Analysis of 2000-2010 Air Quality Trends in DMV Area

February 29, 2020

Dataset: New Pollution (Level 5)

Team: DC20052

Team Members: Nina Wang, Wanyun Yang, Bojing Fan

Mentor: Jennifer Chaffee

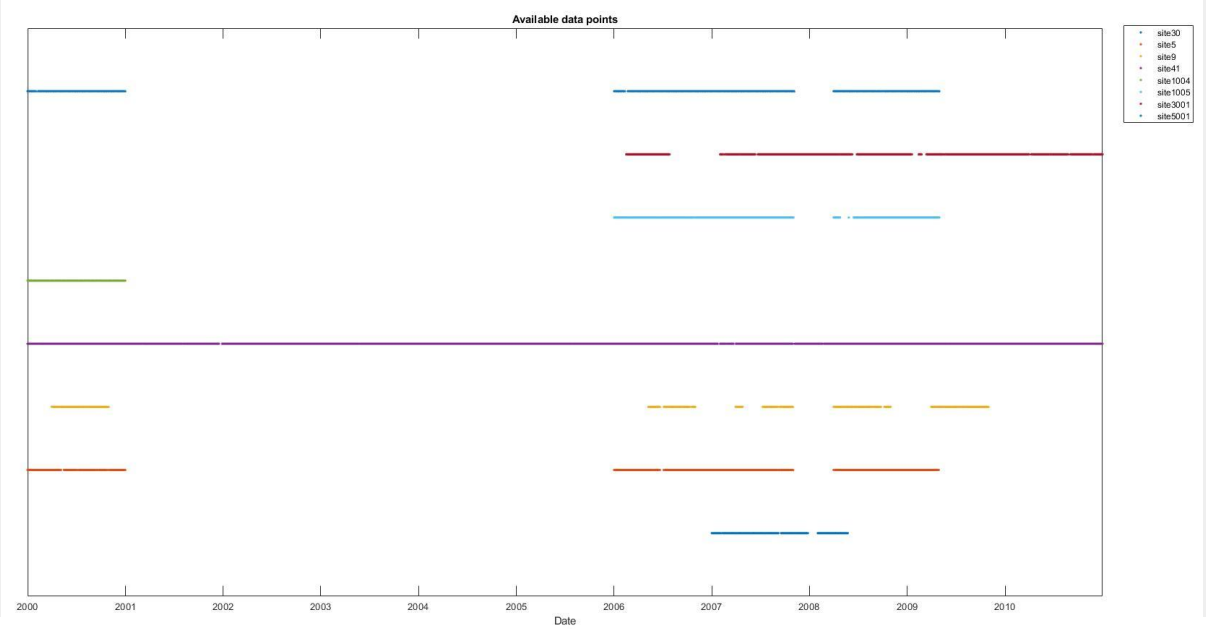
DATA CHALLENGE

Source: Picture from American Lung Association (2020)
https://www.lung.org/our-initiatives/healthy-air/sota/city-rankings/states/district-of-columbia/district-of-columbia.html

Dataset Overview

Dataset Overview

- The provided dataset contained 43,894 rows, spanning from 1/1/2000 to 12/31/2010.



Dataset Overview

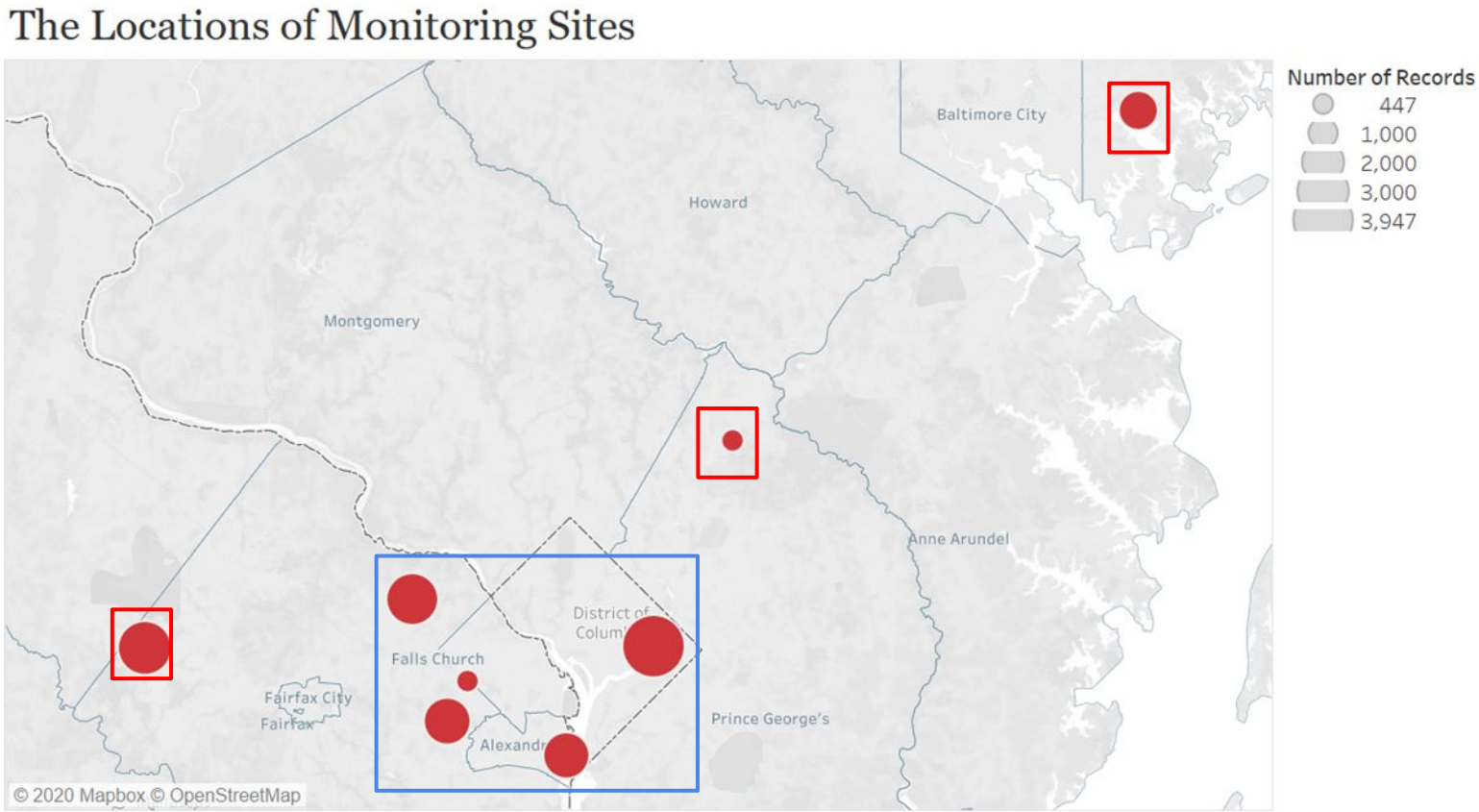
State	County.	Site.	Date.Local	NO2.Max.Value	NO2.Max.Hour	O3.Max.Value	O3.Max.Hour	SO2.Max.Value	SO2.Max.Hour	CO.Max.Value	CO.Max.Hour
11	1	41	1/1/2000	39	17	0.014	23	33	15	4.1	3
11	1	41	1/1/2000	39	17	0.014	23	33	15	3.6	5
11	1	41	1/1/2000	39	17	0.014	23	19	17	4.1	3
11	1	41	1/1/2000	39	17	0.014	23	19	17	3.6	5
11	1	41	1/2/2000	22	0	0.029	11	11	0	2.6	1
11	1	41	1/2/2000	22	0	0.029	11	11	0	2.1	2
11	1	41	1/2/2000	22	0	0.029	11	11	2	2.6	1
11	1	41	1/2/2000	22	0	0.029	11	11	2	2.1	2

