Cognitive Computation (2023) 15:645–671
https://doi.org/10.1007/s12559-022-10070-y

Hybrid Convolutional Neural Network-Multilayer Perceptron Model for Solar Radiation Prediction

Sujan Ghimire[1] · Thong Nguyen-Huy[2,3] · Ramendra Prasad[4] · Ravinesh C. Deo[1] · David Casillas-Pérez[5] · Sancho Salcedo-Sanz[1,6] · Binayak Bhandari[7]

Received: 9 May 2022 / Accepted: 16 October 2022 / Published online: 7 November 2022

Abstract

Urgent transition from the dependence on fossil fuels towards renewable energies requires more solar photovoltaic power to be connected to the electricity grids, with reliable supply through accurate solar radiation forecasting systems. This study proposes an innovative hybrid method that integrates convolutional neural network (CNN) with multi-layer perceptron (MLP) to generate global solar radiation (*GSR*) forecasts. The CMLP model first extracts optimal topological and structural features embedded in predictive variables through a CNN-based feature extraction stage followed by an MLP-based predictive model to generate the *GSR* forecasts. Predictive variables from observed data and global climate models (GCM) are used to predict *GSR* at six solar farms in Queensland, Australia. A hybrid-wrapper feature selection method using a random forest-recursive feature elimination (RF-RFE) scheme is used to eradicate redundant predictor features to improve the proposed CMLP model efficiency. The CMLP model has been compared and bench-marked against seven artificial intelligence–based and seven temperature-based deterministic models, showing excellent performance at all solar energy study sites tested over daily, monthly, and seasonal scales. The proposed hybrid CMLP model should be explored as a viable modelling tool for solar energy monitoring and forecasting in real-time energy management systems.

Keywords Deep learning hybrid models · Convolutional neural network · Multi-layer perceptrons · Solar radiation prediction · Renewable energy · Global climate models

✉ David Casillas-Pérez
david.casillas@urjc.es

Sujan Ghimire
sujan.ghimire@usq.edu.au

Thong Nguyen-Huy
thong.nguyen-huy@usq.edu.au

Ramendra Prasad
ramendrap@unifiji.ac.fj

Ravinesh C. Deo
ravinesh.deo@usq.edu.au

Sancho Salcedo-Sanz
sancho.salcedo@uah.es

Binayak Bhandari
binayak@sis.ac.kr

[1] School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield 4300, QLD, Australia

[2] SQNNSW Drought Resilience Adoption and Innovation Hub, University of Southern Queensland, Toowoomba 4350, QLD, Australia

[3] Centre for Applied Climate Sciences, University of Southern Queensland, Toowoomba 4350, QLD, Australia

[4] Department of Science, School of Science and Technology, The University of Fiji, Saweni, Lautoka, Fiji

[5] Department of Signal Processing and Communications, Universidad Rey Juan Carlos, Fuenlabrada 28942, Madrid, Spain

[6] Department of Signal Processing and Communications, Universidad de Alcalá, Alcalá de Henares 28805, Madrid, Spain

[7] Department of Railroad Engineering, Woosong University, Dongdaejeon-ro 171, Daejeon, Republic of Korea

Introduction

United Nations Member States provided visionary blueprints for clean energy technologies in 2015, through the 2030 Agenda for Sustainable Development. Its key agenda item was SDG-7, which aims to ensure developed and developing countries' accessibility of affordable, reliable, sustainable, and modern energy through global partnerships. This vision was recently reinforced by the 2021 United Nations Conference of Parties (COP 26), where the attempt to limit the global warming to 1.5° reinforces SDG-7 (increasing the penetration of clean energy, particularly solar and wind, into national electricity grids).

Similarly, Australia is also striving to transition from fossil fuel–based electricity generations to renewable resources, with new monitoring and forecasting technologies to support battery storage systems, which ensure consistent availability of electricity from these variable renewable energy sources. It has been demonstrated that solar photovoltaic (PV) is one of the cheapest renewable sources with competitive levelized cost of electricity. Australia's rooftop solar PV installations has expanded recently with the addition of 333,978 new installations, adding a generation capacity of 2.6 GW to the national electricity grid in 2020. This was almost 18% higher than the previous year (2.2 GW and 284,000 installations in 2019) [1]. However, electricity generation from solar PV is intermittent in nature, and may be affected by short- and long-term synoptic weather conditions [2], the solar variability, and the intermittency of transient clouds and aerosols. As a result, the variability in the availability of *GSR* on the surface of PV collectors leads to solar electricity supply instability and issues with frequency response, requiring reactive back-up power generations [3]. With over 330,000 rooftop installations, feeding electricity into the grid at different points in time and with varied capacity, a robust predictive system for *GSR* is a necessity for grid stability. Prior knowledge of *GSR* within its corresponding solar power outputs and inconsistency of supply to meet changing demands can therefore alleviate future risk of supply from renewable energies. This includes solar and wind tecnologies which are subjected to increased curtailment-driven weather and climatic conditions at the concerned sites. In this respect, the development of reliable *GSR* prediction methods, with potential further applications in areas such as wind and wave energy systems, is of vital importance to ensure sustained growth of renewable energy sector.

The traditional *GSR* prediction models, including physical and statistical methods, largely assumed a linear relationship among predictive and target variables [4, 5]. The physical methods utilize numerical weather forecasts (NWP) and adopt satellite data to predict the *GSR*, while the statistical methods are built upon mathematical relationships between predictors (weather data) and target (*GSR*) (such as an exponential smoothing, auto-regressive integrated moving average (ARIMA), or Markov Chains). Therefore, robust and advanced modelling approaches such artificial intelligence (AI) techniques and hybrid methods [6, 7] are necessary to capture the non-linear interactions between predictors and the target. AI-based methods using machine/deep learning (ML/DL) methods allow modelling non-linear relationships between predictors and the target [8, 9] yielding a superior performance in many applications. Several AI-based models have outperformed both physical and statistical models [10, 11] as an alternative methodology in different renewable energy modelling contexts. Following this, the neural networks have been extensively used for solar radiation prediction [12, 13]. Literature reveals a number of algorithms being proposed to predict hourly *GSR*, such as the one in [14], which included adaptive neuro-fuzzy inference system (ANFIS), feed-forward neural network (FFNN), Elman recurrent network (LM), and radial basis function (RBF). Among these, the LM method outperformed the others, and it seems that the inclusion of wind direction as an input led to improved performance of the LM model. Yacef et al. [15] developed and tested an artificial neural network (ANN), Bayesian neural network (BNN), and the Angstrom-Prescot deterministic model for daily *GSR* prediction with dry-bulb air temperature, sunshine duration, relative humidity, and extra-terrestrial irradiation meteorological data over 4 years as predictors. The results showed that the BNN model outperformed the ANN deterministic models. Zou et al. [16] used daily meteorological parameters (precipitation, temperature (minimum, maximum, and mean), sunshine duration hours, water vapor pressure, relative humidity, air pressure, and wind speed) along with ground-based measurements and a daily surface dataset from the Data Assimilation and Modeling Center for Tibetan Multi-spheres (DAM) to predict daily *GSR* and validate the model. The results showed that an ANN method with 9-17-1 input-hidden-output architecture yields a better accuracy than deterministic methods. Lu et al. [17] applied an ANN model to study non-linear relationships between observed *GSR* and spectral information from the Multi-functional Transport Satellite (MTSAT) concluding that the ANN model was a fast, efficient, and accurate method. In spite of these studies supporting the AI methods in *GSR* prediction problems, their conventional learning algorithms could be subjected to a limited ability to deal with relatively complex and stochastic behavior of atmospheric variables, and its inability to extract maximal features in large datasets that are required to attain a superior performance. An improved

form of "learning" model, the deep learning (DL) methods, has recently been recognized as an alternative to the conventional ML models. DL models are simple yet nonlinear models with multiple hidden layers that convert the representation of features into a higher and more abstract level [18]. Having a powerful nonlinear network structure, DL provides a performance edge over traditional learning models that have been challenged for their optimal feature representation [18, 19]. DL improves the prediction by using deep nonlinear network structures, and realizing complex function approximation and deep features from massive data samples [20, 21]. While the initial applications included image processing [22], natural language processing [23], speech processing [24], or medicine [25], DL has also been attractive for time series applications, with utilization of convolutional neural network (CNN) combined with long-short term memory (LSTM) for different prediction problems, including renewable energy applications such as wind speed [26] and also *GSR* [27]. Additionally, in classification problems, the CNN-based model known as RipNet has been used recently for automated rip current identification. This RipNet model has achieved high sensitivity with very few false-positive classifications, thus making it a very practical real-world approach [28].

On the other hand, the hybrid models emerge from integrating different approaches, such as combination of two or more AI approaches to obtain more accurate predictions. These hybrid models combine the advantages of both AI-based models, in such a way that the final model has better generalization performance. Hybrid approaches have been utilized in a number of prediction problems [29]. For example, in *GSR* prediction problems, the following approaches have been noted in literature: particle swarm optimization-extreme learning nachines (PSO-ELM) [30], support vector regression optimized by PSO, bat and whale algorithms [31], maximum overlap discrete wavelet transform (DWT) integrated with SVR (PSO-W-SVR) [32], ANN optimized by PSO (PSO-ANN) [33], CNN integrated with SVR (CNN-SVR) [34], CNN integrated with stacked regression (CNN-REGST) [35], or CNN integrated with extreme gradient boosting and random forest regression (CXGBRFR) [36]. In time series applications in energy, CNN-LSTM models were developed for prediction of residential energy consumption [37], wind speed prediction [38], and *GSR* prediction [39].

One major difficulty in *GSR* prediction with DL methods is the precise requirement for predictor variables to be intrinsically linked to the target. In the context of *GSR*, most significant meteorological and climatological predictors may not be accessible due to remoteness of solar farms, compounded by the expense and maintenance issues associated with solar measuring and monitoring equipment. Fortunately, a global climate models (GCMs) dataset, which are gridded representations of physical variables of a meteorological or climatological origin, is an important resource for solar and wind energy predictive modelling. GCM datasets have been used in crop evapotranspiration [40], rainfall [41], streamflow [42], drought [43], and wind speed [44] forecasting problems. To improve the accuracy of daily *GSR* prediction, a new hybrid model is developed by integrating CNN with multi-layer perceptron model (CMLP) algorithm. The advantage of the superior predictive skills of CNN in extracting detailed spatial features is combined with the superior skills of MLP in extracting time-series features from input-target data matrix while achieving better predictive ability.

The objectives of this study, which define somehow its novelty, are as follows: *(i)* To adopt GCM datasets coupled with available ground-based observation datasets from the most reliable source in the study region as potential variables employed to predict daily *GSR*. *(ii)* To incorporate hybrid-wrapper feature selection, which utilizes a random forest-recursive feature elimination (RF-RFE) approach to filter redundant variables for improved efficacy of the proposed model. *(iii)* To develop a suite of seven AI-based models (i.e., DNN, ANN, ELM, KRR, XGBOOST, BRF, WKNN) and seven temperature-based deterministic models (i.e., TMGO, TMAL, TMAN, TMHAS, TMBC, TMCH, TMHG) as the competing approaches to benchmark the proposed hybrid CMLP model. It is expected that the use of both the AI-based and the temperature-based models will lead to a detailed investigation of the efficacy of the proposed CMLP model. *(iv)* To compare the overall skills of the CMLP model in daily *GSR* prediction by pooling in a total of 123 initial predictive variables across six solar energy sites in Queensland state, Australia. The methodology produced to predict *GSR*, therefore, presents a multi-objective approach, integrating global climate model simulations and hybrid feature mapping methods that are benchmarked against AI-based and deterministic methods used as popular standalone approaches for solar energy predictions.

The rest of this paper has been structured in the following form: the next section introduces the study area and the available dataset, including the information of all the solar farms considered, predictive variables, and objective variables. The "Methods, Study Area, and Datasets" section introduces the most important characteristics of the methods considered in this work, including a brief description of temperature-based approaches, random forest-recursive feature elimination algorithm, convolutionary neural network, and multi-layer percetron models. The "Development of the Proposed CMLP Predictive Model" section explains the proposed hybrid prediction CMLP model, and the "Results" and "Discussion"

sections present the results obtained by the proposed CMLP approach, and a comparison with state-of-the art AI-based and temperature-based algorithms for GSR prediction. Finally, the "Conclusions, Limitations, and Future Research Work" section ends the paper by some key conclusions and remarks on the research implemented.

Methods, Study Area, and Datasets

For the objective hybrid CMLP model, the predictor variables are selected from GCM-based historical and ground data, via random forest-recursive feature elimination (RF-RFE) algorithm, prior to being fed into the CNN algorithm.

Figure 3 illustrates the topological structure of the proposed hybrid CMLP model. The features from the selected predictor variables are extracted by the convolution kernels of the CNN, which are then fed into the multilayer perceptron (MLP) network for the final *GSR* predictions. The CMLP model has been compared with AI-based methods that include ANN, ELM, deep neural networks (DNN), boosting random forest (BRF), kernel ridge regression (KRR), weighted K-nearest neighbor (WKNNR), extreme gradient boosting (XGB), and temperature-based deterministic models that include Hargreaves and Samani (TMHS), Chen (TMCH), Bristow and Campbell (TMBC), Goodin (TMGO), Allen (TMAL), Annandale (TMAN), and Hassan (TMHG) methods. The theoretical background of standalone models (ANN, ELM, DNN, BRF, KRR, WKNNR, and XGB) is described elsewhere [45–53] and derivations for temperature-based deterministic models (TMHS, TMCH, TMBC, TMGO, TMAL, TMAN, and TMHG) are in earlier studies (e.g., [54–56]), so we do not reintroduce them here. However, the primary equations of TM-based deterministic models are as follows:

$$\text{TMHS}: \ GSR = a\Delta T^{b} H_0 \tag{1}$$

$$\text{TMCH}: \ GSR = H_0 \left[a + b \cdot \ln(\Delta T)\right] \tag{2}$$

$$\text{TMBC}: \ GSR = a \left[1 - \exp(-b\Delta T^{c})\right] H_0 \tag{3}$$

$$\text{TMAL}: \ GSR = \mathrm{a}\Delta T^{0.5} H_0 \tag{4}$$

$$\text{TMAL}: \ GSR = a \left[1 - \exp\left(-b\frac{\Delta T^{c}}{G_0}\right)\right] H_0 \tag{5}$$

$$\text{TMHG}: \ GSR = (aT^{b} H_0 + c) H_0 \tag{6}$$

where:

$$\Delta T = T_{\max} - T_{\min} \tag{7}$$

Random Forest-Recursive Feature Elimination

The random forest-recursive feature elimination (RF-RFE) is a hybrid-wrapper type of feature selection method [57] used in this study to optimize the proposed hybrid CMLP model. The RF-RFE initially trains a random forest (RF) model with all predictor variables $(x_{i,1}, x_{i,2}, \ldots, x_{i,p})$, $(i = 1, 2, \ldots, m)$ and then ranks the predictor variables based on their relative importance. The RF model performance is then evaluated based on a diverse list of metrics, repeated with a progressively smaller subset, decreased by *d* features. The best performance model is selected and the combination of predictor variables is determined. In addition, the relative root mean square error (RRMSE) metric is used to evaluate the performance of RF-RFE model. A pseudo-code as in Fig. 1 illustrates the RF algorithm used in improving the overall performance of the proposed CMLP model.

Convolutional Neural Network (CNN)

The proposed hybrid CMLP model is based on a convolutional neural network (CNN) method, which is a feedforward network developed by [58] typically for 2-D image processing. The CNN is constructed by cascading layers comprising a convolutional layer, pooling layer, and fully connected layer [59]. Fundamentally, the CNN technique builds numerous filters to extract hidden features by convoluting and pooling input information, layer by layer, and the derived features are integrated into a fully connected layer while the regression problem is handled by activation functions. In 1-D domain, a filtering kernel plays as a feature extractor with one-dimensional convolution described in [60]:

$$\begin{aligned} a_j^{(l+1)}(\tau) &= \sigma\left(b_j^l + \sum_{f=1}^{F^l} K_{jf}^l(\tau) * a_f^l(\tau)\right) \\ &= \sigma\left(b_j^l + \sum_{f=1}^{F^l}\left[\sum_{p=1}^{p^l} K_{jf}^l(p) a_f^l(\tau - p)\right]\right) \end{aligned} \tag{8}$$

where $a_j^{(l+1)}$ refers to the feature map j in layer $l+1$, σ refers to a nonlinear function, F^l indicates the number of feature

	Algorithm 1: Recursive Feature Elimination (RFE)
1	Train the random forest model with full feature set
2	Evaluate the model performance with RMSE and rank feature importance
3	**For** *i=1 to n*, **do**:
4	Eliminate last *d* features with smallest importance
5	Train the random forest model with tuned subset
6	Evaluate the model performance with RMSE and rank feature importance
7	**End** For
8	Select the optimal feature length and it's feature rank
9	**End** Algorithm

Fig. 1 Random forest-recursive feature elimination (RF-RFE) algorithm

maps in layer$l$, K_{jf}^{l} represents the kernel convolved over feature map f in layer l to build the feature map j in layer $l+1$, and p^{l} denotes the length of kernels in layer l with b_{j}^{l} being a bias vector.

Multi-Layer Perceptron Model

The proposed hybrid CMLP model also comprises of the MLP model constructed by cascading input and hidden and output layers. Furthermore, the learning comes from error-based backpropagation using gradient descent approach with neurons as the main component of an MLP found in network's layers, linked by synaptic weights and their nonlinear modelling behaviors.

Figure 3 shows the schematics and the MLP can be described mathematically as,

$$u_k = \sum_{j=1}^{m} w_{kj} x_j \tag{9}$$

$$y_k = \varphi\left(u_k + b_k\right) \tag{10}$$

where each neuron input x_j is multiplied by a synaptic weight w_{kj} expressing the connection strength between neurons. The results, u_k, are added to bias b_k and that sum is then multiplied by activation function φ to compute the output y_k used as input with b_k and synaptic weight w_{kj} fitted in the training step to predict the output. Further information on MLP can be found in [61].

Study Area and Dataset

To implement the proposed hybrid CMLP model, historical data from 1950 to 2006 from six solar farms across Queensland state, Australia, are selected (Table 1; Fig. 2). These solar farms have energy generation potentials ranging between 19.9 and 1200 MW, yet the *GSR* observations are not publicly available. Hence, the meteorological variables are acquired from the Scientific Information for Land Owners (SILO) database operated by the Department of Environment and Science [62]. The study also used CMIP5 GCM (Coupled Model Intercomparision Phase-5) outputs [63] that include ACCESS1-0 (CSIRO-BOM) [64], Hadley-GEM2-CC (MOHC) [65], and MRI-CGCM3 (MRI) [66] models. The longitude, latitude, time, and atmospheric pressure at eight levels, or near-surface observations, are used to index the historical variables (Table 2).

Development of the Proposed CMLP Predictive Model

To prevent the influence of extremely large or small predictors and target ranges, all input data are normalized before training, using the min-max method. The overall goal is to modify the absolute values of numeric columns in data to a common scale without distorting differences in the ranges of values and transforming the features to be on a similar (relative) scale. This was achieved by applying Eq. (11):

$$X_N = \frac{X - X_{MIN}}{X_{MAX} - X_{MIN}} \tag{11}$$

where X_N refers to the normalized value of the actual value of variable X, and X_{MAX} and X_{MIN} mean the maximum and minimum value, respectively.

Then, feature selection is employed to eliminate redundant features from a total pool of 123 GCM-based and SILO-based atmospheric variables (see Table 2). For this,

Table 1 Statistical description of the target variable, i.e., daily global solar radiation (*GSR*; $MJm^{-2}day^{-1}$) in the six solar farms over Queensland state, Australia

Statistical property	Kidston	Brigalow	Bulli Creek	Baking Board	Chewco	Blackwater
Latitude	18.880Â°S	27 42'S	28 01'S	26 43Â°S	17.047Â°S	23.597Â°S
Longitude	144.151Â°E	151 34'E	150 52'E	150 33'E	145.384Â°E	148.753Â°E
Altitude (m)	521.00	438.00	348.00	337.00	400.00	184.00
Capacity (MW)	50	34.5	1200	19.9	75	150
Median	21.00	20.00	20.00	20.00	20.00	21.00
Mean (MJm^{-2})	20.72	19.60	20.07	20.06	19.97	20.45
Standard deviation	4.53	6.33	6.39	6.33	6.29	5.81
Variance	20.56	40.02	40.88	40.12	39.58	33.80
Maximum	30.00	32.00	32.00	32.00	32.00	32.00
Minimum	7.00	4.00	4.00	4.00	4.00	4.00
Mode	26.00	17.00	16.00	18.00	18.00	18.00
Interquartile range	6.75	10.00	9.00	9.00	9.00	8.00
Skewness	−0.34	−0.13	−0.15	−0.23	−0.22	−0.43
Kurtosis	2.59	2.35	2.39	2.38	2.39	2.85

Table 2 Descriptive information of predictor variables used for daily *GSR* prediction with output variables from global climate model and ground-based observations from scientific information for landowners (SILO) repository

Variable	Description	Units
Global climate model atmospheric predictor variables		
clt	Cloud area fraction	%
hfls	Surface upward latent heat flux	Wm^{-2}
hfss	Surface upward sensible heat flux	Wm^{-2}
hur	Relative humidity	%
hus	Near-surface specific humidity	gkg^{-1}
pr	Precipitation	$kgm^{-2}s^{-1}$
prc	Convective precipitation	$kgm^{-2}s^{-1}$
prsn	Solid precipitation	$kgm^{-2}s^{-1}$
psl	Sea level pressure	pa
rhs	Near-surface relative humidity	%
rhsmax	Surface daily max relative humidity	%
rhsmin	Surface daily min relative humidity	%
sfcWind	Wind speed	ms^{-1}
sfcWindmax	Daily maximum near-surface wind speed	ms^{-1}
ta	Air temperature	K
tas	Near-surface air temperature	K
tasmax	Daily max near-surface air temperature	K
tasmin	Daily min near-surface air temperature	K
ua	Eastward wind	ms^{-1}
uas	Eastern near-surface wind	ms^{-1}
va	Northward wind	ms^{-1}
vas	Northern near-surface wind	ms^{-1}
wap	Omega (Lagrangian tendency of air pressure)	pas^{-1}
zg	Geopotential height	m
Ground-observed variables from SILO		
Tmax	Maximum temperature	K
Tmin	Minimum temperature	K
Rain	Daily rainfall	mm
Evap	Evaporation	mm
VP	Vapor pressure	Pa
RHmaxT	Relative humidity computed at maximum temperature	%
RHminT	Relative humidity computed at minimum temperature	%

the normalized data are processed by means of an RF-RFE feature selection method that selected 65 input variables for the Kidston Solar Project, 66 variables for Brigalow and Bulli Creek solar farm, 46 variables for Baking Board solar farm, and 56 variables for Chewco solar farm. A full list of selected input variables for daily *GSR* prediction is displayed in Table 3. Evidently, the cloud area fraction (clt), precipitation (pr), convective precipitation (prc), near surface air temperature (tas), rainfall (rain), relative humidity, and wind speed are the most important parameters for daily *GSR* predictions.

The inputs based on the RF-RFE selected variables and the target (i.e., daily *GSR*) data matrix are created for predictive modelling purposes. As there is no predefined rule for the data division [12, 50, 67], in this work, we used 54 years of dataset for training (a total of 20,089 data points) of which 20% data (4018 data points) were used for validation, and 1 year data (365 data points) for testing with response data being the daily *GSR* time-series values in an independent test set.

This study has therefore developed a 4-layered CMLP model as shown in Fig. 3, where the first three layers include a 1-D convolution layer to extract data features firstly. The fourth layer in this hybrid approach consists of an MLP-based predictive stage to analyze features and predict the *GSR*. A summarized CMLP algorithm involves the following eight steps:

- Input the data from the prepared dataset after feature selection, and convert the signal vector into a matrix.

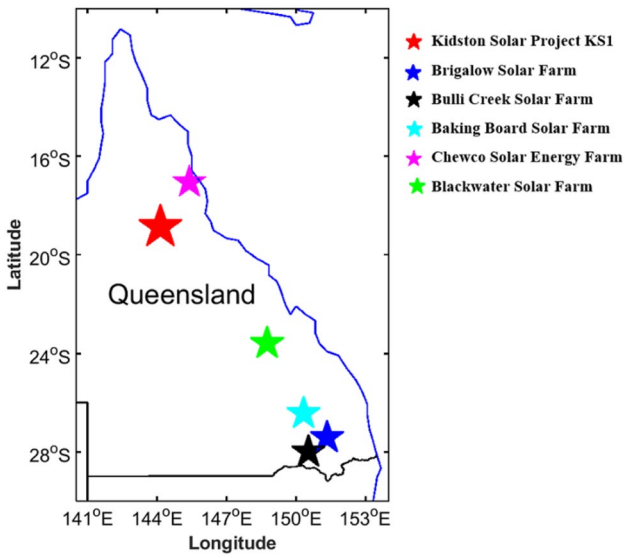

Fig. 2 Six solar energy farms across Queensland, Australia, where the proposed hybrid CMLP model is implemented to predict global solar radiation (*GSR*, $MJm^{-2}day^{-1}$) using global climate model outputs and ground-based measured predictor variables

- Initialize the parameters, such as the total numbers of layers, the maximum epoch of iteration, the learning rate, the optimizer, and the number of filters.
- Implement convolution and pooling.
- Fine-tune the stacked model using the optimized algorithm.
- Finish the optimization and acquire the features of the dataset.
- Use the extracted features for regression by MLP.
- Initialize the MLP parameters, such as the total numbers of neurons, the maximum epoch of iteration, the learning rate, the optimizer, and the activation function.
- Get the final optimized CMLP model and predict on test dataset.

Hyperparameters of the CMLP objective model and all AI-based models benchmark are selected using a Bayesian optimization method (HyperOpt), which comes from an open-source Python library introduced by Komer et al. [68]. It involves Bayesian optimization tree-structured Parzen estimators (BO-TPE) [69, 70] so unlike traditional hyperparameter search algorithms (Grid and Random search) that are only applicable to low dimensional space and may fail to provide optimal solutions, the BO-TPE uses results from previous iteration to choose the next set [71]. The BO-TPE method requires fewer iterations relative to traditional approaches and provides optimal hyperparameter combination [72].

Tables 4 and 5 present the search space for CMLP and AI-based benchmark model hyperparameters, respectively, where the Hyperopt provides a search in all possible combinations of hyperparameters to select their optimal values.

Except for the output layer of CNN, to improve the learning speed and overcome vanishing gradient problems, the rectified linear unit (relu) was used as the activation function [73, 74]. An adaptive moment estimation (Adam) optimizer [75] with a learning rate of 0.001 was used for both CNN and DNN models. The Adam method is an efficient stochastic optimization technique that only needs first-order gradients with less memory requirement [76], and has combined advantage of two popular methods: adaptive gradient algorithm (AdaGrad) [77] that works well with sparse gradients and root mean square propagation algorithm (RMSProp) [78] that has an excellent performance in non-linear models.

All predictive models are built using Keras 2.2.4 [79, 80] on TensorFlow 1.13.1 [81] with backends in Python 3.6. The Hyperopt-sklearn [82] library is used for

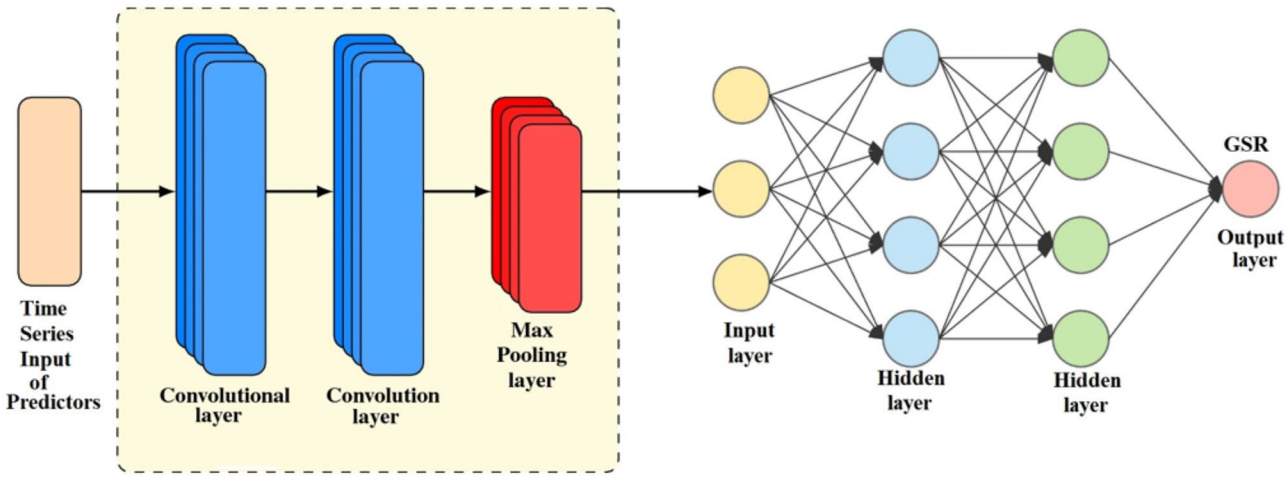

Fig. 3 Topological structure of feature extraction algorithm (convolutional neural network, CNN) that has been integrated with the objective predictive algorithm (multilayer perceptron, MLP) in this study used to construct the hybrid CMLP model for daily *GSR* prediction

Table 3 The hybrid CMLP model's input variables selected using the random forest-recursive feature elimination (i.e., RF-RFE) feature selection process. For a complete description of input variables, see Table 2

Kidston	Brigalow	Bulli Creek	Baking Board	Chewco Solar	Blackwater
clt	clt	clt	clt	clt	clt
hfls	hfls	hfls	hfls	hfls	hfls
hfss	hfss	hfss	hfss	hfss	hfss
huss	huss	huss	huss	huss	huss
prc	pr	pr	pr	prc	prc
psl	prc	prc	prsn	prsn	prsn
rhs	prsn	prsn	psl	psl	psl
rhsmax	psl	psl	rhsmax	rhsmin	rhsmin
rhsmin	rhsmax	rhsmax	rhsmin	tas	tas
tas	tas	tas	tasmax	tasmin	hur_1000
tasmax	tasmax	tasmax	hur_10000	hur_1000	hur_5000
tasmin	tasmin	tasmin	hur_50000	hur_5000	hur_50000
hur_1000	hur_1000	hur_1000	hur_85000	hur_50000	hus_1000
hur_5000	hur_5000	hur_5000	hus_1000	hus_1000	hus_10000
hur_25000	hur_25000	hur_25000	hus_5000	hus_5000	hus_50000
hur_50000	hur_50000	hur_50000	hus_10000	hus_10000	hus_85000
hur_85000	hur_70000	hur_70000	hus_50000	hus_50000	sfcWindmax
hus_1000	hur_85000	hur_85000	hus_70000	hus_70000	ta_1000
hus_5000	hus_1000	hus_1000	sfcWindmax	hus_85000	ta_5000
hus_10000	hus_5000	hus_5000	ta_1000	sfcWindmax	ta_25000
hus_25000	hus_10000	hus_10000	ta_5000	ta_1000	ta_50000
hus_50000	hus_25000	hus_25000	ta_10000	ta_5000	ta_85000
hus_70000	hus_50000	hus_50000	ta_25000	ta_10000	ua_1000
hus_85000	hus_70000	hus_70000	ta_85000	ta_25000	ua_5000
sfcWind	hus_85000	hus_85000	ua_1000	ta_50000	ua_10000
sfcWindmax	sfcWind	sfcWind	ua_10000	ta_70000	ua_70000
ta_1000	sfcWindmax	sfcWindmax	ua_50000	ta_85000	ua_85000
ta_5000	ta_1000	ta_1000	ua_70000	ua_1000	va_50000
ta_10000	ta_5000	ta_5000	ua_85000	ua_5000	va_70000
ta_25000	ta_10000	ta_10000	uas	ua_10000	va_85000
ta_50000	ta_25000	ta_25000	va_5000	ua_50000	wap_1000
ta_70000	ta_50000	ta_50000	va_50000	ua_70000	wap_5000
ta_85000	ta_70000	ta_70000	va_70000	ua_85000	wap_25000
ua_1000	ta_85000	ta_85000	va_85000	va_10000	wap_50000
ua_5000	ua_1000	ua_1000	wap_5000	va_50000	wap_85000
ua_10000	ua_5000	ua_5000	wap_10000	va_70000	zg_1000
ua_25000	ua_10000	ua_10000	wap_25000	va_85000	zg_5000
ua_50000	ua_25000	ua_25000	wap_50000	wap_1000	zg_25000
ua_70000	ua_50000	ua_50000	zg_25000	wap_5000	zg_50000
ua_85000	ua_70000	ua_70000	Tmax	wap_25000	zg_85000
uas	ua_85000	ua_85000	Tmin	wap_50000	Tmax
va_25000	va_1000	va_1000	Rain	wap_85000	Tmin
va_50000	va_5000	va_5000	Evap	zg_1000	Rain
va_70000	va_10000	va_10000	VP	zg_5000	Evap
va_85000	va_25000	va_25000	RHmaxT	zg_10000	VP
wap_1000	va_70000	va_70000	RHminT	zg_25000	RHmaxT
wap_5000	va_85000	va_85000	-	zg_50000	RHminT
wap_10000	wap_1000	wap_1000	-	zg_70000	-
wap_25000	wap_5000	wap_5000	-	zg_85000	-
wap_50000	wap_10000	wap_10000	-	Tmax	-
wap_70000	wap_25000	wap_25000	-	Tmin	-

Table 3 (continued)

Kidston	Brigalow	Bulli Creek	Baking Board	Chewco Solar	Blackwater
wap_85000	wap_50000	wap_50000	-	Rain	-
zg_1000	wap_70000	wap_70000	-	Evap	-
zg_5000	wap_85000	wap_85000	-	VP	-
zg_10000	zg_1000	zg_1000	-	RHmaxT	-
zg_25000	zg_5000	zg_5000	-	RHminT	-
zg_50000	zg_10000	zg_10000	-	-	-
zg_70000	zg_25000	zg_25000	-	-	-
zg_85000	zg_50000	zg_50000	-	-	
Tmax	zg_70000	zg_70000	-	-	-
Tmin	zg_85000	zg_85000	-	-	-
Rain	Tmax	Tmax	-	-	-
Evap	Tmin	Tmin	-	-	-
VP	Rain	Rain	-	-	-
RHmaxT	VP	VP	-	-	-
-	RHminT	RHminT	-	-	-

hyperparameter optimization and the training process is conducted on a computer system with Intel© Core *i*7 central processing unit (CPU) with 32GB random access memory.

Evaluation of Model Performance

To evaluate the CMLP model, the predicted *GSR* are compared with observations and quantitatively evaluated with correlation coefficient (*r*), root mean square error (RMSE), mean absolute error (MAE), mean bias error (MBE), relative root mean square error (RRMSE), mean absolute percentage error (MAPE), Willmott's index (WI), Nash-Sutcliffe equation (NSE), RMSE ratio ($RMSE_{ss}$), and skill score (SS). Next, we describe the characteristics of these performance metrics, together with their mathematical formulation [12, 13, 32, 50, 83–89].

$$r = \frac{\sum_{i=1}^{n}(GSR^m - \langle GSR^m \rangle)(GSR^p - \langle GSR^p \rangle)}{\sqrt{\sum_{i=1}^{n}(GSR^m - \langle GSR^m \rangle)^2}\sqrt{\sum_{i=1}^{n}(GSR^p - \langle GSR^p \rangle)^2}} \tag{12}$$

Table 4 Architecture of deep learning model: convolutional neural networks integrated with multi-layer perceptron (i.e., CMLP) and the deep neural network (DNN) model

Model	Hyperparameters	Range	Kidston	Brigalow	Bulli Creek	Baking Board	Chewco	Blackwater
	Filter1	[50, 80, 100, 200]	100	80	200	80	100	200
	Filter 2	[40, 50, 60, 70, 80]	50	40	60	50	70	40
	Filter 3	[5, 10, 15, 20, 25, 30]	20	25	20	30	15	20
	MLP hidden layer size	[40, 50, 60, 100, 150]	60	100	60	60	50	100
	Epochs	[300, 400, 700, 800, 1000]	400	300	700	400	300	800
	MLP activation function	["purelin," "logistic," "tanh," "relu"]	relu	relu	tanh	relu	tanh	relu
	Solver	["lbfgs," "sgd," "adam"]	lbfgs	lbfgs	lbfgs	lbfgs	lbfgs	lbfgs
CMLP	Batch Size	[5, 10, 15, 20, 25, 30, 40, 50]	10	5	15	10	5	20
	Hiddenneuron 1	[100, 200, 300, 400, 50]	200	100	200	300	200	100
	Hiddenneuron 2	[20, 30, 40, 50, 60, 70]	40	60	40	60	70	40
	Hiddenneuron 3	[10, 20, 30, 40, 50]	20	50	50	20	50	50
	Hiddenneuron 4	[5, 6, 7, 8, 12, 15, 18]	12	18	12	18	15	12
	Activation function	relu						
	Epochs	[300, 400, 700, 800, 1000]	300	400	700	400	700	400
DNN	Batch size	[5, 10, 15, 20, 25, 30, 40, 50]	15	20	25	30	40	50

Table 5 The conventional models used to benchmark the proposed hybrid CMLP model. Benchmark models are ANN, ELM, and KRR methods including ensemble models (BRF, XGBOOST, and WKNNR). Note that ReLU stands for rectified linear units with tansig, logsig, and purelin being hyperbolic tangent transfer function, log-sigmoid transfer function, and linear transfer functions, respectively, and LM, lbfgs, and cgf being the Levenberg-Marquardt, limited memory Broyden-Fletcher-Goldfarb-Shanno, and conjugate gradient backpropagation with Fletcher-Reeves, respectively

Model	Hyperparameters	Selection range	Kidston	Brigalow	Bulli Creek	Baking Board	Chewco	Blackwater
	Kernel	Gaussian	10	30	20	20	30	10
	Cost function	[0.001, 0.01, 0.1, 1, 10, 100]	0.1	0.1	0.01	0.1	0.1	0.001
KRR	Penalty function	[0.001, 0.01, 0.1, 1, 10, 100]	100	10	10	100	10	100
	Hiddenneuron	[20, 30, 40, 50]	40	30	50	20	30	40
ELM	Activation function	[logistic, tanh]	tanh	tanh	tanh	tanh	tanh	tanh
	Hiddenneuron	[10, 20, 30, 40, 60, 80, 100, 200, 300]	100	100	80	80	60	80
	Backpropagation algorithm	["trainlm," trainbfg, traincgf]	trainlm	trainlm	trainbfg	trainlm	trainbfg	trainbfg
ANN	Activation function	[tansig, logsig, purlin]	tansig	tansig	tansig	tansig	tansig	tansig
	Number of neighbors	[5, 10, 20, 30, 50, 100]	20	10	10	20	10	20
	Algorithm used to compute the nearest neighbors	["auto," "ball_tree," "kd_tree," "brute"]	auto	auto	auto	auto	auto	auto
WKNNR	Leaf size passed to BallTree or KDTree	[10, 20, 30, 50, 60, 70]	10	10	20	10	20	10
	Learning rate	[0.01, 0.1, 0.001, 0.005]	0.01	0.001	0.01	0.01	0.01	0.01
	Maximum depth of the individual regression estimators	[5, 8, 10, 20, 25]	8	10	10	10	10	8
	Number of boosting stages to perform	[50, 100, 150, 200]	100	100	150	100	150	100
	Minimum number of samples to split an internal node	[20]						
XGBOOST	Number of features for best split	["auto," "sqrt," "log2"]	auto	auto	auto	auto	auto	auto
	The maximum depth of the tree.	[5, 8, 10, 20, 25]	8	10	10	25	10	8
	The number of trees in the forest.	[50, 100, 150, 200]	100	150	150	100	150	100
	Minimum number of samples to split an internal node	[2, 4, 6, 8, 10]	8	10	8	6	10	8
BRF	The number of features to consider when looking for the best split.	["auto," "sqrt," "log2"]	auto	auto	auto	auto	auto	auto

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(GSR^p - GSR^m)^2} \quad (13)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|GSR^p - GSR^m| \quad (14)$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(GSR^p - GSR^m) \quad (15)$$

$$RRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(GSR^p - GSR^m)^2}}{\langle GSR^m \rangle} \quad (16)$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{|(GSR^m - GSR^p)| \times 100}{GSR^p}\right) \tag{17}$$

$$WI = 1 - \frac{\sum_{i=n}^{n}(GSR^m - GSR^p)^2}{\sum_{i=n}^{n}(|GSR^p - \langle GSR^m\rangle| + |GSR^m - \langle GSR^m\rangle|)^2} \tag{18}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(GSR^m - GSR^p)^2}{\sum_{i=1}^{n}(GSR^m - \langle GSR^m\rangle)^2} \tag{19}$$

$$RMSE_{ss} = \frac{RMSE(p,x)}{RMSE(r,x)} \tag{20}$$

$$SS = 1 - \frac{RMSE(p,x)}{RMSE(r,x)} \tag{21}$$

where GSR^m and GSR^p are the observations and predicted GSR, $\langle GSR^m\rangle$ and $\langle GSR^p\rangle$ are the averages of GSR^m and GSR^p, n is the total number of data points, p stands for model prediction, and x denotes the observations.

The physical interpretation of all performance metrics (Eqs. 12–22) have been discussed in earlier works. For example, one may refer to references [12, 20, 32, 50, 84] for specific solar radiation prediction problems or references [90, 91] for explanation of their characteristics. In summary, the characteristics of the key metrics are as follows:

- The magnitude of r, which measures degree of covariance between the observed and forecasted GSR, should ideally be close to unity, whereas $RMSE$ and MAE including the MBE, which evaluates the overall predicted error in absolute units ($MJm^{-2}day^{-1}$) in testing phase, should both be trivial for the best performing model.
- Reformulated from Eqs. (13–15), the magnitudes of $RRMSE$ and $MAPE$ are the relative (%) representations used to evaluate the proposed CMLP model for geographically diverse study sites. These metrics can enable a comparison of model performance at many different sites as they consider the influence of climatic factors.
- Additional measures, WI and NSE, are non-dimensional (unit-less) metrics where values close to unity are expected to indicate a greater efficiency of the model. The NSE provided a distinct advantage to record the level of agreement between the observed and the predicted GSR while being sensitive to differences in the observed and the forecasted means and variances. By contrast, the WI considered the ratio of the mean square error instead of the square of differences, and therefore, could be preferred over the NSE value.

For checking out the level of congruence between the observed and the predicted solar radiation, we use pr to refer to a perfect prediction for a persistence-based model, and r corresponding to a reference prediction. The persistence model, used more commonly in any GSR prediction problem, considers that the solar radiation at a next time-scale, $t+1$ is described the behavior of solar radiation at t under the assumption of stationary atmosphere, or clear sky conditions. Furthermore, SS metric is applied to compare the predictive capability of the developed model against a persistence model and MBE is used to capture the average bias in the predictions.

In addition to standard metrics (Eqs. 12–22), we have also employed GPI as follows:

$$GPI_i = \sum_{j=1}^{6}\alpha_j(g_j - y_{ij}) \tag{22}$$

where α_j denotes the median of scaled values of statistical indicator j, ($j = 1, 2, 3, 4, 5$) equals to RMSE, MAE, MAPE, RRMSE, and RRMSE, -1 for r; and g_j denotes the scaled value of the statistical indicator j for model i. In terms of its physical interpretation, we note that a large GPI is expected to reveal good performance of the proposed CMLP (and all counterpart) models.

The performance characteristics of the CMLP model were also assessed using the Kling-Gupta efficiency (KGE) [92] and absolute percentage bias (APB; %) [93] expressed as follows:

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\langle GSR^p\rangle}{\langle GSR^m\rangle} - 1\right)^2 + \left(\frac{CV_p}{CV_m}\right)^2} \tag{23}$$

$$APB = \frac{\sum_{i=1}^{n}(GSR^m - GSR^p) * 100)}{\sum_{i=1}^{n} GSR^m}, \tag{24}$$

where r denotes the correlation coefficient between the observed and the forecasted GSR and CV stands for the coefficient of variation among these values in the testing phase. In addition to these, the directional movement was evaluated using a directional symmetry (DS):

$$DS = \frac{1}{n}\sum_{t=2}^{n} d_t \times 100\% \tag{25}$$

An additional evaluation of the CMLP model performance, purely from a statistical viewpoint, was carried using the DM (Diebold-Mariano) and HLN (Harvey, Leybourne, and Newbold) criteria. These criteria aimed to examine the statistical significance of all of the considered AI-based models. To interpret this metric, consider the DM and the

Table 6 The skill score (SS) metric for the proposed hybrid CMLP model (compared with the AI-based and the temperature-based deterministic models) used for *GSR* predictions. For model descriptions, readers should consult Tables 4 and 5. *The highest value of SS should be registered for the best performing model at any given site*

Predictive model	Kidston	Brigalow	Bulli Creek	Baking Board	Chewco	Blackwater
AI-based predictive models						
CMLP — proposed model	0.53	**0.75**	**0.71**	**0.75**	**0.76**	**0.71**
ELM	0.47	0.68	0.58	0.66	0.60	0.63
WKNNR	−0.52	0.01	−0.19	−0.18	−0.24	−0.05
ANN	0.51	0.73	0.53	0.70	0.53	0.65
KRR	0.45	0.38	0.59	0.68	0.68	0.63
XGBOOST	0.36	0.62	0.46	0.63	0.60	0.62
BRF	−0.29	0.29	0.06	0.29	0.08	0.12
DNN	0.07	0.62	0.46	0.35	0.62	0.21
Temperature-based predictive models						
TMAL	−0.13	0.50	0.36	0.65	0.69	−0.02
TMAN	−0.08	0.60	0.37	0.58	0.40	0.55
TMBC	−0.42	0.01	0.02	−0.04	−0.03	0.03
TMBW	−0.44	0.01	0.02	−0.04	−0.07	0.01
TMGO	−0.05	0.63	0.69	0.67	0.35	0.61
TMHAS	−0.24	0.18	0.06	0.54	0.61	0.20
TMHG	−0.45	0.01	0.02	−0.04	−0.08	0.01

HLN statistic to be > 0 if the CMLP model outperforms the comparative model, in accordance with previous works, e.g., [93–95]. Finally, we have utilized the gray relational analysis to calculate the gray relational degree (GRD) for all of the models in order to rank them on the basis of the GRD values [96].

Results

The proposed hybrid CMLP model is evaluated by testing its predictive capability for daily *GSR* simulation at the six solar energy farms over Queensland state, Australia.

Table 7 The performance of the proposed hybrid CMLP model (compared with the AI-based and the temperature-based deterministic models) applied in testing phase as measured by *RRMSE* and *MAPE* values. For model descriptions, readers should consult Tables 4 and 5

Predictive models	Kidston		Brigalow		Bulli Creek		Baking Board		Chewco		Blackwater	
	RRMSE	*MAPE*	*RRMSE*	*MAPE*	*RRMSE*	*MAPE*	*RRMSE*	*MAPE*	*RRMSE*	*MAPE*	*RRMSE*	*MAPE*
AI-based predictive models												
CMLP — proposed model	**9.00%**	**8.40%**	**8.50%**	**10.20%**	**8.60%**	**10.90%**	**8.20%**	**10.70%**	**7.90%**	**10.20%**	**8.40%**	**10.50%**
DNN	10.10%	10.20%	8.50%	10.80%	9.20%	11.50%	8.90%	11.70%	8.90%	11.70%	15.80%	20.70%
ANN	9.20%	8.40%	8.60%	10.70%	9.70%	12.30%	8.90%	11.50%	8.80%	11.20%	9.30%	12.10%
ELM	9.60%	9.00%	9.40%	12.30%	10.30%	13.20%	9.60%	12.70%	8.80%	11.70%	9.50%	12.40%
KRR	9.70%	9.10%	9.20%	11.80%	10.20%	12.80%	9.30%	12.40%	9.00%	11.80%	9.50%	12.40%
XGBOOST	10.40%	9.80%	10.30%	13.90%	11.70%	15.10%	9.90%	13.40%	10.10%	13.40%	9.60%	12.20%
BRF	14.60%	14.20%	14.00%	16.70%	15.40%	19.80%	13.90%	17.90%	15.40%	19.40%	14.60%	18.30%
Temperature-based predictive models												
WKNNR	15.90%	15.70%	16.50%	23.90%	17.40%	24.30%	17.80%	25.20%	18.00%	25.60%	16.10%	22.90%
TMGO	9.80%	9.10%	10.10%	13.10%	9.00%	11.70%	9.30%	11.90%	12.90%	16.20%	9.80%	12.80%
TMAL	10.90%	10.40%	10.50%	13.60%	9.50%	12.30%	10.50%	13.70%	12.90%	16.90%	10.60%	13.80%
TMAN	10.50%	10.10%	10.60%	14.60%	9.70%	13.10%	10.60%	14.50%	12.50%	16.90%	10.60%	14.10%
TMHAS	13.00%	12.90%	13.20%	19.80%	12.70%	18.60%	13.40%	19.70%	14.30%	21.00%	13.20%	18.70%
TMBC	15.30%	15.20%	16.50%	24.50%	15.80%	22.60%	16.60%	24.50%	16.40%	24.30%	15.40%	22.00%
TMCH	15.40%	15.00%	16.50%	24.10%	15.80%	22.30%	16.70%	23.90%	16.70%	24.00%	15.60%	22.40%
TMHG	15.50%	15.10%	16.50%	24.10%	15.80%	22.40%	16.70%	24.00%	16.70%	24.10%	15.60%	22.40%

Table 8 The performance of the proposed hybrid CMLP model (compared with the AI-based and the temperature-based deterministic models) in the testing phase as measured by *WI* and *NSE*. For model descriptions, readers should consult Tables 4 and 5. For model descriptions, readers should consult Tables 4 and 5

Predictive model	Kidston		Brigalow		Bulli Creek		Baking Board		Chewco Solar		Blackwater	
	WI	*NSE*	*WI*	*NSE*	*WI*	*NSE*	*WI*	*NSE*	*WI*	*NSE*	*WI*	*NSE*
AI-based predictive models												
CMLP — proposed model	**0.893**	**0.795**	0.923	**0.860**	**0.929**	**0.859**	**0.931**	**0.869**	**0.933**	**0.875**	**0.933**	**0.875**
DNN	0.832	0.752	**0.926**	0.860	0.915	0.837	0.918	0.847	0.913	0.845	0.913	0.845
ANN	0.888	0.785	0.922	0.855	0.911	0.818	0.918	0.845	0.918	0.848	0.918	0.848
ELM	0.876	0.765	0.909	0.829	0.895	0.796	0.903	0.823	0.917	0.845	0.917	0.845
KRR	0.875	0.758	0.914	0.836	0.898	0.801	0.909	0.834	0.912	0.839	0.912	0.839
XGBOOST	0.859	0.727	0.892	0.794	0.865	0.736	0.899	0.810	0.894	0.798	0.894	0.798
BRF	0.742	0.458	0.812	0.614	0.785	0.544	0.828	0.626	0.772	0.527	0.772	0.527
Temperature-based predictive models												
WKNNR	0.534	0.350	0.649	0.473	0.638	0.425	0.595	0.389	0.569	0.365	0.569	0.365
TMGO	0.887	0.753	0.892	0.800	0.923	0.845	0.912	0.830	0.815	0.667	0.815	0.667
TMAL	0.878	0.698	0.893	0.784	0.922	0.829	0.898	0.784	0.835	0.671	0.835	0.671
TMAN	0.876	0.719	0.885	0.782	0.913	0.822	0.887	0.780	0.833	0.691	0.833	0.691
TMHAS	0.829	0.574	0.841	0.661	0.866	0.698	0.844	0.655	0.814	0.596	0.814	0.596
TMBC	0.668	0.405	0.702	0.482	0.741	0.535	0.697	0.477	0.701	0.483	0.701	0.483
TMCH	0.617	0.399	0.676	0.481	0.722	0.534	0.670	0.476	0.662	0.465	0.662	0.465
TMHG	0.614	0.395	0.677	0.480	0.723	0.534	0.670	0.474	0.663	0.463	0.663	0.463

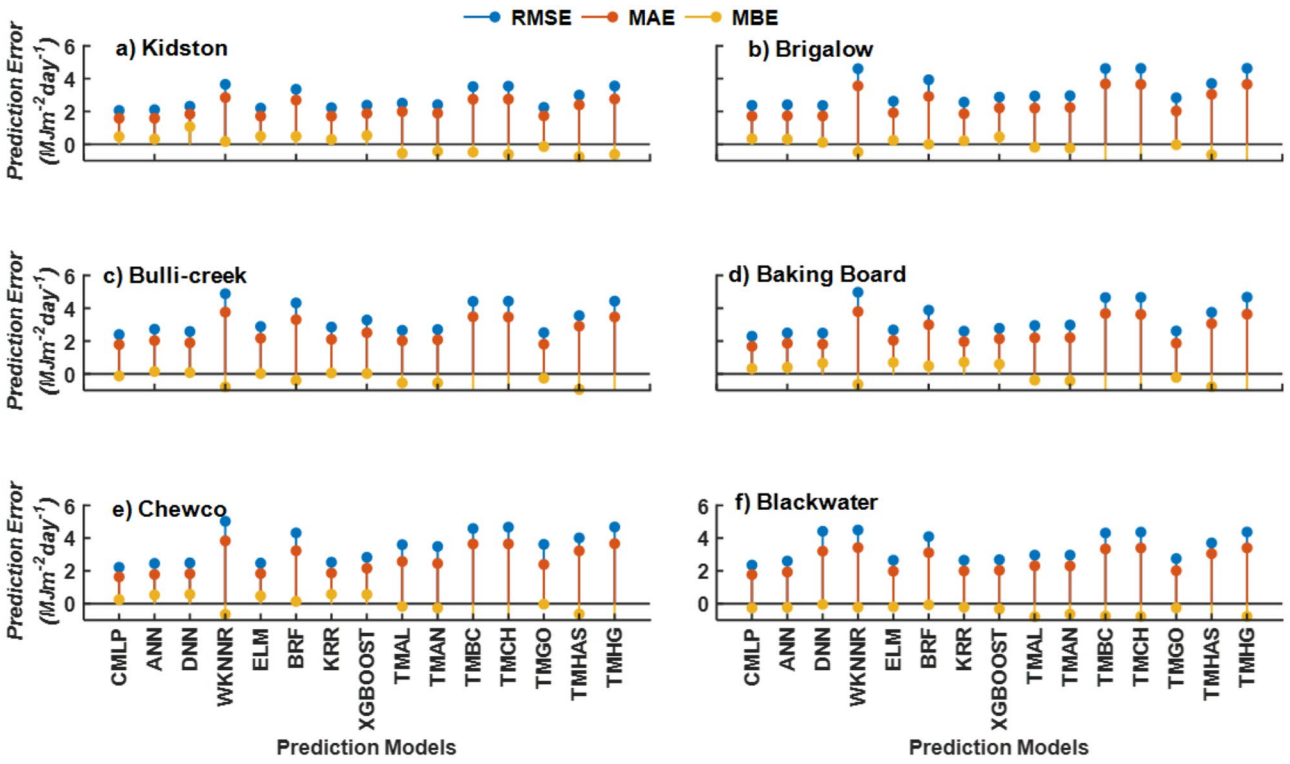

Fig. 4 Evaluation of the hybrid CMLP predictive model relative to its benchmark models using the *RMSE*, *MAE*, and the *MBE*, computed within the testing phase: **a** Kidston Solar farm, **b** Brigalow Solar Farm, **c** Bulli Creek Solar Farm, **d** Baking Board Solar Farm, **e** Chewco Solar Farm, and **f** Blackwater Solar Farm

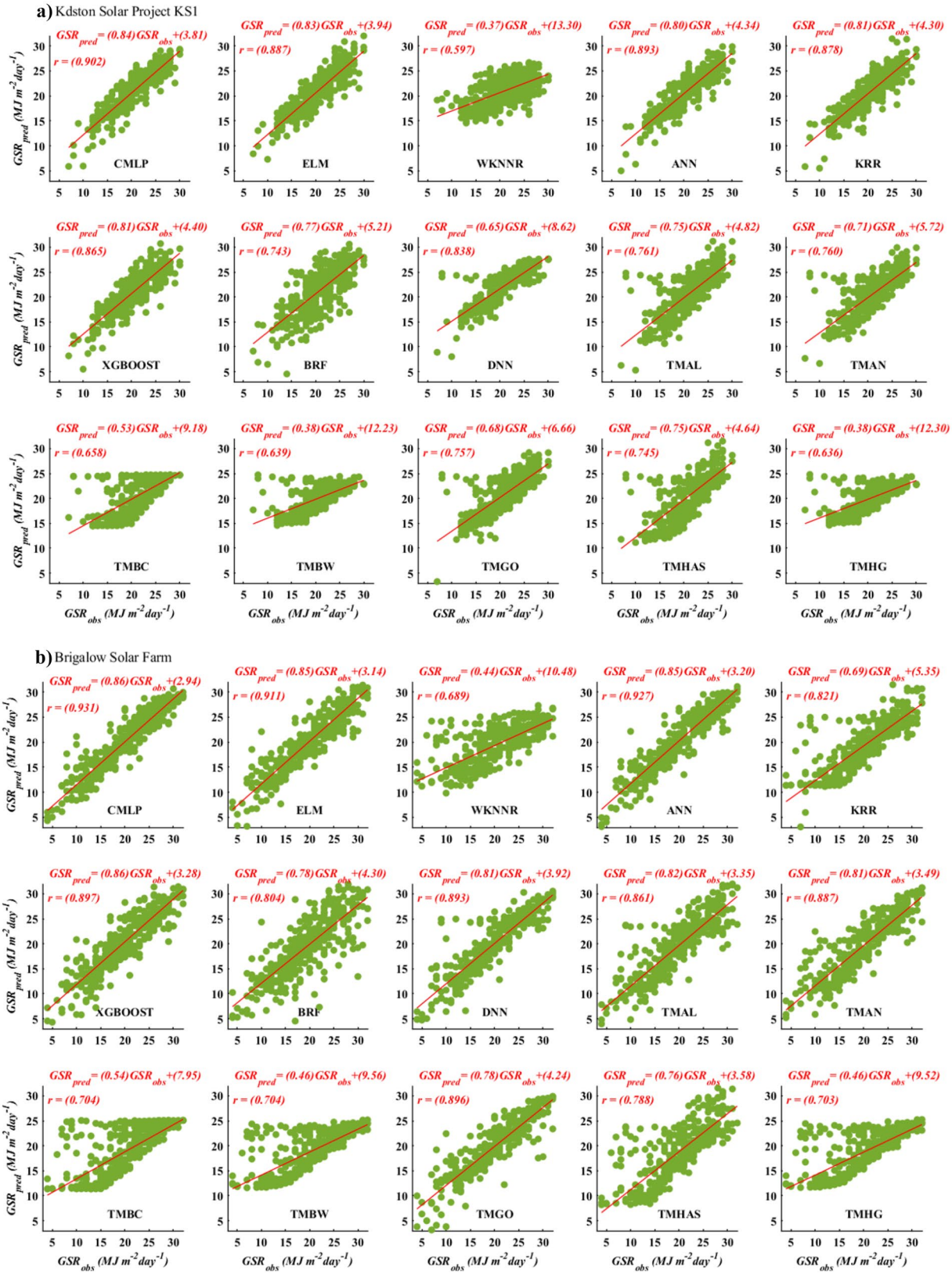

Fig. 5 Scatter plots of the observed (GSR_{obs}) and predicted (GSR_{pred}) daily global solar radiation for all tested region in Queensland region: **a** Kidston Solar farm, **b** Bulli Creek Solar Farm, **c** Brigalow Solar Farm, **d** Baking Board Solar Farm, **e** Chewco Solar Farm, and **f** Blackwater Solar Farm (note: Line in red is the least-squares fit line ($y = mx + c$) to the respective scatter plots, where y = predicted GSR and x = observed GSR)

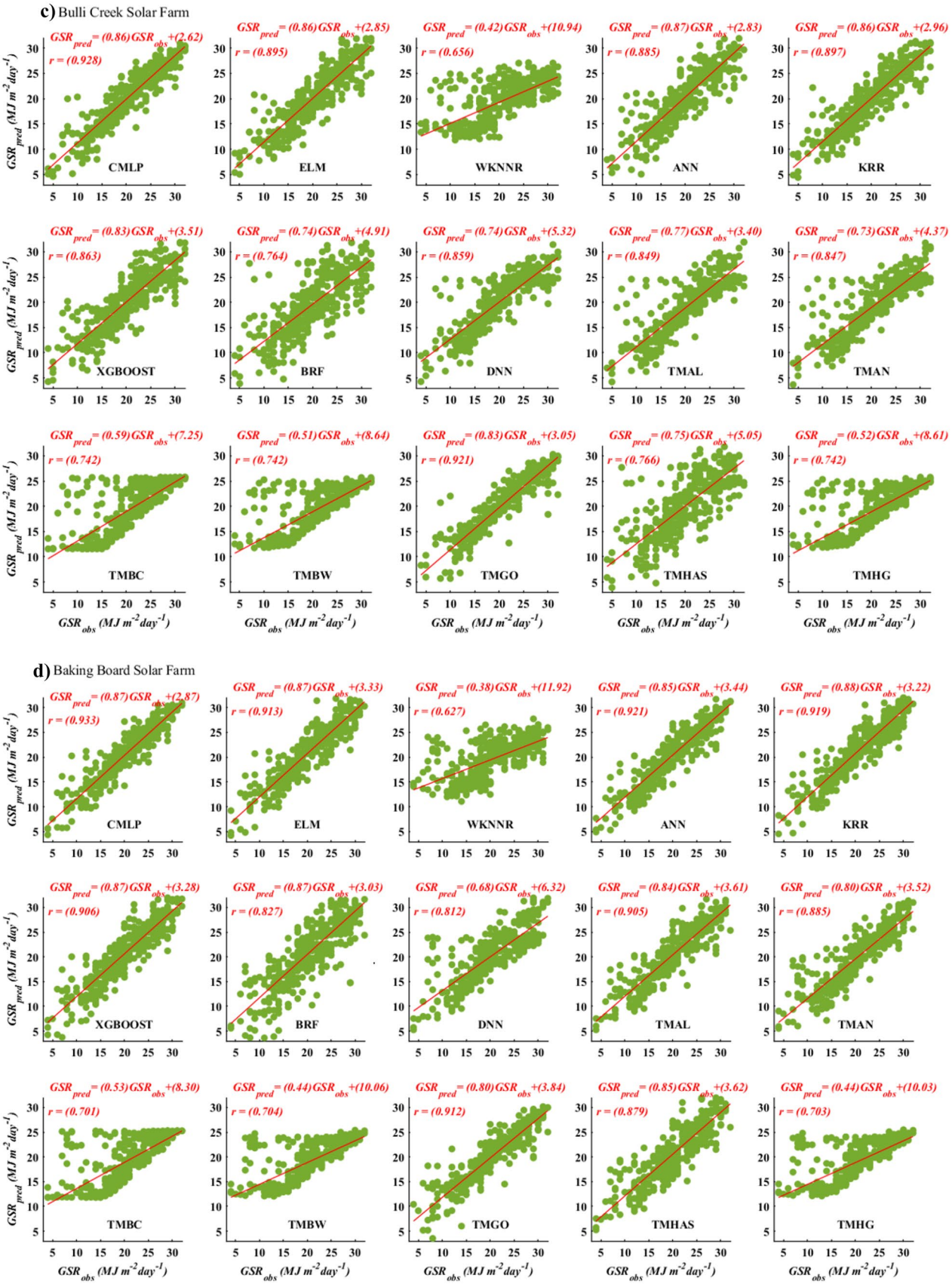

Fig. 5 (continued)

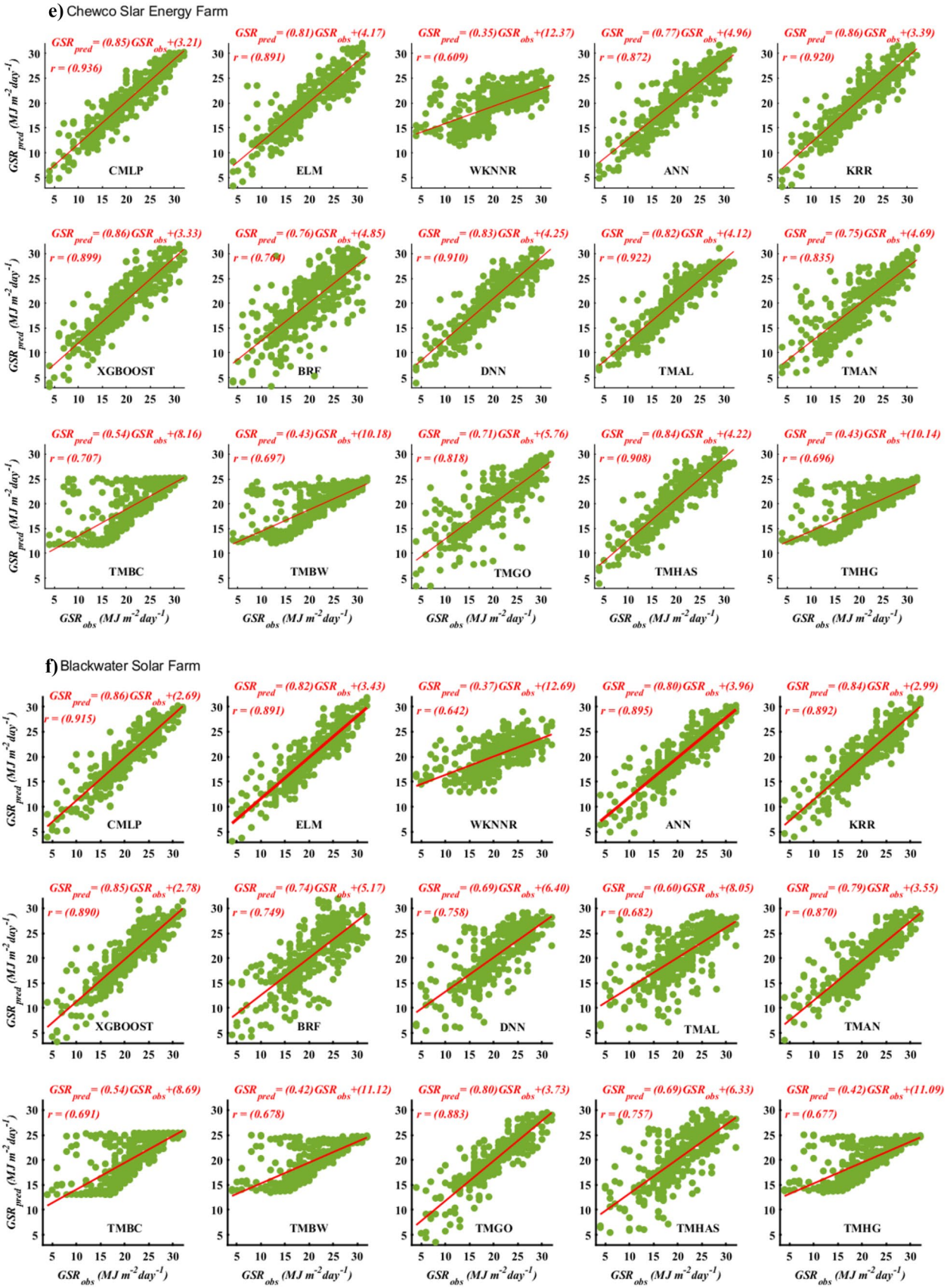

Fig. 5 (continued)

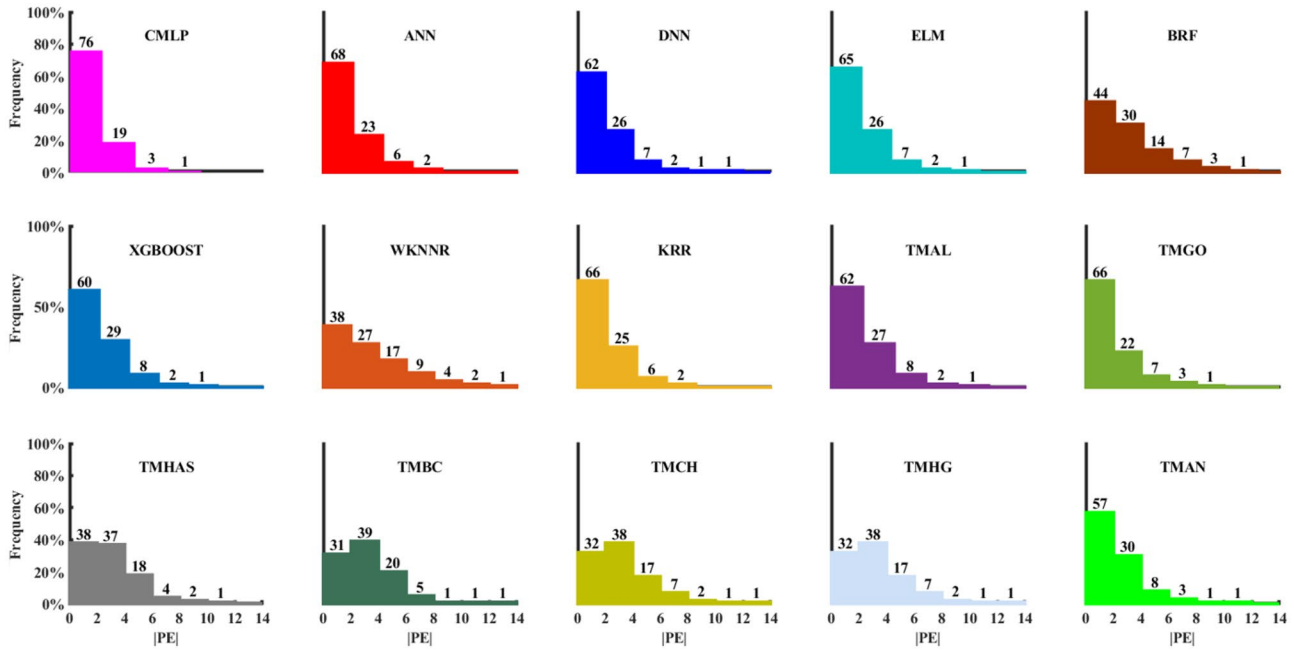

Fig. 6 The cumulative frequency of the daily prediction errors for all tested region in the Queensland pooled together. The percentage error accumulated in each bracket is shown in the respective bar

To fully appraise the performance of the proposed CMLP model, we analyze the tested data with respect to the artificial intelligence (AI)–based models (i.e., ANN, ELM, DNN, BRF, KRR, WKNNR, and XGB) and the temperature-based deterministic (i.e., TMHS, TMCH, TMBC, TMGO, TMAL, TMAN, and TMHG) models. This section reports the statistical measures using Eqs. 13 to 24 while employing the other exploratory analysis approach using histogram, scatter plots, Taylor plot, etc., presented in the following.

We now refer to four statistical measures: *RRMSE*, *MAPE*, *WI*, and *NSE*, including *SS* values presented in Tables 6, 7,

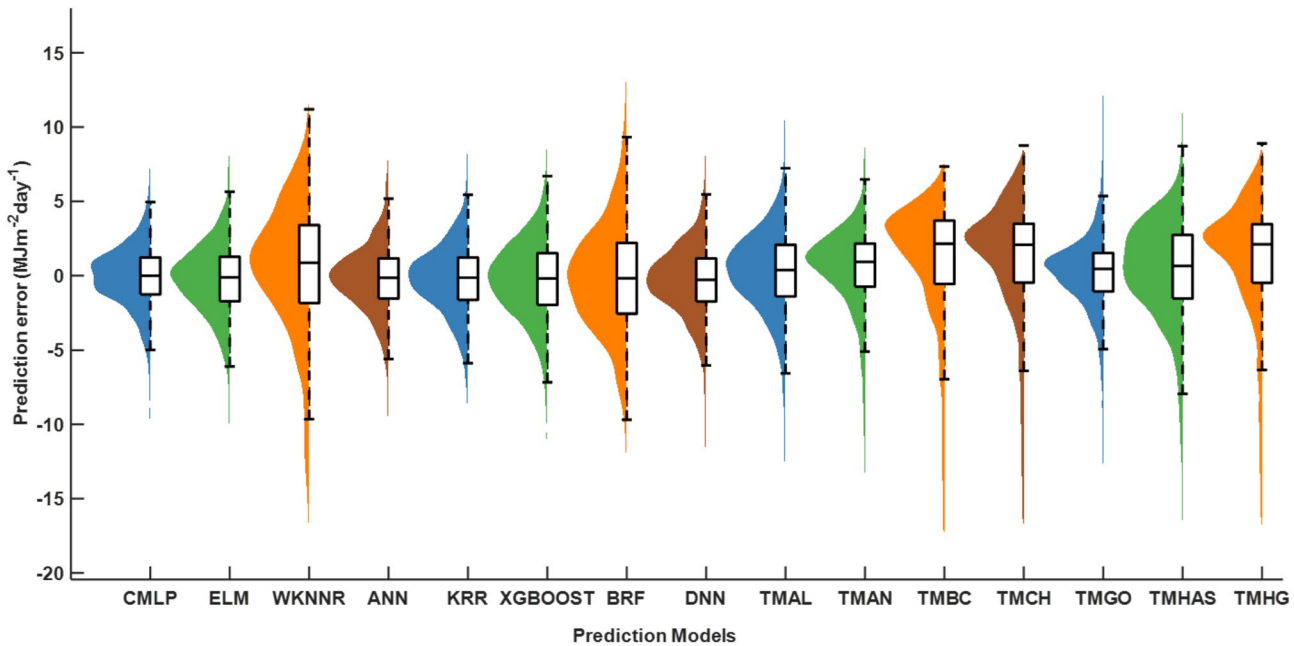

Fig. 7 Violin plots of the prediction error (PE) generated by prediction models during testing phase for daily GSR prediction

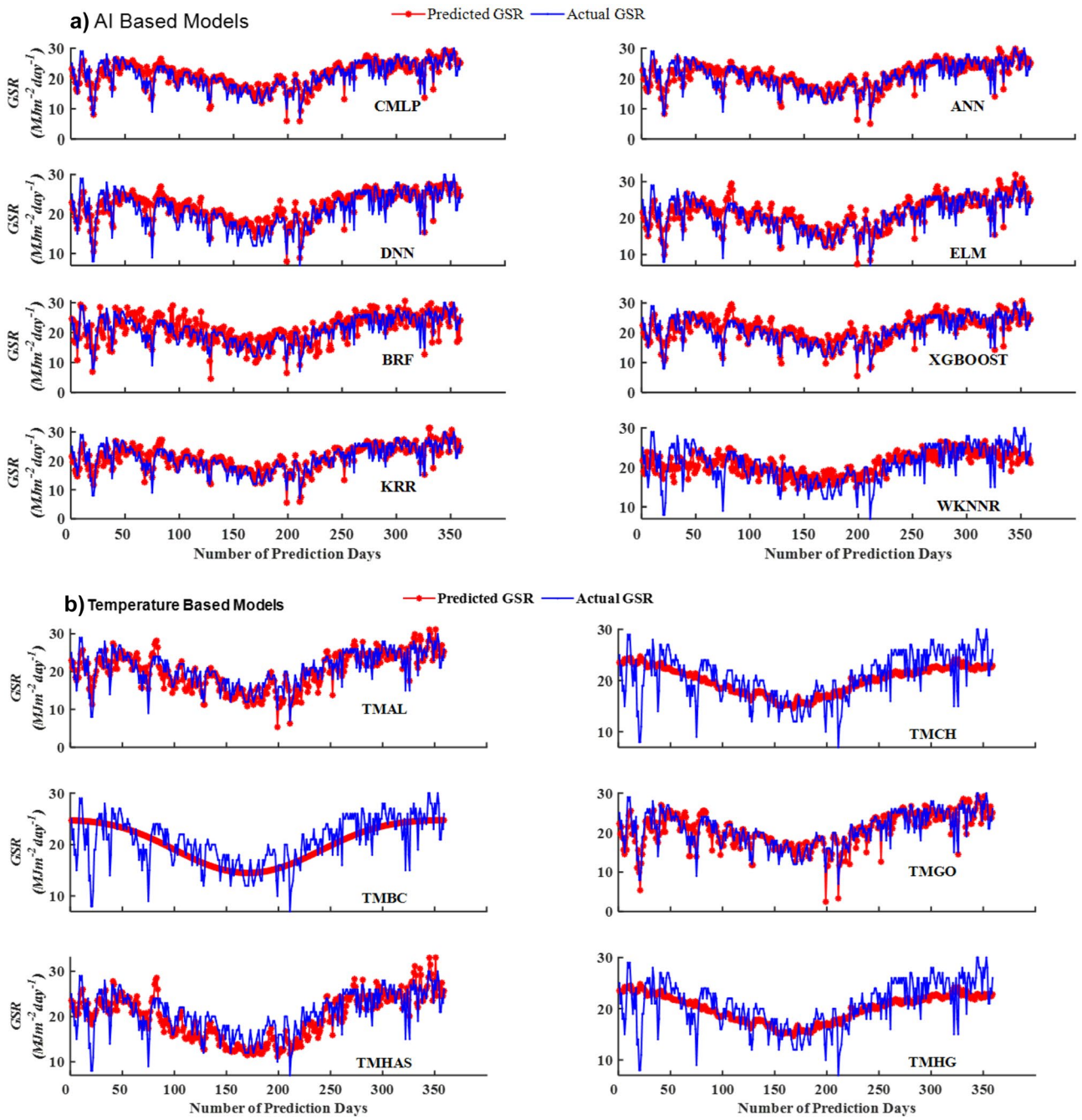

Fig. 8 The predicted vs. the observed daily GSR in the testing phase using CMLP model vs. the other competing **a** artificial intelligence–based models and **b** temperature-based models

and 8 to evaluate the proposed hybrid CMLP against the AI-based and temperature-based counterpart models.

The hybrid CMLP model yields the lowest *RRMSE* and *MAPE* value (≈ [8.60%, 9.00%], *MAPE* ≈ [8.40 %, 10.90 %]) benchmarked with the best AI-based models ANN (*RRMSE* ≈ [8.60 %, 9.30 %], *MAPE* ≈ [8.40 %, 12.30 %]) and temperature-based models TMGO (*RRMSE* ≈ [9.00 %, 12.90 %], *MAPE* ≈ [9.10 %, 16.20 %). The worst performing AI-based model, denoted as WKNNR, generated an *RRMSE* ≈ [15.90 %, 18.00 %], *MAPE* ≈ [15.70 %, 25.60 %] while that in the temperature-based model group is TMBC (*RRMSE* ≈ [15.30 %, 16.60 %], *MAPE* ≈ [15.20 %, 24.50 %]). The results of the objective model, attaining a less than 10 % relative error, clearly suggest that the hybrid CMLP model falls within a category of an excellent model, which also accords to earlier research work, e.g., [97].

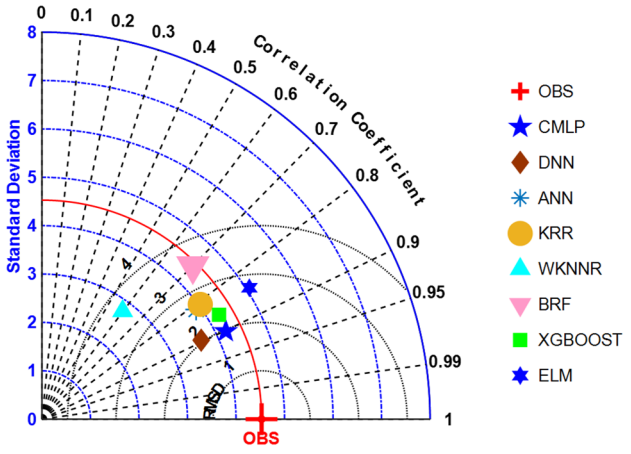

Fig. 9 Taylor diagram for CMLP model versus other artificial intelligence–based models to represent the bias and standard deviation of errors. The azimuthal angle shows the correlation, the radial distance, the standard deviation, and the semicircles centered at the observation "OBS" marker the standard deviation of the errors. The color scale corresponds to the bias (mean of model minus mean of observation)

Similar conclusions can be drawn from *WI* and *NSE* values in Table 8 where the higher values of these measurements, the better is the model performance. Note that these indices are non-dimensional and normalized measures where a value of unity will indicate a predicted *GSR* matching exactly with the observed value. Except for Brigalow Solar Farm, the hybrid CMLP model generally registered the highest *WI* and *NSE* values ($WI \approx [0.893, 0.933]$, $NSE \approx [0.795, 0.875]$), followed by ANN ($WI \approx [0.888, 0.922]$, $NSE \approx [0.785, 0.855]$) and DNN ($WI \approx [0.832, 0.926]$, $NSE \approx [0.752, 0.860]$) in the AI-based group and TMGO ($WI \approx [0.815, 0.923]$, $NSE \approx [0.667, 0.845]$) and TMAL ($WI \approx [0.835, 0.922]$, $NSE \approx [0.671, 0.829]$) in the temperature-based group. The worst performance among the AI-based model is WKNNR ($WI \approx [0.534, 0.649]$, $NSE \approx [0.350, 0.473]$) and TMCH ($WI \approx \approx [0.617, 0.722]$, $NSE \approx [0.399, 0.534]$) and TMHG ($WI \approx [0.614, 0.723]$, $NSE \approx [0.395, 0.534]$). It is therefore clear that the proposed CMLP model far exceeds the performance of all AI-based

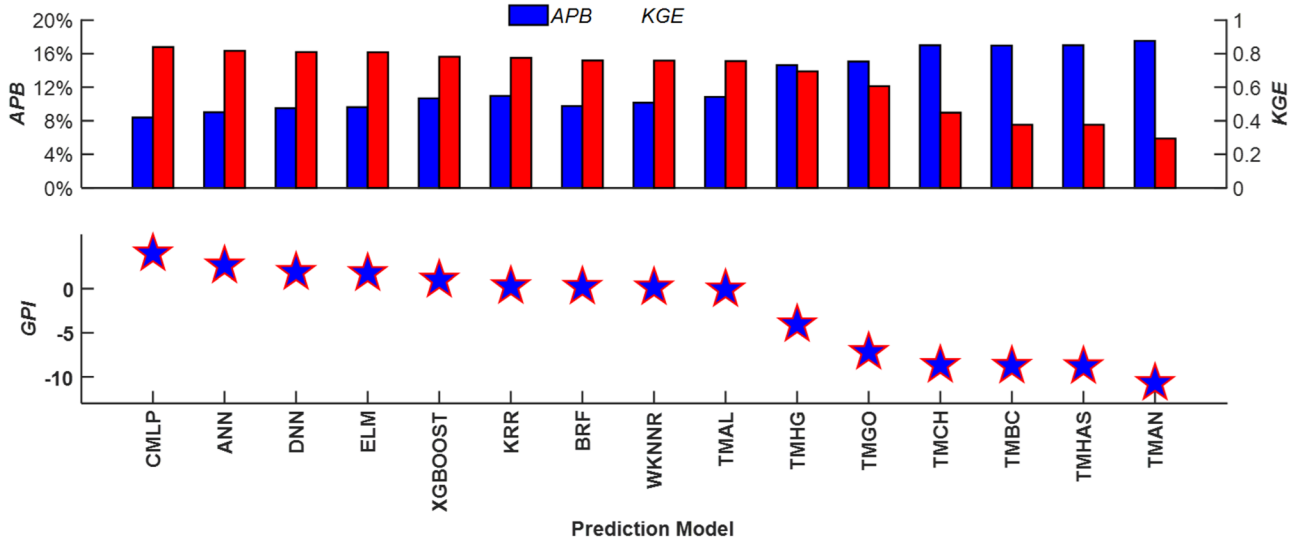

Fig. 10 **a** Bar chart illustrating a comparison of the CMLP model in terms of their absolute percentage bias (APB, %) and Kling-Gupta efficiency (KGE) in the testing phase. **b** Global performance indicator (GPI) of CMLP model compared to other AI-based models and temperature models

Table 9 The Diebold-Mariano (DM) test statistic value where the column is compared to rows. The positive values indicate an outperformance of column-based model relative to the row-based model

Predictive model	CMLP	DNN	ELM	ANN	KRR	BRF	XGBOOST	WKNNR
CMLP	-	8.274	9.440	6.776	8.808	18.757	13.654	21.795
DNN	-	-	−4.220	−5.901	−4.337	11.875	−0.752	17.264
ELM	-	-	-	−4.292	−0.796	16.315	7.564	20.352
ANN	-	-	-	-	3.633	17.437	10.010	20.415
KRR	-	-	-	-	-	16.461	7.237	20.360
BRF	-	-	-	-	-	-	−14.231	6.308
XGBOOST	-	-	-	-	-	-	-	18.462

Table 10 The Harvey, Leybourne, and Newbold statistic value where the column is compared to rows. The positive values indicate an outperformance of column-based model relative to the row-based model

Predictive model	CMLP	DNN	ELM	ANN	KRR	BRF	XGBOOST	WKNNR
CMLP	-	4.297	7.062	5.567	6.190	16.449	9.401	11.012
DNN	-	-	−2.157	−3.046	−2.260	8.639	−0.375	8.331
ELM	-	-	-	−3.085	−0.691	15.297	5.621	10.482
ANN	-	-	-	-	2.712	15.562	6.683	10.369
KRR	-	-	-	-	-	15.787	5.599	10.552
BRF	-	-	-	-	-	-	−13.001	4.035
XGBOOST	-	-	-	-	-	-	-	9.639

Table 11 The performance of the hybrid CMLP model utilizing the ratio of the root mean square error (*r*RMSE). Note: an *r*RMSE greater than unity indicates that the objective model outperforms the comparative model

Predictive model	CMLP	DNN	ELM	ANN	KRR	BRF	XGBOOST	WKNNR
CMLP	-	1.590	1.299	1.179	1.284	3.092	1.535	4.152
DNN	-	-	0.817	0.741	0.807	1.945	0.965	2.611
ELM	-	-	-	0.907	0.988	2.380	1.181	3.196
ANN	-	-	-	-	1.089	2.623	1.302	3.522
KRR	-	-	-	-	-	2.409	1.196	3.235
BRF	-	-	-	-	-	-	0.496	1.343
XGBOOST	-	-	-	-	-	-	-	2.705

and temperature-based models applied to predict *GSR* at six solar farms across Queensland, Australia.

For a broader picture of the error comparison, the values of *RMSE*, *MAE*, and *MBE* (Figs. 4 and 5) and prediction errors (Figs. 6 and 7) of the fifteen models for daily GSR prediction across six solar study sites are illustrated intuitively using vertical histograms, scatter plots overlaid by best-fitting regression lines, frequency analysis, and violin plots (Fig. 7), respectively. In general, several consistent conclusions can be drawn from the detailed analysis of these figures. First, the AI-based models generally produce better performance than the temperature-based models. Second, among the artificial intelligence–based models, the CMLP hybrid predictive model is the best followed by the ANN, ELM, and KRR models. Finally, the TMAL, TMGO, and TMAN models outperform other candidates in the temperature-based group.

Taking Fig. 6 as an example where the daily prediction errors at all tested sites are pooled together, the cumulative

Table 12 The gray relational degree (GRD) computed for each predictive model for the purpose of ranking each predictive model based on the gray relational analysis (GRA) method

Predictive model	Kidston		Brigalow		Bulli Creek		Baking Board		Chewco		Blackwater	
	GRD	*RANK*	*GRD*	*RANK*	*GRD*	*RANK*	*GRD*	*RANK*	*GRD*	*RANK*	*GRD*	*RANK*
CMLP	**0.755**	1	**0.702**	1	**0.715**	1	**0.653**	1	**0.536**	1	**0.585**	1
ELM	0.702	9	0.699	2	0.661	11	0.635	4	0.506	11	0.551	6
WKNNR	0.602	12	0.612	11	0.616	13	0.505	12	0.463	14	0.531	12
ANN	0.733	3	0.685	4	0.693	5	0.627	7	0.520	6	0.547	7
KRR	0.718	7	0.669	6	0.669	10	0.651	2	0.528	4	0.568	4
XGBOOST	0.728	5	0.664	7	0.672	9	0.640	3	0.520	7	0.569	3
BRF	0.725	6	0.652	9	0.698	4	0.631	5	0.528	3	0.566	5
DNN	0.729	4	0.684	5	0.714	2	0.601	10	0.515	8	0.535	10
TMAL	0.703	8	0.657	8	0.683	6	0.624	8	0.510	9	0.536	9
TMAN	0.694	10	0.648	10	0.682	7	0.595	11	0.530	2	0.574	2
TMBC	0.602	13	0.597	12	0.607	15	0.493	13	0.472	12	0.518	13
TMBW	0.553	14	0.566	14	0.616	12	0.477	14	0.463	15	0.493	15
TMGO	0.741	2	0.694	3	0.677	8	0.631	6	0.507	10	0.532	11
TMHAS	0.626	11	0.583	13	0.709	3	0.619	9	0.522	5	0.543	8
TMHG	0.552	15	0.565	15	0.616	14	0.477	15	0.465	13	0.495	14

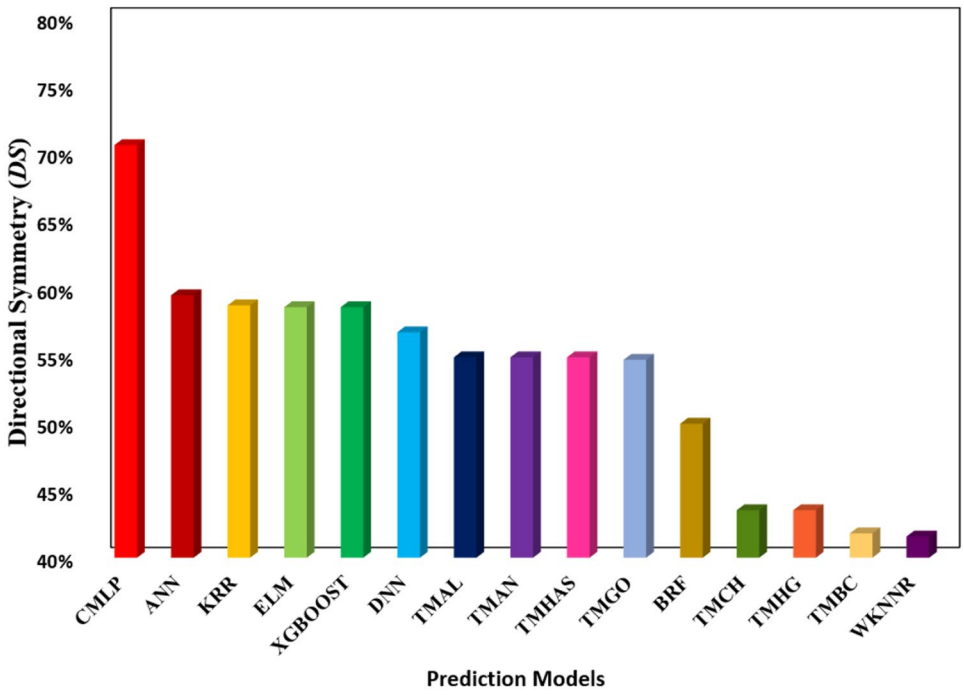

Fig. 11 Performance comparison of CMLP (convolutional neural network model (CNN) integrated with multilayer perceptron (MLP)) model compared to other models using directional symmetry (DS) criteria

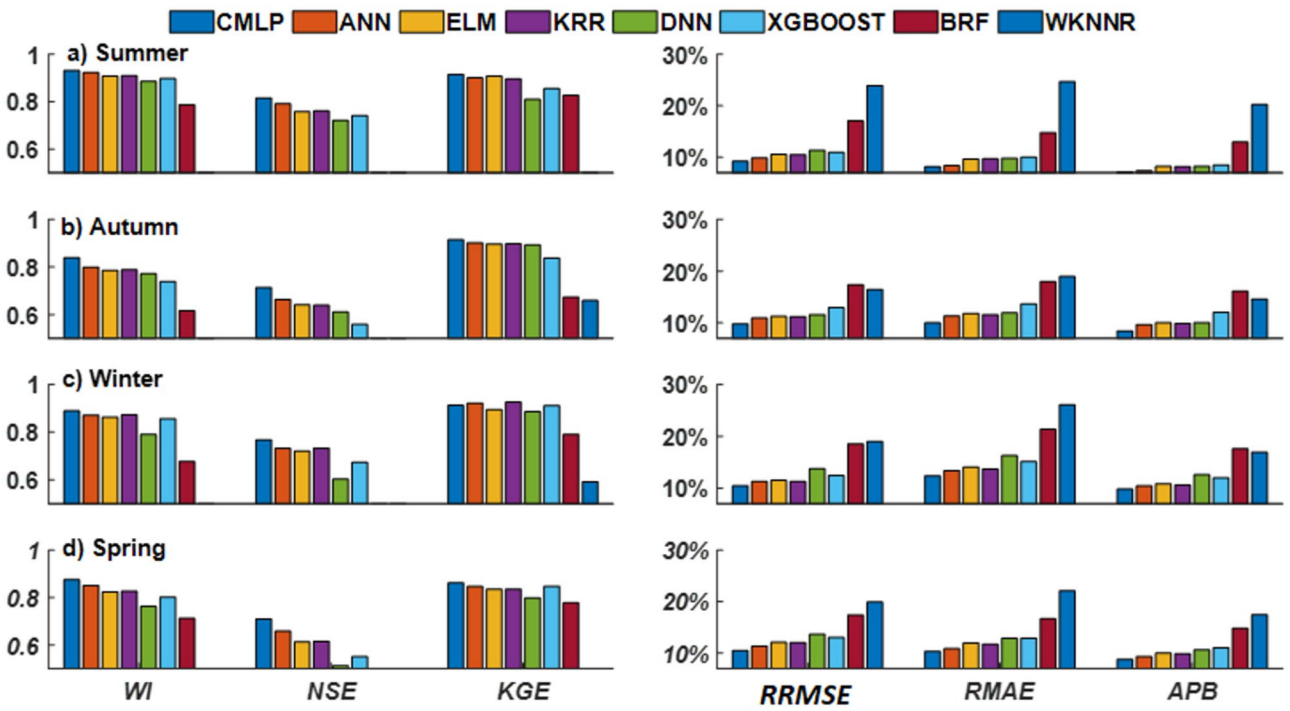

Fig. 12 Evaluation of seasonal performance of the proposed CMLP model relative to other artificial intelligence–based model using Willmott's Index (*WI*), Nash-Sutcliffe coefficient (*NSE*), Kling-Gupta efficiency (*KGE*), relative root mean square error (*RRMSE*, %), relative mean absolute error (*RMAE*, *%orMAPE*, %), and absolute percentage bias (*APB*, %). **a** Summer, **b** Autumn, **c** Winter, and **d** Spring

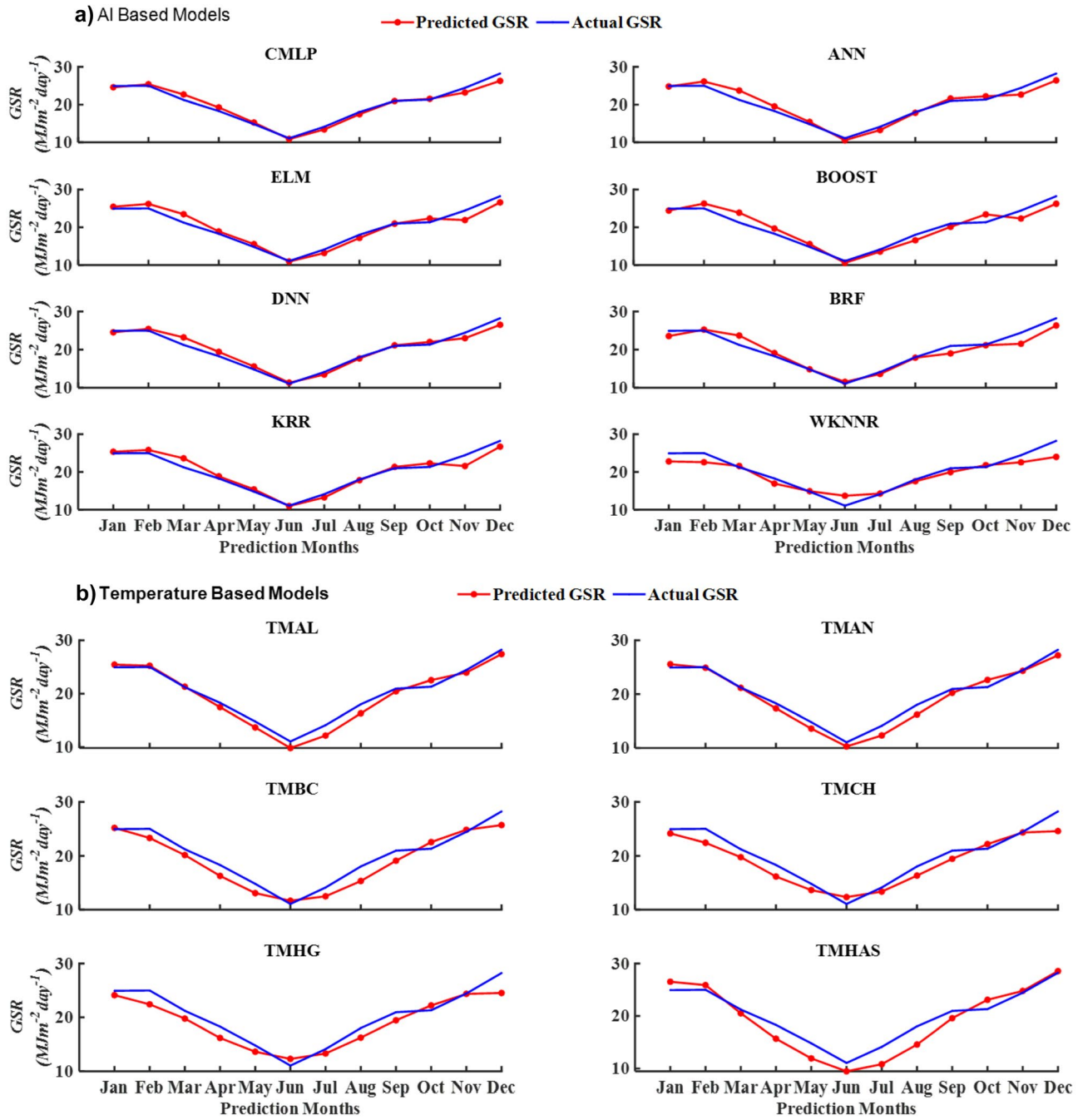

Fig. 13 The predicted vs. the observed monthly GSR in the testing phase using CMLP model vs. the other competing **a** artificial intelligence–based models and **b** temperature-based models

percentage error in [0, 4] MJm^{-2} of the hybrid CMLP model is up to 95 % while this figure of the ANN, ELM, and KRR models is about 91 % and the TMAL, TMGO, and TMAN models between 87 and 89 %. A further diagnostic of model capacity in predicting daily *GSR* in 1 year ahead is performed and shown in stacked line plots in Fig. 8.

The magnitudes of correlation coefficient (*r*) are represented in the form of the Taylor diagram, see Fig. 9, that provides a more detailed evaluation of the model performances. More specifically, the Taylor diagram portrays a more tangible and convincing statistical relationship between the forecasted and observed GSR, depending on correlation coefficients with respect to standard deviations. It can be seen that the simulations from the benchmark model ELM are highly deviated from the observed point. In addition, the proposed hybrid CMLP model is lying closer to

the observed point that confirms the predictive accuracy is higher at all six solar sites than the benchmark models.

Discussion

Other aspects of the prediction capacity in terms of bias, global performance, and directional symmetry are also discussed in this study. Figure 10 shows a comparison of the CMLP model measured by absolute percentage bias (APB, %), the Kling-Gupta efficiency (*KGE*), and global performance indicator (*GPI*) while Fig. 11 shows the findings using directional symmetry (*DS*) criteria. The hybrid CMLP model, as expected, has the lowest value of *APB* and the highest values of *KGE* and *GPI* indicating its outperformance. It is worth pointing out that the CMLP model yields $DS \approx 72$ %, noticeably higher than that figure $DS \approx 60$ % acquired from ANN as the second top model.

Tables 9, 10, and 11 examine the proposed hybrid CMLP model relative to the standalone AI-based models using the DM, HLN, and *r*RMSE performance criterion. Positive values indicate that the proposed model outperforms the comparative models, where the lower the value, the closer the performance. The lowest and highest values of these measurements when comparing CMLP with other models across six solar farms are $DM \approx 6.776, 21.795$ (with ANN and WKNNR), $HLN \approx 4.297, 16.449$ (with DNN and BRF), and $RRMSE \approx 1.179, 4.152$ (with ANN and WKNNR). The rank-based assessment using the GRD method (Table 12) also confirms the supremacy of the proposed hybrid CMLP in its capability to predict the daily *GSR* for all of the solar energy study sites.

Finally, the hybrid CMLP model's skill for seasonal prediction is evaluated using *WI*, *NSE*, *KGE*, *rRMSE*, *RMAE*(*MAPE*), and *APB* (Fig. 12). In general, the CMLP model outperformed the other AI-based and temperature-based models except in winter, when using the *KGE* measurement. These criteria clearly indicate the worst performance of BRF and WKNNR models in all seasons. Moreover, the model performance in monthly prediction is also shown in Fig. 13. It can be seen that the proposed CMLP outperforms other benchmark models in all months except for Nov and Dec where the temperature-based models, TMAL and TMHAS, are slightly better.

Further insights can be gained by checking the *MAPE*(%) of the objective model (CMLP), and the *MAPE*(%) for six solar farms in this study ranges from 8.40 to 10.70 compared to 19.28 and 23.72 with a gated recurrent unit model [98]. Further recently published studies show that the *MAPE*(%) ranges from 10.37 to 13.12, 9.93 to 12.98, and 9.97 to 16.63 for the CNN-LSTM-MLP [39], CNN-stacked regression [35], and CNN-SVR [34] models respectively for daily *GSR* prediction. It is therefore clear that the prescribed CMLP approaches exceeded the performance of earlier studies.

Conclusions, Limitations, and Future Research Work

The prediction of global solar radiation (*GSR*) with minimal forecast error is essential towards integrating the freely available solar energy into a national electricity grid, in agreement with the United Nations Sustainable Development Goal # 7. This research has developed a new hybrid, CMLP algorithm that integrates convolutional neural networks (CNN) and multi-layer perceptron models (MLP) for *GSR* forecasting. In the designed approach, data are initially processed by means of a CNN algorithm to extract the hidden topological structure features that are later used by the MLP model to predict *GSR*. The most influential predictor variables for *GSR* were successfully selected from a large pool of 123 meteorological variables using a random forest-recursive feature elimination (RF-RFE) algorithm prior to applying the MLP method to forecast *GSR*.

Based on the overall results attained that were tested at multiple study sites, several conclusions on the predictive merits of the proposed hybrid model could be drawn through a comparative evaluation using different performance measures. During the testing phase, the hybrid CMLP model was found to outperform the benchmark models at six solar energy sites over daily, monthly, and seasonal time scales. Among the benchmark models, we also noted that the ANN and the DNN models generally yielded the best performance compared to other methods, while the best model candidates for the determinstic (i.e., temperature-based) models were the TMGO and TMAL methods. In accordance with the results obtained, the worst performing model was the WKNNR within the AI-based group and TMCH and TMHG within the temperature-based modelling groups. The results showed that the performance of benchmark models could vary depending upon the tested locations and the forecast time scales, suggesting further investigations to be conducted in future research.

Limitations and Proposals for Future Research Work

In terms of the developed method's merits, the superior performance of the newly designed CMLP model clearly revealed its adaptability as a competitive modelling technique in improving solar radiation forecasting. Therefore, our study showed that the proposed method could be further adopted for supporting clean energy integration into electricity grids, and future planning and management of solar powered renewable energy systems.

Furthermore, regarding the objective model, the CNN has proven to be a highly accurate model for image recognition problems and in detecting the important features without any human supervision. However, it needs a lot of training data. Similarly, the MLP works well with large data, and can be applied to complex non-linear problems, but the proper functioning of the model depends on the quality of the training data. If the model does not work properly, generalization problems may arise [99]. Additionally, regarding the benchmark models, RBF performs faster but the complexity of the model increases with increasing number of neurons. The ordinary RBF has several problems in its structure and training algorithm, so that it is not able to model a strongly nonlinear system [100]. Similarly, the WKNNR, KRR, ELM, and BRF are not suitable for the highly nonlinear data or noisy data [99]. In this study, our novelty lies in integrating multiple climate simulation variables, identifying the best features using random forest-recursive feature elimination and then developing the CMLP model. The added layer of complexity in the modelling methods introduces a more sophisticated method to predict *GSR* accurately.

In spite of the efficacy of CMLP method, there does exist some limitations that could form the subject of further investigation. This study has applied the proposed method on daily and longer-term *GSR*, which generally falls within a medium-scale period for key decisions made by energy companies. For example, daily predictions of solar radiation are useful in understanding the supply of solar energy over a daily basis but more real-time decisions, such as 5-min, or sub-hourly settlements of electricity prices will require short-term predictions and further tests of the CMLP method if used for grid stability and energy pricing analysis. Furthermore, solar energy is highly influenced by cloud cover. Therefore, the datasets in this study limited to cloud area fraction may bound the full assessment of intermittency in energy supply. An independent future study could also consider various cloud chromatic properties, cloud top height, water vapor, ozone, and cloud movements [101, 102] that could be factored to test the overall performance of the CMLP model.

The present study used daily variables from ACCESS1-0, Hadley-GEM2-CC, and MRI-CGCM3 global climate model datasets under CMIP5 that do not have a time-ahead forecast period, but in reality, the knowledge of several day-ahead forecasts of *GSR* may be more attractive for renewable energy industries. Therefore, future studies could employ model outputs from the Global Forecast System (GFS) that has good horizontal resolution of 28 km and a temporal resolution that covers both historical analysis and forecasts up to 16 days ahead period [103]. Furthermore, even at a lower horizontal resolution of 70 km between grid points, GFS can provide forecasts between 1 and 2 weeks. Therefore, the CMLP model could be retrained to generate forecasts anywhere up to 2 weeks in advance to support solar energy monitoring systems.

One other limitation of the present study was that only six solar energy farms across Queensland, Australia, were used to test the proposed hybrid CMLP model, and these sites utilized only the global climate model outputs and ground-based measured predictor variables. While these study sites were somewhat widely distributed in Queensland (see Fig. 2), further testing of the model across other locations in Australia, and particularly, across the high latitude or sub-tropical sites where cloud effects are more prominent, could help evaluate its practical implementation. This new study could also investigate the viability of the proposed CMLP method as a universally deployable solar energy predictive system, e.g., [104] trained with both the observational and the modelled predictor datasets.

Another possibility is to utilize the European Centre for Medium Range Weather Forecasts (ECMWF) as tested already in prior studies, e.g., [12]. In particular, the ECMWF datasets record operational ensemble-based analyses and predictions describing various scenarios and their likelihood of weather occurrence over medium-range (3-hourly, 6-hourly) as well as monthly and seasonal scales and up to a year ahead predictions. The incorporation of ECMWF or other related datasets (e.g., GFS) into the proposed CMLP model may also help better understand how weather model forecasts can be integrated with AI-based models for real-time solar energy forecasting. While these suggestions and approaches are useful considerations to further validate the CMLP model, they await another independent investigation.

Acknowledgements Data were acquired from global climate model archives and ground truth observation from the Scientific Information for Landowners (SILO) repository and Coupled Model Intercomparision Phase-5 global climate models (ACCESS1-0 from CSIRO-BOM, Hadley-GEM2-CC from MOHC, and MRI-CGCM3 from MRI) which are greatly acknowledged.

Funding This research has been partially supported by Spanish Ministry of Science and Innovation (MICINN), through Project Number PID2020-115454GB-C21.

Availability of Data and Materials All the data used in this paper can be freely downloaded from the following sources: (1) Scientific Information for Landowners (SILO) https://www.data.qld.gov.au/dataset/silo-climate-database; (2) Coupled Model Intercomparision Phase-5 global climate models (ACCESS1-0 from CSIRO-BOM, Hadley-GEM2-CC from MOHC, and MRI-CGCM3 from MRI).

Declarations

Ethics Approval This article does not contain any studies with human participants or animals performed by any of the authors.

Competing Interests The authors declare no competing interests.

References

1. Best R, Burke PJ, Nepal R, Reynolds Z. Effects of rooftop solar on housing prices in Australia.Aust J Agric Resour Econ. 2021.
2. Peng T, Zhang C, Zhou J, Nazir MS. An integrated framework of Bi-directional Long-Short Term Memory (BiLSTM) based on sine cosine algorithm for hourly solar radiation forecasting. Energy. 2021;221.
3. Lai CS, Zhong C, Pan K, Ng WW, Lai LL. A deep learning based hybrid method for hourly solar radiation forecasting. Expert Syst Appl. 2021;177.
4. Guermoui M, Melgani F, Danilo C. Multi-step ahead forecasting of daily global and direct solar radiation: a review and case study of Ghardaia region. J Clean Prod. 2018;201:716–34.
5. Gopi A, Sudhakar K, Keng NW, Krishnan AR. Comparison of normal and weather corrected performance ratio of photovoltaic solar plants in hot and cold climates. Energy Sustain Dev. 2021;65:53–62.
6. Ahmed AM, Sharma E, Jui SJJ, Deo RC, Nguyen-Huy T, Ali M. Kernel ridge regression hybrid method for wheat yield prediction with satellite-derived predictors. Remote Sens. 2022;14(5):1136.
7. Balalla DT, Nguyen-Huy T, Deo R. MARS model for prediction of short-and long-term global solar radiation. In: Predictive modelling for energy management and power systems engineering. Netherlands: Elsevier. 2021. p. 391-436.
8. Alkhayat G, Mehmood R. A review and taxonomy of wind and solar energy forecasting methods based on deep learning. Energy and AI. 2021:100060.
9. Deo RC, Yaseen ZM, Al-Ansari N, Nguyen-Huy T, Langlands TAM, Galligan L. Modern artificial intelligence model development for undergraduate student performance prediction: an investigation on engineering mathematics courses. IEEE Access. 2020;8:136697–724.
10. Li B, Zhang J. A review on the integration of probabilistic solar forecasting in power systems. Solar Energy. 2020;210:68–86.
11. Prasad SMM, Nguyen-Huy T, Deo R. Support vector machine model for multistep wind speed forecasting. In: Predictive modelling for energy management and power systems engineering. Netherlands: Elsevier; 2021. p. 335-89.
12. Ghimire S, Deo RC, Downs NJ, Raj N. Global solar radiation prediction by ANN integrated with European Centre for medium range weather forecast fields in solar rich cities of Queensland Australia. Journal of Cleaner Production. 2019;216:288–310.
13. Deo RC, Ghimire S, Downs NJ, Raj N. Optimization of windspeed prediction using an artificial neural network compared with a genetic programming model. In: Research anthology on multi-industry uses of genetic programming and algorithms; 2021. p. 116-47.
14. Sfetsos A, Coonick A. Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques. Solar Energy. 2000;68(2):169–78.
15. Yacef R, Benghanem M, Mellit A. Prediction of daily global solar irradiation data using Bayesian neural network: a comparative study. Renewable Energy. 2012;48:146–54.
16. Zou L, Wang L, Lin A, Zhu H, Peng Y, Zhao Z. Estimation of global solar radiation using an artificial neural network based on an interpolation technique in southeast China. J Atmos Sol Terr Phys. 2016;146:110–22.
17. Lu N, Qin J, Yang K, Sun J. A simple and efficient algorithm to estimate daily global solar radiation from geostationary satellite data. Energy. 2011;36(5):3179–88.
18. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436–44.
19. Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science. 2006;313(5786):504–7.
20. Ghimire S, Deo RC, Raj N, Mi J. Deep learning neural networks trained with MODIS satellite-derived predictors for long-term global solar radiation prediction. Energies. 2019;12(12):2407.
21. Ghimire S, Deo RC, Wang H, Al-Musaylh MS, Casillas-Pérez D, Salcedo-Sanz S. Stacked LSTM sequence-to-sequence autoencoder with feature selection for daily solar radiation prediction: a review and new modeling results. Energies. 2022;15(3):1061.
22. You W, Zhang H, Zhao X. A Siamese CNN for image steganalysis. IEEE Trans Inf Forensics Secur. 2020;16:291–306.
23. Ombabi AH, Ouarda W, Alimi AM. Deep learning CNN-LSTM framework for Arabic sentiment analysis using textual information shared in social networks. Soc Netw Anal Min. 2020;10(1):1–13.
24. Yao Z, Wang Z, Liu W, Liu Y, Pan J. Speech emotion recognition using fusion of three multi-task learning-based classifiers: HSF-DNN. MS-CNN and LLD-RNN Speech Commun. 2020;120:11–9.
25. Dai H, Hwang HG, Tseng VS. Convolutional neural network based automatic screening tool for cardiovascular diseases using different intervals of ECG signals. Comput Methods Prog Biomed. 2021;203.
26. Shen Z, Fan X, Zhang L, Yu H. Wind speed prediction of unmanned sailboat based on CNN and LSTM hybrid neural network. Ocean Eng. 2022;254.
27. Belmahdi B, Louzazni M, El Bouardi A. One month-ahead forecasting of mean daily global solar radiation using time series models. Optik. 2020;219.
28. Rashid AH, Razzak I, Tanveer M, Robles-Kelly A. Ripnet: A lightweight one-class deep neural network for the identification of rip currents. In: International Conference on Neural Information Processing. Springer; 2020. p. 172-9.
29. Ganaie MA, Hu M, etal. Ensemble deep learning: a review. arXiv preprint arXiv:2104.02395. 2021.
30. Zhu B, Feng Y, Gong D, Jiang S, Zhao L, Cui N. Hybrid particle swarm optimization with extreme learning machine for daily reference evapotranspiration prediction from limited climatic data. Comput Electron Agric. 2020;173.
31. Fan J, Wu L, Ma X, Zhou H, Zhang F. Hybrid support vector machines with heuristic algorithms for prediction of daily diffuse solar radiation in air-polluted regions. Renew Energy. 2020;145:2034–45.
32. Ghimire S, Deo RC, Raj N, Mi J. Wavelet-based 3-phase hybrid SVR model trained with satellite-derived predictors, particle swarm optimization and maximum overlap discrete wavelet transform for solar radiation prediction. Renew Sust Energ Rev. 2019;113.
33. Zhang Y, Cui N, Feng Y, Gong D, Hu X. Comparison of BP, PSO-BP and statistical models for predicting daily global solar radiation in arid Northwest China. Comput Electron Agric. 2019;164.
34. Ghimire S, Bhandari B, Casillas-Pérez D, Deo RC, Salcedo-Sanz S. Hybrid deep CNN-SVR algorithm for solar radiation prediction problems in Queensland. Australia Engineering Applications of Artificial Intelligence. 2022;112.
35. Ghimire S, Nguyen-Huy T, Deo RC, Casillas-Perez D, Salcedo-Sanz S. Efficient daily solar radiation prediction with deep learning 4-phase convolutional neural network, dual stage stacked regression and support vector machine CNN-REGST hybrid model. Sustain Mater Technol. 2022;32.
36. Ghimire S, Deo RC, Casillas-Pérez D, Salcedo-Sanz S. Boosting solar radiation predictions with global climate models, observational predictors and hybrid deep-machine learning algorithms. Appl Energy. 2022;316.

