

MARCH 2021



WestFAST News

682 East Vine Street / Suite 7 / Murray, UT 84107 / (801) 685-2555/ www.westernstateswater.org/westfast

Chair – Patrick Lambert; Federal Liaison Officer – Heather Hofman

Water Data Science in 2021

USGS 3/5/21

Where the USGS Water Data Science Branch is headed in 2021

It is an exciting time to be a data science practitioner in environmental science. In the last five years, we've seen massive data growth, modeling improvements, new more inclusive definitions of "impact" in science, and new jobs and duties. The title of "data scientist" has even been formally added as a [job role](#) by the federal government and there are all kinds of data science needs spelled out in the new [USGS science plan](#).

As 2021 progresses, I felt compelled to write up a few of the activities we are focusing on right now, as well as share some ideas we are exploring for the future. If you want to jump ahead, I've included links below to four areas of emphasis our Data Science Branch has this year.

[A renewed focus on water data visualizations](#)

[Advancing machine learning architectures for water prediction](#)

[Forecasting changes to water resources](#)

[Collaborative, reproducible, and efficient data workflows](#)

First, some background

We formally started a data science team in 2014, and this team was converted into the organizational

unit of "Data Science Branch" in the USGS Water Resources Mission Area in late 2017. The scope of this team/branch has evolved over the years but has always included elements of data analysis and prediction, data visualization, reproducibility, and data science training. We've also continued to be motivated by solving problems related to data complexity and data volume, and how we can integrate new data science concepts into more traditional fields like hydrology, ecology, and limnology.

We work for the US Federal Government where change is often slow, but we were able to establish the first formal data science group in the U.S. Geological Survey and continue to benefit from leadership support of this capability. At the same time, what we do and how we do it is pretty different from the way the USGS has operated in the past. Having entire jobs devoted to data visualization or having staff spend the necessary time to make complex data workflows fully reproducible are two examples that have required years of demonstrated usefulness to gain general acceptance. And although machine learning is the main toolset for prediction in many sectors, environmental science has continued to make primary modeling investments in process-based approaches that build on existing knowledge. We've been making progress bridging these modeling approaches by combining them, generating deep learning predictions that are guided by existing theory and are more accurate than either approach alone.

Ok...enough of my rambling and onto the fun stuff.

A renewed focus on water data visualizations



U.S. Geological Survey's VIZLAB (credit E Bechtel)

In 2021 we are stepping up our data visualization game. Step one was hiring Colleen and Ellen to join the team as full-time visualization specialists, as well as Hayley, who does viz, modeling, and workflows. Next, we're holding ourselves to the goal of greater output of high-quality data visualizations. We first started experimenting with data visualization in 2014, visualizing the drought conditions in California with a mix of tech that included D3.js and thoughtful visual design and started using the banner of USGS "Vizlab" (visualization laboratory), going on to release a number of data visualizations on water topics including [microplastics in our waterways](#), [climate change and freshwater fish](#), and [U.S. water use](#). We also have tested different ways to communicate the complexity of flood timing during hurricane events, using both [interactive](#) and [GIFs/videos](#) that were rendered shortly after landfall. But we're making a major leap this year with the way we produce visualizations and we've upped our standards for design. Check out [water sci and mgmt in the Delaware](#), [gages through the ages](#), and [fire hydro](#) for a view into early progress. We're even working on a dataviz that explains how process-guided deep learning models make more accurate predictions (more on that [below](#)). I'm excited to see what's next.

Advancing machine learning architectures for water prediction

There is swirling mix of excitement, fear, misunderstanding, and raw wonder concerning the role of machine learning in water prediction. In the Data Science branch, we're excited about the potential use of ML in all kinds of USGS projects.

