



Chesapeake Community Research Symposium 2026

Session 3: Advancing Chesapeake Bay Water-quality Science and Management: II.
Novel analysis and scientific communication approaches to inform management

Session Leads: Qian Zhang, James Webber, Rebecca Murphy, Kaylyn Gootman

Ashok Jacob (The Pennsylvania State University), Raj Cibin

A Deep Learning Framework for Continuous Stream Nitrate Estimation across the Chesapeake Bay Watersheds

Continuous stream nitrate monitoring is critical for devising effective management strategies and evidence-based policy regulations in the Chesapeake Bay watersheds, where nutrient fluxes strongly influence water quality and ecological conditions. High-frequency in-situ sensors provide detailed information on sub-daily and event-driven nitrate dynamics that are often missed by traditional grab sampling. Such high-frequency observations can be leveraged to generate continuous stream nitrate information at all sites, including grab sampling locations, across the Chesapeake Bay. However, in situ sensor datasets often contain data gaps and discontinuities, which constrain their broader applicability. This study develops a deep learning framework that integrates watershed-specific attributes and hydro-climatic forcings to fill sensor data gaps and generate reliable continuous nitrate datasets. We hypothesize that similarity in watershed responses can be used to group monitoring sites, thereby reducing the data requirements needed to train robust deep learning models. To test this hypothesis, we evaluate multiple modeling schemes, including global, cluster-based, and single-site configurations, to examine trade-offs among data availability, model generalizability, and predictive accuracy. The gap-filled in situ datasets could subsequently be used to generate continuous nitrate estimates at grab sampling sites across the Chesapeake Bay in the future. Overall, the study aims to provide insights that can support the design of adaptive monitoring networks, optimize sensor deployment, and inform policy-relevant nutrient management strategies in the Chesapeake Bay.

Quinn Domanski (Anne Arundel Community College Environmental Center)

Investigating Bidirectional Dynamics in Chesapeake Bay Tributaries Using Long-Term Monitoring Data and Machine Learning

Long-term monitoring across the Chesapeake Bay watershed has documented overall declines in nitrogen loading to the tidal Bay, yet nutrient dynamics within tidal tributaries remain complex. In particular, observations from various tributaries leading into the Bay suggest elevated nitrate and nitrite concentrations near the mouth of tributaries, indicating the potential for bidirectional nutrient exchange between the Bay and its tributaries. This study investigates these patterns using long-term monitoring data and predictive modeling approaches.

For the last 40 years the Maryland Department of Natural Resources Eyes on the Bay program has collected high quality nutrient data from Chesapeake Bay and its tributaries. These data can be analyzed for patterns and used to train predictive models, especially when data harmonization and quality control techniques are performed to enable consistent comparisons with other relevant data sources for weather patterns, flow control patterns, and other heterogeneous ecological data . Preliminary analyses indicate recurring seasonal concentration peaks in several forms of nitrogen, motivating further investigation of longitudinal and tidal influences.

Building on these datasets, machine learning models are being developed and tested to evaluate and predict conditions associated with elevated nitrate and nitrite concentrations using seasonal, hydrologic, and other ecological variables, including the potential influence of the Conowingo Dam discharge rates on downstream nutrient dynamics. By integrating long-term monitoring with predictive modeling, this work aims to improve our understanding of reciprocal nutrient exchange processes and support more informed water-quality assessment and management of the Chesapeake Bay watershed.

Xueting Pu (Pennsylvania State University)

Toward Generalizable and Interpretable Sediment Modeling with AI-Augmented HSPF

Sediment is a major contributor to water quality degradation, ecosystem disturbance, and reservoir siltation in numerous watersheds, including those draining to the Chesapeake Bay. Effective management of sediment dynamics necessitates models that accurately quantify sediment loads and capture their spatial variability across diverse landscapes. Traditional modeling frameworks, such as the Hydrological Simulation Program—FORTRAN (HSPF), often employ spatially static parameterizations, which can obscure critical regional heterogeneity in sediment generation and mobilization processes. Furthermore, the manual and labor-intensive calibration of sediment-related parameters in HSPF introduces subjectivity and significantly contributes to model uncertainty.

