Solar Irradiation Forecasting Technologies: A Review

Anuj Gupta[1,*], Kapil Gupta[1] and Sumit Saroha[2]

[1]*Maharishi Markandeshwar (Deemed to be) University, Mullana-Ambala, India*
[2]*Guru jamdheshwar University of Science and Technology, Hisar, India*
E-mail: annu11gupta@gmail.com
[*]*Corresponding Author*

Received 06 February 2021; Accepted 25 June 2021;
Publication 10 July 2021

Abstract

Renewable energy has received a lot of attention in the previous two decades when it comes to meeting electrical needs in the home, industrial, and agricultural sectors. Solar forecasting is critical for the efficient operation, scheduling, and balancing of energy generation by standalone and grid-connected solar PV systems. A variety of models and methods have been developed in the literature to forecast solar irradiance. This paper provides an analysis of the techniques used in the literature to forecast solar irradiance. The main focus of the study is to investigate the influence of meteorological variables, time horizons, climatic zone, pre-processing technique, optimization & sample size on the complexity and accuracy of the model. Due to their nonlinear complicated problem solving skills, artificial neural network based models outperform other models in the literature. Hybridizing the two models or performing pre-processing on the input data can improve their accuracy even more. It also addresses the various main constituents that influence a model's accuracy. The paper provides key findings based on studied literature to select the optimal model for a specific site. This paper also discusses the metrics used to measure the efficiency of forecasted model. It has been observed that the proper selection of training and testing period also enhance the accuracy of the model.

Strategic Planning for Energy and the Environment, Vol. 39_1–4, 319–354.
doi: 10.13052/spee1048-4236.391413

Keywords: Forecasting techniques, hybrid models, neural network, solar forecasting, error metrics, SVM, Markov chain, numerical weather prediction.

1 Introduction

With rising energy demands and limited availability of fossil fuel all over the world encouraged us to move towards the renewable energy sources such as solar, biomass, geothermal, wind and ocean energy [1] etc. These alternative sources provide a potential solution to meet the huge energy demand. Solar energy is one of the most promising resources of energy that is naturally available on the earth surface [2]. The Earth's surface received approximately 1.5 to 10^{18} KWh/year of solar energy annually that is approximately multiple of 10,000 of total world consumption [3]. All nations are working to build a solar photovoltaic network in an effective way to produce solar power. The number, size and electricity production of photovoltaic plants has increased worldwide, with a combined generation capacity of up to 500 GW [4]. In India, over the past decades, the renewable energy sector has expanded exponentially [5]. In 1992, India has formed a separate renewable energy ministry to encourage the use of renewable energy known as the Ministry of New and Renewable Energy (MNRE). According to the report published by India's government in 2019, the 80GW mark has been crossed with 25 GW of generation only from solar [7]. Solar radiation estimation is related to solar radiation components such as Direct Normal Radiation (DNI), Diffuse Horizontal Radiation (DHI), and Global Horizontal Radiation (GHI) which is the sum of DNI and DHI [8]. However, the measurement of these components is complex due to the climatic and geographical condition of the site [9]. These types of sites need a solar model to estimate these components using time series data [10]. Solar projection modeling includes predicting the precise details of the solar radiation components to decide whether or not to set a plant at a new location. There are many places on the Earth where the measurement of solar radiation is not only a typical task but also sometimes difficult to calculate because of the measuring device costs, upkeep and calibration [11]. Many countries have grid interconnectivity with solar plants and offer the opportunity to sell the excessively generated electricity that opens the door to the common man to earn money [12]. Every country has its schemes and policies to boost its solar market [13]. A large number of researcher continue to work with the solar photovoltaic cell's size, modeling, structure, device, battery and physical parts to efficiently transform solar radiation into electricity but many of the researchers choose to work with the planning of solar PV power plants. Photovoltaic power plant scheduling and planning is a critical task because both operations are carried out under volatile weather conditions, that may result in the poor balancing of load demand and energy production, which further results in the penalty on power producer [14]. Hence it is highly appreciable to build an optimum model for predicting the solar radiation components accurately for both live and offline data [15]. While several researchers have already carried out the solar prediction analyses, taking into account artificial neural network model and hybrid-based technique as well as currently developed model based on pre-processing technique, optimization technique, training and testing period & accuracy evaluation metrics. Various solar forecasting research activities get motivated due to the factors that accurate solar forecasting techniques increase the efficiency of the energy supplied to the grid and mitigate the additional costs associate with weather dependency. Various types of forecasting approaches are introduced based on application; methods and time horizon.

1. For very short time scale various time series models such as an Artificial Neural Network, Autoregressive Integrated Moving Average, and Persistence model used for forecast solar irradiance [6, 8, 13].
2. For short time irradiance forecasting, solar irradiance largely depend on the observation based on the temporal developments of clouds, may be used as a basis.
   - For the sub-hour range, cloud data is collected from sky images ground based with high spatial resolution may be used to predict solar irradiance.
   - For 30 minutes up to 6 hrs solar irradiance depends on cloud motion vector from satellite photos.
3. For long term horizon, from 4–6 hours ahead numerical weather prediction model perform better than the satellite based forecasts [9, 13].
4. There are also integrated techniques to derive an optimized forecast for the different-2 time horizon.

The main aim of this paper is to classify and analyze the forecasting technique based on exogenous and endogenous data. The rest of this paper is organized as follows: The next section discuss motivation for the review. Sections 3 and 4 describe the different techniques used for pre-processing and types of input variables; Section 5 gives a brief about the optimization technique; Sections 6, 7, 8 and 9 present a critical analysis of solar forecasting technique. Section 10 describes the factor affecting the prediction of solar irradiation. Section 11 addresses the measurement criteria for the solar irradiation forecast. The analysis is eventually concluded in Section 12.

2 Motivation for the Review

The ambiguity associated with solar energy due to its dependency on weather parameter can adversely affect the operation of micro & national grid. So, it is necessary to design an efficient planning for predicting solar energy. In order to address this situation, a range of options with a strong emphasis on renewable energy are considered. Over the past two decades a lot of researcher and academics are engaged in developing tools, models and algorithms with varying degree of success. In today's dynamic world, forecasting is a critical part of business planning with greater penetration of renewable energy resources and implementation of power deregulation in industry. Forecasting of solar power has become a major issue in power systems. Following needs of the markets, various techniques are used to forecast the solar radiation. 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 solar power output. The knowledge gain may also help policy makers and energy market participants to make more effective and profitable decision concerning the implementation of solar power system.

3 Input Variables

The extremely unpredictable existence of the solar system originates from uncertainties of its input variables. The better is the selection of input variables, lower is the prediction error. The variables in the input data may be systemic, endogenous and exogenous. On various combinations of input parameter different model behave differently. In most studies, ANN provides importance to meteorological and geographical variables. The increased number of irrelevant metrological parameters degrades the performance of the model. Therefore, the appropriate parameters have to select to increase the performance of a model. M. A. Behrang et al. developed six ANN model using a different combination of metrological parameters to predict solar radiation [20]. However, forecasting solar power is a crucial task; because emission of solar radiation is a natural stochastic process and having some special characteristics like highly uncertain, non-linearity, pre-stationarity

Table 1 The factors affecting solar power and its derivatives

Class	Input Variable
1. Atmospheric Characteristics	Pressure, Temperature, Cloudiness, Rainfall, Cloud formation, Cloud cover Stratification of the atmosphere, Radiations, Humidity, Density, Wind power, Wind speed, Wind direction, Evaporation Sunshine duration, Wind gust, Mean Temperature, Ambient temperature, Minimum temperature, Maximum temperature, Sky information, Average temperature
2. Solar Characteristics	Solar power, Solar irradiation, Solar zenith angle, Global horizontal irradiance, Diffuse horizontal irradiance, Direct normal irradiance, Global solar radiation, Daily solar radiation, Cell temperature, Long wavelength, precipitation, Photovoltaic
3. Geographical Conditions	Latitude, Longitude, Altitude

and high complexity depending on the scale of various physical conditions as set out in Table 1.

3.1 Use of Primary data

Different studies used different variables as a primary data to evaluate their models. Masoud Vakili et al., used relative humidity, wind speed and daily temperature as primary variables for their model [34]. M.A. Behrang et al., used six different combination of input variables: day of the year, daily mean air temperature, relative humidity, sunshine hours, evaporation, and wind speed [20]. However, Vassilis Z. Antonopoulos et al., used air temperature, solar radiation, wind speed & relative humidity for their study [17]. Fermin Rodriguez et al., used atmospheric pressure, relative air humidity & air temperature from a Euskalment Government agency to predict the generation of solar energy from PV generators [29].

3.2 Use of Secondary Data

Some of the studies predicted their target value by numerical methods and by using the time series data can be expressed as secondary data. Fonseca JGDS et al., used the effect of forecast horizon on the accuracy of the PVPF was studied using numerically predicted weather data through SVM [101]. This study compared two versions of ST-PVPF models: analytical PVPF

& multiple layer perceptron PVPF using numerically forecast weather data and historically hourly values for the generation of PV electric power. The value of the RMSE were similar for the both models which is from 11.95%–12.10% [102].

