Criteria Assessment Protocol Workgroup (CAP) Meeting

Wednesday, December 18, 2024
10:30AM-11:45AM

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

Action items:

- Please consider the questions for the Beyond 2025 Outcome review assessment within your organization (questions included in "supporting documents") regarding the Chesapeake Bay Watershed Agreement Water Quality Standards Attainment and Monitoring Outcome.
- Please consider what key topics you want to discuss first in 2025 and send these key topics to Peter Tango (ptango@chesapeakebay.net).
- January 2025 meeting invitations for discussing the Beyond 2025 Outcome review were shared (January 8th, 14th and 22nd). Please contact Peter Tango (ptango@chesapeakebay.net) if you were not added and you would like to be or you did not receive the calendar invitations.
- VA DEQ plans to brief STAR and the Water Quality Goal Implementation Team in 2025 on their plan for Dissolved Oxygen Criteria Assessment (same presentation shared at this CAP meeting).
    - Please share any feedback on the presentation that VA DEQ gave at this meeting with Tish Robertson (Tish.Robertson@deq.virginia.gov).

Participants:

August Goldfischer (CRC), Peter Tango (USGS), Gary Shenk (USGS), Fred Irani (USGS), Karl Blankenship (Bay Journal), Leah Ettema (EPA), Jon Harcum (Tetra Tech), Claire Buchanan (ICPRB), Tish Robertson (VA DEQ), Cindy Johnson (VA DEQ), Amanda Shaver (VA DEQ), Joseph Morina (VA DEQ), Thomas Bryant (VA DEQ), Joseph Wood (Chesapeake Bay Foundation), Rebecca Murphy (UMCES), Tom Parham (MD DNR), Matt Stover (MDE), Becky Monahan (MDE), Sophia Grossweiler (MDE), Melinda Cutler (MDE), Clifton Bell (Brown & Caldwell), Juan Vicenty-Gonzalez (EPA), Gabriel Duran (CRC), Elgin Perry (Independent statistician), Richard Tian (UMCES), Liz McArthur (VA DEQ), Carl Friedrichs (VIMS), Mark Trice (MD DNR), Guido Yactayo (MDE), Jon Mueller (UMD), Adrienne Kotula (CBC), Lewis Linker (EPA), Sara Ramotnik (NWF), Pat Calvert (VCN)

Minutes

10:30 AM **Welcome, Introductions & Announcements – Peter Tango** (USGS), Chair

Upcoming Conferences, Meetings, Workshops and Webinars:

- 14th National Monitoring Conference – March 10-12, 2025, Green Bay, Wisconsin.
- Coastal and Estuarine Research Federation Conference – November 9-13, 2025, Richmond, Virginia.

10:40 AM **Update on Beyond 2025 Phase 2: Outcome Assessment and Priorities List** – *Peter Tango (USGS)*

Summary:

The Reaching 2025 report was completed and published, and Beyond 2025 Phase 1 has moved forward. The Beyond 2025 Phase 1 efforts ask: do we need to shift the expectations of the Watershed Agreement? We are not looking to start from scratch or create a new agreement, but rather adapt and use the Agreement in the way it was developed to apply adaptive management principles. Eastern Research Group (ERG) also conducted an independent evaluation of the Agreement, Outcomes and partnership and evaluated Outcomes using Specific, Measurable, Achievable, Relevant and Time-bound (SMART) principles. This is our opportunity to set the stage for Beyond 2025 Phase 2. Part of this is an outcome assessment, which takes what we learned and how do we start to put into practice directions for the program. We need your input and wisdom on this assessment.

Chesapeake Bay Program leadership has asked us to consider what they are calling the Big Question: What advice do you have for the Management Board on how to consolidate, reduce, update, remove, replace or add new Outcomes? The Goals will remain the same in the Beyond 2025 updated Watershed Agreement.

