Stacked LSTM sequence-to-sequence autoencoder with feature selection for daily solar radiation prediction: a review and new modeling results

Applied Soft Computing Journal

A Hybrid Deep Learning Methodology with Feature Optimization Approach for Daily Solar Radiation Prediction

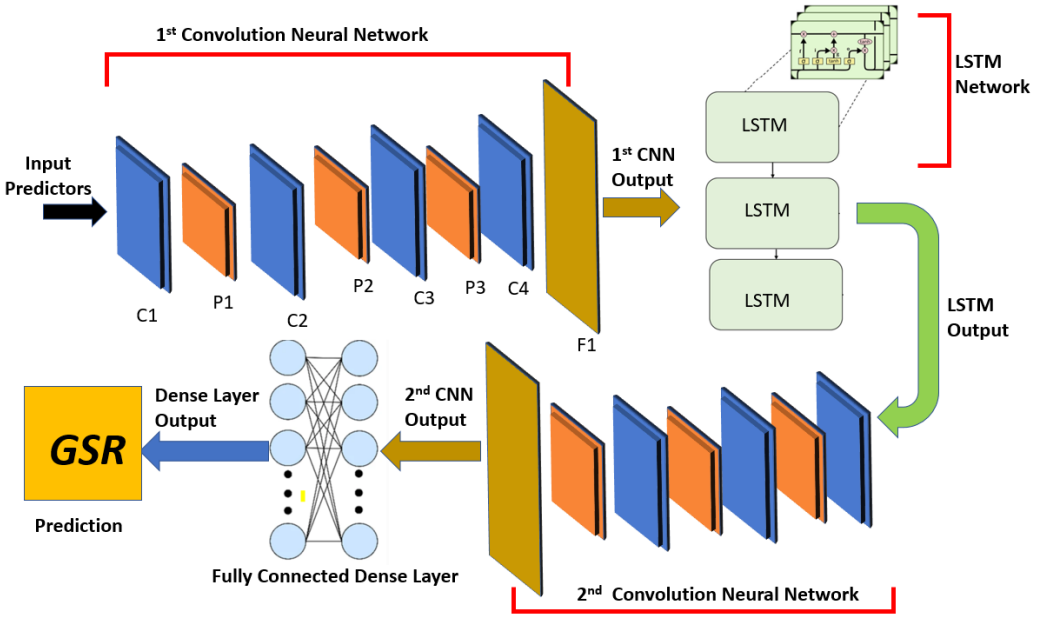
--Manuscript Draft--

Manuscript Number:	
Article Type:	Full Length Article
Keywords:	Global Solar Prediction; Deep Learning networks; Convolutional Neural Networks; Slime Mould Algorithm; Renewable Energy
Corresponding Author:	David Casillas-Perez SPAIN
First Author:	Sujan Ghimire
Order of Authors:	Sujan Ghimire
	Ravinesh Deo, Pr. PhD.
	Hua Wang
	David Casillas-Perez
	Sancho Salcedo-Sanz, Pr. PhD
	Mumtaz Ali
	Ekta Sharma
Abstract:	Global solar radiation (GSR) prediction plays an important role in the planning, controlling, and monitoring of solar power systems. The intermittent and stochastic behaviour of GSR is an important challenge in achieving the satisfactory prediction results. To address this issue, this study proposes a novel hybrid deep learning (DL) model, named SCLC that integrates the Convolutional Neural Network (CNN), Long Short-Term Memory Neural Network (LSTM) and the Multilayer Perceptron (MLP), with a feature selection mechanism by a Slime Mould Algorithm. The resultant model is applied to predict GSR at six solar farms in Queensland, Australia at daily temporal horizons in six distinct time steps. The structure of the final prediction system is the following: first, meteorological data from global climate models (GCM), and ground-based observations were used for GSR prediction. Second, a stochastic optimizer method known as Slime Mould Algorithm (SMA) has been implemented to select the best set of predictive features for this problem. In the next stage, a CNN entails a data processor including feature extractors drawing upon the statistically significant antecedent lagged predictor variables. Then, the model employs four independent LSTM algorithms to encapsulate a feature mapping scheme and another CNN layer is also used on these LSTM-based outputs to further optimize the results. This boosts the overall efficacy and accuracy of the SCLC-based prediction model by extracting effective features, finding the interdependence of data in time series, and detecting the best model suitable for relevant data. The final output of the system consists of a fully connected layer MLP, which emulates the next-day \emph{GSR} prediction. The proposed SCLC is comprehensively benchmarked, outperforming an ensemble of two DL approaches (CNN-LSTM, Deep Neural Network (DNN)) and three machine learning (ML) models (Artificial Neural Network (ANN), Random Forest Regressor (RFR), Self-Adaptive Differential Evolutionary Extreme Learning Machines (SADE-ELM)). A significant efficacy of SCLC in the prediction of next-day GSR has been observed in the model's testing phase for all six solar farms considered. Using detailed error analysis, visual and statistical metrics of GSR simulations relative to observations, this paper establishes the SCLC model's practical utility for the applications in renewable and sustainable energy resource management.
Suggested Reviewers:	David Camacho david.camacho@upm.es
	Javier Del Ser javier.delser@tecnalia.com

Graphical Abstract

A Hybrid Deep Learning Methodology with Feature Optimization Approach for Daily Solar Radiation Prediction

Sujan Ghimire, Ravinesh C. Deo, Hua Wang, David Casillas-Pérez, Sancho Salcedo-Sanz, Mumtaz Ali, Ekta Sharma



Highlights

A Hybrid Deep Learning Methodology with Feature Optimization Approach for Daily Solar Radiation Prediction

Sujan Ghimire, Ravinesh C. Deo, Hua Wang, David Casillas-Pérez, Sancho Salcedo-Sanz, Mumtaz Ali, Ekta Sharma

- A novel deep learning model for Global Solar Radiation prediction is proposed.
- The model integrates deep learning networks with Slime Mould Algorithm optimiser for optimal feature selection.
- Global climate model and meteorological data at solar farms in Australia are considered as predictive variables.
- Error analysis and statistical metrics establish the model's practicality for solar energy management problems.

A Hybrid Deep Learning Methodology with Feature Optimization Approach for Daily Solar Radiation Prediction

Sujan Ghimire[a], Ravinesh C. Deo[a], Hua Wang[b], David Casillas-Pérez[c], Sancho Salcedo-Sanz[d], Mumtaz Ali[b], Ekta Sharma[a]

[a]*School of Sciences, University of Southern Queensland, Springfield, QLD, 4300, Australia.*
[b]*Victoria University, Melbourne, Australia.*
[c]*Department of Signal Processing and Communications, Universidad Rey Juan Carlos, Fuenlabrada, 28942, Madrid, Spain.*
[d]*Department of Signal Processing and Communications, Universidad de Alcalá, Alcalá de Henares, 28805, Madrid, Spain.*

Abstract

Global solar radiation (*GSR*) prediction plays an important role in the planning, controlling, and monitoring of solar power systems. The intermittent and stochastic behaviour of *GSR* is an important challenge in achieving the satisfactory prediction results. To address this issue, this study proposes a novel hybrid deep learning (DL) model, named SCLC that integrates the Convolutional Neural Network (CNN), Long Short-Term Memory Neural Network (LSTM) and the Multilayer Perceptron (MLP), with a feature selection mechanism by a Slime Mould Algorithm. The resultant model is applied to predict *GSR* at six solar farms in Queensland, Australia at daily temporal horizons in six distinct time steps. The structure of the final prediction system is the following: first, meteorological data from global climate models (GCM), and ground-based observations were used for *GSR* prediction. Second, a stochastic optimizer method known as Slime Mould Algorithm (SMA) has been implemented to select the best set of predictive features for this problem. In the next stage, a CNN entails a data processor including feature extractors drawing upon the statistically significant antecedent lagged predictor variables. Then, the model employs four independent LSTM algorithms

*Corresponding author: (Prof. D. Casillas-Pérez) david.casillas@urjc.es

to encapsulate a feature mapping scheme and another CNN layer is also used on these LSTM-based outputs to further optimize the results. This boosts the overall efficacy and accuracy of the SCLC-based prediction model by extracting effective features, finding the interdependence of data in time series, and detecting the best model suitable for relevant data. The final output of the system consists of a fully connected layer MLP, which emulates the next-day *GSR* prediction. The proposed SCLC is comprehensively benchmarked, outperforming an ensemble of two DL approaches (CNN-LSTM, Deep Neural Network (DNN)) and three machine learning (ML) models (Artificial Neural Network (ANN), Random Forest Regressor (RFR), Self-Adaptive Differential Evolutionary Extreme Learning Machines (SADE-ELM)). A significant efficacy of SCLC in the prediction of next-day *GSR* has been observed in the model's testing phase for all six solar farms considered. Using detailed error analysis, visual and statistical metrics of *GSR* simulations relative to observations, this paper establishes the SCLC model's practical utility for the applications in renewable and sustainable energy resource management.

Keywords: Global Solar Prediction, Deep Learning networks, Convolutional Neural Networks, Slime Mould Algorithm, Renewable Energy, Global Climate Models
PACS: 02.70.-c, 07.05.Mh
2000 MSC: 68T05, 68T20

1. Introduction

There is unprecedented momentum to leave fossil fuel age behind us. With the rapid transition to renewable energy and energy efficiency, the Governments globally are trying to turn the tide [1]. Solar power is the key to a clean energy future. With unlimited availability, its utilisation is rapidly increasing around the world, which may relieve the current world energy crisis [2, 3, 4, 5]. The sun offers the most abundant, reliable, and pollution-free power in the world. Being a sustainable, and infinite energy source, solar energy carries the potential to fulfill the energy needs of the entire world. Global solar radiation (*GSR*) identifies the solar power potential [6]. An accurate *GSR* prediction is very important for an effective solar energy utilization, robust planning, decision making, power system operation, management, and investment applications [7]. Furthermore, accurate

GSR predictions are vital for the establishment of reliability and permanency of the electricity grid, and the reduction of risk and costs of energy markets and systems. This will be immensely beneficial not only to the power plants and grid operators but also to traders and Government policymakers [8].

Many researchers have proposed different models to predict *GSR*, including empirical prediction models (EM) [9, 10, 11, 12]. EM are computationally efficient and easy to calculate but because of rapid changes in weather conditions, but they cannot, in general, accurately predict short-term *GSR* [13]. In many cases they result in partially unsatisfactory estimates of *GSR* [14, 15, 16, 17]. Physical [18, 19] and Numerical Weather Prediction (NWP) [20, 21] are other type of *GSR* prediction models quite studied in the literature. In this case, there are challenges such as sourcing and selecting the inputs for the physical or the NWP models [22, 23, 24], and there are also issues related to the high computation cost of the models. There are other types of approaches for *GSR* prediction such as remote sensing retrieval [25], time series-based algorithms [26, 27] and, of course, Machine Learning (ML)-based models.

Artificial Neural Networks (ANN) are probably the most frequently used ML models for *GSR* prediction. They include modalities such as Multilayer Perception Neural Networks (MLP-NNs) [28, 29, 30], recurrent neural networks [31], Radial Basis Function Neural Networks (RBF-NNs) [32], evolutionary neural approaches [33, 34] Generalized Regression Neural Networks (GR-NNs) [35], or Extreme Learning Machines (ELM) [36, 37, 38], among others. Other ML techniques such as Random Forest (RF) Regression (RFR) [39, 40], Support Vector Regression algorithms (SVR) [41, 42, 43, 44], Gaussian Processes [45] or Adaptive Neuro-Fuzzy Inference System (ANFIS) [46] have been also used in *GSR* prediction problems. There are other works which have compared the performance of different ML techniques in *GSR* prediction, such as [47], where MLP, ELM and SVR have been compared in a problem of *GSR* from satellite measurements, [48] where EM, ANN, SVR, Gaussian Process, Genetic Programming (GP), and ARIMA models were compared in a problem of *GSR* prediction in cities of Australia. In [49] ANNs, EM, time series, and mathematical models were compared for *GSR* prediction, and [50] used three ML algorithms (SVM, ANN, and ANFIS) to predict the daily global solar radiation data of six stations in Mexico, among many other works involving shallow ML structures. Moreover, in [44], a wavelet-coupled SVR model for forecasting global incident solar radiation using limited meteorological datasets for Queensland's three solar rich cites

was developed using only the sunshine hours, minimum temperature, maximum temperature, windspeed, evaporation and precipitation as the predictor variables.

On the other hand, Deep Learning (DL) is gaining huge popularity from the past decade. This is due to their robust architecture, powerful nonlinear structure, generalisation capability, and unsupervised feature learning. Unlike shallow ML models, DL models can extract features and latent invariant architectures in data. This makes them a popular choice for areas such as imaging, speech recognition, natural language processing, autonomous driving, or computer vision. Solar prediction with DL technologies is a new and promising research area [51]. An LSTM was used for predicting next-day hourly solar radiance using meteorological features [52]. Additionally, researchers used CNN to extract robust features from predictive variables while LSTM is used to predict *GSR* [53]. In [54] a hybrid DL model for *GSR* prediction that combines Gated Recurrent Unit (GRU) and attention mechanisms was proposed. In [55] a hybrid DL model which combines Deep Belief Networks (DBNs) and Embedded Clustering (ECs) to estimate solar irradiance was introduced, whereas in [56] a novel DL model for *GSR* prediction that incorporates the Sine Cosine Algorithm (SCA), the Bidirectional LSTM network, and the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) was proposed. In [57] a hybrid DL method to optimize deep neural networks (GRU, LSTM, and RNN) for *GSR* prediction based on a genetic algorithm (GA) was proposed, and in [58] a hybrid DL model that combines LSTMs and Choquet Integrals based aggregation functions to predict *GSR* at six different locations in Finland was introduced. In [59] an end-to-end hybrid DL model that incorporates ResNet (Residual Network) and LSTM for short-term *GSR* prediction was proposed. More recently, in [60] a DL-based hybrid method for Global Horizontal Irradiance (GHI) forecasting is proposed. The method consists of a deep learning-based clustering algorithm with a Feature Attention Deep Forecasting deep neural network to generate the GHI forecasts. In [61] a novel methodology to forecast GHI in short- and long-term time-horizons, based on a hybrid DL architecture is proposed. Specifically, the system consists in the combination of a Variational Mode Decomposition algorithm and two CNN, together with RF or LSTM algorithm.

As previously discussed, the single model usage in modelling has disadvantages of intermittent and fluctuating nature of *GSR*. Due to the shortcomings of single models and the need for greater accuracy in *GSR* prediction,

hybrid models have been developed and widely used for predicting *GSR*. However, there are some concerns of these hybrid models too: First, in most studies *GSR* is the only factor considered (or clear sky data taken) during model development, ignoring meteorological factors. In practicality, as the weather varies significantly, these models cannot fully reflect the change in *GSR*. Second, the weather forecast becomes more accurate and convenient, it gets rarely modelled with hybrid models as input parameters. Third, feature selection algorithms are not solely preferred during the modeling process of these hybrid models. It is important to note that, although the resultant hybrid of deep fusion network benefits from both DL and ML, it should alleviate the drawbacks of both the techniques such as computation time and cost. Hybrid models provide more accurate and less computationally expensive solutions when used through Tensorflow, which is Google's open-source platform [62]. Most published literature for *GSR* fails to address these criteria. These are some of the gaps the present study attempts to address.

This paper therefore proposes a novel DL-based hybrid model that overcomes the above limitations, and produces accurate *GSR* predictions. A new hybrid DL model, which process the input data with a sequential application of Slime Mould Algorithm (SMA) for feature selection, CNN, LSTM network, CNN and a final processing with a MLP has been developed in this study, to overcome the shortcomings mentioned above and obtain a more accurate *GSR* prediction. The complete prediction system is called SCLC, and we have tested it by comparison with an ensemble of two alternative DL approaches (CNN-LSTM and Deep Neural Network (DNN)) and three shallow ML models (Artificial Neural Network, Random Forest and a variation of the Extreme Learning Machine), in a problem of *GSR* prediction at six solar farms in Australia.

The remainder of the paper has been structured in the following way: next section summarizes the different methods which form the SCLC prediction system, including, the SMA, CNN and LSTM algorithms. The data and area of the experimental study are described in Section 3. The hybrid SCLC model description and tuning is presented in Section 4. The experimental part and results obtained are summarized in Section 5. Finally, Section 6 closes the paper with some conclusions and remarks on the research carried out and the results obtained.

2. Methods: theoretical overview

In this section we provide a brief theoretical description of the different methods which form the SCLC system for solar radiation prediction. We first describe the SMA approach for feature selection, we then describe the foundations of the CNN algorithm and the LSTM approach.

2.1. Slime Mould Algorithm for feature Selection

In this study, a wrapper feature selection method [63, 64, 65] based upon a meta-heuristic algorithm called Slime Mould Algorithm (SMA) is firstly used to select the optimal features for *GSR* prediction. We have selected SMA based on its recent performance as a metaheuristic algorithm derived from the diffusion and foraging behavior of slime mould [66], and the algorithm's several features with a unique mathematical model that uses adaptive weights to simulate the process of producing positive and negative feedback of the propagation wave of slime mould. It is fundamentally based on bio-oscillator to form the optimal path for connecting food with excellent exploratory ability and exploitation propensity [67]. Mathematically, SMA can be divided into three phases: approach, wrap and grabble food.

- Stage 1 (Approach food): A slime mould approaches food based on its odor in the air, so the following formula mimics the behavior of the slime mould towards food.

$$\mathbf{X}(t+1) = \begin{cases} \mathbf{X}_b(t) + v_b \left(W\mathbf{X}_A(t) - \mathbf{X}_B(t) \right), & r < p \\ v_c \mathbf{X}(t), & r \geq p \end{cases} \tag{1}$$

  where $\mathbf{X}$ is the position of slime mould, and $\mathbf{X}_b$ represents the location of the currently found individual with the highest odor concentration. Two individuals randomly selected from the population are recorded as $\mathbf{X}_A$ and $\mathbf{X}_B$. The parameter v_b varies between the interval $[-a, a]$, v_c linearly decreases from 1 to 0. Also, a is calculated as Equation (2):

$$a = \operatorname{arctanh}\left(-\frac{t}{M_t} + 1 \right), \tag{2}$$

  where t is the current iteration and M_t is the maximum iterations; and p is derived by Equation (3):

$$p = \tanh|S(i) - DF|. \tag{3}$$

Here, $i \in \{1, 2, \cdots, n\}$ and $S(i)$ is the fitness of $\mathbf{X}$. DF is the best fitness obtained in all iterations. The weight of slime mould W is calculated following the next expression:

$$W(SI(i)) = \begin{cases} 1 + r \log\left(\dfrac{bF - S(i)}{bF - wF} + 1\right), & \text{if } S(i) < \text{Med}[S] \\ 1 - r \log\left(\dfrac{bF - S(i)}{bF - wF} + 1\right), & \text{otherwise} \end{cases}, \quad (4)$$

where

$$SI = \text{Sort}[S], \quad (5)$$

and $r \sim U(0, 1)$ is an uniform random variable between 0 and 1. The Med$[\cdot]$ refers to the median operator. Also, the Smell Index (SI) refers to the sequence of fitness values after ascending order in the minimum value problem, $S(i)$ refers to a set of individuals ranked in the top half of the population after sorting by fitness values, bF is the optimal fitness value obtained currently and wF is the worst one.

