4D Interpolator

Large-scale Depth Correlation

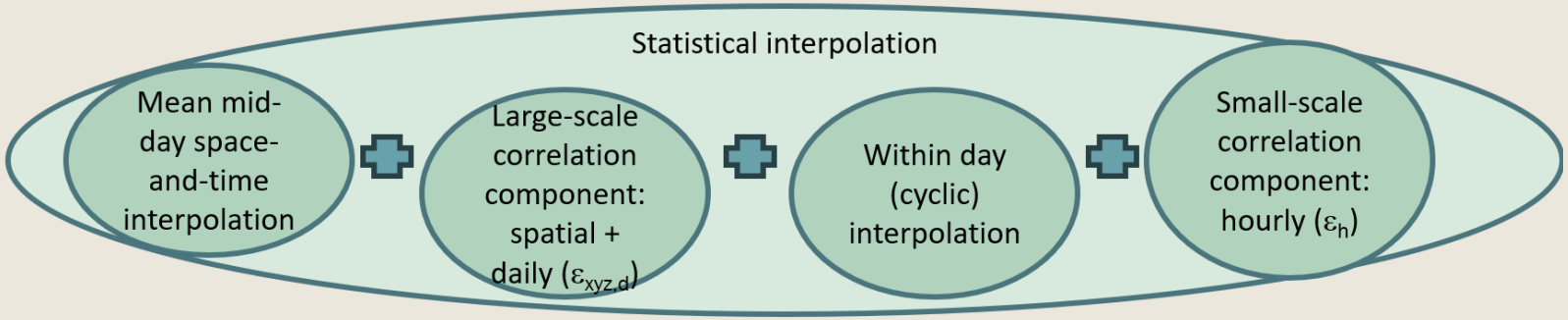
Bay Oxygen Research Group
July 28, 2026

Jon Harcum[1] for development team:
Elgin Perry[2], Rebecca Murphy[3], Breck Sullivan[4], Peter Tango[4]
[1]Tetra Tech, [2]Statistics Consultant, [3]UMCES at CBP, [4]USGS at the CBP

Outline

- Where depth correlation fits within the 4-D interpolator?
- What “depth memory” means and how it affects **simulated DO profiles?**
- How we estimate depth memory from **observed monitoring data?**
- Historical patterns, candidate predictors, and selected parameterization.

Building a Realistic DO Time Series in Components

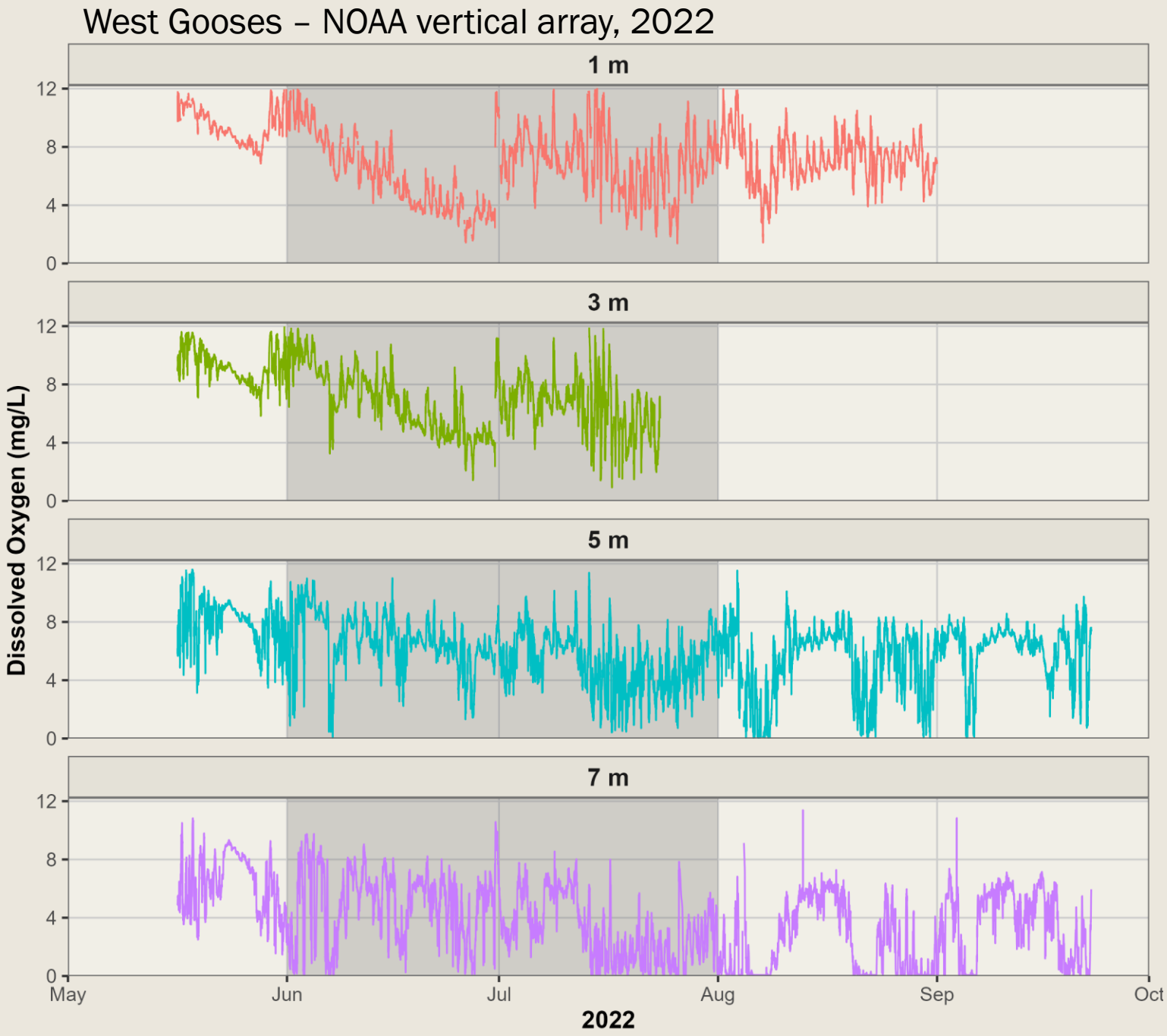


- Begin with the **expected mid-day DO pattern** across season, location, and depth.
- Add **large-scale correlated departures** from that pattern:
  - *day-to-day correlation*
  - *depth correlation*
  - *spatial correlation*
- Then add the **within-day cycle** and **hourly correlated variation**.

Today's focus: How depth correlation gives simulated conditions memory.

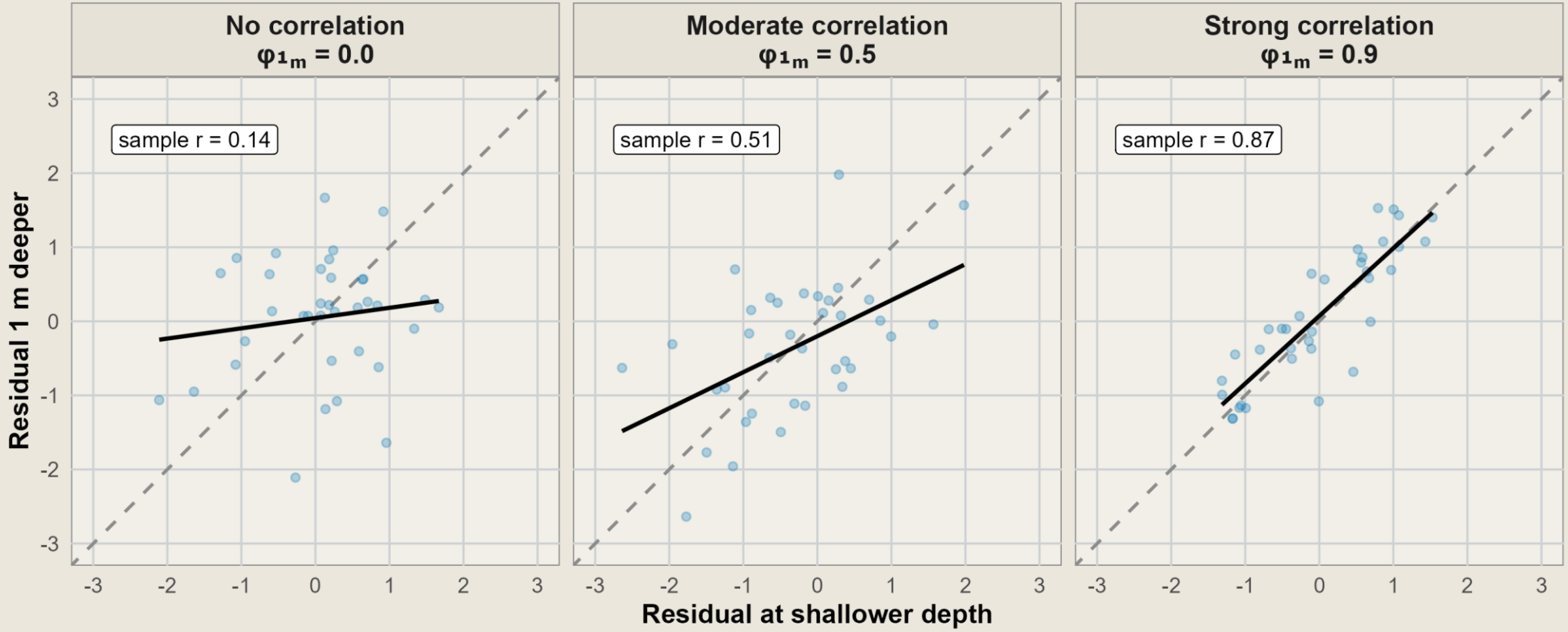
Depths Move Together, But Not Identically

- Nearby days carry useful information, but conditions can change quickly.
- Depths move together, but deeper waters often behave differently.
- These patterns motivate a separate depth-correlation term.
- **TODAY:** we focus on correlation among nearby depths.



