How are you currently using or plan to use the data from the hypoxia stations to support projects?

4-dimensional (4-D) interpolator

Rebecca Murphy (UMCES/CBP)
With key input from Elgin Perry (consultant) and Jon Harcum (Tetra Tech)
Hypoxia Collaborative
Oct. 16, 2024

Goal for BORG team & 4-D tool

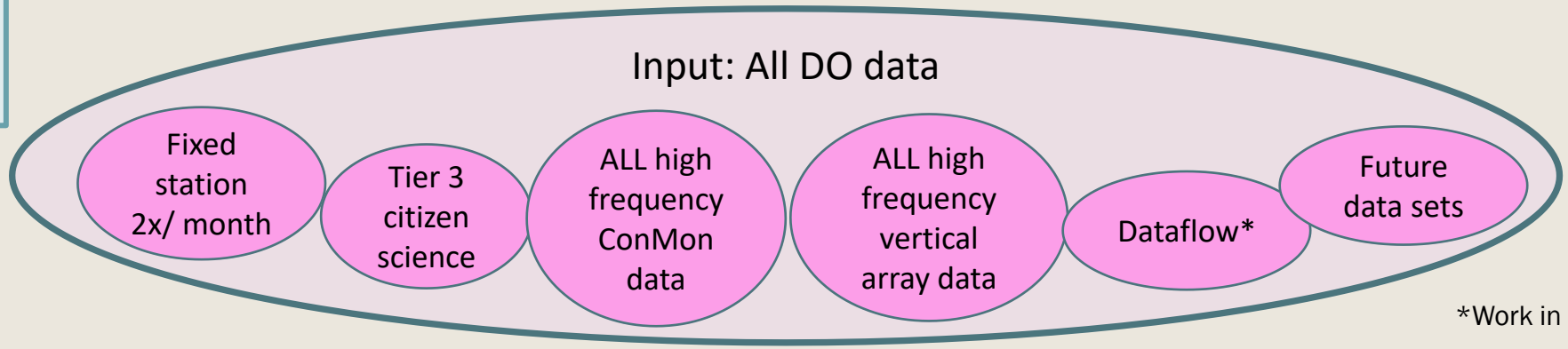
To develop a spatial-and-temporal interpolation tool for water quality monitoring data collected in the tidal waters of the Chesapeake Bay, thus enabling the evaluation of both long- and short-duration water quality criteria.

Specifically, the tool should be able to:

- *Interpolate observed dissolved oxygen in space and time ("4D"),**
- *Provide statistical estimates of uncertainty,*
- *Reproduce daily and hourly variability of the data, and*
- *Allow for post-processing of the interpolation output into designated uses (DU).*

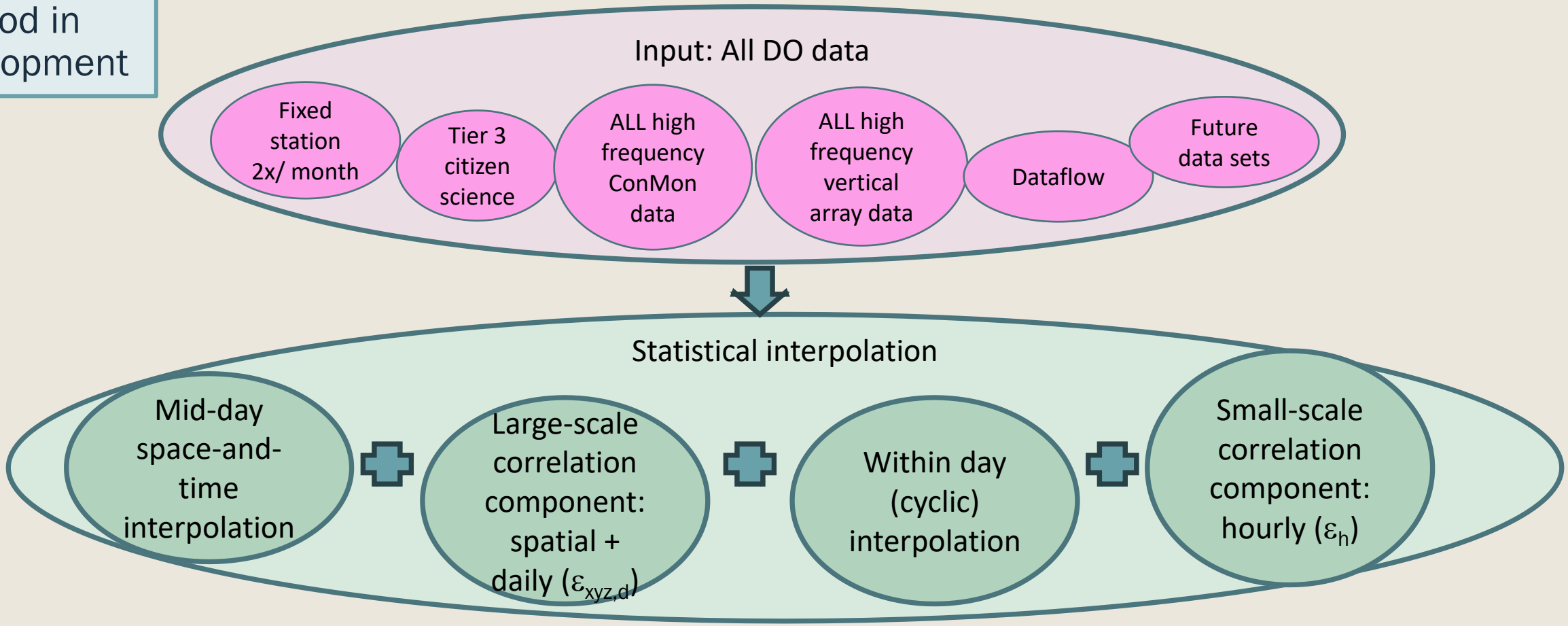
*Note: Focus on development so far has been on dissolved oxygen, but ultimately chlorophyll *a* and clarity may be evaluated as well.

