

Optimization Algorithms As Training Approach With Deep Learning Methods To Develop An Ultraviolet Index Forecasting Model

Abul Abrar Masrur Ahmed (✉ abulabramasrur.ahmed@usq.edu.au)

University of Southern Queensland <https://orcid.org/0000-0002-7941-3902>

Mohammad Hafez Ahmed

West Virginia University

Sanjoy Kanti Saha

Norwegian University of Science and Technology: Norges teknisk-naturvitenskapelige universitet

Oli Ahmed

Leading University

Ambica Sutradhar

Leading University

Research Article

Keywords: Deep Learning, Hybrid Model, Solar Ultraviolet Index, Optimization Algorithms, Public Health

Posted Date: September 21st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-886915/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

1 **Optimization algorithms as training approach with deep learning methods to develop an**
2 **ultraviolet index forecasting model**

3 **A. A. Masrur Ahmed:**

4 School of Sciences, University of Southern Queensland, Springfield, QLD 4300, Australia.

5 Email: masrur@outlook.com.au

6 **Mohammad Hafez Ahmed*:**

7 Department of Civil and Environmental Engineering, West Virginia University, PO BOX 6103,

8 Morgantown, WV 26506-6103, USA. Email: mha0015@mix.wvu.edu

9 **Sanjoy Kanti Saha:**

10 Department of Civil and Environmental Engineering, Norwegian University of Science and

11 Technology, Trondheim, Norway. Email: sanjoyks@ntnu.no

12 **Oli Ahmed:**

13 School of Modern Sciences, Leading University, Sylhet-3112, Bangladesh.

14 Email: oliahmed3034@gmail.com

16 **Ambica Sutradhar:**

17 School of Modern Sciences, Leading University, Sylhet-3112, Bangladesh.

18 Email: ambicasutradhar@gmail.com

19
20 **Corresponding Author*** A. A. Masrur Ahmed (masrur@outlook.com.au)

21 **Abstract:** The solar ultraviolet index (UVI) is a key public health indicator to mitigate the
22 ultraviolet-exposure related diseases. However, in practice, the ultraviolet irradiance
23 measurements are difficult and need expensive ground-based physical models and time-
24 consuming satellite-observed data. Furthermore, accurate short-term forecasting is crucial
25 for making effective decisions on public health owing to UVI related diseases. To this end,
26 this study aimed to develop and compare the performances of different hybridized deep
27 learning models for forecasting the daily UVI index. The ultraviolet irradiance-related data
28 were collected for Perth station of Western Australia. A hybrid-deep learning framework
29 was formulated with a convolutional neural network and long short-term memory called
30 CLSTM. The comprehensive dataset (i.e., satellite-derived Moderate Resolution Imaging
31 Spectroradiometer, ground-based datasets from Scientific Information for Landowners,
32 and synoptic-scale climate indices) were fed into the proposed network and optimized by
33 four optimization techniques. The results demonstrated the excellent forecasting capability
34 (i.e., low error and high efficiency) of the recommended hybrid CLSTM model compared
35 to the counterpart benchmark models. Overall, this study showed that the proposed hybrid
36 CLSTM model successfully apprehends the complex and non-linear relationships between
37 predictor variables and the daily UVI. A complete ensemble empirical mode
38 decomposition with adaptive noise (CEEMDAN)-CLSTM-based is appeared to be an
39 accurate forecasting system capable of reacting quickly to measured conditions. Further,
40 the genetic algorithm is found to be the most effective optimization technique. The study
41 inference can considerably enhance real-time exposure advice for the public and help
42 mitigate the potential for solar UV-exposure-related diseases such as melanoma.

43

44 **Keywords:** Deep Learning; Hybrid Model; Solar Ultraviolet Index; Optimization Algorithms;

45 Public Health.

46

47 **List of abbreviations:**

ACO	Ant Colony Optimization
ACF	Autocorrelation Function
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
AO	Arctic Oscillation
ARPANSA	Australian Radiation Protection and Nuclear Safety Agency
BCC	Basal Cell Carcinoma
BOM	Bureau of Meteorology
CEEMDAN	Complete Ensemble Empirical Mode Decomposition with Adaptive Noise
CEEMDAN-CLSTM	Hybrid Model integrating the CEEMDAN and CNN algorithm with LSTM
CEEMDAN-CGRU	Hybrid Model integrating the CEEMDAN and CNN algorithm with GRU
CEEMDAN-GRU	Hybrid model integrating the CEEMDAN algorithm with GRU
CNN-LSTM (or CLSTM)	Hybrid model integrating the CNN algorithm with LSTM
CNN-GRU (or CGRU)	Hybrid model integrating the CNN algorithm with GRU
CEEMDAN	Complete ensemble empirical mode decomposition with Adaptive Noise
CEEMDAN-DT	Hybrid model integrating the CEEMDAN algorithm with DT
CEEMDAN-MLP	Hybrid model integrating the CEEMDAN algorithm with MLP
CEEMDAN-SVR	Hybrid model integrating the CEEMDAN algorithm with SVR
CNN	Convolutional Neural Network
COVID-19	Coronavirus disease 2019
CCF	Cross-Correlation Function
EEMD	Ensemble empirical mode decomposition
EMD	Empirical Mode Decomposition
DEV	differential evolution
DL	Deep Learning
DT	Decision Tree
DWT	Discrete wavelet Transformation
ECDF	Empirical Cumulative Distribution Function
ELM	Extreme Learning Machine

EMI	El-Nino southern oscillation Modoki indices
ENSO	El Niño Southern Oscillation
FE	Forecasting Error
GA	Genetic Algorithm
GB	Giga Bite
GIOVANNI	Geospatial Online Interactive Visualization and Analysis Infrastructure
GRU	Gated Recurrent Unit
GLDAS	Global Land Data Assimilation System
GSFC	Goddard Space Flight Centre
IMF	Intrinsic mode functions
LM	Legates-McCabe's Index
LSTM	Long- short term memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MARS	Multivariate Adaptive Regression Splines
MDB	Murray-Darling Basin
MJO	Madden-Julian Oscillation
ML	Machine Learning
MLP	Multi-Layer Perceptron
MODWT	Maximum Overlap Discrete Wavelet Transformation
MODIS	Moderate Resolution Imaging Spectroradiometer
MRA	Multi-resolution Analysis
MSE	Mean Squared Error
NAO	North Atlantic Oscillation
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NO	Nitrogen Oxide
NOAA	National Oceanic and Atmospheric Administration
NMSC	Non-melanoma Skin Cancer
NSE	Nash–Sutcliffe Efficiency
PACF	partial autocorrelation function
PDO	Pacific Decadal Oscillation
PNA	Pacific North American
PSO	Particle Swarm Optimization
r	Correlation Coefficient
RMM	Real-time Multivariate MJO series
GA	Genetic Algorithm
BRF	Boruta random forest

RMSE	Root-Mean-Square-Error
RNN	Recurrent Neural Network
RRMSE	Relative Root-Mean-Square Error
SAM	Southern Annular Mode
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SCC	Squamous Cell Carcinoma
SILO	Scientific Information for Landowners
SOI	Southern Oscillation Index
SST	Sea Surface Temperature
SVR	Support Vector Regression
US	United States
UV	Ultraviolet
UVI	Ultraviolet Index
WHO	World Health Organization
WI	Willmott's Index of Agreement

48

49 **1. Introduction**

50 Solar ultraviolet (UV) radiation is an essential component for the sustenance of life on
51 Earth (Norval et al., 2007). The UV irradiance consists of a small fraction (e.g., 5-7%) of the total
52 radiation and produces numerous beneficial effects on human health. It has been in use since the
53 ancient time for improving the body's immune system, such as strengthening of bones and muscles
54 (Juzeniene and Moan, 2012) as well as in treating various hard-to-treat skin diseases such as atopic
55 dermatitis, psoriasis, phototherapy of localized scleroderma (Furuhashi et al., 2020; Kroft et al.,
56 2008), and vitiligo (Roshan et al., 2020). UV-stimulated tanning has been proved as a positive
57 mood changing and relaxing effect for many (Sivamani et al., 2009). Further, UV-induced nitrogen
58 oxide (NO) plays a vital role in reducing human blood pressure (Juzeniene and Moan, 2012;
59 Opländer Christian et al., 2009).

60 UV light has also been widely used as an effective disinfectant in the food and water
61 industry to inactivate disease-producing microorganisms (Gray, 2014). Because of no harmful by-
62 products generation and its effectiveness against protozoa contamination, the use of UV light as a

63 drinking water disinfectant has achieved an increased acceptance (Timmermann et al., 2015). To
64 date, most of the UV-installed public water supplies are in Europe. In the United States (US), its
65 application is mainly limited to groundwater treatment (Chen et al., 2006). However, its use is
66 expected to increase in the future for the disinfection of different wastewater systems. Developing
67 countries worldwide find it useful as it offers a simple, low-cost, and effective disinfection
68 technique in water treatment compared to the traditional chlorination method (Mäusezahl et al.,
69 2009; Pooi and Ng, 2018).

70 The application of UV light has also shown potency in fighting airborne-mediated diseases
71 for a long time (Hollaender et al., 1944; Wells and Fair, 1935). For instance, a recent study
72 demonstrated that a small dose (i.e., 2 mJ/cm² of 222-nm) of UV-C light can efficiently inactivate
73 aerosolized H1N1 influenza viruses (Welch et al., 2018). The far UV-C light can also be applied
74 in sterilizing surgical equipment. Recently, the use of the UV-C light as the surface disinfectant
75 has been significantly increased to combat the global pandemic (COVID-19) caused by
76 coronavirus SARS-CoV2. A recent study also highlighted the efficacy of UV light application on
77 the disinfection of COVID-19 surface contamination (Heilingloh et al., 2020).

78 However, the research on UV radiation has also been a serious concern over the years due
79 to its dichotomy nature. UV irradiance can also have detrimental biological effects on human
80 health, such as skin cancer and eye disease (Lucas et al., 2008; Turner et al., 2017). Chronic
81 exposure to UV light has been reported as a major risk factor responsible for melanoma and non-
82 melanoma cancers (Saraiya et al., 2004; Sivamani et al., 2009) and associated with 50-90% of
83 these diseases. In a recent study, the highest global incidence rates of melanoma were observed in
84 the Australasia region compared to other North American and European parts (Karimkhani et al.,
85 2017). Therefore, it is crucial to provide correct information about the intensity of UV irradiance

86 to the people at risk to protect their health. This information would also be helpful for working
87 people in different sectors (e.g., agriculture, medical sector, water management, etc.).

88 The World Health Organization (WHO) formulated the global UV index (UVI) as a
89 numerical public health indicator to convey the associated risk when exposed to UV radiation
90 (Fernández-Delgado et al., 2014; WHO, 2002). However, the UV irradiance estimation in practice
91 requires ground-based physical models (Raksasat et al., 2021) and satellite-derived observing
92 systems with advanced technical expertise (Kazantzidis et al., 2015). The installation of required
93 equipment (*i.e.*, spectroradiometers, radiometers, and sky images) is expensive (Deo et al., 2017)
94 and difficult for remote regions, primarily mountainous areas. Furthermore, the solar irradiance is
95 also highly impacted by many hydro-climatic factors, e.g., clouds and aerosol (Li et al., 2018;
96 Staiger et al., 2008) and ozone (Baumgaertner et al., 2011; Tartaglione et al., 2020) that can insert
97 considerable uncertainties into the available process-based and empirical models (detail also given
98 in method Section). Therefore, the analysis of sky images may also require extensive bias
99 corrections, *i.e.*, cloud modification (Krzyścin et al., 2015; Sudhibrabha et al., 2006), which creates
100 further technical as well as a computational burden. An application of data-driven models can be
101 useful to minimize these formidable challenges. Specifically, the non-linearity into data matrix can
102 easily be handled using data-driven models that a traditional process-based and/ semi-process-
103 based model fails. Further, the data-driven models are easy to implement, do not demand high
104 process-based cognitions (Qing and Niu, 2018; Wang et al., 2018), and are computationally less
105 burdensome.

106 As an alternative to conventional process-based and empirical models, applying different
107 machine learning (ML) algorithms as data-driven models has proven tremendously successful
108 because of the powerful computational efficiency. With technological advancement,

109 computational efficiency has been significantly increased, and researchers have developed many
110 ML tools. Artificial neural networks (ANNs) are the most common and extensively employed in
111 solar energy applications (Yadav and Chandel, 2014). However, many studies, such as the multiple
112 layer perceptron (MLP) neural networks (Alados et al., 2007; Alfadda et al., 2018), support vector
113 regression (SVR) (Fan et al., 2020; Kaba et al., 2017), decision tree (Jiménez-Pérez and Mora-
114 López, 2016), and random forest (Fouilloy et al., 2018) have also been extensively applied in
115 estimating the UV erythemal irradiance. The multivariate adaptive regression splines (MARS) and
116 M5 algorithms were applied in a separate study for forecasting solar radiation (Srivastava et al.,
117 2019). Further, the deep learning network such as the convolutional neural network (CNN)
118 (Szenicer et al., 2019) and the long short-term memory (LSTM) (Huang et al., 2020; Qing and
119 Niu, 2018; Raksasat et al., 2021) are recent additions in this domain.

120 However, the UVI indicator is more explicit to common people compared to UV irradiance
121 values. Further, only a few data-driven models have been applied for UVI forecasting. For
122 example, an ANN was used in modeling UVI on a global scale (Latosińska et al., 2015). An
123 extreme learning method (ELM) was applied in forecasting UVI in the Australian context (Deo et
124 al., 2017). To date, there have not been many studies that used ML methods to forecast UVI. Albeit
125 the successful predictions of these standalone ML algorithms, they have architectural flaws and
126 predominantly suffer from overfitting efficiency (Ahmed and Lin, 2021). Therefore, the hybrid
127 deep learning models receive increased interest and are extremely useful in predictions with higher
128 efficiency than the standalone machine learning models. Hybrid models such as particle swarm
129 optimization (PSO)-ANN (Isah et al., 2017), wavelet-ANN (Zhang et al., 2019), genetic algorithm
130 (GA)-ANN (Antanasijević et al., 2014), Boruta random forest (BRF)-LSTM (Ahmed et al., 2021;
131 Ahmed et al., 2021b), ensemble empirical mode decomposition (EEMD) (Liu et al., 2015),

132 adaptive neuro-fuzzy inference system (ANFIS)-ant colony optimization (ACO) (Pruthi and
133 Bhardwaj, 2021) and (ACO)-CNN-GRU (Ahmed et al., 2021b) have been applied across
134 disciplines and attained substantial tractions. However, a CNN-LSTM (i.e., CLSTM) hybrid model
135 can efficiently extract inherent features from the data matrix than other machine learning models
136 and has successfully predicted time series air quality and meteorological data (Pak et al., 2018).
137 The application of such a hybrid model for predicting sequence data, i.e., the UVI for consecutive
138 days, can be an effective tool with excellent predictive power. However, the forecasting of UVI
139 with a CLSTM hybrid machine learning model is yet to be explored and was a key motivation for
140 conducting this present study.

141 The conventional MRA, for instance, discrete wavelet transform (DWT), has been
142 implemented for a long time (Deo and Sahin, 2016; Deo et al., 2016; Nourani et al., 2014; Nourani
143 et al., 2009), but there appear to be disadvantages that preclude the full feature extraction, as well
144 as a unique destination of training features into a tested dataset. MODWT, as an advanced DWT,
145 can identify critical issues (Cornish et al., 2006; Prasad et al., 2017; Rathinasamy et al., 2014). In
146 this study, we employed a new model of EMD called complete ensemble empirical mode
147 decomposition with adaptive noise (CEEMDAN) (Prasad et al., 2018). in CEEMDAN-based
148 decomposition, Gaussian white noise with a unit variance is added consecutively at each stage to
149 reduce the forecasting procedure's complexity (Di et al., 2014). Previous studies used CEEMDAN
150 in predicting soil moisture (Ahmed et al., 2021a; Prasad et al., 2018; Prasad et al., 2019). However,
151 a previous version (i.e., EEMD) was used in forecasting streamflow (Seo and Kim, 2016) and
152 rainfall (Beltrán-Castro et al., 2013; Jiao et al., 2016; Ouyang et al., 2016). The only machine
153 learning algorithm used in the study is CLSTM, which has not been coupled with the EEMD or
154 CEEMDAN to produce a UVI forecast system.

155 This study aims to apply a CLSTM hybrid machine learning model, which can exploit the
156 benefits of both convolutional layers (i.e., important feature extraction) and LSTM layers (i.e.,
157 storing sequence data for an extended period) and evaluate its ability to efficiently forecast the
158 UVI for the next day. The model was constructed and fed with hydro-climatic data in association
159 with UV irradiance in the Australian context. The model was optimized using ant colony
160 optimization, genetic algorithm, particle swarm optimization, and differential evolutionary
161 algorithms. The model accuracy (i.e., efficiency and errors involved in UVI estimations) was
162 assessed with the conventional standalone data-driven models' (e.g., SVR, decision tree, MLP,
163 CNN, LSTM, gated recurrent unit (GRU), etc.) performance statistics. The inference obtained
164 from the modeling results was also discussed that could be tremendously useful in building expert
165 judgment to protect public health in the Australian region and beyond.

166 **2. Materials and Methods**

167 **2.1 Study area and UVI data**

168 The study assessed the solar ultraviolet index of Perth (Latitude: -31.93°E and Longitude:
169 115.10°S), Western Australia. The Australian Radiation Protection and Nuclear Safety Agency
170 (ARPANSA) provided the UVI data of Australia (ARPANSA, 2021). Figure 1 shows the monthly
171 UVI data, the location of Perth, and the assessed station. The figure shows that Perth has low to
172 extreme UV concentration over the year between 1979 – 2007. The Summer season (December to
173 February) had the most extreme UV concentration, whereas the Autumn (March to May) has
174 moderate to high, and Winter (June to August) demonstrates lower to moderate, and Spring
175 (September to November) has higher to extreme UVI value in Perth.

176 Malignant melanoma rates in Western Australia are second only to those in Queensland,
177 Australia's most populated state (Slevin et al., 2000). Australia has the highest incidence of NMSC
178 (Non-melanoma skin cancer) globally (Anderiesz et al., 2006; Staples et al., 1998). Approximately
179 three-quarters of the cancer cases have basal cell carcinoma (BCC) and squamous cell carcinoma
180 (SCC) types. These are attributed to the fair-skinned population's high exposure to ambient solar
181 radiation (Boniol, 2016; McCarthy, 2004). As a result, Australia is seen as a world leader in public
182 health initiatives to prevent and detect skin cancer. Programs that have brought awareness of
183 prevention strategies and skin cancer diagnoses have data to show that people act on their
184 knowledge (Stanton et al., 2004). Many research showed that reducing sun security measures is
185 linked to reducing rates of BCC and SCC in younger groups. They might have received cancer
186 prevention messages as children (Staples et al., 2006). Considering the diversified concentration
187 of UVI concentration, this study considers Perth as an ideal study area.

188 **2.2 Datasets of predictor variables**

189 Three distinct data sources were used to collect the predictor variables in this analysis. The
190 Moderate Resolution Imaging Spectroradiometer (MODIS) satellite datasets are used to capturing
191 land surface status and flow parameters at regular temporal resolutions. These are supplemented
192 by ground-based Scientific Information for Landowners (SILO) repository meteorological data for
193 biophysical modeling and climate mode indices to help achieve Sea Surface Temperature (SST)
194 over Australia. Geospatial Online Interactive Visualization and Analysis Infrastructure
195 (GIOVANNI) is a geoscience data repository that provides a robust online visualization and
196 analysis platform for geoscience datasets. It collects data from over 2000 satellite variables (Chen
197 et al., 2010). The MODIS- aqua yielded 8 predictor variables for our study: a high-temporal
198 terrestrial modeling system consisting of a surface state and providing daily products with a high

199 resolution (250 m at nadir). A list of predictors of the MODIS Satellite can be obtained from the
200 National Aeronautics and Space Administration (NASA) database (Giovanni, 2021).

