Forecasting Wind Speed Using Machine Learning ANN Models at 4 Distinct Heights at Different Potential Locations in Pakistan

S.M. AQIL BURNEY[1,2], KONPAL DRAKHSHAN[2], SAADIA KARIM[3]
[1,3] Department of Computer Science
UBIT, University of Karachi
[2] Department of Statistics and Mathematics
CCSIS, Institute of Business Management Karachi
PAKISTAN

Abstract: - The explosive progression in population causes a rapid reduction in the resource of fossil fuel which is the basic supplier of energy in industry and household. This scarcity of fossil fuel is the reason for the costly produced energy. However, pollution is also one of the severe issues occurring due to the burning of gases. Therefore, different researcher worldwide drew their attention to clean and environmentally –friendly energy resources. Wind energy is a renewable source of energy and it is accumulated from renewable resources. Wind speed is one of the most significant parameters used to study the wind energy of any region. This paper presents the fitting of the Artificial Neural Network for the assessment of wind speed in different wind stations in Pakistan. Five Neural Network models have been fitted to the 10-minute mean wind speed data from 2016 to 2018 of each of four distinct heights in 12 different stations in Pakistan. Conventionally used statistical measures are utilized to assess the best-fitted model. The simplest model shows the minimum values of MSE and R2 amongst all other models. The model of one hidden layer with five neurons is the best-fitted model in 12 different stations with four distinct heights in Pakistan. We will be extending this work by applying some other soft computing algorithms such as a random forest with different optimization techniques such as genetic algorithm and swarm optimization algorithms

Key-Words: - Machine Learning Models, Artificial Neural Network (ANN), Wind Speed, Sindh.

Received: May 23, 2022. Revised: July 19, 2023. Accepted: September 2, 2023. Published: October 2, 2023.

1 Introduction

Our planet evolving continuously and as time passes resources will ultimately be insufficient in the provision of energy. Fossil fuels are the basic source that supplies a large amount of energy for industries and households over a decade. Because of the limited resources of fossil fuel, energy is becoming costly [1]. On the other hand, our environment and atmosphere are highly damaged as a consequence of using fossil fuels. One of the severe repercussions of using fossil fuels is called the greenhouse effect. It is the process by which some radiations from the atmosphere warm the earth's surface and gases trap heat in the environment. This trapped heat causes a drastic impact, and it is the major cause of natural disasters. Floods and droughts which occur due to extreme climatic conditions are also the resultant effects of greenhouse gases. Global warming is one of the severe issues caused by gases that trap heat. The melting down of huge glaciers due to global warming elevates the sea level. The greenhouse gases have a significant effect on our health, climate, and our communities [1].

Due to the concerns over the life-threatening effects, environmental pollution, and climate security, several countries are investing in renewable energy. Renewable energy is not reduced by consumption. This term is derived from a wide range of resources that are self-renewing energy sources such as flowing water, sunlight, the earth's internal heat, wind, and biomass such as municipal waste, industrial waste,

agricultural waste, and energy crops [2]. This transfer to environment-friendly energies not only decreased carbon footprint but also reduced climatic change.
Long ago, it was discovered that wind speed has the potential to be utilized for energy production, and the formation of the first-ever windmill was in the year 1887 [3]. Nowadays, it again captivates attention worldwide due to its numerous benefits. It is inexhaustible that it will sustain the atmosphere because it is an emission-free source of production of electricity. More or less 83 countries globally use wind energy for electricity generation [4] and overall, the production capacity has increased by 9.6% in 2018 [5]. Whereas China has expanded its production capacity by wind energy as much as twice annually [6]. The largest producers of wind energy are the US, Germany, Spain, and China [7].

The generation of energy through wind power is comparatively cheaper [8-9], especially in a country like Pakistan which spends 60% of its GDP on imported fuels for the production of electricity [10]. Instead of having the generation potential of 360 GW, it has only produced 106MW from the installed windmills, therefore it ranked 44th in power generation by wind energy [11]. According to the Pakistan Meteorological Department (PMD), 45 different places in Pakistan can be utilized for the harvesting of wind energy [12]. Hence, the appropriate study of wind energy is required for the cost-effective and ecologically harmless generation of electricity.

Wind speed is one of the influential parameters in the production of electricity by wind energy [13]. Therefore, for the assessment of wind energy in any region, it is essentially important to study this parameter by using some appropriate methods of modeling and forecasting. Nowadays ANN is one of the important tools for modeling nonlinear systems and prediction due to its relatively easier and quicker outcome as compared to other practicing methods. ANN is used in a vast range including modeling, assessment, fault diagnostics, identification, prediction, and forecasting of wind energy. ANNs have been popularly adapted techniques by researchers worldwide for the assessment of meteorological parameters fundamentally for wind energy. ANN has been sufficiently utilized for the forecasting of different parameters of wind energy [14].

ANN has been used for the prediction and compared with the analytically obtained results, it has been observed that ANN provided appropriate outcomes [15]. Three different ANN models were applied to the wind speed data collected from Mexico; comparison concluded that the simplest model of ANN outperformed the other two models [16]. ANN is also employed to predict the wind speed data of Hamirpur the mountainous region of India by the input variables, air pressure, solar radiation, temperature, and altitude, it has been observed that the model provides the prediction with greater accuracy [17]. Two different dimensions including temporal and spatial are used in the prediction of wind speed data of Iran by the method of ANN because it is a righteous method for the modeling of complex systems [18]. ANN also assessed and forecasted the capacity of wind energy in Osmaniye the province of Turkey [19]. ANN with modified techniques has been used to analyze the wind speed data such as Radial Basis Function (RBF), Adaptive Linear Element (ALE), and Back-Propagation [20]. It has been stated that RBF is one of the accurately fitted algorithms of NN that provides expeditious forecasting [21].

Hybrid models and systems become popular due to their efficiency and success in solving complex real-world problems. Computational intelligence components are the pillar of the success of this system, including machine learning, Neural Networks (NN), fuzzy logic, and genetic algorithms. These are methods that generate hybrid systems that provide the solution to complex issues by the use of empirical data and domain knowledge and also provide the searching and reasoning methods. Some of the HNNs used in the literature are Genetic algorithm and NN, particle swarm optimization, and NN [22], Wavelet NN [23], Fuzzy NN [24], and Bayesian NN [25].

Some hybrid ANN models are also used for the evaluation of wind speed. A compound model of two Recurrent NN (RNN) with empirical wavelet transformation has been used that provides highly accurate results [26]. A mixture model of RBF and Kohonen's self-organizing map (SOM) has been proposed for the forecasting of short-term power production by both wind direction and wind speed and checked by some popularly utilized statistical measures [27]. A Multi-layer perceptron (MLP) has been trained by a non-dominated sorting genetic algorithm (GA) for the simulation and forecasting of wind energy [28]. The integrated model of chaos and BP-ANN has been proposed and optimized by GA to forecast wind speed [29-30]. The Classical Gaussian Basis Function (CGBF) has been used at every neuron in the improved RBF-ANN model to estimate the wind speed [31]. The hybrid model has been proposed, based on the technique of recurrent wavelet neural network, and compared with the RNN, obtained results showed that the suggested hybrid model has outperformed the conventional model [32]. The convolutional NN has been used and compared with the support vector regression and kernel ridge [33]. Two different types of ANN models have been proposed and compared for the analysis of wind speed by considering the pressure, temperature, and humidity as Input variables for the two distinct suggested models, the obtained results showed that the RNN type NARX model performed better than MLP-BP model [34]. A review of more than a hundred articles has been presented which concluded that ANN has been utilized for multipurpose analysis in the field of wind energy systems [35].

ANN is also combined with the different techniques of other disciplines for the assessment of wind energy to obtain higher accuracy. To reduce the overfitting issues faced due to an extensive number of inputs, a compound model using BP-ANN and Principal Component Analysis (PCA) has been suggested which obtained precise outcomes [36]. To evaluate the seasonal fluctuation in Wind speed data, Weibull distribution has been combined with hybrid ANN. The obtained outcome justified that the proposed model has efficiently illustrated the seasonal impact in the data of wind speed [37].

