



Contents lists available at ScienceDirect

# Engineering Applications of Artificial Intelligence

journal homepage: [www.elsevier.com/locate/engappai](http://www.elsevier.com/locate/engappai)

## Hybrid deep CNN-SVR algorithm for solar radiation prediction problems in Queensland, Australia

Sujan Ghimire<sup>a</sup>, Binayak Bhandari<sup>b</sup>, David Casillas-Pérez<sup>c</sup>, Ravinesh C. Deo<sup>a</sup>,  
Sancho Salcedo-Sanz<sup>d,\*</sup>

<sup>a</sup> School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield, QLD, 4300, Australia

<sup>b</sup> Department of Railroad Engineering, Woosong University, 171, Dongdaejon-ro, Dong-gu, Daejeon, Republic of Korea

<sup>c</sup> Department of Signal Processing and Communications, Universidad Rey Juan Carlos, Fuenlabrada, 28942, Madrid, Spain

<sup>d</sup> Department of Signal Processing and Communications, Universidad de Alcalá, Alcalá de Henares, 28805, Madrid, Spain



### ARTICLE INFO

MSC:

68T05

68T20

Keywords:

Convolutional Neural Networks

Support Vector Regression

Solar energy prediction

Feature selection

Hybrid models

### ABSTRACT

This study proposes a new hybrid deep learning (DL) model, the called CSVN, for Global Solar Radiation (GSR) predictions by integrating Convolutional Neural Network (CNN) with Support Vector Regression (SVR) approach. First, the CNN algorithm is used to extract local patterns as well as common features that occur recurrently in time series data at different intervals. Then, the SVR is subsequently adopted to replace the fully connected CNN layers to predict the daily GSR time series data at six solar farms in Queensland, Australia. To develop the hybrid CSVN model, we adopt the most pertinent meteorological variables from Global Climate Model and Scientific Information for Landowners database. From a pool of Global Climate Models variables and ground-based observations, the optimal features are selected through a metaheuristic Feature Selection algorithm, an Atom Search Optimization method. The hyperparameters of the proposed CSVN are optimized by mean of the HyperOpt method, and the overall performance of the objective algorithm is benchmarked against eight alternative DL methods, and some of the other Machine Learning approaches (LSTM, DBN, RBF, BRN, MARS, WKNNR, GPML and M5TREE) methods. The results obtained shows that the proposed CSVN model can offer several predictive advantages over the alternative DL models, as well as the conventional ML models. Specifically, we note that the CSVN model recorded a root mean square error/mean absolute error ranging between  $\approx 2.172\text{--}3.305 \text{ MJ m}^2/1.624\text{--}2.370 \text{ MJ m}^2$  over the six tested solar farms compared to  $\approx 2.514\text{--}3.879 \text{ MJ m}^2/1.939\text{--}2.866 \text{ MJ m}^2$  from alternative ML and DL algorithms. Consistent with this predicted error, the correlation between the measured and the predicted GSR, including the Willmott's, Nash-Sutcliffe's coefficient and Legates & McCabe's Index was relatively higher for the proposed CSVN model compared to other DL and Machine Learning methods for all of the study sites. Accordingly, this study advocates the merits of CSVN model to provide a viable alternative to accurately predict GSR for renewable energy exploitation, energy demand or other forecasting-based applications.

### 1. Introduction

Following the recent report by the World Energy Outlook (WEO) (Mead, 2017), power plants in future will be dominated by more renewable energy resources, with two-third of the global investments being diverted to such modern systems. Considering also the finite lifetime of fossil fuels, and their major adverse effects on the environment, we are currently trending significantly towards exploiting the enormous potential of renewable energy resources, including but not limited to solar, hydro-power, wind, geothermal, tidal, and nuclear systems. In addition to the benefits of renewable energy-based systems, the WEO 2020 also reports that the production in this sector has somewhat been less affected during and the aftermath of COVID-19 pandemic,

making this as a resilient power generation resource (IEA, 2020). The three most widely used renewable alternative energy sources are hydro-power, wind, and solar systems. Renewable energy-based power plants when connected to a grid should thus be able to generate all required power to meet the real-time consumer demand. Except for hydro-power, the characteristics of the power produced from other sources such as photovoltaic (PV) and wind energy systems depends on stochastic, temporal, and spatially variant weather conditions. This means that the power production supply can be somewhat unreliable in itself without either the sufficient capacity storage devices *e.g.*, batteries, or backup systems like conventional diesel generators (Bhandari et al., 2014b) or a constant supply-demand management system. Two

\* Corresponding author.

E-mail address: [sancho.salcedo@uah.es](mailto:sancho.salcedo@uah.es) (S. Salcedo-Sanz).

<https://doi.org/10.1016/j.engappai.2022.104860>

Received 24 November 2021; Received in revised form 22 March 2022; Accepted 29 March 2022

Available online xxx

0952-1976/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

of the most plausible solutions to this natural problem could be (1) to either design hybrid systems with optimal component sizing or (2) to accurately model, predict and forecast the uncertainty in renewable energy potentials (Bhandari et al., 2014a; Deo et al., 2016). As the need for monitoring and forecasting systems for solar energy technologies continue to increase, improvements in traditional approaches or building new hybrid predictive systems remains an ongoing motivation for research in the solar energy engineering area.

This study focuses on the prediction of daily global solar radiation (GSR), so the rest of the paper will be confined to a discussion of solar renewable energy only. In principle, the radiant energy within solar radiation can be captured and turned into heat and electricity using different technologies (Ghimire et al., 2019a). In the case of a photovoltaic (PV) system, the output power depends on the solar irradiance and temperature (Ghimire et al., 2019c). Therefore, accurate and precise predictions of GSR is expected to assist in the estimation of power production from solar-based systems. However, the prediction of solar radiation is not a straightforward process as it is likely to depend on several atmospheric parameters of geographic origin (e.g., longitude, latitude, altitude), environmental origin (e.g., time of the day, season, weather, landscape, precipitation, air temperature, humidity, sunshine duration, wind speed and direction, daily global irradiation etcetera) (Zhou et al., 2021b; Salcedo-Sanz et al., 2018b). These variables can be obtained from in-situ measurements or simulated through numerical weather prediction models (Salcedo-Sanz et al., 2020), and used to monitor energy generation, supply or financial (*i.e.*) risk modeling for solar energy industries. In spite of this need, existing traditional methods based on deterministic approaches continue to have caveats that limit their performance accuracy for real-time applications.

The history of GSR estimation dates back several decades. In 1984, Brinsfield et al. (1984) developed a mathematical model as a function of the latitude and the total opaque cloud cover. Although they had verified their model efficacy in some cases, it still did not reliably predict the GSR for high or low cloud cover values. By the end of the last decade, researchers were somewhat convinced that empirical models alone may not be suitable for high prediction accuracy, and started to integrate Artificial Intelligence (AI) based models to obtain more accurate GSR predictions (Zhou et al., 2021b). AI has gradually led into more sophisticated Machine Learning (ML) methods to become more popular for GSR prediction problems. In general, ML aims to build a predictive model from the concept of learning patterns in data so GSR prediction methods became dominated by Artificial Neural Networks (ANN) and statistical learning models (Cao and Lin, 2008; Salcedo-Sanz et al., 2014a; Aybar-Ruiz et al., 2016; Jiang, 2017; Cornejo-Bueno et al., 2019; Guijo-Rubio et al., 2020; Zjavka, 2020; Garud et al., 2021; Karaman et al., 2021; Pang et al., 2020). Some recent specific examples are the works by Cornejo-Bueno et al. (2019), where different ML regressors including Multi-Layer Perceptron (MLP), Extreme Learning Machine (ELM) and Support Vector Regressor (SVR) are tested and compared in a problem of solar radiation prediction in Spain. Ağbulut et al. (2021) who has used four different ML models: ANN, SVR, Kernel Nearest-Neighbor ( $K$ -NN) and Deep Neural Network (DNN) for daily GSR prediction by training a model with daily minimum and maximum ambient temperature, cloud cover, daily extraterrestrial solar radiation, day length as the input and GSR as the target. Their study concluded that an ANN model was the best followed by a DNN, SVR and  $K$ -NN, respectively. Using maximum and minimum temperature, sunshine duration, daylight hour, clear-sky solar radiation, and extraterrestrial radiation as the input variables. Ramedani et al. (2014) has developed a ML model using radial Basis Function and Polynomial Based SVR to predict GSR dataset. Alsharif et al. (2019) has developed a Seasonal Autoregressive Integrated Moving Average model on Korean meteorological data to report the accuracy, suitability, and adequacy of the predicted data. Basaran et al. (2019) has investigated an ensemble of SVR and ANN to estimate GSR with hourly meteorological datasets. In spite of these several applications of both ML and deep learning

models, each kind of predictive model on its own, had its merits and constraints, therefore, have not able to capitalize the benefits of their joint capabilities to improve to the GSR prediction.

Besides ANNs and SVR methods, other ML models such as Decision Trees (Jumin et al., 2021), ELM (Hai et al., 2020; Salcedo-Sanz et al., 2014a), Gradient Boosting (Park et al., 2020) or Random Forest (RF) (Zeng et al., 2020) have been pivotal tools in GSR prediction problems in the last years (see Del Ser et al. (2021), a recent review on randomization-based learning approaches in renewable energy). However, in general, these algorithms follow a shallow network principle for learning, which means that their architecture has a single hidden layer. While such shallow networks showed a good performance for small datasets (Sun et al., 2016), they nonetheless are seen to suffer from major drawbacks of over-fitting, gradient disappearance and network training explosions in some cases (Bengio, 2009). These inherent deficiencies in existing GSR predictive models continued to increase the motivation of researchers to develop fusion-based or hybrid models to handle the challenging task of solar radiation prediction in solar renewable energy engineering problems.

To address some of the accuracy and model integrity issues in the more simplistic predictive models, the advent of Deep Learning (DL) methods have recently received significant research attention, and such studies have noted much better performance than a shallow-equivalent (or non-DL) models. One plausible reason for the DL methods to outperform the conventional predictive methods includes their ability to extract the data features more automatically, without knowing any background details of the data, their powerful generalization capability and the ability to interact with multiple huge datasets (Kawaguchi et al., 2017). Several studies on DL models for GSR prediction problems include: Long Short Term Memory (LSTM) (Al-Hajj et al., 2021; Yeom et al., 2020; de Araujo, 2020; Huang et al., 2020; Husein and Chung, 2019; He et al., 2020), Deep Belief Networks (DBN) (Vijayakumar et al., 2021; Zang et al., 2020a), Convolutional Neural Networks (CNN) (Zang et al., 2020b; Rai et al., 2020; Ghimire, 2019), Echo State Networks (ESN) (Alizamir et al., 2021; Li et al., 2020a,b; Del Ser et al., 2021), Recurrent Neural Networks (RNN) (Alizamir et al., 2021; Ahn and Park, 2021; Kumar et al., 2021) and Gated Recurrent Unit (GRU) (Rai et al., 2020; Sivanand et al., 2021; Jaihuni et al., 2021, 2020). These studies have reported the DL-based models to be a superior approach against the conventional prediction methods in GSR prediction problems (Lima et al., 2022). Furthermore, to fully harness the strengths of multiple DL-based models, researchers have also come up with more innovative ideas to implement both hybrid DL and ML models, such as combinations of LSTM and CNN or convolutional LSTM models (Zang et al., 2020b; Prado-Rujas et al., 2021), specific versions of DL algorithms such as bidirectional LSTM (BiLSTM) (Li et al., 2021; Ziyabari et al., 2020) combination of DL with wavelets approaches (Wang et al., 2018; Rodríguez et al., 2022), or hybridization with meta-heuristic approaches (Bendali et al., 2020; Ghimire et al., 2022). In spite of these intense research on hybrid approaches involving DL methods, there are still points which require attention, in order to improve the performance of the models. Specifically, in problems with a huge number of inputs variables, the application of feature selection problems may considerably improve the performance of the prediction approaches. As previously mentioned, some recent approaches have included meta-heuristics algorithms combined with DL algorithms in order to improve the performance of the latter. However, there have not been, to our knowledge, a previous approach in which meta-heuristics and DL are devoted to improve the feature selection process, and a ML algorithm is finally used to obtain the final prediction, in this case to a solar radiation prediction problem.

Based on the above discussion, in this paper we propose a novel DL hybrid model where a meta-heuristic approach and a CNN are used to carry out a robust feature selection, and they are finally integrated with a SVR method to obtain a final prediction of GSR at six solar energy farms in Queensland, Australia. We have named this approach

as CSVR algorithm. To construct the hybrid CSVR model, we have used a large number of meteorological variables derived from Global Climate Models (GCM) as inputs, together with Queensland Government's Scientific Information for Landowners (SILO), and GSR datasets as the target. From these pool of GCM and SILO variables, the most important features are first selected through a metaheuristic Feature Selection (FS) model called Atom Search Optimization (ASO). Then, the most useful features form the ASO are further processed by the CNN, and the SVR algorithm is finally applied, in order to obtain the GSR prediction.

The performance of the proposed DL hybrid CSVR model is benchmarked in this work against eight alternative DL approaches, as well as other ML models (LSTM, DBN, RBF, Boosting RF Regression (BRF), Multivariate Adaptive Regression Splines (MARS), Weighted  $K$ -Nearest Neighbor (WKNNR), Gaussian Processes for ML (GPML) and M5 Regression Tree (M5TREE)), obtaining excellent performance and improving the results of alternative DL and ML in all cases.

The remainder of this paper has been structured in the following way: next section present a brief description of the different methods which form the proposed CSVR approach, and how they are combined to obtain the final proposed algorithm. Section 3 discusses the details of the GSR prediction problems tackled in this paper, including the data description and processing, the hyperparameters optimization of the proposed model, implementation details of the different models considered for comparison and the performance metrics considered. Section 4 shows and discusses the obtained results in six PV solar farms in Queensland, Australia. Section 6 closes the paper with some final conclusions and remarks on the research carried out. Appendix A shows a list of acronyms to facilitate the reading of the article.

## 2. Methods

In this section we fully describe the different methods which form the proposed CSVR model for GSR prediction. We start with the ASO algorithm for FS, we also describe the CNN and the SVR algorithms which process the input data to obtain the final GSR prediction. To build the proposed model, we divide the modeling steps into the following three stages:

- (i) Application of the ASO metaheuristic as a FS technique to extract the important features to predict GSR.
- (ii) Extracted important features are further processed by the CNN.
- (iii) Outputs of CNN are applied as inputs to a SVR approach to finally obtain a GSR prediction.

### 2.1. Atom Search Optimization

In this work we utilized FS model as an efficient data preprocessing technique to find the best subset of significant features, by eliminating redundant and non-informative features (Salcedo-Sanz et al., 2018a; García-Hinde et al., 2018). The main purpose of FS in ML is to improve the prediction accuracy, improve the multicollinearity problem, enable faster training, reduce complexity of the model, and reduce overfitting. FS can be divided into four types: filter, wrapper, hybrid and embedded methods (Salcedo-Sanz et al., 2018a; Chen et al., 2020; Castangia et al., 2021). Filter methods are independent of any ML algorithm, and based on statistical analysis or mutual information, for e.g.,  $t$ -test, Linear Discriminant Analysis, Minimal Redundancy Maximal Relevance Criterion (mRMR), Dynamic Relevance and Joint Mutual Information, Multivariate Relative Discrimination Criterion. Furthermore, filter methods are simple and can be implemented with less computational cost than other FS techniques. Wrapper methods, unlike filter methods, use a learning algorithm as part of the evaluation of the feature subset. Even though the wrapper methods are computationally costly, they have been shown to be superior to filter methods in many cases. Many wrapper methods have used heuristic search algorithms to find an optimal subset of features. These methods usually start with a

randomly generated solution, and in each iteration, they are one step closer to the best subset of the solution. The evolutionary algorithms used in wrapper methods include Genetic Algorithm (GA) (Salcedo-Sanz et al., 2002), Simulated Annealing (SA) algorithm (Yan et al., 2019), Ant Colony Optimization (Ma et al., 2021), Shuffled Frog Leaping Algorithm, Particle Swarm Optimization (PSO) Algorithm (Zhou et al., 2021a), Grey Wolf Search, Moldovan and Slowik (2021) or Coral Reefs Optimization (CRO) approaches (Salcedo-Sanz et al., 2014b,a), among others. Hybrid methods, which incorporate filter and wrapper methods, are another type of FS processes which can be found in the literature (Solorio-Fernández et al., 2016).

Zhao et al. (2019a) has suggested Atom Search Optimization (ASO) as a new metaheuristic algorithm for solving optimization problems. In multiple benchmark studies, ASO outperformed the recent metaheuristic algorithms like PSO, GA, SA, Gravitational Search Algorithm or Wind Driven Optimization. The ASO is based on the atom's movement theorem, which involves characteristic of the potential function, interaction force, and geometric constraint force. Details on the ASO algorithm can be found in Appendix B.

### 2.2. Convolutional Neural Networks

Our primary feature extraction method in this study is the Convolutional neural networks (CNN), originally developed by LeCun et al. (1989). It is a kind of DL network based on multilayer perceptrons. The following benefits can be attributed to the CNNs' unprecedented success: (a) Feature extraction and classification processes are combined into a single CNNs body, allowing CNNs to learn to refine features from raw data during training; (b) Since, CNN neurons are not connected to the previous layer, besides the weights of the filters are shared, they can interact with large datasets more efficiently than MLP (Gao et al., 2020). Suppose the convolutional layer input is  $X \in \mathbb{R}^{A \times B}$ , where  $A$  and  $B$  are the dimensions of the input data. Then the output of the convolutional layer can be calculated as follows:

$$C_{cn} = f(X * W_{cn} + b_{cn}) \quad (1)$$

where  $*$  is an operator of convolution;  $C_{cn}$  is the  $cn$ -th feature map of the convolutional layer, and the number of the filters is  $CN$ ;  $X$  represents the input data matrix;  $W_{cn}$  is the weight matrix of  $cn$ th filter of the current layer; the  $cn$ th bias is denoted with  $b_{cn}$ ; Finally, an activation function  $f$  is applied to the result. Fig. 1 illustrates the schematic diagram of a typical CNN model which is structured by a series of different stages. The first few stages are composed of two combined layers: convolutional layer and the pooling layer, while the last stage of the architecture consists of a fully connected layer and a traditional regression model.

Even though CNN models are often used for image recognition, 1D CNN models have only recently been proposed for prediction tasks involving time series. Another significant feature of the 1D CNN is that due to the simple and compact design of 1D CNNs that perform one-dimensional convolutions, an efficient and low-cost implementation is possible (Cavalli and Amoretti, 2021). A standard CNN is made up of two layers: a feature extraction layer and a fully connected layer, which are all cascaded together (Ju et al., 2019). The feature extraction layer, which comes after the input layer in the architecture, is made up of multiple layers. There are two kinds of layers in the feature extraction layer: convolution layer and pooling layer (Syarifudin et al., 2021). The convolutional layer employs a number of filters that convolve through the data to provide the filters' activation maps. Neurons are directly connected to the input data points in each filter, multiplying the data points by the weights. The weights of all the neurons in a single filter are shared, reducing the time and complexity of CNN optimization. The pooling layer, also known as subsampling, reduces the size of the matrix. Max pooling and average pooling are the two kinds of pooling layers. Max pooling is a reduction in the size of the matrix by taking the largest value or maximum value contained in the matrix. Average

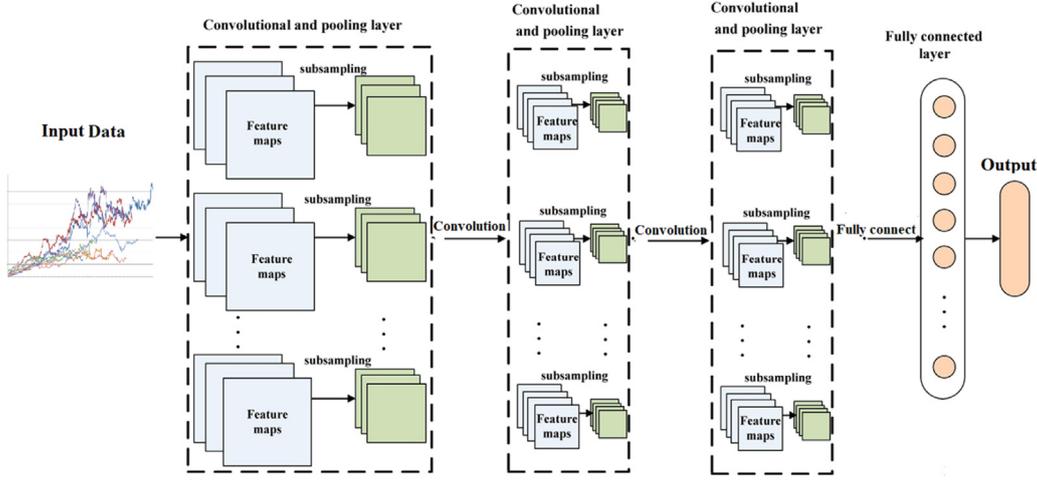


Fig. 1. Structure of the three-layer convolutional neural network algorithm.

pooling is a reduction in the size of the matrix by taking the average value of the matrix (Xie et al., 2020; Kwon et al., 2020).

Finally, these extracted features are combined into the fully connected layer, and the regression problem is solved using an activation function. The backpropagation algorithm (e.g., Stochastic Gradient Descent, Adaptive Gradient Algorithm, Root Mean Square Propagation, Adaptive Moment Estimation (Adam)) is used to train CNNs in a supervised manner. The gradient magnitude (or sensitivity) of each network parameter, such as the weights of the convolution and fully-connected layers, is computed during each iteration of the backpropagation algorithm. The parameter sensitivities are then used to update the CNN parameters iteratively until a certain stopping condition is met. In the 1D domain, a kernel can be regarded as a filter acting as a feature extractor in CNN model. The formula of feature extraction by one-dimensional convolution is described as Ordóñez and Roggen (2016):

$$a_j^{l+1}(\tau) = \sigma \left( b_j^l + \sum_{f=1}^{F^l} K_{jf}^l(\tau) * a_f^l(\tau) \right) = \sigma \left( b_j^l + \sum_{f=1}^{F^l} \sum_{p=1}^{p^l} K_{jf}^l(p) a_f^l(\tau - p) \right) \quad (2)$$

where

- $a_j^{l+1}(\cdot)$  represents the feature map  $j$  in layer  $l$ ,
- $\sigma(\cdot)$  denotes nonlinear function,
- $F^l$  denotes the number of feature maps in layer  $l$ ,
- $K_{jf}^l(\cdot)$  denotes the kernel convolved over feature map  $f$  in layer  $l$  to create the feature map  $j$  in layer  $l + 1$ ,
- $b_j^l$  denotes a bias vector.

