

# 실제 컨버터 출력 데이터를 이용한 특정 지역 태양광 장단기 발전 예측

## Prediction of Short and Long-term PV Power Generation in Specific Regions using Actual Converter Output Data

하은규<sup>1</sup> · 김태오<sup>2</sup> · 김창복<sup>1\*</sup>

<sup>1</sup>가천대학교 에너지 IT학과

<sup>2</sup>가천대학교 IT융합공학과

Eun-gyu Ha<sup>1</sup> · Tae-oh Kim<sup>2</sup> · Chang-bok Kim<sup>1\*</sup>

<sup>1</sup>Department of Energy IT, Gachon University, Gyeonggi-do 13120, Korea

<sup>2</sup>Department of IT Convergence Engineering, Gachon University, Gyeonggi-do 13120, Korea

### [요 약]

태양광 발전은 일사량만 있으면 전기에너지를 얻을 수 있기 때문에, 새로운 에너지 공급원으로 용도가 급증하고 있다. 본 논문은 실제 태양광 발전 시스템의 컨버터 출력을 이용하여 장단기 출력 예측을 하였다. 예측 알고리즘은 다중선형회귀와 머신러닝의 지도 학습 중 분류모델인 서포트 벡터 머신 그리고 DNN과 LSTM 등 딥러닝을 이용하였다. 또한 기상요소의 입출력 구조에 따라 3개의 모델을 이용하였다. 장기 예측은 월별, 계절별, 연도별 예측을 하였으며, 단기 예측은 7일간의 예측을 하였다. 결과로서 RMSE 측도에 의한 예측 오차로 비교해 본 결과 다중선형회귀와 SVM 보다는 딥러닝 네트워크가 예측 정확도 측면에서 더 우수하였다. 또한, DNN 보다 시계열 예측에 우수한 모델인 LSTM이 예측 정확도 측면에서 우수하였다. 입출력 구조에 따른 실험 결과는 모델 1보다 모델 2가 오차가 적었으며, 모델 2보다는 모델 3이 오차가 적었다.

### [Abstract]

Solar photovoltaic can provide electrical energy with only radiation, and its use is expanding rapidly as a new energy source. This study predicts the short and long-term PV power generation using actual converter output data of photovoltaic system. The prediction algorithm uses multiple linear regression, support vector machine (SVM), and deep learning such as deep neural network (DNN) and long short-term memory (LSTM). In addition, three models are used according to the input and output structure of the weather element. Long-term forecasts are made monthly, seasonally and annually, and short-term forecasts are made for 7 days. As a result, the deep learning network is better in prediction accuracy than multiple linear regression and SVM. In addition, LSTM, which is a better model for time series prediction than DNN, is somewhat superior in terms of prediction accuracy. The experiment results according to the input and output structure appear Model 2 has less error than Model 1, and Model 3 has less error than Model 2.

**Key word** : Photovoltaic, linear regression, Support vector machine, Deep neural network, Recurrent neural network.

<https://doi.org/10.12673/jant.2019.23.6.561>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 28 November 2019; Revised 30 November 2019

Accepted (Publication) 21 December 2019 (30 December 2019)

\*Corresponding Author; Chang-bok Kim

Tel: +82-32-446-0695

E-mail: cbkim@gachon.ac.kr

## I. 서론

화석연료는 한정된 자원과 환경오염 문제로 인해 사용을 줄이는 노력이 필요하지만 태양광 및 풍력 발전 등 신재생 에너지는 자원이 무한하고, 환경 친화적이다. 신재생 에너지는 무한한 자원으로 지속 가능한 미래 에너지 관점에서 에너지 공급 안정성을 증대할 수 있고, 환경 친화적이기 때문에 온실가스 및 미세먼지 등의 환경오염 물질을 감소시킬 수 있어 중요성이 날로 커지고 있다[1].

태양광 발전 (solar photovoltaic)은 일사량만 있으면 전기에너지를 얻을 수 있기 때문에, 새로운 에너지 공급원으로 용도가 급증하고 있다. 최근 태양광 발전은 에너지가 필요한 모든 사물에 다각적으로 사용되고 있으며, 비행기, 드론, 전기 자동차 등 이동 사물에 대해서도 화석 연료를 대체하고 있다.

최근 솔라 임펄스 2가 세계 일주 비행에 성공하면서 구글과 페이스북을 포함한 다양한 회사에서 태양광 비행기를 만들어 띄웠으며 나날이 기술이 발전하고 있다. 또한, Airbus의 태양광 비행 드론인 Zephyr S HAPS를 띄웠고 무려 25일 동안의 비행에 성공했다. 이와 같이 태양광 발전은 에너지가 필요한 모든 사물에 다각적으로 사용되고 있다.

그러나 태양광 발전에 필요한 태양광 패널은 특정 장소 및 시간에 따른 일사량, 기온 및 대기 중의 습도, 구름 등 외부 변화에 따라 전압과 전류가 바뀌어 발전 전력이 간헐적이고 예측하기 매우 어려운 불안정한 전력 공급 장치로서 기존의 발전방식들과 비교해서 안정적인 전력을 출력하기가 기술적으로 쉽지 않다. 또한, 태양광 발전은 먼지, 구름, 눈과 같은 외부적 요인의 영향으로 변화가 매우 심하고, 계절적 요인에 의해서도 많은 영향을 받는다. 이러한 태양광 발전 출력전력의 불확실성을 개선하고, 전력계통의 신뢰성을 향상시키며, 전력품질을 향상하기 위해서 태양광 발전 출력전력을 정확히 예측할 필요가 있다[2].

태양광 발전은 중단기 예측, 장기 예측이 있다. 단기 예측은 발전소의 효율적인 운영과 전력계통과의 안정적인 연계에 이용할 수 있다. 장기 예측은 태양광 발전 시스템 설립 시 초기 투자비용이 많이 발생하기 때문에, 발전량 예측을 통한 투자계획을 위해 이용된다[2]-[6].

태양광 예측은 물리적, 통계적, 확률론적 기법 등이 있으며, 예측 알고리즘은 웨이브릿(wavelet), ARIMA (autoregressive integrated moving average), ELM (extreme learning machine), SVR(support vector regression), neuro fuzzy, 인공신경망 (artificial neural network) 등에 대한 연구가 있다[7]-[11].

본 논문은 경기도 연성의 2015년 1월부터 2018년 10월 까지 실제 태양광 발전 시스템의 컨버터 출력을 이용하여 장단기 출력 예측을 하였다. 예측 알고리즘은 여러 개의 독립변수로 결과값을 예측하는 다중선형회귀(multiple linear regression)와 머신러닝의 지도학습 중 분류모델인 SVM(support vector machine), DNN(deep neural network)과 LSTM(long short-term memory)을 이용하였다. 또한 기상요소의 입출력 구조에 따라 3개의 모델

을 이용하였다. 장기 예측은 월별, 계절별, 연도별 예측을 하였으며, 단기 예측은 7일간의 예측을 하였다. 본 논문은 이러한 장단기 예측을 위하여 다양한 데이터 구조를 이용하여 6가지 실험을 통해 효율적인 발전량 예측을 위한 모델을 제안하였다. 본 논문은 2장에 기상요소와 컨버터 발전량의 상관관계와 예측 알고리즘에 대해서 서술하였으며, 3장에 제안 모델과 제안 데이터 구조에 대해서 서술하였다. 또한, 4장에 실험 결과와 검토를 하였으며, 5장에 결론을 맺었다.

## II. 관련 연구

### 2-1 기상요소와 컨버터 발전량

태양광 패널은 복사 에너지에 의해 전력을 생산하기 때문에, 일사량과 일조량이 매우 중요한 요소이며, 반도체 성분이기 때문에 적당한 온도와 습도를 유지해야 한다. 태양광 발전은 기상 상태에 따라서 그 출력이 매우 간헐적이며, 대부분의 기상요소는 태양광 발전에 많은 영향을 주는 요소이다. 따라서 태양광 발전의 출력에 대한 불확실성의 영향을 줄이고 태양광 발전의 경제성을 향상시키기 위하여 기상요소를 이용한 태양광 출력 예측기술이 필수적이다. 표 1에 기상요소와 컨버터 출력 간 상관관계를 나타냈다. 기상요소와 컨버터 출력 데이터는 2015년에서 2018년까지의 기상 데이터와 실제 태양광 발전소의 컨버터 출력 데이터이다.

표에서 태양광 패널의 전력 생산에 영향을 미치는 패널의 표면온도, 주위온도, 일사량, 일조량에서 높은 상관관계가 있음을 알 수 있다. 또한, 강수량은 매우 낮은 상관관계를 보였으며, 이는 강수량이 많을 때 컨버터 출력 예측이 어렵다는 것을 의미한다. 표 2에 기상요소와 컨버터 출력의 연도별 평균값을 나타냈다. 컨버터 출력의 연도별 평균값은 높은 상관관계를 보였던 표면온도, 일사량, 일조량의 평균값에 따라 결정됨을 확인하였다.

표 1. 기상요소와 컨버터 출력과의 상관관계

Table 1. Correlation between meteorological factors and converter output.

Weather elements	2015	2016	2017	2018
Surface temp.	0.661	0.611	0.612	0.612
Ambient temp.	0.3889	0.351	0.325	0.329
Precipitation	-0.059	-0.057	-0.045	-0.069
Wind speed	0.072	0.081	0.132	0.118
Wind direction	0.146	0.1467	0.1388	0.150
Humidity	-0.358	-0.284	-0.389	-0.394
Cloud	-0.041	-0.014	-0.011	-0.068
Radiation	0.958	0.919	0.928	0.936
Sunshine	0.768	0.744	0.767	0.752

**표 2.** 기상요소와 컨버터 출력의 연도별 평균값  
**Table 2.** Yearly average of meteorological factors and converter output.

Weather elements	2015	2016	2017
Surface temp.	18.584	18.433	17.679
Ambient temp.	16.359	16.222	15.622
Humidity	78.720	75.559	65.514
Cloud	3.8463	4.1205	3.6654
Radiation	96.330	60.406	57.183
Sunshine	0.3050	0.2962	0.3154
Power generation	231.47	213.02	237.08

태양광 패널은 반도체 소자이기 때문에 가급적 낮은 온도에서 많은 전력량을 가진다. 또한, 패널의 온도가 25℃가 넘어가면 1℃ 당 0.5%의 효율이 낮아진다. 패널의 표면온도는 주위온도와 일사량에 많은 영향을 받는다. 즉, 주위온도가 높고 일사량이 많으면 표면온도는 높아지게 된다, 따라서 여름이 일조량을 많이 받기는 하지만 주변 온도의 영향으로 인해 발전량이 감소된다. 표에서 2017년도는 낮은 표면온도와 주위온도를 가지며, 일사량은 낮지만 일조량으로 인해 높은 발전량을 가짐을 알 수 있다. 따라서 이러한 기상조건의 차이 때문에 연도별 컨버터 발전량 차이가 있음을 알 수 있다.

**2-2 예측 알고리즘**

다중선형회귀는 단순선형회귀에서 확장하여 설명변수가 다수일 때 사용하는 모형이다.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_q x_{iq} + \epsilon_i \tag{1}$$

여기서  $y_i$ 와  $i = 1, n$  는 종속변수의 관측 값을,  $x_{i1}, x_{i2}, \dots, x_{iq}, i = 1, n$  는  $q$ 개의 설명변수의 관측 값,  $n$ 은 관측 값의 개수를 나타낸다.

$\beta_0 \dots \beta_q$ 는 회귀계수이고  $\epsilon$ 는 잔차 또는 잡음으로써 회귀계수에 의해 설명되지 않는 항이다. 단순회귀모형에 독립변수만을 추가하여 확장된 형태의 모형으로 가정이나 가설, 검증에 있어서 근본적인 차이는 없다.

SVM은 분류와 예측에 응용할 수 있는 지도학습 모델이다. 학습 데이터가 주어졌을 때, 모든 학습 데이터의 목표변수가 주어진 편차 안에 존재하도록 하는 가장 평평한 회귀함수  $f(x)$ 를 찾는 것이다. 회귀함수  $f(x)$ 는 식(2)와 같고 손실함수는 식(3)으로 나타낼 수 있다.

$$f(x) = \langle w, x \rangle + b \tag{2}$$

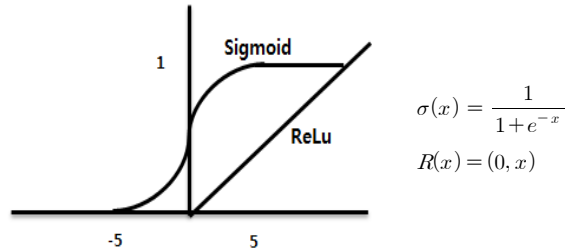
$$R_{reg} = \frac{1}{2} \|w\|^2 + CR_{emp}[f] \tag{3}$$

$w$ 는 가중치 벡터이고  $b$ 는 회귀매개변수들의 임계값이다.

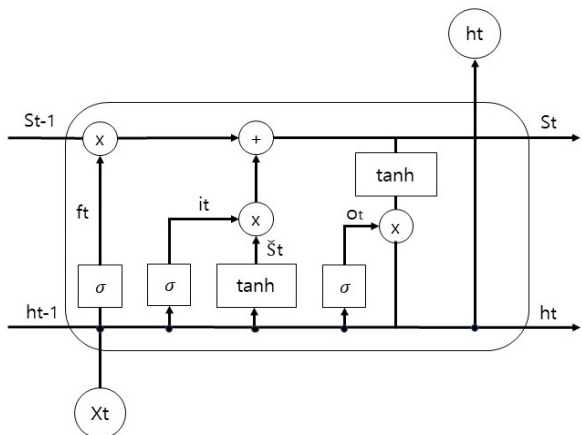
$\frac{1}{2} \|w\|^2$ 는 함수의 평평한 정도를 의미하고  $C$ 는 0보다 큰 상수로 패널티를 의미한다.  $C$ 값을 크게 설정하면, 편차에서 벗어나는 예측 값에 대한 패널티를 높게 주게 된다. SVM은 이 손실함수를 이용해 최적화할 수 있다[12].

DNN은 중간층을 깊게 할 경우 기울기 사라짐의 문제점을 가진 시그모이드(sigmoid) 함수 대신 ReLu(rectified linear Unit) 및 tanh 활성화 함수를 사용하여, 심층에 의한 기울기 사라짐 문제점을 해결하고, 가중치를 학습에 용이하게 초기화하는 RBM(restricted boltzmann machine), Xavier 등의 알고리즘을 통해, 가중치 초기화 기법을 이용해 가중치 초기화 문제를 해결하여 중간층을 깊게 한 심층 신경망이다. 그림 1에 시그모이드와 ReLu 활성화 함수에 대해서 나타냈다.

RNN(recurrent neural network)은 하나의 입출력 패턴을 가진 DNN 병렬 체인 구조로 연결한 형태로 과거 학습 결과를 현재 학습에 사용하며, 시계열 데이터를 처리하는데 효과적이다. RNN은 스탬프(stamp)가 길어지면 이전 스탬프의 학습 값이 사라지는 문제가 있다. LSTM은 전체 체인을 관통하는 셀 스테이트(cell state)를 통해 과거 셀의 학습 결과를 큰 변함없이 전달하는 구조로 이전 스탬프의 학습 값이 사라지는 문제를 해결한다. 그림 2에 LSTM의 구조에 대해서 나타냈다.



**그림 1.** 활성화 함수  
**Fig. 1.** Activation function.



**그림 2.** LSTM 구조  
**Fig. 2.** LSTM Structure.

LSTM은 첫 번째 단계로 시그모이드 함수를 이용해 삭제할 정보를 결정하고, 두 번째 단계로 또 다른 시그모이드 함수와 tanh 함수를 이용해 새로운 정보가 셀 스테이트에 저장될지 결정한다. 세 번째 단계는 셀 스테이트를 업데이트하고 마지막으로 마지막 시그모이드 함수와 셀 스테이트에서 나온 출력을 통과시킨 마지막 tanh 함수로 어떤 출력 값을 출력할지를 결정한다[13].

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{4}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{5}$$

$$\tilde{S}_t = \tanh(W_S \cdot [h_{t-1}, x_t] + b_S) \tag{6}$$

$$S_t = f_t S_{t-1} + i_t \tilde{S}_t \tag{7}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{8}$$

$$h_t = o_t \tanh(S_t) \tag{9}$$

여기서  $f_t$ 는 첫 번째 단계,  $i_t$ 와  $\tilde{S}_t$ 는 두 번째 단계,  $S_t$ 는 세 번째 단계,  $o_t$ ,  $h_t$ 는 마지막 단계의 출력이다. W는 가중치, b는 바이어스, S는 셀 스테이트, t는 타임스텝을 의미한다.

### III. 예측 알고리즘 비교

본 논문은 경기도 연성의 2015년 1월부터 2018년 10월 까지 실제 태양광 발전 시스템의 컨버터 출력을 이용하여 다양한 데이터 구조로 장단기 출력 예측을 하였다. 장기 예측은 월별, 계절별, 연도별 예측을 하였으며, 단기 예측은 2018년 5월을 기준으로 7일간의 예측을 하였다. 예측 알고리즘은 지도 학습 알고리즘이며, 여러 개의 독립변수로 실제 값을 예측하는 다중선형 회귀와 SVM 그리고 DNN과 LSTM을 이용하였다.

본 논문은 독립변수로서 학습의 입력 요소로 태양광 패널의 온도, 주위온도, 강수량, 풍속, 풍향, 습도, 전운량, 일사량, 일조량 등 태양광 발전량과 상관관계가 있는 모든 기상요소를 이용하였다. 또한, 주위온도, 강수량, 풍속, 풍향, 습도, 전운량 등 일사량과 일조량을 제외한 기상요소는 기상청에서 예보하기 때문에 예보된 데이터를 이용한 데이터 구조로 학습하였다. 표 3과 4에 장단기 예측에 대한 학습 및 테스트 데이터에 대해서 나타났다. 또한, 다양한 방법의 컨버터 발전량 예측 성능 평가를 위해 다음과 같은 데이터 구조를 제안하였다.

1. 가장 이상적인 예측 결과를 나타내는 데이터 구조로서 일사량과 일조량도 예보가 된다는 가정으로 모든 시간대가 동일한 실제 기상 데이터 및 컨버터 발전량 예측 데이터 구조.

2. 일사량과 일조량은 예보가 안 되기 때문에 이전시간 대의 일사량과 일조량을 이용한 컨버터 발전량 예측 데이터 구조

3. 일사량 및 일조량은 예보를 안 하기 때문에 주위온도, 강수량, 풍속, 풍향, 습도, 전운량 등 기본적인 기상요소와 패널의 표면온도를 이용한 일사량 및 일조량 예측 데이터 구조

4. 학습 데이터는 실제 모든 기상 데이터 및 컨버터 발전량

데이터를 이용하고, 테스트 데이터는 예측된 일사량 및 일조량을 이용한 컨버터 발전량 예측 데이터 구조

5. 데이터 구조 2에서 상관관계가 낮은 기상요소(강수량)를 제외한 컨버터 발전량 예측 데이터 구조.

단기 예측은 학습 데이터는 이전 실제 데이터를 사용하고 테스트 데이터는 추후 7일간의 일사량 및 일조량 예측 데이터를 이용하여 컨버터 발전량을 예측하였다.

표 5에서 8까지 실험 1에서 4까지의 발전량 예측 데이터 구조에 대해 나타났다. 여기서 입력변수는 주위온도(A), 강수량(B), 풍속(C), 풍향(D), 습도(E), 전운량(F), 표면온도(G), 일사량(H), 일조량(I)이며 출력변수는 컨버터 출력(J)이다. 학습 데이터 구조는 세 가지 모델로 구성하였다. 여기서 숫자는 시간대를 나타내며, 0은 현재 시간을 의미하고 -1은 한 시간 전 -2는 두 시간 전을 의미한다. 모델 1은 현재 시간의 기상요소만을 사용하였고, 모델 2는 현재 시간과 한 시간 전 기상요소를 사용하였으며, 모델 3은 현재 시간과 한 시간 전 그리고 두 시간 전 기상요소를 사용하였다.

표 3. 장기 학습 및 테스트 데이터

Table 3. Long term training and test data.

	Learning data	Test data
Monthly	2015/1 ~ 2018/9	2018/10
Seasonal	2015/1 ~ 2018/ 5	2018/6 ~ 2018/8
Annually	2015/1 ~ 2017/12	2018/1 ~ 2018/10

표 4. 단기 학습 및 테스트 데이터

Table 4. Short term training and test data.

Learning data	Test data
2015/1/1 ~ 2018/4/30	2018/5/1 ~ 2018/5/7
2015/1/1 ~ 2018/4/30	2018/5/11 ~ 2018/5/17
2015/1/1 ~ 2018/4/30	2018/5/21 ~ 2018/5/27
2015/1/1 ~ 2018/4/30	2018/5/25 ~ 2018/5/31

표 5. 학습 데이터 구조 1

Table 5. Training data structure 1.

	input data										Y	
	A-F	G2	G1	G0	H2	H1	H0	I2	I1	I0	J	J
M1	0	/	/	0	/	/	0	/	/	0	0	0
M2	0	/	-1	0	/	-1	0	/	-1	0	0	0
M3	0	-2	-1	0	-2	-1	0	-2	-1	0	0	0

표 6. 학습 데이터 구조 2

Table 6. Training data structure 2.

	input data										Y	
	A-F	G2	G1	G0	H2	H1	H0	I2	I1	I0	J	J
M1	0	/	/	-1	/	/	-1	/	/	-1	0	0
M2	0	/	-2	-1	/	-2	-1	/	-2	-1	0	0
M3	0	-3	-2	-1	-3	-2	-1	-3	-2	-1	0	0

표 7. 학습 데이터 구조 3

Table 7. Training data structure 3.

	input data				label
	A-F	G2	G1	G0	H, I
M11	0	-3	-2	-1	0

표 8. 테스트 데이터 구조 4

Table 8. Testing data structure 4.

	input data										Y
	A-F	$\hat{G}_2$	$\hat{G}_1$	$\hat{G}_0$	$\hat{H}_2$	$\hat{H}_1$	$\hat{H}_0$	$\hat{I}_2$	$\hat{I}_1$	$\hat{I}_0$	J
M1	0	/	/	-1	/	/	0	/	/	0	0
M2	0	/	-2	-1	/	-1	0	/	-1	0	0
M3	0	-3	-2	-1	-2	-1	0	-2	-1	0	0

표 9. 하이퍼 파라미터

Table 9. Hyper parameters.

Parameter	Configuration	
	DNN	LSTM
Number of input nodes	Number of input variables in data structure	
Number of hidden layer nodes	32	32
Number of hidden layer	2	1
Learning rate	0.008	0.01
Number of iterations	15000	1000
Time stamp		24

본 논문에선 인공지능은 하이퍼 파라미터에 따라 예측 결과가 다르기 때문에 파라미터를 수정해가면서 실험한 결과 가장 예측 결과가 좋은 파라미터를 사용하였다. 표 9에 본 논문 실험에서 사용한 하이퍼 파라미터에 대해서 나타냈다.

DNN의 학습률은 0.008로 하였으며, 학습 횟수는 15000번으로 하였다. 기상요소가 입력되는 입력층은 각 데이터 구조의 입력변수 수만큼의 노드를 사용하였고, 중간층은 각 층마다 32 노드로 하여 2층으로 하였으며, 출력층은 1 노드이다. 활성화 함수는 ReLu 함수를 사용하였고, 학습률과 학습속도를 높이기 위해 Xavier 초기화 함수로 가중치를 초기화 하였다. 또한, 최적화 알고리즘으로 AdamOptimizer를 사용하였다.

LSTM은 학습률을 0.01로 하였으며, 입력층은 입력변수의 수만큼의 노드를 사용하였고, 중간층은 32 노드로 하여 1층으로 하였으며, 출력층은 1 노드이다. 또한 학습 셀의 개수는 24개로 하였으며, 학습 횟수는 1000번으로 하였다.

#### IV. 실험 결과 및 검토

본 논문의 실험환경은 Window 10 기반으로 파이썬(python) 3.6기반 텐서플로우(tensorflow) GPU Version을 이용하였다. 일사량 및 일조량 예측을 위한 전체 데이터는 2015년에서 2018년

까지의 기상 데이터로서, 35040개의 레코드로 구성되어 있다. 또한, 장단기 예측을 위해 기간별로 학습 데이터와 테스트 데이터를 나누어 실험하였다.

예측 정확도와 예측 결과 평가를 위해서 다음과 같이 평균 제곱근오차(root mean square error), 평균절대오차(mean absolute error)를 사용하였다. 만약 RMSE >> MAE라면 예측된 값은 실제 값과 높은 편차를 가지며, 비슷한 값을 가지면 예측된 값과 실제 값이 낮은 편차를 가진다는 것을 의미한다.

$$RMSE = \frac{1}{N} \sqrt{\sum_{n=1}^N (x'_n - x_n)^2} \quad (10)$$

$$MAE = \frac{1}{N} \sum_{n=1}^N |x'_n - x_n| \quad (11)$$

표 10부터 16까지 데이터 구조에 대한 실험 결과를 나타냈다. 모든 실험 결과는 5번 반복하여 얻은 결과 값의 평균이다.

표 10. 데이터 구조 1 결과

Table 10. Data structure 1 result.

		model1	model2	model3
Monthly	Regression	0.074	0.072	0.072
	SVM	0.068	0.064	0.062
	DNN	0.059	0.055	0.053
	LSTM	0.057	0.061	0.057
Seasonal	Regression	0.079	0.080	0.080
	SVM	0.078	0.074	0.073
	DNN	0.069	0.066	0.064
	LSTM	0.060	0.065	0.058
Annually	Regression	0.080	0.079	0.079
	SVM	0.075	0.072	0.070
	DNN	0.065	0.063	0.059
	LSTM	0.061	0.061	0.058

표 11. 데이터 구조 2 결과

Table 11. Data structure 2 result.

		model1	model2	model3
Monthly	Regression	0.132	0.117	0.109
	SVM	0.125	0.109	0.103
	DNN	0.114	0.101	0.099
	LSTM	0.086	0.088	0.086
Seasonal	Regression	0.138	0.127	0.116
	SVM	0.131	0.120	0.108
	DNN	0.117	0.103	0.093
	LSTM	0.086	0.081	0.085
Annually	Regression	0.141	0.127	0.116
	SVM	0.136	0.119	0.107
	DNN	0.120	0.099	0.094
	LSTM	0.086	0.082	0.081

표 12. 데이터 구조 3 결과

Table 12. Data structure 3 result.

		일사량	일조량
Monthly	Regression	0.085	0.273
	SVM	0.082	0.273
	DNN	0.074	0.150
	LSTM	0.058	0.129
Seasonal	Regression	0.087	0.284
	SVM	0.082	0.270
	DNN	0.075	0.179
	LSTM	0.061	0.157
Annually	Regression	0.082	0.280
	SVM	0.078	0.268
	DNN	0.074	0.166
	LSTM	0.063	0.149

표 13. 데이터 구조 4 결과

Table 13. Data structure 4 result.

		model1	model2	model3
Monthly	Regression	0.115	0.109	0.105
	SVM	0.115	0.112	0.106
	DNN	0.108	0.105	0.102
	LSTM	0.106	0.103	0.100
Seasonal	Regression	0.115	0.114	0.108
	SVM	0.118	0.121	0.111
	DNN	0.113	0.105	0.101
	LSTM	0.104	0.103	0.096
Annually	Regression	0.115	0.112	0.107
	SVM	0.119	0.114	0.109
	DNN	0.106	0.102	0.097
	LSTM	0.101	0.094	0.093

표 14. 데이터 구조 5 결과

Table 14. Data structure 5 result.

		model1	model2	model3
Monthly	Regression	0.132	0.117	0.109
	SVM	0.125	0.110	0.104
	DNN	0.115	0.101	0.100
	LSTM	0.087	0.086	0.084
Seasonal	Regression	0.138	0.127	0.116
	SVM	0.132	0.120	0.108
	DNN	0.117	0.102	0.094
	LSTM	0.089	0.087	0.080
Annually	Regression	0.141	0.127	0.116
	SVM	0.136	0.118	0.107
	DNN	0.118	0.100	0.094
	LSTM	0.086	0.081	0.077

표 15. 단기 일사량 및 일조량 예측

Table 15. Short term radiation and sunshine forecast.

		일사량	일조량
2018/5/1 -	Regression	0.083	0.264
	SVM	0.077	0.268
	DNN	0.073	0.176
2018/5/7	LSTM	0.062	0.164
	Regression	0.072	0.200
2018/5/11 -	SVM	0.066	0.186
	DNN	0.060	0.142
	LSTM	0.057	0.134
2018/5/21 -	Regression	0.075	0.319
	SVM	0.069	0.283
	DNN	0.064	0.193
	LSTM	0.066	0.156
2018/5/25 -	Regression	0.077	0.322
	SVM	0.071	0.295
	DNN	0.070	0.182
2018/5/31	LSTM	0.061	0.166

표 16. 단기 예측 결과

Table 16. Short term forecast result.

		model1	model2	model3
2018/5/1 -	Regression	0.107	0.107	0.105
	SVM	0.113	0.103	0.099
	DNN	0.112	0.094	0.092
2018/5/7	LSTM	0.090	0.083	0.079
	Regression	0.084	0.083	0.079
2018/5/11 -	SVM	0.094	0.093	0.086
	DNN	0.078	0.073	0.070
	LSTM	0.068	0.067	0.061
2018/5/21 -	Regression	0.109	0.100	0.100
	SVM	0.106	0.095	0.093
	DNN	0.090	0.086	0.079
	LSTM	0.113	0.112	0.111
2018/5/25 -	Regression	0.099	0.100	0.099
	SVM	0.105	0.104	0.101
	DNN	0.099	0.095	0.093
2018/5/31	LSTM	0.094	0.092	0.087

데이터 구조 1 결과에서 월별 데이터 모델 1은 LSTM이 0.057, 모델 2는 DNN이 0.055, 모델 3은 DNN이 0.053으로 가장 우수한 성능을 보였다. 계절별 데이터 모델 1은 LSTM이 0.060, 모델 2는 LSTM이 0.065, 모델 3은 LSTM이 0.058로 가장 우수한 성능을 보였다. 연도별 데이터 모델 1은 LSTM이 0.061, 모델 2는 LSTM이 0.061, 모델 3은 LSTM이 0.058로 가장 우수한 성능을 보였다.

데이터 구조 2 결과에서 월별 데이터 모델 1은 LSTM이 0.086, 모델 2는 LSTM이 0.088, 모델 3은 LSTM이 0.086으로가

장 우수한 성능을 보였다. 계절별 데이터 모델 1은 LSTM이 0.086, 모델 2는 LSTM이 0.081, 모델 3은 LSTM이 0.085로 가장 우수한 성능을 보였다. 연도별 데이터 모델 1은 LSTM이 0.086, 모델 2는 LSTM이 0.082, 모델 3은 LSTM이 0.081로 가장 우수한 성능을 보였다.

데이터 구조 3 결과에서 월별 데이터 일사량 모델은 LSTM이 0.058, 일조량 모델은 LSTM이 0.129로 가장 우수한 성능을 보였다. 계절별 데이터 일사량 모델은 LSTM이 0.061, 일조량 모델은 LSTM이 0.157로 가장 우수한 성능을 보였다. 연도별 데이터 일사량 모델은 LSTM이 0.063, 일조량 모델은 LSTM이 0.149로 가장 우수한 성능을 보였다.

데이터 구조 4 결과에서 월별 데이터 모델 1은 LSTM이 0.106, 모델 2는 LSTM이 0.103, 모델 3은 LSTM이 0.100으로 가장 우수한 성능을 보였다. 계절별 데이터 모델 1은 LSTM이 0.104, 모델 2는 LSTM이 0.103, 모델 3은 LSTM이 0.096으로 가장 우수한 성능을 보였다. 연도별 데이터 모델 1은 LSTM이 0.101, 모델 2는 LSTM이 0.094, 모델 3은 LSTM이 0.093으로 가장 우수한 성능을 보였다.

데이터 구조 5 결과에서 월별 데이터 모델 1은 LSTM이 0.087, 모델 2는 LSTM이 0.086, 모델 3은 LSTM이 0.084로 가장 우수한 성능을 보였다. 계절별 데이터 모델 1은 LSTM이 0.089, 모델 2는 LSTM이 0.087, 모델 3은 LSTM이 0.080으로 가장 우

수한 성능을 보였다. 연도별 데이터 모델 1은 LSTM이 0.086, 모델 2는 LSTM이 0.081, 모델 3은 LSTM이 0.077로 가장 우수한 성능을 보였다.

단기 예측을 위한 일사량 및 일조량 예측 결과에서 2018/5/1-2018/5/7 데이터 일사량 모델은 LSTM이 0.062, 일조량 모델은 LSTM이 0.164로 가장 우수한 성능을 보였다. 2018/5/11-2018/5/17 데이터 일사량 모델은 LSTM이 0.057, 일조량 모델은 LSTM이 0.134로 가장 우수한 성능을 보였다. 2018/5/21-2018/5/27 데이터 일사량 모델은 DNN이 0.064, 일조량 모델은 LSTM이 0.156으로 가장 우수한 성능을 보였다. 2018/5/25-2018/5/31 데이터 일사량 모델은 LSTM이 0.061, 일조량 모델은 LSTM이 0.166으로 가장 우수한 성능을 보였다.

