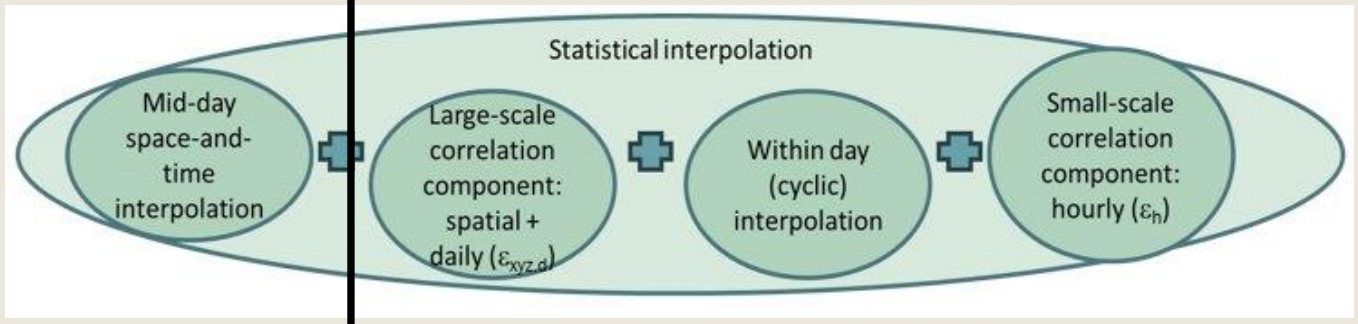
Update on: Segment Interpolation Regions

Bay Oxygen Research Large Group
Aug 18, 2025

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

Focus on this part of the tool

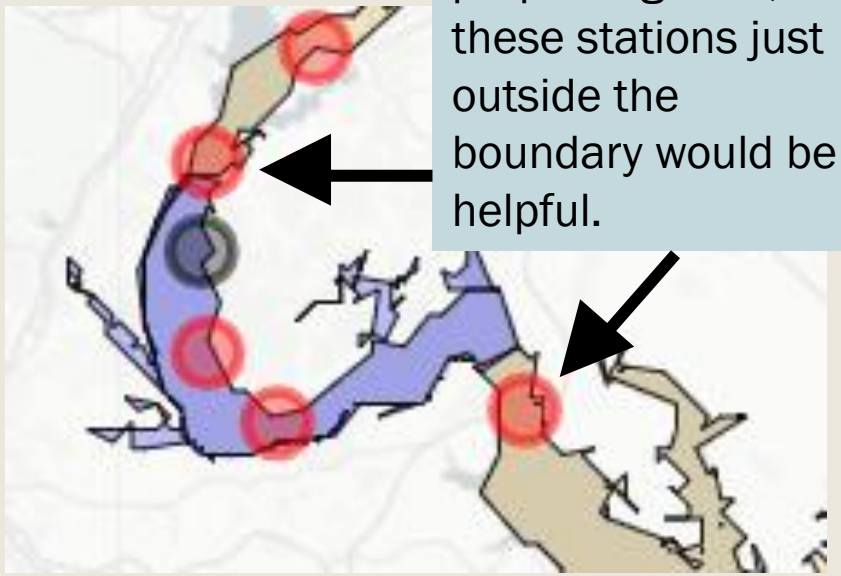


Background

- Current 3-D interpolator uses boundary data outside the segment of interest for interpolations at the edges of segments.
- A similar need exists for the 4-D interpolation for good edge interpolations (to avoid extrapolation).

Background

- Current 3-D interpolator uses boundary data outside the segment of interest for interpolations at the edges of segments.
- A similar need exists for the 4-D interpolation for good edge interpolations (to avoid extrapolation).



Example: If we want to interpolate in the purple segment, these stations just outside the boundary would be helpful.

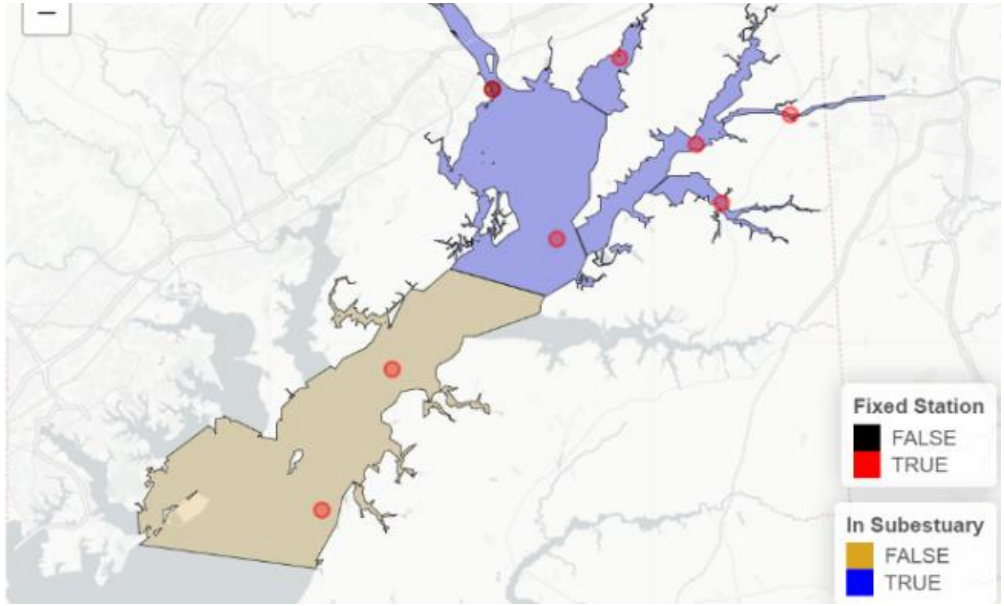
- In addition, the 4-D method requires sufficient data to build the smooth relationships between space, time, and DO. Data from one segment alone is not normally enough.

For more details, see document SegmentGrouping4D_8July2025 which includes maps for 31 currently selected regions and testing results.

Examples:
Segment Interpolation Regions

- One or more segments are grouped together. Each region also includes all touching segments as boundary. Those boundary segments will also be in their own separate region(s).
- Including entire segments as boundary help “future proof” the approach as stations may be added or removed in coming years.

