Check for updates

Journal of KIIT. Vol. 24, No. 10, pp. 93-105, Oct. 31, 2026. pISSN 1598-8619, eISSN 2093-7571
http://dx.doi.org/10.14801/jkiit.2026.24.10.93

계절 적응 기반의 강건 Clear Sky Index 경계와 주간 한정 혼합을 적용한 태양광 발전량 예측 모델 연구

이영학*, 심재창**

Research on a Solar Power Generation Prediction Model Applying Seasonally Adapted Robust Clear Sky Index Boundaries and Day-Limited Mixing

Yeung-Hak Lee*, Jae-Chang Shim**

본 과제(결과물)는 2026년도 교육부 및 경상북도의 재원으로 경북앵커센터의 지원을 받아 국립경국대학교에서 수행된 지역성장 인재양성체계(앵커) 특성화지방대학 30의 결과입니다. (2026-ANCHOR-15-104)

요 약

태양광 발전은 태양 기하에 의한 결정론적 성분과 기상에 의한 확률적 성분이 중첩되어 결정되므로, 두 성분을 분리하는 청천지수(CSI, Clear-Sky Index) 정규화가 예측 정확도에 영향을 미친다. 본 연구는 기존 청천지수를 계절 적응, 경계의 강건화, 주간 한정 혼합복원으로 확장한 A-CSI(Adapted CSI) 알고리즘을 제안한다. 정규화 없음, 기본 CSI, 계절 적응, 제안한 A-CSI의 네 가지 정규화를 다섯 예측 모델과 두 예측 지평(24시간, 168시간)에 적용하여 효과를 검증하였다. 실험 결과, 24시간 예측에서 A-CSI 계열 정규화는 정규화 없음(평균 RMSE 19.26) 대비 예측 정확도를 향상시켰으며, 계절 적응의 평균 RMSE는 16.97로 11.9% 감소하였다. 개선 효과는 딥러닝 계열에서 특히 두드러졌다. 본 연구는 제안한 A-CSI 알고리즘이 딥러닝 계열의 단기 예측에 효과적임을 실증하고, 모델과 예측 지평에 따른 정규화 적용 지침을 제시한다.

Abstract

Solar power generation integrates deterministic solar geometry with stochastic weather patterns. Consequently, Clear-Sky Index (CSI) normalization, which separates these two components, substantially impacts forecasting precision. Here, we propose the Adapted CSI (A-CSI) algorithm, enhancing traditional CSI with seasonal adaptation, robust boundaries, and day-limited mixed restoration. We evaluated four normalization schemes—unnormalized, basic CSI, seasonal adaptation, and proposed A-CSI—across five forecasting models for 24-hour and 168-hour horizons. For 24-hour predictions, A-CSI normalization markedly improved accuracy compared to no normalization (average RMSE 19.26). Notably, seasonal adaptation yielded an average RMSE of 16.97 (an 11.9% error reduction), showing the greatest improvements in deep-learning models. These findings highlight that the proposed A-CSI algorithm is particularly effective for short-term deep-learning-based forecasting and offer practical guidelines for choosing normalization schemes based on model types and prediction horizons.

Keywords

photovoltaic, generation prediction, clear-sky index, seasonal adaptation

* ㈜기븐텍연구소 소장
- ORCID: https://orcid.org/0000-0002-9037-2646
** 국립경국대학교 컴퓨터공학과 교수(교신저자)
- ORCID: https://orcid.org/0009-0000-5392-4512

· Received: Sep. 16, 2026, Revised: Sep. 30, 2026, Accepted: Oct. 03, 2026
· Corresponding Author: Jae-Chang Shim
Computer Science & Engineering, 1375 Gyeongdong-ro, Andong-si, Korea
Tel.: +82-54-820-5645, Email: jcshm@gknu.ac.kr

Ⅰ. 서 론

전 세계 태양광(PV, Photovoltaic) 설비 용량은 급격히 증가하여 2022년 말 누적 1,186GW에 도달하였으며, 각국의 탄소중립 목표 달성을 위해 이러한 확대는 지속될 전망이다[1]. 그러나 태양광 발전이 전력 계통에 대규모로 통합되면서, 태양광 출력의 간헐성과 변동성은 예비력 확보 비용 증가, 급전 가능 발전량 관리, 계통 신뢰도 저하 등 심각한 운영상의 과제를 야기하고 있다[1][2]. 재생에너지 침투율이 30%를 초과하면 통합 비용이 급격히 상승하므로[3], 태양광의 안정적 계통 수용을 위해서는 다양한 시간 규모에 걸친 정확한 발전량 예측이 필수적이다[2]. 그리고 재생에너지 확대에 따라 태양광 발전의 계통 비중이 빠르게 증가하고 있지만, 태양광 출력은 기상에 의존하여 간헐적이므로, 계통의 안정적 운영을 위해서는 정확한 발전량 예측이 필수적이다[4].

태양광 출력은 태양의 위치(일출·일몰, 태양고도, 계절)에 의한 결정론적(Deterministic) 성분과 구름·강수·에어로졸 등 기상에 의한 확률적(Stochastic) 성분이 중첩되어 결정된다[5][6]. 예측 모델이 두 성분을 동시에 학습하기는 어렵기 때문에, 실측 출력을 구름 없는 이상적 조건의 기대 출력(청천 출력)으로 나눈 청천지수(CSI, Clear-Sky Index)로 정규화하여 결정론적 성분을 제거하는 전처리가 널리 사용된다[7]. 청천지수는 일주기·계절 신호를 제거하여 예측 대상을 기상 잔차로 단순화하고, 신경망이 대기 변동 학습에 집중하도록 돕는다[8][9].

그러나 기존 청천지수 정규화는 세 가지 구조적 한계를 가진다. 첫째, 고정된 청천 기준은 잔여 계절 변동을 완전히 제거하지 못한다. 실제로 청천지수와 청명도 시계열은 남은 계절 변동으로 인해 분산이 변화하여 완전히 정상적(Stationary)이지 않고 국소적으로만 정상적임이 보고되었다[10]. 이는 계절별 대기 투과 특성(예: 여름의 높은 습도·에어로졸, 겨울의 맑은 대기)이 고정 기준에 반영되지 못하기 때문으로, 계절에 적응하는 청천 기준의 필요성을 시사한다. 둘째, 청천지수는 상한이 불안정하다. 구름에 의한 과일사(Overirradiance) 현상으로 청천지수가 1을 초과할 수 있으며, 야간 및 아침과 저녁 박명의 작은 청천 기준으로 인해 극단값이 발생한다[11]. 고정 상한을 사용하는 기존 방식은 데이터 특성을 반영하지 못한다. 셋째, 단순 곱셈 복원은 청천 기준 자체의 계통 오차를 그대로 전이시킨다. 곱셈 방식 청천 정규화의 이러한 한계가 지적되면서, 이를 우회하는 대안적 접근이 제안되고 있다[12].

더욱이 청천 정규화의 효과는 보편적이지 않고 예측 모델과 예측 지평에 의존하는 것으로 보고되었다. Silva et al.[9]은 정상화(Stationarization)가 대부분의 기계학습 모델에서 비선형 패턴을 드러내 성능을 향상시키나 RandomForest는 예외적으로 오히려 성능이 저하됨을 관찰하였고, 특히 단기 지평에서 청천 정상화의 이득이 크다고 보고하였다. Yang[7]은 청천 모델의 선택 자체가 예측 결과를 좌우함을 보였으며, 비선형 신경망은 정규화 방식에 상대적으로 둔감한 반면 선형 기법은 정규화에 크게 의존한다는 연구도 있다[10]. 이러한 조건 의존성은 정규화 기법을 모델 계열과 예측 지평에 맞추어 설계·선택해야 함을 의미하나, 다양한 모델과 지평에 걸쳐 체계적으로 검증한 연구는 부족하다.

본 연구는 앞서 지적한 기존 청천지수의 세 가지 한계를 개선하기 위해 A-CSI(Adapted-CSI) 알고리즘을 제안한다. 1) 청천 기준을 월별 계절 스케일로 조정하는 계절 적응(M1)으로 잔여 계절 변동을 제거하고, 2) 학습 데이터 분포로부터 상한을 학습하는 강건 경계(M2) 극단값을 제어하며, 3) 곱셈 복원에 잔차 보정을 결합한 주간 한정 혼합 복원(M3)으로 계통 오차를 교정한다. 본 연구의 목적은 A-CSI 정규화가 태양광 발전량 예측 정확도를 향상시키는지 검증하고, 모델과 예측 지평에 따른 효과를 규명하는 것이다. 주요 기여는 다음과 같다.

첫째, 기존 청천지수의 계절 미적응 및 극단값과 복원 오차 한계를 각각 개선한 A-CSI 알고리즘을 제안한다. 둘째, 정규화 없음(Basic), 기본 CSI, M1, A-CSI의 네 가지 수준을 다섯 예측 모델 LSTM(Long Short-Term Memory), CNN-LSTM (Convolutional Neural Network Long Short-Term Memory), LightGBM(Light Gradient Boosting Machine), XGBoost(eXtreme Gradient Boosting),

N-HiTS(Neural Hierarchical Interpolation for Time Series Forecasting)과 두 예측 지평(24시간, 168시간)에 걸쳐 교차 검증한다. 셋째, RMSE와 R^2를 기준으로 정규화 효과의 모델 계열 및 예측 지평 의존성을 정량화하고, 실무적 정규화 선택 지침을 제시한다.

II. 관련 연구

2.1 태양광 발전량 예측 모델

태양광 예측 모델은 크게 물리 모델, 통계 모델, 기계학습·딥러닝 모델로 구분되며, 최근에는 딥러닝이 주류를 이루고 있다[12]. 발표된 태양광 예측 연구를 분석한 문헌에 따르면, LSTM(31편), CNN(21편)이 가장 널리 사용되었고 BiLSTM과 GRU도 각각 9편에서 채택되었다[13]. 이는 순환 및 합성곱 구조가 태양광 시계열의 시간 의존성과 변수 간 상관을 효과적으로 학습하기 때문이다.

순환신경망 계열에서 장단기 메모리(LSTM)는 기울기 소실 문제를 완화하여 시계열의 장기 의존성을 학습하며, 태양광 예측의 표준 모델로 자리잡았다[4]. 양방향 LSTM(BiLSTM)은 시퀀스를 순방향과 역방향으로 모두 처리하여 시간 순서 정보를 더 풍부하게 포착한다. Herrera-Casanova et al.[14]은 BiLSTM에 베이지안 최적화를 결합하여 1시간 전 태양광 예측에서 우수한 성능을 보고하였고, Guo et al.[15]은 특이 스펙트럼 분석(SSA)으로 분해한 부신호를 병렬 BiLSTM으로 학습하여 예측 정확도를 향상시켰다.

합성곱-순환 혼합 계열의 CNN-LSTM은 합성곱층으로 변수 간 국소 상관(공간 특징)을, LSTM으로 시간 상관을 학습하는 혼합 구조이다. Agga et al.[16],은 CNN-LSTM이 단독 딥러닝 모델보다 단기 태양광 예측에서 우수함을 보였으며 Jakoplić et al.[17]은 하늘 영상과 CNN-LSTM을 결합하여 예측값과 실측값 사이에 74%의 상관을 달성하였다. 나아가 변분 모드 분해(VMD, Variational Mode Decomposition)를 CNN-LSTM에 결합하면 여러 시간 지평에서 성능이 개선되는 것으로 보고되었다 [18].

장기 다중스텝 예측에서는 계층적 보간(Hierarchical interpolation)을 사용하는 N-HiTS가 낮은 계산 비용으로 높은 정확도를 달성한다[19]. N-HiTS는 다중 해상도의 신호 분해와 보간을 통해 긴 예측 구간을 효율적으로 생성하나, 본 연구는 이러한 보간 구조가 특정 정규화와 상호작용할 수 있음을 실험적으로 관찰한다.

부스팅 및 앙상블 트리 계열은 딥러닝과의 비교에서 부스팅 트리와 앙상블 트리는 여전히 강건한 경쟁 기준선으로 남아 있다. XGBoost는 정규화 항을 손실 함수에 포함하여 과적합을 억제하면서 재생에너지의 비선형 패턴을 학습하며[20][21], LightGBM은 히스토그램 기반 분할과 잎 우선(Leaf-wise) 성장으로 대용량 시계열을 빠르고 효율적으로 처리한다[22]. RandomForest는 최소한의 조정만으로도 안정적이고 해석 가능한 예측을 제공하여 실무에서 널리 채택된다[23]. 실제로 여러 에너지 예측 경진대회에서 부스팅 트리가 우수한 성능을 보였으며, 개별 예측 과제의 제한된 학습 데이터 환경에서는 강한 귀납적 편향을 가진 소형 트리 앙상블이 대용량 딥러닝과 경쟁력을 유지한다는 대규모 비교 연구도 있다[24].

본 연구는 다섯 모델(LSTM, CNN-LSTM, LightGBM, XGBoost, N-HiTS)을 채택하여 딥러닝과 트리 계열을 아우르는 비교를 수행한다.

2.2 청천지수 정규화와 변형 기법

태양광·일사 시계열을 정규화하는 방식은 정규화 기준에 따라 구분된다. 가장 널리 쓰이는 청천지수(CSI)는 실측값을 물리적 청천 모델의 기대 출력으로 나눈 것으로, 태양 기하에 의한 일주기·계절 신호를 제거한다[7][10]. 이와 대비되는 청명도 지수(Clearness index)는 대기권 밖 일사(Extraterrestrial irradiance)로 정규화하며, 청천 모델의 대기 파라미터 설정이 불필요하다는 장점이 있으나 일반적으로 청천지수보다 예측 성능이 낮은 것으로 알려져 있다[10]. Lauret et al.[10]은 여섯 개 지점에서 청천지수와 청명도를 비교하여, 청천 모델 기반 정규화가 모델 종류와 무관하게 더 나은 예측을 산출함을 보였다.

그러나 청천 정규화는 여러 실무적 한계를 가진

다. 물리적 청천 모델은 지점별로 정확한 대기 파라미터를 요구하는데, 이 값이 부정확하면 예측 정확도가 저하된다[10]. 또한, 청천 모델 자체의 선택이 예측 결과를 좌우한다[7]. 이에 따라 청천 모델 없이 청천 조건을 암묵적으로 학습하는 대안(예: 극단 학습 기계 기반 비선형 적응 기법)도 제안되어, 곱셈 방식 청천 정규화의 한계를 우회하려는 시도가 이어지고 있다[12]. 최근의 물리 정보 기반(Physics-informed) 모델은 청천 복사 원리를 예측에 통합하여 정확도를 높이기도 한다[20]. 본 연구가 제안하는 A-CSI는 청천 모델을 대체하는 대신, 경험적 청천 기준을 계절 적응(M1), 강건 경계(M2), 혼합 복원(M3)으로 확장하여 위 한계를 개선한다는 점에서 이들과 차별된다.

2.3 정규화 효과의 조건 의존성

청천 정규화의 효과가 예측 모델과 지평에 따라 달라진다는 관찰은 본 연구의 핵심 동기이다. Silva et al.[9]은 초단기 공간-시간 일사 예측에서 정상화(Stationarization)가 대부분의 기계학습 모델에서 비선형 패턴을 드러내 성능을 향상시키나, RandomForest는 예외적으로 오히려 성능이 저하되며, 정상화의 이득이 특히 단기 지평에서 크다고 보고하였다. Lauret et al.[10]은 비선형 신경망이 정규화 방식에 상대적으로 둔감한 반면, 선형 모델은 정규화에 크게 의존함을 보였다. 또한 청천지수와 청명도 시계열이 잔여 계절 변동으로 인해 완전히 정상적이지 않고 국소적으로만 정상적임이 지적되어, 계절에 적응하는 정규화의 필요성이 제기되었다. 이러한 선행 연구들은 정규화를 모델 계열과 예측 지평에 맞추어 설계·선택해야 함을 시사하나, 다양한 모델과 지평에 걸쳐 정규화 변형의 효과를 체계적으로 비교 및 정량화한 연구는 부족하다. 본 연구는 네 수준의 정규화를 다섯 모델과 두 지평에 걸쳐 교차 검증하여 이 부분을 나타내고자 한다.

2.4 예측 성능 평가

태양광 예측의 정확도는 주로 RMSE, MAE, R^2로 평가되며[25][26], 모델의 정확도와 오차의 특성을 측정하여, 실제 전력 계통 운영 환경에서 해당 모델이 얼마나 신뢰할 수 있는지 판단하는 데 사용된다. 본 연구에서 태양광 발전량 예측 정확도를 평가하기 위해 RMSE(Root Mean Squared Error), MAE (Mean Absolute Error), 결정계수(R^2 Score)를 사용하였다. MAE는 오차의 평균적인 크기를 예측하고자 하는 발전량의 단위(예: MWh)로 직접적으로 보여준다. 모든 오차에 대해 동일한 가중치를 부여하므로, 이상치(Outlier)에 비교적 덜 민감하고 직관적인 특징이 있다. 평균 절대 오차는 식 (1)과 같다.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y_i}| \tag{1}$$

RMSE는 MSE값에 제곱근을 취한 값이다. MSE의 단위를 다시 원래의 단위(예: MWh)로 되돌려 MAE와 마찬가지로 오차의 크기를 직관적으로 해석할 수 있게 한다. 오차 분포가 정규분포를 따를 때, RMSE는 오차의 표준편차와 유사하게 해석될 수 있다. 평균 제곱근 오차 식은 식 (2)와 같다.

$$\mathrm{SE} = \sqrt{\frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}(\mathrm{y_i} - \hat{\mathrm{y_i}})^2} \tag{2}$$

R^2(결정계수) 점수는 회귀 모델의 예측이 실제 데이터를 얼마나 잘 설명하는지 나타내는 지표이다. R^2 점수는 0에서 1 사이의 값을 가지며, 1에 가까울수록 모델이 데이터를 잘 설명한다고 평가할 수 있다. R^2 계산 공식은 식 (3)와 같다.

$$R^2 = 1 - \frac{RSS}{TSS} \tag{3}$$

여기서 잔차 제곱합(RSS, Residual Sum of Squares)는 모델의 예측값($\hat{y_i}$)과 실제값(y_i)의 차이(잔차)를 제곱하여 합한 값입니다. 잔차가 작을수록 모델이 데이터를 더 잘 예측하며 식 (4)와 같다.

$$RSS = \sum_{i=1}^{n}(y_i - \hat{y_i})^2 \tag{4}$$

TSS (Total Sum of Squares, 총 제곱합)는 실제값(y_i)과 실제값의 평균($\bar{y}$)의 차이를 제곱하여 합한 값입니다. 이는 종속 변수(y)의 총 변동성을 나타내며 식 (5)와 같다.

$$TSS = \sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{5}$$

III. Adapted-Clear Sky Index

3.1 청천지수 정규화

청천지수(CSI)는 실측 발전량을 청천 기대 출력으로 정규화한 값이다[27]. 본 연구는 물리 기반 청천 모델 대신, 학습 구간의 통계로부터 청천 기준을 경험적으로 산출한다. 먼저 주간(Daytime) 표본을 대기권밖일사가 양수인 시점으로 정의하고, 대기권밖 일사를 $N_{bin} = 12$개의 분위수 구간(Radiation bin)으로 나눈다. 각 시각(Hour) h와 복사 구간 b의 조합에 대해, 학습 구간 발전량의 상위 분위수를 청천 기준으로 취하며 식 (6)과 같다.

$$P_{base}^{CS}(h,b) = Q_q(\{y_t : hour(t) = h, bin(t) = b, t \in T_{fit}\}) \tag{6}$$

여기서 $Q_q(\cdot)$는 분위수(Quantile)연산자, q=0.95는 청천 분위수(Clear-sky quantile)로 맑은 날의 포락선을 근사한다. T_{fit}는 학습 구간, $hour(t)$와 $bin(t)$는 각각 시점 t의 시각과 복사 구간이다. 특정 조합에 표본이 부족한 경우, 시각별 분위수 $Q_q(\{y_t : hour(t) = h\})$로 대체한다.

3.2 계절적응 청천 기준(M1)

M1은 청천 기준에 월별 계절 스케일을 반영하여 계절에 따른 대기 투과 특성 변화[26]를 보정하는 효과를 가진다. 월 m의 스케일 S_m은 학습 구간에서 발전량 대 구동변수 비율의 분위수로 정의되고, 평균으로 정규화한 뒤 경계로 제한하고 이를 청천 기준에 곱하여 계절에 적응시키는 방법이며, 식 (7)과 같다.

$$P_{adj}^{cs}(t) = \max\left(0, P_{base}^{cs}(t) \cdot S_m(t)\right) \tag{7}$$

3.3 강건 CSI 경계(M2)

CSI는 청천 기준이 작은 야간 및 박명 시각에 비정상적으로 큰 값을 가질 수 있다. M2는 고정 상한(예: 1.5) 대신 학습 구간의 주간 CSI 분포로부터 상한 $k_{\max}$를 학습 시키며, 식 (8)과 같다.

$$k_{\max} = Q_{0.995}\left(\left\{\frac{y_t}{P_{adj}^{cs}} : P_{adj}^{cs}(t) > \tau, t \in T_{fit}\right\}\right) \tag{8}$$

여기서 $Q_{0.995}$는 99.5 백분위수로 극단값을 배제한 강건한 상한을 제공하며, τ는 주간 판별 임계값(Clear-sky floor)이다. 본 데이터에서 $k_{\max} \approx 1.425$로 학습되었다. 임계값 τ는 학습 구간 발전량 중앙값의 5%로 설정하며 식 (9)와 같다.

$$\tau = \max\left(1.0, 0.05 \cdot Q_{0.5}(\{y_t : t \in T_{fit}\})\right) \tag{9}$$

정규화된 청천지수는 식 (10)과 같이 상한 제한되어 예측 목표가 된다.

$$k_t = clip\left(\frac{y_t}{P_{adj}^{cs}(t)}, 0, k_{\max}\right) \tag{10}$$

이 방법의 독립적 효과는 예측의 안정성(분산 감소)에 주로 작용하며, 평균 RMSE보다는 시드 간 편차와 최댓값 오차를 줄이는 방향으로 나타난다.