201 The surface UVI is influenced by atmospheric attenuation of incident solar radiation (Deo
202 et al., 2017). The angle subtended from the zenith (θ_s) to the solar disc is another factor that affects
203 the intensity of solar UV radiation (Allaart et al., 2004). The ultraviolet attenuation of clear-sky
204 solar radiation is dependent on ozone and atmospheric aerosol concentrations, along with cloud
205 cover (Deo et al., 2017). This implies that the measurements of biologically effective UV
206 wavelengths are affected by total column ozone concentration. Incident radiation at the Earth's
207 surface is reduced by aerosols such as dust, smoke, and vehicle exhausts (Downs et al., 2016;
208 Román et al., 2013). Moreover, Lee et al. (2009) found a significant correlation between UV solar
209 radiation and geopotential height. Considering the direct influence of the predictors over ultraviolet
210 radiation and UV index, this study collected Ozone total column, aerosol optical depth (550nm
211 and 342.5nm), geopotential height, cloud fraction, and combined cloud optical thickness data from
212 the Geospatial Online Interactive Visualization and Analysis Infrastructure (GIOVANNI)
213 repository.

214 Therefore, meteorological predictor variables (i.e., temperature, u- and v-winds) were
215 significant while modeling UVI (Lee et al., 2009). Moreover, the cloud amount and diurnal
216 temperature range have a strong positive correlation, while rainfall and cloud amount show a
217 strong negative correlation (Jovanovic et al., 2011). Although overall cloud patterns agree with
218 rainfall patterns across Australia, the higher-quality cloud network is too coarse to represent
219 topographic influences accurately. Changes in the amount of cloud cover caused by climate change
220 can result in long-term changes in maximum and minimum temperature. Owing to the relations of
221 hydro-meteorological variables with UVI and their interconnections with cloud cover, the study

222 selected nine meteorological variables from the Scientific Knowledge for Land-Owners (SILO)
223 database to expand the pool of predictor variables, allowing for more practical application and
224 model efficiency. SILO data are managed by Queensland's Department of Environment and
225 Research and can be obtained from <https://www.longpaddock.qld.gov.au/silo>.

226 Aerosol-rainfall relationships are also likely to be artifacts of cloud and cloud-clearing
227 procedures. During the Madden-Julian Oscillation (MJO) wet phase, the high cloud's value
228 increases, the cloud tops rise, and increased precipitation enhances wet deposition, which reduces
229 aerosol mass loading in the troposphere (Tian et al., 2008). The MJO (Lau and Waliser, 2011;
230 Madden and Julian, 1994, 1971) dominates the intra-seasonal variability in the tropical
231 atmosphere. A relatively slow-moving, large-scale oscillation in the deep tropical convection and
232 baroclinic winds exists in the warmer tropical waters in the Indian and western Pacific Oceans
233 (Hendon and Salby, 1994; Kiladis et al., 2001; Tian et al., 2008). The study used the Real-time
234 Multivariate MJO series 1 (RMM1) and 2 (RMM2) obtained from the Bureau of Meteorology,
235 Australia (BOM, 2020). Though RMM1 and RMM2 indicate an evolution of the MJO independent
236 of season, the coherent off-equatorial behavior is strongly seasonal (Wheeler and Hendon, 2004).
237 Pavlakis et al. (2008, 2007) studied the spatial and temporal variation of long surface wave and
238 short wave radiation. A high correlation was found between the longwave radiation anomaly and
239 the Niño3.4 index time series over the Niño3.4 region located in the central Pacific.

240 Moreover, Pinker et al. (2017) investigated the effect of El Niño and La Nina cycles on
241 surface radiative fluxes and the correlations between their anomalies and a variety of El Niño
242 indices. The maximum variance of anomalous incoming solar radiation is located just west of the
243 dateline. It coincides with anomalous SST (Sea surface temperature) gradient in the traditional
244 eastern Pacific El Niño Southern Oscillation (ENSO). However, we derive the Southern

245 Oscillation Index highly correlated with solar irradiance and mean Northern Hemisphere
246 temperature fluctuations reconstructions (Yan et al., 2011). In North America and the North
247 Pacific, land and sea surface temperatures, precipitation, and storm tracks are determined mainly
248 by atmospheric variability associated with the Pacific North American (PNA) pattern. The modern
249 instrumental record indicates a recent trend towards a positive PNA phase, which has resulted in
250 increased warming and snowpack loss in northwest North America (Liu et al., 2017). This study
251 used fifteen climate mode indices to increase the diversity.

252 **2.3 Standalone models**

253 ***2.3.1 Multiple layer perceptron (MLP)***

254 The MLP is a simple feedforward neural network with three layers and is commonly used as a
255 reference model for comparison in machine learning application research (Ahmed and Lin, 2021).
256 The three layers are the input layer, a hidden layer with n-nodes, and the output layer. The input
257 data are fed into the input layer, transformed into the hidden layer via a non-linear activation
258 function (i.e., a logistic function). The target output is estimated, Eq. (1).

$$259 \quad y = f(\sum w^T x + b) \quad (1)$$

260 where w = the vector of weights, x_i = the vector of inputs, b = the bias term; f = the non-linear
261 sigmoidal activation function, i.e., $f(z) = \frac{1}{1+e^{-z}}$.

262 The computed output is then compared with the measured output, and the corresponding loss, i.e.,
263 the mean squared error (MSE), is estimated. The model parameters (i.e., initial weights and bias)
264 are updated using a backpropagation method until the minimum MSE is obtained. The model is
265 trained for several iterations and tested for new data sets for prediction accuracy.

266 2.3.2 Support vector regression (SVR)

267 The SVR is constructed based on the statistical learning theory. In SVR, a kernel trick is applied
268 that transfers the input features into the higher dimension to construct an optimal separating
269 hyperplane as follows (Ji et al., 2017):

$$270 f(x) = w \cdot \varphi(x) + b \quad (2)$$

271 where w is the weight vector, b is the bias, and $\varphi(x)$ indicates the high-dimensional feature space.
272 The coefficients w and b , which define the location of the hyperplane, can be estimated by
273 minimizing the following regularized risk function:

$$274 \text{Minimize: } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\varepsilon_i + \varepsilon_i^*) \quad (3)$$

$$275 \text{Subject to } y_i - w \cdot \varphi(x) - b \leq \varepsilon + \varepsilon_i; w \cdot \varphi(x) + b - y_i \leq \varepsilon + \varepsilon_i^*; \varepsilon_i \leq 0; \varepsilon_i^* \leq 0$$

276 where C is the regularization parameter, ε_i and ε_i^* are slack variables. The Eq. (7) can be solved in
277 a dual form using the Lagrangian multipliers as follows:

$$278 \text{Maximize: } -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (a_i - a_i^*) (a_j - a_j^*) K(x_i, x_j) - \sum_{i=1}^n (a_i - a_i^*) + \sum_{i=1}^n (a_i - a_i^*) y_i \quad (4)$$

$$280 \text{Subject to } \sum_{i=1}^n (a_i - a_i^*) = 0; a_i, a_i^* \in [0, C]$$

281 Where, $K(x_i, x)$ is the non-linear kernel function. In this present study, we used a radial basis
282 function (RBF) as the kernel, which is represented as follows:

$$283 K(x_i, x) = \exp\left(\frac{-\|x - x_i\|^2}{2\sigma^2}\right), \text{ where } \sigma \text{ is the bandwidth of the RBF.}$$

284 2.3.3 Decision tree (DT)

285 A decision tree is a predictive model used for both classification and regression analysis (Jiménez-
286 Pérez and Mora-López, 2016). As our data is continuous, we used it for the regression predictions.

287 It is a simple tree-like structure that uses the input observations (i.e., $x_1, x_2, x_3, \dots, x_n$) to predict

288 the target output (i.e., Y). The tree contains many nodes, and at each node, a test to one of the
289 inputs (e.g., x_1) is applied, and the outcome is estimated. The left/right sub-branch of the decision
290 tree is selected based on the estimated outcome. After a specific node, the prediction is made, and
291 the corresponding node is termed the leaf node. The prediction averages out all the training points
292 for the leaf node. The model is trained using all input variables and corresponding loss; the mean
293 squared error (MSE) is calculated to determine the best split of the data. The number of maximum
294 features is set as the total input features during the partition.

295 **2.3.4 Convolutional neural network (CNN)**

296 The CNN model was originally developed for document recognition (Lecun et al., 1998)
297 and used for predictions. Aside from the input and output layer, the CNN architecture has three
298 hidden layers: the convolutional layers, the pooling layers, and a fully connected layer. The
299 convolutional layers abstract the local information from the data matrix using a kernel. The
300 primary advantage of this layer is the implementation of weight sharing and spatial correlation
301 among neighbors (Guo et al., 2016). The pooling layers are the subsampling layers that reduce the
302 size of the data matrix. A fully connected layer is similar to the traditional neural network added
303 at the final pooling layer after completing an alternate stack of convolutional and pooling layers.

304 **2.3.4 Long short-term memory (LSTM)**

305 An LSTM network is a special form of recurrent neural network that stores sequence data
306 for an extended period (Hochreiter and Schmidhuber, 1997). The LSTM structure has three gates:
307 an input gate, an output gate, and a forget gate. The model regulates all these three gates and
308 determines how much data from previous time steps must be stored and transferred to the next
309 steps. The input gate controls the input data at the current time as follows:

$$310 \quad a_i^t = \sum_{i=1}^I w_{il} x_i^t + \sum_{h=1}^H w_{hl} b_h^{t-1} + \sum_{c=1}^C w_{cl} s_c^{t-1}; b_i^t = f(a_i^t) \quad (5)$$

311 Where x_i^t = the input received from the i^{th} node at time t ; b_h^{t-1} = the result of the h^{th} node at time
312 $t-1$; s_c^{t-1} = the cell state (i.e., memory) of the c^{th} node at time $t-1$. The symbol ‘ w ’ represents the
313 weight between nodes, and the f is the activation function. The output gate transfers the current
314 value from Eq. (5) to the output node, Eq. (6). Then, at the final stage, the current value is stored
315 as the cell state in the forget gate, Eq. (7).

$$316 \quad a_w^t = \sum_{i=1}^I w_{iw} x_i^t + \sum_{h=1}^H w_{hw} b_h^{t-1} + \sum_{c=1}^C w_{cw} s_c^{t-1}; b_w^t = f(a_w^t) \quad (6)$$

$$317 \quad a_\phi^t = \sum_{i=1}^I w_{i\phi} x_i^t + \sum_{h=1}^H w_{h\phi} b_h^{t-1} + \sum_{c=1}^C w_{c\phi} s_c^{t-1}; b_\phi^t = f(a_\phi^t) \quad (7)$$

318 **2.3.5 Gated recurrent unit (GRU) network**

319 The GRU network is an LSTM variant having only two gates, such as reset and update gates (Dey
320 and Salem, 2017). The implementation of this network can be represented by the following
321 equations, Eq. (14-17):

$$322 \quad z = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (8)$$

$$323 \quad r = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (9)$$

$$324 \quad m = \phi(W_m x_t + U_m (h_{t-1} \cdot r) + b_m) \quad (10)$$

$$325 \quad h_t = (1 - z)h_{t-1} + z \cdot m \quad (11)$$

326 where σ = the sigmoidal activation function; x_t = the input value at time t ; h_{t-1} = the output value
327 at time $t-1$; and the W_z , U_z , W_r , U_r , W_m , U_m are the weight matrices for each gate and cell state.
328 The symbols r and z represent the reset and update gates, respectively. ϕ is the activation function,
329 and the dot [\cdot] represents the element-wise dot product.

330 **2.4 The proposed hybrid model**

331 **2.4.1 CLSTM (or CNN-LSTM) Hybrid Model**

332 In this paper, a deep learning method using optimization techniques is constructed on top
333 of a forecast model framework. This study demonstrates how the CNN-LSTM (CLSTM) model,
334 comprised of four-layered CNN, can be effectively used for UVI forecasting. The CNN is
335 employed to integrate extracted features to forecast the target variable (i.e., UVI) with minimal
336 training and testing error. Likewise, the CNN-GRU (CGRU) hybrid model is prepared for the same
337 purpose.

338 **2.5 Optimization techniques**

339 **2.5.1 Ant colony optimization**

340 Ant colony optimization (ACO) algorithm model is the graphical representation of the real
341 ants' behavior. In general, ants live in colonies, and they forage for food as a whole by
342 communicating with each other using a chemical substance, the pheromones (Mucherino et al.,
343 2015). An isolated ant cannot move randomly; they always optimize their way towards the food
344 deposit to their nests by interacting with previously laid pheromones marks on the way. The entire
345 colony optimizes their routes with this communication process and establishes the shortest path to
346 their nests from feeding sources (Silva et al., 2009). In ACO, the artificial ants find a solution by
347 moving on the problem graph. They deposit synthetic hormone pheromones on the graph so that
348 upcoming artificial ants can follow the pattern to build a better solution. The artificial ants calculate
349 the model's intrinsic mode functions (IMFs) anticipation by testing artificial pheromone values
350 against the target data. The probability of finding the best IMFs increases for every ant because of
351 changing pheromones value throughout the IMFs. The whole process is just like ant's behavior of
352 finding the optimal option to reach the target. The probability $p_{fi}(d)$ of selecting the shortest

353 distance between the target and the IMFs of the input variable can be mathematically expressed as
 354 follows (Prasad et al., 2019):

$$355 \quad p_{fi}(d) = \frac{(d_i + \Delta_{fi}(d))^2}{(d_i + \Delta_{fi}(d))^2 + (d_i + \Delta_{ft}(d))^2} \quad (12)$$

356 Where $f \in \{1,2\}$ denotes decision point, i and t express as short and long distance to the target at
 357 an instant d is the total amount of pheromone $\Delta_{ft}(d)$. The probability of the longest path can be
 358 determined where $p_{fi}(d) + p_{ft}(d) = 1$. The testing update on the two branches is described as
 359 follows:

$$360 \quad \Delta_{fi}(d) = \Delta_{fi}(d - 1) + p_{fi}(d - 1)a_f(d - 1) + p_{ki}(d - 1)a_k(d - 1) \quad (13)$$

$$361 \quad \Delta_{ft}(d) = \Delta_{ft}(d - 1) + p_{ft}(d - 1)a_f(d - 1) + p_{kt}(d - r)a_k(d - r) \quad (14)$$

362 Where $f, k \in \{1,2\}$ and the value of r represent the remainder in the model. $a_f(d)$ denotes the
 363 number of ants in the node f at a certain period d is given by:

$$364 \quad a_f(d) = p_{ki}(d - 1)a_k(d - 1) + p_{kt}(d - r)a_k(d - r) \quad (15)$$

365 The ACO algorithm is the most used simulation optimization algorithm where myriad artificial
 366 ants work in a simulated mathematical space to search for optimal solutions for a given problem.

367 The ant colony algorithm is dominant in multi-objective optimization as it follows the natural
 368 distribution and self-evolved simple process. However, with the increase of network information,
 369 the ACO algorithm faces various constraints such as local optimization and feature redundancy
 370 for selecting optimal pathways (Peng et al., 2018).

371 **2.5.2 Differential evaluation optimization**

372 The differential evolution (DEV) algorithm is renowned for its simplicity and powerful stochastic
373 direct search method. Besides, DEV has proven an efficient and effective method for searching
374 global optimal solutions for the multimodal objective function, utilizing N-D-dimensional
375 parameter vectors (Seme and Štumberger, 2011). It does not require a specific starting point, and
376 it operates effectively on a population candidate solution. The constant value N denotes the
377 population; in every module, a new generation solution is determined and compared with the
378 previous generation of the population member. It is a repetition process and runs until it reaches
379 the maximum number of generations (i.e., G_{max}). The G defines the generation number of
380 populations which can be written in mathematical proportional order. If the initial population
381 vector is S_G , then $S_G = i_{1,G}, i_{2,G} \dots \dots, i_{NP,G}$, and $G = 0, \dots, G_{max}$

$$382 \quad i_{n,G}, n = 1, 2, \dots \dots, N$$

383 The initial population $S_{G=0}$ is generated using random within given boundaries, which can be
384 written in the following equation:

$$385 \quad i_{j,0}^n = rand_j[0,1] \left(i_j^{(U)} - i_j^{(L)} \right) + i_j^{(L)}, n = 1, 2, \dots, NP, j = 1, 2, \dots, D \quad (16)$$

386 Where $rand_j[0,1]$ is the uniformly distributed number at the interval $[0,1]$, which is chosen a new
387 for each j . D represents the boundary condition. In contrast, (U) and (L) represents the upper and
388 lower limit of the boundary vector parameters. For every generation, a new random vector is
389 randomly created, selecting vectors from the previous generation from the following manner:

$$390 \quad c_{j,G}^n = \begin{cases} i_{j,G-1}^r + F(i_{j,G-1}^r - i_{j,G-1}^r) if rand_j[0,1] \leq \\ i_{j,G-1}^n otherwise \end{cases} \quad (17)$$

391 Where, r is the number of optimizations, c is the candidate vector, $CR \in [0,1]$ and $F \in [0,2]$
 392 control parameter. k is the randomly selected index that ensures the difference between the
 393 candidate vector and the generation vector. The population for new the new generation S_G will be
 394 assembled from the vector of the previous generation S_{G-1} and the candidate vectors $c_{j,G}^n$ the
 395 following equation can describe selection:

396 $G = 0, \dots, G_{max}; n = 1, 2, \dots, NP$

$$397 \quad I_G^n = \left. \begin{cases} c_G^n & \text{if } f(c_G^n) \leq f(I_{G-1}^n) \\ I_{G-1}^n & \text{otherwise} \end{cases} \right\} \quad (18)$$

398 The process repeats with the following generation population number until it satisfies the pre-
 399 defined objective function.

400 **2.5.3 Particle swarm optimization**

401 The particle swarm optimization (PSO) method was developed for continuous non-linear
 402 functions optimization having roots in artificial life and evolutionary computation (Kennedy and
 403 Eberhart, 1995). The method was constructed using a simple concept that tracks each particle's
 404 current position in the swarm by implementing a velocity vector for particle's previous to the new
 405 position. However, the movement of the particles in the swarm depends on the individuals'
 406 external behavior. Therefore, the process is very speculative, uses each particle's memory for
 407 calculating new position, and gained knowledge by the swarm as a whole. Nearest neighbor
 408 velocity matching and craziness, eliminating ancillary variables and incorporate multidimensional
 409 search and acceleration by distance, were the precursor of PSO algorithm simulation (Eberhart and
 410 Shi, 2001). Each particle in the simulation coordinates in the n-dimensional space calculation and
 411 responds to the two quality factors called '*gbest*' and '*pbest*'. *gbest* represents the best location
 412 and value of particle in the population globally, and *pbest* represents the best-fitted solution

413 achieved by the particle so far in the population swarm. Thus, at each time step in the swarm, the
414 PSO concept stands, each particle changing its acceleration towards its two best quality factor
415 locations. The acceleration process begins by separating random numbers and presenting the
416 optimal 'gbest' and 'pbest' locations. The basic steps for the PSO algorithm are given below,
417 according to (Eberhart and Shi, 2001):

418 1. The process starts with initializing sample random particles with random velocities and
419 locations on n-dimensions in the design space.

420 2. The velocity vector for each particle in the swarm is carried out in the next step as the initial
421 velocity vector value.

422 3. Plot the velocity vector value and compare particle fitness evaluation with particle's *pbest*.
423 If the new value is better than the initial value, update the new velocity vector value as
424 *pbest* and previous location equal to the current location in the design space.

425 4. In this step, compare the fitness evaluation with the particles' overall previous global best.
426 If the current value is better than *gbest*, update it to a new *gbest* value and location.

427 5. The velocity and position of the particle can be changed according to the equations:

$$428 \quad v_{nd} = v_{nd} + m_1 * rand(x) * (p_{nd} - z_{nd}) + m_2 * Rand(x) * (p_{gd} - z_{nd}) \quad (19)$$

$$429 \quad z_{nd} = z_{nd} + v_{nd} \quad (20)$$

430 6. Repeat step 2 and continue until the sufficiently fitted value and position are achieved.

