Bay Oxygen Research Group (BORG)

Monday, September 12th, 2022
12:00 PM – 1:30 PM

Meeting Materials Link

This meeting was recorded for internal use to assure the accuracy of meeting notes.

ACTIONS

- ✓ **Hypoxia Collaborative:** Discuss challenges of previous Rappahannock profiler and how to handle challenges with boats at hypoxia meeting.
  - **ACTION detail: Ensure that we understand places not conducive to placing such fixed sampling gear so that we can select suitable locations for extended deployments of new vertical array continuous monitoring sites**
- ✓ **Water Quality Goal Implementation Team (WQGIT) and Modeling Workgroup:** Consider the Rappahannock for inclusion in the Multiple Tributary Models.
  - **ACTION detail: Include the Rappahannock in the MTM selection process**
- ✓ **BORG topic for next meeting:** At next small working group discuss solutions for technological limitations, utilizing the cloud resources to expand computing capacity, etc. and find solutions to the computational limits experienced.
- ✓ **BORG topic for next meeting:** Discuss incorporating impact of daily flow observations in some regions.
- ✓ **BORG topic for next meeting:** How do we get from multiple simulations to criteria assessment? What's the difference between variability and uncertainty?
  - **ACTION detail: Begin documentation of the proposed assessment method, decision framework needs. This information will eventually need to be published in an Environmental Protection Agency (EPA) Technical Addendum for Chesapeake Bay water quality criteria attainment assessments.**
- ✓ **All BORG members:** Please email Rebecca Murphy at rmurphy@chesapeakebay.net and Peter Tango at ptango@chesapeakebay.net with any additional comments and questions.

MINUTES

12:00 PM Introduction to the proposed method for the 4-Dimensional Interpolator tool – *Gary Shenk (USGS)*

- Gary Shenk: as a reminder, the motivation for the 4-D interpolator is that there are 92 segments with up to 11 designated uses per segment, with 838 combinations of things we should be assessing. We're only assessing 326, however. None of the segments are being assessed for all designated uses. Our thought for assessing all the designated uses

for all time periods is we'll create a 3D representation of water quality through time (the 4th dimension). Specifically, we'll estimate what's happening on March 1st 1993 at midnight (which is the first hour of the Total Maximum Daily Load [TMDL] criteria), and estimate what happens at 1am, all the way to a more current date at 11pm. We'll do estimates for every hour until now, and use that to make our assessment of water quality.

- Gary Shenk: How we'll do this: we start with observed data. We developed models. We use Generalized Additive Models (GAMs) which make a mean daily prediction at all space in time. The next step, which we haven't started yet, is to take those mean daily predictions and use some sort of statistical simulation or other method that will come up with hourly realizations - multiple hourly realizations. This is so we can find the probability of meeting different designated uses. From those multiple hourly realizations we'll do our already standardized set of water quality assessments from those. Then we can have attainment of all criteria and all designated uses assuming all of this works. That is our goal.
- Gary: We're starting with the observed data, and we want to look at all the available observed data we have. When we're talking about data, we're not just talking about the data we're trying to predict (dissolved oxygen, salinity, density, chlorophyll) – we also need things that will predict those things such as salinity, density inflows, wind, tide, etc. Using those observed data Elgin will show where he's gotten creating GAMs that can produce that mean daily prediction. We can't yet produce it everywhere and everywhen, but we're making progress. We're looking at daily predictions for dissolved oxygen (DO), clarity and chlorophyll. In order to get designated uses for oxygen we need to produce the pycnocline and we need salinity and temperature for that.
- Elgin Perry: The statistical simulation will give us results that come out in terms of probabilities, and yet water quality assessment usually has some kind of bright line criteria of pass/fail. Do you have any thought how we'll get from one to the other?
- Gary Shenk: I thought the idea of this statistical simulation came from our conversation. What I was picturing is that in a statistical simulation you use some sort of random number generator to pick numbers that tell you where you are on this distribution. You would take your mean daily prediction and assign a diel curve to it if it's oxygen and you would have some probability of where the actual value is. You would pick values off the distribution randomly and call that one simulation. And you would do that multiple times and have multiple realizations of what the bay could be. What you're saying is before we get to that stuff we're actually producing probabilities that vary through time. If so, we need to figure out how to do a water quality assessment using those things. If that's where you're going that does seem like a computationally less expensive way to do that but whether or not that fits with water quality criteria, that's a good question.
- Elgin Perry: Let's say you do a thousand of these simulations and you find that 10% of them show that there's been a violation in a certain period of time. Are we going to say