3.4 주간 한정 혼합 복원(M3)

모델이 예측한 청천지수 $\hat{k}_t$로부터 발전량을 복원할 때 M3는 곱셈 복원과 가법 잔차 보정을 혼합한다. 이것의 독립적 효과는 주간 시간대의 편향(Bias) 감소에 집중되며, 야간 0값을 보존하여 MAPE 왜곡을 방지한다. 먼저 곱셈 복원은 식 (11)과 같다.

$$\hat{y}_t^{mult} = \max(0, \hat{k}_t \cdot P_{adj}^{CS}(t)) \tag{11}$$

곱셈 복원만으로는 청천 기준의 계통 오차가 그대로 전이되므로, 별도의 잔차 예측 모델(Residual head)이 잔차를 예측하며, 잔차 r_t는 식 (12)와 같다.

$$r_t = y_t - \hat{y}_t^{mult} \tag{12}$$

이 가법 보정은 주간에만 적용하여 야간의 0 값을 보존한다(야간 보정 시 MAPE 폭증 방지). 식 (12)의 곱셈 복원 치우침을 잔차로 교정하되, 그 교정이 야간의 0 값을 훼손하지 않도록 주간에만 적용하여 물리적 타당성과 평가 안정성을 동시에 확보하기위해 식 (13)을 적용한다.

$$\hat{y}_t^{add} = \max\left(0, \hat{y}_t^{mult} + \hat{r}_t \cdot 1\left[P_{adj}^{CS}(t) > \tau\right]\right) \tag{13}$$

여기서 $1[\cdot]$는 지시함수(Indicator function)로, 청천 기준이 임계값을 초과하는 주간에는 1, 그 외 야간에는 0이다. 최종 예측은 두 복원의 볼록 결합으로써 식 (14)와 같이 나타낸다.

$$\hat{y}_t = \max\left(0, \alpha\, \hat{y}_t^{mult} + (1-\alpha)\, \hat{y}_t^{add}\right) \tag{14}$$

여기서 $\alpha = 0.6$은 혼합 계수(Mixing coefficient)로, 곱셈 복원과 가법 보정의 상대 비중을 조절한다. 정규화 없음(Basic)은 y_t를 직접 예측하고, 기본 CSI는 식 (6)의 P_{base}^{CS}만 사용하며, 제안된 알고리즘 A-CSI는 (M1+M2+M3)를 모두 적용한 구성이다.

Ⅳ. 실험 및 결과

4.1 데이터 구성 및 전처리

본 연구는 대한민국 제주도(북위 33.27°, 동경 126.33°)에 위치한 상용 태양광 발전소의 시간 단위(Hourly) 데이터를 사용한다. 전체 기간 중 최근 6년(52,560시간)을 모델링에 활용한다. 목표 변수는 시간당 발전량 Y_t(Solar generation, MWh)이며, 입력은 표 1의 10개 기상 변수(지표면일사, 대기권밖일사, 일조시간, 전운량(10분위), 중하층운량(3분위), 기온, 습도, 강수량, 풍속, 적설)로 구성된다.

결측치는 시계열 선형보간으로 우선 채우고, 잔여 결측은 0으로 대체한다. 입력 변수와 목표 변수 간의 선형 관계는 피어슨 상관계수(PCC, Pearson Correlation Coefficient)로 참고 확인하며, 대기권밖일사(Extra solar radiation)가 발전량의 주요 구동 x변수(Driver)로 사용된다.

데이터는 시간순으로 분할하여, 마지막 예측 지평(24시간과 168시간)을 테스트 구간으로, 그 이전을 학습과 검증 구간으로 사용한다. 학습과 검증 구간은 다시 시간순 90 : 10으로 나눈다. 이러한 시간순 분할은 미래 정보가 과거 학습에 유입되는 정보 누출(Data leakage)을 방지한다.

표 1. 입력 특징 변수(10가지)
Table 1. Input feature variables(10 types)

Items	Features
Radiation	Surface_solar_irradiance, extra_solar_radiation, sunshine,
Cloud cover	cloud(10th), cloud(3rd)
Atmosphere	temperature, humidity, snow, precipitation, wind_speed,

모델 학습에 사용한 하드웨어는 Intel(R) Core(TM) 3.0GHz CPU와 NVIDIA GeForce RTX 3070 GPU이다. LSTM은 은닉 차원 64·32의 2개 LSTM 층에 Dropout(0.2)과 완전 연결층(16→1)을 연결하고, CNN-LSTM은 그 앞에 Conv1D(필터 64, 커널 3)와 MaxPooling1D(2)를 추가하였다. 두 신경망은 모두 Adam 최적화기(학습률 1×10^{-3}), MSE 손실, 배치 64, 최대 200 epoch로 학습하되, 검증 손실 기준 조기 종료(patience=15)를 적용하였다. N-HiTS는 3개 stack(각 block 1개, MLP 유닛 [256,256]), 풀링 커널 [2,2,1], 주파수 다운샘플 [4,2,1], 선형 보간으로 구성하고, Adam(1×10^{-3})으로 최대 500 step 학습하였다. LightGBM과 XGBoost는 $n_{est} = 600$, 학습률 0.03, subsample·colsample_bytree=0.8을 공통으로 두되, LightGBM은 num_leaves=63, min_child_samples=30, XGBoost는 max_depth=6, min_child_weight=5로 설정하였다. 두 부스팅 모델 모두 검증셋 기준 조기 종료(Early stopping rounds=40)를 적용하였다.

4.2 실험 설계

정규화의 조건부 효과를 규명하기 위해, 네 개 수준의 전처리를 5개 모델과 2개 예측 지평에 걸쳐 교차 평가하였다(총 4 × 5 × 2 = 40개 조건). 네 개 전처리 수준은 다음과 같다.

- Basic: 정규화 없음, 발전량 직접 예측(기준선)
- CSI: 기본 청천지수
- M1: 계절 적응 CSI, 월별 스케일 반영
- A-CSI: 계절 적응, 강건 경계 및 혼합복원 결합

평가는 RMSE, MAE, R^2 로 수행하되, 큰 오차에 민감하여 예측 안정성을 대표하는 RMSE를 중심으로 분석한다.

제안된 알고리즘 A-CSI의 작동 원리는 아래와 같다.

A-CSI()는 태양광 출력의 결정론적 성분(태양 기하)과 확률적 성분(기상)을 분리하기 위한 청천지수 정규화의 확장 기법이다. 기본 청천지수가 고정된 청천 기준과 단순 곱셈 복원을 사용하는 것과 달리, A-CSI는 세 구성요소(M1, M2, M3)를 결합하여 계절 변동·극단값·복원 오차를 순차적으로 보정한다. 전체 작동은 학습 단계(청천 기준·경계 산출)와 예측 단계(정규화·학습·복원)로 나눈다.

제안된 A-CSI 알고리즘의 전체 흐름은 그림 1과 같다. 순서는 실측 발전량 y_t를 청천지수 k_t로 정규화(M1, M2)하여 모델이 기상 잔차만 학습하게 하고, 예측된 $\hat{k}_t$를 발전량으로 복원(M3)한다.

4.3 실험 결과

표 2와 표 3은 24시간과 168시간 예측에 대한 평가 지표(RMSE, MAE, R2) 실험 결과를 나타내었다.

4.3.1 24시간 예측 분석

24시간 예측 결과의 모델에 대한 베이스라인, CSI, M1, A-CSI에 대한 분석은 아래와 같으며, 네 가지 방법에 대한 개선율을 그림 1에 나타내었다.

$$y_t \xrightarrow[\text{M1, M2}]{\text{Normalization}} k_t \xrightarrow{\text{Model Training}} \hat{k}_t \xrightarrow[\text{M3}]{\text{Recovery}} \hat{y}_t$$

그림 1. 제안된 알고리즘 흐름도(A-CSI)
Fig. 1. Flowchart of the proposed algorithm(A-CSI)

1) Basic↔CSI: 기본 CSI는 부스팅 트리에서 뚜렷한 개선을 보였으나(LightGBM +10.8%, XGBoost +13.4%), LSTM(−25.4%)과 N-HiTS(−11.4%)에서는 악화되어 평균 −1.3%에 그쳤다. R^2도 평균 −0.0012로 변화가 미미하다. 단순 CSI는 트리 계열에만 선택적으로 유효하다.

2) Basic↔M1: 계절 적응(M1)은 5개 중 4개 모델을 개선하며 평균 +11.9% 향상, R^2 +0.0111을 달성하였다. 특히 N-HiTS(+25.5%)와 CNN-LSTM(+20.1%) 등 딥러닝 계열에서 두드러졌으며, LightGBM만 −3.0%로 소폭 악화되었다.

3) Basic↔A-CSI: A-CSI는 5개 모델 전부를 개선(5/5)한 유일한 방법으로, 평균 +10.0%, R^2 +0.0098을 기록하였다. M1(+11.1%)보다 평균 개선율은 낮으나 모든 모델에서 일관되게 개선한다는 점에서 강건하다. M1에서 악화되던 LightGBM이 A-CSI에서 +0.7%로 개선 전환된 것은 M2와 M3의 결합 효과이다.

4) CSI↔A-CSI: A-CSI는 단순 CSI 대비 평균 +8.6% 개선했으나 계열에 따라 방향이 갈렸다. 딥러닝은 크게 개선(N-HiTS +34.2%, LSTM +27.8%)된 반면 부스팅 트리는 악화(LightGBM −11.4%, XGBoost −10.3%)되었다.

표 2. 24시간 예상 실험 결과
Table 2. Result of 24-hour forecast

Methods	Basic			CSI			CSI-M1			A-CSI		
	RMSE	MAE	R2	RMSE	MAE	R2	RMSE	MAE	R2	RMSE	MAE	R2
LSTM	15.85	9.03	0.97	19.88	10.05	0.95	14.32	8.28	0.97	14.36	8.46	0.97
CNN-LSTM	21.53	10.95	0.94	20.22	11.53	0.95	17.20	10.29	0.96	19.67	11.18	0.95
LightGBM	17.95	13.20	0.96	16.01	8.57	0.97	18.49	10.96	0.96	17.83	10.37	0.96
XGBoost	19.25	13.94	0.95	16.67	9.12	0.96	18.65	10.92	0.95	18.39	10.76	0.96
N-HiTS	21.71	13.60	0.94	24.19	12.60	0.92	16.18	9.21	0.97	15.92	8.73	0.97

표 3. 168시간 실험 결과
Table 3. Result of 168-hour forecast

Methods	Basic			CSI			CSI-M1			A-CSI		
	RMSE	MAE	R2	RMSE	MAE	R2	RMSE	MAE	R2	RMSE	MAE	R2
LSTM	29.87	15.40	0.80	29.60	15.22	0.81	25.40	13.05	0.86	25.81	12.92	0.85
CNN-LSTM	26.92	13.24	0.84	23.65	10.90	0.88	23.74	11.24	0.88	22.42	10.44	0.89
LightGBM	25.99	17.41	0.85	24.42	11.87	0.87	24.44	12.67	0.87	24.53	12.64	0.87
XGBoost	24.96	16.67	0.86	24.85	12.17	0.86	24.89	12.90	0.86	24.74	12.77	0.86
N-HiTS	17.89	10.17	0.93	18.44	8.83	0.92	31.31	15.66	0.78	31.23	15.99	0.78

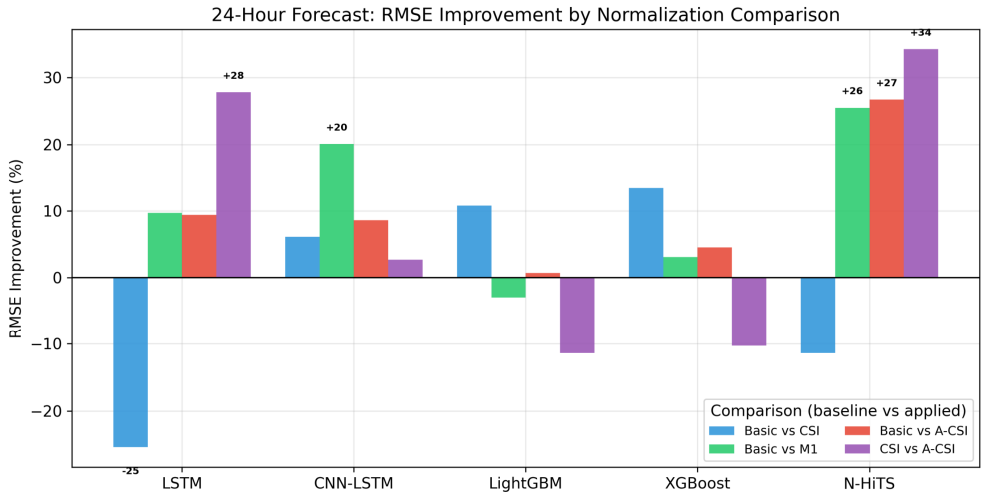


그림 2. 24시간 RMSE 개선: 정규화를 적용한 모델별 RMSE 개선율
Fig. 2. 24-Hour RMSE improvement: RMSE improvement rate by model with normalization applied

그림 2는 24시간 예측의 정규화 비교별 RMSE 개선율을 나타낸다. 여기서 양수는 개선을 음수는 악화를 의미한다.

24시간 실험에서, 단순 CSI는 효과가 제한적(−1.3%)이나 M1(+11.1%)과 A-CSI(+10.0%)는 큰 개선을 이루어, 계절 적응이 개선의 원천임을 보여준다. 특히 제안된 A-CSI는 유일하게 전 모델을 개선하여 가장 강건하며, 전체 최저 RMSE는 LSTM+M1(14.32)에서 관측되었다. 딥러닝 계열이 A-CSI 및 M1의 최대 수혜자이다.

4.3.2 168시간 예측 분석

168시간 예측 결과의 모델에 대한 베이스라인, CSI, M1, A-SCI에 대한 분석은 아래와 같으며, 네 가지 방법에 대한 개선율을 그림 3에 나타내었다.

(1) Basic↔CSI: 단순 CSI는 4/5 모델에서 개선(평균 +3.3%, R^2 +0.0108)되어 24시간보다 안정적이었다. CNN-LSTM(+12.1%)이 가장 크게 개선되었고, 최저 오차 모델인 N-HiTS만 소폭 악화(−3.1%)되었다.

(2) Basic↔M1: LSTM(+15.0%) 등 4개 모델은 개선되었으나, N-HiTS가 −75.1%(R^2 0.93→0.78) 붕괴하여 평균이 −8.4%로 하락하였다.

