Investigating the impacts of streamflow **disturbances on water quality** using a data-driven framework

Delaware River Basin Commission

US Geological Survey

Charuleka Varadharajan

ESS Cyberinfrastructure Meeting

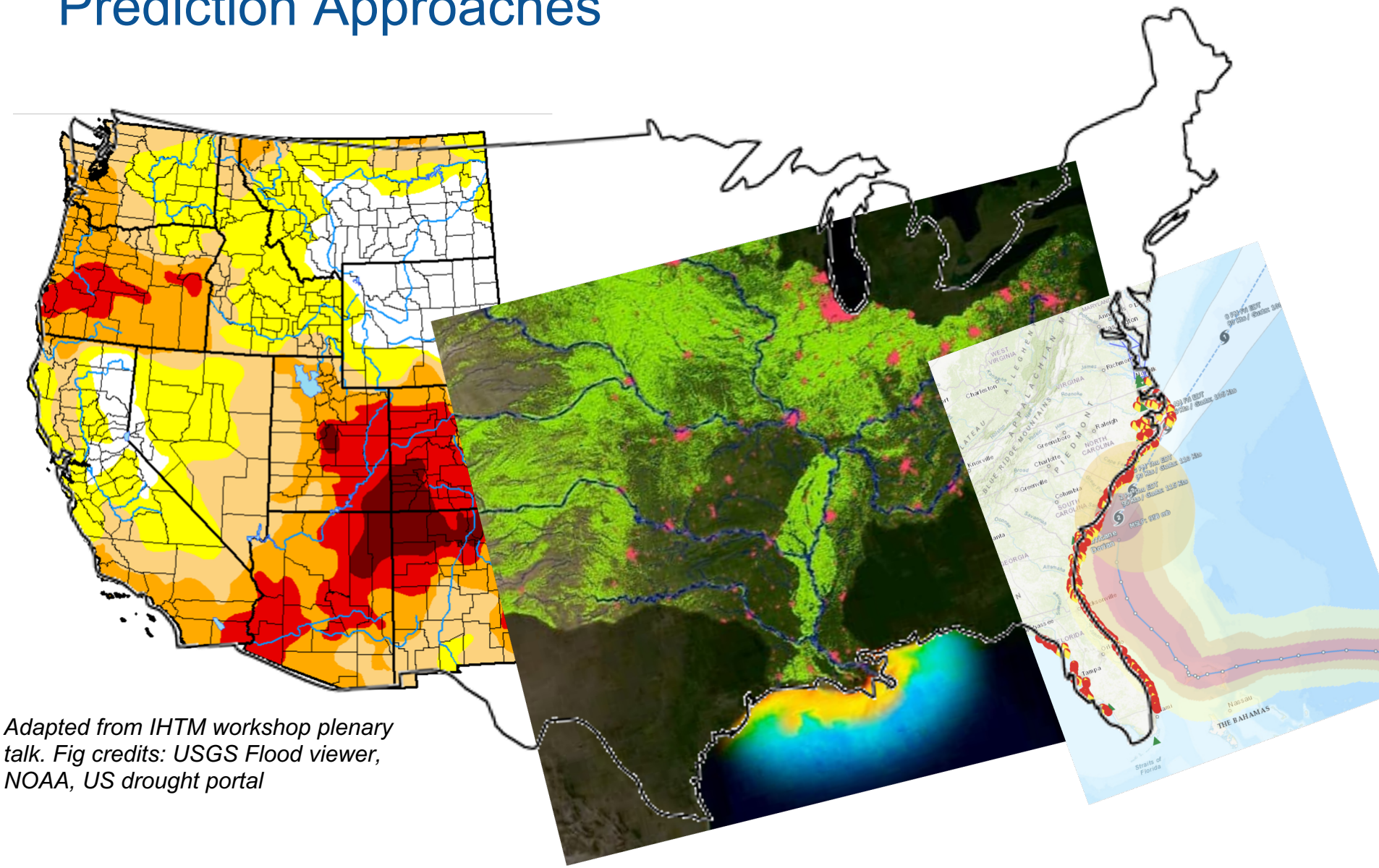
May 11, 2020

U.S. DEPARTMENT OF ENERGY | Office of Science

EARTH & ENVIRONMENTAL SCIENCES

BERKELEY LAB
Lawrence Berkeley National Laboratory

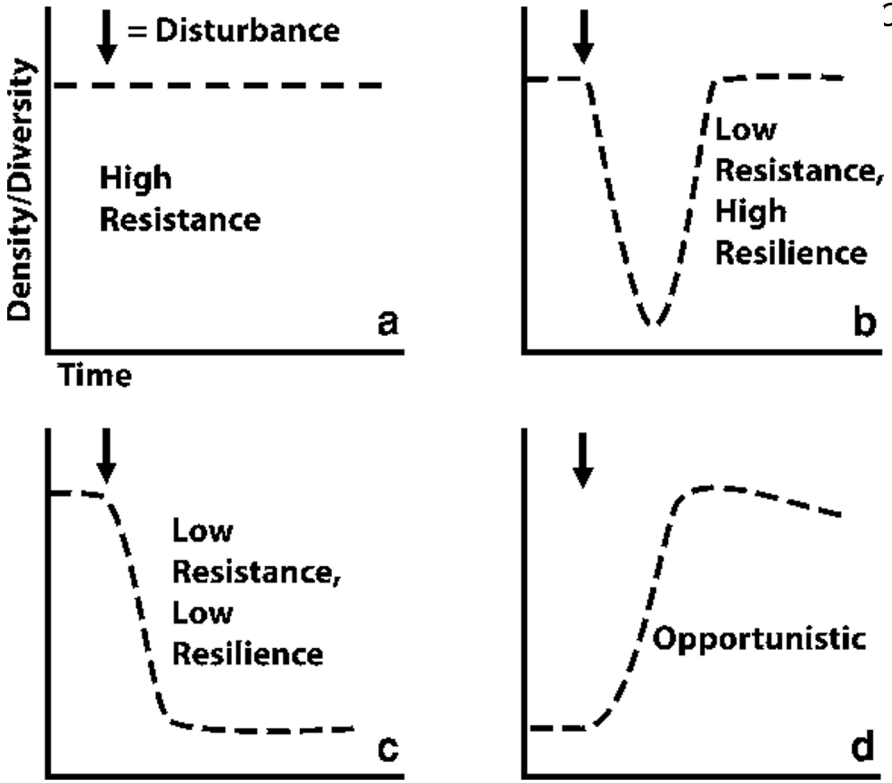
The Nation's Priority Water Challenges Need New Prediction Approaches



Adapted from IHTM workshop plenary talk. Fig credits: USGS Flood viewer, NOAA, US drought portal

Research Objective

To determine watershed emergent properties - ***resistance/resilience -*** to streamflow disturbances, and predict impacts to water quality using a new, open framework for data-driven analyses and modeling



Anderson and Ferrington, Hydrobiologia (2012)

Carey et al., HP (2010, McCluney et al. Fr. Ecology and Env. (2014)

