Open-source AI-ready data for prediction of coastal water and carbon budgets under a changing climate

Authors:
Bhavna Arora (LBNL), Evan Goldstein (UNC Greensboro), John Gardner (Univ. Pittsburgh), Margaret Zimmer (UC Santa Cruz) and Haruko Wainwright (LBNL)

Focal Area(s):
Insights on coastal water and carbon budgets gleaned from complex data (both observed and simulated) using AI

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

a) What is the role of different coastal environments (estuaries, tidal marshes, etc.) in global water and carbon cycles? And, how will this role change under future climate/extreme events?
b) How well is the spatial and temporal heterogeneity in carbon pools and fluxes as well as their multiscale interactions in coastal environments represented in Earth System Models (ESM)?

Rationale:
Coastal ecosystems are a significant carbon sink and control the fate of terrestrial carbon. Climate change is expected to impact the coastal system and its role in terrestrial carbon in several ways —through extensive ecosystem loss as sea levels rise, from impacts to terrestrial inputs through extreme storm events, and through changes in sinks and stores of carbon pools. Further, each of these C pools may be impacted by climate change from landward and seaward directions. For example, extreme precipitation events may increase riverine carbon inputs to the coasts, but sea level rise may alter how that carbon accumulates, is processed at the coastal terrestrial aquatic interface (TAI), or transported to the ocean. In addition, limited data describing the heterogeneous inputs of carbon from rivers to the coast and the multiscale interactions that govern these exports severely restricts our ability to forecast how carbon budgets may change under future climate.

Predictions of the fate and transport of carbon suffer from a lack of adequate data and process understanding of coastal water and carbon budgets. In particular, there are very few depth- and time-resolved measurements of carbon storage and dynamics in the watershed-river-coast continuum, hindering modeling efforts that try to capture the true nature and variability of hydro-biogeochemical response as impacted by climate change (e.g., sea level rise, increased frequency of flooding events, king tides, swell events). Not only are these models poorly constrained, regional and global earth system models only consider one-way feedback from the river to the coast without considering appropriate representations of the multi-directional interactions in these systems[1]. Additionally, most coastal watersheds in the world are poorly gauged, unless they are intensively used or part of a busy navigation network[2]. This implies that we hardly know the length or characteristics of the different coastal environments or the watershed-river-coast continuum through which most material and water travel to reach the coast.

Our Vision
The use of site-calibrated models for predicting carbon fate and transport in coastal environments remains relatively rare, despite advances in methodology, computing technology, and the notable advantages in terms of predictive accuracy. A potential alternative is to link models and datasets using AI/ML approaches. By developing ML techniques that emulate or build upon high-fidelity site-specific models, we may be able to understand how a coastal environment would function under both natural and climate change conditions. Further, ML approaches can help identify key processes/parameters of coastal carbon fate and transport that can then be used to inform regional or global models. Recent advances in AI/ML approaches, including in the limited data space, offer exciting new opportunities for developing data-

DOI: https://doi.org/10.2172/1769758
DOE BER, AI4ESP 2021 #AI4ESP1004 (https://www.ai4esp.org)

driven generalizable models, generating new hypotheses, representing sub grid processes, and improving forecasts.

Existing Gaps in Current Research

Specific knowledge gaps in quantifying coastal water and carbon budgets under current and future climate are described below:

Multi-year prediction and predictability: Developing robust models of carbon transport applicable to a diverse set of coastal environments requires data on C pools and fluxes from a range of settings and across scales. However, data capturing the evolution of C hydro-biogeochemistry under non-stationarity conditions and extreme events (e.g., king tides, flooding events, drought) are fairly limited. This lack of data limits accurate prediction of carbon export to coasts under future extreme events. While integrated predictive assessments using ESM type models may prove beneficial, these models are inadequate in capturing the entirety of input information on heterogeneous carbon pools and fluxes from terrestrial landscapes and dynamic coastal interactions (e.g., tides, sea level rise) due to their limited representation[1].

Interfaces and control points: Today, there is much greater appreciation of TAIs as biogeochemical and ecological control points (Bernhardt et al., 2017). Coastal TAIs are diverse, ecologically complex and represent a turbulent fluid environment that sits between freshwater and marine systems. However, much remains unknown about the role of coastal TAIs in processing and transporting carbon. Without this information, our ability to accurately model and predict carbon behavior and the role of TAIs in the global carbon cycle is incomplete.

Defining the components and extent of coastal environments: While coastal TAIs have high potential for processing and storage of carbon, our understanding of what constitutes a coastal environment and its boundaries is limited. In regional and global models, the shape and extent of coastal TAI environments and the specific components (e.g., estuaries, marshes, barrier islands) may vary significantly depending on the grid resolution. Without an adequate representation of the extent of coastal TAIs, our ability to quantify the link between the terrestrial and ocean carbon cycles and predict how these systems will change in the future is deficient.