431 Particle swarm optimization is well known for its simple operative steps and performance for
432 optimizing a wide range of functions. PSO algorithm can successfully solve the design problem
433 with many local minima and deal with regular and irregular design space problems locally and
434 globally. Although PSO can solve problems more accurately than other traditional gradient-based
435 optimizers, the computational cost is higher in PSO (Ventor and Sobieszczanski-Sobieski, 2003).

436 2.5.4 Genetic algorithm

437 The genetic algorithm (GA) is a heuristic search method based on natural selection and
438 evolution principles and concepts. This method was introduced by John Holland in the mid-
439 seventies, inspired by Darwin's theory of descent with modification by natural selection. To
440 determine the optimal set of parameters GA mimics the reproduction behavior of the biological
441 populations in nature. It has been proven effective for the selection process in solving cutting-edge
442 optimization problems. It can also handle regular and irregular variables, non-traditional data
443 partitioning, and non-linear objective functions without requiring gradient information (Hassan et
444 al., 2004). The basic steps for the PSO algorithm are given below:

- 445 1. The determination of the maximum outcomes from an objective function, the genetic
446 algorithm uses the following function (Mayer and Lobet, 2018):

$$447 \quad f = f((y_1 + y_2), \dots, (y_n + y_{n+1})_n) \quad (21)$$

448 where n is the number of decision variables $y_i \in [y_i^{min}, y_i^{max}]$ with a discretization step δy_i . The
449 initial boundary conditions y_i^{min}, y_i^{max} determined in the beginning of the simulation. δy_i is the
450 determines the physical parameters y_i performances in the experiment. These decision variables
451 are represented by a sequence of binary digits (*GENES*).

- 452 2. The decisions variables are given within initial boundary conditions $y_i = y_i^{min} +$
453 $(GENE\ i) * \delta y_i$, where $GENE\ i \in [0, 2^{n_i} - 1]$ refers to the value of GENES. n_i is the bit
454 length of each GENE, which is the first integer where $y_i^{min} + 2^{n_i} - 1 * \delta y_i \geq y_i^{max}$. The
455 total number of bits in each DNA refers $n_{sum} = \sum_{i=1}^n n_i$. The algorithm process starts with
456 a random selection of objectives. After evaluation of each objective in the fitness function
457 $f = f((y_1 + y_2), \dots, (y_n + y_{n+1})_n)$, and rank them from best to worst.

458 The genetic similarity determines the selection progress indicator. These random individual
459 objectives with rank are transferred to the next generation. The remaining individuals participate
460 in the steps of selection, crossover, and mutation. The individual objective parent selection process
461 can happen several times, and this can be achieved by many different schemes, such as the roulette-
462 wheel ranked method. For any pair of objective parents' selection, crossover, and mutation process
463 of next-generation is defined. After that, the fitness f of all individuals scheduled for the next
464 generation is evaluated. This process repeats from generation to generation until a termination
465 criterion is met.

466 GA methodology is quite similar to another stochastic searching algorithm PSO. Both
467 methods begin their search from a randomly generated population of designs that evolve over
468 successive generations. They do not require any specific starting point for the simulation. The first
469 operator is the "selection" procedure similar to "Survival for the Fittest" principle. The second
470 operator is the "Crossover" operator, which mimics mating in a biological population. Both
471 methods use the same convergence criteria for selecting the optimal solution in the problem space
472 (Hassan et al., 2004). However, GA differs in two ways from the most traditional optimization
473 methods. First, GA does not operate directly on the design parameter vector but a symbolic
474 parameter known as a chromosome. Second, it optimizes the whole design chromosomes at once,
475 unlike other optimization methods single chromosome at a time (Weile and Michielssen, 1997).

476 ***2.5.5 Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN)***

477 The complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN)
478 decomposition approach initiates by discretizing the n -length predictors of any model $\chi(t)$ into
479 IMFs (intrinsic model functions) and residues to conform with tolerability. However, to ensure no
480 information leakage in the IMFs and residues, the decomposition is performed separately by taking

481 training and testing subsets. The actual IMF is produced by taking the mean of the Empirical mode
 482 decomposition (EMD)-grounded IMFs across a trial and combining white noise to model the
 483 predictor-target variables. The CEEMDAN is used in machinery, electricity, and medicine such as
 484 impact signal denoising, daily peak load forecasting, health degradation monitoring for rolling
 485 bearings, friction signal denoising combined with mutual information (Li et al., 2019).

486 The CEEMDAN process is as follows:

487 **Step 1:** The decomposition of p -realizations of $\chi[n] = \varepsilon_1 \omega^p[n]$ using EMD to develop their first
 488 intrinsic approach, as explained according to the equation:

$$489 \quad \widehat{IMF}_1[n] = \frac{1}{p} \sum_{p=1}^p IMF_1^p[n] = \overline{IMF}_1[n] \quad (23)$$

490 **Step 2:** Putting $k = 1$, the 1st residue is computed following Eq. (1).

$$491 \quad Res_1[n] = \chi[n] - \widehat{IMF}_1[n] \quad (24)$$

492 **Step 3:** Putting $k = 2$, the 2nd residual is obtained as

$$493 \quad \widehat{IMF}_2[n] = \frac{1}{p} \sum_{p=1}^p E_1(r_1[n] + \varepsilon_1 E_1(\omega^p[n])) \quad (25)$$

494 **Step 4:** Setting $k = 2 \dots K$ calculates the k^{th} residue as.

$$495 \quad Res_k[n] = Res_{k-1}[n] - \widehat{IMF}_k[n] \quad (26)$$

496 **Step 5:** Now, we decompose the realizations $Res_k[n] + \varepsilon_1 E_1(\omega^p[n])$, Here, $k = 1, \dots K$ until
 497 their first model of EMD is reached; Here the $(k + 1)$ is

$$498 \quad \widehat{IMF}_{(k+1)}[n] = \frac{1}{p} \sum_{p=1}^p E_1(r_k[n] + \varepsilon_k E_k(\omega^p[n])) \quad (27)$$

499 **Step 6:** Now, the k value is incremented, and steps 4–6 are repeated. Consequently, the final
500 residue is achieved:

$$501 \quad RES_k[n] = \chi[n] - \sum_{k=1}^K \widehat{IMF}_k \quad (28)$$

502 Here, K is defined as the limiting case (*i.e.*, the highest number of modes). To comply with the
503 replicability of the earliest input, $\chi[n]$, the following is performed for the CEEMDAN approach.

$$504 \quad \chi[n] = \sum_{k=1}^K \widehat{IMF}_k + RES_k[n] \quad (29)$$

505

506 **2.6 Model implementation procedure**

507 It is crucial to optimize the objective model's architecture to incorporate the relationship between
508 predictors and model. A multi-phase CNN-GRU and GRU model is included, built using Python-
509 based deep learning packages such as *TensorFlow* and *Keras*. A total of nine statistical metrics
510 was used to investigate the forecasting robustness of the models that have been integrated. The
511 model was powered by an Intel i7 @ 3.6GHz processor and 16 GB of memory. Deep learning
512 libraries like *Keras* (Brownlee, 2016; Ketkar, 2017) and *TensorFlow* (Abadi et al., 2016) were
513 used to demonstrating algorithms for the proposed models. In addition, packages like *matplotlib*
514 (Barrett et al., 2005) and *seaborn* (Waskom, 2021) were used for visualization.

515 The determination of the model's valid predictors does not have any precise formula.
516 However, the literature suggests three methods, *i.e.*, trial-and-error, the autocorrelation function
517 (ACF) and partial autocorrelation function (PACF), and the cross-correlation function (CCF)
518 approaches, for selecting lagged UVI memories and predictors to make an optimal model. In this
519 study, the PACF was used to determine significant antecedent behavior in terms of the lag of UVI
520 (Tiwari and Adamowski, 2013; Tiwari and Chatterjee, 2010). Figures 6b and 7b demonstrated the

521 PACF for UVI time series showing the antecedent behavior in terms of the lag of UVI and
522 decomposed UVI (i.e., IMF_n) where antecedent daily delays are apparent. Generally, the CCF
523 selects the input signal pattern based on the predictors' antecedent lag (Adamowski et al., 2012).
524 The CCF determined the predictors' statistical similarity to the target variable (Figures 6a and 7a).
525 A set of significant input combinations was selected after evaluating each predictor's r_{cross} with
526 UVI. The figure shows that the highest correlation between p data and UVI was found for all
527 stations at lag zero (i.e., $r_{\text{cross}} = 0.22 - 0.75$). AOD and GBI both demonstrated significant r_{cross}
528 from 0.65 to 0.80 and 0.68 to 0.75, respectively. Some predictors with insignificant lags such as
529 AO, CT, and OTC were also considered to increase the predictors' diversity. The CCF with UVI
530 with predictors significantly varied for all other stations. However, selecting lags from the cross-
531 correlation function is identical to that used in the objective stations.

532 As mentioned, the CEEMDAN method was used to decompose the data sets. The daily
533 time series of UVI data and predictor variables were decomposed into respective daily IMFs and
534 a residual component using CEEMDAN procedures. The example of the IMFs and the residual
535 component of the respective CEEMDAN is shown in Figure 3. PACF was applied to the daily
536 IMFs and residual component time series generated in the above process. An individual input
537 matrix was created for each IMF, and the residual component was made up based on the significant
538 lagged memory with that of IMF of target UVI. These separate input matrices were used to forecast
539 the respective future IMFs and the residual component. Next, the anticipated IMFs and residuals
540 were combined to produce daily forecasts of UVI values. Note that the CEEMDAN
541 transformations are completely self-adaptive and data-dependent multi-resolution techniques. As
542 such, the number of IMFs and the residual component generated are contingent on the nature of
543 the data.

544 The predictor variables were used to forecast the UVI were normalized between 0 and 1 to
545 minimize the scaling effect of different variables as follows:

$$546 \quad U_{norm} = \frac{U - U_{min}}{U_{max} - U_{min}} \quad (30)$$

547 In Eq. (30), U is the respective predictors, U_{min} is the minimum value for the predictors, U_{max} is
548 the maximum value of the data and U_{norm} is the normalized value of the data. After normalizing
549 the predictor variables, the data sets were partitioned: 70% of the data sets were considered training
550 data, 15% were used for testing, and the remaining 15% of the data sets were considered validation
551 data. The LSTM model was followed by developing a hybrid LSTM model with 3-layered CNN
552 and 4-layered LSTM, as illustrated in Figure 2. Using the conventional models, the traditional
553 antecedent lagged matrix of the daily predictors' variables was applied. The prior application of
554 the optimization algorithm was made before using CCF and PACF and before significant
555 predictors being removed from the model. Table 2 shows the selected predictors using four
556 optimization techniques in association with the UVI.

557 **2.7 Model performance assessment**

558 In this study, the effectiveness of the deep learning hybrid model was assessed using a variety of
559 performance evaluation criteria, *e.g.*, Pearson's Correlation Coefficient (r), root mean square error
560 ($RMSE$), Nash- Sutcliffe efficiency (NS) (Nash and Sutcliffe, 1970), and mean absolute error
561 (MAE). The relative $RMSE$ (denoted as $RRMSE$) and relative MAE (denoted as $RMAE$) were used
562 to explore the geographic differences between the study stations.

563 The exactness of the relationship between predicted and observed values was used to
564 evaluate the effectiveness of a predictive model. When the error distribution in the tested data is
565 Gaussian, the root mean square error ($RMSE$) is a more appropriate measure of model performance

566 than the mean absolute error (MAE) (Chai and Draxler, 2014), but for an improved model
 567 evaluation, Legates-McCabe's (*LM*) Index are used as more sophisticated and compelling
 568 measures (Legates and McCabe, 2013; Willmott et al., 2012). Mathematically, the metrics are as
 569 follows:

570 i) Correlation coefficient (*r*):

$$571 \quad r = \left\{ \frac{\sum_{i=1}^N (UVI_{obs} - \overline{UVI}_{obs})(UVI_{for} - \overline{UVI}_{for})}{\sqrt{\sum_{i=1}^N (UVI_{obs} - \overline{UVI}_{obs})^2 \sum_{i=1}^N (UVI_{for} - \overline{UVI}_{for})^2}} \right\}^2 \quad (31)$$

572 ii) Mean absolute error (*MAE*):

$$573 \quad MAE = \frac{1}{N} \sum_{i=1}^N |UVI_{for} - UVI_{obs}| \quad (32)$$

574 iii) Root mean squared error (*RMSE*):

$$575 \quad RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (UVI_{for} - UVI_{obs})^2} \quad (33)$$

576 iv) Nash-Sutcliffe Efficiency (*NS*):

$$577 \quad NSE = 1 - \left[1 - \frac{\sum_{i=1}^N (UVI_{for})^2}{\sum_{i=1}^N (UVI_{obs} - \overline{UVI}_{for})^2} \right] \quad (34)$$

578 v) Legates-McCabe's Index (*LM*):

$$579 \quad LM = 1 - \left[\frac{\sum_{i=1}^N |UVI_{for} - UVI_{obs}|}{\sum_{i=1}^N ||UVI_{obs} - \overline{UVI}_{obs}||} \right] \quad (35)$$

580 vi) Relative Root Mean Squared Error (*RRMSE*, %):

$$581 \quad RRMSE(\%) = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (UVI_{for} - UVI_{obs})^2}}{\frac{1}{N} \sum_{i=1}^N (UVI_{obs})} \times 100 \quad (36)$$

582 vii) Relative Mean Absolute Error (*RMAE*, %):

$$583 \quad RMAE (\%) = \frac{\frac{1}{N} \sum_{i=1}^N |UVI_{for} - UVI_{obs}|}{\frac{1}{N} \sum_{i=1}^N (UVI_{obs})} \times 100 \quad (37)$$

584 In Eq. (31–37), UVI_{obs} and UVI_{for} represents the observed and forecasted values for i^{th} test
 585 value; \overline{UVI}_{obs} and \overline{UVI}_{for} refer to their averages, accordingly, and N is defined as the number of
 586 observations, while the CV stands for the coefficient of variation.

587 **3. Results**

588

589 This section describes results obtained from the proposed hybrid deep learning model (i.e.,
 590 CEEMDAN-CLSTM) and other hybrid models (i.e., CEEMDAN-CGRU, CEEMDAN-LSTM,
 591 CEEMDAN-GRU, CEEMDAN-DT, CGRU, and CLSTM), and the standalone LSTM, GRU, DT,
 592 MLP, and SVR models. The evolutionary algorithms (i.e., ACO, DEV, GA, and PSO) were
 593 incorporated to obtain the optimum features in model building. Seven statistical metrics from Eqs.
 594 (31) to (37) were used to analyze the models in the testing dataset and visual plots to justify the
 595 forecasted results' justification.

596 **3.1 The evaluation of hybrid and standalone models**

597 The hybrid deep learning model (i.e., CEEMDAN-CLSTM) demonstrated high r and NS
 598 values and low RMSE and MAE compared to their standalone models (Table 3). The best overall
 599 performance was recorded using the CEEMDAN-CLSTM model with the Genetic Algorithm with
 600 the highest correlation (i.e., $r = 0.996$), the highest data variance explained (i.e., $NS = 0.997$), and
 601 the lowest errors (i.e., $RMSE = 0.162$ and $MAE = 0.119$). The performance was followed by the
 602 same model with Partial Swarm Optimisation (i.e., $r \approx 0.996$; $NS \approx 0.992$; $RMSE \approx 0.216$; MAE
 603 ≈ 0.163) and Ant Colony Optimisation (i.e., $r \approx 0.996$; $NS \approx 0.993$; $RMSE \approx 0.220$; $MAE \approx 0.165$).
 604 The single deep learning models (i.e., LSTM and GRU) performed better than the single machine

605 learning models (i.e., DT, SVR, and MLP). Moreover, the hybrid deep learning models without a
606 CNN (i.e., CEEMDAN-GRU and CEEMDAN-GRU) also demonstrated higher forecasting
607 accuracy (i.e., $r = 0.973 - 0.993$; $RMSE = 0.387 - 0.796$) in comparison with standalone deep
608 learning models (i.e., $r \approx 0.959 - 0.981$; $RMSE \approx 0.690 - 0.986$). The following models'
609 performance is then predicted by the CNN-GRU, CEEMDAN-GRU, and GRU models in that
610 order.

611 3.2 The selection of the best model

612 RRMSE and LM for all tested models were used to assess the robustness of the proposed
613 hybrid models as well as for comparisons. The magnitude of RRMSE (%) and LM for the objective
614 model (CEEMDAN-CLSTM) shown in Figure 6 indicates that the proposed hybrid model
615 performed significantly better than other benchmark models. The RRMSE and LM values ranged
616 between 2 and 3.5 percent and between 0.982 and 0.991, respectively. The performance indices
617 (i.e., RRMSE and LM) using four optimization algorithms were higher for the CEEMDAN-CGRU
618 model. Overall, the CEEMDAN-CLSTM model with the GA optimization methods provided the
619 best performance (i.e., $RRMSE = \sim 2.0\%$; $LM = 0.991$), indicating its high efficiency in forecasting
620 the future UV-Index a higher degree of accuracy.

621 A precise comparison of forecasted and observed UVI can also be seen by examining the
622 scatterplot of forecasted (UVI_{for}) and observed (UVI_{obs}) UVI for four optimization algorithms (i.e.,
623 ACO, PSO, DEV, and GA) (Figure 7). Here, scatter plots showed the coefficient of determination
624 (r^2) and a least-square fitting line, along with the equation for UVI and an observed UVI close to
625 the forecasted UVI. As demonstrated in Figure 7, it also appears that the proposed hybrid model
626 performed better when compared with other applied models. However, among the four

627 optimization techniques applied, the hybrid deep learning model (i.e., CEEMDAN-CLSTM)
628 optimized with the GA outperformed the other models in forecasting the UVI. The hybrid
629 CEEMDAN-CLSTM model calculated magnitudes from the GA, which came the closest to unity,
630 with an $m|r^2$ of 0.976|0.995 in pairs. The performance is followed by ACO and DEV algorithms
631 with a potential pair (ACO: 0.975|0.995; DEV: 0.966|0.994). The outliers (i.e., the extremes) are
632 closer to the fitted line, while the y-intercept (i.e., the starting point) is approximately 0.05 units
633 away from zero (0) using the GA method. The other models had outliers, resulting in their
634 intercepts deviating from the ideal value. In conclusion, the CEEMDAN-CLSTM model
635 performed the best for the GA.

636 The proposed hybrid deep learning model (i.e., CEEMDAN-CLSTM) was further assessed
637 employing the ECDFs of absolute forecast error ($|FE|$) (Figure 8). Total 95% forecasted values
638 using the CEEMDAN-CLSTM model with GA demonstrated a small error ranged between 0.01
639 and 0.299, with a substantially larger error for the CCGRU model (i.e., 0.477), followed by the
640 CLSTM model (i.e., 0.626) and CGRU (i.e., 1.104). For the other optimization algorithms, nearly
641 the same level of performance was observed. Predictions ranging between the 95th and 98th
642 percentile were preferred over objective models, which performed best in the current forecast. A
643 hybrid CLSTM model was found to be the most accurate for upcoming extra-terrestrial intelligent
644 beings. However, Figure 9 showed the effect of applying CEEMDAN as a feature extraction
645 method of data on the percent change in RMAE values within the testing phase of UVI forecast
646 incrementally. The contribution of the data decomposition method (i.e., CEEMDAN) was
647 significant in the model implementation. The decrement of RMAE in percent using GA was found
648 between 17% to 63%, whereas the CLSTM showed the highest percentage of decrement (i.e.,
649 63%). Moreover, the PSO optimized model showed that the *RMAE* (%) values with the deep

650 learning model appeared to decrease by ~2 to 60%, and the lowest decreasing RMAE was found
651 for the ACO algorithm with a reduction ~3% to 36%. However, the CLSTM model using four
652 optimization methods showed the highest improvement among all the deep learning approaches
653 that reduced the RMSE from 36% to 63%. It is worth mentioning that the percent increase in
654 *RMAE* was ~83% for the DEV algorithm using the SVR method. Overall, the CEEMDAN, as a
655 data decomposition algorithm for UVI forecasting with four optimization algorithms, showed
656 significant improvement over the testing phase.