2 Material and Methods

This study consists of some basic stages, first is obtaining online information related to wind speed data. Analyze by using suitable methods or techniques of soft computing and statistics and then acquire results, testing the accuracy of the proposed model by statistical testing tools, and at last, attain the conclusion. Figure 1 shows the layout of the research.

Fig. 1: Initial layout of the study

2.1 Data Collection Methodology

The data has been obtained from the renewable energy database of the World Bank which is an online source for acquiring data related to different types of renewable energies. The data is based on 10-minute average wind speed at four distinct heights of anemometer which is a-20m, a-40m, a-60m, and, a-80m across Pakistan from 13/4/2016, 5:39:00 PM-31/5/2018, 23:50:00 PM and other meteorological parameters, including turbulence intensity, temperature, pressure, and humidity. The data from the following 12 stations of Pakistan has been obtained, shown in Figure 2, and given in Table 1 with the longitude and latitude of that area of Pakistan.

Table 1: Selected wind station with its description

Provinces	Station Name	Longitude	Latitude
Punjab	Bahawalpur (B)	71.815556	29.326662
	Chakri (Ck)	72.738250	33.320428
	Sadiqabad(S)	70.008112	28.213358
	Quaidabad(Q)	71.895642	32.346589
Sindh	Sujawal(Sj)	68.188650	24.515563
	Umerkot (U)	69.570337	25.083816
	Tandoghulamali (T)	68.875447	25.123567
	Sanghar(Sg)	69.037532	25.815912
Khyber Pakhtunkhwa	Haripur (H)	73.033111	33.973111
	Peshawar(P)	71.795632	33.922092
Balochistan	Gwadar(Gd)	62.346384	25.279814
	Quetta(Qt)	66.936769	30.271586

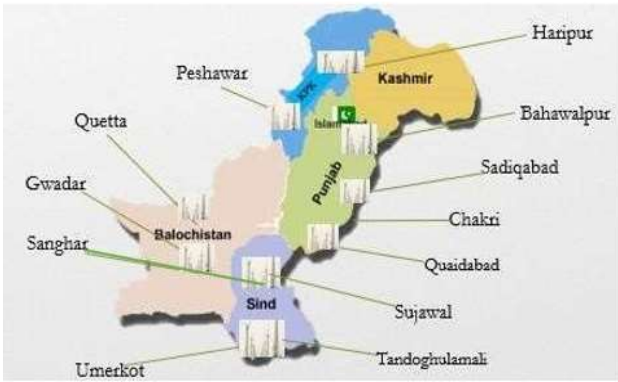


Fig. 2: Selected wind station in Pakistan according to the provinces

2.2 Data Collection Methodology

To assess the wind speed over Pakistan, the ANN technique has been used. ANN is based on a large amount of self-adaptive processing components that are interconnected called neurons. Synchronized computing has been performed to congregate the knowledge by three major steps which are training, learning, and processing of data to find the solution to complex problems and further utilize the information for prediction and forecasting [38-39]. It provides a nonparametric nature of estimators where no predefined assumptions are implemented on the distribution of data [40]. The human neurological system is the inspiration behind the origination of ANN [41]. Neurons in the human brain receive the signals and process them individually likewise neurons in the ANN system send and take the values from the interconnected layers. By repeating the process of providing and receiving values, neurons memorize and learn due to which the system ultimately becomes efficient in making the decisions.

To execute the ANN technique to the wind speed data following steps are to be followed, as mentioned in Figure 3. Each dataset is divided into three subsets, 70% is for training and 30% is for testing and validation.

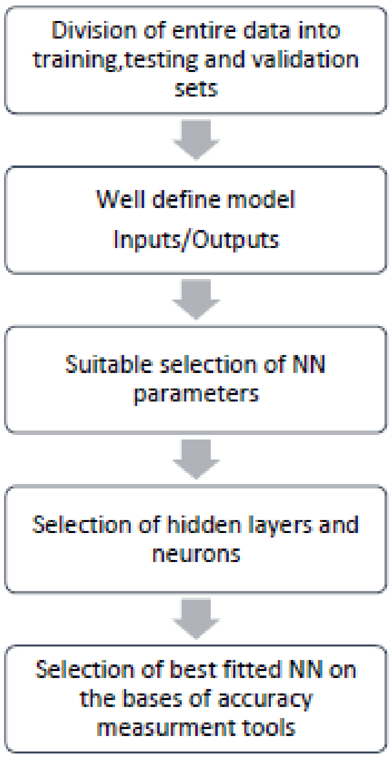


Fig. 3: Methodology of ANN

2.2.1 Design Phase

It is a fundamental stage for the application of this technique. The following basic components are the selection of appropriate input variables for the precise estimation of output, the size of the network including the number of hidden layers and the number of neurons in each hidden layer, the suitable activation function, and the suitable number of experiment selection for acquiring that parameter value which provides the highest accuracy [42].

2.2.1.1 Model of ANN

Wind speed is depending on many metrological parameters. Since this study is based on the estimation of 10-minutes average wind speed. Therefore, the model includes the minimum value and the maximum value for the 10 minutes time duration and the turbulence intensity of the same height. Air pressure, humidity, and temperature are the same for a station for all four heights. The general model of ANN is given in Figure 4.

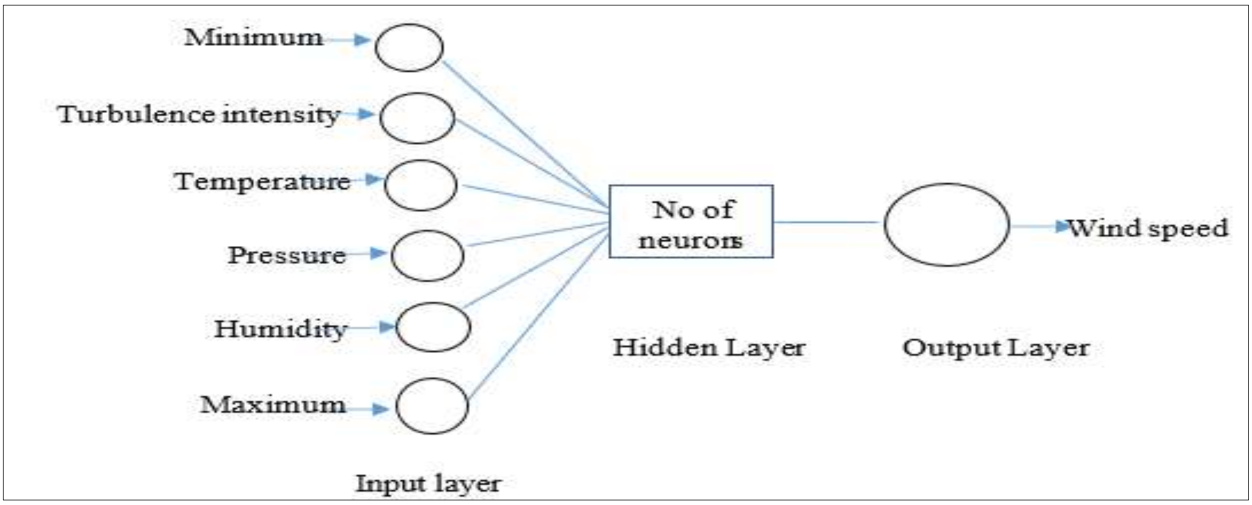


Fig. 4: Methodology of ANN

2.2.1.2 ANN Architecture

The fundamental parameters of ANN are an initial weight function, activation function, methods for encoding the output and the inputs, and iteration required for training sets for obtaining the precise output [43].

The description of the selected network is given in Table 2.

