Current Photovoltaic Research 11(1) 27-33 (2023)
DOI:https://doi.org/10.21218/CPR.2023.11.1.027
pISSN 2288-3274
eISSN 2508-125X

태양광시스템 모델식과 기계학습을 이용한 발전성능 추정

오현규[1,2] · 신우균[2] · 주영철[2] · 배수현[2] · 황혜미[2] · 강기환[2] · 고석환[2]* · 장효식[1]*
[1]에너지과학기술대학원, 충남대학교, 대전, 34134
[2]태양광연구단, 재생에너지연구소, 한국에너지기술연구원, 대전, 34129

Estimation of Power Using PV System Model Formula and Machine Learning

Hyun Gyu Oh[1,2] · Woo Gyun Shin[2] · Young Chul Ju[2] · Soo Hyun Bae[2] · Hye Mi Hwang[2] · Gi Hwan Kang[2] · Suk Whan Ko[2]* · Hyo Sik Chang[1]*
[1]Graduate School of Energy Science & Technology, Chungnam National University, Daejeon, 34134, Korea
[2]Photovoltaics Research Department, Korea Institute of Energy Research, Daejeon, 34129, Korea

Received February 15, 2023; Revised February 28, 2023; Accepted March 6, 2023

ABSTRACT: In this paper, a machine learning model by using a regression algorithm is proposed to estimate the power generation performance of the BIPV system. The physical model formula for estimating the generation performance and the proposed model were compared and analyzed. For the physical model formula, simple efficiency model, temperature correction model, and regressive physics model for changing an irradiance were used. As a result, when comparing the regressive physics model for changing an irradiance and the proposed model with the actual generation measured data, the respective RMSE values are 0.1497 kW, 0.0451 kW and the accuracy values are 86.44%, and 96.56%. Therefore, the proposed model implemented in this experiment can be useful in estimating power generation.

Key words: BIPV, Machine learning, Linear model, Root mean square error

Nomenclature

P_{MP} : maximum power, W

V_{oc} : open circuit voltage, V

I_{sc} : short circuit current, A

V_{MP} : maximum voltage, V

I_{MP} : maximum current, A

η_{STC} : conversion efficiency of the PV module at STC

A : area of the PV array, m^2

G_{POA} : total irradiance incident on the plane of the PV array, W/m^2

η_T : power loss caused by temperature rising thermal loss

γ : power temperature coefficient of the PV module provided by manufacturers, 1/°C

T_{module} : PV module operational temperature, °C

T_{STC} : PV module temperature at STC, °C

T_a : ambient dry-bulb temperature, °C

T_{NOCT} : PV module temperature at NOCT, °C

m_i : measured value for the instance "i"

s_i : simulated value for the instance "i"

N : number of data in the dataset

$\overline{m}$: average value of all measured data.

Subscript

O&M : operation and maintenance

PV : photovoltaic

STC : standard test condition

NOCT : nominal operating cell temperature

PM : physical model

*Corresponding author: korea19@kier.re.kr (Suk Whan Ko); hschang@cnu.ac.kr (Hyo Sik Chang)

ⓒ 2023 by Korea Photovoltaic Society
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.

1. 서 론

정부의 '재생에너지 3020정책'에 따라 태양광 발전 설비가 증가하고 있는 추세[1]를 보이고 있다. 이에 태양광 발전 설비의 효율적인 운영을 위해서는 안정적인 운영 및 유지보수(operation and maintenance, O&M)기술이 요구된다.

O&M 기술은 태양광 발전소의 모듈 및 인버터 등과 같은 구성 기기가 정상적으로 작동되도록 관리하는 것을 목적으로 하고 있다[2]. 발전소 데이터를 사용한 분석은 성능저하 및 고장요인을 감지하고 개선하는 과정으로 진행된다. O&M은 일반적으로 모니터링, 태양광 모듈 및 구조물의 육안 검사, 태양광 접속함 점검 등이 수행된다[3].

