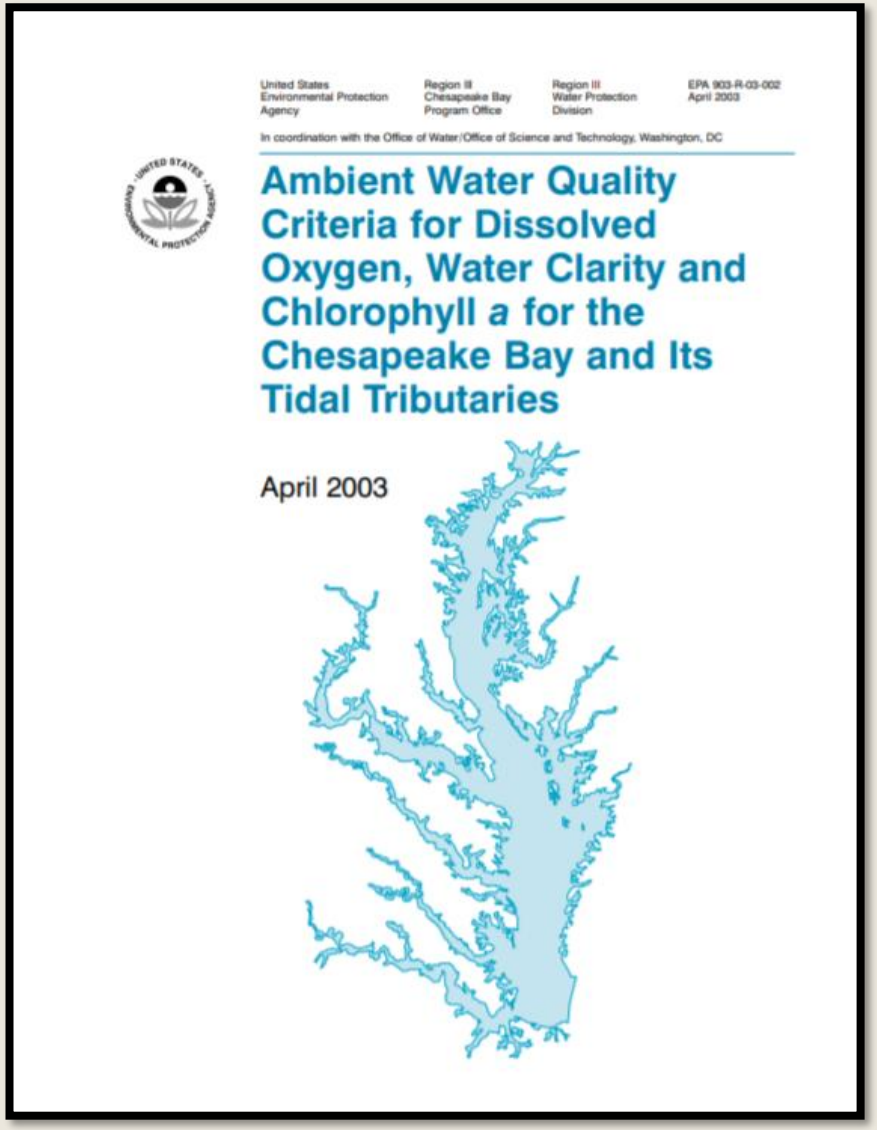
4-dimensional (4-D) interpolator development overview

Bay Oxygen Research Large Group
Feb. 10, 2025

Rebecca Murphy (UMCES/CBP)
Jon Harcum (Tetra Tech), Elgin Perry (statistics consultant),
Breck Sullivan (USGS), and Peter Tango (USGS)

Purpose: New interpolation method for criteria assessment



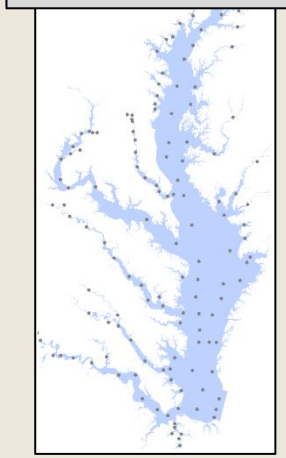
- "…The use of cumulative frequency distributions (CFDs) is recommended for assessing the spatial and temporal water quality criteria exceedances in the Chesapeake Bay".

