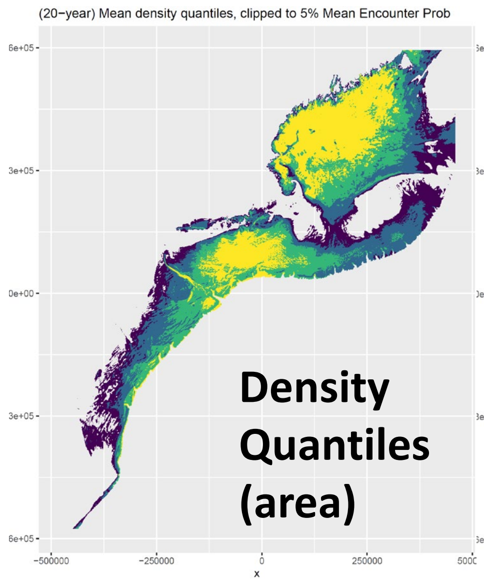
Worked example: Goosefish Juveniles

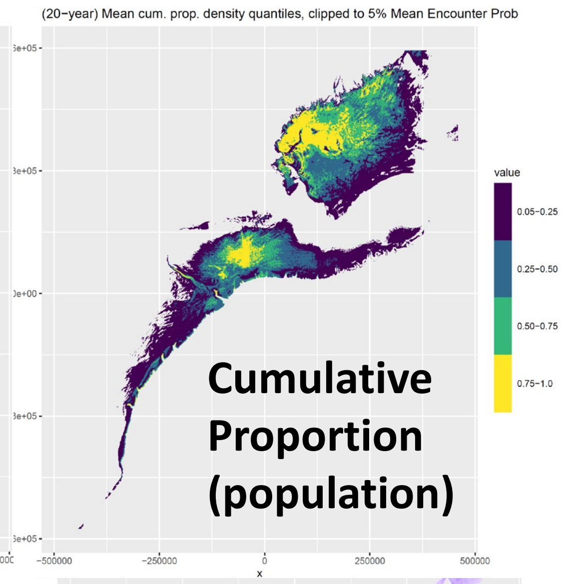
- **Continuous gridded predictions for 20 Yrs**
  - Based on monthly mean conditions
    - Mar-May (Spring)
    - Sept-Nov (Fall))

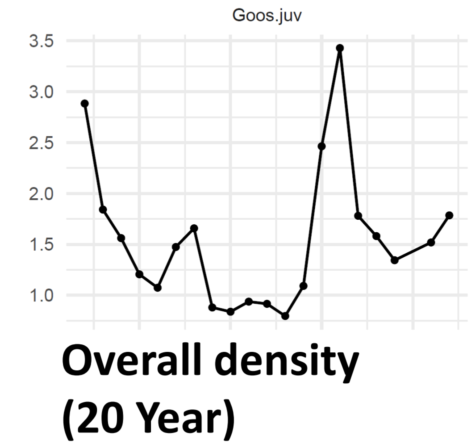


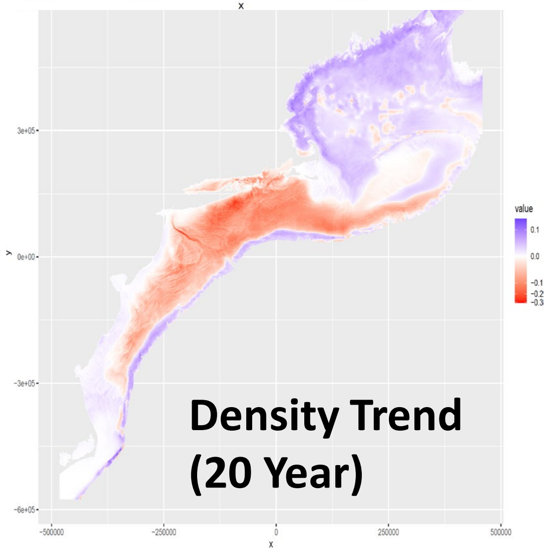
Prediction inference: Species-level

- **More ‘Holistic’ –** pooling environmental and latent processes
- Species ‘footprints’, ‘hotspots’, etc.
  - Density quantiles vs. cumulative proportion
- Distribution shifts
  - Spatially-explicit (cellwise) trends - is density shifting between cells?
  - Center-of-mass - is the center of the overall distribution shifting?
- Changes in total area occupied (i.e., range expansion/contraction)
- Estimated population trends