In this study, we introduce an innovative artificial intelligence (AI)-augmented modeling framework designed to enhance HSPF's sediment simulation capabilities. This novel approach seamlessly integrates neural-networks-based regional parameter learning, partial physics from HSPF and gradient-based optimization to provide interpretable solutions to pollutant transport. With this differentiable HSPF, key sediment-related parameters can now vary flexibly with landscape and climatic conditions, allowing for a more accurate and spatially explicit representation of sediment dynamics. This framework holds the potential to enhance the overall robustness of sediment load estimations in complex watersheds.

Abigail Percich (Virginia Tech), Admin Husic, Allen Gellis, James F. Fox

Watershed-scale sediment source prediction using machine learning

Identifying the dominant sources of sediment erosion in a landscape is essential for effective management, yet existing approaches remain resource-intensive and geographically fragmented. Here, we compile a global dataset of 267 sediment tracing applications and develop an explainable machine learning model that predicts sediment source contributions at a watershed-scale using remotely-sensed watershed characteristics. Despite methodological diversity across applications, four sediment source categories emerge as globally consistent—subsurface, cultivated, non-cultivated, and infrastructure—and our model predicts their contributions with high accuracy ($R^2 = 0.45\text{--}0.52$). Model interpretation shows expected relationships between correlated watershed characteristics and response variables (ie. road density and infrastructure-derived sediment) suggesting that the model captures aspects of physical processes governing sediment sourcing, building credibility in underlying mechanisms. We apply the model to two major watersheds in the United States—the Chesapeake Bay and Upper Mississippi River—and generate source contribution maps that demonstrate its utility for watershed management and reveal spatial differences in sediment sourcing within and across basins. For example, subsurface contributions exceeded 50% in 75% of sub-basins in the US watersheds, with average contributions of 61.8% and 57.2% in the Chesapeake Bay and Upper Mississippi, respectively. Average sediment contributions across the Chesapeake watershed for cultivated, non-cultivated, and infrastructure sources are 19.9%, 11%, and 7.4%, respectively. This synthesis and predictive framework provide the first transferable tool for rapid identification of sediment sources at a large-scale, enabling more efficient watershed management.

Diver Marin Palacio (Occoquan Watershed Monitoring Laboratory, Department of Civil and Environmental Engineering, Virginia Tech, Manassas, Virginia 20110), Chuqiang Chen, Stanley Grant, Admin Husic, Sujay Kaushal

Capturing Event-Driven Salinity Pulses and Nonlinear SC Dynamics in Chesapeake Bay Tributaries using a Deep Learning Model

Inland freshwater salinization is an increasingly pressing challenge, threatening aquatic ecosystems and drinking water resources across the Chesapeake Bay watershed. Most approaches to understand and predict salinization rely on low frequency monitoring data and statistical models to describe the dynamics of specific conductance (SC), a key indicator of freshwater salinity. However, these models may have difficulties capturing nonlinear and event-driven processes governing SC patterns. Recently, deep learning techniques, particularly Long Short-Term memory (LSTM) networks, have emerged as powerful tools for learning temporal dependencies from hydrologic and environmental time series.

Here, we evaluate the applicability of a continental-scale LSTM model, originally developed to simulate daily SC derived from high-frequency aquatic sensors across the contiguous United States (CONUS), within the Chesapeake Bay watershed (CBW). The model uses hydrometeorological forcings, streamflow dynamics and anthropogenic variables to capture complex controls on salinization patterns. We extracted model predictions for 16 monitoring

sites located throughout the CBW and compared predicted daily SC time series against observed records. Results show that the continental LSTM model successfully reproduces SC dynamics across Chesapeake tributaries, capturing season dilution patterns, low flow concentration effects and salinity pulses associated with winter deicing practices. Model performance within the watershed is consistent with continental skill, demonstrating transferability across diverse hydroclimatic and land-use settings. By providing continuous daily SC estimates at monitored locations, continental deep learning models can complement regional monitoring efforts and enhance watershed-scale assessments. These tools offer valuable support for Chesapeake Bay water quality science and management, improving the ability to track salinity dynamics, evaluate anthropogenic stressors, and anticipate emerging risk under a changing environment.