For the Water Quality Standards Attainment and Monitoring (WQSAM) Outcome: We presented the Outcome language as having 4 elements in the Strategy Review System (SRS) process recently. There is the watershed piece, the Bay piece, reporting water quality standards, and capacity to monitor. None of those speak specifically to an ecological outcome in terms of the endpoint; we're saying we're making the progress to support that Outcome evaluation. The Outcome is more about how we get to understanding the goal, than the endpoints of the goal.

According to ERG's evaluation WQSAM did not meet SMART goals criteria. They stated it contains no specific or measurable elements and does not have a time bound aspect. The second aspect of the Outcome (report annually) represents an output. Those elements are fundamental for communicating our results, but not a reflection of an ecological endpoint. Another framework for considering assessment of the Outcome is a logic model framework. When we look at our Outcome and talk about doing the monitoring and producing deliverables those are significant activities we engage in to understand the system. We generate the products (outputs) and we do get to what change we're detecting. The target of that change isn't presently reflected in the outcome text.

In addition, at the biennial Strategy Review System meeting we came up with lessons learned for what makes a good outcome (secret sauce): • Clear in its objective • Measurable • Monitoring program that supports/reinforces the outcome • Partner commitment • Resources identified and/or available to support the efforts necessary to achieve the outcome • Centered around people & living resources (not solely water quality)

Discussion:

Peter Tango: Does anyone feel we should keep the same language for our outcome going forward? (No) Does anyone feel we already have a SMART outcome? (No)

Matt Stover: We'd like a little more time to think about these questions as this is the first time we've seen these.

Peter: For sure. This is not the final date we'll consider this, we have more time. This is the opening discussion and not the deciding meeting.

Tish Robertson: We would like to cosign that. What you're asking us is important and we'd like time to deliberate internally.

Timeline:

- 12/19/24 STAR meeting to discuss Outcome review assessment across all Outcomes.
- 1/8/25 Monitoring Meeting to discuss WQSAM Outcome review assessment.
- 1/14/25 Water Quality Goal Team Beyond 2025 Office Hours for Water Quality Outcomes.
- 1/22/25 Monitoring Meeting to discuss WQSAM Outcome review assessment.
- 2/13/25 Outcome assessment materials due (2-page write-up).
- 2/27/25 Discussion on Outcome with Management Board.

The questions for the Outcome review assessment will be shared alongside the minutes.

Matt Stover: In MD, monitoring duties are split between DNR and MDE. I know we have a lot of thoughts on how to best support this outcome. We've talked quite a bit with Tom Parham and folks to re-think about how we're allocating resources through that cooperative effort. Please make sure to include MDE in any discussions and decisions there. We want to be a big part of that.

Peter Tango: We will make sure to.

We are looking at suggestions to break it out into multiple outcomes or write outcomes more in line with the goal of reducing pollutants and achieving water quality. Think about the connectivity in your suggestions, knowing we are not looking to create brand new monitoring programs.

Adrienne Kotula: Another option is keeping one outcome with multiple indicators.

Peter: Yes.

Homework: Look at the template questions and come back in January with suggestions.

Matt: I'm having trouble understanding the context, what is the framework this falls within? Where does this come from? We know we need to assess all our DO criteria and find a way to monitor it, and in some cases interpolate it. That's what we're focused on. Can you give some context behind what sounds like larger planning discussions and what it is we're commenting on?

Peter: In the Reaching 2025 process, there were a lot of expectations in the Agreement that hinge on achieving the Goals of the Agreement and Outcomes are meant to be diagnostic checkpoints related to Goals. Jurisdictions have signed onto this. Early on it was recognized some Outcomes are more qualifiable and some are more quantifiable. Some groups spent 10 years figuring out their baseline and how they were going to measure something. Some folks said that's not what we should be measuring after 10 years. There is misalignment and changes in expectations. From my perspective working in the Beyond 2025 group there is thought to simplify. 10 goals and 31 outcomes are a lot to manage. Not to say that is exactly what will come out, the expectation is to keep the 10 goals. But simplify 31 outcomes into something easier to measure and track, without rewriting the Agreement. In history there have been a lot of rewrites but this is an adaptation, adjusting outcomes to be more consistent based on input from community.