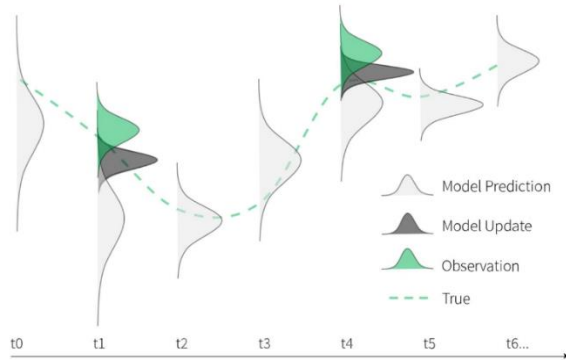
But being part of the USGS, we also understand the positive modeling legacy our agency has, which could be tarnished by misuse or blind adoption of ML. A lot of the fear surrounding ML in environmental sciences comes from a combination of how well it seems to perform in prediction and how seemingly effortless it is to apply ML to the problem du jour seems to be. There is a lot of nuance to each of these assumptions, as we rarely have enough data to build a truly generalizable environmental ML model, and while these models may be faster to spin up, major challenges exist for dealing with data sparsity, estimating uncertainty, and applying knowledge-based constraints that require custom architectures and cutting-edge research. We have partnerships with the Kumar lab at U-MN, the Jia lab at Pitt, and the Shen lab at Penn State to collaborate on the exciting field of "[Knowledge Guided Machine Learning](#)".

We're focusing heavily in 2021 on a subset of KGML - which we're calling Process-Guided Deep Learning - to make improved predictions of water quality or quantity that are guided by physical models and laws (such as conservation of energy, see [here](#)). These are advanced hybrid models that take the best of what our field has to offer for physics-based models and the best machine learning frameworks for spatiotemporally-aware models. The results of these collaborations have been exciting, with large improvements in our predictive accuracy and ability to transfer these models into unseen conditions and still make good predictions. We are currently working on the prediction of water temperature in lakes and streams with these methods, and also working on similar concepts for predicting stream discharge. The place for machine learning in USGS will continue to grow in the next several years and KGML and PGDL will generate some very interesting conversations regarding environmental modeling architectures.

Forecasting changes to water resources

Sequential Data Assimilation

Predict at times t_0-1 , update t_1 , predict t_2-4 , update t_4 , predict $t_5-6...$

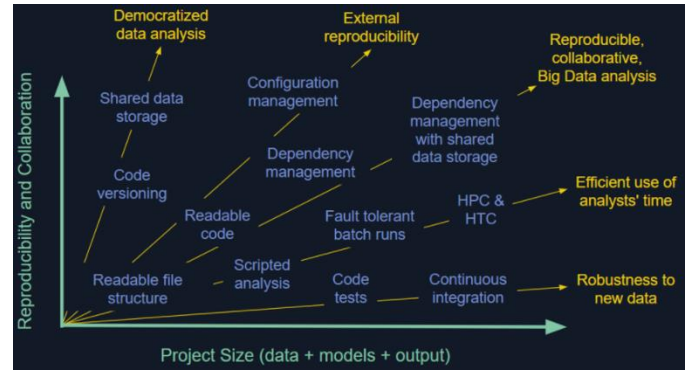


Sequential data assimilation forecasting (credit J Zwart).

Continuing on with the theme of building better predictions with data, we are leading a new project to forecast stream and reservoir temperatures in the Delaware River Basin using data assimilation and machine learning. Data assimilation (DA) is a technique that makes iterative improvements to forecasts by weighting the confidence in new observations (e.g. streamflow and water temperature) with the confidence in model predictions. These improvements or adjustments can include changes to parameters and/or state variables that are internal to the model. DA is an especially powerful tool for environmental forecasts because it can take advantage of the continued accumulation of data to improve predictions and leverage real-time observations for updates, such as those from the thousands of USGS gages that pepper the landscape.

This technique is clearly data-intensive and requires thoughtful automation to perform in a forecast setting. We are able to leverage several of the other Data Science capabilities to strengthen his effort, such as reproducible workflows and data visualization. Jake is currently working on integrating DA with machine learning in this project. Forecasting is relatively new but important territory for USGS, and we're very pleased to be part of leading the simultaneous growth in forecasting and DA in the Water Resources Mission Area.