Lindsey Boyle (U.S. Geological Survey), Kelly Maloney, Rosemary Fanelli

Watershed wide predictions of specific conductance show increasing salinity across half of the Chesapeake Bay watershed

Rising salinity in freshwater streams is a concern for both drinking water and stream health. Within the Chesapeake Bay watershed increasing trends in stream salinity have been documented at many individual monitoring sites, but detailed spatial patterns remain unclear due to sparse monitoring and inconsistent temporal coverage. To fill this gap we utilized machine learning to predict stream salinity, represented by specific conductance (SC), across the watershed. Using the National Hydrography Dataset Plus High-Resolution framework, we compiled annual land-use and land-cover data, monthly climate summaries, lithologic properties, and human-influenced point stressors as predictors. Using over 200,000 SC samples, we calculated yearly seasonal median SC for 2,094 sites, requiring that each site have at least one sample in every month of a given year-season to qualify for inclusion. We developed random forest models to predict seasonal median SC for 348,692 nontidal catchments from 1985 to 2023. We then estimated long-term seasonal trends for each catchment and compared predicted trajectories with observed trends where available. Across all four seasonal models, lithologic calcium and sodium consistently ranked among the top predictors, highlighting the influence of natural sources of SC. Impervious surface was the most important predictor for the spring model, but decreased in relative importance in summer and fall, likely capturing the effect of winter deicer application on stream SC. Around half of all catchments exhibited increasing predicted SC over the 39-year period, and winter exhibited the highest proportion of increasing trends (65 percent of catchments). Forty to 60 percent of catchments had predicted SC levels two or more times higher than expected background (depending on season) when compared to a modeled background SC dataset. These results reveal widespread and persistent salinization across the watershed and demonstrate the value of long-term, spatially comprehensive modeling for identifying vulnerable areas and informing management strategies.

Marina Metes (US Geological Survey)

Predicting Aquatic Physical Habitat Over a 38-Year Period Using Machine Learning

Degraded physical habitat is a common stressor affecting river ecosystems and remains a primary focus of management activities, including stream restoration efforts. To evaluate regional conditions and help prioritize management efforts, there is an ongoing need for assessments of instream physical habitat conditions along continuous spatial and temporal gradients.

We combined over 16,000 unique habitat assessments based on EPA Rapid Habitat Protocols from multiple jurisdictions across the Chesapeake Bay watershed with more than 40 predictive landscape metrics. Using this data, we trained a random forest machine learning model to generate spatially continuous habitat assessment scores on a scale from 0 (poor condition) to 20 (optimal condition) for all nontidal streams in the Chesapeake Bay watershed. Two time-varying metrics, annual precipitation and annual land cover data, were used to generate annual predictions from 1985 to 2023, summarized for each catchment/flowline in the National Hydrography Dataset High Resolution 1:24,000 map scale. We found that currently available landscape metrics lack detailed information on the geomorphic conditions of streams that heavily influence aquatic habitat. Therefore, we are also developing and incorporating new landscape metrics at fine spatial resolutions to include channel shape, stream incision, and valley confinement as additional predictive landscape metrics.

These predictions of habitat assessment scores, and examination of changes over the 38-year period, can be used to help target and prioritize management efforts, as well as be used in conjunction with datasets on other instream assessments (e.g. salinity, water temperature, nutrients, streambank erosion) and living resources (e.g. macroinvertebrates and fish communities) to facilitate understanding of the potential suite of local stressors that may be limiting ecological conditions. This information can then inform decision-making when prioritizing sites and/or stressors to mitigate.

Chuqiang Chen (Virginia Tech), Admin Husic

Increasing event water fraction across the Chesapeake Bay Watershed under climatic and anthropogenic change

Streamflow consists of both event flow and pre-event flow. Accurately quantifying the contribution of event water is critical for understanding hydrologic pathways and the transport of nutrients, sediments, and pollutants, hence advancing in restoring Chesapeake Bay water quality. However, such quantification typically relies on high-resolution tracer data, which are available only for select watersheds in the conterminous United States (CONUS). In this study, we applied specific conductance-based hydrograph separation to 26,357 storm events across 277 sites to determine event water fractions at hourly scales. Two deep learning models were trained at a CONUS scale to predict event flow and total streamflow using catchment attributes,

