Article ID: 1006-8775(2024)03-0327-10

Multimodel Ensemble Forecast of Global Horizontal Irradiance at PV Power Stations Based on Dynamic Variable Weight

YUAN Bin (袁 彬)[1,2], SHEN Yan-bo (申彦波)[1,2,6], DENG Hua (邓 华)[3], YANG Yang (杨 扬)[4], CHEN Qi-ying (陈起英)[5], YE Dong (叶 冬)[1,2], MO Jing-yue (莫景越)[1,2], YAO Jin-feng (姚锦烽)[1,2], LIU Zong-hui (刘宗会)[6]

(1. Public Meteorological Service Center, China Meteorological Administration, Beijing 100081 China; 2. Key Laboratory of Energy Meteorology, China Meteorological Administration, Beijing 100081 China; 3. Guangzhou Institute of Tropical and Marine Meteorology, China Meteorological Administration, Guangzhou 510641 China; 4. Institute of Urban Meteorology, China Meteorological Administration, Beijing 100089 China; 5. Center for Earth System Modeling and Prediction, China Meteorological Administration, Beijing 100081 China; 6. Institute of Desert Meteorology, China Meteorological Administration, Urumqi 830002 China)

Abstract: In the present study, multimodel ensemble forecast experiments of the global horizontal irradiance (GHI) were conducted using the dynamic variable weight technique. The study was based on the forecasts of four numerical models, namely, the China Meteorological Administration Wind Energy and Solar Energy Prediction System, the Mesoscale Weather Numerical Prediction System of China Meteorological Administration, the China Meteorological Administration Regional Mesoscale Numerical Prediction System-Guangdong, and the Weather Research and Forecasting Model-Solar, and observational data from four photovoltaic (PV) power stations in Yangjiang City, Guangdong Province. The results show that compared with those of the monthly optimal numerical model forecasts, the dynamic variable weight-based ensemble forecasts exhibited 0.97%–15.96% smaller values of the mean absolute error and 3.31%–18.40% lower values of the root mean square error (RMSE). However, the increase in the correlation coefficient was not obvious. Specifically, the multimodel ensemble mainly improved the performance of GHI forecasts below 700 W m^{-2}, particularly below 400 W m^{-2}, with RMSE reductions as high as 7.56%–28.28%. In contrast, the RMSE increased at GHI levels above 700 W m^{-2}. As for the key period of PV power station output (02:00–07:00), the accuracy of GHI forecasts could be improved by the multimodel ensemble: the multimodel ensemble could effectively decrease the daily maximum absolute error (AE_{max}) of GHI forecasts. Moreover, with increasing forecasting difficulty under cloudy conditions, the multimodel ensemble, which yields data closer to the actual observations, could simulate GHI fluctuations more accurately.

Key words: GHI forecast; multimodel ensemble; dynamic variable weight; PV power station

CLC number: P456.1 **Document code:** A

Citation: YUAN Bin, SHEN Yan-bo, DENG Hua, et al. Multimodel Ensemble Forecast of Global Horizontal Irradiance at PV Power Stations Based on Dynamic Variable Weight [J]. Journal of Tropical Meteorology, 2024, 30(3): 327–336, https://doi.org/10.3724/j.1006-8775.2024.027

1 INTRODUCTION

Under the carbon peaking and carbon neutrality targets, the development of a new power system that mainly relies on new energy is regarded as an important precondition and necessary direction to achieve the low-carbon transition of modern power systems (Han et al. [1]). China's new energy has entered a large-scale, high-penetration, and market-oriented development stage. By the end of 2023, the country's accumulative installation capacity of power generation have reached approximately 2.92 TW, of which the installation capacity of photovoltaic (PV) power amounted to approximately 610 GW (Nea [2]), accounting for more than 20% of the total amount. Solar power is playing an increasingly important role in new power systems, highlighting the need for grid integration (Jin et al. [3]). In practice, the dependable capacity offered by new energy systems is less than 5% due to power prediction fluctuations (Creei [4]). Therefore, more accurate power predictions are important for solar power generation integration and power grid dispatching.

Global horizontal irradiance (GHI) forecast provides a direct data source for power prediction. Its accuracy determines, to a great extent, the prediction accuracy of the power output of PV power stations, which serves as the essentials for power grids to formulate dispatching plans and ensure energy supply; such a forecast is therefore associated with power system safety and reliability (Singla et al. [5]). The GHI, which is highly dependent on atmospheric conditions, exhibits significant seasonality,

Received 2024-02-27; **Revised** 2024-05-15; **Accepted** 2024-08-15

Funding: Innovation and Development Project of China Meteorological Administration (CXFZ2023J044); Innovation Foundation of CMA Public Meteorological Service Center (K2023002); "Tianchi Talents" Introduction Plan (2023); Key Innovation Team for Energy and Meteorology of China Meteorological Administration

Biography: YUAN Bin, MSc, primarily undertaking research on wind and solar power forecasts.

Corresponding author: SHEN Yan-bo, e-mail: shenyb@cma.gov.cn

volatility, and intermittency (Pardeep et al. [6]). At present, GHI short-term forecast mainly stems from numerical weather prediction (NWP), but existing NWP suffers an obvious defect in the parameterization scheme of cloud microphysical processes, which composes the main factor leading to NWP uncertainty (Huang et al. [7]). Moreover, cloud cover directly affects GHI forecasts, causing the accuracy to not fully meet the assessment requirements of PV power stations. Furthermore, different numerical prediction models exhibit significant variations in terms of their dynamic frameworks, terrain features, parameterization schemes, physical processes, spatial and temporal resolutions, etc., which also lead to notable differences in the prediction results of NWP models (Bougeault et al. [8]). Therefore, it seems necessary to improve the accuracy of GHI forecasts by using various postprocessing techniques. By integrating the effective forecast information of multiple models, a multimodel ensemble, as an effective postprocessing tool, could significantly improve the forecast skill and reliability (Zhu [9]; Rahimi et al. [10]).