Capturing reproducible and efficient data workflows



Reproducibility spectrum (credit A Appling)

Reproducibility is a crisis in some research fields and is only getting harder as science becomes larger, more collaborative, more data-intensive, and more integrated across a greater number of disciplines. We're repro-nerds and think it is a tough ask to trust results or findings that aren't backed by a transparent and repeatable workflow, so we emphasize doing so in all aspects of our work. A lot of researchers can get by without investing too much in reproducibility, but it seems inevitable that this strategy will either result in embarrassment or be a blocker to implementing the next logical approach. Worse, an entire research community that undervalues reproducibility may stunt the growth of the field – for example – accepting the status quo of single numeric predictions instead of expecting more useful results that explicitly quantify uncertainty (but which require workflows due to numerous iterations). Each of our data visualizations includes a GitHub repository that exposes all of the code used to build it, including how we fetched, transformed, and visualized the data. We have large-scale data pipelines for building model-ready data that can be updated quickly, and similar workflows for analyzing traffic to USGS Water websites (new hire Rasha is working on both). We've built some of our own tools to do this that use a make-like tracking of data and code dependencies, so that running the whole processing/modeling workflow only runs steps that have upstream data/code that are out of date, while still ensuring that the entire workflow could be run from scratch if needed.

It's an exciting time and we like what we're doing! Contact us to discuss, collaborate, or even join our team. You can follow [@USGS_DataSci](https://twitter.com/USGS_DataSci) on twitter

to stay up to date. We'll be hiring ~4 people this year alone and are on the lookout for people who would be energized to be part of this team.

Office of the Assistant Secretary of the Army for Civil Works and U.S. Army Corps of Engineers to Conduct Stakeholder Sessions on WRDA 2020 Implementation Guidance

ACOE 3/8/21

Washington - The Office of the Assistant Secretary of the Army for Civil Works and the U.S. Army Corps of Engineers (USACE) will conduct a series of listening sessions to collect input and recommendations on development of implementation guidance for any provision contained in the Water Resources Development of Act of 2020 (WRDA 2020).

The Stakeholder Sessions were announced today in the Federal Register at <https://www.federalregister.gov/documents/2021/03/08/2021-04659/water-resources-development-act-of-2020-comment-period-and-stakeholder-sessions>.

The purpose of the listening sessions is to enable stakeholders and other interested parties to provide input on provisions that may require implementation guidance. Stakeholders with interests in a specific USACE mission area are encouraged to participate in the session aligned with that mission area; however, comments on any provision will be heard at every session. USACE encourages participants to share their thoughts and comments on any problems, challenges, impacts or other factors they see as being critical for consideration during the development of implementation guidance.

The Stakeholder Sessions will take place virtually from 1 p.m. to 3 p.m. Eastern Daylight Time (EDT) on the following dates:

- March 16 - Navigation (Inland and Coastal) provisions.

- March 23 - Flood and Coastal Storm Risk Damage provisions.
- March 30 - Ecosystem Restoration and Nuisance Species provisions.
- April 6 – Water Supply and Hydropower provisions.
- April 13 – Open comments for any provision.

Pursuant to protocols related to combatting the spread of COVID-19, the stakeholder sessions will be conducted by electronic means. All stakeholders are welcome to provide oral or written comments during the sessions and will be selected to speak on a first-come, first-served basis. The sessions will be recorded to capture the public's comments.

The webex/teleconferences may be joined five minutes prior to the scheduled 1 p.m. EDT start time using the web link <https://usace1.webex.com/meet/WRDA2020>. Attendees may also call in at 844-800-2712 and use Access Code 199 937 4287 when prompted.

USACE is also accepting stakeholder comments on WRDA 2020 implementation guidance during the ongoing 60-day comment period that ends May 7, 2021, via the following methods:

- Federal eRulemaking Portal: <http://www.regulations.gov/> (our preferred method). Follow the online instructions for submitting comments.
- Email: WRDA2020@usace.army.mil. Include ID No. COE-2021-0002 in the subject line of the message.
- Mail: U.S. Army Corps of Engineers, ATTN: CECW-P 3F91, U.S. Army Corps of Engineers, 441 G St NW, Washington, DC 20314.
- Hand Delivery/Courier: Due to security requirements, we cannot receive comments by hand delivery or courier.

Information about how to submit comments can be found at <https://www.usace.army.mil/Missions/Civil->

[Works/Project-Planning/Legislative-Links/wrda_2020/](#).

