Trends and gaps in photovoltaic power forecasting with machine learning

Alcañiz, A.; Grzebyk, D.; Ziar, H.; Isabella, O.

DOI
10.1016/j.egyr.2022.11.208

Publication date
2022

Document Version
Final published version

Published in
Energy Reports

Citation (APA)
Alcañiz, A., Grzebyk, D., Ziar, H., & Isabella, O. (2022). Trends and gaps in photovoltaic power forecasting with machine learning. *Energy Reports*, *9*, 447-471. https://doi.org/10.1016/j.egyr.2022.11.208

Energy Reports 9 (2023) 447–471

Contents lists available at ScienceDirect

Energy Reports

journal homepage: www.elsevier.com/locate/egyr

Review article

Trends and gaps in photovoltaic power forecasting with machine learning

Alba Alcañiz *, Daniel Grzebyk, Hesan Ziar, Olindo Isabella

Photovoltaic Materials and Devices Group, Delft University of Technology, Mekelweg 4, Delft, 2628 CD, The Netherlands

A R T I C L E I N F O

Article history:
Received 12 April 2022
Received in revised form 7 November 2022
Accepted 30 November 2022
Available online xxxx

Keywords:
Machine learning (ML)
Deep learning (DL)
PV power forecasting
Solar energy prediction

A B S T R A C T

The share of solar energy in the electricity mix increases year after year. Knowing the production of photovoltaic (PV) power at each instant of time is crucial for its integration into the grid. However, due to meteorological phenomena, PV power output can be uncertain and continuously varying, which complicates yield prediction. In recent years, machine learning (ML) techniques have entered the world of PV power forecasting to help increase the accuracy of predictions. Researchers have seen great potential in this approach, creating a vast literature on the topic. This paper intends to identify the most popular approaches and the gaps in this discipline. To do so, a representative part of the literature consisting of 100 publications is classified based on different aspects such as ML family, location of PV systems, number of systems considered, features, etc. Via this classification, the main trends and gaps can be highlighted while offering advice to researchers interested in the topic.

Contents

* Corresponding author.
E-mail address: A.AlcanizMoya@tudelft.nl (A. Alcañiz).

https://doi.org/10.1016/j.egyr.2022.11.208

1. Introduction

Photovoltaic (PV) energy has the potential to become a major source of electricity worldwide (International Energy Agency, 2021). This renewable energy is abundant, affordable, and easily scalable (Fthenakis et al., 2008), with the unique ability to cover most market segments from small household systems to utility-size power plants (International Energy Agency, 2020). In 2020, PV provided nearly 3.7% of the electricity demand in the world, a proportion that is rapidly increasing every year.

PV follows a rapid growth path, but there are challenges associated with it. Due to the variable nature of weather, the energy supplied by PV modules can be intermittent (Notton et al., 2018). This complicates market equilibrium, reserve capacity planning, and electricity market bidding. Large-scale PV penetration entails a major challenge for electric system operators since it hinders the effective management of the grids. Ineffective management may lead to voltage instability and increased volatility of the grid (Kawabe and Tanaka, 2015).

These issues can be solved with an accurate estimation of the power produced by PV sources (Shivashankar et al., 2016). Accurate forecasts decrease energy yield uncertainty, therefore reducing generation-load mismatch in the power grid, and are essential to ensure PV economic integration (Wan et al., 2016). Solar yield forecasting is an important factor facilitating energy transition and one of the key challenges for massive PV integration.

However, achieving a precise solar power prediction can be extremely difficult considering not only the chaotic nature of weather systems but also uncertainties related to the PV systems' components and location (Wan et al., 2016). The most mature methods rely on physical equations, which usually need detailed system information and are not able to model all uncertainties related to field works (Mayer and Gróf, 2020). Given these shortcomings, a new family of approaches is gaining popularity: machine learning algorithms.

In recent years, the number of manuscripts that use machine learning (ML) techniques for PV power prediction has increased exponentially, as depicted in Fig. 1. Considering this rapid development and the high amount of literature, it is hard to keep track of previous works performed and the recommendations to follow. For this reason, several review papers have already been published on the topic that sum up the progress done so far.

Table 1 summarizes the reviews done on PV power forecasting where ML techniques have been discussed. Most publications focused on PV power forecasting, although a few also included irradiance estimation methods. A brief explanation of the focus of these reviews is provided so the readers can refer to those references in case the topic is of their interest.

Since the first review in 2013, each review has focused on a different aspect of PV power forecasting, showing how ML methods have gained importance and providing a hint at how the main trends have evolved. However, what is lacking from our perspective is an extensive classification of the published literature to highlight what has and has not been done in this discipline. Overall, we want to distinguish from previous reviews in the following points:

1. This work has been limited to scientific articles published from 2015 until 2020. By limiting the time period, we can better identify the trends while covering a considerable percentage of the literature, instead of selecting the most cited papers in a wider period. The literature was selected first by choosing recent and most cited papers on the topic and then tracking back the relevant publications referenced in them.

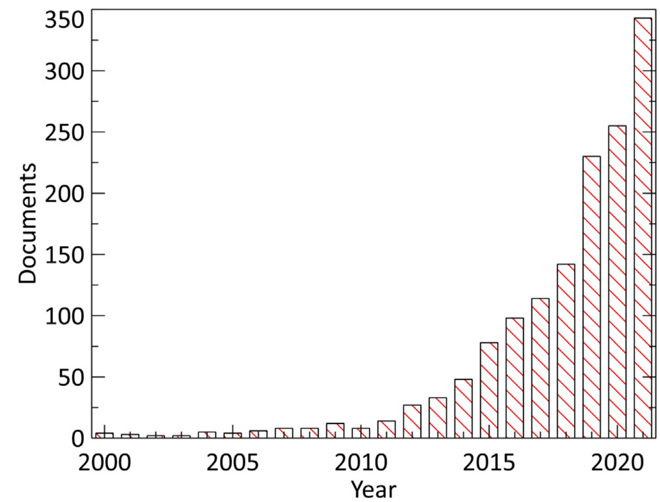


Fig. 1. Number of documents published on PV power forecasting using ML from 2000 until 2021 as of mid-2022. The data was obtained via an advanced literature search including ML techniques and photovoltaic forecasting in Clarivate Web of Science (Clarivate Analytics, 2022).

2. This paper focuses specifically on (1) machine learning employed for (2) PV power forecasting. There are other methods available to predict the PV power, which have been reviewed for instance by Raza et al. (2016), but the focus of this work is on machine learning only. In addition, the considered literature employs only direct PV power forecasting, i.e., the objective is to forecast the PV power produced, not the irradiance (Massaoudi et al., 2021).
3. The 100 reviewed publications on PV power forecasting are classified into: ML family, location, climate, number of systems, timespan, forecast horizon and choice of features. To the best of our knowledge, no previous review has performed such an extensive categorization, neither has classified by climate nor ML family.
4. A section is devoted to discussing special methodologies employed for PV power forecasting, such as online forecasting or global optimization. The intention is to introduce these alternatives that are gaining strength, include some recent examples and refer the reader to corresponding reviews.
5. Finally, we provide recommendations based on the extensive analysis performed. The objective is to be a reference for new researchers that want to explore the world of PV power forecasting, and a sea of ideas for future developments for familiarized researchers.

The structure of the review is as follows. Section 2 focuses on the ML algorithms by classifying them into families. Section 3 explains the highlighted special approaches from the literature. Sections 4, 5 and 6 continue the classification based on system characteristics, forecast horizon and employed features, respectively. Recommendations are given in Section 7, before concluding in Section 8. Overall, we aim to cover all the choices that a researcher needs to make when developing a PV power forecast model. An overview of them can be seen in Fig. 2.

2. Machine learning families

This work starts by classifying the 100 reviewed publications into ML families. Appendix A: Reviewed literature provides a summary of the main characteristics of all documents. Previous reviews mention only the most relevant algorithms, sometimes

Table 1
Previous reviews on PV power forecasting using ML techniques.

Ref	Topic
Inman et al. (2013)	Theoretical basis for most promising methods for solar forecasting. Large focus on irradiance prediction
Ulbricht et al. (2013)	Classification of solar forecasting solutions
Ren et al. (2015)	Wind and solar forecasting with ensemble methods
Antonanzas et al. (2016)	ML and physical methods have the same importance. Focus on economics in forecasting and the difference between point and regional forecasts
Gandoman et al. (2016)	PV power forecasting under cloudy weather
Raza et al. (2016)	Special emphasis on ML algorithms such as ANN (Artificial Neural Network)
Barbieri et al. (2017)	Very short-term PV power forecasting with cloud modelling
Das et al. (2018)	Short-term direct PV power forecasting
Sobri et al. (2018)	Most popular methods for PV power forecasting: advantages and limitations
van der Meer et al. (2018)	Solar power and load probabilistic forecasting
Mosavi et al. (2019)	ML applied to energy systems
Yang (2019)	Provide a set of reporting rules called ROPES to assist scientists in the evaluation of their methods and set guidelines for future research
Ahmed et al. (2020)	Cloud imaging and online PV power forecasting. Methods for pre- and post-processing of data
Massaoudi et al. (2021)	PV power forecasting using deep learning techniques
Mellit et al. (2020)	Focus only on ML techniques
Pazikadin et al. (2020)	ANNs for solar power generation forecasting
Rajagukguk et al. (2020)	Deep learning algorithms for solar irradiance and power forecasting
Feng et al. (2021)	Taxonomical review on the integration of PV using artificial intelligence
Gupta and Singh (2021)	Direct and indirect ML-based PV power forecasting

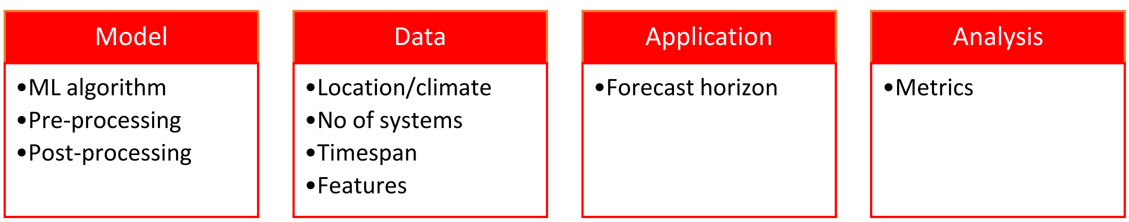


Fig. 2. Choices that a researcher needs to make when developing an ML-based PV power forecast model. Note that post-processing is a topic not extensively covered in this review.

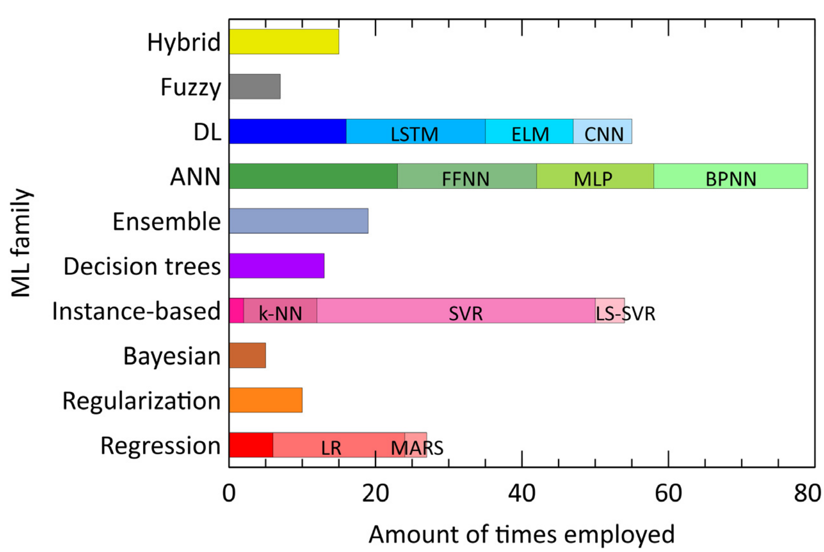


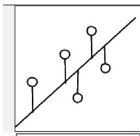
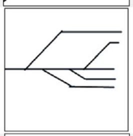
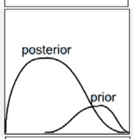
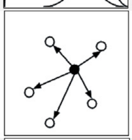
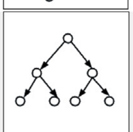
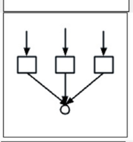
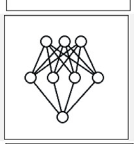
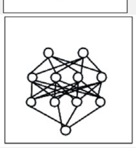
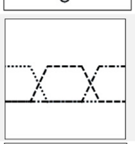
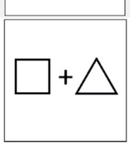
Fig. 3. Amount of times that an ML algorithm from a certain family has been employed in the literature. Bar areas without a specific algorithm indicate algorithms from the ML family different from the most popular ones within it. For colour references refer to the web version of this article.

disregarding promising but not so extensively used approaches. To avoid that, we have grouped the algorithms by their similarity in form or function, following the classification performed by Brownlee (2021). There may be some differences with respect to other works, since there are several classification criteria, and some algorithms belong to more than one category. The distinctive characteristics of each group are explained in Appendix B: ML families, while Table 2 offers a graphical summary of it.

Fig. 3 shows the number of times that an ML algorithm from a certain family has been employed. In this bar graph, the most popular algorithms and their occurrence have been indicated within their corresponding family.

As one can see, and in accordance with literature (Raza et al., 2016), Artificial Neural Networks (ANN) are the most popular family of algorithms in PV power forecasting problems, with Feed-Forward Neural Networks (FFNN), Multi-Layer Perceptron (MLP), and Back-Propagation Neural Network (BPNN) being the

Table 2
Machine learning families and their main characteristics.
Source: Figures from Brownlee (2021).

	Family	MAIN CHARACTERISTICS
	Regression	Starting point in machine learning Composed of linear regression and its improvements
	Regularization	Improvement of regression methods It prevents that the coefficients reach very high values, focusing too much on training data (overfitting)
	Bayesian	Based on Bayes' theorem: the probability distribution of the output depends on that of the input
	Instance-based	Use similarity between the data: build a database and compare new data to it to make the prediction
	Decision tree	Tree-like model of decisions: data space is partitioned, and a prediction model is fitted within each partition
	Ensemble	Combine weak algorithms (e.g. decision trees) into a strong one
	Artificial neural networks	Set of nodes (neurons) interconnected via weights, such that signals can travel through them Very flexible and able to model nonlinear systems
	Deep Learning (DL)	ANN with complex relations between neurons increasing the problem-solving capacity Characterized by having more than one hidden layer
	Fuzzy	Algorithms where binary logic has been substituted by a fuzzy one Higher flexibility than the original ones
	Hybrid	Combine algorithms from two different groups Increased capacity for problem-solving

most popular algorithms within it. The popularity of ANNs lies in their structure which makes them very flexible. Almost tied in second place, one can find the instance-based and Deep Learning (DL) families. From the former family, Support Vector Regression (SVR) is the proposed approach from several groups given its good performance (Abuella and Chowdhury, 2017; Nageem and Jayabarathi, 2017; Zhou et al., 2018; Alfadda et al., 2017), making it the most popular algorithm amongst the reviewed publications. Within DL, Long–Short Term Memory (LSTM) was applied 17 times in the reviewed literature, achieving excellent results in general. Other DL algorithms whose popularity is increasing considerably in the last few years are Extreme Learning Machines (ELM) and Convolutional Neural Networks (CNN). The latter shows excellent performance (Huang and Kuo, 2019) and is the recommended option when dealing with image data.

In the next Section 2.1, we will dive deeper into each family in order to explain the advantages and drawbacks that can explain the reported trends from the previous figure. These arguments will be supported in Section 2.2, where the findings of selected literature that compare several algorithms are explained. The last Section 2.3 addresses a different application of ML algorithms, which consists of performing pre-processing on the data using ML techniques.

2.1. Key features and trends

This section provides a brief overview of each family by explaining the advantages and drawbacks. Moreover, based on the impressions by the authors after the extensive literature review

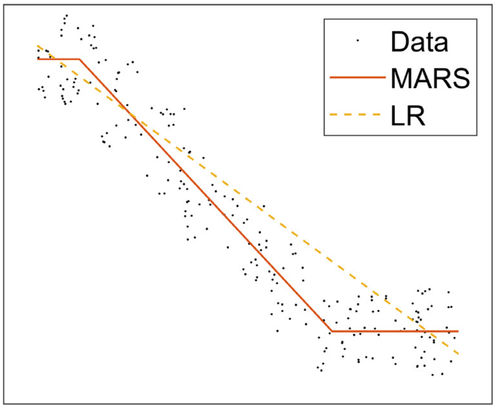


Fig. 4. Graphical difference between MARS and LR to depict the higher flexibility of the former.

performed, we provide our opinions on how will the popularity trends evolve with time.

Starting with the first family, linear regression (LR) is the most common algorithm in the regression family. Although it does not show the best performance (Li et al., 2016a; Mishra et al., 2020; Semero et al., 2018; Ferlito et al., 2017; Lee et al., 2018), it is valued by its simple implementation and ease of interpretation (Touati et al., 2017; Do et al., 2016). It is sometimes employed as a benchmark (Touati et al., 2017; Ramsami and Oree, 2015; Kuzmiakova et al., 2017; Rosato et al., 2017). However, none of the groups proposed it as an option to forecast the PV power. The main limitation of this algorithm lies in the fact that it builds a hyperplane to make the predictions, while this is a too strong reduction of the problem of PV power forecasting.

A recently applied regression algorithm that has shown excellent results is Multivariate Adaptive Regression Splines (MARS). This algorithm improves the simple linear regression by including hinge functions, which allow for more flexibility to fit the data while keeping the interpretability (Friedman, 1991), as seen graphically in Fig. 4. MARS is the algorithm proposed by some groups (Li et al., 2016a; Ferlito et al., 2017; Massidda and Marrocu, 2017), and we think that, because of its interpretability, it will play an important role in the future. Currently, interpretability may only seem relevant when discussing high-stakes predictions that deeply impact human lives (Gosiewska et al., 2021), such as healthcare and criminal justice. However, the negative effects of lack of interpretability have already reached environmental aspects by stating that highly polluted air was safe to breathe (Rudin, 2019). A similar situation may arise in the future after an incorrect ML prediction in a grid with high renewable energy penetration.

Regularization algorithms aim at solving the overfitting problem of regression ones. Overfitting occurs when the hyperplane is too specific and focuses only on the training data, so it is unable to generalize to unseen data. Regularization solves this issue by making sure that the hyperplane stays "simple enough" to be able to generalize. The problem with regularization algorithms is that they still fit a hyperplane, which is unable to explain the complex relationship between meteorological conditions and PV power output. Moreover, overfitting of LR is an issue when the amount of training data is not large enough, but since we are in the era of big data, overfitting of LR is not present anymore. This explains the lack of popularity of regularization algorithms. This family was probably a good actor a decade ago, but it has been displaced by more powerful algorithms.

A similar situation occurs with Bayesian algorithms. Considering the origins of ML, the idea of fitting a probabilistic distribution function to the data seemed the way forward. This approach works well on small datasets and it can provide uncertainty measurements (Sit, 2019). However, some of the probabilistic distributions assume independency between features, which is unrealistic in real life.

The two main members in the instance-based family are k-NN and SVR. Least Squares-SVR (LS-SVR) is just a simplification in the learning procedure of the SVR which makes the solving process more efficient without any effect on accuracy.

The popularity of k-NN lies in its ease of use and flexibility to adapt to the data. It is the most popular classification algorithm and identifies clusters of data based on the nearest neighbours. For forecast applications, the algorithm identifies in the training dataset the cluster that resembles the most to each test condition and creates a forecast value based on that. An identified disadvantage from this explanation is the need to store the whole training set, and that the prediction times may be longer than for other algorithms.

In the case of SVR (and LS-SVR), there are facing opinions in the literature. Some defend it as the algorithm of choice (Abuella and Chowdhury, 2017; De Felice et al., 2015; Rana et al., 2015; Pan et al., 2020), while others report similar or worse results with respect to other algorithms (Huang and Kuo, 2019; Li et al., 2016a; Wang et al., 2018; Sheng et al., 2020; Carrera and Kim, 2020). The reason for this disagreement, according to the authors' opinion, is the difficulty of optimization of SVR. SVR has a wide range of parameters, and the performance of the algorithm is highly sensitive to them (Pawar et al., 2020), so results are jeopardized if the optimum set is not found. Moreover, like k-NN, SVR also relies on similarity to the training samples, so it is not suitable for large datasets (Wang et al., 2018), which may be a limitation in the future. However, with the help of LS-SVR and the popularity seen so far, we think that SVR will be around for some more years.

The working principle of decision trees is promising due to its flexibility and interpretability, but it presents a huge disadvantage: overfitting. While overfitting solutions did not improve the main limitation of regression algorithms, there is an optimum amendment for decision trees: ensemble algorithms. Therefore, in our opinion decision trees will be every time more irrelevant, and they will only be indirectly used through ensemble algorithms.

We think that ensemble algorithms, together with SVR and MARS, will be the only algorithms able to compete with the hegemony of ANNs and DL that is currently present. Ensemble algorithms come at a loss of interpretability with respect to decision trees, but, as opposed to SVR, the difference in performance when the parameters are optimized or not is not that big. This family is a high appeal for new researchers wanting to use ML algorithms.

ANNs are the main protagonist in ML. They are very powerful and able to model any non-linearity relation (Cross et al., 1995). However, they have the same implementation drawback as SVR: the optimal configuration and learning algorithm are hard to obtain (Li et al., 2016a). This can explain some of the lower performance achieved by some researchers with respect to other algorithms (Li et al., 2016a; Sharadga et al., 2020; Zhang et al., 2018; Hossain et al., 2017).

In the explored literature, most proposed algorithms belong to the deep learning family. This is not shown in the graph, since there we included all the algorithms applied in each publication. However, a high percentage of the publications focus on and propose an algorithm of the DL family. These multi-layer ANNs have the benefit of including the impact of historical trends, which increases the accuracy of the predictions. The ANNs' drawback of optimal configuration is still present, and interpretability is challenging. However, people are exploring solutions to these

Table 3
Hybrid algorithms employed in the literature.

Combination	Objective	Ref.
LSTM + CNN	Combine the temporal ability of LSTM with the feature extraction abilities of CNN	Lee et al. (2018), Suresh et al. (2020), Wang et al. (2019b), Lee et al. (2018), Lin et al. (2020) and Li et al. (2016b)
DL + LSTM/DL	Combine the feature learning of the DL algorithm (an AutoEncoder) with the forecasting ability of LSTM/DL algorithm	Alkandari and Ahmad (2020)
DL + LSTM CNN + SVM	Combine the feature learning of a DL (AutoEncoder) with the temporal ability of LSTM Combine the feature extraction ability of CNN with the forecasting ability of SVM	Gensler et al. (2017) Zang et al. (2018)
DL + LSTM	Increase the accuracy of the forecast	Wen et al. (2019) and Jung et al. (2020)
LS-SVM + ANN Ensemble + SVM	Increase the accuracy of the forecast Combine several SVM to increase the accuracy	De Giorgi et al. (2016) Zhou et al. (2018)
Ensemble + ANN	Create a probabilistic forecast	Cervone et al. (2017)
Ensemble of LSTM, SVR, ANN, LR, decision tree	Increase the accuracy of the forecast	Chen and Koprinska (2020)
Ensemble of ANN, SVR, ELM	Increase the accuracy of the forecast	Nayak and Heistrene (2020)
Ensemble of instance-based, ANN	Increase the accuracy of the forecast	Lin et al. (2020)

issues, and we think that the popularity of this family has only begun and that it will gain more ground in the coming years.

Two algorithms use fuzzy logic in the explored literature: fuzzy k-means clustering and Adaptive Neuro Fuzzy Inference System. In the first, fuzzy logic is integrated into k-NN, while in the second, it is integrated with ANN. Some researchers may not have made a different group for these algorithms, but we wanted to highlight them. We think that the flexibility introduced by fuzzy logic can be very powerful, but it may not be advantageous enough. These algorithms are harder to interpret, therefore they compete with DL approaches, which are more popular and with already demonstrated accuracy. Hence we think that, despite being promising, fuzzy algorithms will not be relevant in the future.

The final step in complexity can be taken with hybrid algorithms. Most of them focus on increasing the accuracy of the forecast by combining the abilities of two or more ML algorithms. The main drawback is the increase in complexity and consequent decrease in interpretability. Table 3 gives an overview of the combination structures from the reported literature. One can highlight how ensembles are being applied to strong algorithms, such as ANN, to overcome overfitting and increase the accuracy. More examples of hybrid DL methods for PV power forecasting can be found in the review by Sobri et al. (2018) and Massaoudi et al. (2021). We think that the development of this family will continue as researchers mix and match the ML algorithms as they deem necessary to apply an improvement or overcome a drawback.

2.2. Algorithms comparison

To compare the performance of different ML algorithms, they need to be fairly analysed. The prediction results not only depend on the algorithm employed, but also on the amount of data, the solving method, the location, the metric, etc. Therefore, in this subsection, we are going to present the most relevant studies which have compared the performance of several ML models under the same conditions. Some metrics will be needed for this task, which are defined in Appendix C: Metrics.

We would like to highlight three papers that compare a wide range of algorithms belonging to different ML families. Ferlito et al. forecasted the power of a 1 kW grid-connected PV system with 11 models of different complexity and belonging to the regression, instance-based, decision trees, neural networks, and ensemble families (Ferlito et al., 2017). They found that the best-performing algorithm and optimum dataset length depended on the training and testing years. For example, the Cubist algorithm (ensemble) showed the best performance in a highly variable year such as 2010 with Root Mean Squared Error (RMSE) of around 93 W, while Random Forest (ensemble) had lower prediction errors than other methods in the year 2011. They reported that non-linear models showed up to 2% superior performance with respect to linear ones. Chen et al. implemented 7 base learners from different families and combined their outputs through several ensemble strategies (Chen and Koprinska, 2020). One of the ensembles was identified as the best model, while LSTM was the most accurate base learner, showing competitive results. Finally, Carrera et al. trained 15 ML algorithms to predict the power of a South Korean solar farm (Carrera and Kim, 2020). One can observe the reported RMSE for all algorithms in Fig. 5. In the graph, the red lines represent the medians while the green triangles indicate the means. They reported that ensemble algorithms were the best performing ones, except in terms of bias when Elastic Net (regularization family) outperformed the rest.

The study performed by Sharadga et al. compares different ANN models (Sharadga et al., 2020). The ranking between models depended not only on the type of ANN but also on the training algorithm, on the number of hours ahead forecasting and on the metric. However, recurrent neural networks generally outperformed feed-forward ones. Li et al. reported that the algorithm with the lowest RMSE was MARS (Li et al., 2016a). Although the ranking of these models differed depending on the metric employed, MARS was in the top three position independently on the metric employed and outperformed more complex models such as SVR or ANN in certain metrics.

In Lin et al. (2018), several ANN, SVR and their hybridized forms were compared. In general, the hybrid models outperformed their simpler versions. Without hybridization, RMSE and Mean Absolute Percentage Error (MAPE) showed that BPNN was better than other ANNs. When hybridized, the rank of these models was clearer in both metrics: hybrid Radial Basis Function ANN was the best followed by hybrid BPNN, hybrid Elman and hybrid SVR.

Massaoudi et al. proposed a novel ensemble method and compared it to other algorithms (Massaoudi et al., 2019). They reported that the proposed method showed the best performance both in terms of RMSE and MAE. Finally, in Lin and Pai (2016),

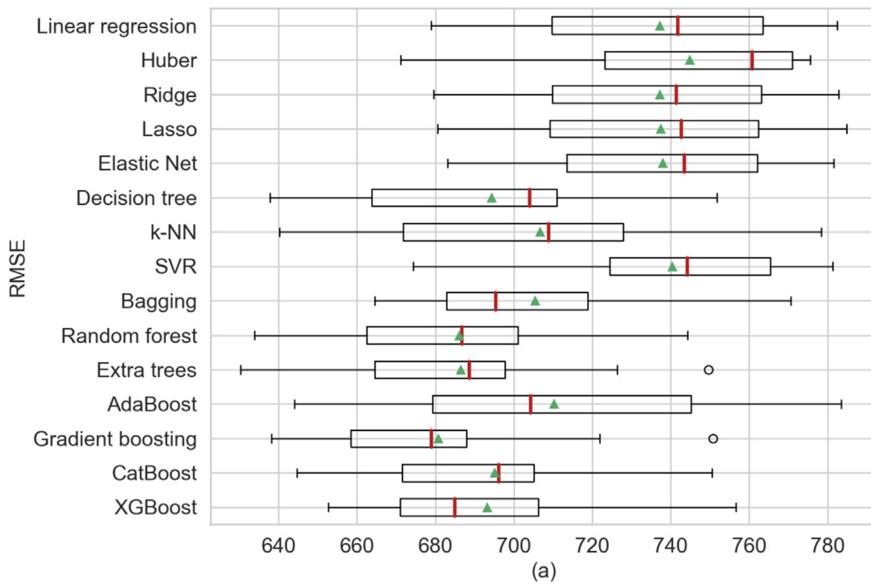


Fig. 5. Forecasting accuracy of several ML methods using 10-fold cross-validation in terms of RMSE.
Source: Figure from Carrera and Kim (2020).

they ranked the models in terms of both RMSE and MAPE. The proposed modified Least-Squares SVR (LS-SVR) model showed the best performance, followed by ANNs.

All these studies show that it is difficult to recommend an algorithm, even a family of algorithms, that is superior. However, it would be useful to create a set of guidelines for researchers to help with decision-making. The personal recommendation of the authors for new researchers is ensemble algorithms, because of their accuracy/interpretation ratio. They are also a good option when dealing with a limited amount of data. If the amount of data is high enough and the researcher is experienced, algorithms such as SVR, LSTM and ELM are recommended. However, it is important to be careful with the optimization of the parameters.

2.3. Data preparation

This subsection describes ML-based pre-processing strategies that reduce the computational cost or complexity of the problem without decreasing the accuracy. Pre-processing can also be achieved with non-ML-based methods, but here we focus on ML-based ones. The focus is also on modifications of the input data once cleaning and normalization are achieved.

It is important to keep in mind that pre-processing requires human resources. The trade-off between the increase in accuracy and the time employed needs to be considered before diving into these strategies. However, these resources are most needed during the training process, so we consider them worthy of their benefits when making predictions. We think that the trend will be to keep applying data-processing algorithms, however, it is hard to tell whether these will be ML-based ones or not. It most likely depends on the pre-processing step. For instance, k-NN is dominant for clustering procedures, but there are several approaches for feature selection.

Post-processing techniques such as reforecasting (Chu et al., 2015) or corrections (Yin et al., 2020) can also be employed to improve the performance. However, due to its limited use in the explored literature, it was not possible to review them, which can be considered a gap that still needs to be explored.

The most common pre-processing step is to reduce the number of features employed, to keep only the relevant ones. These methods are referred to as feature selection and are explained in Section 2.3.1. An alternative way to reduce the input data is to group by similarity, explained in Section 2.3.2. A third strategy is to decompose the input data, explained in Section 2.3.3.

2.3.1. Feature selection

Feature selection or dimensionality reduction consists of selecting the most relevant features for an ML model in order to decrease the dimensionality of the problem (Guyon and Elisseeff, 2003). By doing this, the performance of the model is increased, the computational cost is reduced, and a better understanding of the underlying process is gained. There are several methods available which perform dimensionality reduction. It can even be done by manual inspection (Rosiek et al., 2018; Almeida et al., 2015; Baharin et al., 2016), or with correlation-based strategies (Abuella and Chowdhury, 2017; Wang et al., 2018; Van Tai, 2019; Han et al., 2019). However, it is more effective to employ more complex strategies.

One of the most famous methods for dimensionality reduction is Principal Component Analysis, which was applied in Kuzmiakova et al. (2017) and Pierro et al. (2017). This technique reduces the dimension of the dataset by creating linear combinations of the features that have maximal variance and are mutually uncorrelated. An alternative is to employ ML models used for prediction such as Elastic Net (regularization family) (Massaoudi et al., 2019), Gradient Boost Regression Trees (ensemble) (Massaoudi et al., 2019; Isaksson and Conde, 2018), or SVR (Lee et al., 2019) to perform feature selection. These methods exploit the fact that some ML models assign a weight to each feature, thus its importance is known. Features with weights smaller than a certain threshold are removed for feature selection. The threshold is chosen by the user and depends on the algorithm. For instance, in each split of a gradient boost regression tree, the chosen feature to split on is the one that maximizes the reduction of a certain kind of error (Ratanamahatana and Gunopulos, 2010) (refer to Appendix B: ML families for details on how ensemble algorithms work). By the end of the tree, the divisions are less relevant. Thus, by cutting trees below a certain split (the threshold), one can identify the most important features.

Even though these techniques may not seem relevant at first, they can help boost the performance considerably. In Shang and Wei (2018), feature selection reduced the RMSE of the proposed

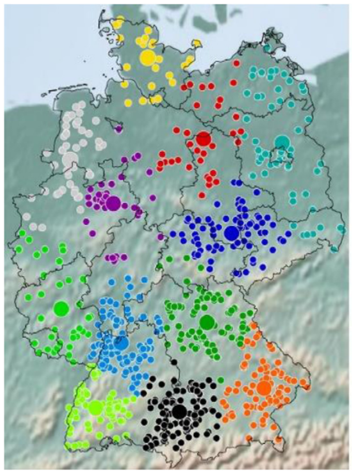


Fig. 6. k-NN applied to several PV systems spread over Germany. Each colour represents a different cluster and small dots represent the individual PV plants. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Source: Figure from Wolff et al. (2016).

model from an average of 5.7% to 4.6%. Similarly, in Isaksson and Conde (2018) the applied feature selection method increased the performance of k-Nearest Neighbour (k-NN, instance-based family) by 2%. Alfadda et al. showed how the error evolves as features are incrementally included in the model (Alfadda et al., 2017). They reported that when all features were considered, the RMSE was higher than when a few were left out. Lee et al. observed a decrease in performance when considering redundant variables in their forecast model (Lee et al., 2019). They reported an average hourly MAPE of 68.2%, compared with a 107.1% MAPE when considering all features.

Overall, the reader is strongly advised to employ a feature selection technique when dealing with a high number of input parameters. Apart from the already mentioned benefits, feature selection can highlight features with high predictive power that are superfluous analytically. However, one also has to be critical of the results, since the selected features could denote errors in the input data. It is recommended to make use of solar engineering knowledge and previous literature to validate the results obtained (refer to Section 6 for the classification of features). In case an unexpected feature is selected, one should explore the input data and make an informed decision.

2.3.2. Clustering

When the amount of input data is large, the algorithms become computationally expensive. This data is often repeated and not all of it is needed. For instance, similar weather conditions yield similar PV power output. The data can be filtered with clustering algorithms such as k-NN and grouped by meteorological characteristics or PV systems. Since cloudy days are harder to forecast than sunny ones, some researchers have developed a different model per type of weather (Baharin et al., 2016; Wang et al., 2020; Li et al., 2015) or season (Yin et al., 2020). Spatial clustering of PV plants is commonly applied to forecast distributed generation (Pierro et al., 2017; Wolff et al., 2016), see Fig. 6.

Some authors integrated clustering methods into their algorithms and considered them a hybrid. Despite having employed a different criterion here, so these algorithms are not considered a hybrid, it is still worth explaining them. Lin et al. developed an algorithm that clustered the data depending on the weather conditions, found similarity of the input data to each of these groups, and used the information on similarity to help the forecasting algorithm with the prediction (Lin et al., 2018). The methodology developed by Wang et al. (2018) clusters the data into groups, trains a forecasting model for each group, and creates an average between the forecasts depending on the similarity to each group.

2.3.3. Time-series decomposition

Another strategy to simplify the learning process of ML algorithms is to decompose the input signal into several frequency series so that each series has better outlines and behaviours. The most commonly employed algorithm to achieve this is Wavelet Decomposition, applied in the works (Mishra et al., 2020; Wang et al., 2017; Raza et al., 2018; Li et al., 2020; Eseye et al., 2018; Soufiane et al., 2020). In Shang and Wei (2018), the effect of decomposing the time series was a reduction of the RMSE from 5.9% to 4.6%. More information on this topic can be found in the review of Ahmed et al. (2020).

3. Special approaches

This section introduces variations that several groups have made to the standard forecasting methodology. We focus on these novel methodologies since they are gaining popularity, and we discuss the advantages and possible difficulties that each one may present.

Probabilistic forecasts

A way to introduce flexibility to the prediction method is to provide a probabilistic forecast. This method is not novel, but it is still worth mentioning due to its ability to help grid operators when managing the grid and considering the energy reserve.

There are three main methods to generate a probabilistic forecast: ensemble algorithms, uncertainty estimation, and quantile regression (Panamtash et al., 2021). Ensemble algorithms, which generate probabilistic intervals intrinsically, are used in Cervone et al. (2017), Panamtash et al. (2021), Alessandrini et al. (2015), Ni et al. (2017) and Takeda (2017). They are generally computationally intensive, affecting their practical implementation. The second method modifies point forecasting algorithms to provide an estimation of the error distribution, using parametric (Pierro et al., 2017; Panamtash et al., 2021) or non-parametric methods (Rana et al., 2015; Almeida et al., 2015). With this methodology, prediction intervals are usually estimated via a normal distribution, which is not reliable for all values (Pierro et al., 2017). Finally, examples of probabilistic quantile regression forecasts can be found in Panamtash et al. (2021) and Alessandrini et al. (2015). Quantile regression is a regression algorithm where each quantile is separately modelled and a different prediction is fitted within each partition. Despite its extended use, this method does not report as good results as the other two (Panamtash et al., 2021).

