Chesapeake Bay Program
Science. Restoration. Partnership.

Modeling Workgroup July Quarterly Review

Day 1 – July 8, 2025

Event webpage: Link

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

Table of Contents (jump to bookmark):

Day 1 – July 7th, 2025:

Skip to Day 2.

9:00 Announcements and Amendments to the Agenda – Mark Bennett, USGS and Dave Montali, Tetra Tech

- New Publication [Link]!

9:05 Proposed New Timeline for P7 Model Suite Development – Lew Linker, EPA-CBPO

Description: *A new timeline for P7 (P7) Model Suite development that accounts for delays in model input data sets will be discussed.*

The key points we will begin and end with are that delays beyond the control of the Modeling Workgroup have affected the P7 model development. There are two main causes of delay: the availability of key datasets and the loss of key model practitioners from the P7 modeling team. Despite these challenges, we are on track to deliver a P7 model product with better performance than the P6 model suite.

The ultimate decisions on the final P7 timeline will be made by the Principals' Staff Committee (PSC), the Management Board (MB), and the Water Quality Goal Implementation Team (WQGIT), though we will propose a timeline. Key datasets in question include inorganic fertilizer, poultry broiler and layer populations which will be made available by the end of February 2026. Other P7 model suite datasets and decisions are on schedule for completion by the end of September 2025 (Slide 3).

Slide 4 lists the team members we have lost. This list of key personnel includes Gary Shenk, Jesse Bash, Isabella Bertani, Alex Gunnerson, and Nicole Kai. Remaining on the team are Gopal Bhatt, Richard Tian, Joseph Delesantro, Lewis Linker, and Carl Cerco. In terms of full-time P7 practitioners, we are now less than half of what we were before.

Slide 5 shows the critical path and start (S) and end (E) times for each task. All decisions must be finalized by the end of September this year, including finalized segmentation, Chesapeake Assessment Scenario Tool (CAST) load source lists, wastewater infiltration determinations, and reconciliation of land-use construction records with state records.

The GIS team, led by Sarah McDonald, along with key practitioners like Jess Riegleman, will take this information and, over two months, prepare the initial 49 land uses plus any new land uses in CAST for initial review by CalCAST and the dynamic watershed model (DWSM). By December 1, 2025, we expect to have the first P7 land use incorporated into the model. This will allow us to test for any surprises and determine necessary adjustments.

Once the three manure datasets are available at the end of February 2026, we will produce the final P7 land use by the end of March 2026 for ongoing modeling work. The timeline then proceeds with CalCAST assessments and analysis, followed by DWSM calibration and final checks beginning May 31, 2026. The Main Bay model (MBM) will then conduct its one-month calibration with the final DWSM model inputs, and the multiple tributary models (MTM) will adjust their boundary conditions accordingly.

This timeline does not reflect all the parallel work occurring in CAST or other peripheral activities but provides a simplified view leading to the end of November 2026.

At every quarterly modeling meeting, as planned and implemented, a new tranche of DWSM results will be produced for the MBM and MTM, lead by Zhengui Wang and his team. We remain committed to incremental improvements and have demonstrated measurable progress with each quarterly update.

By the end of this year, we will have a functional, operational P7 model suite, including at least one month with the initial P7 land use incorporated. This will enable us to integrate final datasets and land use in subsequent phases. Even now, the hybrid P6/P7 model – which uses P6 riverine inputs and P7 dynamic watershed processes – is already performing better than the P6 model in simulating coastal plain loads.

The next slide addresses the year of review and new target development periods. These decisions are outside the MWG's authority and will be determined by the PSC, MB, and WQGIT. A Scientific and Technical Advisory Committee (STAC) peer review and Chesapeake Bay Program Office (CBPO) management review is expected from October

1 to October 31, 2026. The likely application period for new targets will run from November 1 to December 1 over approximately three years and two months.

The challenges we face, delayed key datasets and the loss of critical model practitioners, are beyond the MWG's control. However, with the revised timeline, we remain confident in our ability to deliver a superior P7 model relative to the previous P6 suite.

9:35 Discussion of Proposed New P7 Suite Timeline

Comment from chat: *Scott Heidel:* PA DEP is working with USGS and Resolve Hydro to provide ag land data from the remote sensing pilot project to be used to improve the accuracy of the agricultural land use data layer for P7. This is very important to PA so I wanted to put it on your radar to help ensure it happens as planned.

9:45 Plans for the 2026 Year of P7 Model Review – Lew Linker, EPA-CBPO

Description: The P7 Model review year done with both a STAC led technical peer review and a policy review by CBP decision makers will be discussed.

The review is something requested by the MWG. The year of review will be facilitated through STAC for technical peer review and the MB for policy review. It is an independent review and more of a collaboration than any unilateral decision. We rely on these teams to provide these services.

In Phases 4, 5, and 6, we had extensive STAC model peer reviews. These reviews are extremely useful for building confidence in the tools, increasing knowledge both within the program and STAC, and offering suggestions for future modeling phases. STAC includes many practitioners throughout the Bay region and beyond. Additionally, the Environmental Protection Agency (EPA) requires peer reviews for these models, so it's not only a good idea, but also a requirement.

How do you conduct a peer review? You could start with the Peer Review Handbook, 4th edition, from the Science Advisory Board. It's a thoughtful and comprehensive guide. Other recommendations are in documents like the Guidance for Development, Evaluation, and Application of Environmental Models. The Science Advisory Board emphasizes that scientific models should undergo independent third-party reviews. Reviewers should be selected based on expertise and independence, meaning no direct involvement in the work being reviewed. Reviewers may come from STAC, within the Bay region, or nationally or internationally, depending on budget. A clear purpose and focused charge are essential; reviewers need a defined scope and key scientific and technical issues to address. We do not re-review items already reviewed and documented. Instead, we focus on new elements like shallow water assessments or fine-scale watershed modeling.

The process involves modelers evaluating reviewer comments and using them to improve the final product. Peer review provides guidance, and the P7 model practitioners respond to it to produce a better model suite. Peer review also addresses model application and limitations, like its domain, timeframes, and scientific defensibility. Key review elements could include modeling purpose, theoretical basis, parameter evaluation, data quality, assumptions, performance measures, calibration, documentation, and retrospective and prospective evaluation. Peer review should be integrated into the planning and modeling timeline with adequate schedules and resources.

This is why the discussion is timely. A proposed timeline might start with the MWG and WQGIT requesting review and drafting review questions for STAC in the first quarter of 2026. Selection of watershed and MBM experts would occur through the third quarter of 2026. STAC would decide on structure and guidance during that time. Model documentation should be delivered to STAC by October 1, 2026, or earlier if drafts are acceptable. This allows STAC to prepare for the Year of Review, which begins October 1. The peer review would start with organizational meetings, followed by multi-day meetings for presentations and follow-ups as requested by reviewers. Two peer reviews would be conducted: one for estuary and one for watershed models. The committee would write peer review reports through a consensus process, and STAC would approve and transmit them to the WQGIT and the MWG. The modeling practitioners would then respond to the peer review point by point, and both the peer review and response would be posted publicly.

Management review occurs in parallel with STAC's technical review. It is less structured, focusing on the appropriateness of model applications and findings for management and environmental conditions in the Chesapeake Bay and watershed. This review is broader, more open, and more collaborative. P7 must be as good or better than P6. If P7 is not approved, we would proceed with P6 for Bay Program needs. Any actual errors, technical inconsistencies, or failures to follow science, or what we call "fatal errors", must be corrected regardless of timing or resources.

Sample peer review questions for the watershed model could include:

- Are relative loading rates and methods for determining them appropriate?
- Do sensitivities to nutrient inputs in P6 reflect our best understanding of nutrient load processing and attenuation?
- Are methods for modeling climate change effects on nutrient delivery appropriate and can they be improved?

Peer review questions for the MBM and MTM might address:

- Appropriateness of organic decay rates and opportunities for improvement.
- Scientific rigor of the water quality transport model representing nutrient and sediment loads from the Conowingo Reservoir under various flow conditions.
- Adequacy of shallow water, habitat suitability, and dynamic living resource model performance.
- Representation of sea level rise impacts.

This process reflects past practice and aligns with Science Advisory Board guidance. At this point, we can conclude the presentation and take questions.

