Chesapeake Bay Program
Science. Restoration. Partnership.

Integrated Trends Analysis Team (ITAT) & Factors Joint Meeting

Wednesday, August 27th, 2025

10:00 AM – 11:30 PM

Meeting Materials: Link

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

MINUTES

10:00 – 10:05 AM *Welcome – Breck Sullivan (U.S. Geological Survey, USGS) and Kaylyn Gootman (Environmental Protection Agency, EPA)*

Announcements:

- Newly published Potomac Tributary Summary Report! You can access the report on the ITAT Tributary Summaries Project webpage here.
    - Call for reviewers (Thank you to all that already volunteered for the other reports!):
        1. "Maryland Mainstem Tributary Geonarrative: A summary of trends in tidal water quality and associated factors, 1985 - 2024." (Will be completed ~2 weeks).
        2. "York Tributary Summary: A summary of trends in tidal water quality and associated factors, 1985 - 2024." (Will be completed ~2 weeks).

Conferences:

- Coastal & Estuarine Research Federation (CERF) 28th Biennial Conference, November 9-13th, 2025. Richmond, VA. ***Registration is now open!***
- Alliance for the Chesapeake Bay – Chesapeake Watershed Forum, November 7-9, 2025. National Conservation Training Center, Shepherdstown, WV. ***Proposal for Posters now open!***

10:05 – 10:45 AM *Hypoxia Forecasting for Chesapeake Bay Using Artificial Intelligence, Artificial Intelligence for the Earth Systems*

Presenter(s): *Guangming Zheng (University of Maryland, UMD)*

Description: The abstract for Guangming Zheng's recent publication [link] follows: Seasonal hypoxia is a recurring threat to ecosystems and fisheries in the Chesapeake Bay. Hypoxia forecasting based on coupled hydrodynamic and biogeochemical models has proven useful for many stakeholders, as these models excel in accounting for the effects of physical forcing on oxygen supply, but may fall short in replicating the more complex biogeochemical

processes that govern oxygen consumption. Satellite-derived reflectances could be used to indicate the presence of surface organic matter over the Bay. However, teasing apart the contribution of atmospheric and aquatic constituents from the signal received by the satellite is not straightforward. As a result, it is difficult to derive surface concentrations of organic matter from satellite data in a robust fashion. A potential solution to this complexity is to use deep learning to build end-to-end applications that do not require precise accounting of the satellite signal from the atmosphere or water, phytoplankton blooms, or sediment plumes. By training a deep neural network with data from a vast suite of variables that could potentially affect oxygen in the water column, improvement of short-term (daily) hypoxia forecast may be possible. Here, we predict oxygen concentrations using inputs that account for both physical and biogeochemical factors. The physical inputs include wind velocity reanalysis information, together with 3D outputs from an estuarine hydrodynamic model, including current velocity, water temperature, and salinity. Satellite-derived spectral reflectance data are used as a surrogate for the biogeochemical factors. These input fields are time series of weekly statistics calculated from daily information, starting 8 weeks before each oxygen observation was collected. To accommodate this input data structure, we adopted a model architecture of long short-term memory networks with eight time steps. At each time step, a set of convolutional neural networks are used to extract information from the inputs. Ablation and cross-validation tests suggest that among all input features, the strongest predictor is the 3D temperature field, with which the new model can outperform the state-of-the-art by ∼20% in terms of median absolute error. Our approach represents a novel application of deep learning to address a complex water management challenge.

Guangming Zheng: Here we will present a study where we developed a machine learning model, HypoxAI, to predict dissolved oxygen (DO) in the Chesapeake Bay. The goal was to compare this model with the traditional hydrodynamic model Chesapeake Bay Environmental Forecast System (CBEFS), which is widely used for forecasting in the region. While HypoxAI is now a few years old and somewhat outdated compared to the latest AI methods, it demonstrated clear advantages in prediction accuracy at the time it was developed.

The model was designed using a combination of convolutional neural networks (CNNs) and long short-term memory (LSTM) layers. Inputs included physical variables from CBEFS (such as temperature, salinity, wind and currents), wind data from the European Centre for Medium-Range Weather Forecasts (ECMWF), and phytoplankton indicators derived from MODIS Aqua satellite data, along with in-situ DO measurements from the Chesapeake Bay Program. Rather than relying on single snapshots, the model used weekly averaged time-series data extending up to eight weeks prior to each DO measurement, combined with a spatial “box” around the sampling location to capture surrounding influences. This approach provided richer context for the model to learn from.

