

Article

A Novel Hybrid Model of WRF and Clearness Index-Based Kalman Filter for Day-Ahead Solar Radiation Forecasting

Yuzhang Che ^{1,2,*} , Lewei Chen ³, Jiafeng Zheng ¹, Liang Yuan ¹ and Feng Xiao ²

¹ College of Atmospheric Science, Chengdu University of Information Technology, Chengdu 610225, China; zjf1988@cuit.edu.cn (J.Z.); yuanliang@cuit.edu.cn (L.Y.)

² Department of Mechanical Engineering, Tokyo Institute of Technology, Ookayama, Meguro-ku, Tokyo 152-8550, Japan; xiao.f.aa@m.titech.ac.jp

³ School of Human Settlement and Civil Engineering, Xi'an Jiaotong University, Xi'an 710049, China; chenlewei@stu.xjtu.edu.cn

* Correspondence: che.y.aa@m.titech.ac.jp

Received: 3 September 2019; Accepted: 19 September 2019; Published: 21 September 2019



Abstract: Day-ahead forecasting of solar radiation is essential for grid balancing, real-time unit dispatching, scheduling and trading in the solar energy utilization system. In order to provide reliable forecasts of solar radiation, a novel hybrid model is proposed in this study. The hybrid model consists of two modules: a mesoscale numerical weather prediction model (WRF: Weather Research and Forecasting) and Kalman filter. However, the Kalman filter is less likely to predict sudden changes in the forecasting errors. To address this shortcoming, we develop a new framework to implement a Kalman filter based on the clearness index. The performance of this hybrid model is evaluated using a one-year dataset of solar radiation taken from a photovoltaic plant located at Maizuru, Japan and Qinghai, China, respectively. The numerical results reveal that the proposed hybrid model performs much better in comparison with the WRF-alone forecasts under different sky conditions. In particular, in the case of clear sky conditions, the hybrid model can improve the forecasting accuracy by 95.7% and 90.9% in mean bias error (MBE), and 42.2% and 26.8% in root mean square error (RMSE) for Maizuru and Qinghai sites, respectively.

Keywords: WRF (Weather Research and Forecasting) model; kalman filter; hybrid model; clearness index; solar radiation forecasting

1. Introduction

The utilization of solar energy has increased several fold in recent years due to the drop in pricing of the solar panels and the improvement in the solar panel efficiency. However, integrating this large scale solar yields existing energy supply structures and maximizing the exploitation of such renewable resources are still significant challenges, due mainly to the fluctuating nature of solar radiation at Earth's surface and its high dependence on different atmospheric conditions (e.g., distribution of humidity and clouds). Accurate solar radiation predictions over several time horizons which are essential to grid balancing, real-time unit dispatching, scheduling and trading are required.

Intra-hour forecasts can be determined from advection of clouds observed with ground-based all-sky cameras [1]. The use of a satellite image shows acceptable forecasting skill from 30 min up to six hours [2,3]. For very short-term solar radiation forecasting, statistical approaches that are based on the learning process of historical information have been proposed. These approaches can be divided into two groups, that is, time series and artificial intelligence (AI) based methods. The support vector machines (SVM), linear regressions, autoregressive moving average (ARMA) and the autoregressive

integrated moving average (ARIMA) derived from ARMA that belong to the first group have been widely applied for different applications of solar radiation prediction [4–6]. The AI-based methods, in combination with historical data at the site, is also frequently employed to predict solar radiation. Among them, the artificial neural network (ANN) is more popular and robust to respond to sudden changes dynamically [7]. Al-Alawi and Al-Hinai [8] used it to forecast solar radiation for a target site and a better result compared to an empirical model has been found. Moreover, Benghanem et al. [9] have developed six ANN-based models to predict solar radiation, and the results indicate that the forecasting skill is dependent on sunshine duration and temperature. Recently, ANN and random forest methods have been used to predict the solar radiation [10]. Torres et al. [11] adopted the deep neural network (DNN) to predict the global solar radiation in the Netherlands, showing that the DNN model is much better than the traditional models with low relative RMSE. A long short-term memory (LSTM) model has also been implemented for predicting day-ahead global horizontal irradiance [12,13]. Specific mathematical equations and detailed information related to statistical approaches have been summarized by Raza et al. [7].

Day-ahead forecasting usually is vital to grid integration and decision-making in the energy market. Although Mellit and Pavan [14] proposed a multi-layer perceptron (MLP) model from which the 24-h forecasts of solar irradiance can be obtained, the accuracy of most of the aforementioned statistical approaches decreases rapidly with the extension of the forecasting horizon. Thus, the numerical weather prediction (NWP) model-based approaches become more appropriate and accurate when the prediction horizon is beyond six hours [15,16]. The approaches based on NWP models can generate forecasts of wind speed, direction, surface temperature, humidity and pressure, which can be regarded as indicators for making appropriate strategies. Several NWP models, for example, the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5) [17] and Weather Research and Forecasting (WRF) [16,18–20] are widely used in solar radiation prediction. Lara-Fanego et al. [18] evaluated the reliability of three-day ahead global horizontal irradiance and direct normal irradiance in southern Spain using the WRF model with a spatial resolution of 3 km. Later, Diagne et al. [19] adopted the same horizontal resolution to investigate the ability of the WRF model for forecasting solar radiation at Reunion Island. Both studies show reasonable results that indicate that the WRF model with relatively high-resolution (3 km) can be a better choice in short-term prediction of the solar energy field.

The first objective of this study being to obtain better forecasts with very high resolution under complex terrain conditions, we also employ the WRF as a physical model with the horizontal resolution of 0.5 km rather than 3 km as in the references, as mentioned earlier. However, merely using one physical model that certainly has uncertainties due to inaccurate initial/boundary conditions, simplifications in physics, and numerical approximations [21] might not be able to provide an adequate numerical prediction. It is addressed in Mejia et al. [20] that combining a model output statistics (MOS) approach with WRF forecasts can improve the forecasting skill of solar radiation to some extent. Many existing studies proved that the Kalman filter is much more efficient than the MOS method in the wind prediction field. Thus, our second objective is to develop a novel hybrid model consisting of the high-resolution (0.5 km) WRF model and the clearness index-based Kalman filter, in order to provide day-ahead solar radiation prediction.

Unlike the conventional way of using a Kalman filter [22], we choose the forecasting bias instead of the variable itself as the state variable [23]. More importantly, in our work, the combination of the clearness index and Kalman filter, to the best knowledge of the authors, is proposed to forecast solar radiation for the first time. The clearness index is used to define the different sky conditions (i.e., clear, cloudy and overcast) and then cluster the forecasts series into three groups to conduct Kalman filter separately, to overcome the shortage of a Kalman filter, which is the lack of being able to resolve huge day-to-day changes.

The rest of the paper is organized as follows. Section 2 introduces the model and all the methods we used in this study. The specific information on study areas and the dataset are stated in Section 3,

and Section 4 presents the results and discussions. The paper ends with concluding remarks in Section 5.

2. Model and Methodology

2.1. High-Resolution Physical Model

The latest Advanced Research WRF (ARW) model version 3.9.1, which is a limited-area mesoscale model based on a fully compressible and non-hydrostatic dynamic core, is chosen to predict day-ahead surface solar radiation (hourly interval).