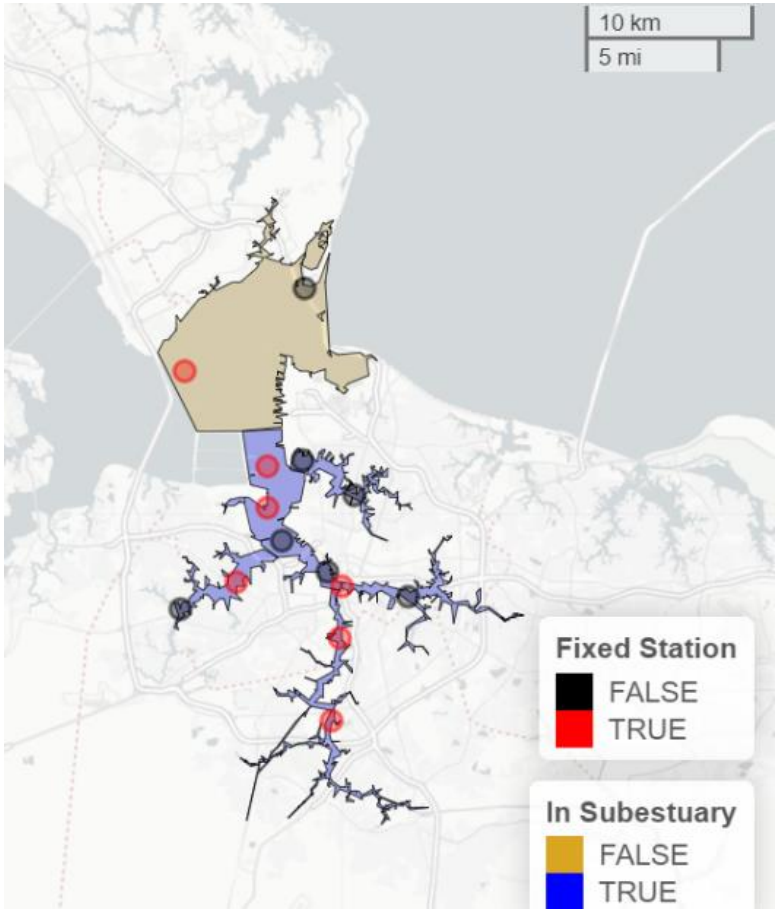
CB1_upE:
BOHOH, C&DOH_DE, C&DOH_MD, CB1TF, ELKOH, NORTF
Boundary: CB2OH



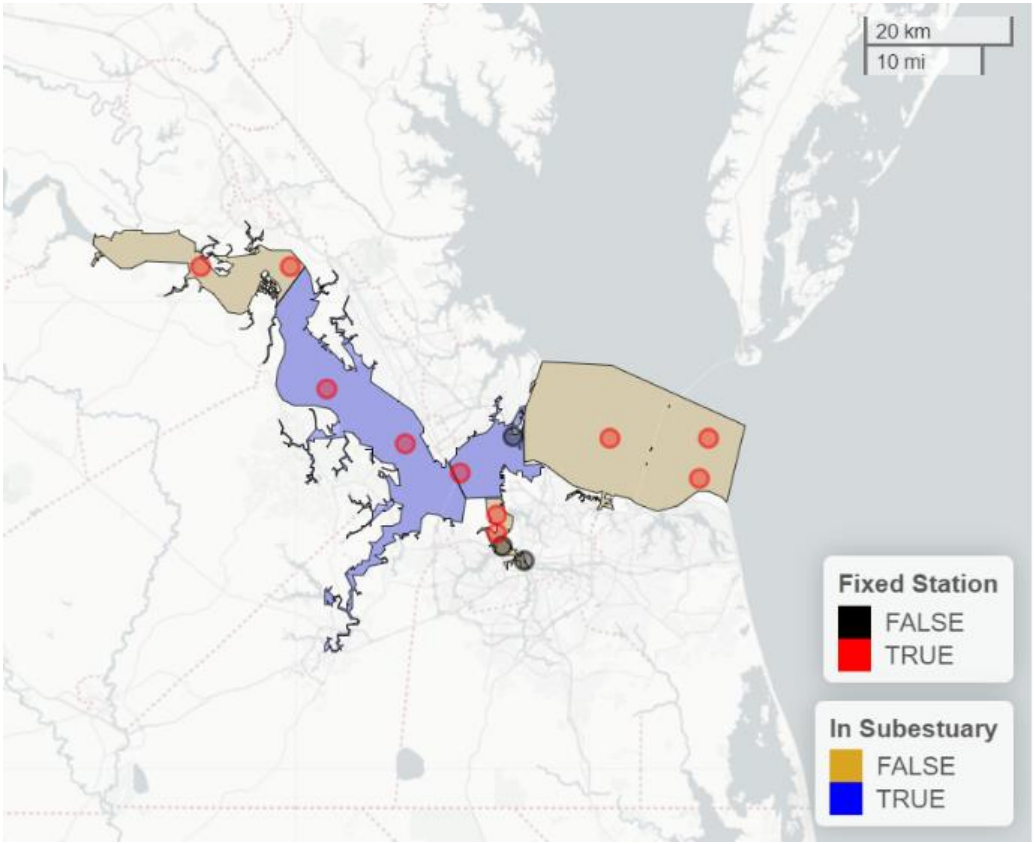
CB2_tribs:
BACOH, BSHOH, CB2OH, GUNOH, MIDOH, SASOH
Boundary: CB1TF and CB3MH

Examples:
Segment Interpolation Regions

ELZ
ELIPH, LAFMH, EBEMH, SBEMH, WBEMH
With boundary: JMSPH



lowJMS
JMSPH, JMSMH
With boundary: CB8PH, ELIPH, JMSOH



Details

- The regions are identified in an input spreadsheet to the 4-D tool, and adjustments are possible as we continue to test.

- Adjustments in the future will be possible too, especially if there are changes to the monitoring program.

- To get to this selected set of Segment Interpolation Regions, multiple options were evaluated.

