A space-time interpolation tool for Chesapeake Bay dissolved oxygen: Parameterizing a 4-dimensional correlation structure

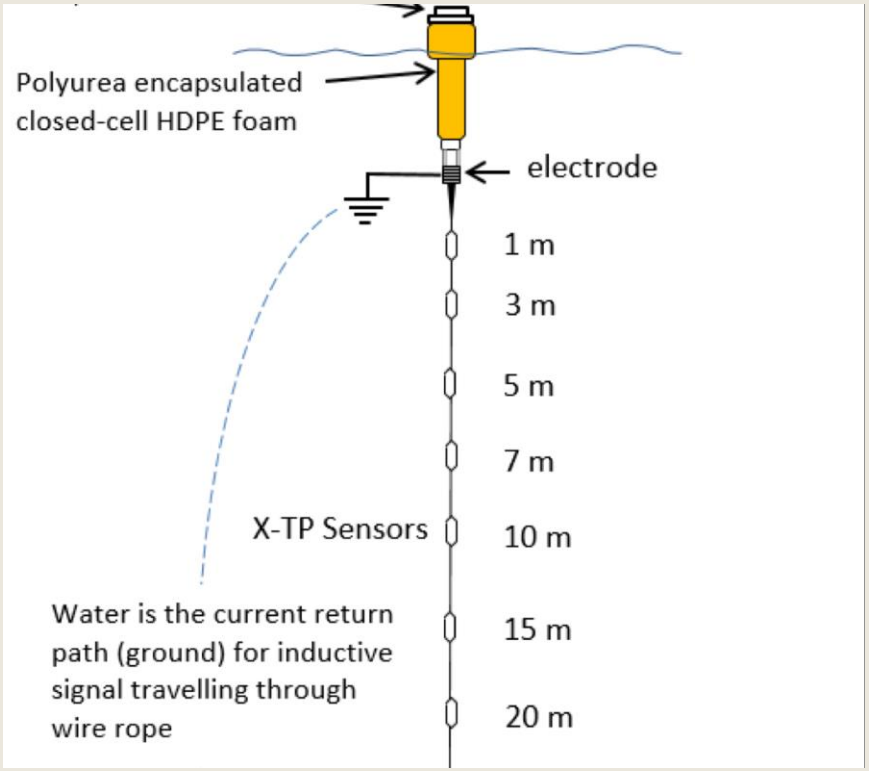
Jon Harcum[1], Elgin Perry[2], Rebecca Murphy[3], and Breck Sullivan[4]

[1]Tetra Tech, [2]Statistics Consultant, [3]UMCES at CBP, [4]USGS at the CBP

CERF
Nov. 12, 2025

Reasons for the space-time interpolator design

- Assessments need **horizontal + vertical + seasonal + within-day** structure together.
- **Ideal world:** vertical arrays every ~1 km, all depths, all days.
- **Reality:** mixed networks (fixed stations, cruises, arrays) ⟶ patchy space and irregular time.
- **Legacy 2-D/3-D** methods miss diurnal × seasonal interactions.
- We need **daily, depth-resolved fields with uncertainty** for assessment.



Our approach: General additive model (GAM) for the **mean mid-day** DO + empirically fitted **horizontal, vertical, and temporal** correlations (MVN errors) ⟶ a **4-D interpolator**.

Outline

- Data landscape
  - *Which data inform mean mid-day DO and correlation structure?*
- Mean mid-day space–time interpolation
- Vertical, temporal, and horizontal correlations
  - *Methodology*
  - *Daily cyclic terms*
- Constructing the 4-D prediction space
- Preliminary Results