(3) Basic↔A-CSI: A-CSI도 N-HiTS를 제외한 4개 모델은 개선(CNN-LSTM +16.7%, LSTM +13.6%)되나 N-HiTS 붕괴(−74.6%)로 평균 −7.6%였다. 다만 A-CSI의 N-HiTS 악화는 M1보다 미세하게 완화되었고, CNN-LSTM은 A-CSI(+16.7%)가 M1(+11.8%)보다 우수하여 M2와 M3의 부분적 보정 효과가 확인된다.

(4) CSI↔A-CSI: A-CSI는 단순 CSI 대비 평균 −10.3%로 열등하였다. LSTM(+12.8%)은 개선되나 N-HiTS 붕괴(−69.4%)가 결정적이어서, 장기 지평에서는 단순 CSI가 안전한 선택이다.

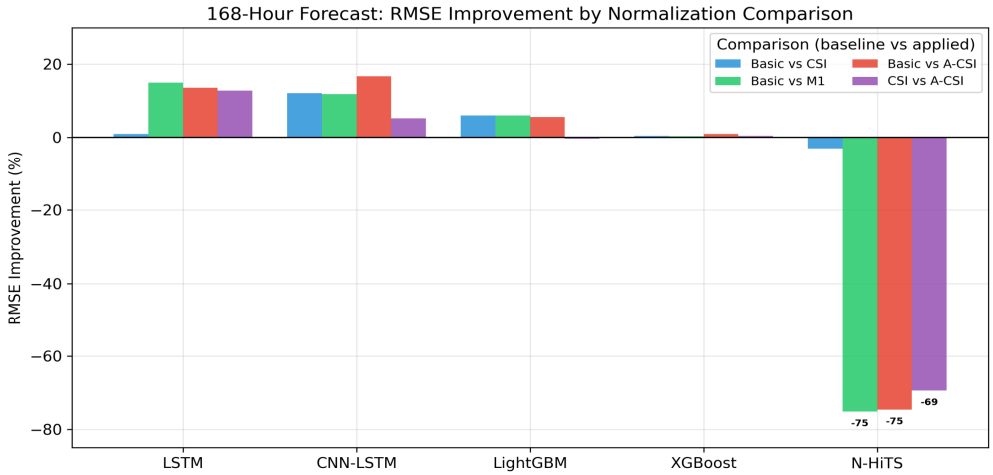


그림 3. 168시간 RMSE 개선: 정규화를 적용한 모델별 RMSE 개선율
Fig. 3. 168-Hour RMSE improvement: RMSE improvement rate by model with normalization applied

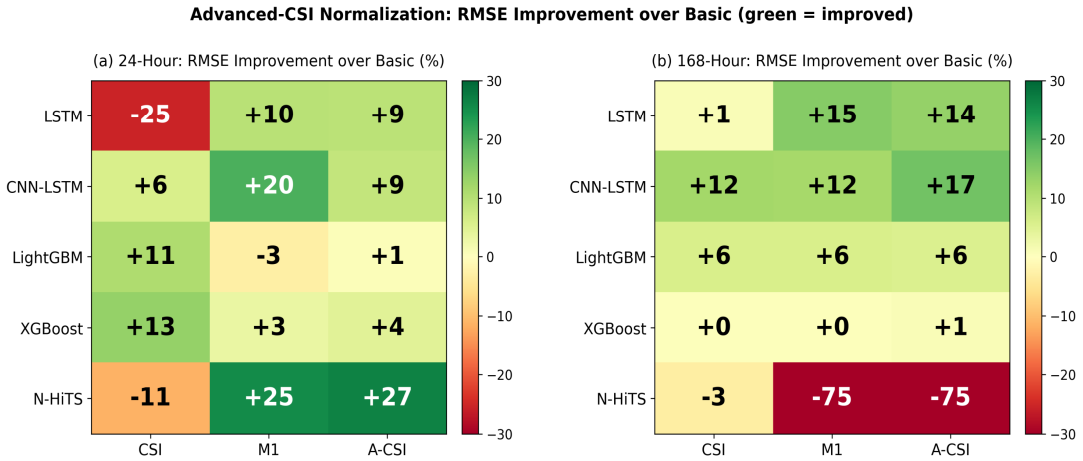


그림 4. 정규화에 따른 RMSE 개선율(%) (a) 24시간, (b) 168시간 (녹색: 개선, 빨간색: 악화)
Fig. 4. RMSE improvement rate (%) by normalization (a) 24 hours, (b) 168 hours (Green: improvement, red: deterioration

168시간 실험결과는 24시간과 정반대로, 단순 CSI만 평균 개선(+3.3%)을 이루었고 M1과 A-CSI는 평균 악화(−8.4%, −7.6%)되었다. 그러나 이는 방법의 결함이 아니라 N-HiTS 단일 모델의 붕괴가 평균을 왜곡한 결과이다. N-HiTS를 제외하면 M1·A-CSI도 나머지 4개 모델에서 +0.3~16.7% 개선된다. 모델별로는 LSTM·CNN-LSTM이 A-CSI에서, N-HiTS가 Basic에서, 부스팅이 CSI에서 최적이다.

N-HiTS는 본 실험에서 지평에 따라 극단적으로 상반된 거동을 보였다. 24시간에서는 M1(+25.5%)과 A-CSI(+26.7%)로 가장 큰 개선을 보인 반면, 168시간에서는 M1·A-CSI 적용 시 약 75% 붕괴(RMSE 17.89→31.31, R^2 0.93→0.78)하였다. 이는 N-HiTS의 구조적 특성에서 기인한다. N-HiTS는 다중 해상도의 계층적 보간(hierarchical interpolation)으로 예측 구간 전체를 한 번에 생성하는데, M1의 월별 계절 스케일은 학습 기간의 계절 패턴에 적합된 값이다. 24시간처럼 짧은 구간에서는 계절 스케일이 국소적으로 일관되어 보간 구조와 조화를 이루지만, 168시간(7일)처럼 긴 구간에서는 보간이 생성하는 매끄러운 다중스텝 궤적과 월별로 이산적으로 변하는 계절 스케일이 충돌한다.

그 결과 복원 과정에서 오차가 누적·증폭되어 붕괴가 발생한다. 반면 순환 구조의 LSTM·CNN-LSTM은 스텝별로 순차 예측하므로 계절 스케일과의 충돌이 적어 168시간에서도 A-CSI의 이득(+13.6%, +16.7%)을 유지한다. 따라서 N-HiTS에는 장기 지평에서 M1 계열 정규화를 적용하지 않

고 정규화 없음(Basic) 또는 단순 CSI를 사용할 것을 권장한다. 이는 정규화 기법을 모델의 예측 생성 방식에 맞추어 선택해야 함을 보여주는 대표적 사례이다.

4.3.3 A-CSI의 계열별 효과

그림 4는 정규화 없음 대비 각 정규화의 RMSE 개선율을 보여준다(녹색=개선). 24시간 예측에서 A-CSI의 계절 적응(M1)은 딥러닝 계열에서 큰 개선을 보였다. N-HiTS는 25.5%, CNN-LSTM은 20.1%, LSTM은 9.7% 개선되었다. 이는 계층적 보간과 순환 구조를 가진 딥러닝 모델이 결정론적 성분 제거의 이득을 크게 받음을 시사한다. 한편 부스팅 트리(LightGBM, XGBoost)는 기본 CSI에서 각각 10.8%, 13.4%의 안정적 이득을 보였다.