Table 2: The network description

Network Type	Feed-Forward backpropagation
Training Function	Gradient Descent backpropagation with an adaptive learning rate (TRAINGDX)
Transfer Function	Log-Sigmoid (LogSig)

There are two elementary architectures of neurons one is feed-forward NN and another is feed-backward NN. For the study, feed-forward with back-propagation algorithms will be utilized.

Feedforward NN (FFNN) was initially developed in the early stages of ANN which is still commonly used as well. In FFNN signals can only be traveled in a single path. Layers can never be affected by the output value because there is no loop from input to output. It may be further classified as Single-Layer FFN (SLFFN) and Multi-Layer FFN (MLFFN). SLFFN has an input layer that is directly connected to the output layer. Whereas MLFFN has another layer named the hidden layer between input and output. The nodes constitute the supplied activation pattern from the input layer to the next hidden layer respectively. The resultant signals act as the input for the next layer and the process is further continued for the remaining network [44].

Training the NN means the selection of that particular model from the entire sets of models that can minimize the cost. There are several algorithms used to train the model of NN but Gradient Descent (GD) is the most commonly utilized training function. For the computation of true gradient, BP has been used by simply finding the derivative of concerning parameters of a network of the defined cost function, and further, the parameters have been changed according to the direction of gradient-related [45].

Generally, there are three types of activation functions: Threshold function, linear and piecewise linear function, and sigmoid function. The entire family of the sigmoid function is formulated of logistic, algebraic, and hyperbolic tangent functions [46]

For the prevention of overfitting and underfitting complications, an appropriate selection of hidden layers and the number of neurons in each hidden layer is essential. There are numerous rules of thumb for the selection of the hidden layer and the number of neurons [47-48]. According to the literature analyzed, 2 hidden layers-based models are employed mostly for the analysis of wind speed, but the good-fitted model can only be attained based on error analysis. Five different models given in Table 3 will be fit to the data and Figure 5 shows the pictorial structure of the first fitted model.

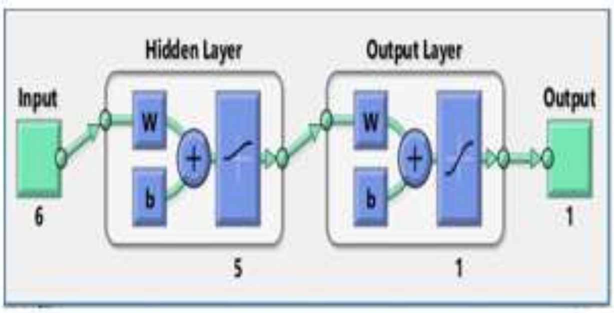


Fig. 5: Structure of the fitted model

Table 3: Five selected models

S/No	No of hidden layer	No neurons in the hidden layer
1	1	5
2	1	7
3	1	9
4	2	7
5	2	9

Models are evaluated by several accuracy measures to find the best-fitted model. For this study following two statistical measures which are Mean Square Error and coefficient of determination 〖(R〗 ^2) shown in Table 4, will be used for the assessment and selection of the model that comes up with the precise output with minimum error [49-50].

Table 4: Performance and accuracy measures

Measures	Formula
Mean Squared Error (MSE)	$\frac{\sum(e_i)^2}{N}$
Coefficient of Determination (R^2)	$1-\frac{SS_r}{SS_t}$

For the computation of proposed models of ANN, MATLAB version 9.4 release name R2018a, open network/data manager, MATLAB tool will be used [51].

2.3 Initially Assessment of Wind Speed

Few popular descriptive measures are calculated to, initially, determine the behavior and pattern of the data. Table 5 shows the measure of central tendency, the measure of dispersion, and the number of observations of each dataset from 2016 to 2018.

Table 5: Descriptive Statistics

Station		**No of observation**	**Min**	**Max**	**Mean**	**Standard deviation**	**Skewness**	**Kurtosis**
B	**a-20**	110946	4e-04	20.0729	3.539682	1.693232	1.105239	5.752392
	a-40	111029	0.0061	20.0729	4.318273	1.930619	0.7649953	4.735393
	a-60	111029	5e-04	23.7493	4.79689	2.220831	0.5737376	3.710448
	a-80	111019	5e-04	24.8228	5.118308	2.47536	0.6026221	3.432719
Ch	**a-20**	119838	5e-04	22.7365	2.440536	1.636908	1.851373	9.013063
	a-40	119831	9e-04	26.1153	2.830441	1.898966	1.6923	8.305313
	a-60	119832	0.001	28.1575	3.126301	2.067429	1.608772	7.887447
	a-80	119844	0.0015	29.0355	3.335638	2.20014	1.559906	7.636192
Gd	**a-20**	81125	0.001	14.5597	3.894323	2.475007	0.9590971	3.425632
	a-40	81130	0.0024	15.8959	4.361159	2.531199	0.7872868	3.265303
	a-60	81126	0.002	16.4991	4.58941	2.566365	0.7137438	3.207104
	a-80	81117	5e-04	16.8967	4.736763	2.607335	0.6915766	3.214347
H	**a-20**	126137	0.0019	16.534	3.06395	1.516315	0.5228371	3.255823
	a-40	125728	5e-04	19.607	3.543264	1.897556	0.6137175	3.243384
	a-60	124782	4e-04	21.2746	3.633065	2.010991	0.7750737	3.90699
	a-80	118709	9e-04	22.3406	3.561241	1.982097	0.9450933	4.932543
P	**a-20**	90508	0.0102	18.7992	2.785071	1.419477	1.467908	8.559319
	a-40	90506	0.002	20.987	2.890474	1.591238	1.560103	8.684283
	a-60	90496	4e-04	22.7723	2.929866	1.701219	1.650405	9.152089
	a-80	89574	5e-04	23.7358	3.009837	1.770817	1.724795	9.627994
Q	**a-20**	91217	0.001	19.8476	2.863788	1.602502	1.717868	9.136759
	a-40	91218	0.0015	23.7222	3.469741	2.006628	1.305507	6.446089
	a-60	91223	0.0026	25.8812	3.80478	2.335359	1.23889	5.591116

	a-80	91219	0.001	27.2414	3.991272	2.548812	1.258608	5.438321
Qt	**a-20**	104316	0.0025	15.9303	3.27863	1.887102	0.9040673	3.782979
	a-40	104327	0.0059	17.9789	3.720266	2.173533	0.8674323	3.751551
	a-60	104326	0.0022	19.1338	3.972312	2.352126	0.8469124	3.707388
	a-80	28152	0.0092	16.0547	4.000859	2.226969	0.7154741	3.249737
S	**a-20**	108982	5e-04	19.6697	3.293894	1.624124	1.154864	5.728112
	a-40	108996	0.0019	23.5003	4.077003	1.900279	0.7960554	4.693372
	a-60	108989	0.0015	26.7719	4.750693	2.341502	0.6757393	3.6171
	a-80	107660	4e-04	28.5486	2.449708	0.639656	0.6396565	3.605485
Sg	**a-20**	96473	0.0202	17.0055	4.484802	2.190856	0.8522987	3.739185
	a-40	96465	0.0014	19.1107	5.296029	2.433241	0.4407148	3.142398
	a-60	96460	0.001	20.6426	5.913478	2.771419	0.2302183	2.546112
	a-80	96464	0.0089	21.5117	6.27292	3.014581	0.1840759	2.318795
Sj	**a-20**	112193	0.0115	16.8196	5.42133	2.355959	0.6366384	3.231811
	a-40	112195	5e-04	18.6522	6.361603	2.471836	0.4164136	3.367307
	a-60	112194	0.0247	19.7977	7.017822	2.608736	0.1523535	3.203628
	a-80	112181	0.0014	20.4691	7.493932	2.740752	-0.0221345	2.990569
T	**a-20**	116045	0.0055	15.8416	4.801472	2.296618	0.7208412	3.080485
	a-40	116047	9e-04	17.36	5.834624	2.517653	0.481283	3.169898
	a-60	116043	0.0015	17.9102	6.512893	2.685257	0.1763456	2.867687
	a-80	116053	0.0014	18.5386	6.961413	2.889245	0.02712907	2.580764
U	**a-20**	91269	5e-04	21.5087	4.176914	2.126058	0.8774143	3.741682
	a-40	89448	0.0277	18.6296	5.018032	2.277336	0.5318844	3.373732
	a-60	91272	0.0126	29.0817	5.674894	2.526999	0.331302	3.088653
	a-80	91275	0.0015	30.6142	6.105267	2.7091	0.2053617	2.782717

The lowest value of all the minimum average value of wind speed is 4e-04 m/s and the highest value of maximum average wind speed is 30.6142 m/s of the selected region over Pakistan during the interval of 4/13/2016, 5:39:00 PM-5/31/2018, 23:50:00 PM. Mean value mostly lies between 2 m/s - 5 m/s. According to the skewness, most datasets are either positively skewed or tend towards symmetricity. The value of kurtosis is diversified, some datasets are highly peaked and moderate, and some are flattened.