Matt Stover: These would go into the new Bay Agreement?

Peter Tango: They will help guide updated expectations in the Watershed Agreement.

Leah Ettema: Is there potential to restructure the Goal Implementation Teams and STAR workgroups based on this?

Peter: Yes. There is a phase of this looking at the structure of the CBP and how it operates. It's in consideration and discussions. As a function of some groups consolidating, looking for synergies that work together that could happen.

Matt: From MD we would second that idea. In our August meeting down in Virginia we suggested restructuring of STAR because there are a lot of workgroups that overlap. It's a lot of the same staff who would cover those. We realized we might be missing some important discussions from those workgroups.

Peter: I don't know the schedule for the structure discussions, but I'll keep you updated when I find out.

11:00 AM **Virginia DEQ's plan for Dissolved Oxygen Criteria Assessment** – *Amanda Shaver and Tish Robertson (VA DEQ)*

Presentation:

Virginia has 11 DO criteria applicable to tidal waters of the Chesapeake Bay and its tributaries. To date we've only assessed 4 out of 11 of these criteria. This is a big deal because we can't show progress towards meeting the Bay Total Maximum Daily Load (TMDL) until we can assess all the designated uses.

These are the Bay criteria organized by designated use (DU). We have 5 DUs in the Bay. The shallow water bay grass use is protected through Submerged Aquatic Vegetation (SAV)/water clarity goals, so not part of this discussion, but others are protected through a suite of DO criteria. The ones we routinely assess are highlighted. DEQ staff looks at this and sees gaps in DUs we're not assessing. We're trying to find a way to assess all criteria

In VA's water quality standards regulation, we have a section that regulates all criteria in VA's part of the Bay. In the same section we have implementation language. We are in the process currently of modifying that language in a fast-track rulemaking process. A conventional rulemaking lasts about 2 years. We decided if we want to meet the goal of having an assessment of all 11 Bay DO criteria in time for 2026, our next assessment report, then a fast-track rulemaking would be in our advantage. There is still a public participation process but it's expedited. We're trying to have assessment of all DO criteria by 2026.

Tish showed a slide with the language VA DEQ is presenting modifications to. The underlined language are insertions, and the language we're pulling out has strike throughs. The language says we have to use the Cumulative Frequency Distribution (CFD) to assess attainment of Bay criteria. We're adding that you can also use other scientifically defensive assessment methods to add more flexibility. You also see references to EPA CBP publications. The documents post-2007 are implementation procedures that reflect work from the CAP workgroup. The most recent technical addendums are from 2017. VA adopted that through fast track rulemaking in 2019.

Why are we modifying the language? To assess all 11 Bay DO criteria. We want to maximize the use of all available datasets and use datasets we haven't been using because they're not compatible with the CFD framework. We want to reduce the need for enhanced monitoring. We've been told the current assessment tool we're using is not compatible with monitoring frequencies of the discrete monitoring program. If we want to assess the 1 or 7 day mean criteria, we would have to go out weekly or daily to do a good job of assessing them using the CFD. That would obligate us to do more enhanced monitoring. That's expensive and impossible for us to do in all segments. In addition, we currently have a very complicated assessment tool. We think it would be good governance to have simpler procedures used alongside the complicated assessment tool for greater transparency. We want the public to be able to replicate what we do and enable easier communication with the public, our own staff and management. We don't have a tool that is easy to communicate currently. We want to use decision tools that are in practice of national EPA guidance. Finally, we want to show progress on meeting DO criteria that are outside pass/fail (being able to show incremental progress) and we currently don't have a way to do so.

The 2017 document was very important, it was the first stab of looking outside the CFD framework. There is a chapter that talks about short duration criteria assessment procedures. There is a conditional probability analysis chapter that gives us a way of assessing the 7 day mean criteria but still requires us to use the interpolator CFD approach. The other set of procedures this workgroup worked on was the sub-segmentation of segments approach, which allows us to take a segment and break it into different zones and through enhanced monitoring analyze the monitoring data outside the CFD framework to form a conclusion about criteria attainment.

