A space-time interpolation tool for Chesapeake Bay dissolved oxygen: Development and preliminary results

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

[1]University of Maryland Center for Environmental Science at the Chesapeake Bay Program (CBP),
[2]Tetra Tech,[3]Statistics Consultant, [4]USGS at the CBP,

CERF

Nov. 12, 2025

Purpose: Build a tool for more complete criteria assessment

Table 1. Chesapeake Bay dissolved oxygen criteria.

Designated Use	Criteria Concentration/Duration	Protection Provided	Temporal Application
Migratory fish spawning and nursery use	7-day mean $\geq$ 6 mg liter^{-1} (tidal habitats with 0-0.5 ppt salinity)	Survival/growth of larval/juvenile tidal-fresh resident fish; protective of threatened/endangered species.	February 1 - May 31
	Instantaneous minimum $\geq$ 5 mg liter^{-1}	Survival and growth of larval/juvenile migratory fish; protective of threatened/endangered species.	
	Open-water fish and shellfish designated use criteria apply		June 1 - January 31
Shallow-water bay grass use	Open-water fish and shellfish designated use criteria apply		Year-round
Open-water fish and shellfish use	30-day mean $\geq$ 5.5 mg liter^{-1} (tidal habitats with 0-0.5 ppt salinity)	Growth of tidal-fresh juvenile and adult fish; protective of threatened/endangered species.	Year-round
	30-day mean $\geq$ 5 mg liter^{-1} (tidal habitats with >0.5 ppt salinity)	Growth of larval, juvenile and adult fish and shellfish; protective of threatened/endangered species.	
	7-day mean $\geq$ 4 mg liter^{-1}	Survival of open-water fish larvae.	
	Instantaneous minimum $\geq$ 3.2 mg liter^{-1}	Survival of threatened/endangered sturgeon species.[1]	
Deep-water seasonal fish and shellfish use	30-day mean $\geq$ 3 mg liter^{-1}	Survival and recruitment of bay anchovy eggs and larvae.	June 1 - September 30
	1-day mean $\geq$ 2.3 mg liter^{-1}	Survival of open-water juvenile and adult fish.	
	Instantaneous minimum $\geq$ 1.7 mg liter^{-1}	Survival of bay anchovy eggs and larvae.	
	Open-water fish and shellfish designated-use criteria apply		October 1 - May 31
Deep-channel seasonal refuge use	Instantaneous minimum $\geq$ 1 mg liter^{-1}	Survival of bottom-dwelling worms and clams.	June 1 - September 30
	Open-water fish and shellfish designated use criteria apply		October 1 - May 31

[1] At temperatures considered stressful to shortnose sturgeon (>29°C), dissolved oxygen concentrations above an instantaneous minimum of 4.3 mg liter^{-1} will protect survival of this listed sturgeon species.

From EPA 2003 Ambient Water Quality Criteria

Purpose: Build a tool for more complete criteria assessment

DO criteria that currently can be evaluated with existing approaches and data

Table 1. Chesapeake Bay dissolved oxygen criteria.

Designated Use	Criteria Concentration/Duration	Protection Provided	Temporal Application
Migratory fish spawning and nursery use *	7-day mean ≥ 6 mg liter^{-1} (tidal habitats with 0-0.5 ppt salinity)	Survival/growth of larval/juvenile tidal-fresh resident fish; protective of threatened/endangered species.	February 1 - May 31
	Instantaneous minimum ≥ 5 mg liter^{-1}	Survival and growth of larval/juvenile migratory fish; protective of threatened/endangered species.	
	Open-water fish and shellfish designated use criteria apply		June 1 - January 31
Shallow-water bay grass use	Open-water fish and shellfish designated use criteria apply		Year-round
Open-water fish and shellfish use	30-day mean ≥ 5.5 mg liter^{-1} (tidal habitats with 0-0.5 ppt salinity)	Growth of tidal-fresh juvenile and adult fish; protective of threatened/endangered species.	Year-round
	30-day mean ≥ 5 mg liter^{-1} (tidal habitats with >0.5 ppt salinity)	Growth of larval, juvenile and adult fish and shellfish; protective of threatened/endangered species.	
	7-day mean ≥ 4 mg liter^{-1}	Survival of open-water fish larvae.	
	Instantaneous minimum ≥ 3.2 mg liter^{-1}	Survival of threatened/endangered sturgeon species.[1]	
Deep-water seasonal fish and shellfish use	30-day mean ≥ 3 mg liter^{-1}	Survival and recruitment of bay anchovy eggs and larvae.	June 1 - September 30
	1-day mean ≥ 2.3 mg liter^{-1}	Survival of open-water juvenile and adult fish.	
	Instantaneous minimum ≥ 1.7 mg liter^{-1}	Survival of bay anchovy eggs and larvae.	
	Open-water fish and shellfish designated-use criteria apply		October 1 - May 31
Deep-channel seasonal refuge use	Instantaneous minimum ≥ 1 mg liter^{-1}	Survival of bottom-dwelling worms and clams.	June 1 - September 30
	Open-water fish and shellfish designated use criteria apply		October 1 - May 31

[1] At temperatures considered stressful to shortnose sturgeon (>29°C), dissolved oxygen concentrations above an instantaneous minimum of 4.3 mg liter^{-1} will protect survival of this listed sturgeon species.

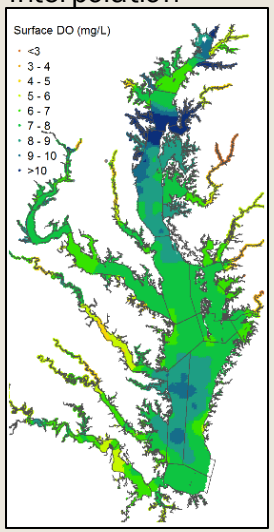
*Note a 30-day mean 6 mg/L MSN value is evaluated for purpose of the WQ indicator.

From EPA 2003 Ambient Water Quality Criteria

Current interpolation:

Inverse distance weighting

Example spatial interpolation



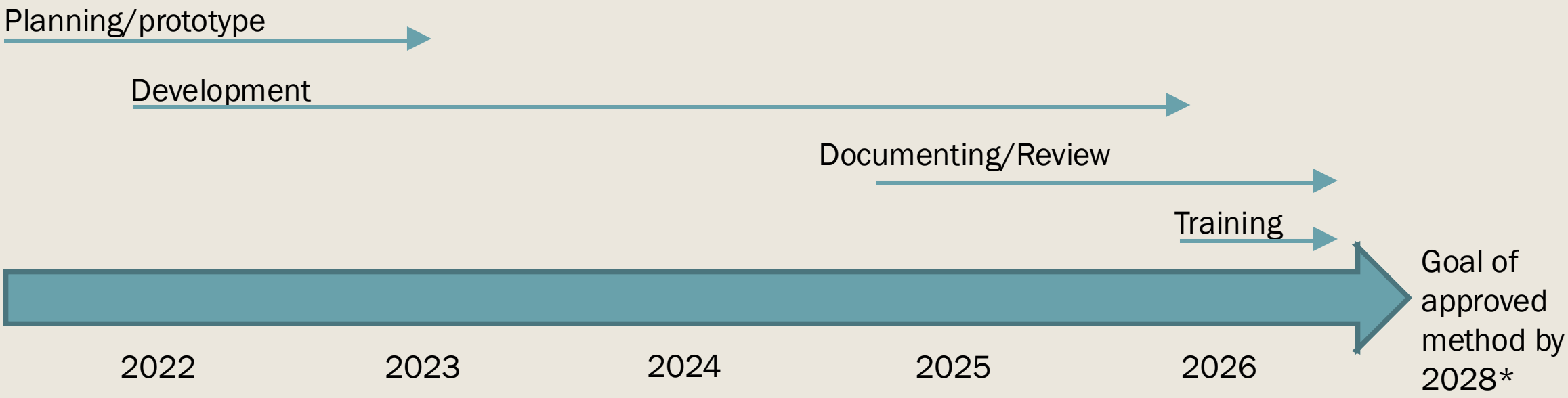
Problems with current interpolation

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

This NEW "4-D" interpolation will:

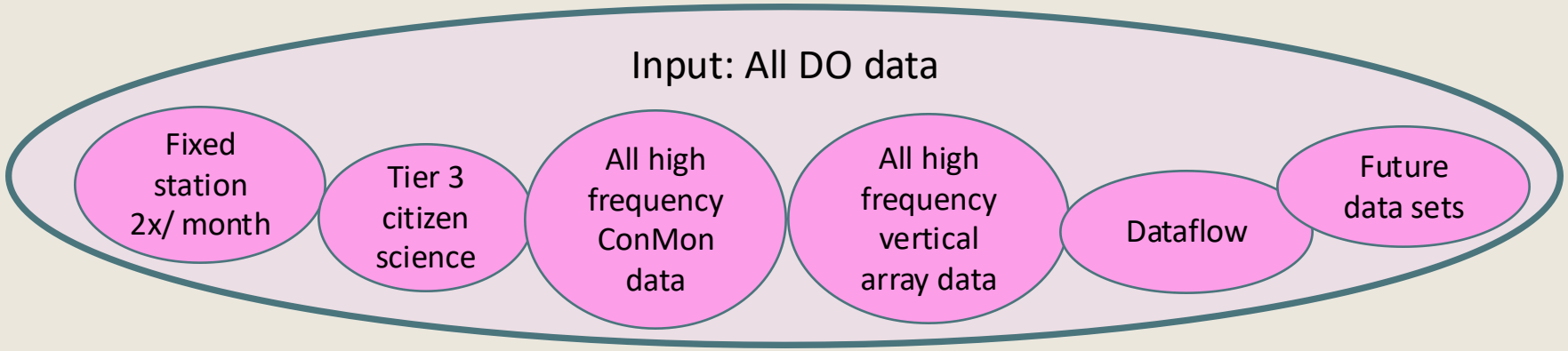
- Use all high frequency data
- Interpolate all data together, not in layers.
- Interpolate in time, so that we do not have to artificially split time periods.
- Statistical – allowing for multiple realizations of interpolation results.