4 Input Data Pre-processing

The quality of input variables plays a crucial role in the accuracy of forecasted model. The data collected from various sites mostly available in raw format and does not have sufficient characteristics to provide appropriate accuracy. So, the data has to be process before processing with the model called pre-processing stage. Here, the pre-processing means scale up or down the input measurement, clean up and define the input data according to the specifications. With the help of pre-processing, we can reduce the improper training problem & operational cost by learning the historical data properly, therefore the model accuracy can be greatly improved by the pre-processing of input data. There are number of pre-processing techniques available in the literature such as Wavelet transform, Kalman filer, Empirical mode decomposition, Self organization map, normalization, trend free time series which were used before the model learning.

4.1 Wavelet Transform

The concept behind the wavelet transform is to decompose input data series into a set of meaningful series. These meaningful series improved forecasting accuracy by reducing input data and provide better presentation than original series. Each component is forecast separately & then final output obtained by aggregating the entire forecasted component into one signal. On the other hand the high frequency signals are evaluate by non-linear filter and low frequency signal are evaluate by linear filters. There are four filter decomposition low and high pass filter, reconstruction low pass & reconstruction high pass filter.

Mohammadi et al., developed a hybrid model for prediction of global solar radiation. The proposed model is based on wavelet transform & SVM (WT-SVM).To evaluate the performance of WT-SVM author used a different combination of weather data. They developed five SVM models corresponding to each combination of weather parameters, than aggregating all the output of SVM-WT models to generate a value of GSR. The output of all SVM-WT models outperforms the conventional models in terms of MAPE, MABE, rRMSE & R^2 [103]. Monjoly et al., used a pre-processing technique WT & EMD with AR & neural network models. The final result is obtained by combining the output of all forecasting models. The obtained result improves the performance of AR & ANN models in terms of rMBE, rMAE & rRMSE [82]. Prasad et al., presented a hybrid model for predicting the solar radiation on the horizontal surface. Empirical mode decomposition is used as a decomposing the data into intrinsic mode function. It observes that the proposed model shown better performance as a comparison to stand alone models [104].

4.2 Kalman Filter

Kalman filter are designed to reduce the systematic error, over fitting & complexity which significantly improve the quality of the solar forecast. Kalman filter has a strong capability of handling unpredictable fluctuations & over trainining of data during the learning process.

Diagne et al., proposed a combination of kalman filter & WRF to improve the forecasting quality of global solar radiation. The performance of proposed model shown that kalman filter improves the forecasting accuracy as compared to only WRF model [105].

Hussain et al., introduced a Bayesian E-kalman filter to adjust the weight of MLP model. The proposed mechanism compared with the benchmark model shown better performance in terms of R^2 & modeling [106]. However, kalman filter also used by the che et al., to improve the forecast accuracy in terms of RMSE & MBE [107].

4.3 SOM

Self organization map are based on clustering based ensemble learning approach in which divide the solar power series into different clusters & each cluster carry a high & low frequency signal. These high and low frequency signals are estimated by the linear & non-linear models & then aggregated to obtain the output signal.

Dong et al., introduced a hybrid model based on SOM to divide the time series into different clusters with similar characteristics. The model achieved categorizing capability of self organization map of the input data [108]. On the other hand, Hady et al., proposed a SOM-SVR-PSO hybrid model. Due to strong capability of dividing the time series data into uniform characteristics, the proposed model showed better performance in terms of nRMSE &

nMBE [109]. However, the SOM Pre-processing were also used by the Wu & wang et al., to improve the forecasting accuracy [110].

4.4 Normalization

Normalization has been utilized to compress the high frequency signals & converse into a smaller range to increase computational economy. Moreover, the missing & distributed values of just often sunrise & just before sunset have to delete due to imaging effects. The correlation among the datasets was also maintained by Alomari et al., in their study using normalization [111].

4.5 Trend Free Time Series

Trend is a systematic change in the series that does not appear periodically for faster modeling and improve model performance. We identifying and understanding the trends information & remove it. Generally, deterministic and stochastic two type of trend exist in the time series. Deterministic trends which deal with the consistently increase or decrease time series whereas, the nature of stochastic trend is inconsistent. A stochastic time series is called non-stationary if a data set does not have a trend. Reikard et al., improved the performance of model by removing the seasonality trend from their datasets [112].

5 Optimization Techniques

Evolutionary techniques have been adopted to overcome the problem of correlation by selecting the input parameter properly. A weak correlation generally produces non-linearity & computational complexity in the system. Various types of evolutionary techniques such as: Firefly, PSO, GA, Artificial Bee colony, Coral Reefs, Cuckoo search, simulated annulling, Biogeography optimization, Glowworm swarm intelligence, Bats warm etc. that has already been used in the literature by several researcher.

Ibrahim et al., used firefly algorithm to optimized the input parameter for improving the forecasting model [113]. However, the PSO optimization is also used to improve the performance with ANFIS technique as describe by Halabi et al. [114]. As per literature, Genetic algorithm work well with ANN & it is best evolutionary for changing the size of hidden layer & to set up a correlation among input parameters. Jovanovic et al., claimed that genetic algorithm improve the performance of ANN model. Genetic algorithm is the most famous & efficient evolutionary amongst other optimization techniques [115]. At present Genetic algorithm used in scale factor & translation factor optimization, input parameter selection, feature selection & build the forecasting model for global solar radiation, adapting the size of hidden layer & in another application as a replacement for genetic algorithm.

6 Classification of Solar Forecasting Techniques

Forecasting of the solar irradiation component is the method of predicting the different component of solar irradiation like GHI, DNI & DHI for a given PV site in advance. However, forecasting these components for a particular location in advance is not an easy task, as it requires predicting at various time horizons which are liable to affect by variable climatic conditions. There are three main methods: method of statistical time series, physical methods and ensemble method. Figure 1 represents the forecasting methods.

6.1 Statistical Time Series Method

The statistical models are dependent on historical data as input and are independent of the internal state of the model [22]. The statistical time series method are constructed using the different techniques which include Artificial neural network model, Regression model, Support vector machine, Markov chain.

6.1.1 Artificial neural network

The ANN technique function is almost comparable with the human brain that makes the decision based on the biological neuron. The neuron in the human brain performs a different type of parallel processing, pattern recognition analysis. The same can be used in solving non-linear mathematics like as forecasting, image processing etc. This technique train the ANN model repeatedly to obtain the best value of weight to map the input and output. The ANN model consists of a three-layer (i) input layer (ii) hidden layer (iii) output layer as shown in Figure 1.

The ANN definition was suggested by the McCulloch and Pits in 1934. The ANN uses the different-2 type of algorithm like as Levenberg Marquardt (LM) algorithm, scaled conjugate gradient, pola-ribiere conjugate gradient to predict the output value [21].

Babak Jahani et al. compared the empirical, ANN, and ANN with a genetic algorithm model to forecast the global solar radiation for the location of Iran. The Genetic algorithm was used in the model to

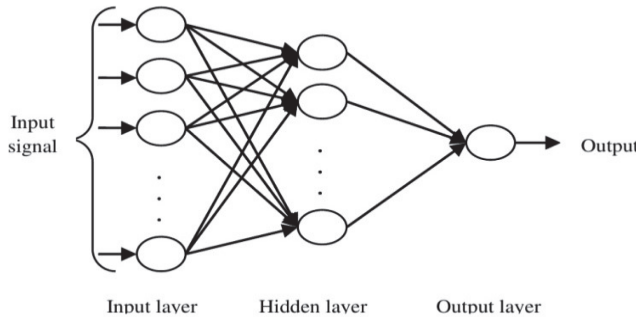


Figure 1 General Architecture of ANN.

reduce the error in predictive results. The ANN+GA model achieved better accuracy in comparison to other models with RMSE, MBE & R^2 of 0.92J/cm^2/day,38.4J/cm^2/day,185.5J/cm/day respectively [30]. Furthermore, the future value of solar irradiance using ANN & Fuzzy logic along with error correction method was predicted by B.Sivaneasan et al., introduced an improved solar forecast algorithm based on artificial neural network model with fuzzy logic. Error correction method was applied to the past value to correct the error acquired by the back propagation algorithm. The solar forecasted output depends on the weather information so, it is necessary to ensemble a pre-processing stage into the artificial neural network. The MAPE obtained for three models ANN, ANN with fuzzy pre-process and ANN with fuzzy pre-processing along with error correction were 46.3%, 43.1%, 29.6% respectively [31]. Mohammed Bou-Rabee et al. estimated the global solar radiation by gradient descent method and LM backpropagation algorithm. The accuracy of this model was determined by the MAPE which was 86.3% for the gradient descent method and 85.6% for LM [25].

