Renewable and Sustainable Energy Reviews 124 (2020) 109792

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: http://www.elsevier.com/locate/rser

A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization

R. Ahmed [a,c,*], V. Sreeram [a], Y. Mishra [a,b], M.D. Arif [c]

[a] The University of Western Australia, School of Electrical, Electronic, and Computer Engineering, Perth, Australia
[b] Queensland University of Technology, School of Electrical Engineering and Computer Science, Brisbane, Australia
[c] International University of Business Agriculture and Technology, College of Engineering and Technology, Dhaka, Bangladesh

ARTICLE INFO

Keywords:
Solar power
Forecasting technique
Wavelet transform
Deep convolutional neural network
Long short term memory
Optimization
Forecast accuracy

ABSTRACT

Integration of photovoltaics into power grids is difficult as solar energy is highly dependent on climate and geography; often fluctuating erratically. This causes penetrations and voltage surges, system instability, inefficient utilities planning and financial loss. Forecast models can help; however, time stamp, forecast horizon, input correlation analysis, data pre and post-processing, weather classification, network optimization, uncertainty quantification and performance evaluations need consideration. Thus, contemporary forecasting techniques are reviewed and evaluated. Input correlational analyses reveal that solar irradiance is most correlated with Photovoltaic output, and so, weather classification and cloud motion study are crucial. Moreover, the best data cleansing processes: normalization and wavelet transforms, and augmentation using generative adversarial network are recommended for network training and forecasting. Furthermore, optimization of inputs and network parameters, using genetic algorithm and particle swarm optimization, is emphasized. Next, established performance evaluation metrics MAE, RMSE and MAPE are discussed, with suggestions for including economic utility metrics. Subsequently, modelling approaches are critiqued, objectively compared and categorized into physical, statistical, artificial intelligence, ensemble and hybrid approaches. It is determined that ensembles of artificial neural networks are best for forecasting short term photovoltaic power forecast and online sequential extreme learning machine superb for adaptive networks; while Bootstrap technique optimum for estimating uncertainty. Additionally, convolutional neural network is found to excel in eliciting a model's deep underlying non-linear input-output relationships. The conclusions drawn impart fresh insights in photovoltaic power forecast initiatives, especially in the use of hybrid artificial neural networks and evolutionary algorithms.

1. Introduction

Electricity is vital for economic development and technological growth [1]. It is a key factor in rapid urbanisation, and industrialisation, to the extent that economic growth is frequently measured in per capita power output of a country [2]. This ever-growing energy demand leads to an increased need for electricity generation and distribution. However, globally, the reliance is on non-renewable pollution-causing fossil fuels for electricity production. Approximately two-thirds of the global carbon dioxide emissions are from such fuel sources whose current share of energy production, if maintained [3], will inevitably lead to a significant rise in average global temperature and other catastrophes. Already, the rise in average global temperature has been correlated with extreme weather patterns [4] namely: increase in violent storms, floods, heavy snowfalls and droughts. The World Meteorological Organization's (WMO) provisional statement on the State of Global Climate mentioned that the year 2019 witnessed one decade of unprecedented elevated global temperature, retreating glaciers and record high sea levels due to greenhouse gas (GHG) emissions. The average global temperatures for the past five (2015–2019) and ten (2010–2019) years were the highest in recorded history.

Surprisingly, the dire necessity for reducing GHG emissions has failed to abate the reliance on fossil fuels due to their perceived economics. The bright side is that the harnessing of renewable energies (RE) has recently brought sweeping changes in the arena of energy generation and demonstrates the potential of clean and limitless energy for the future. Moreover, policies enacted by international organizations and major players in world economy regarding carbon taxation have paved the way for RE. Resolutions such as The United Nations Climate Change

* Corresponding author. The University of Western Australia, School of Electrical, Electronic, and Computer Engineering, Perth, Australia.
E-mail address: razin.ahmed@research.uwa.edu.au (R. Ahmed).

https://doi.org/10.1016/j.rser.2020.109792
Received 27 June 2019; Received in revised form 19 February 2020; Accepted 20 February 2020
Available online 2 March 2020
1364-0321/

Abbreviations

AI	Aerosol Index
APVF	Analytical PVPF
ACO	Ant colony optimization
ASU	Applied Science Private University
ANN	Artificial neural network
AE	Autoencoder
AR	Auto-regressive
ARIMA	Autoregressive integrated moving average
ARMA	Autoregressive moving average
BP	Back Propagation
BPNN	Back propagation neural network
BR	Bayesian Regularization
BS	Bootstrap technique
CAS	Chaotic ant swarm optimization
CFS	Climate Forecast System
CMDV	Cloud motion displacement vector
CMV	Cloud motion vectors
CNN1D	CNN with 1-D convolution layer
CNN2D	CNN with 2-D convolution layer
CEEMD	Complementary ensemble empirical mode decomposition
CSP	Concentrated solar power
CRBM	Conditional restricted Boltzmann machine
CRPS	Continuous Ranked Probability Score
CNN	Convolutional Neural Network
CRrBMs	Convolutional restricted Boltzmann machines
DBN	Deep belief network
DCNN	Deep convolutional neural network
DL	Deep learning
DNN	Deep neural network
DRNN	Deep recurrent neural networks
DRBM	Discriminative restricted Boltzmann machine
ESN	Echo State Networks
ENN	Elman neural network
EMD	Empirical mode decomposition
ECMWF	European Centre for Medium-Range Weather Forecasts
EWMA	Exponentially weighted moving average
FBNN	Feedback neural network
FFNN	Feed-forward neural network
FF	Firefly
FPCT	Fourier phase correlation theory
FT	Fourier Transform
FOA	Fruit fly optimization algorithm
GPR	Gaussian process regression
GRNN	Generalized Regression Neural Network
GWC	Generalized weather classes
GAN	Generative adversarial network
GA	Genetic algorithm
GMS	Geostationary meteorological satellite
GOES	Geostationary operational environmental satellite
GW	Giga Watt
GDAS	Global Data Assimilation System
GFS	Global Forecast System
GHI	Global horizontal irradiance
GLCM	Gray-level co-occurrence matrix
GHG	Greenhouse gas
HEPSO	High Exploration Particle Swarm Optimization
HRRR	High-Resolution Rapid Refresh
IPSI	Image-phase-shift-invariance
IA	Immune algorithm
KDE	Kernel Density Estimation
KNN	k-nearest-neighbours
LVQ	Learning vector quantization
LS SVR	Least-squares support vector regression
LM	Levenberg-Marquardt
LLSR	Linear least square regression
LSTM	Long Short Term Memory
LUBE	Lower Upper Bound Estimation
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MVE	Mean Variance Estimation
MW	Mega Watt
MIF	Meteorological impact factors
MSG	Meteosat second generation
MA	Moving average
MLP	Multilayer perceptron
MLPNN	Multi-layer perceptron neural network
MME	Multi-Model Ensemble
MPVF	Multiplayer perceptron PVPF
MLR	Multiple Linear Regression
NNE	Neural Network Ensemble
NARX	Nonlinear autoregressive exogenous inputs
nMAPE	Normalized Mean Absolute Percentage Error
nRMSE	Normalized Root Mean Square Error
NAM	North American Mesoscale Model
NWP	Numerical Weather Prediction
OS-ELM	Online sequential extreme learning machine
OF	Optimal flow
PIV	Particle image velocimetry
PSO	Particle swarm optimization
PSO-GA	Particle Swarm Optimization- Genetic Algorithm
PV	Photovoltaic
PVPF	Photovoltaic power forecasting
PI	Prediction interval
PICP	Prediction interval coverage probability
PIW	Prediction interval Width
PDF	Probability Density Function
QR	Quantile regression
RBFNN	Radial basis function neural network
RFR	Random forest regression
RAP	Rapid Refresh
RNN	Recurrent neural network
RCN	Recursive convolutional networks
RE	Renewable energies
RBM	Restricted Boltzmann machine
RoBM	Robust Boltzmann machine
RMSE	Root Mean Square Error
SD	Seasonal decomposition
SOM	Self-organized map
ST-PVPF	Short-term PVPF
SLFN	Single layer feed forward networks
SSHS	Solar Space Heating System
SF	Spectral factor
SVM	Support vector machine
SVR	Support vector regression
TGA	Team Game Algorithm
WMO	The World Meteorological Organization's
COP21	United Nations Climate Change Conference
WD	Wavelet decomposition
WNN	Wavelet Neural Network
WT	Wavelet transform
WSPR	Weather statuses pattern recognition
WTHD	Weather type of historical data

Symbols

$P(t+k\|t)$	predicted power
n	number of historic measurements
Δt	step time difference
ΔA	Forecasted change response at spatial location

Y_t	current observation
$\widehat{Y}_t$	predicted value
$\alpha(Y_t - \widehat{Y}_t)$	adjustment factor
p and q	the number of processes
α_i and β_j	coefficients of AR and MA
B	backward shift operator
σ_a	variance
$S_k(t)$	activation function
V_{kq}	connection weight
k	processing node
$g(w_i \cdot x_i + b_i)$	activation function
H	output matrix

Conference (COP21) in 2015 and joint presidential statements by the US and Chinese presidents on climate change signal new policies which hold promise for the implementation of REs [5]. The goals of the European Union are even more ambitious: to curtail GHG emissions by 80% (from a 1990 baseline) and to produce 100% of the required energy from RE by 2050 [6]. Also, contemporary technological advances in extracting energy from wind, and especially solar, have become competitive and viable alternatives to fossil fuels [7,8].

Solar energy is the radiant energy from the sun, which produces massive amounts of electromagnetic energy by thermonuclear fusion of hydrogen gas. The average solar radiation intensity on the earth's surface is 1367 W/m^2 and the total global absorption of solar energy is approximately 1.8×10^{11} MW [9]. This amount of ubiquitous and limitless energy is more than enough to meet all power requirements on a global scale [3]. Fig. 1 illustrates the intensity of solar energy across the globe. Countries located in the geographical zones above 45°N or below 45°S latitudes have tremendous opportunity for harnessing solar energy. Regions in the Middle East, the Mojave Desert (USA), the Chilean Atacama Desert, the Sahara Desert, the Kalahari Desert (Africa) and the North-Western region of Australia are suitable for large scale PV installations.

The current solar power plants are of two types: solar thermal systems and solar photovoltaic (PV). Solar thermal technology concentrates sunlight, thereby increasing the temperature of heat sinks which produce steam. The steam is used in steam turbines for large-scale electricity generation. Thus, countries such as Spain and USA have been implementing concentrated solar power (CSP) for electricity [11]. Nevertheless, CSP requires large installations to be effective. In contrast, photovoltaics use the incident photons in sunlight to excite free electrons in embedded semiconductors, causing a charge build up and yielding electricity. These panels are functional at different scales and are often used on rooftops or open spaces, integrated into buildings or vehicle designs, or arranged in huge arrays in solar power plants. Over the past decades, PVs have been widely installed and their demand has increased as a direct consequence of their mass public appeal and reduction in tariff charge system [12,13]. Fig. 2 compares the global demand for PV and CSP technology.

The graph illustrates the growth rates of PV and CSP technology for the duration 2000 to 2018. It is observed that global PV output tripled from 2005 to 2008, and is growing exponentially ever since (reaching 99 TWh in 2012). In particular, the year 2015 saw total PV output reach 250 TWh with the installation of 185 million units [5], while the contemporary total global CSP capacity was 4.7 TWh. Subsequently, the years 2016–2018 witnessed meteoric increase in PV power, with additions of 97 GW (GW) in 2018; accounting for around half of the total renewable capacity growth. More specifically, solar PV capacity doubled from 2016 to 2017 and surpassed 30% in growth alone in 2018 (571 TWh addition) [14]. This growth in PV power has resulted in its unprecedented contribution to global electricity generation by more than 2% and extrapolation demonstrates its future planet-wide capacity of 1700 GW by 2030 [15]. In contrast, the year 2018 saw the addition of 600 MW of CSP capacity – the greatest annual expansion since 2013 and five times more than in 2017. From 2011 to 2017, CSP capacity increased an average of 25% annually. Finally, in 2018, global CSP output rose by only 8% (estimated) despite record-level additions. Therefore, the statistics succinctly demonstrate the waxing demand for solar energy, especially the preference for PV systems; furthering the requirement for accurate and reliable forecasting of PV power [14].

The remarkable two decades long growth in PV power is mainly due to reduction in the price of PV systems, increase in their efficiency, modularity in installation, low cost of maintenance and operation, long service life, lowering of CO_2 emissions and environmental friendliness [16]. PVs are now competitive with fossil-fired technology in many countries. It is effective in standalone or grid-connected modes, but its efficiency varies considerably due to the fluctuations of incident solar energy on account of environmental conditions and geographical location. The weather of a particular geographical location varies at timescales ranging from minutes to hours to multiple days, as well as across years and decades. PV output also depends on the sun's movement, escalating in the morning, reaching maximum generation during mid-day and falling off at dusk. Thus, PV output is dynamic in nature and reliant on location and climate.

The local solar radiation intensity varies with latitude, season, atmospheric conditions (e.g. rain, snowfall, fog, humidity etc.), air quality

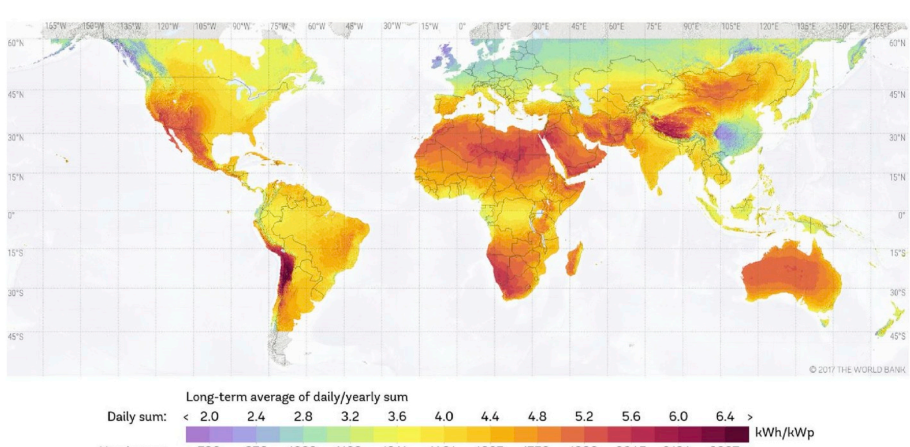
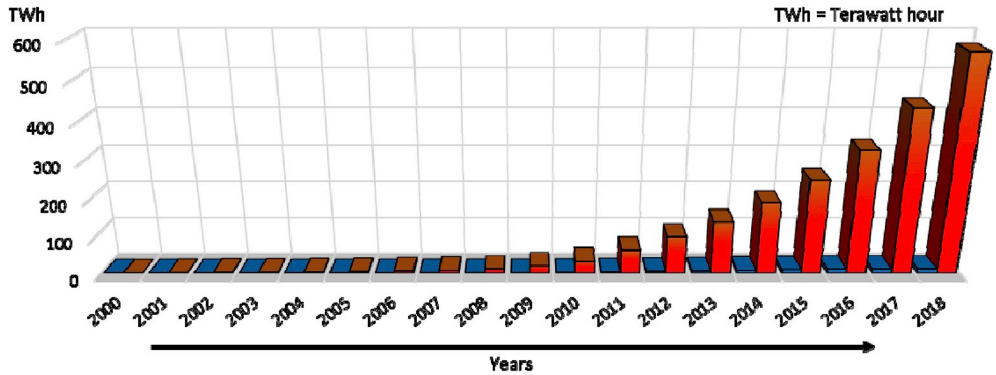


Fig. 1. The world solar resources map [10].



Years	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
CSP(TWh)	0.5	0.6	0.6	0.5	0.6	0.6	0.6	0.7	0.9	0.9	1.6	2.9	4.7	5.9	8.4	9.6	10.5	11	11.9
PV(TWh)	0	1	2	2	3	4	6	8	12	20	32	63	99	139	190	250	328	435	571

Fig. 2. Global installed solar power capacity, 2000–2018 [14].

and pollution index (smog) etc. Therefore, combining solar energy harvesting technologies with existing power systems (hybrid renewable energy systems) poses a significant, and as yet unmitigated, technical and stability challenge. The issues stem from continuous variation in solar irradiance, temperature, PV output, costly energy storage equipment, grid stability and seasonal effects. In addition, integrating PV power network as a back-up electric supply for meeting increased demand without the use of energy reserve devices is not technically viable; such a setup affects grid stability. In short, variations in meteorological states lead to uncertainty in PV output [16] precipitating intermittent penetrations, voltage surges, reverse power flows, variations in frequency harmonic distortion in current and voltage waveforms etc. Such unpredictable output considerably influences the reliability, stability and scheduling of the power system operation and economic dispatch [17]. Reliable PV output forecast will considerably decrease this uncertainty, enhance stability and improve economic viability. Therefore, at present, accurate PV power forecasting (PVPF) is a crucial research arena [18,19].

2. Major factors affecting solar power forecast

Different factors affect PVPF accuracy making such prediction a sophisticated process. It depends on factors such as forecasting horizons, forecast model inputs and performance estimation. To achieve better precision, correlational analysis optimization and uncertainty estimation of PVPF models need to be carried out.

2.1. Forecasting horizons

The future time period for output forecasting or the time duration between actual and effective time of prediction is the forecast horizon [13]. Some researchers prefer three categories of the forecast horizon: short-term, medium-term and long-term, as in Refs. [20,21]. Others have added a “fourth” category [22] based on the requirements of the decision-making process for smart or micro grids [16], aptly termed “very short-term or ultra-short term forecast horizon”. However, as yet there is no universally agreed upon classification criterion [16,23,24].

i Very short-term or ultra-short-term forecasting

Very short-term forecasting is used in power system and smart grid planning with prediction period from seconds to less than 30 min [25]. Others have considered 1 min to 1 or several hours and even up to day ahead forecasting within this category. Such forecasts are highly beneficial to electricity marketing or pricing, power smoothing processes, monitoring of real-time electricity dispatch and PV storage control [13].

ii Short-term forecasting

This is popular in the electricity market, where decisions comprise of economic load dispatch and power system operation. It is also useful in control of renewable energy integrated power management systems. Generally, the temporal horizon is between 30 and 360 min [25]. However, some consider one to several hours, one day, or up to seven days as short-term forecast horizon [13]. For instance Ref. [18], propounded that electric load patterns should be forecasted 2 days ahead for effective scheduling of power plants and for planning transactions.

iii Medium-term forecasting

Medium-term forecasting spans 6–24 h [25]; although, some have considered one day, one week and up to a month ahead as being in this category. It is essential for maintenance scheduling of conventional or solar energy integrated power systems consisting of high-end transformers and different types of electro-mechanical machinery [13].

iv Long term forecasting

Long-term forecasts predict scenarios more than 24 h in advance [25]. Nonetheless, some have categorized periods of a month to a year as long-term forecast [13]. Such prediction horizon is suitable for long term power generation, transmission, distribution and solar energy rationing [26], as well as, for taking into account seasonal trends. However, these models have reduced accuracy because weather fluctuations spanning one or few days cannot be predicted using such long horizons.

The different time horizons add complexity; thus, an alternative is using short-term forecasting methods to extended horizons, i.e. long and medium-term predictions. Although, this may severely degrade the prediction accuracy [27]. Yet researchers [28–30] have developed medium-term and year ahead models for monthly power system scheduling, electricity pricing and load forecasting.

The problems with forecasting as discussed above have prompted many researchers to develop a forecast horizon classification approach specifically for PVPF: intra-hour, intra-day and day ahead. These categories often overlap with the short, medium and long-term categories described previously.

i Intra-hour

Also known as nowcasting, it involves forecast horizons from a few seconds to an hour [31]; thus overlapping with very short-term and short-term horizon categories. Intra-hour forecasts help ensuring grid quality and stability, as well as accurate scheduling of spinning reserves and demand responses. Island grids and low quality power supplies, where high solar penetrations exist, rely on such predictions. It is also used in operation planning at distribution systems to reduce tap operations in transformers and in formulating sub-hourly bids for electricity markets. Nonetheless, most distributed PV plants are generally unaffected by very short-term variability in power production due to their distributed and aggregate nature [32].

ii Intra-day

This forecast horizon spans 1–6 h and overlaps with short and medium-term categories. It finds use in controlling zone specific electric loads and in electricity trading outside the standard grid [33].

While few studies [34–36] utilized numerical weather prediction (NWP) for intra-hour predictions, it is a common tool in intra-day forecasts as it incorporates future atmospheric conditions in PVPF; increasing precision. This is because, NWP's are effective for forecast horizons exceeding 4 h since they lack the required granularity at shorter time scales.

