CBP Modeling Workgroup January Quarterly Review

Day 1 – January 6, 2026

Event webpage: Link

Table of Contents (Jump to bookmark):

Day 1 – January 6, 2026:

Skip to Day 2.

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

10:00 Phase 7 Watershed Model Progress – Gopal Bhatt, Penn State-CBPO

Gopal Bhatt: Here is an update on progress toward completing the Phase 7 (P7) Dynamic Watershed Model (DWSM), emphasizing that all major elements planned for the development phase are now in place. While additional inputs and refinements will still be incorporated as they become available, the model's core structure is now established. There are four main areas: integrating P7 land use (LU), extending the simulation period to cover 40 years, incorporating monitoring data through weighted regressions on time, discharge, and season (WRTDS) into calibration and verification, and revising the Generalized Stream Network (GSN) routing component.

DWSM plays a central role in the broader P7 watershed modeling framework. It receives information from the Chesapeake Assessment Scenario Tool (CAST) and CalCAST, with CAST serving as the management model and CalCAST (calibration) providing estimated parameters such as sensitivities, land-to-water factors, and stream delivery factors. The dynamic model uses these inputs, along with its own hourly hydrologic and water quality simulations, to generate daily loads delivered to the Bay. It is also used to calibrate against monitoring data, support scenario analyses, and provide a foundation for future research and collaboration.

Several important developments from the previous quarter include the addition of organic scour processes, an initial trend component in stream routing, and downscaled P6 Best Management Practice (BMP) information at the NHD catchment scale (National Hydrography Dataset). Performance results from the July 2025 P7 beta version showed that the model was already performing as well as or better than P6 in many major rivers and smaller embayments, based on root mean square deviation (RMSD) comparisons at downstream tidal stations. This strong performance suggested that the model framework was already sound, even as additional components continued to be added.

One of the most significant advances this quarter was the integration of P7 CAST LU data into the DWSM. Although this currently reflects pre-BMP land use, it represents a major improvement over the static 2013 P6 land cover previously used throughout development. A new workflow was developed to transfer these land use data into the model, and modifications were made to the model code and configuration files to accommodate the new inputs. To maintain computational efficiency, the 51 P7 CAST land use classes were aggregated into 17 core categories, a change that preserves catchment-scale loading accuracy while keeping simulation times manageable.

The simulation period has also been extended from 1985–2020 to 1985–2024, creating a full 40-year development-period simulation. This was accomplished using PRISM (Parameter-elevation Regressions on Independent Slopes Model) daily precipitation data disaggregated to hourly values with help from NLDAS-2 (North American Land Data Assimilation System) and DEQ-developed methods (Virginia Department of Environmental Quality). NLDAS-2 continues to provide other meteorological variables such as temperature, evapotranspiration, and wind speed. Workflows were updated to accommodate NLDAS's transition from GRIB to NetCDF, and QA/QC (Quality Assurance/Quality Control) confirmed that the system is functioning properly. Atmospheric nitrogen deposition inputs were also extended through 2024, though the estuarine domain extension is still underway.

Monitoring data continue to play an important role in calibration and verification. Earlier work used Water Quality Data Portal observations to generate Fluxmaster and WRTDS load estimates for stations meeting quality criteria, which were then used to estimate beta parameters and extrapolate them across the stream network using random forest (RF) models. The model is calibrated against both concentration data and WRTDS load estimates where available. More recently, the team began incorporating WRTDS-K, a Kalman filter-based approach, to improve load estimation and support evaluation of routing performance, especially in smaller streams. Calibration is guided primarily by how well simulated and observed concentration frequency distributions match across the full range of values.

The purpose of the GSN routing trend component work is to better align dynamic model load estimates with observed and expected trends and to expand the number of stations where meaningful comparisons can be made. The original conceptual trend approach borrowed from Fluxmaster and WRTDS and attempted to represent long-term changes in concentration as a function of trends in stream inputs. However, the earlier implementation inadvertently double counted hydrologic effects because the trend term still retained a climate-related flow signal that was already represented elsewhere in the routing equation.

To fix this, the team created a repeated-hydrology experiment using the same precipitation year across all years in the simulation, allowing us to isolate concentration trends without confounding them with changing hydrology. This refined trend term improved performance substantially compared with the October beta version and brought results back closer to the strong performance seen in July. Although some issues remain in the latest runs that include P7 pre-BMP LU, the overall conclusion was that the model is performing well and that the major components are now in place. Over the coming quarters, the focus will be on diagnosing remaining issues, incorporating new data as they become available, and continuing calibration and verification so the model can be fully completed by the end of the development period.

10:30 Discussion of the Phase 7 Model Progress

Comment from chat: *Normand Goulet:* UNEC Model: How are estimates of lag-time developed, and did you look at whether that had an impact on model performance?

- ***A:*** *Gopal Bhatt:* estimates of lag time are a fundamental input that informs the simulation. During development, we worked with USGS to quantify groundwater lag time estimates across the watershed. In addition, we empirically estimated lag times for surface runoff and sediment transport based on hydrology and sediment loading characteristics. In terms of evaluating model performance, we examined correlation plots comparing outputs. While we have not yet completed that analysis for P7, we previously compared P6 and P7 model performance. We calibrated the correlation between WRTDS monthly loads and the P6 DWSM monthly loads, and the results showed a much stronger correlation than we had in earlier versions. That improvement gave us confidence that the approach was working well. Moving forward, we plan to conduct a similar correlation analysis using WRTDS-K loads and a longer-term calibrated DWSM for P7.

10:40 Progress on Phase 7 Loading Sensitivities and Transport and Attenuation Factors – Joseph Delesantro, ORISE-CBPO and Conor Keitzer, UMCES

Joseph Delesantro: Here is an update on progress in developing P7 of the Chesapeake Bay watershed model, focusing on loading sensitivities, transport and attenuation factors, and ongoing integration of these components through CalCAST, the statistical calibration framework used alongside CAST. Two main topics are discussed: sensitivities of nutrient loading to input parameters and improvements to transport and attenuation factors, including work related to phosphorus. Many of the results are preliminary and that the current modeling framework is still a hybrid between P6 and P7 components, with additional P7 features planned for incorporation in the coming quarter.

The P7 modeling effort emphasizes incorporating more observational data, scientific literature, and expert knowledge to improve parameter estimation. Prior information is derived from published studies, expert panels from Chesapeake Bay Program (CBP) workgroups (WG), and results from P6 modeling. These priors are then used to inform calibration within CalCAST. The parameters currently being calibrated include loading rates, sensitivities, and coefficients associated with delivery factors. In previous meetings the focus was on loading rate priors, whereas the current discussion focuses primarily on sensitivities.

Sensitivity within the model represents the change in exported nutrient load relative to a change in input load, measured as deviations from the average input for a given source. These sensitivities determine how changes in inputs affect nutrient export from land to water. Over the past year, the modeling team has conducted literature reviews examining sensitivities for cropland fertilizer and manure inputs as well as forest responses to atmospheric nitrogen deposition. Additional work is ongoing to review sensitivities associated with crop nutrient uptake.

One example highlighted is the sensitivity of forested landscapes to atmospheric nitrogen deposition. In P6, this sensitivity was set at 0.02, which is relatively low compared to other watershed sources. However, atmospheric nitrogen deposition has declined substantially over time, and a low sensitivity value may limit the model's ability to capture these trends. Literature suggests higher sensitivity values, and the Forestry WG recommended a baseline estimate of 0.089. When this value was used as a prior in CalCAST, the calibrated result was slightly lower at 0.077, which aligns well with expectations given the CBP's more precise land-use classifications compared to many studies.

Although updating the forest sensitivity value has little effect on overall watershed-scale model performance, it significantly improves the model's ability to capture temporal trends observed at monitoring stations. Using the updated parameter reduces error in trend predictions and improves skill scores by roughly 28%. This demonstrates that while large-scale performance metrics may remain similar, parameter adjustments can meaningfully improve representation of long-term environmental changes. Its recommended that we proceed with CalCAST calibration using the literature-supported upper bound and implementing conditional relationships among sensitivities across other land-use types.

Looking at cropland sensitivities to fertilizer and manure inputs, particularly for nitrogen. A literature review of ~50 studies revealed substantial variability in reported sensitivities. When these values were incorporated as priors in CalCAST, the calibrated parameters tended toward higher values, especially for manure sensitivity. However, model testing showed that several parameter choices produced very similar watershed-scale performance. Because of this limited performance difference, using literature-derived values may be just as defensible as relying solely on calibrated parameters.

A similar analysis was conducted for phosphorus sensitivities, although the structure of the CAST model requires translating literature values into parameters compatible with the APLE (Annual P Loss Estimator) modeling framework, which represents phosphorus transport through several pathways. After combining fertilizer and manure literature estimates and calibrating a scaling factor in CalCAST, the resulting aggregate phosphorus sensitivity decreased slightly from about 0.045 in P6 to ~0.037. Again, improvements in model performance were modest. Since millions of individual fields across the watershed respond differently to nutrient inputs, a limited set of model parameters inevitably represents a simplification of real-world variability.

Next, work is underway to incorporate additional processes affecting phosphorus transport, such as hydrologic connectivity and geogenic phosphorus. Preliminary testing of a landscape connectivity metric, median watershed distance to streams, has already improved phosphorus yield predictions at monitoring stations, increasing model R^2 from about 0.66 to 0.70. The team also evaluated a continuous predictor for geogenic phosphorus derived from sediment geochemistry data, which performs similarly to existing hydrogeomorphic predictors but may reduce the number of variables needed in the model.

Looking ahead, the team plans to conduct more robust feature selection to determine which landscape and stream variables best explain nutrient transport and attenuation. This work is being done in collaboration with researchers at Penn State using RF models combined with partial dependency and accumulated local effects analyses to manage correlated watershed variables. Two modeling strategies will be used: one that predicts downstream nutrient loads directly and another that predicts residual concentrations after accounting for upstream inputs. Together, these approaches aim to identify the most influential predictors and refine the P7 watershed model.

11:20 Future Environmental Conditions and CBP BMP Efficiencies – Maya Struzak, Sarah Fakhreddine, and David Rounce, Carnegie Mellon and Michelle Miro and Krista Grocholski, RAND

Maya Struzak: Here we will present on our progress on how future climate conditions may affect the efficiency of agricultural BMPs in the Chesapeake Bay watershed. The primary goal is to determine how BMP performance and primarily removal efficiencies for nitrogen, phosphorus, and total suspended solids, may change under future environmental conditions. The analysis uses the APEX agricultural hydrologic model (Agricultural Policy/Environmental eXtender), focusing on four physiographic regions (Ridge and Valley, Appalachia, Coastal Plain, and Piedmont) and four representative crop systems (soybeans, corn, wheat, and alfalfa).

The modeling effort currently uses baseline climate conditions from 1990–2000, along with a 10-year spin-up period to stabilize the simulations. Several BMPs are being evaluated across all regions and land uses, including cover crops, no-till practices, nutrient management, manure incorporation, and grass buffers. Most of these practices have been successfully implemented in APEX; however, grass buffer simulations are temporarily delayed due to a bug in the APEX source code that prevents proper simulation when hourly precipitation data are used. The team is working with collaborators at the Water Institute to resolve this issue.

The modeling workflow begins with site characterization, using CBP data, literature sources, and census information to define representative agricultural operations in each physiographic region. Baseline scenarios without BMPs were first developed and calibrated to match CAST P6 water balance components, particularly evapotranspiration and runoff. BMP scenarios were then calibrated to reproduce P6 removal efficiencies across regions and land uses. This calibration process involves sensitivity analysis and grid-search optimization of parameters related to fertilizer application, crop growth characteristics, erosion control, manure inputs, and crop residue. Additional checks ensure that simulated crop yields are realistic and that water balance residuals remain small.

Once the baseline and BMP scenarios are calibrated, the next phase will simulate future climate conditions using CMIP5 climate projections (Coupled Model Intercomparison Project Phase 5) provided by the CBP. These include temperature and precipitation changes under RCP 4.5 and RCP 8.5 scenarios (Representative Concentration Pathway), derived from 30-year centered climate deltas. These changes will be applied to the reference climate for each region, and the models will be rerun to evaluate how altered climate conditions influence BMP performance.

Preliminary results indicate that the calibrated baseline simulations reproduce CAST removal efficiencies for sediment, nitrogen, and phosphorus reasonably well. Early testing also suggests that climate perturbations do influence BMP removal efficiencies, such as increased precipitation and warming. However, the modeling framework is still being finalized. The remaining steps include resolving the grass buffer simulation issue, completing all region and land-use combinations, and running the full set of future climate scenarios to assess how BMP performance may change under projected climate conditions.

11:40 Discussion of CBP BMP Efficiencies Under Changing Environmental Conditions

Comment: *Lewis Linker:* For now, our purpose is to understand whether there is a signal here, i.e., whether BMP responses are likely to change under future temperature and precipitation conditions. We are not at the stage of changing BMP efficiencies in the model. We are trying to understand whether there is a meaningful effect, and if so, how large it is.

Comment: *James Martin:* Are these values actually indicating negative removal efficiency? In other words, would implementation of the BMP be increasing sediment loads under climate change? If so, it is not just a percent change in a small positive removal efficiency. It may actually indicate a negative efficiency, a worst management practice rather than a best management practice for sediment. One thing that might help would be to include a row showing the no-BMP change, so we can understand how much of the apparent shift in removal efficiency is due to an overall increase in loading, as opposed to a change in practice effectiveness itself.

- ***Response:*** *Lewis Linker:* It is important to understand the mechanism behind a result where application of a BMP appears to produce a negative outcome. This is a very useful starting point, and it is also important to recognize that as this work continues through the year across agricultural BMPs and eventually stormwater BMPs, CBP will end up with region- and crop-specific APEX models that can support future sensitivity testing and expert evaluation.