In deterministic forecasting, a multimodel ensemble normally encompasses two schemes: the equal-weight multimodel ensemble, which mainly includes the ensemble mean and bias-removed ensemble mean, and the unequal-weight multimodel ensemble, which mostly includes the superensemble, Kalman filter, and Bayesian model averaging. Compared with the equal-weight multimodel ensemble, the unequal-weight multimodel ensemble assigns a higher weight to models with high prediction skills by determining the prediction ability of each model during the training period in the ensemble process. It takes full advantage of each model and thus has better performance than equal-weight multimodel ensemble (Zhi et al. [11]). The allocation of reasonable weights to each model in the unequal-weight ensemble scheme is thus an important procedure for obtaining final ensemble prediction results (Zhou et al. [12]). Many scholars have applied the unequal-weight multimodel ensemble method in weather forecasting, which has effectively improved forecast accuracy. Wei et al. [13] used the hierarchical optimization weight ensemble forecast method to conduct an ensemble experiment for generating precipitation forecasts in the Pan-Yangtze River region and suggested that the threat scores of the ensemble forecasts were greater than those of the numerical model forecasts at all intervals and all precipitation levels. Sheng et al. [14] compared different objective forecasting methods and optimal ensemble forecasts of temperature, and the results demonstrated that the ensemble forecasting method could improve the accuracy on the basis of various objective forecasts. Tong et al. [15] employed the Bayesian model averaging method to create an ensemble of four models of the China Meteorological Administration (CMA) model system, which effectively reduced the prediction errors of the 2-m temperature, 10-m wind speed, and 2-m relative humidity in the Beijing–Tianjin–Hebei region. Wu et al. [16] used augmented complex extended Kalman filter to conduct multimodel ensemble forecast experiments of the wind speed in East China. They effectively reduced the forecast error and verified that multimodel ensembles were more effective in complex terrain areas (Wu et al. [16]). Zhi et al. [17] developed a multimodel ensemble of surface and upper-air wind fields in East China by using Kalman filter. The results showed that the error significantly decreased after applying the multimodel ensemble, while for the upper-air wind field, the change in the ensemble forecast error associated with height was less than that in the single-model forecast error. Zhao et al. [18] developed an improved ensemble method based on the Markov process and ordered weighted average for day-ahead forecast of local wind speeds. They effectively reduced the uncertainties of numerical simulations and showed that an ensemble with fewer members could generate better results than using a combination of all single members (Zhao et al. [18]).

In Chinese and overseas weather forecasting efforts, although multimodel ensemble techniques have been extensively adopted to improve the accuracy of temperature, precipitation, and wind speed forecasts (Du et al. [19]), their application in GHI forecasting remains inadequate, and currently, available research mainly focuses on the integration of multiple correction methods for a single numerical forecast model. An extensive literature review reveals the following: Sun et al. [20] proposed a decomposition-clustering-ensemble learning method for GHI forecasting, which provided a favorable effect in the Beijing area. Guermoui et al. [21] used the integration technique of multiple machine learning methods to effectively improve the GHI prediction accuracy in Algeria, and the experimental results indicated that the integration technique was superior to the benchmark model across all prediction stages (Guermoui et al. [21]). Baek et al. [22] compared different combinations of NWP scenarios and machine learning algorithms by using weighted integration of various machine learning models. They found that the best model could be obtained from the combinations of multiple prediction machines through weighted averaging and the use of all NWP scenarios (Baek et al. [22]). Jiang et al. [23] suggested that the integrated learning framework could achieve superior performance and improve prediction stability. Basaran et al. [24] evaluated several ensemble models in solar irradiance estimation including random forest, support vector regression, artificial neural network, and decision tree. Despite the above works, there is little research on multimodel ensemble forecasting for GHI, and the effect of multimodel ensemble forecasting at PV power stations need to be verified further.

Based on the above research, from the perspective of multimodel ensemble forecasts of GHI, in this paper, GHI observation data from PV power stations were used to dynamically correct multiple numerical model forecasts,

and the dynamic variable weight multimodel ensemble technique was then employed to establish a forecast model. The variable weight theory could combine the forecasting models preferably, which mainly included statistical methods and artificial intelligence methods (Zhi et al. [17]; Jiang et al. [25]; Chu et al. [26]). Rolling multimodel ensemble experiments were conducted at four PV power stations in Yangjiang City, Guangdong Province, and the ensemble forecast performance was evaluated according to the root mean square error (RMSE), mean absolute error (MAE), correlation coefficient *R*, and absolute error (AE). All the above efforts aimed to eliminate the systematic deviation and maximize the performance of each numerical forecast model to obtain the optimal ensemble forecast. The accuracy of GHI forecasts at PV power stations could thus be improved.

2 DATA AND METHODS

2.1 *Data*

In this paper, Yangjiang City (21.5°–22.7°N, 111.3°–112.4°E), Guangdong Province, was selected as the research area. Located on the southwestern coast of Guangdong Province, Yangjiang experiences a typical subtropical monsoon climate, with long durations of sunshine and abundant heat. The terrain of Yangjiang City is dominated by mountainous and hilly areas, with mountains in its eastern, western, and northern parts, while its southern part faces the South China Sea. The four PV power stations in Yangjiang City, selected as targets in this study, are all situated in relatively flat terrain, with an altitude of less than 30 m. The distribution of the four PV power stations is shown in Fig. 1.

The observational data used were GHI data from the four meteorological stations at the center of each PV power station. These observational data were chosen as they can better represent the GHI observations at each PV power station. The duration of the observation sequence was the entire year of 2022, and the observation frequency was 15 minutes. The GHI observational data obtained from each PV power station were subjected to quality control procedures according to the *Solar Energy Resource Assessment Method GB/T 37526–2019* [27], and any data exceeding 1400 W m^{-2} and data remaining unchanged for longer than five consecutive hours were treated as default values. After the quality control procedure, 53.24%, 96.29%, 96.28%, and 96.16% of the total observational data from the four PV power stations were retained. Notably, data from PV station 1 in April, May, August, September, and December were not available, but the quality of data during the remainder of the year was much better.