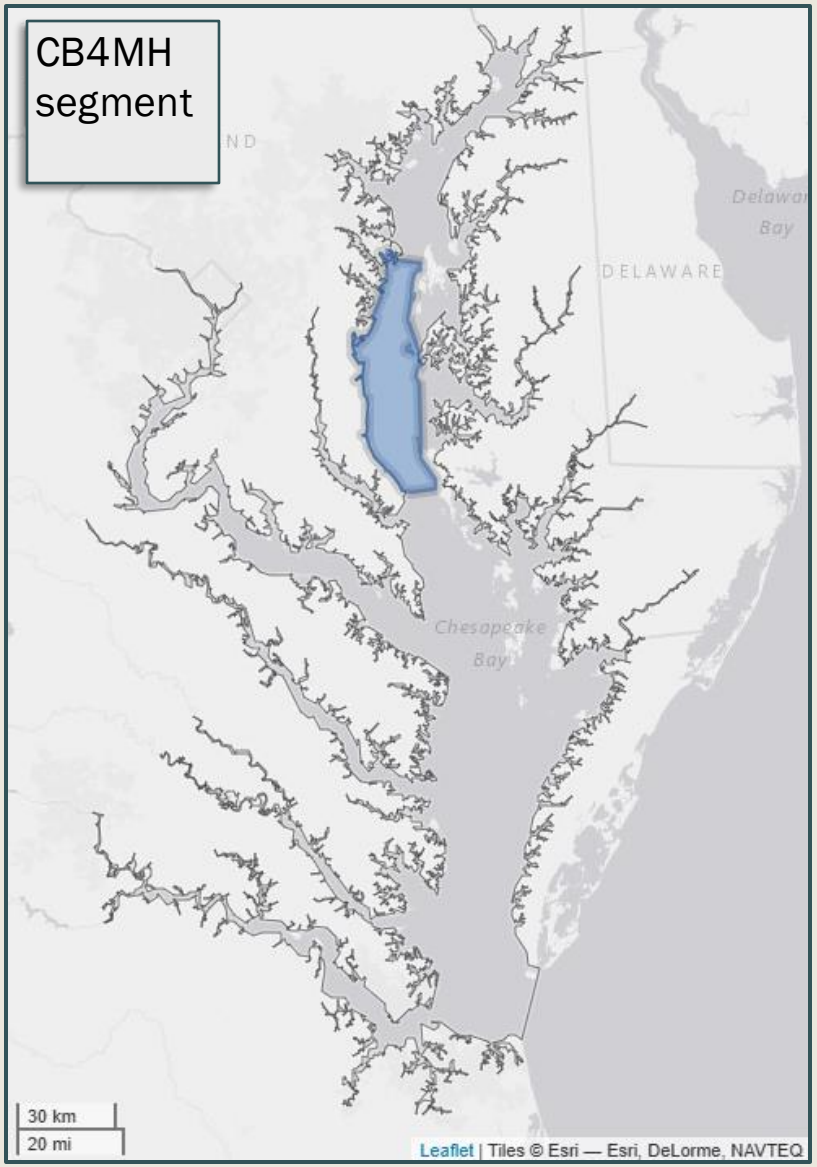
Data Landscape

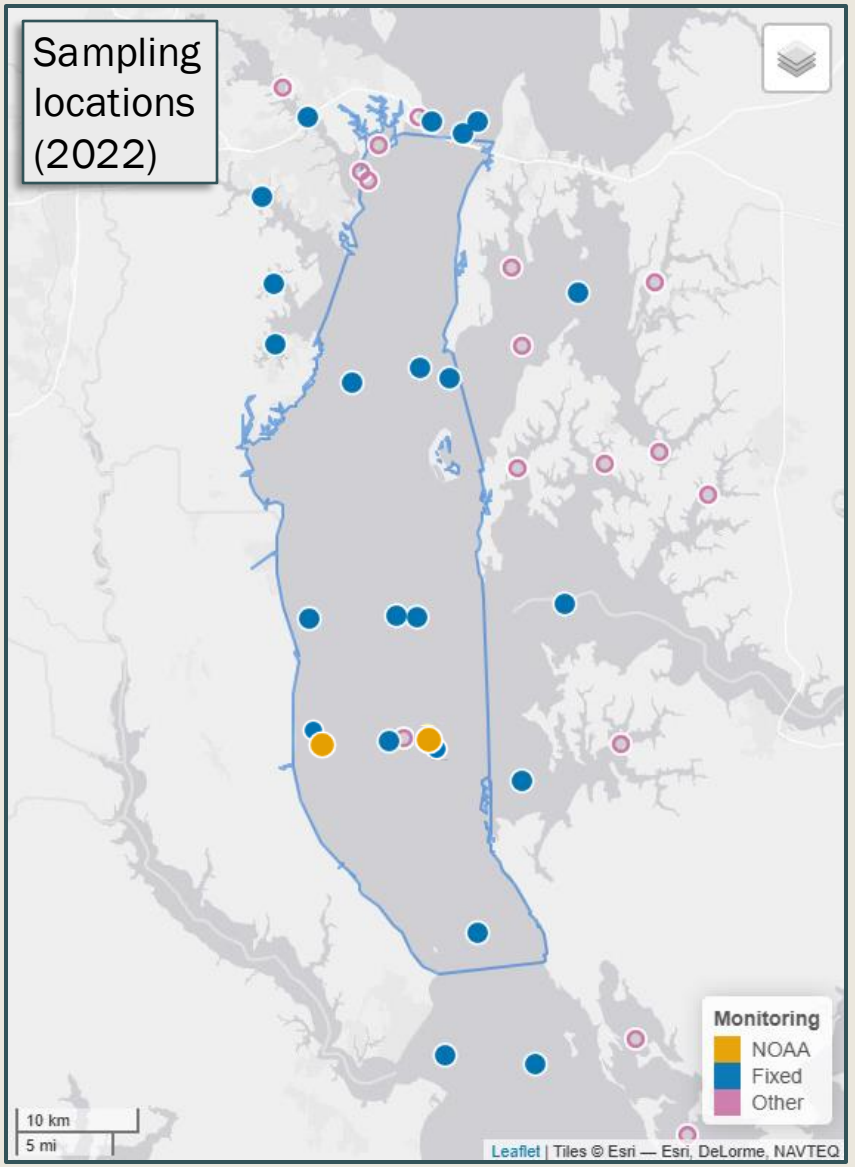
Informs mean mid-day space–time interpolation

Fixed-station network	Cruise-track monitoring (**DataFlow***)	Continuous monitoring (**ConMon***)	Vertical arrays (**NOAA**)	Additional State Agency Collected Data
✓Fixed location; broad spatial coverage ✓Long-term consistency ✓Profiles (every 1-2 meter to bottom) ✓150+ sites ✓1-2x/month	✓Surface mapping continuous data (~0.5 m) ≈ 8-10 sites/yr* ≈ 4-7 cruises per year (Apr-Oct)	✓Fixed location ✓15-minute sampling ✓Fixed depth near surface or bottom ≈ 25-30 sites/yr*; typically, 6-9 months/yr; some year round	✓Fixed location; multi-depth ✓10-minute sampling ✓New since 2022 ≈ 2-3 sites/yr ≈ 5-9 month deployments	✓Expands monitoring breadth ✓1-2x/month **Citizen monitoring (Tier 3)** ✓Expands monitoring breadth ✓1-2x/month
Vertical correlation	Horizontal correlation	Daily correlation; Daily cycle;	Vertical correlation; Daily correlation; Daily cycle	

*Sites commonly rotated about every 3 years to broaden coverage

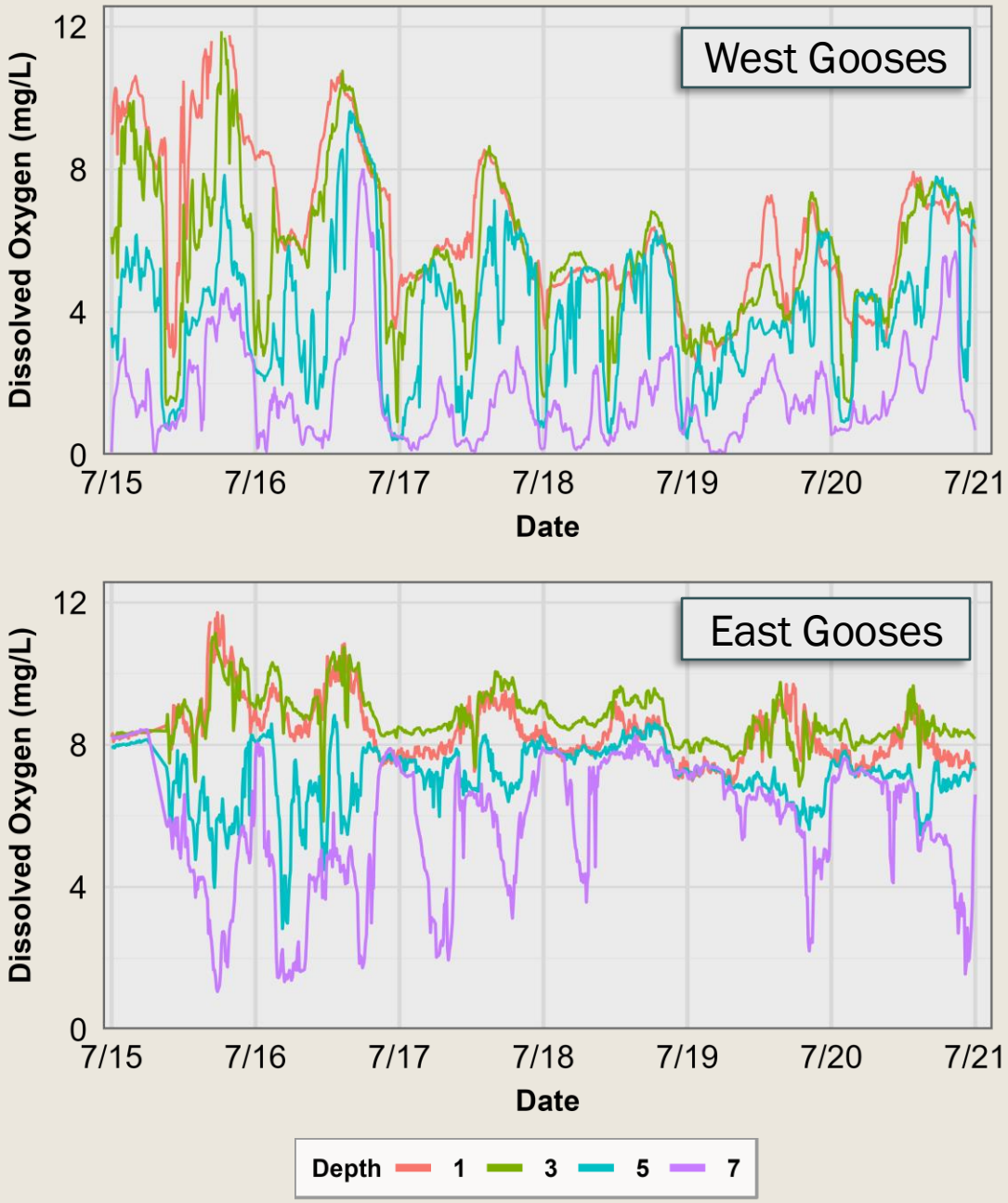
Middle Central Chesapeake Bay (CB4MH)





