energies

MDPI

Article

Prediction of Photovoltaic Power Output at New Energy Bases in the Desert Region During Sandstorm Weather

Shuhao Wang [1,2], Junhan Xu [3], Shi Chen [1,2,*], Jiangping Chen [1,2] and Hongping Yan [1,2]

[1] College of Electrical Engineering, Sichuan University, Chengdu 610065, China; wangshuhao@stu.scu.edu.cn (S.W.); qqcom5083@gmail.com (J.C.); 2023223035219@stu.scu.edu.cn (H.Y.)
[2] Intelligent Electric Power Grid Key Laboratory of Sichuan Province, Sichuan University, Chengdu 610065, China
[3] Faculty of Science and Technology, Beijing Normal-Hong Kong Baptist University, Zhuhai 519000, China; t330024269@mail.bnbu.edu.cn
* Correspondence: chen_shi@scu.edu.cn

Abstract

To address the challenge of forecasting power output from large-scale photovoltaic (PV) bases in desert regions during sand and dust storms, this paper proposes a hybrid data-physics driven prediction method. This approach utilizes satellite remote sensing to obtain regional irradiance data, transforming the traditional one-dimensional time-series forecasting into a two-dimensional spatiotemporal sequence prediction, thereby tracking the dynamic evolution of irradiance intensity under the influence of sand and dust. Firstly, a forecasting model based on a conditional variational autoencoder (CVAE) optimized with a recurrent state-space model (RSSM) is constructed to effectively capture both the deterministic trends and stochastic fluctuations in irradiance variation, providing a reliable input basis for power calculation. Secondly, at the physical modeling level, the model comprehensively considers the isotropic scattering characteristics and changes in sky clarity induced by sand and dust weather, establishing a physical mapping relationship from irradiance to PV output. This mitigates the constraint of scarce historical operational data in desert and sandy regions. This research provides a novel solution for regional-level PV power forecasting under extreme sand and dust weather, contributing to enhanced dispatchability and transmission stability of renewable energy bases during abrupt meteorological changes.

Keywords: solar PV power forecasting; sandstorm; regional solar photovoltaic; satellite remote sensing; conditional generative model

Check for updates

Academic Editor: Silvano Vergura

Received: 10 December 2025
Revised: 25 January 2026
Accepted: 27 January 2026
Published: 4 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

China's desert and Gobi regions are rich in wind and solar resources and host large-scale renewable energy installations. In large-scale new energy bases, PV power generation primarily relies on solar radiation. However, these areas also experience frequent sand and dust weather events, such as floating dust, blowing sand, and sandstorms, which are characterized by early onset, high intensity, and extensive impact. Sand and dust weather can reduce ground-level solar irradiance and significantly decrease PV power generation. In the most severe cases, this could lead to shutdowns of units at new energy bases and interruptions in power transmission, resulting in significant losses [1]. Beyond local PV performance degradation, dust storms can also threaten grid-level energy security. An extreme dust storm in Spain caused nationwide PV output reductions exceeding 80% on a single day and prolonged losses over 50%, leading to significant economic and operational

impacts on the power system. Such results emphasize the importance of developing robust forecasting and mitigation methods for PV power generation under sandstorm weather [2]. Some studies have further linked meteorological conditions with photovoltaic generation characteristics. Reference [3] proposed a historical PV-output characteristic extraction framework based on weather-type classification, indicating that PV output exhibits distinct behaviors under different weather patterns. Nevertheless, these approaches mainly focus on conventional conditions and remain insufficient for rapidly evolving extreme sand and dust weather. Therefore, accurate and timely forecasting of PV output in response to dust weather, considering meteorological uncertainties, can help new energy bases schedule and allocate flexible resources in advance, ensuring stable power transmission.

Existing PV output forecasting methods can be broadly categorized into three types: data-driven, physical modeling, and hybrid approaches. Among them, the data-driven approach primarily utilizes historical data from PV power stations, leveraging the powerful nonlinear fitting capability of neural networks to learn underlying patterns and predict future PV power generation [4–6]. Its limitation lies in the lack of interpretability of the prediction mechanism, as it proves highly unreliable for scenarios outside the distribution of the training data and exhibits strong dependence on data quality. Physical modeling, on the other hand, is based on the physical principles of PV power generation and the technical parameters of PV modules. It offers strong interpretability regarding the physical mechanisms of weather changes [7], and employs forecasted irradiance values as inputs to construct a chain of physical calculation models [8]. After instantiating the model parameters, the prediction accuracy will be determined by the input data. However, because physical models require detailed system parameters and precise numerical weather prediction (NWP) data, they are inherently complex and demand considerable effort to ensure their accuracy. Hybrid forecasting methods combine the advantages of different forecasting models to overcome the limitations of any single model, thereby achieving more accurate and stable prediction results [9,10].

Although renewable energy output forecasting technology has formed a relatively well-established technical system, research under extreme weather conditions still faces the following main challenges: (1) The low probability of extreme weather events leads to scarce historical samples, which severely impacts the performance and development potential of deep learning models, resulting in issues such as overfitting and poor test data performance; (2) When meteorological factors undergo abrupt changes beyond the historical experience boundaries of the model, and renewable energy output characteristics alter drastically, the models struggle to capture such sudden variations, demonstrating poor adaptability.

Significant progress has been made in research on renewable energy output forecasting under extreme weather conditions. Considering the issue of sample deficiency, reference [11] decomposed the power samples during cold waves into baseline and loss sequences, establishing separate forecasting models. By utilizing the characteristic parameters of wind turbine protection control, the method identifies and extracts periods of power loss. Reference [12] defined and established discriminant models for three types of extreme weather—cold waves, typhoons, and icing—based on meteorological factors. Employing the concept of transfer learning with pre-training and fine-tuning effectively enhanced forecasting accuracy. Reference [13], from the perspective of distributed PV power generation, defined extreme weather events and addressed issues such as data scarcity and uncertainty by proposing a few-shot data augmentation method based on generative adversarial network (GAN). The study used an enhanced long short-term memory network (LSTM) optimized by Rime Ice Optimization as the forecasting model, validated under five typical extreme weather conditions. The aforementioned methods effectively define

and discriminate extreme weather from a data perspective, mitigating the issue of data scarcity to a certain extent. Reference [14] employed a truncated normal distribution to model extreme weather intensity and introduced Kalman filtering to dynamically correct wind speed, adapting to the non-Gaussian characteristics of wind speed noise during extreme weather. Reference [15] proposed a short-term power forecasting method for distributed PV that integrates a graph attention network (GAT), convolutional neural network (CNN), and spiking neural p systems (SNPSs) to enhance LSTM. This approach utilizes a comprehensive methodology combining spatial correlation mining, meteorological feature extraction, and time-series modeling. These studies address the poor adaptability of renewable energy output forecasting models under extreme weather conditions by attempting to extract and incorporate extreme weather features, thereby enhancing model adaptability and robustness.

However, unlike the effects of cloud cover on solar radiation, dust storms simultaneously scatter and absorb radiation. Furthermore, the variable particle characteristics make it difficult for models to accurately quantify their impact [16]. NWP typically lacks detailed aerosol data, leading to incomplete input information. Moreover, the rapid dynamic changes and strong spatiotemporal heterogeneity of dust make it challenging to capture its variations effectively using a single station. This necessitates upscaling the prediction to a regional level. Consequently, there remains a lack of research on renewable energy output forecasting that accounts for the impact of sand and dust storms. In regional forecasting studies, Reference [17] divided geographical locations into grid cells and established a mapping relationship between gridded irradiance and PV power output using a 3D CNN for short-term power forecasting of distributed PV station clusters. Reference [18] integrated physical principles with data-driven advantages by incorporating solar time and Earth's tilt angle as physical factors, clustering the study area, and applying black-box modeling to enhance the robustness of both regional and single-site PV forecasting. Reference [8] proposed a scenario-based detector placement algorithm that leverages spatiotemporal correlations between adjacent stations to capture cloud movement. Although these studies improve PV forecasting accuracy through image feature extraction, their predictive performance is constrained in desert scenarios where supporting datasets often fail to meet conditions such as consistent solar time and multi-station observations.

In recent years, satellite remote sensing technology, capable of macroscopically observing cloud cover and movement, has been progressively applied in solar radiation and PV output forecasting studies. Reference [19] used a bi-directional extrapolation method to simulate cloud layer movement, incorporating the predicted cloud imagery as input features into a spatial-temporal graph neural network (ST-GNN) for PV output forecasting. Reference [20] focused on dynamic cloud modeling by introducing effective cloud albedo as a core input. It employed a spatiotemporal autoencoder prediction model based on convolutional long short-term memory (ConvLSTM), enabling both deterministic and probabilistic forecasts without requiring ground-based calibration measurements. Reference [21] proposed a fully convolutional neural network (FCNN) model for interpolating NWP irradiance data, addressing the issue of insufficient temporal resolution caused by storage limitations. This model relies solely on extraterrestrial irradiance and does not require complex atmospheric transmission models. Although these models have demonstrated success in capturing key features such as cloud system movement and spatial correlation, revealing significant potential, no studies have yet applied them to desert and Gobi environments. Irradiance exhibits distinct physical properties and attenuation mechanisms during dust storms and cloud motion. Furthermore, challenges persist at the regional scale, including increased model complexity due to dimensionality and limitations in prediction accuracy.