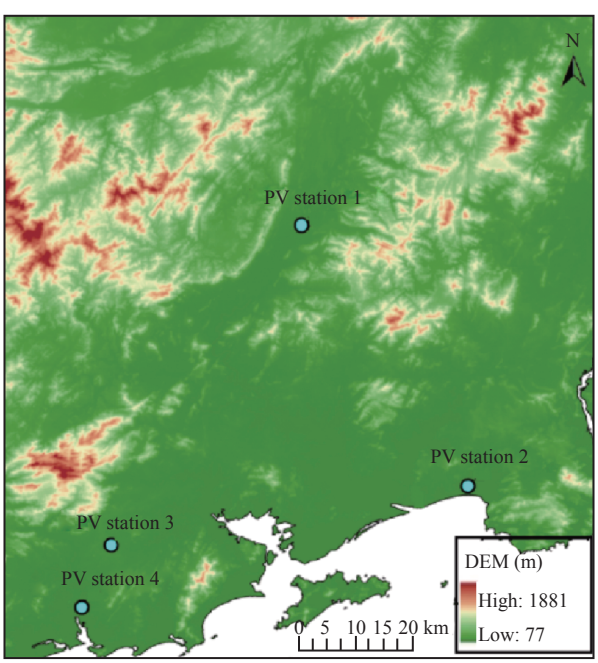

Figure 1. Research area and location of the four PV stations.

The numerical models adopted in this paper were the China Meteorological Administration Wind Energy and Solar Energy Prediction System (CMA-WSP), the Mesoscale Weather Numerical Prediction System of China Meteorological Administration (CMA-MESO), the China Meteorological Administration Regional Mesoscale Numerical Prediction System-Guangdong (CMA-GD), and the Weather Research and Forecasting Model-Solar (WRF-SOLAR). The CMA-WSP, CMA-MESO, and CMA-GD models are wind and solar numerical forecasting models independently developed by the CMA and operated in real time, while the WRF-SOLAR model is an important part of the National Center for Atmospheric Research (NCAR) solar power forecasting system. The WRF-SOLAR model was designed specifically to meet solar forecasting demands, and it is also operated in real time by the CMA. The initial time of these four models is 12:00 UTC. For the PV power stations, since a 24-h ahead forecast is required for assessment (i.e., 24 hours starting at 00:00 on the forthcoming day, Beijing time, BJT), the model forecast period was therefore chosen as 52 hours. Moreover, the forecast sequence duration was still the whole year of 2022. The forecast details of each model used are provided in Table 1.

In this research, the proximal point algorithm was applied to extract GHI forecast data at the target points (the locations of the four PV power stations) from the four numerical models listed in Table 1 for forecast dataset development.

2.2 *Methods*

To conduct GHI forecast multimodel ensemble experiments, the dynamic variable weight ensemble method was adopted for modeling station-by-station forecasts. The key technique was a weighted bias-removed ensemble, where the weights were determined by the reciprocal of the prediction errors of each numerical model within a certain training period that dynamically slid with the rolling updates of the forecasts (Liu et al. [28]). First, with the use of the GHI observations of the PV

Table 1. Details of the numerical prediction models.

Model	Temporal resolution (min)	Spatial resolution (km)	Forecast element	Forecast period (h)
CMA-WSP	15	9	GHI	52
CMA-MESO	60	3		
CMA-GD	15	3		
WRF-SOLAR	15	9		

power stations, deviation calculations and dynamic deviation corrections of different numerical model forecasts were performed. Then, through statistical analysis of the deviation values, the weight of each numerical model was determined, and a dynamic variable weight multimodel ensemble forecast model could be established for each PV power station. In this paper, the training period was set to 10 days, namely, by applying the error analysis and weight coefficient data of the previous ten days, rolling correction and daily ensemble forecasting for a specific day could be fulfilled, and the specific steps are as follows:

By calculating the forecast errors of each numerical model forecast for each PV power station, we set the training period to 10 days:

$$\mathrm{BE}_{mi} = F_{mi} - \mathrm{O}_i \tag{1}$$

where F_{mi} is the GHI forecast value at the i-th forecast time of the m-th numerical model for a single power station, O_i denotes the corresponding GHI observation value, and BE_{mi} is the error of the i-th forecast time of the m-th numerical model.

Sequencing the errors of the m-th numerical model in ascending order at all forecast times during the training period for a single power station, we applied the percentile method to calculate the deviation in the GHI forecasts of each numerical model:

$$\mathrm{BES}_m = \frac{\mathrm{BE}_{m0.25} + 2\mathrm{BE}_{m0.5} + \mathrm{BE}_{m0.75}}{4} \tag{2}$$

where BES_m is the systematic forecast deviation during the training period of the m-th numerical forecast model.

The forecasts of each numerical model could be corrected as follows:

$$\mathrm{FF}_{mi} = F_{mi} - \mathrm{BES}_m \tag{3}$$

where FF_{mi} is the forecast result at the i-th forecast time of the m-th numerical model after deviation correction.

According to the statistical results of the forecast errors of each numerical model during the training period, the ensemble weight of each numerical model could be calculated as follows:

$$W_m = \frac{\frac{1}{A_m}}{\sum_{m=1}^{M}\frac{1}{A_m}} \tag{4}$$

where M is the number of numerical models involved in the ensemble, and A_m is the sum of the AE values during the training period of the m-th numerical prediction model.

The dynamic variable weight multimodel ensemble model could be established as follows:

$$Y_i = \sum_{m=1}^{M} W_m \times \mathrm{FF}_{mi} \tag{5}$$

where Y_i is the multimodel ensemble GHI forecast result at the i-th forecast time.

Given that the temporal resolution of the CMA-MESO model is 1 hour and that of the other three numerical models is 15 minutes, the four models were assembled for each hour of the day, while three models—CMA-WSP, CMA-GD, and WRF-SOLAR—were assembled for the remainder of the time.

The MAE, RMSE, R, and $\mathrm{AE}_{\max}$ (as the temporal resolution was 15 minutes, there were 96 forecast results per day, corresponding to 96 AEs, and the maximum one was defined as $\mathrm{AE}_{\max}$) were used for evaluation. All the evaluations were independent. These metrics could be obtained as follows:

$$\mathrm{AE} = \left|F_i - O_i\right| \tag{6}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|(F_i - O_i)\right| \tag{7}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(F_i - O_i)^2} \tag{8}$$

$$R = \frac{\sum_{i=1}^{n}\left(F_i - \overline{F_i}\right)\left(O_i - \overline{O_i}\right)}{\sqrt{\sum_{i=1}^{n}\left(F_i - \overline{F_i}\right)^2\sum_{i=1}^{n}\left(O_i - \overline{O_i}\right)^2}} \tag{9}$$

where O_i is the observation value, F_i is the forecast value, n is the total number of samples, $\overline{O_i}$ is the average value of the observation samples, and $\overline{F_i}$ is the average value of the forecast samples.

