Incorporating sensor-based high-frequency data in criteria assessment - Discussion

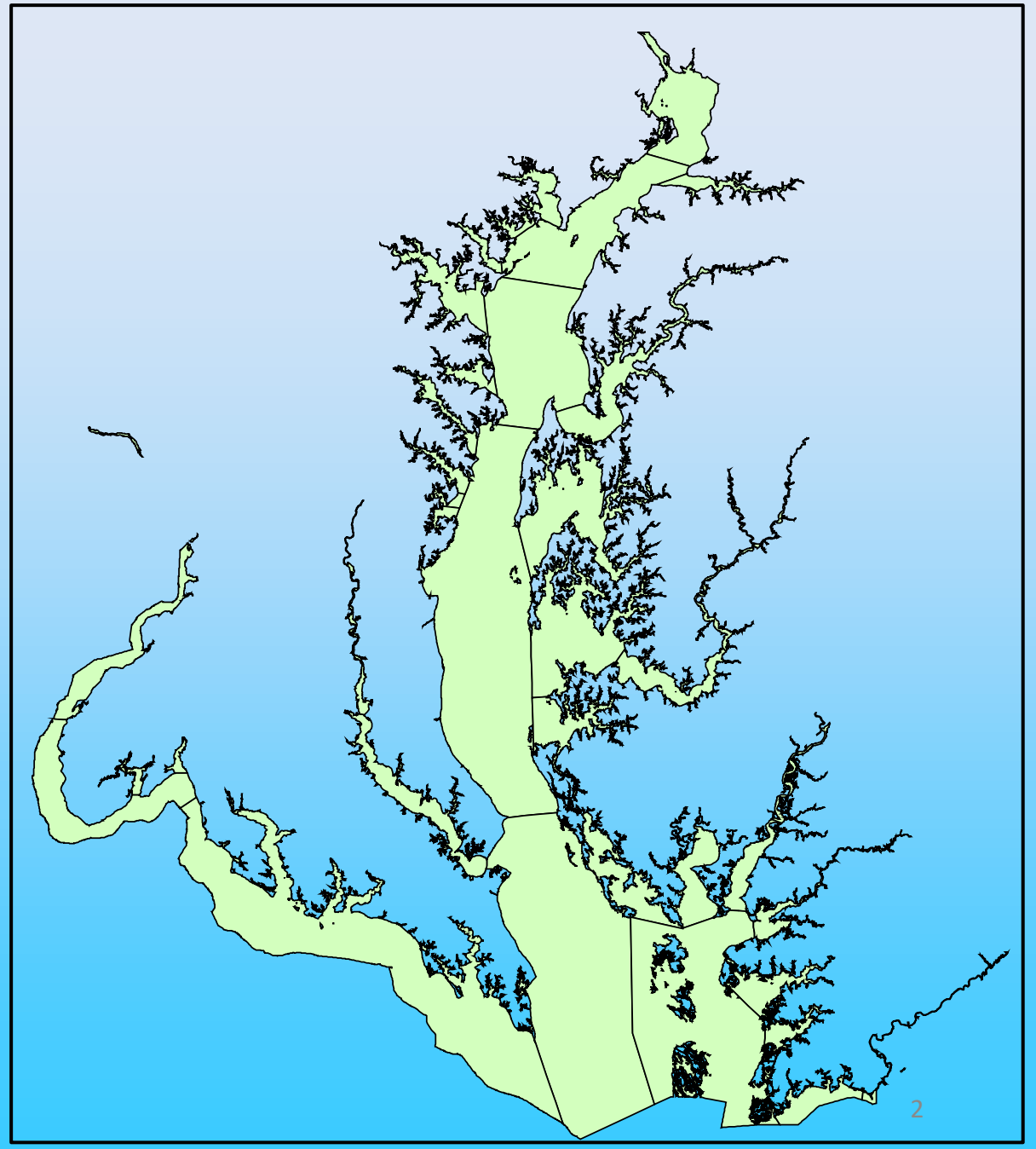
Richard Tian and
modeling team

Modeling Quarterly Review Meeting

Annapolis, Jan. 08, 2025

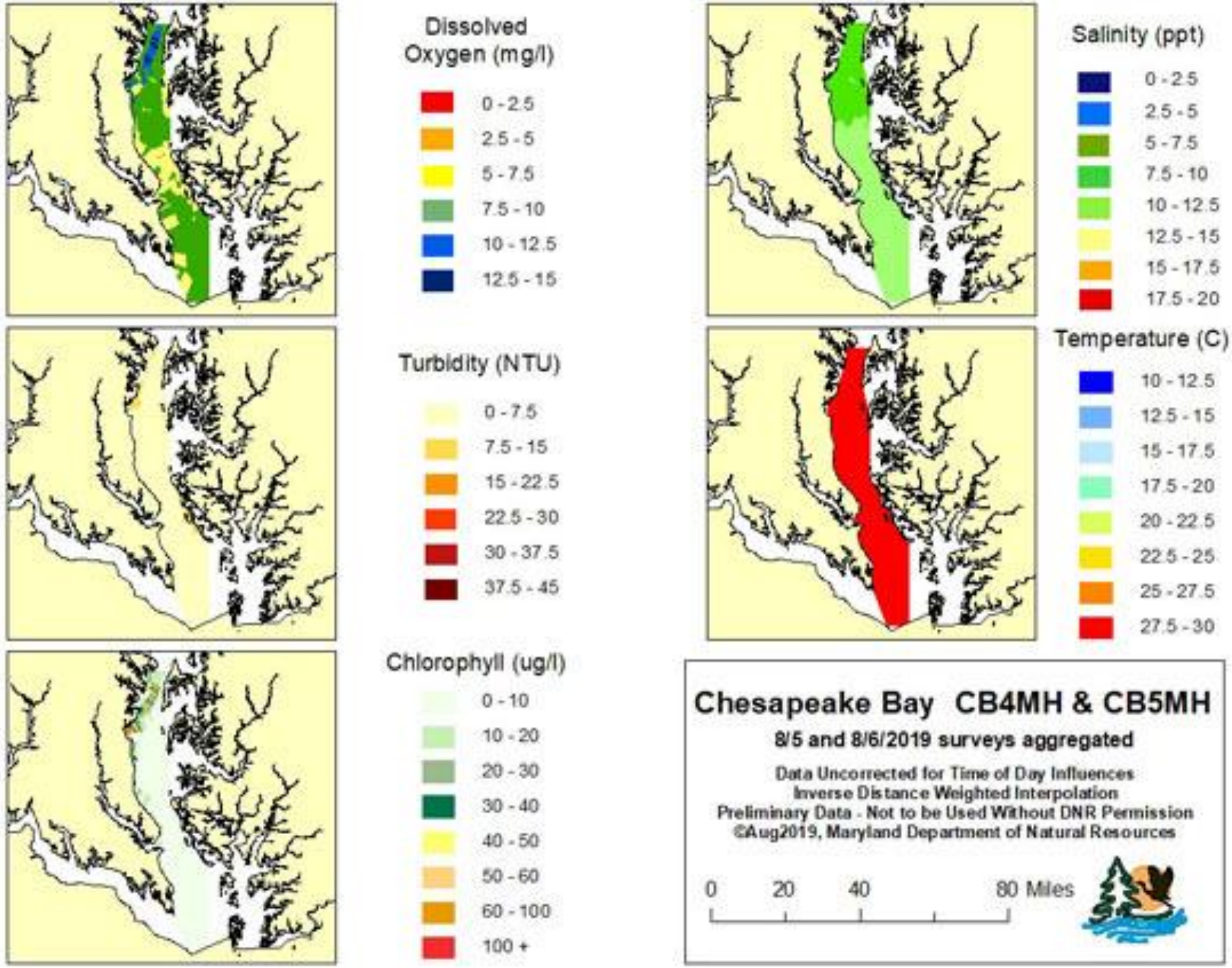
Datafow occupation in the MD portion of Chesapeake Bay