데이터 구조 6 결과는 예측된 일사량과 일조량을 이용한 예측 결과로서 2018/5/1-2018/5/7 데이터 모델 1은 LSTM이 0.090, 모델 2는 LSTM이 0.083, 모델 3은 LSTM이 0.079로 가장 우수한 성능을 보였다. 2018/5/11-2018/5/17 데이터 모델 1은 LSTM이 0.068, 모델 2는 LSTM이 0.067, 모델 3은 LSTM이 0.061로 가장 우수한 성능을 보였다. 2018/5/21-2018/5/27 데이터 모델 1은 DNN이 0.090, 모델 2는 DNN이 0.086, 모델 3은 DNN이 0.079로 가장 우수한 성능을 보였다. 2018/5/25-2018/5/31 데이터 모델 1은 LSTM이 0.094, 모델 2는 LSTM이 0.092, 모델 3은 LSTM이 0.087로 가장 우수한 성능을 보였다.

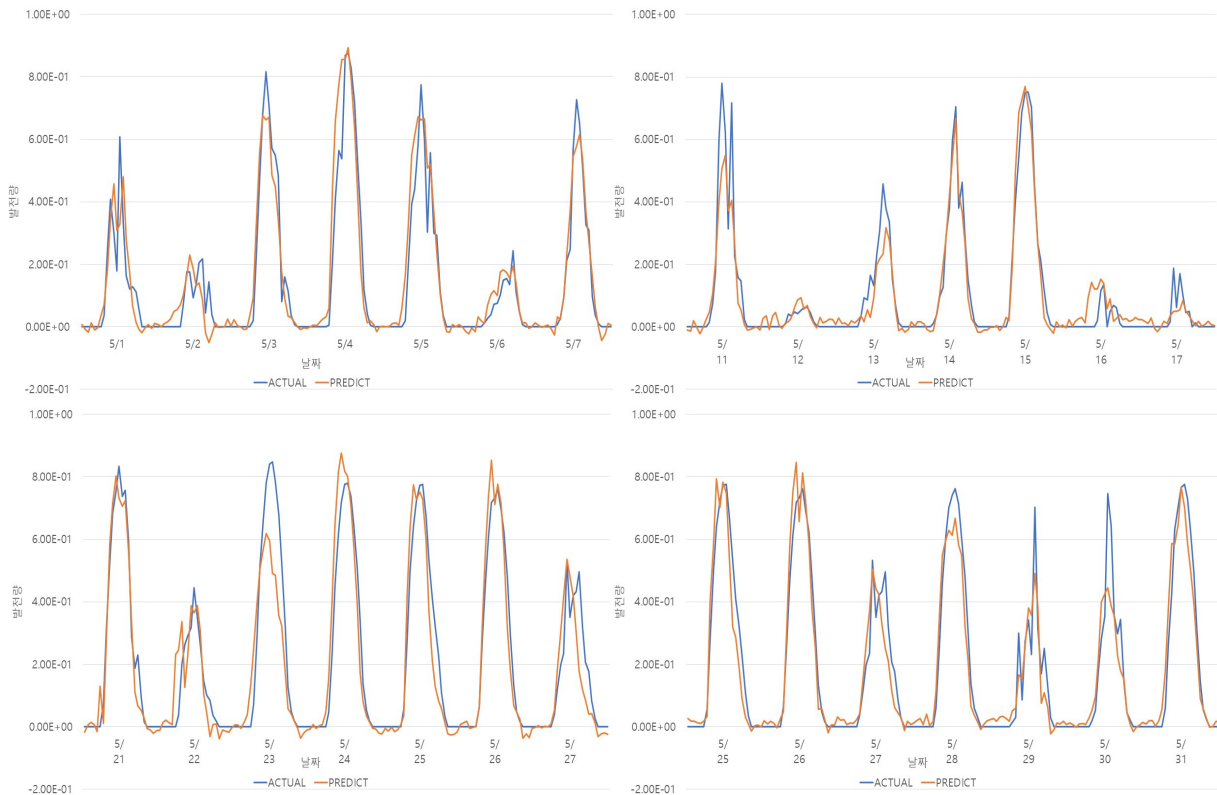


그림 3. 발전량 단기 예측 결과  
Fig. 3. A forecasting result of short-term power generation.

그림 3에 발전량 단기 예측 결과를 그래프로 나타냈다.

이상과 같은 실험 결과로 다중선형회귀와 SVM 그리고 DNN과 LSTM 등과 기상요소의 입출력 구조에 따른 3개의 모델의 장단기 예측 결과를 정리하면 다음과 같다.

(1) 일사량과 일조량 그리고 태양광 패널의 온도를 모두 예보한다는 조건인 데이터 구조 1은 RMSE 측도에 의한 예측 오차가 가장 좋았으며, 예측이 안 되는 일사량과 일조량의 데이터 시간대를 이동한 데이터 구조 2는 RMSE 측도에 의한 예측 오차를 비교해 본 결과 평균 0.02 정도 높았다.

(2) 기상예보가 안 되는 일사량 및 일조량을 예측하여, 학습 데이터는 실제 모든 기상 데이터 및 컨버터 발전량 데이터를 이용하고, 테스트 데이터는 예측된 일사량 및 일조량을 이용한 컨버터 발전량 예측 데이터 구조 4는 RMSE 측도에 의한 예측 오차를 비교해 본 결과 데이터 구조 2보다 평균 0.02 정도 높았다.

(3) 상관관계가 낮은 기상요소인 강수량을 제외한 컨버터 발전량 예측 데이터 구조 5는 RMSE 측도에 의한 예측 오차를 비교해 본 결과 데이터 구조 4보다 평균 0.02 정도 낮았다. 이것은 강수량은 예측에 많은 영향을 끼치지 않는다는 것을 의미한다.

(4) 단기 예측은 데이터 구조 2와 RMSE 측도에 의한 예측 오차가 비슷하였다. 그러나 주일마다 오차 차이가 많았으며, 특히 비가 오는 날, 즉 강수량이 있는 날의 오차가 컸다. 이것은 강수량이 있는 날은 일사량이 적기 때문임을 알 수 있다.