10:00 Discussion of Year of Review Plans

Q from chat: *Olivia Devereux:* If we go ahead with P6, will we add in the new BMPs, new land use categories, new land use, and update the inputs with more recent data? Or keep with the CAST-23 version?

- ***A:*** *Lewis Linker:* It's a great setup, but it's not a question we can answer. It's not in the purview of the MWG. It has a lot to do with the Bay Program's strategic position and what timelines we're operating under and what resources we have to make modifications. New land use categories would require a different calibration. You can't add new land use categories without having a new version of the model.

- ***Comment:*** *Mark Bennett:* P6 is approved as the model of record. Olivia, you're asking about CAST 23, which is a better-than-approved version. But if we go beyond that, does it mean we need to go through another review process? I think we're talking about contingencies. Plan A, and I'd say with 99% probability, is approval of P7, because it is demonstrably and quantifiably better than P6. It would be a convoluted argument not to accept P7 once the MWG completes its work.

- ***Comment:*** *Jeremy Hanson:* Olivia, that's a great point. I know there are a couple of BMPs coming through to the WQGIT in the next few months, so we should at least consider planning versions of those BMPs that could be added to P6.

- ***Response:*** *Olivia Devereux:* Planning versions will be added when the Water Quality GIT approves them, but crediting toward the Bay TMDL will not happen until we have a new version of CAST or a new watershed model.

- ***Comment from chat:*** *Jeremy Hanson:* I had to go check notes to confirm, but part of the Sept 2023 decisions by the PSC was that CAST 2023 would be the last update for P6. Just noting that it would therefore need to be a PSC decision if the partnership wants to change its mind and do another update for P6. [reference]

Q from chat: *Normand Goulet:* How close is CalCAST to being dialed in? In Isabella's last presentation she was still having some issues.

- ***A:*** *Lewis Linker:* it's working reasonably well with P6 land uses. Isabella has taken it as far as possible under P6, but P7 loads and land uses will be the real test. How P7 land use gets incorporated into CalCAST is still to be determined.

10:10 P7 Watershed Model Overview – Gopal Bhatt, Penn State-CBPO

Description: Gopal will provide an updated timeline for completion of the P7 Model in time for the 2026 partnership review.

Gopal Bhatt: In this presentation we will provide an overview of the P7 watershed model structure and the key areas of progress made over the last quarter. Unlike P6, which had two models, P7 consists of three: CalCAST, CAST, and DWSM. CalCAST, a Bayesian statistical model, provides parameters and coefficients for CAST, while CAST generates inputs that inform the DWSM. Together, these models form an integrated system designed to produce the best available loads for downstream estuarine modeling.

The DWSM plays a central role in generating daily load estimates, which are then provided to estuarine models. Because CalCAST operates at the annual scale, the DWSM is used during calibration to validate key watershed processes against observed grab-sample data for nitrogen, phosphorus, and sediment. This ensures consistency and accuracy in model outputs. Progress this quarter has focused on finalizing calibration within the inform–constrain loop and making refinements to watershed process representation.

A second area of focus led by Joseph Delesantro, which is a comprehensive literature review to refine sensitivity estimates for nitrogen and phosphorus. These sensitivities are critical to improving the performance of both CAST and the DWSM. In addition, Joseph Delesantro and Tom Butler have been collaborating with Penn State researchers on machine learning approaches. These techniques are being explored to supplement land-to-water and river delivery factors, as well as to improve downscaling of annual loads to daily time steps within the watershed model.

Finally, there is continued progress on incorporating atmospheric nitrogen inputs. Over the past two quarters, the modeling team has worked with Jesse Bash and others to compile the best available data on atmospheric deposition. This work is feeding into P7 to provide accurate nitrogen inputs for both watershed and Bay domains. These three areas of DWSM calibration, sensitivity analysis, and atmospheric nitrogen inputs, represent the main achievements of the past quarter.

10:35 Progress in P7 DWSM Development – Gopal Bhatt, Penn State-CBPO

Description: *A key theme for the last quarter was an emphasis on the completeness of the DWSM. Gopal will describe key progress areas in completion of the April beta version including incorporation of newly developed beta-parameters describing discharge-concentration relationships, and testing of changes in calibration methods of riverine water quality parameters.*

Gopal Bhatt: We prioritized two major goals this quarter for DWSM: refining the model toward completeness and improving calibration methods. These efforts aim to provide more reliable P7 loads for River Input Monitoring (RIM) stations, which have historically underperformed, thereby supporting both Bay and tributary modeling teams. The DWSM generates loads for the estuarine model, supports calibration, enables scenario analysis, and facilitates research collaborations. Linked with CalCAST, the model integrates monitoring data with process-based techniques to simulate nitrogen, phosphorus, and sediment.

Development of the DWSM began in 2012 with hydrology and expanded to nutrients and sediment. In 2023, refinements linked the model with CalCAST parameters, added small-stream simulation at the finer NHDPlus 100k scale, and improved routing, flow, temperature, and calibration mechanics. The April beta version, completed in May, brought two important updates: replacing North American Land Data Assimilation System (NLDAS) precipitation data with inputs from Virginia Department of Environmental Quality (VA DEQ), improving hydrology, and incorporating a P6 emergent behavior module for nitrate speciation at the land–water interface. Together, these refinements enhanced nitrate delivery estimates and enabled recalibration over a longer historical record (1985–2020).

The team also worked with colleagues at the Virginia Institute of Marine Science (VIMS) to condense roughly 250,000 model output time series into a single NetCDF file. This step streamlined data storage and sharing, while reducing computing demands. With these improvements, the April beta showed stronger performance across both RIM and non-RIM regions. Compared with Weighted Regressions on Time, Discharge, and Season (WRTDS) estimates, P7 results placed average annual flows within 10% of observed values, with efficiency scores trending toward one, indicating strong calibration. Nitrogen, phosphorus, and sediment performance was generally strong, though challenges remain in specific basins such as the Potomac, James, and Conowingo.

Biases persist because DWSM flow is constrained by CalCAST, which uses a longer averaging period than the watershed calibration dataset. Large biases for sediment and phosphorus are often linked to extreme flow events, requiring closer examination of parameter behavior under those conditions. Nevertheless, results show significant progress. The model's finalized step also opens opportunities to refine riverine processes more directly, such as addressing phosphorus underestimation at Marietta and overestimation downstream in the Susquehanna.

Progress this quarter also included further incorporation of beta parameters, particularly for small-stream simulations. These are based on observed relationships between flow and concentration (Q–C relationships) derived from monitoring stations. USGS's FluxMaster tool was used to fit simplified relationships that describe variability with flow, trends, and seasonality. While effective at monitored sites, the main challenge has been extending these parameters to unmonitored streams across the watershed.

To address this, the team-built machine learning models, specifically random forests, with major contributions from Qian Zhang and Isabella Bertani. Isabella compiled watershed attributes and monitoring datasets, while Qian developed models linking nitrogen, phosphorus, and sediment variability to watershed characteristics. With a 70/30 training-validation split, the models performed well and allowed extrapolation of beta parameters across the watershed. Seasonal parameters proved most important for nitrogen, while flow parameters were more critical for phosphorus and sediment. The results showed strong ability to reproduce spatial variability, with nitrogen medians around 0.1 and phosphorus around 0.4, closely matching observed data.

Calibration methods also advanced this quarter. Historically, calibration relied on automated processes using monitoring data, constrained by CAST and CalCAST, while

WRTDS served only as a post-assessment benchmark. A new approach now integrates WRTDS directly into calibration, nudging simulated distributions to match observed loads. Tested in several iterations, like the Choptank River, this strategy improved both scatterplot and CFD agreement. Automated calibration alone resulted in biases as high as +16%, but with WRTDS incorporated, agreement across concentrations, annual variability, and long-term trends strengthened considerably. While sediment results remain mixed and need further vetting, the integration of WRTDS shows promise. For the first time, the DWSM is positioned to deliver complete P7 model loads for both RIM and non-RIM regions to the Bay modeling team.

11:05 Discussion of P7 WSM Development Progress

Q: *Lewis Linker:* This is the first time I've really looked closely at the April results. Based on the current state of the model, I don't think flow has ever been this good. This looks like the best performance I've seen in any version of the watershed model. Is that correct?

- ***A:*** *Gopal Bhatt:* That is my impression as well. I don't have the exact numbers for comparison, but in terms of biases we are very close to P6. Where your point is especially accurate is with efficiency, which measures performance on a year-to-year basis. In that regard, we are doing extremely well, and this improvement carries through to finer time scales as well.
- ***Comment:*** *Lewis Linker:* On both counts, the results are remarkably good. Many basins show biases of less than 1%, which is quite impressive. For example, sediment in the Susquehanna shows a 33% bias, but that will likely be addressed by accounting for special considerations such as sediment dynamics in a large river system and the infill at Conowingo. I expect those refinements will bring it into alignment. Overall, I'm very impressed with the general improvements across the board. Achieving better flow performance will provide significant benefits throughout the entire modeling process.

Q: *Lewis Linker:* TN yield base low index these have important influence on the beta parameters, but what is the 4^{th} – Lithological Magnesium Oxide (MgO). Any idea to what process that might mean? I suppose it might be the amount of lithological magnesium oxide that are around, but is there any light to be shed on that component?

- ***A:*** *Gopal Bhatt:* random forest model is good at finding a relationship if there is some relation to relationship to be found, but sometimes it can lead to kind of head scratcher. And I have to be honest, I had to look this up, so some of the variables behind this graphic are kind of just key terminology that is part of the data set. So, unfortunately, I don't have much to add.
- ***Response:*** *Lewis Linker:* That's fine, and it makes sense. When the random forest model is applied, it shows that lithological magnesium oxide has some level of importance. What that actually means, we don't know, but there seems to be a

regional association. Perhaps it relates to the beta parameters. I understand it's a bit of a black box when applying a random forest model.

- ***Comment from chat:*** *Larry Sanford:* Google says that Lithological MgO is "a key geochemical indicator used to classify and understand the composition and origin of different rock types, particularly in the context of magmatic and sedimentary rocks."

- ***Q:*** *Olivia Devereux:* if this is a little black box, how do you make sure there's not a problem with covariance with another variable? Or does the RF model automate that using a part of covariance?

- ***A:*** *Gopal Bhatt:* We addressed foundational steps upfront, such as cross-checking variables to remove those that were highly correlated and excluding variables with little variance. As Lew pointed out, many of the identified variables make sense, with the exception of lithological magnesium oxide. It may be correlated with another watershed attribute not in our dataset, or it may simply not belong. Still, I take reassurance in the fact that the top three most important variables identified were for total nitrogen, base flow index, and runoff volume. These likely contributed much more strongly than lithological magnesium oxide.

Q: *Lewis Linker:* It sounds like the strategy for sediment is to use HSPF automated calibration with stream routing and beta parameters, followed by any additional calibration needed. This would provide the best representation of sediment. Is that the path forward?

- ***A:*** *Gopal:* Yes, that is exactly right. This isn't full calibration yet, but it establishes completeness by representing missing processes. This becomes the new starting point for further refinements and calibration. HSPF automated calibration, additional hand calibration as needed, and incorporation of stream routing and beta parameters. Applied across the watershed, this should give us as complete and accurate a representation of sediment as possible.

11:15 Progress on P7 Nutrient Inputs and Sensitivities – Joseph Delesantro, ORISE-CBPO with Conor Keitzer and Roshni Nair-Gonzalez, UMCES

Description: *A discussion of the load sensitivity to input values and methods for agricultural fertilizer, manure, and atmospheric deposition to forest will be presented. Also, updates on the sanitary sewer exfiltration estimate method and results will be discussed. This P7 addition will be decided by the Wastewater Treatment Workgroup in July.*

Joseph Delesantro: Sensitivity is defined as the change in export load relative to the change in input load, measured against the mean input across the watershed. Lower sensitivity values result in muted responses, while a sensitivity of zero indicates loads remain constant regardless of inputs. In P6, the sensitivity for true forest nitrogen deposition was set at 0.023, which significantly underestimated the effect of declining deposition. Since forests cover about 60% of the watershed, this discrepancy is important. Observed deposition dropped from 5 in 1985 to 1.5 in recent years, yet CAST captured almost none of this change. Comparisons with other land uses, e.g. harvested forest (0.16), agriculture (0.22), and roads (highest), suggested the true forest value needed reassessment.

To address this, the team conducted a literature review, following recommendations from P6. Led by Roshni Gonzalez and Connor Keitzer, the review examined 43 papers and compiled data from 11 studies, including six field studies and five watershed models. Outliers with significant agricultural or urban influence were excluded, as were extreme values beyond two standard deviations from CAST estimates. Field studies yielded an average sensitivity of 0.09, while watershed-scale models, using SPARROW, produced a mean of 0.16. Both values were substantially higher than P6 and reinforced the need for an update.

This makes sense given that forests, while retentive, often receive deposition during rainfall or high flows, elevating sensitivity compared to soil retention alone. Moving forward, a working group of forestry experts will provide input to ensure the new forest sensitivity value is consistent with other natural load sources. The aim is to finalize and approve this adjustment by the October quarterly meeting.

The review of fertilizer and manure sensitivities followed a similar process but drew from a larger dataset. Values for fertilizer nitrogen and manure nitrogen were slightly lower than those in P6 but still broadly consistent. These results will serve as priors for calibration in P7, either as probability distributions or as central tendencies if timelines and resources require simplification. Phosphorus sensitivities were also reviewed, with CAST's representation, which split phosphorus into extractable, sediment, and stormwater pools, validated against literature that showed lower phosphorus sensitivity compared to nitrogen.

Further refinements are planned, particularly in aggregating phosphorus sensitivities across pools for better comparison with literature values. Crop uptake sensitivity also remains under review and will be a focus in the coming quarter. The Measured Annual Nutrient Loads from AGricultural Environments (MANAGE) dataset has been especially useful in this work, providing extensive field-scale nutrient runoff data and serving as a potential model for databasing future literature reviews to improve transparency and reanalysis.

Additional work was centered on phosphorus transport processes. New datasets have been developed to characterize hydrologic connectivity, biogeochemical controls such as alkaline desorption, and other factors like salinity, road salt, temperature increases, and geogenic sources. While some datasets are ready for immediate use, others, especially alkaline desorption, require more work to represent across space and time. This effort, supported by UMCES, will be a priority for the next quarter.

Finally, progress was made on urban nutrient inputs through modeling sanitary sewer exfiltration. A conservative framework was developed to account for small leaks and defects, with exfiltrated volumes tied to dry weather flow, groundwater conditions, and sewer age. Nutrient concentrations draw from previous expert panel work, while soil and groundwater attenuation methods were modified for urban settings. Early estimates suggest about 100,000 lbs of nitrogen and 10,000 lbs of phosphorus per year in Baltimore and Southeast Virginia, with watershed-wide loads around 2.6 million lbs of nitrogen prior to attenuation. Attenuation is expected to halve this value. In parallel, the Agricultural Modeling Team approved updated crop yields and is advancing fertilizer dataset updates. These refinements together represent important steps toward improving P7 modeling accuracy.

12:15 Discussion of P7 Nutrient Inputs and Sensitivity Progress

Q: *Lewis Linker:* For true forest, the sensitivity will be adjusted up by a factor of 3 whereas manure and fertilizer, there will likely be some down adjustments for nitrogen but I wasn't sure about phosphorus. Lastly, for crop uptake, there will be an increase. Is that correct?

- ***A:*** *Joseph Delesantro:* for the phosphorus, it is too early to say how things will change but we can say that the sensitivity values that we have in CAST are consistent with what we see in the literature. So that will be sort of the final comparison point and that's where we might use that literature review on phosphorus to do the tweaking. It's just the literature just does not break things out by storm water, sediment, etc. It's great that CAST does. But it does mean we don't have that sort of direct one-to-one comparison with the literature like we have for fertilizer nitrogen and manure nitrogen.

Q: *Olivia Devereux:* what is the plan to get sensitivities for the new land uses?

- ***A:*** *Joseph Delesantro:* the sensitivities are for both load sources and input type. We don't really have an answer for that but the status quo, the P6 method was to use the loading ratio to set the sensitivities. For instance, if we split hay into other hay and manage hay, then the sensitivities for those values would be adjusted from the original relative to their new loading rate. It would be proportional to the loading rate. So, in the absence of additional information, that would be the default method.

Q: *Olivia Devereux:* Previously in the model, we have used the averaging multiple models to come up with sector targets. We have had the sector apportionments prior to teasing it out into the specific input in load use sensitivities. Are those sector apportionments for our targeting the same as what we used before or is there a new source data?