Water Quality: Temperature, Dissolved Oxygen, Electrical Conductivity (Salinity)

Variables of critical importance to aquatic life and water managers with high-resolution datasets

Driving Science Questions

S1: How do the intensity, duration, and frequency of streamflow disturbance events change water quality at a given location in the short and long-term?

S2: How far do disturbance-related water quality changes propagate further downstream?

S3: How do the impacts of streamflow disturbances on water quality vary across watersheds with different characteristics of geomorphology, soil properties, climate, land use, and land cover?

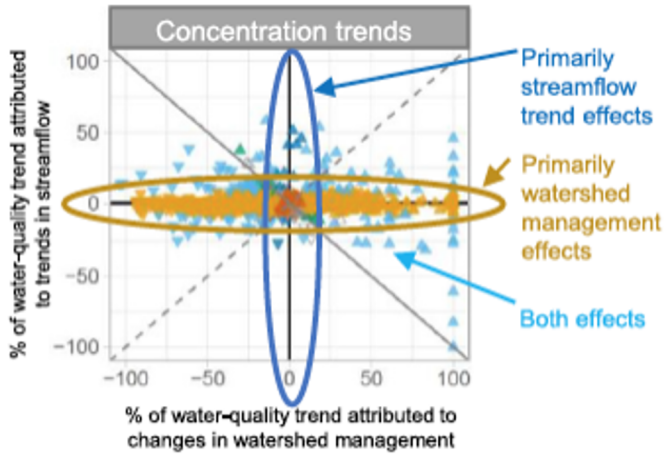
Watershed Resistance/Resilience Index = f (disturbance event, watershed characteristics)

Multiscale Testbeds for 3 River Corridors

Built using USGS NLDI. Yellow and pink dots represent locations of streamflow and water quality observations.