Sampling Dates:
Fishing Bay: Jun 21, 2005
Fishing Bay: May 19, 2005
Fishing Bay: Apr 28, 2005
Fishing Bay: Oct 04, 2004
Fishing Bay: Sep 01, 2004
Fishing Bay: Aug 04, 2004
Fishing Bay: Jul 01, 2004
Fishing Bay: Jun 01, 2004
Fishing Bay: May 03, 2004
Fishing Bay: Apr 07, 2004
Fishing Bay: Oct 23, 2003
Fishing Bay: Sep 22, 2003
Fishing Bay: Aug 27, 2003
Fishing Bay: Jul 31, 2003
Fishing Bay: Jun 09, 2003
Fishing Bay: May 22, 2003

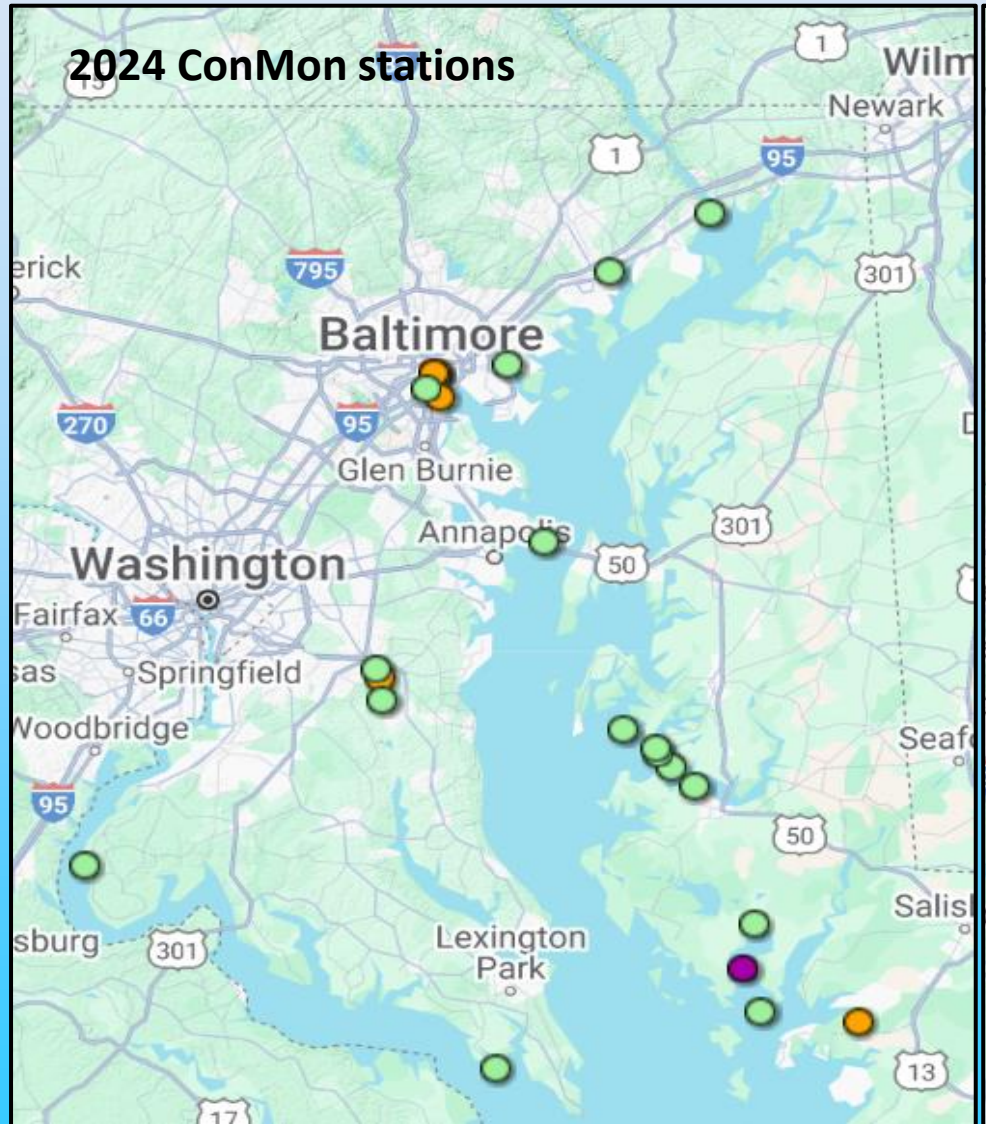


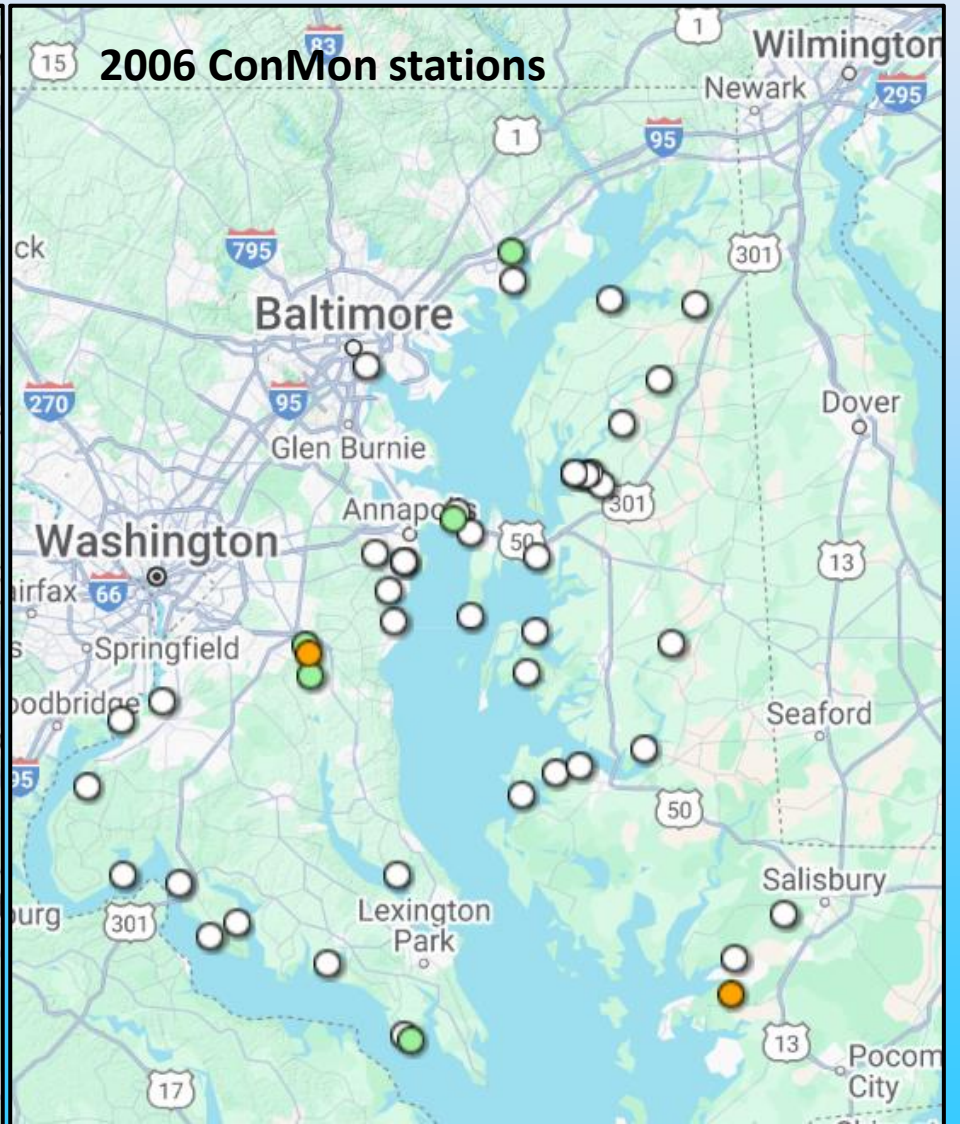
Dataflow in CB4 and CB5

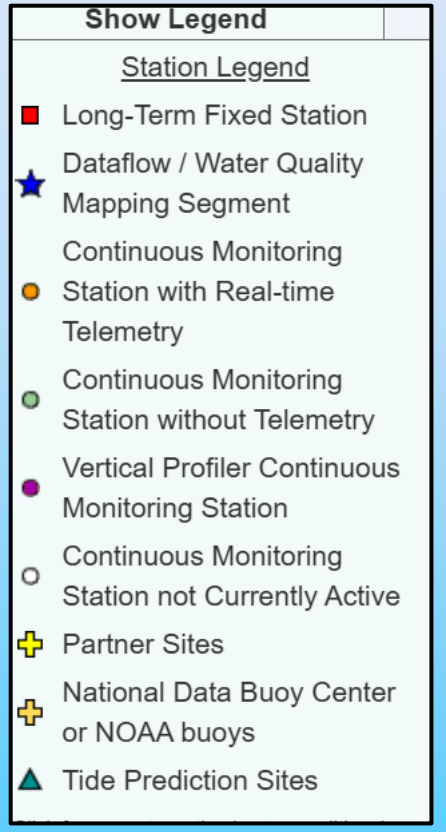
Every 4 seconds
traveling at speeds up to 20 knots



