Air Pollution and Climate

MIDN 1/C Tyler Courtney, 2/C Patrick Gallagher, 2/C Tiffany Johnstone

Presentation Agenda

1. Problem Identification
2. Data
3. Assumptions
4. Models
5. Graphics
6. Conclusion
7. Recommendations
8. Future Work
9. Questions

Air Pollution and Why It Matters

- Large concentrations of air pollution can cause:
  - Diseases
  - Contaminated Water Supply
  - Contaminated Soil
  - Inhabitable Ecosystems
- Air pollution is mostly produced by everyday human actions:
  - Vehicle Emissions
  - Power Production
  - Mineral Refining
  - All other incomplete combustion reactions
- Without keeping our actions in check humanity faces long term effects unless sustainable solutions are implemented to reduce air pollution.

Our Approach to Reducing Air Pollution

Our group decided to see the effect that a location's climate had on the number of air pollutants. If we can figure out which areas are susceptible to larger amounts of air pollution we can focus our time and resources in those areas first. While air pollution is a problem globally we need to start somewhere and our models will help determine which areas are at the greatest risk.

Initial Data

- The data comprised of 4 primary pollutant compounds:
  - NO_2 (Nitrogen Dioxide)
  - O_3 (Ground Level Ozone)
  - SO_2 (Sulfur Dioxide)
  - CO (Carbon Monoxide)
- Each compound had the following information included
  - Mean compound amount
  - Max compound amount
  - Time the max compound amount was collected
- The data was collected daily from the years 2000-2010
- The data contained 43,895 entries

Data Used In Our Model

From the initial data set we used the *mean* values for NO2 and O3. We decided to focus on these two compounds specifically because of the oxidation reaction that connects both of the compounds. We constricted the data to Washington D.C between 2006 and 2010 in order to isolate variables and include additional data to supplement our model.

Additional Data

- We used monthly average temperatures from 2006-2010 to account for the effect of climate on air pollutants
- We also added several high volume combustion reactions that we wanted to control for:
  - DC vehicle traffic
  - DC air traffic
  - DC petroleum power production
- We restricted our data from 2006-2010 because those were the only years we could find traffic data for. This left us with approximately 15,000 data points

Assumptions

- ❖ Uniform distribution of vehicle traffic throughout the year (Year to month conversion)
- ❖ Power plant production initially had an Uniform distribution assumption until we added a weighted scale.
- ❖ Power plant production was adjusted up 15% from NOV-FEB and down 15% from MAY-AUG
- ❖ Chemical compound daily means were averaged together to get a monthly mean

Base Models

We used linear regression models because our data points followed a linear pattern across the time period.

Monthly Mean NO2 = AverageTemp + TotalAir + MegaWattHours + DCTraffic

Monthly Mean O3 = AverageTemp + TotalAir + MegaWattHours + DCTraffic

NO2 Model

Monthly Mean NO2 = -1.15E9(AverageTemp) -1.87E6(TotalAir) -1.12E6(MegaWattHours) +1.22E5(DCTraffic) +2.81E11

```
. reg MonthlyNO2 AverageTemp TotalAir AdjMegaWattHours DCTraffic
```

Source	SS	df	MS
Model	3.2000e+22	4	8.0000e+21
Residual	1.8785e+22	55	3.4155e+20
Total	5.0785e+22	59	8.6076e+20

Number of obs	=	60
F(4, 55)	=	23.42
Prob > F	=	0.0000
R-squared	=	0.6301
Adj R-squared	=	0.6032
Root MSE	=	1.8e+10

MonthlyNO2	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
AverageTemp	-1.15e+09	1.77e+08	-6.50	0.000	-1.51e+09	-7.96e+08
TotalAir	-1878725	930808.3	-2.02	0.048	-3744106	-13343.38
AdjMegaWattHours	-1129987	611694.5	-1.85	0.070	-2355850	95875.98
DCTraffic	122040.8	51615.62	2.36	0.022	18600.77	225480.8
_cons	2.81e+11	6.12e+10	4.59	0.000	1.58e+11	4.04e+11

