



# A comprehensive review of solar irradiation estimation and forecasting using artificial neural networks: data, models and trends

Naima El-Amarty<sup>1</sup> · Manal Marzouq<sup>2</sup> · Hakim El Fadili<sup>2</sup> · Saad Dosse Bennani<sup>1</sup> · Antonio Ruano<sup>3,4</sup>

Received: 20 September 2022 / Accepted: 12 November 2022 / Published online: 23 November 2022  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

## Abstract

Solar irradiation data are imperatively required for any solar energy-based project. The non-accessibility and uncertainty of these data can greatly affect the implementation, management, and performance of photovoltaic or thermal systems. Developing solar irradiation estimation and forecasting approaches is an effective way to overcome these issues. Practically, prediction approaches can help anticipate events by ensuring good operation of the power network and maintaining a precise balance between the demand and supply of the power at every moment. In the literature, various estimation and forecasting methods have been developed. Artificial Neural Network (ANN) models are the most commonly used methods in solar irradiation prediction. This paper aims to firstly review, analyze, and provide an overview of different aspects required to develop an ANN model for solar irradiation prediction, such as data types, data horizon, data preprocessing, forecasting horizon, feature selection, and model type. Secondly, a highly detailed state of the art of ANN-based approaches including deep learning and hybrid ANN models for solar irradiation estimation and forecasting is presented. Finally, the factors influencing prediction model performances are discussed in order to propose recommendations, trends, and outlooks for future research in this field.

**Keywords** Solar irradiation · Climate condition · Feature selection · ANN model · Forecasting horizon · Deep learning

## Introduction

The consumption of electrical energy has recently increased worldwide at a very high rate due to the fast growth of urbanization, improving life quality, and growing population. Meanwhile, facing the increasing depletion of fossil fuels and reducing the massive amount of greenhouse gases, most countries are moving toward the use of renewable

energies as an alternative to generating power, and to ensure pollution-free sustainable global electricity (Das et al. 2018; Jumin et al. 2021; Singla et al. 2022a). The low price and the universal availability of solar energy make it the most important resource among the available renewables. Solar energy comes from the sun to the earth surface in the form of solar irradiation. As known, there are two components of solar irradiation: beam normal irradiance (BNI) and diffuse horizontal irradiation (DHI). These components are the basic information for many solar system applications, including site analysis, site selection, technologies installation, optimal system design, and plant operation. Moreover, it is indispensable for the integration of an important amount of solar energy production technologies in buildings or electrical grids (Castangia et al. 2021a).

Forecasting solar power production is critical in power plant management. It helps greatly grid operators by creating plans to anticipate risks, maintenance schedules, and balancing both generation and demand at every moment (Diagne et al. 2012; Notton et al. 2019), from very short-term to long term going through short term and medium term. However, forecasting accurate solar power system outputs is highly

---

Responsible Editor: Philippe Garrigues

✉ Naima El-Amarty  
naima.elamarty@usmba.ac.ma

<sup>1</sup> SIGER, Faculty of Sciences and Technologies of Fez, Sidi Mohamed Ben Abdellah University, P.O. Box 2202, Fez, Morocco

<sup>2</sup> LIPI, Faculty of Sciences and Technologies of Fez, Sidi Mohamed Ben Abdellah University, P.O. Box 5206, Fez, Morocco

<sup>3</sup> DEEI, Faculty of Science & Technology, University of Algarve, 8005-294 Faro, Portugal

<sup>4</sup> IDMEC, Instituto Superior Técnico, Universidade de Lisboa, 1950-044 Lisbon, Portugal

conditioned by the correct modeling of solar irradiation parameter given as input (Blaga et al. 2019a, b; Huang et al. 2021; Voyant et al. 2017) since the uncertainty and variability of solar irradiation characteristics, caused by weather conditions, climate types, time of the day or night, and seasonal variability (Kumari and Toshniwal 2021b; Lan et al. 2019) lead to the fluctuations in power generation.

However, the availability of accurate data of global solar irradiation and its components present several challenges because the regular updates and maintenance of measurement instruments such as pyranometers and pyrhemometers which are relatively complex and expensive (Ibrahim and Khatib 2017; Kleissl 2013; Zhang et al. 2017). Moreover, solar irradiation data are rarely measured in most meteorological stations over the world. Hence, an alternative way to obtain these data is to propose prediction models for locations with only accessible parameters (Sun et al. 2015).

The estimation and forecasting of solar irradiation is a challenging task, as it depends completely on the geographical position and weather conditions of studied sites. In the literature, the prediction of solar irradiation can be performed using various techniques. These techniques can be mainly categorized into six classes for forecasting: persistence, Cloud motion tracking, Numerical weather predictions (NWP), classical statistical, machine learning (ML), and hybrid methods (Blaga et al. 2019a, b). For estimation, four classes can be considered: empirical, physical, statistical, and machine learning models (Zhou et al. 2021).

Physical models aim to explore the physical state of solar irradiation and other meteorological conditions using mathematical equations (Ramadhan et al. 2021). Their complex structure needs a huge amount of data to calibrate the complex dynamics of the atmosphere.

The statistical methods aim to measure the relationship between the historical value of solar irradiation and weather parameters, by applying the statistical analysis of the different input parameters and make predictions about solar irradiance. Generally, these methods are adopted for short-term forecasting of solar irradiation (Voyant et al. 2017). These models are limited to solve more complex prediction problems in the case of longer forecasting horizons (Diagne et al. 2013).

The empirical models purpose to develop a nonlinear or linear regression equation (Jiang 2009). They are commonly used to correlate solar irradiation with various measured meteorological and geographical parameters. Empirical models are simple and easy; nevertheless, their accuracy is usually limited due to uncertain variables (Gürel et al. 2020).

The persistence model assumes that global irradiance at an hour ahead or a day ahead is the best forecast by its value at previous hour or previous day respectively (Diagne et al. 2013). This model acts as a reference model in the solar irradiation forecasting community, especially for short-term

horizon (Kumar et al. 2020). In the field of solar irradiation forecasting, the persistence model is used to evaluate the quality of the proposed approaches (Blaga et al. 2019a, b).

Cloud motion tracking based-models consist of two processes: cloud detection and solar irradiance forecasting. These approaches based on sky images obtained from either ground or satellite cameras. Because solar irradiation is strongly influenced by cloud pattern, detection of cloud's motion leads to prediction of cloud positions that subsequently allows the forecast of solar irradiance (Kamadinata et al. 2019).

Numerical weather predictions (NWP) simulate the physical state of atmospheric conditions using mathematical models aiming to predict weather in a future times, based on the current climate conditions (Blaga et al. 2019a, b), (Huang et al. 2021).

Recently, machine-learning methods such as Support Vector Machine (SVM), Extreme Learning Machine (ELM), and Artificial Neural Networks (ANN) have been shown to handle linear, non-linear, and non-stationary data forms. They can learn and solve the complex nonlinear relationships between inputs and outputs (Zhou et al. 2021).

Finally, the hybrid methods purpose to combine two or more techniques to design a forecasting model. These models become popular in the last years due to their ability to provide better performances than a standalone model for different prediction problems by combining the strength and the benefits of each model.

The most popular machine learning methods used in solar irradiation estimation and forecasting are ANN models due to their ability to solve complex, uncertain and non-linear problems and require fewer experimental parameters to generate the input/output relationships (Yadav and Chandel 2014), (Blaga et al. 2019a, b). Furthermore, ANN models are capable to deal with many input parameters that make them more accurate and reliable (Qazi et al. 2015).

Lately, many interesting research studies have summarized and studied the literature related to the prediction of solar irradiation based on artificial neural network models. In fact, authors in (Yadav and Chandel 2014) presented a survey on ANN models used to estimate global and daily solar irradiation. Furthermore, this study also reviewed papers dedicated to forecasting solar irradiation in the short-term. This work presented a summary of suitable methods available in the literature in order to identify research gaps.

The research work in (Qazi et al. 2015) presents a systematic review based on an artificial neural network for solar systems design and solar irradiation prediction. The results of this study show that the performance of solar irradiation prediction models depends on input parameters and ANN architectures. In (Rajagukguk et al. 2020), a review of deep learning models to predict solar irradiation and photovoltaic power using time-series data is presented. This study reviewed three

single models long short-term memory (LSTM), recurrent neural network (RNN), gated recurrent unit (GRU), and one hybrid model convolutional neural network-LSTM (CNN-LSTM). The results show that LSTM provides the best performance. However, CNN-LSTM outperforms the three single models. In a review of Kumari and Toshniwal (Kumari and Toshniwal 2021a), the authors presented a detailed study of deep learning models for solar irradiation forecasting, this review proved that deep learning improves prediction performance over other machine learning models. Furthermore, deep hybrid models such as CNN-LSTM outperform standalone models such as LSTM, CNN, and GRU.

Finally, in the review of Guermoui et al. in (Guermoui et al. 2020), a holistic summary is provided of six hybrid model techniques used recently in solar irradiation prediction. The purpose of this study was to compare the different hybrid models proposed in the literature. The aim is to identify the promising and potential approaches for solar radiation prediction and covers an opportunity for future researchers in this area.

From the studied literature and to the author's knowledge, no paper review gathering ANN model types (including deep learning and hybrid ANN models) has been published up to date. Most of the published reviews in the field of solar irradiation prediction focused on a specific approach like ANN, deep learning model, or a general topic such as hybrid models approaches. The purpose of the present paper is not only to make a comprehensive study of the literature but also to conduct, synthetically and clearly, an exhaustive analysis linked to both estimation and forecasting problems based on ANN models (including deep learning and hybrid ANNs).

This systematic review will help researchers to obtain an overview of the various ANN approaches developed over the last 6 years. Our work is devoted to surrounding all-important aspects required in ANN model development for prediction problems in solar irradiation. The aim is to help interested researchers have a clear vision in this particular field. In this context, the main contributions and novelty of our work can be summarized as follows:

- Unlike some important review papers dealing with multiple machine learning techniques (Zhou et al. 2021), (Voyant et al. 2017), our review is focusing on the most used machine learning model, namely ANNs. Our objective is to update the state of art research up to 2022,
- Presenting a clear distinction between estimation and forecasting terminologies. In fact, regarding the existing papers in the literature, several authors often confuse the meaning of “estimation”, “prediction” and “forecasting”. Indeed, most available works in the literature do not indicate any difference between them,
- Providing an overview of different kinds of data types, data horizons, data pre-processing and feature selection,

- Gathering definitions from the literature of the four forecasting horizons: very short-term, short-term, medium-term, long-term,
- Classifying retained papers into single and hybrid models,
- Establishing the highly detailed state of the art with different columns, including ANN architecture, training/testing data percentage, size of the input matrix, forecasting horizon and results of comparison methods,
- Besides single ANN models, hybrid ANNs models are also reviewed and summarized,
- Evaluating the existing challenges, which should be solved for solar irradiation estimation and forecasting, proposing some important recommendations from the analysis of the established state of the art that should be considered in incoming studies and identifying the recent trends.

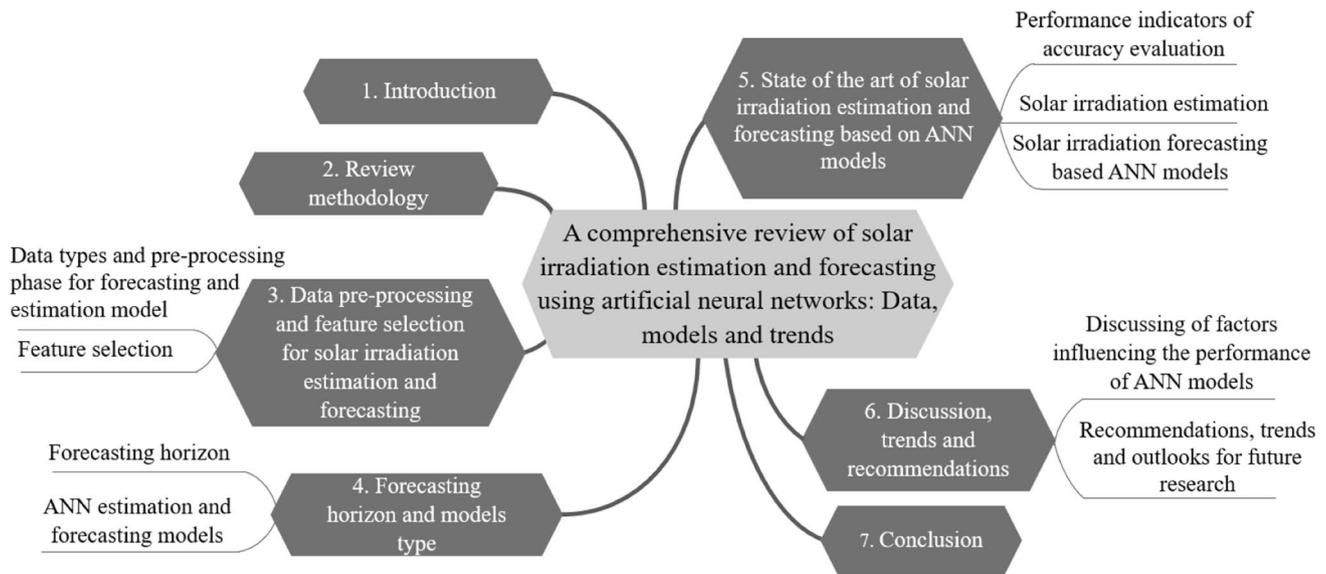
Accordingly, the remainder of this review is organized as follows: in section 2, we illustrate the review methodology. Section 3 includes two subsections: The first one discusses briefly data, including data types, data horizon, and data preprocessing. The second provides an overview of feature selection methods and their application in solar irradiation prediction. Section 4 illustrates a detailed analysis of forecasting horizon and gives a resume of ANN models investigated in our paper. Section 5 presents summaries of different models proposed in the literature for solar irradiation estimation and forecasting based on ANN models and hybrid ANN models.

A detailed discussion of factors influencing the performance of ANN models with recommendations, outlooks, and trends is given in Sect. 6. Finally, the main conclusions are provided in Sect. 7. Figure 1 shows a diagrammatic representation of our review structure.

## Review methodology

The purpose of our review is to study and analyze high quality papers in the field of solar irradiation prediction in order to make a relevant and useful synthesis of previous studies aiming to extract the existing gaps in this particular field. This synthesis will be a valuable opening for future researchers working on solar irradiation estimation and forecasting, particularly when ANN approaches are used as a prediction technique. As shown in Fig. 2, the flowchart of our review consists of seven primary steps.

- **Specification of period time, language, libraries, and keywords for our review:** initially, the reviewed papers



**Fig. 1** Diagrammatic representation of our paper structure

were collected by carrying out online research in well-reputed digital libraries such as IEEE Xplore, Science direct, ResearchGate, Elsevier, and the Google Scholar search engine in order to search for relevant papers published in the years from 2015 to 2022. This primary search was performed using keywords that are illustrated in the last columns of Table 1.

- **Limit the database for our search:** the obtained papers database was limited by searching for keywords in the title of the article, abstract, and investigated keywords. We also included studies from lists of references from found papers.
- **Analyze and select relevant papers:** the second step consists of checking the quality of the obtained papers by the previous step. Those papers are selected based on journal quality, inclusion criteria, and exclusion criteria (see Table 1).
- **Classify papers into estimation and forecasting problems:** in this step, the already chosen papers are then classified into two categories: papers dealing with estimation solar irradiation problems and papers concerning forecasting solar irradiation.
- **Results and discussion:** present the actual state of the art for the two aforementioned categories with a detailed discussion on available approaches in estimation and forecast of solar irradiation.
- **Recommendations, trends and outlooks:** the last step offers useful recommendations and outlooks for futures researches in the field of solar irradiation estimation and forecasting using ANN techniques.

## Data pre-processing and feature selection for solar irradiation estimation and forecasting

In our review, we present an overview of the whole process required in estimation and forecasting models, including the type of input parameters, data pre-processing, feature selection, type of used models, forecasting horizon, and performance indicators.

### Data types and pre-processing phase for forecasting and estimation model

#### Data types

The forecast or estimation of solar irradiation requires data from the studied site. The most used are ground measured and sky images data. Available ground measured data in weather stations are generally geographical and meteorological parameters, solar characteristics and physical parameters. On the other hand, sky images data present the human observation of cloud motion to provide a piece of information about cloud cover, type, speed and level of cloud over a specific spot (Marquez and Coimbra 2013), (Barbieri et al. 2017). There are generally two types of sky images: ground-based sky images (Ferreira et al. 2012) and satellite-based sky images (Blaga et al. 2019a, b). Figure 3 summarizes the two types of input parameters used to estimate and forecast solar irradiation.

