energies

MDPI

Article

Improving Solar Radiation Forecasting in Cloudy Conditions by Integrating Satellite Observations

Qiangsheng Bu [1], Shuyi Zhuang [1], Fei Luo [1], Zhigang Ye [1,*], Yubo Yuan [1], Tianrui Ma [2] and Tao Da [2]

[1] State Grid Jiangsu Electric Power Company Ltd. Research Institute, Nanjing 211103, China; 15105168877@163.com (Q.B.); zhuangshuyi2014@126.com (S.Z.); lf941102@163.com (F.L.); yuanyubo@js.sgcc.com.cn (Y.Y.)

[2] State Grid Jiangsu Electric Power Company Ltd. Zhenjiang Power Supply Branch, Zhenjiang 212002, China; matr1@js.sgcc.com.cn (T.M.); zjg_dt@js.sgcc.com.cn (T.D.)

* Correspondence: yzhggoodluck@hotmail.com

Abstract: Solar radiation forecasting is the basis of building a robust solar power system. Most ground-based forecasting methods are unable to consider the impact of cloud changes on future solar radiation. To alleviate this limitation, this study develops a hybrid network which relies on a convolutional neural network to extract cloud motion patterns from time series of satellite observations and a long short-term memory neural network to establish the relationship between future solar radiation and cloud information, as well as antecedent measurements. We carefully select the optimal scales to consider the spatial and temporal correlations of solar radiation and design test experiments at ten stations to check the model performance in various climate zones. The results demonstrate that the solar radiation forecasting accuracy is considerably improved, particularly in cloudy conditions, compared with purely ground-based models. The maximum magnitude of improvements reaches up to 50 W/m^2 (15%) in terms of the (relative) root mean squared error (RMSE) for 1 h ahead forecasts. The network achieves superior forecasts with correlation coefficients varying from 0.96 at 1 h ahead to 0.85 at 6 h ahead. Forecast errors are related to cloud regimes, of which the cloud amount leads to a maximum relative RMSE difference of about 50% with an additional 5% from cloud variability. This study ascertains that multi-source data fusion contributes to a better simulation of cloud impacts and a combination of different deep learning techniques enables more reliable forecasts of solar radiation. In addition, multi-step forecasts with a low latency make the advance planning and management of solar energy possible in practical applications.

Keywords: solar radiation forecasting; convolutional neural network; cloud amount; temporal and spatial scale; solar energy

Citation: Bu, Q.; Zhuang, S.; Luo, F.; Ye, Z.; Yuan, Y.; Ma, T.; Da, T. Improving Solar Radiation Forecasting in Cloudy Conditions by Integrating Satellite Observations. *Energies* **2024**, *17*, 6222. https://doi.org/10.3390/en17246222

Academic Editors: J. Carlos García-Díaz and Óscar Trull

Received: 19 November 2024
Revised: 4 December 2024
Accepted: 9 December 2024
Published: 10 December 2024

1. Introduction

Promoting the widespread adoption of solar energy is a crucial strategy for reducing greenhouse gas emissions and addressing global climate challenges [1]. However, unlike traditional power generation methods such as hydropower [2] and fuel-based power generation [3], solar photovoltaic (PV) power is characterized by a high variability and intermittency due to the uncertainty of solar radiation [4,5]. To enhance the stability of PV power systems and increase their share in the energy grid, the reliable and accurate forecasting of PV power production is essential [6–8]. Short-term forecasting also facilitates the intelligent management of energy storage systems [9] and solar-integrated grids [10], thereby reducing the operational costs of PV systems [11].

Surface solar radiation is the most critical factor that affects PV power output [12,13]; hence, solar radiation forecasting has been extensively studied. For day-ahead forecasts (24–48 h), numerical weather prediction (NWP) models are preferred [13,14]. These models infer the local cloud probability and solar radiation transmissivity through dynamic atmospheric modeling [15], and tend to be more accurate than statistical methods beyond a

forecast lead time of five to six hours [14,16]. Intra-day solar radiation forecasts (1–6 h) can be addressed using statistical, physical, or hybrid methods. Statistical techniques leverage past solar radiation time series to extract temporal patterns and project them into the future, e.g., the Box–Jenkins model [17], auto-regressive integrated moving average (ARIMA) [18], and exponential smoothing [19]. For sub-hourly forecasts, statistical methods based on ground measurements perform well [14]. Reikard [20] showed that the autoregressive moving average (ARMA) model outperforms other approaches in predicting global horizontal solar radiation at intervals of 5, 15, 30, and 60 min. Statistical methods are simple, but often lack generalizability. In contrast, physical approaches offer a better generalization but are limited by their complexity and computational demands. Hybrid forecasting models have emerged as promising solutions, with machine learning techniques gaining popularity in recent years. For example, Lauret et al. [21] compared the performance of artificial neural networks (ANNs), Gaussian processes (GPs), and support vector regression (SVR) for solar radiation forecasting, finding that machine learning models outperformed their simpler counterparts when the lead times exceeded one hour, especially under unstable sky conditions. More examples can be found in Refs. [22,23]. In general, a substantial body of research demonstrates the superior performance of machine learning methods, which can easily adapt to various locations and climatic conditions.

The stochastic nature of solar radiation disrupts the learning process of machine learning-based models, leading to higher forecasting errors [24]. Long short-term memory (LSTM) networks, which retain historical information for future predictions, help to mitigate this issue to some extent [7,25–27]. For example, Ghimire et al. [25] developed a hybrid framework where a convolutional neural network (CNN) extracts features related to solar radiation changes, and LSTM is used to incorporate these features for half-hourly forecasts, achieving superior results with over 70% of predictive errors falling below 10 W/m^2. Liu et al. [28] believed that a lack of spatial variations in solar radiation hinders the progress of short-term forecasting; hence, they embedded solar radiation and meteorological data from multiple stations into a spatial grid for prediction.

These forecasting methods are predominantly ground-based, relying on local measurements. However, fluctuations in surface solar radiation are not only influenced by the surrounding environment, but also by temporal and spatial variations in the atmosphere [29–31]. Clouds located far from the point of interest can affect local weather conditions within tens of minutes to a few hours [32,33]. Therefore, the integration of real-time spatial data, such as satellite imagery, plays a critical role in enhancing solar radiation forecasting models. Geostationary satellites provide images of cloud fields with refresh rates between 10 and 30 min [29,34]. This is especially important in conditions where clouds rapidly form and move, affecting solar radiation in ways that are difficult to predict using only historical ground data. By incorporating satellite imagery, models can gain a more comprehensive understanding of the spatial and temporal dynamics influencing solar radiation, thereby improving short-term forecasting accuracy. It has been demonstrated that incorporating cloud indices or atmospheric parameters derived from satellite images into models such as ARMA [35,36] and LSTM [37] can improve the quality of solar radiation forecasts. For short-term forecasts under clear skies, biases are relatively low (typically less than 10%); however, for all-sky conditions, the relative root mean square errors range from 15% to 70% [23]. Therefore, improving forecasting accuracy is a continuous pursuit and an open research challenge. Deep learning techniques are favored due to their ability to model sequential data and capture spatial contexts [38]. Several studies have demonstrated the advantages of deep learning approaches (e.g., ANN [21,39,40], CNN [25,28], and LSTM [7,25,37]) for solar radiation forecasting.

Based on these findings, this study aims to develop a predictive network that integrates both satellite observations and ground-based measurements to improve solar radiation forecasting, particularly in cloudy conditions. In this study, a one-dimensional (1D) CNN is employed to extract temporal features from ground-based solar radiation time series, while a three-dimensional (3D) CNN models the long-range spatial connections between

surface solar radiation and clouds distant from the target location based on sequences of geostationary satellite images. Finally, LSTM is utilized to establish the relationships between future solar radiation and its preceding temporal and spatial features across multiple timescales. The main contributions include the following: (1) the fusion of spatial and temporal data from multiple sources, which improves solar radiation forecasts, especially in regions with significant cloud variability; (2) the development of a hybrid CNN–LSTM model that captures both spatial features from satellite images and temporal trends from ground measurements, enabling accurate multi-step forecasts; and (3) extensive validation across various climatic regions, demonstrating a superior performance compared to existing models. This approach provides valuable insights for improving solar radiation forecasting, with implications for renewable energy applications and climate monitoring.

2. Materials and Methods

2.1. Case Studies

Our study was conducted at ten international radiation stations (Table 1) in China, which are distributed in various climatic regions. Each station is representative of the background climate. Hourly measurements of global solar radiation in 2007 and 2008 were collected from China Meteorological Administration (CMA) at http://www.cma.gov.cn/ (accessed on 2 December 2024). The data in 2007 were used to prepare training samples, and the data in 2008 were used for independent model validation. The data quality of ground measurements was checked using a physical threshold test, i.e., records whose clearness index (the ratio of the measured solar radiation to the amount of extraterrestrial solar radiation on a horizontal surface) was less than 0.03 or larger than 0.9 were excluded [41].

Table 1. Basic information of hourly solar radiation datasets of the ten international stations in China.

Station Name	Latitude	Longitude	Altitude (m)	Num. of Training Samples	Num. of Validation Samples
Beijing	39°48′	116°28′	31.3	4073	3730
Shenyang	41°44′	123°31′	49.0	4312	3970
Harbin	45°45′	126°46′	142.3	4280	3958
Shanghai	31°24′	121°27′	5.5	4059	3880
Wuhan	30°37′	114°08′	23.1	3585	3226
Guangzhou	23°10′	113°20′	41.0	3803	3365
Wenjiang	30°42′	103°50′	539.3	3895	3646
Kunming	25°00′	102°39′	1886.5	4084	3766
Yuzhong	35°52′	104°09′	1874.4	4210	3942
Urumqi	43°47′	87°39′	935.0	4159	3847

Spatial observations corresponding to the ground measurements were from Multi-functional Transport Satellites (MTSATs) operated by the Japan Meteorological Agency. MTSAT-1R, launched in February 2005 and positioned at 104° E above the equator, scans the Asia–Pacific region at a spatial resolution of 1 km in the visible channel and 4 km in infrared channels with a time interval of 30 min. The visible channel (0.55–0.80 μm) of MTSAT-1R satellite images was utilized to obtain multi-scale spatial–temporal features of cloud fields. Geo-coordinate mapped images within 70° N–20° S and 70° E–160° E at a 1/20° resolution are available from http://weather.is.kochi-u.ac.jp/ (accessed on 2 December 2024).