This is what I mean by different zones. We have offshore, shallow water, tributaries. Collecting different kinds of datasets in different compartments and rolling up results to summarize. This table summarizes what we came up with for open water use. First 3 bullets on right are tied to the CFD framework. Then below that are procedures independent of the CFD. A lot of work went into this. I see it as the first stab of how to go about doing Bay DO. There are good things, there are also weaknesses. I think we should be treating it as a living document/living work that we have ongoing discussions about.

Going back to rulemaking, the implementation language does reference the 2017 document. We discovered we can't use those procedures we talked about unless we modify the language because it says we have to use the CFD. Also, we are under pressure to use all available data, which is a requirement of the Clean Water Act.

At the August CAP meeting Leah Ettema gave us a great walk through of the requirements that states have to meet when doing their 305b, 303d reporting. If we exclude a dataset from our evaluation we have to provide a rationale. We've been excluding continuous monitoring data for 20 years. The rational we're giving EPA is no longer satisfactory. We have to come up with an assessment tool that does use continuous monitoring.

How much data are we talking about? A lot. Summary of data for 2024 integrated report (IR) (2020-2022) shows there are 28,000 DO observations. There was no good reason for us to not be assessing all of the instantaneous minimum criteria. I don't have a good reason to not use this data to assess those criteria. Those are the easiest criteria in the world to assess.

DEQ staff have developed a tentative plan for assessing the 11 Bay DO criteria. We are here to get feedback from the partnership on this.

The interpolator CFD assessment will continue to be the default assessment method and considered the first line of evidence. What we want to do is layer other lines of information on top of the interpolator. The interpolator is fed through discrete data, but we have multiple data types. Continuous Monitoring data (con-mon) is a huge amount of information. How do we integrate the two types of data? We know there are other analytical approaches. In our inland waters we have a way of assessing waters that does not rely on the CFD. Let's take those tried and true methods and use them in conjunction with the CFD. I'll call this the screening value assessment. A screening value assessment is used in risk assessment to make a rapid determination of a site with a low probability of risk vs one that has some concern. Screening values are easy to use because you apply them to data in its rawest form. Comparing observations directly to a threshold, you make an inference on the state of that water body or site based on the exceedance frequencies you see.

The Bay DO criteria are complicated because of their durations. Those durations are usually not found in states' Water Quality standards. We're saying let's ignore the durations and just look at the magnitude. We'll turn these criteria into instantaneous thresholds and compare them directly to monitoring data, look at exceedance frequencies and make an inference of water q based on that. We settled on a 10% cut-off because it's endorsed by EPA. We think it has good traction, but we're here to have a discussion about it – if you have a different recommendation we will consider it.

For the con-mon data we came up with procedures back in 2017. We'll build on that work. In VA, we don't have any deep channel deep water con-mons, all con-mon data is coming from shallow water. We still want to use these data; we think they're useful. We can make decisions from shallow water con-mon as long as we integrate it with other tools and information we have. We have made some attempts to tell stories from our con-mon data. In our 2024 IR we put together this map showing at each con-mon station what the criteria exceedances look like. We use the procedures from 2017 to do so. This is not an actual assessment or decision-making tool, it's just a narrative, because right now there nis o way for us to integrate con-mon and discrete data and we have the language in the regulation that says we can't use an assessment method outside of the CFD. Our goal one day is to use this outside of the CFD but not in an isolated way.

This is a hypothetical segment using 3 lines of evidence to assess it. We would first look at the interpolator results. Most of the criteria in this table cannot be assessed with the 3-D interpolator/CFD. Then we would move on to screening value exceedance rates. Then we would look at shallow water con-mon exceedance rates. We would look at all these methods to make a decision. Tish showed hypothetical segments with one being not impaired and one being impaired. All lines of evidence for all criteria would be considered to make a decision.

Benefits of an integrated assessment approach, besides being able to assess all data and criteria, are that we recognize there's uncertainty in all evidence and we're not being held back by it. All the methods have weaknesses. We can show the public we are using their resources wisely.

The takeaways are that we're not abandoning the CFD. We put a lot of investment in the interpolator. We will assess segments as impaired if we have any reason to think it's not attaining. We cannot say a segment is meeting its DUs if we cannot do a demonstration showing that, and we'd use all lines of evidence to do so.

Our efforts with CBP are that we are fully committed to partnership. Tish has been an active member of CAP for the past 17 years, along with Cindy and Amanda. We like the work this workgroup does. This past year we had many meetings with CBP, EPA Region 3 and MD talking about Bay DO, talking about what we're doing and aligning on some compromise points. We think we're making good progress on getting people on board to understand our motives. The CAP meeting in August was when we first laid out this vision. We eventually want to brief STAR and the Water Quality GIT as well in 2025**.** We want folks in wider community to see what we're doing.

We have a timeline we're trying to meet. The 2026 IR is on the horizon. We try to shoot for April 2026 for that. Under the Clean Water Act, there's expectations for states to submit IRs in April of even years. We kick off assessment reporting with a review of assessment procedures and guidance manual in spring of 2025. That would be where folks can see our vision in writing and get into the weeds of what we want to do with Bay DO. Under state code we have to have public participation. It's a federal requirement as well. We have to respond to public comments and questions we get. We consider all the feedback we get. I've gotten excellent feedback already that has given me good things to consider for this methodology. VA DEQ hopes to continue that feedback and interaction through the CAP WG.

11:20 AM Discussion

Sara Ramotnik: When does public feedback period end? Tish: The comment period for the rule-making ends January 1. The feedback we have on the intended plan – we don't have a deadline for that. Whenever you get a chance, you can send us feedback.

Matt: I applaud you on all the efforts you've done. MD's biggest thing is to make sure we're consistent as we can be across jurisdictional borders. Your slide #6 stated reasons for these efforts and MD is in full support. We share all those desires, being able to use con-mon data, all existing data etc. We don't want to wait for 4D interpolator in 2028. We want something more easily understandable by the public. We have a lot of tools for assessing inland waters that work fine, all those options should be on the table. We want CBP to support our efforts rather than going into complex models you have to have a PhD in statistics to understand. MD supports this.

Joseph Wood: There's been a lot of concern about this. I appreciate the presentation. I wish it had been included in the regulation release. I think people are concerned because they see small language changes and imagine things outside of that. I think what you described isn't necessarily guaranteed in the regulation change but that's a different conversation for a different day. What you described doesn't cause me concerns, it's an exciting valuable tool and I'm interested in seeing it go forward. But historically CBPO doesn't approve this stuff. Historically addendums have paved the way for but they take a long time to put together. Is there a way for this group to endorse things or provide approval that could give public an assurance outside of technical addendums which takes forever to put together? I think having partnership support makes it a stronger tool in long run. The other thing is not letting perfect be enemy of good. Thinking of the tiered approach floating around the CBP. The big 4 we're already doing, we'll keep doing it and that will set the stringency of the TMDL in long run. This is the fringe thing, the things we haven't done yet that haven't been a focus, and this sheds light on them. I think there is a real opportunity there. I think it's very valuable to shine light on data that have not historically had a light shined on them. Having capacity to do this is really important for the public so we can sell the benefits of why we're doing restoration and we can see benefits in different places. Really appreciate the work and hope we can get to a place where this comes to fruition.

Peter: Regarding endorsements outside technical addendums, we've historically tried to publish short notes or communications that get folded into an addendum. We'll come back to that – it is worthy of further discussion.

Elgin Perry: Do you plan to use data flow in this expanded assessment process?

Tish: Great question. We do have some questions about whether data flow would be the best way to look at DO because it is just a surface monitoring program, but we will seriously consider it. If the partnership thinks that is a valid way of looking at DO, we will include that as an additional line of evidence.

