

Received December 15, 2021, accepted February 20, 2022, date of publication March 4, 2022, date of current version March 11, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3156942

A Review of Machine Learning-Based Photovoltaic Output Power Forecasting: Nordic Context

BERHANE DARSENE DIMD¹, (Graduate Student Member, IEEE), **STEVE VÖLLER**,
UMIT CALI¹, (Member, IEEE), AND **OLE-MORTEN MIDTGÅRD**¹, (Member, IEEE)

Department of Electric Power Engineering, Norwegian University of Science and Technology, 7491 Trondheim, Norway

Corresponding author: Berhane Darsene Dimd (berhane.d.dimd@ntnu.no)

This work was supported by the Department of Electric Power Engineering, Norwegian University of Science and Technology.

ABSTRACT Motivated by factors such as the reduction in cost and the need for a shift towards achieving UN's Sustainable Development Goals, PV (Photovoltaic) power generation is getting more attention in the cold regions of the Nordic countries and Canada. The cold climate and the albedo effect of snow in these regions present favorable operating conditions for PV cells and an opportunity to realize the seasonal matching of generation and consumption respectively. However, the erratic nature of PV brings a threat to the operation of the grid. PV power forecasting has been used as an economical solution to minimize and even overcome this limitation. This paper is therefore a comprehensive review of machine learning-based PV output power forecasting models in the literature in the context of Nordic climate. The impact of meteorological parameters and the soiling effect due to snow, which is unique to this climate, on the performance of a prediction model is discussed. PV power forecasting models in the literature are systematically classified into multiple groups and each group is analyzed and important suggestions are made for choosing a better model for these regions. Ensemble methods, optimization algorithms, time-series decomposition, and weather clustering are identified as important techniques that can be used to enhance performance. And notably, this work proposed two conceptual approaches that can be used to incorporate the effect of snow on PV power forecasting. Future research needs to focus on this area, which is crucial for the development of PV in these regions.

INDEX TERMS Artificial intelligence, cold climate, machine learning, Nordic climate, PV power forecasting.

I. INTRODUCTION

The demand for electrical energy is increasing very significantly because of the continued increase in the world population. In the International Energy Outlook 2019 (IEO2019) Reference case release, the US Energy Information Administration (EIA) projected that world energy usage will grow by nearly 50% between 2018 and 2050 [1]. This increased demand requires more energy generation from coal and other fossil fuels, which not only produce a significant amount of pollution but also ultimately result in the depletion of the limited resources. In addition, more generation in the conventional ways requires up-scaling of the existing

electric grid to accommodate this huge change, which is economically demanding and technically inefficient.

This necessitates the need to look for sustainable alternative measures which can go in line with the UN's Sustainable Development Goals (SDG 7: Affordable and Clean Energy). Motivated by this and the continued advancement in PV (Photo Voltaic) and wind power technology, large wind and PV power plants have been deployed throughout the world. The reduced cost of these technologies is also a huge contributing factor. According to The Sustainable Development Goals Report 2020, the share of renewable energy in total final energy consumption reached 17.3% in 2017, up from 17.0% in 2015 and 16.3% in 2010 [2]. The largest increase in the use of renewables has come from the electricity sector, driven by the rapid expansion of solar and wind power. Unfortunately, these efforts are not on the

The associate editor coordinating the review of this manuscript and approving it for publication was Shunfeng Cheng.

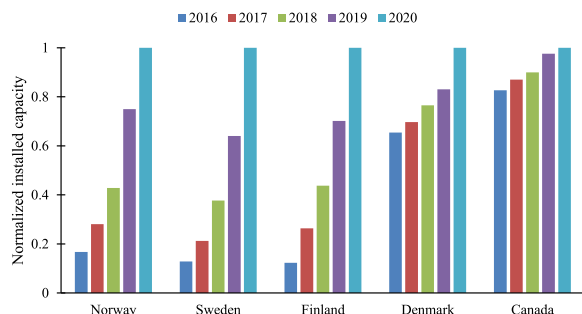


FIGURE 1. PV installation trend in Nordic region and Canada.

scale and more has to be done if it is required to achieve Goal 7 by 2030.

Wind and solar energy have become mainstream electricity sources and are increasingly cost-competitive with fossil fuel power plants. Nearly everywhere in the world, producing electricity from new renewables is more cost-effective than producing it from new coal-fired power plants [3]. Although several key countries and regions, such as China, South and Central Europe, and the United States, have driven these trends and continued to have a large impact in 2019, renewable power, especially PV is growing in all corners of the world.

Even in colder and high latitude climate regions such as the Nordic countries and Canada, solar electricity is gaining huge interest in recent years. Although the focus of this paper is on Nordic regions (especially Norway), Canada is included here since it has relatively similar climatic conditions. According to the data from the statistics bureau of the respective countries and Our World in Data,¹ the PV installation trend in these regions for the last 5 years is shown in Fig 1. PV installed capacity grew by 33%, 56%, 20%, 21%, and 2.5% for Norway, Sweden, Denmark, Finland, and Canada respectively in 2020 alone when compared with the previous year. This can possibly be attributed to the reduced price of PV modules and favorable operating conditions for the solar cells due to the cold temperature. These countries have a combined installed PV power capacity of more than 5800 MW as of 2020 and with the ongoing promising steps of the government policy in these regions, solar electricity will grow even at a more significant rate.

Despite the hydro and wind power dominated power system, solar electricity is gaining momentum in Norway. Solar energy production in Norway is growing rapidly (currently 0.1 TWh annually) and it will be an important component in the power system. Solar electricity has become more attractive both in new zero-emission housing projects and for homeowners who wish to be wholly or partly self-sustained with electricity. According to the Norwegian Water Resources and Energy Directorate (NVE), a total of 40 MW new solar power was installed in Norway during the year 2020 alone [4]. By 2030, the annual energy generation from PV is estimated to be between 1.2 TWh [5] and 4 TWh [6].

¹<https://ourworldindata.org/>

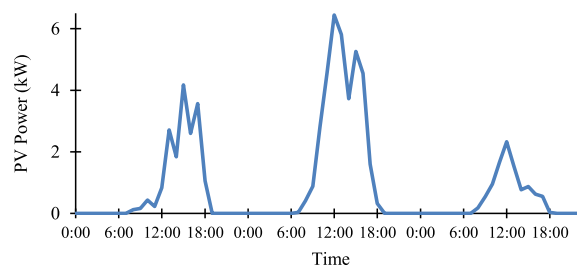


FIGURE 2. Short term PV output variation.

Solar electricity is mainly a distributed resource in Norway, but recent projects are looking more at the use of solar in a larger scale, especially in connection with hydro power reservoirs and larger office buildings. The transition to renewable energy (PV and wind power) however is not smooth.

Solar electricity has significant limitations since it is time reliant and has stochastic nature. This is the result of the high dependence of PV output power on meteorological conditions and parameters. A typical PV output power variation is shown in Fig. 2. This figure shows the PV output power for three consecutive days (from September 24-26, 2020) selected at random for a PV plant located at the rooftop of the Department of Electric Power Engineering, Norwegian University of Science and Technology, Norway. It can be seen from the figure that the power output is highly varying over a short period, which brings a threat to the operation of the power grid. Although solar irradiation and temperature are the two most important meteorological parameters that directly affect the amount of energy generated from PV panels, the contribution of the geographical area where the plant is located is also significant because of a phenomenon called *Soiling Effect*.

The soiling effect can result from snow, dirt, dust, and other particles that cover the surface of a PV module. This effect is significant especially in areas with extreme weather conditions. For a desert type PV (arid and semi-arid areas), for example, a reduction of up to 10 % in energy production was measured as a result of soiling loss due to dust [7]. Similarly, the soiling effect of snow and ice coverage in colder climates is significant on winter days. However, quantification of the exact reduction in the generation is a complicated process due to the complex nature of the optical characteristics of snow [8]. More on factors affecting PV power is covered in Section III.

Due to the concerns discussed above, high penetration of PV power into the grid brings a challenge in the operation of the existing power system. The inherent variability of solar power creates challenges in matching variable load with variable supply. Maintaining the instantaneous balance between production and demand becomes a difficult and expensive task [9]. This will in turn affect the decision-making ability of dispatch centers and energy trading companies on critical issues such as alternate adjustments

for conventional power sources, scheduling arrangements, storage requirements, and overall planning [10].

The unique climate of cold regions makes the task of integrating PV even more challenging. These regions are generally characterized by heavy snow and short sun hours during the winter, long sun hours during the summer, and fast and frequent weather variation over a short period throughout the year due to fast-moving clouds (especially Norway). The consequence of fast-moving clouds can be catastrophic especially when a significant amount of PV plants is integrated with a low voltage grid. This can cause fast and large changes in solar irradiance resulting in a sudden drop in PV output power, which ultimately leads to disruptions and grid instability. Addressing these issue is therefore very important to exploit and realize the full potential of solar electricity in the Nordic region.

The use of storage systems and consumption flexibility are usually suggested as means to overcome these limitations [11], [12]. However, their success is greatly impacted by the high cost of batteries and the lack of clear financial incentives that motivate consumption flexibility. In recent years, there has been an increasing interest in the use of energy informatics technologies to reduce the negative impacts of grid-integrated PV systems.

Energy informatics technologies (AI (Artificial Intelligence) and ML (Machine Learning)) can overcome these PV power limitations and make solar electricity an equal contributor to the energy mix. ML and AI methods can improve the adoption of solar electricity resulting in a modernized electrical grid supporting the reliability and resilience of the overall grid. These technologies are usually used in the forecasting of global irradiance and solar power output [13]. Accurate PV power output forecasting enables power system operators to make proper scheduling and operation planning, allow accurate energy trading decisions in power markets and significantly reduce the cost and size of balancing reserves.

Given the unique weather characteristics in these regions, much uncertainty exists about the design of a PV output power forecasting model. The commonly available AI/ML-based PV power forecasting models in the literature could be inadequate and could result in inaccurate and unstable forecast results. The optimal size and type of informative environmental parameters to use and the specific type of AI algorithm to implement are not fully understood. Very little is also known about the impact of the forecast horizon, the areal scale, and timestep on the performance of a forecast model in these regions. This study offers some important insights into the above critical shortcomings. Our main objective in this paper is therefore to quantitatively and qualitatively assess the role that AI/ML technology play in forecasting the amount and variation of output power focusing on PV plants located in cold regions. This work will potentially serve as a starting point where more area-specific AI-based PV power forecasting models can be designed to further promote the deployment of PV in these regions.

A. OBJECTIVES

The following are some of the important issues that are addressed in this paper by reviewing recent research works:

- 1) Identify relevant parameters or inputs that affect AI-based PV output power forecasting in general and under a cold climatic condition in particular.
- 2) Investigate different PV power forecasting model types and suggest the 'one' that works relatively well for these regions.
- 3) Identify a forecast time step that can capture all the variations in the PV output power and identify a forecast horizon that suits a particular application.
- 4) And finally, identify and suggest techniques and methods that can improve the performance of a forecast model in terms of accuracy and stability.

The rest of this paper is organized as follows. Section II begins by discussing the general steps that are used in the design of an AI-based PV output forecasting model. It will then go on to a brief description of the working principles of the widely used conventional and deep learning AI algorithms for PV power forecasting. The impact of various environmental parameters on PV power forecasting is addressed in section III. In section IV, special focus is given to the case of PV in colder climates. Classification of various PV power forecasting models based on forecast horizon, model type, time step, areal scale, and approach used is covered in Section V. Finally, the important conclusions and observations from this review work are outlined in Section VII.

II. AI FOR PV OUTPUT POWER FORECASTING

This section addresses the generic AI-based PV power forecasting procedures and the common AI algorithms that are used in the design of a typical AI-based PV power output forecasting model.

A. AI PROCESS IN PV OUTPUT POWER FORECASTING

AI algorithms are currently the most widely used PV output power forecasting methods in the literature. Their popularity is due to their ability to effectively map the highly non-linear relationship that exists between environmental input parameters and the PV power. A generic flowchart showing the AI-based PV power forecasting process is shown in Fig. 3. The whole process can be summarized into three phases. A closer look at the figure shows that the forecasting model uses both power and weather parameter measurements as an input. It should be noted however that several high-performing AI models exist in the literature that only requires the measurement of PV power as input, especially for very short-term forecast horizons.

1) INPUT PHASE

AI is a data-driven method, therefore, an important step in AI-based PV power forecasting, is data collection and analysis. This phase involves the collection of input-output data (environmental and PV power) from the site where the plant

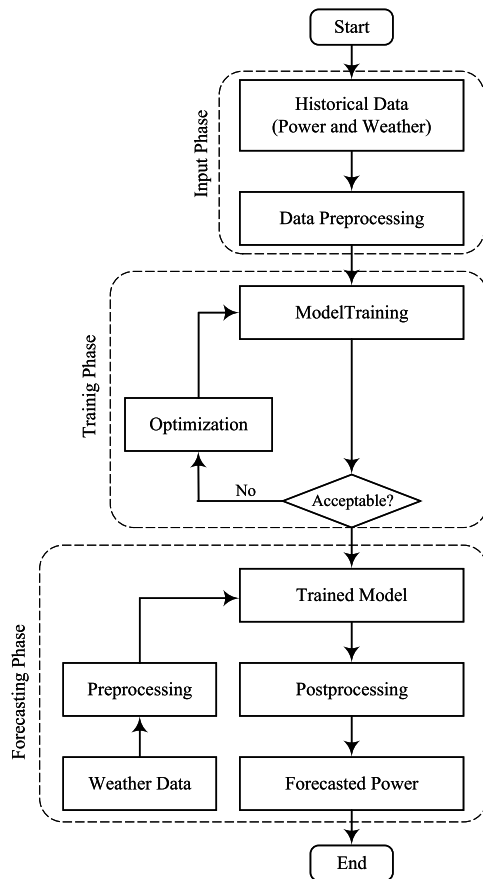


FIGURE 3. Generic AI based PV power forecasting flowchart.

is located and preprocessing of these data. Environmental data including solar irradiance and air temperature are usually gathered from the weather station at or near the plant. Similarly, PV output power records can be obtained from the SCADA (Supervisory Control And Data Acquisition) system in the case of large plants or from data loggers for the case of small PV plants. An equally important step after data collection is the preprocessing of the input data.