To address the aforementioned challenges, this paper establishes a hybrid data-physics driven model for PV output forecasting. Compared with the existing research, the main contributions of this paper are as follows:

1. The transition from single-station to regional spatiotemporal prediction is fully conducted in this paper, covering the shift from one-dimensional time-series data to two-dimensional spatiotemporal image forecasting using satellite-based global irradiance data, and tracking the evolving trajectory of irradiance under sand and dust weather to deduce its variations.
2. To accommodate the increased forecasting complexity and computational cost resulting from this dimensional expansion, a forecasting model based on CVAE is proposed. This model captures both deterministic trends and stochastic fluctuations in temporal irradiance data, providing essential support for subsequent PV power calculation.
3. A physics-based calculation model for meteorological to PV output conversion is constructed in this paper, considering the isotropic distribution tendency of scattered irradiance and variations in sky clarity under dust weather conditions, with irradiance as inputs, combined with the scale parameters of the photovoltaic plant to output the corresponding power magnitude, thereby addressing the issue of insufficient historical output data for new energy bases.

2. Forecasting Model

2.1. Data-Driven Model Based on Conditional Variational Autoencoders

2.1.1. Principles of CVAE

CVAE consist of three components: a prior network, a posterior network, and a generation network. Their fundamental structure and operational principles are illustrated in Figure 1.

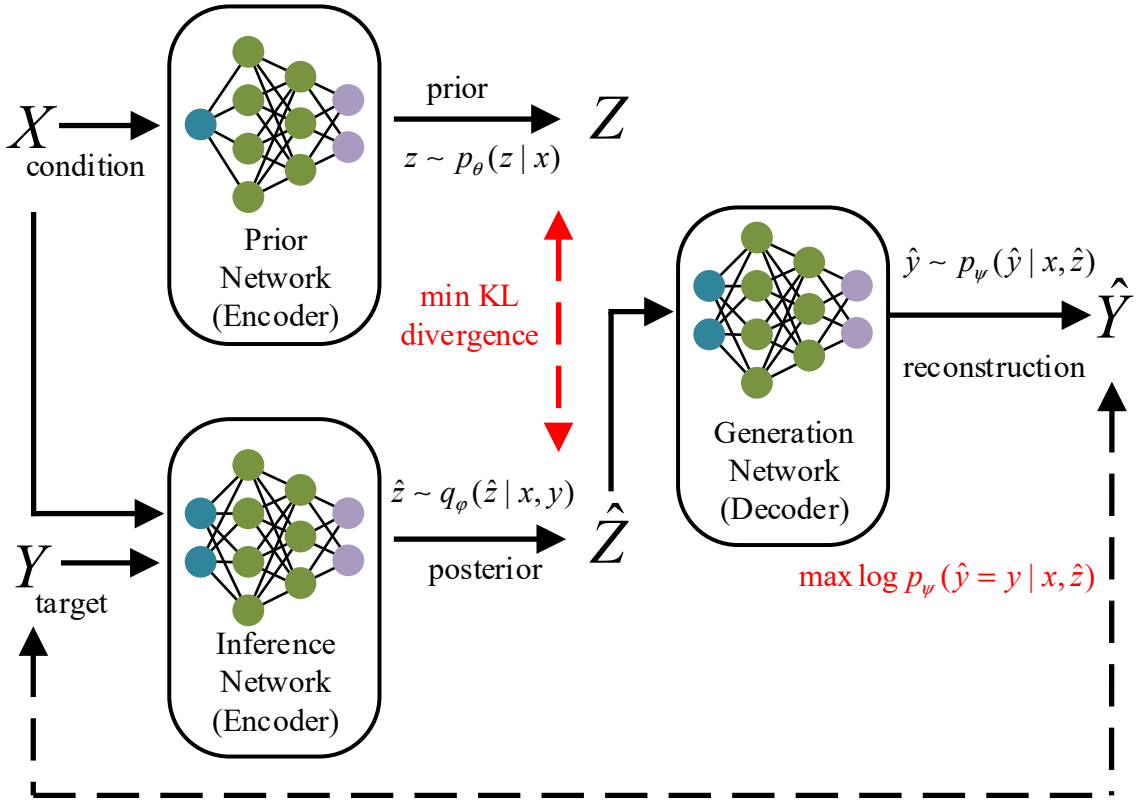


Figure 1. CVAE Structural Principles.

(1) Prior Network

Prior networks define the distribution that latent variables z should follow under a given generative constraint x:

$$z \sim p_\theta(z|x) \tag{1}$$

where p_θ represents the prior network; θ represents the network parameters. In conditional generative models, $p_\theta(z|x)$ is referred to as the prior distribution. The prior network imposes semantic constraints on the latent space, making the latent variable z distribution more concentrated under a given observation x, thereby enhancing the controllability of the generated output.

(2) Inference Network

Inference networks are used to approximate the true posterior distribution, i.e., given observations y and conditions x, they estimate the distribution of latent variables z. Since directly computing the posterior distribution is difficult to handle (typically lacking an analytical solution), CVAE employs a neural network to approximate this distribution:

$$\hat{z} \sim q_\varphi(\hat{z}|x, y) \tag{2}$$

where q_φ represents the inference network; φ represents the network parameters.

(3) Generation Network

Generation network defines the probability distribution of generating observed data y given a latent variable z and a condition x:

$$\hat{y} \sim p_\psi(\hat{y}|x, z) \tag{3}$$

where p_ψ represents the inference network, ψ represents the network parameters. For conditional generative models, the generative network acts as a decoder to model the conditional likelihood, outputting reconstructed data or the distribution parameters of the reconstructed data.

(4) Principle

CVAE aims to learn the conditional distribution $p_\psi(y|x, z)$ of data y given conditional information x. We introduce latent variable z and aim to maximize the conditional log-likelihood $E_{q_\varphi(z|x,y)}\left[\log p_\psi(y|x, z)\right]$. Given the high dimensionality of the integral space, the conditional distribution can become highly complex and analytically intractable, rendering direct optimization infeasible. Therefore, we instead optimize its evidence lower bound (ELBO):

$$\mathrm{ELBO}(\theta, \varphi) = E_{q_\varphi(z|x,y)}\left[\log p_\psi(y|x, z)\right] - \mathrm{KL}(q_\varphi(z|x, y)\|p_\theta(z|x)) \tag{4}$$

where $E_{q_\varphi(z|x,y)}\left[\log p_\psi(y|x, z)\right]$ represents the reconstruction term, denoting the log-likelihood of reconstructing the data y given the conditions x and latent variables z; $-\mathrm{KL}(q_\varphi(z|x, y)\|p_\theta(z|x))$ is the KL divergence term, which encourages the approximate posterior distribution to converge toward the prior distribution. In practical training, the reconstruction term is typically estimated using Monte Carlo methods, transforming the variational lower bound into an empirical lower bound:

$$\hat{L}_{\mathrm{CVAE}}(x, y; \theta, \varphi, \psi) = -\mathrm{KL}(q_\varphi(z|x, y)\|p_\theta(z|x)) + \frac{1}{C}\sum_{c=1}^{C} \log p_\psi\left(y\middle|x, z^{(c)}\right) \tag{5}$$

where C represents the number of samples set during reconstruction.

2.1.2. CVAE Under RSSM Optimization

Although the probabilistic generative framework of CVAE can quantify uncertainty, it exhibits weak dynamic capture capabilities for continuous state spaces, thus tending to lose fine-grained dynamic information due to its reliance on variational posterior approximation. To overcome its limitations in representing continuous states, this paper introduces the RSSM to enhance the temporal dynamic modeling capability of the CVAE, preventing the latent variables from merely capturing discrete or noisy features.

RSSM is a hybrid model combining state space models with deep recurrent networks. It integrates the deterministic dynamic prediction of recurrent neural network (RNN) with the stochastic dynamic prediction of stochastic model (SSM), compressing high-dimensional, complex observational data into a low-dimensional, abstract "state" that encapsulates all critical information. This state accurately represents the true state of observational information while adhering to a simple, predictable dynamic pattern, thereby maintaining high-precision predictive performance. The movement, diffusion, and deposition of dust particles constitute a continuous process with strong spatiotemporal correlations. Irradiance is influenced not only by current dust conditions but also by the historical transport of dust over the preceding hours or even days. The RSSM can encode these high-dimensional observational data into a reduced-dimension state that captures the essential information of the current atmospheric conditions, inferring these key variables from the data. Its structure is shown in Figure 2.