- Stage 2 (Wrap food): A slime mould's search pattern changes based on the quality of food. When there is a high concentration of food near a region, the weight near it will be greater. If the concentration is low, the region's weight will be lower, and it will be forced to explore other locations. The location of the slime mould is updated in the stage based on Equation (6):

$$\mathbf{X}(t+1) = \begin{cases} u(UB - LB) + LB, & r < z \\ \mathbf{X}_b(t) + v_b \left(W\mathbf{X}_A(t) - X_B(t)\right), & z \leq r < p \\ v_c \mathbf{X}(t), & r \geq p \end{cases} \quad (6)$$

where LB and UB are the lower and upper bounds respectively, $u, r \sim U(0, 1)$, and z is a probability used to tradeoff between exploitation and exploration.

- Stage 3 (Grabble food): In this stage, a propagation wave generated by the biological oscillator changes the cytoplasmic flow in veins, so that slime mould moves to better locations for food concentration. W, v_b and v_c is used to mimic the variation of venous width. The variables, v_b and v_c, oscillate between $[-a, a]$ and $[-1, 1]$ respectively. As the iteration number increases, v_b and v_c draw closer to zero. The intuitive and detailed process of SMA is shown in Figure 1.

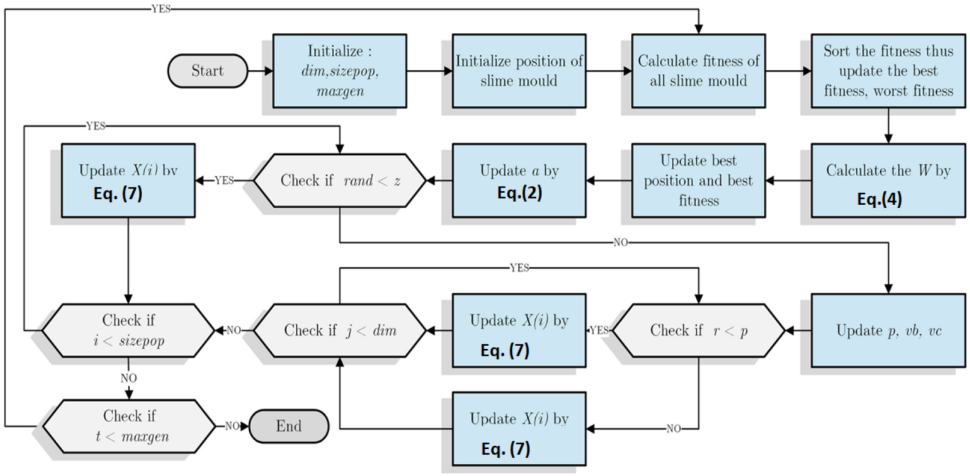


Figure 1: Descriptive flowchart for the relevant steps in the Slime Mould Algorithm (SMA) adopted as a feature selection algorithm for the prediction of *GSR*.

Since, this SMA is wrapper -based method, to implement the SMA algorithm for feature selection, a learning algorithm must be incorporated. This study has utilised K-Nearest Neighbours (KNN) regressor [68] as a learning algorithm for the feature selection (FS) using SMA. The objective of FS is to increase the accuracy while also minimizing the number of features to be selected, therefore, we have chosen fitness value (FV) as the root mean square error (RMSE) complement of regression accuracy and needs to be minimised to get the best feature subset. The three phases of the proposed SMA based FS solution are outlined below:

- Initialization Phase: A SMA produces an initial population of N candidate solutions, where each entity covers a set of features for consideration. The quality and convergence of the optimal solution are critically affected by this step. The population X_0 is randomly generated by Equation (1), and the fitness value is calculated.

- Update Phase: Every new position is evaluated using the fitness function. If the solution quality of a new position is better than that of the current position, the position is updated. The opposite-based learning (OBL) approach [69] is used to update each search agent's position, Equation (4) and Equation (6) are used. To improve the search process by exploring new regions in quest of the optimal solution, increasing

algorithm diversity, avoiding local optima, and confirming whether the new solution is better than the old one, the basic principle of OBL is to consider a solution and its matching opposite solution simultaneously. The FV of the new population is calculated, and then the best solution is determined. Repetition of this process will continue until the termination condition (*i.e.*, the maximum number of function evaluations) is reached. The SMA process returns the best solution obtained in the previous step and only the best features are retained from the original data.

- Termination phase: Until the stopping criteria are satisfied, the maximum number of function evaluations of the proposed algorithm are performed, and the best viable feature subset is discovered.

2.2. Convolutional Neural Network (CNN)

CNN models are a popular choice of the feed-forward network since CNN shares features parameters and enable dimensionality reduction. As CNN enables parameter sharing, the number of parameters gets reduced therefore the computations are also decreased. Compared to its predecessors, the main advantage of CNN is that it automatically detects important spatial features without any human supervision. They also extract hidden features and create filters according to the data patterns. The convolution layers of a CNN are optimized during training so that they extract highly discriminative features, while the latter layers resemble multilayer perceptron's which execute classifying or regression work. There have been very few papers describing the application of CNNs for solar radiation modeling [70], although CNNs have been successfully applied to many practical applications. Figure 2 shows the structure of CNN, which consists of an input layer, convolution layer, pooling layer, full connection layer, and an output layer [71]. The convolution layer utilizes a mathematics operator known as convolution ($*$) to extract features bound to local regions within the input data, a pooling layer is used to reduce the dimensions of the input data, and at the end of the CNN, a fully connected dense layer is employed to predict the output based on the extracted features. A convolutional neural network has two main features: weight sharing and local connections [72]. Convolutional layers target strata on the target variable (GSR) and its associated input variables (meteorological data) to extract spatial patterns, which can be expressed mathematically

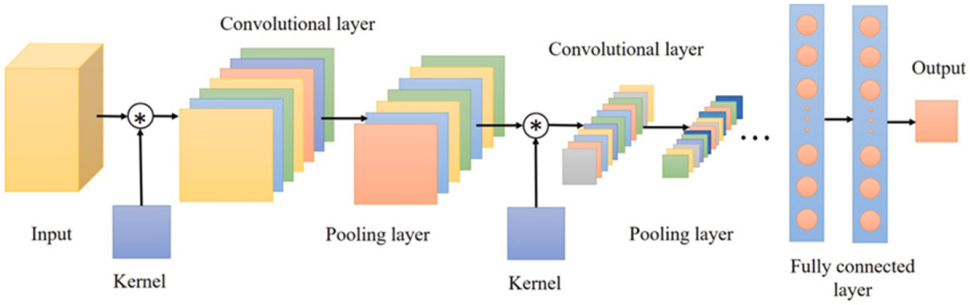


Figure 2: A basic architecture of a Convolutional Neural Network (CNN) model, where $*$ represents the convolution operation and kernel denotes the size of the convolution kernel.

as:

$$h = f(x * W + b) \tag{7}$$

where x is input data, $*$ means convolution operation, W is the weight of the convolution kernel, and b is the offset value. The function $f(\cdot)$ denotes the activation function. An efficient activation function, namely, the Rectified linear unit (ReLu), is employed in this study:

$$f(x) = \max\{0, x\} \tag{8}$$

2.3. Long Short-Term Memory Network (LSTM)

Relative insensitivity to gap length is an advantage of LSTM over RNNs, hidden Markov models, and other sequence learning methods in numerous applications. LSTMs were developed to deal with the vanishing gradient problem that can be encountered when training traditional RNNs. Using LSTM, time series forecasting models can predict future values based on previous, sequential data. This provides greater accuracy for demand forecasters which results in better decision-making for the business. We can say that, when we move from RNN to LSTM (Long Short-Term Memory), we are introducing more and more controlling knobs, which control the flow and mixing of Inputs as per trained Weights. And thus, bringing more flexibility in controlling the outputs.

Different versions of conventional neural networks have been widely used to analyze time-series data, predict trends, and forecast. The accuracy of neural networks in modeling complex relationships has been praised, but

they have been unable to handle historical data dependencies. To deal with historical data dependencies, Recurrent Neural Networks (RNN) have been introduced, which utilize the information accumulated during previous time steps using network loops. In RNN because of network loops, data from previous time steps can be transferred to the current time step to enhance prediction accuracy. Although RNNs are great at handling short-term dependencies, they cannot handle long-term dependencies in time series data due to issues with gradient explosion and disappearance.

To resolve this issue, in [73] a time RNN known as LSTM was introduced. Its architecture incorporates special units called memory cells, which function in place of traditional neurons. Moreover, LSTM employs a gate mechanism that has input, forget, and output gates that enable updating and controlling the information flow within the network. LSTM overcomes the problems associated with gradient disappearance by using internal memory cells and gate mechanism and effectively handles the long-term dependence of data by using internal memory cells. Because LSTM is able to efficiently acquire temporal features from time-series data and handle long-term dependencies, LSTM has been extensively used by researchers for predicting *GSR* [52, 74, 75, 76]. The information distribution of LSTM are exemplified in Figure 3 and mathematically described as follows [77]:

- Forget gate, based on the last hidden state h_{t-1}, and input data x_t, the forget gate f_t, enables the LSTM to determine which information needs to be discarded from the cell state:

$$f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f) \tag{9}$$

  where σ is the sigmoid activation function, w_f is the weight matrices and b_f is the bias vector.

- Input gate, i_t determines which information is saved to each new candidate state $\tilde{C}_t$:

$$\tilde{C}_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c) \tag{10}$$

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \tag{11}$$

  where $\tanh(\cdot)$ demonstrates the hyperbolic tangent function.

- A combination of the previous state C_{t-1} and the new candidate state $\tilde{C}_t$ is used to update the new state, C_t as follows:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{12}$$

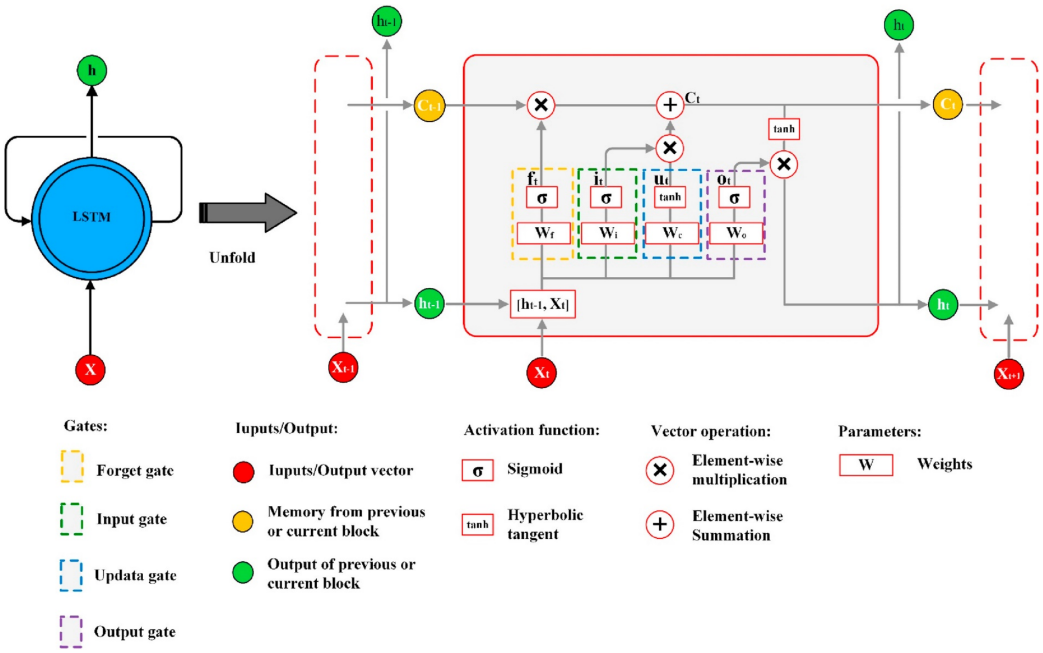


Figure 3: The structure of Long-Short Term Memory (LSTM) Network.

- In the end, o_t a gate is developed to regulate the output of the LSTM cell. Multiplication of o_t and state of the cell C_t, activated by tanh function is the desired output h_t as follows:

$$o_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \tag{13}$$

$$h_t = o_t \tanh(C_t) \tag{14}$$

3. Study Area and Data

Queensland is renowned for being a leader in the Australian solar revolution. With high solar radiation, the region of Western Downs has gained a lot of recognition for its pro-solar movements. Currently, there are 44 large-scale renewable energy projects in Queensland, generating $9.9 billion in investments, 7000 construction jobs, and 5156 megawatts (MW) of renewable energy [78]. This saves about 12 million tons of carbon per year. In Queensland, there are currently 6,200 MW of renewable energy capacity in the state, including rooftop solar panels, which account for 20% of total

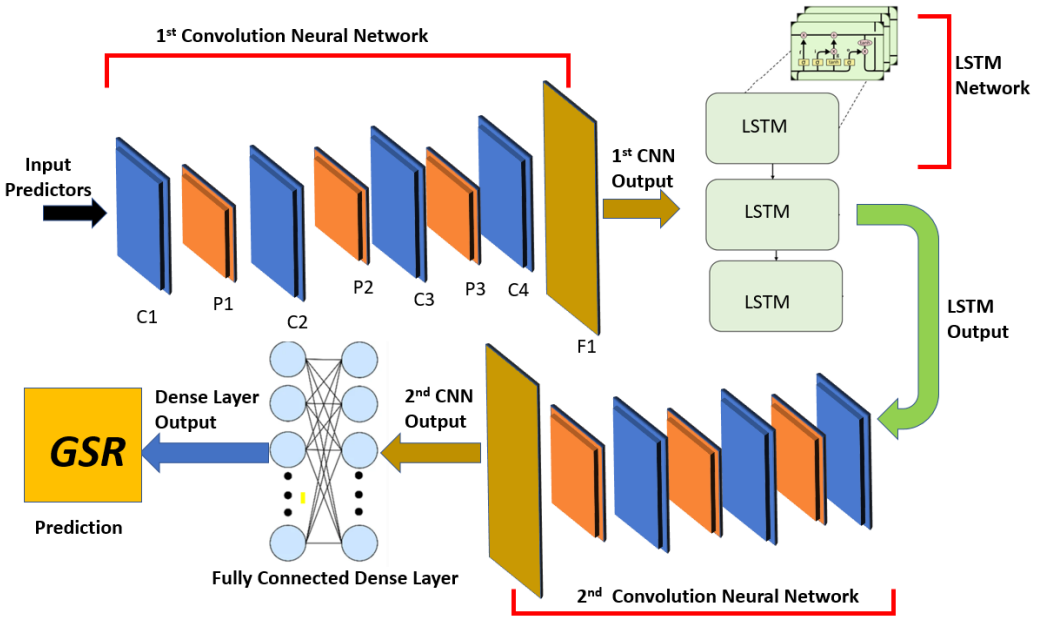


Figure 4: Architecture of the proposed *GSR* prediction method based on the hybrid SMA-CNN-LSTM-CNN-MLP (SCLC) model.

electricity consumption [79]. Six solar farms in Queensland, Australia with power outputs ranging from 55 MW to 148 MW, were chosen for the study.

1. The Cape York Battery Power Plant is the first grid-connected battery power plant in Australia with both solar generation and battery storage. According to the developer, the 20MW/80MWh Fluence battery-based energy storage system plus 55 MW solar generation will provide firm clean energy through a single connection point, using a single power plant controller.

2. The Chinchilla Solar Farm is situated 140 km north of Toowoomba, Australia, near the township of Chinchilla. There will be around 250,000 thin-film photovoltaic (PV) modules installed at the proposed 100 MW project and they will produce enough solar energy to serve approximately 40,000 average Queensland homes.

3. Sun metal solar Farm possesses a 125 MW generation capacity and is located near Townsville, in northern Queensland. Sun Metals is building the farm to secure their zinc refinery there with an uninterruptible

power supply for a lower cost. A total of 1,167,000 solar panels are used and will produce 261 GWh of electricity annually, which accounts for almost 29% of the zinc refinery's current electrical needs.

4. Clermont Solar Farm would build a single-axis tracking solar power plant, located 106km north-northeast of Emerald, Queensland. In total, the site can generate approximately 89 MW and encompasses approximately 497 acres. A total of 205 GWh of electricity will be generated annually using 275,442 PV panels, enough to power approximately 30,996 households.

5. Ubergy has received approval for its Barunggam Solar Farm located on Baking Board, 14km from Chinchilla in Queensland. This solar farm is expected to be 140 MW in size.

6. Cameby solar farm will have a capacity of 148MW and is located in an area of 463ha that presently serves as grazing land with little agricultural potential.

Moreover, it supports the Queensland government's goal of producing 50 percent of its energy from renewable sources by 2030, while establishing Queensland as a leader in renewable energy. Table 1 provides details of the study site (the statistics of *GSR*) and Figure 5 depicts their locations.

A supervised learning process is one in which an example input (predictor) and the desired output (predictand) are presented to a predictive model and a general rule is derived to map inputs to outputs. Since *GSR* prediction is a supervised learning process, we require both predictors and predictands. In this study, we have used global climate models (GCM) meteorological data and ground-based observation data from Scientific Information for Landowners (SILO) to make predictions. The Queensland Climate Change Centre of Excellence (QCCCE), which is part of the Department of Science, Information Technology, Innovation and the Arts (DSITIA), manages the Long Paddock SILO database [80]. A GCM output archive is maintained by the Centre for Environmental Data Analysis (CEDA) as a server for the CMIP5 project's GCM output collection [81]. Daily atmospheric model outputs for historical are sourced from this repository. The models include CSIRO-BOM ACCESS1-0 (grid size 1.25×1.875) [82], MOHC Hadley-GEM2-CC (grid size 1.25×1.875) [83] and the MRI MRI-CGCM3 (grid size 1.12148×1.125) [84]. The GCM model outputs are indexed by dimensions of longitude, latitude,

Table 1: Descriptive statistics of the target variable: daily global solar radiation (GSR; $\mathrm{MJm^{-2}day^{-1}}$) for six solar farms in Queensland, Australia.

