

**산업통상자원부 공공데이터 활용  
비즈니스아이디어 공모전 분석 결과 제출**

**화력발전 배연탈황 공정의 석회석 투입량  
저감 최적화 모델 개발**

이원, 이유송

## 1 명칭

### - 화력발전 배연탈황 공정의 석회석 투입량 저감 최적화 모델 개발

## 2 제안배경

최근 태양광, 풍력 등의 친환경 에너지의 발전에 따라 원래 지속적으로 사용되던 전통적인 화력발전소의 환경 문제에 대한 관심이 커지고 있다. 현재 많은 화력발전소에서 발생하는 황산화물과 이에 따른 미세먼지의 2차 생성을 우려해 우리나라도 환경부 기준 황산화물 배출기준 허용치를 정해 환경 문제를 일으키는 황산화물의 배출 기준을 지정하였다.

그 중 아황산가스는 토양 등의 산성화에 영향을 미치며, 식물의 잎맥 손상을 일으키고, 구조물의 부식을 촉진한다. 또한, 고농도 아황산가스와 함께 미세먼지에 노출될 경우 호흡기 질환과 심장혈관 질환을 악화시킬 수 있다. 아황산가스가 처리 과정 없이 배출될 경우 대기 중의 수증기와 직접 반응하여 산성비가 될 수 있다.

황분이 많은 중유나 석탄을 대량으로 연소시키는 화력발전소나, 공장 등의 대형 보일러의 배연에는 황산화물 함유되어 있다. 이러한 황산화물을 처리하여 기준치 이하로 배출하기 위한 방법으로는 배연탈황 공정이 있다. 배연탈황 공정은 흡수, 산화 및 환원 등의 원리를 이용한 대표적인 황산화물 제거 기술이다.

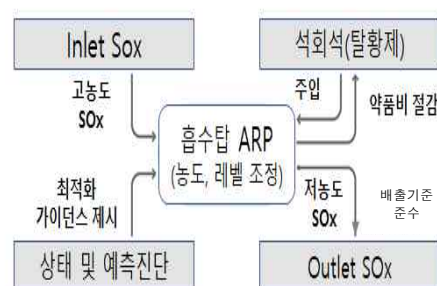
현재 배연탈황 공정은 흡수탑으로 유입된 가스가 흡수, 산화 후 석회석과 반응하며, 이때 생성된 슬러지는 석고로 생성되는 과정으로 구성된다.

1단계: 황산화물 흡수과정( $\text{SO}_2 + \text{H}_2\text{O} \leftrightarrow \text{H}_2\text{SO}_3$ )

2단계: 산화 과정( $\text{H}_2\text{SO}_3 + 1/2\text{O}_2 \rightarrow \text{H}_2\text{SO}_4$ )

3단계: 석회석과 반응하여 중화 과정( $\text{H}_2\text{SO}_4 + \text{CaCO}_3 \rightarrow \text{CaSO}_4 + \text{H}_2\text{O} + \text{CO}_2$ )

4단계: 석고 침전( $\text{CaSO}_4 + 2\text{H}_2\text{O} \rightarrow \text{CaSO}_4 + \text{CO}_2$ )



이러한 배연탈황 공정 중 3단계에서 일어나는 중화 과정에서 투입되는 석회석을 유입된 가

스에 맞게 석회석의 투입량을 저감할 수 있는 방법을 모색해 보았으며 석회석 최소화 및 투입력 저감을 위한 석회석 투입량 저감 최적화 모델을 개발하였다.

석회석 투입량 최적화 모델을 통해 투입량을 절감하여 경제성을 올리고 배연탈황 과정에서 부산물인 석고의 순도 품질을 향상시키는 것을 목표로 하고 있다.

### ③ 분석 내용 및 분석 결과

분석의 목표는 주어진 흡수탑 데이터를 이용해 유입된 가스에 따른 탈황제의 양을 절감 및 최적화 하는 것이다. 이를 위해서 각 데이터 간의 상관관계를 이해할 필요가 있다.

제공되는 흡수탑 데이터는 3, 4, 5, 6 호기 총 4개의 호기로 이루어져 있으며 각 호기마다 날짜, 발전량, 유출가스, 배출정상여부, 유입가스, 흡수탑의 pH농도, 투입된 탈황제의 양 으로 이루어져 있다.

(그림1 : 흡수탑 데이터)

