Modeling Workgroup October Quarterly Review

Day 1 – October 7, 2025

Event webpage: Link

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

Table of Contents (Jump to bookmark):

Day 1 – October 7th, 2025:

Skip to Day 2.

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

10:05 Phase 7 Watershed Model Progress – Gopal Bhatt, Penn State-CBPO

Description: *Gopal will provide an overview of the Phase 7 Watershed Model progress over the last Quarter. Key advances in the Phase 7 Dynamic Watershed Model (DWSM) development will be presented including: (a) implementation of organic nitrogen and phosphorus scour in rivers to better represent high flow TN loads, (b) adding a trend component to the Generalized Stream Network (GSN) routing, and (c) incorporation of BMPs with CAST removal efficiencies aggregated up to key load sources.*

Gopal Bhatt: During the third quarter of 2025, the Dynamic Watershed Model (DWSM) team focused on achieving full integration of all key model components to ensure readiness for fine-tuning and calibration. The quarter's main objectives included finalizing the July 2025 beta release, incorporating organic scour processes for nitrogen and phosphorus, adding a trend component to the stream-network routing model, and introducing best management practices (BMPs) into model calibration. These updates aimed to enhance temporal and spatial performance ahead of final adjustments. The DWSM continues to provide high-resolution, daily watershed load estimates for the Main Bay (MBM) and tributary estuarine (MTM) models, while also serving as a platform for research on emergent watershed behavior and calibration optimization.

The DWSM operates in close conjunction with CalCAST, a statistical model that fuses monitoring data, catchment characteristics, and watershed drivers to estimate loads and delivery factors. CalCAST outputs feed into the DWSM, which disaggregates them temporally into hourly and daily scales. This hybrid approach allows the model to dynamically simulate hydrologic and biogeochemical processes, like sediment, nitrogen, and phosphorus transport, across varying land uses and physiographic regions. Early years focused on model architecture and CalCAST integration, while 2024 and 2025 emphasized calibration and linkage with estuarine models. Over this period, iterative beta releases progressively improved small-stream routing, parameter estimation, and calibration fidelity.

Among the quarter's technical achievements, the July 2025 beta marked a major milestone. The team completed generalized small-stream routing for sediment, complementing earlier flow and nutrient routing functions. Random-forest modeling helped estimate key parameters to represent flow and seasonal variability, substantially improving nitrogen predictions. Enhancements in riverine transport calibration also played a key role: by blending Hydrologic Simulation Program – Fortran (HSPF) native calibration with CalCAST-derived estimates, the team established a two-way feedback system where information from the DWSM improves both upstream and downstream components. Additional constraints from Weighted Regressions on Time, Discharge, and Season (WRTDS) further refined riverine transport parameters, resulting in better load estimates and particularly below the fall line, where direct monitoring data are scarce.

A critical advancement was the introduction of organic scour of nitrogen and phosphorus to improve calibration during high-flow events. Because HSPF lacks an organic-scour module, the team added parameters for organic N and P concentrations associated with scoured sediment, testing two formulations: independent calibration and a biologically constrained version based on the Redfield ratio. The latter provided better performance and biological realism. Results from the James River at Cartersville showed that including the scour process reduced underestimation of high-flow nitrogen peaks and improved the model's alignment with observed data. The addition of scour parameters reduced bias and improved temporal tracking of peak load events, demonstrating their importance in capturing episodic organic nutrient mobilization.

The trend component was added to the generalized stream-network routing to better represent long-term temporal changes in concentrations, especially in smaller NHD-scale streams. By incorporating a $\beta 1$ parameter derived from the trend in inputs to each stream the DWSM now carries forward upstream trend signals into downstream routing. Using a 7 year smoothing window akin to WRTDS, the approach successfully represented increasing concentration trends driven by both rising nutrient inputs and wetter climates, as seen in the Choptank Basin. However, early tests showed mixed outcomes: in some cases (e.g., Choptank), the trend overcorrected, leading to unrealistically sharp declines in recent loads; in others (e.g., Patuxent River), it improved alignment with observed and modeled trends. These variations highlight the need for further refinement of the trend implementation.

The integration of BMPs represented another major milestone. Working with Jess Rigelman, the team incorporated annual, spatially downscaled BMP data compatible with the P7 framework. This allowed simulation of time-varying BMP implementation and removal efficiencies for nitrogen, phosphorus, and sediment. When combined with the scour and trend modules, BMP inclusion improved temporal alignment between modeled and observed data, especially in later years where implementation rates were highest. For instance, in the Choptank, BMPs helped mitigate overcorrections from the trend component, while in the Patuxent, both BMPs and the trend improved calibration accuracy. These results underscore that BMP impacts vary by watershed depending on dominant processes but generally enhance model realism and fit.

Across all metrics, the quarter's findings show consistent patterns. For nitrogen and nitrate, the addition of scour had limited impact, while BMPs modestly improved Nash–Sutcliffe efficiency and bias metrics across multiple watersheds. The trend component offered targeted benefits but requires additional testing. For phosphorus, scour effects were minor, the trend component often degraded performance, and BMPs again provided steady improvement. Taken together, these results mark a point of completeness for the DWSM: all critical elements are now integrated, enabling comprehensive calibration and stronger coupling to estuarine models.

Looking ahead to the final quarter of 2025, the team plans to focus on diagnosing and improving the trend module, incorporating new P7 land-use data, and expanding the use of additional monitoring datasets. These efforts aim to enhance the model's representativeness and predictive capacity before final integration with the MBM and MTM.

10:35 Discussion of the Phase 7 Model Progress

Q: *Lewis Linker:* Back in P6 calibration, weren't our targets ±10% for nitrogen and ±15% for phosphorus? By describing this as "good," are we still in that range?

- ***A:*** *Gopal Bhatt:* During P6, the MWG agreed through expert judgment that a good calibration meant within ±10% for nitrogen and ±15% for phosphorus, given the inherent uncertainty in WRTDS estimates. I should note, however, that P7 differs somewhat. We still use CalCAST and have ported our calibration framework from P6, but now we also use WRTDS to nudge model parameters and adjust loads. This feedback loop didn't exist in P6.
- ***Q:*** *Dave Montali:* I noticed that in the Mattaponi, the flow seems to deviate the most. Does that mean our nitrogen and phosphorus calibration issues might be

tied to flow performance? Is there more work needed there, or do we know why flow is off by about 10%?

- ***A:*** *Gopal Bhatt:* Yes, flow calibration plays a role. We rely on average annual bias to guide calibration, but parameters are defined at the land-segment scale. This means a single parameter adjustment can affect multiple stations. When we started with NLDAS data, we were overestimating flow in the Mattaponi by about 12%. After switching to PRISM, the bias dropped below 10%, which was an improvement, but still a bit high. We haven't yet dedicated effort to further investigate this issue, though it's worth doing. Expect an update next quarter.

- ***Q:*** *Olivia Devereux:* I'm always mindful that when we calibrate the DWSM to match WRTDS, we're still comparing one model to another. How do we know which is more accurate?

- ***A:*** *Gopal Bhatt:* In P6, we always calibrated against direct observations – using cumulative frequency distributions (CFD) of measured concentrations as our gold standard. WRTDS is indeed a model, and while it has limitations, it provides a useful benchmark for estimating loads. We know where it performs well and where it struggles. For instance, WRTDS tends to underestimate during high-flow melt events in the Susquehanna because those samples weren't collected. So, while WRTDS is imperfect, it offers valuable guidance for bracketing load estimates. During calibration, we weigh how closely to match it and what level of deviation is acceptable. Our uncertainty thresholds remain ±10–15% for nitrogen and phosphorus, and possibly larger for sediment. That's the standard we use to evaluate calibration quality.

- ***Q:*** *Guido Yactayo:* How is WRTDS being used here? Is this just for comparing with your model and calculating the statistics?

- ***A:*** *Gopal Bhatt:* We are directly using the WRTDS to change the model parameters so that it is actually informing the calibration. Whenever it doesn't match, we force it. Its similar to SPARROW – we are making WRTDS estimate the loads in all of the monitoring stations and trickling it down. We do a sort of propagating of the parameters upstream and downstream because we have calibrated the monitoring station upstream and the loads are transferred downstream. We are using the WRTDS in what we generally refer to as true condition load over a long period of time, i.e. the average annual load.

10:45 **Progress on Phase 7 Nutrient Inputs and Sensitivities – Joseph Delesantro, ORISE-CBPO with Conor Keitzer and Rosh Nair-Gonzalez, UMCES**
Description: Updates to load sensitivity to input values and key inputs of agricultural fertilizer, manures, and atmospheric deposition (to forest) will be presented. In addition, finalizing the sanitary sewer exfiltration estimate method and overall results will be

discussed as well as phosphorus trend sensitivity to pH, ionic strength, and naturally weathered phosphate by physiographic region.

Joseph Delesantro: In 2024, a comprehensive assessment was conducted using literature reviews, staff reports, and evaluations of the P6 model to identify shortcomings in how phosphorus loading is represented. That work highlighted several key processes missing from P6 that are now being targeted for inclusion in P7 to improve phosphorus predictions across watersheds.

The first major group of missing processes involves hydrologic connectivity, which governs how water, and phosphorus bound to soils, moves across the landscape and reaches streams. These processes include runoff dynamics, topographic controls, hydrogeomorphic position, artificial drainage, and critical source areas. Because phosphorus is strongly sorbed to soils, its transport is highly sensitive to how quickly hydrologic connections form between landscape sources and waterways. Michelle Katoski and the GSAT team have developed new metrics to quantify these properties, which are now being prepared for testing within CalCAST, with initial results expected by December.

The second group of processes focuses on biogeochemical controls on phosphorus mobility, particularly alkaline desorption and salinity-driven desorption. Alkaline desorption occurs as rising pH causes phosphorus previously bound to sediments to be released into the water column. The pH levels in systems such as the Susquehanna River have increased since the 1950s, reaching the critical range where phosphorus desorption becomes significant. Salinity-driven desorption operates through a similar mechanism, where salt ions from salt water intrusion or road deicing, displace phosphate from sediment particles. Since April, efforts have been underway to represent both processes spatially and temporally within CalCAST to test whether they improve phosphorus transport predictions.