Preprocessing consists of (but is not limited to) dealing with missing values, outlier detection, data resampling, data scaling, and time series decomposition. Widely used methods to deal with missing values in PV forecasting include completely ignoring those rows of data from the training process (if the missing data are unimportant and/or small) and data imputation using interpolation techniques. Data resampling is also required when the different input data have different granularity.

The presence of outliers in the training data can lead to inaccurate prediction and may result in requiring a longer time to fully train the model. It is therefore very important to filter out the outliers that exist both in the power series and environmental data. These kinds of data are usually observed in the early periods of sunrise and sunset. In colder climates, snow can be the main reason for outliers during winter. Authors in [14] used the Hampel filter to detect

and remove outliers from the measured PV output power. To benefit from various filtering algorithms, Pan *et al.* [15] used an ensemble filter algorithm to manage abnormally extreme weather inputs.

Data scaling is important to obtain a high-performing model both in terms of accuracy and computational resource requirement. Due to the Gaussian Distribution nature of PV power and other parameters, the most common data scaling technique in PV power forecasting is normalization [16], [17]. This rescales all the input features so that their value is always between 0 and 1. It is given by (1),

$$x_{nor} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

where x is the value in SI unit, x_{min} and x_{max} are the minimum and maximum value in feature X respectively, and x_{nor} is the new scaled value.

Another very significant preprocessing step that is frequently used in AI-based PV power forecasting is time series decomposition. Common tools that are used for treating time-series signals are WPD (Wave Packet Decomposition) [12] and WT (Wavelet Transform) [18]. These methods offer filtering ability and thus result in better performance characteristics.

2) TRAINING PHASE

The training phase involves selecting a particular type of AI algorithm and training the algorithm with the training dataset. Evaluating the model with the validation dataset and fine-tuning the internal parameters of the algorithm to further improve performance is also part of the training process. The common AI algorithms that are widely used for PV forecasting are briefly discussed in Subsections II-B1 and II-B2 below.

For conventional ML algorithms, fine-tuning the internal parameters is usually done using extensive grid searching and optimization algorithms such as PSO (Particle Swarm Optimization) [18], GA (Genetic Algorithm) [19], [20], and ACO (Ant Colony Optimization) [15]. Due to the complex nature of the network, forecasting models based on DL (Deep Learning) algorithms use the trial and error approach to adjust their parameters.

3) FORECASTING PHASE

Once the model is tested and validated to give satisfactory performance, it will be directly used on a new set of input test data for forecasting purposes. The input test data, in this case, is the forecasted weather data (usually from NWP (Numerical Weather Prediction)) unlike historical weather data used to train the model. Sometimes, post-processing can be implemented to further improve the prediction ability of the model.

B. COMMON AI ALGORITHMS

The common AI algorithms that are widely used in the literature for PV power forecasting can be broadly grouped

into conventional ML models and DL models. A brief description of these models is given below.

1) CONVENTIONAL ML MODELS

These are groups of ML algorithms that have limited ability to process data in its original form [21]. These methods require considerable understanding and expert domain knowledge. Consequently, the selection of features is an important step and requires careful engineering. The common conventional ML algorithms that are widely used for PV output power forecasting are discussed below very briefly.

a: SUPPORT VECTOR MACHINES (SVM)

SVM was originally introduced by Vapnik in 1992 to solve classification problems but later extended to handle regression problems too. The input data is first mapped into a high-dimensional feature space through non-linear mapping so that a linear or non-linear regression approximation can be achieved in space. SVM generalization to SVR (Support Vector Regression) is achieved by introducing an ϵ -insensitive region around the function, called the ϵ -tube [22]. An optimization problem is solved to find the tube that best approximates the continuous-valued function, which is a tradeoff between model complexity and prediction error. The ϵ -insensitive loss function penalizes predictions that are farther than ϵ from the desired output. The value of ϵ determines the width of the tube; a smaller value indicates a lower tolerance for error. This ensures that SVR is less sensitive to noisy inputs and the model more robust.

b: ENSEMBLE OF TREES

The ensemble is a technique that allows combining multiple weak machine learning models to create a more powerful model. The two common ensemble models that have proven to be effective on a wide range of datasets including PV power forecasting, both of which use decision trees as their building blocks are random forests and gradient boosted decision trees [23].

The theory behind an RF (Random Forest) is to average multiple DTs (Decision Tree) that are slightly different from each other and that suffer from high variance/overfitting, to build a more robust model that has a better generalization ability and is less prone to overfitting [24]. The DTs in an RF are randomized either by selecting the data points used to build a tree (bootstrapping) or by selecting different features in each split test or both.

In contrast to the RF approach which allows parallel running of DTs, GBDT (Gradient Boosted Decision Tree) works by building trees in a serial manner, where each tree tries to correct the errors of the previous step [23]. The learning rate is the hyperparameter that controls how strongly each tree tries to correct the errors of the previous trees.

c: MULTILAYER PERCEPTRON (MLP)

MLP is a typical example of a feedforward ANN (Artificial Neural Network) where each layer serves as the input to the

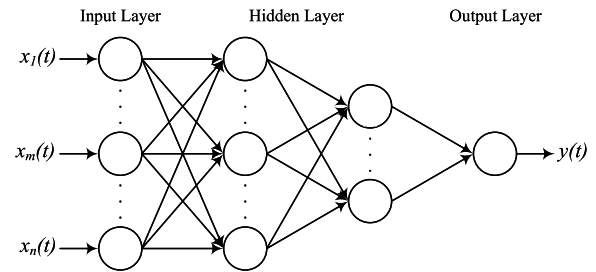


FIGURE 4. The basic structure of a multilayer perceptron architecture.

next layer without loops and it is an entry point towards complex neural networks such as CNN (Convolutional Neural Network). A basic MLP network structure with one input layer, two hidden layers, and one output layer is shown in Fig. 4. The number of hidden layers and units in the MLP are hyperparameters that need to be optimized for a given problem. However, as more layers are added to a network, the error gradient that is calculated using the backpropagation algorithm becomes increasingly small [24]. This vanishing gradient problem makes model learning more challenging.

2) DEEP LEARNING MODELS

The term deep learning refers to artificial neural networks with multiple layers. The interest in having deeper hidden layers has recently begun to surpass the performance of conventional ML methods in various fields [25]. This is also evident from the literature in PV output power forecasting. The commonly used DL models for PV forecasting are CNN and LSTM (Long Short-Term Memory).

a: CNN

CNN can automatically extract high-level features from raw input data, which are much more powerful than human-designed features [26]. It is a multi-level representation learning, where abstract features are learned from raw data through successive non-linear transformations [27]. Consequently, it has brought significant improvement in the performance of DL models in various applications. The basic structure of CNN is shown in Fig. 5 and consists of convolutional layers, pooling layers, and fully connected layers.

- Convolutional layer: A convolutional layer comprises a set of filters (a grid of discrete numbers) which are convolved with a given input to generate an output feature map [28]. The weights of each filter are learned during the training of CNN through multiple iterations.
- Pooling layer: A pooling layer operates on blocks of the input feature map using pooling functions such as average or max. This operation effectively down-samples the input feature map and is useful for obtaining a compact feature representation. It has no learnable parameters.
- Fully connected layer: Fully connected layers are typical feedforward neural network layers like MLP. Each unit is densely connected to all the other units of the previous

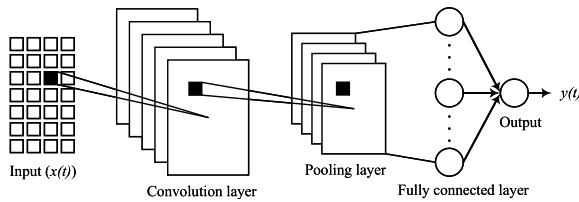


FIGURE 5. Typical architecture of a convolutional neural network.

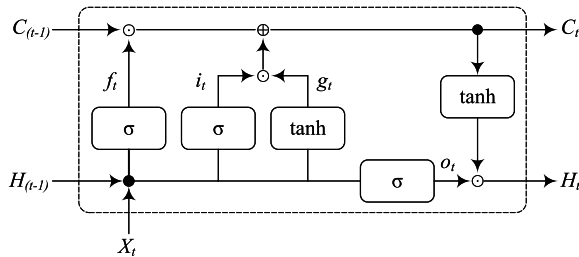


FIGURE 6. The internal structure of LSTM memory cell.

layer. They use non-linear combinations of the extracted features to make the final prediction.

b: LSTM

Unlike feedforward neural networks, RNNs (Recurrent Neural Network) are ideal candidates for modeling time-dependent and sequential data problems, such as stock market prediction, machine translation, and time series prediction. However, conventional RNNs suffer from the problem of vanishing gradients. The gradients become too small, and the weight updates become very insignificant. This makes the learning of long-term dependencies difficult. An LSTM network is a popular RNN architecture that solves this problem of vanishing gradient.

LSTM deal with the vanishing gradient problem by not imposing any bias toward recent observations, but it keeps constant error flowing back through time [29]. This is possible by the introduction of gates (input, forget, and output) into the internal structure of LSTM based neurons (also called memory cells). This structure allows better control of the gradient flow and enables better preservation of long-term dependencies.

The internal structure of a typical LSTM memory cell is shown in Fig. 6 [24]. Here, \odot represents element-wise multiplication and \oplus , element-wise summation. C , H , and X represent the cell state, the hidden unit and the input respectively. Similarly, f_t , i_t , g_t and o_t denotes the forget gate, input gate, input node and output gate respectively. Finally, sigmoid and hyperbolic tangent activation functions are represented by σ and \tanh .

Each gate has a specific functionality. The forget gate decides which hidden unit information to keep or discard from the previous time step. It is calculated using (2),

$$f_t = \sigma(W_{xf}X_t + W_{hf}h_{t-1} + b_f) \tag{2}$$

where W_{xf} is the weight between the input and the forget gate, W_{hf} is the weight between the hidden unit and the forget gate and b_f is the bias term for the forget gate.

The input gate is responsible for updating the current cell state based on the updated cell state of the previous unit by the forget gate (3)-(5),

$$i_t = \sigma(W_{xi}X_t + W_{hi}h_{t-1} + b_i) \tag{3}$$

$$g_t = \tanh(W_{xg}X_t + W_{hg}h_{t-1} + b_g) \tag{4}$$

$$C_t = (C_{t-1} \odot f_t) \oplus (i_t \odot g_t) \tag{5}$$

where W_{xi} is the weight between the input and the input gate, W_{hi} is the weight between the hidden unit and the input gate, W_{xg} is the weight between the input and the input node, W_{hg} is the weight between the hidden unit and the input node and b_i and b_g are the bias terms for the input gate and input node respectively.

Finally, the output gate updates the value of hidden unit from the previous time step and this value is used to compute the hidden unit at the current time step (6)-(7),

$$o_t = \sigma(W_{xo}X_t + W_{ho}h_{t-1} + b_o) \tag{6}$$

$$h_t = o_t \odot \tanh(C_t) \tag{7}$$

where W_{xo} is the weight between the input and the output gate, W_{ho} is the weight between the hidden unit and the output gate and b_o is the bias term for the output gate.

A high-level qualitative comparison between the different AI-based PV power output forecasting algorithms is given in Table 1. In this table, computational demand, explainability level, the amount of required fine-tuning and preprocessing, and time sequence support are demonstrated as key parameters that are used to compare the different AI-based PV power forecasting algorithms. Explainability here is defined as a term that indicates the level of understanding of the internal decision-making rules of a PV output power prediction model as used in [30]. It can vary between the extreme cases of black-box (not interpretable directly) and white-box (easily interpretable) models.

As it can be observed from the table, the conventional ML models generally have the advantages of relatively low computational demand and require a minimum or a moderate level of fine-tuning. However, these methods need a demanding preprocessing step to generate important features from the input. On the other hand, the DL-based methods have a high computational requirement, but minimal preprocessing. They can extract high-level features from the input by themselves. Computational demand or requirement of individual AI-based models or tools depends on the selected AI algorithm, their parametrization details, length of the training and testing datasets, size of the entire dataset, and many other parameters. It is also worth noting that almost all the AI methods in the table have a black-box nature. Furthermore, it is also important to see that only LSTM-based models have the ability to support time sequence in the data. This is one reason that these models and their variations have emerged as particularly attractive algorithms to design a PV

TABLE 1. Qualitative summary of AI-based PV power forecasting algorithms.

Algorithm	Computational demand	Explainability level [30], [31]	Required fine-tuning	Time sequence support	Preprocessing requirement
SVM	High	Blackbox	Minimal	No	Mandatory
RF & GBDT	Low	Blackbox	Minimal	No	Recommended
MLP	Low	Graybox	Moderate	No	Mandatory
CNN	Very high	Blackbox	A lot	No	Minimal
LSTM	Very high	Blackbox	A lot	Yes	Minimal

power forecasting model recently. The choice of a particular algorithm for a given application should however be made considering all the parameters given in the table.

III. ENVIRONMENTAL PARAMETERS AFFECTING PV POWER FORECASTING

Meteorological parameters are the most important factors that directly determine the performance of any PV power forecasting model. Solar irradiance, air temperature, wind speed, wind direction, relative humidity, air pressure, and cloud cover are some of the parameters that are widely used as input to a PV power forecasting model. However, solar irradiance is by far the most significant and universal (not area-specific) of all the parameters and it is directly related to PV output power. Sudden and abrupt changes in PV power production are determined by the movement of clouds. So, in regions where the weather changes quite frequently over a short period, cloud coverage can be an equally important parameter as solar irradiance. Correlation study between meteorological variables and measured PV power is, therefore, an important step in the design of a forecasting model. Selecting an optimum number of informative inputs is crucial to obtain a high-performing forecasting model which uses the smallest computational resources.

The relative importance of most meteorological parameters for PV output power forecasting highly depends on the geographical area where the plant is located [32]. This can be observed from the web chart shown in Fig. 7 for randomly selected locations. This figure shows the correlation of the different meteorological parameters with the PV output power. It is possible to see from the figure that solar irradiance is a very important parameter irrespective of the plant's location. This claim is consistent with the observation that in the majority of literature reviewed in this paper and others, solar irradiance is used as an important input parameter. On the other hand, the importance of wind speed and relative humidity are highly affected by the location of the plant.

In addition to the common meteorological parameters affecting PV output power forecasting, other location-specific factors such as snow and dust cover on PV panels should also be considered for regions with extreme weather conditions. The impact of snow and ice cover on the yield of a PV plant in colder climates during winter days is very complex [8]. It can vary from partial to total obstruction of solar radiation reaching the PV modules, which results in a reduced or no generation at all. And in contrast, during the days where the modules are clear of snow, but the surrounding



FIGURE 7. Relative importance of weather parameters for various locations based on correlation study with respect to the PV output power [12], [33]–[37].

area is covered with snow, the consequence is opposite to the above situation. In this case, the reflectance property of snow (*Albedo Effect*) tends to increase the solar radiation reaching the surface of the PV modules, and hence increasing the PV power generation. If these situations are not taken into consideration in the design of PV output power forecasting model in these regions, they will lead to huge forecast errors. Similarly, in arid and semi-arid areas, issues associated with the accumulation of dust should be taken into consideration. More on the case of PV in a colder climate is covered in the following section (Section IV).

Besides the parameters themselves, seasonal variation of the parameters is also another key aspect worth considering in the design of a PV power forecasting model. This is because the performance of a prediction model is directly affected by the season of the year. This is evident from Table 2 which shows how the prediction performance (measured in terms of *RMSE* (Root Mean Squared Error) and *MAPE* (Mean Absolute Percentage Error)) of a forecast model is dependent on the season of the year. As seen in Table 2, the performance of the prediction models is better during the winter period in terms of *RMSE*. In contrast, the worst performance is observed during the summer. The *RMSE* increased by 77%, 994%, 38%, 80% and 157% from winter to summer for [12], [16], [18], [38], and [39] respectively.

A possible explanation for such observation can be that the weather changes quite frequently in summer and less in winter in these regions. The forecast model is unable to capture the fast and sudden variations. It should be noted however that this conclusion is only made with a limited reference and cannot be claimed to be the case for all regions. Another explanation can be in the definition of *RMSE* itself, which magnifies larger errors calculated between the measured and predicted power. No such conclusive statement

TABLE 2. The effect of seasonal variation on prediction performance.

Author	Metrics	Season			
		Winter	Spring	Summer	Autumn
[12]	MAPE(%)	1.868	2.266	2.888	2.622
	RMSE(kW)	0.1526	0.2454	0.2705	0.2221
[16]	MAPE(%)	15.27	14.2	13.65	14.77
	RMSE(kW)	0.348	0.359	3.809	0.351
[18]	MAPE(%)	5.572	4.145	2.719	4.442
	RMSE(kW)	2.315	2.400	3.208	2.687
[38]	MAPE(%)	27.72	18.13	23.85	21.98
	RMSE(kW)	0.86	1.28	1.55	1.47
[39]	MAPE(%)	2.405	2.398	3.999	2.544
	RMSE(kW)	2.795	3.487	7.193	3.809

can be made from the observation of *MAPE* across the seasons.

IV. PV IN COLDER CLIMATE

Since the focus of this paper is about ML-based PV power forecasting under a Nordic context, it is important to emphasize the case of PV in a cold climate. This section covers how this unique climate offers an ideal condition for PV operation and at the same time how it can also bring its own challenge. Finally, how it is possible to reduce the challenge is also covered in this section.

A. HOW COLD CLIMATE IS 'IDEAL' FOR PV

Many experiments show that solar cells perform very well under cold climate conditions. This is due to the low operating temperature inside the PV modules which is the direct result of the low temperature of the area where the plant is located. Manganiello *et al.* [40] demonstrated that PV modules located in colder regions result in higher seasonal energy yield and negligible energy loss. Their work also indicated that such high latitude climate has the potential to realize seasonal matching of production and consumption. This is also evident from the work of [41]. Their study has shown that by optimally placing PV installation at a relatively higher tilt angle in colder climates, it is possible to bring the temporal production profile of PV into better correlation with typical electricity consumption patterns. They investigated and quantified the potential of PV installations that favor high winter irradiance, high ground-reflected radiation (albedo effect), and steeper panel tilt angles. Adaramola and Vågnes [42] conducted similar work for a PV plant in Ås, Norway. An annual yield of 2.55 kWh/kWp obtained from their analysis suggest that PV installations in this location and similar other locations in Norway is technically feasible.

B. CHALLENGES OF COLD CLIMATE

Despite the opportunities that come along with PV installations in colder climates, it has also a few but important limitations. Lower sun angle, short solar duration hours, thick clouds, and soiling due to snow are the typical challenges that PV installations in high latitude climates face during the winter. The impact of snow is by far the most significant

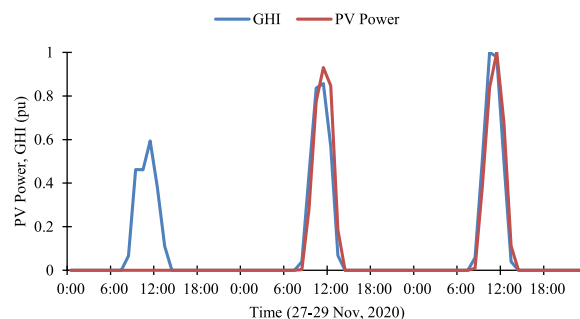


FIGURE 8. The effect of snow on PV power generation (total obstruction).

one. Modeling the impact of snow is widely reported and extensively explored in the literature [43]–[45], but it is still difficult to fully understand and account for snow melting and sliding processes. For this paper and better understanding, the impact of snow on the PV output power can be observed as a three-stage process.

1) DURING SNOWING

During the time of fresh snow, the solar irradiation reaching the surface of the PV modules is also significantly reduced due to the thick cloud. Consequently, the power output from the PV system is close to zero, if not zero already.

2) DAYS AFTER EXTENDED PERIOD OF SNOWING

This is the most challenging period where the impact of snow is difficult to quantify precisely. The snow cover acts as a shading that prevents solar irradiation from entering the solar cells. The effect can vary from full shading resulting in zero generation to partial shading which results in a reduced generation. The impact of snow on yield in such cases depends on various factors such as snow depth, snow weight, the tilt angle of the module, ambient temperature, wind speed, and surface property of the module. Another important factor to account for the snow effect is the snow clearing process and the time required to clear a module completely.

An example of total obstruction of irradiance due to snow accumulation on the surface of the PV module can be seen from Figure 8, which is recorded for the same PV plant described earlier. On 27th November 2020, the PV output power is zero although there is a significant amount of solar radiation (a relatively clear sky day during the month of November in Norway). This can possibly be explained by the accumulation of snow on the PV modules (total obstruction). Such kinds of effects will lead to a huge forecast error if not taken into consideration. For the next two days, it is also possible to see that the snow is cleared from the modules and the plant is operating as it should be.

An early example of research into the impact of snow on the yield of a PV plant includes the work of Becker *et al.* [43]. They have obtained a strong correlation between the incidence of freshly fallen snow and a decrease in yield. An estimated average yearly yield loss between 0.3 and 2.7% was measured for a 1016 kWp PV plant in Munich, Germany

between 1999 and 2004. What was surprising in their finding is the observed sliding procedure of snow from the PV panels. Even for small values of solar irradiation (less than 100 W/m^2) and temperature (less than $0 \text{ }^\circ\text{C}$), the snow on the modules began to slip.

Another important work that quantifies the losses associated with snow for a PV plant in Colorado and Wisconsin is the work of Marion *et al.* [46]. They developed a model that can estimate yield loss due to snow. The model uses daily snow depth, plane-of-array irradiance, air temperature, PV array tilt angle, and the extent of snow coverage on the PV array. The model performed very well mostly and gave a result in agreement with the measured values. Similarly, a simplified two-stage snow modeling method is proposed in [44]. The estimated hourly solar energy output is first calculated using module plane-of-array irradiance, module temperature, and derating factor. And in the second stage, a binary decision-making matrix is used whether to disregard the estimated energy depending on the number of days since last snowfall, snow depth, and ambient temperature. Despite the simplified nature of this approach, it suffers due to the binary decision-making process. It also ignores the fact that snow can be transparent depending on its depth and the time required for clearing the modules is not also considered.

A snow loss prediction model that takes into consideration the transmittance property of snow is proposed in [45]. Snow-covered modules are modeled in such a way that they can generate power based on snow depth. The *Bouguer-Lambert* law was used to estimate the amount of insolation that is received on the surface of uniformly snow-covered PV modules. The results of their study showed that the orientation (portrait or landscape) of the modules plays a significant role in reducing the impact of snow in snowy conditions. This paper only considered snow sliding as the dominant snow removal process and did not consider snow melting and snow removal due to wind. The surface of the PV modules is also assumed to be uniformly covered by snow at every time. Despite the success of this work in bringing significant improvement in the modeling of snow loss, the fact that it ignored other snow clearing procedures and the assumption of uniform snow cover on the modules at all times can be important constraints.

Due to the complex nature of factors such as the extent of snow cover, tilt angle, snow transmittance, and snow clearing processes on snow loss prediction, all the work we have seen above used at least one or two simplifying assumptions. Inspired by this limitation, Hashemi *et al.* [47] proposed a snow loss prediction model based on machine learning algorithms for a PV plant located in Ontario, Canada. They designed an algorithm that can capture all the inherent complexity using the available meteorological data from the location of the plant. A prediction model based on a gradient boosted tree resulted in the lowest mean squared error.

Further details on the impact of snow on yield reduction are beyond the scope of this paper and the reader is directed to Pawluk *et al.* [48] for more information. In this paper, they

have discussed the impact of snow, identified factors that influence the generation loss, examined existing snow impact estimation techniques, and finally concluded by discussing various mitigation strategies to reduce the impact of snow.

3) CLEAR MODULES AFTER SNOWFALL

The effect of snowfall on the performance of a PV plant is not always negative and it can be used to optimize system design in colder climates [49]. This is due to a phenomenon called the *Albedo Effect*. Albedo is the term describing the reflectance property of snow and plays a huge role for PV in high latitude and colder climates. The incoming reflected light from the surrounding can be as high as 3 to 6 times when it is covered in the snow [8]. By properly choosing a location with a strong albedo effect and optimally sizing the tilt angle, it is possible not only to maximize generation but also to shift the temporal production patterns to match the typical demand [41]. Solar electricity generation could effectively be shifted from summer to winter without compromising the total annual yield.

C. HOW TO COPE WITH THE CHALLENGES

PV output power forecasting models that are designed for colder climates should take into consideration all the above limitations. Incorporating short-term hourly snow loss prediction models is very important as they can play a key role in the operational management of electric grids in such locations. Correction strategy is required for PV power forecasting models to incorporate the loss due to snow or snow cover and the albedo effect should be considered as additional input parameters in the design of the forecast model in the first place.

V. CLASSIFICATION OF PV POWER FORECASTING MODELS

Various forecast models have been used effectively for PV output power forecasting. These models can generally be categorized into multiple groups based on forecast horizon, type of prediction model, time step, areal scale, and approach used (Fig. 9). It should be noted however that there is no universally accepted way to classify PV power forecasting models and numerous other approaches have been used widely in the literature.

A. CLASSIFICATION BASED ON FORECAST HORIZON

Based on forecast horizon (length of time into the future for which forecasts are to be made), PV power forecasting models can be categorized into three main groups. Short-term, medium-term, and long-term forecast. Although this classification is widely used, it is also common to find a fourth category in many works of literature (i.e. Ultra short-term forecast). Each forecast horizon is applied for various purposes and applications. It is worth noting that there is a lack of consensus among researchers in defining the boundary between different forecast horizons and usually overlapping happens between each group. Fig. 10 shows this

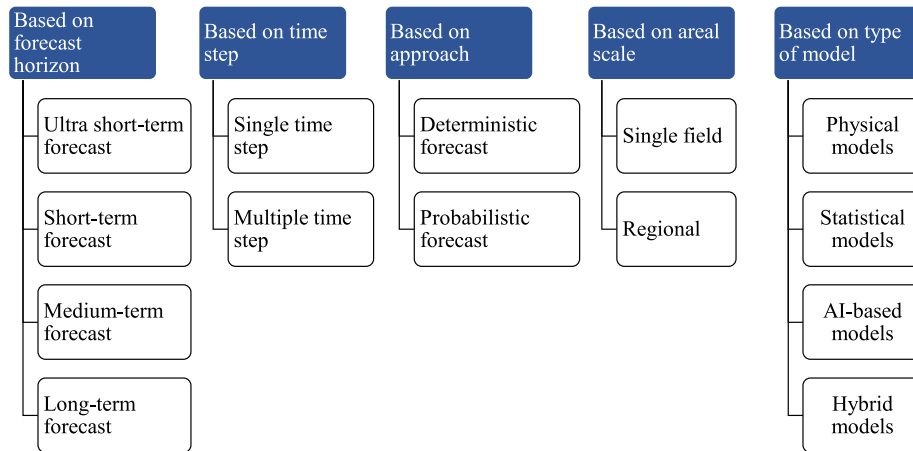


FIGURE 9. Classification of PV power forecasting models.

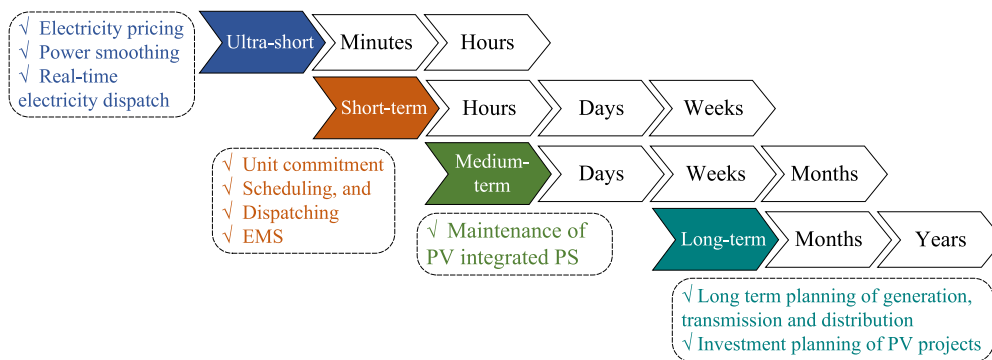


FIGURE 10. Overlapping nature and application areas of forecast horizons.

overlapping nature and the different application areas of each forecast horizon.

1) SHORT-TERM FORECAST

This includes a forecast horizon of up to one hour, several hours, one day, or even a week. This kind of forecast allows power system operators to ensure unit commitment, scheduling and dispatching. It is also crucial in PV-integrated energy management systems and in energy market operation [50]. Ultra short-term forecasts (from minutes to hours) on the other hand are highly beneficial to electricity pricing, power smoothing, and monitoring of real-time electricity dispatch.

2) MEDIUM-TERM FORECAST

Medium-term PV power forecasting is usually done for one week to one month. However, some also consider the forecast between one day and one week ahead as being in this group. This type of forecast is particularly important to schedule the maintenance of PV integrated power systems by considering the availability of generation in the future.

3) LONG-TERM FORECAST

Long-term PV power forecasting includes a forecast horizon of a month and up to a year. Typically used by power system owners and operators for long-term planning of the electricity

generation, transmission, and distribution. The interested reader is directed to the work of [50] for further reading on forecast horizons for PV power output forecasting.

As weather parameters are usually difficult to forecast with acceptable accuracy beyond a certain short period, PV power forecasting also suffers as the horizon is increased. Table 3 shows how the performance of a prediction model degrades as the forecast horizon is increased. It can be seen from the data in the table that there is a significantly noticeable difference in the value of the evaluation metrics as the forecast horizon is increased. For example, in [18], the *RMSE* error increased by 37.5%, 87.5%, and 212.87% when the forecast horizon is increased from 3 hours to 6, 12, and 24 hours respectively.

Such kinds of performance discrepancies can be particularly huge in areas where the weather changes quite frequently over a short period. Norwegian climate can be a good example for this. Due to the frequent movement of clouds; cloud coverage and solar irradiance cannot be forecasted with high accuracy beyond a certain very short period. Poorly forecasted weather parameters directly result in inaccurate PV power output prediction. So, emphasis should be given to the choice of forecast horizon before starting to design a prediction model. Based on this point and closer examination of the discussion above, a short-term, or ultra-short-term forecast could be a good choice for the

TABLE 3. The effect of forecast horizon on the performance of a prediction model.

Author	Evaluation Metrics	Forecast horizon	% Change in metrics value ^a	Reference values
[18]	RMSE	6 hours	37.50	RMSE = 0.8477 kW Time step = 3 hours
		12 hours	87.50	
		24 hours	212.87	
[38]	RMSE	15 minutes	15.12	RMSE = 1.39 kW Time step = 7.5 minutes
		30 minutes	30.21	
		60 minutes	50.36	
[52]	MAE	15 minutes	53.09	MAE = 47.47 kW Time step = 5 minutes
		30 minutes	87.83	
		60 minutes	131.85	
		120 minutes	149.77	
		180 minutes	179.20	

^aCalculated with respect to the reference values given in the table

Nordic climate. This is also evident from the work of [51]. Their result suggest that for a PV plant in Norway, PV output power forecast cannot be made for forecast horizon longer than 1 hour ahead without introducing a significant error.

Comparative study of the performance of PV power prediction models in all the forecast horizon categories has received limited attention in the literature. This issue has been addressed by recent work of [53]. Here the authors evaluated the performance of various ML algorithms such as LR (Linear Regressor), PR (Polynomial Regressor), DTR (Decision Tree Regressor), SVR, RFR (Random Forest Regressor), MLP, and LSTM for short-term (24 hours ahead), medium-term (1 week ahead), and long-term forecast (1 year ahead). Their overall analysis shows that a PV power forecast model based on RFR resulted in a better performance. Such kind of study is very helpful to identify a forecast horizon that specifically matches the weather condition of a given area.

B. CLASSIFICATION BASED ON THE TYPE MODEL USED

Based on the particular type of model they use, PV power forecasting methods can generally be categorized into four groups: Physical, Statistical, AI-based, and Hybrid models. The first group depends on physical modeling of the PV plant and weather system, but the remaining are data-driven methods that rely only on the measurement of PV power and other environmental parameters. It is important to note that there is no universally suitable approach for classifying PV power output forecasting based on the type of models they use. Different approaches have been proposed in different works of literature.

Authors in [50] classified PV power forecasting models into three groups consisting of a physical model, persistence model, and statistical models. The statistical model included both the traditional time-series forecasting methods (ARIMA (Auto Regressive Integrated Moving Average) and SARIMA (Seasonal ARIMA)) and models based on conventional AI/ML algorithms. A simplified classification approach is proposed in [54] where PV power forecasting models are classified into two groups. Model-based and data-driven approaches. An alternative and comprehensive approach which is also similar to the one used in this paper is developed by [32]. Here they classified PV power forecasting

models into physical, statistical, ML, persistence, and hybrid methods. To summarize, PV output power forecast model classification is usually a matter of subjective opinion with which the authors of this paper agree and we believe that the system of classification adopted in this paper satisfies the needs of the study.

1) PHYSICAL MODELS

In this approach, the PV plant is first modeled mathematically with plant-specific parameters such as module inclination angle and module efficiency. Then this model is used directly to calculate the PV generation with forecast information of solar irradiance and temperature obtained from NWP. This approach is particularly important as part of a feasibility study to determine the amount of PV generation before it is constructed. Due to its dependence on the forecast of NWP, which has poor resolution for short periods, this approach is usually used for long-term forecast purposes [55], [56].

2) STATISTICAL MODELS

Statistical models assume that the future value of the target variable is a linear function of past observations and random errors. Popular statistical forecasting methods for PV output prediction include ARIMA and SARIMA. SARIMA is an improved version of ARIMA, and it is designed to support seasonality which is an important characteristic of PV power data.

Kushwaha and Pindoriya [57] implemented SARIMA for very short-term PV power forecast and concluded that despite the satisfactory performance of the model during clear sky days, the performance degraded significantly on cloudy days where frequent weather changes are quite common. To overcome such shortcoming, SARIMA is usually used in a combination with other techniques (Wavelet-SARIMA) and models. Implementing this approach is the work of [58], where SARIMA is used together with ANN in a parallel structure. In doing so, they showed that both the accuracy and resilience of the prediction model are improved. Based on the work of [57] and others, statistical methods are not recommended forecasting models for Nordic climate, given the high variation of weather parameters over a very short period.

3) AI-BASED MODELS

The majority of works in PV output power forecasting in the literature are implemented using AI algorithms. These methods rely on the ability of a model to learn from historical data and to further refine its predictive ability through training. AI-based PV output power forecasting models in this section include both the conventional AI/ML algorithms and the DL algorithms as described in Sections II-B1 and II-B2 respectively. Here we give an overview of how these algorithms are used in the design of a PV power forecasting model. The ensemble method is also discussed as a technique that improves the performance of a prediction model.

The performance of various ML models including ANN, Linear Regression, MSP Decision Tree, and GPR (Gaussian Process Regression) is compared in [59] for a PV plant in Qatar. CFS (Correlation feature selection) and RFS (Relief Feature Selection) techniques are used to identify the relevant features for the models. A forecast model based on the ANN network outperformed all the other models in terms of $RMSE$ and R^2 . The prediction accuracy and error distribution of SVM and ANN are compared in [10]. This information is in turn used to estimate the capacity of the energy storage system required to absorb mismatch in energy trading applications. Both methods gave satisfactory results, but ANN marginally outperformed SVM. Authors in [60] assessed the performance of ANN, SVR, and RT for predicting the power output of a PV plant against a PM (Persistence Model) model. Their comparative analysis shows that the ANN outperformed the other models, resulting in the lowest normalized $RMSE$ and $MAPE$ error of 0.6% and 0.76% respectively.

Three commonly used conventional AI/ML models in PV output power forecasting, i.e., MLR (Multiple Linear Regression), GB (Gradient Boost), and ANN are compared in [61]. These models are tested with different training windows and features for a PV site in the state of Florida. The forecast model based on GB resulted in the lowest $RMSE$ and lowest variance. ANN produced the lowest performance in terms of both accuracy and variance of forecast results. Similar work that implements conventional ML algorithms for PV output power forecasting has also been pursued by others. SVM [62], BNN (Binarized Neural Network), SVR, and RT [63], and ANN [64], [65] are few examples. An important limitation of implementing the above methods arises when they are used for a PV plant located in a region where large and sudden weather variations over a short period are frequent. Both the forecast accuracy and stability of the prediction model suffer in these conditions. To reduce this effect, an ensemble technique is usually applied. The predictions of individual ML models are combined to make a model that has a characteristic of better generalization and robustness.

Ensemble of optimized ANN models is used in [54]. Here, they used the bagging technique to create diversified base estimators that can capture different aspects or characteristics of the PV power series. Statistical aggregation

strategy (median) was used to make the final prediction. The ensemble approach showed superior performance in comparison with the individual ANN models and a smart PM. Conceptually identical work, utilizing DT instead of ANN as a base estimator, was proposed by [66]. Two tree-based ensemble methods, i.e., RF and ET (Extra Trees) were compared with the SVM model. The significance of their ensemble approach is marginal at most in terms of improved accuracy over SVM, but much better forecast stability was achieved.

Similarly, an ensemble of DTs but with a boosting algorithm is implemented in [16]. Unlike bagging which randomly selects subsamples from the data set and train each estimator, boosting uses the same data set but each estimator learns from the last prediction sequentially. This ensemble model outperformed both the statistical model (ARMA) and SVM in all seasons of the year. Other literature that implements an ensemble approach for PV power forecasting includes the work of [67] and [68] where they used NN (Neural Network) and KDE (Kernel Density Estimation) respectively as base estimators.

A closer look at the above works and other pieces of literature, reveals an important gap and shortcoming of PV output power forecasting models based on the conventional ML algorithms. That is, even though these methods have better performance when compared to the statistical approaches, they still suffer from the problem of over-fitting and insufficient generalization to fully capture the highly non-linear characteristics of PV power. They are more effective in a region where the weather stays relatively stable. In other words, they result in a great performance when used only in the cases where the deterministic component (which is explained by the location of the sun) of the PV power is more dominant than the stochastic component (which is explained by the movement of clouds). In an area where the weather changes quite frequently over a short period (i.e. when both the stochastic and deterministic components are equally important), like Norway for example, there is a need for a better approach. Motivated by this, several DL models are proposed for short-term PV power forecasting. A concise summary of some of the work in the literature implementing conventional ML algorithms is given in Table 4 in Appendix A. The key finding, the forecast horizon, and the input parameters are important points to notice here. The term conventional ML algorithms here represent all AI algorithms (including shallow NN) except those based on deep learning networks.

The performance of three DL PV power forecasting models (LSTM, CNN, and CLSTM (Convolutional LSTM)) is compared in [69] for a 23.4 kW PV plant in Alice Springs. For a smaller data set (half-year data), the performance of all three models was poor. A possible explanation for this result can be the lack of adequate training samples to extract the spatial and temporal features required to make a good forecast. The LSTM model was able to extract the temporal features for an input data size of one and half years.

Hence, LSTM outperformed both CNN and the hybrid model (CLSTM). As the size of the data set is further increased (3 years), the CLSTM model which incorporates the advantages of both LSTM (temporal features) and CNN (spatial features) showed a superior performance. An important finding of this work is that the input sequence length (training dataset size) plays a huge role in the performance of a DL prediction model. A longer input sequence doesn't necessarily result in a better performance.

For a location where the weather changes (especially solar radiation) quite frequently over a short period, PV output power prediction becomes even a more challenging task. To address this issue, Zhang *et al.* [70] proposed Autoencoder-LSTM. The autoencoder reduced the uncertainties between perceptron mapping in the training process in response to complex weather conditions. The performance of this architecture was compared with conventional LSTM, FNN (Feedforward Neural Network), and PM models for various locations with different weather conditions. The autoencoder-LSTM based model resulted in the best accuracy for a day ahead PV power forecast on 15-minute time steps.

Another work involving the LSTM network, similar in principle to [70] but using an attention mechanism instead, was proposed in [38]. Here an ensemble of two LSTM neural networks is proposed for power and module temperature time series data. The attention mechanism is added so that the forecast model can adaptively focus only on input features that are more important to the current output. In comparison with other forecast models such as conventional LSTM, PM, ARIMAX, and MLP, the proposed method has the best accuracy for various forecast horizons. The performance improvement is more apparent in a forecast horizon longer than 15 minutes. Similar works involving DL algorithms have also been pursued by others. LSTM [71]–[73], CNN [74], [75], and GRU (Gated Recurrent Unit) [76]. A brief summary of some of the works in PV output power forecasting using DL algorithms is given in Table 5 in Appendix A. One apparent thing to observe from this table is that LSTM is the most widely used DL algorithm for PV output power forecasting.

4) HYBRID MODELS

Due to the complex nature of PV output power forecasting, a single ML/DL model is usually unable to fully capture the highly non-linear characteristics between the inputs and output. Inspired by this, there are many works in the literature where more than one technique or model is used for PV power forecasting. This approach is commonly known in the literature as a hybrid model. The definition of a hybrid model has not been consistent throughout different works of literature. Here it is used to represent a model which can be a combination of the above three PV power forecasting models or techniques from multiple other domains.

One way to form a hybrid model is to use an optimization algorithm to fine-tune the internal parameters of a conventional ML model. Pan *et al.* [15] used SVM as

a base forecasting model and I-ACO (Improved ACO) optimization technique to fine-tune the internal parameters of the model. The R^2 score, $RMSE$, and MAE of the hybrid model are significantly improved to 0.997, 0.1868 kW, and 0.1569 kW respectively. Conceptually identical work, utilizing SVM, is proposed in [20]. Here, the authors used GA instead of I-ACO for optimization purposes and significant performance improvement is achieved as compared with the base SVM model without optimization. The $RMSE$ and $MAPE$ reduced from 680.85 W and 100.47% to 11.23 W and 1.7052% respectively.

The performance of a hybrid model can be further improved by including time series decomposition algorithms on the above approach (i.e. using optimization techniques). A hybrid forecast model combining the SVM model, signal decomposition (WT), and optimization technique (PSO) is proposed in [18] for a 480 kW PV plant in Beijing, China. The WT decomposes and filters historical PV power measurements and weather parameters into subcomponents and these noise-treated subcomponents are used to train the SVM model. The PSO is used to optimally tune the internal parameters of the SVM. Their result demonstrated the adequacy of the proposed method in terms of higher forecasting accuracy. A similar approach that combines WPD and LSTM is implemented in the work of [12]. WPD decomposes the power series into four low (represent the power output trend) and high (represent varying and random output) frequency components. These components are trained by four LSTM neural networks and the final prediction is made by the linear weighting method. Their result showed significant forecast accuracy improvement over other methods such as LSTM, GRU, RNN, and MLP in all four seasons of the year. The $RMSE$ improved by 77.3% as compared with the forecast model involving LSTM without WPD.

Another technique where a hybrid model can be implemented to increase the forecast accuracy is to use a weather classification algorithm at the initial stage of the forecasting process. Based on the classification result, two approaches can be implemented. First, each weather cluster can be trained with a separate model and the final prediction will be the aggregate of the prediction from each model. Second, only a subset of the original data set that is similar to the forecast day will be used to train a common model for all clusters.

Implementing the first approach is the work of authors in [77]. They trained separate RF models for each cluster that is obtained using historical PV power data. Each model predicts on a new test sample, and the final prediction is made by integrating the prediction results from each cluster according to the assigned weight by the ridge regression algorithm. A PV forecasting model that implements the second approach is proposed by Jiang *et al.* [78]. Here, the authors used PCC (Pearson Correlation Coefficient) to measure the similarity between different days based on meteorological variables, and only samples similar to those from the target forecast day are selected as the training set to

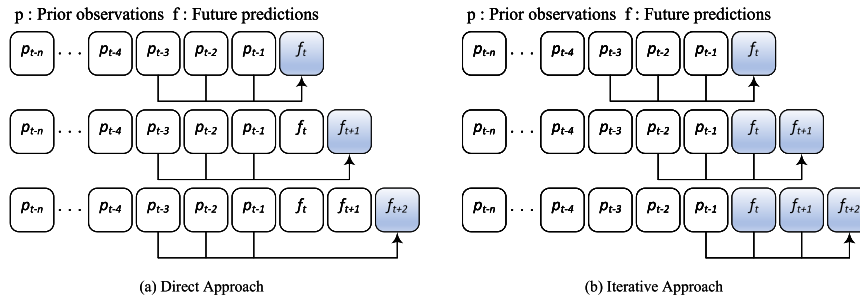


FIGURE 11. Multiple time step forecast [52], [79].

train GA-optimized ELM (Extreme Learning Machine). They were able to achieve an accurate forecast model that is also stable and computationally efficient. What can be seen from their work is that clustering analysis can effectively improve the prediction accuracy and stability of solar output power forecasting. A comprehensive summary of various hybrid PV output power forecasting models is given in Table 6 in Appendix A.

C. CLASSIFICATION BASED ON FORECAST TIME STEP

PV power forecasting models can be divided into two groups depending on the number of future time steps under consideration [52]. A forecast model that requires only the prediction of the next immediate time step is called a single time step forecast. On the other hand, a forecast problem that requires the prediction of more than one time-step is called a multiple time-step forecast. The number of time steps to look into the future depends on the specific application area where the forecasting will be implemented. The more time steps to be predicted into the future, the more challenging the task will be. This is the direct result of both the compounding nature of uncertainty on each forecast time step and the stochastic nature of input weather parameters.

1) SINGLE TIME STEP FORECAST

Most PV power forecasting models in the literature fall in this category. This involves forecasting a single time step into the future on a minute, 5 minutes, 15 minutes, or hourly basis. This category of forecast is relatively simple and accurate. It is more appropriate and ideal especially in areas where the weather changes frequently in a short interval.

2) MULTIPLE TIME STEP FORECAST

In some applications such as scheduling and dispatching, the prediction of the next immediate time step will not be sufficient. In such circumstances, multiple time step forecast becomes very important. The two widely used approaches for multiple time step forecast are direct and iterative methods (Fig. 11). In the direct approach, multiple models are developed for each time step and if it is an NN-based model, multiple output units representing each time step are used in the output layer. In contrast, the iterative approach uses a single model and appends the prediction from a previous

time step as an input for the next step. Although this approach is relatively less complex, it suffers from error accumulation from each previous prediction step. If the true observation of the power is available (which is the case in PV output power forecasting), this value can be used instead of the predicted value as part of the input for making the prediction on the next time step. This way it is possible to eliminate the error propagation problem.

Rana and Rahman [52] proposed a new hybrid approach for multiple time step forecast that integrates data re-sampling technique with an ML algorithm. For every prediction step, the time-series input data is first re-sampled to a new representation that is ideal for that particular step. Then a single step ML model is trained on the new re-sampled data avoiding the error at higher prediction steps. They have shown that their approach is a better alternative to the conventional approaches in the literature for multiple steps ahead prediction. The technique proposed in their work is very inspirational, but an important constraint can be when downsampling the original time series, a piece of important information describing the true variation in PV power can be lost, especially for higher time steps.

The performance of the three multiple time step forecast approaches discussed above was compared in [79] for a PV plant located at the University of Queensland in Brisbane, Australia. They used a heterogeneous ensemble model which dynamically assigns weight to the individual predictions based on the error information from the previous and current time steps. Their result shows that multiple-step forecast based on the approach proposed by [52] result in better accuracy. This approach achieved a 12.2% and 24.8% performance improvement in terms of MAE when compared with the direct and iterative approaches respectively.

D. CLASSIFICATION BASED ON AREAL SCALE

PV output power forecasting models can also be grouped into single-field and regional forecasts based on the number of PV plants and areal coverage considered in the study. All the PV power forecasting models discussed so far in this paper involve a single plant, thus belong to the single-field category. On the other hand, regional forecast involves a group of PV plants spread in a wider area.

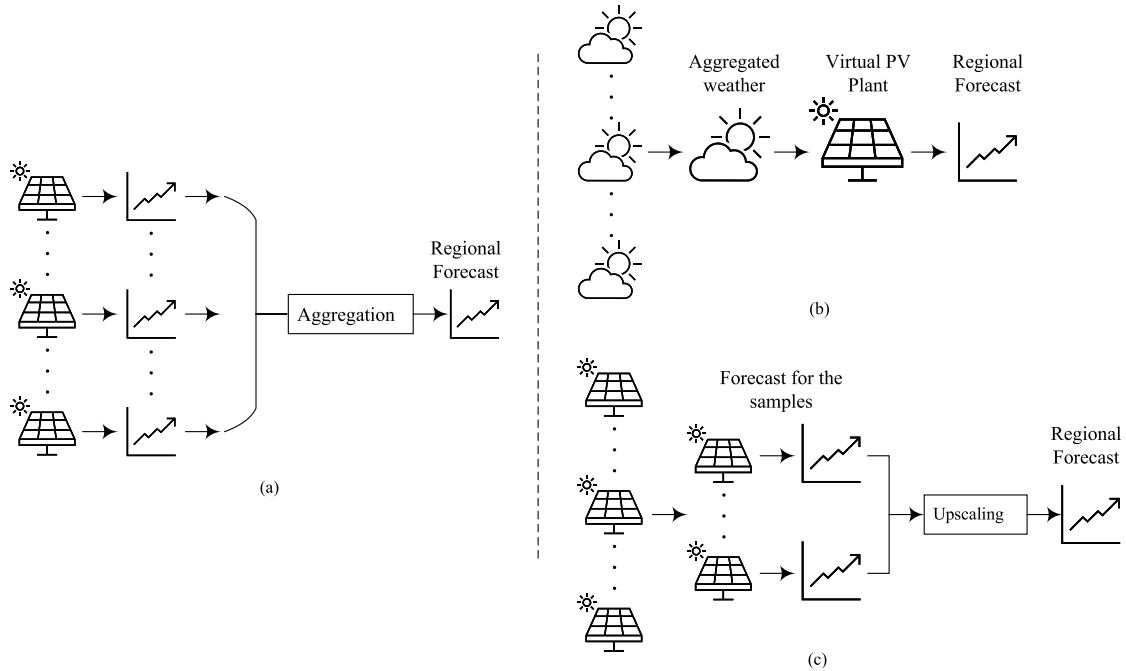


FIGURE 12. Regional forecast. (a) Bottom-up approach (b) Scaling based on weather aggregation and virtual PV plant and (c) Up-scaling based on sample PV plants [81].

This kind of forecast is particularly important for large grid operators for planning unit commitment, determining reserve requirements, contingency analysis, and energy storage dispatch [80].

1) SINGLE-FIELD FORECAST

Single-field forecast implies the prediction of solar output power from a single PV plant. This approach is more accurate and effective because the weather condition of the geographical area where the plant is located can be precisely represented by localized parameters.

2) REGIONAL FORECAST

Integrating PV plants spread over different locations to a grid that connects many regions is a major technical challenge for system operators that have to ensure the balance between production and consumption at all times. This has also major economic implications for energy trading companies operating across multiple regions (ex. Nord Pool in the case of Europe). It is therefore important to have a regional PV power forecast that aggregates the PV power prediction of individual plants located in various geographical areas. Two common approaches exist in the literature for regional PV power output forecasting: Bottom-up and Up-scaling approaches (Fig. 12).

- **Bottom-up Approach:** In this approach, prediction for each PV plant in the regional area under consideration is made first, and then the results are aggregated statistically to obtain the regional forecast. This is shown in Fig. 12 (a). This approach is not widely used practically as it requires large computational resources

and detailed knowledge about each plant in the region. Its application is limited to the cases where there are few PV plants in the area under consideration.

- **Up-scaling Approach:** This can be implemented in two ways. In the first approach, the various PV sites in the forecast region are replaced by a virtual power plant, and prediction is made directly at the regional level by using input data aggregated at a lower level (Fig. 12(b)). In the second approach, representative PV sites are sampled carefully and then the PV power forecast of such plants is re-scaled to obtain the regional power prediction according to the total capacity in the area (Fig. 12(c)). Choosing the right PV plant subsets that represent the spatial distribution of the overall region is equally important as the prediction itself in this approach [82].

A regional PV power forecast implementing a bottom-up approach for PV plants located in Luxembourg is proposed in the work of Koster *et al.* [83]. The hourly regional power generation is predicted using exogenous input (solar irradiance) obtained from NWP and a physical PV performance model. The prediction results from the proposed approach were compared with the true measured values and high accuracy is achieved with an average *RMSE* value of 7.4%.

Authors in [81] compared the performance of different ML-based regional forecasting methods using the up-scaling approach for PV plants located in the regions of Italy and the Netherlands. The forecast model based on an analog ensemble algorithm outperformed all other models including RF and GBDT. Aillaud *et al.* [84] used nationally aggregated

PV generation data and meteorological inputs from NWP for regional PV power output forecast in Germany. A forecast model based on CNN-LSTM gave the best performance when compared with other conventional and DL models. This is directly attributed to the ability of this hybrid model to extract both spatial (CNN) and temporal (LSTM) features from the input.

An up-scaling approach using a sample of reference plants is the practical and economic way of estimating PV output power on a regional level. However, this approach has its own challenges and introduces some uncertainties. These uncertainties arise from the sparsity of reference plants, the number of unknown plants, and most importantly on the difference in characteristics between the reference and unknown plants in the area. Saint-Drenan *et al.* [85] addressed and quantified these uncertainty issues for 366 PV plants in the regions of Germany. Their analysis shows that the *RMSE* decreased as the number of reference plants increased and when the number of unknown PV plants decreased. Another interesting finding of their work is that a variation of *RMSE* between 0.01 and 0.025 kW/kWp is measured depending on the choice of the reference plants.

To improve the performance of the regional PV output power forecast model using the up-scaling approach, authors in [86] used module orientation as an additional input variable. They determined the module orientation of the unknown PV plants using GIS-based (Geographic Information System) data sources and spatial interpolation techniques. Their approach resulted in an improvement of the *RMSE* by 5% when compared with the conventional up-scaling approaches.

The impact of local variability of weather parameters is ‘minimal’ in the regional PV power forecast due to the spatial smoothing effect. However, it can still be very challenging to make regional forecasts for some locations such as Norway. Norway is an elongated country where the southern and western part is fully exposed to the Atlantic ocean and experiences very different weather characteristics than the eastern and northern areas. Furthermore, the inland regions of Norway also have their unique climate. This significant weather variation across the county makes the task of determining sample PV plants that will be used in the up-scaling approach very difficult. Another major bottleneck that makes regional forecasting challenging in Norway can be the uneven distribution of PV plants across the country which again directly affects the choice of reference samples for the up-scaling approach. According to the data from NVE [4] for 2020, this uneven distribution is shown in Fig. 14 which is based on Fig. 13 which shows the Norwegian electricity market. These figures show the spatial distribution of PV plants across the country in different electricity power market zones. From these figures, it is possible to see that almost all ($\approx 98\%$) the PV capacity in Norway is concentrated in the south and west part of the country which would negatively contribute for PV power prediction on regional level.

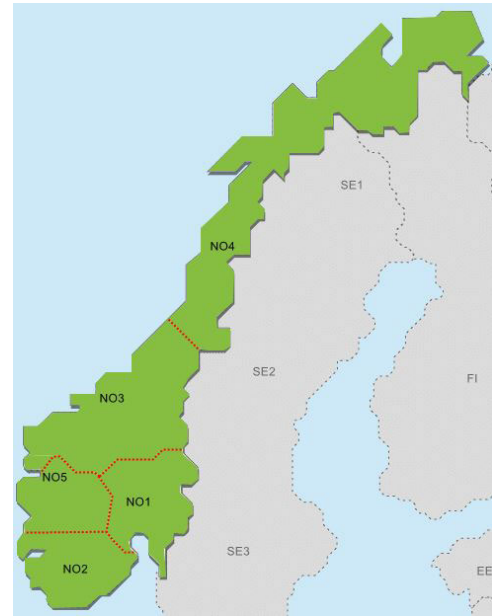


FIGURE 13. Norwegian electricity market.