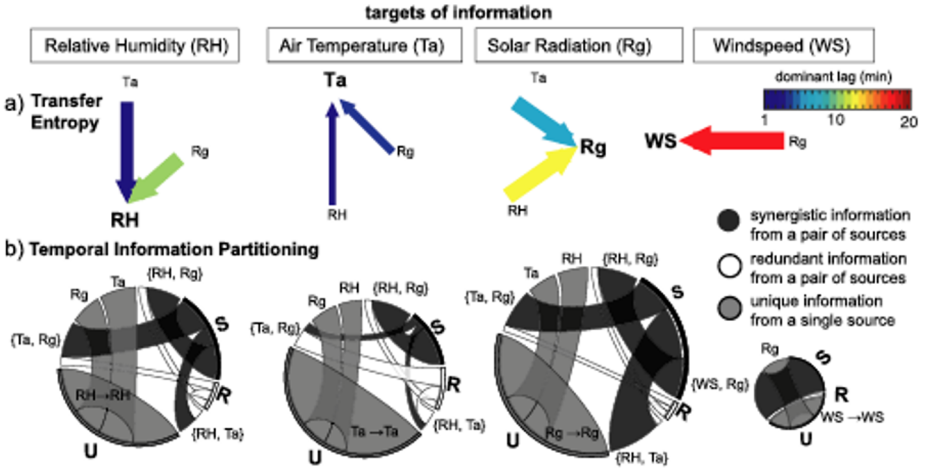
Data-Intensive Approaches Used in Isolation; Opportunity for Better Predictions with Integration

Classical Statistics on Long-Term Monitoring Network Data



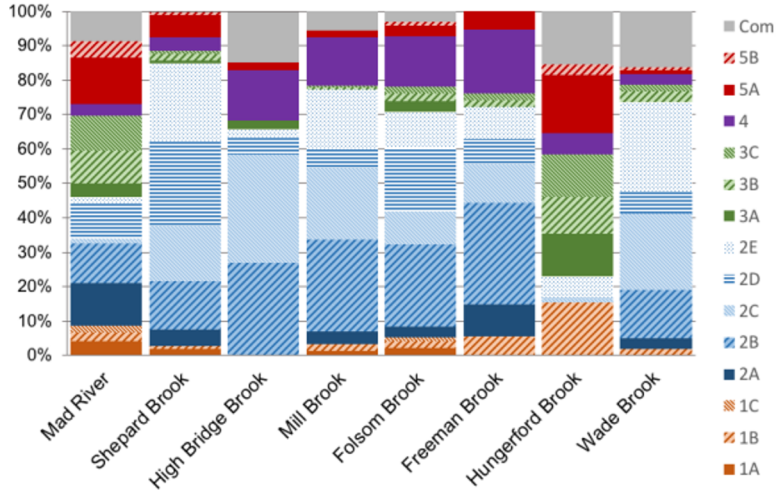
Murphy & Sprague, STOTEN (2019)

Information Theory/Causal Inference



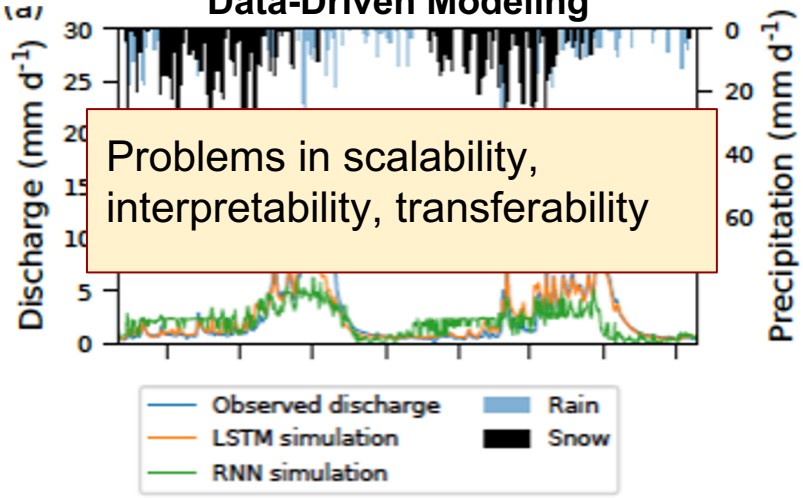
Goodwell Et Al, WRR (2019)