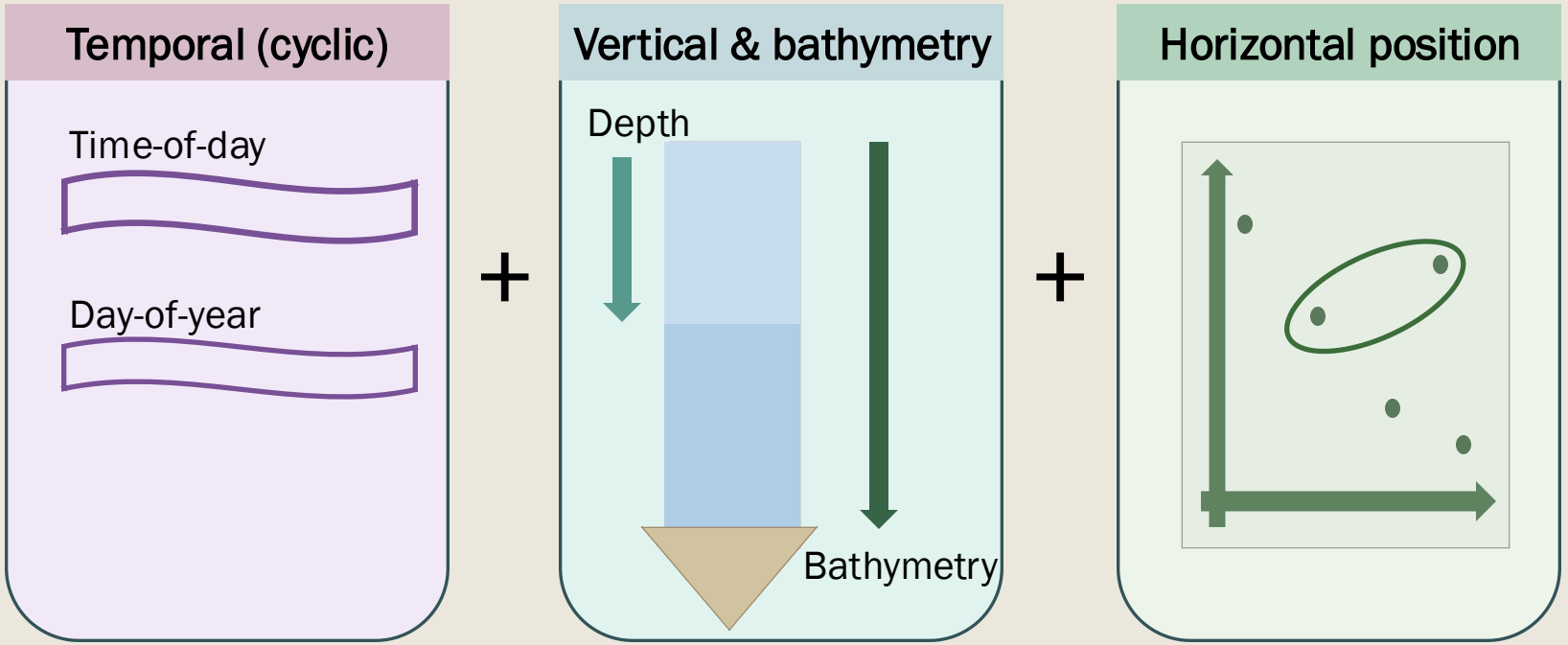
NOAA vertical arrays West & East Gooses: example period

- **Depths 1–7 m** (East reaches **19 m**); 10-min DO sampling.
- In-phase lines ⇒ higher correlation; out-of-phase or different amplitude ⇒ lower correlation.
- Large-scale correlation terms: **depth** and **daily lag**.
- Sites **~9 km** apart ⟶ weaker cross-site correlation.



Mean mid-day space–time interpolation

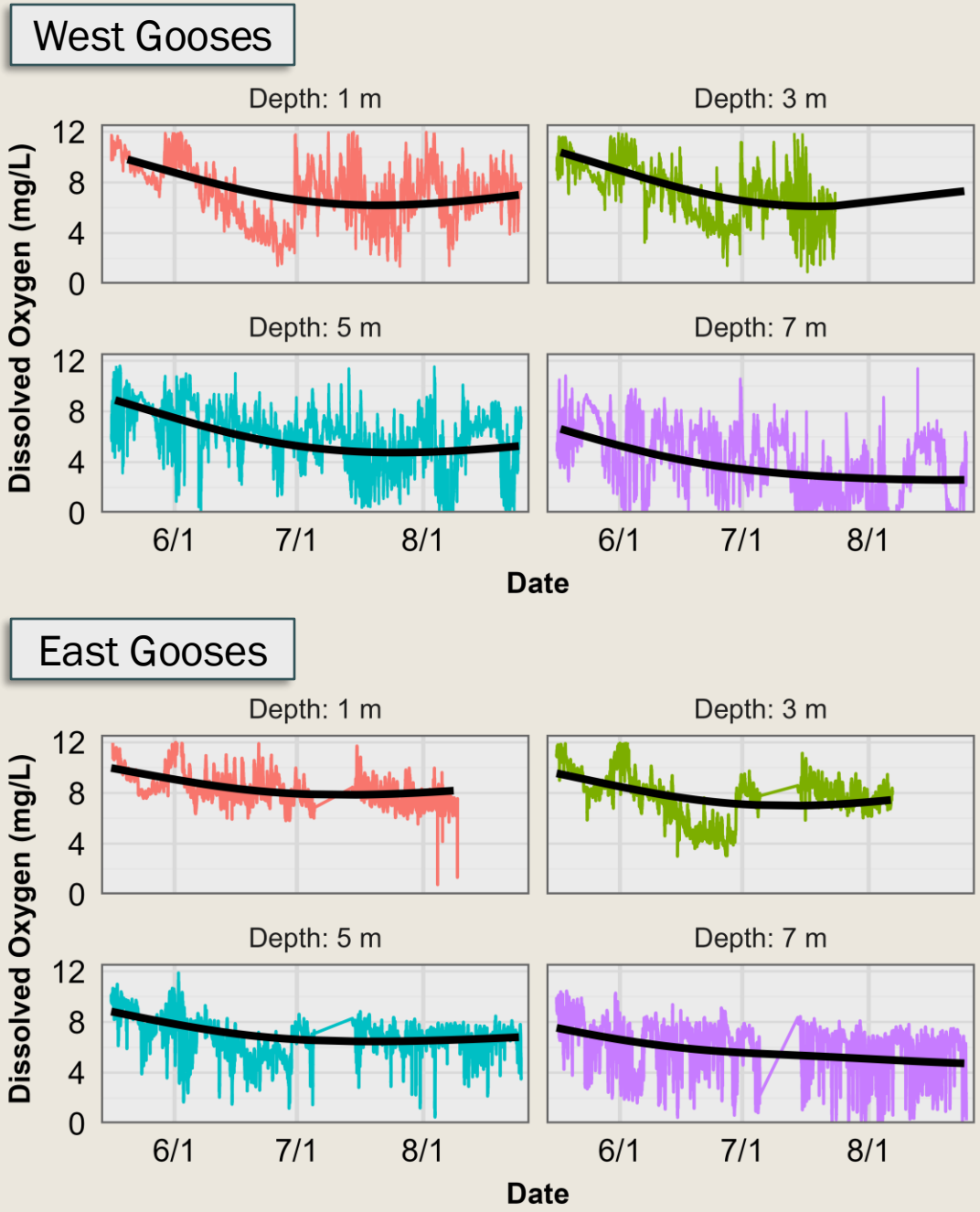
- **Goal:** Capture the central tendency of the data



- **Cyclic smooths (cc)** for time-of-day and day-of-year.
- **Tensor-product smooths (ti)** to model interactions among depth, bathymetry, and horizontal position.

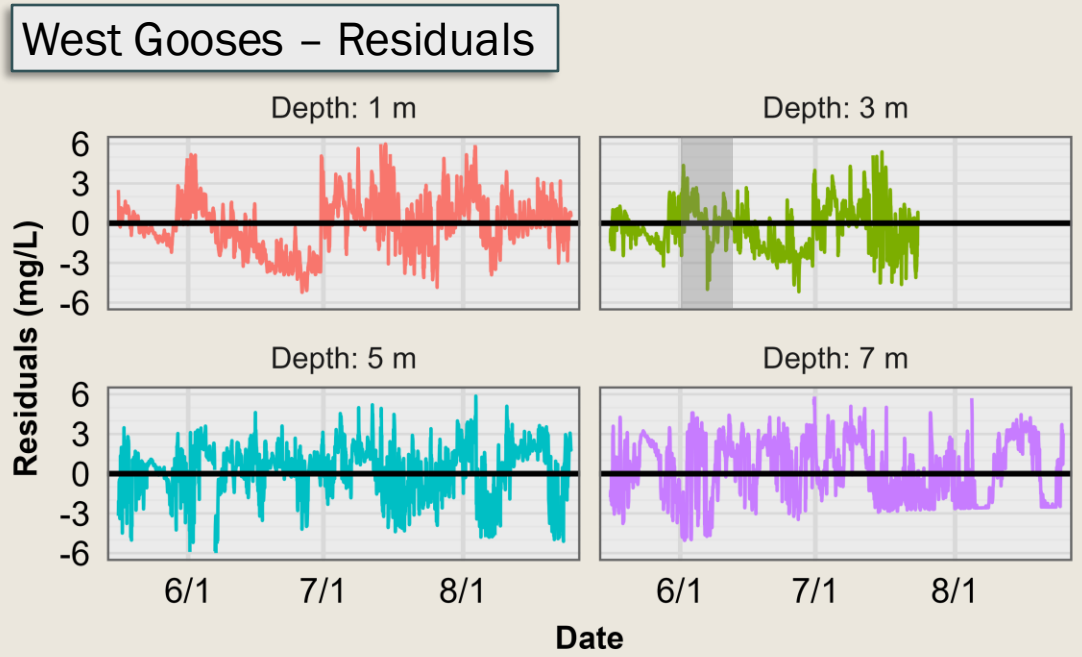
Mean mid-day space–time interpolation

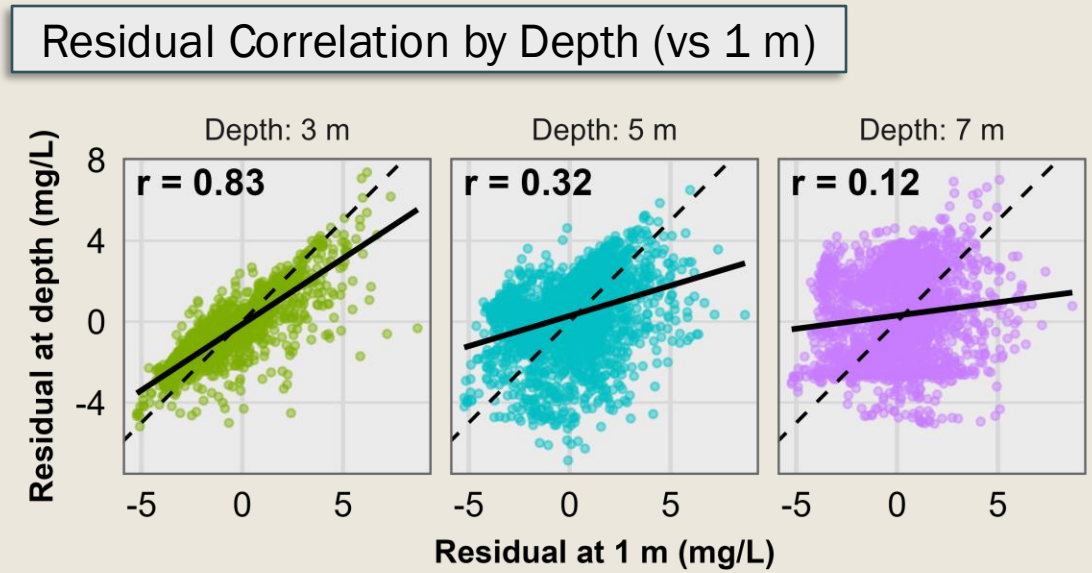
- NOAA vertical arrays West & East Gooses May 15– August 25, 2022.
- 10-min DO; panels show depths 1, 3, 5, and 7 m.
- Black line: mean mid-day central tendency interpolated.



Vertical Correlation (by depth)

- Compute residuals.
  - Removes seasonal signal.
- Correlate each depth vs. other depths at the same site.
  - Correlation measures **co-movement** (tight scatter), not slope.
  - 3, 5, and 7-m vs. 1 m shown here.
- Correlation drops as depth separation grows:
- Same approach for fixed stations with profile data.

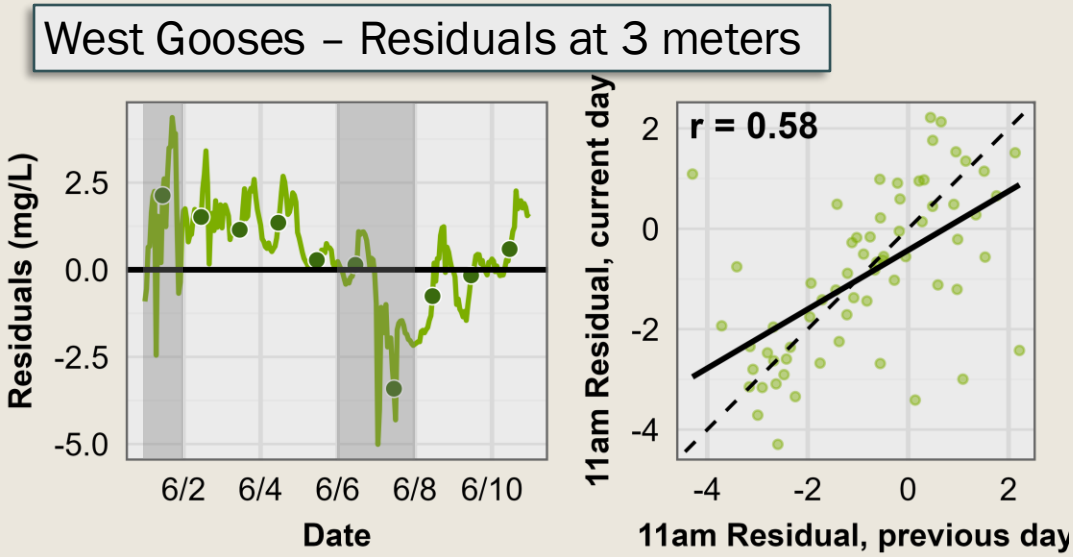




For clarity, plots show observed units (mg/L); analyses were performed with beta–logit–transformed DO, so results may differ slightly though patterns are unchanged.

