Cloud cover bias correction in numerical weather models for solar energy monitoring and forecasting systems with kernel ridge regression

Ravinesh C. Deo[a,*], A. A. Masrur Ahmed[a,b], David Casillas-Pérez[c], S. Ali Pourmousavi[d], Gary Segal[f], Yanshan Yu[f], Sancho Salcedo-Sanz[e]

[a]*School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield, QLD, 4300, Australia.*
[b]*Department of Infrastructure Engineering, The University of Melbourne, Victoria, 3010, Australia.*
[c]*Department of Signal Processing and Communications, Universidad Rey Juan Carlos, Fuenlabrada, 28942, Madrid, Spain.*
[d]*The University of Adelaide, School of Electrical and Electronic Engineering, Australia*
[e]*Department of Signal Processing and Communications, Universidad de Alcalá, Alcalá de Henares, 28805, Madrid, Spain.*
[f]*CS Energy, Level 2, HQ North Tower, 540 Wickham St Fortitude Valley, QLD, 4006, Australia.*

Abstract

Prediction of Total Cloud Cover (TCDC) from numerical weather simulation models, such as Global Forecast System (GFS), can aid renewable energy engineers in monitoring and forecasting solar photovoltaic power generation. A major challenge is the systematic bias in TCDC simulations induced by the errors in the numerical model parameterization stages. Correction of GFS-derived cloud forecasts at multiple time steps can improve energy forecasts in electricity grids to bring better grid stability or certainty in the supply of solar energy. We propose a new kernel ridge regression (KRR) model to reduce bias in TCDC simulations for medium-term prediction at the inter-daily, e.g., 2–8 day-ahead predicted TCDC values. The proposed KRR model is evaluated against multivariate recursive nesting bias correction (MRNBC), a conventional approach and eight machine learning (ML) methods. In terms of the mean absolute error (MAE), the proposed KRR model outperforms MRNBC and ML models at 2-8 day ahead forecasts, with MAE $\approx$ 20–27%. A notable reduction in the simulated cloud cover mean bias error of 20–50% is achieved against the MRNBC and reference accuracy values generated using proxy-observed and non-corrected GFS-predicted TCDC in the model's testing phase. The study ascertains that the proposed KRR model can be explored further to operationalize its capabilities, reduce uncertainties in weather simulation models, and its possible consideration for practical use in improving solar monitoring and forecasting systems that utilize cloud cover simulations from numerical weather predictions.

Keywords: Solar energy generation, Bias correction, Numerical weather models, Global Forecast System, cloud cover study, solar radiation prediction.
PACS: 0000, 1111
2000 MSC: 0000, 1111

1. Introduction

Since its first advent by Richardson in 1922 [1], Numerical Weather Prediction (NWP) models have become the gold standards in real-time weather forecasting. Systematic errors due to physical processes, however, are not addressed correctly in NWP models, and are usually parameterized. This issue induces a significant model bias in several simulated variables such as cloud movements and rainfall. The fidelity of NWP models are largely associated

*Corresponding author: (Prof. Ravinesh C. Deo) ravinesh.deo@usq.edu.au
Email addresses: ravinesh.deo@usq.edu.au (Ravinesh C. Deo), abulmasrur.ahmed@unimelb.edu.au (A. A. Masrur Ahmed), david.casillas@urjc.es (David Casillas-Pérez), a.pourm@adelaide.edu.au (S. Ali

with model design factors, such as incorrectly parameterized physical equations and internal variability of these NWP type models [2]. To utilize NWP simulated variables for operational purposes such as storms or cyclone prediction, climate change and other atmospheric studies, data pre-processing methods are required to significantly reduce the simulated biases [3, 4]. One particular practical use of forecasted cloud cover, particularly over multiple forecast horizons from NWP models, lies in solar irradiance monitoring for a given area, that has in turn applications in rooftop solar and solar farm photovoltaic (PV) power output predictions. Accurate forecasting of solar PV outputs will ensure smooth operation of the electricity grids by allowing effective operational planning with prior information on energy supply intermittencies due to cloud movements. To implement this, NWP-based cloud cover forecasts without significant bias are essential [5].

The Total Cloud Cover (TCDC) is a chief cause of significant intermittency in solar energy supply since a PV panels output can drop down as much as 60% in a few seconds due to a cloud band [6]. This can also happen for the case of the sun travelling across the sky obscured by a passing cloud band, causing major fluctuations in direct normal irradiance reaching a solar PV panel, with the subsequent drop in power generation. Furthermore, a cloudy day can also impact the solar PV output in a much different way as the passing clouds affect solar energy production [6]. Therefore, accurate cloud forecasts over short-term (i.e., sub-hourly, hourly, inter-hourly) and medium-term (i.e., daily or inter-daily) scales have industry implications in solar energy monitoring. To support decisions regarding the sustainability of solar power supply and its integration into electricity grids, reliable forecasts of cloud cover are crucial [7, 8].

Typically, TCDC is defined as the fraction of the sky covered by all visible clouds [9], so, unlike the other weather variables such as temperature and precipitation, the TCDC observational datasets are different in terms of their characteristics [7]. TCDC is also very difficult to monitor over a wide range of spatial scales using physical apparatus, and therefore, are often utilized from NWP model simulations. For example, the movement of clouds over a solar PV panel can be relatively stochastic (i.e., rapidly changing, unpredictable, or intermittent). These uncertain features can no doubt hamper solar energy production and supply rates, so it is highly desirable to construct a better understanding of the features present in total clouds that affect a solar energy generation system.

This paper proposes a new Machine Learning (ML) method to correct bias produced in cloud cover forecasts derived from Global Forecast System (GFS) weather simulation model [10]. Maintained by the National Centre for Environment Prediction, the GFS model is a physics-based system with $0.25^\circ \times 0.25^\circ$ grid resolution with three hourly (3h) temporal resolution for data produced each day. The GFS model simulates the cloud cover, 2-meter height temperature, zonal and meridional wind speed, downward shortwave radiation flux and other atmospheric variables. The GFS model outputs are employed in solar PV prediction modules, for example, in the *pvlib* [11] package that is adopted by electricity industries to monitor their solar generation potentials. In particular, *pvlib* is a python-based community-supported tool with sets of functions and classes to simulate the output of a solar PV system using predicted cloud movements. Developed by the Sandia National Laboratories, *pvlib* [11] provides solar positions, clear sky irradiance, irradiance transposition, direct current power and direct current-to-alternating current power conversions, and therefore, has found applications in the solar energy industry [12, 13]. Apart from *pvlib* [11], there are other types of solar photovoltaic energy prediction software including but not limited to, *Solpy*, *Pandapower*, *Pyleecan*, *Scipy*, *Numpy*, and *Matplotlib* [14]. While these tools could be useful predictive modules in solar energy monitoring systems, they require GFS or other NWP model simulated clouds to estimate the direct normal irradiance. However, significant bias in predicted clouds (or other variables related to solar irradiance) lead to inaccurate prediction of solar energy and therefore, add to generation and demand imbalance in real-time; hence higher electricity prices.

In order to incorporate forecasted cloud cover or weather model variables in solar monitoring systems, reducing the bias in these variables has traditionally focused on correcting the individual variable representations across a single time (e.g., daily, monthly). However, these corrections aim to determine the bias in a statistical or a quantile sense and, therefore, utilise corrected data for future scenarios of solar energy production.

Daily and monthly standardization can address systematic biases in the means and the variances of simulated variables [15, 16] to support renewable energy generation applications. Bias correction with non-parametric approaches such as quantile matching [17, 18, 19, 20] and equidistant quantile [21] was found to be successful methods in reducing errors in weather model variables. Still, a major shortcoming of such techniques is that they tend to examine only the bias in the distribution of GFS (or another model) without considering the impact of its persistence, which continues to influence the accuracy of simulated variables [22].

We refer to the study of Johnson and Sharma [23] that suggests nested bias correction (NBC) approach can reduce the variability and persistence at different time scales. Also, techniques like multivariate bias correction (MBC) [24, 25], copula-based bias correction [26], empirical copula bias correction (EC-BC) [27], distribution transfer methods [20], power transformation methods [28, 29, 30] and local intensity scaling methods [30, 31] have been utilized in many spatial locations to correct bias in weather variables. To the best of the authors' knowledge, no prior method has successfully eliminated the biases, given that relationships between simulated and observed variables are relatively complex [32]. To address this problem, ML has thus been demonstrated as an alternative method to model highly non-linear features in simulated variables relative to observations or proxy-observed variables [33, 34, 35, 25]. Based on their promising performance, ML is therefore becoming a potential tool to correct bias in numerical weather variables [25].

The promise of ML arises from its capability to discover the associations between predictors and a target variable without considering the underlying physical system's operation [36, 37, 38]. This black-box method is advantageous in reducing the mathematical complexity of a physical model by using pattern recognition that is better understood in contrast to a physical model employing partial differential equations with a fixed set of initial conditions [39, 40]. The initial conditions in physical models are somewhat difficult to predict accurately over a wide range of spatial and temporal domains. One type of ML model, the artificial neural networks (ANN), has previously been applied to correct inter-instrument bias [41, 42]. On the other hand, support vector machine (SVM) with its theoretical foundations in statistical learning has also been recognized as a sophisticated ML tool [43, 44] with SVM models using a kernel-based ANN to address the drawbacks of a conventional model [45]. Due to the use of kernel functions, SVMs are therefore quite resilient and efficient in non-linear modelling of noisy data [33, 35].

This study, therefore, adopts an alternative form of ML algorithms known as kernel ridge regression (KRR) for bias correction of the Total Cloud Cover forecasts from the GFS-based numerical weather model. The proposed KRR method [46] integrates kernel functions and ridge regressions to better capture the non-linear correlative features to address regression-based over-fitting issues found in other methods [47]. The KRR method uses a regularized variant of a least-square method to learn the global feature extraction functions; hence, it can potentially predict any target variable with greater accuracy compared to other ML models. Although ML has previously been used in bias correction, the proposed technique remains somewhat under-explored. More generally, the KRR method has been used in other prediction problems, including precipitation [48], drought [49], wind speed [50, 51, 52, 53, 54] and also solar power [55] and thus has offered a significant advantage in terms of computational simplicity relative to a conventional SVM or other ML models.

The novelty of this study is (i) to develop for the first time a KRR-based bias correction model for Total Cloud Cover forecasts (TCDC) at 2-8 day ahead forecast horizons at a solar

energy farm in Queensland, Australia, (ii) to specifically test the capability of a KRR model in reducing the errors in TCDC forecasts found in the GFS-derived TCDC forecasts, (iii) to benchmark the proposed KRR model in respect to the multivariate recursive nesting bias correction approach as a widely used conventional method and the reference values generated by proxy-observed and non-corrected GFS-predicted TCDC in the model's testing phase. To fulfil this aim, we adopt two distinct modelling strategies: Firstly, the KRR model is trained using 2-m height temperature, 10-m zonal (U)-wind, 10-m meridional (V)-wind, downward shortwave radiation flux, and Total Cloud Cover forecasts that are regressed against proxy-observed (i.e., GFS-Analysis) data. Secondly, only the cloud cover data (i.e., $\text{TCDC}_{\text{GFS-Forecast}}$) are incorporated as single inputs (with $\text{TCDC}_{\text{GFS-Analysis}}$ as a target variable) to test the overall performance of this alternative method to particularly reduce the bias in cloud cover forecasts.

To ascertain its practicality, the proposed KRR model is compared with conventional bias correction methods based on multivariate recursive nested bias correction (MRNBC) [25] and ML methods using Bayesian ridge regression (BNR) [56], Decision Tree Regression (DTR) [57], Gradient Boosting Regressor (GBR) [58], Histogram-based Gradient Boosting Regressor (HGBR) [59], k-nearest neighbour regression (KNN) [35], multivariate adaptive regression splines (MARS) [60], extreme gradient boosting (XGB) and random forest (RF) [58] as competing methods to benchmark the KRR model. Finally, the KRR model is tested at inter-daily time horizons using Day 2 to Day 8 cloud cover forecasts using real solar farm data (Columboola Solar Farm in Queensland, Australia) to test the developed predictive system for its application in solar generation monitoring and supporting industry decisions to manage the solar power supply in the national electricity grid.

The rest of the paper has been structured in the following way: the next section presents the materials and methods, which includes a description of the data and study area, a summary of the GFS capabilities and the proposed KRR and its adaptation for bias correction of Total Cloud Cover. Section 3 presents the simulation study, discussing different experiments and comparisons versus alternative ML approaches such as KNN, MARS or Random Forest.

2. Materials and Methods

2.1. Study Area

We implement a newly developed KRR model for cloud cover bias correction for a solar farm in Queensland, referred to as Australia's "Sunshine State", with enormous solar energy potential [61, 62]. Under United Nations Sustainable Development Goal #7 (SDG7) [63], the State government is committed to increasing renewable energy uptake by up to 50% of the overall future energy supply by 2030. These projects represent an investment of $8.5 billion, the creation of 7000 jobs, the installation of 4600 MW of renewable energy production and a reduction of more than 11 million tonnes of CO_2. As of January 2021, Queensland had 6200 MW of renewable plants, including rooftop solar systems. According to the government, renewable energy fulfils 20% of electricity consumed [64], which is expected to increase to 50% by 2030. To improve the existing methodologies that can assist the solar energy producers, this study considers the case of $\text{TCDC}_{\text{GFS-Forecast}}$ obtained at Columboola Solar Farm in Queensland, Australia. This solar farm, with 417,000 solar PV modules, is expected to produce $\approx$ 440 GWh of energy annually after its completion in 2022, provide electricity to 6 % of all homes in the state, create hundreds of regional jobs and produce enough electricity for 75,000 homes for 35 years.

Figure 1 shows the geographic location of the study site where the proposed KRR model for cloud cover bias correction was implemented. Table 1 lists GFS-forecast variables (i.e., 2-metre height temperature, 10-metre wind speed, Total Cloud Cover, and Downward Shortwave Radiation Flux) used as inputs for the proposed model and the GFS analysis variable (i.e., Total Cloud Cover) used as the proxy of the observed data.