Premalatha Neelamegam et al. proposed two ANN model with different combinations of inputs. the accuracy of the model was measured based on MAE, RMSE and R^2. Gradient descent, Levenberg-Marquardt backpropagation (LM), resilient backpropagation (RP) and scaled conjugate gradient (SCG) algorithm were used in the present work to predict the global solar radiation. The two models having inputs from latitude, longitude, altitude, month, mean ambient air temperature, mean station level pressure, wind speed, mean relative humidity and average global solar radiation. The LM algorithm has shown better performance as a comparison to other algorithm used with ANN. The best ANN model for the present work has an MAE, RMSE, R2 value for training and testing data as 0.7800, 1.0416, 0.9545 and

3.0281, 3.6461, 0.9272 respectively [19]. Chao-Rang Chen et al., designed model to forecast solar irradiation using the combination of K-NN and ANN techniques. This method used training and testing data collected on a day for continuous four hours with 5 minutes of the interval from nine different PV plants locations. The K-NN was used as the pre-processing of the input data and then processed by the ANN model. The model performance was determined by MABE (W/m^2) and RMSE (W/m^2). The RMSE for K-NN-ANN was 242(W/m^2) and MABE was 42(W/m^2) whereas RMSE and MABE for K-ANN were 251(W/m^2) and 44(W/m^2) respectively [32]. James Mubiru compared the two forecasting design technique ANN and Empirical model for Solar PV energy. The ANN model was explored as a feed-forward neural network having input parameters of latitude, longitude, altitude, Sunshine hours, maximum temperature and monthly average daily value of global solar radiation. The performance of the ANN and Empirical model was determined by R^2, MBE, and RMSE. The comparison of ANN and Empirical model has shown the superiority of proposed ANN model and the value of correlation coefficient (R^2), MBE and RMSE are 0.998,0.005 (MJ/m^2), 0.197 (MJ/m^2) [76]. Masood Vakili et al. introduced two models for forecasting of global solar radiation. Two combinations of various meteorological variables including particulate matters and without particulate matters were prepared and applied on the neural network. The model performance was calculated by MAP, RMSE, & R^2 which were 3.13, 0.077 & 0.97 respectively [34]. Shah Alam et al. introduced ANN models for four different stations of India: Ahmadabad, Nagpur, Mumbai and Vishakhapatnam. ANN model is used for estimating monthly mean hourly and daily diffuse solar radiation. The performance of ANN models has been shown based on RMSE and MBE. The RMSE and MBE were calculated for all the targeted location of India where maximum RMSE obtained 4.5% among all sites [35].

Fermin Rodriguez et al. developed an ANN model to forecast solar irradiance. The appropriate neuron was selected and kept constant with variation in the input delay to obtain the good accuracy of the ANN model. The RMSE obtained for sunny, partially cloudy and cloudy days by this model was 0.03%, 0.49% and 0.64% respectively [29]. N. Kumar et al., discussed an artificial neural network, an artificial neural network with forwarding unity gain and regression network for the prediction of daily global solar radiation (DGSR) of 10 Indian cities. The data collection considered various parameters as an input to the proposed model like as minimum temperature, average temperature, mean temperature, wind speed, relative humidity, precipitation, extra-terrestrial radiation and hours of sunshine. Statistical measure (RMSE, MBE, and MAPE) were evaluated to assess the accuracy of the forecasting model and were used to compare the effectiveness of the proposed method to recent literature studies. It was found that the proposed method estimates the DGSR with an error of 14.84%, 14.68%, and 16.38% by the ANN, ANN with forwarding unity gain, and ANN with RBF network respectively [27]. Khalil Benmouiza et al. presented research for hourly global horizontal solar radiation prediction based on the combination of unsupervised clustering algorithm K-means and non-linear autoregressive artificial neural networks (ANN). K-means algorithm based on extracting useful data information to modeling the behavior of time series and finding input space pattern by clustering the data. The nonlinear autoregressive neural networks are powerful computational models for non-linear time series modeling and forecasting. Taking advantage of both approaches, a new approach was proposed to produce better predictive outcomes [116]. M.A.Behrang et al. developed a model to predict the DGSR using ANN-based algorithm. The six models have been designed to estimate the DGSR with a different combination of meteorological input variables. The accuracy of the model was observed using the MAPE along with comparison with several conventional models. The MAPE of 5.21% was obtained for the input combination of daily mean air temperature, sunshine hours, wind speed and relative humidity [20].

6.1.2 Support vector machine

It is a form of machine learning introduced in 1995 by Cortes and Vapnik with statistical learning. Firstly this particular approach is developed for pattern recognition and is now enthusiastically used for various technologies such as image retrieval, fault diagnosis, regression computation and forecasting etc. The time series is used to train a model that is as simple as a neural network model and there is no question of the over fitting curve, struck to local minima in SVM [24]. Essentially, it uses the mapping function to map the input vector $(\mathrm{x}_1 + \mathrm{x}_2 + \mathrm{x}_3 + \cdots \mathrm{x}_\mathrm{n})$ to the output $(\mathrm{y}_1 + \mathrm{y}_2 + \mathrm{y}_3 + \cdots \mathrm{y}_\mathrm{n})$ with the mapping function Ø. the SVM Equation (1) can be expressed as [6].

$$y = \sum_{j=1}^{n} W_k \phi_{ij} + b \tag{1}$$

W is weight where Y is output function, and b is bias. The basic architecture of SVM is shown in Figure 2.

To predict the solar PV power output, Jie Shi et al. developed a model based on the support vector machine. This research separated the entire

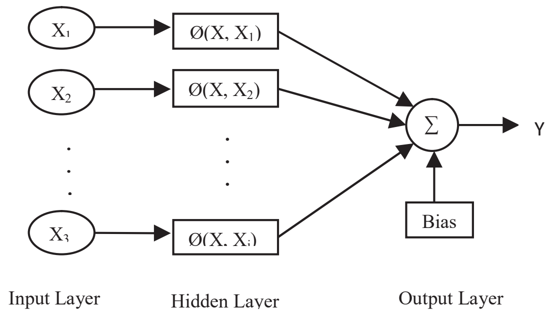


Figure 2 Structure of SVM.

historical and weather data into four categories then these four categories were applied to four different SVM models with different Kernel functions of the radial basis. The RMSE & MRE for the cloudy model was 1.824 (MW) & 12.42% respectively, for the foggy day was 2.52 (MW) & 8.16% respectively, for the rainy day was 2.48 (MW) & 9.12% respectively and for the sunny day was 1.57 (MW) & 4.85% respectively [44]. Han Seung Jang et al. developed a model for solar irradiation forecasting based on satellite image and SVM along with the prediction of cloud quantities at the target site. The atmosphere motion vector scheme was used to extract the motion vector information from the satellite's images. The model performance was determined by RMSE, MSE & R^2 which were 44.1390 (W/m^2), & 7.7329% and 0.9420 respectively [45]. Wolff et al., developed a 15 min to 5-h PV power forecast using a statistical learning model support vector machine (SVM). This model was assessed and served as an alternative to the complement another physical prediction model through PV power calculation, NWP and cloud motion vectors. The PV power was observed using three input sources such as SVR show good performance in prediction up to 1 hour ahead, NWP show better prediction starting at 3 hours ahead and CMVs is best amongst them [43].

6.1.3 Markov chain

It follows a stochastic cycle that uses very short term forecasting of solar irradiance. The Markov chain cycle is essentially dependent on the neighboring states i.e. the current state parameters are dependent on the previous one. Similarly the next state parameter is dependent on the current state [23] as shown in Figure 3.

Markov chain is represented by a series of finite random numbers $Y_1, Y_2, Y_3 \ldots.$

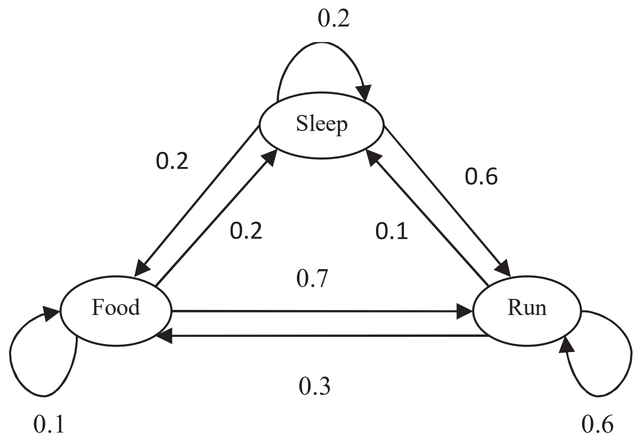


Figure 3 Markov process.

With the Markov property, as shown in the below Equation (2)

$$P_r(Y_{n+1} = y|Y_n = y_n \ldots, Y_1 = y_1) = P_r(Y_{n+1} = y|Y_n = y_n) \quad (2)$$

This equation shows the next state having the dependency on the present state of the series.

Saurabh Bhardwaj et al. used a fuzzy model to forecast solar irradiation. The training of the cluster was performed by the generalized fuzzy model. The input combination of day number, sunshine hours, relative humidity provided RMSE & MPE of 7.9124 & 3.4255 respectively [47]. Sanjari et al. developed a Higher-order Markov chain model to estimate the power of photovoltaic systems. Solar irradiance and ambient temperature were used as input parameters. The performance of proposed model performance was determined using MAE forecast error in which outperforms other approaches i.e. support vector machine, Chronological Probability with an average MAE value of 2.18% [81].

Shuai Li et al. developed a model based on a discrete-time Markov chain to overcome the effect of fluctuation in solar radiation. Based on the clear sky ratio, the model prepared the data in four categories and then clustered using k-means cluster algorithm based on feature vector output. The model performance was calculated by the average percentage error with the comparison of a typical meteorological year (TMY). The synthesized typical solar radiation year have the average maximum and minimum error of 10% and 6% respectively [49].