The training dataset was split by year to ensure independence of training, validation, and testing data. Altogether, the model was trained on ~140,000 samples, validated on 10,000, and tested on 11,000, with each sample representing a specific location, depth, and time.

The architecture allowed CNNs to extract spatial and spectral patterns (e.g., stratification and phytoplankton signals), which were then processed through LSTMs to capture temporal dependencies. Despite its modest size (~22,000 parameters), the model showed notable improvements: when tested against the untouched 2019–2020 dataset, HypoxAI reduced mean absolute error by ~20% compared with CBEFS and achieved higher R^2 values.

Results varied seasonally. In spring, HypoxAI often captured the onset of hypoxia more effectively than CBEFS, though both models sometimes diverged from in-situ measurements. In summer, when hypoxia was well established, the two models performed comparably, with the machine learning model sometimes smoothing sharp vertical gradients. An ablation study revealed that temperature was by far the most important predictor, nearly matching the performance of the full model on its own. This was somewhat surprising, as the team initially expected phytoplankton indicators from satellite data to play a stronger role. However, the use of Rayleigh-corrected MODIS reflectance appeared to introduce too much noise to be consistently useful even though it was chosen to maximize data availability in the turbid Bay.

Looking forward, there are more advanced satellite-derived products, particularly new chlorophyll concentration models built with transformer-based architectures, that could provide better indicators of phytoplankton biomass. Another avenue is to test whether DO can be predicted effectively using satellite-only data, such as sea surface temperature and ocean color, without relying on CBEFS outputs. Overall, the study demonstrated that machine learning can outperform state-of-the-art hydrodynamic models for short-term DO prediction, while highlighting both the promise and challenges of integrating satellite and model-based inputs.

Discussion:

Comment: *Dante Horemans:* In your conclusion, you mentioned using more recent satellite data and not using model-derived temperature results. About three weeks ago, I attended the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) Hackweek in Baltimore, where our group forecasted oxygen in the Chesapeake Bay using only PACE data. That included around 150 reflectance bands, satellite-derived temperature, and Photosynthetically Active Radiation (PAR). We found interesting and similar results, particularly that temperature was the most important predictor, and PACE data was also significant in our application. We focused on model-derived oxygen as the response variable, but the next step would be to use actual observations. Using PACE data along with satellite-derived PAR and temperature, our machine learning model achieved an R-squared of 80%. I would be happy to discuss this further and can send you an email with more details.

- ***Response:*** *Guangming Zheng:* Thank you for sharing this interesting study. It sounds very promising. We should definitely connect after the talk to discuss details, ideas, and potential collaborations.
- ***Q:*** *Dante Horemans:* You used MODIS reflectance data and I wanted to ask how many bands did you use?

- ***A:*** *Guangming Zheng:* We used all available bands. The idea was to allow the model to learn potential patterns. However, that approach is brute force and may not be the best. Feature ranking or feature engineering could help select the most relevant bands. Another approach is to improve the ocean color model so it can better extract chlorophyll information. With a more accurate chlorophyll product, we wouldn't have to worry as much about which features to choose.
- ***Q:*** *Dante Horemans:* Wouldn't chlorophyll information already be embedded in the raw reflectance data? Why would Chlorophyll-*a* be a better predictor than reflectance itself?
- ***A:*** *Guangming Zheng:* Chlorophyll is important because most organic matter in the Chesapeake Bay is produced locally rather than discharged. Chlorophyll directly relates to the organic matter that eventually decomposes and consumes oxygen at the bottom.
- ***Response:*** *Dante Horemans:* We focused on model-derived oxygen because PACE was only launched in 2024, so we didn't have enough observed data. But in the future, it would be valuable to test our model against actual observations. We only had three days to work on our application, so we used what was available. We didn't have time to implement CNNs, so we used neural networks with three to four layers, as well as random forests. The results were similar, but I'd like to explore CNNs further.
- ***Q:*** *Elgin Perry:* Does the Shapley method try to interpret the coefficients inside the neural network layers to see which pathways are important?
- ***A:*** *Dante Horemans:* It comes from game theory and scores the contribution of each variable. It shows the relative importance of input variables, for example, how much temperature influences oxygen concentrations.
- ***Q:*** *Elgin Perry:* So does it work by changing variables randomly to see how results shift through the model?
- ***A:*** *Dante Horemans:* Exactly. It perturbs the variable and measures how much it changes the response.

Q: *Elgin Perry:* Amongst your x variables, you didn't include any freshwater input. Did you consider this or eliminated this as a predictor because of salinity?

- ***A:*** *Guangming Zheng:* Freshwater input is important, but it's already handled by the CBEFS model. CBEFS includes terrestrial inputs as boundary conditions when running. Since our predictor variables come from CBEFS output, the influence of freshwater input is already reflected in the temperature and salinity data.

10:45 - 11:30 PM *RIM Loads and Trends Through Water Year 2024*

Presenter(s): *James Webber (USGS).*

Description: Presenters will provide updated nitrogen, phosphorus, and suspended-sediment loads, and changes in loads, in major rivers across the Chesapeake Bay Watershed that were calculated using monitoring data from the Chesapeake Bay River Input Monitoring (RIM)

Network stations for the period between 1985 through 2024. You can find the [full data release here](#).

James Webber: Here we will present updates on nutrient and sediment monitoring from the River Input Monitoring (RIM) Network, which tracks nitrogen, phosphorus, and suspended sediment (SS) from nine major rivers representing about 78% of the Chesapeake Bay watershed. Data are collected monthly, with additional storm-event sampling, and results are updated annually. To separate management signals from weather-driven variability, long-term (since 1985) and short-term (2015–2024) trends are calculated using "flow-normalized" loads, which remove the influence of wet or dry years.

Overall results show progress but with mixed patterns across locations and parameters. Combined RIM data indicate a 33% reduction in flow-normalized nitrogen loads since 1985 and a 9% reduction over the past decade. Phosphorus and sediment also show decreases relative to 2015, though less consistently and only slightly below 1985 levels. Spatially, the Choptank River continues to deliver the highest nitrogen yield per acre, reflecting its agricultural land use, while the Rappahannock and James rivers stand out for elevated phosphorus and sediment yields.

Trend analysis reveals regional contrasts. Northern stations, including the Susquehanna, Patuxent, and Potomac, generally show improving nitrogen trends, with the Susquehanna improving across all three parameters since 2015. In contrast, many Virginia rivers (e.g., Rappahannock, Mattaponi, Appomattox) exhibit degrading trends, especially for phosphorus and sediment. Areas of concern include the Choptank, where phosphorus loads have been consistently high and continue to rise, and Virginia tributaries where degrading trends persist.

It is important to place monitoring results in the context of Bay Program water-quality goals. When measured relative to 1995 baselines many stations show increased loads of nitrogen, phosphorus, and sediment. When compared with Watershed Implementation Plan (WIP) loads, most rivers remain above their modeled targets, with only the Patuxent currently below its WIP load. These comparisons highlight progress in reducing loads but also underline remaining gaps to reach water-quality goals. You can see the [full data release here](#).

Discussion:

Comment from chat: *Breck Sullivan:* We will have updated tidal trends for 2024 in October! Thank you to our state partners for helping us compile this information!

Q*: Carl Friedrichs:* Great presentation! I had some questions about recent trends, especially in relation to water clarity and submerged aquatic vegetation (SAV). I co-authored a paper that looked at water clarity using the light attenuation coefficient (Kd) along the Bay's main stem. In 2024, Kd showed the clearest water since monitoring began in 1990. Also, the 2024 SAV monitoring found the largest extent of SAV ever mapped in the polyhaline zone, where eelgrass is concentrated, since systematic monitoring began in 1984. This is encouraging, especially since eelgrass is sensitive to rising temperatures. It seems like there is a disconnect between nutrient and sediment input trends from the RIM stations, particularly nitrogen in the Virginia tributaries, and the observed improvements in water clarity and SAV in the main stem. Do you have thoughts on this?

- ***A:*** *James Webber:* I appreciate the connection you're making between water quality and living resources. One possibility is the reduced sediment delivery from the Susquehanna River, which has shown a marked decrease in our monitoring data. Given its size, reductions there could have an outsized effect on the main stem compared to tributary inputs.
- ***Comment from chat:*** *Alex Soroka:* Consider decrease of SS in Susquehanna, and it delivers ~70% of load.
- ***Response:*** *Carl Friedrichs:* That's a good point. Much of the sediment from tributaries doesn't reach the main stem, and the Susquehanna is the largest source. Modeling work by Carl Cerco also suggested shoreline erosion is a major source of sediment in the main stem, and shoreline armoring has reduced erosion steadily.
- ***Q:*** *Carl Friedrichs:* is there a product that combines RIM station data with modeled estimates of additional loads from small rivers, creeks, and shorelines downstream?
- ***A:*** *James Webber:* Yes, that product exists and is uploaded on the Water Quality Standards Chesapeake Progress site. The loads presented here represent about 78% of the watershed area. Additional downstream contributions are estimated through modeling and included in the total load indicator available on the website.
- ***Comment from chat:*** *Kaylyn Gootman:* Indicator with monitoring and modeling link. The second indicator is a combination of monitoring and modeling.
- ***Comment from chat:*** *Qian Zhang:* FYI: For the below-RIM load, Chesapeake Assessment Scenario Tool (CAST) data is processed to "add" interannual variability to obtain the "true-condition" load, which is then combined with the RIM load to get the graph on the WQSAM page for the "Loads to the Bay indicator".