- Some notes in 2003 about this:
  - *OTHER current criteria assessment methods are based only on temporal variation with measurements evaluated at individual monitoring stations.*
  - *Limitation : it is difficult to determine whether an individual sampling location is representative, and there is potential for bias.*
  - *In size of CB, accounting for spatial variation can be very important and CFD approach represents a significant improvement.*

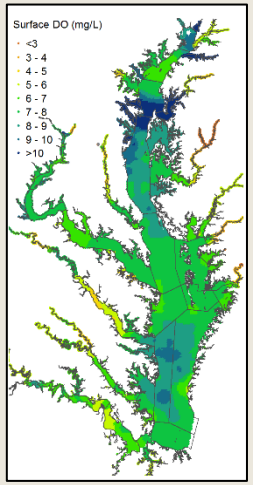
WQS Criterion Assessment

Current approach

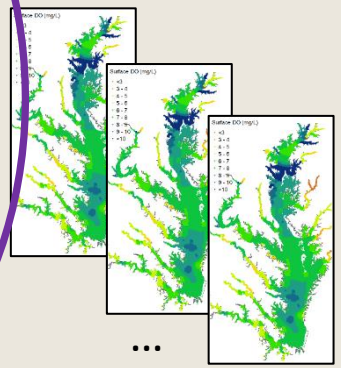
1. Collect bi-weekly to monthly data at each monitoring station



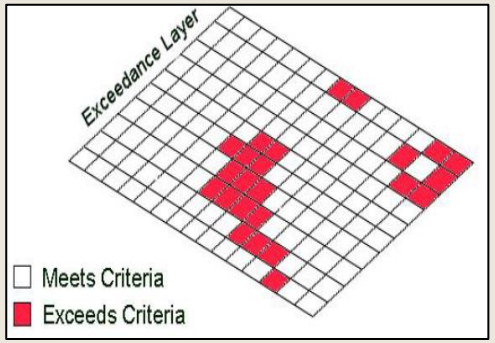
2. Interpolate bay-wide for a single cruise



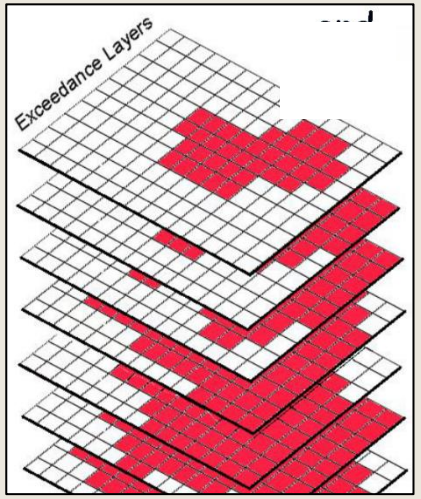
3. Conduct interpolations for a season over the 3-year period



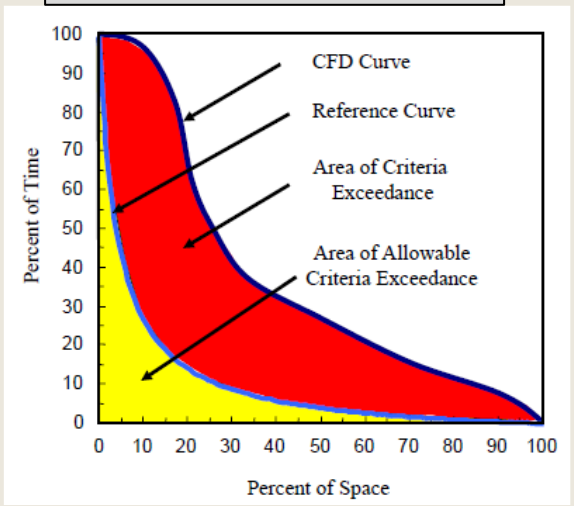
4. Average monthly if 30-day criterion, and compare the interpolated values with appropriate criterion values