In addition to these mechanisms, the team is beginning to incorporate geogenic phosphorus sources derived from mineral weathering, drawing on approaches used in recent SPARROW models. These naturally occurring inputs are expected to further refine phosphorus source representation.

Conor Keitzer: Currently, we are analyzing observed data to understand how pH varies over time and space across the watershed and how those changes relate to phosphorus concentrations. Using data from the Water Quality Portal, they compared coalfield and non-coalfield regions by decade. The analysis showed that pH in coalfield areas increased substantially from below 6 in the 1980s to nearly neutral in the present, while pH in non-coalfield areas remained relatively stable. This pattern supports the idea that legacy mining activity and subsequent recovery are driving long-term pH increases in parts of the watershed.

In summary, sites with paired pH and phosphorus measurements showed higher phosphorus concentrations as pH increased into the 7–8 range, consistent with alkaline desorption of sediment-bound phosphorus. He also reviewed literature on salinity effects, noting strong relationships between specific conductivity, impervious cover, and snow depth. Even low levels of impervious surface were associated with increased conductivity, and higher snow accumulation further elevated salinity, likely due to road salt application. Together, these results reinforce the importance of urbanization and winter road management as drivers of salinity-induced phosphorus release.

11:15 Discussion of Phase 7 Nutrient Inputs and Sensitivity Progress

Comment: *Lewis Linker:* We've been wondering what's going on with phosphorus, and it's clear that in P6 we were missing something. The evidence has always pointed that way. What's interesting is that some areas, particularly those influenced by pH, will benefit from better information and adjustments, such as in the Susquehanna and parts of the Potomac. When you combine snowpack with impervious surfaces, particularly in metropolitan areas like Baltimore and Washington, you can really see the effects. It's fascinating how clues like snowpack, ionic strength, and pH interactions reveal how phosphorus behaves differently across regions. With improved calibration, this helps us understand from a first-principles perspective the chemistry behind phosphorus behavior and how it translates into a more accurate model.

- ***Response:*** *Conor Keitzer:* Rosemary and her team analyzed watershed-wide background connectivity based on geology and other factors. What we found is that this is a widespread issue across the watershed. While that might not surprise anyone here, it's an important factor to account for in the model moving forward.
- ***Comment:*** *Joseph Delesantro:* We're continuing to identify the best geogenic phosphorus dataset. These datasets are commonly used in SPARROW models, so we're reaching out to those modelers to understand their rationale for dataset selection. On the topic of pH, we've developed a spatial representation of where coal mine or acid drainage has occurred upslope of monitoring stations or stream reaches, along with a temporal representation of how pH has changed over time. So, we now have both spatial and temporal dimensions of pH, combined with literature-based information about how alkalinity affects phosphorus desorption. The idea is to integrate these three components. We have some approaches in mind, and you'll see testing and results from that effort in December.

Q: *Dave Montali:* When all this work is complete, how do we plan to portray phosphorus sensitivities? Will we look at management implications or develop ways to represent where actions could make a difference?

- ***A:*** *Joseph Delesantro:* We're considering that carefully. There are multiple ways to address it, and we plan to test some options before bringing recommendations back to this group. The current plan is to treat this as a stream-to-river factor. Essentially, where pH is higher, there will be less sorption and more desorption, leading to greater phosphorus transmission along that stream-to-river continuum. We plan to parameterize a coefficient for this relationship using CalCAST. That's how we envision implementation. Whether or not management actions will directly address this is outside my scope, but it's a potential next step. Even so, any improvement in model accuracy ultimately enhances our ability to manage phosphorus. Especially in sectors where we might currently have misallocations or misinterpretations.

- ***Response:*** *Dave Montali:* So, ultimately, this relates to the concentration of impervious area within a model unit rather than the specific load sources like roads or buildings, correct? Your coefficient would depend on how much impervious area exists in that unit and whether that unit has high de-icing activity.

11:25 **Future Environmental Conditions and CBP BMP Efficiencies – Maya Struzak, Sarah Fakhreddine, and David Rounce, Carnegie Mellon and Michelle Miro and Krista Grocholski, RAND**

Description: Progress will be presented on application of the APEX field scale watershed model under different future climate hydrologic conditions to determine potential change in relative pollutant removal efficiency of current CBP-approved NPS management BMPs. In future 2026 Quarterlies the SWMM stormwater model will be applied at field level scales to examine CBP stormwater BMPs efficiencies to future environmental conditions.

Maya Struzak: The project is focused on evaluating how future climate conditions may affect the efficiency of BMPs across the watershed. The primary objective is to quantify changes in BMP performance under both current and projected hydrologic scenarios. At present, the work concentrates on the agricultural sector using the APEX hydrologic model, with specific attention to BMP effectiveness in reducing nitrogen, phosphorus, and sediment loads.

To represent the diversity of conditions within the Chesapeake Bay watershed, the modeling framework includes four major physiographic regions: Ridge and Valley, Appalachia, Coastal Plain, and Piedmont. Within each region, four agricultural land uses are modeled, resulting in sixteen baseline scenarios. These baseline conditions form the foundation for testing BMP performance under varying weather and climate forcing, allowing for region- and land-use-specific assessments.

The modeling workflow begins with detailed site characterization using CAST data, literature, and census information to define slopes, soils, climate variables, and agricultural practices. Baseline scenarios are first simulated without BMPs to support calibration, followed by BMP scenarios for row crops. Validation is performed by comparing APEX water balance outputs, such as runoff and evapotranspiration, against CAST water balance components for representative hydrologic segments. In parallel, BMP removal efficiencies produced by APEX are compared against Chesapeake Bay Program reference values to ensure consistency.

Current efforts focus on row crop BMP scenarios, beginning with soybean systems. Based on recent discussions with the Bay Program team, the modeling approach may shift to emphasize additional row crops like corn and small grains, in place of forest and pasture land uses to better reflect dominant nutrient loading patterns. Validation results show that APEX runoff estimates align well with CAST outputs, and preliminary BMP simulations demonstrate expected reductions in sediment, nitrogen, and phosphorus across multiple management scenarios.

BMP performance results are evaluated both graphically and in tabular form, directly comparing APEX-derived removal efficiencies with Chesapeake Bay Program values. This process supports ongoing calibration and refinement of the model. Once validation and calibration are completed across all land uses and physiographic regions, the scenarios will be integrated and used to simulate future climate conditions, providing a comprehensive assessment of how BMP efficiency may evolve under changing hydrologic regimes.

11:50 Discussion of CBP BMP Efficiencies Under Changing Environmental Conditions

Comment: *Lewis Linker:* Row crops like corn, soy, and small grains is where the nutrient loads are concentrated. They represent the majority of agricultural loading, so it makes sense to focus there. We have to prioritize what's most important to capture, even if that means reducing other categories. That's the rationale behind this shift, and I welcome discussion on it.

- ***Comment:*** *Joseph Delesantro:* Maybe keep pasture but drop small grains. We arrived at the corn, soy, and small grain combination logically, based on coverage and rotation, but in practice, most nitrogen applications occur on corn and soy, especially considering nitrogen fixation. There are also significant BMP applications on pasture. So, while expanding beyond a single crop is important, we might not need to include small grains right now.
- ***Q:*** *Dave Montali:* Why is hay important in this context?
- ***A:*** *Joseph Delesantro:* I don't recall extensive discussion on hay previously. There's certainly a large land area dedicated to hay, but we can revisit our data on acreage, nutrient application, and BMP coverage to refine that assessment.
- ***Comment:*** *Maya Struzak:* Hay does occupy a large share of land, which might be why it was originally selected, even if its load contribution is lower.
- ***Comment:*** *Kim Van Meter:* I was thinking about hay from the nitrogen perspective. There's significant nitrogen fixation, particularly in alfalfa hay. Although we harvest the aboveground biomass, about half of the fixed nitrogen remains below ground in the roots, which can contribute substantially to nitrogen inputs.
- ***To-do:*** *Lewis Linker:* We'll set up a small group to decide whether to include hay moving forward. It seems small grains may be dropped, and pasture retained.
- ***Comment:*** *Bill Keeling:* The point about legumes is valid, but in Virginia, most hay is actually legume-poor. We had a problem in P6 where we overestimated

fixation – treating fescue as if it fixed 150 pounds of nitrogen per acre. So, if we discuss fixation, we need to ensure it's truly legume-based, not misapplied to grasses.

12:00 LUNCH

1:00 Watershed Modeling Using Machine Learning Techniques – Chaopeng Shen and Kim Van Meter, PSU

Description: Progress in a cooperative Agreement for Penn State PIs to support and collaborate with CBP scientists and modelers using machine learning and/or other appropriate techniques will be described.

Kim Van Meter: This presentation outlines a machine learning initiative designed to modernize the Chesapeake Bay modeling framework by integrating data-driven approaches. Presenters recognize that older Bay models were not optimized for fine spatial scales, despite the growing need to make predictions at those resolutions. Traditional process-based models are computationally demanding when scaled down, but machine learning offers a path toward efficient, high-resolution predictions. By combining extensive geospatial datasets with flexible algorithms, the team seeks to capture nonlinear relationships among water-quality drivers and stream nutrient concentrations and enhance both predictive accuracy and management applicability.

Initial work is centered on developing random-forest models to predict monthly nutrient concentrations and loads, with plans to expand to daily predictions through an LSTM (Long Short-Term Memory) neural-network framework. A key motivation is to test whether patterns learned from the random-forest models can inform the Bay Program's land-to-water factors and whether high-resolution land-use and geomorphic data improve prediction skill.