Daily Lag-1 Correlation

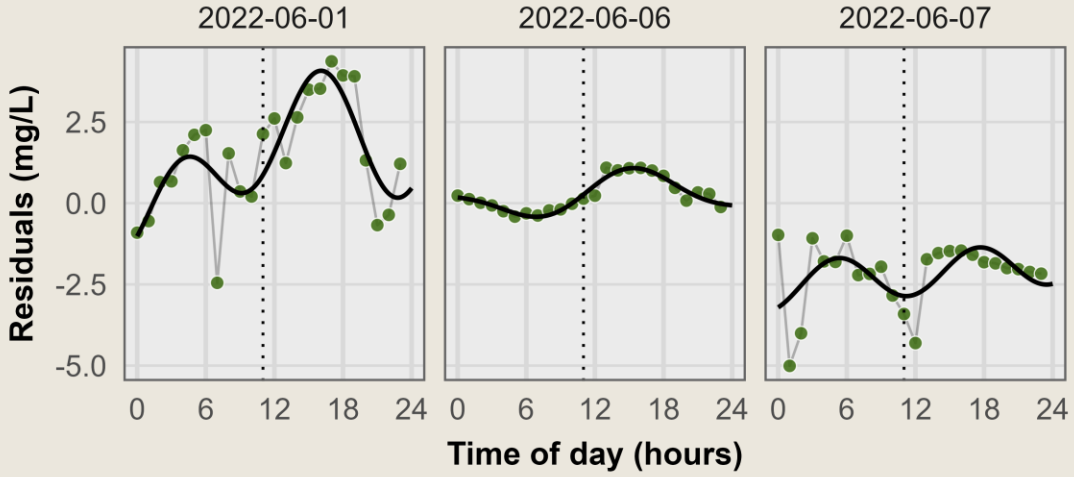
- Sub-sample time series data to daily time series by selecting 11am observation.
- Lag-1 autocorrelation computed.



Daily Cycle

- Sub-sample time series data to hourly time series.
- Fit 1st order Fourier series at 24 h and 12.42 h with linear trend.

$$y(t) = \beta_0 + \beta_1 t + \beta_2 \sin\left(\frac{2\pi t}{24}\right) + \beta_3 \cos\left(\frac{2\pi t}{24}\right) + \beta_4 \sin\left(\frac{2\pi t}{12.42}\right) + \beta_5 \cos\left(\frac{2\pi t}{12.42}\right) + \varepsilon(t)$$



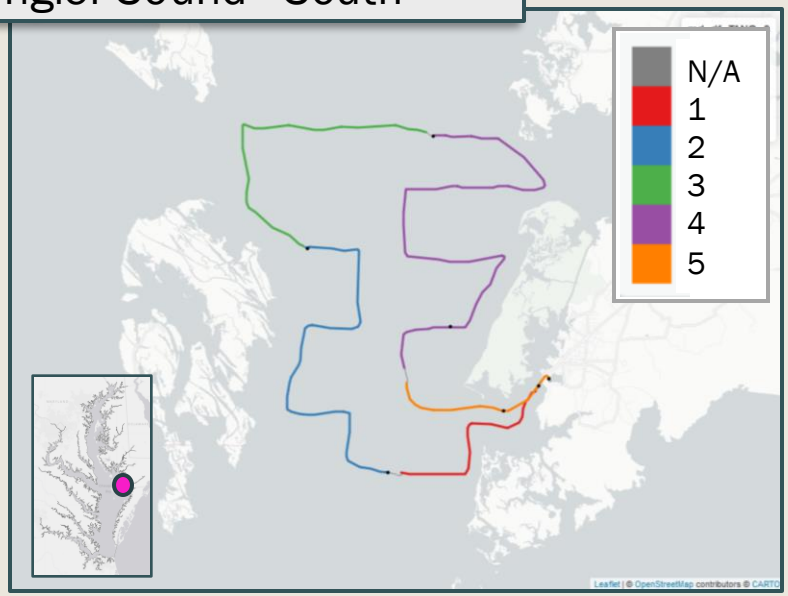
For clarity, plots show observed units (mg/L); analyses were performed with beta–logit–transformed DO, so results may differ slightly though patterns are unchanged.

Correlation: Horizontal (along- and cross-bay)

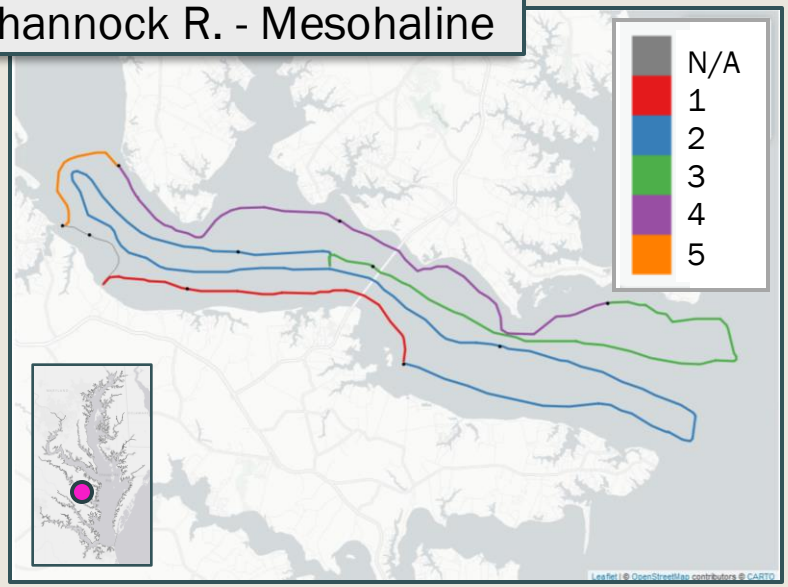
- DataFlow—high frequency underway surveys (~0.5 m intake)
  - **MD DNR Eyes on the Bay (EOTB):** ~63 stations, 2001-2023
  - **Virginia Estuarine and Coastal Observing System (VECOS):** ~36 stations 2003-2024
- Map lines show survey legs (1–5)