6.1.4 Regression model

The model in this category is focused on the mathematical relationship between the dependent and independent variable. Various models based on linearity or non-linearity of the data such as AR, MA, ARMA, ARIMA, SARMIA [36] etc. In 1992 Yule proposed two new approaches for analysis of stationary approaches. Where moving average (MA) was the first and Autoregressive (AR) was the second. Among all models, ARIMA is the most popular model that establishes the relationship between the actual measured output and forecasted outputs [37].

ARIMA Model

There are two methods of testing stationarity of the sequence, one of which is to assess the stationarity of sequence by sequence diagram and auto-correlation diagram characteristics and another is the development of test statistics to test hypotheses. The graph test method is a very simple and widely used method for determining stationarity. However, its drawback is that the discriminant conclusion has a strong subjective colour. Hence the only way to assist the decision is to use the statistical test process. At present, the unit root test is the most widely used stationary statistical test process. ARIMA is an extension of ARMA method [38]. It is a combination of AR & MA and used to find the correlation between input and output time series. AR and MA becomes the ARMA model and mathematically can be expressed as

$$Y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{3}$$

Here p is for autoregressive AR and q is for moving average (MA).

ARIMA based model was developed by Sharif Atique et al., to predict the solar radiation for a given PV panel. The model conceived in the work is using a very simple and sophisticated statistical technique. The non-stationary time series data set was analyzed by the ACF & PACF to transform the data into stationary. The accuracy of the model was determined by MAPE which was 17.70% [39]. Mohammed H. Alsharif et al. developed a model for the estimation of daily and monthly global solar radiation using a seasonal autoregressive integrated moving average model. The non-stationary time series data was first converted into stationary data by analyzing the ACF & PACF [40]. Juan R. Trapero et al., proposed a frequency domain-based approach to estimate the short term solar irradiation. A univariate dynamic harmonic regression model was set up to forecast the global & direct normal irradiance. The model offered the self-adaptation of prediction based on the single-step recursive algorithm. The potential parameter bias problem was efficiently reduced in simultaneous estimation. The rMBE & rRMSE obtained for GHI was 0.21% & 29.66% respectively in case of DHR whereas for DNI it was 3.82% & 46.79% respectively in case of DHR [42].

7 Physical Method

7.1 Numerical Weather Prediction

It is the study of how current weather measurements are practicing to forecast weather future states. The NWP is a perfect solution of one day to many days forecast horizon. Thus, it is a valuable technique to forecast for various applications, such as solar PV forecast, wind forecast.NWP also helps in predicting the transient variation in clouds, which are supposed to be the greatest obstacle for ground solar irradiance. The NWP predicts the future conditions after the assimilation of existing findings [29].

Remco A. Verzijlbergh et al. presented a Model Output Statistics (MOS) routine based on a wide range of meteorological variables available from standard Numerical Weather Prediction output. The approach was based on a stepwise linear regression algorithm that produced a regression model with a collection of variables that better describe the forecast error that has been observed. The resulting average irradiance forecast for the first forecast day over a range of 27 stations corrected with this model reduces the relative root mean square error (rRMSE) to 22.7% compared to a rRMSE of 37.8% of uncorrected forecast [50]. Kilian Bakker et al., compared model output statistical post-processing technique for the probabilistic forecast of NWP for solar global radiation. The model output statistical post-processing method was regression method including parametric and non-parametric method. The NWP data obtained from HARMONIE-AROME (HA) from 2016 to 2018. The error matrices used for the model evaluation were RMSE, RMSE-SS (skill score), continuously ranked probability skill score (CRPSS) [52]. A numerical weather perdition model with a post-processing technique like step-wise regression and Principle component analysis (PCA) was developed by Hadrien Verbois et al., to obtain the one day ahead accurate forecast. In this model, the stepwise regression used for selecting the best explanatory variable. The large number of variables inputs in the WRF and GFS reduced by the PCA in a manner to be uncorrelated with the original levels. The model obtained 169 (W/m^2) of RMSE, 35.7% of rRMSE, 133 (W/m^2)

of MAE, 28.1% of rMAE, −14 (W/m^2) of MBE and 2.9% of rMBE for WRF-solar-PCA [51].

7.2 Empirical Model

Empirical modeling is a generic term for activities that create models by observation and experiments. The first model was developed in 1982 by Hargreaves and Samani [53]. Now a number of models have been evolved by changing the various factor such as latitude, longitude, azimuth angle, elevation angle, air particle dispersion, water vapour content, hours of sunshine, maximum temperature, minimum temperature, cloudiness index, clear sky etc. [15]. For most empirical model the key parameter measured is extraterrestrial solar radiation (H_0). The Empirical model is one of the techniques used to forecast solar irradiance future value by establishing the linear or non-linear relationship between meteorological and solar variables.

Nadjem Bailek et al. discussed 35-empirical models to obtain the accurate diffuse solar radiation by finding the appropriate regression coefficients. Three categories of the model were created based on sunshine duration, clearness index and sunshine duration & clearness index model. The regression coefficient was found for the good fitness in the model by a diffuse fraction and diffuse transmittance. The accuracy of all the models was evaluated using MPE, RMSE, U95 (uncertainty factor), R and t-statistics method (TS) and compared with the performance of eight models discussed in the literature [54]. T.R. Ayodele et al. developed an empirical model to predict the global solar radiation using the proposed regression coefficient of Angstrom-Prescott model, Garcia model and Hargreaves-Sammani model for daily & monthly time horizon. The proposed regression coefficient obtained from the fitting tool was interpolated in these models to obtain good accuracy in the results. The results showed that the Garcia model with quadratic variation performed the best for the daily average global solar radiation with 2.70 (MJ/m^2/day), 1.86 (MJ/m^2/day), 9.34% & 0.68 of RMSE, MAE, MAPE & R^2 respectively whereas 0.0909 (MJ/m^2/day), 0.0733 (MJ/m^2/day), 0.5174% & 0.9974 of RMSE, MAE, MAPE & R^2 respectively for monthly average daily global solar radiation [55].

8 Hybrid Models

These are the most commonly used method for forecasting solar irradiation with greater precision than the isolated ones. Their many factors that are not considered in the individual model by a model needed to perform more accurately [56]. The hybrid approach is about integrating two or more methods for determining the forecast. The hybrid model may combine two or more linear models, two or more non-linear models combined and linear and non-linear models combined. Different pre-processing, post-processing and optimization technologies are used to construct the hybrid models based on the literature.

The ANN & ANFIS (Artificial neuro-fuzzy inference system) model was developed by K. Ranjith Kumar et al. to estimate the solar PV power generation. The ANFIS was the combination of neural network and fuzzy inference system (FIS) that properly tune the fuzzy inference system by applying the neural learning functions. This study showed the %error has lesser in ANN forecast as compared to ANFIS [57]. Shahaboddin Shamshirband et al. developed a hybrid model using SVM and wavelet transform to predict the DSR for the Kerman, Iran. The discrete wavelet transform was used in the model to decompose the input time series and each series then applied to individual SVM model. The developed model was compared with the hybrid structure of SVM and RBF (SVM-RBF), ANN and 3rd-degree empirical model. The SVM + WT model outperforms among all the other hybrid models with MABE of 0.5757 (MJ/m2), RMSE of 0.6940 (MJ/m2) and R^2 of 0.9631 [64]. Yongqi Liu et al. developed a model to forecast the solar irradiation using deep neural network (DNN) by considering the spatial & temporal variation. The study proposed a combination of CNN and gated recurrent unit (GRU) to handle the large dimensions of spatial and temporal variations with the training loss functions. The proposed model used the convolution in the GRU network instead of using the convention multiplications for the spatiotemporal forecasting. This model achieved a mean of defined error metrics RMSE, MEA & NSE for the ConvGRU-VB was 69.5, 34.8, and 0.929 respectively [65]. Hassen Bouzgou et al. used the Extreme Learning Machine (ELM) along with WMIM for forecast GSR whereas L. Cornejo-Bueno et al used it to compare with support vector regression and Gaussian process. WMIM was used with ELM to select the appropriate input variables for training & testing phase of the model. The studies obtained MAPE of 7.4%–10.77% for 1–6 steps ahead forecast for ELM+WMIM & when it compared with support vector regression and Gaussian process, then RMSE was 60.61 (W/m2) for ELM [61]. M.A. Behrang et al. also compared six models of RBF & MLP based on different combinations of meteorological inputs to estimate the GSR [19]. The k-fold cross-validation was used by Mellit et al. to validate the capability of the neural-based forecaster. The maximum value of r obtained for sunny days was 94.14% and minimum RMSE was 32.98% whereas MAE & MBE were 2.75% & −23.25% respectively for this model [69]. Chao Huang et al. used the Jaya base algorithm to optimize the BRT parameters to predict the solar irradiation based on boosted regression trees, ANN, SVM and Least absolute shrinkage and selection operator (LASSO). This study provides nRMSE of 18.4%, 24.3%, 27.9% & 30.6% respectively for time horizon of 30 minutes, 60 minutes, 90 minutes & 120 minutes forecast [63]. Hangxia Zhou et al. developed a method to forecast the PV power using LSTM along with the attention mechanism. The attention mechanism used to observe and select the optimal forecasted output from LSTM. Two LSTM networks were used in this study, one for PV power output forecast and other for temperature forecast. The study showed that the proposed method performed better than other available models for the time horizon 7.5 minutes to 60 minutes ahead [79]. Rachit Srivastava et al. discussed 1–6 days ahead prediction of solar PV plant power output using MARS, CART, M5 model and random forest (RF) model. The performance of the random forest was better than M5, MARS & CART model whereas in cloudy days forecasting results has more errors [80].