Comment: *Olivia Martin:* If we are following the PSC's direction, then the current question is simply whether BMP efficiencies would change, not necessarily what action would be taken if they do. Whether they change a little or a lot is what we are trying to determine. The Phase 5.0 approach introduced hydrologic variability into the dynamic watershed model, but that was quickly found to be not especially useful operationally. This is a different approach because it focuses on BMPs rather than hydrology alone, so it will be interesting to see how it might eventually be incorporated.

11:50 LUNCH

12:50 Progress in CBP Artificial Intelligence and Machine Learning – Joseph Delesantro – EPA ORISE-CBPO

Joseph Delesantro: I will provide a brief overview of a Scientific and Technical Advisory Committee (STAC) workshop held in February 2025 focused on how artificial intelligence (AI) and machine learning (ML) can advance Chesapeake Bay research and management. The workshop brought together over 50 participants from federal, state, and academic institutions, representing a wide range of expertise across the watershed. Its primary goal was to assess the current state of AI/ML applications, identify key challenges and gaps, and explore opportunities for integrating these approaches into Chesapeake Bay science and decision-making.

The workshop produced a published report with four major recommendations. The first emphasizes strengthening data infrastructure and integration, particularly by harmonizing spatial and temporal datasets so they can be easily used across AI/ML workflows. This includes leveraging diverse data sources such as remote sensing and high-frequency monitoring, and designing workflows that better integrate observational data, model outputs, and supporting datasets. A key takeaway is the need to develop standardized, accessible datasets that can serve as inputs for a wide range of AI/ML applications.

The second recommendation focuses on leveraging AI/ML for restoration and decision support. This includes applying these tools to evaluate restoration practices, assess progress, and better understand drivers of change in both tidal and nontidal systems. AI/ML are already being used in some areas, such as feature selection and parameter estimation, but there is significant potential to expand their role. In the future, these methods could also support scenario planning and provide more accessible tools for managers and stakeholders.

The third recommendation highlights the importance of transparency, interpretability, and stakeholder engagement. As AI models can sometimes function as "black boxes," the workshop emphasized the need for explainable AI methods and clear uncertainty frameworks to ensure results are trustworthy and actionable. Engaging managers and decision-makers throughout the development process is critical, as is improving communication tools and visualizations to translate complex AI outputs into practical guidance for restoration and policy decisions.

The fourth recommendation centers on building collaboration and capacity across the Chesapeake Bay community. While there is already a growing network of researchers working in hydro-ML and related fields, there is a clear need to expand collaboration, training, and shared resources to fully integrate these methods into Bay Program workflows. This includes fostering communities of practice and strengthening connections between scientists, modelers, and practitioners.

The presentation also highlights an ongoing project involving the CBP Office and collaborators at Penn State and UMBC, which aims to use machine learning to incorporate high-resolution spatial data, like lidar and remote sensing, into water quality modeling. This effort is closely aligned with the workshop recommendations, particularly around data integration and collaboration. It also supports upcoming P7 modeling work, where machine learning will be used for feature selection and improved representation of transport and attenuation processes. Ultimately, these efforts aim to better identify the most important drivers of nutrient loading, especially for phosphorus, and to reduce uncertainty in model predictions as the program continues to evolve.

1:10 Proposal for a New Hydrologic Long Term & Critical Period – Lew Linker, EPA-CBPO

Lewis Linker: There is growing discussion at CBPO to evaluate and potentially update its hydrologic reference periods, including both the long-term hydrology period and the 3-year critical period used in water quality assessments. While no formal decision has been made, there is increasing interest from the Water Quality Goal Implementation Team (GIT) in pursuing this work. Technically, updating these periods is feasible and well-supported by existing methods, but the policy implications are more complex, as changes in hydrology could redistribute nutrient and sediment loads across jurisdictions, affecting state-level responsibilities and management decisions.

The current framework is grounded in the 2010 Total Maximum Daily Load (TMDL) documentation, which provides strong technical justification for how the original hydrologic periods were selected. The 3-year critical period (1993–1995) was chosen to represent sustained high-flow conditions without extreme events, and it plays a central role in defining the Bay's carrying capacity, particularly for dissolved oxygen. Even if a new critical period is adopted, the 1993–1995 period will likely remain important for continuity and historical comparison, given its long-standing role in CBP decision-making.

Looking ahead, there is both scientific and management interest in updating these hydrologic periods, potentially on an accelerated timeline to align with P7 modeling decisions by the end of 2026. With a much longer hydrologic record now available, there are more options for defining representative long-term and critical conditions. However, selecting appropriate periods requires careful analysis of seasonal flow dynamics, such as the influence of spring freshet through summer, as well as differences among tributaries. Tributary-specific impacts are particularly important because rivers vary in how their nutrient loads affect Bay processes depending on location and transport dynamics.

Previous analyses highlighted in Appendix G of the TMDL documentation, provide a strong foundation for this work. Those analyses used statistical approaches to identify hydrologic conditions that balance realism with avoidance of extreme events. Similar methods will be applied again, potentially supplemented with newer tools such as machine learning, given the expanded data now available. At the same time, there is recognition that different water quality indicators (e.g., dissolved oxygen, chlorophyll, water clarity, submerged aquatic vegetation) may respond to different hydrologic conditions, raising the possibility of exploring multiple critical periods for different ecological responses in the future.

In addition to the critical period, long-term hydrology remains essential for determining controllable nutrient loads and supporting key modeling scenarios, such as geographic influence analyses. Past work showed that long-term hydrology is relatively stable across different candidate periods, particularly when using contiguous 10-year windows that capture antecedent conditions. Moving forward, the effort will involve close coordination among the MWG, the Water Quality GIT, and external experts such as the USGS, with the goal of producing a robust technical assessment that decision-makers can choose to apply or not.

Overall, while updating hydrologic periods is technically achievable and scientifically valuable, it is ultimately a policy-sensitive decision. The analysis itself is expected to be worthwhile regardless of the outcome, providing updated insight into hydrologic influences on water quality. The overall impact on final assessments may be modest, given that existing modeling frameworks already incorporate environmental variability through correction factors. The immediate goal is therefore to complete the analysis and provide clear options to CBP leadership for consideration.

1:30 Discussion of a Proposal for a New Hydrologic Long Term & Critical Period

Q: *Scott Heidel:* I am curious about how you plan to account for the Conowingo Reservoir, especially under extreme flow conditions when legacy sediment stored in the reservoir can become resuspended and transported downstream. Conditions in the upper watershed respond very differently to those loads and flows, so I wanted to get some clarification on how that will be handled.