Time series of ground measurements and satellite observations were employed to forecast m-hour-ahead solar radiation. The forecasting problem can be formally expressed as follows:

$$(\hat{x}_{t+1}, \hat{x}_{t+2}, \cdots, \hat{x}_{t+m}) = f\left(x_t, x_{t-1}, \cdots, x_{t-T+1}, s_t^{h\times w}, s_{t-1}^{h\times w}, \ldots, s_{t-T+1}^{h\times w}\right), \quad (1)$$

where $\hat{x}_{t+i} (i = 1, 2, \cdots, m)$ is the solar radiation to be predicted, with m representing the maximum forecast lead time, x_t is the actual solar radiation at time t, $s_t^{h\times w}$ is the satellite image block of size $h \times w$ at time t, and T denotes the lagged time length.

2.2. Hybrid Forecasting Model

2.2.1. Overall Structure

In this study, we develop a hybrid forecasting network (HModel) to integrate satellite observations and ground measurements for solar radiation forecasting. Satellite images allow us to model the effects of distant clouds on local solar radiation. These clouds, which may be several hundred kilometers away, can significantly affect local radiation levels due to cloud shadowing, reflection, and scattering. By including satellite data in the hybrid model, we capture the non-local spatial correlations between surface solar radiation and clouds. This incorporation of real-time spatial data from satellites enables more accurate short-term solar radiation forecasting, particularly in the presence of dynamic cloud cover. Ground-based measurements offer accurate temporal data. These temporal features, in conjunction with the spatial data from satellites, allow the model to account for both short-term variations in solar radiation and large-scale weather patterns that influence the solar radiation at each location.

HModel mainly consists of a CNN for feature extraction and LSTM for forecasting. Figure 1 shows an example with $T = 6$, $h = w = 16$, and $m = 1$ to illustrate the revolution of extracted features. There are two input sequences, corresponding to the antecedent time series of satellite observations and ground measurements, respectively. Since ground measurements are the average of a time interval (e.g., one hour), while satellite images are instantaneous, two images corresponding to the starting and ending times of the measurement, respectively, are input into HModel for prediction. The number of convolution kernels of the three 3D/1D convolutional layers is 64, 128, and 256, in order. The size of the 3D kernels is set to extract features in 3×3 receptive field across 3 time horizons. The kernels of the 1D convolutional layers also have a time horizon of 3 h. The rectified linear unit (ReLU), which thresholds the non-positive value as zero and keeps the positive value unchanged, is used as the activation function in HModel. ReLU is effective in alleviating the vanishing gradient problem and speeding up the learning process [42]. Max-pooling (MP) is utilized to downsample feature representation by taking the maximum value over a nonoverlapping window of 2×2 in space. Global max-pooling (GAP) is adopted to flatten spatial features into one dimension by calculating the mean of each feature. The dimensionality of the LSTM output space is 100, thus, the number of input neurons in the FC layer equals 100. The dropout rate is set to be 0.3, meaning that dropout randomly sets 30% of hidden neurons in the FC layer to zero during training. The dropped neurons do not contribute to the forward pass and are not used in the back-propagation procedure.

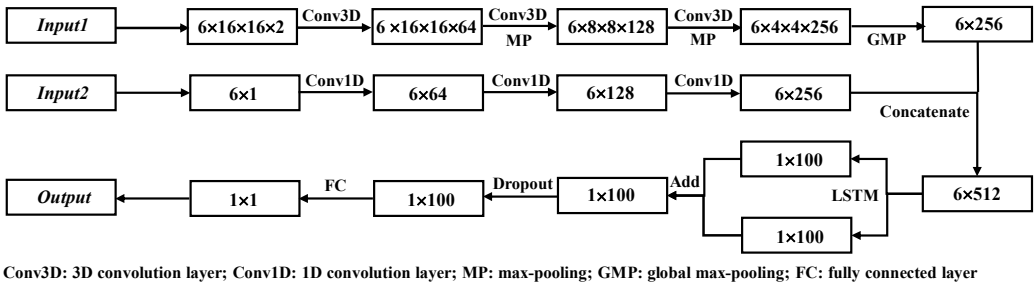

Figure 1. Structure of the proposed HModel. This figure shows an example with $T = 6$, $h = w = 16$, and $m = 1$.

We combine ground measurement data and satellite observations based on their corresponding geographic coordinates (latitude and longitude) and timestamp information. Specifically, we collect hourly measurements for the years 2007 and 2008. HModel is built at each station using measurements taken in 2007 and validated in 2008. All samples are prepared with $T = 12$ and $m = 12$; in this way, models for different forecast horizons or with different input lengths are trained and validated using an equal number of samples to ensure a fair comparison. Samples containing missing or unqualified records are deleted.

The sample number for each station is provided in Table 1. Ground measurements are paired with satellite observations based on the same coordinates, ensuring that each observation corresponds to the correct satellite image and ground measurement at the same time. Satellite input data (i.e., $s_t^{h \times w}$) are transferred into reflectance in advance. Input and target measured data (i.e., x_t and $\hat{x}_t$) are normalized into the range of [0–1] as follows:

$$X_{normalized} = (X_{actual} - X_{min})/(X_{max} - X_{min}) \quad (2)$$

$$X_{actual} = X_{output} \times (X_{max} - X_{min}) + X_{min}, \quad (3)$$

where X denotes the set of all measured x_t and $\hat{x}_t$. During the training process, 80% of the training samples are randomly selected to determine the model parameters, and the remaining are used to identify whether the network is overfitted. Meanwhile, an early stopping mechanism monitors the model performance on the remaining samples and relinquishes the optimization when the performance ceases to sufficiently improve or even degrades with further optimization, thus equipping HModel with the ability to prevent overfitting. The weights and biases of all kernels are initialized by Xavier [43]. The Adam optimizer [44] is used to find the optimal weights and biases that minimize the loss function of the mean squared error (MSE) over the training samples. The learning rate of Adam is originally set as 0.001 and multiplied by 0.5 when encountering a 10 epochs' plateau of the model performance.

HModel is constructed using the Keras (version 2.12.0) library [45] written in Python (version 3.11.3) with a TensorFlow (version 2.12.0) backend in our experiments. The aforementioned configurations of model hyperparameters, except T, h, w, and m, are determined by the tree-structured Parzen estimator (TPE) algorithm based on Bayesian optimization [46], with $T = 6$, $h = w = 16$, and $m = 1$. The search spaces for TPE are listed in Table 2. Because early stopping is adopted, it is not necessary to control the training epoch carefully; hence, a relatively large value of 500 is set. The choice of T, h, w, and m is presented in the following.

Table 2. Parameter configurations of HModel. Except epochs, T, h, w, and m, the optimal parameter is selected among multiple candidates using TPE algorithm. The bold one is finally used.

Parameter Name	Search Space
Numbers of convolution kernels	[64-64-64, 64-128-128, **64-128-256**, 64-256-256, 64-256-512]
Kernel size in space ($p \times q$)	[**3 × 3**, 3 × 5, 5 × 3, 5 × 5, 5 × 7, 7 × 5, 7 × 7]
Kernel size in time (l)	[1, 2, **3**, 4, 5, 6]
Activation function	[**ReLU**, Tanh, Sigmoid]
LSTM cell units	[20, 50, **100**, 200, 512]
Dropout rate	[0.1, 0.2, **0.3**, 0.5]
Optimizer	[SGD, RMSprop, **Adam**, Adagrad]
Learning rate	[Constant value of 0.001, **initialized as 0.001 and multiplied by 0.5 after 10 epochs' plateau of model performance**, initialized as 0.01 and multiplied by 0.2 after 10 epochs' plateau of validation loss, initialized as 0.01 and multiplied by 0.1 after 10 epochs' plateau of validation loss]
Batch size	[**100**, 200, 300, 500]
LSTM cell units	[20, 50, **100**, 200, 512]

2.2.2. Temporal Scale

The first step in applying HModel for forecasting is to determine the most relevant antecedent time series as an input variable. Two statistical methods, i.e., the autocorrelation function (ACF) and partial autocorrelation function (PACF), are adopted to select the lags of solar radiation that have a significant correlation with the prediction target. The

autocorrelation reveals the linear features of the time series, while nonlinear processes are possibly involved in solar radiation variations [25]. The ACF and PACF values with a 95% confidence for the hourly time series are calculated at lags of 0–26 h. Figure 2a,b show an example at Beijing station (the curves for other stations are similar, but with different magnitudes). Both ACF and PACF are periodic. ACF reaches the minimum value and PACF falls into the 95% confidence interval at lag 6 for the first time. Therefore, $T = 6$ might be a rational choice for HModel to predict the value at the next step. It also indicates that future values which have a strong correlation with the current value (x_t) are $\hat{x}_{t+1}, \hat{x}_{t+2}, \cdots, \hat{x}_{t+6}$, that is, it is appropriate to set the maximum forecast horizon at $m = 6$.

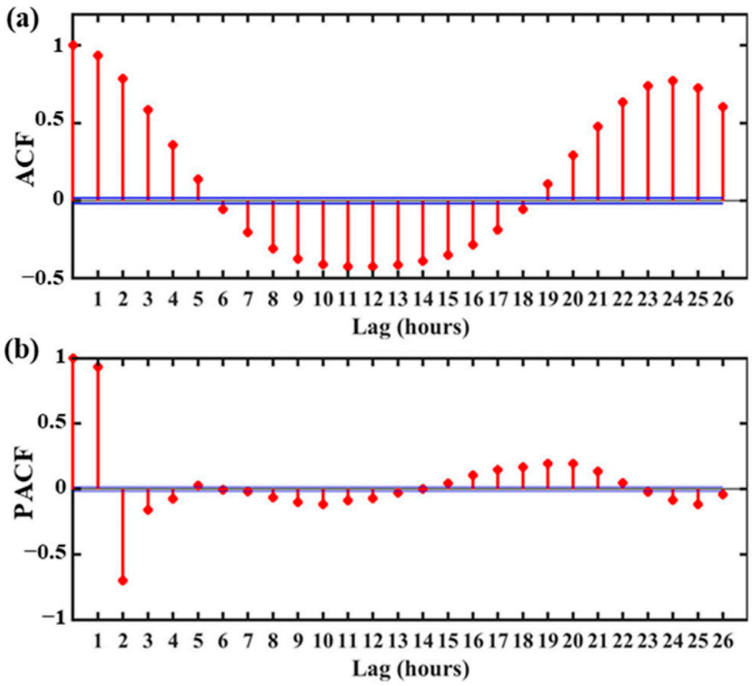

Figure 2. ACF (**a**) and PACF (**b**) of hourly solar radiation time series at Beijing station. Blue lines denote the 95% confidence intervals.

2.2.3. Spatial Scale

Previous studies revealed that integrating spatial information into satellite-based methods can improve the accuracy of solar radiation estimation [28], because it contributes to a better simulation of the spatial effects of clouds [47,48], e.g., nonlocal cloud shadows, cloud-side reflections, and cloud scattering. Therefore, in the design of HModel, satellite image blocks are taken as one of the inputs, expected to simulate the long-range spatial connections between future solar radiation and clouds far from the point of interest. This raises the question of choosing the most optimal spatial scale. We extract satellite time series for a 33×33 grid centered around each station to analyze the correlation of surrounding pixels with the central pixel, which has a direct impact on the surface solar radiation. The time series includes hourly satellite images from 2007 to 2008, comprising a total of 17,544 (731×24) satellite images for analysis. Figure 3a shows that pixels in a certain spatial extent are more relevant to the central pixel, and Figure 3b clearly illustrates the steep decline in correlation with an increasing distance. It also indicates that the optimal scale varies among stations, of which Beijing station has the largest scale. Therefore, we implement a validation experiment at Beijing station. HModel is configured with $T = 6$, $h = w = 32$, and $m = 1$. For each run, the input image blocks are zero-masked to ensure that the valid size ($h' \times w'$) varies from 4×4 to 32×32 with an interval of 4. The model performance in terms of R and RMSE reaches the highest at $h' = w' = 16$ in the validation phase (Figure 3c). Further expansion of the spatial scale seems to have a limited impact on the model performance, like the results in Ref [49]. Based on the experiment, the spatial

scale of HModel is set as $h = w = 16$ (~80 km). This configuration agrees with previous proposal to consider the spatial adjacency effects of solar radiation in a range of about 60 km [49].

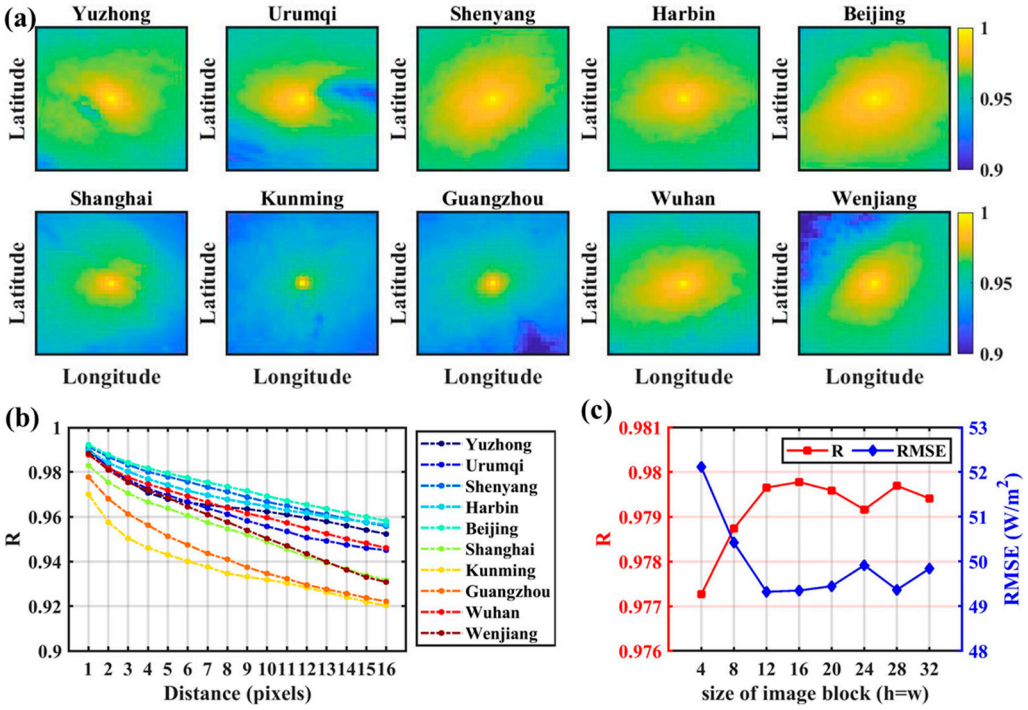

Figure 3. Time series correlation (**a**) of satellite-observed signals between the central and surrounding pixels, average R variations (**b**) as a function of distance to the central pixel, and model performance (**c**) in terms of R and RMSE versus different spatial scales at Beijing station.

2.3. Model Evaluation

To verify the advantage of integrating satellite observations with ground measurements, HModel is compared with three baseline models. The first is the convolutional long short-term memory (CLSTM) neural network proposed by Ghimire et al. [25], where a 1D CNN layer is employed to extract features from time series of ground measurements and an LSTM layer is used to incorporate these features for the prediction of solar radiation. The structure of CLSTM is like that of HModel, without the input sequence of the satellite image block. The second is the two-dimensional convolutional LSTM (ConvLSTM) network developed by Shi et al. [50]. Its structure is like HModel without inputting measured time series, i.e., ConvLSTM forecasts solar radiation based solely on satellite observations. The third is the persistence forecasting model (PFM), which assumes that clouds do not change throughout the forecast lead time and only models the changes occurring due to the deterministic course of the sun. This study uses the clearness-based PFM [51], defined as follows:

$$\hat{x}_{t+1} = K_t \times I_{t+1},\ K_t = x_t / I_t, \tag{4}$$

where I represents the extraterrestrial solar radiation and K is the clearness index, which roughly represents the cloud transmittance of solar radiation. K is nearly zero for overcast conditions and one for cloudless skies.

The performances of different models are measured using several statistical metrics, including R, RMSE, relative RMSE (rRMSE), mean bias error (MBE), and mean absolute error (MAE) between measurements and predictions. Their mathematical expressions are given as follows:

$$\mathrm{R} = \frac{\sum_{t=1}^{n}\left(\hat{x}_t - \overline{\hat{x}}\right)\left(x_t - \overline{x}\right)}{\sqrt{\sum_{t=1}^{n}\left(\hat{x}_t - \overline{\hat{x}}\right)^2}\sqrt{\sum_{t=1}^{n}\left(x_t - \overline{x}\right)^2}}, \tag{5}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{x}_t - x_t)^2}, \mathrm{rRMSE} = \frac{\mathrm{RMSE}}{\overline{x}}, \quad (6)$$

$$\mathrm{MBE} = \frac{1}{n}\sum_{t=1}^{n}(\hat{x}_t - x_t), \quad (7)$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n}|\hat{x}_t - x_t|, \quad (8)$$

where n is the number of data samples, x represents actual measurements whose mean value is $\overline{x}$, and $\hat{x}$ represents forecasting results whose mean value is $\overline{\hat{x}}$.

3. Results

3.1. Accuracy of One-Hour-Ahead Forecasting

Figure 4 shows one-hour-ahead (i.e., $m = 1$) solar radiation forecast errors at ten locations in different climate zones with respect to various statistical metrics. CLSTM maps the preceding fluctuations in surface solar radiation to the future and exhibits an enhanced forecasting skill compared with static PFM. Regarding ConvLSTM, while it does leverage satellite imagery to capture the dynamic changes in cloud cover, its performance is comparatively poorer than CLSTM due to the absence of ground-based measurements. This is a critical limitation, because ConvLSTM relies solely on temporal information from satellite images, which, although helpful for tracking cloud movements, cannot establish a direct and quantitative relationship between cloud variation and surface solar radiation fluctuations. Without this relationship, the model is unable to accurately predict the impact of clouds on solar radiation at the ground level. In contrast, our HModel integrates both satellite imagery and ground-based measurements, allowing it to capture both the spatial and temporal dynamics of solar radiation more effectively. The fusion of these data sources enhances the model's ability to forecast solar radiation. According to Figure 4, HModel has the best performance in all cases, demonstrating that the integration of satellite observations and ground measurements does contribute to improving the predictive capability.

Cross-comparisons of these models reveal that satellite observations assist in modelling the impact of clouds on future solar radiation and ground measurements provide the most important and reliable information for forecasting. In essence, HModel is equivalent to a combination of ConvLSTM and CLSTM, which not only considers the impact of future clouds, but also carries out real-time calibration according to previous time series. All in all, it can be concluded that the predictability of solar radiation decreases with an increased occurrence of clouds, and in parallel, the room for satellite-based improvement increases. The spatially and temporally united forecast rRMSE varied from about 15% for prevailing clear skies to more than 25% for locations dominated by cloudy conditions and decreased by about 10% and 20% compared with CLSTM and ConvLSTM, respectively.

3.2. Forecast Skill Over Multi-Step Horizon

In addition to one-hour-ahead forecasting, it is also vital to explore the potential utility of predictive models over a multi-step horizon. This information is immensely useful in designing a robust solar power system where key decisions regarding the sustainability and future availability of solar energy can be made well in advance [13,22]. We implement a 6 h forecasting test using CLSTM and HModel, that is, the forecast horizon is set as m = 6 (refer to Section 2.2.1). It should be noted that the model simultaneously outputs a sequence of six predicted values, rather than iteratively outputting them one by one (e.g., Refs. [7,25]). Figure 5 shows the average R, RMSE, MBE, and MAE of the ten stations for different forecast lead times, and Table 3 details the specific values for each station. The forecasting skill decreases with an increasing forecast lead time. The R value of HModel (CLSTM) decreases from 0.96 (0.94) at 1 h ahead to 0.85 (0.80) at 6 h ahead. This happens because the spatial–temporal correlation of current and future weather patterns drops off. The rate of error increase slows down gradually, as an increasing number of samples have a value of 0 in the later hours. HModel provides better forecasts for all forecast lead times

at all stations (Table 3), and its margin of error growth is lower than that of CLSTM. This implies that HModel has a stronger and more durable predictive ability. On average, the improved magnitude of HModel in terms of RMSE (rRMSE) increases from 12 W/m^2 (2%) at 1 h ahead to 28 W/m^2 (6%) at 6 h ahead compared with CLSTM. The MBE values indicate that CLSTM tends to overestimate due to a lack of cloud impact, while HModel tends to underestimate, possibly because cloud impacts are slightly exaggerated.

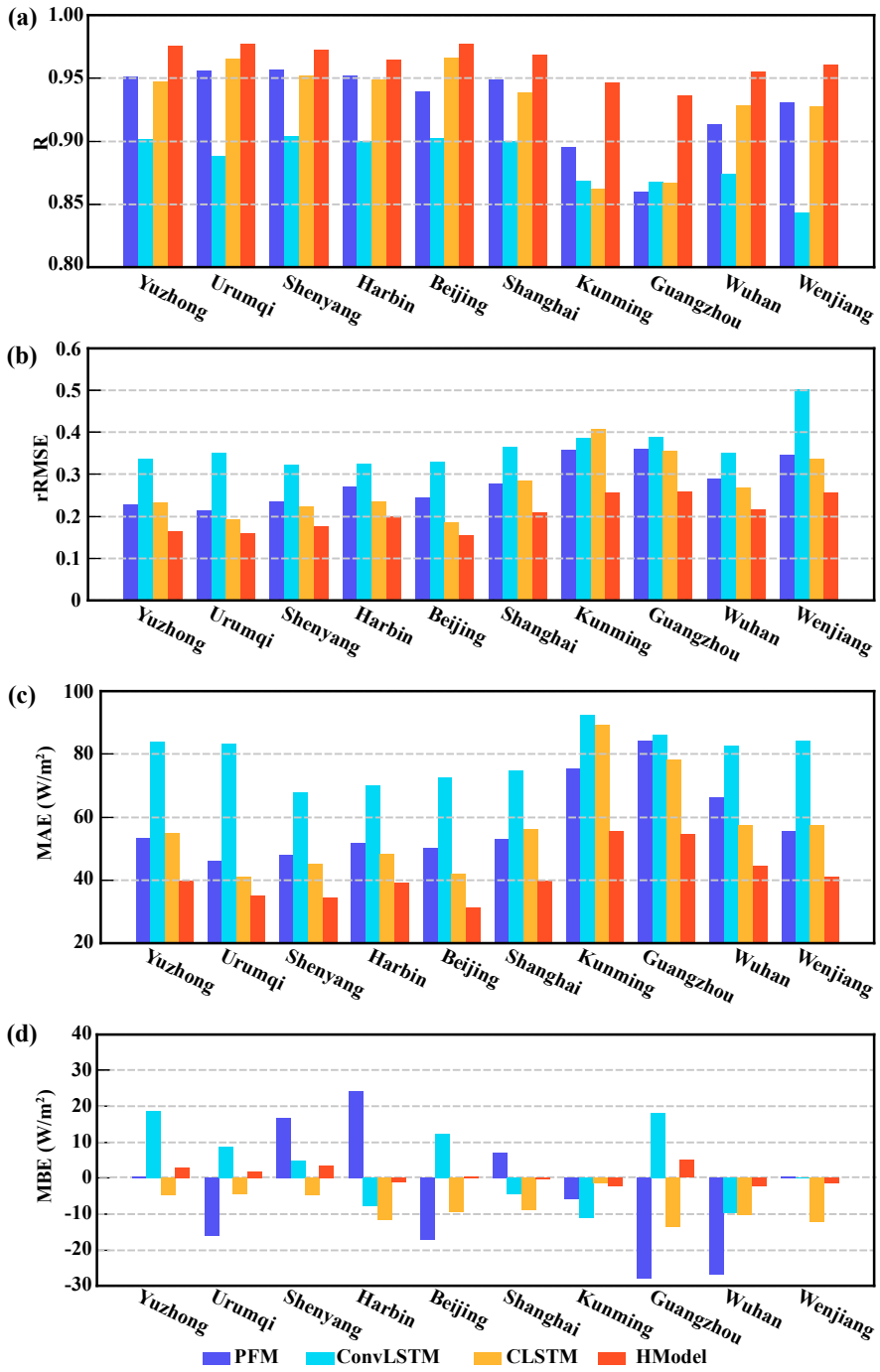

Figure 4. R (**a**), rRMSE (**b**), MBE (**c**) and MAE (**d**) for one-hour ahead forecasting of four compared models at the ten stations.

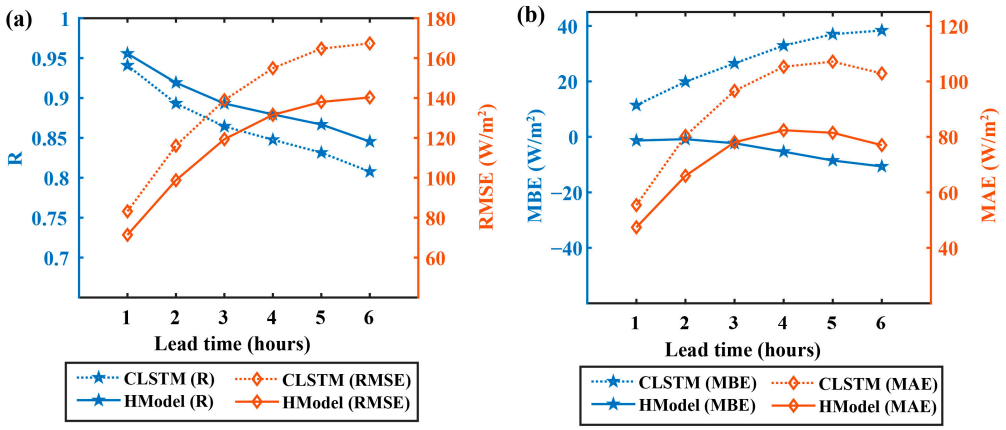

Figure 5. Average R ((**a**), blue), RMSE ((**a**), brown), MBE ((**b**), blue), and MAE ((**b**), brown) of CLSTM (dotted line) and HModel (solid line) forecasts over multi-step horizons at the ten stations.

Table 3. Performance of HModel and CLSTM over multi-step horizons.

Stations	Metrics	HModel						CLSTM					
		1 h	2 h	3 h	4 h	5 h	6 h	1 h	2 h	3 h	4 h	5 h	6 h
Yuzhong	R	0.968	0.938	0.914	0.899	0.884	0.861	0.953	0.911	0.884	0.869	0.855	0.830
	RMSE	70.7	100.2	124.1	140.9	151.7	158.2	85.7	120.9	144.6	160.3	170.2	175.4
	MBE	−2.8	−3.5	−7.5	−13.6	−18.2	−22.2	13.2	13.3	11.1	11.7	14.3	17.5
	MAE	45.9	65.9	80.5	88.1	89.6	86.8	54.7	81.1	98.9	107.1	109.4	107.6
Urumqi	R	0.972	0.948	0.930	0.920	0.914	0.906	0.966	0.929	0.896	0.867	0.837	0.799
	RMSE	60.2	85.2	103.8	115.7	121.0	122.7	66.9	99.5	126.9	150.2	167.8	179.7
	MBE	−0.7	2.8	3.5	2.1	0.1	−0.9	3.7	8.9	15.1	21.5	25.9	28.1
	MAE	39.7	55.8	65.7	70.5	69.6	65.7	41.9	64.6	84.2	100.0	109.1	112.4
Shenyang	R	0.968	0.940	0.908	0.887	0.869	0.843	0.952	0.905	0.874	0.850	0.828	0.799
	RMSE	58.7	82.7	107.3	123.7	133.7	139.1	71.3	103.7	126.7	144.3	156.7	163.0
	MBE	−1.3	−0.9	−2.7	−5.2	−9.1	−11.5	2.6	7.1	11.2	16.8	22.6	28.0
	MAE	38.3	54.5	68.4	75.4	76.8	74.5	46.4	68.8	84.8	96.5	102.2	101.5
Harbin	R	0.958	0.929	0.889	0.860	0.829	0.794	0.945	0.898	0.864	0.838	0.808	0.771
	RMSE	65.9	88.1	114.6	133.4	147.5	154.7	75.5	106.4	130.3	148.9	163.2	171.4
	MBE	−1.5	−5.4	−9.1	−16.9	−25.1	−30.2	7.9	15.3	23.8	31.4	35.7	39.7
	MAE	44.0	59.1	75.7	84.9	89.3	87.6	49.6	73.8	92.3	103.7	108.3	108.1
Beijing	R	0.971	0.940	0.915	0.899	0.882	0.853	0.966	0.927	0.899	0.873	0.845	0.807
	RMSE	56.9	84.6	106.2	120.6	130.3	137.6	62.6	95.1	119.2	139.9	155.9	164.6
	MBE	−2.1	−1.1	−1.8	−3.1	−4.9	−9.1	8.4	15.0	21.5	28.0	32.9	34.2
	MAE	35.6	53.7	66.1	72.2	73.8	73.7	42.0	66.9	84.6	97.5	104.8	105.2
Shanghai	R	0.959	0.922	0.892	0.875	0.861	0.841	0.944	0.893	0.864	0.847	0.829	0.805
	RMSE	71.4	100.6	122.4	135.0	140.3	138.8	84.1	119.4	141.6	155.9	164.0	164.1
	MBE	1.6	1.0	−0.6	−8.2	−11.7	−13.0	13.8	22.3	30.9	37.9	41.1	42.2
	MAE	47.6	66.5	79.6	83.7	82.2	76.3	55.2	81.9	97.8	104.9	105.7	99.9
Kunming	R	0.927	0.873	0.847	0.849	0.848	0.833	0.912	0.853	0.829	0.827	0.832	0.830
	RMSE	98.3	130.7	149.0	153.7	153.7	151.0	109.7	145.4	164.1	174.5	174.8	166.6
	MBE	−0.9	−1.5	−4.2	−7.6	−9.5	−8.8	21.4	35.7	42.3	48.6	50.6	47.3
	MAE	67.5	89.7	100.6	98.7	93.2	84.4	77.7	105.9	118.6	121.2	114.3	100.9
Guangzhou	R	0.929	0.888	0.883	0.885	0.879	0.859	0.907	0.868	0.858	0.862	0.864	0.857
	RMSE	86.9	111.7	121.2	124.5	125.7	125.3	99.9	124.9	140.1	146.6	148.4	143.8
	MBE	−8.2	−10.3	−9.8	−5.2	−7.3	−10.0	13.4	23.9	30.6	38.8	43.3	43.6
	MAE	59.0	76.1	81.2	79.8	75.5	68.8	69.0	88.7	99.0	99.8	95.4	85.6
Wuhan	R	0.953	0.914	0.893	0.885	0.877	0.862	0.935	0.890	0.866	0.855	0.843	0.824
	RMSE	72.4	101.1	120.0	130.3	135.2	135.5	86.7	119.0	143.2	160.6	171.0	171.5
	MBE	3.3	7.4	7.4	5.9	5.4	3.7	14.5	28.3	38.4	47.5	54.3	55.2
	MAE	47.8	68.2	79.7	82.0	79.8	74.9	57.5	82.2	97.8	106.3	108.1	102.8
Wenjiang	R	0.951	0.901	0.860	0.837	0.824	0.802	0.927	0.858	0.811	0.787	0.773	0.754
	RMSE	72.3	102.7	124.9	137.8	141.3	140.1	89.5	126.6	152.5	168.3	175.4	173.8
	MBE	−0.6	3.2	1.9	−1.5	−4.9	−4.5	15.2	29.0	39.9	46.8	49.4	47.9
	MAE	48.4	70.0	83.4	88.5	84.7	77.4	60.0	89.6	108.0	115.7	113.9	104.6

4. Discussion

4.1. Analysis of Cloud Impacts

Clouds (including cloud amount and cloud variability) are recognized as the most important factor affecting the estimation or forecasting of surface solar radiation [30,34,52,53]. Table 4 presents the R and rRMSE of different models under varying cloud amounts. Samples from each station are divided into three groups with equal numbers according to the daily mean of *K*. For all site-specific models, the smaller the cloud amount, the better the predictive ability. The R value rises from less than 0.9 to more than 0.98, and the rRMSE decreases from more than 70% to less than 20%. HModel has the lowest forecasting error under all magnitudes of cloud amount, demonstrating the advantage of ground- and satellite-based integration. Particularly, the forecast accuracy under a heavy cloud amount is significantly improved (cf. the underlined numbers) with an increase of about 0.1 for R and a decline of 25–30% for rRMSE. However, R remains lower than 0.9 and rRMSE remains greater than 50%, which reminds us that more attempts are needed to improve the predictive ability under cloudy conditions.

Table 4. Performance of different models under varying cloud amounts in terms of R and rRMSE. Samples from each station are divided into three groups with equal numbers according to the ascending order of *K*. The bold ones indicate the best forecasts, and the underlined ones have maximum differences between CLSTM and HModel.

Stations	Daily Mean of K	R				rRMSE			
		PFM	ConvLSTM	CLSTM	HModel	PFM	ConvLSTM	CLSTM	HModel
Yuzhong	0.00–0.38	0.801	0.717	0.796	0.886	0.701	0.996	0.824	0.524
	0.38–0.65	0.853	0.832	0.886	0.944	0.288	0.358	0.316	0.205
	0.65–0.93	0.922	0.844	0.930	**0.957**	0.115	0.207	0.115	**0.093**
Urumqi	0.00–0.39	0.853	0.653	0.867	0.899	0.758	1.213	0.715	0.551
	0.38–0.62	0.896	0.823	0.917	0.945	0.218	0.309	0.209	0.170
	0.62–1.00	0.921	0.924	0.977	**0.983**	0.149	0.224	0.103	**0.091**
Shenyang	0.00–0.36	0.806	0.731	0.784	0.880	0.601	0.818	0.799	0.506
	0.36–0.59	0.906	0.811	0.915	0.939	0.248	0.342	0.232	0.197
	0.59–1.00	0.974	0.899	0.965	**0.977**	0.111	0.212	0.119	**0.099**
Harbin	0.00–0.33	0.804	0.737	0.795	0.849	0.669	0.985	0.801	0.610
	0.33–0.61	0.883	0.798	0.879	0.911	0.278	0.318	0.254	0.210
	0.61–1.00	0.978	0.919	0.972	**0.977**	0.124	0.215	0.124	**0.115**
Beijing	0.00–0.34	0.770	0.735	0.847	0.896	0.794	0.997	0.643	0.491
	0.34–0.59	0.765	0.841	0.923	0.952	0.352	0.301	0.213	0.166
	0.59–1.00	0.869	0.928	0.979	**0.983**	0.185	0.213	0.098	**0.090**
Shanghai	0.00–0.23	0.767	0.645	0.766	0.804	0.833	1.401	1.148	0.714
	0.23–0.50	0.868	0.844	0.868	0.921	0.380	0.441	0.415	0.299
	0.50–0.94	0.913	0.787	0.899	**0.936**	0.146	0.225	0.153	**0.120**
Kunming	0.00–0.24	0.712	0.687	0.688	0.799	0.962	1.360	1.285	0.837
	0.24–0.52	0.697	0.771	0.720	0.871	0.477	0.459	0.506	0.324
	0.52–1.00	0.789	0.846	0.846	**0.938**	0.231	0.248	0.250	**0.150**
Guangzhou	0.00–0.25	0.545	0.620	0.708	0.790	1.506	1.109	1.252	0.740
	0.25–0.50	0.586	0.730	0.729	0.832	0.460	0.379	0.411	0.287
	0.50–1.00	0.534	0.884	0.911	**0.946**	0.345	0.246	0.207	**0.160**
Wuhan	0.00–0.28	0.624	0.680	0.752	0.822	1.208	1.190	0.978	0.689
	0.28–0.53	0.602	0.747	0.830	0.890	0.479	0.367	0.305	0.233
	0.53–1.00	0.772	0.911	0.966	**0.968**	0.245	0.224	0.138	**0.132**
Wenjiang	0.00–0.15	0.719	0.630	0.787	0.810	0.904	1.673	1.190	0.785
	0.15–0.35	0.759	0.703	0.817	0.893	0.478	0.634	0.526	0.351
	0.35–0.93	0.866	0.714	0.878	**0.923**	0.217	0.335	0.216	**0.170**

Table 5 quantifies the impact of cloud variability on different models. The cloud amount is restricted to a small range to isolate its effects, and all samples within the range of the highest probability are separated into two equal parts according to the standard deviation of *K*. Generally, as cloud variability increases, the forecasting error increases, provided there are basically equal cloud amounts. Forecasting errors are generally greater when clouds show more drastic changes, and HModel outperforms the other methods for all cases. The cloud variability at the Harbin and Guangzhou stations is significantly higher than that at other locations; hence, the model performance is relatively worse, and the benefit of integrating satellite observations is quite clear.

Table 5. Impact of cloud variability on different forecasting methods in terms of R and rRMSE. For each station, samples within days whose daily mean of *K* is within the range of the highest probability (labelled under the station name) are separated into two equal parts according to standard deviation of *K*.

Stations (*K*)	Daily std. Deviation of *K*	R				rRMSE			
		PFM	ConvLSTM	CLSTM	HModel	PFM	ConvLSTM	CLSTM	HModel
Yuzhong	0.10–0.18	0.966	0.915	0.974	0.979	0.163	0.288	0.160	0.143
(0.5–0.6)	0.18–0.29	0.934	0.886	0.959	0.968	0.269	0.385	0.229	0.203
Urumqi	0.10–0.18	0.938	0.859	0.937	0.964	0.225	0.373	0.253	0.190
(0.4–0.5)	0.18–0.36	0.928	0.847	0.936	0.958	0.293	0.438	0.293	0.237
Shenyang	0.10–0.17	0.969	0.939	0.974	0.977	0.200	0.257	0.171	0.157
(0.4–0.5)	0.17–0.40	0.955	0.902	0.951	0.961	0.237	0.336	0.238	0.214
Harbin	0.14–0.24	0.829	0.803	0.719	0.905	0.385	0.450	0.568	0.317
(0.5–0.6)	0.24–0.66	0.777	0.794	0.710	0.891	0.492	0.510	0.595	0.371
Beijing	0.05–0.18	0.948	0.916	0.946	0.969	0.203	0.274	0.217	0.178
(0.4–0.5)	0.18–0.47	0.935	0.876	0.937	0.966	0.238	0.352	0.251	0.186
Shanghai	0.09–0.15	0.945	0.885	0.951	0.968	0.221	0.332	0.215	0.176
(0.4–0.5)	0.15–0.25	0.936	0.907	0.939	0.967	0.282	0.361	0.289	0.216
Kunming	0.08–0.17	0.932	0.854	0.937	0.963	0.236	0.385	0.254	0.196
(0.3–0.4)	0.17–0.36	0.927	0.851	0.936	0.959	0.303	0.437	0.298	0.236
Guangzhou	0.08–0.20	0.832	0.694	0.836	0.904	0.408	0.726	0.559	0.351
(0.3–0.4)	0.20–0.55	0.791	0.692	0.796	0.883	0.536	0.789	0.650	0.439
Wuhan	0.11–0.18	0.901	0.857	0.911	0.945	0.253	0.348	0.274	0.214
(0.4–0.5)	0.18–0.30	0.876	0.779	0.907	0.922	0.326	0.460	0.308	0.279
Wenjiang	0.02–0.07	0.969	0.945	0.975	0.983	0.151	0.226	0.153	0.128
(0.1–0.2)	0.07–0.16	0.938	0.898	0.947	0.975	0.223	0.312	0.224	0.156

Figures 6–8 compare the forecast results of HModel and CLSTM with actual time series under different (sunny, cloudy, and rainy) weather conditions in temperate (Beijing), subtropical (Wuhan), and tropical (Guangzhou) climate zones, respectively. Only daytime hours with measured solar radiation are displayed. On sunny days, the predicted values change in sync with the actual values, and solar radiation variations are highly consistent. With respect to cloudy days, the forecasting results of CLSTM are severely lagging, because the impact of future clouds is not taken into consideration, for example, forecasts are excessive when a cloud enters the field of view to affect the point of interest because antecedent measurements cannot warn the arrival of cloud, while forecasts are insufficient when a cloud disappears from the field of view, because ground time series alone are unable to warn the departure of cloud. HModel alleviates this difficulty by predicting the next positions of clouds based on the trajectory and velocity inferred from previous satellite images. On overcast days (e.g., 27 June 2008, Beijing, 22/23 July 2008, Wuhan, 21/22 May 2008, Guangzhou), HModel can properly predict solar radiation variations, while if the rainfall is intermittent (e.g., 28 June 2008, Beijing, 21 July 2008, Wuhan, 26/27 May 2008, Guangzhou), its forecasts are relatively poor.

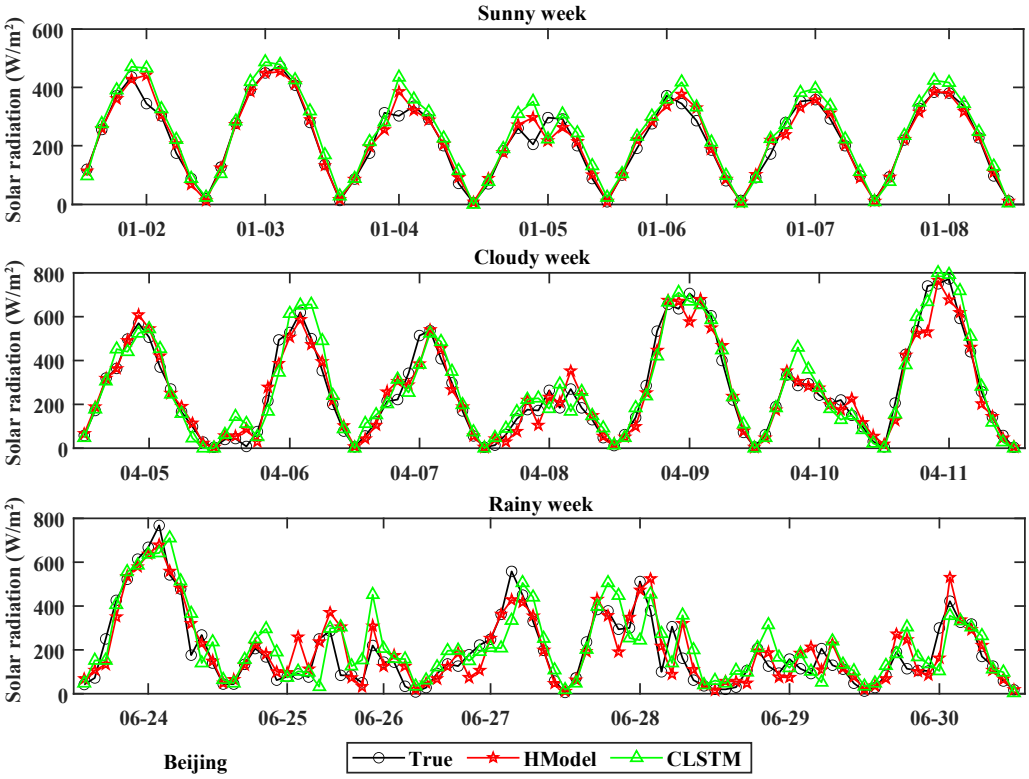

Figure 6. One-hour-ahead forecast results of CLSTM (green) and HModel (red) vs. the measured values (black) under sunny, cloudy, and rainy weather conditions in 2008 at Beijing station (only daytime hours are displayed).

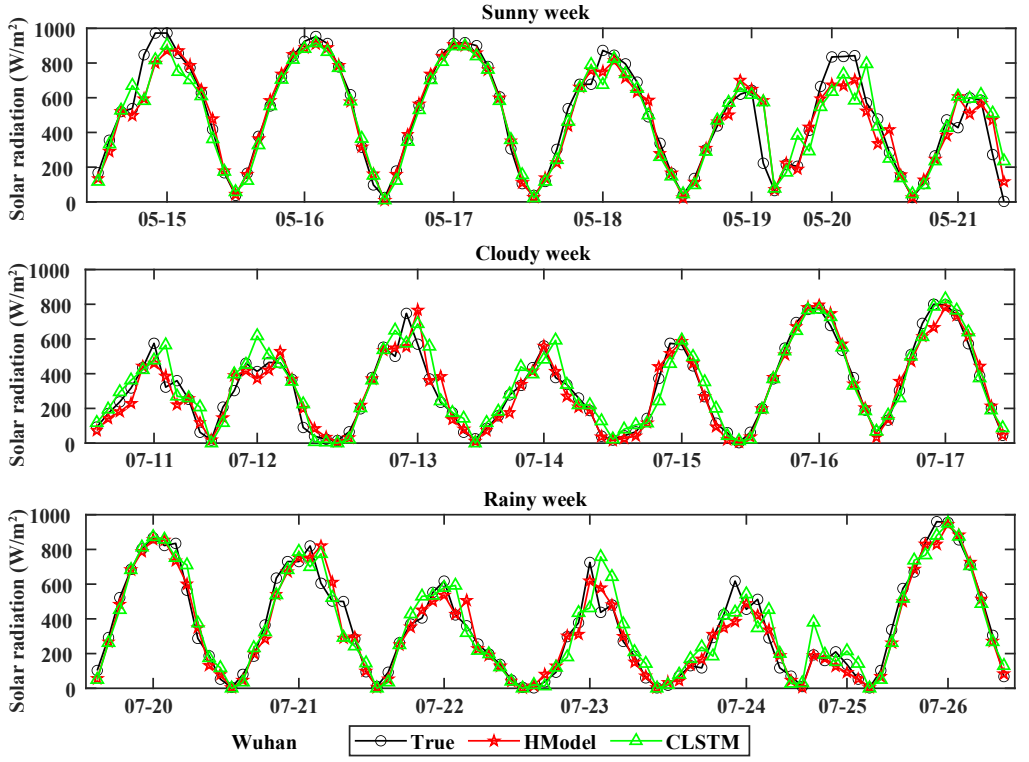

Figure 7. One-hour-ahead forecast results of CLSTM (green) and HModel (red) vs. the measured values (black) under sunny, cloudy, and rainy weather conditions in 2008 at Wuhan station (only daytime hours are displayed).

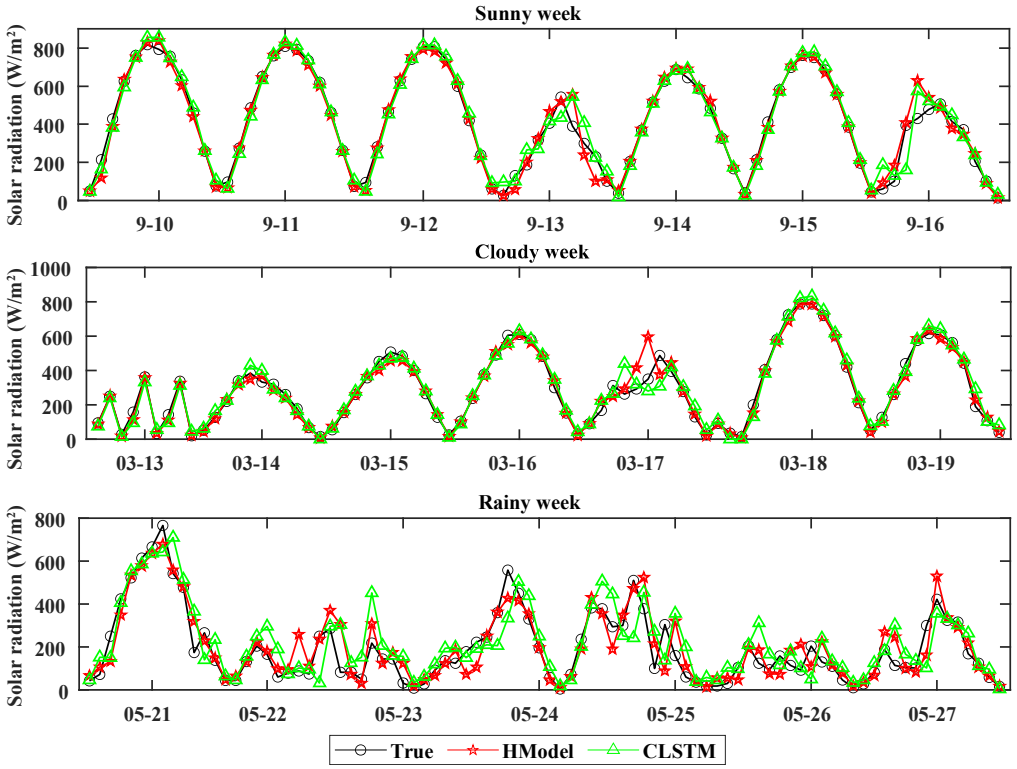

Figure 8. One-hour forecast results of CLSTM (green) and HModel (red) vs. the measured values (black) under sunny, cloudy, and rainy weather conditions in 2008 at Guangzhou station (only daytime hours are displayed).

4.2. Integration of Cloud Information

Accurately capturing and integrating cloud information is critical for enhancing solar radiation forecasting. Traditional approaches primarily rely on ground-based measurements or basic cloud properties extracted from satellite data. While these methods offer insights into solar radiation trends, they often fail to fully capture the dynamic and spatially varying effects of cloud cover. In this study, we address this limitation by developing a hybrid forecasting model that combines satellite observations and ground-based measurements, leveraging advanced machine learning techniques to enhance cloud information integration.

Table 6 provides a comparative analysis of our work with existing methods. Previous studies have explored various approaches for integrating cloud information, but each has notable limitations. For instance, traditional methods, such as those reported by Kosmopoulos et al. (2024) [32] and Wang et al. (2019) [53], rely on cloud motion vectors or cloud physical properties derived from satellite imagery. While effective to some extent, these methods are often restricted to spatial analysis and lack the temporal context provided by ground-based measurements. Similarly, machine learning-based methods, such as those reported by Liu et al. (2019) [28] and Mazorra Aguiar et al. (2015) [40], utilize CNNs for spatial feature extraction from satellite images, but neglect the temporal dynamics essential for accurate short-term predictions.

Our approach advances beyond these limitations by introducing a hybrid CNN–LSTM framework that captures both the spatial and temporal dynamics of cloud impacts on solar radiation. The CNN component extracts spatial cloud features, such as motion trajectories and cloud densities, from satellite imagery, while the LSTM component models the temporal relationships between cloud dynamics and surface radiation fluctuations. This dual approach allows HModel to effectively integrate non-local cloud effects, such as cloud shadowing and scattering, with local ground-level variations, creating a more comprehensive forecasting framework. Moreover, the integration of satellite and ground-

based data ensures that the model captures the quantitative relationships between cloud properties and solar radiation. This is particularly advantageous in regions with a high cloud variability, where traditional models often struggle due to limited spatial or temporal data. By addressing these challenges, our study contributes to the growing body of research on cloud information integration for solar radiation forecasting.

Table 6. Comparison of existing works and our improvements in this study.

