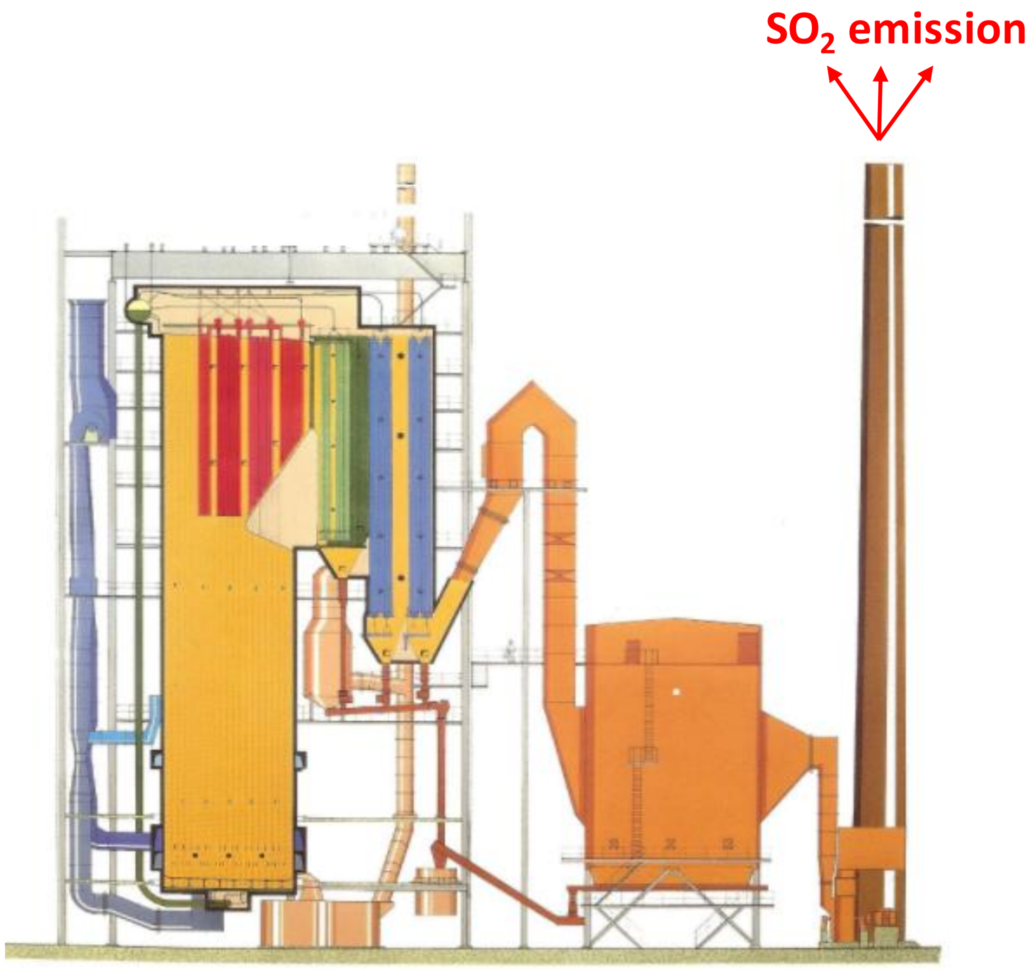
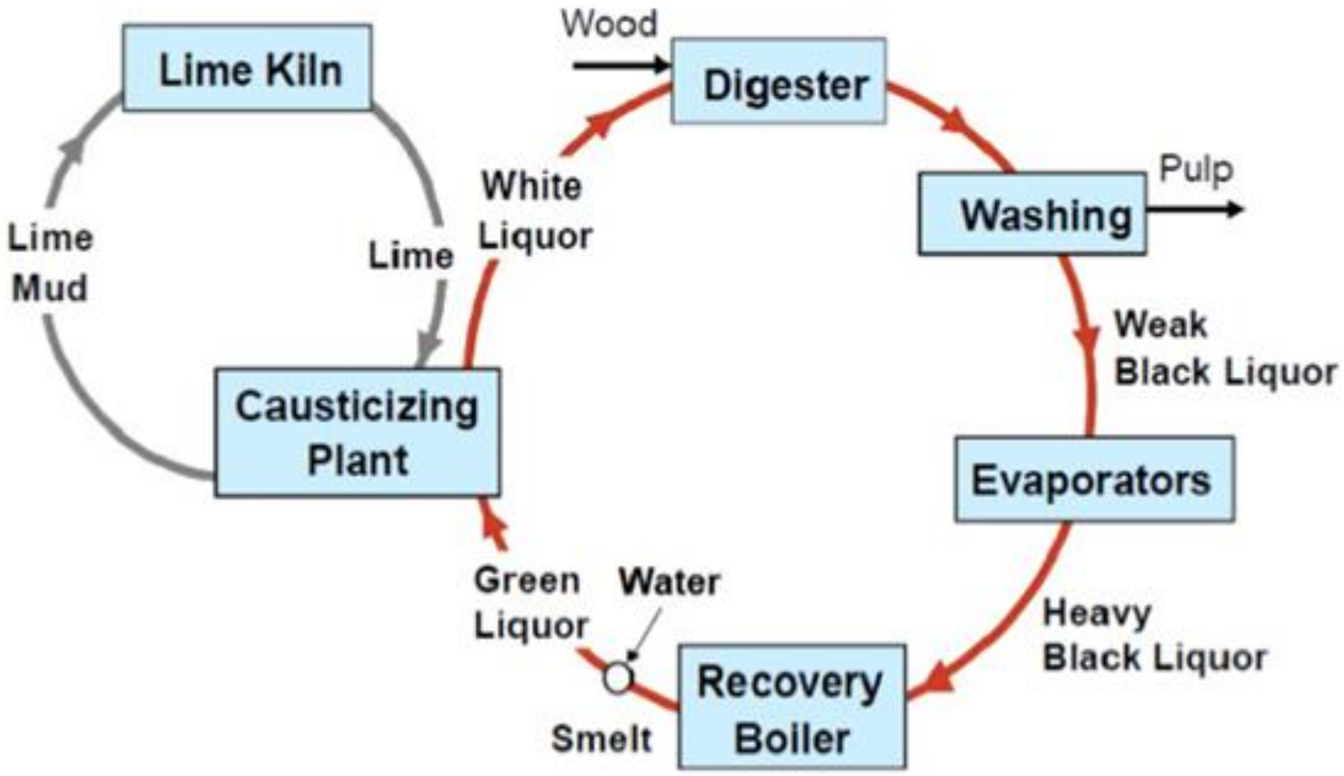
Root Cause Analysis for SO2 Emission in Kraft Recovery Boiler

Daniel Yung Chi Cho

Dec 6, 2018

Recovery Cycle





Problem Statement

- High sulfur dioxide emission at recovery boiler at a pulp mill.
- Sulfur dioxide is a regulated pollutant.
- Company faces penalty for emission beyond the limit.

Objective

- Finding the root cause for sulfur dioxide emission based on historical data.

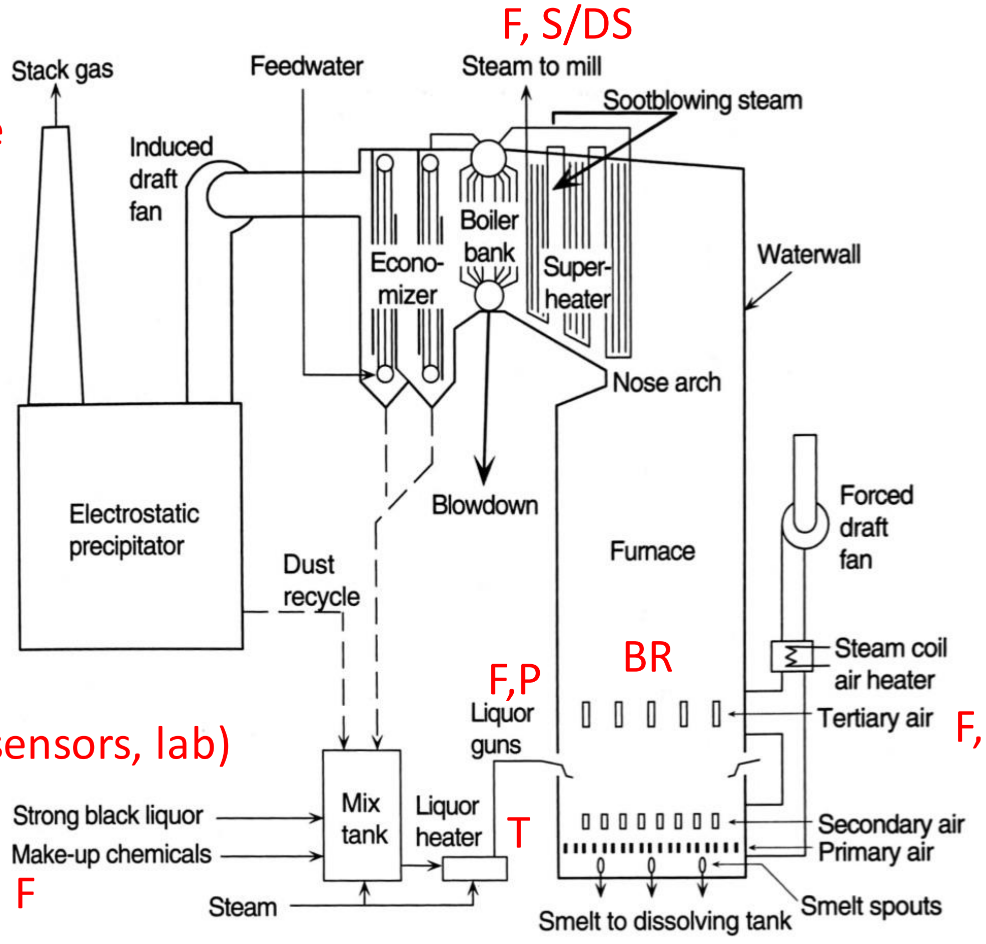
Data Source and Variables

- **Source:** hourly process data from March 2016 to June 2016 provided by an industrial partner.
- **Data Type:** Numerical
- **Variables:** 26
- **Observations:** 2327

Time Stamp	BL Density Dry Solids-B
Primary Air Flow	Primary Air Temp.
Secondary Air Flow	Secondary Air Temp.
Tertiary Air Flow	Steam Flow
Primary Windbox Pressure	Liquor Solids Flow
Secondary Windbox Pressure	ST/DS Ratio
Tertiary Windbox Pressure	SO2 Raw (1min Avg.)
Burn Rate	SO2 Raw (1Hr Avg.)
Sulfidity Calculated	Black Liquor Press
BL Solid 50/50 Test	BL Indirect Heater Temp.
Flue Gas Oxygen	Salt Cake Rotary Feeder Sp.
TRSRAW 1 Min Avg.	Lime Kiln Entry
BL Density Dry Solids-A	Nox (1Hr Avg.)