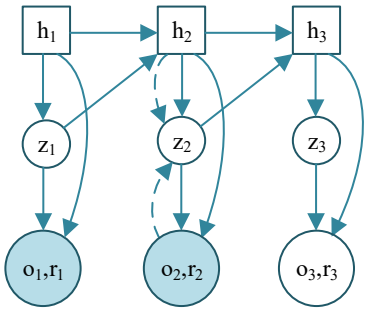


Figure 2. RSSM architecture.

As illustrated in Figure 2, the RSSM decomposes the hidden state into the deterministic component h and the stochastic component z. The deterministic part, analogous to a standard RNN, is updated deterministically based on historical information. The stochastic part captures unpredictable random factors in the environment and is further subdivided into a prior and a posterior. Their expressions are given as follows:

$$RSSM \begin{cases} h_t = f_\phi(h_{t-1}, z_{t-1}) \\ z_t \sim q_\phi(z_t | h_t, o_t) \\ \hat{z}_t \sim p_\phi(\hat{z}_t | h_t) \end{cases} \tag{6}$$

where h_t represents the deterministic state variable at time t; z_t represents the stochastic state variable at time t; o_t represents the observation value at time t. Based on this, we treat the output z from the prior network as input to the inference network, performing RSSM logical optimization. The improved CVAE architecture is shown in Figure 3.

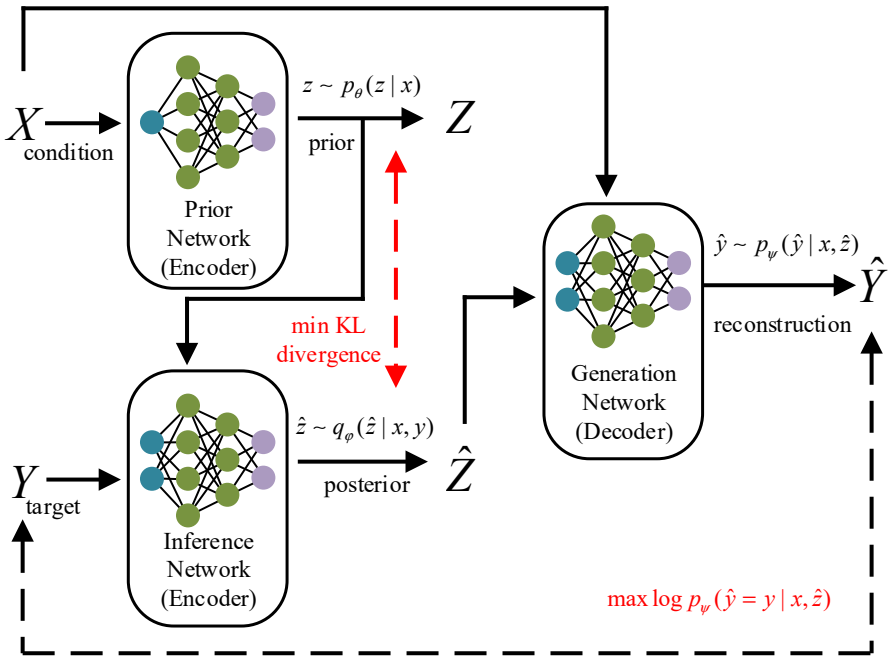


Figure 3. CVAE with RSSM logic optimization.

2.2. Physical Computational Model Based on Photovoltaic Inverters

The power output calculation of a PV system can be formulated as solving a set of coupled equations, treated as a model chain [22] as follows:

$$P_{\mathrm{AC}} = f_{\mathrm{inv}}\left(f_{\mathrm{array}}\left(f_{\mathrm{module}}\left(f_{\mathrm{eff}}\left(f_{\mathrm{poa}}\left(I_{\mathrm{hor}}\right), T_{\mathrm{cell}}\right)\right)\right)\right) \tag{7}$$

This model chain sequentially comprises the calculation of irradiance on the tilted plane, the computation of incident angle losses, the modeling of PV module DC characteristics, the aggregation of PV array power, and the inverter AC conversion.

2.2.1. Calculation of Irradiance on Inclined Surfaces

The global horizontal irradiance data provided by meteorological stations must be converted into the actual irradiance received on the tilted surface of the PV panel. Specifically, the beam irradiance on the tilted surface is calculated through a geometric projection relationship, as follows:

$$\begin{cases} I_{\mathrm{dir,poa}} = DNI \cdot \cos(\mu) \\ \cos(\mu) = \cos(\alpha)\cos(\gamma - \sigma)\sin(\beta) + \sin(\alpha)\sin(\beta) \end{cases} \tag{8}$$

where DNI represents the direct normal irradiance ($\mathrm{W/m^2}$); μ represents the solar incidence angle; α represents the solar elevation angle; β represents the panel tilt angle, σ represents the azimuth angle.

The diffuse irradiance is calculated using the Perez model [23], which decomposes diffuse sunlight into three components: isotropic, circumsolar, and horizon brightness, as follows:

$$I_{\mathrm{dif,poa}} = DHI\left[(1 - F_1)\frac{1 + \cos\beta}{2} + F_1\frac{a}{b} + F_2\cos\beta\right] \tag{9}$$

where DHI represents the diffuse horizontal irradiance ($\mathrm{W/m^2}$); F_1, F_2 are the brightness coefficients, related to sky clarity and brightness; a, b are the geometric factors.

The ground reflected irradiance is calculated using the global horizontal irradiance and the ground albedo, as follows:

$$I_{\text{gnd,poa}} = GHI \cdot \rho_{\text{gnd}} \cdot \frac{1 - \cos\beta}{2} \tag{10}$$

where GHI represents the global horizontal irradiance (W/m^2); ρ_{gnd} is the ground reflection coefficient.

The total irradiance on the tilted surface is the sum of the three component irradiances:

$$I_{\text{poa}} = I_{\text{dir,poa}} + I_{\text{dif,poa}} + I_{\text{gnd,poa}} \tag{11}$$

2.2.2. Calculation of Angle of Incidence Loss

When sunlight does not strike the glass cover plate of the PV module perpendicularly, a portion of the light is reflected away and cannot enter the solar cell to be absorbed for electricity generation. Define IAM as the incident angle modifier, this reflection loss follows the Fresnel equations from optics:

$$\begin{cases} I_{\text{eff}} = I_{\text{poa}} \cdot IAM \\ IAM = 1 - \frac{1}{2}\left(\frac{\sin^2(\mu_r - \mu_i)}{\sin^2(\mu_r + \mu_i)} + \frac{\tan^2(\mu_r - \mu_i)}{\tan^2(\mu_r - \mu_i)}\right) \end{cases} \tag{12}$$

where I_{eff} is the irradiance effectively absorbed by the battery (W/m^2); μ_r is the angle of refraction; μ_i is the angle of incidence.

2.2.3. Modeling of PV Panel DC Characteristics

The DC output power of the PV module under standard test conditions (STC) must be calculated based on the irradiance on the tilted surface, ambient temperature, and wind speed:

$$\begin{cases} T_{\text{cell}} = T_{\text{amb}} + \Delta T \\ \Delta T = \frac{I_{\text{poa}} \cdot e^{a+b \cdot v}}{I_0} \end{cases} \tag{13}$$

where T_{cell} is the cell temperature (°C); T_{amb} is the ambient temperature (°C); ΔT is the temperature rise in the solar cell relative to the ambient environment (°C); a, b are empirical coefficients; v represents wind speed (m/s); I_0 represents reverse saturation current (A).

The parameters of the PV cell model are iteratively solved using the STC parameters and temperature coefficients:

$$\begin{cases} I = I_{\text{ph}} - I_0 \cdot e^{\frac{V + I \cdot R_s}{n N_s V_t} - 1} - \frac{V + I \cdot R_s}{R_{\text{sh}}} \\ V_t = k \cdot T_{\text{cell}} \cdot q \end{cases} \tag{14}$$

where I_{ph} represents photoelectric current (A); R_{s} is the series resistance (Ω); n is the diode factor; V_{t} is the thermal voltage (V); k is the Boltzmann constant; q is the elementary charge.

2.2.4. Power Aggregation of PV Arrays

When N_{mod} PV modules are connected in series, the total voltage equals the sum of the individual module voltages, while the current is limited by the worst-performing module in the string:

$$\begin{cases} V_{\text{string}} = \sum_{i=1}^{N_{\text{mod}}} V_i \\ I_{\text{string}} = \min(I_1, I_2, \ldots, I_N) \end{cases} \tag{15}$$

N_{string} series-parallel connections, with the total current being the sum of the currents in each series connection:

$$\begin{cases} P_{DC} = V_{\text{array}} \cdot I_{\text{array}} \\ I_{\text{array}} = \sum_{j=1}^{N_{\text{string}}} I_j \\ V_{\text{array}} = V_{\text{string}} \end{cases} \tag{16}$$

2.2.5. Inverter AC Conversion

Calculate AC power using the Sandia inverter model:

$$\begin{cases} P_{AC} = P_{DC} \cdot \frac{P_{AC0}}{P_{DC0}} \cdot \left[k_0 + k_1 P'_{DC} + k_2 P'^2_{DC}\right] \\ P'_{DC} = \frac{P_{DC}}{P_{DC0}} \end{cases} \tag{17}$$

where k_1, k_2, k_3 are the efficiency curve coefficient.

2.3. Hybrid Forecasting Framework

Figure 4 illustrates the overall workflow of the proposed hybrid forecasting framework, which is designed to decouple meteorological evolution modeling from PV power conversion. The framework consists of two main stages: (1) spatiotemporal prediction of regional irradiance using the improved RSSM-CVAE model, and (2) physical conversion from irradiance to PV power output. Hourly regional irradiance data cubes, including GHI, DNI, and DHI, are obtained from Fengyun satellite products. Each time step is represented as a 256 × 256 × 3 image, corresponding to the three irradiance components. Prior to training, all irradiance values are quality-controlled and normalized to [0, 1] using Min–Max scaling. A rolling time-window strategy is employed to construct supervised samples, where three consecutive historical irradiance images are used as inputs to predict the irradiance image at the next hour. This input–output configuration balances temporal dependency modeling and computational efficiency, and is suitable for short-term forecasting horizons under rapidly evolving dust conditions. After obtaining the regional irradiance prediction map, the geographic coordinates (longitude and latitude) of each PV plant are projected onto the satellite image grid using the official Fengyun calibration parameters. The corresponding pixel indices are determined through nearest-neighbor mapping. The predicted values of GHI, DNI, and DHI at these pixel locations are then extracted from each forecasted image, forming a one-dimensional future irradiance time series for the target plant. This irradiance sequence is subsequently fed into the physical PV power modeling chain described in Section 2.2. Specifically, the Perez diffuse irradiance decomposition model is applied to account for anisotropic scattering under dust conditions, followed by PV module DC modeling and inverter AC conversion to generate the final power output forecast. By separating irradiance evolution modeling from power conversion, the framework enhances interpretability and allows physical constraints to be explicitly embedded in the forecasting process.

This study focuses on relatively flat and homogeneous desert and Gobi regions where large-scale PV bases cover extensive areas and microclimatic variations within a plant footprint are small compared with the satellite pixel scale. Therefore, a single representative pixel is used to characterize the irradiance conditions of each PV station. For regions with complex terrain, strong surface heterogeneity, or highly dispersed distributed PV clusters, a spatial aggregation strategy based on multiple pixels or weighted averaging would be required, which will be investigated in future work.

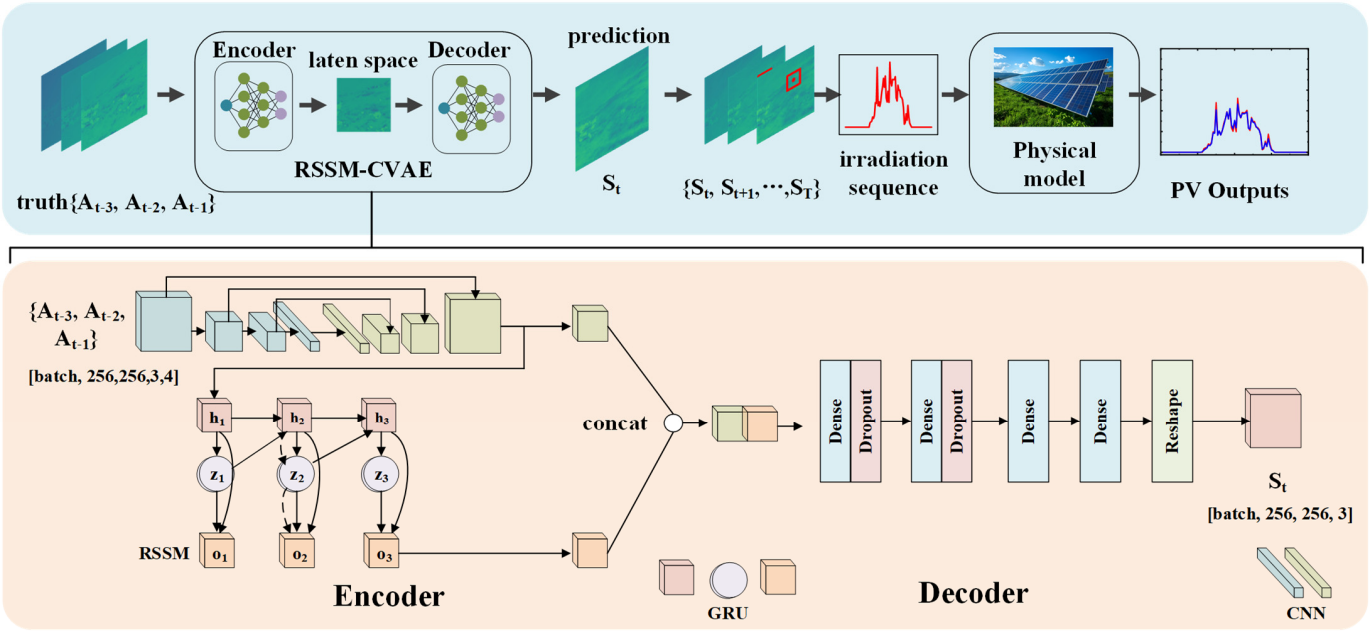


Figure 4. Schematic Diagram of the Hybrid Forecasting Framework.

3. Experimental Results and Analysis

The data-driven meteorological data used in this study are sourced from the National Satellite Meteorological Center of China (NSMC), specifically including AGRI dust detection and AGRI surface incident solar radiation data [24], with a temporal resolution of 1 h per sample. The temporal range of dust event samples was selected from the atmospheric environment bulletin provided by the China Meteorological Administration (CMA) government portal, while the spatial range covers the Xinjiang region of China, which experiences the most frequent dust events. The physics-based PV data are derived from the photovoltaic power output dataset available in the Science Data Bank (ScienceDB) [25], which has a temporal resolution of 15 min per sample and covers PV power generation and measured meteorological data for the period 2018–2019.

3.1. Satellite Data Preprocessing

An initial screening was performed using the provided quality control flags inherent in the satellite data to identify and clean abnormal data points in the meteorological satellite irradiance measurements. These anomalies, including fill values, invalid values, and spatiotemporal inconsistencies caused by sensor malfunctions, were subjected to data cleansing and local neighborhood anomaly detection. Invalid data were removed to minimize their impact on the performance of the forecasting model proposed in this study. The processed dataset was then normalized using the Min-Max normalization method, which linearly scales the data to the range [0, 1] to eliminate the influence of differing dimensional units, as shown in the equation:

$$x_{norm} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{18}$$

3.2. Evaluation Indicators

To validate the performance of the proposed method, this paper employs two standardized metrics: the mean absolute error (MAE) and root mean square error (RMSE). Let $L_i(j)$ and $\hat{L}_i(j)$ denote the true value and predicted value of surface solar irradiance at

spatial position j in the i-th sample, respectively. For a test set containing m samples, each with an image size of h × w, the MAE and RMSE are defined as follows:

$$RMSE = \frac{1}{m}\sum_{i=1}^{m}\sqrt{\frac{1}{w \times h}\sum_{j=1}^{w \times h}\left(L_i(j) - \hat{L}_i(j)\right)^2} \tag{19}$$

$$MAE = \frac{1}{m} \cdot \frac{1}{w \times h}\sum_{i=1}^{m}\sum_{j=1}^{w \times h}\left|L_i(j) - \hat{L}_i(j)\right| \tag{20}$$

To facilitate comparison of relative error levels across different regions or irradiance components, this study simultaneously calculated the normalized mean absolute error (NMAE) and normalized root mean square error (NRMSE). The normalization factor is the average value L_{ref} of all true values in the test set.

$$NRMSE(\%) = \frac{RMSE}{L_{ref}} \times 100\% \tag{21}$$

$$NMAE(\%) = \frac{MAE}{L_{ref}} \times 100\% \tag{22}$$

where L_{ref} is the reference value of irradiance, provided by the Fengyun satellite product documentation, with a value of 1500 W/m^2. This eliminates the influence of irradiance units, allowing the relative magnitude of errors to be expressed as percentages and ensures comparability of results across different studies or conditions.

3.3. Regional Irradiation Forecast

In this study, the optimized CVAE model is employed to predict regional irradiance data. All case studies were implemented in a TensorFlow 2.0 and Python 3.8 environment. The model architecture and parameter settings are summarized in Table 1.

Table 1. Model structure and parameter settings.