Pollutant Standard	Date Local	Arithmetic Mean	1st Max Value	1st Max Hour	AQI
CO 1-hour 1971	1/1/2000	2.0625	4.1	3	
CO 8-hour 1971	1/1/2000	2.005263	3.6	5	41

Dataset Overview: Data collection

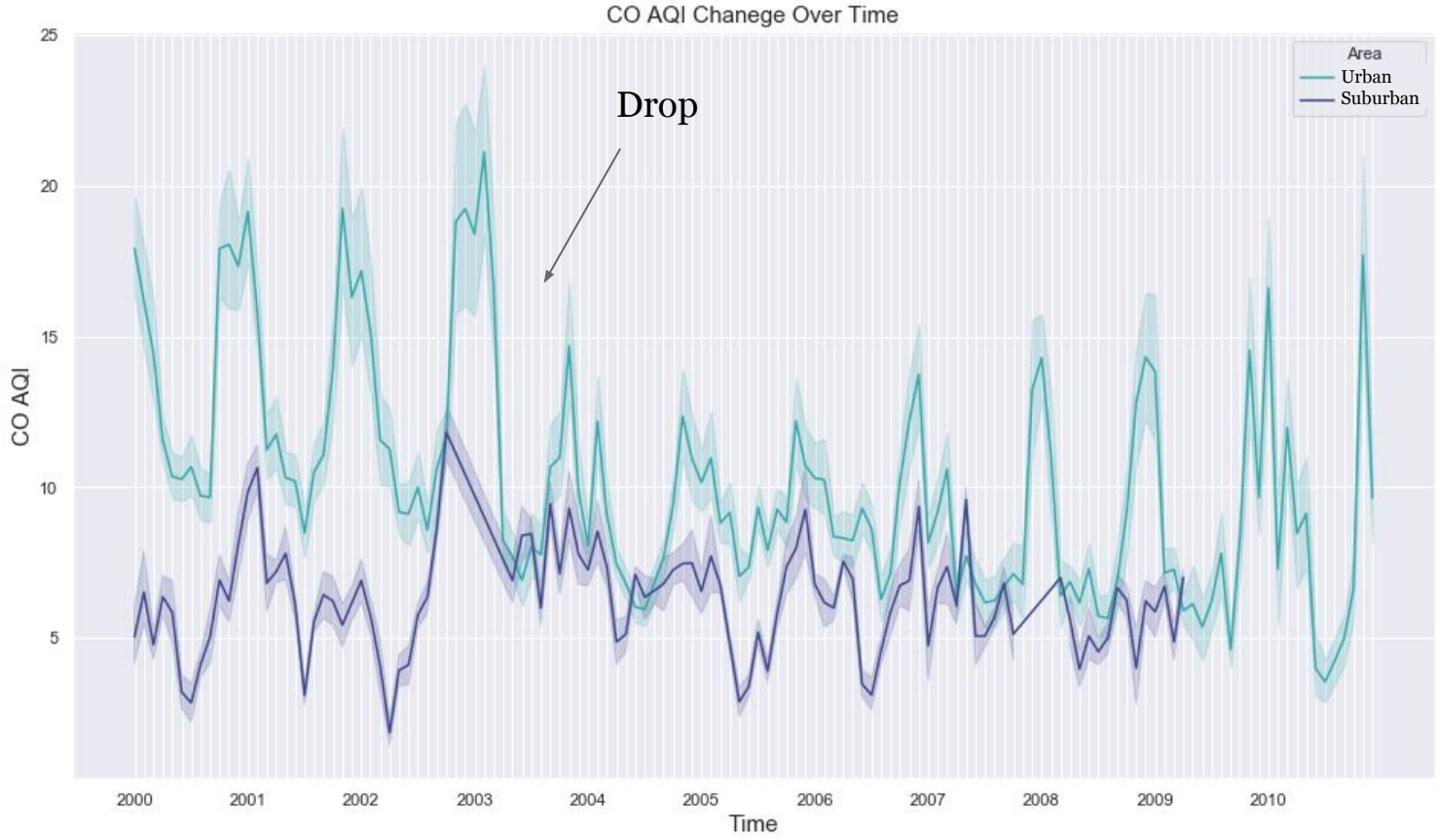
- The collected dataset contained 16,320 rows, spanning from 1/1/2000 to 12/31/2010.
- Additionally, daily average wind speed and precipitation data were retrieved from NOAA.
- Only records with observations of all four pollutants were included, which left us 10 sites in the area.