### 2.3. Support vector regression

This study has adopted Support Vector Regression (SVR) as the final step of the CSVN model, due to its recent success in energy demand and solar radiation prediction problems (Ghimire et al., 2019d; Al-Musaylh et al., 2018a,b; Deo et al., 2016). In general, SVR is built on a statistical learning theory-based ML system and been widely used for solving high-dimensional regression problems, and it works well in situations where there are few training samples and limited computational resources (Salcedo-Sanz et al., 2014c; Piri et al., 2015).

The SVR model's basic idea is to map the original data points from the input space into a higher or even infinite-dimensional feature of space, where an ideal separating hyperplane is constructed (Smola and Schölkopf, 2004). The distance between all data points and the constructed separating hyperplane is the minimum. Hence, only a brief description of the SVR method is given below, together with a

schematic diagram as shown in Fig. C.15, further information on the technique can be found in Smola and Schölkopf (2004). Details on the training process of the SVR algorithm can be found in Appendix C.

### 2.4. Hybrid CNN-SVR model: Construction, merits and constraints

We proposed a new modeling framework (called CNN-SVR) by combining the CNN and SVR methods to provide a novel data-driven, deep learning method for GSR prediction. The decision to develop CNN-SVR model was based on prior success of the approach in computer vision, albeit extending the CNN-SVR applications to the GSR prediction problem. To the best of our knowledge, there has been no such application so in this model, we constructed a hybrid architecture where the top layer of the traditional CNN was engineered into an SVR training system, creating a stacked layer-by-layer platform with convolutional layers and pooling layers within it. As illustrated in Fig. 9, this structure had combined 9 layers totally including the input layer, three convolutional layers (i.e., C1, C2, C3 and C4), three pooling layers (i.e., P1, P2, P3, P4), one flattening layer and an SVR-based top layer. The fully engineered convolutional layer aimed to extract the short-term patterns in time series to better understand the dependence between predictors to produce a homologous feature map. Each convolutional layer consisted of multiple filters ( $FL$ ) of width  $w$  and height  $h$  equal to the number of variables in the sequence. In convolutional layer, there were no connections between neurons, besides the weights of the filters being shared to provide a seamless passage of data features. Therefore, compared to the MLP (Multilayer Perceptron) system with the same layers and neurons, the CNN system can be trained more efficiently. The pooling layer was added to convolutional layer to reduce the output size and prevent over-fitting. The  $k$ th filter scanned the input matrix  $X_{cT}$  to produce:

$$H_{ck} = f(W_{ck} * X_{cT} + b_{ck}) \quad (3)$$

where  $*$  denoted the convolution operation,  $H_{ck}$  was the output vector,  $W_{ck}$  and  $b_{ck}$  was the weight parameter and  $f$  was the activation function. Finally, the output matrix calculated by convolution component was  $d_c \cdot h \cdot k_c$ ;  $d_c$  = number of filters in last convolutional layer. Therefore, the output after the convolution operation is  $V_c \in R^{d_c \cdot h \cdot k_c}$ . The output from CNN were flattened into 1-D arrays and were input into the SVR model as time sequence to predict the GSR values. One of the main challenges in the training of CNN model is the risk of over-fitting, where the model performs very well during training, but it fails to generalize smoothly in the test data. This problem almost occurs when the training data has a large amount of noise or outlier, which does not really represent the actual properties of the regular patterns. To overcome this problem, there have been several methods proposed,

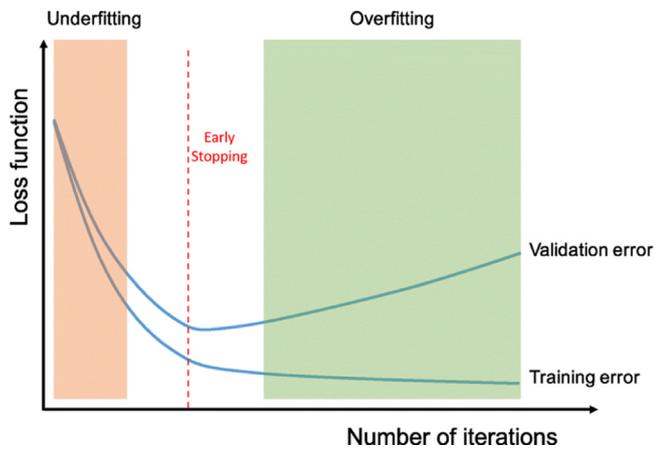


Fig. 2. A routine check to recognize over-fitting by monitoring the loss function on training and validation set during training iterations. If the model performs well on training set compared to the validation set, then the model has been over-fitted to the training set. If the model performs poorly on both training and validation sets, then the model has been under-fit to the data. Although the longer a network is trained, the better it performs on the training set, at some point, the network fits too well to the training data and loses its capability to generalize.

the best solution for reducing over-fitting is to obtain more training data. A model trained on a larger dataset typically generalizes better. The other solutions include early stopping, regularization with dropout or weight decay, batch normalization, and data augmentation, as well as reducing architectural complexity.

In this study we have utilized the early stopping as regularization technique, the available data is divided into training, testing and validation set. The error on the validation set is monitored during the training process. The validation error normally decreases during the initial phase of training, as does the training set error. However, when the network begins to over-fit the data, the error on the validation set typically begins to rise. When the validation error increases for a specified number of iterations the training is stopped, and the weights and biases at the minimum of the validation error are returned. These weight and biases are used in the testing dataset to predict the GSR. This process is exemplified in Fig. 2.

In this study we have engineered a new CNN-SVR framework to produce an efficient and extendable method to predict GSR while testing the overall system at six solar energy farms. We have also adopted an ASO-based as a FS algorithm to select inputs from pool of meteorological and atmospheric variables, fed into a CNN system to extract local pattern/features, as well as common features that recur in the time series at different intervals and subsequently incorporated an SVR regressor to predict the GSR. The proposed CNN-SVR model can predict GSR quickly and accurately by extracting meteorological and atmospheric features that affect GSR. In real-world experiments conducted on Australian solar farm GSR prediction, the proposed CNN-SVR model was compared with LSTM, DBN, RBF, BRf, MARS, WKNNR, GPML and M5TREE models. The results (see Section 4) show the proposed CNN-SVR model can process the time sequence data and effectively extract hidden feature in the datasets. This was mainly because the CNN-SVR topology merges the advantages of the CNN (that extracts data's spatial properties) and the SVR (robust to outliers and has excellent generalization capability, with high prediction accuracy). This set of findings demonstrate that hybrid architectures outperform single models in most scenarios. The CNN-SVR model had the lowest value of Ratio Root Mean Square Error (RRMSE) and Root Mean Square Error (RMSE) and it demonstrates that the proposed CNN-SVR model achieves the best performance among all the eight models. However, even the ameliorated CNN-SVR performs exceptionally well in daily-GSR prediction. Still and all, there are still some improvements to be

made. CNN-SVR can be utilized to predict wind power production, energy costs, and load consumption prediction. Be that as it may, further studies are to be conducted to explore cloud characteristics and apply the collected information to improve GSR prediction model. The main disadvantage with this model is computational time for the selection of hyperparameters, there are seven parameters that need to be selected (Table 4) for CNN-SVR model and we have utilized the Bayesian Optimization (BO) technique to optimize the hyperparameters. One of the reasons behind this was because we have almost 54 years of data ( $20,089 \times 16$ ) for training, which took almost 19 h to get the optimal parameters.

Therefore, reducing the training size with only few optimal input parameters can decrease the computation time but the performance may be hindered. Furthermore, to reduce computational cost while optimizing architecture hyperparameters, researchers often adopt a halving strategy, namely the number of hidden states consecutively reduce by half from lower to higher layer. For instance, if the number of hidden layers  $NL = 3$  and the number of hidden states for the first CNN layer  $Nh1 = 100$ , then the number of hidden states for the second CNN layer  $Nh2 = 50$  and for the third layer  $Nh3 = 25$ . This halving strategy will reduce the number of parameters to be selected and can reduce the computational cost. In our study we have not tested this and used the BO to find the optimal hyperparameters. Moreover, research on the deep learning-based model for GSR prediction is still at an early stage as numerous, more complex and modern deep learning models await further exploration. Finally, it is possible that integration of cloud cover features can also improve predictive power of GSR prediction deep learning models.

The integration of CNN and SVR, as a hybrid CNN-SVR (CSVR) DL neural network, designed for GSR time series prediction in this study, will enable the CNN to extract local pattern features, as well as common features that recur in the time series at different intervals (Tian et al., 2018; Lin et al., 2017). The SVR has replaced the fully connected layer of the CNN to predict the GSR in this case. The inputs after FS using ASO, convolution, and pooling layers of the CNN were set aside for extracting input data functions. The obtained features from CNN were flattened into 1-D arrays and were input into the SVR model as time sequence to predict the GSR values.

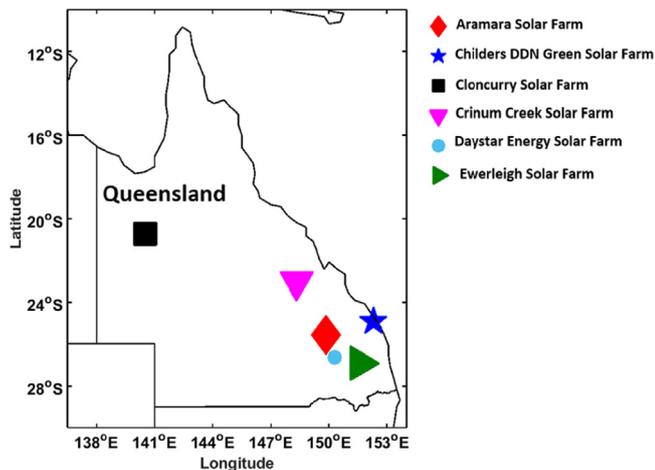
### 3. Data and case study

#### 3.1. Data description

The state of Queensland, often called Australia's "Sunshine state", has abundant solar energy resources (Zahedi, 2010). The Queensland state government has pledged renewable energy sources to account for up to 50% of the overall future energy supply by 2030 in accordance with the United Nations Sustainable Development Goal #7 (SDG7) (Martin, 2019). There are currently 44 large-scale solar energy schemes in Queensland (operating, under construction or financially committed), this equates to an investment of \$ 8.5 billion, 7000 jobs, 4600 MW of renewable electricity, and more than 11 million tons of pollutants avoided. As of January 2021, Queensland has 6200 MW of renewable energy power, which includes rooftop solar, 20% of electricity used in Queensland is produced from renewable energy sources (Works DoEaP, 2021). In this study six solar power generation sites in Queensland, Australia, ranging in size from 30 MW to 150 MW, were selected. The Aramara (140 MW, approved for construction) solar farm project located at Fraser Coast, Aramara, Queensland, Australia on a 323 Ha land and is powered by approximately 500,000 solar PV modules. The Cloncurry solar farm (30 MW) located 120 km East of Mount Isa, Australia, has been announced recently and will be built by Infigen Energy Development. The Childers solar farm (42 MW), located in Bundaberg, Australia has been announced by DDN Green Pty Ltd, using approximately 147,840 PV modules and covering 67 Ha of land. The other three solar farm Cloncurry Solar Farm (100 MW),

**Table 1**  
Descriptive statistics of daily global solar radiation (GSR; MJ m<sup>-2</sup> day<sup>-1</sup>) for six solar energy farms.

Property	Aramara Solar Farm	Childers DDN Green Solar Farm	Cloncurry Solar Farm	Crinum Creek Solar Farm	Daystar Energy Solar Farm	Ewerleigh Solar Farm
<b>Geographic location:</b>						
Latitude	25.572°S	24.923°S	20.706°S	23.031°S	26.643°S	26.913°S
Longitude	149.854°E	152.294°E	140.509°E	148.340°E	150.290°E	150.586°E
Energy Capacity, MW	<b>140</b>	<b>42</b>	<b>30</b>	<b>100</b>	<b>100</b>	<b>150</b>
<b>Daily GSR statistics:</b>						
Median	19.00	20.00	22.00	20.00	20.00	19.00
Mean	19.39	19.33	21.91	19.88	19.86	19.32
Standard deviation	6.32	6.08	5.09	5.83	5.75	6.41
Variance	39.97	37.02	25.94	34.00	33.06	41.15
Maximum	33.00	32.00	32.00	32.00	32.00	33.00
Minimum	4.00	4.00	4.00	4.00	4.00	4.00
Mode	29.00	28.00	28.00	28.00	28.00	29.00
Interquartile range	9.00	8.00	7.00	8.00	8.00	9.00
Skewness	-0.24	-0.34	-0.47	-0.38	-0.39	-0.18
Kurtosis	2.40	2.49	3.11	2.63	2.64	2.33



**Fig. 3.** Location of the six solar energy farms in Queensland, Australia, where CSV model has been implemented to predict the daily GSR data.

Daystar Energy Solar Farm (100 MW) and Ewerleigh Solar Farm (150 MW) are announced and located at Tieri, Columboola and Crossroads respectively (Wikipedia, 2021). The study site details (the statistics of GSR) are shown in Table 1 and their locations are shown in Fig. 3.

Furthermore, the predictor (or input) variables used to predict daily GSR at these study sites were acquired from the meteorological variables produced by a set of GCMs (*i.e.*, cloud parameters, humidity parameters, precipitation, wind speed, etc.) and enriched by the ground-based observation data (*i.e.*, evaporation, vapor pressure, relative humidity at maximum temperature, relative humidity at minimum temperature, rainfall, maximum temperature and minimum temperature) from the Scientific Information for Landowners (SILO) repository. Since, GSR measurements for each solar energy site with an exact latitudinal and longitudinal location are not readily accessible, the ground truth observations of daily GSR are taken from SILO database. The Long Paddock SILO database is operated by the Department of Science, Information Technology, Innovation and the Arts (DSITIA). The GCM outputs are sourced from the web archive at the Centre for Environmental Data Analysis (CEDA), which hosts the CMIP5 project (CEDA, 2020). Daily atmospheric model outputs for historical are sourced from this archive. The models include CSIRO-BOM ACCESS1-0 (grid size 1.25° × 1.875°) (CEDA, 2017), MOHC Hadley-GEM2-CC (grid size 1.25° × 1.875°) (Met Office Hadley Centre, 2012) and the MRI MRI-CGCM3 (grid size 1.121° × 1.125°) (Meteorological Research Institute of the Korean Meteorological Administration, 2013). The runs for the historical outputs span the time range between 1950-01-01T12:00:00

and 2006-01-01T00:00:00. Variables output by the model are indexed by dimensions for longitude, latitude, time, atmospheric pressure (at 8 levels), or as near surface readings. The final dataset obtained in this way was composed of 20 455 records and 75 meteorological variables (20 455 × 75). A brief overview of each meteorological variables in the dataset can be found in Table 2.

In this study we adopt GCMs that depict the climate using a three-dimensional grid over the globe, typically having a horizontal resolution of between 250 and 600 km, 10 to 20 vertical layers in the atmosphere and sometimes as many as 30 layers in the oceans. GCM's employs a mathematical model of the general circulation of a planetary atmosphere or ocean with Navier–Stokes equations on a rotating sphere with thermodynamic terms for energy sources (radiation, latent heat). These equations are the basis for computer programs used to simulate Earth's atmosphere or oceans. Atmospheric and oceanic GCMs (AGCM and OGCM) are key components along with sea ice and land-surface components. In this study, the GCM outputs are sourced from the web archive of the Centre for Environmental Data Analysis (CEDA), which hosts the Coupled Model Intercomparison Project Phase 5 (CMIP5) project's GCM output collection. It is noteworthy that CEDA holds environmental data related to atmospheric and earth observation fields with most popular dataset includes (Climate — *e.g.* HadUK Grid, CMIP, CRU, Composition — *e.g.* CCI, Observations — *e.g.* MIDAS Open, Numerical weather prediction — *e.g.* Met Office NWP, Airborne — *e.g.* FAAM, Satellite data and imagery — *e.g.* Sentinel).

The GCM data (meteorological and atmospheric variables) are extracted for each solar farm locations using latitude and longitude. This requires several of the nearest locations for each variable to be extracted within the region of interest. Similarly past observations of GSR MJ m<sup>-2</sup> (including both direct and indirect radiation) are extracted for each BOM site from the SILO database. The time range available for this data from SILO spans from 1859 up until recent observations, hence historical GCM outputs will need to be aligned with observational data for the time range between 1950 and 2006. 1-dimensional observations are extracted and paired with the daily sequences of GCM outputs for the setting where a 1-dimensional output from the model is required. To reduce the impact of extreme values in different model measures, all data were normalized prior to training the models, a min–max normalization is applied. Furthermore, CMIP5 datasets must pass a series of quality control (QC) checks before they can be published and formally cited. The stages of quality control are: QC1: data and metadata compliance check automatically imposed by CMOR2 and the Earth System Grid data publishing software. QC2: data consistency checks; and QC3: extended checking of data and metadata. In addition to official QC checks, the ACCESS modeling team has also done extensive analysis and manual checking of the data before publishing.

**Table 2**  
Predictor variables and data sources (i.e., atmospheric variables from Global Climate Models and observations from the Scientific Information for Landowners (SILO) repository.

Data repository name	Variable	Description	Units	
Global Circulation Model (GCM) Atmospheric Predictor Variables	clt	Cloud Area Fraction	%	
	hfls	Surface Upward Latent Heat Flux	W m <sup>-2</sup>	
	hfss	Surface Upward Sensible Heat Flux	W m <sup>-2</sup>	
	hur	Relative Humidity	%	
	hus	Near Surface Specific Humidity	g kg <sup>-1</sup>	
	pr	Precipitation	kg m <sup>-2</sup> s <sup>-1</sup>	
	prc	Convective Precipitation	kg m <sup>-2</sup> s <sup>-1</sup>	
	prsn	Solid Precipitation	kg m <sup>-2</sup> s <sup>-1</sup>	
	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	m s <sup>-1</sup>	
	sfcWindmax	Daily Maximum Near-Surface Wind Speed	m s <sup>-1</sup>	
	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	m s <sup>-1</sup>	
	uas	Eastern Near-Surface Wind	m s <sup>-1</sup>	
	va	Northward Wind	m s <sup>-1</sup>	
	vas	Northern Near-Surface Wind	m s <sup>-1</sup>	
	wap	Omega (Lagrangian Tendency of Air Pressure)	Pa s <sup>-1</sup>	
	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	%		

In terms of potential limitations of the database used and its impact on the training of the prediction model and the method for the validation of the prediction tool, we note that CMIP5 data are used extensively in Intergovernmental Panel on Climate Change Assessment Reports (the latest one was IPCC AR5). The use of these data are mostly aimed at: (a) addressing outstanding scientific questions that arose as part of the IPCC reporting process; (b) improving the understanding of the climate system; (c) providing estimates of future climate change and related uncertainties; (d) providing input data for the adaptation to the climate change; (e) examining climate predictability and exploring the ability of models to predict climate on decadal time scales; and (f) evaluating how realistic the different models are in simulating the recent past. Furthermore, a recent study (Jia et al., 2019) focused on assessing the performance of 33 CMIP5 GCMs in simulating temperatures in the Tibetan Plateau. By adopting a multiple-criteria approach, an improved score-based method was used for comprehensive assessment of GCM performance using temperature data gathered from 1961 to 2005. Future temperatures were then projected based on a MME coupled with the Delta method, resulting in near-term (2006–2050) and long-term (2051–2095) temperature projections under RCP4.5 and RCP8.5 scenarios. It was reported that almost all GCMs evaluated in this study (Jia et al., 2019) could reliably reproduce the seasonal temperature pattern of the Tibetan Plateau, with the highest temperature occurring typically in July. However, during winter and spring season the GCM tends to underestimate by an average of  $-2$  °C.

### 3.2. Data preprocessing

The preprocessing of the downloaded meteorological variables was done using the min–max (range between 0 and 1) normalization in order to provide each variables the same order of magnitude and to speed up the training of the ML model (Castangia et al., 2021). The min max normalization method can be expressed as:

$$x_{scaled} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (4)$$

where  $x$  is the vector of values to be scaled,  $x_{scaled}$  is the normalized value of  $x$ ,  $\max(x)$  and  $\min(x)$  are the maximum and minimum values of that vector, respectively.

Furthermore, the ASO FS model was implemented to select the most significant predictor variables by removing unimportant variables from the pool of downloaded 75 meteorological variables. The ASO FS algorithm is proceeded by normalizing all the meteorological variables and running with the following configurations:

- The depth weight  $\alpha = 50$ .
- The multiplier weight  $\beta = 0.2$ .
- The number of atoms in the population:

$$N \in \{5, 10, 35, 80, 100, 150, 250, 350, 400, 550, 700, 900, 1000, 1100, 1200\}.$$

- The number of maximum iterations  $T = 50$ .

The population number  $N$  was varied from 5 to 1200, to find the optimal population which will give the minimum root mean square error (fitness value). As shown in Fig. 4 (Childers DDN Green Solar Farm), when number of atoms  $N$  is 80, after the 23 iterations the fitness value is lower (2.33) when compared to other  $N$  values. Similarly, with increase in number of atoms the fitness value converges faster but the fitness value is high ( $N = 100$  and  $N = 50$ ). It has also been found that higher value of  $N$  increases the computational cost, when  $N = 5$ , the time taken for commutation was 6 min compared to 480 min for  $N = 1200$ . Hence, for Childers DDN Green Solar Farm the number of atoms for the FS process was chosen as 80. With this ASO FS process 15 meteorological variables (Data:  $20\,455 \times 15$ ) are selected for Childers DDN Green Solar Farm, Aramara Solar Farm; Crinum Creek Solar Farm and Daystar Energy Solar Farm, whereas for Cloncurry Solar Farm and Ewerleigh Solar Farm, 16 meteorological variables (Data:  $20\,455 \times 16$ ) are selected. Table 3 shows the final results of the predictors from ASO FS process for the prediction of GSR. Additionally, correlation matrix of predictors and predictands (GSR) is shown in Fig. 5.

Finally, we divided the dataset into a training set and a test set to assess the proposed CSVR model prediction results. In most of

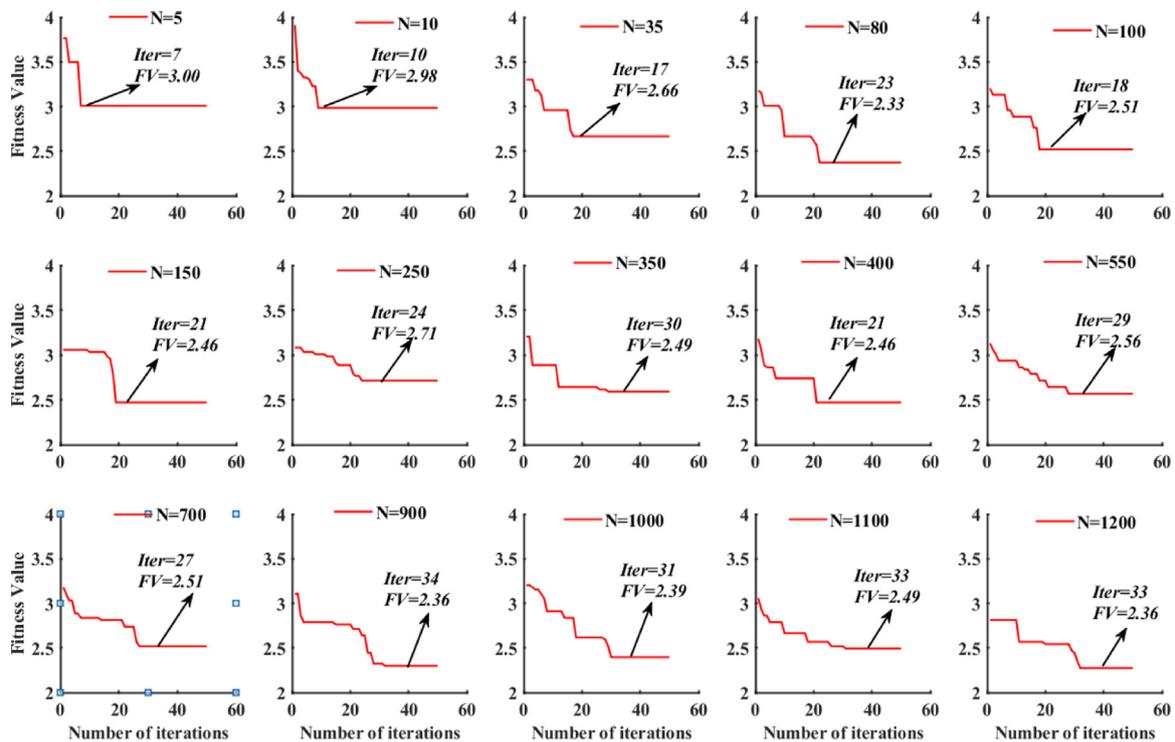


Fig. 4. Convergence curve for the ASO, with increasing number of atoms (N), With N = 80, the fitness value was lower (FV = 2.33) and converges after 23 iterations (Iter).

Table 3

Input variables selected for six solar energy farm sites using a Atom Search Optimization(ASO) feature selection process. Abbreviations as per Table 2 e.g.,  $z_{g1000}$  = Geopotential Height 1000 hPa pressure height.

Aramara Solar Farm	Childers DDN Green Solar Farm	Cloncurry Solar Farm	Crinum Creek Solar Farm	Daystar Energy Solar Farm	Ewerleigh Solar Farm
$z_{g1000}$	Evap	Evap	Evap	Evap	Evap
wap_70000	RHmaxT	RHmaxT	RHmaxT	RHmaxT	RHmaxT
wap_10000	ua_1000	T. Max	hfss	hfss	ua_1000
wap_85000	hfss	hfss	hur_1000	hur_1000	hfss
ua_85000	Rain	ua_5000	ua_5000	wap_1000	hfss
va_85000	hus_5000	hur_1000	wap_1000	ua_5000	hus_5000
ua_50000	$z_{g1000}$	Rain	T. Max	T. Max	wap_1000
va_10000	wap_1000	uas	Rain	Rain	ta_25000
ta_70000	sfcWindmax	wap_1000	va_85000	va_85000	Rain
rhsmx	ua_5000	hur_70000	wap_85000	RHminT	$z_{g1000}$
ua_10000	ta_25000	$z_{g5000}$	$z_{g5000}$	wap_85000	RHminT
tasmx	hur_1000	RHminT	va_50000	$z_{g5000}$	sfcWindmax
prc	va_50000	wap_5000	sfcWindmax	sfcWindmax	T. Max
hus_85000	T. Max	ta_25000	hus_5000	va_50000	ua_5000
va_70000	hfss	va_85000	wap_5000	hus_5000	wap_85000
		va_1000			va_25000

the previous study we found that, 70–30 was used as baseline for training and testing data division, as there are no fixed rule for data division (Ghimire et al., 2019a, 2018, 2019d,c,b; Salcedo-Sanz et al., 2018b), in this study, the 54 years of data used for training (20 089 data points), 20% of data within training set are used for validation (4018 data points) and 1 year of data is used for testing (365 data points), while the target data were the time-series of the daily GSR. It is worth noting that all optimizations were carried out with only the training set to prevent any look-ahead bias. The key phases of our approach are depicted in Fig. 6.

### 3.3. Predictive model development and hyperparameters settings

In this study, the CSVR model is made up of three convolutional layers with pooling operations, the flattening layer's outputs are connected to the SVR to predict the GSR. Fig. 7 shows the proposed CSVR model. The CSVR architecture is defined in more detail as follows.

- Three-layer CNN network (C1, C2 and C3), filter size for each layer is selected using hyperparameter optimization method (HyperOpt).
- Three pooling layers (max-pooling) (P1, P2 and P3) with pool\_size=2.
- One flattening layer, we flatten the output of the convolutional layers to create a single long feature vector to feed as a SVR model input.
- The convolutional layers apply ReLU activation function, this activation function ReLU is used to solve the problem of vanishing gradients, enable models to learn more quickly and perform better (Cavalli and Amoretti, 2021).
- The Adam is selected as the optimization algorithm because it is computationally efficient, has little memory requirements, is invariant to diagonal rescaling of the gradients, and is well suited for large data (Kingma and Ba, 2014). Furthermore, the learning

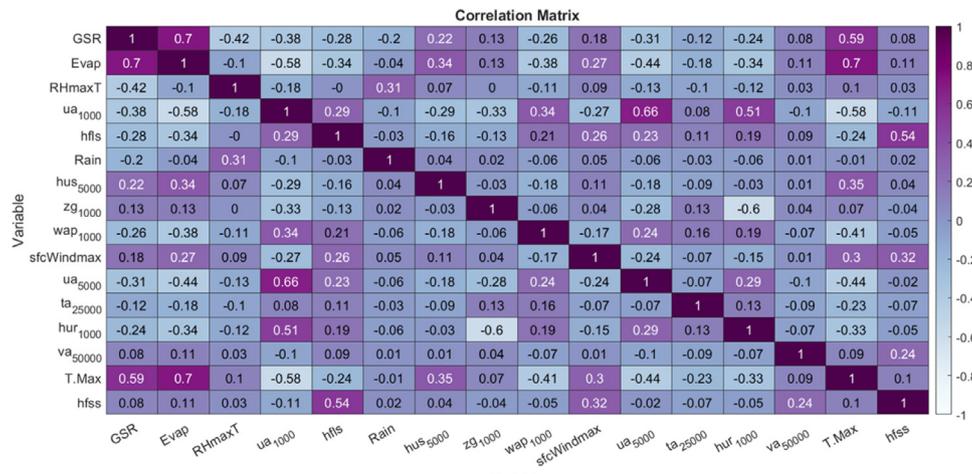


Fig. 5. A matrix of correlation coefficient ( $r$ ) of the predictor variables in respect to the daily GSR data. The variables names are outlined in Table 2.

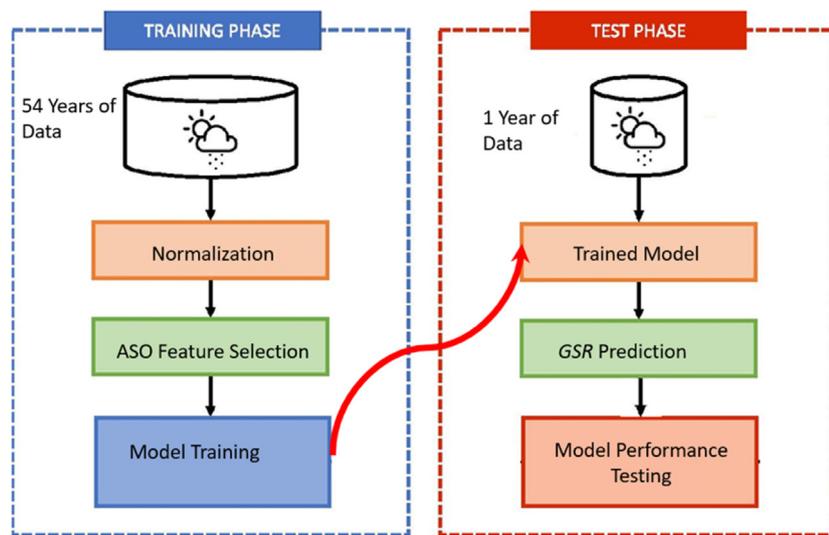


Fig. 6. A schematic of the proposed CSVR methodology employing 54 years of training and 1 year of testing dataset.

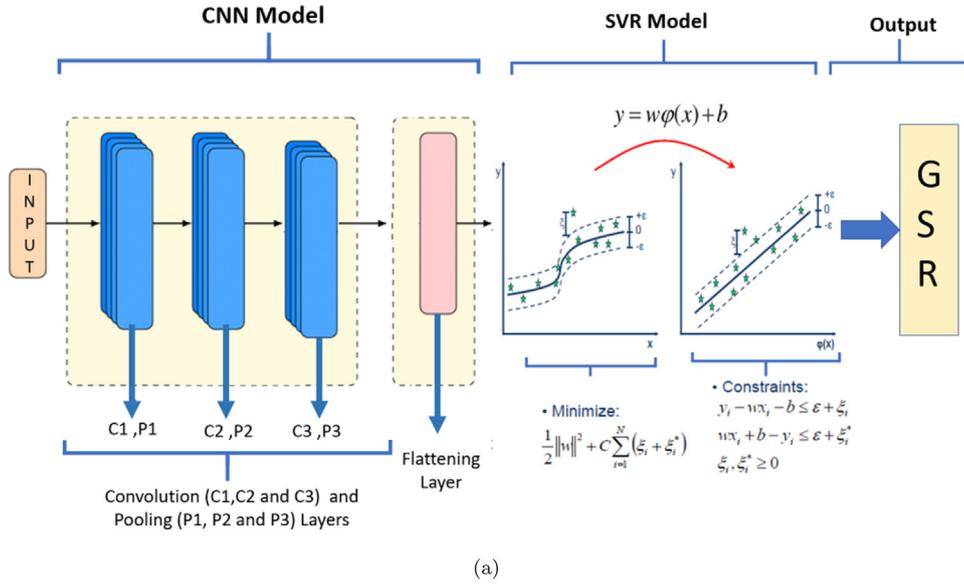
rate of parameters can be dynamically modified in Adam optimizer, as a result, the parameter has a greater chance of escaping the local optimum.

- Number of epoch and batch size is selected using HyperOpt, the epoch specifies how many times the learning algorithm will go through the entire training dataset whereas the batch size is a number of samples processed before the model is updated.
- Radial Basis function is used as kernel for the SVR, the kernel parameter “kernel width” ( $\gamma$ ) and cost function of SVR (C) plays vital role in the model performance and accuracy. These parameters are deduced using HyperOpt.
- The insensitive parameter ( $\epsilon$ ) of SVR is fixed as 0.00001.

### 3.3.1. Hyperparameter optimization

The method of finding the best combination of hyperparameter values in order to achieve optimal output on the data in a reasonable amount of time is called hyperparameter optimization (HPO). HPO is crucial to a ML algorithm’s prediction accuracy. As a result, HPO is regarded as the complicated aspect of developing ML models. In this study, HyperOpt (Bergstra et al., 2013; Komer et al., 2014) is utilized for the HPO of the proposed model as well as benchmark models. Hyperopt is an HPO platform that employs the optimization algorithm based on Bayesian Optimization- Tree-structured Parzen estimator (BO-TPE) in order to tune the ML model to give the good prediction

accuracy. Unlike to some other HPO algorithms (Shen and Rossel, 2021) that only support a single model, HyperOpt can model hierarchical hyperparameters using multiple models. A surrogate function and performance metric are the two key components of a Bayesian optimization algorithm. The surrogate function is a computationally efficient approximation of the real objective function that updates as the optimization is performed. The surrogate function, in combination with the performance metric, provides a new hyperparameter configuration for the next evaluation. The Tree Parzen Estimator is used as a surrogate function, and the expected improvement is used as a performance metric in the TPE algorithm. The TPE algorithm begins with a series of evaluations on hyperparameter configurations that have been randomly sampled. The completed evaluations can then be used to build the optimization history (H), which consists of hyperparameter vector and objective value pairs. The algorithm then employs H to update the surrogate function and choose the next hyperparameter configuration ( $\lambda N$ ). The optimization history H is revised after the evaluation of  $\lambda N$ . Until a stopping criterion is reached, the algorithm continues to propose and evaluate new hyperparameter configurations. The pseudocode of TPE algorithm is shown in Fig. 7(b). The considered search space for the CSVR and the benchmark models hyperparameters is shown in Table 4. HyperOpt considers all possible variations of these values before deciding on the best one. Table 5 shows the optimal hyperparameter obtained by HyperOpt for the objective model CSVR and the benchmarked models.



The TPE (Tree Parzen Estimator) algorithm.

- 1: **procedure** HYPERPARAMETER OPTIMISATION
- 2: Randomly evaluate several independent configurations of hyperparameters
- 3: Establish the optimisation history:  $\mathbf{H}$
- 4: **repeat**
- 5:     *Step 1:* Update the surrogate function based on  $\mathbf{H}$
- 6:     *Step 2:* Select next set of hyperparameters  $\lambda_N$
- 7:     *Step 3:* Evaluate  $\lambda_N$  on the real objective
- 8:     *Step 4:* Update the optimisation history  $\mathbf{H}$
- 9: **until** A stopping criterion reached
- 10: **end procedure**

(b)

Fig. 7. (a) Topological structure of CNN integrated with SVR in the proposed CSVR model used in daily GSR prediction. Mathematical symbols are as per Section 2.2. (b) Pseudocode of The Tree Parzen Estimator (TPE) used in HyperOpt for hyperparameter optimization.

### 3.4. Benchmark models implementations

To comprehensively evaluate the optimal CSVR model for GSR prediction, eight other popular prediction models based on The Long Short-Term Memory (LSTM), Deep Belief Network (DBN), Radial Basis Function Network (RBF), Boosting Random Forest Regression (BRF), Multivariate Adaptive Regression Splines (MARS), Weighted K-Nearest Neighbor (WKNNR), Gaussian Processes For Machine Learning (GPML) and M5 Regression Tree (M5TREE) models were developed. All DL models as well as GPML were built using Keras 2.2.4 (Chollet et al., 2017; Brownlee, 2016) on TensorFlow 1.13.1 (Goldsborough, 2016; Abadi et al., 2016) backend in Python 3.6. Hyperopt-sklearn (Bergstra et al., 2013) library was used for the HPO. The training process of all the models was conducted on a system that has the CPU type of Intel®Core™i7 with 32 GB RAM, remaining models (RBF, BRF, MARS, WKNNR and M5TREE) models were developed using MATLAB software.

### 3.5. Performance metrics

Several approaches for evaluating model efficiency have been used in the past. Since each metric has its own strengths and weaknesses, the current study uses a common collection of statistical metrics (e.g., Correlation Coefficient ( $r$ ), RMSE, Mean Absolute Error (MAE), RRMSE, Relative Mean Absolute Error (RMAE), Willmott's Index (WI), Nash-Sutcliffe's Efficiency (NSE), Legates' Modulus (LM) and Explained Variance Score ( $Evar$ )) to determine model efficiency. These metrics can be

mathematically represented as below (Willmott and Matsuura, 2005; Chai and Draxler, 2014; Moriasi et al., 2007) in Eqs. (5)–(15).

$$r = \frac{\sum_{i=1}^n (GSR^m - \langle GSR^m \rangle)(GSR^p - \langle GSR^p \rangle)}{\sqrt{\sum_{i=1}^n (GSR^m - \langle GSR^m \rangle)^2} \sqrt{\sum_{i=1}^n (GSR^p - \langle GSR^p \rangle)^2}} \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (GSR^p - GSR^m)^2} \quad (6)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |GSR^p - GSR^m| \quad (7)$$

$$RRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (GSR^p - GSR^m)^2}}{\langle GSR^m \rangle} \quad (8)$$

$$RMAE = \frac{1}{n} \sum_{i=1}^n \frac{|GSR^p - GSR^m|}{GSR^p} \quad (9)$$

$$WI = 1 - \frac{\sum_{i=1}^n (GSR^m - GSR^p)^2}{\sum_{i=1}^n (|GSR^p - \langle GSR^m \rangle| + |GSR^m - \langle GSR^m \rangle|)^2} \quad (10)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (GSR^m - GSR^p)^2}{\sum_{i=1}^n (GSR^m - \langle GSR^m \rangle)^2} \quad (11)$$

$$LM = 1 - \frac{\sum_{i=1}^n |GSR^m - GSR^p|}{\sum_{i=1}^n |GSR^m - \langle GSR^m \rangle|} \quad (12)$$

$$Evar = 1 - \frac{\text{Var}(GSR^m - GSR^p)}{\text{Var}(GSR^m)} \quad (13)$$

**Table 4**

Architecture of the CSVR model, with DNN and LSTM models developed for daily GSER prediction. ReLU = Rectified Linear Units; Adam = adaptive moment estimation.

Predictive model	Model Hyperparameters	Hyperparameter Selection	Aramara Solar Farm	Childers DDN Green Solar Farm	Cloncurry Solar Farm	Crinum Creek Solar Farm	Daystar Energy Solar Farm	Ewerleigh Solar Farm
CSVR	Filter1	[50,80,100,200,300,400]	100	80	200	80	100	50
	Filter 2	[40,50,60,70,80]	80	70	80	80	80	80
	Filter 3	[20,10,30,5]	5	5	10	30	20	5
	Batch Size	[25,50,100,400]	50	100	25	100	400	100
	Activation function	ReLU						
	Epochs	[1000,1200,300,400,700]	700	400	300	700	400	400
	Optimization Algorithm	Adam						
	Cost Function	[0.001, 0.01, 0.1, 1, 10, 100]	0.001	10	0.1	100	0.01	100
	Penalty function	[0.001, 0.01, 0.1, 1, 10, 100]	100	0.1	100	0.01	100	0.001
	$\epsilon$	1.00E-05						
SVR_kernel	RBF							
DBN	Hidden Layer Structure	[200-100-50]	200-100-50	200-100-50	200-100-50	200-100-50	200-100-50	200-100-50
	Pre-training iteration	[100,200,300]	100	300	100	200	300	100
	RBM learning rate	[0.0005, 0.001]	0.0005	0.001	0.0005	0.001	0.0005	0.0005
	Samples of pre-training/batch	[75,100,150]	150	75	100	75	150	150
	Reverse fine-tuning epoch	[500,1000,1500]	1000	500	1000	500	1000	500
	Reverse fine-tuning learning rate	0.01						
	Reverse fine-tuning/batch	64						
LSTM	LSTM cell 1	[50, 60,100,200]	100	60	50	200	100	60
	LSTM cell 2	[40,50,60,70,130]	50	70	60	130	70	130
	LSTM cell 3	[20,10,30,5]	5	10	30	20	10	20
	LSTM Cell 4,5 and 6	[Fixed as 30,20,10]						
	Activation function	ReLU						
	Optimization Algorithm	Adam						
	Epochs	[1000,1200,300,400,700]	400	700	1000	400	700	1000
	Drop rate	[0,0.1,0.2]	0	0.1	0	0.1	0	0.1
Batch Size	[25,50,100,400]	40	100	50	25	400	100	

**Table 5**

Architecture of Boosting Random Forest Regression (BRFR), Weighted K-Nearest Neighbor (WKNNR), Gaussian Process of Machine Learning (GPML), Radial Basis Function (RBF) Network And Multivariate Adaptive Regression Splines (MARS).

