Deep neural networks speed up weather and climate models
13 November 2019

Jiali Wang and Rao Kotamarthi, along with Prasanna Balaprakash, were co-authors on the Geoscientific Model Development that focused on the planetary boundary layer. Credit: Argonne National Laboratory

When you check the weather forecast in the morning, the results you see are more than likely determined by the Weather Research and Forecasting (WRF) model, a comprehensive model that simulates the evolution of many aspects of the physical world around us.

"It describes everything you see outside of your window," said Jiali Wang, an environmental scientist at the U.S. Department of Energy's (DOE) Argonne National Laboratory, "from the clouds, to the sun's radiation, to snow to vegetation—even the way skyscrapers disrupt the wind."

The myriad characteristics and causes of weather and climate are coupled together, communicating with one another. Scientists have yet to fully describe these complex relationships with simple, unified equations. Instead, they approximate the equations using a method called parameterization in which they model the relationships at a scale greater than that of the actual phenomena.

Although parameterizations simplify the physics in a way that allows the models to produce relatively accurate results in a reasonable time, they are still computationally expensive. Environmental scientists and computational scientists from Argonne are collaborating to use deep neural networks, a type of machine learning, to replace the parameterizations of certain physical schemes in the WRF model, significantly reducing simulation time.

"With less-expensive models, we can achieve higher-resolution simulations to predict how short-term and long-term changes in weather patterns affect the local scale," said Wang, "even down to neighborhoods or specific critical infrastructure."

In a recent study, the scientists focused on the planetary boundary layer (PBL), or lowest part of the atmosphere. The PBL is the atmospheric layer that human activity affects the most, and it extends only a few hundred meters above Earth's surface. The dynamics in this layer, such as wind velocity, temperature and humidity profiles, are critical in determining many of the physical processes in the rest of the atmosphere and on Earth.

The PBL is a crucial component in the WRF model, but it is also one of the least computationally expensive. This makes it an excellent testbed for studying how more complicated components might be enhanced by deep learning neural networks in the same way.

"We used 20 years of computer-generated data from the WRF model to train the neural networks and two years of data to evaluate whether they could provide an accurate alternative to the physics-based parameterizations," said Prasanna Balaprakash, a computer scientist and DOE Early Career Award recipient in Argonne's Mathematics and Computer Science division and the Argonne Leadership Computing Facility (ALCF), a DOE Office of Science User Facility.
Balaprakash developed the neural network and trained it to learn an abstract relationship between the inputs and outputs by feeding it more than 10,000 data points (8 per day) from two locations, one in Kansas and one in Alaska. The result was an algorithm that the scientists are confident could replace the PBL parameterization in the WRF model.

The scientists demonstrated that a deep neural network that considers some of the underlying structure of the relationship between the input and output variables can successfully simulate wind velocities, temperature and water vapor over time. The results also show that a trained neural network from one location can predict behavior across nearby locations with correlations higher than 90 percent compared with the test data.

"Collaboration between the climate scientists and the computer scientists was crucial for the results we achieved," said Rao Kotamarthi, chief scientist and department head of atmospheric science and climate research in Argonne’s Environmental Science division. "Incorporating our domain knowledge makes the algorithm much more predictive."

The algorithms—called domain-aware neural networks—that consider known relationships not only can predict environmental data more accurately, but they also require training of significantly less data than do algorithms that do not consider domain expertise.

Any machine learning project requires a large amount of high-quality data, and there was no shortage of data for this study. Supercomputing resources at the ALCF and the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility at Lawrence Berkeley National Laboratory, contributed to the production of more than 300 years (700 terabytes) of data describing past, present and future weather and climate in North America.

"This database is unique to climate science at Argonne," said Wang, "and we are using it to conduct further studies in deep learning and determine how it can apply to climate models."

The scientists' ultimate goal is to replace all of the expensive parameterizations in the WRF model with deep learning neural networks to enable faster and higher-resolution simulation.

Currently, the team is working to emulate long-wave and short-wave solar radiation parameterization—two portions of the WRF model that together take up almost 40% of the calculation time of the physics in the simulations.

The results of the study were published in a paper titled "Fast domain-aware neural network emulation of a planetary boundary layer parameterization in a numerical weather forecast model" on October 10 in Geoscientific Model Development 12, 4261–4274, 2019.


Provided by Argonne National Laboratory

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.