Almonacid, et al. [37] developed an intra-day PVPF exploiting ANN. Their model accurately predicted 1 h ahead PV output with global horizontal irradiance (GHI) and cell temperature as inputs. The two inputs were derived from two non-linear autoregressive models which predicted GHI and atmospheric temperature from current and historical measurements. Subsequently, the cell temperature was obtained and the data used in an ANN model to obtain PV output predictions with normalized root mean square error (nRMSE) of 3.38%.

Another intra-day ANN based PVPF used wavelet decomposition of power output combined with GHI and temperature predictions optimized with particle swarm optimization (PSO) [38]. It accurately forecasted 1, 3 and 6 h ahead with uncertainty estimation via bootstrap method.

Zhang, et al. [39] combined intra-day and day-ahead (discussed below) forecast horizons. They analysed four scenarios based on geographical aggregation and location, obtaining good predictions with nRMSE ranging from 2 to 17% for a single and a large ensemble of PV systems (64 and 495 GW), respectively. They demonstrated that spatial averaging minimized errors.

iii Day-ahead

Forecasts spanning 6–48 h are included in this category and overlap with medium and long-term horizons. Such models are used in utilities planning and unit commitment. Chen, et al. [40] proposed a self-organized map (SOM) PVPF model with day-ahead forecasting. Their model classified weather types (sunny, cloudy and rainy) in accordance with meteorological variables (solar irradiance, total and low cloud amounts), and harnessed ANN-RBFN for forecasting with MAPEs of 9.45%, 9.88% and 38.12% for sunny, cloudy and rainy days, respectively.

Frequently, day-ahead forecasts use NWP outputs. Literature search demonstrates that only three intra-hour modelling studies, about 57% of intra-day forecasts and approximately 79% of day-ahead prediction approaches employed NWP variables. The reason: NWP improves accuracy when time horizons increase, feeding future meteorological trends to the models.

Lu, et al. [41] evaluated day-ahead PVPF models which used machine learning and NWP combined with: North American Mesoscale Model (NAM), Rapid Refresh (RAP) and High-Resolution Rapid Refresh (HRRR) methods. Their results verified that the blended model was superior to single models with 30% lower Mean Absolute Error (MAE), attributing it to the cancelation of systematic bias errors.

[39] concluded that the ensemble approach was best for PVPF with hourly forecasts being more precise than day-ahead ones. Another important finding was that the difference in PVPF accuracy for different horizons (hour to day-ahead) increased with the area of distributed systems.

Longer forecast horizons also exist, but PVPF involving 2 days or longer are rarer in extant literature. They are useful for unit commitment, transmission management, trading, hedging, planning and resource optimization [42]. Furthermore, when electricity production is lower than expected, 48 h ahead predictions are popular for economic dispatch and planning plant maintenance.

Even longer horizons were discussed by researchers. Lin and Pai [43] forecasted monthly power output for an ensemble of PV plants in Taiwan. Seasonal decomposition (SD) was used to detrend data and account for seasonal effects. Subsequently, they employed least-squares support vector regression (LS SVR) to select the best input set, similar to Ref. [44] who applied GA for PVPF. The results demonstrated that the SD LS SVR was superior compared to the benchmarks, which consisted of autoregressive integrated moving average (ARIMA), generalized regression NN and LS SVR.

Vaz, et al. [45] also investigated long-term horizons forecasts with nonlinear autoregressive with exogenous inputs (NARX). Their results outperformed the benchmarks which were persistence models for 4, 7 and 15 days ahead forecasting (nRMSE was 20%). However, for longer horizons, i.e. 20 days to a month, the root mean square error (RMSE) of the proposed model increased to 24%; equalling persistence models' performance.

2.1.1. Dependence of PVPF accuracy on time horizon

As discussed, many have studied forecast horizons, and it is observed that keeping forecast model and other parameters constant, forecast accuracy varies with the change in the time span [13]. Lonij, et al. [34] developed a 15–90 min time horizon PVPF and concluded that the model's accuracy varied with horizon's length, keeping all other factors constant. Likewise [46], found that the forecast error increased from 3.2% to 15.5% for forecast horizons ranging from 20 to 180 s for the same dataset. As a consequence, researchers are developing a transcend system which can manipulate short-term temporal forecasting of PV output.

Others have evaluated disparate multi-scale hybrid forecast models for solar irradiance with different forecast horizons. Horizons ranging from 5 to 30 min at 5 min increment and from 1 to 6 h in hourly increments have been studied. It was found that regardless of the type of model, the nRMSE values increased with forecast horizon's duration. Furthermore, it was inferred that shorter time scales of learning data compared to that of test data did not increase prediction accuracy. This meant that increasing statistical information by using shorter time scales did not enhance accuracy; rather it could complicate the learning phase for longer time horizons [47].

The interplay of many inter-dependent factors also affects PVPF accuracy. To exemplify, cloud motion severely affects sunlight intensity; and thus PVPF, especially for minute-scale or ultra-short-term horizons (i.e. intra-hour). Hence, classification and prediction of cloud motion is crucial. Previous works [48–51] did not consider cloud motion, even those which involved 15 min time stamps, thus reducing accuracy in cloudy conditions and failing real-time grid dispatch requirements. Therefore, ultra-short-term PVPF are effective only in combination with weather classification techniques [40,48,52,53]. Such techniques are robust in mapping input-output for prediction in all-weather types.

Another factor in cloud status based PVPF modelling is the use of ground-based sky images which have higher resolution than satellite imagery. Moreover, image processing and statistical techniques, e.g. linear extrapolation, can help predict cloud distribution using cloud motion displacement vectors (CMDVs). Wang, et al. [54] applied image phase shift invariance (IPSI) based CMDV calculation married to Fourier phase correlation theory (FPCT) to develop accurate minute-scale PVPF for different cloud scenarios.

Finally, cloud conditions were also considered in intra-day PVPF models where satellite imagery was exploited. Kühnert, et al. [55] identified sources of such type of images: Meteosat second generation

(MSG), geostationary satellites of the European organization (EUMETSAT); geostationary operational environmental satellite (GOES), for the Americas and geostationary meteorological satellite (GMS), for Japan. From the aggregate images, cloud-index can be estimated and cloud motion vectors (CMV) calculated. Subsequently, cloud movement patterns can be determined to generate predictions of GHI over hourly time scales. In the end, output maps are smoothed to curtail errors and predictions of up to 5 h are possible.

From the above discussion it can be concluded that as the length of the forecast horizon increases the accuracy of any PVPF modelling approach decreases and is significantly affected past a time span of 24–48 h. This is because cloud cover and distribution, which are strongly correlated with solar irradiance, cannot be predicted with considerable precision for extended time periods due to its inherent stochastic nature. Also, intra-hour and intra-day forecast horizons are relevant to PV output predictions while longer time horizon forecasts i.e. day-ahead are more suitable for power system planning. Thus, the current review focused on the efforts of contemporary researchers who have been more interested in the accuracy of PVPF models, which are much higher for shorter time horizons, up to a day-ahead. This is also the reason why there are more works investigating short to medium-term forecast horizons, frequently with weather classification, compared to longer time spans. The fewer studies related to long-term forecast horizons have mostly determined such forecast horizon models to be less accurate as cloud distribution and cover are dynamic and can only be predicted for short durations ahead of time with reasonable degree of accuracy. However, it must be mentioned that, long-term models have found use in predicting seasonal variations and their impacts on photovoltaics; i.e. for PV plant planning.

2.2. Weather classification

PV output is most strongly correlated with solar spectral irradiance; the latter depending on meteorological impact factors (MIF): aerosol distribution, wind speed and direction, humidity and cloud cover. Thus, changing weather status affects PVPF accuracy, and an effective PVPF model must integrate forecasting with weather classification for improving robustness. Indeed, state-of-the-art research demonstrates that weather classification is an indispensable pre-processing step, especially for short-term PVPF (ST-PVPF) [40,48,52].

Yet, weather classification based PVPF models face difficulties concerning insufficiency of training datasets. In Ref. [53] the authors reclassified the standard 33 meteorological weather categories into 10 weather classes by compiling several single weather types to constitute a single new weather type. In contrast, most studies have categorized weather data into less than four general types [40,48,53]. Wang, et al. [53] opined that separate PVPF model for each weather class increases forecast precision and mapping [53,56,57]. To accomplish this task, they employed generative adversarial network (GAN) to augment the training datasets used for each weather types; especially for extreme weathers which are under-represented; to train the individual convolutional neural network (CNN) (base models). Thus, both original and generated solar irradiance data were utilized. The authors claimed that their GAN–CNN–based day-ahead PVPF model was more accurate than established approaches.

Weather patterns are strongly correlated with PV output [53,58]; necessitating the inclusion of weather statuses in PVPF models [9,53,59–61]. Many research studies [40,48,52,53] have concluded that weather classification is an effective pre-processing step for improving prediction accuracy of ST-PVPF, especially for day-ahead forecasting [52,53]. employed SOM and learning vector quantization (LVQ) to classify historical data of PV power output [40,53]. also utilized SOM to categorize local weather types for day-ahead PVPF [53,56]. described a solar irradiance feature extraction and support vector machines (SVM) based weather pattern recognition approach for ST-PVPF [53,57]. studied data distribution in disparate classes of daily weather classification results for better PVPF. However, these researches have not fully investigated the establishment process of a precise and consistent weather classification model. Most considered it an extraneous part of PVPF. In contrast [53], argued that weather classification should be used as a benchmark for selecting the most suitable forecasting strategy for a given region and climate. Furthermore, they showed that single forecasting models (without weather classification) were significantly affected by the historical data of the previous three days and lacked the ability to predict the day-ahead weather type. In short they mentioned, finer the weather classification; better the prediction results.

The literature abounds with comparisons of weather classification and PVPF models; often posing a daunting task for the researcher of PVPF to make the proper selection. In Ref. [53], the authors compared their GAN based weather classification approach, involving 10 weather classes, with five other established prediction models. For objectivity, a confusion matrix was employed [62]. This statistical tool uses a table layout to visualize classifier algorithms' performance for comparison and verifies how often a classifier confuses two adjacent weather categories and mislabels them. Three indices: PA (product's accuracy), UA (user's accuracy) and OA (overall accuracy) calculated the performance. The five common classifiers compared were: CNN with 1-D convolution layer (CNN1D), CNN with 2-D convolution layer (CNN2D), multilayer perceptron (MLP), and SVM and k-nearest neighbor (KNN). For all, weather classification and subsequent data augmentation via GAN for short sample size and data imbalance weather types the accuracy was improved. However, CNN2D, with its deep learning (DL) abilities for eliciting non-linear input-output relationships, performed the best, and so, the authors recommended GAN-CNN2D for PVPF modeling.

Another weather classification and PVPF was detailed by Ref. [56], where simulation and case study based research work was utilized. The authors addressed extreme, short duration weather statuses, which are often missing from weather type of historical data (WTHD). This causes difficulties in model training for data driven modelling approaches. To mitigate, a solar irradiance feather extraction and SVM based weather statuses pattern recognition (WSPR) technique was designed to identity the missing WTHD from time series data. Subsequently, they used this technique to develop a ST-PVPF model using four generalized weather classes (GWC) for covering all weather types; thus, making the classification and prediction processes simpler and computation friendly.

2.3. Forecast model performance

Performance estimation is essential for evaluating a model's forecasting accuracy. Common tools include: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) [63,64] (see Table 1). MAE estimates the average significance of

Table 1
Abbreviations commonly used for forecast performance evaluations.

Abbreviation	Metric	Description
MAE	Mean Absolute Error	Evaluates uniform forecast errors of models
RMSE	Root Mean Square Error	Measures overall accuracy of the forecasting models. However, square order increases the prediction error rate.
nRMSE	Normalized Root Mean Square Error	Measures overall accuracy in a large dataset. However, square order increases the prediction error rate.
MAPE	Mean Absolute Percentage Error	Evaluates uniform forecast errors in percentage for the forecasting models.
nMAPE	Normalized Mean Absolute Percentage Error	Evaluates uniform forecast errors in percentage for large datasets
nMAP	Same as nMAPE	
rMAPE	Same as nMAPE	
NRMSE	Same as nRMSE	

the errors in a dataset of forecasts by averaging the differences between actual observations and predicted results of the entire test sample, giving all individual discrepancies equal weight. Similarly, RMSE estimates the mean value of the error utilizing the square root of the average of squared differences between forecasted values and actual observations. Therefore, it is more robust in dealing with large deviations that are especially undesirable, giving the researcher the ability to identify and eliminate outliers. Also, normalized RMSE (nRMSE) is used in large datasets to evaluate overall deviations. Both (MAE and RMSE) average metrics, however, can range from zero to infinity [65,66]. In contrast, MAPE is a standard prediction technique that measures the accuracy of forecasting and justifies the prediction diversity for real datasets. Likewise, for large datasets, normalized nMAPE is used for forecast evaluation. The equations of these metrics are as follows [26]:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|y_j - t_j\right| \tag{1}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_j - t_j\right)^2} \tag{2}$$

$$\mathrm{nRMSE} = \frac{\sqrt{\sum_{i=1}^{N}\left(y_j - t_i\right)^2}}{\sqrt{\sum_{i=1}^{N} {y_i}^2}} \tag{3}$$

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|y_j - t_j\right|}{y_j} \times 100\% \tag{4}$$

$$\mathrm{nMAPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|y_j - t_j\right|}{\frac{1}{N}\sum_{i=1}^{N} y_i} \times 100\% \tag{5}$$

where y_j and t_j are the measured and corresponding predicted values of PV power and N is the number of test samples [26].

Some researchers [67], however, have opined that the use of statistical error metrics for model's performance assessment is not sufficient. They suggested more application oriented evaluations, i.e. calculating the optimal accuracy of a particular forecast based on system economics and major planning aspects. The domains include: economic dispatch, optimal energy storage size, emerging energy market policies at local and international levels, profit maximization for energy market stakeholders and optimal reserve size determination.

Irrespective of the performance metric employed, the goal of the researcher must be the objective evaluation of a particular forecast model so that practitioners are able to make the best decision in terms of design, installation and utilization of photovoltaics for grid applications.

2.4. Forecast model inputs

Inputs to forecasting models have direct influence on prediction accuracy; a key factor in determining model performance. Generally, imprudent input selection can cause forecast errors which increase time delay, cost and computational complexity. Thus, low accuracy rate could appear on high functional forecast models [16]. The inputs for PV systems are mostly meteorological parameters: solar radiation, atmospheric temperature, module temperature, wind velocity and humidity [68], barometric pressure [56,57,69] and aerosol changes [70–73]; all dependent on climate condition and geographical location. Hence, correlation between PV output and meteorological inputs varies and can be positive or negative; strong or weak [13].

i Use of secondary data

In Ref. [74], the influence of forecast horizon on PVPF accuracy was studied using numerically predicted weather data via support vector regression (SVR) [75]. compared two ST-PVPF models: analytical PVPF (APVF) and multiplayer perceptron PVPF (MPVF), with both models exploiting numerically predicted weather data and past hourly values for PV electric power production. The RMSE values were similar; varying from 11.95% to 12.10%, with forecast horizon being all daylight hours for 24 h ahead predictions. The conclusion was: both approaches are suitable for PV output power prediction.

ii Use of primary data

Alomari, et al. [76] used two years' worth of hourly PV power output data from solar power plant installed at Applied Science Private University (ASU) in Amman, Jordan. Concurrent weather data measured at the same location was also considered. The data were analysed via ANNs optimized by the Levenberg-Marquardt (LM) and Bayesian Regularization (BR) algorithms. The models were able to correlate temperature, solar irradiance and timing with PV power generated for real-time day-ahead predictions. Finally, BR ANNs based PVPF model was selected on the basis of performance.

2.4.1. Solar irradiance

Solar irradiance is the radiant energy from the Sun emitted in the form of electromagnetic radiation. It is directly proportional to the amount of solar power that can be hasrvested and has the strongest correlation with PV power output. De Giorgi, et al. [77] proposed an artificial neural network (ANN) forecast model to analyse the correlations of inputs with model performance. The forecast errors were (NRMSE) 12.57%, 12.60% and 10.91% for input vectors of the historical PV output, solar irradiance and module temperature, respectively. Therefore, study of correlation between meteorological parameters and PV output is essential for forecast model designing.

Fig. 3 represents the pattern of solar irradiance and PV output for a specific day. The experiment was conducted on the roof of an institutional building of the University of Malaya (latitude = 03°09′N; longitude = 101°41′E) Kuala Lumpur, Malaysia [68]. The PV power varied from 7 a.m. to 7 p.m. (day break to dusk) and its magnitude followed the trend of solar irradiance intensity. As the day progressed, the PV output increased and peaked during mid-day; the time when solar irradiance is most intense.

Fig. 4 shows a positive correlation between PV power output and solar irradiance for the same experiment. The correlation coefficient R^2 was 0.988, indicating solar irradiance is a most significant input for the forecast.

2.4.2. Temperature

Temperature is a physical quantity which measures the intensity of heat available in a substance or object or environment. This input parameter is highly likely or unlikely is a coefficient to PV power generation. Some research studies indicated that, there is inadequate

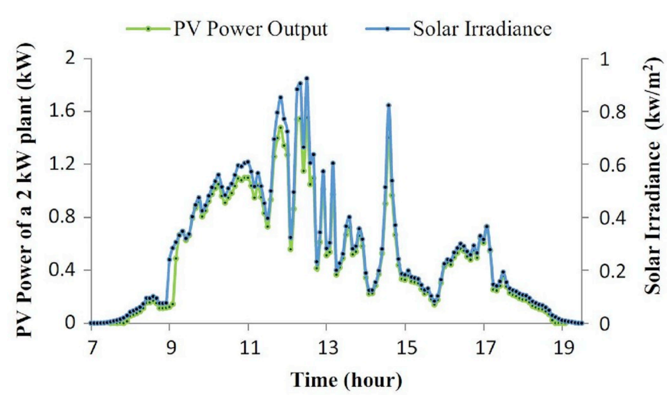


Fig. 3. Solar irradiance and PV output curve for a specific day [13].

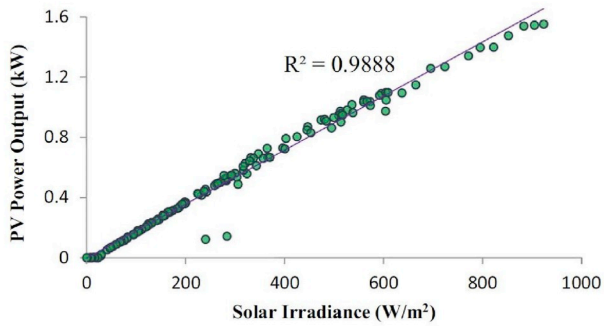


Fig. 4. Correlation between solar irradiance and power output of PV panels [13].

correlation between ambient temperature and PV output power.

Fig. 5 represents the correlation between atmospheric temperature and PV output for the same research work. The correlation coefficient R^2 is less significant at 0.3776. The researchers also found that there was no correlation between atmospheric temperature and PV output during nightfall; for obvious reasons.

2.4.3. Clouds

Cloudiness is one of the most significant factors that determines the intensity of solar radiation incident on terrestrial surface. Furthermore, the fraction of solar energy that passes through the clouds is not constant and depends on the type and amount of cloud formations [78]. Cloud motion, birth, dissipation and deformation, cause fluctuations in sunlight intensity; thus affecting PV output [79,80]. Most investigators have used satellite images for cloud analysis [81–86] which lack precision for regional or low cloud formations due to low spatial and temporal resolutions [87]; especially when applied to ultra-short-term forecasts. This has spurred the use of terrestrial sky imaging techniques.

Raza, et al. [16] investigated PV output variations with solar irradiance during cloudy, partially cloudy and clear sky days, and claimed that during cloudy days more noticeable fluctuations in solar irradiance were perceived compared to partially cloudy and clear sky days. Ogliari, et al. [88] demonstrated that for clear sky conditions, PVPF was highly accurate for day-ahead predictions. They compared the PV output using physical models, based on three and five parameters electric equivalent circuits, and statistical models of ANN [89]. Similarly, Alanazi, et al. [90] used meteorological data of past 15 years, GHI and clear sky irradiance data for training their ANN and acquired good precision.