	date	power_3	outlet_3	normal_3	power_4	outlet_4	normal_4	power_5	outlet_5	normal_5	...	inlet_5	inlet_6	ph_3	ph_4	ph_5	ph_6
0	2023-09-01 0:00	474	0.68	정상	0	1.14	정상	456	15.06	정상	...	185.770000	240.870000	5.20	4.94	6.07	6.13
1	2023-09-01 0:05	486	0.36	정상	0	1.61	정상	473	15.52	정상	...	185.770000	244.570000	5.11	4.86	6.05	6.12
2	2023-09-01 0:10	510	0.4	정상	0	1.74	정상	470	15.71	정상	...	NaN	228.000000	4.96	4.76	6.04	6.11
3	2023-09-01 0:15	518	1.74	정상	0	2.04	정상	454	16.69	정상	...	185.770000	233.270000	4.80	4.64	6.02	6.11
4	2023-09-01 0:20	518	1.52	정상	0	1.49	정상	434	16.13	정상	...	NaN	243.460000	5.06	4.53	6.04	6.11
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
53275	2024-02-29 23:35	418	5.22	정상	0	6.82	정상	481	19.46	정상	...	335.489831	2.521914	6.04	4.83	5.18	6.98
53276	2024-02-29 23:40	418	4.99	정상	0	7.44	정상	482	20.21	정상	...	332.655682	2.531338	6.04	4.83	5.17	6.97
53277	2024-02-29 23:45	418	4.86	정상	0	6.83	정상	476	20.24	정상	...	331.522790	2.450782	6.04	4.83	5.18	6.97
53278	2024-02-29 23:50	419	4.8	정상	0	7.09	정상	473	20.25	정상	...	313.757489	2.571517	6.04	4.83	5.18	6.97
53279	2024-02-29 23:55	420	5.01	정상	0	6.86	정상	470	18.76	정상	...	165.986595	2.433664	6.04	4.83	5.18	6.97

53280 rows x 25 columns

흡수탑 데이터의 전처리를 통해 결측값과 특수문자를 제거하였다. 이를 통해 53280개 였던 데이터가 44627개가 되었다. 또한, 배출이 정상적으로 이루어지는 데이터와 유입가스와 탈황제의 값이 ^음수가 아닌 데이터^만을 추렸다.

(1. 결측값과 특수문자 제거 2. 정상값이랑 음수값 제거 사진 3. 데이터 개수 감소 사진 첨부)

```
df_komipo = df_komipo[~df_komipo.astype('str').apply(lambda row: '-' in row.values, axis = 1)]
df_komipo.dropna(inplace = True)
```

44627 rows x 25 columns

### 3.1 탐색적 분석(기술통계 분석)

흡수탑 데이터 파악을 위해 탐색적 분석(EDA)을 진행하였다. 먼저 각 흡수탑에 기술 통계를 이용하여 데이터의 이상값을 확인하였다. 기술 통계 값의 확인 결과 배출여부가 정상임에도 불구하고 3호기와 5호기에서는 유출가스가 배출제한기준을 넘는값이 발견되었다. 또한, 4호기에서는 전체 데이터에서 발전량이 0으로 측정되었다. 그래서 4호기의 데이터는 비정상적이므로 이번 분석에서 제외하였다.

(그림2 : 기술 통계 사진)

3호기 기술통계

	count	date	power_3	outlet_3 \
mean	2023-12-05	13:51:14.115505664	41210.000000	41210.000000
min		2023-09-01 00:00:00	457.241349	8.010469
25%		2023-10-13 09:06:15	1.000000	0.000000
50%		2023-12-12 06:22:30	496.000000	2.470000
75%		2024-01-20 21:48:45	519.000000	3.610000
max		2024-02-29 23:55:00	524.000000	4.890000
std			1267.600000	32.564189

	inlet_3	ton_3	ph_3
count	41210.000000	41210.000000	41210.000000
mean	199.554032	7.669369	5.989481
min	0.001667	0.000000	4.360000
25%	172.550278	4.640000	5.670000
50%	202.607222	6.469587	6.050000
75%	244.806762	9.602199	6.210000
max	402.878667	72.261812	9.570000
std	70.804215	5.945807	0.666262

4호기 기술통계

	count	date	power_4	outlet_4	inlet_4 \
mean	2023-12-11	20:18:30.838050048	37560.0	37560.000000	37560.000000
min		2023-09-01 00:00:00	0.0	0.000000	0.005032
25%		2023-11-07 21:22:30	0.0	0.000000	67.255154
50%		2023-12-19 23:42:30	0.0	1.250000	207.040697
75%		2024-01-24 06:51:15	0.0	2.870000	283.655381
max		2024-02-29 23:55:00	0.0	409.070000	17115.117810
std			NaN	0.0	5.899987

	ton_4	ph_4
count	37560.000000	37560.000000
mean	5.615441	5.283784
min	0.000000	3.100000
25%	0.000000	4.460000
50%	4.653875	4.600000
75%	7.150000	6.100000
max	67.110000	9.120000
std	4.506379	1.217492

5호기 기술통계

	count	date	power_5	outlet_5 \
mean	2023-12-04	18:03:29.990138800	42592.000000	42592.000000
min		2023-09-01 00:00:00	252.910946	64.650319
25%		2023-10-13 22:58:45	0.000000	0.000000
50%		2023-12-11 04:37:30	0.000000	10.880000
75%		2024-01-19 16:06:15	402.000000	17.690000
max		2024-02-29 23:55:00	473.000000	48.780000
std			535.000000	1684.000000

	inlet_5	ton_5	ph_5
count	42592.000000	42592.000000	42592.000000
mean	176.153545	5.257743	6.093927
min	2.269655	0.013841	0.000000
25%	3.873271	0.023663	5.400000
50%	202.493296	3.568784	5.720000
75%	250.937939	9.926504	7.090000
max	1020.517566	40.060000	7.750000
std	177.295339	6.034058	0.896984