Method in development

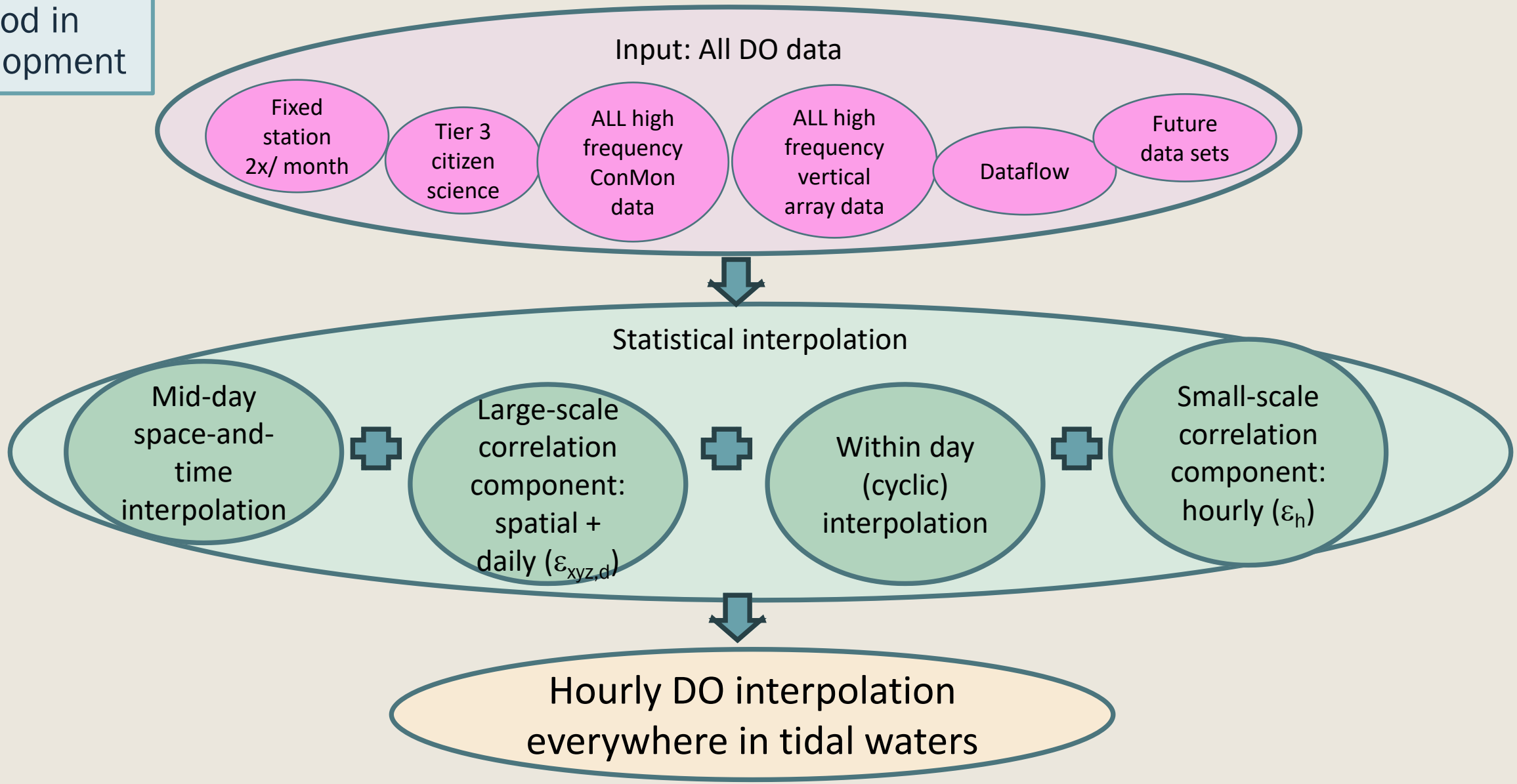


*Work in progress

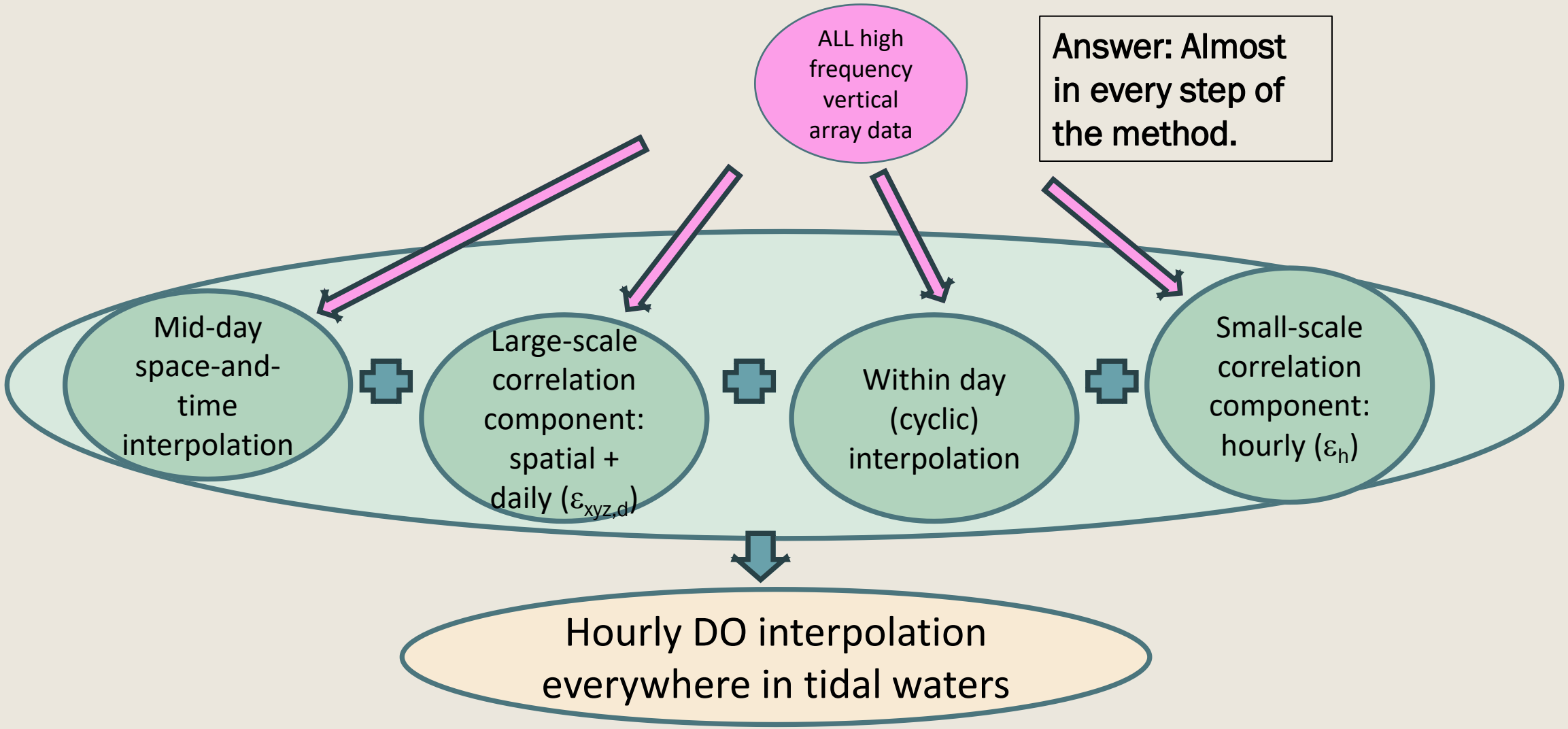
Method in development



Method in development



Question: Where is/will the vertical array data be used?