No.	Category	Reported Methods	Existing Literatures	Our Contributions
1	Traditional methods for ground-based data	Statistical or traditional approaches used for solar radiation prediction based on ground measurements only.	Prasad et al. (2020) [8] Dong et al. (2013) [19] Lauret et al. (2015) [21]	Integration of ground-based measurements with satellite imagery to enhance prediction accuracy by capturing both temporal trends and spatial cloud dynamics.
2	Traditional methods using cloud information	Cloud motion or physical properties used to improve solar radiation forecasts.	Kosmopoulos et al. (2024) [32] Peña-Cruz et al. (2024) [33] Huang et al. (2013) [52] Wang et al. (2019) [53]	A hybrid model combining cloud information from satellite data and temporal data from ground measurements to address both local and non-local cloud effects on solar radiation.
3	Coupled models for meteorological data	Coupled models integrating weather forecasts for solar radiation prediction.	Qing et al. (2018) [26] Ghimire et al. (2019) [54] Wang et al. (2019) [53]	Development of a hybrid model combining real-time meteorological data from satellite observations and ground-based measurements for multi-step solar radiation forecasts.
4	Machine learning methods for ground-based sequences	Advanced machine learning models applied to ground-based data for solar radiation prediction.	Peng et al. (2021) [7] Gairaa et al. (2016) [17] Yagli et al. (2019) [23] Ghimire et al. (2019) [25]	Incorporation of a hybrid CNN–LSTM architecture to utilize both ground-based data and spatially rich satellite imagery, enabling better temporal and spatial feature extraction.
5	CNN-based methods for cloud information	CNN-based approaches applied to satellite imagery for estimating solar radiation but without integrating ground data.	Liu et al. (2019) [28] Mazorra et al. (2015) [40]	Joint application of CNN and LSTM for spatial–temporal feature extraction, ensuring a better simulation of dynamic cloud changes and their impacts on solar radiation at the surface.

5. Conclusions

This study introduces a spatially and temporally united network, mainly consisting of a CNN and LSTM, for solar radiation forecasting. The CNN extracts cloud information such as the speed and trajectory of cloud motion from time series of satellite observations, and together with LSTM, simulates the impacts of clouds on future solar radiation. In parallel, the CNN obtains the actual solar radiation fluctuations caused by clouds from the ground-measured time series to master the rules of how clouds affect surface solar radiation. LSTM predicts future comprehensive states according to the antecedent abstract features extracted by the CNN. This hybrid network (i.e., HModel) enables long-range spatial connections across multiple timescales between future solar radiation and clouds far from the point of interest. The main findings are as follows:

(1) Experiments at ten international stations show that the performance of HModel is related to spatial and temporal scales, and the recommended options are an input length of 6 h, a maximum forecast horizon of 6 h, and a spatial input size of about

16 pixels (~80 km). These configurations provide significant improvements in solar radiation forecasts compared to models that only use ground-based measurements.

(2) HModel is compared with three baseline models. The results show that integrating satellite observations with ground measurements improves its predictive capability. The improvement in terms of RMSE (rRMSE) varies from about 12 W/m^2 (3%) for locations with prevailing clear skies to more than 50 W/m^2 (15%) for locations dominated by clouds with respect to one-hour-ahead forecasting, and from about 12 W/m^2 (2%) at 1 h ahead to 28 W/m^2 (6%) at 6 h ahead with respect to forecasts over a multi-step horizon, demonstrating the model's robustness over varying time scales.

(3) The predictive capability of HModel is affected by the cloud amount and cloud variability. On average, R decreases from more than 0.98 to less than 0.9 and rRMSE increases from less than 20% to more than 70% with an increasing cloud amount. Cloud variability may result in an R (rRMSE) difference of approximately 0.01 (5%) in the case of similar cloud amounts. In addition, the gap between instantaneous satellite observation and 1 h averaged ground measurement also introduces additional forecast errors, particularly in regions with high cloud dynamics.

This study strengthens the meaning and value of multi-source data fusion in solar radiation forecasting. Stochastic fluctuations in surface solar radiation are mainly due to variations in clouds, therefore, satellite observations with rich cloud information are very important. On the other hand, this study proves the applicability of deep learning in integrating multi-scale spatial–temporal features for multi-step forecasts. The combination of CNN and LSTM allows the model to handle both long-range spatial connections and time series trend correlations. Future research can focus on testing the performance of HModel at better temporal resolutions (e.g., 5, 10, or 30 min) prior to being implemented in energy management systems. Further improvements of HModel are also necessary to enable long-term evaluations of solar energy (e.g., at daily or monthly scales). We can also try to optimize the network structure, thereby improving the forecasting skill, by using more advanced techniques such as the popular multiresolution analysis to extract information in the time–frequency domain [54], the deep neural network to model long- and short-term temporal patterns [55], etc.

Author Contributions: Conceptualization, Q.B. and Z.Y.; methodology, S.Z.; software, F.L.; validation, Y.Y., T.M. and T.D.; formal analysis, S.Z.; investigation, Z.Y.; resources, Q.B.; data curation, Z.Y.; writing—original draft preparation, Q.B., S.Z. and F.L.; writing—review and editing, Z.Y.; visualization, F.L.; supervision, Q.B.; project administration, Z.Y.; funding acquisition, S.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Project of State Grid Jiangsu Electric Power Co., Ltd. (Grant number J2023169).

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: We thank the China Meteorological Administration for providing the hourly solar radiation data, and the Japan Meteorological Agency for publicly available Multi-functional Transport Satellites (MTSAT) satellite images.

Conflicts of Interest: Authors Qiangsheng Bu, Shuyi Zhuang, Fei Luo, Zhigang Ye and Yubo Yuan were employed by State Grid Jiangsu Electric Power Company Ltd. Research Institute. Authors Tianrui Ma and Tao Da were employed by State Grid Jiangsu Electric Power Company Ltd. Zhenjiang Power Supply Branch.

Abbreviations and Symbols

Abbreviation	Full Form
PV	Photovoltaic
ARIMA	Auto-Regressive Integrated Moving Average
ARMA	Autoregressive Moving Average

ANN	Artificial Neural Network
GP	Gaussian Processes
SVR	Support Vector Regression
LSTM	Long Short-Term Memory
NWP	Numerical Weather Prediction
CNN	Convolutional Neural Network
CMA	China Meteorological Administration
MTSAT	Multi-Functional Transport Satellites
PFM	Persistence Forecasting Model
ACF	Autocorrelation Function
PACF	Partial Autocorrelation Function
ReLU	Rectified Linear Unit
MP	Max-Pooling
GAP	Global Average Pooling
ANN	Artificial Neural Network
PFM	Persistence Forecasting Model
CLSTM	Convolutional Long Short-Term Memory
ConvLSTM	Convolutional LSTM

Symbol	**Definition**	**Unit**
x_t	Actual solar radiation at time t	W/m^2
x_{t+i}	Forecasted solar radiation at time $t+i$	W/m^2
$s_t^{h \times w}$	Satellite image block of size $h \times w$ at time t	Pixels
K_t	Clearness index at time t	Dimensionless
T	Lagged time length for time series inputs	Hours
m	Forecast lead time	Hours
h, w	Dimensions of satellite image block	Pixels
R	Correlation coefficient between observed and forecasted values	Dimensionless
RMSE	Root-mean square error between observed and forecasted values	W/m^2
rRMSE	Relative root-mean square error	%
MBE	Mean bias error	W/m^2
MAE	Mean absolute error	W/m^2

References

1. Kabir, E.; Kumar, P.; Kumar, S.; Adelodun, A.A.; Kim, K.-H. Solar energy: Potential and future prospects. *Renew. Sustain. Energy Rev.* **2018**, *82*, 894–900. [CrossRef]
2. Nazari-Heris, M.; Mohammadi-Ivatloo, B.; Gharehpetian, G.B. Short-term scheduling of hydro-based power plants considering application of heuristic algorithms: A comprehensive review. *Renew. Sustain. Energy Rev.* **2017**, *74*, 116–129. [CrossRef]
3. Qin, H.; Zhou, J.; Lu, Y.; Wang, Y.; Zhang, Y. Multi-objective differential evolution with adaptive Cauchy mutation for short-term multi-objective optimal hydro-thermal scheduling. *Energy Convers. Manag.* **2010**, *51*, 788–794. [CrossRef]
4. Mühlemann, D.; Folini, D.; Pfenninger, S.; Wild, M.; Wohland, J. Meteorologically-Informed Spatial Planning of European PV Deployment to Reduce Multiday Generation Variability. *Earth's Future* **2022**, *10*, e2022EF002673. [CrossRef]
5. Aldeman, M.R.; Jo, J.H.; Loomis, D.G.; Krull, B. Reduction of solar photovoltaic system output variability with geographical aggregation. *Renew. Sustain. Energy Transit.* **2023**, *3*, 100052. [CrossRef]
6. Guermoui, M.; Melgani, F.; Danilo, C. Multi-step ahead forecasting of daily global and direct solar radiation: A review and case study of Ghardaia region. *J. Clean. Prod.* **2018**, *201*, 716–734. [CrossRef]
7. Peng, T.; Zhang, C.; Zhou, J.; Nazir, M.S. An integrated framework of Bi-directional long-short term memory (BiLSTM) based on sine cosine algorithm for hourly solar radiation forecasting. *Energy* **2021**, *221*, 119887. [CrossRef]
8. Prasad, R.; Ali, M.; Xiang, Y.; Khan, H. A double decomposition-based modelling approach to forecast weekly solar radiation. *Renew. Energy* **2020**, *152*, 9–22. [CrossRef]
9. Shen, Y.; Liu, S.; Mazhar, A.R.; Han, X.; Yang, L.; Yang, X.E. A review of solar-driven short-term low temperature heat storage systems. *Renew. Sustain. Energy Rev.* **2021**, *141*, 110824. [CrossRef]
10. Yan, J.Y.; Yang, Y.; Campana, P.E.; He, J.J. City-level analysis of subsidy-free solar photovoltaic electricity price, profits and grid parity in China. *Nat. Energy* **2019**, *4*, 709–717. [CrossRef]
11. Guermoui, M.; Melgani, F.; Gairaa, K.; Mekhalfi, M.L. A comprehensive review of hybrid models for solar radiation forecasting. *J. Clean. Prod.* **2020**, *258*, 120357. [CrossRef]
12. Antonanzas, J.; Osorio, N.; Escobar, R.; Urraca, R.; Martinez-de-Pison, F.J.; Antonanzas-Torres, F. Review of photovoltaic power forecasting. *Sol. Energy* **2016**, *136*, 78–111. [CrossRef]