- ***A:*** *Joseph Delesantro:* In terms of sensitivity, I would say that we would largely expect them to be consistent because we're only evaluating these sort of select sensitivities as were deemed important. For the agricultural sensitivities, the sensitivities are going to be defined in the final model by both load source and input type. So, the way that those sensitivities will vary across all the different cropland load sources will be based on that loading ratio. But that is sort of downstream of that apportionment.

- ***A:*** *Lewis Linker:* That's a good response, Joseph. This refinement goes beyond what we had in P6, where we didn't have the time for an extensive literature review into crop yields or loading by land uses. P7 has more detailed sensitivities. For example, the forest loading rate appears to be more than double, based on 47 pieces of literature. This reflects the modeling workgroup's incremental improvements. We're a little better in P7, and Joseph's careful work gives us confidence. He has added important analyses such as sewage exfiltration loads. This work is in the same spirit as previous multiple-source approaches, but with the expertise and time in P7, we can re-examine sensitivities and loads in more detail.
- ***Q:*** *Olivia Devereux:* In the last version, we had sector-level calibration targets, and I was wondering if those changed. The relative loading rates and sensitivities affect what happens within a sector, but unless the sector-level allocations change, the loads should remain within that sector. That's what I'm trying to clarify. Gopal, do you have another perspective on this?
- ***Q:*** *Gopal Bhatt:* the answer is detailed and may be best for an ad hoc discussion. I also had a question for Joseph. On that last slide showing agricultural yields, could we use this more sophisticated approach to crop uptake, rather than the census yields we used before? This matters because nutrient fate on cropland is largely determined by crop uptake, so getting this right is critical.
- ***A:*** *Joseph Delesantro:* There are two components being discussed in that slide. First, we have a "bucket" of fertilizer, and where it ends up depends largely on crop yields, with some influence from local manure production. Crop yields define the spatial distribution of fertilizer across the watershed, so we're updating both the fertilizer bucket and the spatial driver of application. The green line shows P6 yields for corn grain in Virginia. The black dots are observed USDA county yields, and the open dots are modeled county values. The red line and red points represent yield expectations that farmers use when deciding fertilizer applications, often an optimistic "best three out of five years" perspective. We also developed a long-term weather-independent yield (orange line) by smoothing out climate variability in the model. This approach, informed by expert input, has been approved by the AMT for P7.
- ***Q:*** *Lewis Linker:* If I understand correctly, most of this representation already existed in P6, but P7 incorporates new crop yields approved by the Agricultural Modeling Team. Is that right?
- ***A:*** *Joseph Delesantro:* Yes. Fertilizer distribution based on crop yields remains the same as P6, but in P7 we will use the red points instead of the green points.

- ***Comment from chat:*** *Olivia Devereux:* Crop yields come from NASS, Ag Census.

12:25 LUNCH

1:00 [Future Environmental Conditions and CBP BMP Efficiencies](#) – Maya Struzak, Michelle Miro and Krista Grocholski, RAND

<u>Description:</u> *Progress will be presented on application of the APEX and SWMM public domain watershed and stormwater models, respectively, under different future climate hydrologic conditions to determine potential change in relative pollutant removal efficiency of current CBP-approved NPS and stormwater management BMPs.*

Maya Struzak: The project focuses on quantifying the impacts of future climate conditions on Chesapeake Bay watershed management, specifically on the efficiencies of agricultural best management practices (BMPs). The team is modeling pollutant removal efficiencies across four land uses and four physiographic regions, yielding 16 baseline scenarios. These scenarios are forced with different hydrologic regimes, beginning with a baseline climate period (1985–2000). Site characterization has been largely completed using real-world data sources, including CBP land segment data, USGS well data, CBP weather datasets, and agricultural practices derived from CAST and USDA census data. Baseline (no-BMP) scenarios were created for each region–land-use combination, forming the foundation for comparison with future BMP applications.

During the proof-of-concept phase, the team has been validating hydrologic outputs against CAST model values to ensure accuracy of the water balance components. However, progress was delayed due to a critical error in the APEX source code. A logic flaw caused the model to generate random weather data when hourly rainfall inputs were used alongside daily ancillary variables. Since hourly precipitation data are essential, this issue disrupted the workflow. With assistance from the APEX forum, a solution was implemented to ensure all daily variables are properly read, though the team is still validating this fix. Achieving accurate baseline simulations is crucial, as BMP removal efficiencies are calculated relative to these baseline conditions.

Baseline generation involves setting the physiographic region, creating baseline files, and then defining BMP operations. Validation to P6 outputs is ongoing, and most hydrologic parameters are already aligning closely with CAST, with water-balance residuals below 1%. Early results demonstrate the relative efficiency of BMPs; for example, applying cover crops to row crops reduced water yield, sediment, nitrogen, and phosphorus export compared to the baseline. These comparisons illustrate how removal efficiencies are expressed as percentage reductions in pollutant outputs.

The team has now completed baseline schedules, operation schedules, and the full set of files for all 16 scenarios. Following the debugging phase, efforts are focused on refining and validating baselines before expanding to BMP modeling and post-processing. BMPs considered so far include cover crops, no-till practices, and manure incorporation for row crops, identified as relatively straightforward to implement. However, there is still uncertainty about the feasibility of BMPs in other land uses such as hayland, pasture, and forest. The ad hoc group will be engaged to refine BMP selection for these categories.

Looking ahead, once baseline validation is complete and BMP operations are fully functional, the team will conduct removal efficiency analyses and advance to future climate modeling. The target year for climate scenario projections is 2055. While current work relies on historical climate data, this next stage will enable exploration of how management practices perform under changing climate conditions.

1:20 Discussion of CBP BMP Efficiencies Under Changing Environmental Conditions

Q: *Lewis Linker:* For the baseline versus cover crops figure, the 1% removal appears to refer to water yield. The Y-axis indicates water yield in millimeters. For the next Y-axis, this likely represents nitrogen in runoff (parts per million), not groundwater. So, the 34% nitrogen removal with cover crops reflects runoff nitrogen only. That seems high, perhaps we need to integrate across the full time series, since a large end-of-period peak may dominate. In any case, this is a proof-of-concept demonstration. We will refine and then explain the drivers of these results. Phosphorus shows 0.00% removal; noted.

- ***A:*** *Maya Struzak:* Yes. In APEX, water yield is total water leaving the subarea (mm). The column is nitrogen in surface runoff (ppm), not groundwater.
- ***A:*** *Sarah Fakhreddine:* To underscore baseline importance: Maya is using phosphorus and nitrogen inputs from our data, and we're seeing phosphorus limitation in crops, which explains the 0% phosphorus removal - there's no excess P in the baseline, which we know isn't realistic. We'd appreciate guidance on data inputs to refine baselines. We also have the BMP efficiencies you provided, which we can use for calibration. One caution: CAST evapotranspiration uses Hamon, which is acceptable for baseline, but for climate scenarios we should use methods with appropriate climate sensitivity (e.g., Gridded meteorology), since Hamon is heavily temperature-weighted. We'll keep that in mind for later meetings.

Q from chat: *Clifton Bell:* One of my standard questions on this: For representing climate change effects on vegetative BMPs, do you plan to adjust anything besides meteorology? e.g., growing season, crop growth rates or N uptake rates?

- ***A from chat:*** *Olivia Devereux:* I think the critical period remains 1991 to 1993 and the average hydrology is 1991 to 2000. Gopal can confirm, but the precipitation years are not changing either.
- ***A from chat:*** *Sarah Fakhreddine:* Correct, we only plan to do meteorological forcing for climate in current project scope. It could be interesting for future work.

- ***Comment from chat:*** *Clifton Bell:* Thanks for the response. You're probably already familiar with some of the background on that question. [reference]

- ***A:*** *Gopal Bhatt:* Thanks Clifton for that question. As Olivia has pointed out, the AHP is still 1991-2000 and 1993-1995 is the critical period but for future changing environmental conditions (CEC) scenarios we alter meteorological inputs and capture hydrologic effects on nutrient and sediment runoff. That said it would be good to have information on other sensitivities that have direct, well understood, effects on nutrient and sediment runoff. The thinking has been that the APEX model currently being developed could be used for asking those sorts of questions.

Q from chat: *Normand Goulet:* When are we going to have a conversation about changing this – the critical period? Both James and I have brought it up in the past and we need to take a good hard look at these dates.

- ***A:*** *Lewis Linker:* This was a topic of the recent workshop on changing environmental conditions, which has just gone through final STAC reviews. It was directly addressed there, and the suggestion is to begin work on this issue, though not as part of the P7 effort. The loads delivered to the Bay are very much tied to hydrology, and hydrology is determined by the period chosen across all partners from New York to Virginia, West Virginia, and Delaware. This is a major technical and policy decision. The modeling workgroup, in its current reduced capacity, doesn't have the bandwidth to take this on. The recommendation is to begin as soon as possible and make it part of the next assessment, perhaps Phase 8.

- ***Comment:*** *Mark Bennett:* It seems more like a policy issue than a technical one. There have been many recommendations from STAC and others to move away from a static critical period, but a policy discussion is needed to replace what we currently have.

- ***Response:*** *Lewis Linker:* Maybe the best next step is to make this a STAC workshop focused on developing recommendations for a replacement approach. Technically, it could be done, but the policy implications make it more complicated. Should it be synthetic, a more recent period, or a longer period? Each option has different consequences.

- ***Comment:*** *Richard Tian:* Another point: the current critical period has never been changed. If we propose to change it, we might need to consult counsel about whether that means we're altering the TMDL itself, which could raise additional challenges. The TMDL framework is very resistant to changes of this kind. This is another layer of complexity.

1:30 Development of Efficient Multi-Objective Optimization Procedures – Ritam Guha, Kalyan Deb, Pouyan Nejadhashemi, and Auden Garrard, MSU

Description: *Progress on the CBP optimization project and plans for upcoming webinars will be presented.*

Ritam Guha: The current focus is to align the optimization system more closely with the CAST framework. Following the approval of a one-year no-cost extension, the revised work plan emphasized Task 4, which includes improving the optimization dashboard, developing surrogate-assisted optimization, and advancing toward robust optimization and sustainable watershed management. Task 3 cleanup work is also being finalized. A major enhancement this quarter was the ability to optimize on top of existing BMPs using historical implementation data from 1982–2023, providing users with more realistic scenarios. Land conversion and efficiency BMPs have been added to the system, while animal and manure transport BMPs are still under development.

A significant portion of work focused on improving the dashboard to mirror CAST, making it more intuitive for existing CAST users. At the same time, progress was made on surrogate-assisted optimization, which uses machine learning to approximate CAST evaluations. This approach enables rapid evaluation of thousands of BMP implementation combinations, filtering out poor solutions and focusing computational effort on the most promising ones. Although surrogate models speed up optimization, all final solutions are still validated with CAST to ensure accuracy. This dual approach improves both efficiency and accuracy as the system evolves.

During recent optimization runs, discrepancies were found between system outputs and CAST data, particularly for wastewater, phosphorus, and sediment. The team identified missing wastewater treatment plant data and corrected it by aggregating CAST source files. They also incorporated the base progress scenario into their framework, which is essential for aligning with CAST processes. Additionally, budget-constrained optimization was added to prevent the generation of infeasible solutions, such as those exceeding cost thresholds (e.g., $40 million). This new feature allows the system to deliver only realistic, budget-compliant solutions.

Other improvements included the addition of a process status bar, enabling users to track optimization phases in real time. The team is exploring email notifications to alert users of progress, as well as options for advanced cost customization, result sharing, and deleting outdated runs to streamline system use. These enhancements improve usability and support collaboration across project partners.

Looking ahead, the team plans to complete integration of animal and manure transport BMPs by October, which will finalize the inclusion of all major BMP categories. They will then convene with the advisory group, including members of the modeling workgroup, to plan a webinar demonstrating the updated process. The first webinar is targeted for December or January to share results and capabilities with stakeholders.

1:50 Progress in P7 Atmospheric Deposition Loads – Gopal Bhatt, Penn State-CBPO

Description: *P7 atmospheric deposition loads are being developed to leverage newly available information from CMAQ and GCAM models while considering previously* *developed estimates of wet deposition from a statistical model and PRISM precipitation. Gopal will provide a progress update on the work for developing atmospheric deposition loads to the WSM and MBM domains.*

Gopal Bhatt: Our goal is to share some results from new datasets and outline next steps. The task is to develop input loads for P7 models that can support calibration (1985–2014), trend analyses for past and future scenarios, and management applications. Key datasets include the P6 air shed model, new CMAQ hindcasts (2002–2019), and GCAM projections for 2035 emissions.

Deposition estimates from P6 captured variability due to both meteorology and emissions. Trends were then extracted by removing meteorological variability, creating detrended loads at fixed emission levels. Long-term averages (1991–2000) were used to calculate trended loads. This process was repeated across all counties and watershed segments, showing how deposition has shifted over time from being the largest nitrogen source to one of several major contributors.

P7 advances this work by incorporating refined CMAQ hindcast data at 12 km resolution, which cover the watershed and coastal ocean, and by including GCAM projections of emissions under 2035 conditions. Results indicate that overall atmospheric deposition is expected to decrease by 2035 compared to both 2016 and 2030 estimates, primarily due to reductions in dry oxidized nitrogen deposition. However, dry reduced nitrogen is projected to increase slightly. These improvements provide more spatial detail and better capture current emission patterns.

A major correction in P7 addresses an assumption from P6 about dry deposition of ammonia. Previously, estimates before 2002 were kept constant due to lack of credible relationships. New analysis, supported by Jesse Bash at EPA, identified a strong inverse linear relationship between oxidized and reduced dry deposition, allowing more realistic reconstruction of historical ammonia deposition. This correction lowered estimated deposition in the early calibration years, aligning better with observed emission trends.

The integration of datasets involved combining P6 long-term trends with CMAQ hindcasts and extending forward using GCAM 2035 projections. Linear fits were applied for consistency, and piecewise linear trend lines now extend from 1985 through 2035. Comparisons show P7 results differ somewhat from P6 but integrated watershed totals remain similar. The updated projections suggest nitrogen deposition loads will be 5 million lbs lower by 2030 and up to 20 million lbs lower by 2035 than previously estimated.

For the Chesapeake Bay domain, P7 harmonized P6 data with CMAQ grid cells (248 at 12 km resolution, including 55 tidal Bay cells). This created consistency between watershed and Bay analyses. Early results show the same general patterns observed for the watershed, though work is ongoing to incorporate PRISM precipitation data to further refine Bay load estimates. These steps ensure deposition estimates are comprehensive across land, tidal Bay, and coastal ocean areas.

In conclusion, P7 now provides a blended atmospheric deposition dataset that integrates P6, CMAQ, and GCAM information. It covers the calibration period, extends hindcast data through 2019, and projects forward to 2035 under IRA policy scenarios. Next steps include completing PRISM-based Bay deposition estimates, extending analysis to the coastal ocean, finalizing QA/QC, and delivering the dataset in netCDF format for modeling applications.

2:20 Discussion of the Merged Atmospheric Deposition Data Set

Comment: *Lewis Linker:* the model back in the 90s was just not that good. By P6, we see the results are really quite good and are constantly improving.

Q: *Dave Montali:* with respect to the loading to the tidal bay, we discussed how dry deposition doesn't happen in open water. I am curious about the tidal stretches from the Main Bay to the Fall line because some of your cells will be a mixture of land and water. How are we going to define the deposition, for instance, way up the Rappahannock.

- ***A:*** *Gopal Bhatt:* The team faced the same challenge when developing P6 data. In P6, dry NOx deposition estimates for the Bay and tributaries were complicated because 12 km CMAQ cells often spanned large areas of land and water. Since deposition rates on land are much higher due to surface friction, raw CMAQ output was not suitable for representing open water. To address this, the team developed a spatial relationship linking each mixed land-water cell to the nearest pure water cell, and applied the water-based value consistently along tributaries. This method was built into the P6 dataset, which serves as the training dataset for P7, so the estimates for 1984–2014 already account for this correction.

