Development and Validation of the R Model for Solar Radiation Short-term Nowcasting: Postprint

Authors: Yao Degui, Liu Weijia, Han Yongxiang, Li Zhe, Liang Yun, Liu Weijia

Date: 2023-02-27T00:00:00+00:00

Abstract

Improving the accuracy of solar radiation short-term nowcasting (<6 h) is an important measure to ensure power grid dispatching and also one of the highly challenging technical bottlenecks. Based on the cloud-radiation relationship, a solar radiation short-term nowcasting model (R model) was constructed using the cloud relative radiation forcing ratio retrieved from ground-based solar irradiance observations, and the forecasting performance of the R model was evaluated using 16 years of irradiance observation data from the Southern Great Plains Central Facility. The results show that: (1) In cases with cloud presence, the forecasting performance of the R model is greatly improved compared to the traditional Simple persistence model (Simple model), and still shows 2%~25% improvement compared to the intelligent persistence models with higher forecasting performance (Smart model or RCRF model). (2) In the overall verification comprising 2.9$×$10^5 cases of eight cloud types over 16 years, when the forecast lead time exceeds 1 h, the forecasting performance of the R model is significantly superior to that of the Simple model and the RCRF model. Compared with the RCRF model, the R model can improve the forecasting performance for global radiation and direct radiation by 25% and 19%, respectively, at a 6 h forecast lead time, with the forecast lead time extended by 1.5 h and 1 h, respectively. (3) The R model provides a benchmark model with higher accuracy for solar radiation short-term nowcasting. Meanwhile, this model can perform forecasting relying solely on short-term ground-based irradiance observation data, providing a new approach or possibility for radiation forecasting at photovoltaic power plants lacking synchronous meteorological element observations.

ChinaRxiv

Full Text

Construction and Validation of R Models for Short-Term Solar Irradiance Forecasting

YAO Degui[1], **LIU Weijia**[2], **HAN Yongxiang**[3], **LI Zhe**[1], **LIANG Yun**[1]

[1]Electric Power Research Institute, State Grid Henan Electric Power Company, Zhengzhou 450000, Henan, China
[2]Hangzhou Meteorological Bureau, Hangzhou 310051, Zhejiang, China
[3]Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, China

Abstract

Improving the accuracy of short-term solar irradiance forecasting represents both a crucial safeguard for ensuring grid dispatching security and one of the most challenging technical bottlenecks in the field. Based on the cloud-radiation relationship, this study constructs a short-term solar irradiance forecasting model (R model) utilizing the cloud relative radiative forcing ratio derived from ground-based solar irradiance observations. The forecasting performance of the R model is evaluated using 16 years of irradiance observations from the Southern Great Plains (SGP) Central Facility site in the United States. The results demonstrate that: (1) In cloudy cases, the R model significantly outperforms the traditional simple persistence model and achieves a 2%–25% improvement over the more advanced smart persistence model (also referred to herein as the cloud relative radiative forcing (RCRF) model). (2) In comprehensive validation across 2.9$\times$10^{5}$ cases encompassing eight cloud types over 16 years, the R model exhibits superior forecasting accuracy compared to both the simple and RCRF models at lead times exceeding 1 hour. Specifically, relative to the RCRF model, the R model improves forecasting performance by 25% for global horizontal irradiance and 19% for direct normal irradiance, while extending the forecast lead time by 1.5 hours and 1 hour, respectively. (3) The R model establishes a more accurate benchmark for short-term solar irradiance forecasting. Moreover, since the model requires only short-term ground-based radiation observations, it offers a novel approach and new possibilities for irradiance forecasting at photovoltaic power plants lacking concurrent meteorological observations.

Keywords: persistence model; short-term irradiance forecast; forecasting performance; photovoltaic power plant

1 Introduction

Photovoltaic power generation constitutes a vital component of clean energy and represents an important renewable energy source for achieving carbon neutrality goals. China's arid northwestern regions possess the most abundant solar energy resources in the country. However, due to the inherent volatility, intermittency, and randomness of solar energy, particularly under cloudy conditions, variations in solar irradiance caused by cloud changes can lead to dramatic fluctuations in photovoltaic power output, threatening the security, stability, and economic operation of the power grid. Short-term solar irradiance forecasting (0–6 hours) serves as a critical guarantee for the stable dispatching of photovoltaic power. Improving forecasting accuracy under cloudy conditions is not only essential for optimizing grid dispatching and rational load distribution but also represents one of the most challenging issues in current solar irradiance forecasting research.