기존에는 단순 측정 검증이 모니터링의 주된 목적이었고[4], 고장 진단은 물리적 모델 방정식만을 이용해 진단하거나 감지했다[5]. 하지만 최근 시스템 손실 분석 및 진단 기능을 갖춘 인공지능(AI) 시스템은 대량의 데이터를 쉽게 적용하고 정확도를 향상시킬 수 있다. 또한 다양한 데이터 세트에 대해 교육할 수 있으므로 다양한 위치, 기후 및 조건에 유연하게 대응할 수 있다. 마지막으로, 추정 및 예측과 관련된 많은 수동 프로세스를 자동화하여 더 빠르고 효율적으로 만들 수 있습니다. 따라서 데이터 기반 기계 학습 모델은 태양광 발전을 위한 과거 데이터, 날씨 패턴 및 기타 관련 변수를 사용하여 정확하게 추정하고 예측할 수 있다.

따라서 본 연구에서는 컬러 모듈을 적용한BIPV 시스템에서 물리모델식(단순 효율 모델, 온도 보정 모델, 일사량 변화에 대한 회귀모델)을 이용하여 발전량을 추정하고 실제 발전량 측정 데이터와 비교하여 가장 정확한 태양광시스템의 물리 모델식을 확인하였다. 또한 물리모델 중 가장 정확도가 높은 모델의 출력값과 환경데이터(일사량,모듈온도)를 회귀 알고리즘을 통한 기계 학습 모델을 제안하여 BIPV 시스템의 발전성능 추정 향상에 적합한 방법을 제안하였다.

2. BIPV 시스템

2.1 개요 및 시스템구성

BIPV 시스템은 Fig. 1과 같이 대전 A 연구원의 남향 외벽에 설치되어 있다. 시스템 구성으로는 3.5 kW급 인버터 3대를 사용하였으며, 본 논문에서는 1대 인버터의 측정 데이터를 이용하여 실험을 수행하였다. 설치된 컬러 BIPV 모듈은 dot 패턴으로 컬러를 구현하며 최대 발전량 2.5 kW로 26장이 직렬로 연결되어 있다. 모듈, 인버터, 시스템의 사양은 Table 3, 4, 5와 같다. 기상환경 데이터를 수집하기 위해 모듈 설치 각도와 동일하게 일사량계를 설치하고, 모듈 후면에 열전대 센서를 부착하여 모듈 온도를 측정하였다. 측정 장비 및 센서의 사양의 설명은 Table 1, 2

Fig. 1. BIPV system applicated building

Table 1. Spec. of pyranometer

Specification	Value
Spectral Range	400 to 1100 nm
Sensitivity	60 to 100 $\mu V/W/m^2$
Response Time	< 500 ns
Non-linearity	0 to 1000 W/m^2
Maximum solar irradiance	2000 W/m^2
Operational Temperature range	-40°C to +80°C
Directional response	< 10 W/m^2 (up to 80° with 1000 W/m^2 beam)
Field of view	180°
Non-stability	< 2% (change/year)

Table 2. Spec. of the K type thermocouple

Specification	Value
Measurement Temperature range	-200°C to 1250°C
Std. Limits of Error	Greater of 2.2°C or 0.75%
Spec. Limits of Error	Greater of 1.1°C or 0.4%

Table 3. Spec. of Color module

$P_{\max}$ [W]	96
I_{MP} [A]	7.34
V_{MP} [V]	13.25
I_{sc} [A]	7.69
V_{oc} [V]	15.53
Area [m^2]	0.9

Table 4. Spec. of system

Series module number	26
Parallel module number	1
Capacity [kW]	2.50

Table 5. Spec. of inverter

DC	MPPT voltage range [V]	100 ~ 400
	Maximum voltage [V]	500
	Maximum current [A]	18
AC	Rated power [kW]	3.5
	Maximum efficiency [%]	97~

와 같다. 기상데이터와 실제 발전량 데이터는 CR 1000 데이터 로거를 통하여 10분 평균에 해당하는 데이터를 취득하였다. 본 연구를 위한 분석 기간은 2021년 3월부터 2022년 9월까지 18개월이다.

3. BIPV 시스템 발전량 추정기법

3.1 데이터 전처리 및 추정기법 프로세스

컬러 모듈을 적용한 BIPV 시스템의 발전량을 추정하기 위해 세 개의 태양광 물리 모델식을 이용했다. 모델들의 발전량을 추정하기 위한 입력데이터로서 일사량과 모듈온도, 면적, 효율의 값이 요구된다. 300 W/m^2 미만의 저일사량 구간에서는 인버터가 간헐적으로 정지되는 경우가 많아 데이터 분석에서 제외하였다. 또한, DC 어레이 전압의 경우도 300 V이상의 값을 사용하였다.

모델 구현은 MATLAB 회귀학습기를 활용하였다. 회귀학습기의 기존 선형 모델에서 일사량, 모듈 온도, 태양광시스템 물리 모델식의 출력값을 입력요소로 사용하여 회귀알고리즘을 통한 기계 학습 모델을 구현한다. 교차검증은 5회 실시되었으며, 훈련데이터와 테스트 데이터의 비율은 70:30이다. 모델 프로시저는 Fig. 2와 같다.

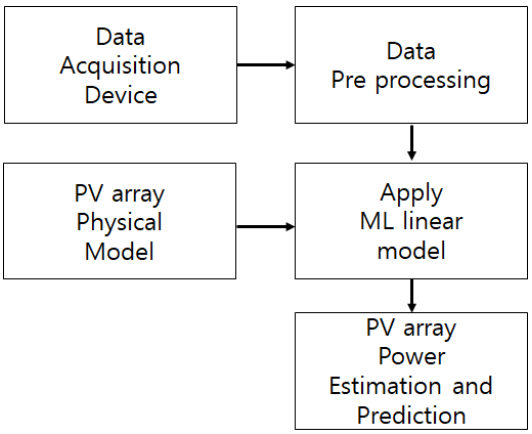


Fig. 2. Procedure of ML model algorithm

3.2 물리모델

태양광시스템의 물리 모델식 총 3개를 사용했다. 단순효율모델과 온도보정모델과 같은 간단한 모델과 일사량 변화에 대한 회귀모델로 모듈온도가 25°C 인 조건에서 일사량에 따라 변화하는 PV 모듈의 MPP (Maximum Power Point) 값을 산출하여 변화된 일사량에 따른 I_{MP}, V_{MP} 값을 도출하고 이를 회귀식으로 나타낸 모델이다[6].

3.2.1 단순 효율 모델

단순 효율 모델은 일사량 변화에 따라 효율이 달라지는 점을 고려한 단순 PV 모듈 모델이다. 모델의 출력은 다음 식과 같이 STC조건에서의 에너지 변환 효율, PV 어레이의 면적, POA 태양 복사 조도를 기준으로 추정된다[7].

$$P_{MP} = \eta_{STC} A G_{POA} \ [\mathrm{W}] \tag{1}$$

여기서 P_{MP}은 전력 출력 [W], η_{STC}는 STC에서 PV 모듈의 변환 효율, A는 PV 어레이의 면적 [m^2], G_{POA}는 PV 어레이의 평면에 입사된 총 방사조도 [W/m^2]이다.

3.2.2 온도 보정 모델

온도 보정 모델은 식 (1)보다 향상된 모델이다. 셀의 온도가 증가함의 따라 PV 모듈의 출력은 낮아진다. 전력 손실은 전력의 온도 계수, STC 작동조건에서의 온도 차이로 설명할 수 있다[8].

$$P_{MP} = \eta_T \eta_{STC} A G_{POA} \ [\mathrm{W}] \tag{2}$$

$$\eta_T = 1 + \gamma (T_{module} - T_{STC}) \tag{3}$$

$$T_{module} = T_a + G_{POA} \left(\frac{T_{NOCT} - 20}{800} \right) [\mathrm{°C}] \tag{4}$$

여기서, η_T는 온도 상승에 따른 열 손실의 대한 전력 손실 효율이고, γ는 제조업체가 제공하는 PV 모듈의 전력 온도 계수이다[1/°C]. T_{module}은 PV 모듈 작동 온도[°C], T_{STC}는 STC에서의 모듈 온도[°C], T_a는 주변 건구 온도[°C], T_{NOCT}는 공칭 작동 셀 온도(NOCT)에서 PV 모듈 온도이다[°C].

$$\begin{aligned} P_{MP} = {} & \eta_{STC} A G_{POA} + \gamma T_a \eta_{STC} A G_{POA} \\ & + \gamma G_{POA} \left(\frac{T_{NOCT} - 20}{800} \right) \eta_{STC} A G_{POA} \\ & - \gamma T_{STC} \eta_{STC} A G_{POA} \ [\mathrm{W}] \end{aligned} \tag{5}$$

총 출력은 (5)와 같다.

3.2.3 일사량 변화에 대한 회귀 물리 모델

온도 조건을 25°C에서 PV 시뮬레이터의 장비를 이용하여 PV 모듈의 I-V curve를 측정한다. 일사량을 200, 400, 600, 800, 1000[W/m^2]로 각 각 구분해 측정을 진행한다. 변화된 일사량에 대한 I_{MP}, V_{MP}값을 산출하고 그 값을 회귀식으로 나타냈다. 그 결과 I_{MP}의 경우 1차 함수, V_{MP}의 경우 로그함수로 나타낼 수 있다.

$$I_{MP} = 0.007 G_{POA} + 0.0103 \ [\mathrm{A}] \tag{6}$$

$$V_{MP} = 0.2745\ln(G_{POA}) + 11.35 \ \text{[V]} \quad (7)$$

$$V_{MP\text{온도보정}} = (0.2745\ln(G_{POA}) + 11.35)(1 - 0.003(T_{module} - 25))\text{[V]} \quad (8)$$

$$P = V \times I \ \text{[W]} \quad (9)$$

식 (7)에 나타난 V_{MP}는 온도조건이 25°C 인 경우에 해당되며, 온도계수 0.3%/°C을 고려해주면 식 (8)과 같이 나타낼 수 있다[6].

3.3 모델 정확성 평가

위 세 모델의 정확성을 검증하기 위해 기상 환경 모니터에 따른 BIPV 시스템에서의 실제 발전량 값과 예측 발전량 값을 비교했다. 추정 발전량은 식 (1), (5), (9)를 통해 구할 수 있다.

모델의 정확도를 판단하기 위해 결정계수(Coefficient of Determination : R^2)와 평균 제곱근 오차(Root Mean Square Error : RMSE)를 이용하여분석하였다.

3.3.1 결정계수

R^2은 시뮬레이션 된 데이터가 측정된 데이터 또는 적합한 회귀선에 얼마나 가까운지 보여준다. R^2의 값은 0부터 1이며 1에 같다는 것은 시뮬레이션 모형 또는 적합한 회귀선이 측정된 데이터와 완벽하게 적합 되는 것을 의미한다.

일반적으로 R^2가 높을수록 모형이 측정된 데이터와 적합하다.

R^2식은 식 (10)을 통해 계산할 수 있다.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(m_i - s_i)^2}{\sum_{i=1}^{N}(m_i - \overline{m})^2} \quad (10)$$

3.3.2 평균 제곱근 오차

RMSE는 실제 편차의 기간별 비교를 통해 시뮬레이션 모델의 성능을 나타낸다. 값이 0이면 측정된 데이터에 완벽하게 적합한 모형을 의미한다. RMSE는 아래의 식 (11)을 통해 계산할 수 있다.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(m_i - s_i)^2}{N}} \quad (11)$$

3.3.3 선형 회귀 분석

선형회귀분석법은 회귀모델의 가장 기본적이고 일반적으로 사용되는 예측모델이다. 주어진 데이터를 나타내는 최적의 직선을 찾아냄으로써 입력(x)과 출력(y)사이의 선형적인 관계를 모델링하는 회귀 분석 기법이다. 단순 선형 회귀 수식을 표현하면 $y = a + bx$과 같다. 이 수식은 예측에 관련된 독립변수 x를 통해 예측하려는 결과 값인 종속변수 y를 도출해내는 것이다[9].

3.4 물리모델과 실제 출력값 비교

Fig. 3은 BIPV 시스템에서 2022년 1월 데이터를 사용하여

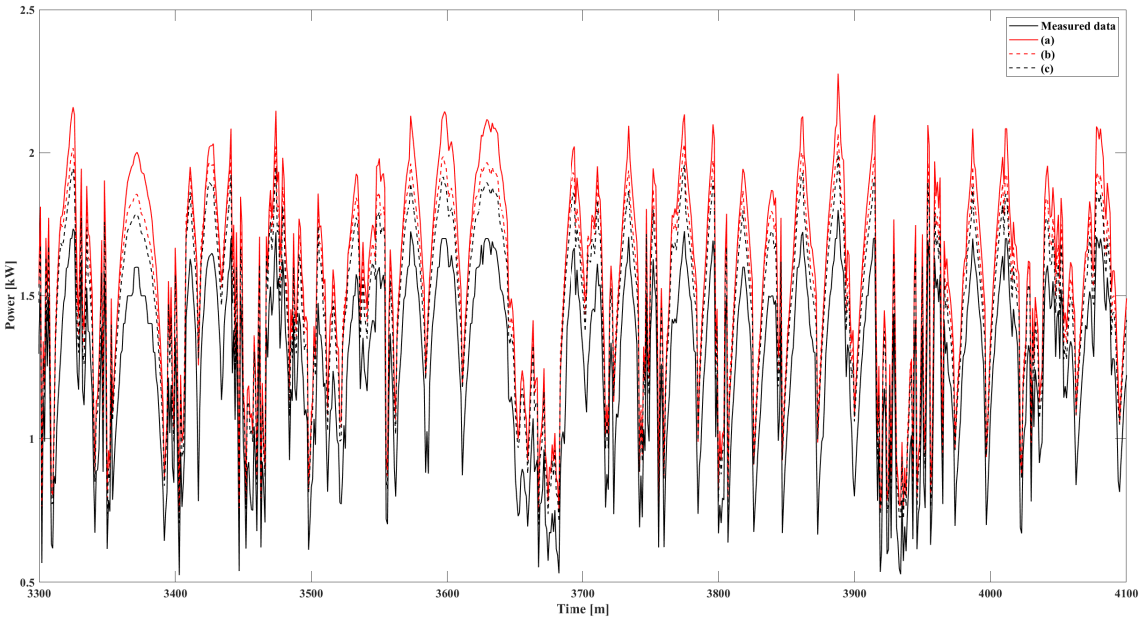
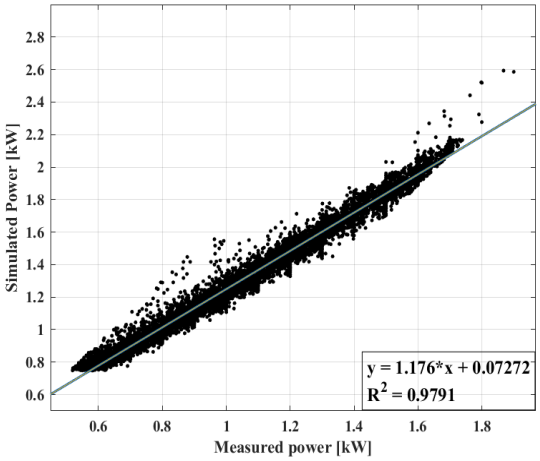
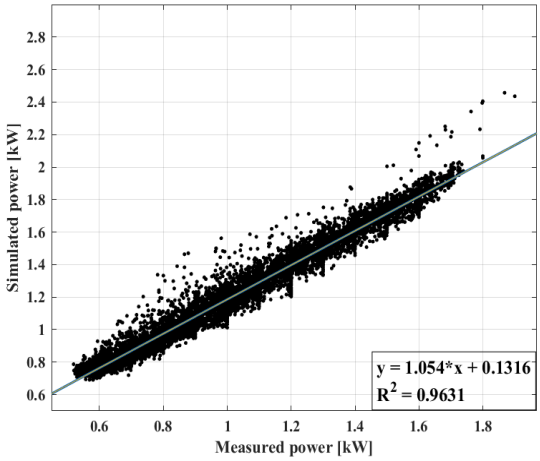