## V. 결론 및 향후과제

본 논문은 2015년 1월 1일 00시부터 2018년 10월 31일까지 3년 10개월의 데이터를 이용하여, 기상요소의 입출력 구조에 따른 3가지 모델로 다중선형회귀와 SVM 그리고 DNN과 LSTM을 이용하여 장단기 출력 예측을 하였다. 또한, 각 알고리즘과 모델에 대한 예측 정확도를 비교하였다. 실험 결과 태양광 발전량의 예측 오차를 RMSE 측도에 의한 예측 오차로 비교해 본 결과 다중선형회귀와 SVM 보다는 딥러닝 네트워크가 예측 정확도가 우수하였다. 그리고 DNN 보다 LSTM 모형이 예측 정확도 측면에서 더 우수하였다. 이것은 본 실험의 데이터 구조가 시계열이기 때문인 것으로 생각된다. 또한, 모델 1보다 모델 2가 오차가 적었으며, 모델 2보다는 모델 3이 오차가 적었다. 이에 입력 데이터 구조에 따라 예측 오차를 줄일 수 있음을 알 수 있다.

## Acknowledgements

본 논문은 산업통상자원부(MOTIE)와 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구 과제입니다. (No. 20194030202290).

## References

- [1] M. A. Islam, M. Hasanuzzaman, N. A. Rahim, A. Nahar, and M. Hosenuzzaman, "Global renewable energy-based electricity generation and smart grid system for energy security," *The Scientific World Journal*, Vol. 2014, pp. 1-13, Aug. 2014.
- [2] H. Yang, C. Huang, Y. Huang, and Y. Pai, "A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output," *IEEE Transactions on Sustainable Energy*, Vol. 5, No. 3, pp. 917-926, Apr. 2014.
- [3] M. Lei, L. Shiyang, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power," *Renewable and Sustainable Energy Reviews*, Vol. 13, No. 4, pp. 915-920, May. 2009.
- [4] A. Costa, A. Crespo, J. Navarro, G. Lizcano, H. Madsen, and E. Feitosa, "A review on the young history of the wind power short-term prediction," *Renewable and Sustainable Energy Reviews*, Vol. 12, No. 6, pp. 1725-1744, Aug. 2008.
- [5] S. Mirasgedis, Y. Sarafidis, E. Georgopoulou, D. Lalas, M. Moschovits, F. Karagiannis, and D. Papakonstantinou, "Models for mid-term electricity demand forecasting incorporating weather influences," *Energy*, Vol. 31, No. 2-3, pp. 208-227, Feb. 2006.
- [6] T. Hong, J. Wilson, and J. Xie, "Long term probabilistic load forecasting and normalization with hourly information," *IEEE Transactions on Smart Grid*, Vol. 5, No. 1, pp. 456-462, Jan. 2014.
- [7] A. J. Conejo, M. A. Plazas, R. Espínola, S. Member, and A. B. Molina, "Day-ahead Electricity Price Forecasting Using the Wavelet Transform and ARIMA Models," *IEEE Transactions On Power Systems*, Vol. 20, No. 2, pp. 1035-1042, May. 2005.
- [8] D. J. Lee, J. P. Lee, C. S. Lee, J. Y. Lim, and P. S. Ji, "Development of PV power prediction algorithm using adaptive neuro-fuzzy model," *The Transactions of the Korean Institute of Electrical Engineers*, Vol. 64, No. 4, pp. 246-250, Dec. 2015.
- [9] C. S. Lee and P. S. Ji, "Development of daily PV power forecasting models using ELM," *The transactions of the Korean Institute of Electrical Engineers*, Vol. 64P, No. 3, pp. 164-168, Sep. 2015.
- [10] S. M. Lee and W. J. Lee, "Development of a system for predicting photovoltaic power generation and detecting using machine learning," *KIPS Transactions on Computer and Communication Systems*, Vol. 5, No. 10, pp. 353-360,



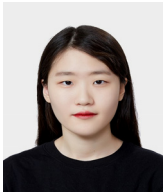
Oct. 2016.

[11] A. Yona, T. Senjyu, T. Funabashi, P. Mandal, and C. H. Kim, "Decision technique of solar radiation prediction applying recurrent neural network for short-term ahead power output of photovoltaic system," *Smart Grid and Renewable Energy*, Vol. 4, No. 6A, pp.32-38, Sep. 2013.

[12] K. H. Lee and W. J. Kim, "Forecasting of 24\_hours Ahead

Photovoltaic Power Output Using Support Vector Regression," *The Journal of Korean Institute of Information Technology*, Vol. 14, No. 3, pp. 175-183, May. 2016.

[13][Internet]. Available: <http://colah.github.io/posts/2015-08-Understanding-LSTMs> [Accessed: Nov.03.2019]



**하 은 규 (Eun-Gyu Ha)**

2016년 3월 ~ 현재 : 가천대학교 IT융합대학 에너지IT학과 재학  
 ※ 관심분야 : 딥러닝, 빅데이터, AI, 사물인터넷



**김 태 오 (Tae-Oh Kim)**

2016년 2월: 청주대학교 전자공학과 공학사  
 2018년 9월 ~ 현재 : 가천대학교 IT융합공학과 석사과정  
 ※ 관심분야 : 딥러닝, 빅데이터, AI, 로봇제어, 안드로이드



**김 창 복 (Chang-Bok Kim)**

1986년 2월: 단국대학교 전자공학과 공학사  
 1989년 2월: 단국대학교 전자공학과 공학석사  
 2009년 2월: 인천대학교 컴퓨터공학과 공학박사  
 1994년 ~ 현재 : 가천대학교 IT대학 에너지IT학과 교수  
 ※ 관심분야 : 데이터 마이닝, 딥러닝, 강화학습, 사물인터넷