Scientists turn to deep learning to improve air quality forecasts

March 31 2021, by Matthew Carroll

Highway traffic in the greater Los Angeles area. Credit: Pixabay Highway traffic in the greater Los Angeles area Credit: Pixabay

Air pollution from the burning of fossil fuels impacts human health but predicting pollution levels at a given time and place remains challenging,

according to a team of scientists who are turning to deep learning to improve air quality estimates. Results of the team's study could be helpful for modelers examining how economic factors like industrial productivity and health factors like hospitalizations change with pollution levels.

"Air quality is one of the major issues within an urban area that affects people's lives," said Manzhu Yu, assistant professor of geography at Penn State. "Yet existing observations are not adequate to provide comprehensive information that may help vulnerable populations to plan ahead."

Satellite and ground-based observations each measure air pollution, but they are limited, the scientists said. Satellites, for instance, may pass a given location at the same time each day and miss how emissions vary at different hours. Ground-based weather stations continuously collect data but only in a limited number of locations.

To address this, the scientists used deep learning, a type of machine learning, to analyze the relationship between satellite and ground-based observations of nitrogen dioxide in the greater Los Angeles area. Nitrogen dioxide is largely associated with emissions from traffic and power plants, the scientists said.

"The problem right now is nitrogen dioxide varies a lot during the day," Yu said. "But we haven't had an hourly, sub-urban scale product available to track air pollution. By comparing surface level and satellite observations, we can actually produce estimates with higher spatial and temporal resolution."

The learned relationship allowed the researchers to take daily satellite observations and create hourly estimates of atmospheric nitrogen dioxide in roughly 3-mile grids, the scientists said. They recently reported their

findings in the journal *Science of the Total Environment.*

"The challenge here is whether we can find a linkage between measurements from earth's surface and satellite observations of the troposphere, which are actually far away from each other. That's where deep learning comes in."

Deep learning algorithms operate much like the human brain and feature multiple layers of artificial neurons for processing data and creating patterns. The system learns and trains itself based on connections it finds within large amounts of data, the scientists said.

The scientists tested two deep-learning algorithms and found the one that compared the ground-based observations directly to the satellite observations more accurately predicted nitrogen dioxide levels. Adding information like meteorological data, elevation and the locations of the ground-based stations and major roads and power plants improved the prediction accuracy further.

Yu said the study could be repeated for other greenhouse gases and applied to different cities or on regional and continental scales, the scientists said. In addition, the model could be updated when new, higher-resolution satellites are launched.

"With a high spatiotemporal resolution, our results will facilitate the study between air quality and health issues and improve the understanding of the dynamic evolution of airborne pollutants," Yu said.