In this study, the GHI forecast error was evaluated for the region (the four PV power stations as a whole) and for individual stations. The evaluation period ranged from 0 to 24 hours a day ahead (i.e., 29–52 hours of each numerical model forecast), and the evaluation duration was the entire year of 2022.

3 RESULTS

3.1 *Monthly forecast error evaluation*

Table 2 provides the results of the performance evaluation of the monthly GHI forecasts for the region. Notably, the CMA-GD model exhibited the smallest error from January to December among the four numerical models. The average MAE was 134 W m^{-2}, and the RMSE

Table 2. Comparison of the performance of the monthly multimodel ensemble and numerical model forecasts.

Month	MAE (W m^{-2})					RMSE (W m^{-2})					*R*				
	ENSEMBLE	CMA WSP	CMA MESO	CMA GD	WRF SOLAR	ENSEMBLE	CMA WSP	CMA MESO	CMA GD	WRF SOLAR	ENSEMBLE	CMA WSP	CMA MESO	CMA GD	WRF SOLAR
1	105	122	131	135	134	180	207	211	216	225	0.79	0.79	0.76	0.72	0.77
2	81	91	106	101	113	139	154	177	163	193	0.90	0.89	0.83	0.86	0.84
3	114	130	132	142	165	176	209	208	217	262	0.87	0.85	0.83	0.81	0.82
4	109	150	121	110	147	164	237	192	170	233	0.92	0.89	0.90	0.92	0.90
5	119	156	151	137	172	184	244	240	217	269	0.86	0.84	0.78	0.81	0.83
6	127	156	158	152	180	193	235	246	237	270	0.85	0.83	0.76	0.77	0.82
7	141	150	142	141	151	202	232	217	209	236	0.91	0.89	0.90	0.90	0.90
8	140	166	161	166	171	206	249	251	249	260	0.83	0.82	0.77	0.77	0.81
9	136	160	131	147	159	199	247	203	226	243	0.89	0.87	0.90	0.86	0.88
10	141	122	130	132	103	191	195	185	180	163	0.94	0.93	0.94	0.94	0.95
11	103	129	126	118	140	161	200	186	174	211	0.86	0.80	0.82	0.84	0.83
12	107	109	113	129	117	172	171	168	191	181	0.88	0.87	0.88	0.84	0.87
Mean	119	137	133	134	146	181	215	207	204	229	0.88	0.86	0.84	0.84	0.85

was 204 W m^{-2} for 12 months. On the other hand, the CMA-WSP model attained the highest *R* value, and the average value for 12 months reached 0.86. There were differences in each month: the CMA-WSP model performed best in January, February, March, and December; the CMA-GD model performed best in April, May, June, July, August, and November; the CMA-MESO model performed best in September; the WRF-SOLAR model performed best in October. The forecast errors of the multimodel ensemble (referred to as ENSEMBLE in the figures and tables) were significantly reduced: the average MAE reached 119 W m^{-2}, and the RMSE was 181 W m^{-2} for 12 months, which were 11.19% and 11.27% lower, respectively, than those of the optimal numerical models, while *R* slightly increased to 0.88. The forecast performance of the multimodel ensemble obviously differed each month. The smallest errors occurred from January to September and also in November, and the MAE and RMSE values decreased by 0.97%–15.96% and 3.31%–18.40%, respectively, compared to those of the monthly optimal numerical model forecasts. However, in October, the forecast error after the multimodel ensemble application was slightly greater than that of the optimal numerical model forecast. In December, the optimal numerical model and multimodel ensemble attained similar forecast performance levels. The improvement in *R* obtained by the multimodel ensemble was not obvious compared with that of the optimal numerical model forecast, and the values ranged from only 1.12% to 2.47% in 8 out of 12 months.

The forecast errors were directly related to the observed GHI values, which exhibited obvious diurnal variations. To analyze the performance of the multimodel ensemble forecasts at different GHI intensities, we considered three GHI intervals in this study: (0, 400), [400, 700], and (700, 1500). Fig. 2 shows the results of the intensity level evaluation of regional GHI forecasts. The performance of each numerical forecast model varied

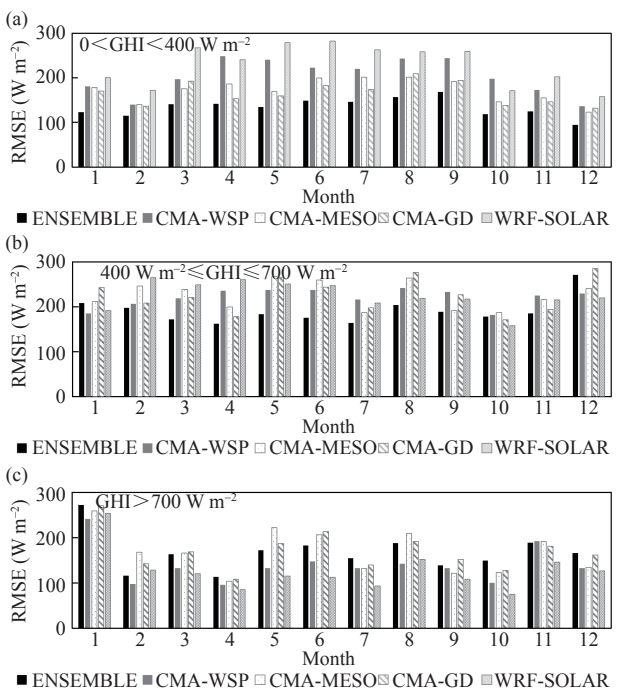

Figure 2. Comparison of the monthly multimodel ensemble and numerical model forecasts at different intensities. RMSE was used for evaluation, and the different GHI intensities are (a) 0 < GHI<400 W m^{-2}, (b) 400 W m^{-2}≤GHI≤700 W m^{-2}, and (c) 700 W m^{-2}<GHI<1500 W m^{-2}.

within different intervals. The multimodel ensemble mainly improved GHI forecasting performance below 700 W m^{-2}. In particular, at GHI levels lower than 400 W m^{-2}, the forecast error of the CMA-GD model, among four numerical models, was the smallest, but the multimodel ensemble error was smaller than that of the CMA-GD model, with a 7.56% to 28.28% reduction in the RMSE value. The CMA-GD and CMA-WSP model forecasts exhibited advantages within the GHI range of 400 to 700 W m^{-2}, while the multimodel ensemble forecast could reduce the RMSE by 4.72% to 26.10% in 9 out of

12 months. At GHI levels greater than 700 W m^{-2}, the forecasts of each numerical forecast model were significantly smaller than the observations. The WRF-SOLAR model had the lowest deviation, and therefore its forecast error was the smallest. However, the RMSE increased after the application of multimodel ensemble. The insufficient samples within intervals above 700 W m^{-2} (Table 3), coupled with the larger forecast error and fluctuation amplitude of the numerical model, partly contributed to an increased or even reversed systematic correction deviation in the process of correcting each numerical model with the rolling deviation over the previous ten days, resulting in an increase in the forecast error (Eq. 3) after multimodel ensemble implementation.