Recovery Boiler schematics

SO2 Hour Average
SO2 Minute Average
Oxygen
Nox



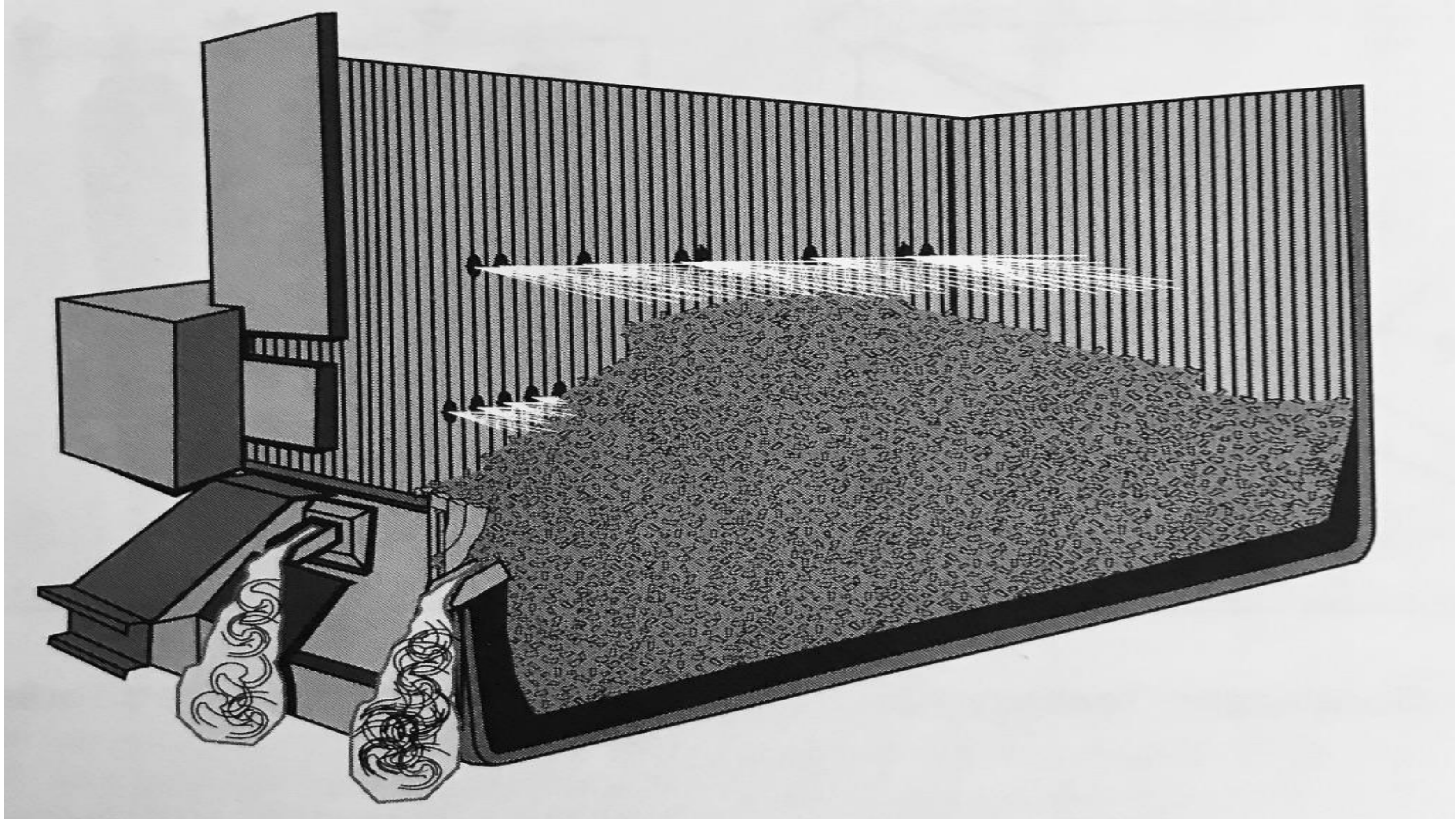
F, T

Density (sensors, lab)
Sulfidity

F, T,P
F, T,P

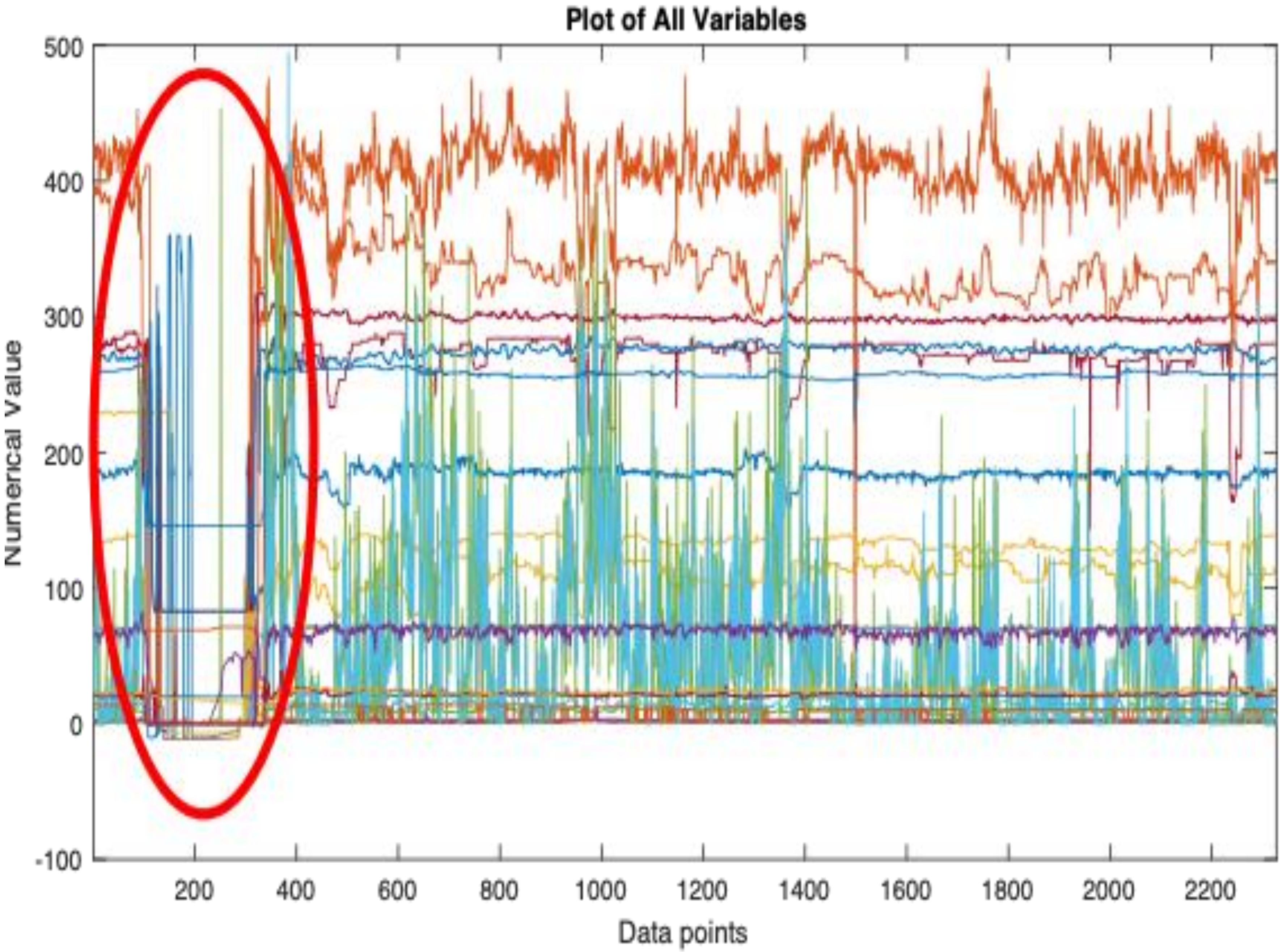
Sulfidity

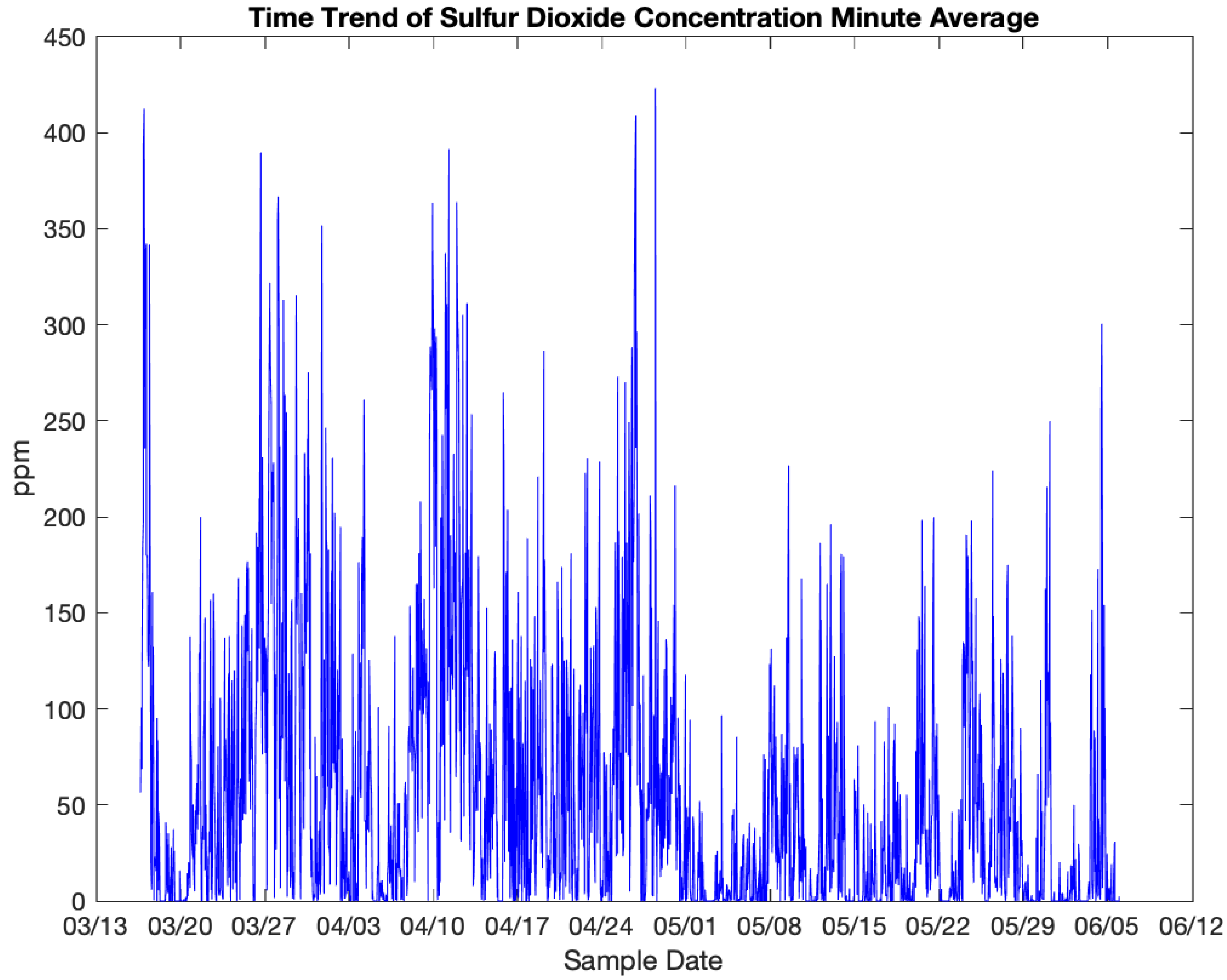
Inside the Recovery boiler



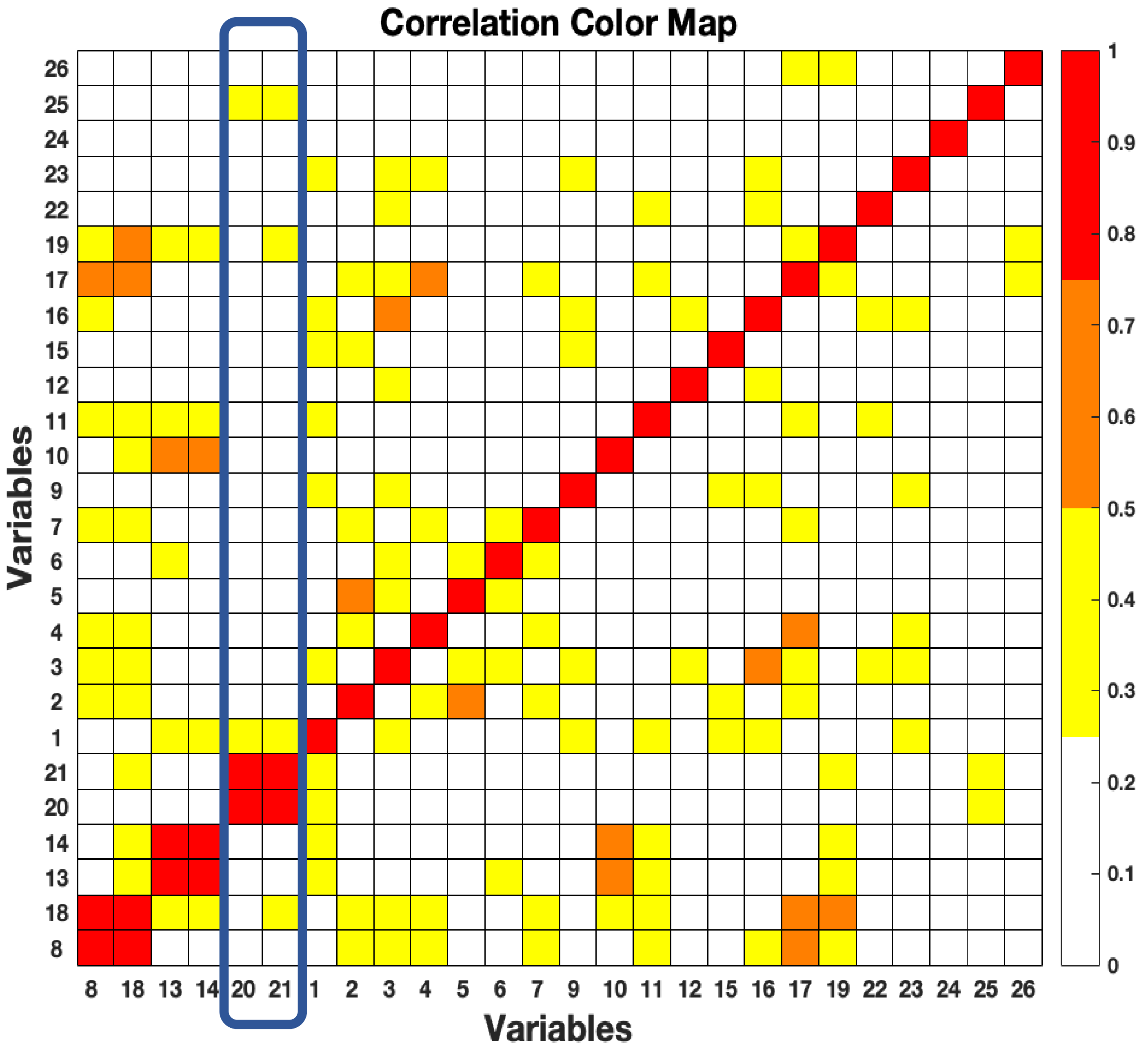
Exploratory Analysis

Correlation Map, High Density plot, Compression Test, Quantification Test

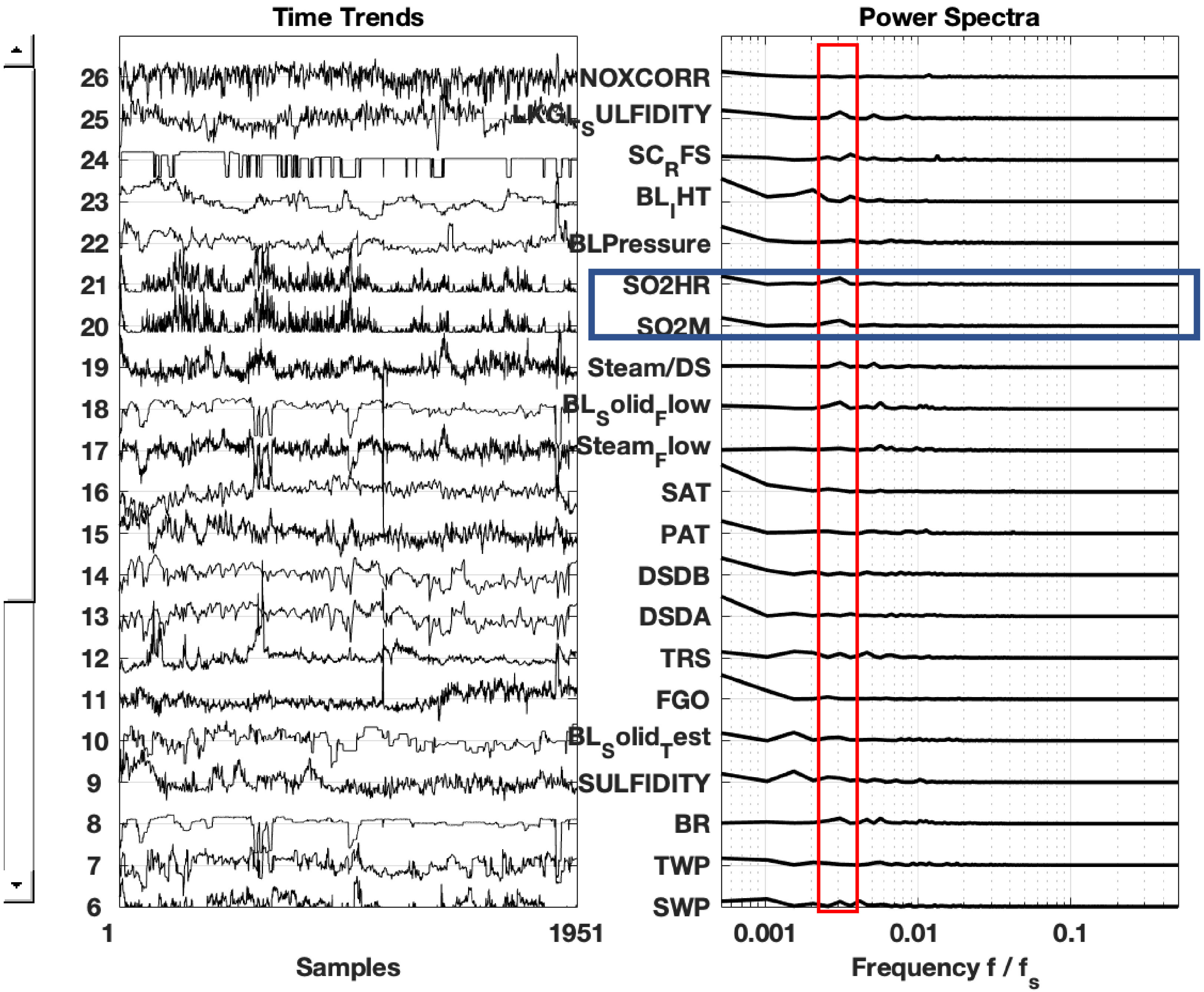

Plot of All Variables
Numerical Value
Data points
500
400
300
200
100
0
-100
200
400
600
800
1000
1200
1400
1600
1800
2000
2200