Another study reported that PV output increased with the rise in ambient temperature, which is again related to cloud cover, by showing high correlation between the input and the output [13]. Raza, et al. [16] investigated the relationship between cloud cover and PV power output. They analysed PV output variation with solar irradiance during cloudy day, partially cloudy day and clear day. They claimed that during cloudy day more noticeable degree of fluctuations with solar irradiance are perceived compared to partially cloudy and clear day. Again, another approach has been the use of digital cameras for taking pictures of clouds. Sky imagers are used to capture cloud pictures from horizon-to-horizon for the purpose of modelling cloud cover and cloud motion [91,92]. The Solar Forecasting Initiative at the University of California Merced was able to use such information for terrestrial surface solar irradiance and power predictions. Very short-term forecasts (minutes ahead) of future cloud patterns over solar generation facilities is possible via such modelling initiatives [93]. Although, cloud cover based PV power output models suffer from two obvious shortcomings: cloud speed and forecast horizon [94]; they can be more accurate than satellite images for cloudy days [95].

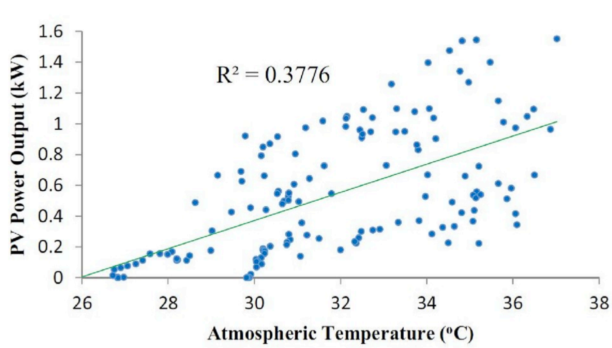


Fig. 5. Correlation between atmospheric temperature and power output of PV panels [13].

However, existing models are not very robust in determining cloud motion from sky images due to clouds' complex motion patterns. Therefore, Zhen, et al. [51] established a new method of classifying sky images and subsequent model optimization. They analysed three mainstream methods, i.e. block matching, optical flow and SURF feature matching algorithms, and proposed a pattern recognition and PSO optimal weights based model for predicting cloud motion. This approach was used for PVPF using real-time sky image data and proved to be very effective.

Wang, et al. [54] utilized historical cloud formation types and distribution for predicting future sky status using CMDVs and FPCT to extract relevant information from sky images. Their method is suitable for very-short-time intervals (<1 min) under fast changing cloud speeds. This partly simplifies the complicated analysis involved in cloud motion studies and to aid in computation speed they applied an IPSI method with their CMDV calculations. Thus, their developed model is very conducive to short-time and ultra-short-time scale PVPF. It proved to be a better approach during simulations with secondary weather data than the original FPCT, optimal flow (OF) and particle image velocimetry (PIV) methods.

Another cloud model was developed by Ishii, et al. [58], where fluctuations of weather based on cloud formation and PV power were considered. Their study was conducted on real-time weather data taken in Japan and using pc-Si, a-Si:H/sc-Si and copper indium gallium (di) selenide modules for PV panels. They claimed that the performance of PVs is influenced by solar spectrum even under identical solar irradiance conditions. Spectral factor (SF) index was used to indicate the solar irradiance between actual solar spectrums and standard AM1-5-G spectrum. Cases of clear sky and cloudy sky conditions were analysed and their effects on PV output quantified using the reciprocal of SF (SF^{-1}). It was found that only in fine weather, the SF^{-1} predicted that solar spectrum has insignificant impact on PV output because of the "offset effect" in PVs. While, in case of extreme weather conditions in Japan, it was found that the performance of Si:H modules can vary by more than ±10%. Thus, their SF^{-1} factor could efficiently predict seasonal and location dependent variations in PV output based on cloud formation, and could also specify the best type of PV panels to be utilized.

Detailed discussions on input selection, pre-processing and parameter selection for PVPF in terms of weather classification are included in section 3.

2.4.4. Wind speed and cell temperature

Wind speed is a crucial input in PVPF models but due to its intermittent nature, it introduces significant uncertainties [96,97]. Its main role is in heat dissipation and reducing PV cell temperature. Generally, the efficiency of PVs depends on the temperature of the module, which rises during operation because of changes in ambient temperature or due to radiation absorbed by it. Therefore, both wind and cell temperature values are important and linked.

In Ref. [98], the authors analysed wind speed and direction data from their ASU solar PV system for building their ensemble of ANNs models for PVPF. They found that wind direction data at 36 m altitude, where their PV plant was installed, demonstrated binary or constant values during the past 1.6 years; and thus, it was excluded. However, wind speed was included and their ensemble model was superior to established approaches: smart persistence and individual ANN models.

Raza, et al. [16] determined by inspection that PV output power pattern does not exactly follow wind speed pattern. During daytime, the PV output power is at a higher level compared to wind speed. Subsequently, this output power decreases with increase in wind speed; however, similar observations were not made by others during the same period. Therefore, according to these researchers, the correlation between PV output and wind speed is weak.

[99,100] utilized cell (module) temperature of PV panels as input for their PVPF models, claiming that it is strongly correlated with PV output [101]. Ting-Chung and Hsiao-Tse [100] modelled cell temperature as a function of solar irradiance, ambient temperature and operating variables related to PV cell technology, and compared it to real world measurements. Also, in contrast to Ref. [16], Schwingshackl, et al. [99] critically analysed different methods for the accurate estimation of PV module temperature and identified wind effects as a major factor. Thus, consideration of cell temperature is vital if PV outputs from different types of modules are to be compared and predicted, as each type of PV varies in its response to changing temperature of its cells. Indeed, such an investigation is recommended by Ref. [98] for aggregating the influences of ambient temperature and global solar radiation on PV module technology. However, in Ref. [98], the researchers did not include cell temperature as an input in their ensemble of models approach to PVPF. Firstly, as cell temperature can be surmised and summarized from wind speed and ambient temperature. Secondly, at the current stage of their research only one PV power plant was considered with identical PV modules, thus negating the need for consideration of cell temperature.

2.4.5. Timestamp

The time stamp of the weather plays a significant role in the prediction accuracy of PVPF model. In Ref. [98] the current time stamp (in hours beginning from the start of each year data, represented as $T(t, d)$) was optimized by trial-and-error method and it was observed that the past $i = 5$ days (used as an embedding dimension) tended to provide the most accurate day-ahead PVPF.

The time stamp (hour of the day studied) introduces uncertainty in the forecast [98] which is significant during mid-day (due to higher weather variations) compared to that at early hours or late evenings. Furthermore, these uncertainties of power predictions change depending on location and season. Therefore, to develop a new PVPF model for a new region, time stamp data must be optimized for obtaining the required optimum time period [98].

ANNs are recommended [98] for this purpose of PVPF as these networks can elucidate the complex input/output relationships, like current time stamp and historical weather data with PV output. In Ref. [76] the authors also utilized time stamp (from the start of the current year $Td(t)$) and weather input vectors ($Wd(t)$) for the same time (t), with the period being past five days prior to day (d), to develop two ANN based PVPF models using LM and BR algorithms.

2.4.6. Summary of model inputs

Solar irradiation has the largest correlation with PV output compared to atmospheric temperature, cloud cover and other meteorological input parameters, such as dew point, relative humidity, precipitable water, air pressure etc. Table 2 lists the correlation between meteorological parameters and PV output power [89].

The accuracy of a PVPF model can be enhanced by using many relevant inputs which are highly dependent on local weather conditions. However, imposing every input vector on a forecast model is neither viable nor computationally expedient. Therefore, a future challenge is designing a PVPF with optimum number of inputs based on their strong correlations with PV output.

Table 2
The correlation between PV output power and meteorological factors.

Meteorological factor	Correlation coefficient
Solar irradiance	0.9840
Air-temperature	0.7615
Cloud type	−0.4847
Dew point	0.6386
Relative humidity	−0.4918
Precipitable water	0.3409
Wind direction	0.1263
Wind speed	0.1970
Air pressure	0.0815

3. Input data processing techniques

The quality of input data is crucial for accurate and reliable forecasting. Many research projects have employed historic time series data of PV output as well as meteorological information of specific power stations and their geographical locations for modelling purposes. However, these datasets often have intermittent static or spike elements caused by weather or seasonal variations, electricity demand fluctuations and power system failures. These are outliers which follow no trend, are influenced by chance events and significantly affect the forecast. Moreover, data may sometimes be corrupted or missing due to sensor defects or erroneous recordings. Therefore, it is imperative to preprocess distorted input data by reconstruction using decomposition, interpolation or seasonal adjustments [102] (i.e. data cleansing and structure change). To accomplish this, several techniques are mentioned in the extant literature such as wavelet transform (WT), trend-free time series, empirical mode decomposition, SOM, normalization and singular spectrum analysis [13], each of which has its own unique strengths and shortcomings. Bacher, et al. [103] and Kemmoku, et al. [104] reported that for solar power prediction, time series technique has been exploited on clear-sky index information for pre-processing input data in several research works. Sfetsos and Coonick [105] however disagreed with the concept and opined that it is random in behaviour and highly sensitive to weather changes. Thus, it yields poor learning rate of the input data which results in high prediction errors. Other studies showed solar irradiance information was more significant and effective in PV output forecast compared to time series of clear-sky index data. Reikard [106] used statistical tool to eliminate the seasonality trend from solar irradiance data. Their study demonstrated amended learning rate with improved prediction accuracy for the developed model. Similarly, Baig, et al. [107] and Kaplanis [108] utilized a trend and de trend technique for solar irradiance dataset as it is quite complicated to precisely determine the trend of daily solar radiation due to everyday weather behaviour. In Refs. [109,110] Boland found low prediction error for periodic pattern of solar irradiance dataset using Fourier series transformation.

Currently, among the data pre-processing techniques normalization [48,111] and WT [38,112] are broadly used as they can convert large input data to a smaller range hence improving computational economy. In normalization, the large number of data are compressed and transformed into a smaller range. The condensed data is confined between 0 and 1 range to limit the regression error and maintain correlation among the datasets. For WT, the concept of a wavelet (miniature wave) with a zero-average value is employed and can produce time-frequency representation of the signal simultaneously and is used for decomposition and reconstruction of signals [113]. It can overcome the problem with non-stationary datasets. Therefore, it has the ability to transform the signals into time-frequency domains and decompose the signals to one approximate set and several detailed sets [114].

Alomari, et al. [76] applied normalization to their data which consisted of global solar irradiance, temperature, PV power and time of year for developing next-day PVPF models. They claimed that normalization helped to achieve homogeneous data and reliable machine learning experiments. Moreover, data filtering and association were utilized for weather and PV power data. This pre-processing filtered out any weather data with its associated PV power value missing or any PV power record with missing weather data [53,115,116]. detailed steps for efficient pre-processing:

1. Negative solar radiation and missing associated production values are often observed in early and late hours of the day, due to solar radiation sensors offset and inverter failures, respectively. It is recommended to set the radiation and PV output values to zero (0).
2. Missing solar radiation, temperature and output power data, during mid-day, can be from solar radiation and temperature sensors malfunction, and inverter or network disruptions, respectively. The recommended step is to exclude these from analysis.
3. Normalization of remaining data [range: 0, 1] will increase computational speed, preserve input correlations and ensure fast convergence of ANNs [26,68].

GAN can also generate missing data or augment the time series data for inputs, such as creating more distinct instances of extreme weather data [53]; serving as a pre-processing technique. Furthermore, the GAN pre-processing step for generating distinct weather data for subsequent modeling by CNN was effective in addressing diminutive sample size of certain weather types during CNN training [53].

To reiterate, weather classification is an effective pre-processing tool; enhancing the accuracy of ST-PVPF [40,48,52]. Moreover, it can help select the optimum PVPF approach for the location and climate [53]. However, only weather classification is not sufficient in training high-capacity PVPF strategies, such as CNN [56,57]. Wang, et al. [57] claimed that small training data size can worsen the performance of most PVPF approaches.

Zhen, et al. [51] discussed pre-processing of cloud images, classifying them based on brightness, size, shape, spectrum, texture features etc. Cloud texture was analysed using gray-level co-occurrence matrix (GLCM) and histogram equalization. Four GLCMs were generated for each sky image incorporating the corresponding energy, correlation, entropy and contrast values as the 4-dimensional texture feature vector of each image. Few others have documented such optimizations of extracting cloud features for efficient classification [53,117–119].

Nevertheless, no single pre-processing technique can be considered to be the best; it depends upon the nature of the datasets. Hence, the efficiency of a pre-processing technique is evaluated by: computational time, resources utilized, accuracy of the processed data, data fidelity, consistency of performance, adaptive nature, robustness and compatibility with established forecast models.

Another important factor is effective design of neural network. Network parameters, such as weights, biases, number of hidden layers and neurons in a layer, are crucial for prediction accuracy. Wang, et al. [53] used statistical loss functions *LG* to set the weights G's of CNN for their technique and claimed that a well-designed network architecture with GAN support is best for deciding on the parameter values of weights and biases for a robust PVPF model.

Parameter selection and optimization (weights and biases of neurons, and network architecture) was also addressed by Zhen, et al. [51]. They propounded that the complexity of cloud formation and motion cannot be done by a single traditional algorithm. Image processing based cloud pattern classification and PSO for optimal weights determination were recommended. This combination approach was applied to three established models: block matching, optical flow and SURF feature matching, for efficiency comparison using actual sky images.

Alomari, et al. [76] obtained excellent PVPF accuracy by using BR and LM back-propagation optimization algorithms for designing two different ANNs. BR technique suggested 28 hidden layers and all weather inputs while LM dictated 23 layers for the data, with the BR-ANN having a lower RMSE.

The LM optimization approach was first reported in Ref. [120] and successfully applied to ANNs in Ref. [121]. It is considered to be the fastest back-propagation supervised algorithm for training feed-forward ANNs. BR algorithm was introduced in Refs. [122,123] and adapted for use in Ref. [124]. Both algorithms determine an ANN's error derivatives regarding weights and biases, and output a Jacobian matrix; this being used to calculate performance by mean squared errors [121].

Finally, post-processing is a requirement prior to any model's forecast evaluation. The two most popular post-processing approaches are anti-normalization [48] and wavelet reconstruction [125]. If normalized data are utilized in the prediction model, then the forecast should first be anti-normalized in order to elucidate the actual forecasted PV power and subsequently to assess the model's performance. Similarly, wavelet reconstruction is employed to obtain the actual forecasted PV power if the model's input data were initially pre-processed using wavelet decomposition (WD). The adaptation of the model based on new real-time or secondary data, as they become available, is a crucial step in improving the prediction accuracy of PVPF models and, hence, a recommended post-processing step further discussed in section Online PVPF and Sequential Extreme Learning Machine [126].

4. Optimization of input parameters

The proper selection of inputs, in terms of number and type, is a prerequisite to improved forecasts. Redundant or weakly correlated inputs will introduce undue complexities in computation, whereas the absence of a major parameter can severely offset the predictions. Thus, optimization algorithms are essential for helping select the most important input parameters.

The literature describes many optimization techniques for inputs of PV output models: PSO [127], grid-search [128], fruit fly optimization algorithm (FOA) [129], firefly (FF) [130], ant colony optimization (ACO) [131], chaotic ant swarm optimization (CAS) [132], chaotic artificial bee colony algorithm [133] and immune algorithm (IA) [134]. While each has its pros and cons, genetic algorithm (GA) based optimization is the most popular and effective in optimizing weights and inputs for forecasting models, and works well with ANN. Tao and Chen [135] optimized the input weights for their back propagation neural network (BPNN) model using GA and obtained better forecast accuracy. Similarly, Pedro and Coimbra [33] claimed that their ANN forecasting model for PV power was enhanced by GA optimization.

In Ref. [51] paper, it was claimed that PSO is similar to GA and is also an iterative optimization algorithm. The system is initialized as a set of random solutions in PSO, and the optimal values are searched iteratively. Compared with GA, it is easier to achieve convergence and does not need to be adjusted for too many parameters [136,137]. At present, PSO is widely used in function optimization, neural network training and parameter selection, fuzzy system control, and in other applications as a substitute for GA. However, PSO has a higher tendency than to get isolated in local extrema.

5. Classification of PV power forecasting techniques

Several modelling approaches: physical, statistical, artificial intelligence (including deep neural network), ensemble and hybrid based prediction models have been utilized for PV output forecasts. Some researchers have compared different forecast models. Almonacid, et al. [138] compared an ANN model for PV output forecast with three conventional mathematical schemes. The forecast accuracy of ANN was much superior. Similarly, another researcher Oudjana, et al. [139] compared forecasting techniques between ANN and regression model for a particular PV power plant in Algeria. Again ANN displayed higher accuracy. Yet others [24] have compared several different forecasting approaches taking into account the techniques, the spatial-temporal horizons, the metrics assessment employed, the input and output parameters modelled, the computational time, the benchmark models referred, the advantages and the disadvantages etc. From such exhaustive comparisons it is clear that PV power output forecasting accuracy depends on the type of modelling utilized, and hence, some of the important ones are discussed below:

5.1. Persistence forecast

Persistence model is popular for very short-term and short-term forecasting. It has less computational cost, low time delay and reasonable accuracy. This technique adopts the concept of today equals tomorrow, in other words, the conditions of climate (i.e. solar irradiance) a day ahead is expected to remain similar to the day before [140]. For instance, if today is a sunny day with 30° Celsius then the model would predict that it will be a sunny day with 30 °C temperature tomorrow. Therefore, it can reckon GHI instantly by decomposing the forecasting GHI information into the computation clear-sky GHI and project the clear-sky index. However, the clear-sky index does not respond to the variation of solar zenith angle due to weather conditions such as cloud, rainfall, storm etc. at the forecasting time window [141]. The equation of forecast model using this approach is described as [140]:

$$P(t+k|t) = \frac{1}{T}\sum_{i=0}^{n-1} P(t - i\Delta t) \tag{6}$$

where k is the forecasting time period and $P(t+k|t)$ is the predicted power for time $t+k$ at time instant t. The forecast interval period is T and n is number of historic measurements. $P(t-i\Delta t)$ is real power measured for time t and time steps i within T and Δt is step time difference of the measured time series.

This model generally endorses the performance of other forecasting paradigms. Hence, different models are benchmarked against it when climate conditions remain persistent up until next day; however, with increase in time horizon, the accuracy of this model decreases drastically.

5.2. Physical model

Physical forecasting involves: air pressure, surface roughness, temperature, orography, impediment, disruption and obstructions, of lower atmosphere for future predictions [142–144]. This technique is generally more reliable in long-term forecasting [102]. Based on this method, the NWP technique amalgamates meteorological information and atmosphere model equations for arriving at predictions. This model is usually categorized into two types based on scale, i.e. mesoscale model and global model. Mesoscale model deals with the atmospheric features for a restricted area such as regions, countries or continents [145]; whereas global model delineates the features of the atmosphere on a global scale. Furthermore, for this global NWP model, there are about 15 weather services available that are active in data acquisition: Global Forecast System (GFS), Climate Forecast System (CFS) and Global Data Assimilation System (GDAS) etc. which are managed by government sponsored organizations such as the US NOAA and European Centre for Medium-Range Weather Forecasts (ECMWF). NWP models can forecast climate status more than 15-days ahead [146] and use a set of numerical equations for the physical state and the dynamic characteristics of the atmosphere, simultaneously. Mathematically NWP can be represented as:

$$\frac{\Delta A}{\Delta t} = F(A) \tag{7}$$

where, ΔA is the change in the value of the forecasted response at a particular spatial location; Δt is the change in time or temporal horizon; and $F(A)$ represents variables which change the value of A.

However, for PV output power prediction, the model employs particular weather characteristics such as GHI, relative humidity, wind speed and direction [147–149] and so the forecast quality is better if the weather variables remain stable [150]. In contrast, erroneous forecasts occur when there exist abrupt changes in values of meteorological variables.

5.3. Statistical techniques

Statistical approaches use historical time series and real-time generated data. It consumes fewer input data compared to DL methods and shows better performance in short-term prediction than NWP models. The statistical models use pure mathematical equations to extract the pattern and correlation from past input data. Generally, the basic algorithm possesses curve fitting, moving average (MA) and autoregressive (AR) models [151]. These models reduce error by estimating the difference between the actual past measured value and the predicted value of PV output. Hence, prediction accuracy depends on the quality and dimension of the data. Statistical techniques can be divided into two groups: machine learning, i.e. artificial intelligence and time series based forecast models [16].

5.3.1. Time series based forecasting techniques

Time series provides statistical information to foresee the nature of the quantified element. These observations are generally recorded over time at successive points in regular intervals such as quarterly, monthly, weekly, daily or even hourly and minutes, depending on the variable's response with time [24]. The motivation of time series analysis is to predict the forthcoming value by evaluating the pattern of past information. For instance, Cornaro, et al. [152] employed statistical methods to establish the correlations between the meteorological past data and hourly solar irradiance. Established techniques are: exponential smoothing, autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA).

i Exponential smoothing

Exponential smoothing method or exponentially weighted moving average (EWMA) is a unique technique that adopts exponential window function for statistical analysis of historical time series data to make predictions. Generally, it allocates an unequal set of weights over equal weights to historical observations, thereby exponentially reducing the data from the most recent to the most distant data points. However, it can easily learn and make determination from assumptions. The technique was first formulated by Brown [153] and has since seen many applications. Overtime, it was extended by Holt in 1957 and by Winter in 1960. It is thus called Holt-Winter's method [154]. The governing equation is as follows:

$$\widehat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\widehat{Y}_t = \widehat{Y}_t + \alpha(Y_t - \widehat{Y}_t) \tag{8}$$

where, current observation is Y_t; predicted value is $\widehat{Y}_t$; and smoothing constant is α, which remains between 0 and 1. Therefore, the forecasting equation outputs the predicted value at $t+1$ which is equal to the sum of the last predicted value $\widehat{Y}_t$ and the forecasted adjustment factor $\alpha(Y_t - \widehat{Y}_t)$.

ii The autoregressive moving average model (ARMA)

ARMA is a time series statistical analysis frequently used in forecasting. The model has been evaluated by many researchers in different applications of forecasting (solar and wind forecasting) and it has consistently performed with good prediction accuracy. The model incorporates two polynomials: AR and MA for forecasting the PV output from historical data [155]. The mathematical expression is as follows [18]:

$$X(t) = \sum_{i=1}^{p} \alpha_i X(t-i) + \sum_{j=1}^{q} \beta_j e(t-j) \quad (9)$$

where, predicted PV output is represented through function $X(t)$ which is a combination of AR and MA functions. p and q indicate the number of processes or the order, while α_i and β_j are the coefficients of AR and MA models, respectively. $e(t)$ is randomly generated white noise; it is not correlated with a model's predictions. ARMA models are very flexible and can represent several different time series by using different orders. Furthermore, this model can recognize when there is an underlying linear auto-correlation structure. Mora-López and Sidrach-de-Cardona [156] in their research adopted ARMA models to investigate the stationary and sequential features of the global irradiation time series and forecast its pattern. Similarly, Hansen [157] found ARMA model was suitable in stationary data analysis. However, the approach requires static time series data which is a significant disadvantage.

iii Autoregressive integrated moving average (ARIMA)

ARIMA is also known as Box-Jenkins model and was developed by George Box and Gwilym Jenkins in 1976 [158]. ARIMA model is an extended version of ARMA and it is a popular time series analysis technique as it supports standard level of forecast accuracy for short term horizon. Moreover, this model has the ability to clip non-stationary values from the analysed data. Its structure consists of autoregression (AR), integration (I) and moving average (MA) to evaluate and predict time series characteristics [118]. The general form of ARIMA (p, d, q) model of time series X_1, X_2, X_3 is as follows:

$$\Phi_p(B)\Delta^d X_t = \Theta_q(B)a_t \quad (10)$$

$$\Phi_p(B) = 1 - \varnothing_1 B - \varnothing_2 B^2 \ldots \varnothing_p B^p \quad (11)$$

$$\Theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 \ldots \theta_q B^q \quad (12)$$

where, the backward shift operator is B; backward difference is $\Delta = 1 - B$ and $BX_y = X_{y-1}$; Φ_p and Θ_q are polynomial numbers of order p and q, respectively. As a result, the ARIMA (p, d, q) model is a composite sum of autoregressive part (p), an integrating part $I(d) = \Delta^{-d}$, and a moving average part (q). The variables in Φ and Θ are precisely selected so that the zeros of all polynomials fall out from the unit circle to evade the creation of interminable processes. To consider the arbitrary disturbance taken from a fixed distribution with zero mean and σ_a variance a_t, a_{t-1}, a_{t-2} are introduced as white noise process. Therefore, the intrinsic characteristics of the time series can be comprehend by the white noise process and backshift operator. Hansen [157] integrated ARIMA to predict global irradiance field and to summarize the coupling procedure between AR and MA as well as how it treats non-stationary series. Reikard [106] used ARIMA method to forecast solar radiation and compared the predicted results with those of an ANN based model. The outcome indicated that the ARIMA model at 24 h time span estimated more accurately than the other models studied. Similarly, Cadenas, et al. [159] found ARIMA model has very close accuracy rate with ANN based models. In another approach, an ensemble technique was used to combine statistical models with other models to generate a hybrid and thereby to exploit the strengths of all the member models. These compound models generally have better performance compared to conventional ANN or statistical models.

5.3.2. *Machine learning forecast techniques*

The other statistical modelling harnesses the advances in machine learning, an approach which is based on computing or artificial intelligence. The method relies on the ability of AI to learn from experience with historical data and to further hone its predictive abilities via training runs. Powerful computers are required to run numerous iterations before a final prediction can be achieved. It can perceive impossible representations without any preordained formulas or equations. Its applications abound: pattern recognition, data mining, classification problems, filtering and forecasting. The main techniques of machine learning are ANN, multi-layer perceptron neural network (MLPNN), recurrent neural network (RNN), feed-forward neural network (FFNN) and Feedback neural network (FBNN).

5.3.2.1. Artificial neural network. ANN mimics the information processing mechanism of the human brain. It has a unique ability to approximate nonlinear functions with high fidelity and accuracy; and is being utilized in such diverse fields as meteorological predictions, finance, physics, engineering and medicine. Fig. 6 is a basic representation of the network.

The basic ANN architecture is divided into three sections: input layer, hidden layer and output layer, comprising artificial neurons and connections. As a similitude of biological neurons, each artificial neuron of a neural network is an activation node where all information processing and decision making activities take place. A particular activation node takes input from a previous node and applies machine learning parameters to generate the weighted sum. This processed information is then passed on to an activation function to compute the composite prediction. This generated prediction is progressively processed until it reaches the desired output. The most frequently used activation functions are Sigmoid, Hyperbolic tangent sigmoid, Gaussian radial basis, Linear, Unipolar step function, Bipolar step function, Unipolar linear function and Bipolar linear function. Along with different activation functions, in the last few decades, ANN forecasting techniques have undergone many modifications to accommodate disparate input-output projections and architectures.

i Multi-layer perceptron neural network (MLPNN)

Many researchers treat MLPNN model as a benchmark [160,161]. It is a technique for elementary and effective ANN approach to designing and prediction. It is so powerful that this network is used in universal approximation, and in nonlinear modelling and complex problems which cannot be solved by an ordinary single layer neural network [162–164]. Generally, MLP is a composite of three or more layers of incoherently activating nodes. These nodes in any layer are connected through a certain amount of weight to other nodes in the next layer. Therefore, it has the capability to correlate the input and output relationship through adequate learning. The correlation between the number of nodes and the hidden layer are essential. Hegazy, et al. [165] observed that a single hidden layer is adequate enough to design a complicated nonlinear function having a sufficient number of hidden

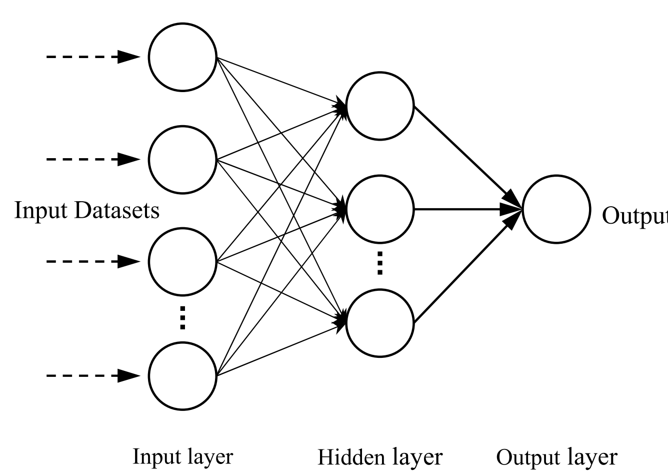


Fig. 6. Basic structure of ANN [24].

nodes. However, overfitting and training difficulties occur due to the rise in the number of nodes [166]. Fig. 7 shows the simplified structure of an MLPNN with eight inputs for solar power prediction.

ii Recurrent neural network (RNN)

RNN is a prominent class of ANN which can learn and process different complex and compound relationships as well as computational structures. This network provisionally relies on time series data by feedback system to inherit the previous time step values; demonstrating temporal dynamic characteristics. The model has a simple structure with built-in feedback loop which allows it to act as a forecasting engine. RNN output of the concerned neural layer is summed with the next input vector and fed back into the same layer which is the only layer in the entire network. Hertz, et al. [167] extensively discussed the basic application of RNN model. The applications are incredibly versatile ranging from speech recognition to driverless cars. In another research, a RNN model is highlighted by Elman [168] where the architecture has a feedback loop that can communicate with the hidden layer of the network and the input layer. Jordan, et al. [169] proposed a slight modification to the network architecture by connecting the feedback loop from the output layer to the input layer. All these feedback loops help to reduce the learning error. In RNN, each active neuron is connected with all the other processing neurons and to itself. Hence, the output result of RNN is inclined to the feedback signal at the previous time step and the input signal. Williams and Zipser [170] conducted a series of experiments on RNN at real time to analyse the learning rate of the model. The activation function of RNN is the weighted sum of input signals and feedback. Thus, the activation function equation is as follows:

$$S_k(t) = \sum_{p=1}^{p+1} \left(W_{ik} X_p(t)\right) \sum_{q=1}^{q} V_{kq} Y_q(t) \tag{13}$$

where, $S_k(t)$ is the activation function at the given time t when processing node k and the connection weight is V_{kq} that corresponds with node q which is connected with node k, and $p+1$ is the value of the bias. Fig. 8 illustrates the relationship.

iii Radial basis function neural network (RBFNN)

RBFNN is a quicker and better approach to machine learning than other ANN approaches. Hence, it is used in approximation, time series prediction, classification and system control. The structure uses radial basis functions as activation functions. This network generally has two layers. The characteristics are merged together with radial basis activation function in the first layer, and then the output of the first layer is used to compute the same output in the next time step. Both layers can be identified through their synaptic weight. The weighted value of the first layer is generated from the input information while the weight of the second layer needs to be determined from the calculation. The RBFNN can learn through unsupervised method as only input data is fed into the network. The network equation is as follows:

$$Y_k(x) = \sum_{j=1}^{M} W_{kj} \varphi_j(x) + W_{ko} \tag{14}$$

Other methods in ANN for forecasting PV output involve: BPNN [171,172] and Elman neural network (ENN) [173,174]. A major condition for accurate and consistent forecasting for all techniques is a reliable historical data set. Therefore, various methods are used to pre-process and post-process input data before forecasting [175], e.g. empirical mode decomposition (EMD) [176], trend-free time series [177], normalization [48], WT, complementary ensemble empirical mode decomposition (CEEMD) [102] etc.

Another versatile ANN approach is the synergy of two or more different methods to exploit the best attributes of each; hybrid modelling technique. Such models have highly accurate and consistent forecasts. To exemplify, WT, ANN and SVM were incorporated together into a hybrid model by Colak and Qahwaji [178], coupling of fuzzy interference model with RNN for solar power generation was done by Yona, et al. [179], PSO amalgamate with enhanced NN was shown by Amjady, et al. [180], hybrid GA and ANN for short-term wind power prediction was discussed by Zameer, et al. [143], and yet other types of hybrid techniques of forecasting were detailed by Qureshi, et al. [64].

iv Extreme learning machine (ELM)

Extreme Learning Machine is an advanced data driven approach for single layer feed forward networks (SLFN) [115]. The network has high enumerate capacity compared to back propagation and gradient descent algorithms with minimum training error and smallest norm of weights.

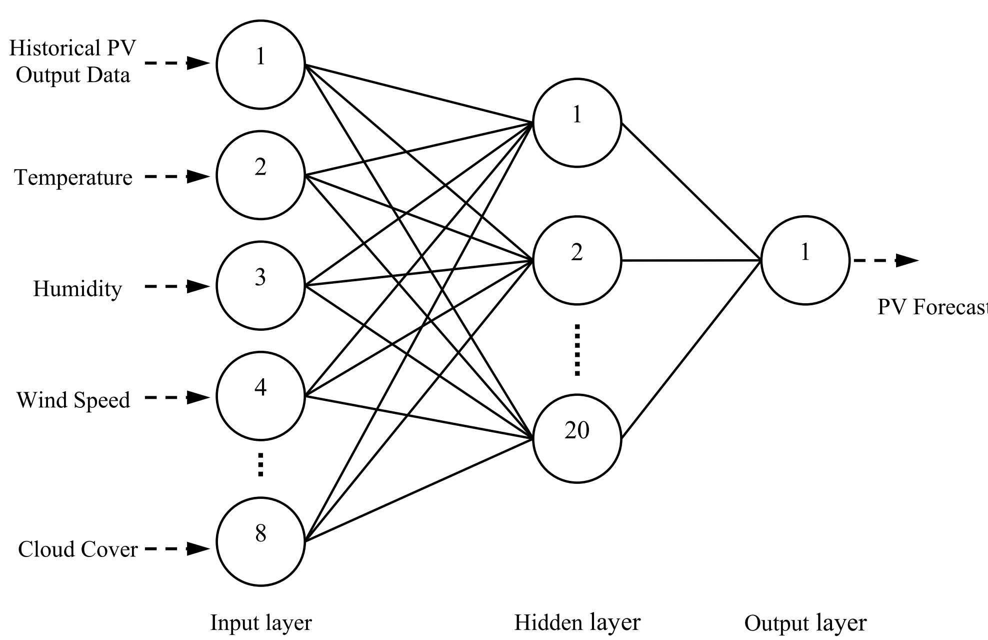


Fig. 7. Multi-layer perceptron neural network [16].

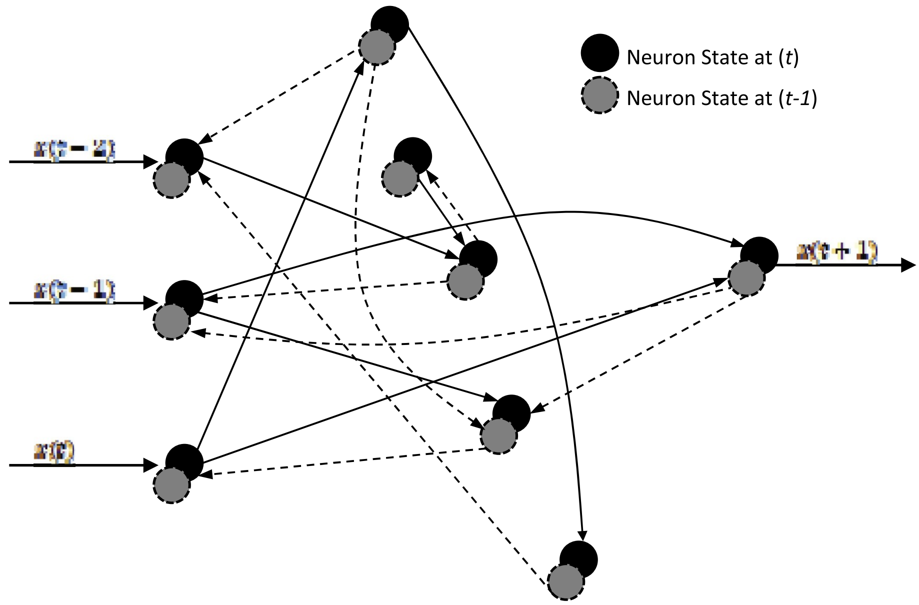


Fig. 8. Basic architecture of Recurrent Neural Network [16].

Therefore, it demonstrates faster learning speed and easier implementation with respect to standard ANNs [181]. Subsequently, it is preferred as a simple algorithm with smaller training input data requirements. Al-Dahidi, et al. [115] used ELM for PVPF due to their simplicity, computational economy and superior generalization ability. Fig. 9 is a typical ELM's flowchart [26].

Huang, et al. [182] proposed modified algorithm of ELM in 2006 [183], where SLFN weights from inputs to the hidden neurons were randomly classified to transform a linear system, while output weights were extracted analytically through Moore Penrose generalized inverse of the hidden layer output matrices. Fig. 10 is a generalized structure of this modified ELM.

The network structure includes input, hidden, and output layers illustrated in Fig. 10. The network output value $y_i \in R^n$ is premeditated based on N training samples $(x_i, t_i) \in R^n \times R^m$, $i = 1, 2, \ldots, N$, K hidden neurons and an activation function $g(w_i \cdot x_i + b_i)$ by the following equation:

$$y_j = \sum_{i=1}^{K} \beta_i g(w_i \cdot x_i + b_i)\ ,\ j = 1, 2, \ldots, N \tag{15}$$

where, w_i and β_i are the connecting weight vectors of input and output neurons to the i^{th} hidden layer neuron respectively, and b_i is the bias of that network. The above equation can be represent as simplified matrix

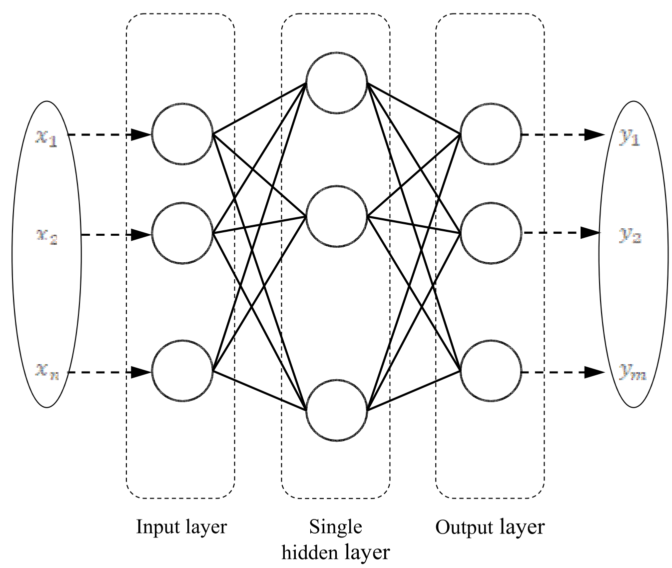


Fig. 10. ELM model structure [184].

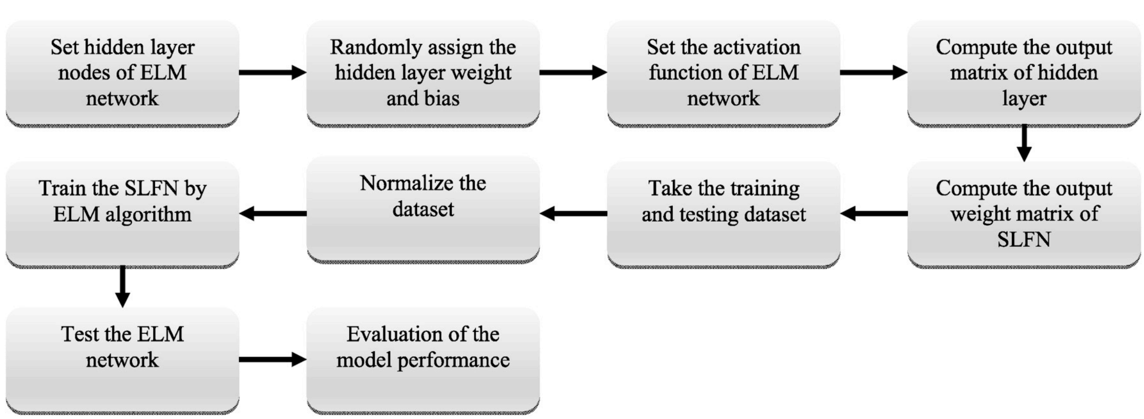


Fig. 9. Flowchart of ELM [26].

format if the number of training set samples is equal to the number of neurons in the hidden layer, that is $y_j = t_j$ [184].

$$H\beta = T \tag{16}$$

The hidden layer output matrix H is

$$H = \begin{bmatrix} h_{(x_1)} \\ \vdots \\ h_{(x_N)} \end{bmatrix} = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) \ldots g(w_K \cdot x_1 + b_K) \\ \vdots \cdots \\ g(w_1 \cdot x_N + b_1) \ldots g(w_K \cdot x_K + b_K) \end{bmatrix}_{N \times K} \tag{17}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_K \end{bmatrix}_{K \times m} \quad T = \begin{bmatrix} T_1 \\ \vdots \\ T_K \end{bmatrix}_{K \times m} \tag{18}$$

β is output weight and T is expected output target.

Hossain, et al. [185] proposed another ELM model which exhibited high forecasting accuracy and learning rate capability compared to SVR and ANN. The study showed that for hourly ahead PV power prediction the RMSE ranged between (55.32–89.55)% in training phase with 0.2496s computational time; while, it ranged between (54.96–90.41)% during validation runs with average 0.0153s computational time. In contrast, for next-day PVPF, the RMSE ranged between (13.83–21.84)% in training with 0.220s computational time and (17.89–35.39)% on validation trials with 0.0335s average computational time.

In Ref. [98], the authors mentioned that PVPF accuracy could be improved with ELMs or echo state networks (ESNs) as base prediction models for the ensemble approach; the former approach being corroborated by Al-Dahidi, et al. [98].

v Online PVPF and sequential extreme learning machine (OS-ELM)

When new data becomes available from same or other closed PV plants, or when the environment in which the PV plant is operating undergoes major changes (evolves); updating the PVPF model is necessary. In the first case, availability of additional patterns can be used for enhancing the prediction accuracy; whereas, in the second case, an evolving environment entails that the prediction model(s) be informed or updated for the new operating conditions [115].

In Ref. [98], the authors discussed the updating of their PVPF model with new data from a PV plant. In some studies, error back propagation (BP) learning algorithm had been exploited to make ANN base-models adaptive to new data as it became available during testing, validation or implementation phase. Ding, et al. [186] had previously developed an improved BP-ANN for more precise 24 h ahead PVPF under different weather patterns. In general, the internal parameters of ANNs (i.e. weights and biases) are initialized randomly and subsequently updated by iteration via BP to reduce error between model prediction and actual PV power production [115,187].

In another study [98], the authors made similar claims as in Ref. [76], stating that LM and BR are popular error BP algorithms for use in model updating with new real-time data, as also corroborated by Ref. [187]. Once such an ANN is trained, it can be used online for PVPF using any test or real world pattern in an adaptive fashion [98]. However, although the availability of new data needs an online learning algorithm, the traditional BP-ANN approach is neither the most effective nor the most efficient solution. It suffers from catastrophic forgetting of previous patterns when the prediction model is re-trained with newly obtained data. Moreover, it is computation intensive and lacks good generalization capabilities.

A better solution was proposed by Al-Dahidi, et al. [115] who employed ELM and optimized its internal parameters using an exhaustive search process for predicting day-ahead PV output. The approach was used, in an online setup, for a 264 kWp PV system installed on the roof of the engineering faculty at the Applied Science Private University (ASU), Amman, Jordan. They compared their forecast accuracy by utilizing three different statistical measures, namely: RMSE, MAE and WMAE; concluding that their predictions were superior to the BP-ANN based PVPF method being currently used at the site [76,116]. Furthermore, the researchers validated using their case study that their ELM approach was better in terms of overall simplicity, computational time and generalization capability. The faster and more accurate prediction accuracy of ELM stems from the way its internal parameters are optimized. In ELM, the parameters, i.e. network weights, are initialized randomly (for input-hidden neural layers) and then calculated analytically (for hidden-output neural layers) as opposed to employing BP algorithm to iteratively set the parameters (weights and biases) in BP-ANN, which increases computational demand. A comprehensive literature search elicits that there are very few studies on the capabilities of ELM data-driven models for online PVPF [115]. [182,188] also corroborated the findings that the ELM approach was superior in terms of generalization and computational economy. Yet another research group, Behera, et al. [26] proposed an enhanced ELM based PVPF method by tuning the network weights using different PSO techniques. Their results surpassed the state-of-the-art BP-ANN approach.

Nevertheless, ELM only takes historical dataset as input for model training, which must therefore be re-trained when new data are acquired. As such, for PVPF at seconds, minutes or hour time-based resolutions, an alternative method with sequential updating capability is required to follow the chaotic and non-stationary PV power patterns [126], with the goal of even better prediction accuracy. Indeed, online sequential extreme learning machine (OS-ELM) has been successfully exploited to address these requirements.

OS-ELM was developed where recursion for the training of new datasets is applied to upgrade the output weights in real time. It feeds the trained model chronologically with new incoming datasets through chunk-by-chunk, block-by-block or one-by-one learning mode [184]. The size of the sequential blocks can be fixed or variable. This algorithm is inspired from ELM, therefore, the number of hidden layer neurons and the corresponding biases are selected arbitrarily; making OS-ELM faster, more accurate and the state-of-the-art compared to its predecessor in online PVPF applications [126,189].

vi Ensemble of models approach to PVPF

The ensemble of models approach, which aggregates the predictions of multiple independent base prediction models, has been effective in PVPF [190–192]. An example of this is the ensemble of ANNs [98]. Weight vectors, internal parameters of ANN, link input nodes to hidden layer, thereby establishing effect of the inputs on the response and biases of the hidden neuronal layer. The output layer, connected to the hidden layer by weights, predicts PV output via activation function. This is parallel processed by a group of ANNs, i.e. base prediction models; the final linear sum being the output of the ensemble approach (for ANN) at a particular time stamp of a specific day.

Moreover, ensemble's forecast accuracy can be improved by generating diversity among its base prediction models [190–192]. This is accomplished by: 1) using disparate prediction methods (SVMs, ANNs etc.); 2) employing the same prediction model, but with differential parameter settings (for example, varying the number of hidden neurons and neuronal layers); and 3) training the individual models by utilizing different datasets [98]. For the last approach, techniques like boosting [193], Bootstrapping AGGregatING (BAGGING) [191,194] and Adaboost [195] are popular.

Single or uniform models lack robustness and flexibility; and therefore, cannot consistently perform accurate PVPF, especially during weather fluctuations [98]. This is the niche for ensemble approach which has consistently shown higher forecasting accuracy while simultaneously quantifying their associated uncertainties [25,196].

In Ref. [98], the authors implemented an Ensemble of ANNs, whose base-models were optimized and diversified using statistical tools to develop a robust and comprehensive day-ahead PVPF model. The model, thus developed, was applied to a PV system in ASU, Jordan and the results compared against established forecasting approaches: smart persistence and single optimized ANN; revealing that the ensemble approach was far superior in prediction and tackling uncertainties. Other researchers include Omar, et al. [197] who developed ensembles of MLP feed-forward ANNs models that receive weather forecasts of the next day as an input and produced more generalized one day-ahead production predictions as an output for a solar facility; and Pierro, et al. [198] who developed a Multi-Model Ensemble (MME) that averaged the 24 h-ahead PVPF obtained by the best different data-driven base models fed with different NWP input data.

5.3.2.1a. Uncertainty prediction and classification of its sources. Uncertainty, a major issue in all PVPF models, must be quantified as a post-processing step for improving accuracy [98]. Various sources of uncertainty affect grid operations [98,199] and can be classified into three types: (1) due to input, i.e. measurement errors related to weather variables; (2) due to intrinsic variability/stochasticity of the fundamental physical processes; and (3) inherent in the model's structure, i.e. parameters selected and training datasets utilized for model building purposes.

For example, during real-time data collection, the authors [98] identified that the variability in individual model's prediction in their ensemble of models approach to PVPF was significant at mid-day because of high uncertainty in weather conditions at the ASU PV plant site. This variability was much less in early mornings and late evenings on account of small variability. This is an instance of the second kind of uncertainty source (i.e. intrinsic variability/stochasticity of the fundamental physical processes).

5.3.2.1b. Uncertainty quantification techniques. Although uncertainty quantification is vital to energy and utilities market stakeholders [192,200], there is scarcity of research in this arena; most PVPF research focusing instead on accuracy of prediction.

In Ref. [98], two established performance metrics for uncertainty quantification were cited: prediction interval coverage probability (*PICP*) and prediction interval Width (*PIW*) [201]. Furthermore, the authors [98] embedded the Bootstrap technique (BS) in their model for quantifying uncertainty, which has proven performance in prediction interval (*PI)* estimation for industrial sector [199,201]. Also, BS is easy to integrate with ensemble of ANNs' architecture. Furthermore, the simplicity, lower computational efforts and ease of interpretation makes BS a very suitable technique for uncertainty quantification, especially in neural network based approaches.

Uncertainty prediction can be enhanced by defining wider PIs which achieve the highest confidence level compared to established approaches [98]. The authors included all three categories of uncertainty measure for high accuracy:

1. Errors in model input
2. Unpredictability inherent from stochasticity of physical processes
3. ANN base model error

Uncertainty quantification involves defining lower and upper limits or prediction intervals (PIs) of PV output within which the true "a priori unknown" output power value is expected to fall with a predefined confidence level of α%. Usually, α value is near 90% [201,202]. In Ref. [98], time stamp was used to calculate the uncertainty as $\widehat{P}ensemble$ (t, d), at time t of day d.

Other well-established uncertainty quantification techniques include: Percentile (10th and 90th percentiles for PI's lower and upper bounds, respectively and $\alpha = 80\%$) [203]; Non-parametric Kernel Density Estimation (KDE) (Probability Density Function (PDF) of the power predictions estimated the base prediction models of the ensemble, i.e. summation of the Gaussian kernel functions assigned to each of the predictions to obtain the final PDF, whose 10th and 90th percentiles are the lower and upper bounds of the *PI*, respectively) [204]; and Mean Variance Estimation (MVE) (uncertainty distribution following Gaussian distribution function whose variance is determined by employing ANN designed following a process similar to that followed in the BS) [205]. In Ref. [98], uncertainty quantified via BS was found superior to the alternatives: KDE and MVE, for obtaining the target coverage level of 0.8 at all-time instances of the day.

Yet other uncertainty estimations exist, such as: Delta [206,207], Lower Upper Bound Estimation (LUBE) [208] etc. which have demonstrated utility under varied industrial applications [201,207,209].

5.3.2.2. Deep learning. Currently, the state-of-the-art in machine learning is deep learning (DL) or deep neural network (DNN). This ANN can learn from voluminous input data and uses improved learning algorithm, better parameter analysing methods and numerous hidden layers [210]. DL is an unsupervised machine learning technique implying that its algorithm can predict the outcome by perceiving the pattern in the input. First proposed by Hinton in 2006 as 'layer-wise greedy learning' [211], DL has enhanced ability to determine local optima and depict assemble rates. It beat the world's Go Game champion Lee Se-dol in Korea in 2016 under the Google deep learning project [212].

Some researchers [65] claimed that deep convolutional neural network (DCNN) consistently outperforms other models in forecasting from different data sets as demonstrated by its statistical evaluations. In terms of forecasting reliability, sharpness and overall continuous ranked probability score (CRPS) skill, forecast models developed from probabilistic combination with WT, DCNN and Quantile regression (QR) models are very effective.

There are mainly four types of deep learning, a restricted Boltzmann machine (RBM), deep belief network (DBN), autoencoder (AE) and deep convolutional neural network (DCNN).

RBM is generative stochastic in nature and uses input data to learn probability distribution. Paul Smolensky in 1986 developed the idea and later Hinton applied it. This method is employed in data classification, collaboration, filtering and dimension reduction [212]. Different hybrids of RBM have been developed: discriminative restricted Boltzmann machines (DRBMs) [213], conditional restricted Boltzmann machines (CRBMs) [214] and robust Boltzmann machine (RoBM) [215].

Further modification by Hinton led to stacked architecture of RBM known as DBN, a MLP with latent variables that can form the Bayesian probability generative model, but its training method is entirely different. In brief, DBN is a combination of RBMs and a classifier that can train hidden layers simultaneously whose attributes are the same as RBM's. Such arrangement ensures that when the first iteration is done its output acts as an input for the next RBM layer. This is a continuous process that progresses until it reaches the output layer [216]; thus, it can learn patterns progressively. When the learning process is complete, it recognizes the distinctive patterns in a data set. Hinton and Nair harnessed this pattern recognition capability of DBN to identify a 3D object [217]. It is still a developing field and research is ongoing on the development of image and speech recognition hybrid models [218–220].

Another DL approach is the autoencoder (AE). Bourlard and Kamp [221] used a MLPNN for information processing in auto-association mode for compressing data dimension. Although an unsupervised learning algorithm, AEs can learn generative models of data for the purposes of forecasting [222]. It achieves this by learning to encode the information from the input layer and later reconstructing it after the decoding process. The main advantage is that AEs can filter out noise during data extraction.

A special subset of DL algorithm, convolutional neural network (CNN) is a novel approach where the connectivity principles between synthetic neurons mimic the organization of animal visual cortex. Besides, it learns to recognize patterns usually through highlighting the edges and pixel behaviors that are generally observed in various images in its layers.

It is currently the best available tool for machine handwriting recognition, face detection, behaviour recognition, speech recognition, recommender systems and image classification. The computational model utilized is called Neocognitron [223] which relies on linear or nonlinear filters to extract the features from an image. To function effectively, CNN consists of several blocks such as convolution, activation and pooling, all functioning together for feature extraction and transformation [224]. It has a deep neural architecture and usually comprises four types of layers [225], namely:

1., Convolutional layer, which extracts feature representations of the inputs.
2., Pooling layer, which connects the previous convolutional layer. The aim of pooling layer is to aggregate the input features by reducing the resolution of feature maps.
3., Fully connected layer, which lies between pooling and logistic regression layer and transports the learned distributed feature representations to one space in order to perform high-level reasoning. All the neurons of previous layer are connected to every single neuron of the current layer.
4., Logistic regression layer is the last layer in CNN. Softmax function is usually employed as the activation function of logistic regression layer for output generation.

Compared with the traditional fully-connected neural networks, the obvious advantage of CNN is the reduced number of parameters to be estimated due to the weight sharing technique. As well, CNN layers consisting of small kernel sizes provide an efficient way for extracting hidden structures and inherent features.

Along with structural performance advantages, CNNs have other practical functionalities, such as excellent data processing with grid topology features [226]. Consequently, CNN is commonly employed in image processing; however, its application in weather classification for PVPF is less common [53]. In this type of approach, a weather classification problem, which is predominantly a pattern recognition task, is transformed into an image processing undertaking; thereby improving the classification accuracy. 1-D solar irradiance and other time series arrays are first converted into a 2-D image for feature extraction, and then reconverted to a 1-D vector for classification purposes.

For even more enhanced performance, hybrid CNNs are being developed, such as recursive convolutional networks (RCNs) [227]. Others, Desjardins and Bengio [228] incorporated RBMs in CNN to develop the convolutional restricted Boltzmann machines (CRrBMs). Jarrett, et al. [229] proposed another novel hybrid model that merges convolution with an AE, and Mathieu et al. introduced the Fourier Transform (FT) on CNNs for fast training procedure [230].

For forecasting the behaviour of atmospheric elements, such as wind intensity, Hu, et al. [63] employed DL with a de-noising AE to pre-train the NN from old wind farm data in order to accurately predict wind power generation for newly developed wind farms. Another researcher, Qureshi, propounded the application of DL for short-term wind power prediction [64]. The forecasting strategy utilized deep belief network for meta-regression and an AE for base regression. It has been aptly named DNN-MRT; another very effective hybrid model. Finally, for PVPF purposes, a variety of approaches have been investigated. Yang, et al. [52] used SOM and LVQ to classify historical PV output data. Chen, et al. [40] adopted SOM and used it to classify local weather type for next-day forecasts from online time series data. Wang, et al. [56] extracted solar irradiance features using SVM and performed pattern recognition of weather statuses for ST-PVPF. Furthermore, SVM and KNN approaches were used for adequate sample size determination. However, all these models suffered from a common issue; shallow learning models. These are unable to determine the deep non-linear relationships between input and output, especially when large quantities of complex data are involved. These can be addressed using CNN which can elucidate the intrinsic abstract features and high-level invariant structures in data [231]. Indeed, previous studies agree that CNN classifications are best achievers in various applications, outperforming classical intelligence methods such as SVM [232]. Table 3 lists ANN based and Table 4 catalogs some DNN based prevalent PV forecasting techniques.

The techniques used in the optimization and evaluation of PV output forecast models are varied and situation specific. The disparate input parameters (location and climate dependent) make comparison among the models complicated; requiring specialized methods to gain useful insights.

The time series data either from primary (experimental observations) or secondary (published data from different meteorological stations) sources frequently contain periodic oscillations and spikes. Such outliers can severely degrade the quality of data and are cleansed using WT method. Another common problem is missing data in the time series due to sensor error and faulty data management system. Such missing data have to be reconstructed before any modelling using similarity technique or interpolation. Hence, along with different models, any comparison has to consider the pre-processing methods utilized to improve data quality.

The data size or length is another important factor and the current literature review performed highlights the need for further investigations into the effect of data length employed in forecasting [211]. Too short or too long data lengths adversely affect a model's precision. In general, the forecast is better with longer and quality training data. Thus, some researchers have used different sized data sets, determined using clearance index, indicating sunny or cloud covered days, to analyse the effect of various data lengths on forecast accuracy. This could be a method for comparing across different data sets, the performance of varied forecast models.

Sufficient length of time series data, which contains the salient features of the forecasted data set and the weather conditions of a particular region, is very essential. A mixture of sunny and cloudy days is more common in winter months than in the summer. Thus, models trained on winter time series data are usually more robust than those trained on summer data for most locations. However, the literature also indicates that longer training data sets are not necessarily comprehensive, i.e. they may not necessarily contain more training samples. Rather, repetition of data and patterns can abound in such long data. Therefore, training models on such data is computationally expensive without any added benefit. This review thus evaluated models based on the use of techniques to ensure quality training data with sufficient sample size.

Another important issue is data resolution; too high data resolution, as in data taken every few minutes usually has inherent high uncertainties which can cause problem in pattern recognition; an important aspect in determining the complex relationships between input and output. Conversely, too low resolution, as in data taken every half hour to an hour has inadequate training patterns and can cause problems in model's convergence, repetition of patterns in data and failure to model the relationship between input and output parameters. The optimum solution is the use of intermediate temporal resolution data sets for effective training, testing and prediction purposes. Therefore, in the evaluation of any model's forecasting ability, the type of data resolution used was assessed.

A prescribed solution to these problems of data length and resolution has been the use of optimization techniques. Many researchers have used statistical or evolutionary algorithm (GA, PSO etc.) based approaches to select optimum training data sets. In the case of neural networks, suitable neural layer architecture and the number of hidden layers have also been optimized. In the current review, therefore, emphasis was given to modelling approaches with optimization.

Yet others have used data normalization techniques to reduce the domain and range of the data for ease of comparison and modelling, usually bringing all data within a range of 0–1. A large data spread in one of the meteorological factors being considered as a model input can cause too much weightage being assigned to it and also improper output of the activation functions used in NN. Normalization during pre-

Table 3
Forecasting techniques using Artificial Neural Network.