Hyperparameters	Value
Batch size	32
Length	256
Width	256
Channel	3
Horizon	4
Learning rate	3×10^{-5}

The frequency of dust storm events from 2019 to 2023 showed a trend of decreasing first and then increasing, as shown in Table 2. Among these years, 2019 and 2023 had the highest frequencies. Only the 2022 dust storm event exhibited characteristics such as a late onset time, weak intensity, and a small affected area; events in other years were similar. Regarding specific indicators, ozone concentrations showed little overall variation across the five years, while PM_{10} concentrations were higher only in 2019 compared to the other four years. Furthermore, considering the common data splitting ratios for training and testing sets and the need for consistency between prediction models and physical verification timelines, this study adopted the following data partitioning strategy: the 2019 sample data was selected as an independent test set, while data from 2020 to 2023 served as the training set.

Table 2. Annual sandstorm indicator statistics.

Year	Times	PM_{10} (μg/m^3)	O_3 (μg/m^3)
2019	15	74	134
2020	10	68	129
2021	13	66	132
2022	10	67	136
2023	17	63	132

The spatial scope covers the area between 42.82° N and 29.13° N latitude, and 73.86° E and 91.65° E longitude, encompassing southern Xinjiang and parts of Tibet. The experimental results are as follows.

Figures 5–7 present the prediction results of three types of regional irradiance—GHI, DNI, and DHI—based on the proposed method. As can be observed, GHI exhibits the highest overall irradiance, with peak values in desert regions reaching up to 1100 W/m^2. This is followed by DNI, which shows a strong correlation with solar elevation angle, peaking at noon when the solar path is shortest and decaying rapidly during early morning or late afternoon when the solar angle is low. In contrast, DHI increases gradually after sunrise as the solar elevation angle rises.

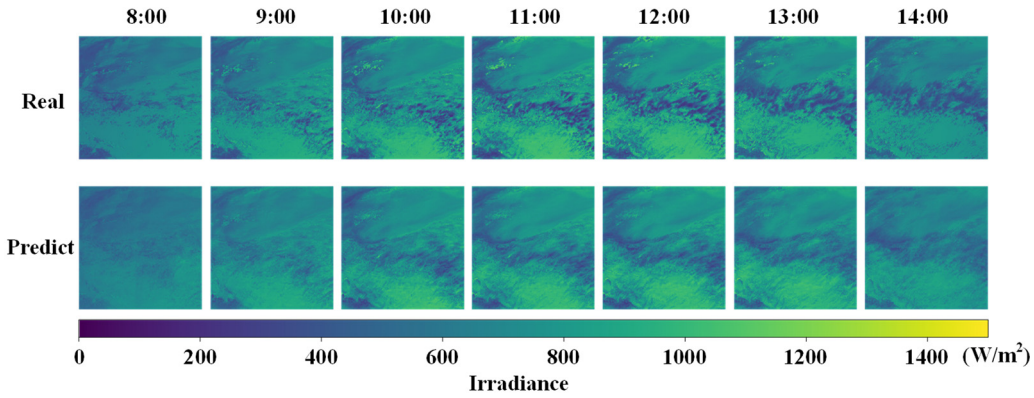


Figure 5. GHI calculation results for 18 May 2019.

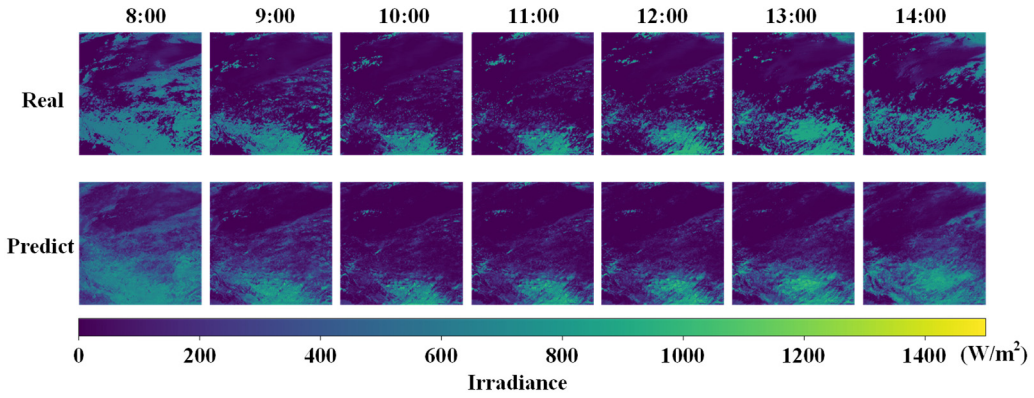


Figure 6. DNI calculation results for 18 May 2019.

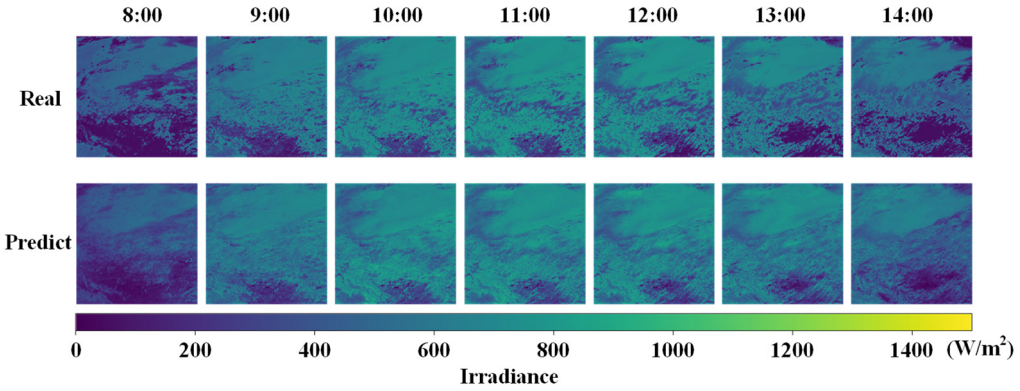


Figure 7. DHI calculation results for 18 May 2019.

In order to quantitatively validate the physical mechanisms of the attenuation and scattering effects of dust aerosols on solar radiation, we conducted a further analysis of dust index products provided by Fengyun satellites during the same period. These data were temporally and spatially matched with irradiance data, with values ranging from 0 to 24, where higher values indicate a greater concentration of dust in the atmospheric column of the corresponding region. Figure 8 shows the spatial distribution of dust monitoring indices at three irradiance measurement points in time. A significant spatial coupling relationship is evident: areas of high dust concentration overlap considerably with regions of low DNI values. Within the Taklamakan Desert (the red-boxed area in Figure 8), dust indices generally exceed 16, corresponding to pronounced troughs in DNI values. This confirms the strong attenuation effect of dust aerosols on direct solar radiation: dust particles reduce DNI while enhancing DHI through multiple scattering. When dust concentrations are sufficiently high, the enhanced scattering radiation effect causes DHI to exceed DNI, which is consistent with the predicted results.

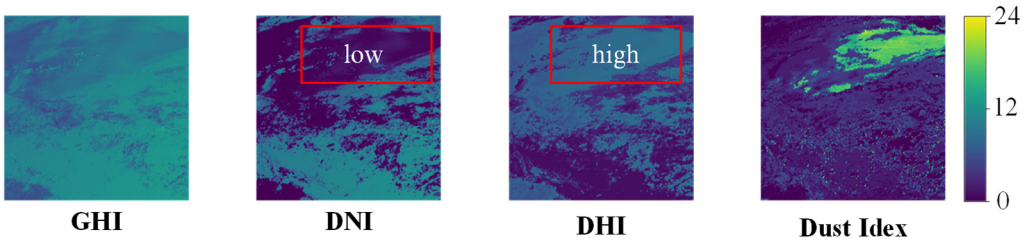


Figure 8. Detail comparison and dust index.

It can be observed that the proposed prediction method effectively captures irradiance variation trends under dust weather conditions. To visually highlight subtle differences that are challenging to detect through direct visual comparison between real and predicted images, a differential calculation is performed between the ground truth and prediction results. This emphasizes discrepant regions and quantifies local errors, with the calculation expressed as follows:

$$diff = I_{\text{real}} - I_{\text{predict}} \tag{23}$$

The differential calculation results are shown in Figure 9, and as shown in Table 3, the average differential values across all regional points are 55.903 W/m^2 for GHI, 96.538 W/m^2 for DNI, and 82.913 W/m^2 for DHI. Based on the valid range of 0–1500 W/m^2 for surface solar incident radiation provided by the NSMC, the percentage error metrics were calculated as shown in Table 3. The validation results indicate that the NMAE for GHI is 3.727%, with the NRMSE of 5.351%; for DNI, the NMAE is 6.436%, with the NRMSE of 10.015%; and for DHI, the NMAE is 5.258%, with the NRMSE of 7.450%. It can be observed

that the predictions for DHI perform the best, followed by GHI. These results demonstrate that the proposed method achieves low error in regional predictions for all three types of irradiances.

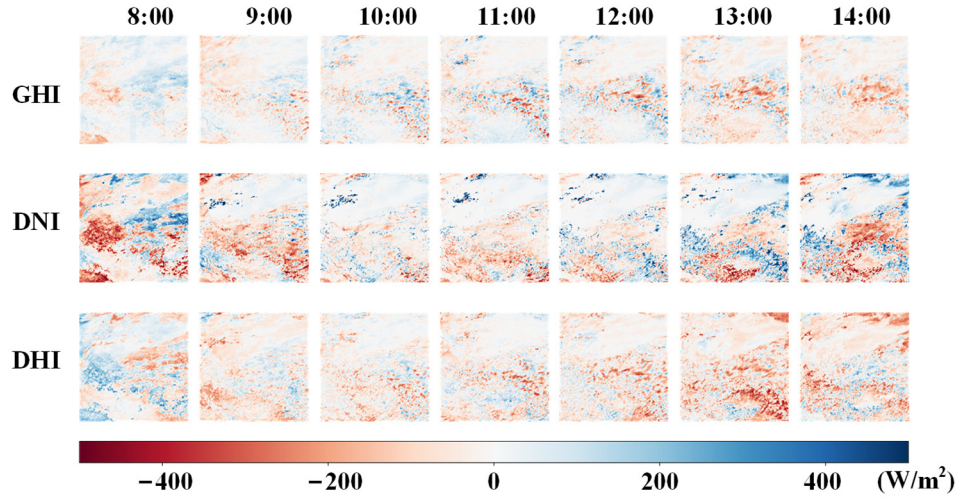


Figure 9. Differential calculations for three irradiation predictions.

Table 3. Three prediction indicators for irradiation forecasting.

Irradiation	MAE (W/m^2)	NMAE/%	RMSE	NRMSE
GHI	55.903	3.727	80.269	5.351
DNI	96.538	6.436	150.218	10.015
DHI	82.913	5.528	111.741	7.450

To further assess the reliability of the model proposed in this paper, we quantified the uncertainty of its predictions. For each test sample, 100 samples were drawn from the model to generate multiple prediction trajectories. For each prediction point, the mean, standard deviation were calculated and compared with different methods. The comparison results are shown in Table 4.

Table 4. Comparison of errors in different models.

model	MAE (W/m^2)		
	GHI	DNI	DHI
Transformer	94.636 ± 0.475	122.158 ± 0.366	131.246 ± 0.549
RSSM-CVAE	55.903 ± 1.667	96.538 ± 1.344	82.913 ± 2.240
model	**RMSE (W/m^2)**		
	GHI	DNI	DHI
Transformer	122.060 ± 0.554	190.271 ± 0.473	178.689 ± 0.687
RSSM-CVAE	80.269 ± 1.818	150.218 ± 1.827	111.741 ± 2.871

In terms of average prediction error, the RSSM-CVAE model achieved lower MAE values than the Transformer model for all three irradiance forecasts, with respective reductions of 40.9%, 21.0% and 36.8%. The mean RMSE was also lower than that of the Transformer model by 34.2%, 21.1% and 37.5%, respectively. In terms of prediction stability, the Transformer model demonstrated high stability by exhibiting minimal fluctuation in multiple prediction results. However, this came at the cost of higher average errors.

3.4. Ablation Experiments and Analysis

The standard CVAE tends to lose fine-grained dynamic information when processing continuous spatiotemporal sequences. The RSSM specifically addresses this issue by introducing separate deterministic and stochastic states. The deterministic state explicitly models the memory effects of physical processes, while the stochastic state captures uncertain fluctuations, such as turbulence. In order to validate the necessity of each core component of the proposed framework and to assess the incremental value of the recurrent state space model relative to standard conditional variational autoencoders, systematic ablation experiments were designed to compare the performance of the full model with that of baseline models from which RSSM components had been removed.

As shown in Table 5, introducing the RSSM module significantly and consistently enhances the performance of all three types of irradiance prediction. The MAE and RMSE of the complete model are lower than those of the baseline model for all metrics. Notably, the model shows particularly significant improvements for GHI and DHI, with reductions in MAE of 31.5% and 33.1%, respectively. This suggests that the RSSM module substantially improves the modeling capability for global radiation trends and complex scattering processes. While DNI's MAE decreased significantly by 19.3%, its RMSE improved by a relatively smaller amount of 7.7%. RMSE penalizes large errors through its squared term, meaning rare but highly erroneous DNI prediction instances contribute disproportionately to RMSE. CVAE-RSSM model significantly improves the overall DNI trend and mean error by leveraging the deterministic nature of RSSM. However, while its predictive capability for these extreme, sudden attenuation events has improved, it cannot eliminate all such occurrences. Consequently, the percentage improvement in RMSE is diluted by these residual, unavoidable large errors.

Table 5. Ablation experiment results.

Irradiance	Model	MAE	RMSE
GHI	CVAE	81.614	104.096
	RSSM-CVAE	55.902	80.268
	Error Reduction Ratio	31.5%	22.9%
DNI	CVAE	119.556	162.728
	RSSM-CVAE	96.538	150.218
	Error Reduction Ratio	19.3%	7.7%
DHI	CVAE	124.012	158.837
	RSSM-CVAE	82.913	111.741
	Error Reduction Ratio	33.1%	29.7%

3.5. Model Robustness Analysis

To evaluate the model's adaptability to dust storm events across different years, annual tests were conducted. The model's structure and hyperparameters were fixed, with each individual year forming the test set and all other years forming the training set. NMAE and NRMSE were then calculated for the model on each test set. The results are presented in Table 6.

The model demonstrates strong predictive performance in years with moderate to high-intensity dust storms and widespread impacts. For example, taking GHI, its NMAE remained stable between 3.25% and 5.71% across the four years, while its NRMSE fluctuated between 4.39% and 7.56%, indicating reasonable variability. However, in 2021, when dust activity was weaker, errors increased slightly. This suggests that the model's predictive capability for less frequent, lower-intensity dust events in the training data could be improved. This is a clear limitation of the model, as its performance depends on the representativeness of the training data for the forecasted weather conditions.

Table 6. Comparison of errors when different years serve as test sets.

Test Year	NMAE/%			NRMSE/%			Intensity of Dust Storm Activity
	GHI	DNI	DHI	GHI	DNI	DHI	
2019	3.727	6.436	5.528	5.351	10.015	7.450	high intensity and wide impact
2020	4.247	7.057	4.955	5.369	9.799	6.732	middle intensity and wide impact
2021	5.705	6.024	6.080	7.564	8.610	8.189	high intensity and wide impact
2022	8.391	9.639	7.8708	10.439	12.731	10.4167	low intensity and limited impact
2023	3.247	7.581	6.474	4.394	10.548	8.870	high intensity and wide impact

3.6. Validation of the Complete Hybrid Forecasting Pipeline

This study selected four dust storm events occurring in the region where the PV power station is located during 2019 from the Chinese meteorological yearbook report [26] for validation. To maintain data authenticity and avoid introducing interpolation errors, we opted for temporal alignment to match higher-resolution photovoltaic power data with satellite data. Although this method loses sub-hour-scale fluctuation information, it effectively captures core dynamic changes. Given that this study focuses on hourly scale forecasting and that dust events typically persist for several hours, this approach constitutes a reasonable choice for the research objectives. The predicted GHI, DNI and DHI data are input into the physical model, which uses the Perez model to calculate photovoltaic power generation under diffuse radiation conditions on an inclined surface. This model provides precise full-sky simulations by decomposing diffuse radiation into three components and integrating sky brightness with the transparency coefficient. This makes the model applicable to various meteorological conditions. In order to quantitatively evaluate the superiority of the proposed hybrid-driven framework over conventional methods, it is compared with two typical benchmark approaches: (1) Pure data-driven method: A spatiotemporal prediction model based on Transformers. (2) Physics-based method: A physical computation chain utilizing NWP irradiance inputs and the Perez model. The computational results are shown in Figure 10 and error calculation results are shown in Table 7.

It can be seen that the hybrid driving method proposed in this paper significantly outperforms both single methods in terms of MAE and RMSE metrics. Compared to the purely data-driven method, the MAE decreased by 0.5548 MW and the RMSE by 0.7309 MW. Compared to the physics-only method, the MAE decreased by 0.9878 MW and the RMSE by 1.5743 MW. The hybrid method exhibits a greater reduction in RMSE than in MAE, indicating a more pronounced improvement in handling extreme errors. The purely physics-driven method yields the highest errors, with the RMSE being particularly pronounced. This is primarily due to significant uncertainties in NWP forecasts of aerosol concentration and distribution during extreme weather events, such as sandstorms. This leads to systematic biases in the input irradiance data. The data-driven approach outperformed the purely physics-driven method, demonstrating its ability to learn effective spatial and temporal patterns from historical data. However, its performance is limited in two ways. First, the scarcity of historical operational data from desert PV plants restricts the expressive power of data-driven models. Second, as a deterministic model, the Transformer struggles to generalize when encountering sudden dust events that are not sufficiently covered in the training data. It also fails to provide quantifiable information on prediction uncertainty.