9 Factors Influencing Solar Radiation Forecasting

There are some other factors/parameters that affect the accuracy of model forecasting directly or indirectly. The solar forecasting depends on forecast horizons, geographical condition, day/night value & normalization, testing period, climatic variability & pre-processing technique.

(a) Forecast Horizon: Time horizon issue is related to the future period for which model is forecasting. This period may be from 1 minute to several hours or days. Based on the literature, there are four categories of time horizon as follows:

   i. Very Short Term forecasting (for 1m to several mins ahead) [84].
   ii. Short term forecasting (for 1 hour to several hour/day ahead) [85].
   iii. Mid- Term Forecasting (1 month to 1 year ahead) [86].
   iv. Long Term Forecasting (1years to several years ahead) [87].

(b) Climatic variability: The variables in the input data may be systemic, endogenous and exogenous. On various combinations of input parameter different model behave differently. In most studies, ANN provides importance to meteorological and geographical variables. The increased number of irrelevant meteorological parameters degrades the performance of the model. Therefore, the appropriate parameters have to select to increases the performance of a model. M. A. Behrang et al. developed six models using ANN using a different combination of a meteorological parameter to predict solar radiation [20].

(c) **Night hour & Normalization:** The solar irradiance is not available in the night hours. Yet grid operators demanded the PV output for all time without interruption. Most of the studies conducted for the day time hours by removing the night hours. Even the time just after the sunrise & just before the sunset also removed from the data set to overcome the effects of false readings. So, the fair comparison should be necessary for the selected time frame [83].

(d) **Preprocessing Techniques:** The model's accuracy could also be enhanced by applying the pre-processing technique to input data sets. The input data sets for a particular targeted site obtained from every entity are extremely unpredictable and abnormal. The preprocessing techniques used on the data to increase or scale down the data element. Many researcher have used wavelet to transform the input series into different constituents and in the same way EMD break down the input series into different frequency [82].

10 Metrics Assessment of Solar Radiation Forecasting Technique

Various evaluation metrics used by numerous researchers to evaluate the forecasting model. These evaluation metrics are also referred to as performance metrics which allow a designer to compare the different models based on error skills, deviation, median etc. The units of different performance metrics are different and generally, the unit used for solar radiation statistical error is W/m^2.

Conventional Statistical Assessment Metrics

Correlation Coefficient: The correlation coefficient is the parameter to set a relationship between the two data sets [133]. Correlation coefficient tells us about the relation between actual and forecasted value but its limitation is not describe curvilinear relationship. it is denote by ρ. The ideal value of correlation coefficient is 1.

Normalized Error: It is denoted by nE and used for finding the outliers in the result [114] but it is not applicable for all data set.

Mean Bias Error (MBE): This metrics used to calculate the average bias in the system or model. It identifies about the underestimation in the results provided by the model [135] but it is less sensitive to small error. The positive value of MBE then the model is overestimating whereas the negative value showed the underestimation.

Mean Absolute Error (MAE): It provides the uniform error in the prediction. This is the measure of difference between two different data sets [134, 135]. The MAE only makes sense for values where divisions and ratios make sense. It doesn't make sense to calculate percentages of temperatures, for instance, so you shouldn't use the MAE to calculate the accuracy of a temperature forecast.

Standard Deviation error (SDE): It is the measure of deviation from the mean [120]. The limitation of standard deviation is it can impact by outliers and extreme values.

Root Mean Square Error (RMSE): It is the measure of largest error in the predicted data set [142]. But its drawbacks are better in terms of reflecting performance when dealing with large error values.

Mean absolute percentage error (MAPE): It is the measure of uniform prediction error in percentage. In simple terms, it is the calculation of MAE in percentage form [143]. The MAPE only makes sense for values where divisions and ratios make sense. It doesn't make sense to calculate percentages of temperatures, for instance, so you shouldn't use the MAPE to calculate the accuracy of a temperature forecast.

Mean Deviation Absolute percentage error (Md-APE): This measure is less affected by the outliers than the MAPE. The mean absolute deviation was used as a measure of dispersion in the past, but then fell into disuse. It has the disadvantage that, unlike the standard deviation (σ), it cannot be readily 'plugged' into the normal distribution formulae.

Clear Sky index: It is the ratio of measure radiation to the clear sky radiation but the limitation of clear sky is not accurate at low elevation angle.

Statistical Metrics

The MAPE, MAE & RMSE only cannot distinguish and separate the two different data sets having the same mean and variance but having a different distribution of symmetry or skewness & kurtosis. However, traditional metrics are required to measure the system but other parameters such as skewness, kurtosis, MASE may affect the real-time process.

Kolmogorov-Smirnov test integral (KSI) and OVER metrics

The Kolmogorov-Smirnov test is a non-parametric test to determine the difference between the two data sets [136]. The critical value depends on the number of points in the estimation of the time sries, measured at a confidence level of 99% [136]. The difference between the two CDFs of real and forecasted power is defined for each interval.

The KSI parameter is defined as the difference between two CDFs expressed as.

Smaller value of KSI interpreted as the real value & forecasted value have similarity. The zero CDF of two data sets represent they are similar [136].

OVER: Unlike KSI OVER is the measure of similarity on the forecast error between predicted and real value [136].

Skewness and Kurtosis

Skewness is the measure of asymmetry in the probability distribution. The advantage of skewness is that it can be either positive or negative or it may even be undefined.

Kurtosis is a measure of the magnitude of the peak of the distribution. The disadvantage is that it will not have negative or undefined form. K is the Kurtosis

Uncertainty Quantification

Renyi entropy of solar forecast error: Renyi entropy is adopted here to measure the uncertainty in solar forecasting and expressed as [137, 138]. Larger is the value of Renyi entropy more uncertainty present in the forecasted result.

Metrics for Ramp Characterization

The ability to accommodate major jumps in solar power output, frequently triggered by cloud fluctuations and severe weather events, is one of the main issues associated with incorporating a significant volume of solar power into the grid. Different time and geographic scales affect solar ramps with varying gravity levels. The forecasting of solar power overcomes the uncertainty level of power supply [139].

In case of Ramp Characterization Florita et al proposed a Swinging door algorithm. The algorithm was a simple method to represent the width of ramp with the help of threshold parameter (ε) [140].

Ramp Detection Index (RDI): It is the measure of ability of a model to forecast ramp in a very short term prediction [125].

Ramp Magnitude (RM): Ramp magnitude is the measure of normalized difference between irradiance at present time and irradiance after small time to clear sky irradiance of present time [141]. The ramp technique circuit is easy to design and its cost is low. Also, the output pulse can be transmitted over long feeder lines. However, the single ramp requires excellent characteristics regarding linearity of the ramp and time measurement. Large errors are possible when noise is superimposed on the input signal. Input filters are usually required with this type of converter.

11 Conclusion

With grid interconnection to solar parks as well as rooftop solar panels, it is critical to anticipate solar irradiance ahead of time. Solar forecasting is also essential to participate in the electricity market and to keep operations and scheduling in balance. This research looked at a variety of statistical, physical, and ensemble models. This study provides a comprehensive evaluation of recent research on solar irradiance forecasting that has been published in reputable journals. According to the findings, statistical models such as ANN are not only widely employed by researchers, but also provide superior accuracy than other solo models. Regardless of the other criteria, the forecasting accuracy worsened with the increase in the time horizon, according to the studies. In ARIMA's forecasting models, the ANN and SVM are employed most of the time to handle linear and nonlinear issues. The hybrid models, when combined with optimization and preprocessing procedures applied to the input data, improve the model's performance. In most experiments, WT and EMD were utilized as preprocessing techniques, whereas GA, PSO, and Firefly were used as optimization approaches. The study also covered the assessment metrics for each category to determine the model's performance. This in-depth look into models, their types, and error metrics not only helps you choose the right model, but it also highlights the crucial elements that can affect the design's performance.

References

[1] S. Sun, S. Wang, G. Zhang, and J. Zheng, "A decomposition-clustering-ensemble learning approach for solar radiation forecasting," *Sol.Energy*, vol. 163, pp. 189–199, Mar. 2018.

[2] A. S. Bahaj, "Means of enhancing and promoting the use of solar energy," *Renew. Energy*, vol. 27, no. 1, pp. 9-105, Sep.2002.

[3] D. I. Barnes, “Understanding pulverized coal, biomass and waste combustion - A brief overview,” *Appl. Therm. Eng.*, vol. 74, pp. 89–95, Jan. 2015.

[4] A. Setel, I. M. Gordan, and C. E. Gordan, “Use of geothermal energy to produce electricity and heating at average temperatures,” in *IET Conference Publications*, 2016, vol. 2016, no. CP711.