Classification/Pattern Recognition



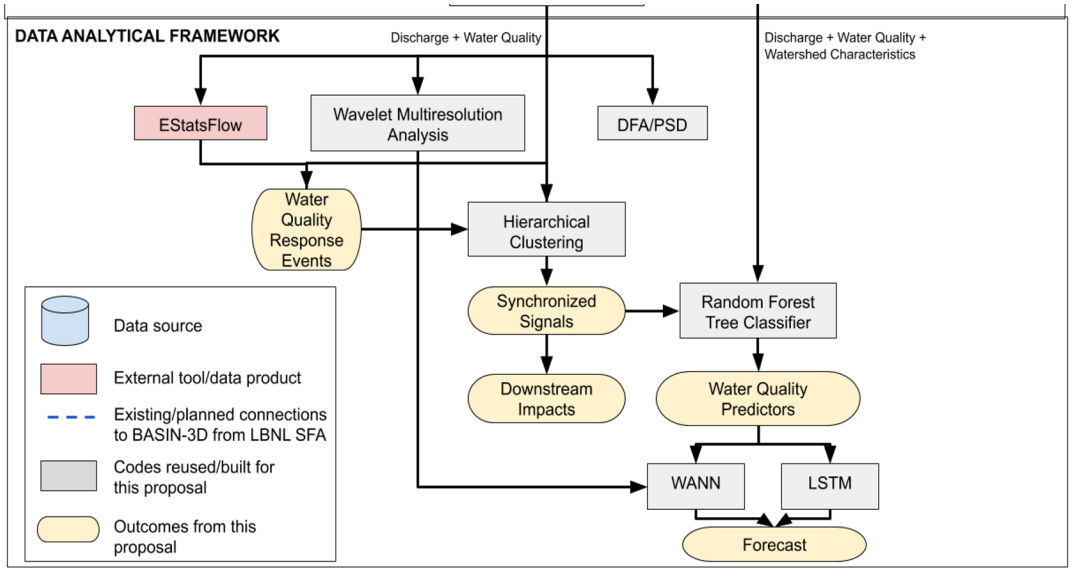
C-Q Storm Event Classification *Hamshaw Et Al. (2019)*

Data-Driven Modeling



Rainfall-runoff modeling using deep neural networks *Kratzert Et Al., HESS (2018)*

My Approach: Integrating data-driven synthesis, analysis and modeling in open framework



Variable(s)	Name of dataset	Source
Discharge	NWIS	USGS
Water Quality	Water Quality Portal, Murphy & Sprague, NWIS	USGS, EPA, USDA, various agencies
Topography, Geomorphology, River network	NHDPlus HR	USGS
Land use/Land Cover	NLCD	USDA
Soil properties	SSURGO/STATSGO2	USDA
Climate	GHCN, PRISM,	NOAA, Oregon State Univ.
Disturbance	LANDFIRE, Hydro Disturbance Index	
Integrated products	StreamCat, CAMELS, GAGES-II	USEPA, UCAR, other datasets

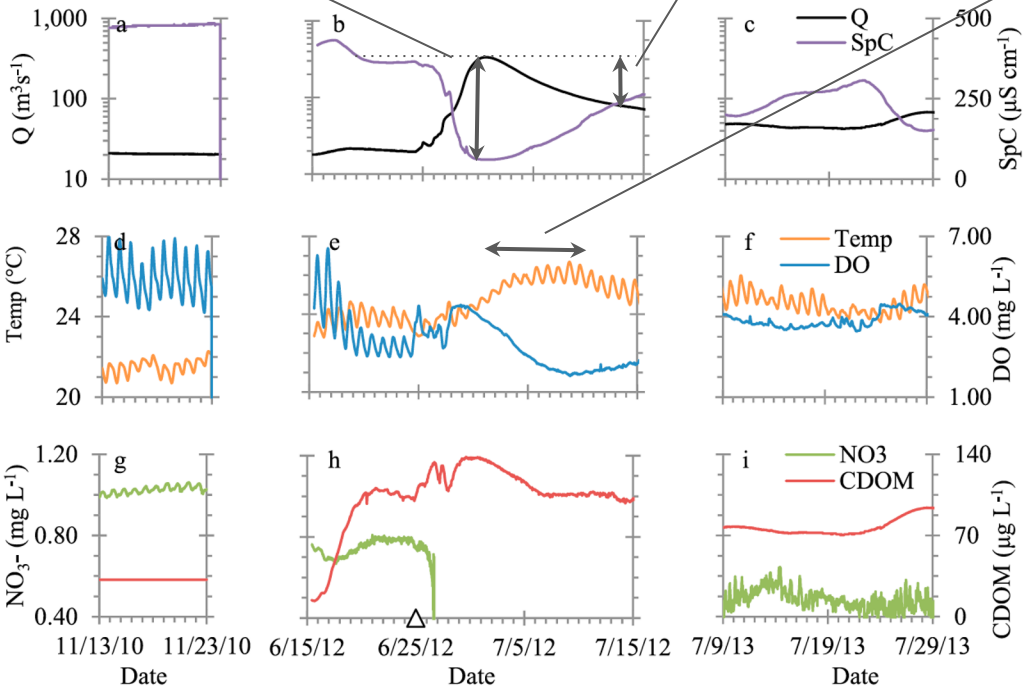
Physical Information into ML Model Derived from Analyses for Science Questions

S1: How do the intensity, duration, and frequency of streamflow disturbance events change water quality at a given location in the short and long-term?