Time Trend of Sulfur Dioxide Concentration Minute Average
ppm
0
50
100
150
200
250
300
350
400
450
03/13
03/20
03/27
04/03
04/10
04/17
04/24
05/01
05/08
05/15
05/22
05/29
06/05
06/12
Sample Date



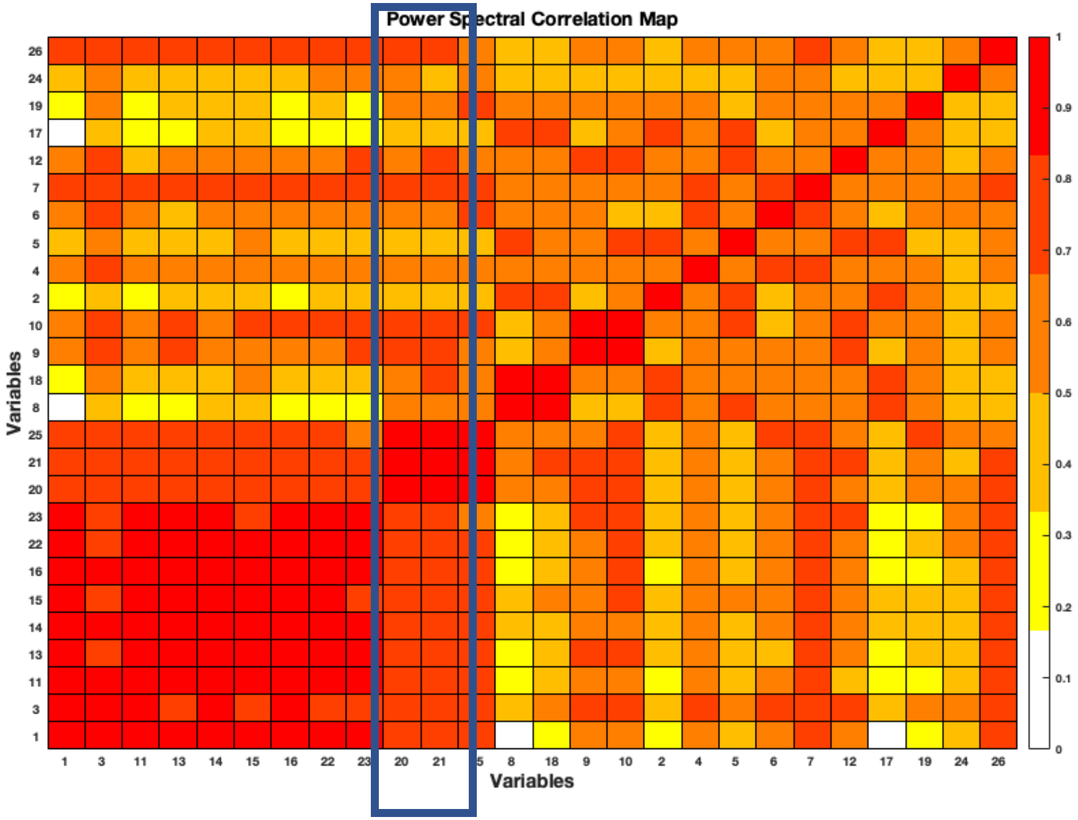
Correlation Color Map
Variables
26
25
24
23
22
19
17
16
15
12
11
10
9
7
6
5
4
3
2
1
21
20
14
13
18
8
Variables
8 18 13 14 20 21 1 2 3 4 5 6 7 9 10 11 12 15 16 17 19 22 23 24 25 26
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0



Time Trends
Power Spectra
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
1
1951
Samples
NOXCORR
LKGL_SULFIDITY
SC_RFS
BL_IHT
BLPressure
SO2HR
SO2M
Steam/DS
BL_Solid_Flow
Steam_Flow
SAT
PAT
DSDB
DSDA
TRS
FGO
BL_Solid_Test
SULFIDITY
BR
TWP
SWP
0.001
0.01
0.1
Frequency f / f_s

Spectral Correlation

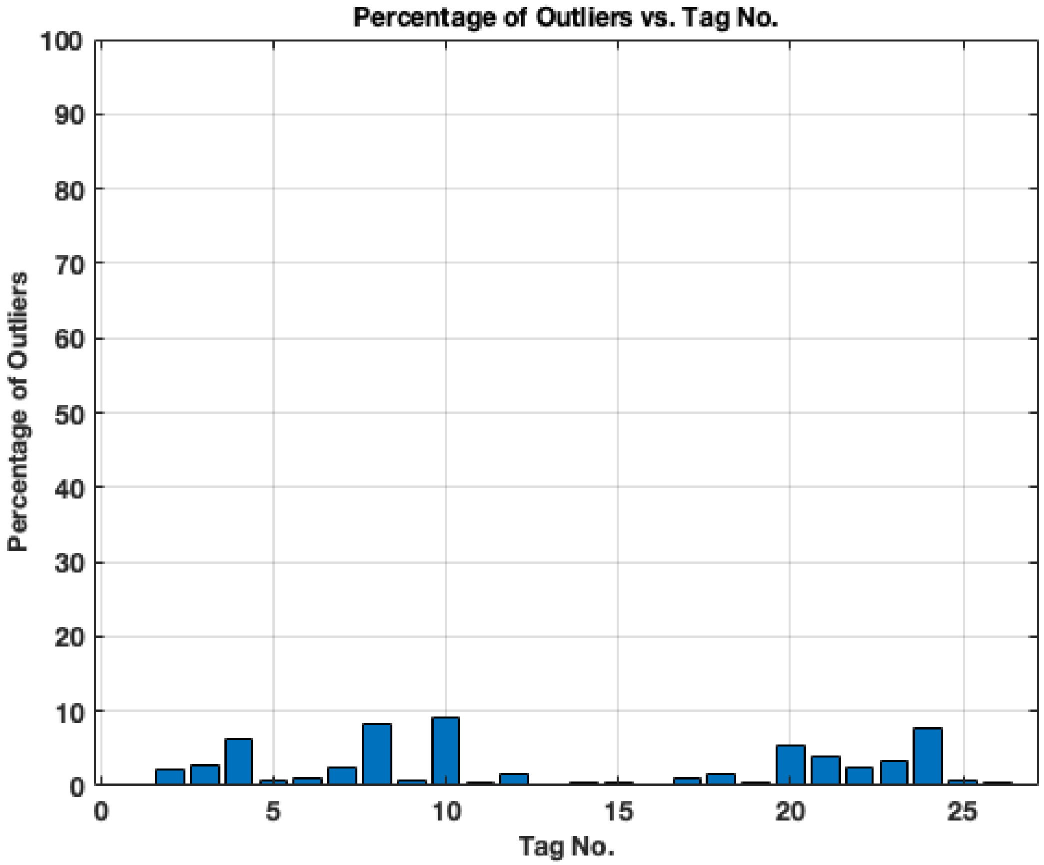
- Secondary Air Flow (Tag 3)
- Flue Gas Oxygen (Tag11)
- Dry Solid Density Transmitter A & B (Tag 13,14)
- Primary Air Temperature (Tag 15)
- Secondary Air Temperature (Tag 16)
- Black Liquor Pressure (Tag 22), Black Liquor Indirect Heater Temperature (Tag 23)
- Green Liquor Sulfidity (tag 25)
- and NOX Correlation (Tag 26)