Clifton Bell: In support of the idea of using other technically defensible methods with brief comments on other things we talked about like alignment with underlying duration components of the criteria and considering borderline cases. The details will be to come. What do you see as the next version of the documentation of this? Will it be that spring document? Or something before then, maybe through this WG where there is more details on decision rules for example?

Tish: We know we'll have a public comment period on the guidance. We'll have more discussions as a partnership. As we move closer to that public comment period those decisions will have been more solidified. I don't know if we will be in a position to share what we want to do before the public comment period. We would let partnership put in two cents during that process because that's how we normally handle our assessment methodology.

Clifton: Will the level of detail you provided today be the level of detail available until that document comes out?

Tish: No. We're prepared to go into the weeds and have more extensive discussions. Normally when we come up with an assessment methodology the form would be during the public comment period where we have it in writing and have it as a proposal.

Gary Shenk: The things you shared on slide 6 make sense. Also agree with what Joe said - what you presented here is a little different than what's written in the regulation. What you presented here is that CFD if available, a specific screening method is different than anything that is scientifically defensible. This is a pretty significant change and probably in a good direction. I'll withhold judgement until I understand that these changes are protective and that's what I would like to see in this process, a demonstration that these changes are protective. To go through that whole process - the partnership process of looking at all these things together, coming up with a technical addendum that formalizes what we're doing, generally having a STAC review - makes sense to me because it allows full review by the partnership of all the details. I know that causes problems because you're interested in doing this quickly and getting it out in next IR. I understand that and that's a difference we have in the moment. I just want to put in the partnership procedures we've followed for past decades.

Tish: We've been waiting 20 years for the partnership to make any kind of progress on this. We've been patient. We're at the moment where we feel we have to take the lead. We have to be more assertive because the hierarchy and slowness of getting things through the partnership are keeping us from meeting our objectives - not just our state objectives, but objectives we have to meet under Federal code. We are always going to adhere to Clean Water Act mandates over tradition which is what you're referring to. We're getting pressure from EPA Region 3 and they provide oversight over our assessment

methods. We're almost at 2025 – it's time for us to show progress toward meeting the Bay TMDL. We want to be good partners, and we want to be good practitioners of the Clean Water Act and sometimes those things are not going to be aligned. It's always going to be the side where there is litigation involved.

Gary: I wasn't advocating tradition for tradition's sake, just for that process allows for a lot of review.

Matt: MD also has been advocating for help with some of these other methods specifically for the Fishing Bay data we've been collecting for the last 3 years. We made a request 4 years ago to have help from the CBP in looking at different methods of assessment rather than the 3D and 4D interpolator but we haven't had a lot of movement on that. VA and MD have been communicating a lot because we are in lockstep in many respects. We would like other methods. I know they take a long time to develop. That's the benefit of the CAP WG that we can come together to discuss these things.

Peter: Thank you. I think we put 2017 documentation together as a means of supporting alternatives to consider and some of those have been put to use in various years. We created the indicator as a means of adopting some of that idea for conditional probability that reflects the lack of data while not necessarily being the rigor of what we have as far as how we've used the 3D for 30 day means and instantaneous minimums. The 4D interpolator evolution is driven by trying to meet these needs and expectations we've been short on. Collectively, really appreciate what MD has brought with Fishing Bay, what DEQ is working to fill the gaps on, what we're doing with the new infrastructure and interpolator. It's a groundswell of recognition we've been short for 20 years and trying to fill that. Respecting the tensions there are in trying to make these changes. Thank you for pushing the envelope forward. Great thinking here and looking forward to tackling what we can in 2025.

Matt Stover in the chat: One thing we'd really like to cover in our next CAP WG meeting is the underpinnings of our different DO criteria to determine the applicability to con-mon data. Essentially we'd like to get back to basics before we go down the road of trying to assess this data in the 4D interpolator. To us, developing the 4D interpolator before settling how we want to assess these kinds of data with respect to our criteria seems like putting the cart before the horse.

What topics do you want to discuss first in 2025? – *homework assignment*

11:45 AM Adjourn