Model	Model Hyperparameters	Hyperparameter Selection	Aramara Solar Farm	Childers DDN Green Solar Farm	Cloncurry Solar Farm	Crinum Creek Solar Farm	Daystar Energy Solar Farm	Ewerleigh Solar Farm
BRFR	The maximum depth of the tree.	[5, 8, 10, 20, 25]	10	20	25	8	10	20
	The number of trees in the forest.	[50, 100, 150, 200, 400, 600, 800]	50	100	50	50	100	100
	Minimum number of samples to split an internal node	[2, 4, 6, 8, 10]	4	6	2	6	8	6
	The number of features to consider when looking for the best split.	['auto', 'sqrt', 'log2']	auto	auto	auto	auto	auto	auto
MARS	Maximum term generated by forward pass	[10, 20, 30]	10	10	20	10	30	10
	Maximum degree of terms generated by forward pass	[5, 10, 15, 20]	10	15	5	10	15	15
WKNNR	Number of neighbors	[5, 10, 20, 30, 50, 100]	20	30	50	50	20	50
	Algorithm used to compute the nearest neighbors	['auto', 'ball_tree', 'kd_tree', 'brute']	auto	auto	auto	auto	auto	auto
GPML	The kernel specifying the covariance function of the Gaussian Process.	[Dot Product, White Kernel, Dot Product + White Kernel, RBF, Matern]	Dot Product	Dot Product + White Kernel	Dot Product + White Kernel	Dot Product	Dot Product + White Kernel	Dot Product + White Kernel
RBF	Spread value	[1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5]	1.5	1.5	2	1.5	3	1.5
	Maximum number of neurons	[50, 100, 150, 200, 250, 300]	150	250	200	100	200	150

$$SS = 1 - \frac{RMSE(p, x)}{RMSE(pr, x)} \tag{14}$$

$$RMSE_r = \frac{RMSE(p, x)}{RMSE(r, x)} \tag{15}$$

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  $\infty$  (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  $RRMSE > 30\%$  (Pan et al., 2013).
- $WI$  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) (Willmott and Matsuura, 2005).
- $NSE$ , compares the variance of observed and predicted  $GSR$  and ranges from  $-\infty$  (the worst fit) to  $1$  (perfect fit) (Nash and Sutcliffe, 1970).
- $LM$ , is a more robust metrics developed to address the limitations of both the  $WI$  and  $ENS$  (Legates and McCabe, 1999) 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) (Despotovic et al., 2015). GPI was calculated using the six metrics.

$$GPI_i = \sum_{j=1}^6 \alpha_j (g_j - y_{ij}) \quad (16)$$

where  $\alpha_j$  denotes the median of scaled values of statistical indicator  $j$ , equals to  $1$  for  $RMSE$ ,  $MAE$ , Mean Absolute Percentage Error (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) (Gupta et al., 2009) and Absolute Percentage Bias (APB; %) (McKenzie, 2011). Mathematically, these metrics are stated as follows:

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\langle GSR^p \rangle}{\langle GSR^m \rangle} - 1\right)^2 + \left(\frac{CV_p}{CV_m}\right)^2} \quad (17)$$

$$APB = \frac{\sum_{i=1}^n (GSR^m - GSR^p) \cdot 100}{\sum_{i=1}^n GSR^m} \quad (18)$$

where  $r$  is the correlation coefficient,  $CV$  is the coefficient of variation,  $GSR^p$  refers to the predicted  $GSR$  ( $MJ\ m^{-2}\ day^{-1}$ ),  $GSR^m$  is the measured  $GSR^p$  ( $MJ\ m^{-2}\ day^{-1}$ ),  $\langle GSR^m \rangle$  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: APB ( $\lambda_{APB}$ ), MAE ( $\lambda_{MAE}$ ), and RMSE ( $\lambda_{RMSE}$ ) (Liu et al., 2018) to compare various models that have been used in the  $GSR$  prediction.

$$\lambda_{APB} = \left| \frac{APB_1 - APB_2}{APB_1} \right| \quad (19)$$

$$\lambda_{MAE} = \left| \frac{RMAE_1 - RMAE_2}{RMAE_1} \right| \quad (20)$$

$$\lambda_{RRMSE} = \left| \frac{RRMSE_1 - RRMSE_2}{RRMSE_1} \right| \quad (21)$$

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

In addition, we have also used supplementary assessment criteria denoted as the Diebold–Mariano (DM) test, the Harvey, Leybourne,

and the Newbold (HLN) to test the statistical significance of all models in this study. These statistical tests are done to further evaluate the proposed 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  $> 0$  and the HLN statistics  $> 0$ . The key steps in implementing the DM and HLN tests are defined in previous literature (Sun et al., 2017; Diebold and Mariano, 2002; Costantini and Pappalardo, 2008).

#### 4. Results and discussion

Using statistical score metrics presented in Section 3.5 aided by several diagnostic plots, the newly developed deep hybrid CSVr model, coupled with the ASO method for FS was evaluated. Over the testing period, the CSVr model was compared with a standalone DL method (i.e., LSTM, DBN) as well as the other ML methods (i.e., RBF, BRf, MARS WKNNR GPML and M5TREE). The objective model (i.e., CSVr) that showed the lowest  $RMSE$ ,  $MSE$ ,  $RRMSE$ ,  $RMAE$ ,  $MAPE$ , and  $APB$  values and the highest  $KGE$ ,  $NSE$ ,  $r$ ,  $LM$ , and  $WI$  was chosen and finally model are ranked on the basis of GPI.

In Table 6, we show the performance results of the deep hybrid CSVr models in terms of the  $r$ ,  $RMSE$ , and  $MAE$  performance metrics used in evaluation of the predictive performance of daily  $GSR$  models at all of the six solar farms. The comparative statistics are also provided for the other DL (LSTM, DBN) model variants along with the other conventional ML models (i.e., RBF, BRf, MARS WKNNR GPML and M5TREE) used as benchmarks for the CSVr. As compared to the other models, the statistical performance metrics for deep hybrid CSVr models were generally larger in magnitude. For instance, CSVr model for Daystar Energy Solar Farm produced higher  $r$  ( $\approx 0.927$ ) and low  $RMSE$  ( $\approx 2.172\ MJ\ m^{-2}\ day^{-1}$ ) and  $MAE$  ( $\approx 1.624\ MJ\ m^{-2}\ day^{-1}$ ) followed by LSTM model ( $r \approx 0.901$ ,  $RMSE \approx 2.514\ MJ\ m^{-2}\ day^{-1}$ ,  $MAE \approx 1.939\ MJ\ m^{-2}\ day^{-1}$ ). The other DL model i.e. DBN model show the worst performance ( $r \approx 0.393$ ,  $RMSE \approx 5.418\ MJ\ m^{-2}\ day^{-1}$ ,  $MAE \approx 4.301\ MJ\ m^{-2}\ day^{-1}$ ). Further for, conventional ML models, MARS ( $r \approx 0.897$ ,  $RMSE \approx 2.647\ MJ\ m^{-2}\ day^{-1}$ ,  $MAE \approx 2.022\ MJ\ m^{-2}\ day^{-1}$ ) shows better performance than RBF, BRf, WKNNR GPML and M5TREE. It was therefore evident that the deep hybrid CSVr model was considerably better than the LSTM and the other comparable conventional ML models.

In Table 7, the results of all DL models as well as conventional ML models developed to predict daily  $GSR$  at six solar farms are summarized based on  $WI$ ,  $NSE$ ,  $LM$  and  $Evar$ . In all six solar farms, the deep hybrid CSVr model performs best with the highest magnitude of (e.g., Daystar Energy Solar Farm)  $WI$  ( $\approx 0.928$ ),  $NSE$  ( $\approx 0.856$ ),  $LM$  ( $\approx 0.651$ ), and  $Evar$  ( $\approx 0.859$ ). These metrics for the other DL model LSTM ( $WI \approx 0.907$ ,  $NSE \approx 0.807$ ,  $LM \approx 0.584$  and  $Evar \approx 0.809$ ) and DBN are in lower magnitude  $WI \approx 0.006$ ,  $NSE \approx 0.107$ ,  $LM \approx 0.076$  and  $Evar \approx 0.108$ . Likewise, for the conventional ML models, RBF model shows better performance for the Aramara Solar Farm, whereas MARS model shows better performance for the Childers DDN Green Solar Farm, Cloncurry Solar Farm, Crinum Creek Solar Farm, Daystar Energy Solar Farm and Ewerleigh Solar Farm. In general for all six solar farms with  $WI > 0.822$ ,  $NSE > 0.856$ ,  $LM > 0.651$  and  $Evar > 0.859$ , the CSVr outperform all the other comparative models. Furthermore, it should also be noted that models can be classified on the basis of NSE metrics as unsatisfactory ( $NSE < 0.800$ ), fairly good ( $0.800 \leq NSE \leq 0.900$ ) and very satisfactory ( $NSE > 0.900$ ).

A further evaluation of the deep hybrid CSVr model is conducted by means of scatter plot diagrams (Fig. 8) of  $GSR_{pred}$  versus  $GSR_{obs}$ , where each scatterplot is generated using a linear regression equation  $GSR_{pred} = m \cdot GSR_{obs} + C$ . The scatter plot also reveals that the CSVr model performs the best since its scatter points lie close to the regression line, whilst the other models are far from the regression line. In line with Tables 6, 7, Fig. 8 confirms that deep hybrid CSVr

**Table 6**

Testing performance of CSVr. (a) correlation coefficient (r), root mean square error RMSE ( $\text{MJ m}^{-2} \text{ day}^{-1}$ ) and mean absolute error MAE (GSRUnitD). (b) relative RMSE and MAE values (%).

Predictive models	Aramara Solar Farm			Childers DDN Green Solar Farm			Cloncurry Solar Farm			Crinum Creek Solar Farm			Daystar Energy Solar Farm			Ewerleigh Solar Farm		
	r	RMSE	MAE	r	RMSE	MAE	r	RMSE	MAE	r	RMSE	MAE	r	RMSE	MAE	r	RMSE	MAE
<b>CSVr</b>	<b>0.863</b>	<b>3.305</b>	<b>2.370</b>	<b>0.893</b>	<b>2.635</b>	<b>1.870</b>	<b>0.863</b>	<b>2.519</b>	<b>1.860</b>	<b>0.908</b>	<b>2.437</b>	<b>1.805</b>	<b>0.927</b>	<b>2.172</b>	<b>1.624</b>	<b>0.924</b>	<b>2.429</b>	<b>1.779</b>
LSTM	0.790	3.879	2.866	0.876	2.862	2.129	0.848	2.703	1.940	0.885	2.772	2.085	0.901	2.514	1.939	0.908	2.673	1.952
DBN	0.494	5.546	4.407	0.711	4.040	3.204	0.598	3.709	2.890	0.486	5.207	4.069	0.393	5.418	4.301	0.444	6.005	4.877
RBF	0.823	3.636	2.642	0.857	3.100	2.307	0.757	3.245	2.401	0.865	3.358	2.577	0.863	3.206	2.404	0.839	3.775	2.915
MARS	0.806	3.761	2.809	0.865	2.923	2.200	0.837	2.693	2.075	0.884	2.737	2.147	0.897	2.647	2.022	0.889	3.031	2.345
WKNNR	0.605	5.010	3.882	0.773	3.594	2.811	0.755	3.075	2.271	0.869	2.983	2.278	0.882	2.858	2.195	0.830	3.539	2.574
GPML	0.607	5.158	3.989	0.610	4.763	3.896	0.706	3.397	2.634	0.805	4.003	3.223	0.803	3.975	3.231	0.753	4.478	3.432
BRFR	0.648	5.202	3.981	0.748	4.188	3.169	0.738	3.470	2.656	0.797	3.659	2.869	0.805	3.520	2.687	0.814	3.914	3.081
M5TREE	0.664	5.251	3.716	0.778	3.986	2.906	0.724	3.567	2.680	0.774	3.895	2.941	0.824	3.424	2.636	0.826	3.792	2.958

**Table 7**

Testing performance of CSVr model in terms of (a) Willmott's Index (WI) and Nash–Sutcliffe coefficients (NS) and (b) Legates and McCabe's index (LM) and explained variance score (Evar).

(a)													
Predictive models	Aramara Solar Farm		Childers DDN Green Solar Farm		Cloncurry Solar Farm		Crinum Creek Solar Farm		Daystar Energy Solar Farm		Ewerleigh Solar Farm		
	WI	NS	WI	NS	WI	NS	WI	NS	WI	NS	WI	NS	
<b>CSVr</b>	<b>0.822</b>	<b>0.726</b>	<b>0.888</b>	<b>0.785</b>	<b>0.807</b>	<b>0.713</b>	<b>0.907</b>	<b>0.822</b>	<b>0.928</b>	<b>0.856</b>	<b>0.926</b>	<b>0.853</b>	
LSTM	0.749	0.617	0.875	0.745	0.781	0.677	0.871	0.773	0.907	0.807	0.903	0.822	
DBN	0.206	0.218	0.651	0.493	0.558	0.355	0.050	0.203	0.006	0.107	0.285	0.116	
RBF	0.782	0.665	0.843	0.704	0.679	0.530	0.807	0.684	0.808	0.702	0.772	0.661	
MARS	0.765	0.641	0.864	0.735	0.816	0.670	0.889	0.775	0.898	0.788	0.880	0.775	
WKNNR	0.579	0.362	0.751	0.596	0.711	0.562	0.834	0.733	0.844	0.751	0.810	0.688	
GPML	0.482	0.325	0.450	0.298	0.560	0.461	0.625	0.518	0.620	0.517	0.661	0.509	
BRFR	0.631	0.306	0.758	0.454	0.722	0.448	0.824	0.599	0.829	0.622	0.824	0.619	
M5TREE	0.632	0.300	0.778	0.506	0.718	0.410	0.794	0.544	0.841	0.642	0.816	0.646	

(b)													
Predictive models	Aramara Solar Farm		Childers DDN Green Solar Farm		Cloncurry Solar Farm		Crinum Creek Solar Farm		Daystar Energy Solar Farm		Ewerleigh Solar Farm		
	LM	Evar	LM	Evar	LM	Evar	LM	Evar	LM	Evar	LM	Evar	
<b>CSVr</b>	<b>0.537</b>	<b>0.744</b>	<b>0.600</b>	<b>0.792</b>	<b>0.500</b>	<b>0.743</b>	<b>0.614</b>	<b>0.823</b>	<b>0.651</b>	<b>0.859</b>	<b>0.662</b>	<b>0.853</b>	
LSTM	0.440	0.624	0.545	0.749	0.478	0.720	0.555	0.785	0.584	0.809	0.629	0.822	
DBN	0.139	0.221	0.315	0.496	0.223	0.356	0.131	0.207	0.076	0.108	0.072	0.118	
RBF	0.484	0.675	0.507	0.716	0.355	0.560	0.449	0.735	0.484	0.741	0.446	0.693	
MARS	0.451	0.650	0.530	0.741	0.442	0.692	0.541	0.777	0.566	0.797	0.554	0.787	
WKNNR	0.242	0.367	0.399	0.596	0.389	0.570	0.513	0.734	0.529	0.755	0.510	0.688	
GPML	0.221	0.329	0.167	0.301	0.292	0.465	0.311	0.519	0.306	0.517	0.347	0.517	
BRFR	0.222	0.307	0.323	0.456	0.286	0.461	0.387	0.602	0.423	0.623	0.414	0.619	
M5TREE	0.274	0.304	0.379	0.510	0.279	0.416	0.372	0.544	0.434	0.643	0.437	0.652	

models demonstrated better and more reliable prediction capabilities throughout all six solar farms.

The model outcomes outlined in Tables 7 and 8 indicate there is a distinct difference in the performance of these models. However, one of the limitations of the above measures (Eqs. (5)–(7), (10)–(12)) is that they do not match models operating at different physical locations. Consequently, the comparative accuracy (Table 8) shows that the CSVr model exhibited the lowest *RRMSE* and *RMAE* compared to LSTM, DBN, RBF, BRFR, MARS WKNNR GPML and M5TREE.

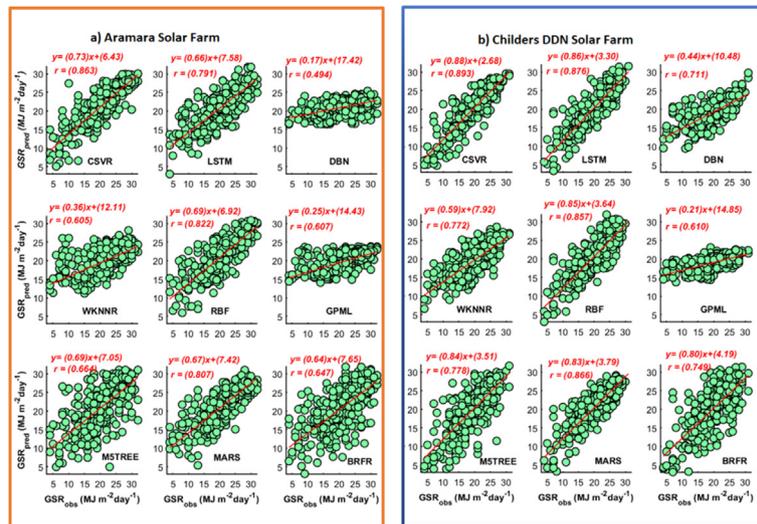
To discuss the objective model more specifically, we group the magnitude of the *RRMSE* and *RMAE* values for the proposed deep hybrid CSVr model against LSTM and MARS models as [*RRMSE*, *RMAE* for CSVr : *RRMSE*, *RMAE* for LSTM : *RRMSE*, *RMAE* for MARS]. For all of the six solar energy farms, these comparative ratios could be enumerated as follows:

- Daystar Energy Solar Farm:  
[10.67%, 9.74% : 12.95%, 11.89% : 13.00%, 11.88%],
- Childers DDN Green Solar Farm:  
[13.52%, 11.91% : 14.69%, 13.83% : 14.99%, 14.17%],
- Cloncurry Solar Farm:  
[11.39%, 9.98% : 12.23%, 10.33% : 12.18%, 10.55%],

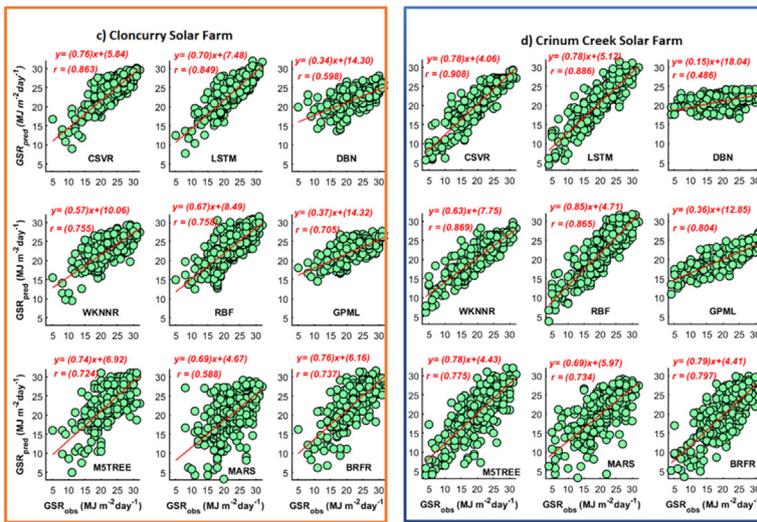
- Crinum Creek Solar Farm:  
[12.10%, 11.08% : 13.66%, 13.14% : 13.49%, 12.89%],
- Ewerleigh Solar:  
[12.06%, 10.71% : 13.27%, 12.91% : 15.04%, 14.17%],  
and
- Aramara Solar Farm:  
[16.41%, 16.38% : 19.26%, 19.69% : 18.67%, 19.79%].

It is thus clear that the proposed hybrid CSVr model was able to produce more accurate results for daily GSR prediction compared to all of the benchmark deep learning (*i.e.*, LSTM, DBN) methods, and also the other conventional ML methods.

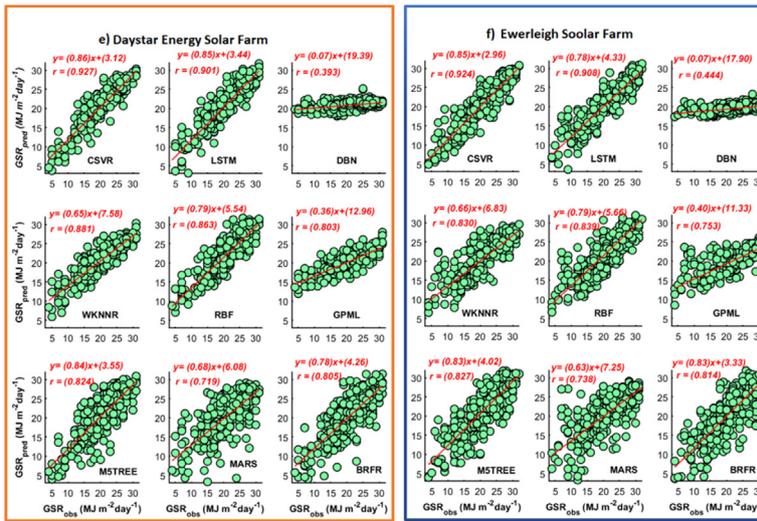
An analysis of the frequency distribution of prediction error  $|PE|$  produced by CSVr compared to benchmark methods is shown in Fig. 9. Each error bin contains the percentage of all tested points, which is shown at the top of each error bar. Each error bin has a size of  $1.5 \text{ MJ m}^{-2} \text{ day}^{-1}$ . Interestingly, GSR predictions using the deep hybrid CSVr model were found to have the greatest frequency of errors within the smallest error bracket ( $1.5 \text{ MJ m}^{-2} \text{ day}^{-1}$ ), which encompassed 55% of the test data. On the other hand, the LSTM, MARS, WKNNR, RBF, M5TREE, BRFR, GPML and DBN accumulated 48%, 44%,



(a) Aramara and Childers DDN solar farms.



(b) Cloncurry and Crinum Creek solar farms.



(c) Daystar energy and and Ewerleigh solar farms.

Fig. 8. Scatter plots of the observed ( $GSR_{obs}$ ) and the predicted ( $GSR_{pred}$ ) daily GSR at six solar energy farms. Red line shows least-square regression  $y = mx + c$  where  $y$  is the  $GSR_{pred}$ ,  $x$  is the  $GSR_{obs}$  and  $r$  is the correlation coefficient. Names for each model are provided in Tables 4 and 5.

