Machine Learning-Based Multi-Scale Multi-Compartment Watershed Characterization across Scales

Haruko Wainwright[1,2], Sebastian Uhlemann[1], Qina Yan[1], Alex Miltenberger[3], Baptiste Dafflon[1], Nicola Falco[1], Burke Minsley[4], Kenneth Williams[1], and Susan Hubbard[1,5]

[1]Lawrence Berkeley National Laboratory, Berkeley, CA;
[2]Massachusetts Institute of Technology, Cambridge, MA;
[3]Stanford University, Stanford, CA;
[4]US Geological Survery, Denver, CO;
[5]Oak Ridge National Laboratory, Oak Ridge, TN

Contact: (suhlemann@lbl.gov)

Project Lead Principal Investigator (PI): Eoin Brodie

BER Program: ESS

Project: Berkeley Lab Watershed Function SFA

Project Website: watershed.lbl.gov

Abstract:

Predictive understanding of watershed functions – such as ecosystem dynamics, water fluxes, nutrient cycling and export – is often hindered by the heterogeneous and multiscale fabric of watersheds. In the Watershed Function SFA, we are developing novel watershed-characterization methodologies to quantify complex watershed systems across scales, while taking advantage of recent advances in airborne remote sensing technologies. Machine learning (ML) plays a significant role in identifying and quantifying the relationships between spatially extensive remote sensing data and sparse measurements, like geophysical or in-situ sampling. ML quantifies the co-variability between above/belowground properties and aids in reducing the multi-dimensional parameter space into a set of *watershed zones* to tractably capture these co-varied properties.

In this presentation, we present several key recent developments and findings:

- The effective integration of multiscale geophysical data (borehole, surface, airborne) through ML has enabled the quantification of correlations between topographic/vegetation metrics (obtained from remote sensing data) and bedrock properties, to estimate bedrock properties across the watershed. In particular, this approach also enabled us to map hydrogeologically important features, such as permeable fracture zones.
- A hybrid model that combines a process-based model and empirical relationships was developed to estimate the spatial heterogeneity of soil thickness at fine spatial resolution across

a hillslope, by combining a process-based soil erosion/diffusion model and empirical relationships between topographic metrics and soil thickness.

- A ML-enhanced probabilistic framework was developed to quantify both the parameter and model uncertainty of geological structures based on geophysical data. This framework was applied to identify a new fault, which is a fast groundwater flow path, and which was later confirmed through drilling and core recovery.

- Watershed zonation was identified and mapped through hillslope clustering of airborne remote sensing data (LiDAR, hyperspectral, and electromagnetic surveys). Hillslopes were identified as an appropriate unit for capturing the key bedrock-through-canopy properties dictated by aspects, elevation and geology. Using independently collected data, we show that the identified zones provide information about zone-based watershed functions, including foresummer drought sensitivity and river nitrogen exports.

Through explicitly bridging information derived from “on the ground" observations made at the East River Watershed and remote sensing data, we quantify fundamental scientific linkages among interacting processes in the watershed.

Miltenberger A, Uhlemann S, Mukerji T, et al (2021) Probabilistic Evaluation of Geoscientific Hypotheses With Geophysical Data: Application to Electrical Resistivity Imaging of a Fractured Bedrock Zone. J Geophys Res Solid Earth 126:1–20. https://doi.org/10.1029/2021JB021767

Uhlemann S, Dafflon B, Wainwright HM, et al (2022) Surface parameters and bedrock properties co-vary across a mountainous watershed: Insights from Machine Learning and Geophysics. Science Advances (accepted).

Wainwright HM, Uhlemann S, Franklin M, et al (2022) Watershed zonation through hillslope clustering for tractably quantifying above- and below-ground watershed heterogeneity and functions. Hydrol. Earth Syst. Sci., 26, 429–444. https://doi.org/10.5194/hess-26-429-2022

Yan Q, Wainwright H, Dafflon B, et al (2021) A hybrid data–model approach to map soil thickness in mountain hillslopes. Earth Surf Dyn 9:1347–1361. https://doi.org/10.5194/esurf-9-1347-2021