Random forests, which aggregate the results of many decision trees built on random subsets of data, have proven both robust and interpretable for this application. The models currently draw on data from roughly 80 non-tidal monitoring stations across the watershed, with plans for further expansion. They predict total nitrogen and total phosphorus, using a combination of static watershed descriptors (land use, soils, geology, geomorphology, watershed area, stream order) and dynamic forcings (streamflow, precipitation, temperature, and nutrient inputs). Nutrient forcing data come from the gTREND dataset, which integrates county-scale census data on manure, fertilizer, biological nitrogen fixation, and wastewater, downscaled to 250 m grids and re-aggregated to individual watersheds.

The models are trained and validated using WRTDS estimates of monthly nutrient loads. Results show very strong performance, with $R^2 \approx 0.96$ and Kling–Gupta Efficiency (KGE) ≈ 0.93 for total nitrogen – indicating that the random forest reproduces observed nutrient concentrations exceptionally well. When compared with the P6 Chesapeake Bay model results, the machine-learning approach performs particularly well at low concentrations and maintains consistent accuracy across years, even under variable hydrologic conditions. Typical of random-forest behavior, residuals are small at low concentrations and increase at higher values; the team plans to test boosted regression trees to further reduce this bias.

A major strength of the approach is its interpretability. Feature-importance analyses reveal that watershed nitrogen surplus is a dominant predictor of total nitrogen concentrations, providing valuable insight into key biophysical drivers. Building on this, the team is producing partial-dependence plots to visualize how predicted nutrient concentrations respond to specific predictors while other factors are held constant. These relationships can help quantify management-relevant sensitivities. For instance, estimating how a reduction in manure inputs from 70 to 50 kg N ha^{-1} might lower total nitrogen concentrations by about 0.2 mg L^{-1}. While such quantitative estimates require caution, even directional understanding is useful for guiding nutrient-reduction strategies.

Looking ahead, the project will broaden its data foundation by incorporating more monitoring stations, finer-scale land-use data, and explicit wastewater information. The team's next goal is to strengthen the link between model sensitivity analyses and practical management applications – translating predictive relationships into actionable insights for land-to-water factor development. In doing so, they aim to ensure that the models are not only powerful predictors but also effective diagnostic tools that clarify how specific management actions can influence nutrient outcomes across the Chesapeake Bay watershed.

1:20 Machine Learning Discussion

Q: *Lewis Linker:* Can the model be analyzed at different spatial scales? For instance, if it could compare basins such as the Susquehanna versus the Potomac, or even finer watershed scales? One of the greatest values of this modeling effort will be its ability to reveal patterns and relationships that are otherwise difficult to observe through traditional data analysis.

- ***A:*** *Kimberly Van Meter:* The goal is to understand what factors make a difference between basins. The current approach identifies which variables the model is most sensitive to overall, but it doesn't directly provide basin-by-basin comparisons. However, the model outputs can be examined at the individual watershed level to explore those finer details. In essence, this analysis sets the stage by highlighting the most important drivers systemwide, which can then guide deeper investigation within specific basins.

Comment: Lewis Linker: For the cause of the nonlinear response, the crop type, manure availability, or regional difference might influence the observed behavior, particularly at higher input levels above 60-70 kg per hectare.

- ***Response:*** *Kimberly Van Meter:* At lower input levels, there may be sufficient cropland available to properly distribute manure, minimizing runoff. In contrast, in areas of intense livestock production with limited cropland, excess manure may not be managed effectively, leading to higher runoff and nutrient concentrations.
- ***Comment:*** *Lewis Linker:* Understanding these mechanisms could help determine whether and how nonlinear relationships from the machine learning models should be integrated into other frameworks, such as the Phase 7 or future Phase 8 models.

Comment: *Joseph Delesantro:* One immediate opportunity is using machine learning as a feature selection. While calibration of coefficients for land-to-water and stream-to-river factors is performed within CAST and CalCAST, those frameworks lack tools for identifying which watershed features are most relevant. Machine learning could help determine which variables should be included as influential factors in these models, such as topography, hydrologic connectivity, or geomorphic characteristics. This could feasibly be developed within the timeframe of the ongoing Phase 7 calibration effort. Distinguishing between nutrient sources and transport/retention processes, since the latter is where feature selection could be most impactful.

- ***Response:*** *Kimberly Van Meter:* The feature importance offers a promising pathway. I suggest we ensure that all potentially relevant predictors are included in the models first so that their relative contributions can be properly evaluated and compared. This approach could identify the combinations of factors most critical for improving land-to-water and stream-to-river representations.

Q: *Samuel Canfield:* Does WRTDS incorporate observed data from the non-tidal network and whether that observed data might be reused during the bagging process in the random forest testing phase?

- ***A:*** *Kimberly Van Meter:* The WRTDS model uses observed data from 81 non-tidal network stations. WRTDS converts sparse observations into daily estimates, which are then aggregated to monthly values used as inputs to the random forest model. 80% of this dataset is used for training, while 20% is held out completely and never introduced into the decision trees. This held-out data serves as an independent test set for model validation.
- ***Response:*** *Samuel Canfield:* I observed that, in random forest and boosted regression tree models (BRT), collinearity is generally less problematic because of the tree-based structure, though it can still affect the magnitude of relationships if the model is applied to new spatial or temporal domains. The main challenge is biological interpretation, understanding which predictors matter most when the model includes dozens of potential features. When I ran BRT models, it was composed of multiple decision trees combined into an averaged "global" model (40–50 iterations) to improve stability and reduce variability.

1:30 Update on 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: The project focuses on modeling the Conowingo Reservoir using an integrated framework that couples hydrodynamics, sediment transport, nutrient dynamics, and water-quality processes. The primary objective is to support evaluations of dredging scenarios as well as extreme-event and climate scenarios. Current efforts are centered on developing a baseline model for the 1991–2000 period, which represents the first major milestone and will be fully documented before scenario analysis begins. Following completion of the baseline, the work will branch into dredging-related analyses and climate-driven simulations, with all scenarios assessed for impacts on sediment transport, sediment quality, and water quality.

A central technical requirement of the project is accurately capturing interactions between hydrodynamics and sediments, including how flow conditions influence water-column chemistry and how sediment changes feed back into water-quality responses. This includes representing physical sediment changes resulting from both natural extreme events and human-driven dredging activities. Over the past six months, significant effort has been devoted to contract development and establishing a computing environment capable of supporting this highly coupled modeling approach.

The model currently relies on P6 hydrologic data to define boundary conditions. Water-quality parameterization is underway and is supported by targeted sediment sampling to inform key sediment-related parameters. The Environmental Fluid Dynamics Code (EFDC) model was selected as the core modeling platform due to its flexibility, compatibility with both desktop and high-performance computing environments, and open-source structure, which ensures long-term transferability of the modeling framework.

Model development included constructing an initially fine-resolution grid, then selectively coarsening it to improve computational efficiency while retaining critical spatial detail. A two-month hydrodynamic and sediment transport simulation was completed on a PC using this grid and now serves as a validation benchmark. With installation on the EFDC high-performance computing system complete, the same simulation is being rerun on the HPC to evaluate differences arising from compilers, numerical libraries, and parallelization. These differences are expected to be minor and are currently being assessed to ensure consistency before advancing to fully coupled baseline simulations.

Sediment data collected in late June play a key role in ongoing parameterization efforts. Sampling locations were designed to characterize both physical transport properties and sediment chemical composition, including carbon fractionation. In parallel, long-term biochemical oxygen demand (BOD) tests are being conducted to estimate sediment oxygen demand. Results to date show highly biologically active sediments, requiring repeated re-oxygenation even in large-volume test bottles. These results will inform preliminary model inputs, with follow-up tests planned to verify findings.

Next steps include completing the linkage between watershed model outputs and EFDC water-quality boundary conditions, building on existing crosswalk approaches previously applied to sediment. With the model now operational on the HPC, dredging scenarios are expected to be set up in the next quarter, followed by extreme-event simulations after full water-quality calibration. Discussions are also underway regarding the timing and implications of transitioning to P7 inputs. Overall, the project's major components are in place, and the team is working toward initiating full calibration of hydrodynamics, sediment, and water quality by the end of the current quarter.

1:40 Discussion of the Conowingo Model Development and Application

Q: *Lewis Linker:* Could oxygen demand also be coming from the water column, not just the sediment, and removing and replacing water during re-oxygenation might remove oxidizable material?

- ***A:*** *Jodi Ryder:* That's a good point. In the earlier re-oxygenation cycles we did not retain the removed water, but in the later stages we could retain and test it separately. Because the water was originally prepared using ultra-filtered water with nutrient additions and inhibitors, it was expected to be bacterially inactive. However, testing the removed water could help confirm whether oxygen demand is dominated by sediment processes.

- ***Comment:*** *Lewis Linker:* performing this test would help address a question that will likely arise later and that any removed oxygen demand could be accounted for quantitatively.

- ***Response:*** *Jodi Ryder:* It would be informative to test that. While the water column was intended to be sterile, the combined sediment–water system remains biologically active even after 90 days, which is noteworthy.

Q: *Lewis Linker:* What does the schedule look like for transitioning to P7 inputs? P7 is performing well in some respects but is not yet fully embedded and reviewed.

- ***A:*** *Jodi Ryder:* We would not be ready to run Phase 7 scenarios until at least this coming summer, so that timing aligns reasonably well with what you're describing.

1:50 **Closing Out Efficient Multi-Objective Optimization Procedures – Kalyan Deb, Pouyan Nejadhashemi, Ritam Guha, and Auden Garrard, MSU**
Description: Last steps and closing out of the CBP optimization project.

Kalyan Deb: The project has been underway since 2020 and, although it will be halted after the first quarter, work has continued using remaining funds to complete closeout activities. While a few planned components could not be finalized, the resulting software is robust and functional. The development team believes it will provide significant value to EPA staff despite these limitations.

The demonstration introduced the Pigeon-C multi-objective optimization tool through a user-facing dashboard. After account creation, users access the "Optimize Scenarios" section, where they can view an overview of the system and initiate new optimization scenarios. The dashboard presents a clear interface for managing and tracking scenarios in various stages of completion.