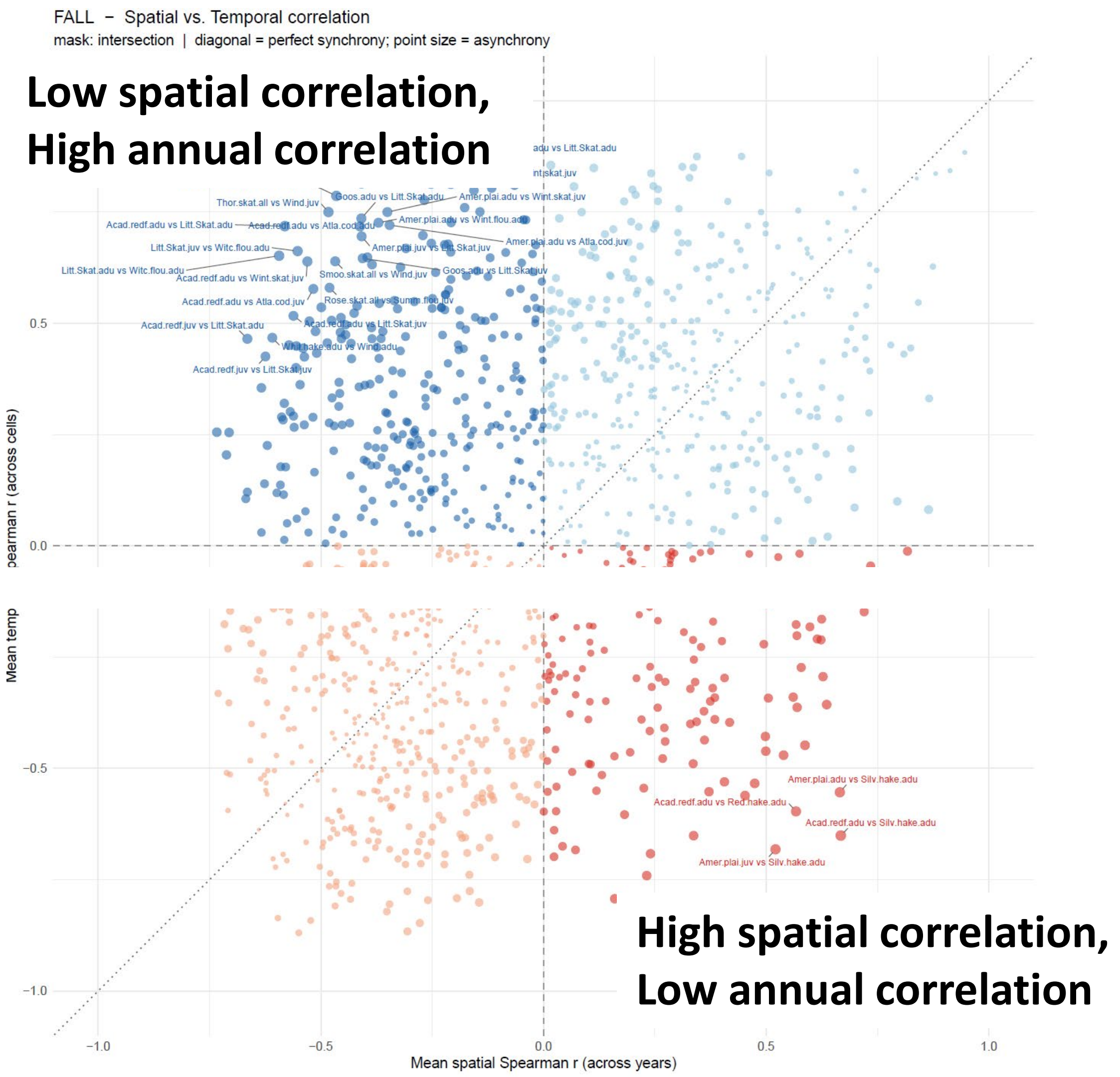






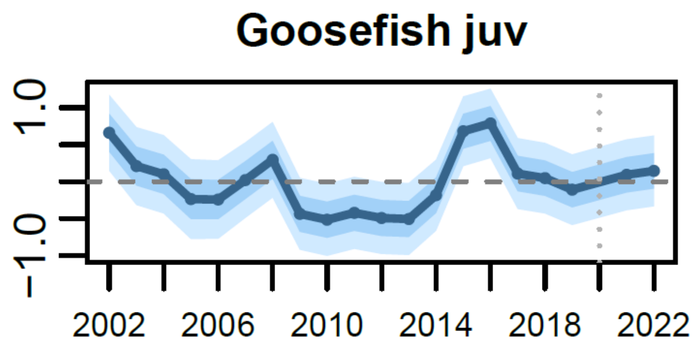
Prediction inference: Community-level

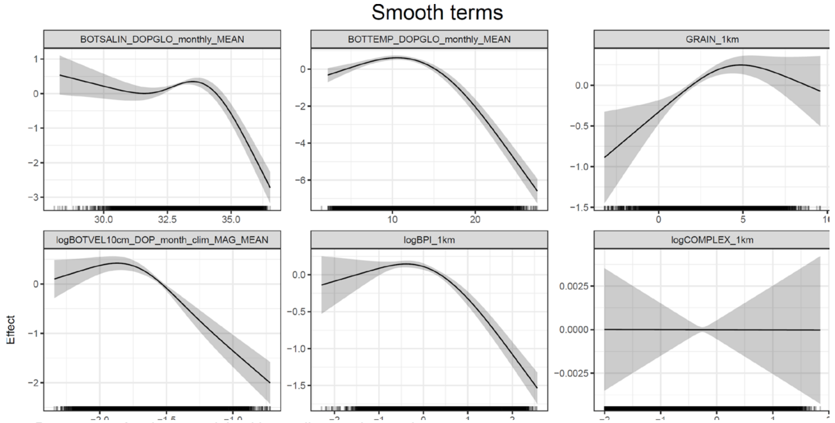
- Pairwise correlations in density (across space; e.g., cooccurrence)
- Spatially-explicit (cellwise) correlations in density over time
- Overall correlations in population size over time
- **Portfolio metrics**
  - Synchrony/asynchrony