2:40 ADJOURN

Attendees:
Richard Tian (UMCES), Lewis Linker (EPA), Jeremy Hanson (CRC), Mark Bennett (USGS), Petra Baldwin (CRC), Olivia Devereux (Devereux Consulting), William Keeling (VADEQ), Samuel Canfield (WVDEP), Normand Goulet (NVRC), Kevin Mclean (VADEQ), Bo Williams (EPA), Larry Sanford (UMCES), Scott Heidel (PADEP), Guido Yactayo (MDE), Joseph Zhang (VIMS), Roshni Nair-Gonzalez (UMCES), Arianna Johns (VADEQ), Wenfan Wu (VIMS), Andy Fitch (USGS), Cathy Wazniak (MDDNR), Clifton Bell (Brown and Caldwell), Gopal Bhatt (Penn State-CBPO), Jian Shen (VIMS), Joseph Delesantro (ORISE-CBPO), KC Filippino (HRPDC), Dave Montali (Tetra Tech), Mukhtar Ibrahim (MWCOG), George Onyullo (DOEE), Sophia Grossweiler (MDE), Joseph Zhang (VIMS), Zhengui Wang (VIMS), Karinna Nunez (VIMS)

Modeling Workgroup July Quarterly Review

Day 2 – July 9, 2025

Event webpage: Link

==

Table of Contents (jump to bookmark):

Day 2 – July 8th, 2025:

Back to Day 1.

10:00 Announcements and Amendments to the Agenda – Dave Montali, Tetra Tech and Mark Bennett, USGS

10:05 Update on Main Bay Model (MBM) Progress – Zhengui Wang, Joseph Zhang, and Jian Shen, VIMS

Description: *Progress in development of the MBM will be presented. The MBM is being readied for initial climate and management sensitivity scenario testing in the coming quarter.*

Zhengui Wang: The past two quarters have seen important advances in coupling the P7 Watershed Model with the MBM. Earlier in the year, the team received the first hybrid version of P7 watershed loading, which used P6 data for the RIM stations and P7 outputs for other tributaries. This dataset was converted into a database, allowing a detailed assessment of new loadings across major rivers and smaller tributaries. Using these inputs, the MBM was set up and tested, producing preliminary results for hydrodynamics, water quality, primary production, and sediment nutrient fluxes.

In the most recent quarter, the focus shifted to refining and calibrating the MBM, with the goal of producing the first calibrated baseline version using P7 loadings. A major achievement was converting the watershed loading data from ASCII to netCDF format, which drastically streamlined processing. Whereas the old ASCII workflow required managing over 220,000 separate files, the netCDF format consolidated all nutrient data into a single structured file. This eliminated the need for post-processing, simplified transfers, and improved organization, with similar conversions underway for shoreline erosion and atmospheric deposition inputs.

Another milestone was the release of the April beta version of P7 watershed loading. Comparison with the January beta showed relatively minor differences in total flow and nutrients, but a sharp increase in nitrate and an 11% decrease in organic nitrogen. These differences, though less impactful in major rivers still tied to P6 loadings, were significant in smaller tributaries and embayments. To assess performance, the team evaluated 16 embayments, where observations are scarce. Results showed that nitrate and phosphorus simulations improved under the April beta, especially in systems such as the Gunpowder, Nanticoke, and Chester, giving confidence in the refinement of the new loadings.

Calibration of the MBM proceeded with a two-step setup: first hydrodynamics, waves, sediment transport, and SAV effects; then water quality, sediment diagenesis, and new resource modules. Performance has been generally satisfactory where salinity and temperature simulations are strong, while water quality metrics are on par with or slightly better than P6. Notably, bottom DO and hypoxia simulations have improved. At the same time, error analysis revealed larger deviations in tributaries than in the mainstem, highlighting both the challenge and the opportunity for the MBM to provide better shallow-water resolution than its predecessors.

The workflow itself has been substantially modernized. By shifting to a Python-based system, the code length was reduced by half, improving efficiency, flexibility, and maintainability. Users can now toggle modules, such as SAV or sediment adjustments, and launch simulations on HPC systems with ease. Supporting databases for initial conditions and parameters have also been created, eliminating the need for lengthy spin-ups and improving accuracy for sediment nutrient states. These innovations reduce computational costs and improve the reliability of results.

Evaluation of sediment processes and biological modules showed encouraging outcomes. Sediment oxygen demand (SOD) simulations captured expected seasonal dynamics under both oxic and anoxic conditions. Nutrient fluxes, such as ammonia, were also well represented. Long-term sediment organic matter concentrations stabilized after a few years, confirming model stability. Early results from new resource modules indicated reasonable seasonal cycles for SAV and oysters, though biomass levels remain underestimated and will require further data for validation.

In summary, the past two quarters have delivered major progress in advancing the MBM with P7 watershed loadings, streamlined workflows, and enhanced calibration. The April beta watershed loading demonstrated improved performance in smaller tributaries, while the MB model now integrates physics, biology, and new resource modules in a flexible, efficient framework. Results are generally on par with or better than P6, with improvements in shallow-water simulations and sediment coupling. Remaining challenges include refining resource modules and validating biomass simulations, but the team is on track to release the first calibrated MBM in the coming months.

10:45 Discussion of the Main Bay Model (MBM) Progress

Comment: *Lewis Linker:* Streamlining these processes is great to see. I did want to mention that for spin-up we need to test some scenarios. For example, if we compared an "all-for" scenario to the base, that might result in an order of magnitude reduction in nutrients. In a no-action scenario, the reduction might only be by a factor of 4 or 5. This means we can really test scenarios, including future climate scenarios, that could significantly change sediment loading. It is always good to spin up under these extreme conditions to confirm the model behaves as expected; otherwise, we must be cautious with spin-up times when conditions change substantially. For sediment diagenesis, conditions usually stabilize after about five years. For next quarter, we should not only test spin-up conditions but also look at small embayments set up with Gopal. Specifically, we should examine differences in outcomes with tidal wetlands off versus on, focusing on variables like DO. Similarly, we should test oysters, SAV, and tidal wetland feedbacks. Sensitivity testing over the next quarter would allow us to assess how living resources influence water quality. It may also be valuable to extend this to larger areas, such as DC stations like CB4, to evaluate changes in DO and nutrient stocks. Overall, the processes developed so far will allow us to perform this testing more easily moving forward.

- ***Comment:*** *Richard Tian:* I found that with the back calculation that I wished at the end of the simulation, the code can dump out the files. In some cases I found that they were not saved at the end of the run, only at specified time steps. The system automatically dumps everything, including segment conditions, at the end of a run. This allows them to be used as initial conditions after spin-up.
- ***Response:*** *Zhengui Wang:* In SCHISM, both the physical and water quality models generate hot-starts, usually every one to three months.
- ***Comment:*** *Lewis Linker:* we have the capability of running SAV in the hydrodynamic sediment model. However, for efficiency, sediment loading must also be included in the water quality model. We can run scenarios with sediment embedded in water quality, and for final scenario runs we can combine everything into a comprehensive setup. Running the full set of processes is time-consuming, so we need to plan carefully.

Comment: *Lewis Linker:* Achieving bottom DO targets in places like CB4 or CB4.4 depends on all of the other processes in the model. I've noticed in both 1992 and 1996, that the current MBM appears slightly later in the onset of hypoxia compared to the 3D model, and it also seems to hold on to hypoxia a little longer. This may actually be more appropriate than what we see in 3D model. It is something we will need to monitor closely as we continue, since deep-water DO will play a major role in the outcomes. That said, the calibration looks very satisfactory at this stage. Regarding SAV, the penultimate slide is important because SAV presence is itself a water quality standard. Over the coming quarters, we all need to focus collective attention on understanding and improving SAV calibration. This includes ensuring that Zhengui has the data required to refine the model. Better calibration of SAV will be an important area of work for the team moving forward.

10:55 Improving MBM Simulation With Physically Based Calibration Approaches – Wenfan Wu, Zhengui Wang, Joseph Zhang, and Jian Shen, VIMS

Description: *An innovative calibration approach to incrementally improve the salinity calibration by assigning roughness to observed bottom types of clay, sand, and gravel and also improving calibration of temperature throughout the water column though the interaction of shortwave radiation and turbidity will be presented.*

Wenfan Wu: Recent work on calibrating the hydrodynamics of the MB model has focused on a physically based approach that incorporates observed parameters of turbidity and sediment types. This calibration strategy demonstrated clear benefits, as turbidity improved temperature simulations in the upper Bay and tributaries, while sediment information enhanced tidal propagation and saltwater intrusion. Overall, the method reduced RMSE for temperature and salinity by roughly 16% compared to earlier Bay modeling studies. The defensible approach also strengthens confidence in the MBM's ability to simulate complex hydrodynamic processes.