Users can review existing or pending scenarios in a central table that displays status, selected counties, base scenarios, and scenario names. In the demonstrated example, a new optimization scenario was created for the state of Delaware with specific assumptions, including the exclusion of forest BMPs, the removal of manure transport from Chester County, Pennsylvania, and a budget cap of $100 million. The scenario used 2023 as the base year and relied on official BMPs, cost profiles, and water datasets.

Once created, the optimization process allows users to define pollutants of interest, select or exclude BMPs and counties, and launch the optimization run. Results are presented as multiple solutions that reflect trade-offs between cost and pollutant reduction. Users can compare constrained and unconstrained scenarios, explore individual solutions based on criteria such as nitrogen load targets, and download detailed output files for further analysis.

The outputs from the optimization tool were validated against the CAST system, showing close agreement and confirming the accuracy of the approach. Final project activities focus on completing documentation, including a technical installation guide and a detailed user manual. Planned technology transfer activities include installation support, live demonstrations, and a potential webinar. The team emphasized that the tool is designed to feel familiar to CAST users and enables informed decision-making by presenting multiple optimization solutions rather than a single prescribed outcome.

2:15 ADJOURN

Attendance:
Gabriel Duran (CRC), Joseph Delesantro (ORISE), Gopal Bhatt (PSU), Richard Tian (UMCES), Lewis Linker (EPA), Dave Montali (Tetra Tech), Samuel Canfield (WVDEP), Wenfan Wu (VIMS), Jian Shen (VIMS), Normand Goulet (NVRC), Clifton Bell (Brown and Caldwell), Kimberly Van Meter (PSU), William Keeling (VADEQ), Tyler Trostle (PADEP), Joseph Zhang (VIMS), Marjorie Zeff (AECOM), Robert Burgholzer (VADEQ), Mukhtar Ibrahim (MWCOG), Carlington Wallace (ICPRB), Larry Sanford (UMCES), Arianna Johns (VADEQ), Steven Bieber (MWCOG), Guido Yactayo (MDE), Marjy Friedrichs (VIMS), Deni Chambers (NGEM), Jodi Ryder (USACE), Cassandra Davis (NY DEC), Zhengui Wang (VIMS), Douglas Bell (EPA), KC Filippino (HRPDC), Jeremy Hanson (CRC), Olivia Devereux (Devereux Consulting), Scott Heidel (PADEP), Karl Blankenship (Bay Journal), Conor Kietzer (UMCES), George Onyullo (DOEE), Xueting Pu (PSU), Sam Merrill (NGEM), Sarah Fakhreddine (CMU), Maya Struzak (CMU), Ashley Hullinger (PADEP), Chaopeng Shen (PSU), Allison Welch (CRC), Christina Lyerly (MDE), Kalyanmoy Deb (MSU), Amirpouyan Nejadhashemi (MSU), Auden Garrard (MSU).

Modeling Workgroup July Quarterly Review

Day 2 – October 8, 2025

Event webpage: Link

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

==

Table of Contents (Jump to bookmark):

Day 2 – October 8th, 2025:

Go back to Day 1.

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

10:05 Assessment of Atlantic Croaker Present and Future Habitat in the Chesapeake– Colin Hawes and Marjy Friedrichs, VIMS

Description: An assessment of Atlantic croaker habitat under present and future conditions in the Chesapeake will be presented. Future directions for the assessment as well as assessment of other Key Chesapeake species such as Bay anchovy and striped bass will be discussed.

Colin Hawes: The work presented is on physiologically based habitat suitability modeling for juvenile Atlantic croaker in the Chesapeake Bay. This work is part of a larger Chesapeake Bay Program funded effort to project habitat suitability under historical, present, and future climate conditions. The project integrates species specific physiological response curves with a three-dimensional hydrodynamic and biogeochemical estuarine model, allowing habitat projections without reliance on historical catch data.

The habitat suitability model is process-based and driven by experimentally derived bioenergetics relationships linking growth to temperature, dissolved oxygen, and salinity. Growth is calculated as grams gained or lost per gram of body weight per day and is derived from bioenergetic components including consumption, respiration, and waste. Habitat is classified as suitable when growth is positive and unsuitable when growth is negative. Environmental inputs are provided by the ROMS-ECB Chesapeake Bay model (Chesapeake Regional Ocean Modeling System – Estuarine Carbon Biogeochemistry), which has high spatial resolution and has been evaluated against long-term water quality monitoring data.

Preliminary results from experiments including a 1990s control scenario, three future climate scenarios representing different combinations of Earth system models and emission pathways, and a management scenario representing achievement of TMDL nutrient reduction goals. Future climate simulations were applied to estuarine conditions only, while watershed climate impacts were examined separately. Analysis focused on bottom habitat across the entire Bay, reflecting the demersal nature of Atlantic croaker and their observed presence at depths up to 20 meters.

Results show that Atlantic croaker cannot occupy the Bay year-round, but future warming consistently increases the duration of suitable habitat, extending both the beginning and end of the seasonal window. Across all future scenarios, the number of habitable days increases, indicating that croaker is likely a climate-tolerant species. The magnitude of this increase varies substantially depending on both Earth system model choice and emission scenario, highlighting the importance of using multiple climate projections to capture uncertainty.

While climate warming increases habitat duration, summer habitat suitability is strongly influenced by DO dynamics. Nutrient reductions associated with TMDL implementation do not significantly alter the seasonal duration of habitat but substantially improve the amount of suitable bottom area during summer months. These improvements can reach up to 700 square kilometers of additional suitable habitat, a change that becomes more apparent when expressed in spatial rather than percentage terms. When nutrient reductions are combined with warming scenarios, habitat conditions improve further due to croaker's high heat tolerance.

In closing, climate change may extend Atlantic croaker occupancy in the Bay by several weeks by the 2070s, but management actions remain critical for mitigating summer oxygen stress and preserving suitable bottom habitat. Future work will include longer simulations to capture interannual variability, additional climate scenarios, and expanded modeling to other species such as striped bass and bay anchovy.

10:35 Discussion of Atlantic Croaker Present and Future Habitat Assessment

Q: *Lewis Linker:* Given the use of ROMS-ECB and the Atlantic croaker bioenergetics model, is it possible to estimate croaker biomass from this analysis?

- ***A:*** *Colin Hawes:* Potentially, yes. One important aspect of the current approach is that croaker body weight is prescribed and held constant. The model calculates potential growth rates for a croaker of a given size in each grid cell. It may be possible to allow body weight to change iteratively over time based on calculated growth rates, but that would require substantially more work. Even then, it would still represent a single theoretical individual, so scaling to biomass would likely require incorporating abundance or catch data.

Q: *Clifton Bell:* I was wondering about food web dynamics within the bioenergetics model. Specifically, how changes in primary productivity, such as increases or decreases in chlorophyll-a under different scenarios, might affect the food web. Are those dynamics explicitly represented in the model, or are they implicitly captured through variables like temperature?

- ***A:*** *Colin Hawes:* At present, they are held constant. The model uses a fixed feeding efficiency. However, there are parameterizations of similar croaker models where feeding efficiency can be calculated based on chlorophyll-a. An example of this approach can be found in Kenny Rose's 2018 work in the Gulf of Mexico using a similar croaker bioenergetics model.

- ***Comment:*** *Lewis Linker:* Following up on Clifton's question, Atlantic croaker are bottom feeders, so benthic infauna would be an important food source. The chlorophyll connection would therefore be indirect – through primary productivity influencing benthic production, which in turn affects croaker. From what I understand, there is currently no explicit feedback involving benthic biomass in the croaker bioenergetics model.

Q: *Joseph Delesantro:* When I see work like this, I sometimes struggle to connect it back to overall productivity or limiting factors. The increased duration of suitable habitat under future scenarios is very interesting, but croaker must be going somewhere during the winter. Is habitat availability within the Bay actually limiting population productivity? More broadly, how should we be thinking about these results within the Bay Program's living resources framework?

- ***A:*** *Colin Hawes:* Addressing population-level productivity would likely require expanding the model to include additional processes such as reproduction. That would help link habitat suitability to population outcomes. One key distinction with this process-based approach, compared to statistical habitat models based on catch data, is that we are modeling the fundamental niche rather than the realized niche. The model shows what croaker could do, not necessarily what they will do. We cannot say for certain whether croaker would enter the Bay earlier or leave later, as that depends on migration cues, but the model suggests it would be physiologically possible. At this point, I am not seeing substantial reductions in available habitat area, so it is difficult to identify a clear limiting factor for croaker based on these results alone.

10:45 Update on Main Bay Model (MBM) Progress – Zhengui Wang, Joseph Zhang, and Jian Shen, VIMS

Description: Progress in many arears of MBM development will be presented. An overall assessment to date of the improvement seen in the sequence of combined watershed and MBMs from Phase 6 to the Phase 7 January 2025, April, and July versions of the models will be reviewed.

Zhengui Wang: Here is an update on recent collaborative work supporting development of the P7 MBM. Over the past quarter, the team tested multiple watershed and airshed nutrient loading scenarios, updated the MBM workflow to incorporate new inputs, and completed initial simulations. Results from these updates are encouraging, and remaining model issues are actively being addressed, with a baseline P7 MBM expected by the next quarterly cycle.

A major workflow improvement involved converting watershed and airshed nutrient loading datasets into a consistent, unified format compatible with the MBM. This consolidation simplified file handling, improved coupling between watershed inputs and the estuarine model, and enabled more efficient processing. Newly delivered P7 atmospheric nutrient loadings were successfully integrated and compared against P6 results. Differences between phases were generally small, while evolving beta versions of P7 watershed loadings showed increasingly promising performance.

Comparisons of atmospheric nutrient deposition between P6 and P7 showed that P7 includes nitrate and other nutrient species over the coastal ocean, whereas P6 did not. Within the Chesapeake Bay, nitrate distributions were largely consistent across phases. Subsequent model comparisons indicated only modest changes in ammonia and nitrate within the Bay and slight reductions in phosphate under P7 conditions, with even smaller differences in the coastal ocean.