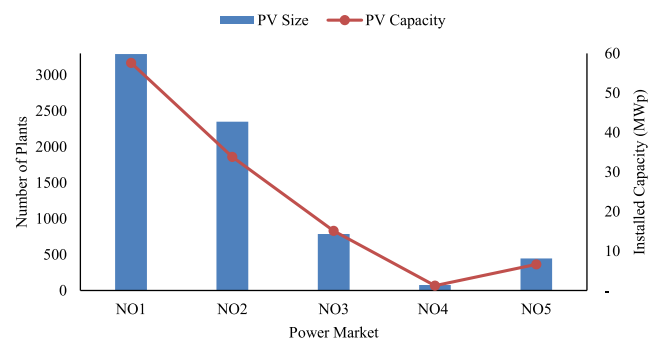


FIGURE 14. Distribution of PV plants in Norwegian electricity market.

E. CLASSIFICATION BASED ON APPROACH

In this classification category, the term ‘approach’ has been used to refer to the nature or type of output a particular PV power prediction model returns. Depending on this, forecast models can be grouped into two categories. Point forecast and Probabilistic forecast.

1) POINT FORECAST

Most works in the literature fall in this group. In this approach, the goal is to determine the precise value of the power production for each forecast step. The limitation of this approach is that it ignores information such as upper and lower bounds of possible forecasts that are very valuable for system operators and energy trading companies [55].

2) PROBABILISTIC FORECAST

The probabilistic forecast provides prediction intervals in addition to precise values with which the forecast is expected to fall with some predefined confidence level or probability. The additional information about the uncertainty of the prediction is very important to decision-makers such as

PV-based electricity market operators [87]. This knowledge is used to provide a precise generation schedule for the day-ahead and real-time market.

A day ahead probabilistic PV output power forecasting model based on QRNN (Quartile Regression Neural Network) is proposed in [87] for a 7 MW PV plant in South-East of Spain. The internal settings of the NN and the size of informative inputs are optimized using GA algorithm. Their analysis shows that a point forecast corresponding to the QRNN for quantile 0.5 has lower *RMSE* compared with a persistence reference model and with the best MLP model.

Authors in [54] used the bootstrap technique to quantify uncertainties associated with each forecast. This technique allowed predictions to be made with a wider prediction interval with a confidence level of 84%. Such kind of forecast is very important for proper planning, scheduling, and generation control of the available energy sources. It is also crucial information for ensuring power system reliability and for an efficient energy market operation.

Thus far, this work has attempted to claim that solar electricity can be a good and economical alternative to conventional ways of generating electricity even in the colder climates of the Nordic regions and Canada. The cold climate can provide a unique opportunity in terms of seasonal matching of generation and consumption, and improving the efficiency of the PV modules. However, the cold climate also brings an important challenge to the operation of PV plants in these regions, such as the soiling effect due to snow. Frequent and abrupt change of weather parameters over a short period is also another important challenge that is worth considering in these regions. Despite these limitations, this work argued that accurate PV output power forecasting that considers the needs of this region's climate can improve the adoption and deployment of PV plants further.

Section II addressed the basic procedures and steps that are commonly followed in the design of AI-based PV power forecasting models. The discussion of the common AI algorithms that are widely used in relation to PV power forecasting was also part of this section. The different environmental parameters that directly affect PV power forecasting are discussed in section III. A more focused discussion on the case of PV in a cold climate is covered in section IV. Both the merits and demerits of the cold climate are thoroughly covered. It is argued that the advantages that come along with the cold climate outweigh the challenges and a PV power forecasting model that takes into consideration this can result in a more efficient system. And finally, the classification of PV power forecasting models based on multiple criteria was presented in section V. In the two sections that follow, a thorough discussion on the main findings from this review paper, together with a final comment and conclusion is provided.

VI. DISCUSSION

In recent years, solar electricity is gaining huge interest in the cold and high latitude climate regions such as the

Nordic countries and Canada. Although the cold climate offers many opportunities, the unique climate in these regions also makes the task of integrating PV into the existing grid more challenging. Accurate PV output power forecasting can overcome this limitation and make PV power an equal contributor to the energy mix in these regions. This paper presented a comprehensive review and evaluation of the state of the art in the use of AI/ML algorithms for PV output power forecasting in the Nordic region. The main goals of this review paper were to identify relevant input parameters and evaluate their importance in PV power forecasting and assess the performance of various PV output power forecasting models. The analysis of the impact of the forecast time step, forecast horizon, and areal scale on the performance of a prediction model was also the point of interest in this review. This review was also aimed at qualitative assessment of some of the techniques that are used to enhance the accuracy and stability of a prediction model. The following are some of the important issues that are discussed in this review work.

The selection of weather parameters as an input to a PV output power forecasting is a vital step to achieving a good model. Proper feature selection and using an optimum number of informative inputs ensure a high-performing model with minimal computational resource requirement. This is one of the important required characteristics of a PV forecast model for practical deployment. Since the Nordic region's weather, especially Norway's, vary very frequently over a short period, it results in fast and abrupt changes in PV output power. This study indicates that such variations can only be precisely accounted for by including solar irradiance and cloud coverage information as inputs in the prediction model.

The impact of the soiling effect due to the accumulation of snow on PV modules is another significant issue that should be taken into consideration in the design of a PV output power forecast model in the Nordic region. This effect can vary from partial to total obstruction of solar irradiance reaching the surface of the PV module resulting in a reduced or zero generation of PV power. In general, prior works are limited to the calculation and prediction of yield reduction due to snow cover. To the author's knowledge, no previous study has investigated how to directly incorporate the impact of snow on PV output power forecasting.

This study suggests two possible strategies that can be used to incorporate this effect in the design of a PV output power forecast model. The first strategy is to consider snow depth and snow cover as additional inputs in the initial stages of the design process of a forecast model. In this way, the algorithm learns how to associate snow depth and snow cover with PV output power. In the second approach, a strategy can be designed in such a way that during the days with significant snow cover, a correction term can be added to the forecast output that is obtained without taking snow accumulation as an input feature. These strategies provide new insight into the problem and can serve as the first step towards a more

TABLE 4. Conventional ML models for PV output power forecasting.

Author	Year	Location	Horizon	Inputs ^b	Method	Key Findings
[10]	2019	Korea	1 day	GHI, SD, CC, SH, RH, Precipitation, Temp, and WS	ANN and SVM	The effect of PV power prediction errors is characterized on energy storage system-based PV power trading markets
[16]	2018	Oregon, USA	1 day	GHI and Temp	GBDT	High accuracy, stable performance, and strong model interpretability
[36]	2018	Jiangsu, China	5 minutes	GHI, Temp, WS and WD	SVR and ARIMA	Weather classification and pattern recognition algorithms increased the forecasting accuracy
[54]	2019	Amman, Jordan	1 day	WS, WD, AP, RH, Precipitation, Tempe, GHI and DHI	Ensemble of ANN	Accurate forecast result and quantification of uncertainties affecting prediction
[58]	2019	Berkeley, USA	1 day	GHI, CC and Temp	SARIMA and ANN	The parallel architecture increased the forecast accuracy and resiliency of the system
[59]	2019	Qatar	1 hour	Temp, RH, GHI, and WS	ANN	Relief feature selection procedure enabled the ANN forecast model to outperform other models
[65]	2018	Amman, Jordan	1 day	solar irradiance lag features (window size = 5)	ANN	Prediction by using only historical irradiance measurements can still result in a very good PV power forecast model
[88]	2019	Milan, Italy	1 day	Temp, GHI, WS, WD,	ANN	Very good and stable prediction performance
[89]	2018	Odisha, India	15 minutes	GHI and Temp	ELM	ELM forecasting technique performs well for different time steps (15,30, and 60 minutes)
[90]	2020	Florida, USA	1 hour	PV Power lag features (window size = 5)	EELM	Simple, fast, and efficient prediction model without meteorological inputs

profound understanding of PV output power forecasting in Nordic regions.

The success of the two proposed approaches above however depends greatly on the precise measurement of snow depth and snow cover on the surface of the PV modules. This can be of major concern and the authors believe that this information may be currently unavailable from most or all PV plants in the Nordic region. This, therefore, suggests a definite need and effort to be directed towards installing a measurement station for monitoring the snow effect. It is also very important to bear in mind that these approaches require addressing the issues related to the snow clearing process due to snow sliding and snow melting. More research is also needed to understand and consider the reflective property of snow and ice (albedo effect) in PV output power prediction in Nordic regions.

The forecast time step and forecast horizon are also important parameters that directly affect the accuracy of a prediction model. These parameters should be chosen considering the stability of the weather condition where the plant is located. This study shows that for a relatively stable weather condition, a prediction model with a forecast time step of up to 1 hour and a forecast horizon of up to one week or more can be designed without significantly compromising the accuracy. On the other hand, if the weather condition of the PV plant's location is frequently changing over a short period, such as in Norway, a prediction model with a forecast time step and a forecast horizon of more than one hour could lead to substantial forecast error.

This review also shows that the choice of a particular type of ML algorithm to apply for PV output power forecasting depends on the weather condition of the area where the plant is located similar to the forecast time step and horizon. For stable weather conditions, the deterministic component (which is explained by the movement of the

sun) of the PV output power is more dominant than the stochastic component (which is explained by the movement of cloud). In such cases, conventional ML algorithms such as RF and SVM can result in a sufficiently good-performing prediction model. However, in areas where the stochastic component is equally important as the deterministic component, the conventional ML algorithms are found to be mostly inadequate. In cases like this, DL algorithms such as LSTM and CNN have been implemented to overcome this inherent limitation. It was found in our analysis that such approaches can fully capture highly complex input-output relationships and result in a high-performing prediction model.

This research has identified several techniques and approaches that have been implemented to improve the performance (accuracy and stability) of a PV power prediction model. One such technique is the use of an ensemble approach. Ensemble techniques such as bagging and boosting which aggregate the prediction ability of multiple base estimators are found to improve the poor generalizing ability and the over-fitting problem of conventional ML methods considerably. Such an approach can be particularly beneficial to enhance the performance of conventional ML-based PV power forecasting models in a region with highly varying weather conditions.

Time-series decomposition is another technique that is mostly applied to improve the performance of a PV power forecast model. WT and WPD are the two commonly utilized and efficient time-series decomposition tools that are used as preprocessing steps to treat time series data against noise. In this way, the algorithm can be trained on the noise-treated data which is found to improve the training speed and prediction accuracy substantially.

This research work has also shown that for PV output power forecasting models based on DL algorithms,

TABLE 5. Deep learning models for PV output power forecasting.

Author	Year	Location	Horizon	Inputs ^b	Method	Key Findings
[14]	2020	China	3 hours	Lag inputs of power (Window size = 8)	BI-LSTM	Accurate prediction in the absence of solar irradiance measurement and other weather parameters
[17]	2020	Germany	1 day	Temp, AP, RH, CC, WS, Precipitation	LSTM	Accurate and transferable predictive model is achieved using publicly available datasets
[34]	2020	Arcadia, Florida	1 day	Temp, WS, RH, ST and GHI	LSTM	Synthetic weather forecast (translating fixed sky categories into dynamic and quantitative irradiance values) enabled significant improvement in forecast accuracy
[38]	2019	China	1 hour	Temp and power time series data	Ensemble of LSTMs	The attention mechanism embedded with the LSTM network enabled to adaptively focus only on the important features at every forecast period, and hence improving accuracy significantly
[69]	2019	Australia	1 day	GHI, DHI, Temp, WD, WS, and RH	LSTM, CNN, and C-LSTM	DL models can result in accurate forecast results, but the input sequence length is very important
[70]	2020	USA and Italy	1 day	Temp, RH, WS, CC and GHI	AE-LSTM	The AE-LSTM resulted in high accuracy during uncertain and unstable weather conditions while the PM gave acceptable prediction on clear sky days
[91]	2019	China	5 minutes	Temp, RH, AP, Precipitation, GHI, DHI and DNI	LSTM	The adaptive hyperparameter adjustment implemented to assist the LSTM enabled to obtain a high performing model under various weather conditions
[92]	2020	Nevada, USA	1 day	DNI, Temp and Lag inputs	LSTM-RNN	Time correlation modification is used to update the forecast output from LSTM-RNN based on the specific pattern in the forecast day. This approach resulted high forecast accuracy.
[93]	2020	Australia	1 day	GHI, DHI, WS, Temp, RH, RF	CNN (ResNet and DenseNet)	Accurate point and probabilistic forecasting results are obtained due to the deep layers of the CNN network which enabled higher representation capabilities
[94]	2020	Korea	Mid-and long-term	Economic factors, Policy factors and Environmental factors	Ensemble of ARIMA, VAR and LSTM	Proper feature selection enabled to prevent overfitting and increase predictive power of the ensemble model that is aggregated based on CNN

TABLE 6. Hybrid models for PV output power forecasting.

Author	Year	Location	Horizon	Inputs ^b	Method	Key Findings
[12]	2020	Australia	1 hour	GHI, DHI, Temp, WS and RH	WPD + LSTM	Accurate and stable performance under high weather variation condition. Linear combination method enhanced the performance
[15]	2020	Australia	1 hour	GHI, DHI, Temp, WD and RH	SVM + I-ACO	Improved forecast accuracy especially during the peak power and nighttime hours
[18]	2018	Beijing, China	1 day	GHI, Temp, CC, RH, AP and WS	WT + PSO + SVM	Improved accuracy due to better mapping between the meteorological inputs and the power as the result of WT
[20]	2019	Australia	1 day	Tempe and GHI	GA + SVM	Increased accuracy due to GA optimization
[33]	2020	Cyprus and USA	1 day	GHI, Elevation angle, RH, Tempe, WS, WD and Solar azimuth	NN+ Weather clustering + postprocessing	Stable and robust method validated for different locations
[39]	2018	Australia	1 day	GHI, DHI, Temp, RH, and WS	k-means + GRA + Elman NN	Comparatively high prediction accuracy
[52]	2020	Australia	3 hours	Lag inputs of power (window size of 6)	Univariate ML Models + Data re-sampling	High level of prediction accuracy for multiple steps ahead forecast
[78]	2020	Australia	1 day	Tempe, GHI and DHI	SDA + GA + ELM	Less computational resource requirement and stable prediction performance due to SDA and GA
[95]	2020	China	1 day	GHI and Temp	Dendritic NN + WT	Improved computational efficiency and accuracy due to the high-low frequency decomposition
[96]	2019	India	1 hour	Lag inputs of power	SARIMA + RVFL + Discrete WT	Improved performance during cloudy/rainy sky conditions when fast variations are frequent

^bGHI: Global horizontal irradiance, SD: Solar duration, CC: Cloud coverage, SH: Solar hour, RH: Relative humidity, Temp: Temperature, WS: Wind speed, WD: Wind direction, AP: Air pressure, RF: Rain fall, DHI: Diffused horizontal irradiance, DNI: Direct normal irradiance

auto-encoder and attention mechanism techniques can be included to improve performance. These approaches denoise the input time series data and ensure that the algorithm adaptively focuses only on the input features that are more important to the current prediction and avoid interference from other features. This enables the prediction model to handle uncertainties in the training process and deal with complex and unstable environmental conditions.

Weather clustering is another approach that this study has identified as a useful tool that has enormous potential in improving the performance of a prediction model. This approach ensures that the model is trained and subsequent prediction is made by considering data only on days which have similar characteristics as the target forecast day. This method not only results in higher accuracy and stability but also needs small computational resources which is a required

attribute in the practical deployment of a PV power prediction model.

VII. CONCLUSION

In general, the findings in this research suggest the need to adapt the available AI/ML algorithms in the literature for the design of PV output power forecasting to match the unique climate of the Nordic region. More emphasis should be given to the impact of the soiling effect due to snow during the winter season in addition to the highly varying weather conditions. This ensures an accurate PV power forecasting model which potentially is a better and economical alternative to other methods to stabilize a grid with high penetration of PV. This work therefore can serve as a groundwork for further research into the design of a high-performing AI/ML PV power forecast model and can contribute in several ways to the limited current literature available for PV in the Nordic region. Although the focus of this work has been in the Nordic regions, it can easily be extended and be applicable in other areas that have similar weather characteristics, for example, Canada.

APPENDIX A

SUMMARY OF REVIEWED PAPERS

See Tables 4–6.

REFERENCES

- [1] U.S. Energy Information Administration. (2019). *International Energy Outlook 2019 With Projections to 2050*. [Online]. Available: <https://www.eia.gov/outlooks/ieo/pdf/ieo2019.pdf>
- [2] U.Nations. (2020). *The Sustainable Development Goals Report 2020*. [Online]. Available: <https://unstats.un.org/sdgs/report/2020/>
- [3] REN21. (2020). *Renewables 2020 Global Status Report*. [Online]. Available: <https://www.ren21.net/gsr-2020/>
- [4] NVE. (2021). *40 MW of Solar Power was Installed in 2020*. Accessed: Oct. 10, 2021. [Online]. Available: <https://www.nve.no/nytt-franve/nyheter-energi/40-mw-solkraft-ble-installert-i-2020/>
- [5] N. W. Resources and E. Directorate. (2017). *Power Market Analysis 2017–2030*. [Online]. Available: <https://www.nve.no/om-nve/publikasjoner-og-bibliotek/publikasjoner/>
- [6] Solenergiklyngen and FME SUSOLTECH. (2020). *Veikart for Den Norske Solkraftbransjen Mot 2030*. [Online]. Available: <https://susoltech.no/files/2020/12/Solenergiklyngen-veikart-F41.pdf>
- [7] A. H. Al-Badi, “Measured performance evaluation of a 1.4 kW grid connected desert type PV in Oman,” *Energy Sustain. Develop.*, vol. 47, pp. 107–113, Dec. 2018.
- [8] E. Andenæs, B. P. Jelle, K. Ramlo, T. Kolås, J. Selj, and S. E. Foss, “The influence of snow and ice coverage on the energy generation from photovoltaic solar cells,” *Sol. Energy*, vol. 159, pp. 318–328, Jan. 2018.
- [9] T. S. Salem, K. Kathuria, H. Ramampiaro, and H. Langseth, “Forecasting intra-hour imbalances in electric power systems,” in *Proc. AAAI Conf. Artif. Intell.*, 2019, vol. 33, no. 1, pp. 9595–9600.
- [10] K. Y. Bae, H. S. Jang, B. C. Jung, and D. K. Sung, “Effect of prediction error of machine learning schemes on photovoltaic power trading based on energy storage systems,” *Energies*, vol. 12, no. 7, p. 1249, Apr. 2019.
- [11] V. Koestler, M. E. Henriksen, M. Sidelnikova, C. A. Veie, I. H. Magnussen, J. Hole, M. Wold, I. Haddeland, E. Skaansar, and A. M. Østenby, “Det svinger mer med fornybar strøm,” Norwegian Water Resour. Energy Directorate, Oslo, Norway, Tech. Rep. 44/2020, 2020.
- [12] P. Li, K. Zhou, X. Lu, and S. Yang, “A hybrid deep learning model for short-term PV power forecasting,” *Appl. Energy*, vol. 259, Feb. 2020, Art. no. 114216.
- [13] U. Cali and C. Lima, “Energy informatics using the distributed ledger technology and advanced data analytics,” in *Cases on Green Energy and Sustainable Development*. Hershey, PA, USA: IGI Global, 2020, pp. 438–481.
- [14] H. Sharadga, S. Hajimirza, and R. S. Balog, “Time series forecasting of solar power generation for large-scale photovoltaic plants,” *Renew. Energy*, vol. 150, pp. 797–807, May 2020.
- [15] M. Pan, C. Li, R. Gao, Y. Huang, H. You, T. Gu, and F. Qin, “Photovoltaic power forecasting based on a support vector machine with improved ant colony optimization,” *J. Cleaner Prod.*, vol. 277, Dec. 2020, Art. no. 123948.
- [16] J. D. Wang, P. Li, R. Ran, Y. B. Che, and Y. Zhou, “A short-term photovoltaic power prediction model based on the gradient boost decision tree,” *Appl. Sci.*, vol. 8, no. 5, p. 689, May 2018.
- [17] N. Maitanova, J.-S. Telle, B. Hanke, M. Grottko, T. Schmidt, K. V. Maydell, and C. Agert, “A machine learning approach to low-cost photovoltaic power prediction based on publicly available weather reports,” *Energies*, vol. 13, no. 3, p. 735, Feb. 2020.
- [18] A. T. Eseye, J. Zhang, and D. Zheng, “Short-term photovoltaic solar power forecasting using a hybrid wavelet-PSO-SVM model based on SCADA and meteorological information,” *Renew. Energy*, vol. 118, pp. 357–367, Apr. 2018.
- [19] W. Jidong, R. Ran, S. Zhilin, and J. Sun, “Short-term photovoltaic power generation forecasting based on environmental factors and GA-SVM,” *J. Elect. Eng. Technol.*, vol. 12, no. 1, pp. 64–71, 2017.
- [20] V. D. William, E. Jamei, G. S. Thirunavukkarasu, M. Seyedmahmoudian, T. K. Soon, B. Horan, S. Mekhilef, and A. Stojcevski, “Short-term PV power forecasting using hybrid GASVM technique,” *Renew. Energy*, vol. 140, pp. 367–379, Feb. 2019.
- [21] N. K. Chauhan and K. Singh, “A review on conventional machine learning vs deep learning,” in *Proc. Int. Conf. Comput., Power Commun. Technol. (GUCON)*, Sep. 2018, pp. 347–352.
- [22] M. Awad and R. Khanna, *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*. Berlin, Germany: Springer, 2015.
- [23] A. C. Müller and S. Guido, *Introduction to Machine Learning With Python: A Guide for Data Scientists*. Newton, MA, USA: O’Reilly Media, 2016.
- [24] S. Raschka and V. Mirjalili, *Python Machine Learning: Machine Learning and Deep Learning With Python, Scikit-Learn, and TensorFlow*. Birmingham, U.K.: Packt Publishing Ltd, 2017.
- [25] S. Albawi, T. A. Mohammed, and S. Al-Zawi, “Understanding of a convolutional neural network,” in *Proc. Int. Conf. Eng. Technol. (ICET)*, Aug. 2017, pp. 1–6.
- [26] S.-H. Feng, J.-Y. Xu, and H.-B. Shen, “Artificial intelligence in bioinformatics: Automated methodology development for protein residue contact map prediction,” in *Biomedical Information Technology*. Amsterdam, The Netherlands: Elsevier, 2020, pp. 217–237.
- [27] L. Lu, Y. Zheng, G. Carneiro, and L. Yang, “Deep learning and convolutional neural networks for medical image computing,” *Adv. Comput. Vis. pattern Recognit.*, vol. 10, pp. 978–983, Jan. 2017.
- [28] S. Khan, H. Rahmani, S. A. A. Shah, and M. Bennamoun, “A guide to convolutional neural networks for computer vision,” *Synth. Lectures Comput. Vis.*, vol. 8, no. 1, pp. 1–207, 2018.
- [29] F. M. Bianchi, E. Maiorino, M. C. Kampffmeyer, A. Rizzi, and R. Jenssen, *Recurrent Neural Networks for Short-Term Load Forecasting: An Overview and Comparative Analysis*, 2017.
- [30] M. Kuzlu, U. Cali, V. Sharma, and O. Guler, “Gaining insight into solar photovoltaic power generation forecasting utilizing explainable artificial intelligence tools,” *IEEE Access*, vol. 8, pp. 187814–187823, 2020.
- [31] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bannetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, “Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI,” *Inf. Fusion*, vol. 58, pp. 82–115, Jun. 2020.
- [32] U. K. Das, K. S. Tey, M. Seyedmahmoudian, S. Mekhilef, M. Y. I. Idris, W. Van Deventer, B. Horan, and A. Stojcevski, “Forecasting of photovoltaic power generation and model optimization: A review,” *Renew. Sustain. Energy Rev.*, vol. 81, pp. 912–928, Jan. 2018.
- [33] S. Theocharides, G. Makrides, A. Livera, M. Theristis, P. Kaimakis, and G. E. Georgioui, “Day-ahead photovoltaic power production forecasting methodology based on machine learning and statistical post-processing,” *Appl. Energy*, vol. 268, Jun. 2020, Art. no. 115023.
- [34] M. S. Hossain and H. Mahmood, “Short-term photovoltaic power forecasting using an LSTM neural network and synthetic weather forecast,” *IEEE Access*, vol. 8, pp. 172524–172533, 2020.
- [35] L. Liu, Y. Zhao, D. Chang, J. Xie, Z. Ma, Q. Sun, H. Yin, and R. Wennersten, “Prediction of short-term PV power output and uncertainty analysis,” *Appl. Energy*, vol. 228, pp. 700–711, Oct. 2018.