37. Kim TY, Cho SB. Predicting the household power consumption using CNN-LSTM hybrid networks. In: International Conference on Intelligent Data Engineering and Automated Learning. Springer; 2018. p. 481-90.
38. Jaseena K, Kovoor BC. Decomposition-based hybrid wind speed forecasting model using deep bidirectional LSTM networks. Energy Convers Manag. 2021;234.
39. Ghimire S, Deo RC, Casillas-Pérez D, Salcedo-Sanz S, Sharma E, Ali M. Deep learning CNN-LSTM-MLP hybrid fusion model for feature optimizations and daily solar radiation prediction. Measurement. 2022:111759.
40. Yan X, Mohammadian A. Forecasting daily reference evapotranspiration for Canada using the Penman-Monteith model and statistically downscaled global climate model projections. Alex Eng J. 2020;59(2):883–91.
41. Gouda K, Nahak S, Goswami P. Evaluation of a GCM in seasonal forecasting of extreme rainfall events over continental India. Weather and Climate Extremes. 2018;21:10–6.
42. Jha MK, Gassman PW. Changes in hydrology and streamflow as predicted by a modelling experiment forced with climate models. Hydrol Proced. 2014;28(5):2772–81.
43. Rhee J, Im J. Meteorological drought forecasting for ungauged areas based on machine learning: using long-range climate forecast and remote sensing data. Agric For Meteorol. 2017;237:105–22.
44. Wang L, Liu X, Brown H. Prediction of the impacts of climate change on energy consumption for a medium-size office building with two climate models. Energy Build. 2017;157:218–26.
45. Huang J, Perry M. A semi-empirical approach using gradient boosting and k-nearest neighbors regression for GEFCom2014 probabilistic solar power forecasting. Int J Forecast. 2016;32(3):1081–6.
46. Chakravorti T, Nayak N, Bisoi R, Dash P, Tripathy L. A new robust kernel ridge regression classifier for islanding and power quality disturbances in a multi distributed generation based microgrid. Renew Energy Focus. 2019;28:78–99.
47. Ali M, Prasad R, Xiang Y, Yaseen ZM. Complete ensemble empirical mode decomposition hybridized with random forest and kernel ridge regression model for monthly rainfall forecasts. J Hydrol. 2020;584.
48. Kiangala SK, Wang Z. An effective adaptive customization framework for small manufacturing plants using extreme gradient boosting-XGBoost and random forest ensemble learning algorithms in an Industry 4.0 environment. Machine Learning with Applications. 2021;4:100024.
49. Callens A, Morichon D, Abadie S, Delpey M, Liquet B. Using Random forest and Gradient boosting trees to improve wave forecast at a specific location. Appl Ocean Res. 2020;104.
50. Ghimire S, Deo RC, Downs NJ, Raj N. Self-adaptive differential evolutionary extreme learning machines for long-term solar radiation prediction with remotely-sensed MODIS satellite and Reanalysis atmospheric products in solar-rich cities. Remote Sens Environ. 2018;212:176–98.
51. Sahu RK, Shaw B, Nayak JR, etal. Short/medium term solar power forecasting of Chhattisgarh state of India using modified TLBO optimized ELM. Eng Sci Technol Int J. 2021.
52. Liu S, Liu X, Lyu Q, Li F. Comprehensive system based on a DNN and LSTM for predicting sinter composition. Appl Soft Comput. 2020;95.
53. Han H, Xu L, Cui X, Fan Y. Novel chiller fault diagnosis using deep neural network (DNN) with simulated annealing (SA). J Int Acad Refrig. 2021;121:269–78.
54. Hargreaves GH, Samani ZA. Estimating potential evapotranspiration. J Irrig Drain Div. 1982;108(3):225–30.
55. Bristow KL, Campbell GS. On the relationship between incoming solar radiation and daily maximum and minimum temperature. Agric For Meteorol. 1984;31(2):159–66.
56. Goodin DG, Hutchinson J, Vanderlip RL, Knapp M. Estimating solar irradiance for crop modeling using daily air temperature data. Agron J. 1999;91(5):845–51.
57. Salcedo-Sanz S, Cornejo-Bueno L, Prieto L, Paredes D, García-Herrera R. Feature selection in machine learning prediction systems for renewable energy applications. Renew Sust Energ Rev. 2018;90:728–41.
58. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, et al. Backpropagation applied to handwritten zip code recognition. Neural Comput. 1989;1(4):541–51.
59. Dong N, Chang JF, Wu AG, Gao ZK. A novel convolutional neural network framework based solar irradiance prediction method. Int J Electr Power Energy Syst. 2020;114.
60. Ordóñez FJ, Roggen D. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. Sensors. 2016;16(1):115.
61. Rodríguez F, Genn M, Fontán L, Galarza A. Very short-term temperature forecaster using MLP and N-nearest stations for calculating key control parameters in solar photovoltaic generation. Sustainable Energy Technol Assess. 2021;45.
62. Morshed A, Aryal J, Dutta R. Environmental spatio-temporal ontology for the Linked open data cloud. In: 2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications. IEEE; 2013. p. 1907-12.
63. CEDA. Centre for Environmental Data Analysis. CEDA Archive; 2020. Available from: https://archive.ceda.ac.uk/.
64. CSIRO. The Commonwealth Scientific and Industrial Research Organisation; Bureau of Meteorology. WCRP CMIP5: The CSIRO-BOM team ACCESS1-0 model output collection.: Centre for Environmental Data Analysis; 2017. Available from: https://catalogue.ceda.ac.uk/uuid/98a933094fa44e8cb886649cf3f5ba4c.
65. MetOffice. Met Office Hadley Centre. WCRP CMIP5: Met Office Hadley Centre (MOHC) HadGEM2-CC model output collection. Centre for Environmental Data Analysis. 2012. Available from: https://catalogue.ceda.ac.uk/uuid/2e4f5b3748874c61a265f58039898ea5.
66. MetKorean. Meteorological Research Institute of the Korean Meteorological Administration. WCRP CMIP5: Meteorological Research Institute of KMA MRI-CGCM3 model output collection. Centre for Environmental Data Analysis. 2013. Available from: https://catalogue.ceda.ac.uk/uuid/e51a799218a3447d8fde4c7d113dc550.
67. Govindasamy TR, Chetty N. Machine learning models to quantify the influence of PM10 aerosol concentration on global solar radiation prediction in South Africa. Cleaner Engineering and Technology. 2021;2.
68. Komer B, Bergstra J, Eliasmith C. Hyperopt-sklearn: automatic hyperparameter configuration for scikit-learn. In: ICML workshop on AutoML. vol.9. Citeseer; 2014. p.50.
69. Li Y, Jiang J, Gao J, Shao Y, Zhang C, Cui B. Efficient Automatic CASH via Rising Bandits. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol.34; 2020. p. 4763-71.
70. Yang L, Shami A. On hyperparameter optimization of machine learning algorithms: theory and practice. Neurocomputing. 2020;415:295–316.
71. Nomura M, Watanabe S, Akimoto Y, Ozaki Y, Onishi M. Warm starting CMA-ES for hyperparameter optimization. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol.35; 2021. p. 9188-96.
72. Park H, Kim J, Kim M, Kim JH, Choo J, Ha JW, etal. VisualHyperTuner: visual analytics for user-driven hyperparameter tuning of deep neural networks. In: Demo at SysML Conference; 2019. .
73. Shawky OA, Hagag A, El-Dahshan ESA, Ismail MA. Remote sensing image scene classification using CNN-MLP with data augmentation. Optik. 2020;221.
74. Ide H, Kurita T. Improvement of learning for CNN with ReLU activation by sparse regularization. In: 2017 International Joint Conference on Neural Networks (IJCNN). IEEE; 2017. p. 2684-91.

75. Zou F, Shen L, Jie Z, Zhang W, Liu W. A sufficient condition for convergences of Adam and RMSProp. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2019. p. 11127-35.
76. Zhang Z. Improved Adam optimizer for deep neural networks. In: 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS). IEEE; 2018. p. 1–2.
77. Ward R, Wu X, Bottou L. AdaGrad stepsizes: sharp convergence over nonconvex landscapes. In: International Conference on Machine Learning. PMLR; 2019. p. 6677-86.
78. Tieleman T, Hinton G. Divide the gradient by a running average of its recent magnitude. coursera: neural networks for machine learning. Technical Report. 2017.
79. Ketkar N. Introduction to Keras. In: Deep learning with Python; 2017. p. 97-111.
80. Chollet F, etal. Keras: the Python deep learning library. Astrophysics Source Code Library. 2018:ascl-1806.
81. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, etal. Tensorflow: a system for large-scale machine learning. In: 12th {USENIX} symposium on operating systems design and implementation ({ OSDI} 16); 2016. p. 265-83.
82. Bergstra J, Yamins D, Cox DD, etal. Hyperopt: a Python library for optimizing the hyperparameters of machine learning algorithms. In: Proceedings of the 12th Python in science conference. vol.13. Citeseer; 2013. p.20.
83. Chand B, Nguyen-Huy T, Deo RC. Artificial neural networks for prediction of Steadman Heat Index. In: Intelligent data analytics for decision-support systems in hazard mitigation. Singapore: Springer; 2021. p. 293-357.
84. Ghimire S, Deo RC, Raj N, Mi J. Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms. Appl Energy. 2019;253.
85. Willmott CJ, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim Res. 2005;30(1):79–82.
86. Prasad R, Charan D, Joseph L, Nguyen-Huy T, Deo RC, Singh S. Daily flood forecasts with intelligent data analytic models: multivariate empirical mode decomposition-based modeling methods. In: Intelligent data analytics for decision-support systems in hazard mitigation. Singapore: Springer; 2021. p. 359-81.
87. Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans ASABE. 2007;50(3):885–900.
88. Ghimire S, Yaseen ZM, Farooque AA, Deo RC, Zhang J, Tao X. Streamflow prediction using an integrated methodology based on convolutional neural network and long short-term memory networks. Sci Rep. 2021;11(1):1–26.
89. Ghimire S. Predictive modelling of global solar radiation with artificial intelligence approaches using MODIS satellites and atmospheric reanalysis data for Australia. University of Southern Queensland; 2019.
90. Willmott CJ. On the validation of models. Phys Geogr. 1981;2(2):184–94.
91. Chai T, Draxler RR. Root mean square error (RMSE) or mean absolute error (MAE)?-Arguments against avoiding RMSE in the literature. Geosci Model Dev. 2014;7(3):1247–50.
92. Gupta HV, Kling H, Yilmaz KK, Martinez GF. Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. J Hydrol. 2009;377(1–2):80–91.
93. McKenzie J. Mean absolute percentage error and bias in economic forecasting. Econ Lett. 2011;113(3):259–62.
94. Diebold FX, Mariano RS. Comparing predictive accuracy. J Bus Econ Stat. 2002;20(1):134–44.
95. Costantini M, Pappalardo C. Combination of forecast methods using encompassing tests: an algorithm-based procedure. Reihe Ökonomie/Economics Series; 2008.
96. Roy DK, Lal A, Sarker KK, Saha KK, Datta B. Optimization algorithms as training approaches for prediction of reference evapotranspiration using adaptive neuro fuzzy inference system. Agric Water Manag. 2021;255.
97. Ertekin C, Yaldiz O. Comparison of some existing models for estimating global solar radiation for Antalya (Turkey). Energy Convers Manag. 2000;41(4):311–30.
98. Faisal AF, Rahman A, Habib MTM, Siddique AH, Hasan M, Khan MM. Neural networks based multivariate time series forecasting of solar radiation using meteorological data of different cities of Bangladesh. Results in Engineering. 2022;13.
99. Akkaya B, Çolakoğlu N. Comparison of multi-class classification algorithms on early diagnosis of heart diseases. 2019.
100. Sharkawy AN. Principle of neural network and its main types. J Adv Appl Computat Math. 2020;7:8–19.
101. Deo RC, Grant RH, Webb A, Ghimire S, Igoe DP, Downs NJ, et al. Forecasting solar photosynthetic photon flux density under cloud cover effects: novel predictive model using convolutional neural network integrated with long short-term memory network. Stoch Env Res Risk A. 2022:1–38.
102. Prasad SS, Deo RC, Downs N, Igoe D, Parisi AV, Soar J. Cloud affected solar UV prediction with three-phase wavelet hybrid convolutional long short-term memory network multi-step forecast system. IEEE Access. 2022;10:24704–20.
103. Chaturvedi D, Isha I. Solar power forecasting: a review. Int J Comput Appl. 2016;145(6):28–50.
104. Deo RC, Şahin M, Adamowski JF, Mi J. Universally deployable extreme learning machines integrated with remotely sensed MODIS satellite predictors over Australia to forecast global solar radiation: a new approach. Renew Sust Energ Rev. 2019;104:235–61.