4-D interpolator development timeline

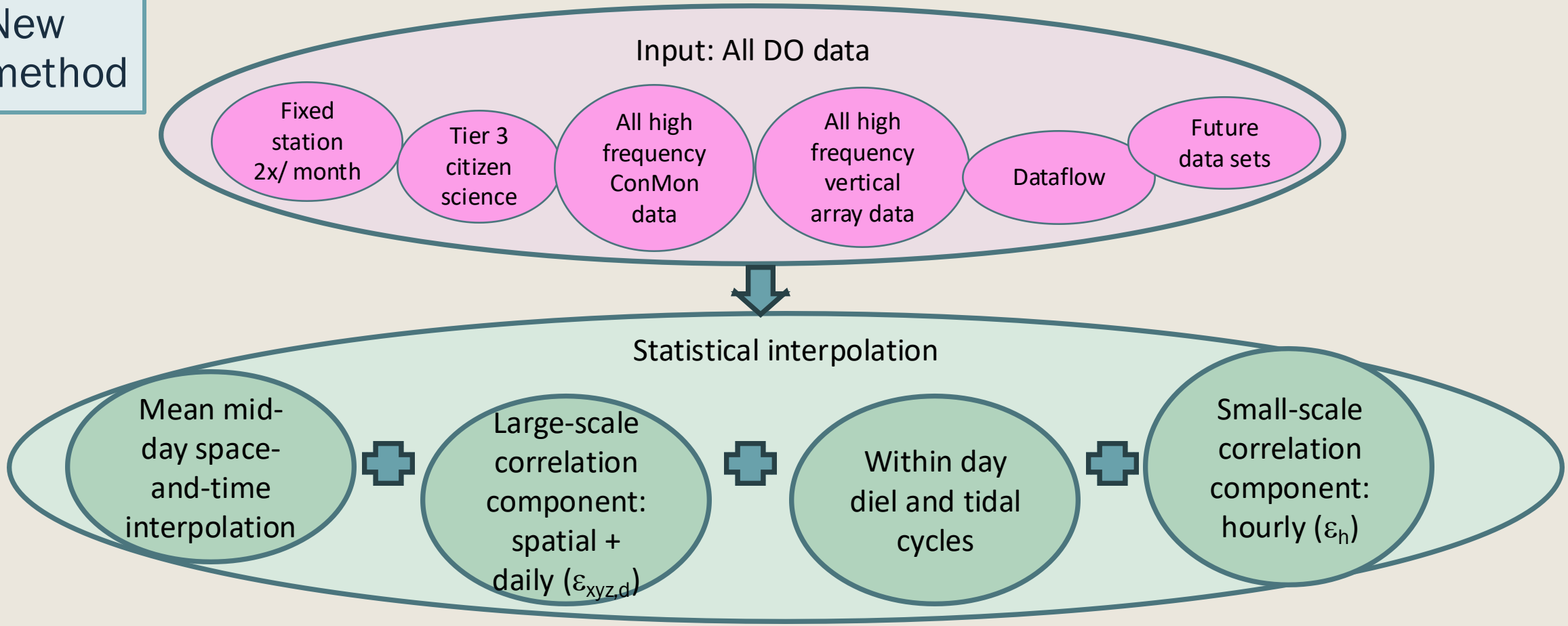


*with 2030 goal of reporting on all criteria

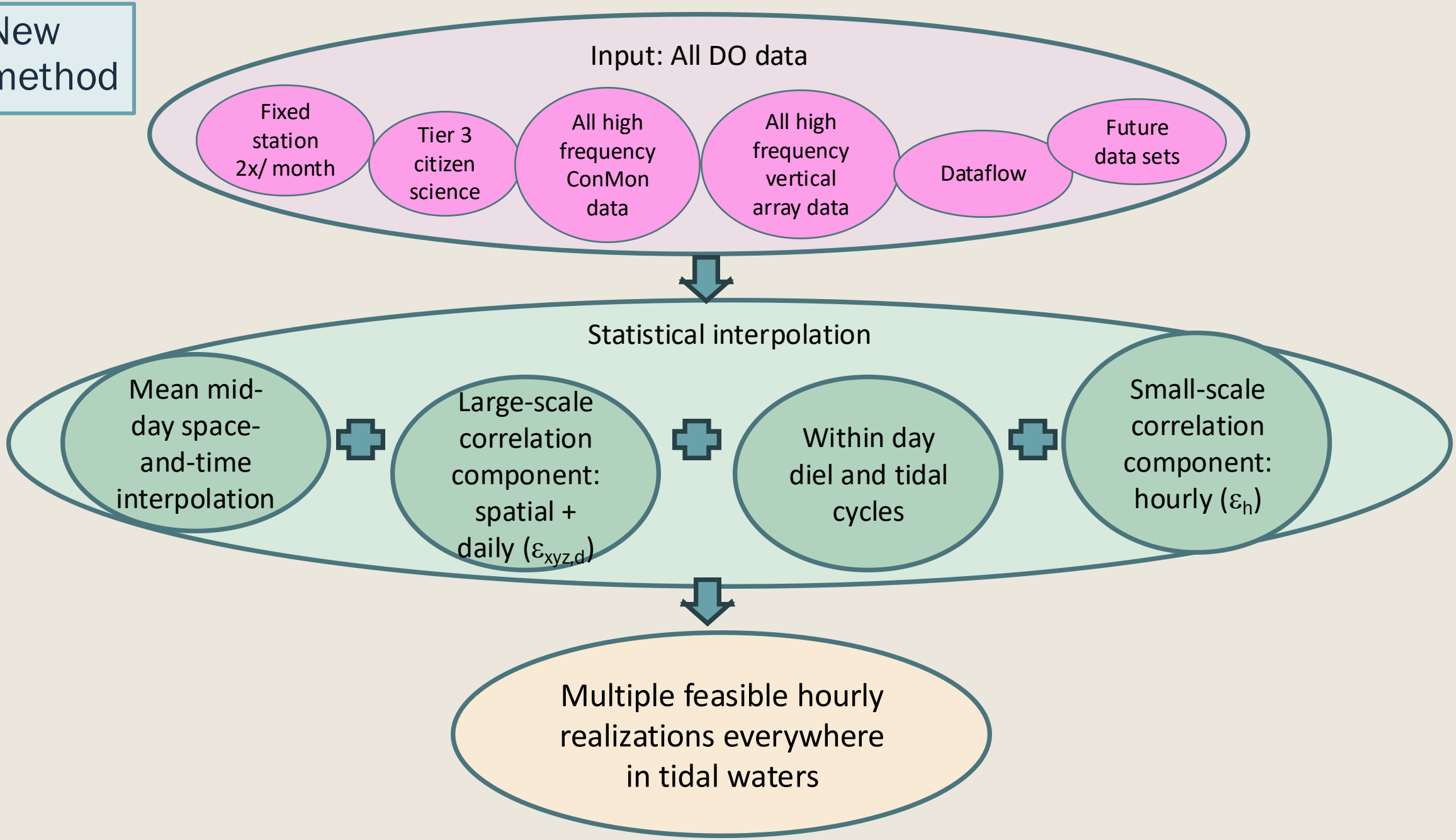
New method

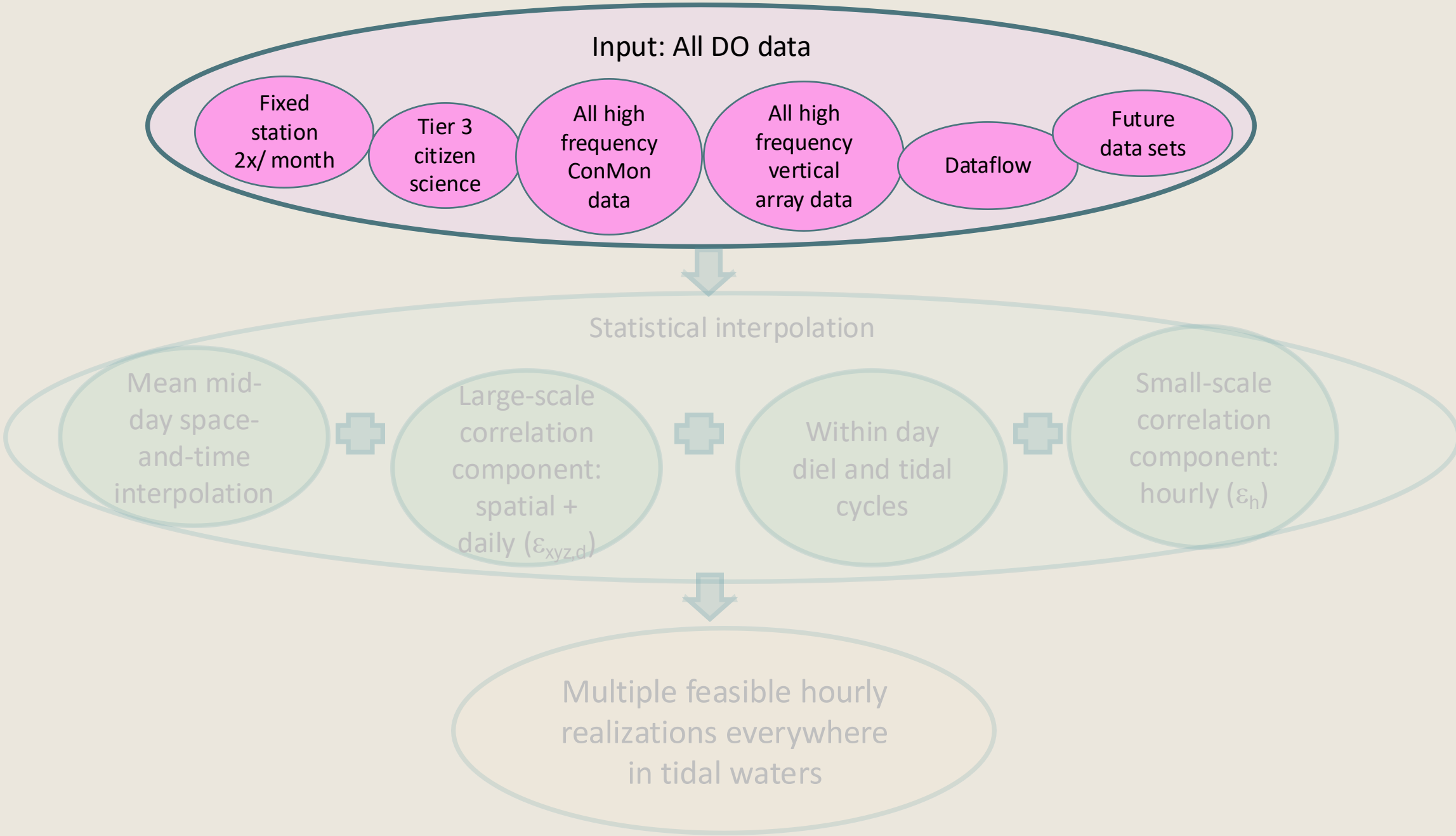

Input: All DO data
Fixed station 2x/ month
Tier 3 citizen science
All high frequency ConMon data
All high frequency vertical array data
Dataflow
Future data sets

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
Mean mid-day space-and-time interpolation
Large-scale correlation component: spatial + daily ($\varepsilon_{xyz,d}$)
Within day diel and tidal cycles
Small-scale correlation component: hourly (ε_h)