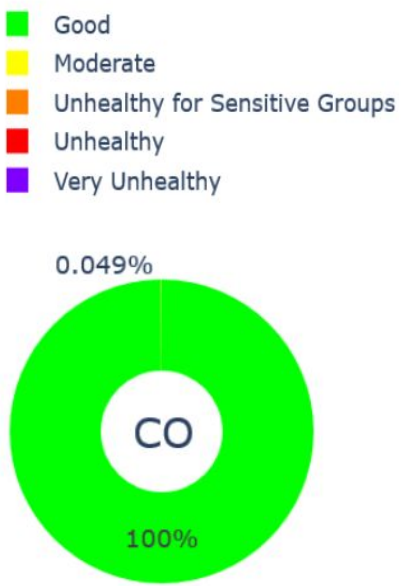
Dataset Overview: Collect raw data from EPA



Time Series Analysis I

Carbon Monoxide





Method	level	trend	seasonality	comment
Simple exponentioal smoothing	V			
Double (Holt's) exponential smoothing	V	V		
Triple (Holt-Winter's) exponential smoothing	V	V	V	
Regression	V	V	V	Assumes stationarity

Since our data has obvious seasonality, we tried Winter's Method:

$$F_{t+n} = L_t + nT_t S_{t+n}$$

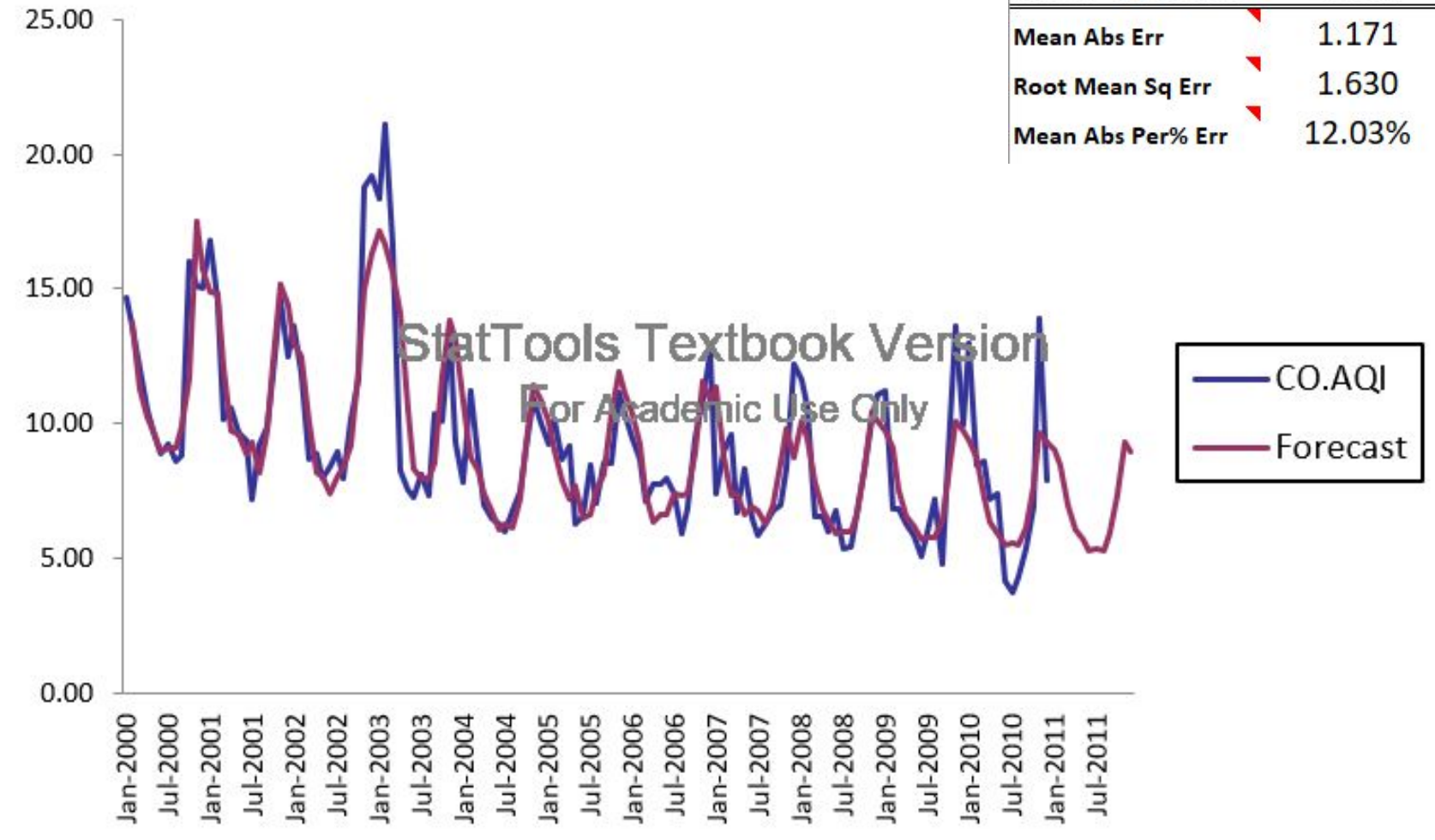
$$L_t = \alpha \frac{Y_t}{S_{t-M}} + (1-\alpha)(L_{t-1} + T_{t-1})$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$$