The MBM spans a wide domain that includes tributaries, the mainstem, and the coastal ocean, while resolving many small-scale topographic features such as creeks and tributary heads. Incorporating these elements requires high-resolution grids and accurate bathymetry, which were achieved by integrating multiple datasets and minimizing smoothing. These detailed structures drive complex dynamics like tidal amplification and resonance, which are strongly influenced by bottom friction. Turbidity and sediment variability further shape hydrodynamics, underscoring the need to move beyond simplified assumptions of constant conditions used in earlier studies.

To test the role of these factors, six calibration runs were designed. The base run applied a moderate, commonly used bottom drag. Runs 2A and 3A examined turbidity effects: one with uniformly high turbidity, and one with spatially varying turbidity derived from satellite Kd data. Runs 2B and 3B tested sediment effects, contrasting constant versus spatially varying Manning's coefficients tied to sediment type. The final experiment, run 4A, combined both turbidity and sediment variability to create a fully calibrated scenario.

The experiments revealed clear patterns. The base run showed significant temperature errors in surface layers and tributary heads, along with salinity errors in bottom layers and upstream stations. Adding uniform turbidity reduced overall bias but did not address tributary-specific issues, while spatially varying turbidity corrected both. For salinity, sediment-dependent roughness produced the strongest improvements, particularly in bottom layers, by more accurately simulating tidal amplification and salinity intrusion. The fully calibrated run (4A) provided the best overall performance, reducing errors across all depths and locations.

Comparisons with P6 confirmed that the calibrated MBM offered notable advancements. Error maps highlighted widespread improvements across the system. This outcome demonstrates the value of using realistic, spatially varying physical parameters rather than uniform approximations. The calibration also shows the importance of integrating bathymetric detail, turbidity, and sediment variability to strengthen cross-scale simulations of Chesapeake Bay hydrodynamics.

While progress has been significant, the study also identified key areas for future development. Seasonal variability in water clarity and bottom drag has not yet been incorporated, nor have processes like sediment stratification, which can reduce turbulence and alter bottom drag. Additionally, SAV and tidal velocity effects, both known to influence hydrodynamics, remain to be added. Addressing these gaps will further refine the MBM, enhancing its accuracy and utility for simulating the Bay's dynamic system.

11:15 Discussion of Physically Based Calibration Approaches

Q: *Carl Cerco:* Do the improved predictions of salinity and temperature in the hydrodynamic model impact the water quality model? Are temperature and salinity values passed directly to the water quality model, or the water quality model calculates them separately?

- ***A:*** *Joseph Zhang:* Most water quality processes are closely linked to temperature, especially in rivers and shallow areas. Improvements in salinity and temperature in those locations lead to improved water quality simulations. In our two-step setup, temperature and salinity are calculated in step one and then passed to step two, which handles water quality. Step two does not calculate them independently.

Q: *Jiabi Du:* When using satellite-derived turbidity to define water types, are temperature improvements due to changes in light penetration, reflection, or both?

- ***A:*** *Wenfan Wu:* We used a spatially constant albedo, so reflection did not change. The improvements are due to how turbidity affects light penetration into the water column, not reflection.
- ***Response:*** *Jiabi Du:* That makes sense. I also appreciate the improvements in tidal range in the Choptank River, which previously had underestimated salinity intrusion. This is a promising direction, and experimenting with different Manning's coefficients may provide additional improvements.

Q: *Lewis Linker:* The P6 model performed reasonably well for temperature and salinity, but even small improvements are meaningful. For example, a 1°C reduction in temperature error represents roughly a 10% increase in accuracy, which doubles reaction rates every 10°C. Improvements in bottom salinity are particularly important since they reflect circulation and mixing. I am thinking about implications, does adjusting bottom friction increase or decrease gravitational circulation?

- ***A:*** *Wenfan Wu:* Lower bottom roughness increases bottom water intrusion, but turbulence and vertical mixing are the dominant factors. Weakening mixing promotes stronger bottom intrusion. Bottom roughness contributes, but mixing plays the larger role.

- ***A:*** *Jiabi Du:* I believe improvements are more notable in tributaries, since small changes in roughness have larger effects in shallow waters, while deep Bay areas are less sensitive.
- ***Q:*** *Dave Montali:* Can you say anything about how this improvement might affect our results of future conditions? I expect increased sea level will strengthen gravitational circulation. Would that amplify or dampen future DO projections under this calibration?
- ***A:*** *Wenfan Wu:* If spatially varying roughness does not account for mud layers, salinity intrusion may still be underestimated. With sea level rise, I believe salinity intrusion will strengthen, particularly in tributaries, making upstream intrusion easier.

Comment: *Lewis Linker:* There's pretty fierce processing of nutrients in these regions of high turbidity and getting an improvement of the temperature correction of the reaction rates through this process is going to be helpful. Thank you!

11:25 Test of Attainment Procedures – Richard Tian, UMCES-CBPO

Description: A preliminary, initial assessment of water quality attainment in the Chesapeake using the P7 MBM will be presented as a test of criteria assessment procedures.

Richard Tian: The preliminary criteria assessment work focused on testing the new P7 model through ten scenarios was designed to explore the assessment procedure, including climate-change cases. The WIP reduction scenario reduced atmospheric deposition to levels projected for 2030 and adjusted nitrogen loadings consistent with P6 calculations. Specifically, watershed nitrogen was reduced by about 11%, atmospheric deposition by 5%, and additional changes from the watershed model were applied. Climate Change Scenario 1 introduced increases in air temperature, sea level, river flow, and river temperature, with open-ocean boundary conditions scaled accordingly. These adjustments were informed by ensemble projections, long-term trends, and established methods for estimating sea-level rise.

To evaluate model skill, hypoxic volume in the Bay's deep channel was compared between observations, P6, and P7. Using interpolated observation data and model results, the tests showed that P7 outperformed P6 in simulating hypoxic volume. At the 1 mg L^{-1} threshold, P7 aligned more closely with observations, while at 3 mg L^{-1} it successfully captured specific hypoxic events. Early comparisons for open-water conditions also confirmed strong performance, providing confidence that P7 better represents Bay dynamics, even though it predicts lower water temperature increases than older models like ICM.

Scenario results highlighted important differences between drivers. Under the WIP scenario, hypoxic volume decreased modestly, while increased flow slightly elevated hypoxia and nutrient load increases produced much larger effects. The ensemble climate-change scenario combined these drivers, showing further increases in hypoxic volume. Decomposition of "heat" effects revealed that reduced oxygen solubility explained most of the response (~53%), followed by smaller contributions from biological rates and stratification. These findings are consistent with results from CH3D, suggesting that both models produce coherent and physically reasonable outcomes.

Assessment tables for the deep channel showed high correlation between different methods of calculating attainment, despite differences in physics between calibration and scenario runs. Importantly, P7's WIP scenario indicated higher non-attainment in specific regions, such as CB4, compared to CH3D. While this may appear as a negative outcome, it reflects P7's more realistic representation of localized hypoxia patterns. For open water, results were again consistent: P7 captured hypoxia and related processes better than P6, while new indicators such as summer chlorophyll and Kd light attenuation performed reasonably well, though some Kd files showed missing bottom-layer data that will need correction.

Overall, P7 represents a clear advancement over P6 and CH3D. It produces more realistic hypoxia simulations, even though it predicts smaller climate-driven temperature increases. This leads to localized areas of greater non-attainment but enhances fidelity to observed conditions. These initial assessments provide a strong foundation for future work, which will include improving bottom roughness and water-quality processes, extending runs to cover longer periods for comparison against long-term trends, and expanding analyses into under-studied tributaries like the Sassafras.

In conclusion, while preliminary, the P7 criteria assessment results are promising. The model successfully integrates nutrient reductions and climate-change drivers into a defensible framework, improves representation of hypoxia, and offers more nuanced insight into localized impairments. With continued refinement and broader scenario testing, P7 is positioned to provide more accurate and policy-relevant assessments for Chesapeake Bay water quality under future conditions.

11:45 Discussion Attainment Procedures Test

Q: *Lewis Linker:* Why does P7 predict lower warming than CH3D under the same forcing?

- ***A:*** *Richard Tian:* I don't have an answer for that but still, the results are promising: P7 appears to better capture hypoxia dynamics and gravitational circulation, while reducing reliance on climate-driven temperature effects.

