Upscaling cross-scale flow and respiration interactions at river sediment interface leveraging observation, numerical models, and machine learning

Authors

Yunxiang Chen, Jie Bao, Jianqiu Zheng, Peiyuan Gao, Qizhi He, James Stegen, Pacific Northwest National Laboratory

Brenda Ng, Lawrence Livermore National Laboratory.

Xiaofeng Liu, Roman Dibiase, Chaopeng Shen, Pennsylvania State University

Focal Area(s)

Predictive modeling of the interactions between flow and respiration at the river sediment interface.

Science Challenge

The multiscale interactions among water, microbial respiration, and sediment under realistic riverbed conditions. The transferability of the scientific understanding of interactions at multiple sites. And the upscaling of these interactions from pore to reach scales (see science challenge in Figure 1).

Rationale

The interactions among flow, sediments, and microbes at the river sediment interface (RSI) control the exchange of water, nutrients, and carbon dioxide between rivers and subsurface sediments. These interactions can be conceptually characterized as the dependencies of riverbed exchange fluxes and microbial respiration rate on riverbed topography and flow conditions (science challenge in Figure 1). Though important for predicting water quality and freshwater carbon cycle [1,2], they are poorly understood and rarely accurately represented in large-scale earth system models.

This is mainly due to the lack of efficient approaches for both cross-scale measurements and modeling that can link the interactions occurring and measurable at pore-to-meter scale to the scales (>100 m) resolvable in large-scale models. The on-site measurements, though important for model calibration and validation, are also hard to capture the representative sediments, exchange fluxes, and microbes/chemical species concentrations at the river sediment interface due to their spatiotemporal fluctuations and hard access to the pore regions in field surveys. Besides, accurate models are missing for the prediction of the interactions among turbulent flow, microbial respiration, and sediments between surface water and groundwater even at small scales due to the lack of high-resolution topography data, which makes it hard to empirically represent the interactions at large scale models. Further, the approaches are not clear that can upscale the interactions obtained at pore-to-meter scales to reach and watershed scales, and the transferability is rarely studied for the upscaling approaches to multiple sites due to the aforementioned issues.

Here we propose an integration of machine-learning (ML) enhanced structure-from-motion (SfM) photogrammetry, a fully-coupled surface-subsurface model, a kilometer-scale surface flow model, on-site observations, and machine learning techniques to solve the above science questions. The ML enhanced SfM approach is to acquire mm-cm scale riverbed sediment topography that is critical for predicting the flow-sediment-microbes interactions. The fully-

DOI: https://doi.org/10.2172/1769792
DOE BER, AI4ESP 2021 #AI4ESP1023 (https://www.ai4esp.org)

coupled surface-subsurface model is to directly resolve the riverbed fluxes and concentrations that are affected by sediments, sediments-generated turbulence, and microbial respiration. The reach-scale model is to provide important hydrodynamics conditions, e.g., velocity and water depth, that drive the interactions at RSI. As the aforementioned processes are usually time-consuming (takes around 10 days for one simulation), an integrated convolution neural network (CNN) [3] and long-short-term-memory (LSTM) is proposed to train a machine learning model using the SfM-derived topography and the reach-scale model generated velocity/depth as inputs and the bed fluxes and concentrations data form the fully-coupled model as training datasets. The resulted ML model can thus efficiently (in seconds or minutes) predict riverbed interactions, i.e., fluxes and concentrations, for any new sites along as images and velocity/depth data are available. Also, as the images represent the information of the pore-meter scale sediment structures, and the velocity/depth data are available at km-scales, the interactions predicted by our ML model leveraging images and reach-scale flow data automatically upscale the interactions at the pore-meter scale to the reach scale to even watershed and earth system scales. The predicted fluxes can be further used for reach-scale subsurface reaction transport models such as Pflotran (Figure 1 upper right). Therefore, the proposed framework solves both the transferability and the cross-scale upscaling issues. Furthermore, as the ML is appliable for any sites, the fluxes and concentrations predicted in the model can provide a guide for where and when to do field observations as long a few images of the potential sites are provided.

Narrative

Using the above physical models and ML model, riverbed fluxes and concentrations at various bed sediment size/distribution/complexity, velocity/depth, and rate constants conditions are available. These data can be used to answer what controls the multiscale interactions among turbulent flow, sediments, and reaction transport. It is needed to be understood how riverbed exchange concentrations and hotspots, i.e., subsurface maximum concentrations, are controlled by sediment-generated turbulence, bed surface porosity, soil porosity, and the microbial respiration rate constants. Due to the complex nature of microbial respiration, quantifying the uncertainty in exchange concentrations and hotspots from the reaction rate constant is also necessary.

Though the upscaling and transferability issues can be solved in principle. The images and flow data at various sites should be collected to train the model and validate the ML's performance in predicting fluxes and concentrations under various conditions. This can be solved using two ways: (a) collecting images by experts and non-scientists using various tools such as UAVs, cameras, 5G stations install near rivers, cellphones, and existing data-sharing network such as the DOE's Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems (WHONDRS). With these images, an ML model that can detect features from 2D images should be developed to remove images that are very different from the others. With these selected images, a semi-automated image conversion workflow, i.e., the SfM photogrammetry workflow, will be developed based on commercial software such as PhotoScan or open-source software CloudCompare, MeshLab, VisualSfM. Such technical development fills the gap of high-resolution river topography data and supports the above science question. (b) Regarding the velocity and depth data, a high-resolution 3D hydrodynamics model can be used to predict the flow if 1 m-resolution river bathymetry data are available. 1D/2D hydrodynamics model can also be used if coarse river bathymetry data are the only choice. The coarse (> 1 m) riverbed topography may be available in USGS, NOAA, and other open data sharing repositories.