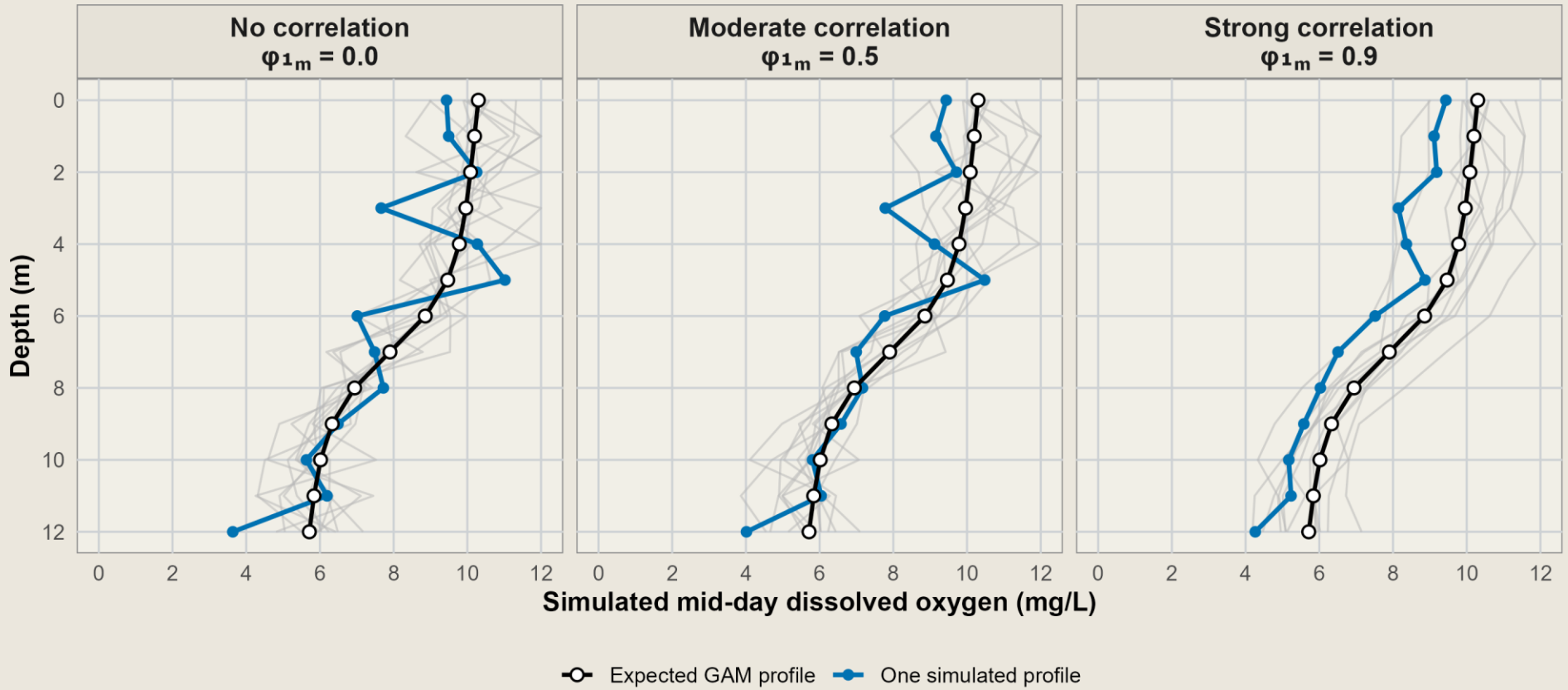
Adjacent-Depth Pairs
Measure 1-m Correlation

- Each point pairs a residual at one depth with the residual 1 m deeper.
- With low correlation, the points form a loose cloud.
- With high correlation, the points align along the diagonal.
- This adjacent-depth relationship is the local 1-m residual correlation.

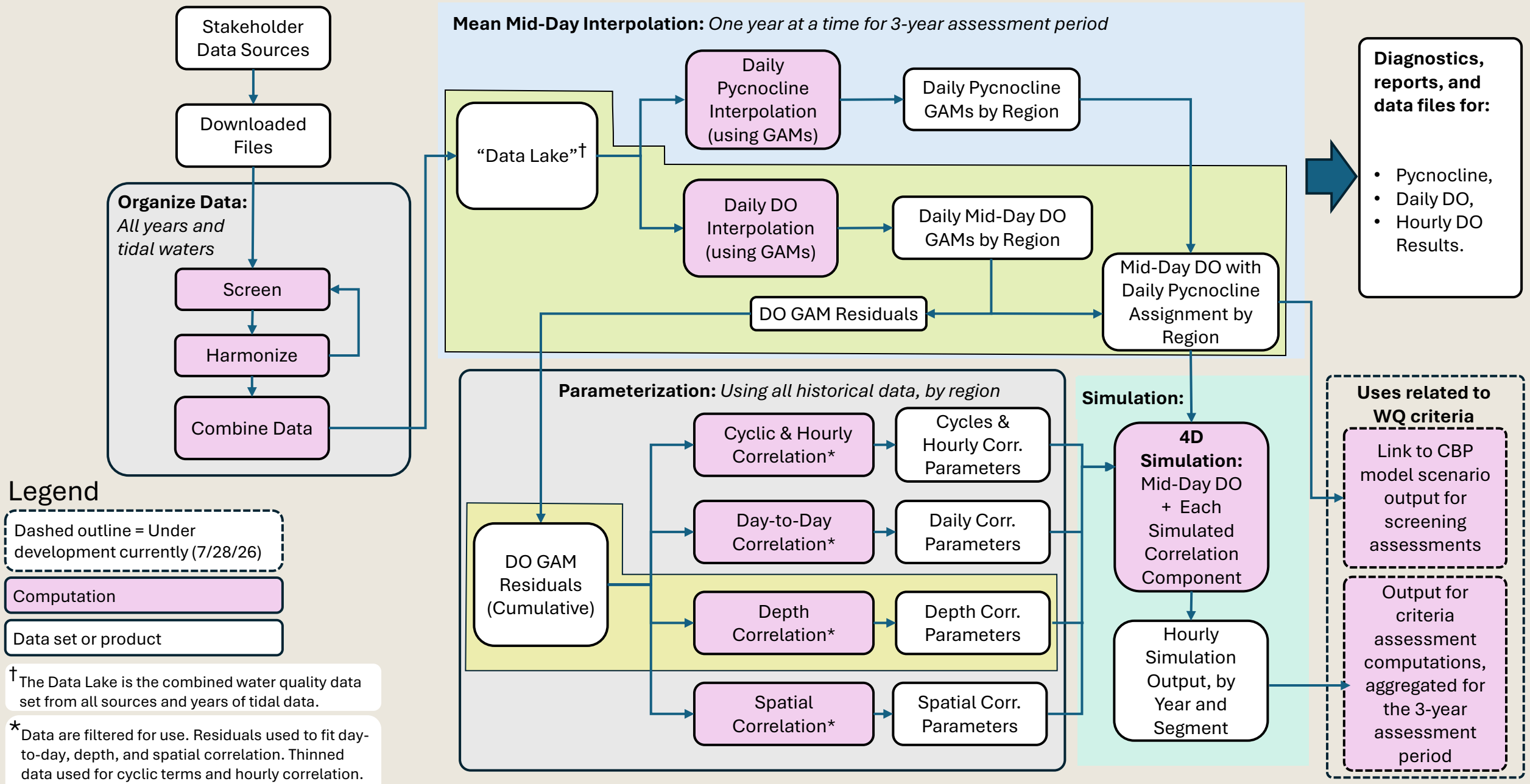


Why Parameterize Depth Correlation?

- The mid-day GAM provides the expected mean DO profile.
- Depth correlation controls how departures behave around that profile.
- Low correlation allows sharp depth-to-depth jumps.
- Higher correlation creates more vertically coherent residual layers.
- Realistic depth correlation is needed to mimic the vertical behavior observed in actual DO profiles.



4-D Interpolator Workflow

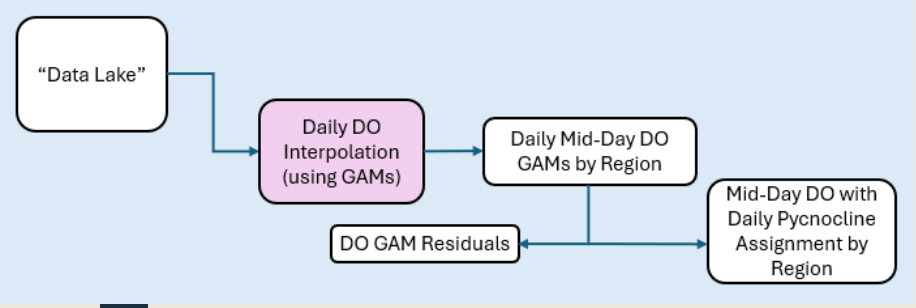


Legend

- Dashed outline = Under development currently (7/28/26)
- Computation
- Data set or product