4.3.4 모델·지평별 최적 정규화

최적 정규화는 모델과 예측 지평에 따라 달랐다. 24시간에서 순환·딥러닝 계열(LSTM, CNN-LSTM, N-HiTS)은 계절 적응(M1) 또는 A-CSI에서, 부스팅 계열(LightGBM, XGBoost)은 기본 CSI에서 최적이었다. 168시간에서는 CNN-LSTM이 A-CSI(22.42), LSTM이 M1(25.40)에서 우수하였으나, N-HiTS는 정규화 없음(17.89)에서 최적을 보였다. 이는 N-HiTS의 계층적 보간이 장기 지평에서 계절 스케일과 충돌할 수 있음을 시사하며, 정규화 방법을 모델과 지평에 맞추어 선택해야 함을 보여준다.

종합하면, A-SCI 정규화는 특히 딥러닝 계열의 24시간 단기 예측에서 정확도를 크게 향상시켰다. 계절 적응(M1)이 개선의 핵심이었으며, 강건 경계·혼합 복원(M2·M3)을 결합한 A-CSI는 안정적 성능을 유지하였다.

4.4 종합 분석

표 4는 네 비교의 지평별 평균 개선율을 요약한다. 첫째, A-CSI 알고리즘은 24시간 단기 예측에서 매우 효과적이며(평균 +10.0%), 유일하게 5개 모델 전부를 개선하여 M1(4/5)보다 강건하다. 둘째, 계절 적응(M1)이 개선의 원천으로, 단순 CSI(24h −1.3%)와의 격차가 이를 입증한다. 셋째, 효과는 모델 계열에 의존하여 딥러닝은 A-CSI·M1에서, 부스팅 트리는 단순 CSI에서 우수하다. 넷째, 효과는 예측 지평에 의존하며 24시간과 168시간이 반대인데, 이는 계절 스케일의 장기 외삽 위험 때문이다. 다섯째, R^2 변화는 모든 비교에서 RMSE 개선율과 방향이 일치하여 두 지표가 상호 검증된다.

실무적으로, 24시간 예측에서는 딥러닝에 A-CSI 또는 M1 방법을, 부스팅 트리에 기본 CSI를 권장한다. 168시간 예측에서는 순환 딥러닝(LSTM, CNN-LSTM)에 A-CSI를, 부스팅 트리에 기본 CSI를, N-HiTS에는 정규화 없음 또는 단순 CSI를 권장한다.

V. 결론 및 향후 과제

본 연구는 기존 청천지수를 계절 적응(M1), 강건 경계(M2) 그리고 혼합 복원(M3)으로 확장한 Adapted-CSI 알고리즘을 제안하고, 제주지역 태양광 발전소의 6년 데이터에 적용하여 그 효과를 검증하였다. 네 가지 정규화를 다섯 모델과 두 예측 지평에 걸쳐 비교한 결과, A-CSI 정규화는 정규화 없음 대비 예측 정확도를 향상시켰다.

표 4. 비교 대상 및 기간별 평균 RMSE 개선율(%) 및 R^2 변화.
Table 4. Average RMSE improvement (%) and R^2 change by comparison and horizon.

Comparison	24h ΔRMSE	24h ΔR^2	168h ΔRMSE	168h ΔR^2
Basic → CSI	-1.3%	-0.001	+3.3%	+0.011
Basic → M1	+11.9%	+0.011	-8.4%	-0.008
Basic → A-CSI	+10.0%	+0.010	-7.6%	-0.006
CSI → A-CSI	+8.6%	+0.011	-10.3%	-0.016

네 가지 정규화를 다섯 모델과 두 예측 지평에 걸쳐 비교한 결과, A-CSI는 정규화 없음 대비 예측 정확도를 향상시켰으며, 그 효과는 모델 계열과 예측 지평에 따라 달랐다. 계절을 반영하는 딥러닝 계열은 A-CSI에서, 부스팅 트리 계열은 기본 CSI에서 우수하여, A-CSI는 특히 딥러닝 기반 단기(24시간) 예측에 효과적임을 확인하였다. 이는 실무자가 모델과 예측 과제에 맞는 정규화를 선택하는 데 지침을 제공한다.

한편, 본 연구는 몇 가지 한계를 가진다. 첫째, 계층적 보간 구조의 N-HiTS는 장기(168시간) 예측에서 계절 적응 기반 정규화 적용 시 성능이 오히려 저하되어, A-CSI가 모든 모델·지평에 보편적으로 유효하지는 않음을 시사한다. 둘째, 단일 지역(제주)의 데이터만을 사용하였으므로 기후·설비 조건이 다른 지역으로의 일반화에는 추가 검증이 필요하다. 향후 연구에서는 다지점 데이터를 통한 일반화 성능 검증과 점 예측을 넘어선 확률 예측으로의 확장을 계획한다.

References

[1] P. D. Leo, A. Ciocia, G. Malgaroli, and F. Spertino, "Advancements and Challenges in Photovoltaic Power Forecasting: A Comprehensive Review", Energies, Vol. 18, No. 8, Art. no. 2108, Apr. 2025. https://doi.org/10.3390/en18082108.

[2] B. Sivadharshini, A. Saxena, G. Saritha, G. Saritha, A. Alawady, S. Lall, and D. Dhabliya, "Solar Energy Forecasting Models for Grid Integration and Power Balancing", E3S Web of Conferences, Vol. 540, Art. no. 10022, Jun. 2024. https://doi.org/10.1051/e3sconf/202454010022.

[3] E. Nejla, K. Saidani, and M. Besbes, "Overcoming the integration bottleneck: a global review of renewable energy and grid adaptation strategies", International Journal of Sustainable Energy, Vol. 44, Art. no. 2569922, Nov. 2025. https://doi.org/10.1080/14786451.2025.2569922.

[4] A. Nadeem, M. Hanif, M. U. Naveed, M. T. Hasan, M. Gul, and A. Zafar, "AI-driven precision in solar forecasting: Breakthroughs in machine learning and deep learning", AIMS Geosciences, Vol. 10, No. 4, pp. 684-734, 2024. https://doi.org/10.3934/geosci.2024035.

[5] G. Wu, Y. Wang, Q. Zhou, and Z. Zhang, "Enhanced Photovoltaic Power Forecasting: An iTransformer and LSTM-Based Model Integrating Temporal and Covariate Interactions", 2024 IEEE the 8th Conference on Energy Internet and Energy System Integration, Shenyang, China, pp. 856-861, Nov.-Dec. 2024. https://doi.org/10.1109/EI264398.2024.10990816.

[6] Y. Sun, V. Venugopal, and A. R. Brandt, "Short-term solar power forecast with deep learning: Exploring optimal input and output configuration", Solar Energy, Vol. 188, pp. 730-741, Aug. 2019. https://doi.org/10.1016/j.solener.2019.06.041.

[7] D. Yang, "Choice of clear-sky model in solar forecasting", Journal of Renewable and Sustainable Energy, Vol. 12, No. 2, Art. no. 026101, Mar. 2020. https://doi.org/10.1063/5.0003495.

[8] Y. Cui, P. Wang, J. F. Meirink, N. Ntantis, and J. S. Wijnands, "Solar radiation nowcasting based on geostationary satellite images and deep learning models", Solar Energy, Vol. 282, Art. no. 112866, Nov. 2024. https://doi.org/10.1016/j.solener.2024.112866.