NO2 Model Robust

Monthly Mean NO2 = -1.15E9(AverageTemp) -1.87E6(TotalAir) -1.12E6(MegaWattHours) +1.22E5(DCTraffic) +2.81E11

```
.
. reg MonthlyNO2  AverageTemp TotalAir AdjMegaWattHours DCTraffic, vce(robust)

Linear regression                               Number of obs     =         60
                                                F(4, 55)          =      33.99
                                                Prob > F          =     0.0000
                                                R-squared         =     0.6301
                                                Root MSE          =    1.8e+10
```

MonthlyNO2	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
AverageTemp	-1.15e+09	1.97e+08	-5.84	0.000	-1.55e+09	-7.56e+08
TotalAir	-1878725	750495.5	-2.50	0.015	-3382751	-374698.2
AdjMegaWattHours	-1129987	509721.9	-2.22	0.031	-2151493	-108481.7
DCTraffic	122040.8	46125.69	2.65	0.011	29602.84	214478.7
_cons	2.81e+11	4.66e+10	6.03	0.000	1.87e+11	3.74e+11

O3 Model

Monthly Mean O3 = 4397.77(AverageTemp) -2.82(TotalAir) +2.89(MegaWattHours) -.19(DCTraffic)+239006.3

```
.
. reg Monthly03  AverageTemp TotalAir AdjMegaWattHours DCTraffic
```

Source	SS	df	MS
Model	2.5807e+11	4	6.4516e+10
Residual	1.7435e+11	55	3.1699e+09
Total	4.3241e+11	59	7.3290e+09

Number of obs	=	60
F(4, 55)	=	20.35
Prob > F	=	0.0000
R-squared	=	0.5968
Adj R-squared	=	0.5675
Root MSE	=	56302

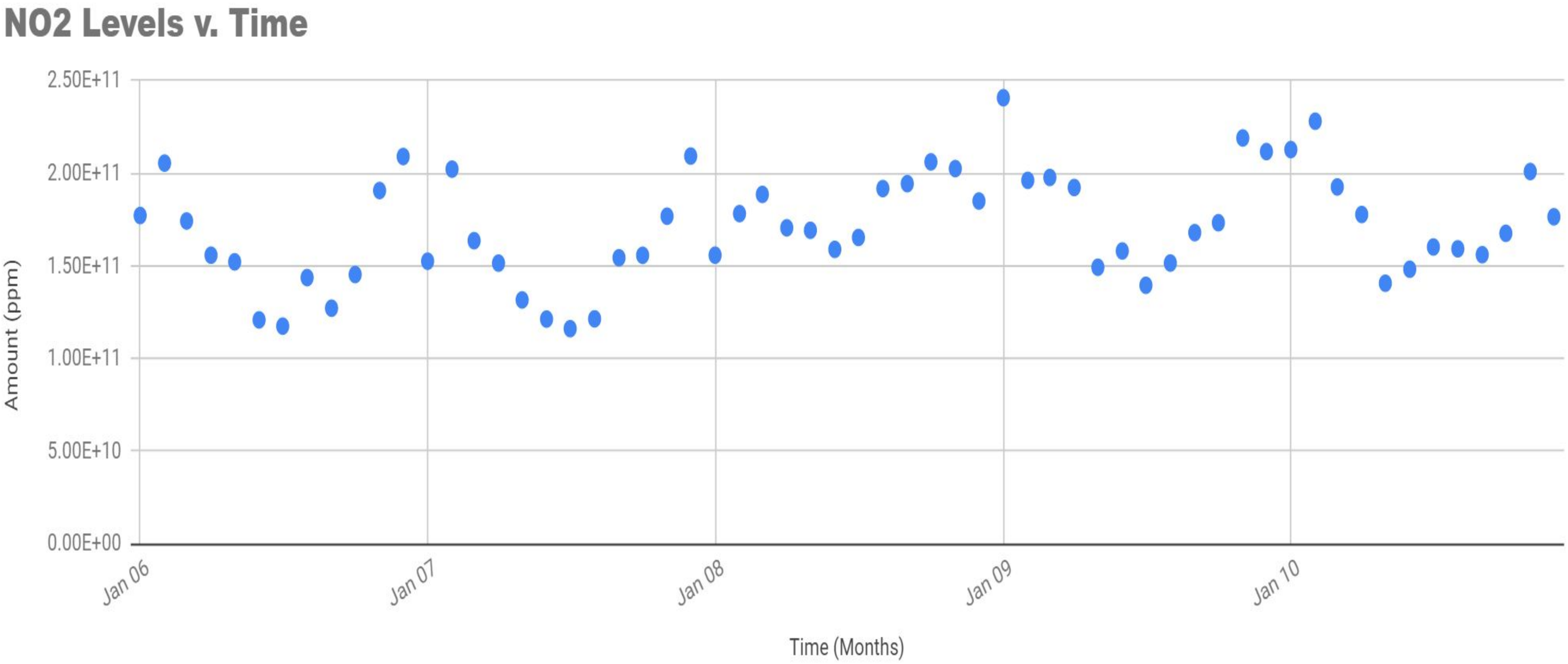
Monthly03	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
AverageTemp	4397.775	539.391	8.15	0.000	3316.812	5478.739
TotalAir	-2.823488	2.835712	-1.00	0.324	-8.506382	2.859405
AdjMegaWattHours	2.899429	1.86353	1.56	0.125	-.8351694	6.634027
DCTraffic	-.1916077	.1572472	-1.22	0.228	-.5067382	.1235228
_cons	239006.3	186517.1	1.28	0.205	-134782.3	612794.8