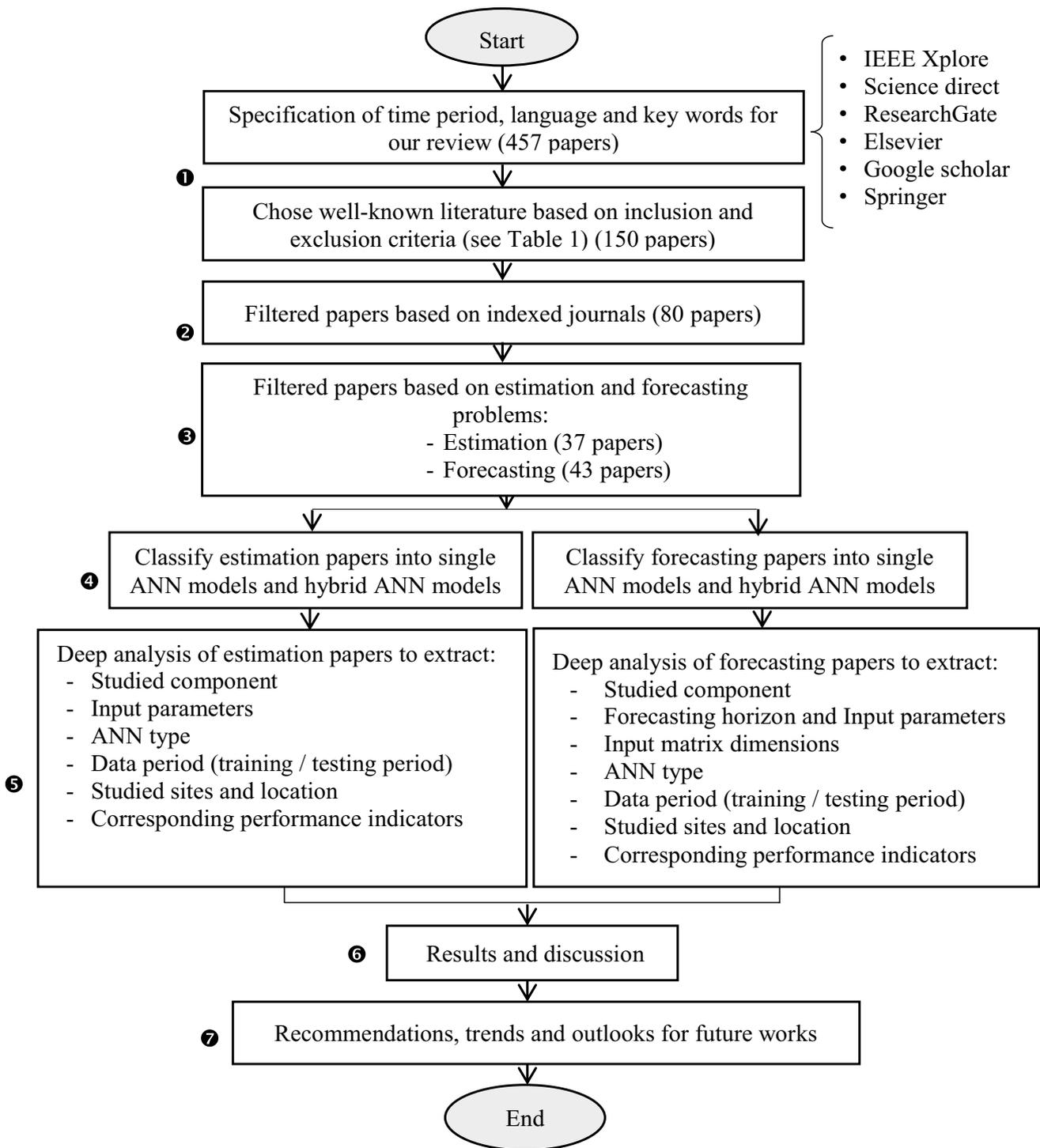


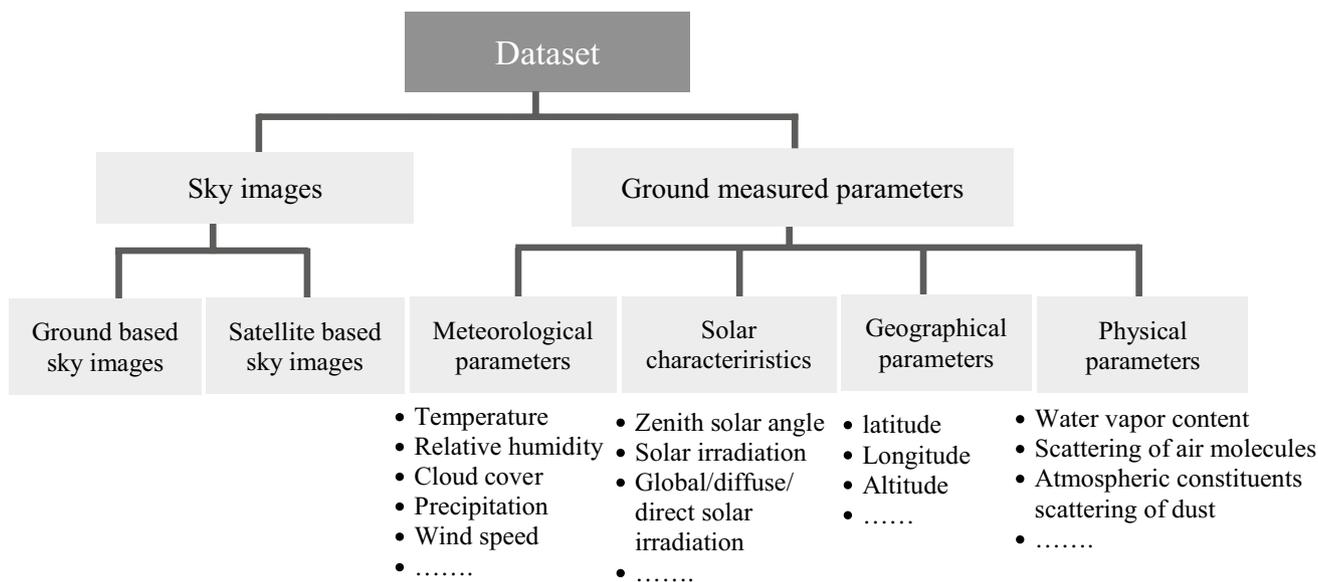
Fig. 2 Flowchart of our review

In prediction problems, these variables can be divided into three types: endogenous, exogenous, and hybrid parameters. The endogenous parameters consider only the historical values of solar irradiation data to predict solar

irradiation. Exogenous inputs include all pre-mentioned parameters (see Fig. 3) besides solar irradiation, and hybrid inputs combine both endogenous and exogenous parameters.

**Table 1** Inclusion, exclusion criteria, and search terms

Category	Inclusion criteria	Exclusion criteria	Keywords
Estimation	<ul style="list-style-type: none"> <li>- Global, diffuse and beam solar irradiation estimation</li> <li>- Monthly, daily and hourly solar irradiation estimation</li> <li>- Hybrid model (ANN combined with other models)</li> <li>- Comparative study between ANN models and other solar radiation estimation models</li> </ul>	<ul style="list-style-type: none"> <li>- Conference papers</li> <li>- Working papers</li> <li>- Commentaries</li> <li>- Review papers</li> <li>- Technical reports</li> <li>- Doctoral and masters theses</li> </ul>	<ul style="list-style-type: none"> <li>- Solar irradiation estimation using artificial neural networks models</li> <li>- Solar radiation estimation using ANNs models</li> <li>- Solar irradiation prediction based on ANNs models</li> <li>- Hybrid ANN model for solar irradiation estimation</li> <li>- Deep learning for solar irradiation estimation</li> </ul>
Forecasting	<ul style="list-style-type: none"> <li>- Very short, short and medium term solar irradiation forecasting</li> <li>- Global, diffuse and beam solar irradiation forecasting</li> <li>- Hybrid model (ANN combined with other models)</li> <li>- Comparative study between ANN models and other solar radiation forecasting models</li> <li>- Hybrid deep learning models</li> </ul>	<ul style="list-style-type: none"> <li>- Conference articles</li> <li>- Working papers</li> <li>- Commentaries</li> <li>- Review papers</li> <li>- Doctoral and masters theses</li> <li>- Long term solar irradiation forecasting</li> </ul>	<ul style="list-style-type: none"> <li>- Solar irradiation forecasting using ANNs models</li> <li>- Artificial neural network for solar radiation forecasting</li> <li>- Solar irradiation prediction based on ANNs models</li> <li>- Hybrid ANN model for solar irradiation forecasting</li> <li>- Deep learning for solar irradiation forecasting</li> </ul>



**Fig. 3** Types of input parameters used for estimation and forecasting solar irradiation

## Data horizon

The measurements of solar irradiation data at ground meteorological stations depend on several factors such as the measuring instruments specifications, calibration method and maintenance of measuring instruments, data acquisition method and accuracy, location, and environmental conditions (Sengupta et al. 2021). Measurement instruments can record solar irradiation values at a certain step or time resolution. Then, these data recorded at a specific time interval can be summed to obtain the desired forecasting horizon. These data can be sub-hourly, hourly, daily, and monthly solar irradiation according to their utility in solar systems (see Fig. 4). Sub-hourly solar irradiation is the value of solar irradiation recorded for one minute or several minutes. Hourly value is the sum of recording values of solar irradiation over a period of one hour. Daily solar irradiation is the total value of solar irradiation during a day, while, Monthly averages or sums are the average value of daily radiation over a month. Generally, the measures of solar irradiation at short time interval such as hourly or sub-hourly are more accurate than long-term interval such as daily and monthly because these data record the change of solar irradiation in details during a smallest time interval (Zhang et al. 2017).

Furthermore, the measuring instruments can be configured to produce output values of solar irradiation or other data at any desired time period (Sengupta et al. 2021). For instance, a value of solar irradiation can be a mean value of more frequent measurement records for a given time period.

## Data pre-processing

The quality of measured data depends on the quality of calibrations, and regular maintenance of measuring instruments. Furthermore, these instruments are greatly affected by time, location, and environmental conditions. In addition, the obtained data are generally presented in raw format and do not have significant characteristics to provide suitable accuracy. Thus, a pre-processing phase is required to prepare and ensure the dataset quality introduced to estimation or forecasting models. The quality of data plays a vital role in the accuracy of predictive models. Therefore, data should be quality controlled and well organized before further use.

Data preprocessing aims to solve these issues and improve the data quality by removing missing, meaningless values, and eliminating outliers caused by the abnormal measurement of data (Lai et al. 2020). Moreover, data pre-processing defines the input data according to the specifications and transforms the united raw format into a simpler data representation that is easier to use for future processing steps.

A number of techniques have been applied in the literature to pre-process the input parameters for forecasting models such as Markov models, linear regression, Kalman filter, Wavelet transform, self-organization map, quality control, empirical mode decomposition, Principal Component Analysis (PCA), and normalization. For instance, the work of Wang et al. in (Wang et al. 2018), used Discrete Wavelet transform for certain weather types (cloudy, rainy and heavy rainy) to decompose the raw solar irradiance sequence into several subsequences. Results showed that Discrete Wavelet transform based solar irradiance sequence decomposition enhances the corresponding forecasting performance of the proposed model. However, Castangia et al. in (2021b) normalized the GHI data by using the clear sky index transformation to introduce the stationarity in solar irradiation time series and to scale each input parameter in the range between 0 and 1. Moreover, in (Husein and Chung 2019a, b), Husein and Chung used linear regression fit to replace the missing values in his database. In (Lan et al. 2019), Lan et al. used PCA to reduce the size of the original database and to identify the essential frequency features given to the Elman-based neural networks.

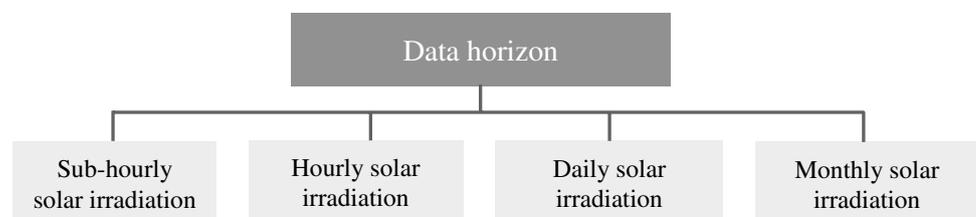
## Feature selection

The procedure of feature selection is a crucial requirement for improving model accuracy. In this section, we will provide an overview of the most used feature selection techniques and study their usefulness and ability to deal with solar irradiation prediction problems.

### An overview of feature selection techniques

Feature selection is an important phase to explore process and analyze data to a given problem. Feature selection techniques aim to explore and analyze the data in order to select the most important and relevant subset of features from a

**Fig. 4** Different horizons of solar irradiation data



large dataset (Garcia et al. 2015; Yahya et al. 2011; Zebari et al. 2020).

There are various objectives of feature selection, the most important ones are: improve models performance, reduce the size of data, reduce model complexity and computational costs, avoid overfitting problem, remove redundant, irrelevant and noisy data, improve data quality, and reduce data storage (Saeys et al. 2007). For instance, some machine learning models do not perform well in the case of high dimensions data. Reducing the size of data can be more efficient to visualize data clearly and reduce models complexity.

Feature selection techniques are generally performed through two major phases. Firstly, subset generation aiming to find the optimal variables from the available data. Followed by a subset evaluation phase to assess and determine whether the generated subset is optimal for the given problem according to a suitable stopping criterion that speeds up the selection and decides which features may be added or removed from generated subset. Moreover, the stopping criteria must be defined according to the purposes of feature selection (Khaire and Dhanalakshmi 2019), (Karagiannopoulos et al. 2007) (see Fig. 5).

There are several feature selection methods available in the literature. The most well-known and used are Filter, Wrapper, and Embedded methods.

#### i. Filter methods

The filter method is a category of statistical methods; it assesses the relevance of parameters by studying the correlation of each feature with the model's output to identify the most relevant parameters (Castangia et al. 2021a).

These methods are used either to produce a features ranking in terms of the relevant measurement by using statistical standards such as the Chi-squared test, mutual information, and Pearson correlation coefficient or combined with algorithms such as forward or backward search to obtain subsets of features (Zhang and Wen 2019).

The filter technique consists of two stages, the first one uses measures such as information, distance, dependence or consistency. The second stage is the learning and testing process with the subsets of relevant features.

The filter approach has several advantages such as dealing with high dimensional data, having better generalizable properties, being faster than other feature selection techniques, and finally the scalability for large-sized data.

Although, they can fail to evaluate complex and profound relationships between the features and the output. In addition, they ignore the dependency between features (Zhang and Wen 2019).

#### ii. Wrapper methods

Wrapper methods use a learning model with different subsets of features to select relevant feature subsets. The purpose to produce a ranking based on the accuracy obtained with each subset. In fact, a various subsets of features are generated and evaluated based on defining stopping criteria. The evaluation of a specific subset is obtained by training and testing the related model.

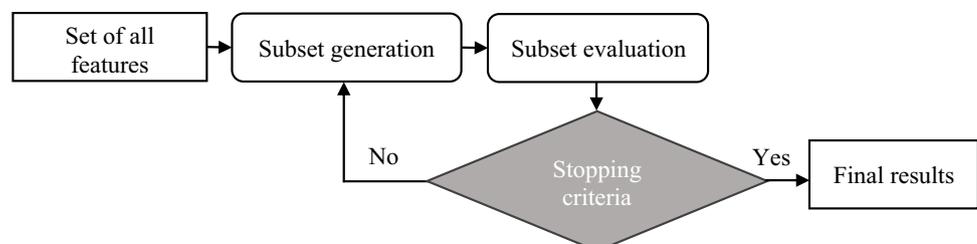
The Wrapper methods can be classified into two main groups, sequential Selection Algorithms and Meta-Heuristic Algorithms. The sequential selection algorithms start with an empty or full set and add or remove features respectively according to chosen criteria until obtaining the best subsets of features that achieve better performance. However, the heuristic search algorithms generate and evaluate different subsets to find the optimal one presenting the best accuracy (Chandrashekar and Sahin 2014). These search methods can be divided in two classes: deterministic and randomized search algorithms.

The advantages of wrapper techniques are their ability to include the interaction between feature subset search and model selection, and their capacity to take into account dependency between features. However, they are highly dependent on machine learning models, which lead to overfitting and computational problems with high dimension datasets (Garcia-Hinde et al. 2016). Also, the wrapped methods produce high time complexity due to the fact that they react after prediction stage.

#### iii. Embedded methods

Those methods are embedded into machine learning models that consider feature selection as a part of the learning process. The main approach of embedded methods is to integrate the feature selection as part of the training process. There are multiple machine learning techniques with embedded feature selection used in estimation such as: decision tree, convolutional neural network models (CNN) (Zhang and Wen 2019), and Multi-Objective Evolutionary Algorithms (Ferreira and Ruano 2011).

**Fig. 5** Major phases of feature selection technique



The embedded method combines the qualities of both filter and wrapper methods. Indeed, they have better computational complexity than wrapper and higher performance accuracy than filter. Furthermore, Embedded methods could provide a faster solution by avoiding reclassifying each subset of features and re-training a predictive model (Garcia et al. 2015).

### Feature selection for solar irradiation prediction using ANN models

The available solar irradiation data, together with meteorological and geographical parameters, result in a large number of variables that can contain unimportant and irrelevant parameters, which leads to a complex database with high dimensions. In fact, solar irradiation prediction models often suffer from a large amount of irrelevant information due to the existence of a high degrees of uncertainty associated with inadequate maintenance of sensors and absence of data quality control (Cebecauer and Suri 2015).

As mentioned above, there is a large number of parameters that are related to solar irradiation. Thus, feature selection is carried out to select the parameters that have a strong correlation with solar irradiation. Note that feature selection algorithms aim to keep the best set of variables in a prediction problem.

According to literature, several studies applied the feature selection techniques to identify the features that are most relevant as inputs and necessary for estimation or forecasting models for solar irradiation. In the case of endogenous inputs, there is no feature selection issue. However, for exogenous and hybrid inputs, choosing relevant ones giving the best accuracy is an important and open problem. For instance, Dahmani et al. in (Dahmani et al. 2016), Castangia et al. in (2021a), Pang et al. in (2020), Ahmad et al. in (Ahmad et al. 2015), Almaraashi in (Almaraashi 2018), and Meenal and Selvakumar in (Meenal and Selvakumar 2018) investigated the use of Filter methods. The wrapped methods were used in (Bouzgou and Gueymard 2017; Jadidi et al. 2018; Marzouq et al. 2019; El Mghouchi et al. 2019a, b; Rao et al. 2018), and the embedded methods were used in (Ghimire et al. 2019), (Zang et al. 2020).

## Forecasting horizon and models type

### Forecasting horizon

In solar system application, each stage of a solar energy system development requires data of solar irradiation at different time scales. The inaccuracy of measured data at specific locations leads to the high financial cost of any solar system project (Sengupta et al. 2021). Furthermore, the operation and management of an intermittent source

of energy produced from a solar system are very difficult when connected to an electrical grid due to their inability to balance between generation and demand at different time scales. In fact, a solar system output could be highly affected by the weather condition (Voyant and Notton 2018), (Kumari and Toshniwal 2021a). Thus, the accurate prediction of solar power requires accurate data of solar irradiation (Rajagukguk et al. 2020). Therefore, there is a crucial need to accurately predict solar irradiation at different forecasting horizons.

The forecast horizon is the time interval for which a model can predict future values. This period can range from few seconds to hours, days, weeks, months, or even years (Gupta et al. 2021). In the case of solar irradiation forecasting, different forecasting horizons can be found in the literature such as very short-term, short-term, medium-term and long-term. However, there is no universal definition of these forecasting horizons.

In fact, in (Yang et al. 2015), very short-term is defined as sub-5 min interval. Caldas and Alonso-Suárez in (2019) define very short as a time interval from 5 min to some hours ahead. In (Kumari and Toshniwal 2021b), it is from few minutes up to few hours, while in (Pedregal and Trapero 2021) it is from few seconds to 1 h.

In short-term horizon, the time interval for forecasting horizon is ranging from few hours up to several hours or days. For instance, in (Diagne et al. 2012), short-term horizon is defined from 1 h up to 5 h forecasts. Other researchers such as Voyant and Notton in (2018) and Marzouq et al. in (2020) considered short term forecast from 1 to 6 h. In addition, the study of Voyant et al. in (2017) considered that short-term can be divided in two sub-classes: now-casting (from 0 to 3 h ahead) that presents the very short time domain, and short-term forecasting (from 3 to 6 h ahead).

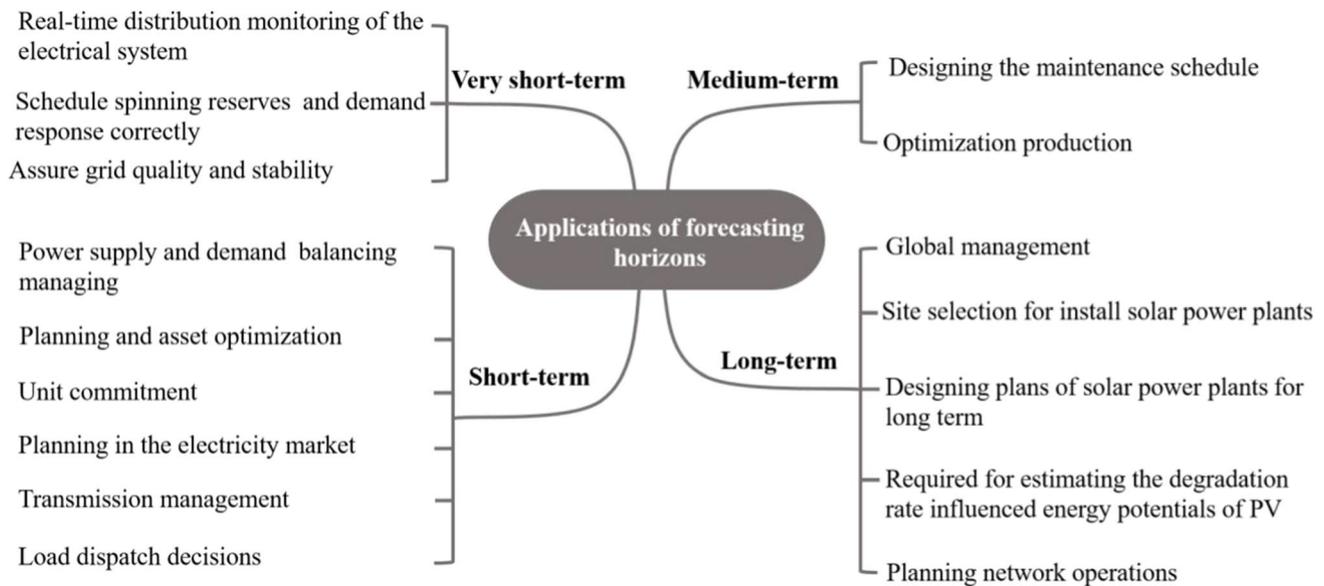
For medium-term, there are few studies carrying out forecasting of solar irradiation. In (Sharma and Kakkar 2018) and (Kumari and Toshniwal 2021b) medium-term is for one week ahead and in (Pedregal and Trapero 2021) it is longer than 48 h ahead.