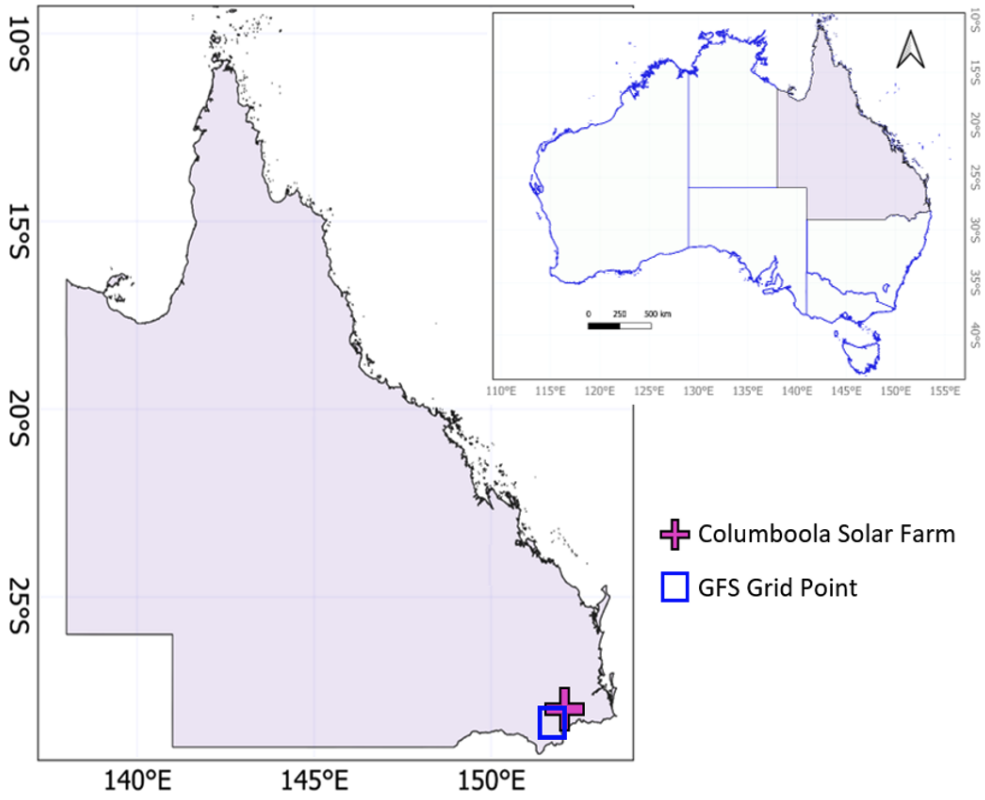

Figure 1: Geographic location of study site: Columboola solar energy farm in Queensland, Australia, where the proposed kernel ridge regression (KRR)-based ML model for bias correction of TCDC was developed utilizing the Global Forecast System (GFS) analysis (i.e., proxy-observed) and forecasted variables.

Table 1: List of Global Forecast System (GFS)-forecast variables (i.e., 2-metre temperature, 10-metre wind speed, Total Cloud Cover, and Downward Short-wave Radiation Flux) used as KRR model inputs, and GFS analysis variable (i.e., Total Cloud Cover used as proxy-observed) in the proposed KRR model used in bias correction problem.

Variable Short Name	**Variable Description**	**Level**	**Units**
KRR Model Inputs: GFS Forecast (Inputs)			
$T2m_{\text{GFS-Forecast}}$	2-metre temperature	Height Above Ground	K
$U_{\text{GFS-Forecast}}$	10-metre U wind component	Height Above Ground	ms^{-1}
$V_{\text{GFS-Forecast}}$	10-metre V wind component	Height Above Ground	ms^{-1}
$TCDC_{\text{GFS-Forecast}}$	Total Cloud Cover	Atmosphere	%
$DSWRF_{\text{GFS-Forecast}}$	Downward short-wave radiation flux	Surface	Wm^{-2}
KRR Model Target: GFS Analysis (proxy-observed)			
$TCDC_{\text{GFS-Analysis}}$	Total Cloud Cover	Atmosphere	%

2.2. Global Forecasting System Cloud Cover and Meteorological Data-sets

We develop KRR model using GFS data-set that are managed by National Oceanic and Atmospheric Administration (NOAA) which aims to deliver an operational set of global weather predictions [65]. The GFS forecast system aims to produce forecast variables up to 16 days in advance with a temporal resolution of 3h and 6h, and a spatial resolution of $0.25^{\text{o}}\times0.25^{\text{o}}$ [66]. The GFS is not a frozen system, so its dynamic core and physical package are modified regularly [67]. For example, after a single-member prediction was replaced by a GFS ensemble mean forecast in late 2001, this method was modified again in late 2003 to properly incorporate the bias-corrected GFS ensemble mean forecast [68, 69].

As this physics-based model is initialised every three hours, newly predicted variables are generated eight times a day at 0 UTC, 3 UTC, 6 UTC, 9 UTC, 12 UTC, 15 UTC, 18 UTC, 21 UTC, and 24 UTC. The GFS utilises Global Data Assimilation System (GDAS) [70] that augments a gridded three-dimensional model space with surface observations, balloon data, wind profiler data, buoy observations, radar observations, or satellite observations. The GDAS model output is generated four times daily and includes projections for the next three hours, six hours, and nine hours.

The present study builds a new modelling strategy to correct the inherent bias in GFS-derived TCDC forecasts (i.e., $\text{TCDC}_{\text{GFS-Forecast}}$) for 3 distinct forecast horizons, which according to Queensland daytime zones (i.e., UTC + 10), are: at 0 UTC (10 AEST), 3 UTC (13 AEST), and 6 UTC (16 AEST). The 3-h GFS experiments, initialized from 0000 UTC compared to AEST (Australian Eastern Standard Time), are illustrated schematically in Figure 3. For comparison, the GFS-analysis Total Cloud Cover ($\text{TCDC}_{\text{GFS-Analysis}}$) is used as a proxy for the observed cloud cover generated by the GFS model. We also utilised temperature ($\text{T2m}_{\text{GFS-Forecast}}$), downward shortwave radiation flux ($\text{DSWRF}_{\text{GFS-Forecast}}$), wind speed ($\text{U}_{\text{GFS-Forecast}}$, and $\text{V}_{\text{GFS-Forecast}}$) to reduce the bias through our newly proposed KRR modelling strategies.

2.3. Theoretical Overview of Kernel Ridge Regression

This section details the proposed KRR model whereas Appendix B shows the details of the conventional bias correction MRNBC method. For details of comparison models, readers can consult several other sources [57, 71, 72, 73, 56, 35, 60, 58, 59]. In general, KRR is a novel algorithm with an unlimited number of non-linear transformations of the independent variables used as regressors [74]. KRR model utilises ML strategy based on kernel and ridge regressions [46] to avoid issues of overfitting found in other regression methods. It, therefore, utilizes regularizations and a kernel technique to capture non-linear connections viz [49].

$$\arg\min \frac{1}{q}\sum_{o=1}^{q}\|f_o - y_o\|^2 + \lambda\|f\|_H^2 \tag{1}$$

$$f_o = \sum_{p=1}^{q}\alpha_p\omega(x_p, x_o) \tag{2}$$

The Hilbert normed space of Equation 1 is defined as $\|\cdot\|_H$ and α is the Lagrange multiplier. For a given $m \times m$ kernel matrix, K is developed by $\omega(x_p, x_o)$ from some fixed predictor variables where y is the input $q \times 1$ regression vector and is the $q \times 1$ unknown situation vector that reduces as follows:

$$y = (K + \lambda qI) \tag{3}$$

$$\tilde{y} = \sum_{p=1}^{q}\alpha_o\omega(x_o, \tilde{x}) \tag{4}$$

In model training stage, KRR technique is applied by solving Equation (3) but utilised to predict the regression of an unknown sample x in Equation (4) in the testing stage. To achieve the highest accuracy possible, linear, polynomial, and Gaussian kernels are employed [47, 75, 76].

2.4. Implementation of Machine Learning (ML)-based Bias Correction

The fundamental idea behind bias correction is to identify a sufficiently adaptable and flexible approach that is capable of learning from available data and then constructing a prediction function that performs well across the projection period (i.e., forecast horizon). To perform robust bias corrections, it was critical first to optimise the architecture of the proposed KRR model, and then to take advantage of the associative links between the bias-corrected TCDC and the fully learned ML model.

An ML-based Python package [77], scikit-learn [78, 79], was thus employed to develop the the proposed KRR and other benchmark models (i.e., BNR, DTR, GBR, HGBR, KNN, MLR, XGB, and RF). For the case of MARS model, we have used the py-earth package, and programming software R for traditional bias correction (i.e., MRNBC) as applied by Yang et al. [25] for correction of bias in global climate models. As we define in Section 2.5, six statistical measures are used to evaluate the experimental outcome of the bias-corrected model, created using Intel i7 processor running at 3.6GHz and 16 GB RAM. Visualisation of bias-corrected TCDC dataset were made through matplotlib [80], seaborn [81] and Microsoft Excel.

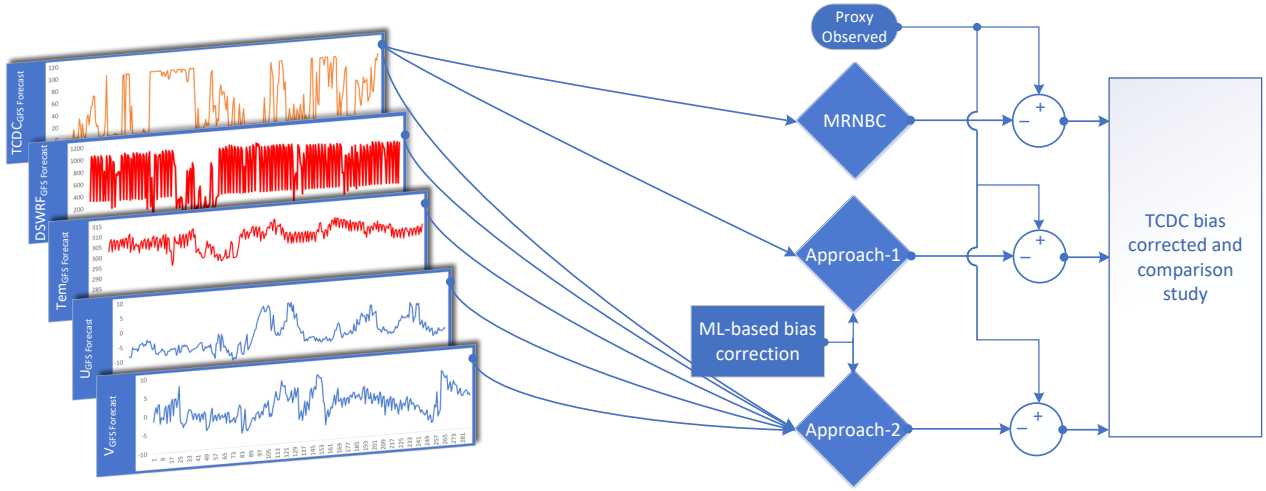

Figure 2: A schematic of the proposed KRR bias correction method benchmarked against conventional MRNBC and nine ML (i.e., BNR, DTR, GBR, HGBR, KNN, MARS, XGB, and RF) models. Interpretive Statement: The proposed KRR bias correction method uses: (i) Approach-1 taking in five GFS outputs: i.e., $\text{TCDC}_{\text{GFS-Forecast}}$, Downward Short-wave Radiation Flux $\text{DSWRF}_{\text{GFS-Forecast}}$, 2-meter temperature ($\text{T2m}_{\text{GFS-Forecast}}$), zonal $\text{U}_{\text{GFS-Forecast}}$ and meridional $\text{V}_{\text{GFS-Forecast}}$ against the Total Cloud Cover $\text{TCDC}_{\text{GFS-Analysis}}$ (or the reference or proxy-observed value) as the target, (ii) Approach 2 taking in $\text{TCDC}_{\text{GFS-Forecast}}$ as an input with $\text{TCDC}_{\text{GFS-Analysis}}$ as a target based on which the bias needs to be corrected.

Figure 2 is a schematic representation of KRR-based bias correction approach including the conventional (i.e., multivariate recursive nested bias correction, MRNBC) methods. In summary, the proposed KRR method is implemented as follows:

1. **Data:** GFS-forecast and GFS-analysis data were downloaded from NCEP repository [82]. As this repository provides 384-hours ahead data at a 3-hr interval, this study has only measured three time periods within the Brisbane daytime zone considering the relevance to solar PV power production at 0 UTC, 3 UTC, and 6 UTC.

   Figure 3 shows a schematic illustration of 3-h GFS forecast experiments initialized at 0000 UTC, compared with the Australian Eastern Standard Time (AEST). We adopted

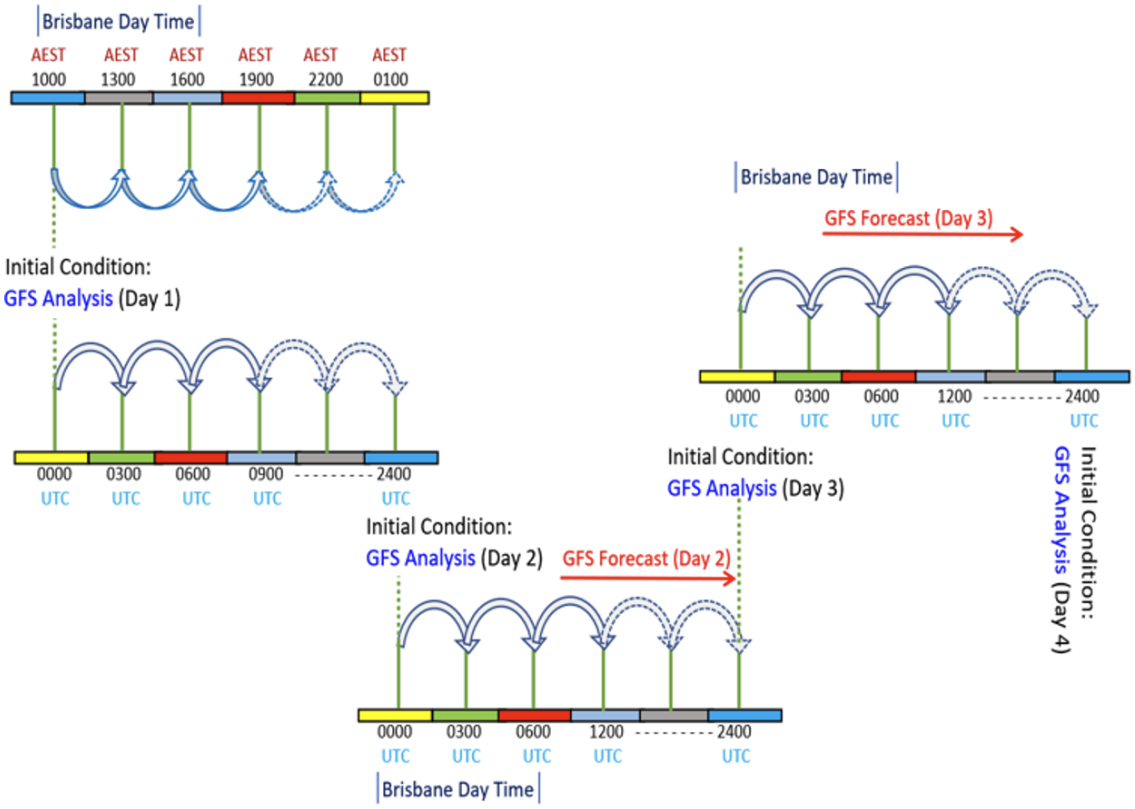

Figure 3: Schematic illustration of the 3-h GFS forecasts initialized at 0000 UTC compared with Australian Eastern Standard Time (AEST) used to develop the proposed KRR bias correction method.

the pygrib python package to extract five selected variables and the datasets were sorted for Day 2 to Day 8 forecast. To apply the bias correction method, we adopted the $\text{TCDC}_{\text{GFS-Analysis}}$ dataset as a proxy for the observation and used these to correct the systemic biases that were present in the $\text{TCDC}_{\text{GFS-Forecast}}$ dataset.

Table 2 shows the descriptive statistics of the GFS forecast and the GFS analysis data-set used to develop the proposed KRR model.

2. **Pre-possessing and post-processing:** Missing values were replaced using the preceding seven data points and all data normalised to be bounded by [0, 1] [83]. As the TCDC dataset has significant zero values as a normal feature of cloud properties due to the presence or absence of cloud, this aspect can affect an ML model's performance. We have therefore used four normalization techniques with the best normalization technique selected based on the minimum mean absolute error (MAE). The normalization techniques trialled were: max-min normalization (T_{MinMax}), maximum absolute normalization (T_{MaxAbs}), z-score normalization (T_{Std}), and robust scaler normalization (T_{Robust}) with their mathematical formulations stated as follows:

   (a) Max-min normalization (T_{MinMax}):

$$T_{MinMax} = \frac{(T_i - T_{min})}{(T_{max} - T_{min})} \tag{5}$$

   (b) z-score normalization (T_{Std}):

$$T_{Std} = \frac{T_i - \bar{T}_l}{Std} \tag{6}$$

Table 2: Descriptive statistics of GFS forecast and GFS analysis (i.e., proxy of the observed) data used to develop the proposed KRR model. Data were acquired from GFS model over January 1, 2019 and April 30, 2020 used for training 70% and testing (30%) where the 15% of the training set is specifically used for model validation.

Variable	Forecast Horizon	Max	Min	Mean	Skewness	Kurtosis
$\mathbf{DSWRF_{GFS}}$ Forecast	Day 2	1100	0.00	601.07	-0.22	-1.38
	Day 3	1100	0.00	605.30	-0.23	-1.46
	Day 4	1100	0.00	595.55	-0.20	-1.47
	Day 5	1100	0.00	595.71	-0.20	-1.46
	Day 6	1100	0.00	599.78	-0.20	-1.39
	Day 7	1090	0.00	604.91	-0.24	-1.44
	Day 8	1100	0.00	605.01	-0.27	-1.42
$\mathbf{TCDC_{GFS}}$ Forecast	Day 2	100	0.00	27.82	1.01	-0.56
	Day 3	100	0.00	29.38	0.91	-0.74
	Day 4	100	0.00	32.80	0.73	-1.04
	Day 5	100	0.00	32.95	0.73	-1.05
	Day 6	100	0.00	32.62	0.70	-1.12
	Day 7	100	0.00	31.88	0.77	-0.96
	Day 8	100	0.00	33.87	0.66	-1.11
$\mathbf{T2m_{GFS}}$ Forecast	Day 2	314.55	285.38	301.64	-0.31	-0.62
	Day 3	314.76	285.36	301.57	-0.35	-0.59
	Day 4	313.59	285.24	301.49	-0.33	-0.67
	Day 5	314.74	284.35	301.45	-0.34	-0.61
	Day 6	314.65	284.76	301.53	-0.33	-0.54
	Day 7	315.22	285.20	301.45	-0.34	-0.55
	Day 8	313.45	285.54	301.70	-0.45	-0.42
$\mathbf{U_{GFS}}$ Forecast	Day 2	10.49	-12.23	-4.25	0.99	0.94
	Day 3	7.38	-13.03	-3.50	0.49	-0.37
	Day 4	8.56	-11.41	-4.37	1.08	1.09
	Day 5	8.80	-12.24	-4.37	1.02	0.95
	Day 6	8.83	-10.67	-4.46	1.13	1.25
	Day 7	10.93	-11.93	-4.52	1.19	1.74
	Day 8	8.85	-13.19	-4.05	0.66	0.01
$\mathbf{V_{GFS}}$ Forecast	Day 2	10.29	-7.74	0.14	0.22	-0.08
	Day 3	10.06	-9.55	-0.70	-0.03	-0.34
	Day 4	8.53	-7.08	0.09	0.25	-0.10
	Day 5	8.65	-7.22	0.12	0.31	-0.03
	Day 6	9.57	-6.64	0.03	0.30	-0.10
	Day 7	8.58	-10.66	-0.07	0.22	0.10
	Day 8	13.70	-7.37	-0.22	0.21	0.35
$\mathbf{TCDC_{GFS}}$ Analysis	Day 2	100	0.00	31.70	0.78	-1.01
	Day 3	100	-5.83	31.82	0.78	-1.02
	Day 4	100	-5.83	31.89	0.77	-1.03
	Day 5	100	-5.83	31.95	0.77	-1.03
	Day 6	100	-5.83	31.95	0.77	-1.03
	Day 7	100	-5.83	31.92	0.77	-1.03
	Day 8	100	-5.83	32.02	0.76	-1.04

(c) Maximum Absolute normalization (T_{MaxAbs}):

$$T_{MaxAbs} = \frac{T_i}{Max(Abs(x))} \tag{7}$$

(d) Robust scaler normalization (T_{Robust}):

$$T_{Robust} = \frac{T_i - T_\omega}{Q_3 - Q_1} \tag{8}$$

where T_i are respective predictors, $\bar{T_l}$ is the average of T_i, T_{min} is the minimum value for predictors, T_{max} is the maximum value and Std is the standard deviation, T_ω is the median of T_i and $(Q_3 - Q_1)$ is the interquartile range between 1st quartile (25th) and 3rd quartile (75th) quantile. As there is no specific rule for data partitioning [83, 84], we used 70% training, 15% testing with a validation set as the last 15% of the training set for all data collected between 1 January 2019 and 30 April 2020.

3. **Implementation of ML-based Bias Correction:** This study has developed a total of 10 different models (i.e., the proposed KRR model along with nine other benchmark models) to correct the bias in $\text{TCDC}_{\text{GFS-Forecast}}$ for data over Day 2 to Day 8 forecasts. Our MARS model considers multivariate data with basis functions to investigate the predictor variable and identifies the predictor and target features [85]. The DTR is a non-parametric, supervised system to approximate a sine curve using 'if-then-else' decision where generally, the deeper the tree, the more complicated a rule could be to fit a model. A prime task of ML is to set hyper-parameters for optimal bias correction method, so an optimum architecture of the KRR model was created using GridSearchCV (regularization strength, $\alpha = 1.5$; gamma parameter is fixed to None, with a degree of the polynomial kernel is 3 and the kernel is linear; see Table 3). The performance of ML bias correction was compared with traditional bias corrections (i.e., MRNBC), and the reference value usually calculated between $\text{TCDC}_{\text{GFS-Forecast}}$ and $\text{TCDC}_{\text{GFS-Analysis}}$ was used with $\text{TCDC}_{\text{GFS-Analysis}}$ considered as the proxy of the observed cloud cover dataset.

4. **Implementation of MRNBC Bias Correction Method:** We now detail the procedure developed to correct bias using the MRNBC method, which is a traditional non-ML approach used previously. We made univariate adjustments followed by multivariate corrections using a time series with appropriate bias correction statistics generated for all variables and locations. Therefore, the MRNBC method corrected the bias in $\text{TCDC}_{\text{GFS-Forecast}}$ by removing the current GFS mean and adding the observed mean. The time series adjusted in Step-2 are standardised, and this residual time series is adapted for bias using auto and cross-correlations for day lag-1 and lag-0. To summarise the corrections necessary at each time scale, a weighting factor may also be computed. The $\text{TCDC}_{\text{GFS-Forecast}}$ daily time series is multiplied by the weighting factor from each time scale to produce the final bias-corrected time series. The MRNBC bias correction procedure is schematized in Figure 4.

5. **Two Different Approaches for Bias Correction** We adopt two different approaches to correct the bias in GFS-based cloud cover predictions. The first approach, denoted as Approach-1 in this paper, integrates five GFS data series comprised of $\text{TCDC}_{\text{GFS-Forecast}}$, $\text{T2m}_{\text{GFS-Forecast}}$, $\text{DSWRF}_{\text{GFS-Forecast}}$, $\text{U}_{\text{GFS-Forecast}}$ and $\text{V}_{\text{GFS-Forecast}}$) that are used as the proposed KRR model's input variables. This approach utilizes the exogenous meteorological variables that are used to reduce the bias in the predicted TCDC. The second approach, denoted as Approach-2, uses a single matrix $\text{TCDC}_{\text{GFS-Forecast}}$ data-set where

Table 3: The optimal hyper-parameters of the proposed KRR model, including that of the other benchmark models include machine learning (i.e., BNR, DTR, GBR, HGBR, KNN, MARS, MLR, and RF)

Model Type	Name	Hyper-parameters	Acronym	Optimum
Objective Model	KRR	Regularization strength	alpha	1.5
		Kernel mapping	kernel	linear
		Gamma parameter	gamma	None
		Degree of the polynomial kernel	degree	3
		Zero coefficient for polynomial and sigmoid kernels	coef0	1.2
Benchmark Machine Learning Models	BNR	Maximum number of iterations	n_iter	200
		Stop the algorithm if w has converged	tol	0.0001
		Shape parameter for Gamma distribution over alpha	alpha_1	1e-05
		Inverse scale parameter over alpha	alpha_2	1e-05
		Shape parameter for Gamma distribution over lambda	lambda_1	1e-06
		Inverse scale parameter for Gamma distribution over lambda	lambda_2	1e-04
		The initial value for alpha	alpha_init	None
	DTR	Maximum depth of the tree	max_depth	None
		Minimum number of samples for an internal node	min_sample_split	2
		Number of features for the best split	max_features	Auto
	GBR	Number of boosting stages	n_estimators	102
		Minimum number of samples for an internal node	min_sample_split	2
		Learning rate	learning_rate	0.1
		Maximum depth of individual regression estimators' estimators	max_depth	3
		Number of features to consider for the best split	max_feature	None
	HGBR	Learning rate	learning_rate	0.1
		Maximum number of iterations	max_iter	120
		maximum number of leaves for each tree	max_leaf_nodes	31
		Maximum number of bins	max_bins	260
	KNN	Number of neighbours	n_neighbors	5
		Weights	Weights	uniform
		The algorithm used to compute the nearest neighbours	algorithm	auto
		Leaf-size passed	leaf_size	30
		Power parameter for the Minkowski metric	p	2
		The distance metric to use for the tree.	metric	minkowski
		Additional keyword arguments for the metric	metric_params	none
		The number of parallel jobs	n_jobs	int
	MARS	maximum degree of terms	max_degree	1
		Smoothing parameter used to calculate GCV	penalty	3.0
	RF	Number of trees in the forest	n_estimators	120
		Maximum depth of the tree	max_depth	2
		Minimum number of samples for an internal node	min_sample_split	2
		Number of features for the best split	max_features	auto

historical patterns and the persistence are used to reduce the bias in the predicted TCDC produced by the GFS model. Both approaches use TCDC analysis data-set as the proxy of the observed variable generated by the GFS Numerical Weather Prediction Model. To arrive at the optimal method used in reducing bias in the predicted TCDC, we have examined 10 models (nine based on ML and MRNBC-based conventional model) to identify the best bias correction performance in comparison with the reference values between $\text{TCDC}_{\text{GFS-Forecast}}$ and $\text{TCDC}_{\text{GFS-Analysis}}$ for the present study site.

2.5. Evaluation of ML-Based Bias Correction Method

The effectiveness of the proposed KRR model, including all of the ML-based and conventional bias correction methods employing the reference value (calculated between $\text{TCDC}_{\text{GFS-Analysis}}$ and $\text{TCDC}_{\text{GFS-Forecast}}$) is evaluated. We adopt a range of performance metrics such as the Pearson's Correlation Coefficient (r), root mean square error (RMSE) and mean absolute error (MAE) in the testing phase where $\text{TCDC}_{\text{GFS-Analysis}}$ (i.e., the proxy-observed) and corrected $\text{TCDC}_{\text{GFS-Forecast}}$ datasets are compared). In its most general sense, the effectiveness of any model is determined by the agreement between the corrected (i.e., TCDC) and the proxy-observed ($\text{TCDC}_{\text{GFS-Analysis}}$) data. While RMSE is a more appropriate measure of performance than MAE when the error distribution is Gaussian [86], for a more persuasive model, the Willmott's Index (WI) [87, 88, 89] and Legates–McCabe's Index (LM) [90, 91, 92] are employed in this study.

Mathematically, these are expressed as follows:

Correlation coefficient (r):

$$r = \frac{\sum_{i=1}^{n}(TCDC_{BC} - \overline{TCDC}_{ANL})(TCDC_{BC} - \overline{TCDC}_{BC})}{\sqrt{\sum_{i=1}^{n}(TCDC_{ANL} - \overline{TCDC}_{ANL})^2}\sqrt{\sum_{i=1}^{n}(TCDC_{BC} - \overline{TCDC}_{BC})^2}} \tag{9}$$

Mean absolute error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|TCDC_{BC} - TCDC_{ANL}| \tag{10}$$

Root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(TCDC_{BC} - TCDC_{ANL})^2} \tag{11}$$

Willmott's Index of Agreement (d):

$$d = 1 - \frac{\sum_{i=n}^{n}(TCDC_{BC} - TCDC_{ANL})^2}{\sum_{i=n}^{n}(|TCDC_{BC} - \overline{TCDC}_{ANL}| + |TCDC_{ANL} - \overline{TCDC}_{ANL}|)^2} \tag{12}$$

Legates –McCabe's Index (LM):

$$LM = 1 - \frac{\sum_{i=1}^{n}|TCDC_{BC} - TCDC_{ANL}|}{\sum_{i=1}^{n}|TCDC_{ANL} - \overline{TCDC}_{ANL}|} \tag{13}$$

Mean Absolute Percentage Deviation (MAPD: %):

$$MAPD = \frac{100}{n}\sum_{i=1}^{n}\frac{|TCDC_{BC} - TCDC_{ANL}|}{TCDC_{ANL}} \tag{14}$$

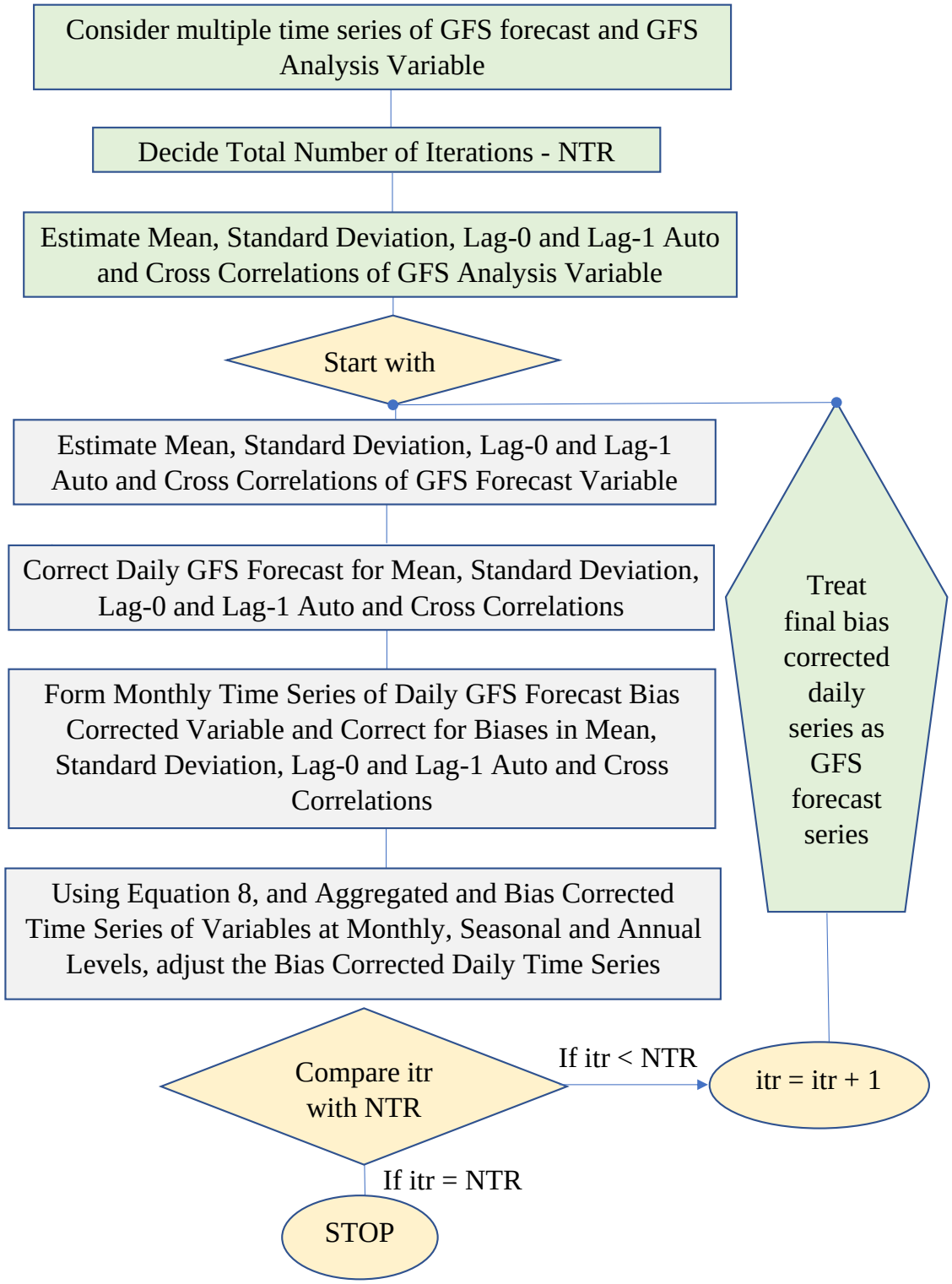

Figure 4: Schematic of the conventional MRNBC method presented in this study as a comparison method against the proposed KRR bias correction method used to correct bias in TCDC.

where $TCDC_{ANL}$ and $TCDC_{BC}$, respectively, represents the proxy of the observed ($TCDC_{GFS-Analysis}$) and bias-corrected data series for i^{th} tested value, and $\overline{TCDC}_{ANL}$ and $\overline{TCDC}_{BC}$ refer to their average values, accordingly. The number of observations is denoted by N, while the coefficient of variation is denoted by CV.

In comparing the different models adopted for this bias correction problem, this study uses promoting percentage of the Legate-McCabe's Index (Δ_{LM} (%)) as a complementary measure

of the model efficiency. The Δ_{LM} (%) is calculated comparing the actual LM obtained using the proposed KRR and LM values generated by the KNN, MARS, and RF models. Mathematically, the Δ_{LM} (%) is computed as follows:

$$\Delta_{LM}(\%) = \frac{LM_{KRR} - LM_{COM}}{LM_{KRR}} \times 100 \tag{15}$$

where LM_{COM} represents the LM value of the benchmark model (e.g., KNN, MARS, or RF).

3. Results and Discussion

The practicality of the proposed KRR model for bias correction is established using two distinct approaches as shown previously in Figure 2. We now evaluate the amount of bias that has been reduced by applying these approaches considering $\text{TCDC}_{\text{GFS-Forecast}}$ data relative to the proxy-observed ($\text{TCDC}_{\text{GFS-Analysis}}$) data using the proposed KRR model. All of the comparative ML models (BNR, DTR, GBR, HGBR, KNN, KRR, MARS, MLR, XGB, and RF) are also assessed using statistical metrics (Equations 10-14), infographics and visualisations to determine the degree of agreement between the corrected $\text{TCDC}_{\text{GFS-Forecast}}$ and the proxy-observed variable ($\text{TCDC}_{\text{GFS-Analysis}}$). Overall, the performance metrics indicate that the proposed KRR model has outperformed all of the alternative models in the testing phase, which is also demonstrated by a superior value of r and d and a low value of RMSE and MAE in the independent testing phase discussed in the following section.

3.1. Boxplots for the Distribution of Errors After Bias Reduction

According to the results presented in Figures 5 and 6, an in-depth examination of Willmott's Index (d) and the root mean squared error (RMSE) provides persuasive evidence that the proposed ML approaches offer substantial benefits in reducing the bias compared with the traditional MRNBC method and the respective reference values tested for all the forecast days over which the GFS Total Cloud Cover forecast is considered. This figure clearly shows the closer distribution of RMSE and d values for the case of ML models using Approach-2 (see Figures 5b and 6b) compared with Approach-1 (Figures 5a and 6a). The lower end of the plot for the value of d is relatively situated within the lower quartile (25^{th}) and the upper quartile (75_{th}) range for the Day 2 GFS forecast data series.

There appears also to be a single outlier found further than the the 75^{th} percentile. However, for Day 3 to Day 8 GFS forecasts, the bias correction of $\text{TCDC}_{\text{GFS-Forecast}}$ time series results in a lesser improvement, except for Day 6 forecasts. This is reasonable as the uncertainties in TCDC are likely to increase with an increment in the forecast horizon. Noticeably, as the forecasting period changes from Day 2 to Day 8, the performance of our bias correction model decreases significantly. Despite this, we can note from Figures 5 and 6 that ML models can be considered the most potent strategy for bias correction at solar farms, at least for the present study site and the suite of models considered.

Further analysis is performed through a boxplot of errors (i.e., RMSE) for results obtained through Approach-2. This shows the bias-corrected Total Cloud Cover vs. $\text{TCDC}^{\text{GFS-Analysis}}$ of all the ML models as illustrated in Figure 5b. For Day 2 $\text{TCDC}_{\text{GFS-Forecast}}$ data series, it is noticeable that the dispersion of RMSE for bias correction methods concerning the quartile values has distinct outliers. The lower end of the boxplot seems to lie precisely between the lower quartile (25^{th} percentile) and upper quartile (75^{th} percentile).

Likewise, the correlation coefficient (d) and RMSE are higher for the other days (Day 2 to Day 8) forecast except for Day 6. Therefore, the improvement of bias using ML methods signifies improved performance compared with the MRNBC and the respective reference values of the $\text{TCDC}_{\text{GFS-Forecast}}$ and $\text{TCDC}_{\text{GFS-Analysis}}$. When data from the other models were compared, the accuracy of KRR-based bias correction outweighed those of the other ML models (see Figure 5).

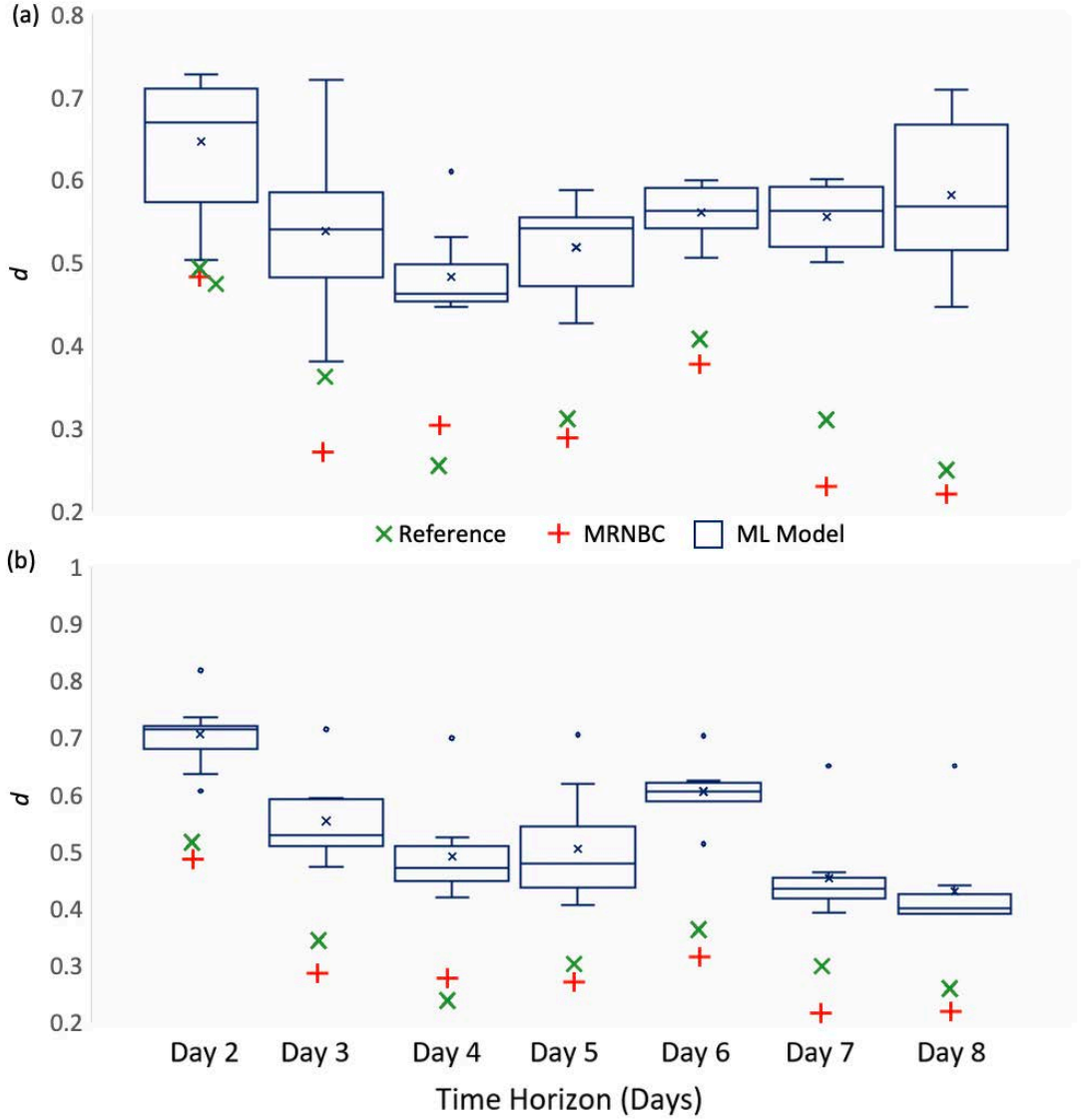

Figure 5: Box plots of the d values calculated for nine ML-bias corrections models (i.e., KRR, BNR, DTR, GBR, HGBR, KNN, MARS, RF, XGB) pooled together including conventional MRNBC method with their respective reference d value calculated from $\text{TCDC}_{\text{GFS-Forecast}}$ and $\text{TCDC}_{\text{GFS-Analysis}}$). (a) Approach-1, (b) Approach-2. [For details on each approach, see Figure 2]

The boxplots of bias-corrected RMSE calculated between data for all the nine ML-based bias correction methods pooled together (i.e., KRR, BNR, DTR, GBR, HGBR, KNN, MARS, RF, XGB), conventional bias correction method (i.e., MRNBC) and along with their respective reference values (RMSE calculated between $\text{TCDC}_{\text{GFS-Forecast}}$ and $\text{TCDC}_{\text{GFS-Analysis}}$) are also shown in Figure 6. When used to correct the TCDC simulations, it appears that the proposed KRR model with Approach-2 (see Figure 2 produces the lowest MAE values compared with the other ML models for the same approach and the reference value method.

For Approach-2, the MAE value generated for Day 2 forecast is bounded by [20.20, 26.75] %, with the best value obtained for the proposed KRR indicating a modest 14% improvement over the reference MAE value. A similar reduction in the cloud cover bias is notable for the cloud cover forecasts generated for the Day 3 over to the Day 7 horizons.

It is imperative to note that Approach-1, which employs a MARS model, was more effective in correcting the TCDC bias for the Day 8 cloud cover forecasts relative to Approach-2. Consequently, the ML-based KRR model outperforms the classic bias correction strategy in correcting the GFS-derived TCDC. In accordance with this result, the four best methods (i.e.,

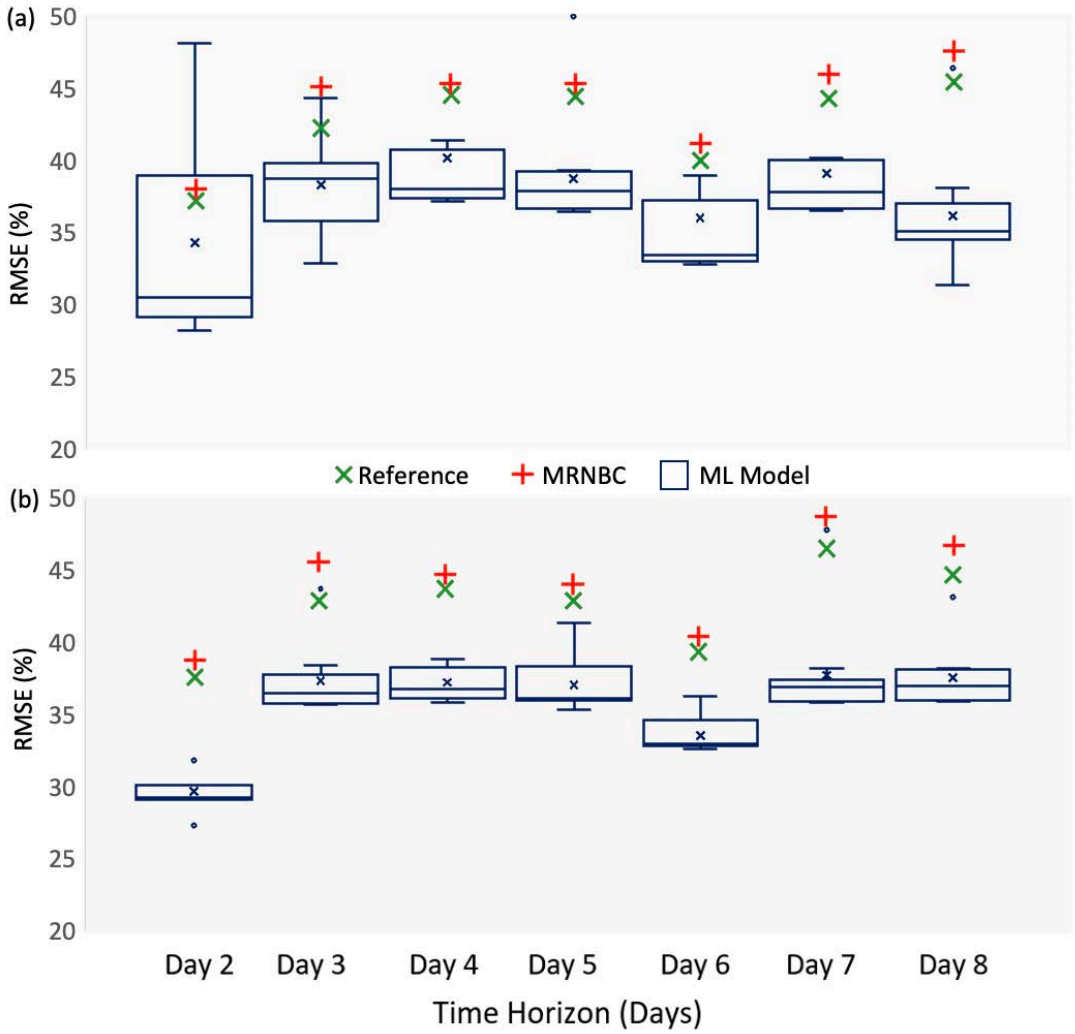

Figure 6: Box plots of the bias-corrected RMSE calculated between data for all ML-based bias correction methods pooled together (i.e., KRR, BNR, DTR, GBR, HGBR, KNN, MARS, RF, XGB), conventional MRNBC method along with their respective reference RMSE calculated between $\text{TCDC}_{\text{GFS-Forecast}}$ and $\text{TCDC}_{\text{GFS-Analysis}}$). (a) Approach-1, (b) Approach-2. [For details on each approach, see Figure 2].

KNN, KRR, MARS, and RF) were then chosen to conduct an in-depth examination of the bias correction approaches utilizing these machine learning models.

To further demonstrate the proposed KRR model's capability to correct the bias in the $\text{TCDC}_{\text{GFS-Forecast}}$ data generated for Day 2-8 forecast horizons, we now show the LM values between corrected cloud cover forecasts and proxy-observed cloud cover forecasts generated by the GFS model. Here, we aim to compare a metric known as the promoting percentage, which is an incremental performance in the model based on the value of LM ($\Delta_L M$,%) derived from the benchmark model against the proposed objective (i.e., KRR) model.

Figure 7 shows the above results of the proposed KRR model against that of the KNN, MARS, and the RF model applied to correct the bias in TCDC data for Day 2 to Day 8 forecast horizons. The bias correction outcomes for the proposed KRR model relative to the other models, is relatively diverse. Notwithstanding this, Figure 7 shows that the effectiveness of the bias correction using the proposed KRR method is more significantly notable by 20% to 65% for all the predicted days. Overall, the highest gain in respect to the accuracy appears to have been reached by ≈70% for the proposed KNN model for the case of 4-day ahead forecasting of Total Cloud Cover.

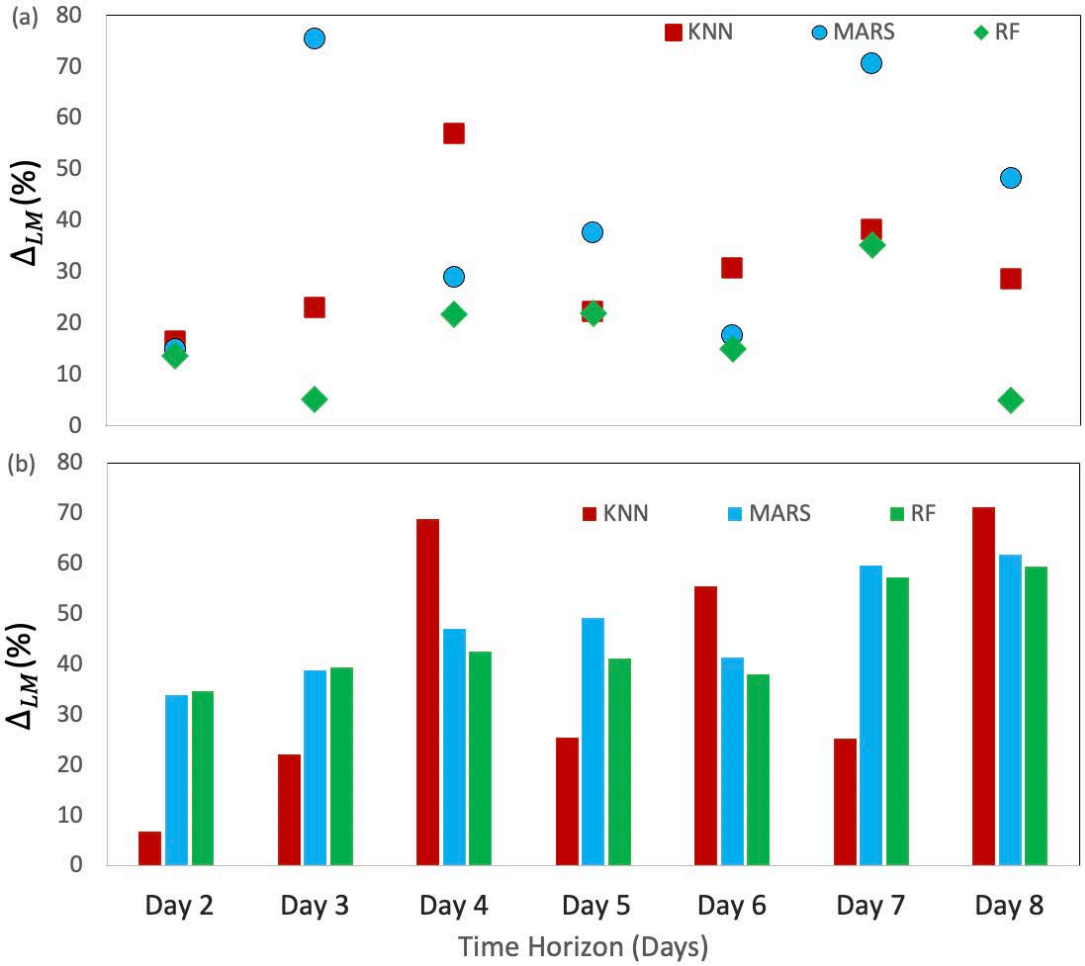

Figure 7: Percentage change in LM that compares its values obtained using the proposed KRR model with respect to KNN, MARS and RF models. (a) Approach-1, (b) Approach-2. Note that: $\Delta_{LM}(\%) = \frac{LM_{KRR} - LM_{COM}}{LM_{KRR}} \times 100$. Note: LM_{COM} represents the LM value of the benchmark (KNN, MARS or RF) model. [For details on each approach, see Figure 2].

3.2. Percentage Reduction in Bias

To investigate the performance of ML-based bias correction and specifically check the performance of the proposed KRR model, the MAE values for all of the tested models is listed in Table 4, along with traditional bias correction method (MRNBC) and the reference value method.

Table 4 shows the MAE (%) computed between the 'proxy-observed' ($\text{TCDC}_{\text{GFS-Analysis}}$) and ML-bias corrected TCDC using the proposed KRR model. Note that here, the conventional bias correction method used is the multivariate recursive nesting bias correction (MRNBC) method, whereas the benchmark ML methods include the BNR, DTR, GBR, HGBR, KNN, MARS, MLR, and the RF model (see Table 3.

It is important to note that in Approach 2, the proposed KRR model outperforms all of the ML, MRNBC and reference value datasets for TCDC forecasts over Days 2-8 forecast horizons based on its lowest error value. For example, for Day 2 forecasts of the predicted TCDC, the proposed KRR model produces an error value that is $\approx$ 13.8 % lower than the reference value comparing the TCDC forecasts and the TCDC analysis variable. Likewise, the bias in TCDC is reduced by $\approx$ 2.9 %, 13.4 %, 9.7 %, 12.3 %, 13.3 % and 13.5 % for Day 3, Day 4, Day 5, Day 6, Day 7 and Day 8, respectively. This shows that the proposed KRR model developed using $\text{TCDC}_{\text{GFS-Forecast}}$ as a predictor with $\text{TCDC}_{\text{GFS-Analysis}}$ as the target variable, which also outperforms the conventional MRNBC method, performs consistently in terms of reducing the

Table 4: The MAE (%) computed between 'proxy-observed' ($TCDC_{GFS\text{-}Analysis}$) and ML-bias corrected TCDC used to evaluate the proposed KRR model. Note Approach-1 uses $T2m_{GFS\text{-}Forecast}$, $V_{GFS\text{-}Forecast}$, $U_{GFS\text{-}Forecast}$, $TCDC_{GFS\text{-}Forecast}$, and $DSWRF_{GFS\text{-}Forecast}$ whereas Approach-2 uses $TCDC_{GFS\text{-}Forecast}$ as a predictor against $TCDC_{GFS\text{-}Analysis}$ as target variable. The reference MAE is computed between $TCDC_{GFS\text{-}Forecast}$ and $TCDC_{GFS\text{-}Analysis}$ data to provide additional benchmarks for the proposed KRR bias correction method. Note: the best bias correction model has been boldfaced.

Model and Method		**GFS Inter-daily Forecast Horizon**						
		Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8
Error comparing $TCDC_{GFS\ Forecast}$ and $TCDC_{GFS\ Analysis}$ datasets	Reference	23.45	29.36	32.93	31.49	27.59	31.68	32.36
Conventional Bias Correction Method	MRNBC	25.90	32.05	32.65	32.76	30.28	33.57	34.50
	Approach-1							
Objective Model	KRR	**25.07**	34.56	32.23	31.33	**27.68**	**30.76**	30.26
Benchmark Models	BNR	25.35	31.90	32.93	32.63	29.08	32.41	31.31
	DTR	35.65	30.47	41.35	37.00	38.24	37.98	34.46
	GBR	32.52	31.68	34.32	32.38	29.85	31.73	28.77
	HGBR	32.45	32.39	34.15	30.95	30.73	33.18	28.77
	KNN	26.76	**29.90**	**30.32**	**30.48**	29.98	32.20	31.31
	MARS	26.60	26.18	33.21	32.77	28.99	33.40	**24.80**
	RF	25.19	32.14	32.84	32.52	28.94	32.27	31.16
	XGB	26.47	30.74	32.96	32.17	28.80	32.08	30.08
	Approach 2							
Objective Model	KRR	**20.20**	**28.75**	**28.52**	**28.44**	**24.20**	**27.47**	**27.99**
Benchmark Models	BNR	25.32	31.63	31.89	31.78	28.77	31.57	31.69
	DTR	26.75	32.22	33.19	31.82	29.23	31.55	32.74
	GBR	25.81	31.73	32.36	31.27	28.52	31.36	31.82
	HGBR	25.91	31.70	32.24	31.55	28.37	31.46	32.19
	KNN	21.22	38.64	33.39	36.67	30.29	41.85	38.18
	MARS	25.36	31.46	31.85	31.75	28.74	31.67	31.66
	RF	25.28	31.60	31.85	31.75	28.74	31.54	31.66
	XGB	25.48	31.50	31.52	31.20	28.36	31.49	31.52

bias in GFS-based predicted cloud cover generated over multiple forecast horizons.

For the case of Approach-1 that that has used meteorological variables such as $T2m_{GFS\text{-}Forecast}$, $V_{GFS\text{-}Forecast}$, $U_{GFS\text{-}Forecast}$, $TCDC_{GFS\text{-}Forecast}$ and $DSWRF_{GFS\text{-}Forecast}$ produced by the GFS model and the $TCDC_{GFS\text{-}Analysis}$ produced as the target variable, the best performance of the proposed KRR model is noted for Day 2, Day 6 and Day 7. This performance in terms of error reduction is relatively inferior to Approach 2 in terms of the MAE value. One possibility for the relatively weaker performance of the proposed KRR model when utilizing these exogenous meteorological variables in Approach-1 could perhaps be attributable to the systematic errors within each individual GFS variable and a potentially weaker relationship with $TCDC_{GFS\text{-}Analysis}$ as the target variable. For Day 3, Day 4 and Day 5, the proposed KNN model appears to be the best for Approach-1, although the proposed KRR model in Approach 2 still remains superior than this model.

We now evaluate the robustness of the four top-performing models, which includes the proposed KRR and the KNN, MARS and RF model by using correlation coefficient (r) for Approach-1 and Approach-2. These are plotted together in Figure 8.

Note that a larger r-value is expected to represent a greater degree of agreement between

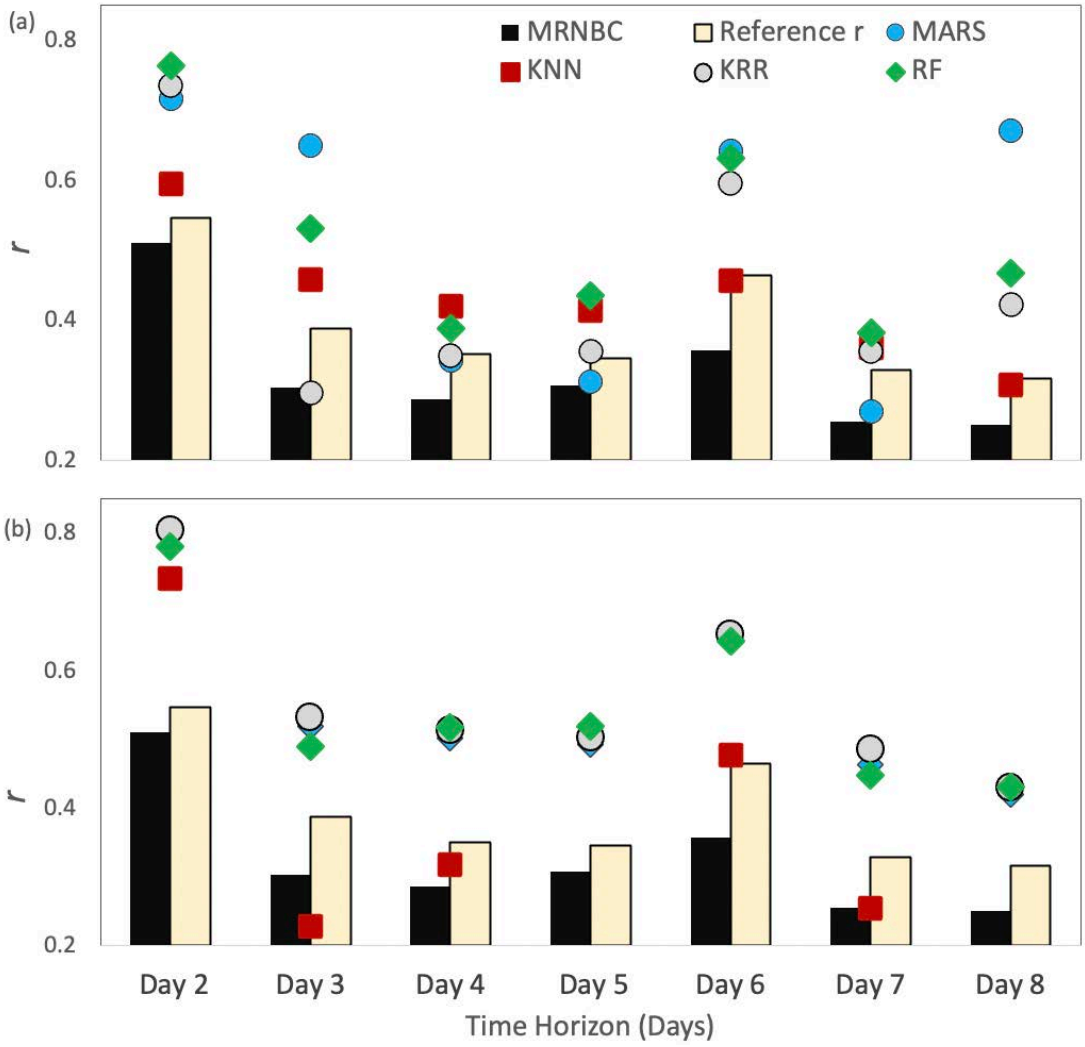

Figure 8: Comparative analysis of selected ML-based bias correction (i.e., KRR, MARS, KNN, RF) methods using correlation coefficient (r) between corrected $\text{TCDC}_{\text{GFS-Forecast}}$ and reference $\text{TCDC}_{\text{GFS-Analysis}}$. Included is a respective reference r-value computed using 'non-corrected' $\text{TCDC}_{\text{GFS-Forecast}}$ and bias-corrected $\text{TCDC}_{\text{GFS-Forecast}}$ using MRNBC method. (a) Approach-1, (b) Approach-2. [For details on each approach, see Figure 2].

corrected $\text{TCDC}_{\text{GFS-Forecast}}$ and reference $\text{TCDC}_{\text{GFS-Analysis}}$. If this is so, the result is expected to show a reduction in the bias within the Total Cloud Cover forecasts generated by the GFS model. Importantly, the results for Approach-2 show consistently higher r-value compared with that of the KRR, MARS, SVR and RF models for all tested forecast horizons from Day 2 to Day 8.