that means those segments fail to meet those criterion or do we say that 10% probability is not enough to say it's failing, we need to see it over 50%? I see the simulation as producing the frequency we interpret as a certain number of failures out of a total number of trials. For each simulation you have a bright line criterion. But for collective there's still not a bright line criteria.

- Gary Shenk: I was thinking that here is a 30 day mean curve. A 30 day mean with 4 months and 3 years has 12 dots in it. We have 12 + 2 dots = 14 dots. We do that because we only have one estimate of the 30 day mean for each of designated use month and year. But if we had 100 simulations, then I was just thinking we put them all into the same distribution. We're increasing the number of dots on the line.
- Elgin: The bright line is still the cumulative distribution function?
- Gary: It's essentially generating a probability and then we're saying it must be at least 90% probable.
- Jim Hagy: I'm thinking about this in terms of separating uncertainty from variability in our simulations. Variability is real fluctuations that affect what percentage is below a particular concentration. The more variability you have, the more data points may be lower or higher relative to the mean prediction. Uncertainty doesn't mean there's more low or high values; it means that we're less confident. Those two things mean something different in terms of how we deal with them. If we're adding variability associated with diel cycles, that's variability, not uncertainty.
- Gary: That's a great thought. Would trying different numbers of simulations and seeing whether they are – actually, let's put a pin on that and come back to it.
    - ACTION: *Bay Oxygen Research Group* put this on next agenda

12:15 PM Update on the Real-Time Vertical Hypoxia Monitoring Program – *Peter Tango (USGS)*

- Peter Tango: You may be aware that recommendations were developed to enhance the monitoring needed to address the short duration criteria and temporal density of information. This is an update on where this is progressing. As a reminder the Principals' Staff Committee (PSC) monitoring report aims to address how to take our monitoring programs from fair to good. The report is still being approved by the United States Geological Survey (USGS) but will be available soon. Since February the recommendations have been put into motion and investments have been made, particularly taking advantage of the infrastructure law money.
- Peter: The Hypoxia Collaborative recommended design includes 3 mainstem monitoring vertical arrays, based on the research of number of sensors that would help with improving estimates of hypoxic volume to compare to the model. The recommendations include lower tributaries with lack of information on fisheries and water quality side. They recommended 2 arrays in lower Potomac which is important for fish migration, and 2 arrays in the Rappahannock, looking at the hypoxia that flows in and out of both

those rivers. Since we can't cover the whole Bay, we're looking for areas that are close to meeting standards but would be helped by having information from a particular segment and its condition, and refining those estimates. It's been suggested to have 3 targeted study arrays to rotate through different areas. Maryland Department of Natural Resources (MD DNR) is organizing a fourth array on their fishing bay study area. These recommendations were included in the PSC monitoring report as recommendations to fund.