- ***A:*** *Lewis Linker:* We always consider what is happening at Marietta as distinct from what is happening at Conowingo, with the understanding that Conowingo is really its own system because of the accumulated sediment in the three lower Susquehanna reservoirs. We expect to use an approach similar to what we used in P6. That means relying on what we know from Hydrologic Engineering Center's River Analysis System (HEC-RAS) and other monitoring work that UMCES has conducted in Conowingo to assess how scour and deposition respond under different flow conditions. It is always a dynamic equilibrium, but as you correctly pointed out, if precipitation increases under future hydrologic conditions, then scour in Conowingo would also be expected to increase. That would be tracked separately through a dedicated Conowingo assessment, much as we did in P6, but now with added information from the Army Corps of Engineers working on the current assessment. I think we will be in better shape in terms of information than we were in P6. Extending the hydrologic record out to 2024 also helps, and we will see how these various products develop as they are brought forward.

Q: *Lawrence Sanford:* If there is a new hydrologic period, would that also become the calibration period, or is the plan to retain the old hydrologic period for calibration of the Phase 7 model?

- ***A:*** *Lewis Linker:* We have found that if we calibrate 10 years of the model, we tend to get a solid calibration across essentially the full temporal domain of the model. So I am fairly confident that we could shift the long-term average hydrology, because once you have 10 years of simulation or even ten years of observations, you have a strong representation of long-term hydrology. There are certainly nuances and slight differences, but I think in general we would be fine. Of course, we would still test and verify everything to make sure it is all correct and robust with any new hydrologic period used in calibration. But my expectation is that we would see similar goodness-of-fit statistics.

- ***Response:*** *Lawrence Sanford:* I have heard bits and pieces of discussion about possibly running a very long model simulation spanning many years. That would be another way to determine whether the predictions based on the old hydrologic period remain valid when you extend the model into a more current period. For example, if a new hydrologic period were something like 2015 to 2025, it would be very convincing to show whether the model calibrated under the older hydrologic period still performs well into the more recent period.

1:40 **Advances in Nutrient Mass Balance Analyses – Robert Sabo, EPA-OW**
Two recent papers advancing a national and a Chesapeake watershed assessment of nutrient mass balances will be presented:
The US EPA's National Nutrient Inventory: Critical Shifts in US Nutrient Pollution Sources from 1987 to 2017. https://pubs.acs.org/doi/10.1021/acs.est.5c08196
And *Net declines in nonpoint source pollution into one of the world's largest estuaries. In preparation.*

Robert Sabo: In this presentation we will discuss our recently published paper on the U.S. National Nutrient Inventory (NNI) and highlight a broader transition toward big data, AI, and ML approaches in water quality research. The NNI provides a comprehensive, mass-balance dataset designed to support stakeholders with clear, interpretable metrics of nutrient inputs, outputs, and trends over time. These datasets are paired with predictive models and decision-support tools to help inform nutrient management strategies, restoration planning, and policy decisions across the United States.

The inventory tracks nutrient dynamics across four major sectors: natural, agricultural, urban, and atmospheric. This is done at the county and subbasin (HUC-12) scales from 1987 to 2017, with some agricultural data extending back to 1950. By integrating these sectors, the NNI enables analysis of how nitrogen and phosphorus sources have changed over time and how they relate to water quality outcomes. Key national findings suggest that despite substantial increases in agricultural production, nutrient surpluses have remained stable or declined, reflecting improvements in nutrient use efficiency. Additionally, atmospheric nitrogen deposition has declined significantly, particularly in the eastern U.S., and wastewater phosphorus loads have decreased even as population has grown, indicating progress in pollution control.

The data also reveal important spatial patterns and sector-specific dynamics. Agricultural inputs have increased substantially to support crop and livestock production, but improvements in crop uptake have helped stabilize nutrient surpluses. For phosphorus, there is evidence of declining surplus and even negative balances in some regions, meaning that legacy phosphorus stored in soils is being drawn down. However, large accumulations of legacy phosphorus remain in certain areas, such as the Delmarva Peninsula, representing both a challenge and an opportunity for improving nutrient management and water quality outcomes.

One way we apply the NNI in tandem with AI/ML was in combination with a neural network model to explain long-term trends in nitrogen export across the watershed. By integrating time series of agricultural surplus, urban inputs, atmospheric deposition, and point-source loads with observed nitrogen export data, the model was able to accurately predict nonpoint-source nitrogen yields across 121 monitoring stations. Results show that total nutrient inputs have declined across major river basins, driven by reductions in atmospheric deposition, improvements in agricultural nutrient management, and decreases in point-source pollution.

The modeling results also provide insight into the relative importance of different drivers of water quality change. Counterfactual analyses suggest that improvements in agricultural nutrient management played a critical role in reducing nitrogen export in major systems like the Susquehanna River, while increases in urban inputs may have offset some gains in other areas. The neural network also revealed nonlinear relationships highlighting the complexity of nutrient transport processes and suggesting opportunities to improve watershed models through better representation of these dynamics, particularly for atmospheric deposition.

Overall, NNI and its associated modeling applications represent a powerful framework for tracking progress, identifying drivers of change, and supporting decision-making. By combining mass-balance datasets with AI-driven models, this work enhances the ability to evaluate nutrient reduction strategies, engage stakeholders, and develop targeted management actions. The integration of these tools into Chesapeake Bay applications demonstrates their potential to improve understanding of long-term trends and guide future restoration efforts.

2:05 Discussion of Advances in Nutrient Mass Balance Analyses

Q: James Martin: My question is about the program's relatively recent focus on soil health, and with that, the increasing use of biological amendments as agricultural inputs. Is the nitrogen fixation associated with those biologicals included in your crop fixation estimates, or are those estimates strictly based on legume acreage?

- ***A:*** *Robert Sabo:* I would say that our estimates probably capture those effects only indirectly and roughly on net. We do not explicitly isolate the effects of biological amendments themselves. Our fixation estimates are not acreage-based; they are yield-based. So, if a producer is growing soybeans and adding biologicals that improve soil health, increase micronutrient mineralization, and ultimately raise yields, then we would hopefully capture the net effect of those practices through the higher crop yield. But that is a very interesting development, and I would be very interested in discussing it further. One of the things we are seeing in states like Illinois, Indiana, and Ohio is that some of these soil-health practices appear to be associated with greater rates of phosphorus drawdown. In other words, producers are improving soil health, drawing down excess phosphorus stored in soils, and still continuing to increase crop yields. It is really remarkable that corn and soybean harvests are still rising in those states while phosphorus surplus has become negative. That makes this a very exciting area for future work.