meteorological drivers, and annual land use data. The event flow model achieved a median Nash–Sutcliffe Efficiency (NSE) of 0.90 under spatial cross-validation. With the models, we reconstructed historical event flow and streamflow from 1986 to 2024 for a broader set of 617 CONUS gages, with 92 located in Chesapeake Bay Watershed (CBW). The median annual mean event water fraction (EWF) across CBW basins is 0.25. Trend analysis revealed that 30% of the CBW sites exhibited a significant ($p < 0.05$) increasing trend in annual mean EWF with a median change rate of 0.029 dec⁻¹. Among these, 99% showed increased urbanization, 79% decreased forested land, 47% increased precipitation, and 93% increased temperatures—indicating that both anthropogenic and climatic factors are contributing to rising EWF. Further CONUS-scale analysis suggests that under urbanization and deforestation, regions with greater warming or more intense precipitation tend to show steeper increases in EWF. This is likely driven by the expansion of impervious surfaces, which enhances surface runoff, and by warming-induced precipitation intensification, which promotes infiltration-excess runoff and reduces antecedent soil moisture—ultimately diminishing groundwater recharge and increasing storm-driven runoff.

Lorena Pinheiro-Silva (University of Maryland Center for Environmental Science), Xiaoxu Guo, Matthew Houser, Greg M. Silsbe

Tracking Nutrient Pollution and Best Management Practice Effectiveness in the Choptank River Using Explainable Machine Learning and Satellite Data

Despite decades of restoration efforts, the Chesapeake Bay fell short of EPA-mandated nitrogen and phosphorus reduction targets by 2025, underscoring the need for improved strategies to address non-point sources of nutrient pollution. High-resolution water quality (WQ) measurements are critical for evaluating the effectiveness of Best Management Practices (BMPs) and informing management decisions. However, quantifying BMP impacts on downstream WQ remains challenging as these practices vary in type and scale, and interact with biogeochemical processes across watersheds. We use machine learning (ML) and publicly available data to estimate nutrient pollution and water clarity throughout Choptank River, the major tributary of Chesapeake Bay from remotely-sensed data. This work builds on previous research from our group demonstrating that Sentinel-3 OLCI data can predict total nitrogen and phosphorus at the estuary scale with a low mean absolute error (MAE). Here, we extend this approach by testing whether model performance improves at smaller spatial scale and assessing how the addition of hydrometeorological and land cover data influences predictions. WQ data were obtained from the Chesapeake Bay Program and ShoreRivers and co-located with satellite imagery after atmospheric correction using best-performing algorithms; and land cover, precipitation and discharge data were obtained from the Chesapeake Conservancy, DayMet and USGS gauges. This presentation describes a Python-based mixture density network for robust and explainable retrieval of total nitrogen, total phosphorus, and water transparency. Performance is evaluated using MAE and bias, and the best model is interpreted with Shapley Additive Explanations and an attention mechanism to identify pollution hotspots and seasonal and spatial variability in nutrient loads. Overall, our results highlight how

combining satellite remote sensing with explainable ML can provide scalable and interpretable WQ monitoring that supports stakeholders and policy makers, complements traditional field sampling, and strengthens future watershed research, particularly in times of reduced funding for environmental monitoring programs.

Nivedita Priyadarshini Kamaraj (Department of Geosciences, Virginia Tech, Blacksburg, VA, USA), Sundarabalan V. Balasubramanian, Manoochehr Shirzaei, Susanna Werth

Multi-Sensor Nutrient Mapping in the Chesapeake Bay

Nutrient enrichment drives eutrophication in the Chesapeake Bay, yet spatial monitoring of total nitrogen (TN) and total phosphorus (TP) remains limited. While in situ datasets provide accurate measurements, they are confined to discrete stations and do not capture continuous, Bay wide spatial variability. The overarching goal of this study is to develop and evaluate a scalable satellite-based framework for generating spatially explicit nutrient maps that can complement existing monitoring programs.

