Solar Energy 215 (2021) 252–265

Contents lists available at ScienceDirect

Solar Energy

journal homepage: www.elsevier.com/locate/solener

Use of physics to improve solar forecast: Physics-informed persistence models for simultaneously forecasting GHI, DNI, and DHI

Weijia Liu [a,b,*], Yangang Liu [a,*], Xin Zhou [a], Yu Xie [c], Yongxiang Han [b], Shinjae Yoo [a], Manajit Sengupta [c]

[a] Brookhaven National Laboratory, Upton, NY, USA
[b] Nanjing University of Information Science and Technology, Nanjing, Jiangsu, China
[c] National Renewable Energy Laboratory, Golden, CO, USA

ARTICLE INFO

Keywords:
Physics-informed persistence models
Solar irradiance forecasting
Cloud predictor
Forecast accuracy

ABSTRACT

Observation-based statistical models have been widely used in forecasting solar energy; however, existing models often lack a clear relation to physics and are limited largely to global horizontal irradiance (GHI) forecasts over relatively short time horizons (<1 h). Incorporating physics into observation-based models, increasing forecast time horizons and developing a model system for forecasting not only GHI but also direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI) remain challenging, especially under cloudy conditions because of complex cloud-radiation interactions. This work attempts to address these challenges by developing a hierarchy of four new physics-informed persistence models that can be used to simultaneously forecast GHI, DNI and DHI. The decade-long measurements (1998 to 2014) at the U.S. Department of Energy's Atmospheric Radiation Measurement (ARM)'s Southern Great Plains (SGP) Central Facility site are used to evaluate the performance of the new models. The results show that the new physics-informed forecast models generally outperform the simple and smart persistence models, and improve the forecast accuracy at lead times from 1.25 h up to 6 h. Further analysis reveals that the forecast error is highly related to the error and temporal variability of the assumed cloud predictor. The best model for forecasting different radiative components can be explained by the relationship between solar irradiances and cloud properties.

1. Introduction

The growing demands for solar energy call for accurate solar resource forecasting to optimize energy management and grid operation (Beltran et al., 2012; Inman et al., 2013; Kleissl, 2013). Observation-based models have been widely used in forecasting solar radiation over short-term timescales, from 5 min up to 6 h (Reikard, 2009). The simplest form of those models, exploiting the temporal persistence of a variable, has been often used as a reference to evaluate more advanced models (Diagne et al., 2013). The simple persistence model, assuming persistent solar irradiance, predicts solar irradiance without knowledge of cloud or weather conditions, but it degrades quickly at lead times > 1 h due to neglecting cloud impacts and sun movement (Martín et al., 2010; Kleissl, 2013). The smart persistence model, which assumes the persistence of clear-sky index defined as the ratio of all-sky global horizontal irradiance (GHI) to clear-sky GHI (Liu and Jordan 1960), improves the simple persistence model by accounting for the overall effect of cloud on solar irradiance. However, it does not differentiate between radiative influences from different cloud properties. Machine learning models and statistical techniques —e.g., Auto-Regressive Integrated Moving Average (ARIMA), and multiple regressions and exponential smoothing—are also applied to improve the solar irradiance forecast (Yang et al., 2012; Lauret et al., 2015; Yang et al., 2015; Voyant et al., 2017; Yang et al., 2017). Most such models rely on statistical relations between prediction and observation with empirically determined weights that are hard to interpret with a clear physical meaning. Furthermore, the observation-based models are often confined to forecast GHI, although a few attempts have been made to infer direct normal irradiance (DNI) or diffuse horizontal irradiance (DHI) (Law et al., 2014; Chu et al., 2015; Bailek et al., 2018).

Solar irradiance forecasting under cloudy conditions is particularly challenging due to complex cloud-radiation interactions (Ramanathan et al., 1989; Rosenfeld, 2006; Matus and L'Ecuyer, 2017) and multiscale variability of cloud properties (Liu, 2019). On one hand, clouds can

* Corresponding authors at: Brookhaven National Laboratory, Upton, NY, USA (W. Liu and Y. Liu).
E-mail addresses: weijialiu90@outlook.com (W. Liu), lyg@bnl.gov (Y. Liu).

https://doi.org/10.1016/j.solener.2020.12.045
Received 8 September 2020; Received in revised form 17 December 2020; Accepted 22 December 2020

change the energy budget by scattering and absorbing solar radiation, which in turn depends on cloud microphysical properties (Kobayashi, 1989; Pfister et al., 2003; Kubar et al., 2009), macroscopic properties (e.g., cloud fraction, cloud thickness), and multiscale structures of a cloud field (Shonk et al., 2010; Shonk and Hogan, 2010). On the other hand, radiative heating and cooling can alter the atmospheric vertical structure and atmospheric dynamics, changing energy redistribution of the cloud-laden atmospheric system, convective process, and cloud activities (Okata et al., 2017).

In search of a theoretical framework that connects cloud properties with solar irradiances for retrievals and model evaluation, we (Liu et al., 2011; Xie and Liu, 2013) developed a set of equations that relates cloud fraction and cloud albedo physically to reduced dimensionless quantities derived from a combination of solar irradiance components. Kumler et al. (2019) recently presented a cloud-optical-depth-based persistence model and showed improved GHI forecast over intra-hour time horizons compared to the simple and smart persistence models.

In this work, we first present a hierarchy of four physics-informed persistence models based on the theoretical formulation under different levels of physical approximations that permit forecasting not only GHI but also its partitioning into DNI and DHI. We then evaluate the performance of the new models, in comparison with the simple and smart persistence models, using decade-long (1998 to 2014) measurements at the U. S. Department of Energy (DOE) Atmospheric Radiation Measurement (ARM)'s Southern Great Plains (SGP) Central Facility site. The reasons underlying the performance discrepancies are dissected as well. The results shed new light on further improving observation-based models to forecast GHI, DNI, and DHI.

The rest of this paper is organized as follows: the theoretical framework for the physics-informed persistence models is described in Section 2; the performances of the new models against the simple and smart persistence models are given in Section 3. Further analyses accounting for the performance behavior are discussed in Section 4. The conclusion and further work are given in Section 5.

2. Theoretical framework for the physics-informed persistence models

2.1. Simple and smart persistence models

The simple and smart persistence models are well known in the solar forecast community; we briefly introduce here to allow the readers to clearly see the connection to our new models as physics is gradually incorporated. Based on the assumptions of the persisted solar irradiance and clear-sky index, the simple persistence model directly projects the historical solar irradiance to the future time steps, and the smart model decomposes the forecasting of GHI into the computation of clear-sky index and clear-sky irradiance. Mathematically, the corresponding forecasting equation for the simple persistence model is given by

$$F^{dn}_{all,i}(t_f) = F^{dn}_{all,i}(t), \tag{1}$$

where F^{dn} represents the downward solar irradiance, the subscript '*all*' denotes the all-sky condition; t_f and t are the target and historical time such that $t_f = t + \Delta t$ with Δt being the forecast lead time. The subscript i can represent GHI, DNI, and DHI (note that $F^{dn}_{all,DNI}$ refers actually to the vertical component of DNI here). The forecast equation for the smart persistence model can be written as

$$F^{dn}_{all,GHI}(t_f) = K(t) \times F^{dn}_{clr,GHI}(t_f), \tag{2a}$$

where the subscript '*clr*' refers to the estimated clear-sky condition, and the clear-sky component is assessed with the method reported in Long and Ackerman (2000) and Long and Gaustad (2004). The clear-sky index K in Eq. (2a) is defined as

$$K = F^{dn}_{all,GHI} / F^{dn}_{clr,GHI} \tag{2b}$$

Despite its utility, the simple persistence model is not capable of forecasting solar irradiances beyond an hour as a result of neglecting the impacts of clouds and sun movement (Martín et al., 2010; Kleissl, 2013). The smart persistence model reduces such effects somewhat, but it is limited to forecasting GHI only. Also, as will be shown later, the smart persistence model lumps together the different effects of cloud fraction and cloud albedo on GHI without clear differentiation.

2.2. Physics-informed persistence model systems

To further improve on the simple and smart persistence models and explore the potentials of incorporating physics and developing a model system that consistently forecasts not only GHI but also DNI and DHI, we build a hierarchy of four physics-informed persistence model systems based on the theoretical framework relating solar irradiances (GHI and DNI) and cloud properties (cloud fraction and cloud albedo) formulated by Liu et al. (2011) and Xie and Liu (2013). This section briefly introduces the theoretical framework in the context of developing the forecast systems and clarifying the underlying physics; details are referred to the original publications.

Based on the concept of cloud radiative forcing widely used in climate studies, Liu et al. (2011) introduced the concept of relative cloud radiative forcing (*RCRF*) for GHI defined as

$$RCRF_{GHI} = (F^{dn}_{clr,GHI} - F^{dn}_{all,GHI}) / F^{dn}_{clr,GHI} \tag{3}$$

They further showed theoretically that $RCRF_{GHI}$ is an approximate product of cloud fraction and cloud albedo. Xie and Liu (2013) extended this work by combining GHI and DNI to derive the following set of theoretical relationships among cloud albedo (α_r), cloud fraction (f), GHI, and DNI given by

$$F^{dn}_{all,GHI} = F^{dn}_{clr,GHI} - \alpha_r \times f \times (F^{dn}_{clr,GHI} - F^{up}_{all,GHI} T^2), \tag{4a}$$

$$F^{dn}_{all,DNI} = [1 - f + f \times \exp(-\tau/\mu_0)] \times F^{dn}_{clr,DNI}, \tag{4b}$$

$$\tau = 2\alpha_r \mu_0 / [(1 - \alpha_r) \times (1 - g)], \tag{4c}$$

where $F^{up}_{all,GHI}$ is the all-sky upwelling flux; $T \approx 2\int_0^1 \mu_0 T^{clr}_{dd}(\mu_0) d\mu_0$ is the transmittance of diffuse radiation of the atmosphere with μ_0 representing the cosine of zenith angle and T^{clr}_{dd} indicating the transmittance of direct radiation that can be computed by a radiative transfer model in clear sky (e.g., Bird and Hulstrom, 1981); τ is the cloud optical depth; the mean value of asymmetry factor g is set as a constant 0.86. A combination of Eq. (4) yields

$$\frac{\alpha_r}{1 - \exp\left[-\frac{2\alpha_r \mu_0}{(1-\alpha_r)\times(1-g)}\right]} = \frac{B1}{B2}, \tag{5a}$$

where $B_1 = \frac{F^{dn}_{clr,GHI} - F^{dn}_{all,GHI}}{F^{dn}_{clr,GHI} - F^{up}_{all,GHI} T^2}$, and B_2 is the RCRF for DNI given by

$$B_2 = RCRF_{DNI} = (F^{dn}_{clr,DNI} - F^{dn}_{all,DNI}) / F^{dn}_{clr,DNI} \tag{5b}$$

Eq. (5) reveals that cloud albedo is essentially a function of the ratio B_1/B_2, and can be well approximated by the following piecewise polynomials

$$\alpha_r = \begin{cases} 0, \quad \text{for } B_1B_2 = 0 \text{ or } 0.07 < \frac{B_1}{B_2} < 0.07872 \\ 1 - 31.1648\frac{B_1}{B_2} + \sqrt{\left(31.1648\frac{B_1}{B_2}\right)^2 - 49.6255\frac{B_1}{B_2}} \\ \quad , \quad \text{for } 0.07872 \le \frac{B_1}{B_2} \le 0.11442 \\ \frac{2.61224B_1 - B_2 + \sqrt{24.2004{B_1}^2 - 9.0098B_1B_2 + {B_2}^2}}{18.3622B_1 - 4B_2} \\ \quad , \quad \text{for } 0.11442 \le \frac{B_1}{B_2} \le 0.185 \\ 0.89412\frac{B_1}{B_2} + 0.02519 \quad , \quad \text{for } 0.185 < \frac{B_1}{B_2} \le 0.23792 \\ \frac{B_1}{B_2} \quad \text{for } 0.23792 < \frac{B_1}{B_2} \le 1.0 \end{cases} \tag{6}$$

Once α_r is determined, cloud fraction f can be estimated by Eq. (4a). Two points are worth highlighting for this study. First, B_1 can be approximated by $RCRF_{GHI}$ because the term $F^{up}_{all,GHI}T^2$ is much smaller than $F^{dn}_{clr,GHI}$ in the B_1 denominator (5%~7%) and thus can be ignored. Thus, the theoretical formulation reveals that cloud albedo is essentially a function of the ratio of the *RCRF* for GHI and DHI defined as

$$R = RCRF_{GHI}/RCRF_{DNI}. \tag{7}$$

In other words, the *RCRF* ratio *R* is essentially determined by cloud albedo. Second, with the approximation of $B_1 = RCRF_{GHI}$, Eq. (4a) is reduced to

$$F^{dn}_{all,GHI} = (1-\alpha_r \times f) \times F^{dn}_{clr,GHI} \tag{8a}$$

$$\text{or } f = B_1/a_r \approx RCRF_{GHI}/a_r \approx RCRF_{DNI}. \tag{8b}$$

Equation (8b) reveals that $RCRF_{DNI}$ is essentially determined by cloud fraction. It is noteworthy that Yang et al. (2012) showed an empirical relationship between cloud fraction and DNI clear-sky index defined as the ratio of all-sky DNI to the clear-sky DNI, providing observational support for the theoretical Eq. (8a).

According to the above theoretical analysis, we can build a hierarchy of persistence model systems to forecast GHI, DNI, and DHI, with different levels of incorporating physics as summarized in Table 1. The corresponding forecast systems are described next.

The 1st level persistent model directly assumes the persistence of GHI, DNI and GHI; an example is the well-known simple persistence model for GHI. The 2nd level forecast model assumes the persistence of clear-sky index or *RCRF*s (RCRF-based persistence model hereafter), and the forecasting equation is given by

$$F^{dn}_{all,i}(t_f) = [1 - RCRF_i(t)] \times F^{dn}_{clr,i}(t_f), \tag{9}$$

where the subscript *i* denotes GHI, DNI, and DHI. It can be readily shown that $RCRF_{GHI}$ has a linear relationship with *K* given by

Table 1
A summary of cloud-radiation relationships at different levels of approximation.

Hierarchy level	Persistent predictor	Cloud physics incorporated
1st level	$F^{dn}_{all,GHI}$, $F^{dn}_{all,DNI}$, $F^{dn}_{all,DHI}$	No direct cloud physics
2nd level	*K or RCRFs*	Overall cloud effects
3rd level	*R*	Approximate separation of radiative effects from cloud albedo and cloud fraction
4th level	α_r,f	Clear separation of radiative effects from cloud albedo and cloud fraction

$$RCRF_{GHI} = 1 - K. \tag{10}$$

Eq. (10) reveals that the smart persistence model is equivalent to the RCRF-based forecast model for GHI, and thus the RCRF-based forecast system (hereafter RCRF-PM) can be viewed as an extension of the smart persistence model to encompass DNI and DHI in addition to GHI. Based on Liu et al. (2011), $RCRF_{GHI}$ equals approximately to cloud fraction times cloud albedo; thus the RCRF-PM essentially assumes persistence of the overall cloud effects on solar irradiances.

The R-based forecast model (hereafter R-PM) takes a further step to apply the ratio of $RCRF_{GHI}$ and $RCRF_{DNI}$ to the forecast as defined in Eq. (7), assuming the persistence of *R* such that

$$F^{dn}_{all,GHI}(t_f) = \left[1 - R(t) \times RCRF_{DNI}(t_f)^*\right] \times F^{dn}_{clr,GHI}(t_f), \tag{11a}$$

$$F^{dn}_{all,DNI}(t_f) = \left[1 - RCRF_{GHI}(t_f)^*/R(t)\right] \times F^{dn}_{clr,DNI}(t_f), \tag{11b}$$

where the variable with a superscript * in Eqs. (11a) and (11b) is estimated with

$$V(t_f) = \frac{\sum_{j=0}^{t}(1-a)^j V(t-j)}{\sum_{j=0}^{t}(1-a)^j}, \tag{12}$$

where *V* denotes the variable to be estimated; *j* denotes the time step in the past 5 time steps; *a* is a smoothing parameter set as 1/3 for the exponential weighted moving average over 5-time steps.

The 4th level forecast system has two models. One assumes the persistence of cloud albedo (cloud albedo-based forecast model and denoted by CA-PM hereafter), and the forecast equation is given by

$$F^{dn}_{all,GHI}(t_f) = [1 - \alpha_r(t) \times f(t_f)^*] \times F^{dn}_{clr,GHI}(t_f) \tag{13a}$$

$$F^{dn}_{all,DNI}(t_f) = [1 - f(t_f)^* + f(t_f)^* \times \exp[-\tau(t_f)/\mu_0(t_f)] \times F^{dn}_{clr,DNI}(t_f) \tag{13b}$$

$$\tau(t_f) = \frac{2\alpha_r(t)\mu_0(t_f)}{(1-\alpha_r(t)) \times (1-0.86)} \tag{13c}$$

where the variable marked with a superscript * also can be approximated by Eq. (12).

The other 4th level forecast model assumes the persistence of cloud fraction (cloud fraction-based forecast system and denoted by CF-PM hereafter), and the forecast equation is given by

$$F^{dn}_{all,GHI}(t_f) = \left[1 - \alpha_r(t_f)^* \times f(t)\right] \times F^{dn}_{clr,GHI}(t_f), \tag{14a}$$

$$F^{dn}_{all,DNI}(t_f) = [1 - f(t) + f(t) \times \exp[(-\tau(t_f)/\mu_0(t_f)] \times F^{dn}_{clr,DNI}(t_f), \tag{14b}$$

$$\tau(t_f) = \frac{2\alpha_r(t_f)^*\mu_0(t_f)}{(1-\alpha_r(t_f)^*) \times (1-0.86)} \tag{14c}$$

Note that the prediction for DNI and DHI of all the new forecast systems are derived by

$$F_{all,DNI} = F^{dn}_{DNI}/\mu_0, \tag{15a}$$

$$F^{dn}_{all,DHI}(t_f) = F^{dn}_{all,GHI}(t_f) - F^{dn}_{all,DNI}(t_f), \tag{15b}$$

where the cosine of the solar zenith angle μ_0 can be calculated with the method reported by Reda and Andreas (2004); the solar zenith angle from the dataset is used in this study.

3. Forecast and evaluation

3.1. Measurement data

We have tested the hierarchy of new forecast systems with the decade-long (from 1998 to 2014) and quality-controlled measurements of GHI, DNI, and DHI with the 15-minute resolution at the U. S. DOE ARM SGP Central Facility site. The site is also chosen for its availability of co-located measurements of detailed data on atmospheric processes relevant to clouds and solar radiation (www.arm.gov). The clear-sky components are assessed based on the method proposed by Long and Ackerman (2000) and Long and Gaustad (2004).

3.2. Overall performance

For convenience of discussion, various quantities assumed to be persistent (e.g., *RCRF*s, *R*, cloud albedo, and cloud fraction) are generically referred to as cloud predictors. A common criterion 1 $W\cdot m^{-2} < F_{all}(t_f) < 1361$ $W\cdot m^{-2}$ (mean solar constant) (Kopp and Lean, 2011; Kopp, G., 2014) is used to determine the valid data for the models. As indicated in Table 2, the maximum percentages of the outliers at all the lead times examined, are less than 0.47%, 8.8% and 5.2% for GHI, DNI and DHI, respectively. These results suggest a minimal influence of the excluded outliers on the evaluation and subsequent analysis. The higher percentage for DNI and R-PM may result from the operation of division and approximations involved (Eqs. (11)–(14)). To further assure the consistency of comparison across the different models and radiative components, the same set of data are used for performance evaluation that produce valid predictions of GHI, DNI and DHI in all the models.

Fig. 1 compares the overall performance of different forecast systems in terms of percentage error (PE) defined as the root mean squared error normalized by the mean measurement at lead times from 15 min up to 6 h. A few points are evident. First, for GHI, although the simple persistence model has a comparable PE at lead times less than 1 h, all the physics-informed persistence models outperform the simple persistence model at all lead times, with significant improvements especially over longer forecasting horizons beyond 1 h. These results are consistent with the previous studies on the GHI forecast (Martín et al., 2010; Kleissl, 2013). Second, all the other physics-informed forecast systems have smaller PEs than the smart persistence model (note the equivalence of RCRF-PM for forecasting GHI with the smart persistence model) beyond 1.25 h, indicating better accuracies of the new forecast systems in forecasting GHI. The performances of the fourth level models, CF-PM and CA-PM, are even better than that of the third level R-PM in forecasting GHI despite the minor differences among those three models. Third, the physics-informed persistence models successfully partition GHI into DNI and DHI with forecasting accuracies comparable to the conventional smart persistence model for GHI. Generally, CF-PM has the best performance in forecasting GHI, but CA-PM outperforms the other models in predicting DNI and DHI at lead times longer than 1 h. Finally, the forecast accuracy for GHI is generally better than those for DNI and DHI at the same lead time.

To better illustrate the quantitative improvement over the simple and smart persistence model, we calculate the PE skill score (*S*) for all the models with lead times from 15 min to 6 h. The PE skill score is defined as (Yang et al., 2020)

$$s_{reference} = \left(1 - \frac{PE_{model}}{PE_{reference}}\right)*100\%, \tag{16}$$

where the subscript "model" denotes one of the new models, and the subscript "reference" indicates that the reference persistence model is the simple (S_{simple}) or smart persistence model (S_{RCRF}). Note that RCRF-PM is equivalent to the smart persistence model for GHI, and can be regarded as the extension of the smart persistence model for DNI and DHI. Positive and negative values of *S* indicate that the new model improves and degrades the forecast relative to the reference persistence model, respectively. The magnitude of the skill score quantifies the degree of forecast improvement or degradation as compared to the reference persistence model. Fig. 2 shows the PE skill scores as a function of lead time for different models in forecasting GHI (a,d), DNI (b,e) and DHI (c, f) relative to the simple and smart (RCRF) persistence model. Several points are noteworthy. First, for GHI, all the new models outperform the simple and smart persistence model at all the lead times examined (from 15 min to 6 h), and the improvement increases with increasing lead times. CF-PM has the best performance with improvement up to 68% and 29% at 6-hour lead time compared to the simple and smart persistence model, respectively. It is noteworthy that the difference between CA-PM and CF-PM is negligible, and both models have slightly better performance than the third-level model, R-PM. The results clearly demonstrate the enhancement of the skill scores in forecasting GHI resulting from incorporating physics into the forecast models at different levels of the hierarchy, and the model at the higher level tends to have better performance. Second, for DNI, the improvements of the third and fourth level models increase with increasing lead times similar to GHI, and they all outperform the simple persistence model and RCRF-PM from 0.75-hour and 1.25-hour lead time, respectively. CA-PM performs the best, with improvements up to 47% and 15% relative to the simple model and RCRF-PM. It is noted that the performance of R-PM is worse than the simple persistence model and RCRF-PM in forecasting DNI ($s < 0$) at lead times less than 0.75 and 1.25 h, respectively. Third, for DHI, the improvement starts from 1-hour lead time for all but R-PM. Despite the general increase of the skill score with increasing lead times as well, the improvements are less than 30% for DHI based on S_{simple}, smaller than those for GHI and DNI forecasts, which is likely related to the error propagation/enhancement, since it is essentially calculated as a difference between forecasted GHI and DNI. R-PM starts to forecast DHI better than the simple persistence model at a later lead time of 3 h, and it is postponed to 4 h regarding S_{RCRF}. The relatively poor performance of R-PM for DNI and DHI might be related to the above-mentioned approximations involved in deriving the *R* ratio.

Another metric to gauge forecast models lies in the improvements in the forecast lead time given a reference forecast accuracy. It is recognized that the simple persistence models generally have acceptable forecasting accuracy within the 1-hour forecasting horizon (Diagne et al., 2013), which is also shown in Fig. 1. Thus, we choose the PEs of simple persistence models at the lead time of 1 h as the reference to identify the lead time of the other models with similar PEs and quantify the improvement in forecast lead time over the simple persistence model. The results are summarized in Table 3. Relative to the simple persistence model, RCRF-PM, R-PM, CA-PM and CF-PM extend the lead times from 3 h up to 6 h (longest lead time examined) for GHI; the corresponding extensions for DNI are 1.75 h, 2 h, 2.5 h, and 2 h, respectively. The improvements in lead time for DHI are not as significant as the other two components with all being smaller than 1.25 h. Similarly, Table 4 lists the improvements in forecast lead times of the third and fourth level models over the smart (RCRF) persistence model.

To further examine the error components determining PE, we modify the Taylor diagram (Taylor, 2001) to show the individual error components (correlation coefficient *r* between the measurement and forecast, normalized standard deviation $\sigma = \sigma_{pre}/\sigma_{obs}$ (σ_{pre} and σ_{obs} represent the standard deviation of model and measurement, respectively), and the centered root-mean-squared error (CRMSE) ($1 + \sigma^2 - 2\sigma^* r$) at all lead times (Fig. 3). The shorter the distance between each point and the reference point 'Ref', the smaller the CRMSE. The modification lies in the additional representation of the magnitude of mean bias between the measured and predicted solar irradiances by different colors and lead

Table 2
Maximum percentage (%) of the outliers for individual models.

Model	GHI	DNI	DHI
Simple	0.0029	2.4	0.0012
RCRF-PM (smart)	0.019	2.2	0.048
R-PM	0.47	8.8	5.2
CA-PM	0.055	0.62	0.14
CF-PM	0.055	2.7	0.11

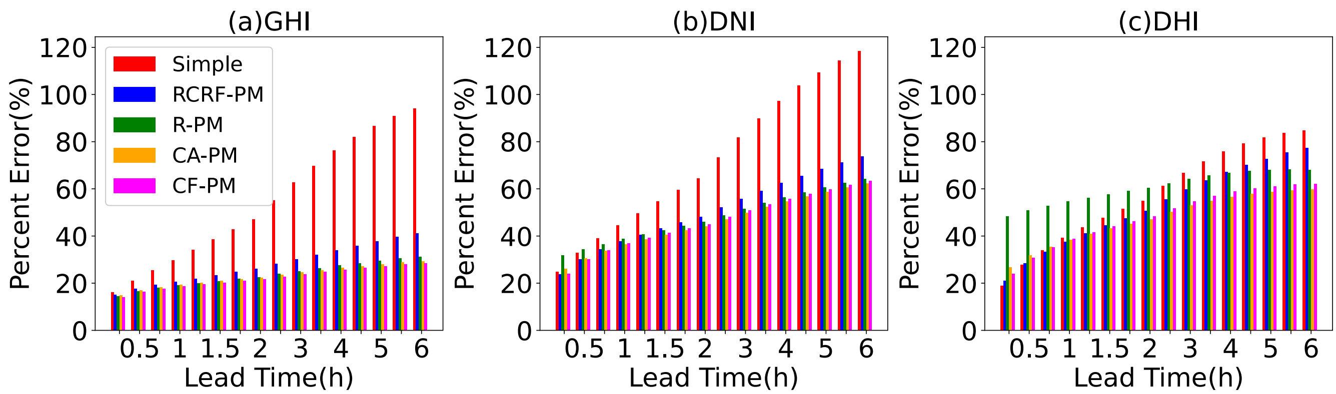

Fig. 1. Overall performance of all forecast systems as a function of forecast lead time.

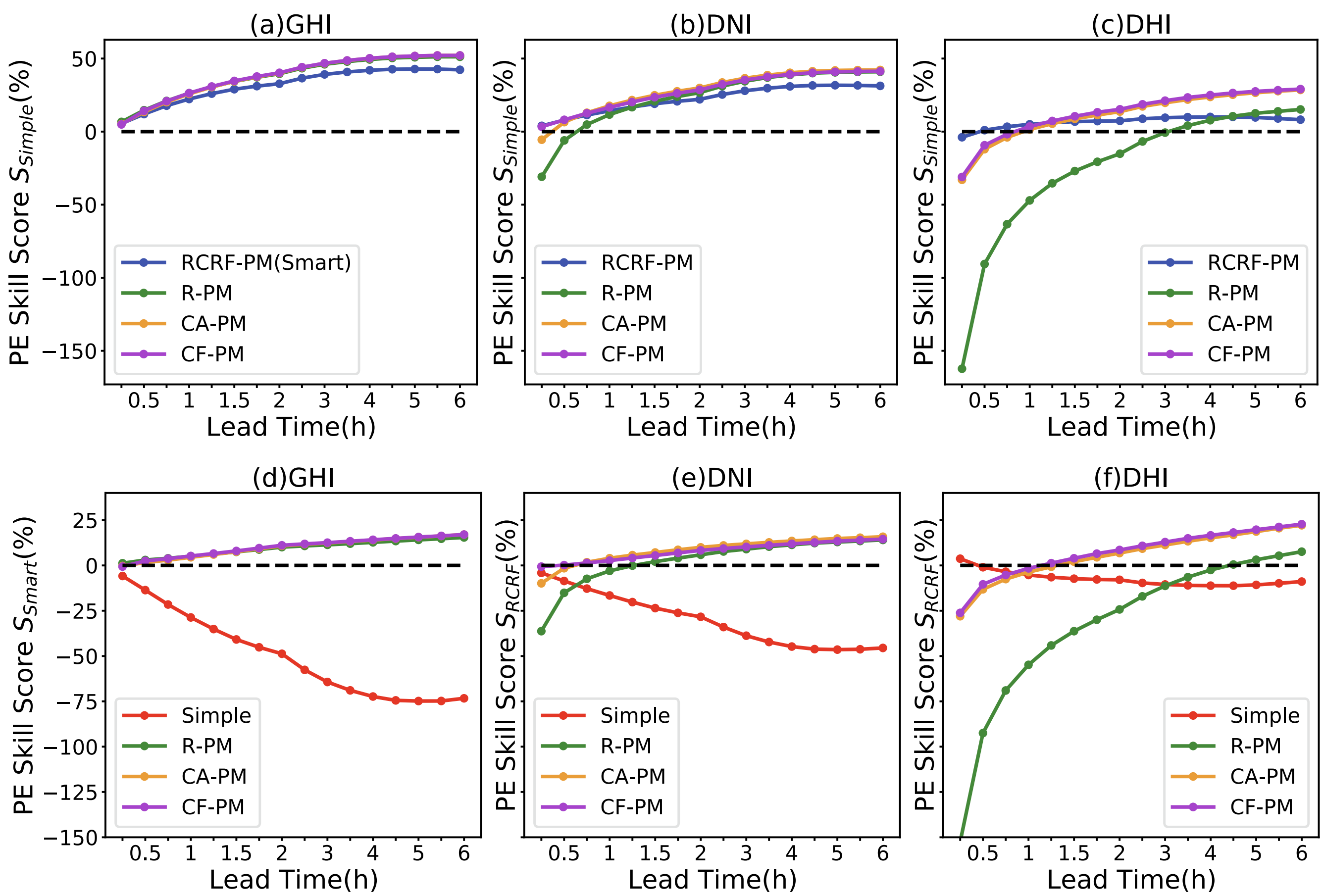

Fig. 2. PE skill score as a function of lead time for different models. S_{simple} and S_{RCRF} mean the PE skill scores are referenced by the simple and smart (and RCRF) persistence models, respectively.

Table 3
Improvement in forecast lead times of the new models over simple persistence model.

	RCRF-PM	R-PM	CA-PM	CF-PM
GHI	3 h	5.5 h	6 h	6 h
DNI	1.75 h	2 h	2.5 h	2 h
DHI	1.25 h	0 h	1.25 h	1.25 h

Table 4
Improvement in forecast lead times of the new models over smart (RCRF) persistence model.

	R-PM	CA-PM	CF-PM
GHI	1.5 h	1.5 h	1.75 h
DNI	1 h	1.25 h	1.25 h
DHI	1 h	1 h	1 h

times by different symbol sizes. With all the error components of PE represented, the modified Taylor diagram helps to determine the individual factors resulting in the differences in the performances among various models. Some features are noteworthy. First, model performance gradually degrades with an increasing CRMSE as the lead time increases. With the exception of RCRF-PM, the increase of CRMSE is

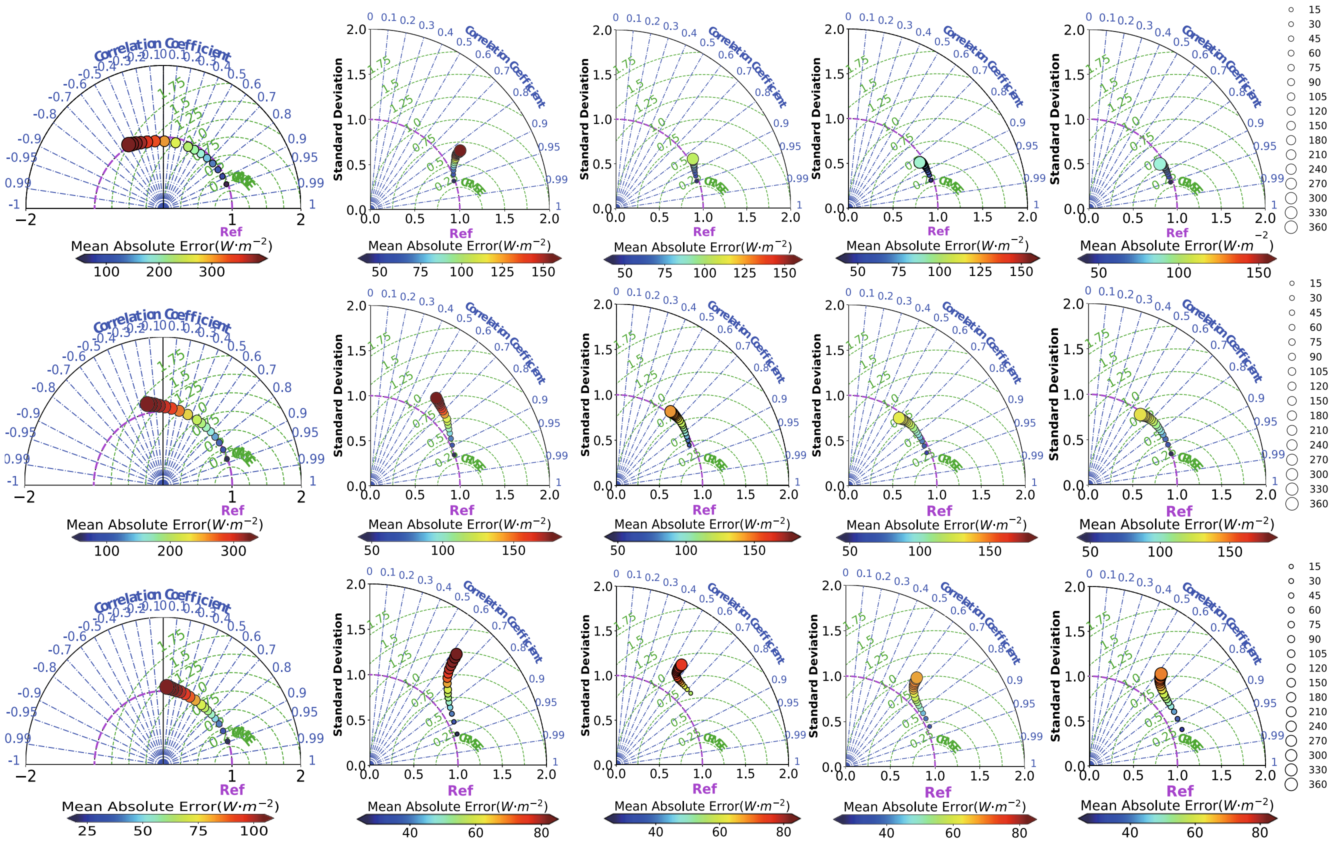

Fig. 3. Taylor diagram illustrating the model performances at all lead times. The top, middle, and bottom rows are for GHI, DNI, and DHI; different columns are for different models; different symbol sizes and colors are for various lead times and the mean absolute errors between the measured and forecasted solar irradiances, respectively. The radial distance from the origin is proportional to the normalized standard deviation. The centered RMSE (green dotted lines) between the forecast and observation field is proportional to their distance apart. The blue dotted lines indicate the correlation coefficient between the forecast and reference field. Note the error bar for the simple persistence model is not the same as the other models. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

mainly caused by the decrease of the correlation coefficient, since the variation of the normalized standard deviation with lead time is minimal. Similarly, the smaller CRMSE for GHI compared to DNI and DHI results mainly from the higher correlation as well. As for RCRF-PM, a distinction from the other models is embodied in the increasing normalized standard deviation with lead time, which is responsible for the growing CRMSE as well. Second, at the lead times with positive S_{simple} as shown in Fig. 2, the physics-informed forecast systems outperform the simple persistence models mainly due to smaller mean biases and higher correlation. At the lead times with positive S_{smart}, the physics-informed forecast systems have better performance mainly resulting from smaller mean biases and smaller standard deviation rather than higher correlation; the correlation coefficients are similar for RCRF-PM and the other physics-informed models. Nevertheless, at the lead times with negative S, relative to the simple or smart persistence model, the worse performance of the physics-informed models stems from a lower correlation and/or a higher mean bias. For example, R-PM has negative S_{RCRF} at the lead times less than 1.25 h and 4 h for DNI and DHI forecasts, respectively, and the Taylor diagram suggests that the lower correlation coefficient and larger mean bias are responsible for the poor performance in R-PM at those lead times. Third, the fourth level-based models have the best performance in forecasting irradiances due to their improvements in forecasting correlation, mean bias as well as low standard deviation.

In addition to PE (or PE skill score) and its various error components, Yang et al. (2020) recently recommended the use of the Murphy–Winkler framework to conduct distribution-oriented forecast verification whereby the marginal occurrence probability distributions of observations and forecasts can be compared, and more detailed joint distribution can be evaluated. Fig. 4 further shows an example of comparing different forecasts at the lead time of 3 h against the corresponding measurements. As can be seen that the scatters of the simple persistence model are much broader than the other models for all the solar radiative components. In general, the fourth-level based models, CA-PM and CF-PM, outperform the other models with the data points centering around the perfect one-to-one line. Without referencing the modified Taylor diagram or PE, it is difficult to determine which model is the best due to the visually similar scatters between these two. The second and third level-based models, RCRF-PM and R-PM, tend to overestimate the high-value irradiances, with the high occurrence frequency zone shifting upward from the one-to-one line. Moreover, the histograms display the marginal distributions of measurement (on the top) and forecast (on the right). For clarity, the forecasted and measured marginal cumulative distributions are further compared in Fig. 5 for GHI (a), DNI (b), and DHI (c). Visually, the new persistence models can well capture the observed GHI distribution. A larger discrepancy between the forecasted and observed marginal distributions exists under low DNI and DHI conditions ($<$100 W·m^{-2}) for most models, suggesting the limitations of low irradiance conditions and plausible stronger effects of clouds (see Section 4 for more detailed analyses). The relatively larger discrepancy in DNI and DHI between the forecast and observed marginal distributions also supports the preceding result that the improvement in forecasting DNI and DHI stems mainly from correlation (or joint PDF). Note that the discussion above is based on the percent error gauging the overall model performance; careful inspection of Figs. 4 and 5 suggests that CF-PM better captures the events of first DNI bins (DNI $<$ 20

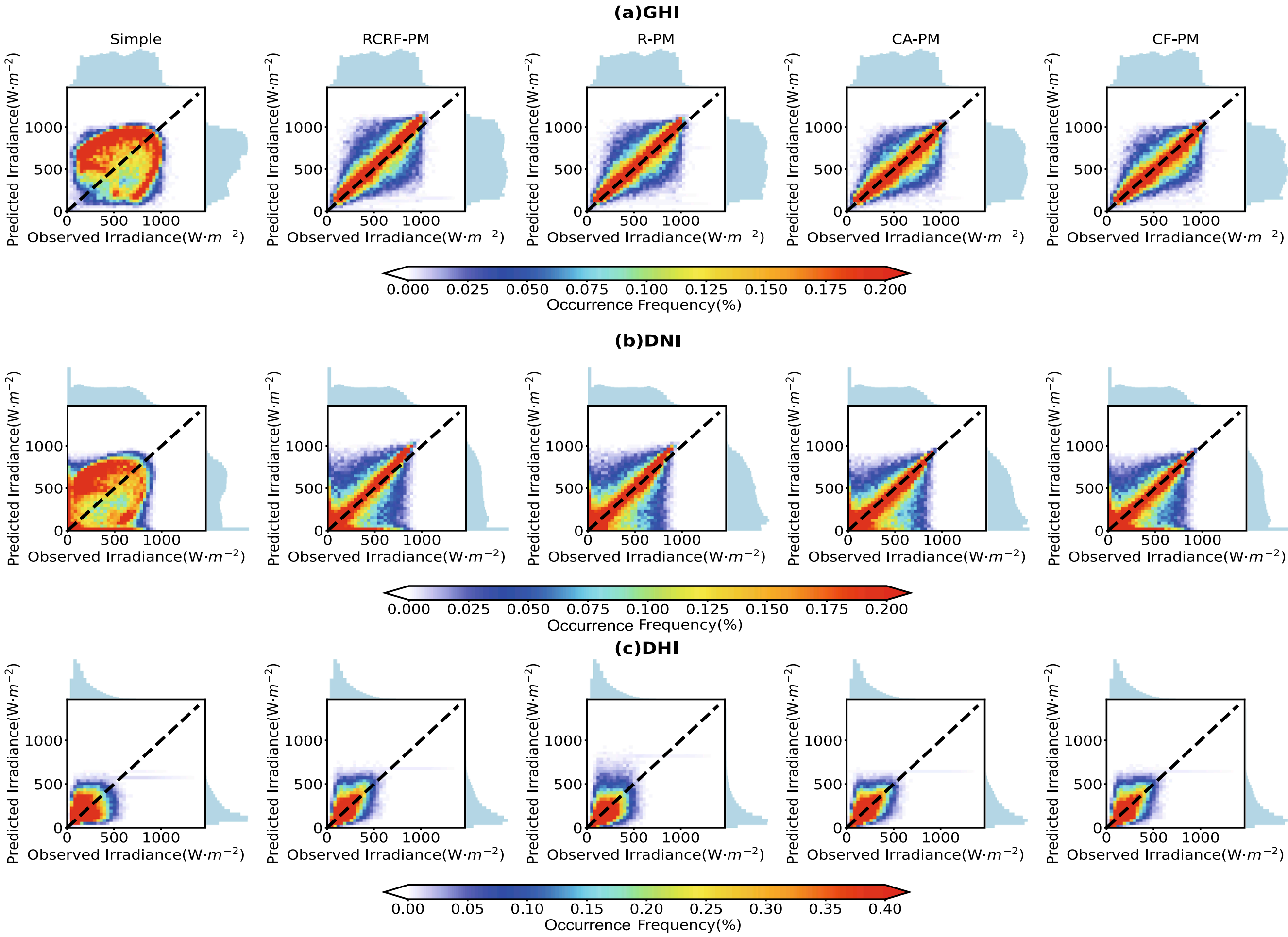

Fig. 4. Joint and marginal distributions of the predicted and observed solar irradiances at the lead time of 3 h. Different columns are for different models as marked on the center top of the plots in the top row. The top, middle and bottom rows are for GHI, DNI, and DHI, respectively; the plots in the same row share the same colorbar displayed at the end; the black dashed line is the one-to-one line; the top and right marginal distributions are for observations and forecasts, respectively.

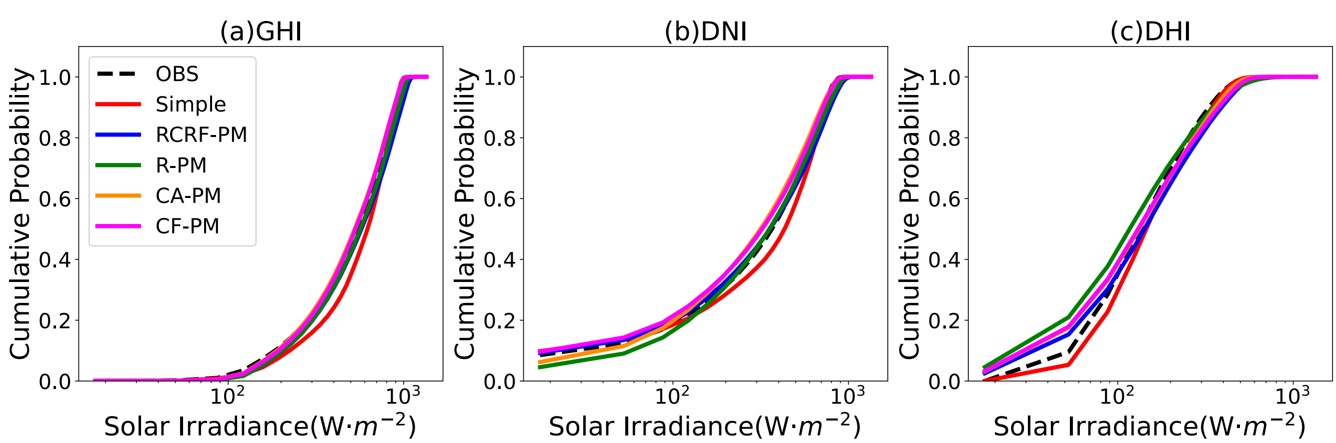

Fig. 5. Comparison between the observed marginal cumulative distributions and those forecasted by the different models. The radiative components are given on top of plots.

W·m^{-2}).

Together, the analyses demonstrate that the new physics-informed forecast systems are capable of predicting GHI, DNI, and DHI and outperform the simple and smart persistence model, especially over longer time horizons (>1.25 h) by reducing measurements of GHI and DNI into dimensionless cloud predictors that are related to different cloud properties. Furthermore, regardless of the detailed differences for various solar irradiances, the fourth-level models (CF-PM and CA-PM) perform the best. This makes physical sense considering that the third-level R-model framework is an approximation of the fourth-level framework, the second-level RCRF system considers the overall cloud radiative effects without separation, and the first level models consider solar radiation directly without explicitly considering any cloud effects.

4. Further analysis

As qualitatively illustrated in Fig. 6, large forecast errors are often related to the cloudy episodes with a dramatic temporal variation of cloud properties, suggesting the possible linkage between the variability of cloud property and model performance. This section further explores the quantitative relationships between them.

To quantify the temporal variation, we introduce the magnitude of the relative variability (ε) defined as

$$\varepsilon = \left|\frac{V(t+\Delta t) - V(t)}{V(t)}\right| \times 100\%, \tag{17}$$

where V denotes the variable in question, t is the time, and Δt is the time difference between the steps (hereafter referred to as interval time for convenience). The use of the relative variability permits the comparison of quantities with different physical units; taking the absolute value allows us to focus on the magnitude of the variation to better show the trend of the statistical mean ε with interval time without being influenced by the sign. Fig. 7 shows an example of the temporal variations of ε for solar irradiances and cloud properties at the interval time of 3 h in

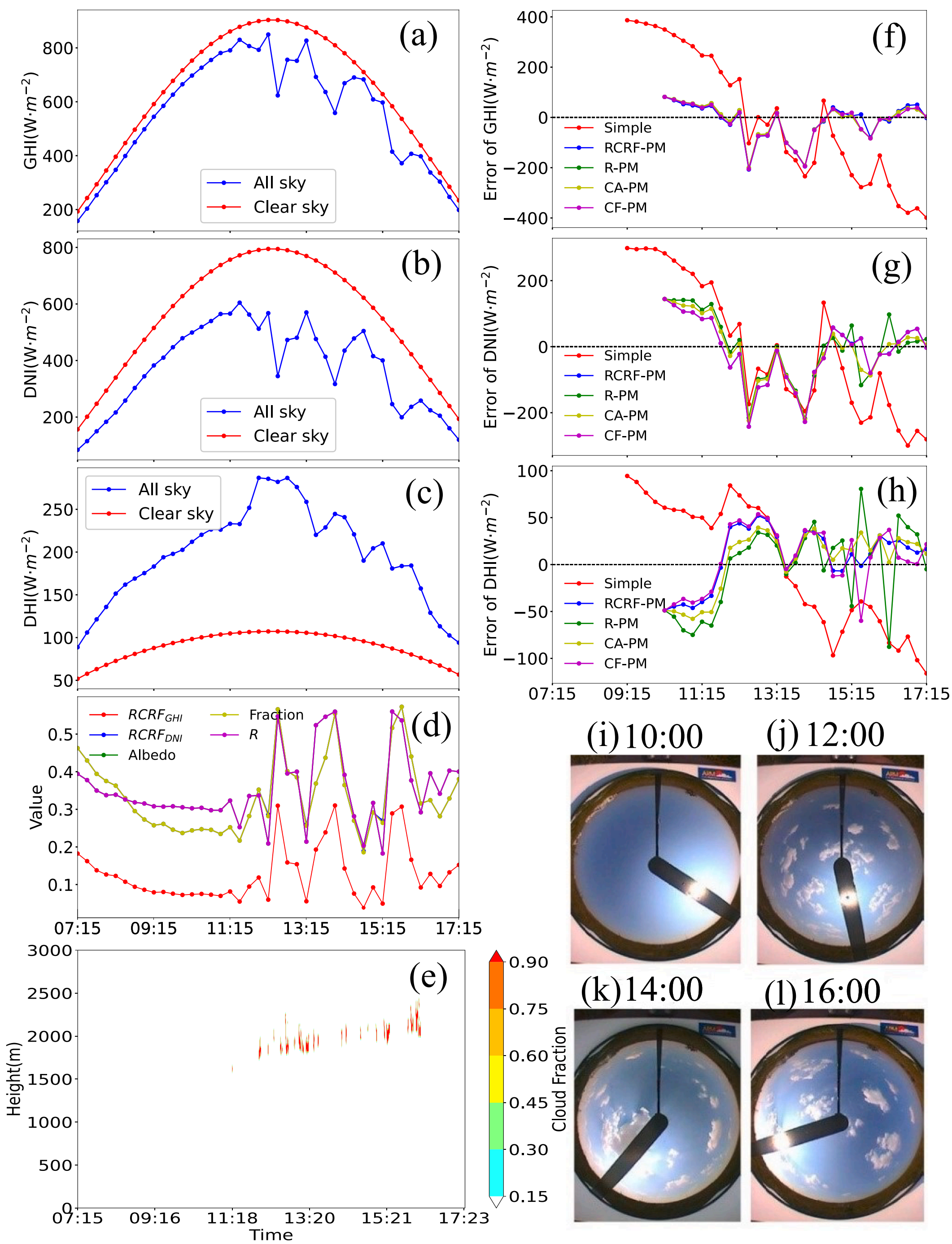

Fig. 6. A case on 09/07/2002; (a) ~ (e) show the temporal variation of GHI, DNI, DHI, cloud properties, and ground-based-observed cloud fraction; (f) ~ (h) are the forecasted errors from all the systems for GHI, DNI, and DHI at the lead time of 2 h. Four sky images from the top left to the bottom right correspond to the local time 10:00, 12:00, 14:00, and 16:00, respectively.

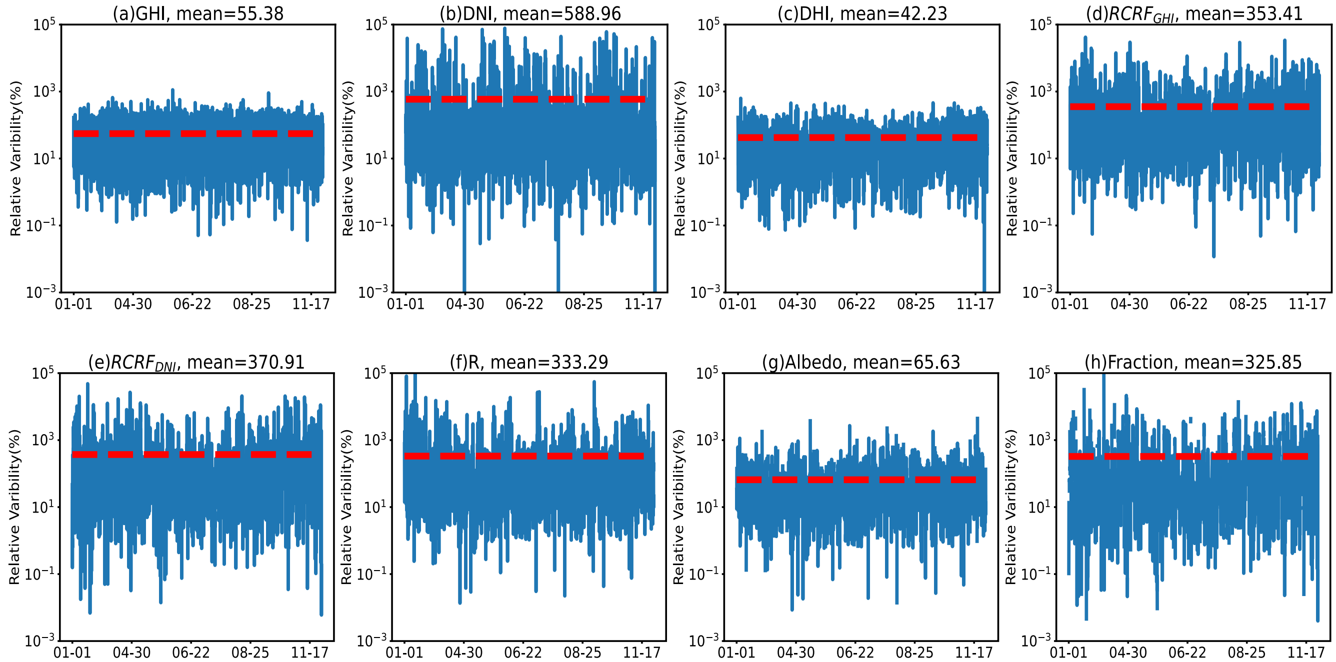

Fig. 7. Temporal variation of the magnitude of relative variability ε for radiative components and cloud predictors at the interval time of 3 h. The red line represents the mean ε. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2014. The mean ε for GHI is much smaller than that of DNI, indicating a gentle variation in GHI from t to $t + \Delta t$, which is potentially associated with the best performance among three radiative components by using one type model to forecast GHI (Figs. 1–3). Even though the smallest ε in DHI, it is derived in an indirect way with the forecast accuracy being both affected by the signature of GHI and DNI, leading to worse performance in DHI compared to GHI. Furthermore, the fourth-level-properties (cloud fraction and cloud albedo) have smaller mean εthan the other level ones (*R*, *RCRFs*), supporting that the fourth-level-model systems outperform the other models in predicting individual radiative components. To illustrate the connections between temporal variability and model performance extended to all interval times examined and the long-term periods, Fig. 8 relates the mean ε (from 1998 ~ 2014) of the cloud predictor with the PE from the corresponding predictor-based model when the lead time is equal to the interval time, showing that PE positively correlates with the mean ε for all the models, and both generally increase with lead time. In general, the smaller mean ε associated with better performance in GHI and the fourth-level-based systems agrees well with Fig. 7; however, the minimum mean ε in cloud albedo only contributes the least to PE for DNI and DHI, but not for GHI, and meantime model performances are relatively independent of the mean ε among R-PM, CF-PM and CA-PM, since the discrepancy between the mean ε at the same lead time does not affect the PE significantly, both of which possibly imply other factors influencing the model performance in addition to the temporal variability. Moreover, though a smaller ε in GHI than the other cloud predictors, the worst performance of the simple model could be possibly explained as by the physical understanding that this predictor does not embody a specific type of cloud influence, but the irradiance itself. These findings highlight the linkages among solar radiation changes, temporal variability of cloud properties, and model performances, and also suggest the minimal ε principle in general for choosing the cloud predictor among different hierarchy levels to build the corresponding forecast system.

The increasing mean ε with the increasing interval time essentially indicates the growing forecast errors with lead time, which makes the assumption of the persistence of the key predictors used in the model is more questionable for long lead times. Based on the assumptions, a larger error between the actual and the assumed persistent cloud predictor (*RCRF*, *R*, cloud albedo, and cloud fraction) is likely translated into a larger error in the solar irradiance forecast (see the corresponding forecast equations in Section 2). To confirm this, Fig. 9 shows the relationship of the errors between cloud predictors and GHI (top panel), DNI (middle panel), and DHI (low panel) at the lead time of 2 h. Also shown are the PEs of the models. Several points are noteworthy. First, with

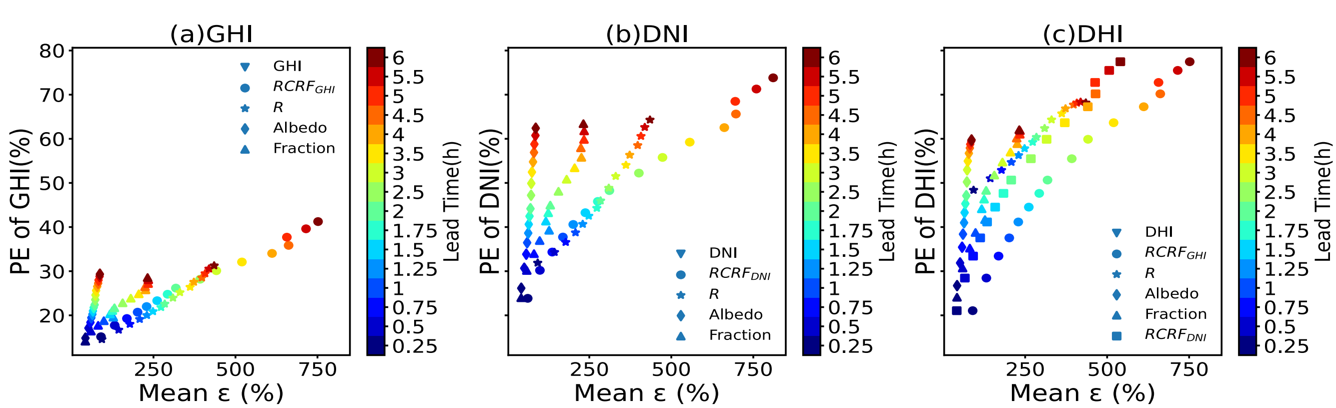

Fig. 8. Percent error (PE) as a function of the mean ε for the cloud predictors at all the lead times.

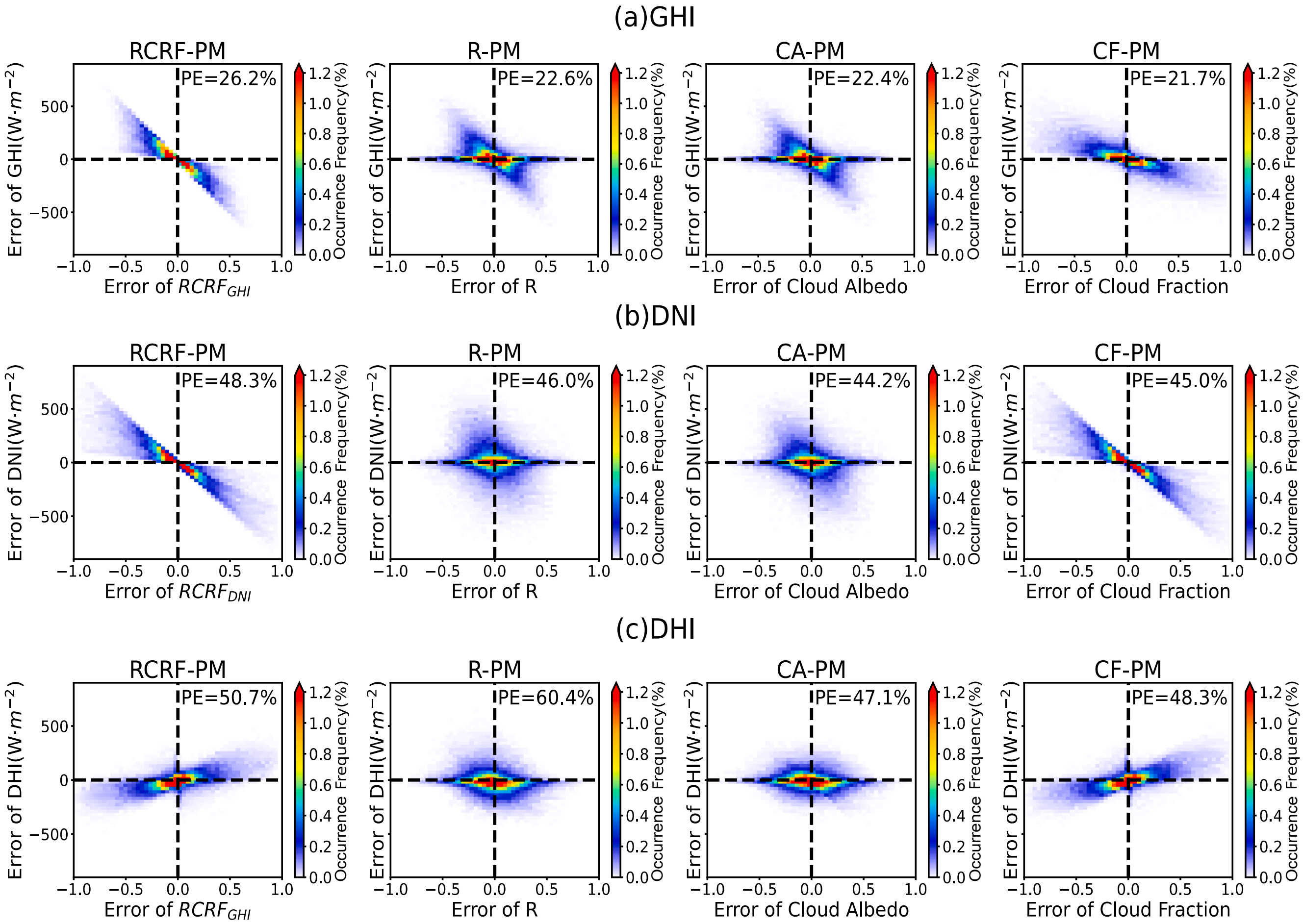

Fig. 9. Relationships between the errors in cloud predictors and solar irradiances with the corresponding PEs at the lead time of 2 h. The top, middle, and bottom rows are for GHI, DNI, and DHI, respectively; different columns represent different models as marked on the top in the first row.

three exceptions. All the error correlations are negative for GHI and DNI, suggesting that underestimated (overestimated) cloud properties (e..g, cloud fraction and cloud albedo) lead to overestimated (underestimated) solar irradiances. One exception is the relationship between the error in cloud fraction (and RCRF$_{DNI}$, RCRF$_{DNI\approx f}$) and the error in forecasted DHI, which is positive, suggesting that an underestimated (overestimated) cloud fraction (and RCRF$_{DNI}$) leads to underestimated (overestimated) DHI. These results seem consistent with the physical understanding that clouds normally enhance DHI, but reduce DNI and GHI. The other two exceptions are the relationships of the error in the forecasted DNI to the errors in *R* and cloud albedo, both of which are largely independent. This error independence seems to accord with the physical understanding that DNI is determined primarily by cloud fraction (Yang et al., 2012; Xie and Liu, 2013). Second, the details of the error relationships vary among different cloud predictors and radiative components. For example, on one hand, for individual cloud predictor, the error contributes a wider variation range to the error in DNI compared to GHI; on the other hand, for the same radiative component, the error in cloud albedo contributes the least to the error in DNI as well as in DHI with the smallest PE compared to the other three cloud predictors (*RCRF*, *R* and cloud fraction) but not shown in the GHI forecast. All these results indicate the different sensitivities of the radiation error to the cloud predictor error.

The sensitivity of the radiation error to the cloud preditor error basically refers to the propagation strength from the predictor error to the solar irradiance error. As can be seen from Fig. 9, cloud albedo and R have relatively weaker propagation strength due to the visually independent relationship between the errors in cloud predictor and radiation. To illustrate the point, Fig. 10 (a) ~ (c) show the slope (m) of the linear-fitting equation between the error in the cloud predictor (x) and the error in the solar irradiance (y) with the expression of $y = mx$, m is used as an indicator of propagation strength; the intercept is set as zero to eliminate its influence on the results when comparing different m at various lead times. Fig. 10 (e) ~ (f) show the correlation coefficient (r) between x and y. The negative m and r for all the physical models in GHI and DNI forecasts further confirm that the error in solar radiation negatively correlates with the error in cloud predictor at all the lead times as supplementary to Fig. 9. Meanwhile, regardless of the details on the variation in m and r with lead time, a consistent ranking order between them among different cloud predictors also indicates that the magnitude of m can roughly represent the correlation r between predictor error (x) and solar radiation error (y), and vice versa. The magnitudes in m and r vary among cloud predictors, manifesting their different propagation strengths from predictor error to the solar irradiance error and thus regulate the model performance by its integration with the error in cloud predictor. For example, the relatively larger error in *RCRF*s (Fig. 8) and the stronger propagation strength both contribute to a large radiation error (y) degrading performance with the largest PE in RCRF-PM except for the simple model. Conversely, a plausible reason for the best performance of CA-PM in DNI and DHI forecasts is due to the smallest error (Fig. 8) in cloud albedo as well as the relatively weaker propagation strength (Fig. 10). Moreover, the combined effects of the propagation strength (or correlation) and the error in cloud predictor are possibly responsible for the independent relationship of PE with the mean ε among R-PM, CA-PM and CF-PM as shown in Fig. 8, the difference in the ranking order between the mean ε and correlation among R,

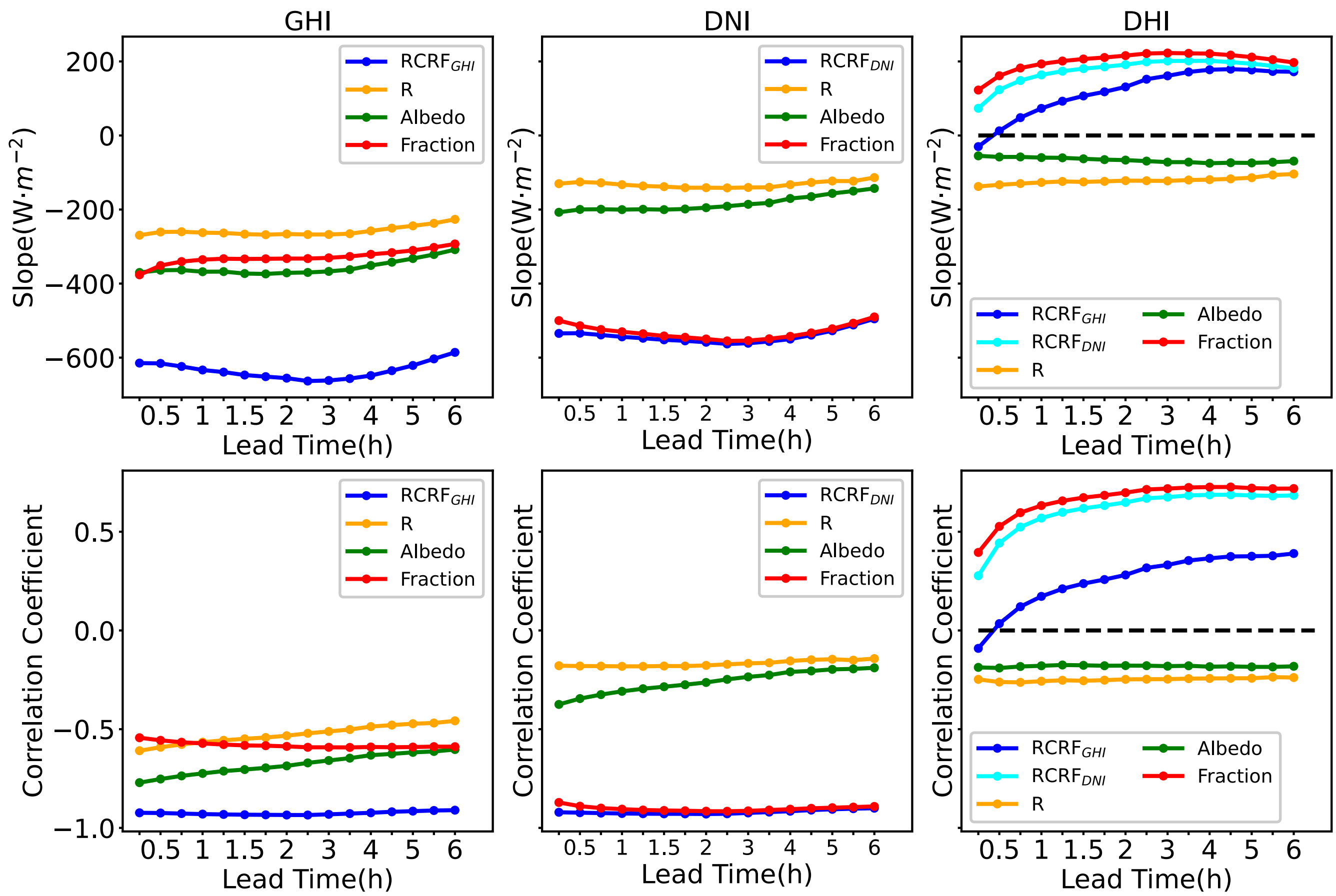

Fig. 10. (a) ~ (c) shows the slope (m) of the linear regression equation, $y = mx$, as a function of lead time, where y and x, respectively, denote the error of the solar irradiance and cloud predictor for each model. (d) ~ (f) show the correlation coefficient (r) between x and y as a function of lead time.

cloud albedo and cloud fraction makes it possible to result in similar model performance.

The preceding analysis discussed the relationship between the error in the key cloud predictor and the model performance without differentiating the impact of the potential error from the other assumed variables introduced by applying the exponential technique (i.e., assumed *RCRFs* in R-PM, assumed cloud fraction in CA-PM, and assumed cloud albedo in CF-AM). The joint impact of the errors from both variables on the solar radiation error is further investigated given by Fig. 11. Note that the x-axis denotes R or cloud albedo due to the relationship of $R \approx \alpha_r$ for better comparison regardless of the assumption technique (persistence or exponential moving average) applied to cloud predictors. Different patterns are shown for various radiative components. In GHI forecast, large radiation errors mainly occur in the first and third quadrant with either negative or positive error in two cloud predictors. The underestimated (overestimated) solar radiations mainly occur when the error in cloud fraction or $RCRF_{GHI}$ is greater (less) than 0, but the variation in the sign of the cloud albedo error does not cause a significant variation in the DNI error. In DHI forecast, the overestimation and underestimation in DHI are separated by a split line across the first and third quadrant with an angle to the x-axis around 45°, the methodology in deriving DHI makes the relationship complicated and not straightforward. That DNI error mainly determined by the cloud fraction error is consistent with the results shown by Figs. 9 and 10, and agrees with Yang et al. (2012).

To further investigate the physical reasons underlying the model behaviors, Fig. 12 shows the relationships between solar irradiances and cloud predictors based on the all measurements examined. Both GHI (top panel) and DNI (middle panel) irradiances decrease with increasing values of the cloud predictors. DHI (lower panel) exhibits more complicated relationships; it increases with cloud fraction, but it increases with $RCRF_{GHI}$, R and cloud albedo only when their values are less than 0.4, and then DHI decreases with them as they further increase. These results seem consistent with the physical understanding that clouds normally enhance DHI but reduce DNI and GHI, and that also accounts for the underestimated cloud predictors causing an overestimated GHI and DNI as well as underestimated cloud fraction leading to an underestimated DHI, as shown in Fig. 9.

These findings also help explain the performance among different radiative components and different models. It is clear that the contrasting dependences of DNI and DHI on cloud fraction somewhat cancel each other as the cloud fraction varies, which only leads to a slightly decreasing trend in GHI with cloud fraction (first row, last column in Fig. 11) as GHI theoretically is the sum of DHI and the vertical component of DNI. The contrasting dependencies of DNI and DHI on the other predictors (*RCRF, R*, and cloud albedo), however, occur only when the value of the cloud predictor is less than 0.4; beyond that, obvious descending trends on GHI are shown because of the reductions both on DNI and DHI with *RCRF, R* and cloud albedo. The feature infers that GHI has a gentle variation with the increasing cloud fraction causing a smaller error in GHI when using CF-PM to make a forecast compared to the other models. Conversely, DNI is determined primarily by cloud fraction, and a small error in cloud fraction will introduce a noticeable error in DNI. However, DNI is relatively independent of cloud albedo or R, which not only supports the above-mentioned largely independent relationships of the error in forecasted DNI to the errors of R and cloud albedo, but also indicates a better performance on DNI when R-PM and CA-PM are used. Despite the similar relationships between solar radiation and cloud albedo and R, the smaller error in cloud albedo than in R at the same lead time (Fig. 9, Fig. 11) seems to account for the better

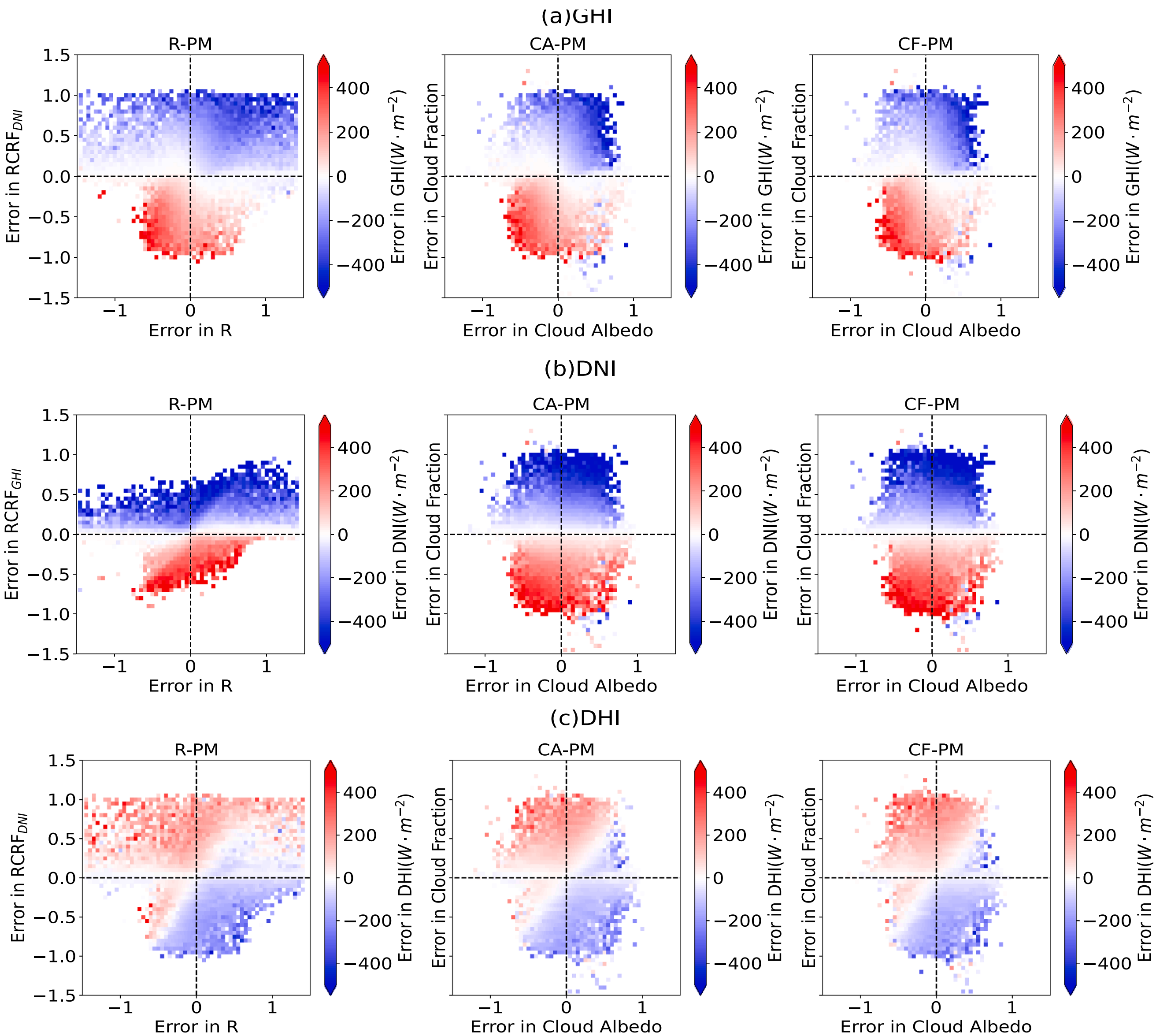

Fig. 11. Relationships between the errors in two estimated cloud predictors in R-PM (first column), CA-PM (second column) and CF-PM (third column) at lead time of 3 h. The colorbar shows the error in solar irradiance; different rows are for different radiative components.

performance in CA-PM than R-PM in forecasting DNI.

5. Concluding remarks

A hierarchy of four new physics-informed persistence models is presented to improve the ability to forecast both GHI and its partitioning into DNI and DHI by incorporating clear physics into the persistence models based on the theoretical framework that connected solar irradiances with cloud properties (Liu et al., 2011; Xie and Liu, 2013). The decade-long measurements at the ARM SGP Central Facility site are utilized to evaluate the performance of these models and compared them with the commonly used simple and smart persistence models. An in-depth analysis is conducted as well to assess the model performance and associate it with specific cloud predictors.

Our results show that the new physics-informed models outperform the simple and smart persistence models and improve the forecasting accuracy of GHI, DNI, and DHI at long lead times (>1.25 h). Generally, CF-PM demonstrates the best performance in predicting GHI, and CA-PM is the best forecast system for DNI and DHI forecast based on the overall performance evaluated by the percent error. Regardless of the detailed differences for various solar components, the fourth-level forecast systems (CA-PM and CF-PM) have the overall best performance among the models. Further analysis shows that model performance is related to the temporal variability of the cloud predictor assumed to be persistent: a larger variability between the actual and the assumed persistent cloud predictor will generally translate into a larger error in the solar irradiance forecast.

The results clearly demonstrate the importance and utility of incorporating physics into developing observation-based forecast models. A few points are noteworthy. First, this study is primarily focused on the overall cloud influences without the separation of different cloud types. In the future, it is desirable to test the forecast systems under different cloud types. Also desirable is to evaluate the models in different climate zones and locations. Second, in this study, the prediction for DHI is obtained by use of the equality DHI = GHI-DNI*μ_0, and thus the errors from GHI and DNI will both affect predicting DHI irradiance. A direct method is more desirable. Third, as the first proof of concept, this study focuses primarily on the persistence models, advanced approaches (e.g., machine learning) in time-series forecasting merits exploration in the future. Last, the theoretical framework presented in Liu et al. (2011) and