6호기 기술통계

	count	date	power_6	outlet_6 \
mean	2023-12-09	17:06:26.493973248	40071.000000	40071.000000
min		2023-09-01 00:00:00	1.000000	0.000000
25%		2023-10-24 11:50:00	432.000000	12.490000
50%		2023-12-16 06:00:00	472.000000	18.490000
75%		2024-01-22 06:22:30	475.000000	23.050000
max		2024-02-29 23:55:00	547.000000	78.320000
std			166.881409	7.602777

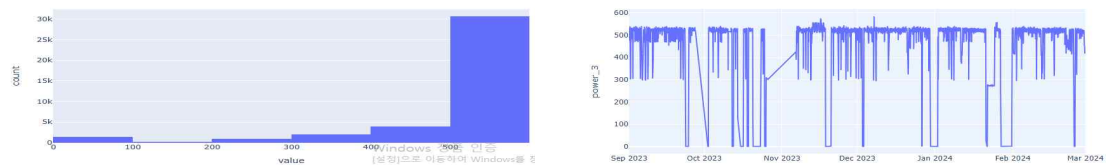
	inlet_6	ton_6	ph_6
count	40071.000000	40071.000000	40071.000000
mean	225.116209	12.617490	5.626828
min	0.000000	0.000000	3.300000
25%	195.101700	0.049944	5.310000
50%	242.743729	6.796316	5.500000
75%	295.567863	18.619363	5.790000
max	570.277212	40.337962	8.250000
std	109.943766	14.766050	0.583040

### 3.2 이상치 제거 및 발전량 제한 기준 설정

기술통계로 파악한 이상값을 제거하기 위해 대쉬보드를 통해 데이터를 시각화하였다. 먼저, 날짜별 발전량 그래프 확인 시 각호기의 발전량의 움직임이 공통적으로 발전량이 300Mw 이상 구간과 0Mw 구간으로 구분되었다. 또한, 발전량의 히스토그램을 확인하면 분포가 대부분 값이 200Mw 이상부터 존재함을 알 수 있다.

(그림3 : 값을 제한하지 않은 날짜별 발전량 그래프, 발전량 히스토그램 사진)

3호기 발전량 히스토그램, 날짜별 발전량 그래프



5호기 발전량 히스토그램, 날짜별 발전량 그래프

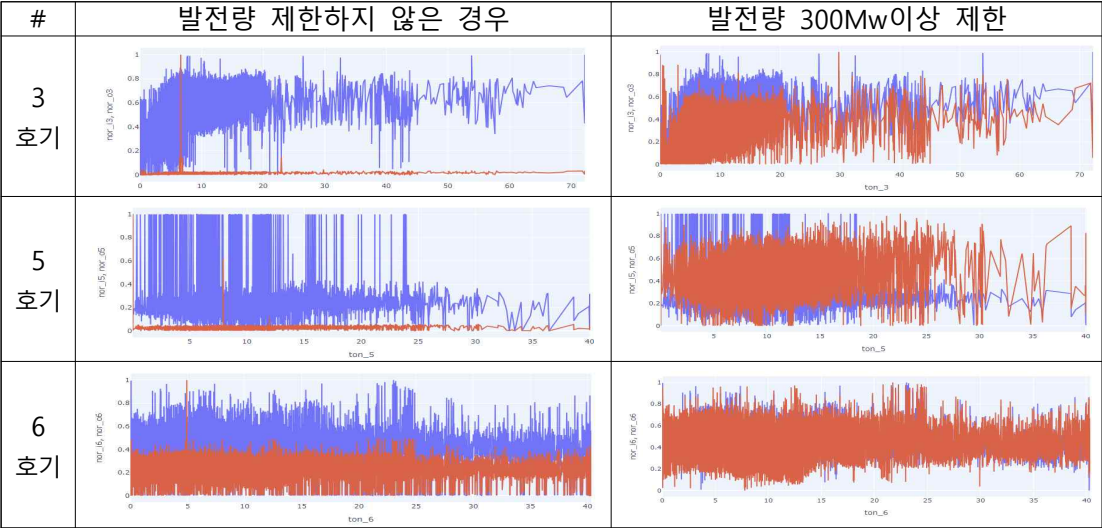


6호기 발전량 히스토그램, 날짜별 발전량 그래프



위 그래프의 결과를 통해 발전 기동정지 값을 제외하기 위해 **발전량의 제한값을 300Mw로** 기준을 두었다. 발전량 제한 값을 통해 실제 이상값이 존재하는 유출가스에 대한 변화를 확인하였다.

(그림4 : [발전량 제한하지 않은 경우, 발전량 제한]-탈황제에 따른 inlet, outlet(정규화) 그래프)



발전량 제한전의 그래프를 보면 유출가스에서 유입가스보다 높은 값을 가지는 데이터가 존재함을 알 수 있다. **제한하지 않은 그래프와 제한한 그래프를 비교해보면 이상값이 제거된 것을 확인할 수 있다.** 유입된 Inlet SOx가 탈황제를 투입하여 반응한 Outlet SOx 배출 시간 차이가 문의 결과 10분여 발생한다고 하여 **발전량, Inlet SOx, 탈황제를 투입량과 10분후 Outlet SOx 값으로 분석 데이터셋을** 확정하였다.