We establish a cross-sensor, model comparative approach by integrating optical remote sensing observations from Sentinel-2 Multispectral Imager (MSI) (20 m), Landsat-8 Optical Land Imager (OLI) (30 m), Sentinel-3 Ocean and Land Color Instrument (OLCI) (300 m), and PACE Ocean Color Instrument (OCI) (~1.2 km). Because TN and TP are optically inactive and lack direct spectral absorption features, we implement data-driven models that learn nonlinear relationships between surface reflectance and field-measured nutrient concentrations. Our objective is to (1) develop and compare machine learning models, including Partial Least Squares Regression (PLSR), Support Vector Regression (SVR), Random Forest (RF), and Extreme Gradient Boosting (XGBoost), for remotely predicting nutrient variability, and (2) assess the effectiveness of different remote sensing sensors in capturing Bay wide nutrient variability. Models are trained using satellite–field matchups and validated across different sensors. The best performing models are then selected and applied pixel-wise to generate TN and TP nutrient maps across the Chesapeake Bay at multiple spatial scales. By systematically comparing nutrient patterns from meter- to kilometer-scale observations, this framework assesses the capability of optical sensors to capture nutrient gradients and supports the integration of remote sensing into nutrient reduction tracking and restoration assessment efforts in the Chesapeake Bay.

Breck Sullivan (USGS @ Chesapeake Bay Program), Jon Harcum, Elgin Perry, Rebecca Murphy, Peter Tango

Filling the Gaps: A space-time interpolation tool for Chesapeake Bay dissolved oxygen

Our collaborative team is building a new analysis tool to support evaluation of Chesapeake Bay dissolved oxygen (DO) water quality criteria. To date, the full suite of water quality criteria that

apply to Chesapeake Bay tidal waters cannot be assessed due to insufficient data frequencies and/or analysis techniques to reasonably interpolate DO at time scales shorter than a monthly average. We are developing a new spatial-temporal interpolation tool that combines information from multiple datasets, including high-frequency spatial and temporal data, to develop hourly estimates of DO throughout a requested evaluation period and region. This presentation will provide an overview of the multi-part statistical approach and initial results. The framework consists of using generalized additive models to interpolate mid-day daily DO coupled with within day simulations to capture diel and tidal patterns which together capture the environmental structure of DO. Large- and small-scale correlated error components enable simulation of DO uncertainty that reflects measurement noise and environmental structure. This tool will help answer an assessment need that has existed for over two decades in the Chesapeake Bay and has the potential to be applied in other coastal and estuarine systems with multiple types of water quality data.

Jeremy Testa (UMCES Chesapeake Biological Laboratory), Amir Azarnivand, Damian Brady, Walter Boynton, Lora Harris, Carl Friedrichs, Casey Hodgkins

The diversity of patterns and controls on oxygen depletion in Chesapeake Bay tributaries

Numerous studies have investigated the controls on seasonal hypoxia formation and persistence in many of the world's large estuaries and coastal hypoxic zones (including Chesapeake Bay), but fewer studies have examined patterns and controls on oxygen depletion in smaller and shallower environments. As monitoring and research-based investigations have expanded over time in the Chesapeake Bay ecosystem, the full diversity of oxygen depletion scales and their respective controls has come into focus. This presentation will draw from a range of recent studies to synthesize how oxygen depletion in shallow-water systems responds to both local and regional influences. The synthesis will include (1) insights from a quantitative analysis of 181 stations with dissolved oxygen concentrations collected at shallow water sites (primarily < 2 m) at high-frequency (15-minutes), (2) select case studies of shallow estuaries with multi-decade oxygen time-series, and (3) local, habitat-specific hotspots for oxygen depletion. Although the analyses reveal site-specific external controls on oxygen depletion that may be event-driven, general patterns of oxygen variability responses to riverine flow and temperature emerge that mirror those controls on seasonal hypoxia. Multiple shallow-water locations indicate responses to external nutrient loading and phytoplankton productivity that suggest that meeting TMDL goals will help alleviate diel-cycling hypoxia. These new insights can help inform updated oxygen models to support the modeling and management of shallow-water estuaries in the face of managed nutrient reductions and climate change.