MD ConMon stations





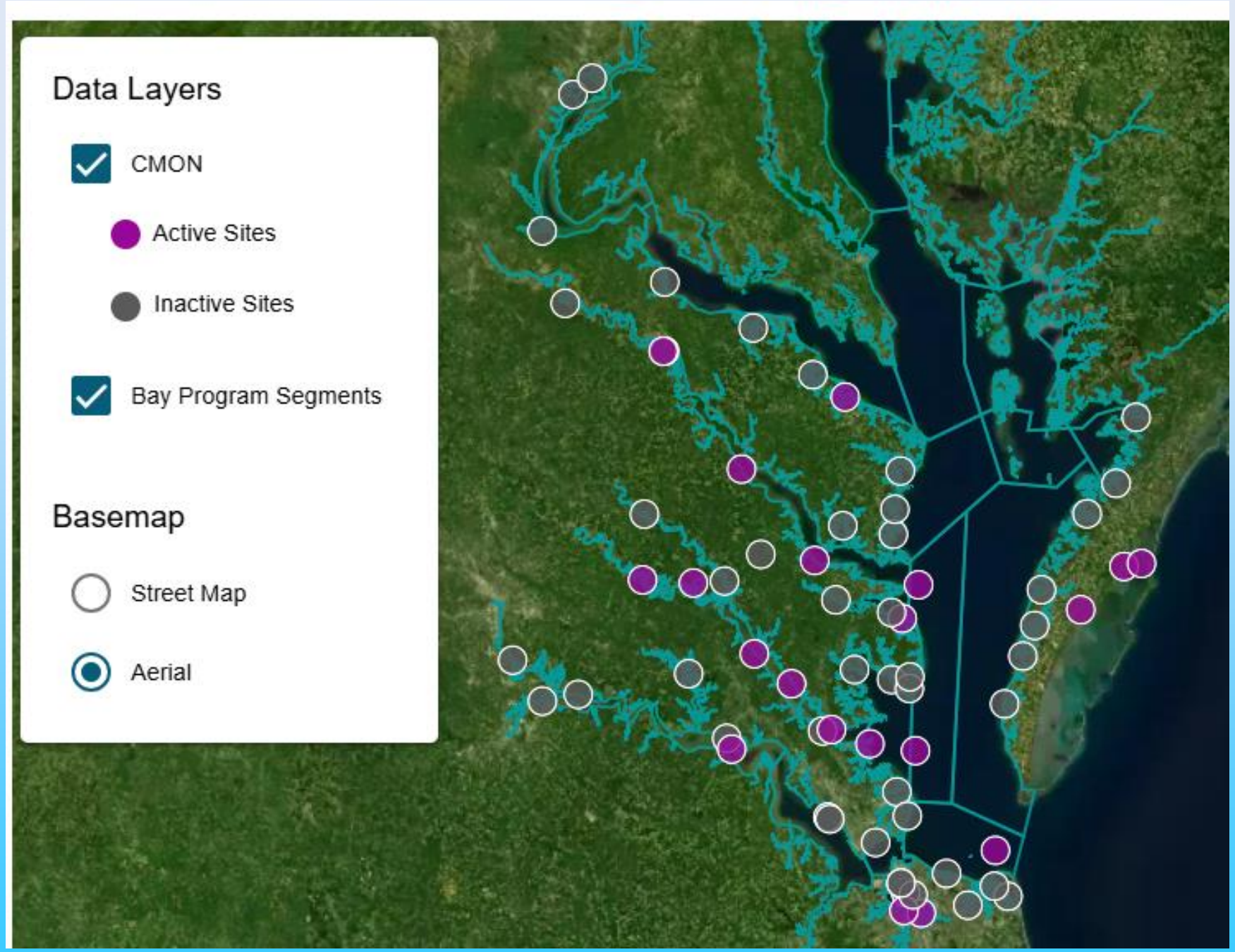


Eyes on the Bay

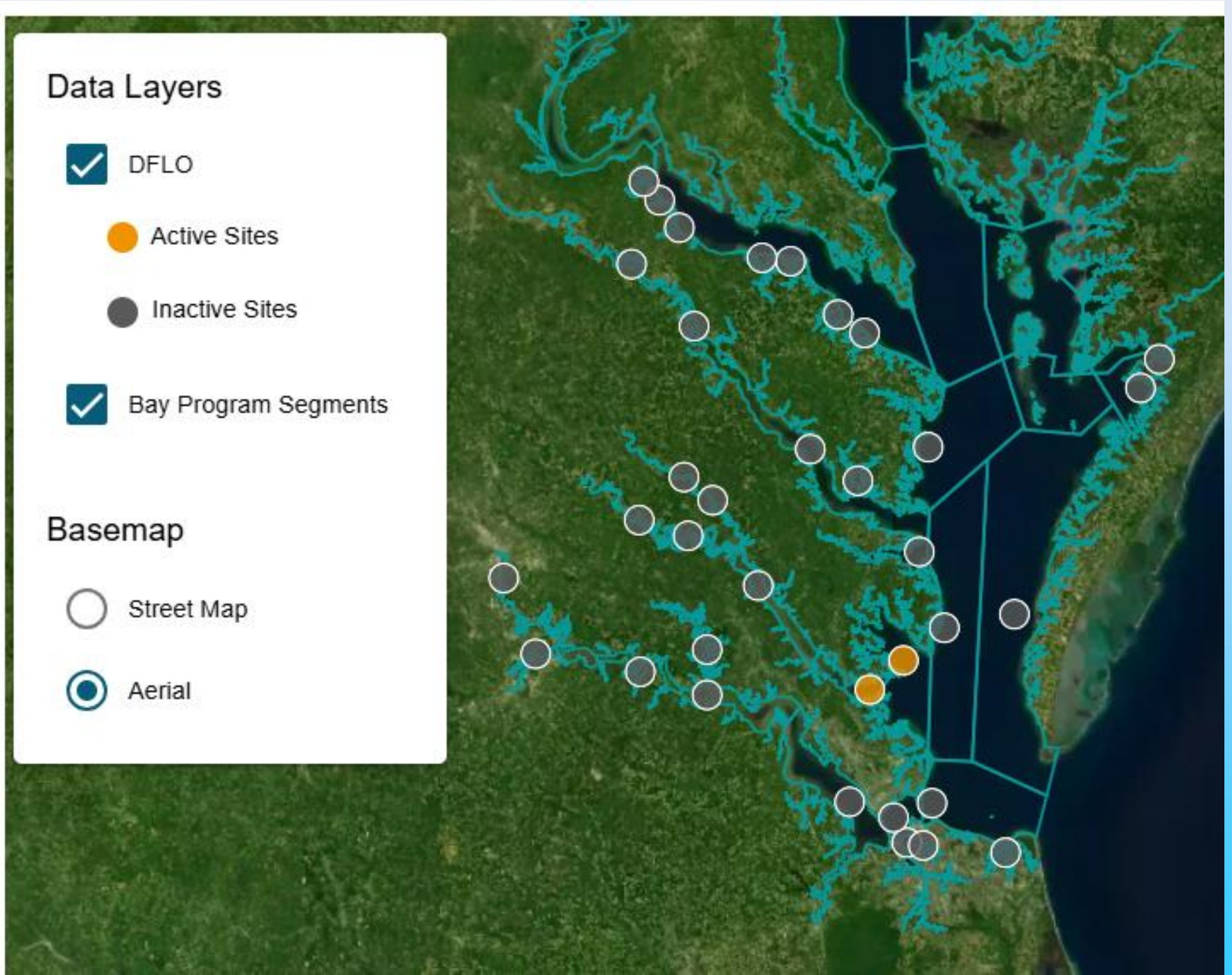
MD ConMon data collection

- Each instrument is programmed to record seven parameters every 15 mins.
- Water temperature, salinity, dissolved oxygen (DO) saturation, DO concentration, pH, turbidity, and fluorescence (a measure of chlorophyll-a present in the water column).