3 Fuzzy Time Series Modeling and Forecasting

Fuzzy time series have been extensively used to make predictions of weather, road accidents, academic enrollments, population, and stock prices. In this paper, we have introduced an improved fuzzy time series forecasting model. This new model is applied in forecasting the University of Alabama student enrollments. Latera comparison has been done with some of the existing fuzzy time series forecasting methods being carried out on the same data set for university student enrollments. It has been observed that the proposed model has improved forecasting accuracy as well as reduced model complexity compared to other methods.

Fig.6: (S-O-T-1-F-T-S) Model Algorithms of fuzzy times series

4 Results and Discussion

Selected ANN models mentioned in Table 3 are fit to the data and their training operations were investigated systematically. Statistical performance measures are employed to evaluate the model, the higher the value of R^2 and the lower the value of MSE are considered as the good fitted models. Table 6 exhibits the summary of models fitted to the data commensurate with the conventionally used accuracy tools.

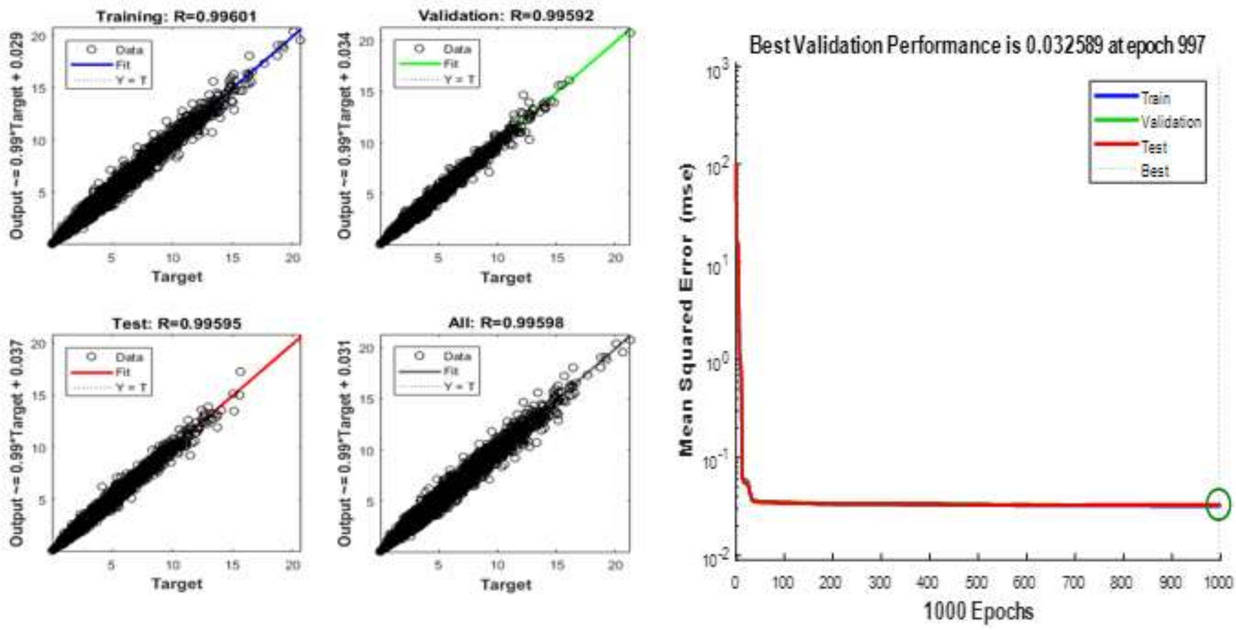
Table 6: 〖R〗^2 and MSE values for the suggested models

S/No	**hidden layer**	**neurons in the hidden layer**	**MSE**	R^2	**MSE**	R^2	**MSE**	R^2	**MSE**	R^2
Station B			**a-20**		**a-40**		**a-60**		**a-80**	
M-1	1	5	0.11507	0.97931	0.13306	0.98177	**0.12112**	**0.98754**	0.14468	0.98831
M-2	1	7	0.12729	0.91877	0.094685	0.98687	0.38238	0.95884	0.12396	0.98993
M-3	1	9	0.13178	0.97883	0.093567	0.98819	0.18579	0.98095	**0.08516**	**0.99329**
M-4	2	7	**0.11263**	**0.97927**	**0.080331**	**0.98883**	0.225	0.97656	0.16123	0.98689
M-5	2	9	0.13622	0.97546	0.31675	0.9582	0.12748	0.98706	0.20133	0.98349
Station Ck			**a-20**		**a-40**		**a-60**		**a-80**	
M-1	1	5	0.10466	0.98093	**0.087211**	**0.98803**	**0.10981**	**0.98702**	0.13581	0.98716
M-2	1	7	0.087207	0.98364	0.15515	0.39787	0.153	0.98197	0.18984	0.98074
M-3	1	9	0.14206	0.9736	0.13323	0.98246	0.15933	0.98158	0.15395	0.9848
M-4	2	7	0.13245	0.97629	0.13694	0.98102	0.16827	0.98036	**0.10821**	**0.9884**
M-5	2	9	**0.085646**	**0.98423**	0.1707	0.9772	0.13581	0.98398	0.17186	0.9825
Station Gd			**a-20**		**a-40**		**a-60**		**a-80**	
M-1	1	5	0.15494	0.98743	0.12583	0.98993	**0.06288**	**0.99506**	0.14693	0.9883
M-2	1	7	0.15905	0.98705	0.078611	0.99393	0.11101	0.99145	0.1465	0.98521
M-3	1	9	**0.13057**	**0.98921**	**0.072972**	**0.99413**	0.31583	0.97526	0.11885	0.98737
M-4	2	7	0.16582	0.98663	0.13466	0.98934	0.11367	0.99144	0.13984	0.98622
M-5	2	9	0.16217	0.98674	0.13899	0.98918	0.15524	0.98817	**0.10883**	**0.98876**
Station H			**a-20**		**a-40**		**a-60**		**a-80**	
M-1	1	5	**0.029898**	**0.99384**	**0.1046**	**0.98559**	**0.03258**	**0.99598**	0.1829	0.97732
M-2	1	7	0.11101	0.97576	0.14252	0.98016	0.08635	0.98828	0.16824	0.97793
M-3	1	9	0.16379	0.9675	0.10521	0.98577	0.14351	0.98287	**0.10603**	**0.98631**
M-4	2	7	0.16991	0.96213	0.22276	0.9693	0.10918	0.98727	0.11497	0.9841
M-5	2	9	0.13367	0.96998	0.15031	0.97934	0.12921	0.98468	0.31723	0.96003
Station P			**a-20**		**a-40**		**a-60**		**a-80**	
M-1	1	5	0.16203	0.96021	**0.073585**	**0.9850**	0.10908	0.98018	0.13198	0.97941
M-2	1	7	**0.07477**	**0.98124**	0.10125	0.9794	0.11993	0.97999	0.12496	0.98098
M-3	1	9	0.09679	0.97693	0.11081	0.9787	0.21395	0.96314	0.09886	0.98404
M-4	2	7	0.10953	0.97191	0.076405	0.9844	**0.06499**	**0.98854**	**0.08214**	**0.98699**
M-5	2	9	0.08297	0.97896	0.13642	0.9741	0.1476	0.97424	0.19448	0.96876
Station Q			**a-20**		**a-40**		**a-60**		**a-80**	
M-1	1	5	0.08883	0.98349	**0.10766**	**0.9860**	**0.10384**	**0.90901**	0.15153	0.98664
M-2	1	7	**0.065177**	**0.98773**	0.13457	0.9851	0.23191	0.9788	0.1804	0.98638
M-3	1	9	0.1175	0.97697	0.1176	0.986	0.13053	0.98769	0.15224	0.98864
M-4	2	7	0.074855	0.98492	0.18394	0.9769	0.15939	0.98572	**0.14395**	**0.98939**
M-5	2	9	0.093661	0.98235	0.28651	0.9647	0.18081	0.9834	0.1958	0.98487
Station Qt			**a-20**		**a-40**		**a-60**		**a-80**	