(그림5 : [최종확정 분석데이터셋 3호기 예시])

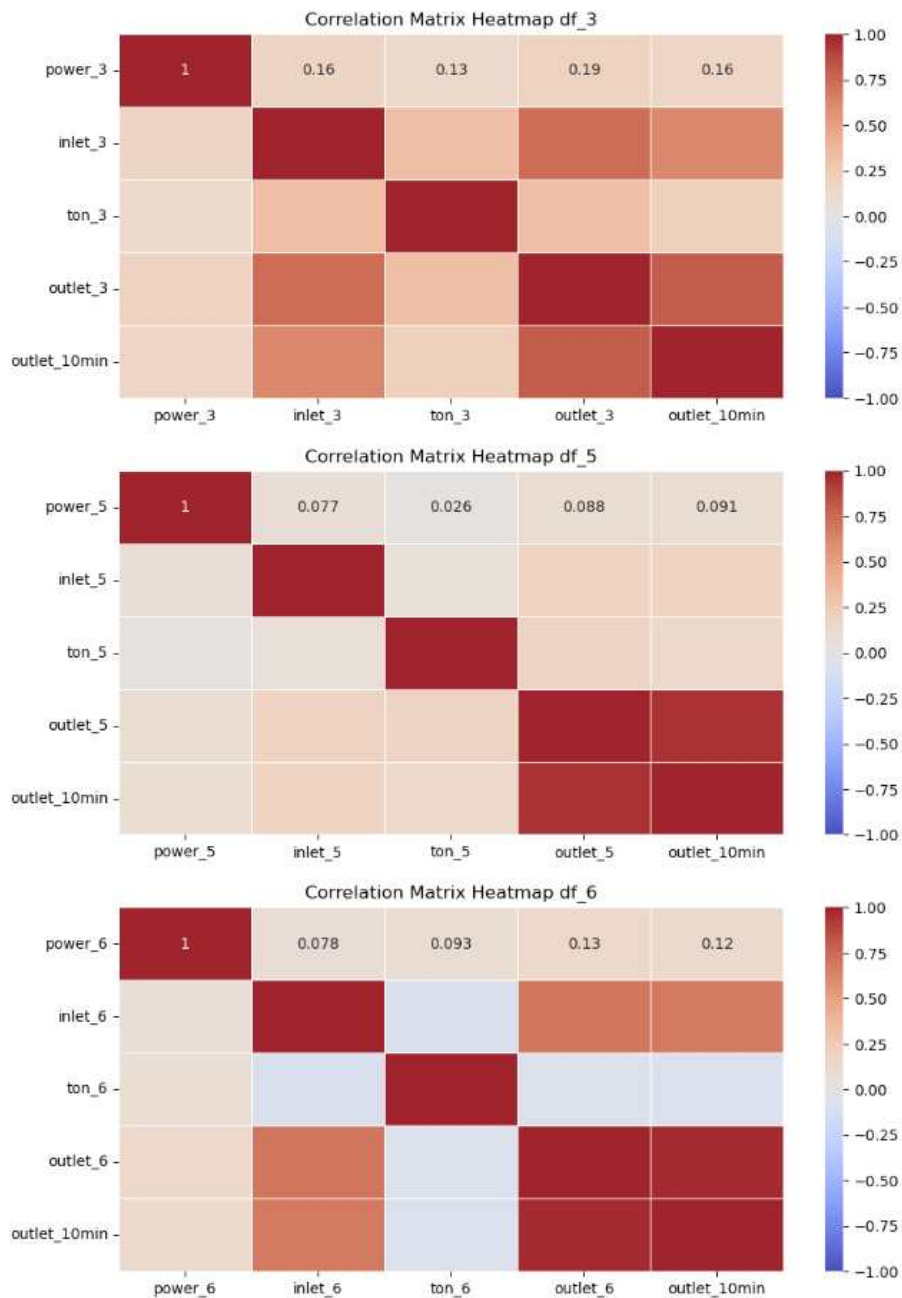
	date	power_3	inlet_3	ton_3	normal_3	outlet_3	outlet_10min
0	2023-09-01 00:00:00	474.0	189.45	5.71	정상	0.68	0.40
1	2023-09-01 00:05:00	486.0	188.30	3.07	정상	0.36	1.74
2	2023-09-01 00:10:00	510.0	186.30	3.43	정상	0.40	1.52
3	2023-09-01 00:15:00	518.0	188.13	10.62	정상	1.74	0.82
4	2023-09-01 00:20:00	518.0	191.39	11.99	정상	1.52	0.83

### 3.3 발전량 구간화를 통한 탈황제의 추이

#### 3.3.1 분석 데이터셋 변수간 상관관계 분석

머신러닝 알고리즘을 통해 모델을 개발하려 한다. 이를 위해 데이터간의 상관성 파악을 위한 상관계수와 히트맵을 확인했다.