	A	B	C	D	E	F	G	O
1	objectid	cbseg_92	cbpseg_st	state	cb_303d_segment	seg_int_reg	bndy_data	seg_int_reg_ID
2	11	CB2OH	CB2OH_MD	MD	Upper Chesapeake Bay	CB1_upE	TRUE	1
3	10	CB1TF	CB1TF_MD	MD	Northern Chesapeake Bay	CB1_upE	FALSE	1
4	31	ELKOH	ELKOH_MD	MD	Elk River	CB1_upE	FALSE	1
5	9	C&DOH_MD	C&DOH_MD	MD	C&D Canal, MD	CB1_upE	FALSE	1
6	8	C&DOH_DE	C&DOH_DE	DE	C&D Canal, DE	CB1_upE	FALSE	1
7	6	BOHOH	BOHOH_MD	MD	Bohemia River	CB1_upE	FALSE	1
8	52	NORTF	NORTF_MD	MD	Northeast River	CB1_upE	FALSE	1
9	10	CB1TF	CB1TF_MD	MD	Northern Chesapeake Bay	CB2_tribs	TRUE	2
10	12	CB3MH	CB3MH_MD	MD	Upper Central Chesapeake Bay	CB2_tribs	TRUE	2
11	75	SASOH	SASOH_MD	MD	Sassafras River	CB2_tribs	FALSE	2
12	11	CB2OH	CB2OH_MD	MD	Upper Chesapeake Bay	CB2_tribs	FALSE	2
13	4	BACOH	BACOH_MD	MD	Back River	CB2_tribs	FALSE	2
14	44	MIDOH	MIDOH_MD	MD	Middle River	CB2_tribs	FALSE	2
15	33	GUNOH	GUNOH_MD	MD	Gunpowder River	CB2_tribs	FALSE	2
16	7	BSHOH	BSHOH_MD	MD	Bush River	CB2_tribs	FALSE	2
17	11	CB2OH	CB2OH_MD	MD	Upper Chesapeake Bay	CB3	TRUE	3
18	13	CB4MH	CB4MH_MD	MD	Middle Central Chesapeake Bay	CB3	TRUE	3
19	24	CHSMH	CHSMH_MD	MD	Lower Chester River	CB3	TRUE	3
20	53	PATMH	PATMH_MD	MD	Patapsco River	CB3	TRUE	3
21	41	MAGMH	MAGMH_MD	MD	Magothy River	CB3	TRUE	3
22	12	CB3MH	CB3MH_MD	MD	Upper Central Chesapeake Bay	CB3	FALSE	3
23	53	PATMH	PATMH_MD	MD	Patapsco River	PAT_MAG	FALSE	4
24	41	MAGMH	MAGMH_MD	MD	Magothy River	PAT_MAG	FALSE	4
25	12	CB3MH	CB3MH_MD	MD	Upper Central Chesapeake Bay	PAT_MAG	TRUE	4

Evaluation

Some alternatives considered:

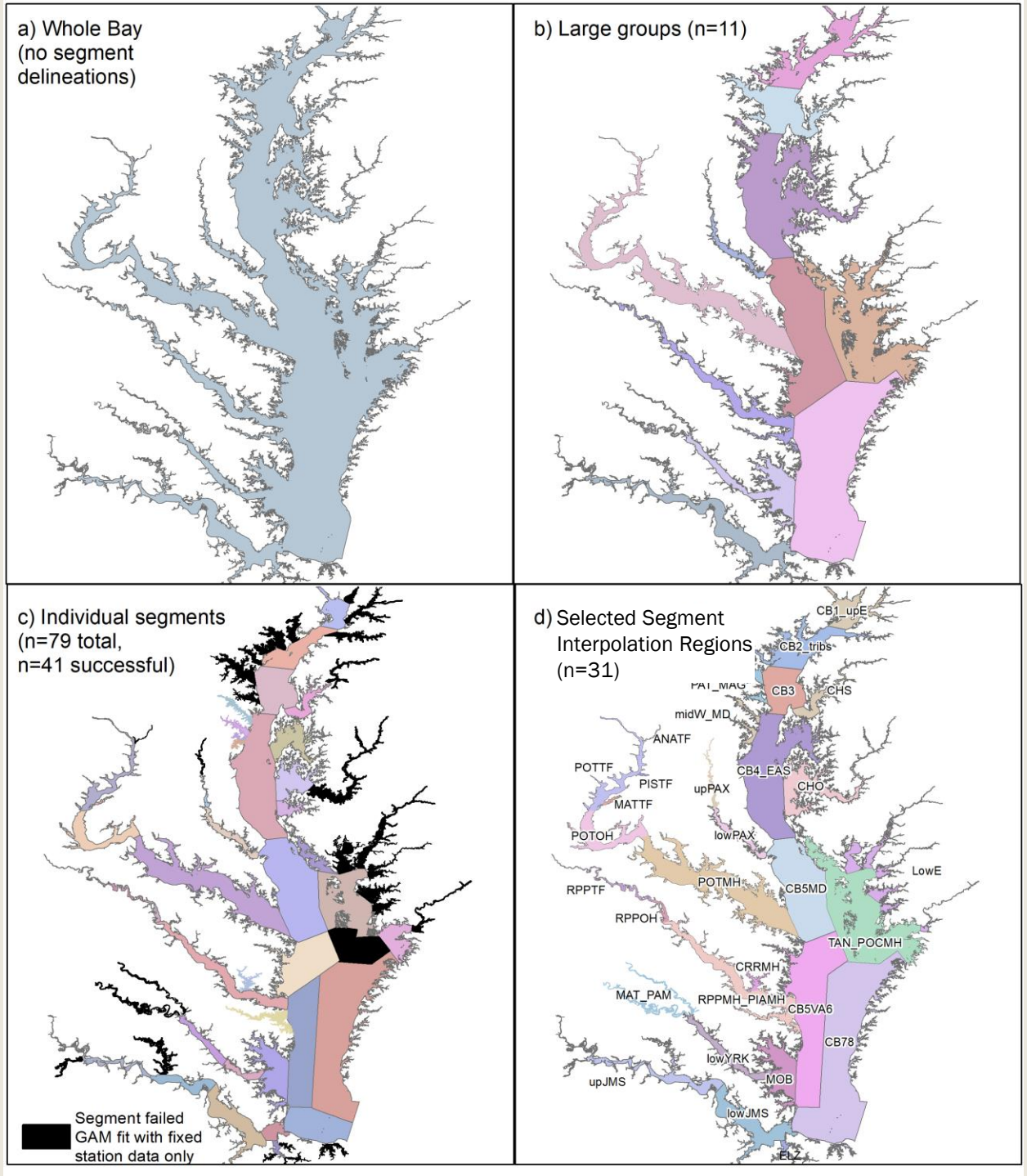
A. Whole bay together, no segment delineations.

B. 11 relatively large segment groups.

C. 79 individual or small segment groups *(failures indicated with black).*

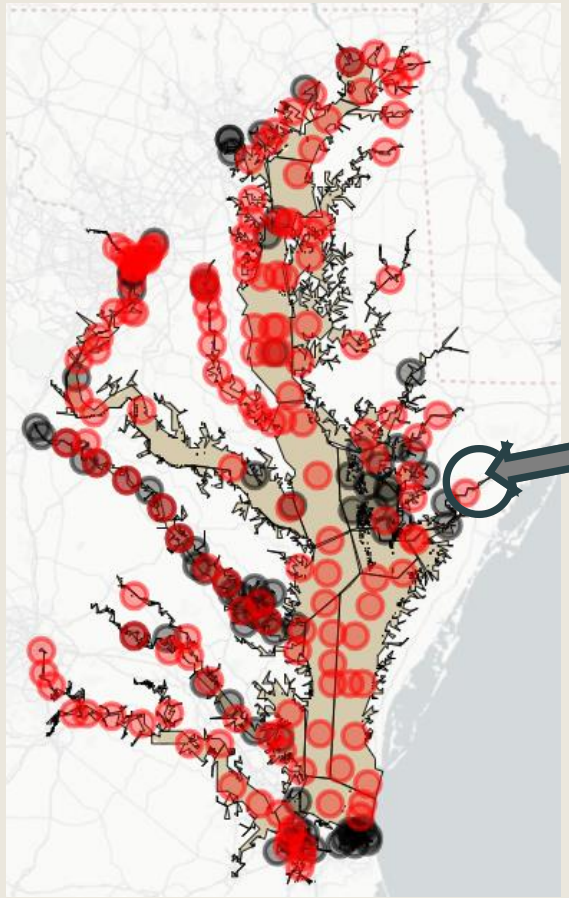
D. 31 selected Segment Interpolation Regions–small, but balancing need to have enough data.