- Every two weeks, the water quality meters need to be exchanged for cleaning and calibration.
- 141 rotational ConMon sensors-stations.
- Mostly 1m below surface, some 0.3-0.5m above the bottom.
- 8 rotational profilers, every 1-3 hours data collection at 1-2m vertical intervals.

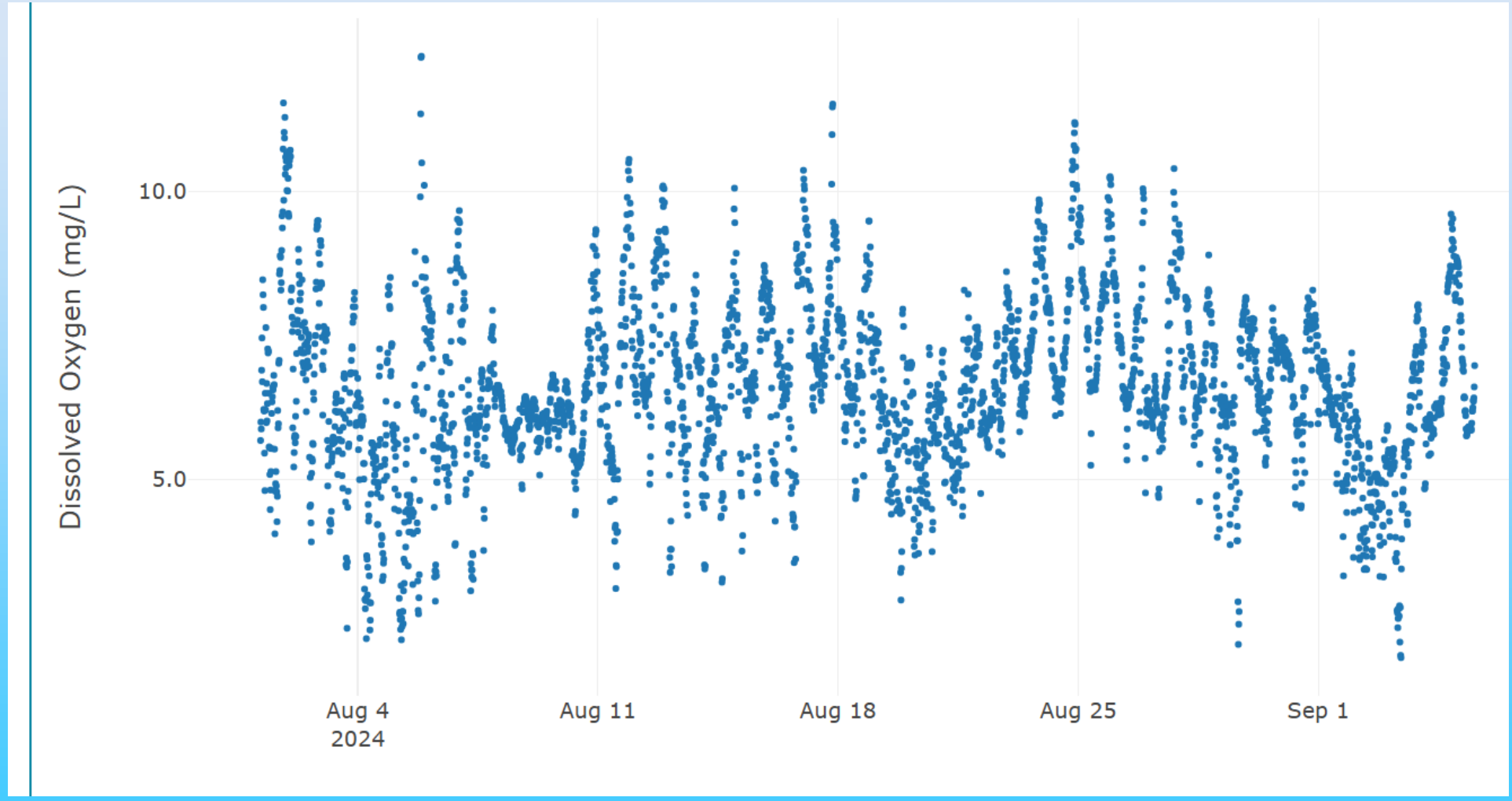
VECOS ConMon data 67 rotational stations