- Peter: National Oceanic and Atmospheric Administration (NOAA) had invested in and is operating 2 vertical arrays in the mainstem, near the mouth of the Choptank and a site across the Bay on the other shoulder. Those 2 have been in operation, except for one issue where one was disconnected recently, but that's being fixed. Funding was approved by EPA for the purchase of all additional 8 water quality sensor arrays. NOAA is funding the operations and maintenance for the 10 arrays. The River Input Monitoring (RIM) site monitoring continuous monitors (con-mons) have been approved for funding and is being coordinated by EPA and USGS. A lot of the information and data will support the operation of the 4-D Interpolator and other analyses for fish habitat and hypoxia monitoring.
- Peter: We are still looking at where exactly to place the instruments following the pilot, which was to get the operation and details worked out. Key drivers of location selection are key objectives with hypoxia monitoring, water quality standards attainment, criteria assessment, 4D data needs, and fish habitat assessment.
- Jim Hagy: What parameters will be monitored at the river input stations? Continuous nutrients?
- Peter Tango: The offshore sensors measure temperature, DO and salinity. On the watershed side it looks at 6 parameters: temperature, DO, conductivity, pH, turbidity and nitrate. These are all real time.
- Peter: Back in 2003 EPA said the recommended monitoring level was to place a vertical continuous water quality monitoring station at all 156 long-term water quality monitoring locations in the bay and its tidal tributaries.
- Peter: There are some near term needs to refine criteria for the decisions. We have highly uncertain, poorly monitored areas. We need to decide which places will be fixed and which will have mobile deployments and the duration of deployments. Depending on the distribution will affect how many sensors we can put on any particular buoy, looking at that vertical resolution. We're putting in a proposal for grant funding for additional assistance to put together a sampling design that will meet our key objectives with the distribution of the sites. Some of those decisions will need to be made in the very near term in order to get permits from the Coast Guard. Stay tuned on that. When we think about the ROAR proposal and shallow water zones, there's opportunity to link shallow water objectives in here and have complimentary shallow water monitoring.
- Richard Tian: What's the rationale behind the Rappahannock location?

- Peter Tango: It's one of the areas that's still affected by hypoxia that we see very little data from and have little to reference and understand the lower end of that tributary system. It's relevant to fisheries as well.
- Richard Tian: I ask because we're doing the Multiple Tributary Modeling (MTM) and Rappahannock is not part of it. I hope we can coordinate and use this data to categorize this and think about that. Secondly, what kind of sensor do you plan to use? You mentioned a couple of them but didn't mention fluorescence chlorophyll.
- Peter Tango: Thus far that was not part of the sensor package we had planned to use.
- Richard Tian: Most continuous monitoring sensors have chlorophyll as part of it.
- Peter: Certainly plankton density and movements may be a good reason to bring in chlorophyll.
- Tish Robertson: I second the support for the Rappahannock because it has all the designated uses. It has deep channel, it has the deep water use. You had in that slide there would be two arrays in the Potomac and in the Rappahannock. Would one be near shore and one mid channel deployment? How would the two differ?
- Peter: That is to be decided. These particular units were designed to get vertical profilers in the deeper, off-shore waters. That was the purpose, to separate them from our shallow water con-mons. We had more mid channel expectations for these, mid channel and adjacent to deep water without affecting boat traffic.
- Tish: In the Rappahannock we used to have a vertical profiler Virginia Institute for Marine Science (VIMS) managed for us under the shallow water monitoring program. We got cool data out of it but it was logistically very challenging. It was deployed in the shipping channel and it was constantly getting knocked over and dragged. We got it to work eventually but it was challenging.
- Peter: We should talk more about that with the hypoxia folks.
- Tish: Carl's group would have more expertise. Ken was in charge back then and may have some insight as well.
- Alex Gunnerson: MTM selection has not been finalized. It could very well be possible that the WQ GIT decides to recommend the Rappahannock be included in the RFA.
- Peter: The inputs on the designated use representation and other opportunities is still important, yes.
    - ACTIONS:
    - *Hypoxia Collaborative:* Discuss challenges of Rappahannock profiler and how to handle challenges with boats at hypoxia meeting.
    - *WQGIT and Modeling WG:* Consider the Rappahannock for inclusion in the Multiple Tributary Models.

12:30 PM Overview of updates to the GAM and cross validation work – *Elgin Perry (consultant)*