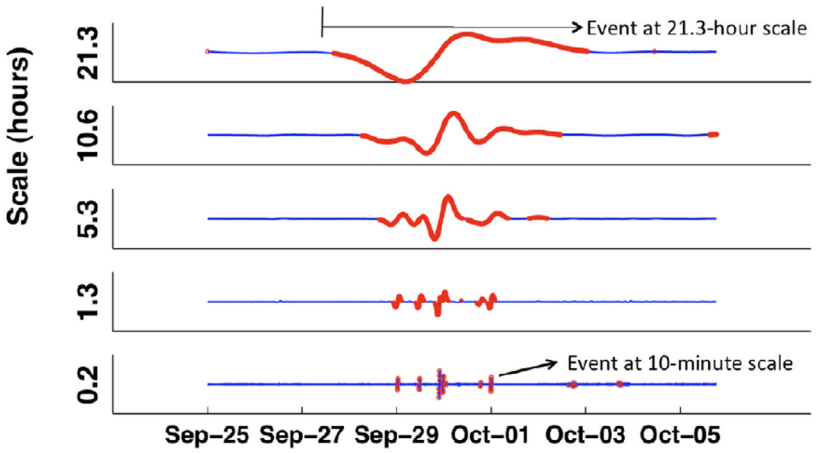
H1: Resistance (Chemostasis)

H2: Resilience

H3: System Lag/Memory Effects



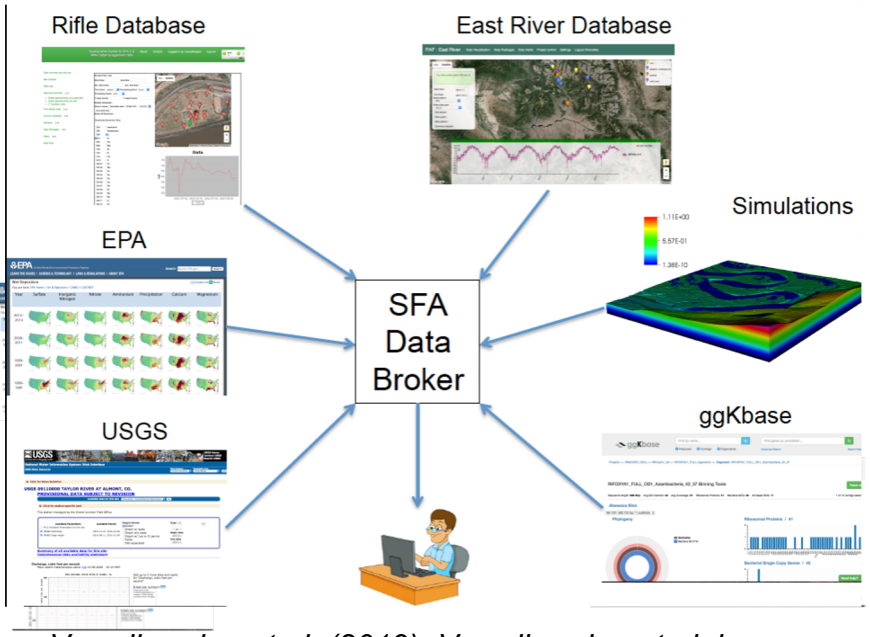
Observed impacts of Tropical Storm Debby in the Santa Fe river, FL. Hensley et al. (2019)



Wavelet-based event identification. Varadharajan and Hemond (2012)