Amongst these studies, we want to highlight the work performed by Rana et al. (2015) where the interval forecasts are in 2D, providing not only a range of expected values in power but also in time. A graphical explanation of the difference between point (non-probabilistic), interval (probabilistic in power) and 2D-interval (probabilistic in power and time) forecasts is shown in Fig. 7. This method can provide even more flexibility to the forecast, which can be especially useful for risk management applications.

More information on PV power probabilistic forecasting can be found in the reviews (van der Meer et al., 2018; Panamtash et al., 2021).

Meta-heuristic algorithms

As previously mentioned, some algorithms; such as SVR or ANN, are hard to optimize and that can have negative effects. To overcome this disadvantage, several groups have improved

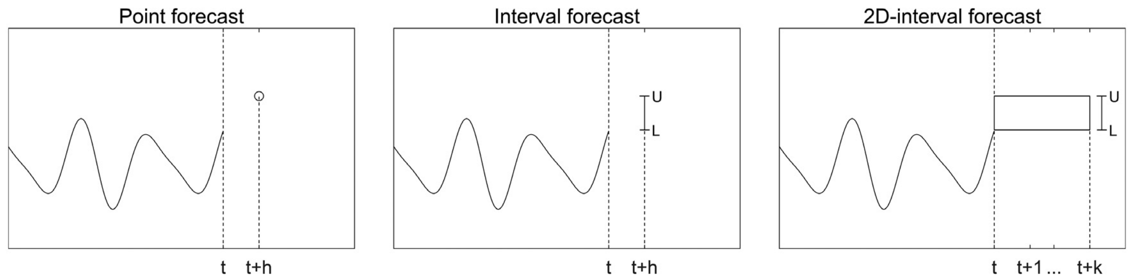


Fig. 7. Graphical explanation of the difference between point, interval and 2-D forecasts.
Source: Figure redrawn from Rana et al. (2015).

the performance of an ML algorithm by using optimizers to find the optimum set of parameters. Global optimizers are required for the process to cover all possibilities. Amongst these techniques, we want to highlight the meta-heuristic algorithms. These are ML global optimization algorithms, suitable for non-linear parameters.

In the reviewed literature, several groups have employed meta-heuristic algorithms for optimization in PV power forecasting problems, such as Shuffled Frog Leaping Algorithm (Asrari et al., 2017), Ant Colony Optimization (Pan et al., 2020) or Multiverse Optimization (Li et al., 2019b). The most popular optimization algorithms are Genetic Algorithm (GA), inspired by the theory of evolution (Semero et al., 2018; Lin and Pai, 2016; Chu et al., 2015; Eseye et al., 2018; Asrari et al., 2017); and Particle Swarm Optimization (PSO), inspired by bird flocks' motion (Semero et al., 2018; Shang and Wei, 2018; Eseye et al., 2018; Li et al., 2019b).

The use of these algorithms can considerably improve the performance. Eseye et al. (2018) reduced the Normalized Mean Absolute Error (NMAE) of BPNN from 1.3% to 1.2% using GA and to 1.0% using PSO, and the NMAE of SVR from 0.8% to 0.7% with GA and to 0.5% with PSO. This increase in performance comes again at a cost of simplicity, but we think that global optimizers are required if the objective is to keep decreasing the errors and employing complex algorithms.

More information on this topic can be found in the review of Ahmed et al. (2020).

Online forecasting

One last forecasting approach worth mentioning is online forecasting. In online training mode, the algorithm updates the weights for each training phase for each batch of results (Theodoridis and Koutroumbas, 2009). The main advantages are that the training dataset required is minimum and that the algorithm can rapidly adapt to major changes in the environment. Its main drawback is the continuous need for training.

This method was applied in Ferlito et al. (2017) with several ML algorithms. The results showed that the RMSE of offline training was higher than that of online, and that the ranking of the algorithms depended on the training method. Al-Dahidi et al. (2020) also developed a probabilistic very short-term model which was ready to be used for online or real-time applications. In Sheng et al. (2020), an online adaptive learning framework was implemented so that the model could adjust to the complex dynamics of the weather. The model could selectively accumulate or forget knowledge to respond to climate changes. Similarly, the models proposed in Golestaneh et al. (2016) and Agoua et al. (2019) could be used for online predictions, which can be especially suitable for Internet-of-Things systems (Kraemer et al., 2020).

More information on online PV power forecasting can be found in the review by Ahmed et al. (2020).

4. System characteristics

This section focuses on the main characteristics of the PV systems employed, namely location, number of systems and timespan.

4.1. Location

Fig. 8 represents the number of publications whose studied systems are located in each country. Most of them are in Europe and no PV system has been forecasted in South America. The country in which more systems are located is the USA (13 studies) closely followed by Australia (12), Italy (10) and China (9). Considering the intensity of solar energy across the globe, areas located at latitudes 45 degrees from the equator have a tremendous opportunity for harnessing solar energy (Ahmed et al., 2020). Regions such as the Middle East, most of Australia and the deserts are suitable for large-scale PV installations. However, most of the reviewed literature focuses on areas with relatively small solar intensity (except for Australia), not on these high-potential zones.

Characterizing the PV systems' location by country is not representative of the weather conditions that they are subjected to. Therefore, we have classified the reviewed papers by type of climate, as reported in Fig. 9. The Köppen climate classification has been employed for this purpose (Chen and Chen, 2013) since it divides the globe with a single metric into five categories: tropical, dry, mild temperate, continental and polar.

Around half of the forecasted systems from the reviewed publications are in mild temperate climates. Roughly 25% of the systems are in dry climates, while only 8 are located in tropical and 10 in snow climates. No system is in polar climate since mostly the Arctic, Antarctica and tops of mountains are representative of this climate.

Some studies have been left out from one or both figures since they consider more than one system under different climates or countries. Do et al. (2016) forecasted two French PV systems, one in Guadeloupe (geographically in North America, tropical climate) and the other in Lille (Europe, mild temperate climate). They found that due to the seasonal variation of the mild temperate climate, a longer training duration for the system in Lille is needed than for the one in Guadeloupe (6 months instead of 3). However, once the model is trained, the algorithm shows higher performance in the temperate climate than in the tropical one. Zhang et al. (2015) forecasted the PV power of three distant PV systems: San Diego, USA; Braedstrup, Denmark; and Catania, Italy. Although the three systems are in a mild temperate climate, they presented diverse daily weather distribution. These influenced the modelling parameters and features that yielded the best forecasting result. Golestaneh et al. (2016) trained ELM in two PV systems located in Australia (mild temperate) and Singapore (tropical). They found that the climate conditions of each

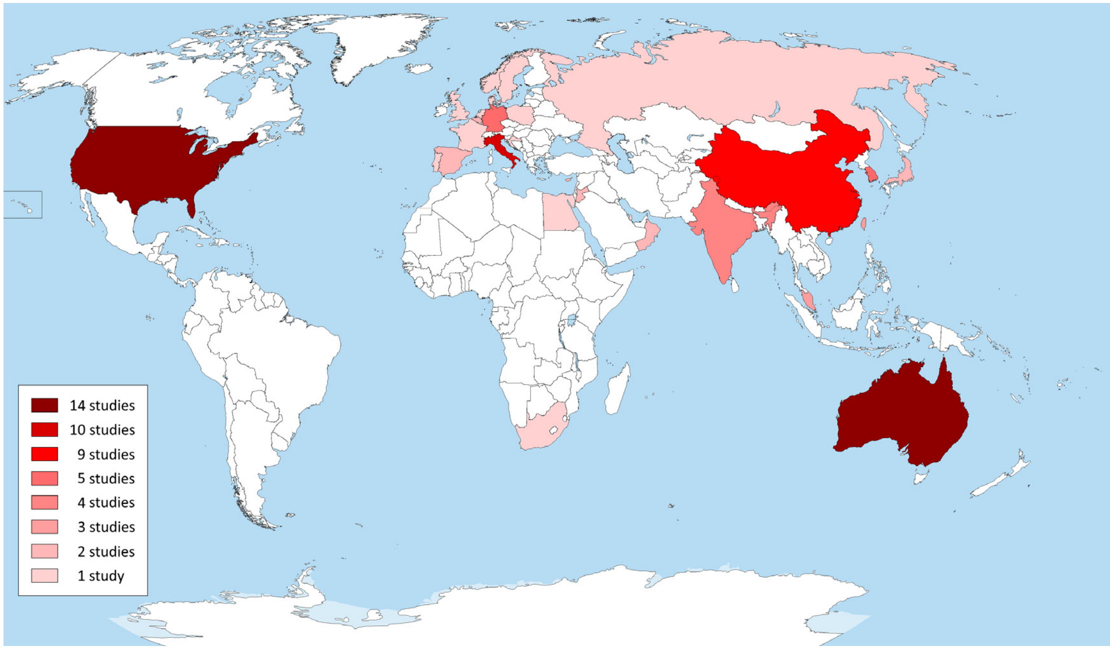


Fig. 8. World map with the countries highlighted depending on the number of studies published on the topic. Only the mainland has been coloured. Since some countries' contribution might not be visible on the map, the raw data is; 14 studies: Australia, USA; 10 studies: Italy; 9 studies: China; 5 studies: Germany, South Korea; 4 studies: India; 3 studies: Malaysia, Taiwan; 2 studies: Belgium, Cyprus, Japan, Jordan, Oman, Spain; 1 study: Croatia, Denmark, Egypt, France, Netherlands, Norway, Poland, Portugal, Qatar, Russia, Singapore, South Africa, Sweden, UK. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

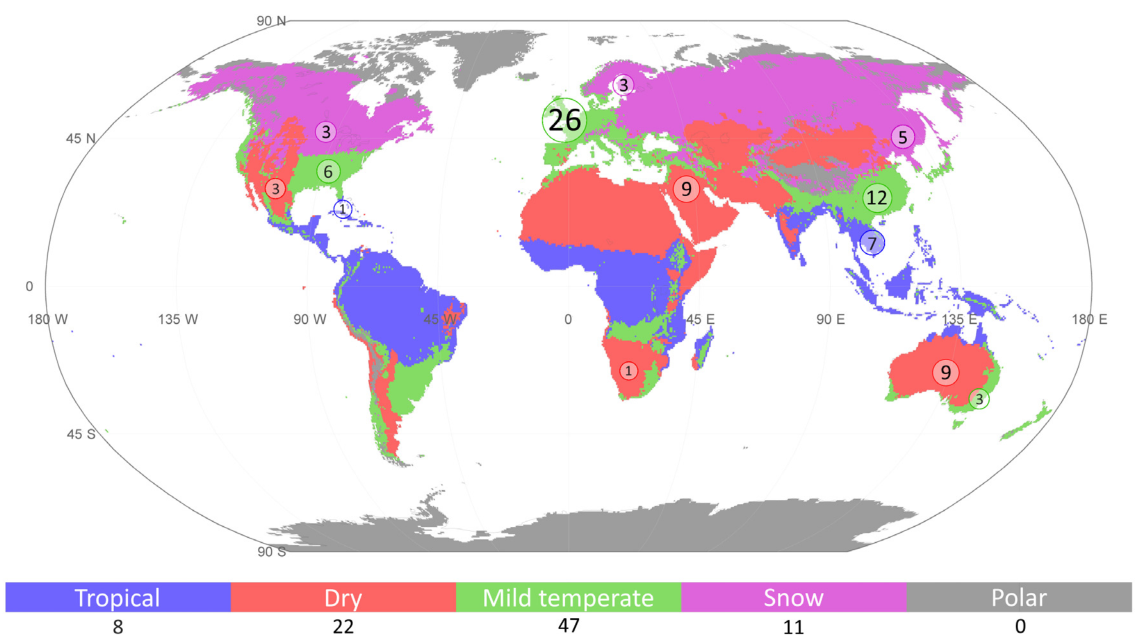


Fig. 9. World map indicating the major Köppen territories in different colours (Chen and Chen, 2013) and the number of systems in each region. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

location affected the optimum parameters of the model. Han et al. (2019) forecasted the PV power from two systems in USA and China and found that the station in the USA was easier to forecast. Similar results were found for four systems spread around the world by Soufiane et al. (2020). Even when considering systems in the same country, such as what Lin et al. and Chen et al. did with two Australian systems, differences are found in the optimum hyperparameters (Lin et al., 2020) and the ranking of the ML algorithms (Chen and Koprinska, 2020). In Sheng et al. (2020) these climatic differences were considered, so an adaptive learning algorithm was implemented for two different systems, able to adjust to different climates and periods. Jung et al. and Lee et al. focused on several sites spread over South Korea (Lee et al., 2018; Jung et al., 2020), but since the country is under mild temperate and cold climates, the studies were left out of Fig. 9.

As just presented, the performance of a PV system is conditioned by the type of climate. Depending on its location, certain features are more important than others and the optimum parameters of the model are affected. This high bias for mild temperate climates entails a problem of generalization. If most

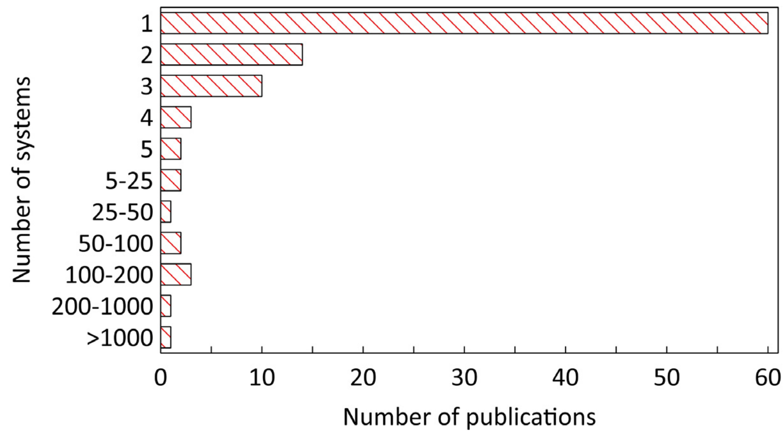


Fig. 10. Number of publications as a function of the number of systems considered in each one.

of the forecasted systems are in mild temperate climates, when a system operator located in another type of climate tries to employ the developed model, the performance will most probably be lower because of the lack of generalization.

4.2. Number of systems

The next classification considers the number of systems used in each publication, Fig. 10. Most studies used data from only one system, while 10% considered more than 10 systems to forecast the PV power.

When more than 5 PV systems were employed, their PV power production was generally aggregated to perform a regional forecast. This was the case of the 16 PV systems employed by Lin and Pai (2016), the 136 by Agoua et al. (2019), the 137 by Wang et al. (2019c) and the 65 by De Felice et al. (2015). In this last study, the systems were differentiated between the North and South of Italy before being aggregated. The systems in the South of Italy showed better performance than those in the North, which was attributed to lower weather variability in the southern part of the country (De Felice et al., 2015).

Wolff et al. (2016) grouped 921 PV systems depending on their location and then divided the clusters into sub-clusters depending on the system specifications. In the end, 10% of each sub-cluster was randomly selected, resulting in a test set of 92 PV systems. The predictions obtained from these single PV systems were employed to produce a regional forecast. When considering the difference between the single site and regional forecasts, the errors in the latter were lower due to averaging effects.

Spatial clustering was also applied in the 1985 PV systems by Pierro et al. (2017) to test two different averaging approaches. In the first approach, the power generation was calculated for each cluster and then averaged to obtain the regional prediction. In the second approach, the regional prediction was directly obtained by inputting the data from each cluster centroid. The second model gave slightly better results.

Exceptions to the regional approach were the forecasts performed in 71 systems by Lee et al. (2018) and in 21 PV systems by Gensler et al. (2017). In the latter, each facility was considered individually and the forecasting error was higher or lower depending on the target system. Unfortunately, the origins of this unpredictability were not reported. Another exception is the study by Jung et al., whose objective was to evaluate the most suitable locations for PV plants in a wide area considering terrain and weather conditions (Jung et al., 2020). They concluded that "the data from one site might not adequately capture the circumstances at other sites with different geographical and topological features".

There is one study which has been left out of Fig. 10, the publication by Takeda (2017). This study did not employ any specific PV system, but rather installed PV capacity. By making use of the monthly installed PV capacities for each type of supplier and weather data, the hourly PV power was estimated for several prefectures in Japan.

4.3. Timespan

In this section, the systems are classified by timespan, defined as the time period of input data employed to develop a model. A summary of the timespan of the reviewed studies is shown in Fig. 11. Most studies employ a timespan of one year. However, a long timespan is essential for accurate training, since the algorithm needs enough data to identify the seasonal variation throughout the year. This is especially important for places having mild temperature climates (Do et al., 2016) where most of the systems are located. It is hence recommended to employ more than one year of data so that the training data covers all seasons. Moreover, a high amount of data is generally required to train a complex algorithm (Wang et al., 2019a).

The main limitation of this requirement is the high computational power required. This gets worse as the time resolution of the input data increases. Training algorithms with such a large amount of data can be unfeasible. Possible solutions could be to use reinforcement or online learning. Alternatives employed in literature are, instead of using the first months of the year, selecting the first period of each month as training data (Wang et al., 2017; Paulescu et al., 2017), or random days along the period (Ramsami and Oree, 2015; Zhang et al., 2018).

These models may also require to be retrained in order to forecast the PV power of unseen systems. Since it is not desirable to repeat the whole training process for each new system, the works (Wolff et al., 2016; Graditi et al., 2016) proposed selecting a few representative days for each new PV system and training the algorithm with this reduced number of samples. This way the generalization to new systems would be done faster.

5. Forecast horizon

The forecast horizon is the time interval into the future for which PV power is predicted. It differs depending on the requirements of the decision-making process for which the PV power is forecasted (Wan et al., 2016). This significantly affects the performance and choice of the prediction algorithm (Antonanzas et al., 2020).