$$S_t = \gamma \frac{Y_t}{L_t} + (1-\gamma)S_{t-M}$$

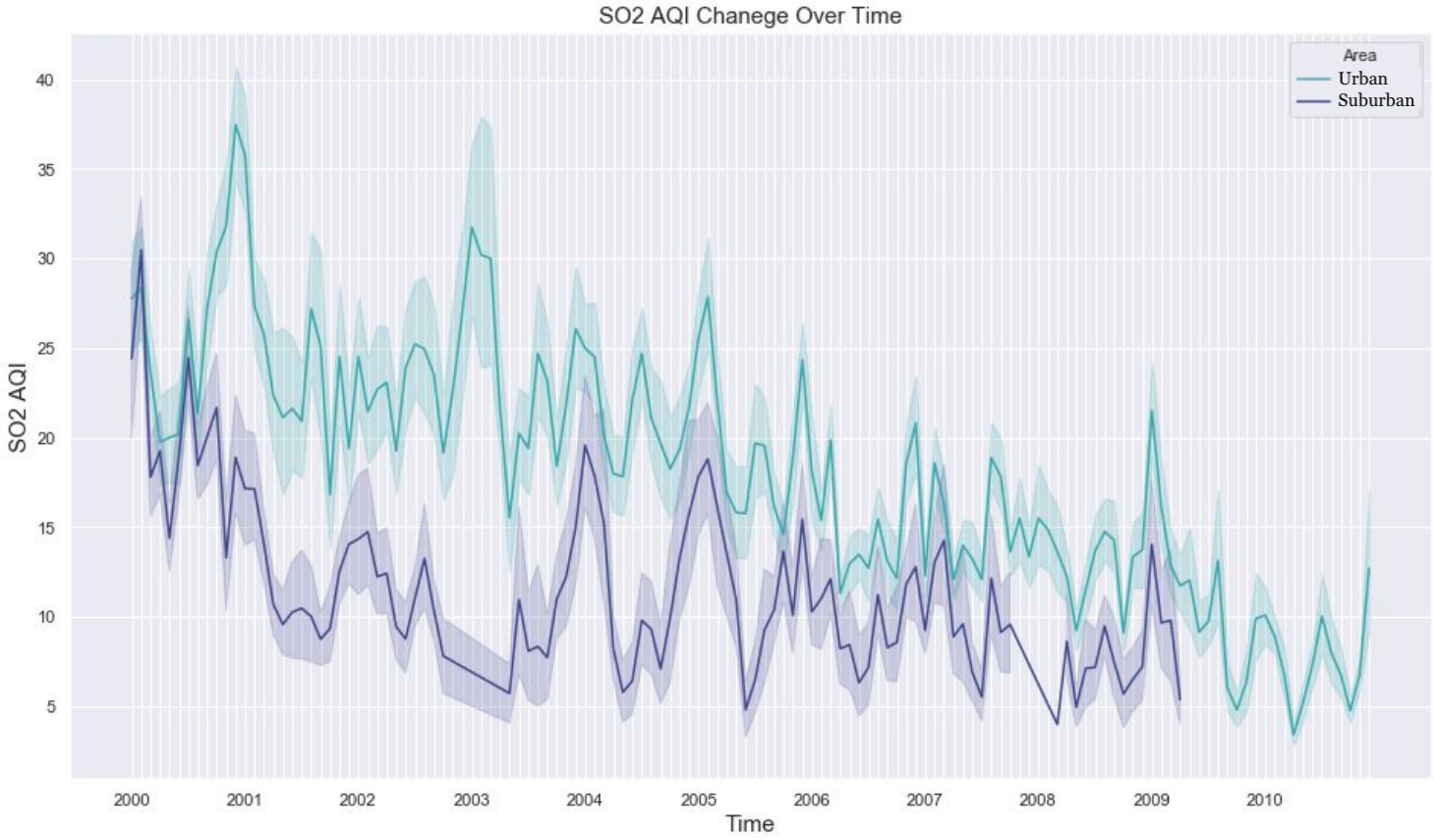
Forecast and Original Observations

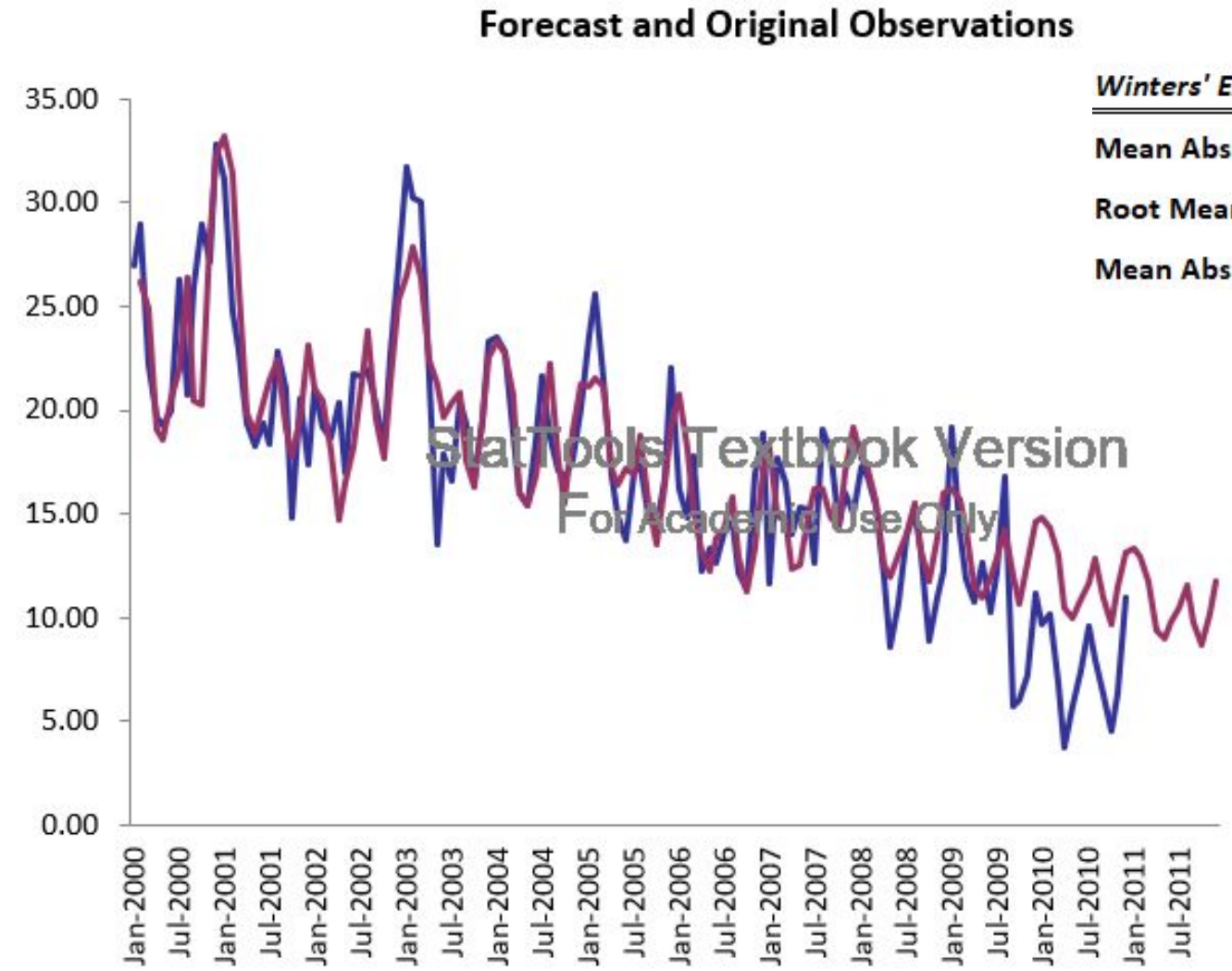
Winters' Exponential	Estimation Period	Holdouts Period
Mean Abs Err	1.171	1.127
Root Mean Sq Err	1.630	1.490
Mean Abs Per% Err	12.03%	14.90%



