

ORIGINAL RESEARCH

Open access

Advancements in Solar Radiation Prediction: GAN-Solar Model for High-Precision Forecasting in Photovoltaic Systems

Reem Al-Harbi^{1*}, Omar Al-Qahtani², Salma Nasser¹

Abstract

Reliable short-term forecasting of solar radiation is crucial to ensuring the smooth operation and optimal scheduling of photovoltaic (PV) power systems. At present, cutting-edge encoder–decoder frameworks that leverage satellite remote sensing imagery are the dominant strategy for accomplishing this task. Yet these approaches reveal clear weaknesses, particularly when using longer prediction windows. Under extended horizons, the resulting maps commonly display blurred spatial patterns and noticeable inaccuracies in radiation levels. Such flaws substantially undermine forecast sharpness and trustworthiness, rendering them inadequate for applications requiring very high accuracy. To address these shortcomings, we introduce GAN-Solar — a novel post-processing architecture for improving short-term solar radiation predictions that leverages Generative Adversarial Networks (GANs). The framework adopts a conditional ED-AttUNet as its generator. A carefully designed discriminator equipped with residual blocks, stepwise resolution reduction, and auxiliary conditional signals is trained to separate genuine from synthetic radiation fields. Through this competitive training — steered by a composite loss and a discriminator acting as an adaptive quality metric — the model progressively sharpens prediction fidelity. Evaluations of real summer solar radiation datasets show that GAN-Solar delivers substantial gains in output quality. Compared with the plain ED-AttUNet baseline, it reduces the Root Mean Square Error by about 3.2% and increases the Structural Similarity Index from 0.84 to 0.87. Overall, the method effectively corrects for texture smearing and intensity warping, yielding noticeably crisper, more faithful solar radiation maps.

Keywords Solar radiation, Short-term forecast, Generative adversarial networks, Photovoltaic systems

*Correspondence:

Reem Al-Harbi
rharbi@ksu.edu.sa

¹ Department of Smart Electrical Systems, Faculty of Engineering, King Saud University, Riyadh, Saudi Arabia

² Department of Energy Data Systems, Faculty of Engineering, Qatar University, Doha, Qatar

Introduction

Today's global energy system is rapidly pivoting toward low-carbon and renewable energy sources, and solar power has established itself as a cornerstone of this transition [1]. Photovoltaic (PV) technology, in particular, provides an efficient and increasingly affordable means of converting sunlight into electricity, thereby curbing greenhouse gas emissions and bolstering energy independence [2]. The accelerating rollout of PV installations — spanning massive solar parks to individual building-integrated arrays — underscores solar energy's indispensable role in building a decarbonized energy system [3].

Because solar radiation fluctuates strongly due to variable weather patterns, integrating large volumes of PV electricity into power grids remains technically challenging [4]. High-quality short-term solar radiation forecasts, therefore, constitute a prerequisite for successful grid control, allowing better load balancing, enhanced reliability of supply [5], and fuller exploitation of available solar generation [6]. Given the sparse and uneven distribution of surface observation stations, satellite-based remote sensing has become an essential complementary data stream, providing consistent, wide-area information [7]. Meanwhile, deep learning algorithms [8, 9] have quickly become the preferred tool for processing these voluminous satellite-derived irradiance records, proving highly effective at learning the intricate, nonlinear space-time behavior of solar radiation [10, 11].

Even though deep learning has driven remarkable improvements, the most elaborate models — especially those built around encoder–decoder pipelines — still struggle with intrinsic constraints [12]. The difficulty becomes particularly evident at longer lead times: generated radiation maps tend to lose fine-grained spatial structure, leading to important local details fading or vanishing entirely [13]. At the same time, systematic biases often appear in the predicted irradiance magnitudes, thereby compromising downstream calculations of expected power output [14]. These combined effects erode forecast credibility and limit usability. The root causes are commonly traced to progressive error buildup during autoregressive generation and to insufficient modeling of subtle atmospheric phenomena [15]. As a consequence, current predictions often fall short of the stringent fidelity standards demanded by critical real-world use cases [16].

In response to these persistent quality bottlenecks in short-term solar radiation forecasting, the present work proposes GAN-Solar, an original GAN-based refinement framework. GAN-Solar operates as a powerful post-processing stage that enhances the coarse outputs produced by a strong existing spatiotemporal predictor (namely ED-AttUNet). The ED-AttUNet backbone is retained as the generator and is conditioned on relevant contextual cues to produce more informed initial estimates. The distinguishing feature of GAN-Solar is its custom discriminator, which embeds residual pathways, gradual feature downsampling, and conditional guidance. This discriminator learns to recognize authentic high-resolution radiation distributions and thereby provides constructive feedback to the generator. By means of sustained adversarial optimization — regulated through a blended loss term and with the discriminator serving as a differentiable quality proxy — the entire system learns to suppress spatial blurring and intensity artifacts, ultimately delivering solar radiation forecasts that are simultaneously more accurate, visually coherent, and realistically detailed. The key novel contributions of the paper are the following:

This study presents GAN-Solar, an original GAN-based framework specifically designed to refine and improve the quality of short-term solar radiation predictions. The approach directly targets and corrects the prevalent problems of spatial texture deterioration and misrepresentation of radiation intensity that frequently occur in the outputs of forecasting models.

Introduction of a sophisticated GAN design that incorporates an improved ED-AttUNet architecture serving as the generator, paired with a purpose-built discriminator. This discriminator integrates residual pathways, stepwise feature downsampling, and conditional guidance, thereby operating as an adaptive, learnable criterion that steers the refinement process toward greater realism and precision.

Thorough performance assessment using authentic satellite-based solar radiation records, which confirms that GAN-Solar delivers marked improvements in both numerical evaluation metrics and the perceptual visual sharpness of the resulting solar radiation maps relative to a robust baseline model.

Related Work

Techniques for solar radiation forecasting have advanced considerably, incorporating insights from multiple fields. They can generally be grouped into empirical/statistical techniques, numerical weather prediction (NWP) approaches, and, more recently, deep learning-based strategies.

Empirical and statistical forecasting methods

Initial efforts in solar radiation prediction relied primarily on conventional statistical methods, including linear regression and classical time-series methods. Later studies shifted toward models capable of representing the nonlinear behavior typical of solar radiation time series. This evolution prompted the use of more advanced nonlinear techniques such as Support Vector Regression (SVR) [17, 18] and Artificial Neural Networks (ANNs) [19, 20]. These approaches have demonstrated stronger performance in extracting complex nonlinear relationships, thereby improving forecast skill for short- to medium-range horizons. For example, Hejase and Assi [21] investigated time-series regression methods, including Autoregressive Moving Average (ARMA) models, for predicting daily global solar radiation. Similarly, Alsharif et al. [22] developed seasonal autoregressive integrated moving-average (SARIMA) models for both daily and monthly solar radiation estimation, achieving acceptable levels of accuracy.

In summary, although statistical techniques offer adaptability and can yield useful predictions, they suffer from intrinsic constraints, including reliance on simplifying assumptions and limited generalization to unseen scenarios. As a result, research has increasingly shifted toward machine learning and artificial intelligence solutions that aim to address these weaknesses and improve both accuracy and applicability.

Numerical weather prediction-based forecasting methods

Among current numerical models, the Weather Research and Forecasting system extended with solar radiation capabilities (WRF-Solar) remains one of the most commonly applied for irradiance forecasting. Various investigations have tested and refined WRF-Solar for better solar prediction performance. Prasad and Kay [23] examined its effectiveness in reproducing diverse components of solar irradiance under highly variable conditions at the Mildura site in Australia. In a separate effort, Evans et al. [24] evaluated different combinations of WRF physics parameterizations when simulating East Coast Lows—intense rainfall systems—affecting Australia's southeast coastal region. Their findings indicated that the choice of physics scheme becomes especially decisive during heavy precipitation events.

Additionally, Kim et al. [25] formulated an optimized procedure for introducing random perturbations to selected variables in WRF-Solar. They subsequently employed the model's irradiance outputs, together with satellite measurements, to train multiple fully convolutional neural network architectures. This combined strategy also illustrated the strong promise of deep fully convolutional networks for short-term solar radiation prediction.

NWP techniques are appreciated for their grounding in physical principles and their capacity to cover broad geographic regions and extended forecast periods. Nevertheless, they demand considerable computational resources, rely on extensive input datasets, and remain sensitive to uncertainties arising from initial states, physical parameterizations, and model assumptions—factors that particularly degrade accuracy for short-lead, localized predictions involving evolving cloud fields rapidly.

Deep learning-based forecasting methods

The strong success of Generative Adversarial Networks (GANs) [26] has spurred their widespread use across tasks such as image synthesis, style transfer, and image restoration. In solar-related applications, GANs often generate high-fidelity imagery, whereas convolutional LSTM units capture the temporal evolution of radar or satellite image sequences. Assaf et al. [27] developed a data-augmentation framework that combined GANs, CNNs, and SVR; their system successfully generated novel yet statistically consistent samples, markedly boosting the precision of solar radiation forecasts. Son et al. [28] fused LSTM networks with GANs to extract temporal patterns from image sequences and predict forthcoming cloud fields. The generated images were subsequently converted into quantitative estimates and combined with weather variables and historical PV records to improve photovoltaic power forecasts. In a different direction, Zhang et al. [29] presented SolarGAN, a specialized GAN variant designed to fill gaps in multivariate solar energy time series. By adapting the GAN input format, the technique effectively manages largely independent solar sequences. It surpasses prior imputation methods across a range of missingness levels, thereby supporting more reliable PV forecasting. Anbo et al.

[30] addressed the challenges of short-term solar radiation prediction for recently commissioned wind-solar hybrid sites and introduced a multi-gradient evolutionary deep network. Their solution integrates time-series GANs with multivariate variational mode decomposition to sustain forecast quality despite constrained training data.

Despite these notable advances, short-term solar radiation forecasting remains hampered by widespread spatial texture degradation and intensity distortion in model outputs. Although many contemporary deep learning approaches can produce superficially realistic predictions, the results often lack critical high-frequency spatial information and accurate magnitude representation. This lack of detail reduces the practical value of the forecasts and increases the risk of erroneous interpretations or suboptimal decisions in real-time energy system operations. The persistence of these quality issues reveals an important limitation in existing methods. It underscores the urgent need for more powerful refinement strategies to restore detail, precision, and overall trustworthiness to solar radiation predictions.

Short-Term Solar Radiation Forecast Quality Optimization Method Based on GANs

This section presents the proposed GAN-Solar model—an innovative framework developed to significantly improve the fidelity of short-term solar radiation predictions [31]. By leveraging Generative Adversarial Networks combined with residual learning blocks, our method specifically targets and corrects typical artifacts encountered in forecasting outputs. Problems such as progressive loss of spatial texture detail and systematic shifts in radiation intensity commonly result from cumulative errors in conventional models, particularly as prediction horizons increase. As illustrated in Figure 1, the overall GAN-Solar framework consists of a generator and a discriminator that are jointly refined through adversarial training.

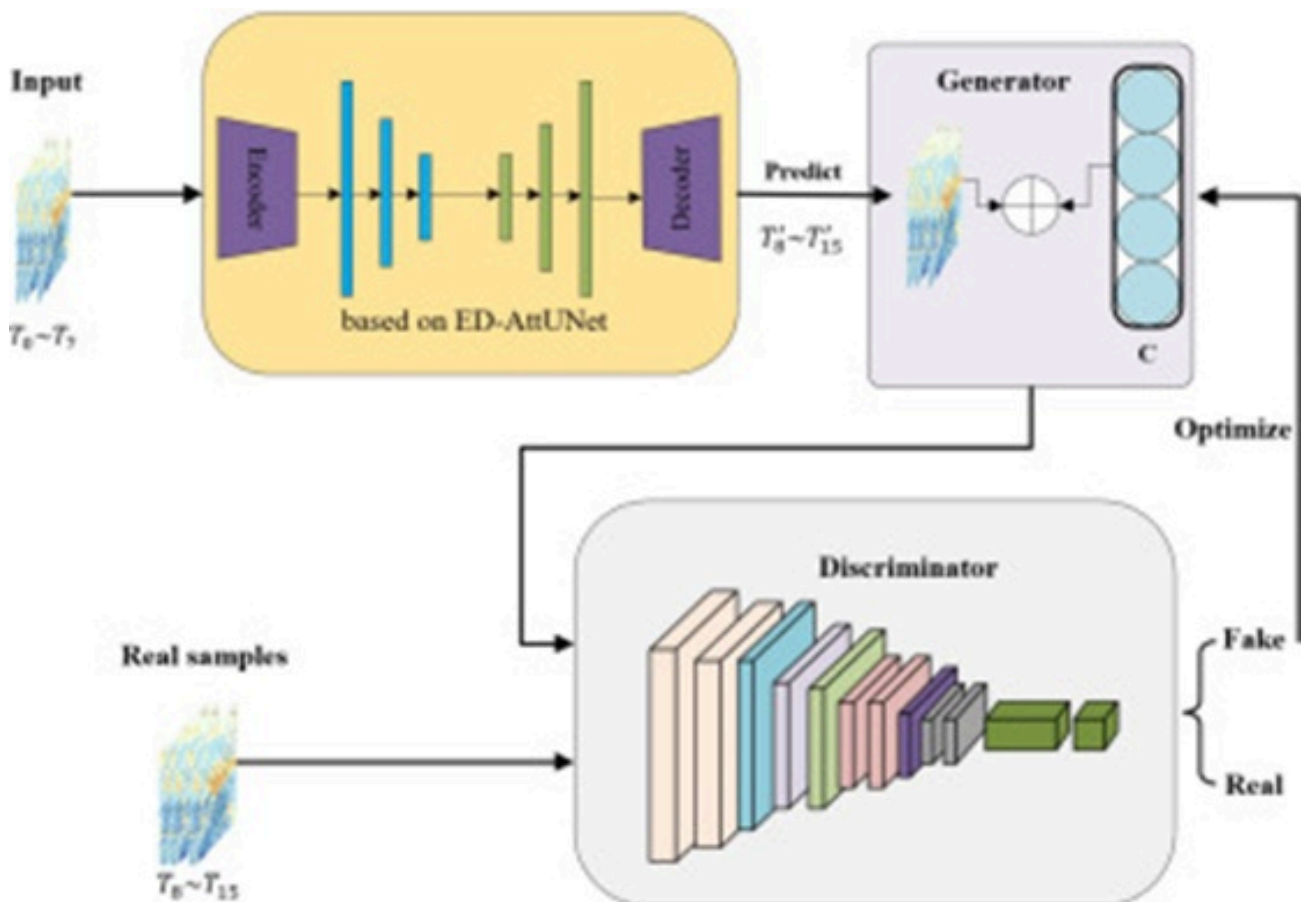


Figure 1. GAN-Solar model architecture.

The generator is constructed on top of the ED-AttUNet backbone and is responsible for sharpening initially coarse or blurred forecast maps. It handles the forecast sequence frame-wise: the output from the current time step is fed back as a conditional signal to guide the generation of the next frame, thereby preserving temporal consistency across the sequence. Rather than merely performing binary real/fake classification, the discriminator functions as a trainable, differentiable quality evaluator. It provides detailed gradient information that guides the generator toward outputs that achieve not only high numerical precision but also structural and textural characteristics virtually identical to those of actual satellite-derived ground truth.

Within the GAN framework, the generator G and discriminator D participate in a continuous minimax contest. The generator strives to generate increasingly lifelike samples, while the discriminator aims to improve at detecting synthetic samples. Formally, their opposing goals are captured by the following value function:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z)))], \quad (1)$$

Where $\mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)]$ denotes the expected log-probability assigned by the discriminator to real samples drawn from the true data distribution $p_{\text{data}}(x)$. Here, $D(x)$ represents the probability that the discriminator assigns to sample x being real; maximizing this expectation encourages the discriminator to assign values close to 1 for genuine data. Conversely, $\mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z)))]$ concerns generated samples $G(z)$ produced by the generator from latent noise z . The term $D(G(z))$ is the discriminator's estimated probability that $G(z)$ is real; minimizing this expression pushes the generator to craft samples that the discriminator increasingly classifies as authentic (i.e., $D(G(z))$ approaches 1). Through this competitive dynamic, the two networks iteratively improve during training until an equilibrium is approached in which generated samples closely match the statistics of the real distribution.

Thus, $1 - D(G(z))$ quantifies the discriminator's belief that a generated sample is synthetic. The generator's objective — minimizing the second term — effectively seeks to maximize $D(G(z))$, deceiving the discriminator into accepting synthetic outputs as real. The standard GAN training procedure consists of two alternating steps:

(1) Discriminator update: With the generator held fixed, optimize the discriminator D to maximize $V(D, G)$, thereby strengthening its capacity to differentiate real from fake samples.

(2) Generator update: With the discriminator held fixed, optimize the generator G to minimize $V(D, G)$, enabling it to produce more convincing outputs that evade detection.

The detailed training procedure for the GAN-Solar model is outlined in [Algorithm 1](#).

```

1: Input: Training set  $D$ , 8 samples per batch, adversarial training iterations  $T$ , discriminator iterations per round  $K$ , gradient penalty parameter  $\gamma$ 
2: Output: Generator  $G(z, \theta)$ 
3: Randomly initialize  $\theta, C$ 
4: for epochs do
5:   Train discriminator
6:   for  $t$  do
7:     Get training samples
8:     Get predicted 2-hour radiation data  $\{T'_1, T'_2, \dots, T'_8\}$  from ED-AttUNet
9:     Input to generator  $G$  combined with random noise  $C$ 
10:    Input corresponding 2-hour real radiation data  $\{T_1, T_2, \dots, T_8\}$  from the training set
11:    Calculate the mean of the discriminator's output for fake data
12:    Update discriminator parameters using discriminator optimization loss function  $loss_D$ 
13:  end for
14:  Again, get predicted 2-hour radiation data  $\{T'_1, T'_2, \dots, T'_8\}$  from ED-AttUNet
15:  Update generator loss function  $loss_G$ 
16: end for

```

Algorithm 1. GAN-Solar Network training process.

GAN-SOLAR generator network structure design

This study adopts a GAN-based framework in which the generator relies on an existing, well-established architecture rather than introducing an entirely new one. Specifically, the generator reuses the ED-AttUNet model. As illustrated in **Figure 2**, this network receives a sequence of historical solar radiation maps covering the previous two hours (denoted $T_0 \sim T_7$) as input. Employing an encoder–decoder backbone augmented with spatio-temporal attention modules, it produces a refined prediction sequence $T'_8 \sim T'_{15}$ frame by frame, forecasting solar radiation over the subsequent two-hour period. The input consists of solar radiation data from 8 consecutive time steps spanning two hours, and the output correspondingly predicts the next 8 time steps over the same duration. Thanks to its specialized design, ED-AttUNet excels at modeling sequential data and effectively captures extended temporal dependencies within time series.

Nevertheless, despite the strong performance of ED-AttUNet, its forecasts still exhibit noticeable spatial texture degradation and radiation intensity distortion, issues that warrant further improvement. Building directly on this foundation, the present work extends the model by incorporating supplementary conditional information C . This enhancement enables finer-grained control over the generation process.

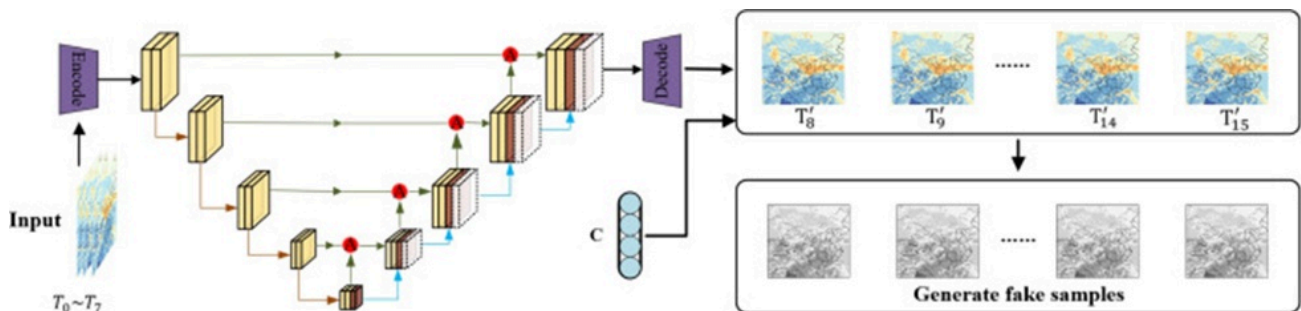


Figure 2. GAN-SOLAR generator network structure.

More precisely, the conditional information C is fused with the feature maps generated at the final stages of the ED-AttUNet decoder. This strategic integration point is deliberate: by applying the conditioning signal late in the pipeline, the generator can provide high-level contextual guidance to modulate the final output without interfering with the encoder–decoder core's essential spatio-temporal representations. The result is a more precise, condition-sensitive refinement of the forecasts.

The incorporation of auxiliary conditioning draws inspiration from Conditional Generative Adversarial Networks (cGANs) [32]. In the cGAN paradigm, both the generator and the discriminator receive additional conditioning information, enabling more directed, controllable data synthesis. By supplying C as a conditioning input, the generator is steered to generate predictions that remain highly realistic while simultaneously aligning closely with the provided external contextual cues.

GAN-SOLAR discriminator network structure design

Conventional GAN discriminators typically rely on a stacked convolutional pyramid architecture [33]. However, this standard design suffers from three primary limitations: (1) excessive depth often leads to vanishing gradients and increases the risk of mode collapse, particularly with high-resolution inputs; (2) static convolutional filters have difficulty adapting to highly varied and complex data distributions; and (3) common normalization techniques tend to dilute the influence of conditional signals, hindering effective incorporation of auxiliary information such as meteorological variables. Short-term solar radiation forecasting involves rich, high-dimensional spatio-temporal patterns, imposing stringent requirements on the discriminator. It must robustly extract and discriminate these intricate features while remaining responsive to conditional inputs, thereby enabling accurate judgment or generation under diverse contextual scenarios. To address these challenges, integrating residual network blocks with conditional information C substantially enhances training stability and model performance. Residual connections mitigate degradation in very deep networks through identity mappings. At the same time, conditional guidance C equips the discriminator with the adaptability needed to tailor its feature analysis and classification decisions to specific conditions.

As depicted in **Figure 3**, the GAN-SOLAR discriminator adopts a progressive downsampling strategy to capture multi-scale spatio-temporal characteristics from solar radiation map sequences. This architecture offers two key benefits: effective feature fusion across different resolutions and a relatively lightweight parameter footprint. Each residual block comprises two consecutive 3×3 convolutional layers, followed by batch normalization and LeakyReLU nonlinearities. Skip connections within each block ensure smooth gradient propagation, markedly improving the trainability of deep architectures. At the same time, the embedded conditional information C enables dynamic modulation of feature processing and decision boundaries based on the provided context. In cases where input and output channel dimensions differ within a residual unit, a 1×1 convolution is inserted to align dimensions. To progressively extract richer representations, several residual blocks are stacked in series. Downsampling is achieved via strided convolutions (stride = 2) that gradually reduce the spatial resolution. After the feature extraction stages, global average pooling is applied to compress the spatial dimensions, reduce the parameter count, and mitigate overfitting. One or more fully connected layers then aggregate the pooled features to produce the final classification decision — distinguishing genuine short-term solar radiation forecasts from those synthesized by the generator. At the output, a single fully connected layer followed by a Sigmoid activation produces a scalar score in the $[0, 1]$ range, representing the estimated probability that the input is authentic.