Source: https://eyesonthebay.dnr.maryland.gov/

Tangier Sound - South



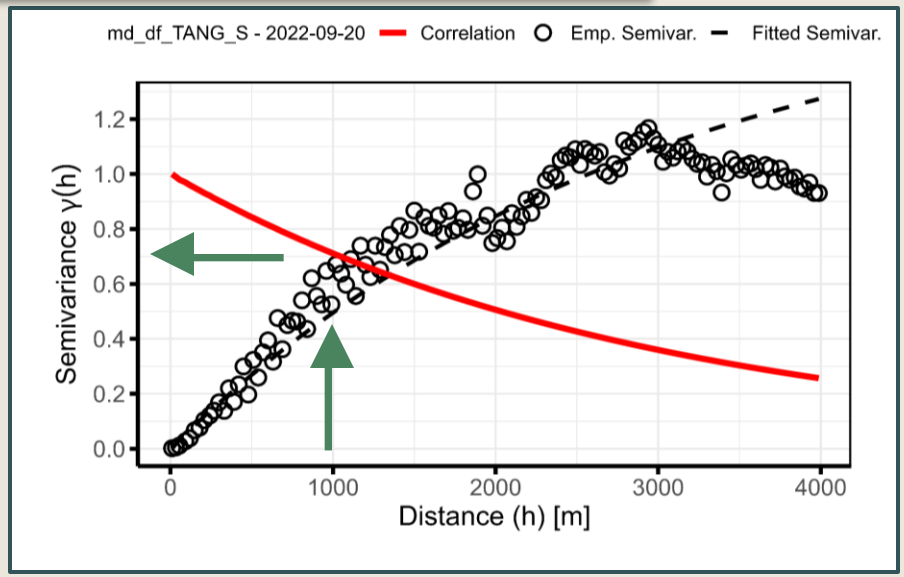
Rappahannock R. - Mesohaline



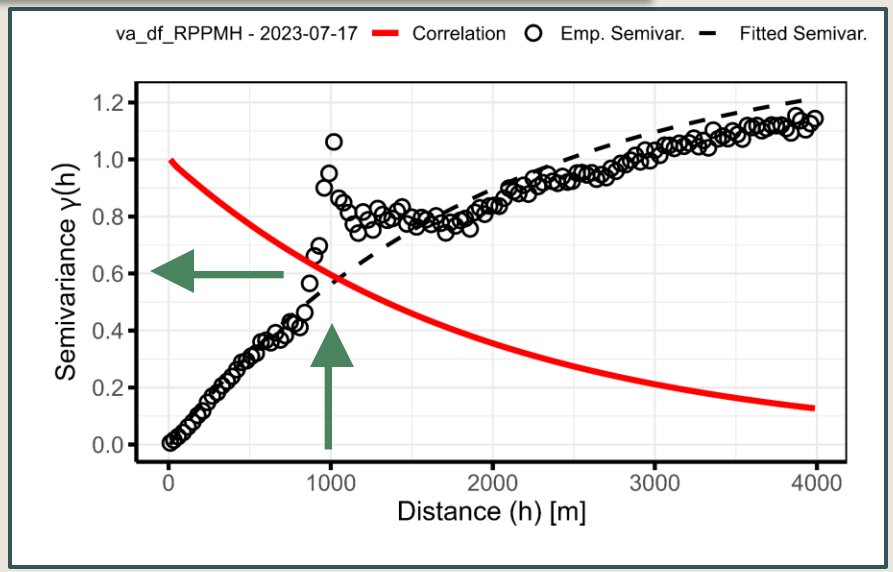
Correlation: Horizontal (along- and cross-bay)

- For each cruise, compute a variogram cloud by chunk (all point pairs within a chunk).
- Bin the clouds (30 m bins) and compute the empirical semivariogram.
- Fit exponential variogram.
- Use pairs out to 4,000 m maximum distance.
- Plot guide:
  - open circles = empirical semivariance.
  - dashed = fitted semivariogram.
  - red line = correlation decay vs. distance.

Tangier Sound - South

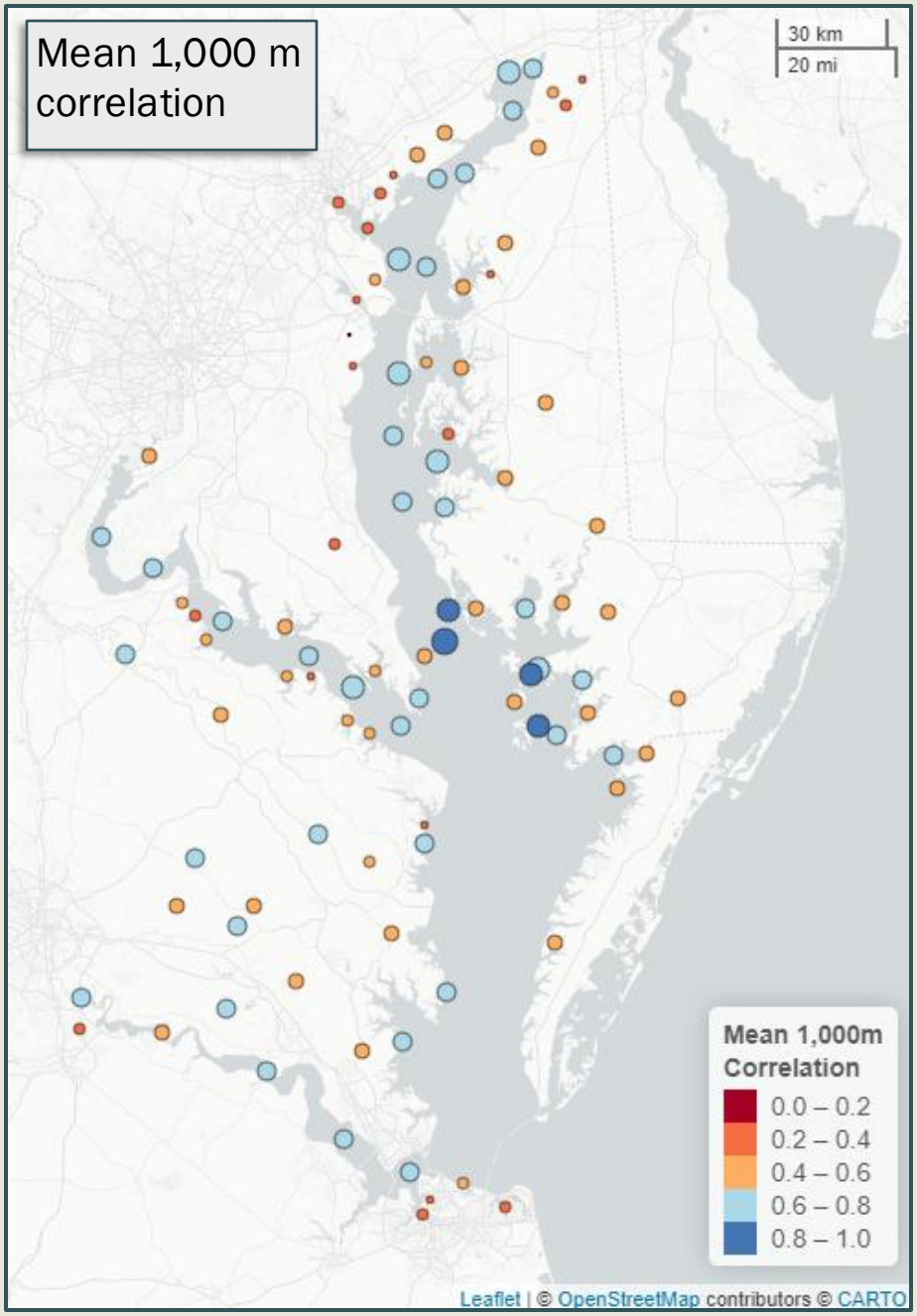


Rappahannock R. - Mesohaline



Correlation: Horizontal (along- and cross-bay)

