Event-scale predictions of water and nitrogen exports in coastal watersheds

Authors: Bhavna Arora (LBNL), Chaopeng Shen (PSU), Margaret Zimmer (UCSC), Gemma Anderson (LLNL), and Dwivedi Dipankar (LBNL)

Focal Area(s):

The purpose of this white paper is to call for the improvement of the predictability and explainability of the magnitude and timing of water and nutrient exports in coastal watersheds. In this regard, hybrid machine learning approaches and process-based numerical models provide new opportunities to maximize predictability, emulate process-based model components, and explicitly address model uncertainty associated with event-scale predictions in coastal watersheds.

Rationale:

Coastal watersheds provide numerous benefits and services that are essential for climate change adaptation, including protection from storms and sea level rise, prevention of shoreline erosion, increased biodiversity, and food security for many coastal communities[1]. Additionally, they modulate the global cycles of elements, including climatically significant carbon, nitrogen, and sulfur, and form a critical part of the water quality of coastal ecosystems. Despite these benefits, threats to coastal watersheds from sea level rise and storm surge have increased globally in the 21st century, and losses of these valuable ecosystems are expected to become more frequent, extreme, and spatially extensive in the future.

Currently, one of the biggest impediments to developing robust water quality knowledge in coastal watersheds is an incomplete understanding of event-scale dynamics and lack of sampling at timescales over which nutrients are removed, retained, and transported. Historical event-scale observations alone are inadequate to predict how future extreme events such as increased storm surges, saltwater intrusion, and changes in temperatures and precipitation regimes, will affect water quality and global biogeochemical cycles. This is because future event patterns may be unique, and their impacts more extreme and/or spatially extensive than they have been historically. Undoubtedly, the complexity and diversity of coastal watersheds – from freshwater to saltwater wetlands, from tidally managed systems to natural bogs, and across land-ocean connectivity of 1 to 10s of kilometers – play a pivotal role in nutrient retention and water quality management within these systems. Given this background, it becomes urgent to improve and develop predictive models of water and nutrient exports with a wider spatial coverage and higher temporal fidelity.

Below, we document some key needs in these topic areas. Then, we highlight where artificial intelligence and machine learning (AI/ML) approaches provide the most benefit to coastal watershed research. With this white paper, we aim to impress upon our readers that hybrid-ML and process-based approaches will significantly enhance our ability to predict the impact of future events within an acceptable range of uncertainty and across a diversity of coastal environments.

Knowledge Gaps in Coastal Watersheds

Specific knowledge gaps of water and nitrogen exports in coastal watersheds include prediction in ungauged catchments, identification of drivers and mechanisms of nutrient retention, impacts of extreme events, and upscaling strategies.

DOI: https://doi.org/10.2172/1769706
DOE BER, AI4ESP 2021 #AI4ESP1005 (https://www.ai4esp.org)

Event-scale predictions of water and nitrogen exports in coastal watersheds

Prediction in ungauged catchments: Our predictive capability and understanding of timing and magnitude of nitrogen loading and flow are best in basins with gaging stations or periodic sampling. Spatial and temporal data gaps limit our ability to fully quantify material export to coasts. There is a need to extend our process understanding from data-rich basins to basins with limited data (e.g ungauged basins, grab samples, periodic sampling).

Drivers and mechanisms of nutrient cycling: As material transport and transformations are highly variable in time and space, there is a need for improved understanding of the spatial and temporal nitrogen sources and sinks as well as what triggers delivery to and through rivers. Such an understanding would help reduce uncertainty in predicting these transformations under an uncertain future climate.

Impacts of Extreme events: The accelerating water cycle due to climate change coupled with land use/cover changes has made nutrient and water export to coasts difficult to predict. There is an urgent need for quantification of how sea level rise and extreme hydrologic events couple to impact nitrogen retention and processing (e.g., chronic versus episodic export regimes).

Upscaling strategies: Most historical biogeochemical and hydrologic datasets are point-scale measurements from gaging stations. While these data integrate upstream and watershed-scale processes and sources, they restrict our ability to identify sources and track changes. There is thus a need to upscale local data (e.g., grab data, further inland datasets) to tease out important watershed processes, drivers, and sources of material export to coasts.