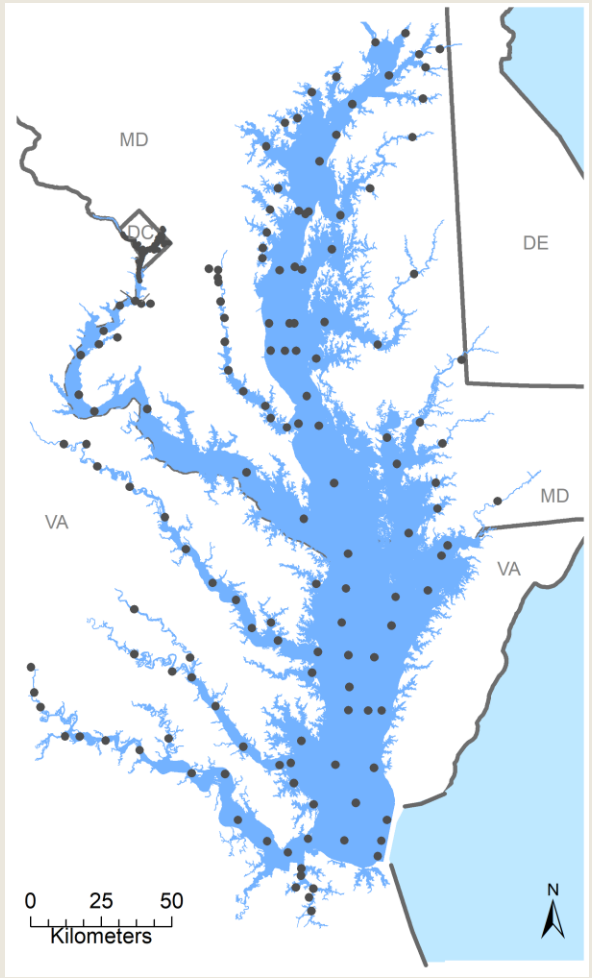
New method


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
Mean mid-day space-and-time interpolation
Large-scale correlation component: spatial + daily ($\varepsilon_{xyz,d}$)
Within day diel and tidal cycles
Small-scale correlation component: hourly (ε_h)
Multiple feasible hourly realizations everywhere in tidal waters


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
Mean mid-day space-and-time interpolation
Large-scale correlation component: spatial + daily ($\varepsilon_{xyz,d}$)
Within day diel and tidal cycles
Small-scale correlation component: hourly (ε_h)
Multiple feasible hourly realizations everywhere in tidal waters

Multiple types of data being used

Bi-weekly long-term sampling (DOEE, MDDNR, VADEQ, CBP)

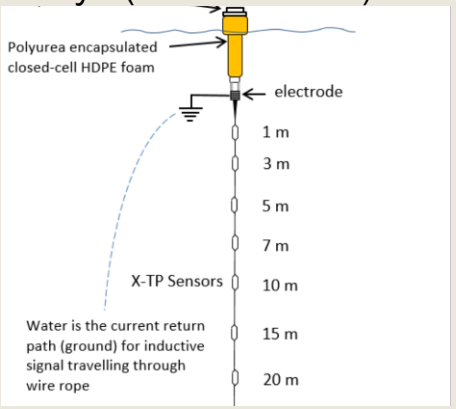


Shallow water continuous monitoring (MDDNR and VECOS)

From http://vecos.vims.edu/

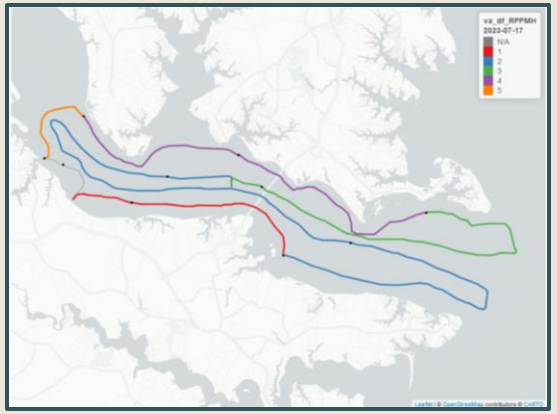
From https://eyesonthebay.dnr.maryland.gov/

New continuous vertical arrays (NOAA & CBP)



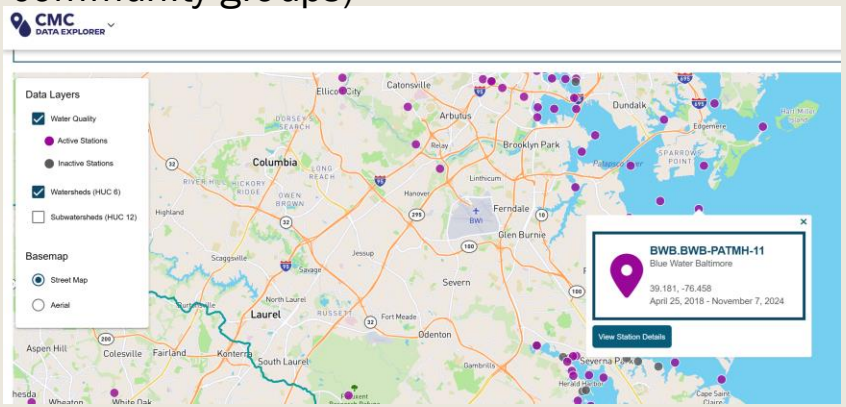
https://www.chesapeakebay.net/who/group/hypoxia-collaborative-team

Dataflow (MDDNR and VECOS)

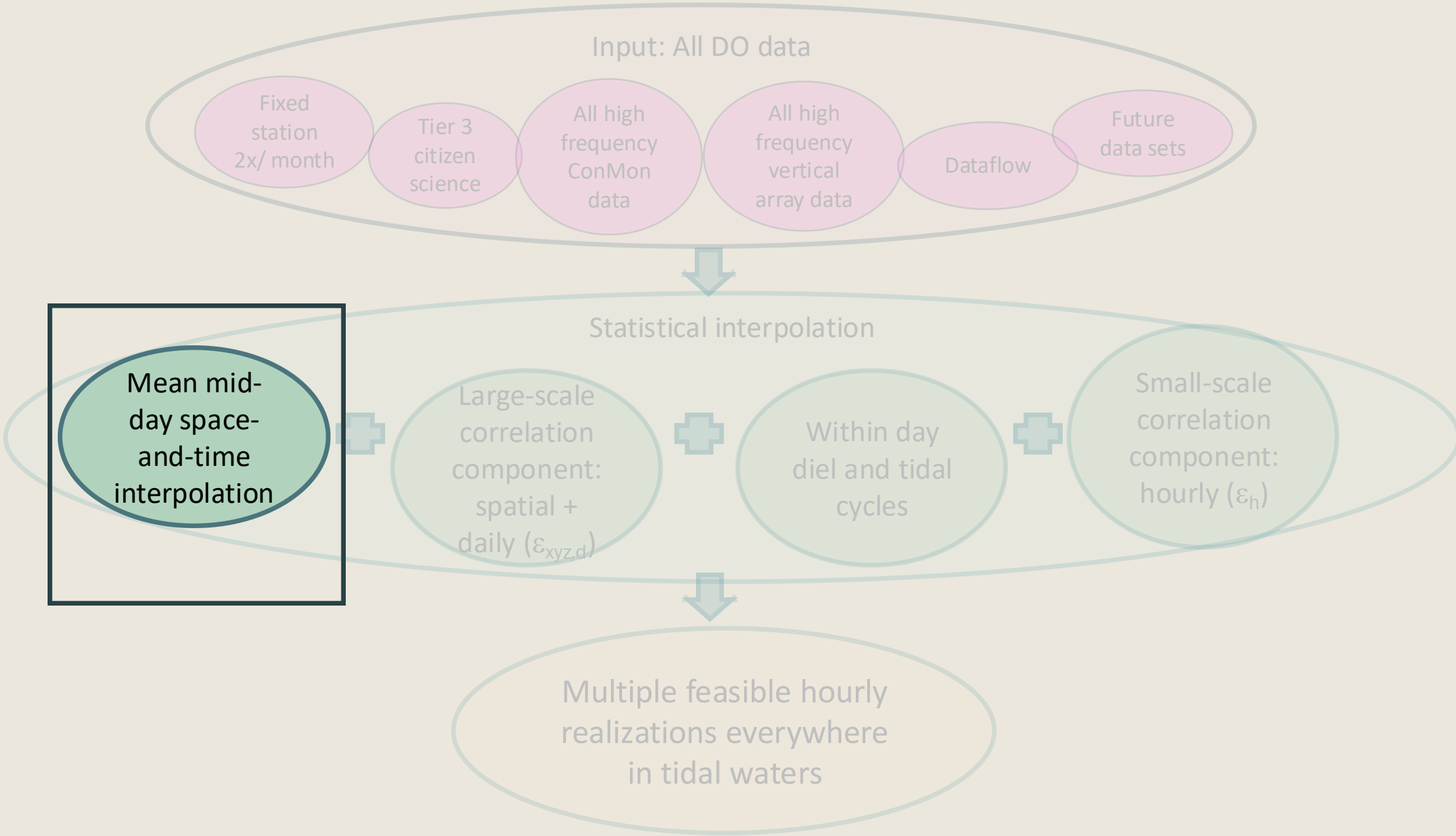


From http://vecos.vims.edu/

Citizen monitoring (CMC and multiple community groups)