2:15 ADJOURN

Attendees:

Gabriel Duran (CRC)
Anna Shats (KDHE)
Dave Montali (Tetra Tech)
Wenfan Wu (VIMS)
Normand Goulet (NVRC)
Samuel Canfield (WVDEP)
Lewis Linker (EPA)
Clifton Bell (Brown and Caldwell)
Gopal Bhatt (Penn State)
Richard Tian (UMCES)
Scott Heidel (PADEP)
Tyler Trostle (PADEP)
Guido Yactayo (MDE)
Zhengui Wang (VIMS)
Jian Shen (VIMS)
Mukhtar Ibrahim (MWCOG)
KC Filippino (HRPDC)
William Keeling (VADEQ)
Robert Burgholzer (VADEQ)
Pierre St-Laurent (VIMS)
Sushanth Gupta (MWCOG)
Carlington Wallace (ICPRB)
George Onyullo (DOEE)
Jeremy Testa (UMCES)
Joseph Zhang (VIMS)
Olivia Martin (Devereux Consulting)
Ashley Hullinger (PADEP)
Kevin Mclean (VADEQ)
Alex Gunnerson (Koniag)
Cassandra Davis (DEC)
Andrew Sekellick (USGS)
Arianna Johns (VADEQ)
Joseph Delesantro (ORISE)
Anna Kasko (MDE)
Steven Bieber (MWCOG)
Conor Keitzer (UMCES)
Karl Blankenship (Bay Journal)
Carl Friedrichs (VIMS)
James martin (DCR)
Krista Romita Grocholski (RAND)
Maya Struzak (CMU)
Hannah Nuest (KDHE)
Sarah Fakhreddine (CMU)
Jeremy Hanson (CRC)
Alisha Mulkey (MD DA)
Qian Zhang (UMCES)
Lawrence Sanford (UMCES)
Rebecca Murphy (UMCES)
Robert Sabo (EPA)
Zhaoying Wei (UMCES)
James Webber (USGS)
Alexander Soroka (USGS)
Robin Glas (USGS)
Marjy Friedrichs (VIMS)
Sam Merrill (NGEM)
Andrew Stoddard (DS, LLC)
Andrew Leight (MD DA)

Chesapeake Bay Program
Science. Restoration. Partnership.

Modeling Workgroup January Quarterly Review

Day 2 – January 7, 2026

Event webpage: Link

==

Table of Contents (Jump to bookmark):

Day 2 – January 7, 2026:

Back to Day 1.

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

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

Zhengui Wang: Here we will focus on evaluating how the P7 Main Bay Model (MBM) performs after transitioning from a hybrid watershed loading framework to a fully P7 loading system. There are two model configurations: MBM6, which uses P6 loads for major rivers and P7 elsewhere, and MBM7, which applies P7 loading consistently across atmospheric deposition, watershed inputs, and shoreline erosion. The purpose of the analysis is to determine whether this transition affects calibration and model skill, and to compare the updated model not only to the earlier hybrid setup but also to the legacy CH3D (Curvilinear-grid Hydrodynamics 3D model) benchmark.

The assessment has two main components. First, the team compares MBM7 against MBM6 to evaluate the effects of switching to fully P7 nutrient loading. Second, we compare MBM7 against the CH3D benchmark to determine whether the new model represents an improvement over the previous generation of the Bay model. Overall, the P7 MBM has now been calibrated using fully P7 nutrient inputs, and its performance is strong. In broad terms, it is consistent with both MBM6 and the CH3D benchmark. Also, the model includes both a physical component, including hydrodynamics, waves, and sediment transport, and a biogeochemical component, including SAV, marsh, and oyster modules, with sediment transport now activated to support future scenario analyses.

The first major set of results examines watershed nutrient loading differences between P6 and P7. Using nutrient concentrations and comparing them to downstream monitoring data, we founds that P7 loading is generally as good as or better than P6 for most variables, including TN, TP, and TSS) Ammonia performs similarly under both versions, while nitrate is one area where P6 appears slightly better in some smaller tributaries. Even so, the overall conclusion is that P7 loading provides a comparable or improved representation of watershed inputs, supporting its adoption in the MBM.

We compare full model outputs from MBM6 and MBM7 over a 10-year simulation period (1991–2000) across 150 monitoring stations. We found that the two versions are highly similar. Statistical metrics such as RMSE and bias show only minor differences between them. Most variables perform nearly identically, although MBM6 is slightly better for surface TP, and MBM7 shows slightly worse performance for phosphate, which is an area needing further work. Despite these differences, the larger conclusion is that moving to P7 loading does not degrade performance and in some cases offers benefits.

Time series comparisons for chlorophyll-a and DO in the Upper Bay, Mid Bay, Lower Bay, and tributaries reinforce this message. Both MBM6 and MBM7 reproduce similar temporal patterns, and in some cases MBM7 shows modest improvements in DO simulation. The P7 model does a reasonable job capturing seasonal bloom dynamics, especially in the Middle and Lower Bay, though it can overestimate winter–spring blooms in the Upper Bay. These bloom dynamics are important because they influence dissolved organic carbon (DOC) production and, in turn, the timing and severity of hypoxia. Recent parameter adjustments were intentionally designed to better represent these interactions.

A more detailed spatial analysis shows where model errors are concentrated. For chlorophyll-a, errors are generally small, though the model tends to overestimate in parts of the upper Bay and some upper tributaries, while performing well in the lower Bay. For DO, surface DO is somewhat underestimated in the mainstem and slightly overestimated in tributaries, while bottom DO is generally well simulated along the main channel but somewhat overestimated in tributaries and shallow regions. These patterns are important because living-resource modules such as SAV can affect DOC production and oxygen demand, meaning that model adjustments must account for strong interactions across the system.

For nutrients, the P7 model performs well overall, particularly for TN, which shows relatively small RMSE and bias in the mainstem. TP is more problematic, with a tendency toward underestimation in the mainstem and mixed bias elsewhere. For individual nutrient species, nitrate is generally well represented, though some tributaries are underestimated. Ammonia shows overestimation in parts of the main Bay at depth and underestimation in shallow and tributary areas. Phosphate stands out as the clearest weakness, with overestimation in a number of locations and a need for additional refinement; this be a priority issue for future model improvement.

The P7 model performs especially well in many examples across the Bay. Along the main Bay channel, the model captures bottom chlorophyll-a winter–spring bloom peaks quite well, which appears to help reproduce hypoxia timing. In the James River, the model successfully captures strong mid-reach chlorophyll blooms and the upstream-to-downstream spatial gradient. This can be attributed in part to the higher spatial resolution of the P7 model relative to CH3D, which better resolves channels, shallow areas, and nutrient-biological dynamics. Also, when comparing bottom DO fields and vertical DO profiles in the Chesapeake Bay and Potomac River, the model captures the general spatial structure of hypoxia reasonably well, even if some upper-Bay underestimation remains.

The fully P7 loading framework maintains or improves model performance relative to both the earlier hybrid configuration and the legacy CH3D benchmark. The updated model aligns well with observations, performs satisfactorily for most variables, and shows notable improvements in some tributaries and shallow regions. Remaining work will focus on documenting the model and addressing specific weaknesses, especially phosphate and DO in certain locations. The team will also continue working closely with the watershed modeling group to test new beta versions, but overall the P7 MBM appears to be a strong and credible advancement for future Chesapeake Bay scenario analyses.