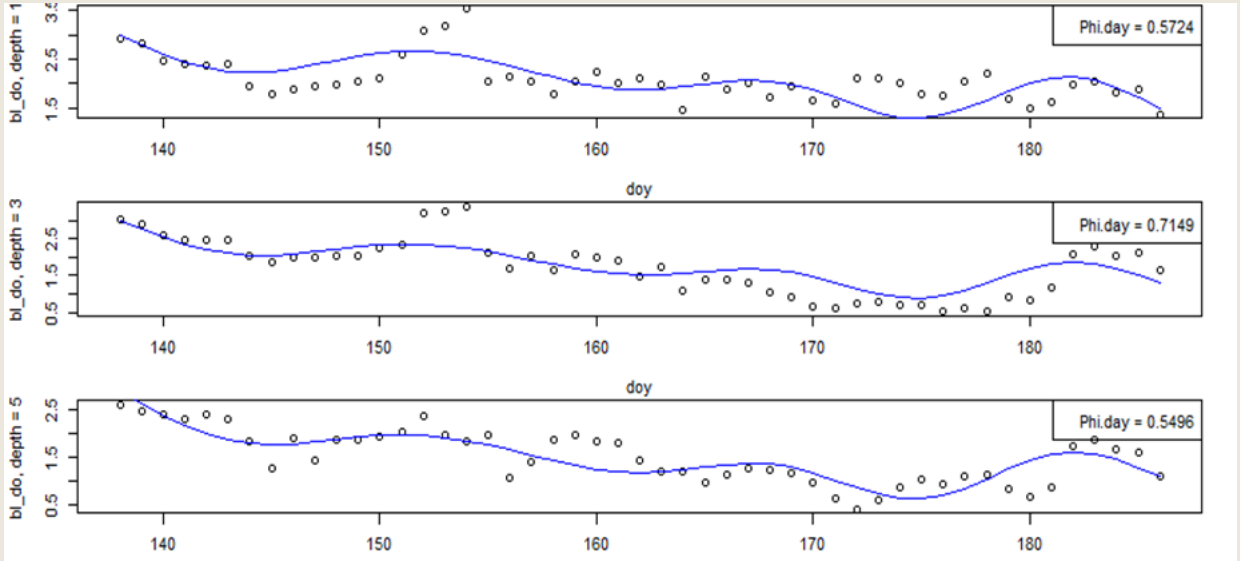


Uses/plans for vertical array data to-date:

1. Development and Testing
2. Parameterization
3. Implementation

1. Development and Testing

- We've compiled all data we have (including vertical arrays) into a data set that has been used to decide on the right statistical approaches for each part of the tool.
  - *Daily: Testing of the GAM to get the right equation*



From Elgin Perry

CBP DataHub
- 1984-2022
- 835 stations
- 819k obs.

EOTB: Eyes on the Bay
- 2001-2022
- 126 stations
- 11,916k obs.

VECOS: Virginia Estuarine and Coastal Observing System
- 2003-2022
- 54 stations
- 6,776k obs.

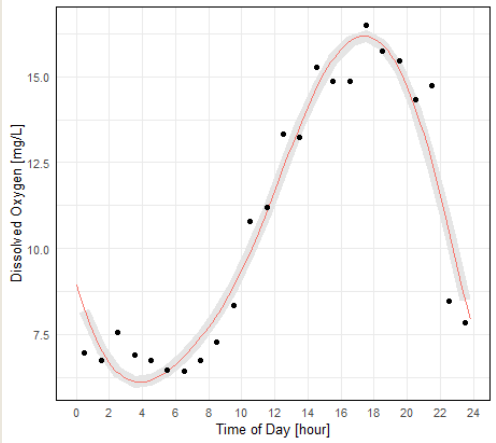
NOAA vertical array
- 2022-2023
- 5 stations
- 440k obs.

DATAFLOW (pilot)
- 2007-2008 (Potomac)
- 576k obs.

From Jon Harcum, Tetra Tech

1. Development and Testing

- We've compiled all data we have (including vertical arrays) into a data set that has been used to decide on the right statistical approaches for each part of the tool.
  - *Daily: Testing of the GAM to get the right equation*
  - *Hourly: Finding cyclic parameters for hourly variability, figuring out how to represent them with depth, season, space*



From Jon Harcum, Tetra Tech

CBP DataHub
- 1984-2022
- 835 stations
- 819k obs.

EOTB: Eyes on the Bay
- 2001-2022
- 126 stations
- 11,916k obs.

VECOS: Virginia Estuarine and Coastal Observing System
- 2003-2022
- 54 stations
- 6,776k obs.

NOAA vertical array
- 2022-2023
- 5 stations
- 440k obs.

DATAFLOW (pilot)
- 2007-2008 (Potomac)
- 576k obs.

1. Development and Testing

- As a test case location for putting all 4 pieces of the tool together (Elgin's work)