657 After additional analysis, the forecasted-to-observed UVI and absolute forecasting errors
658 are displayed in Figure 10. The absolute forecasted error has a maximum dispersion of ($|FE| =$
659 $|UVI_{for} - UVI_{obs}|$). The box plot demonstrated the data dispersal of the observed and forecasted
660 UVI from the proposed deep learning approaches and other comparing models. Figure 10 provides
661 a clear visualization of the data concerning quartiles distinctly outliers. The lower end of the plot
662 lies between the lower quartile (25th percentile) and upper quartile (75th percentile). It is evident
663 that the median of the forecasted and the observed UVI for the CEEMDAN-CLSTM model with
664 the GA optimization.

665 Moreover, the DEV-based CEEMDAN-CLSTM model showed identical forecasting to the
666 GA-based CEEMDAN-CLSTM model with a slight variation. A more in-depth inspection of the
667 absolute forecasted error ($|FE|$) from the hybrid CEEMDAN-CLSTM model for two optimizations
668 (i.e., GA and DEV) further strengthens the suitability of the hybrid CLSTM approach in
669 forecasting the UVI of Perth station of Australia with the narrowest distribution in comparison
670 with other models. A significant percentage (98%) of the $|FE|$ in the first error brackets
671 ($0 < |FE| < 0.15$) was observed for the GA-based CEEMDAN-CLSTM model, while for the DEV-
672 based model, the percentage is 95%.

673 With the help of a time series plot, we can better understand forecasting ability and refine
674 the proposed model, taking it from standalone to hybrid model. The time series plot of forecasted
675 and observed UVI using CEEMDAN-CLSTM optimized by four optimization methods is depicted
676 in Figure 11. The results showed that the proposed GA-based CEEMDAN-CLSTM model is close
677 to the observed UVI, indicating that the model has high predictive accuracy. The application of
678 the GA in the model optimization resulted in a significant improvement in forecasted UVI. For
679 other algorithms that use the CEEMDAN-CLSTM model, it is discovered that the forecasted UVI
680 is accurate when compared to the other optimization methods.

681 Finally, Figure 12 presents a comprehensive interpretation by illustrating the absolute
682 forecasting error frequency distributions ($|FE|$) using all GA-based models for Perth stations of
683 Australia. It is apparent from Figure 12 that the CEEMDAN-CLSTM model provided significantly
684 improved distributions with the maximum 98% forecasting error ($|FE|$) within the first error
685 brackets ($0 < |FE| < 0.10$). It is also noteworthy that the CEEMDAN-CGRU model showed a higher
686 percentage of $|FE|$ between 0 and 0.25 of all forecasting yielded a considerably small error and the
687 remaining 15% of simultaneously produced forecasting error between 0.25 and 1.0. The highest
688 forecasted error was found for machine learning models when the $|FE|$ value for all models (i.e.,
689 SVR, MLP, and DT) was considered.

690

691 **4. Discussion**

692 The establishment of robust predictive modelling of the UV index and physical
693 interpretation is critical for various practical applications, such as helping policymakers in their

694 daily health impact assessment. These systems emulate how a human expert would solve a
695 complex forecasting problem by reasoning through a set of UVI-related predictors rather than
696 through conventional or procedural methods. These methods warrant continuous measurement of
697 irradiance or radiative transfer models, which are tedious (as discussed in the introduction) and
698 often inaccurate. This study demonstrated the efficacy of hybrid deep learning methods in
699 forecasting UVI on a near real-time horizon. The study site was in Perth, Western Australia,
700 Australia, where skin cancer is significantly high. An accurate forecasting system in this region is
701 therefore essential.

702 To function effectively, alert systems must generate accurate irradiance forecasts, but UVI
703 is generally determined by many factors (i.e., the solar zenith, altitude, cloud fraction, aerosol and
704 optical properties, albedo, and vertical ozone profile) (Deo et al., 2017). The study extensively
705 utilized four optimization techniques (i.e., GA, ACO, DEV, and PSO) to have optimum predictors
706 used in UVI forecasting. The incorporated predictors from three distinct data sets (i.e., SILO,
707 MODIS, and CI) were optimized. The optimization techniques selected a diversified list of
708 variables except for RMM1 and RMM2, as four algorithms select them both. The predictors like
709 ozone total column, AOD, and cloud fraction were significant using the GA algorithm. In most
710 cases, the hydro-meteorological variables were insignificant by all four algorithms that agree with
711 UV concentration's general concept. The objective algorithm (i.e., GA) selected SOI, GBI, AAO,
712 Nino4, Nino12, RMM1, and RMM2 as potential predictors as well. The ground-based
713 measurements and modelling studies are essential (Alados et al., 2007, 2004), but are challenging
714 to implement in practice. Furthermore, secondary factors affecting UV levels (i.e., clouds or
715 aerosols) are rarely known with sufficient precision. Considering the practical feasibility, an
716 algorithm that is data-efficient, simple to develop, flexible, and user-friendly should be considered

717 as a viable alternative for information (Igoe et al., 2013a, 2013b; Parisi et al., 2016). Therefore,
718 our developed forecasting model will play a vital role in the decision-makers to adopt prompt
719 measures without difficulties.

720 The proposed hybrid deep learning network (i.e., CEEMDAN-CLSTM) for predicting
721 surface UV radiation also demonstrated low errors in forecasting, i.e., showing around 10% error
722 for the next-day forecast and 13–16% error for 7-day up to the 4-week forecast. This further affirms
723 that the quantitative UV forecast is appropriate for heliotherapy applications, which tolerates up
724 to 10–25% error levels. The CEEMDAN-CLSTM's performance is competitive on UV data from
725 multiple regions. Thus, the CEEMDAN-CLTSM model can be adapted to forecast other useful
726 UV action spectra, such as vitamin D production and erythematous UV index. A key limitation of
727 machine learning is its overfitting tendency on the training dataset and often does not generalize
728 well to other datasets from different distributions. In the context of UV forecasting, this dictates
729 that the model must be retrained with data from the weather station to be used for that geographic
730 region. In a geographical region with the highly variable weather condition, such as London in
731 2019, artificial neural network models' performance dropped significantly (Raksasat et al., 2021).
732 This capability of the model to extract seasonal patterns may also explain why the addition of
733 ozone, cloud fraction, and AOD information significantly improved the performance of
734 CEEMDAN-CLSTM, particularly when the GA algorithm was applied.

735 **5. Conclusion**

736 This study conducted a daily UV Index forecasting at Perth station using aggregated
737 significant antecedent satellite-driven variables associated with UV irradiance. The forecasting
738 was made using a novel hybrid deep learning model (i.e., CEEMDAN-CLSTM) and compared

739 with other benchmark models such as LSTM, GRU, DT, SVR, etc. Four optimization methods
740 were employed to extract the crucial features of the response variable (i.e., the UVI). After
741 applying the proposed model and benchmarked models, the model's merits were evaluated using
742 different statistical metrics (e.g., r , WI, MAE, NS, and LM), graphical plots, and relevant
743 discussions. The key findings are summarized as follows:

- 744 • The CEEMDAN-CLSTM hybrid model demonstrated excellent forecasting ability
745 compared to its counterpart models.
- 746 • The GA optimization algorithm is appeared to be an attractive option for selecting
747 mechanistically meaningful features of the dependable variable compared to the other three
748 optimization techniques.
- 749 • The performance metrics showed that the GA and CEEMDAN-optimized models had
750 better performance and higher efficiency metrics (i.e., r , NS, and LM) and lower error
751 metrics (i.e., MSE and RMSE).
- 752 • However, in UVI forecasting, the standalone models' (i.e., LSTM, GRU, DT, and SVR)
753 performances were poor compared to the proposed hybrid model.

754 Adapted to an Australian climate in the sub-tropics during peak summer-time conditions,
755 applying a CLSTM model to forecast the UVI is a novel deep learning approach. The forecasts
756 derived from our data were within one UVI unit of the actual measured values indicating the
757 remarkable forecasting capability. Therefore, this data-driven model would be of tremendous help
758 for the decision-makers to promptly protect public health without delay. It has the tremendous
759 potential to be adopted by a more significant segment of the community, particularly children and

760 the elderly facing a greater risk of developing skin cancer (i.e., melanoma) in the Australian region
761 and worldwide.

762 **Acknowledgement**

763 We would like to thank the Australian Radiation Protection and Nuclear Safety Agency for
764 providing the data. This study did not receive any external funding.

765

766 **Credit authorship contribution statement**

767 **A. A. Masrur Ahmed:** Writing - original draft, Conceptualization, Methodology, Software,
768 Model development, and application. **Mohammad Hafez Ahmed:** Conceptualization, Model
769 Application, Writing – draft, review & editing. **Sanjoy Kanti Saha:** Writing - review & editing.
770 **Oli Ahmed:** Data Collection, Writing - review & editing, **Ambica Sutradhar:** Data Collection,
771 Writing - review & editing

772

773 **Reference**

- 774 Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving,
775 G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner,
776 B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., 2016.
777 TensorFlow: A System for Large-Scale Machine Learning. Presented at the 12th
778 {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16),
779 pp. 265–283.
- 780 Adamowski, J., Chan, H.F., Prasher, S.O., Ozga-Zielinski, B., Sliusarieva, A., 2012. Comparison
781 of multiple linear and non-linear regression, autoregressive integrated moving average,
782 artificial neural network, and wavelet artificial neural network methods for urban water
783 demand forecasting in Montreal, Canada. *Water Resources Research* 48.
784 <https://doi.org/10.1029/2010WR009945>
- 785 Ahmed, A.A.M., Deo, R.C., Ghahramani, A., Raj, N., Feng, Q., Yin, Z., Yang, L., 2021. LSTM
786 integrated with Boruta-random forest optimiser for soil moisture estimation under
787 RCP4.5 and RCP8.5 global warming scenarios. *Stoch Environ Res Risk Assess.*
788 <https://doi.org/10.1007/s00477-021-01969-3>
- 789 Ahmed, M.H., Lin, L.-S., 2021. Dissolved oxygen concentration predictions for running waters
790 with different land use land cover using a quantile regression forest machine learning
791 technique. *Journal of Hydrology* 597, 126213.
792 <https://doi.org/10.1016/j.jhydrol.2021.126213>

793 Alados, I., Gomera, M.A., Foyo-Moreno, I., Alados-Arboledas, L., 2007. Neural network for the
794 estimation of UV erythemal irradiance using solar broadband irradiance. *International*
795 *Journal of Climatology* 27, 1791–1799. <https://doi.org/10.1002/joc.1496>
796 Alados, I., Mellado, J.A., Ramos, F., Alados-Arboledas, L., 2004. Estimating UV Erythemal
797 Irradiance by Means of Neural Networks. *Photophys. phot* 80, 351–358.
798 <https://doi.org/10.1562/2004-03-12-RA-111.1>
799 Alfadda, A., Rahman, S., Pipattanasomporn, M., 2018. Solar irradiance forecast using aerosols
800 measurements: A data driven approach. *Solar Energy* 170, 924–939.
801 <https://doi.org/10.1016/j.solener.2018.05.089>
802 Allaart, M., Weele, M. van, Fortuin, P., Kelder, H., 2004. An empirical model to predict the UV-
803 index based on solar zenith angles and total ozone. *Meteorological Applications* 11, 59–
804 65. <https://doi.org/10.1017/S1350482703001130>
805 Anderiesz, C., Elwood, M., Hill, D.J., 2006. Cancer control policy in Australia. *Aust New*
806 *Zealand Health Policy* 3, 12. <https://doi.org/10.1186/1743-8462-3-12>
807 Antanasijević, D., Pocajt, V., Perić-Grujić, A., Ristić, M., 2014. Modelling of dissolved oxygen
808 in the Danube River using artificial neural networks and Monte Carlo Simulation
809 uncertainty analysis. *Journal of Hydrology* 519, 1895–1907.
810 <https://doi.org/10.1016/j.jhydrol.2014.10.009>
811 ARPANSA, 2021. Australian Radiation and Nuclear Protection Agency 2021 Realtime UV
812 levels [WWW Document]. ARPANSA. URL [https://www.arpansa.gov.au/our-](https://www.arpansa.gov.au/our-services/monitoring/ultraviolet-radiation-monitoring/ultraviolet-radiation-index)
813 [services/monitoring/ultraviolet-radiation-monitoring/ultraviolet-radiation-index](https://www.arpansa.gov.au/our-services/monitoring/ultraviolet-radiation-monitoring/ultraviolet-radiation-index) (accessed
814 7.9.21).
815 Barrett, P., Hunter, J., Miller, J.T., Hsu, J.-C., Greenfield, P., 2005. matplotlib -- A Portable
816 Python Plotting Package 347, 91.
817 Baumgaertner, A.J.G., Seppälä, A., Jöckel, P., Clilverd, M.A., 2011. Geomagnetic activity
818 related NO_x enhancements and polar surface air temperature variability in a chemistry
819 climate model: modulation of the NAM index. *Atmospheric Chemistry and Physics* 11,
820 4521–4531. <https://doi.org/10.5194/acp-11-4521-2011>
821 BOM, 2020. Australia's official weather forecasts & weather radar - Bureau of Meteorology
822 [WWW Document]. URL <http://www.bom.gov.au/> (accessed 7.9.21).
823 Boniol, M., 2016. Descriptive epidemiology of skin cancer incidence and mortality, in: *Skin*
824 *Cancer Prevention*. CRC Press, pp. 221–242.
825 Brownlee, J., 2016. *Deep Learning With Python: Develop Deep Learning Models on Theano and*
826 *TensorFlow Using Keras*. Machine Learning Mastery.
827 Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? –
828 Arguments against avoiding RMSE in the literature. *Geoscientific Model Development* 7,
829 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>
830 Chen, C., Jiang, H., Zhang, Y., Wang, Y., 2010. Investigating spatial and temporal
831 characteristics of harmful Algal Bloom areas in the East China Sea using a fast and
832 flexible method, in: 2010 18th International Conference on Geoinformatics. Presented at
833 the 2010 18th International Conference on Geoinformatics, pp. 1–4.
834 <https://doi.org/10.1109/GEOINFORMATICS.2010.5567490>
835 Chen, J.P., Yang, L., Wang, L.K., Zhang, B., 2006. Ultraviolet Radiation for Disinfection, in:
836 Wang, L.K., Hung, Y.-T., Shammas, N.K. (Eds.), *Advanced Physicochemical Treatment*
837 *Processes, Handbook of Environmental Engineering*. Humana Press, Totowa, NJ, pp.
838 317–366. https://doi.org/10.1007/978-1-59745-029-4_10

839 Deo, R.C., Downs, N., Parisi, A.V., Adamowski, J.F., Quilty, J.M., 2017. Very short-term
840 reactive forecasting of the solar ultraviolet index using an extreme learning machine
841 integrated with the solar zenith angle. *Environmental Research* 155, 141–166.
842 <https://doi.org/10.1016/j.envres.2017.01.035>

843 Dey, R., Salem, F.M., 2017. Gate-variants of Gated Recurrent Unit (GRU) neural networks, in:
844 2017 IEEE 60th International Midwest Symposium on Circuits and Systems
845 (MWSCAS). Presented at the 2017 IEEE 60th International Midwest Symposium on
846 Circuits and Systems (MWSCAS), pp. 1597–1600.
847 <https://doi.org/10.1109/MWSCAS.2017.8053243>

848 Downs, N., Butler, H., Parisi, A., 2016. Solar Ultraviolet Attenuation during the Australian (Red
849 Dawn) Dust Event of 23 September 2009. *Bulletin of the American Meteorological
850 Society* 97, 2039–2050. <https://doi.org/10.1175/BAMS-D-15-00053.1>

851 Eberhart, Shi, Y., 2001. Particle swarm optimization: developments, applications and resources,
852 in: *Proceedings of the 2001 Congress on Evolutionary Computation (IEEE Cat.
853 No.01TH8546)*. Presented at the Proceedings of the 2001 Congress on Evolutionary
854 Computation (IEEE Cat. No.01TH8546), pp. 81–86 vol. 1.
855 <https://doi.org/10.1109/CEC.2001.934374>

856 Fan, J., Wu, L., Ma, X., Zhou, H., Zhang, F., 2020. Hybrid support vector machines with
857 heuristic algorithms for prediction of daily diffuse solar radiation in air-polluted regions.
858 *Renewable Energy* 145, 2034–2045. <https://doi.org/10.1016/j.renene.2019.07.104>

859 Fernández-Delgado, M., Cernadas, E., Barro, S., Ribeiro, J., Neves, J., 2014. Direct Kernel
860 Perceptron (DKP): Ultra-fast kernel ELM-based classification with non-iterative closed-
861 form weight calculation. *Neural Networks* 50, 60–71.
862 <https://doi.org/10.1016/j.neunet.2013.11.002>

863 Fouilloy, A., Voyant, C., Notton, G., Motte, F., Paoli, C., Nivet, M.-L., Guillot, E., Duchaud, J.-
864 L., 2018. Solar irradiation prediction with machine learning: Forecasting models
865 selection method depending on weather variability. *Energy* 165, 620–629.
866 <https://doi.org/10.1016/j.energy.2018.09.116>

867 Furuhashi, T., Torii, K., Ikumi, K., Kato, H., Nishida, E., Morita, A., 2020. Ultraviolet A1
868 phototherapy for the treatment of localized scleroderma. *The Journal of Dermatology* 47,
869 792–795. <https://doi.org/10.1111/1346-8138.15368>

870 Giovanni [WWW Document], 2021. URL <https://giovanni.gsfc.nasa.gov/giovanni/> (accessed
871 7.9.21).

872 Gray, N.F., 2014. Chapter Thirty-Four - Ultraviolet Disinfection, in: Percival, S.L., Yates, M.V.,
873 Williams, D.W., Chalmers, R.M., Gray, N.F. (Eds.), *Microbiology of Waterborne
874 Diseases (Second Edition)*. Academic Press, London, pp. 617–630.
875 <https://doi.org/10.1016/B978-0-12-415846-7.00034-2>

876 Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., Lew, M.S., 2016. Deep learning for visual
877 understanding: A review. *Neurocomputing, Recent Developments on Deep Big Vision*
878 187, 27–48. <https://doi.org/10.1016/j.neucom.2015.09.116>

879 Hassan, R., Cohanin, B., Weck, O. de, Venter, G., 2004. A Comparison of Particle Swarm
880 Optimization and the Genetic Algorithm, in: 46th AIAA/ASME/ASCE/AHS/ASC
881 Structures, Structural Dynamics and Materials Conference. American Institute of
882 Aeronautics and Astronautics. <https://doi.org/10.2514/6.2005-1897>

883 Heilingloh, C.S., Aufderhorst, U.W., Schipper, L., Dittmer, U., Witzke, O., Yang, D., Zheng, X.,
884 Sutter, K., Trilling, M., Alt, M., Steinmann, E., Krawczyk, A., 2020. Susceptibility of

885 SARS-CoV-2 to UV irradiation. *American Journal of Infection Control* 48, 1273–1275.
886 <https://doi.org/10.1016/j.ajic.2020.07.031>

887 Hendon, H., Salby, M., 1994. The Life Cycle of the Madden–Julian Oscillation.
888 [https://doi.org/10.1175/1520-0469\(1994\)051<2225:TLCOTM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1994)051<2225:TLCOTM>2.0.CO;2)

889 Hochreiter, S., Schmidhuber, J., 1997. Long Short-Term Memory. *Neural Computation* 9, 1735–
890 1780. <https://doi.org/10.1162/neco.1997.9.8.1735>

891 Hollaender, A., Buy, H.G.D., Ingraham, H.S., Wheeler, S.M., 1944. Control of Air-Borne
892 Microorganisms by Ultraviolet Floor Irradiation. *Science* 99, 130–131.
893 <https://doi.org/10.1126/science.99.2563.130>

894 Huang, X., Zhang, C., Li, Q., Tai, Y., Gao, B., Shi, J., 2020. A Comparison of Hour-Ahead Solar
895 Irradiance Forecasting Models Based on LSTM Network [WWW Document].
896 *Mathematical Problems in Engineering*. <https://doi.org/10.1155/2020/4251517>

897 Igoe, D., Parisi, A., Carter, B., 2013a. Smartphones as tools for delivering sun-smart education to
898 students. *Teaching Science* 59, 36–38.

899 Igoe, D., Parisi, A., Carter, B., 2013b. Characterization of a smartphone camera’s response to
900 ultraviolet A radiation. *Photochem Photobiol* 89, 215–218.
901 <https://doi.org/10.1111/j.1751-1097.2012.01216.x>

902 Isah, O.R., Usman, A.D., Tekanyi, A.M.S., 2017. A Hybrid Model of PSO Algorithm and
903 Artificial Neural Network for Automatic Follicle Classification 16.

904 Ji, X., Shang, X., Dahlgren, R.A., Zhang, M., 2017. Prediction of dissolved oxygen
905 concentration in hypoxic river systems using support vector machine: a case study of
906 Wen-Rui Tang River, China. *Environmental Science and Pollution Research* 24, 16062–
907 16076. <https://doi.org/10.1007/s11356-017-9243-7>

908 Jiménez-Pérez, P.F., Mora-López, L., 2016. Modeling and forecasting hourly global solar
909 radiation using clustering and classification techniques. *Solar Energy* 135, 682–691.
910 <https://doi.org/10.1016/j.solener.2016.06.039>

911 Jovanovic, B., Collins, D., Braganza, K., Jakob, D., Jones, D.A., 2011. A high-quality monthly
912 total cloud amount dataset for Australia. *Climatic Change* 108, 485–517.
913 <https://doi.org/10.1007/s10584-010-9992-5>

914 Juzeniene, A., Moan, J., 2012. Beneficial effects of UV radiation other than via vitamin D
915 production. *Dermato-Endocrinology* 4, 109–117. <https://doi.org/10.4161/derm.20013>

916 Kaba, K., Kandirmaz, H.M., Avci, M., 2017. Estimation of daily sunshine duration using support
917 vector machines. *International Journal of Green Energy* 14, 430–441.
918 <https://doi.org/10.1080/15435075.2016.1265971>

919 Karimkhani, C., Green, A.C., Nijsten, T., Weinstock, M.A., Dellavalle, R.P., Naghavi, M.,
920 Fitzmaurice, C., 2017. The global burden of melanoma: results from the Global Burden
921 of Disease Study 2015. *British Journal of Dermatology* 177, 134–140.
922 <https://doi.org/10.1111/bjd.15510>

923 Kazantzidis, A., Smedley, A., Kift, R., Rimmer, J., Berry, J.L., Rhodes, L.E., Webb, A.R., 2015.
924 A modeling approach to determine how much UV radiation is available across the UK
925 and Ireland for health risk and benefit studies. *Photochem. Photobiol. Sci.* 14, 1073–
926 1081. <https://doi.org/10.1039/C5PP00008D>

927 Kennedy, J., Eberhart, R., 1995. Particle swarm optimization, in: *Proceedings of ICNN’95 -*
928 *International Conference on Neural Networks*. Presented at the *Proceedings of ICNN’95 -*
929 *International Conference on Neural Networks*, pp. 1942–1948 vol.4.
930 <https://doi.org/10.1109/ICNN.1995.488968>

931 Ketkar, N., 2017. Introduction to Keras, in: Ketkar, N. (Ed.), *Deep Learning with Python: A*
932 *Hands-on Introduction*. Apress, Berkeley, CA, pp. 97–111. [https://doi.org/10.1007/978-1-](https://doi.org/10.1007/978-1-4842-2766-4_7)
933 [4842-2766-4_7](https://doi.org/10.1007/978-1-4842-2766-4_7)

934 Kiladis, G.N., Straub, K.H., Reid, G.C., Gage, K.S., 2001. Aspects of interannual and
935 intraseasonal variability of the tropopause and lower stratosphere. *Quarterly Journal of*
936 *the Royal Meteorological Society* 127, 1961–1983.
937 <https://doi.org/10.1002/qj.49712757606>

938 Kroft, E.B.M., Berkhof, N.J.G., Kerkhof, P.C.M. van de, Gerritsen, R.M.J.P., Jong, E.M.G.J. de,
939 2008. Ultraviolet A phototherapy for sclerotic skin diseases: A systematic review. *Journal*
940 *of the American Academy of Dermatology* 59, 1017–1030.
941 <https://doi.org/10.1016/j.jaad.2008.07.042>

942 Krzyścin, J.W., Guzikowski, J., Czerwińska, A., Lesiak, A., Narbutt, J., Jarosławski, J.,
943 Sobolewski, P.S., Rajewska-Więch, B., Wink, J., 2015. 24 hour forecast of the surface
944 UV for the antipsoriatic heliotherapy in Poland. *Journal of Photochemistry and*
945 *Photobiology B: Biology* 148, 136–144. <https://doi.org/10.1016/j.jphotobiol.2015.04.002>

946 Latosińska, J.N., Latosińska, M., Bielak, J., 2015. Towards modelling ultraviolet index in global
947 scale. Artificial neural networks approach. *Aerospace Science and Technology* 41, 189–
948 198. <https://doi.org/10.1016/j.ast.2014.12.013>

949 Lau, W.K.-M., Waliser, D.E., 2011. *Intraseasonal variability in the atmosphere-ocean climate*
950 *system*. Springer Science & Business Media.

951 Lecun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to
952 document recognition. *Proceedings of the IEEE* 86, 2278–2324.
953 <https://doi.org/10.1109/5.726791>

954 Lee, S.-W., Hwang, S.-J., Lee, S.-B., Hwang, H.-S., Sung, H.-C., 2009. Landscape ecological
955 approach to the relationships of land use patterns in watersheds to water quality
956 characteristics. *Landscape and Urban Planning* 92, 80–89.
957 <https://doi.org/10.1016/J.LANDURBPLAN.2009.02.008>

958 Legates, D.R., McCabe, G.J., 2013. A refined index of model performance: a rejoinder.
959 *International Journal of Climatology* 33, 1053–1056. <https://doi.org/10.1002/joc.3487>

960 Li, J., Jiang, Y., Xia, X., Hu, Y., 2018. Increase of surface solar irradiance across East China
961 related to changes in aerosol properties during the past decade. *Environ. Res. Lett.* 13,
962 034006. <https://doi.org/10.1088/1748-9326/aaa35a>

963 Li, Youxing., Chen, X., Yu, X., 2019. Processes | Free Full-Text | A Hybrid Energy Feature
964 Extraction Approach for Ship-Radiated Noise Based on CEEMDAN Combined with
965 Energy Difference and Energy Entropy [WWW Document]. URL
966 <https://www.mdpi.com/2227-9717/7/2/69> (accessed 6.19.21).

967 Liu, B., Wang, D., Fu, S., Cao, W., 2017. Estimation of Peak Flow Rates for Small Drainage
968 Areas. *Water Resources Management* 31, 1635–1647. [https://doi.org/10.1007/s11269-](https://doi.org/10.1007/s11269-017-1604-y)
969 [017-1604-y](https://doi.org/10.1007/s11269-017-1604-y)

970 Liu, H., Tian, H., Li, Y., 2015. Four wind speed multi-step forecasting models using extreme
971 learning machines and signal decomposing algorithms. *Energy Conversion and*
972 *Management* 100, 16–22. <https://doi.org/10.1016/j.enconman.2015.04.057>

973 Lucas, R.M., McMichael, A.J., Armstrong, B.K., Smith, W.T., 2008. Estimating the global
974 disease burden due to ultraviolet radiation exposure. *International Journal of*
975 *Epidemiology* 37, 654–667. <https://doi.org/10.1093/ije/dyn017>

- 976 Madden, R.A., Julian, P.R., 1994. Observations of the 40–50-Day Tropical Oscillation—A
977 Review. *Monthly Weather Review* 122, 814–837. [https://doi.org/10.1175/1520-0493\(1994\)122<0814:OOTDTP>2.0.CO;2](https://doi.org/10.1175/1520-0493(1994)122<0814:OOTDTP>2.0.CO;2)
- 978
979 Madden, R.A., Julian, P.R., 1971. Detection of a 40–50 Day Oscillation in the Zonal Wind in the
980 Tropical Pacific. *Journal of the Atmospheric Sciences* 28, 702–708.
981 [https://doi.org/10.1175/1520-0469\(1971\)028<0702:DOADOI>2.0.CO;2](https://doi.org/10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2)
- 982 Mäusezahl, D., Christen, A., Pacheco, G.D., Tellez, F.A., Iriarte, M., Zapata, M.E., Cevallos, M.,
983 Hattendorf, J., Cattaneo, M.D., Arnold, B., Smith, T.A., Colford, J.M., 2009. Solar
984 drinking water disinfection (SODIS) to reduce childhood diarrhoea in rural Bolivia: a
985 cluster-randomized, controlled trial. *PLoS Med* 6, e1000125.
986 <https://doi.org/10.1371/journal.pmed.1000125>
- 987 Mayer, A., Lobet, M., 2018. UV to near-infrared broadband pyramidal absorbers via a genetic
988 algorithm optimization approach, in: *Metamaterials XI*. Presented at the *Metamaterials*
989 *XI, International Society for Optics and Photonics*, p. 1067127.
990 <https://doi.org/10.1117/12.2303823>
- 991 McCarthy, W.H., 2004. The Australian experience in sun protection and screening for
992 melanoma. *J Surg Oncol* 86, 236–245. <https://doi.org/10.1002/jso.20086>
- 993 Mucherino, A., Fidanova, S., Ganzha, M., 2015. Ant Colony Optimization with environment
994 changes: an application to GPS surveying. Presented at the 2015 Federated Conference
995 on Computer Science and Information Systems, pp. 495–500.
996 <https://doi.org/10.15439/2015F33>
- 997 Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A
998 discussion of principles. *Journal of Hydrology* 10, 282–290.
999 [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- 1000 Norval, M., Cullen, A.P., de Gruijl, F.R., Longstreth, J., Takizawa, Y., Lucas, R.M., Noonan,
1001 F.P., van der Leun, J.C., 2007. The effects on human health from stratospheric ozone
1002 depletion and its interactions with climate change. *Photochem Photobiol Sci* 6, 232–251.
1003 <https://doi.org/10.1039/b700018a>
- 1004 Opländer Christian, Volkmar Christine M., Paunel-Görgülü Adnana, van Faassen Ernst E., Heiss
1005 Christian, Kelm Malte, Halmer Daniel, Mürtz Manfred, Pallua Norbert, Suschek
1006 Christoph V., 2009. Whole Body UVA Irradiation Lowers Systemic Blood Pressure by
1007 Release of Nitric Oxide From Intracutaneous Photolabile Nitric Oxide Derivates.
1008 *Circulation Research* 105, 1031–1040.
1009 <https://doi.org/10.1161/CIRCRESAHA.109.207019>
- 1010 Pak, U., Kim, C., Ryu, U., Sok, K., Pak, S., 2018. A hybrid model based on convolutional neural
1011 networks and long short-term memory for ozone concentration prediction. *Air Qual*
1012 *Atmos Health* 11, 883–895. <https://doi.org/10.1007/s11869-018-0585-1>
- 1013 Parisi, A.V., Downs, N., Turner, J., Amar, A., 2016. Online educative activities for solar
1014 ultraviolet radiation based on measurements of cloud amount and solar exposures.
1015 *Journal of Photochemistry and Photobiology B: Biology* 162, 434–440.
1016 <https://doi.org/10.1016/j.jphotobiol.2016.07.015>
- 1017 Pavlakis, K.G., Hatzianastassiou, N., Matsoukas, C., Fotiadi, A., Vardavas, I., 2008. ENSO
1018 surface shortwave radiation forcing over the tropical Pacific. *Atmospheric Chemistry and*
1019 *Physics* 8, 5565–5577. <https://doi.org/10.5194/acp-8-5565-2008>
- 1020 Pavlakis, K.G., Hatzidimitriou, D., Drakakis, E., Matsoukas, C., Fotiadi, A., Hatzianastassiou,
1021 N., Vardavas, I., 2007. ENSO surface longwave radiation forcing over the tropical

1022 Pacific. Atmospheric Chemistry and Physics 7, 2013–2026. <https://doi.org/10.5194/acp->
1023 7-2013-2007

1024 Peng, H., Ying, C., Tan, S., Hu, B., Sun, Z., 2018. An Improved Feature Selection Algorithm
1025 Based on Ant Colony Optimization. *IEEE Access* 6, 69203–69209.
1026 <https://doi.org/10.1109/ACCESS.2018.2879583>

1027 Pinker, R.T., Grodsky, S., Zhang, B., Chen, W., 2017. ENSO Impact on Radiative Fluxes as
1028 Observed from Space. *Journal of Geophysical Research - Oceans* 122.
1029 <https://doi.org/10.1002/2017JC012900>

1030 Pooi, C.K., Ng, HY, 2018. Review of low-cost point-of-use water treatment systems for
1031 developing communities. *npj Clean Water* 1, 1–8. <https://doi.org/10.1038/s41545-018->
1032 0011-0

1033 Prasad, R., Ali, M., Kwan, P., Khan, H., 2019. Designing a multi-stage multivariate empirical
1034 mode decomposition coupled with ant colony optimization and random forest model to
1035 forecast monthly solar radiation. *Applied Energy* 236, 778–792.
1036 <https://doi.org/10.1016/j.apenergy.2018.12.034>

1037 Pruthi, D., Bhardwaj, R., 2021. Modeling air quality index using optimized neuronal networks
1038 inspired by swarms. *Environmental Engineering Research* 26.
1039 <https://doi.org/10.4491/eer.2020.469>

1040 Qing, X., Niu, Y., 2018. Hourly day-ahead solar irradiance prediction using weather forecasts by
1041 LSTM. *Energy* 148, 461–468. <https://doi.org/10.1016/j.energy.2018.01.177>

1042 Raksasat, R., Sri-iesaranusorn, P., Pemcharoen, J., Laiwarin, P., Buntoung, S., Janjai, S.,
1043 Boontaveeyuwat, E., Asawanonda, P., Sriswasdi, S., Chuangsuwanich, E., 2021.
1044 Accurate surface ultraviolet radiation forecasting for clinical applications with deep
1045 neural network. *Sci Rep* 11, 5031. <https://doi.org/10.1038/s41598-021-84396-2>

1046 Román, R., Antón, M., Valenzuela, A., Gil, J.E., Lyamani, H., Miguel, A.D., Olmo, F.J., Bilbao,
1047 J., Alados-Arboledas, L., 2013. Evaluation of the desert dust effects on global, direct and
1048 diffuse spectral ultraviolet irradiance. *Tellus B: Chemical and Physical Meteorology* 65,
1049 19578. <https://doi.org/10.3402/tellusb.v65i0.19578>

1050 Roshan, D.R., Koc, M., Abdallah, A., Martin-Pomares, L., Isaifan, R., Fountoukis, C., 2020. UV
1051 Index Forecasting under the Influence of Desert Dust: Evaluation against Surface and
1052 Satellite-Retrieved Data. *Atmosphere* 11, 96. <https://doi.org/10.3390/atmos11010096>

1053 Saraiya, M., Glanz, K., Briss, P.A., Nichols, P., White, C., Das, D., Smith, S.J., Tannor, B.,
1054 Hutchinson, A.B., Wilson, K.M., Gandhi, N., Lee, N.C., Rimer, B., Coates, R.C., Kerner,
1055 J.F., Hiatt, R.A., Buffler, P., Rochester, P., 2004. Interventions to prevent skin cancer by
1056 reducing exposure to ultraviolet radiation: a systematic review. *Am J Prev Med* 27, 422–
1057 466. <https://doi.org/10.1016/j.amepre.2004.08.009>

1058 Seme, S., Štumberger, G., 2011. A novel prediction algorithm for solar angles using solar
1059 radiation and Differential Evolution for dual-axis sun tracking purposes. *Solar Energy* 85,
1060 2757–2770. <https://doi.org/10.1016/j.solener.2011.08.031>

1061 Silva, C.A., Sousa, J.M.C., Runkler, T.A., Sá da Costa, J.M.G., 2009. Distributed supply chain
1062 management using ant colony optimization. *European Journal of Operational Research*
1063 199, 349–358.

1064 Sivamani, R.K., Crane, L.A., Dellavalle, R.P., 2009. The benefits and risks of ultraviolet (UV)
1065 tanning and its alternatives: the role of prudent sun exposure. *Dermatol Clin* 27, 149–vi.
1066 <https://doi.org/10.1016/j.det.2008.11.008>

1067 Slevin, T., Clarkson, J., English, D., 2000. Skin Cancer Control Western Australia: Is it Working
1068 and What Have we Learned? *Radiat. Prot. Dosim.* 91.
1069 <https://doi.org/10.1093/oxfordjournals.rpd.a033225>

1070 Srivastava, R., Tiwari, A.N., Giri, V.K., 2019. Solar radiation forecasting using MARS, CART,
1071 M5, and random forest model: A case study for India. *Heliyon* 5, e02692.
1072 <https://doi.org/10.1016/j.heliyon.2019.e02692>

1073 Staiger, H., den Outer, P.N., Bais, A.F., Feister, U., Johnsen, B., Vuilleumier, L., 2008. Hourly
1074 resolved cloud modification factors in the ultraviolet. *Atmospheric Chemistry and*
1075 *Physics* 8, 2493–2508. <https://doi.org/10.5194/acp-8-2493-2008>

1076 Stanton, W.R., Janda, M., Baade, P.D., Anderson, P., 2004. Primary prevention of skin cancer: a
1077 review of sun protection in Australia and internationally. *Health Promot Int* 19, 369–378.
1078 <https://doi.org/10.1093/heapro/dah310>

1079 Staples, M., Marks, R., Giles, G., 1998. Trends in the incidence of non-melanocytic skin cancer
1080 (NMSC) treated in Australia 1985-1995: are primary prevention programs starting to
1081 have an effect? *Int J Cancer* 78, 144–148. [https://doi.org/10.1002/\(sici\)1097-0215\(19981005\)78:2<144::aid-ijc3>3.0.co;2-z](https://doi.org/10.1002/(sici)1097-0215(19981005)78:2<144::aid-ijc3>3.0.co;2-z)

1082
1083 Staples, M.P., Elwood, M., Burton, R.C., Williams, J.L., Marks, R., Giles, G.G., 2006. Non-
1084 melanoma skin cancer in Australia: the 2002 national survey and trends since 1985. *Med*
1085 *J Aust* 184, 6–10. <https://doi.org/10.5694/j.1326-5377.2006.tb00086.x>

1086 Sudhibrabha, S., Harold Buchanan Exell, R., Sukawat, D., 2006. Ultraviolet forecasting in
1087 Thailand. *ScienceAsia* 32, 107. <https://doi.org/10.2306/scienceasia1513-1874.2006.32.107>

1088
1089 Szenicer, A., Fouhey, D.F., Munoz-Jaramillo, A., Wright, P.J., Thomas, R., Galvez, R., Jin, M.,
1090 Cheung, M.C.M., 2019. A deep learning virtual instrument for monitoring extreme UV
1091 solar spectral irradiance. *Science Advances* 5, eaaw6548.
1092 <https://doi.org/10.1126/sciadv.aaw6548>

1093 Tartaglione, N., Toniazzo, T., Orsolini, Y., Otterå, O.H., 2020. Impact of solar irradiance and
1094 geomagnetic activity on polar NO_x, ozone and temperature in WACCM simulations.
1095 *Journal of Atmospheric and Solar-Terrestrial Physics* 209, 105398.
1096 <https://doi.org/10.1016/j.jastp.2020.105398>

1097 Tian, B., Waliser, D.E., Kahn, R.A., Li, Q., Yung, Y.L., Tyranowski, T., Geogdzhayev, I.V.,
1098 Mishchenko, M.I., Torres, O., Smirnov, A., 2008. Does the Madden-Julian Oscillation
1099 influence aerosol variability? *Journal of Geophysical Research: Atmospheres* 113.
1100 <https://doi.org/10.1029/2007JD009372>