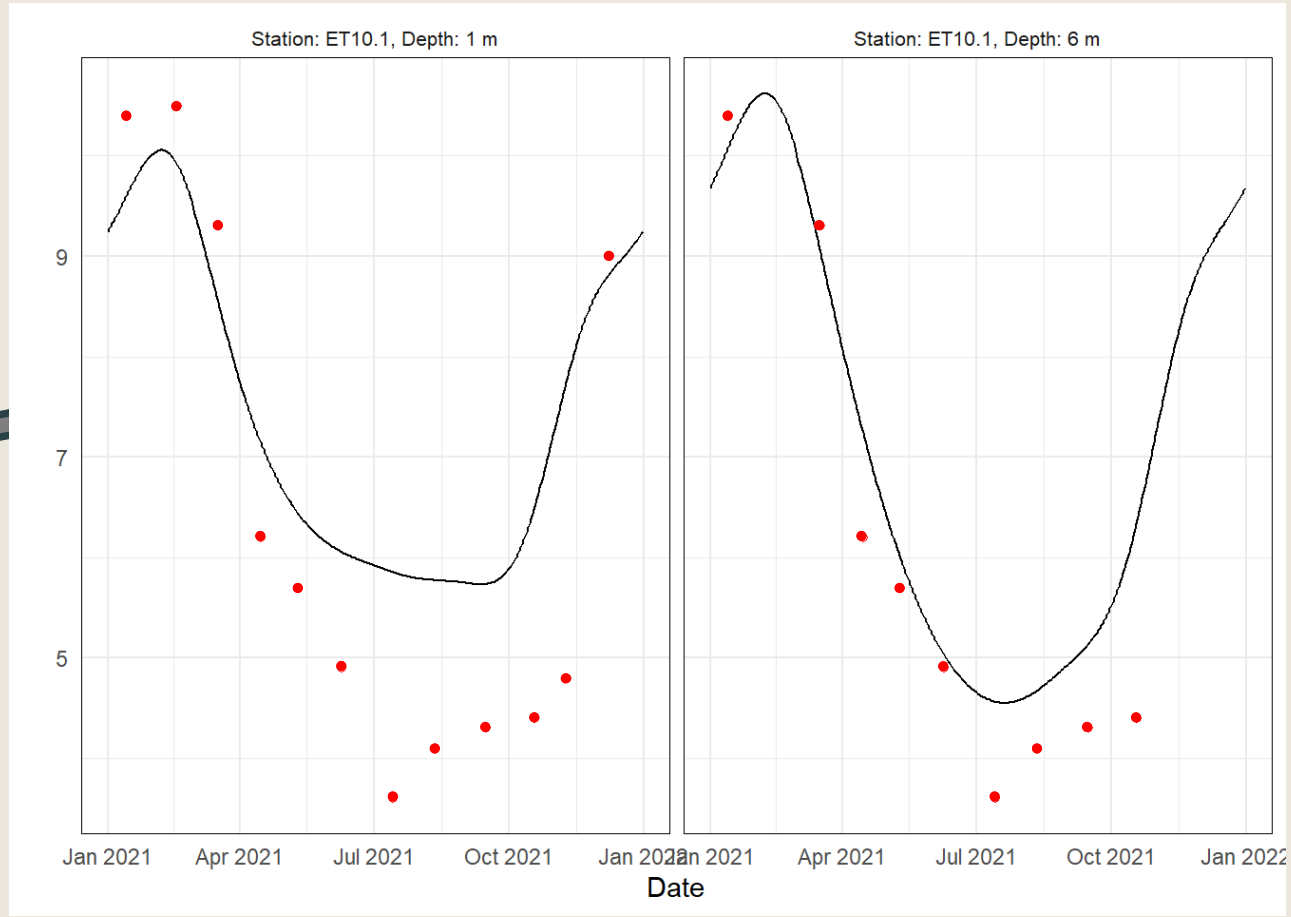
Note: in all options except A, adjacent segments are used as boundary.



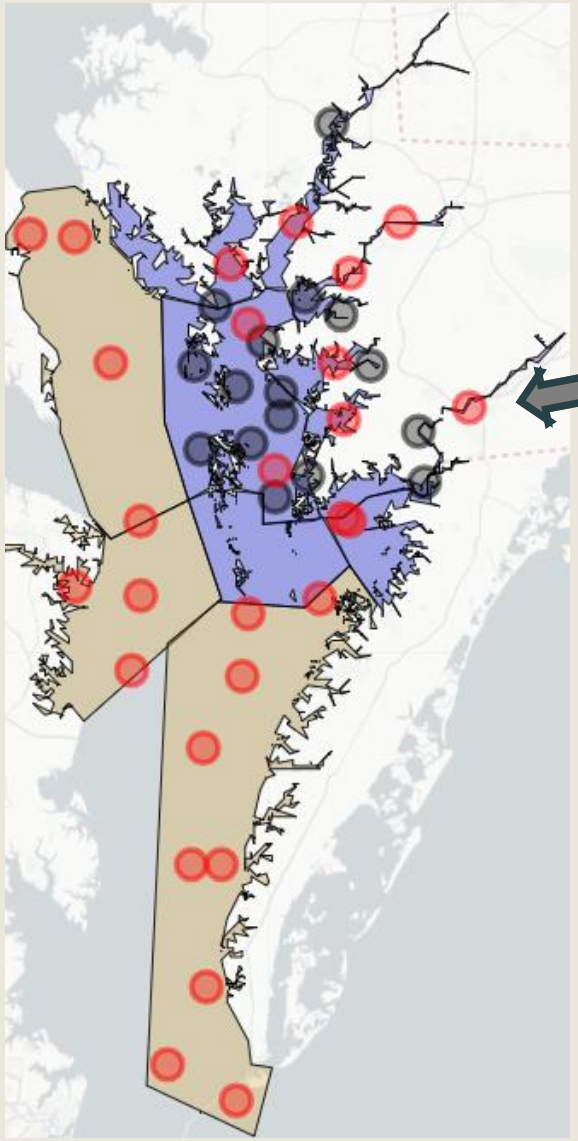
Example: impact at one station of region size