StatTools Textbook Version
For Academic Use Only

Sulfur Dioxide

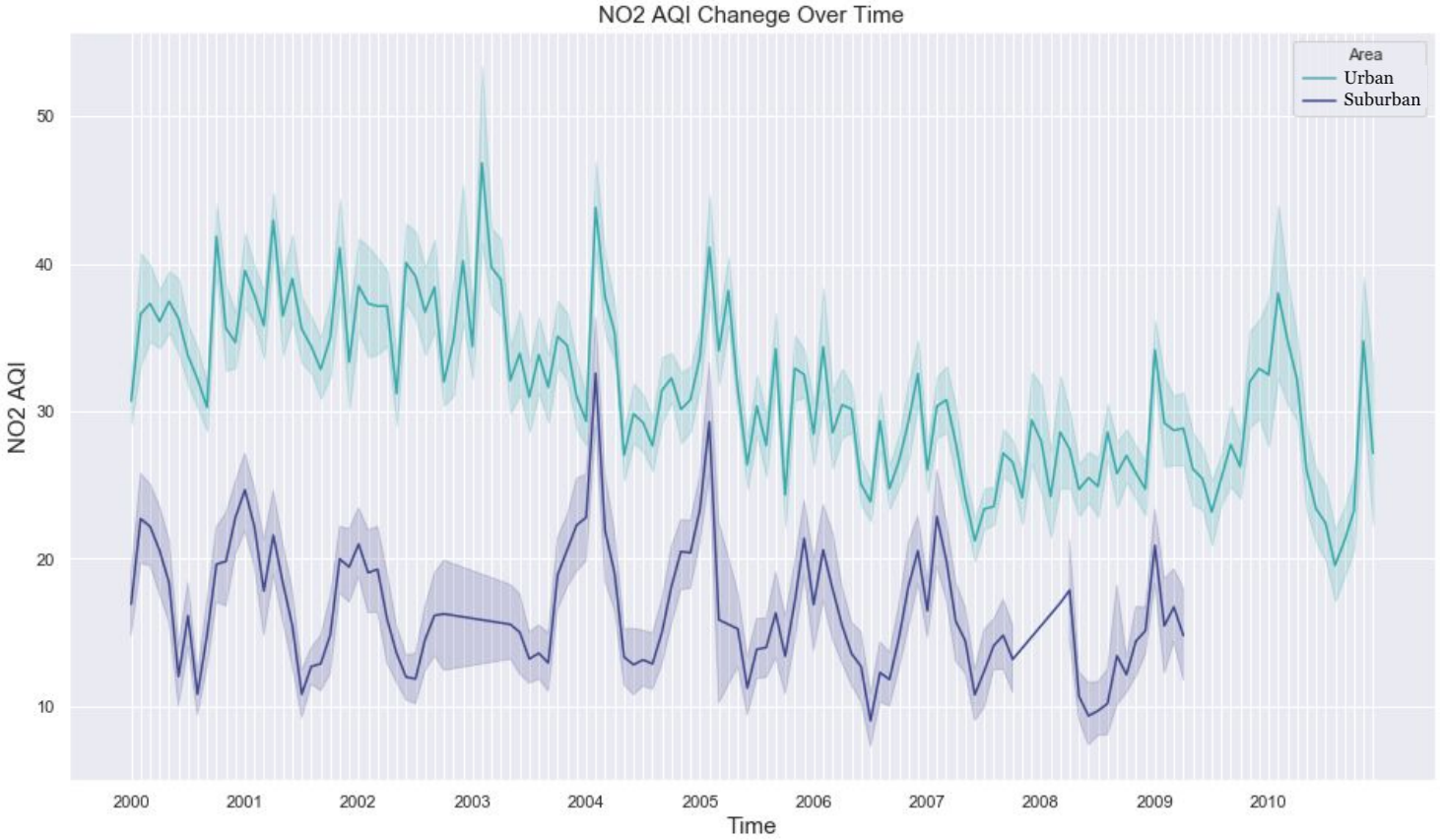


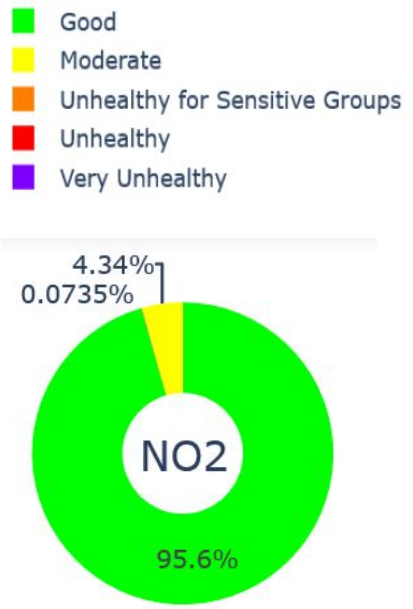


Winters' Exponential	Estimation Period	Holdouts Period
Mean Abs Err	2.039	2.918
Root Mean Sq Err	2.764	3.511
Mean Abs Per% Err	10.85%	39.33%