Property	Barunggam Solar Farm	Cameby Solar Farm	Cape York Solar Storage	Chinchilla Solar Farm	Clermont Solar Farm	Sun Metals Solar Farm
Latitude	26.685°S	26.682°S	15.898°S	26.670°S	22.839°S	19.437°S
Longitude	150.765°E	150.510°E	144.857°E	150.793°E	147.581°E	146.696°E
Capacity (MW)	140	148	55	100	75	125
Median ($\mathrm{MJm^{-2}}$)	19.00	19.00	20.00	19.00	20.00	20.00
mean ($\mathrm{MJm^{-2}}$)	19.23	19.28	19.45	19.21	20.03	19.88
Standard deviation ($\mathrm{MJm^{-2}}$)	6.36	6.43	4.84	6.35	5.85	5.55
Variance ($\mathrm{MJm^{-2}}$)	40.49	41.34	23.38	40.27	34.18	30.77
Maximum ($\mathrm{MJm^{-2}}$)	33.00	32.00	29.00	32.00	32.00	31.00
Minimum ($\mathrm{MJm^{-2}}$)	4.00	4.00	5.00	4.00	4.00	4.00
Mode ($\mathrm{MJm^{-2}}$)	29.00	28.00	24.00	28.00	28.00	27.00
Interquartile range ($\mathrm{MJm^{-2}}$)	9.00	9.00	6.00	9.00	8.00	7.00
Skewness	-0.18	-0.18	-0.51	-0.18	-0.38	-0.54
Kurtosis	2.34	2.34	2.83	2.35	2.65	2.71

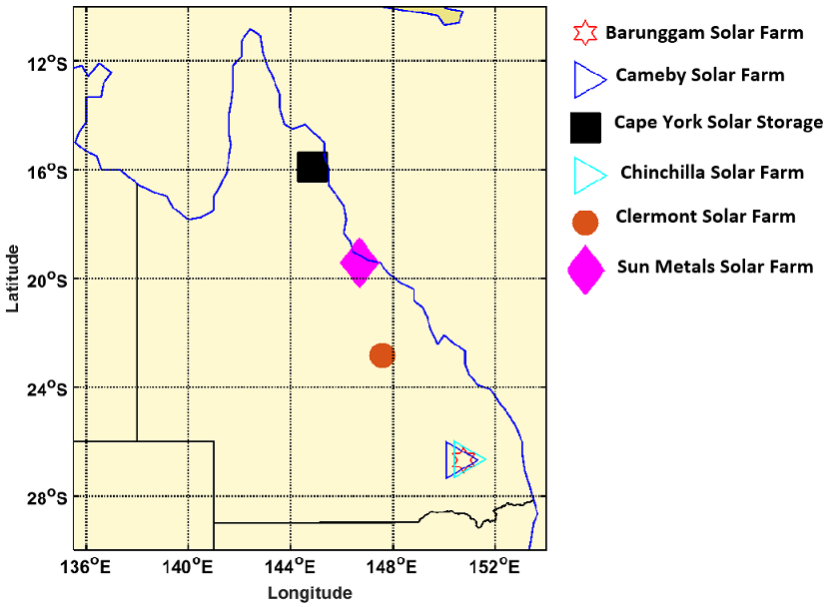


Figure 5: Study sites in Queensland, Australia where the proposed deep hybrid SCLC model was implemented.

time, atmospheric pressure (at 8 levels), or near-surface readings. The historical outputs span the period 1950-01-01T12:00:00 to 2006-01-01T00:00:00. Table 2 provides a brief overview of each of the meteorological variables comprised in the dataset. This final dataset contained 20455 records and 75 meteorological variables (20455×75).

4. The Hybrid SCLC Model Development

The general method for predicting *GSR* introduced in this study is illustrated in Figure 6.

A hybrid model generally performs better than a single/standalone model. Considering the usefulness of LSTMs and CNNs, this study has proposed a new functional model for extracting temporal and spatial features to predict *GSR* with greater accuracy. In this study, a hybrid model, called Hybrid SCLC referred to as a connection of SMA, CNN, LSTM, and CNN is proposed to model daily *GSR*, as illustrated in Figure 4. The proposed model is exceptionally capable of predicting *GSR* by extracting complex features and patterns from meteorological variables. The model is developed in the

Table 2: Description of the global pool of predictor variables used in daily *GSR* prediction, including atmospheric variables from global climate models and Ground-based observational climate data from scientific information for landowners (SILO).

Data Repository Name	**Variable**	**Description**	**Units**
Global Circulation Model (GCM) Atmospheric Predictor Variables	clt	Cloud Area Fraction	%
	hfls	Surface Upward Latent Heat Flux	wm^{-2}
	hfss	Surface Upward Sensible Heat Flux	wm^{-2}
	hur	Relative Humidity	%
	hus	Near Surface Specific Humidity	gkg^{-1}
	pr	Precipitation	$kgm^{-2}s^{-1}$
	prc	Convective Precipitation	$kgm^{-2}s^{-1}$
	prsn	Solid Precipitation	$kgm^{-2}s^{-1}$
	psl	Sea Level Pressure	pa
	rhs	Near Surface Relative Humidity	%
	rhsmax	Surface Daily Max Relative Humidity	%
	rhsmin	Surface Daily Min Relative Humidity	%
	sfcWind	Wind Speed	ms^{-1}
	sfcWindmax	Daily Maximum Near-Surface Wind Speed	ms^{-1}
	ta	Air Temperature	K
	tas	Near Surface Air Temperature	K
	tasmax	Daily Max Near-Surface Air Temperature	K
	tasmin	Daily Min Near-Surface Air Temperature	K
	ua	Eastward Wind	ms^{-1}
	uas	Eastern Near-Surface Wind	ms^{-1}
	va	Northward Wind	ms^{-1}
	vas	Northern Near-Surface Wind	ms^{-1}
	wap	Omega (Lagrangian Tendency of Air Pressure)	pas^{-1}
	zg	Geopotential Height	m
Ground-based SILO	T.Max	Maximum Temperature	K
	T.Min	Minimum Temperature	K
	Rain	Rainfall	mm
	Evap	Evaporation	mm
	VP	Vapor Pressure	Pa
	RHmaxT	Relative Humidity at Maximum Temperature	%
	RHminT	Relative Humidity at Minimum Temperature	%

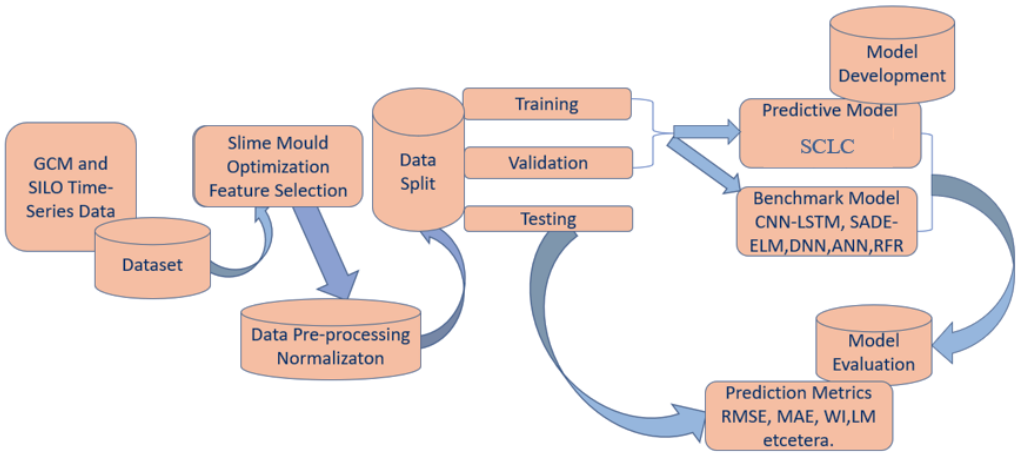


Figure 6: Workflow diagram detailing the necessary steps taken to design proposed deep hybrid SCLC model for daily *GSR* prediction at six solar farms of Queensland, Australia.

following stages:

- Stage 1: Initially, meteorological data from the global climate models (GCM), and ground-based observations from Scientific Information for Landowners (SILO) were used for *GSR* prediction.
- Stage 2: Secondly, Slime Mould Algorithm was adopted as a stochastic optimiser in the study to extract features. The proposed SMA has several new features with a unique mathematical model. It uses adaptive weights to simulate the process of producing positive and negative feedback of the propagation wave of slime mould. This is based on bio-oscillator to form the optimal path for connecting food with excellent exploratory ability and exploitation propensity.
- Stage 3: In the next stage, the CNN model is introduced as the first extraction layer of deep learning. The introduction of CNN helps in reducing dimensionality and thereby reduces the computation time. CNN entails a data processer including feature extractors drawing upon statistically significant antecedent lagged predictor variables.
- Stage 4: In the next stage, the model building employed four independent LSTMs to encapsulate feature mapping schemes.

- Stage 5: For further model building stage, another CNN layer was employed on the LSTM outputs. The convolutional layers in the second CNN model apply the convolution operation on time series data (input from LSTM) to extract spatial patterns and intrinsic characteristics from diverse meteorological variables. CNNs typically consist of several levels of convolutional– pooling layers and there are several convolutions runs are performed on each layer to collect useful information. The second layer boosts the efficacy and accuracy of resultant hybrid SCLC by extracting effective features, finding the interdependence of data in time series, and detecting the best mode suitable for relevant data. During this process, the CNN uses weights for the meteorological parameter based on its effect on *GSR*.

- Stage 6: Lastly, the final output from stage five goes through the fully connected (FC) layer (MLP) resulting in next-day *GSR* prediction. The proposed hybrid SMA-CNN-LSTM-CNN-MLP (SCLC) final layer has fully connected dense layers and can predict *GSR* over some time.

The CNN unit's output is a flattened feature vector $h^l = (h_1, h_2, \ldots, h_l)$ where l represents the number of filters in CNN. The Equation (15) represents the equation deployed at that level. σ is a non-linear activation function, w is the weight of the ith node for layer $l-1$ and jth node for layer l, and b_i^{l-1} represent bias.

$$d_i^l = \sum_j w_{ji}^{l-1} \left(\sigma(h_i^{l-1}) + b_i^{l-1}\right) \tag{15}$$

The proposed hybrid SCLC model is depicted in Figure 1. The first CNN layer is composed of four convolutional layers followed by a pooling layer, then the results are flattened. Convolutional layer 1 reads through input data (predictor) and displays results as feature maps, layer 2 and layer 3 perform the same operation on feature maps created by layer 1 and layer 2 respectively, and layer 4 repeats the process to amplify any salient features. Feature maps that are extracted after layer 4 are then flattened and fed to the 3-layer stacked LSTM model. The next step involves transferring the LSTM extracted temporal information from predictor variables to the input layer of the second CNN model (3 Convolutional layers and 3 pooling layers). The feature maps extracted after the pooling layer of the second CNN are then flattened into a long vector (1-dimensional array). Lastly, we use a fully connected layer (*i.e.*, dense) to aggregate the data and predict the *GSR* by

analyzing the extracted features. In this architecture, spatial and temporal features are extracted independently with CNNs and LSTMs, hence using the positive aspects of both CNNs and LSTMs and producing a robust model.

4.1. Data Normalisation

In the prediction task, numerical values with different scales must be normalized, ignoring this step hinders gradient descent-based algorithms, resulting in slower convergence speeds and distorting prediction results [85]. Thus, this study has utilised the Z-score normalization method to the predictor dataset intending to scale all the variables to a similar range. Let $X = \{x_1, x_2, \ldots, x_L\}$ be the considered time-series input data with L component. Each sample of X was normalized with a center of 0 and a standard deviation of 1 by following Equation (16):

$$\tilde{X} = \frac{x - \mu}{\sigma} \tag{16}$$

where μ and σ are the mean and standard deviation of X, respectively. Finally, the scaled data is represented with $\tilde{X} = \{\tilde{x}_1, \ldots, \tilde{x}_L\}$ and the subset of the input can be prepared. Since the normalization is invertible, the results are unaffected.

4.2. Feature Selection

This study has utilised the meta-heuristic (SMA) as search algorithm and K-Nearest Neighbor Regressor (KNNR) as a machine learning algorithm for the selection of optimal input for the prediction of *GSR* at six solar farms of Queensland, Australia. SMA feature selection process involves the partitioning of the normative matrix of predictors and predictands (*GSR*) into training sets and testing sets (e.g., 80% for training, 20% for testing in 5-fold cross-validation) and running the KNNR on the selected features in the dataset. Each feature subset considered in the SMA FS is trained by the KNNR and its performance is evaluated by measuring the generalization performance on the original data. Feature subsets with minimum RMSE are considered the optimal subset. The SMA feature selection is run with the below configuration:

- Population $(N) = [10, 20, 50, 80, 100, 200, 300, 500]$.
- The number of maximum iterations $(T) = 50$.

- Number of k in K-nearest neighbor $(K) = 5$
- Probability of exploration and exploitation capability $(z) = 0.03$

This study also demonstrated that population size affects SMA feature selection performance as measured by root mean square error (Fitness Value, FV). Hence, this study evaluated the proposed SMA FS for populations of 10, 20, 50, 80, 100, 200, 300, and 500; the convergence curve was plotted to show the optimal fitness value in the slime mould during the iteration process. Based on the convergence curve (Figure 6), it can be concluded that population increases are not always beneficial for FV, at Cape York solar storage when population size (N) was increased from 50 to 300, there was an only minimum change in the fitness value (with N=50, FV=2.05 and N=300, FV=1.98). Additionally, the higher population size is computationally inefficient, therefore for all other five solar farms, the population size is set to 300 to balance the FV with the algorithm computation time. With this SMA FS process, 17 meteorological predictors from the pool of 75 (predictor matrix: 20455×17) are selected for Barunggam solar farm, Cameby solar farm, and Chinchilla solar farm. Whereas for Cape York solar storage and Sun metals solar farm only 12 meteorological predictors (predictor matrix: 20455×12) are selected. Similarly, for the Clermont solar farm, only 13 meteorological predictors (predictor matrix: 20455×16) are selected. The predictors from the SMA feature selection process for the prediction of GSR for all six solar farms are shown in Table 3, along with the correlation matrix for predictors and the predictands (GSR) in Figure 8.

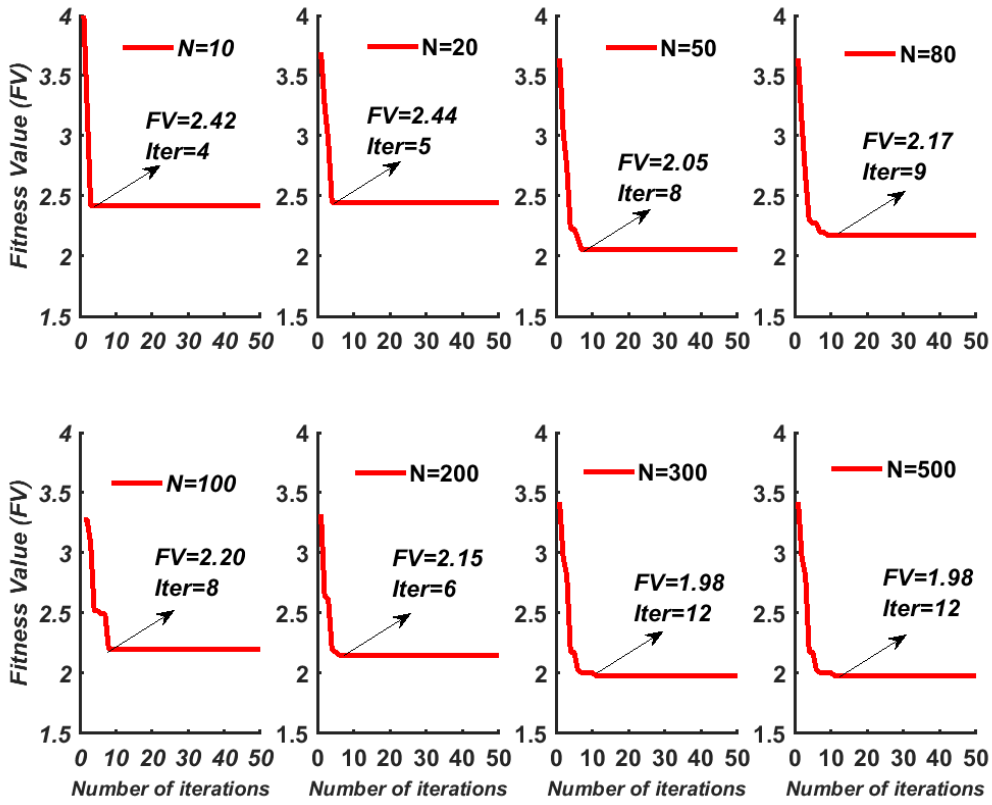


Figure 7: Convergence curve for SMA feature selection on predictors of Cape York solar storage.

Table 3: The selected predictor (input) variable for six solar energy farms using the Slime Mould Optmization (SMA) feature selection. For abbreviations, readers should refer to Table 2 (Example: hur_{1000} is the relative humidity 1000 hPa pressure height).

Barunggam Solar Farm	**Cameby Solar Farm**	**Cape York Solar Storage**	**Chinchilla Solar Farm**	**Clermont Solar Farm**	**Sun Metals Solar Farm**
Evap	Evap	Evap	Evap	Evap	Evap
RHmaxT	RHmaxT	RHmaxT	RHmaxT	RHmaxT	RHmaxT
ua_1000	ua_1000	hfss	ua_1000	hfss	hfss
hfls	hfls	hur_1000	hfls	hur_1000	Rain
hfss	hfss	Rain	hus_5000	T.Max	hur_1000
hus_5000	hus_5000	T.Max	hfss	Rain	hfls
wap_1000	ta_25000	hur_85000	wap_1000	ua_5000	T.Max
ta_25000	wap_1000	hus_1000	ta_25000	wap_1000	ua_5000
wap_85000	sfcWindmax	wap_1000	sfcWindmax	va_85000	va_50000
zg_1000	zg_1000	va_70000	zg_1000	RHminT	zg_5000
sfcWindmax	Rain	uas	wap_85000	wap_85000	wap_1000
ua_5000	RHminT	RHminT	Rain	zg_5000	hur_70000
RHminT	wap_85000		RHminT	va_50000	ta_25000
Rain	ua_10000		ua_5000	hfls	
T.Max	hur_1000		hur_1000	sfcWindmax	
va_25000	psl		va_25000	hus_5000	
hur_1000	va_25000		T.Max		

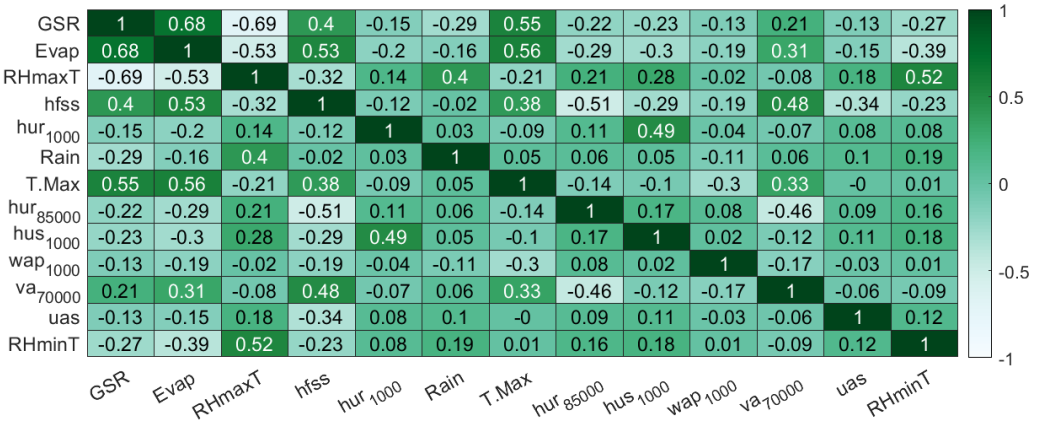


Figure 8: A Correlation matrix showing the correlation coefficient (r) of the predictor variables with the target variable (*GSR*) for Cape York solar storage. Note: -The variables' names are outlined in Table 2.