40%, 40%, 36%, 32%, 28% and 24% respectively. In accordance with Fig. 8 and Tables 6–8, the CSVR model generates most of its error predictions within the lowest magnitude band, making it more accurate

in predicting GSR at the six solar farms. Additionally, violin plot of prediction error  $|PE|$  for CSVR as well as the other DL models and the conventional ML models are created in Fig. 10. The  $|PE|$  error

**Table 8**  
Testing performance of CSVr in terms of relative RMSE and MAE values (%).

Predictive models	Aramara Solar Farm		Childers DDN Green Solar Farm		Cloncurry Solar Farm		Crimum Creek Solar Farm		Daystar Energy Solar Farm		Ewerleigh Solar Farm	
	RRMSE	RMAE	RRMSE	RMAE	RRMSE	RMAE	RRMSE	RMAE	RRMSE	RMAE	RRMSE	RMAE
<b>CSVr</b>	<b>16.405</b>	<b>16.373</b>	<b>13.521</b>	<b>11.913</b>	<b>11.395</b>	<b>9.979</b>	<b>12.014</b>	<b>11.084</b>	<b>10.668</b>	<b>9.474</b>	<b>12.056</b>	<b>10.712</b>
LSTM	19.257	19.693	14.686	13.828	12.226	10.328	13.664	13.143	12.345	11.890	13.265	12.907
DBN	27.531	30.967	20.728	21.672	16.773	15.226	25.666	29.296	26.612	29.492	29.805	31.397
RBF	18.049	18.531	15.903	14.286	14.677	12.984	16.555	15.900	15.745	15.169	18.736	19.072
MARS	18.669	19.785	14.996	14.167	12.178	10.549	13.494	12.886	13.003	11.881	15.043	14.165
WKNNR	24.866	25.872	18.439	18.461	13.906	11.959	14.703	14.634	14.036	13.876	17.563	17.373
GPML	25.602	26.944	24.437	25.929	15.362	14.076	19.734	21.412	19.522	21.069	22.225	22.106
BRFR	25.821	25.551	21.485	18.788	15.694	13.499	18.036	16.567	17.290	15.192	19.426	18.254
M5TREE	26.066	24.947	20.453	17.621	16.135	13.768	19.198	16.548	16.816	14.548	18.822	17.505

**Table 9**  
The promoting percentage metric,  $\lambda$  for the comparison models against objective (i.e., CSVr) model in the testing phase. Note that  $\lambda_{RMAE}$  = promoting percentage of the relative mean absolute error,  $\lambda_{RRMSE}$  = promoting percentages of relative root mean square error, and  $\lambda_{APB}$  = promoting percentages of absolute percentage bias.

Predictive models	Aramara Solar Farm			Childers DDN Green Solar Farm			Cloncurry Solar Farm			Crimum Creek Solar Farm			Daystar Energy Solar Farm			Ewerleigh Solar Farm		
	$\lambda_{RRMSE}$	$\lambda_{RMAE}$	$\lambda_{APB}$	$\lambda_{RRMSE}$	$\lambda_{RMAE}$	$\lambda_{APB}$	$\lambda_{RRMSE}$	$\lambda_{RMAE}$	$\lambda_{APB}$	$\lambda_{RRMSE}$	$\lambda_{RMAE}$	$\lambda_{APB}$	$\lambda_{RRMSE}$	$\lambda_{RMAE}$	$\lambda_{APB}$	$\lambda_{RRMSE}$	$\lambda_{RMAE}$	$\lambda_{APB}$
LSTM	17%	20%	21%	9%	16%	14%	7%	3%	4%	14%	19%	15%	16%	25%	19%	10%	20%	10%
DBN	58%	74%	71%	49%	71%	63%	44%	51%	53%	100%	139%	109%	129%	168%	138%	134%	160%	159%
RBF	6%	7%	6%	11%	11%	14%	20%	20%	19%	18%	16%	19%	19%	19%	18%	22%	27%	23%
MARS	13%	18%	17%	9%	16%	14%	5%	4%	9%	9%	11%	13%	15%	16%	17%	16%	18%	19%
WKNNR	45%	48%	54%	33%	46%	43%	21%	19%	20%	20%	28%	22%	26%	37%	28%	37%	47%	34%
GPML	37%	41%	42%	59%	76%	72%	29%	34%	34%	53%	71%	62%	63%	84%	73%	58%	66%	64%
BRFR	37%	34%	40%	33%	27%	33%	28%	25%	30%	31%	26%	33%	34%	27%	33%	33%	34%	38%
M5TREE	37%	34%	34%	32%	30%	33%	30%	28%	31%	40%	33%	40%	36%	33%	38%	35%	37%	38%

distribution acquired by the deep hybrid CSVr model for all sites is observed to be smaller as compared to standalone LSTM, MARS, WKNNR, RBF, M5TREE, BRFR, GPML and DBN. Consequently, the frequency distribution plot and the violin plot clearly shows that the CSVr accomplished the best predictive accuracy for all six solar farms. An analysis of the frequency distribution of prediction error  $|PE|$  produced by CSVr compared to benchmark methods is shown in Fig. 9. Each error bin contains the percentage of all tested points, which is shown at the top of each error bar. Each error bin has a size of  $1.5 \text{ MJ m}^{-2} \text{ day}^{-1}$ . Interestingly, GSR predictions using the deep hybrid CSVr model were found to have the greatest frequency of errors within the smallest error bracket ( $1.5 \text{ MJ m}^{-2} \text{ day}^{-1}$ ), which encompassed 55% of the test data. On the other hand, the LSTM, MARS, WKNNR, RBF, M5TREE, BRFR, GPML and DBN accumulated 48%, 44%, 40%, 40%, 36%, 32%, 28% and 24% respectively. In accordance with Fig. 8 and Tables 6–8, the CSVr model generates most of its error predictions within the lowest magnitude band, making it more accurate in predicting GSR at the six solar farms. Additionally, violin plot of prediction error  $|PE|$  for CSVr as well as the other DL models and the conventional ML models are created in Fig. 10. The  $|PE|$  error distribution acquired by the deep hybrid CSVr model for all sites is observed to be smaller as compared to standalone LSTM, MARS, WKNNR, RBF, M5TREE, BRFR, GPML and DBN. Consequently, the frequency distribution plot and the violin plot clearly shows that the CSVr accomplished the best predictive accuracy for all six solar farms.

Furthermore, to determine the enhanced performance of the models, promoting percentages are calculated as per Eqs. (19)–(21). Concurrent with earlier findings (Tables 6–8), the deep hybrid CSVr for all solar farms shows the better performance. For instance, at Daystar Energy Solar Farm the deep hybrid CSVr can significantly improve  $RRMSE$  over LSTM, DBN, RBF, MARS, WKNNR, GPML, BRFR, and M5TREE by 16%, 129%, 19%, 15%, 26%, 63%, 34% and 36%, respectively. Therefore, with positive promoting percentage errors, the deep hybrid CSVr model appears to possess advanced predictive capabilities that allow it to provide reliable prediction of GSR (see Table 9).

The effectiveness of the deep hybrid CSVr model was verified using  $KGE$  and the  $APB$ ,  $KGE$  is based on the decomposition of NSE into its component parts (correlation, variability bias and mean bias), and addresses several perceived shortcomings with NSE (Knoben et al., 2019). With a relatively high  $KGE$  and a comparatively low  $APB$ , the

results show that the deep hybrid CSVr predictive model outperforms the counterpart models significantly, as illustrated Fig. 11a. More accurately, the magnitudes of  $KGE$  and  $APB$  when comparing the deep hybrid CSVr model with the subsequent best prediction model, LSTM and MARS, in the grouping [CSVr: LSTM: MARS] were  $[0.900, 9.229 : 0.886, 10.536 : 0.886, 11.097]$  Furthermore, the GPI metrics are used, GPI captures the individual effects of all statistical tests and helps rank the developed GSR model. The GPI score for the developed GSR models lies between 2.721 and  $-4.271$ . The maximum value indicates the best model. Based on GPI, the best suitable model is deep hybrid CSVr ( $GPI \approx 2.721$ ) which ranks first among all the develop GSR models. Model LSTM ( $GPI \approx 1.774$ ) and MARS ( $GPI \approx 1.566$ ) comes in second and third respectively, Fig. 11b. Indeed, the highest value of GPI for the deep hybrid CSVr model cemented even further its accuracy in predicting daily GSR.

Additionally, to enables a more detailed assessment of the model results, a Taylor plot (Fig. 12) was established to compare the models' performance based on centered root-mean-square (RMS) difference and correlation coefficient ( $r$ ). Based on Taylor plot, the most efficient model is the deep hybrid CSVr, with the lowest centered RMS difference and a high  $r$ . In other words, the CSVr produced the best results among all the DL as well as conventional ML models developed. Even though comparing standard deviations between different models in this study is not significant since their average values differ, the CSVr model demonstrated acceptable variation of predicted values.

Furthermore, all predictive models were tested by using several statistical techniques, namely Diebold–Mariono (DM), Harvey, Leybourne, and Newbold (HLN) tests, for which the statistical significance of all of the predictive models was analyzed. The purpose of these tests is to deduce to see if the deep hybrid CSVr prediction model is more accurate than the predictions made by other comparison models (Table 10(a) and (b)). It should be noted that when comparing models in these tables, the model in the column is compared with the model in the row, and if there is a positive outcome, the model in the column would most likely outperform the one in the row while vice versa if the result is negative. Congruent with the previous findings and taken together the results of DM and HLN tests, we aver that the deep hybrid CSVr model can predict the daily GSR data more accurately than the other models.

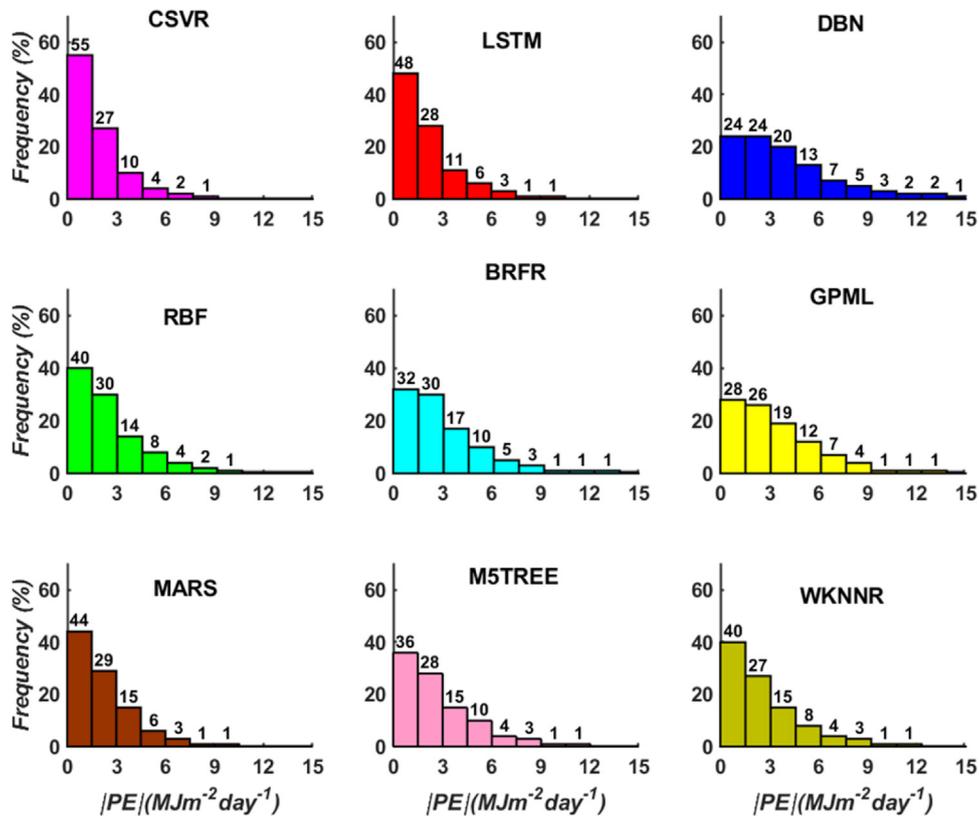


Fig. 9. Cumulative frequency of the forecast error (in  $\pm 1.5 \text{ MJ m}^{-2} \text{ day}^{-1}$  bracket) for all tested solar energy farms pooled together.

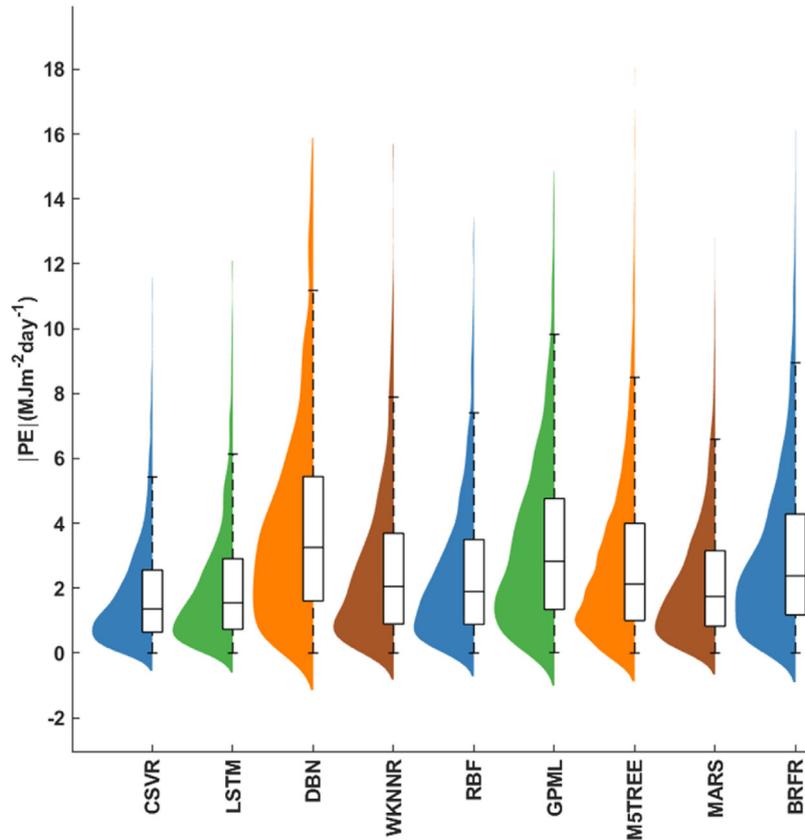


Fig. 10. Violin plot of the absolute forecast error of daily GSR generated by the proposed CSVR model in respect to eight other comparative models (LSTM, DBN, WKNNR, RBF, GPML, M5TREE, MARS and BRFR) within the testing phase.

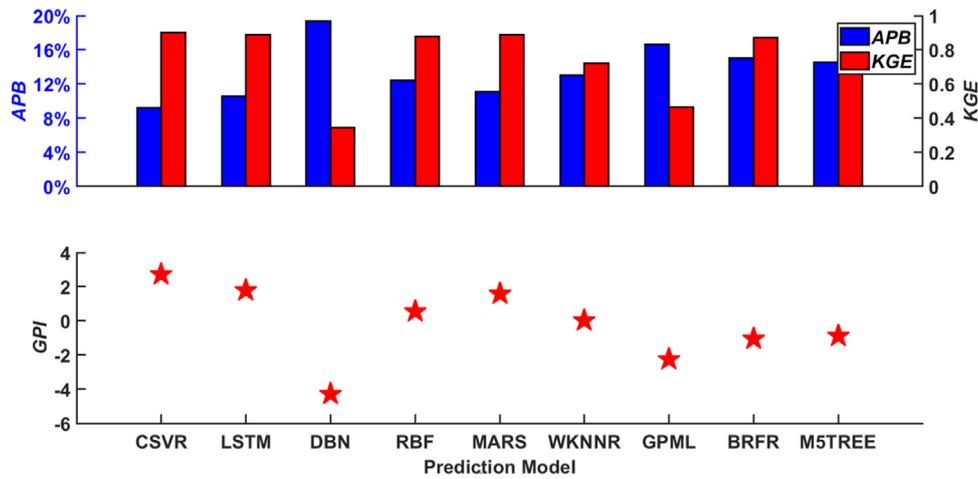


Fig. 11. (a) Bar chart comparing the efficacy of the proposed CSVR model in terms of the tested absolute percentage bias (*APB*, %) and Kling–Gupta efficiency (*KGE*). (b) Global performance indicator (*GPI*) used to evaluate the proposed CSVR model relative to eight other benchmarked models.

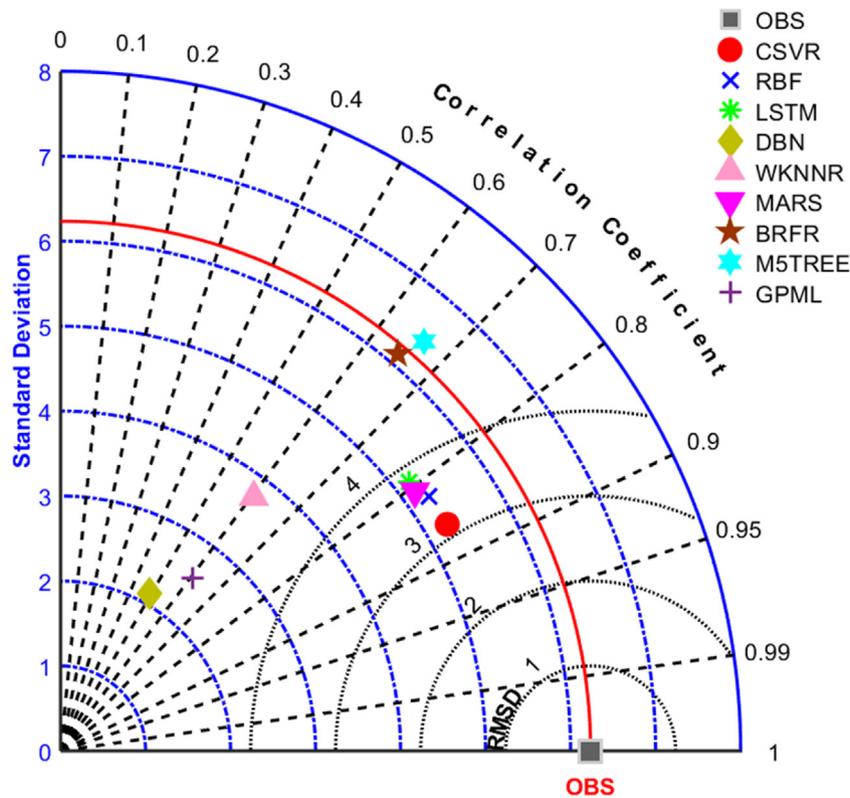


Fig. 12. Taylor diagram evaluating the proposed CSVR model in terms of bias and standard deviation of model errors. The azimuthal angle represents the correlation, the radial distance shows the standard deviation, and the semicircles centered at the observation “OBS” marker show the standard deviation of the errors. Color scale shows the bias (*i.e.*, mean of model minus mean of observation) and the names of each model are provided in Tables 3–5.

The *RMSE* values of the deep hybrid CSVR model, are now compared with the *RMSE* values of the model developed using only clear-sky index persistence measure (Marquez and Coimbra, 2013), denoted as the Skill Score (*SS*). Likewise, ratio of the *RMSE* of deep hybrid model CSVR with the *RMSE* of the other DL models and conventional ML models are computed as per Eq. (15) (Yang et al., 2020). Notably, all the comparative models appear to have a significantly lower *SS* and *RMSEr* relative to the deep hybrid CSVR predictive model as shown in Table 11a and b. It should be noted that, if  $SS > 0$  the prediction model have a smaller *RMSE* than that of the persistence model, otherwise,  $SS \leq 0$  indicates that the model of fails to outperform the persistence model. Similarly,  $RMSEr > 1$  would indicate that the objective model outperforms the comparative

model and  $RMSEr < 1$  indicates the comparative model is better. The proposed deep hybrid CSVR model appears to perform exceptionally well for all six solar farms, whereas the DBN, the WKNNR, the BRFR, and the GPML models produce very poor results.

As an additional evaluation of the deep hybrid CSVR predictive model, the data of all study sites are divided into four distinct seasons and the simulations are repeated for all models. Fig. 13 is a representation of the model in terms of the performance measures of *WI*, *NSE* and *KGE*, *RRMSE*, *RMAE* and *APB* for all four seasons. Concurrent with previous deductions for daily *GSR* predictions, the proposed deep hybrid SAELSTM model appears to register the best seasonal performance, with a lower value of *RRMSE*, *RMAE* and *APB* and a higher value of *WI*, *NSE* and *KGE* compared with equivalent metrics for the LSTM,

**Table 10**

Evaluation of CSV model against comparison models in terms of: (a) The Diebold–Mariano (DM) test statistic, (b) The Harvey–Leybourne–Newbold (HLN) test statistic. The column of the table is compared with the rows, and if the result is positive, the model in the rows outperforms the one in the column; on the contrary, if it is negative, then the one in the column is superior.

(a)									
Predictive model	CSV	LSTM	DBN	M5TREE	BRFR	RBF	GPML	MARS	WKNNR
<b>CSV</b>		<b>5.534</b>	<b>10.663</b>	<b>11.326</b>	<b>13.786</b>	<b>8.304</b>	<b>10.456</b>	<b>6.758</b>	<b>6.974</b>
LSTM			9.799	9.293	11.975	5.536	9.558	0.944	5.741
DBN				-4.815	-5.040	-8.043	-4.798	-9.304	-7.312
M5TREE					0.054	-4.496	2.125	-8.772	-3.876
BRFR						-5.167	2.251	-10.469	-4.063
RBF							6.050	-5.166	1.318
GPML								-8.364	-6.962
MARS									-4.632

(b)									
Predictive model	CSV	LSTM	DBN	M5TREE	BRFR	RBF	GPML	MARS	WKNNR
<b>CSV</b>		<b>5.600</b>	<b>10.791</b>	<b>11.462</b>	<b>13.951</b>	<b>8.403</b>	<b>10.582</b>	<b>6.839</b>	<b>7.057</b>
LSTM			9.917	9.404	12.119	5.602	9.673	0.955	5.810
DBN				-4.873	-5.100	-8.140	-4.855	-9.415	-7.400
M5TREE					0.055	-4.550	2.150	-8.877	-3.922
BRFR						-5.229	2.278	-10.594	-4.111
RBF							6.122	-5.228	1.334
GPML								-8.465	-7.046
MARS									-4.688

**Table 11**

Evaluation of CSV model and all other comparison models in terms of (a) Skill Score Metric (SS) 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). (b) The performance of the CSV model with comparative benchmark models in the test period measured by the ratio of root mean square error ( $RMSE_t$ ). The column of the table is compared with the rows, and if the result is  $<1$ , the model in the row outperforms the one in the column; on the contrary, if it is  $>1$ , then the one in the column is superior.