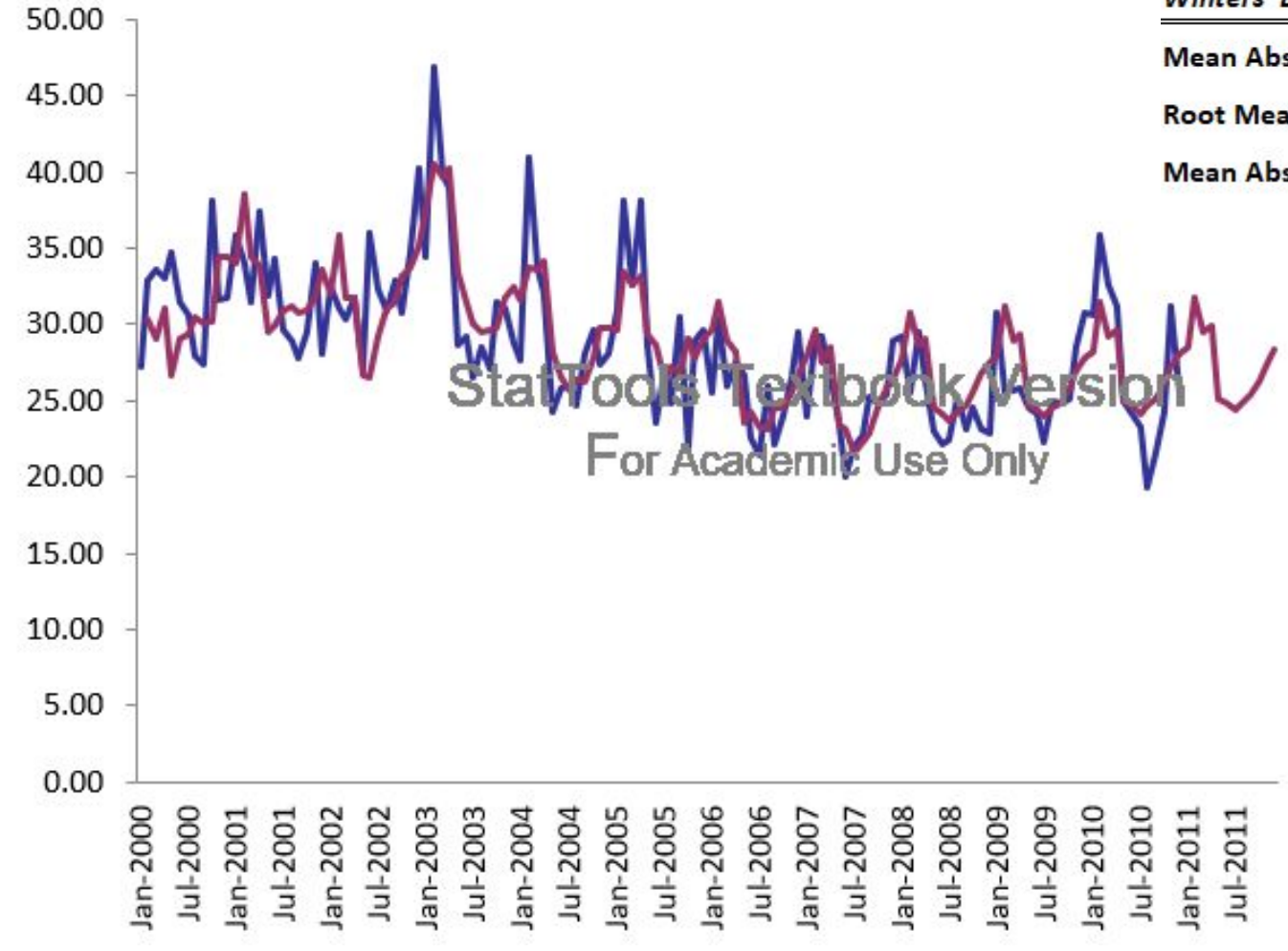
Relatively big

Nitrogen Dioxide





Forecast and Original Observations

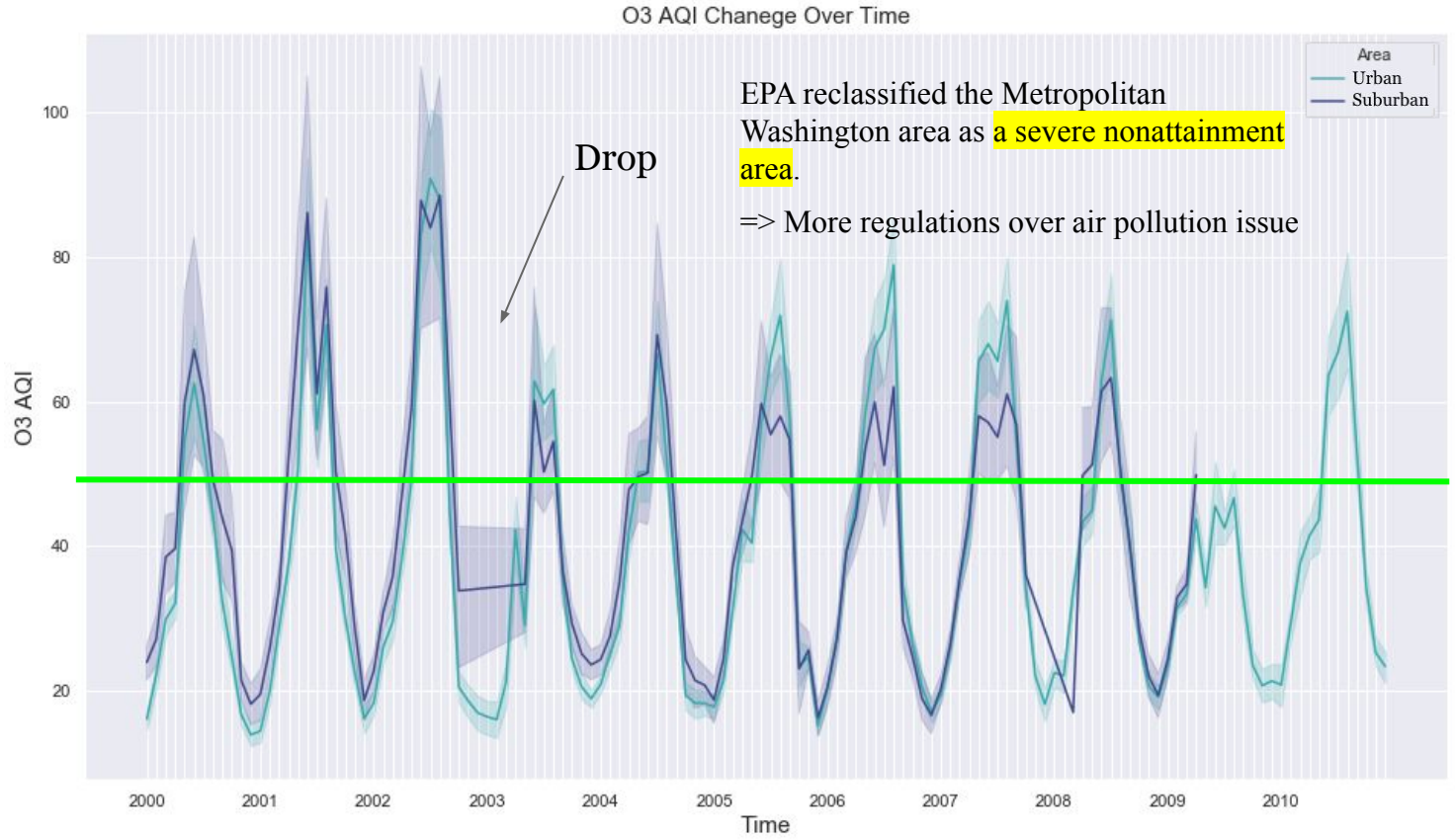


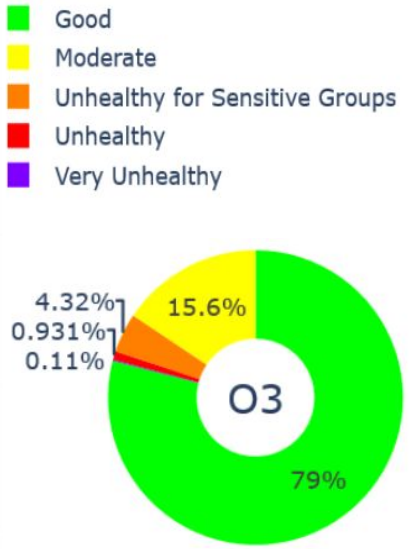
StatTools Textbook Version
For Academic Use Only

Winters' Exponential	Estimation Period	Holdouts Period
Mean Abs Err	2.515	2.249
Root Mean Sq Err	3.187	2.749
Mean Abs Per% Err	8.39%	8.78%

Small

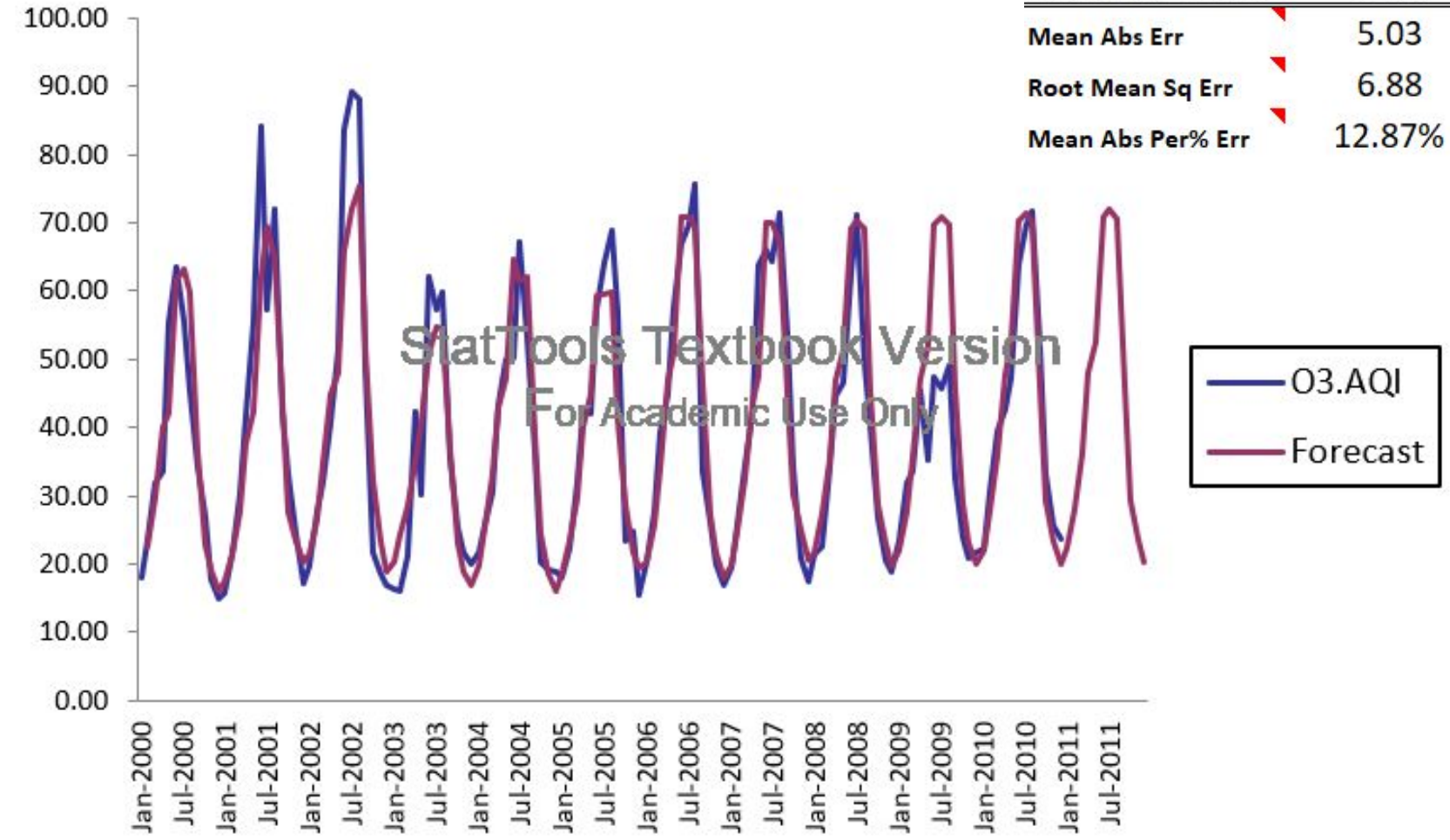
Ozone


O3 AQI Chanege Over Time
Area
Urban
Suburban
Drop
EPA reclassified the Metropolitan Washington area as a severe nonattainment area.
=> More regulations over air pollution issue
O3 AQI
100
80
60
40
20
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
Time


Good
Moderate
Unhealthy for Sensitive Groups
Unhealthy
Very Unhealthy
4.32%
0.931%
0.11%
15.6%
O3
79%