A. Whole bay together: results in estimates at ET10.1 clearly mis-matched to the data

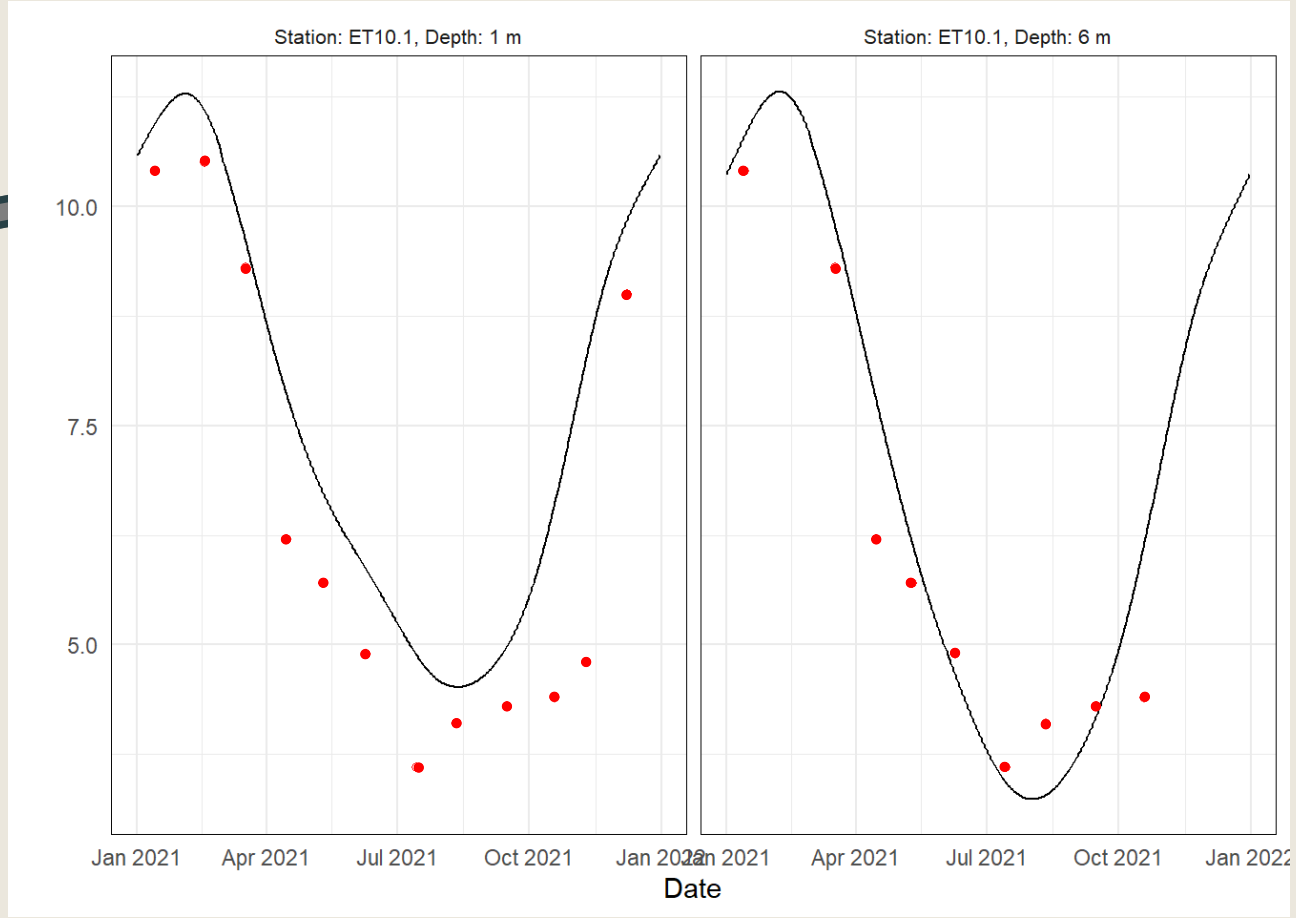




Example: impact at one station of region size