Gabriel Duran (Chesapeake Research Consortium), Jon Harcum, Elgin Perry, Breck Sullivan, Kaylyn S. Gootman, Rebecca Murphy, and Allison Welch

Utilizing Cluster Analysis to Assess Water Quality Trends in the Chesapeake Bay

Effective restoration and protection of the Chesapeake Bay depend on a strong understanding of local and regional water quality conditions. To support this need, the Chesapeake Bay Program (CBP), in collaboration with state partners, has conducted coordinated water quality monitoring of more than 130 stations throughout the mainstem and tidal portions of the Bay since the mid-1980s. Long- and short-term water quality trends are evaluated using a Generalized Additive Modeling (GAM) framework known as baytrends. This modeling approach accounts for variability driven by factors such as meteorological conditions, implementation of Best Management Practices (BMPs), and other external influences. While baytrends effectively characterizes temporal patterns at individual stations, it does not explicitly identify locations that share similar patterns. To better understand spatial coherence in trends, CBP partners developed a clustering methodology that groups monitoring stations across different timescales according to similarities in log-transformed or mean-adjusted water quality profiles. This analytical step enables assessment of whether management actions, including BMP implementation, are producing regionally consistent responses, and distinguishes between areas demonstrating measurable improvement from those where degrading trends persist and may warrant additional intervention. This presentation will summarize the clustering framework, highlight patterns identified in select tributaries, and describe how these results inform CBP reporting and inform management.

David Parrish (Virginia Institute of Marine Science), Carl Friedrichs, William Reay, and Erin Shields

Recent Shifts in Water Clarity Across Virginia's Lower Chesapeake Tributaries: Evidence from Four Decades of Kd Observations

Water clarity is a primary control on submerged aquatic vegetation habitat and a key indicator of broader estuarine ecosystem condition in the Chesapeake Bay. In Virginia's lower tributaries, including the mesohaline and polyhaline regions of the James, York, and Rappahannock Rivers, multiple long-term monitoring programs have collected optical and water-quality observations over several decades extending back to the mid-1980s. Here we synthesize diffuse attenuation coefficient (Kd) records assembled from monitoring conducted by DEQ, HRSD, ODU, and multiple groups at VIMS (Analytical Services, Biological Sciences, CBNERR-VA), largely in close partnership with the EPA Chesapeake Bay Program.

Our analysis focuses on characterizing long-term change in water clarity while evaluating evidence for recent accelerated improvements with a specific emphasis on observations collected in shallow water areas (depths $\leq \sim 2$ m). We harmonize data from multiple monitoring program sources and standardize temporal alignment across programs. Trends are evaluated using statistical time-series approaches that separate seasonal structure from long-term trends and uncertainty. To support inference about recent acceleration, we compare models that allow gradual change over time with alternatives that permit shifts in the rate of change during recent years. This work provides an updated, tributary-scale assessment of water clarity trajectories

using a flexible framework for interpreting multi-agency monitoring records in the context of long-term change and emerging improvements in the estuarine light conditions.

Peichen Huang (Virginia Institute of Marine Science, William & Mary, Gloucester Point, VA, United States), Dante M.L. Horemans, Marjorie A.M. Friedrichs

The Importance of Mixotrophy for Phytoplankton Production and Nutrient Management

Mixotrophy, which is the ability of phytoplankton to supplement inorganic nutrient uptake with organic nutrient sources, is a potentially important process affecting phytoplankton production and corresponding nutrient cycling in estuarine environments. Understanding the balance between growth supported by inorganic versus organic nutrients is crucial to improving the representation of phytoplankton dynamics in water quality models, such as those used by the Chesapeake Bay Program to assess the impacts of nutrient management scenarios. In this study, we integrate multi-decadal taxonomic in situ cell count observations from the Chesapeake Bay with machine learning and numerical modeling. Taxonomy data will be paired with environmental conditions derived from a 3D coupled hydrodynamic–biogeochemical model. These paired data will be analyzed to identify the spatiotemporal distribution of mixotrophs in the Bay and machine learning will be applied to identify environmental conditions associated with mixotroph occurrence. These insights will be incorporated into a recently developed mechanistic mixotrophic growth model implemented within the 3D model. This coupled framework will be used to quantify the potential impact of mixotrophy on hypoxia with and without nutrient reductions. Incorporating mixotrophy into biogeochemical models enables phytoplankton growth under conditions of low inorganic nutrient availability, potentially leading to increased hypoxia. We hypothesize that neglecting phytoplankton growth using organic nutrients in water quality models leads to an increase of model sensitivity to inorganic nutrient availability. As a result, models excluding mixotrophy may overestimate the effectiveness of nutrient reduction scenarios in mitigating future hypoxia. This research highlights the need for modelers to account for mixotrophy to better capture sensitivity of phytoplankton growth and hypoxia to inorganic nutrients.