- Pairwise overlaps in area occupied
  - Overall, or 'core'
  - Trends in overlap?

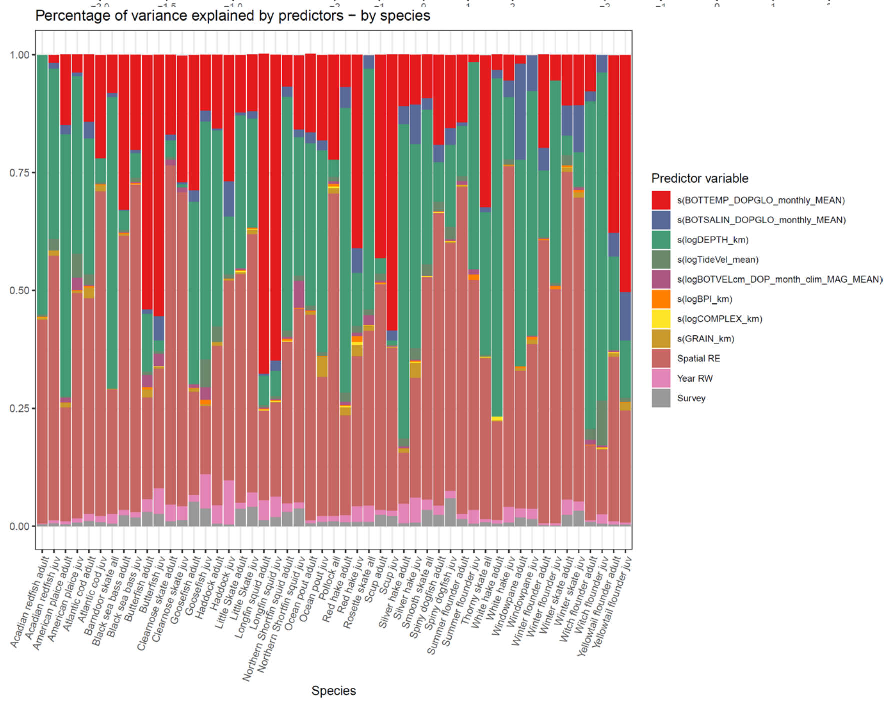


Model inference: Species-level

- **More 'Granular'** – partitioning out environmental and latent processes
- Species-environment relationships
- Relative importance of environmental drivers (via variance decomposition)
- Spatial & Temporal trends not attributable to measured environmental variables