For Ozoegwu, long term is a period of few months to years (Ozoegwu 2019). In (Aslam et al. 2019) it is up to one year ahead. However, in (Sharma and Kakkar 2018) and in (Kumari and Toshniwal 2021b) long-term is forecasting for months or years ahead.

The requirement for the forecasting horizon changes with applications. A synthesis of the studied papers of the literature can be summarized in Fig. 6, which shows the utility of each horizon from the very short to the very long term.

### ANN estimation and forecasting models

The estimation of solar irradiation is defined as the process in which the solar irradiation data can be predicted from the



**Fig. 6** Applications of different forecasting horizons

knowledge of other parameters of different nature. The collected data up to, and excluding time  $t$  are given as an input to an estimation model and it predicts the value of solar irradiation at the same time  $t$  as an output. The forecasting process consists in predicting future values of solar irradiation based on present and historical data. The collected data given to a forecasting model as inputs can be times series data that consider only current and past values of solar irradiation, structural data that present other variables of a different nature, or hybrid data that combine between time series data and structural data.

There are numerous types of ANN models used in the estimation and forecasting of solar irradiation, they can be broadly classified into single and hybrid ANN models. Single ANN models can be feedforward neural networks such as Multi-Layer Perceptron (MLP), Radial Basis Function Network (RBFNN), Generalized Regression Neural Network (GRNN), Extreme Learning Machine (ELM) and Recursive Neural Networks as Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM). The single ANN models can be used in different configuration for instance, Non-linear Autoregressive Neural Network with eXogenous inputs (NARX), Non-linear Autoregressive Neural Network (NAR), Time-Delay Neural Network (TDNN), and Wavelet Neural Network (WNN).

On the other hand, the hybrid approach consists of combining two or more models. In our review, we have only focused on hybrid models that combine ANN and other models such as statistical, empirical, or machine learning models. Even though these hybrid models are

complex, they have been widely used because of their ability to combine the advantages of different models, thus achieving higher accuracy for estimation and forecasting.

In Fig. 7, we synthesize existing ANN single models and some of the most used hybrid ANN models found in the literature for the prediction of solar irradiation.

### State of the art of solar irradiation estimation and forecasting based on ANN models

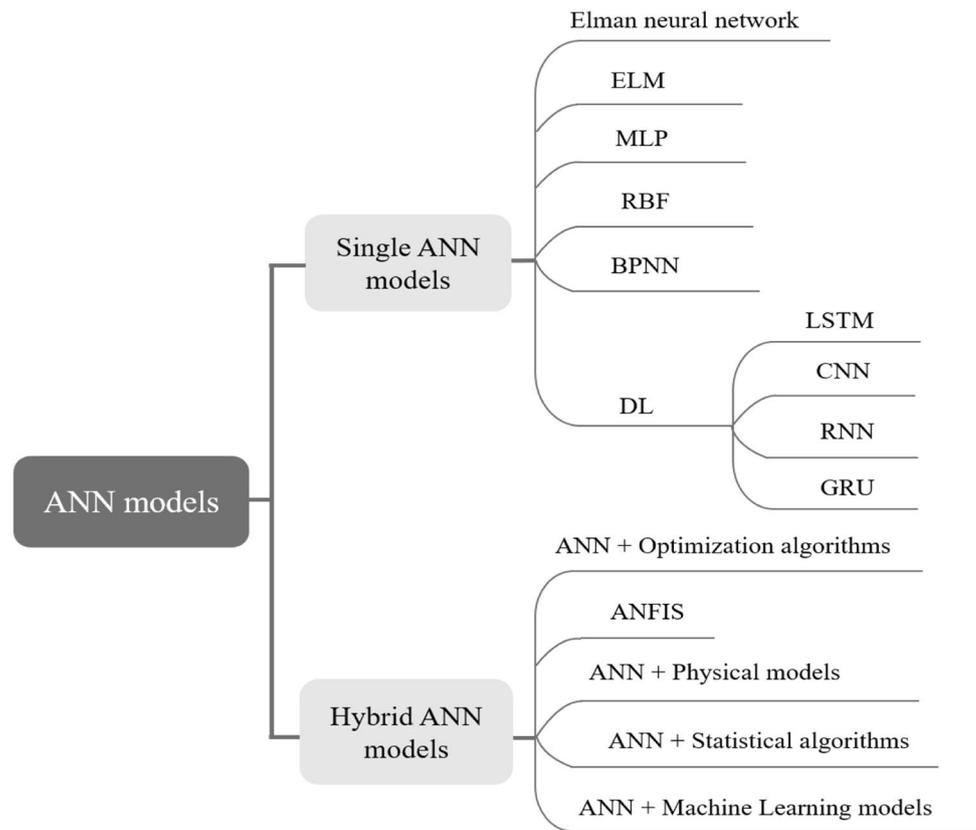
This section presents notable findings in estimation and forecasting models of solar irradiation after a deep analysis of the published articles based on all types of ANN models (including the hybrid ANNs and deep learning). In addition, a brief introduction to performance indicators investigated in the literature studies is briefly presented.

### Performance indicators of accuracy evaluation

Since the estimation and forecasting accuracy is a critical factor in selecting any prediction model, many well-known indicators were used in the literature. The most used are Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), normalized Mean Absolute Error (nMAE), and normalized Root Mean Square Error (nRMSE) (see Table 2).

Where  $N$  is the number of samples in the database,  $Y_{pred}$  is the  $i^{th}$  predicted value and  $Y_{meas}$  is the  $i^{th}$  measured value.

**Fig. 7** Single and hybrid ANN models found in the literature



**Table 2** Performance indicators used in our review study

Performance indicator	Formula
- Mean Absolute Percentage Error (MAPE)	$\frac{100}{N} \times \sum_{i=1}^N \left  \frac{Y_{pred} - Y_{meas}}{Y_{meas}} \right $
- Normalized Root Mean Square Error (nRMSE/nRMSD)	$\frac{100}{\frac{1}{N} \sum_{i=1}^N Y_{meas}} \times \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_{pred} - Y_{meas})^2}$
- Root-Coefficient of Variation Mean-Squared Error (CVRMSE)	$1 - \left( \frac{\sum_{i=1}^N (Y_{pred} - Y_{meas})^2}{\sum_{i=1}^N (Y_{meas} - \frac{1}{N} \sum_{i=1}^N Y_{meas})^2} \right)$
- Coefficient of determination (R <sup>2</sup> )	$1 - \left( \frac{\sum_{i=1}^N (Y_{pred} - Y_{meas})^2}{\sum_{i=1}^N (Y_{meas} - \frac{1}{N} \sum_{i=1}^N Y_{meas})^2} \right)$
- Root Mean Square Error (RMSE/RMSD)	$\sqrt{\frac{1}{N} \sum_{i=1}^N (Y_{pred} - Y_{meas})^2}$
- Normalized Mean Bias Error (nMBE)	$\frac{\sum_{i=1}^N (Y_{meas} - Y_{pred})}{\sum_{i=1}^N Y_{meas}} \times 100$
- Normalized Mean Absolute Error (nMAE/MAD)	$\frac{100}{\sum_{i=1}^N Y_{meas}} \times \sum_{i=1}^N  Y_{pred} - Y_{meas} $
- Mean Absolute Error (MAE)	$\frac{1}{N} \sum_{i=1}^N  Y_{pred} - Y_{meas} $
- Mean Absolute Bias Error (MABE)	$\frac{1}{N} \sum_{i=1}^N  Y_{pred} - Y_{meas} $
- Mean Percentage Error (MPE)	$\frac{100}{N} \times \sum_{i=1}^N \frac{(Y_{pred} - Y_{meas})}{Y_{meas}}$
- Mean Bias Error (MBE)	$\frac{1}{N} \sum_{i=1}^N (Y_{meas} - Y_{pred})$

**State of the art of solar irradiation estimation**

Our analysis of the retained papers from the literature dealing with the estimation of solar irradiation using an ANN-based model is classified and summarized into single and hybrid models presented respectively in Tables 3 and 4. These tables have been listed in a descendant

chronological order according to publication year by specifying the estimated component, the inputs parameters, the types of estimation model, the ANN architecture, data period and training/testing percentage, the number of studied sites and location, the corresponding performance indicators, and finally the results of compared methods.

**Table 3** State of the art of solar irradiation estimation based on single ANN models

Comp	Ref	Input parameters	ANN type	ANN architecture	Performance indicators	Training/testing dataset size	N of sites & locations	Results with compared methods
GHI	(Geetha et al. 2022)	$T_{in}$ , $W_S$ , hour, Long, $Y$ , $M_{ip}$ , Lat	FFNN	One hidden layer (7–12–1)	$R = 0.93$ $MAPE = 0.25$	1 year 5 sites / 1 site	1 site in India	N/A
GHI	(Martinez-Castillo et al. 2021)	Lat, Long, Alt, $M_{ip}$ , $RH_{mr}$ , $RH_{max}$ , $RH_{min}$	FFNN	One hidden layer (7-N/A-1)	$R^2(\%)$ : from 93.5 to 97.1 $RMSE(KJ/m^2 \cdot D)$ : from 201.3 to 209.4	14 years 3 sites / 2 sites	5 sites in Spain	FFNN > RF > SVM > MLR
GHI	(Ağbulut et al., 2021)	$T_{min}$ , $T_{max}$ , CC, $H_o$ , D	FFNN	One hidden layer (5–14–1); (5–13–1); (6–8–1); (5–16–1);	$nRMSE(\%)$ : from 14.10 to 18.67 $MBE(MJ/m^2)$ : from 0.23 to -0.02 $MABE(MJ/m^2)$ : from 1.5447 to 2.05	2 years; 70% / 30%	4 sites in Turkey	FFNN > DL > SVM > KNN
GHI	(Gürel et al. 2020)	SD, T, P, RH, WS	FFNN	One hidden layer (5–10–1)	$MBE(MJ/m^2 \cdot D) = -0.13$ $RMSE(MJ/m^2 \cdot D) = 0.78$ $MAPE(\%) = 4.92$	10 years / 1 year	4 sites in Turkey	ANN > Holt-Winters > RSM > Empirical models
GHI	(El Mghouchi et al. 2019a, b)	$K_p$ , $f(T_m)$ , $H_o$	BPNN	2 hidden layers (3–2–1–1)	$MBE$ , $nRMSE$ , $MAPE$ -0.203 0.055 0.002	2 years; 25 sites / 10 sites	27 sites in Morocco & 8 site in Algeria, Spain, Mauritania	N/A
GHI Daily Weekly Monthly	(Rocha et al. 2019)	DoY, $T_{max}$ , $T_{min}$ , SD, RF, CC, $H_o$ , RH, evap, WS	MLP	One hidden layer (10–40–1) One hidden layer (10–50–1) One hidden layer (10–30–1)	$nRMSE(\%)$ , $MAPE(\%)$ 6.29 14.87 6.57 18.17 6.42 20.83	14 years; 70% / 30%	1 site in Brazilian Northeast	N/A
GHI GTI $\alpha = 36.8^\circ$	(Notton et al. 2019)	GHI: Ta, RH, P, $W_S$ , SD, $H_o$ GTI: GHI, $\delta$ , $\theta_z$ , $H_o$	FFNN	One hidden layer (6-N/A-1) One hidden layer (4–4–1)	$nMAE(\%)$ , $nRMSE(\%)$ , $MBE(W.hour/m^2)$ 16.18 19.35 -0.081 6.63 8.81 -0.31	2 years; 80% / 20%	2 sites in France	N/A
GHI	(Rao et al. 2018)	$\Delta T$ , $H_o$ , $SD_{th}$	FFNN	One hidden layer (3–20–1)	$nRMSE(\%)$ , $MAPE(\%)$ , $MPE(\%)$ 3.96 2.61 1.32	2 years; 70% / 30%	1 site in India	N/A
GHI	(Meenal and Immanuel Selvakumar 2018)	SD, $T_{max}$ , $T_{min}$	ANN	One hidden layer (3–10–1)	$RMSE(MJ/m^2 \cdot D)$ : from 0.58 to 0.98 $MPE(\%)$ : from 0.11 to 1.01	6 years / 3 years 6 years / 3 years 4 years / 2 years 6 years / 3 year	8 sites in India	ANN > SVM > Empirical model
GHI	(Guermoui et al. 2018)	SD, RH, $T_{min}$	RBF MLP	One hidden layer (3–6–1) One hidden layer (3–9–1)	$MABE(MJ/m^2)$ , $nRMSE(MJ/m^2)$ 0.70 5.23 0.73 5.61	One year	1 site in Algeria	GPR > RBF > MLP

Table 3 (continued)

Comp	Ref	Input parameters	ANN type	ANN architecture	Performance indicators	Training/testing dataset size	N of sites & locations	Results with compared methods
GHI	(Kaba et al. 2018)	$H_o$ , SD, CC, $T_{min}$ , $T_{max}$ , DoY	DL	N/A	$MAE(MJ/m^2)=0.61$ $RMSE(MJ/m^2)=0.78$	7 years for train and test	34 sites in Turkey	N/A
GHI	(Marzo et al. 2017)	$T_{min}$ , $T_{max}$ , $H_o$	MLP	One hidden layer (3–5–1)	$nRMSE(\%) = 13$ $nMBE(\%) = 4$	NREL, BSRN 7 sites for training and testing	13 site in Chile, Israel, Saudi Arabia, South Africa and Australia	N/A
GHI	(Zou et al. 2016)	SD, $\Delta T$ , $T_{max}$ , WS, $T_{min}$ , RH, RF, P, $P_{wv}$	MLP	One hidden layer (9–17–1)	$RMSE(\%) = 11.25$ $MBE(\%) = -0.90$	50 years; 6 sites / 4 sites	10 sites in Southeast China	MLP > Empirical models
GHI	(Premalatha and Arasu 2016)	Lat, Long, Alt, Y, RH, $T_{min}$ , Pm, $M_{hr}$ , WS <sub>m</sub>	FFNN	One hidden layer (9–24–1)	$MAE(\%) = 3.03$ $RMSE(\%) = 3.65$	10 years; 5 sites / 5 sites	5 sites in India	N/A
GHI	(Dahmani et al. 2016)	Ta, RH, WS, WD, Pr, SD, P, $H_o$ , $\delta$ , $\beta_z$	MLP	One hidden layer (10–12–1)	$nRMSE(\%) = 13.33$ $nMAE(\%) = 12.72$	33 months 60% / 40%	1 site in Algeria	N/A
GHI	(Dahmani et al. 2016)	Ta, RH, WS, WD, Pr, SD, P, $\delta$ , $\beta_z$	MLP	One hidden layer (9–7–1)	$nRMSE(\%) = 18.65$ $nMAE(\%) = 15.50$	33 months 60% / 40%	1 site in Algeria	N/A
GI	(Yadav et al. 2015)	$T_m$ , $T_{max}$ , $T_{min}$ , Alt, SD	FFNN	One hidden layer (5–N/A–1)	$MAPE(\%) = 6.89$	22 sites / 4 sites	76 sites in India	FFNN > RBFNN > GRNN
GI	(Citakoglu 2015)	$M_{hr}$ , $H_o$ , $T_m$ , RH <sub>m</sub>	ANN	One hidden layer (4–6–1)	$MAE(\%) RMSE(\%)$ 1.253 1.650	20 years: 75% / 25%	163 sites in Turkey	ANN > ANFIS > Empirical models > MLR
GHI	(Shams Shirband et al. 2015a, b)	SD <sub>m</sub> , $\Delta T$ , $T_m$ ; RH	ELM	(4–N/A–1)	$MAPE(\%) = 2.25$ $nRMSE(\%) = 2.98$	7 years / 3 years	1 site in Iran	ELM > GP > SVM > ANN

