Graz air pollution

Methods

1. Base additive
2. Base multiplicative
3. Base multicative cutoffs
4. Features (f3)
5. f3_cutoffs
6. Features2 (f4)
7. f4_cutoffs
8. features3
9. crossed_feaures

```
features=['Temp', 'RH', 'Pressure',
          'Winddirection', 'Windspeed', 'Precip']

features2=['weekday_Friday', 'weekday_Monday',
       'weekday_Saturday', 'weekday_Sunday', 'weekday_Thursday',
       'weekday_Tuesday', 'weekday_Wednesday']

features3 = ['season_fall', 'season_spring',
             'season_summer', 'season_winter'
]
```

7. feature cross

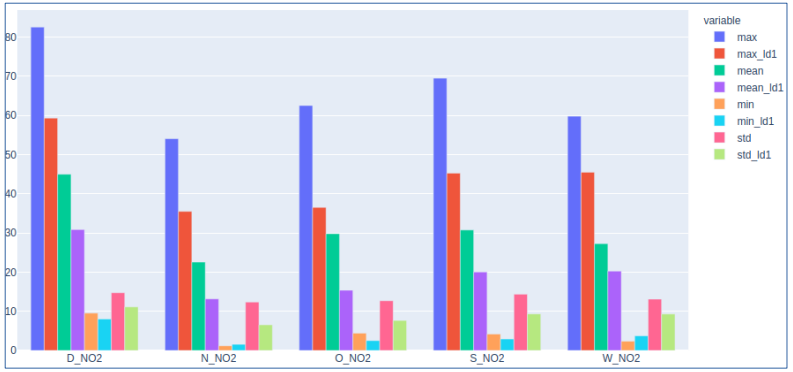
```
weekday_convert = {
    'weekday_Sunday': 0,
    'weekday_Monday': 1,
    'weekday_Tuesday': 2,
    'weekday_Wednesday': 3,
    'weekday_Thursday': 4,
    'weekday_Friday': 5,
    'weekday_Saturday': 6
}
for c in features2:
    df_org[c] = [weekday_convert[c] if i else i for i in df_org[c]]
df_org['dayoftheweek'] = df_org[features2].sum(axis=1)
df_org[['dayoftheweek']]
```

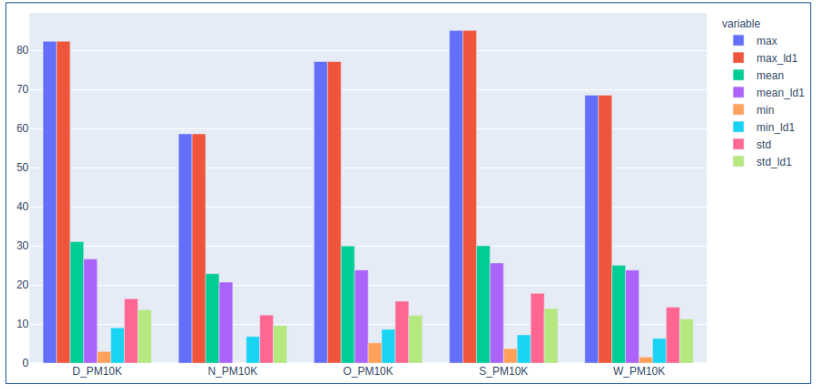
```
season_convert = {
    'season_spring': 1,
    'season_summer': 2,
    'season_fall': 3,
    'season_winter': 4
}

for c in season_convert.keys():
    df_org[c] = [season_convert[c] if i else i for i in df_org[c]]
df_org['season'] = df_org[season_convert.keys()].sum(axis=1)
df_org[['season']]
```

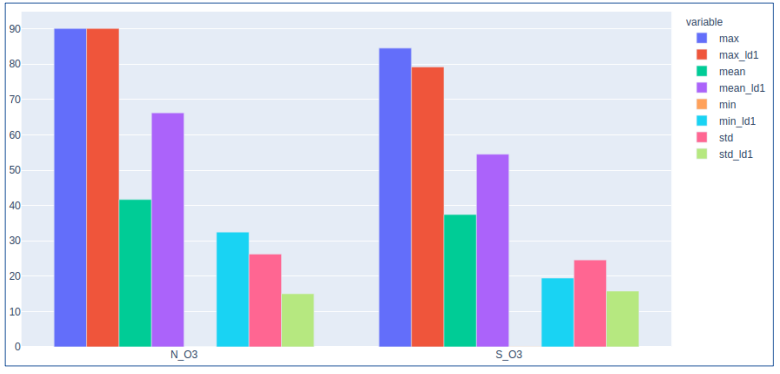
```
df_org['cross1'] = df_org['season']*df_org['dayoftheweek']
df_org['cross2'] = df_org['season']**2 + df_org['dayoftheweek']**2
df_org['cross3'] = df_org['season']**2
df_org['cross4'] = df_org['dayoftheweek']**2
crosses = ['cross1', 'cross2', 'cross3', 'cross4']
```