- Elgin: When we last talked, we talked about the need to develop cross validation tools. I've been working on that. We also talked about a need to try to determine whether there were certain regions in time or space where the model doesn't perform well. I've worked on graphics to look at those. I'll also talk about getting the model to fit better at certain regions, and increasing the flexibility of the smoothing functions.
- Elgin: We're working with data from the fixed station network that comes from the central portion of the Bay and covers the deep trench. The time period is from 1990-2010 and we have either once or twice monthly observations. I had gone through a model variable selection process and ended up with the model I'm showing the results of with an r square of 0.859. That was the best model I developed at that point. But after that, Rebecca came up with a model that had an r square of 0.863. The results I'll show today will use this model which I call Model gs7. This is my base model for looking at things like cross-validation and trying to increase the flexibility of it.
- Elgin: I used the 1990-2010 data for the cross validation effort, which is a process of identifying and setting aside certain observations to use for validating the model and the remainder of the data you retain for training data for training the model. In going through this exercise I considered a station date as an event. If I decided to put a station date in the cross validation data set, I took the entire set of observations from the surface to the bottom and moved those over to cross validation, I did not work with the individual depth measurements and split them up. I identified the months where we had two or more of these station data events within the month, and randomly choose one for the validation data, and retained the others for the training data. We will have certain dates into the cross validation data where at that station we don't have any observations from that date. So this is a test of how well the model is interpolating through time. At some point in the future, I think we'll work on a different cross validation that addresses the issue of how well the model performs when it's interpolating its base or extrapolating its base. Jon is looking at ways to get spatial data beyond the fixed station network.
- Elgin: This chart shows the number of observations, and these are individual depth observations. It shows the number in training, the number put into observation, and the percentages. About 30% of the data got set aside for validation, but it varies by station. The simple channel stations had around 30%; lateral stations (which are monitored in the season where we have around 2 observations a month) can be more like 40%, and there are some Eastern Embayment stations which are down around 20%.
- Elgin: The next chart is similar except we're looking at months. The winter months aren't well represented in the validation data. This is because we don't sample a lot in the winter. The spring time months are over-represented in the validation data. It's unbalanced but I don't see it as a problem because the model performs well in the winter; DO isn't that dynamic in the winter. The springtime is the time the model currently has greater variability. Having more data and validation then isn't a bad thing.

The next slide shows the fit of the model Rebecca developed. The table at the bottom shows the statistics that I'm computing from the validation data that we can compare to similar statistics from the training data. There's the root mean square error, the median absolute deviation, and the r square. These numbers tend to come out where RMSE and MAD you get a slightly larger value for the validation data than you do for the training data. For r squared you get a slightly smaller value for the validation data than you do for the training data.

- Elgin: These are plots of the residuals. On the left the 4 plots are of the training data and on the right they are for the validation data. I find it encouraging that the distribution of the residuals for these two are pretty similar. As Gary was telling us we're visualizing a two-step modeling process, and the first step is to estimate the means. Then we'll put on top of that a stochastic simulator. The residuals from the first step will do a lot to inform the stochastic simulator. The fact that the residuals from the training data exhibit similar patterns to the residuals from the validation data, which are completely independent, is encouragement that process might work well. These are more plots comparing training data on the left to validation on the right for observed vs predicted DO. Curiously there are some larger residuals in the training data set than there are for the validation data; we'd expect it to be the other way around. The two bottom figures are about whether or not we meet the homogenous variances assumption of v squared, so we're plotting the residuals vs the observed DO for both training and validation data. They look pretty similar.
- Elgin: Another thing that I did was to take that series of models that I've gone through from last time and re-fit all of that for training and cross-validation to see if looking at the cross-validation statistics we would reach pretty much the same decisions that we made when we were doing it just looking at the training statistics. It worked out that when you saw improvement in the training data, you'd see a similar degree of improvement in the cross-validation data. That leads me to believe it wasn't a bad decision to just go with the training data.
- Elgin: Here are some graphics I've been working on to look for places where the model isn't doing a good job of estimating the mean of the data. What clued Rebecca in on how to improve the model that I developed was to recognize that there were certain stations that had larger negative residuals than others. And so these plots are looking for those. You can see that station EE2.2 has some pretty large negative residuals and the box in the middle of the graph is pretty wide. That's an area where the model isn't doing great. There is a persistent trend that the model is doing better in the center of the Bay than at the lateral stations. I do the same thing with months. April and May have broader distributions of residuals, although when you look at things averaged over stations these residuals are looking pretty unbiased. Looking at plots for an individual month, we can see things like whether or not an individual station in an individual month is having biased observations. For example, CB5.4 has observed observations

above the predicted observations, so we're biased low in that region. Another station, EE2.1, has small range of residuals, but that center box is almost completely to the right, suggesting that 75% of the time we're underestimating the DO of that station. That's the kind of information we're hoping to glean from this which will guide us towards making a better model.

