

<Solar Energy Production Analysis>

컴퓨터공학부 202001581 신소정

주제 – Solar Energy Production Analysis(태양 에너지 생성 분석)

- 주제 선정 배경

지구 온난화의 가속화에 따라 기온의 변화가 계절성을 파괴하고 있다.

그렇다면, 태양광 에너지의 변화는 시간의 흐름에도 여전히 계절성을 동일하게 띠 것인가?

계절성을 제외한다면 어떤 변화가 일어나는가?

- 정의

시계열 데이터(y)는 추세(trend) + 계절성(seasonality) + 나머지(remainder)로 분해(decomposition)된다.

- 데이터 셋 (Data Set)

<https://www.kaggle.com/datasets/ivnlee/solar-energy-production/data>

해당 데이터 셋은 한 시간 마다 Calgary 에 위치한 태양광 패널에서 취득되는 전력량을 5 년 간(2017 년~2022 년) 측정.

- 데이터 크기: 3MB
- 데이터 형식: csv 형식

1. <그림 1>과 같은 형태로 5 개의 고유한 id 를 가진 태양광 패널들의 <표 1>과 같은 정보를 제공.
2. 태양광 패널들은 지역의 위치(위도, 경도 상)에 따라 고유한 id 를 갖음.

Solar_Energy_Production							
name	id	address	date	kWh	public_url	installationDate	uid
Calgary Fire Hall Headquarters	314106	1212 42 AV SE	2017/09/11 08:00:00 AM	1.130	https://monitoring.solaredge.com/solaredge-web/p/public?name=Fire%20HQ	2016/11/07	3141062017-09-11 08:00:00
Calgary Fire Hall Headquarters	314106	1212 42 AV SE	2017/09/11 09:00:00 AM	2.340	https://monitoring.solaredge.com/solaredge-web/p/public?name=Fire%20HQ	2016/11/07	3141062017-09-11 09:00:00
Calgary Fire Hall Headquarters	314106	1212 42 AV SE	2017/09/11 10:00:00 AM	3.656	https://monitoring.solaredge.com/solaredge-web/p/public?name=Fire%20HQ	2016/11/07	3141062017-09-11 10:00:00
Calgary Fire Hall Headquarters	314106	1212 42 AV SE	2017/09/11 11:00:00 AM	4.577	https://monitoring.solaredge.com/solaredge-web/p/public?name=Fire%20HQ	2016/11/07	3141062017-09-11 11:00:00
Calgary Fire Hall Headquarters	314106	1212 42 AV SE	2017/09/11 12:00:00 PM	6.506	https://monitoring.solaredge.com/solaredge-web/p/public?name=Fire%20HQ	2016/11/07	3141062017-09-11 12:00:00
Calgary Fire Hall Headquarters	314106	1212 42 AV SE	2017/09/11 01:00:00 PM	7.031	https://monitoring.solaredge.com/solaredge-web/p/public?name=Fire%20HQ	2016/11/07	3141062017-09-11 13:00:00
Calgary Fire Hall Headquarters	314106	1212 42 AV SE	2017/09/11 02:00:00 PM	9.218	https://monitoring.solaredge.com/solaredge-web/p/public?name=Fire%20HQ	2016/11/07	3141062017-09-11 14:00:00
Calgary Fire Hall Headquarters	314106	1212 42 AV SE	2017/09/11 03:00:00 PM	9.018	https://monitoring.solaredge.com/solaredge-web/p/public?name=Fire%20HQ	2016/11/07	3141062017-09-11 15:00:00
Calgary Fire Hall Headquarters	314106	1212 42 AV SE	2017/09/11 04:00:00 PM	5.266	https://monitoring.solaredge.com/solaredge-web/p/public?name=Fire%20HQ	2016/11/07	3141062017-09-11 16:00:00

<그림 1>

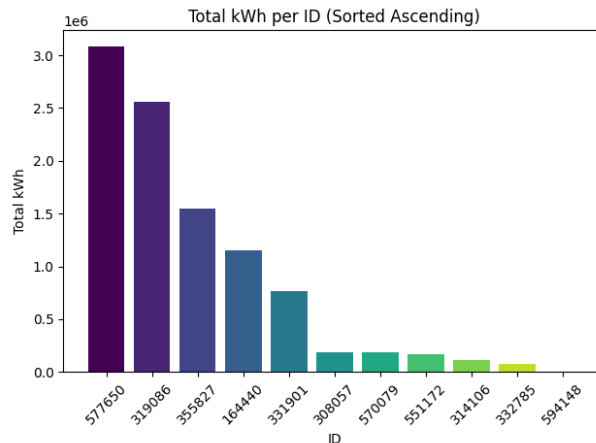
Name	Id	Address	Date	kWh	Public_url	Installation	Uid
패널 이름	지역 고유 id	패널 설치 주소	측정 일자	측정 전력	측정 데이터 원본	설치 일자	정보

<표 1. Header 정보>

- 데이터 전처리 (Data cleaning)

1. Id 별로 그룹화하고, 에너지 총 측정 량이 많은 태양광 패널 id 선택

본 데이터는 id 별로 데이터가 무작위로 섞여 있기에 그룹화하여 전력 측정 량이 많은 순으로 정렬한다. id 별로 시계열 데이터 분석을 진행해야 하기에, 측정량이 가장 많은 id=577650, Bearspaw Water Treatment Plant 의 태양광 에너지 생산량 데이터를 사용한다.



<그림 2>

2. 데이터의 일관성 확인을 위한 누락된 데이터 확인

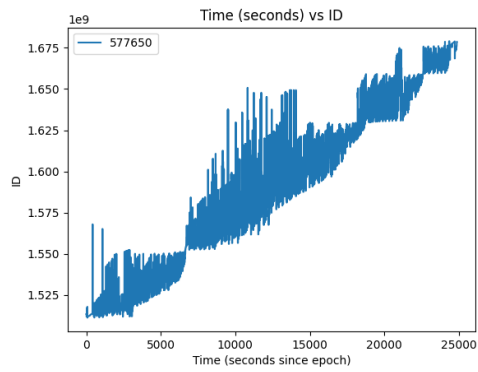
2-1. <표 1>의 4 번째 열 데이터를 이용하여 year-month-day 형식의 time 을 UTC(유닉스 타임 스탬프 시스템) 기준 초단위로 변환하여 0 번째 열 벡터 추가.

seconds	name	id	address	kWh				
1511261100.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/21 10:45:00 AM	0.970	https://monitoring.solaredge.com/solaredge-web/p/public?name=Bearspaw	2017/10/31	5776502017-11-21 10:45:00
1511264700.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/21 11:45:00 AM	23.927	https://monitoring.solaredge.com/solaredge-web/p/public?name=Bearspaw	2017/10/31	5776502017-11-21 11:45:00
1511268300.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/21 12:45:00 PM	69.659	https://monitoring.solaredge.com/solaredge-web/p/public?name=Bearspaw	2017/10/31	5776502017-11-21 12:45:00
1511271900.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/21 01:45:00 PM	59.172	https://monitoring.solaredge.com/solaredge-web/p/public?name=Bearspaw	2017/10/31	5776502017-11-21 13:45:00
1511275500.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/21 02:45:00 PM	18.061	https://monitoring.solaredge.com/solaredge-web/p/public?name=Bearspaw	2017/10/31	5776502017-11-21 14:45:00
1511340300.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 08:45:00 AM	12.285	https://monitoring.solaredge.com/solaredge-web/p/public?name=Bearspaw	2017/10/31	5776502017-11-22 08:45:00
1511341200.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 09:00:00 AM	20.961	https://monitoringpublic.solaredge.com/solaredge-web/p/site/public?name=Bearspaw	2017/10/31	5776502017-11-22 09:00:00
1511343900.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 09:45:00 AM	42.792	https://monitoring.solaredge.com/solaredge-web/p/public?name=Bearspaw	2017/10/31	5776502017-11-22 09:45:00
1511344800.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 10:00:00 AM	50.230	https://monitoringpublic.solaredge.com/solaredge-web/p/site/public?name=Bearspaw	2017/10/31	5776502017-11-22 10:00:00
1511347900.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 10:45:00 AM	166.551	https://monitoring.solaredge.com/solaredge-web/p/public?name=Bearspaw	2017/10/31	5776502017-11-22 10:45:00

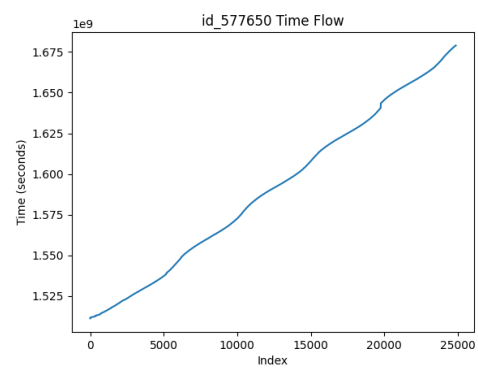
2-2. <그림 3>와 같이 시간의 변화(x 축)에 따라 측정 시간(y 축)이 선형적이지 않으므로 시간이 뒤섞여 있음.

2-3. 시간을 기준(초 단위)으로 <그림 4>과 같이 그래프 표현이 가능.

하지만 이 역시 선형적이지 않고, 5182(2018/10/04 10:00:00 AM)-5183(2018/10/09 02:00:00 PM) 인덱스와 19760(2021/12/28 02:00:00 PM) -19761(2022/01/29 08:00:00 AM) 인덱스 부분 등에서 측정이 되지 않은 시간대가 발생하여 데이터 누락이 확인됨.

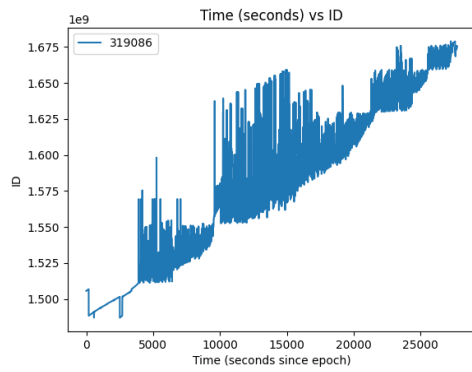


<그림 3>

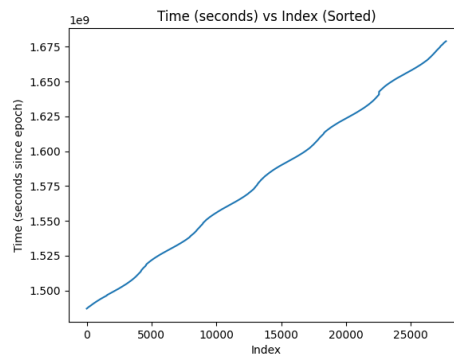


<그림 4>

**** <그림 5>와 <그림 6>과 같이 'id = 319086' Whitehorn 의 경우에도 전처리과정이 필요.**

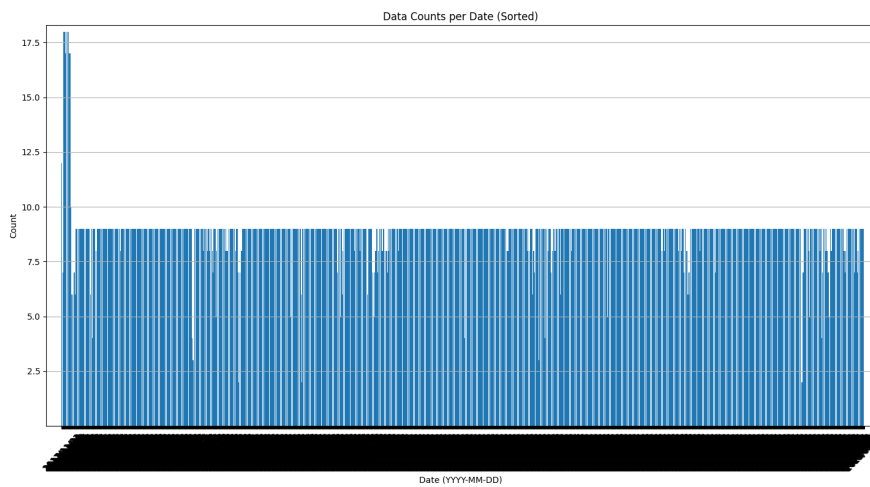


<그림 5>



<그림 6>

3. Id=577650 인 태양광 패널에서 하루에 에너지를 측정한 빈도수 확인



2017 년도 앞 쪽 데이터만 이후 데이터에 비해 높은 빈도수로 데이터가 추출된 것으로 확인. 이후 데이터 빈도수는 상대적으로 고름.

4. Raw data 확인

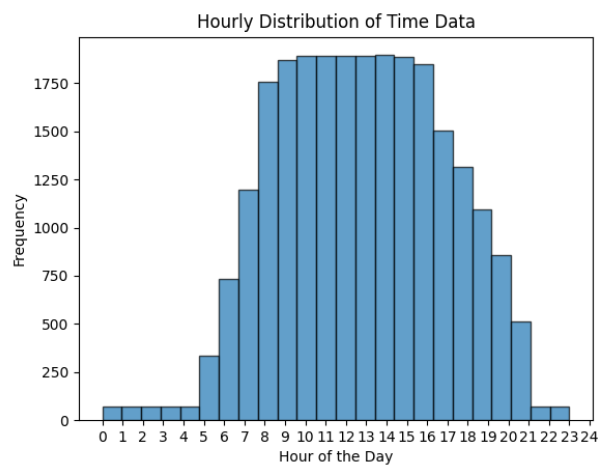
<표 2>와 같이 한 달 가량 15 분, 45 분 단위로 데이터를 측정하다가 1 시간 단위로 데이터 측정하는 모습을 볼 수 있음. 따라서 2017-11-21 부터 약 한 달 가량의 데이터를 제외하고, 정각에 취득된 데이터만 취득.

1511341200.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 09:00:00 AM	20.961	2017/10/31	5776502017-11-22 09:00:00
1511343900.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 09:45:00 AM	42.792	2017/10/31	5776502017-11-22 09:45:00
1511344800.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 10:00:00 AM	50.23	2017/10/31	5776502017-11-22 10:00:00
1511347500.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 10:45:00 AM	166.551	2017/10/31	5776502017-11-22 10:45:00
1511348400.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 11:00:00 AM	236.6	2017/10/31	5776502017-11-22 11:00:00
1511351100.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 11:45:00 AM	380.227	2017/10/31	5776502017-11-22 11:45:00
1511352000.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 12:00:00 PM	399.535	2017/10/31	5776502017-11-22 12:00:00
1511354700.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 12:45:00 PM	399.868	2017/10/31	5776502017-11-22 12:45:00
1511355600.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 01:00:00 PM	398.994	2017/10/31	5776502017-11-22 13:00:00
1511358300.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 01:45:00 PM	289.615	2017/10/31	5776502017-11-22 13:45:00
1511359200.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/22 02:00:00 PM	185.018	2017/10/31	5776502017-11-22 14:00:00
1511426700.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/23 08:45:00 AM	6.638	2017/10/31	5776502017-11-23 08:45:00
1511427600.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/23 09:00:00 AM	9.015	2017/10/31	5776502017-11-23 09:00:00
1511430300.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/23 09:45:00 AM	5.205	2017/10/31	5776502017-11-23 09:45:00
1511434000.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/11/23 10:00:00 AM	0.898	2017/10/31	5776502017-11-23 10:00:00
...							
151377500.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/12/20 01:45:00 PM	1.339	2017/10/31	5776502017-12-20 13:45:00
1513778400.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/12/20 02:00:00 PM	0.837	2017/10/31	5776502017-12-20 14:00:00
1513781100.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/12/20 02:45:00 PM	0.133	2017/10/31	5776502017-12-20 14:45:00
1513782000.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/12/20 03:00:00 PM	0.06	2017/10/31	5776502017-12-20 15:00:00
1513846800.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/12/21 09:00:00 AM	0.004	2017/10/31	5776502017-12-21 09:00:00
1513850400.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/12/21 10:00:00 AM	0.433	2017/10/31	5776502017-12-21 10:00:00
1513854000.0	Bearspaw Water Treatment Plant	577650	11444 Bearspaw Dam RD NW	2017/12/21 11:00:00 AM	0.854	2017/10/31	5776502017-12-21 11:00:00

<표 2. Raw data 확인>

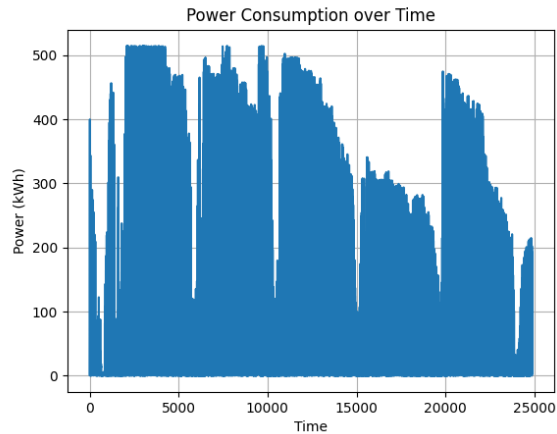
5. 5 년 간 시간별 데이터 측정 빈도 수 비교

8 시-16 시(8 시간)의 데이터가 균일한 정도로 데이터가 취득이 됐으므로 해당 시간대 측정 량만 사용.



- Exploratory 데이터 분석

1. Id=577650 패널의 시간 별 kWh 양 시각화



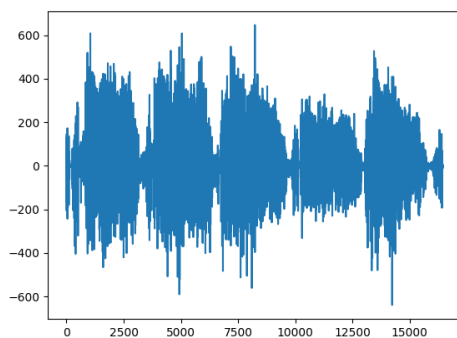
2. Arima 모델을 사용하여 2 차 차분을 구하고, stationarity 확인.

7-1. stationary 한 데이터는 arima 분석 가능.

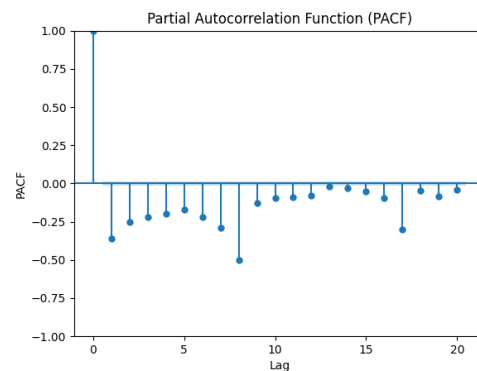
7-2. <그림 5>와 같이 차분을 2 회 실시하여 그래프로 나타냄.

7-3. <그림 6>과 같이 PACF 를 이용하여 stationary 를 확인. 신뢰구간에 들어오는 상관관계가 없음.

7-4. 따라서 본 데이터는 seasonality 의 성질이 강하므로 Arima 분석 어려움.



<그림 5>



<그림 6>

3. 추세 그래프 시각화 (Trend plotting)

3-1. Moving average (pink) = 8 사용. (하루에 8 시간씩 데이터 수집)

3-2. Loess = 0.1 사용 (red)

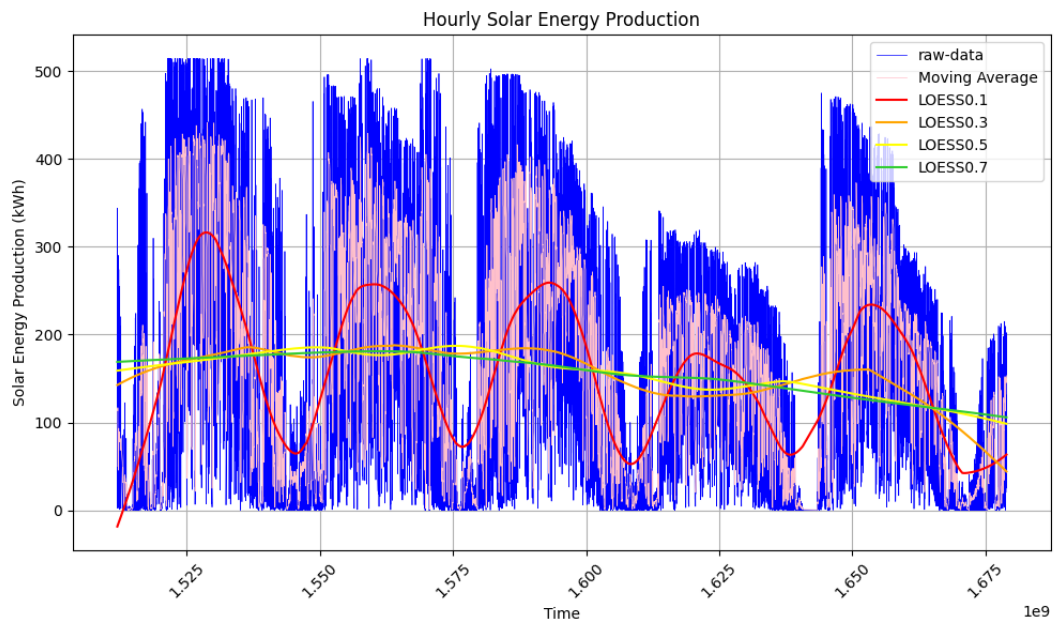
3-3. Loess = 0.3 사용 (orange)

3-4. Loess = 0.5 사용 (yellow)

3-5. Loess = 0.7 사용 (green)

5 년 간의 전체적인 추이를 보기에는 Loess = 0.7 이 가장 적합.

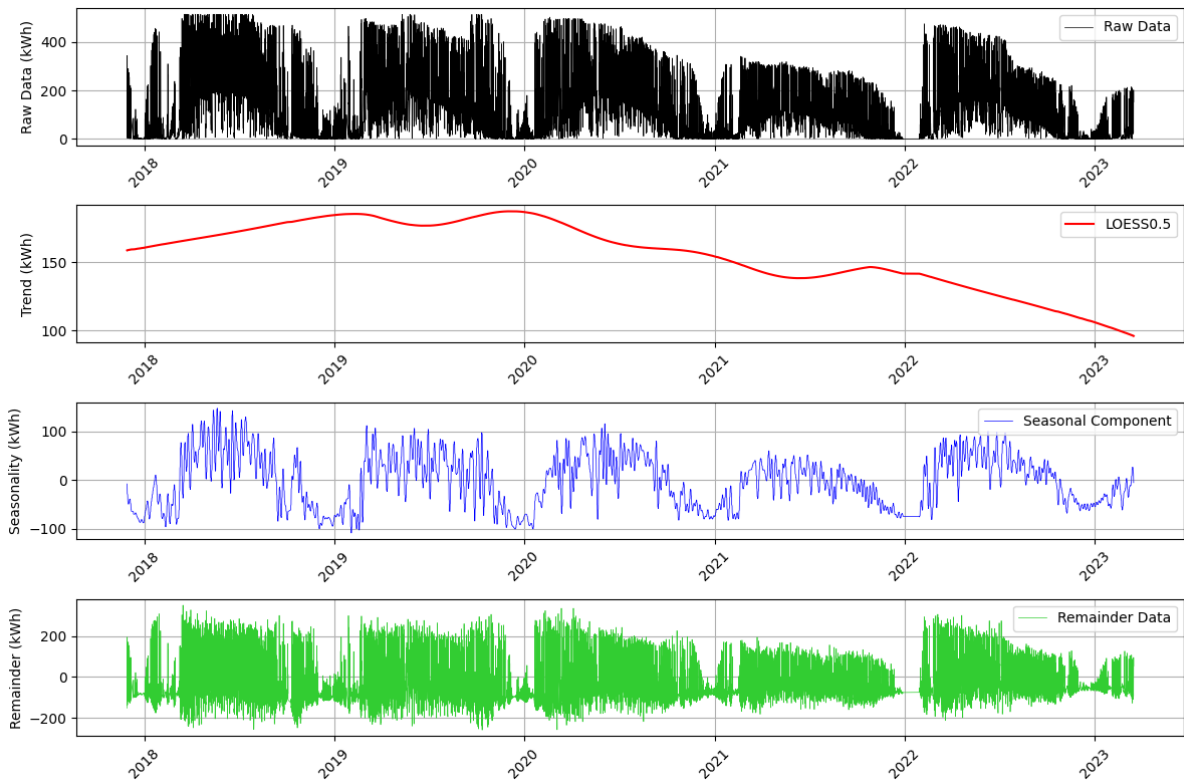
태양광 에너지가 년 간 약 10kWh 정도 감소.



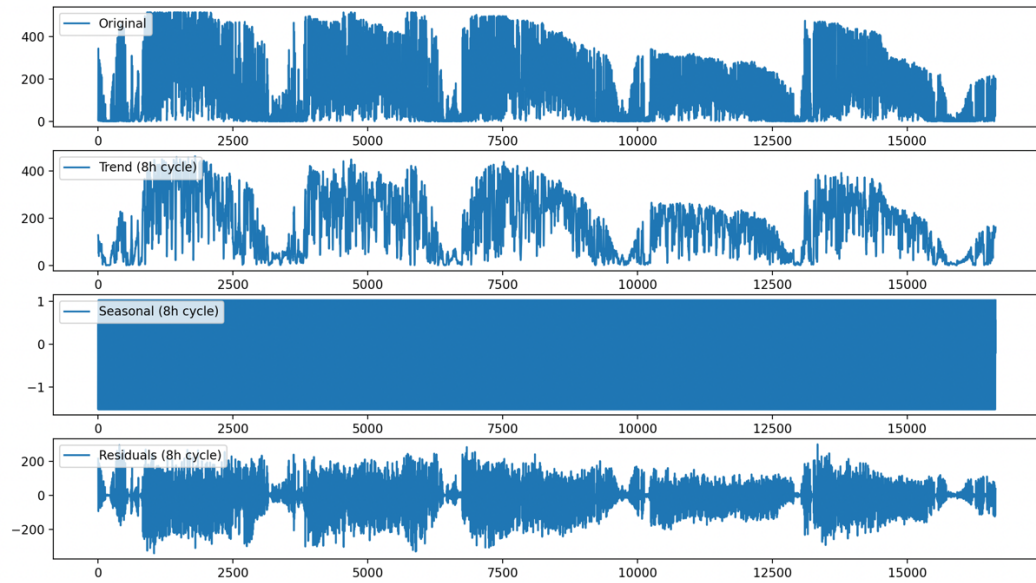
4. 성분 분해 (Decomposition)

4-1. loess and FFT 사용 직접 성분 분해

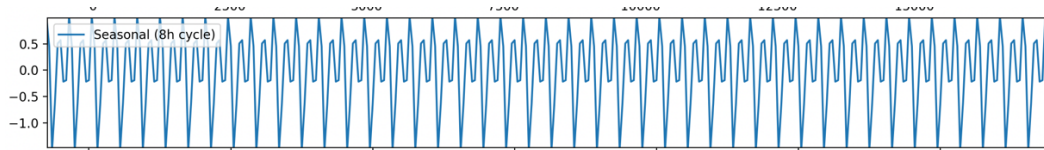
- 추세(Trend): Loess=0.5 이용
해가 지날 수록 태양에너지 측정량이 10kWh/year 감소.
- 계절성(Seasonality): FFT (Fast Fourier Transform) 함수를 이용하여 연간 주기를 갖는 데이터의 주파수 대역 선택
계절에 따라 변화하는 모습이 사인파와 유사.
- 나머지(Remainder): raw-data 에서 추세와 계절성 그래프를 뺀 나머지.



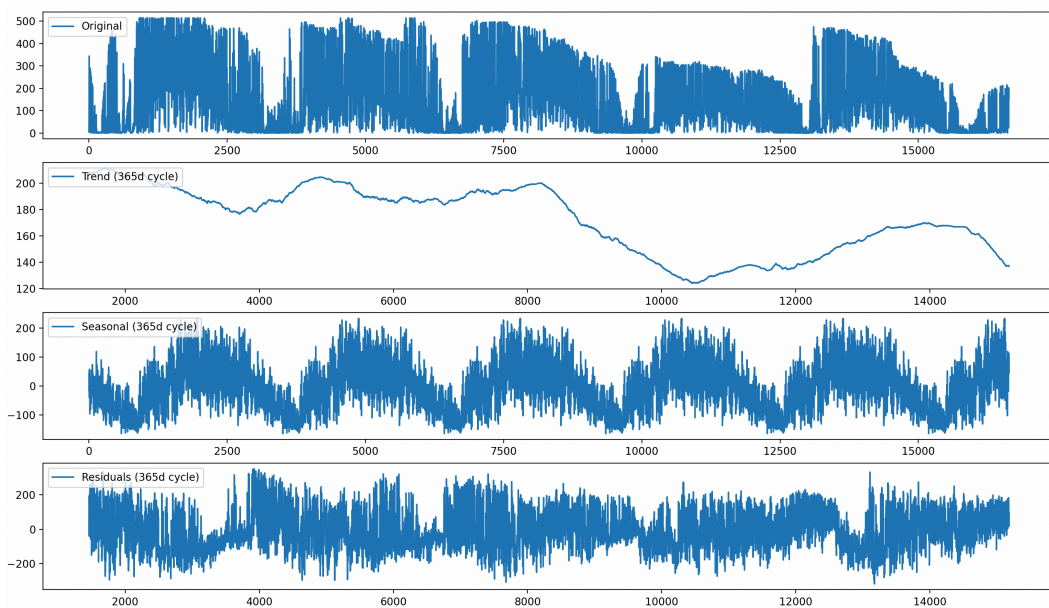
4-2. stldecompose 라이브러리 사용 성분 분해



<8h(하루 측정 시간 8 시~16 시) 주기>



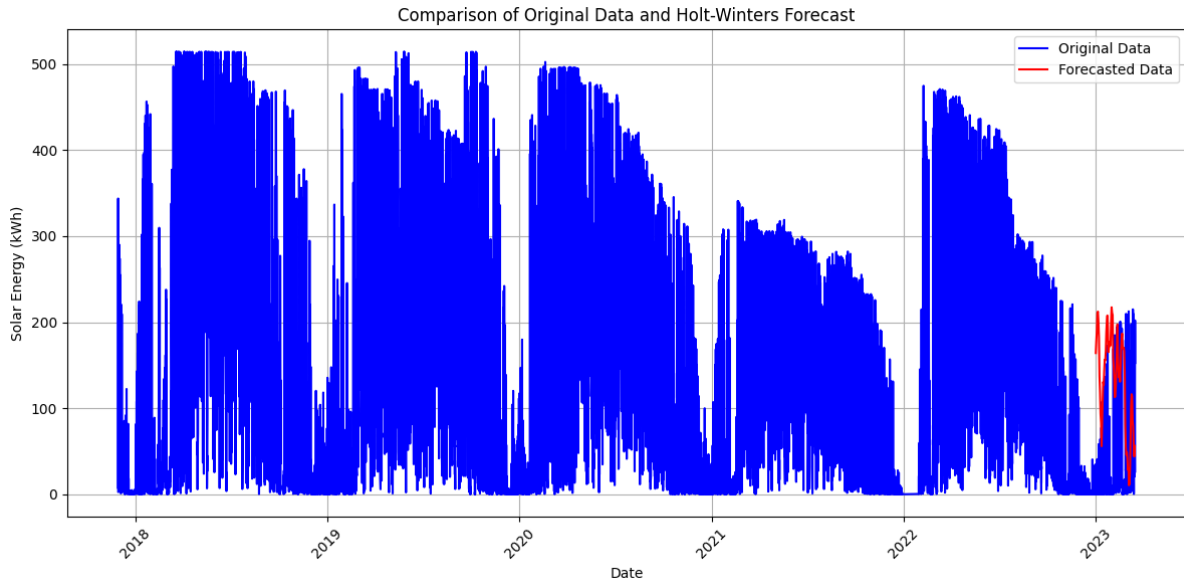
<seasonal (8h cycle) 확대 (2950~3250)>



<1 년 (8h*12month*30days) 주기>

- 데이터 예측 (Forecast)

- 4-1. 에서 사용한 decomposition 을 토대로 Holt-Winters 모델 사용하여 2023-01-01 ~ 2023-03-16 데이터 예측.



- 결론

본 연구는 태양 에너지 생산의 계절성 및 추세에 대한 분석을 수행하였다.

다음은 주요 결론이다:

1. 계절성의 존재: 데이터 분해를 통해 태양 에너지 생산량에는 명확한 계절적 변동이 존재함을 확인하였다. 이는 해가 움직이는 패턴과 유사한 주기성을 보였으며, FFT 를 통해 확인된 주기성은 사인파와 유사한 형태를 띠었다.
2. 추세의 변화: 태양 에너지 생산량의 추세 분석 결과, 연간 약 10kWh 씩 감소하는 추세를 확인하였다. 이러한 감소는 시간이 흐를수록 뚜렷해지며, 환경 변화나 태양 활동 등의 영향이 있을 수 있다.
3. 데이터 전처리의 중요성: 데이터 전처리는 분석의 정확성에 중대한 영향을 미친다. 특히 데이터의 누락 및 불일치를 처리하는 과정에서 신중함이 요구된다. 우리는 누락된 데이터를 확인하고 적절한 처리를 통해 신뢰할 수 있는 분석 결과를 얻을 수 있었다.
4. 시간별 측정 빈도: 데이터 측정 빈도는 시간에 따라 다를 수 있으며, 분석에 앞서 데이터의 일관성을 확인하는 것이 중요하다. 분석에 사용된 데이터는 8 시간 간격으로 측정된 것으로 확인되었으며, 이를 바탕으로 추세 및 계절성 분석을 진행하였다.
5. 라이브러리와 직접 분해 비교: 4-2. stldecompose 라이브러리를 사용하여 성분 분해한 그래프가 4-1. loess 와 FFT 를 사용하여 직접 성분 분해한 그래프와 유사하다.

이러한 결과들은 태양 에너지 생산의 변화를 이해하고, 향후 에너지 정책 및 시스템에 대한 결정에 도움을 줄 수 있다.

Remainder 에 해당하는 그래프와 값들은 계절성이나 트렌드를 따르지 않는 요소들로 오존층 두께, 태양광 패널 위의 이물질, 환경적 변화 등 여러 요인들을 추측하고 가설화 할 수 있다.