3.2 *Diurnal variation in forecast error evaluation*

The GHI value at midday, the key period for the PV power station output, is high, and the accuracy of GHI forecasting is more important for this period. To examine the performance of the multimodel ensemble forecasts at different times of one day, the regional GHI forecasts were compared over time. Fig. 3 shows the evaluation results of the RMSE for each month. It suggests that the improvement effect of the multimodel ensemble was the greatest at midday when the RMSE reached the highest value. Meanwhile, the multimodel ensemble performance varied month by month. In months with larger numerical model forecast errors, the improvement effect of the multimodel ensemble was greater, but the multimodel ensemble yielded a limited improvement effect in months with smaller numerical model forecast errors. Specifically, compared with the optimal numerical model forecast, the multimodel ensemble provided the best GHI forecast

Table 3. Monthly sample size at different GHI intensities.

GHI (W m^{-2})	Sample size											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
(0, 400)	9653	7311	9121	5104	6563	9244	7765	5950	4591	5868	8893	3293
[400, 700]	1739	950	1969	1389	1268	2279	2498	1253	1484	2809	1795	744
(700, 1500)	489	677	1636	1500	1111	1922	3166	1295	1622	3007	698	577

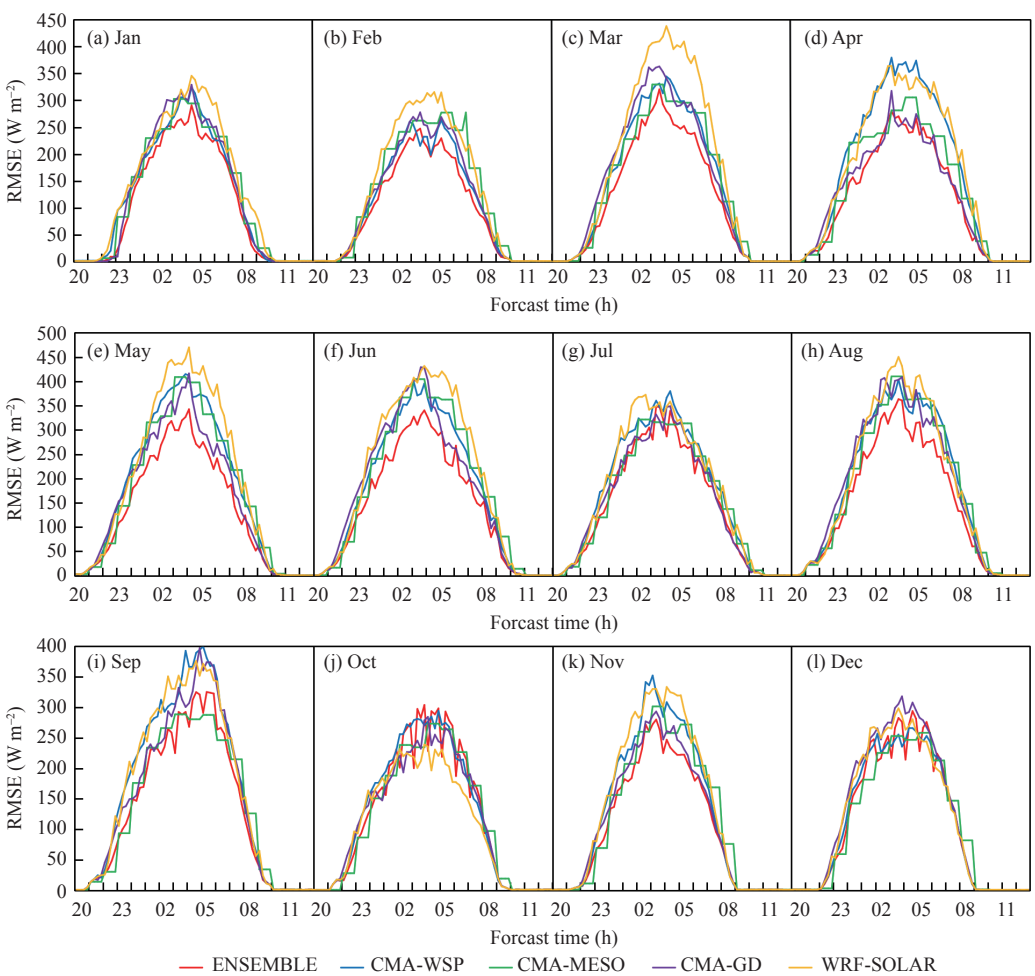

Figure 3. Monthly diurnal variations in the multimodel ensemble and numerical model forecasts. RMSE was used for evaluation from 20:00 UTC to 12:00 UTC the next day.

improvement in March, May, June, July, and August, and the RMSE of the multimodel ensemble forecast could be reduced by 33–85 W m^{-2} from 02:00 UTC to 07:00 UTC. In January, February, April, September, and November, the RMSE values of the multimodel ensemble forecasts were slightly smaller than those of the optimal numerical model forecasts. Nevertheless, the performance of the multimodel ensemble forecast in October and December was worse. In general, for the key output period at midday, the multimodel ensemble forecast attained the lowest RMSE and the most stable effect. Therefore, the effectiveness of the multimodel ensemble in improving the GHI forecast accuracy during the key period of the PV output could be verified.

3.3 *Forecast error evaluation at each PV station*

Figure 4 shows the forecast errors for each PV power station. In the scope of the numerical model, at PV station 1, the forecast error of the CMA-WSP model was the smallest, and the *R* of the WRF-SOLAR model was the highest. At PV station 2, the forecast error of the CMA-MESO model was the smallest, and the *R* of the CMA-WSP model was the highest. At PV station 3, the forecast errors of the CMA-GD and CMA-WSP models were relatively small, and the *R* of the CMA-WSP model was the highest. At PV station 4, the forecast errors of the CMA-GD and CMA-MESO models were relatively small, and the *R* of the CMA-WSP model was the highest. Moreover, the numerical model forecast error at PV station 2 was the smallest, indicating that the numerical model prediction capability significantly differed at various PV power stations.

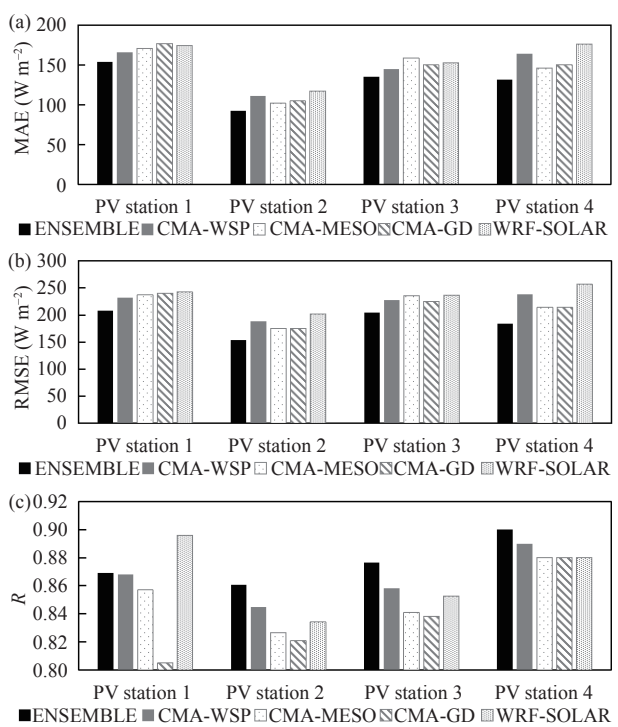

Figure 4. Comparison of the multimodel ensemble and numerical model forecasts at each PV power station. (a) MAE, (b) RMSE, and (c) *R* were used for evaluation.

To analyze the applicability and duplicability of the multimodel ensemble model at the different PV power stations, the ensemble effect at each PV power station for the whole year was evaluated. The results illustrated that the multimodel ensemble improved the forecast effect at these four PV power stations to different extents, as reflected by the smallest forecast errors and highest *R* values. Compared with the optimal numerical model forecasts, the MAE and RMSE values of the multimodel ensemble at the four PV power stations decreased by 13–19 W m^{-2} and 20–30 W m^{-2}, respectively. The *R* values at the three stations, except for PV station 1, increased by 0.01–0.02.

3.4 *Maximum forecast error evaluation at the PV stations*

The fluctuation in the PV power output is the greatest challenge for grid integration. In addition to the overall monthly forecast accuracy, power stations and power grids notably consider the transition in the PV power curve (Yu et al. [29]), which can be assessed through the monthly average value of the AE_{max} of GHI forecasts. The assessment can reflect the precision of the starting/ending time and the magnitude of weather transition determined by numerical models, especially GHI fluctuations due to cloud cover variations, which constitutes a complex issue and bottleneck in the high-temporal-resolution numerical model forecasting. Fig. 5 shows the evaluation results for PV station 4 as an example. The monthly mean of AE_{max} of each numerical model forecast is usually very large. The AE_{max} of the CMA-WSP model ranged from 293 W m^{-2} to 621 W m^{-2}, that of the CMA-MESO model ranged from 281 W m^{-2} to 504 W m^{-2}, that of the CMA-GD model ranged from 327 W m^{-2} to 607 W m^{-2}, and that of the WRF-SOLAR model ranged from 314 W m^{-2} to 669 W m^{-2}. The AE_{max} exhibited obvious seasonal variation, with high values in summer and low values in winter. Although the AE_{max} of the CMA-MESO model was the smallest, it could not reflect GHI fluctuations within an hour and therefore led to a limited reference since the temporal resolution of this model is 1 hour. The multimodel ensemble facilitated a reduction in the AE_{max} to 313–516 W m^{-2}. Moreover, compared with the monthly optimal numerical model forecasts (except for the CMA-MESO), the AE_{max} from January to September could be reduced by 5.72%–15.90%, while it increased by 2.04%–7.46% from October to December. It should be emphasized that the multimodel ensemble generated a

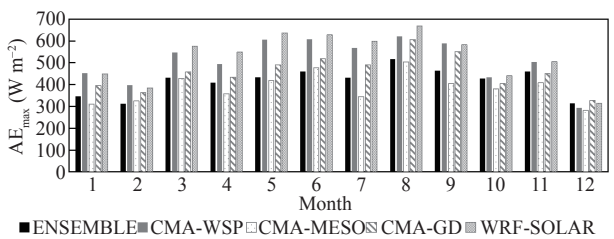

Figure 5. Comparison of the multimodel ensemble and numerical model forecasts at PV station 4. The AE_{max} was used for evaluation.

positive effect overall.

3.5 *Forecast performance under complex weather conditions*

The accuracy of GHI forecasts is significantly affected by weather conditions. On sunny days, the forecast accuracy is greater, while on cloudy days, due to the uncertainty in cloud cover, the forecasting difficulty increases, and the forecast accuracy decreases accordingly (Da et al. [30]). The forecasts and observations of PV station 4 from May 8 to 12 and from 21:00 UTC to 12:00 UTC the next day were compared to analyze the performance of the multimodel ensemble under variable cloudy weather conditions. Cloud cover observational data were obtained from the Yangjiang National Meteorological Station. The comparison results are shown in Fig. 6. From May 8 to 12, the daytime was mainly cloudy, the cloud cover fluctuated between 80% and 100%, and the GHI fluctuation was significant, which led to notable differences in the magnitude and fluctuation phase between the numerical model forecasts and observations. The forecasts of the CMA-WSP and WRF-SOLAR models were significantly greater than the observations, and the errors could reach above 500 W m^{-2} at midday. The forecasts of the CMA-GD and CMA-MESO models were lower than the observations, and the variation trend greatly differed from real situation. Although the forecasts of the CMA-MESO model were close to the observations, they could not reflect the refined GHI change due to the lack of 15-minute forecasts. Therefore, the CMA-MESO model could only offer a basic reference. Compared with the numerical model forecasts, the multimodel ensemble forecasts were closer to the observations and provided a better performance with regard to GHI fluctuations. On May 9 and 10, all numerical model forecasts failed to accurately capture GHI changes, and the multimodel ensemble achieved the optimal forecast improvement effect. Although the multimodel ensemble was not the optimal forecast at every moment, it remained the most stable and reliable from a long-term perspective.