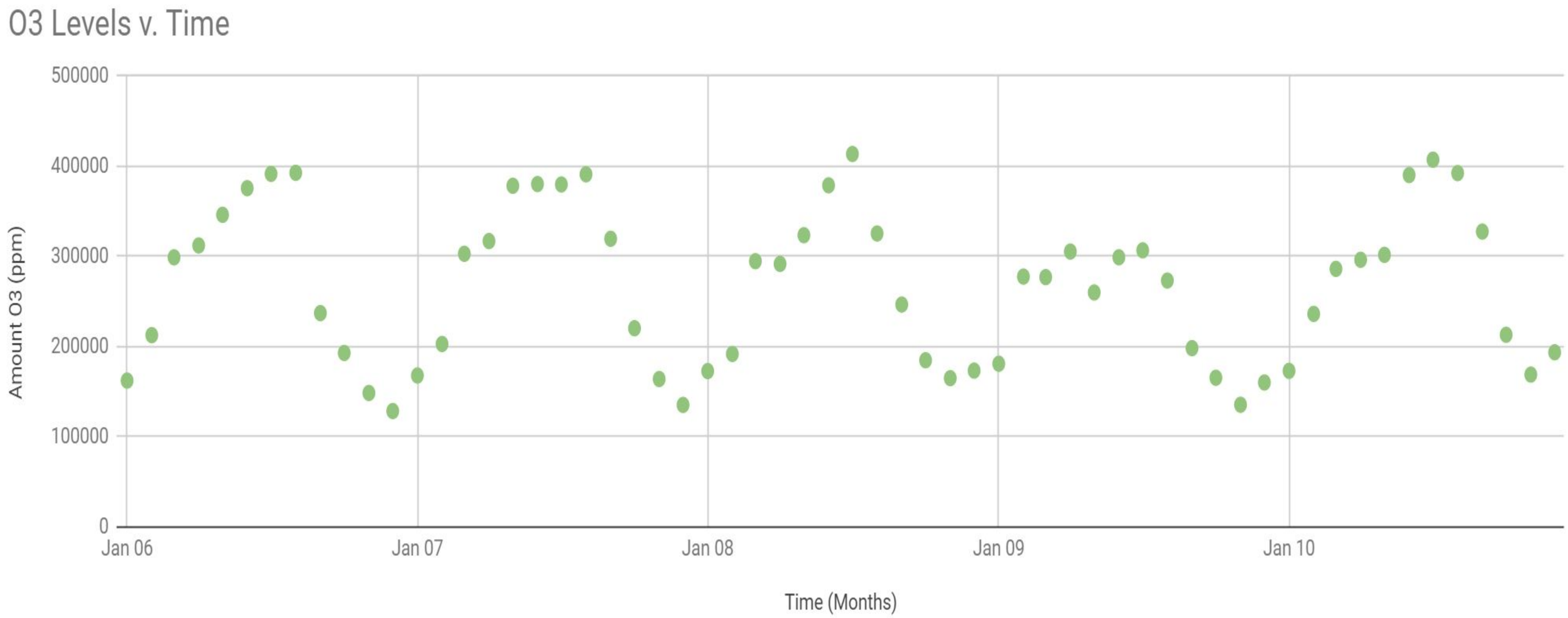
O3 Model Robust

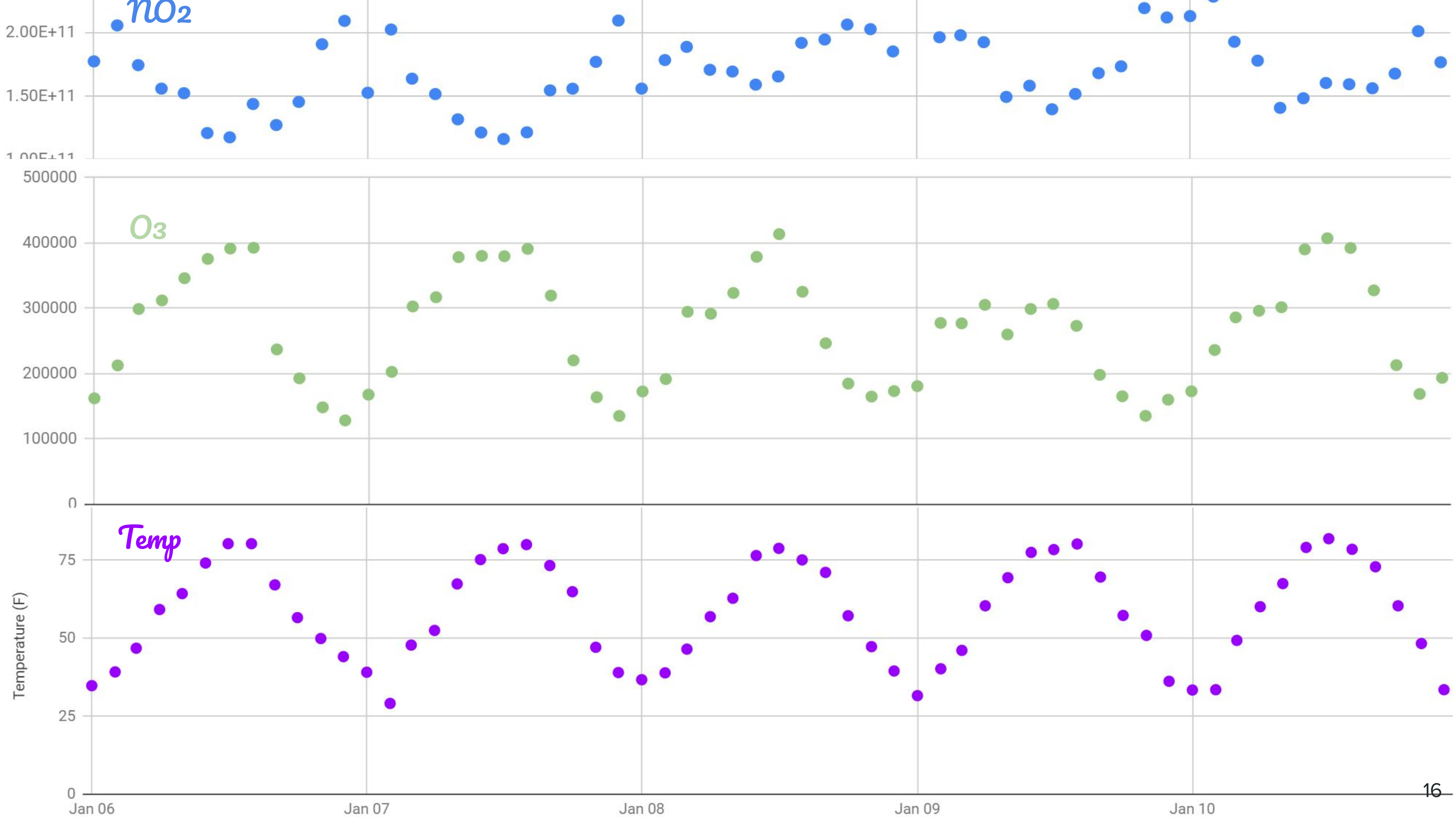
Monthly Mean O3 = 4397.77(AverageTemp) -2.82(TotalAir) +2.89(MegaWattHours) -.19(DCTraffic)+239006.3

```
. reg MonthlyO3  AverageTemp TotalAir AdjMegaWattHours DCTraffic, vce(robust)

Linear regression                               Number of obs     =         60
                                                F(4, 55)          =      31.90
                                                Prob > F          =     0.0000
                                                R-squared         =     0.5968
                                                Root MSE          =      56302
```