Author	Year	Forecasting Model	Description	Accuracy	Input Selection	Input Correlation Analysis	Data pre-processing	Parameter Selection	Forecast Horizon
Qing and Niu [233]	2018	LSTM	A LSTM model developed for day ahead prediction of solar irradiance from weather data. Its performance is better than persistence, LLSR and multilayered FFNN-BPNN models.	RMSE 18.34%	Temperature, Dew Point, Humidity, Visibility, Wind speed, Weather type	The hourly solar irradiance is positively and strongly correlated with temperature and negatively correlated with humidity. Other weather variables have low correlation values	Normalization	Weights and biases are adjusted by using their gradients.	hourly
Sivaneasan, et al. [234]	2017	ANN and fuzzy logic	An ANN was coupled with triple layer BP and fuzzy pre-processing approach for solar power forecasting. Its performance was found to be superior to pure ANN, coupled ANN-fuzzy and ANN-fuzzy with error correction factor.	MAPE 29.6%	Temperature, Dew point, Wind speed, Gust wind, Solar irradiance	The outputs solar power generation and forecast error changes with solar irradiation or temperature variation	A fuzzy pre-processing toolbox	8 inputs in the input layer, 1 output in the output layer and 25 hidden neurons are used. A tangent sigmoid is used as the activation function for all the neurons in the neural network	monthly
Sharma, et al. [235]	2016	WNN	A feed-forward ANN with WT model was developed with Morlet and Mexican hat type activation functions. Its predictions were compared with persistence, ETS, ARIMA and ANN for temporal horizons of 15 min and 24 h; and found to be superior.	nRMSE (%) 9.42 to 15.41	Solar irradiance	PV output found to be positively correlated with solar irradiance	wavelet transform	5 input neurons and 11 Morlet wavelet based hidden neurons and 1 output is used	hourly
Gutierrez-Corea, et al. [236]	2016	ANN	Different architectures and parameters of ANN based on MLPNN with BP algorithm were investigated to forecast global solar irradiance. The forecast horizon was 1 to 6-h and data from a neighbouring weather station was used as input.	nRMSE (%) below 20	Global solar irradiance (GSI), Atmospheric temperature (AT), Relative air humidity (RAH), Wind direction (WD), Wind speed (WS), Horizontal extraterrestrial solar irradiance (HESI), the Clearness index (K_T), Distance from solar noon (DSN), Zenith angle (ZA)	K_T obtained based on GSI and HESI; similarly, GSI and HESI will have a greater value when DSN is close to zero and, all these variables are correlated with the Zenith angle	Spatial DataBases, Standardized	100 and 10 neurons are used in the first and second hidden layers, respectively	Between 1 and 6 h
Pedro and Coimbra [237]	2015	ANNs and KNN	ANN was combined with k-nearest-neighbours (kNN) optimization to forecast global solar irradiance for time horizons from 15 min to 2 h. The outcome was better than other simpler forecasting schemes.	RMSE(%) <15.00	GSI and Clear-sky, Cloud pattern	Cloudy conditions are highly correlated with GSI	Normalization	10 neurons are used in single hidden layer with hyperbolic tangent sigmoid transfer function and single output layer with linear transfer function	15 min to 2 h
Chu, et al. [238]	2015	ANN with GA Optimization	A hybrid ANN with GA optimization forecasting scheme was developed for real-time output prediction of a 48 MWe PV plant. The results were superior to those of deterministic physical model based on cloud tracking technique, ARMA and kNN models.	MAE 21.02%	Cloud position, Clear sky index, Solar output power, solar irradiance	Cloud position is positively correlated with irradiance	pseudo-Cartesian transform	single hidden layer with 7–10 neurons	5, 10 and 15 min
Ramsami and Oree [239]	2015	GRNN, FFNN and MLR	A hybrid approach for day-ahead power output forecasting of a PV installation was developed by integrating GRNN with coupled FFNN and MLR. Stepwise regression was employed for selecting the most strongly correlated meteorological parameters as inputs for this model.	RMSE 2.74%	Atmospheric pressure (a), Humidity (h), Temperature (t), Wind direction (d), Wind speed (w), Rainfall (r), In-plane solar irradiance (s) and Sunshine hours (sh).	Temperature, Wind direction, Wind speed, Rainfall, Solar irradiance are positively correlated to average energy generated by the PV system	Normalization	FFNN has 8 input neurons, 12 hidden neurons and 1 output neuron. SR-FFNN has 2 input neurons, 10 hidden neurons and 1 output neuron.	24-h

(*continued on next page*)

Table 3 (*continued*)

Author	Year	Forecasting Model	Description	Accuracy	Input Selection	Input Correlation Analysis	Data pre-processing	Parameter Selection	Forecast Horizon
Liu, et al. [240]	2015	BP and ANN	A BP with ANN model was utilized to forecast day-ahead photovoltaic output. Aerosol Index was an input, as the researchers claimed it was strongly (linearly) correlated with the attenuation of incident solar radiation. The model demonstrated forecasting accuracy surpassing conventional ANN.	MAE 11%	Aerosol index (AI), Temperature, Humidity, and Wind speed data	Aerosol index (AI) has strong linear correlation with solar radiation attenuation, and can potentially influence the power generated by PV panels	Normalization	number of input layer neurons 20, hidden layer neurons 9, output layer neurons 12	24-h
Raza, et al. [241]	2018	NNE	The NNE model comprised FNNs, trained via PSO, for day-ahead PV output predictions in smart grids. Varying resolution, length and quality data were used for model training. Model performance was measured against five forecasting techniques and results indicated that the NNE method was superior.	median MAPE 9.75%	Historical power output of PV, Solar irradiance, Wind speed, Temperature and Humidity	Solar irradiance and temperature largely affect PV power output.	Wavelet Transformation	1st and 2nd structures have 10 and 15 hidden layer neurons, respectively and so on. The last structure contains 60 number of neurons.	24 h
Lu and Chang [242]	2018	RBFNN	A hybrid RBFNN model with data regularity scheme was developed for photovoltaic output predictions. Benchmarking was done against ARIMA, BPNN and RBFNN for day ahead forecasting and the hybrid performed best.	MAPE (%) 3.71 RMSE (%) 4.65	Consecutive two days of measured PV power generation data in time series form				24 h
Jamali, et al. [243]	2019	ANN and PSO-GA	They optimized an ANN forecast model using PSO-GA to predict performance of a SSHS. Benchmarking was done against HEPSO and TGA showing that their PSO-GA-ANN hybrid was more accurate and reliable.	RMSE <24%	Ambient temperature, Solar radiation, Outlet oil temperature of the PTC, Water temperature of the storage tank	Highly positively correlated with solar radiation	Normalization	10 neurons in the hidden layer	
İzgi, et al. [244]	2012	Artificial Neural Network – Levenberg Marquardt Back-Propagation algorithm	Levenberg-Marquardt Back-propagation algorithms were exploited to optimize the network parameters (weights) of an ANN. The coupled approach produced the best mapping among the inputs and the response	RMSE = 19.9515 to 54.1102	750 Wp solar PV panel Ambient and cell temperature Diffuse solar irradiation				0–45 min
Notton et al. Notton, et al. [245]	2013	Artificial Neural Network – Multi Layer Perceptron architecture with Feed-Forward Back-Propagation	The technique, involving ANN, helped extract reliable input data for optimization and modelling of both types of solar energy harnessing systems, i.e. PV and CSP. Sensitivity analysis aided in minimizing the input data volume without affecting the optimum network architecture. The approach is simple and computationally economical	RMSE is around 9%	Zenith angle −10-min of extra-terrestrial horizontal irradiance and horizontal global irradiation Inclination angle			6, 12, 24 hidden neurons	10 min
Dahmani et al. Dahmani, et al. [246]	2014	ANN-MLP	The method helped to address missing tilted solar radiation data by analysing horizontal solar radiation inputs. When the ANN is sufficiently trained, this approach	RMSE 8.81%	Horizontal extra-terrestrial irradiance, Solar declination, Zenith angle, Azimuth angle, Horizontal irradiation			1 to 8 hidden neurons	5 min

(*continued on next page*)

Table 3 (*continued*)

Author	Year	Forecasting Model	Description	Accuracy	Input Selection	Input Correlation Analysis	Data pre-processing	Parameter Selection	Forecast Horizon
			can obviate the need of an inclined pyranometer						
Marquez et al. Marquez, et al. [247]	2013	Satellite imagery analysed via ANN	Stochastic learning, Ground and/or Remote sensing image processing, and ground telemetry data were combined to develop a robust and accurate solar radiation forecast model	RMSE 5–25%	Velocimetry, Cloud indexing (from satellite imagery), Solar irradiance		Normalization	5 hidden neurons	30, 60, 90 and 120 min
Paoli et al. Paoli, et al. [248]	2011	ANN	Pre-processed data fed into an optimized ANN model to generate high accuracy predictions	nRMSE < 2%	Global daily solar irradiation data	Positive correlation with global radiation data	Normalization	1–3 hidden layers, and 1–50 neurons	Seasonal forecast 4 months ahead
Huang et alHuang, et al. [249]	2010	Neural Network and statistical method	two power forecasting methods for PV systems, physical method and statistical method, are studied. A physical model based on the construction of PV systems and a NN statistical model based on historical data are set up	nRMSE 10% and 13%	Solar irradiance, Air temperature, Cloud, Humidity and Position of the sun			11-15 hidden neurons	One week

processing of input data can help to reduce these problems and make comparison among different sized data sets possible; nevertheless, the normalized data need to be de-normalized in post-processing in order to make sense of the forecasts, an added complication. Other approaches have applied equal weight aggregation methods, which average the final forecast of various neural networks, and trimming aggregation techniques, to excise extreme data points. The same techniques can also be used to compare the performance of various models developed from different data sets; however, there are no exact methods in determining the values of trimming parameters making the process less amenable to objective inquiry.

Another key issue has been the use of appropriate evaluation techniques for model performance. A common approach is benchmarking against established models and data. This has been exploited by numerous researchers with the persistent model being a favourite. Others have compared their results with meteorological and physical models used in weather forecasting, and some have developed experimental setups to make direct objective comparison of their developed models. A more proven model evaluation approach is the application of statistical measures. Many statistical tools are available, among these, RMSE is very robust and effective due to its ability to handle outliers and spikes in data.

Thus, comparing the performance of different forecast models is a complicated affair; nonetheless, this paper has highlighted the major concerns in the various models discussed and has taken an objective approach. It is hoped that the contrasts made among the different models, although based on different data sets and criteria, will help give future researchers a starting point in deciding which type of approach to use based on their specific requirements.

6. Motivation for the review

The uncertainty inherent in harnessing solar energy due to its dependence on meteorological parameters makes PVPF essential for efficient planning and integration of PVs in micro and national grids. From the comprehensive and critical review above it is lucid that over the past two decades many different approaches to the modelling and forecasting of PV output have been attempted with varying degrees of success. Nevertheless, the conventional and statistical methods in forecasting have not been up to the par in terms of reliability, accuracy and computational economy. The obvious choice is the machine learning approach in the form of ANN or its hybrids, which can handle large amounts of data and generate accurate predictions for short to medium temporal horizons, without the need for complex mathematical relationships or physical representations. Thus, the demand on the researcher's and the practitioner's skills and effort is reduced; making the ANN approach easily reproducible for different sets of initial scenarios and locations.

Not surprisingly, researchers have exploited the flexibility and robustness of ANN based forecasting and the literature abounds with descriptions of various modelling approaches. Among these, DNN is promising for its ability to fathom non-linear and highly complex relationships between numerous inputs and the forecasted response. More specialized forms such as RNN and CNN are fast gaining popularity. These approaches have already proven their versatility in other disparate applications.

Even more sophisticated approaches are recently available such as DCNN and deep long short term memory (DLSTM); the latter is a state-of-the-art technique which exploits the benefits of memory loops integrated in a DNN, allowing the neural net to learn from data sets at a faster pace. Thus, it is hoped that the comprehensive review will aid future researchers as well as utilities operators to gain valuable insight into the need and the modes of forecasting for PV power output. The knowledge gained may also help policy makers and energy market participants to make more effective and profitable decisions concerning the implementation of hybrid PV integrated electric grids for local or

Table 4
Forecasting techniques using Deep Neural Network.

Author	Year	Forecasting Model	Description	
Wen, et al. [250]	2019	DNN and LSTM	A coupled DNN-LSTM approach was used to predict the load and PV output in a micro-grid. For better prediction, the inputs were optimized with PSO. The hybrid was better than MLPNN and SVM in terms of total forecasting cost and reliability.	MAPE 7.43%
Siddiqui, et al. [251]	2019	DCNN and LSTM	A twin stage DNN model forecasted solar-irradiance from cloud cover videos. The first stage was a DCNN which encoded the sky images. Then, LSTM as the second stage used metrological parameters for predictions. The model outperformed GFS and ECMWF.	nMAP<22%
Lee, et al. [252]	2018	DCNN and LSTM	Twin DNN model was employed to generate day-ahead solar power forecast from time series data obtained from PV inverters and national weather centre reports. The two DCNN with different filter sizes extracted short-time local patterns and a LSTM captured the long-time features in the data. The approach outperformed linear regression, RFR, SVR and other traditional models.	RMSE 9.87%
Zhang, et al. [253]	2018	DCNN	Several DCNNs were exploited to predict solar power generation. High-resolution weather data with various spatial and temporal connectivities were used to understand cloud movement and its correlation with solar energy utilization. The prediction accuracy was better than the persistent and SVR models.	rMAPE 11.8%
Haixiang, et al. [254]	2018	DCNN and Variational mode decomposition	A hybrid approach based on DCNN for short term PV output forecasting. A variational mode decomposition technique extracted different frequencies from historical data to form two-dimensional datasets that correlated both daily and hourly timescales through convolution kernels. The model outperformed SVR, GPR and RFR techniques.	RMSE< 4%
Srivastava and Lessmann [255]	2018	LSTM	A DNN married to LSTM predicted day-ahead global horizontal irradiance from satellite data. The model outperformed Gradient Boosting Regression, FFNN and Persistence methods, with an average forecast skill of 52.2% over the persistence model.	RMSE <29.26 Wm^{-2}
Zhang, et al. [31]	2018	LSTM	Three DNNs were analysed: deep MLPNN, DCNN and deep LSTM in terms of forecasting of PV output on minute scale. The LSTM based model demonstrated the best performance.	RMSE< 21%
Wang, et al. [65]	2017	DCNN, WT and QR	A hybrid approach merging WT, DCNN and QR was used to enhance the prediction capability over BPNN, SVM and SVM integrated WT methods.	RMSE<15%
Alzahrani, et al. [66]	2017	DRNN	DRNN was employed for forecasting solar irradiance. Inputs were cloud cover, scattering of sunlight, overcast sky and clear-sky datasets. The model outperformed SVR and FFNN approaches.	RMSE 8.6%
Gensler, et al. [256]	2016	AE and LSTM	A hybrid DL algorithm comprising AE and LSTM was developed which outperformed MLPNN and physical forecasting models. Comparison also showed its superiority over DBF and deep LSTM.	RMSE <7%

national consumption.

7. Conclusion

The dynamic nature of solar irradiance affects the reliability of PV incorporated systems. Sporadic penetrations due to sunlight intensity variations lead to voltage and power fluctuations; disrupting utility companies, energy markets and power distribution. Reliable forecasting is the solution and can be classified from ultra-short term to long term forecasts; with current research and industry focus mostly on transcend systems for short-term next-day predictions as these are more accurate and can address the vagaries of cloud cover. On the other side, long term forecast have gained importance for long duration power system planning purposes.

However, precise predictions are challenging and disparate approaches involving physical, statistical, artificial intelligence, ensemble and hybrid models have been investigated. Among these ANN, especially CNN or its hybrid forms, often in ensemble arrangement, hold most promise for forecast accuracy, especially for short-term forecast horizons.

Moreover, forecast horizons, model's performance estimation, input selection and its correlation analysis with optimization, data pre and post-processing, time stamp and weather classification, network parameter optimization, adaptive neural architecture, and uncertainty quantification must be considered for designing an effective PVPF. Forecast horizons impact a model's prediction accuracy. In general, the longer the forecast horizon, the greater is the chance of forecast error. In contrast, ultra-short-term forecasting suffers most due to the vagaries of nature, i.e. cloud cover and its turbulent movements. Alongside a model's performance estimations by employing statistical measures: MAE, MAPE or RMSE; economic analysis metric may be incorporated for more practical assessment of PVPF models. There is no single performance measure for all situations, rather the researcher or practitioner must decide on the best evaluation method being cognisant of the strengths and weaknesses of each.

Also, model inputs must be amenable to correlational analysis, with strongly correlated inputs being most important. Improper inputs cause high forecast errors leading to time delays, cost overruns and computational complexities. Common inputs include meteorological data, for instance: solar irradiance, atmospheric temperature, wind speed and direction, and humidity. Furthermore, PV module (cell) temperature is another, albeit system related, input. Among these, solar irradiance is most positively correlated with PV output.

Nevertheless, including many inputs considerably increases the complexity and computational time of a model. Hence, optimizing for the most strongly correlated inputs is essential for efficient PVPF and is usually done using evolutionary algorithms.

Also of concern is pre and post-processing of data. Cleansing and restructuring of historical datasets is mandatory as is its augmentation, since they contain intermittent static, spikes or missing instances precipitated by weather fluctuations, seasonal variations, electricity demand shifts or sensor failures. WTs and normalization are popular, while GAN is gaining ground for data augmentation purposes.

Another pre-processing technique, weather classification, has been recommended by scholars. It accounts for change of solar irradiance due to change in weather and 33 major types have been categorized, with recommendations for merging these into 10 classes being presently under consideration. This is also inseparable from cloud motion and categorization studies, which exploits image processing based pattern recognition. Such techniques are generally employed with time stamp of day and month of the year.

For optimization of network parameters (weights and biases), GA and PSO have gained prominence. In addition, design of adaptive network features to handle newly available data from same or different PV plants for online PVPF applications has been solved using OS-ELM, which is more efficient than the state-of-the-art BP-ANN. In this regard, ensemble of base prediction models comprising many ANNs or ELMs, have also shown potential.

Subsequently, post-processing to interpret the predictions and to make it adaptive to newly available data for online applications is vital. Both primary (from experiments) and secondary (from NWP) data have been used in this regard. Often researchers have used a single PV plant, while some have used aggregate readings from distributed PV systems spanning large areas.

Finally, all the various components of a PVPF model must have synergy in order to implement the predictions and derive benefits.

8. Scope of future work

The vagaries of weather give the impression that RE based, especially PV integrated, hybrid electric grids are difficult to realize. However, advances in mathematical modelling, physical representations, statistical analysis and computing power have made forecasting a viable option. Many agencies have shown initiative to support research in this direction and, as a response, considerable research findings are available. Still, the existent forecast models are either too specific to a spatial-temporal horizon or are circumscribed for a particular region. The current need, therefore, is for a more versatile forecasting approach which is unbounded by such limitations and can be reproduced for varying initial conditions at different geo-climatic conditions. The obvious contender for this is ANN, more specifically the DNN approach, along with its derivatives such as DLSTM and DCNN. More research is needed to investigate the utilization of these techniques in solving the conundrum of PVPF.

References

[1] Guo ZF, Zhou KL, Zhang C, Lu XH, Chen W, Yang SL. Residential electricity consumption behavior: influencing factors, related theories and intervention strategies. Renew Sustain Energy Rev Jan 2018;81:399–412.

[2] Mocanu E, Nguyen PH, Gibescu M, Kling WL. Deep learning for estimating building energy consumption. Sustain Energy Grids; Networks Jun 2016;6:91–9.

[3] Ssen Z. Solar energy in progress and future research trends. Prog Energy Combust Sci 2004;30(4):367–416.

[4] Varotsos CA, Efstathiou MN, Christodoulakis J. Abrupt changes in global tropospheric temperature. Atmos Res Mar 2019;217:114–9.

[5] Zervos A. Renewables 2016 global status report, vol. 60; 2016.

[6] Arthouros Zervos CL, Muth Josche. RE-thinking 2050-A 100% renewable energy vision for the European union. 2010.

[7] Elhadidy M, Shaahid SM. Parametric study of hybrid (wind plus solar plus diesel) power generating systems. Renew Energy Oct 2000;21(2):129–39 (in English).

[8] Khare V, Nema S, Baredar P. Solar–wind hybrid renewable energy system: a review. Renew Sustain Energy Rev 2016;58:23–33.

[9] Shah A, Yokoyama H, Kakimoto N. High-precision forecasting model of solar irradiance based on grid point value data analysis for an efficient photovoltaic system. Ieee Trans Sustain Energy Apr 2015;6(2):474–81 (in English).

[10] World energy resources. World Energy Council; 2016.

[11] Cojocaru EG, Bravo JM, Vasallo MJ, Santos DM. Optimal scheduling in concentrating solar power plants oriented to low generation cycling. Renew Energy 2019/05/01/ 2019;135:789–99.

[12] Comello S, Reichelstein S, Sahoo A. The road ahead for solar PV power (in English). Renewable & sustainable energy reviews. vol. 92; Sep 2018. p. 744–56. Review.

[13] Das UK, et al. Forecasting of photovoltaic power generation and model optimization: a review (in English) Renew Sustain Energy Rev Jan 2018;81: 912–28. Review.

[14] IEA., May. Tracking power. 2019. Available: https://www.iea.org/reports/tracking-power-2019.

[15] Hoeven Mv d. Technology roadmap solar photovoltaic energy. 2014. France.

[16] Raza MQ, Nadarajah M, Ekanayake C. On recent advances in PV output power forecast. Sol Energy 2016;136:125–44.

[17] Shivashankar S, Mekhilef S, Mokhlis H, Karimi M. Mitigating methods of power fluctuation of photovoltaic (PV) sources - a review (in English) Renew Sustain Energy Rev Jun 2016;59:1170–84. Review.

[18] Diagne M, David M, Lauret P, Boland J, Schmutz N. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids (in English) Renew Sustain Energy Rev Nov 2013;27:65–76. Review.

[19] Voyant C, et al. Machine learning methods for solar radiation forecasting: a review (in English) Renew Energy May 2017;105:569–82. Review.

[20] Raza MQ, Khosravi A. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. Renew Sustain Energy Rev 2015;50:1352–72.

[21] de Marcos RA, Bello A, Reneses J. Electricity price forecasting in the short term hybridising fundamental and econometric modelling. Elec Power Syst Res 2019; 167:240–51.

[22] Amral N, Ozveren CS, King D. Short term load forecasting using Multiple Linear Regression. In: 42nd international universities power engineering conference; 2007. p. 1192–8. 2007.

[23] Nespoli A, et al. Day-ahead photovoltaic forecasting: a comparison of the most effective techniques. Energies 2019;12(9):1621.

[24] Sobri S, Koohi-Kamali S, Rahim NA. Solar photovoltaic generation forecasting methods: a review. Energy Convers Manag 2018;156:459–97.

[25] Ren Y, Suganthan PN, Srikanth N. Ensemble methods for wind and solar power forecasting—a state-of-the-art review. Renew Sustain Energy Rev 2015;50:82–91.