5. Repeat Step 4 throughout the 3-year period



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

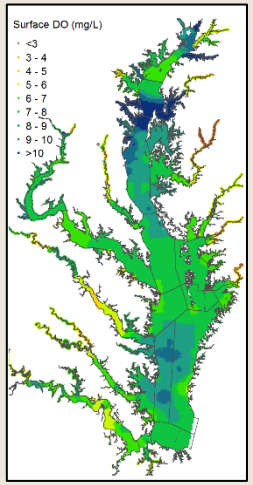


7. Quantify attainment status (attaining or not) for the 3-year period

From Peter Tango

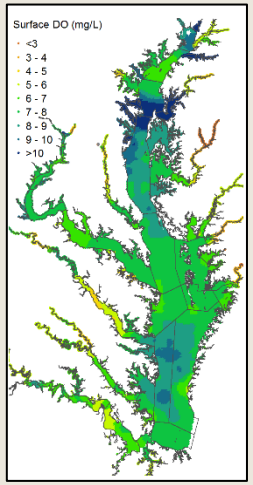
Current interpolation

2. Interpolate bay-wide for a single cruise



Current interpolation

2. Interpolate bay-wide for a single cruise



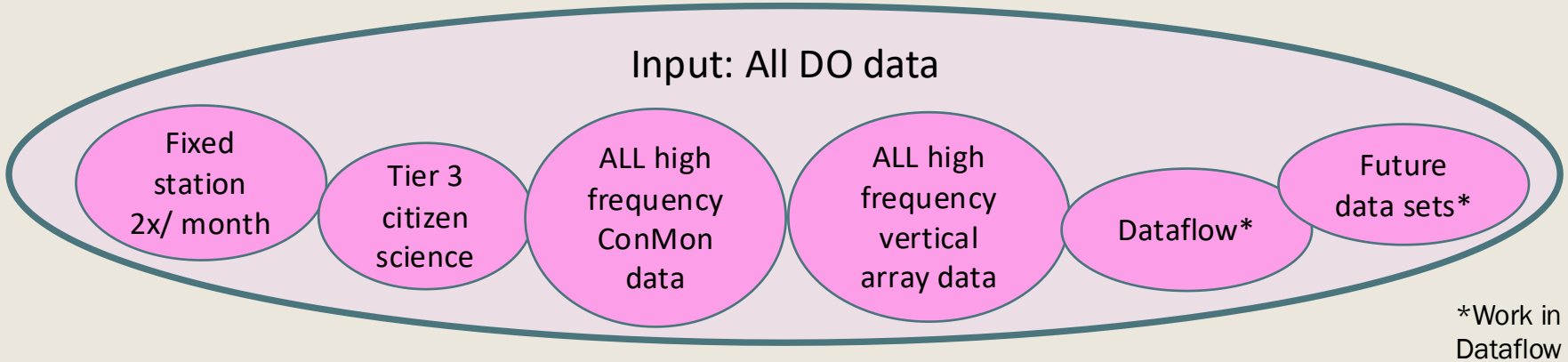
Problems with current interpolation

- 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:

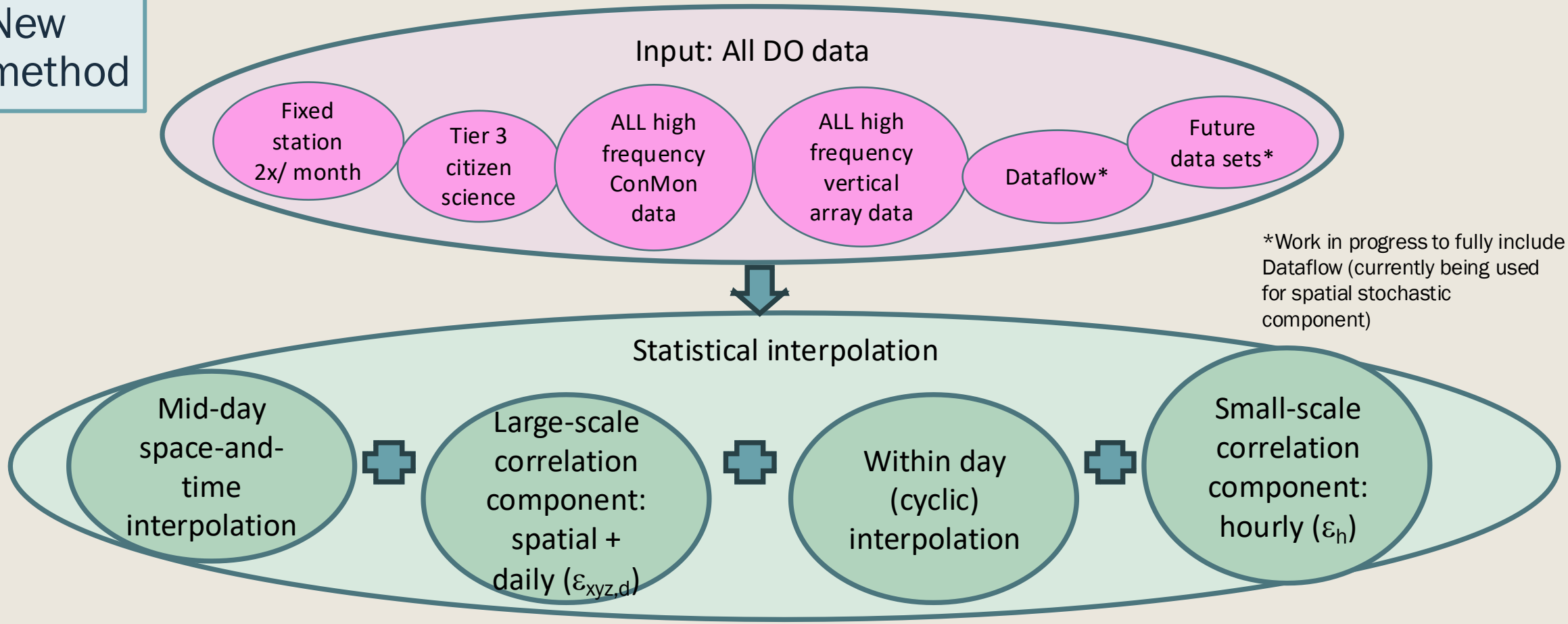
- Use **ALL high frequency data** (ConMon and vertical array)
- Interpolate all data together, not in layers.
- Interpolate in time, so that we do not have to artificially split time periods.
- Statistical – allowing for uncertainty bounds if needed.

New method



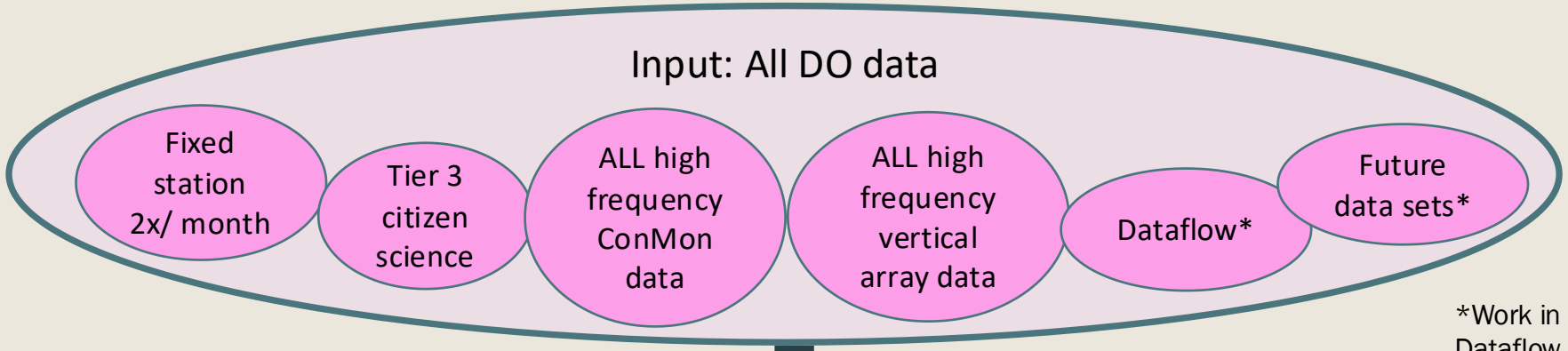
*Work in progress to fully include Dataflow (currently being used for spatial stochastic component)