VECOS:
35 Rotational DATAFLOW occupations



Gloucester Point ConMon data



ConMon data inventory

(Jhon Harcum, 2024)

Table 1. Number of stations and DO observations for CBP DataHub (1984-2022), Eyes on the Bay (2001-2022), and VECOS (2003-2022).

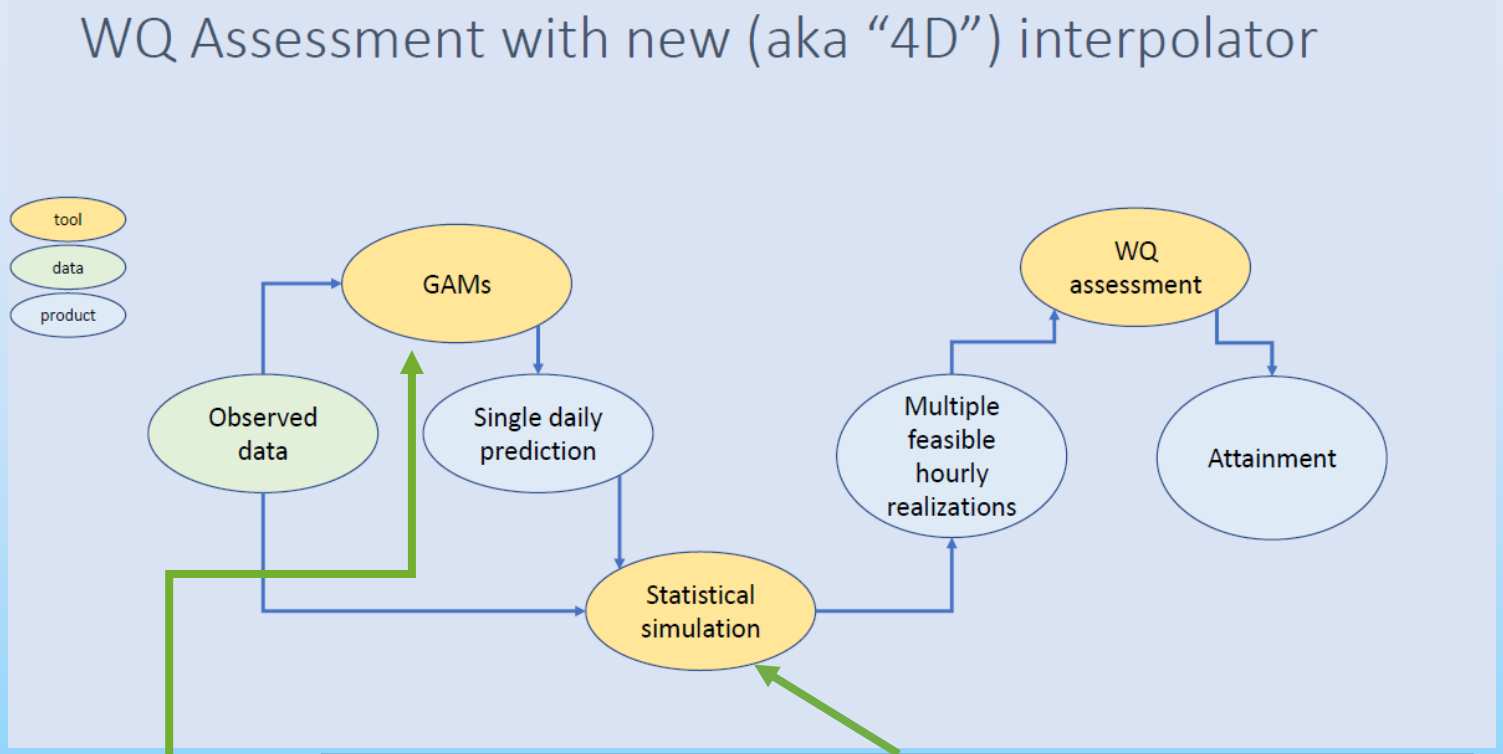
Source	Date	Stations	Observations		
			All	Unique station/date/depth	Unique station/date
DataHub (fixed stations)	1984-2022	156	672,000	669,000	n/a
DataHub (non-fixed stations)	1984-2022	679	148,000	146,000	n/a
Eyes on the Bay (EOTB)	2001-2022	126	11,917,000	n/a	*131,000
VECOS	2003-2022	52	6,640,000	n/a	71,000
NOAA vertical arrays	2022	2	185,000	1,400	230

*EOTB includes a few stations with two sensors placed nominally at `bottom` and `surface`. These stations were summarized at the station/date/layer level.

How to incorporate these data in criteria assessment?