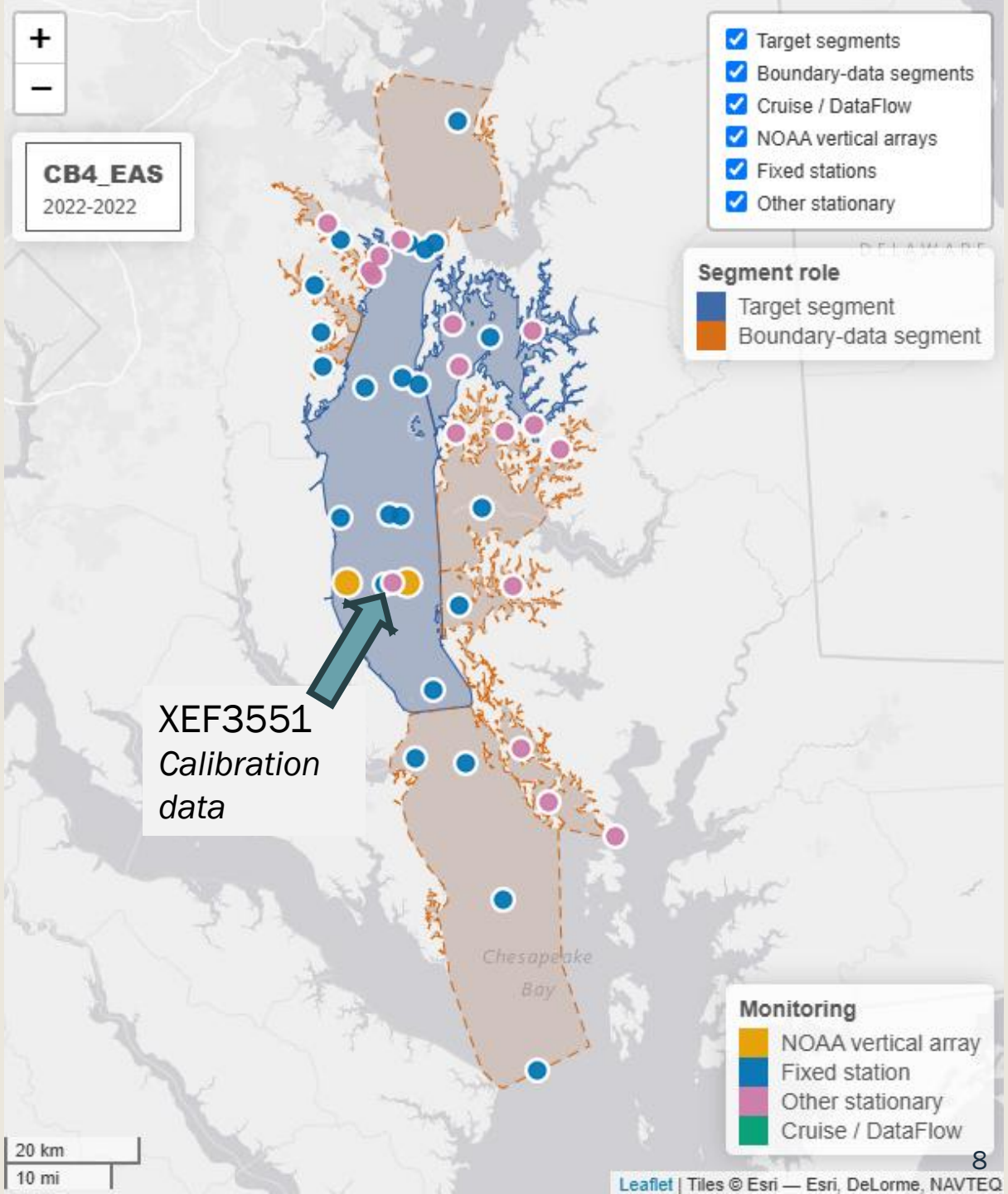
† The Data Lake is the combined water quality data set from all sources and years of tidal data.

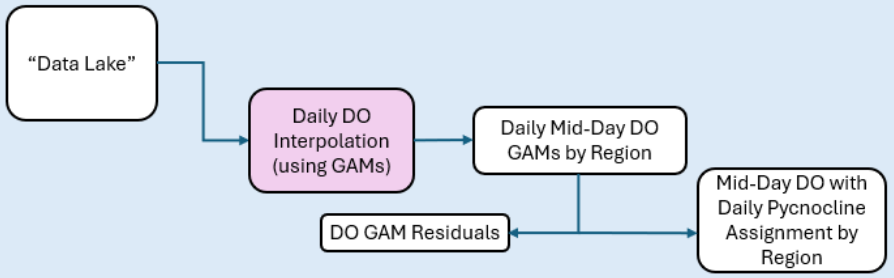
* Data are filtered for use. Residuals used to fit day-to-day, depth, and spatial correlation. Thinned data used for cyclic terms and hourly correlation.



Mid-Day GAMs Are Fit by Interpolator Region

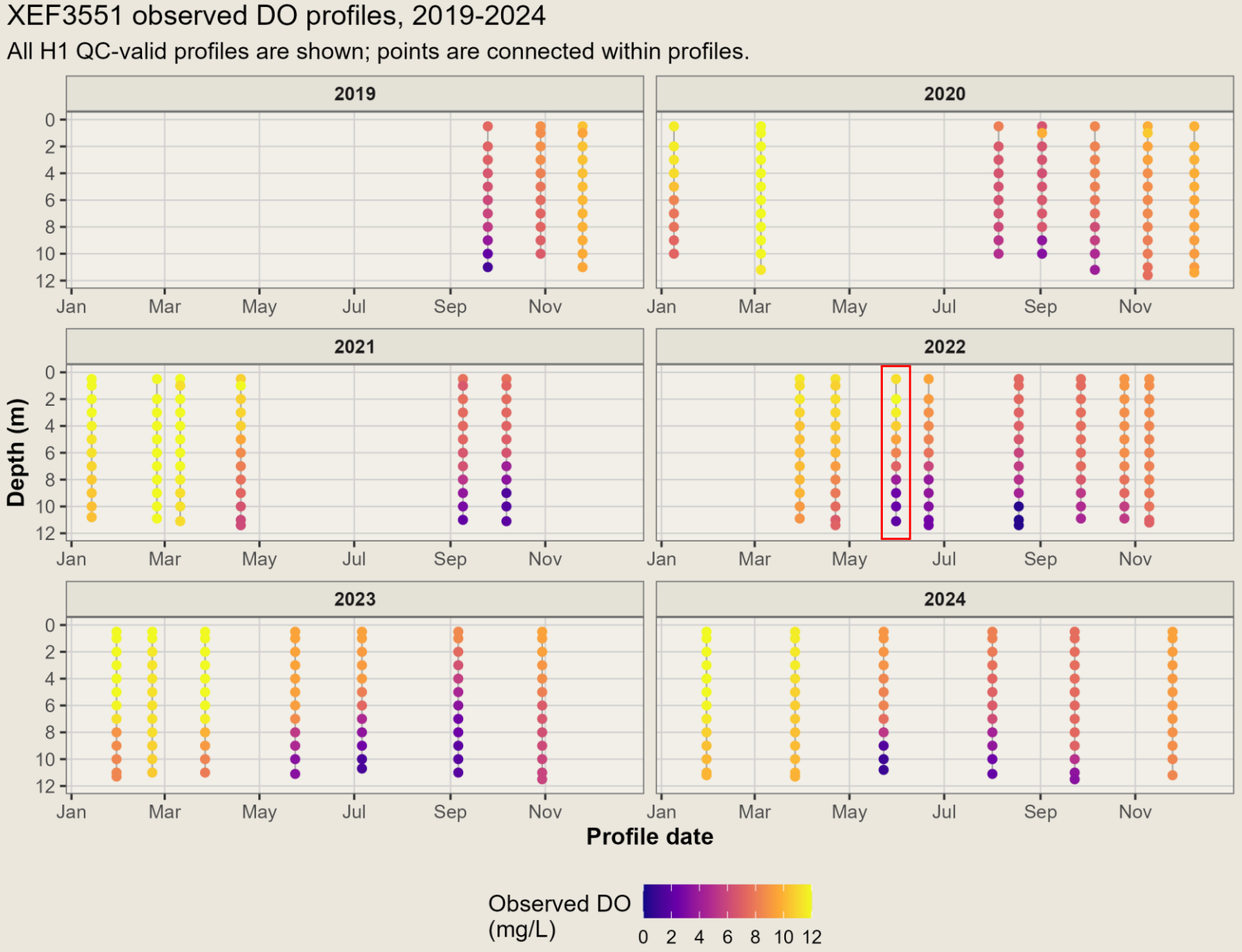
- **27 overlapping segment interpolator regions.**
- Target segments define the primary interpolation area.
- Boundary-data segments provide supporting observations near regional edges.
- XEF3551 is used as the worked example within **CB4_EAS**.