13. Das, U.K.; Tey, K.S.; Seyedmahmoudian, M.; Mekhilef, S.; Idris, M.Y.I.; Van Deventer, W.; Horan, B.; Stojcevski, A. Forecasting of photovoltaic power generation and model optimization: A review. *Renew. Sustain. Energy Rev.* **2018**, *81*, 912–928. [CrossRef]
14. Diagne, M.; David, M.; Lauret, P.; Boland, J.; Schmutz, N. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. *Renew. Sustain. Energy Rev.* **2013**, *27*, 65–76. [CrossRef]
15. Choi, J.; Son, S.-W.; Lee, S.; Park, S. Advancing global solar photovoltaic power forecasting with sub-seasonal climate outlooks. *Renew. Energy* **2024**, *237*, 121803. [CrossRef]
16. Perez, R.; Kivalov, S.; Schlemmer, J.; Hemker, K.; Renné, D.; Hoff, T.E. Validation of short and medium term operational solar radiation forecasts in the US. *Sol. Energy* **2010**, *84*, 2161–2172. [CrossRef]
17. Gairaa, K.; Khellaf, A.; Messlem, Y.; Chellali, F. Estimation of the daily global solar radiation based on Box–Jenkins and ANN models: A combined approach. *Renew. Sustain. Energy Rev.* **2016**, *57*, 238–249. [CrossRef]
18. Hassan, J. ARIMA and regression models for prediction of daily and monthly clearness index. *Renew. Energy* **2014**, *68*, 421–427. [CrossRef]
19. Dong, Z.; Yang, D.; Reindl, T.; Walsh, W.M. Short-term solar irradiance forecasting using exponential smoothing state space model. *Energy* **2013**, *55*, 1104–1113. [CrossRef]
20. Reikard, G. Predicting solar radiation at high resolutions: A comparison of time series forecasts. *Sol. Energy* **2009**, *83*, 342–349. [CrossRef]
21. Lauret, P.; Voyant, C.; Soubdhan, T.; David, M.; Poggi, P. A benchmarking of machine learning techniques for solar radiation forecasting in an insular context. *Sol. Energy* **2015**, *112*, 446–457. [CrossRef]
22. Voyant, C.; Notton, G.; Kalogirou, S.; Nivet, M.-L.; Paoli, C.; Motte, F.; Fouilloy, A. Machine learning methods for solar radiation forecasting: A review. *Renew. Energy* **2017**, *105*, 569–582. [CrossRef]
23. Yagli, G.M.; Yang, D.; Srinivasan, D. Automatic hourly solar forecasting using machine learning models. *Renew. Sustain. Energy Rev.* **2019**, *105*, 487–498. [CrossRef]
24. Duan, J.; Zuo, H.; Bai, Y.; Chang, M.; Chen, X.; Wang, W.; Ma, L.; Chen, B. A multistep short-term solar radiation forecasting model using fully convolutional neural networks and chaotic aquila optimization combining WRF-Solar model results. *Energy* **2023**, *271*, 126980. [CrossRef]
25. Ghimire, S.; Deo, R.C.; Raj, N.; Mi, J. Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms. *Appl. Energy* **2019**, *253*, 113541. [CrossRef]
26. Qing, X.; Niu, Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy* **2018**, *148*, 461–468. [CrossRef]
27. Zhu, L.; Huang, X.; Zhang, Z.; Li, C.; Tai, Y. A novel U-LSTM-AFT model for hourly solar irradiance forecasting. *Renew. Energy* **2025**, *238*, 121955. [CrossRef]
28. Liu, Y.; Qin, H.; Zhang, Z.; Pei, S.; Wang, C.; Yu, X.; Jiang, Z.; Zhou, J. Ensemble spatiotemporal forecasting of solar irradiation using variational Bayesian convolutional gate recurrent unit network. *Appl. Energy* **2019**, *253*, 113596. [CrossRef]
29. Letu, H.; Yang, K.; Nakajima, T.Y.; Ishimoto, H.; Nagao, T.M.; Riedi, J.; Baran, A.J.; Ma, R.; Wang, T.; Shang, H.; et al. High-resolution retrieval of cloud microphysical properties and surface solar radiation using Himawari-8/AHI next-generation geostationary satellite. *Remote Sens. Environ.* **2020**, *239*, 111583. [CrossRef]
30. Tang, W.; Qin, J.; Yang, K.; Liu, S.; Lu, N.; Niu, X. Retrieving high-resolution surface solar radiation with cloud parameters derived by combining MODIS and MTSAT data. *Atmos. Chem. Phys.* **2016**, *16*, 2543–2557. [CrossRef]
31. Wei, Y.; Zhang, X.; Li, W.; Hou, N.; Zhang, W.; Xu, J.; Feng, C.; Jia, K.; Yao, Y.; Cheng, J.; et al. Trends and Variability of Atmospheric Downward Longwave Radiation Over China From 1958 to 2015. *Earth Space Sci.* **2021**, *8*, e2020EA001370. [CrossRef]
32. Kosmopoulos, P.; Dhake, H.; Melita, N.; Tagarakis, K.; Georgakis, A.; Stefas, A.; Vaggelis, O.; Korre, V.; Kashyap, Y. Multi-Layer Cloud Motion Vector Forecasting for Solar Energy Applications. *Appl. Energy* **2024**, *353*, 122144. [CrossRef]
33. Peña-Cruz, M.I.; Díaz-Ponce, A.; Sánchez-Segura, C.D.; Valentín-Coronado, L.; Moctezuma, D. Short-term forecast of solar irradiance components using an alternative mathematical approach for the identification of cloud features. *Renew. Energy* **2024**, *237*, 121691. [CrossRef]
34. Jiang, H.; Lu, N.; Qin, J.; Tang, W.; Yao, L. A deep learning algorithm to estimate hourly global solar radiation from geostationary satellite data. *Renew. Sustain. Energy Rev.* **2019**, *114*, 109327. [CrossRef]
35. Dambreville, R.; Blanc, P.; Chanussot, J.; Boldo, D. Very short term forecasting of the Global Horizontal Irradiance using a spatio-temporal autoregressive model. *Renew. Energy* **2014**, *72*, 291–300. [CrossRef]
36. Marchesoni-Acland, F.; Alonso-Suárez, R. Intra-day solar irradiation forecast using RLS filters and satellite images. *Renew. Energy* **2020**, *161*, 1140–1154. [CrossRef]
37. Narvaez, G.; Giraldo, L.F.; Bressan, M.; Pantoja, A. Machine learning for site-adaptation and solar radiation forecasting. *Renew. Energy* **2021**, *167*, 333–342. [CrossRef]
38. Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N.; Prabhat, F. Deep learning and process understanding for data-driven Earth system science. *Nature* **2019**, *566*, 195–204. [CrossRef] [PubMed]
39. Ghimire, S.; Deo, R.C.; Downs, N.J.; Raj, N. Global solar radiation prediction by ANN integrated with European Centre for medium range weather forecast fields in solar rich cities of Queensland Australia. *J. Clean. Prod.* **2019**, *216*, 288–310. [CrossRef]
40. Mazorra Aguiar, L.; Pereira, B.; David, M.; Díaz, F.; Lauret, P. Use of satellite data to improve solar radiation forecasting with Bayesian Artificial Neural Networks. *Sol. Energy* **2015**, *122*, 1309–1324. [CrossRef]

41. Yao, W.; Li, Z.; Xiu, T.; Lu, Y.; Li, X. New decomposition models to estimate hourly global solar radiation from the daily value. *Sol. Energy* **2015**, *120*, 87–99. [CrossRef]
42. Nair, V.; Hinton, G.E. Rectified linear units improve restricted Boltzmann machines. In Proceedings of the 27th International Conference on International Conference on Machine Learning, Haifa, Israel, 21–24 June 2010; pp. 807–814.
43. Glorot, X.; Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. *J. Mach. Learn. Res.* **2010**, *9*, 249–256.
44. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
45. Chollet, F. Keras. Available online: https://github.com/fchollet/keras (accessed on 2 December 2024).
46. Bergstra, J.; Yamins, D.; Cox, D.D. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In Proceedings of the 30th International Conference on International Conference on Machine Learning—Volume 28, Atlanta, GA, USA, 17–19 June 2013; pp. I-115–I-123.
47. Wyser, K.; O'Hirok, W.; Gautier, C. A simple method for removing 3-D radiative effects in satellite retrievals of surface irradiance. *Remote Sens. Environ.* **2005**, *94*, 335–342. [CrossRef]
48. Ryu, Y.; Jiang, C.; Kobayashi, H.; Detto, M. MODIS-derived global land products of shortwave radiation and diffuse and total photosynthetically active radiation at 5 km resolution from 2000. *Remote Sens. Environ.* **2018**, *204*, 812–825. [CrossRef]
49. Jiang, H.; Lu, N.; Huang, G.; Yao, L.; Qin, J.; Liu, H. Spatial scale effects on retrieval accuracy of surface solar radiation using satellite data. *Appl. Energy* **2020**, *270*, 115178. [CrossRef]
50. Shi, X.; Chen, Z.; Wang, H.; Yeung, D.-Y.; Wong, W.-K.; Woo, W.-C. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. *arXiv* **2015**, arXiv:1506.04214.
51. Inman, R.H.; Pedro, H.T.C.; Coimbra, C.F.M. Solar forecasting methods for renewable energy integration. *Prog. Energy Combust. Sci.* **2013**, *39*, 535–576. [CrossRef]
52. Huang, H.; Xu, J.; Peng, Z.; Yoo, S.; Yu, D.; Huang, D.; Qin, H. Cloud motion estimation for short term solar irradiation prediction. In Proceedings of the 2013 IEEE International Conference on Smart Grid Communications (SmartGridComm), Vancouver, BC, Canada, 21–24 October 2013; pp. 696–701.
53. Wang, P.; van Westrhenen, R.; Meirink, J.F.; van der Veen, S.; Knap, W. Surface solar radiation forecasts by advecting cloud physical properties derived from Meteosat Second Generation observations. *Sol. Energy* **2019**, *177*, 47–58. [CrossRef]
54. Ghimire, S.; Deo, R.C.; Raj, N.; Mi, J. Wavelet-based 3-phase hybrid SVR model trained with satellite-derived predictors, particle swarm optimization and maximum overlap discrete wavelet transform for solar radiation prediction. *Renew. Sustain. Energy Rev.* **2019**, *113*, 109247. [CrossRef]
55. Lai, G.; Chang, W.-C.; Yang, Y.; Liu, H. Modeling Long- and Short-Term Temporal Patterns with Deep Neural Networks. In Proceedings of the The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, Ann Arbor, MI, USA, 8–12 July 2018; pp. 95–104.