Machine learning our ways to better air quality forecasts

Mat Evans[1,2], Peter Ivatt[1]
[1]*Wolfson Atmospheric Chemistry Laboratories, Department Chemistry, University of York*
[2]*National Centre for Atmospheric Sciences, University of York*

Sophisticated tools to measure and to model air pollution exist. However, providing strong links between the two are difficult. Improving the initial state of the model can be achieved through data assimilation approaches. However, there are few tools to deal with biases in the forward predictions of the model. Here we describe a new methodology which uses a machine learning approach to correct the biases seen in a model of atmospheric chemistry. Observations of ozone (a critical pollution) made at the ground- and balloon-based platforms are compared to the predictions made by a chemistry transport model (GEOS-Chem) and the bias evaluated. A machine learning (Gradient Boosted Regression Tree) algorithm is used to predict the bias, based on the meteorology and the concentration of other pollutants in the model. We evaluate the veracity of this bias prediction again a number of independent datasets and find that the bias predictor provides a robust improvement of the model performance. By applying this bias correction to the model predictions, we can evaluate the impact of the bias corrector for important derived quantitates such as the total ozone in the atmosphere and the probability distribution. Again, the predictions of the bias corrected model appear much more consistent with observations than the model alone. This machine learning bias corrector methodology gives a new tool to link air quality forecasts with observations and so may provide improved air quality forecasts in the future.