M-1	1	5	0.10133	0.9862	**0.12616**	**0.9892**	**0.11602**	**0.98654**	0.10887	0.98874
M-2	1	7	**0.099054**	**0.98589**	0.12548	0.9868	0.19773	0.98223	0.11081	0.98914
M-3	1	9	0.11798	0.98315	0.20998	0.9778	0.14992	0.98627	0.33548	0.9668
M-4	2	7	0.13617	0.98047	0.15087	0.9838	0.17241	0.98402	**0.09089**	**0.9906**
M-5	2	9	0.15232	0.97791	0.15688	0.9834	0.20319	0.98144	0.12534	0.98755
Station S			**a-20**		**a-40**		**a-60**		**a-80**	
M-1	1	5	0.10101	0.98161	0.072932	0.9897	0.17427	0.98464	**0.09873**	**0.99258**
M-2	1	7	0.085139	0.98437	0.090355	0.9882	**0.13754**	**0.98703**	0.14309	0.98868
M-3	1	9	0.078468	0.98582	0.036634	0.9952	0.30534	0.96918	0.22355	0.98246
M-4	2	7	**0.070228**	**0.98663**	**0.03468**	**0.9951**	0.22482	0.97807	0.17363	0.98627
M-5	2	9	0.13428	0.97542	0.10689	0.9855	0.23174	0.97717	0.32009	0.97359
Station Sg			**a-20**		**a-40**		**a-60**		**a-80**	
M-1	1	5	**0.077652**	**0.9918**	**0.13607**	**0.98874**	**0.0763**	**0.9947**	**0.15084**	**0.99211**
M-2	1	7	0.091397	0.99057	0.144405	0.98777	0.1552	0.9911	0.17118	0.99058
M-3	1	9	0.12366	0.98745	0.14123	0.98788	0.1414	0.9906	0.16088	0.99108
M-4	2	7	0.11486	0.98809	0.20093	0.98335	0.2744	0.9825	0.18521	0.99043
M-5	2	9	0.27057	0.97178	0.15437	0.9868	0.2571	0.9831	0.17489	0.99005
Station Sj			**a-20**		**a-40**		**a-60**		**a-80**	
M-1	1	5	0.10245	0.98933	0.13793	0.98834	0.1457	0.9905	0.18782	0.98746
M-2	1	7	0.13753	0.98572	**0.10799**	**0.99063**	0.1388	0.991	0.18086	0.9884
M-3	1	9	**0.096375**	**0.98973**	0.15351	0.98713	0.1347	0.9910	0.22609	0.98485
M-4	2	7	0.14827	0.98443	0.12919	0.9892	0.1693	0.9887	**0.1522**	**0.99035**
M-5	2	9	0.12384	0.98688	0.1491	0.98743	**0.1229**	**0.9918**	0.22224	0.98559
Station T			**a-20**		**a-40**		**a-60**		**a-80**	
M-1	1	5	0.12589	0.98683	**0.09876**	**0.99117**	0.1430	0.9896	0.20226	0.98738
M-2	1	7	0.12311	0.98704	0.15166	0.9869	0.2456	0.9813	**0.13898**	**0.99162**
M-3	1	9	0.2322	0.97579	0.13719	0.98781	0.1121	0.9918	0.20338	0.98784
M-4	2	7	**0.086782**	**0.9909**	0.11568	0.98975	0.9866	0.1811	0.15024	0.99095
M-5	2	9	0.10659	0.98876	0.17431	0.98558	**0.1069**	**0.9919**	0.15596	0.99033
Station U			**a-20**		**a-40**		**a-60**		**a-80**	
M-1	1	5	0.13355	0.98543	**0.083314**	**0.99106**	0.2690	0.9798	0.12938	0.99143
M-2	1	7	0.13146	0.98561	0.18039	0.98265	0.5000	0.9609	0.27528	0.98168
M-3	1	9	**0.12644**	**0.986**	0.30914	0.96964	0.1750	0.9865	0.39789	0.97247
M-4	2	7	0.25204	0.97144	0.18036	0.98263	**0.1459**	**0.9878**	0.28094	0.98135
M-5	2	9	0.18124	0.98026	0.98267	0.18017	0.2734	0.9787	**0.11668**	**0.99131**

According to the criteria mentioned above, models are selected (the bolded values) for each of the four different heights of several 12 distinct stations in four different provinces of Pakistan. For the proposed methodology, 240 ANN models have fitted to the data and 48 models have been selected for each dataset.

It has been observed that MD-1 (one hidden layer with 5 neurons) is the most suitable model appropriately fitted in a total of 19 datasets after that MD-4(two hidden layers with 7 neurons) is also accurately fitted in 12 datasets. The MD-5 has appeared the least in the best-fitted models and seems well fitted in only 5 datasets amongst 48. The ratio of is similar and each of the models satisfied 6 datasets. It has been found that Station Sg is the only station that shows MD-1 as the best-fitted model for all four distinct heights. Stations U, Sj, and T are those stations that show different fitted models for all four distinct heights. The model MD-1 fitted to station H at the height of a-60m shows the highest R2 value amongst all fitted models and MD-1 fitted to station H at the height of a-20m shows the minimum MSE value amongst the entire fitted model. The plots of MSE and *R*2have also been analyzed for all three sets, training, testing, and validation, following five best-selected models having the highest R2 and lowest MSE of MD-1is shown in Figure 7(a)-7(e).