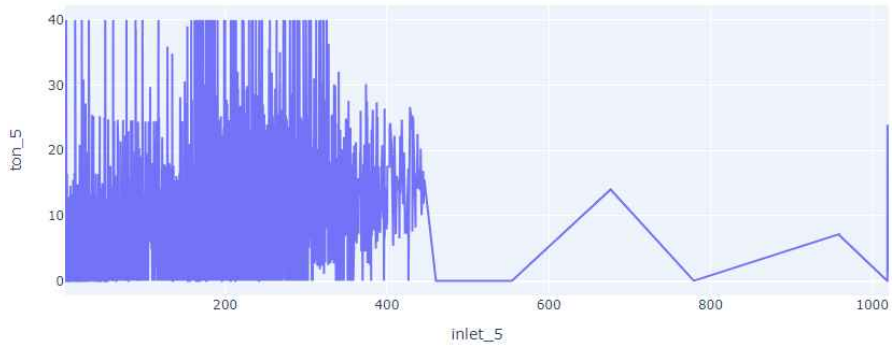
(그림6 : 3, 5, 6호기의 SOx 유입·유출량, 탈황제투입량 관계 히트맵)



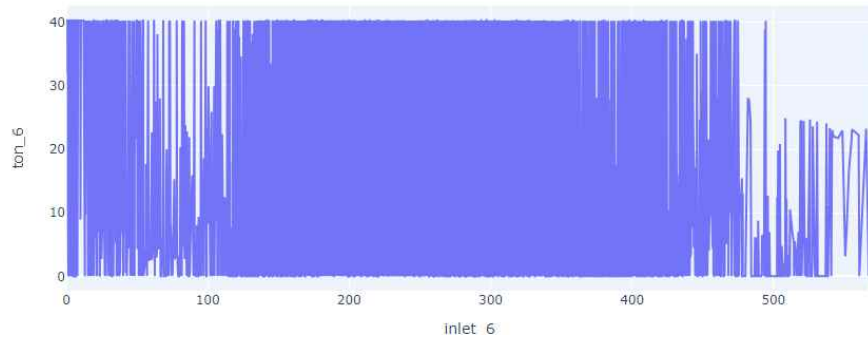
300Mw이상 발전량을 기준으로 각호기의 상관계수와 히트맵을 확인한 결과 알고리즘에 사용하려 했던 유입 가스, 유출 가스, 탈황제 데이터간의 상관계수가 5, 6호기에서 특히 낮아 선형적인 특성을 이용한 알고리즘을 사용할 수 없다는 것을 파악했다.

(그림7 : 5, 6호기의 SOx 유입량 대비 탈황제투입량 그래프)

Daily for ton\_5 (5gg)



Daily for ton\_6 (6gg)

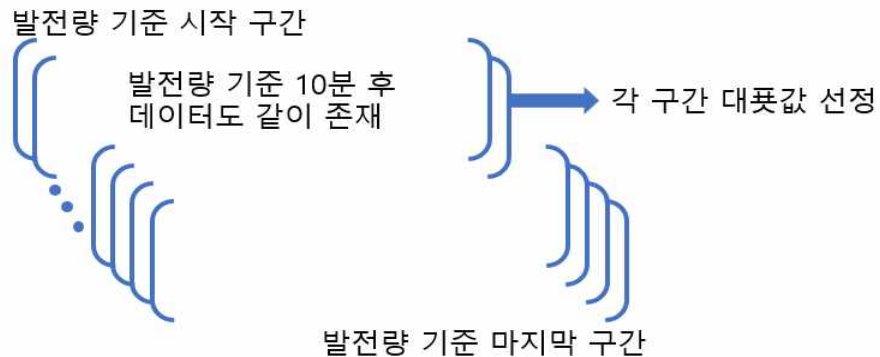


5, 6호기의 SOx 유입량 대비 탈황제투입량 그래프에서 보듯 Inlet SOx의 유입량과의 관계 없이 탈황제 투입이 이루어진 것으로 보여 SOx 유입량 또는 발전량을 단계별로 나누어 탈황제 투입하는 방식으로 결정 하였다.

### 3.3.2 발전량 단계별 탈황제 투입량 대푯값 구하기

그림5의 각호기의 데이터를 이용하여 발전량을 기준으로 구간을 나누고 대푯값을 찾아 추이를 분석하였다. 적절한 구간의 개수 및 대푯값을 구하기 위해 구간을 10, 30, 50, 70개로 나누었다. 그 후 구간마다 4분위 수로 나누어 75% 값을 대푯값으로 선정하였다.

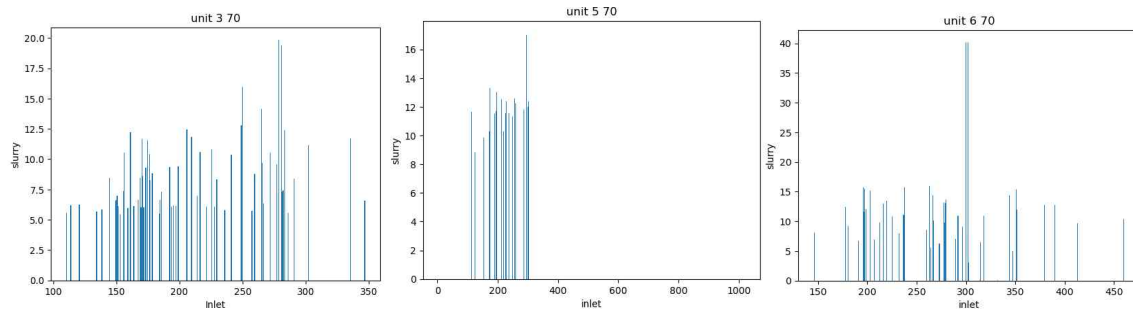
(그림8 : 선정 방식 설명)