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

Q: *Lewis Linker:* Zhengui, it is encouraging to see those bottom chlorophyll peaks early in the year. They really look like the fresh, seasonal chlorophyll peaks we would expect, and the MBM is doing a very good job of representing them. As I recall, earlier versions of the model did not capture those bottom chlorophyll peaks nearly as well. What changed?

- ***A:*** *Zhengui Wang:* There were two main changes. The first involved settling velocity. The winter species are mainly diatoms, which have a relatively high settling velocity, so during winter and spring they tend to sink more readily. Once they settle to the bottom, we assume that some of them accumulate there. The second change involved respiration. Because temperatures are low in winter and spring, respiration is weaker than in summer, so the diatoms are able to persist and accumulate near the bottom. In other words, both adjustments work toward the same goal: allowing the diatoms to accumulate in the bottom layer. That seems consistent with the observations, and you can also see some interannual variability at the Middle Bay station that suggests the model is representing the process reasonably well.

Q: *Carl Friedrichs:* For DO at CB3.2, the model is producing values that are quite low. In fact, I see occasions where surface DO drops to around 3 mg/L, which is clearly not right. I am concerned that this may lead us into situations where the model predicts violations that do not actually occur. My impression is that your vertical diffusivity may be too high. If you then look at the bottom DO, you now have very strong hypoxia, perhaps even overestimated, especially between 1999 and 2000. I suspect that what is happening is that low-oxygen water from the bottom is being mixed upward, or the oxygen demand at the bottom is drawing oxygen down from the surface. These plots are one indication of that, and I could point to other examples as well.

- ***A:*** *Zhengui Wang:* After we reviewed these results, we realized that the surface DO was too low. After we sent you the plots, we continued working to improve this. We know that our earlier DO simulation was not quite right. In trying to strengthen oxygen demand, we increased DOC and adjusted other components, but in doing so we seem to have overcorrected, both at the surface and bottom. At the bottom it was not as problematic, but at the surface we clearly saw unrealistically low DO. So we agree with your assessment. This is an issue, and we are working on it. We have already made some additional adjustments and the results have improved, although we have not yet sent you those updated runs. We will continue trying to improve this further.

- ***Comment:*** *Richard Tian:* I just noticed that there is a lot of high-frequency variability in the results. I was wondering whether profile data or sensor data could help us verify whether that kind of variability is realistic. I think there are a couple of stations in the mainstem Bay, and if we could obtain those data and compare the patterns, that might help justify whether this aspect of the simulation is realistic.

- ***Q:*** *Lewis Linker:* At some point we are going to want to bring in those high-frequency stations. I am fairly sure there is at least one deep-water station. I remember the earlier discussion about how to position it so that it would not get run over by shipping traffic while still sampling deep water. So I do think there is something there worth pursuing. The question is: while we are looking at 1991 through 2000 for good reason, because that is our long-term and critical period together, do we also have a simulation running from 1985 up to the more recent period where we now have those high-frequency DO data? Gopal mentioned yesterday that the watershed side has now been extended to 2024. Do we have the capacity in the main Bay model to simulate recent years so that we can compare against those newer high-frequency observations?

- ***Response:*** *Zhengui Wang:* We are working on that now. We have completed the physical model from 1985 through 2020, which is the period for which we currently have watershed loading. The next step is to launch a 36-year run for the biogeochemical model, and then we will see what we can get from that.

10:55 **Testing Spin Up Requirements for the MBM and A Bluefish Benchmark for the MBM – Richard Tian, UMCES-CBPO**

Richard Tian: This presentation will focus on improving model spin-up strategies for Chesapeake Bay scenario simulations, with an emphasis on how the system adjusts to new environmental forcing conditions. When a scenario is applied, the Bay is effectively shifted into a new regime, and both the ecosystem and model require time to equilibrate. This is especially important for sediment-driven processes, which respond slowly and retain memory of past conditions, making proper spin-up critical for realistic simulation results.

The traditional P6 approach used a double-run method (1991–2000), where the first run served as spin-up and the second as the actual simulation. However, this method has limitations. The system does not return to its original state due to legacy effects, and repeating the same time period reduces interannual variability. More broadly, ocean modeling practices suggest repeating initial conditions over multiple years, raising the key question of how long spin-up should last to achieve near-equilibrium conditions.

To address this, the team tested 3-, 5-, and 10-year spin-up durations, using hypoxic volume as a key indicator of convergence. The results show that spin-up is essential: even a 3-year spin-up dramatically reduces hypoxic volume compared to no spin-up. However, differences between 3- and 5-year runs remain substantial, and even 10 years does not fully converge with shorter runs. This indicates that short spin-up periods are insufficient, and longer durations are needed to adequately stabilize the system.

The rate of convergence varies across regions. Open water equilibrates relatively quickly, with 5- and 10-year spin-ups producing nearly identical results. In contrast, deep channel and especially deep water regions converge more slowly, reflecting stronger sediment legacy effects. Quantitatively, deep water shows large annual changes (~12%) after 3 years, still notable changes (~4%) after 5 years, and much smaller changes (~1%) after 10 years. Additional comparisons between 9- and 10-year spin-ups show only minimal differences (~0.5%), suggesting that 10 years is a reasonable practical compromise, even if full equilibrium is not perfectly achieved.

The updated spin-up approach also improves upon the P6 method. Differences of up to 13% in deep-channel hypoxic volume indicate that the new method achieves more stable and realistic conditions. This improvement is partly due to differences in hydrologic sequencing: the earlier approach included higher-loading years in the spin-up, delaying equilibration, whereas the updated method begins with a lower-loading year, allowing the system to adjust more efficiently. As a result, the new approach better captures system dynamics and reduces bias from legacy conditions.

In addition to spin-up improvements, we looked at a major challenge in modeling SAV. Currently, the model cannot reliably predict SAV acreage, which limits its use for direct assessment of SAV criteria. As an alternative, the team is exploring ML approaches. An initial test using a RF model to predict SAV distribution based on water quality outputs shows promising results, with strong agreement between observed and predicted patterns ($R^2 \approx 0.98$). Key predictors include sediment composition (especially fine sediments), light availability, suspended solids, chlorophyll, bathymetry, salinity, and temperature.

However, challenges remain in applying this approach to future scenarios, since ML models like RF can struggle with extrapolation beyond observed conditions. The team is still in early stages of exploring how to address this limitation and develop a robust method for predicting SAV under changing environmental conditions. This represents an important area for future research and collaboration.

In conclusion, we recommend adopting a 10-year spin-up period as the best balance between accuracy and computational cost, with potential improvements such as initializing scenarios from partially equilibrated states. The model has also been successfully run on cloud platforms, though consistency between calibration and scenario runs is important. At the same time, developing a reliable method for SAV assessment remains a priority, with ML approaches offering a promising but still evolving pathway.

11:15 Discussion of Spin Up Test & Bluefish Benchmark

Comment: *Carl Friedrichs:* What we are doing here is different from what we did decades ago. Back then, when we determined spin-up needs, we did not look at predicted hypoxic volume and how it approached equilibrium. Instead, we focused on the bottom sediments, because they take years to equilibrate. That is how we originally arrived at the 10-year spin-up period: by looking at sediment behavior rather than the water column. I do not think there is anything wrong with the approach being presented now; it is simply different. If there are implications, I am not sure exactly what they are, but one thing I would say is that because sediments change much more slowly than the water column, there is a good argument for using the full 10-year period to represent sediment equilibration rather than repeatedly spinning up only the initial year. That may be a better way to integrate the slower sediment response over time. If the group decides that this newer method makes sense, that is perfectly fine.

- ***Response:*** *Lewis Linker:* One option is to take 1991 and hyper-equilibrate the sediments around that year, and you have given us good arguments for that. Another option is to use a more generalized, decade-long sediment condition that represents both dry years like 1991 and wetter years, perhaps even including extreme events like the January 1996 Conowingo melt event. In other words, do we want a sediment state that is highly tuned to the initial year of the run, or a more generalized, long-term equilibrium state to use when launching scenarios? I think Carl's point in the end is that there may not be a single correct answer and that these are simply two different approaches. But we will need to decide which one we want to use and be clear about why. There are real implications either way. As you have shown here, it takes time to equilibrate the sediments, and that matters not only for what was originally assessed in terms of sediment diagenesis and state variables, but also for a management-relevant endpoint like hypoxic volume. One immediate implication of the method you are showing is that long-term hypoxia would generally be reduced under this approach. This will require further discussion.

- ***Comment:*** *Zhengui Wang:* I think one idea we could try in the future is to build a baseline database from a long run like 40 years and then, if we later introduce a new grid, we could use that database to estimate initial conditions for both the sediment and the water column. If we have a 40-year run beginning in 1985, then by 1991 we would already have 5-6 years of effective spin-up. Right now, when we do a 10-year spin-up and then use the end state from 2000 to reinitialize conditions in 1991, I worry that we may be overlooking some interannual variability. When I run the model, I can see that conditions in the early 1990s, when loading was relatively high, are not necessarily the same as those in the late

1990s or around 2000. So if we use the 2000 end state to initialize 1991, we may be smoothing over some of that natural variability. For example, the first 3 years may actually be more similar to 1991 than the year 2000 is.

- ***Response:*** *Lewis Linker:* that is a very good idea and something Richard has also mentioned: developing a library of equilibrated sediment conditions that we can pull from to pre-equilibrate future runs. If we know we are going to run many WIP 3-type scenarios, or variations of them, then having those stored sediment states available could give us a good head start. I have to say, our computational times are going to be a serious constraint, so whatever tricks we can use to make these runs more efficient are going to be very welcome.

~~11:25 Update on Conowingo Model Development – Earl Hayter, Jodi Ryder, CoE-ERDC and Matt Rowe, MDE~~
~~Progress in the development of the Conowingo Model as well as initial results of the Conowingo sediment sampling will be reviewed.~~

~~11:45 Discussion of the Conowingo Model Development and Application~~

11:55 LUNCH

12:50 Progress on Patapsco/Back MTM – Harry Wang, Breanna Maldonado, VIMS and Jeremy Testa, UMCES
Harry Wang: Here we provide a progress review of the second year of the Patapsco–Back River Multiple Tributary Model (MTM), with a focus on improving calibration of TSS. The work is divided between Jeremy Testa, who developed a method to estimate inorganic suspended solids (ISS), and the broader calibration effort. TSS is emphasized because it is a critical integrator of hydrodynamics, sediment transport, and biogeochemistry, and it strongly influences light availability, which in turn affects nutrients, oxygen dynamics, and SAV.

A key challenge in calibrating TSS is that it consists of both inorganic (ISS) and organic (volatile suspended solids, VSS) components. Without separating these, calibration becomes a nonlinear problem where parameters interfere with one another. We developed a method to estimate ISS from existing monitoring data. Using relationships between particulate carbon (PC) and VSS, derived from stations where both variables were measured, and we were able to estimate VSS from PC and then calculate ISS by subtracting VSS from TSS. Although this method relies on assumptions about spatial consistency, it provides a practical way to generate ISS estimates and enables a more efficient, sequential calibration approach.

PC and VSS are reasonably well correlated, with VSS consisting of roughly 40% carbon. Using this relationship, we estimated ISS at several stations in and near the Patapsco–Back River domain. Results indicate that large TSS peaks are primarily driven by inorganic material, rather than organic components. These ISS estimates provide a new and valuable dataset that can be directly compared to model outputs during the physical calibration stage, improving confidence in sediment dynamics before moving to water quality calibration.

Building on this, ISS was used in iterative calibration of TSS. Calibration in the Patapsco and Back River system is particularly challenging because these environments are dominated by fine, cohesive, organic-rich sediments influenced by wastewater inputs and frequent resuspension from ship traffic. However, the current model framework uses a noncohesive sediment formulation, requiring extensive parameter tuning. Initial calibration attempts showed underprediction of TSS, which led to insufficient resuspension and unrealistically high DO levels in some locations.

To explore the parameter space, we tested an over-calibration scenario by lowering critical shear stress and settling velocity, and increasing the fraction of fine sediments. This resulted in excessive resuspension, leading to overprediction of TSS, chlorophyll-a, TN, and TP, as well as overly strong anoxic conditions. These two cases of under- and over-calibration, effectively bracket the solution, demonstrating that the correct parameterization lies somewhere in between.

Through iterative adjustments, we identified intermediate parameter combinations that better match observed conditions. Comparisons with ISS estimates show that the MTM can achieve more realistic sediment concentrations, particularly in Baltimore Harbor. The analysis also highlights differences between the MTM and the MBM, including the impact of model resolution and geometry. For example, features like Pleasure Island are resolved in the MTM but not in the coarser MBM, affecting flow patterns and sediment resuspension. These structural differences can influence calibration outcomes and may require additional refinements.

The results also reinforce the importance of TSS calibration for the broader system. Changes in TSS directly affect light penetration and particulate nutrient dynamics, which in turn influence chlorophyll, DO, TN, and TP. This interconnectedness means that accurate sediment calibration is essential for achieving realistic water quality simulations. The use of ISS estimates helps isolate the physical sediment processes, making it easier to diagnose and correct model behavior.

In summary, the presentation highlights three key advances. First, the development of an ISS estimation method transforms TSS calibration into a more linear and efficient process. Second, iterative calibration guided by ISS, allows us to systematically refine model parameters in a complex, sediment-dominated system. Third, the approach reduces computational burden by enabling earlier validation at the physical modeling stage, rather than waiting for full water quality simulations. Moving forward, we plan to continue refining calibration, address remaining discrepancies, and transition toward P7 loading conditions, with the expectation that this improved framework will enhance model performance in the Patapsco–Back River system.