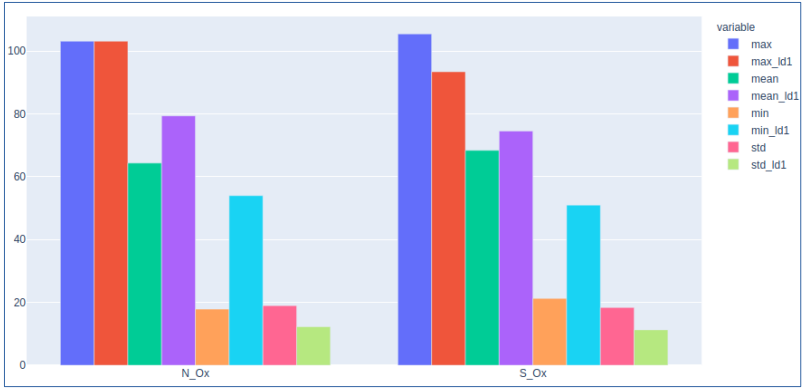
Lockdown1 vs. train

Lockdown1 vs. train

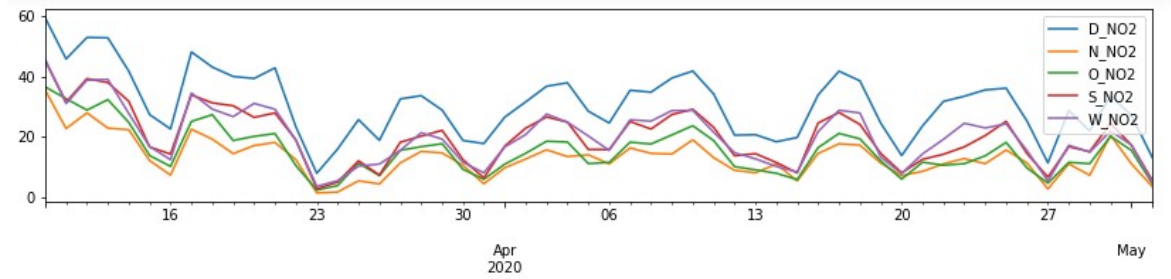


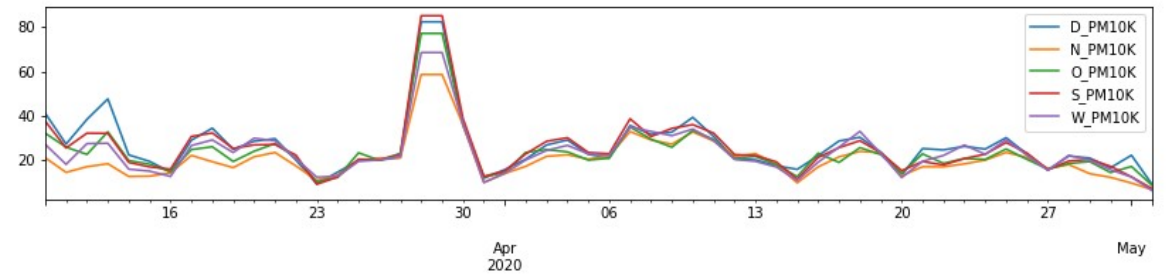


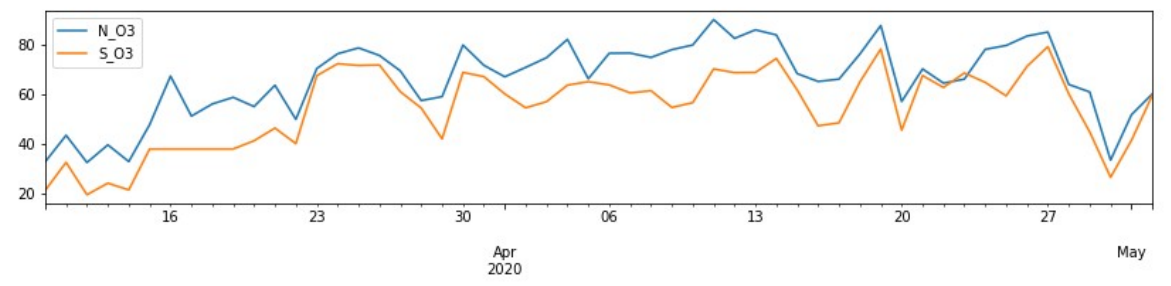


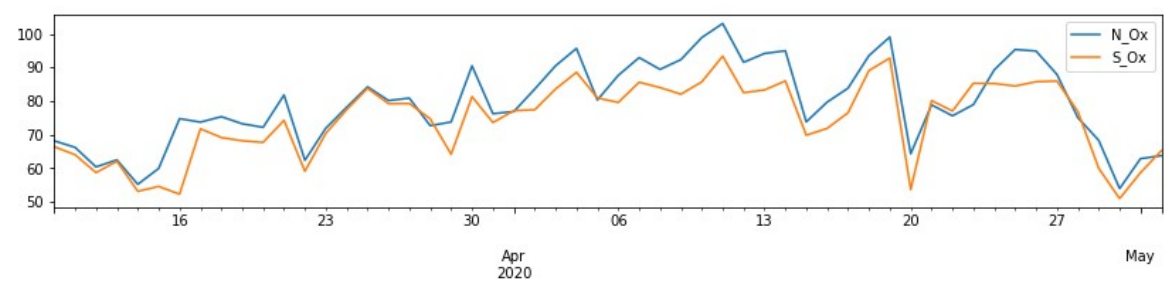


- Test data for NO2 is more simple than PM10
  - max_test ~ max_train


D_NO2
N_NO2
O_NO2
S_NO2
W_NO2
60
40
20
0
16
23
30
06
13
20
27
Apr
2020
May

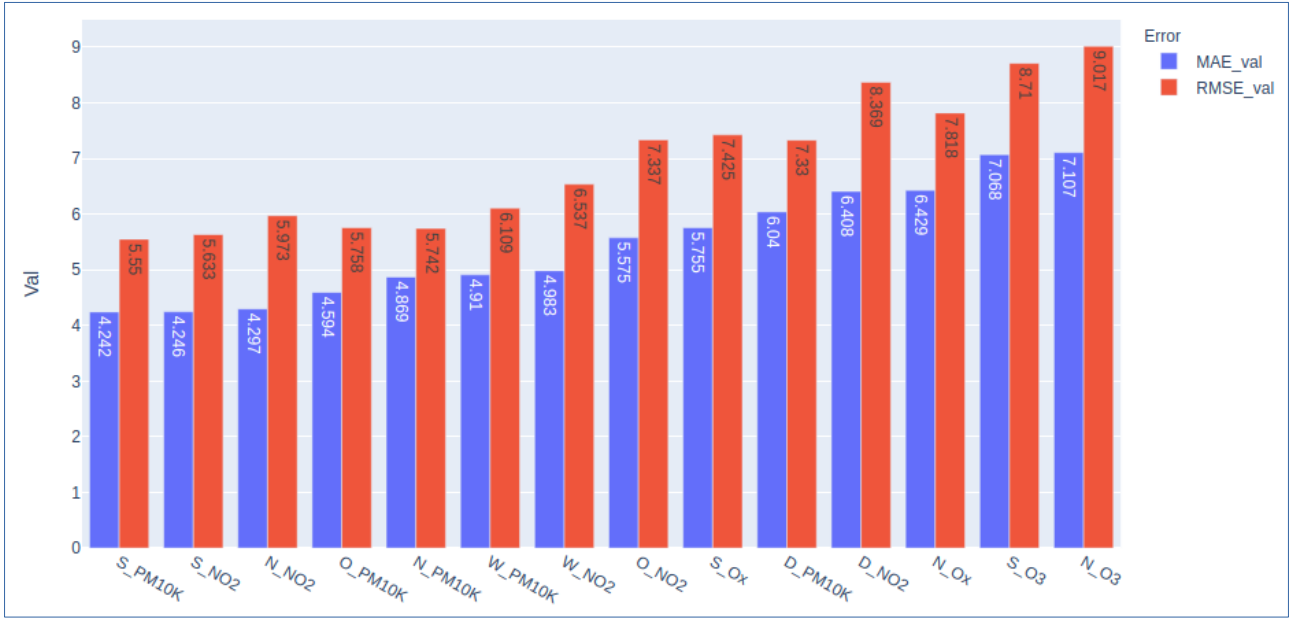

D_PM10K
N_PM10K
O_PM10K
S_PM10K
W_PM10K
80
60
40
20
16
23
30
06
13
20
27
Apr
2020
May


N_O3
S_O3
80
60
40
20
16
23
30
06
13
20
27
Apr
2020
May


N_Ox
S_Ox
100
90
80
70
60
50
16
23
30
06
13
20
27
Apr
2020
May

Min

	RMSE	MAE	RMSE_val	MAE_val
D_NO2	RMSE_f4	MAE_f4	8.368847	6.407923
N_NO2	RMSE_new_base_add	MAE_base_add	5.973273	4.297227
O_NO2	RMSE_new_base_add	MAE_new_base_add	7.337199	5.575182
S_NO2	RMSE_f4	MAE_f4	5.633155	4.246237
W_NO2	RMSE_f4	MAE_f4	6.536686	4.982993
D_PM10K	RMSE_new_base_multi	MAE_new_base_multi	7.329732	6.039981
N_PM10K	RMSE_crossed_f5	MAE_crossed_f5	5.742252	4.869489
O_PM10K	RMSE_new_crossed_f	MAE_new_crossed_f	5.758059	4.593578
S_PM10K	RMSE_new_crossed_f	MAE_new_crossed_f	5.549522	4.242137
W_PM10K	RMSE_new_crossed_f	MAE_new_crossed_f	6.108925	4.909600
N_O3	RMSE_crossed_f5	MAE_features2_multi	9.016616	7.106509
S_O3	RMSE_crossed_f5	MAE_crossed_f5	8.710238	7.068301
N_Ox	RMSE_f3_cutoffs	MAE_f3_cutoffs	7.817738	6.428780
S_Ox	RMSE_tunning_params	MAE_new_crossed_f	7.425076	5.755069