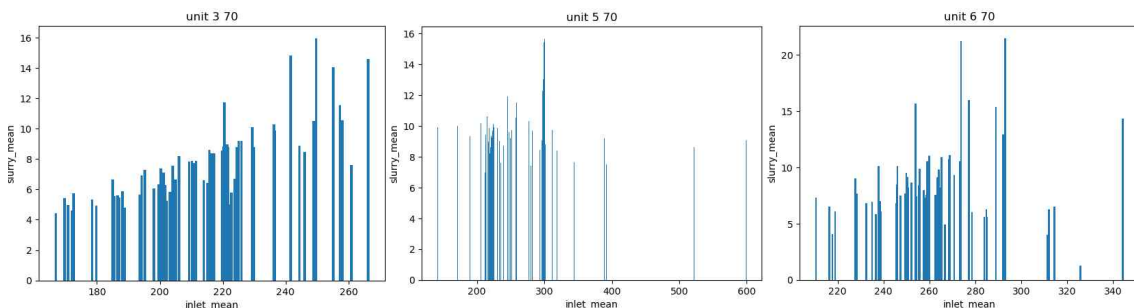
먼저 대포값을 75%로 선정한 경우 구간별 유입가스에 따른 탈황제 양의 상승이 비선형적이므로 대포값으로서 적절하지 않음을 알 수 있다. 구간 70개를 기준으로 보면 전체적으로 선형적인 모양을 취하는 3호기와 전체적으로 값이 비슷하게 나오는 5호기와는 다르게 6호기에서는 앞선 두 호기와 같은 특징을 찾기 어렵다.

(그림7 : 구간 70 3, 5, 6호기 quantile 0.75% 그래프)



위에 구한 대포값을 대신하여 각 구간의 평균을 대포값으로 선정하여 각 구간의 유입가스 별 그래프를 제작했다. 이렇게 제작한 그림8의 그래프를 보면 차이는 있으나 유입가스에 따른 탈황제 양의 상승이 비교적 선형적임을 알 수 있다.

(그림8 : 구간 70 3, 5, 6호기 평균 대포값 그래프)



대포값을 선정하기 위하여 70 구간별 유입 SOx와 탈황제 투입량 평균값과 Quantile 75%를 회귀 분석하여  $R^2$  값을 구하여 설명력이 높은 평균값을 대표값으로 선정 하였다.

(표1 : 구간 70 회귀분석  $R^2$  값 비교 )

구간70	quantile 0.75	평균
	$R^2$	$R^2$
3호기	0.16727	0.67558
5호기	0.00001	0.00129
6호기	0.00085	0.03641

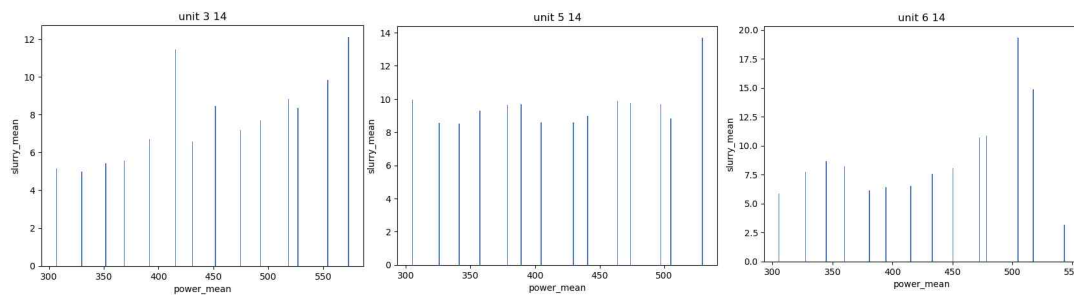
평균 대포값을 기준으로 하여 각호기의 구간의 개수를 선정하여 다시 그래프를 그렸으나 각 호기의 발전량 300Mw로 제한하면서 데이터 수는 각각 3호기 41398개, 5호기 23341개, 6호기 35202개로 동일한 기준의 단계를 적용하게 되면 상이한 구간별 데이터 건수로 인하여 왜곡이 발생 하여 데이터 건수가 1을 초과하는 구간값이 존재하는 구간은 14개 구간으로 3호기는 구간값 19Mw, 5호기는 구간값 16Mw, 6호기는 17Mw구간이 적합하다고 판단했다.



(표2 : 호기별 구간별 데이터 건수)

구간	3호기 (19Mw)	5호기 (16Mw)	6호기 (17Mw)
1	834	176	312
2	282	149	225
3	1,051	309	858
4	276	195	253
5	138	460	798
6	488	174	180
7	399	220	106
8	475	1,753	3,035
9	1,912	936	911
10	1,378	1,937	22,539
11	21,915	14,870	5,679
12	11,589	168	287
13	648	1,977	2
14	13	17	17