[5] L. Alhmoud and B. Wang, “A review of the state-of-the-art in wind-energy reliability analysis,” Elsevier Ltd, Jan. 2018.

[6] IRENA, *Renewable capacity statistics 2019, International Renewable Energy Agency (IRENA).*

[7] S. Sobri, S. Koohi-Kamali, and N. A. Rahim, “Solar photovoltaic generation forecasting methods: A review,” *Energy Conversion and Management*, vol. 156. Elsevier Ltd, pp. 459–497, 15-Jan-2018.

[8] S. Mohanty, P. K. Patra, S. S. Sahoo, and A. Mohanty, “Forecasting of solar energy with application for a growing economy like India: Survey and implication,” *Renewable and Sustainable Energy Reviews*, vol. 78. Elsevier Ltd, pp. 539–553, 2017.

[9] International Energy Agency, “Snapshot of Global Photovoltaic Markets – 2018,” *Rep. IEA PVPS T1-332018*, pp. 1–16, 2018.

[10] “Photovoltaic Power System Technology Collaboration Program Snapshot of Global PV Markets: Report IEA PVPS T1-35:2019,” 2019.

[11] IRENA, Rrnewable *capacity statistics 2010, International Renewable Energy Agency (IRENA), Abu Dhabi.* 2019.

[12] “Annual Report, Ministry of New and Renewable Energy, Government of India,” 12–13, 2017.

[13] R. K. Singh, “India’s renewable energy capacity crosses 80GW-mark,” *The Economic Times*, 2016. [online]. Available: https://economictimes.indiatimes.com/industry/energy

[14] G. Masson and M. Brunisholz, “2015 Snapshot of global photovoltaic markets,” *Iea Pvps T1-292016*, pp. 1–19, 2016.

[15] S. A. Kalogirou, “Global Photovoltaic Markets,” *McEvoy’s Handb. Photovoltaics*, pp. 1231–1235, 2016.

[16] J. G. da Silva Fonseca, T. Oozeki, T. Takashima, G. Koshimizu, Y. Uchida, and K. Ogimoto, “Use of support vector regression and numerically predicted cloudiness to forecast power output of a photovoltaic power plant in Kitakyushu, Japan,” *Prog. Photovoltaics Res. Appl.*, vol. 20, no. 7, pp. 874– 882, Nov. 2012.

[17] V. Z. Antonopoulos, D. M. Papamichail, V. G. Aschonitis, and A. V. Antonopoulos, "Solar radiation estimation methods using ANN and empirical models," *Comput. Electron. Agric.*, vol. 160, pp. 160– 167, May 2019.

[18] A. Khosravi, R. O. Nunes, M. E. H. Assad, and L. Machado, "Comparison of artificial intelligence methods in estimation of daily global solar radiation," *J. Clean. Prod.*, vol. 194, pp. 342–358, Sep. 2018.

[19] N. Premalatha and A. Valan Arasu, "Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms," *J. Appl. Res. Technol.*, vol. 14, no. 3, pp. 206–214, Jun. 2016.

[20] M. A. Behrang, E. Assareh, A. Ghanbarzadeh, and A. R. Noghrehabadi, "The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data," *Sol. Energy*, vol. 84, no. 8, pp. 1468–1480, Aug. 2010.

[21] A. Fouilloy *et al.*, "Solar irradiation prediction with machine learning: Forecasting models selection method depending on weather variability," *Energy*, vol. 165, pp. 620–629, Dec. 2018.

[22] L. Mazorra-Aguiar and F. Díaz, "Solar Radiation Forecasting with Statistical Models," 2018, pp. 171–200.

[23] Y. Liu, S. Shimada, J. Yoshino, T. Kobayashi, Y. Miwa, and K. Furuta, "Ensemble forecasting of solar irradiance by applying a mesoscale meteorological model," *Sol. Energy*, vol. 136, pp. 597–605, Oct. 2016.

[24] A. Ghanbarzadeh, A. R. Noghrehabadi, E. Assareh, and M. A. Behrang, "Solar radiation forecasting based on meteorological data using artificial neural networks," in *IEEE International Conference on Industrial Informatics (INDIN)*, 2009, pp. 227–231.

[25] M. Bou-Rabee, S. A. Sulaiman, M. S. Saleh, and S. Marafi, "Using artificial neural networks to estimate solar radiation in Kuwait," *Renewable and Sustainable Energy Reviews*, vol. 72. Elsevier Ltd, pp. 434– 438, 2017.

[26] Y. Yu, J. Cao, and J. Zhu, "An LSTM Short-Term Solar Irradiance Forecasting under Complicated Weather Conditions," *IEEE Access*, vol. 7, pp. 145651–145666, 2019.

[27] N. Kumar, U. K. Sinha, S. P. Sharma, and Y. K. Nayak, "Prediction of daily global solar radiation using Neural Networks with improved gain factors and RBF Networks," *Int. J. Renew. Energy Res.*, vol. 7, no. 3, pp. 1235–1244, 2017.

[28] G. Notton, C. Voyant, A. Fouilloy, J. L. Duchaud, and M. L. Nivet, “Some applications of ANN to solar radiation estimation and forecasting for energy applications,” *Appl. Sci.*, vol. 9, no. 1, pp. 1–21, 2019.

[29] F. Rodríguez, A. Fleetwood, A. Galarza, and L. Fontán, “Predicting solar energy generation through artificial neural networks using weather forecasts for microgrid control,” *Renew. Energy*, vol. 126, pp. 855–864, Oct. 2018.

[30] B. Jahani and B. Mohammadi, “A comparison between the application of empirical and ANN methods for estimation of daily global solar radiation in Iran,” *Theor. Appl. Climatol.*, vol. 137, no. 1–2, pp. 1257–1269, Jul. 2019.

[31] B. Sivaneasan, C. Y. Yu, and K. P. Goh, “Solar Forecasting using ANN with Fuzzy Logic Pre- processing,” in *Energy Procedia*, 2017, vol. 143, pp. 727–732.

[32] C. R. Chen and U. T. Kartini, “k-nearest neighbor neural network models for very short-term global solar irradiance forecasting based on meteorological data,” *Energies*, vol. 10, no. 2, 2017.

[33] Z. Li, S. Rahman, R. Vega, and B. Dong, “A Hierarchical Approach Using Machine Learning Methods in Solar Photovoltaic Energy Production Forecasting,” *Energies*, vol. 9, no. 1, p. 55, Jan. 2016.

[34] M. Vakili, S. R. Sabbagh-Yazdi, S. Khosrojerdi, and K. Kalhor, “Evaluating the effect of particulate matter pollution on estimation of daily global solar radiation using artificial neural network modeling based on meteorological data,” *J. Clean. Prod.*, vol. 141, pp. 1275–1285, Jan. 2017.

[35] R S. Alam, S. C. Kaushik, and S. N. Garg, “Assessment of diffuse solar energy under general sky condition using artificial neural network,” *Appl. Energy*, vol. 86, no. 4, pp. 554–564, 2009.

[36] P. M. Reilly, “Probability and Statistics for Engineers and Scientists,” *Can. J. Stat.*, vol. 6, no. 2, pp. 283-284, 1978.

[37] *CSEE J. Power Energy Syst.*, vol. 1, no. 4, pp. 38–46, Jan. 2016.

[38] G.U.Yule, “On the Time –Correlation Problem,witj Especial Reference to the Variate-Difference Correlation Method,” J.R.Stat.Soc., vol. 84, no. 4, p. 497, Jul. 1921.

[39] “Economteric Modeler App Overview-MATLAB & Simulink-MathWorks India,”[Online]. Available: http://in.mathworks.com/help/econ/econometric-modeler-overview.html.

[40] S. Atique,S. Noureen,V. Roy,V. Subburaj,S. Bayne,and J. MacFie, "Forecasting of total daily solar energy generation using ARIMA:A case study,"in 2019 IEEE 9th Annual computing and Communication Workshop and Conference, CCWC 2019, 2019, pp. 114–119.

[41] M. Alsharif, M. Younes, and J. Kim, "Time Series ARIMA Model for Prediction of Daily and Monthly Average Global Solar Radiation: The Case Study of Seoul, South Korea," Symmetry (Basel)., vol. 11, no. 2, p. 240, Feb. 2019.

[42] J. R. Trapero, N. Kourentzes, and A. Martin, "Short-term solar irradiation forecasting based on dynamic harmonic regression," *Energy*, vol. 84, pp. 289–295, May 2015.

[43] P. F. Jiménez-Pérez and L. Mora-López, "Modeling and forecasting hourly global solar radiation using clustering and classification techniques," *Sol. Energy*, vol. 135, pp. 682–691, Oct. 2016.

[44] J. Shi, W. J. Lee, Y. Liu, Y. Yang, and P. Wang, "Forecasting power output of photovoltaic systems based on weather classification and support vector machines," in *IEEE Transactions on Industry Applications*, 2012, vol. 48, no. 3, pp. 1064–1069.

[45] H. S. Jang, K. Y. Bae, H. S. Park, and D. K. Sung, "Solar Power Prediction Based on Satellite Images and Support Vector Machine," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 1255–1263, Jul. 2016.