MD-2 and MD-3

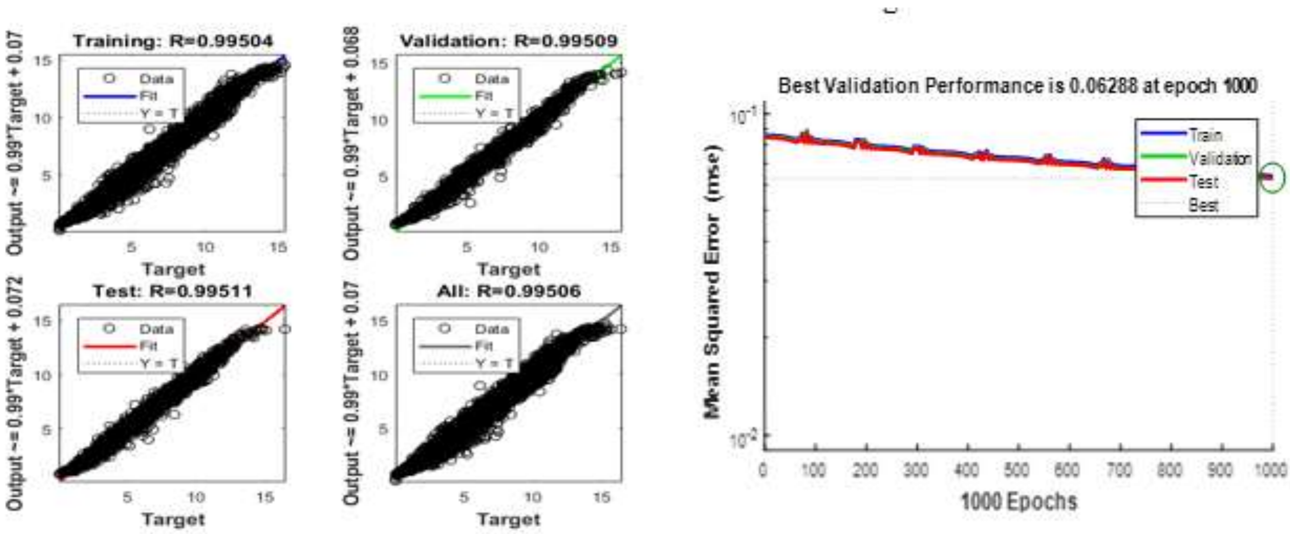


Fig. 7 (b): best-fitted MD1 on station Gd of a-60m

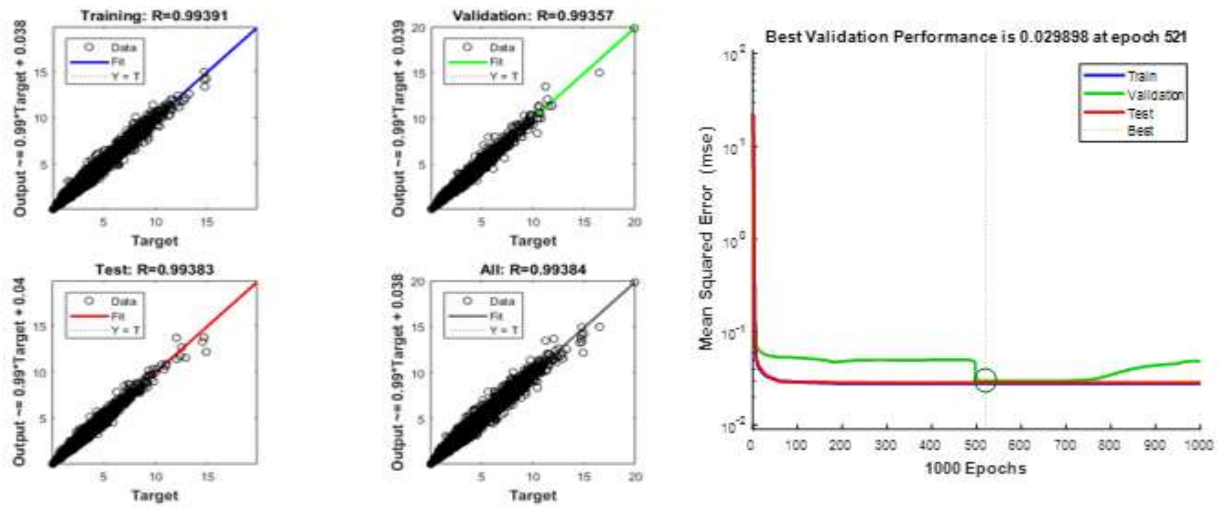


Fig. 7 (c): best-fitted MD1 on station H of a-60m

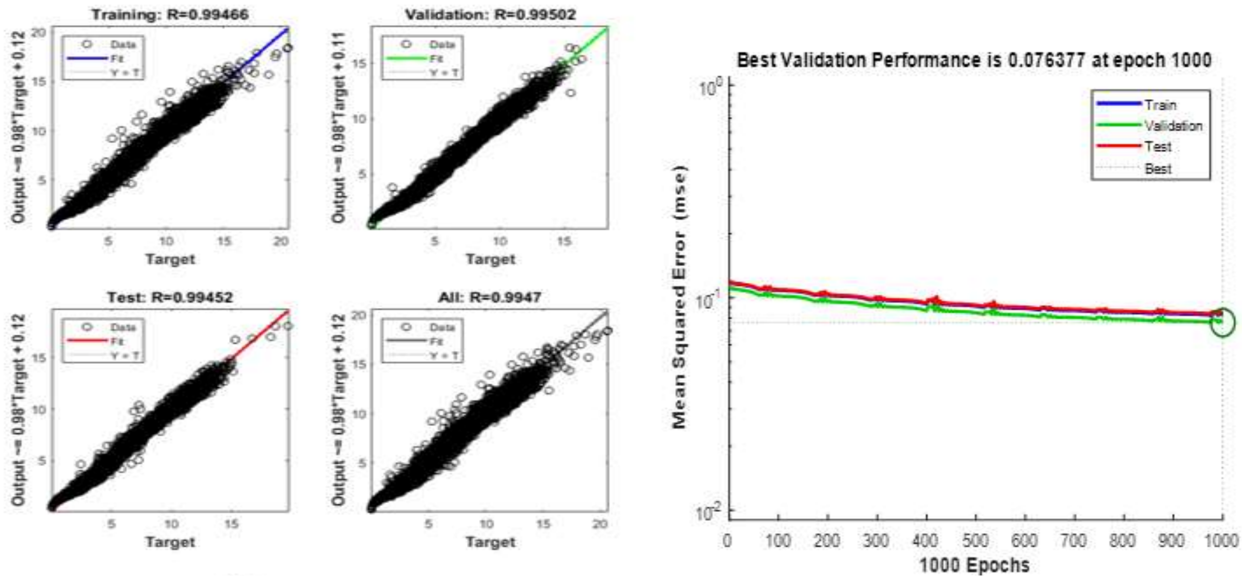


Fig. 7 (d): best-fitted MD1 on station Sg of a-60m

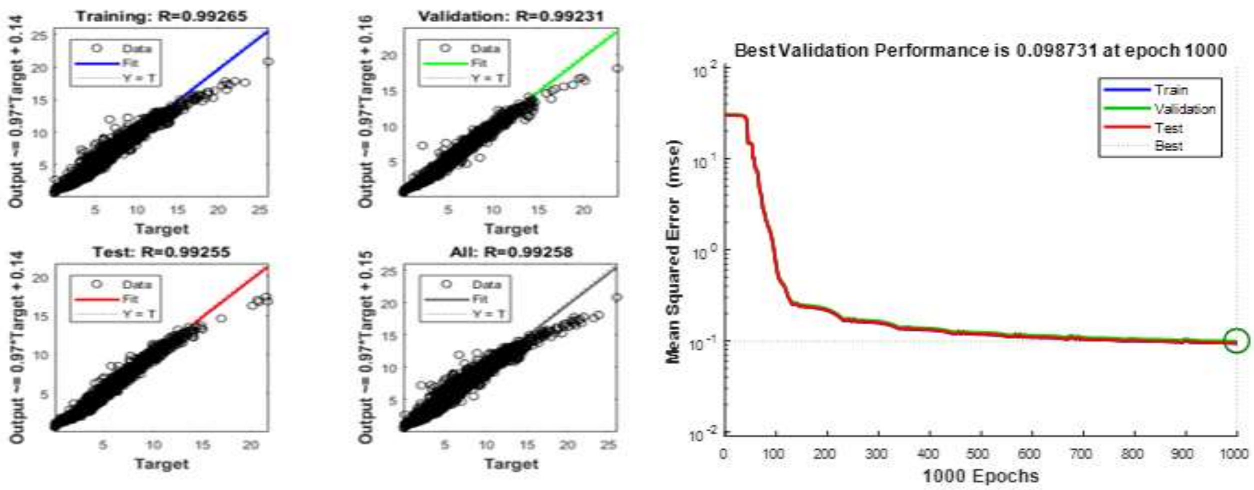


Fig. 7 (e): best-fitted MD1 on station S of a-60m

The scatter plots provide the values of R2 at the training, testing, and validation steps and also provide the explained variation for the entire procedure. The line graph shows the decline pattern of MSE for the training, testing, and validation sets and also the epochs required to approach the optimization point while attaining the minimum value of MSE.