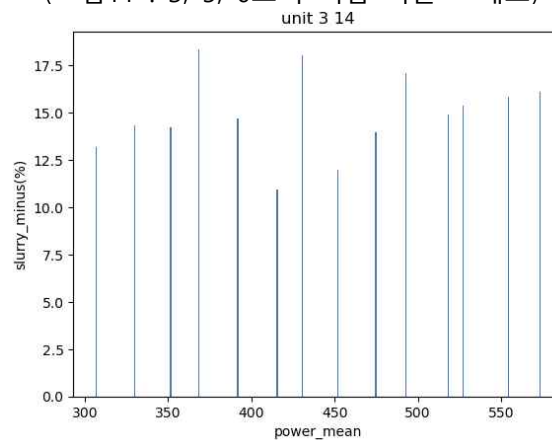
(그림10 : 구간 선정 후 3, 5, 6호기 평균 대포값 그래프)

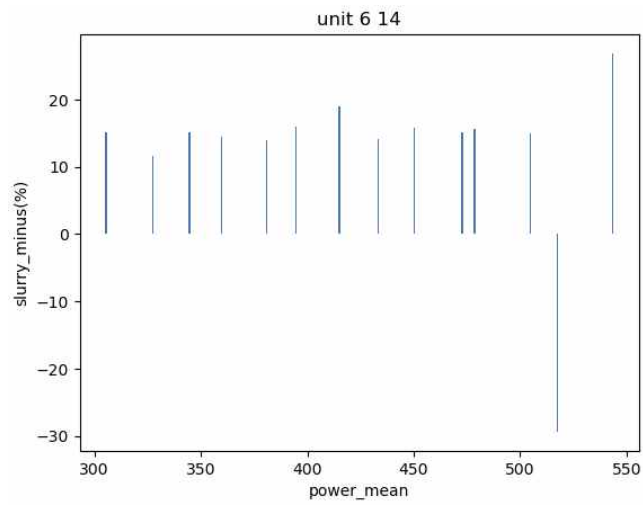
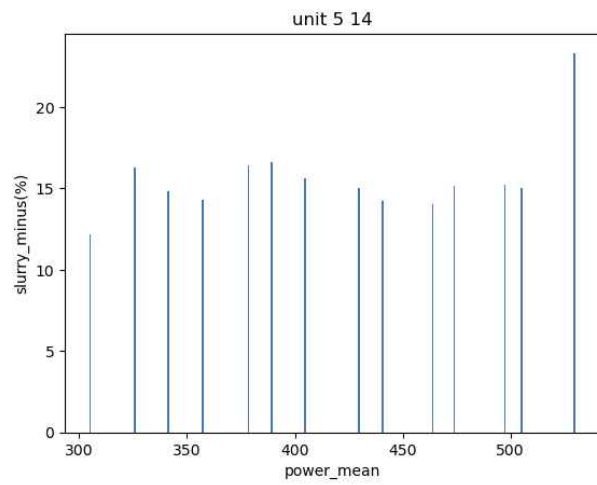


[그림10] 처럼 각호기에 맞게 단계 구간의 개수를 구하고 각 구간의 대포값을 설정하였다.

이후 각 호기의 구간 별 실제 투입량값 대비 선정한 대포값의 차이를 비교하여 얼마나 탈황제의 저감을 할 수 있을지 알아보았다.

(그림11 : 3, 5, 6호기 저감 비율 그래프)





※  $slurry\_minus(\%) = \frac{(\text{각구간의실제투입값합계}) - (\text{대푯값의85\%}) * (\text{구간의데이터개수})}{(\text{각구간의실제투입값합계})} \times 100$

#### 4. 결론

최종적으로 발전량 300Mw이상을 기준으로 한 데이터를 발전량 기준으로 각 호기의 맞게 구간을 나누어 평균을 그 구간에 대푯값으로 선정하였다.

선정한 대푯값을 기준으로 85%만을 탈황제로 투입하였을 경우 실제로 넣은 탈황제의 양을 평균적으로 11.94% 저감할 수 있다. 또한, 유입가스에 따른 차이값을 그래프로 그려보면 3호기는 200ppm 이상 240ppm 이하, 5호기와 6호기는 240ppm 이상 300ppm 이하에서 많은 양의 데이터를 가지며 그 차이값이 크므로 탈황제의 양을 줄일 시 큰 효과를 볼 수 있다.

그림9는 위 저감 퍼센트 식을 통해 실제 투입값과 대푯값의 85%를 투입하였을때의 저감된 탈황제의 비율을 확인했다.

(그림10 : 3, 5, 6호기 평균 저감 비율)

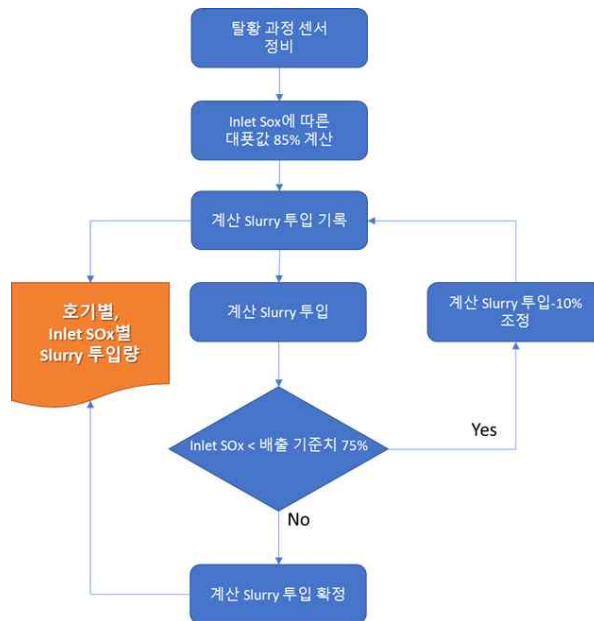
3호기 평균 저감 비율: 12.892118175285832  
5호기 평균 저감 비율: 14.534196362832619  
6호기 평균 저감 비율: 8.398908886682385