[46] V. Eniola, T. Suriwong, C. Sirisamphanwong, and K. Ungchittrakool, "Hour-ahead forecasting of photovoltaic power output based on Hidden Markov Model and Genetic Algorithm," *Int. J. Renew. Energy Res.*, vol. 9, no. 2, pp. 933–943, 2019.

[47] S. Bhardwaj *et al.*, "Estimation of solar radiation using a combination of Hidden Markov Model and generalized Fuzzy model," *Sol. Energy*, vol. 93, pp. 43–54, Jul. 2013.

[48] A. Wibun and P. Chaiwiwatworakul, "An Estimation of Thailand's Hourly Solar Radiation Using Markov Transition Matrix Method," *Appl. Mech. Mater.*, vol. 839, pp. 29–33, Jun. 2016.

[49] S. Li, H. Ma, and W. Li, "Typical solar radiation year construction using k-means clustering and discrete-time Markov chain," *Appl. Energy*, vol. 205, pp. 720–731, Nov. 2017.

[50] R. A. Verzijlbergh, P. W. Heijnen, S. R. de Roode, A. Los, and H. J. J. Jonker, "Improved model output statistics of numerical weather prediction based irradiance forecasts for solar power applications," *Sol. Energy*, vol. 118, pp. 634–645, Aug. 2015.

[51] H. Verbois, R. Huva, A. Rusydi, and W. Walsh, "Solar irradiance forecasting in the tropics using numerical weather prediction and statistical learning," *Sol. Energy*, vol. 162, pp. 265–277, Mar. 2018.
[52] K. Bakker, K. Whan, W. Knap, and M. Schmeits, "Comparison of statistical post-processing methods for probabilistic NWP forecasts of solar radiation," *Sol. Energy*, vol. 191, pp. 138–150, Oct. 2019.
[53] G. H. Hargreaves and Z. A. Samani, "Estimating Potential Evapotranspiration," *J. Irrig. Drain. Div.*, vol. 108, no. 3, pp. 225–230, 1982.
[54] E. Quansah *et al.*, "Empirical Models for Estimating Global Solar Radiation over the Ashanti Region of Ghana," *J. Sol. Energy*, vol. 2014, pp. 1–6, Jan. 2014.
[55] T. R. Ayodele and A. S. O. Ogunjuyigbe, "Performance assessment of empirical models for prediction of daily and monthly average global solar radiation: the case study of Ibadan, Nigeria," *Int. J. Ambient Energy*, vol. 38, no. 8, pp. 803–813, Nov. 2017.
[56] Y. Xie, "Values and limitations of statistical models," *Res. Soc. Stratif. Mobil.*, vol. 29, no. 3, pp. 343–349, 2011.
[57] K. R. Kumar and M. S. Kalavathi, "Artificial intelligence based forecast models for predicting solar power generation," in *Materials Today: Proceedings*, 2018, vol. 5, no. 1, pp. 796–802.
[58] L. M. Aguiar, B. Pereira, P. Lauret, F. Díaz, and M. David, "Combining solar irradiance measurements, satellite-derived data and a numerical weather prediction model to improve intra-day solar forecasting," *Renew. Energy*, vol. 97, pp. 599–610, Nov. 2016.
[59] S. Shamshirband et al., "Estimating the diffuse solar radiation using a coupled support vector machine- wavelet transform model," *Renewable and Sustainable Energy Reviews*, vol. 56. Elsevier Ltd, pp. 428–435, 01-Apr-2016.
[60] N. Dong, J. F. Chang, A. G. Wu, and Z. K. Gao, "A novel convolutional neural network framework based solar irradiance prediction method," *Int. J. Electr. Power Energy Syst.*, vol. 114, Jan. 2020.
[61] H. Bouzgou and C. A. Gueymard, "Fast short-term global solar irradiance forecasting with wrapper mutual information," *Renew. Energy*, vol. 133, pp. 1055–1065, Apr. 2019.
[62] S. Ghimire, R. C. Deo, N. J. Downs, and N. Raj, "Global solar radiation prediction by ANN integrated with European Centre for medium range weather forecast fields in solar rich cities of Queensland Australia," *J. Clean. Prod.*, vol. 216, pp. 288–310, Apr. 2019.

[63] C. Huang, L. Wang, and L. L. Lai, "Data-Driven Short-Term Solar Irradiance Forecasting Based on Information of Neighboring Sites," *IEEE Trans. Ind. Electron.*, vol. 66, no. 12, pp. 9918–9927, Dec. 2019.

[64] X. Huang, J. Shi, B. Gao, Y. Tai, Z. Chen, and J. Zhang, "Forecasting Hourly Solar Irradiance Using Hybrid Wavelet Transformation and Elman Model in Smart Grid," *IEEE Access*, vol. 7, pp. 139909–139923, Sep. 2019.

[65] Y. Liu et al., "Ensemble spatiotemporal forecasting of solar irradiation using variational Bayesian convolutional gate recurrent unit network," *Appl. Energy*, vol. 253, Nov. 2019.

[66] M. Q. Raza, M. Nadarajah, J. Li, K. Y. Lee, and H. B. Gooi, "An Ensemble Framework For Day-Ahead Forecast of PV Output in Smart Grids," *IEEE Trans. Ind. Informatics*, 2018.

[67] L. Benali, G. Notton, A. Fouilloy, C. Voyant, and R. Dizene, "Solar radiation forecasting using artificial neural network and random forest methods: Application to normal beam, horizontal diffuse and global components," *Renew. Energy*, vol. 132, pp. 871–884, Mar. 2019.

[68] A. Heydari, D. Astiaso Garcia, F. Keynia, F. Bisegna, and L. De Santoli, "A novel composite neural network based method for wind and solar power forecasting in microgrids," *Appl. Energy*, vol. 251, Oct. 2019.

[69] A. Mellit and A. M. Pavan, "A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy," *Sol. Energy*, vol. 84, no. 5, pp. 807–821, May 2010.

[70] S. X. Chen, H. B. Gooi, and M. Q. Wang, "Solar radiation forecast based on fuzzy logic and neural networks," *Renew. Energy*, vol. 60, pp. 195–201, Dec. 2013.

[71] H. Lan, C. Zhang, Y. Y. Hong, Y. He, and S. Wen, "Day-ahead spatiotemporal solar irradiation forecasting using frequency-based hybrid principal component analysis and neural network," *Appl. Energy*, vol. 247, pp. 389–402, Aug. 2019.

[72] F. Wang, Z. Mi, S. Su, and H. Zhao, "Short-Term Solar Irradiance Forecasting Model Based on Artificial Neural Network Using Statistical Feature Parameters," *Energies*, vol. 5, no. 5, pp. 1355–1370, May 2012.

[73] D. Liu and K. Sun, "Random forest solar power forecast based on classification optimization," *Energy*, vol. 187, Nov. 2019.

[74] W. Zhang, H. Dang, and R. Simoes, "A new solar power output prediction based on hybrid forecast engine and decomposition model," *ISA Trans.*, vol. 81, pp. 105–120, Oct. 2018.
[75] Y. Dong and H. Jiang, "Global Solar Radiation Forecasting Using Square Root Regularization-Based Ensemble," *Math. Probl. Eng.*, vol. 2019, 2019.
[76] James Mubiru "Using Artificial Neural Networks to predict direct solar irradiation," *Hindawai Publishing Corporation.,* Volume 2011, Article ID 142054, 6 pages.
[77] A. T. Eseye, J. Zhang, and D. Zheng, "Short-term photovoltaic solar power forecasting using a hybrid wavelet-PSO-SVM model based on SCADA and Meteorlogical information," Renew. Energy, vol. 118, pp. 357-367, Apr. 2018.
[78] W. VanDeventer et al., "Short-term PV power forecasting using hybrid GASVM technique," Renew. Energy, vol. 140, pp. 367–379, Sep. 2019.
[79] A. Shadab, S. Said, and S. Ahmad, "Box-Jenkins multiplicative ARIMA modeling for prediction of solar radiation: a case study," *Int. J. Energy Water Resour.*, Sep. 2019.
[80] J. Dong et al., "Novel stochastic methods to predict short-term solar radiation and photovoltaic power," *Rener. Energy*, vol. 145, pp. 333–346, Jan. 2020.
[81] Sanjari et al., "Probabilistic forecast of PV power generation based on higher order markov chain", *IEEE Access,* vol. 32, pp. 2942–2952, July 2017.
[82] S. Monjoly, M. Andre, R. Calif, and T. Soubdhan, "Hourly forecasting of global solar radiation based on multiscale decomposition methods: A hybrid approach," Energy, vol. 119, pp. 228–298, 2017.
[83] M. Paulescu and E. Paulscu, "Short term forecasting of solar irradiance," Renew. Energy, vol. 143, pp. 985–994, Dec. 2019.
[84] C. Voyant and G. Notton, "Solar irradiation now casting by stochastic persistence:A new parsimonious simple and efficient forecasting tool, *Renew. Sustain Energy Rev.*, vol. 92, 343–352, Sep. 2018.
[85] F. Jiang et al., "Artificial intelligence in healthcare: past, present and future, "Stroke vasa.Neural., vol. 2, no. 4, pp. 230–243, 2017.
[86] H. Bowang, J. Ning Xiong, & C. Yinzhao, "The mid-term forecast method of solar radiation index," Chinese Astron. Astrophys., vol. 39, no. 2, pp. 198–211, Apr. 2015.