Early Work: Applying a Data Broker For Repeatable Synthesis





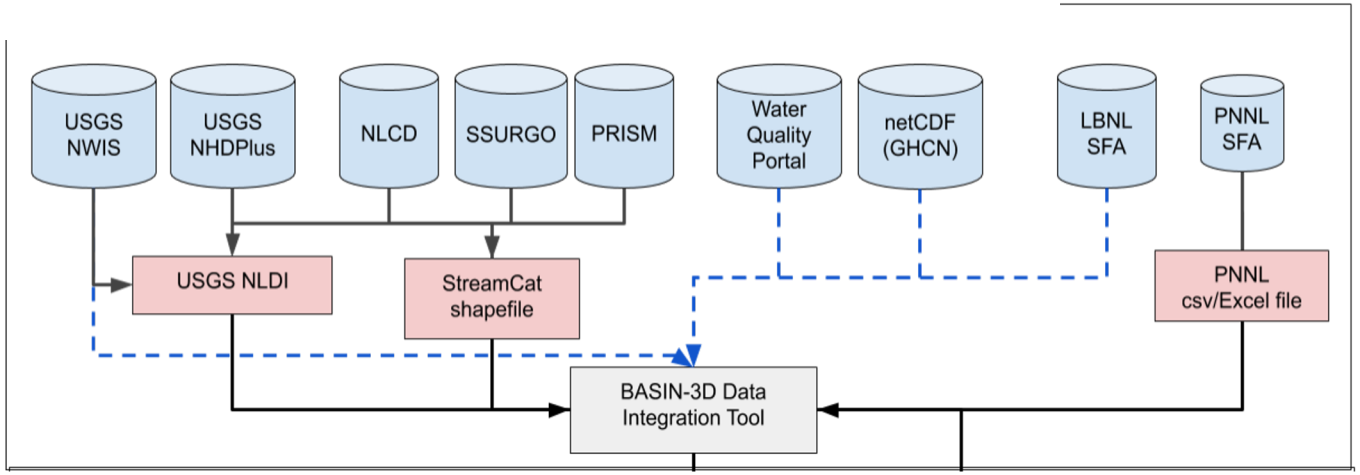
Varadharajan et al. (2019), Varadharajan et al. in prep.

BASIN-3D :
Data Integration Broker that transforms data and metadata to standardized format for synthesis

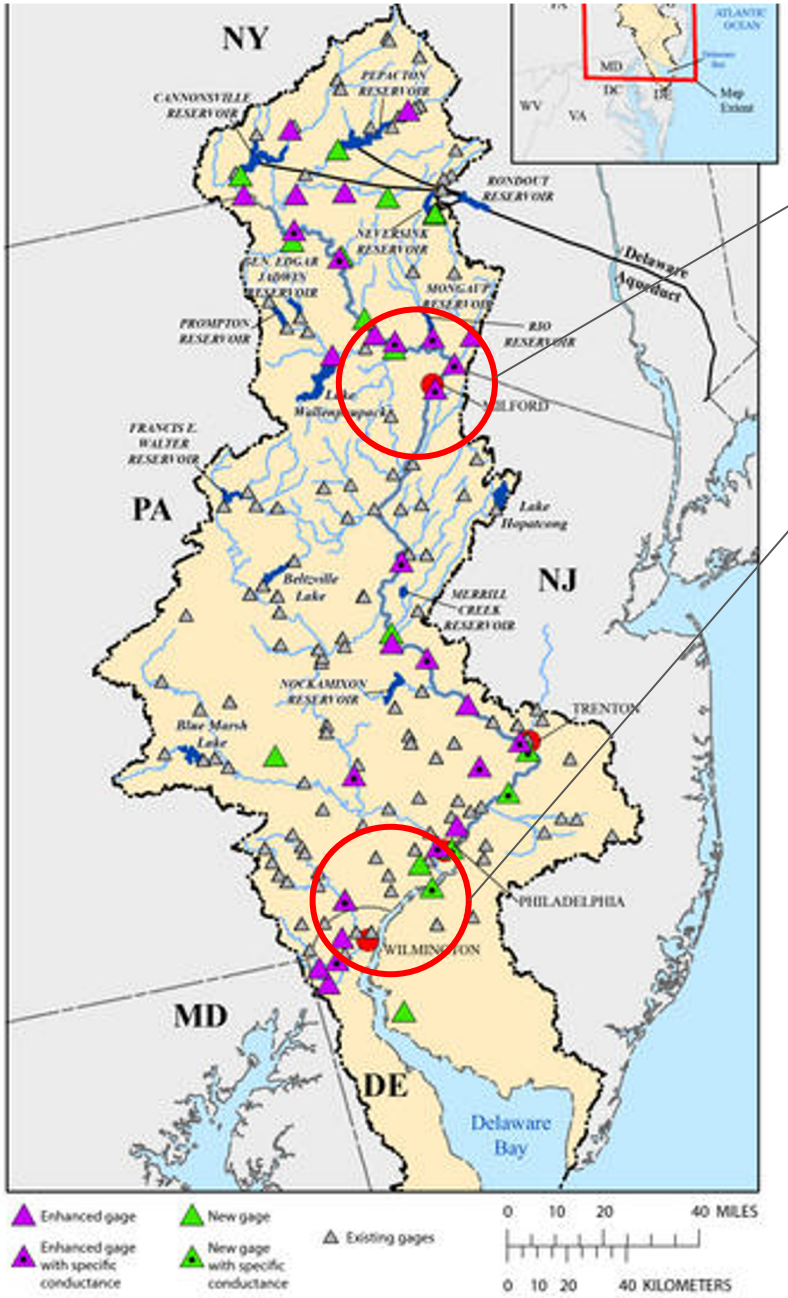
BASIN-3D is different from other data integration tools