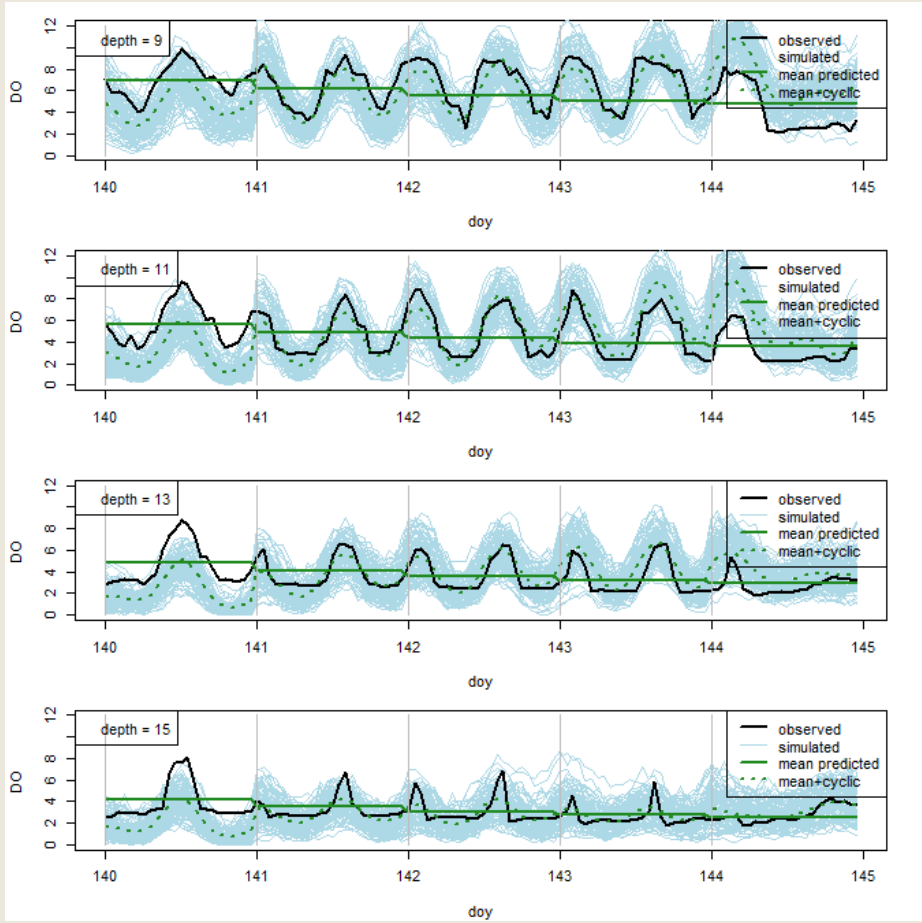
West Gooses Profiler Data Simulation Overview
Presentation to BORG
9/16/2024, Elgin Perry

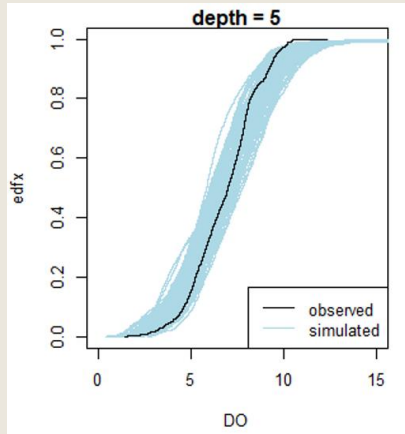
Goals:

Test if the interpolation will mimic within day cycles and hour-to-hour correlation.

Test concepts for simultaneously predicting temporal and spatial correlation.

Test if a cloud of multiple interpolations covers the observed data.

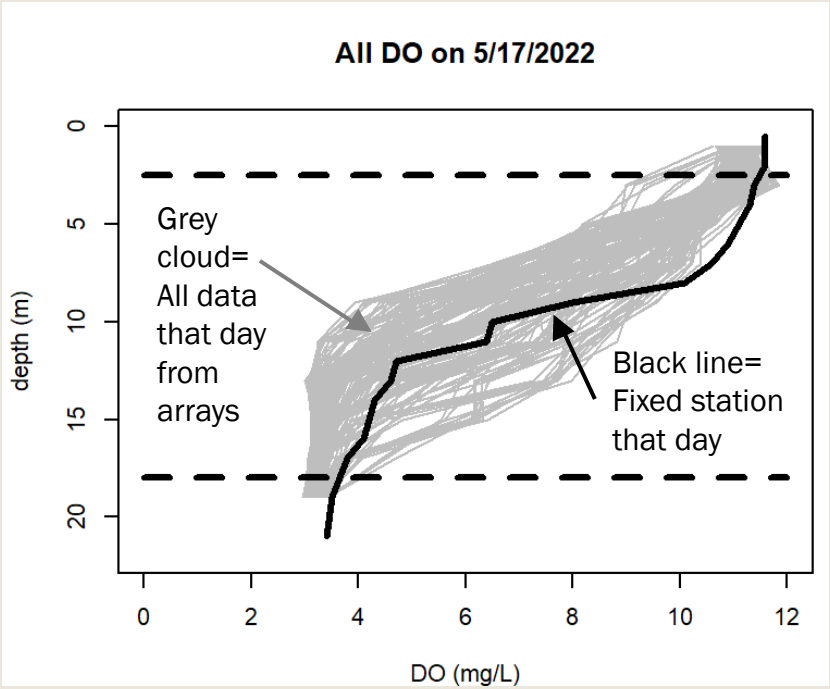




From Elgin Perry, see: BORG Sept meeting

1. Development and Testing

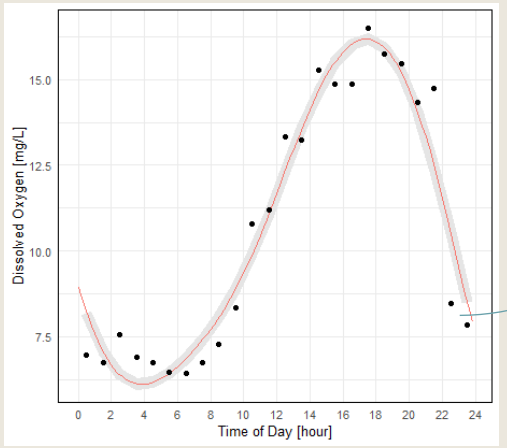
- Pycnocline movement at high frequency: TBD
  - *Will use the salinity and temperature data collected*