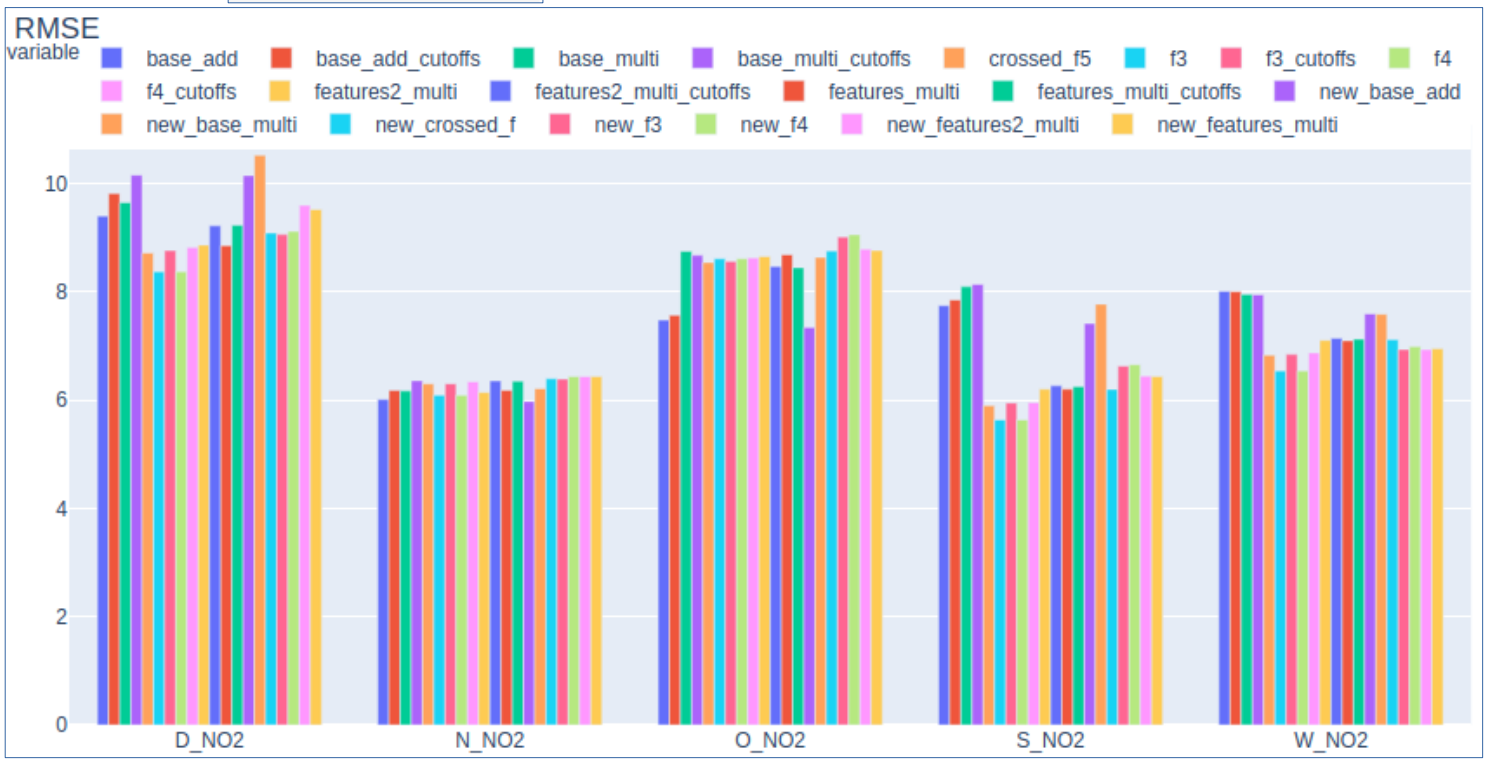
- Methods
  - 50%:50 for new and old data
  - *crossed_features* dominate
    - S_PM10
    - N_PM10
    - O_PM10
    - W_PM10
  - *base_add* fits well for N_NO2, O_NO2 and *base_mul* for D_PM10K
  -

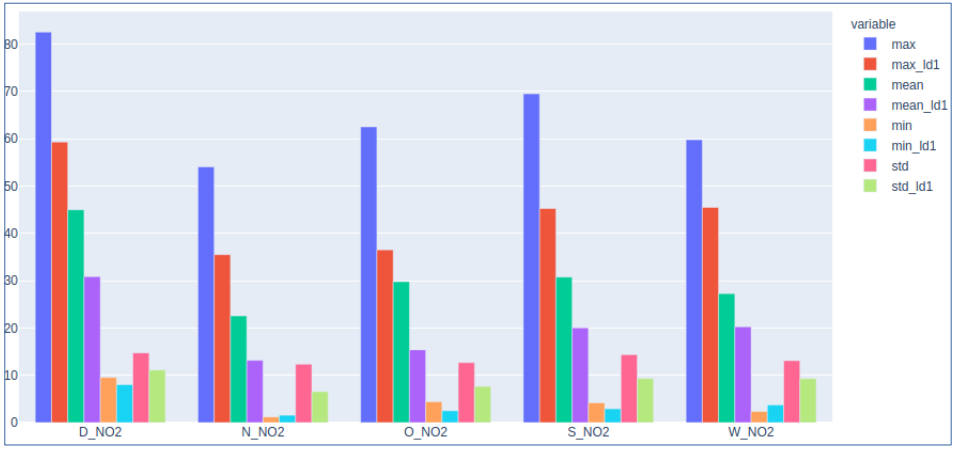
- Values
  - S_PM10K

RMSE

	0	1
D_NO2	8.369	f3
N_NO2	5.973	new_base_add
O_NO2	7.337	new_base_add
S_NO2	5.633	f3
W_NO2	6.537	f3

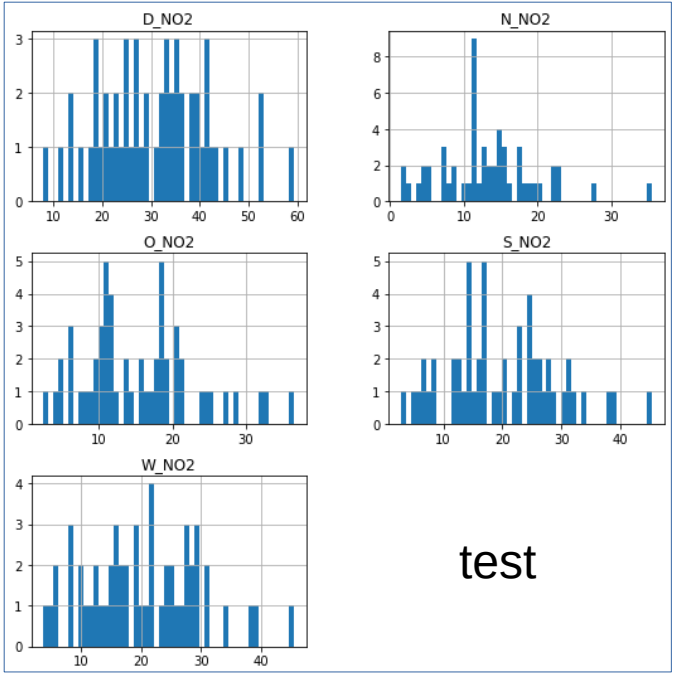
- N_NO2
  - is already low error compared to other NO2
- N_NO2 and O_NO2: more features do not help
  - Best at new base additive
    → *tune the parameters of the algorithm/preprocess more data to get better results*
- D_, S_, W_: best at f3 (=f4) from old data