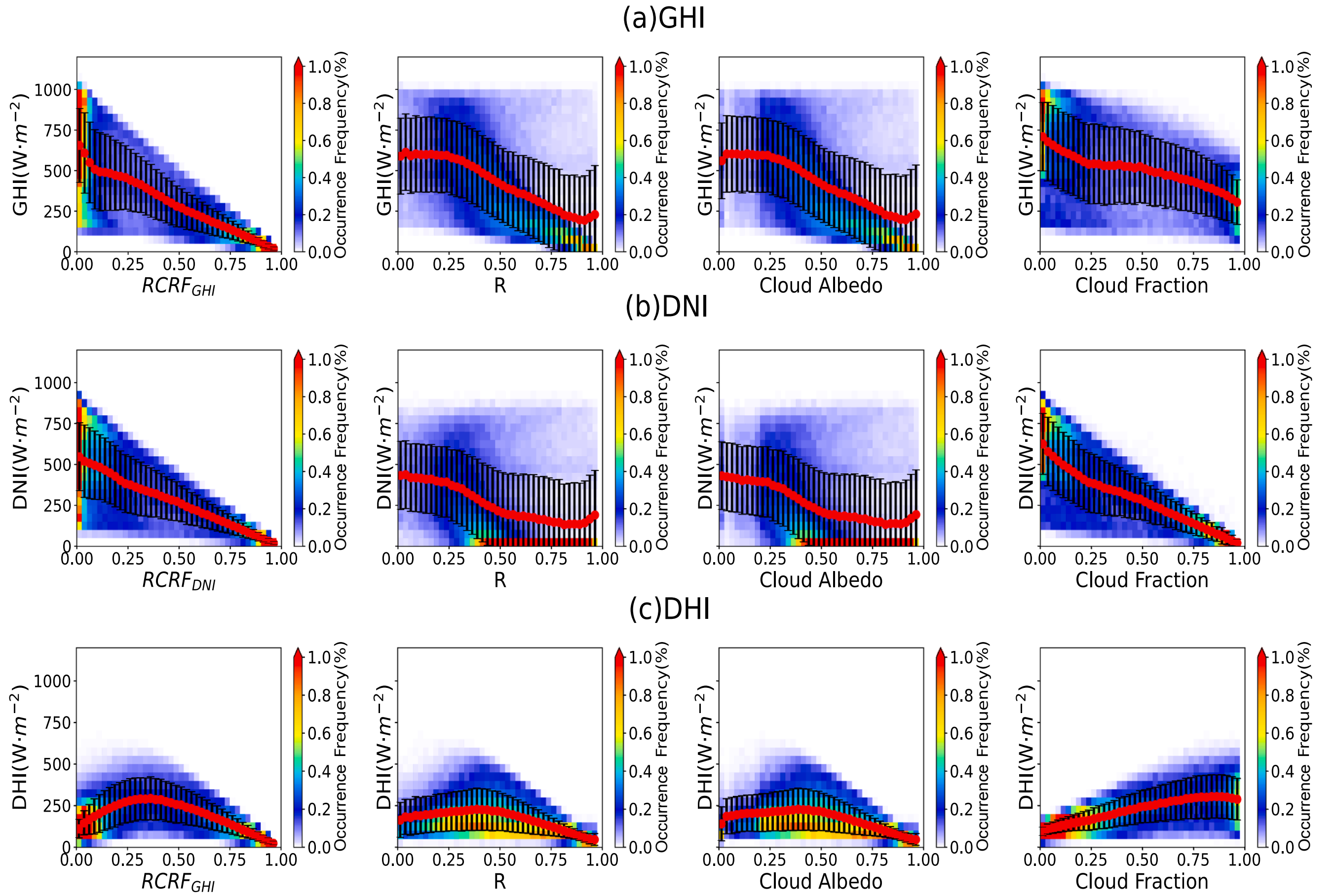

Fig. 12. Relationships between cloud predictors and solar irradiances The top, middle, and bottom rows are for GHI, DNI, and DHI, respectively; different columns show the relationships between solar irradiances with *RCRF*, *R*, cloud albedo, and cloud fraction, respectively. The red line with the black error bar indicates the mean ± standard deviation of solar irradiance. The plots are based on all the measurements examined and the color bar shows the occurrence frequency of corresponding data points. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Xie and Liu (2013) have inherent assumptions that may not hold for ambient clouds. For example, the homogeneous cloud assumption cannot accurately represent 3-D radiative transfer in some clouds (e.g., deep convective clouds). Careful use of this method in such specific scenarios is needed. The impacts of the overlapped clouds, aerosols and the 3-D cloud effect are not yet considered in our model. The model is expected to be more accurate when considering these factors, and associated works in the future are needed to be explored.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are grateful to Dr. Dazhi Yang Singapore at the Institute of Manufacturing Technology, and anonymous reviewers for their valuable comments. The data are downloaded from the DOE ARM website (www.arm.gov). The work is supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under Solar Energy Technologies Office (SETO) [grant number 33504] and Atmospheric System Research (ASR). The Brookhaven National Laboratory is operated by the Brookhaven Science Associates, LLC (BSA), for the U.S. Department of Energy under Contract No. DE-SC0012704. The National Renewable Energy Laboratory is operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy under Contract No. DE-AC36-08GO28308. The views expressed in the article do not necessarily represent the views of the Department of Energy or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

References

Bailek, N., Bouchouicha, K., Al-Mostafa, Z., El-Shimy, M., Aoun, N., Slimani, A., Al-Shehri, S., 2018. A new empirical model for forecasting the diffuse solar radiation over Sahara in the Algerian Big South. Renewable Energy 117, 530–537.
Beltran, H., Perez, E., Aparicio, N., Rodríguez, P., 2012. Daily solar energy estimation for minimizing energy storage requirements in PV power plants. IEEE Trans. Sustainable Energy 4 (2), 474–481.
Bird, R.E., Hulstrom, R.L., 1981. Simplified clear sky model for direct and diffuse insolation on horizontal surfaces (No. SERI/TR-642-761). Solar Energy Research Inst., Golden, CO (USA).
Chu, Y., Li, M., Pedro, H.T., Coimbra, C.F., 2015. Real-time prediction intervals for intra-hour DNI forecasts. Renewable Energy 83, 234–244.
Diagne, M., David, M., Lauret, P., Boland, J., Schmutz, N., 2013. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renew. Sustain. Energy Rev. 27, 65–76.
Inman, R.H., Pedro, H.T., Coimbra, C.F., 2013. Solar forecasting methods for renewable energy integration. Prog. Energy Combust. Sci. 39 (6), 535–576.
Kleissl, J., 2013. Solar energy forecasting and resource assessment. Academic Press.
Kobayashi, T., 1989. Radiative properties of finite cloud fields over a reflecting surface. J. Atmos. Sci. 46 (14), 2208–2214.
Kopp, G., 2014. An assessment of the solar irradiance record for climate studies. J. Space Weather Space Clim. 4, A14. https://doi.org/10.1051/swsc/2014012.

Kopp, G., Lean, J.L., 2011. A new, lower value of total solar irradiance: Evidence and climate significance. Geophys. Res. Lett. 38 (1), L01706. https://doi.org/10.1029/2010GL045777.
Kubar, T.L., Hartmann, D.L., Wood, R., 2009. Understanding the importance of microphysics and macrophysics for warm rain in marine low clouds. Part I: Satellite observations. J. Atmos. Sci. 66 (10), 2953–2972.
Kumler, A., Xie, Y., Zhang, Y., 2019. A Physics-based Smart Persistence model for Intra-hour forecasting of solar radiation (PSPI) using GHI measurements and a cloud retrieval technique. Sol. Energy 177, 494–500.
Lauret, P., Voyant, C., Soubdhan, T., David, M., Poggi, P., 2015. A benchmarking of machine learning techniques for solar radiation forecasting in an insular context. Sol. Energy 112, 446–457.
Law, E.W., Prasad, A.A., Kay, M., Taylor, R.A., 2014. Direct normal irradiance forecasting and its application to concentrated solar thermal output forecasting–A review. Sol. Energy 108, 287–307.
Liu, B.Y.H., Jordan, R.C., 1960. The interrelationship and of direct, diffuse and characteristic distribution total solar radiation. Sol. Energy 4 (3), 1–19.
Liu, Y., 2019. Introduction to the special section on fast physics in climate models: parameterization, evaluation, and observation. J. Geophys. Res.: Atmos. 124 (15), 8631–8644.
Liu, Y., Wu, W., Jensen, M.P., Toto, T., 2011. Relationship between cloud radiative forcing, cloud fraction and cloud albedo, and new surface-based approach for determining cloud albedo. Atmos. Chem. Phys. 11 (14), 7155–7170.
Long, C.N., Ackerman, T.P., 2000. Identification of clear skies from broadband pyranometer measurements and calculation of downwelling shortwave cloud effects. J. Geophys. Res.: Atmos. 105 (D12), 15609–15626.
Long, C.N., Gaustad, K.L., 2004. The Shortwave (SW) Clear-Sky Detection and Fitting Algorithm: Algorithm Operational Details and Explanations." Atmospheric Radiation Measurement Program Technical Report, ARM TR-004, Available via https://www.arm.gov/publications/tech_reports/arm-tr-004.pdf.
Martín, L., Zarzalejo, L.F., Polo, J., Navarro, A., Marchante, R., Cony, M., 2010. Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning. Sol. Energy 84 (10), 1772–1781.
Matus, A.V., L'Ecuyer, T.S., 2017. The role of cloud phase in Earth's radiation budget. J. Geophys. Res.: Atmos. 122 (5), 2559–2578.
Okata, M., Nakajima, T., Suzuki, K., Inoue, T., Nakajima, T.Y., Okamoto, H., 2017. A study on radiative transfer effects in 3D cloudy atmosphere using satellite data. J. Geophys. Res.: Atmos. 122 (1), 443–468.
Pfister, G., McKenzie, R.L., Liley, J.B., Thomas, A., Forgan, B.W., Long, C.N., 2003. Cloud coverage based on all-sky imaging and its impact on surface solar irradiance. J. Appl. Meteorol. 42 (10), 1421–1434.
Ramanathan, V.L.R.D., Cess, R.D., Harrison, E.F., Minnis, P., Barkstrom, B.R., Ahmad, E., Hartmann, D., 1989. Cloud-radiative forcing and climate: Results from the earth radiation budget experiment. Science 243 (4887), 57–63.
Reda, I., Andreas, A., 2004. Solar position algorithm for solar radiation applications. Sol. Energy 76 (5), 577–589.
Reikard, G., 2009. Predicting solar radiation at high resolutions: A comparison of time series forecasts. Sol. Energy 83 (3), 342–349.
Rosenfeld, D., 2006. Aerosol-cloud interactions control of earth radiation and latent heat release budgets. In: Solar Variability and Planetary Climates. Springer, New York, NY, pp. 149–157.
Shonk, J.K., Hogan, R.J., Edwards, J.M., Mace, G.G., 2010. Effect of improving representation of horizontal and vertical cloud structure on the Earth's global radiation budget. Part I: Review and parametrization. Q. J. R. Meteorolog. Soc. 136 (650), 1191–1204.
Shonk, J.K., Hogan, R.J., 2010. Effect of improving representation of horizontal and vertical cloud structure on the Earth's global radiation budget. Part II: The global effects. Q. J. R. Meteorolog. Soc. 136 (650), 1205–1215.
Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res.: Atmos. 106 (D7), 7183–7192.
Voyant, C., Notton, G., Kalogirou, S., Nivet, M.L., Paoli, C., Motte, F., Fouilloy, A., 2017. Machine learning methods for solar radiation forecasting: A review. Renewable Energy 105, 569–582.
Xie, Y., Liu, Y., 2013. A new approach for simultaneously retrieving cloud albedo and cloud fraction from surface-based shortwave radiation measurements. Environ. Res. Lett. 8 (4), 044023.
Yang, D., Alessandrini, S., Antonanzas, J., Antonanzas-Torres, F., Badescu, V., Beyer, H. G., David, M., 2020. Verification of deterministic solar forecasts. Sol. Energy. https://doi.org/10.1016/j.solener.2020.04.019.
Yang, D., Jirutitijaroen, P., Walsh, W.M., 2012. Hourly solar irradiance time series forecasting using cloud cover index. Sol. Energy 86 (12), 3531–3543.
Yang, D., Quan, H., Disfani, V.R., Rodríguez-Gallegos, C.D., 2017. Reconciling solar forecasts: Temporal hierarchy. Sol. Energy 158, 332–346.
Yang, D., Sharma, V., Ye, Z., Lim, L.I., Zhao, L., Aryaputera, A.W., 2015. Forecasting of global horizontal irradiance by exponential smoothing, using decompositions. Energy 81, 111–119.