4.3. Data Partition

Finally, the SMA selected predictor matrix is merged with predictands (*GSR*) to get the input-target data for supervised machine learning. Before integrating data into the ML models, training, validation, and testing data are created to predict daily *GSR* at six solar farms of Queensland, Australia. The models are calibrated on the training set while the validation set doesn't participate in training and helps to tune the models during the model development phase. The test set is only used after a model has been trained (using train and validation sets), mostly for model evaluation. In this study, for training, 54 years of data are used (20089 data points), validation uses 20% of the data in the training set (4018 data points) and testing uses 1 year of data (365 data points).

4.4. Benchmark Model Development

A comparison of the SCLC model with five popular forecast models, such as CNN-LSTM, Deep Neural Network (DNN), Artificial neural network (GBM), SADE-ELM, and Random Forest Regression (RFR), was done to validate its efficacy. All proposed (Hybrid SCLC), as well as benchmark models, were built using Python under the framework of Keras 2.2.4 [86, 62] on TensorFlow 1.13.1 [87, 62]. The training process of all the models was conducted on a system that has the CPU type of Intel®Core™i7 with 32GB RAM.

4.5. Model Tuning

All the ML models have different hyperparameters that determine both the network structure (e.g., number of filters, neurons) and how the network models are trained (e.g., type of optimizer, activation function) [88]. As such, the performance of an ML model can vary greatly depending on its chosen set of hyperparameters therefore to achieve optimal performance hyperparameters should be selected cautiously [89]. In this study, a grid search method [90, 91] based on five-fold cross-validation was utilized to optimize all the hyperparameters in the model and was evaluated based on its average RMSE on the validation set for each set of hyperparameters. In grid search, all possible combinations of hyperparameters are tried for a dataset to find the best hyperparameter. Furthermore, during deep learning model (SCLC and CLSTM) training, rectified linear unit (ReLu) activation function is used in all except the last layer. ReLu performs better than sigmoid and hyperbolic tangent activation functions and does not have the vanishing gradient problem [92]. This study has also used Adam as the optimization algorithm with a constant learning rate of (lr) 0.001; decay rate $\beta_1 = 0.9$ and $\beta_2 = 0.9999$ and epsilon (ϵ) of 10^{-8}. Adam optimization is an adaptive learning rate optimization algorithm designed to train neural networks [93]. The name Adam comes from adaptive moment estimation [94] and uses a quadratic gradient to change the learning rate, as well as the momentum based on the moving average of the gradient. The Adam algorithm is memory-efficient, invariant to diagonal scaling of gradients, and well suited to problems with a large number of data [94]. Furthermore, this study also employed the following regularization technique when developing a robust deep learning model for *GSR* prediction.

- During training, the dropout technique was employed to prevent overfitting and enhance performance. The dropout involves dumping a certain number of neurons randomly on the network. The connections of the dropped neurons, therefore, are ignored [95, 96, 97].

- ReduceLROnPlateau regularization was employed to monitor the improvement of validation loss (root mean square error; RMSE), and in the case that no improvement is verified for a 'patience' number of 10 iterations, the learning rate (lr) is reduced at the factor of 0.2($lr_{new} = lr \times 0.2$). Consequently, when the learning process stagnates, this reducing strategy could be of significant benefit to the model [98, 99].

- The early stopping (es) regularisation method monitors the loss of the validation set and stop stops the training when the validation loss (root mean square error; RMSE) is no longer decreasing for a certain number of epochs [100, 101]. Hence, with es the training can be stopped when no further important improvements can be achieved or when the validation loss starts to increase due to overfitting [102, 103]. In this study, the training was stopped after the loss stopped decreasing for 15 consecutive epochs.

It should be also noted that in this study the es and ReduceLROnPlateau were not used along with grid search because the programming code did not permit to integrate them. Therefore, these regularization methods were only used during the training of the final model with optimal parameters. Tables 4 and 5 list the search space and optimized results for the hybrid SCLC as well as other benchmark models. Figure 9 illustrates that the training and validation losses of the hybrid SCLC model with optimum parameters (Cape York solar storage); both losses gradually decrease as the epoch increases, indicating the satisfactory performance of the SCLC training.

Table 4: (a) The objective deep learning hybrid model (SCLC) with other deep learning models (i.e., Convolution Neural Network integrated with Long-term short-term Memory), CLSTM, and Network Deep Neural Network (DNN).

Predictive Deep Learning Models	**Model Hyperparameters**	**Hyperparameter Selection**	**Barunggam Solar Farm**	**Cameby Solar Farm**	**Cape York Solar Storage**	**Chinchilla Solar Farm**	**Clermont Solar Farm**	**Sun Metals Solar Farm**
SCLC	Filter1 (for first CNN)	[50, 80,100,200]	100	200	50	80	50	100
	Filter 2 (for first CNN)	[40,50,60,70,80]	70	40	50	80	40	60
	Filter 3 (for first CNN)	[20,10,30,5]	30	30	20	30	30	30
	LSTM cell 1	[100]						
	Drop rate	[0,0.1,0.2]	0	0.1	0	0.2	0.1	0
	LSTM cell 2	[80]						
	LSTM cell 3	[50]						
	Filter1 (for second CNN)	[50, 80,100,200]	80	50	50	80	200	50
	Filter 2 (for second CNN)	[40,50,60,70,80]	70	40	60	60	80	40
	Filter 3 (for second CNN)	[20,10,30,5]	20	30	30	30	20	10
	Epochs	[300,400,500,600,700]	500	400	600	700	500	400
	Batch Size	[5,10,15,20,25,30]	10	15	5	20	25	5
CLSTM	Filter1	[50, 80,100,200]	80	50	100	100	80	80
	Filter 2	[40,50,60,70,80]	40	60	50	60	80	70
	Filter 3	[20,10,30,5]	10	5	30	10	20	30
	LSTM cell 1	[50, 60,100,200]	50	60	100	60	50	100
	LSTM cell 2	[40,50,60,70,130]	40	50	70	40	50	40
	LSTM cell 3	[20,10,30,5]	20	10	10	30	20	10
	LSTM Cell 4,5 and 6	[Fixed as 30,20,10]						
	Activation function	ReLu						
	Epochs	[300,400,500,600,700]	500	400	300	600	300	400
	Drop rate	[0,0.1,0.2]	0.1	0	0	0.1	0	0
	Batch Size	[5,10,15,20,25,30]	5	10	5	15	25	10
DNN	Hidden neuron 1	[100,200,300,400,50]	100	200	50	80	50	100
	Hidden neuron 2	[20,30,40,50,60,70]	70	40	50	80	40	60
	Hidden neuron 3	[10,20,30,40,50]	30	30	20	30	30	30
	Hidden neuron 4	[5,6,7,8,12,15,18]	12	15	15	12	5	8
	Activation function	ReLu						
	Epochs	[300,400,500,600,700]	400	400	600	500	700	300
	Drop rate	0.1						
	Batch Size	[5,10,15,20,25,30]	15	20	25	5	10	10

Table 5: (b) Comparison of machine learning models used in the study. Note that ReLU is Rectified Linear Units. Tansig, logsig and purelin refer to Hyperbolic tangent transfer function, Log-sigmoid transfer function, and linear transfer function respectively. LM, lbfgs, cgf, rp, br, scg refer to Levenberg-Marquardt, limited memory Broyden–Fletcher–Goldfarb–Shanno, Conjugate gradient backpropagation with Fletcher-Reeves, resilient backpropagation, Bayesian regulation backpropagation, and one-step secant backpropagation algorithm respectively.

Predictive Conventional Models	**Model Hyperparameters**	**Hyperparameter Selection**	**Barunggam Solar Farm**	**Cameby Solar Farm**	**Cape York Solar Storage**	**Chinchilla Solar Farm**	**Clermont Solar Farm**	**Sun Metals Solar Farm**
ANN	Hidden neuron	[10,20,30,40,60,80,100,200,300]	200	100	60	80	100	200
	Training Function	['trainlm', trainbfg, traingdx]	trainlm	trainlm	trainbfg	trainbfg	trainbfg	trainlm
	Activation function	[tansig, logsig, purlin]	tansig	tansig	logsig	logsig	logsig	tansig
SADE-ELM	Hidden neuron	[20,40,80,100]	80	100	80	40	80	100
	Activation function	[sig]						
	Crossover ratio	0.5						
	Population	100						
	Amplification factor	0.5						
	Maximum Generation	100						
RFR	The maximum depth of the tree.	[5,8,10,20,25]	10	20	10	10	25	10
	The number of trees in the forest.	[50,100,150,200]	100	150	50	150	100	150
	Minimum number of samples to split an internal node	[2,4,6,8,10]	8	8	6	10	6	8
	The number of features to consider when looking for the best split.	['auto', 'sqrt', 'log2']	auto	auto	auto	auto	auto	auto

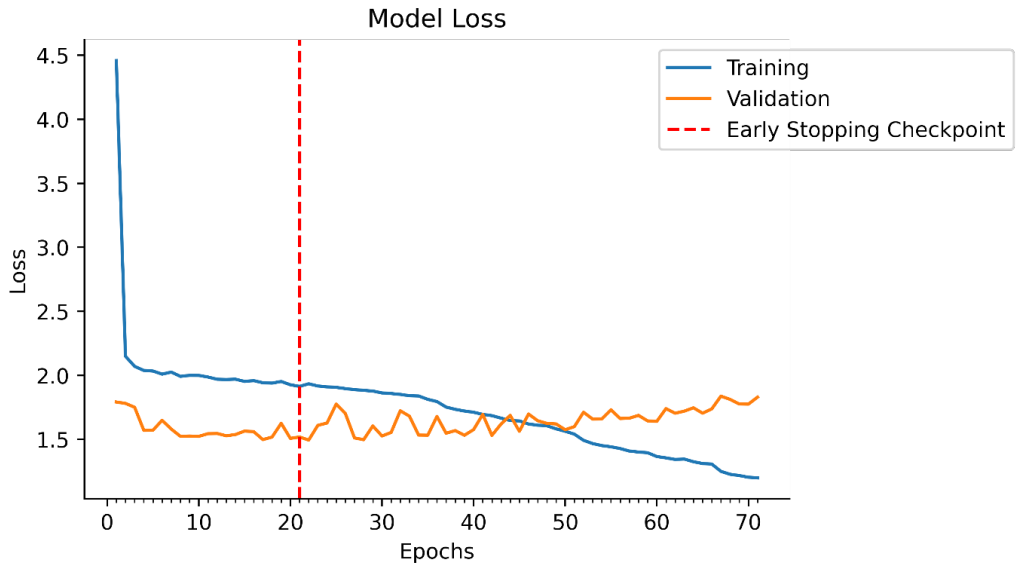


Figure 9: Training and validation loss (mean absolute error)) during the predictive model development phase for the prediction of *GSR*. The early stopping callback stops the model if there is no improvement in the loss for a certain number of predefined epochs.

4.6. Performance Evaluation Metrics

Statistical metrics based on earlier approaches [35, 38, 45, 122-128] were employed to assess the performance of the hybrid CXGBRFR model.

$$r = \frac{\sum_{i=1}^{n}(GSR^m - \langle GSR^m \rangle)(GSR^p - \langle GSR^p \rangle)}{\sqrt{\sum_{i=1}^{n}(GSR^m - \langle GSR^m \rangle)^2}\sqrt{\sum_{i=1}^{n}(GSR^p - \langle GSR^p \rangle)^2}} \tag{17}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(GSR^p - GSR^m)^2} \tag{18}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|GSR^p - GSR^m| \tag{19}$$

$$RRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(GSR^p - GSR^m)^2}}{\langle GSR^m \rangle} \tag{20}$$

$$RMAE = \frac{1}{n}\sum_{i=1}^{n}\frac{|GSR^p - GSR^m|}{GSR^p} \tag{21}$$

$$WI = 1 - \frac{\sum_{i=n}^{n}(GSR^m - GSR^p)^2}{\sum_{i=n}^{n}(|GSR^p - \langle GSR^m \rangle| + |GSR^m - \langle GSR^m \rangle|)^2} \tag{22}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(GSR^m - GSR^p)^2}{\sum_{i=1}^{n}(GSR^m - \langle GSR^m \rangle)^2} \tag{23}$$

$$LM = 1 - \frac{\sum_{i=1}^{n}|GSR^m - GSR^p|}{\sum_{i=1}^{n}|GSR^m - \langle GSR^m \rangle|} \tag{24}$$

$$E_{var} = 1 - \frac{\mathrm{Var}(GSR^m - GSR^p)}{\mathrm{Var}(GSR^m)} \tag{25}$$

$$SS = 1 - \frac{RMSE(p,x)}{RMSE(pr,x)} \tag{26}$$

$$RMSE_r = \frac{RMSE(p,x)}{RMSE(r,x)} \tag{27}$$

where GSR^m and GSR^p are the observed and predicted value of GSR, $\langle GSR^m \rangle$ and $\langle GSR^p \rangle$ are the observed and predicted mean of GSR, p stands for the model prediction, x for the observation, pr for perfect prediction (persistence), and r for the reference prediction. The persistence model considers that the solar radiation at $t+1$ is equal to the solar radiation at t. It assumes that the atmospheric conditions are stationary (clear sky condition).

For a better model performance,

- r can be in the range of -1 and $+1$, MAE, RMSE all range from 0 (perfect fit) to ∞ (the worst fit);
- RRMSE and RMAE ranges from 0% to 100% and model evaluation, a model's precision level is excellent if $RRMSE < 10\%$, good if $10\% < RRMSE < 20\%$, fair if $20\% < RRMSE < 30\%$, and poor if $RRMSE > 30\%$ [104].
- WI which is improvement to RMSE and MAE and overcomes the insensitivity issues as the differences between the observed and predicted values are not squared, ranges from 0 (the worst fit) to 1 (perfect fit) [105].
- NSE, compares the variance of observed and predicted GSR and ranges from $-\infty$ (the worst fit) to 1 (perfect fit) [106].
- LM, is a more robust metrics developed to address the limitations of both the WI and ENS [107] and the value ranges between 0 and 1 (ideal value).

- E_{var}; uses biased variance for explaining the fraction of variance and ranges from 0 to 1.

Furthermore, the overall model performance was ranked using the Global Performance Indicator (GPI) [108]. GPI was calculated using the six metrics.

$$\mathrm{GPI}_i = \sum_{j=1}^{6} \alpha_j (g_j - y_{ij}) \tag{28}$$

where α_j denotes the median of scaled values of statistical indicator j, equals to 1 for RMSE, MAE, MAPE, RRMSE and RRMSE ($j = 1, 2, 3, 4, 5$), -1 for r; g_j denotes the scaled value of the statistical indicator j for model i. Greater GPI value indicates the corresponding model has better performance. This study also evaluated the performance of the models using the Kling-Gupta Efficiency (KGE) [109] and Absolute Percentage Bias (APB; %) [110]. Mathematically, these metrics are stated as follows:

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\langle GSR^p \rangle}{\langle GSR^m \rangle} - 1\right)^2 + \left(\frac{CV_p}{CV_m}\right)^2} \tag{29}$$

$$APB = \frac{\sum_{i=1}^{n} (GSR^m - GSR^p) * 100)}{\sum_{i=1}^{n} GSR^m}, \tag{30}$$

where r is the correlation coeffecient, CV is the coefficient of variation, GSR^p refers to the predicted GSR ($\mathrm{MJm^{-2}day^{-1}}$), GSR^m is the measured GSR^p ($\mathrm{MJm^{-2}day^{-1}}$), $\langle GSR^m \rangle$ is the average value of the GSR^m, $\langle GSR^p \rangle$ is the average value of the GSR^p and finally n is the number of actual values.

Furthermore, this study also use the Promoting Percentage of: Absolute Percentage Bias (λ_{APB}), Mean Absolute Error (λ_{MAE}), and Root Mean Square Error (λ_{RMSE}) [111] to compare various models that have been used in the GSR prediction.

$$\lambda_{APB} = \left| \frac{APB_1 - APB_2}{APB_1} \right| \tag{31}$$

$$\lambda_{MAE} = \left| \frac{RMAE_1 - RMAE_2}{RMAE_1} \right| \tag{32}$$

$$\lambda_{RRMSE} = \left| \frac{RRMSE_1 - RRMSE_2}{RRMSE_1} \right| \tag{33}$$

where, APB_1, $RRMSE_1$ and $RMAE_1$ refers to the objective model (*i.e.*, SCLC) performance metrics and APB_2, $RRMSE_2$ and $RMAE_2$ refers to the benchmark model performance metrics.

Additionally, the performance to prediction direction of movement was measured by a Directional Symmetry (DS) as follows:

$$DS = \frac{1}{n}\sum_{t=2}^{n} d_t \times 100\% \tag{34}$$

where

$$d_t = \begin{cases} 1, & \text{if}(GSR_t^m - GSR_{t-1}^m)(GSR_t^p - GSR_{t-1}^m) > 0 \\ 0, & \text{otherwise} \end{cases} \tag{35}$$

An assessment criterion is known as the Diebold-Mariano (DM) test, the Harvey, Leybourne, and the Newbold (HLN) were used to test the statistical significance of all models in this study, these statistical tests are done to further evaluate the propsoed model prediction performance and the directional prediction performance from a statistical standpoint. When comparing such models, the alternative model is expected to outperform the comparative model when the DM statistics is greater than 0 and the HLN statistics is greater than 0 too. The key steps in implementing the DM and HLN tests are defined in previous literature [112, 113, 114].

5. Results and Discussion

The deep hybrid SCLC model used for GSR prediction was able to produce a high r-value and lower MAE and RMSE values for the Barunggam Solar Farm ($r \approx 0.930$, RMSE ≈ 2.338 MJm^{-2}day^{-1}, MAE ≈ 1.69 MJm^{-2}day^{-1}). This contrasted the results of the deep learning model CNN-LSTM model ($r \approx 0.916$, RMSE ≈ 2.538 MJm^{-2}day^{-1}, MAE ≈ 1.911 MJm^{-2}day^{-1}) and the DNN model ($r \approx 0.914$, RMSE ≈ 2.633 MJm^{-2}day^{-1}, MAE ≈ 1.946 MJm^{-2}day^{-1}). Likewise, for the conventional ML models (ANN, SADE-ELM, and RFR) the r-value is lower than that of the SCLC model, RMSE, and MAE, both metrics are higher than that of the SCLC model. Additionally, the SCLC model of the other five solar farms produced substantially healthier *GSR* prediction than other DL as well as conventional ML models (Table 6 and Figure 10). This result shows that the SCLC is a possible

choice to be implemented as a well-designed forecasting approach for *GSR* predictions in comparison to DL-based models (CLSTM and DNN), and the conventional ML models.