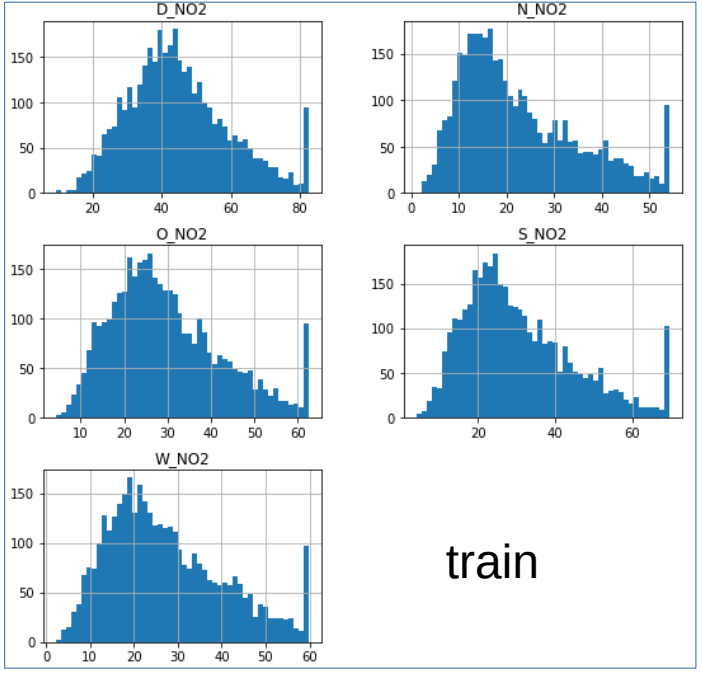


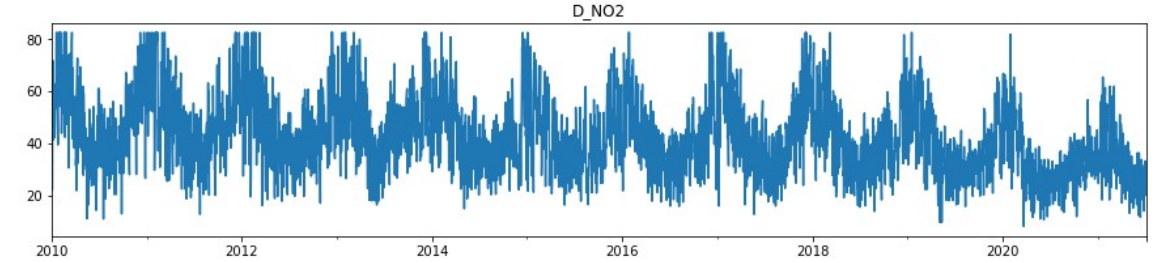


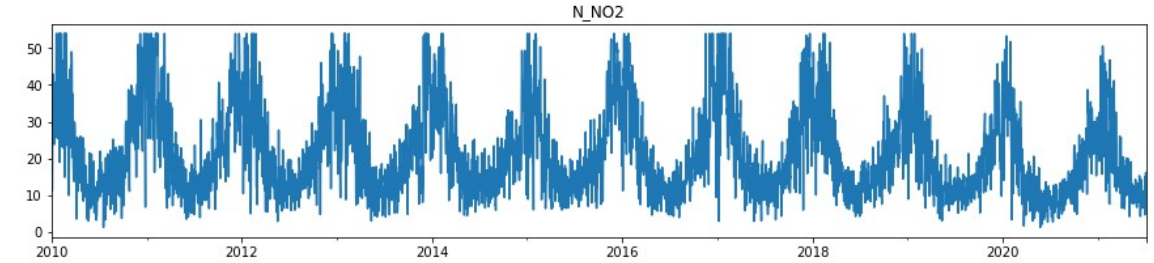
Explanation

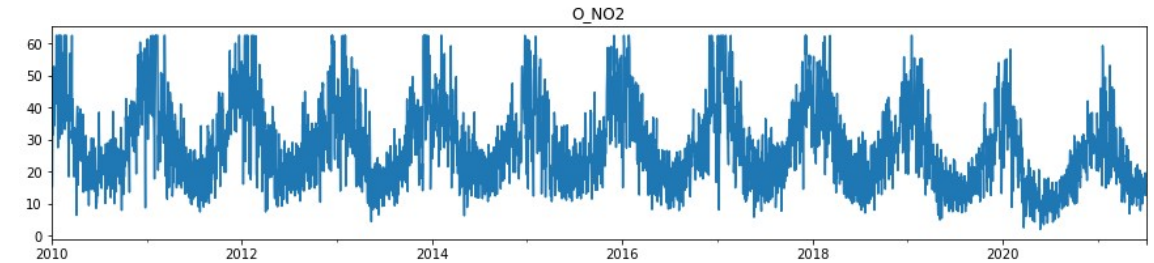
- N_NO2 and S_NO2 with lower error:
- D_NO2 bigger error
-

