Evan Coulson and Lauren West
Prof. Weiqing Gu
Math189Z
6 May 2020

Preliminary Findings

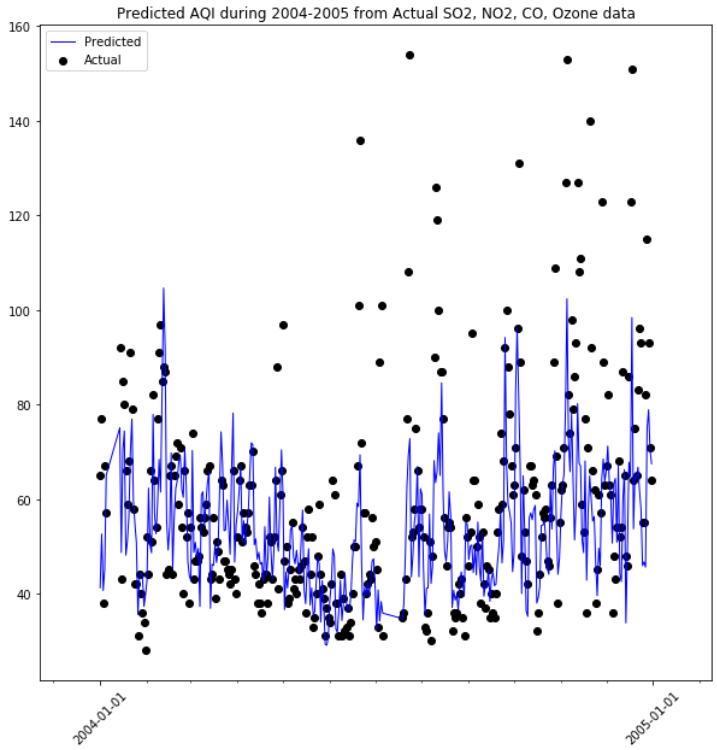
Our original idea was to see what the future would look like air quality-wise if the recent drops hadn't happened. We also want to predict future air-quality given the recent drops. We cleaned up air quality data from the US, and we are hoping to compare predictions for this year with data pre-COVID-19 and for data incorporating recent trends(aka decreases due to lack of industry pollution).

Our question: Is expected air quality for the next year higher than models informed with pre COVID 19 data? In other words, will the new air quality numbers generate a prediction with higher air quality for the Future?

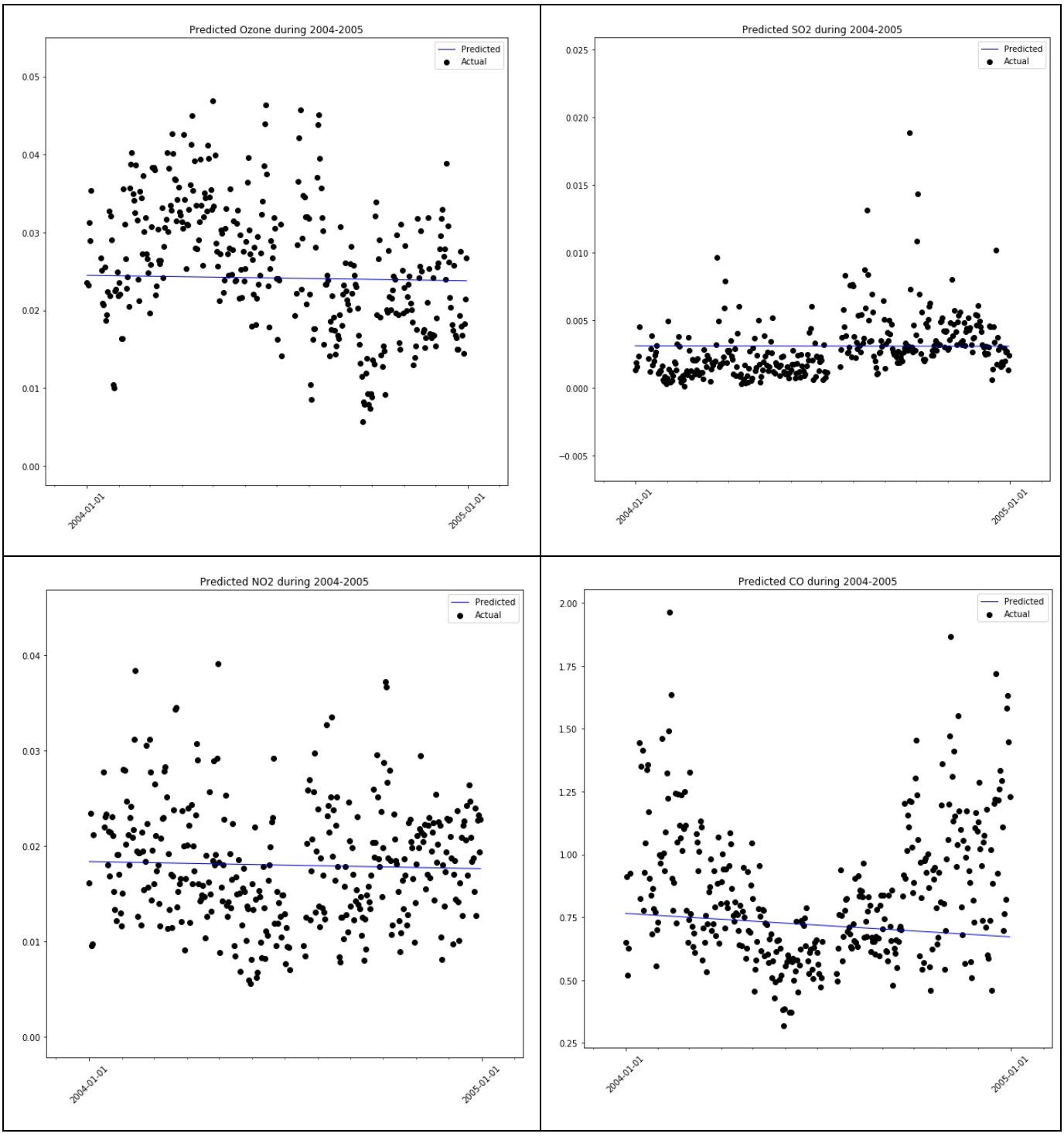
To answer our research question, we used a linear regression on the respective data sets (the data before impact due to COVID-19 and the data including and following COVID-19 impact).