Shifting to submerged aquatic vegetation (SAV) modeling within the MBM framework, focusing on shallow-water processes. The SAV modeling approach couples hydrodynamic, wave, sediment, and biological components using a two-step strategy. Observed SAV coverage from 1990–2020 shows patchy distributions concentrated in shallow regions, often near model boundaries and only partially overlapping grid cells. To address this, the team developed a subgrid-based method to compute SAV coverage fractions for each MBM cell, preserving spatial patterns while enabling efficient integration. The physical SAV module dynamically simulates interactions between vegetation, currents, waves, and sediments, using fixed structural parameters and spatially varying SAV density derived from coverage fractions.

Results demonstrate that SAV significantly alters hydrodynamics, particularly in large beds such as the Susquehanna Flats. SAV reduces current speeds within beds, redirects flow into adjacent channels, lowers diffusivity, and attenuates wave energy. These physical changes lead to substantial reductions in total suspended solids (TSS), especially downstream of large SAV beds. Model–data comparisons at monitoring stations show that including SAV markedly improves agreement for both surface and bottom TSS, while effects in smaller SAV regions are present but weaker.

The biological SAV model simulates biomass dynamics for leaves, stems, roots, and tubers, producing realistic spatial and seasonal patterns. SAV biomass is highest in the upper Bay and exhibits a clear seasonal cycle, though modeled leaf biomass variability is smaller than observed due to limited calibration data. SAV also influences bottom dissolved oxygen, generally reducing DO through export of dissolved organic carbon (DOC), though results are highly sensitive to how biomass is routed among carbon pools.

During development, the team identified unrealistic bottom-water oxygen supersaturation and excessive DOC in some regions, particularly the upper Potomac. These issues were traced to SAV photosynthesis and organic matter routing. Model refinements allowing excess oxygen to vent to the atmosphere and redistributing organic matter between particulate and dissolved pools substantially improved agreement with observations. While further validation is needed, results indicate that SAV beds can export meaningful amounts of organic matter, especially in dissolved form.

In conclusion, the Phase 7 MBM now incorporates a sophisticated SAV component that integrates physical and biological processes, improving representation of shallow-water dynamics in the Chesapeake Bay. SAV reduces currents, waves, diffusivity, and sediment resuspension, exhibits strong seasonal biomass cycles, and plays a significant role in oxygen and carbon dynamics. Future work will focus on incorporating additional observational data, refining biological parameterizations, and expanding analyses of shallow-water processes and living resource interactions under future scenarios.

11:25 Discussion of the Main Bay Model (MBM) Progress

Q: *Lewis Linker:* SAV typically thrives in shallow areas and is often located outside of our primary model grid. As a result, we frequently observe SAV occurring beyond the defined subgrid. The approach described here involves collecting all SAV area, including portions outside the subgrid, assigning those areas to the appropriate subgrid, and then incorporating that subgrid into the MBM. This process ensures that the full SAV extent across the Bay is ultimately represented within the modeling framework. I want to confirm that this accurately reflects how SAV extent is being collected and assigned.

- ***A:*** *Zhengui Wang:* At present, most SAV located outside our model domain is likely being ignored. This is a limitation that we could address in the future. If we were to incorporate SAV outside the current domain, we would need to carefully consider how that integration affects the physical processes represented in the model.

- ***Comment:*** *Joseph Zhang:* While the physics would certainly be interesting to examine, it is not the primary determinant in this context. The most important factor is that SAV presence itself is a water quality standard. Every unit of SAV contributes directly to meeting water quality standards that support living resources and water clarity. This makes accurate representation of SAV critically important. In the coming quarters, we will need to examine SAV and water clarity in much greater depth, and this detailed presentation provides a strong foundation. Representing SAV everywhere, both in base conditions and future scenarios, is essential. I agree that extending the MBM grid to explicitly cover these areas is preferable to forcing them into the existing grid structure. This would allow us to maintain physical consistency while incorporating SAV more accurately. We believe this can be done efficiently without dramatically increasing mesh size or computational burden.

Comment: *Lewis Linker:* The seasonal variation in SAV presence and its wintertime absence is clearly visible, along with corresponding effects on TSS. An interesting extension of this work would be to introduce a tracer representing particulate organic carbon, phosphorus, or nitrogen. This raises the question of whether these particulates would be entrained and trapped within SAV beds due to reduced wave energy and increased settling, or whether they would bypass SAV beds through higher-velocity flow around them. This is an important and testable question.

- ***Response:*** *Zhengui Wang:* To some extent, the effects on settling are already represented. Flow fields change due to SAV blocking, and velocities are reduced within SAV beds, allowing organic matter to remain in place longer. However, in the current SAV model, some parameters in the water column remain constant. There are therefore two competing processes, and further refinement may be needed. In summary, profile changes capture some effects on settling, but because certain values in the sediment model are not dynamically varying, future improvements could enhance realism. So, because waters within SAV beds are more quiescent we would expect increased settling of TSS and particulate organic matter once material enters the bed.

Comment: *Larry Sanford:* A master's student working with Cindy Palinkas and me published an analysis very similar to this several years ago. That study focused on 2011, which included back-to-back events from Hurricane Irene and Tropical Storm Lee. One was dominated by wave energy and the other by runoff, yet the conclusions were very similar to what you have shown here, which is encouraging. One important distinction is that during the modeled period, there was no SAV present; recovery occurred only at the very end of that time period. In the 1990s, there was essentially no SAV in that area. Using an average SAV condition is acceptable for scenario analysis, but it introduces uncertainty for calibration. One particularly interesting point is that SAV may actually increase scour in adjacent channels. It is not only reducing deposition; it may also enhance scour at choke points where channels converge, especially during large storms. This process is not well captured in the model, but it could increase suspended sediment loads locally. Going forward, capturing feedbacks between physical processes, water clarity, and SAV recovery would be extremely valuable.

Q: *Jiabi Du:* you showed the difference in sediment concentration with and without SAV near the river inflow. I am wondering why the sediment concentration near the Susquehanna River inflow appears so low. We would expect a large sediment load from major rivers, yet the values near the inflow appear close to zero. My interpretation is that the high sediment concentrations observed in the upper Bay are primarily induced by sediment resuspension from the seabed rather than direct river input. This seems like something that should be addressed, although I am not certain what is driving this pattern. My concern is that something may not be represented correctly with respect to sediment loads from the major rivers.

- ***A:*** *Zhengui Wang:* This largely depends on timing. The Susquehanna River exhibits a wide range of observed TSS concentrations. During low-flow conditions, upstream reservoirs significantly reduce TSS, resulting in concentrations of only a few milligrams per liter. During high-flow events, concentrations can exceed one gram per liter. The snapshot shown here represents a quiescent, low-flow period.

Q: *Carl Cerco:* In previous modeling efforts, we represented three SAV communities: freshwater, mesohaline, and polyhaline. Does the current model represent three distinct communities, or are we using a single community across the entire system? And, do the current parameters correspond to one of the historical community types from the old Bay model, or is this a new formulation?

- ***A:*** *Zhengui Wang:* At present, we are using a single SAV community. This is an area where the model could be improved in the future. To your second question, not exactly. The SAV model design differs somewhat from the earlier 3D Bay model. However, the spatially varying parameter capability gives us flexibility to apply different species or community types in different regions. We would likely need your guidance to identify appropriate parameter values for each community.

11:35 **Assessment of Sea Level Rise Scenario Methods for the MBM – Richard Tian, UMCES-CBPO**
Description: An initial assessment of sea level rise (SLR) scenarios using the Main Bay Model (MBM) will be presented including an assessment of residual currents under SLR conditions and approaches to changing ocean boundary conditions for SLR scenarios.

Richard Tian: Here I provide a preliminary assessment of SLR impacts using the P7 watershed model. The analysis focused on comparing a baseline Watershed Implementation Plan (WIP) scenario with a SLR scenario, emphasizing that all results are temporary and subject to change. Initial findings show that the two scenarios are largely similar, which aligns with expectations that SLR would alter transport volumes by roughly 10%. Two primary hypotheses were identified to explain the limited differences observed: either gravitational circulation under SLR is not yet fully responsive in the model, or the open boundary conditions require refinement.

To investigate these hypotheses, the team analyzed multiple transects across Chesapeake Bay segments and an offshore transect, examining bottom residual currents, vertically integrated flow, and perpendicular currents. Across most transects, results showed coherent and physically consistent patterns, including downstream surface flow driven by freshwater inputs and enhanced upstream bottom flow associated with gravitational circulation. Difference fields between the baseline and SLR scenarios revealed small but detectable increases in upstream bottom flow, particularly within deeper channels. Although the magnitudes were modest, the spatial consistency across transects suggests a real signal related to SLR.

More complex circulation patterns were observed near tributary transitions and offshore regions, where eddy-like structures and bidirectional bottom flows emerged. These patterns highlight the sensitivity of circulation to channel geometry and offshore dynamics, reinforcing the importance of accurately specifying boundary conditions. Comparisons with historical observations from the late 1990s showed similar spatial flow distributions, though the model generally underestimates current magnitudes. Despite this, the modeled directionality of surface and bottom flows aligns well with observations.

The analysis also examined salinity responses, showing that the largest changes occur in frontal zones near freshwater–saltwater interfaces. Under SLR conditions, saltwater intrusion extended farther upstream in several tributaries, in some cases by up to 20 kilometers. When averaged across the system, the increase in intrusion distance was ~4 kms. These changes represent potentially meaningful ecological impacts.

A key finding emerged from sensitivity analyses of boundary condition nudging. Adjustments to offshore concentration nudging produced substantial changes in results, and removing nudging altogether caused roughly two-thirds of the differences between scenarios to disappear. This indicates that boundary condition treatment exerts a dominant influence on modeled responses and must be handled carefully. Although differences between baseline and SLR scenarios remain, their interpretation depends heavily on how nudging is implemented.

In summary, no formal conclusions should be drawn at this stage, given the preliminary nature of the work. The hydrodynamic response to SLR appears physically sound and well represented in the model, suggesting that remaining uncertainties are more likely tied to boundary conditions rather than core physics.