[26] Behera MK, Majumder I, Nayak N. Solar photovoltaic power forecasting using optimized modified extreme learning machine technique. Eng Sci Technol, Int J 2018;21(3):428–38.

[27] García-Martos C, Rodríguez J, Sánchez MJ. Forecasting electricity prices and their volatilities using Unobserved Components. Energy Econ 2011;33(6):1227–39.

[28] Torbaghan SS, Motamedi A, Zareipour H, Tuan LA. Medium-term electricity price forecasting. In: 2012 North American power symposium (NAPS); 2012. p. 1–8.

[29] Vehviläinen I, Pyykkönen T. Stochastic factor model for electricity spot price—the case of the Nordic market. Energy Econ 2005;27(2):351–67.

[30] Hong T, Wilson J, Xie J. Long term probabilistic load forecasting and normalization with hourly information. IEEE Trans Smart Grid 2014;5(1): 456–62.

[31] Zhang J, Verschae R, Nobuhara S, Lalonde J-F. Deep photovoltaic nowcasting. Sol Energy 2018;176:267–76.

[32] Antonanzas J, Osorio N, Escobar R, Urraca R, Martinez-de-Pison FJ, Antonanzas-Torres F. Review of photovoltaic power forecasting. Sol Energy 2016;136: 78–111.

[33] Pedro HTC, Coimbra CFM. Assessment of forecasting techniques for solar power production with no exogenous inputs. Sol Energy 2012;86(7):2017–28.

[34] Lonij VPA, Brooks AE, Cronin AD, Leuthold M, Koch K. Intra-hour forecasts of solar power production using measurements from a network of irradiance sensors. Sol Energy 2013;97:58–66.

[35] Lorenz E, Kuehnert J, Wolff B, Hammer A, Kramer O, Heinemann D. PV power predictions on different spatial and temporal scales integrating PV measurements, satellite data and numerical weather predictions. 2014.

[36] Li Z, Rahman SMM, Vega R, Dong B. A hierarchical approach using machine learning methods in solar photovoltaic energy production forecasting. Energies 2016;9(1):55.

[37] Almonacid F, Pérez-Higueras PJ, Fernández EF, Hontoria L. A methodology based on dynamic artificial neural network for short-term forecasting of the power output of a PV generator. Energy Convers Manag 2014;85:389–98.

[38] AlHakeem D, Mandal P, Haque AU, Yona A, Senjyu T, Tseng T. A new strategy to quantify uncertainties of wavelet-GRNN-PSO based solar PV power forecasts using bootstrap confidence intervals. In: 2015 IEEE power & energy society general meeting; 2015. p. 1–5.

[39] Zhang J, et al. A suite of metrics for assessing the performance of solar power forecasting. Sol Energy 2015;111:157–75.

[40] Chen C, Duan S, Cai T, Liu B. Online 24-h solar power forecasting based on weather type classification using artificial neural network. Sol Energy 2011;85 (11):2856–70.

[41] Lu S, et al. Machine learning based multi-physical-model blending for enhancing renewable energy forecast - improvement via situation dependent error correction. In: 2015 European control conference (ECC); 2015. p. 283–90.

[42] Zamo M, Mestre O, Arbogast P, Pannekoucke O. A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, part I: deterministic forecast of hourly production. Sol Energy 2014; 105:792–803.

[43] Lin K-P, Pai P-F. Solar power output forecasting using evolutionary seasonal decomposition least-square support vector regression. J Clean Prod 2016;134: 456–62.

[44] De Giorgi MG, Congedo PM, Malvoni M, Laforgia D. Error analysis of hybrid photovoltaic power forecasting models: a case study of mediterranean climate. Energy Convers Manag 2015;100:117–30.

[45] Vaz AGR, Elsinga B, van Sark WGJHM, Brito MC. An artificial neural network to assess the impact of neighbouring photovoltaic systems in power forecasting in Utrecht, The Netherlands. Renew Energy 2016;85:631–41.

[46] Lipperheide M, Bosch JL, Kleissl J. Embedded nowcasting method using cloud speed persistence for a photovoltaic power plant. Sol Energy 2015;112:232–8.

[47] Monjoly S, André M, Calif R, Soubdhan T. Forecast horizon and solar variability influences on the performances of multiscale hybrid forecast model. Energies 2019;12(12):2264.

[48] Shi J, Lee WJ, Liu Y, Yang Y, Wang P. Forecasting power output of photovoltaic systems based on weather classification and support vector machines. IEEE Trans Ind Appl 2012;48(3):1064–9.
[49] Mellit A, Massi Pavan A, Lughi V. Short-term forecasting of power production in a large-scale photovoltaic plant. Sol Energy 2014;105:401–13.
[50] Almeida MP, Perpiñán O, Narvarte L. PV power forecast using a nonparametric PV model. Sol Energy 2015;115:354–68.
[51] Zhen Z, et al. Pattern classification and PSO optimal weights based sky images cloud motion speed calculation method for solar PV power forecasting. IEEE Trans Ind Appl 2019;55(4):3331–42.
[52] Yang H, Huang C, Huang Y, Pai Y. A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output. IEEE Trans Sustain Energy 2014;5 (3):917–26.
[53] Wang F, et al. Generative adversarial networks and convolutional neural networks based weather classification model for day ahead short-term photovoltaic power forecasting. Energy Convers Manag 2019;181:443–62.
[54] Wang F, et al. Image phase shift invariance based cloud motion displacement vector calculation method for ultra-short-term solar PV power forecasting. Energy Convers Manag 2018;157:123–35.
[55] Kühnert J, Lorenz E, Heinemann D, Kleissl J. Chapter 11 - satellite-based irradiance and power forecasting for the German energy market. In: Solar energy forecasting and resource AssessmentBoston. Academic Press; 2013. p. 267–97.
[56] Wang F, Zhen Z, Mi Z, Sun H, Su S, Yang G. Solar irradiance feature extraction and support vector machines based weather status pattern recognition model for short-term photovoltaic power forecasting. Energy Build 2015;86:427–38.
[57] Wang F, Zhen Z, Wang B, Mi Z. Comparative study on KNN and SVM based weather classification models for day ahead short term solar PV power forecasting. Appl Sci 2017;8(1):28.
[58] Ishii T, Otani K, Takashima T, Xue Y. Solar spectral influence on the performance of photovoltaic (PV) modules under fine weather and cloudy weather conditions. Prog Photovoltaics Res Appl 2013;21(4):481–9.
[59] Lima FJL, Martins FR, Pereira EB, Lorenz E, Heinemann D. Forecast for surface solar irradiance at the Brazilian Northeastern region using NWP model and artificial neural networks. Renew Energy 2016;87:807–18.
[60] Engerer NA. Minute resolution estimates of the diffuse fraction of global irradiance for southeastern Australia. Sol Energy 2015;116:215–37.
[61] Wang F, Zhen Z, Liu C, Mi Z, Shafie-khah M, Catalão JPS. Time-section fusion pattern classification based day-ahead solar irradiance ensemble forecasting model using mutual iterative optimization. Energies 2018;11(1):184.
[62] Stehman SV. Selecting and interpreting measures of thematic classification accuracy. Rem Sens Environ 1997;62(1):77–89.
[63] Hu Q, Zhang R, Zhou Y. Transfer learning for short-term wind speed prediction with deep neural networks. Renew Energy 2016;85:83–95.
[64] Qureshi AS, Khan A, Zameer A, Usman A. Wind power prediction using deep neural network based meta regression and transfer learning. Appl Soft Comput 2017;58:742–55.
[65] Wang H, et al. Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network. Energy Convers Manag 2017;153: 409–22.
[66] Alzahrani A, Shamsi P, Dagli C, Ferdowsi M. Solar irradiance forecasting using deep neural networks. Procedia Comput Sci 2017;114:304–13.
[67] Ahmed A, Khalid M. A review on the selected applications of forecasting models in renewable power systems. Renew Sustain Energy Rev 2019;100:9–21.
[68] Das UK, et al. SVR-based model to forecast PV power generation under different weather conditions. Energies 2017;10:876. Art. no. 7.
[69] Rana M, Koprinska I, Agelidis VG. Univariate and multivariate methods for very short-term solar photovoltaic power forecasting. Energy Convers Manag 2016; 121:380–90.
[70] Kasten F, Czeplak G. Solar and terrestrial radiation dependent on the amount and type of cloud. Sol Energy 1980;24(2):177–89.
[71] Stefan N, Carol R. Solar spectral irradiance under clear and cloudy skies: measurements and a semiempirical model. J Appl Meteorol 1991;30(4):447–62.
[72] Kaskaoutis DG, Kambezidis HD, Jacovides CP, Steven MD. Modification of solar radiation components under different atmospheric conditions in the Greater Athens Area, Greece. J Atmos Sol Terr Phys 2006;68(10):1043–52.
[73] Sun Y, Wang F, Wang B, Chen Q, Engerer NA, Mi Z. Correlation feature selection and mutual information theory based quantitative research on meteorological impact factors of module temperature for solar photovoltaic systems. Energies 2017;10(1):7.
[74] Fonseca JGdS, Oozeki T, Takashima T, Koshimizu G, Uchida Y, Ogimoto K. Photovoltaic power production forecasts with support vector regression: a study on the forecast horizon. In: 2011 37th IEEE photovoltaic specialists conference; 2011. p. 2579–83.
[75] Monteiro C, et al. Short-term forecasting models for photovoltaic plants: analytical versus soft-computing techniques. 2013. In: Mathematical problems in engineering. vol. 9; 2013. Art. no. 767284.
[76] Alomari MH, Younis Ola, Hayajneh SMA. A predictive model for solar photovoltaic power using the levenberg-marquardt and bayesian regularization algorithms and real-time weather data. Int J Adv Comput Sci Appl 2018;9. Art. no. 1.
[77] De Giorgi MG, Congedo PM, Malvoni M. Photovoltaic power forecasting using statistical methods: impact of weather data. IET Sci Meas Technol 2014;8(3): 90–7.
[78] Mohammadi K, Goudarzi N. Study of inter-correlations of solar radiation, wind speed and precipitation under the influence of El Niño Southern Oscillation (ENSO) in California. Renew Energy 2018;120:190–200.
[79] Wang Z, Wang F, Su S. Solar irradiance short-term prediction model based on BP neural network. Energy Procedia 2011;12:488–94.
[80] S. Yujing et al., "Research on short-term module temperature prediction model based on BP neural network for photovoltaic power forecasting," in 2015 IEEE power & energy society general meeting, 2015, pp. 1-5.
[81] Mukherjee DP, Acton ST. Cloud tracking by scale space classification. IEEE Trans Geosci Rem Sens 2002;40(2):405–15.
[82] Rolf S, Martin R, Ehrhard P. An improvement of the IGMK model to derive total and diffuse solar radiation at the surface from satellite data. J Appl Meteorol 1990;29(7):586–603.
[83] Escrig H, et al. Cloud detection, classification and motion estimation using geostationary satellite imagery for cloud cover forecast. Energy 2013;55:853–9.
[84] Peng Z, Yoo S, Yu D, Huang D. Solar irradiance forecast system based on geostationary satellite. In: 2013 IEEE international conference on smart grid communications. SmartGridComm); 2013. p. 708–13.
[85] Martínez-Chico M, Batlles FJ, Bosch JL. Cloud classification in a mediterranean location using radiation data and sky images. Energy 2011;36(7):4055–62.
[86] Peng Z, Yu D, Huang D, Heiser J, Yoo S, Kalb P. 3D cloud detection and tracking system for solar forecast using multiple sky imagers. Sol Energy 2015;118: 496–519.
[87] Peng Z, Yu D, Huang D, Heiser J, Kalb P. A hybrid approach to estimate the complex motions of clouds in sky images. Sol Energy 2016;138:10–25.
[88] Ogliari E, Dolara A, Manzolini G, Leva S. Physical and hybrid methods comparison for the day ahead PV output power forecast. Renew Energy 2017; 113:11–21.
[89] Liu L, et al. Prediction of short-term PV power output and uncertainty analysis. Appl Energy 2018;228:700–11.
[90] Alanazi M, Alanazi A, Khodaei A. Long-term solar generation forecasting. In: 2016 IEEE/PES transmission and distribution conference and exposition (T&D); 2016. p. 1–5.
[91] Tuohy A, et al. Solar forecasting: methods, challenges, and performance. IEEE Power Energy Mag 2015;13(6):50–9.
[92] Zhen Z, Xuan Z, Wang F, Sun R, Duić N, Jin T. Image phase shift invariance based multi-transform-fusion method for cloud motion displacement calculation using sky images. Energy Convers Manag 2019;197:111853.
[93] Sophie Pelland JR, Jan Kleissl, Oozeki Takashi, De Brabandere Karel. Photovoltaic and solar forecasting: state of the art. International Energy Agency (IEA) Photovoltaic Power System Programme (PVPS); 2013.
[94] Chow CW, et al. Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. Sol Energy 2011;85(11):2881–93.
[95] Alonso-Montesinos J, Batlles FJ, Portillo C. Solar irradiance forecasting at one-minute intervals for different sky conditions using sky camera images. Energy Convers Manag 2015;105:1166–77.
[96] Reddy SS. Optimal scheduling of thermal-wind-solar power system with storage. Renew Energy 2017;101:1357–68.
[97] Reddy SS, Bijwe PR. Real time economic dispatch considering renewable energy resources. Renew Energy 2015;83:1215–26.
[98] Al-Dahidi S, Ayadi O, Alrbai M, Adeeb J. Ensemble approach of optimized artificial neural networks for solar photovoltaic power prediction. IEEE Access 2019;7:81741–58.
[99] Schwingshackl C, et al. Wind effect on PV module temperature: analysis of different techniques for an accurate estimation. Energy Procedia 2013;40:77–86.
[100] Ting-Chung Y, Hsiao-Tse C. The forecast of the electrical energy generated by photovoltaic systems using neural network method. In: 2011 international conference on electric information and control engineering; 2011. p. 2758–61.
[101] Paulescu M, Brabec M, Boata R, Badescu V. Structured, physically inspired (gray box) models versus black box modeling for forecasting the output power of photovoltaic plants. Energy 2017;121:792–802.
[102] Yang Z, Wang J. A hybrid forecasting approach applied in wind speed forecasting based on a data processing strategy and an optimized artificial intelligence algorithm. Energy 2018;160:87–100.
[103] Bacher P, Madsen H, Nielsen HA. Online short-term solar power forecasting. Sol Energy 2009;83(10):1772–83.
[104] Kemmoku Y, Orita S, Nakagawa S, Sakakibara T. Daily insolation forecasting using a multi-stage neural network. Sol Energy 1999;66(3):193–9.
[105] Sfetsos A, Coonick AH. Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques. Sol Energy 2000;68(2):169–78.
[106] Reikard G. Predicting solar radiation at high resolutions: a comparison of time series forecasts. Sol Energy 2009;83(3):342–9.
[107] Baig A, Akhter P, Mufti A. A novel approach to estimate the clear day global radiation. Renew Energy 1991;1(1):119–23.
[108] Kaplanis SN. New methodologies to estimate the hourly global solar radiation; Comparisons with existing models. Renew Energy 2006;31(6):781–90.
[109] Boland J. Time series modelling of solar radiation. In: V B, editor. Modeling solar radiation at the earth's surface. Berlin, Heidelberg: Springer; 2008.
[110] Boland J. Time-series analysis of climatic variables. Sol Energy 1995;55(5): 377–88.
[111] Yang C, Thatte AA, Xie L. Multitime-scale data-driven spatio-temporal forecast of photovoltaic generation. IEEE Trans Sustain Energy 2015;6(1):104–12.
[112] Aghajani A, Kazemzadeh R, Ebrahimi A. A novel hybrid approach for predicting wind farm power production based on wavelet transform, hybrid neural networks and imperialist competitive algorithm. Energy Convers Manag 2016;121:232–40.
[113] Hadi SJ, Tombul M. Monthly streamflow forecasting using continuous wavelet and multi-gene genetic programming combination. J Hydrol 2018;561:674–87.
[114] Du Z, Qin M, Zhang F, Liu R. Multistep-ahead forecasting of chlorophyll a using a wavelet nonlinear autoregressive network. Knowl Base Syst 2018;160:61–70.