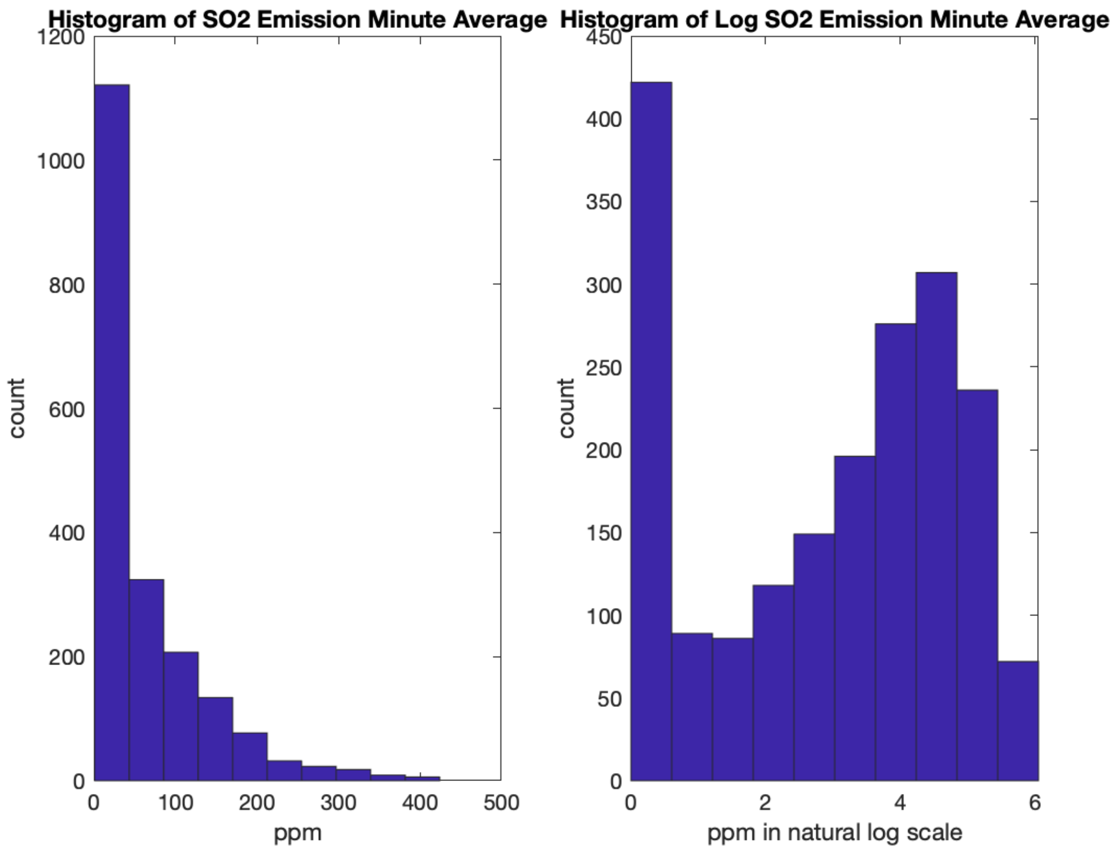
Pre-Processing

Outliner, Transformation, lag adjustment, MA filter, Normalization

Outlier Treatment Using DVA tool



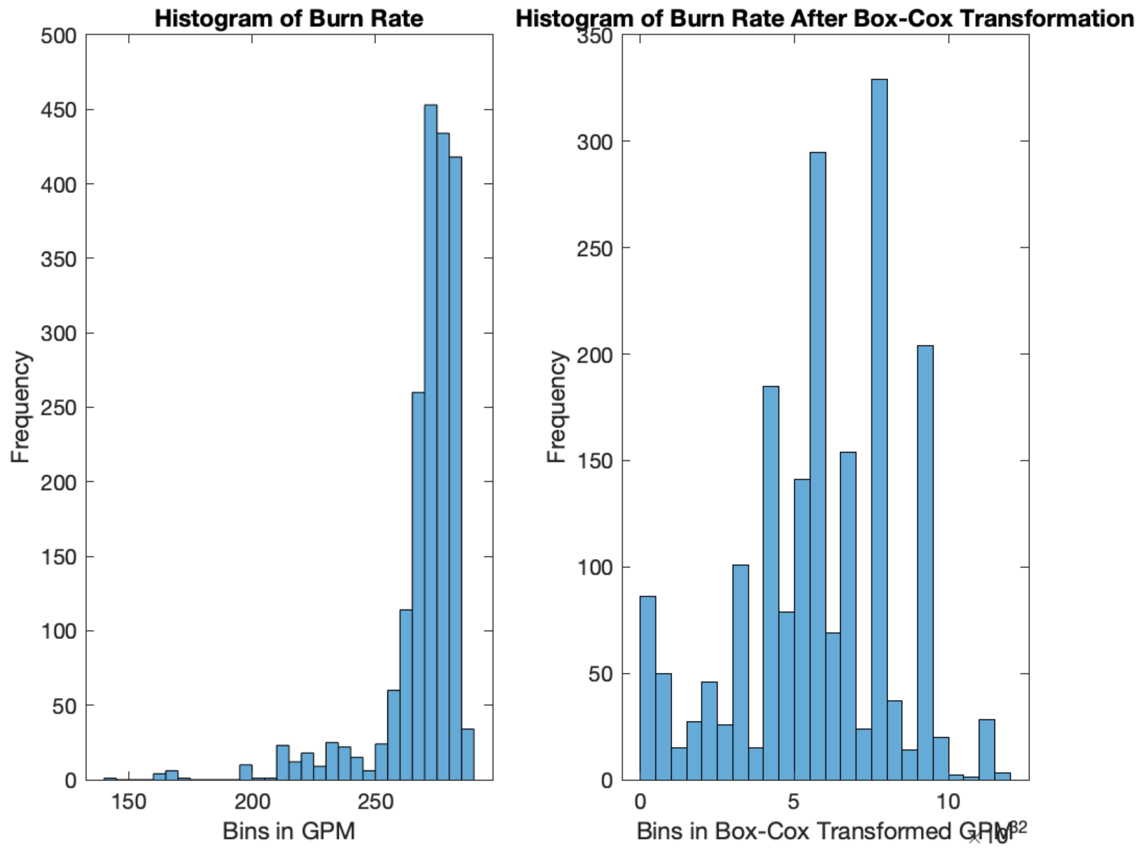
Log Transformation of SO2 Emission



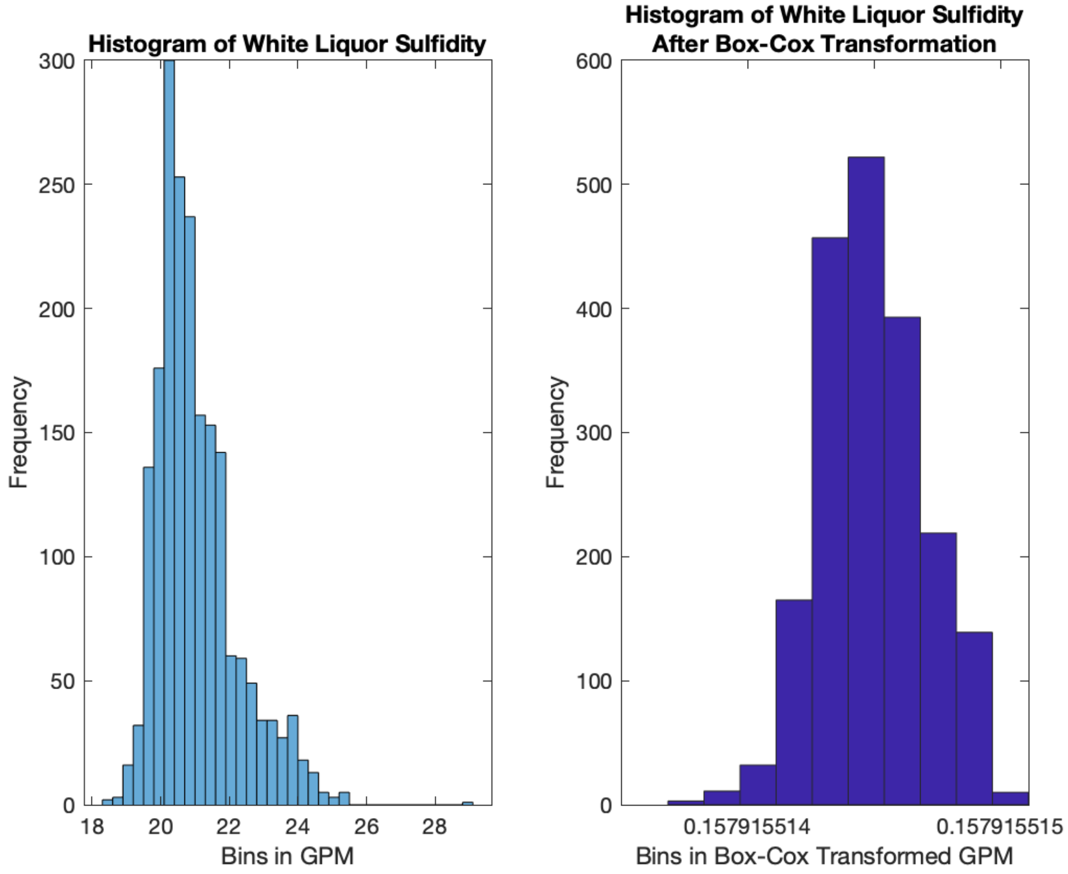
Box-Cox Transformation of burn rate

Definition of Box-Cox

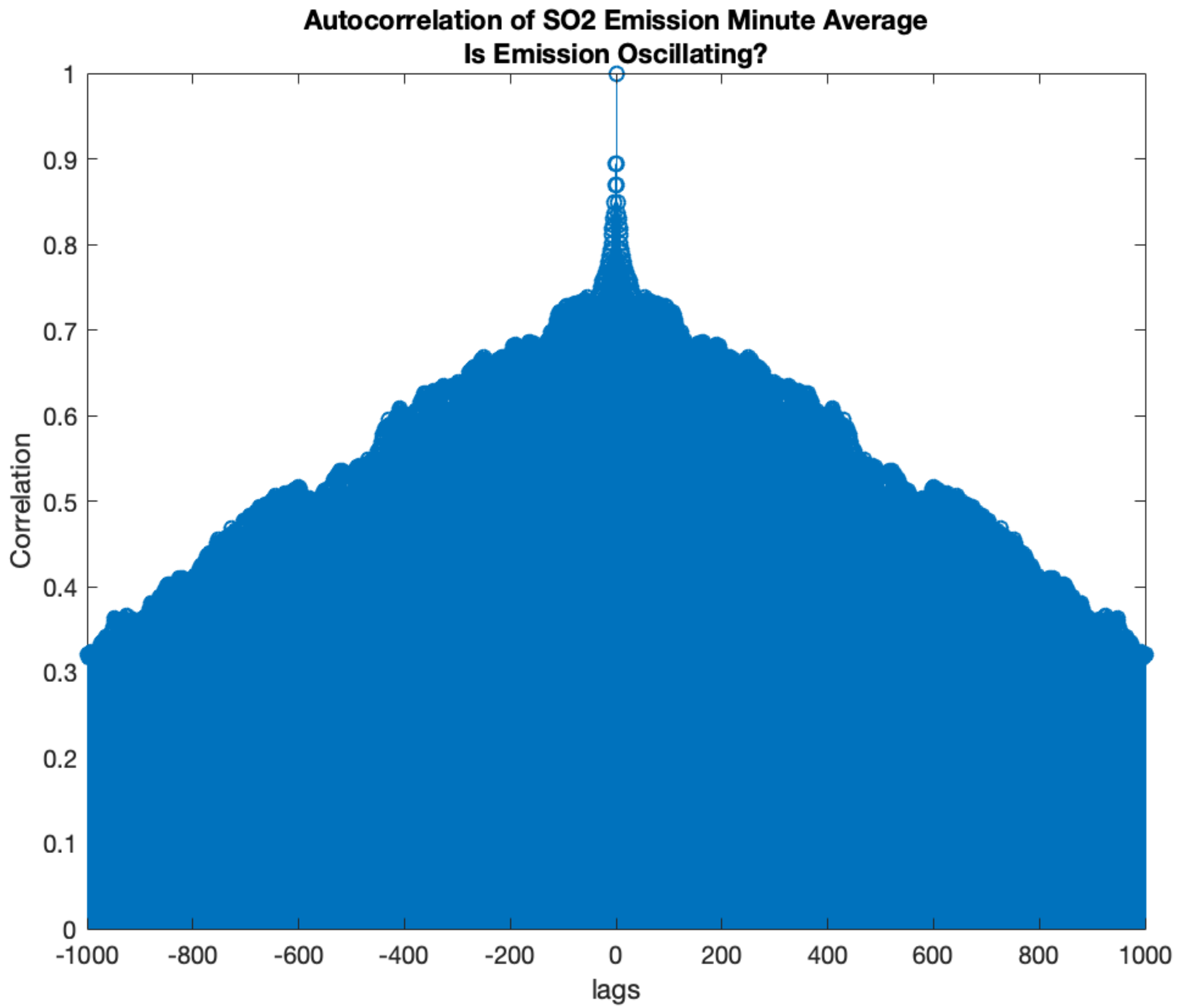
$$Data(\lambda) = \frac{Data^{\lambda} - 1}{\lambda}$$



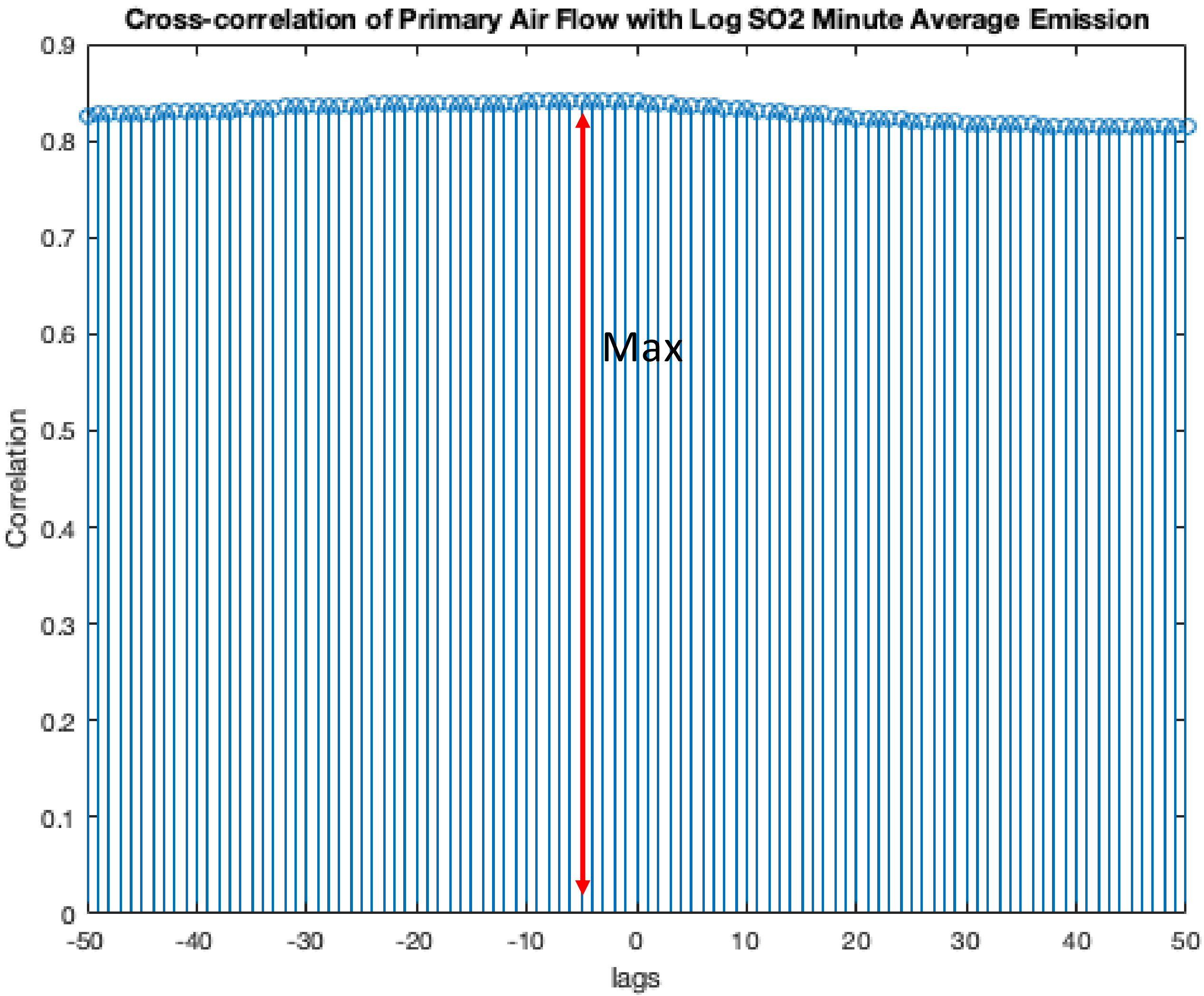
Box-Cox Transformation of Sulfidity



Is Emission Oscillating?



Cross-correlation for all variables with SO2

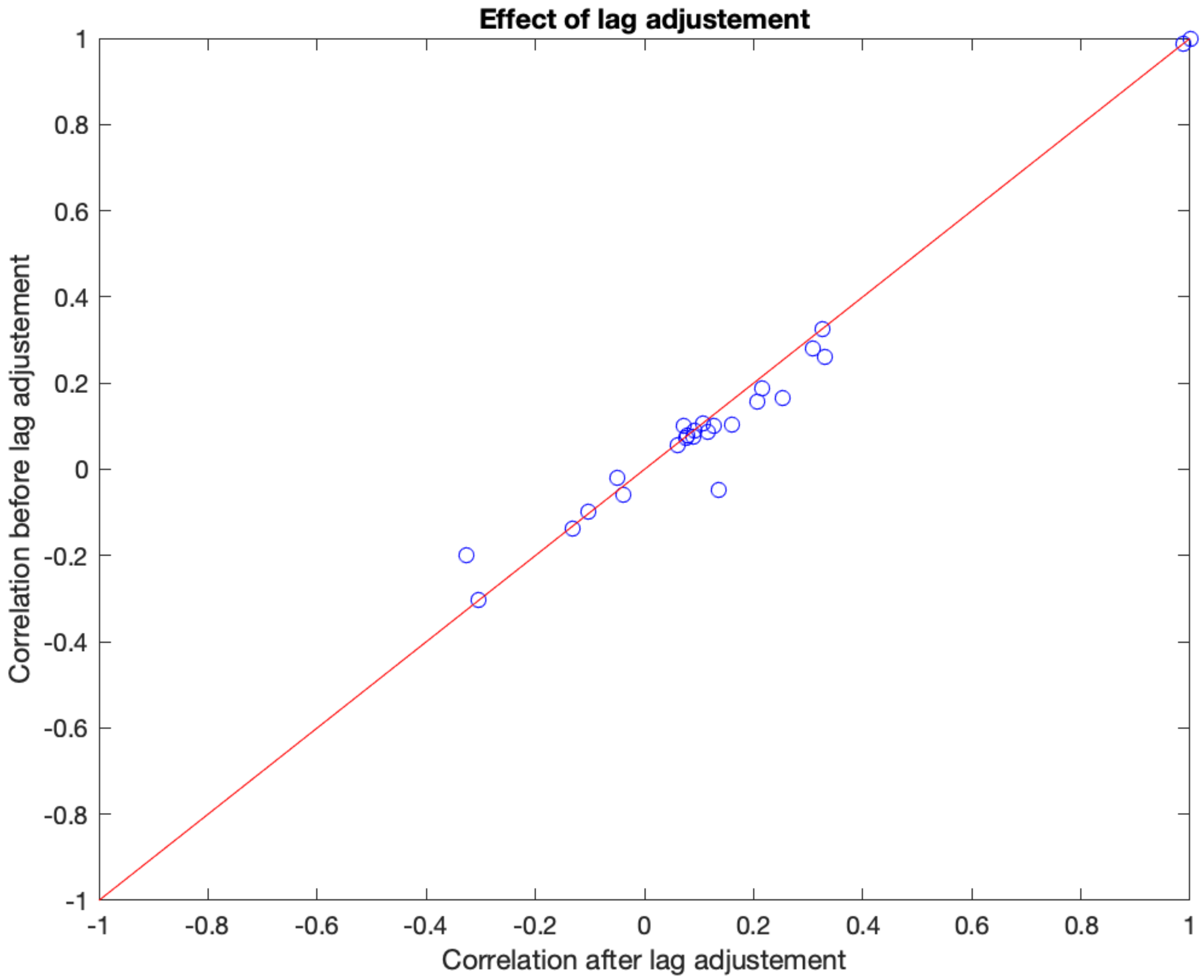


Lag time found by maximizing Cross-Correlation

Variables	Lag time with Log
Time stamp	0
PAF	-5
SAF	-5
TAF	-9
PWP	-5
SWP	0
TWP	-26
BR	-10
Sulfidity	-1
BL_Solid_Test	-9
FGO	-21
TRS	-30
DSDA	-10