11:55 Discussion SLR Scenario Methods

Comment: *Lewis Linker:* I want to offer a preliminary conclusion framed around "big things and little things," meaning we should keep the end goal in mind as we draw conclusions. Under P6 climate change analyses, temperature had a dramatic and generally negative influence on water quality. Increased watershed loads driven by greater intensity and precipitation volume were also important. Higher temperatures influenced respiration, increased stratification, and contributed to lower dissolved oxygen saturation. SLR was also a major finding. You systematically demonstrated that the hydrodynamic representation is not only good, but excellent, and the SLR implementation appears very strong. That suggests the remaining issues may not be physical processes themselves, but something else. Returning to "big things and little things," boundary conditions are now a major focus: we need to ensure they are correct under key scenarios and across all scenario operations.

Q: *Harry Wang:* I have a question about boundary conditions. In SCHISM, the domain has been extended further out onto the continental shelf. With that extension, do you think the open boundary conditions need to be adjusted? Or do you expect the system to adjust on its own because the open boundary is now farther offshore and spans both east–west and north–south boundaries?

- ***A:*** *Richard Tian:* For Scenario 1, when nudging is applied, the open boundary conditions may matter less because conditions are largely overwritten or constrained by the nudging zone. However, if we remove nudging, then open boundary conditions become something we need to consider more carefully. Even if the boundary is farther offshore and potentially less important, it may still influence the physics and, importantly, water quality variables. Under some scenarios, the open boundary could change due to factors like atmospheric deposition, since offshore waters would reflect those changes. The magnitude may be large, or it may already be captured by the circulation patterns, but I do not have a definitive answer at this point. This could be particularly important near the northern end, where material may enter from adjacent systems such as Delaware Bay.
- ***Comment:*** *Joseph Zhang:* a quick clarification - the current nudging applies to water quality variables only. It does not apply to temperature, salinity, or hydrodynamics. There are always uncertainties associated with nudging, and Jian has run many simulations exploring this. The coastal ocean includes processes we do not fully understand, so we need to test nudging carefully. Conceptually, nudging is a practical method to constrain model water quality using available observations, including the water quality observations compiled and used in prior work. It is particularly important for DO, and we have published on this approach. For other variables, the effects may be smaller. For example, the apparent reduction in hypoxic area may reflect changes around the 2 mg/L threshold, where small absolute shifts can have large effects on mapped hypoxic extent.
- ***Comment:*** *Jian Shen:* This is a very good test case. If you look at the nudging zone relative to the Bay plume, much of the plume extends beyond the nudging region. The return flow from the Bay can be strong and pushes material outward. When you remove nudging and rely only on the original open boundary condition, the incoming water tends to be too "clean," and the model will not reproduce dissolved oxygen patterns near the mouth very well. In our testing, we adjusted the open boundary conditions substantially so that calibration could still match observations near the Bay mouth. We do not fully know the true source contributions, especially from the northern portion, or contributions from the Eastern Shore or coastal processes. However, when we adjust the open boundary

while maintaining calibration under current conditions, the net changes under scenarios may not be large. The justification for those coastal boundary adjustments is still difficult, because coastal sources and processes remain uncertain. If we eliminate nudging, we likely need to implement a different approach near the true open boundary and adjust boundary conditions accordingly. We should treat this as a recalibration problem. We have long suspected that Eastern Shore loadings could be high, but it may not be a simple loading issue. There are extensive wetlands, potentially low DO waters, groundwater interactions, and other processes that we do not fully understand. We also lack strong measurements for DOC and particulate organic carbon in these coastal areas. Our working assumption is that something is contributing, but we can only treat it as sensitivity testing at this stage and adjust the boundary so the model remains calibrated.

12:05 LUNCH

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

Description: Progress on the Patapsco/Back MTM and linkage with Phase 6 will be reviewed. Plans to move forward with water quality calibration using the latest version the Phase 7 Watershed Model will be discussed.

Harry Wang: This is an update on the Patapsco–Back River MTM program. We have three goals: describe the nonpoint source loading being used, summarize 10-year (1991–2000) water quality simulations conducted with an offline approach, and highlight model issues that are specific to the Patapsco–Back River system, considering specific locations like Baltimore Harbor, which has unique characteristics that require special consideration.

For inputs, the team is using a hybrid of P6 and P7 nonpoint source loading provided by Zhengui. The Susquehanna River contributes a dominant flow signal, and inputs include key freshwater sources to the Baltimore Inner Harbor such as Jones Falls, Gwynns Falls, and the Patapsco freshwater branch. One notable gap is the lack of modeled loading from Herring Run into Back River, although Back River itself includes a major wastewater treatment plant that contributes substantial loading.

The water quality setup includes 23 modeled variables driven by the source input (21 variables when temperature is excluded). The offline workflow runs hydrodynamics first, then waves, then sediment transport within the hydrodynamic framework; outputs are saved every 30 minutes and used to drive the water quality model for the full period. This represents a shift from earlier modeling approaches where sediment was not explicitly modeled in the hydrodynamics; in SCHISM, sediment transport is integrated into the hydrodynamic component. This makes total suspended solids (TSS) a key variable to track, as it strongly influences downstream water quality dynamics.

Hydrodynamic performance was generally strong. Water surface elevations compared well with observations at Fort McHenry and Annapolis and temperature captured the expected seasonal cycle with good agreement across multiple Bay and harbor stations. Salinity required more calibration, but results were ultimately considered acceptable. Surface salinity was generally modeled better than bottom salinity, particularly at WT 5.1 where surface R^2 values were higher than bottom R^2 values, consistent with the challenges of representing salt intrusion dynamics near the bed.

Water quality results were broadly within observed ranges. Dissolved oxygen simulations captured the emergence of hypoxia and anoxia moving down-bay from CB 2.2 through CB 3 stations. In Baltimore Harbor, bottom anoxia timing was captured well though surface dissolved oxygen appeared biased low. Chlorophyll-a was generally reproduced within expected ranges (typically ~10–20 µg/L with occasional bloom spikes), with the harbor showing a pattern of reasonable surface performance but overprediction at depth, likely tied to sediment representation.

Nutrient results were also mostly reasonable: TN and TP means were captured across Upper Bay stations, though WT 5.1 exhibited notable TN spikes around 1994–1995 and late 1996 that need investigation. TP showed some underprediction at CB 3.3, and it should be noted that TP behavior is closely linked to TSS, reinforcing the central importance of sediment dynamics. Overall, the model shows no major red flags for hydrodynamics and core water quality variables, but Baltimore Harbor stands out as an area needing targeted refinement.

The dominant remaining challenge is TSS, which is systematically underpredicted across stations, especially in middle and bottom layers at WT 5.1. The current formulation computes TSS as fixed inorganic solids plus a biologically derived component (including algal biomass terms and particulate organic carbon), following MBM-recommended guidance. The sediment transport model uses four non-cohesive sediment classes (two clays, one silt, one sand) with grain-size-based settling velocities and critical shear stresses applied uniformly across the domain. This framework likely breaks down in Baltimore Harbor, where evidence suggests very low critical shear stress and cohesive behavior driven by high clay loads, ship-induced resuspension, biofilm-enhanced flocculation, and low mid-water turbulence due to stratification. Flocculation can increase effective settling velocities substantially yet these cohesive processes are not represented, contributing to TSS underprediction and related biases such as depth chlorophyll-a overprediction. Cohesive sediment processes must be incorporated or parameterized more realistically for Baltimore Harbor, and that future calibration should be spatially explicit given variability in seabed properties and local dynamics.

1:15 Discussion of Patapsco/Back MTM Progress & P7 linkage

Comment: *Jeremy Testa:* I appreciate the focus on improving sediment and resuspension processes in the models, but my broader takeaway is that stratification at station WT 5.1 appears to be substantially underestimated. When thinking about primary drivers of the model behavior you are seeing, this is the first issue I would prioritize because stratification strongly controls vertical structure and water quality outcomes. For example, is chlorophyll elevated because it is being mixed throughout the water column instead of remaining concentrated near the surface? Is surface-water oxygen unrealistically low because of excessive vertical mixing? The model results showing near-anoxic conditions in surface waters at that station appear highly divergent from reality, which suggests a larger underlying issue.

- ***Response:*** *Harry Wang:* For WT 5.1 stratification, surface salinity is captured reasonably well, but bottom salinity can still be improved. In the comparisons I have for 1991–1995 and 1991–2000, we have tried to reproduce salt intrusion as best we can. You are right that bottom salinity mismatches remain in some periods. However, stratification is essentially the difference between bottom and surface salinity, and overall that structure is present, though it may still be weaker than observed. I agree that some chlorophyll should remain more concentrated near the surface, but in the model it appears to behave differently. My view is that settling velocity likely plays a major role in this outcome.
- ***Response:*** *Jeremy Testa:* What caught my attention is that while the system is stratified and surface salinity appears reasonable, there are extended periods where bottom salinity differs from observations by roughly 5–6 units. I am trying to understand how much that discrepancy may allow unrealistic vertical behavior that drives some of what we are seeing. I also do not have much experience with the flocculation parameterization used here, so it is possible there are aspects I am missing due to limited familiarity.
- ***Comment:*** *Lewis Linker:* Across the MTM, TSS remains difficult to simulate, including in the Back River, Patapsco, and other tributaries. Many stations shown here are also represented in the main Bay model, where TSS calibration appears substantially better. It may be useful to coordinate with Zhengui and Joseph to understand what differs between the two implementations. This could be something as simple as a unit conversion or output-processing issue, but operationally we need to determine what is driving the discrepancy. Since both models use the same code base and ICM, it is notable that one achieves satisfactory calibration while the other does not. In terms of sequencing, I agree with Jeremy that we should resolve the broader TSS behavior first before making targeted changes for Baltimore Harbor. The "fluffy layer" and flocculent material described in the literature seems to be a common phenomenon and not unique to the Patapsco. Therefore, we should first improve TSS across the MTM domain, then examine whether additional refinements is needed for processes such as flocculation.