PANGEO

INTAKE

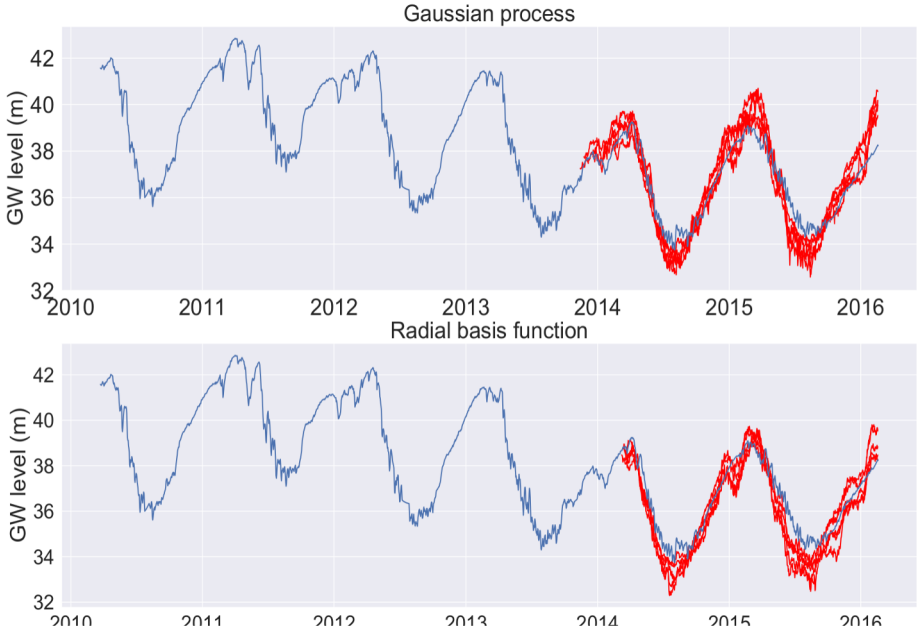


Early Work: ML for Delaware River Corridor



Reservoir/thermal plant impacts on stream temperature

Salt Wedge movement to Flood/Drought and relationship to groundwater

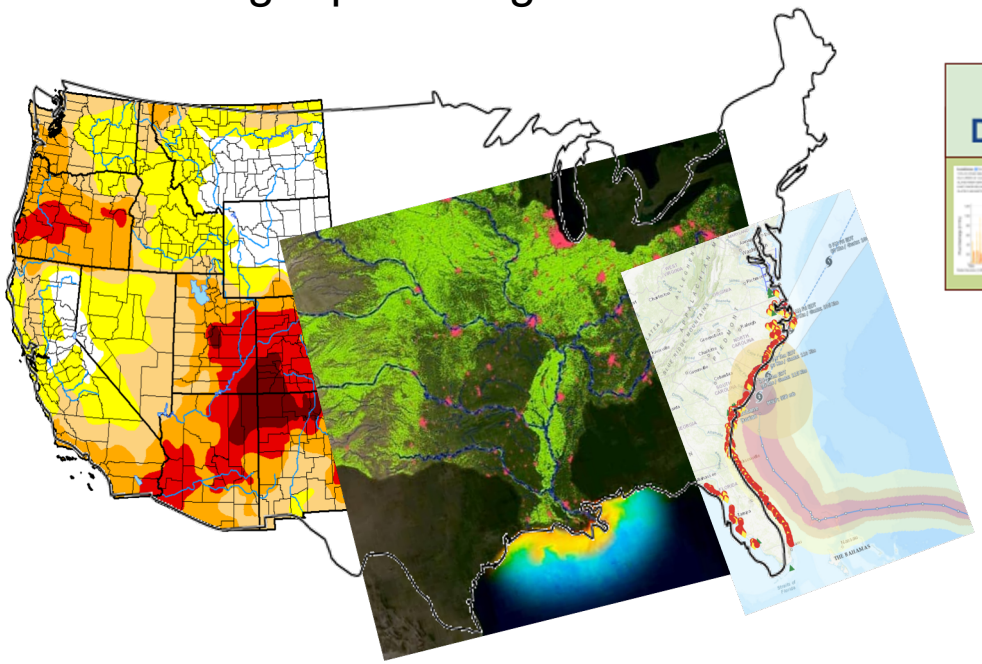


Predictions of groundwater level in Butte County, CA using optimized neural networks (MLP)