- Elgin: The box plots work well when we're dealing with the fixed station network because we can make a box for each station and compare pretty easily but I think we'll have to work on a different tool when we start doing spatial validation. That data is much more dispersed.
- Richard: My feeling is that it's pretty successful for the cross validation, but when I look at the number of residuals, it's uncommon up to 5, that makes me worried. This is the full range of criteria assessment up to 5. We need to bear that in mind that the residual looks high in terms of the criteria assessment.
- Elgin: They are large residuals and that's one of the things we'll rely on the second stage model to capture for us. You're right. The fact that there are large residuals here when we have a model that's pushing r squared of 0.9 suggests that looking at the mean value is not a good way to evaluate criteria and it emphasizes the importance of stage two of the model which will be simulating the variability that these residuals represent.
- Richard: What're the time steps of the output? Is it every day or every hour?
- Elgin: With this model we're planning to use it to generate observations every day. We can have it generate observations every hour but we're thinking the hourly part of it is something we will build into the stage 2 part of the model so that when we look at within day variability we might have a deterministic diel component in addition to the random component.
- Richard: Thinking asking criteria for 7 day mean. I'm thinking I'd like to see some variation based on 7 day mean. I'm hoping the number will be much better in terms of a residual.
- Elgin: I would expect it to be. That's something we could test once we start getting data from these profilers that Peter talked about. At this point we don't have a lot of data to test that. I would expect if you were to take observed data collected continuously over 7 days and compare that to what this model is predicting the residuals should be much smaller. Although for instance at CB5.4 the model is looking somewhat biased, and in that case I wouldn't expect a lot of improvement. We'll just have to see.

12:45 PM Review and summary of data sources utilized for the 4-Dimensional Interpolator tool – *Jon Harcum (TetraTech)*

- Jon Harcum: TetraTech's role so far has been to compile the data from the Chesapeake Bay Data Hub. That included our long term monitoring stations with monthly or bimonthly sampling. The Data Hub also has data from the CMON and DFLOW sites for their calibration events when the boats are stationary. We combined that together. We

also got metadata level information from the VIMS website Virginia Estuarine and Coastal Observing System (VECOS) which has the Virginia portion of the watershed, and Eyes on the Bay which is maintained by MD DNR. Shout out to Dave Parrish for providing me the meta level information for the VECOS sites, and Mark Trice for getting the meta information for the Eyes on the Bay sites. Since that time I was asked to drill in and follow on the work that Rebecca and Elgin have been doing, is eventually adding additional data into the pilot model they're working on. So far they've been focusing on the observations from the long term monitoring data, and it represents about 120,000 observations, which go from about the Bay Bridge down to just north of the mouth of the Rappahannock. The next layer of data put on the list was to look at the con-mon sites. For purposes of down selecting it, I chose a few sites focusing in on the location where the pilot model is set up now. I probably should have gone a little further into these eastern embayment sites. We picked about 28 sites, picked 3 sites highlighted as a first batch. Compiled the data, organized it a bit and passed it over to the modeling group.

- Jon: The 3 sites highlighted – one is by Sandy Point South Beach. This is a neat one because we have data back to 2004, and keep in mind that the data sets currently being used, they had about 120,000 observations, so at that one station we're almost tripling that by adding in the one station. The Gooses site is both surface and bottom sites are available from 2010-2020. That's adding another 210,000 observations. The VECOS site is Dividing Creek, a little north of the Rappahannock. That one has more recent data, 2013-2015. We packaged up 3 sites worth of data, moved that over to Elgin and Rebecca to navigate putting it into the pilot setting. We were weighing the options of do we want to expand the pilot model framework and start going up some of the nearshore, shallow waters, or do we want to just figure out how to get the data into the existing pilot model. The decision was made to figure out how to get this additional large amount of data into this modeling framework. The other site we have other data, the profiler data, there's a couple of sites at the East and West Gooses sites. They've been online this summer. The type of data that's available is at 10 minute intervals, at 1, 3, 5 meters all the way down to the bottom, and about another 45,000 observations per month. The NOAA site does a nice job of displaying that data and making it easy to download.
- Peter: Excellent, we've had significant temporal investment in the con-mon program, shallow water and multiple uses for that coming up with the Bay model moving more into the shallows and our own analysis that needs to accommodate that.
- Breck Sullivan: I'm curious about your third option from VECOS. Were you wanting to choose a site that had a short time period of data or were you looking for one that has more? Because I could provide you with some examples.
- Jon: Great question. Based on the metadata that I have available to me I was more interested in picking sites looking to get exposure of data from Eyes on the Bay and the