1:25 Progress on the Rappahannock MTM & P7 Linkage – Qubin Qin, East Carolina University and Jian Shen, Zhengui Wang, and Pierre St-Laurent, VIMS
Description: Progress on the Rappahannock MTM will be reviewed by the Rappahannock MTM Team. The application of using merged high-resolution

Rappahannock River and the main Bay models will be presented. Preliminary water quality model calibration using Phase 6 loading will be discussed.

Jian Shen: This is a summary of recent progress and presenting hydrodynamic model results for the Rappahannock River tributary system. The team implemented a revised modeling workflow developed by Zhengui that simplifies the tributary approach while leveraging the latest hydrodynamic framework. This updated setup significantly improves Bay-wide salinity and temperature representation and embeds the refined tributary grid directly within the full Chesapeake Bay model.

At this stage, the focus has been on hydrodynamics rather than water quality, as the team is waiting for the next water quality model version. Previous results using a refined tributary grid were already close to the MBM, with only modest improvements. Remaining discrepancies are believed to stem from missing lateral inputs from marshes and SAV. Rather than recalibrating boundary conditions, which is required in traditional nested approaches, the team opted for a fully embedded grid to avoid this complexity.

Embedding the fine-resolution tributary grid directly into the Bay-wide model offers several advantages. It eliminates the need to repeatedly pass boundary conditions and enables bidirectional coupling, allowing feedbacks between the tributary and the Bay. This is especially important for the Rappahannock River, where Bay–tributary interactions are strong. The approach also ensures consistency, as the MTM inherits the same parameterization as the Bay model from Zhengui's centralized database.

The only required adjustment under this framework is remapping SAV coverage to the refined grid. Otherwise, hydrodynamic, water quality, and suspended sediment parameterizations remain unchanged. Localized tuning is avoided unless clearly justified, maintaining consistency with the Bay model. Although running the entire Bay model is computationally slower than running an isolated tributary, this tradeoff is considered worthwhile because scenarios can be run seamlessly once the Bay simulation is complete, and computational resources are no longer a major constraint.

Hydrodynamic results from the new merged grid are very strong. Salinity aligns well with observations across stations near the river mouth, downstream, and farther outside the tributary, with weak stratification correctly reproduced. Additional 10-year simulations (1996–2000) confirm consistent performance, even where observational data are limited. Comparisons between the old and new merged grids show only minor differences, mostly in deeper waters related to mixing, with no impact on salinity intrusion into the Rappahannock River.

Temperature performance is also excellent, with modeled values matching observations closely at all stations. These results provide high confidence in the new configuration. The next step is to integrate the latest water quality model, which includes updated representations of SAV, marshes, and oysters. After careful remapping to the new grid, the team will rerun water quality simulations and expects further improvements, particularly at the few stations that previously showed discrepancies.

1:40 Discussion of Rappahannock MTM Progress

Comment: *Lewis Linker:* We need the MBM to run at a pace that can keep up with decision-making for the 2035 future-condition assessments and other analyses that Bay decision-makers will require. The MTM are not under the same runtime constraints. Your approach with using boundary conditions that allow two-way exchange and effectively "carry" the MBM along to provide what the detailed Rappahannock model needs is promising. I wonder whether this could also be applied to the Choptank, the Patapsco/Back River, and other tributary model configurations. It is worth considering whether this approach should be adopted more broadly, and I would welcome thoughts from the group.

- **Response:** Jian Shen: When we started with the tributary discussions, the open boundary on the left-hand side and the river boundary conditions were initially set up in the simplest way. However, we still saw differences when we ran the tributary model separately. Velocity, momentum, and bidirectional exchange matter, and whenever you define new boundary conditions, some approximation is unavoidable, either through nudging or imposed boundary constraints. The SCHISM model provides fine spatial resolution, but this is not a universal solution; it applies to a single tributary and does not solve the broader problem everywhere. We adopted this approach because it resolves multiple issues at once like the TSS issues Harry raised, wave-related issues, and loading-related issues. With this boundary approach, those problems are reduced or avoided.

1:50 Progress on the Choptank MTM – Jian Zhao, William Nardin, Elizabeth North, Larry Sanford, Jeremy Testa, UMCES and Jiabi Du, Texas A&M

Description: Progress on the Choptank MTM and P7 linkage will be described by the Choptank MTM Team and plans to move toward water quality calibration with the latest version the Phase 7 Watershed Model will be discussed.

Jian Zhao: Jian Zhao presented an update on Choptank River modeling on behalf of a large, collaborative team, outlining progress in both hydrodynamic and biogeochemical (ICM) simulations. He framed the work around two main components: improving hydrodynamic performance through grid refinement and diagnosing remaining challenges in biogeochemical behavior, particularly those linked to salinity intrusion, nutrients, light, and sediments.

For hydrodynamics, the team tested multiple grid configurations to better represent the Choptank River system. Early grids captured the overall river but poorly resolved the central channel, prompting gradual increases in spatial resolution and experiments with extending the open boundary farther into the Bay. While overall results across grids were similar, persistent underestimation of salt intrusion in the upper estuary remained a key issue. Efforts focused on balancing computational efficiency with improved physical realism, and version 1.6 of the model is currently used for both hydrodynamic and biogeochemical simulations.

Temperature performance across the estuary is strong, with the model successfully reproducing seasonal cycles and interannual variability at CBP monitoring stations in the upper, middle, and lower estuary. Salinity is also reasonably simulated near the estuary entrance and in the middle reaches, capturing expected spring–fall variability driven by river discharge. However, the model significantly underestimates salt intrusion into the upper estuary, where deeper saline water from the lower Bay should penetrate farther upstream – an issue with important implications for biogeochemical processes.

Through sensitivity testing, the team identified freshwater inflow as a major control on salt intrusion. Replacing watershed-model flows with USGS-based discharge substantially increased simulated intrusion though extreme events were still not fully captured. Additional experiments modifying local bathymetry by artificially deepening the upper estuary channel produced only minor improvements, leading to the conclusion that river inflow, rather than bathymetry, is the dominant limiting factor for upper-estuary salt intrusion.

On the biogeochemical side, the team compared fully coupled hydrodynamic–ICM simulations with an offline standalone ICM configuration. Because both approaches produced very similar results, the standalone ICM is now favored due to its much larger allowable time steps and significantly improved computational efficiency. Results were evaluated at stations near the entrance, in the middle estuary, and in the upper estuary. Near the entrance, temperature and salinity match observations very closely, while chlorophyll-a shows strong spring blooms linked to nutrient-rich, low-salinity conditions but is persistently overestimated in fall and winter, indicating calibration needs.

In the middle estuary, nutrients are generally underestimated, likely due to strong upstream consumption, which may explain underpredicted summer chlorophyll-a. At the same time, some chlorophyll-a overestimation suggests possible advection from downstream rather than local production. In the upper estuary, nutrient concentrations are much higher and drive strong spring chlorophyll-a responses, but organic carbon is underestimated and sediment processes are not yet activated. This led the team to investigate light limitation by introducing spatially varying light attenuation based on salinity as a proxy for turbidity.

Applying the light attenuation scheme substantially reduced chlorophyll-a across the system, including during peak bloom periods, and slightly cooled modeled temperatures. Effects extended beyond the upper estuary into the middle and lower estuary, altering salinity and nutrient dynamics in seasonally variable ways. Overall, the results indicate that light attenuation and sediment-related processes strongly influence biogeochemical performance throughout the Choptank system. Going forward, the team plans to further refine riverine nutrient inputs and activate sediment dynamics to better capture coupled physical–biogeochemical behavior.

2:05 Discussion of Choptank MTM Progress

Comment: *Lewis Linker:* If I understand correctly, the strategy is that Gopal and Zhengui work together each quarter to update inputs for both MBM and the MTM. Once the workflow is established, the updates should be relatively straightforward. There have been improvements each quarter, and I believe this is the quarter when we will move fully into P7. Water quality representation at tidal stations has improved significantly, as demonstrated by P7 results. These inputs should be ready and available, and Zhengui would be your primary contact.

2:15 **Criteria Assessment of modeled scenarios based on the 3-D Interpolator and Potential Phase 7 Application of 4-D Interpolator – Richard Tian and Rebecca Murphy, UMCES-CBPO**

Description: A presentation on a) how the current 3D Interpolator is used to assess CBP WQ standards; b) what the 4D Interpolator is; c) how the 4D Interpolator could be compared to the current 3D Interpolator in the assessment of Chesapeake Bay water quality standards; and d) thoughts on how to integrate the 3D and 4D Interpolators.

Richard Tian: DO assessments are organized by designated uses (deep channel, deep-water, and open-water) each tied to species protection with distinct criteria by life stage and assessment frequency. For spawning/nursery habitats, four assessment types are specified (7-day mean, 30-day mean, instantaneous, continuous). To date, only some of these can be evaluated; 7-day mean and instantaneous assessments remain infeasible because of data gaps.

Both assessment and modeling rely on data quality. The team uses only Tier 3 data (highest quality) for regulatory assessments and model applications; Tier 2 supports planning and Tier 1 supports education. A major constraint is the TMDL hydrologic reference period (1991–2000): it lacks high-frequency sensor data.

The assessment workflow proceeds as follows: assemble data; interpolate across space and time; compare interpolated values to DO criteria cell-by-cell; aggregate to three-year periods with four seasonal segments per year; and compute attainment versus non-attainment. For modeling, the emphasis shifts to modifying datasets to reflect scenario-specific sensitivities (e.g., to nutrient/sediment loads).

Four methods are used to create scenario-adjusted datasets: (1) Regression: map calibration to scenario and apply the function to observations (good for baseline/scenario comparisons but less robust under climate change); (2) Percentile change: apply percentile-specific deltas to each observation (handles non-linear shifts better under changing conditions); (3) Direct extraction: pull model values at the observation’s location/time; and (4) Direct model use: for volume-based analyses. Integrating observations, interpolations, and model outputs (P6/7) helps address CESR-noted mismatches, with interpolated datasets often aligning better than raw model fields. The goal is full integration with the 4D interpolator to deliver accurate, continuous DO assessments under baseline and scenario conditions.