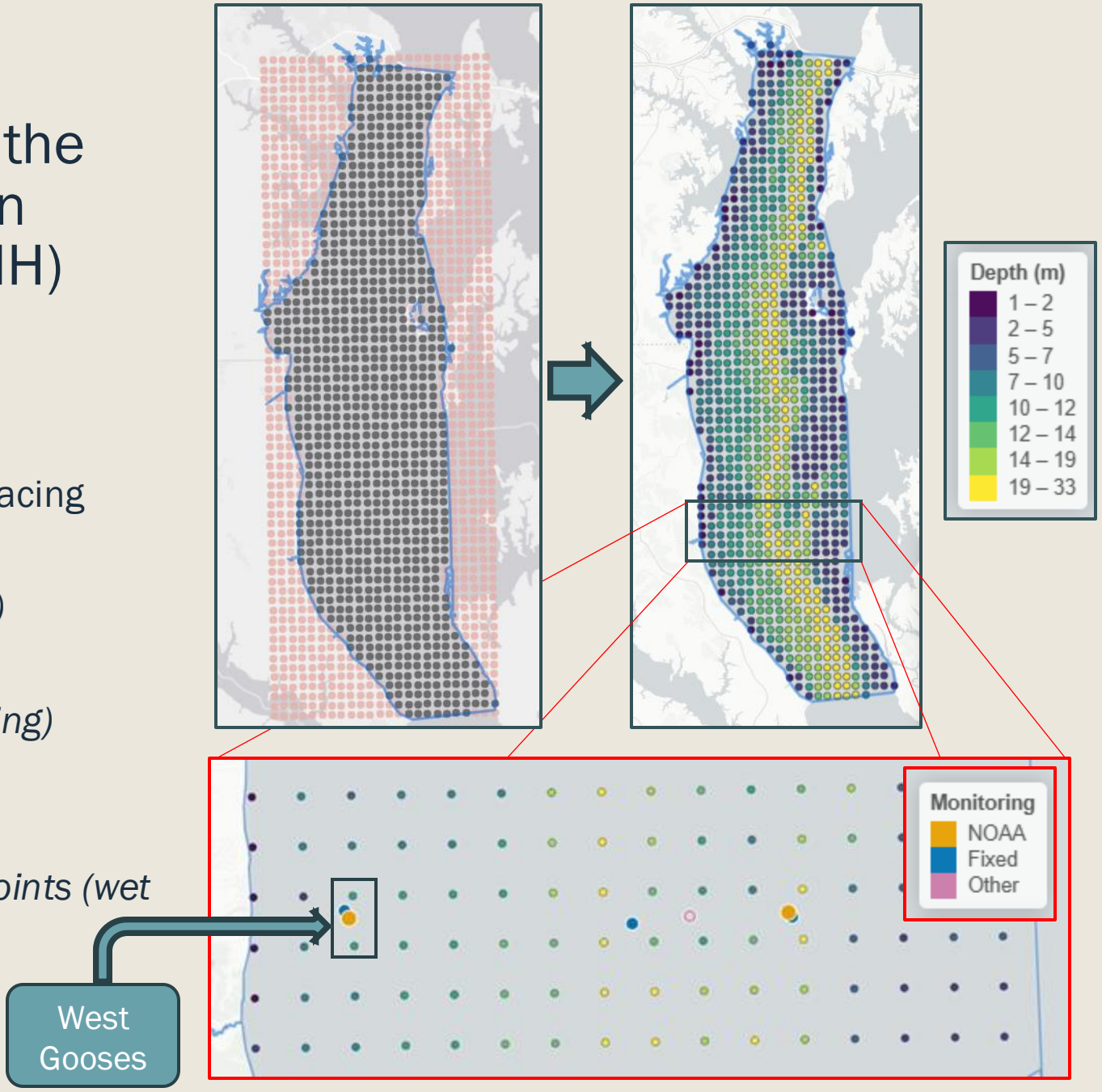
- **Highest (0.8–1.0):** mid-Bay near CB5MH_MD and Tangier Sound.
- **Moderately high (0.6–0.8):** widespread in the upper mainstem and larger tributaries.
- **Lower (0.2–0.4):** common in upper tributaries, small embayments, and nearshore.
- Across ~100 stations, most means fall 0.4–0.7.
- Spatial differences dominate variability more than temporal effects.



Constructing the 4-D prediction space (CB4MH)

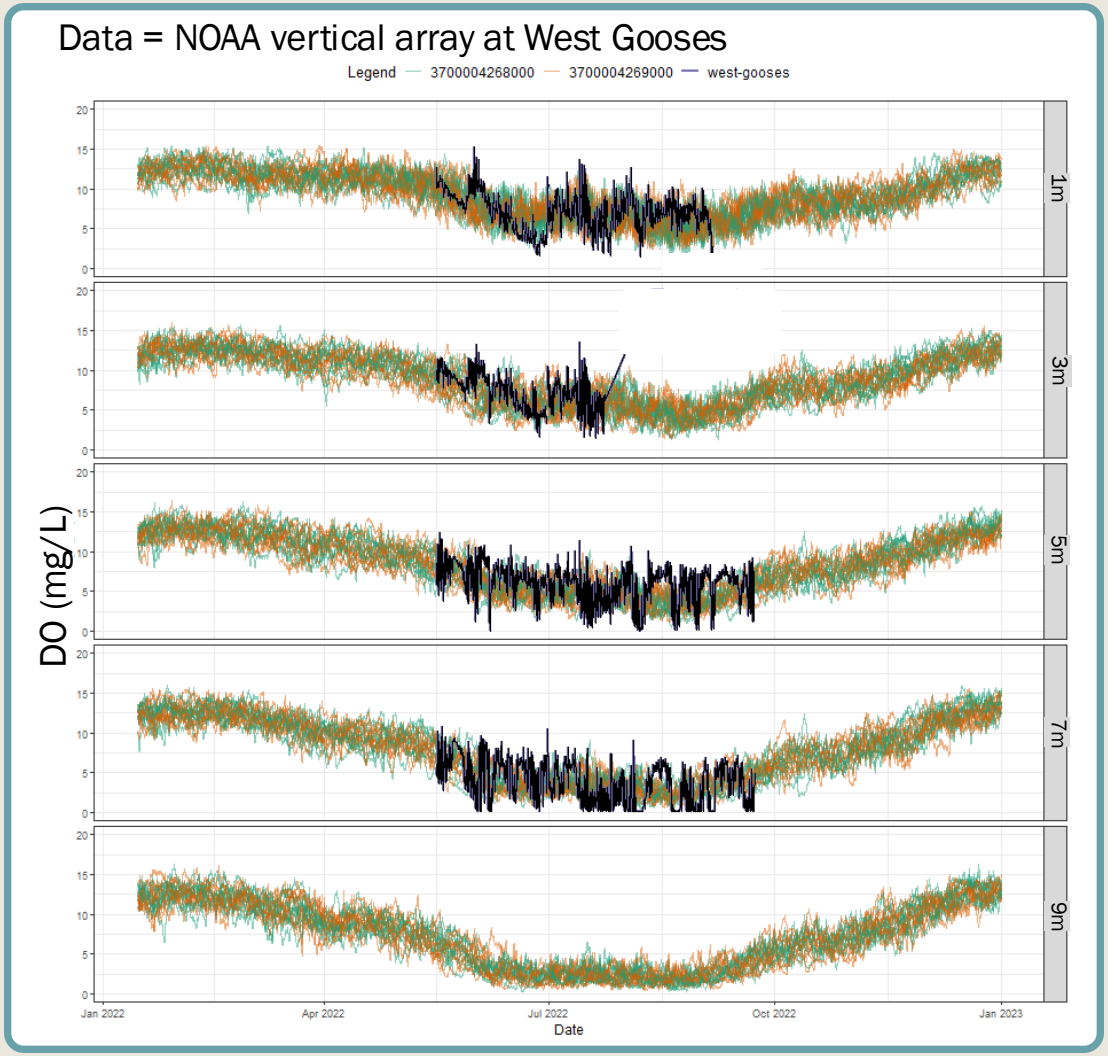
- **Horizontal grid**
  - *69 x 24*
    - 1 km x 1 km spacing
- **Depth levels**
  - *33 (1 m spacing)*
- **Time**
  - *365 (1 day spacing)*