**Table 4** State of the art of solar irradiation estimation based on hybrid ANN models

Comp	Ref	Input parameters	ANN type	ANN architecture	Performance indicators	Training/testing dataset size	N of sites & locations	Results with compared methods
GHI	(Kiliç et al. 2021)	Input1: Long, RH, SD <sub>day</sub> , SD, SD <sub>asr</sub> , RF, WS Input2: SD <sub>day</sub> , Lat, T <sub>max</sub> , SD, SD <sub>ast</sub>	MLP-HEA	One hidden layer (7–33–1) One hidden layer (5–34–1) Three hidden layer (3–N/A–1)	MAE(%) 3.36 RMSE(%) 8.2 RMSE(%) 10.95	60% / 40%	Switzerland, Europe and USA	MLP-EATE > MLP > DL
GHI	(Tao et al. 2021)	T <sub>max</sub> , T <sub>min</sub> , T <sub>m</sub>	ANFIS-muSG	N/A	RMSE(%) from 25.7 to 54.8	9 years 70% / 30%	5 sites in USA	ANFIS-mSG > ANFIS-GA > ANFIS-PSO > ANFIS-SSA > ANFIS-GWO > ANFIS-DA > ANFIS-GOA > ANFIS
GHI	(Kuhe et al. 2021)	T <sub>max</sub> , T <sub>min</sub> , RH, SD, WS	FFNN + RBFNN + GRNN	N/A	RMSE=0.8461 MAPE(%)=5.4425	11 years / 4 years	1 site, Nigeria	FFNN + RBFNN + GRNN > GRNN > FFNN > RBFNN
GHI	(Jalil et al. 2020)	H, T, RH, WS, RF	Multi neural approach of MLP	Two hidden layer (5–25–15–1)	R(%) = 96.24 – 98.41 RMSE(%) = 15.2 – 22.5	7 years of data	2 sites in Morocco	N/A
GHI	(Guijo-Rubio et al. 2020)	VIS 0.6, VIS 0.8, C <sub>s</sub> GI, CI, CAMS GI, SolarGIS GI	EANN	One hidden layer (8–N/A–1)	MBE(W/m <sup>2</sup> )=1.09 RMSE(W/m <sup>2</sup> )=51.82 MAE(W/m <sup>2</sup> )=33.46	1 year 80% / 20%	1 site in Spain	N/A
GHI	(Rabehi et al. 2020)	T <sub>m</sub> , T <sub>min</sub> , T <sub>max</sub> , H <sub>o</sub> , SD	LR-MLP-BDT	One hidden layer (5–N/A–1)	nRMSE=0.033 MAE=0.021585	3 years 70% / 30%	1 site in Algeria	MLP > LR-MLP > LR-BDT > BDT
GHI	(Feng et al. 2019)	T <sub>max</sub> , T <sub>min</sub> , H <sub>o</sub> , DoY	MEA-BPNN	One hidden layer (4–10–1)	MAE(MJ/m <sup>2</sup> .D): from 2.11 to 2.68 nRMSE: from 0.19 to 0.238	15 years/5 years	4 sites in China	MEA-BPNN > WNN > BPNN > RF > Empirical models
GHI	(Marzouq et al. 2019)	RF, WD, ΔT, H <sub>o</sub>	EANN	One hidden layer (4–9–1)	nRMSE(%)=17.85 MAPE(%)=18.46	7 years 75% / 25%	1 site in Morocco	EANN > KNN > Empirical models
GHI	(Jahani and Mohammadi 2019)	SD, diurnal air temperature	MLP-GA	One hidden layer (1–N/A–1)	MBE=38.4 RMSE(J/cm <sup>2</sup> .D)=185.5	3 years / 1 year	10 sites in Iran	MLP-GA > MLP
GHI	(Hou et al. 2018)	WS, T <sub>min</sub> , H <sub>max</sub> , H <sub>min</sub> , Evap, V <sub>fp</sub>	FOS-ELM	One hidden layer (6–N/A–1)	nRMSE(%)≈1.30 nMAE(%)≈0.69	15 years 4977 days / 500 days	1 site in Burkina Faso	FOS-ELM > ELM
GHI	(Chimire et al. 2018)	Satellite and Re-analysis data	SaDE-ELM	One hidden layer (20–40–1)	RMSE(MJ/m <sup>2</sup> ): from 0.47 to 0.50 MAE(MJ/m <sup>2</sup> ): from 0.14 to 0.43	15 years 130 months / 46 months	2 sites in Australia	SaDE-ELM > Hybrid-SVR > Hybrid-ANN > GP
GHI	(Mousavi et al. 2017)	T <sub>min</sub> , T <sub>max</sub> , T <sub>max</sub> , RH, P, WS, E	SAN-MLP	One hidden layer (7–10–1)	RMSE(KW.hour/m <sup>2</sup> .D)=0.90 MAE(KW.hour/m <sup>2</sup> .D)=0.67	15 years / 5 years	1 site in Iran	SAMLP > MLP > SVM
GHI	(Quej et al. 2017)	T <sub>min</sub> , T <sub>max</sub> , RF, H <sub>o</sub>	ANFIS	One hidden layer (4–N/A–1)	RMSE(MJ/m <sup>2</sup> .D)=2.78 MAE(MJ/m <sup>2</sup> .D)=2.15	Data collected from different Time period; 70% / 30%	6 sites in Yucatán Peninsula, México	G-SVM > ANFIS > FFNN
DHI	(Xue 2017)	MoY, SD, T <sub>min</sub> , RF, WS, RH, GHI	PSO-BPNN	One hidden layer (7–5–1)	RMSE=0.824 MAE=0.703	10 years	1 sites in china	N/A
GHI	(Hussain and AlAjili 2017)	Ta, SD, WS, RH	DWT-GRNN	N/A	RMSE(%)=2.78 MAPE(%)=1.43 MBE(%)=-0.0077	10 years 70% / 30%	1 site in United Arab Emirates	DWT-GRNN > DWT-NARX- DWT-MLP > DWT-ANFIS

**Table 4** (continued)

Comp	Ref	Input parameters	ANN type	ANN architecture	Performance indicators	Training/testing dataset size	N of sites & locations	Results with compared methods
GHI	(Shamshirband et al. 2016)	SD, SD <sub>max</sub>	ANN-ARX	N/A	MAPE(%)=9.45 nRMSE(%)=9.90	9 years / 4 years	1 site in Iran	ANN-ARX > ANFIS
GHI	(Mohammadi et al., 2016a)	Isfahan, Kerman: SD, SD <sub>max</sub> , T <sub>max</sub> Tabass: SD <sub>max</sub> , SD, H <sub>0</sub>	ANFIS	N/A	MABE(MJ/m <sup>2</sup> ) 1.47 2.35 RMSE(MJ/m <sup>2</sup> )=0.81	7 years / 3 years; 7 years / 3 years 5 years / 2 years	3 sites in Iran	N/A
DHI	(Mohammadi et al. 2016a, b)	GHI, H <sub>0</sub> , SD	ANFIS	3 hidden layers (2-3-6-10-1)	MABE(MJ/m <sup>2</sup> )=0.81 RMSE(MJ/m <sup>2</sup> )=1.10	4 years of data 2 years/2 years	1 site in Iran	N/A
GHI	(Shamshirband et al. 2015a, b)	T <sub>max</sub> , ΔT	KELM	N/A	MABE(MJ/m <sup>2</sup> ) 1.35 11.25 nRMSE(%)	9 years / 4 years	1 site in Iran	ELM > SVR

**Estimation of solar irradiation based on single ANN models**

In the literature, several ANN modeling approaches were used to estimate solar irradiation. Table 3 presents our state of the art of single ANN models for solar irradiation estimation. From this table, we can see that even if there is a large choice of approaches in neural networks models, only four models were widely used in the literature: MLP with different configurations and ELM, which demonstrates their effectiveness in this field.

**Estimation of solar irradiation based on hybrid ANN models**

The hybrid ANN approach combines an extension of the ANN model with another approach aiming to overcome the disadvantages of a single ANN model. These models have been developed and broadly used in the estimation of solar irradiation, especially in the last years. We can see from Fig. 7 that hybrid models can be a combination of two or more models or optimization of a single ANN model. Table 4 lists our state of the art of solar irradiation estimation using hybrid ANN models.

**Solar irradiation forecasting based on ANN models**

The published papers during the last seven years based on ANN techniques for solar irradiation forecasting were analyzed and summarized in Tables 5 and 6. These tables have been classified in a descendant chronological order according to the year of publication by specifying the forecasting component and horizon, the inputs parameters, the forecasting model type, the input matrix dimension, data period and training/testing percentage, the number of studied sites, location, the performance evaluation indicators, and results of compared methods if available.

**Forecasting using ANN-based single model**

Solar irradiation forecasting is a challenging task. In fact, there are several proposed algorithms aiming to maximize prediction accuracy. Regarding ANN models, MLP with different architecture, and LSTM are the most used to solve this issue (Additional details will be given in Sect. 6). In Table 5, we provided a deep analysis of the published works.

**Forecasting using ANN-based hybrid models**

From the literature, we can note that the most published studies in the last years for solar irradiation forecasting are based on hybrid models. Those methods adopt either combination or optimization techniques to increase the ANN-based single model accuracy. For instance in (Kumari and Toshniwal 2021c), authors combine CNN and LSTM. While

Table 5 State of the art of works using single ANN models to forecast solar irradiation

Comp Ref	Input parameters	Horizon	ANN type	Input matrix size	Performance indicators	Training / testing dataset size	N of sites & locations	Results with compared methods
GI (Eşjik et al. 2022)	sky images	5 min ahead	LSTM	Previous 5 min	Result for four days: RMSE: 47.58, 53.83, 68.10 and 92.89 MAE: 41.20, 32.88, 48.07 and 60.74 $RMSE(W/m^2) = 52.58$	1 month / 2 days / 4 days	1 site in Turkey	LSTM > Others models based on sky images
GI (Etxegarai et al. 2022)	Season, time range, GI	10 min ahead	CNN	24 previous hours		2 years / 1 year	1 site in Basque Country	CNN > LSTM > Per
GHI (Castangia et al. 2021a)	UV, CC, Ta, GHI, RH, DPT, WD, SD, $H_o$	4 h ahead	LSTM	3 previous observations	$MAD$ : from 10.74 to 24.51 $RMSE$ : from 29.81 to 55.19	4 years / 1 year	1 site in Italy	LSTM > FFNN > RF > CNN > ESN
GHI (Ghémou et al. 2021)	GHI	10 min, 1 h and 4 h ahead	LSTM	Previous time steps	$nRMSE(\%)$ : from 0.20 to 0.40	1 year / 1 year	1 site in France	LSTM > MLP > GPR > SVR > Per
GHI/DHI (Kheifit et al. 2020)	$D$ , $T_{min}$ , $T_{max}$ , $T_m$ , $\delta$ , $RH_{min}$ , $DD$ , $SR$ , $RH_{max}$ , $Pm$ , $RH_m$ , $P_{min}$ , $P_{max}$ , $SD$ , $ME$ , $DHI$ , $GHI$	3 days ahead	MLP RBFNN	2 previous days	$nRMSE$ : from 34.37 to 38.39 $MABE$ : from 4.73 to 5.87 $nRMSE$ : 33.39 from to 39.59 $MABE$ : from 4.94 to 6.01	3 years; 1.5 years / 1.5 years	1 site in Algeria	N/A
GHI (Rangel et al. 2020)	GI, Ta	1 h ahead	NARX	1 previous hour	$MAE(W/m^2)$ : from 29.1 to 46.8 $RMSE(W/m^2)$ : from 38.6 to 83.0	1 year / 1 year	3 sites in Mexico	NARX > SP
GHI (Huynh et al. 2020)	GI	1, 5, 10, 15 and 30 min ahead	LSTM	4 previous time steps	$RMSE(W/m^2)$ : From 40.91 to 731.75 $MAE(W/m^2)$ : from 21.64 to 409.72	10 month; 80% / 20%	1 site in Vietnam	LSTM > DNN > MLP > ARIMA > SVR
GI (Jeon and Kim 2020)	Ta, RH, WS, RF, GI, SC	Day ahead	LSTM	2 previous days	$RMSE(W/m^2) = 30.21$ $CVRMSE(\%) = 21.64$	5 years; 5 sites / 1 site	6 sites in Korea	N/A
GHI (Yu et al. 2019)	GHI, Ta, RF, RH, $\theta_z$ , WS, WD, CC, DPT	1 h and 24 h ahead	LSTM	24 previous hours	1 h ahead: $RMSE(W/m^2)$ : from 41.37 to 66.69 $MAE(w/m^2)$ : from 30.19 to 46.04 24 h ahead: $MAPE(\%)$ : from 8.0 to 11.2	4 years / 1 year	3 sites in USA	LSTM > RNN > ARIMA > SVR > BPNN > CNN
GHI/DNI (Notton et al. 2019)	GHI DNI	from 1 to 6 h ahead	MLP	6 previous hours for GHI and 7 for DNI	$nRMSE(\%)$ : $nMAE(\%)$ : from 22.57 from 16.16 to 34.85 to 26.28 from 38.23 from 30.27 to 61.88 to 53.82	2 years; 80% / 20%	2 sites in France	MLP > SP
GI (Wojtkiewicz et al. 2019)	GHI, RH, $Ta$ , $\theta_z$	1 h ahead	LSTM	48 previous hours	$RMSE(W/m^2) = 66.57$ $MAPE(\%) = 25.37$	11 years; 80% / 20%	1 site in USA	LSTM > GRU
GHI (Aslam et al. 2019)	GI, GHI in clear sky	1 h and 1 day ahead	LSTM	10 previous values	1 h ahead: $RMSE = 0.35-0.42$ 24 h ahead: $RMSE = 4.77-5.57$	10 years / 1 year	2 sites in South Korea	LSTM > GRU > FFNN > RNN > SVR > RPR
GHI/BNI/DHI (Benali et al. 2019)	GHI, BNI and DHI in clear sky	From 1 to 6 h ahead	MLP	Using auto mutual information to select historical data	$nRMSE(\%)$ : $nMAE(\%)$ : from 22.57 from 16.16 to 34.85 to 26.28 -from 38.23 from 30.27 to 61.88 to 53.82 -from 40.61 from 29.32 to 66.41 to 47.94	3 years; 80%/20%	1 site in France	RF > MLP > SP
GI (Qing and Niu 2018)	DoM, $D_p$ , $Ta$ , DPT, RH, WS, WT	1 day ahead	LSTM	1 previous day	$RMSE(W/m^2) = 126.31$ $RMSE(W/m^2) = 76.24$	-2 years / half-year / -9 years / 1 year	2 sites in USA	LSTM > BPNN > Per > LSR

Table 5 (continued)

Comp	Ref	Input parameters	Horizon	ANN type	Input matrix size	Performance indicators	Training / testing dataset size	N of sites & locations	Results with compared methods
GHI	(Fouillloy et al. 2018)	CS indexes	from 1 to 6 h ahead	MLP	8 previous hours for 2 sites and 6 for 1 site	$nRMSE(\%)$ : from 18.26 to 33.84 $MAE(W/m^2)$ : from 60.63 to 103.31	3 years for 2 sites & 1 year for 1 site 80% / 20%	2 sites in France and 1 site in Greece	MLP > ARMA > Pruned RT > Boosted RT > Bagged RT > R > GPV > Per
GI	(Bou-Rabee et al. 2017)	GI	Day ahead	FFNN	1 previous day	$MAPE(\%) = 85.6$	4 year / 1 year	5 sites in Kuwait	N/A
GI	(Wei 2017)	P, WS, RF, Ta, SA, RH, Az, GI, $\delta$ , $\theta_z, \omega$	from 1 to 12 h ahead	MLP	1 previous hour	$nMAE$ : from 0.20 to 0.40 $nRMSE$ : from 0.39 to 0.66	6 years / 1 year	1 site in Taiwan	MLP > RF > KNN > LR
GI	(Sharma et al. 2016)	GI	15 min and 1 h ahead	WNN	The number of input change with season	$MBE$ : from $nRMSE(\%)$ : from 13.27 to 14.36 9.42 to 12.25 From 6.91 From 15.41 to 8.84 to 20.14	12 month; 2 month / 1 month in 6 test period	10 sites in Singapore	WNN > MLP > ETS > ARMA > Per
GHI	(Gutierrez-Corea et al. 2016)	Ta, RH, GI, Kc, WS, WD, $H_o$ , $\theta_z$ , DSN	From 1 to 6 h ahead	MLP	10 previous Time frames	$MAE(\%)$ : from 26.21 to 47.99 $nRMSE(\%)$ : from 13.38 to 18.27	18 000 000 samples 70% / 30%	1 site in Spain	N/A
GHI	(Ahmad et al. 2015)	$T_{max}$ , $T_{min}$ , P, RH, $\theta_z$ , Az, RF	day ahead	NARX	1 previous day	$nRMSE(MJ/m^2) = 0.0495$	3 years; 70% / 30%	10 sites in New Zealand	NARX > ARMA > MLP > Per
GHI	(Mazorra Aguiar et al. 2015)	GHI and satellite data	from 1 to 6 h	BPNN	60 previous minutes	$nRMSE(\%)$ : from 15.3 to 36.2	1 year; 8 month / 4 month	2 sites in Spain	BPNN > ANN > SP > CLI > Per

in (Marzouq et al. 2020), authors used an optimization method in the form of evolutionary algorithm to increase the ANN model accuracy. More discussion will be given in the next section. A detailed analysis of the state of the art is presented in Table 6.

### Discussion, trends and recommendations

From the previously presented states of the art, a detailed discussion will be given about research directions and open problems extracted from the related papers. Finally, our review is concluded with compilation of the main findings and interesting outlooks in solar irradiation prediction based on ANN models.

### Discussion of factors influencing the performance of ANN models

During our review, we noticed that forecasting or estimating solar irradiation with a good accuracy is a difficult task. In fact, as we can see from Tables 3, 4, 5, and 6, the performances of ANN models are highly dependent on several influential factors including forecasting horizon, climate type, used ANN model, input data, optimization algorithms, input matrix dimension, and forecast component (see Fig. 8).

In the next subsections, we will present a detailed discussion about all pre-mentioned factors.

### Impact of ANN model type

Despite the fact that single ANN models provide high accuracy in predicting solar irradiation, many researchers are still working to improve it by combining ANN models with other algorithms in order to form a more accurate hybrid model.

As we can see from Tables 3, 4, 5, and 6, there are various approaches of ANN models to estimate and forecast solar irradiation. In our review, we classified these methods into two main classes. The first class is dedicated to single ANN models such as MLP, RBFNN, LSTM, and CNN. The second class is devoted to hybrid ANN models, which combine ANN with other statistical or machine learning approaches, for instance, evolutionary artificial neural networks (EANN), LSTM-CNN, and ANFIS.

The aim of hybrid models is to maximize the model accuracy by overcoming the weakness of a single ANN approach (Benmouiza and Chekane 2016), (Pazikadin et al. 2020). Indeed, each extension of ANN models has strengths and limitations to forecast and estimate solar irradiation. Several studies proved that ANN as a single model has difficulties to deal with the intermittency and fluctuations of solar irradiance (Huang et al. 2021). In order to overcome the shortcomings of a single model to aim higher accuracy, authors

**Table 6** States of the art of works using hybrid ANN models to forecast solar irradiation

Comp	Ref	Input parameters	Horizon	ANN type	Input matrix size	Performance indicators	Training/testing size	N of site & location	Results of compared methods
GHI	(Gallo et al. 2022)	Satellite images, GHI in clear sky	From 15 min to 6 h ahead	3D-CNN ConvLSTM	-Previous 24 and 2 h of GHI And images respectively	MAD(%): from 18.0 to 26.57 nRMSE: from 0.39 to 0.57	6 years; 75% / 25%	1 site in Italy	ConvLSTM > 3D-CNN For longer forecasting horizons
GHI	(Bhatt et al. 2022b)	GHI, $T_m$ , $RH_m$	From 15 min to 1 h and 30 min	CNN-LSTM	12 previous steps (average 15 min interval)	RMSE( $W/m^2$ ): from 5.79 to 34.89 MAE( $W/m^2$ ): from 3.83 to 20.07	2 years; 80% / 20%	1 site in Thailand	CNN-LSTM > CNN > LSTM > Per
GHI	(Singla et al. 2022b)	GHI	24 h ahead solar	WT-Bi-LSTM	15 previous hours	RMSE( $W/m^2$ ) = 45.61 MAPE(%) = 6.48	1 year; 80% / 20%	1 site in India	WT-BiLSTM > LSTM > LSTM > GRU > Native
GI	(Jaihani et al. 2022)	GI, $T_a$ , RH, RF, WS, WD	5 min ahead	Bi-GRU	Previous 5 min	$R^2=0.958$ RMSE( $W/m^2$ )=46.1	26 month For training and testing	1 site in South Korea	Bi-GRU > Bi-LSTM > RNN > LSTM > GRU
GHI	(Ali-Ou-Salah et al. 2021)	H, $T_a$ , RH, WS	1 h ahead	FCM- FFNN	1 previous hour	RMSE( $W/m^2$ )=106.4	3 years / 2 year	1 site in Portugal	FCM-FFNN > FFNN
GHI	(Kumari and Toshniwal 2021c)	T, WS, RH, RF, $\theta_s$ , CC, Cs, GHI, DPT, P	1 h	LSTM-CNN	12 previous hours	MAE( $w/m^2$ ): from 27.38 to 37.02 RMSE( $w/m^2$ ): from 49.61 to 58.12	4 years; 4 years / 1 year	3 sites in USA	LSTM-CNN > CNN-LSTM > LSTM > ANN > SVM > CNN > SP
GHI	(El Alani et al. 2021)	Sky images, GHI, $T_a$ , RH, WS, WD	15 min ahead	CNN-MLP	15 previous minutes	nRMSE( $W/m^2$ ): from 16.54 to 120.83 MBE( $W/m^2$ ): from -0.87 to -7.66	01/09/2021 to 20/04/2021 80% / 20%	1 site in morocco	CNN-MLP > Per
GTI $\alpha=36.8^\circ$	(Amiri et al. 2021)	$\delta$ , SA, $H_o$ , GHI, $\alpha$	1 h ahead	SUNN	1 previous hour	-Ground data: nRMSE(%) = 18.40 nMAE(%) = 11.08 - Satellite and ground data: nRMSE(%) = 21.83 nMAE(%) = 14.09	-2 years of Ground data and 35 month of Satellite data 60%/40%	1 site in Algiers	SUNN > RBFNN > PUNN
GI	(Huang et al. 2021)	$H_o$ , GI, $T_a$ , RH, WS	1 h ahead	WPD-CNN-LSTM-MLP	24 previous hours	MAE(%) : from 14.35 to 20.38 nRMSE(%) : from 15.48 to 23.91	Denver: 4 years / 1 year; Clark: 3 years / 1 year; Folsom: 2 years / 1 year	3 sites in USA	WPD-CNN-LSTM-MLP > CNN-LSTM-MLP > LSTM-MLP
GHI	(Liebermann et al. 2021)	GI, RH, WS, $T_a$ , P, RF	24 h ahead	LSTMconv	24 previous hours	RMSE( $W/m^2$ ): from 71.59 to 95.25 nMAE(%) : from 74.78 to 79.37	3 years/ 1 year	1 site: in Spain Germany, England and Finland	LSTMconv > convLSTM > LSTM > CNN

Table 6 (continued)

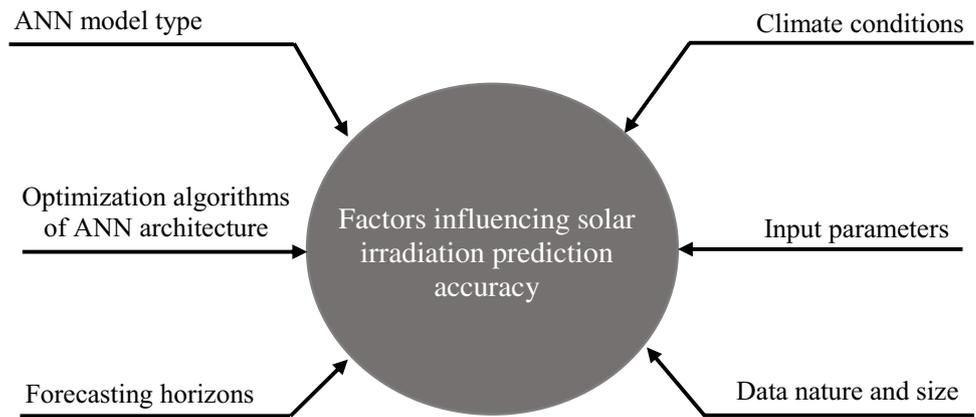
Comp	Ref	Input parameters	Horizon	ANN type	Input matrix size	Performance indicators	Training/testing size	N of site & location	Results of compared methods
GI	(Rodríguez et al. 2021)	Td, RP, GI, season	10 min ahead	FFNNST	24 previous hours	$RMSE(W/m^2)=50.80$	2 years/ 1 year	1 site in Spain	FFNNST > FNN > RNN > SVM
GI	(Gao et al. 2020)	GI	1 h ahead	CEEMDAN-CNN-LSTM	24 previous hours	$nRMSE(\%)$ : from 13.10 to 20.69 $MAE(W/m^2)$ : from 14.91 to 23.08	4 years / 2 years	4 sites in USA	CEEMDAN-CNN-LSTM > CEEMDAN-LSTM > CEEMDAN-BP > SVM > CEEMDAN-LSTM > CEEMDAN-ARIMA > LSTM > BPNN > SVM > ARIMA
GHI	(Marzouq et al. 2020)	GHI	from 1 to 6 h ahead	HAEMLP	The matrix dimension change with each site	$nRMSE(\%)$ : from 7.59 to 12.49 $nMAE(\%)$ : from 4.41 to 8.12	1 year; 18 sites /10 sites	28 sites in Morocco	HAEMLP > RF > SP > RT
GHI	(Zang et al. 2020)	GHI, DPT, RF, RH, $\theta_z$ , WS, WD, Ta	1 h ahead	CNN + LSTM	10 previous hours	$nMAE(\%)$ : from 37.20 to 52.00 $nRMSE(\%)$ : from 16.58 to 22.85	3 dataset; 7 years /1 year	34 sites in USA	CNN-LSTM > ANN-LSTM > CNN-ANN > LSTM > ANN > CNN > SVM > Per
GHI	(Lan et al. 2019)	GHI	24 h ahead	DFT-PCA-Elman	24 previous hours	$RMSE$ : from 72.95 to 191.33 $MAE$ : from 39.46 to 118.67	1 year; 5 sites/1 site	6 sites in china	DFT-PCA-Elman > DFT-PCA-BPNN > PCA-BPNN > ARIMA > Per
GHI	(Husein and Chung 2019a, b)	DBT, DPT, RH, H, MoY, GI	Day ahead	LSTM-RNN	2 previous days	$MAE(W/m^2)$ : from 36.90 to 64.36 $RMSE(W/m^2)$ : from 60.31 to 108.52	Six dataset From different time period	6 sites in USA Germany, South Korea, Switzerland	LSTM-RNN > FFNN > Per
BNI	(Zhao et al. 2019)	DNI, DNI <sub>th</sub> , in clears sky, images	10 min ahead	3D-CNN + MLP	3 previous values	$nRMSE(\%)=30.82$ $nMAE(\%)=17.91$ $nMBE(\%)=0.21$	NREL dataset; 1 year/ 1 year	USA	3D-CNN + MLP > MLP > Per
GHI	(Lan et al. 2018)	GHI	Day ahead	EEMD-SOM-BP	1 previous day	$MAE(W/m^2)$ from: 42.89 to 89.61 $RMSE(W/m^2)$ from: 72.84 to 137.85	SolarGIS dataset; 50% / 50%	4 sites in china	EEMD-SOM-BPNN > RBFNN > BPNN
GI	(Wang et al. 2018)	GI	1 day ahead	CNN-LSTM	3 previous days	$RMSE$ : from 32.41 to 180.09 $MAE$ : from 16.57 to 150.19	- 5 years - 4 years	2 sites in USA	DWT-CNN-LSTM > CNN-LSTM > CNN > LSTM > ANN > ARIMA > Per
GHI	(Jadidi et al. 2018)	$T_{mp}$ , WCT, $AA$ , ACRT, $PWS_n$ , $P_{m,\theta_{Zm}}$ , DPT <sub>mp</sub> , RH <sub>mp</sub> , $Az_m$ , GHI, $WS_n$	1 h ahead	GA-MLP	N/A	$nRMSE = 0.0458$ $nMAE = 0.0238$	4 years; 70% / 30%	North Carolina, USA	GA-MPLNN > PSOMPLNN

Table 6 (continued)

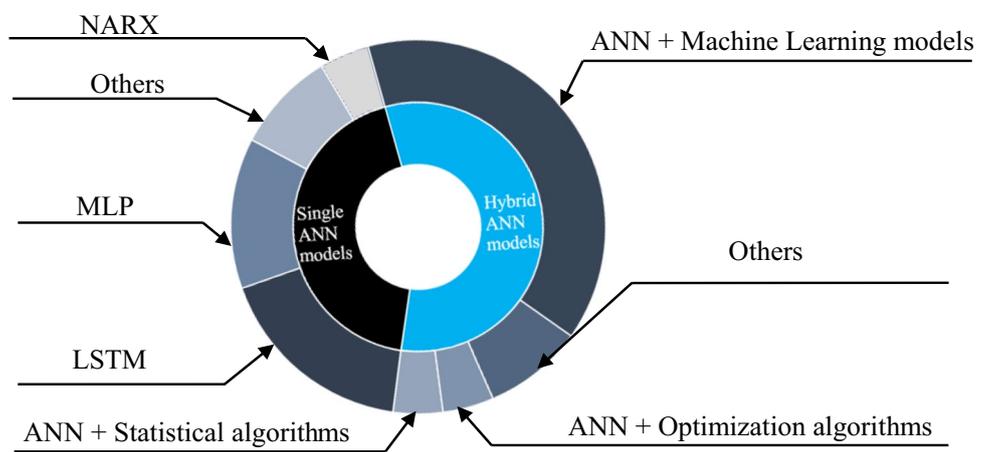
Comp	Ref	Input parameters	Horizon	ANN type	Input matrix size	Performance indicators	Training/testing size	N of site & location	Results of compared methods
GHI	(Monjoly et al. 2017)	Clear sky index	1 h ahead	DWT-AR + MLP	Mutual information	$nMBE(\%)$ : from -2.06 to 0.02 $nMAE(\%)$ : from 2.76 to 6.64 $nRMSE(\%)$ : from 3.80 to 8.31	1 year / 1 year	1 site in France	DWT-AR + MLP > EEMD-AR + MLP > EMD-AR + MLP > Per
GHI	(Benmouiza and Chekmane 2016)	GHI	915 h	ARMA + NAR	Previous values	$nRMSE$ : from 0.19 to 0.21	1 year for training and testing	1 site in Algeria 1 site in England	ARMA + NAR > NAR > ARMA
GHI	(Gairaa et al. 2016a)	GI	1 day ahead	ARMA-NAR	-4 previous hours; -2 previous hours	$nRMSE$ $nMBE$ 0.298 -0.0113 0.1119 -0.0054	2 years; 80% / 20%	1 site in Spain 2 sites in Algeria	ARMA- NAR > ARMA
GI	(Azimi et al. 2016)	GI	1, 24, 48 h ahead	TB K-means + MLP	15 previous hour	$nRMSE(\%)= 7.27-11.945-17.90$	NREL and Mesonet; 80% / 20%	3 sites in USA	TB K-means + MLPNN > Per
GHI BNI	(Ghofrani et al. 2016)	Ta, WS, WD, $H_o$	1, 24, 48 h ahead	K-means + MLP	N/A	$nRMSE$ : from 0.07 to 0.15 from 0.036 to 0.07 -from 0.11 from 0.06 to 2.27 to 0.19	NREL; 80% / 20%	U.S.A	N/A
GHI	(Pedro and Coimbra 2015)	GHI	From 15 to 120 min	KNN + MLP	Actual value	$MAE$ : from 17.0 to 63.2 $RMSE$ : from 41.3 to 90.2	1 year; 50% / 50%	5 sites in Unit state	ANN > KNN > Per
GHI	(Cornaro et al. 2015)	GHI, Ta	1 day ahead	MOSMLP	1 previous day	$RMSE(W/m^2)= 106$ $MAE(W/m^2) = 69$	4 years; 60% / 40%	1 site in Italy	MOSMLP > STML > Per

\*N/A: Not Available

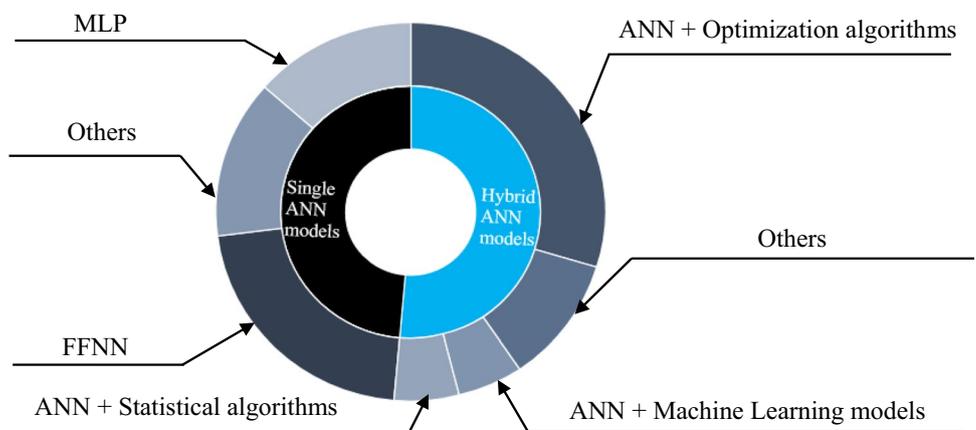
**Fig. 8** Factors influencing solar irradiation prediction accuracy



**Fig. 9** Most used approaches in solar irradiation forecasting



**Fig. 10** Most used approaches in solar irradiation estimation



focused on testing hybrid models to improve forecasting accuracy and improve model reliability, especially in bad weather conditions (Azimi et al. 2016), (Ali-Ou-Salah et al. 2021), (Benmouiza and Chekane 2016). As we can see in (Gairaa et al. 2016a), authors conducted a study to forecast solar irradiation by combining Auto Regressive Moving

Average (ARMA) with ANN: ARMA was used to capture the linearity information and the ANN was capturing the nonlinear characteristics of data.

Regarding the studied literature, we deduced that models as MLPNN, RBFNN are limited in high dimensional datasets and sky images, which directly influences the prediction

model accuracy by overfitting and extrapolation. Therefore, further studies with more focus on other extensions of ANNs such as RNN, LSTM, CNN, and their combination (Huynh et al. 2020), (Liebermann et al. 2021), (Husein and Chung 2019a, b), (Pang et al. 2020) were proposed to overcome this limitation. These models become popular in the last years, as they present some advantages, such as the ability of capturing the short and the long-term dependencies within the solar irradiation data series patterns to accurately predict it in the future at different time horizons and also the ability of performing a more general feature extraction. In addition, these models perform well in different weather conditions. As demonstrated in the study of (Gbémou et al. 2021), LSTM presents good performances for cloudy days. Moreover, the hybrid models that combine between LSTM and CNN are more robust under diverse climatic conditions seasonal and sky conditions (Liebermann et al. 2021), (Zang et al. 2020), (Kumari and Toshniwal 2021c). The performance of these models exceed other extension of ANN models; however, they have some limitations. In fact, these models require high dimensional datasets and complex inputs parameters such as images. In addition, research in this area is very recent and other studies are expected to deeply explore more performances from future models. Figure 9 and Fig. 10 show a synthesis of studied papers. We can note that the number of papers dealing with forecasting problem is slightly higher than papers devoted to estimation. We can also see that FFNN is the most used algorithm in the case of irradiation estimation and LSTM in the case of forecasting for single ANN models. For hybrid, models optimized ANN and machine learning combined with ANNs are the most used in the case of estimation and forecasting respectively.

### Optimization algorithms of ANN architecture

In all presented models, the selection of appropriate parameters is crucial for improving the model accuracy. In ANN models, the selected parameters are usually, the number of hidden layers, the number of neurons in each hidden layer, and the number of epochs and the initial weights. These parameters are associated with the data patterns, input parameters, training, and testing datasets and training algorithms.

During our inspection of the literature, we observed that various studies determined the pre-mentioned parameters experimentally based on the trial-and-error method to select parameters that enhance the accuracy of estimation and forecasting models. For instance, Citakoglu in (2015) and Quej et al. in (2017) tested the ANN model under several architectures to choose the best ANN structure. However, this method produces high computational complexity.

On the other hand, there are few studies carrying out automatic methods to optimize these parameters. In this

context, optimization algorithms have been adopted as the best solution to overcome this problem by identifying and selecting the appropriate model parameters. Various types of optimization techniques have been used in the literature by several researchers such as Genetic Algorithms (GA), Simulated Annealing (SAN) and Particle Swarm Optimization (PSO). In fact, in (Xue 2017), two optimization techniques, genetic algorithms (GA) and particle swarm optimization (PSO), have been used to improve the efficiency of back-propagation neural network model to estimate daily diffuse solar radiation. Results show that the proposed model optimized by PSO performs better than the one optimized by GA. Marzouq et al. proposed evolutionary artificial neural networks for estimation and forecasting solar irradiation, where the authors selected automatically the input parameters in (Marzouq et al. 2019) and ANN architecture (size of input matrix, number of hidden layers and hidden neurons in each layer) in (Marzouq et al. 2020).

### Impact of forecasting horizon

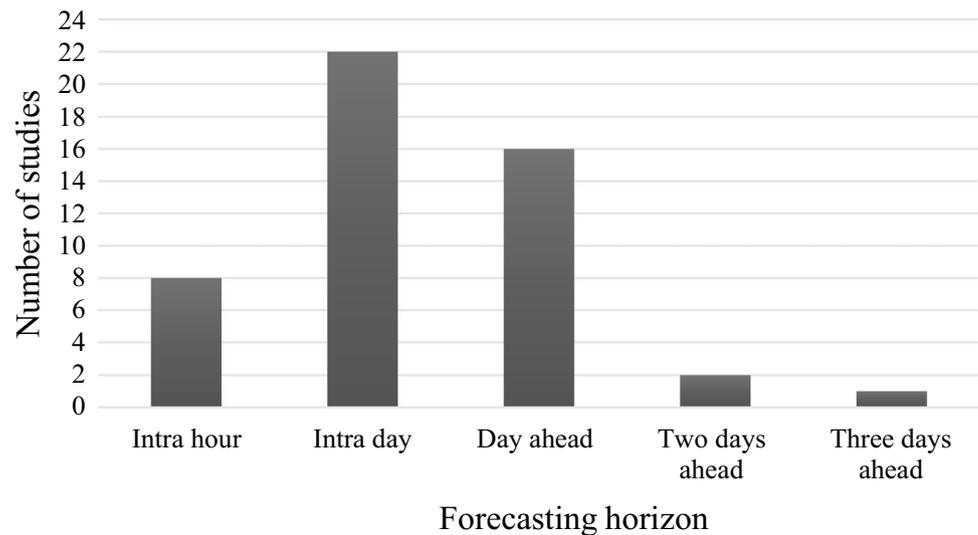
According to published literature, we noticed that forecasting horizon could be divided into five time intervals: intra-hour, intra-day, day ahead, two days ahead, and three days ahead. Every time scale is relevant according to its application in solar systems. Figure 11 shows the number of papers devoted to forecasting in each time interval from the studied literature.

As previously noticed in Tables 5 and 6, forecasting performances depend on the prediction horizon. In fact, forecasting errors rise with the increase in time horizon as shown in the work of Marzouq et al. in (2020) where the values of nRMSE significantly increased from one to six hours ahead. This can be also noticed in the study of Ghofrani et al. in [132] where the forecast was performed from one hour to two days ahead.

### Climate conditions

The availability of solar irradiation data is heavily influenced by clouds motion and weather conditions. In fact, the cloudiness affects the amount of solar irradiation received at the earth's surface (Besharat et al. 2013). As demonstrated in the literature, the forecasting accuracy can change drastically depending on the site climate (Liebermann et al. 2021), (Hong et al. 2020). Certainly, most proposed models perform well on sunny days (Mohammadi et al. 2016a), while on cloudy days the forecasting accuracy decreases significantly. In addition, ANN models present a limitation in forecasting in extremely bad weather conditions (Benmouiza and Chekane 2016). As we can see in the study of Pedro et al. in (Pedro and Coimbra 2015), the forecast error doubles between Merced

**Fig. 11** Number of papers in each time interval for solar irradiation forecasting



(a location with low GHI variability) and Ewa Beach (a location with high GHI variability). These results indicate that low forecast errors are more difficult to obtain for locations where clouds are formed locally. Another study in (Kumari and Toshniwal 2021c) proposed a hybrid LSTM–CNN model where the authors proved that this model presents highly prediction accuracy in different sky conditions (overcast, cloudy, mixed, and sunny) for different climatic locations.

Accordingly, some researchers suggest classifying weather before prediction. In fact, this procedure is essential to increase the forecasting performances and the robustness of used models. As proved by McCandless et al. in (2016), identifying the different cloud regimes by applying the K-means clustering based on surface weather and irradiance data can improve significantly forecasting accuracy. Also, the study of Wang et al. in (2018) forecast solar irradiation in four different weather types (sunny days, cloudy days, rainy days, and heavy days). For sunny days, the hybrid CNN- LSTM model performs well, while for cloudy and rainy days, they used Discrete Wavelet transform as a pre-processing step to decompose data given to the CNN-LSTM model.

On the other hand, the season has a major impact on forecasting performances. In the literature, several authors tested ANN models under each season. In (Yu et al. 2019), the authors evaluated their models under four seasons: spring, summer, fall, and winter. They found that the forecast error decreases in winter and summer and increases in spring and fall. Another study in (Kumari and Toshniwal 2021c) proved that the hybrid LSTM–CNN model outperforms other models such as SP, SVM, MLP, CNN, and LSTM in all seasons for three different locations. Furthermore, this study revealed that the seasonal errors of the proposed model are highly influenced by the site climate.

### Impact of feature selection, data nature and size

The estimation and forecasting ANN models require accurate dataset from studied sites. From our investigation, we noticed that some studies do not provide specifications about used dataset such as total samples and data recording intervals. This information should be considered in future works in order to evaluate prediction models properly. Furthermore, the combination of input parameters is important to enhance model's performance. Therefore, a pre-processing phase followed by a feature selection procedure are necessary to solve this issue and identify the best subset of data fed to the ANN model. In fact, research shows that the performance of ANN models increases when the input parameters have high correlation factor with the solar irradiation.

Regarding input nature, the estimation approaches proposed in the literature are mostly based on exogenous inputs. However, in the forecasting process, we found that authors adopt endogenous, exogenous and hybrid inputs for their models.

The best choice of input parameters in estimation models plays an important role in increasing the prediction performances. Thus, many works used different combinations of parameters to choose input variables that present better prediction accuracy. Based on the finding results of (Jahani and Mohammadi 2019), (Kaba et al. 2018) and (Marzo et al. 2017), sunshine duration was the most relevant parameter and most significant in the estimation process compared to other parameters. However, in (Citakoglu 2015) authors proved that the ANN model with four input parameters: month number (M), extraterrestrial radiation, average air temperature ( $T_m$ ), and average relative humidity ( $RH_m$ ) presents the best result. In addition, this study proved that the month number is the most significant one among the other

variables. On the other hand, (Guijo-Rubio et al. 2020) proposed an approach to estimate solar irradiation based only on satellite data in order to avoid the use of ground data obtained with expensive measuring instruments. This approach can be used to predict solar irradiation in locations with similar data.

Regarding the forecasting procedure, Husein and Chung in (2019a, b) forecast solar irradiation using only exogenous inputs. They proved that the prediction accuracy of the proposed algorithm is dependent on the combination of weather predictor variables and training algorithms. On the other hand, the studies in (Notton et al. 2019), (Huynh et al. 2020), (Bou-Rabee et al. 2017), (Sharma et al. 2016), (Benmouiza and Cheknane 2016), (Azimi et al. 2016), (Pedro and Coimbra 2015) and (Gairaa et al. 2016b) only used endogenous parameters to forecast solar irradiation. Additionally, various authors used mixed parameters (endogenous and exogenous). Indeed, Castangia et al. in (2021a) and Jadidi et al. (Jadidi et al. 2018) proved that adding exogenous inputs to solar irradiation parameters can significantly improve the forecasting performance for prediction horizons greater than 15 min.

Furthermore, in the study of Ahmad et al. in (2015), several combinations of endogenous and exogenous inputs to forecast solar irradiation for 24 h in advance have been used. In addition, Wojtkiewicz et al. in (2019) proved that adding cloud cover as input improves the prediction performance. However, the main challenge of using both endogenous and exogenous inputs is to develop a model with the simplest architecture and the smallest possible number of inputs that can achieve a better accuracy of forecasting (Diez et al. 2020).

From the above studies, the historical input values can also have a significant impact on forecasting accuracy. Therefore, authors need to choose the optimal input matrix dimension. As indicated by Wojtkiewicz et al. in (2019), they tested from one to three historical values as an inputs matrix. The authors found that two previous days of input data give the best performance (using LSTM and GRU) to forecast solar irradiation for one hour ahead. However, there are only a few studies carrying out an automatic selection of previous values used as inputs to their models. For instance, in the study of Marzouq et al. in (2020), the authors used an evolutionary approach to determine the optimal historical data needed. In (Notton et al. 2019), (Fouilloy et al. 2018), (Benali et al. 2019), the authors used the auto mutual information method for this purpose.

### Estimation and forecasting of solar irradiation using ANN versus other Models

In order to show the ability of the proposed approach to forecast and estimate solar irradiation, authors have compared

their results with other machine learning, statistical, and empirical models. In fact, in the studies of Martínez-Castillo et al. in (2021) and of Gürel et al. in (2020), ANNs reveal better accuracy than RF, SVM, MLR, Holt-Winters, RSM, and empirical models. Most studies show that ANN models performances exceed all other methods (see Tables 3, 4, 5, and 6).

On the other hand, the deep learning models such as LSTM, RNN, and GRU exceed other types of ANN models such as MLP and BPNN in terms of performance. As showed in (Castangia et al. 2021a), (Aslam et al. 2019), (Qing and Niu 2018), the LSTM model exceeds BPNN and FFNN. In addition, deep learning models are found to accurately forecast solar irradiation in different climatic conditions (Jallal et al. 2020). The study of Yu et al. in (2019) compared LSTM with ARIMA, SVR, and ANN models and found that LSTM has a strong competitive advantage in forecasting in cloudy days and mixed days in different seasons.

### Recommendations, trends and outlooks for future research

From our review, we noticed that the performances of estimation and forecasting models depend on several factors (see Fig. 8). In this subsection, we will provide concluding remarks, recommendations, and trends deduced from our investigation of the studied papers.

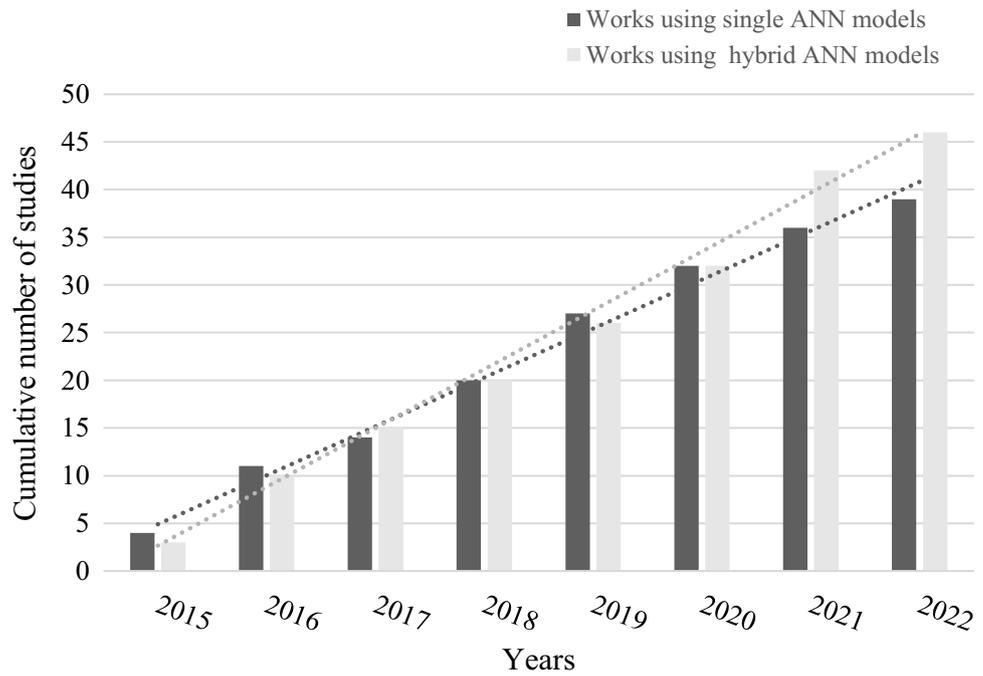
Figure 12 shows the cumulative number of publications per year using single or hybrid ANN models from 2015 to 2022.

- As we can see from Fig. 12, the use of hybrid models increased in comparison with single ANNs, proving that the predictions generated by single ANN models tend to be limited in terms of performances. The key findings of the studied works highlighted the capacity of the hybrid ANN models to significantly improve the performances of the solar irradiation prediction for different climates and seasons.
- To prove the generalization ability of proposed models, researchers should test their models under different weather conditions, seasons, and locations. Nevertheless, in the literature only few works consider data collected from different climates and few compare the model performance under different weather conditions and seasons. These gaps should be taken into consideration in future studies.

Figure 13 shows the cumulative number of works using deep learning models in the studied literature for both estimation and forecasting models.

Arising trend in the use of deep learning models in the solar irradiation forecasting field can be clearly noticed

**Fig. 12** Cumulative number of publications that used ANN models for solar irradiation prediction with single and hybrid ANN models



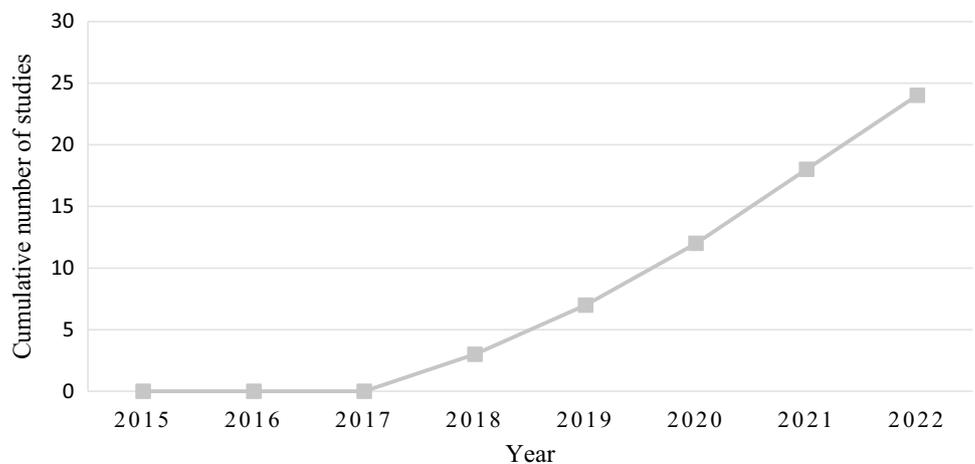
in Fig. 13, where the number of publications using deep learning models increased significantly from 2015 to 2022, which demonstrates that deep learning models proved their ability to perform well under bad weather conditions. Thus, the research community proposed to extend their attention toward deep learning models for forecasting and estimating solar irradiation and its components.

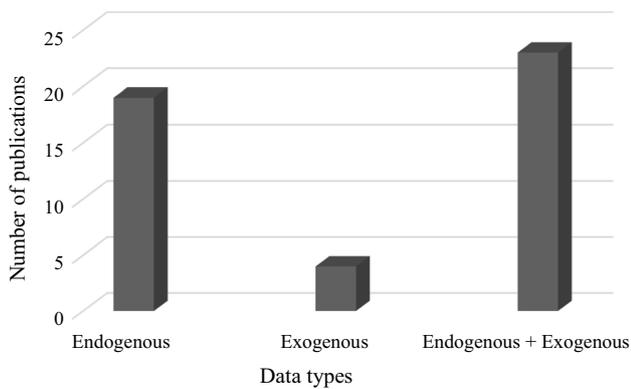
Another important finding through the literature study is that recent works use the cloud cover, temperature and humidity as inputs to their neural network predictor (Wojtkiewicz et al. 2019), which may indicate their effectiveness and pertinence as input features. These variables should be taken into account in the forecasting process, especially in sites with high solar irradiation variability. In fact, many authors found that including these exogenous inputs

improves the forecasting performances (Castangia et al. 2021a),(Jadidi et al. 2018), which explains the trends toward using simultaneously endogenous and exogenous inputs. This trend is clearly shown in Fig. 14 detailing the number of publications versus the types of used inputs. However, the correlation between such parameters and the forecasting output changes from one site to another. Thus, more comparative studies should be carried out to show the effect of including meteorological parameters on the performance of the proposed models to draw a conclusion on this research topic.

- An effective approach that we have identified from our literature study includes investigations on the clouds effect during model development, starting by classifying

**Fig. 13** Number of studies using Deep Learning models for solar irradiation prediction from the studied literature





**Fig. 14** Used data types in studied literature for solar irradiation forecasting

the sky conditions before following up by the forecasting task (McCandless et al. 2016). This decomposition can be performed using data-preprocessing techniques such as K-means clustering and Discrete Wavelet transform.

- From the literature review, we noticed that all proposed works focused only on a specific issue, which is either estimation or forecasting. In (Amiri et al. 2021), Amiri et al. presented for the first time a new approach that could both estimate and forecast solar irradiation. This study is based on a hybrid model with two outputs: one estimates the irradiation at the current instant and another predicts it for the next hour. Despite the complexity of the addressed problem, obtained results are good compared to the literature as claimed by the authors. This approach can be useful for monitoring a solar system in real-time and forecasting future energy production. Further studies simultaneously estimating and forecasting solar irradiation should be addressed to evaluate this proposed approach. This can be a challenging direction in future research.
- In solar irradiation prediction problems, ANN models need to be trained with an accurate and preferably a large database. Moreover, the method to set the model architecture, the number of input features, and the input matrix size to enhance the prediction accuracy remains an open problem, especially with different types of data and different time horizons. As a synthesis, using proper data pre-processing methods, optimization algorithms,

and feature selection techniques has a remarkable impact on the model performance and should be considered as key success factors of incoming models.

## Conclusion

The understanding of solar irradiation predictive methods is of great interest to control and operate solar power generation. In this paper, we provided a comprehensive and in-depth review of the recent studies on estimation and forecasting solar irradiation using ANN models in order to reveal the existing gaps and future suggestions in this field.

Findings show that the performance of predicting models depends on several aspects:

- Pre-processing phase and feature selection procedures used at the first steps of the prediction process,
- selection of ANN architecture, and type of employed model,
- desired forecasting horizon which can lead to higher accuracies for very short horizons,
- climate type of the location of interest,

In most cases, the model prediction accuracy is influenced by the forecast horizon and the climatic conditions. In fact, ANN models performed well under the clear sky while the accuracy decreases significantly under rapid and varying weather conditions. Moreover, it is worth noting that solar irradiation phenomenon is extremely complex and is not necessarily well modeled with simple approaches, especially when high performances are required. Thus, combining several approaches might be necessary to seek greater accuracy. Further, in the last few years, several studies moved toward the use of deep neural networks such as LSTM and RNN to overcome some limitations of conventional ANNs.

The state of the art of solar irradiation prediction presented in this paper can enrich and chart a path toward promising and suitable directions for future planners and forecasting professionals to develop an appropriate ANN predictor. Additionally, the offered directions will inspire future researches that can improve the prediction models toward more accuracy and efficiency.

## Appendix

### Nomenclature

ACRT	ACR data logger temperature	DA	Dragonfly Algorithm
ANFIS	Adaptive neuro fuzzy inference system	DBT	Dry bulb temperature
Alt	Altitude	E	Earth skin temperature
$\alpha$	Angle of inclination	ESN	Echo State Network ( Recurrent Neural Network (RNN))
ANN	Artificial Neural Networks	EMD	Empirical Mode Decomposition
P	Atmospheric pressure	EEMD	Ensemble empirical mode decomposition
ARIMA	Auto regressive integrated moving average models	Evap	Evaporation
AR	Auto regressive	EATE	Evolutionary algorithms with tournament selection and elitism
ARMA	Autoregressive and moving average model	ETS	Exponential trend smoothing
ARX	Autoregressive exogenous	$H_o$	Extraterrestrial radiation ( $KW.hour/m^2.D$ )
AA	Average airmass	ELM	Extreme Learning Machine
$Az_m$	Average/Mean azimuth angle	FFNN	Feedforward Neural Network
$DPT_m$	Average dew point temperature	$F(T_m)$	Function of $T_m$
$Az$	Azimuth angle	FCM	Fuzzy c-means algorithm
$PWS_m$	Average peak wind speed	GPV	Gaussian process vector
$RF_m$	Average/Mean rainfall	GRU	Gated recurrent unit
$RH_m$	Average relative humidity	GRNN	Generalized regression neural network
ART	Average roof temperature	GA	Genetic algorithm
$SD_m$	Average/Mean sunshine duration	GP	Genetic programming
$T_m$	Average/ Mean temperature	GHI	Global horizontal solar irradiation
$WCT_m$	Average wind chill temperature	BNI	Global normal irradiation
$WS_m$	Average/Mean wind speed	GTI	Global Tilted Irradiation/Irradiance
$\theta_{zm}$	Average zenith angel	GBMs	Gradient boosting machines
ANFIS-muSG	ANFIS-SSA-GOA	GOA	Grasshopper Optimization Algorithm
BPNN	Backpropagation neural network	GWO	Grey Wolf Optimizer
BSRN	Baseline Surface Radiation Network	H	Hours
Bi-LSTM	Bidirectional long short term memory	$\omega$	Hour angle
BDT	Boosted decision tree	$CI_h$	Hourly clearness index
Cs	Clear-sky	$D_h$	Hours of day
$K_t$	Clearness index	KELM	Kernel extreme learning machine
CC	Cloud cover	KNN	k-Nearest Neighbours
CI	Cloud index	Lat	Latitude
Comp	Component	LM	Levenberg–Marquardt
CEEMDAN	Complete ensemble empirical mode decomposition adaptive noise	Long	Longitude
CNN	Convolutional neural network	LSTM	Long short-term memory
DD	Day duration	ME	Maximum elevation (ME)
$GI_d$	Daily values of total global radiation	$P_m$	Mean station level pressure
D	Day	MEA	Mind evolutionary algorithm
DoM	Day of month	$P_{min}$	Minimum pressure
$\delta$	Declination angle	$M_{th}$	Month
DPT	Dew point temperature	MoY	Month of the year
$\Delta T$	Difference of daily maximum and minimum temperature	MLP	Multilayer perceptron

## Nomenclature

DHI	Diffuse horizontal irradiation	NREL	National Renewable Energy Laboratory
DFT	Discrete Fourier transform	$P_{\max}$	Maximum pressure
DWT	Discrete wavelet transformation	$SD_{\max}$	Maximum sunshine duration
DSN	Distance from Solar Noon	$T_{\max}$	Maximum temperatures
MOSMLP	Model Output Statistics multilayer perceptron based on ANN and Numerical Weather Prediction	SAN	Simulated annealing
$T_{\min}$	Minimum temperatures	SC	Sky cover
NAR	Nonlinear autoregressive neural network	SP	Smart persistence
N	Number	SA	Solar altitude angle ( $\alpha$ )
PSO	Particle swarm optimization	GI	Solar irradiation/irradiance/ radiation
Per	Persistence	STMLP	Statistical model based on multilayer perceptron
P	Pressure	SD	Sunshine duration
PCA	Principal component analysis	SR	Sunshine Ratio
PUNN	Product Unit Neural Network	$SD_{\text{day}}$	Sunshine duration per day (hour)
RBF	Radial basis function	SVM	Support vector machine
RBFNN	Radial Basis Function Neural Network	$T_a$	Temperature ambient
RF	Rainfall	$BNI_{\text{th}}$	Theoretical Global normal irradiation
RFR	Random forest regression	$SD_{\text{th}}$	Theoretical sunshine duration
RNN	Recurrent neural network	$T_d$	Time of day
VIS 0.6 and VIS 0.8	Reflectivity	TB K-means	Transformation based K-means algorithm
RT	Regression trees	UV	Ultraviolet index
FOS-ELM	Regularized online sequential extreme learning machine with variable forgetting factor	$V_{\text{PD}}$	Vapor pressure deficiency
$RH_{\max}$	Relative humidity Maximum	$P_{\text{wv}}$	Water vapor pressure
$RH_{\min}$	Relative humidity minimum	WNN	Wavelet neural network
RP	Relative position among target and chosen locations	WPD	Wavelet Packet Decomposition
RBP	Resilient back propagation	WT	Weather type
SSA	Salp Swarm Algorithm	WD	Wind direction
SCG	Scaled conjugate gradient	WS	Wind speed
SaDE-ELM	Self-adaptive differential evolutionary ELM	$\theta_z$	Zenith angle
SUNN	Sigmoidal Unit Neural Network	LSR	Linear least square regression

**Acknowledgements** The authors thank the anonymous reviewers for their helpful comments.

**Author contribution** Naima El-Amarty: Conceptualization, Writing—original draft, Methodology, Writing—review & editing.