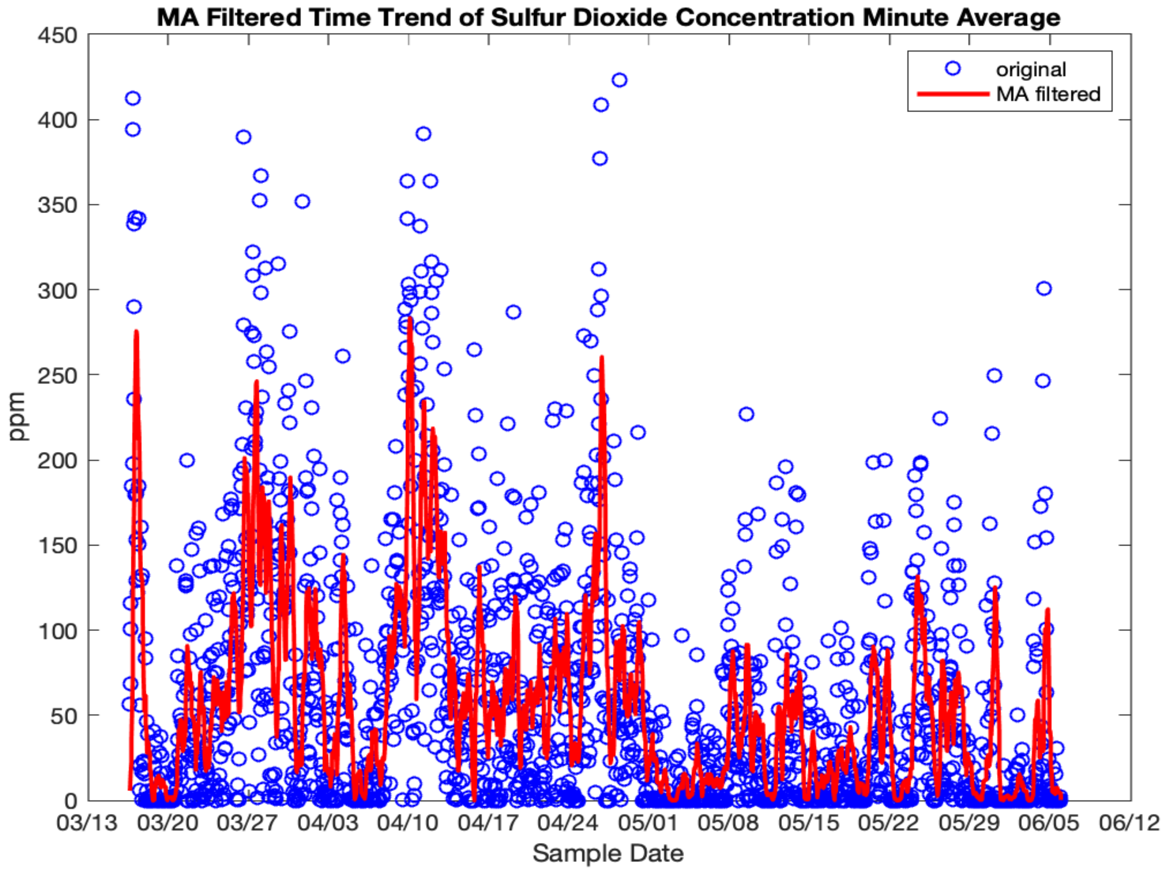
Variables	Lag time
DSDB	-10
PAT	0
SAT	0
Steam_Flow	-2
BL_Solid_Flow	-10
Steam/DS	-2
SO2M	0
SO2HR	-1
BLPressure	-10
BL_IHT	0
SC_RFS	-34
LKGL_SULFIDITY	-9
NOXCORR	-15

Effect of Lag adjustment
Changes in Correlation with SO2 emission



Moving Average filter

Window size = 20



Normalization

- Data are normalized and centered to balance the scale of the variables and make variance comparable.

Creation of training and test set

- Training set and test set were created using random sampling without replacement.
- 1366 records goes to training and 585 records goes to testing

Training and Testing

OLS, PCR,PLS, Regression Trees

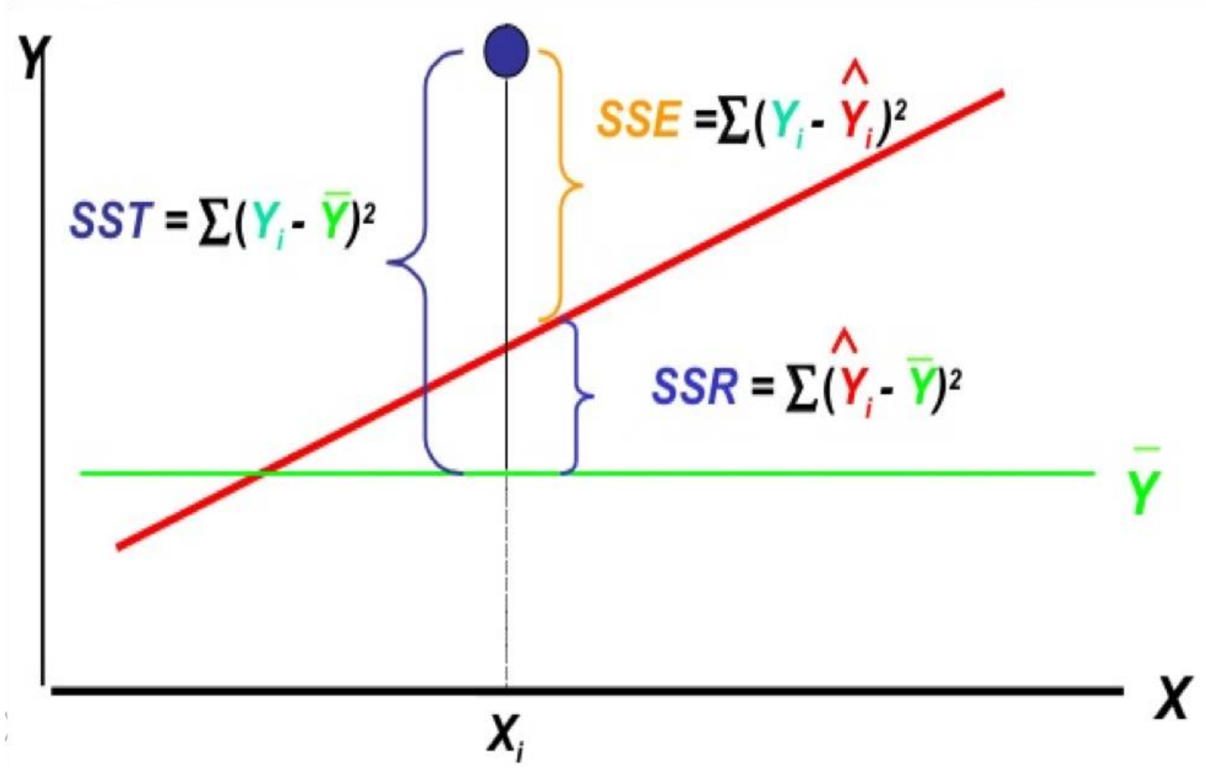
Ordinary Least Square Regression

Model

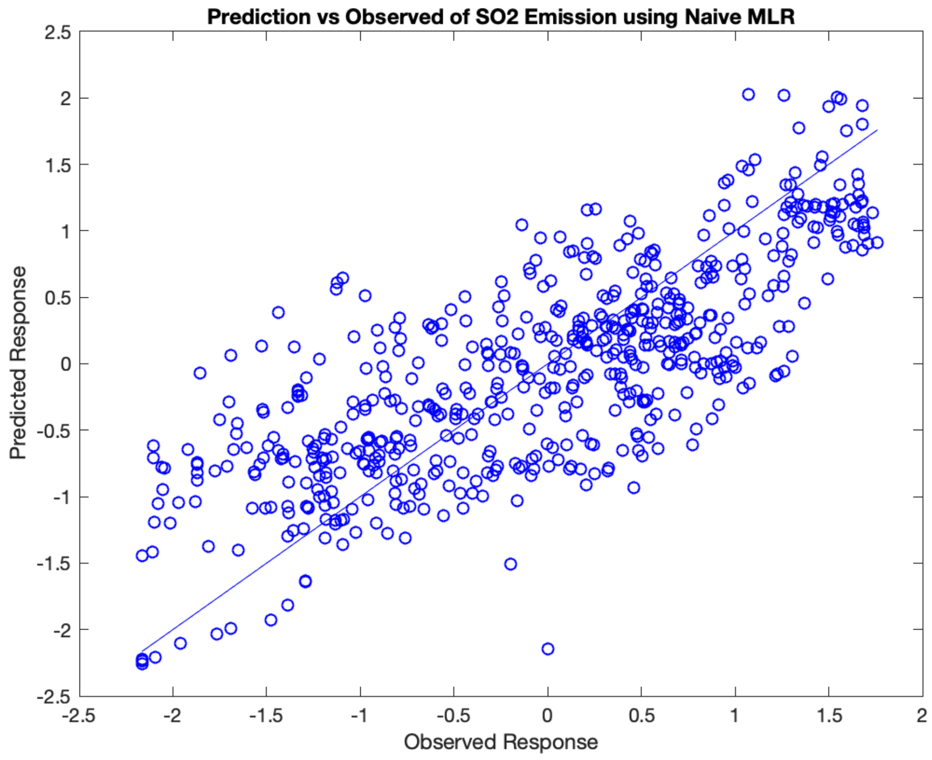
$\hat{\beta} = (X_t^T X_t)^{-1} X_t^T Y_t$

Testing

$\hat{Y} = \hat{\beta} X_{test}$



Ordinary Least Square Testing

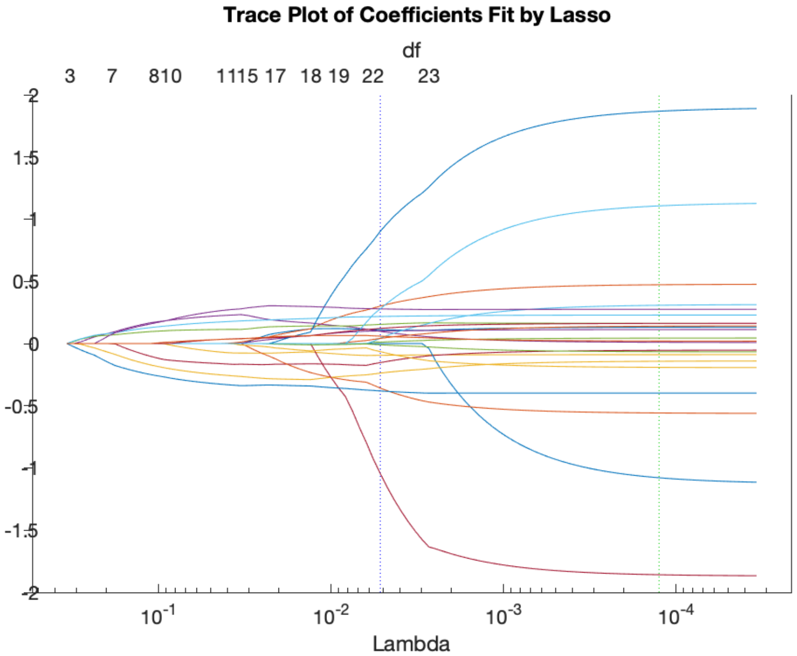


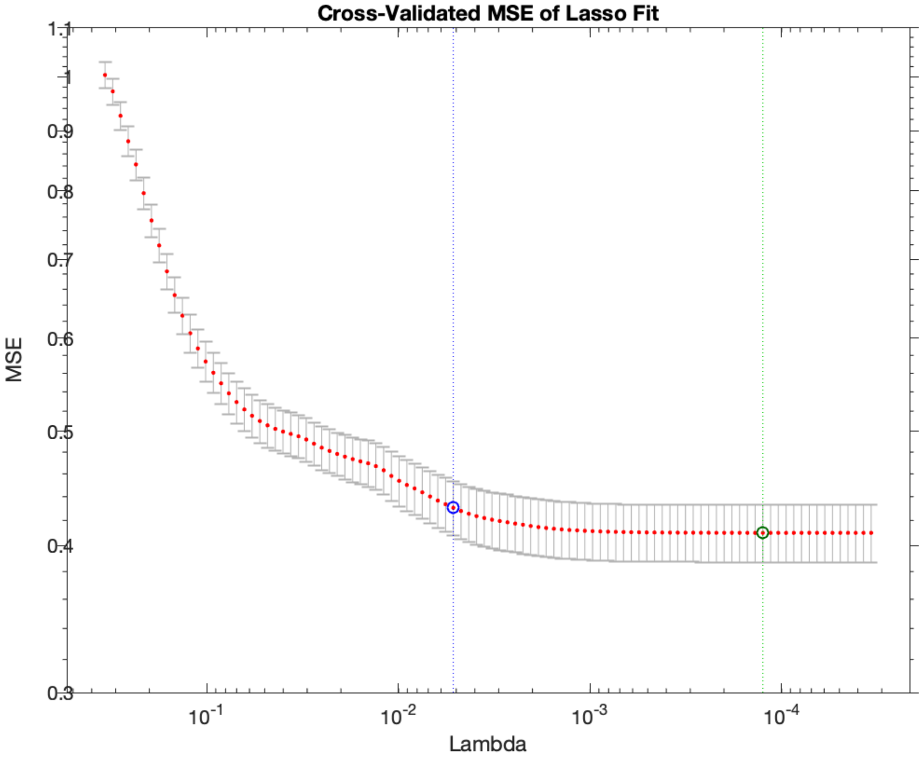
- Regress with all 23 variables
- SSE is 237.8710
- RMSE was 0.6512
- R^2 is 0.5866.