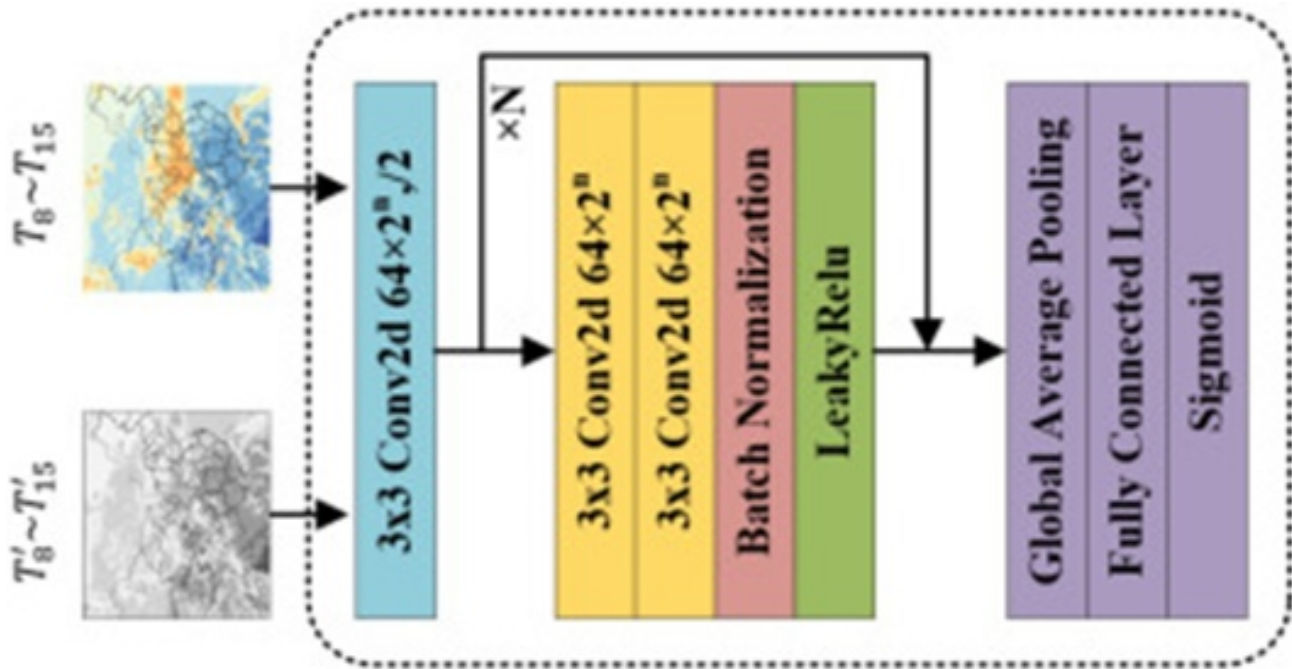


Figure 3. GAN-SOLAR discriminator network structure.

In our design, the conditional information C is first reshaped and then concatenated with the intermediate feature maps in the discriminator. This deliberate fusion strategy empowers the discriminator to conduct conditional discrimination: it assesses not merely whether a given forecast appears realistic in isolation, but whether it is realistic under the particular condition C . Such conditioning effectively discourages the generator from producing outputs that may seem superficially plausible yet fail to align with the provided contextual constraints.

To reliably distinguish authentic from synthetic prediction results, the discriminator requires a sufficiently expansive receptive field. This broad field of view is essential for accurately gauging overall similarity between predicted and observed solar radiation maps. Achieving this with conventional CNN-based architectures typically demands either larger convolutional kernels or deeper stacks of layers—both of which substantially inflate the model's parameter count. However, the massive storage footprint of geostationary satellite-derived radiation datasets already imposes strict limits on allowable model complexity. To resolve this tension, GAN-Solar introduces an innovative multi-scale discriminator architecture that constructs a feature pyramid, thereby optimizing computational cost while preserving strong discriminative capability. As illustrated in **Figure 4**, the original full-resolution observation image (left) is spatially downsampled by a factor of 2 to produce a lower-resolution version (right). This operation dramatically reduces computational overhead and simultaneously enlarges the discriminator's effective receptive field, enabling efficient evaluation of the forecast's global structural properties. Although downsampling inevitably discards some fine-grained local details, its primary objective is to provide a complementary global perspective that augments the detailed local scrutiny of the original high-resolution input. For instance, a 2×2 convolutional kernel operating at the original resolution captures the same spatial context as a 1×1 kernel applied after $2 \times$ downsampling. Leveraging this principle, the present work employs small convolutional kernels throughout the discriminator, preserving robust local feature modeling while markedly decreasing the total parameter volume. This multi-scale approach offers a practical and efficient solution for discriminating high-resolution spatio-temporal fields without sacrificing judgment quality.

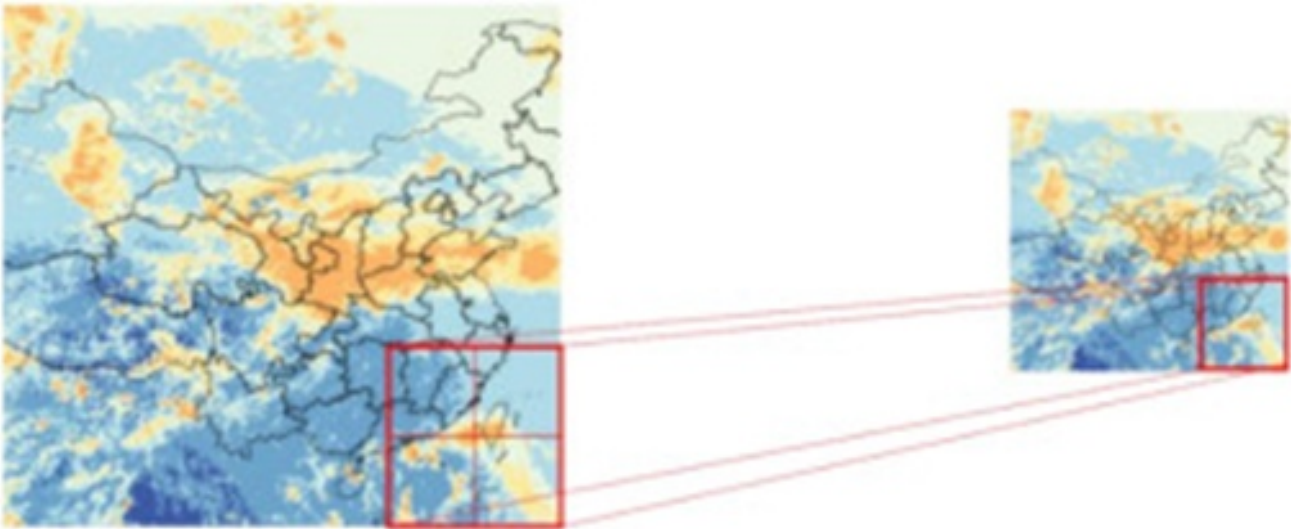


Figure 4. Downsampling.

The overall depth of the discriminator is governed by the layer configuration specified in [Table 1](#). Progressive downsampling is achieved via a sequence of strided convolutional layers (stride = 2). Together with accompanying normalization and activation operations, these layers form the fundamental residual building blocks described earlier. The resulting network architecture achieves a sufficiently large receptive field, allowing the discriminator to evaluate not only fine local textures but also the large-scale structural consistency and coherence across the entire forecast map.

Table 1. Discriminator network parameter information.

Layer	Type	Kernel size/Stride	Output channels	Padding
1	Convolutional layer	3×3/step = 1	64	1
2	Convolutional layer	3×3/step = 2	64	1
3	Layer normalization	–	–	–
4	LeakyReLU	–	–	–
5	Convolutional layer	3×3/step = 2	128	2
6	Layer normalization	–	–	–
7	LeakyReLU	–	–	–
8	Convolutional layer	3×3/step = 1	128	2
9	Convolutional layer	3×3/step = 2	256	2
10	Layer normalization	–	–	–
11	LeakyReLU	–	–	–
12	Convolutional layer	3×3/step = 1	256	2
13	Convolutional layer	3×3/step = 2	512	2
14	Layer normalization	–	–	–
15	LeakyReLU	–	–	–

16	Global average pooling	–	–	–
17	Fully connected layer	–	–	–
18	Sigmoid	–	–	–

GAN-SOLAR loss function design

Training GANs encounters several technical difficulties [34]. The fundamental challenge stems from maintaining equilibrium between the generator and the discriminator within a constantly evolving competitive framework. Achieving this dynamic equilibrium demands that both networks sustain a state of rivalry while jointly improving overall performance through adversarial training. Nevertheless, the procedure involves considerable complications: the discriminator's primary role is to maximize the distinguishability between genuine and generated samples, whereas the generator seeks to approximate the true data distribution by reducing distributional divergence. This opposing dynamic produces the following behaviors during training: excessive optimization of the generator may cause the discriminator to develop an indistinct decision boundary; on the other hand, an overly dominant discriminator can restrict the generator's capacity for exploration.

To enhance the quality of generated outputs, this work adopts a carefully designed hybrid loss function that enables fine-grained control over generator training. In the GAN-SOLAR model, the generator's loss is formulated as a multi-objective optimization structure, incorporating an adversarial loss component, a Mean Squared Error (MSE) term for structural fidelity, and a Mean Absolute Error (MAE) term as a robustness constraint. A gradient-weighted fusion strategy is employed to harmonize the dual goals of adversarial learning and physical realism, thereby overcoming the game imbalance issues commonly encountered in conventional GAN training.

The generator's objective function can be expressed as:

$$L_G = \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))] + \omega_{MSE} \cdot MSE + \omega_{MAE} \cdot MAE, \quad (2)$$

Where ω_{MSE} and ω_{MAE} represent essential hyperparameters that regulate the relative contributions of the adversarial loss and the pixel-level fidelity terms. Notably, these weights remain constant throughout each training run and are not dynamically adjusted. The term refers to our systematic approach of determining these weights via multiple preliminary experiments to identify an effective configuration.

Consequently, the complete GAN-SOLAR model objective function is given by:

$$\theta(D, G) = \mathbb{E}_{x \sim P_{data}(x)} [\log D(x|c)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z|c)))], \quad (3)$$

$$\min_G \max_D V(D, G) = \theta(D, G) + MSE + MAE. \quad (4)$$

The incorporation of these additional metrics aims to holistically account for the diversity, realism, and structural similarity of the generated outputs relative to the target data, thereby further improving both training stability and overall generation performance.

Experimental Environment and Results

Dataset

The experimental dataset in this study is derived from the surface incident solar radiation product from the Fengyun-4A geostationary satellite, obtained from the National Satellite Meteorological Center. This product is generated using data from the advanced geostationary radiation imager. Surface solar incident radiation denotes the shortwave solar energy received per unit time and per unit area on a horizontal plane at the Earth's surface. This parameter has broad applications in meteorology, climate studies, land-surface modeling, the solar energy sector, agriculture, and related domains.

The data utilized in this experiment consists of summer radiation records spanning early June to late August 2022. The detailed steps for dataset preparation are illustrated in **Figure 5**. The process includes the following stages:

First, convert all downloaded NC-format files into NPY format.

Second, remove corrupted files and those lacking radiation values.

Third, extract the China region using the satellite-projection transformation algorithm provided by the National Satellite Meteorological Center.

Fourth, apply bilinear interpolation to achieve a 15-minute temporal resolution.

Fifth, normalize the radiation values to the range [0, 1].

Sixth, construct the dataset employing a sliding window approach.

Seventh, partition the data into training, validation, and test sets. Specifically, the initial 70% of the data was assigned to the training set, the subsequent 15% to the validation set (for hyperparameter optimization and model selection), and the remaining 15% to the test set (reserved solely for final performance evaluation).

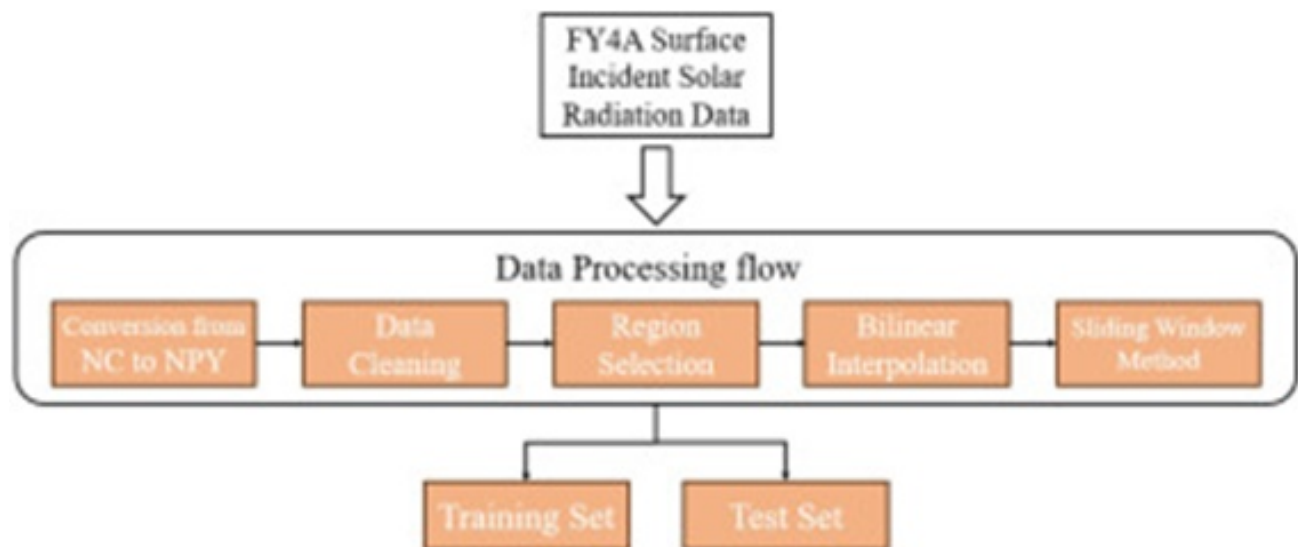


Figure 5. Dataset creation process.

Experimental setup

The experiments in this section employ the following model configurations and training procedures: The batch size is fixed at 16. Training utilizes the Adam optimizer with an initial learning rate of 0.001. The model takes as input a sequence

of 8 historical time steps and is tasked with forecasting the subsequent 8 consecutive time steps. For the ED-AttUNet architecture, the number of stacked layers in both the encoder (NE) and decoder (ND) is initially set to 4. The proposed GAN-SOLAR model is evaluated against a baseline that incorporates residual networks.

During training, the following protocol is adopted: Following 10,000 adversarial training iterations between the generator and discriminator, a model checkpoint is saved every additional 2,000 iterations. At the same time, performance metrics are simultaneously logged for both the training and validation sets. Comprehensive details regarding the software and hardware environment are provided in [Table 2](#).

Table 2. Software and hardware environment information.

Experimental environment	Equipment model and version information
Operating system	CentOS 7.9.2009
CPU	Xeon(R) CPU E5-2620 v4
GPU	NVIDIA TITAN V
Development language	PyTorch 1.12.0
Development tool	Python 3.8.10

Results and analysis

This section presents an in-depth analysis building upon the ED-AttUNet framework. The experiments use summer surface solar radiation data from the Fengyun-4A geostationary satellite, spanning early June to late August 2022. To rigorously assess the effectiveness of the proposed contributions, a comprehensive set of ablation studies was conducted.

Four progressively enhanced model variants were defined and systematically compared against the ED-AttUNet baseline to quantify the impact of individual components:

- GAN-Solar-V1: A vanilla GAN implementation lacking both conditional guidance and residual connections. This serves as the minimal baseline for ablation analysis.
- GAN-Solar-V2: An extension of V1 that incorporates conditional constraint information.
- GAN-Solar-V3: Further augments GAN-Solar-V2 by introducing residual structures.
- GAN-Solar-V4: The complete proposed model, which jointly integrates conditional information and residual connections to deliver optimal performance.

In this ablation study, ED-AttUNet serves as the reference benchmark for assessing improvements in generation quality. As presented in [Table 3](#), GAN-Solar-V1 — without conditional information or residual connections — exhibits modestly inferior performance relative to ED-AttUNet, confirming that the plain GAN underperforms the benchmark on this task.

GAN-Solar-V3, which adds the residual structure, yields noticeable gains, particularly in MSE and SSIM, achieving values of 88.75 and 0.87, respectively. The fully equipped GAN-Solar-V4, which combines both conditional constraints and residual connections, consistently outperforms all preceding variants across all evaluated metrics. Most strikingly, nRMSE is reduced to 27.30.

These findings strongly validate the synergistic benefit of integrating conditional guidance with residual architectures: the approach not only elevates generation quality but also markedly strengthens the model's capacity to capture and represent intricate data distributions.

Table 3. Ablation experiment result comparison.