Upscaling cross-scale flow and respiration interactions at river sediment interface leveraging observation, numerical models, and machine learning

With the required images and flow data, the ML learning training procedure is as follows. For convenience, we take the riverbed vertical exchange flux as an example. The flux is a 2D data structure and has different values at different simulation times. For each time step, such data will be first reduced to a lower-dimension latent space using the CNN-encoder technique. Such new lower-dimensional data with the lower dimensional data at a few old-time steps are used as the input of the LSTM, which will predict a new lower-dimensional data for the next time step. The CNN auto-decoder will then be used to convert this low dimensional data into the original dimension. Repeating the above loops for each time step until the ML model can predict a good accuracy. The accuracy could be evaluated by the standard deviation of the difference between the predicted and the input data as the fluctuation magnitude is a key quantity for evaluating the importance of exchange flux. The comparison of the probability density function of the flux between the predicted and input data are desirable considering the randomness nature of the exchange flux. Other variables such as horizontal exchange fluxes, pressure, and concentrations can be trained similarly.

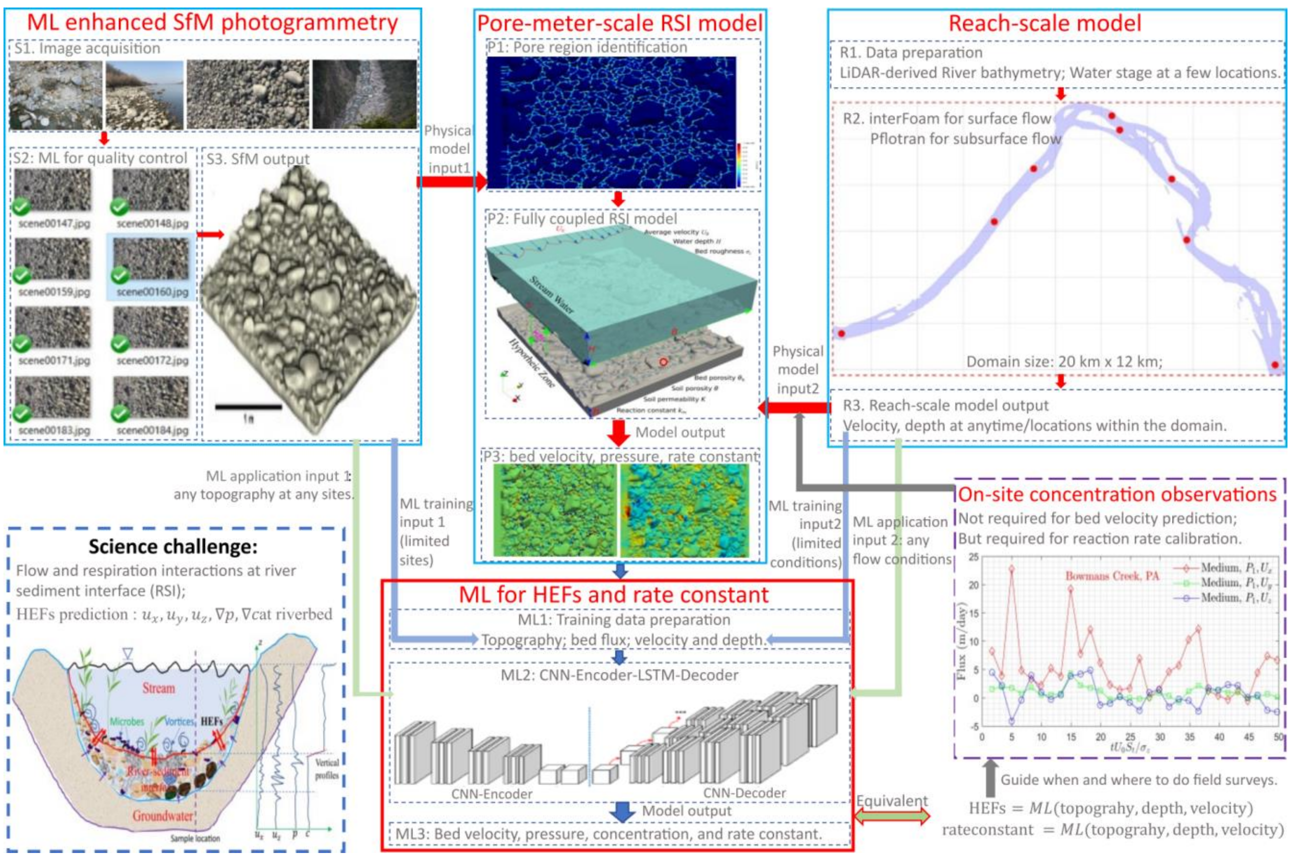

Figure 1: The key science challenge (lower left) and the communications among the ML enhanced SfM photogrammetry (upper left), pore-to-meter scale river sediment interface model (upper middle), reach scale model (upper right), and the resulted ML framework (lower middle) as well as on-site observations (lower right).

References (Optional)

[1] Gomez-Velez, J., et al. Denitrification in the Mississippi River network controlled by flow through river bedforms. Nature Geosci 8, 941–945 (2015).

[2] Biddanda, B. A., Global significance of the changing freshwater carbon cycle, Eos, 98, (2017).

[3] B. Kim, et. al, Deep fluids: A generative network for parameterized fluid simulations, Computer Graphic Forum, Proceeding of Eurographics 2019.