LASSO

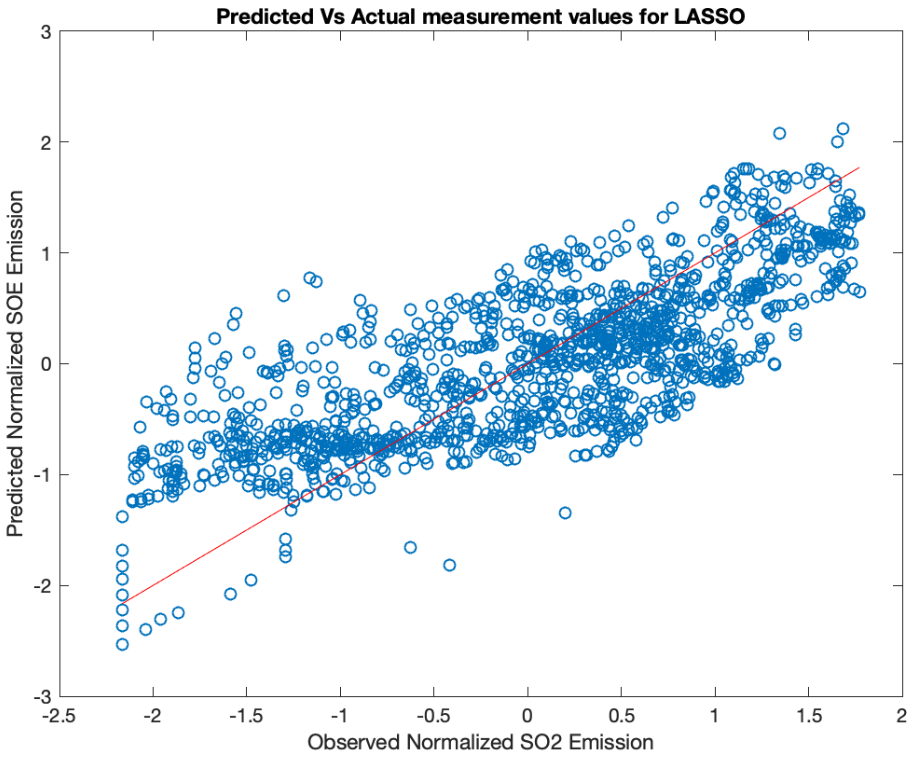
Objective Function

$$\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{N} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1 \right\}$$



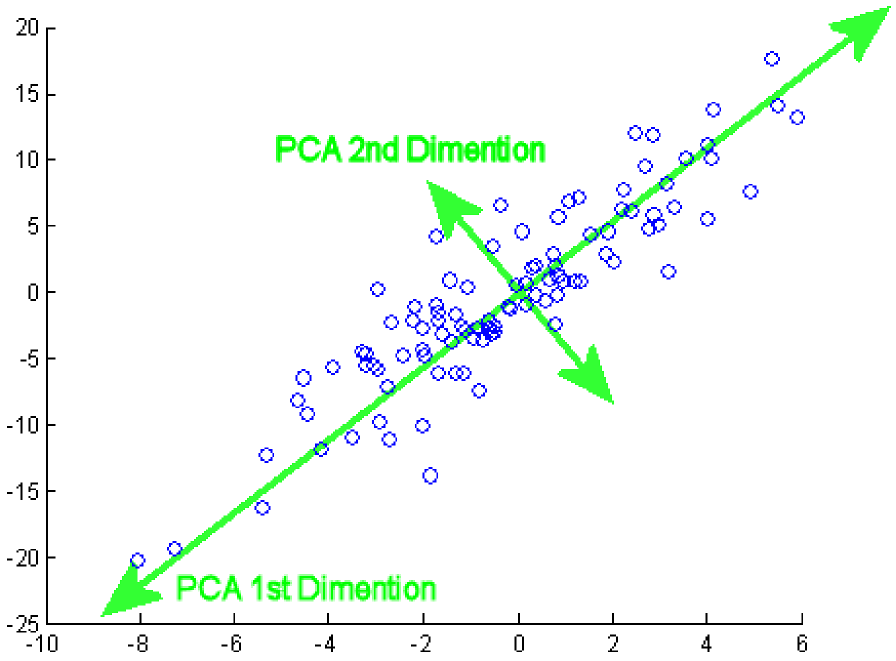


Testing LASSO



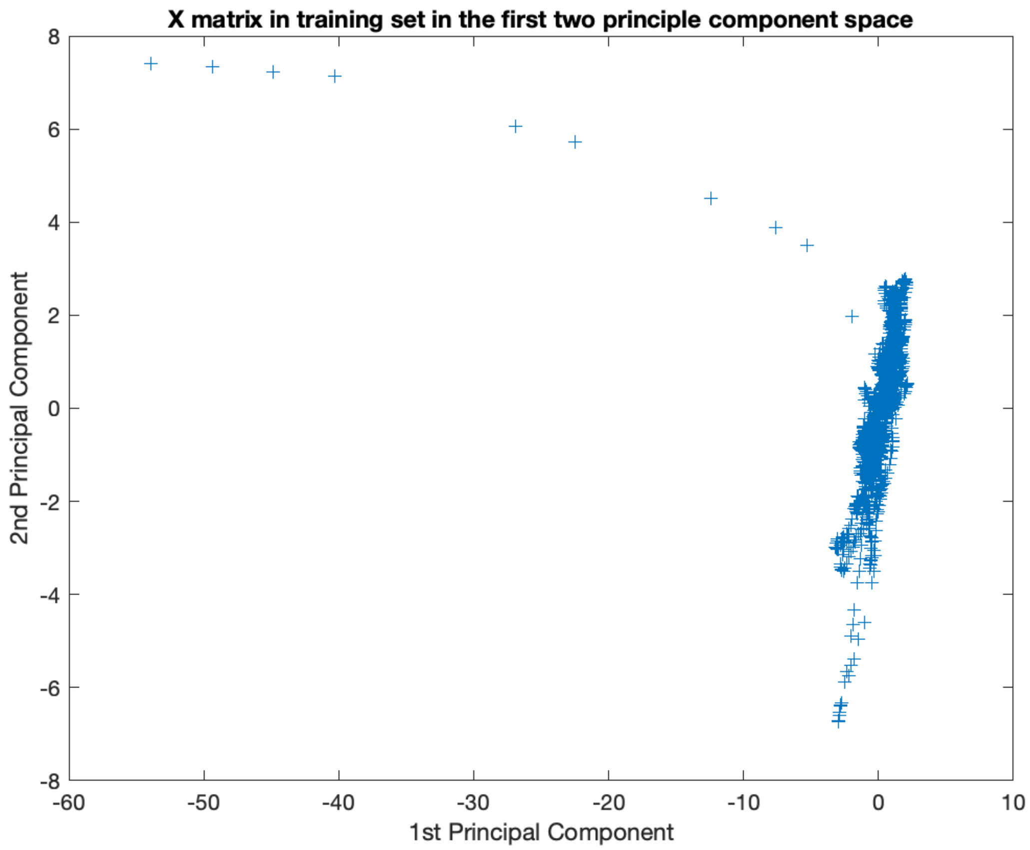
- Regress with all except Sulfidity
- RSS = 251.8956
- RSE = 0.6695
- Rsq = 0.5623

Principle Component Analysis

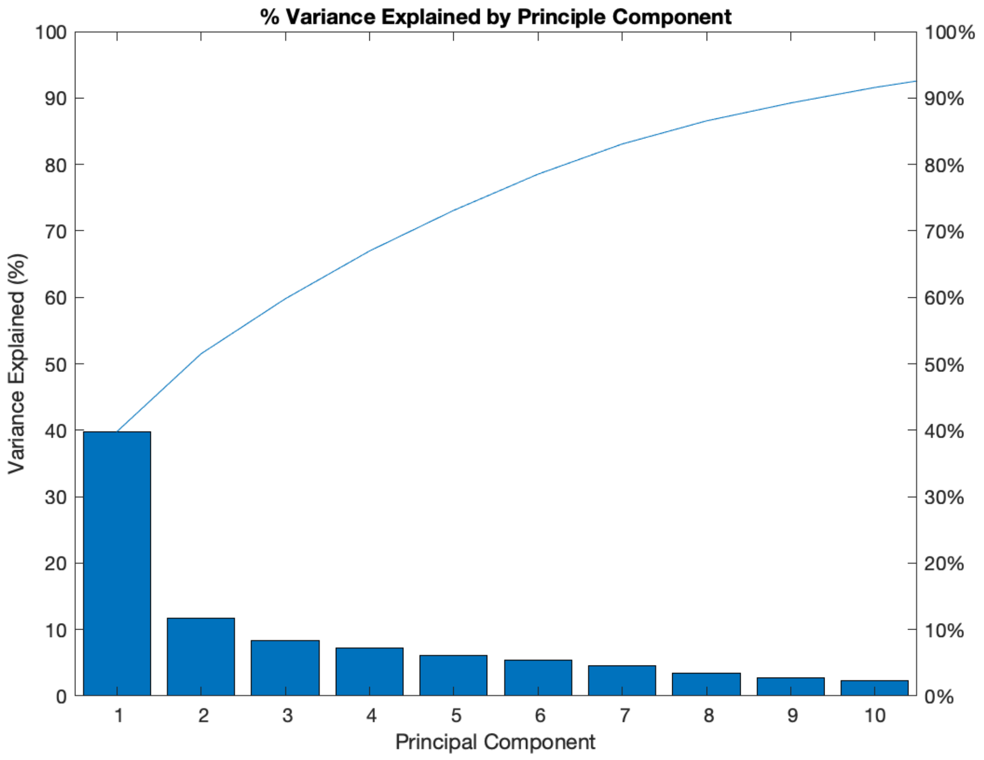


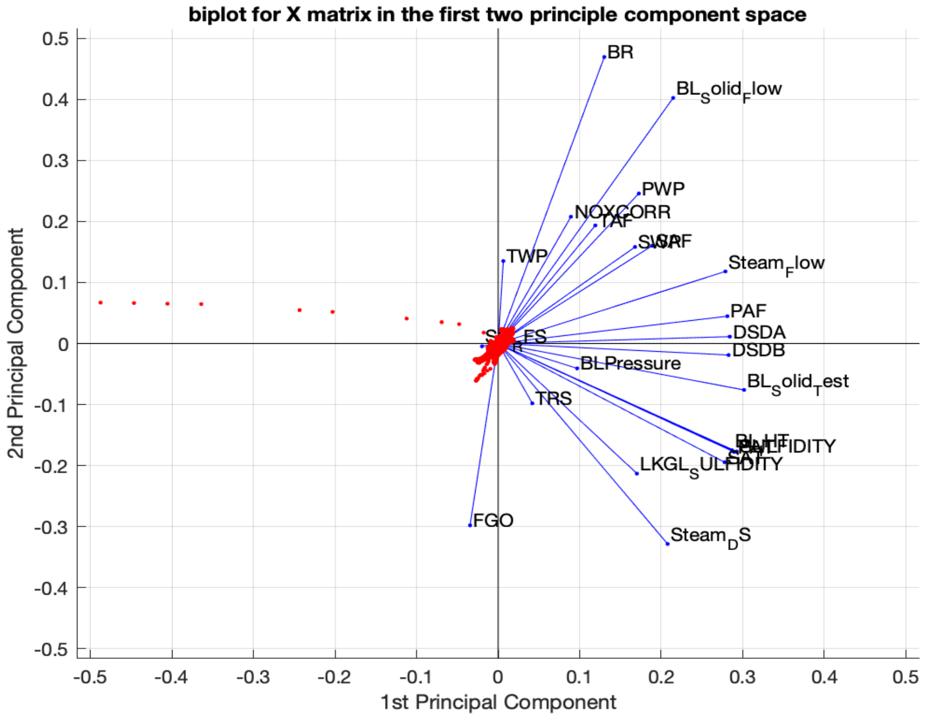
- Dimension reduction by creating latent variables that are linear combination of the original variables.
- PCA finds efficient ways to describe data with less variable and with minimum data loss