VIMS site to get a flavor for what the data looks like in terms of how they store it and any concerns I might have going from the two different types of data. Of those 3 sites, all had data from 2013-2015. So it was pick one of the 3. I didn't recognize the eastern embayments that Elgin was indicating, because there were some sites that remain of the long-term monitoring network of eastern embayment sites, perhaps going over to eastern shore Maryland there might have been some additional sites with a longer record but I hadn't picked those up on first pass of things.

- Breck: This might be a better question for the hypoxia collaborative team but I was hoping for clarification on how many vertical arrays are being funded and how many you're looking for location sites? Thought it was 10 arrays, but one would be left out in case one got snipped again or needed to come in for maintenance.
- Peter: We need at least 10 sites and redundancy was factored in to have machines off shore. There's a potential 11th site (Fishing Bay site is being looked into by MD DNR to assess a particular segment) and there are reasons to believe others are interested in expanding the numbers of vertical profilers. It would be to our benefit to look beyond 10 sites. Permitting may not go as planned, may need a plan b for a particular site. And to expect that other folks may be deploying and they want to be part of the network and facilitate data collections in line with our objective than it would help to have more than 10 sites chosen as part of our network. We have 11 arrays funded.

1:15 PM Questions/Feedback

- Jim Hagy: Looking at Elgin's models, I noticed bigger residuals happened on edges. What sounds like a terrible error such as a residual of -5, it could be water that you thought that was 2 meters lower is actually in a particular spot. Where those residuals occur also depends on the time of the year. You're not going to see big residuals associated with pycnocline tilting if there's not much of a DO gradient above and below the pycnocline, in January for example. The same edges occur in the latitude sense and I think that's what's going on in 5.4 where oxygen is declining as you go into deep part of Bay and if that's not where you expect it to be, you have a residual. What is the impact of that on assessment? It's not a huge issue however, because these are transient conditions that reflect these edges. In the end, time and space attainment plot may not be affected by them nearly as much as we think but it would be nice to be sure of that.
- Peter: Throughout the water column, did you give it for one depth, or were you consolidating the entirety of the residuals, treating each depth as part of the summary package?
- Elgin: I was giving you the picture of whole water column in each of those residual plots. I could slice the data by depth as I sliced it by month. I would like to add to Jim's comment - another thing could be going on at CB5.4 is that it's right at the edge of prediction space and these GAM functions tend to be wonky at the edge of the

prediction space. That emphasizes the idea that if we're trying to fit a GAM for a particular segment, we would want to include stations from neighboring segments to stabilize the wonkiness of the edge of the prediction space.