[115] Al-Dahidi S, Ayadi O, Adeeb J, Alrbai M, Qawasmeh BR. Extreme learning machines for solar photovoltaic power predictions. Energies 2018;11(10):2725.
[116] AlOmari Mohammad H, Adeeb Jehad, Younis O. Solar photovoltaic power forecasting in Jordan using artificial neural networks. Int J Electr Comput Eng 2018;8:497–504.
[117] Zhen Z, et al. SVM based cloud classification model using total sky images for PV power forecasting. In: 2015 IEEE power & energy society innovative smart grid technologies conference. ISGT); 2015. p. 1–5.
[118] Wang F, et al. Daily pattern prediction based classification modeling approach for day-ahead electricity price forecasting. Int J Electr Power Energy Syst 2019;105: 529–40.
[119] Wang F, Mi Z, Su S, Zhang C. A practical model for single-step power prediction of grid-connected PV plant using artificial neural network. In: 2011 IEEE PES innovative smart grid technologies; 2011. p. 1–4.
[120] Donald WM. An algorithm for least-squares estimation of nonlinear parameters. J Soc Ind Appl Math 1963;11(2):431–41.
[121] Hagan MT, Menhaj MB. Training feedforward networks with the Marquardt algorithm. IEEE Trans Neural Network 1994;5(6):989–93.
[122] David JCM. Bayesian interpolation. Neural Comput 1992;4(3):415–47.
[123] Foresee FD, Hagan MT. Gauss-Newton approximation to Bayesian learning. In: Proceedings of international conference on neural networks (ICNN'97). vol. 3; 1997. p. 1930–5. 3.
[124] MathWorks. t. rainbr: Bayesian regularization backpropagation. 2019-11-02. Available: https://www.mathworks.com/help/nnet/ref/trainbr.html.
[125] Zhu HL, Li X, Sun Q, Nie L, Yao JX, Zhao G. A power prediction method for photovoltaic power plant based on wavelet decomposition and artificial neural networks. Energies Jan 2016;9(1).
[126] Golestaneh F, Hoay Beng G. Batch and sequential forecast models for photovoltaic generation. In: 2015 IEEE power & energy society general meeting; 2015. p. 1–5.
[127] Bashir ZA, El-Hawary ME. Applying wavelets to short-term load forecasting using PSO-based neural networks. IEEE Trans Power Syst 2009;24(1):20–7.
[128] Bao Y, Liu Z. A fast grid search method in support vector regression forecasting time series. Berlin, Heidelberg: Springer Berlin Heidelberg; 2006. p. 504–11.
[129] Li H, Guo S, Zhao H, Su C, Wang B. Annual electric load forecasting by a least squares support vector machine with a fruit fly optimization algorithm. Energies 2012;5(11):1–16.
[130] Haque AU, Nehrir MH, Mandal P. Solar PV power generation forecast using a hybrid intelligent approach. In: 2013 IEEE power & energy society general meeting; 2013. p. 1–5.
[131] Niu D, Wang Y, Wu DD. Power load forecasting using support vector machine and ant colony optimization. Expert Syst Appl 2010;37(3):2531–9.
[132] Hong W-C. Application of chaotic ant swarm optimization in electric load forecasting. Energy Pol 2010;38(10):5830–9.
[133] Hong W-C. Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm. Energy 2011;36(9): 5568–78.
[134] Hong W-C. Electric load forecasting by support vector model. Appl Math Model 2009;33(5):2444–54.
[135] Tao Y, Chen Y. Distributed PV power forecasting using genetic algorithm based neural network approach. In: Proceedings of the 2014 international conference on advanced mechatronic systems; 2014. p. 557–60.
[136] Wang F, Zhou L, Ren H, Liu X. Search improvement process-chaotic optimization-particle swarm optimization-elite retention strategy and improved combined cooling-heating-power strategy based two-time scale multi-objective optimization model for stand-alone microgrid operation. Energies 2017;10(12):1936.
[137] Wang F, Zhou L, Wang B, Wang Z, Shafie-khah M, Catal̃ão JoPS. Modified chaos particle swarm optimization-based optimized operation model for stand-alone CCHP microgrid. Appl Sci 2017;7(8):754.
[138] Almonacid F, Rus C, Hontoria L, Muñoz FJ. Characterisation of PV CIS module by artificial neural networks. A comparative study with other methods. Renew Energy 2010;35(5):973–80.
[139] Oudjana SH, Hellal A, Mahammed IH. Power forecasting of photovoltaic generation. Int J Electr Comput Energetic, Electron Commun Eng 2013;7. Art. no. 6.
[140] Dutta S, et al. Load and renewable energy forecasting for a microgrid using persistence technique. Energy Procedia 2017;143:617–22.
[141] Kumler A, Xie Y, Zhang Y. A Physics-based Smart Persistence model for Intra-hour forecasting of solar radiation (PSPI) using GHI measurements and a cloud retrieval technique. Sol Energy 2019;177:494–500.
[142] Zhao J, Guo Z-H, Su Z-Y, Zhao Z-Y, Xiao X, Liu F. An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed. Appl Energy 2016;162:808–26.
[143] Zameer A, Arshad J, Khan A, Raja MAZ. Intelligent and robust prediction of short term wind power using genetic programming based ensemble of neural networks. Energy Convers Manag 2017;134:361–72.
[144] Lei M, Shiyan L, Chuanwen J, Hongling L, Yan Z. A review on the forecasting of wind speed and generated power. Renew Sustain Energy Rev 2009;13(4):915–20.
[145] Monteiro C, Santos T, Fernandez-Jimenez LA, Ramirez-Rosado IJ, Terreros-Olarte MS. Short-term power forecasting model for photovoltaic plants based on historical similarity. Energies 2013;6(5):2624–43.
[146] Lorenz E, Heinemann D, Sayigh A. Prediction of solar irradiance and photovoltaic power. In: Comprehensive renewable EnergyOxford. Elsevier; 2012. p. 239–92.
[147] Mathiesen P, Kleissl J. Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States. Sol Energy 2011;85(5):967–77.
[148] Lorenz E, Scheidsteger T, Hurka J, Heinemann D, Kurz C. Regional PV power prediction for improved grid integration. Prog Photovoltaics Res Appl 2011;19 (7):757–71.
[149] Pelland S, Galanis G, Kallos G. Solar and photovoltaic forecasting through post-processing of the Global Environmental Multiscale numerical weather prediction model. Prog Photovoltaics Res Appl 2013;21(3):284–96.
[150] Soman SS, Zareipour H, Malik O, Mandal P. A review of wind power and wind speed forecasting methods with different time horizons. In: North American power symposium; 2010. p. 1–8. 2010.
[151] Firat U, Engin SN, Saraclar M, Ertuzun AB. Wind speed forecasting based on second order blind identification and autoregressive model. In: 2010 ninth international conference on machine learning and applications; 2010. p. 686–91.
[152] Cornaro C, Pierro M, Bucci F. Master optimization process based on neural networks ensemble for 24-h solar irradiance forecast. Sol Energy 2015;111: 297–312.
[153] Sbrana G, Silvestrini A. Random switching exponential smoothing and inventory forecasting. Int J Prod Econ 2014;156:283–94.
[154] Ferbar Tratar L, Strmčnik E. The comparison of Holt–Winters method and Multiple regression method: a case study. Energy 2016;109:266–76.
[155] David M, Ramahatana F, Trombe PJ, Lauret P. Probabilistic forecasting of the solar irradiance with recursive ARMA and GARCH models. Sol Energy 2016;133: 55–72.
[156] Mora-López L, Sidrach-de-Cardona M. Multiplicative ARMA models to generate hourly series of global irradiation. Sol Energy 1998;63(5):283–91.
[157] Hansen BE. Time series analysis James D, vol. 11. Hamilton Princeton University Press; 1994. p. 625–30. 1995- Econom Theor.
[158] Haiges R, Wang YD, Ghoshray A, Roskilly AP. Forecasting electricity generation capacity in Malaysia: an auto regressive integrated moving average approach. Energy Procedia 2017;105:3471–8.
[159] Cadenas E, Rivera W, Campos-Amezcua R, Heard C. Wind speed prediction using a univariate ARIMA model and a Multivariate NARX model. Energies 2016;6:109.
[160] Bui K-TT, Tien Bui D, Zou J, Van Doan C, Revhaug I. A novel hybrid artificial intelligent approach based on neural fuzzy inference model and particle swarm optimization for horizontal displacement modeling of hydropower dam. Neural Comput Appl June 01 2016;29(12):1495–506.
[161] Tien Bui D, et al. Hybrid intelligent model based on least squares support vector regression and artificial bee colony optimization for time-series modeling and forecasting horizontal displacement of hydropower dam. In: Handbook of neural computation. Academic Press; 2017. p. 279–93.
[162] Sadowski Ł, Hoła J, Czarnecki S, Wang D. Pull-off adhesion prediction of variable thick overlay to the substrate. Autom ConStruct 2018;85:10–23.
[163] Pham BT, Tien Bui D, Prakash I, Dholakia MB. Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. Catena 2017;149: 52–63.
[164] Goetzke-Pala A, Hoła J, Sadowski Ł. Non-destructive neural identification of the moisture content of saline ceramic bricks. Construct Build Mater 2016;113: 144–52.
[165] Hegazy T, Fazio P, Moselhi O. Developing practical neural network applications using back-propagation 1994;9(2):145–59.
[166] Tien Bui D, Nhu V-H, Hoang N-D. Prediction of soil compression coefficient for urban housing project using novel integration machine learning approach of swarm intelligence and Multi-layer Perceptron Neural Network. Adv Eng Inf 2018;38:593–604.
[167] Hertz J, Krogh A, Palmer R, Jensen RV. Introduction to the theory of neural computation. Baca Raton CRC Press Taylor and Francis Group; 1994.
[168] Elman JL. Finding structure in time. Cognit Sci 1990;14(2):179–211.
[169] Jordan MI, Donahoe JW, Packard Dorsel V. Serial order: a parallel distributed processing approach. In: Advances in psychology. vol. 121. North-Holland; 1997. p. 471–95.
[170] Williams RJ, Zipser D. Experimental analysis of the real-time recurrent learning algorithm. Connect Sci 1989;1(1):87–111.
[171] Wang D, Luo H, Grunder O, Lin Y, Guo H. Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm. Appl Energy 2017;190: 390–407.
[172] Ren C, An N, Wang J, Li L, Hu B, Shang D. Optimal parameters selection for BP neural network based on particle swarm optimization: a case study of wind speed forecasting. Knowl Base Syst 2014;56:226–39.
[173] Liu H, Tian H-q, Liang X-f, Li Y-f. Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks. Appl Energy 2015;157:183–94.
[174] Wang J, Qin S, Zhou Q, Jiang H. Medium-term wind speeds forecasting utilizing hybrid models for three different sites in Xinjiang, China. Renew Energy 2015;76: 91–101.
[175] Dong Q, Sun Y, Li P. A novel forecasting model based on a hybrid processing strategy and an optimized local linear fuzzy neural network to make wind power forecasting: a case study of wind farms in China. Renew Energy 2017;102: 241–57.
[176] Wang S, Zhang N, Wu L, Wang Y. Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method. Renew Energy 2016;94:629–36.
[177] Azadeh A, Ghaderi SF, Sohrabkhani S. Forecasting electrical consumption by integration of Neural Network, time series and ANOVA. Appl Math Comput 2007; 186(2):1753–61.

[178] Colak T, Qahwaji R. Automatic sunspot classification for real-time forecasting of solar activities. In: 2007 3rd international conference on recent advances in space technologies; 2007. p. 733–8.
[179] Yona A, Senjyu T, Funabashi T, Kim C. Determination method of insolation prediction with fuzzy and applying neural network for long-term ahead PV power output correction. IEEE Trans Sustain Energy 2013;4(2):527–33.
[180] Amjady N, Keynia F, Zareipour H. Wind power prediction by a new forecast engine composed of modified hybrid neural network and enhanced particle swarm optimization. IEEE Trans Sustain Energy 2011;2(3):265–76.
[181] Behera MK, Nayak N. A comparative study on short-term PV power forecasting using decomposition based optimized extreme learning machine algorithm. Eng Sci Technol, Int J 2019.
[182] Huang G-B, Zhu Q-Y, Siew C-K. Extreme learning machine: theory and applications. Neurocomputing 2006;70(1):489–501.
[183] Shukla S, Raghuwanshi BS. Online sequential class-specific extreme learning machine for binary imbalanced learning. Neural Network 2019;119:235–48.
[184] Wang X, Yang K, Kalivas JH. Comparison of extreme learning machine models for gasoline octane number forecasting by near-infrared spectra analysis. Optik 2020; 200:163325.
[185] Hossain M, Mekhilef S, Danesh M, Olatomiwa L, Shamshirband S. Application of extreme learning machine for short term output power forecasting of three grid-connected PV systems. J Clean Prod 2017;167:395–405.
[186] Ding M, Wang L, Bi R. An ANN-based approach for forecasting the power output of photovoltaic system. Procedia Environ Sci 2011;11:1308–15.
[187] Lawrence Steve, Lee Giles C, Tsoi AC. What size neural network gives optimal generalization? Convergence properties of backpropagation. 1998. p. 1–37. Networks, no. UMIACS-TR-96-22 and CS-TR-3617.
[188] Yang Z, Baraldi P, Zio E. A comparison between extreme learning machine and artificial neural network for remaining useful life prediction. In: Prognostics and system health management conference. PHM-Chengdu); 2016. p. 1–7. 2016.
[189] Zhang S, Tan W, Li Y. A survey of online sequential extreme learning machine. In: 2018 5th international conference on control. Decision and Information Technologies (CoDIT); 2018. p. 45–50.
[190] Bonissone PP, Xue F, Subbu R. Fast meta-models for local fusion of multiple predictive models. Appl Soft Comput 2011;11(2):1529–39.
[191] Polikar R. Ensemble based systems in decision making. IEEE Circ Syst Mag 2006;6 (3):21–45.
[192] Al-Dahidi S, Baraldi P, Zio E, Zio E, Legnani E. A dynamic weighting ensemble approach for wind energy production prediction. In: 2017 2nd international conference on system reliability and safety. ICSRS); 2017. p. 296–302.
[193] Schapire RE. The strength of weak learnability. Mach Learn, J Artic June 01 1990; 5(2):197–227.
[194] Breiman L. Bagging predictors. Mach Learn, J Artic August 01 1996;24(2): 123–40.
[195] Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci 1997;55(1):119–39.
[196] Ahmed Mohammed A, Aung Z. Ensemble learning approach for probabilistic forecasting of solar power generation. Energies 2016;9(12):1017.
[197] Omar M, Dolara A, Magistrati G, Mussetta M, Ogliari E, Viola F. Day-ahead forecasting for photovoltaic power using artificial neural networks ensembles. In: 2016 IEEE international conference on renewable energy research and applications (ICRERA); 2016. p. 1152–7.
[198] Pierro M, et al. Multi-Model Ensemble for day ahead prediction of photovoltaic power generation. Sol Energy 2016;134:132–46.
[199] Al-Dahidi S, Baraldi P, Zio E, Lorenzo M. Quantification of uncertainty of wind energy predictions. 2018.
[200] Brancucci Martinez-Anido C, et al. The value of day-ahead solar power forecasting improvement. Sol Energy 2016;129:192–203.
[201] Khosravi A, Nahavandi S, Creighton D, Atiya AF. Comprehensive review of neural network-based prediction intervals and new advances. IEEE Trans Neural Network 2011;22(9):1341–56.
[202] Ni Q, Zhuang S, Sheng H, Wang S, Xiao J. An optimized prediction intervals approach for short term PV power forecasting. Energies 2017;10(10):1669.
[203] Schwarz C. Sampling, regression, experimental design and analysis for environmental scientists. Biologists, and Resource Managers; 2011.
[204] Botev ZI, Grotowski JF, Kroese DP. Kernel density estimation via diffusion. Ann Stat 2010;38(5):2916–57.
[205] Nix DA, Weigend AS. Estimating the mean and variance of the target probability distribution. In: Proceedings of 1994 IEEE international conference on neural networks (ICNN'94). vol. 1; 1994. p. 55–60. 1.
[206] Hwang JTG, Ding AA. Prediction intervals for artificial neural networks. J Am Stat Assoc 1997;92(438):748–57.
[207] Ho SL, Xie M, Tang LC, Xu K, Goh TN. Neural network modeling with confidence bounds: a case study on the solder paste deposition process. IEEE Trans Electron Packag Manuf 2001;24(4):323–32.
[208] Khosravi A, Nahavandi S, Creighton D, Atiya AF. Lower upper bound estimation method for construction of neural network-based prediction intervals. IEEE Trans Neural Network 2011;22(3):337–46.
[209] Rigamonti M, Baraldi P, Zio E, Roychoudhury I, Goebel K, Poll S. Ensemble of optimized echo state networks for remaining useful life prediction. Neurocomputing 2018;281:121–38.
[210] Deng L, Yu D. Deep learning for signal and information processing. In: Microsoft research monograph; 2013.
[211] Geoffrey EH, Simon O, Yee-Whye T. A fast learning algorithm for deep belief nets. Neural Comput 2006;18(7):1527–54.
[212] Liu W, Wang Z, Liu X, Zeng N, Liu Y, Alsaadi FE. A survey of deep neural network architectures and their applications. Neurocomputing 2017;234:11–26.
[213] Papa JP, Rosa GH, Marana AN, Scheirer W, Cox DD. Model selection for discriminative restricted Boltzmann machines through meta-heuristic techniques. J Comput Sci 2015;9:14–8.
[214] Xie C, Lv J, Li Y, Sang Y. Cross-correlation conditional restricted Boltzmann machines for modeling motion style. Knowl Base Syst 2018;159:259–69.
[215] Tang Y, Salakhutdinov R, Hinton G. Robust Boltzmann Machines for recognition and denoising. In: 2012 IEEE conference on computer vision and pattern recognition; 2012. p. 2264–71.
[216] Hinton GE, Dayan P, Frey BJ, Neal RM. The "wake-sleep" algorithm for unsupervised neural networks. Science 1995;268(5214):1158.
[217] Nair V, Hinton GE. 3D object recognition with deep belief nets. In: Culotta YBaDSaJD LaCK IWaA, editor. Advances in neural information processing systems. Curran Associates, Inc.; 2009. p. 1339–47.
[218] Deng L, Yu D, Platt J. Scalable stacking and learning for building deep architectures. In: 2012 IEEE international conference on acoustics, speech and signal processing. ICASSP); 2012. p. 2133–6.
[219] Arel I, Rose DC, Karnowski TP. Deep machine learning - a new frontier in artificial intelligence research [research frontier]. IEEE Comput Intell Mag 2010;5(4):13–8.
[220] Liao B, Xu J, Lv J, Zhou S. An image retrieval method for binary images based on DBN and softmax classifier. IETE Tech Rev 2015;32(4):294–303.
[221] Bourlard H, Kamp Y. Auto-association by multilayer perceptrons and singular value decomposition. Biol Cybern 1988;59(4):291–4.
[222] Deng L. Three classes of deep learning architectures and their applications: a tutorial survey. APSIPA Transactions on Signal and Information Processing; 2012.
[223] Fukushima K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern 1980;36(4):193–202.
[224] Jiang X, Pang Y, Sun M, Li X. Cascaded subpatch networks for effective CNNs. IEEE Trans Neural Network Learn Syst 2018;29(7):2684–94.
[225] Gu J, et al. Recent advances in convolutional neural networks. Pattern Recogn 2018;77:354–77.
[226] Wang Y, Li H. A novel intelligent modeling framework integrating convolutional neural network with an adaptive time-series window and its application to industrial process operational optimization. Chemometr Intell Lab Syst 2018;179: 64–72.
[227] Eigen D, Rolfe J, Fergus R, Lecun Y. Understanding deep architectures using a recursive convolutional network. In: International conference on learning representations (ICLR2014). CBLS; 2014.
[228] Desjardins G, Bengio Y. Empirical evaluation of convolutional RBMs for vision. In: Département d'Informatique et de Recherche Opérationnelle. Université de Montréal; 2008.
[229] Jarrett K, Kavukcuoglu K, Ranzato M, LeCun Y. What is the best multi-stage architecture for object recognition?. In: 2009 IEEE 12th international conference on computer vision; 2009. p. 2146–53.
[230] Mathieu M, Henaff M, Lecun Y. Fast training of convolutional networks through FFTs. In: Presented at the international conference on learning representations (ICLR2014). CBLS; 2014.
[231] He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE international conference on computer vision. vol. 2015. International Conference on Computer Vision, ICCV; 2015. p. 1026–34. 2015.
[232] Sezer OB, Ozbayoglu AM. Algorithmic financial trading with deep convolutional neural networks: time series to image conversion approach. Appl Soft Comput 2018;70:525–38.
[233] Qing X, Niu Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy 2018;148:461–8.
[234] Sivaneasan B, Yu CY, Goh KP. Solar forecasting using ANN with fuzzy logic pre-processing. Energy Procedia 2017;143:727–32.
[235] Sharma V, Yang D, Walsh W, Reindl T. Short term solar irradiance forecasting using a mixed wavelet neural network. Renew Energy 2016;90:481–92.
[236] Gutierrez-Corea F-V, Manso-Callejo M-A, Moreno-Regidor M-P, Manrique-Sancho M-T. Forecasting short-term solar irradiance based on artificial neural networks and data from neighboring meteorological stations. Sol Energy 2016; 134:119–31.
[237] Pedro HTC, Coimbra CFM. Short-term irradiance forecastability for various solar micro-climates. Sol Energy 2015;122:587–602.
[238] Chu Y, Urquhart B, Gohari SMI, Pedro HTC, Kleissl J, Coimbra CFM. Short-term reforecasting of power output from a 48 MWe solar PV plant. Sol Energy 2015; 112:68–77.
[239] Ramsami P, Oree V. A hybrid method for forecasting the energy output of photovoltaic systems. Energy Convers Manag 2015;95:406–13.
[240] Liu J, Fang W, Zhang X, Yang C. An improved photovoltaic power forecasting model with the assistance of aerosol index data. IEEE Trans Sustain Energy 2015; 6(2):434–42.
[241] Raza MQ, Nadarajah M, Li J, Lee KY, Gooi HB. An ensemble framework for day-ahead forecast of PV output in smart grids. IEEE Trans Ind Inf 2018. 1-1.
[242] Lu HJ, Chang GW. A hybrid approach for day-ahead forecast of PV power generation. IFAC-PapersOnLine 2018;51(28):634–8.
[243] Jamali B, Rasekh M, Jamadi F, Gandomkar R, Makiabadi F. Using PSO-GA algorithm for training artificial neural network to forecast solar space heating system parameters. Appl Therm Eng 2019;147:647–60.
[244] İzgi E, Öztopal A, Yerli B, Kaymak MK, Şahin AD. Short–mid-term solar power prediction by using artificial neural networks. Sol Energy 2012;86(2):725–33.

[245] Notton G, Paoli C, Ivanova L, Vasileva S, Nivet ML. Neural network approach to estimate 10-min solar global irradiation values on tilted planes. Renew Energy 2013;50:576–84.
[246] Dahmani K, Dizene R, Notton G, Paoli C, Voyant C, Nivet ML. Estimation of 5-min time-step data of tilted solar global irradiation using ANN (Artificial Neural Network) model. Energy 2014;70:374–81.
[247] Marquez R, Pedro HTC, Coimbra CFM. Hybrid solar forecasting method uses satellite imaging and ground telemetry as inputs to ANNs. Sol Energy 2013;92:176–88.
[248] Paoli C, Voyant C, Muselli M, Nivet M-L. Forecasting of preprocessed daily solar radiation time series using neural networks. Sol Energy 2010;84(12):2146–60.
[249] Huang Y, Lu J, Liu C, Xu X, Wang W, Zhou X. Comparative study of power forecasting methods for PV stations. In: International conference on power system technology; 2010. p. 1–6. 2010.
[250] Wen L, Zhou K, Yang S, Lu X. Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting. Energy 2019;171:1053–65.
[251] Siddiqui TA, Bharadwaj S, Kalyanaraman S. A deep learning approach to solar-irradiance forecasting in sky-videos. In: 2019 IEEE winter conference on applications of computer vision. WACV); 2019. p. 2166–74.
[252] Lee W, Kim K, Park J, Kim J, Kim Y. Forecasting solar power using long-short term memory and convolutional neural networks. IEEE Access 2018;6:73068–80.
[253] Zhang R, Feng M, Zhang W, Lu S, Wang F. Forecast of solar energy production - a deep learning approach. In: 2018 IEEE international conference on big knowledge. ICBK); 2018. p. 73–82.
[254] Haixiang Z, et al. Hybrid method for short-term photovoltaic power forecasting based on deep convolutional neural network (IET Generation, Transmission & Distribution). Institution of Engineering and Technology; 2018. p. 4557–67.
[255] Srivastava S, Lessmann S. A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data. Sol Energy 2018;162:232–47.
[256] Gensler A, Henze J, Sick B, Raabe N. Deep Learning for solar power forecasting — an approach using AutoEncoder and LSTM Neural Networks. In: 2016 IEEE international conference on systems, man, and cybernetics. SMC); 2016. p. 2858–65.