4D interpolator

Lead by Tango and Murphy




```
Formula:DO ~ s(cyear, k = 20) + s(doy, bs = "cc") +
s(wDepth) +     s(LonKm) + s(bDepth) + ti(LonKm,
wDepth) + ti(bDepth, wDepth) +     ti(cyear, wDepth) +
ti(cyear, LonKm) + ti(cyear, bDepth) +     ti(doy,
wDepth, bs = c("cc", "tp")) + ti(doy,     LonKm, bs =
c("cc", "tp")) + +ti(doy, bDepth,     bs = c("cc",
"tp")) + ti(cyear, doy, bs = c("tp",     "cc"))
```

2. Development & testing

Uses Fourier analysis to estimate the expected deviation from the mid-day DO at each hour of the day

Estimate the Hourly DO Deviations

Example: Fourier analysis with just daily cycle to fit hourly DO (DO_h)

$$DO_h = lc * h + sc * sin\left(\frac{2\pi * h}{24}\right) + cc * cos\left(\frac{2\pi * h}{24}\right) + \tau$$

coefficients

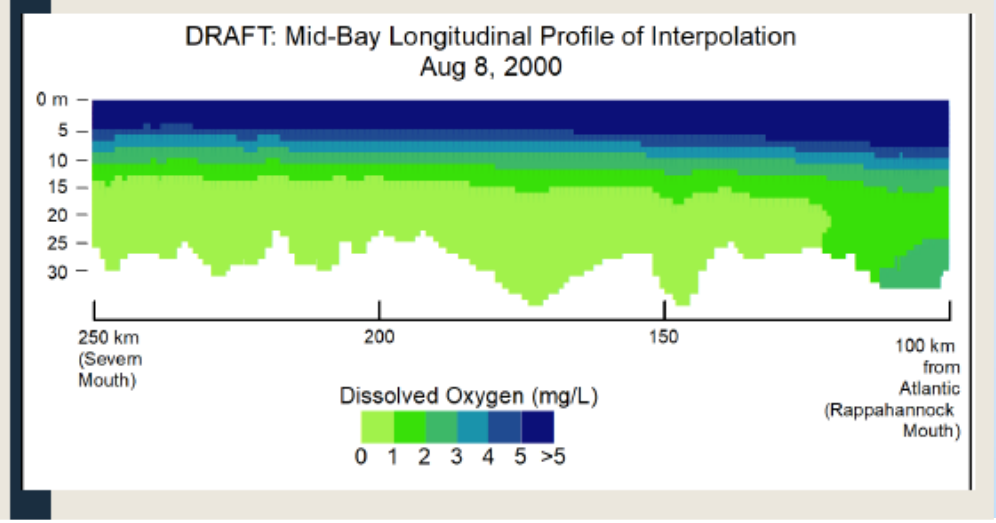
h = hour 1:24

- **EOTB, VECOS, and NOAA high frequency ConMon and Vertical array data, hourly subset**
  - *Subset to hourly only if collected at higher frequency*

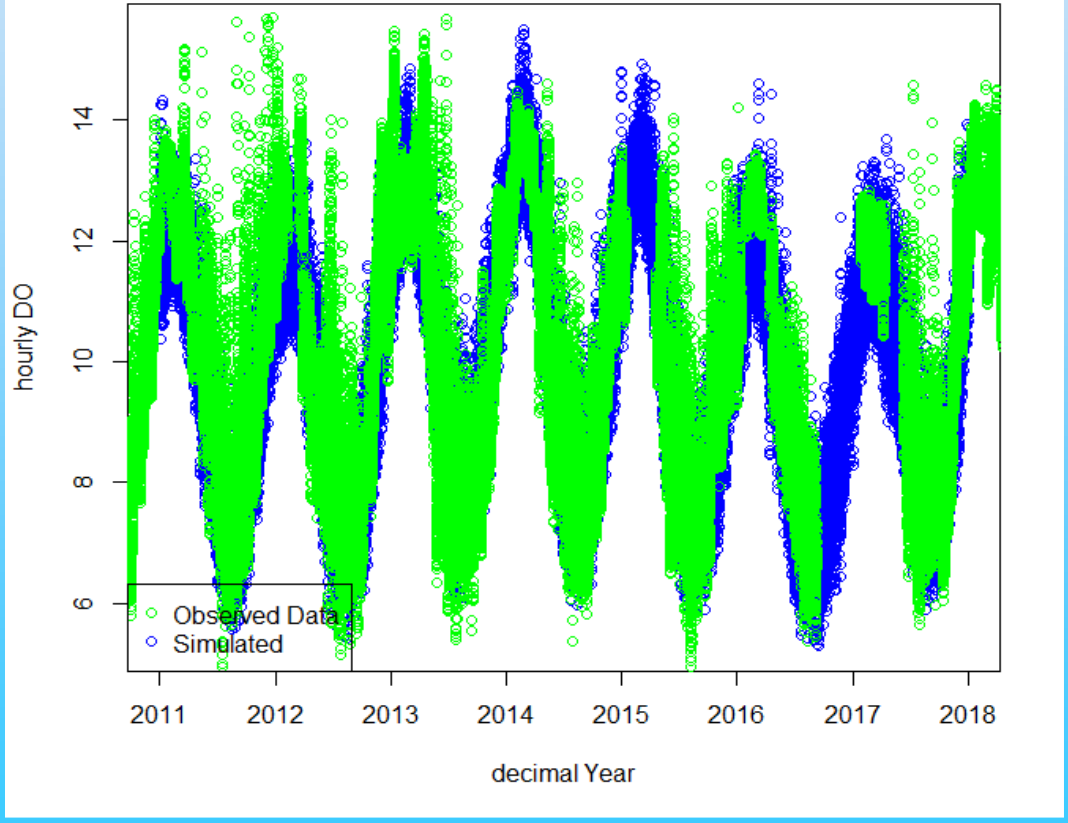
Tango and Murphy, 2024

Examples of the 4D interpolation

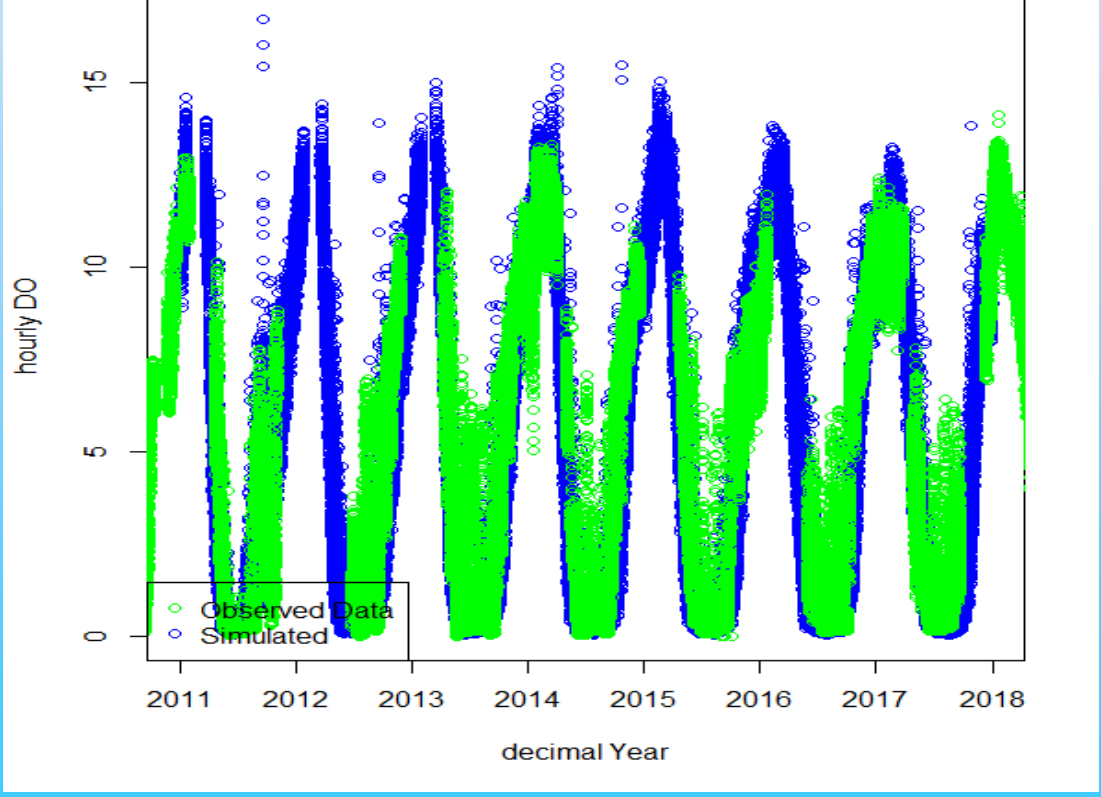
(Elgin Perry, 2023)