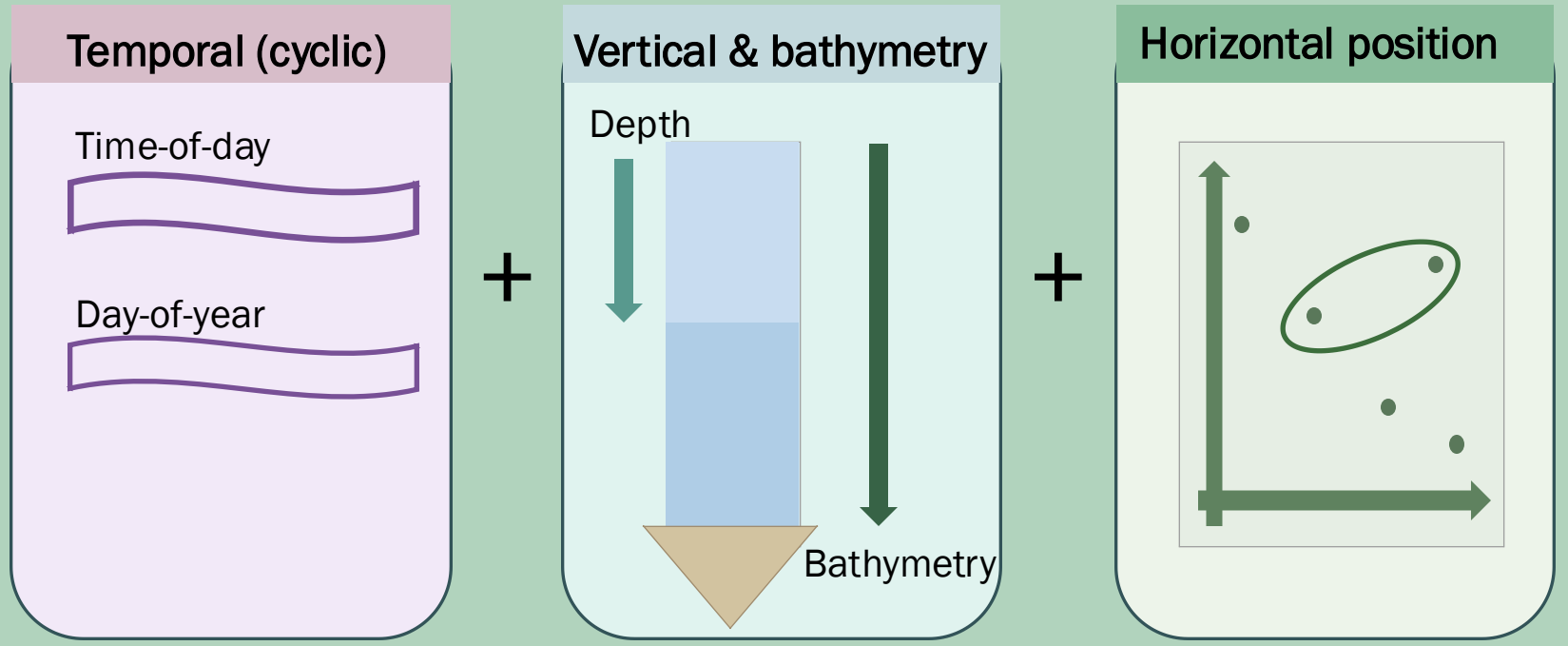


Chesapeake Monitoring Cooperative: https://www.chesapeakemonitoringcoop.org/


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
Mean mid-day space-and-time interpolation
Large-scale correlation component: spatial + daily ($\varepsilon_{xyz,d}$)
Within day diel and tidal cycles
Small-scale correlation component: hourly (ε_h)
Multiple feasible hourly realizations everywhere in tidal waters

Mean mid-day space-and-time interpolation

- **Goal:** Capture the central tendency of the data. Output an estimate every day, everywhere, to add to hourly variability.



Uses Generalized Additive Models (GAMs) with:

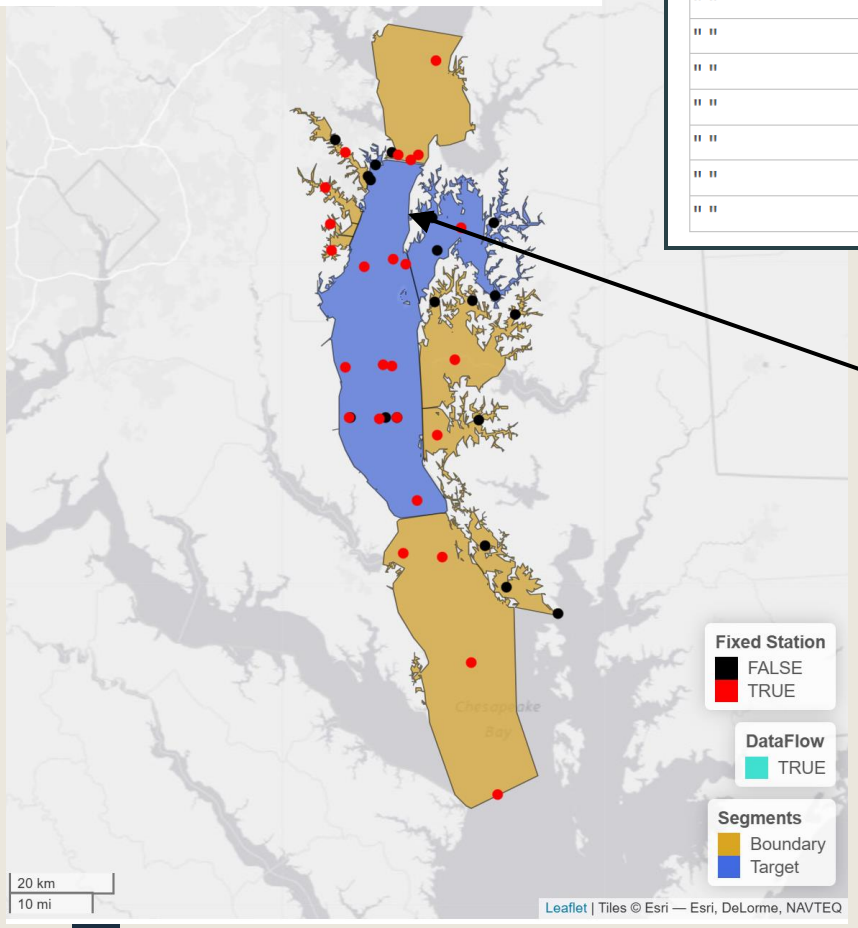
- **Cyclic smooths (cc)** for time-of-day and day-of-year.
- **Tensor-product smooths (ti)** to model interactions among depth, bathymetry, and horizontal position.

Mean mid-day space-and-time interpolation

Example region: mid-bay

Purple are target segments

Tan are boundary segments



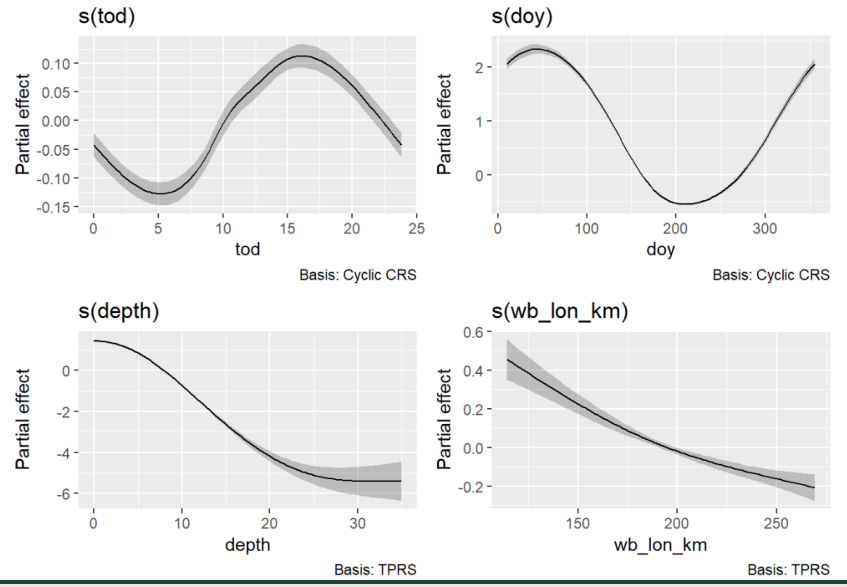
Multiple smooth terms are fit to the data

Tabular output

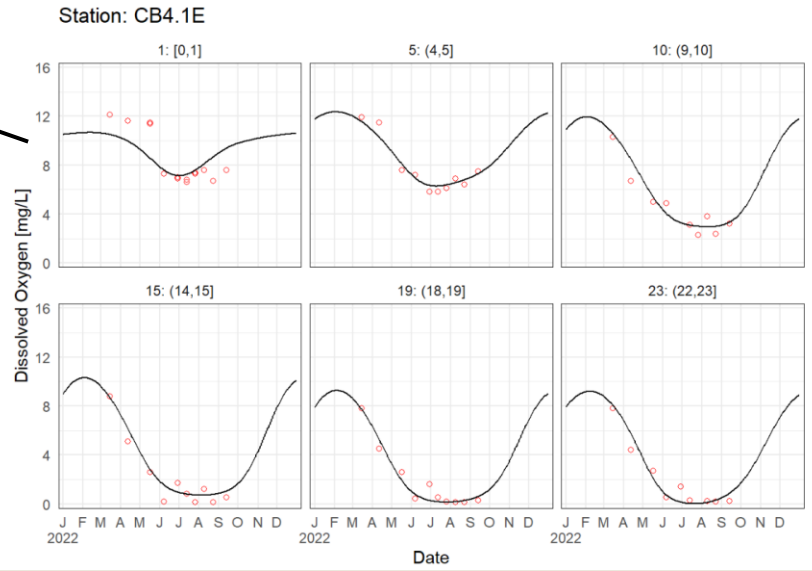
Approximate significance of smooth terms:

type	source	df	F	p.value
parametric terms	NA			-
smoothed terms	s(tod)	3.47	102.7910	<0.0001
" "	s(doy)	4.00	2164.2063	<0.0001
" "	s(depth)	2.46	2229.1807	<0.0001
" "	s(wb_lon_km)	2.15	36.8366	<0.0001
" "	s(wb_lat_km)	2.68	35.6186	<0.0001
" "	s(depth_b)	3.00	78.5430	<0.0001
" "	ti(wb_lat_km,depth,wb_lon_km)	5.64	16.0468	<0.0001
" "	ti(depth,depth_b)	4.00	142.3053	<0.0001
" "	ti(depth,doy)	5.99	406.9135	<0.0001
" "	ti(wb_lon_km,doy)	4.59	18.8503	<0.0001
" "	ti(wb_lat_km,doy)	5.17	57.9243	<0.0001
" "	ti(depth_b,doy)	5.59	11.9146	<0.0001
" "	ti(doy,tod)	3.44	1.4268	0.0019

Graphical output of partial smooth effects



Example GAM fit at a sample location



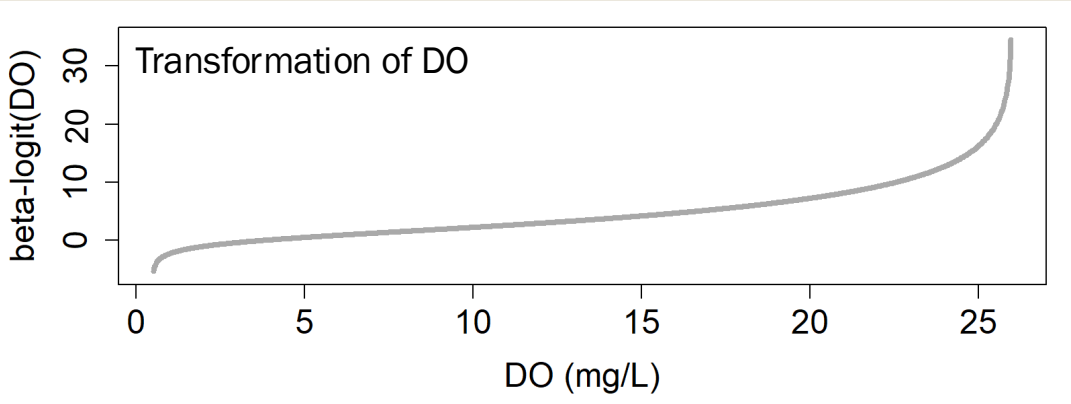
Some findings related to GAMs for interpolation