1101 Timmermann, L.F., Ritter, K., Hillebrandt, D., Küpper, T., 2015. Drinking water treatment with
1102 ultraviolet light for travelers – Evaluation of a mobile lightweight system. *Travel*
1103 *Medicine and Infectious Disease* 13, 466–474.
1104 <https://doi.org/10.1016/j.tmaid.2015.10.005>

1105 Tiwari, M.K., Adamowski, J., 2013. Urban water demand forecasting and uncertainty assessment
1106 using ensemble wavelet-bootstrap-neural network models. *Water Resources Research* 49,
1107 6486–6507. <https://doi.org/10.1002/wrcr.20517>

1108 Tiwari, M.K., Chatterjee, C., 2010. A new wavelet-bootstrap-ANN hybrid model for daily
1109 discharge forecasting. *Journal of Hydroinformatics* 13, 500–519.
1110 <https://doi.org/10.2166/hydro.2010.142>

1111 Turner, E.C., Manners, J., Morcrette, C.J., O'Hagan, J.B., Smedley, A.R.D., 2017. Toward a
1112 New UV Index Diagnostic in the Met Office's Forecast Model. *Journal of Advances in*
1113 *Modeling Earth Systems* 9, 2654–2671. <https://doi.org/10.1002/2017MS001050>
1114 Ventor, G., Sobieszczanski-Sobieski, J., 2003. Particle Swarm Optimization | *AIAA Journal*
1115 [WWW Document]. URL <https://arc.aiaa.org/doi/10.2514/2.2111> (accessed 6.19.21).
1116 Wang, F., Yu, Y., Zhang, Z., Li, J., Zhen, Z., Li, K., 2018. Wavelet Decomposition and
1117 Convolutional LSTM Networks Based Improved Deep Learning Model for Solar
1118 Irradiance Forecasting. *Applied Sciences* 8, 1286. <https://doi.org/10.3390/app8081286>
1119 Waskom, M.L., 2021. seaborn: statistical data visualization. *Journal of Open Source Software* 6,
1120 3021. <https://doi.org/10.21105/joss.03021>
1121 Weile, D., Michielssen, E., 1997. Genetic algorithm optimization applied to electromagnetics: a
1122 review. <https://doi.org/10.1109/8.558650>
1123 Welch, D., Buonanno, M., Grilj, V., Shuryak, I., Crickmore, C., Bigelow, A.W., Randers-
1124 Pehrson, G., Johnson, G.W., Brenner, D.J., 2018. Far-UVC light: A new tool to control
1125 the spread of airborne-mediated microbial diseases. *Scientific Reports* 8, 2752.
1126 <https://doi.org/10.1038/s41598-018-21058-w>
1127 Wells, W.F., Fair, GM, 1935. Viability of B. Coli Exposed to Ultra-Violet Radiation in Air.
1128 *Science* 82, 280–281. <https://doi.org/10.1126/science.82.2125.280-a>
1129 Wheeler, M.C., Hendon, H.H., 2004. An All-Season Real-Time Multivariate MJO Index:
1130 Development of an Index for Monitoring and Prediction. *Monthly Weather Review* 132,
1131 1917–1932.
1132 WHO, 2002. Global solar UV index: a practical guide : a joint recommendation of World Health
1133 Organization, World Meteorological Organization, United Nations Environment
1134 Programme, International Commission on Non-Ionizing Radiation Protection. Geneva,
1135 WHO.
1136 Willmott, C.J., Robeson, S.M., Matsuura, K., 2012. A refined index of model performance.
1137 *International Journal of Climatology* 32, 2088–2094. <https://doi.org/10.1002/joc.2419>
1138 Yadav, A.K., Chandel, SS, 2014. Solar radiation prediction using Artificial Neural Network
1139 techniques: A review. *Renewable and Sustainable Energy Reviews* 33, 772–781.
1140 <https://doi.org/10.1016/j.rser.2013.08.055>
1141 Yan, H., Sun, L., Wang, Y., Huang, W., Qiu, S., Yang, C., 2011. A record of the Southern
1142 Oscillation Index for the past 2,000 years from precipitation proxies. *Nature Geosci* 4,
1143 611–614. <https://doi.org/10.1038/ngeo1231>
1144 Zhang, J., Zhang, X., Niu, J., Hu, B.X., Soltanian, M.R., Qiu, H., Yang, L., 2019. Prediction of
1145 groundwater level in seashore reclaimed land using wavelet and artificial neural network-
1146 based hybrid model. *Journal of Hydrology* 577, 123948.
1147 <https://doi.org/10.1016/j.jhydrol.2019.123948>
1148

Table 1. Description of global pool of 24 predictor variables used to design and evaluate hybrid CEEMDAN-CNN-LSTM predictive model for the daily maximum UV Index forecasting.

MODIS-Satellite		
OTC	Ozone Total Column	DU
GH	Geopotential Height (Daytime)	-
AO	Aerosol Optical Depth 550 nm	-
AOD2	Aerosol Optical Depth 342.5 nm	-
TCW	Total Column Water Vapour (Daytime)	Kg/m ²
CF	Cloud Fraction (Daytime)	-
CP	Cloud Pressure (Daytime)	hPa
CCO	Combined Cloud Optical Thickness (Mean)	-
SILO (Ground- Based Observations)		
T.Max	Maximum Temperature	°C
T.Min	Minimum Temperature	°C
Rain	Rainfall	mm
Evap	Evaporation	mm
Radn	Radiation	MJ m ⁻²
VP	Vapour Pressure	hPa
RHmaxT	Relative Humidity at Temperature T.Max	%
RHminT	Relative Humidity at Temperature T.Min	%
FAO56	Morton potential evapotranspiration overland	mm
SYNOPTIC-SCALE (Climate Mode Indices)		
Nino3.0	Average SSTA over 150°–90 °W and 5 °N–5 °S	NONE
Nino3.4	Average SSTA over 170 °E–120 °W and 5 °N–5 °S	
Nino4.0	Average SSTA over 160 °E–150 °W and 5 °N–5 °S	
Nino1+2	Average SSTA over 90°W–80°W and 0°–10°S	
AON	Arctic Oscillation	
AAO	Antarctic Oscillation	
EPO	East Pacific Oscillation	
GBI	Greenland Blocking Index (GBI)	
WPO	Western Pacific Oscillation (WPO.)	
PNA	Pacific North American Index	
NAO	North Atlantic Oscillation	
SAM	Southern Annular Mode index	
SOI	Southern Oscillation Index, as per Troup (1965)	
RMM1	Real-Time Multivariate MJO indices 1	
RMM2	Real-Time Multivariate MJO indices 1	

Table 2. List of selected input variables prior applying in the proposed model using four optimization techniques (i.e., ACO, DEV, GA and PSO).

	ACO	DEV	GA	PSO
OTC	✓	✗	✓	✗
GH	✗	✓	✗	✓
AOD	✓	✓	✓	✗
AOD2	✗	✓	✗	✗
TCW	✓	✗	✓	✗
CF	✗	✗	✓	✓
CP	✓	✗	✗	✗
CCO	✗	✓	✗	✓
T.Max	✓	✗	✗	✗
T.Min	✗	✗	✗	✓
Rain	✗	✓	✗	✓
Evap	✗	✗	✓	✗
Radn	✗	✗	✗	✓
VP	✗	✗	✗	✗
RHmaxT	✗	✗	✗	✗
RHminT	✗	✗	✓	✗
FAO56	✗	✗	✗	✗
SOI	✗	✗	✓	✗
EPO	✓	✗	✗	✗
GBI	✓	✗	✓	✗
AAO	✗	✗	✓	✓
AO	✓	✗	✓	✗
NAO	✗	✓	✓	✗
PNA	✓	✗	✗	✓
Nino3	✓	✗	✗	✗
NINO4	✗	✓	✓	✗
NINO12	✗	✓	✓	✗
NINO34	✗	✓	✗	✗
RMM1	✓	✓	✓	✓
RMM2	✓	✓	✓	✓

Table 3. Evaluation of hybrid CEEMDAN-CLSTM vs. benchmark (CNN-GRU, CNN-LSTM, CEEMDAN-GRU, CEEMDAN-LSTM, GRU and LSTM) models for Perth observation sites. The correlation coefficient (r), root mean square error ($RMSE$), mean absolute error (MAE) and Nash-Sutcliffe coefficient (NS) are computed between forecasted and observed UVI for 7 Day ahead periods in testing phase. The optimal model is **boldfaced** (blue).

	ACO				DEV				GA				PSO			
	r	NS	$RMSE$	MAE												
CEEMDAN-CGRU	0.995	0.991	0.317	0.261	0.994	0.981	0.460	0.392	0.994	0.994	0.260	0.198	0.995	0.990	0.343	0.281
CEEMDAN-CLSTM	0.996	0.992	0.216	0.163	0.995	0.986	0.258	0.148	0.996	0.997	0.162	0.119	0.996	0.993	0.220	0.165
CEEMDAN-GRU	0.973	0.945	0.796	0.598	0.993	0.987	0.387	0.300	0.974	0.948	0.768	0.579	0.993	0.985	0.413	0.312
CEEMDAN-LSTM	0.982	0.963	0.648	0.483	0.983	0.967	0.615	0.447	0.992	0.981	0.463	0.354	0.978	0.949	0.766	0.563
CEEMDAN-DT	0.960	0.920	0.952	0.672	0.986	0.971	0.575	0.420	0.959	0.918	0.968	0.685	0.984	0.968	0.599	0.425
CEEMDAN-MLP	0.945	0.890	1.117	0.846	0.969	0.938	0.839	0.576	0.957	0.914	0.987	0.705	0.964	0.927	0.911	0.688
CEEMDAN-SVR	0.988	0.974	0.541	0.384	0.993	0.987	0.387	0.300	0.990	0.978	0.504	0.370	0.991	0.991	0.385	0.223
CNN-GRU	0.987	0.987	0.317	0.361	0.988	0.974	0.541	0.384	0.988	0.968	0.601	0.453	0.981	0.962	0.663	0.457
CNN-LSTM	0.986	0.982	0.307	0.334	0.977	0.950	0.758	0.548	0.986	0.979	0.556	0.292	0.991	0.972	0.567	0.458
GRU	0.973	0.945	0.796	0.598	0.959	0.913	0.986	0.767	0.968	0.933	0.875	0.644	0.975	0.948	0.771	0.559
LSTM	0.977	0.950	0.758	0.548	0.980	0.954	0.721	0.508	0.975	0.945	0.793	0.537	0.981	0.958	0.690	0.485
DT	0.884	0.766	1.631	1.112	0.878	0.753	1.675	1.161	0.877	0.752	1.678	1.148	0.950	0.897	1.082	0.752
MLP	0.935	0.872	1.207	0.905	0.933	0.838	1.359	1.072	0.943	0.889	1.124	0.814	0.957	0.907	1.030	0.773
SVR	0.957	0.914	0.990	0.655	0.939	0.878	1.178	0.809	0.948	0.896	1.090	0.733	0.970	0.940	0.826	0.557

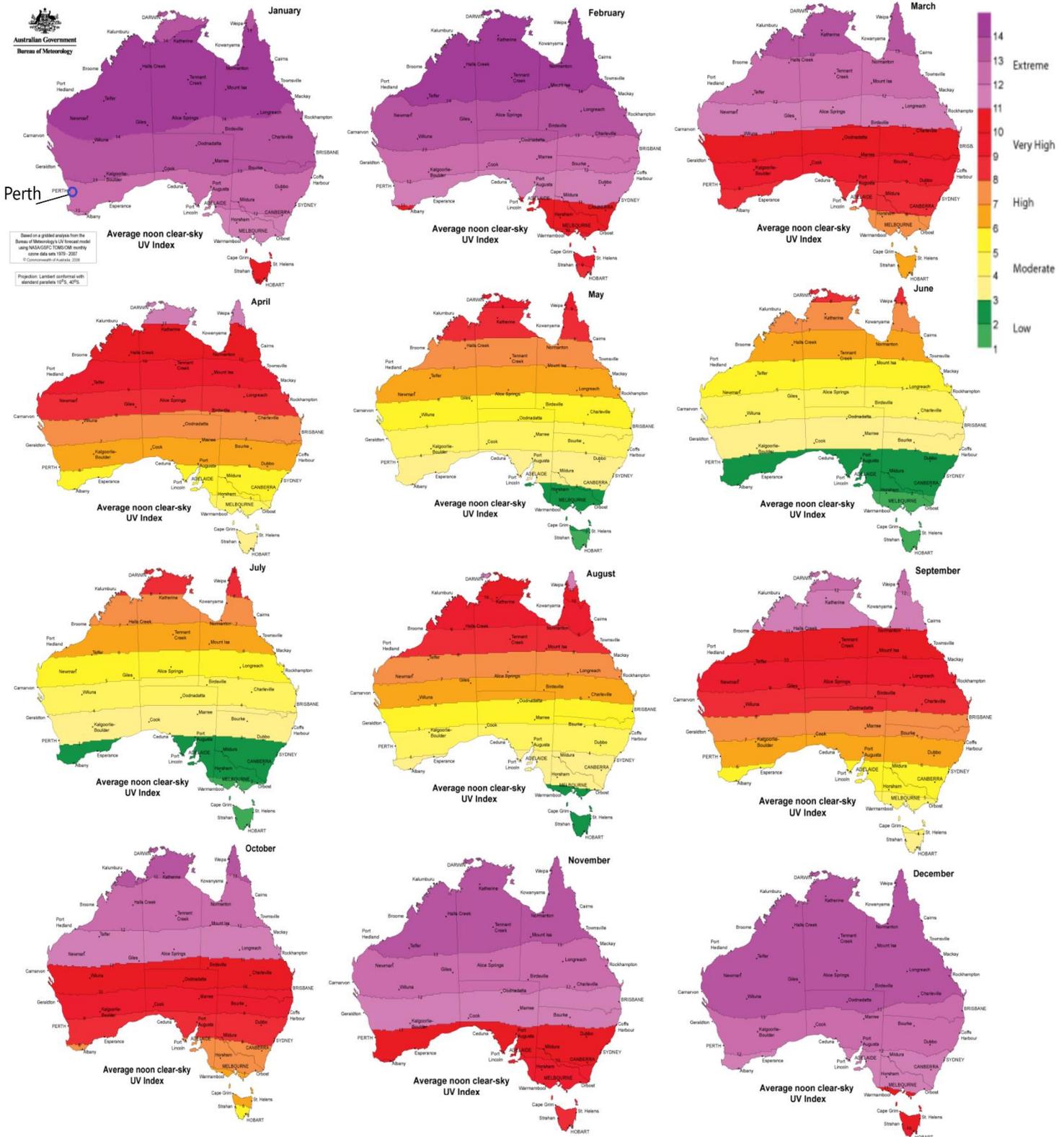


Figure 1 Study site (Perth, Australia) of the work and monthly average noon clear-sky UV index based on gridded analysis from the Bureau of Meteorology's UV forecast model using NASA/GFSC TOMS OMI monthly ozone data sets between 1979 - 2007

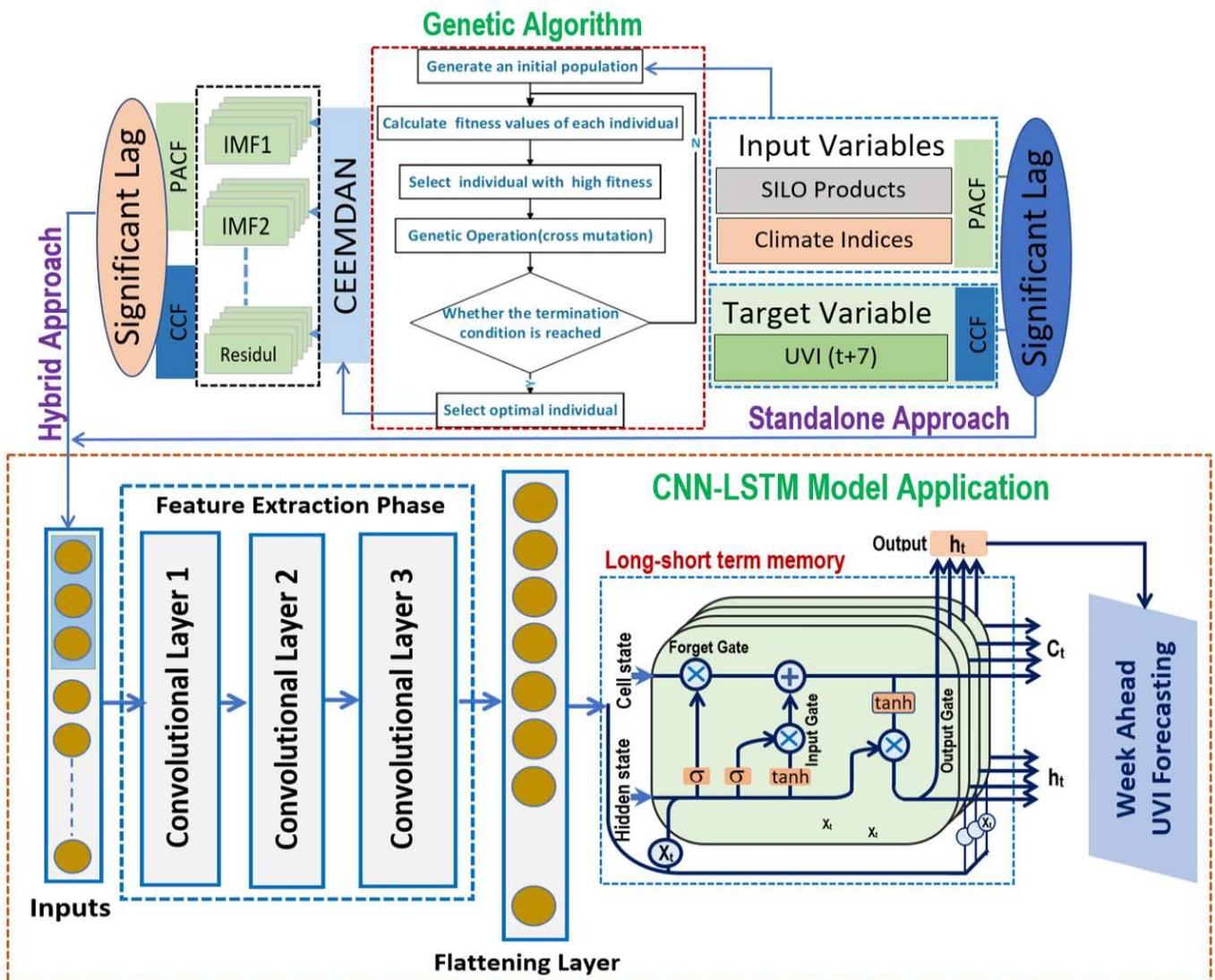


Figure 2 The developed model architecture of (Convolutional Neural Network, CNN) with the 4 layered long short term memory for a hybrid CNN-LSTM (CLSTM) model to forecast a week daily maximum UV Index with Genetic Algorithm.

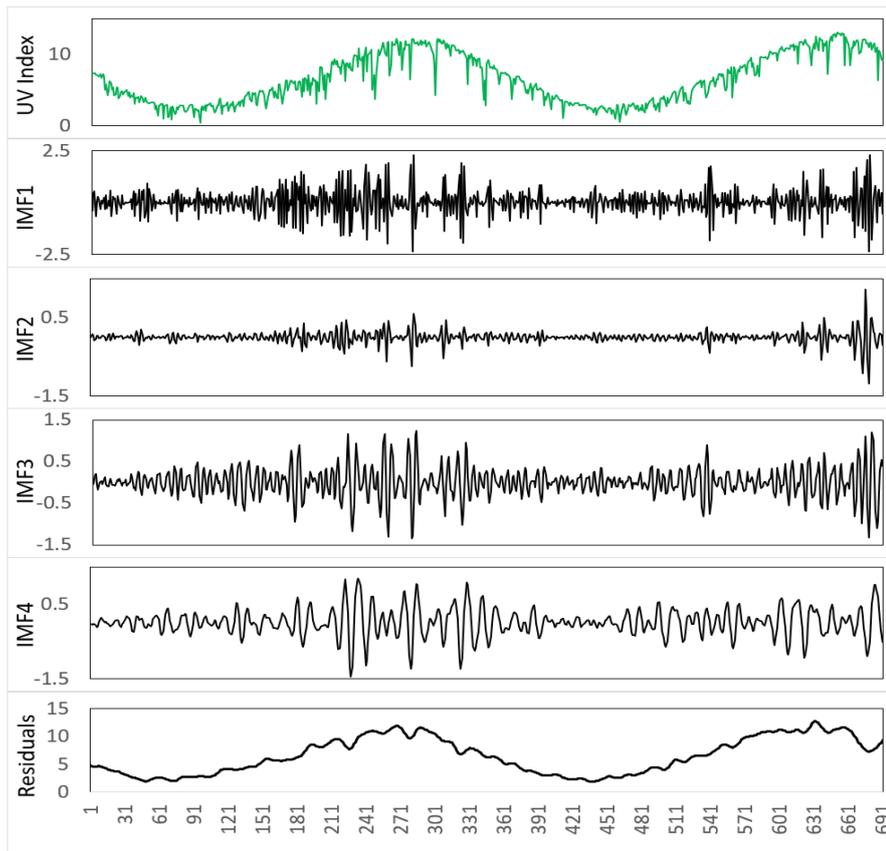


Figure 3 An example time-series showing data features in IMFs and residuals produced by the CEEMDAN transformation of daily maximum UV Index for the case of Perth study site.