New method

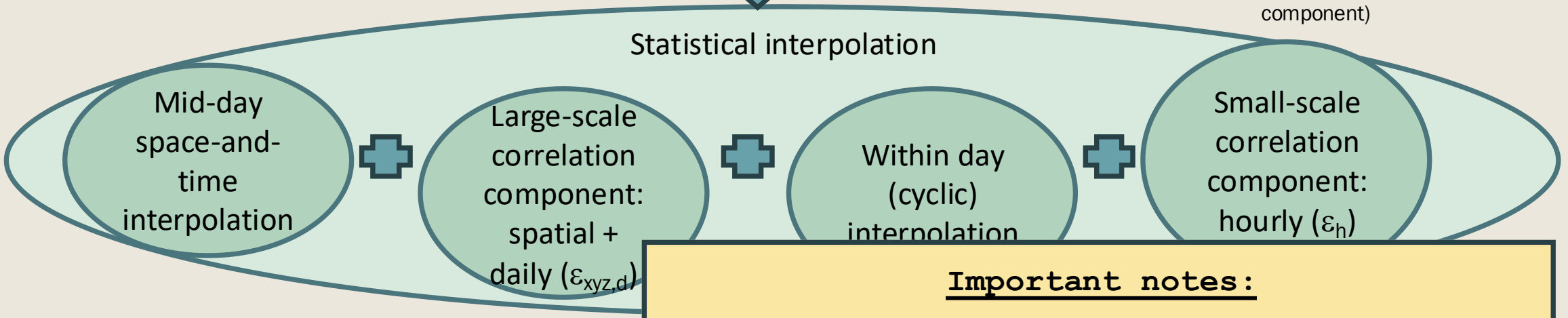


*Work in progress to fully include Dataflow (currently being used for spatial stochastic component)