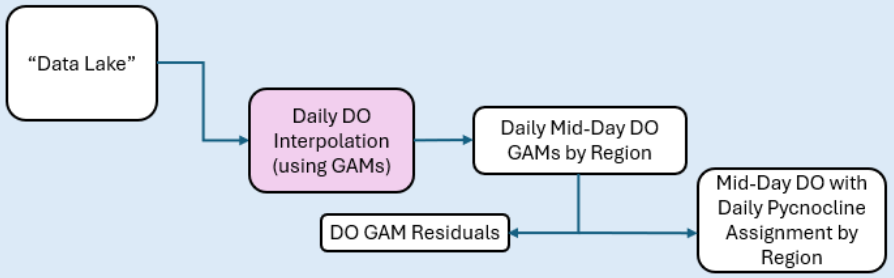




Observed XEF3551 Profiles Show the Raw Vertical Pattern

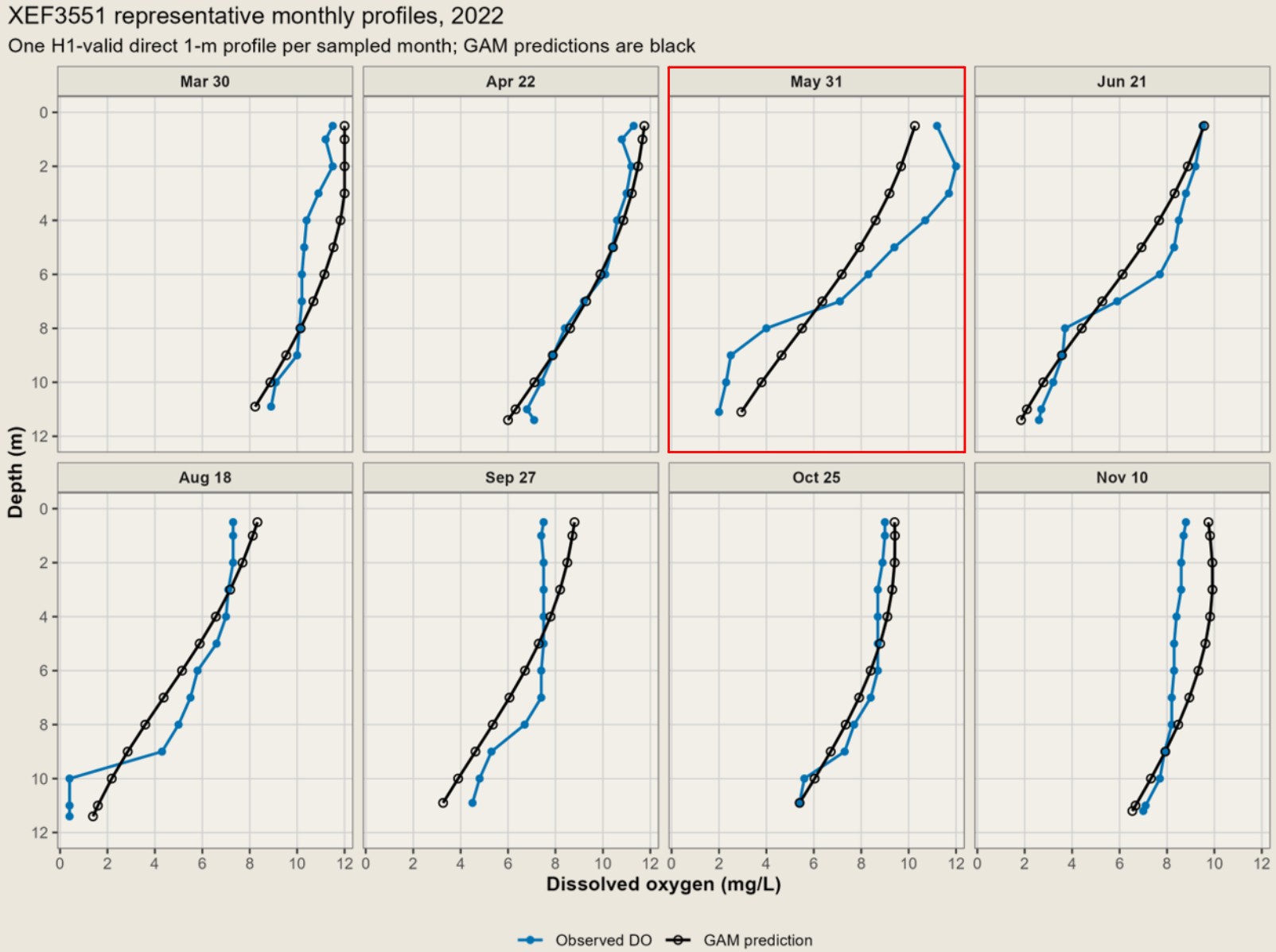
- XEF3551 provides repeated profile observations.
- Profiles generally span the upper 10–12 m of the water column.
- Seasonal changes in DO are clear.
- Vertical gradients are also clear, especially during lower-oxygen periods.

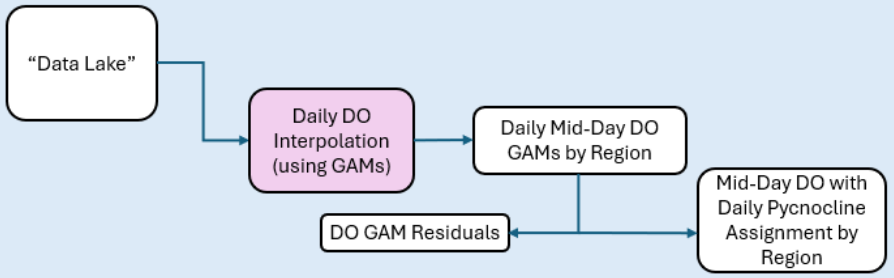




The GAM Captures the Expected Profile Shape

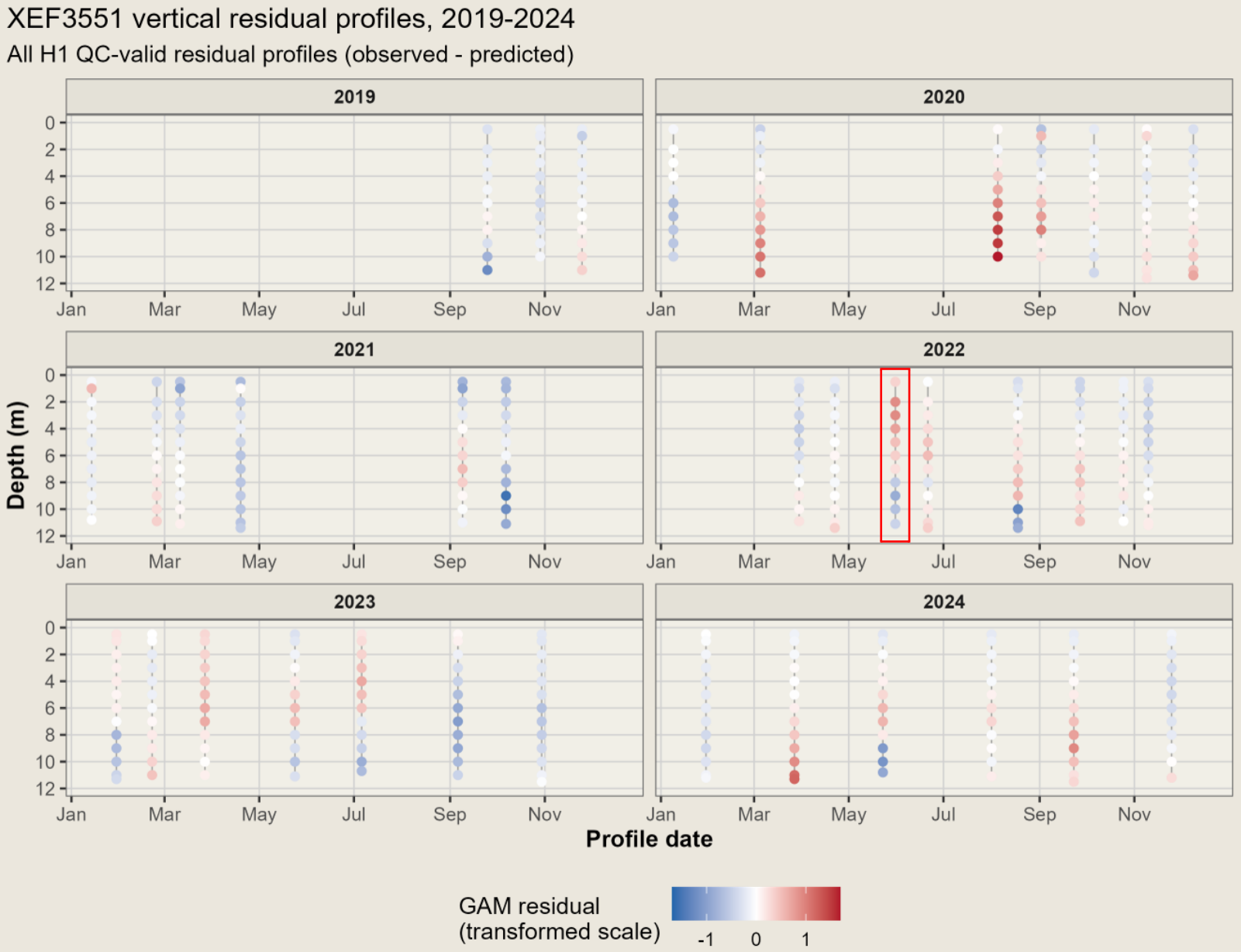
- Blue profiles show observed DO.
- Black profiles show expected mid-day DO from the regional GAM.
- The GAM captures broad seasonal and vertical structure.
- Differences from the GAM expectation often persist across several neighboring depths, indicating strong vertical dependence in the residuals.
- Remaining differences between observed and expected profiles become residuals.