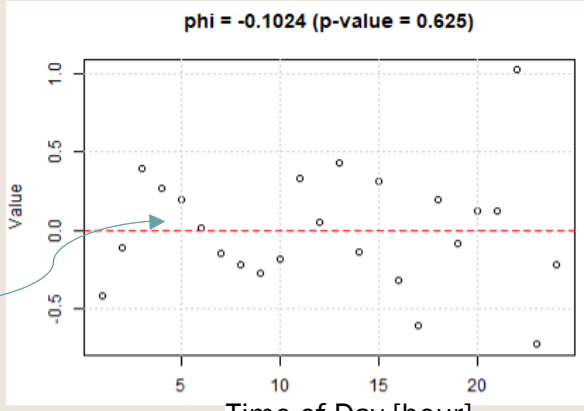


2. Parameterization

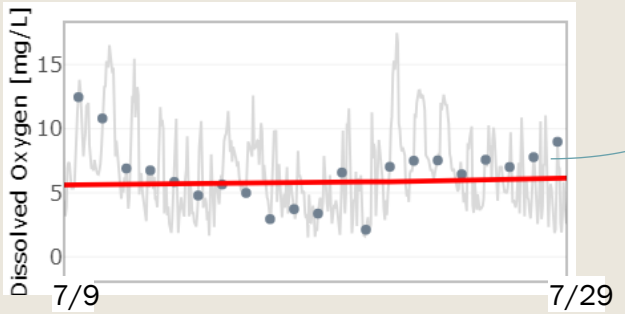
- Correlation components of the model are being parameterized with available data. Vertical array data is very useful for this:

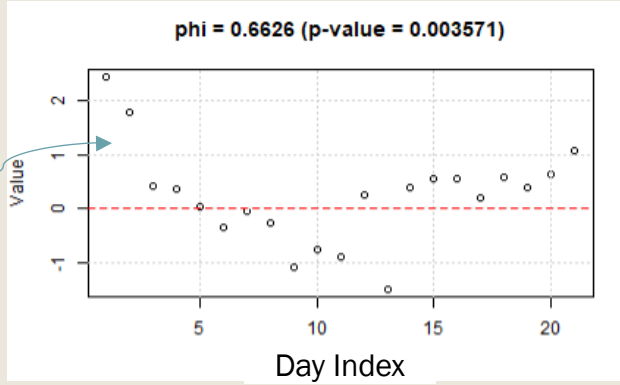
Hour-by-hour autocorrelation





Day-to-day autocorrelation





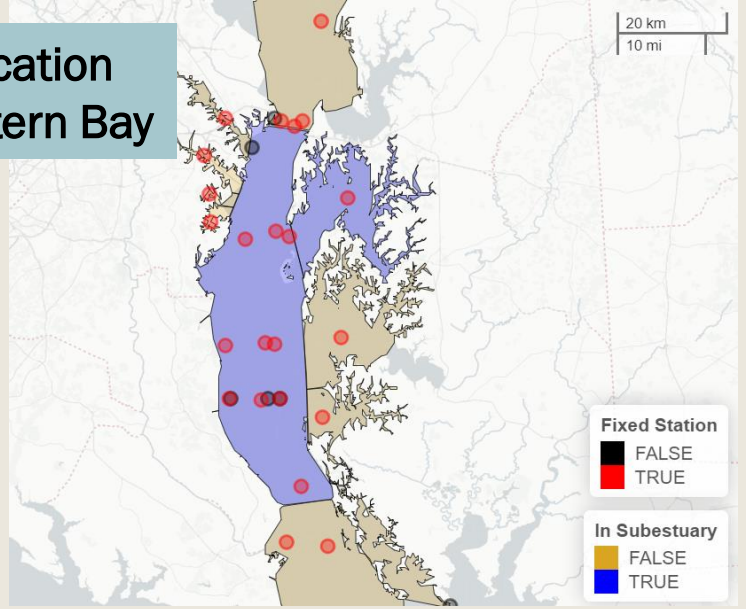
From Jon Harcum, see: BORG Sept meeting

And depth-to-depth residual correlation

3. Implementation

- Any segment with vertical array data would have that data included in its interpolation for criteria assessment analysis in a particular year.