New method



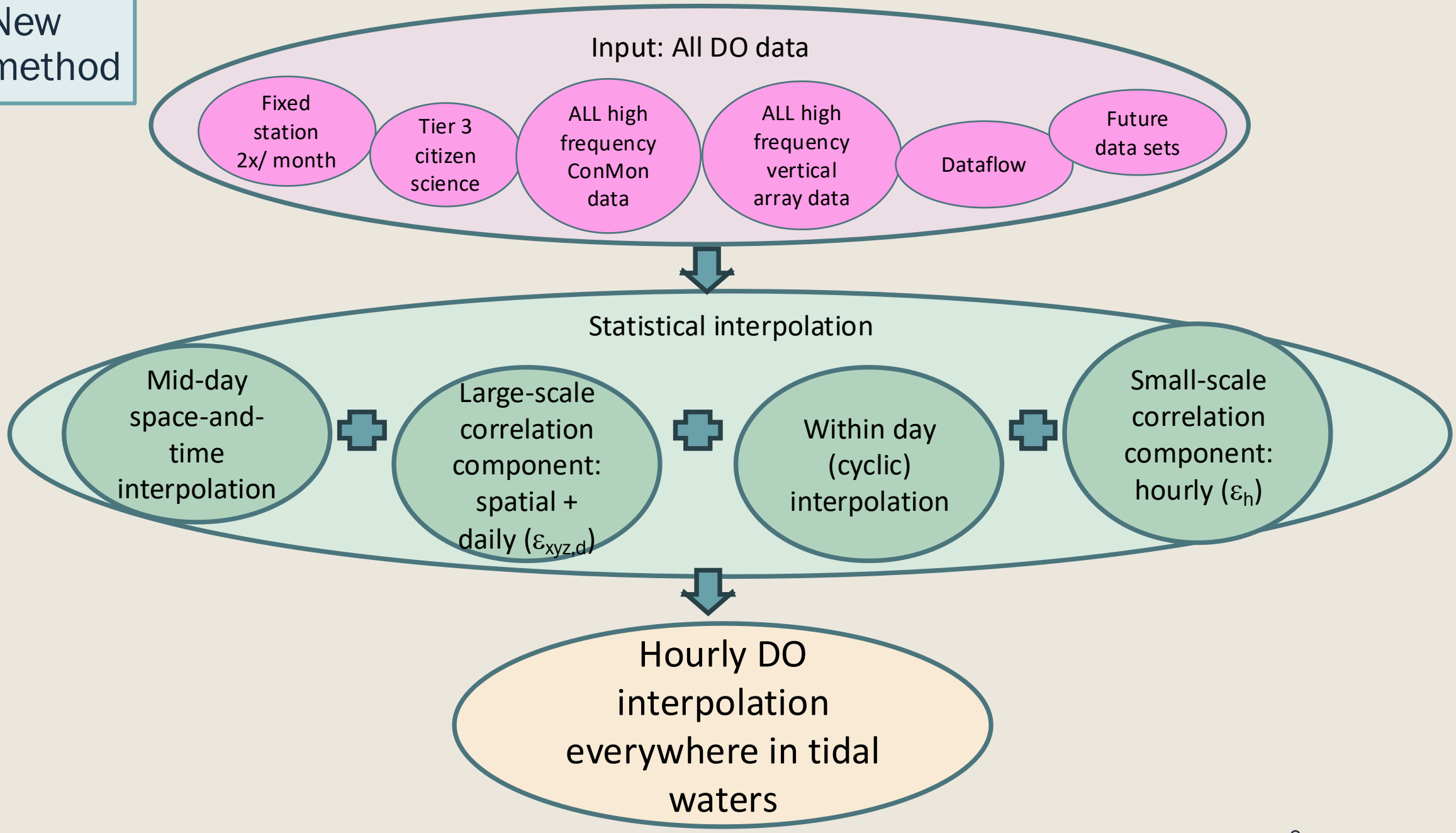
*Work in progress to fully include Dataflow (currently being used for spatial stochastic component)



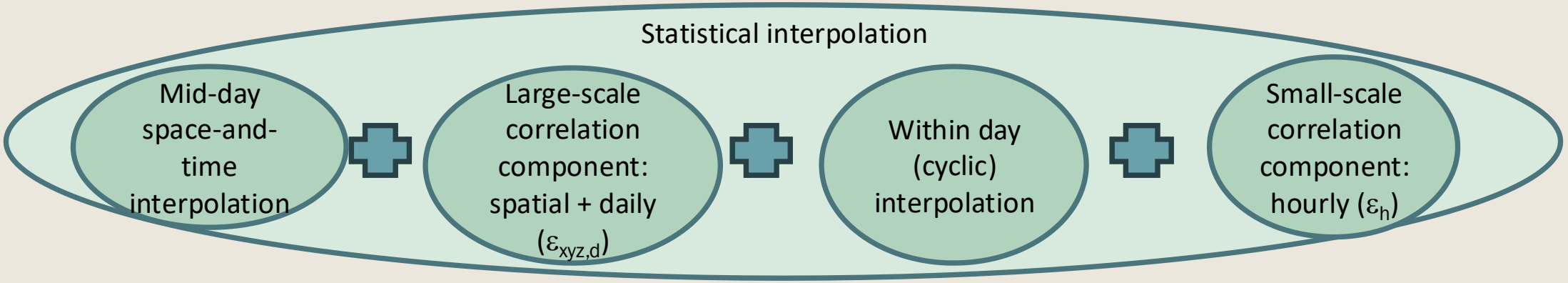
Important notes:

- **The 4 interpolation pieces listed here were designed to represent different spatial and temporal features of the DO data.**
- **Ultimately, it is the SUM of these 4 components that will be used, NOT any incremental results.**

New method

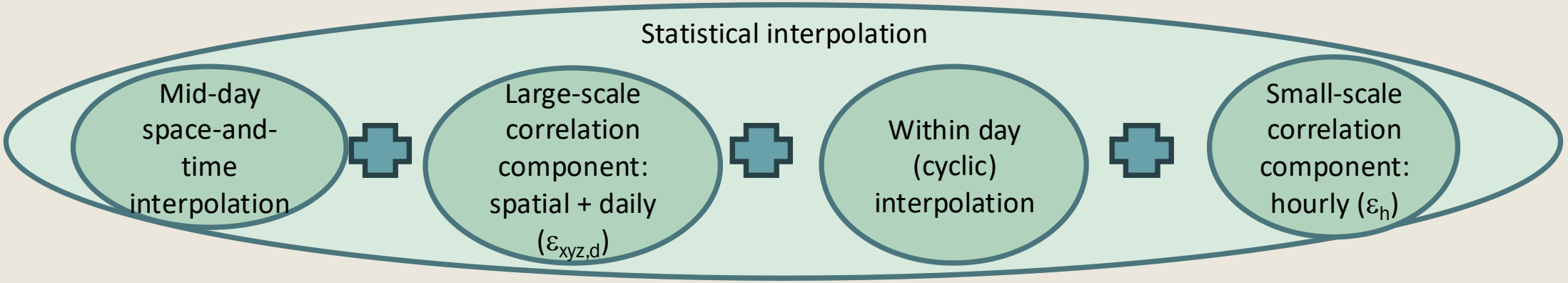

Input: All DO data
Fixed station 2x/ month
Tier 3 citizen science
ALL high frequency ConMon data
ALL high frequency vertical array data
Dataflow
Future data sets
Statistical interpolation
Mid-day space-and-time interpolation
Large-scale correlation component: spatial + daily ($\varepsilon_{xyz,d}$)
Within day (cyclic) interpolation
Small-scale correlation component: hourly (ε_h)
Hourly DO interpolation everywhere in tidal waters

Today: Piecing it all together



- Progress has been made on all components
  - *Necessary parameterizing and testing will continue in 2025 so that the simulations interpolate the data appropriately.*

Today: Piecing it all together



- Progress has been made on all components
  - *Necessary parameterizing and testing will continue in 2025 so that the simulations interpolate the data appropriately.*

- Jon will show the combined 4-D approach applied to one segment and year
  - *Huge success that it worked!*
  - *Tests adjusting various "knobs" show expected results.*
  - *Comparison to high frequency East Goose vertical array DO in 2022.*
  - *Keep in mind, this is DRAFT and not expected to be the best match to the data.*

4-D interpolator development timeline

This year:

- Still developing
- More documentation
- Linking to criteria assessment

Current and next development steps

- **Criteria assessment link:** Work with CAP team on how results would be used in criteria assessment (starting now)
  - *Settle on method for interpolating pycnocline (and build in code).*
  - *Assess if any changes are needed to prediction grid.*
  - *Brainstorm options for using 4D results in CFD or other approaches.*
  - *Conduct test cases of various options.*
- **Method development:** Continue necessary parametrization of all 4 parts of the 4-D tool (continuing through entire 2025)
  - *Continue work on cyclic interpolation.*
  - *Parameterize vertical correlation.*
  - *Expand Dataflow analysis beyond Potomac for horizontal correlation.*
  - *Scale up simulation to bay-wide, examine results, and validate with additional or hold-out data.*
  - *Consider non-stationarity in correlation parameters.*
- **Software development** (continuing through entire 2025 into 2026)
  - *Continue data compilation and use of all data sets.*
  - *Determine file management and storage options.*
  - *Work with future users on features.*