In fact, compared with reference value derived from the 'non-corrected' $\text{TCDC}_{\text{GFS-Forecast}}$ and bias-corrected $\text{TCDC}_{\text{GFS-Forecast}}$, there appears to be a dramatic reduction of 52.2 % in these biases as measured by an increase in r-value for Day 2, which is $\approx$ 38.9 % - 60.1 % for Day 3 to Day 8 forecasts. When compared with the conventional bias correction using MRNBC, we note that the proposed KRR model has generated an increased r-value by $\approx$ 85.1 - 112.6% for Day 3 to Day 8 forecasts.

While the other three ML models have also led to an reduction in the bias in Total Cloud Cover, the magnitude of bluethis change in r-value remains lower when compared with both the MRNBC and the reference r-values. When the results are closely inspected for Approach-1, the proposed KRR model has led to an increase in the r-value (compared against MRNBC) by $\approx$ 55.8-13.8 % for Day 2 to Day 7. However, when compared against then reference r-values, the proposed KRR model increases the r-value by 44.5 % for Day 2, 15.7 % for Day 3, 2.3 %

for Day 5 and 20.8 % for Day 6.

Overall, it is evident that the proposed KRR model developed using Approach-2 where $TCDC_{GFS\text{-}Forecast}$ and $TCDC_{GFS\text{-}Analysis}$ are used in model construction is superior for all forecast horizons and against all of the ML and conventional methods used to reduce the overall bias in Total Cloud Cover forecasts. Because the benchmark models performed poorly, as demonstrated in Figure 8, the newly proposed KRR model is therefore reaffirmed as superior for the present research study site.

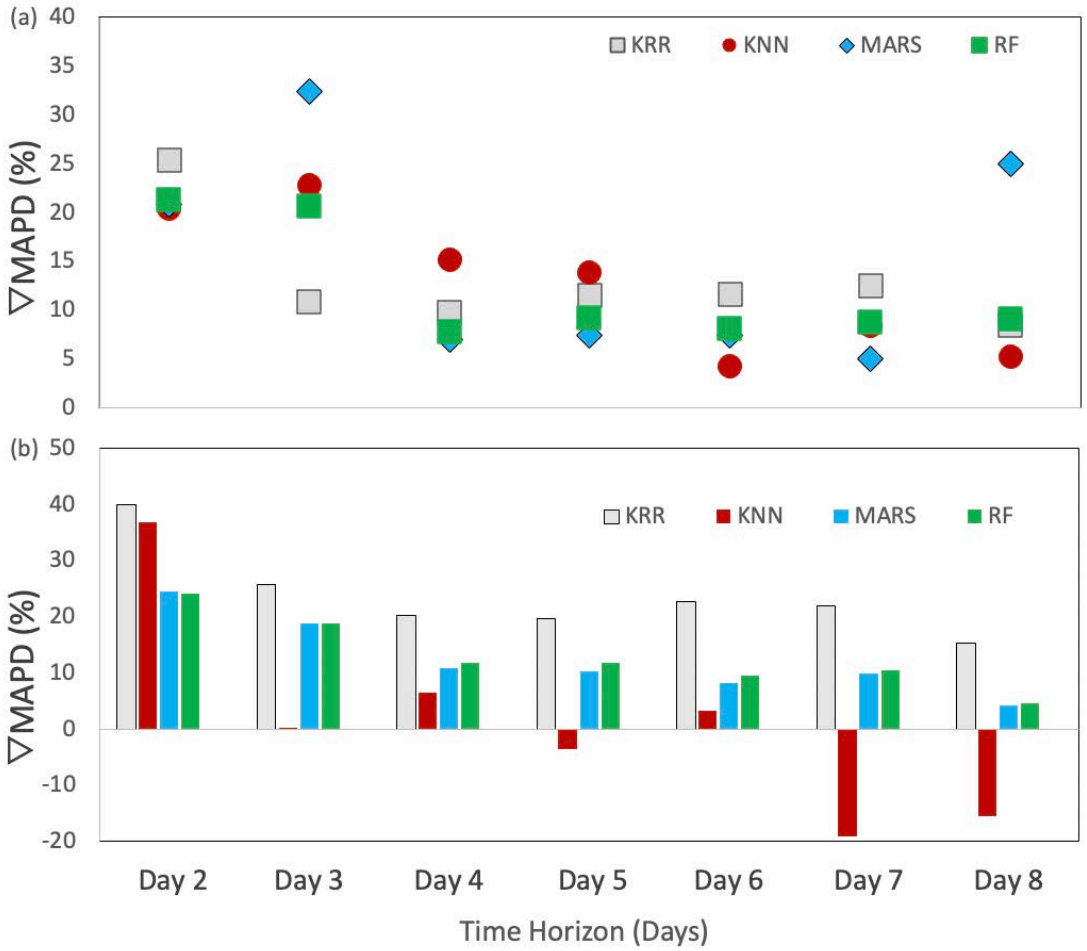

Figure 9: Change (∇) in mean absolute percentage error, MAPD (%) generated by proposed KRR bias correction method against a reference value of MAPD deducted from $TCDC_{GFS\text{-}Forecast}$ and $TCDC_{GFS\text{-}Analysis}$. (a) Approach-1, and (b) Approach-2. [For details on each approach, see Figure 2]. Interpretive statement: a positive change is used to show the objective model outperforms benchmark models.

The change (∇) in MAPD (%) generated by the proposed KRR method compared to the reference value deduced from $TCDC_{GFS\text{-}Forecast}$ and $TCDC_{GFS\text{-}Analysis}$ is presented in Figure 9. A positive change shows the proposed objective model (i.e., KRR) outperforming the benchmark model.

For both approaches, ∇MAPD (%) is significant for Day 2 GFS forecast, whereas Approach 2 with KRR shows the lowest value at $\approx$48%. For Approach-1, the MAE value from an SVR model is $\approx$17.5% higher, whereas ∇MAPD range from [5, 35] % for Day 3 to Day 8 forecasts in Approach-2 with some deviation noted for the KNN model. In a rational sense, the proposed KRR model demonstrates the most significant improvement in MAPD (∇MAPD; %) ranging from 15% to 14% for Day 2 to Day 8 with respect to a reduction in bias for the TCDC dataset. Accordingly, we can ascertain that our newly developed KRR model appears to fall within the criterion of an acceptable predictive model that can correct the bias in GFS-derived Total Cloud

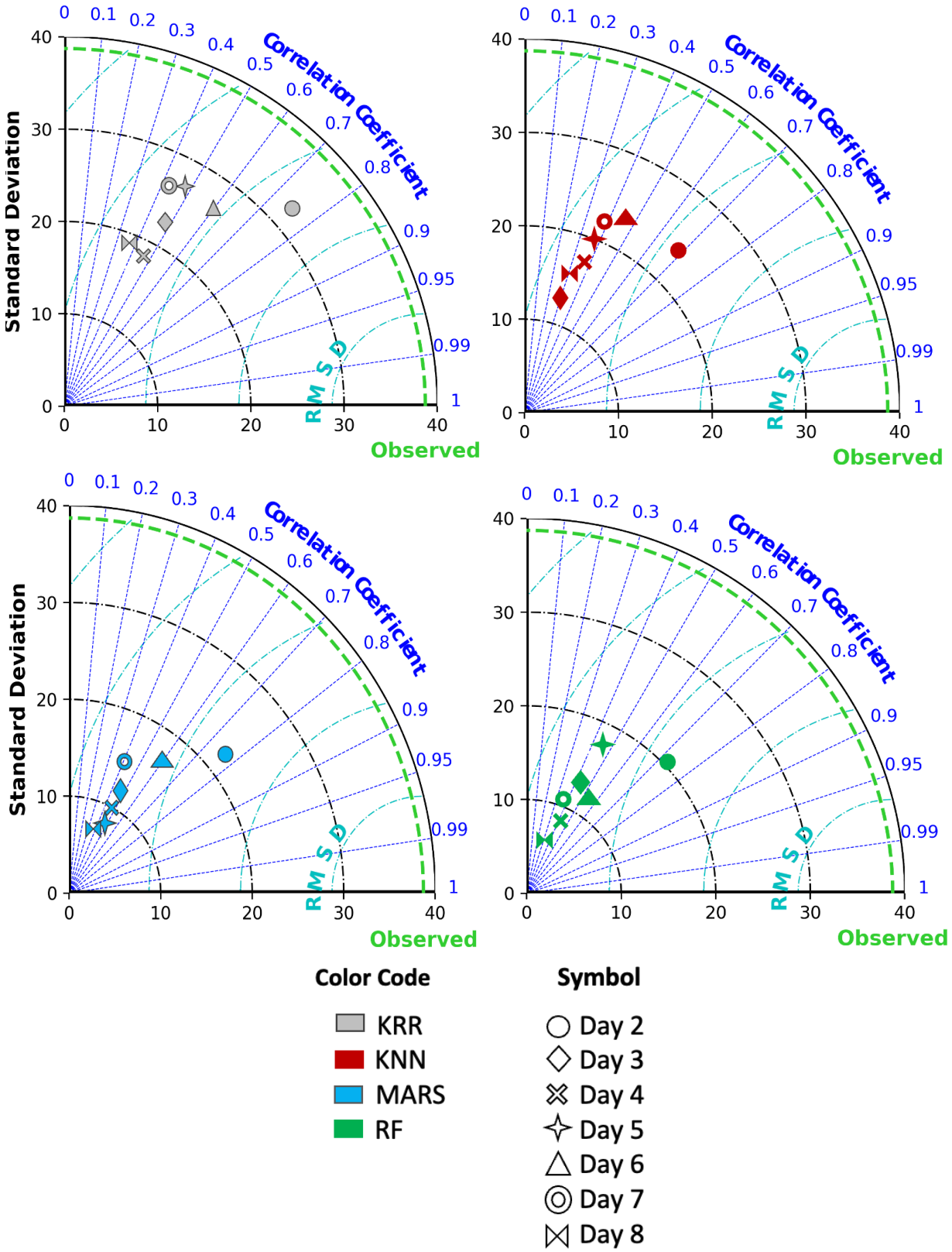

Figure 10: Taylor diagram showing correlation coefficient, standard deviation, and root mean square centered difference (RMSD). (a) The proposed KRR model is compared with: (b) KNN, (c) MARS, (d) RF) for the most accurate approach (i.e., Approach-2). [For details on each approach, see Figure 2].

Cover forecasts. Therefore, it may be a useful tool for solar energy monitoring and forecasting systems.

3.3. Evaluation of Proposed Model Using a Taylor Diagram

We now revert to a Taylor diagram that provides a way of graphically summarizing how closely the model performance matches the observations. Figure 10 is an alternative representation of proposed KRR model's performance compared to the benchmark models using a Taylor diagram [93]. In this case, a significant correlation seems to exist between bias-corrected TCDC and the proxy-observed variable ($\text{TCDC}_{\text{GFS-Analysis}}$) for the case of the proposed KRR model.

It is clear that the bias corrected TCDC data produced from the proposed KRR model is a close match to the proxy of the observed TCDC data relative to the other competing ML models. Therefore, in a nutshell, based on the statistical performance measures, we can ascertain that the newly developed KRR model has the predictive skills to reduce the overall bias in Total Cloud Cover generated by the weather simulation model used in this study.

4. Conclusions, Limitations and Future Research Insights

4.1. Conclusion

This paper utilised ML-based bias correction (i.e., KRR) method to reduce the bias in Total Cloud Cover generated by the GFS numerical weather model at a solar energy farm in Queensland, Australia. To demonstrate the feasibility of the developed KRR model, data from Columboola solar energy farm located in Queensland, Australia, were used. The results indicated a superior performance of the proposed model compared to several machine learning and conventional bias correction methods. We learned that the ML-based bias correction approach had a solid potential to significantly reduce, if not eradicate, the bias in TCDC, by utilising cloud cover, temperature, wind speed and downward solar radiation flux forecasts as covariates for TCDC that provide adequate predictive features and relationships in observed cloud cover variables. Precisely, the KRR model's capability to correct the bias in TCDC dataset was established in terms of the percentage improvement in mean bias error that for this study site has ranged from $\sim$20% to $\sim$50% using the traditional MRNBC method for Day 2 to Day 8 forecast.

The study showed that training a ML model using a single GFS predictor variable (i.e., $\text{TCDC}_{\text{GFS-Forecast}}$ as well as integrating multiple predictor variables (i.e., $\text{T2m}_{\text{GFS-Forecast}}$, $\text{V}_{\text{GFS-Forecast}}$, $\text{U}_{\text{GFS-Forecast}}$ and $\text{DSWRF}_{\text{GFS-Forecast}}$. against the proxy-observed GFS variable (i.e., $\text{TCDC}_{\text{GFS-Analysis}}$) successfully corrected the bias in Total Cloud Cover, albeit with a varying degree of accuracy. These GFS-based predictor variables provided historical information on the cloud evolution against the respective meteorological variables and their lagged stochastic behaviour. Nonetheless, we contend that biases in individual predicted variables from GFS may also affect the accuracy of cloud cover bias correction task. In our study, we found using a single set of model input variables (i.e., $\text{TCDC}_{\text{GFS-Forecast}}$) was better suited compared to the multi-variable approach, such that the results have established high predictive potency of employing a single variable to resolve the bias-related problem for this solar energy site.

These results have shown that the performance of ML-based bias correction for longer-term forecast horizon (i.e., Day 8) was much better in Approach-1 (where multiple predictor variables: $\text{TCDC}_{\text{GFS-Forecast}}$, $\text{T2m}_{\text{GFS-Forecast}}$, $\text{DSWRF}_{\text{GFS-Forecast}}$, $\text{U}_{\text{GFS-Forecast}}$, and $\text{V}_{\text{GFS-Forecast}}$ were incorporated in the KRR model's input matrix). This outcome appears to reveal the interactions of these variables with the proxy-observed cloud cover over the passage of time. This led to an improved overall performance, i.e., for a longer-term Day 8 bias correction result although the multi-variable approach (i.e., Approach-1) registered comparatively large bias

compared with the single variable approach (Approach-1). While the results of this pilot study may not be explicitly conclusive and may require further investigation, one possible explanation for comparatively large bias could be the interference of disproportionately embedded biases within each of these forecast variables that could hinder the correlation among such biases to affect further TCDC produced by the GFS model.

4.2. Limitations and Future Research

In spite of the success of the proposed KRR model in reducing the bias in Total Cloud Cover forecasts generated by GFS model over Day 2 to Day 8 horizons, there remain some limitations. Firstly, this study has tested a single solar energy farm in Queensland, Australia. Further tests of the model including relevant parameter tuning and application at more diverse locations are warranted to fully explore its potential in reducing the bias in cloud cover predictions. Secondly, such tests should also include integrating the bias-corrected cloud cover forecasts into a solar PV monitoring software such as pvlib, Solpy, Pandapower, Pyleecan, Scipy, Numpy, or Matplotlib [11, 11, 14] to check the impact of more accurate forecasts on solar generation monitoring and related economic (e.g. solar energy price bidding) or other benefits. Thirdly, a future study could deep learning algorithms that have exceptional capabilities in terms of extracting more complex data features may offer better performance in correcting bias in real-time weather models used for solar energy monitoring. Some relevance may be drawn from recent studies where deep learning was broadly implemented, for example, in hydrology [37, 39] and solar energy studies [94, 95, 96].

Therefore, in future studies, subject to availability of big atmospheric datasets, a deep learning hybrid approach could be adopted as a bias correction method both for solar power production monitoring and power failure risk analysis when solar energy is integrated into the energy grid. Finally, the exact positioning of the spatial grid over a specific solar farm remains a major challenge if we are to use the bias corrected cloud cover forecasts for solar PV power monitoring as evident in this study where the Columboola solar energy farm was located slightly off-grid from the GFS model. Therefore, exploring other types of NWP models with finer grids, or exploring an ensemble of NWP forecasts to correct the bias in their cloud cover remains an open problem of interest to the solar energy community. Our group's next step in future research is to adopt the Global Ensemble Forecast System or the Australian Community Climate and Earth-System Simulator (ACCESS)-S2/S3) that are NWP model candidates to be used by solar energy companies in the USA or Australia for their intra-daily and inter-daily solar generation capacity prediction, including its effect on electricity sale bidding price in smart grids or their solar-conventional energy supply-demand models.

Credit authorship contribution statement

A. A. Masrur Ahmed: Writing - original draft, Conceptualization, Methodology, Software, Data Curation, Formal analysis, Investigation, Model development and application. Ravinesh C. Deo: Conceptualization, Investigation, Project administration, Writing - review & editing, Investigation, Supervision. Sancho Salcedo-Sanz: Writing - review & editing, Investigation, Supervision. David Casillas Pérez: Writing - review & editing, Investigation. Ali Pourmousavi Kani: Writing - review & editing, Investigation. Gary Segal: Conceptualization. Yanshan Yu: Conceptualization.

Acknowledgement

This study was supported by CS Energy under Australian Postgraduate Research Program (APR.intern) program mentored by Professor Ravinesh Deo (UniSQ), Gary Segal and

Dr Yanshan Yu (CS Energy). We thank Australian Mathematical Science Institute for administration of research partnership between UniSQ and CS Energy. This research has been partially supported by project PID2020-115454GB-C21595 of the Spanish Ministry of Science and Innovation (MICINN) that involves Prof Deo (UniSQ, Australia) and Prof Salcedo-Sanz (Universidad de Alcala, Spain).

References

[1] L. F. Richardson, Weather prediction by numerical process, University Press, 1922.

[2] G. Evin, B. Hingray, J. Blanchet, N. Eckert, S. Morin, D. Verfaillie, Partitioning uncertainty components of an incomplete ensemble of climate projections using data augmentation, Journal of Climate 32 (8) (2019) 2423–2440.

[3] J. H. Christensen, F. Boberg, O. B. Christensen, P. Lucas-Picher, On the need for bias correction of regional climate change projections of temperature and precipitation, Geophysical Research Letters 35 (20) (2008).

[4] P. Vaittinada Ayar, M. Vrac, A. Mailhot, Ensemble bias correction of climate simulations: preserving internal variability, Scientific Reports 11 (1) (2021) 1–9.

[5] J. Zhang, R. Verschae, S. Nobuhara, J.-F. Lalonde, Deep photovoltaic nowcasting, Solar Energy 176 (2018) 267–276.

[6] A. Mills, M. Ahlstrom, M. Brower, A. Ellis, R. George, T. Hoff, B. Kroposki, C. Lenox, N. Miller, J. Stein, et al., Understanding variability and uncertainty of photovoltaics for integration with the electric power system, Tech. rep., Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States) (2009).

[7] Á. Baran, S. Lerch, M. El Ayari, S. Baran, Machine learning for total cloud cover prediction, Neural Computing and Applications 33 (7) (2021) 2605–2620.

[8] D. Matuszko, Influence of the extent and genera of cloud cover on solar radiation intensity, International Journal of Climatology 32 (15) (2012) 2403–2414.

[9] W. M. O. WMO, International cloud atlas: Manual on the observation of clouds and other meteors. WMO-No. 407 (1975).

[10] E. M. Center, The GFS atmospheric model, National Centers for Environmental Prediction Office Note 442 (2003) 14.

[11] W. F. Holmgren, C. W. Hansen, M. A. Mikofski, pvlib python: A python package for modeling solar energy systems, Journal of Open Source Software 3 (29) (2018) 884.

[12] X. Li, F. Wagner, W. Peng, J. Yang, D. L. Mauzerall, Reduction of solar photovoltaic resources due to air pollution in China, Proceedings of the National Academy of Sciences 114 (45) (2017) 11867–11872.

[13] X. Li, D. L. Mauzerall, M. H. Bergin, Global reduction of solar power generation efficiency due to aerosols and panel soiling, Nature Sustainability 3 (9) (2020) 720–727.

[14] R. Sivapriyan, D. Elangovan, K. S. Lekhana, Review of python for solar photovoltaic systems, Evolutionary Computing and Mobile Sustainable Networks (2021) 103–112.

[15] R. L. Wilby, S. P. Charles, E. Zorita, B. Timbal, P. Whetton, L. O. Mearns, Guidelines for use of climate scenarios developed from statistical downscaling methods, Supporting material of the Intergovernmental Panel on Climate Change, available from the DDC of IPCC TGCIA 27 (2004).

[16] S.-H. Chen, S.-C. Yang, C.-Y. Chen, C. van Dam, A. Cooperman, H. Shiu, C. MacDonald, J. Zack, Application of bias corrections to improve hub-height ensemble wind forecasts over the tehachapi wind resource area, Renewable Energy 140 (2019) 281–291.

[17] J. Chen, F. P. Brissette, D. Chaumont, M. Braun, Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America, Water Resources Research 49 (7) (2013) 4187–4205.

[18] C. Piani, J. Haerter, E. Coppola, Statistical bias correction for daily precipitation in regional climate models over Europe, Theoretical and Applied climatology 99 (1) (2010) 187–192.

[19] A. W. Wood, L. R. Leung, V. Sridhar, D. Lettenmaier, Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, Climatic Change 62 (1) (2004) 189–216.

[20] C. Teutschbein, J. Seibert, Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, Journal of Hydrology 456 (2012) 12–29.

[21] H. Li, J. Sheffield, E. F. Wood, Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching, Journal of Geophysical Research: Atmospheres 115 (D10) (2010).

[22] F. Johnson, A. Sharma, What are the impacts of bias correction on future drought projections?, Journal of Hydrology 525 (2015) 472–485.

[23] F. Johnson, A. Sharma, A nesting model for bias correction of variability at multiple time scales in general circulation model precipitation simulations, Water Resources Research 48 (1) (2012).

[24] A. J. Cannon, Multivariate bias correction of climate model output: Matching marginal distributions and intervariable dependence structure, Journal of Climate 29 (19) (2016) 7045–7064.

[25] L. Yang, Q. Feng, Z. Yin, X. Wen, R. C. Deo, J. Si, C. Li, Application of multivariate recursive nesting bias correction, multiscale wavelet entropy and ai-based models to improve future precipitation projection in upstream of the Heihe river, Northwest China, Theoretical and Applied Climatology 137 (1) (2019) 323–339.

[26] G. Mao, S. Vogl, P. Laux, S. Wagner, H. Kunstmann, Stochastic bias correction of dynamically downscaled precipitation fields for germany through copula-based integration of gridded observation data, Hydrology and Earth System Sciences 19 (4) (2015) 1787–1806.

[27] M. Vrac, P. Friederichs, Multivariate—intervariable, spatial, and temporal—bias correction, Journal of Climate 28 (1) (2015) 218–237.

[28] R. Leander, T. A. Buishand, Resampling of regional climate model output for the simulation of extreme river flows, Journal of Hydrology 332 (3-4) (2007) 487–496.

[29] R. Leander, T. A. Buishand, B. J. van den Hurk, M. J. de Wit, Estimated changes in flood quantiles of the river meuse from resampling of regional climate model output, Journal of Hydrology 351 (3-4) (2008) 331–343.

[30] P. Smitha, B. Narasimhan, K. Sudheer, H. Annamalai, An improved bias correction method of daily rainfall data using a sliding window technique for climate change impact assessment, Journal of Hydrology 556 (2018) 100–118.

[31] J. Schmidli, C. Frei, P. L. Vidale, Downscaling from gcm precipitation: a benchmark for dynamical and statistical downscaling methods, International Journal of Climatology 26 (5) (2006) 679–689.

[32] S. H. Pour, S. Shahid, E.-S. Chung, X.-J. Wang, Model output statistics downscaling using support vector machine for the projection of spatial and temporal changes in rainfall of Bangladesh, Atmospheric Research 213 (2018) 149–162.

[33] Y. Shi, L. Song, Z. Xia, Y. Lin, R. B. Myneni, S. Choi, L. Wang, X. Ni, C. Lao, F. Yang, Mapping annual precipitation across mainland China in the period 2001–2010 from trmm3b43 product using spatial downscaling approach, Remote Sensing 7 (5) (2015) 5849–5878.

[34] Z. Sa'adi, S. Shahid, E.-S. Chung, T. bin Ismail, Projection of spatial and temporal changes of rainfall in sarawak of borneo island using statistical downscaling of CMIP5 models, Atmospheric Research 197 (2017) 446–460.

[35] M. Devak, C. Dhanya, A. Gosain, Dynamic coupling of support vector machine and k-nearest neighbour for downscaling daily rainfall, Journal of Hydrology 525 (2015) 286–301.

[36] A. M. Ahmed, S. M. A. Shah, Application of artificial neural networks to predict peak flow of surma river in sylhet zone of bangladesh, International Journal of Water 11 (4) (2017) 363–375.

[37] A. M. Ahmed, R. C. Deo, Q. Feng, A. Ghahramani, N. Raj, Z. Yin, L. Yang, Deep learning hybrid model with boruta-random forest optimiser algorithm for streamflow forecasting with climate mode indices, rainfall, and periodicity, Journal of Hydrology 599 (2021) 126350.

[38] Z. M. Yaseen, S. M. Awadh, A. Sharafati, S. Shahid, Complementary data-intelligence model for river flow simulation, Journal of Hydrology 567 (2018) 180–190.

[39] A. M. Ahmed, R. C. Deo, N. Raj, A. Ghahramani, Q. Feng, Z. Yin, L. Yang, Deep learning forecasts of soil moisture: convolutional neural network and gated recurrent unit models coupled with satellite-derived MODIS, observations and synoptic-scale climate index data, Remote Sensing 13 (4) (2021) 554.

[40] A. Ahmed, R. C. Deo, A. Ghahramani, N. Raj, Q. Feng, Z. Yin, L. Yang, LSTM integrated with boruta-random forest optimiser for soil moisture estimation under RCP4. 5 and RCP8. 5 global warming scenarios, Stochastic Environmental Research and Risk Assessment 35 (9) (2021) 1851–1881.

[41] S. S. Haykin, S. S. Haykin, Kalman filtering and neural networks, Vol. 284, Wiley Online Library, 2001.

[42] D. J. Lary, L. Remer, D. MacNeill, B. Roscoe, S. Paradise, Machine learning and bias correction of modis aerosol optical depth, IEEE Geoscience and Remote Sensing Letters 6 (4) (2009) 694–698.

[43] A. J. Smola, B. Schölkopf, A tutorial on support vector regression, Statistics and Computing 14 (3) (2004) 199–222.

[44] S. Salcedo-Sanz, J. L. Rojo-Álvarez, M. Martínez-Ramón, G. Camps-Valls, Support vector machines in engineering: an overview, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 4 (3) (2014) 234–267.

[45] S. Tripathi, V. Srinivas, R. S. Nanjundiah, Downscaling of precipitation for climate change scenarios: a support vector machine approach, Journal of Hydrology 330 (3-4) (2006) 621–640.

[46] Y. Zhang, J. Duchi, M. Wainwright, Divide and conquer kernel ridge regression, in: Conference on learning theory, PMLR, 2013, pp. 592–617.

[47] Y. You, J. Demmel, C.-J. Hsieh, R. Vuduc, Accurate, fast and scalable kernel ridge regression on parallel and distributed systems, in: Proceedings of the 2018 International Conference on Supercomputing, 2018, pp. 307–317.

[48] M. Ali, R. Prasad, Y. Xiang, Z. M. Yaseen, Complete ensemble empirical mode decomposition hybridized with random forest and kernel ridge regression model for monthly rainfall forecasts, Journal of Hydrology 584 (2020) 124647.

[49] M. Ali, R. C. Deo, T. Maraseni, N. J. Downs, Improving spi-derived drought forecasts incorporating synoptic-scale climate indices in multi-phase multivariate empirical mode decomposition model hybridized with simulated annealing and kernel ridge regression algorithms, Journal of Hydrology 576 (2019) 164–184.

[50] F. Douak, F. Melgani, N. Benoudjit, Kernel ridge regression with active learning for wind speed prediction, Applied Energy 103 (2013) 328–340.

[51] J. Naik, R. Bisoi, P. Dash, Prediction interval forecasting of wind speed and wind power using modes decomposition based low rank multi-kernel ridge regression, Renewable energy 129 (2018) 357–383.

[52] M. A. Alalami, M. Maalouf, T. H. EL-Fouly, Wind speed forecasting using kernel ridge regression with different time horizons, in: International Conference on Time Series and Forecasting, Springer, 2019, pp. 191–203.

[53] S. Zhang, T. Zhou, L. Sun, C. Liu, Kernel ridge regression model based on beta-noise and its application in short-term wind speed forecasting, Symmetry 11 (2) (2019) 282.

[54] J. Naik, P. K. Dash, S. Dhar, A multi-objective wind speed and wind power prediction interval forecasting using variational modes decomposition based multi-kernel robust ridge regression, Renewable energy 136 (2019) 701–731.

[55] P. Dash, I. Majumder, N. Nayak, R. Bisoi, Point and interval solar power forecasting using hybrid empirical wavelet transform and robust wavelet kernel ridge regression, Natural Resources Research 29 (5) (2020) 2813–2841.

[56] W. Xu, X. Liu, F. Leng, W. Li, Blood-based multi-tissue gene expression inference with bayesian ridge regression, Bioinformatics 36 (12) (2020) 3788–3794.

[57] S. Kumar, C. Ojha, M. K. Goyal, R. Singh, P. Swamee, Modeling of suspended sediment concentration at kasol in india using ann, fuzzy logic, and decision tree algorithms, Journal of Hydrologic Engineering 17 (2) (2012) 394–404.

[58] J. Cai, K. Xu, Y. Zhu, F. Hu, L. Li, Prediction and analysis of net ecosystem carbon exchange based on gradient boosting regression and random forest, Applied energy 262 (2020) 114566.

[59] A. Guryanov, Histogram-based algorithm for building gradient boosting ensembles of piecewise linear decision trees, in: International Conference on Analysis of Images, Social Networks and Texts, Springer, 2019, pp. 39–50.

[60] S. Shabani, S. Samadianfard, M. T. Sattari, A. Mosavi, S. Shamshirband, T. Kmet, A. R. Várkonyi-Kóczy, Modeling pan evaporation using gaussian process regression k-nearest neighbors random forest and support vector machines; comparative analysis, Atmosphere 11 (1) (2020) 66.