The high-resolution model is operated with the WRF Single-Moment 6-class (WSM6) microphysics parameterization [24], the Kain–Fritsch convective parameterization [25], the Noah land surface model (LSM) [26], and the planetary boundary layer scheme we selected is the YSU (Yonsei University) scheme [27]. The specific information of the short- and longwave radiation schemes we have chosen is discussed in Section 4.1. We re-initialized the WRF model using “cold-start” mode at 6:00 p.m. UTC each day and each re-initialization runs for 30-h. The outputs during initial 6-h are excluded because of the consideration of spin-up time. It is worthwhile to note that, in this study, UTC of WRF output is converted to the local time of the study site, in order to compare with in situ observations conveniently.

2.2. Kalman Filter

The Kalman filter is an estimation algorithm named after Rudolf E Kálmán, which operates recursively on streams of input data (containing random variations) to produce a statistically optimal estimate of the underlying system state. It is over 50 years old but is still one of the most essential and standard data fusion algorithms in use nowadays. The great success of the Kalman filter is due to its adaptive, recursive, optimal characteristics and small computational requirement. Therefore, it is widely used in various fields from radar and computer vision to meteorological purposes. Similar to our previous study [23], the specific set of mathematical equations, which are used to estimate and remove the errors in the solar radiation forecasts, are described as follows.

In this work, the forecasting bias, rather than the variable itself as used in [28,29], is chosen as the state variable in the Kalman filter equations. The forecasting bias x_t at time t is related to the state at previous time step $t - \delta t$:

$$x_{t|t-\delta t} = x_{t-\delta t|t-2\delta t} + \eta_{t-\delta t}, \tag{1}$$

where δt is a time lag, $t|t - \delta t$ indicates that the value of a variable at time t depends on values at time $t - \delta t$, and η is the white noise that has the variance ($\sigma_{\eta,t}^2$) being uncorrected in time. Due to some of the unavoidable factors including numerical noises, the lack of accuracy in the description of sub-grid physical processes and complex geographic conditions, the Kalman filter algorithm further assumes that the forecast error y_t could be stated as follows:

$$y_t = x_t + \epsilon_t = x_{t|t-\delta t} + \eta_t + \epsilon_t, \tag{2}$$

where ϵ_t is normally distributed with zero-mean and variance $\sigma_{\epsilon,t}^2$.

Following the idea of Kalman [30], x_t can be written as a combination of the previous predicted bias and the previous forecast error with a parameter of Kalman gain (G_t):

$$\hat{x}_{t+\delta t|t} = \hat{x}_{t|t-\delta t} + G_t(y_t - \hat{x}_{t|t-\delta t}), \tag{3}$$

$$G_t = \frac{p_{t-\delta t} + \sigma_{\eta,t}^2}{p_{t-\delta t} + \sigma_{\eta,t}^2 + \sigma_{\epsilon,t}^2}, \tag{4}$$

$$p_t = (p_{t-\delta t} + \sigma_{\eta,t}^2)(1 - G_t), \tag{5}$$

where the hat (^) notation indicates the estimation of the variable, and p is the expected mean square error. Given a reasonable initial guess of p_0 and G_0 , this Kalman filter will quickly converge. The time lag δt of 24-h is used in this study. In addition, the estimation of two white noises $\sigma_{\eta,t}^2$ and $\sigma_{\epsilon,t}^2$ is crucial to the implementation of the Kalman filter procedure. The detailed information can be found in Appendix A of Delle Monache et al. [31].

2.3. The Clearness Index Based Kalman Filter

In general, the Kalman filter is implemented on a time-sequence of forecasts to reduce the systematical and random forecasting errors. As mentioned earlier, however, Kalman filter is not able to handle substantial day-to-day changes in the forecasting error. For the sake of improving the performance of the Kalman filter for day-ahead solar prediction, a new framework of implementing Kalman filter is proposed in this study based on the clearness index. It is a dimensionless number between 0 and 1, which is usually used to measure the clearness of the atmosphere, and is defined as the surface radiation divided by the horizontal extraterrestrial radiation,

$$K_t = \frac{\sum_{i=1}^{24} I}{H_0}, \tag{6}$$

where I and H_0 stand for the solar radiation (hourly) and the extraterrestrial horizontal radiation per day, respectively. H_0 is the horizontal insolation without any atmosphere effects and usually calculated for specific sites from the following formulation [32],

$$H_0 = \frac{24 \times 3600 G_{sc}}{\pi} \left(1 + 0.033 \cos \frac{360n}{365}\right) \times \left(\cos \phi \cos \delta \sin \omega_s + \frac{\pi \omega_s}{180} \sin \phi \sin \delta\right), \tag{7}$$

where G_{sc} is the solar constant (1367 W/m^2), n is the day of the year, ϕ is the latitude of the location in degree, δ is the declination angle in degree and ω_s is the sunset hour angle in degree. The horizontal extraterrestrial radiation H_0 is a function of the day of year and latitude. Thus, the H_0 at two target sites (i.e., Maizuru and Qinghai) in Japan and China for all years can be easily obtained (Figure 1) and then used to calculate clearness index that is used to define the sky conditions in this study.

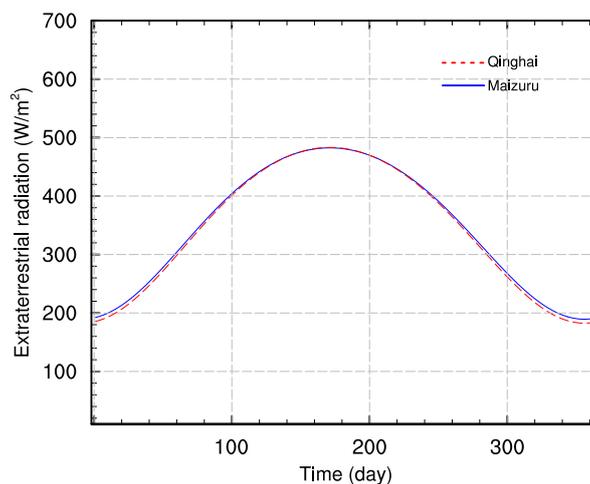


Figure 1. The extraterrestrial horizontal radiation H_0 (W/m^2) series for two sites.

We choose the clearness index defined in Equation (6) to rearrange the time-sequences of forecasts and observations to three new kinds of data series that might have similar forecasting error characteristics, in order to conduct a Kalman filter efficiently. To be specific, three kinds of sky conditions, i.e., clear ($K_t > 0.65$), cloudy ($0.4 \leq K_t \leq 0.65$) and complete overcast ($K_t < 0.4$), are defined firstly based on the values of the clearness index. Then, we implement the Kalman filter on three new data series under different sky conditions. Such a framework for implementing the Kalman

filter in solar radiation forecasting, to the best knowledge of the authors, is proposed for the first time. We compare the performance of the clearness index-based Kalman filter against the results of merely implementing the Kalman filter to the time-sequence of forecasts, WRF raw forecasts and the corresponding observed data in Section 4.

