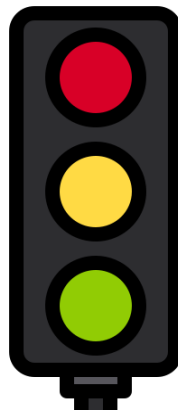


신재생e 설비(접속반) 건전성 평가 모델 분석

2022-08



이원(명지대)
임혁진(강원대)
이유송(강원대)

목 차

I . 분석 필요성 및 목적.....	1
1. 분석 목적	1
II . 분석 내용.....	1
1. 분석 범위 및 대상.....	1
2. 분석 목표	1
III . 분석 결과.....	2
1. 데이터 가공절차.....	2
2. 데이터 처리.....	3
3. 정규화(Normalization).....	4
4. 데이터 탐색.....	5
5. 분석모델.....	7
6. 분석모델 평가.....	8
7. 결론.....	9

I 분석 필요성 및 목적

1. 분석 목적

○ 접속반 설비에 대한 건정성 지표 개발

- 태양광 발전 설비는 최근 3년간 전기적 요인에 의해 발생한 화재에서 발화지점은 분전반 및 접속함이 147건에 이르고 있고 화재 원인으로는 접속반의 높은 온도 상승 및 결로 현상을 통해 접속반 내부의 스파크 발생이 원인으로 추정되고 있다.

	분전반 및 접속반	배선 (접속단자)	인버터	모듈 및 어레이	기타	합계
2016년	36	12	14	11	0	73
2017년	44	26	10	6	0	86
2018년	67	35	28	13	2	145
합계	147	73	52	30	2	304
비율	48%	24%	17%	10%	1%	

표1 2016년~2018년까지 태양광설비 화재 현황표 조사[1]

- 태양광 발전소의 지속적인 증가는 화재사고의 증가로 이어지고 있으며 태양광 설비에 대한 화재의 관리에 필요한 지표 개발이 필요하다.

II 분석 내용

1. 분석 범위 및 대상

○ 시간적 범위 : 2021년 1월~2022년 3월

○ 분석 Tool

- 환경 : MS-Azure
- Tool : Python(numpy, pandas, sklearn, matplotlib)

○ 분석 데이터

Data Set	설명	건수	수집주기	기간
th_connectbox_day	시간당 접속반 Data	92,044	1분	21/01/01/05:00~ 22/03/30/09:00
th_connectbox	분당 접속반 Data	4,227,340	1시간	21/01/01/05:00~ 22/03/30/09:55

Data Set	설명	건수	수집주기	기간
th_digitalmeter	분당 차단기 Data	1,300,970	1분	21/01/01/05:00~ 22/03/30/09:55
th_digitalmeter_day	시간당 차단기 Data	28,766	1시간	21/01/01/05:00~ 22/03/30/09:00
th_event	이벤트 Data	5,463	발생시	21/01/01/05:00~ 22/03/30/09:55
th_inverter	분당 인버터의 Data	878,236	1분	21/01/01/05:00~ 22/03/30/09:55
th_inverter_day	시간당 인버터의 Data	15,133	1시간	21/01/01/05:00~ 22/03/30/09:00
th_temp	분당 변압기의 Data	845,468	1분	21/01/01/05:00~ 22/03/30/09:55
th_temp_day	시간당 변압기의 Data	15,133	1시간	21/01/01/05:00~ 22/03/30/09:00
th_weather	분당 기상 Data	422,734	1분	21/01/01/05:00~ 22/03/30/09:55
th_weather_day	시간당 기상 Data	15,133	1시간	21/01/01/05:00~ 22/03/30/09:00

2. 분석 목표

○ 접속반 내의 단자대 온도를 활용한 건전성 지표 개발

- 분당 접속반 데이터(th_connectbox) 건수가 400만건이 넘어 분석에 필요한 대표 접속반 선정을 위한 기술통계
- 단자대 온도에 영향을 미치는 요소 상관 분석
- 현재/5/10분전 데이터로 10분후 단자대 온도 예측 모델 구현
- 10분후 예측 단자대 온도에 대하여 단측 검정값으로 건전성 판단 지표 개발
- 98.5% 노랑(50.9)
- 99.995% 주황(54.2)
- 설비기준 빨강(70)