1. A transformation is helpful for capturing key features of low DO

2. Spatial and temporal limits needed to fit smooth functions for purpose of interpolation

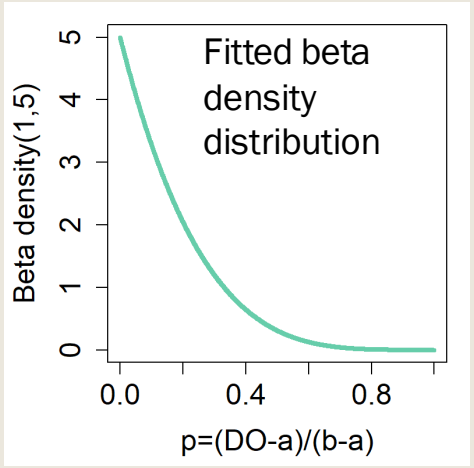
1. DO transformation

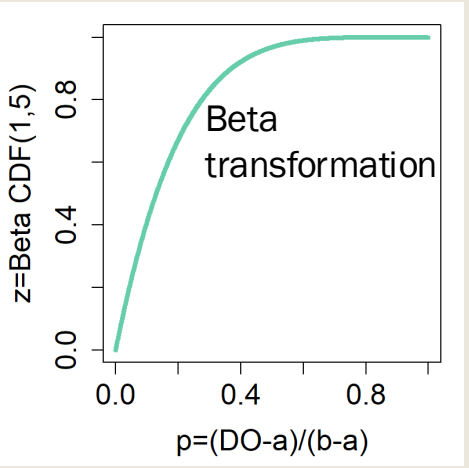
- Reasoning:
  - *Predictions will likely be improved if a transformation is used to yield approximately normal residuals.*
  - *A transformation can help avoid negative DO estimates, and*
  - *It can be targeted to the lower tail of the distribution, important for criteria assessment.*

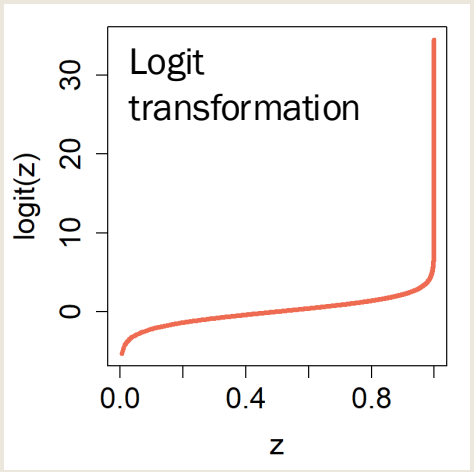


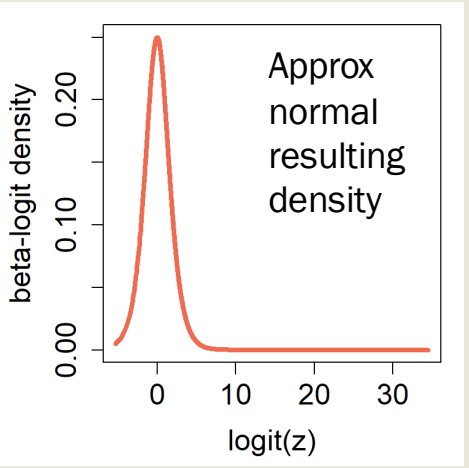
1. DO transformation

- Approach: Beta-logit
  - *Scale data 0 to 1.*
  - *Beta: Apply beta distribution to account for skewness of DO.*
  - *Logit: Use logit transformation to create an approximately normal distribution.*





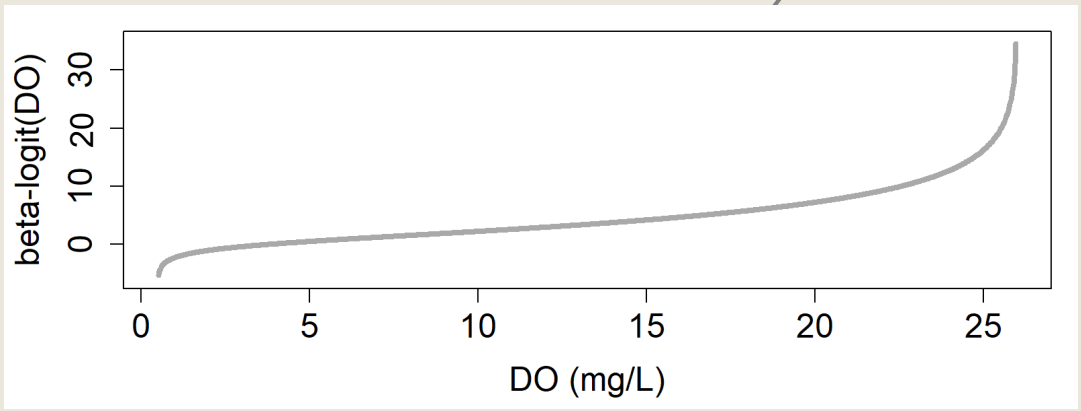


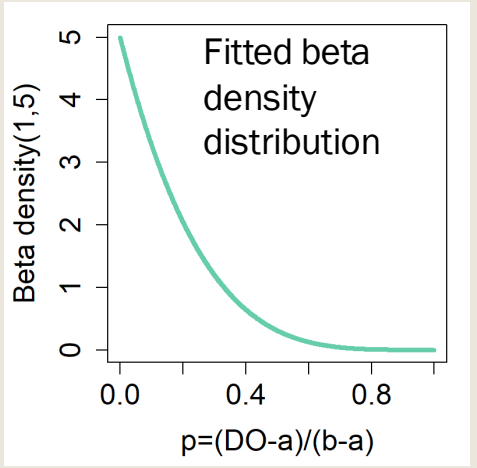


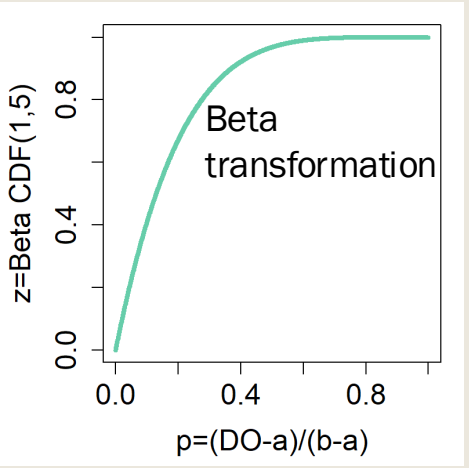
Method development from Elgin Perry

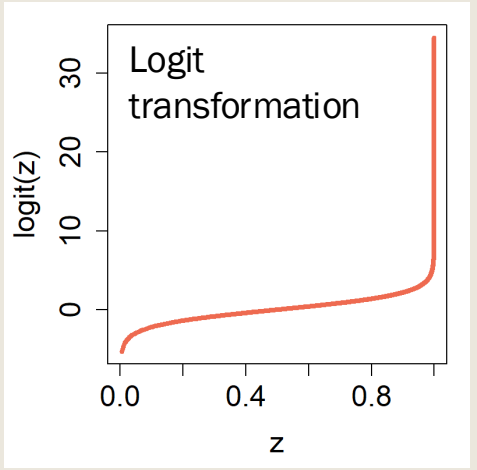
1. DO transformation

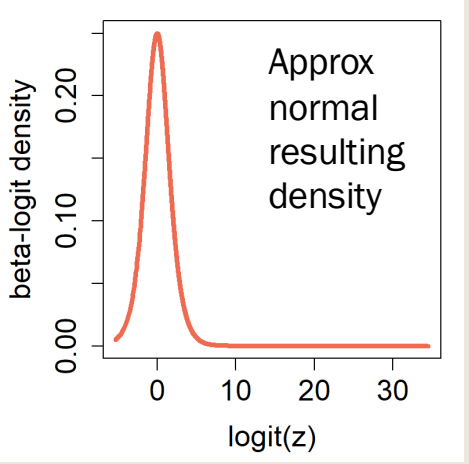
- Approach: Beta-logit
  - *Scale data 0 to 1.*
  - *Beta: Apply beta distribution to account for skewness of DO.*
  - *Logit: Use logit transformation to create an approximately normal distribution.*







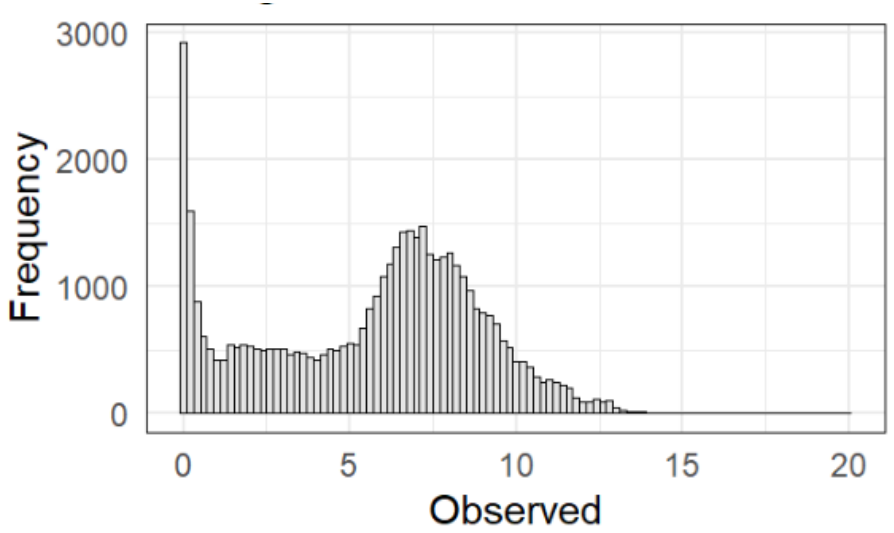




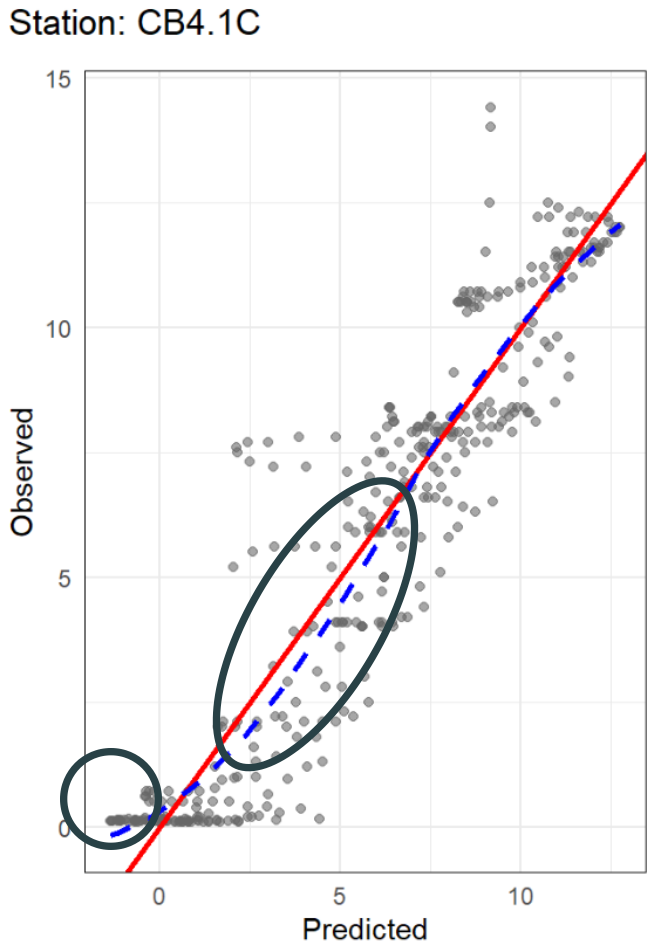
Method development from Elgin Perry

1. DO transformation: example

2022 DO in a mid-bay region:
Data is both skewed and has high count of zeros

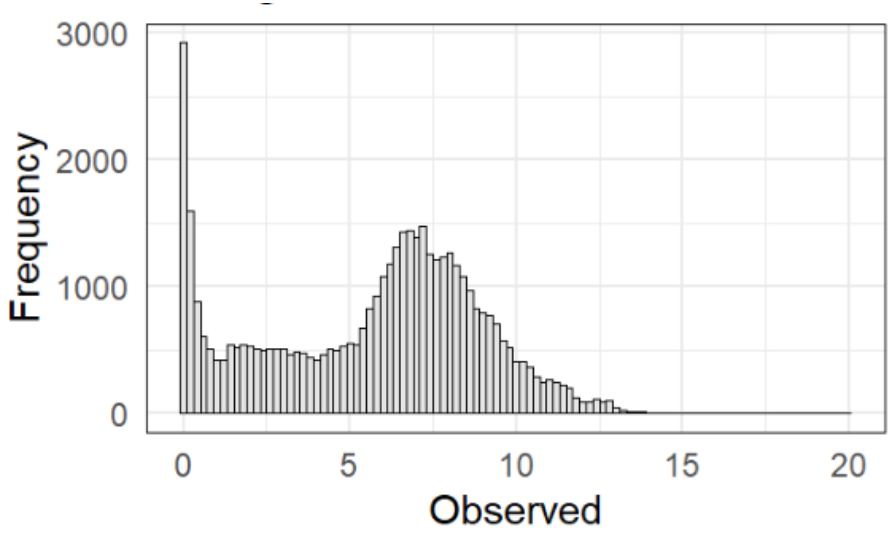


One station: Observed vs. predicted with no transformation

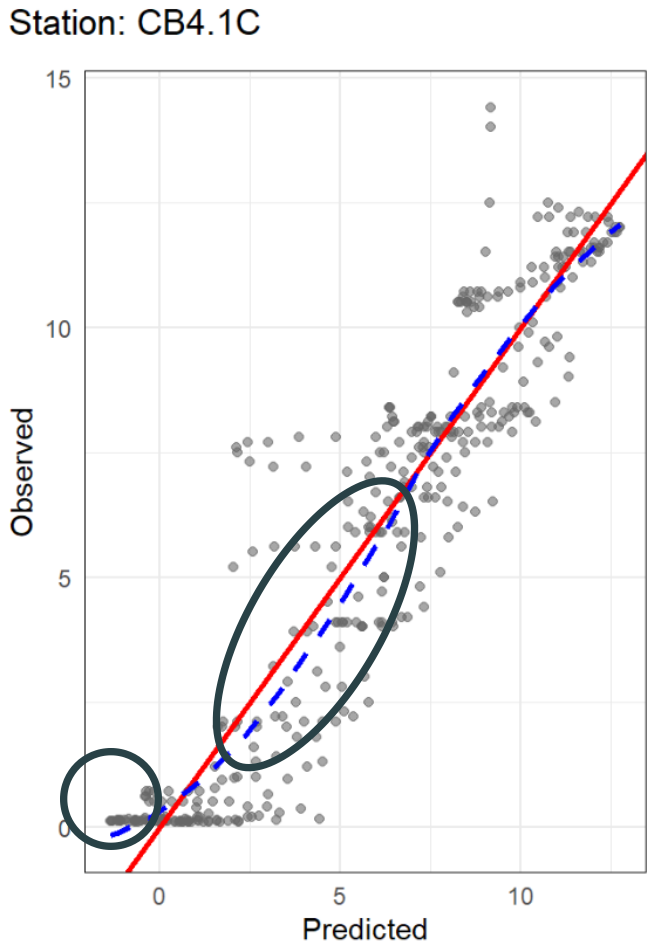


1. DO transformation: example

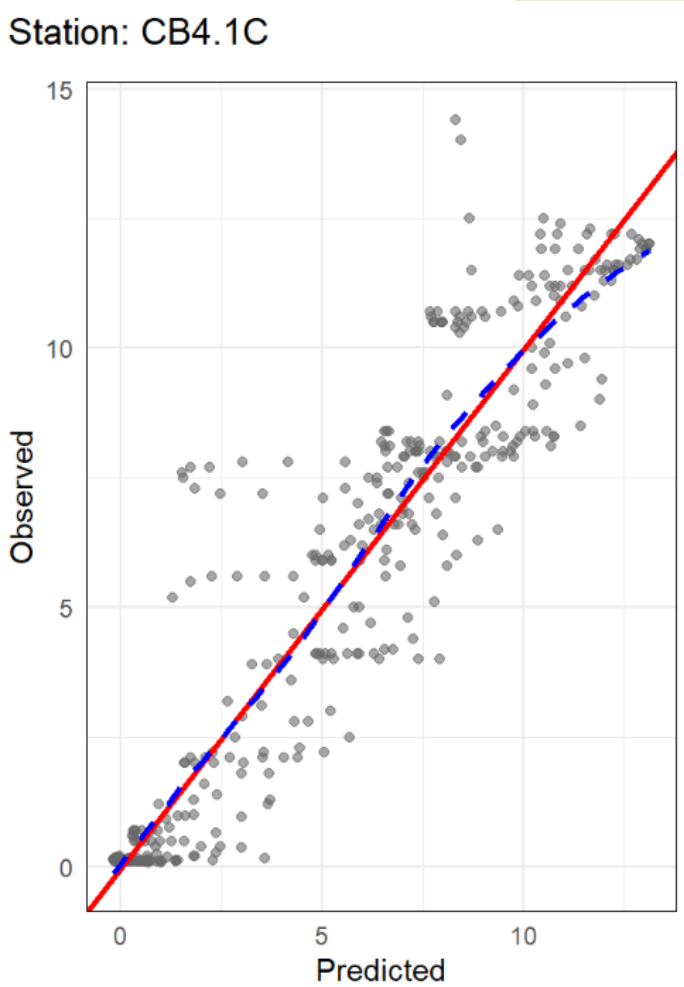
2022 DO in a mid-bay region:
Data is both skewed and has high count of zeros



One station: Observed vs. predicted with no transformation



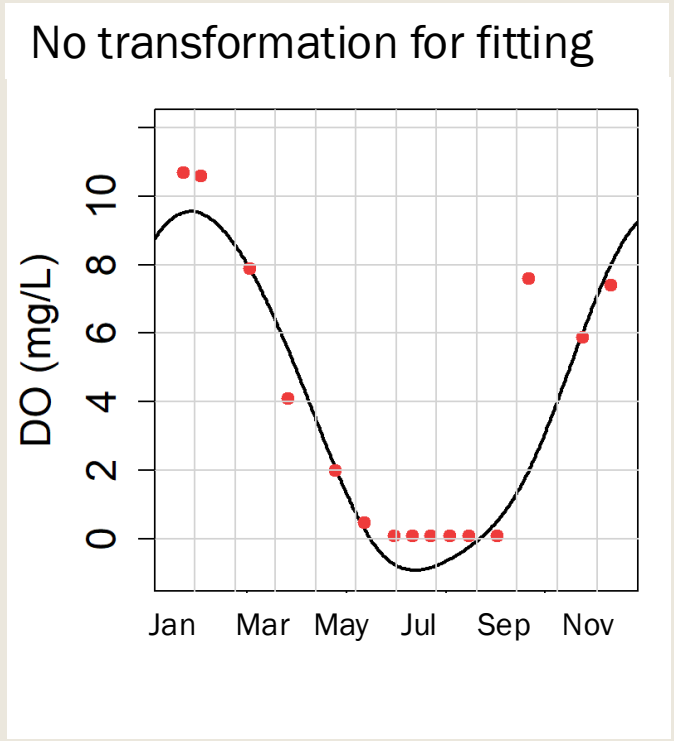
With transformation

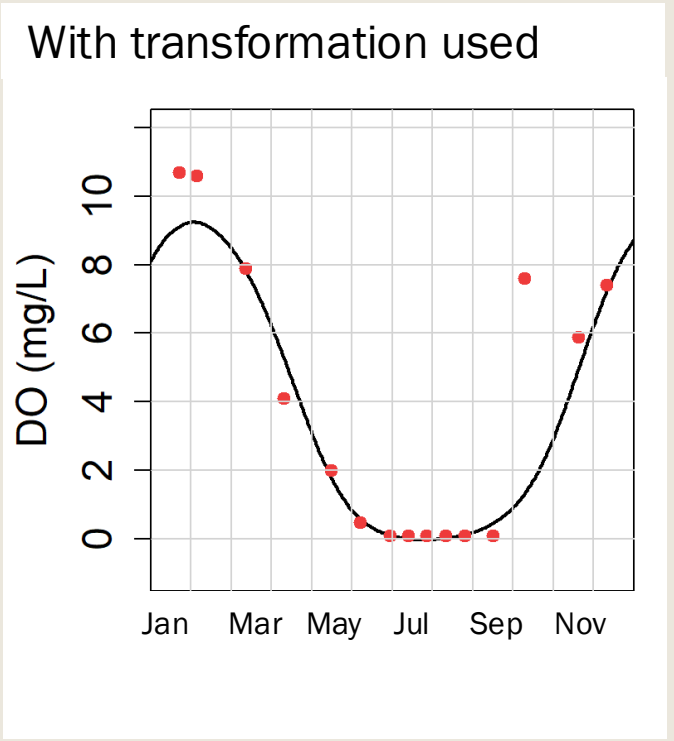


1. DO transformation: example

2022 CB4.1C (mid-bay) 18-19m

With **no** transformation, negative estimates are generated from the mean mid-day interpolation



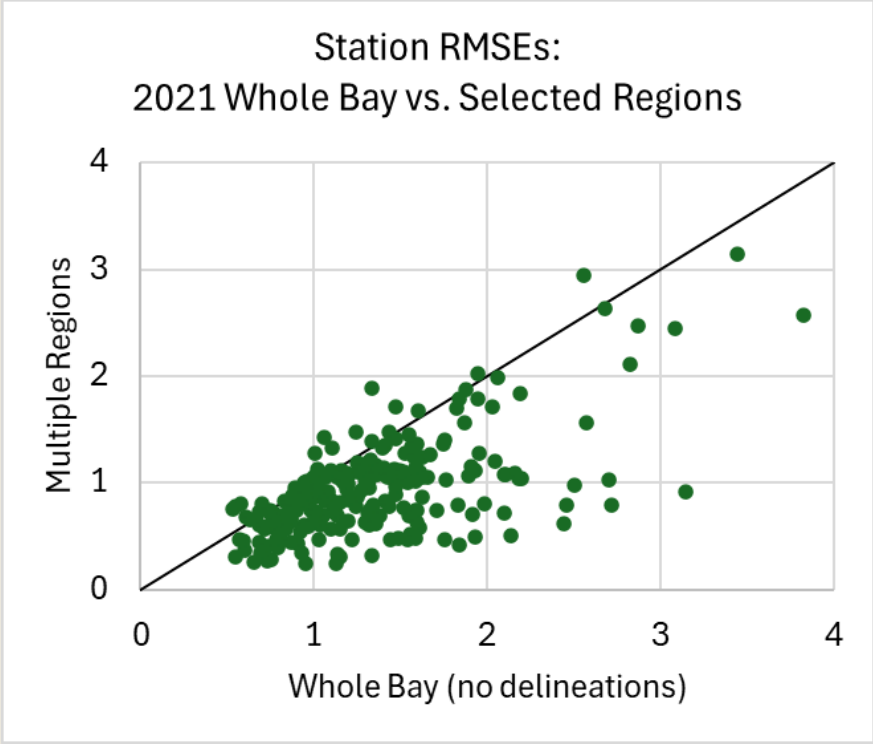


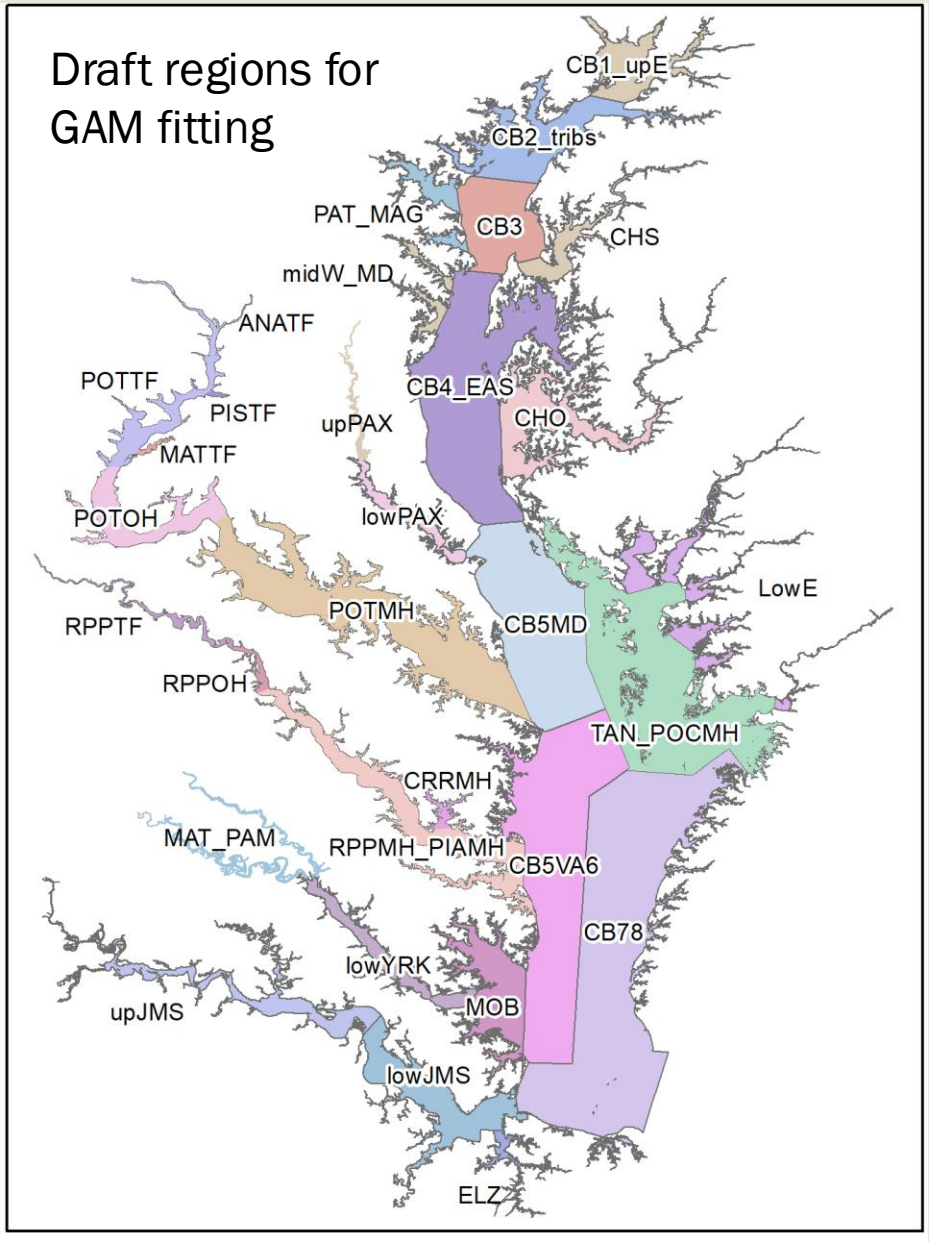
With **beta-logit** transformation, results match the lowest DO better

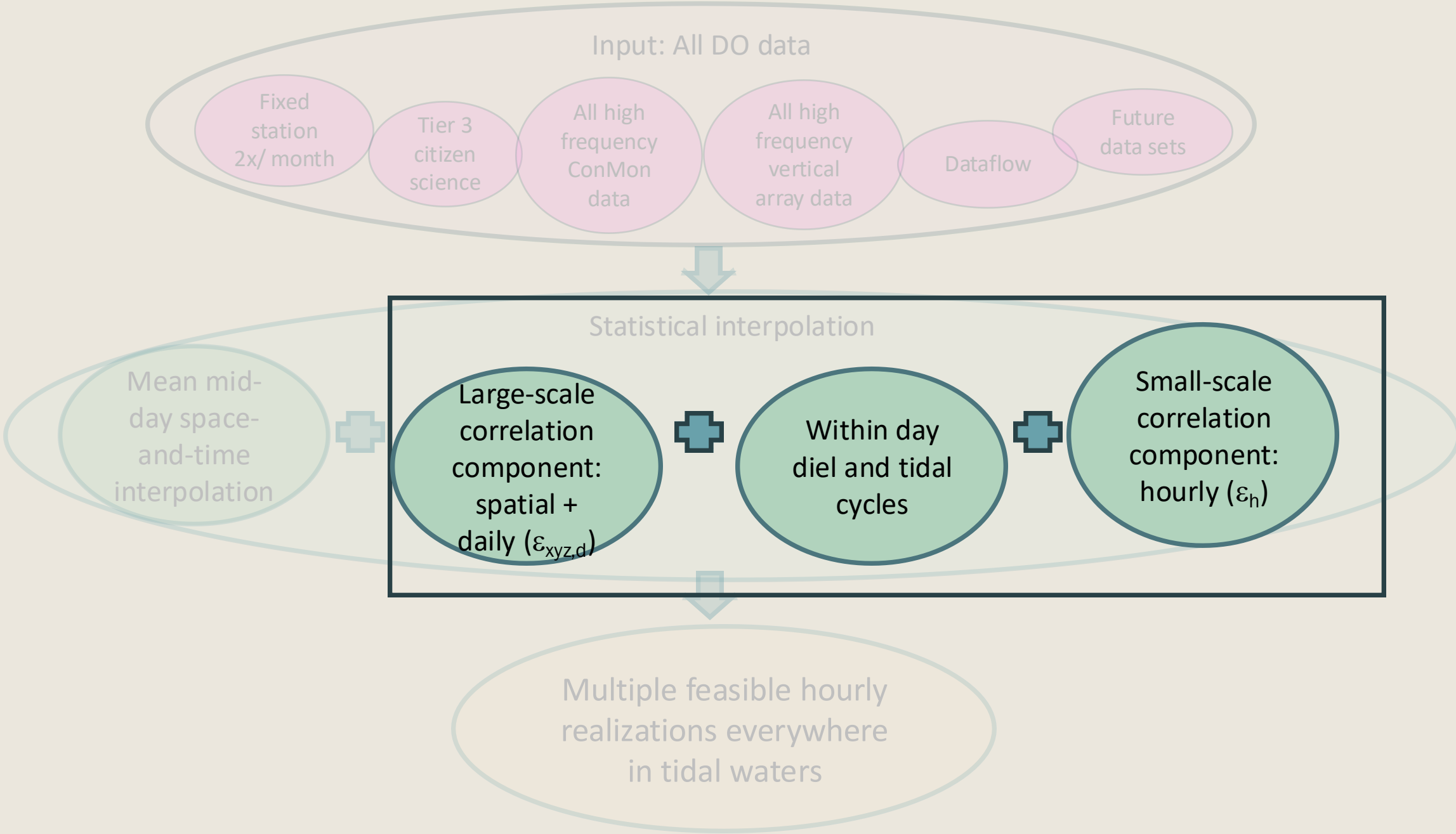
2. Spatial and temporal limits for GAMs

One GAM fit to entire tidal waters for multiple years is not optimal for this complex system. We settled on:

- *One year at a time, and*
- *Split the bay into regions.*

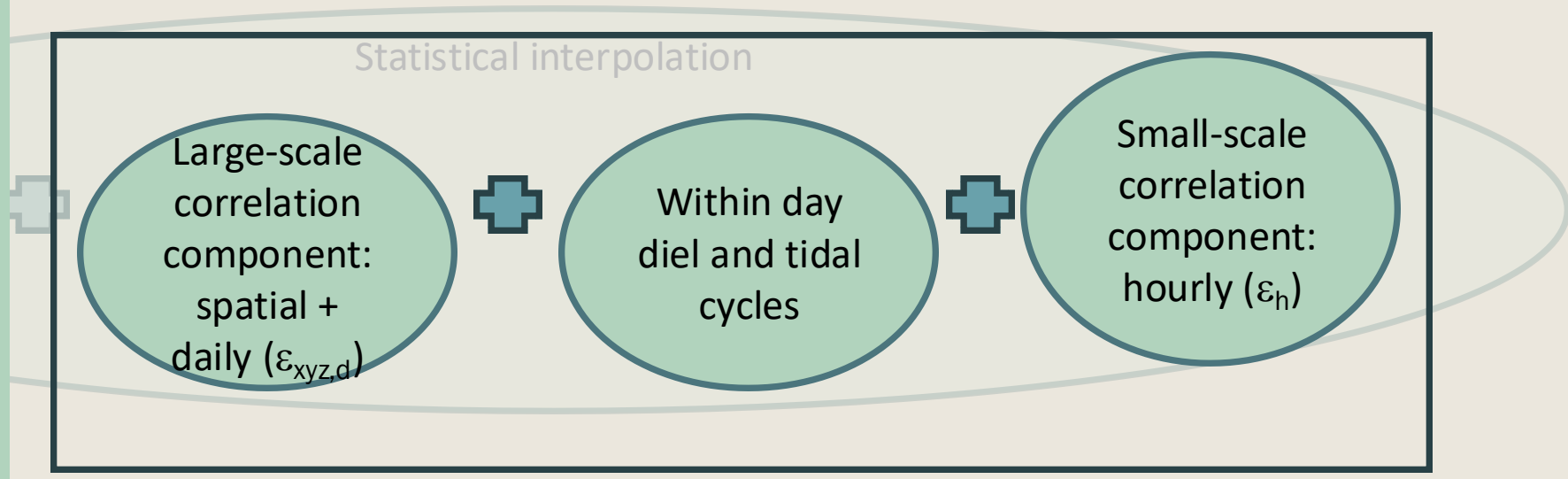





Input: All DO data
Fixed station 2x/ month
Tier 3 citizen science
All high frequency ConMon data
All high frequency vertical array data
Dataflow
Future data sets
Statistical interpolation
Mean mid-day space-and-time interpolation
Large-scale correlation component: spatial + daily ($\varepsilon_{xyz,d}$)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

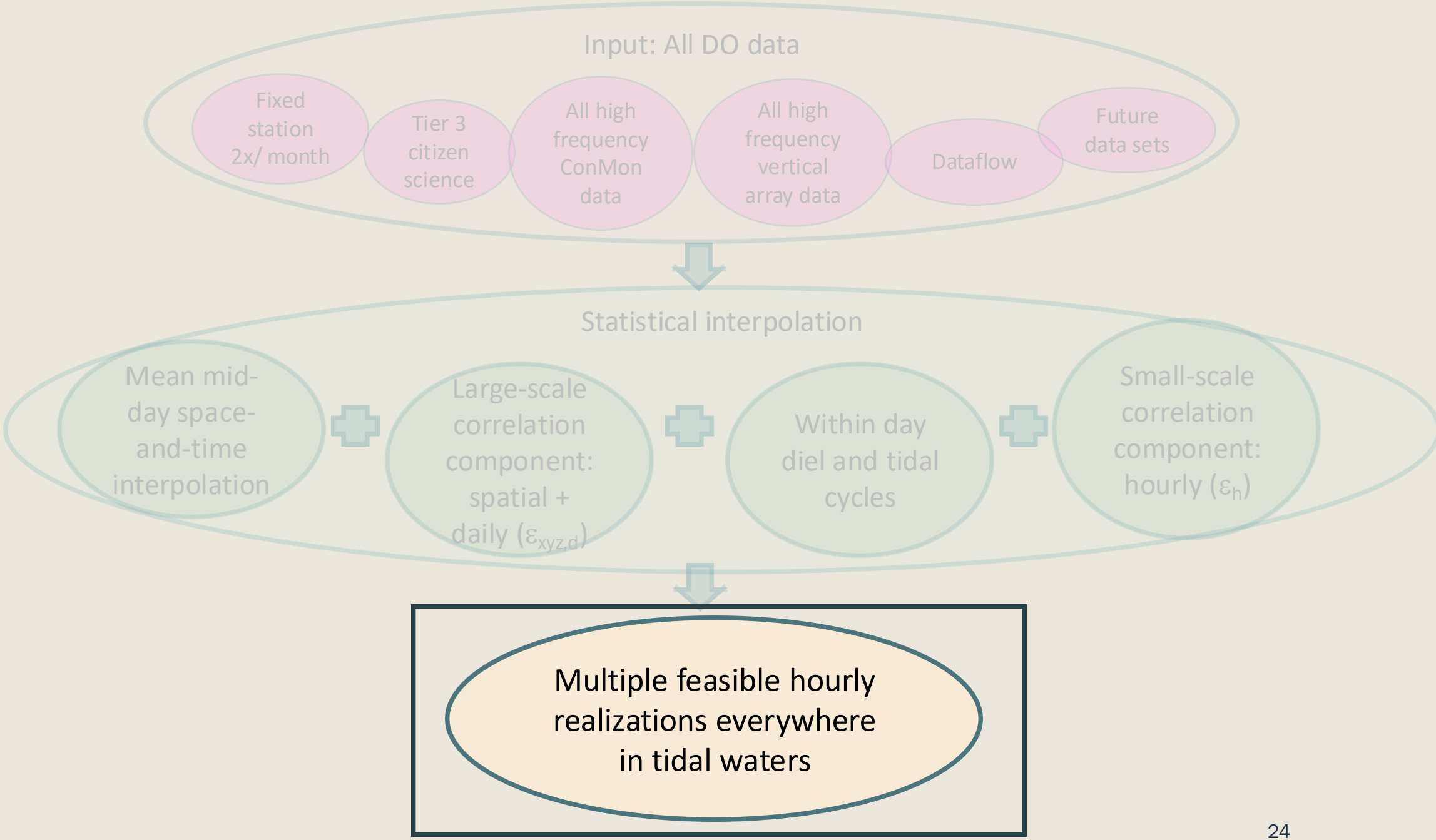
Evaluate high-frequency data for within-day (Conmon) and location-to-location (Dataflow) variability.

Apply those patterns to generation hourly simulation results.



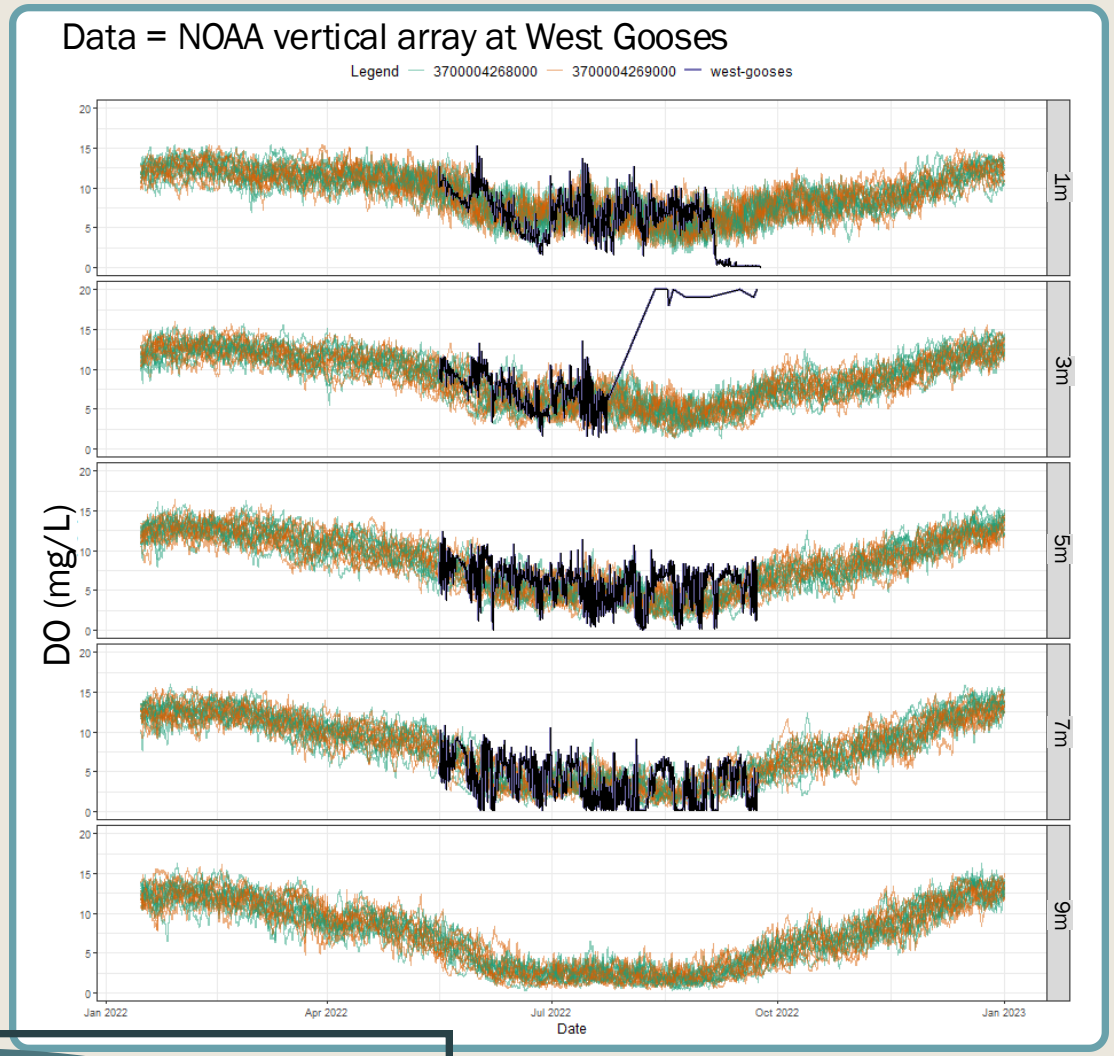
Hear more about this in the next presentation!

A space-time interpolation tool for Chesapeake Bay dissolved oxygen: Parameterizing a 4-dimensional correlation structure
Jon Harcum, Tetra Tech


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
Mean mid-day space-and-time interpolation
Large-scale correlation component: spatial + daily ($\varepsilon_{xyz,d}$)
Within day diel and tidal cycles
Small-scale correlation component: hourly (ε_h)
Multiple feasible hourly realizations everywhere in tidal waters

100 realizations of hourly estimates everywhere for a year.

This example was an early test at one location, compared to observed high frequency data (black lines).



Multiple feasible hourly realizations everywhere in tidal waters

Thank you!

For more info: email Rebecca Murphy, rmurphy@chesapeakebay.net

Data sources

- Fixed-station network:
  - *DOEE, MDDNR, VADEQ, CBP*
  - *https://datahub.chesapeakebay.net/*
- Cruise-track monitoring (dataflow) and Continuous monitoring (ConMon)
  - *VECOS, MDDNR*
  - *http://vecos.vims.edu/ and https://eyesonthebay.dnr.maryland.gov/*
- Vertical arrays
  - *NOAA*
  - *https://sensors.ioos.us/*
- Additional State Agency data
  - *MDE, VADEQ, DNREC*
  - *Some in https://datahub.chesapeakebay.net/*
- Citizen monitoring
  - *Coordinated through CMC, with many community providers*
  - *https://cmc-dev.vims.edu/data-explorertoring*

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).