(a)										
Locations	CSV	LSTM	DBN	RBF	MARS	WKNNR	GPML	BRFR	M5TREE	
Aramara Solar Farm	0.462	0.369	0.098	0.408	0.388	0.185	0.161	0.154	0.146	
Childers DDN Green Solar Farm	0.498	0.455	0.231	0.410	0.444	0.316	0.094	0.203	0.241	
Cloncurry Solar Farm	0.549	0.516	0.336	0.419	0.518	0.449	0.392	0.378	0.361	
Crinum Creek Solar Farm	0.553	0.492	0.046	0.384	0.498	0.453	0.266	0.329	0.286	
Daystar Energy Solar Farm	0.591	0.527	-0.021	0.396	0.501	0.462	0.251	0.337	0.355	
Ewerleigh Solar Farm	0.625	0.587	0.073	0.417	0.532	0.454	0.309	0.396	0.415	

(b)										
Predictive model	CSV	LSTM	DBN	M5TREE	BRFR	RBF	GPML	MARS	WKNNR	
<b>CSV</b>		<b>1.268</b>	<b>3.762</b>	<b>2.390</b>	<b>2.397</b>	<b>1.696</b>	<b>2.764</b>	<b>1.315</b>	<b>1.890</b>	
LSTM	0.789		2.967	1.885	1.891	1.338	2.180	1.037	1.490	
DBN	0.266	0.337		0.636	0.637	0.451	0.735	0.350	0.502	
M5TREE	0.418	0.530	1.574		1.003	0.710	1.156	0.550	0.791	
BRFR	0.417	0.529	1.569	0.997		0.708	1.153	0.549	0.788	
RBF	0.590	0.747	2.217	1.409	1.413		1.629	0.775	1.114	
GPML	0.362	0.459	1.361	0.865	0.867	0.614		0.476	0.684	
MARS	0.760	0.964	2.860	1.817	1.823	1.290	2.102		1.437	
WKNNR	0.5292	0.6709	1.9905	1.2649	1.2685	0.8977	1.4627	0.696		

RBF, MARS, WKNNR, GPML, BRFR, and M5TREE. In accordance with this finding, we content that the deep hybrid CSV predictive model is deemed suitable for both daily and seasonal GSR predictions.

**5. Further discussion and predictive model comparison**

Accurate prediction of GSR is pivotal to the monitoring of electricity supply, management and utilization of solar energy, and it is also of paramount importance to the renewable energy industries. Although computational prediction of GSR has made much progress recently, the accuracy remains to be improved. In this study, we introduced a novel and interpretable deep learning framework named CNN-SVR for GSR prediction. Specifically, CNN works as a trainable feature extractor and SVR performs as a GSR predictor. Compared with other deep learning and ML models, CNN-SVR achieved the best performance. So

by engineering the two methods (SVR & CNN together), the resulting CNN-SVR framework was not only able to automatically extract GSR features, but also led to improved generalization capability of CNN and regression method’s accuracy.

Previous studies have suggested that hybrid models (Agga et al., 2022; Livera et al., 2018) that incorporate multiple DL and ML architectures can achieve higher accuracy than a single learner model. Inspired by this strategy, instead of using a single convolution network to train the model, we have engineered the CNN and SVR model. Compared with several state-of-the-art learning-based methods, CNN-SVR can effectively exploit deep features of the GSR time series. Experimental results demonstrated the predictive power of the proposed CNN-SVR for GSR predictions. Comparing deep hybrid CSV model with benchmark models, the accuracy of the proposed hybrid CSV model is notably higher based on a lower relative error and high-performance metric.

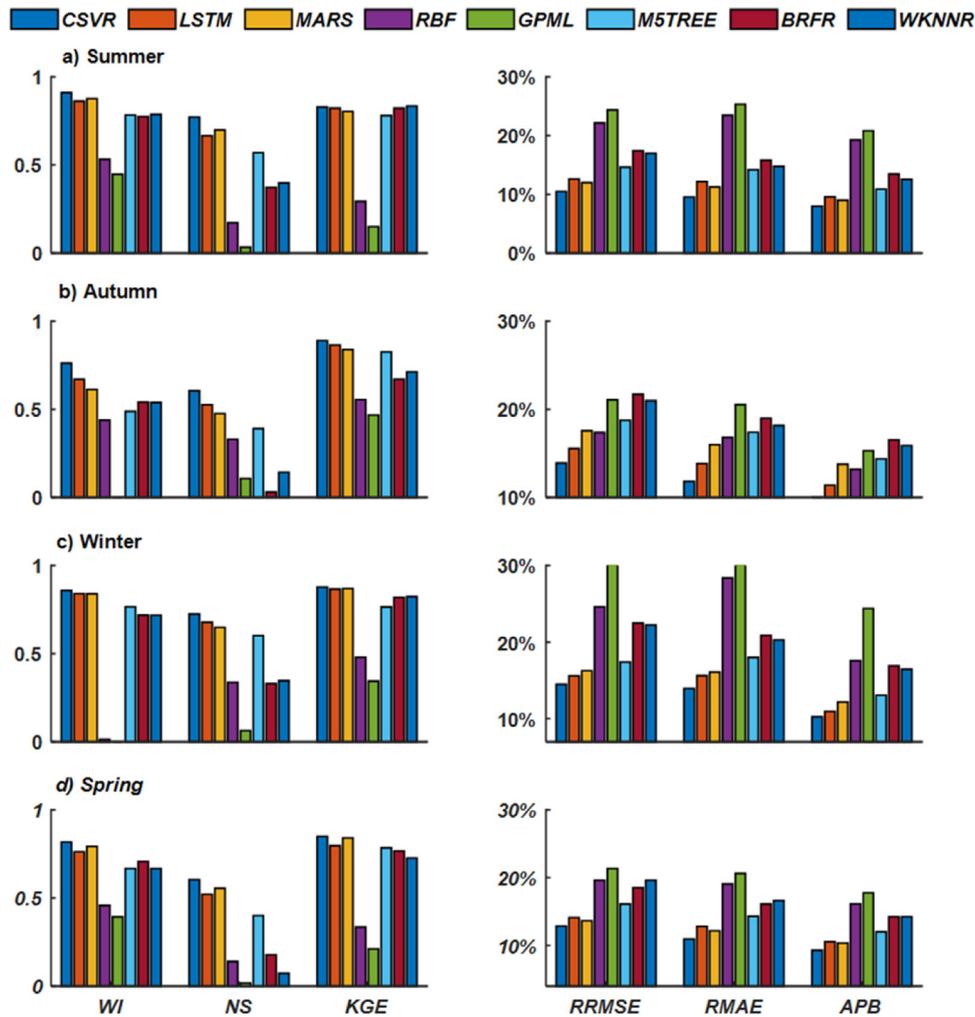


Fig. 13. Seasonal performance of the proposed CSV model compared with benchmark models in terms of Willmott's Index (*WI*), Nash-Sutcliffe Coefficient (*NS*), Kling Gupta efficiency (*KGE*), normalized root mean square error (*nRMSE*, %), Relative mean absolute error (*RMAE*, %) and absolute percentage bias (*APB*, %).

The deep hybrid CSV also demonstrated the lowest *RMSE*, *APB* and *MAE* for all study sites, as well as highest magnitude of the normalized performance metrics (*r*, *WI*, *NSE*, *KGE*, *LM*, *Evar*) thereby outperforming the DL based model LSTM and DBN, as well as the conventional ML models. Furthermore, the CSV model outperformed LSTM, DBN, RBF, BRFR, MARS, WKNNR, GPML, and M5TREE models, obtaining the lowest magnitude of *RRMSE* and *RMAE* at the six solar farms considered. The overall results revealed that the CNN-SVR model, using Atmospheric and Meteorological data, had a superior performance compared to the other ML as well as deep learning models.

To further ascertain the predictive capability of the proposed CSV, we now compare these results with recent studies conducted in Turkey and some of the other regions. Since the aim of our study was to estimate the daily GSR, we have considered only the daily basis studies from literature. In Salcedo-Sanz et al. (2018b) a robust hybrid modeling mechanism was developed for Brisbane and Sunshine Coast in Australia with the Interim-ERA European Centre for Medium-Range Weather Forecasting (ECMWF) Reanalysis data employed to train and cross-validate the model formulated by an evolutionary-type algorithm: CRO integrated with ELM model. The hybrid CRO-(ELM) algorithm is applied in two stages: first for FS process guided by an ELM algorithm (a class of fast training neural network tool) as the model's fitness function to screen an optimal set of predictor variables and second, for the estimation of the solar radiation using the optimally screened variables by the final hybrid CRO-(ELM)-ELM system. In terms of root mean square error and Willmott's Index, the hybrid CRO-(ELM)-ELM

was seen to outperform the non-hybrid ELM, MARS, SVR and MLR-based models with *RMSE* = 3.1877 MJ m<sup>-2</sup>, *RRMSE* = 16.0277 MJ m<sup>-2</sup> and *WI* = 0.8249 vs. 3.5990, 19.5328, 0.8072 for the ELM-based model, 3.6172, 19.6052, 8.122 for the MARS-based model, 3.4609, 19.0729, 0.8048 for the SVR-based model and 3.6787, 20.1047, 0.8024 for the MLR-based model.

The recent study of Zang et al. (2020a) proposed a deep learning method for estimating daily GSR which is constituted by embedding clustering (EC) and functional deep belief network (DBN) and reported to have *RMAE* of 13.71%. Furthermore, in Quej et al. (2017) ANN, SVM and Adaptive neuro-fuzzy interference system (ANFIS) models to estimate the daily GSR for five stations which are located in Yucatán Peninsula in Mexico. It was reported that SVM model had better performances than the ANN and ANFIS. The best values of *MAE*, *RMSE* and *R<sup>2</sup>*, respectively, were found to be as 1.97 MJ m<sup>-2</sup>, 2.68 MJ m<sup>-2</sup> and 0.689. Similarly in Bulut and Büyükalaca (2007) a simple model based on a trigonometric function was developed and it has only one dependent parameter, namely, day of the year, for estimating the daily global radiation. The correlation coefficient (*r*), *MAE* and *RMSE* were in the range of 0.620 and 0.900, 2.13 MJ m<sup>-2</sup> and 4.45 MJ m<sup>-2</sup> and 2.89 MJ m<sup>-2</sup> and 5.45 MJ m<sup>-2</sup>, respectively. Additionally, in Ghimire et al. (2019a), ANN model was developed to predict GSR at five different location of Australia. The study concluded that the performance of ANN was better than the other models (SVR, GPML, GP, TM), resulting in lower *RMSE* (1.715–2.27 MJ m<sup>-2</sup> relative to 2.14–5.90 MJ m<sup>-2</sup>), relative *RMSE* (9.07–12.47% vs. 10.98–29.15 m

%), relative *RMAE* (7.97–11.74% vs. 9.27–33.96%) and larger *WI*, *NS* (0.938–0.967 vs. 0.462–0.955, 0.935–0.872 vs. 0.355–0.915, 0.672–0.783 vs. 0.252–0.740). In conclusion, our results are better than the rest of above studies, except the study of Ghimire et al. (2019a) to which our results are comparable.

The results presented so far have indicated an excellent capability of the proposed CSVN model to predict solar radiation. It is noteworthy that the solar energy industry relies on accurate predictions of the atmospheric state considering the cloud movement, including factors like humidity, temperature, rainfall and cloud fraction that affect the ground solar radiation. Solar energy received at solar energy farms or their PV systems can be highly intermittent due to these atmospheric factors — therefore, accurate forecast models are required by solar industries. In this study, a daily-step CSVN model is trained with GCM predictor variables utilizing meteorological variables produced by a set of GCMs (i.e., cloud parameters, humidity parameters, precipitation, wind speed, etc.) and enriched by the ground based observation data (i.e., evaporation, vapor pressure, relative humidity at maximum temperature, relative humidity at minimum temperature, rainfall, maximum temperature and minimum temperature). Therefore, this trained model can be used to monitor solar radiation at industrial scale solar farms, at inter-daily, or weekly scales whereby accurate prediction of solar radiation can help energy companies in decision-making regarding the best ways to integrate solar energy into electricity grids.

## 6. Conclusion and future research directions

### 6.1. Conclusion

In this study, a robust version of a Convolutional Neural Network (CNN) integrated with Support Vector Regression (SVR) model has been developed to predict daily *GSR*. A comprehensive comparison of the *GSR* prediction performance of CSVN and other DL methods as well as conventional ML models is carried out. The CNN model is used to extract local pattern features as well as common features that recur in the time series at different intervals, and the SVR replaced the fully connected layer of the CNN to predict the daily *GSR* at six solar farms of Australia.

In order to develop the hybrid CSVN model and other comparative models, we have used the meteorological variables as input from the GCM and the SILO. Additionally, from the pool of GCM and SILO meteorological variables, optimal features for the input to develop the models are selected through a metaheuristic FS algorithm called ASO. The hyperparameter selection of the CSVN as well as other DL based models are carry out via HyperOpt. The performance of proposed deep hybrid CSVN model is benchmarked against eight other DL as well as other ML models (LSTM, DBN, RBF, BRN, MARS, WKNNR, GPML, and M5TREE).

Comparing deep hybrid CSVN model with benchmark models, the accuracy of the proposed deep hybrid CSVN model has been higher based on a low relative error and high-performance metric. Deep hybrid CSVN demonstrated the lowest *RMSE*, *APB* and *MAE* for all study sites, as well as highest magnitudes for normalized performance metrics (*r*, *WI*, *NSE*, *KGE*, *LM*, *Evar*) thereby outperforming the DL based models LSTM and DBN, as well as conventional ML models. Furthermore, the deep hybrid CSVN model outperformed the LSTM, DBN, RBF, BRN, MARS, WKNNR, GPML, and M5TREE models, obtaining the lowest magnitude of *RRMSE* and *RMAE* at the six solar farms considered.

A comparison of the model's prediction error when all six solar farms were pooled, revealed that the CSVN model generated the largest proportion of predicted error ( $\approx 55\%$ ) in the smallest ( $\pm 0.5 \text{ MJ m}^{-2} \text{ day}^{-1}$ ) error range, although the LSTM model also recorded about 48% and MARS model generated only 44% of all errors in this particular error bracket. Additionally, the promoting percentage error demonstrates that CSVN model was able to improve the predictive accuracy,

for instance, at Daystar Energy Solar Farm, the deep hybrid CSVN can significantly improve *RRMSE* over LSTM, DBN, RBF, MARS, WKNNR, GPML, BRN, and M5TREE by 16%, 129%, 19%, 15%, 26%, 63%, 34% and 36%, respectively and *RMAE* by 25%, 168%, 19%, 16%, 37%, 84%, 27%, and 33%, respectively. In this study, we clearly endorse the use of a deep hybrid CSVN model as an effective alternative tool for *GSR* prediction, and the tool may therefore be explored for use as an AI tool to advance energy exploration and planning. In the future, we would like to see whether the proposed model can predict solar radiation at real-time resolutions (e.g., sub-hourly or hourly) in addition to testing other kinds of hybrid time sequencing methods, such as Deep net approaches.

### 6.2. Limitations, additional scope, and opportunities for future research

It is crucial to mention that the evaluation of the proposed GSR prediction models based on past climate observations could have some important limitations such as being limited to those variables and phenomena for which observations exist. In many cases, the lack or insufficient quality of long-term observations, be it a specific variable, an important processes, or a particular region (e.g., polar areas, the upper troposphere/lower stratosphere (UTLS), and the deep ocean), remains an impediment. In addition, owing to observational uncertainties and the presence of internal variability, the observational record against which models are assessed is 'imperfect'. These limitations can be reduced, but not entirely eliminated, through the use of multiple independent observations of the same variable as well as the use of model ensembles. Elimination of these limitation is out of scope for this study and the approach to GSR prediction model evaluation taken in this study reflects the need for climate models to represent the observed behavior of past climate as a necessary condition to be considered a viable tool for future predictions. Additionally in the previous studies (Ghimire et al., 2019d; Ghimire, 2019) it has been reported that with the use of Aerosol and cloud properties the performance can be improved, so including such parameters during training can improve the accuracy of proposed model.

The increasing capacity to use freely-available solar energy into the renewable energy mix creates new problems in area of solar PC power systems engineering. The problem arises due to the increasing role in everyday electricity usage systems, with the proliferation of roof-top solar PV panels as well as new industrial-scale solar farms being built to integrate more energy into power grids. One main challenge for optimizing solar energy contributions is the intermittency in energy supply with well-documented reports suggesting the need for better forecasting of cloud cover impact. As solar radiation is significantly affected by cloud characteristics, the testing of the developed CSVN model under various cloud cover conditions remains to be seen. The need to monitor cloud cover in respect to ground-based solar radiation was identified in an Australian Government CSIRO report (Sayeef et al., 2012) stating that: "Without such a forecast system, wind and solar renewable energy generation will be subjected to increasing level of curtailment that can undermine their viability and significant contributions towards greenhouse gas reduction". It is therefore desirable that the proposed CSVN model is tested in a real-time PV power system to examine its capability to monitor the supply of electricity into power grids.

Although the proposed CSVN model was developed and tested for six solar farms, its ability to model the *GSR* in respect to the worst-case, rapidly changing cloud patterns, including cumulus (i.e., fast-moving, well-defined clouds with the clear sky in between) or a squall (i.e., solid line of dark clouds moving across clear sky) can help solar engineers to better enable an optimization of solar energy in electrical power grids. This is a key problem in solar engineering area where poor prediction of cloud cover shifts is a challenge to maximize its contribution to energy needs as stipulated in a recent energy industry report (Brinsmead et al., 2014). If the proposed CSVN model can also be employed to test its

capability to connect with solar PV systems for real-time cloud cover monitoring, there could be major economic benefits arising from the improved solar energy forecasting. One practical example could be applying the CSVR model in future renewable energy investment, such as the Australia–Asia Power Link (AAPowerLink) (ASEANPowerLink, 2020), a proposed electricity infrastructure planned to include the world’s largest solar plant, battery and longest submarine power cable in Australia. The AAPowerLink has the world’s largest solar farm in the Northern Territory in Australia with a 4500 km transmission system planned to supply solar power Darwin, Singapore, and the ASEAN power export markets with competitively-priced renewable electricity. Therefore, a future testing of the CSVR model at such large-scale solar farms could lead to a better understanding of the capability of our system to monitor cloud cover effects on the transported electricity in this system.

Recently, there has been some emphasis on using sky images and sequence-to-sequence models to capture not only the cloud cover effects but also the influence of water vapor, aerosols, ozone changes, and particulate matter, or dust on the electrical energy received at a solar PV panel (Ghimire et al., 2022; Prasad et al., 2022). To increase the scope of the proposed CSVR model and to better incorporate these effects into solar radiation forecast models or in energy demand monitoring, a future research should consider sky images as inputs to build more explainable predictive system whereby for example, determine how a particular image pixel with a specific cloud feature can contribute to an outcome and to what extent it influences the solar energy captured by a solar PV panel. A future study can therefore use deep Taylor decomposition methods to decompose the CSVR predictions of GSR in terms of contributions of individual input features at a pixel level capturing water vapor, aerosols, ozone, particulate matter, or dust effects. These effects captured at the input relevance level (e.g. cloud shifts) from the output layer of the CSVR model can help understand contributing neurons in terms of total relevance of each image feature to produce an explainable CSVR model for solar PV engineering and related energy forecasting problems.

### CRedit authorship contribution statement

**Sujan Ghimire:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. **Binayak Bhandari:** Conceptualization, Validation, Writing – review & editing. **David Casillas-Pérez:** Conceptualization, Supervision, Writing – review & editing. **Ravinesh C. Deo:** Conceptualization, Supervision, Validation, Investigation, Writing – review & editing. **Sancho Salcedo-Sanz:** Conceptualization, Supervision, Investigation, Writing – review & editing.

### Declaration of competing interest

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.

### Acknowledgments

The data were obtained from Scientific Information for Landowners (SILO) repository managed by Queensland Government (Queensland Climate Change Centre of Excellence) in Department of Science, Information Technology, Innovation and the Arts (DSITIA). The GCM data were sourced from Centre for Environmental Data Analysis (CEDA) that hosts the CMIP5 project. Daily atmospheric model outputs for historical period were sourced for the GCM models (CSIRO-BOM ACCESS1-0, MOHC Hadley-GEM2-CC and MRI MRI-CGCM3) all of 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), Spain.

**Table A.12**

List of acronyms of the metrics.

Acronym	Full name
APB	Absolute Percentage Bias
DM	Diebold–Mariano
GPI	Global Performance Indicator
HLN	Harvey, Leybourne, and the Newbold
KGE	Kling–Gupta Efficiency
LM	Legates’ Modulus
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
NSE	Nash–Sutcliffe’s Efficiency
$r$	Correlation Coefficient
RMAE	Relative Mean Absolute Error
RMSE	Root Mean Square Error
RRMSE	Relative Root Mean Square Error
WI	Willmott’s Index

**Table A.13**

List of acronyms.

Acronym	Full name
Adam	Adaptive Moment Estimation
AI	Artificial Intelligence
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Networks
ASO	Atom Search Optimization
BILSTM	Bidirectional Long Short-Term Memory
CNN	Convolutional Neural Network
CSVR	Our hybrid approach CNN with SVR
DBNs	Deep Belief Networks
DL	Deep Learning
DNN	Deep Neural Network
ELM	Extreme Learning Machine
ESN	Echo State Networks
FS	Feature Selection
GA	Genetic Algorithm
GCM	Global Climate Models
GRU	Gated Recurrent Unit
GSR	Global Solar Radiation
K-NN	Kernel Nearest-Neighbor
LSTM	Long Short-Term Memory Networks
MLP	Multilayer Perceptrons
ML	Machine Learning
PV	PhotoVoltaic
PSO	Particle Swarm Optimization
RF	Random Forest
RNN	Recurrent Neural Networks
SA	Simulated Annealing
SILO	Queensland Government’s Scientific Information for Landowners
SVR	Support Vector Regression
WEO	World Energy Outlook

## Appendix A. Acronyms

Table A.12 shows a list of alphabetically ordered acronyms used for the metrics used in the present paper. Table A.13 shows a list of alphabetically ordered acronyms for the methods that appear in this paper.