[61] S. Salcedo-Sanz, R. C. Deo, L. Cornejo-Bueno, C. Camacho-Gómez, S. Ghimire, An efficient neuro-evolutionary hybrid modelling mechanism for the estimation of daily global solar radiation in the sunshine state of Australia, Applied Energy 209 (2018) 79–94.

[62] A. Zahedi, Australian renewable energy progress, Renewable and Sustainable Energy Reviews 14 (8) (2010) 2208–2213.

[63] D. A. Martin, Linking fire and the united nations sustainable development goals, Science of the Total Environment 662 (2019) 547–558.

[64] Works DoEaP., Achieving our renewable energy targets. Queensland Government; 2021. (2021).

[65] C. Arena, 'australian solar energy forecasting system final report: Project results and lessons learnt, Commonwealth Sci. Ind. Res. Org., Canberra, ACT, Australia, Tech. Rep (2016).

[66] R. Kistler, E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, et al., The NCEP–NCAR 50-year reanalysis: monthly means cd-rom and documentation, Bulletin of the American Meteorological society 82 (2) (2001) 247–268.

[67] Y. Fan, H. van den Dool, Bias correction and forecast skill of NCEP GFS ensemble week-1 and week-2 precipitation, 2-m surface air temperature, and soil moisture forecasts, Weather and Forecasting 26 (3) (2011) 355–370.

[68] H. Van den Dool, J. Huang, Y. Fan, Performance and analysis of the constructed analogue method applied to us soil moisture over 1981–2001, Journal of Geophysical Research: Atmospheres 108 (D16) (2003).

[69] J. Huang, H. M. van den Dool, K. P. Georgarakos, Analysis of model-calculated soil moisture over the united states (1931–1993) and applications to long-range temperature forecasts, Journal of Climate 9 (6) (1996) 1350–1362.

[70] P. A. Blankenau, Bias and other error in gridded weather data sets and their impacts on estimating reference evapotranspiration (2017).

[71] Y. Feng, N. Cui, D. Gong, Q. Zhang, L. Zhao, Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling, Agricultural Water Management 193 (2017) 163–173.

[72] L. Breiman, Random forests, Machine learning 45 (1) (2001) 5–32.

[73] M. S. Al-Musaylh, R. C. Deo, J. F. Adamowski, Y. Li, Short-term electricity demand forecasting using machine learning methods enriched with ground-based climate and ECMWF reanalysis atmospheric predictors in southeast queensland, australia, Renewable and Sustainable Energy Reviews 113 (2019) 109293.

[74] P. Exterkate, Model selection in kernel ridge regression, Computational Statistics & Data Analysis 68 (2013) 1–16.

[75] C. Saunders, A. Gammerman, V. Vovk, Ridge regression learning algorithm in dual variables (1998).

[76] A. Alaoui, E. Willimann, K. Jasper, G. Felder, F. Herger, J. Magnusson, R. Weingartner, Modelling the effects of land use and climate changes on hydrology in the ursern valley, switzerland, Hydrological Processes 28 (10) (2014) 3602–3614.

[77] M. F. Sanner, et al., Python: a programming language for software integration and development, J Mol Graph Model 17 (1) (1999) 57–61.

[78] O. Kramer, Scikit-learn, in: Machine learning for evolution strategies, Springer, 2016, pp. 45–53.

[79] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Scikit-learn: Machine learning in python, Journal of Machine Learning Research 12 (2011) 2825–2830.

[80] P. Barrett, J. Hunter, J. T. Miller, J.-C. Hsu, P. Greenfield, matplotlib–a portable python plotting package, in: Astronomical data analysis software and systems XIV, Vol. 347, 2005, p. 91.

[81] M. L. Waskom, Seaborn: statistical data visualization, Journal of Open Source Software 6 (60) (2021) 3021.

[82] E. Metzger, O. Smedstad, S. Carroll, User's manual for the global ocean forecast system (gofs) version 3.0 (03 2009).

[83] A. M. Ahmed, Prediction of dissolved oxygen in surma river by biochemical oxygen demand and chemical oxygen demand using the artificial neural networks (anns), Journal of King Saud University-Engineering Sciences 29 (2) (2017) 151–158.

[84] R. C. Deo, X. Wen, F. Qi, A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset, Applied Energy 168 (2016) 568–593.

[85] R. C. Deo, N. Downs, A. V. Parisi, J. F. Adamowski, J. M. Quilty, Very short-term reactive forecasting of the solar ultraviolet index using an extreme learning machine integrated with the solar zenith angle, Environmental Research 155 (2017) 141–166.

[86] T. Chai, R. R. Draxler, Root mean square error (rmse) or mean absolute error (mae)?–arguments against avoiding rmse in the literature, Geoscientific Model Development 7 (3) (2014) 1247–1250.

[87] C. J. Willmott, K. Matsuura, Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Climate Research 30 (1) (2005) 79–82.

[88] C. J. Willmott, S. M. Robeson, K. Matsuura, A refined index of model performance, International Journal of Climatology 32 (13) (2012) 2088–2094.

[89] C. J. Willmott, S. G. Ackleson, R. E. Davis, J. J. Feddema, K. M. Klink, D. R. Legates, J. O'donnell, C. M. Rowe, Statistics for the evaluation and comparison of models, Journal of Geophysical Research: Oceans 90 (C5) (1985) 8995–9005.

[90] D. R. Legates, R. E. Davis, The continuing search for an anthropogenic climate change signal: Limitations of correlation-based approaches, Geophysical Research Letters 24 (18) (1997) 2319–2322.

[91] D. R. Legates, G. J. McCabe Jr, Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, Water Resources Research 35 (1) (1999) 233–241.

[92] D. R. Legates, G. J. McCabe, A refined index of model performance: a rejoinder, International Journal of Climatology 33 (4) (2013) 1053–1056.

[93] K. E. Taylor, Summarizing multiple aspects of model performance in a single diagram, Journal of Geophysical Research: Atmospheres 106 (D7) (2001) 7183–7192.

[94] S. Ghimire, T. Nguyen-Huy, R. C. Deo, D. Casillas-Perez, S. Salcedo-Sanz, 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, Sustainable Materials and Technologies 32 (2022) e00429.

[95] S. Ghimire, B. Bhandari, D. Casillas-Pérez, R. C. Deo, S. Salcedo-Sanz, Hybrid deep cnn-svr algorithm for solar radiation prediction problems in queensland, australia, Engineering Applications of Artificial Intelligence 112 (2022) 104860.

[96] S. Ghimire, R. C. Deo, D. Casillas-Pérez, S. Salcedo-Sanz, Boosting solar radiation predictions with global climate models, observational predictors and hybrid deep-machine learning algorithms, Applied Energy 316 (2022) 119063.

[97] A. Sarhadi, D. H. Burn, F. Johnson, R. Mehrotra, A. Sharma, Water resources climate change projections using supervised nonlinear and multivariate soft computing techniques, Journal of Hydrology 536 (2016) 119–132.

[98] J. D. Salas, G. Q. Tabios III, P. Bartolini, Approaches to multivariate modeling of water resources time series 1, JAWRA Journal of the American Water Resources Association 21 (4) (1985) 683–708.

[99] R. Mehrotra, A. Sharma, Correcting for systematic biases in multiple raw gcm variables across a range of timescales, Journal of Hydrology 520 (2015) 214–223.

[100] R. Mehrotra, F. Johnson, A. Sharma, A software toolkit for correcting systematic biases in climate model simulations, Environmental Modelling & Software 104 (2018) 130–152.

[101] G. G. Pegram, et al., A nested multisite daily rainfall stochastic generation model, Journal of Hydrology 371 (1-4) (2009) 142–153.

Appendix A. Further Analysis of Bias Reduction Results

Tables A.5 and A.6 show the percentage increase in the level of agreement between bias-corrected Total Cloud Cover forecasts versus the non-corrected values generated by the GFS model over an eight day forecast horizon. Here, we show the two approaches and present the change in r-value against conventional bias correction method and the reference value (without any bias corrections applied). It is evident that all ML models lead to a significant reduction in the bias in cloud cover forecasts for Approach-2. When results for Approach-1 are considered (Table A.5), there is some discrepancy for Day 5 and Day 8. In spite of this, the present study shows a strong potential utility of ML methods for bias correction of cloud cover forecasts generated by the GFS numerical weather prediction model.

Table A.5: The percentage change in correlation coefficient (r) between $\text{TCDC}_{\text{GFS-Forecast}}$ and $\text{TCDC}_{\text{GFS-Analysis}}$ after applying bias corrections for Approach-1. Note that a positive change, indicated in blue font, represents a reduction in the bias of Total Cloud Cover forecasts. The best model for bias reduction is boldfaced.

Day Ahead Forecast Horizon	**Relative to MRNBC**			
	KRR	MARS	RF	SVR
Day 2	***55.8***	*52.4*	*52.4*	*55.1*
Day 3	*54.2*	***67.8***	***67.8***	*45.3*
Day 4	*3.8*	*12.0*	*12.0*	***29.6***
Day 5	17.6	-23.4	-23.4	42.8
Day 6	63.1	71.4	71.4	**66.5**
Day 7	13.8	**40.4**	**40.4**	20.3
Day 8	-17.2	-41.8	-41.8	-2.9
Day Ahead Forecast Horizon	**Relative to reference value**			
Day 2	**44.5**	41.4	41.4	43.8
Day 3	**15.7**	25.9	25.9	9.0
Day 4	-18.4	-11.9	-11.9	1.9
Day 5	**2.3**	-33.4	-33.4	24.3
Day 6	**20.8**	26.9	26.9	23.3
Day 7	-16.2	3.4	3.4	-11.4
Day 8	-37.6	-56.1	-56.1	-26.8

Appendix B. Multivariate recursive nesting bias correction

The MRNBC corrects the seasonal and non-seasonal time series based on multivariate auto-regressive modelling. First introduced by Mehrotra et al. (2018), the MRNBC aims to incorporate the Recursive Nested Bias Correction (RNBC). The method has been used previously [25]. So, in this approach, the $\text{TCDC}_{\text{GFS-Forecast}}$ simulations are nested into the observed data for all timescales of interest. Before applying the nesting, seasonal and non-seasonal time series are standardised to a mean of zero and a standard deviation of 1.

With m predictor variables at an i time step for a $Z(m \times t)$ matrix, the lag-one auto-correlation and the lag-one and lag-zero cross-correlation in $\text{TCDC}_{\text{GFS-Forecast}}$ simulations can be modified to match the observed correlations in the time and space [97]. The multivariate autoregressive order 1 (MAR1) model for $\text{TCDC}_{\text{GFS-Forecast}}$ data and observed variables is therefore expressed as follows [98]:

$$\hat{Z}_i^h = C\hat{Z}_{i-1}^h + D_{\epsilon i} \tag{B.1}$$

Table A.6: The percentage change in correlation coefficient (r) between $\text{TCDC}_{\text{GFS-Forecast}}$ and $\text{TCDC}_{\text{GFS-Analysis}}$ after applying bias corrections for Approach-2. Note that a positive change, indicated in blue font, represents a reduction in the bias of Total Cloud Cover forecasts. The best model for bias reduction is boldfaced.

Day Ahead Forecast Horizon	**Relative to MRNBC**			
	KRR	MARS	RF	SVR
Day 2	***64.0***	*61.0*	*54.5*	*59.7*
Day 3	*85.1*	***85.0***	***75.8***	*86.0*
Day 4	*98.3*	*96.3*	*96.3*	***96.5***
Day 5	84.0	76.0	76.0	82.7
Day 6	97.9	79.3	95.6	**93.8**
Day 7	112.6	**101.2**	**101.2**	105.4
Day 8	89.7	88.6	83.6	85.9
Day Ahead Forecast Horizon	**Relative to reference value**			
Day 2	**52.2**	49.4	43.3	48.1
Day 3	**38.9**	38.8	31.9	39.6
Day 4	**55.9**	54.4	54.4	54.6
Day 5	**60.1**	53.2	53.2	59.0
Day 6	**46.5**	32.8	44.9	43.5
Day 7	56.6	48.1	48.1	51.2
Day 8	42.9	42.1	38.4	40.1

$$\hat{Z}_i^g = E\hat{Z}_{i-1}^g + F_{\epsilon i} \tag{B.2}$$

where Z^h represents the observations and Z^g is the $\text{TCDC}_{\text{GFS-Forecast}}$ data. Data are standardised to construct a periodic time series $\hat{Z}_i^g$ to be modified to match the observation $\hat{Z}_i^h$, where ϵi is a mutually independent vector with random variation having zero mean value and an identity covariance matrix. C and D are lag-zero and lag-one cross-correlation coefficient matrices for observation $\hat{Z}_i^h$ and the coefficients E and F are calculated for the standardised $\text{TCDC}_{\text{GFS-Forecast}}$ output.

Equations (B.1) and (B.2) are rearranged and modified $\hat{Z}_i^g$ along with lag-zero and lag-one correlation matrices such as C and D to $\hat{Z}_i^g$ have the desired dependence properties [98].

$$\hat{Z}_i'^h = CZ_{i-1}'^g + DF^{-1}\hat{Z}_i^g - DF^{-1}E\hat{Z}_{i-1}^g \tag{B.3}$$

For correction of periodic parameters, let vectors $Z_{t,i}^h$ and $Z_{t,i}^g$ represent the observations and the $\text{TCDC}_{\text{GFS-Forecast}}$ outputs, respectively, with m variables for month i and year t. The standardised periodic time series with a mean of zero and a unit variance is denoted as $\hat{Z}_{t,i}$. Following Equation (B.3), the series $\hat{Z}_{t,i}'^g$ maintains the observed lag-one serial and cross dependence as follows [98]:

$$\hat{Z}_{t,i}'^g = C_i Z_{t,i-1}'^g + D_i F_i^{-1}\hat{Z}_{t,i}^g - D_i F_i^{-1}E_i\hat{Z}_{t,i-1}^g \tag{B.4}$$

where $Z_{t,i-1}'^g$ is the corrected time series from a previous month in year t. After corrections, the resulting time series Z'^g is rescaled by the observed mean and standard deviation to yield the final corrected time series $\bar{Z}^g$, details of which can be found in [97, 99, 100].

After correcting the monthly time series, Z is combined to produce a seasonal sequence and the periodic correction. This time series is connected to an annual time series and the correlation, standard deviation, and mean are corrected to form Ag (A is the matrix of yearly

data, $p \times \frac{n}{12}$). Subsequently, each time, aggregation corrections can be applied to daily time series to create a simple correction step [101]:

$$\bar{Z}^{g}_{i,j,s,t} = \left(\frac{\bar{Y}^{g}_{j,s,t}}{Y^{g}_{j,s,t}}\right) \times \left(\frac{\bar{S}^{g}_{s,t}}{S^{g}_{s,t}}\right) \times \left(\frac{\bar{A}^{g}_{t}}{A^{g}_{t}}\right) \times Z^{g}_{i,j,s,t} \tag{B.5}$$

where $\bar{Y}^{g}_{j,s,t}$, $\bar{S}^{g}_{s,t}$ and $\bar{A}^{g}_{t}$ indicate the monthly, seasonally, and annually corrected values, respectively. $Y^{g}_{j,s,t}$, $S^{g}_{s,t}$ and A^{g}_{t} represent the accumulated monthly, seasonal, and annual values, respectively, in day i and j, season s, and year t. The three-step bias correction technique confirms that future variation is not influenced by the bias correction procedure utilised to correct $\text{TCDC}_{\text{GFS-Forecast}}$ [99].