

MATH334 Group Project: California Pollution¹

N.Hofmann, A.Nazarovs, A.Reid, R.Harwood

February 12, 2021

¹Group 3, IDs: 34747826,34864008,34697217,34734945

What is our Data?

- From 2000-2016 pollution data was sampled in America
- Four pollutants were monitored: NO_2 , O_3 , CO and SO_2

Motivation to perform time series analysis

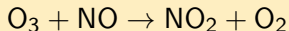
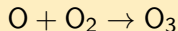
Why should we monitor the levels of pollutants?

- Studies have shown a link between pollution levels and hospital admissions related to respiratory problems
- The World Health Organisation set out guidelines to help reduce pollution including ideal levels
- We want to make sure that pollution levels are dropping and are expected to fall to or below the targets.

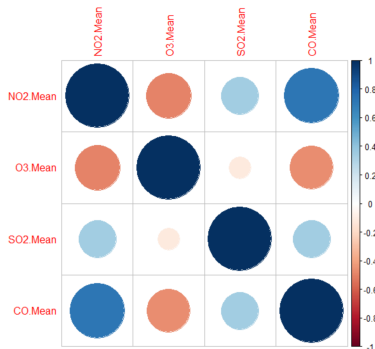
An interesting bit of Chemistry

We found that NO_2 and O_3 might be two interesting data streams to look at due to the chemical reactions which take place between them.

The Reaction Between NO_2 and O_3



Correlation!



This shows what we expect based on our chemistry knowledge, a negative correlation between NO_2 and O_3

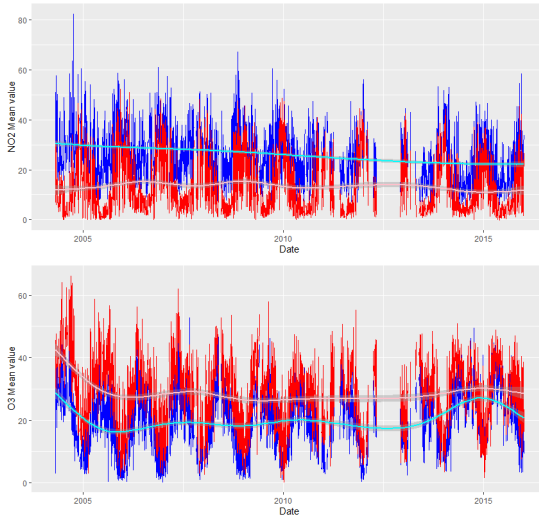
Deciding on data streams

We have already decided we specifically what to focus on NO_2 and O_3 , but where should we focus on?

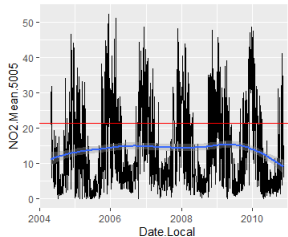
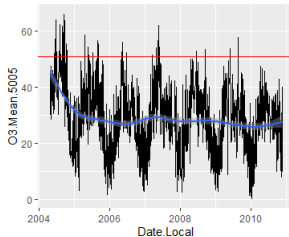
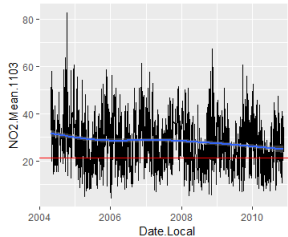
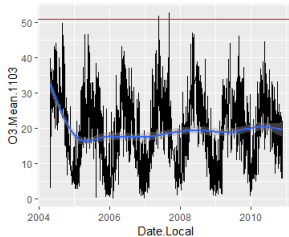
Our original data was for California so we looked there.

We tried to find sites with a lot of data, we settled on LA because it had two sites with lots of data.

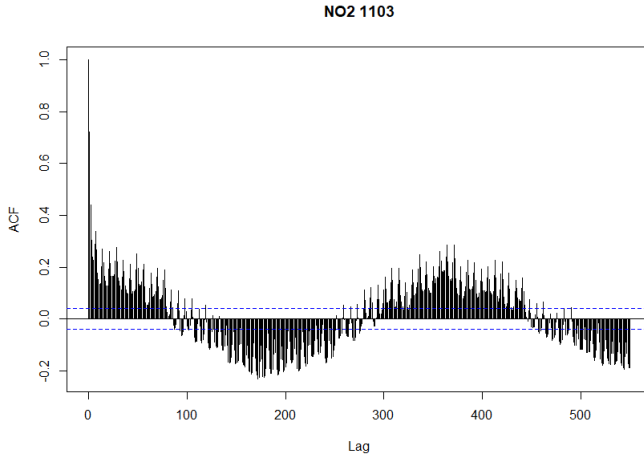
Comparing two sites: 1103 & 5005



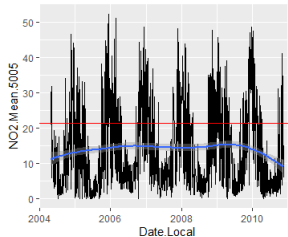
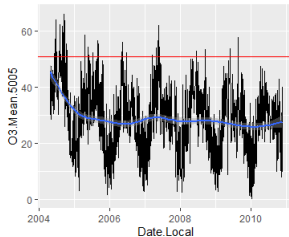
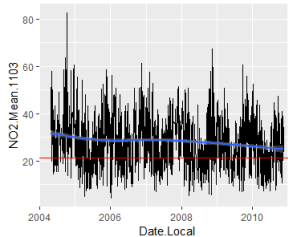
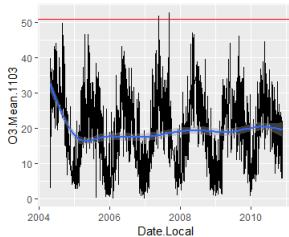
The 4 Time Series