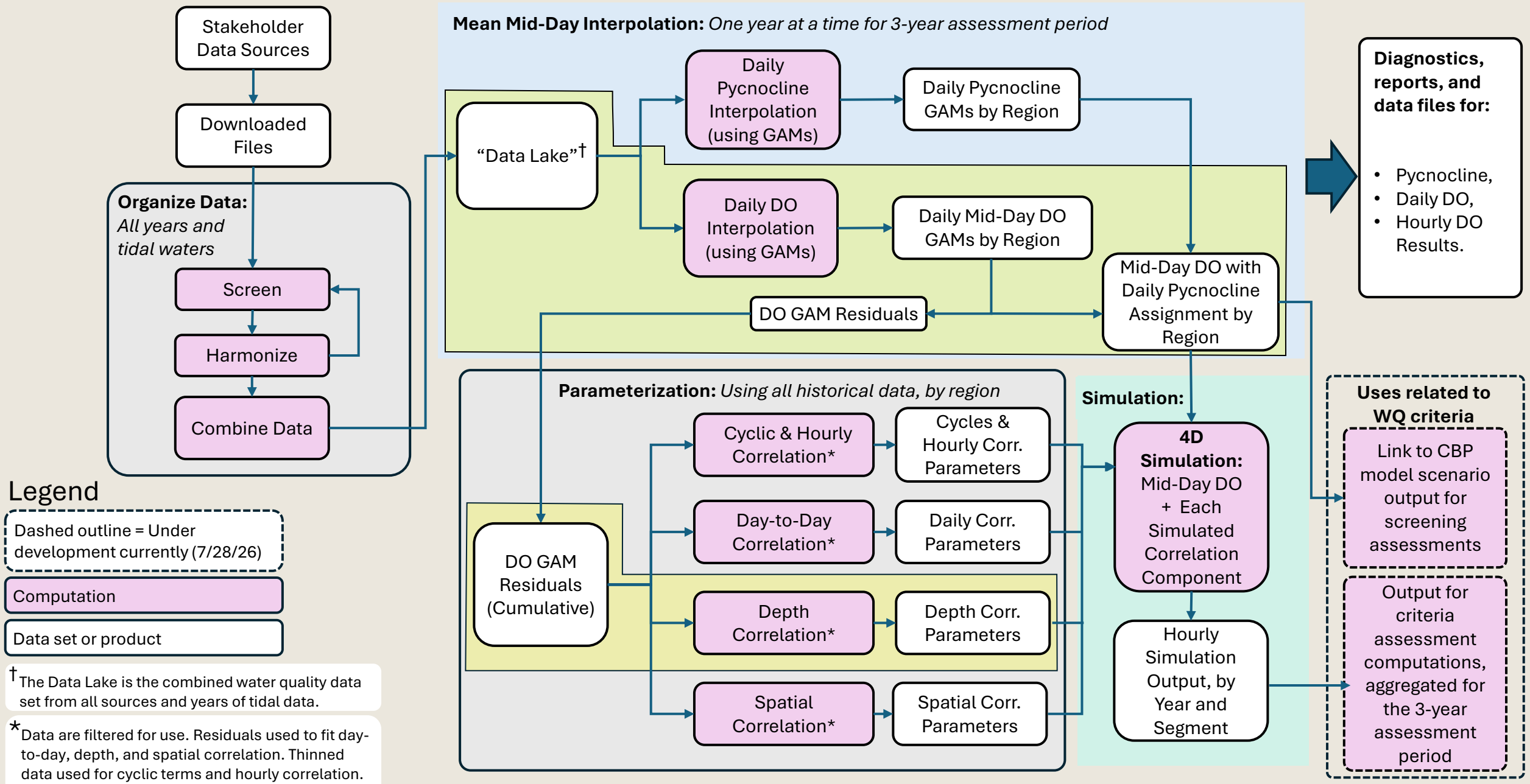


Residual Profiles Retain Vertical Structure

- Residuals are observed DO minus the GAM expectation.
- Positive residuals are above expected conditions.
- Negative residuals are below expected conditions.
- Adjacent residuals often move together vertically.
- The vertical structure of the residuals is what depth correlation measures.



4-D Interpolator Workflow

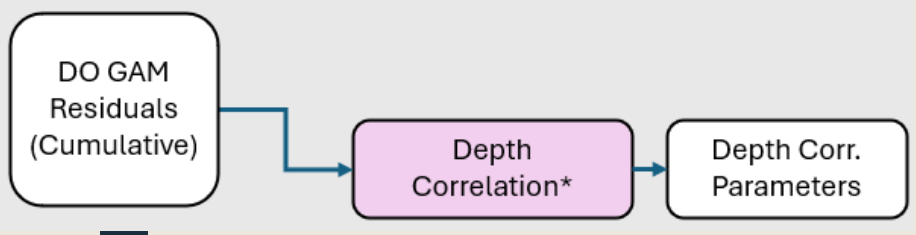


Legend

- Dashed outline = Under development currently (7/28/26)
- Computation
- Data set or product

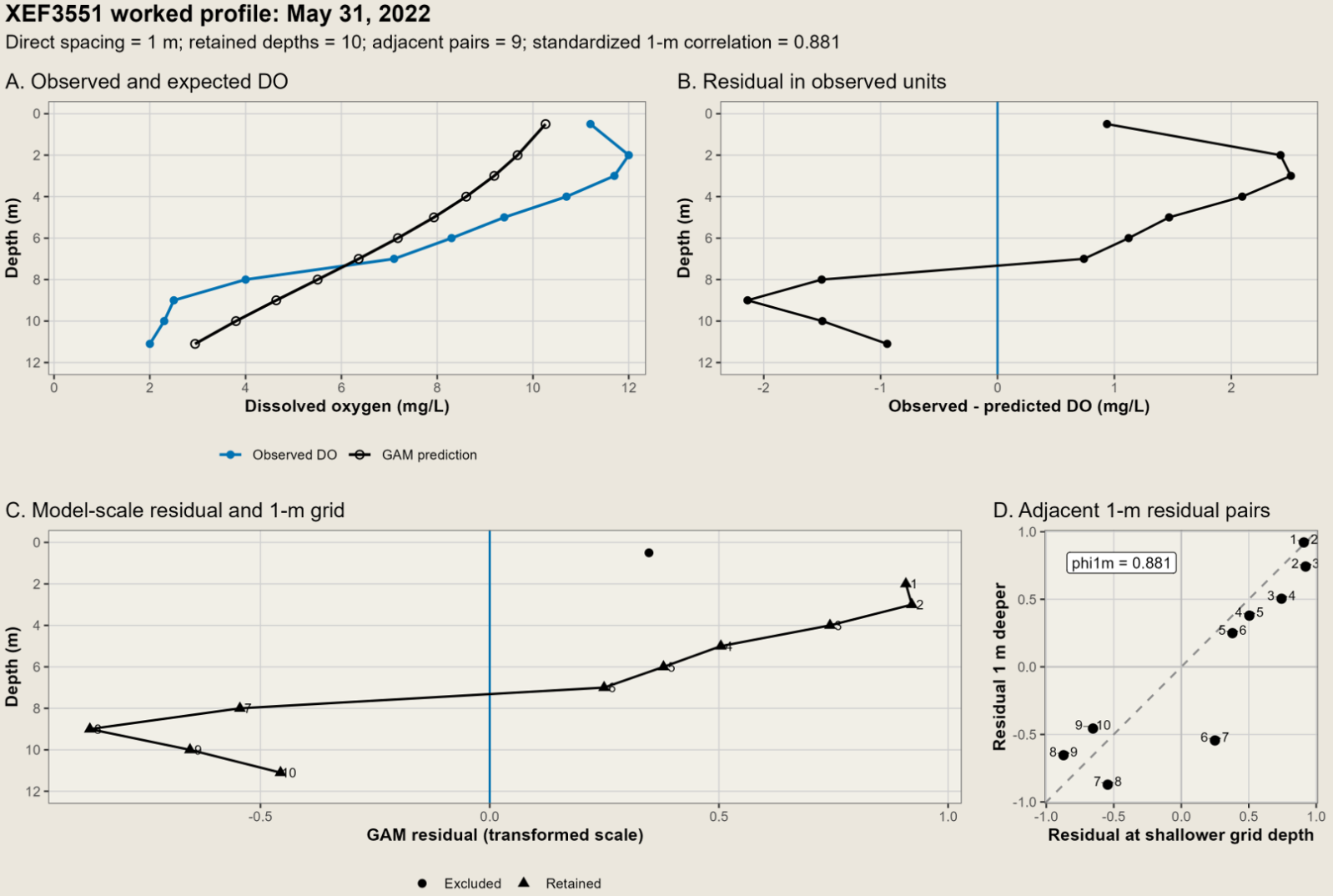
† The Data Lake is the combined water quality data set from all sources and years of tidal data.

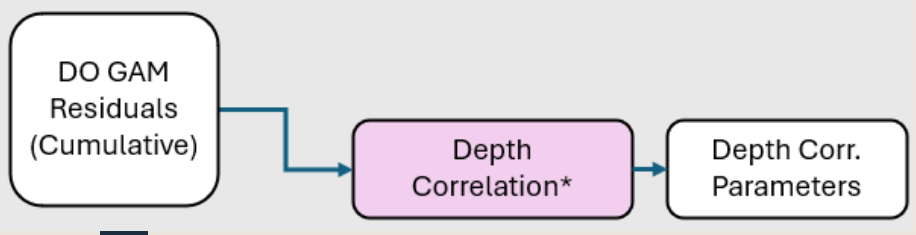
* Data are filtered for use. Residuals used to fit day-to-day, depth, and spatial correlation. Thinned data used for cyclic terms and hourly correlation.



From One Profile to One Depth-Correlation Estimate

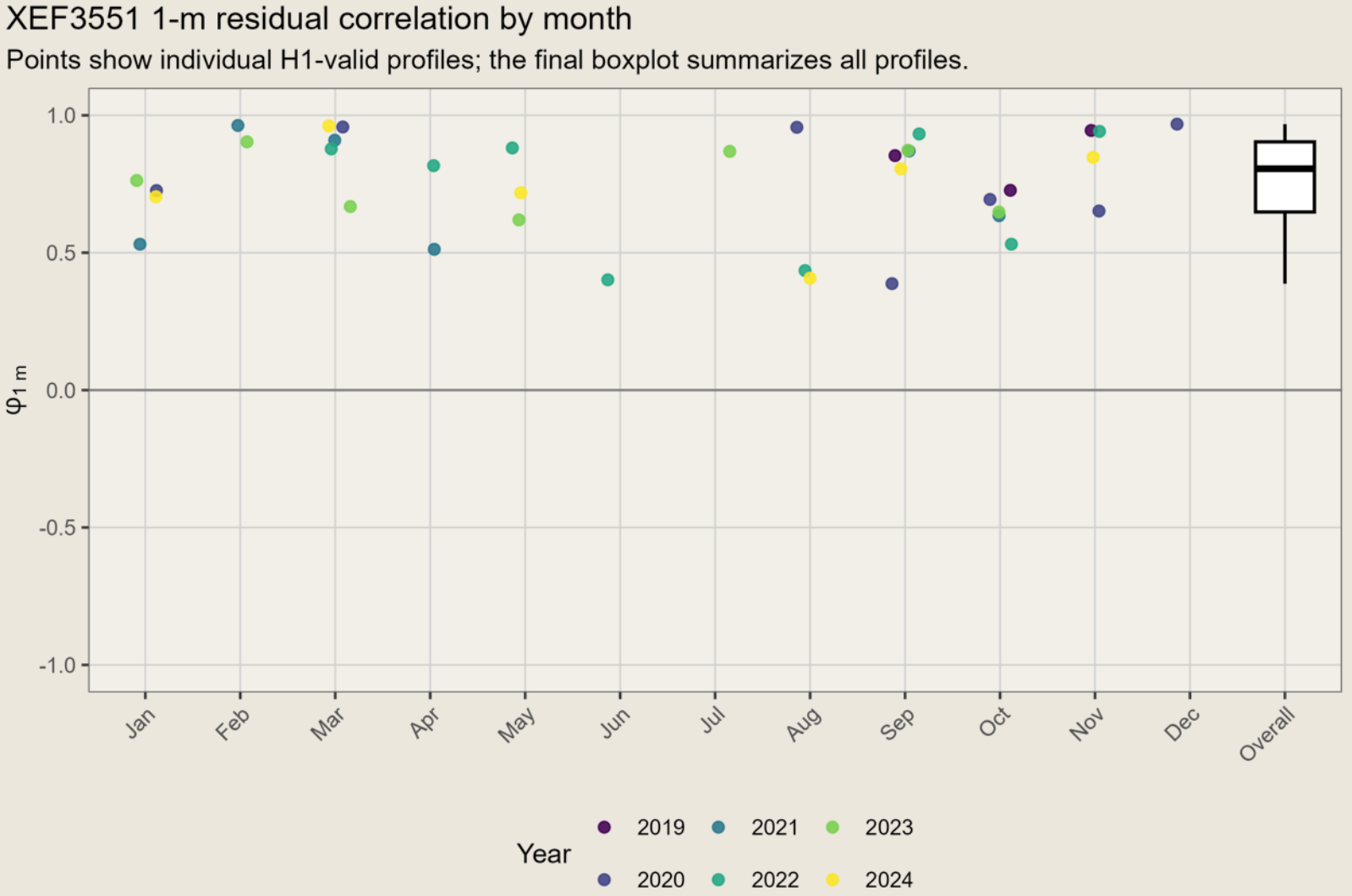
- May 31, 2022 provides a clean worked example.
- First compare observed DO with the GAM-expected profile.
- Then compute the residual profile.
- Finally pair adjacent retained depths.
- This profile gives a strong 1-m correlation: φ_{1m} = 0.881.