III 분석 결과

1. 데이터 가공

□ 접속반 데이터(th_connectbox) 기술통계

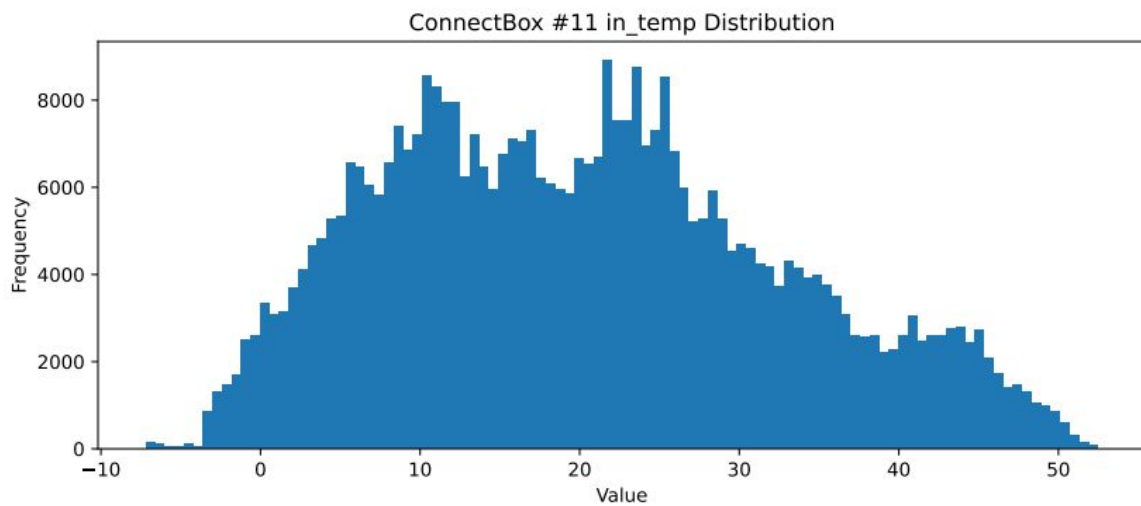
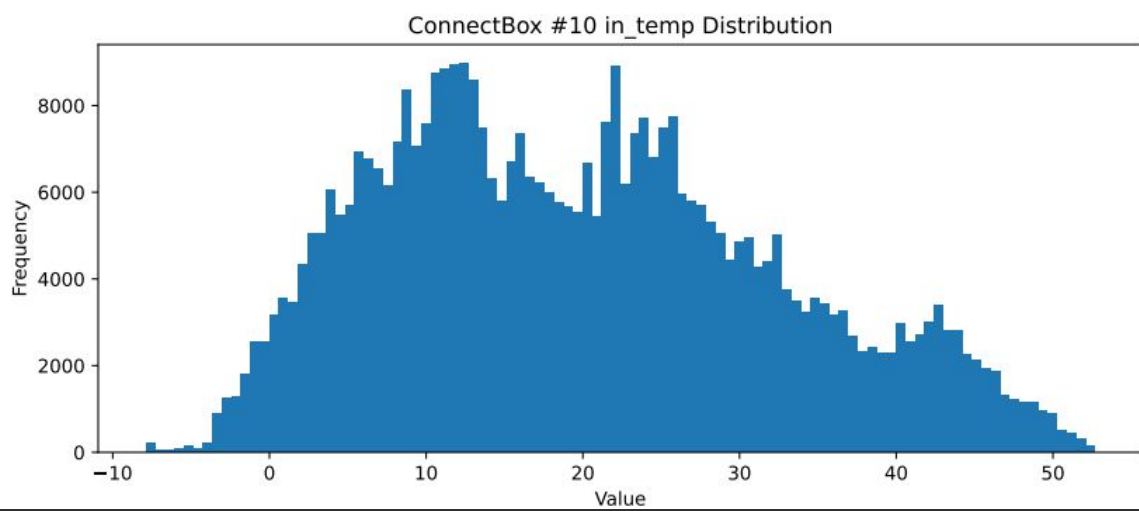
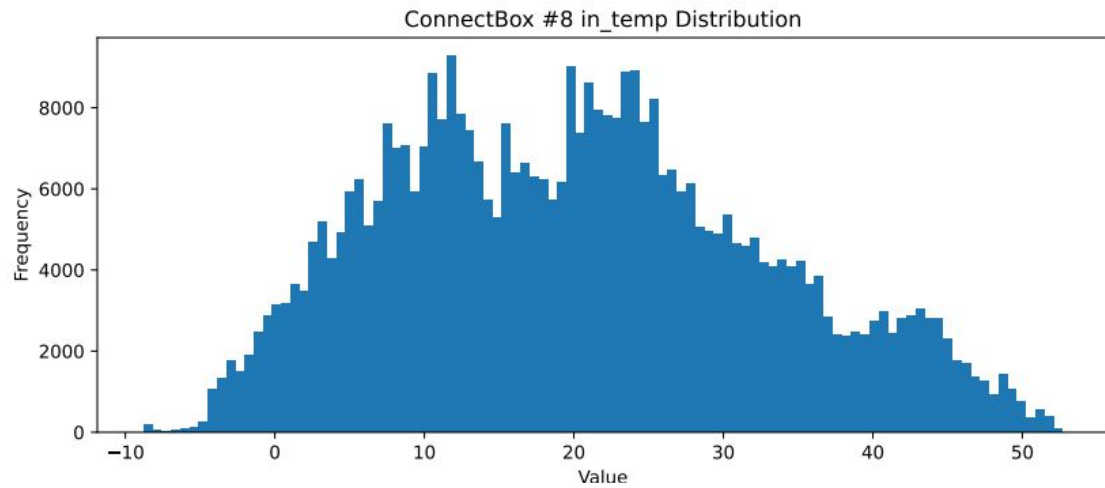
○ 각 접속반별 기술통계 값

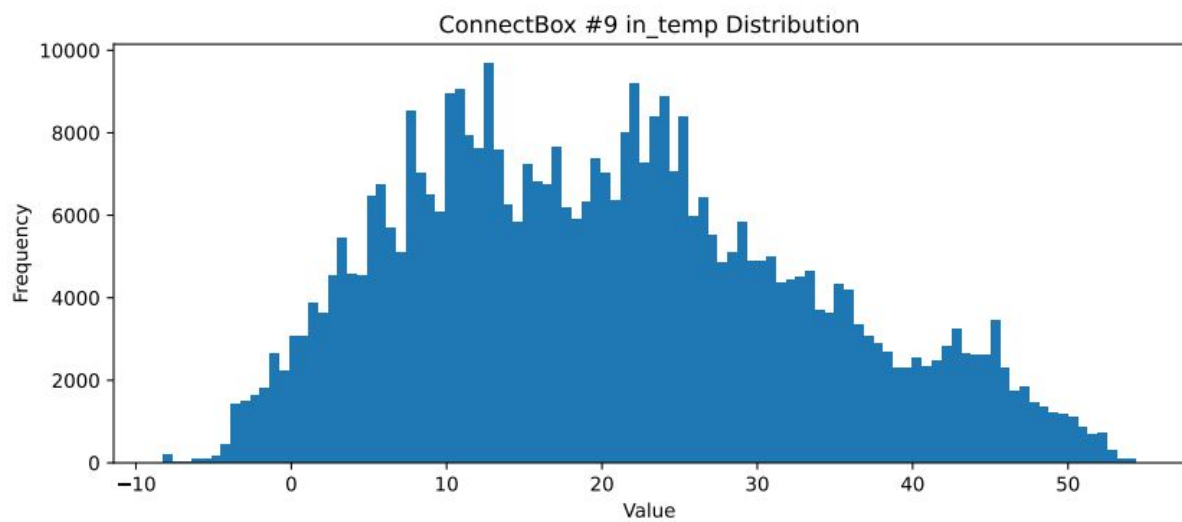
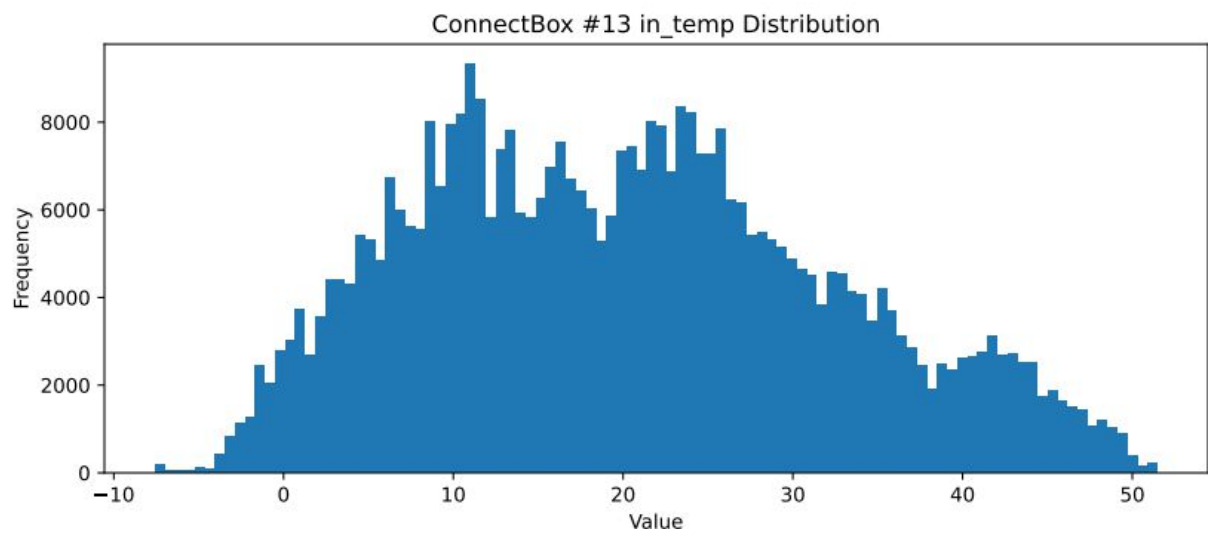
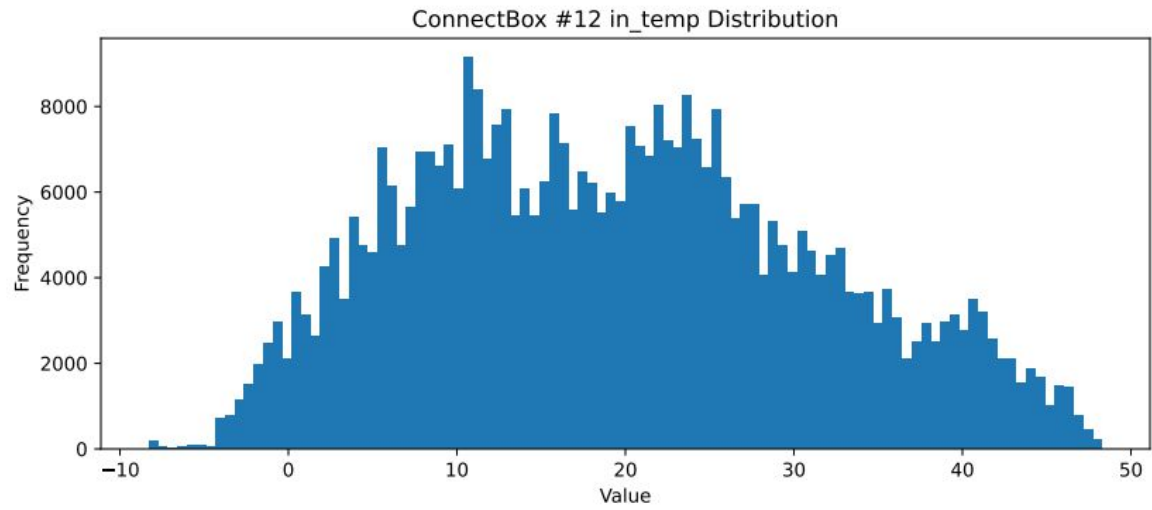
	접속반8	접속반9	접속반10	접속반11	접속반12	접속반13	접속반14	접속반15
건수	425,511	422,734	422,734	422,734	422,734	422,734	422,734	422,734
평균	20.2	20.5	19.9	20.3	19.5	20.2	20.1	20.3
표준편차	12.5	12.8	12.6	12.4	11.9	12.2	12.1	12.4
최소값	-	-	-	-	-	-	-	-
최소값	8.8	8.3	7.9	7.2	8.3	7.6	7.9	7.8
25%	10.5	10.5	10.0	10.5	10.2	10.6	10.6	10.6
50%	19.9	19.7	18.7	19.6	19.0	19.7	19.6	19.9
75%	28.7	29.2	28.4	28.7	27.8	28.6	28.4	28.8
최대값	52.7	54.4	52.7	52.5	48.3	51.5	51.1	52.1