Surface



Bottom



1. Should the 4D interpolation be applied to modeling scenarios?

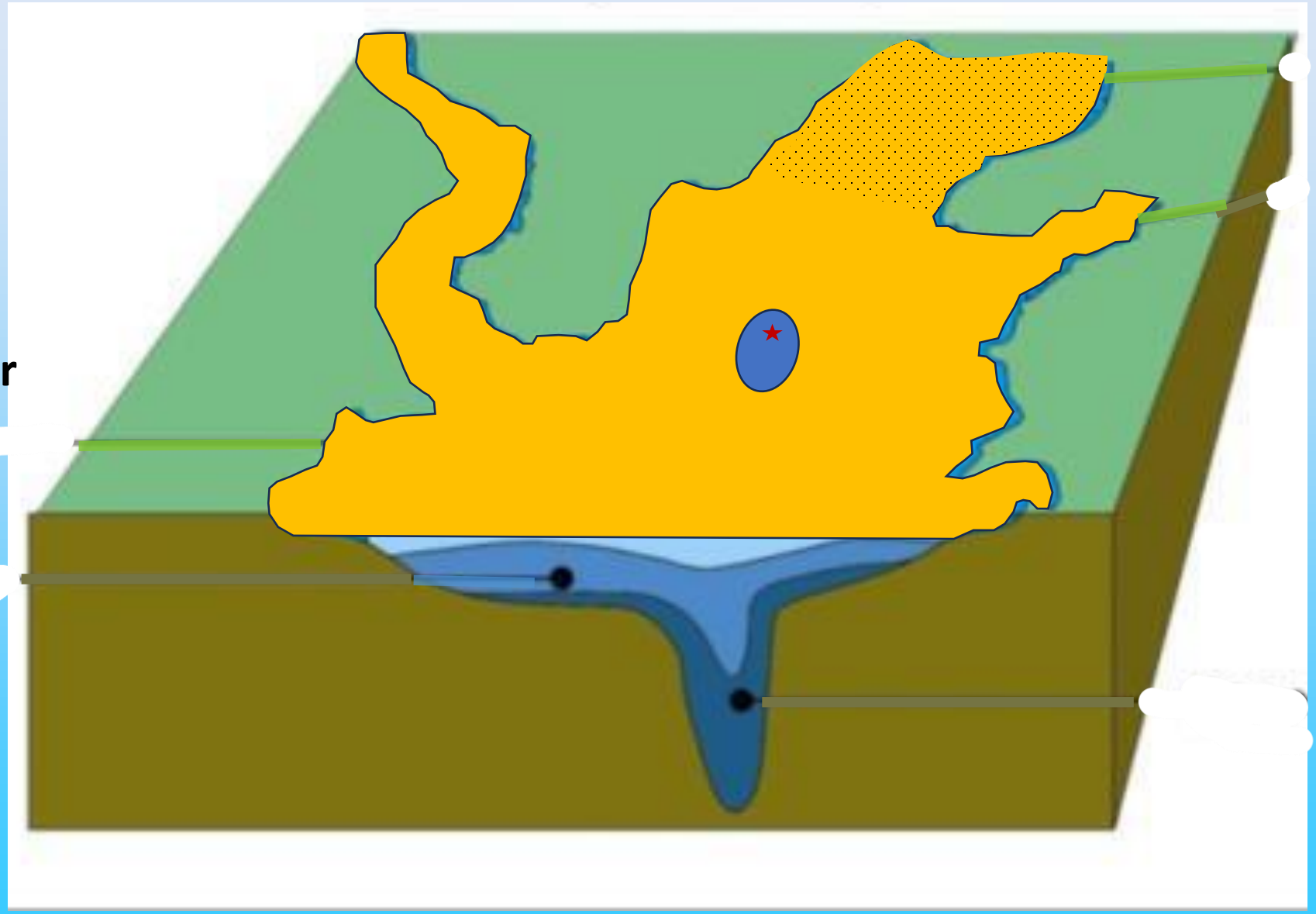
- First thought is yes.
- But the nutrient reduction allocation can be changed due to change in methodology.
- Millions of sensor-based data to be scenarioed (modification based on model sensitivity).
- Alternatively, modifying the 4D interpolator prediction also involved millions of data (76255 surface cells and 296224 cells in total).