## Appendix B. Theoretical details — Atom Search Optimization

Like other metaheuristic algorithms (PSO, GA, SA), ASO generates an initial population of solutions called atoms. In ASO, each atom maintains two vectors, namely position and velocity as follows:

$$X_i = (X_i^1, X_i^2, \dots, X_i^D) \quad (\text{B.1})$$

$$V_i = (V_i^1, V_i^2, \dots, V_i^D) \quad (\text{B.2})$$

where  $X_i$  and  $V_i$  are the position and velocity of the  $i$ th atom, and  $D$  is the maximum number of dimensions. The acceleration of an atom has a major impact on its movement. The atom's acceleration can be expressed mathematically as:

$$a_i = \frac{F_i + G_i}{m_i} \quad (\text{B.3})$$

where  $F$  is the interaction force,  $G$  is the constraint force, and  $m$  is the mass of the atom.

The interaction force between the  $i$ th atom and  $j$ th atom  $d$ th dimension at  $t$  time is drawn from the Lennard-Jones (L-J) potential as follow (Zhao et al., 2019b):

$$F_{ij}(t) = \frac{24\epsilon(t)}{\sigma(t)} \left( 2 \left( \frac{\sigma(t)}{r_{ij}(t)} \right)^{13} - \left( \frac{\sigma(t)}{r_{ij}(t)} \right)^7 \right) \frac{r_{ij}(t)}{r_{ij}^d(t)} \quad (\text{B.4})$$

$$F'_{ij}(t) = \frac{24\epsilon(t)}{\sigma(t)} \left( 2 \left( \frac{\sigma(t)}{r_{ij}(t)} \right)^{13} - \left( \frac{\sigma(t)}{r_{ij}(t)} \right)^7 \right) \quad (\text{B.5})$$

where  $F'$  is the model of the interaction force,  $\epsilon$  is the depth of potential,  $\sigma$  is the length scale,  $r$  is the distance between two atoms (atom  $i$  and atom  $j$ ), and  $d$  is the dimension of search space. As for optimization, a simplified version of Eq. (B.3) is designed as:

$$F'_{ij}(t) = -\eta(t) \left( 2 (h_{ij}(t))^{13} - (h_{ij}(t))^7 \right) \quad (\text{B.6})$$

where the function  $h$  is the scaled distance between two atoms (Hekimoğlu, 2019) and  $\eta$  is the depth function to regulate the attraction or repulsion region, and defined as:

$$h_{ij}(t) = \begin{cases} h_{min} & \frac{r_{ij}(t)}{\sigma(t)} < h_{min} \\ \frac{r_{ij}(t)}{\sigma(t)} & h_{min} \leq \frac{r_{ij}(t)}{\sigma(t)} \leq h_{max} \\ h_{max} & \frac{r_{ij}(t)}{\sigma(t)} > h_{max} \end{cases} \quad (\text{B.7})$$

$$\eta = \alpha \left( 1 - \frac{t-1}{T} \right)^3 e^{-\frac{20t}{T}} \quad (\text{B.8})$$

where  $\alpha$  is the depth weight,  $T$  is the maximum number of iterations, and,  $h_{max}$  and  $h_{min}$  are the upper boundary and lower boundary of  $h$ , respectively (Zhao et al., 2019b), calculated as:

$$\begin{cases} h_{min} = g_0 + g(t) \\ h_{max} = u \end{cases} \quad (\text{B.9})$$

where  $g_0$  is the lowest limit set to 1.1 and  $u$  is the upper limit set to 1.24, and  $g(t)$  is the drift factor to make the algorithm capable of drifting from exploration to exploitation, which is given as:

$$g(t) = 0.1 \sin\left(\frac{\pi t}{2T}\right); \quad \text{where } t \text{ is the current iteration.} \quad (\text{B.10})$$

The length scale  $\sigma$  is calculated as:

$$\sigma(t) = \left\| x_{ij}(t) \frac{\sum_{j \in \mathcal{K}_{best}} x_{ij}(t)}{K(t)} \right\|_2 \quad (\text{B.11})$$

where,  $\mathcal{K}_{best}$  is a subset of an atom population, which is made up of the first  $K$  atoms with the best function fitness values.

Then the sum of components with random weights in the  $d$ th dimension acting on the  $i$ th atom from the other atoms can be considered as a total force, which is expressed as

$$F_i^d(t) = \sum_{j \in \mathcal{K}_{best}} U_j F_{ij}^d(t) \quad (\text{B.12})$$

where  $U_j$  is a random number uniformly chosen over the interval  $[0, 1]$ .

According to Newton's third law, the force on the  $j$ th atom for the same pairwise interaction is the opposite of force on the  $i$ th atom:

$$F_{ij} = -F_{ji} \quad (\text{B.13})$$

Furthermore, the geometric constraint of the  $i$ th atom and constraint force which is the weighted position difference between each atom and the best atom can be expressed as follows:

$$G_i^d(t) = \lambda(t) (x_{best}^d(t) - x_i^d(t)) \quad (\text{B.14})$$

where  $x_{best}^d(t)$  is the position of the best atom in  $d$ th dimension and  $\lambda(t)$  is the Lagrangian multiplier, which is defined as:

$$\lambda(t) = \beta e^{-\frac{20t}{T}} \quad (\text{B.15})$$

where  $\beta$  is the multiplier weight. At last, the acceleration of the atom can be written as:

$$a_i^d(t) = \frac{F_i^d(t)}{m_i^d(t)} + \frac{G_i^d(t)}{m_i^d(t)} \quad (\text{B.16})$$

$$a_i^d(t) = -\alpha \left( 1 - \frac{t-1}{T} \right)^3 e^{-\frac{20t}{T}} \sum_{j \in \mathcal{K}_{best}} \frac{U_j \left( 2 (h_{ij}(t))^{-13} - (h_{ij}(t))^{-7} \right) (x_j^d(t) - x_i^d(t))}{m_i(t) \|\vec{x}_i(t), \vec{x}_j(t)\|_2} + \beta e^{-\frac{20t}{T}} \frac{(x_{best}^d(t) - x_i^d(t))}{m_i(t)} \quad (\text{B.17})$$

where  $m_i^d(t)$  is the mass of the  $i$ th atom in  $d$ th dimension at iteration  $t$ , and calculated by its fitness function value as follows:

$$M_i(t) = e^{-\frac{Fit_i(t) - Fit_{best}(t)}{Fit_{worst}(t) - Fit_{best}(t)}} \quad (\text{B.18})$$

$$m_i(t) = \frac{M_i(t)}{\sum_{j=1}^N M_j(t)}$$

where  $Fit_i(t)$  is the fitness function value of  $i$ th atom at iteration  $t$ ,  $Fit_{best}(t)$  and  $Fit_{worst}(t)$  are the fitness values of the best and worst atoms at iteration  $t$ , respectively, and defined as:

$$Fit_{best}(t) = \min_{i \in \{1, 2, \dots, N\}} Fit_i(t) \quad (\text{B.19})$$

$$Fit_{worst}(t) = \max_{i \in \{1, 2, \dots, N\}} Fit_i(t)$$

Finally, the velocity and position update of  $i$ th atom at iteration  $(t+1)$  are defined as:

$$v_i^d(t+1) = U_i^d v_i^d(t) + a_i^d(t) \quad (\text{B.20})$$

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1)$$

where  $x_i^d$  and  $v_i^d$  are the position and velocity of the  $i$ th atom,  $a$  is the acceleration,  $d$  is the dimension of search space,  $U_i^d$  is a random vector uniformly chosen in  $[0, 1]$ , and  $t$  is current iteration.

In ASO, the number of best atoms in subset  $K$  is used to balance the exploration and exploitation phase.

$$k = N - (N-2) \sqrt{\frac{t}{T}} \quad (\text{B.21})$$

where  $N$  is the number of atoms in the population. Initially, a higher value of  $k$  enables the atoms to explore the untried areas. At the end of the iteration, a lower value of  $k$  promotes exploitation, which performs the search around the best solutions.

The flowchart of ASO is shown in Fig. B.14.

### Appendix C. Theoretical details — Support vector regression

For a training dataset  $\{(x_1, y_1), \dots, (x_n, y_n) | x_i \in \mathbb{R}^D, y_i \in \mathbb{R}\}$ , where  $\mathbb{R}^D$  is a  $D$ -dimensional real input vector,  $y_i \in \mathbb{R}$  is the corresponding

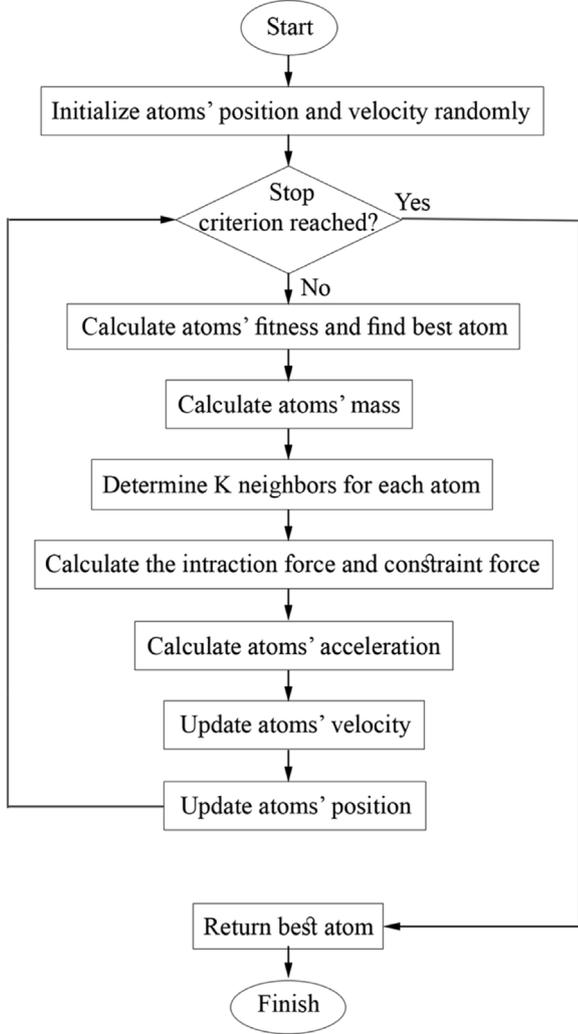


Fig. B.14. Flowchart of the ASO FS used to select predictands to train the objective CSVR model.

target value, and  $n$  is the total number of data patterns, the regression function of the SVR model is expressed as follows:

$$f(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b \quad (\text{C.1})$$

where  $\mathbf{w} \in \mathbb{R}^D$  is a weight vector,  $T$  stands for the transpose operator. The term  $b$  is a bias,  $\phi(\cdot)$  is a nonlinear transfer function mapping the input vectors into a high dimensional feature space.

The slack variables  $\xi_i$  and  $\xi_i^*$  are defined to address infeasible constraints. The SVR algorithm's optimization problem can be expressed using Eq. (2):

$$\text{Subject to } \begin{cases} f(\mathbf{x}_i) - y_i \leq \epsilon + \xi_i \\ y_i - f(\mathbf{x}_i) \leq \epsilon + \xi_i^* \\ \xi_i \geq 0 \quad i = 1, 2, \dots, n \\ \xi_i^* \geq 0 \quad i = 1, 2, \dots, n \end{cases} \quad (\text{C.2})$$

Again, transform the objective function into the unconstrained Lagrange objective function as Eq. (C.1):

$$L(\mathbf{w}, b, \alpha, \alpha^*, \xi, \xi^*, \nu, \nu^*) = \frac{1}{2} \mathbf{w}^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) - \sum_{i=1}^n \mu_i \xi_i - \sum_{i=1}^n \mu_i^* \xi_i^* + \sum_{i=1}^n a_i (f(\mathbf{x}_i) - y_i - \epsilon - \xi_i) + \sum_{i=1}^n a_i^* (y_i - f(\mathbf{x}_i)) - \epsilon - \xi_i^* \quad (\text{C.3})$$

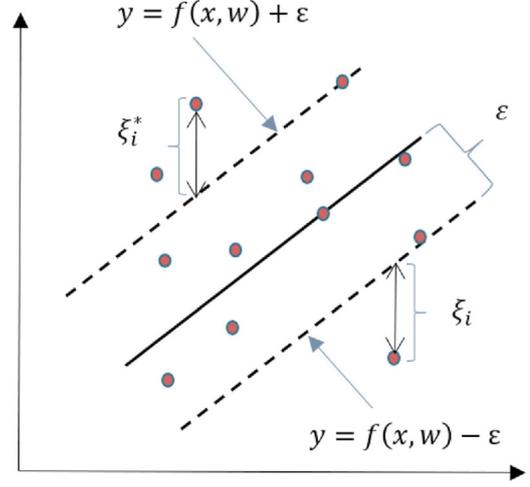


Fig. C.15. Schematic representation of the SVR algorithm showing the slack variable ( $\xi$ ). The mathematical symbols are outlined in Section 2.3 along with the SVR model equations.

where the Lagrange multipliers are  $a_i \geq 0$ ,  $a_i^* \geq 0$ ,  $\mu_i \geq 0$ , and  $\mu_i^* \geq 0$ .  $C$  is the punishment factor for the SVR.

Let the partial derivative of Eq. (C.3) be 0 and introduce the solution back into Eq. (C.3), the dual SVR algorithm problem can be expressed by using Eq. (C.4).

$$\max_{\alpha, \alpha^*} \left( \sum_{i=1}^n y_i (a_i^* - a_i) - \epsilon (a_i^* + a_i) - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (a_i^* - a_i) (a_j^* - a_j) \mathbf{x}_i^T \mathbf{x}_j \right) \quad (\text{C.4})$$

$$\text{s.t. } \begin{cases} \sum_{i=1}^n (a_i^* - a_i) = 0 \\ 0 \leq a_i, a_i^* \leq C \end{cases} \quad (\text{C.5})$$

The dual problem should satisfy the Karush–Kuhn–Tucker condition as follows:

$$\text{s.t. } \begin{cases} a_i (f(\mathbf{x}_i) - y_i - \epsilon - \xi_i) = 0 \\ a_i^* (y_i - f(\mathbf{x}_i) - \epsilon - \xi_i^*) = 0 \\ (C - a_i) \xi_i = 0, (C - a_i^*) \xi_i^* = 0 \\ a_i a_i^* = 0 \end{cases} \quad (\text{C.6})$$

Finally, the SVR solution can be expressed using Eq. (C.7),

$$f(\mathbf{x}) = \sum_{i=1}^n (a_i^* - a_i) \mathbf{x}_i^T \mathbf{x} + b. \quad (\text{C.7})$$

The inner product  $\mathbf{x}_i^T \mathbf{x}$  can be replaced by the so-called kernel function  $K(x_i, x)$  under Mercer's condition. Therefore, the final form of SVR function can be expressed using Eq. (C.8).

$$f(\mathbf{x}) = \sum_{i=1}^n (a_i^* - a_i) K(\mathbf{x}_i, \mathbf{x}) + b. \quad (\text{C.8})$$

The kernel function, as can be seen from Eq. (C.8), plays a critical role in the SVR algorithm. In SVR model, polynomial, sigmoid, linear and radial basis function (RBF) can be used as kernel function. In this study RBF kernel function was chosen because of its *a*) capability of modeling nonlinear relationships by mapping data points from the input space into a high dimensional feature space in a nonlinear fashion, *b*) compare to polynomial and sigmoid kernels, RBF needs less customizable parameters, making it straightforward and functional (Keerthi and Lin, 2003) and *c*) RBF's superior performance has been demonstrated in several literature (Keerthi and Lin, 2003; Dibike et al., 2001).

The kernel function RBF is expressed as Eq. (C.9):

$$K(\mathbf{x}_i, \mathbf{x}) = e^{-\frac{\|\mathbf{x}_i - \mathbf{x}\|^2}{2\sigma^2}} \quad (\text{C.9})$$

where  $\sigma$  is variance and  $\|x_i - x\|$  is the Euclidean distance ( $L_2$ -norm) between two points  $x_i$  and  $x$ .

$$K(x_i, x) = e^{-\gamma \|x_i - x\|^2} \quad (\text{C.10})$$

The RBF Kernel Support Vector Machines has two hyperparameters associated with it,  $C$  for SVR and  $\gamma$  for the RBF Kernel. Here,  $\gamma$  is inversely proportional to  $\sigma$  and can be expressed as below.

$$\gamma = \frac{1}{2\sigma^2} \quad (\text{C.11})$$

## References

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al., 2016. Tensorflow: A system for large-scale machine learning. In: 12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16). pp. 265–283.
- Ağbulut, U., Gürel, A.E., Biçen, Y., 2021. Prediction of daily global solar radiation using different machine learning algorithms: Evaluation and comparison. *Renew. Sustain. Energy Rev.* 135, 110114.
- Agga, A., Abbou, A., Labbadi, M., El Houm, Y., Ali, I.H.O., 2022. CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production. *Electr. Power Syst. Res.* 208, 107908.
- Ahn, H.K., Park, N., 2021. Deep RNN-based photovoltaic power short-term forecast using power IoT sensors. *Energies* 14 (2), 436.
- Al-Hajj, R., Assi, A., Fouad, M., Mabrouk, E., 2021. A hybrid LSTM-based genetic programming approach for short-term prediction of global solar radiation using weather data. *Processes* 9 (7), 1187.
- Al-Musaylh, M.S., Deo, R.C., Adamowski, J.F., Li, Y., 2018a. Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in queensland, Australia. *Adv. Eng. Inform.* 35, 1–16.
- Al-Musaylh, M.S., Deo, R.C., Li, Y., Adamowski, J.F., 2018b. 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. *Appl. Energy* 217, 422–439.
- Alizamir, M., Kim, S., Zounemat-Kermani, M., Heddam, S., Shahrabadi, A.H., Gharabaghi, B., 2021. Modelling daily soil temperature by hydro-meteorological data at different depths using a novel data-intelligence model: deep echo state network model. *Artif. Intell. Rev.* 54 (4), 2863–2890.
- Alsharif, M.H., Younes, M.K., Kim, J., 2019. Time series ARIMA model for prediction of daily and monthly average global solar radiation: The case study of Seoul, South Korea. *Symmetry* 11 (2), <http://dx.doi.org/10.3390/sym11020240>, URL <https://www.mdpi.com/2073-8994/11/2/240>.
- ASEANPowerLink, 2020. Australia–ASEAN Power Link Project Environment Protection Act (NT) Referral, 754-MELEN232735. Report Prepared by Sun Cable Pty Ltd.
- Aybar-Ruiz, A., Jiménez-Fernández, S., Cornejo-Bueno, L., Casanova-Mateo, C., Sanz-Justo, J., Salvador-González, P., Salcedo-Sanz, S., 2016. A novel grouping genetic algorithm–extreme learning machine approach for global solar radiation prediction from numerical weather models inputs. *Sol. Energy* 132, 129–142.
- Basaran, K., Özçift, A., Kılınc, D., 2019. A new approach for prediction of solar radiation with using ensemble learning algorithm. *Arab. J. Sci. Eng.* 44 (8), (Springer Science & Business Media BV).
- Bendali, W., Saber, I., Bourachdi, B., Boussetta, M., Mourad, Y., 2020. 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, pp. 1–8.
- Bengio, Y., 2009. *Learning Deep Architectures for AI*. Now Publishers Inc.
- Bergstra, J., Yamins, D., Cox, D.D., et al., 2013. Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. In: *Proceedings of the 12th Python in Science Conference*, Vol. 13. Citeseer, p. 20.
- Bhandari, B., Lee, K.-T., Lee, C.S., Song, C.-K., Maskey, R.K., Ahn, S.-H., 2014a. A novel off-grid hybrid power system comprised of solar photovoltaic, wind, and hydro energy sources. *Appl. Energy* 133, 236–242.
- Bhandari, B., Poudel, S.R., Lee, K.-T., Ahn, S.-H., 2014b. Mathematical modeling of hybrid renewable energy system: A review on small hydro-solar-wind power generation. *Int. J. Precis. Eng. Manuf.-Green Technol.* 1 (2), 157–173.
- Brinsfield, R., Yaramanoglu, M., Wheaton, F., 1984. Ground level solar radiation prediction model including cloud cover effects. *Sol. Energy* 33 (6), 493–499.
- Brinsmead, T.S., Hayward, J., Graham, P., 2014. Australian Electricity Market Analysis Report to 2020 and 2030. CSIRO Technical Report No. EP141067.
- Brownlee, J., 2016. Time series prediction with lstm recurrent neural networks in python with keras. *Mach. Learn. Mastery*.
- Bulut, H., Büyükalaca, O., 2007. Simple model for the generation of daily global solar-radiation data in Turkey. *Appl. Energy* 84 (5), 477–491.
- Cao, J., Lin, X., 2008. Application of the diagonal recurrent wavelet neural network to solar irradiation forecast assisted with fuzzy technique. *Eng. Appl. Artif. Intell.* 21 (8), 1255–1263.
- Castangia, M., Aliberti, A., Bottaccioli, L., Macii, E., Patti, E., 2021. A compound of feature selection techniques to improve solar radiation forecasting. *Expert Syst. Appl.* 178, 114979.
- Cavalli, S., Amoretti, M., 2021. CNN-based multivariate data analysis for bitcoin trend prediction. *Appl. Soft Comput.* 101, 107065.
- 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.
2020. Centre for Environmental Data Analysis. CEDA Archive.
- Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* 7 (3), 1247–1250.
- Chen, C.-W., Tsai, Y.-H., Chang, F.-R., Lin, W.-C., 2020. Ensemble feature selection in medical datasets: Combining filter, wrapper, and embedded feature selection results. *Expert Syst.* 37 (5), e12553.
- Chollet, F., et al., 2017. Keras (2015).
- Cornejo-Bueno, L., Casanova-Mateo, C., Sanz-Justo, J., Salcedo-Sanz, S., 2019. Machine learning regressors for solar radiation estimation from satellite data. *Sol. Energy* 183, 768–775.
- Costantini, M., Pappalardo, C., 2008. Combination of Forecast Methods using Encompassing Tests: An Algorithm-Based Procedure. *Tech. rep., Reihe Ökonomie/Economics Series*.
- de Araujo, J.M.S., 2020. Performance comparison of solar radiation forecasting between WRF and LSTM in Gifu, Japan. *Environ. Res. Commun.* 2 (4), 045002.
- Del Ser, J., Casillas-Perez, D., Cornejo-Bueno, L., Prieto-Godino, L., Sanz-Justo, J., Casanova-Mateo, C., Salcedo-Sanz, S., 2021. Randomization-based machine learning in renewable energy prediction problems: Critical literature review, new results and perspectives. *arXiv preprint arXiv:2103.14624*.
- Deo, R., Wen, X., Qi, F., 2016. A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset. *Appl. Energy* 168, 568–593.
- Despotovic, M., Nedic, V., Despotovic, D., Cveticanovic, S., 2015. Review and statistical analysis of different global solar radiation sunshine models. *Renew. Sustain. Energy Rev.* 52, 1869–1880.
- Dibike, Y.B., Velickov, S., Solomatine, D., Abbott, M.B., 2001. Model induction with support vector machines: introduction and applications. *J. Comput. Civ. Eng.* 15 (3), 208–216.
- Diebold, F.X., Mariano, R.S., 2002. Comparing predictive accuracy. *J. Bus. Econom. Statist.* 20 (1), 134–144.
- Gao, B., Huang, X., Shi, J., Tai, Y., Zhang, J., 2020. Hourly forecasting of solar irradiance based on CEEMDAN and multi-strategy CNN-LSTM neural networks. *Renew. Energy* 162, 1665–1683.
- García-Hinde, O., Terrén-Serrano, G., Hombrados-Herrera, M., Gómez-Verdejo, V., Jiménez-Fernández, S., Casanova-Mateo, C., Sanz-Justo, J., Martínez-Ramón, M., Salcedo-Sanz, S., 2018. Evaluation of dimensionality reduction methods applied to numerical weather models for solar radiation forecasting. *Eng. Appl. Artif. Intell.* 69, 157–167.
- Garud, K.S., Jayaraj, S., Lee, M.-Y., 2021. A review on modeling of solar photovoltaic systems using artificial neural networks, fuzzy logic, genetic algorithm and hybrid models. *Int. J. Energy Res.* 45 (1), 6–35.
- Ghimire, S., 2019. Predictive Modelling of Global Solar Radiation with Artificial Intelligence Approaches using MODIS Satellites and Atmospheric Reanalysis Data for Australia (Ph.D. thesis). University of Southern Queensland.
- Ghimire, S., Deo, R.C., Downs, N.J., Raj, N., 2018. Self-adaptive differential evolutionary extreme learning machines for long-term solar radiation prediction with remotely-sensed MODIS satellite and Reanalysis atmospheric products in solar-rich cities. *Remote Sens. Environ.* 212, 176–198.
- Ghimire, S., Deo, R.C., Downs, N.J., Raj, N., 2019a. Global solar radiation prediction by ANN integrated with European Centre for medium range weather forecast fields in solar rich cities of Queensland Australia. *J. Cleaner Prod.* 216, 288–310.
- Ghimire, S., Deo, R.C., Raj, N., Mi, J., 2019b. Deep learning neural networks trained with MODIS satellite-derived predictors for long-term global solar radiation prediction. *Energies* 12 (12), 2407.
- Ghimire, S., Deo, R.C., Raj, N., Mi, J., 2019c. Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms. *Appl. Energy* 253, 113541.
- Ghimire, S., Deo, R.C., Raj, N., Mi, J., 2019d. Wavelet-based 3-phase hybrid SVR model trained with satellite-derived predictors, particle swarm optimization and maximum overlap discrete wavelet transform for solar radiation prediction. *Renew. Sustain. Energy Rev.* 113, 109247.
- Ghimire, S., Deo, R.C., Wang, H., Al-Musaylh, M.S., Casillas-Pérez, D., Salcedo-Sanz, S., 2022. Stacked LSTM sequence-to-sequence autoencoder with feature selection for daily solar radiation prediction: A review and new modeling results. *Energies* 15 (3), 1061.
- Goldsborough, P., 2016. A tour of tensorflow. *arXiv preprint arXiv:1610.01178*.
- Guijo-Rubio, D., Durán-Rosal, A., Gutiérrez, P., Gómez-Orellana, A., Casanova-Mateo, C., Sanz-Justo, J., Salcedo-Sanz, S., Hervás-Martínez, C., 2020. Evolutionary artificial neural networks for accurate solar radiation prediction. *Energy* 210, 118374.
- Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.* 377 (1–2), 80–91.