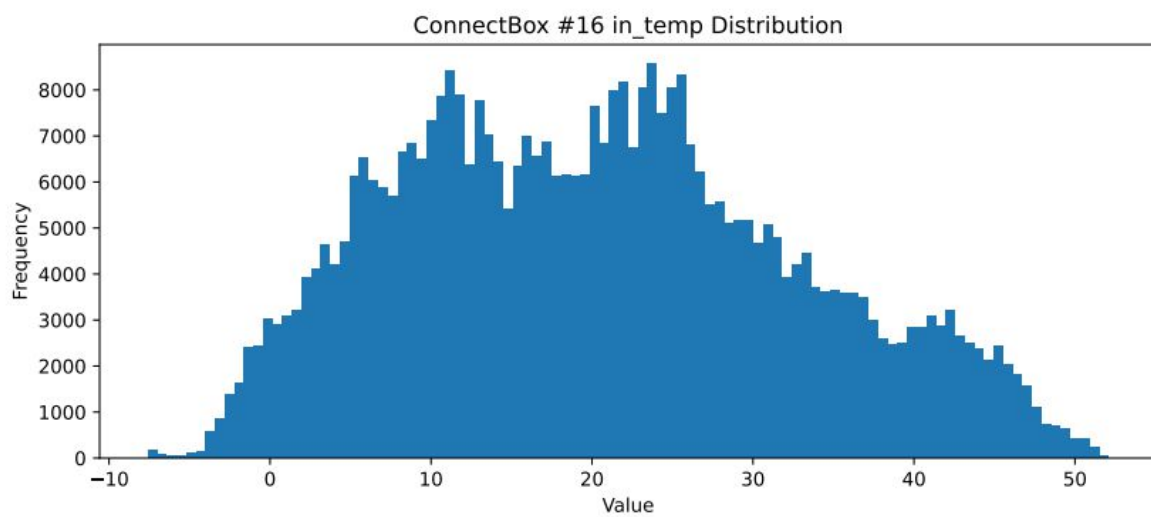
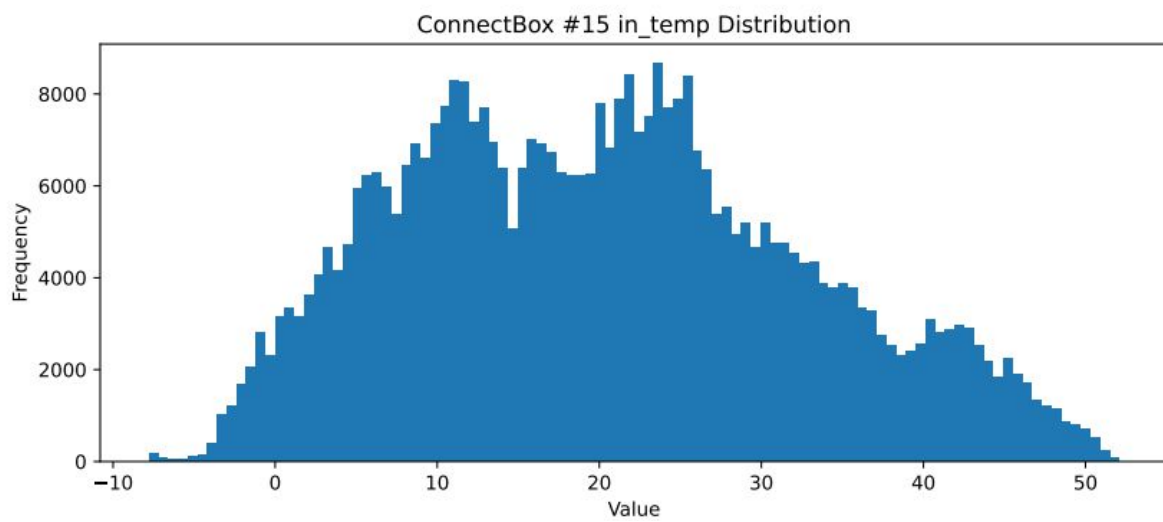
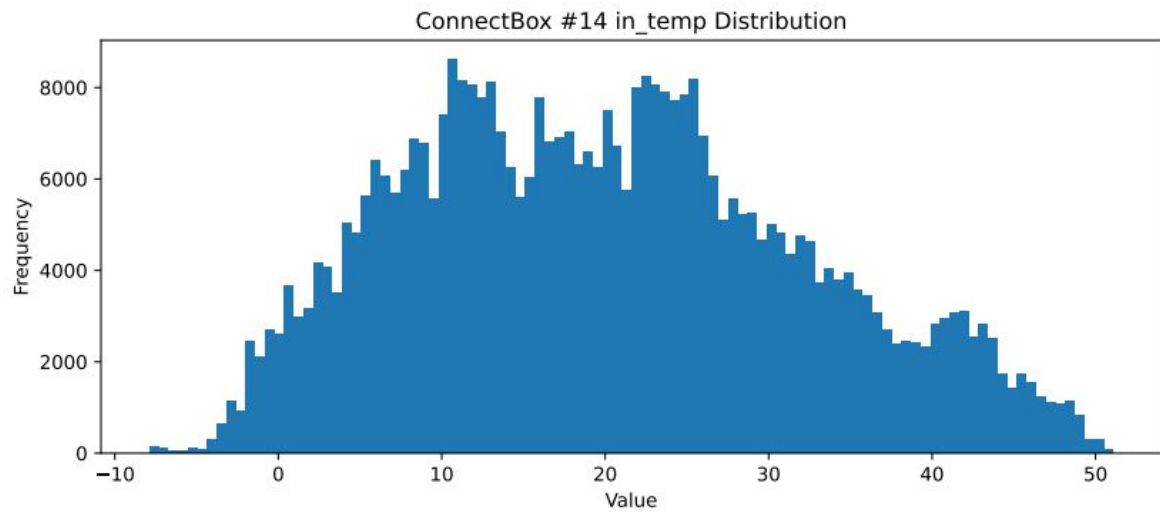
	in_temp	v_total	a_total	s_radiation	h_radiation	mod_temp	out_temp
in_temp	1.00						
v_total	0.31	1.00					
a_total	0.64	0.36	1.00				
s_radiation	0.62	0.39	0.97	1.00			
h_radiation	0.57	0.41	0.96	0.99	1.00		
mod_temp	0.93	0.22	0.47	0.43	0.38	1.00	
out_temp	0.94	0.36	0.77	0.75	0.73	0.90	1.00