Manal Marzouq: Conceptualization, Writing—original draft, Methodology, review and Validation.

Hakim El Fadili: Conceptualization, Writing—original draft, Methodology, Supervision, review and Validation.

Saad Dosse Bennani: Conceptualization, Investigation, Supervision, Validation.

Antonio Ruano: Conceptualization, Investigation, Supervision, Validation.

**Funding** The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

**Data availability** Not applicable.

## Declarations

**Ethics approval** Not applicable.

**Consent to participate** Not applicable.

**Consent for publication** Not applicable.

**Competing interests** The authors declare no competing interests.

## References

- Ağbulut Ü, Gürel AE, Biçen Y (2021) Prediction of daily global solar radiation using different machine learning algorithms: evaluation and comparison. *Renew Sustain Energy Rev* 135(March 2020). <https://doi.org/10.1016/j.rser.2020.110114>

- Ahmad A, Anderson TN, Lie TT (2015) Hourly global solar irradiation forecasting for New Zealand. *Sol Energy* 122:1398–1408. <https://doi.org/10.1016/j.solener.2015.10.055>
- Alani El, Omaima MA, Ghennioui H, Ghennioui A, Ikenbi I, Dahr F-E (2021) Short term solar irradiance forecasting using sky images based on a hybrid CNN–MLP model. *Energy Rep* 7(May):888–900. <https://doi.org/10.1016/j.egy.2021.07.053>
- Ali-Ou-Salah H, Oukarfi B, Tlemcani M (2021) Short-Term solar radiation forecasting using a new seasonal clustering technique and artificial neural network. *Int J Green Energy* 18(10):1–11. <https://doi.org/10.1080/15435075.2021.1946819>
- Almaraashi M (2018) Investigating the impact of feature selection on the prediction of solar radiation in different locations in Saudi Arabia. *Appl Soft Comput J* 66:250–263. <https://doi.org/10.1016/j.asoc.2018.02.029>
- Amiri B, Gómez-Orellana AM, Gutiérrez PA, Dizène R, Hervás-Martínez C, Dahmani K (2021) A novel approach for global solar irradiation forecasting on tilted plane using hybrid evolutionary neural networks. *J Clean Prod* 287:125577. <https://doi.org/10.1016/j.jclepro.2020.125577>
- Aslam M, Lee JM, Kim HS, Lee SJ, Hong S (2019) Deep learning models for long-term solar radiation forecasting considering microgrid installation: a comparative study. *Energies* 13(1). <https://doi.org/10.3390/en13010147>
- Azimi R, Ghayekhloo M, Ghofrani M (2016) A hybrid method based on a new clustering technique and multilayer perceptron neural networks for hourly solar radiation forecasting. *Energy Convers Manage* 118:331–344. <https://doi.org/10.1016/j.enconman.2016.04.009>
- Barbieri F, Rajakaruna S, Ghosh A (2017) Very short-term photovoltaic power forecasting with cloud modeling: a review. *Renew Sustain Energy Rev* 75(August 2015):242–63. <https://doi.org/10.1016/j.rser.2016.10.068>
- Benali L, Notton G, Fouilloy A, Voyant C, Dizene R (2019) Solar radiation forecasting using artificial neural network and random forest methods: application to normal beam, horizontal diffuse and global components. *Renew Energy* 132:871–884. <https://doi.org/10.1016/j.renene.2018.08.044>
- Benmouiza K, Chekane A (2016) Small-scale solar radiation forecasting using ARMA and nonlinear autoregressive neural network models. *Theoret Appl Climatol* 124(3–4):945–958. <https://doi.org/10.1007/s00704-015-1469-z>
- Besharat F, Dehghan AA, Faghih AR (2013) Empirical models for estimating global solar radiation: a review and case study. *Renew Sustain Energy Rev* 21:798–821. <https://doi.org/10.1016/j.rser.2012.12.043>
- Bhatt Ankit, WeerakornOngsakul M, Madhu Nimal, Singh Jai Govind (2022) Sliding window approach with first-order differencing for very short-term solar irradiance forecasting using deep learning models. *Sustain Energy Technol Assess* 50(November 2021):101864. <https://doi.org/10.1016/j.seta.2021.101864>
- Blaga R, Sabadus A, Stefu N, Dughir C, Paulescu M (2019a) A current perspective on the accuracy of incoming solar energy forecasting. *Prog Energy Combust Sci* 70:119–144. <https://doi.org/10.1016/j.pecs.2018.10.003>
- Blaga R, Sabadus A, Stefu N, Dughir C, Paulescu M, Badescu V (2019b) A current perspective on the accuracy of incoming solar energy forecasting. *Prog Energy Combust Sci* 70:119–144. <https://doi.org/10.1016/j.pecs.2018.10.003>
- Bou-Rabee M, Sulaiman SA, Saleh MS, Marafi S (2017) Using artificial neural networks to estimate solar radiation in Kuwait. *Renew Sustain Energy Rev* 72(January):434–438. <https://doi.org/10.1016/j.rser.2017.01.013>
- Bouzgou H, Gueymard CA (2017) Minimum redundancy – maximum relevance with extreme learning machines for global solar radiation forecasting: toward an optimized dimensionality reduction for solar time series. *Sol Energy* 158(September):595–609. <https://doi.org/10.1016/j.solener.2017.10.035>
- Caldas M, Alonso-Suárez R (2019) Very short-term solar irradiance forecast using all-sky imaging and real-time irradiance measurements. *Renew Energy* 143:1643–1658. <https://doi.org/10.1016/j.renene.2019.05.069>
- Castangia M, Aliberti A, Bottaccioli L, Macii E, Patti E (2021a) A compound of feature selection techniques to improve solar radiation forecasting. *Expert Syst Appl* 178(March). <https://doi.org/10.1016/j.eswa.2021a.114979>
- Castangia M, Aliberti A, Bottaccioli L, Macii E, Patti E (2021b) A compound of feature selection techniques to improve solar radiation forecasting. *Expert Syst Appl* 178(January). <https://doi.org/10.1016/j.eswa.2021b.114979>
- Cebecauer T, Suri M (2015) Typical meteorological year data: Solar-GIS Approach. *Energy Procedia* 69:1958–1969. <https://doi.org/10.1016/j.egypro.2015.03.195>
- Chandrashekar G, Sahin F (2014) A survey on feature selection methods. *Comput Electr Eng* 40(1):16–28. <https://doi.org/10.1016/j.compeleceng.2013.11.024>
- Citakoglu H (2015) Comparison of artificial intelligence techniques via empirical equations for prediction of solar radiation. *Comput Electron Agric* 118:28–37. <https://doi.org/10.1016/j.compag.2015.08.020>
- Cornaro C, Pierro M, Bucci F (2015) Master optimization process based on neural networks ensemble for 24-h solar irradiance forecast. *Sol Energy* 111:297–312. <https://doi.org/10.1016/j.solener.2014.10.036>
- Dahmani K, Notton G, Voyant C, Dizene R, Nivet ML, Paoli C, Tamas W (2016) Multilayer perceptron approach for estimating 5-min and hourly horizontal global irradiation from exogenous meteorological data in locations without solar measurements. *Renew Energy* 90:267–282. <https://doi.org/10.1016/j.renene.2016.01.013>
- Das UK, Tey KS, Seyedmahmoudian M, Mekhilef S, Idris MYI, Van Deventer W, Horan B, Stojcevski A (2018) Forecasting of photovoltaic power generation and model optimization: a review. *Renew Sustain Energy Rev* 81(June 2017):912–28. <https://doi.org/10.1016/j.rser.2017.08.017>
- Diagne HM, David M, Lauret P, Boland J (2012) Solar Irradiation Forecasting: State-of-the-Art and Proposition for Future Developments for Small-Scale Insular Grids. *World Renewable Energy Forum, WREF 2012, Including World Renewable Energy Congress XII and Colorado Renewable Energy Society (CRES) Annual Conferen* 6:4802–9
- Diagne M, David M, Lauret P, Boland J, Schmutz N (2013) Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. *Renew Sustain Energy Rev* 27:65–76. <https://doi.org/10.1016/j.rser.2013.06.042>
- Diez FJ, Navas-Gracia LM, Chico-Santamarta L, Correa-Guimaraes A, Martínez-Rodríguez A (2020) Prediction of horizontal daily global solar irradiation using artificial neural networks (ANNs) in the Castile and León Region, Spain. *Agronomy* 10(1). <https://doi.org/10.3390/agronomy10010096>
- Eşlik AH, Akarslan E, Hocaoglu FO (2022) Short-term solar radiation forecasting with a novel image processing-based deep learning approach. *Renew Energy* 200(May):1490–1505. <https://doi.org/10.1016/j.renene.2022.10.063>
- Etxegarai G, López A, Aginako N, Rodríguez F (2022) Energy for sustainable development an analysis of different deep learning neural networks for intra-hour solar irradiation forecasting to compute solar photovoltaic generators' Energy Production. *Energy Sustain Dev* 68:1–17. <https://doi.org/10.1016/j.esd.2022.02.002>
- Feng Y, Gong D, Zhang Q, Jiang S, Zhao L, Cui N (2019) Evaluation of Temperature-based machine learning and empirical models for predicting daily global solar radiation. *Energy Convers Manage*

- 198(April):111780. <https://doi.org/10.1016/j.enconman.2019.111780>
- Ferreira PM, Gomes JM, Martins IAC, Ruano AE (2012) A neural network based intelligent predictive sensor for cloudiness, solar radiation and air temperature. *Sensors (switzerland)* 12(11):15750–15777. <https://doi.org/10.3390/s121115750>
- Ferreira PM, Ruano AE (2011) Evolutionary multiobjective neural network models identification: evolving task-optimised models. *Stud Comput Intell* 372:21–53. [https://doi.org/10.1007/978-3-642-11739-8\\_2](https://doi.org/10.1007/978-3-642-11739-8_2)
- Fouilloy A, Voyant C, Notton G, Motte F, Paoli C, Nivet ML, Guillot E, Duchaud JL (2018) Solar irradiation prediction with machine learning: forecasting models selection method depending on weather variability. *Energy* 165:620–629. <https://doi.org/10.1016/j.energy.2018.09.116>
- Gairaa K, Khellaf A, Messlem Y, Chellali F (2016a) Estimation of the daily global solar radiation based on Box-Jenkins and ANN models: a combined approach. *Renew Sustain Energy Rev* 57:238–249. <https://doi.org/10.1016/j.rser.2015.12.111>
- Gairaa K, Khellaf A, Messlem Y, Chellali F (2016b) Estimation of the daily global solar radiation based on Box – Jenkins and ANN models : a combined approach. *Renew Sustain Energy Rev* 57:238–249. <https://doi.org/10.1016/j.rser.2015.12.111>
- Gallo R, Castangia M, Macii A, Macii E, Patti E, Aliberti A (2022) Solar radiation forecasting with deep learning techniques integrating geostationary satellite images. *Eng Appl Artif Intell* 116(March):105493. <https://doi.org/10.1016/j.engappai.2022.105493>
- 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. <https://doi.org/10.1016/j.renene.2020.09.141>
- Garcia-Hinde O, Gomez-Verdejo V, Martinez-Ramon M, Casanova-Mateo C, Sanz-Justo J, Jimenez-Fernandez S, Salcedo-Sanz S (2016) “Feature selection in solar radiation prediction using bootstrapped SVRs”. 2016 IEEE Congress on Evolutionary Computation. CEC 2016:3638–3645. <https://doi.org/10.1109/CEC.2016.7744250>
- García S, Luengo J, Herrera F (2015) Data preprocessing in data mining, vol 72. Springer International Publishing, Cham, Switzerland, pp 59–139
- Gbémou S, Eynard J, Thil S, Guillot E, Griou S (2021) A comparative study of machine learning-based methods for global horizontal irradiance forecasting. *Energies* 14(11). <https://doi.org/10.3390/en14113192>
- Geetha A, Santhakumar J, Mohana Sundaram K, Usha S, Thamiz TM, Thentral CS, Boopathi RR, Sathyamurthy R (2022) Prediction of hourly solar radiation in tamil nadu using ANN model with different learning algorithms. *Energy Rep* 8:664–671. <https://doi.org/10.1016/j.egy.2021.11.190>
- Ghimire S, Deo RC, Raj N, Mi J (2019) Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms. *Appl Energy* 253(June):113541. <https://doi.org/10.1016/j.apenergy.2019.113541>
- Ghimire S, Deo RC, Downs NJ, 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(April):176–198. <https://doi.org/10.1016/j.rse.2018.05.003>
- Ghofrani M, Ghayekhloo M, Azimi R (2016) A novel soft computing framework for solar radiation forecasting. *Appl Soft Comput J* 48:207–216. <https://doi.org/10.1016/j.asoc.2016.07.022>
- Guermoui M, Gairaa K, Rabehi A, Djafer D, Benkaciali S (2018) Estimation of the daily global solar radiation based on the gaussian process regression methodology in the saharan climate. *Eur Phys J Plus* 133(6). <https://doi.org/10.1140/epjp/i2018-12029-7>
- Guermoui M, Melgani F, Gairaa K, Mekhalif ML (2020) A comprehensive review of hybrid models for solar radiation forecasting. *J Clean Prod* 258. <https://doi.org/10.1016/j.jclepro.2020.120357>
- Guijo-Rubio D, Durán-Rosal AM, Gutiérrez PA, Gómez-Orellana AM, Casanova-Mateo C, Sanz-Justo J, Salcedo-Sanz S, Hervás-Martínez C (2020) Evolutionary artificial neural networks for accurate solar radiation prediction. *Energy* 210. <https://doi.org/10.1016/j.energy.2020.118374>
- Gupta A, Gupta K, Saroha S (2021) A review and evaluation of solar forecasting technologies. *Mater Today: Proc* 47:2420–2425. <https://doi.org/10.1016/j.matpr.2021.04.491>
- Gürel AE, Ağbulut Ü, Biçen Y (2020) Assessment of machine learning, time series, response surface methodology and empirical models in prediction of global solar radiation. *J Clean Prod* 277. <https://doi.org/10.1016/j.jclepro.2020.122353>
- Gutiérrez-Corea FV, Manso-Callejo MA, Moreno-Regidor MP, Manrique-Sancho MT (2016) Forecasting short-term solar irradiance based on artificial neural networks and data from neighboring meteorological stations. *Sol Energy* 134:119–131. <https://doi.org/10.1016/j.solener.2016.04.020>
- Hong YY, Martinez JFF, Fajardo AC (2020) Day-ahead solar irradiation forecasting utilizing gramian angular field and convolutional long short-term memory. *IEEE Access* 8:18741–53. <https://doi.org/10.1109/ACCESS.2020.2967900>
- Hou M, Zhang T, Weng F, Ali M, Al-Ansari N, Yaseen ZM (2018) Global solar radiation prediction using hybrid online sequential extreme learning machine model. *Energies* 11(12). <https://doi.org/10.3390/en11123415>
- Huang X, Li Q, Tai Y, Chen Z, Zhang J, Shi J, Gao B, Liu W (2021) Hybrid deep neural model for hourly solar irradiance forecasting. *Renewable Energy* 171:1041–1060. <https://doi.org/10.1016/j.renene.2021.02.161>
- Husein M, Chung I-Y (2019a) Day-ahead solar irradiance forecasting for microgrids using a long short-term memory recurrent neural network : a deep learning approach. <https://doi.org/10.3390/en12101856>
- Husein M, Chung IY (2019b) Day-Ahead solar irradiance forecasting for microgrids using a long short-term memory recurrent neural network: a deep learning approach. *Energies* 12(10). <https://doi.org/10.3390/en12101856>
- Hussain S, AlAlili A (2017) A hybrid solar radiation modeling approach using wavelet multiresolution analysis and artificial neural networks. *Appl Energy* 208(September):540–550. <https://doi.org/10.1016/j.apenergy.2017.09.100>
- Huynh ANL, Deo RC, An-Vo DA, Ali M, Raj N, Abdulla S (2020) Near real-time global solar radiation forecasting at multiple time-step horizons using the long short-term memory network. *Energies* 13(14). <https://doi.org/10.3390/en13143517>
- Ibrahim IA, Khatib T (2017) A novel hybrid model for hourly global solar radiation prediction using random forests technique and firefly algorithm. *Energy Convers Manage* 138:413–425. <https://doi.org/10.1016/j.enconman.2017.02.006>
- Jadidi A, Menezes R, De Souza N, Lima ACDC (2018) A hybrid GAMLPPNN model for one-hour-ahead forecasting of the global horizontal irradiance in Elizabeth City, North Carolina. *Energies* 11(10):1–18. <https://doi.org/10.3390/en11102641>
- Jahani B, Mohammadi B (2019) A comparison between the application of empirical and ANN methods for estimation of daily global solar radiation in Iran. *Theoret Appl Climatol* 137(1–2):1257–1269. <https://doi.org/10.1007/s00704-018-2666-3>
- Jaihuni M, Basak JK, Khan F, Okyere FG, Sihalath T, Bhujel A, Park J, Lee DH, Kim HT (2022) A novel recurrent neural network approach in forecasting short term solar irradiance. *ISA Trans* 121:63–74. <https://doi.org/10.1016/j.isatra.2021.03.043>