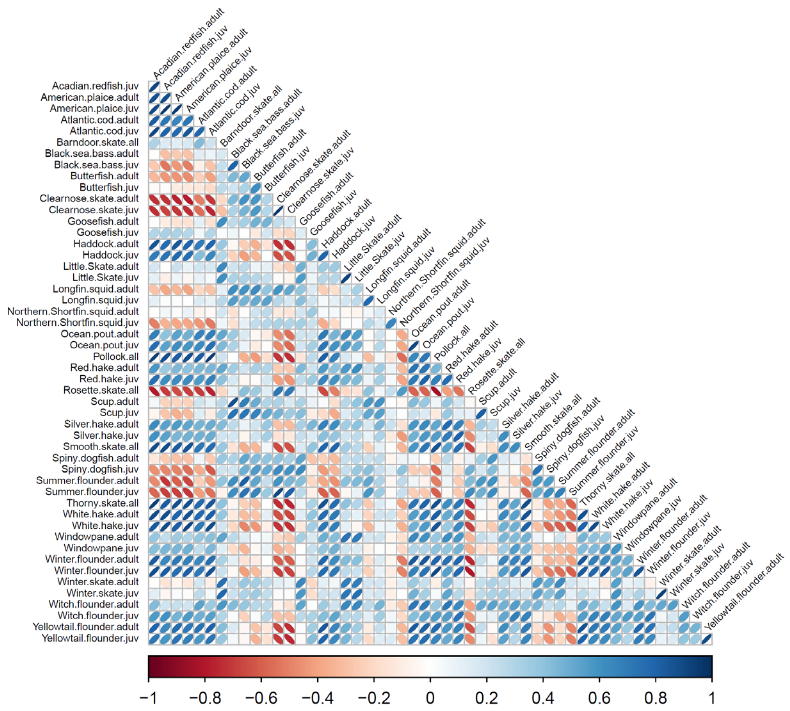


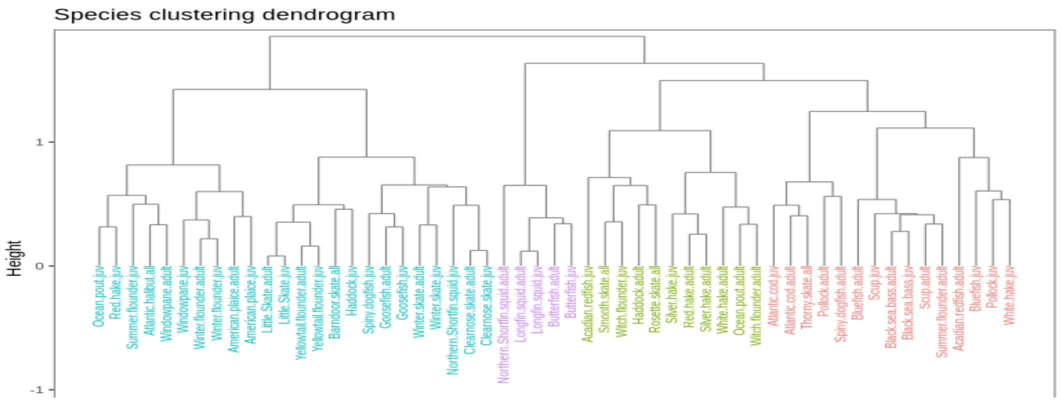




Model inference: Community-level

- Correlations in response to environment & spatiotemporal REs
  - Ecological interactions?
- Hierarchical clustering by:
  - Environmental response
  - Spatial & Temporal REs
  - Group taxa with similar responses to environmental or latent processes





Discussion Questions

- Which of these model outputs are ready for near-term management use? Which remain exploratory / under development?
- What projects can we apply these results to?
  - Essential Fish Habitat designation
  - Ecosystem Component Species evaluation
  - Portfolio Analysis
  - Others?
- What are the limitations or uncertainties the Steering Committee should keep in mind before encouraging broader use?

Acknowledgements

Recent funding for this work has been provided by the New England Fishery Management Council and the Mid-Atlantic Fishery Management Council.

Data were provided by NOAA Fisheries, Maine Dept. Marine Resources, New Hampshire Dept. Fish and Game, Massachusetts Div. Marine Fisheries, and others.

MID-ATLANTIC FISHERY MANAGEMENT COUNCIL

MONMOUTH UNIVERSITY

New England Fishery Management Council