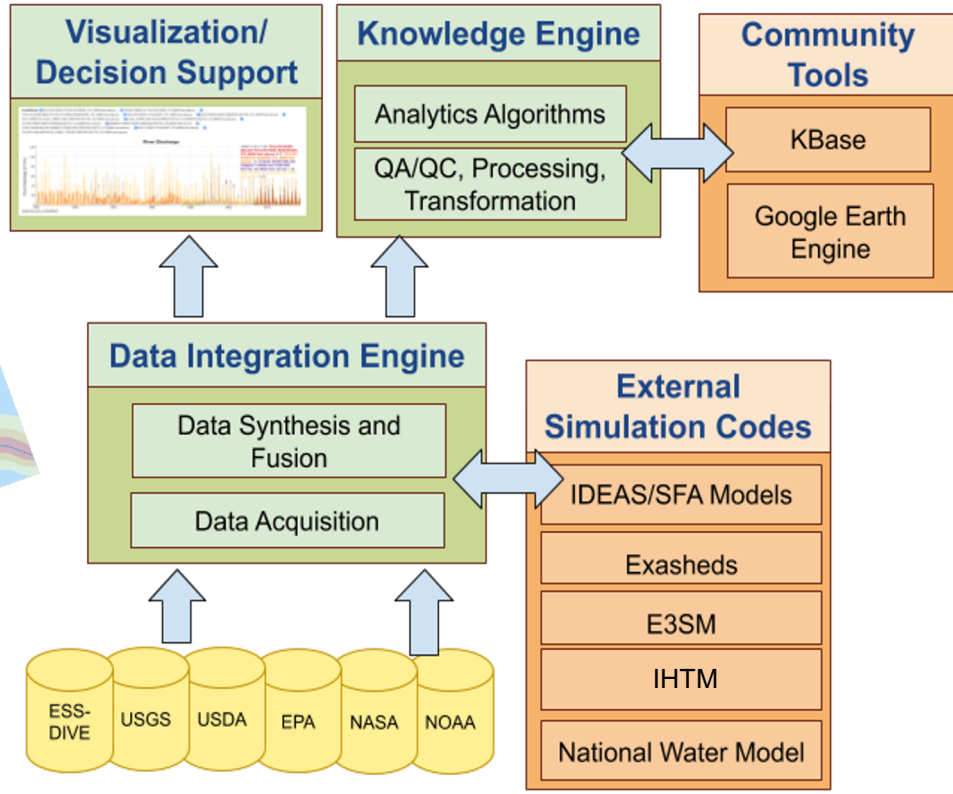
Mueller et al. (2019), Sahu et al. in prep, Berkeley Lab LDRD

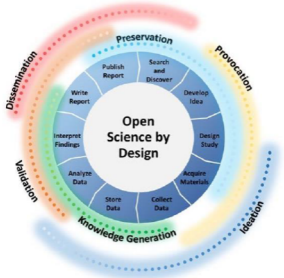
Scientific, Technical, and Social Vision

Transferable concepts of impacts of flow disturbances on water quality to guide predictions in other flood/drought-prone regions



Open data-ML framework, reproducible benchmark datasets will be a foundation for an interagency water quality prediction capability





U.S. DOE (2019)

WHONDRS

Watershed Function

IDEAS Watersheds

ESS-DIVE

ILAMB

USGS

USDA

NSF

NASA

IHTM

Acknowledgements

Funding

U.S. DEPARTMENT OF ENERGY | Office of Science

Facilities

NERSC

ESS-DIVE

BERKELEY LAB SCIENCE-IT

Team S

Watershed Scientific Focus Area Function


Juliane Mueller


Reetik Sahu


Jangho Park


Bhavna Arora


Boris Faybishenko


Deb Agarwal

Watershed Surrogate Modeling Berkeley Lab LDRD

Collaborators

David Moulton (IDEAS)
Forrest Hoffman (ILAMB)
Ken Williams (LBNL SFA)
Tim Scheibe (PNNL SFA)
Brian Pellerin (NGWOS Delaware, USGS)
Jennifer Murphy (USGS water quality assessment)
Kate Maher (Stanford)
Jordan Read+USGS Data Science Group
Julianne Mueller (LBNL CRD)

Mentors

Eoin Brodie
Deb Agarwal
Boris Faybishenko
Rosemary Carroll
Peter Nico