There is no standard classification on how to divide the forecast horizons (Raza et al., 2016; Barbieri et al., 2017; Ahmed

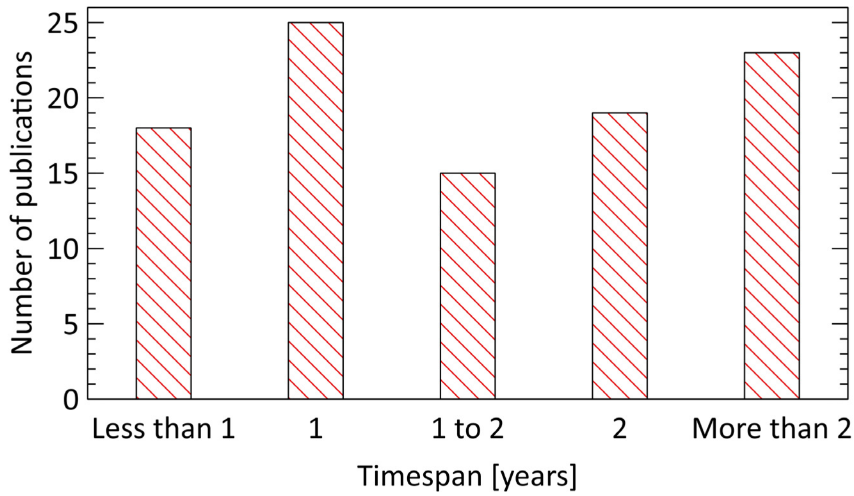


Fig. 11. Number of publications as a function of the timespan employed for PV power predictions.

et al., 2020; Mishra et al., 2020). In this review, we define four categories: *long-term* for predictions longer than a week; *day-ahead* when forecasting between one week and 5 h; *hour-ahead* for predictions between 15 min and 5 h; and *nowcasting* for forecasts smaller than 15 min. These horizons can be related to energy markets: auction (day-ahead), intraday (hour-ahead) and balancing (nowcasting).

Long-term predictions are generally employed for maintenance scheduling, pilot installations and better bankability studies. The two latter objectives deal with PV plants before construction, hence they are usually obtained through physical methods. However, with the use of ML techniques, a forecast of the purchased PV power for the following years on a monthly basis can be reported (Lin and Pai, 2016; Takeda, 2017).

In the day-ahead energy market, the economic dispatch problem is solved by knowing how much electricity will be produced by PV sources, which determines the day-ahead price. Trading outside cross-border zones also occurs in this market segment (Poplavskaya et al., 2020). Day-ahead forecasting can also be used for Internet-of-Things (Kraemer et al., 2020). When knowing the harvestable energy, the demand can be adjusted to match the power availability.

Since weather conditions can change during the day, the predictions may not be as accurate as desired or unplanned outages can occur. In the case of these events, energy can be traded in the intraday markets (Ortner and Totschnig, 2019). With the increased share of renewable energy sources, these markets have gained importance as a complement to day-ahead markets (Maciejowska et al., 2019). Hour-ahead predictions obtain a higher accuracy so that final remarks can be made on the PV power output to account for recent errors.

The accuracy of these predictions may still not be perfect, or some last minutes changes may occur so that the demand does not meet the supply. To avoid PV power curtailment and system instability, nowcasting can be employed. Nowcasting is defined as predicting the present or very near future (here considered as 15 min or less). This technique can be employed to provide last minutes corrections which are used by utility operators to bring spinning reserves online (Rosiek et al., 2018). Nowcasting can also be employed for early anomaly detection (Torabi et al., 2017).

An overview of the forecasting ranges and their main applications is shown in Table 4.

Following this criterion, like in previous sections, the analysed literature has been classified as shown in Fig. 12. A fifth group has been included for studies that forecast the PV power in time

Table 4
Classification of the time horizons employed in this review and their most common application.

Category	Time horizon	Application
Long-term	More than one week	Maintenance scheduling Pilot installations Bankability studies
Day-ahead	5 h to one week	Economic dispatch Day-ahead price determination
Hour-ahead or intra-day	15 min to 5 h	Unplanned outages compensation Unit commitment
Nowcasting	Less than 15 min	Power balance Electricity market clearing Early anomaly detection

horizons belonging to more than one category. The results show that most publications focus on day-ahead forecasting, while only a few belong to the nowcasting and long-term categories. Given that the forecast depends on the application, the authors do not recommend any specific category.

Some researchers studied the behaviour of ML models under different time horizons. Generally, they found that the higher the time horizon, the higher the error (Eseye et al., 2018; Paulescu et al., 2017; Li et al., 2018; Chang and Lu, 2020; Li et al., 2016b). De Giorgi et al. (2016) found that LS-SVR achieved an NMAE of around 6% when predicting the PV power 1 h ahead, and the value increased to around 21% when predicting 12 h ahead. They stated that when predicting longer times, the PV power peaks were underestimated, while the valleys were overestimated. Similarly, in Pierro et al. (2017) an RMSE of 5%–7% was obtained for the intra-day forecast (1 to 4 h) while this value increased to 7% and 7.5% for one and two-day forecasts, respectively. One exception was found in Massidda and Marrocu (2017), where the day-ahead forecasting had a higher accuracy than when estimating 15 min, 1 h and 3 h ahead. They claimed this was obtained for the day-ahead prediction because of compensation along the day.

The forecast horizon also affects the ranking of algorithms. In Sharadga et al. (2020) the performance of several algorithms was tested under one-, two- and three-hour ahead forecasting for a 20 MW grid-connected PV station in China. For one- and two-hour ahead, LSTM showed the best performance with RMSE of 0.841 and 1.102 MW, while ANN achieved 0.961 and 1.395 MW respectively. However, for three-hour ahead forecasting, ANN achieved lower (1.805 MW) RMSE than LSTM (1.824 MW). Similar results were obtained in Wolff et al. (2016).

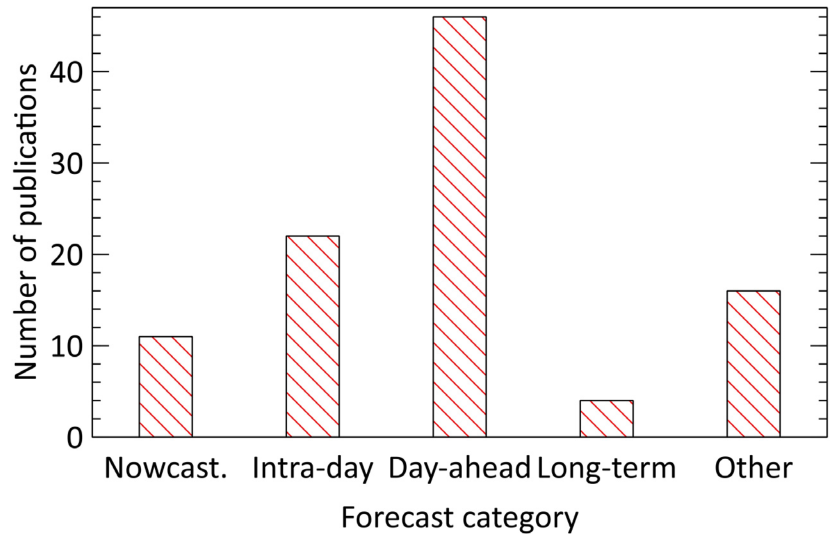


Fig. 12. Number of publications as a function of the forecast category that they belong to. A new group has been included for publications which belong to more than one category.

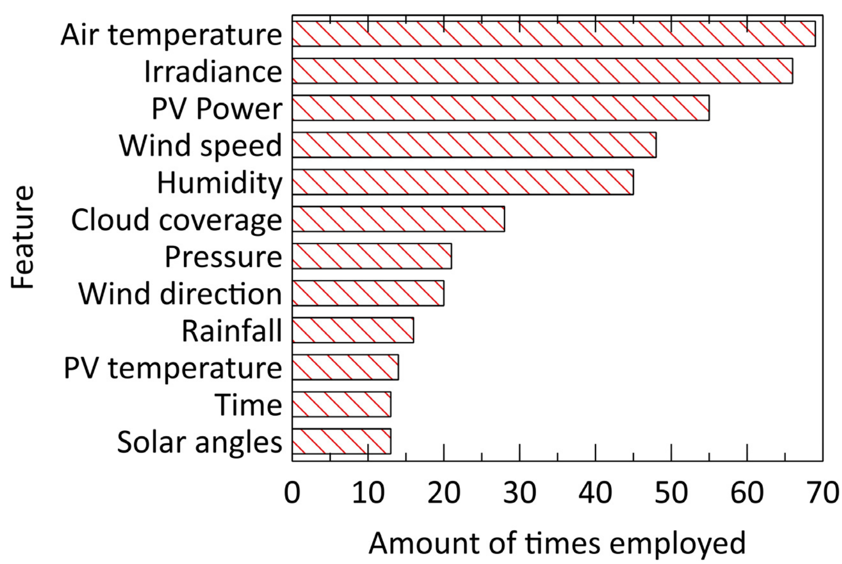


Fig. 13. Number of times that each feature has been employed for predictions in the publications reviewed.

6. Choice of features

In this section, the inputs employed for PV power forecasting, called features in a machine learning model, are explored. In Fig. 13, the most common features have been depicted as a function of the number of publications in which they have been employed. Here, the irradiance bar includes several forms such as plane of array or diffuse irradiance.

As expected from a physics-based approach, irradiance and temperature are the most important features. However, none of these two features is essential for ML predictions. Several works have been performed which rely only on previous PV power outputs, such as Sharadga et al. (2020), Li et al. (2015), Asrari et al. (2017), Chang and Lu (2020) and Li et al. (2019a). These methods are more flexible to apply in practice. The use of previous PV power measurements in combination with weather data can considerably increase the performance of the algorithms (De Giorgi et al., 2016). It is also important to consider that some of the ML methods consider previous PV power values intrinsically (such as LSTM).

One surprising fact from a solar engineering perspective is the large number of papers that employ humidity as a feature. According to Fig. 13, this parameter is more popular than cloud coverage, which is usually essential for accurate physical predictions. Researchers in Li et al. (2019b) found that humidity is negatively correlated with PV power. An increase in moisture implies an increase in the absorption, reflection and refraction of sunlight, hence it reduces the radiation received by PV panels.

One would think that the larger the number of features, the better. However, as already mentioned in Section 2.3.1, extra variables can be detrimental to the model's performance and can increase the demand for computational power. For instance, in Kraemer et al. (2020) pressure was detrimental to the ML model, while humidity and rainfall did not influence performance. Feature selection identifies the most relevant features for each particular model. Literature results show that the choice of features depends on the feature selection algorithm employed (Touati et al., 2017; Massaoudi et al., 2019), on location and even on specific events. For instance, in Kuzmiakova et al. (2017), wind speed was not as relevant in November as in the rest of the months due to exceptional hurricane-force winds. Similar results were reported by Abuella and Chowdhury (2017), who showed that the correlation between parameters has seasonal dependency. On the importance of location, Touati et al. (2017) provided a good example. They predicted the PV power output of a system located in Qatar, a desert-like climate. The results from the feature selection algorithm showed that dust was one of the most relevant features, while it was not employed in any of the other studies included in this review.

To conclude, there is no optimum set of features for all cases. When in presence of a new model, the reader is recommended to check previous literature and gather as many relevant features as possible. Fig. 13 can give a hint about the latter. Despite some models being able to identify relevant features (such as the ones explained in Section 2.3.1, and some DL algorithms (Massaoudi et al., 2021)), it is still recommended to do at least feature exploration. The use of less important features (e.g., sunshine duration) is not recommended unless no more relevant features are available, since they can still provide the information indirectly.

Use of 2D features

A special mention needs to be made in this section regarding the use of 2D features. A source of uncertainty in PV power predictions is cloud coverage information. Rapidly changing cloud coverage has a significant influence on PV modules output and is a big challenge in solar yield forecasting. Cloudy days are harder to forecast than sunny ones (Ferlito et al., 2017; Sharadga et al., 2020; De Giorgi et al., 2016; Rosiek et al., 2018; Baharin et al., 2016; Wolff et al., 2016).

To overcome this, some groups have centred their efforts on employing satellite images to make an accurate estimation of cloud movement. This approach, which is extensively used for irradiance forecasting, is now gaining popularity for the prediction of PV power. Examples are the works by Wolff et al. (2016), Rosiek et al. (2018) and Pierro et al. (2017). However, satellite images have either high temporal or high spatial resolution but never both Dev et al. (2016), while both characteristics are required for PV power prediction, especially in the short-term (Pazikadin et al., 2020).

An alternative method to satellite imaging is employing sky images. Ground-based sky imagers are gaining popularity since they have both high temporal and high spatial resolution. Sky images were employed in Anagnostos et al. (2019) to detect the cloud type. These images can also be employed for comprehending cloud movement, as in Zhang et al. (2018) and Sun et al. (2019). Chu et al. (2015) uniquely used these images in their work. First, they employed an ANN to forecast the PV power, as done by other researchers. Then, applying simpler methods together with sky images, they re-forecasted the PV power to increase the accuracy. The main disadvantage of sky images is that they are local, hence cannot be employed for wide-area forecasting.

For more information on satellite and sky imaging, the reader is referred to the reviews by Ahmed et al. (2020) and Feng et al. (2021).

In this section, we have not discussed the origin of the weather data. The input data can come from different sources such as numerical weather predictions or neighbouring PV systems. The reader is referred to Antonanzas et al. (2016) and Feng et al. (2021) if more information on the topic is yearned for.

7. Outlook

This section presents the main recommendations given to the readers based on the discussion performed in previous sections.

Machine learning algorithms

Starting with the ML algorithms, we have seen in Section 2 that there is a wide range of families to choose from. Some algorithms are too simple to properly forecast the PV power output, but most of them can yield reasonable results. Indeed, linear regression approaches are unable to correctly model the non-linearities of PV power forecasting, but once this step is overcome, the difference in performance between MARS and a DL hybrid algorithm is not that high (Li et al., 2016a; Ferlito et al., 2017). In these cases, developing a model that would be applied in practice is more relevant. This can imply that the algorithm is interpretable, that the requirements in terms of training data are not high, that the model is not computationally expensive, and that it is flexible enough to be adapted to unseen PV systems with different characteristics and climates.

Generalization

Regarding the last point, most ML methods are usually tested on a few systems under the same climate so their capacity to adapt to unseen PV plants is unknown. This can be an issue when these models are put into practice.

Starting with the number of PV systems (Section 4.2) most researchers motivate their study by contributing to improved generator dispatch, power quality effects mitigation, and reducing secondary reserve capacity (Theocharides et al., 2018), but if their results are not validated for multiple systems, they are not reliable enough for upscaling. This hinders the practical use of the developed methods since they are not fully tested or verified under different conditions. Moreover, there is a need for more regional forecast studies, which will provide the basis for grid management.

A similar issue is found with the location of PV systems (Section 4.1). Most studies focus on systems located in the same climate and country. We are still missing a study which employs PV systems from different places on the earth, which analyses the differences in performance between the climates and can adapt to any type of PV system. One of the barriers to achieving this is that large-scale data acquisition is costly and might be difficult for research institutions. Lack of coordination between researchers and the industry might result in an unnecessarily long search for optimal solutions and little implementation.

Data acquisition

Another issue related to data acquisition is its availability and quality. There is generally enough available data, but frequently not with the desired resolution. This affects, for instance, the choice of the forecast horizon. Nowcasting is currently the most demanded forecast horizon because of cloud-related instability (Antonanzas et al., 2016), but the amount of data available with the required time resolution is limited. This explains the small percentage of studies focusing on this forecast horizon (see Fig. 12). Similarly, the lack of sky images and high-resolution cloud coverage information is a major concern.

Data quality is another important factor, even more than the chosen algorithm. Incorrect data will negatively affect the algorithm, without consideration of its complexity (Escalante, 2005). This is also the main source of error in complex physical forecasting models. Unless the objective of the forecasting model is to predict the power produced by a specific PV plant with a meteorological station nearby, we recommend not using local weather parameters such as wind speed and direction. These lowly correlated parameters (Li et al., 2017) can only add noise if they are not measured on-site. Other highly correlated features with lower spatial sensitivity such as temperature will not be so detrimental to the model.

It is also important to keep in mind that up to two hours ahead the previous PV power is more important than numerical weather predictions (Abuella and Chowdhury, 2017). Although

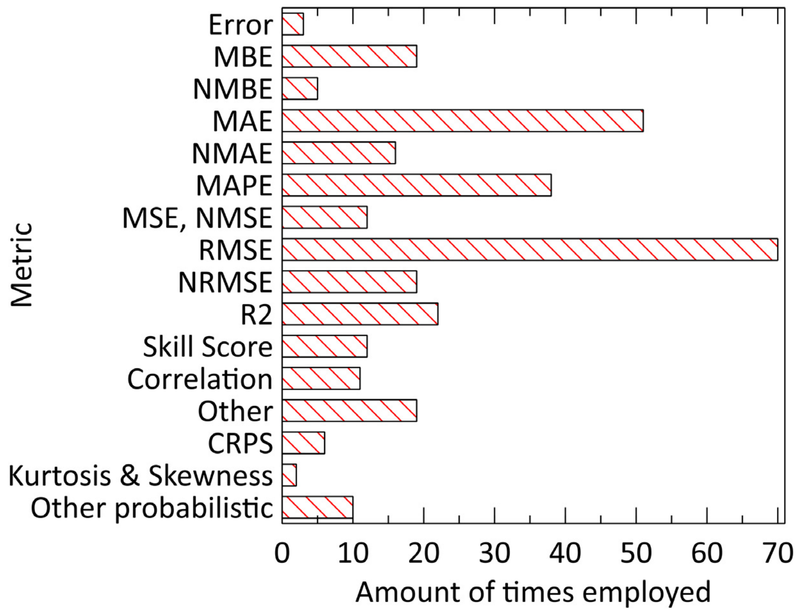


Fig. 14. Metrics as a function of the number of times employed in the literature explored in this review.

we recommend PV power as a feature, with only that the accuracy of the predictions is very limited, especially during cloudy days. From our point of view, the best approach is to complement previous PV power and clear-sky irradiance with cloud movement extracted from images. Unfortunately, the use of imaging is currently limited by the low geographical resolution of satellites and the small coverage and price of ground-based sky images.

Metrics

Due to the extensive literature covering metrics for PV power forecasting (Yang, 2019; Antonanzas et al., 2020; Zhang others, 2015; Armstrong and Collopy, 1992) this is not a focus of this work. Appendix C: Metrics describes the main metrics for completeness. However, we still want to provide our insights based on Fig. 14, which represents the popularity of the main metrics.

RMSE is by far the most popular measure, commonly preferred in many disciplines as it provides good sensitivity (Armstrong and Collopy, 1992). Other popular metrics are MAE and its normalizations. Although not included, visual inspection is also an important representation commonly employed by researchers.

Although it can be considered for comparison purposes, popularity should not be the determining factor when choosing a metric. The most suitable metric depends on the characteristics of the system such as geographic location, the forecast timescales and the objective of the study (Zhang others, 2015; Armstrong and Collopy, 1992). All these factors complicate the existence of a standard metric or group of metrics. As a general recommendation, it is important to employ more than one measure of prediction error and to make an analysis of the type of error made. This will display the flaws of the proposed approach for further improvement. We also recommend the use of normalized metrics for easier comparison and understanding.