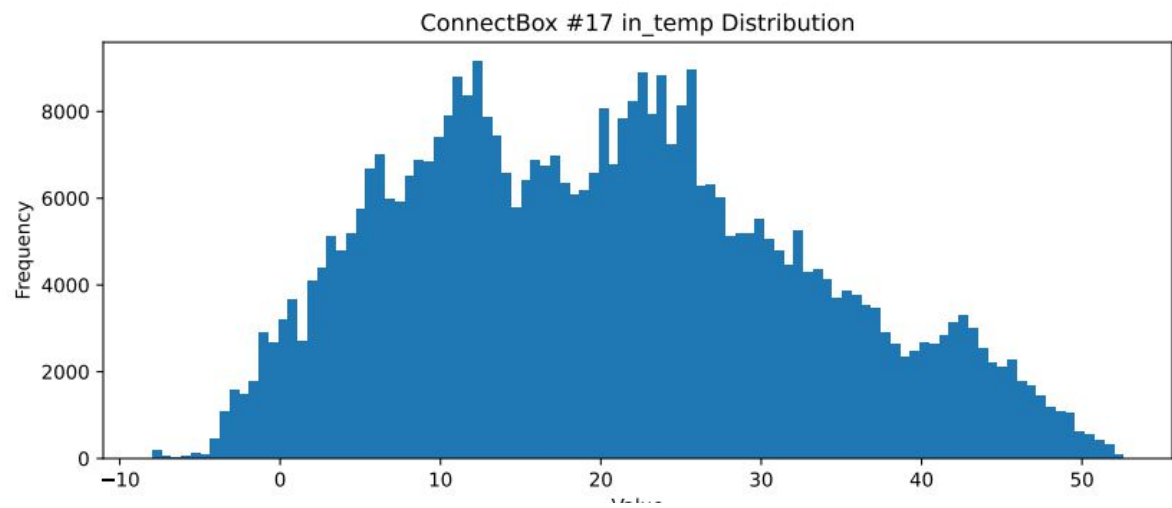
- 기상자료개방포털¹⁾ 및 동업날씨²⁾

1) 기상자료개방포털(<http://data.kma.go.kr>.)









□ 신재생e 발전량 데이터

○ 신보령 태양광 에너지 발전량(단위: Kwh)

- 2020년 10월부터 2020년 10월까지의 분석

day	T	SITE_COD	UNIT_ID	발전본부	발전소	발전량 합산(kw/h)	예측발전량 합산(kw/h)	오차율
2020-10-01	0	7	1	신보령발전본부	태양광 1호기	0	0.96	
2020-10-01	1	7	1	신보령발전본부	태양광 1호기	0	0.96	
2020-10-01	2	7	1	신보령발전본부	태양광 1호기	0	0.96	
2020-10-01	3	7	1	신보령발전본부	태양광 1호기	0	0.96	
2020-10-01	4	7	1	신보령발전본부	태양광 1호기	0	0.96	
2020-10-01	5	7	1	신보령발전본부	태양광 1호기	0	0.96	
2020-10-01	6	7	1	신보령발전본부	태양광 1호기	1.96	0.96	51.04
2020-10-01	7	7	1	신보령발전본부	태양광 1호기	90.81	0.96	98.94
2020-10-01	8	7	1	신보령발전본부	태양광 1호기	393.3	271.27	31.03
2020-10-01	9	7	1	신보령발전본부	태양광 1호기	451.41	226.02	49.93
2020-10-01	10	7	1	신보령발전본부	태양광 1호기	1,053.46	393.15	62.68
2020-10-01	11	7	1	신보령발전본부	태양광 1호기	1,295.16	758.69	41.42
2020-10-01	12	7	1	신보령발전본부	태양광 1호기	1,454.09	1,024.12	29.57
2020-10-01	13	7	1	신보령발전본부	태양광 1호기	1,542.72	1,061.21	31.21
2020-10-01	14	7	1	신보령발전본부	태양광 1호기	1,043.70	1,021.58	2.12
2020-10-01	15	7	1	신보령발전본부	태양광 1호기	767.33	777.58	1.34
2020-10-01	16	7	1	신보령발전본부	태양광 1호기	320.56	389.4	21.48
2020-10-01	17	7	1	신보령발전본부	태양광 1호기	91.55	194.52	112.48

그림 14 태양광 발전량 데이터 예시

- 일자, 시간별로 예측발전량과 발전량, 그리고 오차율이 표시되어 있음.