- Hai, T., Sharafati, A., Mohammed, A., Salih, S.Q., Deo, R.C., Al-Ansari, N., Yaseen, Z.M., 2020. Global solar radiation estimation and climatic variability analysis using extreme learning machine based predictive model. *IEEE Access* 8, 12026–12042.
- He, H., Lu, N., Jie, Y., Chen, B., Jiao, R., 2020. Probabilistic solar irradiance forecasting via a deep learning-based hybrid approach. *IEEE Trans. Electr. Electron. Eng.* 15 (11), 1604–1612.
- Hekimoğlu, B., 2019. Optimal tuning of fractional order PID controller for DC motor speed control via chaotic atom search optimization algorithm. *IEEE Access* 7, 38100–38114.
- Huang, X., Zhang, C., Li, Q., Tai, Y., Gao, B., Shi, J., 2020. A comparison of hour-ahead solar irradiance forecasting models based on LSTM network. *Math. Probl. Eng.* 2020.
- Husein, M., Chung, I.-Y., 2019. Day-ahead solar irradiance forecasting for microgrids using a long short-term memory recurrent neural network: A deep learning approach. *Energies* 12 (10), 1856.
- International Energy Agency (IEA), 2020. *World Energy Outlook Paris. Annual Report*, <https://www.iea.org/reports/world-energy-outlook-2020>.
- Jaihuni, M., Basak, J.K., Khan, F., Okyere, F.G., Arulmozhi, E., Bhujel, A., Park, J., Hyun, L.D., Kim, H.T., 2020. A partially amended hybrid bi-GRU—ARIMA model (PAHM) for predicting solar irradiance in short and very-short terms. *Energies* 13 (2), 435.
- Jaihuni, M., Basak, J.K., Khan, F., Okyere, F.G., Sihalath, T., Bhujel, A., Park, J., Lee, D.H., Kim, H.T., 2021. A novel recurrent neural network approach in forecasting short term solar irradiance. *ISA Trans.*
- Jia, K., Ruan, Y., Yang, Y., You, Z., 2019. Assessment of CMIP5 GCM simulation performance for temperature projection in the Tibetan Plateau. *Earth Space Sci.* 6 (12), 2362–2378.
- Jiang, H., 2017. A novel approach for forecasting global horizontal irradiance based on sparse quadratic RBF neural network. *Energy Convers. Manage.* 152, 266–280.
- Ju, Y., Sun, G., Chen, Q., Zhang, M., Zhu, H., Rehman, M.U., 2019. A model combining convolutional neural network and LightGBM algorithm for ultra-short-term wind power forecasting. *IEEE Access* 7, 28309–28318.
- Jumin, E., Basaruddin, F.B., Yusoff, Y.B.M., Latif, S.D., Ahmed, A.N., 2021. Solar radiation prediction using boosted decision tree regression model: A case study in Malaysia. *Environ. Sci. Pollut. Res.* 28 (21), 26571–26583.
- Karaman, O.A., Ağır, T.T., Arsel, I., 2021. Estimation of solar radiation using modern methods. *Alex. Eng. J.* 60 (2), 2447–2455.
- Kawaguchi, K., Kaelbling, L.P., Bengio, Y., 2017. Generalization in deep learning. *arXiv preprint arXiv:1710.05468*.
- Keerthi, S.S., Lin, C.-J., 2003. Asymptotic behaviors of support vector machines with Gaussian kernel. *Neural Comput.* 15 (7), 1667–1689.
- Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Knoben, W.J., Freer, J.E., Woods, R.A., 2019. Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrol. Earth Syst. Sci.* 23 (10), 4323–4331.
- Komer, B., Bergstra, J., Eliasmith, C., 2014. Hyperopt-sklearn: automatic hyperparameter configuration for scikit-learn. In: *ICML Workshop on AutoML*, Vol. 9. Citeseer, p. 50.
- Kumar, D., Mathur, H., Bhanot, S., Bansal, R.C., 2021. Forecasting of solar and wind power using LSTM RNN for load frequency control in isolated microgrid. *Int. J. Model. Simul.* 41 (4), 311–323.
- Kwon, S., et al., 2020. A CNN-assisted enhanced audio signal processing for speech emotion recognition. *Sensors* 20 (1), 183.
- LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D., 1989. Backpropagation applied to handwritten zip code recognition. *Neural Comput.* 1 (4), 541–551.
- Legates, D.R., McCabe, Jr., G.J., 1999. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* 35 (1), 233–241.
- Li, Q., Wu, Z., Ling, R., Feng, L., Liu, K., 2020a. Multi-reservoir echo state computing for solar irradiance prediction: A fast yet efficient deep learning approach. *Appl. Soft Comput.* 95, 106481.
- Li, Q., Wu, Z., Zhang, H., 2020b. Spatio-temporal modeling with enhanced flexibility and robustness of solar irradiance prediction: a chain-structure echo state network approach. *J. Cleaner Prod.* 261, 121151.
- Li, C., Zhang, Y., Zhao, G., Ren, Y., 2021. Hourly solar irradiance prediction using deep BiLSTM network. *Earth Sci. Inform.* 14 (1), 299–309.
- Lima, M.A.F., Fernández Ramírez, L.M., Carvalho, P., Batista, J.G., Freitas, D.M., 2022. A comparison between deep learning and support vector regression techniques applied to solar forecast in Spain. *J. Solar Energy Eng.* 144 (1).
- Lin, T., Guo, T., Aberer, K., 2017. Hybrid neural networks for learning the trend in time series. In: *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*. pp. 2273–2279.
- Liu, H., Mi, X., Li, Y., 2018. Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network. *Energy Convers. Manage.* 166, 120–131.
- Livera, A., Theristis, M., Makrides, G., Sutterluetti, J., Georghiou, G.E., 2018. Advanced diagnostic approach of failures for grid-connected photovoltaic (PV) systems. In: *35th European Photovoltaic Solar Energy Conference (EU PVSEC)*.
- Ma, W., Zhou, X., Zhu, H., Li, L., Jiao, L., 2021. A two-stage hybrid ant colony optimization for high-dimensional feature selection. *Pattern Recognit.* 116, 107933.
- Marquez, R., Coimbra, C.F., 2013. Proposed metric for evaluation of solar forecasting models. *J. Solar Energy Eng.* 135 (1), 011016.
- Martin, D.A., 2019. Linking fire and the United Nations sustainable development goals. *Sci. Total Environ.* 662, 547–558.
- McKenzie, J., 2011. Mean absolute percentage error and bias in economic forecasting. *Econom. Lett.* 113 (3), 259–262.
- Mead, I., 2017. *International Energy Outlook 2017*. US Energy Information Administration.
- Met Office Hadley Centre, 2012. WCRP CMIP5: Met Office Hadley Centre (MOHC) HadGEM2-CC Model Output Collection. Centre for Environmental Data Analysis. 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.
- Moldovan, D., Slowik, A., 2021. Energy consumption prediction of appliances using machine learning and multi-objective binary grey wolf optimization for feature selection. *Appl. Soft Comput.* 111, 107745.
- Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* 50 (3), 885–900.
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* 10 (3), 282–290.
- Ordóñez, F.J., Roggen, D., 2016. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors* 16 (1), 115.
- Pan, T., Wu, S., Dai, E., Liu, Y., 2013. Estimating the daily global solar radiation spatial distribution from diurnal temperature ranges over the Tibetan Plateau in China. *Appl. Energy* 107, 384–393.
- Pang, Z., Niu, F., O’Neill, Z., 2020. Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons. *Renew. Energy* 156, 279–289.
- Park, J., Moon, J., Jung, S., Hwang, E., 2020. Multistep-ahead solar radiation forecasting scheme based on the light gradient boosting machine: A case study of Jeju Island. *Remote Sens.* 12 (14), 2271.
- Piri, J., Shamshirband, S., Petković, D., Tong, C.W., ur Rehman, M.H., 2015. Prediction of the solar radiation on the earth using support vector regression technique. *Infrared Phys. Technol.* 68, 179–185.
- Prado-Rujas, I.-I., García-Dopico, A., Serrano, E., Pérez, M.S., 2021. A flexible and robust deep learning-based system for solar irradiance forecasting. *IEEE Access* 9, 12348–12361.
- Prasad, S.S., Deo, R.C., Downs, N., Igoe, D., Parisi, A.V., Soar, J., 2022. Cloud affected solar UV predictions with three-phase wavelet hybrid convolutional long short-term memory network multi-step forecast system. *IEEE Access*.
- Quej, V.H., Almorox, J., Arnaldo, J.A., Saito, L., 2017. ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment. *J. Atmos. Sol.-Terr. Phys.* 155, 62–70.
- Rai, A., Shrivastava, A., Jana, K.C., 2020. A CNN-BiLSTM based deep learning model for mid-term solar radiation prediction. *Int. Trans. Electr. Energy Syst.* e12664.
- Ramedani, Z., Omid, M., Keyhani, A., Shamshirband, S., Khoshnevisan, B., 2014. Potential of radial basis function based support vector regression for global solar radiation prediction. *Renew. Sustain. Energy Rev.* 39, 1005–1011.
- Rodríguez, F., Azcárate, I., Vadillo, J., Galarza, A., 2022. Forecasting intra-hour solar photovoltaic energy by assembling wavelet based time-frequency analysis with deep learning neural networks. *Int. J. Electr. Power Energy Syst.* 137, 107777.
- Salcedo-Sanz, S., Casanova-Mateo, C., Pastor-Sánchez, A., Sánchez-Girón, M., 2014a. Daily global solar radiation prediction based on a hybrid coral reefs optimization—Extreme learning machine approach. *Sol. Energy* 105, 91–98.
- Salcedo-Sanz, S., Cornejo-Bueno, L., Prieto, L., Paredes, D., García-Herrera, R., 2018a. Feature selection in machine learning prediction systems for renewable energy applications. *Renew. Sustain. Energy Rev.* 90, 728–741.
- Salcedo-Sanz, S., Del Ser, J., Landa-Torres, I., Gil-López, S., Portilla-Figueras, J., 2014b. The coral reefs optimization algorithm: a novel metaheuristic for efficiently solving optimization problems. *Sci. World J.* 2014.
- Salcedo-Sanz, S., Deo, R.C., Cornejo-Bueno, L., Camacho-Gómez, C., Ghimire, S., 2018b. An efficient neuro-evolutionary hybrid modeling mechanism for the estimation of daily global solar radiation in the Sunshine State of Australia. *Appl. Energy* 209, 79–94.
- Salcedo-Sanz, S., Ghamisi, P., Piles, M., Werner, M., Cuadra, L., Moreno-Martínez, A., Izquierdo-Verdiguier, E., Muñoz-Marí, J., Mosavi, A., Camps-Valls, G., 2020. Machine learning information fusion in Earth observation: A comprehensive review of methods, applications and data sources. *Inf. Fusion* 63, 256–272.
- Salcedo-Sanz, S., Prado-Cumplido, M., Pérez-Cruz, F., Bousño-Calzón, C., 2002. Feature selection via genetic optimization. In: *International Conference on Artificial Neural Networks*. Springer, pp. 547–552.
- Salcedo-Sanz, S., Rojo-Álvarez, J.L., Martínez-Ramón, M., Camps-Valls, G., 2014c. Support vector machines in engineering: an overview. *Wiley Interdiscip. Rev.: Data Min. Knowl. Discov.* 4 (3), 234–267.
- Sayeeef, S., Heslop, S., Cornforth, D., Moore, T., Percy, S., Ward, J., Berry, A., Rowe, D., 2012. Solar intermittency: Australia’s clean energy challenge. Characterising the effect of high penetration solar intermittency on Australian electricity networks.

- Shen, Z., Rossel, R.V., 2021. Automated spectroscopic modelling with optimised convolutional neural networks. *Sci. Rep.* 11 (1), 1–12.
- Sivanand, R., Singh, J., Ali, A., Khan, A., Hoda, M.F., 2021. A comparative study of different deep learning models for mid-term solar power prediction. In: 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE). IEEE, pp. 519–524.
- Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. *Stat. Comput.* 14 (3), 199–222.
- Solorio-Fernández, S., Carrasco-Ochoa, J.A., Martínez-Trinidad, J.F., 2016. A new hybrid filter-wrapper feature selection method for clustering based on ranking. *Neurocomputing* 214, 866–880.
- Sun, S., Chen, W., Wang, L., Liu, X., Liu, T.-Y., 2016. On the depth of deep neural networks: A theoretical view. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 30.
- Sun, S., Qiao, H., Wei, Y., Wang, S., 2017. A new dynamic integrated approach for wind speed forecasting. *Appl. Energy* 197, 151–162.
- Syarifudin, M.A., Novitasari, D.C.R., Marpaung, F., Wahyudi, N., Hapsari, D.P., Supriyati, E., Farida, Y., Amin, F.M., Nugraheni, R.D., Nariswari, R., et al., 2021. Hotspot prediction using 1D convolutional neural network. *Procedia Comput. Sci.* 179, 845–853.
- Tian, C., Ma, J., Zhang, C., Zhan, P., 2018. A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network. *Energies* 11 (12), 3493.
- Vijayakumar, P., Kumaresan, G., Gokul Karthik, A., 2021. Optimization and prediction of pulsating heat pipe compound parabolic solar collector performances by hybrid DBN based bald eagle search optimizer. *Environ. Prog. Sustain. Energy* e13740.
- Wang, F., Yu, Y., Zhang, Z., Li, J., Zhen, Z., Li, K., 2018. Wavelet decomposition and convolutional LSTM networks based improved deep learning model for solar irradiance forecasting. *Appl. Sci.* 8 (8), 1286.
2021. List of Solar Farms in Queensland - Wikipedia. [Enwikipediaorg](https://en.wikipedia.org/wiki/List_of_Solar_Farms_in_Queensland).
- Willmott, C.J., Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* 30 (1), 79–82.
- Works DoEaP, 2021. Achieving Our Renewable Energy Targets. Queensland Government.
- Xie, G., Shangguan, A., Fei, R., Ji, W., Ma, W., Hei, X., 2020. Motion trajectory prediction based on a CNN-LSTM sequential model. *Sci. China Inf. Sci.* 63 (11), 1–21.
- Yan, C., Ma, J., Luo, H., Patel, A., 2019. Hybrid binary coral reefs optimization algorithm with simulated annealing for feature selection in high-dimensional biomedical datasets. *Chemometr. Intell. Lab. Syst.* 184, 102–111.
- Yang, D., Alessandrini, S., Antonanzas, J., Antonanzas-Torres, F., Badescu, V., Beyer, H.G., Blaga, R., Boland, J., Bright, J.M., Coimbra, C.F., et al., 2020. Verification of deterministic solar forecasts. *Sol. Energy* 210, 20–37.
- Yeom, J.-M., Deo, R.C., Adamowski, J.F., Park, S., Lee, C.-S., 2020. Spatial mapping of short-term solar radiation prediction incorporating geostationary satellite images coupled with deep convolutional LSTM networks for South Korea. *Environ. Res. Lett.* 15 (9), 094025.
- Zahedi, A., 2010. Australian renewable energy progress. *Renew. Sustain. Energy Rev.* 14 (8), 2208–2213.
- Zang, H., Cheng, L., Ding, T., Cheung, K.W., Wang, M., Wei, Z., Sun, G., 2020a. Application of functional deep belief network for estimating daily global solar radiation: A case study in China. *Energy* 191, 116502.
- Zang, H., Liu, L., Sun, L., Cheng, L., Wei, Z., Sun, G., 2020b. Short-term global horizontal irradiance forecasting based on a hybrid CNN-LSTM model with spatiotemporal correlations. *Renew. Energy* 160, 26–41.
- Zeng, Z., Wang, Z., Gui, K., Yan, X., Gao, M., Luo, M., Geng, H., Liao, T., Li, X., An, J., et al., 2020. Daily global solar radiation in China estimated from high-density meteorological observations: A random forest model framework. *Earth Space Sci.* 7 (2), e2019EA001058.
- Zhao, W., Wang, L., Zhang, Z., 2019a. Atom search optimization and its application to solve a hydrogeologic parameter estimation problem. *Knowl.-Based Syst.* 163, 283–304.
- Zhao, W., Wang, L., Zhang, Z., 2019b. A novel atom search optimization for dispersion coefficient estimation in groundwater. *Future Gener. Comput. Syst.* 91, 601–610.
- Zhou, Y., Lin, J., Guo, H., 2021a. Feature subset selection via an improved discretization-based particle swarm optimization. *Appl. Soft Comput.* 98, 106794.
- Zhou, Y., Liu, Y., Wang, D., Liu, X., Wang, Y., 2021b. A review on global solar radiation prediction with machine learning models in a comprehensive perspective. *Energy Convers. Manage.* 235, 113960.
- Ziyabari, S., Du, L., Biswas, S., 2020. A spatio-temporal hybrid deep learning architecture for short-term solar irradiance forecasting. In: 2020 47th IEEE Photovoltaic Specialists Conference (PVSC). IEEE, pp. 0833–0838.
- Zjavka, L., 2020. Photo-voltaic power daily predictions using expanding PDE sum models of polynomial networks based on operational calculus. *Eng. Appl. Artif. Intell.* 89, 103409.