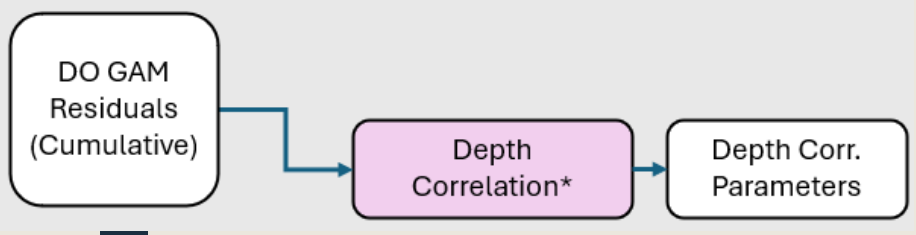




Multiple years of data at XEF3551 Support the Worked Example

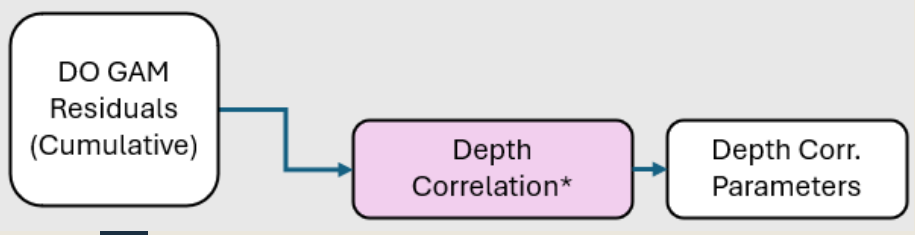
- The worked profile is not an isolated case.
- Most 1-m residual correlations are positive and strong.





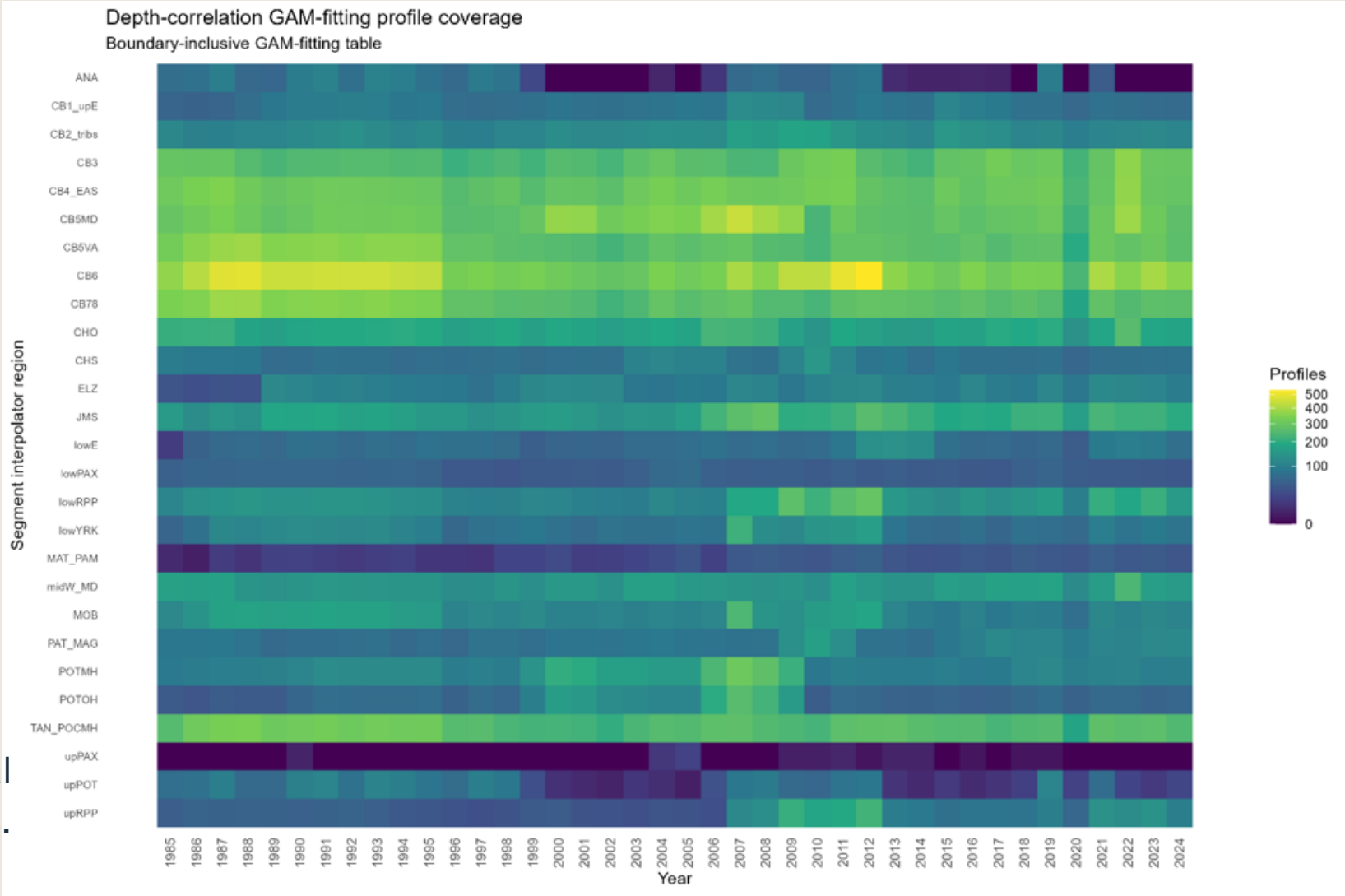
Build Eligible Station-Date Profiles

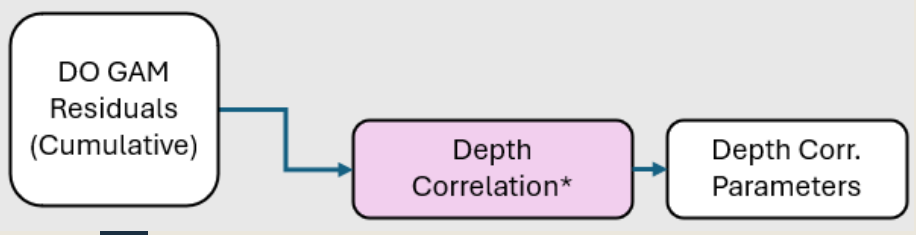
- Construct candidate profiles within each source, type, station, and date.
- Retain CBP DataHub profiles as observed daily profiles; for profiler and vertical-array data, select the profile nearest 11:00 within ±2 hours.
- Require at least **5 finite raw-depth observations** in each profile.
- Test regular grids in order: fixed 1 m, fixed 2 m, flexible 1 m, then flexible 2 m; retain the first grid that passes.
- Require at least **5 retained regular depths** and **4 adjacent grid-index pairs**.
- Retain 1-m estimates directly; convert accepted 2-m estimates to implied 1-m correlation before Fisher-Z transformation.



Historical Profiles Provide Regional Support

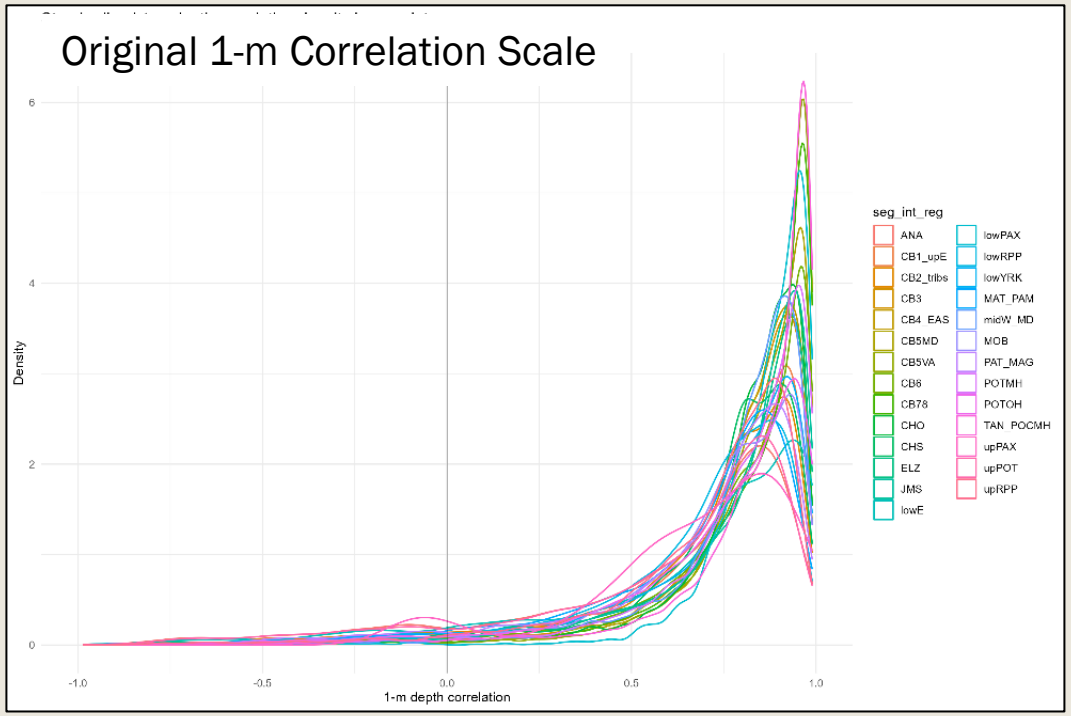
- The target-only table contains 52,498 eligible profile estimates.
- The boundary-inclusive fitting table contains 159,930 profile-region records.
- All 27 interpolator regions are represented.
- The record spans 40 years.
- Boundary support improves regional model stability near segment edges.