A look into the future

To finalize, we would like to give our view on future developments in the field. We believe that in the near future ML-based methods for PV power forecasting will be more extended at the industry level. There are currently some factors hindering their application in real life, such as robustness and interpretability. These can affect the choice of the algorithm and could distance from highly potential models such as neural networks, which have lower robustness due to their inherent randomness and are hard to optimize (Li et al., 2016a). We think that interpretability strategies such as explainable AI will most likely gain ground since in practice model decisions need to be justified. Probabilistic approaches will most likely also be important, especially in the day-ahead time frame, since they provide flexibility to the predictions. We also hope that the time and spatial resolution of satellite images will increase fast enough so that they can be more highly employed in the near future.

There are some other aspects to take into consideration regarding the expansion of this field. For instance, the velocity at which PV capacity is deployed. Grid limitations as well as administrative regulations slow down the installation of new facilities. The lower the importance of PV in the electricity generation, the lower the demand for accurate forecasts. Storage development will probably also affect the deployment in practice of AI since in presence of large and cheap storage solutions, supply–demand balancing will become easier, and there will be fewer incentives to have accurate prediction solutions.

Overall, it is important to keep in mind that no model is going to yield 100% accurate predictions: not only because it is unfeasible to collect all the necessary data, but also because of errors in the data collection which are swept along in all the phases of the prediction. The most important concern should be to focus on the application one is willing to achieve, and give the best possible result knowing that there will be some uncertainty in it.

8. Conclusion

In this review, we have provided an overview of the state of the art of PV power forecasting employing machine learning techniques. The main contribution of this work with respect to previous ones published on the topic is the extensive categorization of the literature. We have classified according to the characteristics of the PV systems employed for forecasting, as well as the forecast horizon and input features. This has highlighted gaps that have not been reported previously, such as the lack of studies in certain parts of the world or the focus on day-ahead predictions.

The literature has also been classified according to the forecasting technique employed, giving special emphasis on grouping the ML algorithms into categories. We have given an overview of all ML families, instead of focusing on the most popular algorithms, in order to provide alternatives to researchers wanting to deviate from the standard. A special mention has been made of alternative approaches such as cloud imaging and probabilistic forecasting.

One of the limitations of this study is the specific range of years that it focuses on. By looking from 2015 until 2020, we are not covering the most up-to-date literature and we are neither looking at long-term trends. The process employed to obtain that literature was also biased by the most cited papers. However, the large number of publications considered reduces this bias.

Finally, the outlook has summarized the main outcomes of each section while providing recommendations to everyone interested in the field. Although the objective was to make a review as impartial as possible, the authors' opinions and point of view needed to be expressed across the paper, especially in this last section. This could lead to researchers being misled, but it can also be a valuable contribution to the field.

Declaration of competing interest

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

Data availability

No data was used for the research described in the article.

Acknowledgements

The research leading to these results has received funding from the Horizon 2020 Programme, under Grant Agreement 952957, Trust-PV project.

Appendix A. Reviewed literature

See Table A.1.

Appendix B. ML families

This appendix provides the main distinctive characteristics of each of the 10 ML families considered in this work. The objective is not to provide a detailed explanation, but an overview of the differences between families.

Regression

With their origin in statistics, regression methods have been the starting point of machine learning thus they have been adopted as part of the discipline (Theodoridis and Koutroumbas, 2009). This group is mainly composed of the linear regression method and its variations. It is important to distinguish between regression and auto-regressive methods. The main differences are how the linear coefficients are obtained and that auto-regressive methods include an error term.

The equation below depicts an example of the estimation of the PV power $\widehat{P}(t)$ at time step t, using a simple linear regression model. Here we consider only the irradiance $I(t-1)$ and the temperature $T(t-1)$ at the previous time step as inputs to the model.

$$\widehat{P}\left(t, \overrightarrow{\alpha}\right) = \alpha_0 + \alpha_1 I(t-1) + \alpha_2 T(t-1) \tag{B.1}$$

The coefficients of the linear regression model $\vec{\alpha} = [\alpha_0, \alpha_1, \alpha_2]$, like in most machine learning algorithms, are found by minimizing the loss function $\mathcal{L}$ between estimated and real PV power $P(t)$. The loss function can take several forms, being the quadratic loss amongst the most popular ones.

$$\min_{\overrightarrow{\alpha}} \left(\sum_t \mathcal{L}\left(\widehat{P}\left(t, \overrightarrow{\alpha}\right) - P(t)\right) \right) \tag{B.2}$$

It is also relevant to highlight the Multivariate Adaptive Regression Splines (MARS) algorithm in this family. This approach improves the linear regression by including hinge functions (Friedman, 1991). They allow non-linear relations between the input and the output, which increases flexibility at a cost of complexity (Li et al., 2016a).

Regularization

A common issue encountered when solving machine learning problems is *data overfitting*. This phenomenon occurs when the algorithms fit the testing set so much that they are unable to fit unseen data. To solve this issue, regularization techniques were developed in order to limit the size of the parameters to be optimized. This method is generally applied to linear regression models.

For instance, the normalization of the linear regression algorithm from Eq. (B.1) can be applied by limiting the size of the coefficients $\vec{\alpha}$. This is imposed by adding a penalization term $\mathcal{R}(\vec{\alpha})$ to the function to be minimized. Hence, the objective function in Eq. (B.2) becomes Eq. (B.3). When the algorithm performs the minimization, it is motivated to keep the coefficients low.

$$\min_{\overrightarrow{\alpha}} \left(\sum_t \mathcal{L}\left(\widehat{P}\left(t, \overrightarrow{\alpha}\right) - P(t)\right) + \mathcal{R}\left(\overrightarrow{\alpha}\right) \right) \tag{B.3}$$

Depending on how the penalization of the parameters is defined, one encounters several algorithms: Least Absolute Shrinkage and Selection Operator, 1-norm (absolute value) penalization; Ridge Regression, 2-norm (root squared value) penalization; and Elastic Net, 1-norm and 2-norm penalizations.

Bayesian algorithms

Bayesian algorithms use Bayes' theorem as a basis to develop their models (Brownlee, 2021). Bayes' theorem relates the probability of a certain event to the knowledge a priori of conditions related to that event. In terms of the previous example, Bayes' theorem finds $p(P|I,T)$, which is the probability to produce the PV power P given a certain irradiance I and temperature T (time dependencies have been removed for easier understanding). $p(P|I,T)$ is related to the prior, which is the distribution of the PV power $p(P)$ before knowing the meteorological conditions, and to the likelihood, which is the probability to measure certain weather conditions knowing the produced PV power $p(I,T|P)$. Eq. (B.4) depicts this relation, where $p(I,T)$ is the distribution of meteorological conditions, obtained via normalization of the product of probabilities.

$$p(P|I,T) = \frac{1}{p(I,T)} p(P)\, p(I,T|P) \tag{B.4}$$

Depending on the distribution assumed for the terms of Bayes' theorem, different Bayesian algorithms arise. If the probability distribution of the produced PV power is assumed to be Gaussian, we encounter the Gaussian process regression (Schulz et al., 2018). If independence between features is also assumed, we encounter the Gaussian Naïve Bayes algorithm (Cervone et al., 2017; Brownlee, 2021). However, this last assumption is strong and unlikely in real data.

Instance-based

These algorithms find important points within the training data set that describe the problem at hand (Brownlee, 2021). One of the most known algorithms inside this block is the k-Nearest Neighbour (k-NN). In presence of a new input, k-NN looks at the k nearest neighbours of the input in order to decide the output.

The main character in this family is the Support Vector Machine (SVM). This algorithm identifies the most relevant data

Table A.1
Summary of the main characteristics of the reviewed literature. Below the table, one can find the meaning of the abbreviations.

Ref.	Country	Köppen	Features	ML algorithms
Abdel-Nasser and Mahmoud (2019)	Egypt	B	pow	LR dt FFNN LSTM
Abuella and Chowdhury (2017)	Australia	B	pow irr ta h cc r ws wd p t	LR SVR MLP dr
Agoua et al. (2019)	France	C	pow irr ta h wd	bay regu ELM
Al-Dahidi et al. (2020)	Jordan	B	pow irr ta h ws t	FFNN ELM
Alessandrini et al. (2015)	Italy	C	pow irr ang ta cc	regr FFNN ens
Alfadda et al. (2017)	USA	C	pow ta h cc ws wd t	LR regu SVR
Alkandari and Ahmad (2020)	–	–	irr ta h r ws wd	LSTM DL hyb
Almeida et al. (2015)	Spain	C	irr ta h cc ws wd p	dt dr
Alomari et al. (2019)	Jordan	B	pow irr ta t	BPNN
Anagnostos et al. (2019)	Germany	C	irr ang tpv cc	MLP
Asrari et al. (2017)	USA	C	pow	FFNN GA opt
Baharin et al. (2016)	Malaysia	A	irr ta h ws	SVR dr
Carrera and Kim (2020)	South Korea	C	irr ang ta h cc r ws wd p	LR regu k-NN SVR dt ens
Cervone et al. (2017)	Italy	C	pow irr ang ta cc	BPNN ens hyb
Chang and Lu (2020)	Taiwan	C	pow	SVR BPNN ANN DL
Chen and Koprinska (2020)	Australia	–	pow	LR SVR dt FFNN LSTM ens hyb
Chu et al. (2015)	USA	B	pow	k-NN FFNN GA
Das others (2017)	Malaysia	A	pow irr ta ws	SVR BPNN
De Felice et al. (2015)	Italy	C	irr ta	SVR
De Giorgi et al. (2016)	Italy	C	pow irr ta tpv	LS-SVR DL hyb
Do et al. (2016)	–	–	pow ta cc	regr FFNN
Eseye et al. (2018)	China	D	pow irr ta h cc ws p	SVR BPNN pre GA PSO
Ferlito et al. (2017)	Italy	C	pow irr ta cc	LR MARS k-NN SVR dt ANN ELM ens
Gensler et al. (2017)	Germany	C		MLP LSTM DL hyb
Golestaneh et al. (2016)	–	–	pow	ELM
Graditi et al. (2016)	Italy	C	irr tpv	MLP LSTM pre
Gulin et al. (2017)	Croatia	C	irr ta h ws wd p	MLP
Han et al. (2019)	–	–	pow ta h p	SVR LSTM dr
Hossain et al. (2017)	Malaysia	A	pow irr ta tpv ws	SVR FFNN ELM
Huang et al. (2016)	USA	D	irr ang tpv	bay SVR FFNN
Huang and Kuo (2019)	Taiwan	C	pow irr ta	SVR dt MLP LSTM CNN ens
Isaksson and Conde (2018)	Sweden	D	pow irr ta h cc r ws wd p	regu k-NN BPNN ens dr
Jung et al. (2020)	South Korea	–	irr ta h cc rwsr ws t	LSTM DL
Kazem and Yousif (2017)	Oman	B	irr ta	SVR FFNN MLP DL
Konstantinou et al. (2021)	Cyprus	B	pow	LSTM
Kraemer et al. (2020)	Norway	D	ang ta h cc r p	ANN DL ens
Kumar and Kalavathi (2018)	India	A	pow irr ta h rwsr ws p	FFNN fuz
Kuzmiakova et al. (2017)	USA	D	ta h cc ws p	regr LSTM ens dr
Lee et al. (2019)	South Korea	D	ang ta tpv h cc rwsr ws wd	SVR dt FFNN DL ens dr
Lee et al. (2018)	South Korea	–	pow irr ta h rwsr ws	LR regu regu SVR ens hyb
Lee and Kim (2019)	South Korea	D	irr ta h cc t	BPNN LSTM
Li et al. (2019b)	Australia	B	irr ta h	SVR ANN PSO opt
Li et al. (2016a)	China	C	ta h r ws p	LR MARS k-NN SVR dt FFNN
Li et al. (2015)	China	C	pow	BPNN ELM
Li et al. (2020)	Australia	B	pow irr ta h ws	MLP DL LSTM DL pre
Li et al. (2018)	South Africa	B	pow ta h	BPNN DL
Li et al. (2016b)	USA	C	pow ang ta ws wd	SVR FFNN
Li et al. (2019a)	Belgium	C	pow	SVR BPNN ANN DL LSTM
Lin and Pai (2016)	Taiwan	C	pow	LS-SVR ANN GA
Lin et al. (2018)	Australia	B	pow irr ta h ws	LS-SVR BPNN ANN DL hyb pre
Lin et al. (2020)	Australia	–	pow irr ta r ws	ANN hyb
Maitanova et al. (2019)	Germany	C	pow ta h cc r	LSTM opt
Majumder et al. (2020)	USA	–	pow	SVR ANN ELM opt
Massaoudi et al. (2019)	Australia	B	pow irr ta h ws wd t	bay regu k-NN ens dr
Massidda and Marrocu (2017)	Germany	C	irr ta cc ws p	MARS
Mishra et al. (2020)	USA	D	pow ang ta h cc ws p	LR regu LSTM pre opt
Nageem and Jayabarathi (2017)	India	A	irr ta h ws p t	SVR
Nayak and Heistrene (2020)	India	B	irr	SVR dt ANN ELM hyb
Ni et al. (2017)	Singapore	A	pow irr ta tpv ws wd	ELM
Ogliari et al. (2018)	Italy	C	irr ta cc rwsr ws wd p t	hyb
Pan et al. (2020)	Australia	B	irr ta h wd t	SVR opt
Panamtash et al. (2021)	USA	A	pow	BPNN
Paulescu et al. (2017)	Italy	C	irr ta cc	LR regr fuz
Pawar et al. (2020)	Australia	C	irr	SVR
Pierro et al. (2017)	Italy	C	pow irr ang ta h ws	k-NN MLP pre dr
Pujić et al. (2020)	Denmark	C	irr ta h cc ws wd p	LR k-NN SVR FFNN ens
Ramsami and Oree (2015)	UK	C	irr ta h rwsr ws wd p	LR FFNN ANN dr
Rana et al. (2015)	Australia	C	pow irr ta h ws	SVR MLP
Raza et al. (2018)	Australia	C	pow irr ta h ws	FFNN ANN DL pre
Rosato et al. (2017)	Italy	C	pow	LR ANN fuz
Rosiek et al. (2018)	Spain	C	irr ang ta h ws wd p	BPNN pre dr
Semero et al. (2018)	China	D	irr ta tpv h cc p	LR BPNN fuz dr GA PSO
Shang and Wei (2018)	USA	B		MLP BPNN ANN fuz pre dr PSO

(*continued on next page*)

Table A.1 (*continued*).

Ref.	Country	Köppen	Features	ML algorithms
Sharadga et al. (2020)	China	C	pow	MLP ANN LSTM DL fuz pre
Sheng et al. (2020)	–	–	irr ta h ws t	bay SVR ANN ELM opt
Soufiane et al. (2020)	–	–	pow	ANN pre
Sun et al. (2019)	USA	C	pow cc	CNN
Suresh et al. (2020)	Poland	C	irr ta tpv ws	LR CNN hyb
Takeda (2017)	Japan	C	pow irr ta ws	ens
Tang et al. (2016)	USA	C	irr ta h ws	ANN ANN ELM
Theocharides et al. (2018)	Cyprus	B	irr ang ta h ws wd	SVR dt FFNN
Torabi et al. (2017)	Portugal	C	pow irr	BPNN
Touati et al. (2017)	Qatar	B	irr ta tpv h ws	LR dt dr
Van Tai (2019)	Russia	D	irr tpv ws	LR FFNN fuz dr
Wang et al. (2019b)	Australia	B	irr ta h ws wd	LSTM CNN hyb
Wang et al. (2018)	Australia	B	irr ta h cc r ws wd p	SVR BPNN LSTM DL hyb dr
Wang et al. (2020)	USA	B	irr ta	k-NN BPNN LSTM
Wang et al. (2017)	Belgium	C	pow	regr SVR BPNN CNN pre
Wang et al. (2019c)	USA	C	pow	CNN
Wang et al. (2016)	China	C	irr ta h rwsr ws p	LR regr MLP ANN
Wen et al. (2019)	Australia	B	irr ta h ws t	SVR MLP hyb
Wolff et al. (2016)	Germany	C	pow irr ta cc	SVR
Yadav et al. (2018)	India	A	irr tpv	ANN
Yang et al. (2020)	Netherlands	C	pow	ANN opt
Yin et al. (2020)	China	D	irr ta h ws	SVR BPNN ELM
Yousif et al. (2017)	Oman	B	irr ta	SVR MLP DL
Zang et al. (2018)	China	C	pow irr ta ws	bay SVR BPNN CNN ens hyb pre
Zhang et al. (2015)	–	–	irr ang ta tpv cc ws	k-NN LS-SVR ANN
Zhang et al. (2018)	Japan	C	pow cc	MLP LSTM CNN
Zhou et al. (2018)	China	C	pow tpv t	SVR pre

The five Köppen categories are: A—tropical, B—dry, C—mild temperate, D—continental and E—polar.
The abbreviations for the features are: pow—PV power; irr—irradiance; ang—solar angles; ta—air temperature; tpv—PV temperature; h—humidity; cc—cloud coverage; r—rainfall; ws—wind speed; wd—wind direction; p—pressure; t—time. Be aware that only the most relevant features have been mentioned, but some authors have included more.
The abbreviations for the ML algorithms have been mentioned in the text. If one or more algorithms from a certain family have been employed and are not one of the popular algorithms, the abbreviation of the family is mentioned. The abbreviations correspond to: regr—regression; regu—regularization; bay—Bayesian algorithms; dt—decision trees; ens—ensemble algorithms; ANN—Artificial Neural Networks; DL—Deep Learning; fuz—fuzzy algorithms; hyb—hybrid algorithms. In the next appendix, the main characteristics of each family are briefly mentioned.

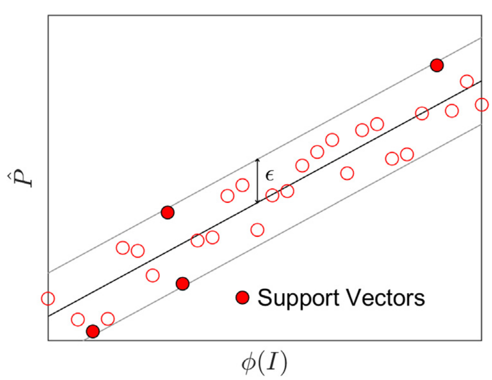


Fig. B.1. Schematic example of Support Vector Regression applied for the prediction of the PV power as a function of the transformed irradiance.

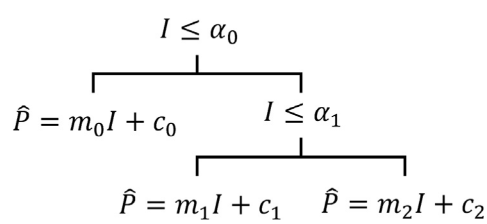


Fig. B.2. Example decision tree for PV power prediction.

points during training and employs this subset for prediction (Bishop, 2006). The method was originally developed for classification problems, but it was recently extended for regression under the name of Support Vector Regression (SVR).