Impacts of Extreme events: In a world of extremes - both wetter and drier - carbon inputs from terrestrial landscapes and their fate in coastal waters will change. Extreme flow events will cause water residences times in coastal environments, and potentially TAI extent, to oscillate between high and low conditions, which coupled with increasing temperatures may lead to changes in C burial, transformation, and flux to the ocean. In order to understand how an accelerating water cycle may alter the role of TAI in storing, transforming, or transporting carbon to the global ocean, we must understand event-scale dynamics.

Proposed Research Directions

To move forward, we need a synthesis and integration of available data and model predictions to capture the true nature and variability of carbon fluxes from rivers to the coast as impacted by climate change (e.g., sea level rise, increased frequency of flooding events, king tides and swell events). Here, we propose our community must ask the following key questions:

- Can we generate a list of observational datasets that are available, and also list those that are critical, but not yet used or available?
- Can we use models to fill in the identified data gaps? Which parameters/processes cause the most uncertainty in carbon budgets?

Observational opportunities:

Satellites have been used for decades to understand earth systems processes and generate critical inputs for earth system models. Advances in cloud storage, cloud computing, and machine learning have made these tools widely accessible, which has supercharged the potential of satellite observations to generate global, high spatial resolution data products and new insight developed by both space agencies and unaffiliated government or academic groups.

Current and historical satellite missions alone include various global, multi-temporal data products useful for understanding the global carbon cycle, such as products of terrestrial vegetation, terrestrial gross primary productivity, land cover, surface temperature, evapotranspiration, ocean primary productivity, snow and ice cover, burned land area, concentrations of atmospheric pollutants (e.g. sensors aboard Landsat 4-8, Terra, Aqua, VIIRS/Suomi-NPP/NOAA-20 satellites), precipitation (GPM), and soil moisture (SMAP). Geostationary satellites (e.g. GEOS-R) provide additional high temporal resolution (~minutes) to observe the evolution of weather and hydrologic events over specific regions. Future government and private satellite missions will provide additional earth observation capacity.

AI/ML-based opportunities:

There are several exciting opportunities that AI/ML-based techniques offer for predictions of carbon fate and transport in coastal environments. First, AI/ML methods can tap into imagery from ongoing missions with continuous long-term records (e.g., Landsat), or upcoming missions that provide high temporal resolution of direct observations (e.g., aggregating Cubesats with *in situ* sensors) and/or indirectly derived parameters (e.g., NASA SWOT for surface water elevation). Here, ML approaches can also be used for optimizing *in situ* sensor network placement. Second, we propose the development of ML-based emulations of common models of carbon fate and transport in coastal TAIs. After calibrating physics-based models on monitored systems using observational data, we can develop deep learning emulations using recurrent neural networks (i.e., long short-term memory, gated recurrent units, 1D convolutional models (e.g.,Wavenet)). Such an approach can be beneficial to preserving multi-scale features of C fate and transport and thus applicable to new sites. Third, AI/ML approaches can look for systems with similar carbon fate and transport characteristics (leveraging explainability techniques, e.g. *Jeyakumar et al.*[4]). Finding analogous systems can be used to generate hypotheses: What other places are likely to be similar to my field site? Why? Will they respond similarly to climate change events?

Future 10-year Vision

In the absence of measurements of the carbon cycle at relevant time scales in coastal environments, AI/ML is poised to resolve critical questions that we don't know the answers to – impact of king tides, extreme events, and sea level rise – on global water and carbon budgets. These solutions will yield high-confidence insights and answers to the most critical problems and challenges for the USA in the next decade – i.e., assessing the global carbon budgets in a future uncertain climate. New datasets (e.g. remote sensing, greater optical sensor coverage) will undoubtedly expand our temporal and spatial observations beyond what is possible today. This, coupled with rapid advancements in AI/ML and an increasing focus on FAIR principles, will open new doors to unraveling the ultimate fate of terrestrial carbon inputs to understand the role of coastal TAIs in the global carbon cycle in a world that is increasingly stochastic and extreme.

Acknowledgment

This AI4ESP white paper is inspired from a breakout session organized as part of the Early Career Critical Zone Network-of-networks Workshop during the 2020 AGU Fall Meeting.

References

1. Ward, Nicholas D., et al., Nature communications 11.1 (2020): 1-14.
2. Savenije, Hubert H. Water Resources Research 51.4 (2015): 2464-2476.
3. Bernhardt, Emily S., et al. Ecosystems 20.4 (2017): 665-682.
4. Jeyakumar, Jeya V., et al. Advances in Neural Information Processing Systems 33 (2020).