Model	MSE↓	RMSE↓	SSIM↑	nRMSE↓
ED-AttUNet	90.06	135.19	0.86	29.87
GAN-Solar-V1	91.32	136.12	0.83	31.07
GAN-Solar-V2	90.02	135.37	0.85	30.06
GAN-Solar-V3	88.75	133.78	0.87	28.97
GAN-Solar-V4	87.56	132.76	0.88	27.30

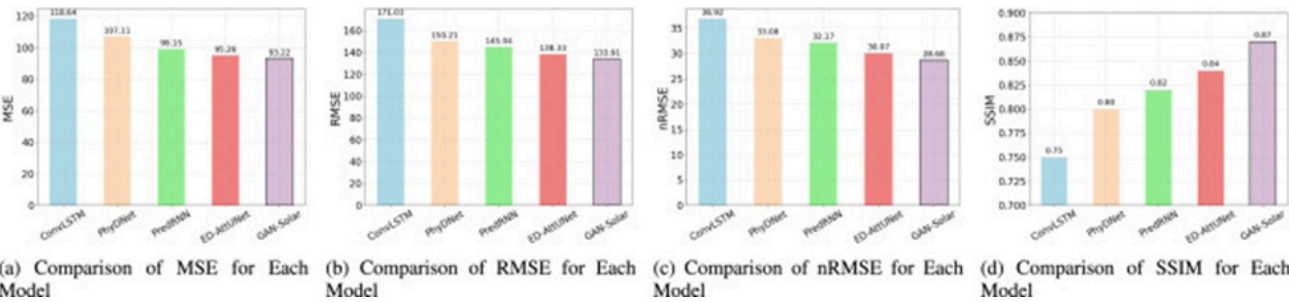


Figure 6. Comparison of all statistical metrics for each model.

To further strengthen the rigor of the study and enhance confidence in the findings, this section implements a more extensive comparative evaluation to thoroughly demonstrate the superiority of the proposed model. In particular, the forecasting performance of the GAN-Solar model is benchmarked against several well-established deep learning models widely used for spatiotemporal prediction, namely ConvLSTM [35], PhyDNet [36], and PredRNN [37]. These baseline models were deliberately selected to represent three fundamentally different paradigms in spatiotemporal forecasting, thereby ensuring a comprehensive and balanced comparison. ConvLSTM acts as the classic recurrent convolutional baseline, PredRNN exemplifies a more sophisticated, fully data-driven recurrent architecture, and PhyDNet represents the class of hybrid approaches that incorporate physics-informed constraints, enabling our method to be evaluated against an alternative methodological lineage.

As illustrated in Table 4 and Figure 6, the proposed GAN-Solar model substantially surpasses the competing models across multiple statistical performance indicators. Subplots (a), (b), and (c) display the Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and normalized RMSE (nRMSE), respectively; for these error-based metrics, lower values reflect superior performance. Subplot (d) shows the Structural Similarity Index (SSIM), with higher values indicating better results.

Concretely, the GAN-Solar model records the lowest values of MSE, RMSE, and nRMSE among all compared methods, indicating that its predictions are closest to the ground truth and exhibit the highest overall forecasting accuracy. Concurrently, GAN-Solar achieves the highest SSIM score of 0.87, demonstrating that the short-term solar radiation forecast images it generates possess the greatest structural fidelity to the real observations and most faithfully preserve the spatiotemporal patterns of solar radiation.

By contrast, the ConvLSTM model performs worst across all metrics, underscoring its limitations in both quantitative accuracy and visual quality. While PhyDNet and PredRNN deliver improvements over ConvLSTM in certain aspects, their overall effectiveness remains noticeably below that of the GAN-Solar model. These observed differences arise from intrinsic shortcomings of the baseline approaches. ConvLSTM, constrained by spatially uniform memory mechanisms, struggles to adequately model localized cloud dynamics, frequently producing excessively smoothed or blurred outputs. Despite its architectural advancements, PredRNN still degrades high-frequency details and sharpness, particularly as the forecast horizon increases. The effectiveness of PhyDNet, in turn, heavily depends on the accuracy of its embedded physical model; any discrepancy between this simplified representation and the true complexity of atmospheric processes can introduce systematic biases and propagate errors over time.

Table 4. Comparison of all statistical metrics for each Model from June to August 2022.

Model	ConvLSTM	PhyDNet	PredRNN	ED-AttUNet	GAN-Solar
MSE	118.64	107.11	99.15	95.26	93.22
RMSE	171.03	150.21	145.04	138.33	133.91
nRMSE	36.92	33.08	32.17	30.07	28.68
SSIM	0.75	0.80	0.82	0.84	0.87

In particular, by refining the ED-AttUNet foundation, the GAN-Solar model delivers a clear further boost to the quality of the generated forecasts. This outcome provides strong evidence of both the practical value and the progressive nature of our GAN-driven refinement strategy for elevating short-term solar radiation prediction performance. Beyond that, it supplies a robust technical foundation for increasing the efficiency, reliability, and real-world applicability of solar photovoltaic installations.

Table 5, together with **Figure 7**, tracks how key performance indicators evolve across various models as the forecast lead time increases. Examining the RMSE progression during multi-step prediction reveals that all models exhibit cumulative error buildup, reflected in a consistent upward trend in RMSE as the number of future steps increases. While GAN-Solar fully retains the solid forecasting strength demonstrated by ED-AttUNet, it also achieves additional improvements. By contrast, ConvLSTM sees its error balloon to 233.19 by Step 8—a dramatic 133% rise from Step 1. The physics-guided PhyDNet model struggles notably, posting higher errors than purely data-driven approaches beyond Step 5. This pattern suggests a fundamental mismatch between the model's prescribed differential equations and the observed dynamics, exposing the constraints inherent in traditional physics-embedded designs when handling extended temporal sequences. GAN-Solar and ED-AttUNet stand out for their superior ability to restrain error growth. Over the medium- to longer-range window (Steps 5 through 8), RMSE rises by roughly 41% for GAN-Solar and 37.8% for ED-AttUNet—gains that substantially exceed those of the competing approaches. That said, even though GAN-Solar holds the lowest overall RMSE at distant horizons (reaching only 181.26 at Step 8, a level comparable to where ConvLSTM and PhyDNet stand around Step 6), its error accumulation pace edges slightly ahead of ED-AttUNet's. PredRNN excels in the very short range (Steps 1–4), keeping RMSE impressively low at 89.29 and 96.48 across the first two steps, well below the field. Yet after Step 5, its accuracy falls off steeply, with RMSE climbing by 48% from Step 5 to Step 8, a markedly sharper deterioration than seen in either ED-AttUNet or GAN-Solar.

Table 5. Comparison of root mean square error metrics for each model at different time steps from June to August 2022.

Model	Step1	Step2	Step3	Step4	Step5	Step6	Step7	Step8
Conv-LSTM	100.08	109.65	121.06	133.25	154.93	183.92	211.48	233.19
Phy-DNet	96.52	103.82	115.22	128.47	151.91	178.44	195.59	227.88

Pred-RNN	89.29	96.48	114.61	127.23	145.37	169.23	187.04	211.19
ED-AttUNet	90.47	97.84	109.37	124.74	139.93	158.37	175.74	192.93
GAN-Solar	88.97	93.31	100.22	114.94	128.51	144.71	167.74	181.26