그림10에서는 현재 3호기와 6호기에는 실제 투입값과 대표 투입값의 85%의 차이가 매우 큰 값이 존재하며 그 값이 6호기에서 특히 그 값이 커 그 값을 -10000% 이상으로 제한하여 저감 비율을 구했다. 이를 통해 평균을 이용한 대푯값을 이용하여 그 값을 이용하면 최저 8.4%에서 최대 14.53%를 저감할 수 있다.

## ⑤ 사업화방안 및 기대효과

탈황 제거 공정을 분석이 처음 시도되는 데이터로 판단되며 Inlet SOx값에 영향을 미치는 발전용량과 투입되는 탄의 종류등 다양한 요인으로 SOx의 농도가 결정 된다. 본 분석과제에서는 유입되는 SOx값에 따른 최적의 탈황제를 산출하는 것으로 변화를 줄 수 있는 변수는 탈황제 투입량이다.

현장에서 본 결과물을 적용하려면 아래오 같은 절차를 수행해야 할것으로 판단 됩니다.



[ 본 분석과제 결과물 적용 절차도 ]

단계별 대포값의 85%를 적용하여 탈황제를 투입한다면 아래 표와 같은 투입량 저감과 경제적인 성과를 기대 합니다.

[ 3호기 Slurry 저감 기대량 ]

구간	실제투입량 (A)	대포값의85% (B)	저감량 (A-B)	slurry_minus (%)
1	4,198.3	3,645.3	553.0	13.2%
2	1,393.9	1,194.1	199.8	14.3%
3	5,637.9	4,834.8	803.1	14.2%
4	1,601.8	1,307.7	294.2	18.4%
5	920.9	785.5	135.4	14.7%
6	5,332.2	4,749.0	583.1	10.9%
7	2,710.9	2,221.8	489.1	18.0%
8	3,881.3	3,417.2	464.2	12.0%
9	13,555.6	11,663.4	1,892.1	14.0%

10	10,871.3	9,011.1	1,860.2	17.1%
11	193,278.2	164,464.6	28,813.6	14.9%
12	97,299.5	82,338.2	14,961.3	15.4%
13	6,439.6	5,418.2	1,021.4	15.9%
14	159.5	133.8	25.7	16.1%
<b>합계</b>	<b>347,280.9</b>	<b>295,184.5</b>	<b>52,096.3</b>	<b>14.9%</b>

[ 5호기 Slurry 저감 기대량 ]

구간	실제투입량 (A)	대포깎의85% (B)	저감량 (A-B)	slurry_minus (%)
1	1,692.7	1,486.2	206.5	12.2%
2	1,294.7	1,083.9	210.8	16.3%
3	2,624.7	2,235.2	389.5	14.8%
4	1,797.9	1,540.5	257.4	14.3%
5	4,508.6	3,768.2	740.4	16.4%
6	1,718.3	1,432.4	285.8	16.6%
7	1,907.7	1,609.8	297.9	15.6%
8	15,064.3	12,799.6	2,264.7	15.0%
9	8,340.7	7,153.0	1,187.7	14.2%
10	18,873.4	16,228.5	2,644.9	14.0%
11	145,140.9	123,155.0	21,985.9	15.1%
12	1,633.2	1,384.9	248.3	15.2%
13	17,417.8	14,799.1	2,618.6	15.0%
14	258.2	198.0	60.3	23.3%
<b>합계</b>	<b>222,273.1</b>	<b>188,874.4</b>	<b>33,398.7</b>	<b>15.6%</b>

[ 6호기 Slurry 저감 기대량 ]

구간	실제투입량 (A)	대포깎의85% (B)	저감량 (A-B)	slurry_minus (%)
1	1,836.6	1,559.4	277.2	15.1%
2	1,676.5	1,482.6	193.9	11.6%
3	7,457.3	6,334.9	1,122.4	15.1%
4	2,065.2	1,766.6	298.6	14.5%
5	4,831.2	4,158.3	672.9	13.9%
6	1,169.6	982.2	187.4	16.0%

7	725.2	588.0	137.2	18.9%
8	22,690.0	19,471.5	3,218.5	14.2%
9	7,403.3	6,233.6	1,169.8	15.8%
10	241,022.9	204,761.0	36,261.9	15.0%
11	62,016.5	52,318.4	9,698.1	15.6%
12	5,549.5	4,716.5	833.0	15.0%
13	19.5	25.3	-5.7	-29.4%
14	62.8	46.0	16.9	26.8%
합계	358,526.2	304,444.3	54,081.9	12.7%

[ 3~6호기 합계 저감 기대량 ]

	3호기(ton)	5호기(ton)	6호기(ton)	합계
저감량	52,096	33,399	54,082	139,577
저감비율	15%	16%	13%	14%