Journal of Science and Technology – Smart Systems and Devices
Volume x, Issue x, Month Year, ppp–ppp

Short-Term Photovoltaic Power Forecasting for Photovoltaic Systems Using Temporal Convolutional Networks

Hung Nguyen Quang, Cuong Do Minh, Anh An Thi Hoai Thu*
Faculty of Electrical-Electronic Engineering,
University of Transport and Communications, Ha Noi, Vietnam
**Corresponding author email: htanh.ktd@utc.edu.vn*

Abstract

Short-term Photovoltaic (PV) power forecasting plays a pivotal role in ensuring the safe and stable operation and optimal dispatch of transmission power grids. This paper proposes a Temporal Convolutional Network (TCN) model integrated with a clear-sky index (K_t) modelling technique to forecast PV power output 15 minutes ahead for the Yulara solar power plant (Australia) with a rated capacity of 1,500 kW. A 10-year dataset (04/2016–02/2026) underwent a rigorous five-step preprocessing pipeline, including 5-minute time-frame synchronisation, physical boundary clipping, missing-data imputation, Ineichen astronomical modelling, and trigonometric encoding, yielding a 14-feature input space. Rather than directly predicting power output, which exhibits highly nonlinear fluctuations, the TCN forecasts the Kt index over a 2-hour historical sliding window via a 3-layer Temporal Block architecture featuring dilated causal convolutions, residual connections, and a Huber Loss function. Independent evaluation on 12,096 real-time samples from 2026 shows that the TCN achieves optimal overall performance, with an MAE of 38.94 kW, a 7.20% reduction in absolute error compared to the best-performing Machine Learning algorithm (CatBoost), and outperformance over recurrent time-series networks (LSTM and GRU). Real-time power profile analysis confirms that the TCN accurately reproduces clear-sky daily peak power (~1,224.25 kW) and captures sharp power drops induced by sudden cloud cover with near-zero phase lag, providing a reliable scientific foundation for renewable energy dispatching operations.

Keywords: Forecast, photovoltaic (PV) power output, temporal convolutional network (TCN).

1. Introduction

In recent years, renewable energy sources, particularly solar and wind power, have drawn increasing global attention because they can mitigate greenhouse gas emissions and meet growing energy demand. The rapid growth of these energy sources has driven a swift expansion of photovoltaic (PV) system installations worldwide. Notably, research by Alireza Bahadori and Chikezie Nwaoha showed that Australia has abundant solar energy resources, with an estimated annual incident solar radiation of approximately 58 million petajoules (PJ) across its territory, highlighting its potential for large-scale PV deployment [1].

However, unlike conventional power sources, the power output of photovoltaic (PV) systems directly depends on solar irradiance and meteorological conditions such as cloud cover, temperature, and other natural factors. Consequently, PV power output is highly volatile, particularly over short-term horizons. This intermittency complicates power system planning, dispatch, and operation, especially as PV penetration continues to rise [2]. Therefore, PV power forecasting, particularly short-term forecasting, plays a critical role in supporting stable power system operation and enhancing solar energy's grid integration capability.

In this context, numerous forecasting methods have been developed, ranging from statistical and regression methods to Machine Learning and Deep Learning approaches. While some studies focus on directly exploiting historical power or irradiance data, recent research trends favour incorporating additional information on weather conditions, sky images, and satellite imagery.

Hewage [3] developed a lightweight deep learning framework for fine-grained weather forecasting by investigating LSTM and Temporal Convolutional Network (TCN) models and comparing them with conventional machine learning methods, statistical forecasting approaches, a dynamic ensemble model, and the Weather Research and Forecasting (WRF) model. The proposed framework used 10 interrelated surface weather parameters and considered both multi-input multi-output and multi-input single-output forecasting strategies. The results showed that the proposed deep learning models could provide accurate short-and medium-term forecasts and achieved better overall performance than the WRF model for horizons up to 12 hours. However, the study used weather data from only January to May, and the authors noted that increasing the training dataset and incorporating additional atmospheric variables could further improve

p-ISSN 3093-3285
e-ISSN 3093-3315
https://doi.org/10.51316/jst.xxx.ssad.xxxx.xx.x.x
Received: Sep 28, 2026; Revised: Oct 5, 2026;
Accepted:Oct 7, 2026; Online: Nov 23, 2026

In press

the model, although this would also increase model complexity.

Building on this direction, Hewage [4] focused specifically on the Temporal Convolutional Network (TCN) for localised weather forecasting using time-series observations from a local weather station. The authors investigated TCN's ability to capture long-term temporal dependencies through causal and dilated convolutions and evaluated both MIMO-TCN and MISO-TCN configurations. MISO-TCN achieved strong short-term forecasting performance, while a fine-tuned TCN using weights learned from short-term forecasting further improved results for longer horizons of 2–24 hours. Nevertheless, the forecasting accuracy generally decreased as the prediction horizon increased, with larger errors observed for parameters such as wind direction, and the model was ultimately intended for localised weather forecasting rather than direct photovoltaic power prediction. In addition, the authors used only 93 days of training data, which they identified as a limitation that could be addressed with a larger dataset.

In addition, historical data-driven forecasting methods utilise solar irradiance or PV power values collected at consecutive time intervals to establish the relationship between past data and future target values [5–9]. Jin Xu [5] developed a Local Vector Autoregressive Ridge (LVARR) framework that exploits local spatio-temporal correlations among multiple irradiance sensors to forecast solar irradiance over 1–5 minute horizons, while using Ridge regularisation to mitigate overfitting. Sakshi Mishra [6] employed Recurrent Neural Networks (RNNs) for solar irradiance forecasting across multiple forecast horizons. According to Yujia Ge [7], a Long Short-Term Memory (LSTM) network was integrated with Empirical Mode Decomposition (EMD) to improve solar irradiance forecasting. Srivastava and Lessmann [8] evaluated LSTM's capability for day-ahead solar irradiance forecasting using multi-location datasets. Notably, R. Azimi [9] proposed a hybrid forecasting method that combines a transformation-based *K*-means clustering technique, time-series analysis, a cluster selection algorithm, and a multilayer perceptron neural network (MLPNN) to forecast solar radiation over horizons ranging from 1 to 48 hours. The proposed method was evaluated using multiple solar datasets with different radiation characteristics and showed improved forecasting performance compared with several established forecasting models. However, the method relies on a clustering-based preprocessing procedure and a fixed lagging window, and its forecasting performance decreases as the forecasting horizon lengthens, particularly under highly variable weather conditions. Generally, these methods leverage historical data without requiring complex data sources such as sky or satellite imagery. However, their forecasting performance can be limited when irradiance or PV power fluctuates rapidly due to abrupt weather changes.

To incorporate cloud status information, studies [10–16] used sky images combined with satellite imagery for cloud detection, classification, and feature extraction before forecasting solar irradiance. Clara Arbizu-Barrena [10] evaluated nowcasting methods based on sky images and satellite imagery. In [11], Quentin Paletta combined All-Sky images (ASIs) with satellite imagery to exploit local and wide-area information simultaneously. Wang and Zhiming Xuan [12] used real-time sky images for minute-level irradiance forecasting, while Lasanthika H. Dissawa [13] integrated sky images with cloud motion information for short-term irradiance forecasting. Convolutional Neural Networks (CNNs) were employed in the research by Cong Feng and Jie Zhang [14] to extract features from sky images for solar power forecasting. Furthermore, satellite imagery was used to forecast cloud cover by E. Nikitidou [15] and to classify clouds by Vesta Afzali Gorooh [16]. Although these methods incorporate spatial information that historical data cannot provide, a single image at a specific timestamp still struggles to fully capture temporal cloud dynamics.

Another approach involves modelling cloud motion. Methods such as Cloud Motion Vectors (CMVs), optical flow, and motion-tracking techniques are used to determine cloud displacement direction and speed from sky images or satellite imagery [10, 17–19]. Although these methods enable prediction of short-term variations in cloud cover and solar irradiance, they primarily focus on cloud displacement and still struggle to characterise cloud formation, growth, deformation, and dissipation.

Recently, researchers have investigated hybrid and multi-source data methods to exploit multiple information types simultaneously. Studies [17, 19–21] combine historical data with meteorological information, Numerical Weather Prediction (NWP) models, satellite imagery, or sky images. In [20], Johan Mathe employed a localised convolutional recurrent neural network architecture to simultaneously extract spatial and temporal information from NWP data for PV power forecasting. Hybrid approaches such as CNN-LSTM and Conv-LSTM also enable leveraging the feature extraction capabilities of CNNs alongside the temporal dependency learning abilities of LSTMs. Although they can exploit multiple information sources, these models typically have complex architectures and require larger input datasets.

In general, direct comparisons across existing studies remain challenging because of differences in input data, features, forecast horizons, weather conditions, and evaluation metrics. Consequently, a systematic evaluation of Deep Learning models under identical experimental conditions remains critical to clarify each

In press

algorithm's capabilities, advantages, and limitations for PV power forecasting.

On this basis, this study evaluates and compares Machine Learning and Deep Learning algorithms for short-term PV power forecasting. The algorithms are trained and evaluated using identical forecast horizons, evaluation metrics, and a shared power dataset from the Yulara solar power plant in Australia. Selecting a real-world PV system in Australia not only provides a dataset characterised by substantial solar irradiance but also lays the foundation for assessing the applicability of data-driven forecasting methods to PV systems under diverse climatic conditions. The findings can provide additional reference insights for developing PV power forecasting methodologies in Vietnam, where solar energy is playing an increasingly vital role in renewable energy expansion. Applying it to Vietnam would require local operational and meteorological data, along with recalibrating the clear-sky reference model to account for local climatic conditions.

2. Data and Methodology

The dataset used to evaluate the forecasting model was obtained from the Yulara solar power plant in Australia. The plant provides public access to download meteorological datasets measured from April 2016 to February 2026. The primary variables extracted from this site include timestamps, solar irradiance (W/m²), active power output to the grid (kW), ambient temperature (°C), solar panel temperature (°C), wind speed (m/s), and auxiliary solar panel temperature (°C).

Solar power forecasting models are currently developed primarily using deep learning methods based on historical datasets [5–9]. In this paper, we propose a Temporal Convolutional Network (TCN) algorithm for short-term power forecasting. Because of its fully parallel processing capability and effective handling of long-term sequences, the TCN algorithm is well suited for predicting future power output from extensive historical data. The proposed model framework consists of two main stages: data preprocessing and TCN algorithm execution. The mathematical notations, physical symbols, and units used in this paper are summarised in Table A1 (Appendix A).

2.1. Data Preprocessing and Input Generation

Feature extraction plays a decisive role in the TCN model's accuracy. After cleaning and reshaping the raw data, the model extracts features to select the optimal weather variables for the algorithm. This process consists of five main steps:

- Time-series restructuring and missing data handling.
- Astronomical modelling, date-time encoding and clear-sky index (K_t) calculation.
- Trigonometric encoding of temporal cycles.
- Sliding window partitioning into 3D Tensors and time-continuity verification.
- Time-series dataset splitting and Global 3D Z-Score normalisation.

In Step 2, the geographical parameters of the Yulara solar plant are configured with Latitude $\phi = -25.241^{\circ}$, Longitude $\lambda = 130.988^{\circ}$, and time zone Australia/Darwin. To account for solar position geometry and atmospheric optical attenuation, the Ineichen clear-sky optical model, implemented via the pvlib library, is used to compute key astronomical variables. Through this stage, a 14-feature input space is established and categorised into three core data groups: acquisition measurements, including actual active power output, actual solar irradiance, ambient temperature, PV module temperature, and wind speed; astronomical parameters, comprising:

Solar zenith angle: The angular distance between the direct solar ray path and the local vertical zenith line, reflecting the optical path length traversed by solar radiation through the atmosphere:

$$\cos(\theta_z) = \sin(\phi)\sin(\beta) + \cos(\phi)\cos(\beta)\cos(h_s) \quad (1)$$

where θ_z is the solar zenith angle ($^{\circ}$); ϕ is the geographical latitude of the Yulara solar plant ($\phi = -25.241^{\circ}$); β is the solar declination angle ($^{\circ}$), representing the tilt angle between the Earth-Sun vector and the equatorial plane; h_s is the solar hour angle ($^{\circ}$), quantifying the solar angular displacement relative to the local meridian.

Solar Declination Angle(β):

$$\beta = 23.45^{\circ} \times \sin\left(\frac{360^{\circ}}{N_{\text{days}}} \times (DOY - 81)\right) \quad (2)$$

where 23.45° is Earth's maximum axial tilt relative to the orbital plane normal, $N_{\text{days}} \in \{365, 366\}$ is the total days in the year, $DOY \in [1, 366]$ is the day of the year, and 81 denotes the vernal equinox (around March 21) when $\beta = 0^{\circ}$.

Solar Hour Angle (h_s):

$$h_s = 15^{\circ} \times (H(t) - 12) \quad (3)$$

where $H(t) \in [0, 24)$ is the fractional hour of the day and 15°/hour represents Earth's angular rotation velocity.

Solar elevation angle: The solar height above the horizontal horizon:

In press

$$\alpha_s = 90^{\circ} - \theta_z \tag{4}$$

where α_s determines the geometric incident angle of solar radiation on the PV module surface.

Theoretical cloud radiation: Derived via Haurwitz's optical formulation:

$$GHI_{\text{clear}}(t) = 1098 \times \cos(\theta_z) \times \exp(-0.057 \times m) \tag{5}$$

where 1098 W/m^2 is the theoretical extraterrestrial solar baseline at ground level, 0.057 is the optical attenuation coefficient, and $m \approx 1/\cos(\theta_z)$ is the relative optical air mass traversed by light rays.

Theoretical cloud power:

$$P_{\text{clear}}(t) = \text{Max}(P_{\text{active}}) \times \frac{GHI_{\text{clear}}(t)}{\text{Max}(GHI_{\text{clear}})} \tag{6}$$

where $\text{Max}(P_{\text{active}}) = 1500$ kW is the plant's rated peak capacity and $\text{Max}(GHI_{\text{clear}})$ is the peak theoretical clear-sky irradiance over the entire dataset.

To preserve continuous temporal transitions across midnight boundaries (23:55 to 00:00) and year-end boundaries (December 31 to January 1), integer temporal indices are mapped onto a continuous unit circle using trigonometric sine and cosine transformations:

Diurnal Cycle Encoding:

$$\text{Time_sin}(t) = \sin\left(\frac{2\pi \times M(t)}{1440}\right) \tag{7}$$

$$\text{Time_cos}(t) = \cos\left(\frac{2\pi \times M(t)}{1440}\right) \tag{8}$$

where $M(t) \in [0, 1439]$ represents the minute of the day ($\text{Hour} \times 60 + \text{Minute}$) and 1440 is the total minutes in 24 hours.

Seasonal Cycle Encoding:

$$\text{DayOfYear_sin}(t) = \sin\left(\frac{2\pi \times DOY}{365.25}\right) \tag{9}$$

$$\text{DayOfYear_cos}(t) = \cos\left(\frac{2\pi \times DOY}{365.25}\right) \tag{10}$$

where $DOY \in [1, 366]$ is the day of the year and 365.25 accounts for leap years.

Upon completing step 3, an input vector comprising 14 features is obtained. The clear-sky index K_t is defined as the ratio of the actual PV power output to the corresponding clear-sky power:

$$K_t(t) = \frac{P_{active}(t)}{P_{clear}(t)} \tag{11}$$

where $P_{active}(t)$ is the actual PV power output and $P_{clear}(t)$ is the corresponding clear-sky power. By applying the sliding window method in step 4, these 14 features are transformed into a 3D data block structure $X_{\text{seq}} \in \mathbb{R}^{N \times 24 \times 14}$, enabling the TCN model to learn the inertia of variation in past weather conditions. First, the model extracts the 2D input matrix features:

$$X^{(n)} = \begin{bmatrix} x(\text{t-23}\Delta\text{t}) \\ x(\text{t-22}\Delta\text{t}) \\ \vdots \\ x(t) \end{bmatrix} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,14} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,14} \\ \vdots & \vdots & \ddots & \vdots \\ x_{24,1} & x_{24,2} & \cdots & x_{24,14} \end{bmatrix} \tag{12}$$

Subsequently, the sliding window moves step by step progressively to obtain a set of N valid 2D matrices:

$$S = \{X^{(1)}, X^{(2)}, \ldots, X^{(N)}\} \tag{13}$$

Finally, by stacking these N 2D matrices along the sample data axis, we obtain an input 3D tensor $X_{\text{seq}} \in R^{N \times 24 \times 14}$

$$X_{\text{seq}} = \begin{bmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(N)} \end{bmatrix} \in \mathbb{R}^{N \times 24 \times 14} \tag{14}$$

Combined with a continuous time-check condition, this ensures the complete exclusion of any missing values within the past 24 steps (120 minutes).

To rigorously evaluate the model's generalisation capability and mitigate most of the risk of time-series data leakage, all N valid data samples are partitioned into three sets: training (from April 4, 2016, to December 31, 2024), validation (from January 1, 2025, to December 31, 2025), and testing (from January 1, 2026, to February 11, 2026).

However, another challenge arises because the 14 features have distinct units and scales. Consequently, Z-score normalisation transforms the data into a standard normal distribution with a mean of 0 and a variance of 1.

$$\mu_f = \frac{1}{N_{\text{train}} \times 24} \sum_{n \in N_{\text{train}}} \sum_{l=1}^{24} X_{n,l,f} \tag{15}$$

$$\sigma_f = \sqrt{\frac{1}{N_{\text{train}} \times 24} \sum_{n \in N_{\text{train}}} \sum_{l=1}^{24} \left(X_{n,l,f} - \mu_f\right)^2} \tag{16}$$

In press

2.2. Temporal Convolutional Network

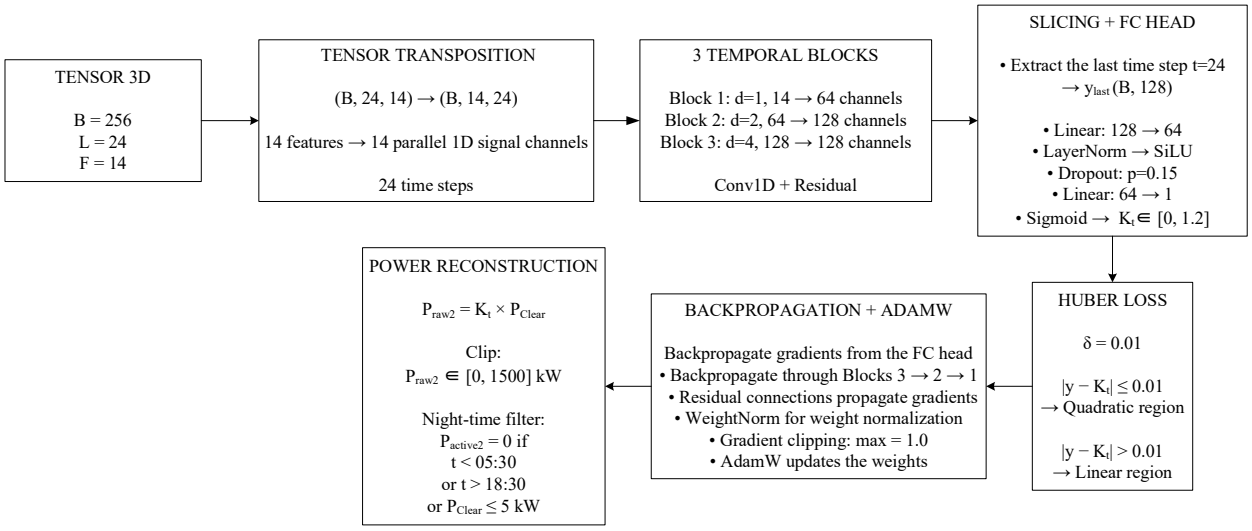


Fig. 1. Overview of the temporal convolutional network (TCN) algorithm

The Z-score-normalised 3D tensor $X_{raw} \in \mathbb{R}^{B\times 24\times 14}$ is subjected to an axis permutation resulting in:

$$X = \text{Permute}(X_{raw}, (0,2,1)) \in \mathbb{R}^{B\times 14\times 24} \tag{17}$$

This operation transforms the 14 features into 14 parallel 1D signal channels, enabling the 1D convolutional filters to scan simultaneously across all channels along the 24-step time axis.

The tensor X is then propagated through 3 sequential Temporal Blocks with an increasing dilation factor $d_l \in \{1,2,4\}$ to progressively expand the historical receptive field:

Temporal Block 1 ($d_1 = 1$, $14 \to 64$ channels): Samples consecutive time steps with a dilation factor of $[t, t-1, t-2]$ (short-range horizon of 25 minutes). The input is processed through a Left Zero-Pad ($P_1 = 2$), Conv1D ($k = 3$, WeightNorm), a Chomp1D layer to truncate future data, SiLU, and a Dropout layer (20%). The residual path employs a 1×1 convolution to project the number of channels from 14 up to 64:

$$X^{(1)} = \text{SiLU}\left(\text{Net}_1(X) + \text{Conv1D}_{1\times 1}(X)\right) \tag{18}$$

Temporal Block 2 ($d_2 = 2$, $64 \to 128$ channels): The filters sample at a dilation factor of $[t, t-2, t-4]$ (10-minute intervals), expanding the receptive field to 45 minutes. The residual path employs a 1×1 convolution to project the number of channels from 64 up to 128, obtaining $X^{(2)} \in \mathbb{R}^{B\times 128\times 24}$.

Temporal Block 3 ($d_3 = 4$, $128 \to 128$ channels): The filters sample at a dilation factor of $[t, t-4, t-8]$ (20-minute intervals), fully covering the 24-step historical sequence (2 hours). The residual path uses a direct transmission (Identity connection), obtaining $X^{(3)} \in \mathbb{R}^{B\times 128\times 24}$.

Temporal Slicing: Extracting the final time-step column solely $t = L = 24$ from the output of Block 3 to condense the entire 2-hour meteorological inertia:

$$y_{last,b} = X^{(3)}_{b,:,24} \in \mathbb{R}^{128} \tag{19}$$

FC Head Network: The vector y_{last} is propagated through the nonlinear transformation layers

$$z_1 = y_{last} \cdot W_1 + b_1 \in \mathbb{R}^{64} \tag{20}$$

$$\hat{z}_1 = LayerNorm(z_1) \tag{21}$$

$$h_1 = Dropout_{0.15}(SiLU(\hat{z}_1)) \in \mathbb{R}^{64} \tag{22}$$

$$u_b = h_1 \cdot W_2 + b_2 \in \mathbb{R} \tag{23}$$

$$\hat{y}_b = \hat{K}_{t,b} = \sigma(u_b) = \frac{1.2}{1+e^{-u_b}} \in [0,1.2] \tag{24}$$

To overcome the limitation of the standard Sigmoid function, which strictly restricts the output to the interval $(0,1)$ and fails to capture high-intensity cloud enhancement events, a scaling factor of 1.2 is introduced. This modification scales the activation output to the interval $(0,1.2)$, thereby effectively accommodating transient situations where the clear-sky index exceeds unity ($K_t > 1.0$) due to cloud-edge reflections.

The Huber loss function ($\delta = 0.01$) computed over a batch $B = 256$:

$$L_{Huber} = \frac{1}{B}\sum_{b=1}^{B} l_b \tag{25}$$

In press

$$l_b = \begin{cases} \frac{1}{2}(y_b - \hat{y}_b)^2 & \text{if} |\ y_b - \hat{y}_b| \le \delta \\ 0.01 \cdot |\ y_b - \hat{y}_b| - \frac{1}{2}(0.01)^2 & \text{if} |\ y_b - \hat{y}_b| > \delta \end{cases} \quad (26)$$

Gradient Propagation: The error signal

$$\delta_b = \frac{\partial l_b}{\partial u_b} \quad (27)$$

from the Sigmoid output backpropagates through the FC Head to obtain the gradient at y_{last}. This gradient is injected into column $t = 24$ of Tensor $(B, 128, 24)$ and flows backwards through the 3 Blocks.

Gradient Stabilisation: The Residual connections allow gradients to flow directly without vanishing. The Weight Norm decomposition

$$W = g \frac{V}{\mathrm{P}V\ \mathrm{P}_2} \quad (28)$$

prevents gradient explosion.

Gradient Clipping & AdamW: Clips the gradient norm if $g_{norm} > 1.0$. Weights are updated according to the AdamW algorithm ($\eta = 0.001, \psi = 10^{-4}$).

The predicted index $\hat{K}_t(t+15\text{min})$ is inversely converted back into active power output in kW:

$$P_{raw}(t+15\text{min}) = \hat{K}_t(t+15\text{min}) \times P_{clear}(t+15\text{min}) \quad (29)$$

$$P_{active}(t+15\text{min}) = \text{Clip}\left(P_{raw}(t+15\text{min}), 0, 1500\right) \quad (30)$$

Applying a night-time filter to eliminate physical noise:

$$P_{active}(t+15) = 0(kW) \begin{cases} t_{hour} < 5:30 \\ t_{hour} > 18:30 \\ P_{clear}(t+15\,\text{min}) \le 5kW \end{cases} \quad (31)$$

The fixed time thresholds of 05:30 and 18:30 are retained as conservative safeguards to suppress residual near-sunrise and near-sunset noise, even when the clear-sky power estimate slightly exceeds 5 kW; the $P_{clear}(t + 15\text{ min}) \le 5\text{kW}$ condition provides the primary physical nighttime criterion. The final result yields the accurate 15-minute-ahead predicted active power output at the Yulara plant.

3. Simulation Results

The 15-minute-ahead solar power forecasting performance of the Temporal Convolutional Network (TCN) is evaluated through a comprehensive benchmark comparison against six representative AI algorithms—categorised into Deep Learning approaches (TCN, LSTM, GRU) and Machine Learning methods (CatBoost, XGBoost, LightGBM)—as well as a Smart K_t Persistence baseline model. The evaluation is conducted on an independent test dataset comprising 12,096 real-time samples (spanning January 1, 2026, to February 11, 2026) from the Yulara solar power plant in Australia ($P_{rated} = 1{,}500$ kW). However, the independent test period covers approximately 42 days and represents only a limited seasonal period. Therefore, it may not fully capture seasonal variations and diverse weather conditions, which will be further investigated using longer multi-season test periods in future work.

3.1. Quantitative Evaluation on the Independent 2026 Test Dataset

Table 1. Quantitative benchmark performance across seven forecasting models on the 2026 test dataset

Algorithm	Overall MAE (kW)	Overall nMAE (%)	Overall RMSE (kW)	Overall nRMSE (%)	R^2	Forecast Skill SS_MAE (%)	Daytime MAE (kW)	Daytime nMAE (%)	Daytime RMSE (kW)	Daytime nRMSE (%)
K_t Persistence	43.5647	2.9043%	110.4715	7.3648%	0.9367	0.0000%	80.5376	5.3692%	150.3376	10.0225%
LightGBM	43.2377	2.8825%	97.6991	6.5133%	0.9505	+0.7505%	79.9321	5.3288%	132.9549	8.8637%
XGBoost	42.8870	2.8591%	97.8410	6.5227%	0.9504	+1.5556%	79.2825	5.2855%	133.1480	8.8765%
CatBoost	41.9590	2.7973%	96.9293	6.4620%	0.9513	+3.6858%	77.5637	5.1709%	131.9072	8.7938%
GRU	40.8906	2.7260%	102.1209	6.8081%	0.9459	+6.1382%	75.5850	5.0390%	138.9729	9.2649%
LSTM	39.5086	2.6339%	100.2402	6.6827%	0.9479	+9.3105%	73.0254	4.8684%	136.4133	9.0942%
TCN	38.9372	2.5958%	98.6626	6.5775%	0.9495	+10.6222%	71.9671	4.7978%	134.2662	8.9511%

To prevent artificial metric dilution caused by zero-power night-time samples, performance metrics are reported under both global test set conditions (12,096 samples) and daytime-only operating conditions ($P_{clear} > 5.0$ kW). Normalised errors ($nMAE$ and $nRMSE$) are calculated relative to the plant's rated capacity ($P_{rated} = 1{,}500$ kW). Forecast skill (SS_{MAE}) evaluates the relative error reduction achieved over the Smart K_t Persistence model.

3.1.1. Superiority in absolute error and forecast skill

As shown in Table 1, the proposed TCN model achieves the highest overall forecasting precision, obtaining the lowest overall MAE (38.94 kW, $nMAE = 2.60\%$) and daytime MAE (71.97 kW, $nMAE_{Day} = 4.80\%$). TCN achieves a 10.62% Forecast Skill improvement (SS_{MAE}) over the Smart K_t Persistence baseline. Compared to CatBoost (the best-performing Machine Learning model), TCN yields

In press

a 7.20% reduction in absolute error (38.94 kW vs. 41.96 kW). Furthermore, TCN outperforms recurrent networks, reducing MAE by 1.44% relative to LSTM (39.51 kW) and 4.77% relative to GRU (40.89 kW).

3.1.2. Analysis of deep learning vs. machine learning groups

The three Deep Learning architectures (TCN, LSTM, GRU) consistently outperform the decision-tree models (CatBoost, XGBoost, LightGBM) in terms of MAE and Forecast Skill. This dominance stems from the capacity of 1D convolutional and recurrent layers to extract temporal inertia directly from 3D tensor sequences. 24 steps×14 features), whereas tree models flatten sequential dependencies into static tabular features.

Conversely, decision-tree models achieve lower RMSE values (96.93–97.70 kW) and slightly higher R^2 scores (0.9504–0.9513). Because tree-splitting criteria minimise mean squared errors (MSE), decision trees heavily penalise large deviations, resulting in conservative predictions that smooth out extreme peaks. By incorporating the scaled activation function (1.2×Sigmoid) and physical nighttime filtering, TCN substantially reduces squared error penalties, bringing its RMSE down to 98.66 kW ($nRMSE = 6.58\%$). This bridges the RMSE gap with CatBoost (96.93 kW) while maintaining superior responsiveness during cloud transients.

3.1.3. Significance of daytime-only metrics

Evaluating daytime samples ($P_{clear} > 5.0$ kW) reveals un-inflated operational error levels. While global $nMAE$ values appear low (2.60%–2.90%) due to nighttime zero values, the actual daytime normalised error $nMAE_{Day}$ ranges from 4.80% to 5.37%, providing a realistic accuracy baseline for grid dispatchers.

3.2. Detailed Profile Analysis of Real-Time Power Curves

To visually evaluate dynamic tracking capabilities, predictions across models were analysed for two distinct meteorological profiles: a clear-sky day (January 16, 2026) and a highly volatile cloudy day (January 17, 2026).

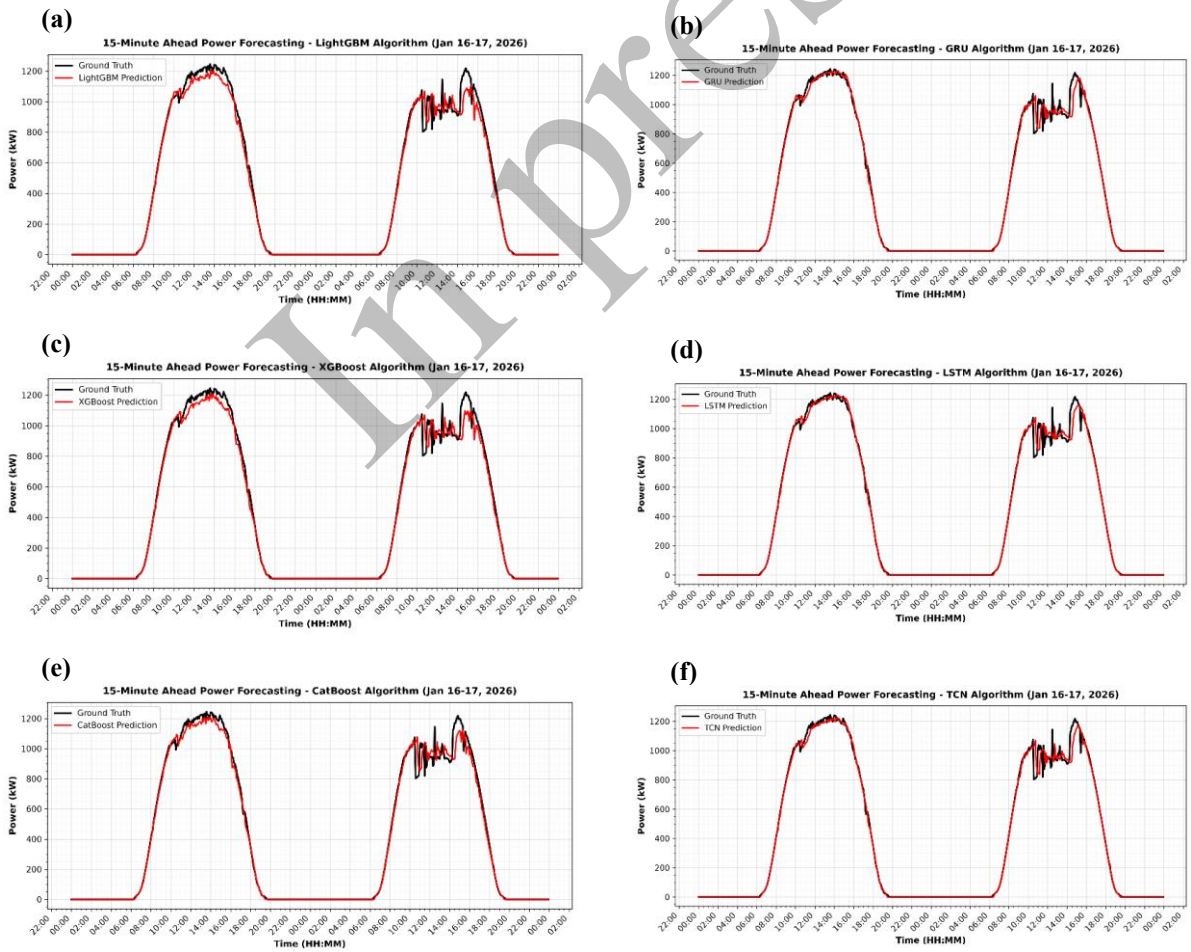


Fig. 2. Real-time 15-minute ahead PV power forecasting profiles across six algorithms compared against Ground Truth (Jan 16–17, 2026): (a) LightGBM, (b) GRU, (c) XGBoost, (d) LSTM, (e) CatBoost and (f) TCN

In press

Table 2. Detailed comparison of MAE metrics and curve morphologies among six algorithms (January 16–17, 2026)

Algorithm	Clear-Sky MAE (Jan 16, kW)	Peak Power Deviation (Jan 16)	Volatile Day MAE (Jan 17, kW)	Transient Response Lag (Jan 17)
TCN	10.29	-1.67% (1,224.25 kW)	20.90	0 min (Zero phase lag)
LSTM	11.09	-1.85% (1,221.94 kW)	20.89	5–10 min lag (Hysteresis)
GRU	10.98	-1.60% (1,225.06 kW)	21.39	5–10 min lag (Hysteresis)
CatBoost	15.72	-3.04% (1,207.16 kW)	26.43	15 min lag & step artifacts
XGBoost	19.52	-3.26% (1,204.50 kW)	27.23	15–20 min lag & step artifacts
LightGBM	19.85	-3.31% (1,203.80 kW)	27.60	15–20 min lag & step artifacts

3.2.1. Clear-sky day profile (January 16, 2026)

On clear days, active power rises smoothly along an astronomical parabolic curve, peaking at $1,245.03$ kW at 13:00. The Deep Learning models simulate the solar trajectory accurately, with TCN achieving the lowest MAE (10.29 kW) and predicting a noon peak of $1,224.25$ kW (a minor 1.67% deviation). In contrast, Machine Learning models suffer from "Peak Attenuation," suppressing peak predictions down to $1,204.50$ kW (a 40.53 kW underestimation). The discrete step-wise splitting of decision trees restricts their ability to model continuous parabolic peaks smoothly.

3.2.2. Volatile cloudy day profile (January 17, 2026)

Between 10:00 and 14:30, passing cumulus clouds induce sudden power ramp-down events, causing power to plunge rapidly from $1,218.59$ kW to 802.43 kW. TCN tracks these power drops with high responsiveness ($\text{MAE} = 20.90$ kW) and near-zero phase lag. Parallel processing across the 24-step receptive field enables TCN to detect instantaneous trend shifts immediately. Conversely, LSTM and GRU show slight phase delay (hysteresis) due to sequential hidden-state updates, while decision tree models exhibit heavy phase lags of 15–20 minutes alongside stepped profile noise.

4. Conclusion

This paper presented an integrated short-term solar power forecasting framework combining a Temporal Convolutional Network (TCN) architecture with an indirect Ineichen clear-sky index (K_t) decomposition technique. Evaluated on an independent 10-year dataset from the Yulara solar power plant in Australia (12,096 test samples from early 2026), the proposed TCN model achieved optimal overall absolute precision, obtaining an MAE of 38.94 kW ($nMAE = 2.60\%$) and a daytime MAE of 71.97 kW ($nMAE_{Day} = 4.80\%$). This represents a 10.62% Forecast Skill improvement over the Smart K_t Persistence baseline and a 7.20% error reduction relative to the top Machine Learning baseline (CatBoost). By utilizing dilated causal convolutions and a scaled activation output ($1.2 \times \text{Sigmoid}$), TCN effectively eliminates noon-peak attenuation on clear-sky days and maintains near-zero phase lag during sharp cloud ramp events.

In press

References

[1] A. Bahadori and C. Nwaoha, A review on solar energy utilisation in Australia, Renewable and Sustainable Energy Reviews, vol. 18, pp. 1–5, Feb. 2013. https://doi.org/10.1016/j.rser.2012.10.003

[2] J. Shi, W.-J. Lee, Y. Liu, Y. Yang, and P. Wang, Forecasting power output of photovoltaic system based on weather classification and support vector machine, in Proceedings of the IEEE Industry Applications Society Annual Meeting (IAS), 2011, pp. 1–6. https://doi.org/10.1109/IAS.2011.6074294

[3] P. Hewage, M. Trovati, E. Pereira, and A. Behera, Deep learning-based effective fine-grained weather forecasting model, Pattern Analysis and Applications, vol. 24, no. 1, pp. 343–366, Feb. 2021. https://doi.org/10.1007/s10044-020-00898-1

[4] P. Hewage, A. Behera, M. Trovati, E. Pereira, M. Ghahremani, F. Palmieri, and Y. Liu, Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station, Soft Computing, vol. 24, pp. 16453–16482, Nov. 2020. https://doi.org/10.1007/s00500-020-04954-0

[5] J. Xu, S. Yoo, J. Heiser, and P. Kalb, Sensor network based solar forecasting using a local vector autoregressive ridge framework, in Proceedings of the 31st Annual ACM Symposium on Applied Computing, Pisa, Italy, Apr. 2016, pp. 2113–2118. https://doi.org/10.1145/2851613.2853124

[6] S. Mishra and P. Palanisamy, Multi-time-horizon solar forecasting using recurrent neural network, in

Proceedings of the IEEE Energy Conversion Congress and Exposition (ECCE), Portland, OR, USA, 2018, pp. 18–24. https://doi.org/10.1109/ECCE.2018.8558187

[7] Y. Ge, Y. Nan, and L. Bai, A hybrid prediction model for solar radiation based on long short-term memory, empirical mode decomposition, and solar profiles for energy harvesting wireless sensor networks, Energies, vol. 12, iss. 24, Dec. 2019, Art. no. 4762. https://doi.org/10.3390/en12244762

[8] S. Srivastava and S. Lessmann, A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data, Solar Energy, vol. 162, pp. 232–247, Mar. 2018. https://doi.org/10.1016/j.solener.2018.01.005

[9] R. Azimi, M. Ghayekhloo, and M. Ghofrani, A hybrid method based on a new clustering technique and multilayer perceptron neural networks for hourly solar radiation forecasting, Energy Conversion and Management, vol. 118, pp. 331–344, Jun. 2016. https://doi.org/10.1016/j.enconman.2016.04.009

[10] F. J. Rodríguez-Benítez, M. López-Cuesta, C. Arbizu-Barrena, M. M. Fernández-León, M. Á. Pamos-Ureña, J. Tovar-Pescador, F. J. Santos-Alamillos, and D. Pozo-Vázquez, Assessment of new solar radiation nowcasting methods based on sky-camera and satellite imagery, Applied Energy, vol. 292, Jun. 2021, Art. no. 116838. https://doi.org/10.1016/j.apenergy.2021.116838

[11] Q. Paletta, G. Arbod, and J. Lasenby, Omnivision forecasting: Combining satellite and sky images for improved deterministic and probabilistic intra-hour solar energy predictions, Applied Energy, vol. 336, Apr. 2023, Art. no. 120818. https://doi.org/10.1016/j.apenergy.2023.120818

[12] F. Wang, Z. Xuan, Z. Zhen, Y. Li, K. Li, L. Zhao, M. Shafie-khah, and J. P. S. Catalão, A minutely solar irradiance forecasting method based on real-time sky image-irradiance mapping model, Energy Conversion and Management, vol. 220, Sep. 2020, Art. no. 113075. https://doi.org/10.1016/j.enconman.2020.113075

[13] L. H. Dissawa, R. I. Godaliyadda, P. B. Ekanayake, A. P. Agalgaonkar, D. Robinson, J. B. Ekanayake, and S. Perera, Sky image-based localized, short-term solar irradiance forecasting for multiple PV sites via cloud motion tracking, International Journal of Photoenergy, vol. 2021, Jul. 2021, Art. no. 9973010. https://doi.org/10.1155/2021/9973010

[14] C. Feng and J. Zhang, SolarNet: A sky image-based deep convolutional neural network for intra-hour solar forecasting, Solar Energy, vol. 204, pp. 71–78, Jul. 2020. https://doi.org/10.1016/j.solener.2020.03.083

[15] E. Nikitidou, A. Zagouras, V. Salamalikis, and A. Kazantzidis, Short-term cloudiness forecasting for solar energy purposes in Greece, based on satellite-derived information, Meteorology and Atmospheric Physics, vol. 131, pp. 175–182, Apr. 2019. https://doi.org/10.1007/s00703-017-0559-0

[16] V. Afzali Gorooh, S. Kalia, P. Nguyen, K.-L. Hsu, S. Sorooshian, S. Ganguly, and R. R. Nemani, Deep neural network cloud-type classification (DeepCTC) model and its application in evaluating PERSIANN-CCS, Remote Sensing, vol. 12, iss. 2, Jan. 2020, Art. no. 316. https://doi.org/10.3390/rs12020316

[17] H. Shaohua, F. Xin, W. Xiuru, Z. Chenyu, Z. Fuju, and W. Jiaming, Distributed photovoltaic power output prediction based on satellite cloud map video frames, Frontiers in Energy Research, vol. 11, Oct. 2023, Art. no. 1247304. https://doi.org/10.3389/fenrg.2023.1247304

[18] A. Prasad and M. Kay, Prediction of solar power using near-real time satellite data, Energies, vol. 14, iss. 8, Sep. 2021, Art. no. 5865. https://doi.org/10.3390/en14185865

[19] Z. Si, M. Yang, Y. Yu, and T. Ding, Photovoltaic power forecast based on satellite images considering effects of solar position, Applied Energy, vol. 302, Nov. 2021, Art. no. 117514. https://doi.org/10.1016/j.apenergy.2021.117514

[20] J. Mathe, N. Miolane, N. Sebastien, and J. Lequeux, PVNet: A LRCN architecture for spatio-temporal photovoltaic power forecasting from numerical weather prediction, 2019, arXiv:1902.01453.

[21] E. Ekinci, A comparative study of LSTM-ED architectures in forecasting day-ahead solar photovoltaic energy using weather data, Computing, vol. 106, no. 5, pp. 1611–1632, Feb. 2024. https://doi.org/10.1007/s00607-024-01266-1

In press

Appendix A. Nomenclature

Table A1. Description of mathematical notations, physical symbols, and units

Symbol	Mathematical Definition / Physical Meaning	Unit / Value Range
$P_{active}(t)$	Actual active power output of the plant at time t	kW
$GHI(t)$	Measured Global Horizontal Irradiance at ground level	W/m²
$GHI_{clear}(t)$	Theoretical Global Horizontal Irradiance under clear-sky conditions	W/m²
$P_{clear}(t)$	Maximum theoretical power output of the plant under clear-sky conditions	kW
$K_t(t)$	Actual clear-sky index at time t	[0.0, 1.2]
$\hat{K}_t$	Clear-sky index predicted by the TCN at a future time step	[0.0, 1.2]
θ_z, α_s	Solar zenith angle (θ_z) and solar elevation angle (α_s)	Degrees (°)

Symbol	Mathematical Definition / Physical Meaning	Unit / Value Range
L, H	Historical window length ($L = 24$ steps) and forecast horizon ($H = 3$ steps)	Time step (5 min)
N, B	Total number of valid data samples and training batch size ($B = 256$)	Samples
F, F_{in}, F_{out}	Number of input features ($F = 14$), number of input channels, and number of output channels	Channels
k, d_l, P_l	Kernel size ($k = 3$), dilation rate, and padding size at layer l	Integer
W, V, g, b	Weight matrix, direction vector, magnitude scalar, and bias vector	Real-valued
y_{last}	Hidden-state vector at the last time step $t = L$ after the Slicing operation	$\mathbb{R}^{128}$
l_b, L_{Huber}	Loss of sample b and mean Huber loss over batch B	Real-valued
δ	Threshold separating the MSE and MAE regions in the Huber loss ($\delta = 0.01$)	Real-valued
η, ψ	Learning rate ($\eta = 0.001$) and weight-decay coefficient ($\psi = 10^{-4}$)	Real-valued
$X_{n,l,f}$	Value of the f -th feature at time step l for sample n in the raw 3D tensor	Real-valued
n	Index of the sliding-window data sample ($n \in [1, N]$)	Integer
l	Time-step index within the historical window ($l \in [1, 24]$), corresponding to 2 hours	Step (5 min)
f	Input-feature column index ($f \in [1,14]$)	Integer
μ_f	Arithmetic mean of feature f , calculated exclusively from the training set	Real-valued
σ_f	Standard deviation of feature f , calculated exclusively from the training set	Real-valued
g_{norm}	Global L_2 norm of the gradient vector used for Gradient Clipping	Real-valued
S	Set of N valid 2D input matrices generated by the sliding-window operation	24x14 per matrix
X_{seq}	3D input tensor of all samples	$\mathbb{R}^{N\times 24\times 14}$
P_{rated}	Rated power capacity of the PV plant	kW
$nMAE$	Normalized mean absolute error	%
$nRMSE$	Normalized root mean square error	%
SS_{MAE}	Relative MAE reduction compared with the Smart K_t Persistence baseline	%
Δ_t	Sampling interval	5 min

In press