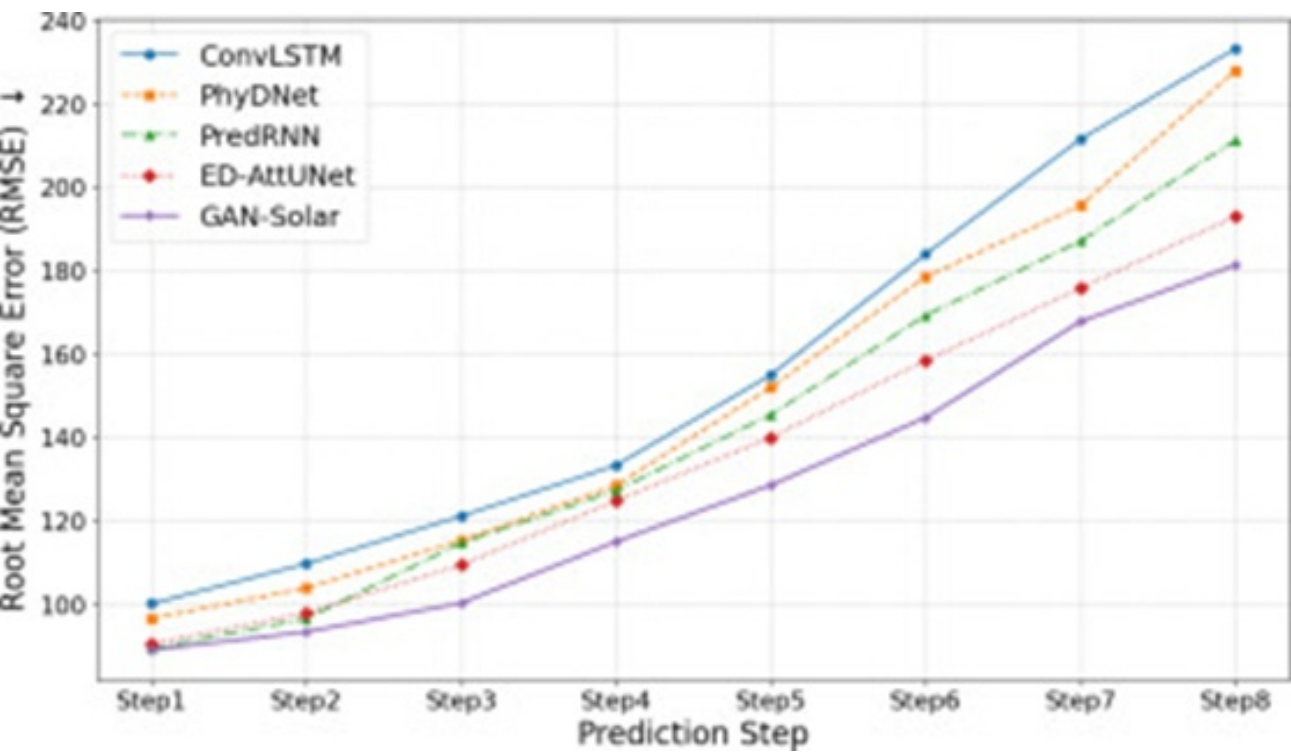


Figure 7. Comparison of Root Mean Square Error Metrics for each model at different time steps.

The SSIM trends presented in Table 6 and Figure 8 reinforce this picture: across the board, structural similarity scores decline as prediction length increases, signaling a gradual loss of detail fidelity over longer horizons. ConvLSTM consistently ranks last in SSIM at every step, underscoring its persistent weakness in preserving image structure. PredRNN holds a competitive edge through Step 4 but then suffers a noticeable drop-off between Steps 5 and 8, compared with the two architectures proposed here. Meanwhile, GAN-Solar—while capitalizing on the baseline stability of ED-AttUNet—introduces targeted enhancements that allow it to more faithfully reconstruct the complex spatiotemporal patterns inherent in solar radiation fields. Together, these observations highlight the clear superiority and real-world promise of a GAN-based framework like GAN-Solar for refining short-term solar radiation forecasts.

Taken as a whole, deep learning methods—especially the GAN-Solar model—prove particularly well suited to the challenges of short-term solar radiation prediction. The model not only markedly sharpens forecast clarity and numerical precision but also leverages powerful representation learning to achieve greater robustness and consistency in the face of the intricate, rapidly changing nature of solar radiation patterns. In doing so, it furnishes substantially stronger support for optimized scheduling, dispatch, and overall management of solar photovoltaic energy systems.

Table 6. Comparison of structural similarity index metrics for each model at different time steps from June to August 2022.

Model	Step1	Step2	Step3	Step4	Step5	Step6	Step7	Step8
-------	-------	-------	-------	-------	-------	-------	-------	-------

Conv-LSTM	0.79	0.77	0.75	0.72	0.69	0.65	0.60	0.53
Phy-DNet	0.81	0.77	0.75	0.73	0.70	0.65	0.58	0.51
Pred-RNN	0.84	0.83	0.82	0.79	0.75	0.70	0.64	0.59
ED-AttUNet	0.87	0.86	0.84	0.81	0.78	0.73	0.67	0.61
GAN-Solar	0.90	0.87	0.85	0.82	0.79	0.75	0.70	0.65

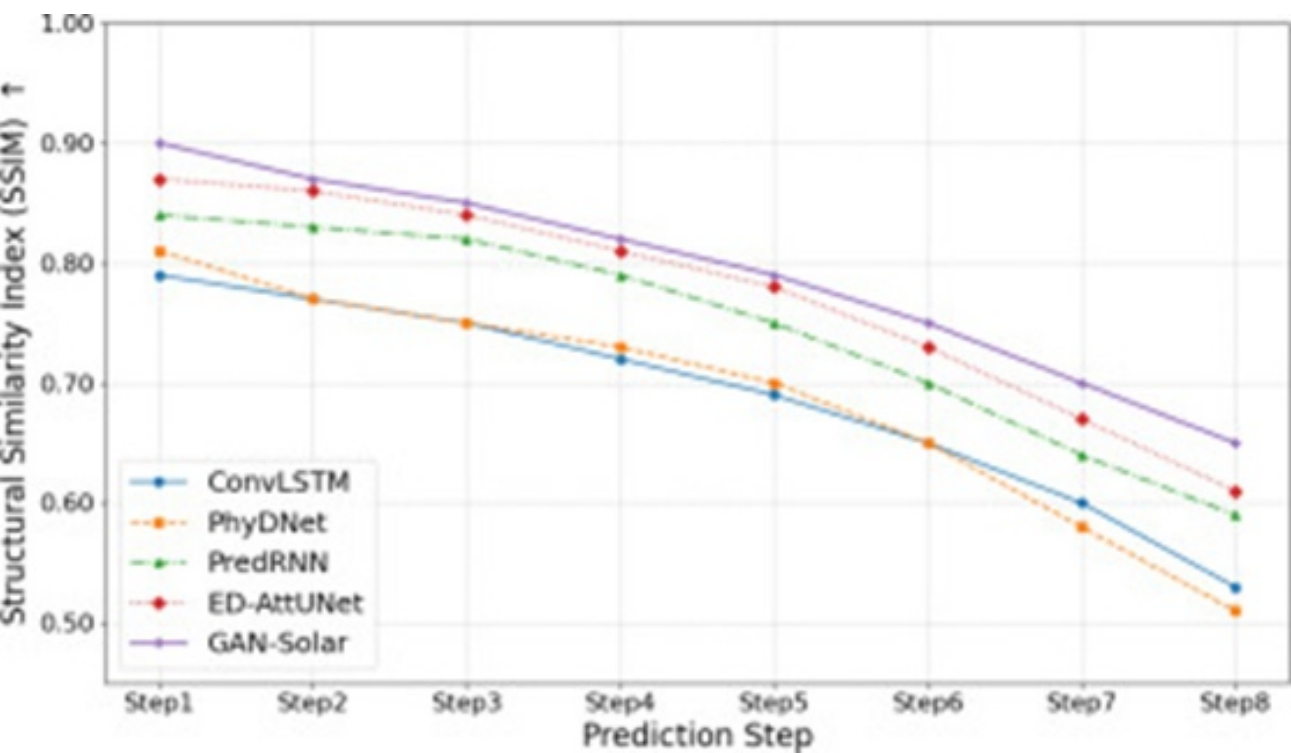


Figure 8. Comparison of Structural Similarity Index Metrics for each model at different time steps.

Conclusion

In this study, our primary focus was on investigating GAN-based techniques to enhance the quality of short-term solar radiation forecasting outputs. We began by examining discrepancies between predictions from current short-term solar radiation models and corresponding ground-truth observations, and pinpointing the underlying sources of these inaccuracies. Building on this analysis, we introduced the GAN-Solar model, which leverages the principles of generative adversarial networks. In this framework, the discriminator functions as an adaptive, learnable loss objective; by re-designing the optimization process from the standpoint of the objective function itself, the model achieves substantial improvements in forecast quality. Through the competitive interplay between the generator and the discriminator, the predicted data distribution is progressively driven closer to the true empirical distribution. As a result, the generated forecasts exhibit markedly greater sharpness and fidelity, effectively mitigating the common issue of fine detail degradation to a meaningful degree.

Extensive experimental evaluation demonstrated that the GAN-based approach yields considerable performance gains in short-term solar radiation prediction tasks. The refined outputs are visually crisper and quantitatively more aligned with real measured values, thereby elevating both the precision and trustworthiness of the forecasts. These findings highlight the substantial promise of GANs for improving the quality of short-term solar radiation predictions and open new directions and methodological pathways for subsequent research and practical applications in this domain.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China under Grant [62372242](#) and Grant [92267104](#), the Jiangsu Provincial Major Project on Basic Research of Cutting-edge and Leading Technologies, China, under Grant no. [BK20232032](#).

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 29 Mar 2023

Revised: 10 May 2023

Accepted: 20 Jun 2023

Published online: 20 September 2023

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

- Nagaraja MR, Biswas WK, Selvan CP. Advancements and challenges in solar photovoltaic technologies: Enhancing technical performance for sustainable clean energy – a review. *Sol Energy Adv.* 2024;100084.
- Song T, Jeswani HK, Azapagic A. Assessing the environmental and economic sustainability of emerging tandem photovoltaic technologies in China. *Energy Convers Manag.* 2024;318:118890.
- Sweeney C, Bessa RJ, Browell J, Pinson P. The future of forecasting for renewable energy. *Wiley Interdiscip Rev Energy Environ.* 2019;9(2):e365.
<https://doi.org/10.1002/wene.365>.

Osman M, Qureshi I. Review of photovoltaic and concentrated solar technologies including their performance, reliability, efficiency and storage. *Results Eng.* 2025;104424.

Xu X, Meng K, Xiang H, Cui G, Xia X, Dou W. Blockchain-enabled secure, fair, and scalable data sharing in zero-trust edge-end environment. *IEEE J Sel Areas Commun.* 2025.
<https://doi.org/10.1109/JSAC.2025.3560007>.

Tohsing K, Phaisathit D, Janjai S. Applying a semi-empirical model for estimating hourly diffuse solar radiation at the main regions of Thailand. *J Phys Conf Ser.* 2025;2934(1):012032.

Mirzabe AH, Hajiahmad A, Keyhani A. Assessment and categorization of empirical models for estimating monthly, daily, and hourly diffuse solar radiation: A case study of Iran. *Sustain Energy Technol Assess.* 2021;47:101330.

Kumari P, Toshniwal D. Deep learning models for solar irradiance forecasting: a comprehensive review. *J Clean Prod.* 2021;318:128566.

Zhou Y, Liu Y, Wang D, Liu X, Wang Y. A review on global solar radiation prediction with machine learning models in a comprehensive perspective. *Energy Convers Manag.* 2021;235:113960.

Khan HA, Alam M, Rizvi HA, Munir A. Solar irradiance forecasting using deep learning techniques. *Eng Proc.* 2023;46(1):15.
<https://doi.org/10.3390/engproc2023046015>.

Voyant C, Notton G, Kalogirou S, Nivet ML, Paoli C, Motte F, et al. Machine learning methods for solar radiation forecasting: A review. *Renew Energy.* 2017;105:569-82.

Dou W, Wang K, Shan S, Zhang K, Wei H. Day-ahead correction of numerical weather prediction solar irradiance forecasts based on similar day analysis. In: 2023 35th Chinese Control and Decision Conference (CCDC). IEEE; 2023. p. 2554-9.

Chisale SW, Lee HS, Soto Calvo MA, Jeong JS, Aljber M, Williams Z, et al. Advanced solar energy potential assessment in Malawi: Utilizing high-resolution WRF model and GIS to identify optimal sites for solar PV generation. *Renew Energy.* 2025;239:122084.

Barhmi K, Heynen C, Golroodbari S, van Sark W. A review of solar forecasting techniques and the role of artificial intelligence. *Solar.* 2024;4(1):99-135.
<https://doi.org/10.3390/solar4010005>.

Xu X, Dong H, Xiang H, Hu X, Li X, Xia X, et al. C2IRec: causal contrastive learning for user cold-start recommendation with social variables. *ACM Trans Inf Syst.* 2025.

Benavides CL, Amaro e Silva R, Manso Callejo MA, Cira CI. Review on spatio-temporal solar forecasting methods driven by in situ measurements or their combination with satellite and numerical weather prediction (NWP) estimates. *Energies.* 2022;15(12):4341.
<https://doi.org/10.3390/en15124341>.

Bangdiwala SI. Regression: simple linear. *Int J Inj Control Saf Promot.* 2018;25(1):113-5.

Martínez-Merino LI, Puerto J, Rodríguez-Chía AM. Ordered weighted average support vector regression. *Expert Syst Appl.* 2025;274:126882.

Abdolrasol MGM, Hussain SMS, Ustun TS, Sarker MR, Hannan MA, Mohamed R, et al. Artificial neural networks based optimization techniques: a review. *Electronics.* 2021;10(21):2689.
<https://doi.org/10.3390/electronics10212689>.

Maier HR, Galelli S, Razavi S, Castelletti A, Rizzoli A, Athanasiadis IN, et al. Exploding the myths: An introduction to artificial neural networks for prediction and forecasting. *Environ Model Softw.* 2023;167:105776.
<https://doi.org/10.1016/j.envsoft.2023.105776>.

Hejase HAN, Assi AH. Time-series regression model for prediction of mean daily global solar radiation in Al-Ain, UAE. *Int Sch Res Not*. 2012;2012:412471.

Alsharif MH, Younes MK, Kim J. Time series ARIMA model for prediction of daily and monthly average global solar radiation: the case study of Seoul, South Korea. *Symmetry*. 2019;11(2):240.

Prasad AA, Kay M. Assessment of simulated solar irradiance on days of high intermittency using WRF-solar. *Energies*. 2020;13(2):385.

Evans JP, Ekström M, Ji F. Evaluating the performance of a WRF physics ensemble over Southeast Australia. *Clim Dyn*. 2012;39:1241-58.

Kim JH, Jimenez PA, Dudhia J, Yang J, Sengupta M, Xie Y, et al. Probabilistic Forecast of All-Sky Solar Radiation Using Enhanced WRF-Solar: Tech. Rep. National Renewable Energy Lab.(NREL), Golden, CO, United States; 2020.

Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. *Commun ACM*. 2020;63(11):139–44.

Assaf AM, Haron H, Abdull Hamed HN, Ghaleb FA, Dalam ME, Eisa TAE, et al. Improving solar radiation forecasting utilizing data augmentation model generative adversarial networks with convolutional support vector machine (GAN-CSVR). *Appl Sci*. 2023;13(23):12768.

Son Y, Zhang X, Yoon Y, Cho J, Choi S. LSTM–GAN based cloud movement prediction in satellite images for PV forecast. *J Ambient Intell Humaniz Comput*. 2023;14:12373–86.

Zhang W, Luo Y, Zhang Y, Srinivasan D. SolarGAN: multivariate solar data imputation using generative adversarial network. *IEEE Trans Sustain Energy*. 2020;12(1):743–6.

Meng A, Zhang H, Yin H, Xian Z, Chen S, Zhu Z, et al. A novel multi-gradient evolutionary deep learning approach for few-shot wind power prediction using time-series GAN. *Energy*. 2023;283:129139.

Xu X, Wu F, Bilal M, Xia X, Dou W, Yao L, et al. XRL-SHAP-Cache: An explainable reinforcement learning approach for intelligent edge service caching in content delivery networks. *Sci China Inf Sci*. 2024;67:170303.

Boroujeni SPH, Razi A. IC-GAN: An improved conditional generative adversarial network for RGB-to-IR image translation with applications to forest fire monitoring. *Expert Syst Appl*. 2024;238:121962.

Xiang H, Zhang X, Xu X, Beheshti A, Qi L, Hong Y, et al. Federated learning-based anomaly detection with isolation forest in the IoT-edge continuum. *ACM Trans Multimed Comput Commun Appl*. 2024;22(1).

Xiang H, Zhang X, Xu X, Beheshti A, Qi L, Hong Y, et al. OptiForest: optimal isolation forest for anomaly detection. *arXiv preprint arXiv:2306.12703*. 2023.

Shi X, Chen Z, Wang H, Yeung DY, Wong WK, Woo WC. Convolutional LSTM network: a machine learning approach for precipitation nowcasting. *Adv Neural Inf Process Syst*. 2015;28.

Le Guen V, Thome N. Disentangling physical dynamics from unknown factors for unsupervised video prediction. In: *Proc IEEE/CVF Conf Comput Vis Pattern Recognit*. 2020. p. 11474–84.

Wang Y, Wu H, Zhang J, Gao Z, Wang J, Yu PS, et al. PredRNN: A recurrent neural network for spatiotemporal predictive learning. *IEEE Trans Pattern Anal Mach Intell*. 2022;45(2):2208–25.