- [36] F. Mei, Y. Pan, K. Zhu, and J. Zheng, "A hybrid online forecasting model for ultrashort-term photovoltaic power generation," *Sustainability*, vol. 10, no. 3, p. 820, Mar. 2018.
- [37] A. I. Khalyasmaa, S. A. Eroshenko, V. A. Tashchilin, H. Ramachandran, T. P. Chakravarthi, and D. N. Butusov, "Industry experience of developing day-ahead photovoltaic plant forecasting system based on machine learning," *Remote Sens.*, vol. 12, no. 20, p. 3420, Oct. 2020.
- [38] H. Zhou, Y. Zhang, L. Yang, Q. Liu, K. Yan, and Y. Du, "Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism," *IEEE Access*, vol. 7, pp. 78063–78074, 2019.
- [39] P. Lin, Z. Peng, Y. Lai, S. Cheng, Z. Chen, and L. Wu, "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.*, vol. 177, pp. 704–717, Dec. 2018.
- [40] P. Manganiello, J. Govaerts, I. T. Horvath, M. G. Chowdhury, G. H. Yordanov, H. Goverde, B. Aldalali, I. Beausoleil-Morrison, S. Valkealahti, K. Lappalainen, and J. Poortmans, "Tuning electricity generation throughout the year with PV module technology," *Renew. Energy*, vol. 160, pp. 418–427, Nov. 2020.
- [41] A. Kahl, J. Dujardin, and M. Lehning, "The bright side of PV production in snow-covered mountains," *Proc. Nat. Acad. Sci. USA*, vol. 116, no. 4, pp. 1162–1167, Jan. 2019.
- [42] M. S. Adaramola and E. E. Vågnes, "Preliminary assessment of a small-scale rooftop PV-grid tied in Norwegian climatic conditions," *Energy Convers. Manage.*, vol. 90, pp. 458–465, 2015.
- [43] G. Becker, B. Schiebelsberger, W. Weber, C. Vodermyer, M. Zehner, and G. Kummerle, "An approach to the impact of snow on the yield of grid connected PV systems," in *Proc. Eur. PVSEC*. Citeseer, 2006.
- [44] L. B. Bosman and S. B. Darling, "Performance modeling and valuation of snow-covered PV systems: Examination of a simplified approach to decrease forecasting error," *Environ. Sci. Pollut. Res.*, vol. 25, no. 16, pp. 15484–15491, Jun. 2018.
- [45] E. Mohammadi, J. Khodabakhsh, G. Moschopoulos, and R. Fadaeinedjad, "A study on the performance of PV modules in snowy conditions considering orientation of modules," in *Proc. IEEE Can. Conf. Electr. Comput. Eng. (CCECE)*, Aug. 2020, pp. 1–4.
- [46] B. Marion, R. Schaefer, H. Caine, and G. Sanchez, "Measured and modeled photovoltaic system energy losses from snow for Colorado and Wisconsin locations," *Sol. Energy*, vol. 97, pp. 112–121, Nov. 2013.
- [47] B. Hashemi, A.-M. Cretu, and S. Taheri, "Snow loss prediction for photovoltaic farms using computational intelligence techniques," *IEEE J. Photovolt.*, vol. 10, no. 4, pp. 1044–1052, Jul. 2020.
- [48] R. E. Pawluk, Y. Chen, and Y. She, "Photovoltaic electricity generation loss due to snow—A literature review on influence factors, estimation, and mitigation," *Renew. Sustain. Energy Rev.*, vol. 107, pp. 171–182, Jun. 2019.
- [49] R. W. Andrews, A. Pollard, and J. M. Pearce, "The effects of snowfall on solar photovoltaic performance," *Sol. Energy*, vol. 92, pp. 84–97, Jun. 2013.
- [50] R. Ahmed, V. Sreeram, Y. Mishra, and M. D. Arif, "A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization," *Renew. Sustain. Energy Rev.*, vol. 124, May 2020, Art. no. 109792.
- [51] B. Dimd, S. Völler, O. Midtgård, and T. Zenebe, "Ultra-short-term photovoltaic output power forecasting using deep learning algorithms," in *Proc. IEEE MELECON*, 2022.
- [52] M. Rana and A. Rahman, "Multiple steps ahead solar photovoltaic power forecasting based on univariate machine learning models and data re-sampling," *Sustain. Energy, Grids Netw.*, vol. 21, Mar. 2020, Art. no. 100286.
- [53] K. Mahmud, S. Azam, A. Karim, S. Zobaed, B. Shanmugam, and D. Mathur, "Machine learning based PV power generation forecasting in Alice springs," *IEEE Access*, vol. 9, pp. 46117–46128, 2021.
- [54] S. Al-Dahidi, O. Ayadi, M. Alrbai, and J. Adee, "Ensemble approach of optimized artificial neural networks for solar photovoltaic power prediction," *IEEE Access*, vol. 7, pp. 81741–81758, 2019.
- [55] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Sol. Energy*, vol. 136, pp. 78–111, Oct. 2016.
- [56] S. Sobri, S. Koohi-Kamali, and N. A. Rahim, "Solar photovoltaic generation forecasting methods: A review," *Energy Convers. Manage.*, vol. 156, pp. 459–497, Jan. 2018.
- [57] V. Kushwaha and N. M. Pindoriya, "Very short-term solar PV generation forecast using SARIMA model: A case study," in *Proc. 7th Int. Conf. Power Syst. (ICPS)*, Dec. 2017, pp. 430–435.
- [58] E. Vrettos and C. Gehbauer, "A hybrid approach for short-term PV power forecasting in predictive control applications," in *Proc. IEEE Milan PowerTech*, Jun. 2019, pp. 1–6.
- [59] A. Khandakar, M. E. H. Chowdhury, M. K. Kazi, K. Benhmed, F. Touati, M. Al-Hitmi, and S. P. Gonzales, "Machine learning based photovoltaics (PV) power prediction using different environmental parameters of Qatar," *Energies*, vol. 12, no. 14, p. 2782, 2019.
- [60] S. Theocharides, G. Makrides, G. E. Georghiou, and A. Kyprianou, "Machine learning algorithms for photovoltaic system power output prediction," in *Proc. IEEE Int. Energy Conf.*, Jun. 2018, pp. 1–6.
- [61] R. F. Widjaja, H. Panamtash, Q. Zhou, and D. Li, "Solar power forecasting with model selection analysis," in *Proc. Clemson Univ. Power Syst. Conf. (PSC)*, Sep. 2018, pp. 1–6.
- [62] W. Buwei, C. Jianfeng, W. Bo, and F. Shuanglei, "A solar power prediction using support vector machines based on multi-source data fusion," in *Proc. Int. Conf. Power Syst. Technol.*, Nov. 2018, pp. 4573–4577.
- [63] S. Theocharides, M. Theristis, G. Makrides, M. Kynigos, C. Spanias, and G. E. Georghiou, "Comparative analysis of machine learning models for day-ahead photovoltaic power production forecasting," *Energies*, vol. 14, no. 4, p. 1081, Feb. 2021.
- [64] F. Touati, A. Khandakar, M. E. Chowdhury, S. Antonio, Jr., C. K. Sorino, and K. Benhmed, "Photo-voltaic (PV) monitoring system, performance analysis and power prediction models in Doha, Qatar," in *Renewable Energy-Technologies and Applications*. London, U.K.: IntechOpen, 2020.
- [65] M. H. Alomari, J. Adee, and O. Younis, "Solar photovoltaic power forecasting in Jordan using artificial neural networks," *Int. J. Elect. Comput. Eng.*, vol. 8, no. 1, p. 497, 2018.
- [66] M. W. Ahmad, M. Mourshed, and Y. Rezgui, "Tree-based ensemble methods for predicting PV power generation and their comparison with support vector regression," *Energy*, vol. 164, pp. 465–474, Dec. 2018.
- [67] M. Rana, I. Koprinska, and V. G. Agelidis, "Univariate and multivariate methods for very short-term solar photovoltaic power forecasting," *Energy Convers. Manage.*, vol. 121, pp. 380–390, Aug. 2016.
- [68] M. Lotfi, M. Javadi, G. J. Osório, C. Monteiro, and J. P. S. Catalão, "A novel ensemble algorithm for solar power forecasting based on kernel density estimation," *Energies*, vol. 13, no. 1, p. 216, Jan. 2020.
- [69] K. Wang, X. Qi, and H. Liu, "A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network," *Appl. Energy*, vol. 251, Oct. 2019, Art. no. 113315.
- [70] Y. Zhang, C. Qin, A. K. Srivastava, C. Jin, and R. Sharma, "Data-driven day-ahead PV estimation using autoencoder-LSTM and persistence model," *IEEE Trans. Ind. Appl.*, vol. 56, no. 6, pp. 7185–7192, Nov./Dec. 2020.
- [71] M. Abdel-Nasser and K. Mahmoud, "Accurate photovoltaic power forecasting models using deep LSTM-RNN," *Neural Comput. Appl.*, vol. 31, no. 7, pp. 2727–2740, Oct. 2017.
- [72] M. Gao, J. Li, F. Hong, and D. Long, "Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM," *Energy*, vol. 187, Nov. 2019, Art. no. 115838.
- [73] F. Harrou, F. Kadri, and Y. Sun, "Forecasting of photovoltaic solar power production using LSTM approach," in *Advanced Statistical Modeling, Forecasting, and Fault Detection in Renewable Energy Systems*. London, U.K.: IntechOpen, 2020, p. 3.
- [74] C.-J. Huang and P.-H. Kuo, "Multiple-input deep convolutional neural network model for short-term photovoltaic power forecasting," *IEEE Access*, vol. 7, pp. 74822–74834, 2019.
- [75] H. Aprillia, H.-T. Yang, and C.-M. Huang, "Short-term photovoltaic power forecasting using a convolutional neural network–salp swarm algorithm," *Energies*, vol. 13, no. 8, p. 1879, 2020.
- [76] K. Yan, H. Shen, L. Wang, H. Zhou, M. Xu, and Y. Mo, "Short-term solar irradiance forecasting based on a hybrid deep learning methodology," *Information*, vol. 11, no. 1, p. 32, Jan. 2020.
- [77] C. Pan and J. Tan, "Day-ahead hourly forecasting of solar generation based on cluster analysis and ensemble model," *IEEE Access*, vol. 7, pp. 112921–112930, 2019.
- [78] Y. Zhou, N. Zhou, L. Gong, and M. Jiang, "Prediction of photovoltaic power output based on similar day analysis, genetic algorithm and extreme learning machine," *Energy*, vol. 204, Aug. 2020, Art. no. 117894.

- [79] I. Koprinska, M. Rana, and A. Rahman, "Dynamic ensemble using previous and predicted future performance for multi-step-ahead solar power forecasting," in *Proc. Int. Conf. Artif. Neural Netw.* Cham, Switzerland: Springer, 2019, pp. 436–449.
- [80] X. Zhang, Y. Li, S. Lu, H. F. Hamann, B.-M. Hodge, and B. Lehman, "A solar time based analog ensemble method for regional solar power forecasting," *IEEE Trans. Sustain. Energy*, vol. 10, no. 1, pp. 268–279, Jan. 2019.
- [81] IEA-PVPS. (2020). *Regional Solar Power Forecasting*. [Online]. Available: <https://iea-pvps.org/key-topics/regional-solar-power-forecasting-2020/>
- [82] E. Lorenz, T. Scheidsteiger, J. Hurka, D. Heinemann, and C. Kurz, "Regional PV power prediction for improved grid integration," *Prog. Photovolt., Res. Appl.*, vol. 19, no. 7, pp. 757–771, 2011.
- [83] D. Koster, F. Minette, C. Braun, and O. O'Nagy, "Short-term and regionalized photovoltaic power forecasting, enhanced by reference systems, on the example of Luxembourg," *Renew. Energy*, vol. 132, pp. 455–470, Mar. 2019.
- [84] P. Aillaud, J. Lequeux, J. Mathe, L. Huet, C. Lallemand, O. Liandrat, N. Sebastien, F. Kurzrock, and N. Schmutz, "Day-ahead forecasting of regional photovoltaic production using deep learning," in *Proc. 47th IEEE Photovolt. Spec. Conf. (PVSC)*, Jun. 2020, pp. 2688–2691.
- [85] Y. M. Saint-Drenan, G. H. Good, M. Braun, and T. Freisinger, "Analysis of the uncertainty in the estimates of regional PV power generation evaluated with the upscaling method," *Sol. Energy*, vol. 135, pp. 536–550, Oct. 2016.
- [86] S. Killinger, P. Guthke, A. Semmig, B. Müller, B. Wille-Hausmann, and W. Fichtner, "Upscaling PV power considering module orientations," *IEEE J. Photovolt.*, vol. 7, no. 3, pp. 941–944, May 2017.
- [87] L. A. Fernandez-Jimenez, S. Terres-Olarte, M. Mendoza-Villena, E. Garcia-Garrido, E. Zorzano-Alba, P. M. Lara-Santillan, P. J. Zorzano-Santamaria, and A. Falces, "Day-ahead probabilistic photovoltaic power forecasting models based on quantile regression neural networks," in *Proc. Eur. Conf. Electr. Eng. Comput. Sci. (EECS)*, Nov. 2017, pp. 289–294.
- [88] A. Nespoli, E. Ogliairi, S. Leva, A. M. Pavan, A. Melliti, V. Lughi, and A. Dolara, "Day-ahead photovoltaic forecasting: A comparison of the most effective techniques," *Energies*, vol. 12, no. 9, p. 1621, Apr. 2019.
- [89] M. K. Behera, I. Majumder, and N. Nayak, "Solar photovoltaic power forecasting using optimized modified extreme learning machine technique," *Int. J. Eng. Sci. Technol.*, vol. 21, no. 3, pp. 428–438, Jun. 2018.
- [90] S. Mishra, L. Tripathy, P. Satapathy, P. K. Dash, and N. Sahani, "An efficient machine learning approach for accurate short term solar power prediction," in *Proc. Int. Conf. Comput. Intell. Smart Power Syst. Sustain. Energy (CISPSSSE)*, Jul. 2020, pp. 1–6.
- [91] M. Chai, F. Xia, S. Hao, D. Peng, C. Cui, and W. Liu, "PV power prediction based on LSTM with adaptive hyperparameter adjustment," *IEEE Access*, vol. 7, pp. 115473–115486, 2019.
- [92] F. Wang, Z. Xuan, Z. Zhen, K. Li, T. Wang, and M. Shi, "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.*, vol. 212, May 2020, Art. no. 112766.
- [93] H. Zang, L. Cheng, T. Ding, K. W. Cheung, Z. Wei, and G. Sun, "Day-ahead photovoltaic power forecasting approach based on deep convolutional neural networks and meta learning," *Int. J. Electr. Power Energy Syst.*, vol. 118, Jun. 2020, Art. no. 105790.
- [94] H. Eom, Y. Son, and S. Choi, "Feature-selective ensemble learning-based long-term regional PV generation forecasting," *IEEE Access*, vol. 8, pp. 54620–54630, 2020.
- [95] T. Zhang, C. Lv, F. Ma, K. Zhao, H. Wang, and G. M. P. O'Hare, "A photovoltaic power forecasting model based on dendritic neuron networks with the aid of wavelet transform," *Neurocomputing*, vol. 397, pp. 438–446, Jul. 2020.
- [96] V. Kushwaha and N. M. Pindoriya, "A SARIMA-RVFL hybrid model assisted by wavelet decomposition for very short-term solar PV power generation forecast," *Renew. Energy*, vol. 140, pp. 124–139, Sep. 2019.



BERHANE DARSENE DIMD (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering from Hawassa University, Ethiopia, in 2011, and the M.Sc. degree in electric power engineering from the Norwegian University of Science and Technology (NTNU), in 2015. He is currently pursuing the Ph.D. degree in digital solar electricity with the Department of Electric Power Engineering, NTNU. He is an Electric Power Engineer with a special interest in renewable energy technologies. His main research interests include PV systems, artificial intelligence and machine learning, energy optimization, and smart grids.



STEVE VÖLLER received the Ph.D. degree in energy technology from the University of Wuppertal, Germany, in 2010. From 2010 to 2019, he worked as a Postdoctoral Researcher and a Researcher at the Norwegian University of Science and Technology (NTNU) and SINTEF Energy, respectively, mainly on energy systems, offshore wind, energy transmission, CO₂ emissions, micro-grids, electric/hydrogen vehicles, integration of renewables, and hydropower scheduling. Since 2019, he has been an Associate Professor of sustainable electrical energy with the Department of Electric Power Engineering, NTNU. He is an Electrical Engineer with a specialization in energy systems. His current research interests include energy system analysis, solar energy, wind power, energy storage, and hydrogen systems.



UMIT CALI (Member, IEEE) received the Ph.D. degree in electrical engineering and computer science from the University of Kassel, Germany. He is currently an Associate Professor with the Department of Electric Power Engineering, Norwegian University of Science and Technology. He is the first author of *Digitalization of Power Markets and Systems Using Energy Informatics*. His research interests include energy informatics, artificial intelligence, blockchain technology, renewable energy systems, and energy economics. He is the Chair of IEEE TEMS Special Interest Group on Energy DLT and the Vice Chair of IEEE Blockchain in Energy Standards WG (P2418.5).



OLE-MORTEN MIDTGÅRD (Member, IEEE) received the M.Sc. and Ph.D. degrees from the Department of Electric Power Engineering, Norwegian University of Science and Technology (NTNU), in 1992 and 1997, respectively. He has been a Professor with the Department of Electric Power Engineering, NTNU, since 2012, where he was the Head of the Department, from 2017 to 2021. He is the Vice Dean of Sustainability and Innovation with the Faculty of Information Technology and Electrical Engineering, NTNU. His current research interests include electric aviation, smart grids, PV systems, artificial intelligence, power electronics, and distributed energy management systems.

...