The list of WRDA 2020 provisions can be found at

<http://www.usace.army.mil/Missions/Civil-Works/Project-Planning/Legislative-Links/>.

For questions related to the stakeholder sessions, please contact Mr. Gib Owen at 571-274-1929 or gib.a.owen.civ@mail.mil, or Ms. Amy Frantz at 202-761-0106 or WRDA2020@usace.army.mil.

Media inquiries should be directed to Ms. Robyn Colosimo of the Office of the Assistant Secretary of the Army for Civil Works at (703) 209-6354 or Robyn.S.Colosimo.civ@mail.mil.

Contact

Gene Pawlik
202-761-7690
Eugene.A.Pawlik@usace.army.mil

or

Doug Garman
202-761-1807
Doug.M.Garman@usace.army.mil

Upcoming Meetings

[WSWC/ICWP/NWSA Washington, DC Roundtable \(virtual\)](#)

Upcoming Events

[Crop Condition and Soil Moisture Analytics \(Crop-CASMA\) System for Agriculture](#)

Other Federal News

USBR 3/2/21. [Reclamation awards \\$3.6 million to improve desalination technologies](#)

EPA 3/3/21. [EPA Announces Availability of Up to \\$6 Million in Environmental Justice Grants](#)

DOI 3/3/21. [Interior Department Takes Steps to Strengthen Scientific Integrity and Evidence-Based Policymaking](#)

USFS 3/8/21. [USDA Invests \\$285 Million to Improve National Forest and Grassland Infrastructure](#)

USGS 3/9/21. [How we work on the Water Data for the Nation Team](#)

FWS 3/12/21. [American Rescue Plan Provides Critical Support for U.S. Fish and Wildlife Service Programs](#)

EPA 3/15/21. [WaterSense Challenges Homeowners: Take 10 Minutes to Find and Fix a Leak this Week](#)

NRCS 3/15/21. [USDA Invests \\$28 Million in New Projects to Help Restore Lost Wetland Functions, Benefits on Agricultural Landscapes](#)

NRCS 3/16/21. [USDA Seeks Innovative Partner-led Projects Delivering Sustainable Agricultural Solutions](#)

USGS 3/17/21. [New Report Highlights Declining Sagebrush Ecosystem, Provides Foundation for Next Generation of Conservation and Management](#)

USGS 3/17/21. [Parts of US's Southernmost States will "Tropicalize" as Climate Changes](#)

USBR 3/17/21. [Reclamation launches prize competition to improve reliability in hydropower facilities](#)

USGS 3/18/21. [Drought May Lead to Elevated Levels of Naturally Occurring Arsenic in Private Domestic Wells](#)

NRCS 3/18/21. [USDA Announces \\$218 Million Investment in Land and Water Conservation](#)

EPA 3/22/21. [EPA Celebrates World Water Day, Announces \\$2.7 Billion to Improve Water Infrastructure](#)

USBR 3/25/21. [Reclamation provides \\$2.1 million for collaborative watershed projects in seven western states](#)

USGS 3/29/21. [Recreating the U.S. River Conditions animations in R](#)

NRCS 3/29/21. [USDA Announces Expansion of DamWatch® to Serve Forest Service Dams](#)

NRCS 3/30/21. [USDA, Partners Unveil New Fire Mapping Tool with National Wildfire Management Implications](#)

EPA 3/31/21. [Administrator Regan Directs EPA to Reset Critical Science-Focused Federal Advisory Committees](#)

People

EPA 3/11/21. [Michael S. Regan Sworn in as 16th EPA Administrator](#)

DOI 3/17/21. [Secretary Haaland Hits the Ground Running on Day One at the Department of the Interior](#)

The WESTERN STATES FEDERAL AGENCY SUPPORT TEAM (WestFAST) is a collaboration between 12 Federal agencies with water management responsibilities in the West, including: BLM, DOD, EPA, FWS, NASA, NOAA, NPS, NRCS, Reclamation, USACE, USFS, and USGS. WestFAST was established to support the Western States Water Council and the Western Governors' Association in coordinating Federal efforts regarding water issues.