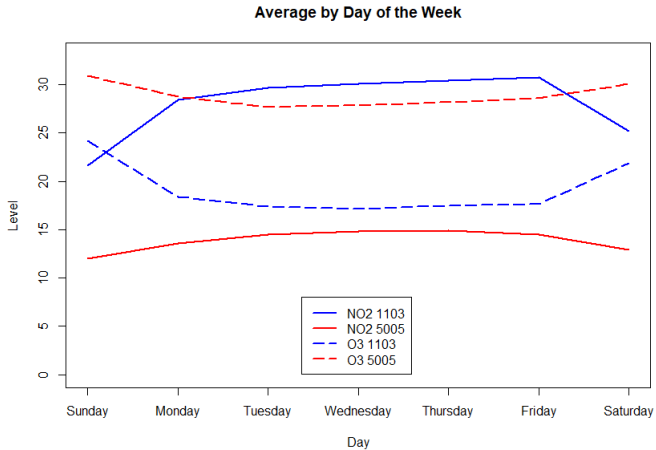
Double Seasonality



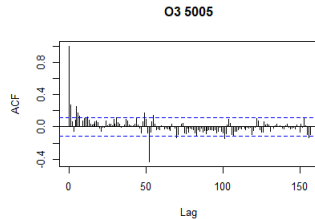
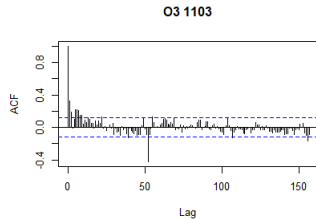
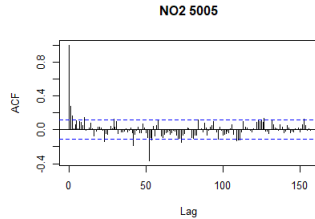
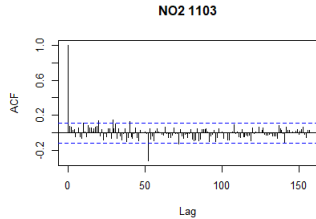
The 4 Time Series



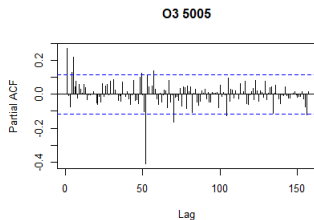
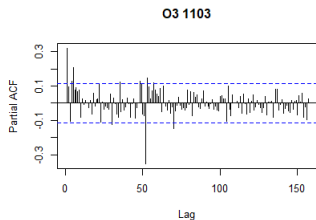
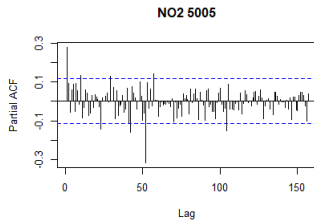
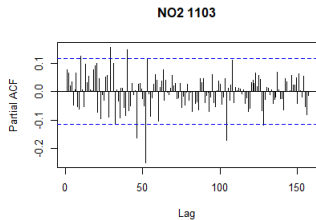
An Interesting Result



ACF Models



PACF Models



An AIC Matrix Example: O3 site 1103

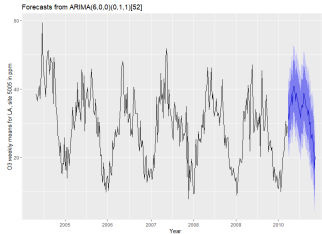
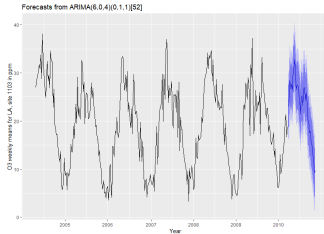
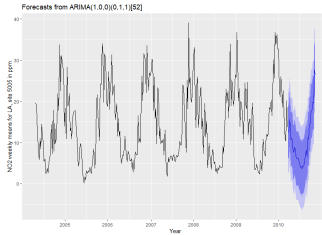
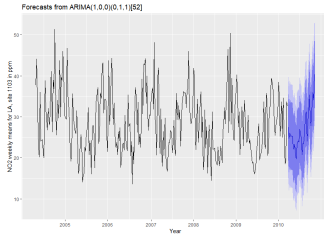
	MA=0	MA=1	MA=2	MA=3	MA=4	MA=5
AR=0	-2320.748	-2350.851	-2367.295	-2365.706	-2364.071	-2365.614
AR=1	-2364.550	-2373.273	-2375.469	-2376.438	-2377.151	-2378.189
AR=2	-2369.534	-2367.756	-2369.594	-2375.475	-2377.280	-2376.205
AR=3	-2367.551	-2365.860	-2373.049	-2377.041	-2376.203	-2374.242
AR=4	-2370.119	-2378.765	-2377.892	-2376.099	-2373.420	-2373.954
AR=5	-2375.646	-2378.004	-2376.200	-2374.570	-2372.640	-2373.963
AR=6	-2376.132	-2376.004	-2374.454	-2375.072	-2379.760	-2371.881

The models we chose

Models

- NO₂ 1103 SARIMA(1,0,0, 0,1,1, 52)
- NO₂ 5005 SARIMA (1,0,0, 0,1,1, 52)
- O₃ 1103 SARIMA(6,0,4, 0,1,1, 52)
- O₃ 5005 SARIMA(6,0,0, 0,1,1, 52)

Forecasting



Any Questions?