□ 분석 분석 절차

○ 발전량 데이터

- 시계열 제공 데이터: 시 단위로 나누어져 있으며, 11월 · 12월의 데이터가 다수 누락
- 추가 제공받은 데이터로 누락된 부분 추가

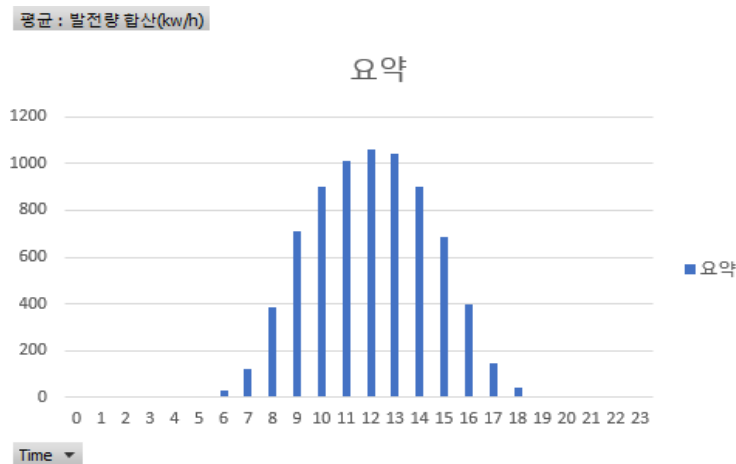
○ 기후 데이터

- 전운량이 문자로 표기되어 계산시에 오류가 발생하므로 데이터 형식을 숫자로 변경함.

2. 데이터 처리

□ 발전량 데이터 가공 및 기후 데이터와의 병합

○ 시계열 제공 데이터



- 추가데이터인 2020년 11월, 12월 데이터의 20시부터 23시, 0시부터 4시까지의 발전량은 0kwh이고, 오전 5-6시, 오후 20시 또한 발전량 30kwh 이하이므로 오전 7시부터 오후 7시부터의 값만이 유의미하다고 판단하여, 이외의 시간은 제외

- 2020년 11월 · 12월 데이터 누락본을 보정하기 위해 분 단위 데이터 추가

- Hour 함수 이용하여 분 단위의 2020년 11월, 12월 데이터를 시간별로 재분류하고, 피벗 테이블을 통해 인버터(CID)별/시간별 발전량 도출

- 텍스트 나누기 기능으로 기후 데이터의 날짜와 시간을 분리하여 기존 데이터와 동일한 형식으로 바꾼 후, Lookup 함수를 통해 날짜, 시간별로 기후 데이터와 발전량 데이터 매치하고 발전량 데이터가 기록되지 않는 날짜에는 생산량 평균 삽입

#. 사용한 함수

$$IFERROR(LOOKUP(1, 1/((\#1[[day]] = \$A2) * (\#1[Time] = \$B2)), \#1[발전량합산(kw/h)]), VLOOKUP([time], \$U\$1:\$V\$24, 2, 1))$$

- 용이한 분석처리를 위해 피벗테이블을 이용하여 주단위로 변경

3. 정규화(Normalization)

□ 데이터 정규화

○ 정규화 처리의 장점

- 기후데이터와 발전량데이터 간에 단위차이가 크기 때문에, 변화를 명확히 판별하기 어려움.

- 정규화를 통해 모든 데이터를 0과 1 사이의 값으로 변환시켜, 기후데이터가 발전량에 미치는 영향에 대해 비교분석

○ 정규화 수식

$$- \frac{X - \min X}{\max X - \min X} = (\text{기후데이터값} - \text{최솟값}) / (\text{최댓값} - \text{최솟값})$$

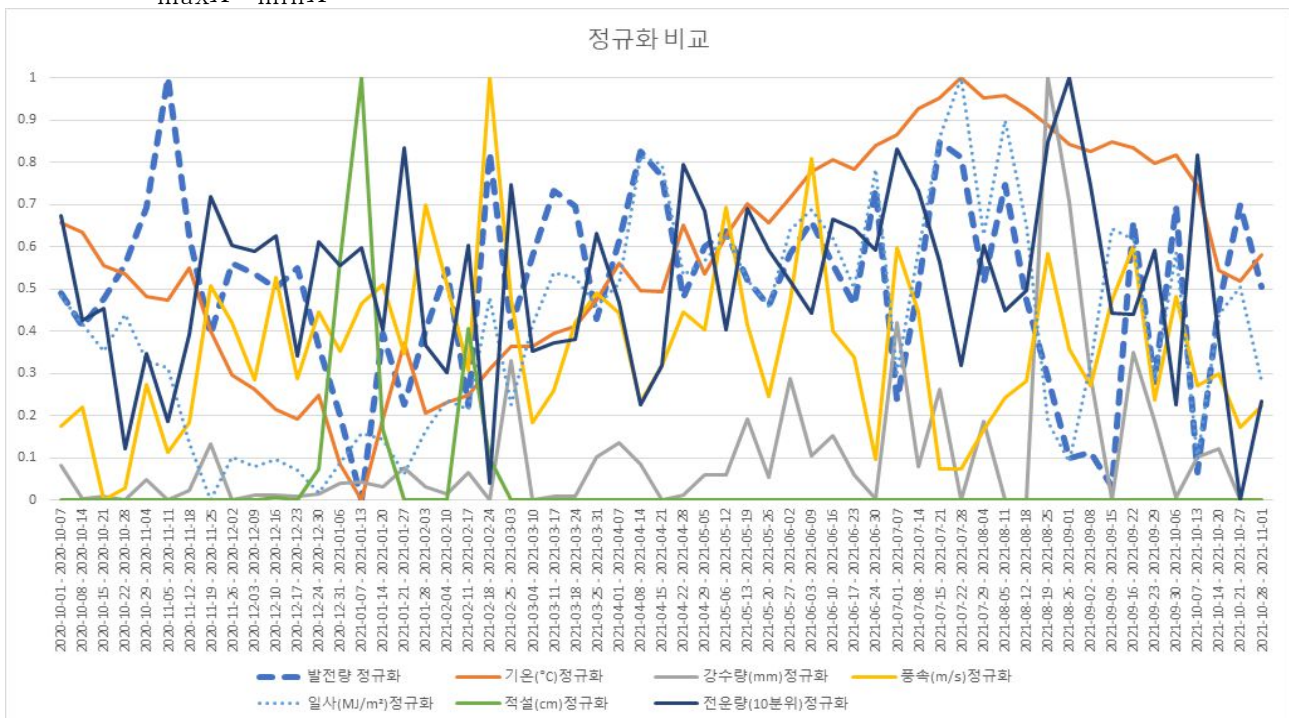


그림 19 기후데이터와 발전량 정규화 수치 그래프

- 정규화 그래프를 비교해보면 발전량의 그래프와 일사의 그래프가 특정 부분을 제외하고 비례하여 변화한다는 것을 확인할 수 있음.