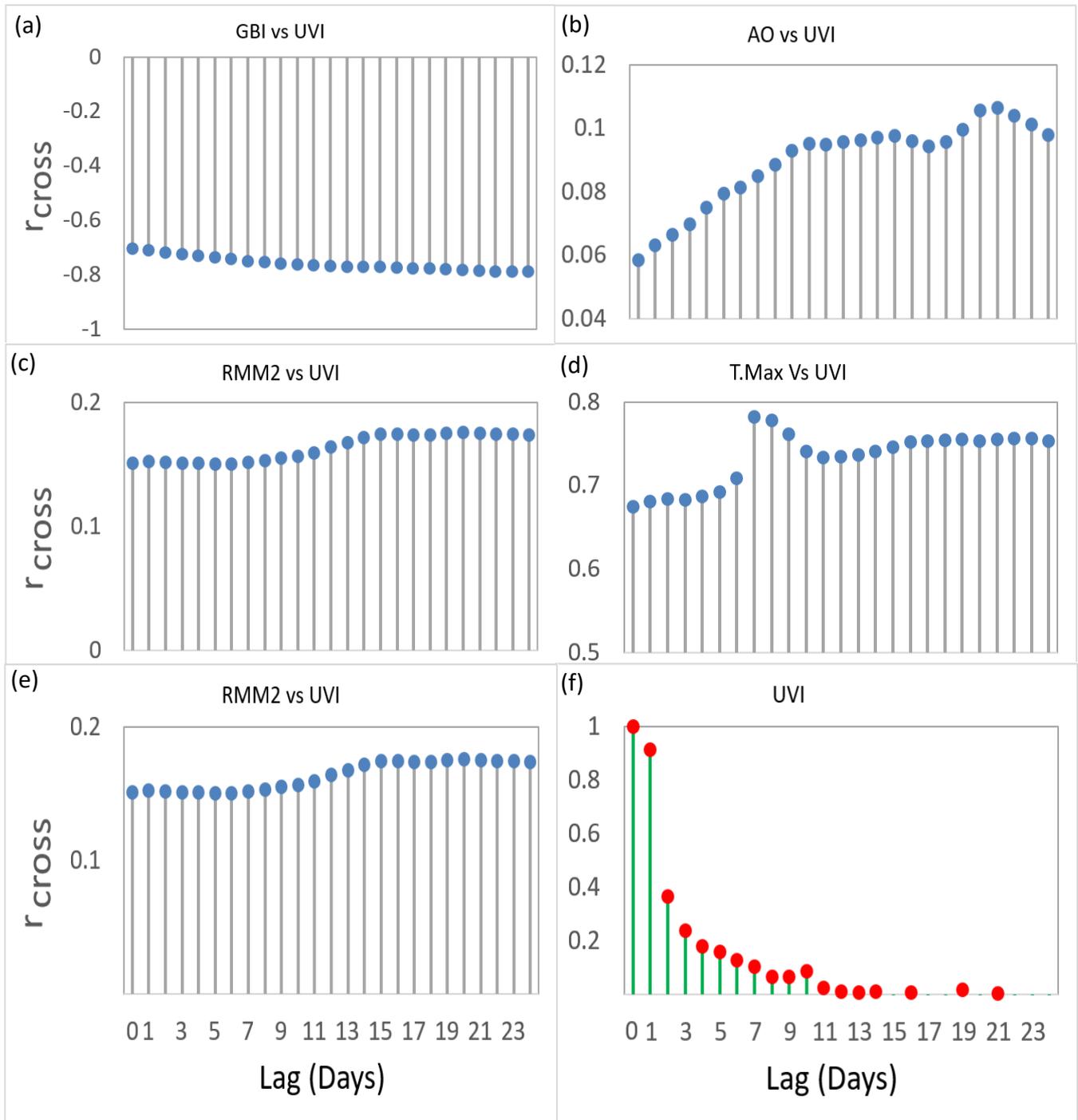


Figure 4 (a-e) Correlogram showing the covariance between the objective variable (UVI) and the predictor variables in terms of the Cross-correlation coefficient (r_{cross}) and (f) Partial autocorrelation function (PACF) plot of the UVI time series exploring the antecedent behavior in terms of the lag of UVI every day.

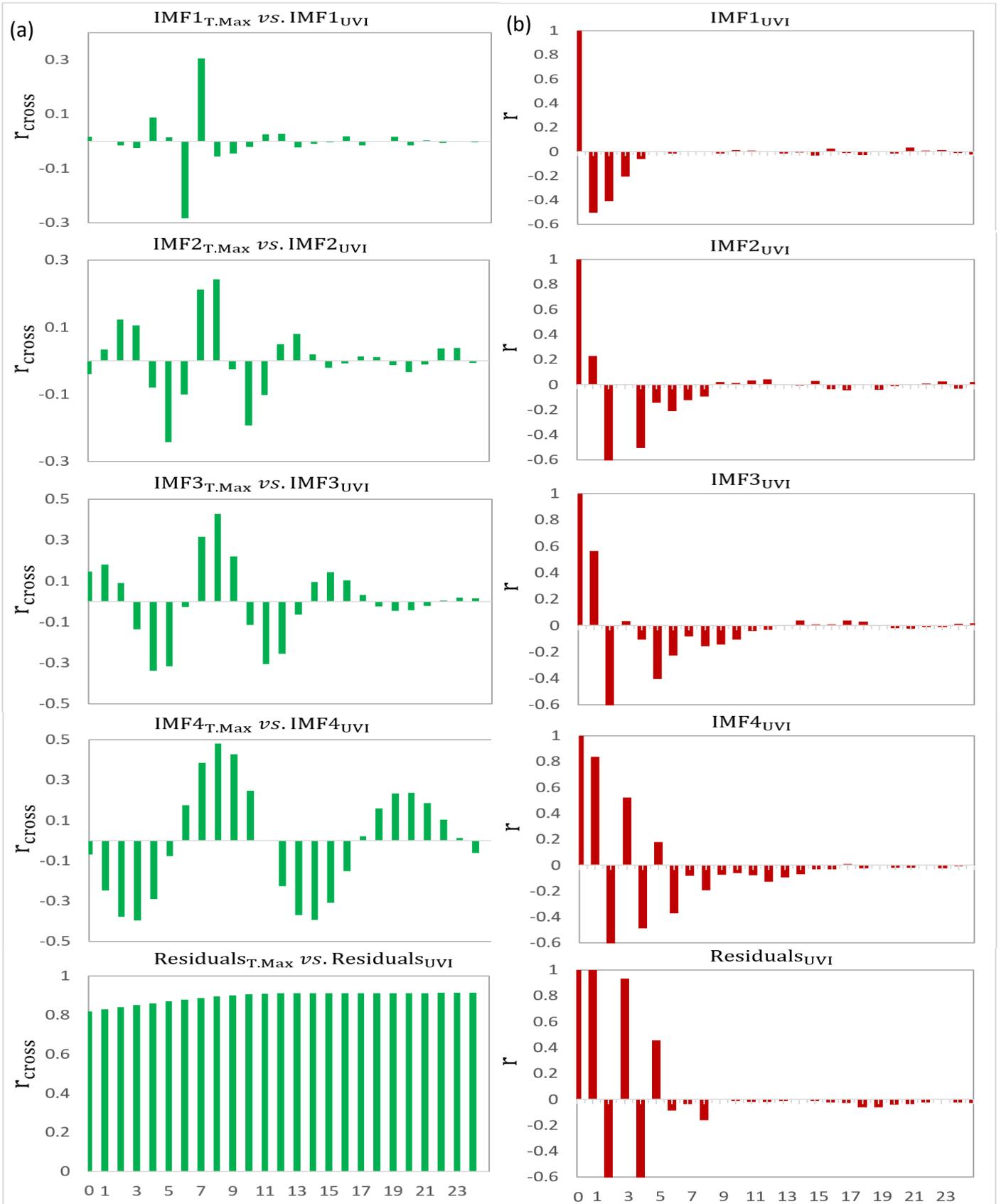


Figure 5 (a) Correlogram showing the covariance between the objective variable (UVI) and the CEEMDAN decomposed T.Max (IMF1_{T.Max} to Residuals_{T.Max}) in terms of the Cross-correlation coefficient (r_{cross}) and (b) Partial autocorrelation function (PACF) plot of the CEEMDAN decomposed UVI time series exploring the antecedent behavior in terms of the lag of UVI every day.

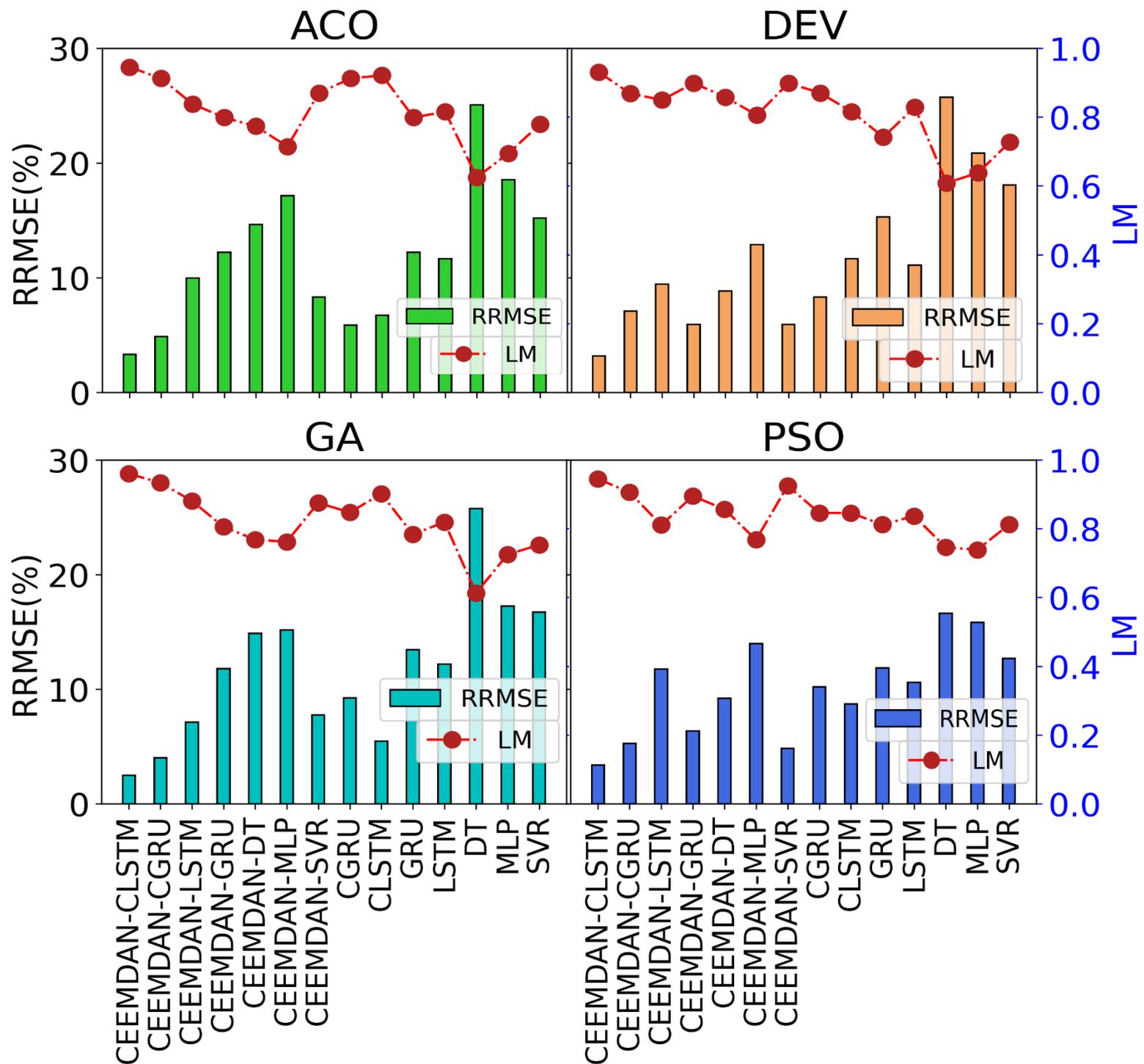


Figure 6 Comparison of the forecasting skill for all of the proposed models in terms of the relative error: RRMSE (%) and Legate McCabe Index (LM) within the testing period.

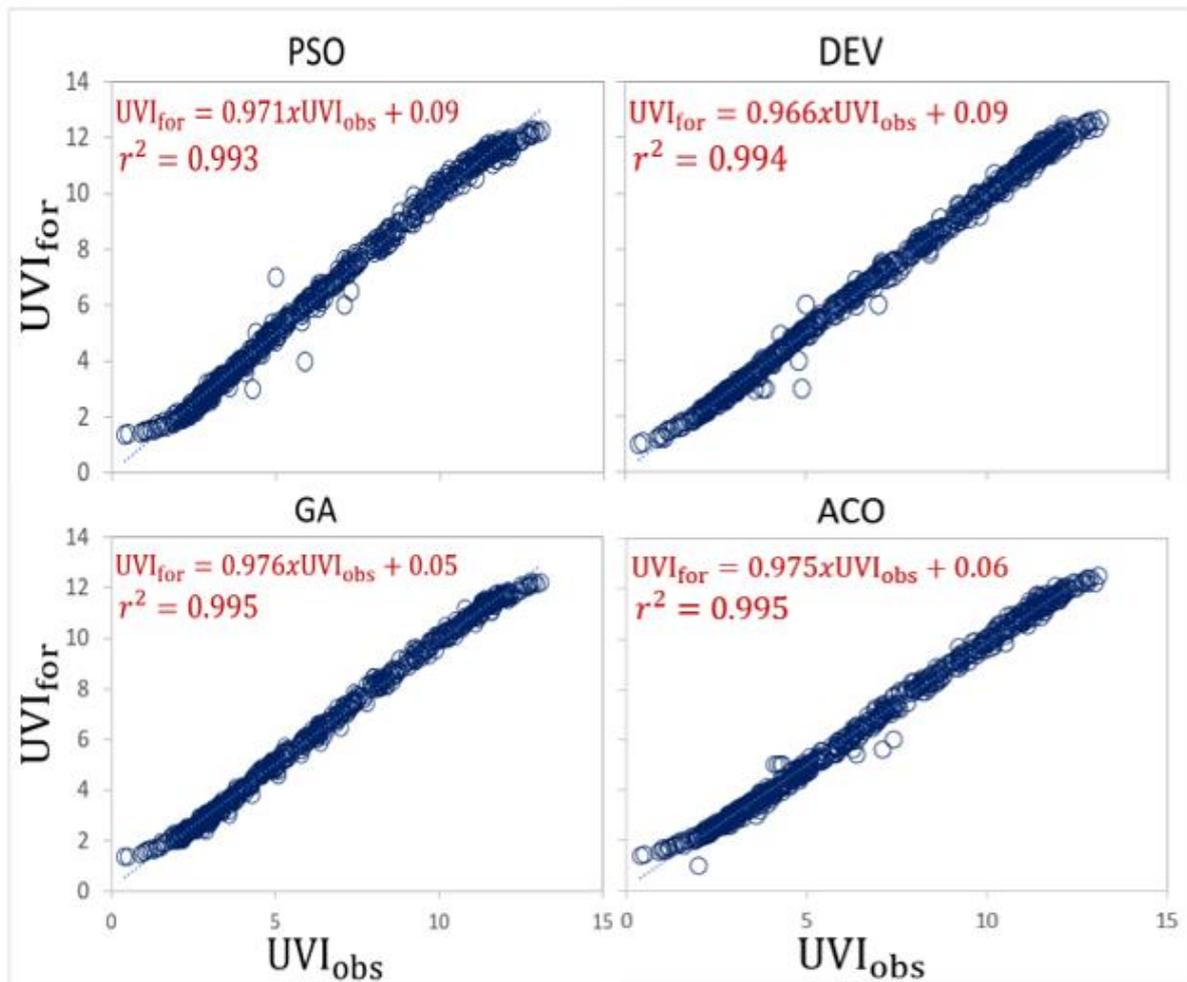


Figure 7 Scatter plot of forecasted with observed UVI (UVI) of Perth station CEEMDAN-CNN-GRU (CEEMDAN-CLSTM) model. A least square regression line and coefficient of determination (R^2) with a linear fit equation are shown in each sub-panel.

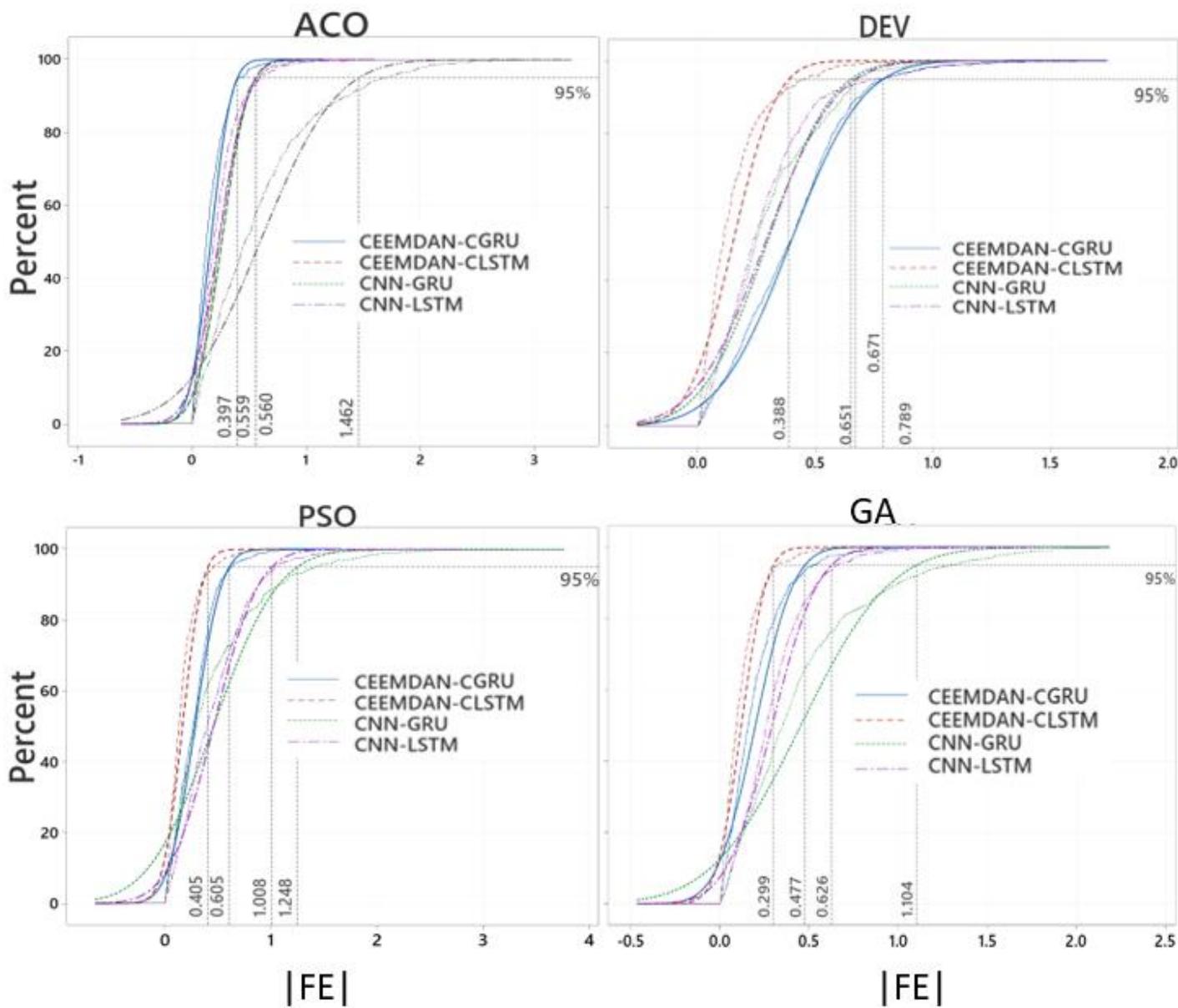


Figure 8 Empirical cumulative distribution function (CDF) in terms of forecasting error $|FE|$ for CEEMDAN-CNN-GRU, CEEMDAN-CNN-LSTM, CNN-GRU, and CNN-LSTM model, shown for the 95 percentile on ECDF.