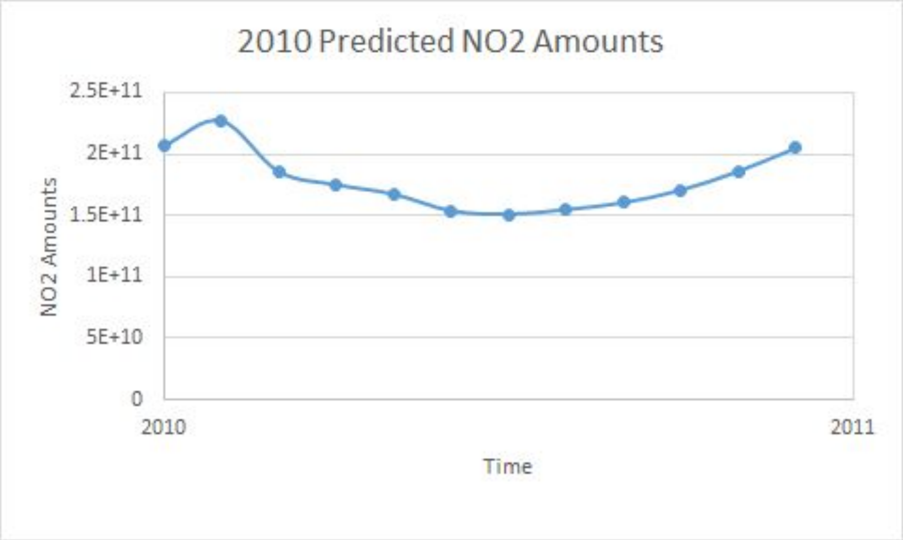
MonthlyO3	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
AverageTemp	4397.775	458.975	9.58	0.000	3477.969	5317.582
TotalAir	-2.823488	2.748788	-1.03	0.309	-8.332182	2.685205
AdjMegaWattHours	2.899429	1.936861	1.50	0.140	-.982127	6.780984
DCTraffic	-.1916077	.1425702	-1.34	0.184	-.4773248	.0941094
_cons	239006.3	160924.3	1.49	0.143	-83493.14	561505.7


NO2 Levels v. Time
2.50E+11
2.00E+11
1.50E+11
1.00E+11
5.00E+10
0.00E+00
Amount (ppm)
Jan 06
Jan 07
Jan 08
Jan 09
Jan 10
Time (Months)