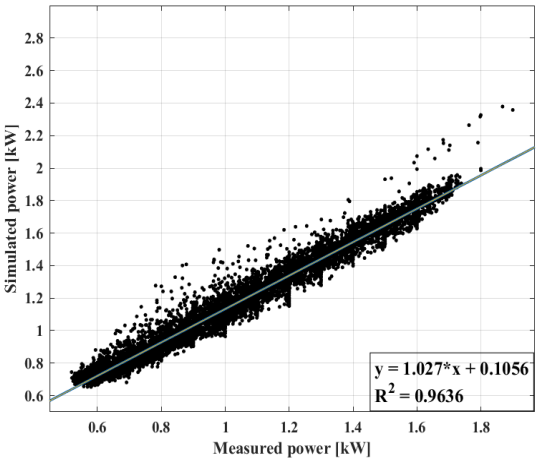
Fig. 3. Comparison between physical models result and measured data for January 2022
(a) Simple efficiency model, (b) Temperature correction model, (c) Regressive physics model for changing an irradiance



(a) Simple efficiency model



(b) Temperature correction model



(c) Regressive physics model for changing an irradiance

Fig. 4. Linear regression analysis between physical models result and measured data

Table 6. Summary of the physical models

	Slope	R-Square	RMSE (kW)	Accuracy (%)
(a)	1.176	0.9791	0.2746	74.86
(b)	1.054	0.9631	0.2019	81.15
(c)	1.027	0.9636	0.1497	86.44

3.2절 물리모델들의 출력값과 실제 측정 발전량 데이터를 비교했다. 일사량 변화에 대한 회귀모델, 온도보정모델, 단순효율모델 순으로 측정데이터에 근접함을 보였다. Fig. 4는 물리모델들과 측정데이터의 적합함을 확인하기 위해 선형회귀분석을 실시했다. Table 6에는 기울기와 결정계수, 측정데이터와의 RMSE, 정확도를 나타내었다. 그 결과 세 모델 모두 결정계수가 1에 가까운 값을 나타냈고 각 각 물리모델의 출력값과 측정데이터의 RMSE를 계산한 결과로는 단순효율모델은 0.2746 kW, 온도보정모델은 0.2019 kW, 일사량 변화에 대한 회귀모델이 0.1497 kW로 일사량에 변화에 대한 회귀모델이 가장 오차 적음을 확인하였다. 또한 모델들의 정확도는 각 각 74.86%, 81.15%, 86.44% 임을 확인하였다.

따라서 일사량 변화에 대한 회귀모델이 가장 작은 RMSE값과 1에 가까운 결정계수, 가장 높은 정확도로 인하여 물리모델 중 가장 정확한 모델임을 입증하였다.

3.5 물리모델과 인공지능 기술이 결합된 출력 추정값 비교

Fig. 5의 결과는 BIPV 시스템에서 2022년 1월 데이터를 사용하여 물리모델 중 가장 정확했던 일사량 변화에 대한 회귀모델과 3.1절 회귀알고리즘을 이용한 기계학습 모델을 구현 후 비교하였다. 제안하는 모델이 일사량에 변화에 대한 회귀모델보다 실제 발전량 측정데이터와 근접함을 확인하였다. 제안하는 모델 또한 선형회귀분석을 실시하여 Fig. 6에 나타냈다. 세부값은 Table 7에 나타냈다. 제안한 모델의 결정계수가 1에 가까운 값을 보였으며 측정데이터와의 RMSE는 0.0451 kW, 정확도는 96.56% 임을 확인하였다. 3.4절 일사량 변화에 대한 회귀모델은 물리모델 중 가장 정확했지만, 제안한 모델은 0.1046 kW 감소된 RMSE, 10.12% 향상된 정확도를 보여줬다. 이에 따라 제안한 모델이 발전량 추정에 있어 유용하게 사용될 것이다.