Table 6: The performance of the deep hybrid SCLC predictive model in respect to its counterpart models in terms of correlation coefficient (r) and root mean square error (*RMSE*, MJm-2day-1) in the model's testing phase. The Objective model SCLC is blue bold-faced.

Predictive Models	**Barunggam Solar Farm**		**Cameby Solar Farm**		**Cape York Solar Storage**		**Chinchilla Solar Farm**		**Clermont Solar Farm**		**Sun Metals Solar Farm**	
	r	*RMSE*	*r*	*RMSE*	*r*	*RMSE*	*r*	*RMSE*	*r*	*RMSE*	*r*	*RMSE*
SCLC	**0.930**	**2.338**	**0.930**	**2.386**	**0.880**	**2.207**	**0.933**	**2.314**	**0.905**	**2.502**	**0.935**	**2.792**
CLSTM	0.916	2.538	0.910	2.653	0.863	2.441	0.910	2.638	0.881	2.761	0.920	3.091
SADE-ELM	0.895	2.849	0.895	2.911	0.853	2.539	0.893	2.833	0.878	2.793	0.920	3.106
DNN	0.914	2.633	0.901	2.772	0.881	2.384	0.898	2.787	0.891	2.647	0.895	3.389
ANN	0.899	3.815	0.910	2.705	0.864	2.507	0.833	3.471	0.881	2.778	0.892	3.627
RFR	0.681	5.033	0.694	5.492	0.578	4.515	0.516	6.135	0.686	4.574	0.839	4.462

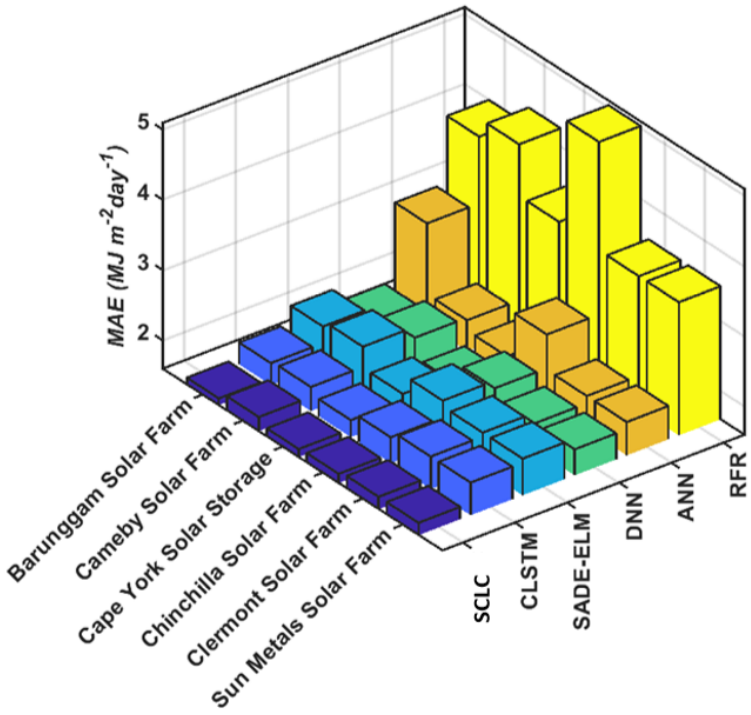


Figure 10: Evaluation of Deep hybrid SCLC predictive model, in respect to its counterpart models, as measured by mean absolute error (MAE, MJm-2 day-1) in the testing phase. (Note: Names for each model are provided in Table 3 and Tables 4 and 5).

In Table 7, to compare the SCLC model against CLSTM, DNN, ANN, SADE-ELM, and RFR models, we utilized multiple criteria based on WI and NSE. According to these model penalisation metrics, the results produced by the SCLC model for the case of Barunggam Solar Farm yielded a value of (WI $\approx$ 0.926, NSE $\approx$ 0.862), followed by those for a CLSTM (WI $\approx$ 0.913, NSE $\approx$ 0.8377), the DNN model (WI $\approx$ 0.904, NSE $\approx$ 0.826), the ANN model (WI $\approx$ 0.723, NSE $\approx$ 0.6331), the SADE-ELM model (WI $\approx$ 0.886, NSE $\approx$ 0.795) and the RFR model (WI $\approx$ 0.564, NSE $\approx$ 0.390). Similarly, other model penalization metrics like LM and Evar (Table 9) were also utilised and for the Barunggam Solar Farm, the SCLC model with high LM and Evar (LM $\approx$ 0.674, Evar $\approx$ 0.864) outperform all other DL models as well as the conventional ML models. Furthermore, the SCLC model of other five solar farms (Cameby Solar Farm, Cape York Solar Storage, Chinchilla Solar Farm, Clermont Solar Farm, and Sun Metals Solar Farm) performed

substantially better. Compared to r, RMSE, and MAE, these higher-order metrics demonstrate that the deep hybrid SCLC model has superior predictive abilities to deliver accurate prediction of *GSR*.

Table 7: As per Table 5 but measured in terms of the Willmott's Index (*WI*) and Nash–Sutcliffe coefficients (*NSE*).

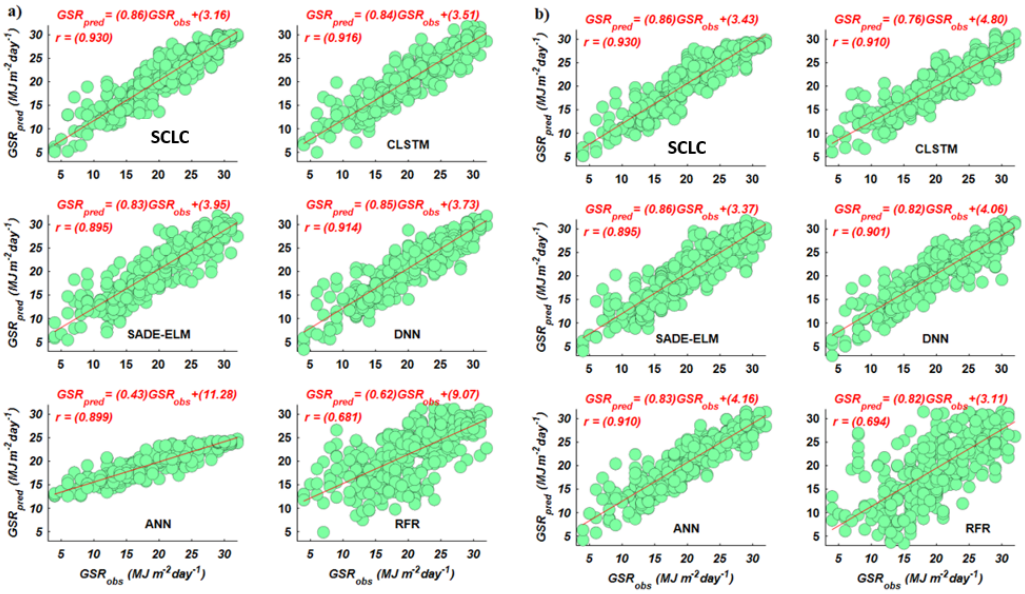
Predictive Models	**Barunggam Solar Farm**		**Cameby Solar Farm**		**Cape York Solar Storage**		**Chinchilla Solar Farm**		**Clermont Solar Farm**		**Sun Metals Solar Farm**	
	WI	*NSE*	*WI*	*NSE*	*WI*	*NSE*	*WI*	*NSE*	*WI*	*NSE*	*WI*	*NSE*
SCLC	**0.926**	**0.862**	**0.922**	**0.858**	**0.873**	**0.769**	**0.928**	**0.865**	**0.906**	**0.816**	**0.874**	**0.756**
CLSTM	0.913	0.837	0.900	0.822	0.830	0.724	0.903	0.824	0.886	0.775	0.820	0.703
SADE-ELM	0.886	0.795	0.888	0.788	0.808	0.704	0.891	0.796	0.888	0.770	0.814	0.701
DNN	0.904	0.826	0.897	0.807	0.829	0.741	0.896	0.803	0.900	0.793	0.808	0.640
ANN	0.723	0.631	0.896	0.818	0.795	0.710	0.821	0.694	0.883	0.772	0.700	0.592
RFR	0.564	0.390	0.690	0.245	0.443	0.134	0.438	0.062	0.606	0.407	0.727	0.376

Table 8: As per Table 5 but measured in terms of the Legates and McCabes index (LM) and explained variance score (E_{var}).

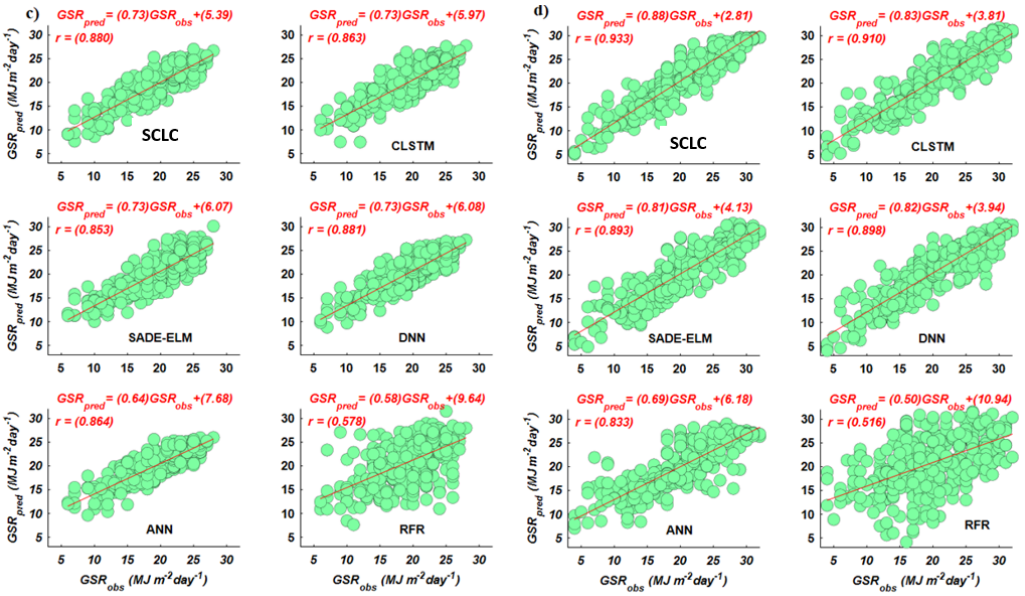
Predictive Models	**Barunggam Solar Farm**		**Cameby Solar Farm**		**Cape York Solar Storage**		**Chinchilla Solar Farm**		**Clermont Solar Farm**		**Sun Metals Solar Farm**	
	LM	E_{var}	LM	E_{var}	LM	E_{var}	LM	E_{var}	LM	E_{var}	LM	E_{var}
SCLC	**0.674**	**0.864**	**0.653**	**0.865**	**0.539**	**0.772**	**0.665**	**0.870**	**0.623**	**0.819**	**0.525**	**0.757**
CLSTM	0.632	0.839	0.626	0.822	0.505	0.744	0.620	0.828	0.562	0.776	0.491	0.710
SADE-ELM	0.586	0.800	0.570	0.796	0.481	0.728	0.574	0.798	0.549	0.771	0.490	0.707
DNN	0.625	0.836	0.600	0.811	0.515	0.774	0.595	0.807	0.582	0.793	0.444	0.641
ANN	0.413	0.631	0.607	0.828	0.490	0.733	0.502	0.694	0.559	0.774	0.389	0.599
RFR	0.230	0.411	0.179	0.246	0.052	0.150	0.027	0.064	0.258	0.424	0.223	0.376

To overcome the limitation of objective metrics in *GSR* prediction, diagnostic plots were used to improve the suitability of the deep hybrid SCLC model. Figure 10 shows scatterplots of the observed and predicted *GSR* during the testing phase from DL models as well as conventional ML models at all six solar farms. For better illustration, both the linear fit equation and the Correlation coefficient (r) [Range=(0,+1);Idealvalue=+1], which gives a measure on the adequacy of the model [115], have been included. As can also be seen by the scatter plot, the SCLC model performs the best since the scatter points are close to the y = mx + C line, in comparison to the other models which are scattered farther from the y = mx + C line. The scatterplot concurs with the results of r, RMSE, MAE, LM, NSE, WI, and Evar metrics as well.

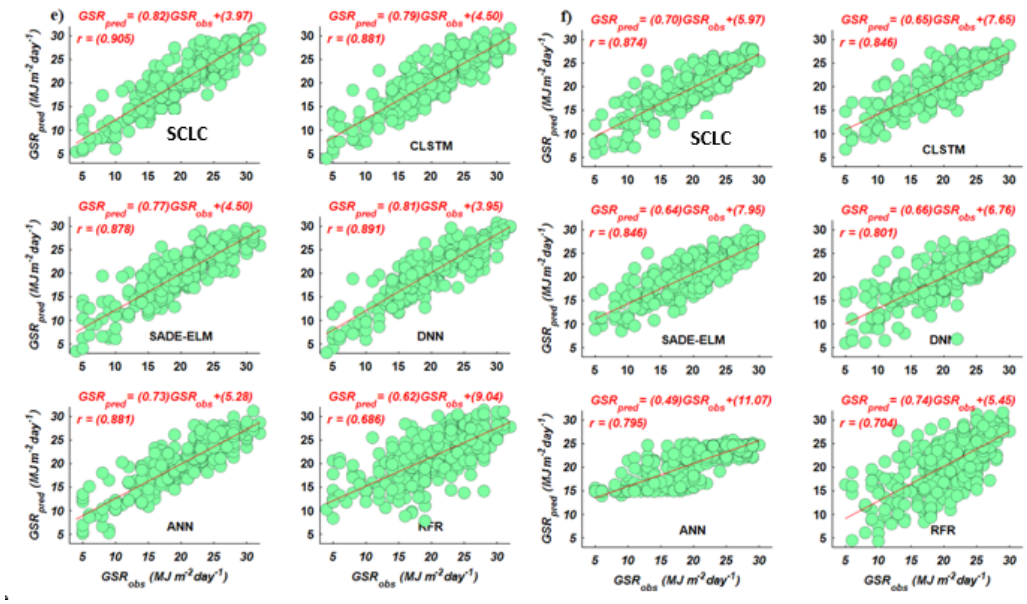
To compare the model performances in the prediction of *GSR* at the sites that differ geographically, physically, and climatically, alternative relative metrics like RRMSE and RMAE were used. Comparing the models (Table 9) showed that the deep hybrid SCLC model had the lowest RRMSE and RMAE when compared to the CLSTM, DNN, ANN, SADE- ELM and RFR approaches for all six solar farms. SCLC model at Cape York Solar Storage produces the lowest (RRMSE $\approx$ 11.46%, RMAE$\approx$ 10.37%) relative metrics compared to other DL models as well as conventional ML models. In all six sites, the deep hybrid SCLC model resulted in the lowest values of both RRMSE and RMAE, and they were lower than those of other comparative models, indicating that the SCLC is undoubtedly the best option.


a)
GSR_pred = (0.86)GSR_obs +(3.16)
r = (0.930)
SCLC
GSR_pred = (0.84)GSR_obs +(3.51)
r = (0.916)
CLSTM
GSR_pred = (0.83)GSR_obs +(3.95)
r = (0.895)
SADE-ELM
GSR_pred = (0.85)GSR_obs +(3.73)
r = (0.914)
DNN
GSR_pred = (0.43)GSR_obs +(11.28)
r = (0.899)
ANN
GSR_pred = (0.62)GSR_obs +(9.07)
r = (0.681)
RFR
b)
GSR_pred = (0.86)GSR_obs +(3.43)
r = (0.930)
SCLC
GSR_pred = (0.76)GSR_obs +(4.80)
r = (0.910)
CLSTM
GSR_pred = (0.86)GSR_obs +(3.37)
r = (0.895)
SADE-ELM
GSR_pred = (0.82)GSR_obs +(4.06)
r = (0.901)
DNN
GSR_pred = (0.83)GSR_obs +(4.16)
r = (0.910)
ANN
GSR_pred = (0.82)GSR_obs +(3.11)
r = (0.694)
RFR
GSR_pred (MJ m-2 day-1)
GSR_obs (MJ m-2 day-1)

(a)


c)
GSR_pred = (0.73)GSR_obs +(5.39)
r = (0.880)
SCLC
GSR_pred = (0.73)GSR_obs +(5.97)
r = (0.863)
CLSTM
GSR_pred = (0.73)GSR_obs +(6.07)
r = (0.853)
SADE-ELM
GSR_pred = (0.73)GSR_obs +(6.08)
r = (0.881)
DNN
GSR_pred = (0.64)GSR_obs +(7.68)
r = (0.864)
ANN
GSR_pred = (0.58)GSR_obs +(9.64)
r = (0.578)
RFR
d)
GSR_pred = (0.88)GSR_obs +(2.81)
r = (0.933)
SCLC
GSR_pred = (0.83)GSR_obs +(3.81)
r = (0.910)
CLSTM
GSR_pred = (0.81)GSR_obs +(4.13)
r = (0.893)
SADE-ELM
GSR_pred = (0.82)GSR_obs +(3.94)
r = (0.898)
DNN
GSR_pred = (0.69)GSR_obs +(6.18)
r = (0.833)
ANN
GSR_pred = (0.50)GSR_obs +(10.94)
r = (0.516)
RFR
GSR_pred (MJ m-2 day-1)
GSR_obs (MJ m-2 day-1)

(b)



(c)

Figure 11: Scatter plots of the observed (GSRobs) and predicted (GSRpred) daily global solar radiation for all tested regions in Six solar farms of Queensland. a) Barunggam Solar Farm, b) Cameby Solar Farm, c) Cape York Solar Storage, d) Chinchilla Solar Farm, e) Clermont Solar Farm and f) Sun Metals Solar Farm. (Note: Line in red is the least-squares fit line ($y = mx + c$) to the respective scatter plots, where y is the predicted GSR and x is the observed GSR. Names for each model are provided in Table 3 and Tables 4 and 5).

Table 9: As per Table 5 but measured in terms of the relative root mean square error ($RRMSE$, %) and relative mean absolute error ($RMAE$, %) in the testing phase.