X training data in the first two principle component space

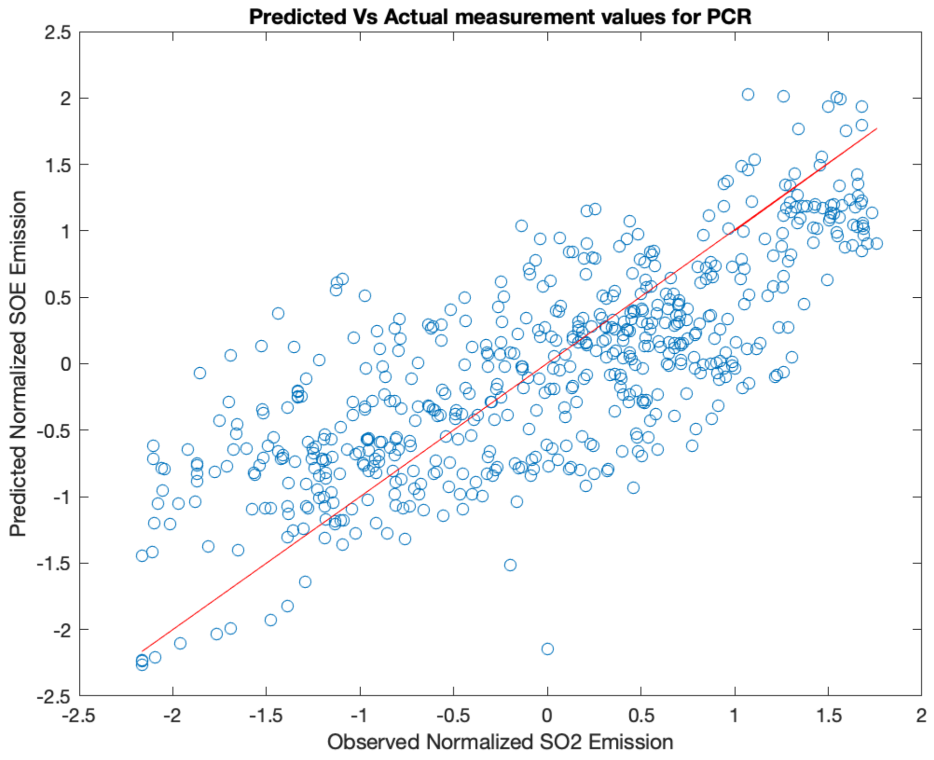


Variance explained by principle components



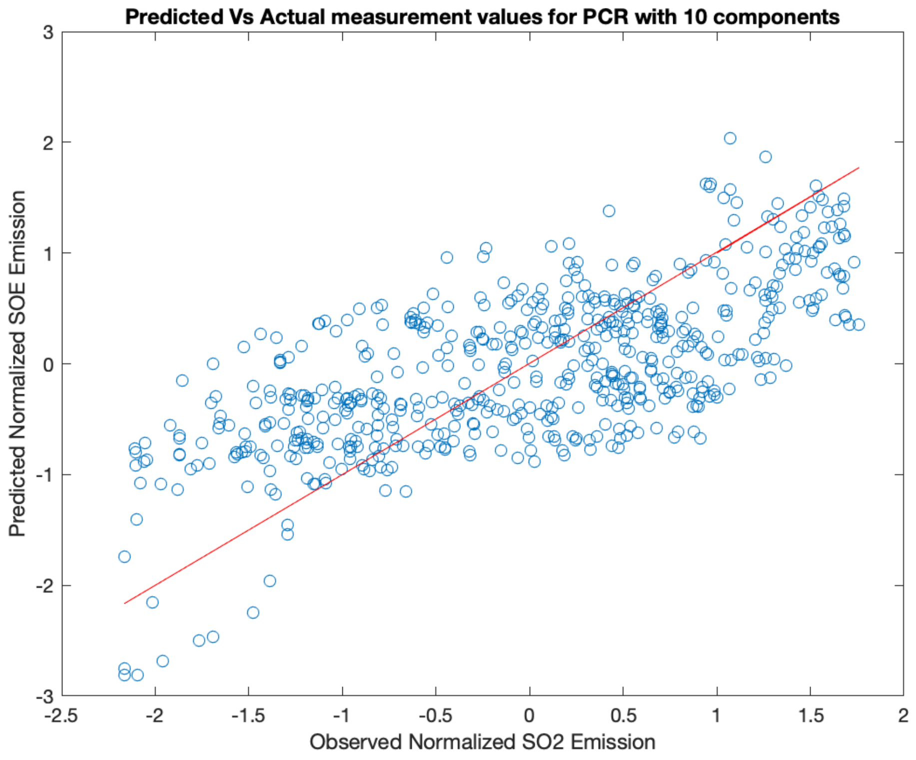


Principle Component Regression



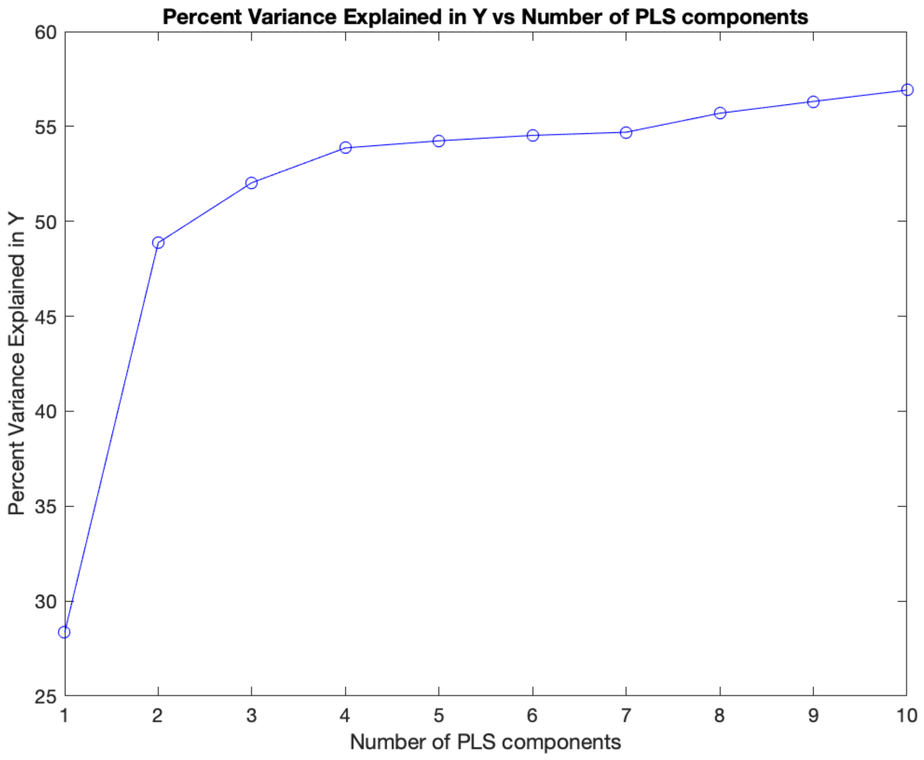
- Regression with all variables
- SSE is 237.87
- RMSE is 0.6512
- Rsq = 0.5866
- Identical to OLS when regressing all variables

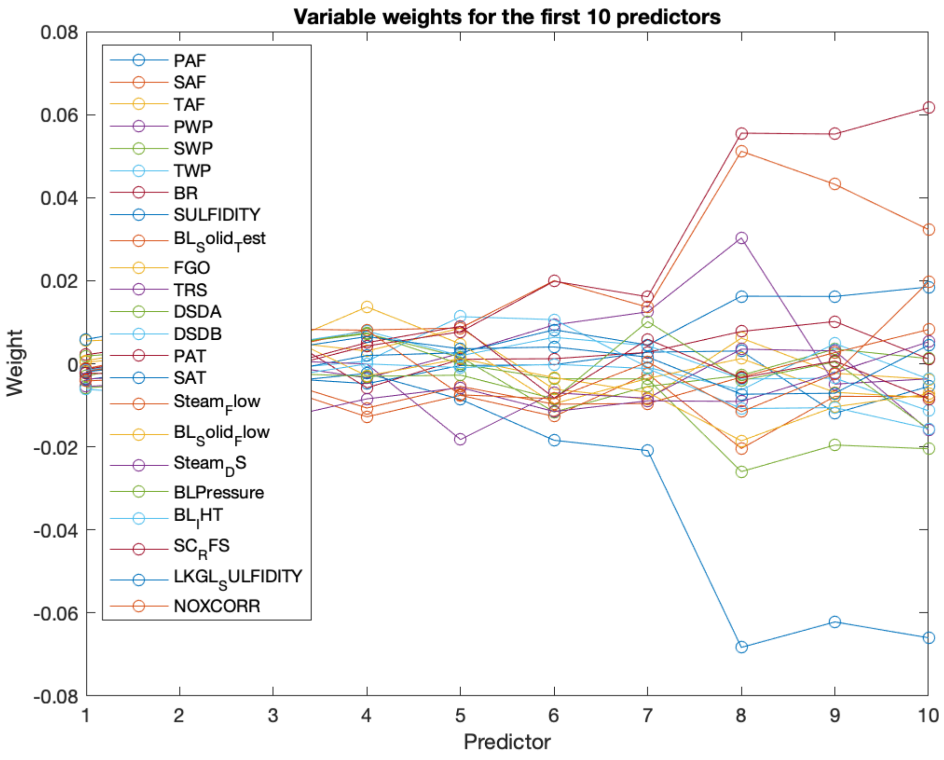
Principle Component Regression



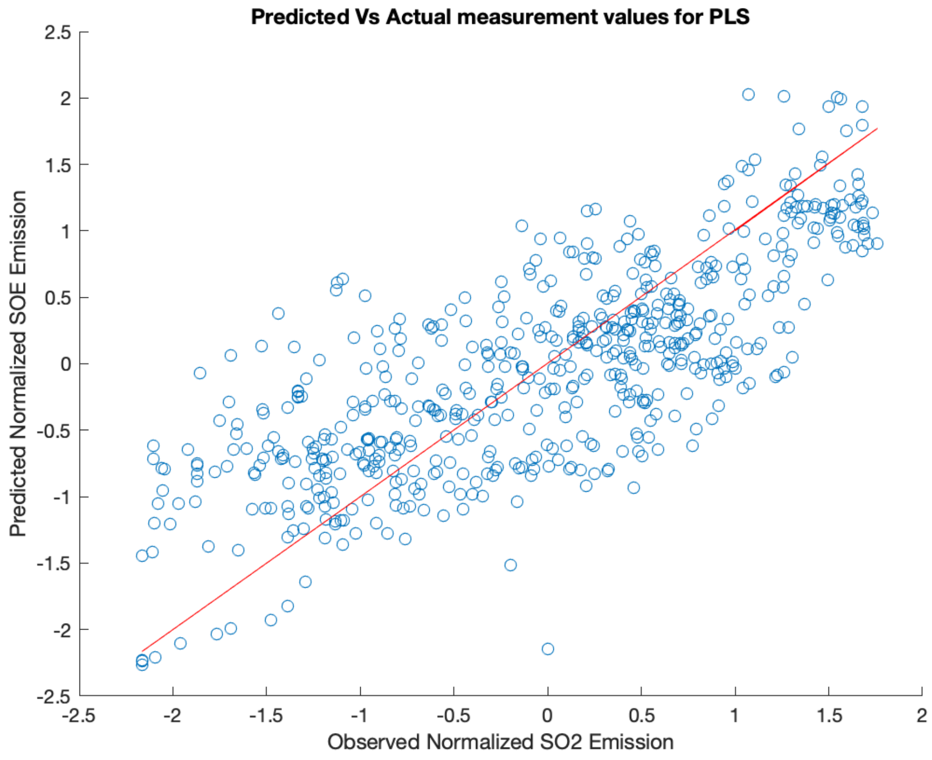
- Regress with top 10 components
- SSE = 304.6549
- RMSE = 0.7369
- R-sq = 0.4706
- Over estimated in the low end and underestimated in the high end

Partial Least Square



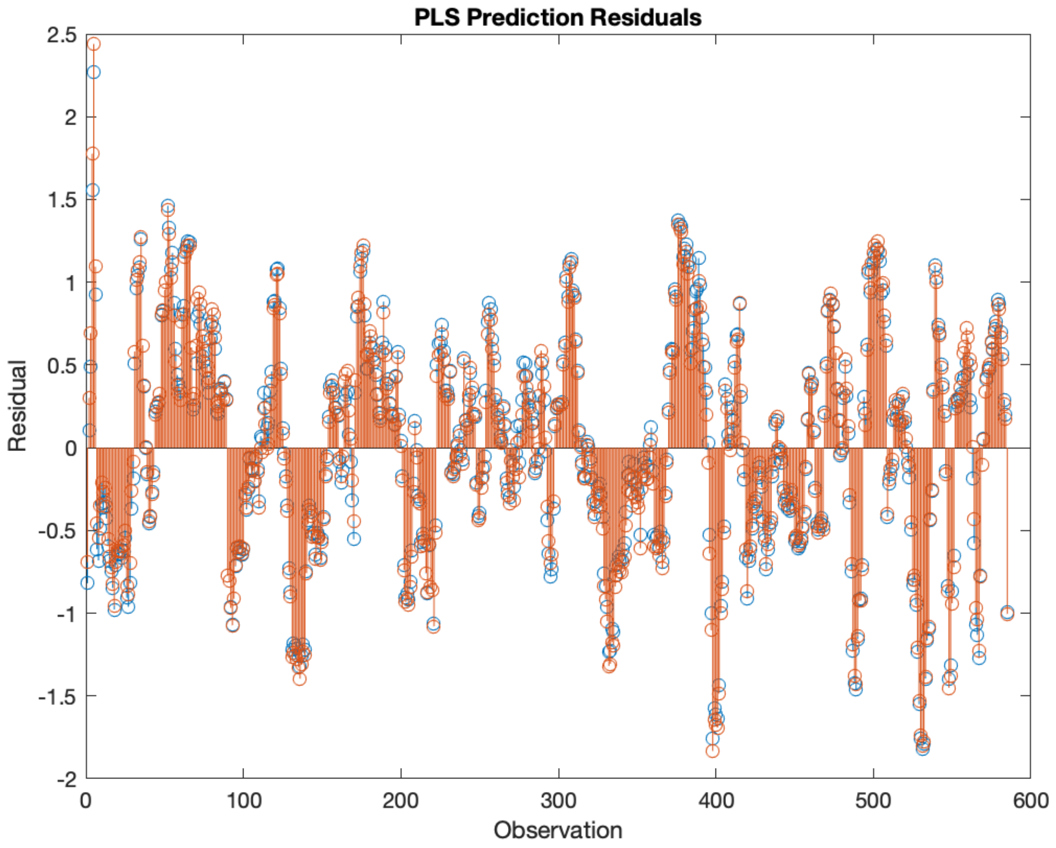


Partial Least Square Testing

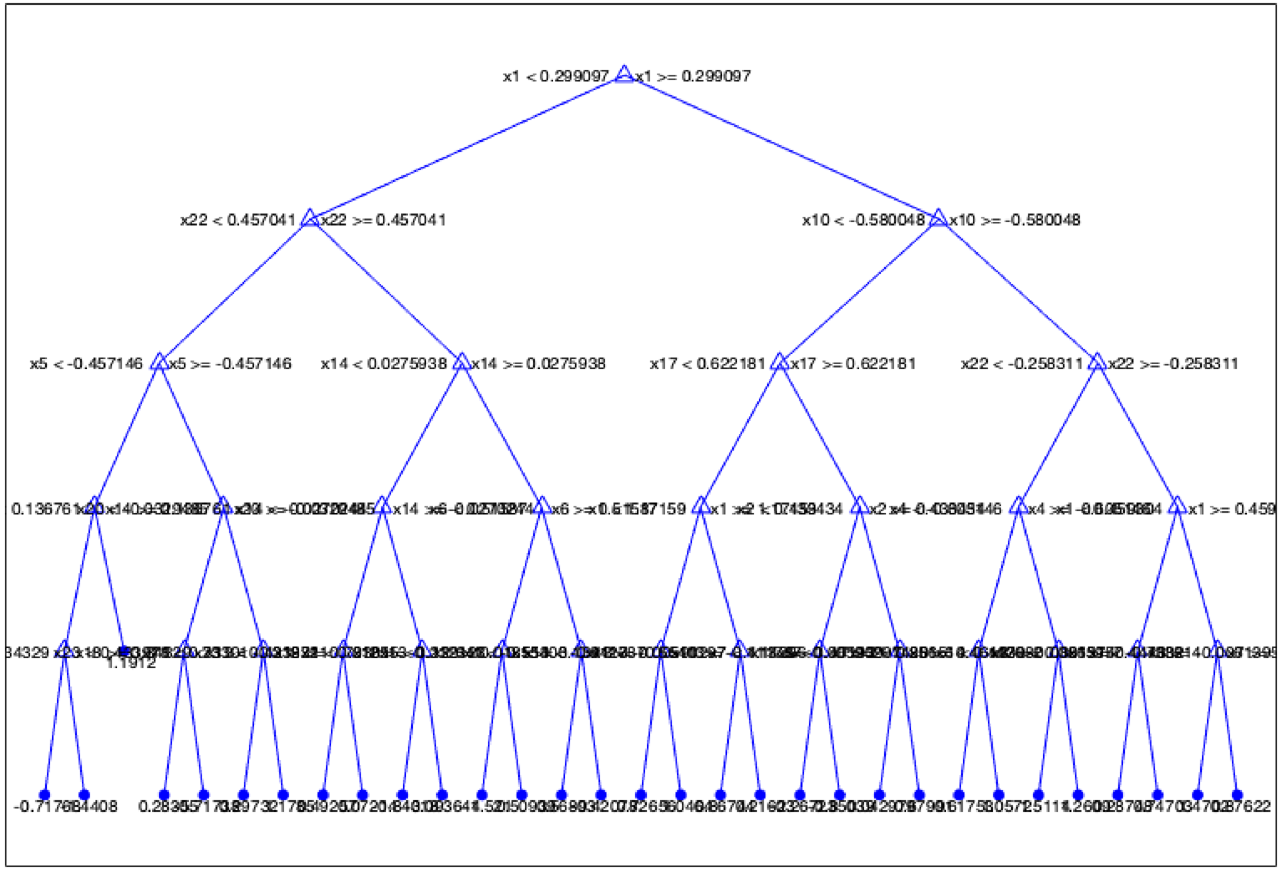


- Regress with to 10 components & 10 fold crossvalidation
- SSE = 257.56
- RMSE = 0.66
- R square = 0.55
- Same accuracy as OLS with fewer components