- **Active cells**
  - *~850,000 4-D points (wet cells only)*



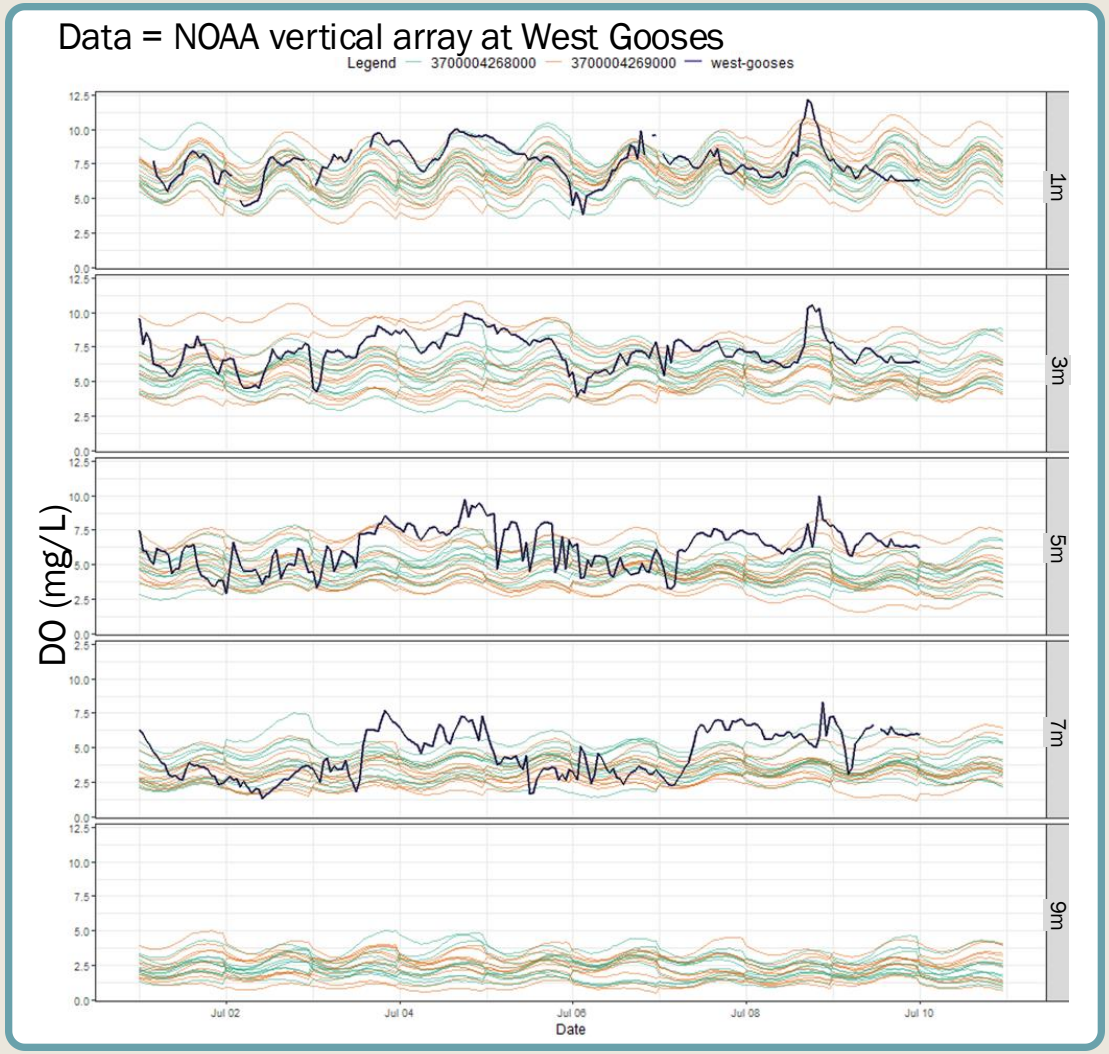
Statistical Interpolation – comparison with high frequency data

- Combine the pieces together to get one realization
- Generate multiple realizations to account for the range of possible conditions
- Example 10 realizations from grid cells near “West Gooses” vertical array
  - *2022*

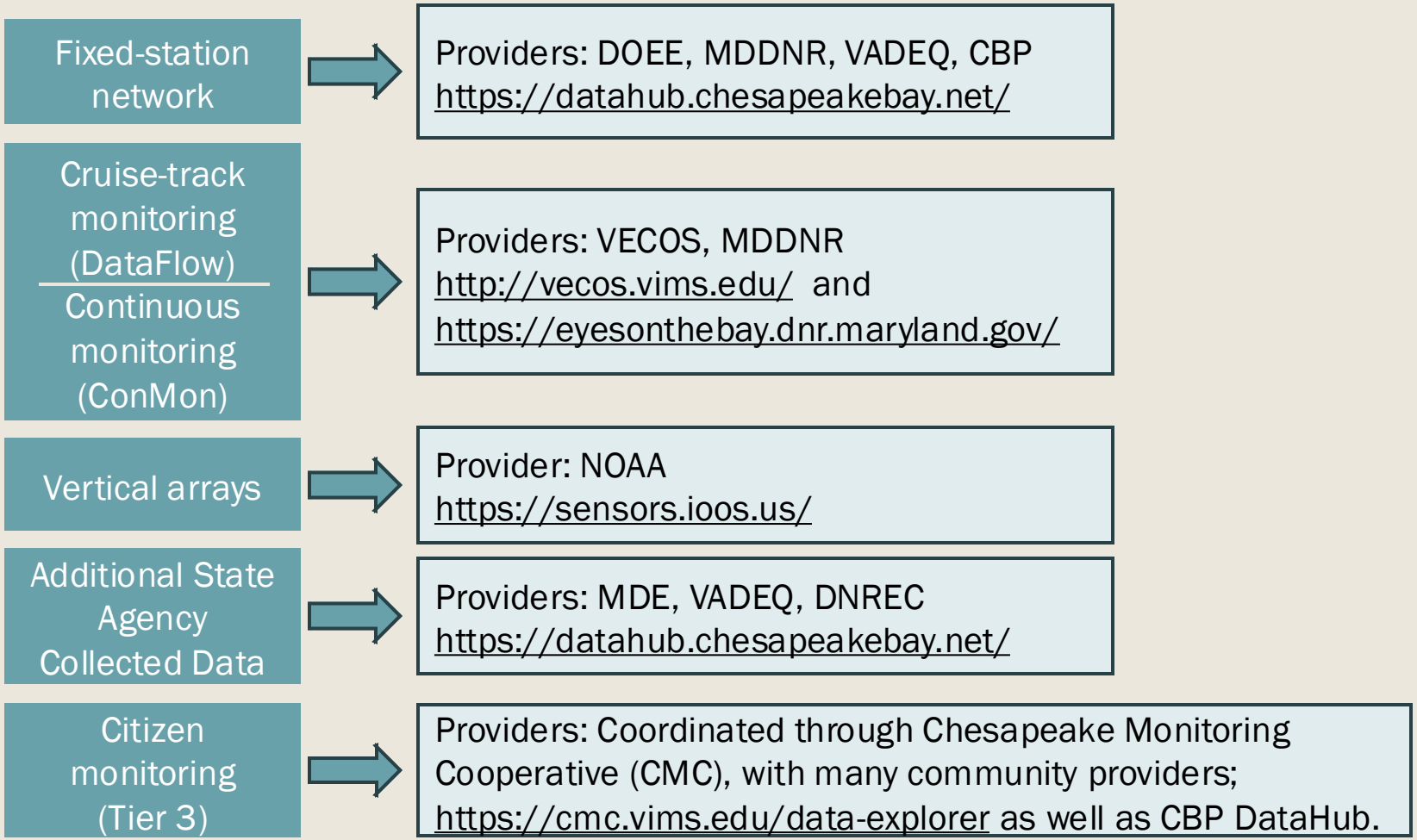


Statistical Interpolation – comparison with high frequency data

- Zoom in: 10 realizations from grid cells near “west-gooses” vertical array
  - *July 1-10, 2022*



Data sources



Data compilation thanks: Mike Mallonee (ICPRB); Mark Trice and Rebecca Burrell (MDDNR); David Parrish (VIMS) and Carl Friedrichs (VIMS); Jay Lazar and CJ Pellerin (NOAA); Liz Chudoba (Alliance for CB).