4. 결 론

본 논문에서는 2021년 3월부터 2022년 9월까지 태양광 시스템 모델식 3가지를 사용하여 측정데이터와 비교를 해봤다. 그 결과는 단순효율모델의 RMSE값은 0.2746 kW, 온도보정모델의 RMSE값은 0.2019 kW, 일사량 변화에 대한 회귀모델의 RMSE값은 0.1497 kW임을 확인하였고, 모델들의 정확도는 각 각 74.86%, 81.15%, 86.44% 임을 확인하였다. 결과 값에 따라서 일사량 변화에 대한 회귀 물리모델이 측정데이터와 가장 적합했던 태양광시스템의 모델식임을 확인하였다.

물리모델 중 가장 정확했던 일사량 변화에 대한 회귀 모델의 출력값과 기상데이터(일사량, 모듈온도)를 회귀알고리즘으로 학습하여 기계 학습 모델을 구현 후 제안하였다. 제안한 모델을 실제 발전량 측정데이터와 비교 분석한 결과 RMSE값은 0.0451 kW, 정확도는 96.56%를 확인하였다. 따라서 가장 정확했던

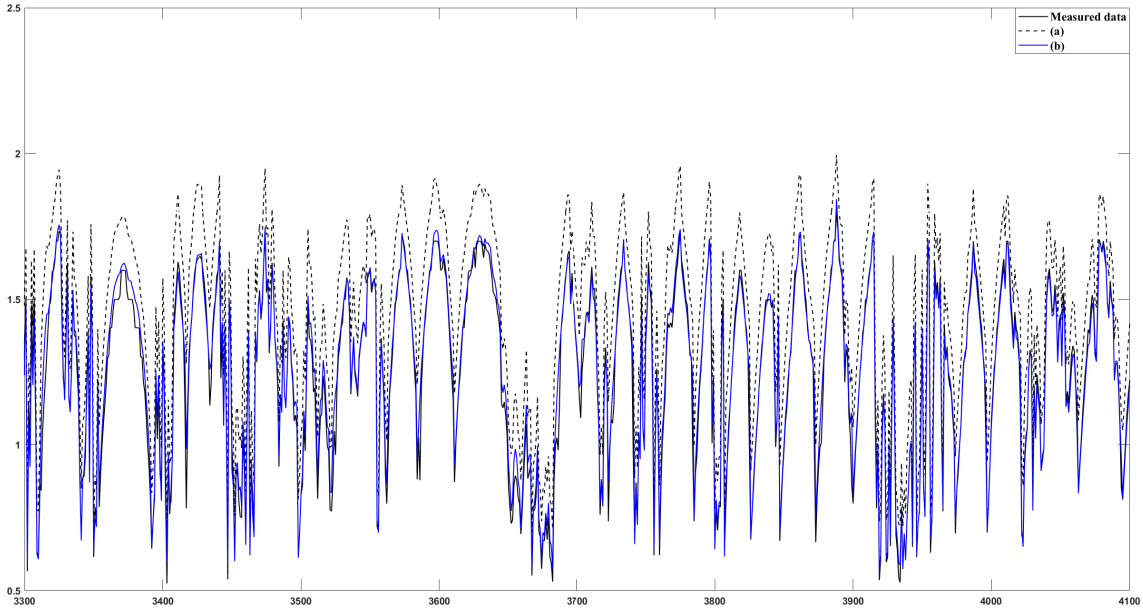


Fig. 5. Comparison between models result and measured data for January 2022
(a) Regressive physics model for changing an irradiance, (b) proposed model