- 발전량 그래프와 전운량 그래프는 상반되는 추이를 보이고 있음.

4. 데이터 탐색

□ 상관계수

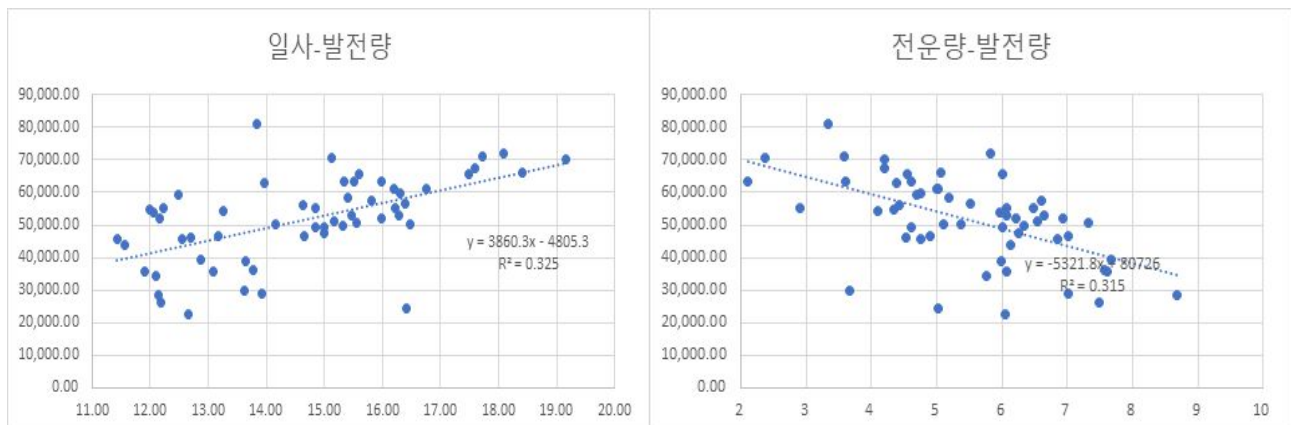
	합계 : 발전량(kw/h)-결측치 보완
합계 : 발전량(kw/h)-결측치 보완	1
평균 : 기온(°C)	0.100781723
합계 : 강수량(mm)	-0.291572919
평균 : 풍속(m/s)	-0.110892771
평균 : 일사(MJ/m²)	0.570064029
합계 : 적설(cm)	-0.385248722
평균 : 전운량(10분위)	-0.561207432

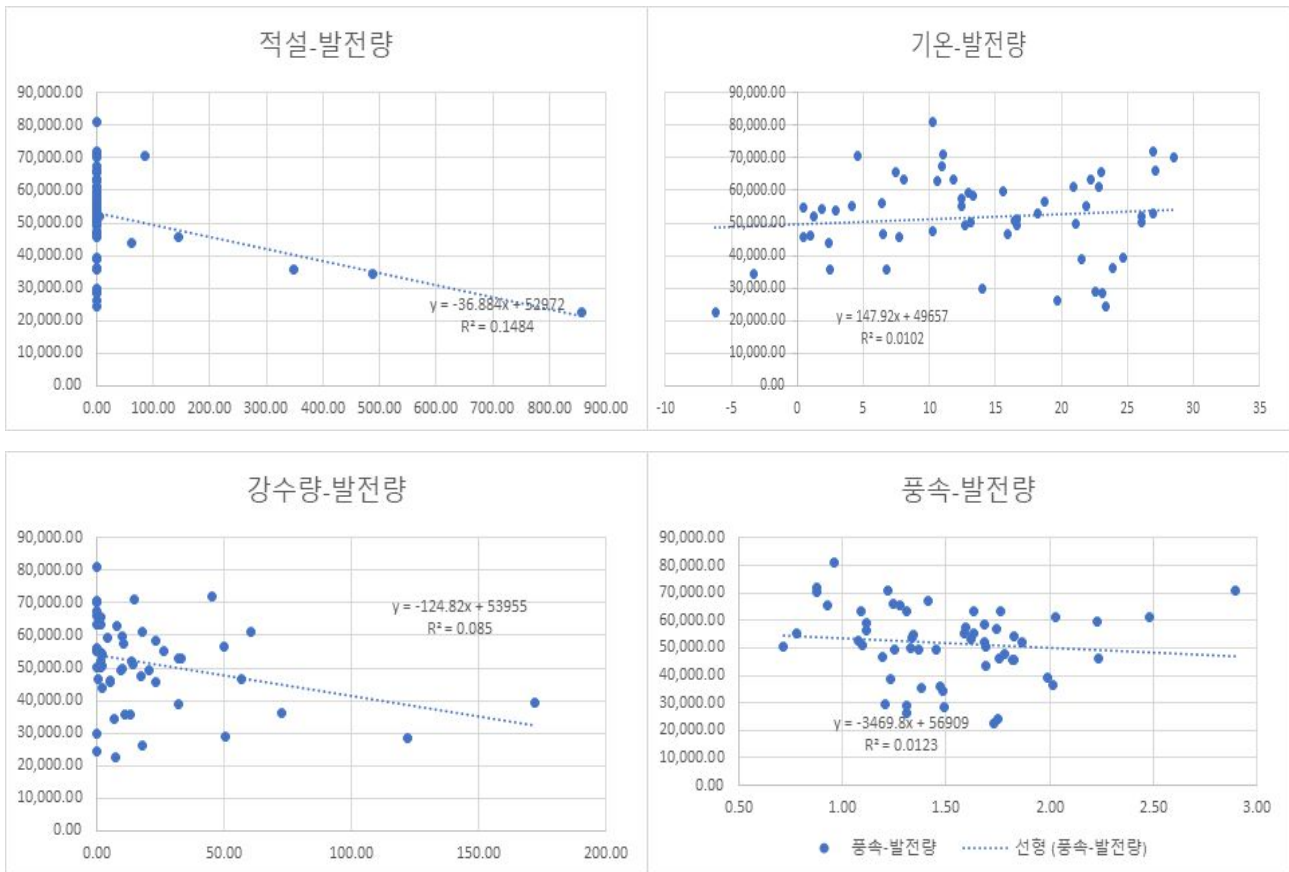
그림 21 발전량과 각 기후데이터의 상관계수

- 일사량과 발전량의 상관계수는 0.5 이상으로 강한 양의 상관관계를 보이고 있으며, 전운량은 -0.5 이하로 강한 음의 상관관계를 보이고 있음. 강수량과 적설은 발전량과 약한 음의 상관관계를 보임.

□ 단순회귀분석

○ 단순회귀식





○ 단순회귀분석 결과

	결정계수	p값
일사	0.3249로 유의수준(60%)에 미치지 못함.	3.6775E-06으로 유의미한 분석
전운량	0.315로 유의수준(60%)에 미치지 못함.	5.59E-06으로 유의미한 분석
적설	0.1484로 유의수준(60%)에 미치지 못함.	0.003084로 유의미한 분석
기온	0.01015로 유의수준(60%)에 미치지 못함. 매우 낮은 결정계수를 보임.	0.45571으로 귀무가설 채택
강수량	0.0805로 유의수준(60%)에 미치지 못함. 매우 낮은 결정계수를 보임.	0.45571으로 귀무가설 채택
풍속	0.01229로 유의수준(60%)에 미치지 못함. 매우 낮은 결정계수를 보임.	0.0411526으로 귀무가설 채택

표 11 단순회귀분석 결과 결정계수와 P값

5. 분석모델

□ 다중회귀분석

- 기후 데이터의 표본 증가로 다중회귀분석 결과의 결정계수(R^2)값이 높게 측정됨.

회귀분석 통계량									
다중 상관계수	0.843419								
결정계수	0.711356								
조정된 결정계수	0.706904								
표준 오차	1910.45								
관측수	396								
분산 분석									
	자유도	제곱합	제곱 평균	F 비	유의한 F				
회귀	6	3.5E+09	5.83E+08	159.7801	1.1E-101				
잔차	389	1.42E+09	3649820						
계	395	4.92E+09							
	계수	표준 오차	t 통계량	P-값	하위 95%	상위 95%	하위 95.0%	상위 95.0%	
Y 절편	401.0453	911.7408	0.439868	0.660277	-1391.51	2193.602	-1391.51	2193.602	
평균 : 기온(°C)	-102.52	14.88583	-6.88708	2.3E-11	-131.787	-73.2532	-131.787	-73.2532	
평균 : 일사(MJ/m²)	723.9655	51.60924	14.02783	1.83E-36	622.4975	825.4334	622.4975	825.4334	
평균 : 전운량(10분위)	-373.713	53.20774	-7.02366	9.71E-12	-478.324	-269.102	-478.324	-269.102	
합계 : 적설(cm)	-36.9152	4.449399	-8.29666	1.78E-15	-45.663	-28.1673	-45.663	-28.1673	
합계 : 강수량(mm)	2.251248	11.93331	0.188652	0.850464	-21.2106	25.7131	-21.2106	25.7131	
평균 : 풍속(m/s)	43.14511	120.5634	0.357862	0.72064	-193.892	280.1826	-193.892	280.1826	

그림 30 다중회귀분석 결과

- 다중회귀분석 결과 강수량과 풍속은 P값이 유의수준에 미치지 못하므로 제외하여 재회귀분석을 진행

요약 출력									
회귀분석 통계량									
다중 상관계수	0.843339								
결정계수	0.711221								
조정된 결정계수	0.708267								
표준 오차	1906.003								
관측수	396								
분산 분석									
	자유도	제곱합	제곱 평균	F 비	유의한 F				
회귀	4	3.5E+09	8.75E+08	240.7442	4.9E-104				
잔차	391	1.42E+09	3632848						
계	395	4.92E+09							
	계수	표준 오차	t 통계량	P-값	하위 95%	상위 95%	하위 95.0%	상위 95.0%	
Y 절편	539.892	848.9634	0.635943	0.525186	-1129.21	2208.996	-1129.21	2208.996	
평균 : 기온(°C)	-101.585	14.21039	-7.14867	4.33E-12	-129.524	-73.647	-129.524	-73.647	
평균 : 일사(MJ/m²)	718.5183	48.82047	14.71756	2.51E-39	622.5349	814.5018	622.5349	814.5018	
평균 : 전운량(10분위)	-374.024	53.06775	-7.04804	8.25E-12	-478.357	-269.69	-478.357	-269.69	
합계 : 적설(cm)	-36.7627	4.421543	-8.31445	1.54E-15	-45.4557	-28.0697	-45.4557	-28.0697	

그림 31 강수량과 풍속을 제외한 다중회귀분석 결과

- 다중회귀식= $EO = -101.5853T + 718.5183I - 374.0235CA - 36.7627S + 539.8920$ ³⁾
- 기온, 일사, 전운량, 적설의 P값이 모두 0.05보다 작으므로 해당

3) EO(Energy Output): 발전량, I(Insolation): 일사, CA(Cloud Amount): 전운량, S(Snow): 적설량

회귀식 채택

6. 분석모델 평가

○ 예측 발전량

일시	평균 : 기온(°C)	합계 : 강수량(mm)	평균 : 풍속(m/s)	평균 : 일사량(MJ/m)	합계 : 적설(cm)	평균 : 전운량(10분위)	발전량(kw/h)	예상발전량(kw/h)	오차율
10월31일	14.26	0.00	0.86	13.79	0.00	4	5760.540	7334.572	4.04%
11월1일	13.39	0.00	0.59	13.12	0.00	3	7516.850	7343.827	0.44%
11월2일	12.07	0.00	0.76	14.08	0.00	5	8174.780	7668.909	1.3%
11월3일	11.54	0.00	0.77	14.50	0.00	4	7552.970	8399.539	2.17%
11월4일	13.13	7.50	1.06	14.36	0.00	4	8210.450	7869.629	0.87%
11월5일	13.58	0.00	1.27	14.11	0.00	4	7504.910	7961.609	1.17%
11월6일	14.85	0.00	1.57	14.41	0.00	2	7700.460	8651.415	2.44%
11월7일	16.83	0.00	1.69	14.27	0.00	4	7495.650	7600.947	0.27%
11월8일	11.72	37.60	1.98	9.86	0.00	9	2580.340	3178.793	1.53%
11월9일	9.31	2.30	2.16	10.69	0.00	8	3281.780	4359.799	2.76%
11월10일	7.25	17.00	1.13	10.12	0.00	8	2235.380	4006.491	4.54%
11월11일	7.95	1.40	1.38	10.16	0.00	7	2419.420	4286.726	4.79%
11월12일	6.17	8.40	0.90	10.50	0.00	9	2747.850	4247.100	3.84%
11월13일	8.88	0.00	1.24	13.73	0.00	4	7274.910	7957.615	1.75%
11월14일	12.45	0.00	1.58	13.41	0.00	3	7037.600	7900.623	2.21%
11월15일	10.58	0.00	0.96	13.54	0.00	4	7042.000	7726.023	1.75%
11월16일	9.09	0.00	0.99	14.22	0.00	1	7850.630	9307.833	3.74%
11월17일	8.39	0.00	0.87	14.22	0.00	2	7510.480	8986.899	3.79%
11월18일	10.79	0.00	2.69	12.73	0.00	5	4403.560	6561.824	5.53%
11월19일	13.05	0.00	1.72	11.46	0.00	7	4284.480	4783.083	1.28%
11월20일	10.19	0.00	0.83	12.37	0.00	6	4753.200	5993.978	3.18%
11월21일	10.64	0.00	0.74	10.41	0.00	9	2272.950	3606.406	3.42%
11월22일	6.39	6.00	1.91	11.39	0.00	9	4149.210	4883.026	1.88%
11월23일	2.79	3.20	1.40	10.83	0.40	8	2339.120	5111.101	7.11%
11월24일	7.53	2.60	1.60	13.74	0.00	4	8855.470	8278.457	1.48%
11월25일	10.66	0.00	1.20	13.09	0.00	2	7205.560	8285.947	2.77%
11월26일	5.58	0.00	0.87	13.93	0.00	1	7956.060	9715.537	4.51%
11월27일	4.62	0.00	1.09	14.07	0.00	0	7527.590	10167.446	6.77%

그림 33 발전량과 예측발전량 비교

- 전력시장운영규칙의 오차율 산식: $\frac{|발전량 - 예측량|}{설비용량} \times 100$ 4)

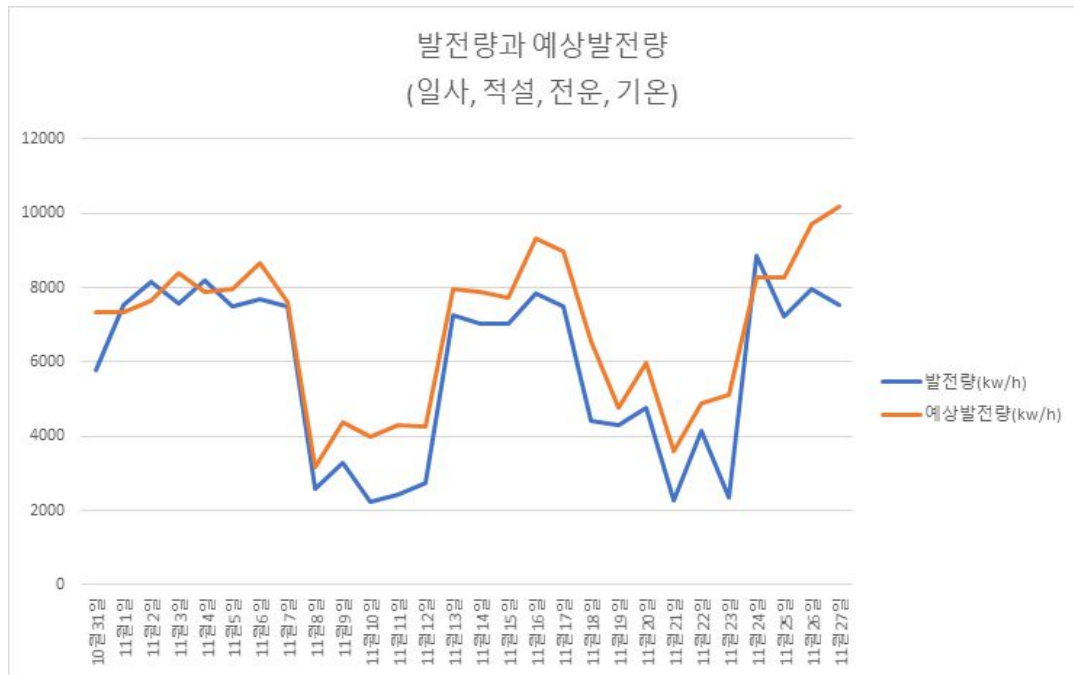
- 다중회귀분석을 통해 도출된 다중회귀식에 2021년 11월 기후데이터를 적용하여 예측 발전량을 계산하자 다음과 같은 결과 도출

○ 결론

- 회귀분석을 통한 예측 발전량과 실제 발전량의 그래프의 추이가 비슷한 것으로 볼 때, 기후가 태양광 발전의 효율을 좌우한다는 것을 알 수 있음.

- 일사량과 발전량의 상관계수 0.57, 결정계수 0.3249인 것으로 미루어 보아 기후요소 중 일사량이 발전량에 가장 큰 영향을 미친다는 것을

4) 신보령 태양광 1호기의 설비용량은 3,000kw/h, 일단위(7시부터 19시)로 계산하여 12시간을 곱한 36,000으로 계산



알 수 있음.

- 기온과 발전량이 음의 상관관계에 있으므로 기온이 지나치게 높으면 태양광 발전패널의 효율이 떨어져 발전량이 줄어듦. 따라서 여름이 아닌 봄과 가을에 태양광 발전 효율이 가장 좋음.

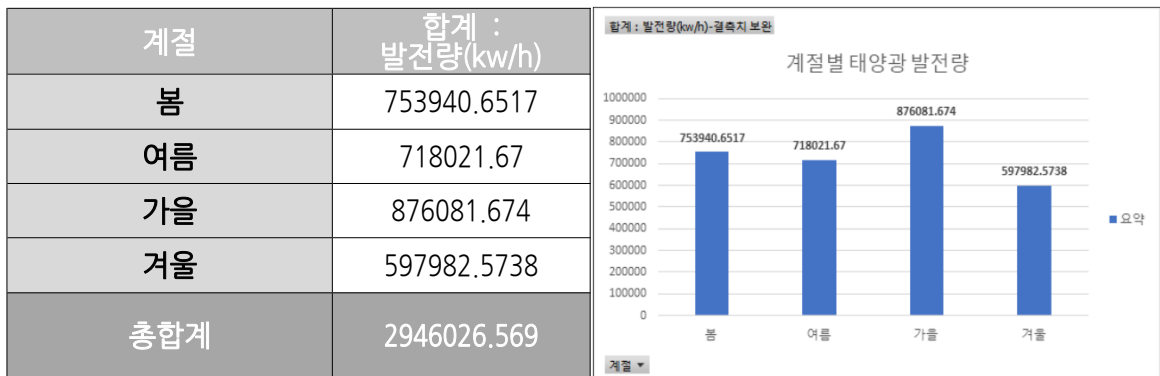


표 12 계절별 발전량

그림 34 계절별 발전량 가시화

○ 개선점

- 태양광 패널 고장 및 이상 등, 실제 운용에 대한 정확한 정보가 뒷받침 됐다면 더 정확한 예측치가 도출될 수 있었을 것

- 발전량 데이터 가공 시에 분별력 있는 기준으로 결측치, 이상치를 제외하지 못한 점이 오차율을 높였음.

- 참고 문헌

1. Kwang-Muk Park, Jae-Hyun Kim, Jin-Yeong Park, Seong-Jun Hong, Sun-Bae Bang, "Fire Statistics and Case Study of Photovoltaic System through Analysis of Fire Status Survey", Proceedings of the Korean Electrical Society, 2020.7, 2207-2208.