O3 Levels v. Time
500000
400000
300000
200000
100000
0
Amount O3 (ppm)
Jan 06
Jan 07
Jan 08
Jan 09
Jan 10
Time (Months)


NO2
2.00E+11
1.50E+11
1.00E+11
500000
O3
400000
300000
200000
100000
0
Temp
75
50
25
0
Temperature (F)
Jan 06
Jan 07
Jan 08
Jan 09
Jan 10

2010 Predictive Modeling Comparison



Conclusions

Our model shows that there is a relationship between the amount of heat an area gets and the amount of pollutants that are present. We can follow the model and the graphics which show that as an area gets warmer we will see an increase in the amount of ozone and a decrease in nitrogen dioxide. The reason this occurs is because the bond between the nitrogen and oxygen breaks easier at higher temperatures. This creates a lone oxygen atom and nitrogen oxide. The lone oxygen reacts with O_2 to create O_3. This falls in line with the results we obtained.

Recommendations

- ❖ Areas in hotter climates should receive aid and install new policies first.
- ❖ NO2 should be the focal point of our air pollution reduction programs
- ❖ NO2 is a by product of incomplete combustion reactions which means some possible solutions to reduce those could be:
  - ➢ Federal Emission Testing Policies (15-25$ per test)
  - ➢ Selective Catalytic Reduction technology required for all Diesel Vehicles (400-600$ Per system)
  - ➢ Selective Catalytic Reduction technology required for all merchant vessels (700,000$ for largest ship engines)
  - ➢ Zero emission electric vehicle tax breaks
  - ➢ Adoption of Ultra Low Emission Zones to help pay for reduction costs

Future Analysis

- ❖ Obtain more precise vehicle data
- ❖ Obtain population density data for a year by year basis
- ❖ Find similar data in different locations to reproduce our results
- ❖ Obtain monthly data about power plant production
- ❖ Find data on more combustion reactions to control for

Questions?