1:10 Discussion of Patapsco/Back MTM Progress & P7 linkage

Q: *Jeremy Testa:* What really struck me is that in Back River, at WT4.1, the MTM result shows these massive peaks, likely associated with big flow events in some of those wet years. Because the axis is truncated, it is hard to tell just how large they are. It alternates between periods with almost no ISS and then very large amounts of ISS. I am still trying to make sense of that, especially in the context of spatially varying shear stress. That might be a more muted and realistic way to address the issue of representing inorganic sediments better, because if the bed is primarily silty and very easily resuspended in a shallow area, then perhaps changing the bed composition alone is causing unrealistically large swings.

- ***A:*** *Harry Wang:* To improve that, I think the regular tidal velocity needs to be higher so that during the non-event periods there is still some resuspension. Right now, those big red peaks at WT4.1 are tied mainly to events, but during quiescent periods the sediment is not resuspending enough. That is what drew my attention to the tidal configuration at the mouth and why I noticed that the MBM, without Pleasure Island, has roughly twice the velocity there.

- ***Response:*** *Jeremy Testa:* If the tidal velocity is higher, then you get more regular resuspension, and therefore fewer extremes between quiet periods and event-driven peaks. But then in the MTM you are also getting those very large resuspension events, which might reflect how we are characterizing the sediment composition itself.

- ***Comment:*** *Harry Wang:* The distribution of sediment classes definitely matters, because in the current setup I have to increase the proportion of clay 1 to get a higher concentration. But if we have the option for spatially varying critical shear stress, then the calibration becomes much easier. We could simply specify that Back River should have a lower critical shear stress and let the model respond accordingly. Right now, we cannot do that directly. The model gives us velocities, but we do not have measurements there, so we do not know which model is more correct.

1:20 **Transport and Retention Processes and Their Water Quality Implications in the Middle–Lower Rappahannock River – Qubin Qin, East Carolina University and Jian Shen, Zhengui Wang, and Pierre St-Laurent, VIMS**

Qubin Qin: This presentation summarizes a modeling study examining how hydrodynamics influence transport and water quality in the Rappahannock River, conducted as part of ongoing P7 model integration efforts. The work leverages a calibrated hydrodynamic model combined with a particle tracking framework to better understand how physical processes control the movement and retention of materials in the river. The study is motivated in part by the limited historical research on this system, with prior work primarily focused on hypoxia, benthic communities, and submerged aquatic vegetation.

A key feature of the Rappahannock is a shallow region in the middle-lower estuary, which previous studies identified as an important divider between different ecological regimes: one associated with hypoxia and another with elevated chlorophyll concentrations. This study revisits that finding by asking why the shallow region matters and whether its influence is driven by hydrodynamic processes. To test this, the team developed two scenarios: a baseline case and a dredging scenario in which the shallow region is deepened, allowing for direct comparison of how depth affects circulation and transport.

The first set of results focuses on salinity distributions. Across both spring and summer conditions, the dredging scenario shows a clear upstream shift in salinity zones, indicating stronger saltwater intrusion. This behavior is consistent with known estuarine physics: increasing water depth enhances gravitational circulation, strengthening exchange flow and allowing saltwater to penetrate farther upstream. These findings demonstrate that even modest deepening can significantly alter large-scale estuarine structure.

The study then uses particle tracking simulations to examine how materials originating from the watershed move through the river system. Particles are released from the headwaters and tracked over time, with scenarios representing both passive (non-settling) and active (settling) materials. For passive particles, dredging generally leads to reduced particle age, meaning faster downstream transport. This effect is most pronounced during spring high-flow conditions and reflects enhanced exchange flow caused by the deeper channel.

However, when particles are assigned settling velocities, the transport behavior becomes more complex. Instead of a consistent pattern of faster downstream movement, the results show a mix of increased and decreased transport rates depending on location and conditions. This complexity arises from interactions between vertical particle movement and stratified flow, introducing nonlinear dynamics that make it difficult to generalize the effects of dredging across all material types.

Additional analysis of particle number and spatial distribution shows that dredging alters transport patterns throughout the entire river, not just within the modified region. In some cases, particle concentrations decrease within the dredged area while increasing upstream or downstream, indicating system-wide impacts. These findings suggest that localized physical changes can propagate through the estuarine system in non-intuitive ways.

Overall, the study demonstrates that modifying bathymetry can significantly influence hydrodynamics, salinity intrusion, and material transport in the Rappahannock River, especially in shallow regions. These changes have important implications for water quality, including the transport of sediments, nutrients, carbon, and DO. The work reinforces the importance of accurately representing hydrodynamics as the foundation for reliable water quality modeling.

In conclusion, this analysis highlights how physical structure and circulation processes shape estuarine behavior, with even subtle changes in depth producing measurable impacts. The findings provide valuable insight for future modeling efforts and management considerations, emphasizing that understanding and accurately simulating hydrodynamics is essential for predicting water quality outcomes in systems like the Rappahannock River.

1:40 Discussion of Rappahannock MTM Progress

Comment: *Lewis Linker:* There are two things I really appreciate about your presentation. First, the ability to confidently test scenarios like deepening or enlarging a channel and directly evaluate how that would influence water quality. Second, the modeling tools themselves, allow us to assess those implications in a meaningful way. This kind of approach could be useful not only for the Rappahannock but also for other areas of the Bay. For example, thinking about the Rappahannock Shoal region, there may be cases where deepening a channel could help ventilate deeper waters like those in CB4.

1:50 ADJOURN

Attendees:

Gabriel Duran (CRC)
Lewis Linker (EPA)
Andrew Stoddard (DS, LLC)
Dave Montali (Tetra Tech)
Richard Tian (UMCES)
Neil Kamal Ganju (USGS)
Qubin Qin (ECU)
Tyler Trostle (PADEP)
Carl Cerco (USACE)
Wenfan Wu (VIMS)
Gopal Bhatt (Penn State)
Clifton Bell (Brown and Caldwell)
John Lancaster (PA)
Scott McLaughlin (NGEM)
Joseph Delesantro (ORISE)
Ashley Hullinger (PADEP)
Zhengui Wang (VIMS)
Mukhtar Ibrahim (MWCOG)
Samuel Canfield (WVDEP)
Joseph Zhang (VIMS)
Guido Yactayo (MDE)
Normand Goulet (NVRC)
KC Filippino (HRPDC)
Carlington Wallace (ICPRB)
Cathy Wazniak (MDDNR)
Sophia Grossweiler (MDE)
Rebecca Murphy (UMCES)
Lawrence Sanford (UMCES)
George Onyullo (DOEE)
Jiabi Du (TAMU)
Jeremy Testa (UMCES)
Carl Friedrichs (VIMS)
William Keeling (VADEQ)
Pierre St-Laurent (VIMS)
Arianna Johns (VADEQ)
Kevin Mclean (VADEQ)
Harry Wang (VIMS)
Marjorie Zeff (AECOM)
Christina Lyerly (MDE)
Steven Bieber (MWCOG)
Alexander Soroka (USGS)
Early Hayter (USACE)
Jian Shen (VIMS)
Hannah Nuest (KDHE)