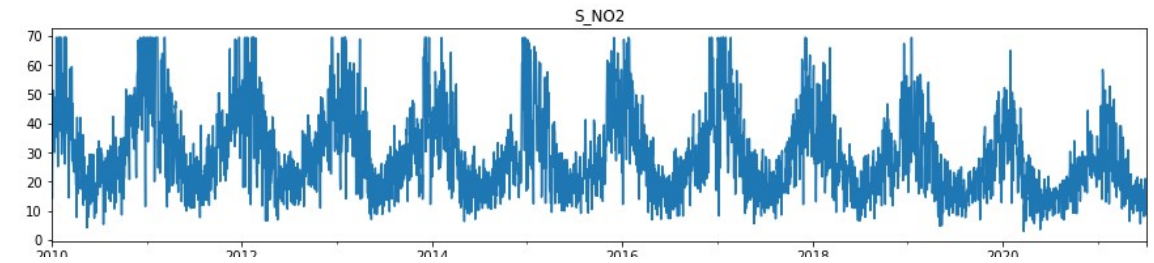












Explanation

RMSE

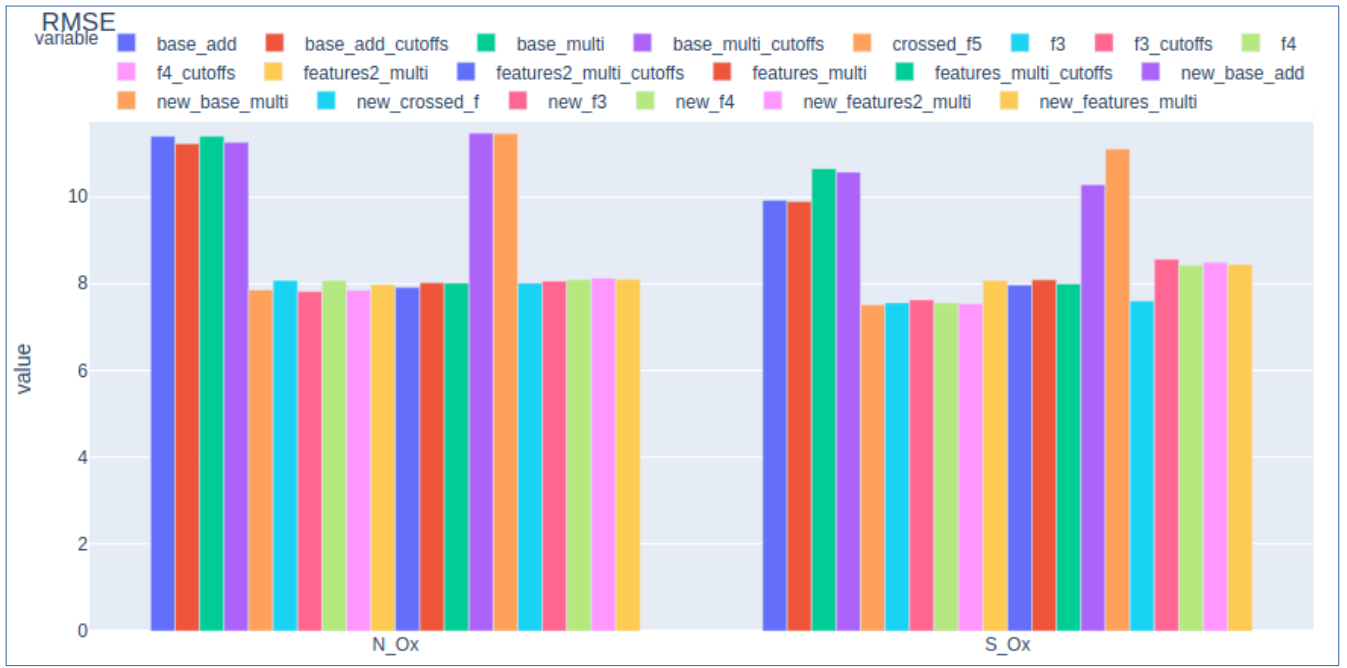
	0	1
D_PM10K	7.330	new_base_multi
N_PM10K	5.742	crossed_f5
O_PM10K	5.758	new_crossed_f
S_PM10K	5.550	new_crossed_f
W_PM10K	6.109	new_crossed_f

RMSE

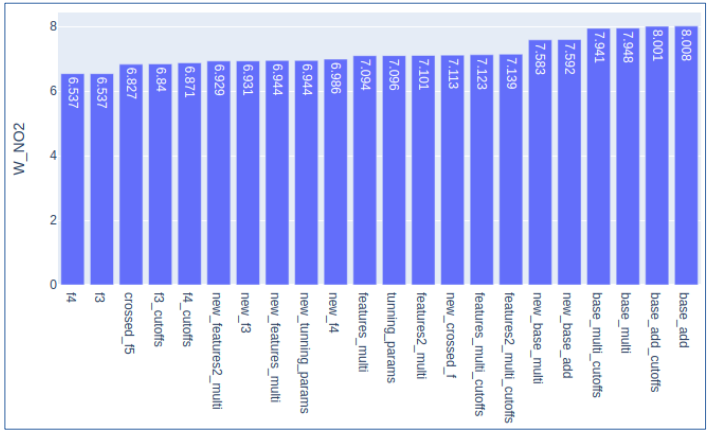
	0	1
N_O3	9.017	crossed_f5
S_O3	8.710	crossed_f5

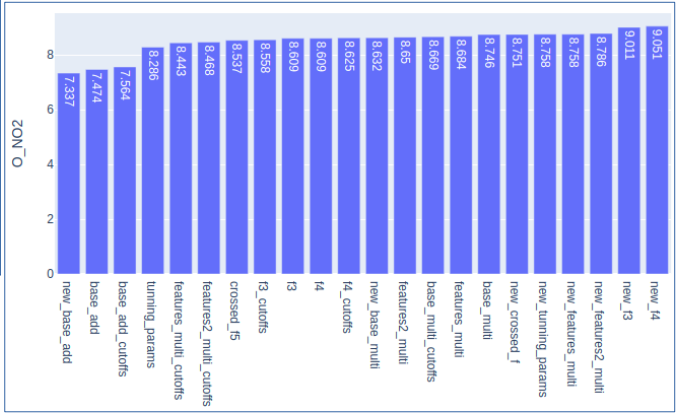
RMSE

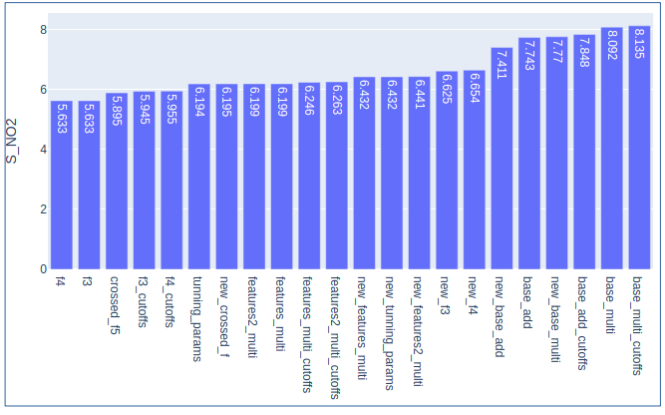
	0	1
N_Ox	7.818	f3_cutoffs
S_Ox	7.510	crossed_f5