2. Can anything be done with the 3D interpolator?

Including all DATAFLOW data in the interpolation

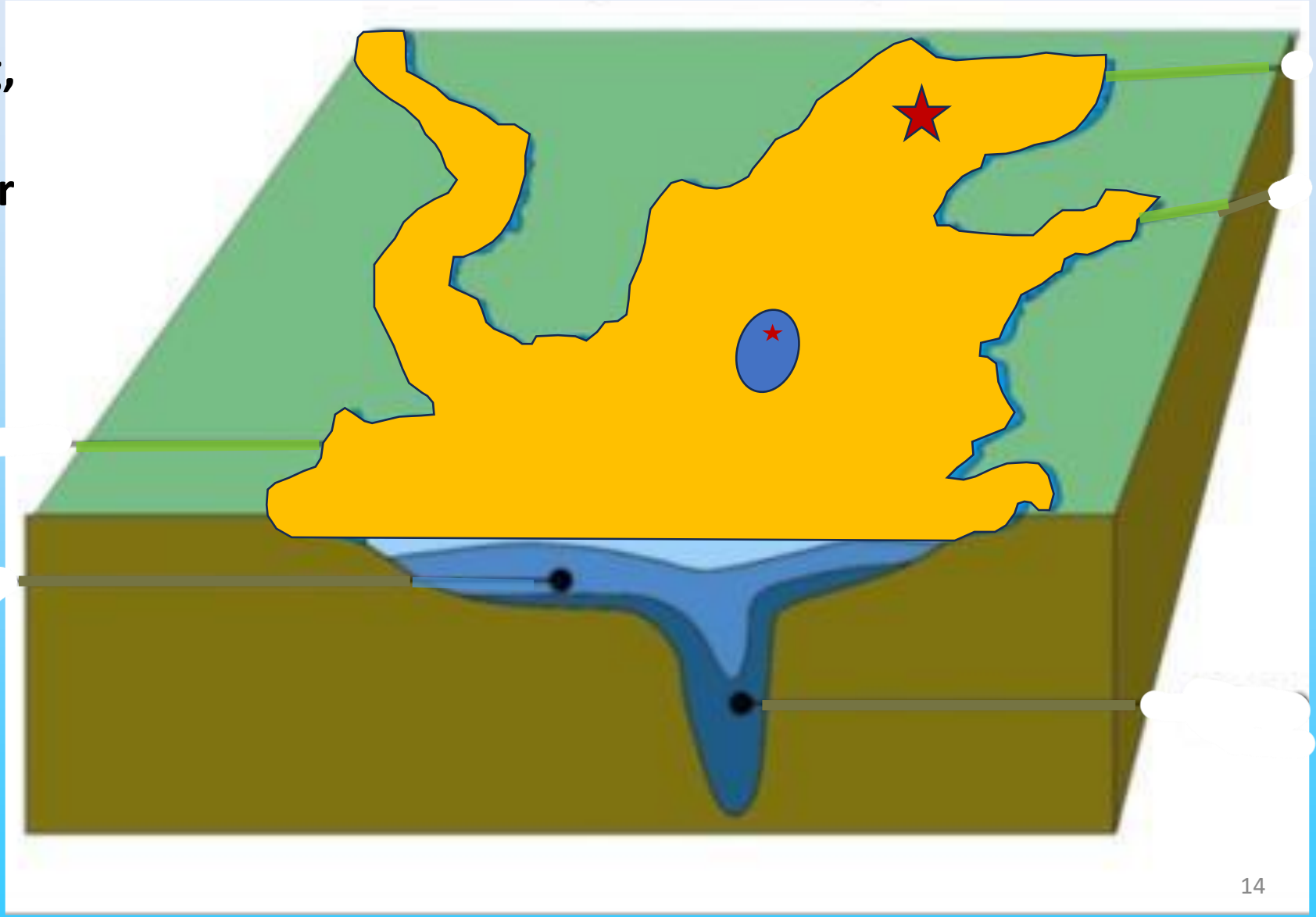
Hypothetic segment

- DATAFLOW dominating, with transitional zone
- Only in the surface layer
- May need subsampling or binning
- May cause variability



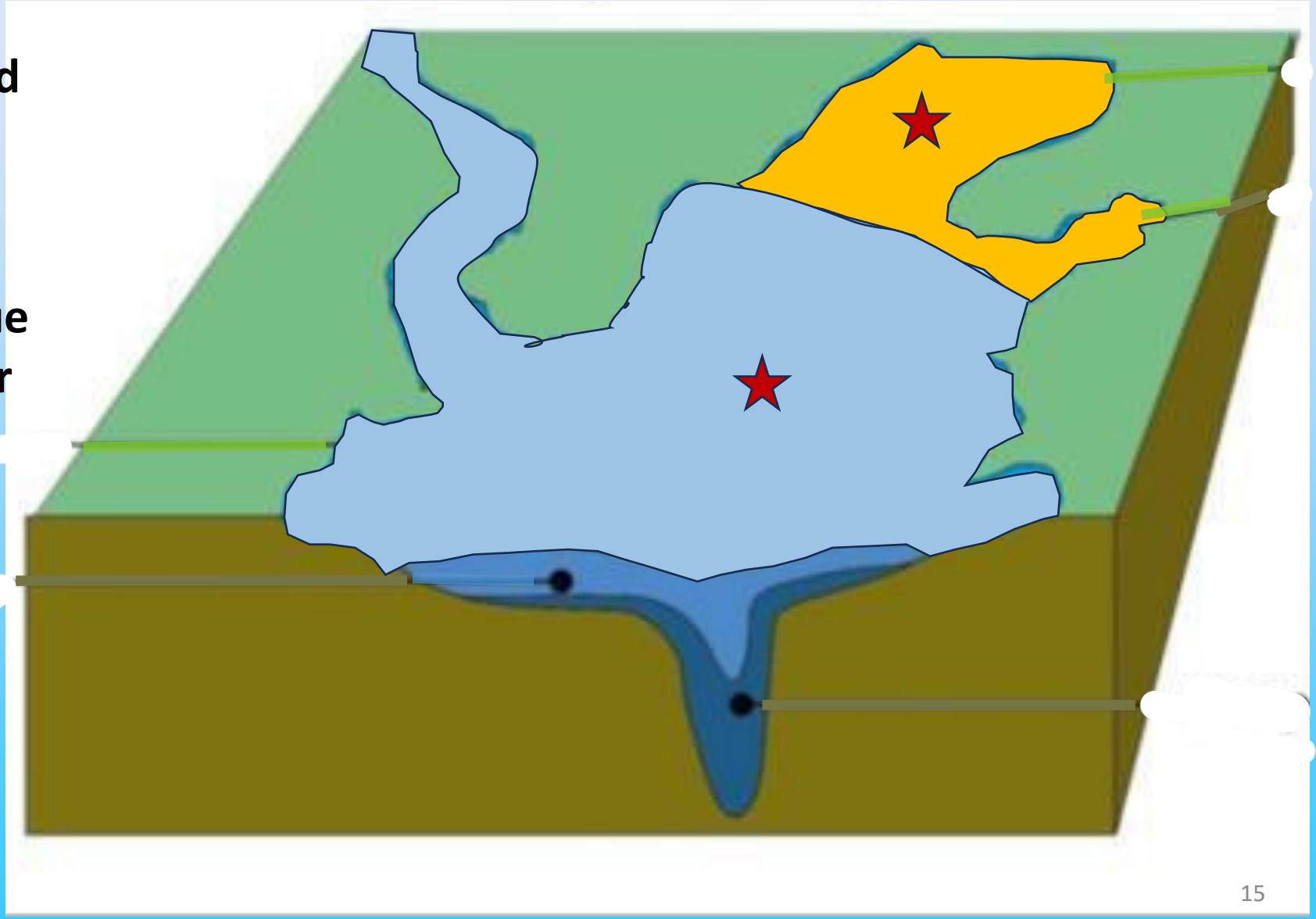
Including all CMON data in the interpolation

- CMON data dominating, with transitional zone
- Only in the surface layer if there is only one sensor
- Limited to depth of the deepest sensor.
- May need subsampling or binning
- May cause variability, less than dataflow



Including biweekly average CMON data in the interpolation

- CMON data more limited locally
- Limited to the deepest sensor depth
- May cause variability due to sensor interruption or rotation.



Summary questions

- Should the 4D interpolator be used for modeling scenarios?
  - Management concern
  - Technical concern

- How high frequency sensor data be included in the 3D interpolator and the CFD-based criteria assessment?
  - Representativeness of CMON stations and dataflow samplings
  - Variability resulted from rotational deployment and discontinuity in data collection.