Predictive Models	**Barunggam Solar Farm**		**Cameby Solar Farm**		**Cape York Solar Storage**		**Chinchilla Solar Farm**		**Clermont Solar Farm**		**Sun Metals Solar Farm**	
	RRMSE	*RMAE*	*RRMSE*	*RMAE*	*RRMSE*	*RMAE*	*RRMSE*	*RMAE*	*RRMSE*	*RMAE*	*RRMSE*	*RMAE*
SCLC	**11.66%**	**10.71%**	**11.84%**	**11.12%**	**11.46%**	**10.37%**	**11.56%**	**10.61%**	**12.17%**	**10.85%**	**13.67%**	**13.12%**
CLSTM	12.65%	12.28%	13.17%	12.68%	12.67%	11.61%	13.17%	12.34%	13.43%	12.54%	15.13%	14.96%
SADE-ELM	14.20%	13.75%	14.45%	13.59%	13.18%	12.23%	14.15%	13.61%	13.59%	12.92%	15.20%	15.41%
DNN	13.13%	12.15%	13.76%	12.49%	12.37%	11.52%	13.92%	12.55%	12.88%	11.58%	16.59%	16.11%
ANN	19.02%	20.09%	13.43%	13.18%	13.01%	12.41%	17.33%	16.77%	13.52%	12.58%	17.75%	19.10%
RFR	25.09%	26.83%	27.27%	26.86%	23.44%	21.53%	30.63%	32.18%	22.25%	23.10%	21.84%	21.40%

The predictability of the deep hybrid SCLC model is evaluated by comparing Promoting Percentages, presented via incremental performance (λ) of the objective model over competing approaches, where, for example, λ=RMAECLC-RMAECLSTM, is evaluated to estimate the difference in the relative mean absolute error of SCLC and CLSTM model. Table 10 compares the deep hybrid SCLC model with the other models tested during the testing phase. Deep hybrid SCLC performs better than DL models, CLSTM and DNN, and other conventional ML models ANN, SADE-ELM, and RFR.

Table 10: The promoting percentage of comparison, against the objective (i.e., SCLC) predictive model where the magnitude of the percentage value indicates the improvement of the objective over the benchmark model. Note that λ_{RMAE} is the promoting percentages of relative mean absolute error, λ_{RRMSE} is the promoting percentages of relative root mean square error, and λ_{APB} is the promoting percentages of absolute percentage bias.

Predictive Models	**Barunggam Solar Farm**			**Cameby Solar Farm**			**Cape York Solar Storage**			**Chinchilla Solar Farm**			**Clermont Solar Farm**			**Sun Metals Solar Farm**		
	λ_{RRMSE}	λ_{RMAE}	λ_{APB}	λ_{RRMSE}	λ_{RMAE}	λ_{APB}	λ_{RRMSE}	λ_{RMAE}	λ_{APB}	λ_{RRMSE}	λ_{RMAE}	λ_{APB}	λ_{RRMSE}	λ_{RMAE}	λ_{APB}	λ_{RRMSE}	λ_{RMAE}	λ_{APB}
CLSTM	9%	15%	13%	11%	14%	8%	11%	12%	7%	14%	16%	14%	10%	16%	16%	11%	14%	7%
SADE-ELM	22%	28%	27%	22%	22%	24%	15%	18%	13%	22%	28%	27%	12%	19%	19%	11%	17%	7%
DNN	13%	13%	15%	16%	12%	15%	8%	11%	5%	20%	18%	21%	6%	7%	11%	21%	23%	17%
ANN	63%	88%	80%	13%	18%	13%	14%	20%	11%	50%	58%	49%	11%	16%	17%	30%	46%	29%
RFR	115%	151%	136%	130%	141%	137%	105%	108%	106%	165%	203%	191%	83%	113%	97%	60%	63%	64%

A graphical analysis of model performance is as important as numerically evaluating the model. Figure 12 shows the boxplots of the deep hybrid SCLC model and other comparative DL as well as conventional ML models. As shown in the figure, the + symbols represent the outliers of the extreme absolute prediction error ($|PE| = GSR_{obs-} - GSRpred$) of the testing data, along with their upper quartile, median, and lower quartile. The distribution of the —PE— error acquired by the deep hybrid SCLC model for all sites is confirmed to exhibit a much smaller quartile followed by the standalone SADE-ELM, CLSTM, DNN, ANN, and RFR. Additionally, the kernel density estimate (KDE) plots of the standardized residuals were also plotted in Figure 13 to get a clearer picture of the residual distributions. The KDE plot of the standardized residuals for the SCLC model is close to the standard normal. We have not performed any correlation tests, but with such a large sample size, a hypothesis of correlated residuals is unlikely to be rejected. Hence, the box plot (Figure 10) and KDE plot of standardized residuals (Figure 11) further confirm SCLC's superior accuracy in *GSR* prediction compared to other competing models.

To broadly gauge the efficiency, a comprehensive and unbiased assessment of models is carried out by plotting a Taylor graph [116]. Figure 14 illustrates the statistical association between predicted and actual *GSR* based on r and standard deviation. By comparing r to standard deviation, it is shown that RFR, ANN, and SADE-ELM are not proper as their r to standard deviation was extremely far from the observed *GSR*, whereas deep learning model DNN and CLSTM overlap and are closer to observation. The SCLC model closely matched the actual *GSR* approving the prediction was better. To provide further insight into the prediction capability of the proposed modelling systems for *GSR* prediction, Figure 15 shows the plot for KGE, APB, and GPI. With high KGE (≈ 0.888) and low APB (≈ 9.035), the performance of the deep hybrid SCLC model far exceeds that of the counterpart models. Furthermore, the ranking of models is done according to their prediction efficiency using the GPI metrics. The GPI varies from -7.172 to 1.199 (Figure 15(b)). The highest value of GPI of 1.199 is for the proposed deep hybrid SCLC model, further cementing the advanced modelling capabilities of the proposed deep hybrid SCLC model in *GSR* prediction.

Additionally, statistical test DM, HLN, and DS were used to validate whether the prediction generated by the deep hybrid SCLC model is significantly more accurate than the prediction from other comparative models. Table 11 and Table 12 below show the statistics of the DM and HLN test

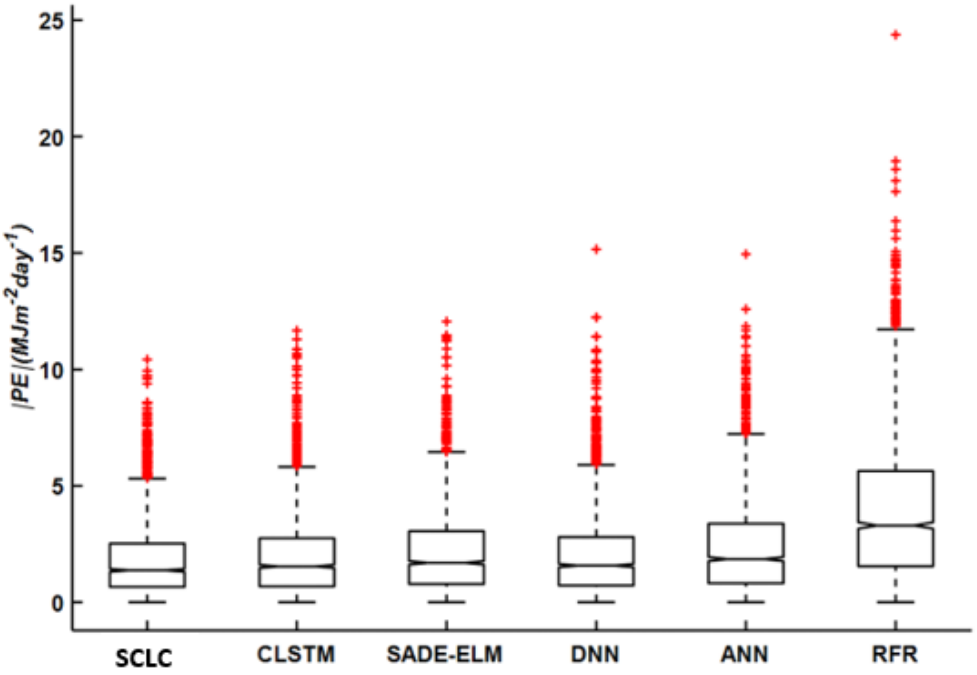


Figure 12: Box plots of the prediction error (PE) generated by prediction models during the testing phase for daily *GSR* prediction. (Note: Names for each model are provided in Table 3 and Tables 4 and 5).

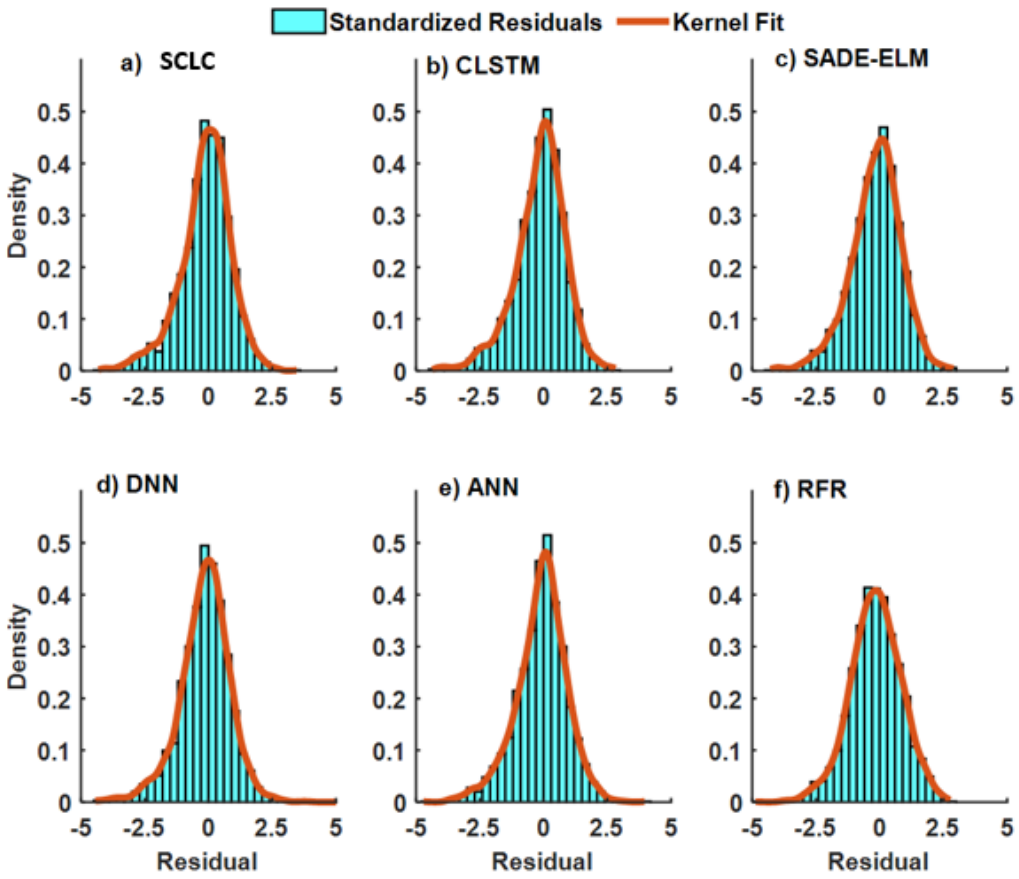


Figure 13: Marginal kernel density estimate for the standardized residuals of the model, along with a standard normal density in red.

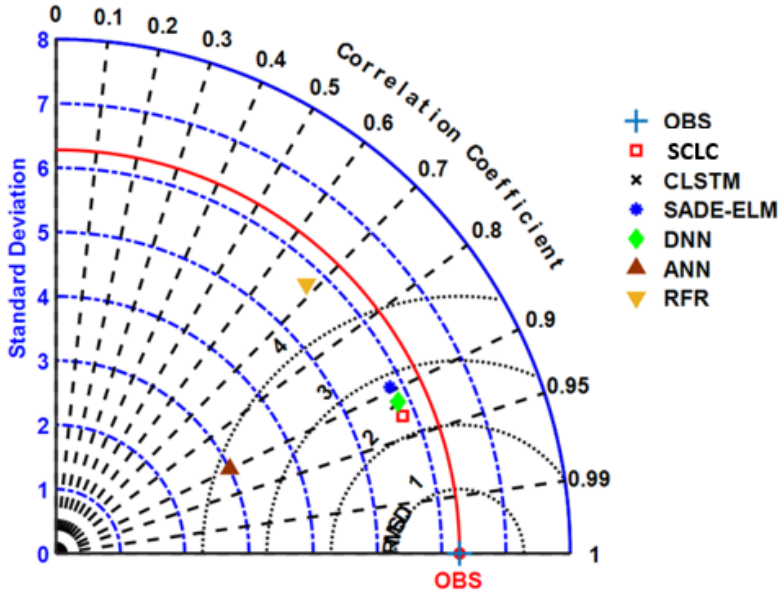


Figure 14: Taylor plots for the *GSR* prediction models.

results for all models. The models in the column of the table are compared with the model in the rows, and if the result is positive, the model in the column outperforms the one in the row; on the contrary, if it is negative, then the one in the row is superior. Similarly, Figure 16 shows that deep hybrid SCLC model DS (*i.e.*, directional prediction accuracy) is greater than the other five models, with an average of 60.86 %. Congruent with previous findings, DM, HLN, and DS test provides consistent results, which indicate that deep hybrid SCLC predicts *GSR* more accurately than other models.

Moreover, the RMSE of the deep hybrid as well as the comparative model prediction was compared with RMSE of the model using only clear-sky index persistence [117], thereby providing prediction skill or skill score (SS). In addition to this, the comparison of the deep hybrid SCLC model with other comparative models was done using RMSE ratio (RMSEr) [118]. All comparative models have a significantly lower SS and RMSEss than the deep hybrid SCLC model (Table 12 and Table 14).

To further evaluate the model performance for practical acceptance, data of each site are divided into four seasons, and simulations were conducted for the models. Figure 17 presents the model WI, NSE, KGE, RRMSE, RMAE,

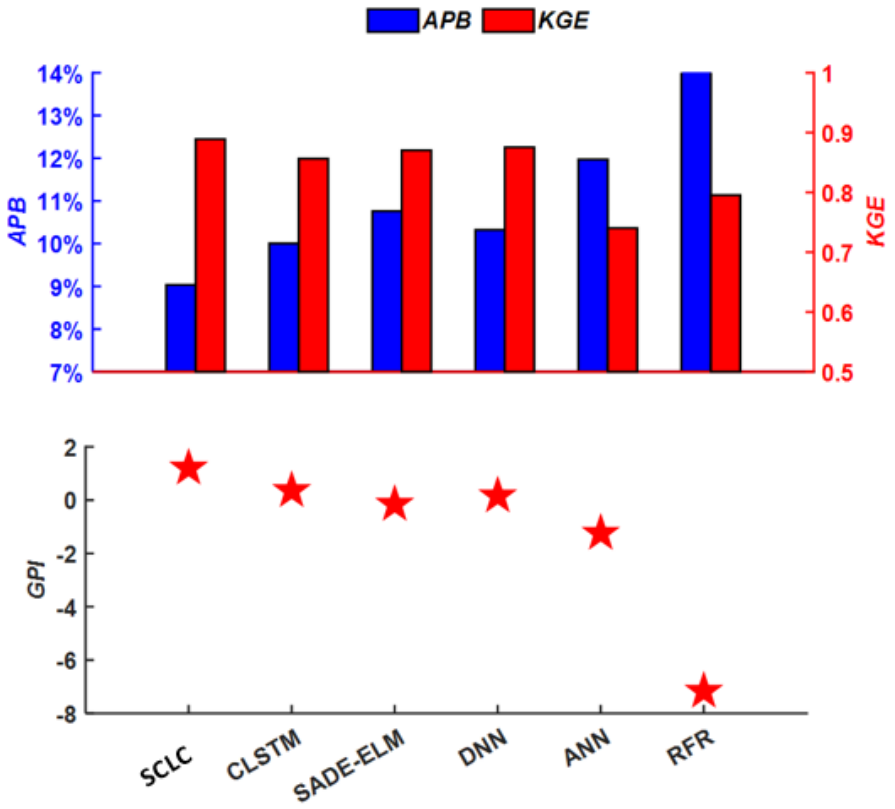


Figure 15: a) Bar chart showing a comparison of the SCLC model in terms of their absolute percentage bias (APB, %) and the Kling–Gupta efficiency (KGE) in the testing phase. b) Global performance indicator (GPI) of SCLC model compared with other artificial intelligence-based models. (Note: Names for each model are provided in Table 3 and Tables 4 and 5).

Table 11: The Diebold-Mariano (DM) test statistics (The column of the table is compared with the rows, and if the result is positive, the model in the column outperforms the one in the row; on the contrary, if it is negative, then the one in the row is superior).

Predictive Models	**SCLC**	**CLSTM**	**SADE-ELM**	**DNN**	**ANN**	**RFR**
SCLC		**3.070**	**2.971**	**3.726**	**3.897**	**7.002**
CLSTM			0.225	1.691	2.910	5.545
SADE-ELM				1.728	2.993	5.694
DNN					1.249	4.646
ANN						2.993

Table 12: Statistics of Harvey, Leybourne, and Newbold test. The column of the table is compared with the rows, and if the result is positive, the model in the column outperforms the one in the row; on the contrary, if it is negative, then the one in the row is superior.

Predictive Models	**SCLC**	**CLSTM**	**SADE-ELM**	**DNN**	**ANN**	**RFR**
SCLC		**3.216**	**3.112**	**3.903**	**4.083**	**7.335**
CLSTM			0.236	1.771	3.048	5.809
SADE-ELM				1.810	3.136	5.965
DNN					1.308	4.867
ANN						3.135

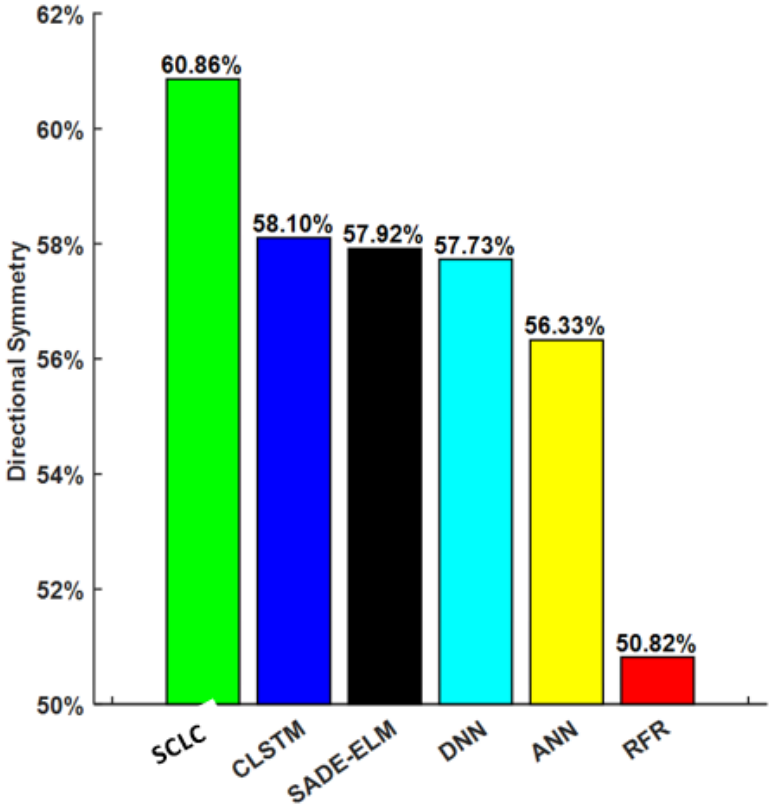


Figure 16: Performance comparison of deep hybrid SCLC model compared to other models under study in terms of directional symmetry (DS) criteria. (Note: Names for each model are provided in Table 3 and Tables 4 and 5).