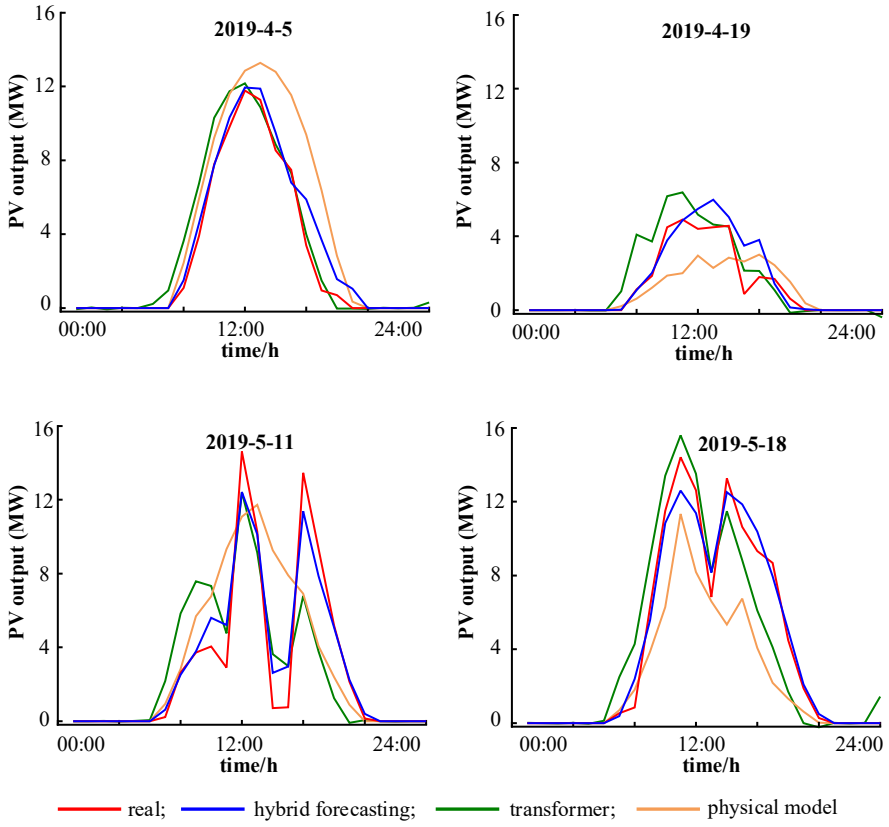


Figure 10. Comparison of Hybrid Forecasting with Data-Driven and Physics-based Models.

Table 7. Errors of hybrid forecasting with data-driven and physics-based models.

Method	MAE/MW	RMSE/MW
transformer	1.0528	1.6095
physical model	1.4858	2.4532
hybrid forecasting	0.4980	0.8789

3.7. Verification of Scattered Irradiance on Inclined Surfaces

To analyze and compare the completeness of the physical meaning and the generalizability of the Perez model, the corresponding PV power under the Haydavies, Klucher, and Reindl scattering irradiance decomposition and conversion models was verified through comparison. The comparison results are shown in Table 8. Specifically, Haydavies divides DHI into isotropic and circadian components, offering simplified calculations with significant improvements over isotropic models. Reindl builds upon Haydavies by incorporating anisotropic treatment of ground-reflected radiation, while Klucher enhances Haydavies with a horizon ring scattering correction, demonstrating strong performance under moderate scattering conditions.

Table 8. Evaluation metrics for each model under dust storm.

Model	MAE/MW	MAPE/%	RMSE/MW
Perez	0.4980	53.9607	0.8789
Klucher	0.4922	57.0695	0.8965
Haydavies	0.5130	55.9739	0.9006
Reindl	0.5041	54.3634	0.8827

Under dust storm conditions, it can be observed that the Perez model demonstrates optimal performance in terms of MAPE and RMSE compared to other models. Meanwhile, the Klucher model exhibits superior performance in MAE. Taking into account the physical universality of the models and their performance across multiple metrics, this study selects the Perez model as the physical core.

4. Conclusions

This study addresses the challenges of historical data scarcity, abrupt changes in meteorological factors, and strong spatiotemporal heterogeneity in forecasting the performance and base PV output in desert and Gobi regions during sand and dust storms. A hybrid data-physics driven regional PV output forecasting method is proposed. The main conclusions are as follows:

(1) The proposed hybrid-driven framework accurately predicts the spatiotemporal evolution of irradiance through an improved data-driven CVAE model, and subsequently converts irradiance to PV output using a physics-based model, effectively combining the strengths of both paradigms. This framework not only enhances model robustness in data-scarce scenarios but also improves interpretability and adaptability to the specific physical processes of sand and dust weather by incorporating physical mechanisms.

(2) By integrating a RSSM into the CVAE to form the core temporal forecasting component, the non-stationarity and high stochasticity of irradiance sequences under sand and dust weather are successfully addressed.

(3) By establishing a physics-based calculation model that accounts for sand and dust weather conditions, the predicted irradiance is reliably mapped to PV output. This provides a feasible technical solution for the stable early-stage operation of renewable energy bases in desert and Gobi regions.

5. Discussions

Compared to existing studies on the forecasting of photovoltaic power generation, this framework offers several significant advantages in the event of a dust storm. Most prior research has focused on cloud-dominated attenuation mechanisms, primarily relying on ground-based observational data or purely data-driven learning strategies [13,15,19,20]. While these methods perform well under normal or cloudy conditions, they exhibit significantly reduced robustness when aerosol scattering and absorption dominate radiation transport processes, particularly in desert regions where there is limited historical operational data.

As shown in Table 9, existing regional-scale methods based on satellite imagery primarily simulate cloud motion and spatial correlations, rather than explicitly incorporating the physical processes that convert irradiance into photovoltaic power generation. In contrast, this study integrates a generative spatiotemporal forecasting model with a physics-based photovoltaic power calculation chain. The RSSM-CVAE module simultaneously captures the deterministic transport patterns and random fluctuations of dust aerosols. Meanwhile, the physical model explicitly accounts for isotropic scattering tendencies and variations in sky transparency under dusty conditions. This hybrid design ensures the framework's effectiveness even when historical power samples are scarce.

The experimental results further validate this advantage. As shown in Table 7, during dust storm events, the proposed hybrid framework achieved reductions in MAE and RMSE of 52.7% and 45.4%, respectively, compared to the purely data-driven Transformer model. Compared to the purely physical model, the reductions were 66.5% and 64.2%, respectively. Furthermore, the annual robustness assessment in Table 6 shows that, during severe

dust years, the NMSE for GHI consistently remained below 6%, demonstrating stable performance under intense aerosol perturbations. These results confirm that integrating probabilistic spatial-temporal forecasting with physical power conversion significantly improves robustness in extreme dust conditions. From an applicability perspective, this method requires only two input datasets that are readily available: (i) regional-scale, time-series meteorological data, such as satellite-derived irradiance products; (ii) the fundamental physical parameters of the photovoltaic power plant. Consequently, this approach is well-suited to large-scale PV bases in arid and semi-arid regions, where ground-based observations and historical operational records are limited. In these regions, dust aerosols dominate the radiation attenuation process and their scattering and absorption mechanisms align with the physical assumptions adopted in this study.

Table 9. Comparison between related works and the proposed method.

Comparison Dimension	Method	Target Scenario	Main Advantages	Limitations/Challenges
[13]	GAN data augmentation + LSTM prediction	Multiple extreme events for power forecasting	1. Effectively mitigates sample scarcity for extreme weather	1. Generated data quality depends on training 2. Difficult to ensure physical consistency
[15]	Integrates GAT, CNN, SNPSs to enhance LSTM for spatial-temporal-feature fusion	Distributed PV under multiple extreme weather types	1. Comprehensive feature fusion 2. Enhanced robustness for multiple extreme weather types	1. Requires extensive labeled data from distributed sites 2. Complex model integration may increase computational cost
[19]	Bi-directional cloud motion extrapolation + ST-GNN	PV forecasting under cloud movement	1. Explicitly models cloud layer movement 2. ST-GNN effectively captures spatiotemporal correlations between PV sites	1. Performance highly dependent on accuracy of cloud motion extrapolation 2. Requires historical data from multiple stations to establish spatial correlations
[20]	ConvLSTM-based satellite image sequence prediction	General cloud movement for PV forecasting under various weather	1. Excels at capturing spatiotemporal features of cloud movement 2. Well-established technique	1. Requires substantial historical data 2. Performance degrades during abrupt weather changes
This work	CVAE-RSSM + physical model	Dust storms in desert regions for regional PV power forecasting	1. Balances accuracy and physical interpretability 2. Addresses data scarcity issue 3. Provides uncertainty information	Generalization to weak dust events needs improvement

However, certain limitations warrant attention. The Perez diffuse irradiance model used in this study is inherently empirical and was calibrated using statistical observational data. If the dominant aerosol type in the target region differs significantly from desert dust, for example, in coastal areas with marine aerosols or industrial zones with complex anthropogenic particulates, the empirical coefficients may no longer be applicable. In such cases, the Perez model would require regional recalibration or the adoption of more sophisticated radiative transfer models. Furthermore, integrating aerosol optical depth products with multispectral satellite data in future research could enhance the physical consistency of the irradiance-to-power conversion process under various extreme meteorological conditions.

