2
Preliminaries, Predictive Models and Related Work

2.1 SOLAR ENERGY: BASICS

Renewable energy sources are capable of developing sustainable power plants as these energy sources are abundant, eco-friendly and provide energy with negligible emissions of air pollutant and greenhouse gases. Sun and wind are most potential known renewable energy sources. The part of sun radiations travels in a straight line from the sun to the earth surface and falls at the right angle to the plane of incidence is known as direct normal irradiance (DNI). DNI is used as input to concentrated solar power (CSP) based technologies. Some portion of solar radiation gets scattered, before reaching on the earth's horizontal surface, is termed as diffuse horizontal irradiance (DHI). Total amount of radiations received by a horizontal plane on earth is known as Global Horizontal Irradiance (GHI). On any surface, global irradiance is a sum of direct and diffuse irradiance and is given by $GHI = DNI \times COS(SZA) + DHI$, where SZA is solar zenith angle. GHI and DNI components of solar radiation are taken into account for flat plate solar collectors design and concentrating solar collectors design respectively. The solar photovoltaic (SPV) cells of flat plate solar collector directly convert solar radiation into DC electricity. The electricity produced through PV can be used directly or it can be stored in the battery for further use. Solar PV is a cost-effective mode of electrification and is economical in remote areas where electricity supply from the grid is costly to meet the user demand [Bijarniya *et al.*, 2016]. Hence contribution of solar PV is more than that of CSP in renewable energy generation. This is one of the reasons of rapid growth of PV technology over the past decade.

The intermittency of energy generation output is higher for solar PV compared to CSP. CSP plants can be considered to have better inertia against cloud produced transients and provide less intermittent output because of some storage inherent in all CSP systems [Sayeef *et al.*, 2012]. Solar PV is highly influenced by the cloud and produces highly variable output and needs to be appropriately predicted and integrated into the grid to provide stable and quality of generation. That is the reason irradiance forecasting becomes extremely important for solar PV generation.

The total power output from solar PV is proportional to the solar irradiance at the earth's surface and given by multiplication of the solar irradiation, the collector area (A), and the efficiency of the PV module (η) [Luque and Hegedus, 2011]. A and η are the characteristics of PV and differ for various modules. Therefore, in the thesis, solar irradiation has been used as the primary variable for SPV output power forecasting.

2.2 INTRODUCTION

As mentioned in the previous chapter, the solar radiation forecasting can be categorized into two broad categories; physical models and data-driven models [Ren *et al.*, 2015]. The physical models are based on science of physics and depend on interactions between the physical state and dynamical motion of the Sun. These physical models consist of three submodels (i) Numerical Weather Prediction (NWP), (ii) Ground-based Total Sky Imagery (TSI) and (iii) Satellite Imagery. The data-driven models use past recorded data to predict future value based on the previous pattern already present in the data. The data-driven models consist of mainly two sub-models (i) autoregressive models and (ii) machine learning based models. A comparison between physical

models and data-driven/statistical models is listed in Table 2.1.

Table 2.1 : Data-driven and physical models

Technique	Strengths	Limitations
Data-driven/Statistical	1. Computation is much faster with real time data	1. Requires historical data
	2. Can model non linear patter in the data	2. Non- transparent and confined
	3. Often more accurate	3. Difficult to generalize
Deterministic	1. Rely upon the science of physics	1. Difficult to model real science
	2. Transparent and training data is not required	2. Data unavailability of building properties
	3. Easy to generalize	3. Not very accurate

The selection of the model mainly depends on the prediction horizon because forecast accuracy varies for different forecast horizons. Any combination of physical and statistical models constitute a hybrid model. The main motivation behind construction of hybrid framework is to combine various models with a unique feature to address the limitations of the individual models and improve the prediction accuracy. Hybrid models have ability to model linear as well as nonlinear patterns present in data in the more convincing way. Table 2.2 shows the time horizon, interval and corresponding suitable applications of various forecasting methods. In addition, various forecasting approaches and their corresponding examples and suitability are presented in Table 2.3. Various solar energy generation techniques and their forecast requirements are listed in Table 2.4. Stakeholders and operators and their various forecast requirements are listed in Table 2.5. This chapter presents a review of different statistical models and highlights the strengths and weaknesses of these statistical models which can be useful in understanding the extent under which a model should be preferred. The chapter also elaborates various hybrid methods and the motivation behind their ensembling.

2.2.1 Objectives of the review

Reviews on solar resource forecasting deliver detailed accounts of well established time series forecasting methods and related work that has been done by various researchers in recent past. Inman et al. and Wan et al. reviewed resource forecasting methods and their configurations for multiple time horizons [Inman *et al.*, 2013; Wan *et al.*, 2015]. In the same year, Diagne et al also reviewed various irradiance forecasting methods [Diagne *et al.*, 2013]. The review focuses on the classification of the models and their suitability for various time horizons. Kashyap et al. and Yadav and Chandel presented reviews on solar irradiance forecasting based on Artificial Intelligence (AI) techniques [Kashyap *et al.*, 2015; Yadav and Chandel, 2014]. Various ensemble methods were highlighted by [Ren *et al.*, 2015]. A recent review by Voyant et al., Li et al. and Jimenez et al. discussed various machine learning methods for solar irradiance forecasting [Voyant *et al.*, 2017; Li *et al.*, 2016; Jiménez-Pérez and Mora-López, 2016]. All these reviews offer broader understanding of forecasting methods on various temporal and spatial horizons. This chapter aims to explore the existing time series forecasting techniques along with their advantages, challenges and suitability. Review on hybrid models is also presented since hybrid models have significant contribution in the field. We provide a basis of comparison among various available techniques

Table 2.2 : Relation between forecast horizons, forecast interval and related applications

Time horizon	Interval	Some applications
Intra-hour	< 2 h	Short term ramps, variability related to operations
Intra-day	$1 - 6$ h	Load following
Day ahead	$1 - 3$ days	Unit commitment, transmission scheduling, day ahead market

Table 2.3 : Categorization of solar irradiance predictive model based on various approaches

Approaches	Examples	Suitability
Physical	(i) NWP	(i) NWP is more accurate for day-ahead
	(ii) TSI	(ii) TSI is more accurate for intra-hour
	(iii) Satellite imagery	(iii) Satellite imagery is more accurate for intra-hour and intra-day.
Statistical/Data driven	(i) Regressive models (AR, MA, ARMA, ARIMA)	(i) Regressive models are more accurate for intra hour
	(ii) Machine learning (KNN, SVM, ANN)	(ii) Machine learning based models are more accurate for intra-hour and intra-day
Hybrid	Mixture of two or more	Suitable for all time horizons

Table 2.4 : Solar generation techniques and their forecast requirements

Solar generation techniques	Intra-hour	Intra-day	Day ahead
Large concentrated solar thermal	√	X	X
distributed PV	√	√	X
Large grid connected PV	√	√	X
Small non concentrated solar thermal	√	√	X

mentioned here. In nutshell, the objectives of this chapter are:

1. To present an overview in order to facilitate configuration and selection of appropriate forecast model according to needs, applications and horizons of prediction.

2. To discuss various hybrid models along with their motivation of ensembling.

3. To present a comparative analysis of various forecasting methods that include both qualitative and quantitative aspects of these techniques.

The chapter is organized as follow. In Section 2.3, statistical and learning based forecasting approaches are reviewed. In this section, strength and weakness of models are discusses along with their configurations. Section 2.4, concludes the chapter.

2.3 DATA-DRIVEN MODELS

Statistical models, more specifically regressive models, have been used for time series forecasting in the field of renewable energy for many years. Time series models use historical observed values for prediction of solar radiation. The traditional regressive time series models such as autoregressive (AR), moving average (MA) and autoregressive moving average (ARMA) are useful when the underlying data series is stationary [Tsay, 2005]. Modelling process flow diagram of regression based time series models is shown in Figure 2.1.

Table 2.5 : Solar operators and stakeholders and their forecast horizons

Stakeholders	Intra-hour	Intra-day	Day ahead
Energy market	√	√	√
T & D planning	√	√	√
Operations	√	√	X
Financial planning	X	X	√

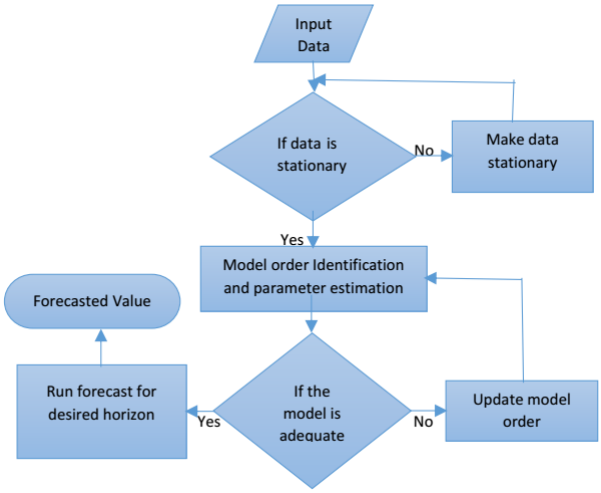

Figure 2.1 : Modelling process flow diagram of regressive based time series models

However, solar radiation data does not satisfy the stationarity condition in most of the cases. This leads to the second approach of using learning based techniques, such as artificial neural networks (ANN), support vector machine (SVM), K-nearest neighbor (KNN), decision tree (DT), Markov chain etc. Modelling process flow diagram of machine learning based predictors is shown in Figure 2.2. In the literature, many predictor models have been proposed by the

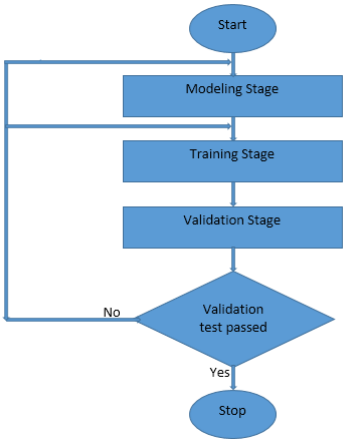

Figure 2.2 : Modelling process flow diagram of machine learning based model

researchers for forecasting of solar resources. Here, we present eight most frequently used regressive and machine learning based predictive models in the field of solar forecasting. The strengths and limitations of each of these predictors, in brief, are presented in Table 2.6.

2.3.1 Autoregressive Moving Average (ARMA)

One of the most extensively used classical time series models is the ARMA model. ARMA is popular because of its capability to model useful statistical properties and the adoption of the well-known Box-Jenkins methodology [Box *et al.*, 2015]. ARMA is flexible as it can represent several models by simply changing its order suitably. Model is more effective and competent when time

series posses an underlying linear auto correlation structure. The mathematical model is given as:

$$x_t = \phi_0 + \sum_{i=1}^{p} \phi_i x_{t-i} + a_t - \sum_{i=1}^{q} \theta_i a_{t-i}, \tag{2.1}$$

where, x_t is forecasted value at time t, a_t is a white noise, ϕ_i for $i = 1, 2, ..., p$ are AR parameters and θ_i for $i = 1, 2, ..., q$ are MA parameters. p and q are the orders of AR and MA processes respectively.

The basic foundation of time series analysis is that it must obey stationarity condition. A time series is weak stationary if the mean of the time series is constant and the covariance with its lagged series is time invariant [Tsay, 2005]. Mathematically, a time series X_t is weak stationary if $E(X_t) = \mu$, which does not depend on t and $Cov(X_t, X_{t-l}) = \gamma_l$, which only depends on l. In practice, the weak stationarity implies that the time plot of data fluctuate with constant variation around a fixed label. Another approach through which we can determine weather a time series is stationary or not is by the use of spectral analysis of time series. The frequency spectrum of a time series is basically the frequency component (spectral component) of that time series. The frequency spectrum of a time series shows what frequency exist in the time series. Time series whose frequency content do not change with time, is called stationary. However, stationarity test using Augmented Dickey–Fuller (ADF) test confirms that solar radiation series usually found to be non-stationary. Thus, to convert solar time series into weak stationary series, a phase of detrending is needed, which is usually achieved by differencing methods with the assumption that the d_th difference of the time series may be considered as a stationary process. After detrending, stationarity of the time series can be checked through ADF test. Order of the ARMA model is determined by plotting autocorrelation function (ACF) and Partial ACF (PACF). Decreasing exponential or alternate in sign or decreasing sinusoidal patterns ensure the time series to be stationary and the number of its value greater than the significance level determines the number of MA and AR parameters respectively. Thus ACF and PACF are used to determine preliminary order of ARMA. Tsay and Tiao [Tsay, 2005] proposed a new approach called extended autocorrelation function (EACF) to identify the preliminary order of ARMA process. Finally, ACF and PACF of the residual series and Akaike information criterion (AIC) of different orders are checked to obtain the optimal ARMA model [Akaike, 1998].

Reikard [Reikard, 2009] compared ARMA models with various other nonlinear models, including neural networks and hybrid models, at resolutions of 5, 15, 30 and 60 minutes using GHI, and concluded that, in nearly all the cases, performance of ARMA model was better. Perdomo et al. used daily solar radiation measured data obtained from Bogotá, Columbia, between 2003 to 2009 for predicting daily mean GHI using linear time series model [Perdomo *et al.*, 2010]. ARMA model recently found many applications in the construction of hybrid systems [Ji and Chee, 2011; Voyant *et al.*, 2012; Bouzerdoum *et al.*, 2013; David *et al.*, 2016]. Hybrid of ARMA and time delay neural networks (TDNN) along with several detrending models for hourly solar radiation prediction was introduced by [Ji and Chee, 2011]. The data was from Nanyang Technological University, Singapore with sampling interval of 10 minutes frequency. Bouzerdoum et al. proposed hybrid of seasonal ARIMA and SVM for short-term power forecasting of a small-scale grid-connected photovoltaic plant [Bouzerdoum *et al.*, 2013]. Solar irradiance forecasting with recursive ARMA and GARCH models for very short-term solar forecasting, from 10 min to 1 h, applied for six different location, was introduced by David et al [David *et al.*, 2016].

2.3.2 Exponential smoothing (ES)

The simple exponential smoothing works on the principal of continually revision of a forecast in the light of more recent experience. The predicted values are calculated using weighted averages of past observations with exponentially decreasing weights. Recent observations are given relatively more weights than the previous observations. The model can be mathematically

represented as

$$S_t = \alpha \times X_t + (1 - \alpha) \times S_{(t-1)}, \tag{2.2}$$

where, X_t is the given time series at time t, S_t is the predicted value of the time series at time t, α is the smoothing parameter. The weights decrease exponentially and it depends on the value of parameter α, $0 \leq \alpha \leq 1$. An α close to 1 means all the previous observations are ignored entirely and α equal to 0 means the current observation is ignored entirely. More details about the construction and fundamentals of exponential smoothing can be found in [Winters, 1960].

Recently, ES is used extensively in the construction of hybrid models with other models in the field of renewable energy [Dong *et al.*, 2013; Yang *et al.*, 2015; Dong *et al.*, 2014]. Dong et al. used exponential smoothing state space model for short-term solar irradiance forecasting [Dong *et al.*, 2013]. The study employ two sets of data from meteorological station Singapore and from a rooftop station Colorado, USA. Authors proposed Fourier trend model to stationarize the solar irradiance data and compared the performance with other candidate models using residual analysis and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) stationary test. The candidate models that are used for performance comparison are ARIMA, linear exponential smoothing (LES), simple exponential smoothing (SES) and random walk (RW). Yang et al. in 2015 proposed three hybrid frameworks based on STL decompositions and ETS [Yang *et al.*, 2015]. In the first model, STL was used to decompose the GHI and residual series was forecasted using ETS. Forecast of GHI was obtained by aggregation of these results and the seasonal components. In the second model, DNI and DHI are decomposed by STL and forecast of two residual series are separately accomplished by ETS and final forecast of GHI is obtained by closure equation of aggregation of two forecast results and their respective seasonal components. In the third proposed method, cloud cover index was considered to forecast GHI. Dong et al. [Dong *et al.*, 2014] proposed hybrid prediction algorithm comprising of satellite image analysis and exponential smoothing state space (ESSS)/ANN. Self-organizing maps (SOM) were used to classify cloud cover index and the cloud cover index was forecasted using ESSS. Multilayer perceptron (MLP) was used to derive the solar irradiance from cloud cover index. The author finally compared the results of the proposed model with ARMA, linear exponential smoothing (LES), simple exponential smoothing (SES) and random walk (RW).

2.3.3 Artificial neural networks (ANN)

ANN has been one of the most popular methods in various renewable energy applications and forecasting [Kalogirou, 2001; Mellit and Kalogirou, 2008; Yadav and Chandel, 2014; Kashyap *et al.*, 2015]. ANN is extensively used in the realm of time series to model data containing nonlinear patterns. ANNs are data-driven methods that can efficiently perform a nonlinear mapping between sets of input and output variables. A neural network connects the input variables to one or more output variables through interconnected nodes, called neurons. An input connection has two values associated with it, an input value and a weight. Neurons assign weights to input variables and, through an activation function, gives an output. Each neuron learns through iterative learning cycles and produces the optimum value of weight parameters. The learning algorithm minimizes an error function. The error function depends on the weights associated to different interconnections. At each iteration, the input values are multiplied by an appropriate value of the weight parameter, often within the range (-1 to 1). Before giving inputs to ANN, input values are often normalized to values in the range 0 and 1 to adapt the activation function to the weight values. The error is generally calculated by the square difference between the predicted and the observed values of the output. Input data is usually divided into two sets, 70 percent data is used for training and remaining 30 percent is for testing and validation. There are different ways, discussed in the literature, to determine the number of neurons in the hidden layer [Mellit and Kalogirou, 2008].

In Figure 2.3, x_1, x_2,..., x_D are inputs, z_1, z_2,..., z_M are M hidden neurons and y_1, y_2,..., y_K are K output neurons. x_0 and z_0 are the biases provided at input and hidden neurons respectively.

Each input to hidden neuron is linear combination of the input variables x_1, x_2,..., x_D and

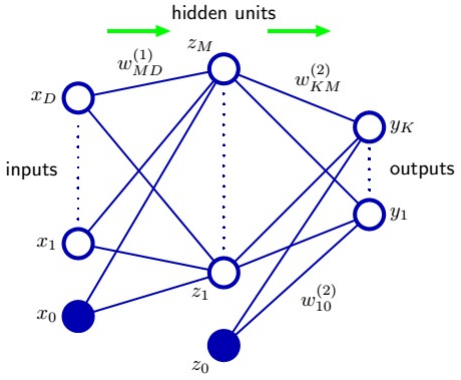

Figure 2.3 : Neural network schematic diagram

M such linear combinations are constructed for $j = 1, 2, ..., M$,

$$a_j = w^{(1)}_{j0} + \sum_{i=1}^{D} w^{(1)}_{ji} x_i, \tag{2.3}$$

where, parameters $w^{(1)}_{ji}$ are the weights and $w^{(1)}_{j0}$ are the biases. The quantity a_j is known as activations of hidden neurons. The hidden neuron uses a non-linear activation function h(.) to transform the input,

$$z_j = h(a_j). \tag{2.4}$$

The output unit linearly combines the input from the hidden layer to give output layer activations and K such linear combination are constructed, for $k = 1, 2, ..., K$, as

$$d_k = w^{(2)}_{k0} + \sum_{j=1}^{M} w^{(2)}_{kj} z_j, \tag{2.5}$$

where, $w^{(2)}_{k0}$ is the bias parameter. Now, each output unit produces the network output $\hat{y}_k$,

$$\hat{y}_k = \Theta(a_k), \tag{2.6}$$

where, $\Theta(.)$ is activation function for the output layer.
The error of prediction is defined by,

$$E(w) = 1/2 \sum_{k=1}^{N} \| \hat{y}_k - t_k \|^2, \tag{2.7}$$

where, $\hat{y}_k$ is network output and t_k is the corresponding target value. Network training functions are used to update weights and bias values.

Many researchers have used the family of ANNs as a predictor for solar irradiance forecasting. Early use of ANN by Al-Alawi and Al-Hinai was seen for climatological variables as inputs to forecast monthly values of GHI over a year [Al-Alawi and Al-Hinai, 1998]. Sfetsos

and Coonick proposed a hybrid framework for mean hourly GHI prediction using feedforward, recurrent and radial basis ANNs and compared the results with traditional linear methods [Sfetsos and Coonick, 2000]. The improved results produced by ANN motivated many researchers to implement it in the field of renewable energy [Paoli *et al.*, 2010; Gutierrez-Corea *et al.*, 2016; Guarnieri *et al.*, ????; Wang *et al.*, 2012; Sfetsos and Coonick, 2000; Mellit and Pavan, 2010; Hocaoğlu *et al.*, 2008].

ANN also found many applications in the construction of hybrid framework with several other methods in recent years [Cao and Cao, 2006, 2005; Mellit *et al.*, 2006; Cao and Lin, 2008; Theodosiou, 2011; Ji and Chee, 2011; Voyant *et al.*, 2012; Yacef *et al.*, 2012; Benmouiza and Cheknane, 2013; Aguiar *et al.*, 2015; Dong *et al.*, 2015; Sharma *et al.*, 2016; Azimi *et al.*, 2016; Ghofrani *et al.*, 2016; Monjoly *et al.*, 2017].

Use of ANN in a hybrid framework for solar irradiance forecasting was introduced by Cao et al. in 2005 and 2006. In these two consecutive works, they proposed a hybrid of discrete wavelet transform (DWT) and artificial network (ANN) [Cao and Cao, 2005, 2006]. In the next work, Cao et al. proposed hybrid of diagonal recurrent wavelet neural network (DRWNN) and the fuzzy network for hourly solar irradiance prediction [Cao and Lin, 2008]. After that many researchers used the hybrid of family of ANN with wavelet transform for prediction of solar resources [Mellit *et al.*, 2006; Sharma *et al.*, 2016; Ji and Chee, 2011]. Later on researchers introduced the use of machine learning techniques for data mining which laid foundation of hybrid framework of ANN with many other machine learning techniques [Benmouiza and Cheknane, 2013; Wu and Chan, 2013; Wu *et al.*, 2014; Azimi *et al.*, 2016; Ghofrani *et al.*, 2016; Hassan *et al.*, 2017; Jiménez-Pérez and Mora-López, 2016; Monjoly *et al.*, 2017]. K-means clustering and nonlinear autoregressive (NAR) neural network models were combined to forecast hourly global horizontal irradiance by Benmouiza et al. [Benmouiza and Cheknane, 2013]. A multi-model framework (MMF) was proposed, based on clustering and an appropriate predictor model by [Wu and Chan, 2013]. In similar work Wu et al. proposed genetic approach of combining multi-model framework for prediction of solar irradiance [Wu *et al.*, 2014]. Azimi et al. proposed a hybrid framework using TB K-means clustering and a multilayer perceptron neural network (MLPNN) [Azimi *et al.*, 2016]. In another similar work, Ghofrani et al. in the same year proposed a hybrid framework using clustering technique, a classification method, a cluster selection algorithm and MLPNN [Ghofrani *et al.*, 2016].

2.3.4 K-nearest neighbour (KNN)

KNN is an instance-based learning approach and it works on the homogeneity of individuals for a given group [Voyant *et al.*, 2013]. KNN stores all available instances present in the data and predict values on a similarity measure (for example, distance functions). The objects for which the class/property values (for classification/regression) are known in advance are chosen as the candidate neighbours. All data points in a group must have similar properties, is the basis of learning rule for the classification. This learning rule assigns a class to all the unclassified data points and classifies to the nearest for a set of previously classified labels. Machine learning tools often learn a model based on the available information in the database. The KNN does not require any explicit learning as the training set itself is considered as the model. A KNN algorithm is characterized mainly by; selection of metric to measure similarity between datasets and determination of number of nearest neighbours [Dasarathy, 1991]. The choice of a metric to measure similarity mostly depends on the nature of the time series. Frequently used metric is the square of the Euclidean distances. Often, the number of neighbours (K) is estimated by minimizing the error metric of the training data. KNN is useful for classification as well as regression. KNN assigns weights such that the closest neighbours are given more weights than the distant ones.
KNN has been extensively used in the field of solar time series forecasting in many applications.

Table 2.6 : Forecasting methods, their strengths and limitations

Forecasting methods	Strengths	Limitations
ARMA	1. The model uses lag and shifts of historic observation	1. Accurate model estimation is not easy.
	2. Autoregressive model with a moving average	2. Does not suit for long-term forecasting
		3. Hardly capture nonlinear patterns present in the time-series
SVM	1. Good for classification or numeric prediction problem	
	2. Not overly influenced by noisy data and not very prone to overfitting	1. Finding the best model requires testing of various combination of kernels and model parameters
	3. Deals nonlinearity and arbitrarily structured data with the help of kernel function	2. Estimation of optimum parameters is computationally challenging and is proportional to number of parameters and size of dataset
ANN	1. Efficiently map input and output relationships	1. "Black box" nature
	2. The availability of multiple training algorithms	2. Initialization of weight value
	3. More general and flexible	3. Challenges of local minima
		4. The problem of Overfitting
KNN	1. Explicit training is not needed	1. Many a times function is approximated locally
	2. Intuitive and easy to implement	2. Determination of the appropriate number of nearest neighbours is challenging
ES	1. Simple mathematical model	1. Suitable for linear time series data
	2. Larger weights are attached to more recent observations	2. Estimation of weight value
Decision tree	1. Domain knowledge is not required	1. Size of decision tree
	2. Can handle multidimensional data	
Markov chain	1. Stochastic process with the Markov property	1. Probabilistic model
Hybrid	1. Compliment the strength of techniques involved	1. Challenging to identify which methods to combine
	2. Can handle complex problems and generally enhance the forecast accuracy	2. Computational cost
		3. Increased model complexity

Pedro et al. used KNN methodology for intra-hour GHI and DNI prediction for horizons ranging from 5 min up to 30 min, and also estimated the corresponding prediction intervals[Pedro and Coimbra, 2015]. Lora et al. used KNN for market price forecasting [Lora *et al.*, 2007]. In the applications of short-term load forecasting, Sudheer and Suseelatha used weighted KNN to forecast one of the wavelet decomposed subseries of electric load data [Sudheer and Suseelatha, 2015].

2.3.5 Support vector machine (SVM)

SVM is a kernel-based machine learning technique introduced by Vapnik that can be used for classification tasks and regression problems [Vapnik, 2013]. The concepts of SVM are based on determination of hyperplanes classifying data into two classes.

As we can see in Figure 2.4, the determination of hyperplane separating two classes is based on finding the largest margin between two classes. Here, maximum margin is interpreted as largest separation of the plane parallel to the hyperplane that does not contain any interior data points. The theoretical details of SVM can be found in Vapnik et al. [Vapnik *et al.*, 1997]. When SVM is used for regression purpose it is termed as support vector regression (SVR) [Lauret *et al.*, 2015]. For regression problems, generally, the target is to fit a function that must not deviate from the measured outputs beyond an error term for each input value. Suitable kernel function is used to map the input data into high-dimensional feature space in case of nonlinear regression problems and therefore, linear classification of data becomes a possibility. The performance of the SVR relies upon the choice of kernel function and selection of the kernel parameter [Zhang *et al.*, 2016]. Many a times SVR is used in the construction of hybrid models with in the field of solar resource forecasting [Bouzerdoum *et al.*, 2013; Mohammadi *et al.*, 2015; Dong *et al.*, 2015; Jiménez-Pérez and Mora-López, 2016]. A hybrid of seasonal ARIMA (SARIMA) and SVM is proposed for short-term

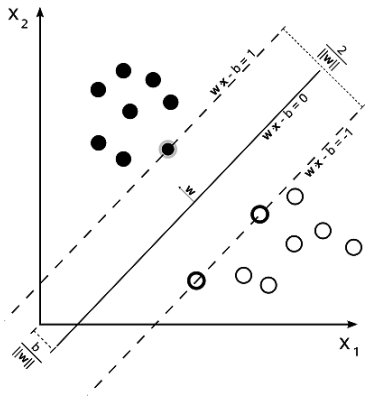

Figure 2.4 : Support vector machine classification

power forecasting of a grid-connected SPV plant [Bouzerdoum *et al.*, 2013]. Hybrid of SVM and wavelet transform (WT) is introduced by Mohammadi et al. to predict horizontal global solar radiation [Mohammadi *et al.*, 2015]. Dong et al. in 2015 introduced a construction of hybrid framework based on SOM, SVR and PSO (particle swarm optimization) to forecast hourly solar irradiance [Dong *et al.*, 2015]. SOM was applied to partition the input dataset into several disjoint sub-datasets of different characteristic information. PSO is implemented to select the parameters of SVR and finally, SVR is used as a predictor. Four hybrid frameworks: decision tree (DT) and ANN; DT and support vector machine regression(SVM-R); support vector machine clustering (SVM-C) and ANN; SVM-C and SVM-R for hourly solar irradiance forecast were reported by Jimenez et al. [Jiménez-Pérez and Mora-López, 2016].

2.3.6 Decision tree

The decision tree is one of the most recently used machine learning tools in the field of solar resource forecasting. A decision tree is a simple non-parametric hierarchical model. It utilizes a tree-like structure to model and predict values. It applies recursive-partitioning regression to split observations into similar observations in smaller regions of the input space [Breiman, 2017]. The decision tree can predict a finite set of values or continuous values known as classification trees and regression trees respectively. The decision tree is characterized by internal nodes, branches, leaf nodes, and the root node. The uppermost node in a tree denotes the root node, each internal node performs a check on the attribute and an outcome of a check is represented by branch and each leaf node represents a class label [**?**]. The advantages of decision tree are; No requirement of domain knowledge for its construction and, its capability of handling multidimensional data. Along with that, the incorporation of acquired knowledge in a tree-like structure is intuitive. The learning of a decision tree is simple, fast and its accuracy is good.

The working of regression tree is straightforward. A response from the number of inputs is predicted by developing a binary tree. A test condition on each input is applied at each internal (decision) node. Based on the outcome of the test either left or the right sub-branch of the tree is selected. The predicted value is provided to the leaf node. The prediction is average of training data points which reaches that leaf. The Error metric is used to determine the best split for each of the variables. Regression tree based ensemble methods for irradiance modelling are discussed in [Hassan *et al.*, 2017]. A regression tree is used for estimation of prediction interval for global irradiance by [Voyant *et al.*, 2018]. Variability of solar irradiance is forecasted using model tree by [McCandless *et al.*, 2015].

2.3.7 Markov chain

Researchers have recently started using Markov processes, more specifically, the Markov chain in the domain of time series modelling and forecasting. A Markov process follows the Markov property which states that given the previous states, the next state of process is independent of all other previous states and depends only on the current state. The present state of the process fully captures all the information for future evolution of the Markov process. Hence, future states can be achieved via a probabilistic process. The transitions between various states are probabilistic with probabilities called transition probabilities. The Markov process can be fully defined by the triple; a state space, a transition matrix and an initial state. Markov chain can model a time series having serial dependence only among adjacent states. Thus Markov process can be effectively used to model systems containing sequence connected events. Markov chains employ to convert the measured data (observations) into states to correspond the data between set intervals. In solar radiation forecasting domain, the sequences of measured irradiance values are first transformed into discrete states. After that these states are used as radiation values between intervals that make the transition over the time [Aguiar *et al.*, 1988]. Hocaoglu and Serttas developed a hybrid of Mycielski and Markov method for hourly prediction of solar radiation [Hocaoglu and Serttas, 2017]. The Mycielski algorithm finds the longest repeating sequence in the past that is present in solar radiation dataset. The Markov transition probabilities decide the foremost probable historical sub-pattern among all sub-patterns obtained by the use of Mycielski algorithm. The most probable historic sub-pattern gives the final value of forecast. Sahin and Sen used Markov models to model and predict the wind speed of ten different cities on hourly time scale at north-western region of Turkey [Sahin and Sen, 2001]. The work of several researchers suggest that Markov models are often used for generation of wind speed time series data [Shamshad *et al.*, 2005; Carpinone *et al.*, 2015; Hocaoğlu *et al.*, 2010].

2.3.8 Hybrid models

As mentioned earlier, hybrid models are any combination of physical and statistical models. The main motivation is to combine various models with a unique feature to complement the advantages of techniques involved to get better forecast accuracy. Hybrid models have the ability to model the linear as well as nonlinear patterns present in the data to span all relevant areas of interest and improve the forecast accuracy. For this reason, hybrid models are used for high-fidelity and robust forecasting to get better forecast accuracy on several spatial and temporal resolutions. Some of the hybrid models and their motivation behind various combinations are presented in brief in the Table 1.3.

2.4 CONCLUSIONS

Various predictive models for solar resource forecasting are reviewed. We also examine the strengths and limitations of forecasting models and discussed various challenges of modelling time series data. We also provided insights that can help in configuration and selection of appropriate forecast model as per needs, applications, and horizon of prediction. We reviewed various hybrid models present in literature along with their motivation of ensembling. Although from various research it is prominent that hybrid approach is suitable for all time horizons and generally gives better forecast accuracy. From the discussions on various forecast models, it is also evident that the choice of the appropriate forecasting models significantly depends on forecast horizon and the variability of underlying data.

…