[87] D. Yang, "A Guideline to solar forecasting research practice: Reproducible, operational, probabilistic or physically based, ensemble and skills (ROPES)," *Journal of Renewable and Sustainable Energy*, vol. 11, no. 2. American Institute of physics Inc., 01-Mar-2019.

[88] R. J. Hyndman and A. B. Koehler, "Another look at measure of forecast accuracy," Int. J. Forecast., vol. 22, no. 4, pp. 679–688, Oct. 2006.

[89] J. Zhang, B. M. Hodge, A. Florita, S. Lu, H. F. Hamann, and V. Banunarayanan, "Metrics for evaluating the accuracy of solar power forecasting," in 3rd International Workshop on Integrations of Solar Power into Power Systems, 2013, vol. no. 17436, October, pp. 1–10.

[90] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean Square error (RMSE) in assessing average model performance," Clim.Res., vol. 30, no. 1, pp. 79–82, 2005.

[91] Mills. A., and R. Wiser. (2010), "Implications of Wide-Area Geographic Diversity for Short-Term Variability of Solar Power", *Report on Environmental Energy Technologies Division, Ernest Orlando Lawrence Berkeley National Laboratory.*

[92] A. Florita, B. M. Hodge, and K. Orwig, "Identifying wind and solar ramping events," in IEEE Green Technologies Conference, 2013, no. January, pp. 147–152.

[93] Y. Chu, H. T. C. Pedro, M. Li, and C. F. M. Coimbra, "Real-time forecasting of solar irradiance ramps with smart image processing," Sol. Energy. vol. 114, pp. 91–104, Apr. 2015.

[94] P. Lauret, C. Voyant, T. Soubdhan, M. David, and P. Poggi, "A benchmarking of machine learning techniques for solar radiation forecasting in an insular context," Sol. Energy, vol. 112, pp. 446–457.

[95] C. J. Wilmot and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," Clim. Res., vol. 30, no. 1, pp. 79–82, 2005.

[96] C. Feng, M. Cui, B. Hodge, S. Lu, H. Hamann, and J. Zhang, "Unsupervised Clustering-Based Short-Term Solar Forecasting," *IEEE Trans. Sustain. Energy*, 2018.

[97] N. Basurto, A. Arroyo, R. Vega, H. Quintian, J.L. Calvo-Rolle, and A. Herrero, "A Hybrid Intelligent System to forecast solar energy production," Comput. Electr. Eng., vol. 78, pp. 373–387, Sep. 2019.

[98] H. Zhou, Y. Zhang, L. Yang, Q. Liu, K. Yan, and Y. Du, "Short-Term photovoltaic power forecasting based on long short term memory

neural network and attention mechanisum," IEEE Access, vol. 7, pp. 78063–78074, 2019.

[99] V. Kushwaha and N.M. Pindoriya , "A SARIMA-RVFL hybrid model assisted by wavelet decomposition for very short-term solar PV power generation forecast," *Renew. Energy*, vol. 140, pp. 124–139, Sep. 2019.

[100] Y. Liu, Y. Zhou, Y. Chen, D. Wang, Y. Wang, and Y. Zhu, "Comparison of support vector machine and copula-based non-linear quantile regression for estimating the daily diffuse solar radiation: A case study in China," Renew. Energy, vol. 146, pp. 1101–1112, Feb. 2020.

[101] Fonseca JGDS, Oozeki T, Takashima T, Koshimizo G, Uchida Y, Ogimoto K, Photovoltaic power production forecast with support vector regression: a study on the forecast horizon. In: 2011 37th IEEE photovoltaic specialists conference: 2011, pp. 2579–2583.

[102] Montteitro 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.

[103] Mohammadi, K. Shamshirband, S. Tong, C. W., Arif, M., Petkovic, D., & Ch, S. A new hybrid support vector machine-wavelet transform approach for estimation of horizontal global solar radiation. Energy Conversion and Management, 92, 162–171. http://doi.org/10.1016/j.enconman.2014.12.050

[104] Prasad, R., Ali, M., Kwan, P., & Khan, H. Designing a multistage multivariate empirical model decomposition coupled with ant colony optimization and random forecast model to forecast monthly solar radiation. Applied energy, 236, 778–792. http://doi.org/10.1016/j.apenergy.2018.12.034.

[105] Diagne, M., David, M., Boland, J., Schmutz, N., & Lauret, P.(2014). Post-processing of solar irradiance forecasts from WRF model at Reunion Island. Solar Energy, 105, 99–108. https://doi.org/10.1016/j.solener.2014.03.016.

[106] Hussain, S., & AlAlili, A. Online sequential Learning of Neural Network in Solar Radiation Modeling Using Hybrid Bayesian Hierarchical Approach. Journal of Solar Energy Engineering, 138(6), 061012. https://doi.org/10.1115/1.4034907.

[107] Che, Y., Chen, L., Zheng, J., Yuan, L., & Xiao, F. A Novel Hybrid Model of WRF and Clearness Index-Based Kalman Filter for Day-Ahead Solar Radiation Forecasting. Applied Sciences, 9(19), 3967. http://doi.org/10.3390/app9193967.

[108] Dong, Z., Yang, D., Reindl, T., & Walsh, W.M. A novel hybrid approach based on self-organizing maps, support vector regression and particle swarm optimization to forecast solar irradiance. Energy, 82, 570–577. http://doi.org/10.1016/j.energy.2015.01.066

[109] Hady, M. F. A., Schwenker, F., & Palm, G. Semi-supervised learning for tree-structured ensemble of RBF networks with co-training. Neural Network, 23(4), 497–509. http://doi.org/10.1016/j.neunet.2009.09.001

[110] Wu, Y., & Wang, J. A novel hybrid model based on artificial neural network for solar radiation prediction. Renewable Energy, 89, 268–284. https://doi.org/10.1016/j.renene.2015.11.070.

[111] Alomari MH, Youuis Ola, Hayainesh SMA: A predictive mode 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.

[112] Reikard G. Predicting solar radiation at high resolution: a comparison of time series forecast. Sol Energy 2009; 83(3):3-12-9.

[113] Ibrahim, I.A., & Khatib, T. A novel hybrid model for hourly global solar radiation prediction using random forests techniques and firefly algorithm. Energy Conversion and Management, 138, 413–425. http://doi.org/10.1016/j.enconman.2017.02.006

[114] Halabi, L.M., Mekhilef, S., & Hossain, M. Performance evaluation of hybrid adaptive neuro-fuzzy inference system models for predicting monthly global solar radiation. Applied energy, 213, 247–261. http://doi.org/10.1016/j.apenergy.2018.01.035

[115] Jovanovic, R., Pomares, L.M., Mohieldeen, Y.E., Perez-Astudillo, D., & Bachour, D. An evolutionary method for creating ensembles with adaptive size neural networks for predicting hourly solar irradiance. In 2017 International Joint Conference on Neural Networks(IJCNN) (pp. 1962–1967).IEEE. DOI: 10.1109/IJCNN.2017.796609

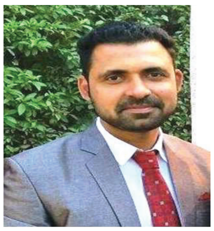
[116] K. Benmouiza and A. Cheknane, “Forecasting hourly global solar radiation using hybrid k-means and nonlinear autoregressive neural network models,” *Energy Convers. Manag.*, vol. 75, pp. 561–569, Nov. 2013.

[117] K. Wang, X. Qi, and H. Liu, “A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network,” *Appl. Energy*, vol. 251, Oct. 201934

Biographies

Anuj Gupta recieved the B.Tech in Electronics and Communication Engineering from Kurukshetra University, M.Tech in Electronics and Communication Engineering from Kurukshetra University, Kurukshetra. Presently he is Assistant Professor in EEE Deptt. at Asia Pacific Institute of Information Technology, Panipat and pursuing Ph.D. in the area of solar irradiance forecasting from Electronics and Communication Engineering Department, Maharishi Markandeshwar University, Mullana, Ambala, India. His reasearch area are deregulated electricity market, solar irradiance forecasting. He has more than five years teaching and research experience.

Kapil Gupta received his B.E. (HONS) degree in Electronics & Communication Engineering in 2003 from Rajasthan University and M.E. (HONS) degree in Digital Communication from MBM Engineering College Jodhpur, Rajasthan in 2008. He earned Ph.D. degree in 2013 from MITS University, Rajasthan. Presentlyhe is Associate Professor in the Department of Electronics and Communication Engineering, M.M.E.C, Maharishi Markandeshwar (Deemed to be University) Mullana, Ambala and has more than 15 years of experience in teaching. His research interests are in . . . . . . , Wireless Sensor Networks, Wireless Communication, Diversity Techniques and Error Correction Coding.



Sumit Saroha is currently working as Assistant Professor in the Department of Electrical Engineering, Guru Jambheshwar University of Science & Technology, Hisar, India. He received Ph.D. in the area of forecasting issues in present day power systems. His research interests are transformer design, electricity markets, electricity forecasting, neural networks, wavelet transform, fractional order systems and Multi agent systems.