Example Application
CB4MH + Eastern Bay



Sample Count (2022, individual days with data)

Observation count by station and depth

station	in_subestuary	fixed	reg	cbseg_92	num_rec_tot	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
CB4.1C	TRUE	TRUE	CB4eas	CB4MH	371	15	15	15	15	7	15	10	15	12	15	13	15	13	15	13	15	12	15	8	15	5	15	1	15	1	15	1	15		15	2	13	5			
CB4.4	TRUE	TRUE	CB4eas	CB4MH	362	15	15	15	15	5	15	10	15	12	15	12	15	12	15	11	15	10	15	8	15	4	15	1	15	1	15	1	15		15	2	13	5			
CB4.2C	TRUE	TRUE	CB4eas	CB4MH	331	15	15	15	15	4	15	8	15	10	15	11	15	11	15	11	15	11	15	11	15	8	15	3	15	2	15	7	8	1							
CB4.3C	TRUE	TRUE	CB4eas	CB4MH	316	15	15	15	15	6	15	10	15	10	15	11	15	11	15	12	15	10	15	8	15	4	15	2	15	1	15	6									
CB4.1E	TRUE	TRUE	CB4eas	CB4MH	207	11	10	10	10	5	10	8	10	9	10	9	10	9	10	8	10	8	10	6	10	2	9		9	4											
CB4.3E	TRUE	TRUE	CB4eas	CB4MH	173	9	9	9	9	2	9	6	9	7	9	8	9	9	9	9	9	8	9	4	9	4	8														
EE1.1	TRUE	TRUE	CB4eas	EASMH	156	12	12	12	12	12	12	12	12	12	12	13	11	12																							
CB4.1W	TRUE	TRUE	CB4eas	CB4MH	91	10	10	10	10	10	10	10	10	9	2																										
CB4.3W	TRUE	TRUE	CB4eas	CB4MH	90	9	9	9	9	9	9	9	9	9	9																										
CB4.2E	TRUE	TRUE	CB4eas	CB4MH	84	9	9	9	9	9	9	9	9	9	3																										
CB4.2W	TRUE	TRUE	CB4eas	CB4MH	83	9	9	9	9	9	9	9	9	9	2																										
east-gooses	TRUE	FALSE	CB4eas	CB4MH	905		78		75		100		100		100		100		97		78		77		100															
west-gooses	TRUE	FALSE	CB4eas	CB4MH	467		131		76		130		130																											
XEF3551	TRUE	FALSE	CB4eas	CB4MH	98	8	6	8	8	8	8	8	8	8	8	8	12																							
XGF7832	TRUE	FALSE	CB4eas	CB4MH	54	8	6	8	8	8	8	8																												

Summary

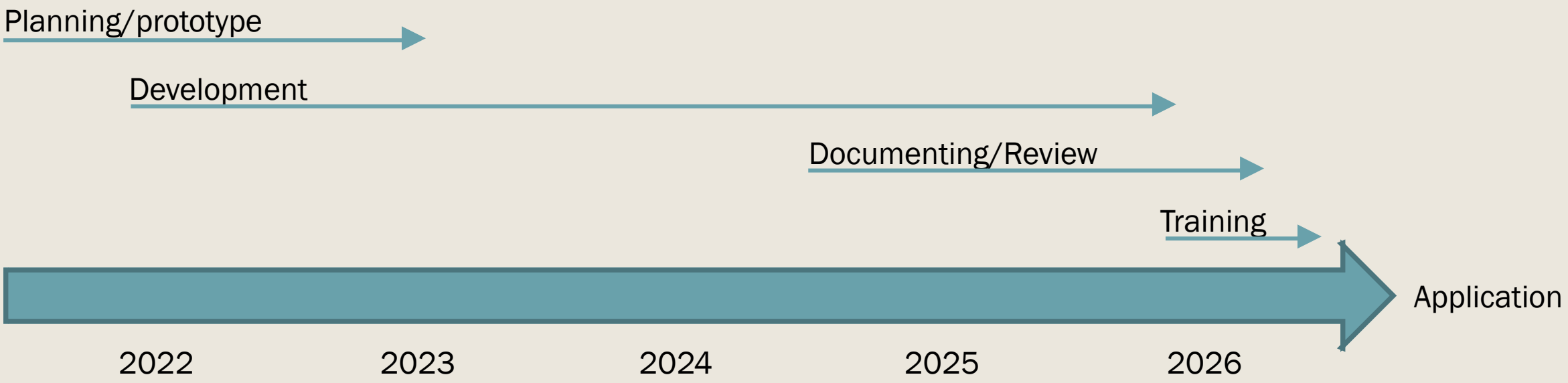
- Vertical array data has already been used in development and testing for all parts of the 4-D tool.

- Parameterization is underway, and current data collection will very likely be used to inform the correlation part of the interpolation going forward.

