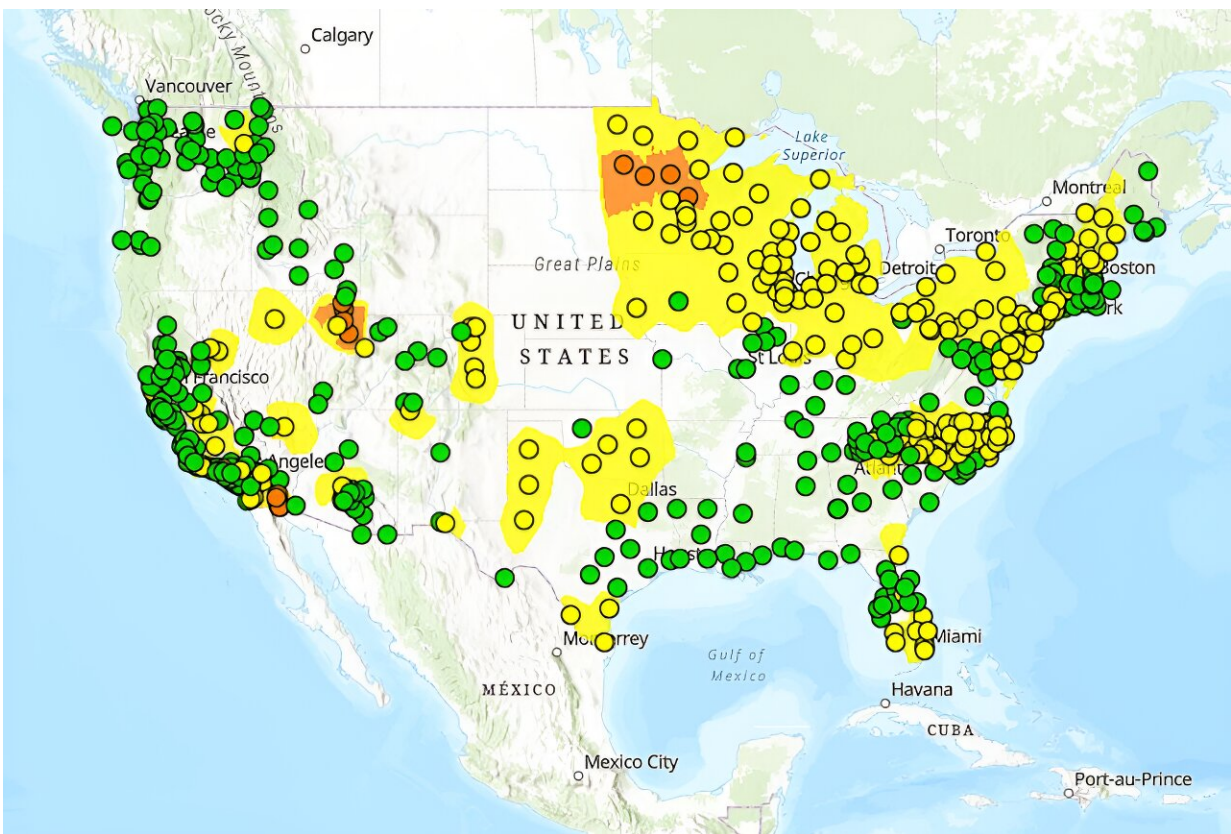# AI can help forecast air quality, but freak events like 2023's summer of wildfire smoke require traditional methods too

August 9 2023, by Joshua S. Fu

The EPA's AirNow air pollution forecasts use machine learning. During wildfire events, a smoke-transport and dispersion model helps to simulate the spread of smoke plumes. This map is the forecast for Aug. 9, 2023. Yellow indicates moderate risk; orange indicates unhealthy air for sensitive groups. Credit: AirNow.gov

Wildfire smoke from [Canada's extreme fire season](#) has left a lot of people thinking about air quality and wondering what to expect in the days ahead.

All air contains gaseous compounds and small particles. But as air quality gets worse, these gases and particles can trigger asthma and exacerbate heart and respiratory problems as they enter the nose, throat and lungs and even circulate in the bloodstream. When wildfire smoke turned New York City's skies orange in early June 2023, [emergency room visits](#) for asthma doubled.

In [most cities,](#) it's easy to find a daily [air quality index score](#) that tells you when the air is considered unhealthy or even hazardous. However, predicting air quality in the days ahead isn't so simple.

I work on air quality forecasting as a [professor of civil and environmental engineering](#). Artificial intelligence has improved these forecasts, but research shows it's much more useful when paired with traditional techniques. Here's why:

How scientists predict air quality

To predict air quality in the near future—a few days ahead or longer—scientists generally rely on two [main methods](#): a [chemical transport model](#) or a machine-learning [model](#). These two models generate results in totally different ways.

Chemical transport models use lots of known chemical and physical formulas to calculate the presence and production of air pollutants. They use data from emissions inventories reported by local agencies that list pollutants from known sources, such as wildfires, traffic [or factories,](#) and data from meteorology that provides atmospheric information, such as wind, precipitation, temperature and solar radiation.

These models simulate the flow and chemical reactions of the air pollutants. However, their simulations involve multiple variables with huge uncertainties. Cloudiness, for example, changes the incoming solar radiation and thus the photochemistry. This can make the results less accurate.

Machine-learning models instead learn patterns over time from [historical data](#) to predict future air quality for any given region, and then apply that knowledge to current conditions to predict the future.

The downside of [machine-learning models](#) is that they do not consider any chemical and physical mechanisms, as chemical transport models do. Also, the accuracy of machine-learning projections under [extreme conditions](#), such as [heat waves](#) or wildfire events, can be off if the models weren't trained on such data. So, while machine-learning models can show where and when high pollution levels are most likely, such as during rush hour near freeways, they generally cannot deal with more random events, like wildfire smoke blowing in from Canada.

Which is better?

Scientists have determined that neither model is accurate enough on its own, but using the [best attributes of both](#) models together [can help better predict the quality](#) of the air we breathe.

This combined method, known as the machine-learning—measurement model fusion, or ML-MMF, has the ability to provide science-based predictions with [more than 90% accuracy](#). It is based on known physical and chemical mechanisms and can simulate the whole process, from the air pollution source to your nose. Adding [satellite data](#) can help them inform the public on both air quality safety levels and the direction pollutants are traveling with greater accuracy.

We recently [compared predictions from all three models](#) with actual pollution measurements. The results were striking: The combined model was 66% more accurate than the chemical transport model and 12% more accurate than the [machine-learning model](#) alone.

The [chemical](#) transport model is still the most common method used today to predict air quality, but applications with [machine-learning](#) models are becoming more popular. The regular [forecasting method](#) used by the U.S. Environmental Protection Agency's [AirNow.gov](#) relies on machine learning. The site also compiles air quality forecast results from state and local agencies, most of which use [chemical transport models](#).

As information sources become more reliable, the combined models will become more accurate ways to forecast hazardous [air quality](#), particularly during unpredictable events like wildfire smoke.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation