

전력 수요 예측을 위한 시계열모형 비교

충북대학교 정보통계학과

2019068041 이준구

2019068067 유윤종

1. 서론

본 연구의 목적은 한국전력거래소의 일별 최대 전력 수요 데이터를 활용하여 전력 수요 예측 모델을 개발하고, 주요 설명 변수인 온도, 주말 여부, 공휴일 여부를 포함한 다양한 모델의 예측 성능을 비교하는 것이다. 이를 통해 전력 수요에 영향을 미치는 요인들을 파악하고, 효과적인 예측 모델을 제시하고자 한다.

이러한 예측을 위해 한국전력공사에서 제공하는 2010년부터 2023년까지의 전력수요 데이터를 바탕으로 2가지의 자기회귀모형 그리고 XGBoost 모형을 이용해 2023.12.1. ~ 2023.12.31. 한 달 간의 전력수요 예측을 수행하였으며, 예측값과 실제값을 비교하여 모형의 성능을 평가하였다.

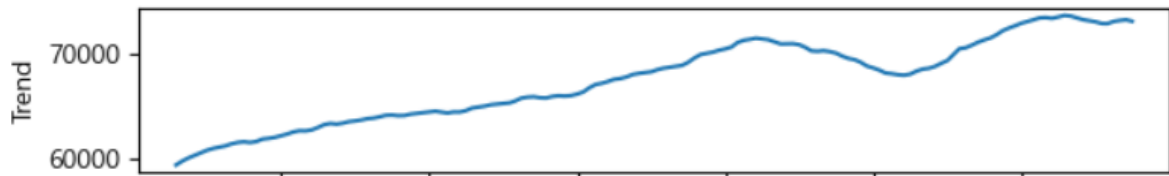
2. 데이터

2.1 변수 구성

변수명	설명	값의 의미
date	날짜	관측 날짜 (형식: YYYY-MM-DD)
watt	일별 최대 전력 수요 (단위: MW)	하루 동안 관측된 최대 전력 수요
temp	일별 평균 온도 (단위: °C)	해당 날짜의 평균 온도
temp_abs	온도 절대 편차 ($ temp - 15 $)	평균 온도와 기준 온도(15°C) 간의 절대 차이
day	요일 정보	1: 일요일, 2: 월요일, ..., 7: 토요일
weekend	주말 여부	0: 평일, 1: 주말
holiday	공휴일 여부	0: 비공휴일, 1: 공휴일
time	시간 지표	0부터 시작하는 순차적인 숫자 (시간순 데이터 인덱스)

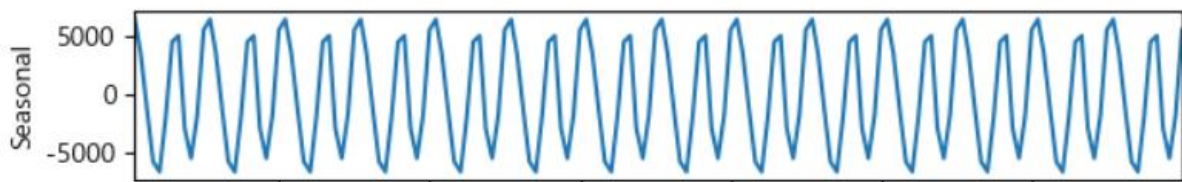
2.2 시계열 분해

2.2.1 추세



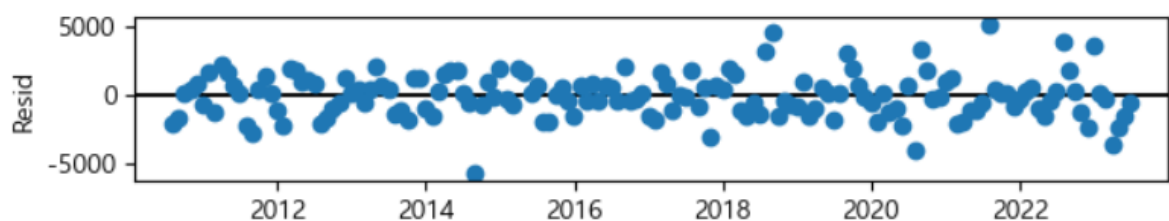
데이터의 장기적인 추세를 보여준다. 2010년부터 2020년대 초까지 꾸준히 증가하는 추세를 보이다가 약 2022년경부터 전력 수요의 증가세가 완화되거나 정체되는 경향이 나타난다.

2.2.2 계절성



데이터의 계절적 변동을 나타낸다. 주기가 12개월로 설정되어, 월별 전력 수요의 반복적인 패턴을 확인할 수 있다. 계절적 패턴은 여름과 겨울에 전력 수요가 증가하고, 봄과 가을에 감소하는 주기를 보여준다. 이는 계절성이 뚜렷하며, 연중 반복되는 규칙적인 변동을 볼 수 있다.

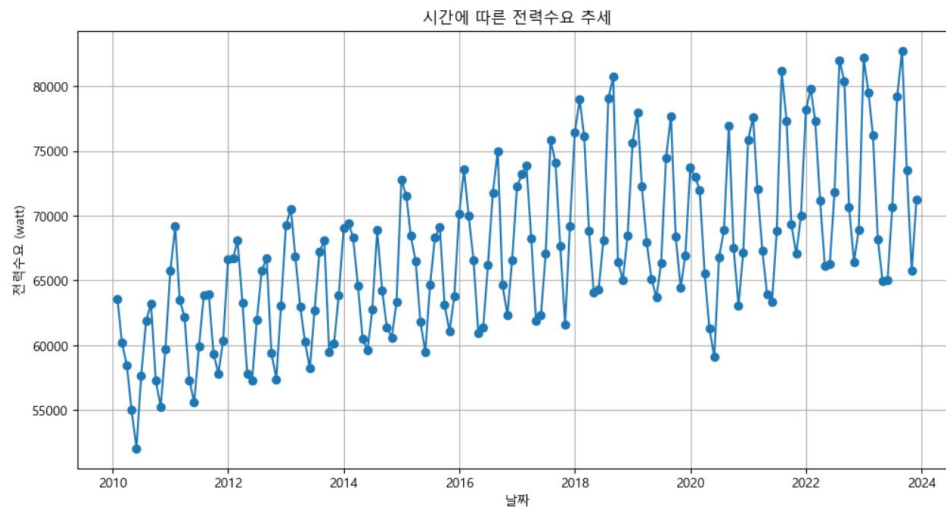
2.2.3 잔차



원 시계열 데이터에서 추세와 계절성을 제거한 후 남은 불규칙한 변동(잡음)을 보여준다. 전반적으로 평균이 0에 가까운 분포를 보이는데, 특정 시점 2014년 또는 2020년에서 잔차가 크게 변동하는 구간이 있다. 이는 예상치 못한 이벤트(정전, 극단적인 날씨 등)와 관련되는 것으로 보인다.

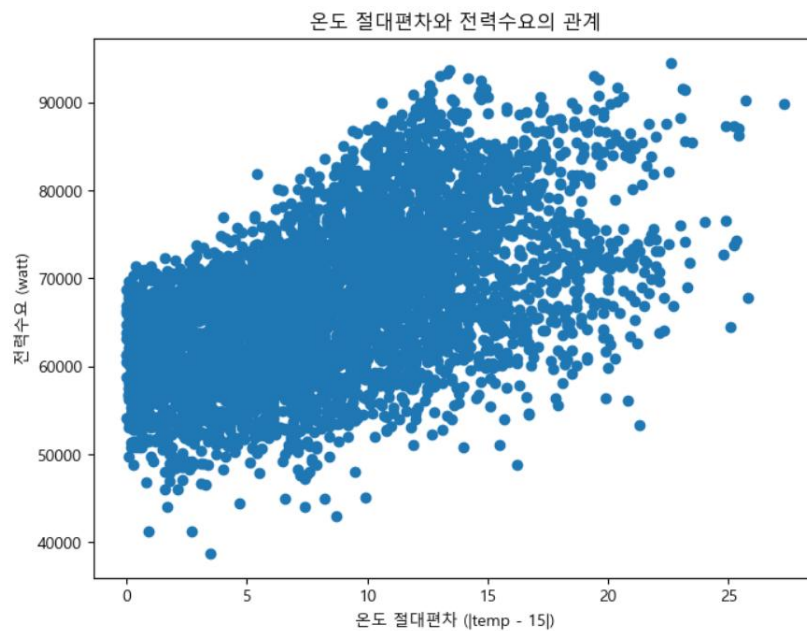
2.3 전력수요와 변수들 간의 관계

2.3.1 시간(time)과 전력 수요(watt)



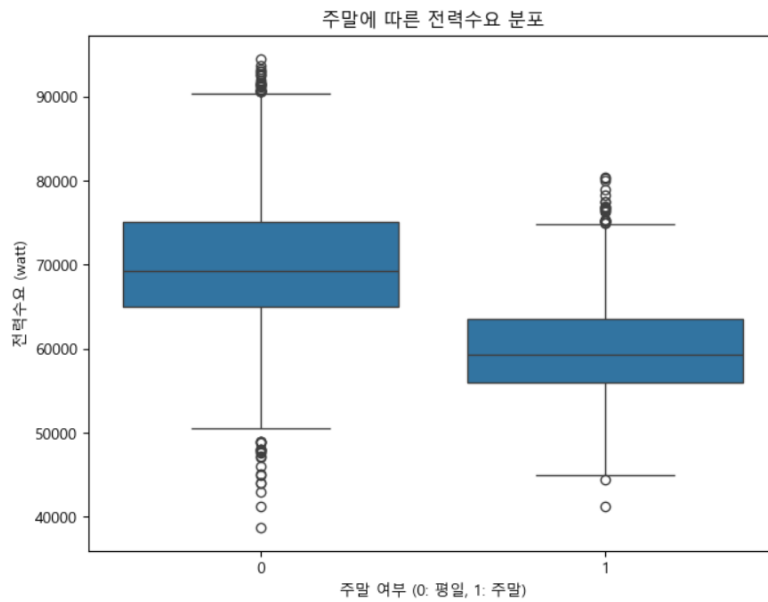
해가 지날수록 전력수요가 증가하는 추세를 보인다.

2.3.2 온도절대편차(temp_abs)와 전력 수요(watt)



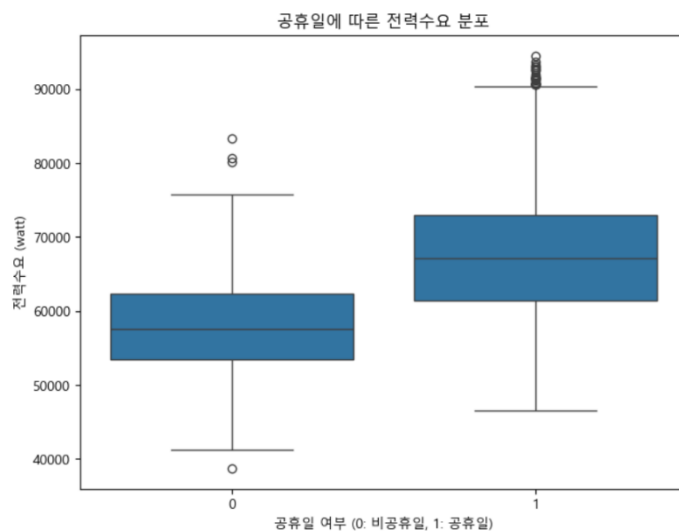
온도와 전력 수요간에 유의미한 상관관계가 있음을 볼 수 있는데, 상관계수를 계산해 보면 상관계수가 0.5377, $p\text{-value} < 0.0001$ 이므로 상관계수가 통계적으로 유의함을 볼 수 있다.

2.3.3 주말(weekend)과 전력 수요(watt)



상자그림을 그려본 결과 주말에 따라 전력수요에 차이가 남을 볼 수 있는데, t 검정 결과 $p\text{-value} < 0.0001$ 이므로 주말에 따른 전력수요 차이가 유의미하게 있음을 볼 수 있다.

2.3.4 공휴일(holiday)과 전력 수요(watt)

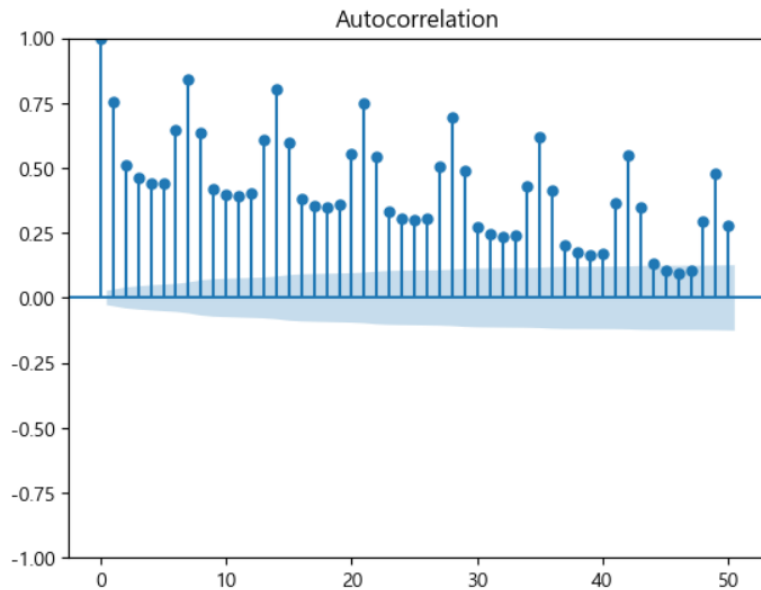


상자그림을 그려본 결과 공휴일에 따라 전력수요에 차이가 남을 볼 수 있는데, t 검정 결과 $p\text{-value} < 0.0001$ 이므로 주말에 따른 전력수요 차이가 유의미하게 있음을 볼 수 있다.

따라서 시간, 온도절대편차, 휴일, 공휴일의 변수는 전력 수요 예측에 영향을 주는 주요 요인이라 보고 이 변수들을 설명변수로 한 모델들로 예측을 수행한다.

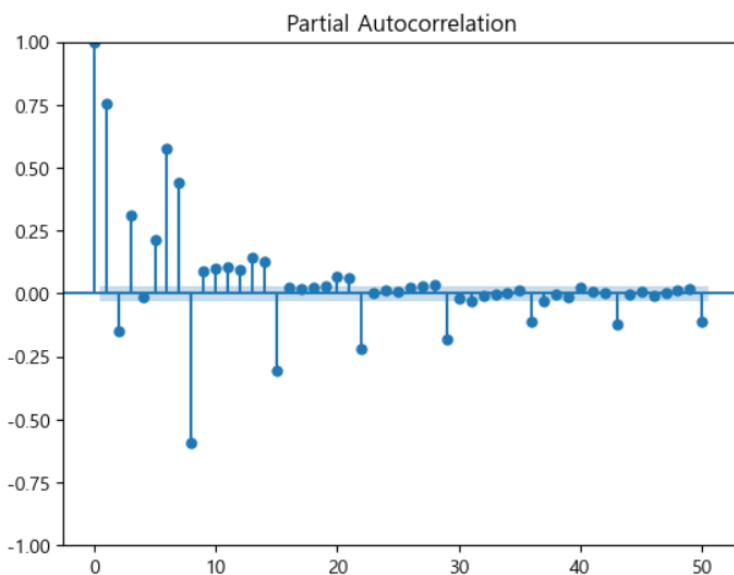
3. 모형 적합, 예측

3.1 ACF와 PACF 플롯



ACF 플롯은 전력 수요의 시계열 데이터에서 각 시차(lag)에 따른 자기상관계수를 보여준다.

시차가 증가함에 따라 자기상관계수가 서서히 감소하는 주기적인 패턴이 보인다. 이는 시계열 데이터가 어느 정도 자기상관을 가지고 있다는 의미이다.



PACF 플롯은 시계열 데이터에서 특정 시차(lag)의 순수한 자기상관만을 계산한다.

lag=8 일때 급격하게 줄어드는 것을 볼 수 있다. 이로 인해 해당 시차 까지만 고유한 상관성이 존재한다고 판단하고, AR(8)모형을 적합해본다.

3.2 설명변수를 사용한 회귀모형 1

$$watt_t = \beta_0 + \beta_1 watt_{t-1} + \beta_2 time + \beta_3 temp_abs + \beta_4 weekend + \beta_5 holiday + \epsilon_t$$

$$\epsilon_t = \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_8 \epsilon_{t-8} + \nu_t$$

3.2.1 SARIMAX 모형 적합 결과

SARIMAX Results						
=====						
Dep. Variable:	watt	No. Observations:	5112			
Model:	SARIMAX(8, 0, 0)	Log Likelihood	-46680.979			
Date:	Fri, 13 Dec 2024	AIC	93391.957			
Time:	01:07:56	BIC	93490.048			
Sample:	0	HQIC	93426.298			
	- 5112					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]

const	2.91e+04	0.005	5.48e+06	0.000	2.91e+04	2.91e+04
lagwatt	0.2071	0.009	24.315	0.000	0.190	0.224
time	23.0594	14.453	1.595	0.111	-5.268	51.387
temp_abs	445.0267	14.509	30.672	0.000	416.589	473.464
weekend	-9357.5095	107.321	-87.192	0.000	-9567.854	-9147.165
holiday	8219.3842	81.126	101.317	0.000	8060.381	8378.388
ar.L1	0.4061	0.013	31.165	0.000	0.381	0.432
ar.L2	0.1517	0.013	11.535	0.000	0.126	0.178
ar.L3	0.0816	0.013	6.407	0.000	0.057	0.107
ar.L4	0.0033	0.013	0.250	0.803	-0.023	0.029
ar.L5	0.0434	0.014	3.203	0.001	0.017	0.070
ar.L6	0.0113	0.013	0.838	0.402	-0.015	0.038
ar.L7	0.4589	0.011	43.315	0.000	0.438	0.480
ar.L8	-0.1577	0.012	-13.260	0.000	-0.181	-0.134
sigma2	5.184e+06	0.021	2.47e+08	0.000	5.18e+06	5.18e+06
=====						
Ljung-Box (L1) (Q):		0.00	Jarque-Bera (JB):	3981.20		
Prob(Q):		0.95	Prob(JB):	0.00		
Heteroskedasticity (H):		1.77	Skew:	-0.06		
Prob(H) (two-sided):		0.00	Kurtosis:	7.32		
=====						

p값이 큰 AR(4)와 AR(6)을 제외한 회귀식은 다음과 같다.

$$\hat{watt}_t = 2.91 \times 10^4 + 0.2071 \cdot lagwatt_{t-1} + 23.0594 \cdot time_t + 445.0267 \cdot temp_abs_t - 9357.5095 \cdot weekend_t + 8219.3842 \cdot holiday_t$$

$$+ 0.4061 \cdot \epsilon_{t-1} + 0.1262 \cdot \epsilon_{t-2} + 0.0816 \cdot \epsilon_{t-3} + 0.0343 \cdot \epsilon_{t-5} + 0.4589 \cdot \epsilon_{t-7} - 0.1577 \cdot \epsilon_{t-8} + \nu_t$$

3.3 log(watt)에 설명변수를 사용한 회귀모형 2

$$\log(watt_t) = \beta_0 + \beta_1 \log(watt_{t-1}) + \beta_2 time + \beta_3 temp_abs + \beta_4 weekend + \beta_5 holiday + \epsilon_t$$

$$\epsilon_t = \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_8 \epsilon_{t-8} + \nu_t$$

3.3.1 SARIMAX 모형 적합 결과

SARIMAX Results						
=====						
Dep. Variable:	log_watt	No. Observations:	5112			
Model:	SARIMAX(8, 0, 0)	Log Likelihood	10091.622			
Date:	Fri, 13 Dec 2024	AIC	-20153.243			
Time:	01:14:24	BIC	-20055.176			
Sample:	0	HQIC	-20118.908			
	- 5112					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]

const	7.2794	0.087	83.553	0.000	7.109	7.450
log_lagwatt	0.3231	0.008	40.145	0.000	0.307	0.339
time	2.902e-05	1.41e-06	20.638	0.000	2.63e-05	3.18e-05
temp_abs	0.0095	0.000	52.827	0.000	0.009	0.010
weekend	-0.1432	0.002	-68.262	0.000	-0.147	-0.139
holiday	0.1318	0.001	107.346	0.000	0.129	0.134
ar.L1	0.2757	0.012	22.062	0.000	0.251	0.300
ar.L2	0.1012	0.013	7.881	0.000	0.076	0.126
ar.L3	0.0330	0.013	2.552	0.011	0.008	0.058
ar.L4	-0.0283	0.013	-2.174	0.030	-0.054	-0.003
ar.L5	0.0331	0.014	2.433	0.015	0.006	0.060
ar.L6	-0.0048	0.013	-0.359	0.720	-0.031	0.022
ar.L7	0.4779	0.010	46.037	0.000	0.458	0.498
ar.L8	-0.1358	0.013	-10.736	0.000	-0.161	-0.111
sigma2	0.0011	1.42e-05	79.275	0.000	0.001	0.001
=====						
Ljung-Box (L1) (Q):	0.03	Jarque-Bera (JB):	3594.15			
Prob(Q):	0.85	Prob(JB):	0.00			
Heteroskedasticity (H):	1.38	Skew:	-0.06			
Prob(H) (two-sided):	0.00	Kurtosis:	7.11			

p값이 가장 큰 AR(6)를 제외한 회귀식은 다음과 같다.

$$\log_watt_t = 7.2794 + 0.3231 \cdot \log_lagwatt_{t-1} + 2.902 \times 10^{-5} \cdot time_t + 0.0095 \cdot temp_abs_t - 0.1438 \cdot weekend_t + 0.1318 \cdot holiday_t$$

$$+ 0.2757 \cdot \epsilon_{t-1} + 0.1211 \cdot \epsilon_{t-2} + 0.0330 \cdot \epsilon_{t-3} - 0.0283 \cdot \epsilon_{t-4} + 0.0348 \cdot \epsilon_{t-5} + 0.4779 \cdot \epsilon_{t-7} - 0.1353 \cdot \epsilon_{t-8} + \nu_t$$

3.4 회귀모형 1, 2에 의한 예측

2023년 12월 01일부터 31일까지 한 달 동안의 데이터를 예측 후, 하나의 데이터프레임으로 정리하였다.

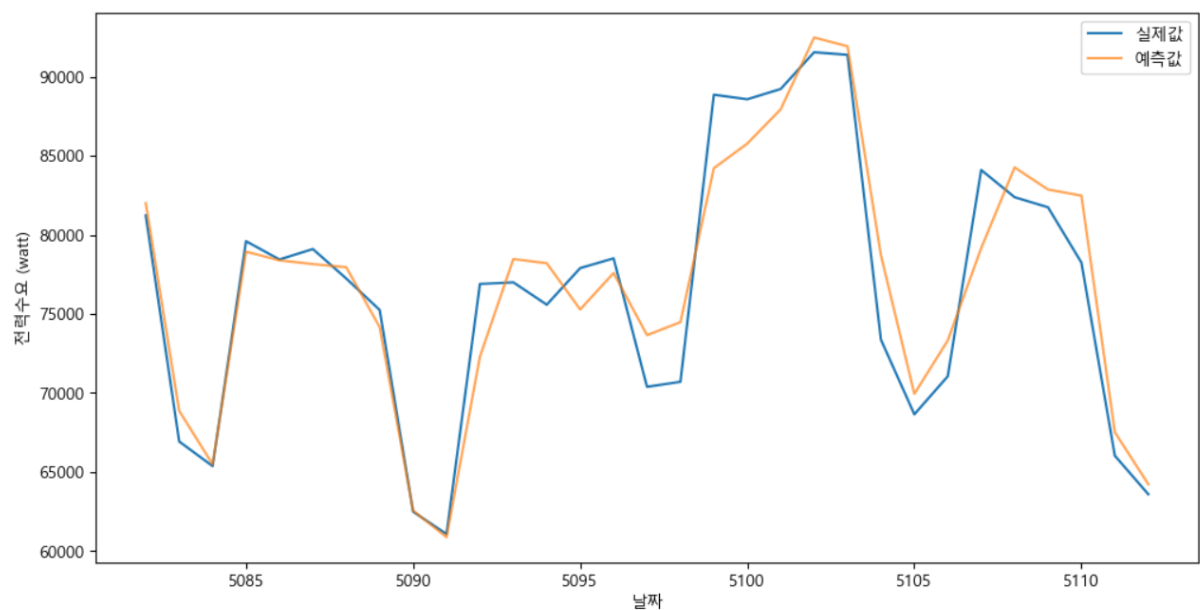
Date	Forecast_Watt	Forecast_Log_Watt
2023-12-01 00:00:00	81699.2693	83071.39621
2023-12-02 00:00:00	69232.80403	68347.74579
2023-12-03 00:00:00	65202.65797	64897.84958
2023-12-04 00:00:00	78394.00692	79780.97733
2023-12-05 00:00:00	78553.91014	79191.53823
2023-12-06 00:00:00	78694.91973	78683.17285
2023-12-07 00:00:00	78553.33402	77669.72726
2023-12-08 00:00:00	75120.68961	73528.35673
2023-12-09 00:00:00	63354.6091	62089.70428
2023-12-10 00:00:00	61207.8403	59977.26807
2023-12-11 00:00:00	72619.86233	71101.8508
2023-12-12 00:00:00	77918.93745	78630.01585
2023-12-13 00:00:00	78075.5762	78605.17602
2023-12-14 00:00:00	75354.84928	75050.36935
2023-12-15 00:00:00	77092.19146	77580.08545
2023-12-16 00:00:00	71477.86144	73212.37025
2023-12-17 00:00:00	72049.86137	74675.52807
2023-12-18 00:00:00	82838.37021	87317.32132
2023-12-19 00:00:00	85497.06818	87484.02912
2023-12-20 00:00:00	86991.79348	89972.91564
2023-12-21 00:00:00	90861.98679	95810.06251
2023-12-22 00:00:00	90615.74072	94666.99588
2023-12-23 00:00:00	78863.54434	77748.58826
2023-12-24 00:00:00	70680.38112	69020.44282
2023-12-25 00:00:00	73723.01778	71288.54644
2023-12-26 00:00:00	79361.0728	79034.41636
2023-12-27 00:00:00	84881.08078	85010.84988
2023-12-28 00:00:00	83627.35199	82704.7869
2023-12-29 00:00:00	82891.61332	81948.42542
2023-12-30 00:00:00	67571.81487	67122.47849
2023-12-31 00:00:00	64480.25152	63614.48401

3.4.1 회귀모형 1로 산출한 예측값과 실제값들의 비교

예측값들을 실제값과 비교하여 오차를 확인해보면 다음과 같다.

Date	Forecast_Watt	Actual_Watt	Error_Watt	Absolute_Error_Watt	Absolute_Error_Percentage_Watt
2023-12-01 00:00:00	82007.23027	81228	779.2302732	779.2302732	0.009593124
2023-12-02 00:00:00	68861.90945	66924	1937.909455	1937.909455	0.028956868
2023-12-03 00:00:00	65482.65562	65368	114.6556154	114.6556154	0.001754002
2023-12-04 00:00:00	78937.00613	79599	-661.9938666	661.9938666	0.00831661
2023-12-05 00:00:00	78386.96979	78441	-54.03021248	54.03021248	0.000688801
2023-12-06 00:00:00	78150.00473	79102	-951.9952703	951.9952703	0.012035034
2023-12-07 00:00:00	77950.39914	77237	713.3991427	713.3991427	0.009236495
2023-12-08 00:00:00	74159.48104	75245	-1085.518956	1085.518956	0.01442646
2023-12-09 00:00:00	62554.35995	62483	71.35994827	71.35994827	0.00114207
2023-12-10 00:00:00	60870.92943	61067	-196.0705704	196.0705704	0.003210745
2023-12-11 00:00:00	72293.42956	76893	-4599.570441	4599.570441	0.059817804
2023-12-12 00:00:00	78470.32164	76995	1475.321638	1475.321638	0.019161266
2023-12-13 00:00:00	78203.56993	75581	2622.569929	2622.569929	0.034698799
2023-12-14 00:00:00	75276.90117	77896	-2619.09883	2619.09883	0.033623021
2023-12-15 00:00:00	77588.08454	78513	-924.9154601	924.9154601	0.011780412
2023-12-16 00:00:00	73653.75142	70383	3270.751424	3270.751424	0.046470759
2023-12-17 00:00:00	74476.76128	70703	3773.761281	3773.761281	0.05337484
2023-12-18 00:00:00	84216.74787	88869	-4652.252131	4652.252131	0.05234955
2023-12-19 00:00:00	85775.6232	88580	-2804.376804	2804.376804	0.031659255
2023-12-20 00:00:00	87942.84192	89230	-1287.158078	1287.158078	0.014425172
2023-12-21 00:00:00	92488.38267	91556	932.3826687	932.3826687	0.010183742
2023-12-22 00:00:00	91934.44142	91391	543.4414168	543.4414168	0.005946334
2023-12-23 00:00:00	78713.30757	73369	5344.307565	5344.307565	0.072841494
2023-12-24 00:00:00	69940.43261	68643	1297.432613	1297.432613	0.018901164
2023-12-25 00:00:00	73318.97889	71050	2268.978886	2268.978886	0.03193496
2023-12-26 00:00:00	79157.68656	84106	-4948.313441	4948.313441	0.05883425
2023-12-27 00:00:00	84266.14772	82383	1883.147718	1883.147718	0.02285845
2023-12-28 00:00:00	82868.9757	81745	1123.975697	1123.975697	0.013749779
2023-12-29 00:00:00	82481.83361	78244	4237.833613	4237.833613	0.054161771
2023-12-30 00:00:00	67510.10519	66025	1485.105195	1485.105195	0.022493074
2023-12-31 00:00:00	64225.28047	63589	636.2804733	636.2804733	0.010006141

오차의 절대값, 오차의 절댓값 / 실제값 비율도 하나의 데이터프레임에 담았다.



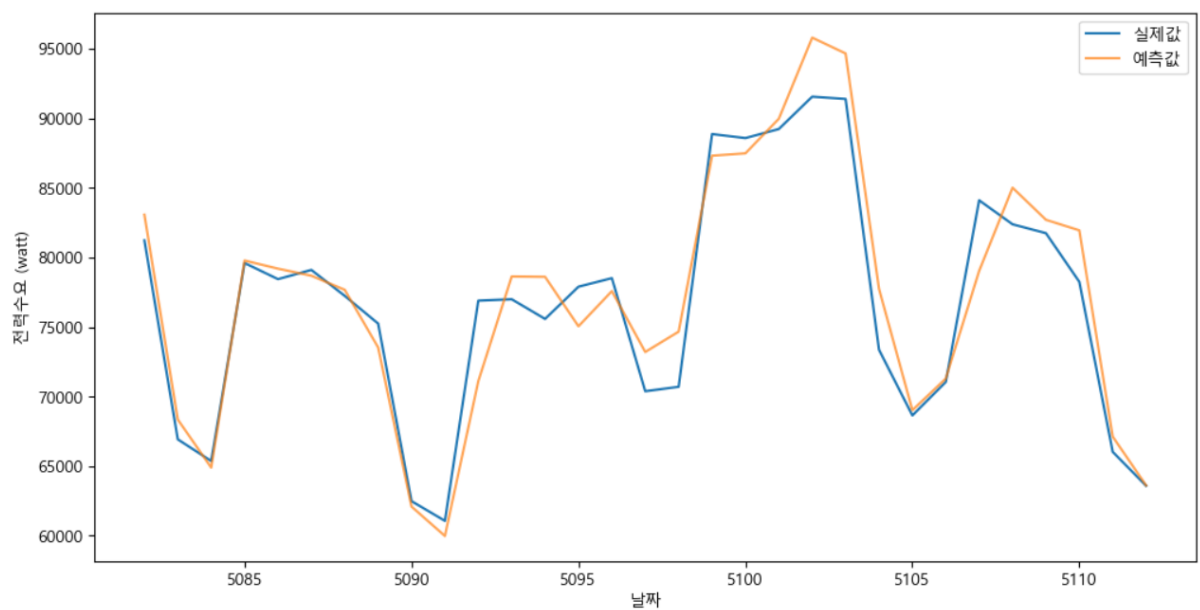
실제값과 예측값이 유사한 패턴을 보이며, 전반적인 추세를 잘 반영하고 있음을 볼 수 있다.

3.4.2 회귀모형 2로 산출한 예측값과 실제값들의 비교

예측값들을 실제값과 비교하여 오차를 확인해보면 다음과 같다.

Date	Forecast_Log_Watt	Actual_Watt	Error_Log_Watt	Absolute_Error_Log_Watt	Absolute_Error_Percentage_Log_Watt
2023-12-01 00:00:00	83069.25333	81228	1841.253329	1841.253329	0.022667717
2023-12-02 00:00:00	68346.90318	66924	1422.903176	1422.903176	0.021261478
2023-12-03 00:00:00	64896.62994	65368	-471.3700598	471.3700598	0.007211022
2023-12-04 00:00:00	79779.3577	79599	180.357701	180.357701	0.002265829
2023-12-05 00:00:00	79193.02687	78441	752.0268747	752.0268747	0.009587166
2023-12-06 00:00:00	78685.7428	79102	-416.2572043	416.2572043	0.005262284
2023-12-07 00:00:00	77671.75771	77237	434.7577069	434.7577069	0.005628879
2023-12-08 00:00:00	73531.05912	75245	-1713.940878	1713.940878	0.022778136
2023-12-09 00:00:00	62091.95167	62483	-391.0483309	391.0483309	0.006258476
2023-12-10 00:00:00	59979.14608	61067	-1087.85392	1087.85392	0.017814105
2023-12-11 00:00:00	71103.51816	76893	-5789.481841	5789.481841	0.075292703
2023-12-12 00:00:00	78629.23898	76995	1634.23898	1634.23898	0.021225261
2023-12-13 00:00:00	78605.65728	75581	3024.657275	3024.657275	0.040018752
2023-12-14 00:00:00	75050.92616	77896	-2845.073836	2845.073836	0.036524004
2023-12-15 00:00:00	77579.25737	78513	-933.7426309	933.7426309	0.011892841
2023-12-16 00:00:00	73204.19402	70383	2821.19402	2821.19402	0.040083458
2023-12-17 00:00:00	74667.14291	70703	3964.142912	3964.142912	0.056067535
2023-12-18 00:00:00	87312.77711	88869	-1556.222893	1556.222893	0.017511426
2023-12-19 00:00:00	87483.24451	88580	-1096.755486	1096.755486	0.012381525
2023-12-20 00:00:00	89967.07767	89230	737.0776664	737.0776664	0.008260424
2023-12-21 00:00:00	95800.52981	91556	4244.529805	4244.529805	0.046359931
2023-12-22 00:00:00	94659.63471	91391	3268.634713	3268.634713	0.03576539
2023-12-23 00:00:00	77747.5646	73369	4378.564599	4378.564599	0.059678674
2023-12-24 00:00:00	69021.10134	68643	378.1013406	378.1013406	0.005508229
2023-12-25 00:00:00	71291.62636	71050	241.6263575	241.6263575	0.003400793
2023-12-26 00:00:00	79032.85619	84106	-5073.143806	5073.143806	0.060318453
2023-12-27 00:00:00	85013.05658	82383	2630.056585	2630.056585	0.031924749
2023-12-28 00:00:00	82706.94341	81745	961.9434096	961.9434096	0.011767612
2023-12-29 00:00:00	81947.70852	78244	3703.70852	3703.70852	0.047335368
2023-12-30 00:00:00	67120.79344	66025	1095.79344	1095.79344	0.016596644
2023-12-31 00:00:00	63613.32962	63589	24.32961815	24.32961815	0.000382607

오차의 절대값, 오차의 절댓값 / 실제값의 비율도 하나의 데이터프레임에 담았다.



실제값과 예측값이 유사한 패턴을 보이며, 전반적인 추세를 잘 반영하고 있음을 볼 수 있다.

3.5 XGboost 모형

XGBoost(Extreme Gradient Boosting)는 그래디언트 부스팅(Gradient Boosting) 알고리즘을 기반으로 한 머신러닝 라이브러리로, 대용량 데이터와 복잡한 모델에서도 빠르고 효율적인 성능을 자랑하는 도구이다. 주로 결정 트리를 기반으로 약한 학습기를 반복적으로 추가하여 예측 성능을 향상시킨다. 회귀(regression), 분류(classification), 랭킹(ranking) 등 다양한 머신러닝 작업에 사용된다.

XGBoost 모형식은 기본적으로 다음과 같이 표현된다.

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), \quad f_k \in \mathcal{F}$$

이는 주어진 데이터셋 (x_i, y_i) , $i=1, \dots, n$ 에서 y_i 를 예측하는 함수를 학습하는 것으로 여러 트리 $f_k(x)$ 들을 합산하여 예측값을 출력한다. $f_k(x)$ 는 k 번째 트리에서 학습된 예측함수이고, \mathcal{F} 는 트리 함수의 공간, K 는 학습할 트리의 개수이다.

XGBoost는 기존 모델에 새 트리를 반복적으로 학습시켜 잔차를 줄이는데 이는 다음의 손실 함수 L 을 최소화함으로 수행된다. 손실함수는 예측값 \hat{y}_i 와 실제값 y_i 간의 차이를 측정하는 함수로 다음과 같이 표현된다.

$$L(\phi) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k)$$

$l(y_i, \hat{y}_i)$ 는 예측값과 실제값 간의 차이를 의미하고 $\Omega(f_k)$ 는 정규화 항으로 트리의 복잡도를 조절하고 과적합을 방지하는 항이다. 정규화 항은 또 다음과 같이 표현된다.

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2$$

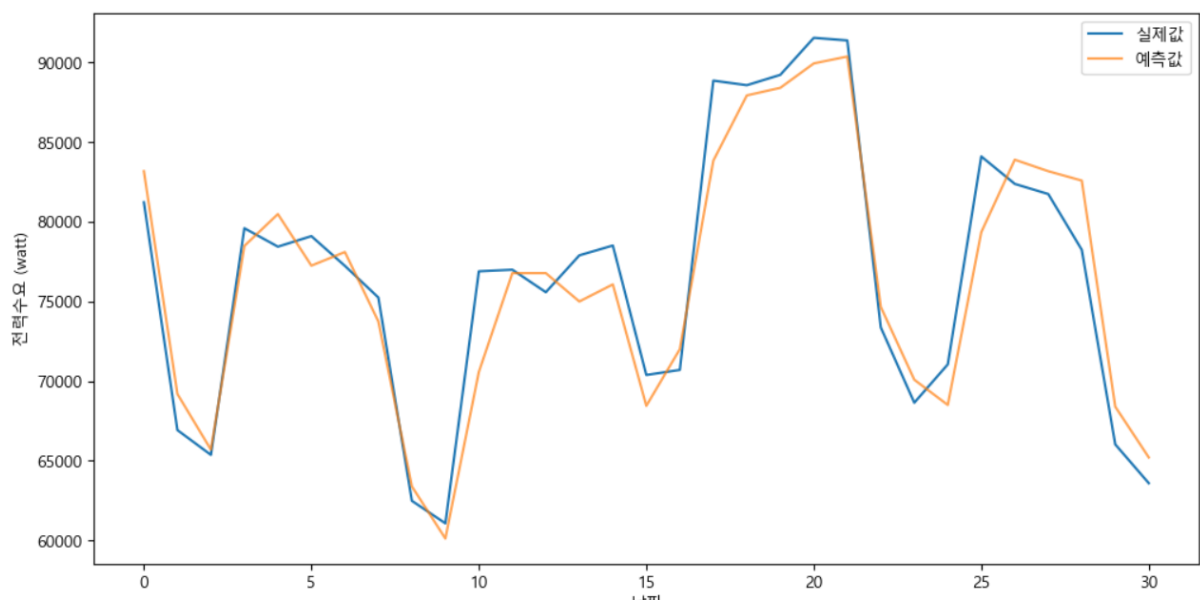
정규화 항은 트리의 리프 개수와 가중치 크기를 제한하여 모형의 과적합을 방지한다. 여기서 T 는 트리의 리프 노드 개수, w_j 는 리프 노드 j 의 가중치, γ 는 리프 개수에 대한 패널티를 조절하는 매개변수, λ 는 가중치의 크기에 대한 패널티를 조절하는 매개변수이다.

3.5.1 XGBoost 모형으로 예측, 비교

예측 후, 실제값과 비교하여 오차를 확인해보면 다음과 같다.

Date	Forecast_XGBoost_Watt	Actual_Watt	Error	Absolute_Error	Absolute_Error_Percentage
2023-12-01 00:00:00	83180.14063	81228	1952.141	1952.140625	0.024032854
2023-12-02 00:00:00	69193.09375	66924	2269.094	2269.09375	0.033905531
2023-12-03 00:00:00	65695.875	65368	327.875	327.875	0.005015833
2023-12-04 00:00:00	78488.03125	79599	-1110.97	1110.96875	0.013957069
2023-12-05 00:00:00	80487.5625	78441	2046.563	2046.5625	0.026090469
2023-12-06 00:00:00	77249.875	79102	-1852.13	1852.125	0.023414389
2023-12-07 00:00:00	78105.35938	77237	868.3594	868.359375	0.01124279
2023-12-08 00:00:00	73743.75781	75245	-1501.24	1501.242188	0.019951388
2023-12-09 00:00:00	63388.72266	62483	905.7227	905.7226563	0.014495505
2023-12-10 00:00:00	60114.21094	61067	-952.789	952.7890625	0.015602356
2023-12-11 00:00:00	70575.875	76893	-6317.13	6317.125	0.082154748
2023-12-12 00:00:00	76778.15625	76995	-216.844	216.84375	0.002816335
2023-12-13 00:00:00	76778.15625	75581	1197.156	1197.15625	0.015839381
2023-12-14 00:00:00	74998.40625	77896	-2897.59	2897.59375	0.037198235
2023-12-15 00:00:00	76072.20313	78513	-2440.8	2440.796875	0.031087806
2023-12-16 00:00:00	68447.5	70383	-1935.5	1935.5	0.027499538
2023-12-17 00:00:00	72005.5	70703	1302.5	1302.5	0.018422132
2023-12-18 00:00:00	83839.29688	88869	-5029.7	5029.703125	0.056596824
2023-12-19 00:00:00	87937.70313	88580	-642.297	642.296875	0.007251037
2023-12-20 00:00:00	88415.59375	89230	-814.406	814.40625	0.009127045
2023-12-21 00:00:00	89948.58594	91556	-1607.41	1607.414063	0.017556622
2023-12-22 00:00:00	90380.63281	91391	-1010.37	1010.367188	0.011055434
2023-12-23 00:00:00	74656.875	73369	1287.875	1287.875	0.017553394
2023-12-24 00:00:00	70085.66406	68643	1442.664	1442.664063	0.021016915
2023-12-25 00:00:00	68499.84375	71050	-2550.16	2550.15625	0.035892417
2023-12-26 00:00:00	79350.92188	84106	-4755.08	4755.078125	0.056536729
2023-12-27 00:00:00	83901.96094	82383	1518.961	1518.960938	0.018437796
2023-12-28 00:00:00	83180.14063	81745	1435.141	1435.140625	0.017556311
2023-12-29 00:00:00	82589.91406	78244	4345.914	4345.914063	0.055543097
2023-12-30 00:00:00	68393.00781	66025	2368.008	2368.007813	0.035865321
2023-12-31 00:00:00	65206.89453	63589	1617.895	1617.894531	0.025442994

오차의 절대값, 오차의 절댓값 / 실제값의 비율도 하나의 데이터프레임에 담았다.



실제값과 예측값이 유사한 패턴을 보이며, 전반적인 추세를 잘 반영하고 있음을 볼 수 있다.

3.6 예측 성능 평가

	Model	MAE	MAPE
0	Linear Regression	1912.810923	2.479459
1	Log(Watt) AR Model	1906.928675	2.448489
2	XGBoost	1952.266885	2.542446

회귀모형 1, 회귀모형 2, XGBoost 모형을 MAE, MAPE의 지표로 예측의 정확성을 평가한 결과, 회귀모형 2가 가장 좋은 예측력을 보인다.

4. 결론

본연구는 전력수요 예측을 위해 전력수요와 연관성이 높은 설명변수 온도, 주말, 공휴일 변수를 추가하여 예측비교를 수행하였다. 설명변수를 사용한 회귀모형 1, log(watt)에 설명변수를 사용한 회귀모형 2 그리고 XGBoost 모형을 이용해 전력수요에 대한 예측을 수행하였고, 그 결과 MAE와 MAPE값이 가장 작은 회귀모형2가 가장 우수한 성능을 보였다고 할 수 있다.