[9] R. A. Silva, L. B. Cesar, M. Á. M. Callejo, and C. Cira, "Impact of Stationarizing Solar Inputs on Very-Short-Term Spatio-Temporal Global Horizontal Irradiance (GHI) Forecasting", Energies, Vol. 17, No. 14, Art. no. 3527, Jul. 2024. https://doi.org/10.3390/en17143527.

[10] P. Lauret, R. Alonso-Suárez, J. Salle, and M. David, "Solar Forecasts Based on the Clear Sky Index or the Clearness Index: Which Is Better?", Solar, Vol. 2, No. 4, pp. 432-444, Oct. 2022. https://doi.org/10.3390/solar2040026.

[11] L. Varaschin and D. Silva, "Benchmarking Deep Learning Methods for Irradiance Estimation from Sky Images with Applications to Video Prediction-Based Irradiance Nowcasting", Solar Energy, Vol. 302, Art. no. 114000, Dec. 2025. https://doi.org/10.1016/j.solener.2025.114000.

[12] C. Voyant, M. Despotovic, G, Notton, Y, Saint-Drenan, and M, Asloune, "On the Importance of Clearsky Model in Short-Term Solar Radiation Forecasting", Solar Energy, Vol. 294, Art. no. 113490, Jul. 2025. https://doi.org/10.1016/j.solener.2025.113490.

[13] H. Mauladdawilah, E. J. Gago, H. Balfaqih, and M. C. Pegalajar, "Towards energy efficiency: A comprehensive review of deep learning-based photovoltaic power forecasting strategies", Heliyon, Vol. 10, No. 13, Art. no. e33419, Jul. 2024. https://doi.org/10.1016/j.heliyon.2024.e33419.

[14] R. Herrera-Casanova, A. Conde, and C. Santos-Pérez, "Hour-Ahead Photovoltaic Power Prediction Combining BiLSTM and Bayesian Optimization Algorithm, with Bootstrap Resampling for Interval Predictions", Sensors, Vol. 24, No. 3, Art. no. 882, Jan. 2024. https://doi.org/10.3390/s24030882.

[15] X. Guo, Y. Mo, and K. Yan, "Short-term photovoltaic power forecasting based on historical information and deep learning methods", Sensors, Vol. 22, No. 24, Art. no. 9630, Dec. 2022. https://doi.org/10.3390/s22249630.

[16] A. Agga, A. Abbou, M. Labbadi, Y. Houm, and I. H. Ali, "CNN-LSTM: An Efficient Hybrid Deep Learning Architecture for Predicting Short-Term Photovoltaic Power Production", Electric Power Systems Research, Vol. 208, Art. no. 107908, Jul. 2022. https://doi.org/ 10.1016/j.epsr.2022.107908.

[17] A. Jakoplić, D. Franković, J. Havelka, and H. Bulat, "Short-Term Photovoltaic Power Plant Output Forecasting Using Sky Images and Deep Learning", Energy, Vol. 16, No. 14, pp. 1-18, Jul. 2023. https://doi.org/10.3390/en16145428.

[18] Y. Yang, B. Ma, and P. Li, "VMD-LSTM-Based Model Predictive Control for Hybrid Energy Storage Systems with Auto-Tuning Weights and Constraints", Energies, Vol. 18, No. 21, pp. 1-23, Oct. 2025. https://doi.org/10.3390/en18215559.

[19] C. Challu, K. G. Olivares, B. N. Oreshkin, F. Garza, M. Mergenthaler-Canseco, and A. Dubrawski, "N-HiTS: Neural hierarchical interpolation for time series forecasting", Proc. AAAI Conf. Artificial Intelligence, Online, Vol. 37, No. 6, pp. 6989-6997, Feb. 2023 https://doi.org/10.1609/aaai.v37i6.25854.

[20] W. Peng, T. He, Y. Ma, and Y. Zheng, "A physics-informed machine learning method for short-term solar radiation forecast with satellite-derived cloud optical depth", Applied Energy, Vol. 406, Art. no. 127284, Mar. 2026. https://doi.org/10.1016/j.apenergy.2025.127284.

[21] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system", Proc. 22nd ACM SIGKDD Int. Conf. KDD, pp. 785-794, Aug. 2016. https://doi.org/10.1145/2939672.29397.

[22] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree", Advances in Neural Information Processing Systems 30, Long Beach, CA, USA, Vol. 30, pp. 3149-3157, Dec. 2017.

[23] L. Breiman, "Random Forests, Machine Learning", Machine Learning, Vol. 45, No. 1, pp. 5-32, Oct. 2001. https://doi.org/10.1023/A:1010933404324.

[24] L. Grinsztajn, E. Oyallon, and G. Varoquaux, "Why do tree-based models still outperform deep learning on typical tabular data?", arXiv preprint, arXiv:2207.08815, Jul. 2022. https://doi.org/10.48550/arXiv.2207.08815.

[25] F. Robrini and B. Amrouche, "Exploring Forecasting and Prediction Processes for

Decision-Making to Promote the Photovoltaic Energy Integration into the Grid: A Mini Review", 2023 International Conference on Decision Aid Sciences and Applications, Annaba, Algeria, pp. 180-186, Sep. 2023. https://doi.org/10.1109/DASA59624.2023.10286623.

[26] Y. H. Lee, "A Study on Solar Power Generation Forecasting using Weighting Schemes based on Meteorological Data Correlation", Journal of KIIT, Vol. 24, No. 4, pp. 27-38, Apr. 2026. http://dx.doi.org/10.14801/jkiit.2026.24.4.27.

[27] M. Renner, A. Kleidon, and M. Wild, "Estimating Shortwave Clear-Sky Fluxes From Hourly Global Radiation Records by Quantile Regression", Earth and Space Science, Vol. 6, No. 10, pp. 1532-1546, Aug. 2019. https://doi.org/10.1029/2019EA000686.

저자소개

이 영 학 (Yeung-Hak Lee)



2003년 8월 : 영남대학교 전자공학과(박사)
1991년 1월 ~ 1995년 9월 : LG정밀(주) 용인연구소 주임연구원
2005년 10월 ~ 2006년 10월 : 학술진흥재단 박사후 연구원(Cardiff Univ.)
2010년 3월 ~ 2017년 2월 : 경운대학교 항공전자공학과 부교수
2019년 9월 ~ 2019년 9월 : 국립안동대학교 SW융합교육원 교수
2025년 3월 ~ 현재 : 국립경국대학교 교양교육원 교수
2025년 9월 ~ 현재 : ㈜기븐텍 연구소장
관심분야 : 인공지능, 영상처리, 패턴 및 생체 인식, 로봇비전, 뉴럴네트워크

심 재 창 (Jae-Chang Shim)



1994년 8월 : 경북대학교 전자공학과(박사)
1997년 2월 ~ 1998년 2월 : IBM. T. J Watson 연구소, 국비 박사후 연구원
1999년 2월 : IBM. T. J Watson 연구소 AI팀 연구원
2005년 12월 ~ 2007년 2월 : 프린스턴 대학교 전자공학과 Visiting Fallow Professor
2021년 2월 ~ 2022년 2월 : 네브레스카 대학교 컴퓨터학과 Visiting Professor
1994년 3월 ~ 현재 : 국립경국대학교 컴퓨터공학과 교수
관심분야 : 영상처리, 패턴인식, 컴퓨터비전, AI, 임베디드 시스템, 소프트웨어 교육