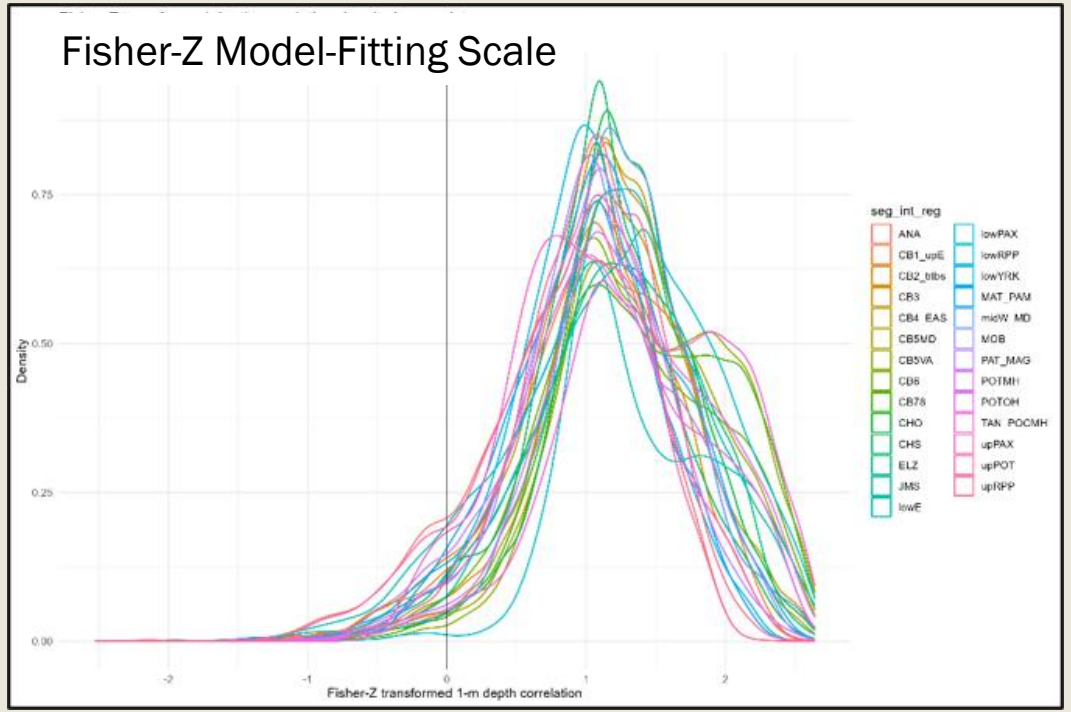


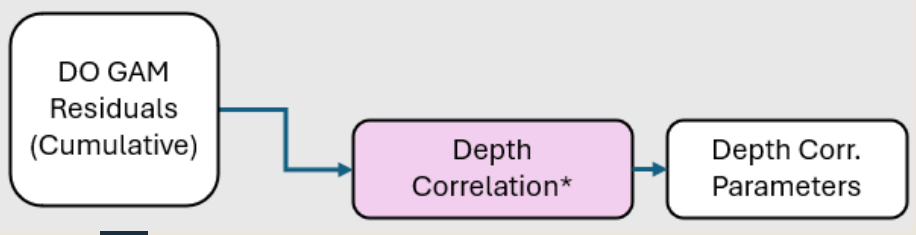


Depth Correlation Is Generally Strong and Positive

- All regional distributions are centered between approximately 0.75 and 1.0.
- This indicates strong vertical persistence over a 1-m separation.
- Distributions vary among regions.
- That variation supports regional parameterization rather than one Bay-wide value.

