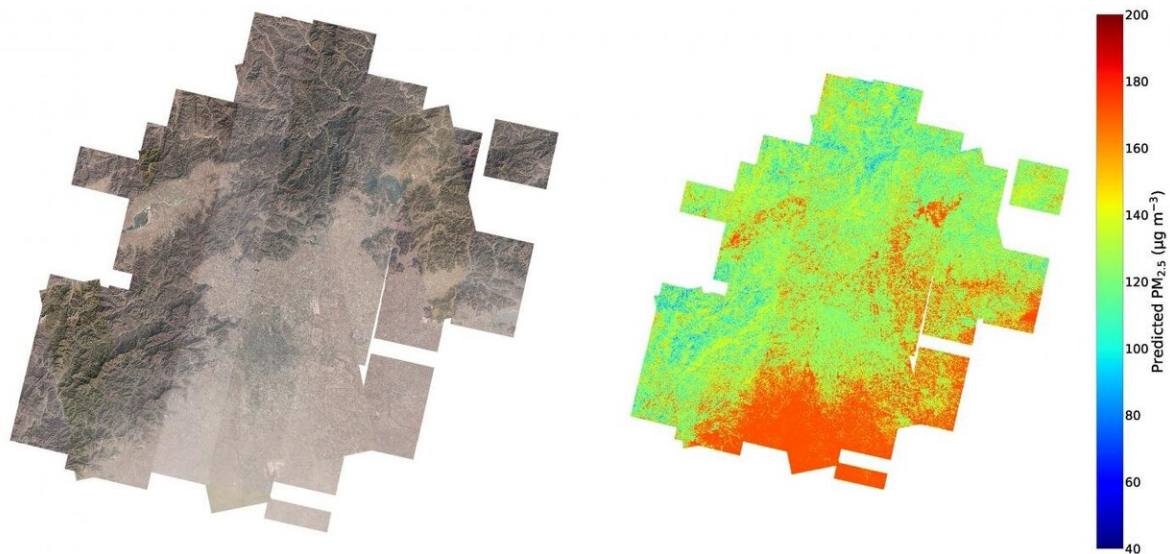


Spotting air pollution with satellites, better than ever before

April 22 2020, by Ken Kingery



A detailed map of the pollution levels in Beijing and its surrounding areas using a new machine learning algorithm for satellite images and weather. Credit: Tongshu Zheng, Duke University

Researchers from Duke University have devised a method for estimating the air quality over a small patch of land using nothing but satellite imagery and weather conditions. Such information could help researchers identify hidden hotspots of dangerous pollution, greatly improve studies of pollution on human health, or potentially tease out the effects of unpredictable events on air quality, such as the breakout of an

airborne global pandemic.

The results appear online in the journal *Atmospheric Environment*.

"We've used a new generation of micro-satellite images to estimate ground-level air pollution at the smallest spatial scale to date," said Mike Bergin, professor of civil and environmental engineering at Duke.

"We've been able to do it by developing a totally new approach that uses AI/machine learning to interpret data from surface images and existing [ground stations](#)."

The specific air quality measurement that Bergin and his colleagues are interested in is the amount of tiny airborne particles called PM_{2.5}. These are particles that have a diameter of less than 2.5 micrometers—about three percent of the diameter of a human hair—and have been shown to have a dramatic effect on [human health](#) because of their ability to travel deep into the lungs.

For example, PM_{2.5} was globally ranked as the fifth mortality risk factor, responsible for about 4.2 million deaths and 103.1 million years of life lost or lived with disability, by the 2015 Global Burden of Disease study. And in a recent study from the Harvard University T.H. Chan School of Public Health, researchers found that areas with higher levels of PM_{2.5} also are associated with higher death rates due to COVID-19.

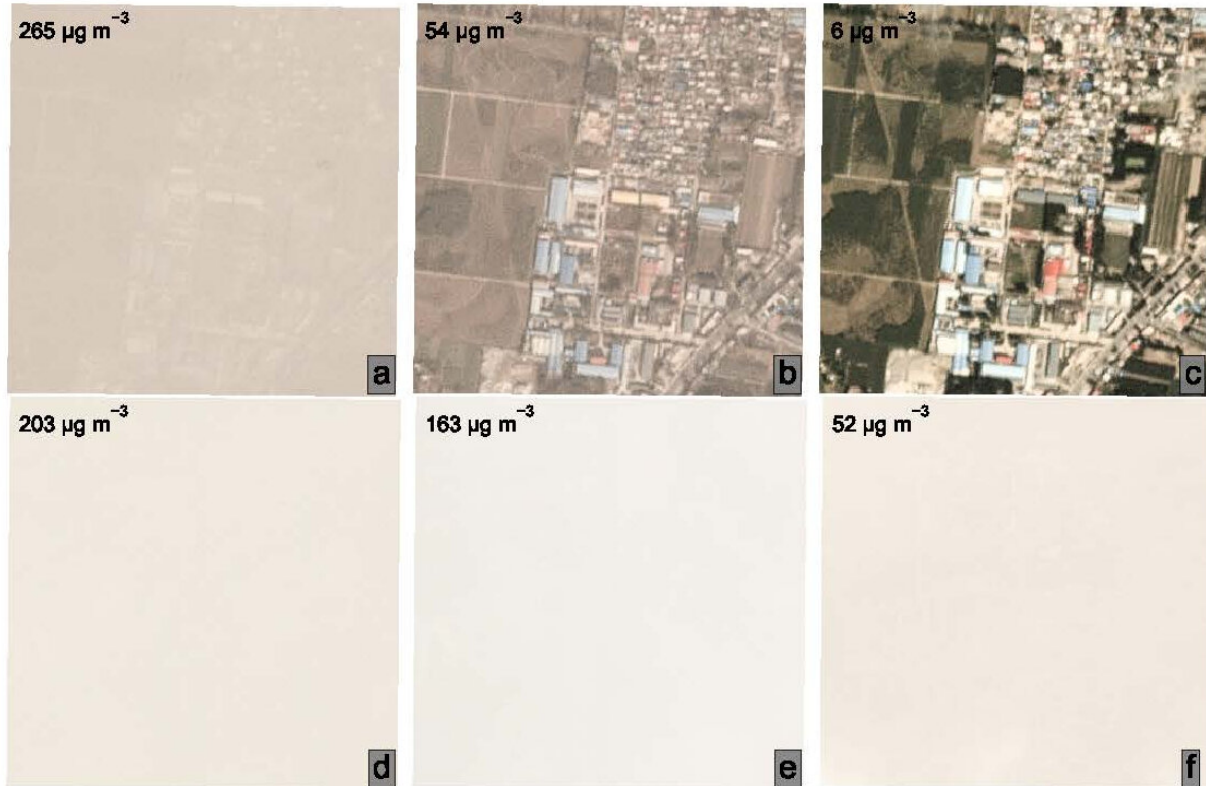
Current best practices in remote sensing to estimate the amount of ground-level PM_{2.5} use satellites to measure how much sunlight is scattered back to space by ambient particulates over the entire atmospheric column. This method, however, can suffer from regional uncertainties such as clouds and shiny surfaces, atmospheric mixing, and properties of the PM particles, and cannot make accurate estimates at scales smaller than about a square kilometer. While ground pollution monitoring stations can provide direct measurements, they suffer from