Science Questions

With this paper, we want to bring forward the following questions:

1. How do coastal watersheds respond to extreme water cycle changes, both in terms of their structure (e.g. shift inland with rising sea level) and processes (e.g. dominant processes shift due to changes in oxygen gradients with accelerating water cycle)? and
2. How do we ensure accuracy and reliability of predictions in ungauged watersheds?

The focus here is on developing ML approaches that are transferable and interpretable.

Proposed Research Directions

Hybrid machine-learning and process-based models have the potential to overcome many of the limitations that have, until now, hindered a more widespread adoption of machine learning to hydro-biogeochemical datasets in complex environments. The hybrid approaches have the ability to implement a multi-scale strategy fusing process-resolving simulations and including machine learning, data mining, genetic algorithms, and other techniques enabling process fidelity with computational tractability. In other words, the hybrid approaches take advantage of process-based and machine learning approaches, using them to complement and enrich one another. Below, we provide specific examples of how and where hybrid approaches can enhance our ability to address ungauged watersheds, similarity and functioning of diverse coastal environments, hydro-biogeochemical scaling, as well as impacts of extreme events.

Significance of hybrid approaches in data-poor environments:
ML-enabled spatial generalization methods can help bridge the gap between intensively sampled sites and infrequently-sampled sites. For example, we can train models in data-intensive regions and migrate it to infrequently-sampled regions using transfer learning to leverage limited local observations, e.g., *Ma et al.*[2]. Additionally, predictions using traditional models could be "elevated" using transfer learning

approaches. The traditional model could be replaced by an ML surrogate model that can subsequently "learn" from observations. Initial training of the surrogate is performed in the data-rich model world and then "retrained" on limited observational data, e.g., *Kustowski et al.*[3]

Advancing understanding of drivers and mechanisms:
One limitation of earth system models is that their predictions are strongly dependent on model parameters, which may or may not be obtainable at relevant spatial and temporal scales. Traditional optimization methods require thousands of computationally-expensive model runs to calibrate these parameters. In this context, ML methods offer the advantage of being highly efficient at extracting information and learning parameterizations, e.g., *Tsai et al.* [4]

Predictions of "Extreme" water cycles: We expect ML methods to offer specific advantages to the study of extreme water cycle and weather events. First, hybrid approaches will be able to generate hypotheses regarding generalizable trends of water and nutrient exports across a range of coastal watershed settings. Second, ML approaches offer the opportunity to exploit and be exploited. This implies that these approaches will be able to identify regions where data-driven approaches alone may not be sufficient to capture responses to event-scale dynamics. Third and last, the development of transferable and interpretable ML approaches will enable predictions of water and nutrient exports across event, seasonal, and decadal scales.

Upscaling strategies: Deep learning downscaling techniques can be used to assess the impact of future climate change on water and nutrient exports. This would involve leveraging future projections of sea level rise and other relevant variables from global climate models at large spatial scales and learning relationships between these variables and water and nutrient exports at local scale using historical observations to obtain predictions for a variety of future climate scenarios.

Future (10-year) Vision

The quality of coastal watersheds is one of the most important concerns for the future, as recognized by the U.S. Departments of Energy and Defense, among other organizations. Coastal watersheds are complex systems with steep redox gradients and spatially heterogeneous composition, making these systems both valuable for the global biogeochemical cycles, but also difficult to model and predict. Ultimately, this paper calls for the: (1) development of hybrid ML-reactive transport modeling approaches to predict water and N exports in coastal environments; and (2) use of these hybrid approaches to reduce predictive uncertainty regarding changes in water and N exports due to sea level rise and other extreme events across a diversity of settings.

References

1. Gittman, Rachel K., et al. "Living shorelines can enhance the nursery role of threatened estuarine habitats." *Ecological Applications* 26.1 (2016): 249-263.
2. Ma, Kai, et al. "Transferring hydrologic data across continents--leveraging US data to improve hydrologic prediction in other countries." (2020).
3. Kustowski, Bogdan, et al. "Transfer learning as a tool for reducing simulation bias: application to inertial confinement fusion." *IEEE Transactions on Plasma Science* 48.1 (2019): 46-53.
4. Tsai, Wen-Ping, et al. "From parameter calibration to parameter learning: Revolutionizing large-scale geoscientific modeling with big data." *arXiv preprint* arXiv:2007.15751 (2020).