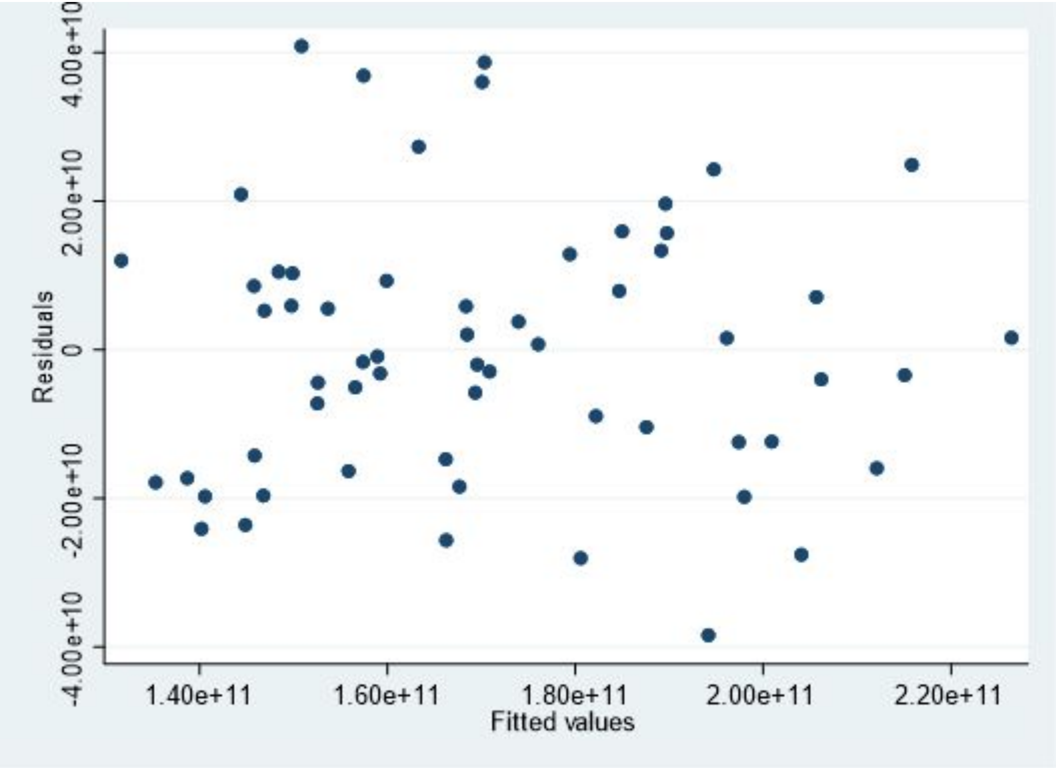
Correlation Test

	Mont-NO2	Monthl-3	Mont-SO2	Monthl-O	Averag-p	RRairt-c	Dulles-c	DCTraf-c	MegaWa-s	AdjMeg-s	timest-a	TotalAir
MonthlyNO2	1.0000											
MonthlyO3	-0.6384	1.0000										
MonthlySO2	0.4441	-0.2682	1.0000									
MonthlyCO	0.1323	-0.2872	0.4338	1.0000								
AverageTemp	-0.7030	0.7583	-0.5399	-0.2359	1.0000							
RRairtraffic	-0.5282	0.3405	-0.2862	0.0313	0.5545	1.0000						
Dullesairt-c	-0.5335	0.2196	-0.1202	0.2402	0.3269	0.5794	1.0000					
DCTraffic	0.2704	0.0059	-0.2973	-0.2908	0.0064	-0.1409	-0.5617	1.0000				
MegaWattHo-s	0.0015	0.1426	-0.3326	-0.1071	0.0058	-0.0728	-0.1816	0.5931	1.0000			
AdjMegaWat-s	0.1322	-0.0299	-0.2458	-0.0674	-0.1673	-0.1807	-0.2467	0.5783	0.9751	1.0000		
timestatdata	0.2908	-0.0626	-0.2992	-0.2597	0.0555	-0.0954	-0.5502	0.9666	0.4468	0.4356	1.0000	
TotalAir	-0.5879	0.2848	-0.1906	0.1928	0.4405	0.7868	0.9589	-0.4744	-0.1629	-0.2497	-0.4498	1.0000

NO2 Tests

```
. ovtest

Ramsey RESET test using powers of the fitted values of MonthlyNO2
       Ho:  model has no omitted variables
                  F(3, 52) =      1.11
                  Prob > F =      0.3537

. vif

    Variable |       VIF       1/VIF
-------------+----------------------
   DCTraffic |      2.10    0.475561
    TotalAir |      1.78    0.562350
AdjMegaWat~s |      1.62    0.618846
 AverageTemp |      1.44    0.694556
-------------+----------------------
    Mean VIF |      1.73

. hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of MonthlyNO2

         chi2(1)      =     0.03
         Prob > chi2  =   0.8727

. dwstat

Durbin-Watson d-statistic(  5,    60) =  1.491261
```