Author Contributions: Conceptualization, S.C. and S.W.; methodology, S.W. and J.X.; investigation, S.W. and S.C.; resources, S.W., J.C. and H.Y.; writing—original draft preparation, S.C., S.W. and J.X.; writing—review and editing, S.C. and S.W.; visualization, S.C.; supervision, S.C.; project administration, S.C. and S.W.; funding acquisition, S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (52577129) and the National Key Research and Development Program of China (2025ZD0807300).

Data Availability Statement: The raw data supporting this study are from the following public third-party sources: 1. Satellite-derived irradiance and dust detection data: FengYun-4A AGRI product, obtained from the National Satellite Meteorological Center (NSMC) of China (http://satellite.nsmc.org.cn/ (accessed on 2 December 2025)). 2. Dust event chronology: Selected from the annual "Atmospheric Environment Bulletin" published by the China Meteorological Administration (https://www.cma.gov.cn/en/ (accessed on 2 December 2025)). 3. Photovoltaic power and meteorological measurements: PVOD v1.0 dataset, hosted on the Science Data Bank (ScienceDB) (https://www.scidb.cn/en (accessed on 9 December 2025)). The processed datasets generated during the study, including the spatiotemporally aligned irradiance-dust-PV power dataset, the extracted features for model input, and the model prediction results, are not publicly available due to their large size and integration complexity. However, they can be made available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Alanzi, S.S.; Aldalali, B.; Kamel, R.M. Effects of Sandstorms on Hybrid Renewable Energy Sources and Load Demand in Arid Desert Climates: A Case Study. *Energy Sustain. Dev.* **2024**, *81*, 101473. [CrossRef]
2. Micheli, L.; Almonacid, F.; Bessa, J.G.; Fernández-Solas, Á.; Fernández, E.F. The Impact of Extreme Dust Storms on the National Photovoltaic Energy Supply. *Sustain. Energy Technol. Assess.* **2024**, *62*, 103607. [CrossRef]
3. Zheng, L.; Su, R.; Sun, X.; Guo, S. Historical PV-Output Characteristic Extraction Based Weather-Type Classification Strategy and Its Forecasting Method for the Day-Ahead Prediction of PV Output. *Energy* **2023**, *271*, 127009. [CrossRef]
4. Liu, Z.-F.; Li, L.-L.; Tseng, M.-L.; Lim, M.K. Prediction Short-Term Photovoltaic Power Using Improved Chicken Swarm Optimizer—Extreme Learning Machine Model. *J. Clean. Prod.* **2020**, *248*, 119272. [CrossRef]
5. Raza, M.Q.; Nadarajah, M.; Ekanayake, C. On Recent Advances in PV Output Power Forecast. *Sol. Energy* **2016**, *136*, 125–144. [CrossRef]
6. Pazikadin, A.R.; Rifai, D.; Ali, K.; Malik, M.Z.; Abdalla, A.N.; Faraj, M.A. Solar Irradiance Measurement Instrumentation and Power Solar Generation Forecasting Based on Artificial Neural Networks (ANN): A Review of Five Years Research Trend. *Sci. Total Environ.* **2020**, *715*, 136848. [CrossRef]
7. Ameur, A.; Berrada, A.; Loudiyi, K.; Aggour, M. Forecast Modeling and Performance Assessment of Solar PV Systems. *J. Clean. Prod.* **2020**, *267*, 122167. [CrossRef]
8. Pfenninger, S.; Staffell, I. Long-Term Patterns of European PV Output Using 30 Years of Validated Hourly Reanalysis and Satellite Data. *Energy* **2016**, *114*, 1251–1265. [CrossRef]
9. Wang, B.; Lv, Y.; Chen, Z. Hybrid Mechanism-Data-Driven Short-Term Power Forecasting of Distributed Photovoltaic Considering Information Time Shift. *Autom. Electr. Power Syst.* **2022**, *46*, 67–74.
10. Li, Y.; Song, L.; Zhang, S.; Kraus, L.; Adcox, T.; Willardson, R.; Komandur, A.; Lu, N. A TCN-Based Hybrid Forecasting Framework for Hours-Ahead Utility-Scale PV Forecasting. *IEEE Trans. Smart Grid* **2023**, *14*, 4073–4085. [CrossRef]
11. Ye, L.; Li, Y.; Pei, M.; Li, Z.; Xu, X.; Lu, J. Combined Approach for Short-Term Wind Power Forecasting under Cold Weather with Small Sample. *Proc. CSEE* **2023**, *43*, 543–554.
12. Li, Y.; Chen, F.; Yan, J.; Ge, C.; Han, S.; Liu, Y. Adaptive Short-Term Wind Power Forecasting for Extreme Weather Based on Transfer Learning and AutoEncoder. *Autom. Electr. Power Syst.* **2025**, *49*, 85–95.
13. Ma, Y.; Huang, Y.; Shi, W.; Yuan, Y. A Short-Term Prediction Method for Distributed Photovoltaic Power Considering Extreme Weather Events. *Sol. Energy* **2025**, *302*, 114078. [CrossRef]
14. Zhang, Y.; Su, H.; Wang, R.; Deng, J.; Wang, Y.; Guo, W.; Li, R. Short-Term Forecast Method of Wind Power Output Based on Multi-Scale CNN-LSTM in Extreme Weather. *Int. J. Electr. Power Energy Syst.* **2025**, *172*, 111191. [CrossRef]
15. Guan, X.; Han, X.; Wang, J.; Wang, T. A Novel Short-Term Prediction Method for Distributed Photovoltaic Power Generation Considering Extreme Weather. *Eng. Appl. Artif. Intell.* **2025**, *162*, 112540. [CrossRef]

16. Wang, H.; Sun, F.; Liu, W. Spatial and Temporal Patterns as Well as Major Influencing Factors of Global and Diffuse Horizontal Irradiance over China: 1960–2014. *Sol. Energy* **2018**, *159*, 601–615. [CrossRef]
17. Qiao, Y.; Sun, R.; Ding, R.; Li, S.; Lu, Z. Distributed Photovoltaic Station Cluster Short-Term Power Forecasting Part II: Gridding Prediction. *Power Syst. Technol.* **2021**, *45*, 2210–2218.
18. Zhang, X.; Li, Y.; Lu, S.; Hamann, H.F.; Hodge, B.-M.; Lehman, B. A Solar Time Based Analog Ensemble Method for Regional Solar Power Forecasting. *IEEE Trans. Sustain. Energy* **2019**, *10*, 268–279. [CrossRef]
19. Cheng, L.; Zang, H.; Wei, Z.; Ding, T.; Sun, G. Solar Power Prediction Based on Satellite Measurements—A Graphical Learning Method for Tracking Cloud Motion. *IEEE Trans. Power Syst.* **2022**, *37*, 2335–2345. [CrossRef]
20. Nielsen, A.H.; Iosifidis, A.; Karstoft, H. IrradianceNet: Spatiotemporal Deep Learning Model for Satellite-Derived Solar Irradiance Short-Term Forecasting. *Sol. Energy* **2021**, *228*, 659–669. [CrossRef]
21. Zech, M.; Hammer, A.; Von Bremen, L. A Fully Convolutional Neural Network to Interpolate Solar Irradiation NWP Ensemble Forecasts. *Sol. Energy* **2025**, *301*, 113865. [CrossRef]
22. Holmgren, W.F.; Hansen, C.W.; Mikofski, M.A. Pvlib Python: A Python Package for Modeling Solar Energy Systems. *JOSS* **2018**, *3*, 884. [CrossRef]
23. Perez, R.; Seals, R.; Ineichen, P.; Stewart, R.; Menicucci, D. A New Simplified Version of the Perez Diffuse Irradiance Model for Tilted Surfaces. *Sol. Energy* **1987**, *39*, 221–231. [CrossRef]
24. National Satellite Meteorological Center of China. Available online: http://nsmc.org.cn/nsmc/cn/home/index.html (accessed on 2 December 2025).
25. PVOD v1.0: A Photovoltaic Power Output Dataset. Available online: https://www.scidb.cn/en/detail?dataSetId=f8f3d7af144f441795c5781497e56b62 (accessed on 9 December 2025).
26. China Atmospheric Environment Meteorological Bulletin. 2019. Available online: https://www.cma.gov.cn/zfxxgk/gknr/qxbg/202301/t20230119_5273437.html (accessed on 2 December 2025).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.