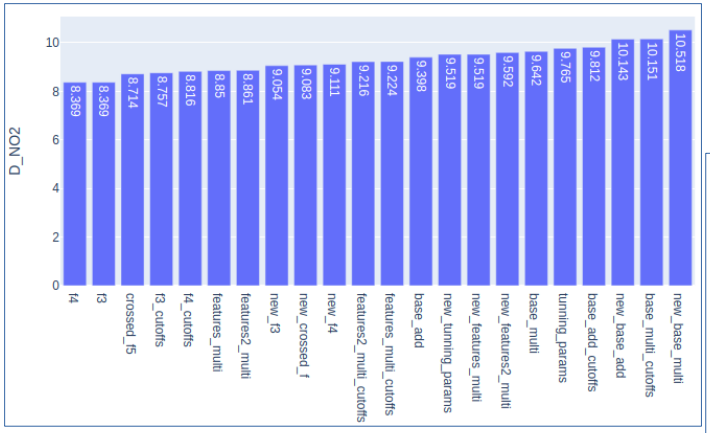


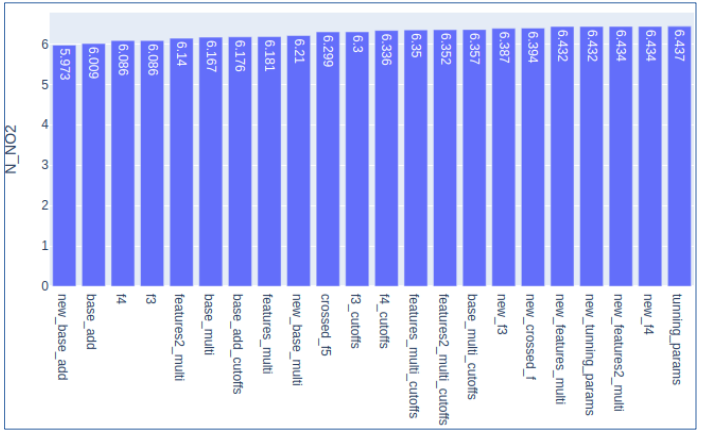
RMSE


W_NO2
base_add 8.008
base_add_cutoffs 8.001
base_multi 7.948
base_multi_cutoffs 7.941
new_base_add 7.592
new_base_multi 7.583
features2_multi_cutoffs 7.139
features_multi_cutoffs 7.123
new_crossed_f 7.113
features2_multi 7.101
tunning_params 7.096
features_multi 7.094
new_f4 6.986
new_tunning_params 6.944
new_features_multi 6.944
new_f3 6.931
new_features2_multi 6.929
f4_cutoffs 6.871
f3_cutoffs 6.84
crossed_f5 6.827
f3 6.537
f4 6.537


O_NO2
new_f4 9.051
new_f3 9.011
new_features2_multi 8.786
new_features_multi 8.758
new_tunning_params 8.758
new_crossed_f 8.751
base_multi 8.746
features_multi 8.684
base_multi_cutoffs 8.669
features2_multi 8.65
new_base_multi 8.632
f4_cutoffs 8.625
f4 8.609
f3 8.609
f3_cutoffs 8.558
crossed_f5 8.537
features2_multi_cutoffs 8.468
features_multi_cutoffs 8.443
tunning_params 8.286
base_add_cutoffs 7.564
base_add 7.474
new_base_add 7.337


S_NO2
base_multi_cutoffs 8.135
base_multi 8.092
base_add_cutoffs 7.848
new_base_multi 7.77
base_add 7.743
new_base_add 7.411
new_f4 6.654
new_f3 6.625
new_features2_multi 6.441
new_tunning_params 6.432
new_features_multi 6.432
features2_multi_cutoffs 6.263
features_multi_cutoffs 6.246
features_multi 6.199
features2_multi 6.199
new_crossed_f 6.195
tunning_params 6.194
f4_cutoffs 5.965
f3_cutoffs 5.945
crossed_f5 5.895
f3 5.633
f4 5.633


D_NO2
new_base_multi 10.518
base_multi_cutoffs 10.151
new_base_add 10.143
base_add_cutoffs 9.812
tunning_params 9.765
base_multi 9.642
new_features2_multi 9.592
new_features_multi 9.519
new_tunning_params 9.519
base_add 9.398
features_multi_cutoffs 9.224
features2_multi_cutoffs 9.216
new_f4 9.111
new_crossed_f 9.083
new_f3 9.054
features2_multi 8.861
features_multi 8.85
f4_cutoffs 8.816
f3_cutoffs 8.757
crossed_f5 8.714
f3 8.369
f4 8.369


N_NO2
tunning_params 6.437
new_f4 6.434
new_features2_multi 6.434
new_tunning_params 6.432
new_features_multi 6.432
new_crossed_f 6.394
new_f3 6.387
base_multi_cutoffs 6.357
features2_multi_cutoffs 6.352
features_multi_cutoffs 6.35
f4_cutoffs 6.336
f3_cutoffs 6.3
crossed_f5 6.299
new_base_multi 6.21
features_multi 6.181
base_add_cutoffs 6.176
base_multi 6.167
features2_multi 6.14
f3 6.086
f4 6.086
base_add 6.009
new_base_add 5.973

Next Steps

1. Cutoffs
2. Optimization for each area
3. Lockdown 2, 3?

Thank you!