- Vertical array data will be important for interpolation each year, especially at deep depths where we have little other high frequency data.

extras

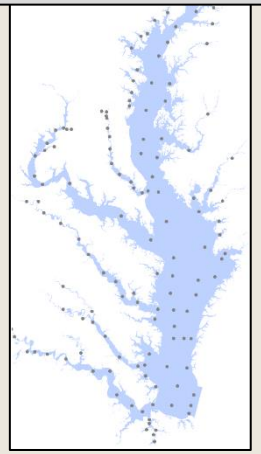
4d interpolator development timeline



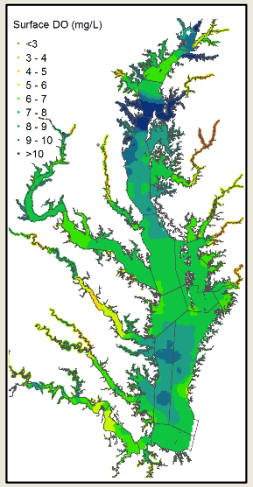
WQS Criterion Assessment

Current approach

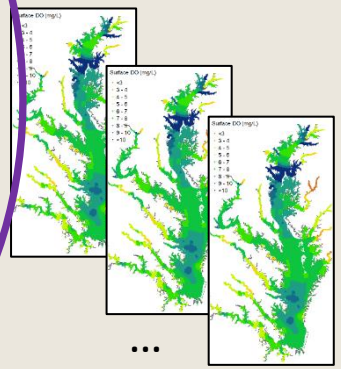
1. Collect bi-weekly to monthly data at each monitoring station



2. Interpolate bay-wide for a single cruise

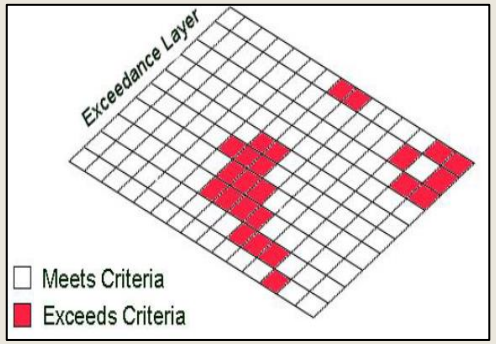


3. Conduct interpolations for a season over the 3-year period

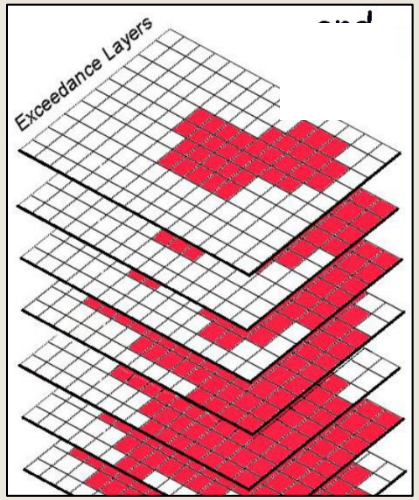


...

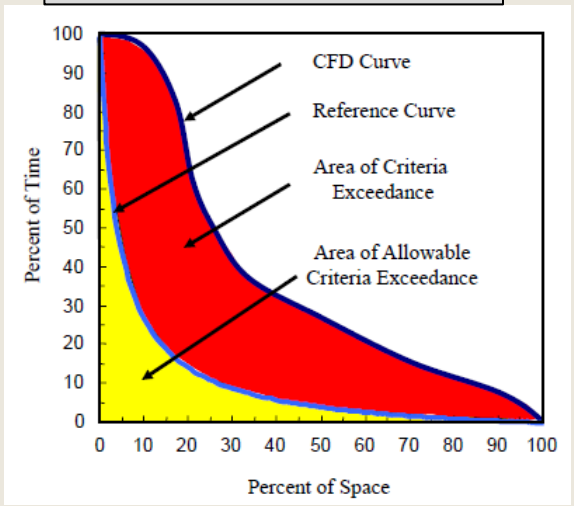
4. Average monthly if 30-day criterion, and compare the interpolated values with appropriate criterion values



5. Repeat Step 4 throughout the 3-year period



6. Calculate the cumulative frequency distribution (CFD) for extent of exceedance in space and time

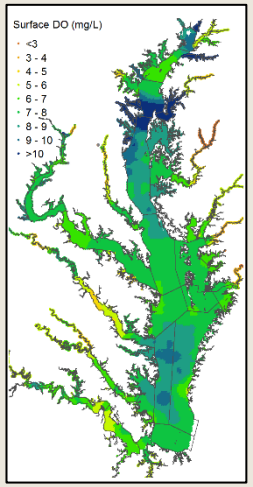


7. Quantify attainment status (attaining or not) for the 3-year period

From Peter Tango

Current interpolation

2. Interpolate bay-wide for a single cruise



Problems with current interpolation

- Does not use the high frequency data (except the calibration data).
- Vertical layers interpolated horizontally and stacked;
- One cruise at a time, meaning a 2-week period assumed static; and
- Not statistical.

This NEW interpolation will:

- Use **ALL high frequency data** (ConMon and vertical array)
- Interpolate all data together, not in layers.
- Interpolate in time, so that we do not have to artificially split time periods.
- Statistical – allowing for uncertainty bounds if needed.