For our Preliminary Results, we honed in on the situation in Seattle, WA.

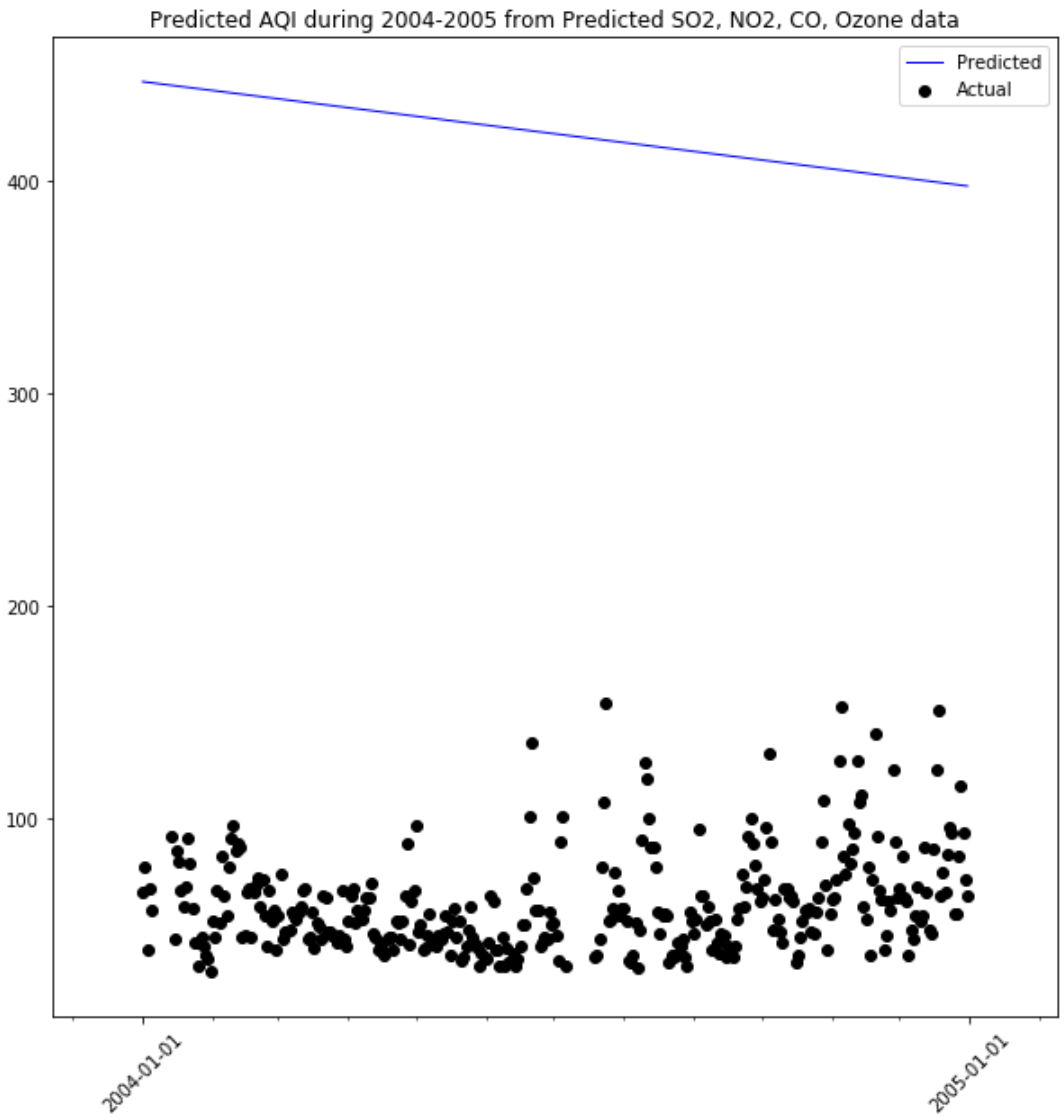
First, we created a linear regression on past data to predict the Air Quality Index, and our predictions mirrored the actual values rather well. So, we decided we could trust linear regression to predict Air Quality Index to a certain extent. These are our Preliminary Results:



Next, we ran a linear regression on ozone, CO, NO_2 and SO_2 to predict their future values based on data from the past. These are our Preliminary Results:



Obtaining those predictions, we were able to apply what we had first learned to predict the Air Quality Index for the future using linear regression. These are our Preliminary Results:



Looking Forward to the Final Project:
We will not use the same procedure in our final project. Evidently, this is nowhere near accurate. Our procedure did not account for the variation in SO2, NO2, CO, and Ozone as they vary on a day to day basis. To resolve this issue, a neural network would be helpful to predict the variance of the pollutant data, unlike a linear regression. We plan to make predictions from data not impacted by COVID-19, and then also make predictions that include more recent air quality trends. In the end, we will have a better picture of what various cities' air quality would have looked like without COVID-19 impacting human activity, and what it will look like taking into account recent trends.

SOURCES:

The data we plan to use provides historical data of air qualities of counties across the U.S.: aqs.epa.gov/aqsweb/documents/data_api.html