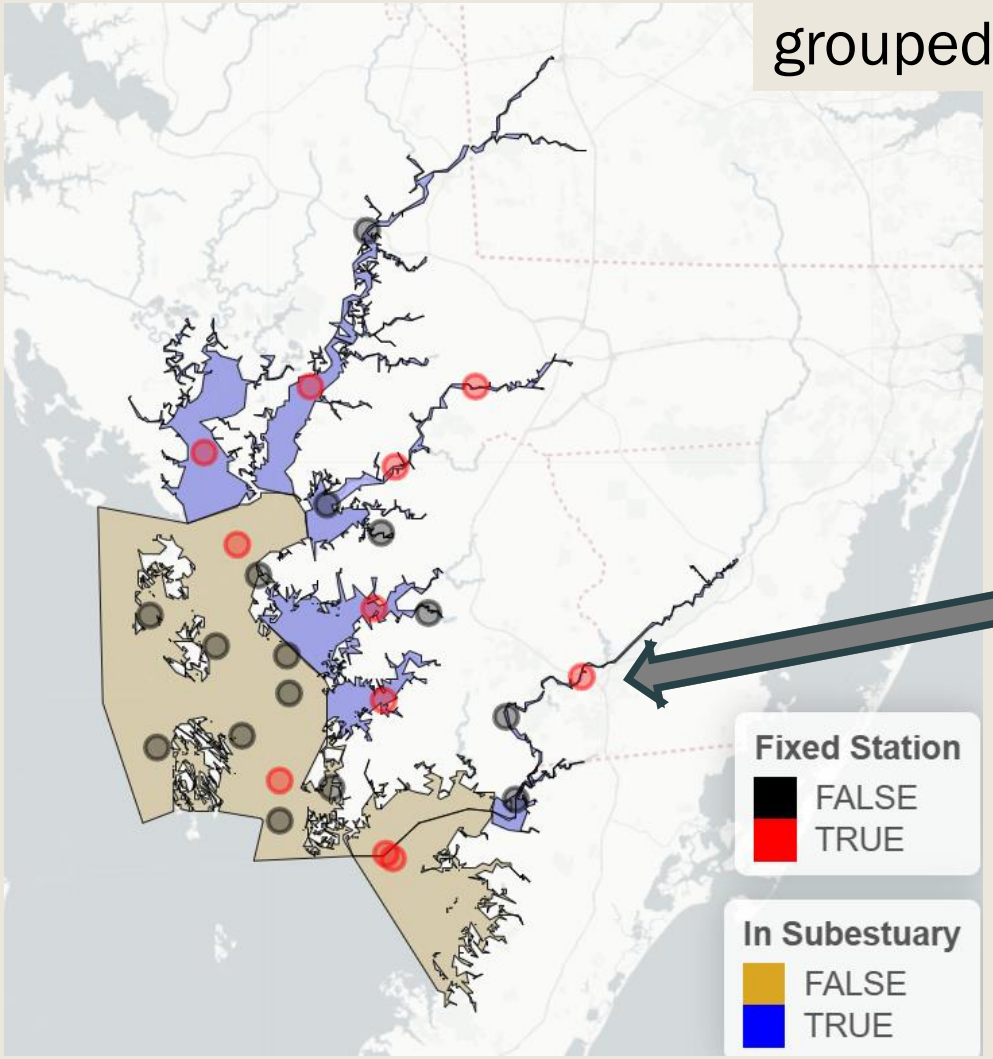


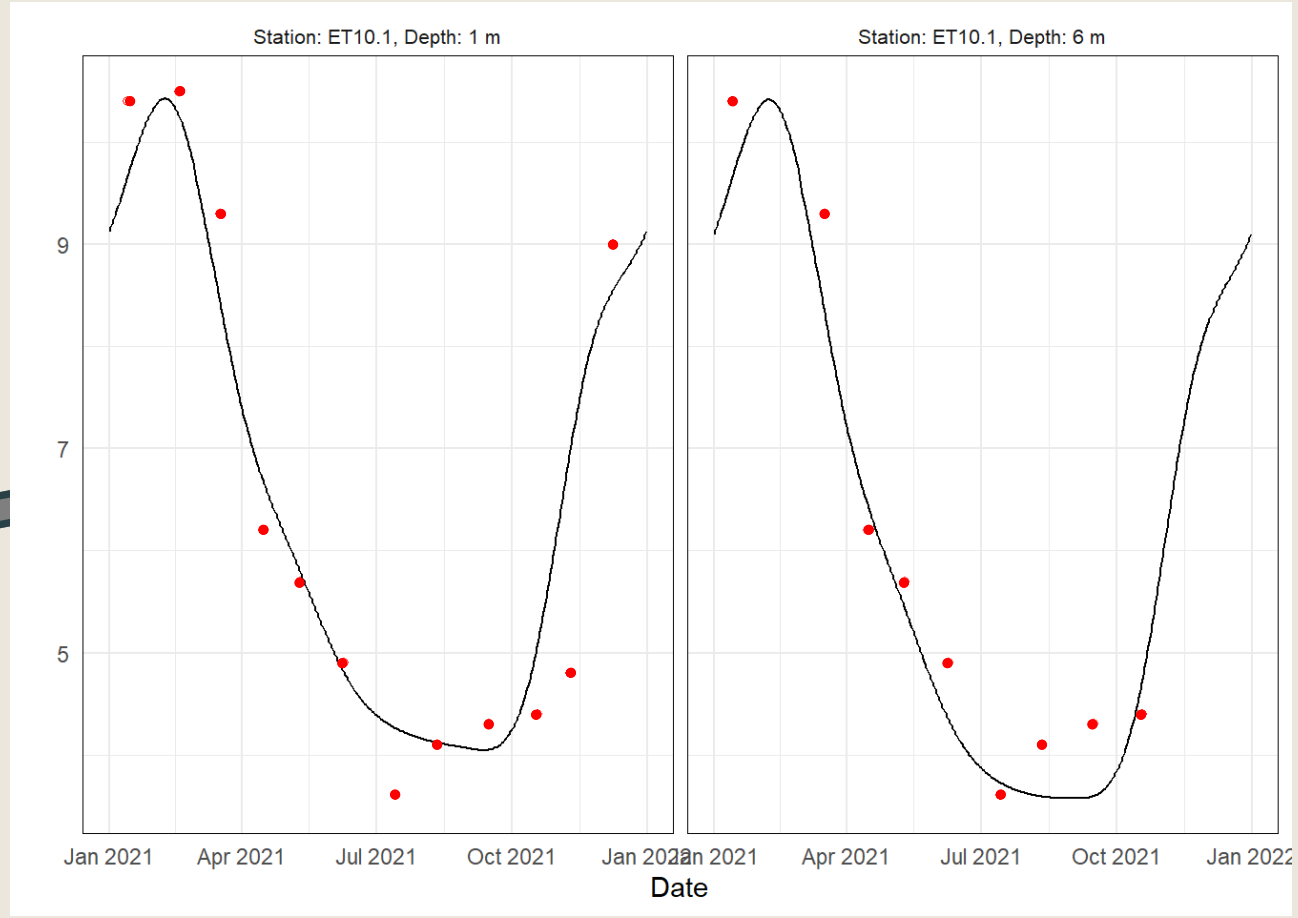
B. Large groups – segment is with Tangier and other eastern shore tribs, but also has boundary in main bay