their own host of drawbacks and are only sparsely located around the world.

"Ground stations are expensive to build and maintain, so even large cities aren't likely to have more than a handful of them," said Bergin. "Plus they're almost always put in areas away from traffic and other large local sources, so while they might give a general idea of the amount of PM2.5 in the air, they don't come anywhere near giving a true distribution for the people living in different areas throughout that city."

In their search for a better method, Bergin and his doctoral student Tongshu Zheng turned to Planet, an American company that uses micro-satellites to take pictures of the entire Earth's surface every single day with a resolution of three meters per pixel. The team was able to get daily snapshot of Beijing over the past three years.

The key breakthrough came when David Carlson, an assistant professor of civil and environmental engineering at Duke and an expert in machine learning, stepped in to help.



Six different photos taken of the same area of Beijing on different days experiencing different levels of air pollution. While the naked human eye can clearly tell that some days are more polluted than others, a new machine learning algorithm can make reasonably accurate estimates air pollution at ground level. Credit: Tongshu Zheng, Duke University

"When I go to machine learning and artificial intelligence conferences, I'm usually the only person from an environmental engineering department," said Carlson. "But these are the exact types of projects that I'm here to help support, and why Duke places such a high importance on hiring data experts throughout the entire university."

With Carlson's help, Bergin and Zheng applied a convolutional neural network with a random forest algorithm to the image set, combined with

meteorological data from Beijing's weather station. While that may sound like a mouthful, it's not that difficult to pick your way through the trees.

A random forest is a standard machine learning algorithm that uses a lot of different decision trees to make a prediction. We've all seen decision trees, perhaps as an internet meme that uses a series of branching yes/no questions to decide whether or not to eat a burrito. Except in this case, the algorithm is looking through decision trees based on metrics such as wind, relative humidity, temperature and more, and using the resulting answers to arrive at an estimate for PM2.5 concentrations.

However, random forest algorithms don't deal well with images. That's where the convolutional neural networks come in. These algorithms look for common features in images such as lines and bumps and begin grouping them together. As the algorithm "zooms out," it continues to lump similar groupings together, combining basic shapes into common features such as buildings and highways. Eventually the algorithm comes up with a summary of the image as a list of its most common features, and these get thrown into the random forest along with the weather data.

"High-pollution images are definitely foggier and blurrier than normal images, but the human eye can't really tell the exact pollution levels from those details," said Carlson. "But the algorithm can pick out these differences in both the low-level and high-level features—edges are blurrier and shapes are obscured more—and precisely turn them into air quality estimates."

"The [convolutional neural network](#) doesn't give us as good of a prediction as we would like with the images alone," added Zheng. "But when you put those results into a random forest with weather data, the results are as good as anything else currently available, if not better."

In the study, the researchers used 10,400 images to train their model to predict local levels of PM2.5 using nothing but satellite images and weather conditions. They tested their resulting model on another 2,622 images to see how well it could predict PM2.5.

They show that, on average, their model is accurate to within 24 percent of actual PM2.5 levels measured at reference stations, which is at the high end of the spectrum for these types of models, while also having a much higher spatial resolution. While most of the current standard practices can predict levels down to 1 million square meters, the new method is accurate down to 40,000—about the size of eight football fields placed side-by-side.

With that level of specificity and accuracy, Bergin believes their method will open up a wide range of new uses for such models.

"We think this is a huge innovation in satellite retrievals of [air quality](#) and will be the backbone of a lot of research to come," said Bergin.

"We're already starting to get inquiries into using it to look at how levels of PM2.5 are going to change once the world starts recovering from the spread of COVID-19."

More information: Tongshu Zheng et al, Estimating ground-level PM2.5 using micro-satellite images by a convolutional neural network and random forest approach, *Atmospheric Environment* (2020). [DOI: 10.1016/j.atmosenv.2020.117451](#)

Provided by Duke University

Citation: Spotting air pollution with satellites, better than ever before (2020, April 22) retrieved 28 June 2024 from <https://phys.org/news/2020-04-air-pollution-satellites.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.