Fig. B.1 exemplifies the use of SVR to predict the PV power $\widehat{P}$ as a function of the kernel of the irradiance $\phi(I)$. The kernel is a transformation made to the input so that a linear relationship can be found between the transformed input and output. This linear relationship is found by minimizing the margin of tolerance ϵ so that all the points fall within the line with the margin. The points that are within the borders of the margin of tolerance determine the parameters of the linear relationship and are called the support vectors.

A simplification in the learning procedure of the SVR leads to the Least-Squares SVR (LS-SVR), making the solving process more efficient. This results in a lower computational cost, without any effect on accuracy.

Decision trees

The algorithms in this class predict using a tree-like model of decisions. The data space is recursively partitioned depending on its value, and a simple prediction model is then fitted within each partition (Loh, 2011). Fig. B.2 shows a very simple decision tree that predicts the PV power as a function of irradiance. The difference between the models within this family lies in their underlying algorithms which determine for example the type of prediction model fitted within each partition (Loh, 2011).

Ensemble algorithms

Ensemble algorithms combine many weak learners (also called predictors) into one strong learner (Brownlee, 2021). A weak learner is defined as one whose performance is slightly better than random chance. For instance, random forests and gradient boosts regression trees are made of a combination of decision trees. An example of their structure can be seen in Fig. B.3, where the result of three different decision trees is combined to improve the estimation of the output. One can also combine deterministic forecasts to create a probabilistic one. This is the working method of the Analog Ensemble. This algorithm starts by producing a

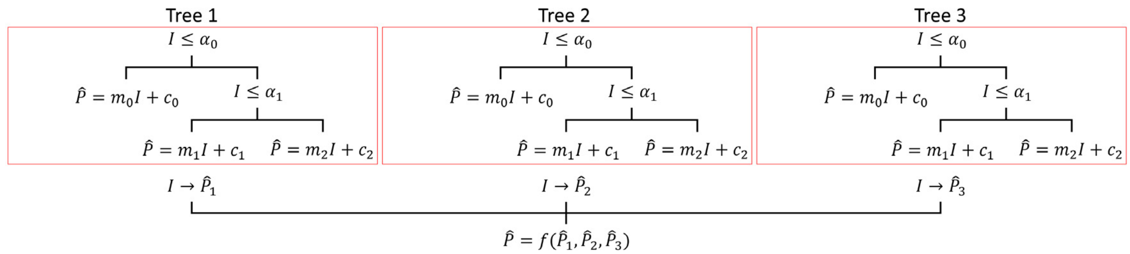


Fig. B.3. Schematic of random forest, where the output of three different decision trees is combined to improve the prediction.

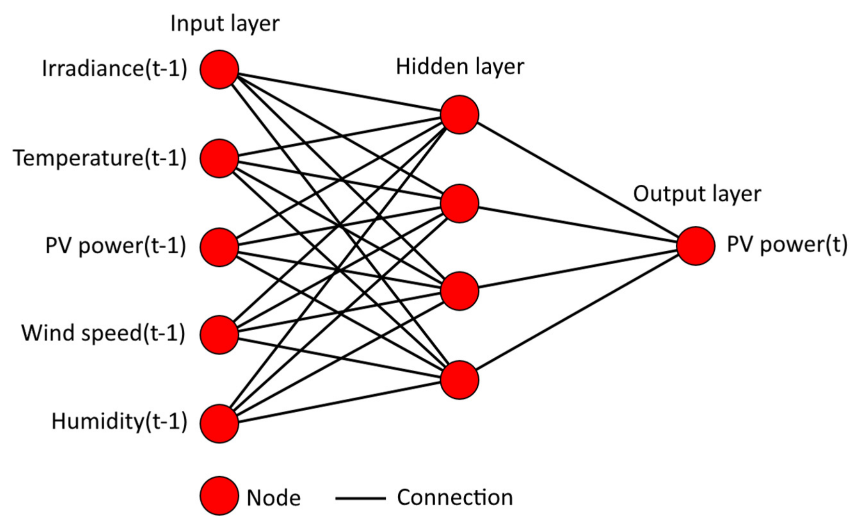


Fig. B.4. MLP or FFNN structure for a PV power forecasting problem.

short-term deterministic forecast with training data. When new data is provided, a deterministic forecast is computed, and its result is compared to the most similar past forecasts. These are then combined to generate a bias-adjusted probabilistic forecast.

Artificial neural networks

Artificial Neural Networks (ANN) were inspired by the synaptic connections in the nervous system, hence their name (Wang, 2003). An ANN consists of an input layer of nodes (neurons), several hidden layers and a final layer of output neurons (Cross et al., 1995). The nodes are interconnected via a set of weights such that signals can travel through them (Fausett, 2006). The output of each node is the result of an activation function, applied to the aggregated weighted values of the node's input (Wang, 2003; Dongare et al., 2008). The purpose of the activation function is to introduce nonlinearity to the structure so its flexibility is increased, making this configuration very powerful (Cross et al., 1995). They have an implementation drawback since the optimal configuration and learning algorithm are hard to obtain a priori (Li et al., 2016a).

The weights assigned to each node are modified by experience while training. The most common method to train these structures is using back-propagation (Gardner and Dorling, 1998), hence these ANN are referred to as back-propagation neural networks (BPNN).

If the nodes of the ANN never form a cycle, the ANN is called a feed-forward neural network (FFNN) or multilayer perceptron (MLP) (Gardner and Dorling, 1998). This is the first and simplest type of ANN (Wang, 2003). Fig. B.4 shows graphically its structure for a PV power forecasting problem. All the nodes are fully connected between consequent layers and the information flows from the input to the hidden and finally the output.

Several characteristics can be tuned to create different types of ANN (Tch, 2017), such as applying a different activation function in each node, implemented by the Radial Basis Function and generalized regression ANNs.

Deep learning

Deep Learning (DL) is a subset of ML, in particular a subset of ANNs. These ANNs are characterized by having a deeper number of processing layers (Shrestha and Mahmood, 2019) so that their problem-solving capacity is increased, and have higher extraction capabilities (Wang et al., 2019a). Famous examples of the use of DL are face recognition (Jamil et al., 2001) and autonomous driving (Kocić et al., 2019). This increased complexity requires huge amounts of data which implies a high computational time and a hard, sometimes impossible, interpretability (Chen and Meng, 2020).

One example is the Recurrent Neural Network (RNN). These structures, as opposed to FFNN, are characterized by having at least one loop (Chang and Lu, 2020; Shrestha and Mahmood, 2019), for instance, by making that each hidden cell receives its output with a fixed delay (Tch, 2017), vide Fig. B.5. If furthermore, the hidden layer is upgraded by including "memory" cells, we encounter the long–short term memory (LSTM) RNN.

An issue of FFNNs is their time-consuming learning process due to slow training algorithms applied several times (Ding et al., 2015). This problem was solved for single-hidden layers by Huang et al. with the Extreme Learning Machine (ELM) (Bin Huang et al., 2006). The algorithm reduces the computational burden of FFNN by selecting the hidden nodes randomly and calculating the output weights only once.

Another DL structure very popular especially when dealing with images is Convolutional Neural Network (CNN). Its name originates from the use of the mathematical operation of convolution. CNNs are made of several layers, including convolutional, pooling and fully connected ones.

Fuzzy algorithms

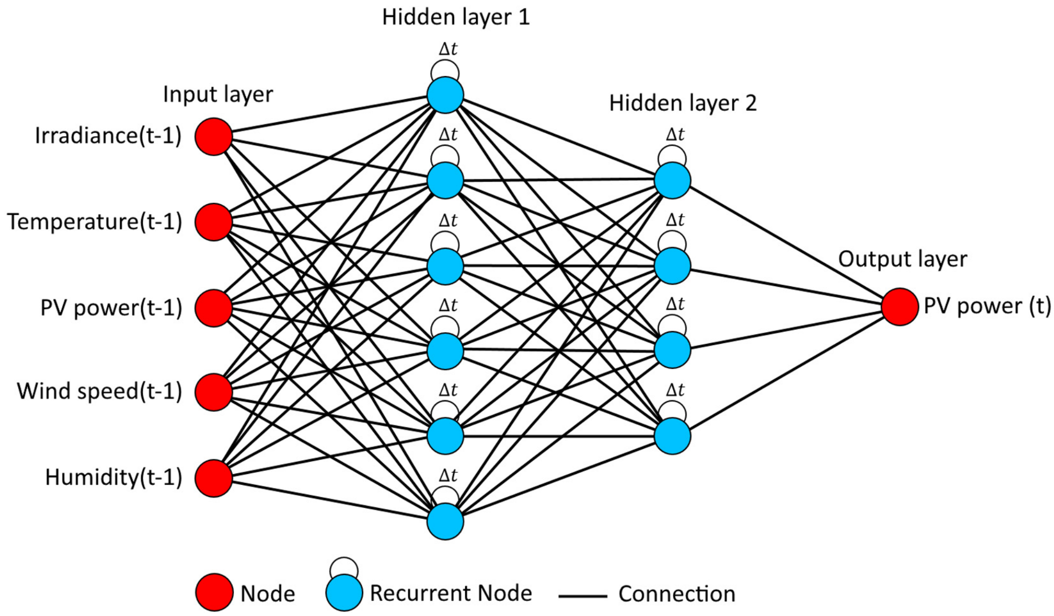


Fig. B.5. Recurrent deep learning neural network structure. For colour references refer to the web version of this article.

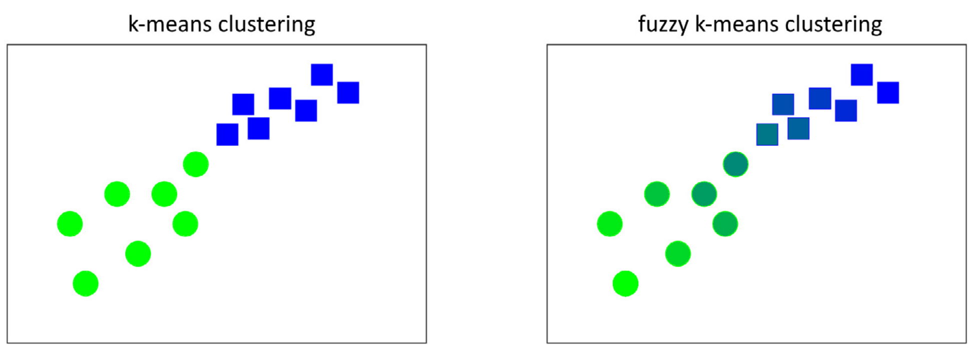


Fig. B.6. Comparison between k-means and fuzzy k-means clustering, where the association to a group is defined by a probability rather than a certainty. For colour references refer to the web version of this article.

The algorithms in this category employ fuzzy logic methods to forecast the PV power. Fuzzy logic can be understood as an extension of binary logic. In the latter, the only possible values are 0/False or 1/True. Fuzzy code extends this logic by considering that the components take values between 0 and 1, which can be interpreted as a probability of a fact occurring. By making use of this logic, Takagi and Sugeno developed a model (Takagi and Sugeno, 1985) which has been employed to modify several ML algorithms. Two algorithms belonging to this group have been applied for PV power forecasting: fuzzy k-means clustering and Adaptive Neuro Fuzzy Inference System.

k-means clustering is a highly employed unsupervised ML algorithm that divides the space into k groups or clusters where each element of the space has an assigned group. When fuzzy logic is applied to this method, the elements can belong simultaneously to more than one cluster up to some degree (Theodoridis and Koutroumbas, 2009), as represented in Fig. B.6. Fuzzy logic can also be integrated into ANN yielding to the Adaptive Neuro Fuzzy Inference System (Jang, 1993; Konstantinou et al., 2021). Because of the use of fuzzy logic, this method is more flexible than standard ANN.

Hybrid algorithms

Hybrid algorithms combine two or more ML methods to improve their performance, and sometimes to reduce the extensive data and computational requirements (Massaoudi et al., 2021). These algorithms can be combined by concatenating two forecasting algorithms, optimizing the parameters of the forecasting model, or modifying the input data so it is more easily interpretable. The third case is not considered a hybrid in this work, but a pre-processing step.

Appendix C. Metrics

Metrics are employed to measure how close predicted values are to real ones. The most basic metric is the error, defined as the difference between real and forecasted values. This measure can be made unit free through normalization by the nominal capacity of the power plant, the real value (percentage error) or the average of the real values, among others.

Time series are not practical to work with, since they do not provide a single value and hence are hard to compare. Some transformations can be applied to this error to express it as one figure, such as:

- Averaging, giving place to the Mean Bias Error (MBE) or simply Bias
- Taking the absolute and averaging, Mean Absolute Error (MAE)
- Squaring and averaging, Mean Squared Error (MSE)
- Squaring, averaging, and taking the square root, Root Mean Squared Error (RMSE)

Table C.1
Most employed metrics for PV power forecasting with their equations and main characteristics.

Metric	Equation	Main characteristics
Error	$x_i - \hat{x}_i$	Time series
NMBE	$\frac{1}{\bar{x}} \frac{\sum_i (\hat{x}_i - x_i)}{N}$	Measure of central tendency Indicated over and under estimations
NMAE	$\frac{1}{\bar{x}} \frac{\sum_i \lvert x_i - \hat{x}_i \rvert}{N}$	Measure of average error Same importance to all values
NMSE	$\frac{1}{\bar{x}^2} \frac{\sum_i (\hat{x}_i - x_i)^2}{N}$	Measure of average of the squares of the errors Higher importance to outliers
NRMSE	$\frac{1}{\bar{x}} \sqrt{\frac{\sum_i (x_i - \hat{x}_i)^2}{N}}$	Measure of average error Higher importance to outliers
MAPE	$\frac{1}{N} \sum_i \left\lvert \frac{x_i - \hat{x}_i}{x_i} \right\rvert$	Measure of prediction accuracy as a percentage of the error
R^2	$1 - \frac{\sum_i (\hat{x}_i - x_i)^2}{\sum_i (\bar{x} - x_i)^2}$	Measure of variability Represents how the model fits the observed data
Skill score	$1 - \frac{RMSE_{proposed}}{RMSE_{reference}}$	Represents an improvement over the reference model (persistence, usually)
Correlation	$\frac{\sum_i (x_i - \bar{x}) \left(\hat{x}_i - \bar{\hat{x}}\right)}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i \left(\hat{x}_i - \bar{\hat{x}}\right)^2}}$	Measure of linear relationship between variables
CRPS	$\frac{1}{N} \sum_i \int_{-\infty}^{\infty} \left(F_i^f(x) - F_i^0(x)\right)^2 dx$	Measure of the difference between the real value and the prediction interval
Kurtosis	$\frac{1}{N} \frac{\sum_i \left(\hat{x}_i - \bar{\hat{x}}\right)^4}{\sigma^4}$	Measure of asymmetry of the distribution
Skewness	$\frac{1}{N} \frac{\sum_i \left(\hat{x}_i - \bar{\hat{x}}\right)^3}{\sigma^3}$	Measure of sharpness of the distribution

All these metrics can be further normalized, giving Normalized Mean Bias Error (NMBE) and equivalent measures (NMAE, NMSE and NRMSE). When the MAE is normalized by the real value of the data series, it gives the Mean Absolute Percentage Error (MAPE), which is a special case of NMAE. Normalization gives unit-free errors and allows for comparison of the performance of PV power plants of different nominal power.

Each of these metrics gives different information about the type of error performed. MBE indicates over- and underestimations, but positive errors are cancelled by negative ones. MAE describes the quantitative error of the forecasts, giving the same importance to all values (Rosiek et al., 2018). This last fact is sometimes undesirable in real applications of PV power since larger errors are the most costly. MSE and RMSE solve this issue at a cost of giving higher importance to outliers. Between the two, RMSE is preferred since the scale remains intact.

Other measures that are not so straightforward to derive but also extensively employed for PV power forecasting are the coefficient of determination or R-squared (R^2), the skill score (SS) and the correlation. R^2 indicates the variability between real and predicted values; it represents how well future outcomes are likely to be predicted (Rosiek et al., 2018). The skill score compares the error of the proposed method with reference to the error of a standard method (such as persistence). Correlation determines the forecast performance even if any systematic correction or re-scaling has occurred.

There is a separate group of metrics which are specific to probabilistic PV power forecasting. These measure the error made on prediction intervals, not on single points. The most employed metric in this category is the Continuous Ranked Probability Score (CRPS), which computes the difference between the real value and the prediction interval (Carney and Cunningham, 2006). It combines deterministic outcomes with probabilistic predictions, and it is equivalent to MAE for deterministic forecasts. An alternative to CRPS is the measure of kurtosis and skewness, which are usually computed together. They represent how different the shape of the prediction is with respect to a normal distribution. While the skewness measures the asymmetry of the distribution, the kurtosis computes its sharpness (Theodoridis and Koutroumbas, 2009).

Table C.1 shows each of the mentioned metrics with their equation and main characteristics. Here, x_i are the measured values, $\hat{x}_i$ are the predicted values, $\bar{x}$ is the average true prediction value, N represents the number of samples, $F_i^f(x)$ is the cumulative distribution function (CDF) of the ith probabilistic forecast, $F_i^0(x)$ is the CDF of the ith observation and σ is the variance of the prediction distribution.

References

Abdel-Nasser, M., Mahmoud, K., 2019. Accurate photovoltaic power forecasting models using deep LSTM-RNN. Neural Comput. Appl. 31 (7), 2727–2740. http://dx.doi.org/10.1007/s00521-017-3225-z.

Abuella, M., Chowdhury, B., 2017. Solar power forecasting using support vector regression.

Agoua, G., Girard, R., Kariniotakis, G., 2019. Probabilistic models for spatio-temporal. IEEE Trans. Sustain. Energy 10 (2), 780–789.

Ahmed, R., Sreeram, V., Mishra, Y., Arif, M.D., 2020. A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. Renew. Sustain. Energy Rev. 124, 109792. http://dx.doi.org/10.1016/j.rser.2020.109792, Elsevier Ltd.

Al-Dahidi, S., Louzazni, M., Omran, N., 2020. A local training strategy-based artificial neural network for predicting the power production of solar photovoltaic systems. IEEE Access 8, 150262–150281. http://dx.doi.org/10.1109/ACCESS.2020.3016165.

Alessandrini, S., Delle Monache, L., Sperati, S., Cervone, G., 2015. An analog ensemble for short-term probabilistic solar power forecast. Appl. Energy 157, 95–110. http://dx.doi.org/10.1016/j.apenergy.2015.08.011.

Alfadda, A., Adhikari, R., Kuzlu, M., Rahman, S., 2017. Hour-ahead solar PV power forecasting using SVR based approach. In: 2017 IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. ISGT 2017. http://dx.doi.org/10.1109/ISGT.2017.8086020.

Alkandari, M., Ahmad, I., 2020. Solar power generation forecasting using ensemble approach based on deep learning and statistical methods. Appl. Comput. Inform. http://dx.doi.org/10.1016/j.aci.

Almeida, M.P., Perpiñán, O., Narvarte, L., 2015. PV power forecast using a nonparametric PV model. Sol. Energy 115, 354–368. http://dx.doi.org/10.1016/j.solener.2015.03.006.

Alomari, M.H., Adeeb, J., Younis, O., 2019. PVPF tool: An automated web application for real-time photovoltaic power forecasting. Int. J. Electr. Comput. Eng. 9 (1), 34–41. http://dx.doi.org/10.11591/ijece.v9i1.pp34-41.

Anagnostos, D., Schmidt, T., Cavadias, S., Soudris, D., Poortmans, J., Catthoor, F., 2019. A method for detailed, short-term energy yield forecasting of photovoltaic installations. Renew. Energy 130, 122–129. http://dx.doi.org/10.1016/j.renene.2018.06.058.

Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., Martinez-de Pison, F.J., Antonanzas-Torres, F., 2016. Review of photovoltaic power forecasting. Sol. Energy 136, 78–111. http://dx.doi.org/10.1016/j.solener.2016.06.069, Elsevier Ltd.

Antonanzas, J., Perpinan-Lamigueiro, O., Urraca, R., Antonanzas-Torres, F., 2020. Influence of electricity market structures on deterministic solar forecasting verification. Sol. Energy 210, 44–46. http://dx.doi.org/10.1016/J.SOLENER.2020.04.017.

Armstrong, J.S., Collopy, F., 1992. Error measures for generalizing about forecasting methods: Empirical comparisons. Int. J. Forecast. 8 (1), 69–80. http://dx.doi.org/10.1016/0169-2070(92)90008-W.

Asrari, A., Wu, T.X., Ramos, B., 2017. A hybrid algorithm for short-term solar power prediction - sunshine state case study. IEEE Trans. Sustain. Energy 8 (2), 582–591. http://dx.doi.org/10.1109/TSTE.2016.2613962.

Baharin, K.A., Abdul Rahman, H., Hassan, M.Y., Gan, C.K., 2016. Short-term forecasting of solar photovoltaic output power for tropical climate using ground-based measurement data. J. Renew. Sustain. Energy 8 (5), 053701. http://dx.doi.org/10.1063/1.4962412.

Barbieri, F., Rajakaruna, S., Ghosh, A., 2017. Very short-term photovoltaic power forecasting with cloud modeling: A review. Renew. Sustain. Energy Rev. 75, 242–263. http://dx.doi.org/10.1016/j.rser.2016.10.068, Elsevier Ltd.

Bin Huang, G., Zhu, Q.Y., Siew, C.K., 2006. Extreme learning machine: Theory and applications. Neurocomputing 70 (1–3), 489–501. http://dx.doi.org/10.1016/j.neucom.2005.12.126.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer.

Brownlee, Jason, 2021. Machine learning mastery. https://machinelearningmastery.com/ (accessed Aug. 30, 2021).

Carney, M., Cunningham, P., 2006. Evaluating Density Forecasting Models. Dep. Comput. Sci., Trinity Coll. Dublin.

Carrera, B., Kim, K., 2020. Comparison analysis of machine learning techniques for photovoltaic prediction using weather sensor data. Sensors 20 (3129), http://dx.doi.org/10.3390/s20113129.

Cervone, G., Clemente-Harding, L., Alessandrini, S., Delle Monache, L., 2017. Short-term photovoltaic power forecasting using Artificial Neural Networks and an Analog Ensemble. Renew. Energy 108, 274–286. http://dx.doi.org/10.1016/j.renene.2017.02.052.

Chang, G.W., Lu, H.J., 2020. Integrating gray data preprocessor and deep belief network for day-ahead PV power output forecast. IEEE Trans. Sustain. Energy 11 (1), 185–194. http://dx.doi.org/10.1109/TSTE.2018.2888548.

Chen, D., Chen, H.W., 2013. Using the Köppen classification to quantify climate variation and change: An example for 1901–2010. Environ. Dev. 6 (1), 69–79. http://dx.doi.org/10.1016/j.envdev.2013.03.007.

Chen, Z., Koprinska, I., 2020. Ensemble Methods for Solar Power Forecasting.

Chen, K., Meng, X., 2020. Interpretation and understanding in machine learning. J. Comput. Res. Dev. 57 (9), 1971. http://dx.doi.org/10.7544/ISSN1000-1239.2020.20190456.

Chu, Y., Urquhart, B., Gohari, S.M.I., Pedro, H.T.C., Kleissl, J., Coimbra, C.F.M., 2015. Short-term reforecasting of power output from a 48 MWe solar PV plant. Sol. Energy 112, 68–77. http://dx.doi.org/10.1016/j.solener.2014.11.017.

Clarivate Analytics, 2022. Web of science. https://www.webofscience.com/wos/woscc/basic-search (accessed Aug. 08, 2022).

Cross, S.S., Harrison, R.F., Kennedy, R.L., 1995. Introduction to neural networks. Lancet 346, 1075–1079. http://dx.doi.org/10.1016/S0140-6736(95)91746-2.

Das, U.K., et al., 2018. Forecasting of photovoltaic power generation and model optimization: A review. Renew. Sustain. Energy Rev. 81, 912–928. http://dx.doi.org/10.1016/j.rser.2017.08.017, Elsevier Ltd.

Das others, U.K., 2017. SVR-based model to forecast PV power generation under different weather conditions. Energies 10 (7), 876. http://dx.doi.org/10.3390/EN10070876.

De Felice, M., Petitta, M., Ruti, P.M., 2015. Short-term predictability of photovoltaic production over Italy. Renew. Energy 80, 197–204. http://dx.doi.org/10.1016/j.renene.2015.02.010.

De Giorgi, M.G., Malvoni, M., Congedo, P.M., 2016. Comparison of strategies for multi-step ahead photovoltaic power forecasting models based on hybrid group method of data handling networks and least square support vector machine. Energy 107, 360–373. http://dx.doi.org/10.1016/j.energy.2016.04.020.

Dev, S., Wen, B., Lee, Y.H., Winkler, S., 2016. Ground-based image analysis: A tutorial on machine-learning techniques and applications. IEEE Geosci. Remote Sens. Mag. 4 (2), 79–93. http://dx.doi.org/10.1109/MGRS.2015.2510448.

Ding, S., Zhao, H., Zhang, Y., Xu, X., Nie, R., 2015. Extreme learning machine: algorithm, theory and applications. Artif. Intell. Rev. 44 (1), 103–115. http://dx.doi.org/10.1007/s10462-013-9405-z.

Do, M.T., Soubdhan, T., Robyns, Benoît, 2016. A study on the minimum duration of training data to provide a high accuracy forecast for PV generation between two different climatic zones. Renew. Energy 85, 959–964. http://dx.doi.org/10.1016/j.renene.2015.07.057.

Dongare, A.D., Kharde, R.R., Kachare, A.D., 2008. Introduction to artificial neural network. Certif. Int. J. Eng. Innov. Technol. 9001 (1), 2277–3754.

Escalante, H.J., 2005. A comparison of outlier detection algorithms for machine learning.

Eseye, A.T., Zhang, J., Zheng, D., 2018. Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and meteorological information. Renew. Energy 118, 357–367. http://dx.doi.org/10.1016/J.RENENE.2017.11.011.

Fausett, L.V., 2006. Fundamentals of Neural Networks: Architectures, Algorithms and Applications. Pearson Education, India.

Feng, C., Liu, Y., Zhang, J., 2021. A taxonomical review on recent artificial intelligence applications to PV integration into power grids. Int. J. Electr. Power Energy Syst. 132, 107176. http://dx.doi.org/10.1016/J.IJEPES.2021.107176.

Ferlito, S., Adinolfi, G., Graditi, G., 2017. Comparative analysis of data-driven methods online and offline trained to the forecasting of grid-connected photovoltaic plant production. Appl. Energy 205, 116–129. http://dx.doi.org/10.1016/j.apenergy.2017.07.124.

Friedman, J.H., 1991. Multivariate adaptive regression splines. Ann. Statist. 19 (1), 1–141.

Fthenakis, V.M., Kim, H.C., Alsema, E., 2008. Emissions from photovoltaic life cycles. Environ. Sci. Technol. 42 (6), 2168–2174. http://dx.doi.org/10.1021/ES071763Q.

Gandoman, F.H., Raeisi, F., Ahmadi, A., 2016. A literature review on estimating of PV-array hourly power under cloudy weather conditions. Renew. Sustain. Energy Rev. 63, 579–592. http://dx.doi.org/10.1016/j.rser.2016.05.027, Elsevier Ltd.

Gardner, M.W., Dorling, S.R., 1998. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmos. Environ. 32 (14–15), 2627–2636. http://dx.doi.org/10.1016/S1352-2310(97)00447-0.

Gensler, A., Henze, J., Sick, B., Raabe, N., 2017. Deep learning for solar power forecasting - An approach using AutoEncoder and LSTM Neural Networks. In: 2016 IEEE Int. Conf. Syst. Man, Cybern. SMC 2016 - Conf. Proc.. pp. 2858–2865. http://dx.doi.org/10.1109/SMC.2016.7844673.

Golestaneh, F., Pinson, P., Gooi, H.B., 2016. Very short-term nonparametric probabilistic forecasting of renewable energy generation - With application to solar energy. IEEE Trans. Power Syst. 31 (5), 3850–3863. http://dx.doi.org/10.1109/TPWRS.2015.2502423.

Gosiewska, A., Kozak, A., Biecek, P., 2021. Simpler is better: Lifting interpretability-performance trade-off via automated feature engineering. Decis. Support Syst. 150, 113556. http://dx.doi.org/10.1016/J.DSS.2021.113556.

Graditi, G., Ferlito, S., Adinolfi, G., 2016. Comparison of photovoltaic plant power production prediction methods using a large measured dataset. Renew. Energy 90, 513–519. http://dx.doi.org/10.1016/j.renene.2016.01.027.

Gulin, M., Pavlović, T., Vašak, M., 2017. A one-day-ahead photovoltaic array power production prediction with combined static and dynamic on-line correction. Sol. Energy 142, 49–60. http://dx.doi.org/10.1016/j.solener.2016.12.008.

Gupta, P., Singh, R., 2021. PV power forecasting based on data-driven models: a review. http://dx.doi.org/10.1080/19397038.2021.1986590.

Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. J. Mach. Learn. Res. 3 (Mar), 1157–1182.

Han, S., hui Qiao, Y., Yan, J., qian Liu, Y., Li, L., Wang, Z., 2019. Mid-to-long term wind and photovoltaic power generation prediction based on copula function and long short term memory network. Appl. Energy 239, 181–191. http://dx.doi.org/10.1016/J.APENERGY.2019.01.193.

Hossain, M., Mekhilef, S., Danesh, M., Olatomiwa, L., Shamshirband, S., 2017. Application of extreme learning machine for short term output power forecasting of three grid-connected PV systems. J. Clean. Prod. 167, 395–405. http://dx.doi.org/10.1016/j.jclepro.2017.08.081.

Huang, C., Bensoussan, A., Edesess, M., Tsui, K.L., 2016. Improvement in artificial neural network-based estimation of grid connected photovoltaic power output. Renew. Energy 97, 838–848. http://dx.doi.org/10.1016/j.renene.2016.06.043.

Huang, C.J., Kuo, P.H., 2019. Multiple-input deep convolutional neural network model for short-term photovoltaic power forecasting. IEEE Access 7, 74822–74834. http://dx.doi.org/10.1109/ACCESS.2019.2921238.

Inman, R.H., Pedro, H.T.C., Coimbra, C.F.M., 2013. Solar forecasting methods for renewable energy integration. Prog. Energy Combust. Sci. 39 (6), 535–576. http://dx.doi.org/10.1016/j.pecs.2013.06.002, Pergamon.

International Energy Agency, 2020. Snapshot of Global PV Markets. Accessed: May 06, 2021. [Online]. Available: www.iea-pvps.org.

International Energy Agency, 2021. Snapshot of Global PV Markets. Accessed: May 06, 2021. [Online]. Available: www.iea-pvps.org.

Isaksson, E., Conde, M.K., 2018. Solar Power Forecasting with Machine Learning Techniques. KTH Royal Institute of Technology.

Jamil, N., Lqbal, S., Iqbal, N., 2001. Face recognition using neural networks. In: Proceedings - IEEE Lnternational Multi Topic Conference 2001: Technology for the 21st Century, IEEE INMIC 2001. pp. 277–281. http://dx.doi.org/10.1109/INMIC.2001.995351.

Jang, J.S.R., 1993. ANFIS: Adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man Cybern. 23 (3), 665–685. http://dx.doi.org/10.1109/21.256541.

Jung, Y., Jung, J., Kim, B., Han, S.U., 2020. Long short-term memory recurrent neural network for modeling temporal patterns in long-term power forecasting for solar PV facilities: Case study of South Korea. J. Clean. Prod. 250, 119476. http://dx.doi.org/10.1016/J.JCLEPRO.2019.119476.

Kawabe, K., Tanaka, K., 2015. Impact of dynamic behavior of photovoltaic power generation systems on short-term voltage stability. IEEE Trans. Power Syst. 30 (6), 3416–3424. http://dx.doi.org/10.1109/TPWRS.2015.2390649.

Kazem, H.A., Yousif, J.H., 2017. Comparison of prediction methods of photovoltaic power system production using a measured dataset. Energy Convers. Manage. 148, 1070–1081. http://dx.doi.org/10.1016/j.enconman.2017.06.058.

Kocić, J., Jovičić, N., Drndarević, V., 2019. An end-to-end deep neural network for autonomous driving designed for embedded automotive platforms. Sensors (Switzerland) 19 (9), http://dx.doi.org/10.3390/s19092064.

Konstantinou, M., Peratikou, S., Charalambides, A.G., 2021. Solar photovoltaic forecasting of power output using LSTM networks. Atmos. 12 (1), 124. http://dx.doi.org/10.3390/ATMOS12010124.

Kraemer, F.A., Palma, D., Braten, A.E., Ammar, D., 2020. Operationalizing solar energy predictions for sustainable, autonomous IoT device management. IEEE Internet Things J. 7 (12), http://dx.doi.org/10.1109/JIOT.2020.3002330.

Kumar, K.R., Kalavathi, M.S., 2018. Artificial intelligence based forecast models for predicting solar power generation. Mater. Today Proc. 5 (1), 796–802. http://dx.doi.org/10.1016/J.MATPR.2017.11.149.

Kuzmiakova, A., Colas, G., McKeehan, A., 2017. Short-Term Memory Solar Energy Forecasting at University of Illinois. Univ. Illinois, pp. 1–6, no. December.

Lee, D., Jeong, J., Yoon, S.H., Chae, Y.T., 2019. Improvement of short-term BIPV power predictions using feature engineering and a recurrent neural network. Energies 12 (17), 3247. http://dx.doi.org/10.3390/EN12173247.

Lee, D., Kim, K., 2019. Recurrent neural network-based hourly prediction of photovoltaic power output using meteorological information. Energies 12 (2), 215. http://dx.doi.org/10.3390/EN12020215.

Lee, W., Kim, K., Park, J., Kim, J., Kim, Y., 2018. Forecasting solar power using long-short term memory and convolutional neural networks. IEEE Access 6, 73068–73080. http://dx.doi.org/10.1109/ACCESS.2018.2883330.

Li, L.L., Cheng, P., Lin, H.C., Dong, H., 2017. Short-term output power forecasting of photovoltaic systems based on the deep belief net. 9 (9) 1–13. http://dx.doi.org/10.1177/1687814017715983.

Li, Y., He, Y., Su, Y., Shu, L., 2016a. Forecasting the daily power output of a grid-connected photovoltaic system based on multivariate adaptive regression splines. Appl. Energy 180, 392–401. http://dx.doi.org/10.1016/j.apenergy.2016.07.052.

Li, Z., Mahbobur Rahman, S.M., Vega, R., Dong, B., 2016b. A hierarchical approach using machine learning methods in solar photovoltaic energy production forecasting. Energies 9 (1), 55. http://dx.doi.org/10.3390/EN9010055.

Li, G., Wang, H., Zhang, S., Xin, J., Liu, H., 2019a. Recurrent neural networks based photovoltaic power forecasting approach. Energies 12 (13), 2538. http://dx.doi.org/10.3390/EN12132538.

Li, L.L., Wen, S.Y., Tseng, M.L., Wang, C.S., 2019b. Renewable energy prediction: A novel short-term prediction model of photovoltaic output power. J. Clean. Prod. 228, 359–375. http://dx.doi.org/10.1016/J.JCLEPRO.2019.04.331.

Li, Q., Wu, Z., Xia, X., 2018. Estimate and characterize PV power at demand-side hybrid system. Appl. Energy 218, 66–77. http://dx.doi.org/10.1016/j.apenergy.2018.02.160.

Li, Z., Zang, C., Zeng, P., Yu, H., Li, H., 2015. Day-ahead hourly photovoltaic generation forecasting using extreme learning machine. In: 2015 IEEE Int. Conf. Cyber Technol. Autom. Control Intell. Syst. IEEE-CYBER 2015. pp. 779–783. http://dx.doi.org/10.1109/CYBER.2015.7288041.

Li, P., Zhou, K., Lu, X., Yang, S., 2020. A hybrid deep learning model for short-term PV power forecasting. Appl. Energy 259, 114216. http://dx.doi.org/10.1016/J.APENERGY.2019.114216.

Lin, Y., Koprinska, I., Rana, M., Troncoso, A., 2020. Solar power forecasting based on pattern sequence similarity and meta-learning. In: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). In: LNCS, vol. 12396, pp. 271–283. http://dx.doi.org/10.1007/978-3-030-61609-0_22/TABLES/5.

Lin, K.P., Pai, P.F., 2016. Solar power output forecasting using evolutionary seasonal decomposition least-square support vector regression. J. Clean. Prod. 134 (Part B), 456–462. http://dx.doi.org/10.1016/j.jclepro.2015.08.099.

Lin, P., Peng, Z., Lai, Y., Cheng, S., Chen, Z., Wu, L., 2018. Short-term power prediction for photovoltaic power plants using a hybrid improved Kmeans-GRA-Elman model based on multivariate meteorological factors and historical power datasets. Energy Convers. Manage. 177, 704–717. http://dx.doi.org/10.1016/j.enconman.2018.10.015.

Loh, W.Y., 2011. Classification and regression trees. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 1 (1), 14–23. http://dx.doi.org/10.1002/widm.8.

Maciejowska, K., Nitka, W., Weron, T., 2019. Day-Ahead vs. Intraday—Forecasting the price spread to maximize economic benefits. Energies 12 (4), 631. http://dx.doi.org/10.3390/en12040631.