Example: impact at one station of region size

D. Selected Seg. Interp. Regions: Eastern shore tributaries grouped together, with only nearby boundary



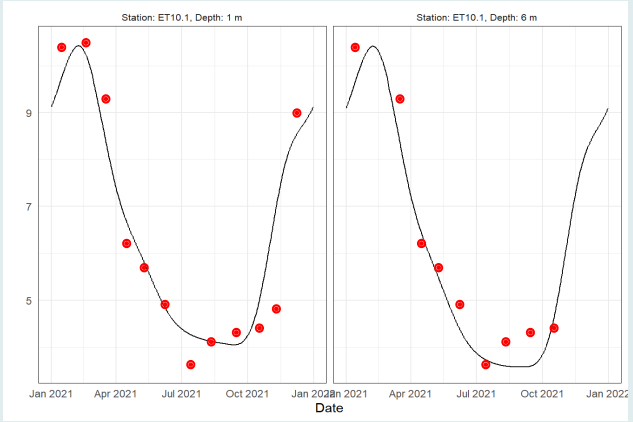


Computing Root Mean Squared Error (RMSE) difference

ET10.1 in 2021

D. Selected Seg. Interp. Reg.

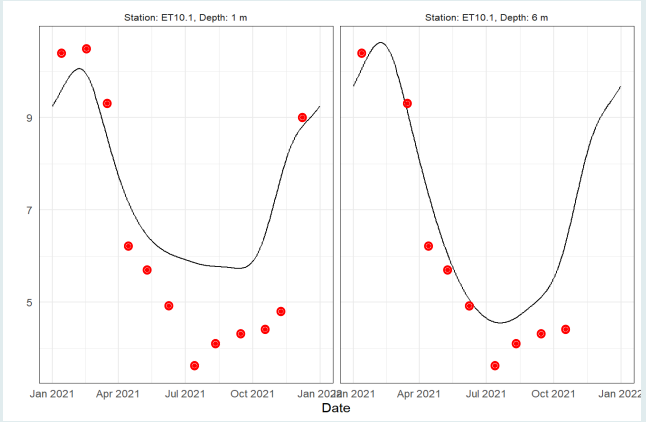
RMSE 0.69



minus

A. Bay-wide GAM

RMSE 1.4



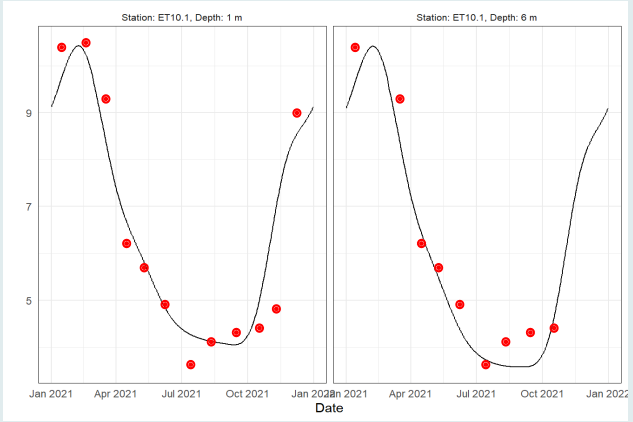
= -0.71

Computing Root Mean Squared Error (RMSE) difference

ET10.1 in 2021

D. Selected Seg. Interp. Reg.

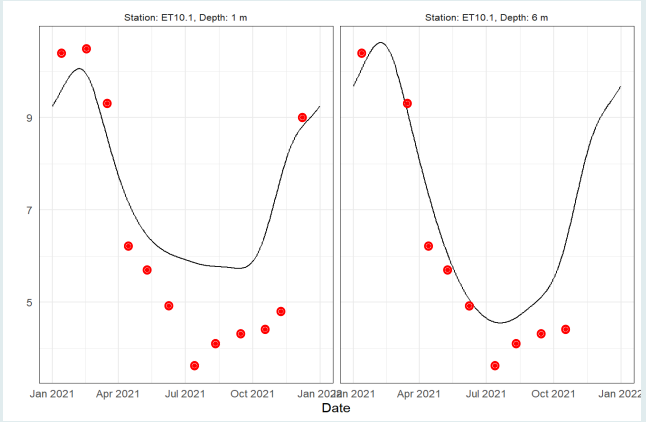
RMSE 0.69



minus

A. Bay-wide GAM

RMSE 1.4



= -0.71

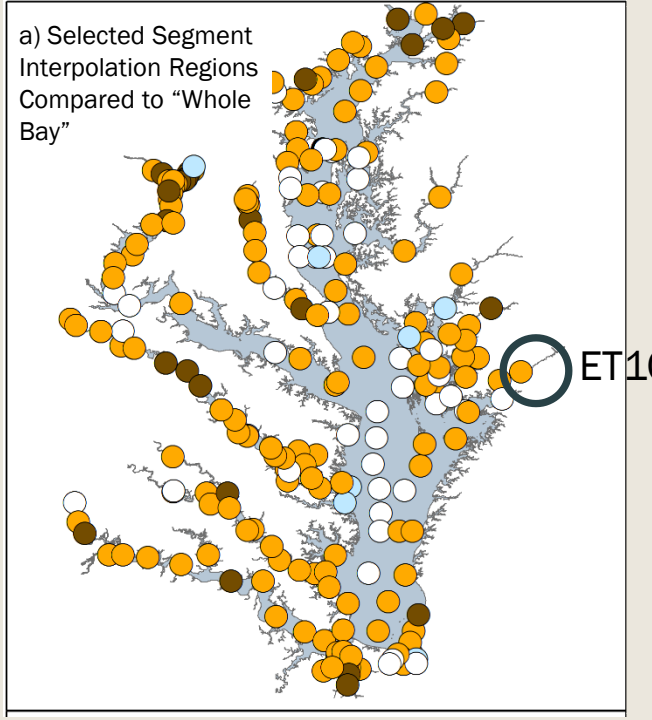
Each station's RMSE difference indicated with a symbol:

- Brown and orange = The selected 31 Seg. Interp. Regions are LOWER error
- White = almost the same
- Blue = Selected Regions are HIGHER error
- Black = GAM unable to be fit

All stations

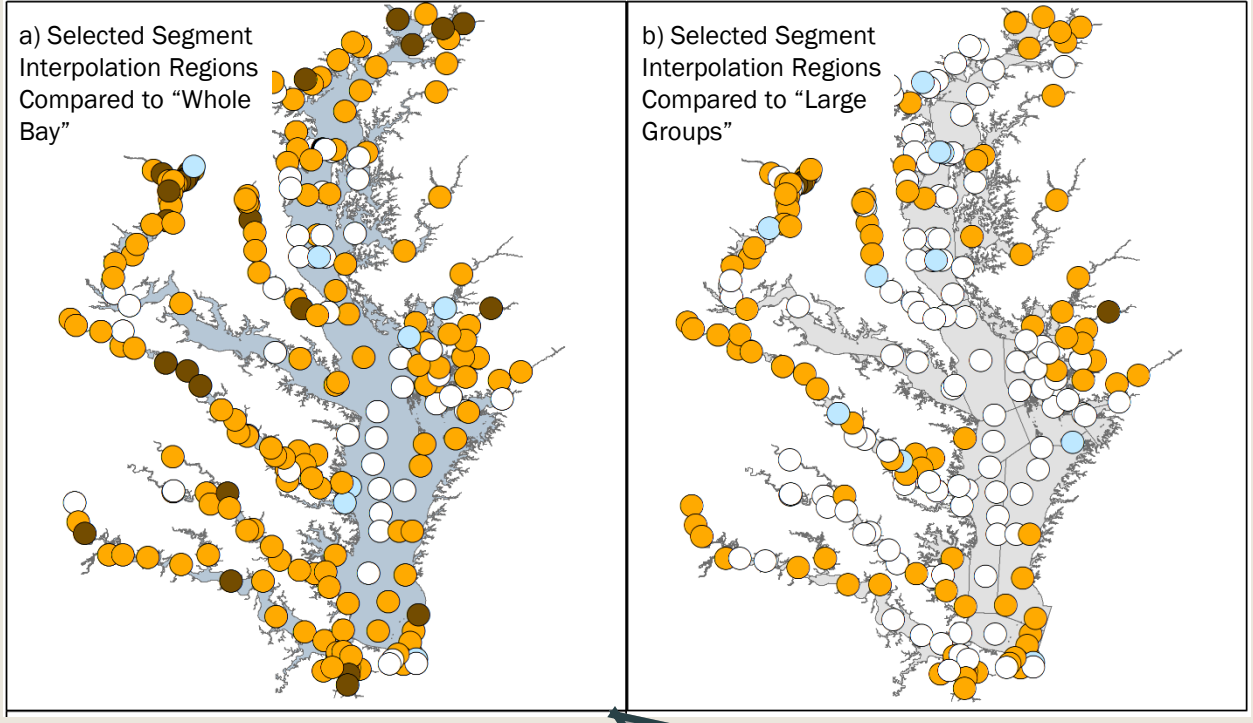
(D) Selected 31 Seg. Interp. Regions **—** (A) Whole bay together **=**