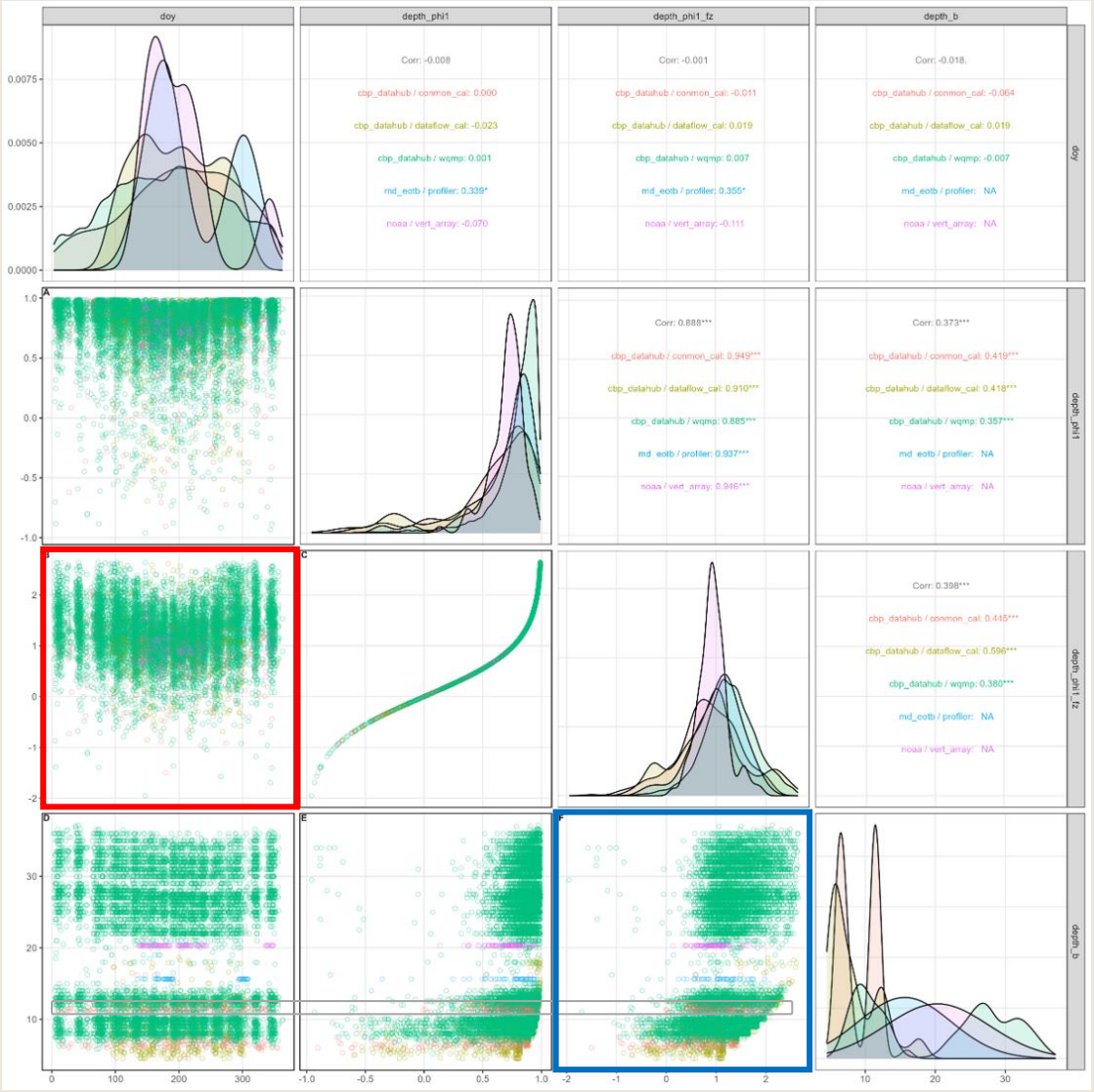


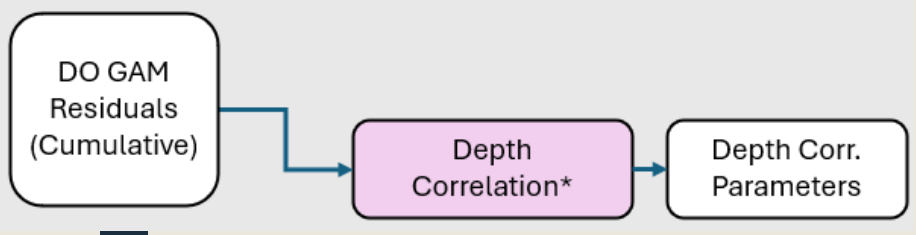




CB4_EAS Shows Seasonal and Bottom-Depth Structure

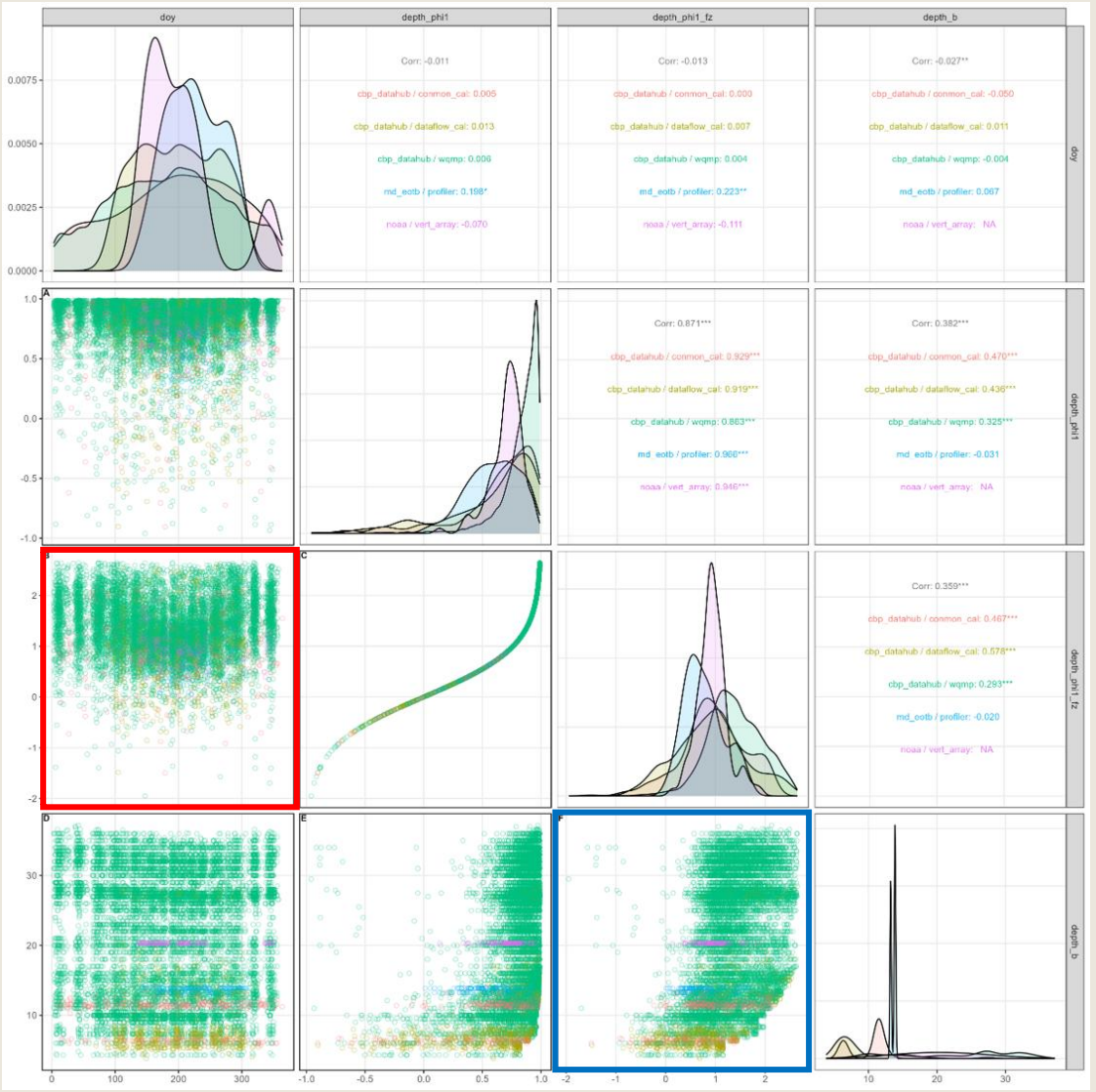
- CB4_EAS is the region used for the XEF3551 example.
- Exploratory plots show seasonal structure.
- They also show bottom-depth structure.
- This supports considering both doy and depth_b as production predictors.

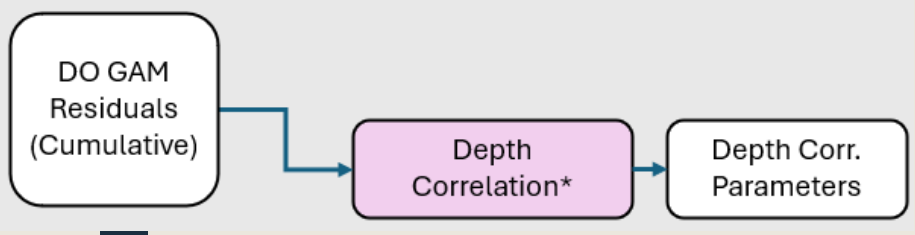




CB5MD Reinforces the Bottom-Depth Signal

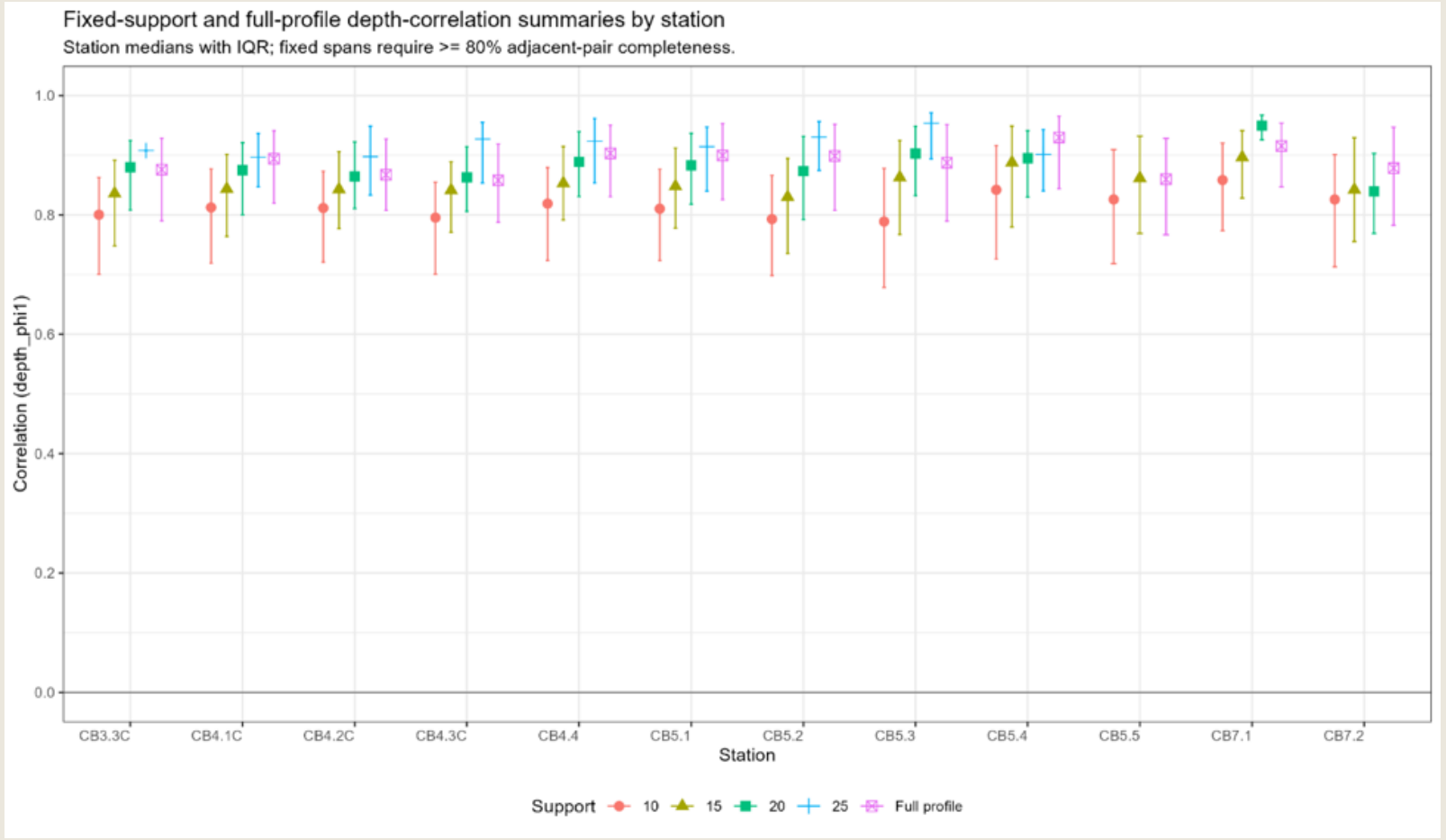
- CB5MD provides a second regional example.
- Depth correlation remains generally strong.
- The relationship with bottom depth is visible.
- Seasonal variation is also present.
- This strengthens the case that bottom depth is not just a local artifact.

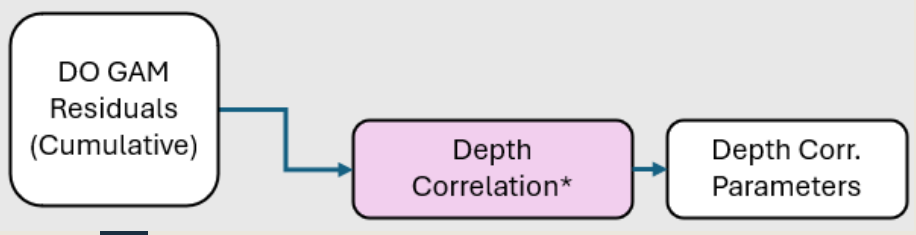




Sensitivity Analysis: Fixed-Support Results Across Mainstem Stations

- Deeper profiles generally contain more observations and adjacent 1-m pairs.
- The analysis tests whether the apparent bottom-depth relationship is mainly a vertical-support or sample-size artifact.
- Correlations were recalculated using fixed spans of 10, 15, 20, and 25 adjacent 1-m pairs.
- Shorter fixed spans generally produced somewhat lower correlations than full-profile estimates.
- Among matched profiles, the median differences were approximately −0.048, −0.020, and −0.007 for the 10-, 15-, and 20-pair spans.
- Correlations remained strong across all evaluated stations and support levels.
- Vertical support length affects the estimated magnitude but does not explain the broader depth-correlation pattern.





Selected Depth-Correlation Parameterization

- Models are fit separately by interpolator region on the Fisher-Z scale.
  - *depth_phi1_fz ~ s(doy, bs = "cc", k = 6) + s(depth_b, bs = "cs")*
- Day of year and bottom depth account for uneven historical sampling across seasonal and depth conditions.
- The fitted GAM provides a balanced regional summary rather than allowing densely sampled conditions to dominate.
- For the current simulation, predictions are back-transformed, clamped, and summarized to one regional depth-correlation value.
- Depth correlation is therefore not currently varied by season or bottom depth within the simulation.