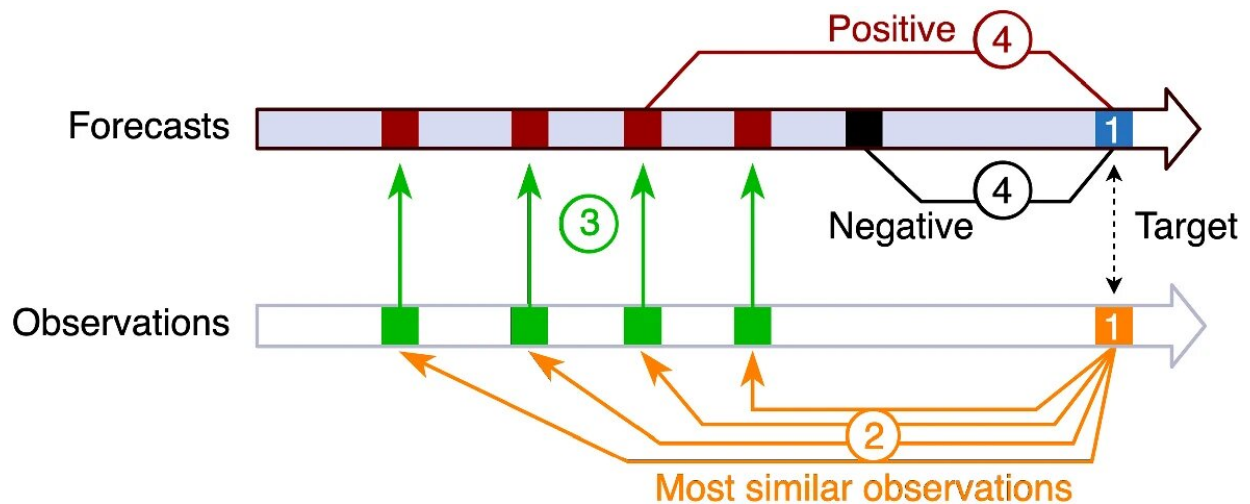# Machine learning technology boosts analog weather forecasting

May 24 2023, by Matthew Carroll

Schematic of the reverse analog for constructing triplet samples. Credit: *Boundary-Layer Meteorology* (2023). DOI: 10.1007/s10546-022-00779-6

Machine learning technology that can recognize human faces may also help to improve weather forecasts, according to a team of scientists.

"The idea behind this work comes from Google's FaceNet, but instead of comparing your picture to images of faces in a database, we are comparing weather to historical forecasts," said Weiming Hu, a machine learning scientist at the University of San Diego and a former doctoral student at Penn State.

The scientists applied a [deep learning algorithm](link) to [analog](link) weather forecasting, which uses past weather conditions to make future forecasts. They found that analyzing surface wind speed and solar irradiance forecasts in Pennsylvania from 2017 to 2019 using machine learning improved analog forecasting accuracy in this [case study](link).

"You want to understand how much energy you can expect for the day ahead," said Hu, who received his doctorate in geography from Penn State. "You want to understand the risk—no matter if you over predict or under predict there are going to be penalties like power shortages or overproduction. Our work shows we can improve the accuracy of these wind and solar forecasts."

Analog forecasting is an alternative to numerical weathering prediction (NWP), which uses computer models to simulate how initial weather conditions will evolve in the days or weeks ahead. NWP has led to great advances in forecasting over the last several decades, but uncertainties remain.

Those uncertainties are addressed in part by running a number of simulations, called ensembles, which show a range of possible future atmospheric states but are also computationally intensive and expensive to produce, the scientists said.

"Analog forecasting, however, can generate ensembles without expensive, repeated model runs," Hu said. "It works by searching for historical forecasts that are most similar to the target forecast. And then the past observations associated with the most similar past forecasts make up the ensemble members."

The analog ensembles are produced by combining a deterministic forecast—a highly detailed single run of an NWP model—with past weather observations—like temperature, pressure and humidity—from

past forecasts that are similar to the current one.

The best analogs are chosen based on a similarity metric that weighs individual weather [forecast](link) predictors, but this process uses a constrained exhaustive search that limits the number of predictors that can be used and does not consider the relationships among predictors.

"That has been the limitation for analog ensemble forecasting," Hu said. "This paper tries to address that by introducing a machine learning approach to learn the intricacy among predictors."

The machine learning technique takes all the weather variables—like temperature, pressure and humidity—and transforms them into a latent space, or a clustered pattern that is helpful for selecting the ideal forecasts and analogs, the scientists said.

"This approach tries to identify the most helpful features to look for to improve analog forecasts," Hu said. "Simply put, it is clustering the candidates, and that gives you the most accurate forecasts and pushes away the less similar points of data from less similar forecasts."

The machine learning technique overcomes the computational limit posed by optimizing predictor weights in traditional analog ensemble forecasting, the scientists said.

"Machine learning has been used operationally for many years to speed up or improve the accuracy of predictions, however its role was mainly limited to post processing or data preparation," said Guido Cervone, Penn State professor of geography and meteorology and [atmospheric science](link), Hu's adviser and a co-author of the paper. "It is really during the last year or so that machine learning has been used as a central core of algorithms, often even replacing numerical model solutions."

The results of the study, published in the journal *Boundary-Layer Meteorology*, suggest machine learning will enable more predictors to be used and will generate predictions with higher accuracy.

"Our work shows that a machine learning model can be used for looking at complex features even in a geosciences field," Hu said. "In geosciences, we are dealing with hundreds of variables. In this search, we had more than 300. And most of the time they have a lot of correlations. We show machine learning can actually detect all those relationships from this large dataset."

More information: Weiming Hu et al, Machine Learning Weather Analogs for Near-Surface Variables, *Boundary-Layer Meteorology* (2023). DOI: 10.1007/s10546-022-00779-6

Provided by Pennsylvania State University