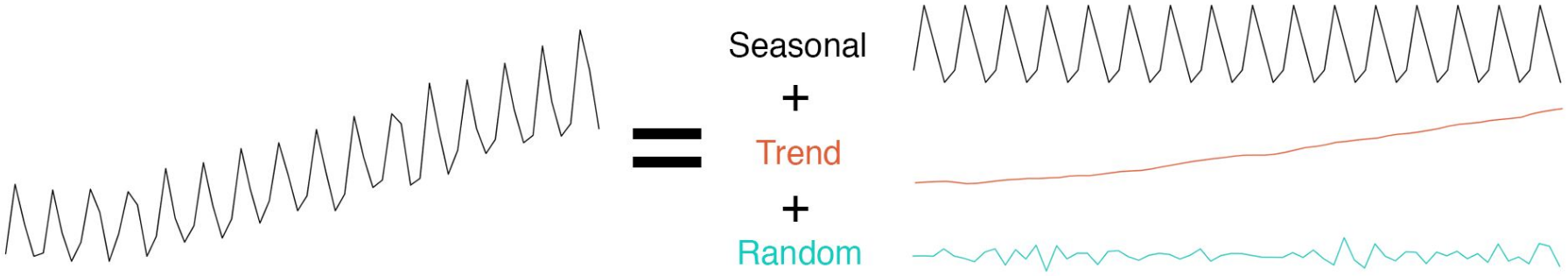
Forecast and Original Observations

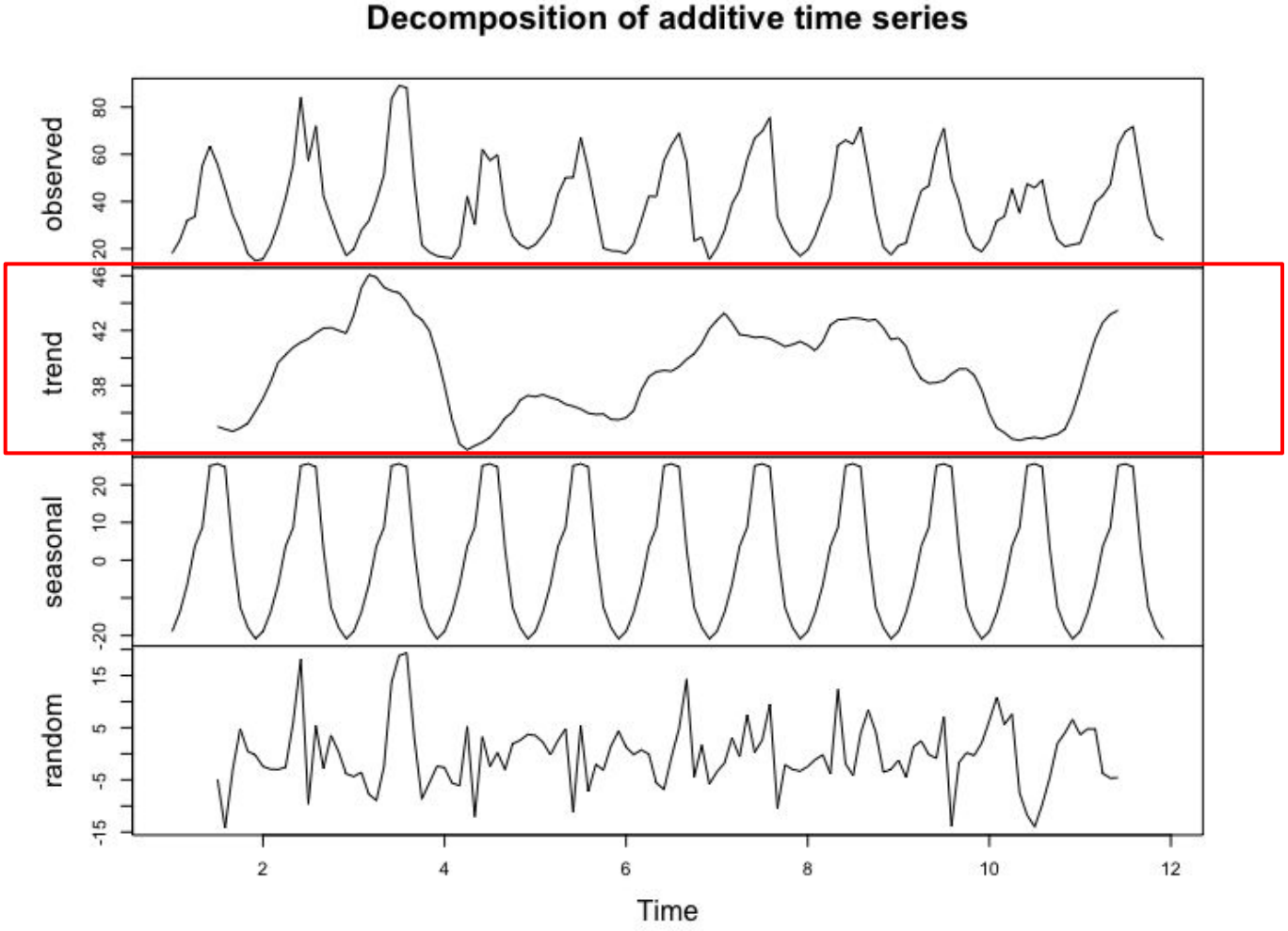
Winters' Exponential	Estimation Period	Holdouts Period
Mean Abs Err	5.03	5.84
Root Mean Sq Err	6.88	8.78
Mean Abs Per% Err	12.87%	14.92%



StatTools Textbook Version
For Academic Use Only

Time Series Analysis II




Decomposition of additive time series
observed
trend
seasonal
random
Time
2
4
6
8
10
12

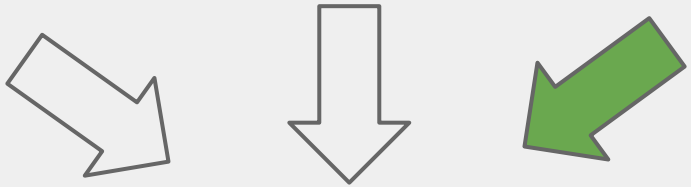
Model outcomes

Semi-log model of Ozone (Adj R^2:0.3795)

Significant factors	Interpretation
State.Code	VA (3.45941) > MD (1.49926)> DC
Season	Summer > Spring > Autumn > Winter
Daytimes	Morning > Afternoon > Evening > Night
CO	Positive Correlation (0.26661)
NO2	Positive Correlation (0.27092)
SO2	Positive Correlation (0.06627)
Wind Speed	Negative Correlation (-0.09780)

Suggestion for Reducing Ozone Emission

NOx ✗ + Sunlight ✗ + VOC ✓



Ozone

Reduce VOC (Volatile Organic Compounds)

Implements

- Avoid using aerosol consumer products such as hairsprays, air fresheners, deodorants, and insecticides that often use VOCs as their propellants.
- Avoid using VOC-containing products such as organic cleaning solvents.
- Store VOC-containing products in air-tight containers.
- Buy products with less packaging as the printing of packaging materials generates VOCs.
- Drive less, share rides and use public transportation.

Source: EPA(2011). Volatile Organic Compound (VOC) Control Regulations. https://www3.epa.gov/region1/airquality/voc.html

Thank you!