- Jallal MA, Chabaa S, Zeroual A (2020) A New artificial multi-neural approach to estimate the hourly global solar radiation in a semi-arid climate site. *Theoret Appl Climatol* 139(3–4):1261–1276. <https://doi.org/10.1007/s00704-019-03033-1>
- Jeon BK, Kim EJ (2020) Next-day prediction of hourly solar irradiance using local weather forecasts and LSTM trained with non-local data. *Energies* 13(20). <https://doi.org/10.3390/en13205258>
- Jiang Y (2009) Computation of monthly mean daily global solar radiation in China using artificial neural networks and comparison with other empirical models. *Energy* 34(9):1276–1283. <https://doi.org/10.1016/j.energy.2009.05.009>
- Jumin E, Basaruddin FB, Yusoff YBMD, Latif SD, Ahmed AN (2021) Solar radiation prediction using boosted decision tree regression model: a case study in Malaysia. *Environ Sci Pollut Res* 28(21):26571–83. <https://doi.org/10.1007/s11356-021-12435-6>
- Kaba K, Sarigül M, Avcı M, Mustafa Kandırmaz H (2018) Estimation of daily global solar radiation using deep learning model. *Energy* 162:126–135. <https://doi.org/10.1016/j.energy.2018.07.202>
- Kamadinata JO, Ken TL, Suwa T (2019) Sky image-based solar irradiance prediction methodologies using artificial neural networks. *Renew Energy* 134:837–845. <https://doi.org/10.1016/j.renene.2018.11.056>
- Karagiannopoulos M, Anyfantis D, Kotsiantis SB, Pintelas PE (2007) Feature Selection for regression problems.” The 8th Hellenic European Research on Computer Mathematics & Its Applications, HERCMA 2007 (i):20–22
- Khaire UM, Dhanalakshmi R (2019) Stability of feature selection algorithm: a review. *J King Saud Univ-Comput Inf Sci* (xxxx). <https://doi.org/10.1016/j.jksuci.2019.06.012>
- Khelifi R, Guermoui M, Rabehi A, Lalmi D (2020) Multi-step-ahead forecasting of daily solar radiation components in the saharan climate. *Int J Ambient Energy* 41(6):707–715. <https://doi.org/10.1080/01430750.2018.1490349>
- Kılıç F, Yılmaz İH, Kaya Ö (2021) Adaptive co-optimization of artificial neural networks using evolutionary algorithm for global radiation forecasting. *Renew Energy* 171:176–190. <https://doi.org/10.1016/j.renene.2021.02.074>
- Kleissl J (2013) Solar energy forecasting and resource assessment. Academic Press
- Kuhe A, Achirgenda VT, Agada M (2021) Global solar radiation prediction for Makurdi, Nigeria, using neural networks ensemble. *Energy Sour Part A: Recove Util Environ Effects* 43(11):1373–1385. <https://doi.org/10.1080/15567036.2019.1637481>
- Kumar DS, Yagli GM, Kashyap M, Srinivasan D (2020) Solar irradiance resource and forecasting: a comprehensive review. *IET Renew Power Gener* 14(10):1641–1656. <https://doi.org/10.1049/iet-rpg.2019.1227>
- Kumari P, Toshniwal D (2021a) Deep learning models for solar irradiance forecasting: a comprehensive review. *J Clean Prod* 318(August):128566. <https://doi.org/10.1016/j.jclepro.2021.128566>
- Kumari P, Toshniwal D (2021b) Extreme gradient boosting and deep neural network based ensemble learning approach to forecast hourly solar irradiance. *J Clean Prod* 279:123285. <https://doi.org/10.1016/j.jclepro.2020.123285>
- Kumari P, Toshniwal D (2021c) Long short term memory–convolutional neural network based deep hybrid approach for solar irradiance forecasting. *Appl Energy* 295(January):117061. <https://doi.org/10.1016/j.apenergy.2021.117061>
- Lai JP, Chang YM, Chen CH, Pai PF (2020) A survey of machine learning models in renewable energy predictions. *Appl Sci (Switzerland)* 10(17). <https://doi.org/10.3390/app10175975>
- Lan H, Yin H, Hong YY, Wen S, Yu DC, Cheng P (2018) Day-ahead spatio-temporal forecasting of solar irradiation along a navigation route. *Appl Energy* 211(May 2017):15–27. <https://doi.org/10.1016/j.apenergy.2017.11.014>
- Lan H, Zhang C, Hong YY, He Y, Wen S (2019) Day-ahead spatio-temporal solar irradiation forecasting using frequency-based hybrid principal component analysis and neural network. *Appl Energy* 247(January):389–402. <https://doi.org/10.1016/j.apenergy.2019.04.056>
- Liebermann S, Um J, Hwang Y, Schlüter S (2021) Performance evaluation of neural network-based short-term solar irradiation forecasts. *Energies* 14:1–21. <https://doi.org/10.3390/en1413030>
- Marquez R, Coimbra CFM (2013) Intra-hour DNI forecasting based on cloud tracking image analysis. *Sol Energy* 91:327–336. <https://doi.org/10.1016/j.solener.2012.09.018>
- Martinez-Castillo C, Astray G, Mejuto JC (2021) Modelling and prediction of monthly global irradiation using different prediction models. *Energies* 14(8):1–15. <https://doi.org/10.3390/en14082332>
- Marzo A, Trigo M, Alonso-Montesinos J, Martínez-Durbán M, López G, Ferrada P, Fuentealba E, Cortés M, Batlles FJ (2017) Daily global solar radiation estimation in desert areas using daily extreme temperatures and extraterrestrial radiation. *Renew Energy* 113:303–311. <https://doi.org/10.1016/j.renene.2017.01.061>
- Marzouq M, Bounoua Z, Hakim El Fadili A, Mechaqrane KZ, Lakhliai Z (2019) New Daily global solar irradiation estimation model based on automatic selection of input parameters using evolutionary artificial neural networks. *J Clean Prod* 209:1105–1118. <https://doi.org/10.1016/j.jclepro.2018.10.254>
- Marzouq M, El Fadili H, Zenkour K, Lakhliai Z, Amouzg M (2020) Short Term solar irradiance forecasting via a novel evolutionary multi-model framework and performance assessment for sites with no solar irradiance data. *Renew Energy* 157:214–231. <https://doi.org/10.1016/j.renene.2020.04.133>
- Mazorra Aguiar L, Pereira B, David M, Díaz F, Lauret P (2015) Use of satellite data to improve solar radiation forecasting with bayesian artificial neural networks. *Sol Energy* 122:1309–1324. <https://doi.org/10.1016/j.solener.2015.10.041>
- McCandless TC, Haupt SE, Young GS (2016) A regime-dependent artificial neural network technique for short-range solar irradiance forecasting. *Renew Energy* 89:351–359. <https://doi.org/10.1016/j.renene.2015.12.030>
- Meenal R, Immanuel Selvakumar A (2018) Assessment of SVM, empirical and ANN Based solar radiation prediction models with most influencing input parameters. *Renew Energy* 121:324–343. <https://doi.org/10.1016/j.renene.2017.12.005>
- El Mghouchi Y, Chham E, Zemmouri EM, El Bouardi A (2019) Assessment of different combinations of meteorological parameters for predicting daily global solar radiation using artificial neural networks. *Build Environ* 149(September 2018):607–22. <https://doi.org/10.1016/j.buildenv.2018.12.055>
- El Mghouchi Y, Chham E, Zemmouri EM, El Bouardi A (2019) Assessment of different combinations of meteorological parameters for predicting daily global solar radiation using artificial neural networks. *Build Environ* 149(September 2018):607–22. <https://doi.org/10.1016/j.buildenv.2018.12.055>
- Mohammadi K, Shamshirband S, Kamsin A, Lai PC, Mansor Z (2016) Identifying the most significant input parameters for predicting global solar radiation using an ANFIS selection procedure. *Renew Sustain Energy Rev* 63:423–34. <https://doi.org/10.1016/j.rser.2016.05.065>
- Mohammadi K, Shamshirband S, Petković D, Khorasanizadeh H (2016b) Determining the most important variables for diffuse solar radiation prediction using adaptive neuro-fuzzy methodology; case study: city of Kerman, Iran. *Renew Sustain Energy Rev* 53:1570–1579. <https://doi.org/10.1016/j.rser.2015.09.028>
- Monjoly S, André M, Calif R, Soubdhan T (2017) Hourly Forecasting of global solar radiation based on multiscale decomposition methods: a hybrid approach. *Energy* 119:288–298. <https://doi.org/10.1016/j.energy.2016.11.061>

- Mousavi SM, Mostafavi ES, Jiao P (2017) Next generation prediction model for daily solar radiation on horizontal surface using a hybrid neural network and simulated annealing method. *Energy Convers Manage* 153(August):671–682. <https://doi.org/10.1016/j.enconman.2017.09.040>
- Notton G, Voyant C, Fouilloy A, Duchaud JL, Nivet ML (2019) Some applications of ANN to solar radiation estimation and forecasting for energy applications. *Appl Sci (Switzerland)* 9(1). <https://doi.org/10.3390/app9010209>
- Ozoegwu CG (2019) Artificial neural network forecast of monthly mean daily global solar radiation of selected locations based on time series and month number. *J Clean Prod* 216:1–13. <https://doi.org/10.1016/j.jclepro.2019.01.096>
- 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. <https://doi.org/10.1016/j.renene.2020.04.042>
- Pazikadin AR, Rifai D, Ali K, Malik MZ, Abdalla AN, Faraj MA (2020) Solar irradiance measurement instrumentation and power solar generation forecasting based on artificial neural networks (ANN): a review of five years research trend. *Sci Total Environ* 715:136848. <https://doi.org/10.1016/j.scitotenv.2020.136848>
- Pedregal DJ, Trapero JR (2021) Adjusted combination of moving averages: a forecasting system for medium-term solar irradiance. *Appl Energy* 298(June):117155. <https://doi.org/10.1016/j.apenergy.2021.117155>
- Pedro HTC, Coimbra CFM (2015) Short-term irradiance forecastability for various solar micro-climates. *Sol Energy* 122:587–602. <https://doi.org/10.1016/j.solener.2015.09.031>
- Premalatha N, Arasu AV (2016) Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms. *J Appl Res Technol* 14(3):206–214. <https://doi.org/10.1016/j.jart.2016.05.001>
- Qazi A, Fayaz H, Wadi A, Raj RG, Rahim NA, Khan WA (2015) The artificial neural network for solar radiation prediction and designing solar systems: a systematic literature review. *J Clean Prod* 104:1–12. <https://doi.org/10.1016/j.jclepro.2015.04.041>
- Qing X, Niu Y (2018) Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy* 148:461–468. <https://doi.org/10.1016/j.energy.2018.01.177>
- Quej VH, Almorox J, Arnaldo JA, Saito L (2017) ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment. *J Atmos Solar Terr Phys* 155(February):62–70. <https://doi.org/10.1016/j.jastp.2017.02.002>
- Rabehi A, Guermoui M, Lalmi D (2020) Hybrid models for global solar radiation prediction: a case study. *Int J Ambient Energy* 41(1):31–40. <https://doi.org/10.1080/01430750.2018.1443498>
- Rajagukguk RA, Ramadhan RAA, Lee HJ (2020) A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power. *Energies* 13(24). <https://doi.org/10.3390/en13246623>
- Ramadhan RAA, Heatubun YRJ, Tan SF, Lee HJ (2021) Comparison of physical and machine learning models for estimating solar irradiance and photovoltaic power. *Renew Energy* 178:1006–1019. <https://doi.org/10.1016/j.renene.2021.06.079>
- Rangel E, Cadenas E, Campos-Amezcuca R, Tena JL (2020) Enhanced prediction of solar radiation using NARX models with corrected input vectors. *Energies* 13(10):1–22. <https://doi.org/10.3390/en13102576>
- Rao K, Krishna SDV, Premalatha M, Naveen C (2018) Analysis of different combinations of meteorological parameters in predicting the horizontal global solar radiation with ANN approach: a case study. *Renew Sustain Energy Rev* 91(July 2017):248–58. <https://doi.org/10.1016/j.rser.2018.03.096>
- Rocha PAC, Fernandes JL, Modolo AB, Lima RJP, da Silva MEV, Bezerra CAD (2019) Estimation of daily, weekly and monthly global solar radiation using ANNs and a long data set: a case study of Fortaleza, in Brazilian Northeast Region. *Int J Energy Environ Eng* 10(3):319–34. <https://doi.org/10.1007/s40095-019-0313-0>
- Rodríguez F, Martín F, Fontán L, Galarza A (2021) Ensemble of machine learning and spatiotemporal parameters to forecast very short-term solar irradiation to compute photovoltaic generators' Output Power. *Energy* 229. <https://doi.org/10.1016/j.energy.2021.120647>
- Saeys Y, Inza I, Larrañaga P (2007) A review of feature selection techniques in bioinformatics. *Bioinformatics* 23(19):2507–2517
- Sengupta M, Habte A, Wilbert S, Gueymard C, Remund J (2021) Best practices handbook for the collection and use of solar resource data for solar energy applications (No. NREL/TP-5D00-77635). National Renewable Energy Lab.(NREL), Golden, CO
- Shamshirband S, Mohammadi K, Chen HL, Samy GN, Petković D, Ma C (2015a) Daily global solar radiation prediction from air temperatures using kernel extreme learning machine: a case study for Iran. *J Atmos Solar Terr Phys* 134:109–117. <https://doi.org/10.1016/j.jastp.2015.09.014>
- Shamshirband S, Mohammadi K, Piri J, Petković D, Karim A (2016) Hybrid auto-regressive neural network model for estimating global solar radiation in Bandar Abbas, Iran. *Environ Earth Sci* 75(2):1–12. <https://doi.org/10.1007/s12665-015-4970-x>
- Shamshirband S, Mohammadi K, Yee PL, Petković D, Mostafaeipour A (2015b) A comparative evaluation for identifying the suitability of extreme learning machine to predict horizontal global solar radiation. *Renew Sustain Energy Rev* 52:1031–1042. <https://doi.org/10.1016/j.rser.2015.07.173>
- Sharma A, Kakkar A (2018) Forecasting daily global solar irradiance generation using machine learning. *Renew Sustain Energy Rev* 82(August 2017):2254–69. <https://doi.org/10.1016/j.rser.2017.08.066>
- Sharma V, Yang D, Walsh W, Reindl T (2016) Short term solar irradiance forecasting using a mixed wavelet neural network. *Renew Energy* 90:481–492. <https://doi.org/10.1016/j.renene.2016.01.020>
- Singla P, Duhan M, Saroha S (2022a) A comprehensive review and analysis of solar forecasting techniques. *Frontiers in Energy* 16(2):187–223. <https://doi.org/10.1007/s11708-021-0722-7>
- Singla P, Duhan M, Saroha S (2022b) An ensemble method to forecast 24-h ahead solar irradiance using wavelet decomposition and BiLSTM Deep Learning Network. *Earth Sci Inf* 15(1):291–306. <https://doi.org/10.1007/s12145-021-00723-1>
- Sun H, Zhao Na, Zeng X, Yan D (2015) Study of solar radiation prediction and modeling of relationships between solar radiation and meteorological variables. *Energy Convers Manage* 105:880–890. <https://doi.org/10.1016/j.enconman.2015.08.045>
- Tao H, Ewees AA, Al-Sulttani AO, Beyaztas U, Hameed MM, Salih SQ, Armanuos AM, Al-Ansari N, Voyant C, Shahid S, Yaseen ZM (2021) Global solar radiation prediction over North Dakota using air temperature: development of novel hybrid intelligence model. *Energy Rep* 7:136–157. <https://doi.org/10.1016/j.egy.2020.11.033>
- Voyant C, Notton G (2018) Solar irradiance nowcasting by stochastic persistence: a new parsimonious, simple and efficient forecasting tool. *Renew Sustain Energy Rev* 92:343–352. <https://doi.org/10.1016/j.rser.2018.04.116>
- Voyant C, Notton G, Kalogirou S, Nivet ML, Paoli C, Motte F, Fouilloy A (2017) Machine learning methods for solar radiation forecasting: a review. *Renew Energy* 105:569–582. <https://doi.org/10.1016/j.renene.2016.12.095>
- Wang F, Yili Yu, 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* (switzerland) 8(8):1–29. <https://doi.org/10.3390/app8081286>
- Wei CC (2017) Predictions of surface solar radiation on tilted solar panels using machine learning models: A case study of Tainan City, Taiwan. *Energies* 10(10). <https://doi.org/10.3390/en10101660>
- Wojtkiewicz J, Hosseini M, Gottumukkala R, Chambers TL (2019) Hour-Ahead solar irradiance forecasting using multivariate gated recurrent units. *Energies* 12(21):1–13. <https://doi.org/10.3390/en12214055>
- Xue X (2017) Prediction of daily diffuse solar radiation using artificial neural networks. *Int J Hydrogen Energy* 42(47):28214–28221. <https://doi.org/10.1016/j.ijhydene.2017.09.150>
- Yadav AK, Chandel SS (2014) Solar radiation prediction using artificial neural network techniques: a review. *Renew Sustain Energy Rev* 33:772–781. <https://doi.org/10.1016/j.rser.2013.08.055>
- Yadav AK, Malik H, Chandel SS (2015) Application of rapid miner in ANN based prediction of solar radiation for assessment of solar energy resource potential of 76 sites in Northwestern India. *Renew Sustain Energy Rev* 52:1093–1106. <https://doi.org/10.1016/j.rser.2015.07.156>
- Yahya AA, Osman A, Ramli AR, Balola A (2011) Feature selection for high dimensional data: an evolutionary filter approach. *J Comput Sci* 7(5):800–820. <https://doi.org/10.3844/jcssp.2011.800.820>
- Yang D, Ye Z, Lim LHI, Dong Z (2015) Very short term irradiance forecasting using the lasso. *Sol Energy* 114:314–326. <https://doi.org/10.1016/j.solener.2015.01.016>
- Yu Y, Cao J, Zhu J (2019) An LSTM short-term solar irradiance forecasting under complicated weather conditions. *IEEE Access* 7:145651–145666. <https://doi.org/10.1109/ACCESS.2019.2946057>
- Zang H, Liu L, Sun Li, Cheng L, Wei Z, Sun G (2020) Short-term global horizontal irradiance forecasting based on a hybrid CNN-LSTM model with spatiotemporal correlations. *Renew Energy* 160:26–41. <https://doi.org/10.1016/j.renene.2020.05.150>
- Zebari R, Abdulazeez A, Zeebaree D, Zebari D, Saeed J (2020) A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. *J Appl Sci Technol Trends* 1(2):56–70. <https://doi.org/10.38094/jastt1224>
- Zhang J, Zhao L, Deng S, Weicong X, Zhang Y (2017) A critical review of the models used to estimate solar radiation. *Renew Sustain Energy Rev* 70(July 2016):314–29. <https://doi.org/10.1016/j.rser.2016.11.124>
- Zhang L, Wen J (2019) A systematic feature selection procedure for short-term data-driven building energy forecasting model development. *Energy Build* 183:428–442. <https://doi.org/10.1016/j.enbuild.2018.11.010>
- Zhao X, Wei H, Wang H, Zhu T, Zhang K (2019) 3D-CNN-Based feature extraction of ground-based cloud images for direct normal irradiance prediction. *Solar Energy* 181(September 2018):510–18. <https://doi.org/10.1016/j.solener.2019.01.096>
- Zhou Y, Liu Y, Wang D, Liu X, Wang Y (2021) A review on global solar radiation prediction with machine learning models in a comprehensive perspective. *Energy Convers Manage* 235(13):113960. <https://doi.org/10.1016/j.enconman.2021.113960>
- Zou L, Wang L, Lin A, Zhu H, Peng Y, Zhao Z (2016) Estimation of global solar radiation using an artificial neural network based on an interpolation technique in Southeast China. *J Atmos Solar Terr Phys* 146:110–122. <https://doi.org/10.1016/j.jastp.2016.05.013>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.