Maitanova, N., et al., 2019. Machine learning approach to a low-cost day-ahead photovoltaic power prediction based on publicly available weather reports for automated energy management systems. In: Proc. 36th Eur. Photovolt. Sol. Energy Conf. Exhib., no. September. pp. 1711–1718.

Majumder, I., Dash, P.K., Bisoi, R., 2020. Short-term solar power prediction using multi-kernel-based random vector functional link with water cycle algorithm-based parameter optimization. Neural Comput. Appl. 32, http://dx.doi.org/10.1007/s00521-019-04290-x.

Massaoudi, M., Chihi, I., Abu-Rub, H., Refaat, S.S., Oueslati, F.S., 2021. Convergence of photovoltaic power forecasting and deep learning: State-of-art review. IEEE Access 9, 136593–136615. http://dx.doi.org/10.1109/ACCESS.2021.3117004.

Massaoudi, M., Chihi, I., Sidhom, L., Trabelsi, M., Refaat, S.S., Oueslati, F.S., 2019. PV power forecasting using weighted features for enhanced ensemble method.

Massidda, L., Marrocu, M., 2017. Use of multilinear adaptive regression splines and numerical weather prediction to forecast the power output of a PV plant in Borkum, Germany. Sol. Energy 146, 141–149. http://dx.doi.org/10.1016/j.solener.2017.02.007.

Mayer, M.J., Gróf, G., 2020. Extensive comparison of physical models for photovoltaic power forecasting. Appl. Energy 116239. http://dx.doi.org/10.1016/j.apenergy.2020.116239.

Mellit, A., Pavan, A.M., Ogliari, E., Leva, S., Lughi, V., 2020. Advanced methods for photovoltaic output power forecasting: A review. Appl. Sci. 10.2 (487), http://dx.doi.org/10.3390/app10020487.

Mishra, M., Byomakesha Dash, P., Nayak, J., Naik, B., Kumar Swain, S., 2020. Deep learning and wavelet transform integrated approach for short-term solar PV power prediction. Measurement 166, 108250. http://dx.doi.org/10.1016/J.MEASUREMENT.2020.108250.

Mosavi, A., Salimi, M., Ardabili, S.F., Rabczuk, T., Shamshirband, S., Varkonyi-Koczy, A.R., 2019. State of the art of machine learning models in energy systems, a systematic review. Energies 12 (7), 1301. http://dx.doi.org/10.3390/en12071301, MDPI AG.

Nageem, R., Jayabarathi, R., 2017. Predicting the power output of a grid-connected solar panel using multi-input support vector regression. Procedia Comput. Sci. 115, 723–730. http://dx.doi.org/10.1016/J.PROCS.2017.09.143.

Nayak, A., Heistrene, L., 2020. Hybrid machine learning model for forecasting solar power generation; hybrid machine learning model for forecasting solar power generation. In: 2020 Int. Conf. Smart Grids Energy Syst.. http://dx.doi.org/10.1109/SGES51519.2020.00167.

Ni, Q., Zhuang, S., Sheng, H., Kang, G., Xiao, J., 2017. An ensemble prediction intervals approach for short-term PV power forecasting. Sol. Energy 155, 1072–1083. http://dx.doi.org/10.1016/j.solener.2017.07.052.

Notton, G., et al., 2018. Intermittent and stochastic character of renewable energy sources: Consequences, cost of intermittence and benefit of forecasting. Renew. Sustain. Energy Rev. 87, 96–105. http://dx.doi.org/10.1016/j.rser.2018.02.007, Elsevier Ltd.

Ogliari, E., Niccolai, A., Leva, S., Zich, R., 2018. Computational intelligence techniques applied to the day ahead PV output power forecast: PHANN, SNO and mixed. Energies 11 (6), 1487. http://dx.doi.org/10.3390/en11061487.

Ortner, A., Totschnig, G., 2019. The future relevance of electricity balancing markets in Europe - A 2030 case study. Energy Strateg. Rev. 24, 111–120. http://dx.doi.org/10.1016/j.esr.2019.01.003.

Pan, M., et al., 2020. Photovoltaic power forecasting based on a support vector machine with improved ant colony optimization. J. Clean. Prod. 277, 123948. http://dx.doi.org/10.1016/J.JCLEPRO.2020.123948.

Panamtash, H., Mahdavi, S., Zhou, Q., 2021. Probabilistic solar power forecasting: A review and comparison. In: 2020 52nd North Am. Power Symp. NAPS 2020. http://dx.doi.org/10.1109/NAPS50074.2021.9449746.

Paulescu, M., Brabec, M., Boata, R., Badescu, V., 2017. Structured, physically inspired (gray box) models versus black box modeling for forecasting the output power of photovoltaic plants. Energy 121, 792–802. http://dx.doi.org/10.1016/j.energy.2017.01.015.

Pawar, P., Mithulananthan, N., Raza, M.Q., 2020. Solar PV power forecasting using modified SVR with Gauss–Newton method. In: Proc. - 2020 IEEE 2nd Glob. Power, Energy Commun. Conf. GPECOM 2020. pp. 226–231. http://dx.doi.org/10.1109/GPECOM49333.2020.9247935.

Pazikadin, A.R., Rifai, D., Ali, K., Malik, M.Z., Abdalla, A.N., Faraj, M.A., 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. http://dx.doi.org/10.1016/J.SCITOTENV.2020.136848.

Pierro, M., et al., 2017. Data-driven upscaling methods for regional photovoltaic power estimation and forecast using satellite and numerical weather prediction data. Sol. Energy 158, 1026–1038. http://dx.doi.org/10.1016/j.solener.2017.09.068.

Poplavskaya, K., Totschnig, G., Leimgruber, F., Doorman, G., Etienne, G., de Vries, L., 2020. Integration of day-ahead market and redispatch to increase cross-border exchanges in the European electricity market. Appl. Energy 278, 115669. http://dx.doi.org/10.1016/j.apenergy.2020.115669.

Pujić, D., Jelić, M., Tomašević, N., 2020. Comparison between different ML approaches for PV and STC production forecasting using real world data. In: Proc. ICIST.

Rajagukguk, R.A., Ramadhan, R.A.A., Lee, H.-J., 2020. A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power. Energies 13 (24), 6623. http://dx.doi.org/10.3390/EN13246623.

Ramsami, P., Oree, V., 2015. A hybrid method for forecasting the energy output of photovoltaic systems. Energy Convers. Manage. 95, 406–413. http://dx.doi.org/10.1016/j.enconman.2015.02.052.

Rana, M., Koprinska, I., Agelidis, V.G., 2015. 2D-interval forecasts for solar power production. Sol. Energy 122, 191–203. http://dx.doi.org/10.1016/j.solener.2015.08.018.

Ratanamahatana, C., Gunopulos, D., 2010. Feature selection for the naive bayesian classifier using decision trees. 17 (5–6) 475–487. http://dx.doi.org/10.1080/713827175.

Raza, M.Q., Mithulananthan, N., Summerfield, A., 2018. Solar output power forecast using an ensemble framework with neural predictors and Bayesian adaptive combination. Sol. Energy 166, 226–241. http://dx.doi.org/10.1016/j.solener.2018.03.066.

Raza, M.Q., Nadarajah, M., Ekanayake, C., 2016. On recent advances in PV output power forecast. Sol. Energy 136, 125–144. http://dx.doi.org/10.1016/j.solener.2016.06.073, Elsevier Ltd.

Ren, Y., Suganthan, P.N., Srikanth, N., 2015. Ensemble methods for wind and solar power forecasting - A state-of-the-art review. Renew. Sustain. Energy Rev. 50, 82–91. http://dx.doi.org/10.1016/j.rser.2015.04.081, Elsevier Ltd.

Rosato, A., Altilio, R., Araneo, R., Panella, M., 2017. Embedding of time series for the prediction in photovoltaic power plants. Energies 10 (1003), http://dx.doi.org/10.3390/en10071003.

Rosiek, S., Alonso-Montesinos, J., Batlles, F.J., 2018. Online 3-h forecasting of the power output from a BIPV system using satellite observations and ANN. Int. J. Electr. Power Energy Syst. 99, 261–272. http://dx.doi.org/10.1016/j.ijepes.2018.01.025.

Rudin, C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat. Mach. Intell. 1 (5), 206–215. http://dx.doi.org/10.1038/s42256-019-0048-x.

Schulz, E., Speekenbrink, M., Krause, A., 2018. A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. J. Math. Psych. 85, 1–16. http://dx.doi.org/10.1016/J.JMP.2018.03.001.

Semero, Y.K., Zhang, J., Zheng, D., 2018. PV power forecasting using an integrated GA-PSO-ANFIS approach and Gaussian process regression based feature selection strategy. CSEE J. Power Energy Syst. 4 (2), 210–218. http://dx.doi.org/10.17775/CSEEJPES.2016.01920.

Shang, C., Wei, P., 2018. Enhanced support vector regression based forecast engine to predict solar power output. Renew. Energy 127, 269–283. http://dx.doi.org/10.1016/j.renene.2018.04.067.

Sharadga, H., Hajimirza, S., Balog, R.S., 2020. Time series forecasting of solar power generation for large-scale photovoltaic plants. Renew. Energy 150, 797–807. http://dx.doi.org/10.1016/j.renene.2019.12.131.

Sheng, H., Ray, B., Chen, K., Cheng, Y., 2020. Solar power forecasting based on domain adaptive learning. IEEE Access 8, http://dx.doi.org/10.1109/ACCESS.2020.3034100.

Shivashankar, S., Mekhilef, S., Mokhlis, H., Karimi, M., 2016. Mitigating methods of power fluctuation of photovoltaic (PV) sources - A review. Renew. Sustain. Energy Rev. 59, 1170–1184. http://dx.doi.org/10.1016/j.rser.2016.01.059, Elsevier Ltd.

Shrestha, A., Mahmood, A., 2019. Review of deep learning algorithms and architectures. IEEE Access 7, 53040–53065. http://dx.doi.org/10.1109/ACCESS.2019.2912200.

Sit, Hilarie, 2019. Quick start to Gaussian process regression. Towards Data Sci. https://towardsdatascience.com/quick-start-to-gaussian-process-regression-36d838810319 (accessed Nov. 27, 2020).

Sobri, S., Koohi-Kamali, S., Rahim, N.A., 2018. Solar photovoltaic generation forecasting methods: A review. Energy Convers. Manage. 156, 459–497. http://dx.doi.org/10.1016/j.enconman.2017.11.019, Elsevier Ltd.

Soufiane, G., Ouafia, F., Ahmed, A., Author, C., 2020. Solar power time series prediction using wavelet analysis. Int. J. Renew. Energy Res. 10 (4).

Sun, Y., Venugopal, V., Brandt, A.R., 2019. Short-term solar power forecast with deep learning: Exploring optimal input and output configuration. Sol. Energy 188, 730–741. http://dx.doi.org/10.1016/j.solener.2019.06.041.

Suresh, V., Janik, P., Rezmer, J., Leonowicz, Z., 2020. Forecasting solar PV output using convolutional neural networks with a sliding window algorithm. Energies 13 (3), 723. http://dx.doi.org/10.3390/EN13030723.

Takagi, T., Sugeno, M., 1985. Fuzzy identification of systems and its applications to modeling and control. IEEE Trans. Syst. Man Cybern. SMC-15 (1), 116–132. http://dx.doi.org/10.1109/TSMC.1985.6313399.

Takeda, H., 2017. Short-term ensemble forecast for purchased photovoltaic generation. Sol. Energy 149, 176–187. http://dx.doi.org/10.1016/j.solener.2017.03.088.

Tang, P., Chen, D., Hou, Y., 2016. Entropy method combined with extreme learning machine method for the short-term photovoltaic power generation forecasting. Chaos Solitons Fractals 89, 243–248. http://dx.doi.org/10.1016/J.CHAOS.2015.11.008.

Tch, A., 2017. The mostly complete chart of neural networks, explained. Towards Data Sci. https://towardsdatascience.com/the-mostly-complete-chart-of-neural-networks-explained-3fb6f2367464 (accessed Sep. 02, 2021).

Theocharides, S., Makrides, G., Georghiou, G.E., Kyprianou, A., 2018. Machine learning algorithms for photovoltaic system power output prediction. In: 2018 IEEE International Energy Conference. ENERGYCON, pp. 1–6.

Theodoridis, S., Koutroumbas, K., 2009. Pattern Recognit..

Torabi, R., Rodrigues, S., Cafofo, N., Morgado-Dias, F., 2017. Development of an ANN model to predict the electricity produced by small scale roof-top PV systems in Madeira Island. In: Energy Sustain. Small Dev. Econ. ES2DE 2017 - Proc.. pp. 2–7. http://dx.doi.org/10.1109/ES2DE.2017.8015344.

Touati, F., et al., 2017. Long-term performance analysis and power prediction of PV technology in the State of Qatar. Renew. Energy 113, 952–965. http://dx.doi.org/10.1016/j.renene.2017.06.078.

Ulbricht, R., Fischer, U., Lehner, W., Donker, H., 2013. First steps towards a systematical optimized strategy for solar energy supply forecasting. In: Proc. Eur. Conf. Mach. Learn. Princ. Pract. Knowl. Discov. Databases (ECMLPKDD 2013), Vol. 2327.

van der Meer, D.W., Widén, J., Munkhammar, J., 2018. Review on probabilistic forecasting of photovoltaic power production and electricity consumption. Renew. Sustain. Energy Rev. 81, 1484–1512. http://dx.doi.org/10.1016/j.rser.2017.05.212, Elsevier Ltd.

Van Tai, D., 2019. Solar photovoltaic power output forecasting using machine learning technique. J. Phys. Conf. Ser. 1327 (1), 012051. http://dx.doi.org/10.1088/1742-6596/1327/1/012051.

Wan, C., Zhao, J., Song, Y., Xu, Z., Lin, J., Hu, Z., 2016. Photovoltaic and solar power forecasting for smart grid energy management. CSEE J. Power Energy Syst. 1 (4), 38–46. http://dx.doi.org/10.17775/cseejpes.2015.00046.

Wang, Sun-Chong, 2003. Interdisciplinary Computing in Java Programming Language. Springer Science+Business Media, New York.

Wang, Y., Liao, W., Chang, Y., 2018. Gated recurrent unit network-based short-term photovoltaic forecasting. Energies 11 (8), 2163. http://dx.doi.org/10.3390/EN11082163.

Wang, K., Qi, X., Liu, H., 2019a. A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network. Appl. Energy 251, 113315. http://dx.doi.org/10.1016/J.APENERGY.2019.113315.

Wang, K., Qi, X., Liu, H., 2019b. Photovoltaic power forecasting based LSTM-convolutional network. Energy 189, 116225. http://dx.doi.org/10.1016/J.ENERGY.2019.116225.

Wang, G., Su, Y., Shu, L., 2016. One-day-ahead daily power forecasting of photovoltaic systems based on partial functional linear regression models. Renew. Energy 96, 469–478. http://dx.doi.org/10.1016/J.RENENE.2016.04.089.

Wang, F., Xuan, Z., Zhen, Z., Li, K., Wang, T., Shi, M., 2020. A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework. Energy Convers. Manage. 212, 112766. http://dx.doi.org/10.1016/J.ENCONMAN.2020.112766.

Wang, H., et al., 2017. Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network. Energy Convers. Manage. 153, 409–422. http://dx.doi.org/10.1016/j.enconman.2017.10.008.

Wang, K., et al., 2019c. Multiple convolutional neural networks for multivariate time series prediction. Neurocomputing 360, 107–119. http://dx.doi.org/10.1016/J.NEUCOM.2019.05.023.

Wen, L., Zhou, K., Yang, S., Lu, X., 2019. Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting. Energy 171, 1053–1065. http://dx.doi.org/10.1016/J.ENERGY.2019.01.075.

Wolff, B., Kühnert, J., Lorenz, E., Kramer, O., Heinemann, D., 2016. Comparing support vector regression for PV power forecasting to a physical modeling approach using measurement, numerical weather prediction, and cloud motion data. Sol. Energy 135, 197–208. http://dx.doi.org/10.1016/j.solener.2016.05.051.

Yadav, A.K., Sharma, V., Malik, H., Chandel, S.S., 2018. Daily array yield prediction of grid-interactive photovoltaic plant using relief attribute evaluator based Radial Basis Function Neural Network. Renew. Sustain. Energy Rev. 81, 2115–2127. http://dx.doi.org/10.1016/j.rser.2017.06.023, Elsevier Ltd.

Yang, D., 2019. A guideline to solar forecasting research practice: Reproducible, operational, probabilistic or physically-based, ensemble, and skill (ROPES). J. Renew. Sustain. Energy 11 (2), 22701. http://dx.doi.org/10.1063/1.5087462, American Institute of Physics Inc.

Yang, Z., Mourshed, M., Liu, K., Xu, X., Feng, S., 2020. A novel competitive swarm optimized RBF neural network model for short-term solar power generation forecasting. Neurocomputing 397, 415–421. http://dx.doi.org/10.1016/J.NEUCOM.2019.09.110.

Yin, W., Han, Y., Zhou, H., Ma, M., Li, L., Zhu, H., 2020. A novel non-iterative correction method for short-term photovoltaic power forecasting. Renew. Energy 159, 23–32. http://dx.doi.org/10.1016/J.RENENE.2020.05.134.

Yousif, J.H., Kazem, H.A., Boland, J., 2017. Predictive models for photovoltaic electricity production in hot weather conditions. Energies 10 (7), 971. http://dx.doi.org/10.3390/EN10070971.

Zang, H., et al., 2018. Hybrid method for short-term photovoltaic power forecasting based on deep convolutional neural network. IET Gener. Transm. Distrib. 12 (20), 4557–4567. http://dx.doi.org/10.1049/iet-gtd.2018.5847.

Zhang, Y., Beaudin, M., Taheri, R., Zareipour, H., Wood, D., 2015. Day-ahead power output forecasting for small-scale solar photovoltaic electricity generators. IEEE Trans. Smart Grid 6 (5), 2253–2262. http://dx.doi.org/10.1109/TSG.2015.2397003.

Zhang, J., Verschae, R., Nobuhara, S., Lalonde, J.F., 2018. Deep photovoltaic nowcasting. Sol. Energy 176, 267–276. http://dx.doi.org/10.1016/j.solener.2018.10.024.

Zhang others, J., 2015. A suite of metrics for assessing the performance of solar power forecasting. Sol. Energy 111, 157–175. http://dx.doi.org/10.1016/j.solener.2014.10.016.

Zhou, H., Zhang, Y., Yang, L., Liu, Q., 2018. Short-term photovoltaic power forecasting based on stacking-SVM. In: Proc. - 9th Int. Conf. Inf. Technol. Med. Educ. ITME 2018. pp. 994–998. http://dx.doi.org/10.1109/ITME.2018.00221.