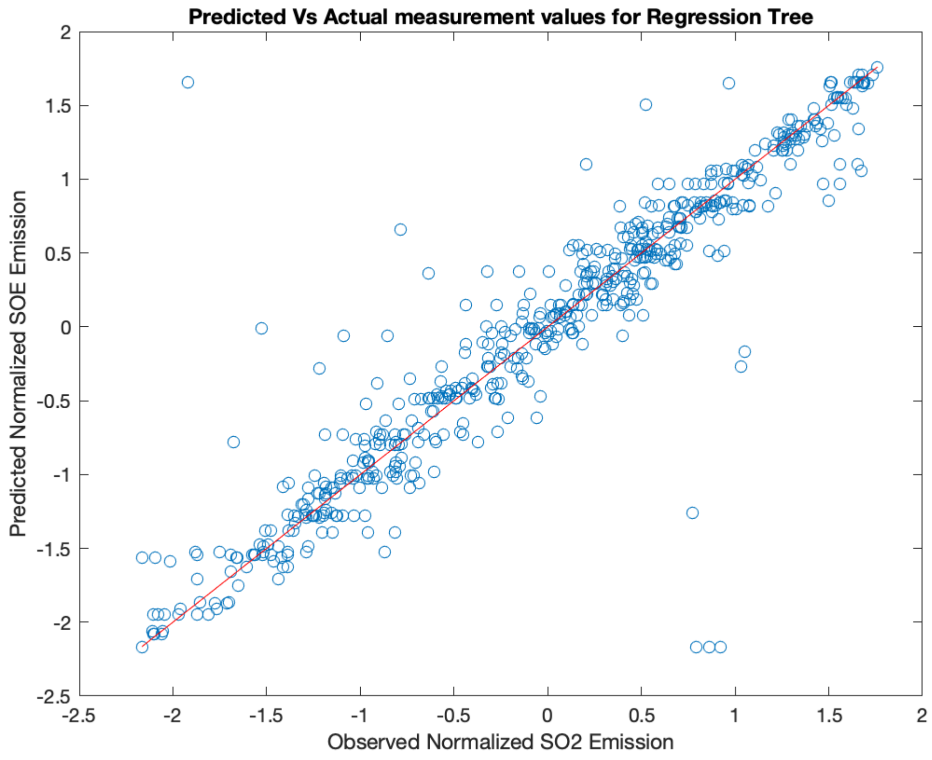
PLS Residuals



Regression Trees

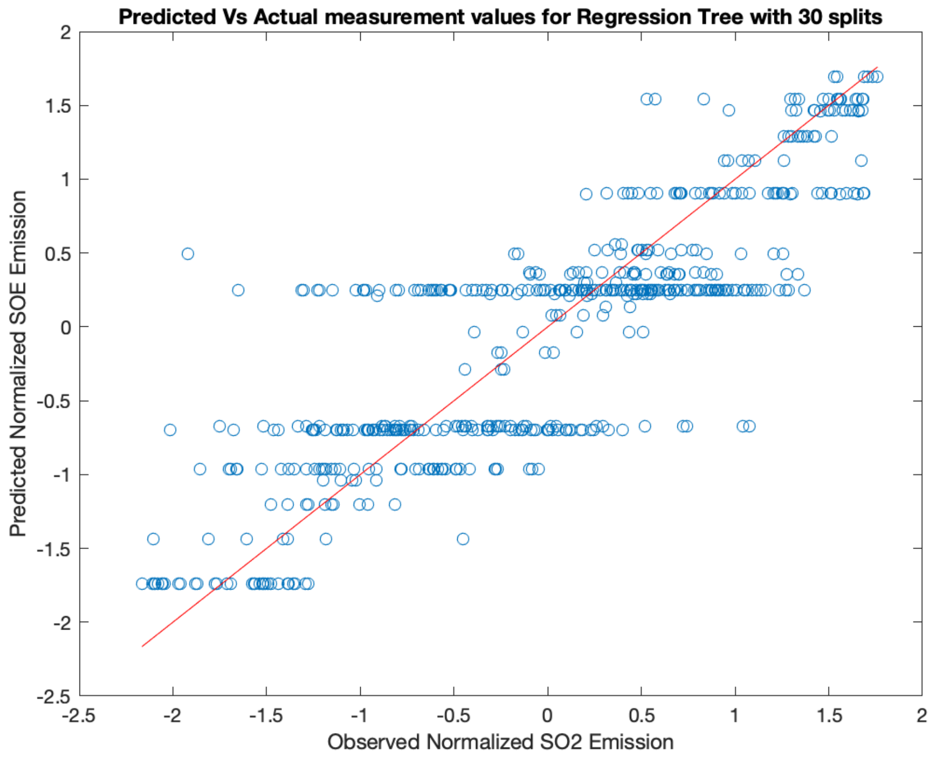


Regression Tree Testing



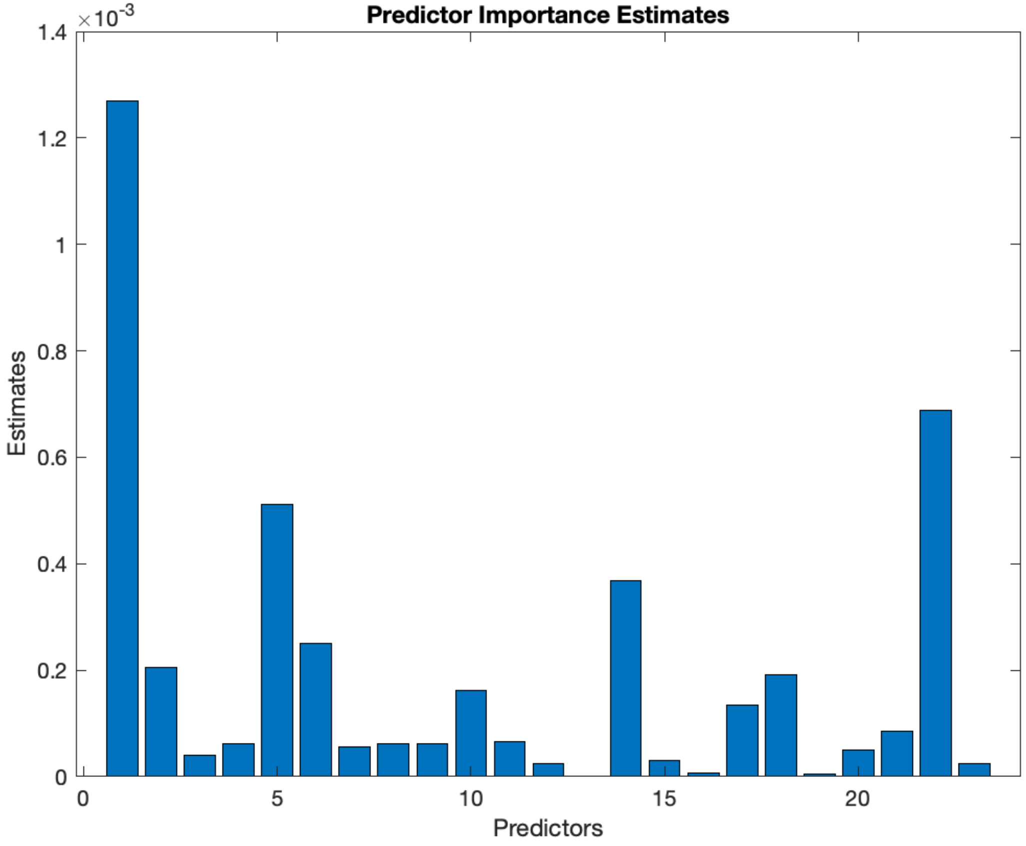
- Regression with all variables with unlimited branches and 10 fold cross validation.
- SSE = 81.5320
- RMSE = 0.3812
- R sq = 0.8583

Regression Tree Testing



- With 30 splits only, 10 fold cross validation, and all variables
- SSE = 150.8188
- RMSE = 0.5126
- R sq = 0.7379

Regression Tree Importance of parameters



Summary of Method Prediction Performance

Method	RMSE	R square	number of variables/ components/ split for trees
OLS	0.6512	0.5866	23
LASSO	0.6695	0.5623	22
PCR	0.6512	0.5866	23
PCR	0.7369	0.4706	10
PLS	0.6699	0.5524	10
Regression Trees	0.3669	0.8583	max
Regression Trees	0.5056	0.7451	30

Root Cause Analysis
Top 6 Variables in beta magnitude

	OLS Beta	PCR Beta	PLS Beta with 10 component	Decision Tree
Intercept		-0.0033	-0.002	
Primary Air Flow	0.13	0.1338	0.1736	1.25E-03
Secondary Air Flow	**0.48**	0.4791	0.3014	8.33E-05
Tertiary Air Flow	-0.09	-0.0891	-0.106	1.24E-05
Primary Windbox Pressure	0.01	0.0063	0.0692	4.35E-05
Secondary Windbox Pressure	0.16	0.1629	0.1517	4.62E-04
Tertiarary Windbox Pressure	0.23	0.2291	0.2278	2.92E-04
Burn Rate	-0.05	-0.0517	-0.2344	1.10E-05
White Liquor Sulfidity	**-1.15**	-1.1528	-0.2019	3.55E-05
Black Liquor Solid 50./50 Test	**-0.56**	-0.5635	-0.5889	1.15E-04
Flue Gas Oxygen	-0.14	-0.1381	-0.2517	6.98E-05
Total Reduced Sulfur	0.11	0.1131	0.1305	5.07E-05
Dry Solid Density Transmitter A	0.04	0.0446	0.2063	1.25E-05
Dry Solid Density Transmitter B	0.32	0.3174	-0.0241	1.12E-05
Primary Air Temperature	**-1.88**	-1.8789	-0.6623	4.04E-04
Secondary Air Temperature	**1.91**	1.9104	0.7832	6.65E-05
Steam Flow Rate	0.14	0.1429	0.103	1.87E-05
Black Liquor Solid Flow Rate	-0.19	-0.1942	-0.0042	8.37E-05
Steam to Dry Solid Ratio	0.28	0.2757	0.3191	1.62E-04
Black Liquor Pressure	-0.07	-0.0679	0.0106	4.80E-05
Black Liquor Indirect Heater Temperature	**1.15**	1.1518	0.1438	1.75E-04
Salt Cake Rotary Feeder Speed	0.16	0.163	0.1107	7.59E-07
Lime Kiln Green Liquor Sulfidity	-0.4	-0.3978	-0.3204	7.19E-04
Nitrogen Oxides correlation	0.02	0.0184	0.0697	1.39E-04

Root Cause Analysis

	SPECTRAL ANALYSIS	REGRESSION	REACHABILITY
PRIMARY AIR FLOW		Yes	
SECONDARY AIR FLOW	Yes	Yes	Yes
TERTIARY AIR FLOW			
PRIMARY WINDBOX PRESSURE			
SECONDARY WINDBOX PRESSURE			
TERTIARARY WINDBOX PRESSURE			
BURN RATE			
WHITE LIQUOR SULFIDITY	Yes	Yes	Yes
BLACK LIQUOR SOLID 50./50 TEST		Yes	
FLUE GAS OXYGEN	Yes		
TOTAL REDUCED SULFUR			
DRY SOLID DENSITY TRANSMITTER A	Yes		
DRY SOLID DENSITY TRANSMITTER B	Yes		
PRIMARY AIR TEMPERATURE	Yes	Yes	Yes
SECONDARY AIR TEMPERATURE	Yes	Yes	Yes
STEAM FLOW RATE			
BLACK LIQUOR SOLID FLOW RATE			
STEAM TO DRY SOLID RATIO			
BLACK LIQUOR PRESSURE	Yes		
BLACK LIQUOR INDIRECT HEATER TEMPERATURE	Yes	Yes	Yes
SALT CAKE ROTARY FEEDER SPEED			
LIME KILN GREEN LIQUOR SULFIDITY	Yes	Yes	
NITROGEN OXIDES CORRELATION			

Conclusion

- Primary air temperature is the root cause of SO2 emission
- Regression tree has the highest accuracy in predicting SO2 emission from data
- Emission can be reduced if the mill increase primary air temperature and lower black liquor indirect heat temperature
- There is no oscillation in SO2 emission in the data provided.

Next Step

- Sensitivity analysis on moving average filter window size
- Random forest maybe able to give better insights into importance of parameters
- Convert normalized prediction back to its units