- Jim: In situations when the entire box in the box plot sits to one side of the predictions, those seem like something should be a greater focus initially to figure out what's going on with that.
- Elgin: I agree.
- Jim: Maybe what you just said is part of the solution, to expand the spatial domain. The term used in simulation modeling is a burn in. The model is out there but you ignore the edges of that.
- Rebecca: I agree. One thing Elgin didn't mention but is a next step that we've played around with a little bit already, is that in some of those regions, maybe the more northern or southern part of the prototype region, is incorporating the impact of daily flow observations that we have to some extent. We've fought with it because it is a challenging thing to do, but that is one tool we have if we have some stubborn places where we can't get a regular pattern of oxygen. Maybe there's some relationship with the wet and dry periods that could help explain that. It is a challenge because the relationship with flow is different, all these depths and times we're dealing with, but it might help build a target in on some of those more stubborn problems.
- Jim: For thinking about where in time you deploy continuous sensors, is think about areas with higher variability of residuals, those edge places as being places where you want to have observations to better quantify that variability, not necessarily so you can eliminate it, although it would be nice to, but if the cost to do that is massive increase in model complexity, then potentially you just want to make sure you have a good understanding of what it is and why it is what it is so you can explain it.
- Peter: Yes, hope we can leverage work done here to inform those decisions on locations.
- Jim: There's a lot of competing objectives and interests in citing these things.
- Tish Robertson: Can the profiler data be used?
- Rebecca: Yes, it can.
- Peter: The investment that has just been made for the 10 instruments is this fixed center depth approach but other folks may introduce a different form of technology which is going to be a slightly different dataset which presents another challenge in terms of the data ingestion and use. Hopefully not too much more complicated than fixed sensors. But that's another one we can be sure we work with given that other agencies and institutions may make use of the profiler as an alternative instrument.
- Tish: I was thinking it might be useful for model calibration.
- Elgin: I agree with Tish. We just don't have a lot of information about diel cycles, at least not in deep water (we have a lot from con-mon, but that's shallow). It will be useful when we get to stage two of the modeling process and try to understand what the

deterministic variability is on shorter scale in deep water. There's the profiler data from the Rappahannock, there's a few EMAP data sets that get at that, but there's not a lot around.

- Peter: Let's touch bases about the challenges of the Rappahannock.
- Elgin: The idea about increasing model flexibility - there's a parameter in the GAM called k which controls the dimension of the spline basis you're trying to fit, and that's one way to increase flexibility in the model. When you start playing with that parameter, the execution time of the model fit goes up dramatically. I was able to tweak k for one term in the model and go from a 0.86 r square to a 0.9 and that was a great jump but it took a half hour. The next time I tweaked another parameter, it took over an hour. The third parameter I tried to tweak it told me I couldn't allocate enough space. We're up against the limits of the technology and the size of the machine I'm working with. Other thing I didn't look at is trying to use a distribution other than normal distribution. The Tweedie distribution is good for fitting data that are not negative, that seemed like a good choice because of DO. When I put that in for the distribution using the same terms that Rebecca used it, took almost 5 hours to reach a solution. Then when you look at the results, I couldn't see any improvement over what we're doing with the normal distribution. This is not surprising because we have such a huge dataset it's not very sensitive. If we were working with 100 observations I think we'd see some differences, but the thousands of observations we have both methodologies produce the same answer, so I'm sticking with the normal distribution.
- Peter:
    - ACTION: *BORG* - Come back to this in a small team discussion. Where is the breakpoint for capacity, what do we have for resources to go bigger than a laptop and home computer?
- Isabella Bertani: We knew originally that we may run into computational issues. We do have cloud resources. Maybe this is them we talk about moving from Elgin's computer to our cloud resources which will increase capacity.
- Peter: Thank you to everyone for your presentation contributions and discussion contributions. We'll have more information on deployments and new sensor units soon. It feels monumental after the years of thinking about having data on short time scales in habitats of interest, it is phenomenal to have this new set of investments which will benefit research and management. To see where we're headed with modeling applications and improvements has been very helpful today and we'll continue to improve our efforts.
- Rebecca Murphy: If you have any additional thoughts or comments, please email your input. Thank you!

1:30 PM Adjourn

Participants:

Amy Goldfischer (CRC), Jon Harcum (TetraTech), Rebecca Murphy (UMCES), Elgin Perry (independent consultant) Diana Domotor (MD DNR), Breck Sullivan (USGS), Isabella Bertani (UMCES), Kaylyn Gootman (EPA), Peter Tango (USGS), Richard Tian (UMCES), Tom Parham (MD DNR), Mark Trice (MD DNR), Matt Stover (MDE), Tish Robertson (VADEQ), Gary Shenk (USGS), Mark Nardi (USGS), Jim Hagy (EPA), Amanda Shaver (VADEQ), Angie Wei (UMCES), Leah Ettema (EPA), Lee McDonnell (UMCES), Alexander Gunnerson (CRC)