4 CONCLUSION AND DISCUSSION

With the use of the CMA-WSP, CMA-MESO, CMA-GD, and WRF-SOLAR model forecasts and GHI observational data from four PV power stations in Yangjiang City, Guangdong Province, in 2022, the dynamic variable weight multimodel ensemble method was adopted to conduct rolling error correction and ensemble experiments. The main conclusions are as follows:

The multimodel ensemble could effectively reduce the MAE and RMSE of GHI forecasts, but the ensemble performance varied greatly from month to month. Compared with those of the monthly optimal numerical model forecast, the MAE could be reduced by 0.97%–15.96%, and the RMSE could be reduced by 3.31%–18.40%. However, improvement in R obtained by the multimodel ensemble was not obvious, and the values ranged from only 1.12% to 2.47% in 8 out of 12 months. From an intensity level evaluation perspective, the multimodel ensemble provided improved GHI forecasts below 700 W m^{-2}, and the effect was remarkable, particularly at GHI levels below 400 W m^{-2}, with a 7.56%–28.28% decrease in the RMSE compared with that of the optimal numerical model forecast in each month. On the contrary, at GHI levels greater than 700 W m^{-2}, the RMSE increased after multimodel ensemble application.

During the key period of the PV power output (02:00 UTC to 07:00 UTC), the multimodel ensemble generated improved GHI forecast performance. Compared with that of the optimal numerical model forecast, the RMSE could be reduced by 33–85 W m^{-2} in March, May, June, July, and August, while in January, February, April, September, November, and December, the decline of RMSE was slight. Notably, the performance worsened in October. In general, the effect of the multimodel ensemble forecast was optimal and remained the most stable.

Regarding the AE_{max}, for which PV power stations and power grids have special concerns, the multimodel ensemble yielded a certain improvement. Compared with that of the optimal numerical model forecast for each month, the errors from January to September were reduced by 5.72%–15.90%, while those from October to December increased by 2.04%–7.46%. Overall, the multimodel ensemble provided positive effects. When dealing with

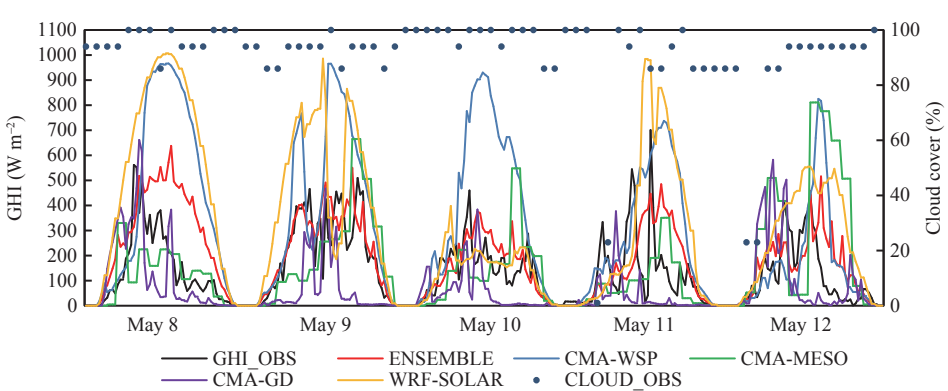

Figure 6. Comparison of the multimodel ensemble and numerical model forecasts at PV station 4 from May 8 to May 12. The blue dots illustrate the cloud cover observations per hour.

the forecasting difficulties under cloudy conditions, the multimodel ensemble had results that were closer to the observations, and it achieved a better performance regarding GHI fluctuations.

In this study, the four PV power stations were modeled separately. However, due to the relatively uniform underlying surface conditions of the four PV stations, as well as the small terrain shielding effect, the advantages of the WRF-SOLAR model were not fully reflected. The WRF-SOLAR model may have greater potential for forecasting under complex terrain conditions, which needs to be tested in other areas. The forecast results of the four PV power stations after the multimodel ensemble application improved to various levels. The forecast error was always the smallest, and the *R* value was the greatest, and therefore the multimodel ensemble exhibited great application potential.

However, the research method of this paper exhibited several limitations. This method only considered the GHI while ignoring the influence of other variables, such as albedo and precipitation. Moreover, the development of separate models for each PV power station failed to consider spatial impacts. Furthermore, the length of the training period was set to 10 days. Some studies have shown that the length of the training period can impact multimodel ensemble forecasts and should be investigated further. Therefore, in subsequent research, further experiments can be conducted from multiple perspectives to improve the accuracy of GHI forecasts.

REFERENCES

[1] HAN F J, WANG X H, QIAO J, et al. Review on artificial intelligence based load forecasting research for the new-type power system [J]. Proceedings of the CSEE, 2023, 43(22): 8569–8591, in Chinese with English abstract, https://doi.org/10.13334/j.0258-8013.pcsee.221560

[2] National Energy Administration. National Energy Administration 2023 National Power Industry Statistics [EB/OL]. [Available at: https://www.nea.gov.cn/2024-01/26/c_1310762246.htm, 2024-01-26/2024-02-02]

[3] JIN Q S, WANG X, NI Y L, et al. Development research and outlook on photovoltaic industry under carbon peaking and carbon neutrality goals [J]. Environmental Protection, 2022, 50(1): 44–50, in Chinese with English abstract.

[4] China Electric Power Planning & Engineering Institute. Report on China Energy Development 2023 [R]. Beijing: China Electric Power Planning & Engineering Institute, 2023.

[5] SINGLA P, DUHAN M, SAROHA S. A comprehensive review and analysis of solar forecasting techniques [J]. Frontiers in Energy, 2022, 16(25): 187–223, https://doi.org/10.1007/s11708-021-0722-7

[6] PARDEEP S, MANOJ D, SUMIT S. A point and interval forecasting of solar irradiance using different decomposition based hybrid models [J]. Earth Science Informatics, 2023, 16(30): 2223–2240, https://doi.org/10.1007/S12145-023-01020-9

[7] HUANG Q, QIAN Y. Comparative analysis of single-moment and double-moment microphysics schemes in WRF on the heavy precipitation process of the macroscale and microscale characteristics of the cloud [J]. Transactions of Atmospheric Sciences, 2021, 44(4): 615–625, in Chinese with English abstract.