11:55 Conowingo Model Development – Earl Hayter, Jodi Ryder, CoE-ERDC and Matt Rowe, MDE

Description: *Progress in development of the Conowingo Model will be presented.*

Jodi Ryder: A key milestone this quarter was the long-awaited sampling event at Conowingo, which took place at the end of June after access and weather-related delays. Two types of samples were collected: particle size distributions to support shear stress and erosion assessments, and biological samples for sediment carbon fractionation (G1, G2, G3).

The carbon analysis focuses on evaluating the G1–G3 fractions. One sample contained large debris such as mussel shells and stones, which provided a "sacrificial" sample for initial testing. The fine, mucusy sediment texture and strong odor matched expectations for material collected downstream of the reservoir. A BOD5 trial was launched to measure oxygen consumption rates over 90 days, ensuring oxygen depletion remained within safe bounds. If oxygen consumption proved too slow, nutrient seeding was planned. This trial concludes today and will help finalize the parameters for the long-term incubation.

The sampling covered four sites, each with replicates, using the Ekman dredge method. This successfully targeted the top 10–20 cm of sediment without disturbing deeper layers, providing representative material. Nitrification inhibition has been applied, and the runs are expected to last through October 7 if the full 90-day schedule is maintained. Depending on oxygen levels, testing may extend beyond this date.

Because the laboratory results will not be ready until October or later, the team plans to use G1–G3 fractions from Palinkas (2019) for initial parameterization. These values will serve as a placeholder for both water quality and sediment processes, with refinements made once the new Conowingo-specific results are available. Preliminary drafts of these parameterizations are already underway, allowing progress to continue in parallel with the ongoing incubation trials.

Looking forward, once the 90-day tests conclude in early October, data processing will begin. In parallel, efforts are underway to transition the model from its current PC-based setup to an HPC (high-performance computer) environment. A software contract is in place to configure the supercomputer for this work, which is expected to delay the schedule by about a quarter. To minimize disruptions, some tasks are being held until the HPC system is operational, with the expectation that progress will accelerate once the transition is complete. Overall, the Conowingo sampling effort has provided a solid foundation for the next phase of model refinement and calibration.

Earl Hayter: Recent sediment sampling included grab samples collected with a Ponar grab for grain-size analysis, as well as larger volumes gathered in five-gallon buckets for erosion testing with the Sedflume. At four sites, sediments were consolidated into cores for varying durations before being tested in the Sedflume to determine erosion rates under different bed shear stresses. These updated erosion rates, reflecting both grain-size distribution and bulk density, will replace the outdated values currently used in the sediment transport model. Importantly, this study provides fresh data on how mixed-grain materials of sand and fines respond to stress, offering a more accurate foundation for modeling.

The new erosion data will be directly incorporated into the sediment transport model. To prepare, a multi-month simulation is underway to test long-term model performance, with an emphasis on sediment conservation. The model simulates two fine-grain size classes along with five coarser classes, tracking their transport and ensuring conservation of both individual classes and overall sediment mass. This long-term trial serves as a key test of the model's accuracy in handling sediment dynamics and its ability to maintain both sediment and water mass balance. Results from this exercise are expected within the next couple of weeks.

Currently, these simulations are running on a workstation, but hardware limitations constrain performance, particularly in memory-intensive operations. The HPC will allow use of MPI parallelization, which is not feasible on the workstation. This shift is expected to accelerate simulations by at least a factor of five, significantly improving efficiency and enabling more comprehensive modeling once confirmed through testing.

In summary, the new Sedflume data will modernize erosion rate inputs, strengthening the sediment transport model with site-specific, updated information. Coupled with the upcoming move to HPC resources, these improvements will enhance both the accuracy and efficiency of simulations, positioning the project for more robust long-term sediment modeling.

12:25 LUNCH

1:00 Progress on Patapsco/Back MTM – Harry Wang, Breanna Maldonado, VIMS and Jeremy Testa, UMCES

Description: *Progress on the Patapsco/Back MTM water quality calibration will be discussed.*

Harry Wang: The presentation began with a characterization of P6 hybrid 6-7 nonpoint source loading in the Chesapeake Upper Bay domain. Major contributions come from the Susquehanna River, with smaller but still important inputs from shoreline tributaries. These sources are incorporated into the model by specifying flow data and concentrations of 23 water quality variables, with flux calculated as the product of flow and concentration. This approach integrates nonpoint source loading into the broader modeling framework.

The second focus was on direct coupling of the SCHISM, ICM, WorldCore, and Marine results, which run hydrodynamics and water quality simultaneously. The Patapsco–Back River MTM uses fine resolution grids (50–100 m), capturing features such as ship channels and side channels that the MBM, with coarser 200–500 m grids, cannot resolve. Calibration efforts between 1991–1996 at station WT 5.1 demonstrated the strong influence of Susquehanna River inflows, which periodically reduce salinity to near zero and increase turbidity. Adjustments to bottom drag coefficients improved salinity representation and model performance.

Results from stations CB 2.1, 3.2, and 3.3 illustrate downstream gradients in water quality. At CB 2.1, salinity remains low, DO is above 5 mg/L, and nutrients are relatively high due to Susquehanna input. At CB 3.2, closer to Baltimore Harbor, salinity rises and hypoxia begins to appear, while chlorophyll levels fluctuate with bloom events. At CB 3.3 near the Bay Bridge, salinity can reach 20 ppt and anoxia is evident, with unusual chlorophyll distributions suggesting circulation and nutrient effects. Overall, comparisons show strong agreement between modeled and observed conditions.

Within Baltimore Harbor, Susquehanna-driven freshwater inputs and three-layer circulation contribute to strong stratification and internal drivers of low DO. Anoxia occurs earlier inside the harbor than outside, while chlorophyll concentrations are extremely high at the surface. In the Back River, nutrient loading from the sewage treatment plant drives highly eutrophic conditions, with chlorophyll often exceeding 100 mg/L, though the shallow depth prevents prolonged hypoxia. Additional results from Green Branch confirm intermittent hypoxia and high nutrient concentrations, which the model captured well due to fine-scale grid resolution.

The study represents the first long-term comparison of water quality modeling in the Patapsco and Back Rivers. Direct coupling with SCHISM and WorldCore was successful, producing reasonable results for dissolved oxygen, chlorophyll, TN, and TP across multiple monitoring stations. However, indirect coupling (offline setup) presented challenges. Questions remain about appropriate time steps for finer grids, the stability of models with reduced friction, and the handling of wave–current interactions. These issues must be resolved before 10-year offline simulations can be confidently implemented.

In conclusion, the modeling demonstrates strong capability in representing hydrodynamics and water quality through direct coupling. Key successes include accurate representation of salinity intrusions, nutrient gradients, and DO dynamics, with reasonable statistical agreement with monitoring data. The next steps will focus on refining indirect coupling approaches, addressing numerical and stability issues, and ensuring robust long-term simulations.

1:30 ADJOURN

Attendees:
Richard Tian (UMES), Lewis Linker (EPA), Mark Bennett (USGS), Gabriel Duran (CRC), Dave Montali (Tetra Tech), Jian Shen (VIMS), Petra Baldwin (CRC), Carl Cerco (Arlluk Technologies), Cathy Wazniak (MDDNR), Kachapond Chettanawanit (East Carolina University), Clifton Bell (Brown and Caldwell), Jiabi Du (Texas A&M), Earl Hayter (USACE-ERDC), Ray Fenstermacher, Gopal Bhatt (Penn State-CBPO), Harry Wang (VIMS), Ashley Hullinger (PADEP), Jodi Ryder (USACE-ERDC), Joseph Delesantro (ORISE-CBPO), KC Filippino (HRPDC), William Keeling (VADEQ), Len Schugam (MDE), Larry Sanford (UMCES), Kevin Mclean (VADEQ), Mukhtar Ibrahim (MWCOG), Normand Goulet (NVRC), George Onyullo (DOEE), Pierre St-Laurent (VIMS), Brenda Rashleigh (EPA-ORD), Rebecca Murphy (UMCES), Samuel Canfield (WVDEP), Sophia Grossweiler (MDE), Sam Merrill (NGEM), Steven Bieber (MWCOG), Wenfan Wu (VIMS), Joseph Zhang (VIMS), Marjorie Zeff, Zhengui Wang (VIMS), Hassan Mirsajadi (DNREC), Matthew Rowe (MDE), John Lancaster (PADEP).