Table 13: Skill Score Metric (SS) for deep hybrid SCLC as well as other deep learning and comparative models compared with persistence models in the testing phase. Note: The persistence model considers that the solar radiation at $t+1$ is equal to the solar radiation at t. It assumes that the atmospheric conditions are stationary (clear sky condition).

Solar Energy Farms	**SCLC**	**CLSTM**	**SADE-ELM**	**DNN**	**ANN**	**RFR**
Barunggam Solar	**0.728**	0.679	0.596	0.655	0.276	0.261
Cameby Solar	**0.719**	0.653	0.582	0.621	0.639	0.488
Cape York Solar Storage	**0.582**	0.489	0.446	0.512	0.460	0.450
Chinchilla Solar	**0.731**	0.651	0.597	0.610	0.395	0.352
Clermont Solar	**0.697**	0.631	0.623	0.661	0.627	0.215
Sun Metals Solar	**0.613**	0.525	0.521	0.429	0.346	0.315

Table 14: The performance of the SCLC model with comparative benchmark models in the test period measured by the ratio of root mean square error (RMSEss). The column of the table is compared with the rows, and if the result is lower than 1, the model in the row outperforms the one in the column; on the contrary, if it is greater than 1, then the one in the column is superior.

Predictive Models	**SCLC**	**CLSTM**	**SADE-ELM**	**DNN**	**ANN**	**RFR**
SCLC		**1.226**	**1.238**	**1.474**	**1.687**	**2.554**
CLSTM			1.010	1.202	1.377	2.084
SADE-ELM				1.191	1.363	2.064
DNN					1.145	1.733
ANN						1.514

and APB for four seasons. Concurrent with previous deductions, the deep hybrid SCLC model shows the best performance with lower RRMSE, RMAE, and APB (spring, summer, autumn, and winter) and higher WI, NSE, and KGE compared to DNN, DBN, ANN, and MARS model. Additionally, the deep hybrid SCLC model produces the lower RMSE for the spring season ($\approx$ 2.171 $MJm^{-2}day^{-1}$), followed by Autumn ($\approx$ 2.334 $MJm^{-2}day^{-1}$), Summer ($\approx$ 2.451 $MJm^{-2}day^{-1}$), and Winter ($\approx$ 2.734 $MJm^{-2}day^{-1}$) (Figure 18). Hence, it can be contended that the deep hybrid SCLC model can be deemed suitable for seasonal *GSR* prediction.

6. Conclusion and Discusion

The goal of this research was to develop a new deep learning-based hybrid model that can be adopted to simulate the *GSR* across six solar farms in Queensland, Australia by an ordered integration of the CNN, LSTM, and finally, another CNN algorithm thus making the overall hybrid SCLC-based predictive model. The Slime Mould Algorithm (SMA) feature selection process was implemented to screen out the most optimal predictive features (in terms of best input variables) from global climate model (GCM) meteorological data sets, and ground-based observation data sets. The CNN network was employed to extract the spatial information from feature-selected input variables, and those extracted information from CNN was then used as potential inputs for the LSTM predictive network. Through LSTM, temporal features were extracted, and the resulting data were input to the layer of the second CNN model. Finally, *GSR* prediction was done via a fully connected (MLP) dense layer. To validate the proposed SCLC prediction model, five different well-established AI models (*i.e.*, CLSTM, DNN, ANN, SADE-ELM, RFR) were implemented.

Using statistical metrics and diagnostic plots, the resultant SCLC model has been validated for effectiveness and reliability. In terms of different assessments, including MAE, RMSE, and r, with high r and low RMSE and MAE, SCLC presents considerable enhancements over the comparative models for all six solar farms. Furthermore, compared with the other DL and conventional ML models, the deep hybrid model (SCLC) can not only significantly improve the accuracy of *GSR* prediction but also ensure the highest prediction accuracy at different sites. In detail, compared with CLSTM, DNN, ANN, SADE-ELM, and RFR, the RRMSE of the proposed deep hybrid model SCLC is improved by 9 %, 13 %, 63 %, 22 %, and 115 %, respec-

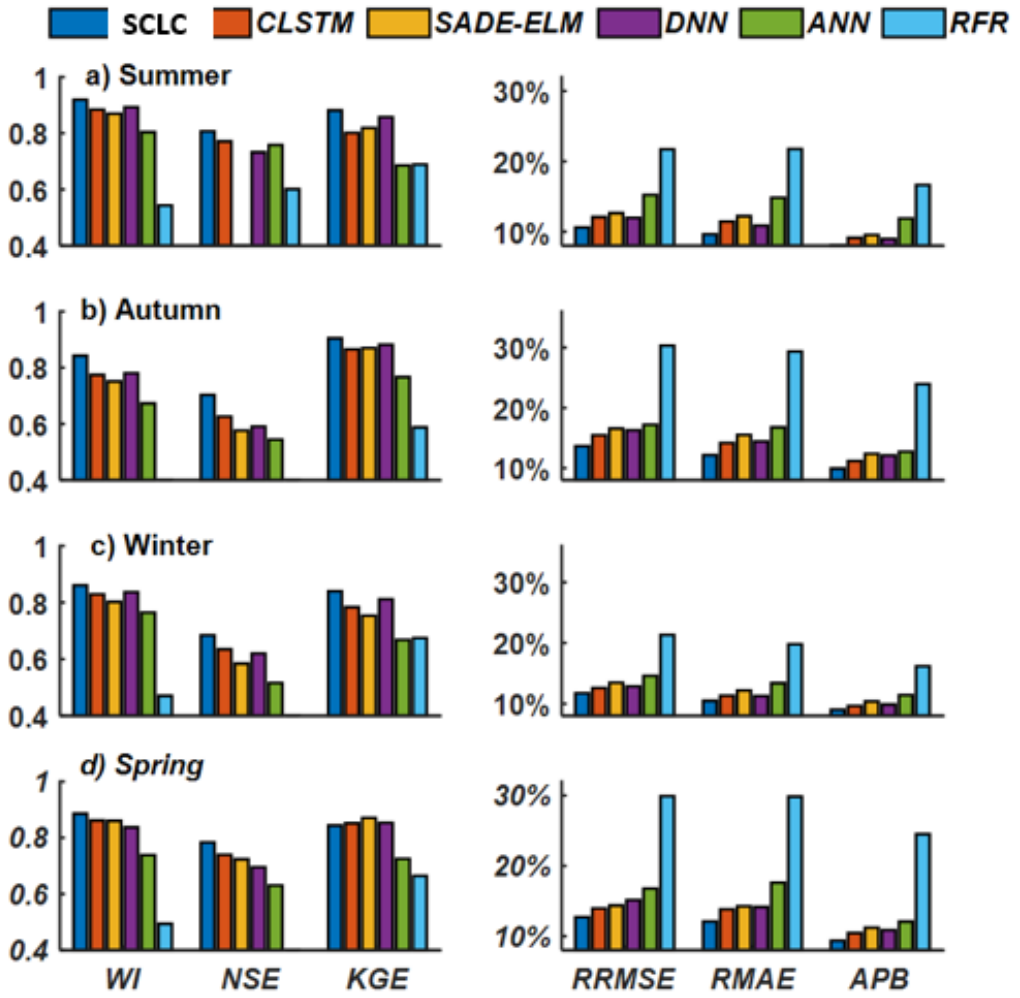


Figure 17: Seasonal performance evaluation of deep hybrid SCLC model compared to other artificial intelligence-based models in terms of Willmott's Index (WI), Nash–Sutcliffe Coefficient (NSE), Kling Gupta efficiency (KGE), relative root mean square error (RRMSE, %), Relative mean absolute error (RMAE, %) and absolute percentage bias (APB, %). a) Summer, b) Autumn, c) Winter, and d) Spring. (Note: Names for each model are provided in Table 3 and Tables 4 and 5).

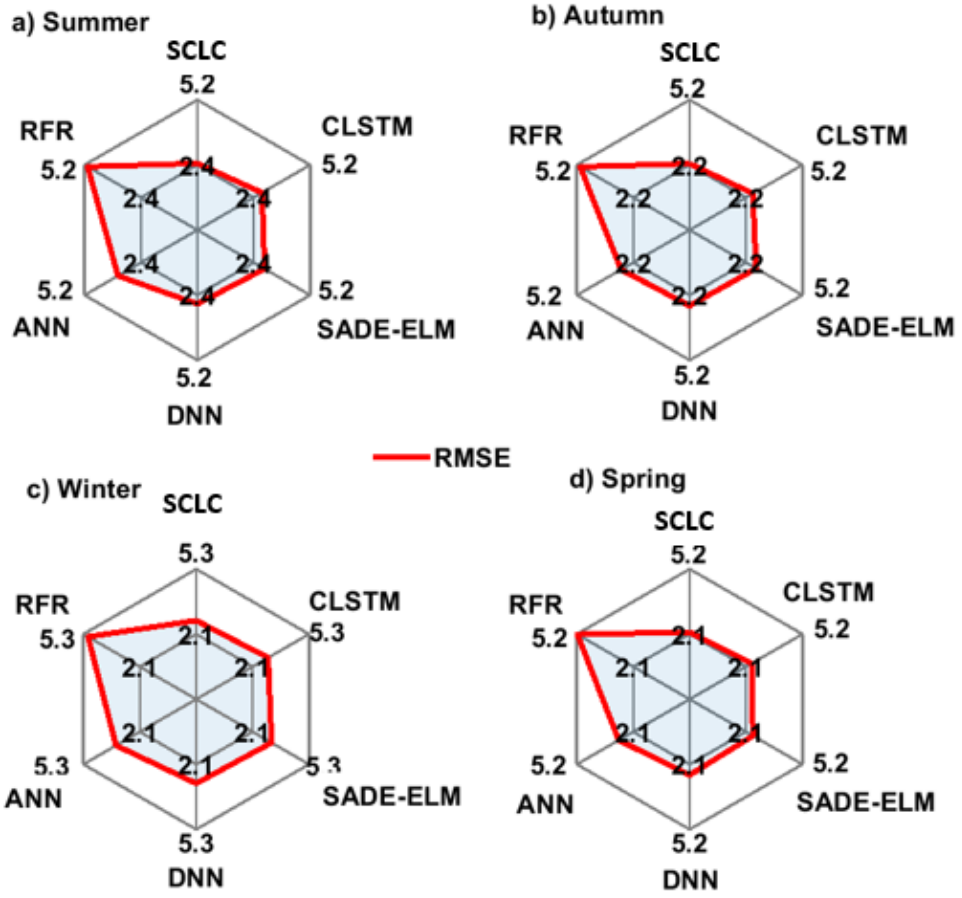


Figure 18: Seasonal performance evaluation of deep hybrid SCLC model compared to other artificial intelligence-based models in terms of root mean square error (RMSE). (Note: Names for each model are provided in Table 3 and Tables 4 and 5).

tively for Barunggam Solar Farm. Additionally, the proposed hybrid method achieved high performance on *GSR* prediction with the high value of KGE (≈ 0.888), GPI (≈ 1.199), and the low value of APB (≈ 9.035).

Thus, the present research avers that by combining the strengths of these two promising DL (CNN and LSTM) methods, the final results attained were superior than the benchmark methods such that the performance of the proposed deep hybrid SCLC model applied for *GSR* prediction had relatively high accuracy, and fast forecasting speed, low volatility of prediction results, and an excellent adaptability to the problem of solar radiation prediction. Further improvements in the deep learning methodology could incorporate design of predictor data decomposition methods such as wavelet analysis and empirical mode decomposition [44, 119, 120] to screen best features using the SMA or other feature selection inputs prior to emulating the global solar radiation, windspeed, air quality and other variables of interest.

Acknowledgement

Data were acquired from (i) Queensland Climate Change Centre of Excellence (QCCCE), which is part of the Department of Science, Information Technology, Innovation, and the Arts (DSITIA) (ii) the Centre for Environmental Data Analysis (CEDA) as a server for the CMIP5 project's GCM output collection for CSIRO-BOM ACCESS1-0, MOHC Hadley-GEM2-CC and the MRI MRI-CGCM3 – which are greatly acknowledged. This work has been partially supported by the project PID2020-115454GB-C21 of the Spanish Ministry of Science and Innovation (MICINN).

References

[1] J. Fan, L. Wu, X. Ma, H. Zhou, F. Zhang, Hybrid support vector machines with heuristic algorithms for prediction of daily diffuse solar radiation in air-polluted regions, Renewable Energy 145 (2020) 2034–2045.

[2] F. J. Rodríguez-Benítez, C. Arbizu-Barrena, J. Huertas-Tato, R. Aler-Mur, I. Galván-León, D. Pozo-Vázquez, A short-term solar radiation forecasting system for the iberian peninsula. part 1: Models description and performance assessment, Solar Energy 195 (2020) 396–412.

[3] A. Khanlari, A. Sözen, C. Şirin, A. D. Tuncer, A. Gungor, Performance enhancement of a greenhouse dryer: Analysis of a cost-effective alternative solar air heater, Journal of Cleaner Production 251 (2020) 119672.

[4] S. A. Kalogirou, Solar energy engineering: processes and systems, Academic Press, 2013.

[5] A. Khanlari, A. Sözen, F. Afshari, C. Şirin, A. D. Tuncer, A. Gungor, Drying municipal sewage sludge with v-groove triple-pass and quadruple-pass solar air heaters along with testing of a solar absorber drying chamber, Science of The Total Environment 709 (2020) 136198.

[6] Ü. Ağbulut, A. E. Gürel, A. Ergün, İ. Ceylan, Performance assessment of a v-trough photovoltaic system and prediction of power output with different machine learning algorithms, Journal of Cleaner Production 268 (2020) 122269.

[7] Ü. Ağbulut, A. E. Gürel, Y. Biçen, Prediction of daily global solar radiation using different machine learning algorithms: Evaluation and comparison, Renewable and Sustainable Energy Reviews 135 (2021) 110114.

[8] N. Dong, J.-F. Chang, A.-G. Wu, Z.-K. Gao, A novel convolutional neural network framework based solar irradiance prediction method, International Journal of Electrical Power & Energy Systems 114 (2020) 105411.

[9] B. Jahani, Y. Dinpashoh, A. R. Nafchi, Evaluation and development of empirical models for estimating daily solar radiation, Renewable and Sustainable Energy Reviews 73 (2017) 878–891.

[10] J. Fan, B. Chen, L. Wu, F. Zhang, X. Lu, Y. Xiang, Evaluation and development of temperature-based empirical models for estimating daily global solar radiation in humid regions, Energy 144 (2018) 903–914.

[11] J.-L. Chen, L. He, H. Yang, M. Ma, Q. Chen, S.-J. Wu, Z.-l. Xiao, Empirical models for estimating monthly global solar radiation: A most comprehensive review and comparative case study in china, Renewable and Sustainable Energy Reviews 108 (2019) 91–111.

[12] B. Jamil, E. Bellos, Development of empirical models for estimation of global solar radiation exergy in india, Journal of Cleaner Production 207 (2019) 1–16.

[13] R. Chang, L. Bai, C.-H. Hsu, Solar power generation prediction based on deep learning, Sustainable Energy Technologies and Assessments 47 (2021) 101354.

[14] Y. Feng, D. Gong, Q. Zhang, S. Jiang, L. Zhao, N. Cui, Evaluation of temperature-based machine learning and empirical models for predicting daily global solar radiation, Energy Conversion and Management 198 (2019) 111780.

[15] A. Khosravi, R. Nunes, M. E. H. Assad, L. Machado, Comparison of artificial intelligence methods in estimation of daily global solar radiation, Journal of cleaner production 194 (2018) 342–358.

[16] S. S. Sharifi, V. Rezaverdinejad, V. Nourani, Estimation of daily global solar radiation using wavelet regression, ann, gep and empirical models: A comparative study of selected temperature-based approaches, Journal of Atmospheric and Solar-Terrestrial Physics 149 (2016) 131–145.

[17] Y. Jiang, Prediction of monthly mean daily diffuse solar radiation using artificial neural networks and comparison with other empirical models, Energy policy 36 (10) (2008) 3833–3837.

[18] F. Zhang, J.-R. Yan, J. Li, K. Wu, H. Iwabuchi, Y.-N. Shi, A new radiative transfer method for solar radiation in a vertically internally inhomogeneous medium, Journal of the Atmospheric Sciences 75 (1) (2018) 41–55.

[19] I. M. Galván, J. Huertas-Tato, F. J. Rodríguez-Benítez, C. Arbizu-Barrena, D. Pozo-Vázquez, R. Aler, Evolutionary-based prediction interval estimation by blending solar radiation forecasting models using meteorological weather types, Applied Soft Computing (2021) 107531.

[20] H. Gamarro, J. E. Gonzalez, L. E. Ortiz, On the assessment of a numerical weather prediction model for solar photovoltaic power forecasts in cities, Journal of Energy Resources Technology 141 (6) (2019).

[21] C. Voyant, M. Muselli, C. Paoli, M.-L. Nivet, Numerical weather prediction (NWP) and hybrid ARMA/ANN model to predict global radiation, Energy 39 (1) (2012) 341–355.

[22] K. Yang, T. Koike, B. Ye, Improving estimation of hourly, daily, and monthly solar radiation by importing global data sets, Agricultural and Forest Meteorology 137 (1-2) (2006) 43–55.

[23] S. Salcedo-Sanz, P. Ghamisi, M. Piles, M. Werner, L. Cuadra, A. Moreno-Martínez, E. Izquierdo-Verdiguier, J. Muñoz-Marí, A. Mosavi, G. Camps-Valls, Machine learning information fusion in earth observation: A comprehensive review of methods, applications and data sources, Information Fusion 63 (2020) 256–272.

[24] O. García-Hinde, G. Terrén-Serrano, M. Hombrados-Herrera, V. Gómez-Verdejo, S. Jiménez-Fernández, C. Casanova-Mateo, J. Sanz-Justo, M. Martínez-Ramón, S. Salcedo-Sanz, Evaluation of dimensionality reduction methods applied to numerical weather models for solar radiation forecasting, Engineering Applications of Artificial Intelligence 69 (2018) 157–167.

[25] A. Hammer, D. Heinemann, C. Hoyer, R. Kuhlemann, E. Lorenz, R. Müller, H. G. Beyer, Solar energy assessment using remote sensing technologies, Remote Sensing of Environment 86 (3) (2003) 423–432.