5 Conclusion

ANN models have fitted to the data. Data sets are divided into three sets 70% training set, 15% testing sets, and 15% validation sets. 5 different models are used since generally, the criteria are, that the model is considered to best model if the errors are minimum and maximum variation has explained by the model based on the selection criteria the most appropriately selected model's heights are given in that Table 7. For a height of a-20m, the least suitable is MD-5 and all three models MD-2, MD-3, and MD-4 are equally likely for the analysis of wind speed at a-20m height for all the stations. At a-40m and a-60m for all 12 stations, MD-1 has significantly appeared at a larger number of times, indicating that at a-40m and a-60m MD-1 has been observed to be the best model fitted to this height.

Table 7: Summary of selected ANN models

Stations	Heights			
	a-20	**a-40**	**a-60**	**a-80**
B	M-4	M-4	M-1	M-3
Ck	M-5	M-1	M-1	M-4
Gd	M-3	M-3	M-1	M-5
H	M-1	M-1	M-1	M-3
P	M-2	M-1	M-4	M-4
Q	M-2	M-1	M-1	M-4
Qt	M-2	M-1	M-1	M-4
S	M-4	M-4	M-2	M-1
Sg	M-1	M-1	M-1	M-1
Sj	M-3	M-2	M-5	M-4
T	M-4	M-1	M-5	M-2
U	M-3	M-1	M-4	M-5

At a-80m MD-4 was found to be the best model to assess the wind speed at this height though it has not very significantly fitted in most of the stations.

Overall, it can be concluded that the simplest model has by far the most repeated model different stations-wise and distinct height-wise which means the MD-1 has most appropriately explained the data of wind speed in Pakistan. The least fitted model is MD-5, a model based on 2 hidden layers and 9 neurons.

6 Future Work

This study is based on the fitting of different ANN models and the assessment of their fitness by different measures. To evaluate the behavior of wind speed in Pakistan by changing the location and the height of an anemometer. The obtained results may be further used for predictions and forecasting. Furthermore, some hybrid models and some defined optimization techniques as discussed in the literature review can be used to analyze the wind speed pattern in Pakistan with more places from Sindh coastal areas of Pakistan. This study will be extended to forecasting using fuzzy time series modeling and neuro-fuzzy time series modeling for forecasting when a subjective approach is preparable, [52] and for Continuous support and development across various domains of the energy sector are required to achieve sustainability targets following [55] forecasting approaches using soft intelligent computing and generative AI approach will be dominating in our future work following recent trends [56] keeping in the view demand side of energy from all workable resources [56].

Acknowledgement:
The authors acknowledge the advice and cooperation of Dr Mudassir Professor of Statistics University of Karachi and Prof. Dr. Feroz Ahmed of the Physics Dept, University of Karachi as an expert in renewable energy. Dr Khuram Iqbal of CCSIS 2018-2021.

References:

[1] S. R. Bhusare, "Renewable energy: An eco-friendly alternative?" *Quest-The Journal of UGC-HRDC Nainital*, vol. 9, no. 2, p. 152, 2015. doi:10.5958/2249-0035.2015.00024.8
[2] J. H. Perkins, "Special report on renewable energy sources and climate change mitigation: 2011. Intergovernmental Panel on Climate Change, Working Group III—mitigation of climate change. Cambridge University Press, Cambridge, England. 1,088 pp. $100.00

hardcover (ISBN13: 9781107607101). also available for free at http://srren.ipcc-wg3.de/report (ca. 1,544 pp.).," *Environmental Practice*, vol. 14, no. 3, pp. 237–238, 2012. doi:10.1017/s1466046612000233.

[3] T. J. Price, "James Blyth — Britain's first modern wind power pioneer," *Wind Engineering*, vol. 29, no. 3, pp. 191–200, 2005. doi:10.1260/030952405774354921

[4] REN21, "Renewables 2011: Global Status Report", 2011.

[5] N. Gas, Wind power "Global Installed Capacity in 2018, GWEC", 2018.

[6] Z. Tan et al., "Potential and policy issues for sustainable development of wind power in China," *Journal of Modern Power Systems and Clean Energy*, vol. 1, no. 3, pp. 204–215, 2013. doi:10.1007/s40565-013-0037-8

[7] J. A. Peters and J. G. Olivier, Trends in CO2 emissions, 1990-2012 report, 2015. doi:10.1787/saeo-2015-graph46-en

[8] A. D. Mills and K. Porter, The cost of transmission for wind energy: A review of transmission planning studies (No. LBNL-1471E), *Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States,* 2009.

[9] D. Gielen, Renewable energy technologies: cost analysis series. Sol Photovolt, vol. 1, no. 1, pp. 52, 2012.

[10] M. A. Sheikh, Energy and renewable energy scenario of Pakistan, *Renewable and Sustainable Energy Reviews*, vol. 14, no. 1, pp. 354-363, 2010.

[11] A. W. Bhutto, A. A. Bazmi, G. Zahedi, Greener energy: Issues and challenges for Pakistan—wind power prospective, *Renewable and Sustainable Energy Reviews*, vol. 20, pp. 519-538, 2013.

[12] K. Harijan, M. A. Uqaili, M. Memon, and U. K. Mirza, "Forecasting the diffusion of wind power in Pakistan," *Energy*, vol. 36, no. 10, pp. 6068–6073, 2011. doi: 10.1016/j.energy.2011.08.009

[13] W. Tong, Fundamentals of wind energy, In Wind power generation and wind turbine design, WIT Press, pp. 23, 2010.

[14] O. Özgönenel, & D. W. Thomas, Short-term wind speed estimation based on weather data, Turkish Journal of Electrical Engineering & Computer Sciences, vol. 20, no. 3, pp. 335-346.

[15] A. N. Celik, & M. Kolhe, Generalized feed-forward based method for wind energy prediction, Applied Energy, vol. 101, pp. 582-588, 2013.

[16] E. Cadenas, & W. Rivera, Short-term wind speed forecasting in La Venta, Oaxaca, México, using artificial neural networks, *Renewable Energy*, vol. 34, no. 1, pp. 274-278, 2009.

[17] P. Ramasamy, S. S. Chandel, & A. Yadav, A, Wind speed prediction in the mountainous region of India using an artificial neural network model, Renewable Energy, vol. 80, pp. 338-347, 2015.

[18] Y. Noorollahi, M. A. Jokar, & Kalhor, Using artificial neural networks for temporal and spatial wind speed forecasting in Iran. Energy Conversion and Management, vol. 115, pp. 17-25, 2016.

[19] S. Tasdemir, B. Yaniktepe, & A. B. Guher, Determination of Wind Potential of a Specific Region using Artificial Neural Networks, International Journal of Intelligent Systems and Applications in Engineering, vol. 5, no. 3, pp. 158-162, 2017.

[20] G. Li, & J. Shi, on comparing three artificial neural networks for wind speed forecasting, Applied Energy, vol. 87, no. 7, pp. 2313-2320, 2010.

[21] X. Wu, B. Hong, X. Peng, F. Wen, & J. Huang, Radial basis function neural network based short-term wind power forecasting with Grubbs test, In 2011 4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), pp. 1879-1882, 2011.

[22] H. Liu, H. Q. Tian, C. Chen, & F. Y. Li, An experimental investigation of two Wavelet-MLP hybrid frameworks for wind speed prediction using GA and PSO optimization, International Journal of Electrical Power & Energy Systems, vol. 52, pp. 161-173, 2013.

[23] H. Esen, F. Ozgen, M. Esen, & A. Sengur, A, Artificial neural network and wavelet neural network approaches for modeling of a solar air heater, Expert systems with applications, vol. 36, no. 8, pp. 11240-11248, 2009.