With increasing demand for solar photovoltaic power generation, solar forecasting methods have advanced rapidly, evolving from early statistical models (such as autoregressive moving average and exponential smoothing) to numerical weather prediction systems (e.g., WRF-Solar), persistence models, and machine learning approaches. These various forecasting models have been widely applied in solar irradiance prediction, demonstrating high accuracy under clear-sky conditions. However, each has distinct advantages and disadvantages. For single-site, hourly solar irradiance predictions, persistence models and time series methods remain the primary approaches, while mesoscale numerical models dominate for extended spatial scales and longer lead times (0–6 hours). Nevertheless, under cloudy conditions, all these models suffer from low forecasting accuracy.

Notably, except for persistence models, other approaches require extensive observational data support. In contrast, persistence models can forecast future time steps using only a single observation, offering a remarkably simple principle with very high accuracy for short-term clear-sky forecasting. Consequently, persistence models typically serve as benchmark models for evaluating the performance of other forecasting approaches. Furthermore, since photovoltaic power stations are usually equipped only with radiation observation instruments and have limited historical radiation data, and because the reference value of nearby meteorological station data decreases significantly when the distance between meteorological stations and photovoltaic plants exceeds 10 km, persistence models represent a more practical choice for solar irradiance forecasting at photovoltaic power stations.

If a persistence model could be constructed for cloudy conditions with improved forecasting accuracy, it would undoubtedly become a valuable tool for short-term solar irradiance forecasting. Currently, traditional persistence models mainly include the simple persistence model and the smart persistence model. The simple persistence model, which neglects cloud effects and solar movement, typically provides acceptable forecasts only within 15 minutes. The smart persistence model preliminarily accounts for cloud radiative forcing effects by defining the clear-sky index as the ratio of total radiation to clear-sky total radiation. While the smart persistence model shows significant improvement over the simple model under cloudy conditions, it can only forecast global horizontal irradiance (GHI) and cannot effectively predict direct normal irradiance (DNI).

Liu Weijia developed the RCRF model based on cloud relative radiative forcing, which can be considered an extension of the smart persistence model capable of forecasting both GHI and DNI. However, the RCRF model only considers the combined effects of clouds on radiation without accounting for more detailed macroscopic features such as cloud albedo. According to radiative transfer equations, the relative radiative forcing ratio derived from ground-based irradiance observations can approximately represent cloud albedo. Building upon this principle and employing a persistence approach, this study constructs a new solar persistence model (R model) and evaluates its performance using 16 years of observations from the SGP Central Facility site, aiming to provide a more accurate benchmark model for short-term solar irradiance forecasting.

1.1 Data Sources

The data used in this study are obtained from the Atmospheric Radiation Measurement (ARM) program' s Southern Great Plains (SGP) Central Facility site (https://www.arm.gov/). The dataset spans 16 years (in Central Standard Time) with a temporal resolution of 15 minutes. This data product is derived from 1-minute observations that have undergone averaging and rigorous quality control procedures, including fitting and other validation processes. The dataset, directly available from the website, contains 2.9$×10^{5}$ samples across eight cloud types: shallow cumulus, other low clouds, dense cumulus, deep convective clouds, altocumulus, altostratus, and cirrostratus. The data have undergone strict quality control and are widely used in atmospheric cloud research. The forecast lead time in this study is 15 minutes.

1.2 Construction of the Relative Radiative Forcing Ratio-Based Persistence Model and Performance Evaluation

1.2.1 Model Construction In the simplified radiative transfer equation, relative radiative forcing (RCRF) is defined as:

$$RCRF_i = 1 - \frac{F_{all,i}}{F_{clr,i}}$$

where the subscript i denotes the parameter identifier, which can represent either global horizontal irradiance (GHI) or direct normal irradiance (DNI). $F_{all,i}$ and $F_{clr,i}$ represent the downward irradiance and clear-sky irradiance, respectively. The cloud relative radiative forcing ratio (R) is defined as:

$$R = \frac{RCRF_{GHI}}{RCRF_{DNI}}$$

where $RCRF_{GHI}$ and $RCRF_{DNI}$ are the relative radiative forcings for GHI and DNI, respectively. For convenience, the vertical component is hereafter referred to as $F_{all,DNI}$.

The persistence model assumes that the value of a variable (V) at a future time (t_f) is a projection of its value at a historical time (t), i.e., $V(t_f) = V(t)$. Therefore, the model based on the persistence of the cloud relative radiative forcing ratio (R) satisfies:

$$R(t_f) = R(t)$$

By substituting the terms from the definition of R into the equation for time t_f and replacing $R(t_f)$ with $R(t)$, we obtain:

$$R(t) = \frac{RCRF_{GHI}(t_f)}{RCRF_{DNI}(t_f)}$$

Rearranging this equation yields:

$$RCRF_{GHI}(t_f) = R(t) \times RCRF_{DNI}(t_f)$$

Substituting the definition of $RCRF_{GHI}$ into this equation and rearranging gives:

$$1 - \frac{F_{all,GHI}(t_f)}{F_{clr,GHI}(t_f)} = R(t) \times RCRF_{DNI}(t_f)$$

Further substitution and rearrangement produce the forecasting equation based on cloud relative radiative forcing ratio (R):

$$F_{all,GHI}(t_f) = \left[1 - R(t) \times RCRF_{DNI}(t_f)\right] \times F_{clr,GHI}(t_f)$$

$$F_{all,DNI}(t_f) = \left[1 - RCRF_{DNI}(t_f)\right] \times F_{clr,DNI}(t_f)$$

The clear-sky irradiance $F_{clr,i}$ in these equations is estimated using a five-point weighted smoothing method:

$$RCRF(t_f) = \sum_{j=1}^{t} a_j \times RCRF(t-j)$$

where j represents the j-th time step in the moving process, t represents the total number of past time steps used, and a is the smoothing coefficient (typically set to 1/5 for five-point weighted averaging).

For comparison, the forecasting formulas for the simple persistence model and the smart persistence model (as an extension of the RCRF model) are also provided:

$$F_{all,i}(t_f) = F_{all,i}(t)$$

$$F_{all,i}(t_f) = [1 - RCRF_i(t)] \times F_{clr,i}(t_f)$$

1.2.2 Forecast Performance Evaluation Metrics This study employs the percentage root mean square error (PE) to evaluate model forecasting performance, defined as:

$$PE = \frac{RMSE}{\bar{y}_{obs}} \times 100\%$$

where $RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(y_{obs,k} - y_{fcst,k})^2}$, k and n represent the k-th and total number of time steps, respectively; $y_{obs,k}$ and $y_{fcst,k}$ are the observed and forecasted values; and $\bar{y}_{obs}$ is the mean observed value. PE is a dimensionless quantity, with smaller values indicating better model performance.

Additionally, this study introduces the PE-SRCRF metric to characterize the performance improvement of the new model relative to the reference model:

$$S_{reference} = \frac{PE_{model} - PE_{reference}}{PE_{reference}} \times 100\%$$

where the subscripts "model" and "reference" denote the evaluated new model and reference model, respectively. Positive values indicate performance improvement, while negative values indicate inferior performance relative to the reference model.

---

2 Results and Analysis

2.1 Case Study of Radiation Components Forecasting Under Cloudy Conditions

Using radiation data from September 7, 2002, at the SGP Central Facility site as an example, which includes both clear-sky and cloudy periods, [Figure 1: see original paper] shows cloud observations from the total sky imager and corresponding irradiance measurements. Before 10:00, the sky was cloudless, and both GHI and DNI showed continuous upward trends approaching peak values of approximately 800 $\mathrm{W \cdot m^{-2}}$ and 600 $\mathrm{W \cdot m^{-2}}$, respectively, consistent with clear-sky irradiance characteristics. From approximately 12:00 to 16:00, scattered cloud patches appeared over the station, causing significant fluctuations in observed irradiance during this period and resulting in observed values substantially lower than clear-sky irradiance, demonstrating that clouds simultaneously reduce both GHI and DNI.

The simple persistence model, RCRF model, and newly constructed R model were used to forecast GHI and DNI at lead times of 1, 2, and 3 hours. Since the RCRF model can be considered an extension of the smart persistence model across radiation components, the smart persistence model is hereafter represented by RCRF. [Figure 2: see original paper] presents the observed GHI and DNI values alongside forecasts from the three models at different lead times. Visual inspection reveals that the R model outperforms both the simple and RCRF models, with both R and RCRF showing significantly better performance than the simple model. To quantitatively assess model performance, PE values were calculated for all three lead times, as shown in . The results indicate that PE values increase with lead time for all models, reflecting growing forecast errors and decreasing performance. However, the PE values for the R model are substantially lower than those for the simple and RCRF models across all lead times. Specifically, at the 1-hour lead time, the R model reduces PE by 2.9$ $\times 10^{\{2\}\times 10}$\{2\}$% for DNI. Even when compared to the better-performing RCRF model, the R model achieves improvements of 2%–18% across the three lead times, with greater improvement for GHI than DNI. Although all models show increasing PE values with lead time, the R model maintains superior performance, with its advantages becoming more pronounced at longer lead times.

[Figure 3: see original paper] displays scatter plots of predicted versus observed values at the 3-hour lead time. The results show that for both GHI and DNI, the simple and RCRF models exhibit wider scatter distributions, with high-frequency regions located away from the diagonal line, indicating substantial deviations between observations and predictions. In contrast, the R model shows high-frequency scatter concentrated along the diagonal, demonstrating smaller deviations. While both simple and RCRF models tend to overestimate high values, their performance remains significantly inferior to the R model. These results suggest that the R model offers particular advantages at longer lead times.

2.2 Overall Sample Forecast Performance Validation

Although the case study demonstrates superior forecasting performance, conclusions drawn from a single case may be subject to considerable randomness. Therefore, this study further validates the new model using 16 years of daytime observations from the SGP Central Facility site. The dataset comprises

2.9$×10^{5}$ samples across eight cloud types: shallow cumulus, other low clouds, dense cumulus, deep convective clouds, altocumulus, altostratus, and cirrostratus.

[Figure 4: see original paper] illustrates the variation of PE values with forecast lead time. For both GHI and DNI, when the lead time is within 1 hour, the R model' s performance is slightly lower than that of the simple model. However, beyond 1 hour, as lead time increases, the R model not only significantly outperforms the simple model but also surpasses the RCRF model. Compared to the better-performing RCRF model, the R model improves forecasting performance by 25% for GHI and 19% for DNI, with improvements increasing gradually with lead time.

Another critical metric for evaluating forecast performance is the extension of lead time. Since the simple model demonstrates good prediction accuracy within 1 hour, its PE value serves as a reference for determining whether other models extend or shorten the forecast lead time. The method involves identifying the lead time at which other models first exceed the simple model' s PE reference value. In [Figure 4: see original paper], the horizontal red dashed line represents the simple model' s PE value. The results show that compared to the simple model, the RCRF model extends the lead time from 1 hour to 1.5 hours for GHI and from 1 hour to 1.75 hours for DNI. The R model further extends these lead times to 4.5 hours and 2.75 hours, respectively—representing extensions of 3.5 hours for GHI and 1.75 hours for DNI relative to the simple model, and 3 hours for GHI and 1 hour for DNI relative to the RCRF model.

[Figure 5: see original paper] presents the PE-SRCRF score of the R model relative to the RCRF model across forecast lead times. Positive values indicate superior R model performance, while negative values indicate inferior performance. The results demonstrate that the R model outperforms the RCRF model across all lead times, with performance improvements gradually increasing as lead time extends. Notably, this study derives the core forecasting factor—the cloud relative radiative forcing ratio—based on the relationship between clouds and radiation [Equation (1)]. This ratio simultaneously considers the combined effects of clouds on both GHI and DNI. The RCRF model only utilizes $RCRF_{GHI}$, while the R model incorporates both $RCRF_{GHI}$ and $RCRF_{DNI}$, suggesting that introducing additional radiation information enhances forecasting accuracy. According to the forecasting equations, the R model requires only short-term ground-based radiation observations without auxiliary observations (such as cloud measurements), providing a new approach for irradiance forecasting at photovoltaic power plants lacking concurrent meteorological data.

---

3 Discussion

The persistence model constructed based on relative radiative forcing ratio is a mathematical model with physical significance. In lightly polluted regions such as the SGP Central Facility site, the relative radiative forcing ratio can approximate cloud albedo. Validation across 2.9$×10^{5}$ cases and eight cloud types demonstrates that its short-term forecasting performance is significantly superior to the current best-performing smart persistence model. However, in heavily polluted areas, theoretical derivation suggests that the relative radiative forcing ratio contains information about both cloud albedo and pollutants. For instance, northwestern China’ s desert regions consistently experience massive dust aerosol loading, requiring further observational data to validate the model’ s universality.

Additionally, the theoretical derivation assumes a single, uniform cloud layer, while the actual atmosphere may present more complex scenarios such as multi-layer clouds, horizontal cloud inhomogeneity, three-dimensional effects, horizontal photon transport, and influences from aerosol particles and water vapor. These factors warrant more in-depth investigation in future research.

4 Conclusions

This study constructs a short-term solar irradiance forecasting system (R model) using the cloud relative radiative forcing ratio derived from ground-based irradiance observations. The forecasting performance of the newly developed R model is evaluated using both a 16-year cloudy case from the SGP site and 2.9$×10^{5}$ samples across eight cloud types, with comparisons made against traditional persistence models (simple and RCRF). The main conclusions are as follows:

1. For cloudy cases, the R model demonstrates substantial improvements over the simple persistence model, with PE values decreasing by 2.9$×10^{2}$% for both GHI and DNI. Even when compared to the better-performing RCRF model, the R model achieves 2%–25% improvement, with greater enhancement for GHI than DNI.
2. In comprehensive validation using 2.9$×10^{5}$ samples, the R model shows slightly lower performance than the simple model within 1-hour lead times. However, beyond 1 hour, its performance becomes significantly better than both the simple and RCRF models. Compared to the RCRF model, the R model improves forecasting performance by 25% for GHI and 19% for DNI, extending lead times by 1.5 hours and 1 hour, respectively. Relative to the simple model, the R model extends lead times by up to 4.5 hours for GHI and 2.75 hours for DNI.
3. The superior performance of the R model implies that incorporating more cloud information enhances forecasting accuracy. The R model provides a more accurate benchmark for solar irradiance forecasting under cloudy conditions. Since it requires only short-term ground-based radiation observations, it offers new possibilities for short-term solar irradiance fore-

casting at photovoltaic power plants lacking concurrent meteorological observations.

References

[1] Hu Angang. China' s goal of achieving carbon peak by 2030 and its main approaches[J]. Journal of Beijing University of Technology (Social Sciences Edition), 2021, 21(3): 1-15.

[2] Zou Caineng, Xiong Bo, Xue Huaqing, et al. The role of new energy in carbon neutral[J]. Petroleum Exploration and Development, 2021, 48(2): 411-420.

[3] Cui Rongqiang. Solar photovoltaic power: Hope of low carbon economy in China[J]. Chinese Journal of Nature, 2010, 32(3): 149-154.

[4] Liu Jia, He Qing, Liu Rui, et al. Solar radiation character and solar energy resource in Xinjiang[J]. Arid Meteorology, 2008, 26(4): 61-66.

[5] Sang Jianren, Liu Yulan, Lin Li. Characteristic of solar radiation in Ningxia and integrated evaluation on utilization potential of solar energy[J]. Journal of Desert Research, 2006, 26(1): 122-125.

[6] Fouad M M, Shihata L A, Morgan E I. An integrated review of factors influencing the performance of photovoltaic panels[J]. Renewable and Sustainable Energy Reviews, 2017, 80: 1499-1511.

[7] Zhang Xueli, Liu Qihui, Ma Huimeng, et al. Analysis of influencing factors of output power of photovoltaic power plant[J]. Advances of Power System & Hydroelectric Engineering, 2012, 28(5): 75-81.

[8] Alsharif M H, Younes M K, Kim J. Time series ARIMA model for prediction of daily and monthly average global solar radiation: The case study of Seoul, South Korea[J]. Symmetry, 2019, 11(2): 240.

[9] Jimenez P A, Hacker J P, Dudhia J, et al. WRF Solar: Description and clear sky assessment of an augmented NWP model for solar power prediction[J]. Bulletin of the American Meteorological Society, 2016, 97(7): 1249-1264.

[10] Diagne M, David M, Lauret P, et al. Review of solar irradiance forecasting methods and a proposition for small scale insular grids[J]. Renewable and Sustainable Energy Reviews, 2013, 27: 65-76.

[11] Aggarwal S K, Saini L M. Solar energy prediction using linear and linear regularization models: A study on AMS (American Meteorological Society) 2013-14 solar energy prediction contest[J]. Energy, 2014, 78: 247-256.

[12] Voyant C, Notton G, Kalogirou S, et al. Machine learning methods for solar radiation forecasting: A review[J]. Renewable Energy, 2017, 105: 569-582.

[13] Ma Jinyu, Luo Yong, Shen Yanbo, et al. A review on methods of solar energy forecasting and its application[J]. Resources Science, 2011, 33(5): 829-837.

[14] Jiang Junxia, Gao Xiaoqing, Lü Qingquan. Study on cloud tracking and solar irradiance ultra short term forecasting based on TSI images[J]. Acta Energiae Solaris Sinica, 2020, 41(5): 351-358.

[15] Lara Fanego V, Ruiz Arias J A, Pozo Vázquez D, et al. Evaluation of the WRF model solar irradiance forecasts in Andalusia (southern Spain)[J]. Solar Energy, 2012, 86(8): 2200-2217.

[16] Zhao Dong. Long-term changes and assessment models of solar energy over China[D]. Nanjing: Nanjing University of Information Science and Technology, 2009.

[17] Liu B Y H, Jordan R C. The interrelationship and characteristic distribution of direct, diffuse and total solar radiation[J]. Solar Energy, 1960, 4(3): 1-19.

[18] Kumler A, Xie Y, Zhang Y. A physics model for intra-hour forecasting of solar radiation (PSPI) using GHI measurements and a cloud retrieval technique[J]. Solar Energy, 2019, 177: 494-500.

[19] Xie Y, Liu Y. A new approach for simultaneously retrieving cloud albedo and cloud fraction from surface based shortwave radiation measurements[J]. Environmental Research Letters, 2013, 8(4): 044023.

[20] Long C N, Gaustad K L. The shortwave (SW) clear sky detection and fitting algorithm: Algorithm operational details and explanations[R]. United States: DOE Office of Science Atmospheric Radiation Measurement (ARM) User Facility, 2004.

[21] Liu Y, Wu W, Jensen M P, et al. Relationship between cloud radiative forcing, cloud fraction and cloud albedo, and new surface based approach for determining cloud albedo[J]. Atmospheric Chemistry and Physics, 2011, 11(14): 7155-7170.

[22] Liu Weijia. Construction and evaluation of cloud property based solar radiation short term forecast systems[D]. Nanjing: Nanjing University of Information Science and Technology, 2021.

[23] Reda I, Andreas A. Solar position algorithm for solar radiation applications[J]. Solar Energy, 2004, 76(5): 577-589.

[24] Yang D, Alessandrini S, Antonanzas J, et al. Verification of deterministic solar forecasts[J]. Solar Energy, 2020, 210: 20-37.

[25] Yu Zhixiang, Li Xia, Yu Xiaojing, et al. Spatiotemporal variation characteristics of aerosol optical depth in Xinjiang from 2003 to 2019[J]. Arid Land Geography, 2022, 45(2): 346-358.

[26] Sun Xiaoyun, Fang Yanjie, Zhao Jingfeng, et al. Spatial and temporal distribution characteristics of sand drift potential in Taklimakan Desert[J]. Arid

Land Geography, 2020, 43(1): 38-47.

[27] Liu Chang, Li Chengzhi, Li Shenghui, et al. Desertification analysis based on grid accumulation method in Tarim Basin, China[J]. Arid Land Geography, 2021, 44(1): 197-207.

Note: Figure translations are in progress. See original paper for figures.