[26] H. Sun, D. Yan, N. Zhao, J. Zhou, Empirical investigation on modeling solar radiation series with ARMA-GARCH models, Energy Conversion and Management 92 (2015) 385–395.

[27] K. Gairaa, A. Khellaf, Y. Messlem, F. Chellali, Estimation of the daily global solar radiation based on box–jenkins and ANN models: A combined approach, Renewable and Sustainable Energy Reviews 57 (2016) 238–249.

[28] M. Ghofrani, M. Ghayekhloo, R. Azimi, A novel soft computing framework for solar radiation forecasting, Applied Soft Computing 48 (2016) 207–216.

[29] A. Jadidi, R. Menezes, N. De Souza, A. C. de Castro Lima, A hybrid ga–mlpnn model for one-hour-ahead forecasting of the global horizontal

irradiance in elizabeth city, North Carolina, Energies 11 (10) (2018) 2641.

[30] W. Sun, Y. Wang, Short-term wind speed forecasting based on fast ensemble empirical mode decomposition, phase space reconstruction, sample entropy and improved back-propagation neural network, Energy conversion and Management 157 (2018) 1–12.

[31] J. Cao, X. Lin, Application of the diagonal recurrent wavelet neural network to solar irradiation forecast assisted with fuzzy technique, Engineering Applications of Artificial Intelligence 21 (8) (2008) 1255–1263.

[32] H. Jiang, A novel approach for forecasting global horizontal irradiance based on sparse quadratic RBF neural network, Energy Conversion and Management 152 (2017) 266–280.

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

[34] D. Guijo-Rubio, A. Durán-Rosal, P. Gutiérrez, A. Gómez-Orellana, C. Casanova-Mateo, J. Sanz-Justo, S. Salcedo-Sanz, C. Hervás-Martínez, Evolutionary artificial neural networks for accurate solar radiation prediction, Energy 210 (2020) 118374.

[35] M. M. Lotfinejad, R. Hafezi, M. Khanali, S. S. Hosseini, M. Mehrpooya, S. Shamshirband, A comparative assessment of predicting daily solar radiation using bat neural network (BNN), generalized regression neural network (GRNN), and neuro-fuzzy (NF) system: A case study, Energies 11 (5) (2018) 1188.

[36] S. Salcedo-Sanz, C. Casanova-Mateo, A. Pastor-Sánchez, M. Sánchez-Girón, Daily global solar radiation prediction based on a hybrid coral reefs optimization–extreme learning machine approach, Solar Energy 105 (2014) 91–98.

[37] A. Aybar-Ruiz, S. Jiménez-Fernández, L. Cornejo-Bueno, C. Casanova-Mateo, J. Sanz-Justo, P. Salvador-González, S. Salcedo-Sanz, A novel

grouping genetic algorithm–extreme learning machine approach for global solar radiation prediction from numerical weather models inputs, Solar Energy 132 (2016) 129–142.

[38] S. Ghimire, R. C. Deo, N. J. Downs, N. Raj, Self-adaptive differential evolutionary extreme learning machines for long-term solar radiation prediction with remotely-sensed modis satellite and reanalysis atmospheric products in solar-rich cities, Remote Sensing of Environment 212 (2018) 176–198.

[39] J. Chen, W. Zhu, Q. Yu, Estimating half-hourly solar radiation over the continental united states using GOES-16 data with iterative random forest, Renewable Energy 178 (2021) 916–929.

[40] L. Benali, G. Notton, A. Fouilloy, C. Voyant, R. Dizene, Solar radiation forecasting using artificial neural network and random forest methods: Application to normal beam, horizontal diffuse and global components, Renewable energy 132 (2019) 871–884.

[41] J.-L. Chen, H.-B. Liu, W. Wu, D.-T. Xie, Estimation of monthly solar radiation from measured temperatures using support vector machines–a case study, Renewable Energy 36 (1) (2011) 413–420.

[42] J.-L. Chen, G.-S. Li, Evaluation of support vector machine for estimation of solar radiation from measured meteorological variables, Theoretical and applied climatology 115 (3) (2014) 627–638.

[43] M. Alrashidi, M. Alrashidi, S. Rahman, Global solar radiation prediction: Application of novel hybrid data-driven model, Applied Soft Computing 112 (2021) 107768.

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

[45] S. Salcedo-Sanz, C. Casanova-Mateo, J. Muñoz-Marí, G. Camps-Valls, Prediction of daily global solar irradiation using temporal gaussian processes, IEEE Geoscience and Remote Sensing Letters 11 (11) (2014) 1936–1940.

[46] J. Piri, O. Kisi, Modelling solar radiation reached to the earth using ANFIS, NN-ARX, and empirical models (case studies: Zahedan and bojnurd stations), Journal of Atmospheric and Solar-Terrestrial Physics 123 (2015) 39–47.

[47] L. Cornejo-Bueno, C. Casanova-Mateo, J. Sanz-Justo, S. Salcedo-Sanz, Machine learning regressors for solar radiation estimation from satellite data, Solar Energy 183 (2019) 768–775.

[48] S. Ghimire, R. C. Deo, N. J. Downs, N. Raj, Global solar radiation prediction by ann integrated with european centre for medium range weather forecast fields in solar rich cities of queensland australia, Journal of cleaner production 216 (2019) 288–310.

[49] A. E. Gürel, Ü. Ağbulut, Y. Biçen, Assessment of machine learning, time series, response surface methodology and empirical models in prediction of global solar radiation, Journal of Cleaner Production 277 (2020) 122353.

[50] V. H. Quej, J. Almorox, J. A. Arnaldo, L. Saito, ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment, Journal of Atmospheric and Solar-Terrestrial Physics 155 (2017) 62–70.

[51] A. G. Salman, B. Kanigoro, Y. Heryadi, Weather forecasting using deep learning techniques, in: 2015 international conference on advanced computer science and information systems (ICACSIS), Ieee, 2015, pp. 281–285.

[52] X. Qing, Y. Niu, Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM, Energy 148 (2018) 461–468.

[53] S. Ghimire, R. C. Deo, N. Raj, J. Mi, Deep learning neural networks trained with modis satellite-derived predictors for long-term global solar radiation prediction, Energies 12 (12) (2019) 2407.

[54] K. Yan, H. Shen, L. Wang, H. Zhou, M. Xu, Y. Mo, Short-term solar irradiance forecasting based on a hybrid deep learning methodology, Information 11 (1) (2020) 32.

[55] H. Zang, L. Cheng, T. Ding, K. W. Cheung, M. Wang, Z. Wei, G. Sun, Application of functional deep belief network for estimating daily global solar radiation: A case study in china, Energy 191 (2020) 116502.

[56] T. Peng, C. Zhang, J. Zhou, M. S. Nazir, An integrated framework of bi-directional long-short term memory (biLSTM) based on sine cosine algorithm for hourly solar radiation forecasting, Energy 221 (2021) 119887.

[57] W. Bendali, I. Saber, B. Bourachdi, M. Boussetta, Y. Mourad, Deep learning using genetic algorithm optimization for short term solar irradiance forecasting, in: 2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS), IEEE, 2020, pp. 1–8.

[58] M. Abdel-Nasser, K. Mahmoud, M. Lehtonen, Reliable solar irradiance forecasting approach based on choquet integral and deep LSTMs, IEEE Transactions on Industrial Informatics 17 (3) (2020) 1873–1881.

[59] S. Ziyabari, L. Du, S. Biswas, A spatio-temporal hybrid deep learning architecture for short-term solar irradiance forecasting, in: 2020 47th IEEE Photovoltaic Specialists Conference (PVSC), IEEE, 2020, pp. 0833–0838.

[60] C. S. Lai, C. Zhong, K. Pan, W. W. Ng, L. L. Lai, A deep learning based hybrid method for hourly solar radiation forecasting, Expert Systems with Applications 177 (2021) 114941.

[61] D. Cannizzaro, A. Aliberti, L. Bottaccioli, E. Macii, A. Acquaviva, E. Patti, Solar radiation forecasting based on convolutional neural network and ensemble learning, Expert Systems with Applications 181 (2021) 115167.

[62] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., Tensorflow: A system for large-scale machine learning, in: 12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16), 2016, pp. 265–283.

[63] S. Salcedo-Sanz, L. Cornejo-Bueno, L. Prieto, D. Paredes, R. García-Herrera, Feature selection in machine learning prediction systems for

renewable energy applications, Renewable and Sustainable Energy Reviews 90 (2018) 728–741.

[64] M. Almaraashi, Investigating the impact of feature selection on the prediction of solar radiation in different locations in saudi arabia, Applied Soft Computing 66 (2018) 250–263.

[65] M. Castangia, A. Aliberti, L. Bottaccioli, E. Macii, E. Patti, A compound of feature selection techniques to improve solar radiation forecasting, Expert Systems with Applications 178 (2021) 114979.

[66] S. Li, H. Chen, M. Wang, A. A. Heidari, S. Mirjalili, Slime mould algorithm: A new method for stochastic optimization, Future Generation Computer Systems 111 (2020) 300–323.

[67] F. L. Howard, The life history of physarum polycephalum, American journal of botany (1931) 116–133.

[68] N. S. Altman, An introduction to kernel and nearest-neighbor nonparametric regression, The American Statistician 46 (3) (1992) 175–185.

[69] G. Vashishtha, S. Chauhan, M. Singh, R. Kumar, Bearing defect identification by swarm decomposition considering permutation entropy measure and opposition-based slime mould algorithm, Measurement 178 (2021) 109389.

[70] Y. Sun, G. Szűcs, A. R. Brandt, Solar pv output prediction from video streams using convolutional neural networks, Energy & Environmental Science 11 (7) (2018) 1811–1818.

[71] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, et al., Recent advances in convolutional neural networks, Pattern Recognition 77 (2018) 354–377.

[72] D.-X. Zhou, Universality of deep convolutional neural networks, Applied and computational harmonic analysis 48 (2) (2020) 787–794.

[73] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural computation 9 (8) (1997) 1735–1780.

[74] J.-M. Yeom, R. C. Deo, J. F. Adamowski, S. Park, C.-S. Lee, Spatial mapping of short-term solar radiation prediction incorporating geostationary satellite images coupled with deep convolutional LSTM networks for south korea, Environmental Research Letters 15 (9) (2020) 094025.

[75] A. Gensler, J. Henze, B. Sick, N. Raabe, Deep learning for solar power forecasting—an approach using autoencoder and LSTM neural networks, in: 2016 IEEE international conference on systems, man, and cybernetics (SMC), IEEE, 2016, pp. 002858–002865.

[76] S. Srivastava, S. Lessmann, A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data, Solar Energy 162 (2018) 232–247.

[77] A. Graves, Generating sequences with recurrent neural networks, arXiv preprint arXiv:1308.0850 (2013).

[78] Works DoEaP., Powering Queensland Plan: an integrated energy strategy for the state; Queensland Government;2021. (2021).

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

[80] G. Stone, R. Dalla Pozza, J. Carter, G. McKeon, Long paddock: climate risk and grazing information for australian rangelands and grazing communities, The Rangeland Journal 41 (3) (2019) 225–232.

[81] Centre for Environmental Data Analysis. CEDA Archive. 2020. (2020).

[82] The Commonwealth Scientific and Industrial Research Organisation; Bureau of Meteorology (2017). WCRP CMIP5: The CSIRO-BOM team ACCESS1-0 model output collection.: Centre for Environmental Data Analysis; 2017. (2017).

[83] Met Office Hadley Centre (2012). WCRP CMIP5: Met Office Hadley Centre (MOHC) HadGEM2-CC model output collection.: Centre for Environmental Data Analysis; 2012. (2012).

[84] Meteorological Research Institute of the Korean Meteorological Administration (2013). WCRP CMIP5: Meteorological Research Institute of

KMA MRI-CGCM3 model output collection.: Centre for Environmental Data Analysis; 2013. (2013).

[85] S. Ghimire, Z. M. Yaseen, A. A. Farooque, R. C. Deo, J. Zhang, X. Tao, Streamflow prediction using an integrated methodology based on convolutional neural network and long short-term memory networks, Scientific Reports 11 (1) (2021) 1–26.

[86] F. Chollet, et al., Keras (2015) (2017).

[87] P. Goldsborough, A tour of tensorflow, arXiv preprint arXiv:1610.01178 (2016).

[88] H. Xie, L. Zhang, C. P. Lim, Evolving cnn-LSTM models for time series prediction using enhanced grey wolf optimizer, IEEE Access 8 (2020) 161519–161541.

[89] N. Reimers, I. Gurevych, Optimal hyperparameters for deep LSTM-networks for sequence labeling tasks, arXiv preprint arXiv:1707.06799 (2017).

[90] G. Xie, A. Shangguan, R. Fei, W. Ji, W. Ma, X. Hei, Motion trajectory prediction based on a CNN-LSTM sequential model, Science China Information Sciences 63 (11) (2020) 1–21.

[91] I. Priyadarshini, C. Cotton, A novel LSTM–cnn–grid search-based deep neural network for sentiment analysis, The Journal of Supercomputing (2021) 1–22.

[92] S. Cavalli, M. Amoretti, CNN-based multivariate data analysis for bitcoin trend prediction, Applied Soft Computing 101 (2021) 107065.

[93] M. Čokina, V. Maslej-Krešňáková, P. Butka, Š. Parimucha, Automatic classification of eclipsing binary stars using deep learning methods, Astronomy and Computing 36 (2021) 100488.

[94] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).

[95] Y. Qi, Q. Li, H. Karimian, D. Liu, A hybrid model for spatiotemporal forecasting of pm2. 5 based on graph convolutional neural network and

long short-term memory, Science of the Total Environment 664 (2019) 1–10.

[96] Y.-Y. Hong, T. R. A. Satriani, Day-ahead spatiotemporal wind speed forecasting using robust design-based deep learning neural network, Energy 209 (2020) 118441.

[97] Y.-Y. Hong, C. L. P. P. Rioflorido, A hybrid deep learning-based neural network for 24-h ahead wind power forecasting, Applied Energy 250 (2019) 530–539.

[98] Y. Sun, Z. Xie, Y. Chen, X. Huang, Q. Hu, Solar wind speed prediction with two-dimensional attention mechanism, Space Weather 19 (7) (2021) e2020SW002707.

[99] S. Tiwari, R. Sabzehgar, M. Rasouli, Short term solar irradiance forecast based on image processing and cloud motion detection, in: 2019 IEEE Texas Power and Energy Conference (TPEC), IEEE, 2019, pp. 1–6.

[100] J.-D. Jang, A. Viau, F. Anctil, Neural network estimation of air temperatures from avhrr data, International Journal of Remote Sensing 25 (21) (2004) 4541–4554.

[101] L. Prechelt, Early stopping-but when?, in: Neural Networks: Tricks of the trade, Springer, 1998, pp. 55–69.

[102] N. Lu, J. Qin, K. Yang, J. Sun, A simple and efficient algorithm to estimate daily global solar radiation from geostationary satellite data, Energy 36 (5) (2011) 3179–3188.

[103] A. A. Jiménez, C. Q. G. Muñoz, F. P. G. Marquez, L. Zhang, Artificial intelligence for concentrated solar plant maintenance management, in: Proceedings of the tenth international conference on management science and engineering management, Springer, 2017, pp. 125–134.

[104] T. Pan, S. Wu, E. Dai, Y. Liu, Estimating the daily global solar radiation spatial distribution from diurnal temperature ranges over the tibetan plateau in china, Applied Energy 107 (2013) 384–393.

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

[106] J. E. Nash, J. V. Sutcliffe, River flow forecasting through conceptual models part i—a discussion of principles, Journal of hydrology 10 (3) (1970) 282–290.

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

[108] M. Despotovic, V. Nedic, D. Despotovic, S. Cvetanovic, Review and statistical analysis of different global solar radiation sunshine models, Renewable and Sustainable Energy Reviews 52 (2015) 1869–1880.

[109] H. V. Gupta, H. Kling, K. K. Yilmaz, G. F. Martinez, Decomposition of the mean squared error and nse performance criteria: Implications for improving hydrological modelling, Journal of hydrology 377 (1-2) (2009) 80–91.

[110] J. McKenzie, Mean absolute percentage error and bias in economic forecasting, Economics Letters 113 (3) (2011) 259–262.

[111] H. Liu, X. Mi, Y. Li, Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network, Energy Conversion and Management 166 (2018) 120–131.

[112] S. Sun, H. Qiao, Y. Wei, S. Wang, A new dynamic integrated approach for wind speed forecasting, Applied energy 197 (2017) 151–162.

[113] F. X. Diebold, R. S. Mariano, Comparing predictive accuracy, Journal of Business & economic statistics 20 (1) (2002) 134–144.

[114] M. Costantini, C. Pappalardo, Combination of forecast methods using encompassing tests: An algorithm-based procedure, Tech. rep., Reihe Ökonomie/Economics Series (2008).

[115] J. Hora, P. Campos, A review of performance criteria to validate simulation models, Expert Systems 32 (5) (2015) 578–595.

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

[117] R. Marquez, C. F. Coimbra, Proposed metric for evaluation of solar forecasting models, Journal of solar energy engineering 135 (1) (2013) 011016.

[118] D. Yang, S. Alessandrini, J. Antonanzas, F. Antonanzas-Torres, V. Badescu, H. G. Beyer, R. Blaga, J. Boland, J. M. Bright, C. F. Coimbra, et al., Verification of deterministic solar forecasts, Solar Energy 210 (2020) 20–37.

[119] S. Ghimire, R. C. Deo, N. Raj, J. Mi, Wavelet-based 3-phase hybrid svr model trained with satellite-derived predictors, particle swarm optimization and maximum overlap discrete wavelet transform for solar radiation prediction, Renewable and Sustainable Energy Reviews 113 (2019) 109247.

[120] M. S. Al-Musaylh, R. C. Deo, Y. Li, J. F. Adamowski, Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting, Applied energy 217 (2018) 422–439.

Cover Letter

David Casillas Pérez

Universidad Rey Juan Carlos (URJC)

Department of Signal Theory and Communications, ETSIT

david.casillas@urjc.es

Dear Editor:

I would be interested in publishing our article entitled “A Hybrid Deep Learning Methodology with Feature Optimization Approach for Daily Solar Radiation Prediction”.

Please find attached the paper we submit to be considered in as a Full Regular Paper in *Applied Soft Computing*. We state that this is our original work and it has not been submitted to any other publication forum at the same time. Should you have any doubt or question about this submission please do not hesitate in contacting us.

Yours sincerely,

David Casillas Pérez.

Corresponding author

Research Highlights

- A novel deep learning model for Global Solar Radiation prediction is proposed.
- The model integrates deep learning networks with Slime Mould Algorithm optimiser for optimal feature selection.
- Global climate model and meteorological data at solar farms in Australia are considered as predictive variables.
- Error analysis and statistical metrics establish the model's practicality for solar energy management problems.

Declaration of interests

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

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:
