

New machine learning methods could improve environmental predictions

22 June 2021



Credit: CC0 Public Domain

Machine learning algorithms do a lot for us every day—send unwanted email to our spam folder, warn us if our car is about to back into something, and give us recommendations on what TV show to watch next. Now, we are increasingly using these same algorithms to make environmental predictions for us.

A team of researchers from the University of Minnesota, University of Pittsburgh, and U.S. Geological Survey recently published a new study on predicting flow and temperature in [river networks](#) in the 2021 Society for Industrial and Applied Mathematics (SIAM) International Conference on Data Mining (SDM21) proceedings.

The research demonstrates a new machine learning method where the algorithm is "taught" the rules of the physical world in order to make better predictions and steer the algorithm toward physically meaningful relationships between inputs and outputs.

The study presents a model that can make more

accurate river and stream temperature predictions, even when little data is available, which is the case in most rivers and streams. The model can also better generalize to different time periods.

"Water temperature in streams is a 'master variable' for many important aquatic systems, including the suitability of aquatic habitats, evaporation rates, greenhouse gas exchange, and efficiency of thermoelectric energy production," said Xiaowei Jia, a lead author of the study and assistant professor in the University of Pittsburgh's Department of Computer Science at University in the School of Computing and Information. "Accurate [prediction of water temperature](#) and streamflow also aids in decision making for resource managers, for example helping them to determine when and how much water to release from reservoirs to downstream rivers.

A common criticism of machine learning is that the predictions aren't rooted in physical meaning. That is, the algorithms are just finding correlations between inputs and outputs, and sometimes those correlations can be "spurious" or give false results. The model often won't be able to handle a situation where the relationship between inputs and outputs changes.

The new method published by Jia, who is also a 2020 Ph.D. graduate of the University of Minnesota Department of Computer Science and Engineering in the College of Science and Engineering, and his colleagues uses "process-guided or knowledge-guided machine learning." This method is applied to a use case of water temperature prediction in the Delaware River Basin (DRB) and is designed to overcome some of the common pitfalls of prediction using machine learning. The method informs the machine learning model with a relatively simple process—correlation through time, the spatial connections between streams, and energy budget equations.

Data sparsity and variability in stream temperature dynamics are not unique to the Delaware River Basin. Relative to most of the continental United States, the Delaware River Basin is well-monitored for water temperature. The Delaware River Basin is therefore an ideal place to develop new methods for stream temperature prediction.

[An interactive visual explainer released by the U.S. Geological Survey](#) highlights these model developments and the importance of water temperature predictions in the DRB. The visualization demonstrates the societal need for water temperature predictions, where reservoirs provide drinking water to more than 15 million people, but also have competing water demands to maintain downstream flows and cold-water habitat for important game fish species. Reservoir managers can release cold water when they anticipate water temperature will exceed critical thresholds and having accurate water temperature predictions is key to using limited water resources only when necessary.

The recent study builds on a collaboration between [water](#) scientists at the U.S. Geological Survey and University of Minnesota Twin Cities computer scientists in Professor Vipin Kumar's lab in the College of Science and Engineering's Department of Computer Science and Engineering, where researchers have been developing knowledge-guided machine learning techniques.

"These knowledge-guided machine learning techniques are fundamentally more powerful than standard machine learning approaches and traditional mechanistic models used by the scientific community to address environmental problems," Kumar said.

These new generation of machine learning methods, funded by NSF's Harnessing the Data Revolution Program, are being used to address a variety of environmental problems such as improving lake and stream temperature predictions.

In [another new NSF-funded study on predicting water temperature dynamics of unmonitored lakes](#) in the American Geophysical Union's *Water Resources Research* led by University of

Minnesota Department of Computer Science and Engineering Ph.D. candidate Jared Willard, researchers show how knowledge-guided machine learning models were used to solve one of the most challenging environmental prediction problems—prediction in unmonitored ecosystems.

Models were transferred from well-observed lakes to lakes with few to no observations, leading to accurate predictions even in lakes where [temperature](#) observations don't exist. Researchers say their approach readily scales to thousands of lakes, demonstrating that the method (with meaningful predictor variables and high-quality source models) is a promising approach for many kinds of unmonitored systems and environmental variables in the future.

More information: Xiaowei Jia et al, Physics-Guided Recurrent Graph Model for Predicting Flow and Temperature in River Networks, *Proceedings of the 2021 SIAM International Conference on Data Mining (SDM)* (2021). [DOI: 10.1137/1.9781611976700.69](#)

Provided by University of Minnesota

APA citation: New machine learning methods could improve environmental predictions (2021, June 22) retrieved 21 October 2021 from <https://techxplore.com/news/2021-06-machine-methods-environmental.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.