2.4. System Overview

We summarize the hybrid day-ahead solar radiation forecasting model described in Figure 2. The procedure mainly involves two steps, that is, the solar radiation from WRFv3.9.1 model given the best configuration of radiation schemes and the implementing Kalman filter considering the clearness index. In order to validate whether the forecasts from the novel hybrid model can be improved, a controlled experiment is conducted as displayed in Figure 2:

- KF_TS: Implementing Kalman filter on the time-sequence of forecasts.
- KF_Kt: Implementing the clearness index based Kalman filter on three rearranged forecasts data series stated in Section 2.3.

The total length of the sequence of forecasts is 12 months and the proportions of the three rearranged data series are mentioned in Section 4.2. Two weeks are chosen as a training period for implementing the Kalman filter properly for both cases.

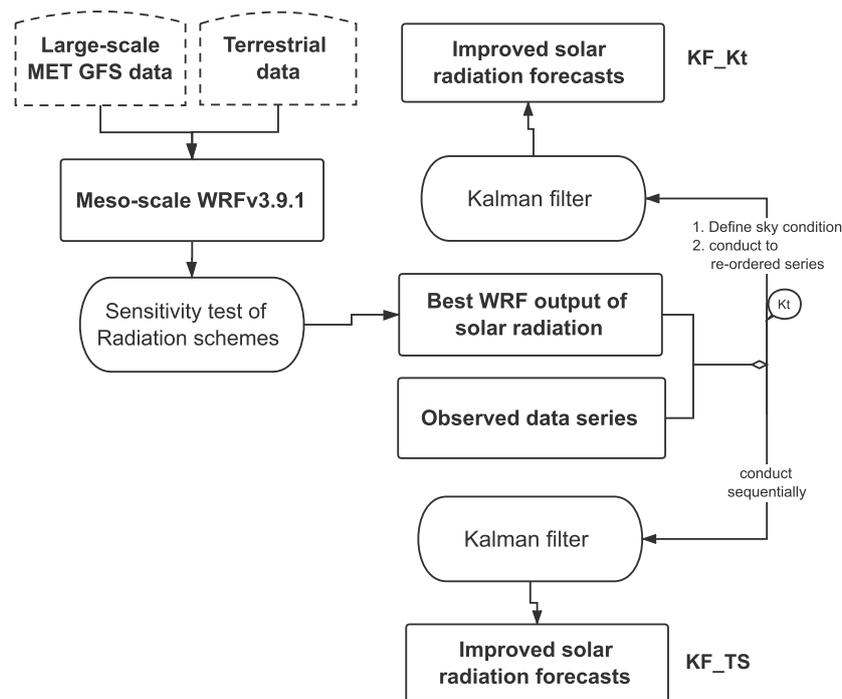


Figure 2. A schematic diagram of the hybrid system of forecasting day-ahead solar radiation.

3. Study Area, Dataset and Evaluation Metrics

3.1. WRF Domain Configuration

Two entirely different study areas (as shown in Figure 3) are involved in this study to validate the forecasting ability of the proposed hybrid model, as summarized in Figure 2. One is the area that covers the site of Maizuru located in the middle of Japan; the other one is located in Qinghai, China.

The WRF domain configuration follows the steps recommended by Warner [33]. For the Maizuru site, it consists of a parent domain and two nested domains with horizontal resolutions of 6.0 km, 1.5 km and 0.5 km, respectively, in order to obtain high-resolution predictions. For the Qinghai site, there are four domains with horizontal resolutions of 12.5 km, 4.5 km, 1.5 km, and 0.5 km, respectively.

The resolutions of the parent domains for both sites are based on GFS (Global Forecast System) real-time forecasts.

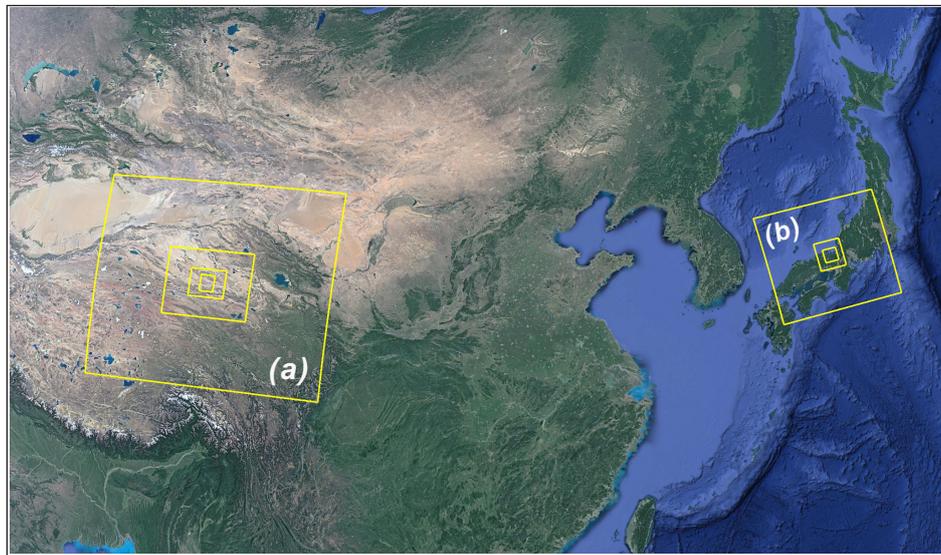


Figure 3. Domain configurations of the WRF (Weather Research and Forecasting) model for Qinghai (a) and Maizuru (b) sites in Japan and China, respectively. This image is taken from Google Earth.

3.2. Dataset

GFS real-time forecasts are chosen as initial/boundary conditions for the high-resolution WRF model. The horizontal resolution is 0.25×0.25 degrees for the Maizuru site while 0.5×0.5 degrees for the Qinghai site. The study periods are also different, i.e., 31 December 2017 to 31 December 2018 and 31 December 2016 to 31 December 2017 for Maizuru and Qinghai sites, respectively.

Numerical results of the present study are evaluated based on ground data of solar radiation collected at two radiometric stations (35.5N, 135.4E; 36.4N, 95.2E), which are all in the centre of our study regions. The data quality has been controlled using the method introduced in [34]. The length of the observed data series adopted in the study is the same as the WRF simulation period.

3.3. Evaluation Metrics

To evaluate the proposed hybrid model quantitatively, the following set of statistical metrics are used.

Mean bias error (MBE):

$$MBE = \frac{1}{N} \sum_{i=1}^N (fore_i - obs_i), \tag{8}$$

where i is the time point and N is the total number of verification time points. $fore$ and obs represent the forecasts and observations, respectively.

Mean absolute error (MAE):

$$MAE = \frac{1}{N} \sum_{i=1}^N |fore_i - obs_i|, \tag{9}$$

Root mean square error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (fore_i - obs_i)^2}{N}}. \tag{10}$$

Normalized root mean square error (nRMSE):

$$\text{nRMSE} = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (\text{fore}_i - \text{obs}_i)^2}}{\text{Max}\{\text{obs}\} - \text{Min}\{\text{obs}\}}. \quad (11)$$

4. Results and Discussion

In this section, we have firstly tuned the WRF model for the solar radiation prediction in two target regions via a set of sensitivity tests of radiation parameterization schemes. Then, the overall results of the comparison between the raw WRF solar radiation forecasts and the corresponding observations are presented based on the statistical parameters described in Section 3.3. Finally, the performances of KF_TS and KF_Kt for improving the raw solar radiation forecasts of the WRF-alone model have been investigated.

4.1. Sensitivity Experiments of Radiation Schemes

Solar radiation is one of the most critical drivers to planetary boundary layer dynamics. There are a couple of short- and longwave schemes to handle the surface downward short- and longwave radiation for the ground heat budget in the current WRF model. The impact of different schemes on surface solar radiation prediction in winter and summer is needed to understand given a target region. The setup of the numerical experiments for inter-comparison of shortwave parameterization schemes is summarized in Table 1. The longwave scheme we adopt in this test is the Rapid Radiative Transfer Model (RRTM) [35]. Two-week in January and July is chosen to represent winter and summer, respectively, with consideration of relatively limited computation resources.

Table 1. Shortwave schemes included in the sensitivity analysis.

Experiment	Short Wave Scheme
Dudhia	Dudhia scheme [36]
Goddard	Goddard shortwave scheme [37]
CAM_SW	Community Atmospheric Model [38]
RRTMG_SW	Rapid Radiative Transfer Model for general circulation model [39]
New_Goddard	New Goddard shortwave scheme [40]

Figure 4 illustrates the variations of the predicted solar radiation from the five sets of experiments with different shortwave parameterization schemes, referred to as Dudhia, Goddard, CAM_SW, RRTMG_SW, and New_Goddard. It is easy to find that outputs of the WRF model underestimate the solar radiation in winter while overestimate it in summer for the Maizuru site. Opposite phenomena can be found for the Qinghai site. Among five shortwave schemes, the case of RRTMG_SW has the smallest normalized standard deviation (NSD) in both winter and summer for the Maizuru site. For the Qinghai site, the case of Dudhia has the smallest NSD and the most considerable correlation. In addition, we have also examined the impact of longwave parameterization schemes; however, no significant difference has been found. Thus, the RRTM longwave scheme is adopted for both two sites; the schemes of the RRTMG shortwave and Dudhia shortwave are chosen for Maizuru and Qinghai sites, respectively.

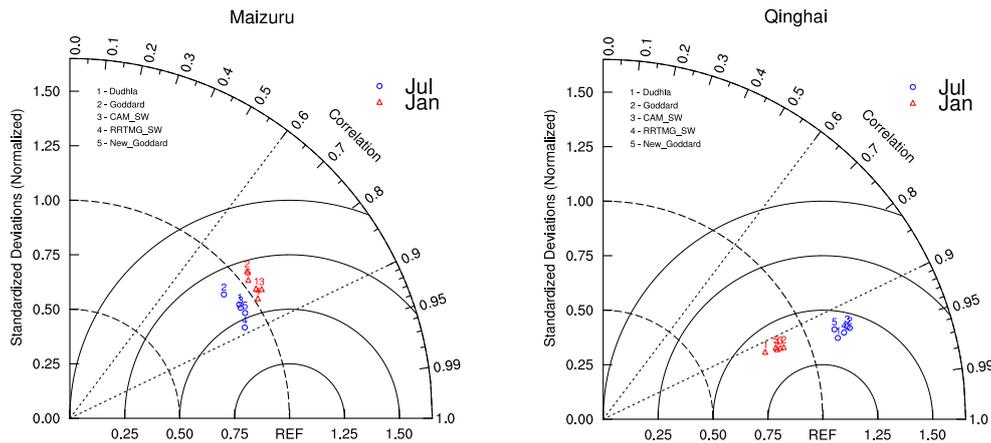


Figure 4. Taylor diagrams for Maizuru (left) and Qinghai (right) sites. The diagrams show normalized standard deviations and correlations of solar radiation forecasts for five experiments referred to observations (REF) in winter (January) and summer (July). The number of samples for each experiment is 336.

4.2. WRF-Alone Forecasts of Solar Radiation

With proper configurations of the WRF model, a 12-month time series of the solar radiation forecasts are generated for two sites, respectively. The forecasting skills of WRF model have been verified with observed data. Figure 5a,c depict the comparison of diurnal variations (over one-year) of predicted and observed solar radiation for the Maizuru and Qinghai sites, respectively. It is observed that the WRF raw forecasts of solar radiation can capture well the pattern of the corresponding observations for both sites. Overestimations of WRF forecasts have been found.

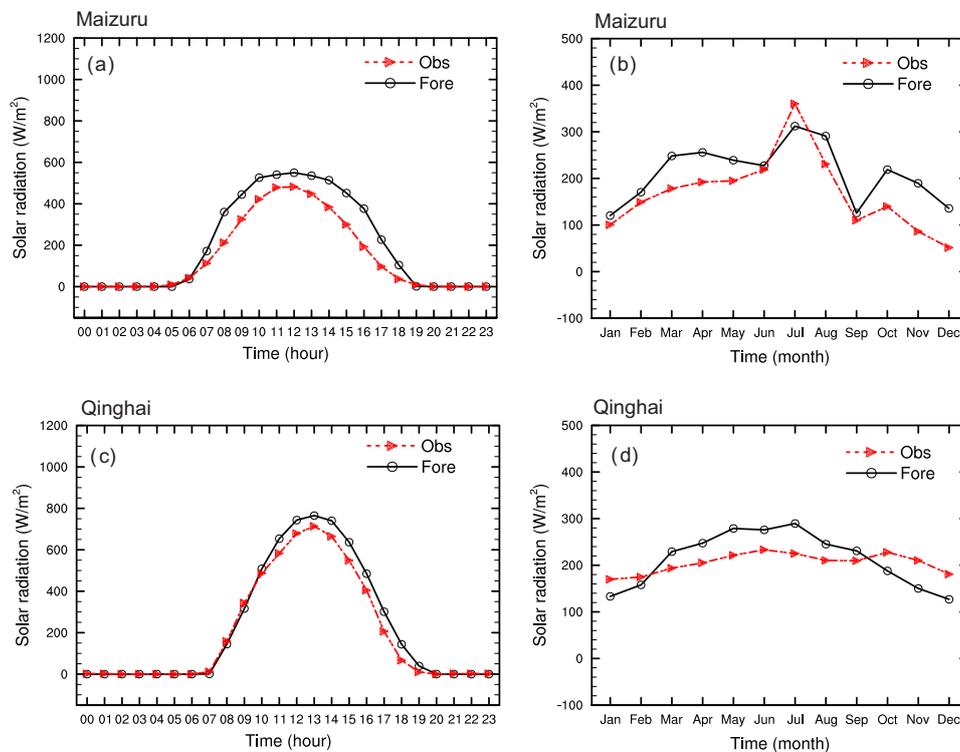


Figure 5. The 12-month mean solar radiation (W/m^2) series from 12:00 a.m. to 11:00 p.m. (a) and the monthly averaged series of 2018 solar radiation (b) for the Maizuru site; (c) and (d) display similar results but for the Qinghai site in 2017. Note that the time mentioned is local time. The marked black and red lines stand for forecasts and the corresponding observations, respectively.

The monthly variations of predicted and observed solar radiation are plotted in Figure 5b,d for the Maizuru and Qinghai sites, respectively. From Figure 5b, we can find slight overestimation for each month except for July. Similarly, the forecasts of the Qinghai site also meet the observed data well and have overestimations for most of the months.

Furthermore, the forecasting ability of WRF model under different sky conditions (defined in Section 2.3) is investigated, and the results are summarized in Table 2. Note that the statistical performance of hourly solar radiation forecasts under different sky conditions is calculated based on their occurrence frequency in the whole year, which is displayed in the second row of Table 2.

Table 2 reveals that the WRF model can predict well the solar radiation for all three sky conditions since the relatively small values of MBE and RMSE can be obtained for both Maizuru and Qinghai sites. It is also observed that the nRMSE of the whole time series (denoted by “All-year”) reaches a relatively small value of 0.2 and 0.1 for Maizuru and Qinghai sites, respectively. Considering all the above results, we may conclude that the WRF model has a relatively high ability of forecasting short-term solar radiation while there are slight differences among different sky conditions and different locations. Such differences can also be found in Figure 6.

Table 2. Performance statistics of hourly WRF (Weather Research and Forecasting) raw forecasts of solar radiation (W/m^2) under different sky conditions from 1 January to 31 December 2018 for the Maizuru site, and from 1 January to 31 December 2017 for the Qinghai site, respectively.

Site	Parameters	Clear-Day	Cloudy-Day	Overcast-Day	All-Year
Maizuru	Frequency of occurrence	57%	17%	26%	—
	MBE	124.7	47.1	−18.1	70.4
	MAE	145.8	116.2	62.9	117.4
	RMSE	247.6	210.3	130.0	213.3
	nRMSE	0.2	0.2	0.2	0.2
Qinghai	Frequency of occurrence	41%	39%	20%	—
	MBE	89.9	0.33	−1.1	24.9
	MAE	113.7	161.9	132.8	96.1
	RMSE	204.7	298.7	251.2	181.1
	nRMSE	0.2	0.3	0.23	0.1

A period of five-day forecasts of the Maizuru site is chosen as an example to show the different performances under three sky conditions. It is also used to explain the reason that we conduct the clearness index-based Kalman filter for the sake of improving the raw forecasts. Figure 6a–c show comparative forecasts of solar radiation (hourly interval) for clear, cloudy and overcast sky conditions, respectively. It can be found that the forecasting patterns under clear sky conditions are very similar. Relatively large variation has been observed in WRF forecasts of solar radiation as well as observations under cloudy sky conditions (Figure 6b). In addition, it can be observed in Figure 6c that the values of solar radiation forecasts are much smaller than the values under the clear and cloudy sky conditions. Such a significant difference in forecasting errors among three kinds of sky conditions might limit the performance of KF_TS for improving the forecasting accuracy. In other words, searching analogues using the clearness index will enhance the role of the Kalman filter in reducing forecasting errors. In the following sections, the performance of the Kalman filter on time-sequence (KF_TS) and clearness index-based Kalman filter (KF_Kt) are discussed.

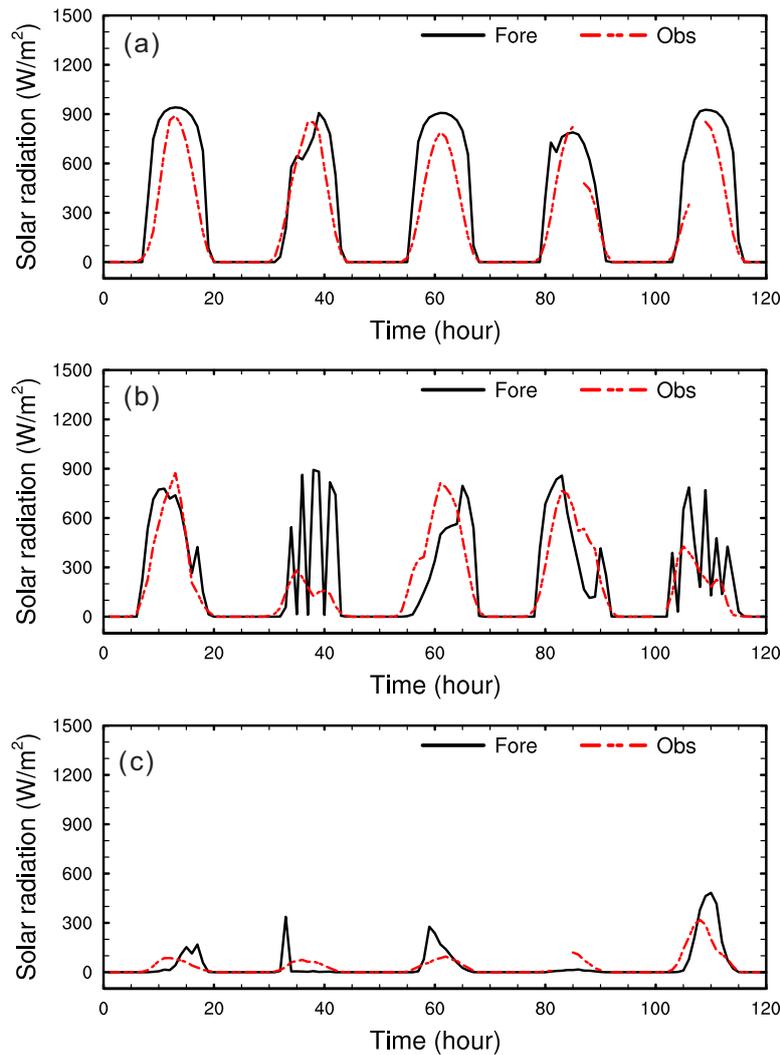


Figure 6. The solar radiation (W/m²) series over five days of clear (a), cloudy (b) and overcast (c) sky conditions for the Maizuru site. The solid black and dotted red lines stand for forecasts and the corresponding observations, respectively.

4.3. Kalman Filtered Forecasts of Solar Radiation

Having confirmed the forecasting ability of the WRF model for the solar radiation prediction, we further apply the Kalman filter to improve the raw prediction. This section is concluded with an analysis of the advantages of KF_Kt with respect to KF_TS.

4.3.1. KF_TS Forecasts

The statistics of MBE, MAE, RMSE, and nRMSE are listed in Table 3 to quantitatively assess the effects of applying the Kalman filter on the time-sequence of forecasts, in comparison with the WRF raw forecasts. It is worthwhile to point out that the values in the second and fourth rows of Table 3 are slightly different with the values in the last column of Table 2. The reason is that we excluded two-week period data from the original data series for training the Kalman filter. Examining the third row of Table 3, the positive value of MBE is observed. It indicates that the KF_TS forecasts of solar radiation are still overestimated for the Maizuru site; however, compared to the raw predictions, the MBE of KF_TS is reduced by 80.2%. The value of RMSE of KF_TS forecasts is much smaller compared to raw estimates, with a relative error reduction of 22.9%. Similarly, for the Qinghai site, smaller MBE value of KF_TS indicates that conducting the Kalman filter on a time-sequence has reduced part of systematic

errors in the raw forecasts. Consistent with MBE, MAE, RMSE, and nRMSE of raw forecasts are all improved to some extent.

Given these results, we can affirm that KF_TS can provide apparent improvement in the situation where the raw forecast presents large errors.

Table 3. Performance statistics of hourly WRF raw and KF_TS forecasts of solar radiation, respectively.

Site	Case	MBE (W/m ²)	MAE (W/m ²)	RMSE (W/m ²)	nRMSE
Maizuru	Raw	70.2	117.3	213.8	0.21
	KF_TS	13.9	87.9	164.8	0.17
Qinghai	Raw	27.1	97.6	183.4	0.08
	KF_TS	−10.0	79.4	158.1	0.06

4.3.2. KF_Kt Forecasts

The comparison of the WRF model and KF_Kt forecasts of solar radiation under clear sky conditions is plotted in Figure 7. It presents hourly WRF raw forecasts (black line) and observed solar radiation at the target site (red line) as well as the KF_Kt predictions (blue line) for a period of 14 days. It is observed from Figure 7a that the WRF model demonstrates the capability of predicting the day-ahead solar radiation at the Maizuru site. Nevertheless, the KF_Kt can improve the raw prediction of the WRF model significantly, with reductions of MBE by 90.1% and RMSE by 43.4%. As expected, from Figure 7b, we also can see the positive impact of KF_Kt on the WRF raw solar prediction for the Qinghai site.

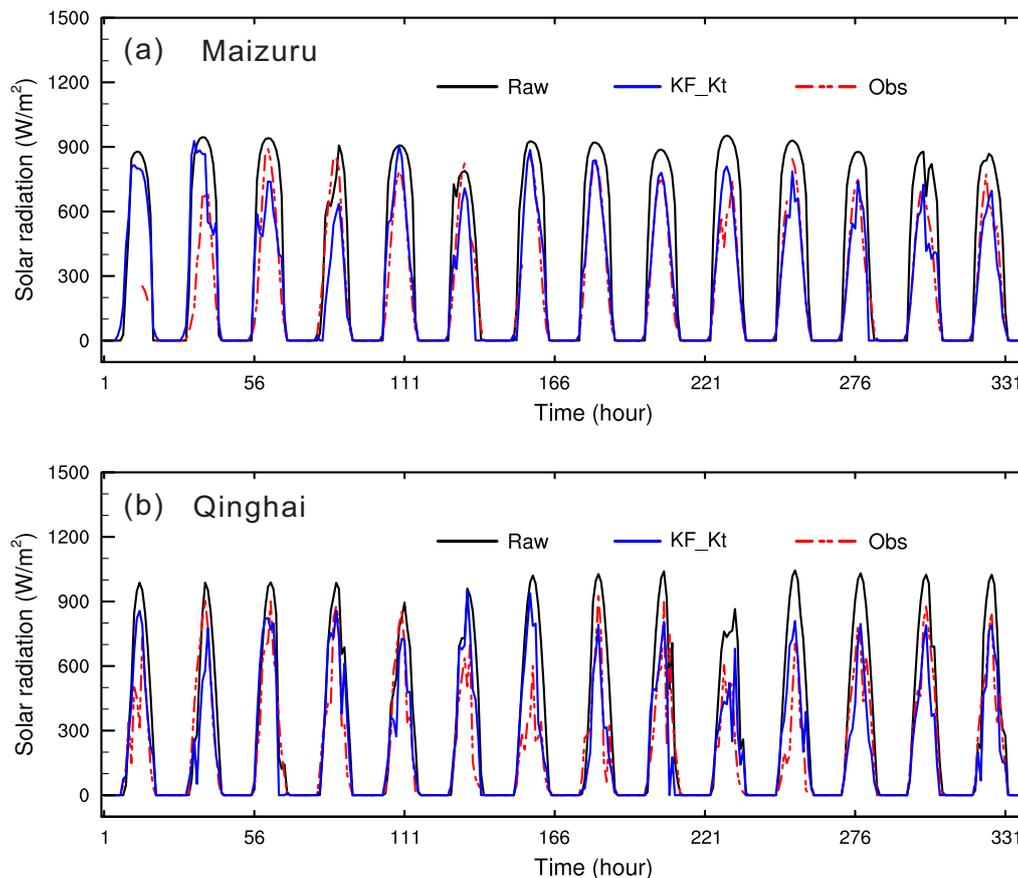


Figure 7. The solar radiation (W/m²) series over 14 days under clear sky condition for Maizuru (a) and Qinghai sites (b). The solid black, blue and dotted red lines stand for WRF raw forecasts, KF_Kt forecasts and the corresponding observations, respectively.

Based on the above discussions, we can primarily conclude that KF_TS and KF_Kt both can make the performance of WRF raw forecasts better. To help to illustrate the priority between those two, the comparison of MBE and RMSE under three sky conditions for both ways of conducting the Kalman filter, are plotted in Figures 8 and 9.

Figure 8a depicts the MBE of the KF_TS and KF_Kt forecast concerning the WRF raw prediction for three sky conditions at the Maizuru site. It can be easily found that the WRF model overestimates the solar radiation under clear and cloudy sky conditions while it underestimates the solar radiation of an overcast day. The most massive MBE value of raw forecasts is found for clear sky conditions (124.8 W/m^2), followed by the cloudy day. Both the KF_TS and KF_Kt can mainly alleviate the systematic error tendency in the raw forecasts of the solar radiation. In the case of clear sky conditions, the reduction of MBE from the KF_TS correction around 54.5%, while that of the KF_Kt is 95.7%, which indicates that KF_Kt performs significantly better than KF_TS under clear sky conditions in removing the systematic errors. Nearly the same positive effects for both KF_TS and KF_Kt can be found for cloudy sky conditions. Note that the Kalman filter without considering the clearness index hardly corrects the systematic errors of solar radiation forecasts under overcast sky conditions.

Figure 8b displays the results of RMSE for the Maizuru site. Similar to MBE, the largest value of RMSE (247.6 W/m^2) and meanwhile the smallest value (130.0 W/m^2) are found for clear and overcast sky conditions, respectively. As expected, for clear sky conditions, the KF_Kt reduces RMSE by 42.2%, which is more significant compared to that of KF_TS forecasts (23.4%). The improvements in RMSE of both KF_TS and KF_Kt are around 12.0% for the cloudy days. KF_Kt can improve the RMSE under overcast sky conditions by 3.2%; however, a decrease of 7.1% can be found for KF_TS. This means that neither Kalman filter nor clearness index-based Kalman filter can improve the day-ahead solar radiation largely under overcast conditions, due maybe to the complexity of WRF outputs.

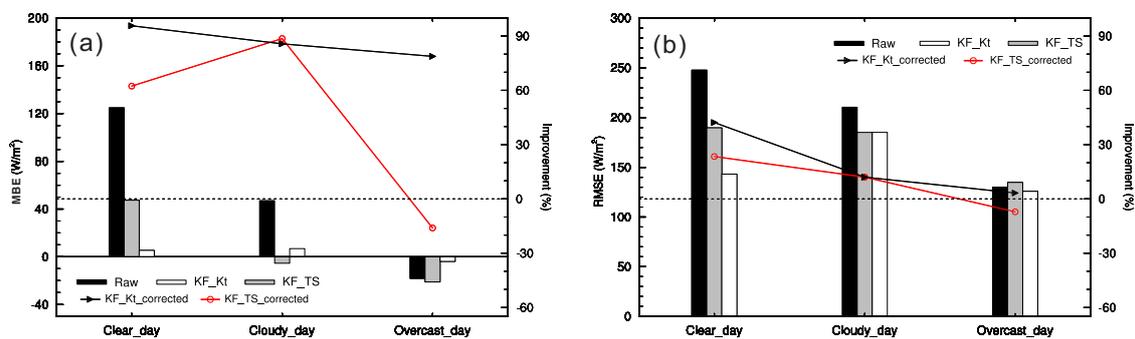


Figure 8. A comparison of the MBE (a) and RMSE (b) of KF_TS (solid gray bar) and KF_Kt (solid white bar) predictions with respect to the raw WRF predictions (solid black bar) for three sky conditions. The marked lines stand for the relative improvement of the KF_TS (red) and KF_Kt (black) against the raw forecasts of the WRF model at the Maizuru site.

The characteristics of MBE and RMSE for the Qinghai site are summarized in Figure 9. The significant advantage of KF_Kt against KF_TS for reducing the systematic errors can be found in Figure 9a. The corrections of MBE by KF_Kt are all around 90% for three sky conditions. As in the case of MBE, RMSE (Figure 9b) of KF_Kt are consistently much smaller than KF_TS under different sky conditions.

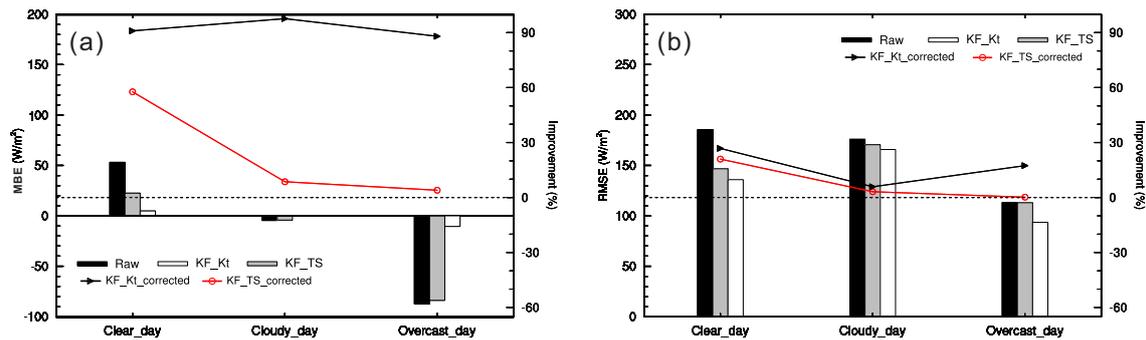


Figure 9. The same as in Figure 8, but for the Qinghai site.

Overall, although the forecasting errors of solar radiation can be reduced by using both KF_TS and KF_Kt, the impact of the latter is much more apparent for both sites, especially the site of Qinghai. This might indicate that the clearness index-based Kalman filter can overcome the shortcomings to some extent when the Kalman filter resolves the data series with substantial day-to-day changes.

5. Conclusions

In the present work, we have established a novel hybrid model for day-ahead solar radiation prediction, based on the mesoscale meteorological WRF model and the clearness index-based Kalman filter post-processing procedure. The hybrid model has been validated at two study sites, i.e., Maizuru and Qinghai, in Japan and China, respectively.

The WRFv3.9.1 model, with a high horizontal resolution of $0.5 \text{ km} \times 0.5 \text{ km}$, is chosen. A sensitivity test of the radiation parameterization schemes was firstly conducted to tune the WRF model for solar radiation forecasts. Based on the results of sensitivity tests, the RRTM longwave scheme is adopted for both sites; the schemes of RRTMG shortwave and Dudhia shortwave are chosen for Maizuru and Qinghai sites, respectively.

Then, the forecasting ability of the WRF model for solar radiation has been validated at two sites. The results indicate that the WRF model has a relatively high ability of forecasting short-term solar radiation, though there are differences among different sky conditions and different locations.

For the sake of improving the raw WRF forecasts of solar radiation, the KF_Kt is conducted, and the result was compared with the corresponding observations, as well as the results of KF_TS. The KF_Kt is able to largely improve MBE by around 80% and RMSE by 20% in day-ahead solar radiation under clear and cloudy sky conditions at the Maizuru site. The performance of KF_Kt is better at the Qinghai site than the Maizuru site. The result also shows the notable advantage of KF_KT against KF_TS for both sites. A relatively small impact of KF_KT under overcast conditions can be found at both study sites.

The hybrid solar radiation model proposed in this present work can be expected as an effective tool for day-ahead operational control for any target sites. It also indicates some new directions worthy of further investigations, for example, development of a more reliable approach which is able to search and group the solar radiation forecasts with similar characteristics of forecasting errors.

Author Contributions: Conceptualization, Y.C. and F.X.; Methodology, Y.C.; Formal analysis, Y.C.; Investigation, L.Y. and L.C.; Validation, L.C., Y.C. and J.Z.; Supervision, F.X.; Writing—original draft, Y.C.; Writing—Review and Editing, F.X. and Y.C.

Funding: This research was funded by the National Natural Science Foundation of China (Grant Nos. 41905084 and 2018YFC1507000) and the Scientific Research Foundation of Chengdu University of Information Technology (Grant No. KYTZ201813).

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

Abbreviations

The following abbreviations are used in this manuscript:

NWP	Numerical weather prediction
WRF	Weather research and forecasting
GFS	Global forecasting system
MBE	Mean bias error
MAE	Mean absolute error
RMSE	Root mean square error
nRMSE	Normalized root mean square error
UTC	Coordinated universal time
MOS	Model output statistics
LSTM	Long short-term memory

References

- Peng, Z.; Yu, D.; Huang, D.; Heiser, J.; Yoo, S.; Kalb, P. 3D cloud detection and tracking system for solar forecast using multiple sky imagers. *Sol. Energy* **2015**, *118*, 496–519. [[CrossRef](#)]
- Hammer, A.; Heinemann, D.; Lorenz, E.; Lücke, B. Short-term forecasting of solar radiation: A statistical approach using satellite data. *Sol. Energy* **1999**, *67*, 139–150. [[CrossRef](#)]
- 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](#)]
- Olatomiwa, L.; Mekhilef, S.; Shamsirband, S.; Mohammadi, K.; Petković, D.; Sudheer, C. A support vector machine-firefly algorithm-based model for global solar radiation prediction. *Sol. Energy* **2015**, *115*, 632–644. [[CrossRef](#)]
- Reikard, G. Predicting solar radiation at high resolutions: A comparison of time series forecasts. *Sol. Energy* **2009**, *83*, 342–349. [[CrossRef](#)]
- Zhang, G.P. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* **2003**, *50*, 159–175. [[CrossRef](#)]
- Raza, M.Q.; Nadarajah, M.; Ekanayake, C. On recent advances in PV output power forecast. *Sol. Energy* **2016**, *136*, 125–144. [[CrossRef](#)]
- Al-Alawi, S.; Al-Hinai, H. An ANN-based approach for predicting global radiation in locations with no direct measurement instrumentation. *Renew. Energy* **1998**, *14*, 199–204. [[CrossRef](#)]
- Benghanem, M.; Mellit, A.; Alamri, S. ANN-based modelling and estimation of daily global solar radiation data: A case study. *Energy Convers. Manag.* **2009**, *50*, 1644–1655. [[CrossRef](#)]
- Benali, L.; Notton, G.; Fouilloy, A.; Voyant, C.; Dizene, R. Solar radiation forecasting using artificial neural network and random forest methods: Application to normal beam, horizontal diffuse and global components. *Renew. Energy* **2019**, *132*, 871–884. [[CrossRef](#)]
- Torres, J.F.; Troncoso, A.; Koprinska, I.; Wang, Z.; Martínez-Álvarez, F. Deep Learning for Big Data Time Series Forecasting Applied to Solar Power. In *International Joint Conference SOCO'18-CISIS'18-ICEUTE'18*; Graña, M., López-Guede, J.M., Etxaniz, O., Herrero, Á., Sáez, J.A., Quintián, H., Corchado, E., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 123–133.
- Srivastava, S.; Lessmann, S. A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data. *Sol. Energy* **2018**, *162*, 232–247. [[CrossRef](#)]
- Qing, X.; Niu, Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy* **2018**, *148*, 461–468. [[CrossRef](#)]
- Mellit, A.; Pavan, A.M. A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. *Sol. Energy* **2010**, *84*, 807–821. [[CrossRef](#)]
- Perez, R.; Lorenz, E.; Pelland, S.; Beauharnois, M.; Van Knowe, G.; Hemker, K., Jr.; Heinemann, D.; Remund, J.; Müller, S.C.; Traunmüller, W.; et al. Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe. *Sol. Energy* **2013**, *94*, 305–326. [[CrossRef](#)]
- Verbois, H.; Huva, R.; Rusydi, A.; Walsh, W. Solar irradiance forecasting in the tropics using numerical weather prediction and statistical learning. *Sol. Energy* **2018**, *162*, 265–277. [[CrossRef](#)]

17. Zamora, R.J.; Dutton, E.G.; Trainer, M.; McKeen, S.A.; Wilczak, J.M.; Hou, Y.T. The accuracy of solar irradiance calculations used in mesoscale numerical weather prediction. *Mon. Weather Rev.* **2005**, *133*, 783–792. [[CrossRef](#)]
18. Lara-Fanego, V.; Ruiz-Arias, J.; Pozo-Vázquez, D.; Santos-Alamillos, F.; Tovar-Pescador, J. Evaluation of the WRF model solar irradiance forecasts in Andalusia (southern Spain). *Sol. Energy* **2012**, *86*, 2200–2217. [[CrossRef](#)]
19. Diagne, M.; David, M.; Boland, J.; Schmutz, N.; Lauret, P. Post-processing of solar irradiance forecasts from WRF model at Reunion Island. *Sol. Energy* **2014**, *105*, 99–108. [[CrossRef](#)]
20. Mejia, J.F.; Giordano, M.; Wilcox, E. Conditional summertime day-ahead solar irradiance forecast. *Sol. Energy* **2018**, *163*, 610–622. [[CrossRef](#)]
21. Al-Yahyai, S.; Charabi, Y.; Gastli, A. Review of the use of Numerical Weather Prediction (NWP) Models for wind energy assessment. *Renew. Sustain. Energy Rev.* **2010**, *14*, 3192–3198. [[CrossRef](#)]
22. Soubdhan, T.; Ndong, J.; Ould-Baba, H.; Do, M.T. A robust forecasting framework based on the Kalman filtering approach with a twofold parameter tuning procedure: Application to solar and photovoltaic prediction. *Sol. Energy* **2016**, *131*, 246–259. [[CrossRef](#)]
23. Che, Y.; Xiao, F. An integrated wind-forecast system based on the weather research and forecasting model, Kalman filter, and data assimilation with nacelle-wind observation. *J. Renew. Sustain. Energy* **2016**, *8*, 053308. [[CrossRef](#)]
24. Hong, S.Y.; Lim, J.O.J. The WRF single-moment 6-class microphysics scheme (WSM6). *Asia-Pac. J. Atmos. Sci.* **2006**, *42*, 129–151.
25. Kain, J.S. The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteorol.* **2004**, *43*, 170–181. [[CrossRef](#)]
26. Chen, F.; Dudhia, J. Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Weather Rev.* **2001**, *129*, 569–585. [[CrossRef](#)]
27. Hong, S.Y.; Noh, Y.; Dudhia, J. A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Weather Rev.* **2006**, *134*, 2318–2341. [[CrossRef](#)]
28. Al-Hamadi, H.; Soliman, S. Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model. *Electr. Power Syst. Res.* **2004**, *68*, 47–59. [[CrossRef](#)]
29. Louka, P.; Galanis, G.; Siebert, N.; Kariniotakis, G.; Katsafados, P.; Pytharoulis, I.; Kallos, G. Improvements in wind speed forecasts for wind power prediction purposes using a Kalman filtering. *J. Wind Eng. Ind. Aerodyn.* **2008**, *96*, 2348–2362. [[CrossRef](#)]
30. Kalman, R.E. A new approach to linear filtering and prediction problems. *J. Fluids Eng.* **1960**, *82*, 35–45. [[CrossRef](#)]
31. Delle Monache, L.; Nipen, T.; Deng, X.; Zhou, Y.; Stull, R. Ozone ensemble forecasts: 2. A Kalman filter predictor bias correction. *J. Geophys. Res.* **2006**, *111*, D05308. [[CrossRef](#)]
32. Duffie, J.A.; Beckman, W.A.; Worek, W.M. *Solar Engineering of Thermal Processes*, 4th ed.; Wiley: Hoboken, NJ, USA, 2003; Volume 116, p. 67.
33. Warner, T.T. Quality assurance in atmospheric modeling. *Bull. Am. Meteorol. Soc.* **2011**, *92*, 1601–1610. [[CrossRef](#)]
34. Long, C.N.; Shi, Y. An automated quality assessment and control algorithm for surface radiation measurements. *Open Atmos. Sci. J.* **2008**, *2*, 23–37. [[CrossRef](#)]
35. Mlawer, E.J.; Taubman, S.J.; Brown, P.D.; Iacono, M.J.; Clough, S.A. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.* **1997**, *102*, 16663–16682. [[CrossRef](#)]
36. Dudhia, J. Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.* **1989**, *46*, 3077–3107. [[CrossRef](#)]
37. Chou, M.D.; Suarez, M.J. An Efficient Thermal Infrared Radiation Parameterization for Use in General Circulation Models. Available online: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.26.4850> (accessed on 3 September 2019).
38. Collins, W.D.; Rasch, P.J.; Boville, B.A.; Hack, J.J.; McCaa, J.R.; Williamson, D.L.; Kiehl, J.T.; Briegleb, B.; Bitz, C.; Lin, S.J.; et al. Description of the NCAR community atmosphere model (CAM 3.0). *NCAR Tech. Note NCAR/TN-464+ STR* **2004**, *226*, 102–117.

39. Iacono, M.J.; Delamere, J.S.; Mlawer, E.J.; Shephard, M.W.; Clough, S.A.; Collins, W.D. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.* **2008**, *113*. [[CrossRef](#)]
40. Chou, M.D.; Suarez, M.J. A solar radiation parameterization for atmospheric studies. *NASA Tech.* **1999**, *15*, 1–35.



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).