Majority of stations have better interpolation results with smaller regions.



- Brown and orange = The selected 31 Seg. Interp. Regions are LOWER error
- White = almost the same
- Blue = Selected Regions are HIGHER error
- Black = GAM unable to be fit

All stations



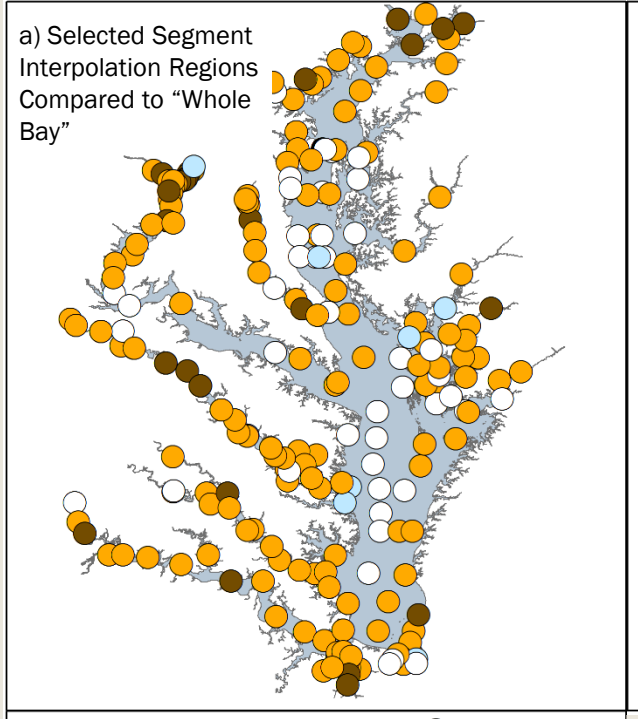
(D) Selected 31 Seg. Interp. Regions — (B) 11 Large groups =

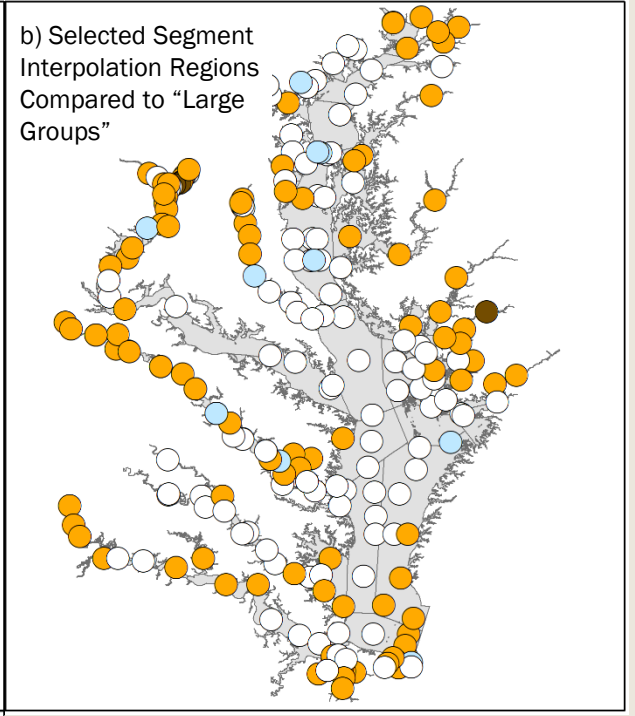
- Brown and orange = The selected 31 Seg. Interp. Regions are LOWER error
- White = almost the same
- Blue = Selected Regions are HIGHER error
- Black = GAM unable to be fit

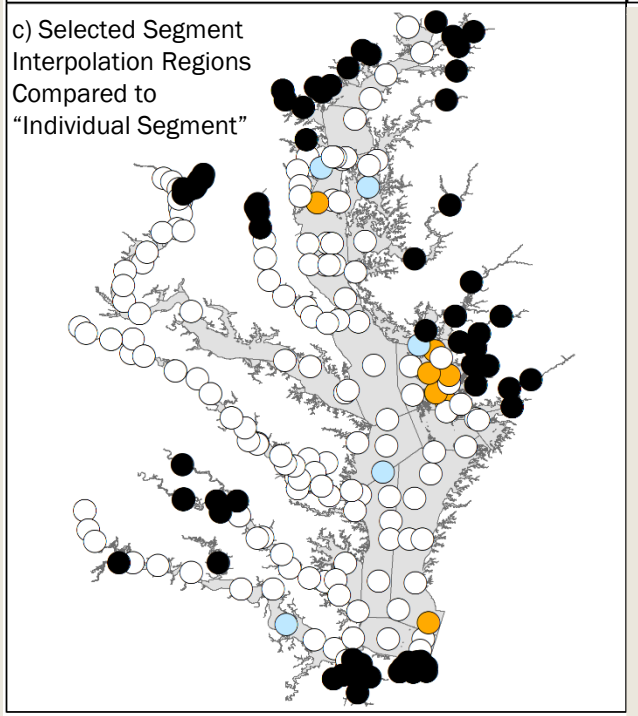
All stations

D) Selected 31 Seg. Interp. Regions — (C) 79 Individual Segments

- Brown and orange = The selected 31 Seg. Interp. Regions are LOWER error
- White = almost the same
- Blue = Selected Regions are HIGHER error
- Black = GAM unable to be fit







Conclusion: Small regions are best, but they need to be large enough to fit the interpolation equation.

Summary/discussion

- Evaluation of multiple options resulted in current set of 31 Segment Interpolation Regions. Regions contain on average 3 segments grouped together (range =1 to 11), plus boundary.
- Possible discussion
  - *Impact of needing data across state lines.*
  - *The tool will identify if an individual segment does not have data in a year, since in that case it should not be assessed.*