Q: *Elgin perry:* @Carl Friedrichs, since degradation is happening more in the Virginia tributaries, which are closer to the mouth of the Bay, could the shorter residence time of their contributions explain why their effects seem less persistent compared to the Susquehanna?

- ***A:*** *Carl Friedrichs:* For fine sediment, residence times can be very long. In the York River, geological analysis shows a residence time of about 70 years, unless a major event like Hurricane Agnes flushes material out. Sediment transport modeling indicates net transport is actually from the Bay into the York due to estuarine circulation. Similar patterns hold for the Rappahannock and James Rivers. Day-to-day, net sediment transport at their mouths is upstream. Even the ocean is a net source of sediment into the Chesapeake Bay.
- ***Q from chat:*** Alex Soroka: How much of water clarity is driven by biotic activity vs. sediment?
- ***A:*** *Carl Friedrichs:* This relates to work by Jesse Turner, one of my former PhD students. Her research looked at the roles of biotic activity versus inorganic sediment in water clarity. Secchi depth is strongly affected by biotic activity, while

SAV responds more to light penetration affected by inorganic sediment. For years, these two measures diverged: SAV clarity improved while Secchi clarity worsened due to algae growth from available nutrients. Since 2009, both measures have improved, reflecting reductions in nitrogen input. Chlorophyll levels closely match Secchi depth trends, rising until 2009 and then declining since. These patterns align with the Phase 6 model results, showing that both biotic activity and sediment play roles, but in different ways depending on the metric used.

- ***Comment from chat:*** Breck Sullivan: This might be the paper by Jessie Turner (link).
- ***Response from chat:*** Alex Soroka: Thanks Carl, we took visiting colleagues around the Baltimore harbor yesterday and marveled at the luscious brown water from a bloom.
- ***Comment from chat:*** *Rebecca Murphy:* Very generally speaking, consistent with what Carl said, we're seeing better short-term trends than long-term trends in Secchi across the tidal waters. Stay tuned to the tidal trends!

Comment: Breck Sullivan: One of your slides compared monitored loads with modeled WIP loads. A new outcome is being proposed called "Reducing Excess Nitrogen, Phosphorus, and Sediment", which calls for demonstrating reductions using multiple lines of evidence, including monitoring. The results you shared could directly support this outcome.

- ***Response:*** *James Webber:* We want these results to be useful and informative, and I agree they should be supported with multiple lines of evidence. Adding context and reference lines helps us better communicate the trends, even if the comparisons aren't perfect.

11:30 PM Adjourn

Next Meeting: Wednesday September 24th, 2025, from 10 AM – 12 PM

Attendees:

Breck Sullivan (USGS), Kaylyn Gootman (EPA), Gabriel Duran (CRC), Guangming Zheng (UMD), James Webber (USGS), Alex Soroka (USGS), Elgin Perry (CBP Contractor), Allison Welch (CRC), Cynthia Johnson (VA DEQ), Michael Lane (ODU), Chris mason (USGS), Rebecca Murphy (UMCES), Doug Moyer (USGS), Marjy Friedrichs (VIMS), Mukhtar Ibrahim (MWCOG), Andrew Keppel (MD DNR), Anthony Timpano (VADEQ), Qian Zhang (UMCES), Tish Robertson (VA DEQ), Robert Hirsch (USGS), Carl Friedrichs (VIMS), Helen Golimowski (Devereux Consulting), Richard Tian (UMCES), Nicholas Santoro (USGS), George Onyullo (DOEE), Rikke Jepsen (ICPRB), Jeremy Cox (Bay Journal), Dante Horemans (W&M), Efeturi Oghenekaro (DOEE), Jeremy Hanson (CRC), and Andrew Sekellick (USGS).