[8] BOUGEAULT P, TOTH Z, BISHOP C, et al. The THORPEX interactive grand global ensemble [J]. Bulletin of the American Meteorological Society, 2010, 91(8): 1059–1072, https://doi.org/10.1175/2010BAMS2853.1

[9] ZHU Y J. Ensemble forecast: A new approach to uncertainty and predictability [J]. Advances in Atmospheric Sciences, 2005, 22(6): 781–788, https://doi.org/10.1007/BF02918678

[10] RAHIMI N, PARK S, CHOI W, et al. A comprehensive review on ensemble solar power forecasting algorithms [J]. Journal of Electrical Engineering & Technology, 2023, 18: 719–733, https://doi.org/10.1007/s42835-023-01378-2

[11] ZHI X F, JI X D, ZHANG J. Multimodel ensemble forecasts of surface air temperature and precipitation over China by using Kalman filler [J]. Transactions of Atmospheric Sciences, 2019, 42(2): 197–206, in Chinese with English abstract.

[12] ZHOU H, QIN H, JI L Y, et al. Research progresses of multimodel ensemble forecast of surface meteorological elements [J]. Transactions of Atmospheric Sciences, 2022, 45(6): 815–825, in Chinese with English abstract.

[13] WEI G F, LIU H J, WU Q S, et al. Multi-model consensus forecasting technology with optimal weight for precipitation intensity levels [J]. Journal of Applied Meteorological Science, 2022, 31(6): 668–680, in Chinese with English abstract.

[14] SHENG C Y, FAN S D, RONG Y M, et al. Comparison of several objective methods and optimal consensus forecast study of temperature [J]. Meteorological Monthly, 2020, 46 (10): 1351–1361, in Chinese with English abstract.

[15] TONG H, ZHANG Y T, QI Q Q, et al. The multi-model blending forecasts of near-surface parameters based on CMA model system [J]. Meteorological Monthly, 2022, 48(12): 1539–1549, in Chinese with English abstract.

[16] WU B Y, ZHI X F, CHEN C H, et al. Multi-model ensemble forecasts of wind over East China by using augmented complex extended Kalman filter [J]. Meteorological Monthly, 2022, 48(4): 393–405, in Chinese with English abstract.

[17] ZHI X F, WU B Y, LUO Z H, et al. Multimodel ensemble forecast of high-resolution surface and high-level wind forecasts over East China [J]. Transactions of Atmospheric Sciences, 2023, 46(6): 917–927, in Chinese with English abstract.

[18] ZHAO J, WANG J Z, GUO Z H, et al. Multi-step wind speed forecasting based on numerical simulations and an optimized stochastic ensemble method [J]. Applied Energy, 2019, 255(1): 1–16, https://doi.org/10.1016/j.apenergy.2019.113833

[19] DU J, BERNER J, BUIZZA R, et al. Ensemble methods for meteorological predictions [M]// DUAN Q Y, PAPPENBERGER F, WOOD A, et al. (eds), Handbook of Hydrometeorological Ensemble Forecasting. HEIDELBERG: Springer Berlin Press, 2019: 99–149.

[20] SUN S L, WANG S Y, ZHANG G W, et al. A decomposition clustering ensemble learning approach for solar radiation forecasting [J]. Solar Energy, 2018, 163(15): 189–199, https://doi.org/10.1016/j.solener.2018.02.006

[21] GUERMOUI M, BENKACIALI S, GAIRAA K, et al. A novel ensemble learning approach for hourly global solar

radiation forecasting [J]. Neural Computing and Applications, 2021, 34(4): 2983–3005, https://doi.org/10.1007/s00521-021-06421-9

[22] BAEK M K, LEE D. Spatial and temporal day-ahead total daily solar irradiation forecasting: ensemble forecasting based on the empirical biasing [J]. Energies, 2017, 11(1): 70, https://doi.org/10.3390/en11010070

[23] JIANG F, YANG J W. Short-term solar radiation forecast based on ensemble learning of multi-objective optimization [J]. Journal of Yunnan University: Natural Sciences Edition, 2021, 43(3): 451–461, in Chinese with English abstract.

[24] BASARAN K, ÖZCIFT A, KIlINC D. A new approach for prediction of solar radiation with using ensemble learning algorithm [J]. Arabian Journal for Science and Engineering, 2019, 44: 7159–7171, https://doi.org/10.1007/s13369-019-03841-7

[25] JIANG P, LIU Z K. Variable weights combined model based on multi-objective optimization for short-term wind speed forecasting [J]. Applied Soft Computing, 2019, 82: 105587, https://doi.org/10.1016/j.asoc.2019.105587

[26] CHU J C, YUAN L, PAN L, et al. NWP combination correction model based on variable-weight stacking algorithm [J]. Energy Procedia, 2019, 158: 6309–6314, https://doi.org/10.1016/j.egypro.2019.01.408

[27] State Administration for Market Regulation. Solar Energy Resource Assessment Method GB/T 37526–2019 [Z]. State Administration for Market Regulation, 2019.

[28] LIU Z J, LIU B X, WANG R, et al. Research on the gale prediction methods in the Yellow Sea and Bohai Sea based on traditional and deep learning technologies [J]. Marine Forecasts, 2022, 39(6): 34–43, in Chinese with English abstract.

[29] YU X X, SONG F L, LI J, et al. Power supply security improvement of power grid with high proportion of renewable energy under extreme weather events [J]. Modern Electric Power, 2023, 40(3): 303–313, in Chinese with English abstract.

[30] DA X F, LI Z R, WANG X Y, et al. Correction technology of short-time solar radiation forecast based on cloud cover [J]. Journal of Arid Meteorology, 2021, 39(6): 1006–1016, in Chinese with English abstract.

Citation: YUAN Bin, SHEN Yan-bo, DENG Hua, et al. Multimodel Ensemble Forecast of Global Horizontal Irradiance at PV Power Stations Based on Dynamic Variable Weight [J]. Journal of Tropical Meteorology, 2024, 30(3): 327–336, https://doi.org/10.3724/j.1006-8775.2024.027