[24] S. A. Burney, T. A. Jilani, & A. Saleemi, Optimizing neuronal complexity using wavelet-based multiresolution analysis for Type-I fuzzy neural networks, Proc. of 4th WSEAS-CIMMACS, pp. 210-216, 2005.

[25] R. M. Neal, Bayesian learning for neural networks, Springer Science & Business Media, vol. 118, 2012.

[26] H. Liu, X. W. Mi, & Y. F. Li, Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short-term memory neural network and Elman neural network, *Energy conversion and management*, vol. 156, pp. 498-514, 2018.

[27] M. Dadkhah, M. J. Rezaee, & A. Chavoshi, A. Z, Short-term power output forecasting of hourly

operation in power plant based on climate factors and effects of wind direction and wind speed, Energy, vol. 148, pp. 775-788, 2018.
[28] R. Ak, Y. F. Li, V. Vitelli, & E. Zio, Adequacy assessment of a wind-integrated system using neural network-based interval predictions of wind power generation and load, International Journal of Electrical Power & Energy Systems, vol. 95, pp. 213-226, 2018.
[29] D. Z. Huang, R. X. Gong, & S. Gong, Prediction of wind power by chaos and BP artificial neural networks approach based on genetic algorithm, Journal of Electrical Engineering & Technology, vol. 10, no. 1, pp. 41-46, 2015.
[30] P. Zhang, Y. Wang, L. Liang, X. Li, & Q. Duan, Short-Term Wind Power Prediction Using GA-BP Neural Network Based on DBSCAN Algorithm Outlier Identification, Processes, vol. 8, no. 2, pp. 157, 2020.
[31] G. Chang, H. J. Lu, Y. R. Chang, & Y. D. Lee, An improved neural network-based approach for short-term wind speed and power forecast, Renewable energy, vol. 105, pp. 301-311, 2017.
[32] P. P. Pradhan, & B. Subudhi, Wind speed forecasting based on wavelet transformation and recurrent neural network. International Journal of Numerical Modelling: Electronic Networks, Devices and Fields, vol. 33, no. 1, 2020.
[33] Z. Wang, J. Zhang, Y. Zhang, C. Huang, & L. Wang, Short-Term Wind Speed Forecasting Based on Information of Neighboring Wind Farms, IEEE Access, vol. 8, pp. 16760-16770, 2020.
[34] Y. Amellas, O. El Bakkali, A. Djebil, & A. Echchelh, Short-term wind speed prediction based on MLP and NARX network models, Indonesian Journal of Electrical Engineering and Computer Science, vol. 18, no.1, 150-157, 2020.
[35] R. Ata, Artificial neural networks applications in wind energy systems: a review, Renewable and Sustainable Energy Reviews, vol. 49, pp. 534-562, 2015.
[36] D. He, & R. Liu, Ultra-short-term wind power prediction using ANN ensemble based on PCA, In Proceedings of the 7th International Power Electronics and Motion Control Conference, vol. 3, pp. 2108-2112, 2012.
[37] A. A. Kadhem, N. Wahab, I. Aris, J. Jasni, & A. Abdalla, Advanced wind speed prediction model based on a combination of Weibull distribution and an artificial neural network, Energies, vol.10, no. 11, pp. 1744, 2017.
[38] I. A. Basheer, & M. Hajmeer, Artificial neural networks: fundamentals, computing, design, and application, Journal of microbiological methods, vol. 43, no. 1, pp. 3-31, 2000.
[39] S. M. Burney, M. S. A. Khan, M. S. A., & T. A. Jilani, Feature deduction and ensemble design of parallel neural networks for an intrusion detection system, IJCSNS, vol. 10, no. 10, pp. 259, 2010.
[40] S. Walczak, Artificial neural networks, In Advanced Methodologies and Technologies in Artificial Intelligence, Computer Simulation, and Human-Computer Interaction, IGI Global, pp. 40-53, 2019.
[41] E. Alpaydin, Introduction to machine learning, MIT Press, 2009.
[42] S. L. Muhammad, W. A. Abidin, W. Y. Chai, A. Baharun, & T. Masri, Development of wind mapping based on artificial neural network (ANN) for energy exploration in Sarawak, International Journal of Renewable Energy Research, vol. 4, no. 3, pp. 618-627, 2014.
[43] I. G. Damousis, M. C. Alexiadis, J. B. Theocharis, & P. S. Dokopoulos, A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation, IEEE Transactions on Energy Conversion, vol. 19, no. 2, pp. 352-361, 2004.
[44] S. Haykin, Neural networks, New York: Prentice Hall, vol. 2, 1994.
[45] M. Forouzanfar, H. R. Dajani, V. Z. Groza, M. Bolic, & S. Rajan, Comparison of Feed-Forward Neural Network Training Algorithms for Oscillometric Blood Pressure Estimation (PDF). 4th Int. Workshop Soft Computing Applications, Arad, Romania: IEEE, 2010.
[46] D. Kriesel, A brief introduction on neural networks, 2007. https://www.dkriesel.com/_media/science/neuronalenetze-en-zeta2-2col-dkrieselcom.pdf
[47] F. S. Panchal, & M. Panchal, Review on methods of selecting a number of hidden nodes in an artificial neural network, International Journal of Computer Science and Mobile Computing, vol. 3, no. 11, pp. 455-464, 2014.
[48] R. Anderson Sprecher, Model comparisons and R^2, The American Statistician, vol. 48, no. 2, pp. 113-117, 1994.
[49] S. Makridakis, S. C. Wheelwright, & R. J. Hyndman, Forecasting methods and applications, John Wiley & sons,2008.
[50] H. Demuth, & M. Beale, Neutral network toolbox for use with Matlab (No. BOOK). The Math Works, 1995.
[51] T. A. Jilani, & S. M. A. Burney, M-factor high order Fuzzy Time Series forecasting for road accident data, Analysis and design of intelligent

systems using soft computing techniques, Springer Berlin Heidelberg, pp. 246-254, 2007.
[52] S. M. A. Burney, S. M. Ali, & S. Burney, A survey of soft computing applications for decision making in supply chain management, Engineering Technologies and Social Sciences (ICETSS), Third International Conference 2017, pp. 1-6, 2017.
[53] S. M. A. Burney, S. M. Ali, and M. S. Khan, A Novel High Order Fuzzy Time Series Forecasting Method with Higher Accuracy Rate, IJCSNS, vol. 18, no. 5, pp. 9, 2018.
[54] W. Tushar, M. I. Azim, M. R. Alam, C. Yuen, R. Sharma, T. Saha, H. V. Poor, Achieving the UN's sustainable energy targets through dynamic operating limits, iScience, vol. 26, no. 7, 2023, doi:10.1016/j.isci.2023.107194.
[55] Hong, P. Pinson, Y. Wang, R. Weron, D. Yang, & H. Zareipour, Energy Forecasting: A Review and Outlook, in *IEEE Open Access Journal of Power and Energy*, vol. 7, pp. 376-388, 2020, doi:10.1109/OAJPE.2020.3029979.
[56] F. Ahmed, S. M. A. Burney, S. A. Husain, Monthly average daily global beam and diffuse solar radiation and its correlation with hours of bright sunshine for Karachi, Pakistan. Renew Energy, vol. 1, no. 1, pp. 115-8, 1991. http://dx.doi.org/10.1016/0960- 1481(91)90111-2

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

Professor Dr S M Aqil Burney contribution is conceiving planning for data collection and suggested modelling approach. Konpal contribution is follow for computing as suggested by Professor Dr S M Aqil Burney. The contribution of Saadia Karim is to check computing and finalize script for publication keeping in view Statistics and soft computing approach with data verification and validation of results with proper referencing.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself
No funding was received for conducting this study.

Conflict of Interest

No conflict of interest of any author