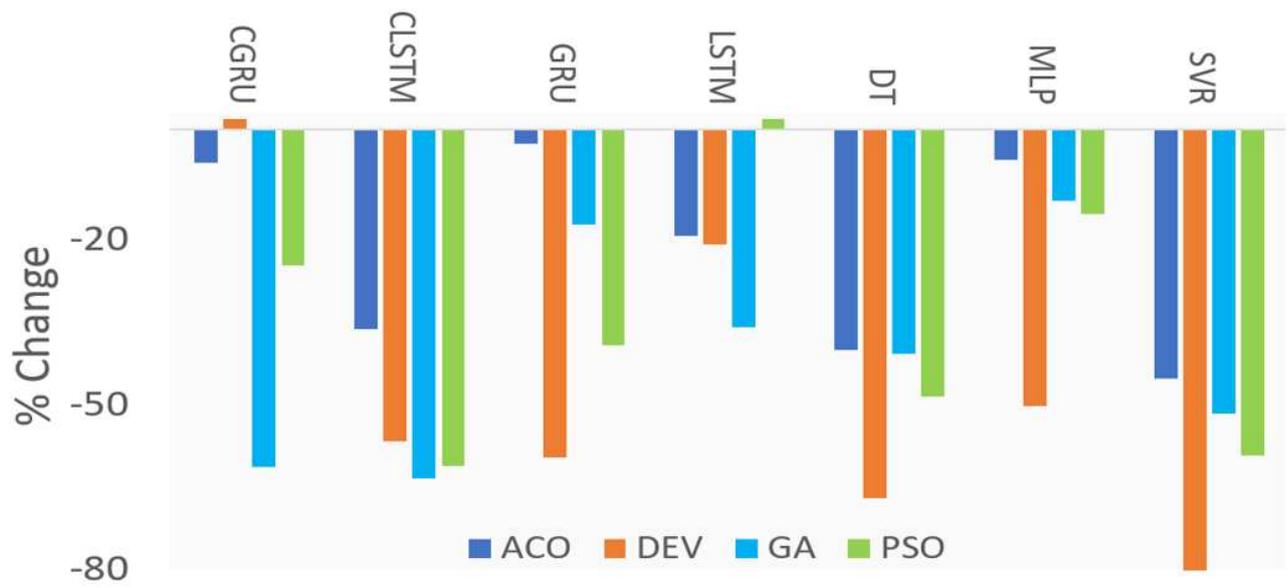


Figure 9 Effect on per cent change (%) of RMAE using CEEMDAN as a feature extraction approach in forecasting UVI at Perth station using Genetic Algorithm (GA), Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), and Differential Evolution (DEV)

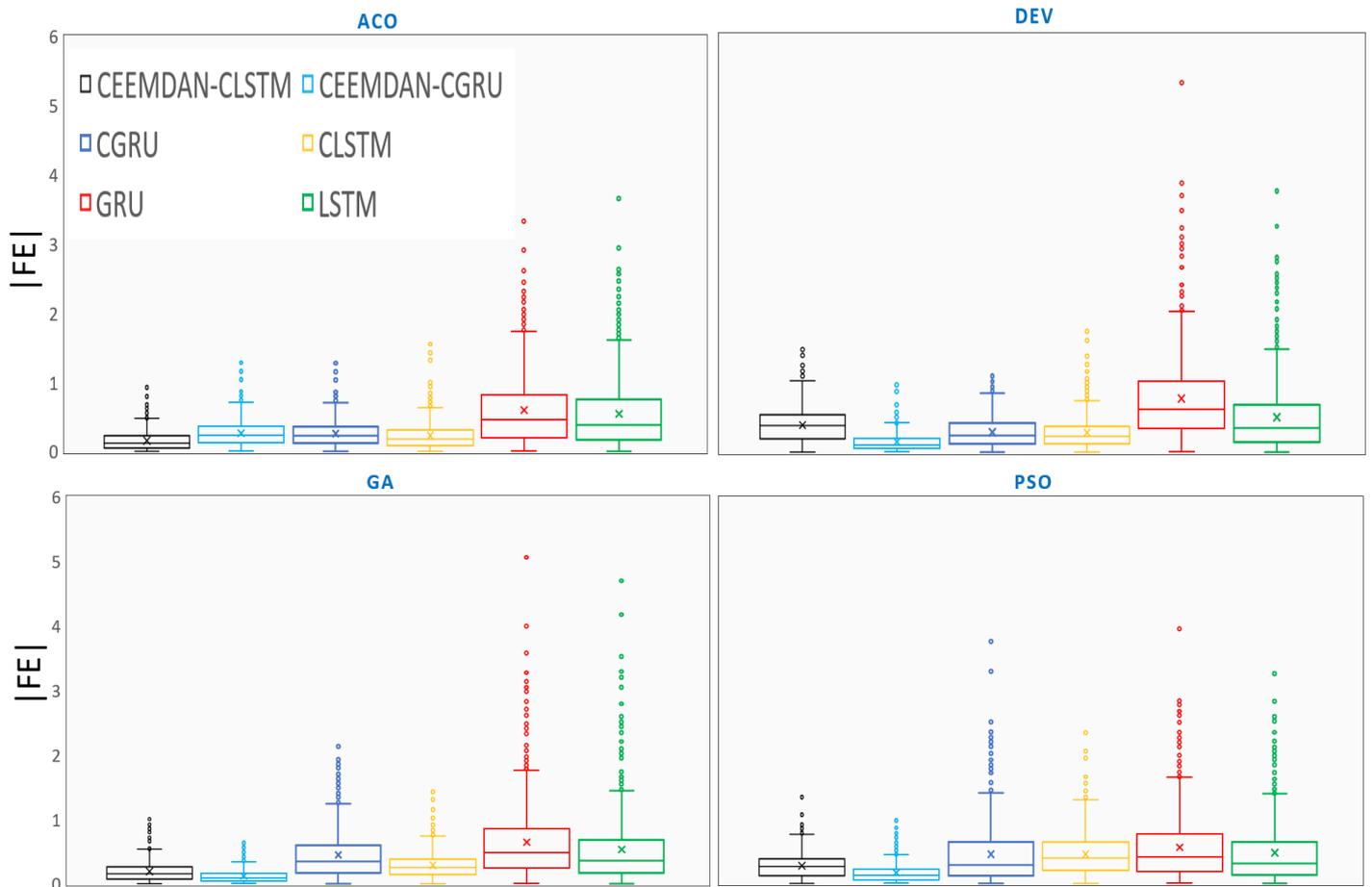


Figure 10 Evaluation of the performance of the proposed hybrid deep learning, CEEMDAN-CLSTM model with the comparative benchmark models based on the absolute forecasted error $|FE|$ using four optimization techniques.

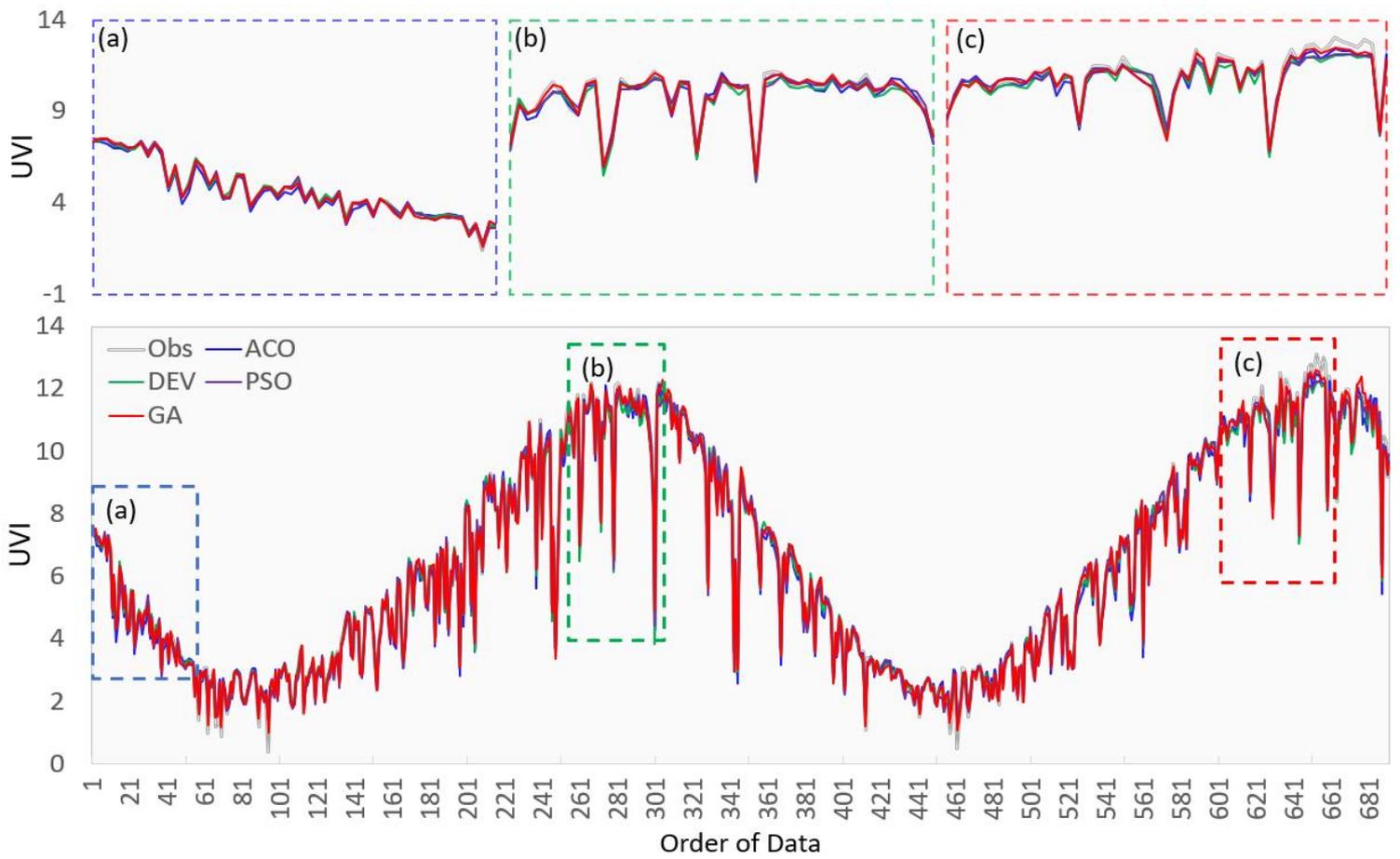


Figure 11 Time series of daily maximum UV index (UVI) for observed UVI and forecasted UVI for the objective model, CEEMDAN-CLSTM using four optimization approaches.

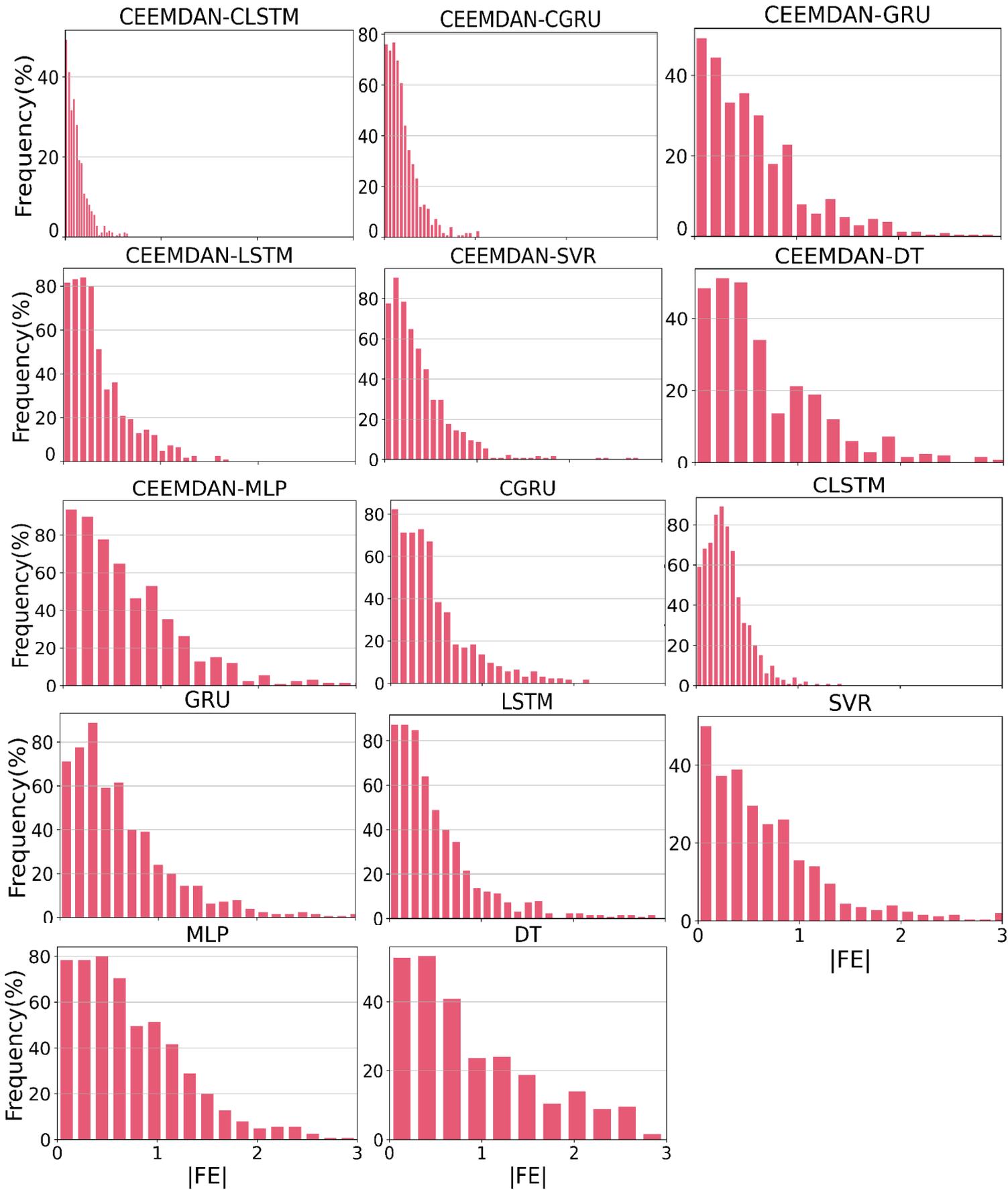


Figure 12 Illustration of the frequency of absolute value of estimation errors ($|EE|$) of the proposed hybrid deep learning CEEMDAN-CNN-LSTM model and comparing models using Genetic Algorithm (GA).

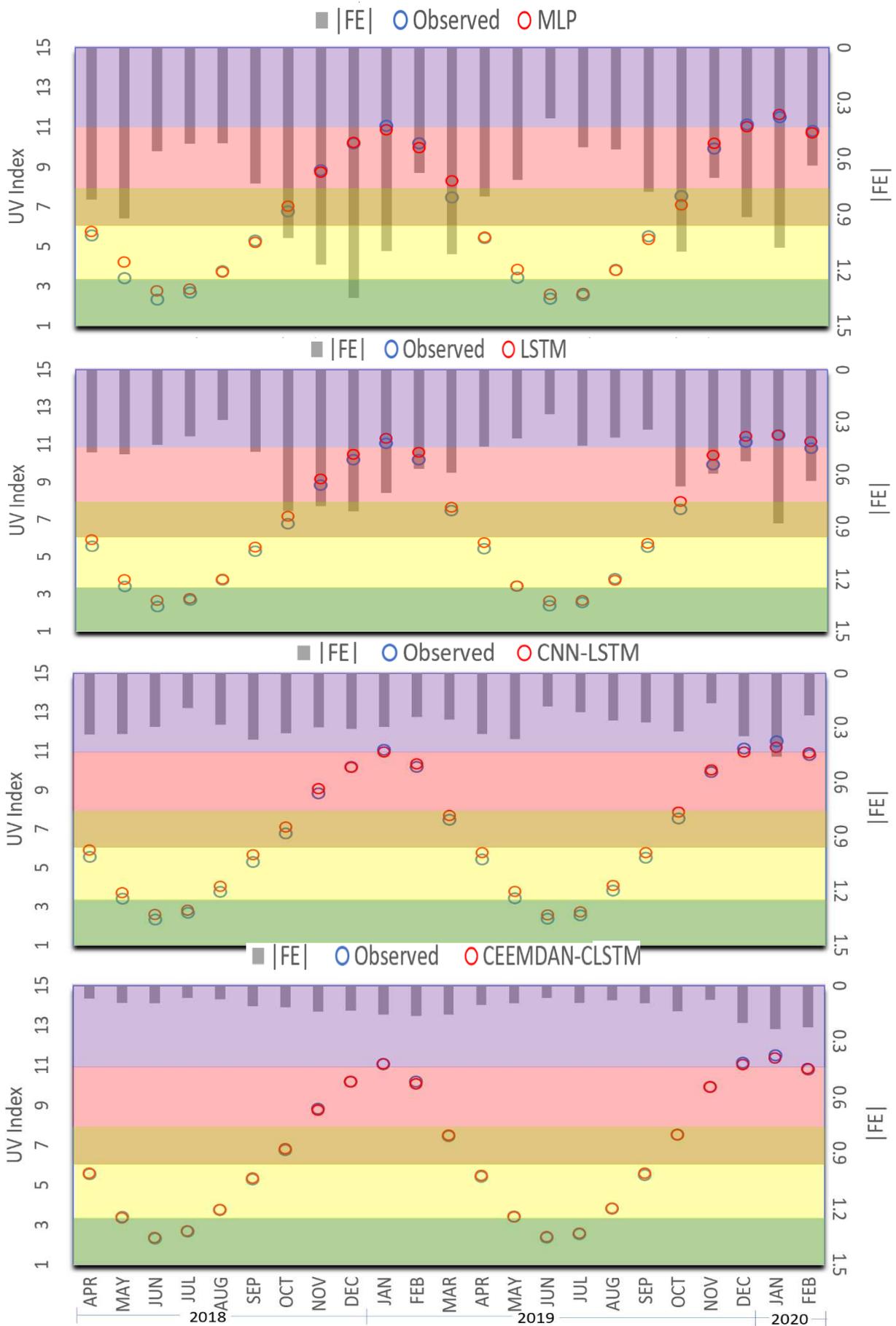


Figure 13 Monthly distribution of observed UVI and forecasted UVI using proposed CEEMDAN-CNN-LSTM model with three benchmark models (i.e., CNN-LSTM, LSTM, MLP) with respective absolute forecasting error ($|FE|$) for the testing phase. The color bar indicates the severity of UV values as described in Figure 1.