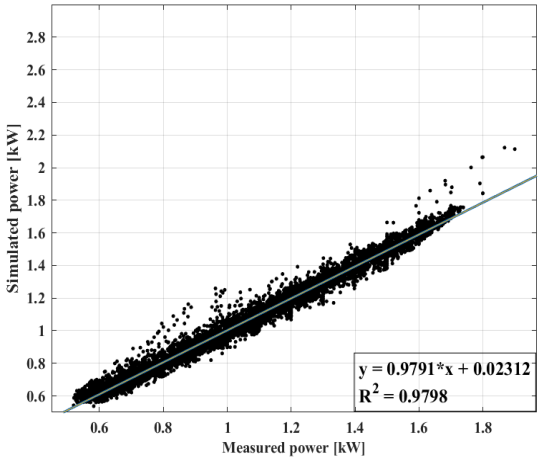


Fig. 6. Linear regression analysis between proposed model and measured data

Table 7. Summary of the proposed model

Slope	R-Square	RMSE (kW)	Accuracy (%)
0.9791	0.9798	0.0451	96.56

물리모델 보다 RMSE값은 0.1046 kW 감소하였고, 정확도는 10.12% 향상되었다. 따라서 본 논문에서는 물리모델만을 사용하여 발전량을 추정하는 기법보다 인공지능이 결합된 추정기법을 이용한다면 보다 정확한 추정 및 예측이 가능한 것을 입증하였다. 추 후 제안한 모델의 정밀도를 향상시키기 위해서는 양질의 데이터를 대량으로 투입하고 더 세분화된 형태의 빅데이터를 구축한다면 보다 향상된 결과를 확인 할 것이다.

후 기

본 연구는 산업통상자원부와 한국에너지기술평가원의 지원을 받아 수행한 연구 과제입니다(No:20223030010200).

References

1. Ministry of Trade, Industry and Energy, Renewable Energy 3020 Implementation Plan, Ministry of Trade, Industry and Energy (2017).
2. Di Lorenzo, Gianfranco, G., Araneo, R., Mitolo, M., Niccolai, A., & Grimaccia, F., Review of O&M practices in PV plants: Failures, solutions, remote control, and monitoring tools, IEEE Journal of Photovoltaics, 10(4), 914-926 (2020).
3. Klise, G., & Balfour, J. A Best Practice for Developing Availability Guarantee Language in Photovoltaic (PV) O&M Agreements, Office of Scientific and Technical Information (OSTI) (2015).
4. Ko, S.-W., So, J.-H., Hwang, H.-M., Ju, Y.-C., Song, H.-J., Shin, W.-G., … Kang, I.-C. The Monitoring System with PV Module-level Fault Diagnosis Algorithm. Journal of the Korean Solar Energy Society. The Korean Solar Energy Society (2018, June 1).
5. Shin, W. G., Oh, H. G., Bae, S. H., Ju, Y. C., Hwang, H. M., & Ko, S. W. Fault Diagnosis of PV String Using Deep-Learning and I-V Curves. Current Photovoltaic Research, 10(3), 77-83 (2022).
6. Shin, J.-Y., Ko, S.-W., Shin, W.-G., Hwang, H.-M., Ju, Y., Kang, G.-H., & Chang, H.-S. Proposal of Power Estimation Model of Color BIPV System. Journal of the Korean Solar Energy

Society. The Korean Solar Energy Society (2021, October 1).
7. Makrides, G., Zinsser, B., Schubert, M., & Georghiou, G. E. Energy yield prediction errors and uncertainties of different photovoltaic models. Progress in Photovoltaics: Research and Applications. Wiley (2011, November 15).
8. Roberts, J. J., Mendiburu Zevallos, A. A., & Cassula, A. M. Assessment of photovoltaic performance models for system simulation. Renewable and Sustainable Energy Reviews. Elsevier BV (2017, May).
9. Lee, Y. K., Shin, W.-G., Ju, Y.-C., Hwang, H.-M., Kang, G.-H., Ko, S.-W., & Chang, H.-S. Estimation of PV Power Generation by Linear Regression Model Using Voltage and Current Data. Journal of the Korean Solar Energy Society. The Korean Solar Energy Society (2021, October 1).