Rebecca Murphy: This work builds on what Richard presented. We’re collaborating closely with Richard because interpolation is a key step in both the scenario assessments on your side and other related applications. The goal is to enable a more complete criteria assessment for dissolved oxygen in tidal waters. As Richard mentioned, several high-frequency criteria haven't been fully assessed yet, mainly because in the past there wasn't enough high-frequency data and because the existing assessment methodology wasn't designed to handle it and especially the interpolation component.

The current interpolation tool was developed more than twenty years ago. It's been useful but was never designed to interpolate through time or to accommodate short-term water-quality criteria. Now that we have much denser and higher-frequency data, we can fill temporal gaps more accurately. Still, even with improved data, interpolation remains necessary to avoid spatial or temporal bias. Our new method aims to fill those gaps in both space and time for more accurate criteria assessments.

We've been working on this for several years and expect to have a working tool by the end of 2025, with examples applied to multiple segments. We're collaborating closely with the Criteria Assessment Protocol (CAP) workgroup to refine how to approach high-frequency criteria. As we move into 2026–27, we'll be documenting, reviewing, and adding post-processing functions that link the tool to criteria assessment, while coordinating with the modeling teams.

The process for determining whether a segment meets a specific criterion begins with monitoring data, which are interpolated to fill spatial and temporal gaps. The results are compared to the criteria, and cumulative frequency distributions (CFDs) are built to determine attainment. For modeling scenarios, regression-based model products are substituted for observed data and run through the same steps to evaluate whether criteria would be met under future or alternative conditions.

Our 4D team's role is to replace the interpolation component of that workflow. The other steps will largely remain, though we'll refine the assessment of high-frequency criteria. The current interpolator has several limitations: it doesn't use high-frequency data effectively, treats each vertical layer independently, and performs only two-dimensional interpolation. Our new approach interpolates in full 3D space plus time (4D). Currently, data from a two-week "cruise" are assumed static, but the 4D interpolator will model continuous changes over time. It also incorporates statistical methods, unlike the older purely geometric approach.

The 4D interpolator will capture:

1. Mid-day space-and-time interpolation, producing daily estimates for every grid cell and depth.
2. Large-scale correlation, capturing day-to-day and spatial variability based on data-flow surveys.
3. Within-day cyclic interpolation, accounting for diel oxygen cycles driven by light and tides.
4. Small-scale correlation, representing fine-scale residual variability.

In a test case at station CB4, located in the mainstem Bay, we used a 1-km horizontal × 1-m vertical grid. The model uses generalized additive models (GAMs) to fit smooth relationships among oxygen concentration, location, and time. The mid-day interpolation produces realistic vertical and seasonal gradients. Adding the large-scale correlation component captures day-to-day variability, producing a realistic range of simulated oxygen levels. The within-day cyclic interpolation adds predictable daily oscillations, derived from 25 years of high-frequency data. Finally, random residual variability is added to reflect natural short-term fluctuations.

The combined output produces an hourly, depth-resolved interpolation that matches observed data well. Ten simulation realizations reproduce the observed range of hourly oxygen measurements with acceptable error. Zooming into just 10-days, we can confirm that the simulated profiles track the observed diel and depth variability accurately.

Scaling up, the 4D interpolator is being applied Bay-wide, generating hourly dissolved-oxygen fields across all grid cells. To assess criteria by designated use, we must also compute the pycnocline depth daily. We're updating that interpolation method to ensure consistency with the 4D approach. Once pycnoclines are defined, oxygen results are partitioned into the designated-use layers and evaluated using the existing CFD framework, now adapted for high-frequency data.

Our next steps include finalizing parameterization, comparing 4D and 3D interpolator outputs in case-study regions, refining criteria-assessment procedures, and continuing collaboration with the modeling team. We've had preliminary discussions about how the 4D interpolator could process estuary-model outputs in future scenario assessments. Conceptually, it could follow the same workflow as the current 3D method but incorporate the added temporal dynamics derived from high-frequency observations. These conversations are ongoing, and we'll revisit them in January as we refine both the technical design and integration plans.

2:45 Discussion of Potential Phase 7 Application of 4-D Interpolator

Q: *Lewis Linker:* for practitioners working with the MBM and the MTM, and thinking beyond the calibration stations where there may be other areas worth examining during calibration to support practitioners, how is simulated water quality performing there? The interpolator can provide estimates, and shallow-water continuous monitoring could help quantify confidence. There are applications beyond scenario analysis that we should consider as we complete the Main Bay and Multiple Tributary Model development this year. When do you anticipate scaling to every hour of a full year, and which years will the interpolator cover?

- ***A:*** *Rebecca Murphy:* We've tested across multiple years, with a focus on 2022 because NOAA's vertical array provided a key new high-frequency dataset, and they've continued sampling in subsequent years. For criteria assessment, that work is handled by the states rather than the modeling team. For our Bay Program indicator, we update annually and interpolate once new data are available, such as for 2025. We've also discussed the need to cover the 1990s period, and I've tested portions of that timeframe. The monitoring network was fairly consistent, but we lacked the high-frequency data we now use in the 4D interpolator. Even so, that variability is informed by all historical data, so while sampling locations rotate each year, we're drawing from the entire record. When we run the 1990s period, we use information on daily cycles and correlations derived from high-frequency data collected in the 2000s. There's no strict restriction on which years can be interpolated.

Q: *Richard:* Once this tool is fully developed, we might use it even for modern calibration where data are limited. I'm curious about the time realizations in the model. Each realization seems random – what exactly do those represent?

- ***A:*** *Rebecca Murphy:* We run multiple simulations because there are many variable components throughout the tool. We fit a statistical model where each parameter has some variability. If you fit that regression across many days of data, the slopes would vary slightly because of data uncertainty and natural variability. We sample from that distribution of fitted parameters derived from multiple days and sites. We also build correlation structures across space and time, since we don't have perfect data everywhere. The goal is to estimate a reasonable range for how often DO drops below thresholds like 1 mg/L. Using multiple realizations helps us capture that uncertainty, drawing on all available data sources. So, essentially variability in parameter values.
- ***Response:*** Elgin Perry: We introduce variability at several stages. One is at the daily prediction step, where the GAM model estimates a parameter vector with an associated variance–covariance matrix. We treat that vector as the mean of a multivariate normal distribution and make random draws from it for each simulation's daily mean prediction. Beyond that, we account for temporal autocorrelation, vertical (depth-based) autocorrelation, and spatial correlation across latitude and longitude. We also include deterministic cycles, like tidal and diel patterns identified through wavelet analysis. Altogether, these elements introduce realistic variability into the simulations. Unlike past interpolation methods that produced a single smooth value, we treat results as random variables, better reflecting uncertainty in the data.
    - ***Q:*** *Lewis Linker:* But interpolation results are usually unique once data are fixed. The range among realizations here seems large. How do we interpret that spread and avoid bias when combining results?
    - ***A:*** *Elgin Perry:* We don't intend to collapse all realizations into one "answer." Each simulation provides an estimate of the proportion of Bay waters not meeting the criterion. The spread across simulations shows uncertainty, helping determine whether results are confidently compliant or non-compliant.

Comment: *Jian Shen:* My concern involves lateral transport. In shallow water, flow moves southward, while deeper water moves northward, so there's limited exchange. Interpolating laterally can be tricky. The 4D interpolator's incorporation of tidal and daily cycles is promising, though. As Richard mentioned earlier, sea-level rise intensifies bottom currents and residual circulation, which presents challenges for the 4D interpolator.

- ***Response:*** *Rebecca Murphy:* One major improvement in the 4D interpolator is its ability to capture these east–west differences when supported by data. The GAM component models smooth functions laterally, longitudinally, and with depth, allowing interactions among these dimensions. In our tests, interpolation results differ between the east and west sides of the Bay, matching observed data. It's not perfect, but it's a big step beyond the older 3D approach, which relied purely on inverse-distance weighting and treated all directions equally.

Q: *Jian Shen:* For chlorophyll assessments, we currently use regression methods. Could this 4D approach eventually apply to chlorophyll or turbidity?

- ***A:*** *Rebecca Murphy:* At present, we're only applying it to DO. Extending to chlorophyll or turbidity is possible conceptually, but it would require additional planning with USGS partners. Once DO methods are finalized and reviewed, we can explore other parameters.

2:55 ADJOURN

Attendance:
Richard Tian (UMCES), Rebecca Murphy (UMCES), Dave Montali (Tetra Tech), Lewis Linker (EPA), Gopal Bhatt (PSU), Joseph Delesantro (ORISE), Robert Burgholzer (VADEQ), Colin Hawes (VIMS), Wenfan Wu (VIMS), Samuel Canfield (WVDEP), Marjy Friedrichs (VIMS), Brend Rashleigh (EPA), Clifton Bell (Brown and Caldwell), Kachapond Chettanawanit (ECU), Normand Goulet (NVRC), Carl Cerco (USACE), Jian Shen (VIMS), Larry Sanford (UMCES), George Onyullo (DOEE), William Keeling (VADEQ), Pierre St-Laurent (VIMS), Joseph Zhang (VIMS), Rebecca Murphy (UMCES), Harry Wang (VIMS), Breanna Maldonado (VIMS), Jeremy Testa (UMCES), Zhengui Wang (VIMS), Mukhtar Ibrahim (MWCOG), KC Filippino (HRPDC), Jiabi Du (TAMUG), Tyler Trostle (PADEP), Zhaoying Wei (UMCES), Dante Horemans (VIMS), Carlington Wallace (ICPRB), Alisha Mulkey (MDDOA), Cassandra Davis (DEC), John Lancaster (PADEP), Ashley Hullinger (PADEP), Scott Heidel (PADEP), Jian Zhao (UMCES), Deni Chambers (NGEM), Jon Harcum (Tetra Tech), Elgin Perry (Consultant), Allison Welch (CRC).