NO2 Residuals plot



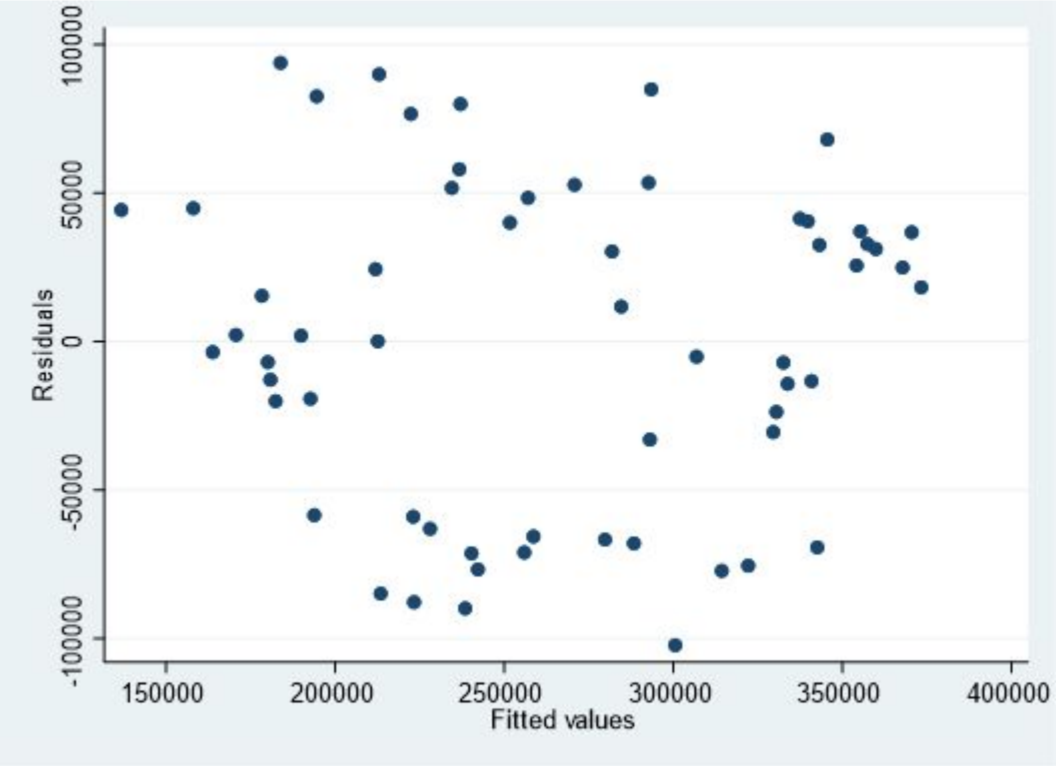
O3 Tests

```
. ovtest

Ramsey RESET test using powers of the fitted values of MonthlyO3
       Ho:  model has no omitted variables
                 F(3, 52) =       1.64
                 Prob > F =       0.1906

. vif
```

Variable	VIF	1/VIF
DCTraffic	2.10	0.475561
TotalAir	1.78	0.562350
AdjMegaWat~s	1.62	0.618846
AverageTemp	1.44	0.694556
Mean VIF	1.73	

```
. hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of MonthlyO3

         chi2(1)      =     0.48
         Prob > chi2  =   0.4893

. dwstat

Durbin-Watson d-statistic(  5,    60) =  .5891645
```

O3 Residuals Plot



Data Sources

Air Traffic

https://www.mwaa.com/about/reagan-air-traffic-statistics

https://www.mwaa.com/about/dulles-air-traffic-statistics

Weather

https://w2.weather.gov/climate/xmacis.php?wfo=lwx

Vehicle Traffic

https://opendata.dc.gov/datasets/2006-traffic-volume?page=4

https://opendata.dc.gov/datasets/2010-traffic-volume

Power Production

https://www.eia.gov/state/search/#?1=101&2=188

https://www.eia.gov/state/?sid=DC#tabs-1

Recommendation Sources

Naval SCR Tech

https://nepis.epa.gov/Exe/ZyNET.exe/P100GPCR.txt?ZyActionD=ZyDocument&Client=EPA&Index=2006%20Thru%202010&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&UseQField=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=&File=D%3A%5CZYFILES%5CINDEX%20DATA%5C06THRU10%5CTXT%5C00000033%5CP100GPCR.txt&User=ANONYMOUS&Password=anonymous&SortMethod=h%7C-&MaximumDocuments=1&FuzzyDegree=0&ImageQuality=r75g8/r75g8/x150y150g16/i425&Display=hpfr&DefSeekPage=x&SearchBack=ZyActionL&Back=ZyActionS&BackDesc=Results%20page&MaximumPages=1&ZyEntry=50

Vehicle SCR Tech

https://theicct.org/sites/default/files/publications/ICCT_costs-emission-reduction-tech-HDV_20160229.pdf

Ultra Low Emission Zone

https://tfl.gov.uk/modes/driving/ultra-low-emission-zone